United States
Environmental Protection
Agency

EPA/600/R-08/076F | June 2009 | www.epa.gov

Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines

Office of Research and Development, Washington, DC 20460
National Center for Environmental Assessment

EPA/600/R-08/076F
June 2009

Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines

Global Change Research Program
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

ABSTRACT

Climate and land-use change are major components of global environmental change with feedbacks between these components. The consequences of these interactions show that land use may exacerbate or alleviate climate-change effects. Based on these findings it is important to use land-use scenarios that are consistent with the specific assumptions underlying climate-change scenarios. The Integrated Climate and Land-Use Scenarios (ICLUS) project developed land-use outputs that are based on the Intergovernmental Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) social, economic, and demographic storylines and adapted these to the United States. ICLUS outputs are derived from a pair of models. A demographic model generates population estimates that are distributed by the spatial allocation model as housing density (HD) across the landscape; land-use outputs were developed for the four main SRES storylines and a base case. The model is run for the conterminous United States and output is available for each scenario by decade to 2100. In addition to maps of HD across the conterminous United States, this project also generated maps of impervious surface (IS) cover based on the HD projections. This report describes the modeling methodology for, some initial analyses using ICLUS outputs, and recommendations for further research.

Preferred Citation:
U.S. Environmental Protection Agency (EPA). (2009) Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines. Global Change Research Program, National Center for Environmental Assessment, Washington, DC; EPA/600/R-08/076F. Available from: National Technical Information Service, Springfield, VA, and online at http://www.epa.gov/ncea.

CONTENTS

CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS AND ACRONYMS

A1	The A1 storyline in the *Special Report on Emissions Scenarios*
A2	The A2 storyline in the *Special Report on Emissions Scenarios*
B1	The B1 storyline in the *Special Report on Emissions Scenarios*
B2	The B2 storyline in the *Special Report on Emissions Scenarios*
EPA	U.S. Environmental Protection Agency
GCM	General Circulation Model
GHG	greenhouse gas
GIS	Geographic Information System
HD	Housing Density
HUC	Hydrologic Unit Code
ICLUS	Integrated Climate and Land-Use Scenarios
IPCC	Intergovernmental Panel on Climate Change
IS	impervious surface
IS_{PUI}	percent urban impervious surface
MIGPUMA	Migration Public-Use Microdata Areas
NCHS	National Center for Health Statistics
NLCD	National Land Cover Database
NRCS	Natural Resources Conservation Service
NRI	Natural Resources Inventory
PUMA	Public-Use Microdata Areas
PUMS	Public-Use Microdata Samples
SERGoM	Spatially Explicit Regional Growth Model
SRES	Special Report on Emissions Scenarios
UPPT	unprotected, private protected, public protected, and tribal/native lands
USDA	U.S. Department of Agriculture
VMT	vehicle miles traveled

PREFACE

This report was prepared jointly by ICF International, Colorado State University, and the Global Change Research Program (GCRP) in the National Center for Environmental Assessment (NCEA) of the Office of Research and Development (ORD) at the U.S. Environmental Protection Agency (U.S. EPA). The report has undergone internal, public, and external peer review. Changes and edits between the draft and final report reflect comments made by reviewers. The report describes the methodology used to develop and modify the models that constitute the ICLUS project. The scenarios and maps resulting from this effort are intended to be used as benchmarks of possible land-use futures, focusing on residential housing, that are consistent with socioeconomic storylines used in the climate change science community. The two-way feedbacks that exist between climate and land use are not yet fully understood and have consequences for air quality, human health, water quality, and ecosystems. This report lays the foundation for characterizing and assessing the effects of these feedbacks and interactions by developing residential HD scenarios and deriving IS cover from them. These outputs facilitate future integrated assessments of climate and land-use changes that make consistent assumptions about socioeconomic and emissions futures. Outputs will also be distributed through the Web and a Geographic Information System (GIS) tool that allows some customization of scenarios. EPA's intention is to use the results of this first phase of modeling to inform and facilitate investigation of a broader set of impacts scenarios and potential vulnerabilities in areas such as water quality, air quality, human health, and ecosystems. More specifically, this research will enable more sophisticated model runs that will evaluate the effects of projected climate changes on demographic and land-use patterns and the results of these changes on endpoints of concern.

AUTHORS AND REVIEWERS

The GCRP, within NCEA, ORD, is responsible for publishing this report. This document was prepared by ICF International under Contract No. GS-10F-0234J, U.S. EPA Order No. 1101. Dr. Chris Pyke and Dr. Britta Bierwagen served as successive Technical Project Officers. Drs. Pyke and Bierwagen provided overall direction and technical assistance, and Dr. Bierwagen contributed as an author.

AUTHORS

U.S. EPA
Britta Bierwagen, John Thomas, Chris Pyke[*]

ICF International, Washington, DC
Anne Choate, Jonathan Cohen, Philip Groth

Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO
David M. Theobald

REVIEWERS

This report benefited from the comments and suggestions of the following reviewers.

EPA Reviewers:
Cynthia Gage, Ellen Cooter, Sandy Bird, Laura Jackson, Matt Dalbey, Brooke Hemming, Henry Lee

External Reviewers:
Dan Brown, University of Michigan; Steven Manson, University of Minnesota; Dawn Parker, George Mason University; David Skole, Michigan State University.

ACKNOWLEDGMENTS

The authors thank the Global Change Assessment Staff in NCEA, particularly Chris Weaver and Philip Morefield, for their input, advice, and assistance on this project. Randy Freed also provided valuable input and advice. We also thank John Norman for his assistance. Comments on earlier versions of this report by EPA and external reviewers greatly improved the final product.

[*]Present affiliation: CTG Energetics, Washington, DC

EXECUTIVE SUMMARY

Climate and land-use change are major components of global environmental change. Assessments of impacts associated with these changes often show interactions and two-way feedbacks between climate and land use. The consequences of these interactions also show that land use could exacerbate or alleviate climate-change effects. Based on these observations it is important to use land-use scenarios that are consistent with the assumptions underlying recognized international climate-change scenarios.

The ICLUS project developed outputs that are based on the IPCC SRES social, economic, and demographic storylines. We adapted these storylines to the United States and modified U.S. Census Bureau population and migration projections to be consistent with these storylines. ICLUS outputs are derived from a pair of models: a demographic model that generates population projections and a spatial allocation model that distributes projected population into housing units across the landscape. The demographic model is at the county scale and the spatial allocation model is at the 1 ha pixel scale (though results of analyses could be aggregated to larger scales). Each scenario is run for the conterminous United States to 2100 by decade.

The results of this first phase of the project provide a foundation for evaluation of a broader set of impacts scenarios and potential vulnerabilities in areas such as water quality, air quality, human health, and ecosystems. More specifically, these scenarios will underlie more sophisticated model runs that will evaluate the effects of projected climate changes on demographic and land-use patterns and the results of these changes on endpoints of concern. The products generated in this first phase are consistent with socioeconomic storylines used by the climate change modeling community, but climate-change variables are not integrated into the models in this phase. Outputs are also not specifically designed to feed into climate change models.

The ICLUS project uses the SRES storylines because these storylines are utilized as direct inputs into general circulation models developed by the climate change science community. These storylines were selected to facilitate future, more integrated assessments of climate and land use at national or regional scales, because the broad underlying assumptions are the same. The SRES describe storylines along two major axes, economic versus environmentally-driven development (A-B) and global versus regional development (1-2); the four quadrants defined by these axes comprise the four storylines, A1, A2, B1, and B2. We adapted these storylines for the United States to inform our changes to variables such as fertility, domestic and international migration, household size, and travel times from the urban core.

The demographic model is composed of five components, fertility, mortality, domestic in-migration, domestic out-migration, and net international migration, which are calculated using a cohort-component model and a gravity model. The population is divided into cohorts that are age-, gender-, and race/ethnicity-specific. Changes due to these five components of change are estimated over time as each cohort is tracked separately. The gravity model is used to track domestic in- and out-migration by county. Components of the gravity model include certain county amenities and functional distance that connects counties based on population locations and transportation infrastructure. The resulting county-based population projections are the inputs to the spatial allocation model.

The spatial allocation model distributes the population into housing units across the country at a 1 ha pixel scale. The model used in this project is the Spatially Explicit Regional Growth Model (SERGoM). SERGoM uses five main base datasets: housing units and population based on the 2000 U.S. Census, undevelopable lands, road and groundwater well density, commercial/industrial land use from the National Land Cover Database, and county population projections from the demographic model described above. Household size and travel times are adjusted to reflect the assumptions of the different U.S.-adapted SRES storylines. These modifications result in different spatial allocations among the scenarios such that the B storylines show more compact growth focused around urban areas and the A storylines show less compact growth overall and more housing in suburban and exurban densities.

The scenarios result in a range of projected increases in urban and suburban area across the United States. The smallest increase is 60% for the B1 scenario and the largest increase is 164% for A2. These increases in housing can be translated to changes in IS cover, which can be used to examine impacts on water quality, for example. Our results show that there could be a significant increase of watersheds (8-digit Hydrologic Unit Codes [HUCs]) that are likely to be stressed from IS coverage of at least 5%. These changes will vary regionally across the country.

This report describes the modeling methodology for the ICLUS project and some initial analyses using the outputs. There are many additional modifications that are possible to explore additional land-use futures and there are many options for further research. Model modifications can be made to further explore policy and planning alternatives such as Smart Growth or other low-impact development patterns. The demographic and spatial outputs can be used in numerous analyses examining potential future impacts on air quality and human health, water quality, traffic and associated emissions, and regional growth rates.

1. INTRODUCTION

Climate change and land-use change are global drivers of environmental change. Impact assessments frequently show that interactions between climate and land-use changes can create serious challenges for aquatic ecosystems, water quality, and air quality. In many cases, it is impossible to determine the impact of climate change without consideration of land use and land cover dynamics. While land use can exacerbate climate impacts, land-use planning, policy, and management can also create important adaptation opportunities to increase the resilience of sensitive socioeconomic or ecological systems.

Integrated assessments of both climate and land-use changes currently are limited by fragmented information on potential future land use. In many cases individual municipal areas have conducted extensive analyses, but it is impossible to place results in regional or national contexts. Moreover, the studies are often based on inconsistent or poorly documented socioeconomic storylines. The motivation for the U.S. Environmental Protection Agency (EPA) Integrated Climate and Land-Use Scenarios (ICLUS) project was derived from the recognition of the complex relationships between land-use change and climate change impacts and the absence of an internally consistent set of land-use scenarios that could be used to assess climate change effects.

This report describes the first phase of this project, which is designed to provide a foundation for the evaluation of a broader set of impacts scenarios and potential vulnerabilities in areas such as water quality, air quality, human health, and ecosystems. More specifically, the land-use scenarios described here can be used as inputs to more sophisticated model runs that evaluate the effects of projected climate changes on demographic and land-use patterns and the results of these changes on endpoints of concern. The products generated in this first phase are consistent with socioeconomic storylines used by the climate change modeling community, but this first phase does not explicitly integrate climate change variables into the models. Outputs are also not specifically designed for incorporation into climate change models.

The ICLUS project developed scenarios of housing density (HD) and derived impervious surface (IS) cover from these scenarios, for the entire conterminous United States for each decade through 2100. These scenarios are based on the Intergovernmental Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) social, economic, and demographic storylines (Nakićenović, 2000), because these storylines are used as direct inputs into general circulation models (GCMs) used by the climate change science community. This link facilitates any future, more integrated assessments of climate and land use at national or regional scales, because the broad underlying assumptions are the same. The ICLUS scenarios are developed using a combination of models representing demography, including domestic and international migration, and spatial allocation of housing (Figure 1-1). The resulting scenarios (1) will enable us, our partners, and our clients to conduct assessments of both climate and land-use change effects across the United States; (2) provide consistent benchmarks for local and regional land-use change studies; and (3) can be used to identify areas where climate-land-use interactions may exacerbate impacts or create adaptation opportunities.

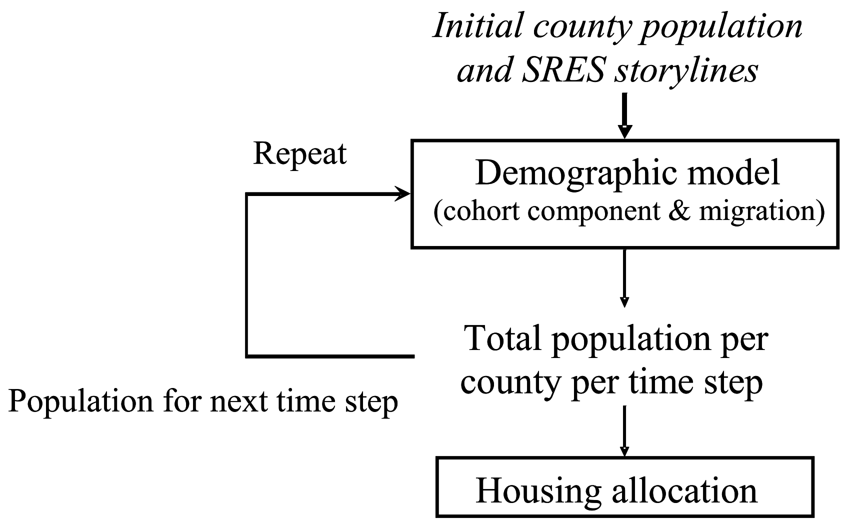

Figure 1-1. Model and information flow within the ICLUS project.

EPA's Global Change Research Program in the National Center for Environmental Assessment of the Office of Research and Development began investigating the availability of state and county-level population projections in 2004. Initial efforts evaluated the availability, sources, and extent of state and county-level population projections with an emphasis on identifying projections for the time period from 2050 through 2100. These efforts yielded numerous datasets, but very few sources projected populations to 2050 and beyond, particularly at the county level. Demographic projections for these later years are particularly relevant when considering the impacts of climate change on ecosystems, water infrastructure, transportation infrastructure, and land protection efforts. Population and land-use projections based on economic factors such as regional income and employment growth can shift dramatically over time. Projections based on fundamental demographic drivers such as fertility and mortality are somewhat more stable, particularly over longer time frames. Therefore, the ICLUS project uses demographic projections as the basis for modeling changes in HD.

The modeling framework (Figure 1-1) presented in this report uses demography to drive the number and migration of people, while a spatial allocation model governs the distribution of people on the landscape in housing units. The demographic model has two parts, a cohort-component model and a gravity model. The spatial allocation is conducted using an established Geographic Information System (GIS)-based model, Spatially Explicit Regional Growth Model (SERGoM) (Theobald, 2001, 2003, 2005). The study was designed to provide county-level demography and HD for the conterminous United States with HD allocations for each 1 ha pixel for each decade through 2100.

With countless variables and unforeseen events affecting demographic and land-use change patterns, it is impossible to quantify the level of uncertainty associated with these projections. In addition to general concerns about model accuracy, this approach does not capture some important attributes of land-use change that will likely confront the U.S. landscape in the future, such as regional development restrictions, local zoning laws, potential regional growth limitations due to water scarcity, changes in national immigration policy, and unforeseen socioeconomic limitations to development. Therefore, the scenarios presented in this study are designed to explore a range of possible futures. Given the scope of this task, the authors acknowledge the limitations and possibilities for additional evaluation in future analysis.

2. ADAPTING THE SRES TO THE UNITED STATES

The socioeconomic storylines in the SRES are derived from anticipated demographic, economic, technological, and land-use changes data for the 21st century, and are highly aggregated into four world regions (Nakićenović et al., 2000). The SRES describe linkages between physical changes in climate and socioeconomic factors, because they link development pathways with greenhouse gas (GHG) emissions levels used as inputs to GCMs (Rounsevell et al., 2006). There have been other scenario-development exercises to assess future impacts on a range of endpoints, including the Millennium Ecosystem Assessment (MEA, 2005), which also describes economic and environmental conditions in the future. The benefit of using the SRES storylines is their direct link to GCMs. This enables us to link our land-use projections to emissions and future climate scenarios for integrated assessments of the effects of land-use change and climate change in a consistent way.

Modeling and projecting human activity into the future is a challenge for many reasons. Any attempt to project demographic and economic changes over time must contend with large degrees of uncertainty. Furthermore, taking discontinuous events into consideration—ones that would have a profound effect on any anticipated trajectory—is even more difficult. Nevertheless, a forward-looking approach to environmental and economic problems encourages us to look into the future to attempt to better understand the challenges that lie ahead, and to better prepare our society to confront those challenges.

By taking a scenario approach to such modeling efforts, we acknowledge this inherent uncertainty and consider a variety of possible trajectories. This approach results in a range of outputs. No single output may be the "right" one, but together they paint a picture of a likely range of possible futures. The primary challenge to the scenario approach lies in developing a reasonable range of scenarios that can be used in multiple modeling efforts.

The emissions storylines in the SRES cover a wide range of possible paths for the primary social, economic, and technological drivers of future emissions. These storylines have since become the standard input of socioeconomic information for GCMs and other land-use change modeling efforts (e.g., Solecki and Olivieri, 2004; Reginster and Rounsevelle, 2006; Rounsevelle et al., 2006) providing a reasonable set of scenarios to bound the potential futures with respect to climate change. By using the SRES storylines as the basis for the scenarios investigated by this project, the results may be put into the context of widely available and peer-reviewed climate-change model output (e.g., IPCC, 2001, 2007).

2.1. OVERVIEW OF THE SRES STORYLINES

The development of SRES consisted of three main steps, beginning with qualitative "storylines" that describe broad economic, environmental, technological, and social development patterns that could unfold over the 21st century (Nakićenović et al., 2000). Next, particular quantitative paths for the fundamental driving forces of emissions, including population and gross domestic product, were selected for each storyline. Finally, six different modeling teams produced quantitative interpretations of the storylines, using the quantitative paths for driving forces as inputs, resulting in 40 different storylines for energy use, land use, and associated GHG emissions over the next 100 years (Nakićenović et al., 2000). The SRES describe storylines along two major axes, economic versus environmentally-driven development (A-B) and global versus regional development (1-2), which make up the four combinations of storylines, A1, A2, B1, and B2 (Figure 2-1). There are between 6 and 18 emissions scenarios within each of the four SRES storylines. Table 2-1 provides a summary of the qualitative fertility, mortality, and migration assumptions made by the SRES authors for each storyline for the industrialized country regions and the developing country regions; these qualitative assumptions served as the framework for the more quantitative inputs for the scenarios.

Based on descriptions in the SRES report (primarily in Sections 4.3 and 4.4 of the SRES report), a summary of the reasoning for these assumptions is provided below:

- In the A1 storyline, rapid economic development, associated with improved education and reduced income disparities, is assumed to drive a relatively rapid fertility decline in the high fertility regions. Global population is expected to rise until peaking in the middle of the century, after which fertility is generally below replacement level. Fertility in industrialized regions is assumed to follow a medium path at least in part so that, relative to the developing regions, the storyline is consistent with the assumption that social and economic convergence will lead to demographic convergence as well. For mortality, it is assumed that the conditions leading to low fertility are also consistent with relatively low mortality, so mortality is assumed to be low in all regions. No explicit discussion of migration is provided, although the projection eventually adopted assumes medium migration levels.

- In the A2 storyline, the regional orientation and slower rate of economic growth, limited flow of people and ideas across regions, and orientation toward family and community values was judged to be consistent with a relatively high fertility in all world regions. Mortality was assumed to be high as well, based on the assumption that conditions leading to high fertility would also lead to relatively high mortality in all regions. Although the storyline describes a limited flow of people across regions, the storyline authors assumed medium migration flows, as in all other storylines.

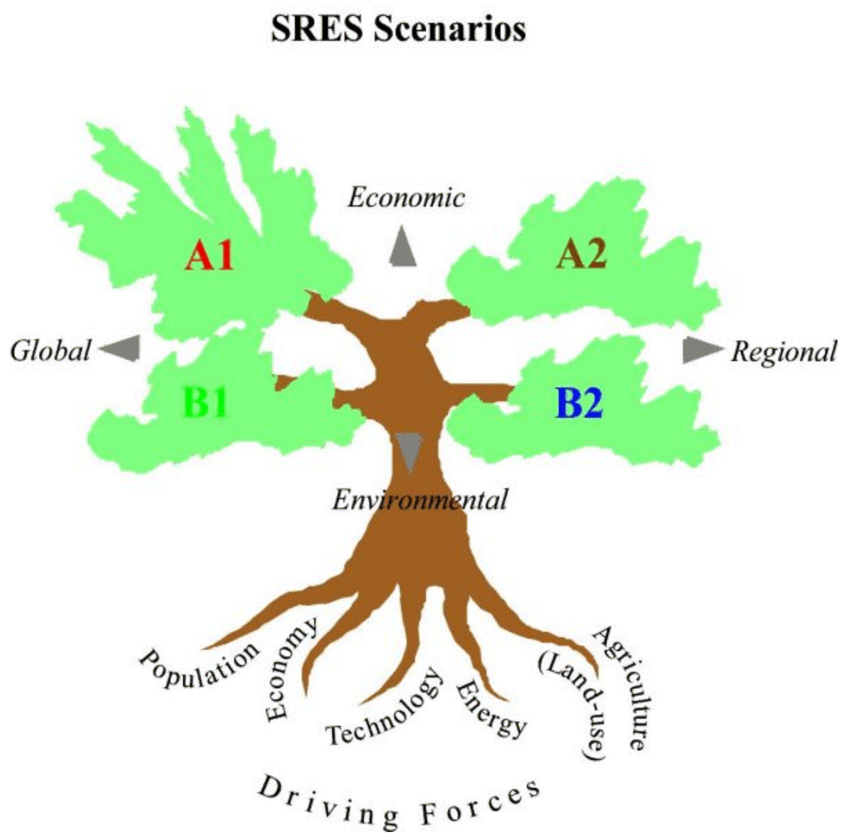

Figure 2-1. Illustration of the SRES families along two dimensions that indicate the relative orientation of the storylines along the axes of global or regional development and economic or environmental concerns (reprinted from Nakićenović et al., 2000).

Table 2-1. Qualitative demographic assumptions in global SRES storylines

Storyline	Fertility	Mortality	Migration	Projection source
A1/B1	IND: medium DEV: low	IND: low DEV: low	IND: medium DEV: medium	Lutz et al., 1996
A2	IND: high DEV: high	IND: high DEV: high	IND: medium DEV: medium	Lutz et al., 1996
B2	IND: medium DEV: medium	IND: medium DEV: medium	IND: medium DEV: medium	UN, 1998

IND = Industrialized country regions; DEV = Developing country regions.

The "high", "medium", and "low" descriptions are interpreted as relative to the overall outlook within each region (i.e., high fertility in the IND region means the high end of the plausible range for that region, but may in fact be lower than low fertility paths for the DEV region, which occupy the low end of the plausible range for the DEV region).

- The B1 storyline shares the same population projection as the A1 storyline, although for somewhat different reasons. Rapid social development, particularly for women, and an emphasis on education drives a relatively rapid decline in fertility in developing country regions (as opposed to the A1 storyline, in which economic development is seen as the main driver). Reasoning for fertility in industrialized countries, and for mortality and migration assumptions, are the same as in A1.

- In the B2 storyline, economic development is moderate, particularly in the developing country regions. However education and welfare programs are pursued widely and local inequity is reduced through strong community support networks. The mix of moderate economic development and strong but heterogeneous social development results in an assumption of medium fertility and mortality paths. Migration is again assumed to be medium, with no explicit discussion of this choice.

Demographic assumptions in SRES are intended to be consistent with storylines and with other driving force assumptions. Consistent relationships among these factors mean that the demographic assumptions occur in the context of other socioeconomic trends in a way that is not at odds with established theory, the weight of historical experience, or current thinking in the literature on determinants of demographic trends. For example, a key factor differentiating population assumptions across SRES storylines is the assumed speed of the transition from high to low fertility in regions with relatively high current fertility. The transition occurs faster in the A1 and B1 storylines, and slowest in the A2 storyline. These choices are based on the rationale that there are a range of conditions that contribute to fertility transitions, including economic development, education, labor force opportunities for women, and the spread of ideas about modern lifestyles (Lee, 2003). These factors are present, and stronger, in the A1 and B1 storylines (in different combinations) and absent, or weaker, in the A2 storyline. Thus, in this case, there is a clear notion of consistency in which storyline elements can be said to favor preferentially a particular demographic outcome.

However, consistency does not mean that the assumed demographic trends are the only possible outcomes, or in some cases even the most likely outcomes, conditional on a particular storyline. In some cases, storylines serve only as weak constraints on demographic futures, and a wide range of demographic assumptions might all be consistent with the broader development trends. For example, the demographic transition reasoning just described applies only to countries with relatively high fertility (e.g., substantially above replacement level of about two births per woman). These conditions occur for only about half the current population of the world, and for only the first half (or even less) of the 21st century, according to projections (Lee, 2003). Once the transition to low fertility is complete, there is little theoretical basis for linking subsequent fertility changes or cross-country differences to particular socioeconomic trends; a wide range of outcomes is possible (O'Neill, 2005a). Yet SRES storylines link the pace of economic growth with fertility outcomes (i.e., demographic transition-type reasoning) for all regions of the world, and for the entire century. It is certainly possible that rapid economic growth will be associated with relatively low fertility (and slow growth with high fertility) in post-transition societies, and it is not inconsistent in the sense of contradicting established theory (though this is at least partly because there is no single established theory to contradict). Completely opposite associations (e.g., low fertility with slow economic growth) are equally plausible (and equally consistent) for post-transition societies.

Thus in considering the implications of SRES storylines for demographic outcomes at the national and sub-national level, it must be kept in mind that it is possible that a wide range of different assumptions could be judged to be consistent with SRES. Indeed other studies have developed alternative demographic assumptions for SRES storylines, with different quantitative outcomes (Hilderink, 2004), or have quantified a range of plausible outcomes associated with storylines (O'Neill, 2004, 2005a,b).

2.2. INTERPRETING AND ADAPTING THE SRES STORYLINES

The SRES storylines do not provide a clear blueprint for downscaling to the local or even the national level. In incorporating the SRES storylines into county-level projections for the United States, an effort was made to be consistent in qualitative terms with the global SRES storylines. Given the wide range of potential interpretations, this consistency was understood to imply that the qualitative trends do not contradict established theory, historical precedent, or current thinking. It was also a goal to model a wide a range of assumptions, while remaining consistent with the SRES and U.S. demographic patterns. Rationales connected to SRES storylines are discussed briefly for each scenario. For each of the storylines adapted to the United States, the fertility assumptions are exactly consistent with the global assumptions, while domestic and international migration patterns leave more room for interpretation and are more specifically adapted to the United States. The low U.S. Census scenario for mortality was chosen for all storylines used in the modeling (see Appendix B for more information). These model inputs were varied to develop the different scenarios rather than to investigate the relative importance of each of the inputs.

2.2.1. A1 Storyline Adapted for the United States

A1 represents a world of fast economic development, low population growth, and high global integration. In this storyline fertility is assumed to decline and remain low in a manner similar to recent and current experience in many European countries (Sardon, 2004). A plausible rationale would be that the rapid economic growth in this storyline leads to continuing high participation of women in the workforce, but it becomes increasingly difficult to combine work with childbearing due to inflexibilities in labor markets. At the same time, social changes in family structures lead to increasing individuation, a rise in divorce rates, a further shift toward cohabitation rather than marriage, later marriages and delayed childbearing, all of which contribute to low fertility. Substantial aging resulting from the combination of low birth rates and continued low death rates raises the demand for immigration. Meanwhile, economic growth throughout the world and an increasingly unified global economy encourage the free movement of people across borders. Domestic migration is anticipated to be relatively high as well, as economic development encourages a flexible and mobile workforce.

2.2.2. B1 Storyline Adapted for the United States

This storyline represents a globally-integrated world similar to A1, but with greater emphasis on environmentally sustainable economic growth. Like A1, fertility is assumed to be low due to higher incomes and economic development. International migration is expected to be high due to widespread economic development and freer global flows. Domestic migration however is lower due to a combination of factors. First, an increased focus on sustainability leads to a reduction of subsidies for development in previously rural counties with significant natural amenities. Second, the information oriented economy increases demand for specialized labor pools and increases the number of high paying jobs in traditional large urban centers.

2.2.3. A2 Storyline Adapted for the United States

The A2 storyline represents a world of continued economic development, yet with a more regional focus and slower economic convergence between regions. Fertility is assumed to be higher than in A1 and B1 due to slower economic growth, and with it, a slower decline in fertility rates. International migration is assumed to be low because a regionally-oriented world would result in more restricted movements across borders. Domestic migration is high because, like in A1, the continued focus on economic development is likely to encourage movement within the United States.

2.2.4. B2 Storyline Adapted for the United States

The B2 represents a regionally-oriented world of moderate population growth and local solutions to environmental and economic problems. Fertility rate is assumed to be medium, while international and domestic migration are low due to the local emphasis, focus on sustainability, and increasing number of jobs in urban centers. International migration is low due to the regional orientation, as with A2. Domestic migration is low due to the more environmental orientation, as with B1.

The SRES storylines adapted to the United States were integrated into our modeling framework. We modified both the demographic and spatial allocation models for each of the four scenarios.

3. DEMOGRAPHIC PROJECTIONS

3.1. OVERVIEW

The ICLUS demographic model utilizes a cohort-component methodology. The cohort-component methodology is a common technique for projecting population on the basis of five independent variables or components of population change: fertility, mortality, domestic in- and out-migration, and net international migration. The population is divided into cohorts that are age-, gender-, and race/ethnicity-specific. Changes due to these five components of change are estimated over time as each cohort is tracked separately, hence the term "cohort-component" (Siegel and Swanson, 2004).

The methodology is flexible in that different assumptions can be applied to each component of population change. For example, fertility—or the number of births—is estimated by multiplying cohort-specific fertility rates (births per woman) times the population of each cohort of women. Different rates can be applied to women of different ages and ethnicities. Furthermore, these rates can change over time or between different scenarios. In all cases, the fundamental method stays the same while changes in the rates can be used to simulate different SRES storylines.

The population of a county in any year (*t*) as estimated by the model is determined using Equation 3-1:

$$P_t = P_{t-1} + B - D + NDM + NIM \tag{3-1}$$

Where:

P_t = Population in year t

P_{t-1} = Population in the previous year

B = Births in year t

D = Deaths in year t

NDM = Net Domestic Migration in year t

NIM = Net International Migration in year t

Beginning with an initial set of populations, annual components of change are applied in the following process, which are repeated annually until the desired end year is reached.

1) Add births by cohort
2) Deduct deaths by cohort

3) Add net international migration
4) Add net domestic migration by combining domestic in-migration and out-migration (Estimated every fifth year, as discussed below.)
5) Age population one year and repeat for the next year

This methodology is illustrated in Figure 3-1 below. The cycle begins with an initial Year 2005 population and is repeated until reaching Year 2100.

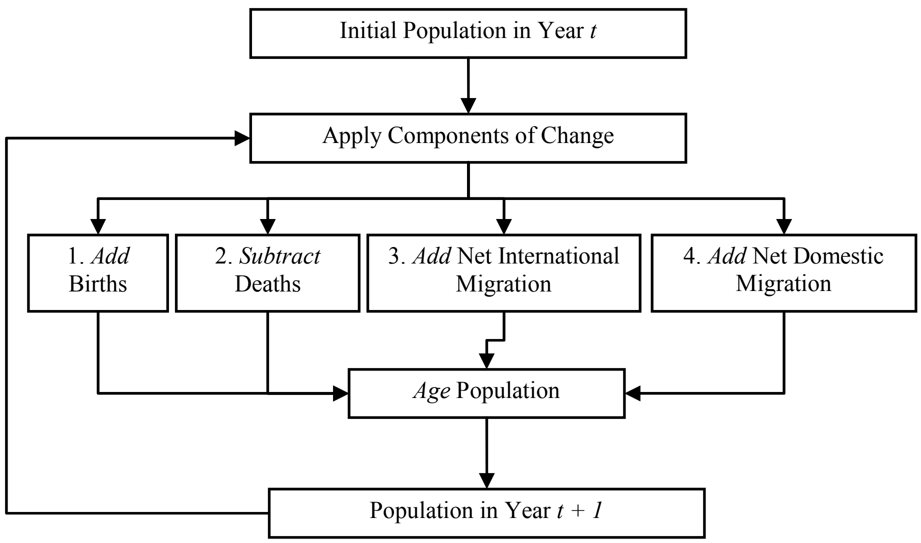

Figure 3-1. Demographic model flow.

Table 3-1 provides a qualitative description of how we modified the global storylines to the United States. As a general observation, note that the proposed fertility and mortality assumptions follow the SRES assumptions, with the exception that the A1 storyline assumes low fertility (rather than medium as in SRES). In addition, the international migration assumptions assume high and low immigration in storylines B1 and A2, respectively (rather than medium as in SRES). These choices are made in order to explore a fuller range of demographic trends for the United States than in SRES, since the SRES storylines contained only a limited range of projections for the North America region (O'Neill, 2005a). It should also be noted that these assumptions are not designed to cover the widest possible range of population size outcomes; e.g., the combinations resulting in the highest and lowest population sizes are not explored here. In all cases the mortality rate was kept constant across all scenarios. This was partly due to data availability—the U.S. Census did not release alternate scenarios for mortality rates. Experiments with adjusting these mortality rates (shown in Figure B-4 in Appendix B) demonstrated relatively little change in total national population due to variations in the mortality rate.

Table 3-1. Summary of qualitative adjustments to the demographic projections

Storyline	Demographic Model		
	Fertility	**Domestic migration**	**Net international migration**
A1	Low	High	High
B1	Low	Low	High
A2	High	High	Medium
B2	Medium	Low	Medium
Baseline	Medium	Medium	Medium

In the following sections, the methods and data used for the initial population and each component are discussed in greater detail. A summary of the data sources and the major assumptions used in the demographic model is presented in the Appendix in Table F-1.

3.2. INITIAL POPULATION

In order to use the rates for components of change provided by the U.S. Census Bureau (U.S. Census Bureau, 2000) (discussed below), it was necessary to begin with an initial Year 2005 population dataset that was disaggregated using the same cohorts. These cohorts in the rates data are divided into two genders (male and female), 100 age groups (0−99 in one year increments), and five racial/ethnic—or "bridged race"—groups (Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic American Indian or Alaskan Native, and non-Hispanic Asian or Pacific Islander).[1] This represents 1,000 distinct population cohorts (2 genders x 100 ages x 5 bridged race groups).

County populations using bridged race and one-year age cohorts were most readily accessible using the Bridged-Race Vintage 2006 dataset for July 1, 2005 provided by the

[1]In general, the U.S. Census Bureau considers the primary racial categories to be American Indian/Alaskan Native, Asian/Pacific Islander, Black, and White. "Hispanic origin" is considered an ethnic category. The race and ethnicity categories used by the Census have changed over time. In the 2000 Census, participants were allowed to identify with two or more racial groups for the first time. This project utilized the "bridged race" categories listed above as a way of making data collected with one set of categories consistent with data collected using another set of categories.

National Center for Health Statistics.[2] After downloading and parsing the data, two manipulations were required. First, the dataset provided eight bridged race categories—the non-Hispanic populations of the four race groups and the Hispanic populations of the four race groups. The Hispanic populations were summed and combined into a single bridged race group. This group, along with the four non-Hispanic race groups, comprised the five bridged race groups used throughout the model.

Second, the 2000 Census dataset provided one-year age groups from ages 0−84 and combined all others in the 85+ age group. In order to extend the age distribution through 99 years, the 85+ population data was disaggregated into ages 85−99. We used the national age 85−99 populations by race and gender from the 2000 Census to allocate the 85+ age group from the NCHS data. We assumed an identical race- and gender-specific age distribution in each county for the Year 2005 base population. Although this step likely led to some distortions, the effects are not long-lasting in the model as the initial 85−99 age groups "age out" of the model by 2015.

3.3. FERTILITY AND MORTALITY

Fertility and mortality followed a simple methodology. For fertility, the number of children born equals the number of women in a given cohort times the average number of children born annually to every 1,000 women in that cohort divided by 1,000. Because virtually all births occur to women between the ages of 10 and 49, only those cohorts are considered in this model. These births are summed by race and used to create a new age zero cohort. To allocate births between males and females, we calculated a historic ratio of 1,046 males born for every 1,000 females born and assumed that this ratio holds steady (Matthews and Hamilton, 2005).

Similarly, mortality is estimated by multiplying the number of people in a given cohort times the cohort-specific mortality rates. The resulting number of deaths is then subtracted from the cohort. Unlike fertility, all cohorts are subject to mortality. Therefore, mortality rates are applied to each cohort. Although an increasing number of Americans is living to the age of 100 or more, the model assumes 100% mortality after age 99 for the sake of computational efficiency. Even with continued rates of survivorship past this age, the 100+ age group will remain a miniscule portion of the population.

---

[2] NCHS (National Center for Health Statistics). (2007) Bridged-race Vintage 2006 postcensal population estimates for July 1, 2000 - July 1, 2006, by year, county, single-year of age, bridged-race, Hispanic origin, and sex. Released August 16, 2007. Centers for Disease Control and Prevention (CDC), Atlanta, GA. Available online at http://www.cdc.gov/nchs/about/major/dvs/popbridge/datadoc.htm#vintage2006.

For fertility and mortality rates, the U.S. Census Bureau's "Component Assumptions of the Resident Population by Age, Sex, Race, and Hispanic Origin" were used (U.S. Census Bureau, 2000). These are the same data used in Census projections. These components of change are associated with the 1990 National Projections and are used in both the 1990 State Projections (Campbell, 1996) and the 2000 National Projections (Hollman et al., 2000). While it would be preferable to use more recent data, at this time components of change based on the 2000 Census have not yet been released. While the rates are national averages, county differences that arise in fertility and mortality are a reflection of each county's unique age, sex, and racial composition.

For both fertility and mortality, the so-called Middle Series of component information was used as the baseline. Fertility rates are provided in a single file; mortality rates for each component are provided in three different tables, for the years 1999−2010, 2015−2055, and 2060−2100. Projected fertility rates are provided for each year to 2100, but beginning with 2010, mortality rates are provided in five year increments only. We assumed that 2010 mortality rates held steady from 2010−2014, 2015 mortality rates held steady from 2015−2019, and so on.

The Census Bureau also provides a low and high scenario for fertility. Alternative scenarios for mortality were not available. While the middle series was used in this "base case," the low and high series were used when developing projections specific to the SRES storylines. The low data set was used for the A1 and B1 scenarios, the medium set was used for the B2 scenario, and the high set was used for the A2 scenarios.

3.4. NET INTERNATIONAL MIGRATION

The projections for net international migration utilized a simple method based on the U.S. Census Bureau's international migration projections for the entire country. These files contain the projected net international migration for each gender, age, and race cohort for the years 2000−2100. Like the fertility and morality rates, these data are provided by the Census (U.S. Census Bureau, 2000).

Since the tables "Foreign-born Net Migration to the United States" contain only national level data, it was necessary to allocate the national migrants to the counties. Using 2000 Census data (U.S. Census Bureau, 2007, Summary File 3, Table P22), we determined each county's share of the total population of recent immigrants (i.e., those who entered within the last five years). These county shares were then used to allocate each cohort of immigrants among the nation's counties. The estimated number of immigrants in each cohort was then added to the existing county population of each cohort. This method assumes a constant distribution of recent immigrants based on Year 2000 immigration patterns. While we anticipate that many of the current "gateway counties" will continue to draw a large share of new immigrants, it is likely that new settlement patterns will develop in the future. A more complex international migration component that allows for changes in the distribution pattern of immigrants may be utilized in future improvements to the model.

The Census Bureau provides a low, medium, and high series for net international immigration. In the base case, the middle series was used. The medium and high series were used when developing SRES storyline-specific projections. The high series was used for the A1 and B1 storylines, while the medium series was used for the A2 and B2 storylines. The low set was not used because the values projected under the U.S. Census Bureau's low set were so far below current levels of immigration (Figure 3-2, U.S. Census Bureau, 2008) that they were considered to be unrealistic for these purposes.

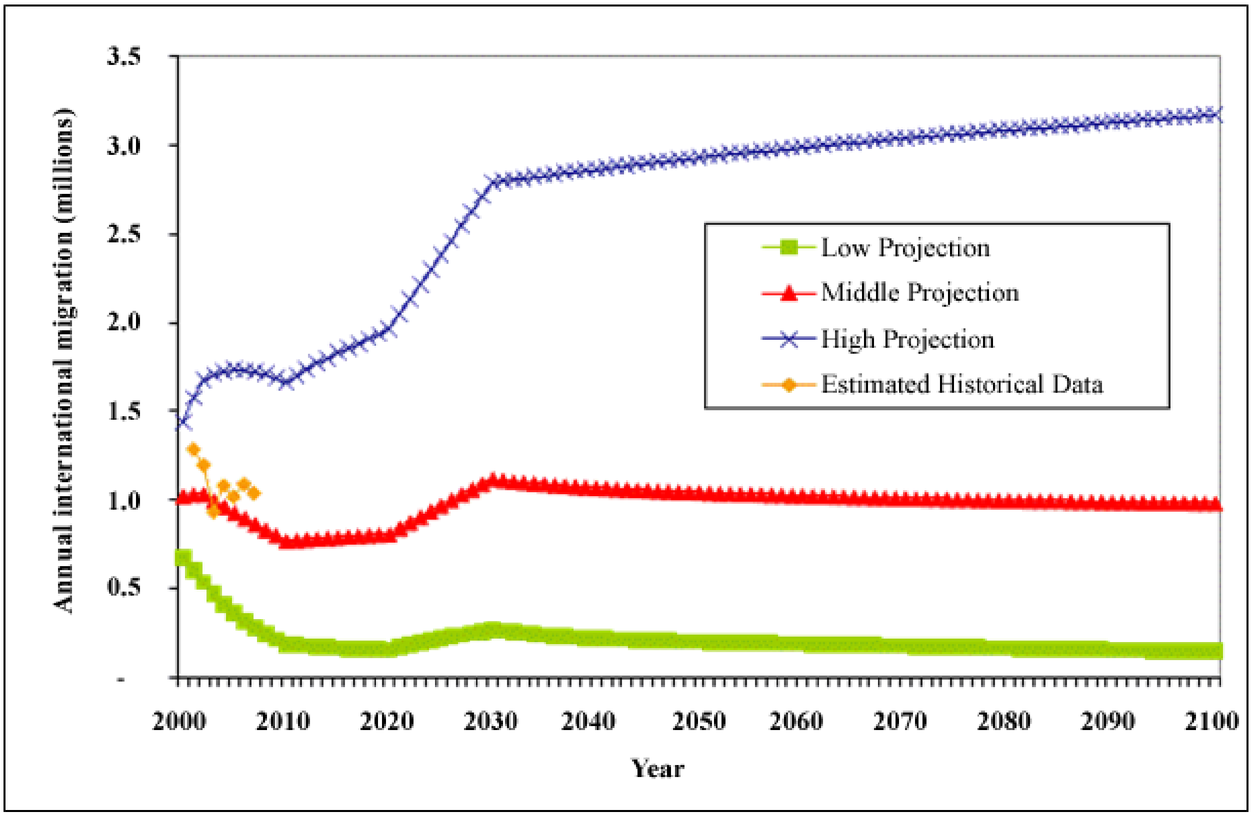

Figure 3-2. Projected net international migrations under U.S. Census Bureau Scenarios.

3.5. DOMESTIC MIGRATION AND GRAVITY MODELING

Domestic migration is the most complex component of change included in this population model. In contrast to straightforward processes like fertility and mortality, which may be predicted in aggregate with reasonable confidence, domestic migration is much more difficult to predict. People move within the United States with relative frequency—according to did in 1995 (Perry and Schachter, 2003). These migrations occur for a wide variety of personal and economic reasons that are difficult to predict. Unlike international immigration, which can be thought of as a single external source of immigrants, domestic migration works two ways—migrants must choose to leave one place and resettle in another place.

To estimate domestic migration, this project utilized a "gravity model" approach. Gravity models are a common type of spatial interaction model in which a dependent variable (in this case, migration flow between a pair of counties) is estimated as a function of a series of independent variables (Rodrigue et al., 2006). Common independent variables include population and distance (Haynes and Fotheringham, 1984); natural amenities, such as climate variables and bodies of water, have also been shown to affect migration patterns, particularly in rural areas (McGranahan, 1999). Among these, temperate summers and warm winters have the strongest correlation with population growth ($R^2 = 0.38$ and $R^2 = 0.27$, respectively), a finding that holds true even when correcting for macro-trends such as the general movement from the Northeast and Midwest to the South and West (McGranahan, 1999). Other work has also shown that amenity-rich mountain and coastal areas are highly successful at attracting migrants, particularly those who are affluent and/or retired (Manson and Groop, 2000). While a wide variety of independent variables were tested, the final gravity model used in this effort included the following independent variables: population of the origin and destination counties, selected climate variables of the origin and destination counties, water surface area of the origin and destination counties, ratio of 2000 county population to 1980 population of the destination county, and the distance between counties. This first phase of the modeling does not attempt to project changes in climate variables over time.

The final model form and model coefficients were estimated by analyzing 1995 to 2000 migration data reported in Public-Use Microdata Samples (PUMS) (U.S. Census Bureau, 2003). Using county-to-county migration data as the dependent variable and the various county attributes as dependent variables, we ran a series of stepwise regression analyses to select the most statistically significant set of dependent variables and estimate the model coefficients. Below, we discuss how each data source for the regression analyses was developed, the resulting model form, and the way this model was implemented in the overall demographic model.

3.5.1. 1995−2000 Migration Data

The 2000 Census PUMS data provides detailed records of individual domestic migrations and characteristics of these migrants between 1995 and 2000 (U.S. Census Bureau, 2003). However, the raw data does not readily provide county-to-county migration counts. Because the migrations are organized by the destination state, in-migrants to any given state can be analyzed using that state's PUMS file, but analysis of out-migration requires analyzing data from all

50 states, the District of Columbia, and Puerto Rico. This project used a national migrant file created by the New York State Data Center from the individual state and area files.[3]

The second challenge with the PUMS data is that it is organized by Public-Use Microdata Areas (PUMAs). Counties and PUMAs overlap in a non-systematic way. Large urban counties consist of multiple PUMAs, while small rural counties are often combined into a single PUMA. To further complicate the issue, the spatial area that PUMS uses to capture migration origins (where respondents lived in 1995), the so-called Migration PUMA (MIGPUMA), is not the same spatial area as the one used to capture where they lived in 2000 (the PUMA). The MIGPUMA is a larger area, and there can be one or more PUMAs contained within the MIGPUMA.

Because we required a database that contains migration, attraction factors, production factors, and distance information for pairs of areas, introducing the MIGPUMA into our spatial framework would have required re-calculating all of the other variables at the MIGPUMA level, and aggregating the migration data (destinations) to the MIGPUMA level, losing considerable local detail. Instead, we elected to proportionately assign MIGPUMA origins to its constituent PUMAs, on an average basis.

Furthermore, retaining a county-based framework is practical because the amenity data is available only for counties and counties are a more popularly understood and conceptualized units than PUMAs. For the sake of consistency and simplicity, we then proportionately assigned PUMA origins and destinations into counties based on the proportion of each PUMA's population belonging to one or more counties. In the case of large metropolitan and suburban counties, several PUMAs were typically aggregated to a single county, while in smaller suburban counties and rural areas, PUMAs were often split among two or more counties. Because the largest migration flows involve large urban and suburban population centers where PUMAs were aggregated up to the county level, it was assumed that this step did not introduce significant error into the migration analysis. By converting from MIGPUMA to PUMA to county, the migration data set was used to estimate county-to-county flows needed for this study.

These county-to-county flow data were then aggregated into various age groups. Since migration behavior is hypothesized to vary for different age groups, different gravity models were calibrated for different age cohorts. In initial runs, the population was divided into ten-year cohorts, from 0–9 to 90 and up, and a different gravity model was estimated for each. We observed some broad differences in behavior between the older (over 50) age groups and the younger age groups, presumably due to differences in migration patterns due to retirement. However, the differences between the ten individual models were not found to be significant

---

[3] The website for the New York State Data Center is http://www.empire.state.ny.us/nysdc/default.asp.

enough to warrant ten different age groups. As a result, the dataset was recoded into just two age groups: 0−49 and 50+.

3.5.2. County Attribute Data

After recoding the migration data set from the original geographies to counties, we then added the county data that would be included as independent variables in the analysis. These variables include county environmental amenities, population, and 1980−2000 population growth, which serves as a proxy for economic growth.

The source for the amenity information used in the regression was found in a database compiled by the U.S. Department of Agriculture (USDA) (McGranahan, 1999), which utilizes data originally collected from the Center for National Health Statistics, U.S. Department of Health and Human Services. The amenity index includes a range of climatic and amenity factors that are thought to influence migration. The data used from this set include:

- Mean temperature for January, 1941−70;
- Mean hours of sunlight for January, 1941−70;
- Mean temperature for July, 1941−70;
- Mean relative humidity for July, 1941−70; and
- Percent water area.

County population data were drawn from the U.S. Census Bureau's *City and County Data Book*, 2000 edition. The population over age 50—used in analyses of the 50−99 age group—was determined from data downloaded from the 2000 U.S. Census (Summary File 1, Table P12). The 2000 county population data, combined with 1980 and 1990 county populations provided by the Census Bureau (U.S. Census Bureau, 1992), were used to calculate the growth rate of counties between 1980 and 2000. This term was added after initial model runs resulted in declining populations in medium-sized counties that exhibited strong growth in the 1980s and 1990s.

3.5.3. Distance Matrix

A full county-to-county distance matrix was the final data input used in developing the gravity model. Typical methods of migration movement between geographic locations (e.g., city to city or county to county) assume that interaction can be estimated as straight-line county centroid-to-centroid distance (Conley and Topa, 2002). Two main deficiencies are that: (1) often the geographic centroid of a county poorly represents the population-weighted centroid of a county; and (2) humans do not move around on the ground optimally using a straight-line strategy, rather we are constrained to transportation infrastructure which in turn has evolved in response to major topographic and water features.

Because of these deficiencies, we developed a "functional" estimate of county-to-county interaction. The main assumptions incorporated into the county-to-county functional distance calculations are that the amount of interaction, and more specifically migration, between counties are better approximated by examining: (a) where in a county population is located, and (b) how much ground- (and water-) based transportation infrastructure is in place.

The geographic centroids of each of 25,150 U.S. Census Bureau-defined places (within conterminous United States) were found and assigned to the county in which they were located. The population (2000 Census) was then used to weight each point, and the population-weighted centroid for each county was determined. For roughly a half-dozen counties, the population-weighted centroid was calculated to be outside the boundaries of its respective county. For these, the centroid was manually moved back into its respective county. We also generated a centroid for a couple of counties in Nevada that did not have places in them (visually in the center).

We generated a cost-weight surface using major roads such that the weights were assigned the assumed average speed limit assigned by road type. Cost distance was then computed using the population-weighted centroids as the "seeds," computing minutes travel time (T) along the roads. For each adjacent (first-order neighbor) county-county pair (i,j) we adjusted the travel time to account for k multiple roads (multiple connections between population centroids) to compute an interaction weight (W). Equation 3-2 shows the calculation for the interaction weight (W) to generate the cost-weight surface for county pairs.

$$W_{ij} = 1/\left(\sum_k (1/T_k)\right) \tag{3-2}$$

Where:

W_{ij} = Interaction weight for the county-county pair i, j

T_k = Travel time for k multiple roads

We then generated a network and used a network-based least-cost path algorithm to compute effective distances along paths of pair-wise segments using the FunConn tools (Theobald et al., 2006). We developed a distance matrix of pairwise functional distances (W_{ij}) computed in minutes travel time. We exported this matrix from the GIS database in a list of from-county, to-county, weight format (over 9 million records for 3109 counties times 3108 destination counties) to the gravity model data table. We also manually added linkages between counties that are served with regular ferry service as delineated in the U.S. Transportation Atlas 2006 (e.g., across Lake Michigan and Nantucket Island Sound).

3.5.4. Model Specification

After collecting the above data, it was reorganized into a single data table. Each record in the data table contained the number of migrations from one county to another, the attributes of the origin county, the attributes of the destination county, and the functional distance between them. In keeping with the traditional functional form of gravity models, which states migration is proportionate to attraction and production variables, and inversely proportional to distance, the functional form of the model used is shown in Equation 3-3:[4]

$$F_{ij} = \alpha \times P_i^{(\beta 1 + \beta 2 \times Pi)} \times Jt_i^{\beta 3} \times Js_i^{\beta 4} \times St_i^{\beta 5} \times Sh_i^{\beta 6} \times W_i^{\beta 7} \times P_j^{(\beta 8 + \beta 9 \times Pj)} \times Jt_j^{\beta 10} \times Js_j^{\beta 11} \times St_j^{\beta 12} \times Sh_j^{\beta 13} \times W_j^{\beta 14} \times G_j^{\beta 15} \times d_{ij}^{\beta 16} \tag{3-3}$$

Where:

F_{ij}	=	the migration of the population from county *i* to county *j* from 1995 to 2000
P	=	the population of a county in 2000. For the 0−49 age group, total population was used. For over-50 age group only the over-50 population was used
Jt	=	the mean January temperature for a county
Js	=	the mean January hours of sunlight for a county
St	=	the mean July temperature for a county (S = summer)
Sh	=	the mean July humidity for a county (S = summer)
W	=	the percent water area for a county
G	=	county growth rate, expressed as the ratio of 2000 population to 1980 population
d_{ij}	=	the functional distance between *i* and *j*
α, $\beta 1$, ... $\beta 16$	=	parameters estimated by the regression model.

[4]The parameters estimated by the regression model (provided below) are positive for those variables that are directly proportional to migration (e.g., population) and negative for those variables that are inversely proportional to migration (e.g., distance).

In order to estimate this equation using multiple regression, a logarithm of both sides is taken, which yields Equation 3-4, the linear form of Equation 3-3.

$$\begin{aligned}\log(F_{ij}) = {} & \log(\alpha) + \beta 1 \times \log(P_i) + \beta 2 \times P_i \times \log(P_i) + \beta 3 \times \log(Jt_i) + \\ & \beta 4 \times \log(Js_i) + \beta 5 \times \log(St_i) + \beta 6 \times \log(Sh_i) + \beta 7 \times \log(W_i) + \\ & \beta 8 \times \log(P_j) + \beta 9 \times P_j \times \log(P_j) + \beta 10 \times \log(Jt_j) + \beta 11 \times \log(Js_j) + \\ & \beta 12 \times \log(St_j) + \beta 13 \times \log(Sh_j) + \beta 14 \times \log(W_j) + \beta 15 \times \log(G_j) + \\ & \beta 16 \times \log(d_{ij})\end{aligned} \tag{3-4}$$

We also ran a variety of regionally-specific experimental runs to consider if different effects were observed on a regional basis. The four Census-defined regions are the Northeast, South, Midwest, and West. Although some differences were observed between the different regions, the marginal improvement in model fit offered by regionally-specific models was outweighed by the complexity of developing a separate model for each pair-wise combination of regions.

In the equations above, population is treated differently from other variables. Initial model runs used the form log(migration) = $a + b \times$ log(population) + other terms. Thereby, migration is proportional to (population)^b. Under this model, log(migration) always increases with log(population) at the same rate (assuming $b > 0$), all else being equal. When running the model over many decades, this caused the model to predict that most of the population of many small counties would migrate to nearby large counties. However, such a scenario is unrealistic as large counties grow increasingly crowded and the differential between land prices in the urban core and land prices in suburban/exurban areas grows. We then modified the model so that the slope (b) is not constant but varies slowly with the size of the population. The new model used $b = c + d \times$ population, providing the following form: log(migration) = $a + c \times$ log(population) + $d \times$ population × log(population) + other terms. With this revised equation, migration is proportional to (population)$^{(c + d \times \text{population})}$. As expected, the fitted c coefficient was positive while the d coefficient was small and negative, so that for small populations, the new model looks like the old model, but for large populations, log(migration) increases with log(population) at a slower rate. This model resulted in projections more consistent with expectations: with more modest growth, the largest populations did not grow in extreme proportions, while the suburban and exurban counties (particularly those in between two relatively close cities) grew the fastest.

Table 3-2. Gravity model results

Variable	Age group[a]	
	0−49	**$50+$[b]**
Adj. R^2	0.665	0.6591
Intercept [log(α)]	5.74014	3.14215
Production Population	0.7756	0.78429
$(\text{Production Population})^{\text{Production Population}}$	−9.12E−09	−5.63E−08
Production January Temp.	0.07301	N/A
Production January Sun	0.04666	−0.15858
Production July Temp.	−1.28061	−0.85029
Production July Humidity	−0.4482	−0.22306
Production Water Surface	0.01199	−0.00422
Attraction Population	0.85084	0.84382
$(\text{Attraction Population})^{\text{Attraction Population}}$	−1.58E−08	−8.83E−08
Attraction January Temp.	0.01362	0.09291
Attraction January Sun	0.06784	0.09263
Attraction July Temp.	−0.78317	−0.78248
Attraction July Humidity	−0.38647	−0.28732
Attraction Water Surface	0.0192	0.01334
Attraction 1980−2000 Growth Rate	0.30131	0.54938
Distance	−0.98919	−0.83684

[a]The variable result values correspond to the β-values in Equations 3-3 and 3-4.
[b]Population over age 50 was used in place of total population in the Age 50+ model.

All parameters were significant at the $p < 0.0001$ level, with the exception of Production January Temp. in the 50+ model, which was not significant.

Table 3-3. Ranking of the contribution of independent variables to explanatory power of the models

Rank	Ages 0−49	Partial R-square for each variable	Cumulative model R-square	Ages 50+*	Partial R-square for each variable	Cumulative model R-square
1	Attraction Population	0.3150	0.3150	Attraction Population	0.3315	0.3315
2	Production Population	0.2130	0.5280	Production Population	0.1976	0.5291
3	Distance	0.1231	0.6511	Distance	0.1143	0.6433
4	Production July Humidity	0.0052	0.6563	Attraction 1980−2000 Growth Rate	0.0062	0.6496
5	Attraction July Humidity	0.0030	0.6593	$(\text{Attraction Population})^{\text{Attraction Population}}$	0.0023	0.6518
6	$(\text{Attraction Population})^{\text{Attraction Population}}$	0.0023	0.6615	Attraction July Humidity	0.0025	0.6543
7	Production July Temp.	0.0010	0.6625	$(\text{Production Population})^{\text{Production Population}}$	0.0020	0.6563
8	Attraction 1980−2000 Growth Rate	0.0010	0.6636	Production July Temp.	0.0016	0.6579
9	$(\text{Production Population})^{\text{Production Population}}$	0.0007	0.6643	Production July Humidity	0.0007	0.6586
10	Attraction July Temp.	0.0004	0.6647	Production January Sun	0.0002	0.6588
11	Attraction Water Surface	0.0001	0.6648	Attraction July Temp.	0.0001	0.6589
12	Production January Temp.	0.0001	0.6649	Attraction January Temp.	0.0001	0.6590
13	Production Water Surface	0.0001	0.6650	Attraction Water Surface	0.0001	0.6591
14	Attraction January Sun	0.0000	0.6650	Attraction January Sun	0.0001	0.6591
15	Production January Sun	0.0000	0.6650	Production Water Surface	0.0000	0.6591
16	Attraction January Temp.	0.0000	0.6650			

*Population over age 50 was used in place of total population in the Age 50+ model.

3.5.5. Stepwise Regression Results

Table 3-2 provides a summary of the stepwise regression results, including the model R^2 and the individual parameters. The R^2 indicates the overall goodness-of-fit of the regression models, or what percentage of the variance of the logarithm of the flow is explained by the independent variables. The estimates for each parameter in the models are also provided. The regression modeling was performed using a stepwise approach, such that at each step a term was added if the significance level was <0.05 and terms were removed if the significance level was $\geq$0.05. For these models, all the regression parameters except for one regression parameter in the 50+ model were statistically significant at each step ($p < 0.0001$), i.e., the improvement in goodness-of-fit was statistically significant. The relative explanatory power of each variable is provided in Table 3-3. These results show that population and distance represent the majority of the model's explanatory power. While certain amenity variables may be more important in rural counties, when considered with much larger urban and suburban counties, the additional contribution of these variables is smaller, though still significant.

We also ran a correlation on each of the independent variables to determine the prevalence of collinearity among these variables. In nearly all cases, all pairs of variables were found to be statistically collinear at the $p < 0.0001$ level, although many of the estimated squared correlations (R^2) were small. With such an extremely large sample size (n = 2,397,007, or the number of migration records in PUMS), it is relatively easy for even very low R^2 values to be statistically significant. In this case, having a very significant *p*-value simply suggests that the data are inconsistent with a true correlation of zero, but they could be consistent with some very small, but non-zero, correlation.

3.5.6. Model Flow

Having estimated the gravity model parameters, the gravity model was incorporated into the overall demographic model to estimate domestic migration. Because the underlying migration data used in this analysis measured migration over a five-year period, the migration was estimated at five-year intervals.

Using a pair of nested loops, the model cycles through each pair of counties, estimating the migration from each county to every other county. Using the same amenity and distance data discussed in Section 3.5.1−3.5.3 and the current estimated county populations as calculated by the model, it enters the terms into the gravity model equation provided above for each of sets of model parameters for each of the two age groups.

Because the gravity model does not consider race, gender, or age beyond the two broad age groups, every estimation of migration from county A to county B must be applied to the individual age, gender, and race groups. The model assumes that all groups within one migration model are equally likely to move, so it allocates the total migrants from one county to another based on the relative populations of each group in the origin county. In reality, it is likely that migration patterns vary by factors such as age and race. Immigrants are also found to display different migration patterns than native-born citizens (Rogers and Henning, 1999; Perry and Schachter, 2003). However, the initial data set used here did not provide the level of specificity needed to add greater detail to the gravity model analysis.

Once migrants from one county to another are estimated by age, race, and gender, these migrations are then subtracted from the origin county and added to the destination county. This process is then repeated for every pair of counties. One major exception is that for counties that fall in the least populated quintile of counties (approximately 620 counties), with a maximum population in 2000 of 9,350, the gravity model is not applied, and no domestic migration is assumed for those counties. The counties were included in the regression modeling, and analyses indicated that the smallest counties are less likely to conform to the modeled behavior. This higher error rate, combined with the powerful attraction of large counties, caused many small counties to decline precipitously. Although excluding them from the gravity model (i.e., effectively estimating zero net domestic migration for these counties) likely leads to some distortions in these counties, collectively they account for roughly 1% of the national population. Therefore it is unlikely that their exclusion has much effect on the remainder of U.S. counties.

3.5.7. Gravity Model and U.S.-Adapted SRES Storylines

The gravity model provides many potential levers for adjusting migration patterns to account for various future scenarios. To model the high and low migration patterns needed for the SRES storylines adapted to the United States, total migrations were scaled in order to estimate greater flows of people. In the A1 and A2 storylines, where domestic migration is assumed to be higher than in the base case, all migrations were increased by 50%. In the B1 and B2 storylines, where domestic migration is assumed to be lower than in the base case, all migrations were reduced by 50%. Appendix B includes a discussion on testing other migration assumptions.

3.6. MODEL RESULTS

The demographic model was run for a base case (all parameters set to “medium”) and for the four SRES-compatible scenarios. A variety of sensitivity tests with a wider array of model inputs were also carried out. These tests and the results are discussed in greater detail in Appendix A. Model runs calculated population on an annual basis for 2005−2100, with county totals reported for five-year intervals. These outputs were then used as inputs to the spatial allocation model (Chapter 4). While the detailed model output is too large to present here, a summary discussion of the demographic model results follows.

Figure 3-3 shows the total population for the conterminous United States for the years 2005−2100, as modeled in the base case and four SRES-compatible scenarios. A2, with high fertility and high net international migration represents the highest population scenario. The base case and scenario B2 are the middle scenarios, with medium fertility and medium international migration. The difference between these two scenarios lies in the domestic migration, where the base case assumes middle-of-the-road migration flows, while B2 assumes low domestic migration flows. As a result of this distinction, the county populations in urban and suburban areas generally grow faster than in rural areas in the base case, but the experiences of individual counties vary. A1 and B1, with low fertility and high international migration are the lowest of the population scenarios. The primary difference between these scenarios occurs at the domestic migration level, with an assumption of high domestic migration under A1 and low domestic migration under B1. The effect of different migration assumptions becomes evident in the spatial model when the population is allocated into housing units across the landscape. A more extensive discussion of the regional differences in population growth is included in Section 5.1.

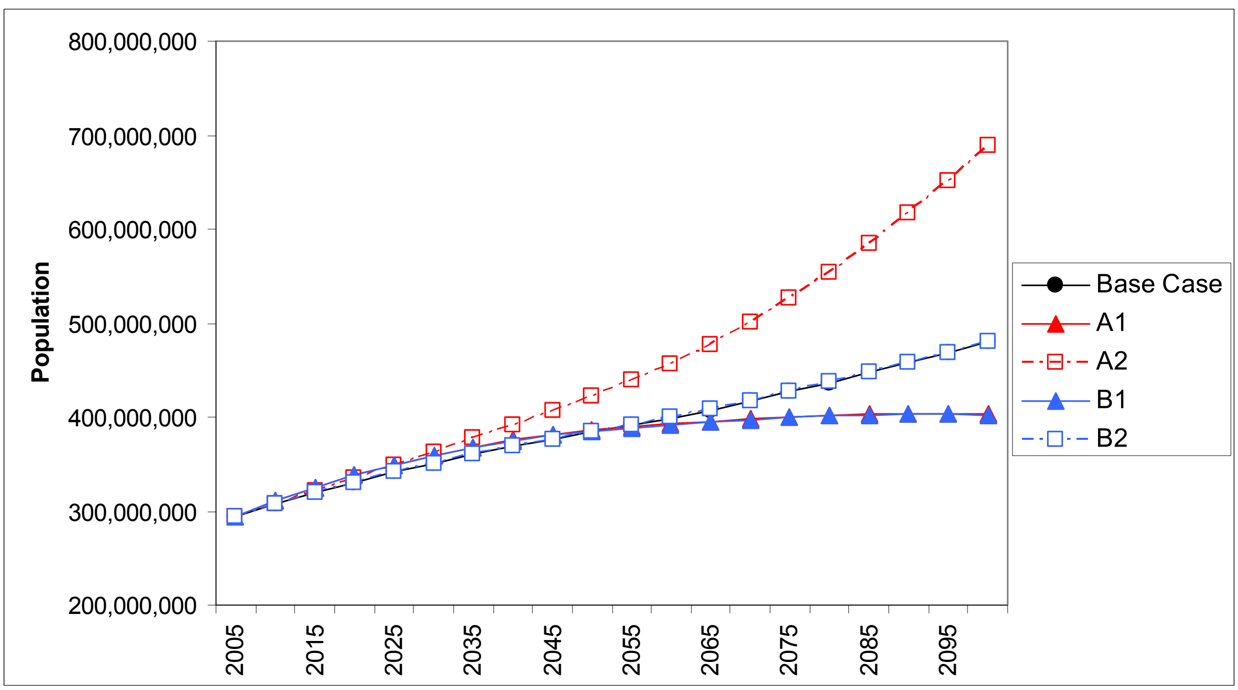

Figure 3-3. Total population under five ICLUS scenarios. Scenario B2 and the base case have the same population trajectories, as do scenarios A1 and B1.

In general, the need for additional development is directly proportional to population growth. As shown in Table 5-3 below, the growth in the extent of urban and suburban areas is greatest in the scenario with the highest projected population growth (A2). The growth in urban and suburban areas in the other scenarios is analogous to their population growth: B2 and the Base Case are in the middle, while A1 and B1 are the lowest. With ISs, the results are somewhat different (Table 5-5), indicating that the allocation and housing parameters adjusted in SERGoM such as household size and commute travel time are a greater driver of changes in IS than population alone. In this case the A1 and A2 scenarios have the highest growth in IS, while the B1 and B2 scenarios have the lowest. This suggests that decisions about which lands are utilized are as important as actual population growth in driving land-use impacts. The impacts of the population scenarios on land use and the methodologies used to model this are explored in the following chapters in greater detail.

3.7. COMPARISON OF DEMOGRAPHIC MODEL WITH EXISTING PROJECTIONS

In order to substantiate the results of the demographic model, we compared the projected county populations for five states with population projections developed by the states themselves. The five states used were California, Colorado, Florida, Minnesota, and New Jersey. All state projections were for the year 2030, except for New Jersey, which was for the year 2025. These states were selected based on data availability and regional diversity.

Differences between the state-developed projections and the ICLUS projections were anticipated. The ICLUS projections were developed using a single national model to estimate county populations, while the individual state methodologies, while not reviewed for this report, were likely developed using state-specific methods and information not available on the national scale.

The results of these comparisons are shown in Figure 3-4 through Figure 3-8. For each of the five states, the graphs depict the difference between the ICLUS projections and the state projections expressed as a percentage of the state projection on the Y-axis and the log of the county population on the X-axis. Each data point represents one county. While the results are mixed, the position of each trend line below the X-axis indicates that ICLUS estimates were lower than state estimates, on average. This is likely due to a combination of several factors:

- Higher estimated fertility, domestic in-migration, or international in-migration in the state estimates than in the ICLUS estimates;
- Differences in the state methodology that likely do not fully estimate migration patterns in other states; and

- Local knowledge about specific areas targeted for new development that was not included in the ICLUS demographic model.

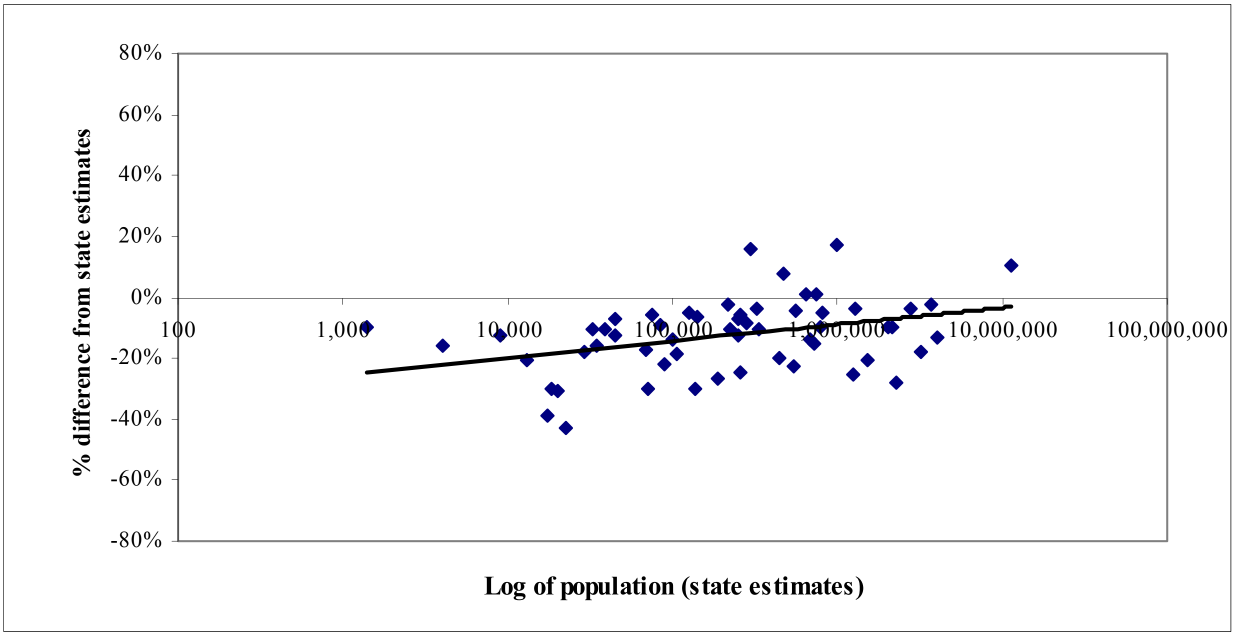

Figure 3-4. Comparison of California and ICLUS (base case) 2030 projections.

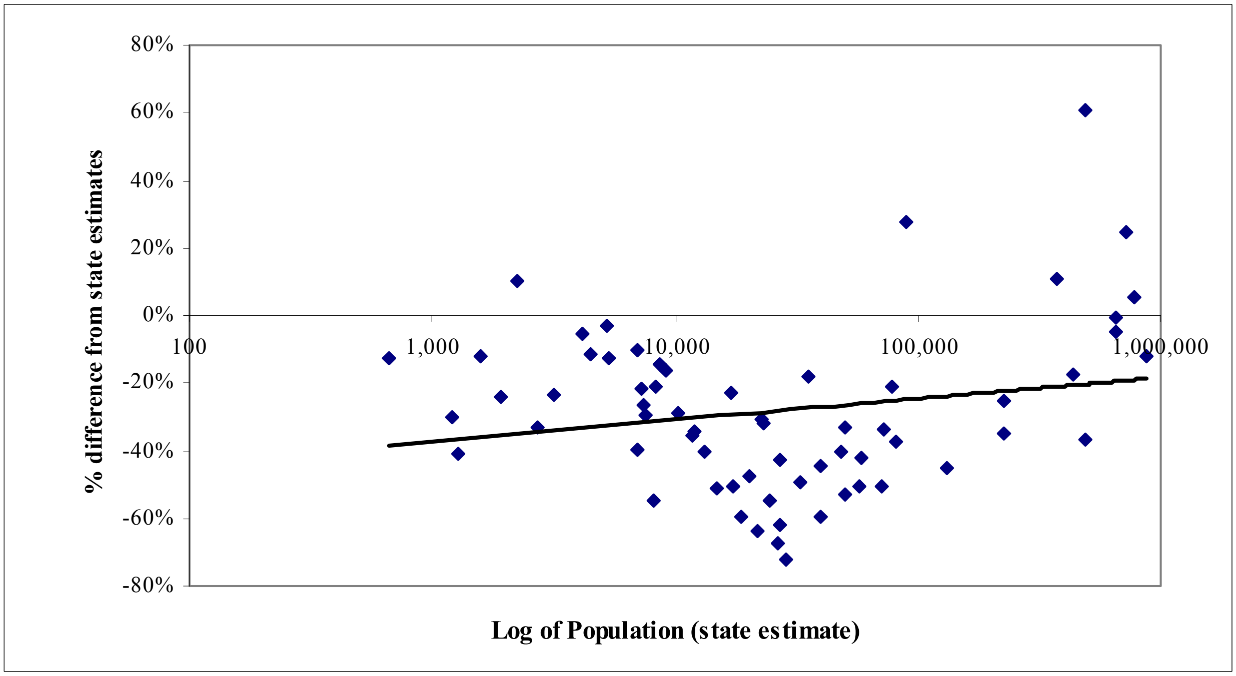

Figure 3-5. Comparison of Colorado and ICLUS (base case) 2030 projections.

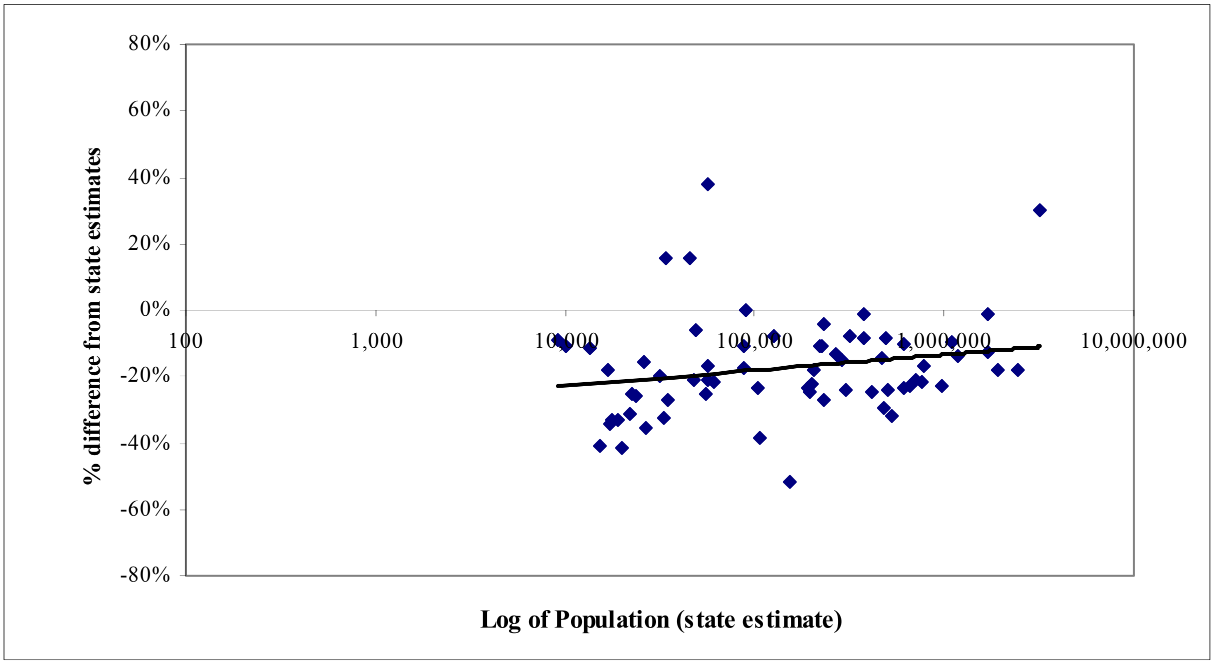

Figure 3-6. Comparison of Florida and ICLUS (base case) 2030 projections.

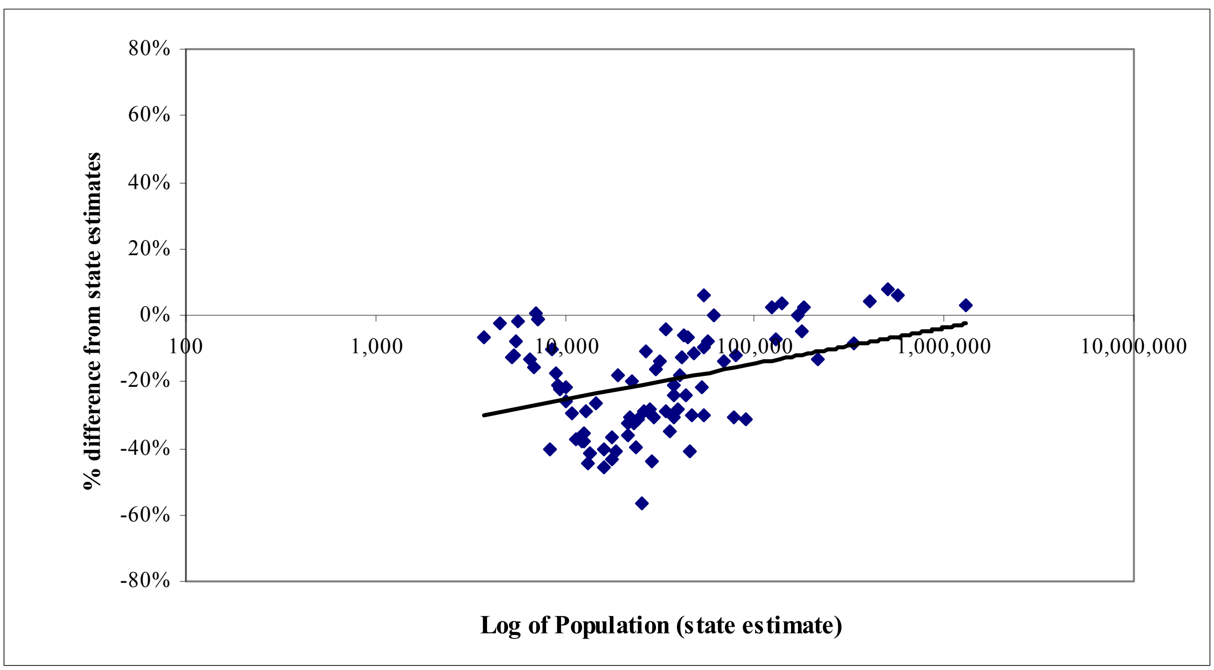

Figure 3-7. Comparison of Minnesota and ICLUS (base case) 2030 projections.

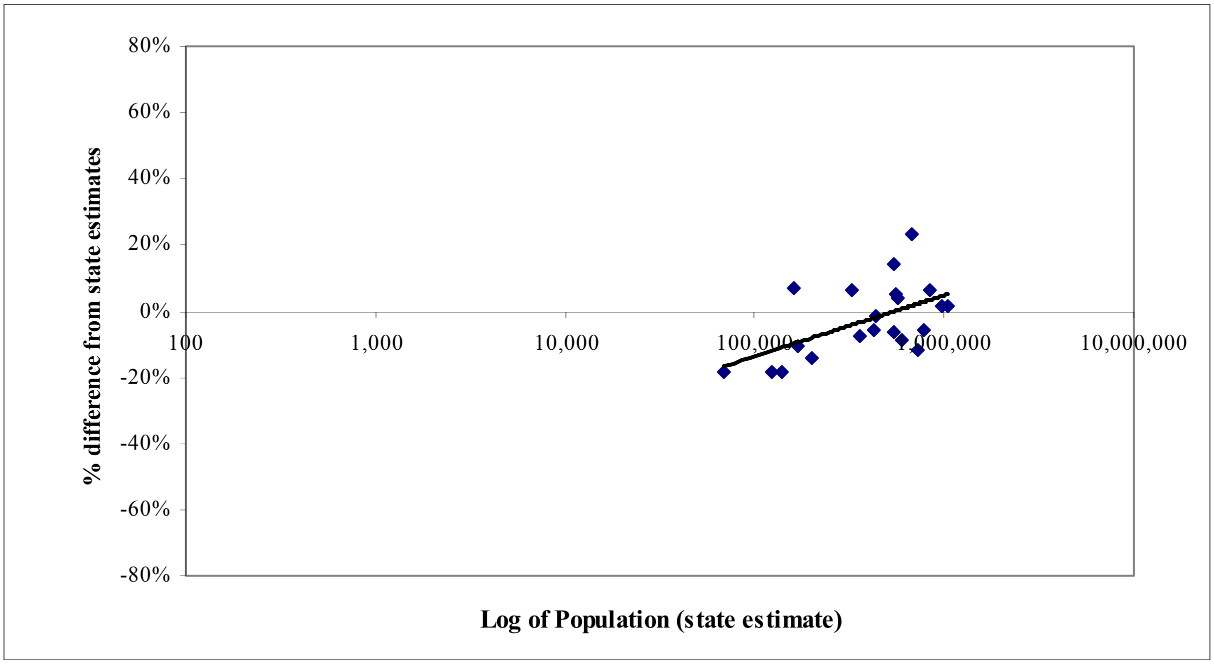

Figure 3-8. Comparison of New Jersey and ICLUS (base case) 2025 projections.

A second important observation is the slight positive slope in each trend line. This indicates that the ICLUS model is likely to underestimate the population of smaller counties more than larger counties. This result was anticipated because of the large power of population in the gravity model—in original runs, large counties grew enormous and small counties were reduced to nothing over several decades. As a result, the lowest quintile of counties were excluded from the gravity model and a limiting term was added that helped slow growth in the largest counties. The downside of this is that the ICLUS model is not able to predict population growth due to migration in small rural counties with high natural amenities. McGranahan (1999) discusses the role of natural amenities in driving rural population growth, and while the ICLUS model included natural amenities so as to model this trend, the stronger predictive power of population overwhelmed the impact of natural amenities.

In terms of overall population though, the trend line indicates that the larger the county, the more likely it was to be closer to the X-axis. This indicates that the percentage difference between state projections and ICLUS projections is smaller for large counties, where a majority of the population lives. With the exception of Florida, where the difference between the total state population projection and ICLUS was 11%, all of the other state projections discussed here differed from ICLUS by less than 7%. Given the heterogeneity of data and methods inherent in the state-specific projections as compared to the national approach applied at the county level for ICLUS, we were satisfied with the performance of the ICLUS demographic model.

4. SPATIAL ALLOCATION MODEL

We selected a spatial allocation model to distribute the population into housing units across the country. SERGoM (Theobald, 2005) was used to develop the land-use projections for this effort. This section begins with a discussion of the reasons for selecting SERGoM, and continues with a discussion of SERGoM's methodology. We then conclude this section with a discussion of how the demographic model outputs were incorporated into SERGoM and how the model was adjusted for the SRES-compatible scenarios.

4.1. RATIONALE FOR THE SELECTION OF SERGoM

SERGoM, unlike the majority of land-use change models, allocates a full continuum of HD, from urban to rural. This allows a more comprehensive examination of growth patterns, since exurban/low-density development generally has a footprint 10 times as large as urban areas and is growing at a faster rate than urban areas (Theobald, 2005). Hence, it is an important aspect of possible future growth scenarios. Other modeling approaches have recognized the importance of including land-use changes beyond the urban fringe, but these are based on modeling assumptions that require very detailed (spatially and thematically) data at the parcel level, such as zoning and socioeconomic considerations. A comparison of the differences between existing models and SERGoM is provided in Theobald (2001, 2003, 2005), but we provide a brief comparison to other existing models here for completeness. Other modeling approaches use county-level data (such as Natural Resources Conservation Service [NRCS] Natural Resource Inventory) and are useful for very coarse-level planning, but they do not provide spatially explicit forecasts of land-use change (Alig et al., 2004; Nowak and Walton, 2005). Cellular automata modeling is a particularly popular form of modeling land-use change (e.g., Batty, 1997; Theobald and Hobbs, 1998), and in a direct comparison between SLEUTH (Clarke and Gaydos, 1998) and SERGoM in the Chesapeake Bay Watershed, no practical difference between the models were found (even after making the SERGoM model coarser to fit the level of resolution for the SLEUTH model that was needed because of computational limits; Claggett et al., 2004). There are other more spatially-explicit modeling approaches, particularly from the economic geography literature, but the data to parameterize these models simply do not exist in an easily available format across the United States (e.g., Landis and Zhang, 1998; Irwin and Bockstael, 2002; Waddell, 2002; Jantz et al., 2003)—which is the primary reason why there are no other national-extent, spatially-explicit maps of forecasted HD and/or land use. Other national-extent models, mostly from Europe, are useful approaches, but have data and computational limits that preclude them from being applied to a large extent such as the United States (Verburg et al., 2002, 2006). The land-use change modeling field has been very active and new methods continue to be developed and explored, including Engelen et al.'s (2003) coupling of cellular automata model with SRES storylines and Verburg et al.'s use of CLUE-S model in Europe (2006). Parker et al. (2008) overviews a number of theoretical issues that are important for stronger inclusion of human agency into models.

An advantage of modeling all types of HD, especially low density rural development, is that links to GHG emissions by HD can be made. In addition, SERGoM forecasts housing development by establishing a statistical relationship between neighboring HD, population growth rates, and transportation infrastructure (Theobald, 2005). The model is hierarchical, being parameterized at national, state, and county scales. It is dynamic in that as new urban core areas emerge, the model re-calculates travel time from these areas. Although for our model runs we assumed a static transportation infrastructure, the travel times from urban areas do change as a function of the emergence of new urban cores. For this modeling effort the any changes in functional connectivity that could result from such emerging urbanization were not fed back into the functional connectivity calculations used to calculate domestic migration (Section 3.5.3 above). SERGoM also incorporates a detailed layer of developable/un-developable areas that incorporates public protected lands as well as private protected (e.g., through conservation easements) lands. Finally, SERGoM was designed to forecast HD growth for large, broad (regional to national) extents. Population forecasts are a principal driver of SERGoM; in the model, population growth is converted to housing units, which are spatially allocated in response to the spatial pattern of previous growth and transportation infrastructure. An important technical advantage of this model is that it produces seamless, nationwide maps at 1-ha resolution. The benefit of this approach is that there are fewer (internal to conterminous United States) discrete differences across artificial analytical boundaries imposed by "piecing" individual model runs into a nationwide map, although the allocation of new housing units is restricted to counties. Growth rates and many model parameters are specified spatially-explicitly, so different regions (even census tracts or neighborhoods) have different parameters. Although not exercised in this project, some additional model parameters could be made spatially-explicit so they too could vary regionally—these might include different HD thresholds for "urban" or "exurban" and relative changes in household size.

SERGoM has been evaluated by comparing SERGoM-based projections to estimated "real" conditions from 1990 and 2000 (starting from "real" 1980 conditions) and summarized in Theobald (2005). The evaluation used multi-resolution methods advocated by Pontius (2002) and had reasonably good performance (from 79% to 99% accuracy). It is exceedingly difficult to evaluate land-use change models, and Verburg et al. (2006) provide some useful considerations. For example, they differentiate the inductive versus deductive role of theory, introduce pixel versus agent-based representations of space, and consider the key issues of scale and level of analysis. They conclude that among the most pressing issues that need to be addressed are: validation of models and particularly multi-agent models; stronger linkage between process and pattern-based modeling approaches; and deeper understanding of the interactions and sensitivities of model results to scale of spatial data. It is important to remember that the goal of the land-use change modeling effort here is to project a range of likely changes under certain assumptions, rather than to predict the future. Compared to many other modeling approaches, SERGoM is based on very simple assumptions that incorporate only basic land-use drivers (population growth and transportation infrastructure)—but the trade-off is that an initial estimate of forecasted housing patterns has been provided. A number of other models also consider what types of cover (e.g., forest, agriculture, etc.) are being converted to residential land use.

4.2. METHODOLOGY

The spatial database generated by SERGoM provides historical, current, and future estimates of HD for the conterminous United States. That is, it represents residential land uses, which are the major types of development and intensification of land use related to urbanization in the United States (note that we also recognize and map commercial/industrial land use—however, these are static layers and we do not explicitly represent other "development" in the form of cropland or forestry developments). HD (number of housing units per acres) was computed for each 1-ha cell (100 m x 100 m raster; 2.47 acres). There are five main input spatial datasets used to estimate HD. These inputs are discussed below, while a summary of the data sources and the major assumptions used in SERGoM is presented in the Appendix in Table F-2.

1. **2000 Census.** Data were compiled from the 2000 Census on the number of housing units and population for each block and the geography or polygon boundary for each census block using the 2000 Census geography (from the SF1 dataset). Block-groups, which are a coarser-level aggregation of block polygons, and attributes of the number of housing units built by decade were used to estimate the historical number of housing units in each block. An operating assumption in estimating historical housing units is that they have not declined over time, so that the number of housing units in any past decade (back to 1940) did not exceed the number of units in any subsequent decade (up to 2000). Reservoirs, lakes, and wide rivers that were identified as "water blocks" were also removed, so that no housing units were placed in these undevelopable areas.

2. **Undevelopable lands.** Spatial data on land ownership were compiled from a variety of sources to create the most current and comprehensive dataset—called the UPPT (unprotected, private protected, public protected, and tribal/native lands). The UPPT dataset was generated by starting from the Conservation Biology Institute's PAD v4 database (CBI, 2008). We updated the PAD dataset with more current data for 21 states. The operating assumption is that housing units do not occur on publicly owned lands (e.g., national parks, forests, state wildlife areas, etc.) or on privately-owned, protected

lands. Some state lands in the western United States (the so-called "school lands" sections, but not "stewardship" lands) were kept in the developable category because they are in practice sold to generate revenue for state school systems. Also, tribal lands are often considered federal (public), but here we included tribal lands as developable (except for known tribal parks). The portions of blocks that overlapped with public (and other non-developable lands) were deleted to create a modified or refined block. All housing units associated with each block are then assumed to be located in the refined (developable) portion of the blocks. Housing units were apportioned within the refined block using a dasymetric mapping approach described below. The final product is a raster dataset that represents the developable and undevelopable lands for SERGoM, which is called dev20080611 and is available through the ICLUS tools.

3. **Road and groundwater well density.** The existence of major roads (interstates, state highways, county roads) was used to better allocate the location of housing units within a block. In a previous version of SERGoM (v1, v2), housing units were spread evenly throughout the refined blocks. Here, in v3 of SERGoM, housing units were disproportionately weighted to areas according to fine-grained land use/cover data from the 2001 National Land Cover Database (NLCD 2001). Because major road infrastructure is included in the NLCD 2001 (actually burned in as values 21, 22, and some 23), road density per se was not included. Also, in the western United States where the rural blocks are particularly large, groundwater well density was included to refine the allocation of units. Also, the analytical hierarchy process (Saaty, 1980) was used to provide an estimate of logical consistency during the development of the weights (the consistency index was 0.035, which is less that then 0.15 threshold, showing that these estimates were logically consistent). Note also that these weights are applied in a relative, not absolute context. That is, the number of units that will be distributed in a given area is specified by the census block and so units are allocated in proportion to the weights found within a given block. This is robust in the face of potential misclassification of land cover types, because all the known housing units will be allocated to a given block, regardless of the land cover (but note that the undevelopable—water and public lands—portions of the blocks are excluded). That is, the number of units are specified at the block level, and the allocation of the units are in proportion to the land cover weights. This guarantees that the correct number of units will always be allocated within each block, although the spatial distribution of the units is controlled by other factors such as roads, land cover, and groundwater well density.

4. **County population projections.** Population projections for each county, discussed in Chapter 3, are used to drive the growth forecasts. Additional housing units were computed by determining the number of new housing units needed to meet the needs of the additional population, assuming the same (in 2000) population to housing unit ratio in each tract, using 2000 U.S. Census of Housing data.

5. **Commercial/industrial land use.** We also mapped locations with land uses that would typically preclude residential development (increased HD), especially commercial, industrial, as well as transportation land uses. Using urban/built-up categories of NLCD 2001 (not open space developed), we identified locations (1-ha cells) that had >25%

urban/built-up land cover but that had also had lower than suburban levels of HD (because high-density residential areas would otherwise be included in the urban/built-up land cover categories). Although some re-development of central business districts ("gentrification") is occurring, SERGoM works from the operating assumption that these are relatively smaller portions of the landscape and typically brown-field settings.

The functional flow of SERGoM using data from these five sources is illustrated in Figure 4-1.

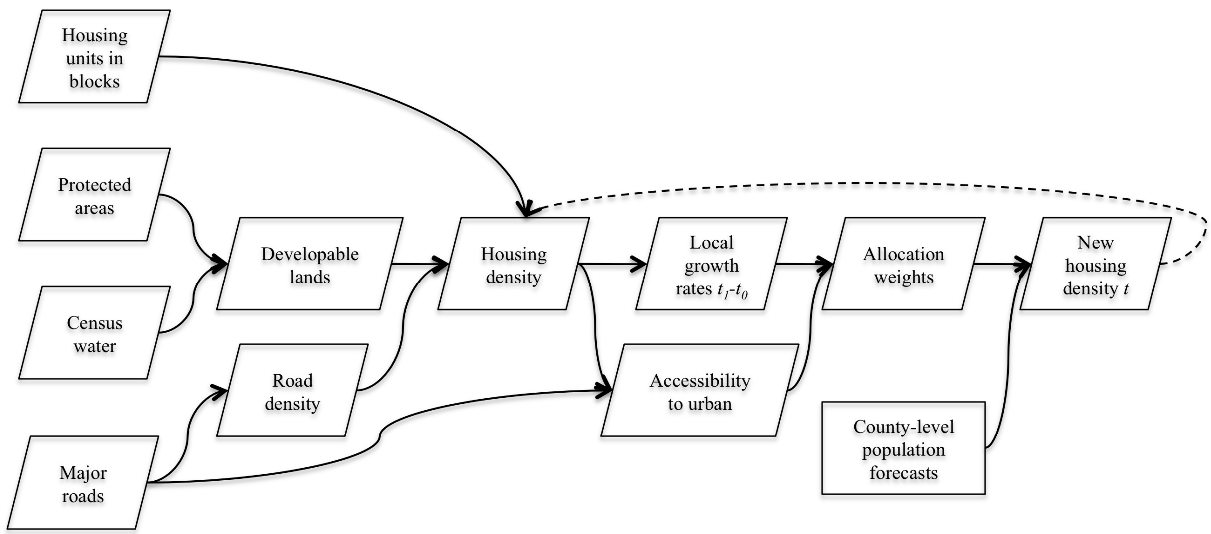

Figure 4-1. SERGoM functional flow.

HD for each decade from 2010 to 2100 was forecast using the SERGoM v3. SERGoM is a demand/allocation/supply model, where the number of new housing units needed for the next decade is computed to meet the demands of the projected population, computed here for each county (but could be other analytical unit boundaries). The average growth rate for each state-HD class is computed from the previous to current time step (e.g., 1990 to 2000). These average growth rates are computed using a moving neighborhood (radius = 1.6 km, 500-m cell) for 12 development classes. These classes are formed by overlaying three HD classes—urban/suburban, exurban, and rural—with four accessibility classes measured as travel time (minutes one way) from the nearest urban/suburban core along (existing) major roads: 0−10, 10−30, 30−60, and >60 minutes. The resulting combination creates a "surface" of raster values that reflect historical patterns of growth—called allocation weights—and are used to allocate the new housing units for a given time step.

Based on the Census definition of urban areas, we defined urban housing densities as less than 0.1 ha per unit and suburban as 0.1−0.68 ha per unit. We defined exurban density as

0.68−16.18 ha per unit (to 40 acres) to capture residential land use beyond the urban/suburban fringe that is composed of parcels or lots that are generally too small to be considered productive agricultural land use (though some high-value crops such as orchards are a notable exception). Rural is defined as greater than 16.18 ha per unit where the majority of housing units support agricultural production.

The strength of SERGoM is that it provides a comprehensive, consistent, and nationwide estimate of HD. It uses the most fine-grained data set currently available, has performed reasonably well in assessments (79 to 99% accuracy rates), and compares favorably to parcel-level and aerial photography data during ad hoc analyses in a variety of locations in the United States (Theobald, 2005). It assumes that growth rates and patterns are likely to be similar to recent times (1990s to 2000). The SERGoM outputs provide much more spatial detail as compared to the USDA Census of Agriculture,[5] which are county-based and only provide data on land in farms. One other common datatset used to estimate the extent and trend of urbanization in the United States is the NRCS's Natural Resources Inventory (NRI).[6] It is based on relatively fine-grained aerial photo analysis, but because they are sampled data, they are aggregated up to coarse analysis units (either county, watershed 8-digit Hydrologic Unit Codes [HUCs], or Major Land Resource Area). NRI also categorizes urban areas into only two classes―either as "small" or "large" development―resolving housing densities at urban and roughly 1 per 10 acre densities.

4.3. INCORPORATING U.S.-ADAPTED SRES INTO SERGoM

In addition to changes in population that resulted from the various demographic assumptions associated with each SRES-compatible scenario developed for the ICLUS project, the spatial location of growth was modified using SERGoM in two ways, through household size and travel time (Table 4-1). With SERGoM, household size is expected to reflect demographic changes due to changes in fertility and socioeconomic changes that affect household formation. Travel time from urban "central city" locations is used to help express how the evolution of the urban form might by affected by changing priorities and increases in the cost of transportation. Table 4-1 also shows how travel times are translated into this urban form. Housing units are allocated to a surface and the allocation is weighted by the accessibility to the transportation network, thereby influencing urban form over time to create a more "compact" form of development when allocations near urban centers are weighted more favorably.

---

[5] http://www.agcensus.usda.gov/.
[6] http://www.ncgc.nrcs.usda.gov/products/nri/.

Table 4-1. Summary of adjustments to SERGoM v3 for SRES storylines

Storyline	Household size	Travel time (minutes) <5; 10; 20; 30; 45; >45 *Weighting (in percent)*	Form
A1	Smaller (−15%)	75; 75; 85; 90; 95; 100	NC
B1	Smaller (−15%)	90; 95; 85; 90; 95; 100	Slight compact
A2	Larger (+15%)	75; 75; 85; 90; 95; 100	NC
B2	NC	90; 95; 85; 90; 95; 100	Slight compact
Baseline	NC	75; 75; 85; 90; 95; 100	NC

NC = No change from the U.S. Census Bureau's estimates.

The first SERGoM modification changed assumptions about households, particularly household size (roughly family size), defined as the number of people living in a single housing unit. Population projections from the U.S. Census assume that the ratio of population per unit, computed at the tract level from the 2000 U.S. Census data, is static. We modified this ratio to reflect assumptions in the SRES storylines to adjust for assumed changes in demographic characteristics. For example, SRES A1 and B1 assume smaller household sizes (reduction by 15%), whereas scenarios B2 and baseline are not changed and A2 assumes a 15% increase in household size (Jiang and O'Neill, 2007). The changes in household size correspond to changes in fertility rates that are assumed under the different storylines. Under A1 and B1, where fertility is lowest, smaller average household sizes are also expected. Conversely, A2 has the highest fertility rates, so an increase in household sizes is expected. In B2, which uses the medium fertility rates, household sizes are not changed.

The second modification involves changing travel times by adjusting weighting values as a function of distance away (travel time) from urban cores. Urban area (<5 minutes) weights can be lowered by a given percentage to reflect a carrying capacity or saturation of an area, specified by zoning perhaps; or raised to reflect increased desire for urban living (lofts, gentrification, etc.). Exurban area weights (~30−60 minutes) can be lowered to reflect assumptions of lower rates of development due to increased fuel prices or can be used as a surrogate for lower land availability because of increased conservation purchases (or easements). It can also be raised for exurban areas to reflect increased "urban flight" of baby-boomer retirees and rural amenities. This weighting surface is re-computed at each time step. We modified the weights of travel times for the B1 and B2 storylines to model a "compact" growth scenario (see Table 4-1). Given the environmental orientation of the B1 and B2 storylines, we assumed that growth patterns in these scenarios would place a greater emphasis on promoting denser growth patterns closer to existing urban centers.

We parameterized SERGoM to reflect the SRES storylines in the following ways. First, the A1 and B1 scenarios were modeled to reflect a 15% decline in average household size. A2 was modeled to show a 15% increase in average household size. B2 was modeled with no change in household size. This modification changes the resulting number of housing units because housing units in SERGoM are a product of the population growth and household size. That is, number of housing units depends on the number of new people needed to be housed, along with the assumed number of people per housing unit. Second, to model the "compact" growth scenarios for the B1 and B2 scenarios, SERGoM was run with modifications to the spatial allocation of new housing units as a function of travel time from urban cores (Table 4-1). That is, the allocation weights were modified by multiplying the modification factor (e.g., 75% for <5 minutes travel time for A1) times the overall weight developed from past growth patterns (as a function of HD and travel time away from urban areas).

4.4. INTEGRATION OF DEMOGRAPHIC, SERGoM, AND IMPERVIOUS MODELS

The demographic inputs from each of the U.S.-adapted SRES storylines were fed into SERGoM, along with the SRES-specific adjustments in SERGoM for a given SRES storyline. The outputs of SERGoM were then mapped for each decade from 2000 to 2100 (Appendix A). We also used the IS model to convert the SRES HD estimates to the total percent IS cover (Section 5.3). Chapter 5 discusses some preliminary analyses and potential future developments of the ICLUS modeling framework.

5. IMPACTS AND INDICATORS ANALYSIS

5.1. RATES OF GROWTH IN DIFFERENT REGIONS

The growth rates of the different regions of the United States under the various SRES storylines provides some interesting insight into the potential relative growth patterns in the coming decades. For this analysis of regional growth patterns, we used the U.S. Census regions (listed in Table 5-1). A separate analysis using EPA regions is presented in 0. The populations of each of the Census regions and scenarios for 2005, 2030, 2060, and 2090, as well as the growth rates for each intervening period are presented in Table 5-2. These data are then displayed graphically to compare the different regions and scenarios (Figure 5-1 to Figure 5-13).

Table 5-1. U.S. Census regions

Census region	States
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
Midwest	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia
West*	Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

*Alaska and Hawaii were excluded from the West region in this analysis.

Figure 5-1 through Figure 5-5 compare the population in the four Census regions under each of the four SRES storylines and the base case. In all five sets of population projections, the South region remains the most populous region, while growing faster than most other regions in most scenarios. The West region, which begins with approximately the same population as the Northeast and Midwest regions, outpaces those two regions in all scenarios except for B1. Across the board, all regions experience growth in all scenarios, with the exception of the Midwest region in scenarios A1 and B1.

Table 5-2. Projected regional populations and growth rates

Census region	Population				Growth Rate (%)		
	2005	2030	2060	2090	2005–2030	2030–2060	2060–2090
Base case							
Northeast	54,679,292	63,384,211	70,279,618	79,600,361	16%	11%	13%
Midwest	65,936,398	71,250,888	73,287,983	76,745,425	8%	3%	5%
South	108,981,468	134,649,231	158,547,450	187,417,301	24%	18%	18%
West	64,973,375	81,705,117	96,926,417	114,073,784	26%	19%	18%
A1 Storyline							
Northeast	54,679,292	66,910,792	73,137,911	76,159,296	22%	9%	4%
Midwest	65,936,398	69,265,132	65,378,806	59,355,485	5%	−6%	−9%
South	108,981,468	140,717,741	162,287,975	174,443,647	29%	15%	7%
West	64,973,375	82,571,676	91,877,085	94,284,391	27%	11%	3%
A2 Storyline							
Northeast	54,679,292	64,718,449	78,041,141	101,339,355	18%	21%	30%
Midwest	65,936,398	71,810,622	80,335,490	98,476,101	9%	12%	23%
South	108,981,468	141,466,245	186,892,292	262,941,111	30%	32%	41%
West	64,973,375	85,129,358	111,597,232	153,721,857	31%	31%	38%
B1 Storyline							
Northeast	54,679,292	68,481,437	76,906,738	82,327,874	25%	12%	7%
Midwest	65,936,398	72,973,243	72,166,976	67,641,537	11%	−1%	−6%
South	108,981,468	136,176,035	152,179,001	159,582,292	25%	12%	5%
West	64,973,375	81,468,120	90,471,745	93,288,362	25%	11%	3%
B2 Storyline							
Northeast	54,679,292	64,097,466	72,006,459	82,638,846	17%	12%	15%
Midwest	65,936,398	73,134,039	77,070,686	82,236,752	11%	5%	7%
South	108,981,468	132,513,626	153,888,715	179,498,472	22%	16%	17%
West	64,973,375	81,182,232	96,254,307	113,554,597	25%	19%	18%

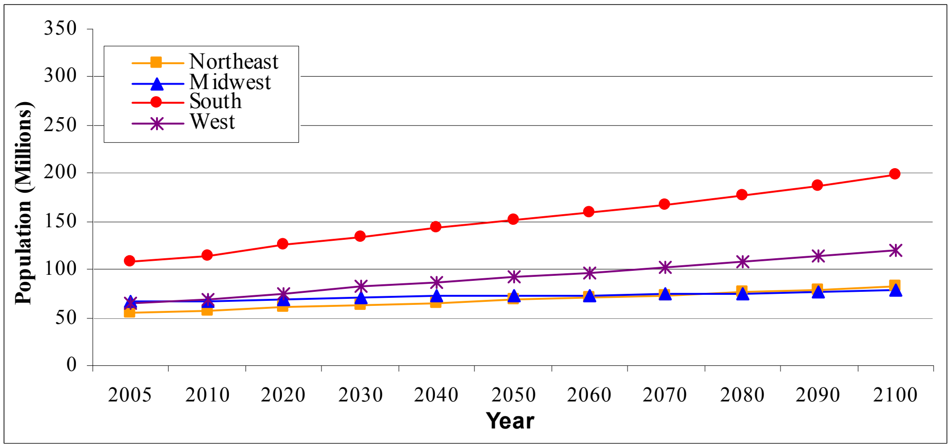

Figure 5-1. Base case population by Census region.

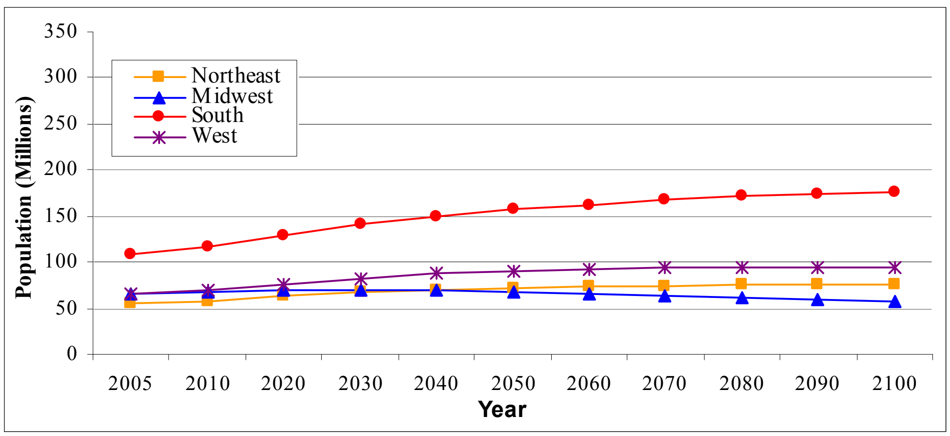

Figure 5-2. A1 storyline population by Census region.

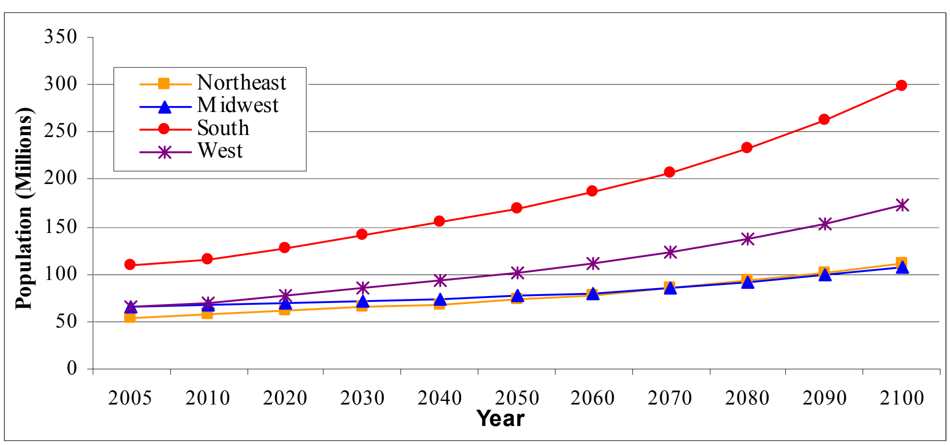

Figure 5-3. A2 storyline population by Census region.

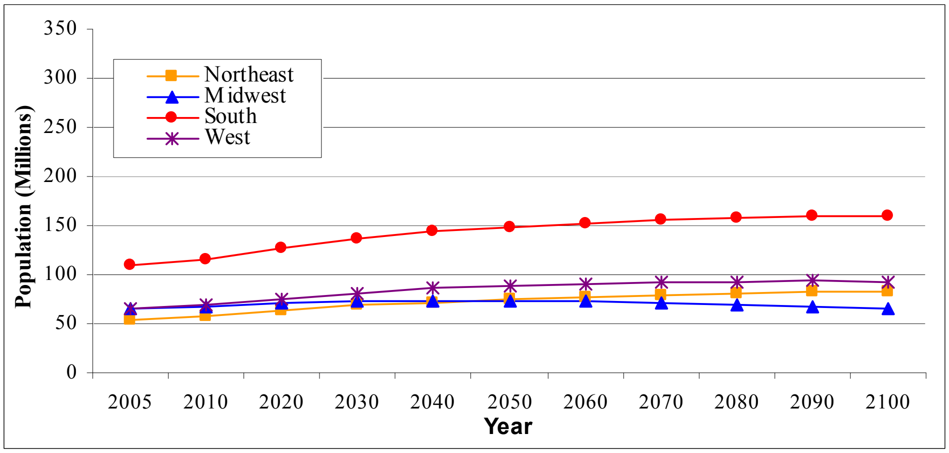

Figure 5-4. B1 storyline population by Census region.

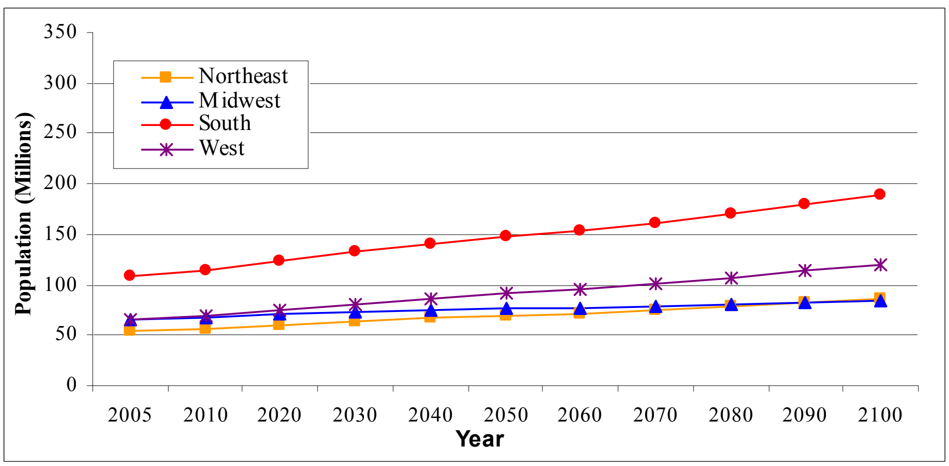

Figure 5-5. B2 storyline population by Census region.

Figure 5-6 through Figure 5-10 compare the average annual growth rates during each modeled decade under the different scenarios (e.g., a growth rate of 1.01 indicates growth of 1%). Under the base case and B2 storyline, population growth rates are highest during the first time period, then level off for the remaining two periods. The A1 and B1 storylines, by comparison, generally decline in growth rate throughout the 21st century. The A2 storyline, which has the highest overall population growth, is the only one that exhibits steady increasing rates of population growth over the next century.

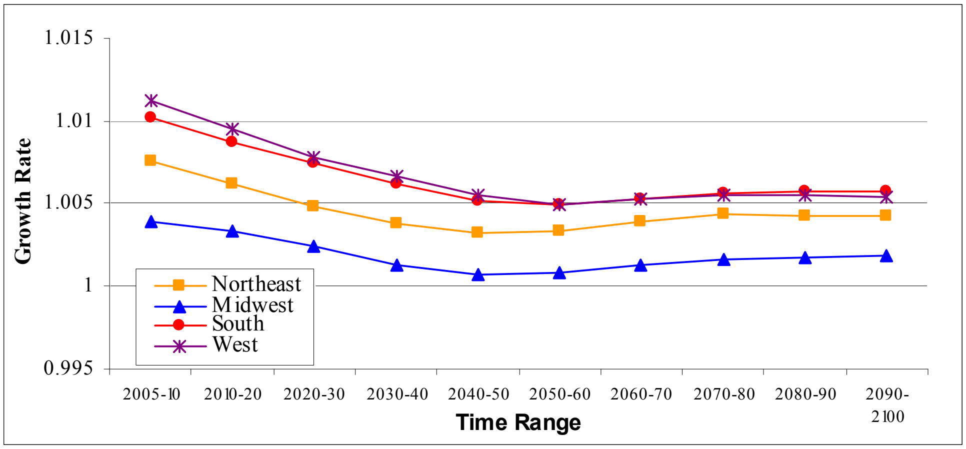

Figure 5-6. Base case annual population growth rates by region.

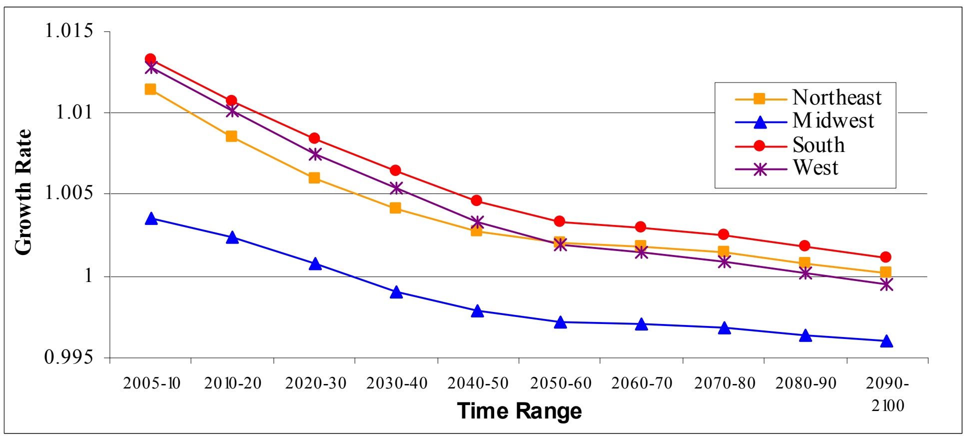

Figure 5-7. A1 storyline annual population growth rates by Census region.

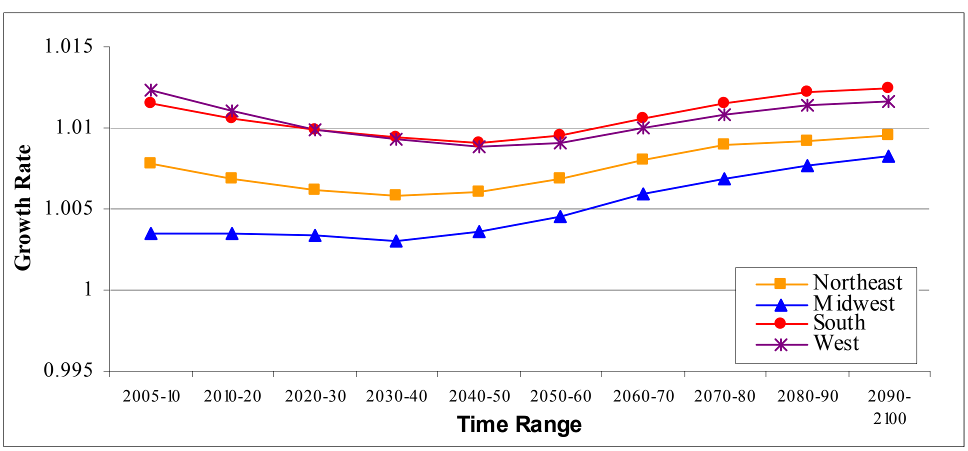

Figure 5-8. A2 storyline annual population growth rates by Census region.

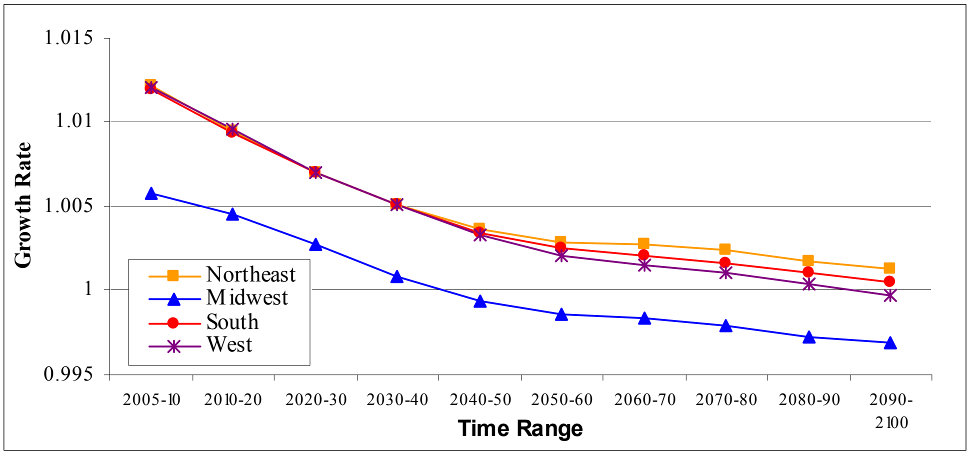

Figure 5-9. B1 storyline annual population growth rates by Census region.

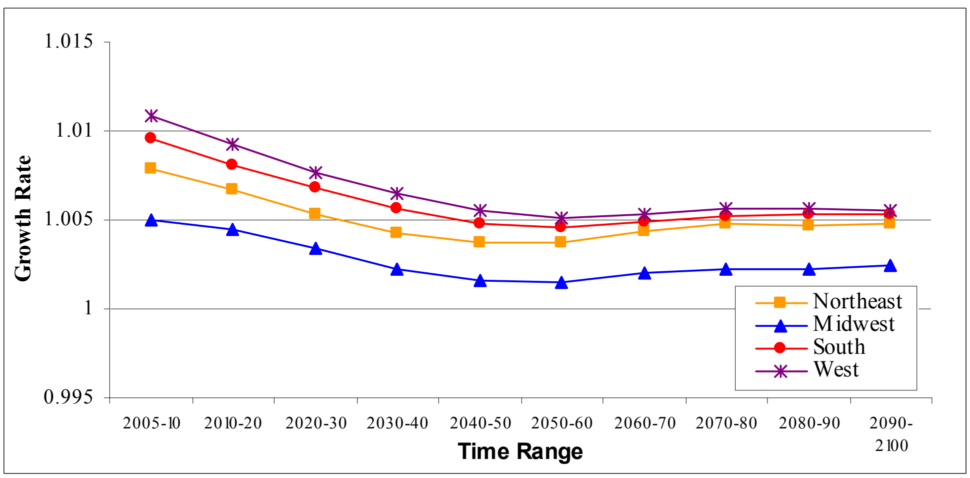

Figure 5-10. B2 storyline annual population growth rates by Census region.

Figure 5-11 through Figure 5-14 provide comparisons of the storylines for each of the four Census regions. A2 produces the highest population in each region, confirming our expectations based on our interpretation of the SRES storylines. A1 produces the lowest population in the Northeast and Midwest regions, while B1 produces the lowest population in the South and West regions. This is because domestic migration, which is expected to continue to trend toward the South and West (and away from the Midwest and Northeast), is set to “high” under A1 and “low” under B1. Again, all regions grow under all scenarios, with the exception of the Midwest region under A1 and B1.

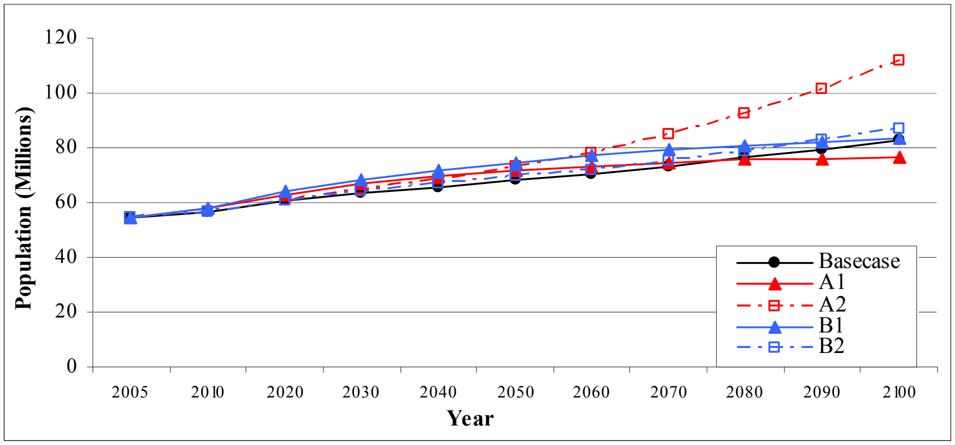

Figure 5-11. Northeast region population by storyline.

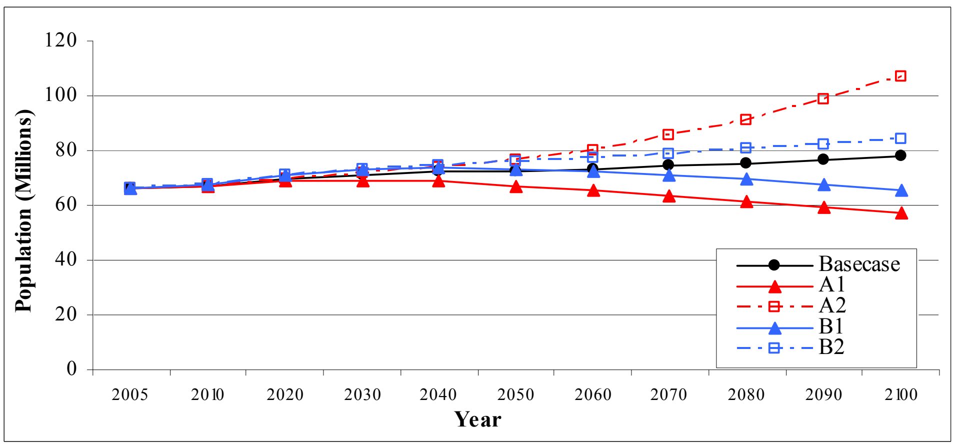

Figure 5-12. Midwest region population by storyline.

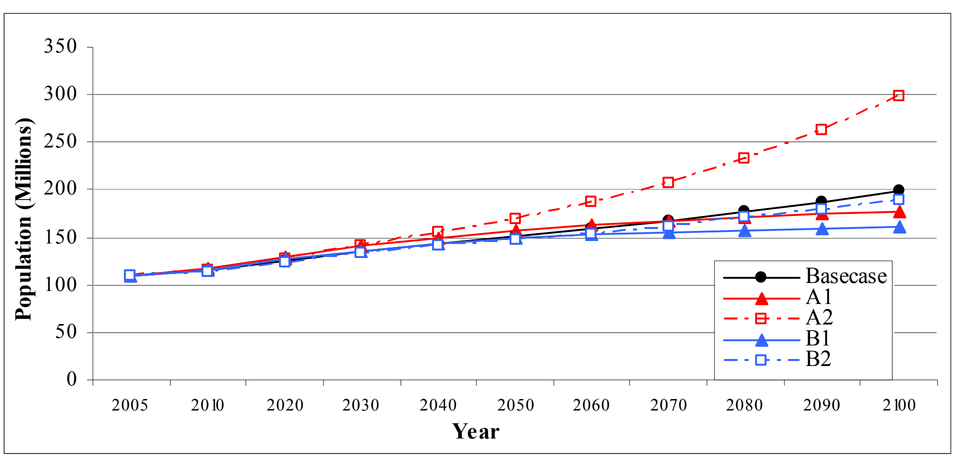

Figure 5-13. South region population by storyline.

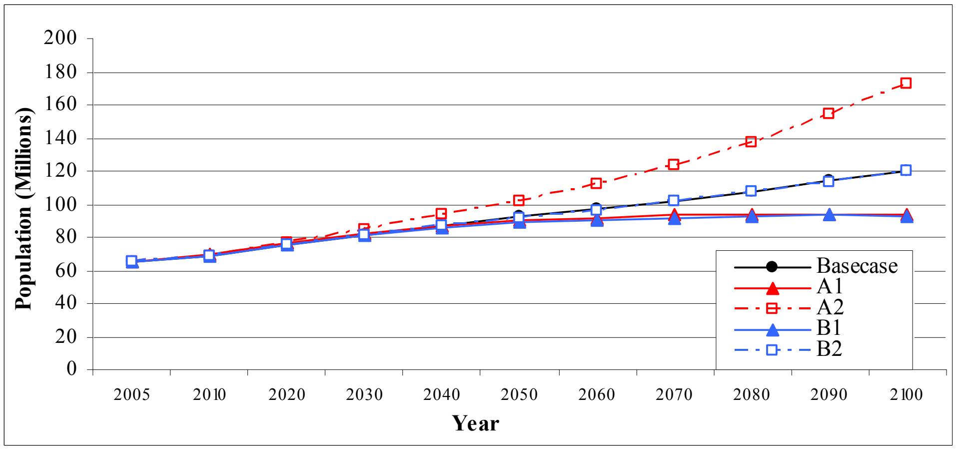

Figure 5-14. West region population by storyline.

5.2. HOUSING DENSITY TRENDS

The projected growth in population and HD is anticipated to lead to corresponding impacts on environmental attributes such as water quality and air quality. Since the challenges are expected to be greater in urban and suburban areas, we used the model outputs to estimate the growth in these higher density areas. Under all modeled scenarios, urban and suburban areas are expected to increase between 60 and 164%. For this analysis, urban and suburban areas are defined as those areas with more than ~0.6 units per acre (or less than 1.68 acres per unit). This land class is expected to increase the most in the A2 scenario, adding over 193,000 km^2 over the next century, or 164% more than 2000 levels (about 118,000 km^2) for a total of over 300,000 km^2 of urban/suburban area in 2100 (Table 5-3). Other increases are expected to be significant, but more moderate than the A2 scenario, with B1 having the smallest increase (60%) (Table 5-3; Figure 5-15). The non-intuitive result that B2 has a higher amount of urban/suburban area as compared to the base case may be the result of the net trade-off of negatively weighting growth in regions beyond suburban areas (i.e., exurban and rural areas); this growth results in a greater extent of the land surface containing urban/suburban densities, compared with the base case where those housing units are more frequently in exurban or rural categories. Note that we do not include in our estimates of urban/suburban housing densities the over 32,300 km^2 of estimated commercial/industrial land cover for 2000.

We also examined which broad land cover types were likely to be most affected by the projected development patterns (Table 5-4). To do this, we quantified the spatial overlap of the urban, suburban, and exurban housing densities (>1 unit per 40 acres) on the existing major land cover type as characterized by NLCD 2001 Anderson Level I coding. We did this by aggregating the 30-m resolution data to 90-m cells using the majority class filter. The largest impacts for all scenarios, both in terms of percentage and total area, are estimated to be on agricultural (cropland) cover where approximately 33% of area is converted into housing in these scenarios. Although wetlands cover less land area, our scenarios convert 30−36% of wetlands to housing. Shrublands are similar in total area converted, although the scenarios present more of a range in the amount converted (25−34%).

Table 5-3. Projected urban and suburban area increases in modeled scenarios, 2000−2100 (km^2)

Scenario	2000	2050	2100	% Increase, 2000−2100
Base case	118,468	177,066	263,315	99%
A1	118,468	194,312	223,272	88%
A2	118,468	192,878	312,426	164%
B1	118,468	174,063	189,649	60%
B2	118,468	181,857	244,080	106%

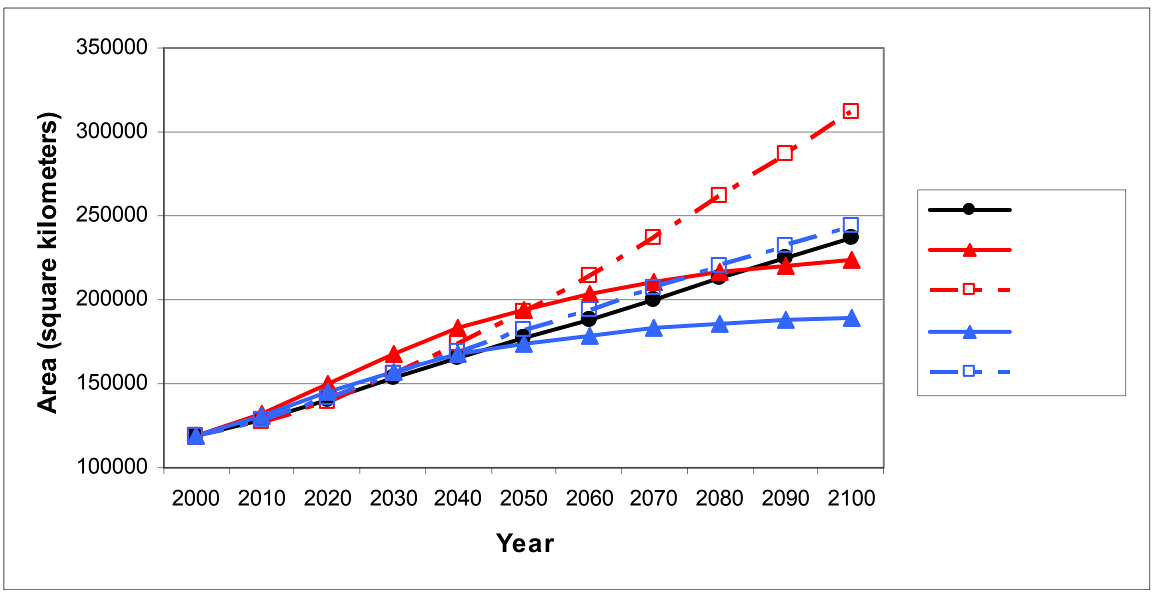

Figure 5-15. Urban/suburban housing land-use trends for ICLUS SRES storylines.

Table 5-4. Projected area (km^2) effects of urban, suburban, and exurban housing densities on NLCD 2001 land cover types in modeled scenarios for 2050

Scenario	Forest	Shrubland	Grassland	Agriculture (cropland)	Wetland
Current	505,280	56,258	83,257	188,316	51,493
Total area change between current and 2050					
Base Case	−54,656	−20,491	−18,959	−58,054	−18,035
A1	−57,267	−21,276	−19,755	−59,188	−18,787
A2	−55,867	−20,956	−19,549	−58,887	−18,852
B1	−55,389	−19,161	−17,920	−57,613	−16,958
B2	−50,560	−18,156	−17,132	−51,884	−16,466
Percent change between current and 2050					
Base Case	10.8%	36.4%	22.8%	30.8%	35.0%
A1	11.3%	37.8%	23.7%	31.4%	36.5%
A2	11.1%	37.2%	23.5%	31.3%	36.6%
B1	11.0%	34.1%	21.5%	30.6%	32.9%
B2	10.0%	32.3%	20.6%	27.6%	32.0%

5.3. IMPERVIOUS SURFACES

ISs such as pavement and roofs degrade water quality, and by collecting pollutants, increasing run-off during storm events, and absorbing heat they also contribute to the heat island effect (Frazer, 2005). Although there are a variety of ways that IS plays a role in emissions or climate, in this document we pursued only the use of IS as a general indicator―not specifically tied to possible changes in carbon cycling, emissions, or heat island effects.

To develop national estimates of likely future IS, we developed a statistical relationship between 2000 HD and the NLCD 2001 Percent Urban Imperviousness (IS_{PUI}) data set.[7] We aggregated the 30 m IS_{PUI} data set to 900-m^2 cells and then resampled using bilinear interpolation to 1-km^2 resolution. We aggregated the SERGoM 1 ha HD up to 1 km^2 as well. Based on a spatially-balanced random sample of 200,000 points from across the conterminous United States, we developed a relationship using the *cv.tree* function in S-Plus (Insightful Corporation, Seattle, WA) between IS_{PUI} and HD_{2000}. To generate a map of IS based on the HD

[7] http://www.mrlc.gov/multizone_download.php?zone=4; accessed 12 February 2007.

(IS_{HD}), we converted the tree into a set of if-then-else conditional statements in ArcGIS. A brief comparison of our modeled IS to estimates of "real-world" IS (computed from interpretation of high-resolution aerial photography) provide evidence that our modeling effort worked reasonably well—resulting in R^2 between 0.69 and 0.96. For example, we obtained an $R^2 = 0.69$ for 80 "real-world" points used by Elvidge et al. (2004). Also, we obtained an $R^2 = 0.69$ and $R^2 = 0.96$ using "real-world" data points generated in Frederick County, Maryland and Atlanta, Georgia (Exum et al., 2005). Because our estimates of IS are produced on a per 1 km^2 pixel basis, we were able to roughly compare our estimates to Exum et al.'s (2005) by averaging pixels that fell within their 10−12-digit HUCs. It would be useful to compare our results to other aerial-photography-based estimates, but to our knowledge the spatial datasets of these estimates of IS are not readily obtainable. The detailed methods and results are described in Theobald et al. (2009) and in Appendix C.

Because we estimated IS as a function of HD, the modifications to urban form and consequently HD computations for scenarios B1 and B2 carry through. That is, the estimated IS is a function of the HD specifically, and not just simply the number of housing units. This captures the non-linear (decrease) in IS percentage on a per housing unit basis.

5.3.1. Impervious Surface Calculations

The total percent IS (computed at 1 km^2) for the United States in 2000 was 80,094 km^2. We developed a regression model (described in Appendix C) that relates HD estimates in 2000 to estimates from the IS_{PUI} from the NLCD 2001 dataset. Based on that statistical relationship, we were able to forecast likely changes to IS for different future patterns of land use that reflect our SRES growth scenarios (Table 5-5; Figure 5-16). To aid interpretation and mapping of our results, we classified our continuous, quantitative estimates of IS into five legend classes: 0−0.9% (unstressed), 1−4.9% (lightly stressed), 5−9.9% (stressed), 10−24.9% (impacted), and >25% (degraded), following Slonecker and Tilley (2004) and Elvidge et al. (2007) and Theobald et al. (2009). Note that the use of legend labels such as "stressed" are for cartographic purposes only and should be used as a general indicator of the level of IS.

5.3.2. Percent of Watersheds Over 5% Impervious

Our estimates here do not include IS for known commercial/industrial lands (in 2000)—that added an additional 13,430 km^2 of IS area. All HD classes were included when estimating the IS. In 2000, urban/suburban areas (<1.68 acres per unit) comprised 49.4% of the total IS (accounting for different percent IS), exurban areas (1.68−40 acres per unit) comprise 34.2%, and rural comprised 16.3%. We estimate that in 2000 there were 136 8-digit HUCs that were stressed or higher (at least 5% IS), and this will likely increase to between 208 to 231 in 2050 and to between 225 and 326 in 2100.

In general, there are fairly large differences between the amount of IS that likely will result from different growth scenarios—from a 5.5% increase (A1) from base case to a 1.2% decline (B1) from base case in 2050. Figure 5-17 to Figure 5-26 show the IS patterns for each U.S.-adapted SRES storyline and the relative differences between these scenarios and the base case.

Table 5-5. Impervious surface estimates for SRES storylines

SRES	Year 2050: Total impervious surface (km^2)	Year 2050: Number of stressed 8-digit HUCs (of 2101)	Year 2100: Total impervious surface (km^2)	Year 2100: Number of stressed 8-digit HUCs (of 2101)
Base case	106,174	221	124,835	257
A1	112,052	231	122,533	256
A2	110,574	229	149,754	326
B1	104,946	208	110,949	225
B2	105,252	217	122,933	248

"Stressed" level is defined as at least 5% impervious surface.

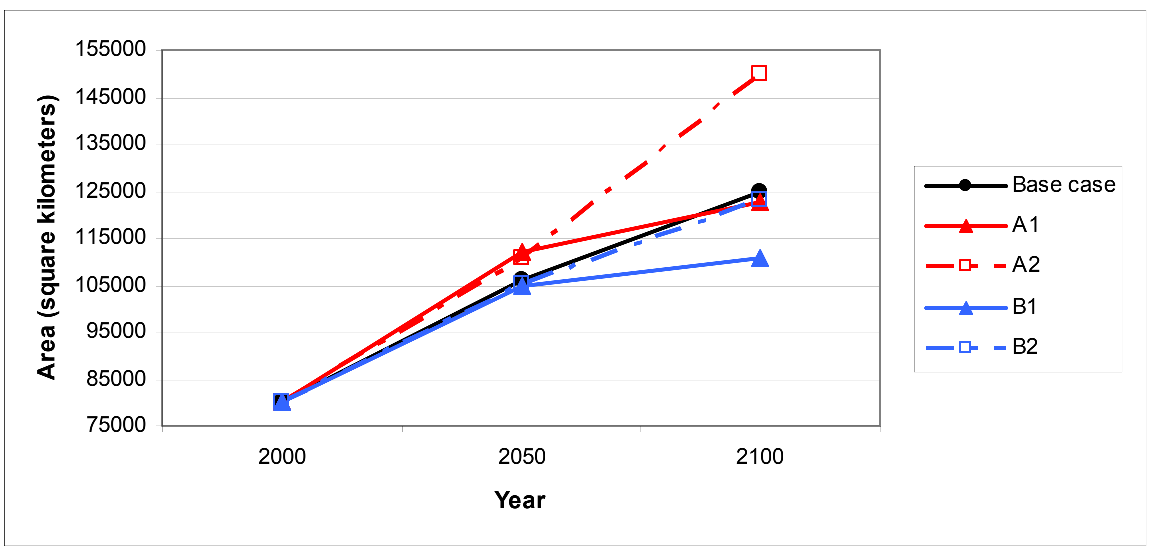

Figure 5-16. Impervious surface area estimates, 2000−2100.

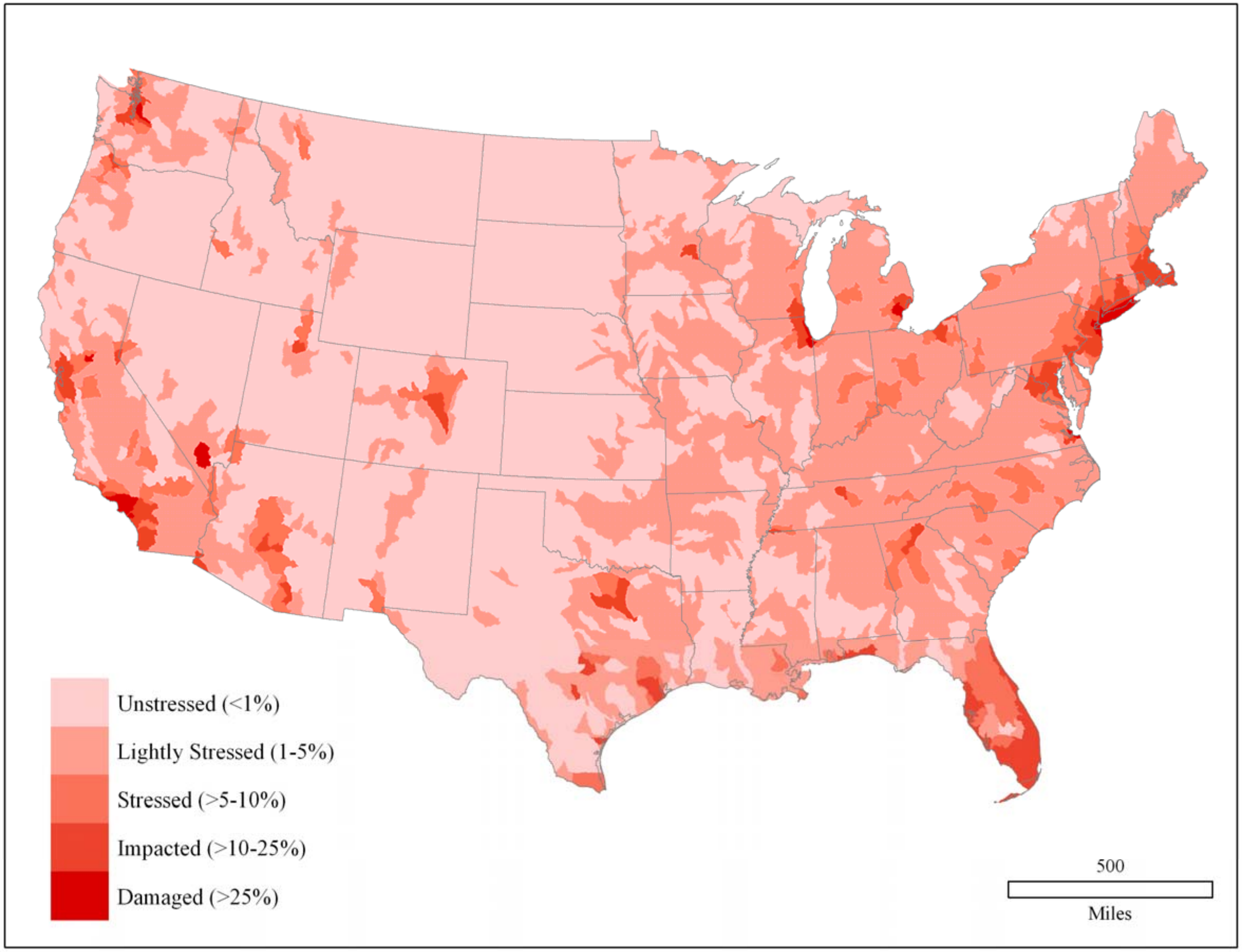

Figure 5-17. 2050 estimated percent impervious surface, base case.

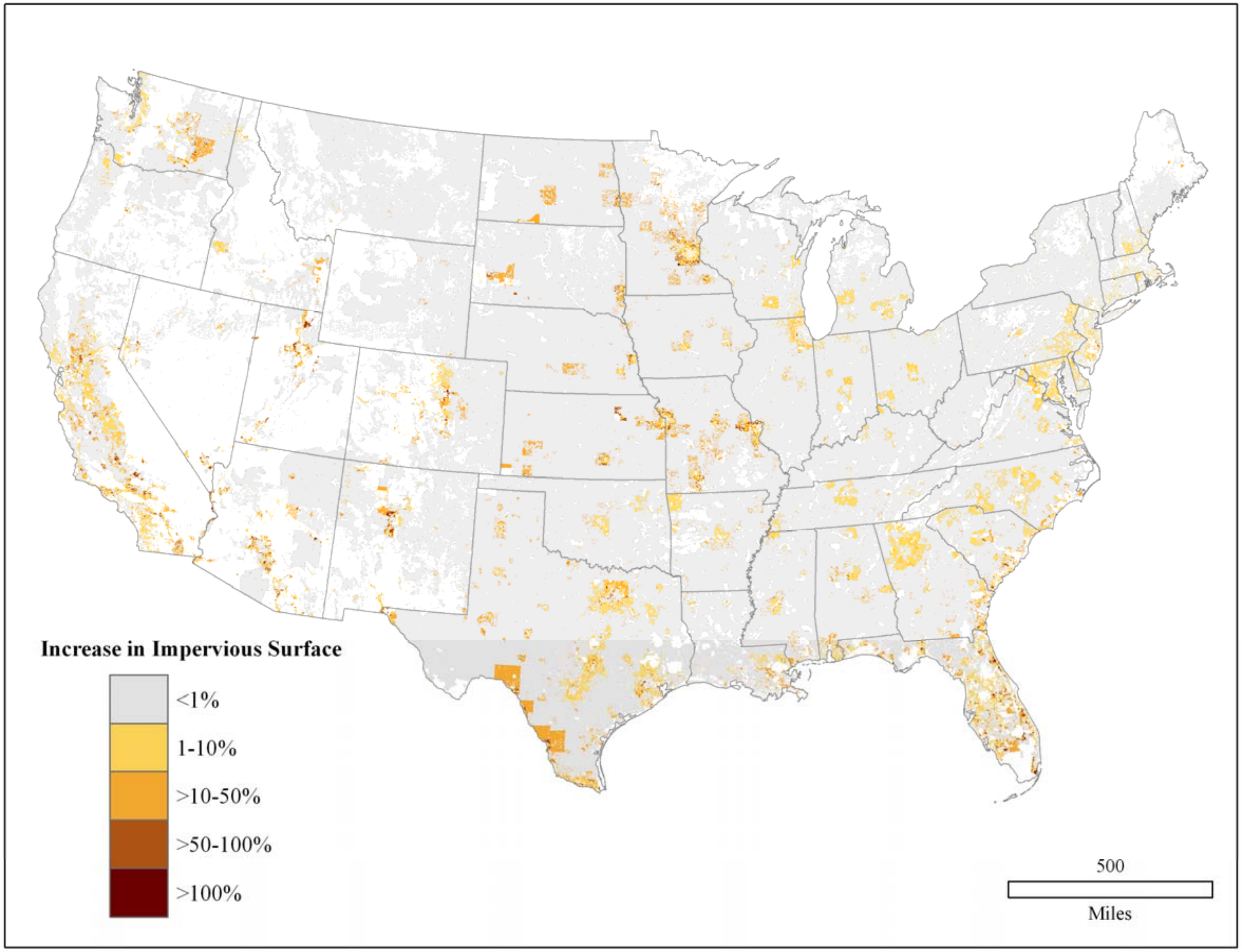

Figure 5-18. 2000−2050 relative change in impervious surface, base case.

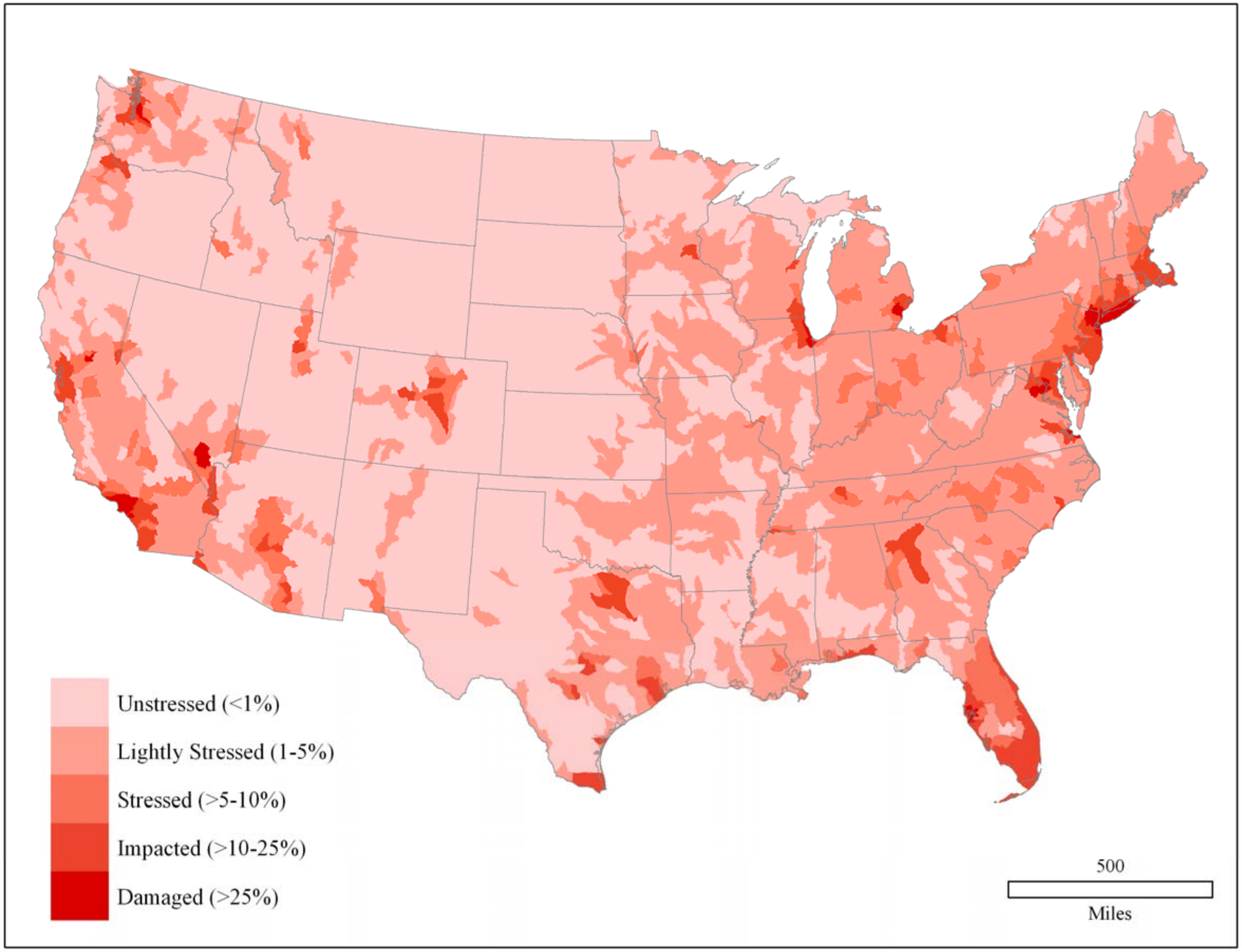

Figure 5-19. 2050 impervious surface, A1 storyline.

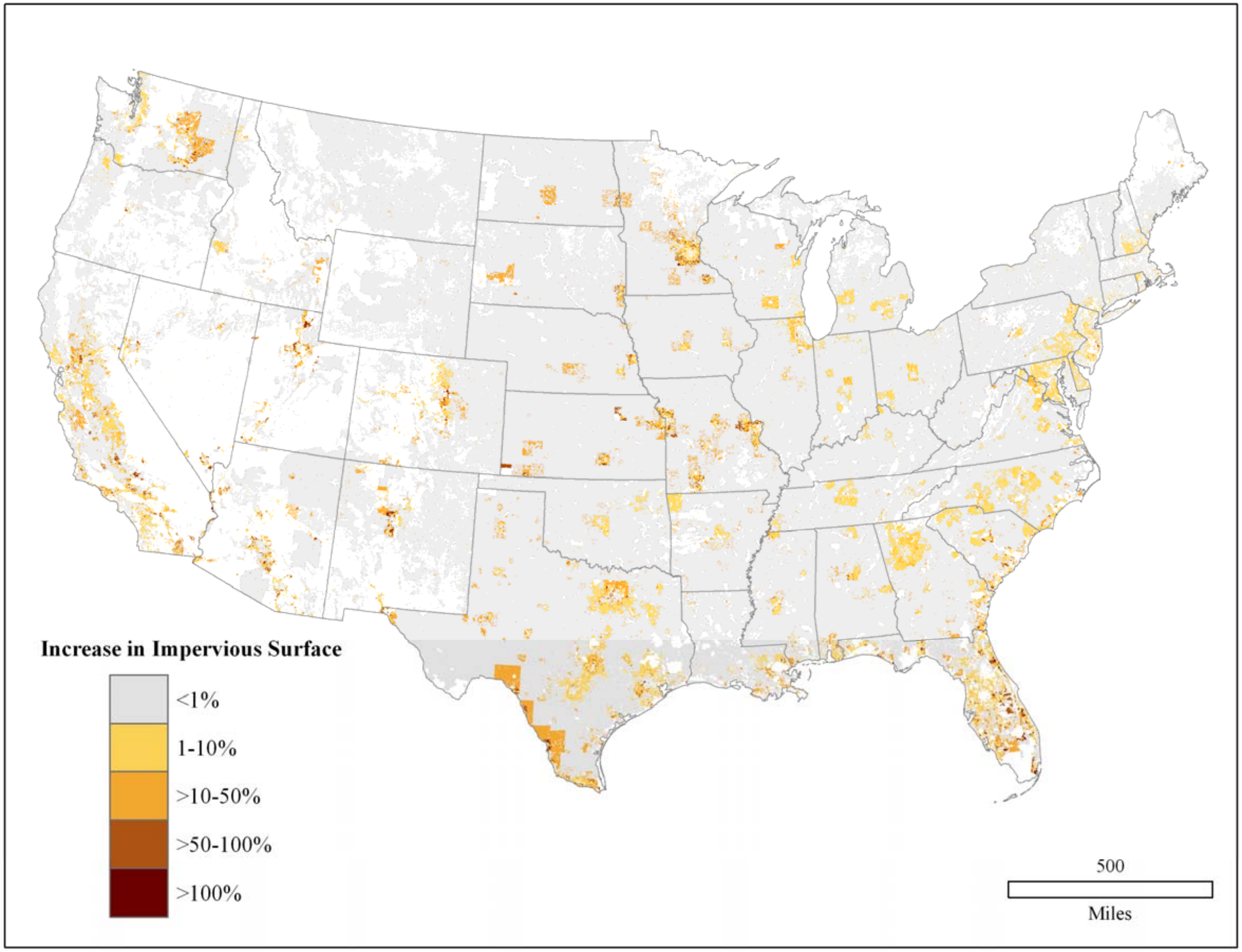

Figure 5-20. 2000−2050 relative change in impervious surface, A1 storyline.

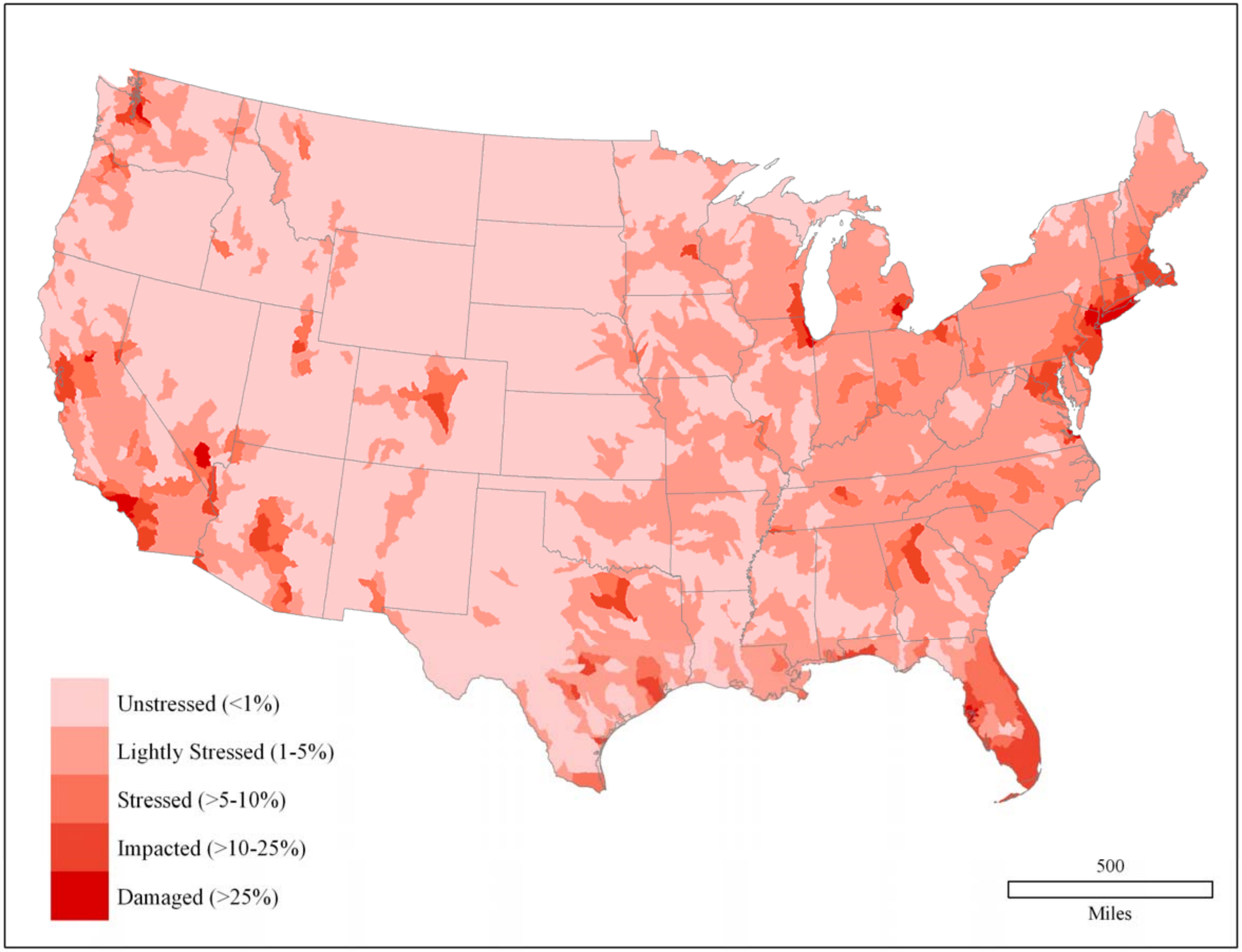

Figure 5-21. 2050 impervious surface, A2 storyline.

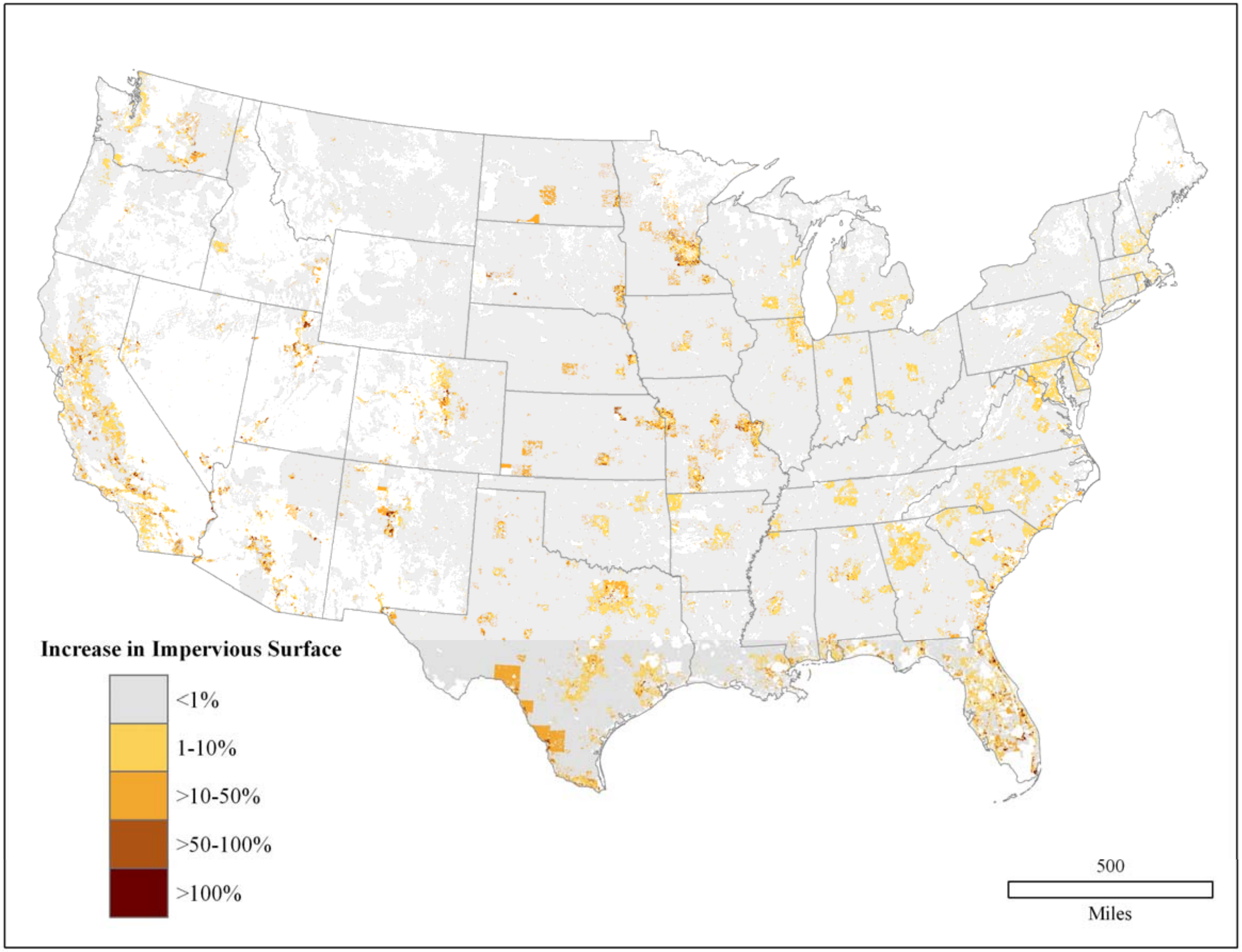

Figure 5-22. 2000−2050 relative change in impervious surface, A2 storyline.

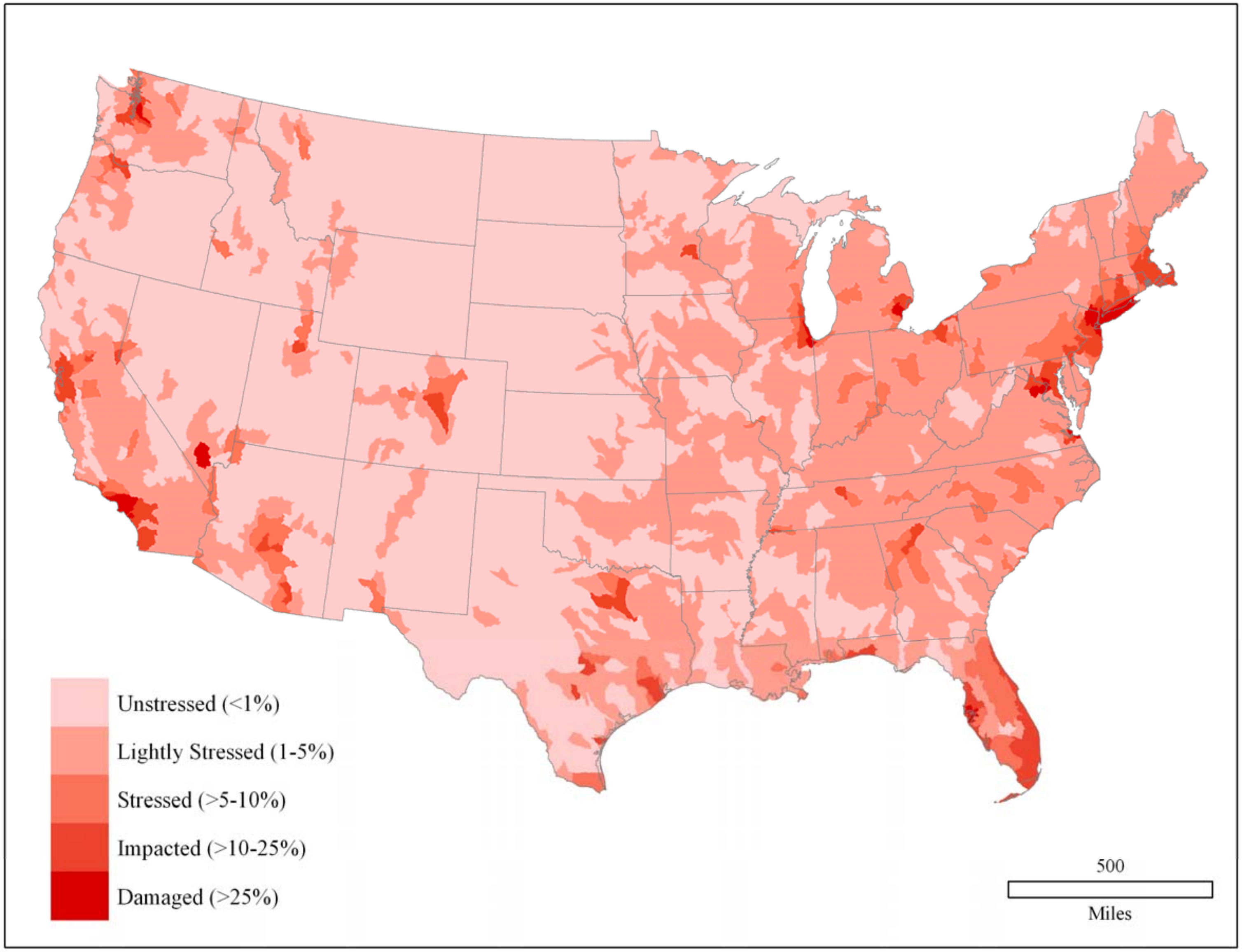

Figure 5-23. 2050 impervious surface, B1 storyline.

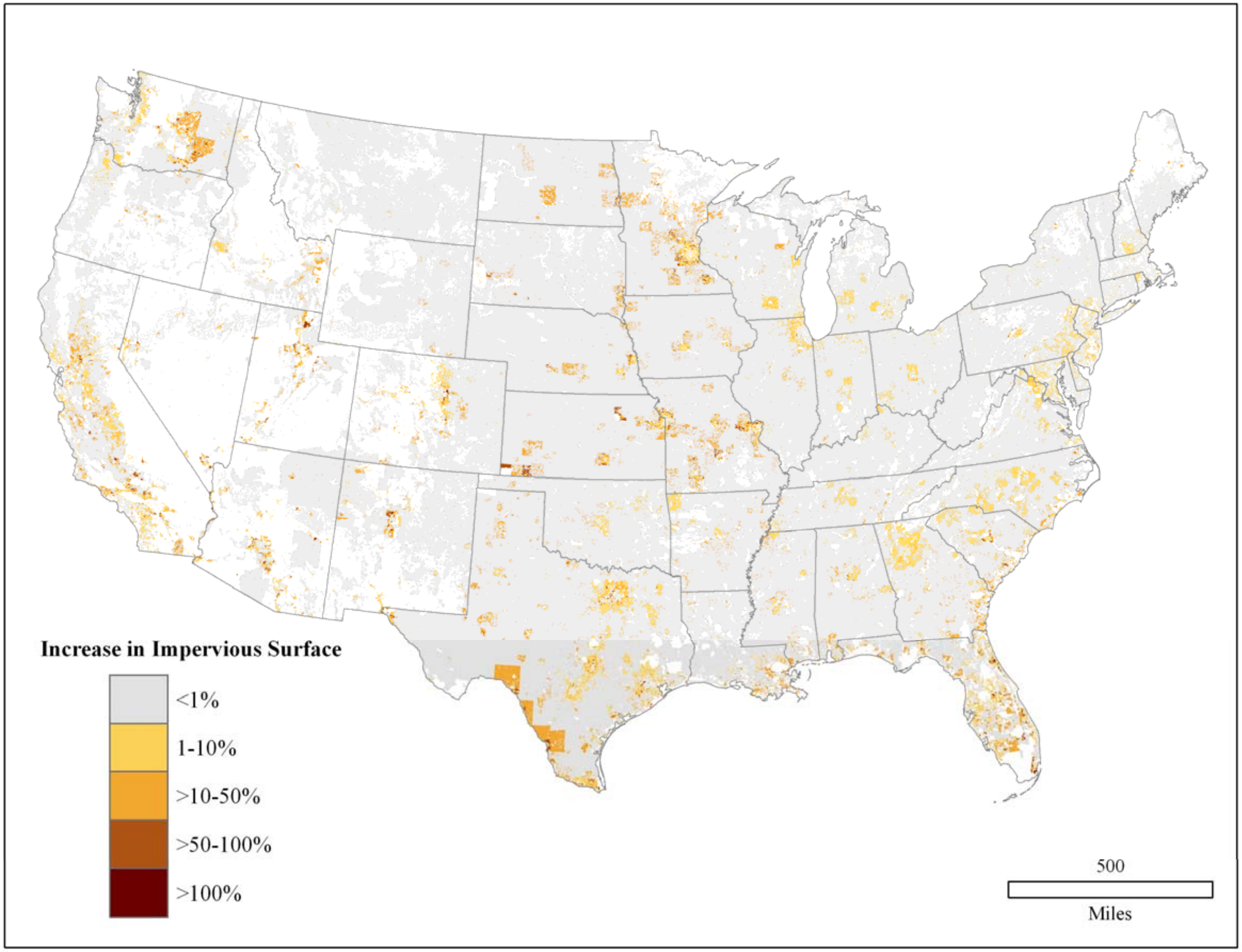

Figure 5-24. 2000−2050 relative change in impervious surface, B1 storyline.

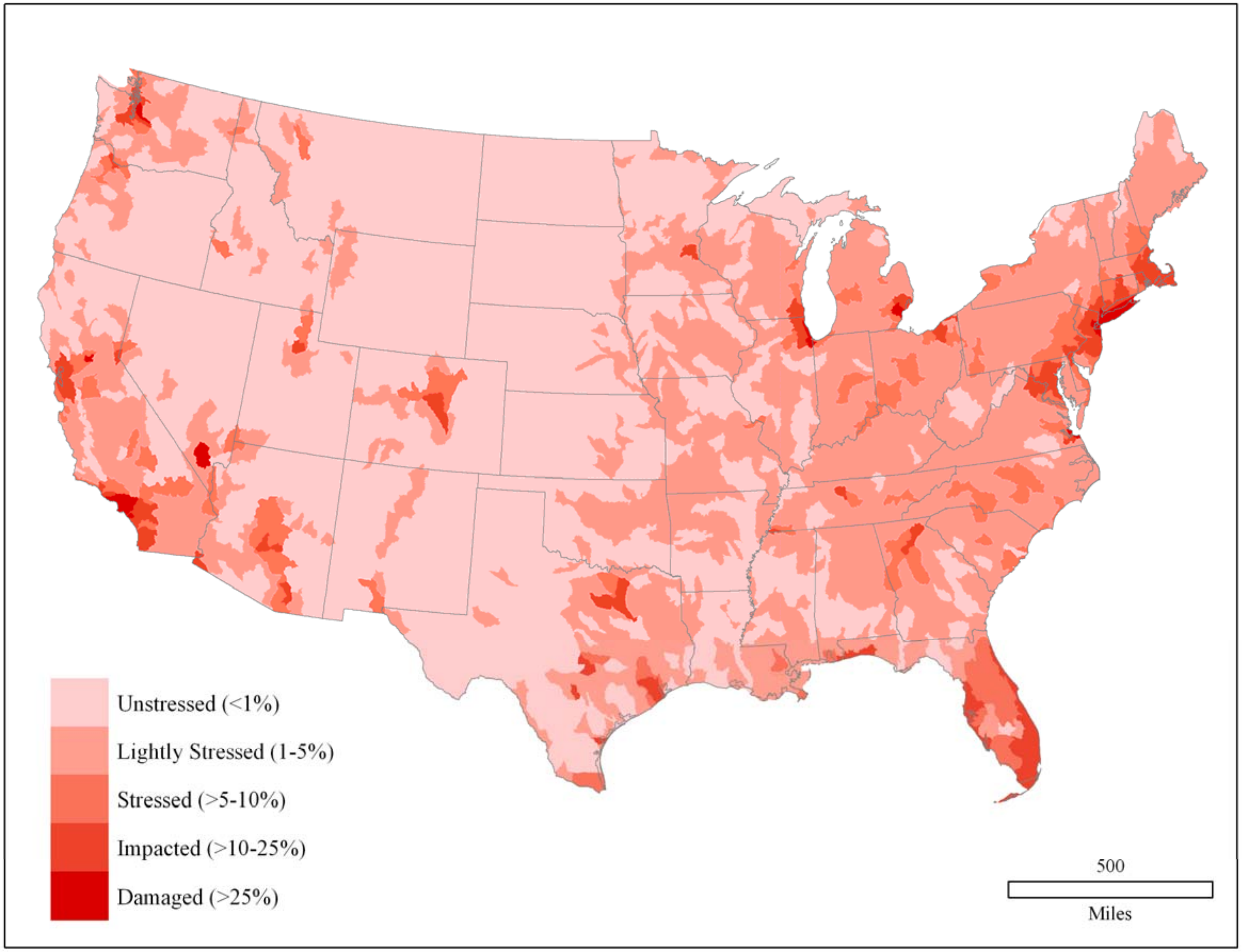

Figure 5-25. 2050 impervious surface, B2 storyline.

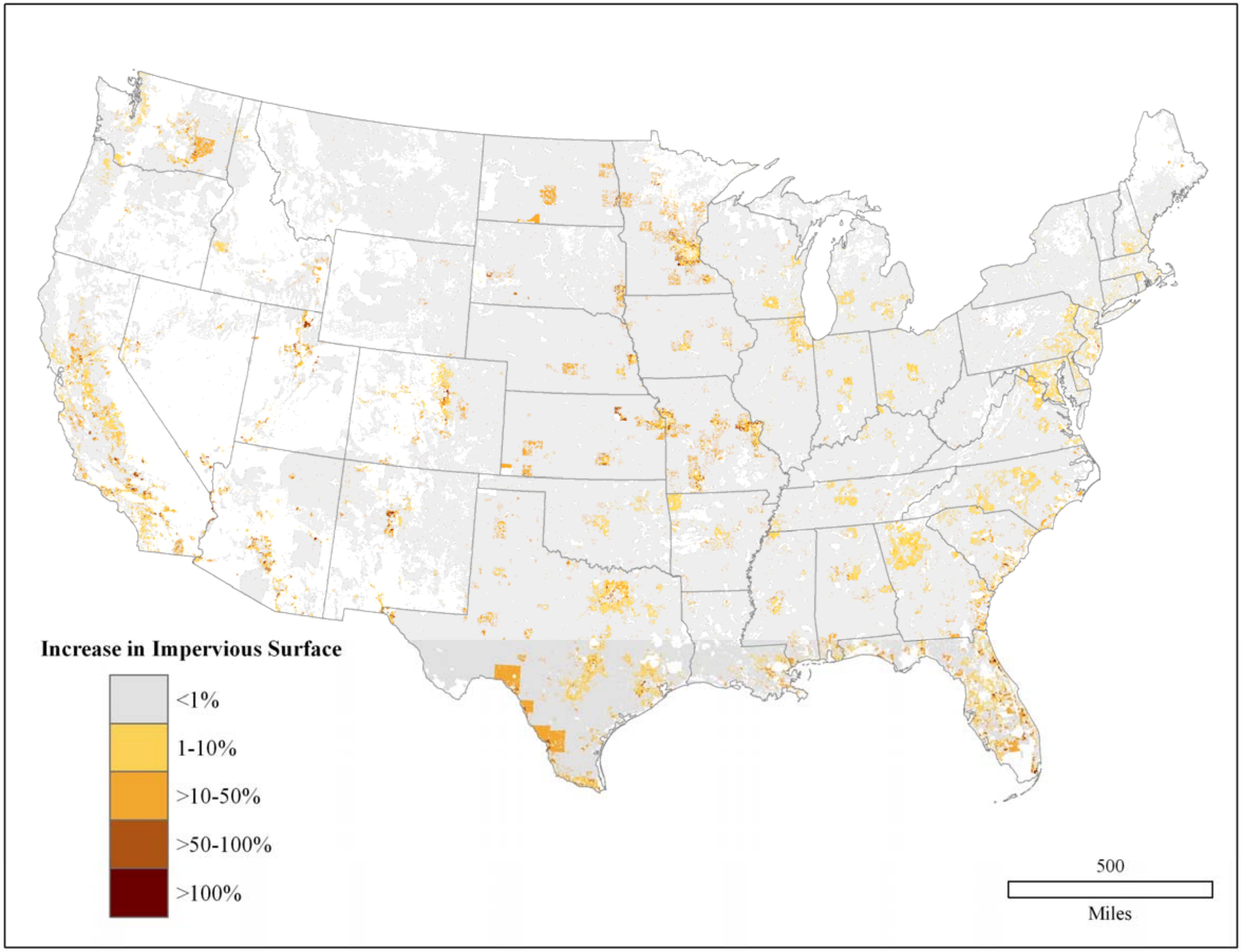

Figure 5-26. 2000−2050 relative change in impervious surface, B2 storyline.

5.4. OPTIONS FOR FUTURE STUDY

IS calculations and regional growth rates are just the first set of many possible analyses using results from this project. The HD and population projections can inform modeling exercises that consider such diverse areas of research as traffic volumes, air quality, and water quality. As a part of this first phase of the ICLUS project we will distribute map products on the Web and through an ArcGIS tool. This tool will allow users to customize maps to a certain extent, by varying some of the allocation parameters (i.e., travel times) and by selecting smaller spatial extents (i.e., several states or a single state).

In subsequent phases of this project the demographic and housing allocation models could be further modified to incorporate climate change variables, incorporate additional factors affecting population change and housing patterns, or consider specific policy responses such as an emphasis on Smart Growth or other low-impact development patterns. While efforts here benefit from a whole-country approach, regional differences in planning efforts and local zoning laws are likely to have a significant impact on land use. A future expansion on this work could include regional variation in how populations grow. The current set of scenarios also does not reflect the effects of climate change on development patterns. Some climate change effects, e.g., sea level rise, are likely to have significant impacts on development patterns by the end of the century. By integrating projected changes in sea level, projected population growth, and land development patterns, the resulting scenarios would provide valuable information to planners interested in coastal development, transportation infrastructure vulnerability, HD, water quality (e.g., salt water intrusion), and other endpoints of concern. Such integration is a possible next step in the development of this project in order to begin to incorporate climate change variables into the models to facilitate more comprehensive assessments of the combined effects of climate change and land-use change. Additional options for incorporating climate change variables include modifications to the gravity model to allow climate variables to change over time.

Another possible modification might examine changes in HD regionally. Current HD class ranges are static and the same throughout the country. This can be easily modified geographically, particularly west versus east to explore different development patterns regionally, such as larger ranches or farms in the western United States (e.g., 10−40 acres exurban versus ~2−10 acres exurban in the east). Also, the amount of developable land is currently assumed to be static, but that could be modified to remove additional protected lands. However, an approach would first need to be developed to identify where these specific lands may be protected and what the relationship to new housing allocations and densities might be.

The driving factor in SERGoM is population, so that forecasts are exogenous variables that are input to the model. Other population scenarios could also be modeled. In the current version of SERGoM housing units do not move across boundaries if an analytical unit is saturated. One reason for this is that it would require coupling SERGoM with the demographic model (that is, at each decade SERGoM would have to pass back to the ICLUS demographic model the actual number of people/households that were placed in a county). Rather, we endeavored to make the ICLUS demographic model more responsive to broader-scale (county and up) trends by using the amenity-based gravity model. A future refinement could be to couple the demographic model and SERGoM to ensure explicit distribution of people/houses. Another way to handle this challenge, if data were available and assumptions were reasonable, is to estimate the "build-out" (or carrying capacity) of a county based on zoning, for example.

HD currently is spread from one location to another (even across county or state analytical units) as a function of distance away (travel time) from urban core areas. Urban "core areas" are identified by a high HD level (e.g., urban) of a size/population/area to approximate providing general services. In the ICLUS scenarios these parameters are specified by two user-defined values that do not vary across the study area (e.g., nationwide model). These could be easily changed so that they are spatially-explicit parameters, which would allow regional variation to occur. Also note that as growth continues through time, eventually some new urban core areas will "emerge", creating a new "hot spots" or concentrations of growth.

Some options for future research include:

- **Identify new urban clusters.** Currently, new urban core areas are modeled by SERGoM in the spatial allocation process. An analysis of the detailed SERGoM outputs would help identify new or expanded urban clusters. Such new urban clusters would likely experience dramatic changes in air quality, water quality, and traffic. Identifying these growth areas could help improve regional analyses and planning.

- **Estimate traffic demands.** Vehicle miles traveled (VMT) can be estimated for each grid cell based on the number of housing units and available estimates on the number of automobile trips generated per day that vary based on HD class. When combined with average trip lengths, these projections can be used to estimate fuel consumption, travel demand, and other factors. Combining the HD maps with the road network layer of SERGoM can allow more sophisticated traffic analyses.

- **Model air quality changes.** The VMT outputs above, along with other layers of data about stationary emissions sources, can be used with existing air quality models to estimate projected air quality under the various scenarios.

- **Analyze effect of IS on water quality.** The IS analysis can be used to consider the quality of stormwater runoff and its impact on water quality in key watersheds. Groisman et al. (2005) suggest that one potential impact of climate change is an increase in the intensity of individual storm events. Since these events are responsible for the majority of impacts to water quality from stormwater runoff, examining the possible extent of ISs become even more important given the anticipated impacts of climate change.

- **Develop alternative Smart Growth scenarios.** Alternative SERGoM runs could be used to project HD under alternative development patterns that reflect Smart Growth goals. One example would be to model denser development along existing transportation infrastructure. The results could also be combined with some of the other suggested analyses to estimate the performance of such strategies. Another use would be to examine the amount of growth that could be accommodated in brown/greyfield sites, versus greenfield sites.

- **Analyze effect of urban versus exurban growth on IS thresholds.** Another interesting question to ask in the future would be the degree to which growth in urban/suburban versus exurban classes is the main cause for watersheds to cross over a threshold (e.g., >5%) into the stressed IS classification. It would also be interesting to explore the consequences of assumptions about how population per housing unit would vary both between urban and rural areas, and through time.

- **Evaluate forecasted housing density patterns to more local/regional land-use models.** It would be useful to evaluate the SERGoM estimated housing density patterns against more regional and structural models of land-use change (e.g., CLUE-S; SLEUTH; etc.).

- **Explore a stronger coupling of agricultural land conversion to residential.** Typically agricultural land (either rangeland or cropland) is converted when undergoing increases in residential land use, which has important implications for environmental endpoints. It would be useful in future efforts to explore a stronger coupling of residential land-use change with agricultural land-use change in order to understand trade-offs between these land uses.

6. DISCUSSION AND CONCLUSIONS

The models and results described in this report represent a first step in creating national-scale scenarios of HD patterns that are broadly consistent with IPCC emissions storylines. The IPCC emissions storylines were adapted to the United States by modifying Census demographic projections, by incorporating county-to-county connectivity and amenity variables, and by modifying the spatial growth pattern of housing. The resulting scenarios provide benchmarks for possible future HD patterns. The scenarios are unique in terms of presenting these results for the conterminous United States to the year 2100 in a manner consistent with IPCC emissions storylines. Other available land-use models are still spatially more constrained or require data not available consistently at the national scale.

Validation of models projecting a future condition is generally challenging. We compared our demographic projections at the county level with selected state estimates to show that our model is in the range of other projections, and therefore can be used to generate plausible population scenarios. The differences underscore that there are many approaches that can generate scenarios, and that these approaches may use more detailed or finer-scale information. However, the ICLUS methodology does produce demographic results within the range projected by other efforts and is therefore useful in benchmarking scenarios within established GHG emissions storylines. Subsequent model improvements will attempt to use other datasets for additional validation studies, particularly of domestic migration patterns that contribute to the demographic projections.

The preliminary results presented in this report call attention to the expected spatial variability of land-use effects and their possible intersection with regional climate changes. For examples, population growth rates in the South and West may increase the vulnerability of these regions to water quantity and quality issues if precipitation decreases. This is one way that these land-use scenarios can facilitate integrated assessments of climate change and land-use change. While there is a high degree of uncertainty associated with the specific projections, the comparison of the scenarios in these types of assessments may still be useful in highlighting potential vulnerabilities.

Our preliminary results also show differences in the impacts on various land-cover classes, which, along with increases in IS cover, are likely to translate to effects on air quality, human health, water quality, and ecosystems. For example, IS cover may serve as a surrogate for estimating certain types of emissions associated with buildings and roads and as an input to stormwater runoff models. Further assessments of effects from changes in HD and IS cover, including more detailed spatial analyses of which watersheds and regions may be more vulnerable to these changes; which wetland types and in which watersheds and regions may be more impacted by land conversion; and which regions may benefit more from policies and planning that includes Smart Growth development patterns, will be important next steps. The explicit integration of climate change into the next phase of modeling will further facilitate these assessments.

In conclusion, the U.S.-adapted SRES storylines produce a range of outcomes both in the demographic model of the ICLUS project and in the spatial allocation model. The range of population projections, housing densities, and IS cover allows for a broad examination of trends and impacts to a variety of endpoints. The ICLUS methodology also allows for future modifications that can incorporate more explicit climate-change information, feedbacks to domestic migration patterns from emerging growth centers, and a variety of regional changes to housing densities and allocation preferences. The current outputs from the ICLUS project can be used in a variety of assessments that include effects on air quality, water quality, and any other endpoints that are modeled using either population or land use as an input, especially when these assessments compare differences among scenarios.

REFERENCES

Alig, RJ; Kline, JD; Lichtenstein, M. (2004) Urbanization on the US landscape: looking ahead in the 21st century. Landscape Urban Plan 69:219−234. Available online at http://www.fsl.orst.edu/lulcd/Publicationsalpha_files/Alig_etal_2004_LUP.pdf.

Batty, M. (1997) Cellular automata and urban form: a primer. J Am Plan Assoc 63(2):266−274.

Breiman, L; Friedman, J; Olshen, R; et al. (1984) Classification and regression trees. New York: Chapman and Hall; 358 pp.

Campbell, PR. (1996) Population projections for states by age, sex, race, and Hispanic origin: 1995 to 2025" U.S. Census Bureau. Population Division, PPL-47. Available online at http://www.census.gov/population/www/projections/ppl47.html.

Claggett, PR; Jantz, CA; Goetz, SJ; et al. (2004) Assessing development pressure in the Chesapeake Bay watershed: an evaluation of two land-use change models. Environ Monit Assess 94(1−3):129−146.

Clarke, KC; Gaydos, LJ. (1998) Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int J Geogr Inf Sci 12:699−714.

Conley, TG; Topa, G. (2002) Socioeconomic distance and spatial patterns in unemployment. J Appl Econom 17(4):303−327.

CBI (Conservation Biology Institute). (2008) Protected Areas Database v4. CBI, Corvallis, OR. Available online at http://www.consbio.org/what-we-do/protected-areas-database-pad-version-4.

De'ath, G; Fabricius, KE. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecol*.,* 81(11):3178−3192. Available online at http://161.111.161.171/estadistica2004/e0081death.pdf.

Elvidge, CD; Milesi, C; Dietz, JB; et al. (2004) US constructed area approaches the size of Ohio. EOS Tran Am Geophy Union 85(24):233−240. Available online at http://www.co.suffolk.ny.us/stormwater/FAQs/ImperviousArticle.pdf.

Elvidge, CD; Tuttle, BT; Sutton, PC; et al. (2007) Global distribution and density of constructed impervious surfaces. *Sensors* 7: 1962−1979. Available online at http://www.ngdc.noaa.gov/dmsp/pubs/ISAglobal_20070921-1.pdf.

Engelen, G; White, R; De Nijs, T. (2003) Environment explorer: a spatial support system for the integrated assessment of socioeconomic and environmental policies in the Netherlands. Integrat Assess 4(2):97−105.

Exum, LR; Bird, SL; Harrison, J; Perkins, CA. (2005) Estimating and projecting impervious cover in the southeastern United States. U.S. Environmental Protection Agency, Ecosystems Research Division, Athens, GA; EPA/600/R-05/061. Available online at http://www.epa.gov/athens/publications/reports/Exum600R05061EstimatingandProjectingImpervious.pdf.

Frazer, L. (2005) Paving paradise: the peril of impervious surfaces. Environ Health Perspect 113(7):A457−A462.

Friedl, MA; Brodley, CE. (1997) Decision tree classification of land cover from remotely sensed data. Rem Sens Env 61(3):399−409.

Groisman, PY; Knight, RW; Easterling, DR; et al. (2005) Trends in intense precipitation in the climate record. J Clim 18(9):1326−1350.

Manson, GA; Groop, RE. (2000) U.S. intercounty migration in the 1990s: people and income move down the urban hierarchy. Prof Geogr 52(3):493−504.

Haynes, KE; Fotheringham, AS. (1984) Gravity and spatial interaction models. Beverly Hills: Sage Publications.

Hilderink, H. (2004) Population and scenarios: worlds to win? RIVM Report 550012001. Bilthoven, The Netherlands. Available online at http://www.rivm.nl/bibliotheek/rapporten/550012001.pdf.

Hollman, FW; Mulder, TJ; Kallan, JE. (2000) Methodology and assumptions for the population projections of the United States: 1999 to 2100. Population division working paper No. 38. Bureau of the Census. Washington, DC. Available online at http://www.census.gov/population/www/documentation/twps0038.html.

Homer, C; Huang, C; Yang, L; et al. (2004) Development of a 2001 national landcover database for the United States. Photogramm Eng Rem Sens 70(7):829−840. Available online at http://www.mrlc.gov/pdf/July_PERS.pdf.

IPCC (Intergovernmental Panel on Climate Change). (2001) Climate change 2001: impacts, adaptation and vulnerability. Contribution of Working Group II to the Third Assessment Report of the IPCC. JJ; McCarthy, OF; Canziani, NA; et al., eds., Cambridge, UK: Cambridge University Press. Available online at http://www.grida.no/publications/other/ipcc_tar/?src=/climate/ipcc_tar/wg2/index.htm.

IPCC (Intergovernmental Panel on Climate Change). (2007) Climate change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the IPCC. Parry, ML; Canziani, OF; Palutikof, JP; et al., eds. Cambridge UK: Cambridge University Press; 976pp. Available online at http://www.gtp89.dial.pipex.com/intro.pdf.

Irwin, EG; Bockstael, NE. (2002). Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Pattern. Journal of Economic Geography 2(1), January: 31-54.

Jantz , CA; Goetz, SJ; Shelley, MK. (2003) Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. Environ Plan B: Plan Design 31(2):251-271.

Jiang, L; O'Neill, BC. (2007) Impacts of demographic trends on US household size and structure. Popul Dev Rev 33(3):567−591.

Landis, J; Zhang, M. (1998) The second generation of the California urban futures model. Part 1: Model logic and theory. Environ Plan B: Plan Design 25:657-666.

Lee, R. (2003) The demographic transition: three centuries of fundamental change. J Econ Persp, 17(4):167−190. Available online at http://www.soc.upenn.edu/courses/2005/spring/soc796_hpkohler/readings/lee03t.pdf.

Lutz, W; Sanderson, WC; Scherbov, S; et al. (1996) World population scenarios for the 21st century. In: Lutz, W, ed. The future population of the world. What Can We Assume Today? W. Lutz. London: Earthscan; pp. 361−396.

Matthews, TJ; Hamilton, B. (2005) Trend analysis of the sex ratio at birth in the United States. Natl Vital Stat Rep 53(20):1−17.

McGranahan, D. (1999) Natural amenities drive rural population change. Agricultural Economic Report No. AER781. U.S. Department of Agriculture. Available online at http://www.ers.usda.gov/Publications/AER781/.

MEA (Millennium Ecosystem Assessment). (2005) Ecosystems and Human Well-Being: Volume 2 Scenarios: Findings of the Scenarios Working Group. Washington: Island Press. 596 pp.

Nakićenović, N; Alcamo, J; Davis, G; et al. (2000) Special report on emissions scenarios. Intergovernmental Panel on Climate Change (IPCC). Cambridge UK: Cambridge University Press. Available online at http://www.grida.no/publications/other/ipcc_sr/?src=/climate/ipcc/emission/.

Nowak, D; Walton, JT. (2005) Projected urban growth (2000-2050) and its estimated impact on the US forest resource. J Forest. 103(8):383−389.

O'Neill, BC. (2004) Conditional probabilistic population projections: an application to climate change. Int Stat Rev 72(2)167−184. Available online at http://www.iiasa.ac.at/Research/PCC/pubs/sci-pol/ONeill2004_ISR.pdf.

O'Neill, BC. (2005a) US socioeconomic futures. International network to advance climate talks, options for future climate policy: transatlantic perspectives October 2005. Available online at http://www.iiasa.ac.at/Research/PCC/pubs/dem-emiss/ONeill2005_INTACT_final.pdf.

O'Neill, BC. (2005b) Population scenarios based on probabilistic projections: an application for the Millennium Ecosystem Assessment. Popul Environ 26(3):229−254.

Parker, D; Hessl, A; Davis, SC. (2008) Complexity, land-use modeling, and the human dimension: fundamental challenges for mapping unknown outcome spaces. Geoforum 39:789−804.

Perry, M; Schachter, J. (2003) Migration of natives and the foreign born: 1995 to 2000. Census 2000 special reports. U.S. Census Bureau. Available online at http://www.census.gov/prod/2003pubs/censr-11.pdf.

Pontius, RG, Jr. (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm Eng Rem Sens 68(10):1041−1049.

Reginster, I; Rounsevell, M. (2006) Scenarios of future urban land use in Europe. Environ Plann B: Plann Design 33(4):619−636.

Rodrigue, JP; Comtois, C; Slack, B. (2006) The geography of transport systems. New York: Routledge.

Rogers, A; Henning, S. (1999) The internal migration patterns of the foreign-born and native-born populations in the United States: 1975−80 and 1985−90. Int Migrat Rev 33(2)403−429.

Rounsevell, MDA; Reginster, I; Araujo, MB. (2006) A coherent set of future land-use change scenarios for Europe. Agri Eco Environ 114:57−68.

Sardon, JP; Robertson, DA. (2004) Recent demographic trends in the developed countries. Population 59(2):263−314.

Saaty, TL. (1980) The Analytic Hierarchy Process. New York, NY: McGraw-Hill.

Slonecker, ET; Tilley, JS. (2004) An evaluation of the individual components and accuracies associated with the determination of impervious area. GISci Rem Sens 41(2):165−184.

Siegel, JS; Swanson, DA; eds. (2004) The methods and materials of demography, second edition. San Diego: Elsevier Academic Press.

Solecki, WD; Olivieri, C. (2004) Downscaling climate change scenarios in an urban land-use change model. J Environ Mod 74:105−115.

Theobald, DM. (2001) Land-use dynamics beyond the American urban fringe. Geogr Rev 91(3):544−564. Available online at http://www.amergeog.org/gr/jul01/Theobald-1st.pdf.

Theobald, DM. (2003) Targeting conservation action through assessment of protection and exurban threats. Conserv Biol 17(6):1624−1637.

Theobald, DM. (2005) Landscape patterns of exurban growth in the USA from 1980 to 2020. Ecol Soc 10(1):32. Available online at http://www.ecologyandsociety.org/vol10/iss1/art32/.

Theobald, DM; Hobbs, NT. (1998) Forecasting rural land-use change: a comparison of regression- and spatial transition-based models. Geogr Environ Mod 2(1):65−82.

Theobald, DM; Norman, JB; Sherburne, MR. (2006) FunConn v1 user's manual: ArcGIS tools for functional connectivity modeling. Natural Resource Ecology Lab, Colorado State University. Available online at http://www.nrel.colostate.edu/projects/starmap/FUNCONN%20Users%20Manual_public.pdf.

Theobald, DM; Stevens, Jr, DL; White, D; et al. (2007) Using GIS to generate spatially-balanced random survey designs for natural resource applications. Environ Manag 40(1):134−146.

Theobald, DM; Goetz, SJ; Norman, J; et al. (2009) Watersheds at risk to increased impervious surface in the conterminous US. J Hydrol Eng 14(4):362−268.

UN (United Nations). (1998) World Population Projections to 2150. Department of Economic and Social Affairs, Population Division, United Nations, New York.

United States Census Bureau. (1992) 1980 to 1990 demographic components of change file of the U.S., states, and counties. Available online at http://www.census.gov/popest/archives/1980s/.

United States Census Bureau. (2000) Component assumptions of the resident population by age, sex, race, and Hispanic origin: lowest, middle, and highest series, 1999 to 2100. Available online at http://www.census.gov/population/www/projections/natdet-D5.html.

United States Census Bureau. (2003) Census 2000: 5-Percent public use microdata sample (PUMS) files. Available online at http://www.census.gov/Press-Release/www/2003/PUMS5.html.

United States Census Bureau. (2008) Population, population change and estimated components of population change: April 1, 2000 to July 1, 2007 (NST-EST2007-alldata). Available online at http://www.census.gov/popest/datasets.html.

United States Census Bureau. (2007) Census 2000. Available online at http://factfinder.census.gov.

Verburg, PH; Soepboer, W; Veldkamp, A; et al. (2002) Modeling the spatial dynamics of regional land use: the CLUE-S model. Environ Manag 30(3):391−405.

Verburg, P; Kok, K; Pontius, RG. (2006) Modeling land-use and land-cover change. In: Lambin, E; Geist, H; eds, Land-use and land-cover change: local processes, global impacts. New York, NY: Springer.

Waddell, P. (2002) UrbanSim: modeling urban development for land use, transportation and environmental planning. J Am Plann Assoc 68:297−314.

APPENDIX A

MAPS FOR ICLUS SCENARIOS

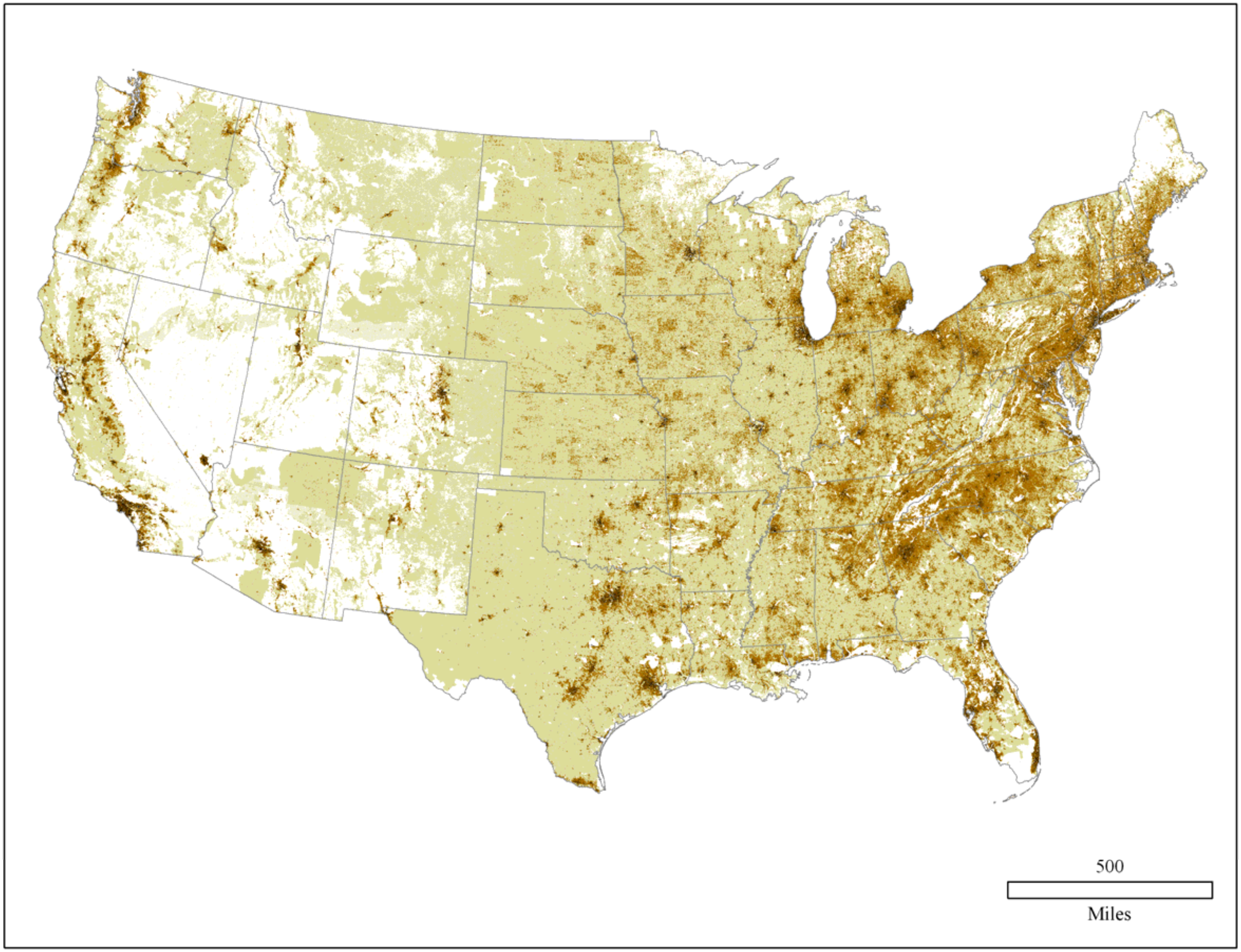

Figure A-1. Base case, year 2010 housing density map.

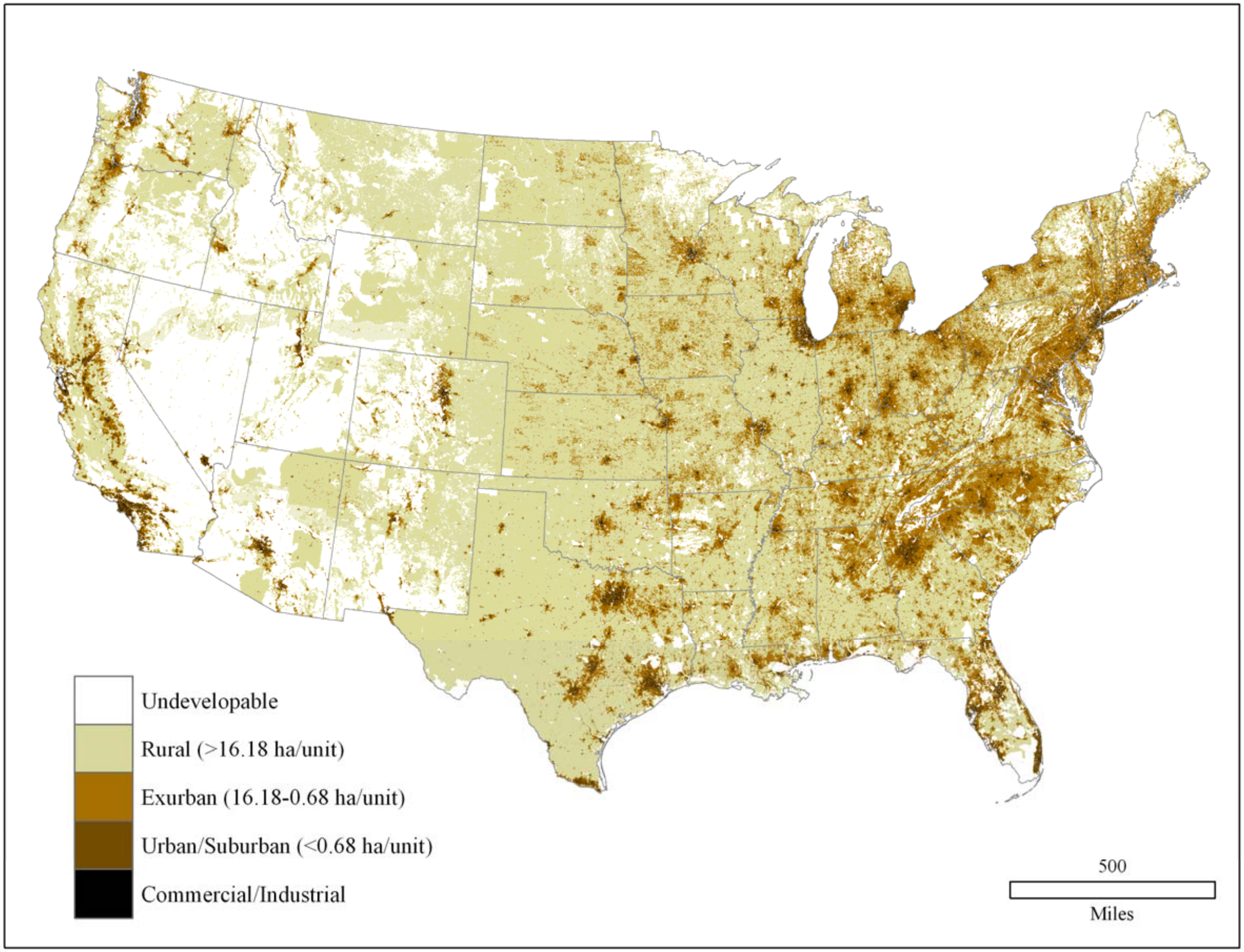

Figure A-2. Base case storyline, year 2050 housing density map.

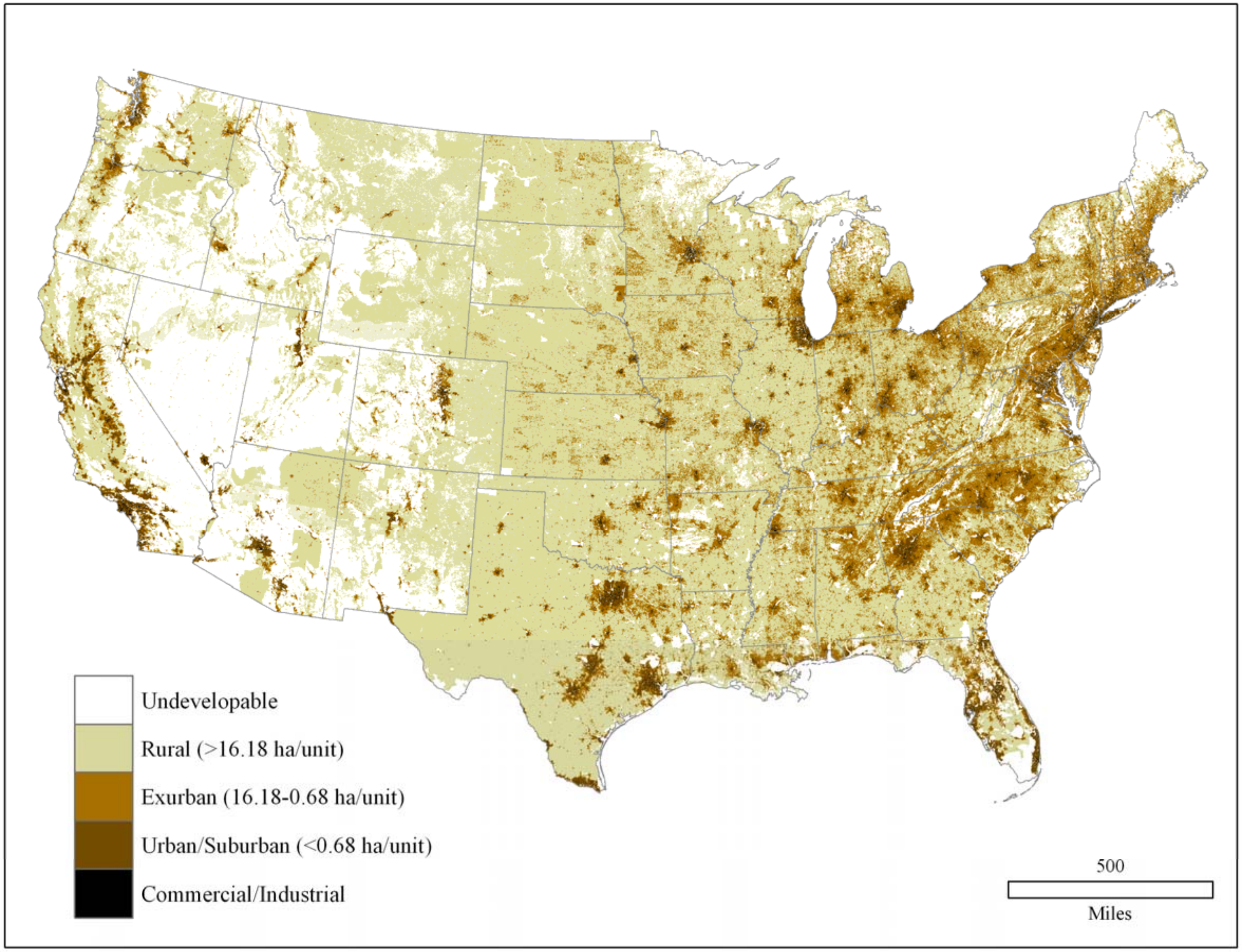

Figure A-3. Base case, year 2100 housing density map.

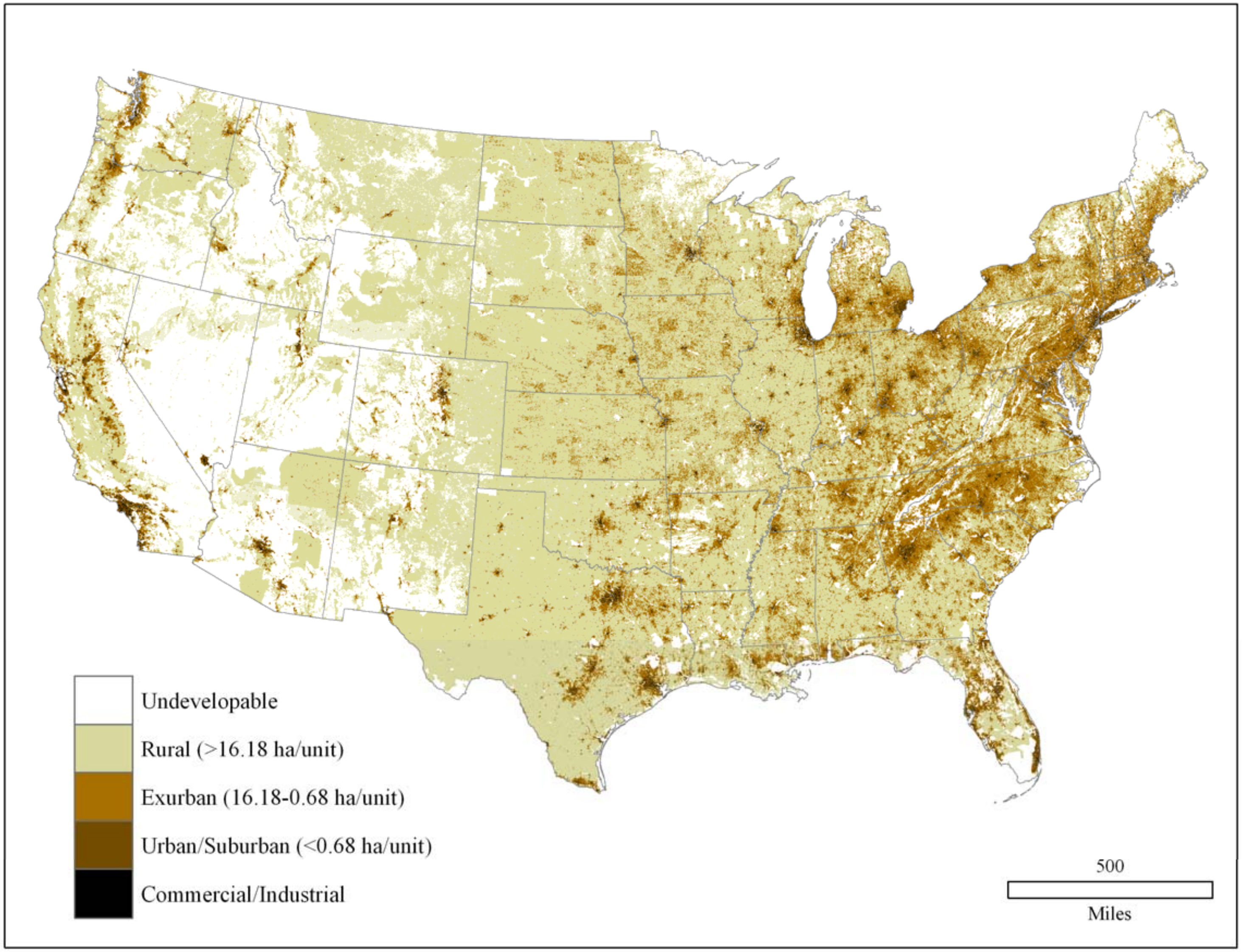

Figure A-4. A1 storyline, year 2010 housing density map.

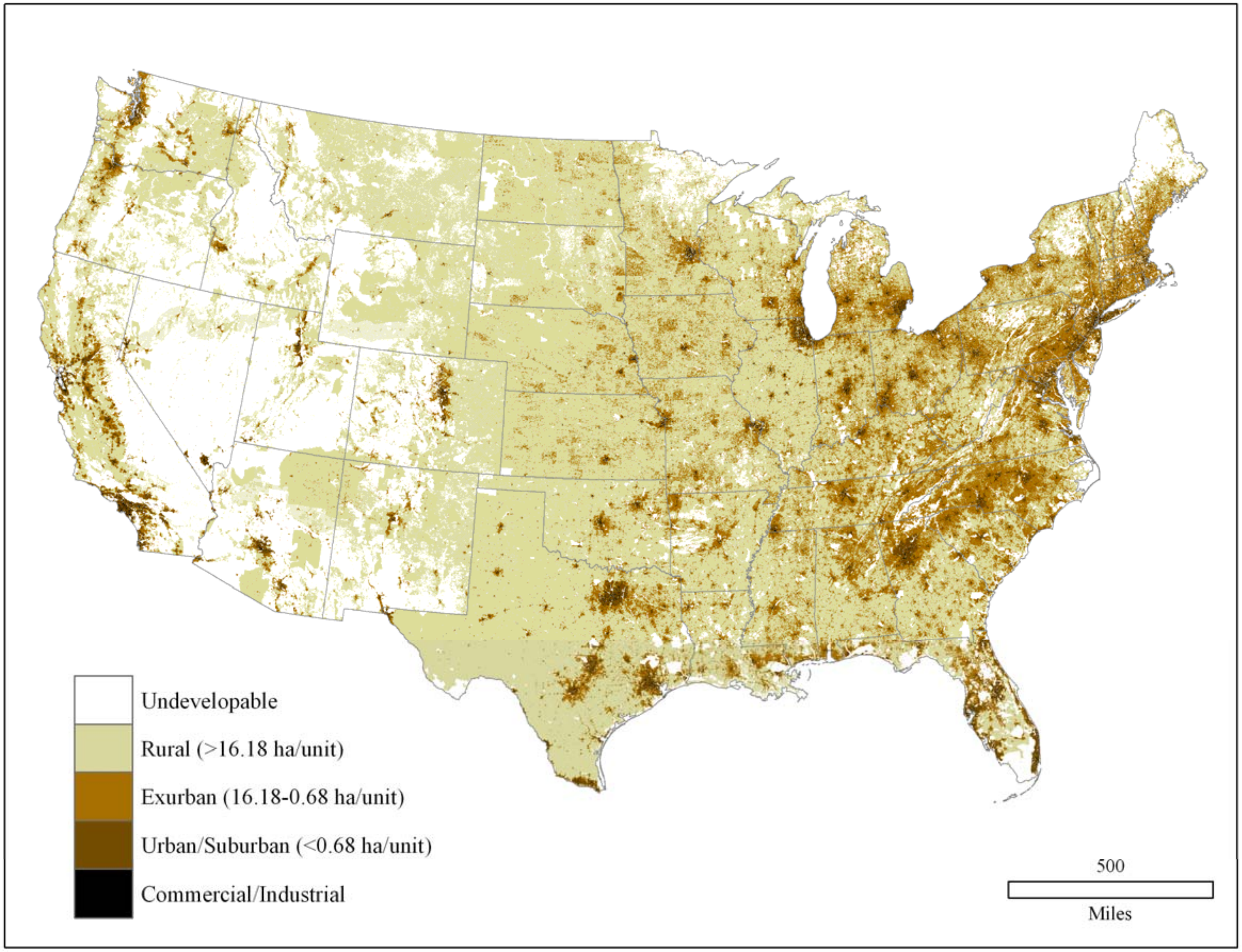

Figure A-5. A1 storyline, year 2050 housing density map.

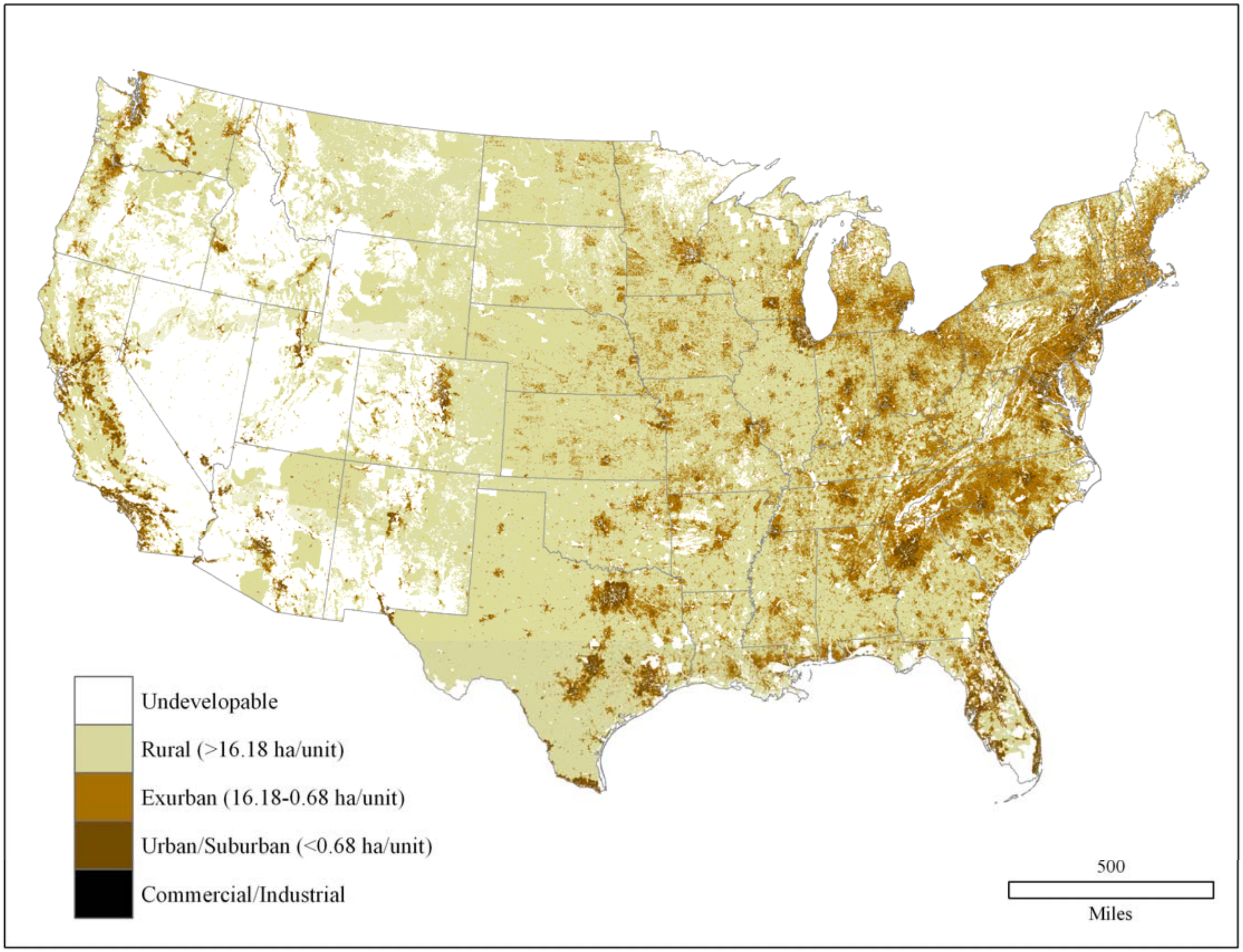

Figure A-6. A1 storyline, year 2100 housing density map.

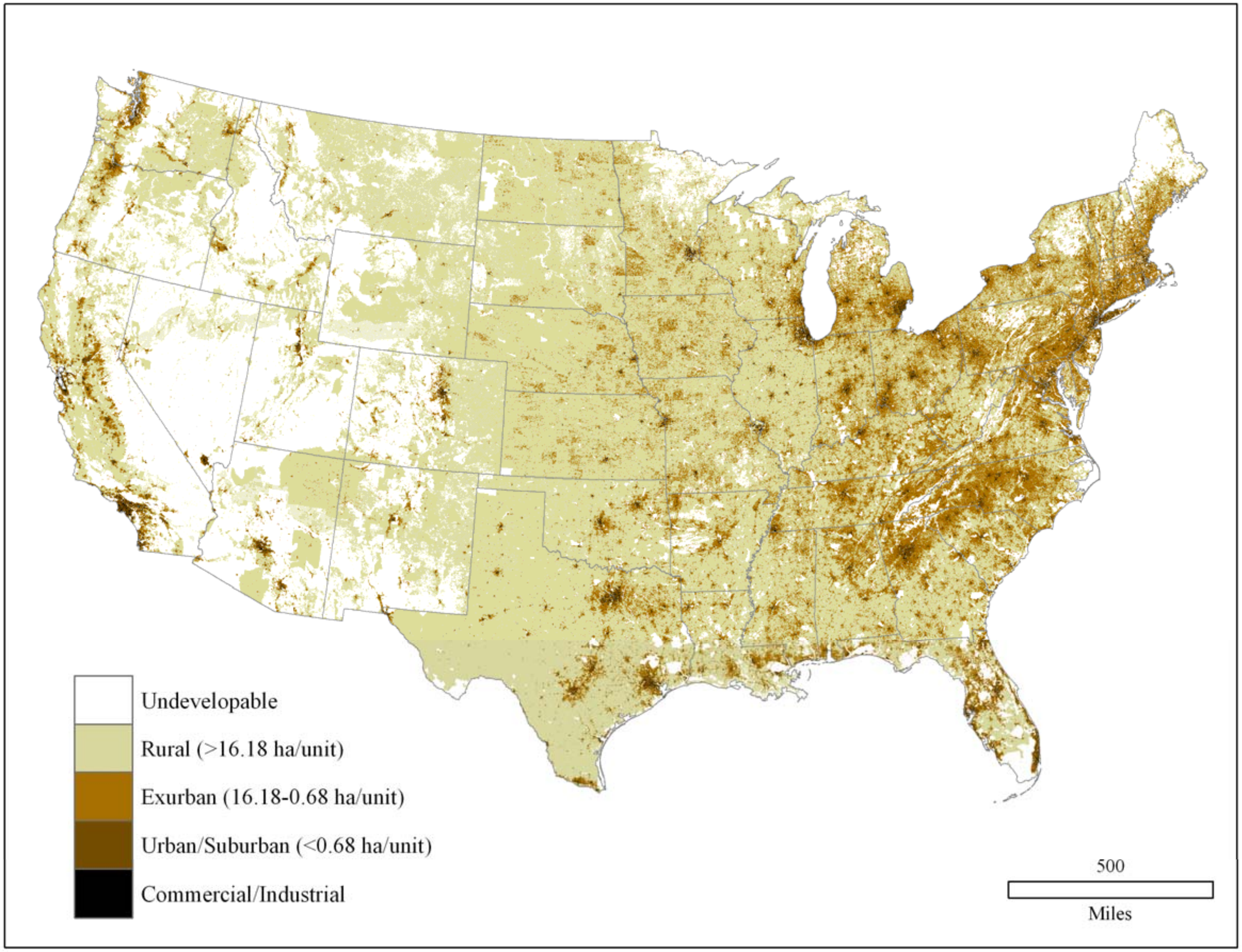

Figure A-7. A2 storyline, year 2010 housing density map.

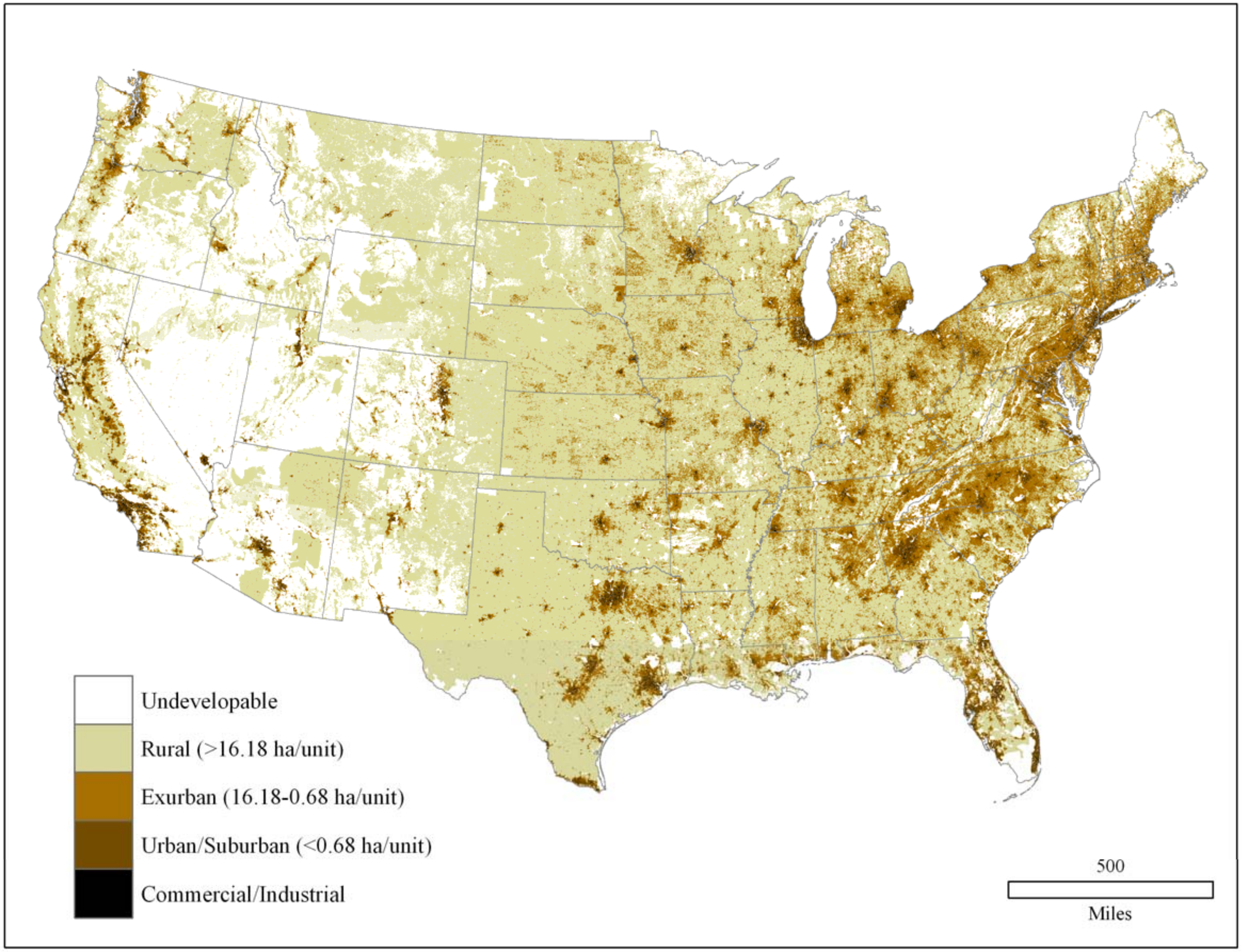

Figure A-8. A2 storyline, year 2050 housing density map.

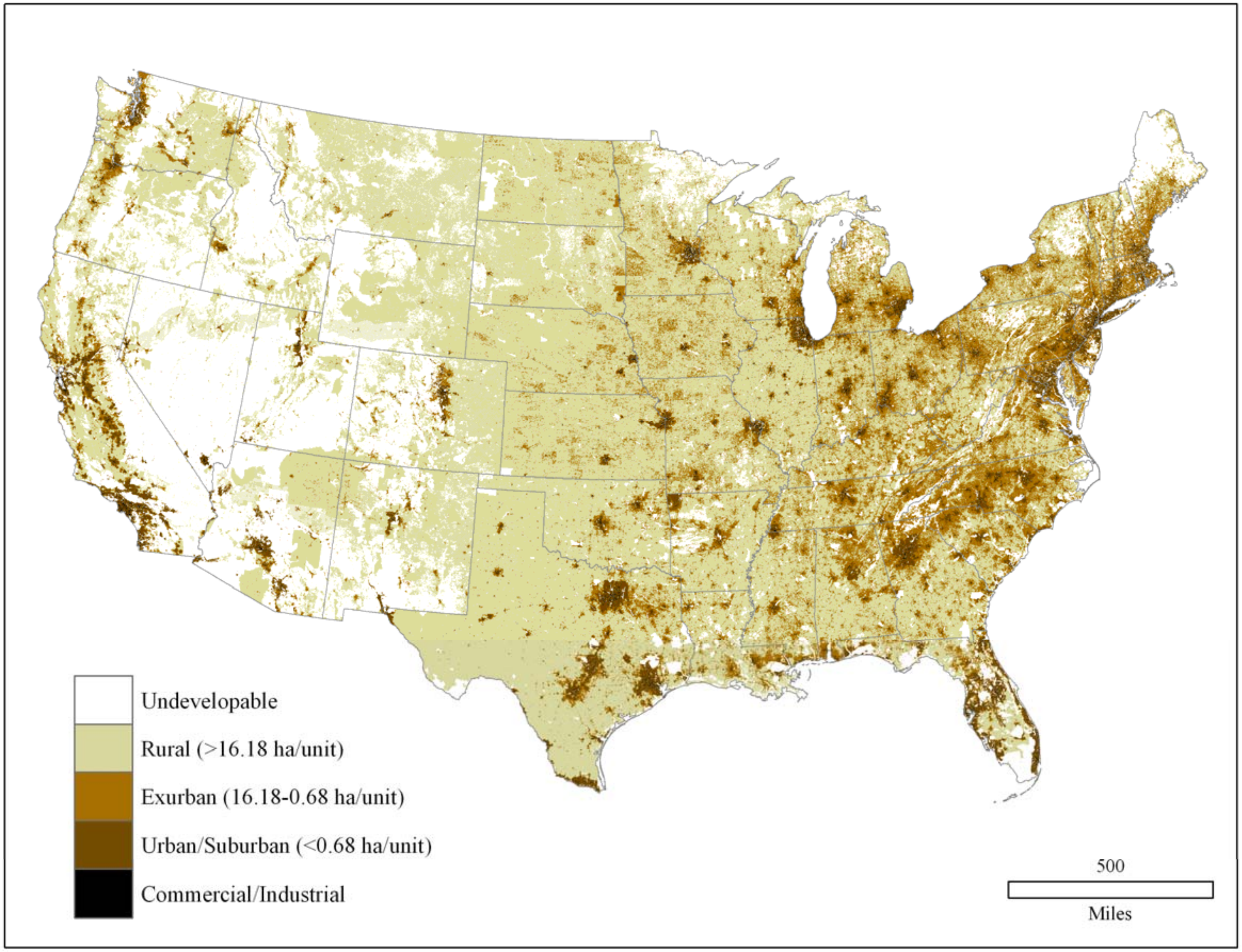

Figure A-9. A2 storyline, year 2100 housing density map.

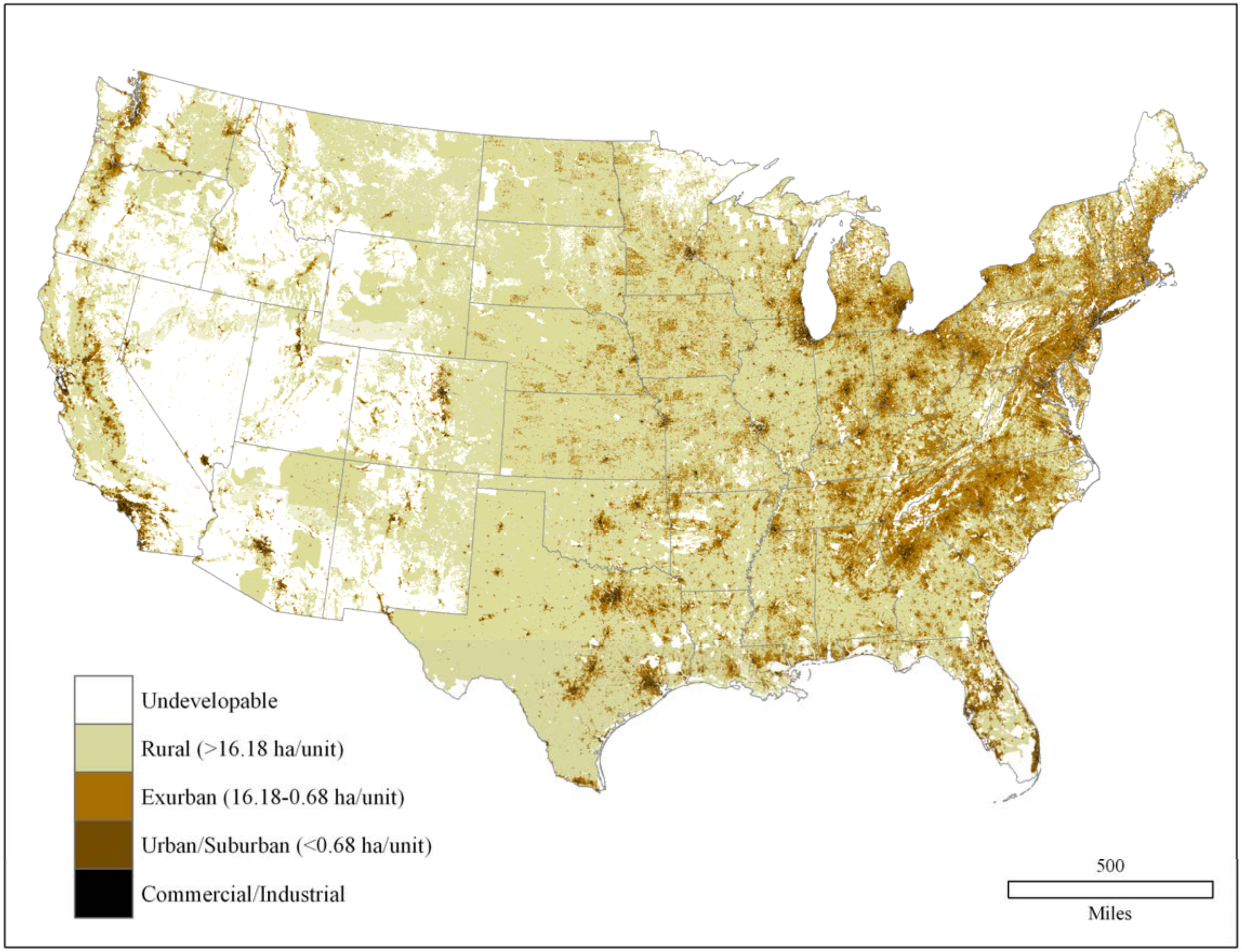

Figure A-10. B1 storyline, year 2010 housing density map.

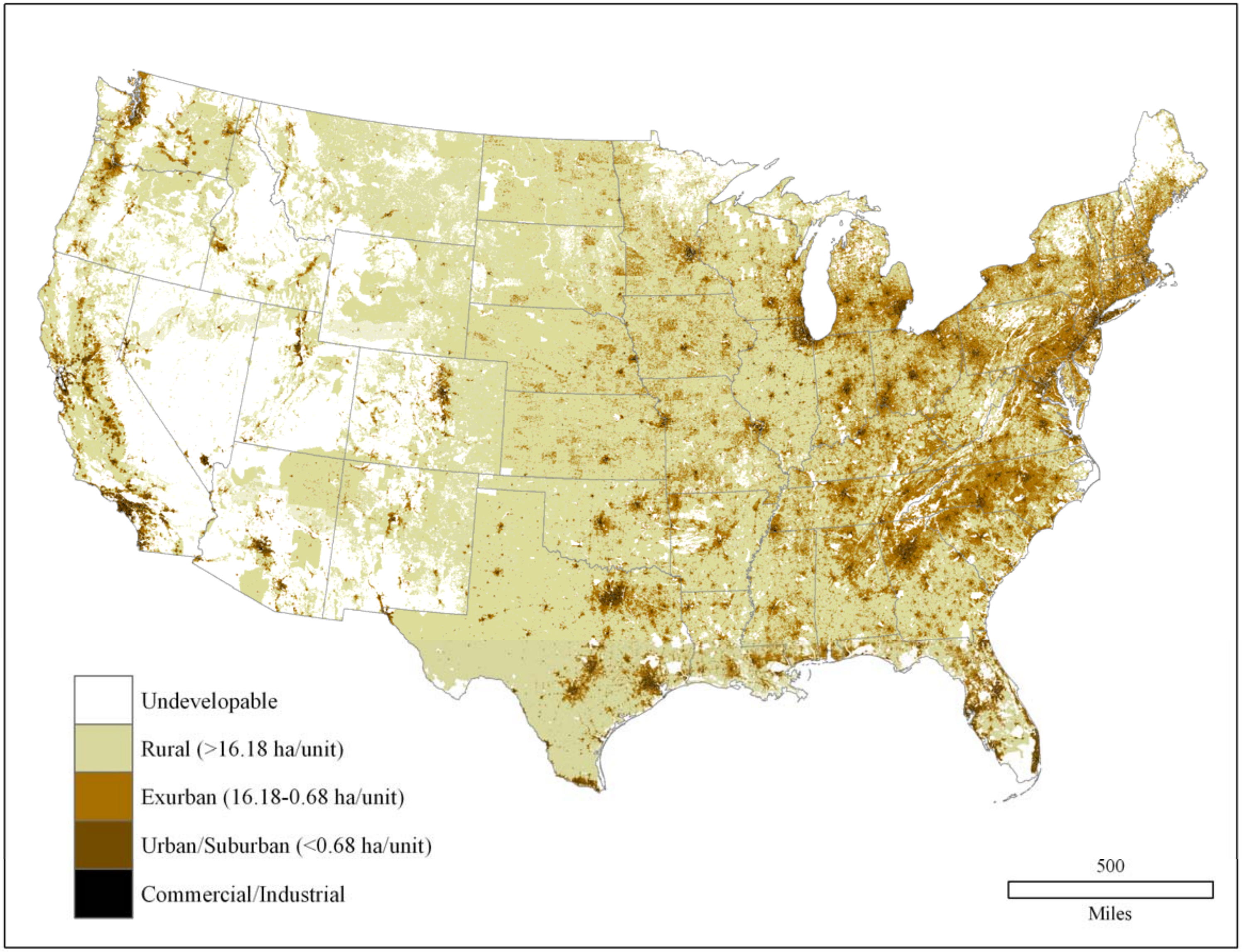

Figure A-11. B1 storyline, year 2050 housing density map.

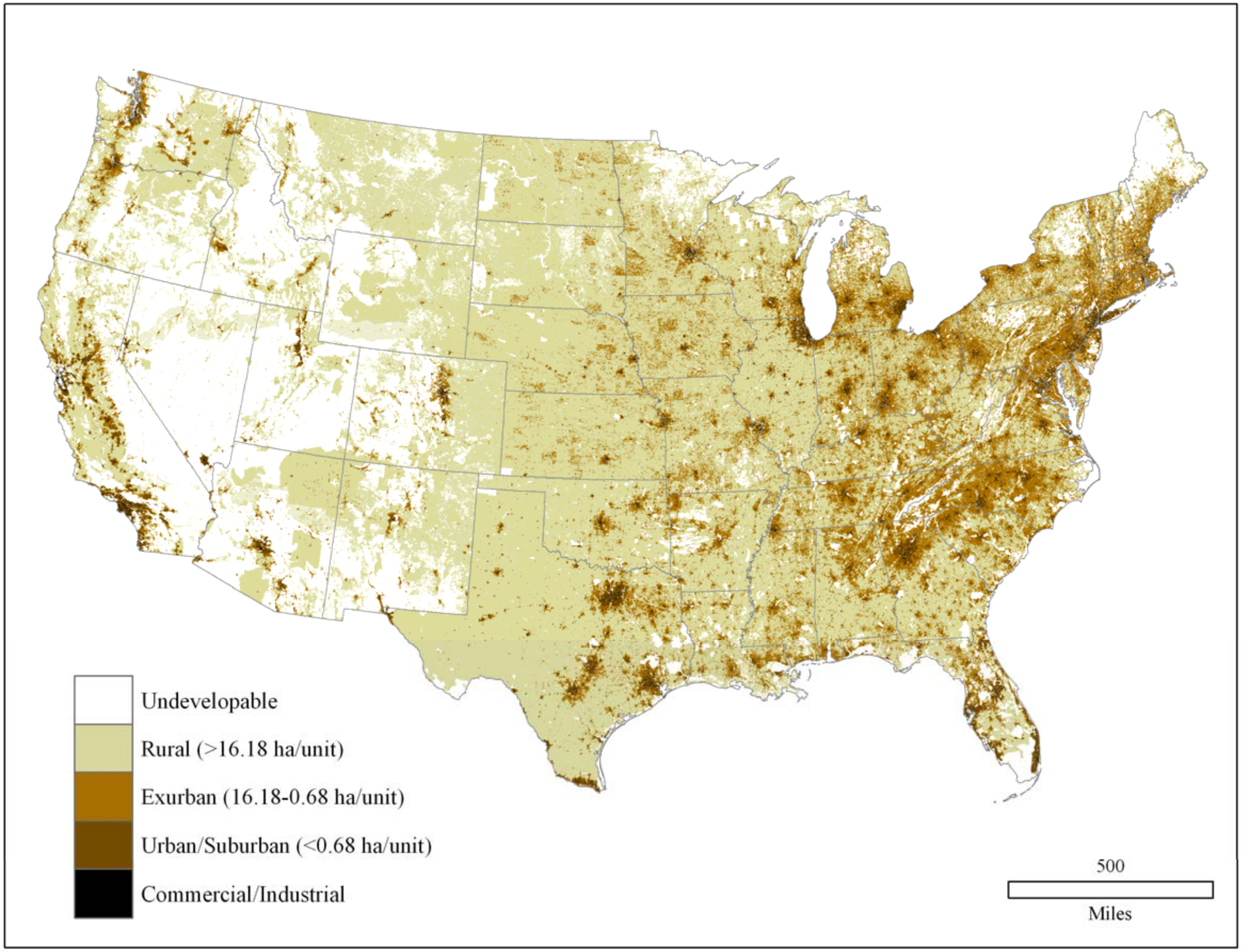

Figure A-12. B1 storyline, year 2100 housing density map.

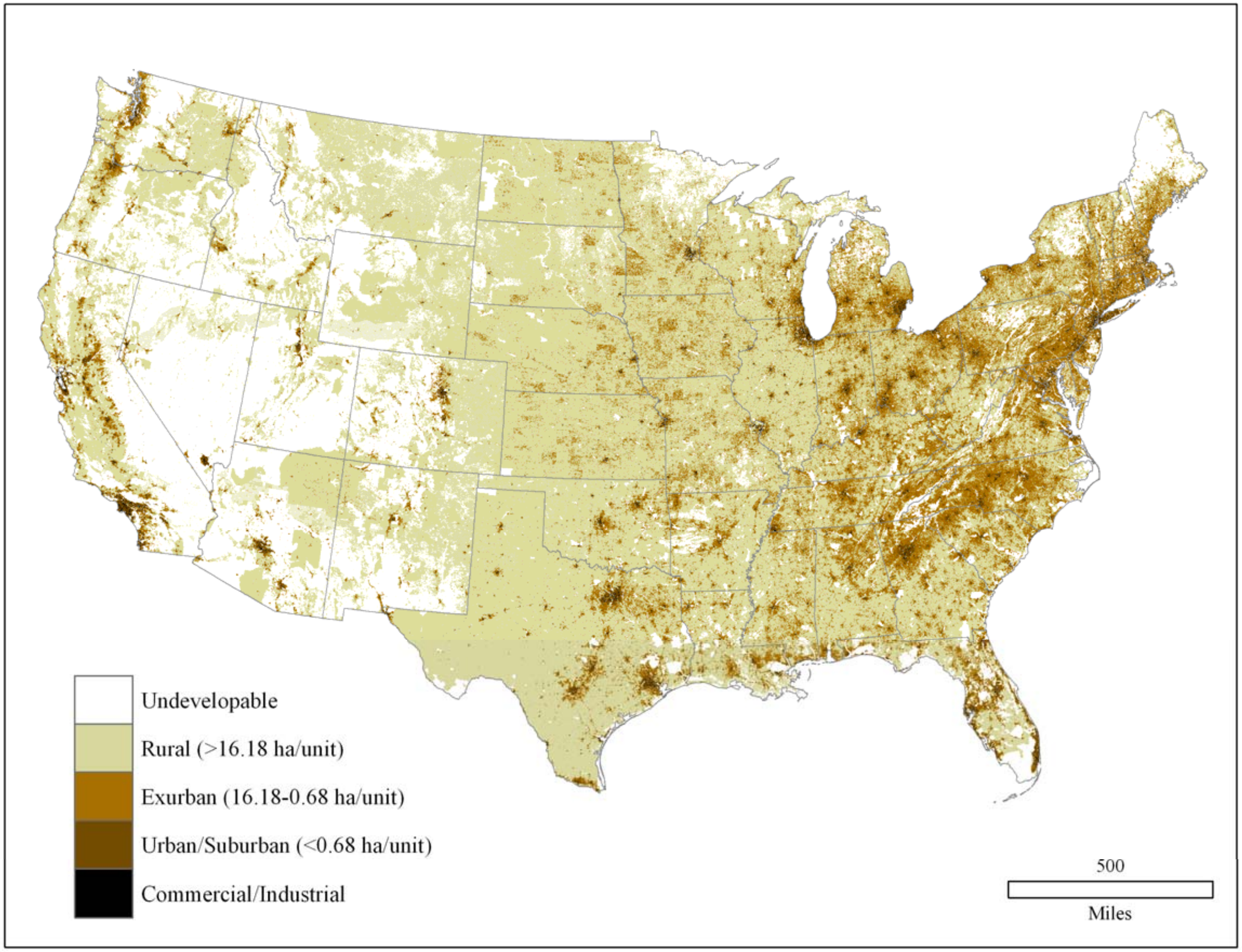

Figure A-13. B2 storyline, year 2010 housing density map.

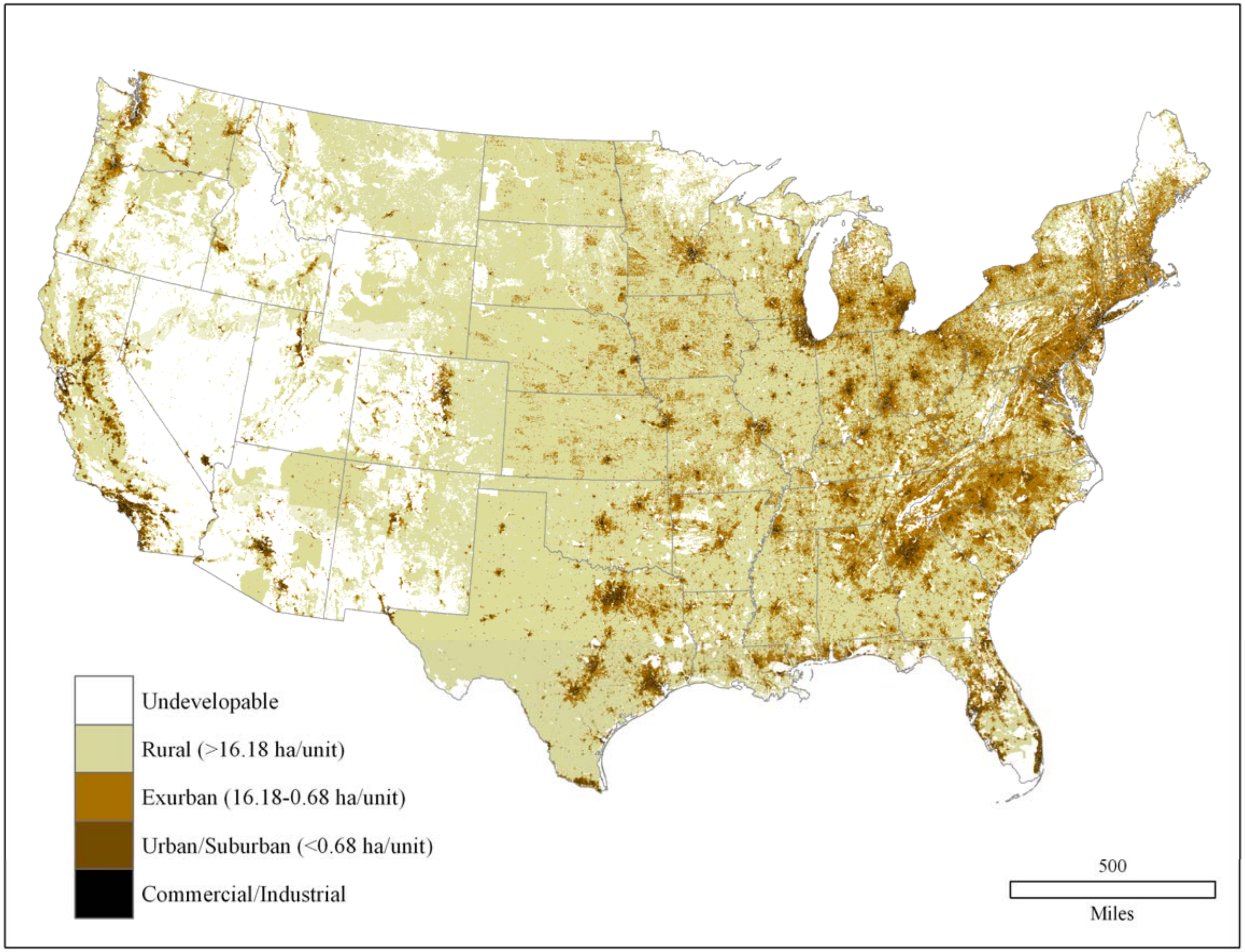

Figure A-14. B2 storyline, year 2050 housing density map.

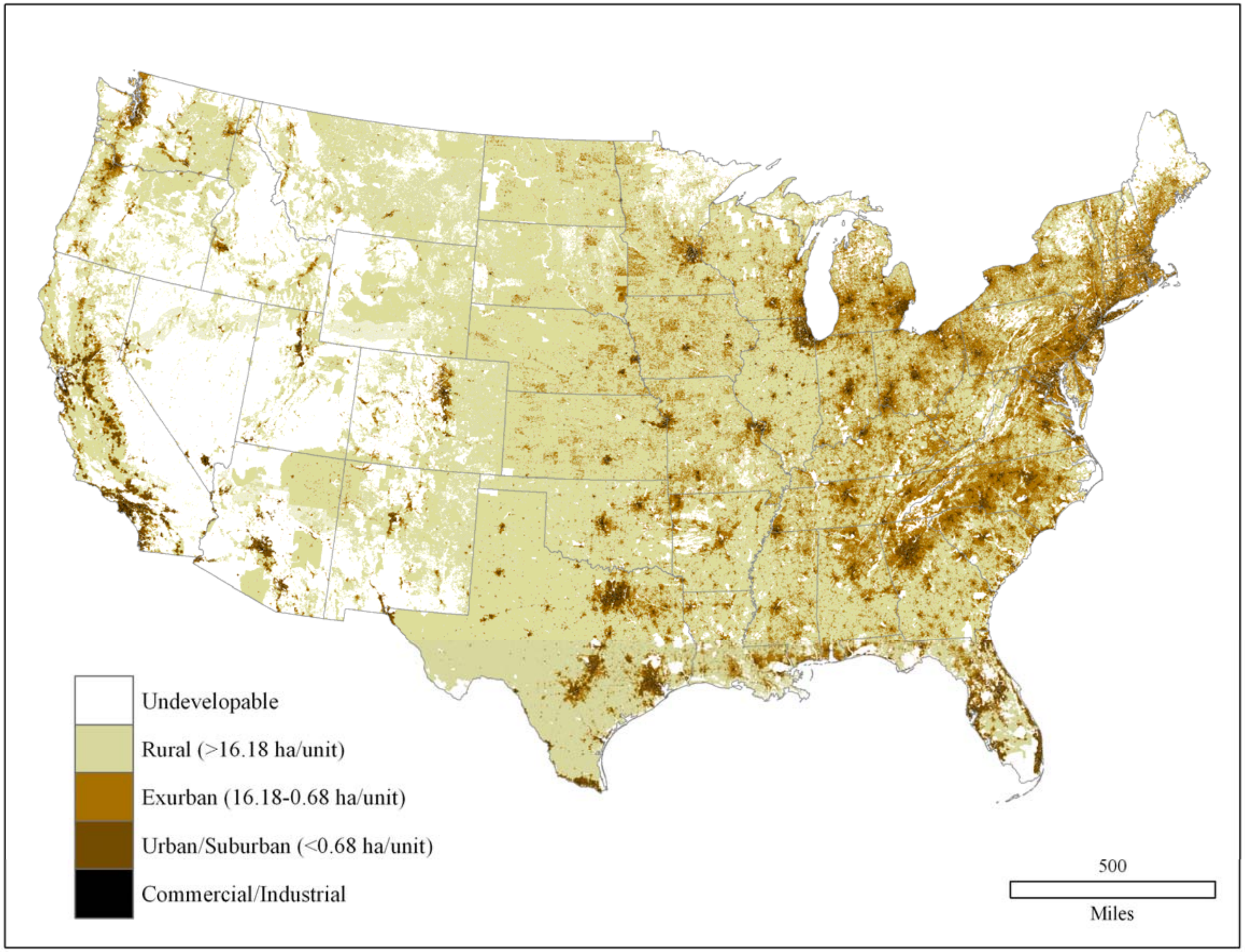

Figure A-15. B2 storyline, year 2100 housing density map.

APPENDIX B

DEMOGRAPHIC MODEL SENSITIVITY TESTING

In order to explore the demographic model's response to changes to its inputs and a variety of potential scenarios, we ran a series of sensitivity tests that paid particular attention to the behavior of the gravity model. These tests were used to improve the model, as well as to develop the downscaling approach to the Special Report on Emissions Scenarios (SRES). We also compared the outputs for several states to county-level projections produced by those states. While this project could not expect to match the level of detail and state-specific methodology produced for each state's estimates, this did provide us with a useful benchmark for comparison with the Integrated Climate and Land-Use Scenarios (ICLUS) outputs.

The first set of testing involved the fertility rate. International migration was held at medium, while domestic migration was set to zero, since we were interested only in looking at total national population at this time. The model was then run using the low, medium, and high fertility rate scenarios provided by the U.S. Census Bureau. Figure B-1 below compares the impact of the fertility rate scenarios on national population. Under these settings, low fertility predicts a mid-century peak in population followed by a small decline, with the medium and high scenarios result in steadily rising population.

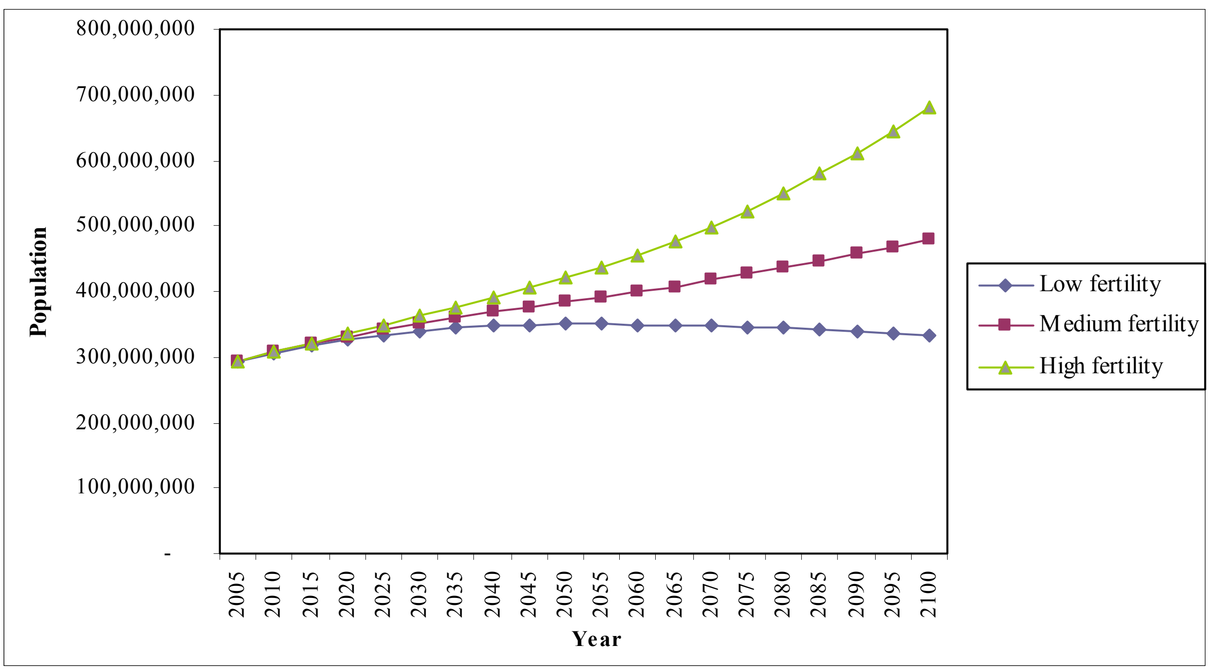

Figure B-1. Effect of fertility rate on national population.

Next, we considered the impact of international migration on total population. This time, fertility was held constant at medium and domestic migration was set to zero. The model was run using the low, medium, and high international migration scenarios provided by the U.S. Census Bureau. Figure B-2 below displays the results of these scenarios. All scenarios result in steadily rising national population, but since medium fertility produces a small steady increase in the population, this increase is at least partially due to fertility. Figure B-2, which presents a wider range of scenarios, shows that the combination of low fertility and low international migration presents the lowest possible population trajectory given our current inputs. In Figure B-3, the outputs for the base case and four SRES-compatible scenarios are presented, as well as the maximum and minimum scenarios (calculated using high fertility and high immigration, and low fertility and low immigration, respectively).

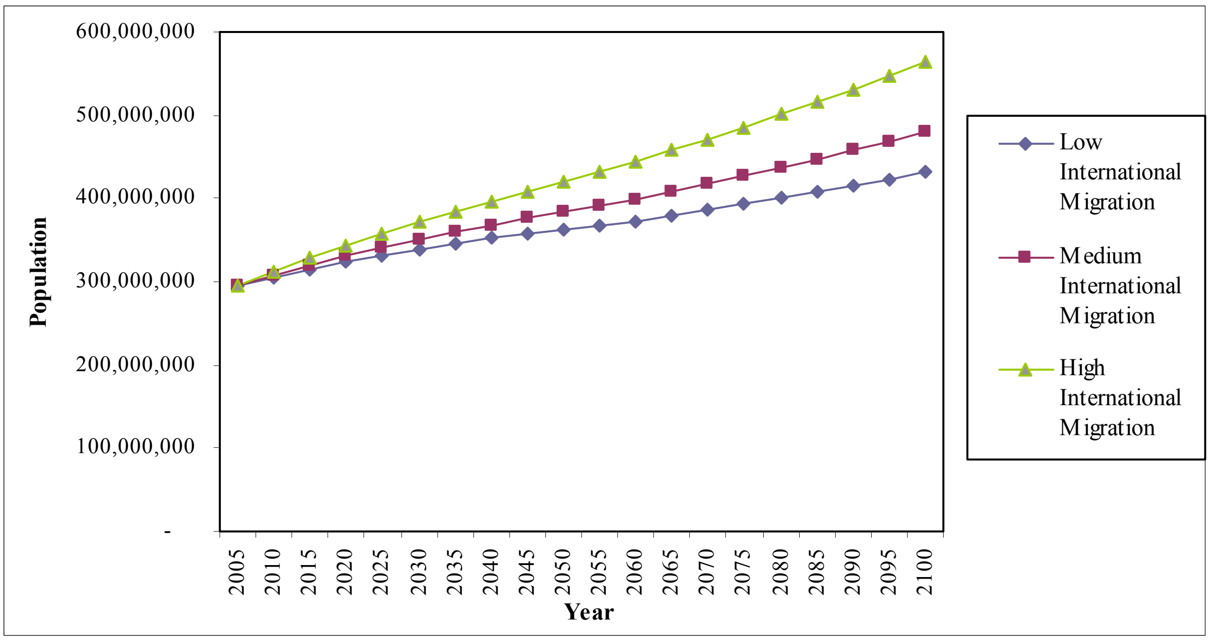

Figure B-2. Effect of international migration on national population.

We also ran tests with the mortality rate, even though the only available data set for mortality rates was the Census middle case, and we elected to use this set in all SRES storylines. To explore the effect of mortality on national population, we used the Census middle case to create additional sets of mortality rates by adjusting the Census medium scenario by +/−1%, +/−1.5%, and +/−2% per decade so that by 2100, the high cases were 10, 15, and 20% higher than middle and the low cases was 10, 15, and 20% lower than middle, respectively. Each was run with fertility and international migration set to medium and domestic migration set to zero. As Figure B-4 shows, the effect of even the strongest change was relatively small compared to changes in the other components of change.

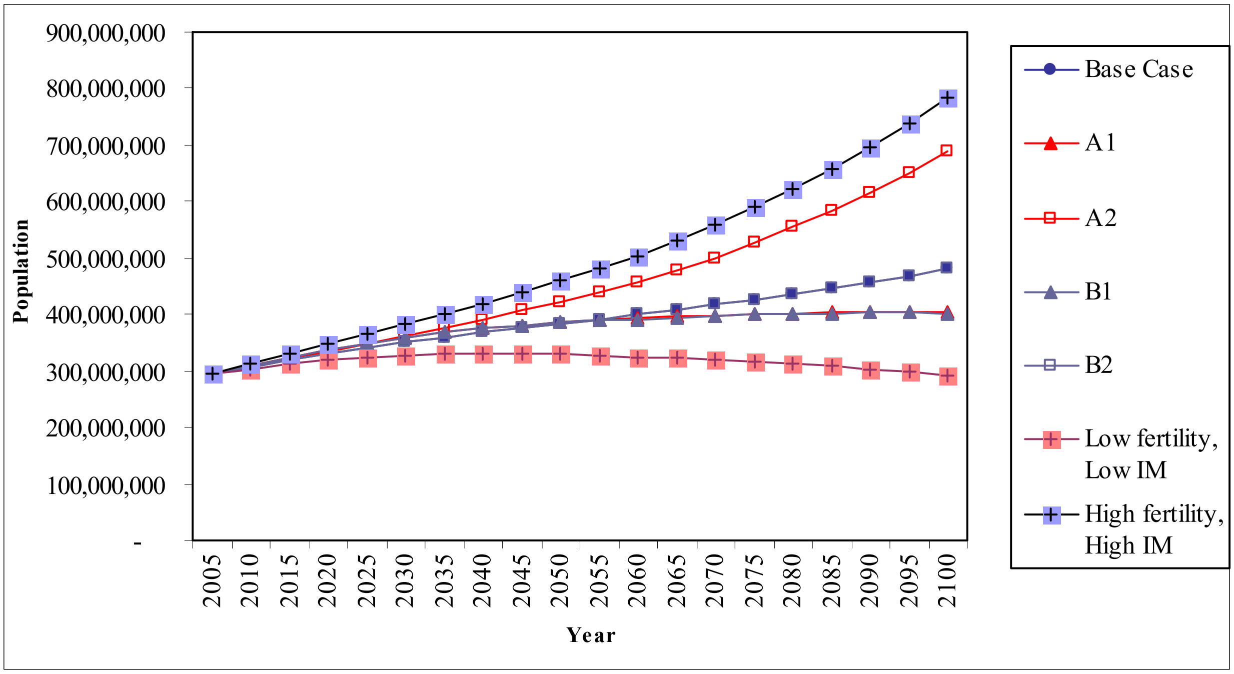

Figure B-3. Comparison of a broad range of scenarios.

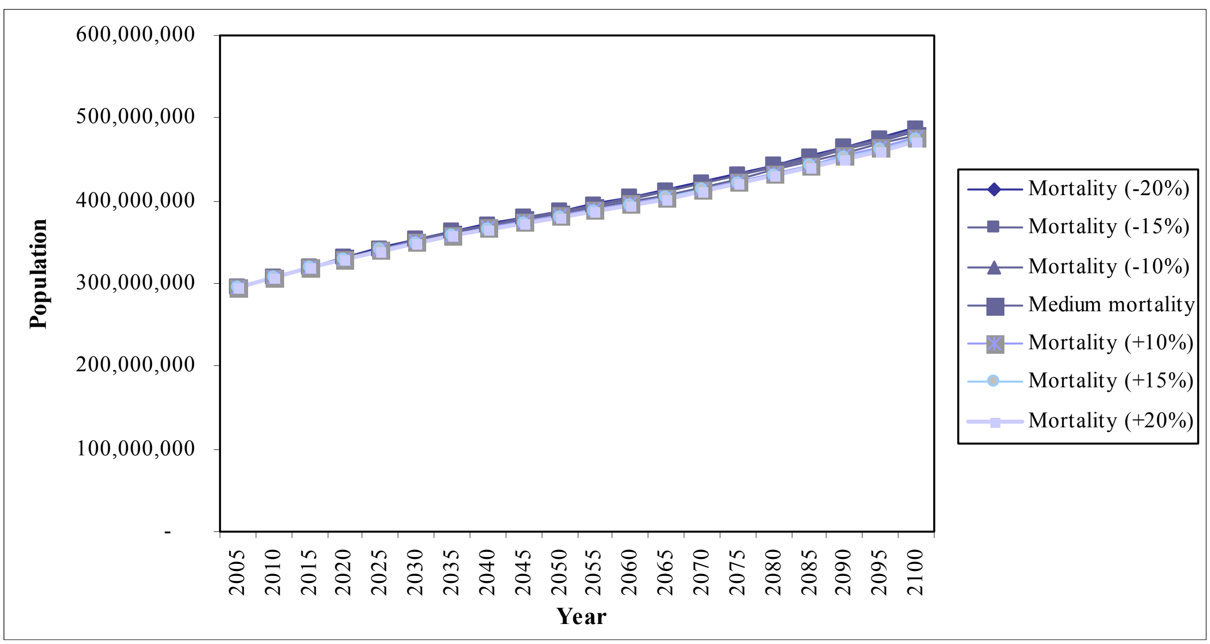

Figure B-4. Effect of mortality on national population.

Due to the complexity of the gravity model, there were many possible adjustments that could be made to change the magnitude of domestic migration. The simplest change involved the introduction of a scaling factor. Under this adjustment, the gravity model calculations would proceed as normal, but all calculated migrations would be scaled upward or downward. For example, if the normal model estimated 10 migrations from county A to county B, with a 50% scaling factor to cut migrations in half, only 5 migrations would occur. Other approaches could involve adjusting the model coefficients and/or y-intercept. These approaches would allow more fine tuning by increasing the attraction of large cities or increasing the friction of distance, for example.

In this analysis, we ran the gravity model with the following nine alternative scenarios:

1) Scaling all migrations by 50%. This has the practical effect of reducing migrations.

2) Scaling all migrations by 150%. This has the practical effect of increasing migrations.

3) Increasing the production population coefficient by 20%. Since the production population coefficient is positive, this has the practical effect of increasing migrations.

4) Decreasing the production population coefficient by 20%. Since the production population coefficient is positive, this has the practical effect of decreasing migrations.

5) Increasing the attraction population coefficient by 20%. Since the attraction population coefficient is positive, this has the practical effect of increasing migrations.

6) Decreasing the attraction population coefficient by 20%. Since the attraction population coefficient is positive, this has the practical effect of increasing migrations.

7) Increasing the distance coefficient by 20%. Since the distance coefficient is negative, this has the practical effect of reducing migrations.

8) Decreasing the distance coefficient by 20%. Since the distance coefficient is negative, this has the practical effect of increasing migrations.

9) Increasing the 1980−2000 growth rate coefficient by 20%. Since the growth rate coefficient is positive, this has the practical effect of increasing migrations. We did not analyze a similar decrease because adding this coefficient was deemed to be a model improvement and the areas of concern so far have been in the rapidly growing counties.

Because domestic migration can only be considered from a county perspective, we compared the outputs from five states with state projections and the revised base case projections from October. The five states—California, Colorado, Florida, Minnesota, and New Jersey—were selected for their geographic diversity and availability of suitable county-based projections. We compared the 2030 state projections (2025 for New Jersey) with the ICLUS base case projections and with outputs from the scenario tests.

We used the outputs of these tests to help refine the demographic projections. For example, early runs showed that urban counties were growing much faster in the ICLUS projections than anticipated by state projections. This led to changes in how we modeled the attraction of migrants to urban centers (see Section 3.5.4). We also found that counties currently identified by the states as fast-growing areas did not grow as quickly in the ICLUS model as they did in the state projections. Since the ICLUS model was designed to be a relatively simple national model, it was not possible to include some of the specialized local details that the states included in their projections. Therefore, divergences from the state projections were expected. This observation did lead us to include 1980−2000 growth rates as a term in the migration model. As a result, those fast-growing areas continued their relatively rapid growth rates in our projections.

APPENDIX C

STATISTICAL RELATIONSHIP BETWEEN HOUSING DENSITY AND IMPERVIOUS SURFACE COVER

C.1. BACKGROUND

The goal of this analysis was to develop a model to statistically relate housing density (HD) estimated by Spatially Explicit Regional Growth Model (SERGoM) to impervious surface (IS) cover. To do this, we examined how IS from the NLCD 2001 (MRLC, 2001) related to HD and ancillary variables including transportation (highways, secondary, local roads), and neighborhood density of urban/built-up land uses. Statistical relationships were developed by using regression tree models. This approach is useful because regression trees do not assume linear relationships in the independent variables, can handle a range of response variable types, are easy to construct and robust, and results are straightforward to interpret (De'ath and Fabricius, 2000). Regression trees have been increasingly used for broad-scale remote sensing and Geographic Information System (GIS) datasets (e.g., Friedl and Brodley, 1997; Homer et al., 2004). We also investigated other regression-based approaches that estimate imperviousness from land cover (e.g., NLCD 2001) and/or population data. We felt these were limited or not appropriate for our purposes primarily because they use population data rather than housing data, and because population is tied to primary residence, they underestimate the actual landscape effects of housing units. We also explored adding the locations of commercial/industrial land use from NLCD 2001 to the IS estimates for HD because they are likely to have very high IS levels as well. Note that we developed our IS model as a function of estimated housing densities from SERGoM to NLCD 2001 IS estimates. One alternative would be to use "real-world" estimates developed from either field data or aerial photo interpreted data, but these data simply do not exist (e.g., Elvidge et al., 2004 rely on 80 non-probability based sample points). Rather, we used ad hoc datasets on "real-world" IS as a way to validate our model only, not to generate the estimates of IS.

C.2. METHODS

We downloaded the Percent Urban Imperviousness (PUI) dataset from the Multi-Resolution Land Characteristics Consortium website.[8] The PUI dataset is produced using a Regression Tree that includes satellite imagery and roads (Homer et al., 2004). We aggregated the 30-m resolution to roughly 1-km^2 resolution (990-m) to compute the average PUI for each 1-km^2 cell. We then resampled the average PUI at 0.98-km^2 cell to 1 km^2 using bilinear interpolation. The 1-km^2 resolution is a commonly used resolution to develop national estimates of imperviousness (e.g., Elvidge et al., 2004).

We aggregated the HD estimates for the year 2000 from the SERGoM (Theobald, 2005) from 1-ha to 1-km^2 resolution. This provided us with the average HD for each 1 km^2. We

[8]http://www.mrlc.gov/multizone_download.php?zone=4; accessed 12 February 2007.

generated a sample of 200,000 random points (generated in a spatially-balanced way using the Reversed Recursive-Quadrant Randomized Raster algorithm; Theobald et al., 2007) from across the conterminous United States. We extracted the values of both the PUI and HD at each point, and used a Classification and Regression Tree model to develop a regression equation to develop a relationship between PUI and HD.

We generated the percent IS for current HD using the *tree* function in S-Plus (Insightful Corp, Seattle, Washington). See Brieman et al. (1984) for a review of categorical and regression tree methods. The resulting regression tree (Figure C-1) was then evaluated using a cross validation (*cv.tree* S-Plus function) technique to investigate if the tree over-fitted the data. As the number of terminal nodes increase, the overall deviance decreases (Figure C-2), indicating that the original tree is not over-fitting the data. If the deviance were to start increasing after some point within the cross validation analysis, then the tree would need to be pruned to a size that would minimize deviance. Because this was not the case, we decided to develop the percent IS with the original tree. The large tree size (58 nodes) can be explained best by the relatively poor non-parametric relationship between PUI and HD ($R^2 = 0.38$), meaning that there was not a simple, linear fit, as shown in Figure C-3.

The model resulted in a Residual Mean Deviance (which is the sum of the square differences between the actual values and the predicted values divided by the sample size) of 4.671. This is equivalent to the standard error in linear regression, which is the spread of error (in impervious units) for a given observation. The distribution of the residuals (error associated with the IS observations) has a minimum value of −62.35, a mean of −1.132e−14 (about 0), and a max value of 89.5. This distribution was not unexpected because there are areas where impervious values (independent variable) had values of 0 but have positive values of HD (dependent, response variable). Figure C-1 shows the decision backbone of the full regression tree with the length of a limb indicating deviance. The top ten nodes within the tree minimized deviance the most, with the remaining nodes making small adjustments for non-parametric small grain instances. Figure C-4 shows the top ten nodes within the full regression tree and the HD thresholds used to estimate percent impervious.

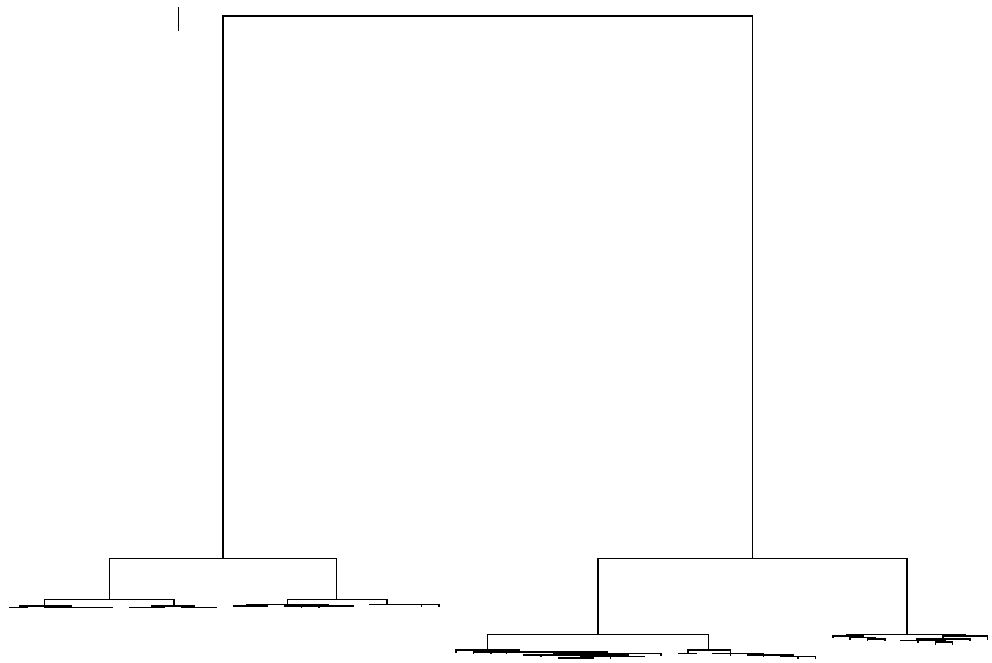

Figure C-1. Full regression tree backbone (58 terminal nodes) without labels.

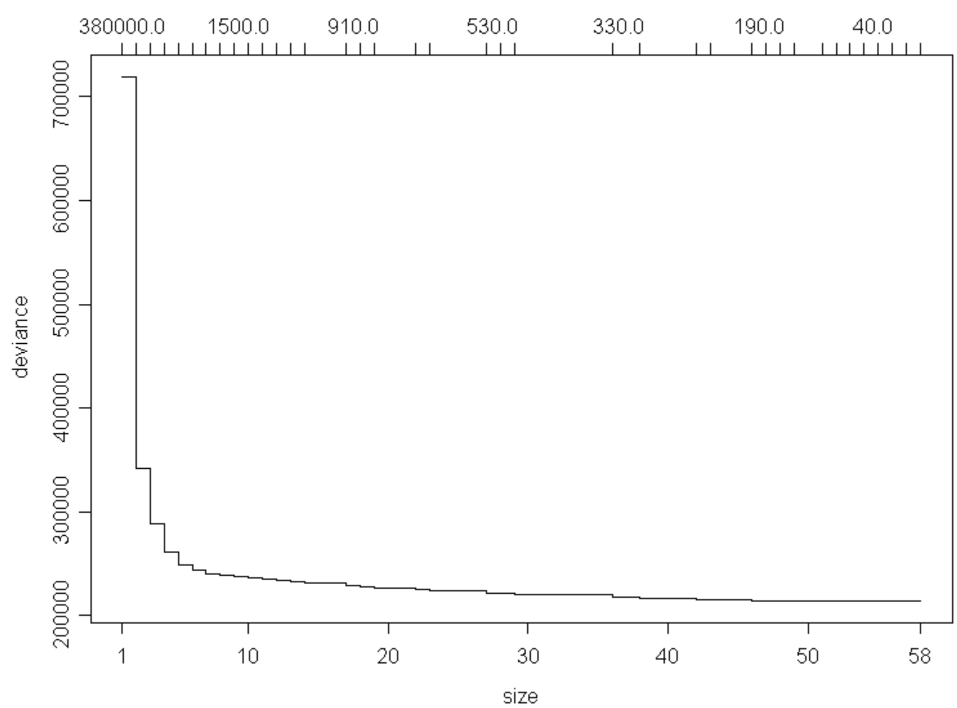

Figure C-2. Cross validation results for the full regression tree.

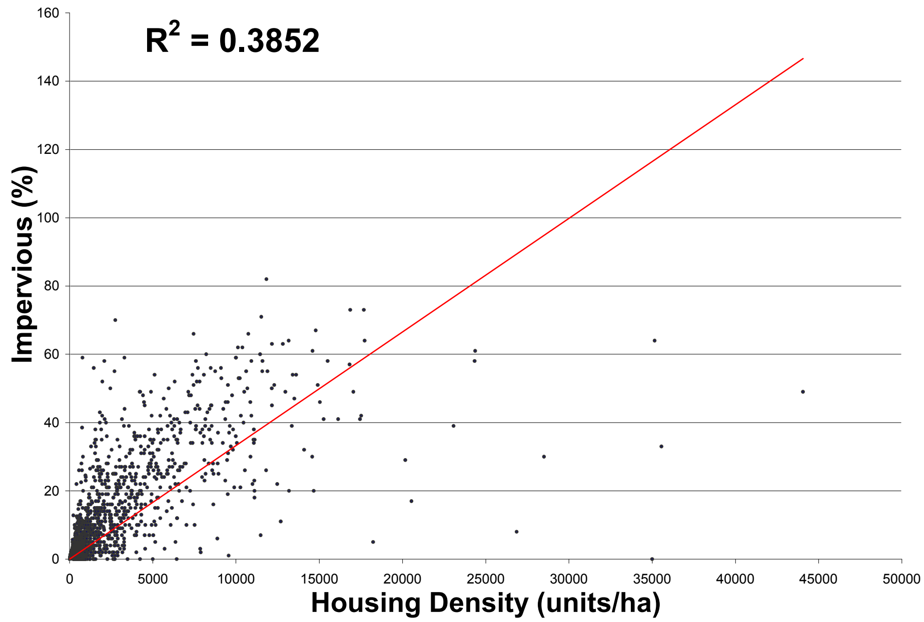

Figure C-3. The relationship between percent impervious and housing density.

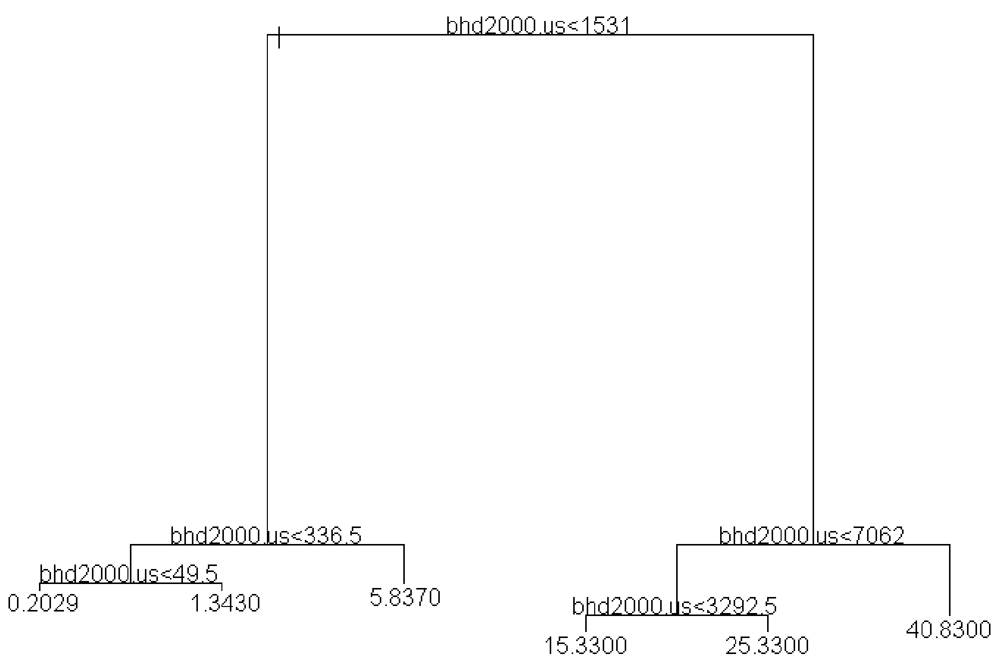

Figure C-4. Top ten terminal nodes within full regression tree with housing density labels and percent impervious estimates (terminal nodes).

We then converted the tree into a “con” (conditional) statement that can be input into ArcGIS as a Map Algebra statement. This statement then will then apply the regression model to generate a spatially-explicit map (note that HD is in units of housing units × 1000 per hectare; and “x” is HD for the cell being processed):

```
con(x < 1855, con(x < 410.5, con(x < 67.5, con(x < 1.5, 0.009477, con(x < 13.5, 0.358900, con(x <
49.5, con(x < 25.5, 0.529100, con(x < 41.5, con(x < 26.5, 0.706300, 0.633000), 0.592000)), 0.766400))),
con(x < 207.5, con(x < 138.5, con(x < 105.5, 0.950700, 1.191000), con(x < 158.5, 1.397000, 1.721000)),
con(x < 281.5, con(x < 220.5, 2.492000, 2.235000), con(x < 289.5, 3.624000, con(x < 321.5, 2.516000,
3.137000))))), con(x < 1087, con(x < 545.5, con(x < 417.5, 5.245000, con(x < 473.5, 3.968000, 4.651000)),
con(x < 647.5, con(x < 585.5, 7.035000, 5.772000), con(x < 918.5, con(x < 692.5, 8.319000, con(x <
900.5, con(x < 884.5, con(x < 859.5, con(x < 826.5, 6.770000, 8.867000), 5.696000), 9.420000),
5.504000)), 8.126000))), con(x < 1429.5, con(x < 1417.5, con(x < 1321.5, 9.220000, 11.580000),
5.473000), con(x < 1684.5, con(x < 1677.5, con(x < 1521.5, 13.830000, 11.730000), 31.220000),
11.010000)))), con(x < 6800.5, con(x < 3106.5, con(x < 2666.5, con(x < 2519, con(x < 2473.5, con(x <
1883.5, 16.710000, con(x < 1909.5, 11.010000, con(x < 2219.5, 15.140000, 13.710000))), 21.680000),
13.450000), con(x < 2990, 18.220000, 14.670000)), con(x < 4620.5, con(x < 3141.5, 31.230000, con(x <
3824.5, con(x < 3803, con(x < 3165.5, 10.520000, con(x < 3748, 20.880000, 26.350000)), 10.280000),
23.640000)), con(x < 4639.5, 41.870000, con(x < 6031.5, con(x < 4796, 29.470000, con(x < 5193.5,
24.020000, 26.770000)), 29.800000)))), con(x < 13884.5, con(x < 9541, con(x < 9405, con(x < 7383.5,
32.750000, con(x < 7463, 44.370000, con(x < 8153, 34.070000, 36.610000))), 25.960000), con(x < 13713,
con(x < 9619, 54.960000, con(x < 10690.5, 37.480000, 41.710000)), 24.620000)), con(x < 32918, con(x <
19997.5, 51.850000, 48.040000), 68.890000))))))
```

C.3. RESULTS AND DISCUSSION

Using 2000 SERGoM v3 HD, we estimated 80,094 km^2 of IS (Figure C-5). Our estimated extent of IS (80,094 km^2) is fairly similar to other nationwide estimates. The NLCD 2001 Urban Imperviousness layer estimated an IS of 95,746 km^2. The impervious correlation coefficient against a ground-truth dataset was 0.83, 0.89, and 0.91 (Homer et al., 2004). Note that we were not able to develop “bracket” or “bookend” models that incorporate under- and over-estimations based on deviations around published error statistics because they were not provided for the NLCD 2001 IS cover. Also, the precision of estimated PUI below 20% is believed to diminish so that low % PUI estimates from PUI were difficult to obtain (Homer et al., 2004).

Elvidge et al. (2004) estimated an IS area of 113,260 km^2 (they actually report 112,610; +/−12,725 km^2) based on nighttime lights radiance, road density, and NLCD 1992 urban land cover classes. Thus, we were within 12% of NLCD 2001. Because SERGoM has NODATA values on public (non-developable) lands, there were some cases where IS occurs on

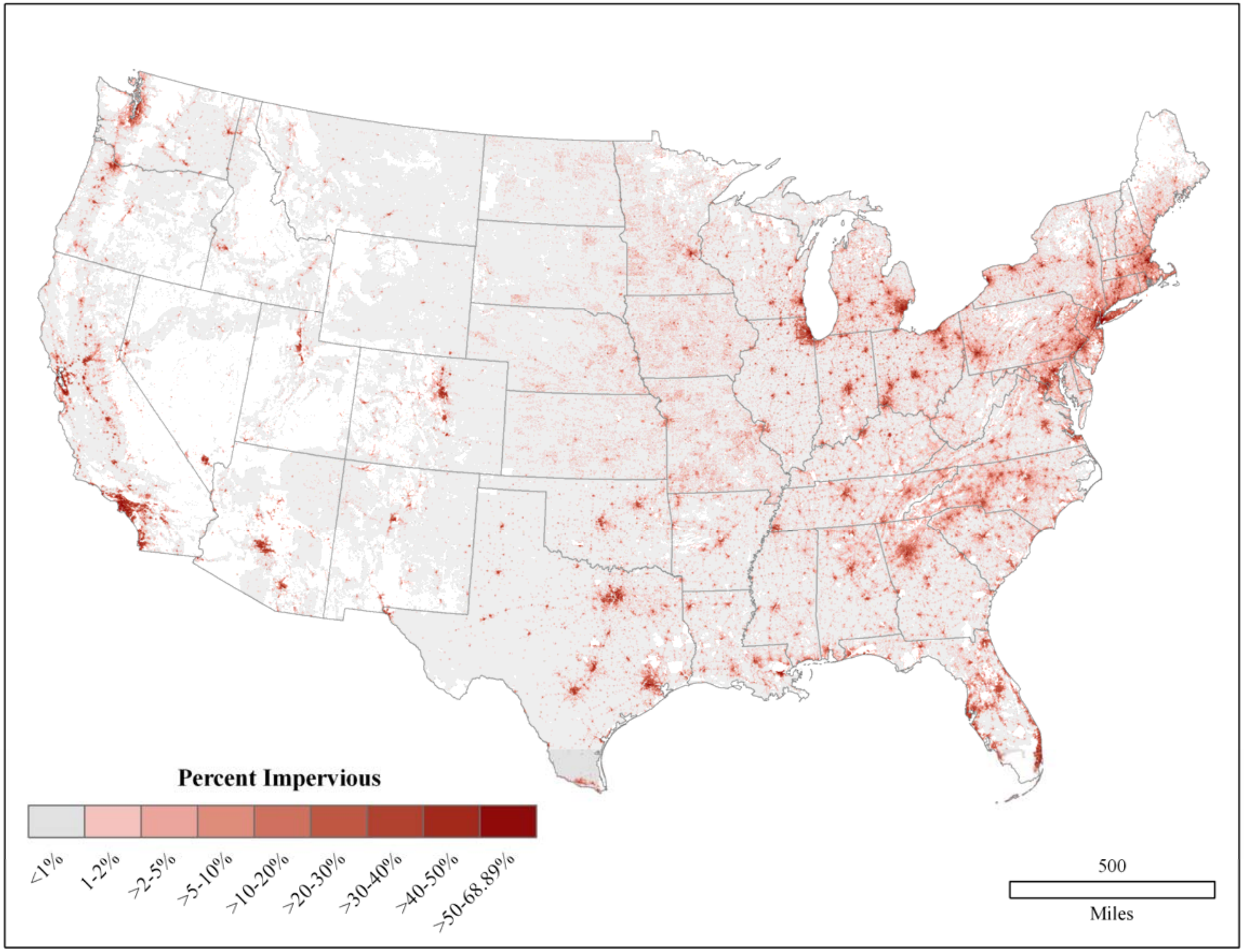

Figure C-5. Estimated national impervious surface, 2000.

non-developable lands, such as military bases, airports, developed portions (visitor centers) of national parks, interstates, etc. When we filled in areas of the SERGoM-based IS that had no HD (mostly public lands) with NLCD 2001 urban imperviousness, the total IS area increased slightly to 83,846 km^2. Thus, we recommend using this combination of datasets to better represent total IS that gets at roads and commercial/industrial areas as well.

We also conducted an additional validation step by developing a simple linear regression of the SERGoM-based IS against 80 data points generated from high-resolution aerial photography of 1-km^2 "chips" and used to generate the Elvidge et al. (2004) product. The resulting R^2 was 0.694. This result was better than expected, as the 80 data points were not randomly selected, rather purposively targeted to capture a gradient of urbanization, and as a result these points were selected to pick up much of the commercial and industrial land cover types.

We also generated a difference map to compare the NLCD 2001-derived estimates against the SERGoM estimates (Figure C-6 through Figure C-8). In general, NLCD 2001 estimated higher imperviousness in urban areas (shown in red), and under-represented imperviousness in lower-density, suburban/exurban land-use areas (shown in blue). Our estimates of imperviousness are likely underestimated in urban areas because they do not include commercial and industrial land uses. Because it is difficult to identify low HD land uses beyond the suburban fringe using NLCD 2001 data, it is likely that NLCD 2001 PUI slightly underestimates IS in exurban and rural areas.

Finally, in Figure C-9 we show the results of using the SERGoM HD projections for 2030—base case and the estimated IS as a function of HD. We will need to consider how to incorporate commercial/industrial contributions to IS estimates for future projections. For now, we recommend reporting just HD-based IS, realizing that it is a conservative estimate, and that future efforts should better represent urban/industrial land-use growth as a function of population/HD growth.

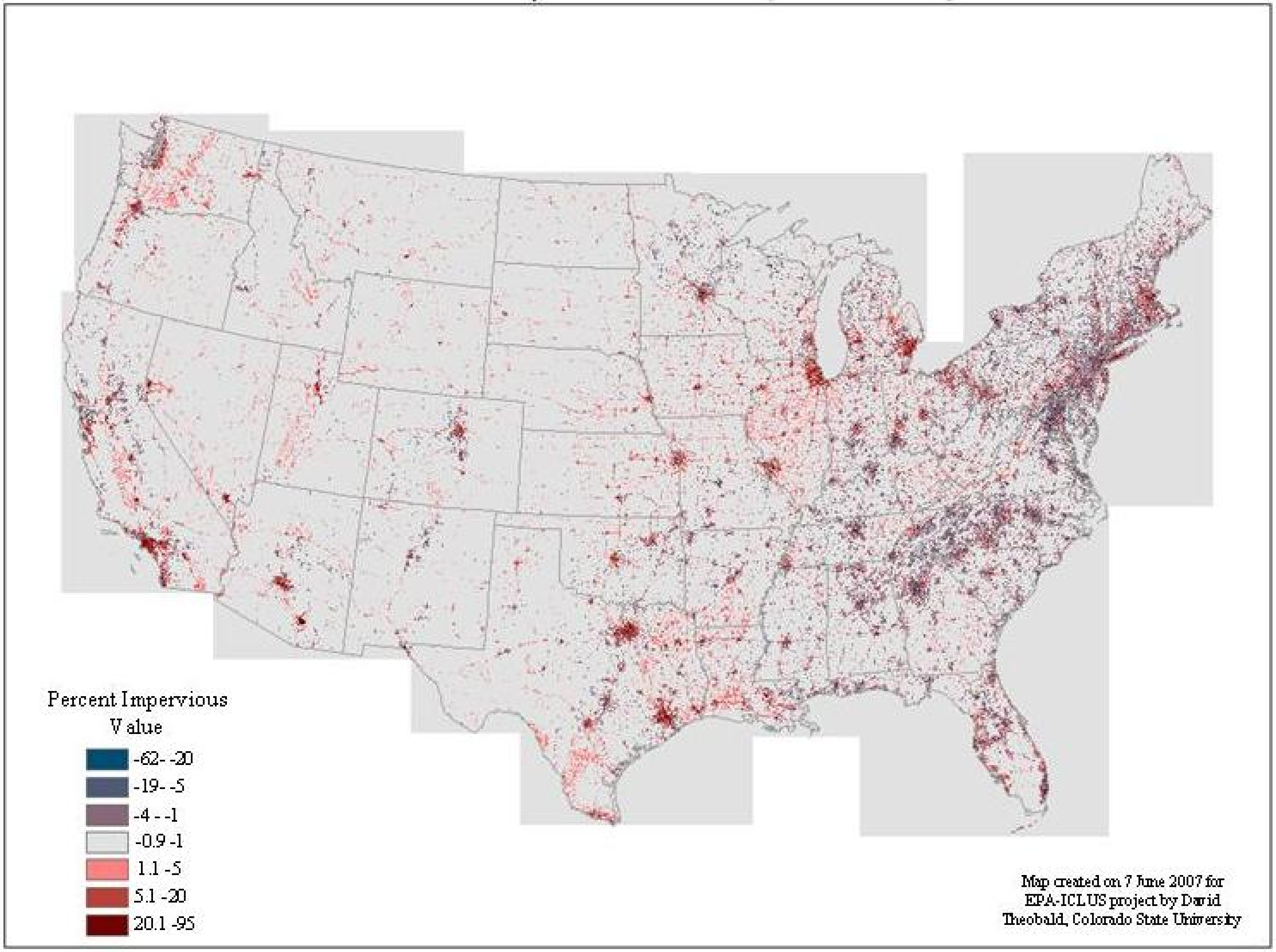

Figure C-6. Difference in impervious surface, United States.

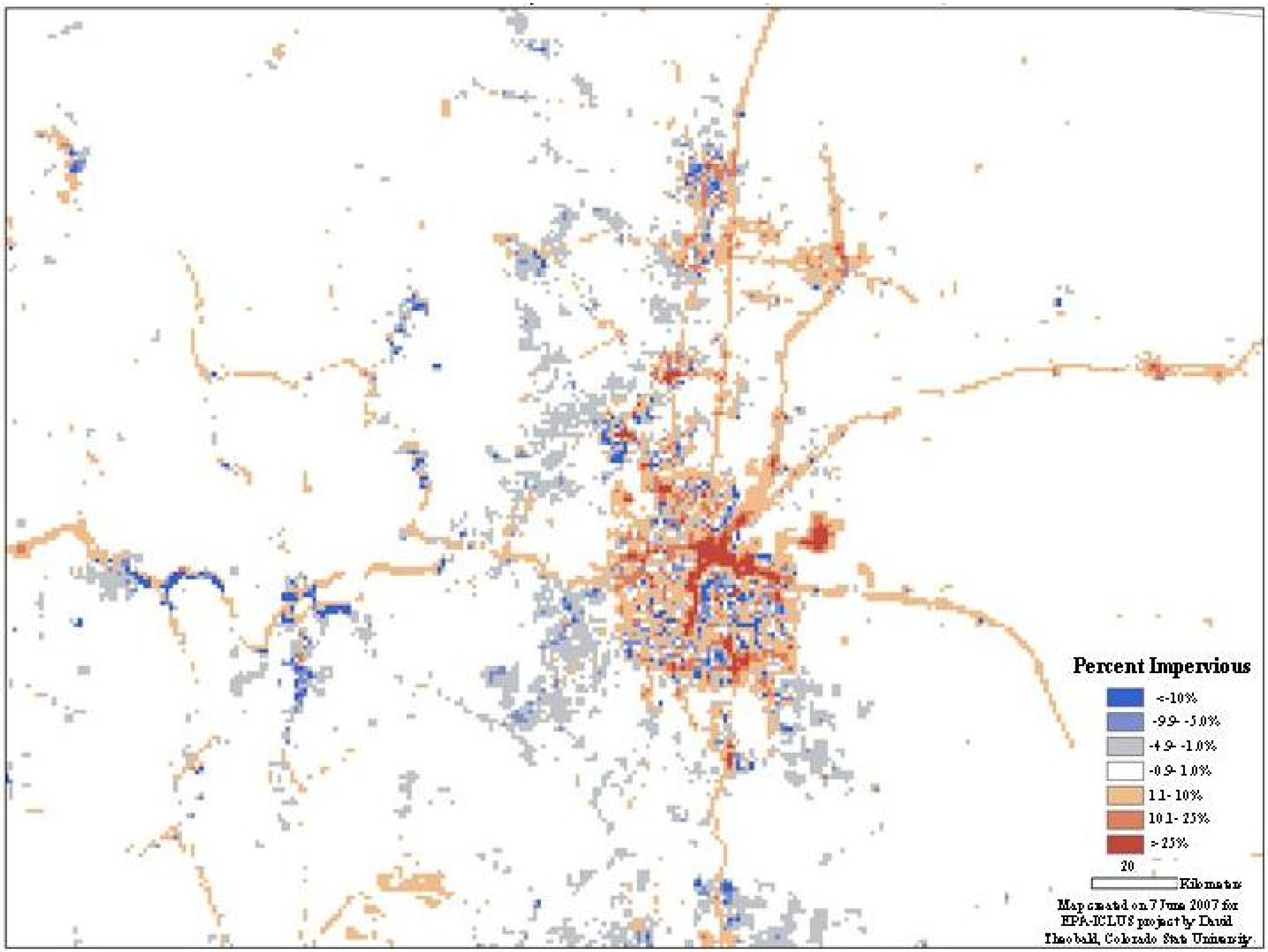

Figure C-7. Difference in impervious surface, Colorado.

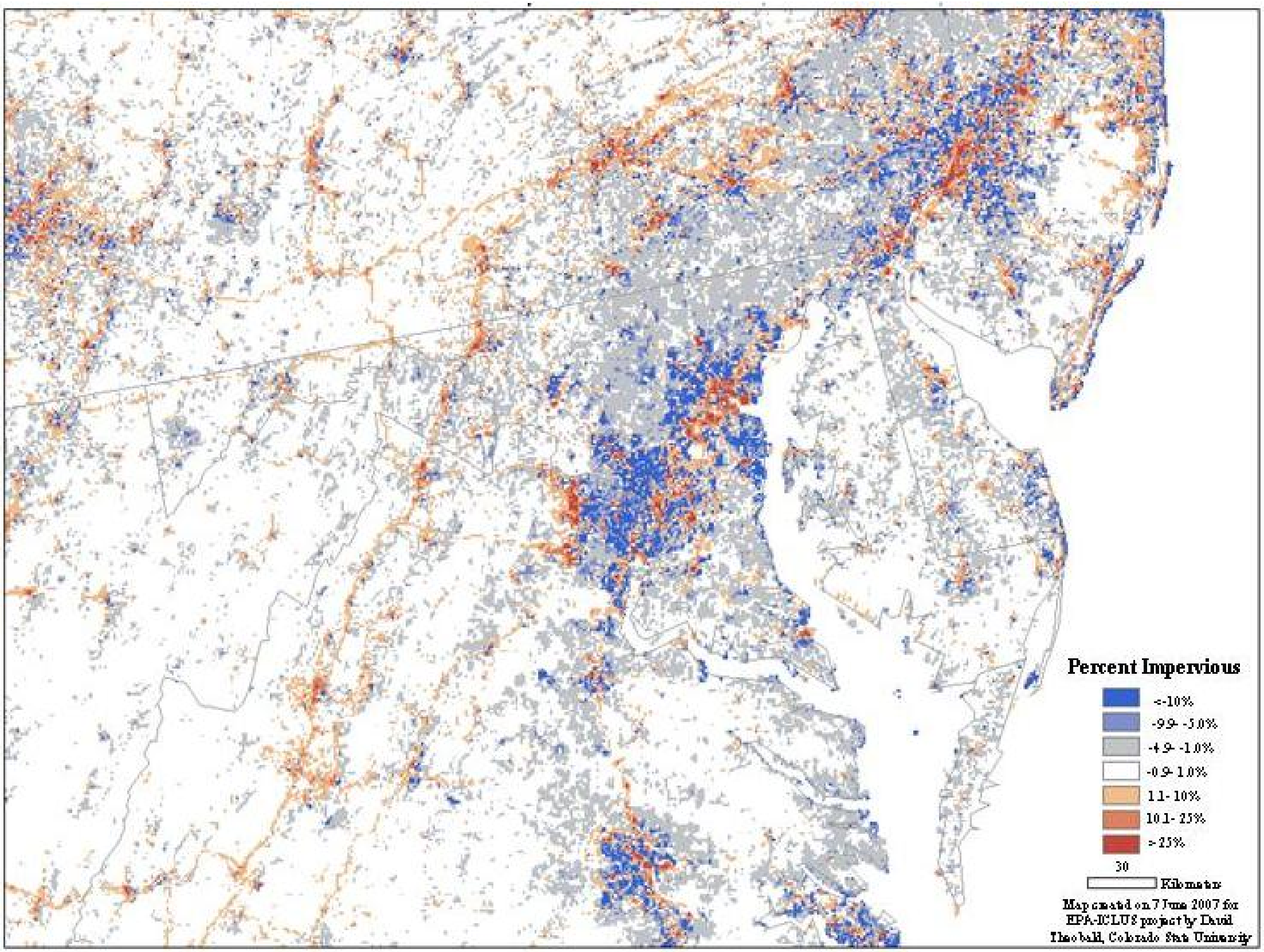

Figure C-8. Difference in impervious surface, Mid-Atlantic region.

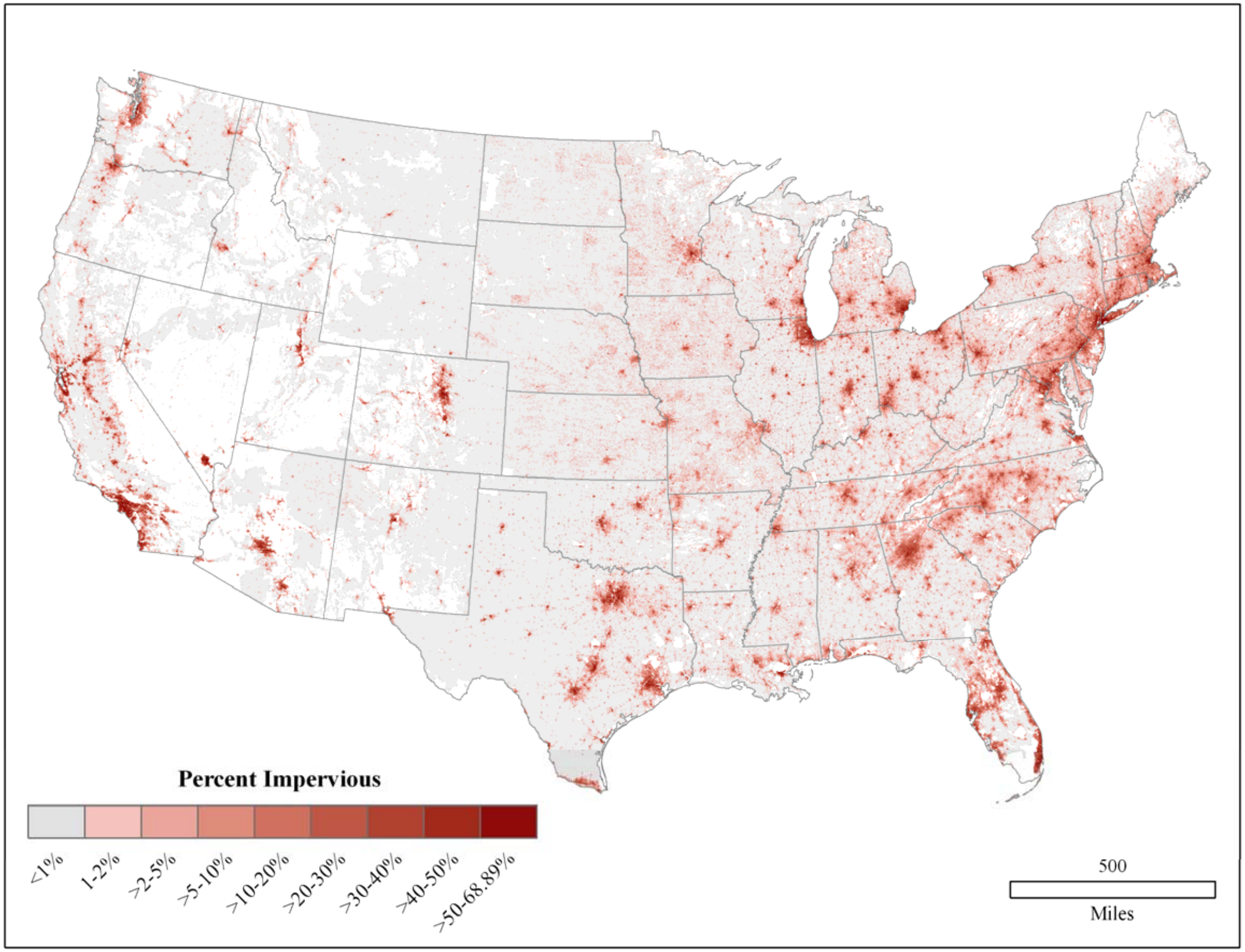

Figure C-9. Estimated impervious surface, base case 2030.

APPENDIX D

REGIONAL POPULATION GROWTH RATES AND PROJECTIONS BASED ON EPA REGIONS

Table D-1 below provides a list of the U.S. Environmental Protection Agency (EPA) regions that were used for the analysis of regional differences in population growth. Table D-2 provides the projected populations for each of the regions and each of the modeled storylines for 2005, 2030, 2060, and 2090.

Table D-1. EPA regions

Region	Region description	States
Region 1	Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Region 2	Mid-Atlantic	New Jersey, New York[a]
Region 3	Mid-east	Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, West Virginia
Region 4	Southeast	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
Region 5	Mid-west	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Region 6	Southwest	Arkansas, Louisiana, New Mexico, Oklahoma, Texas,
Region 7	Cornbelt	Iowa, Kansas, Missouri, Nebraska
Region 8	Mountain-west	Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
Region 9	Pacific-south	Arizona, California, Nevada[b]
Region 10	Northwest	Idaho, Oregon, Washington[c]

[a]Puerto Rico and the U.S. Virgin Islands were not included in Region 2 for this analysis.
[b]Hawaii, American Samoa, Guam, Northern Mariana Islands, and Trust Territories were not included in Region 9 for this analysis.
[c]Alaska was not included in Region 10 for this analysis.

Table D-2. Projected population and growth rate by scenario and EPA region

EPA region	Population				Growth rate		
	2005	2030	2060	2090	2005–2030	2030–2060	2060–2090
Base case							
1	14,255,073	15,519,657	15,902,916	16,887,110	9%	2%	6%
2	28,018,871	35,109,989	41,602,295	49,258,590	25%	18%	18%
3	28,797,147	31,978,301	34,424,923	38,637,662	11%	8%	12%
4	57,405,312	70,474,771	82,641,728	97,900,593	23%	17%	18%
5	51,257,348	55,592,269	57,423,505	60,486,700	8%	3%	5%
6	35,680,974	45,708,352	55,363,006	65,879,226	28%	21%	19%
7	13,269,562	14,224,959	14,504,537	15,010,238	7%	2%	3%
8	10,006,652	12,930,609	15,592,590	18,402,596	29%	21%	18%
9	44,519,456	56,426,295	67,555,641	79,977,083	27%	20%	18%
10	11,360,137	13,024,245	14,030,327	15,397,072	15%	8%	10%
Storyline A1							
1	14,255,073	15,141,889	14,141,988	12,659,207	6%	−7%	−10%
2	28,018,871	39,117,688	46,902,172	51,792,819	40%	20%	10%
3	28,797,147	32,682,229	34,693,103	36,115,978	13%	6%	4%
4	57,405,312	73,800,430	85,111,046	92,041,933	29%	15%	8%
5	51,257,348	54,146,181	51,539,278	47,200,686	6%	−5%	−8%
6	35,680,974	47,647,066	55,559,004	59,071,696	34%	17%	6%
7	13,269,562	13,780,804	12,800,038	11,445,472	4%	−7%	−11%
8	10,006,652	13,473,510	15,807,016	16,909,652	35%	17%	7%
9	44,519,456	56,021,194	61,616,885	62,219,659	26%	10%	1%
10	11,360,137	13,654,349	14,511,247	14,785,715	26%	10%	1%

Table D-2 (continued)

EPA region	Population				Growth rate		
	2005	2030	2060	2090	2005–2030	2030–2060	2060–2090
Storyline A2							
1	14,255,073	15,987,287	18,051,181	22,566,774	12%	13%	25%
2	28,018,871	35,917,932	45,986,762	61,228,288	28%	28%	33%
3	28,797,147	32,999,574	40,088,688	54,511,203	15%	21%	36%
4	57,405,312	74,318,414	97,906,372	138,195,518	29%	32%	41%
5	51,257,348	56,019,601	63,041,628	77,561,440	9%	13%	23%
6	35,680,974	47,833,571	64,361,337	90,250,013	34%	35%	40%
7	13,269,562	14,369,471	15,919,641	19,525,569	8%	11%	23%
8	10,006,652	13,630,155	18,236,722	25,157,767	36%	34%	38%
9	44,519,456	58,911,077	78,044,717	108,283,362	32%	32%	39%
10	11,360,137	13,137,594	15,229,105	19,198,490	16%	16%	26%
Storyline B1							
1	14,255,073	15,141,889	14,141,988	12,659,207	6%	−7%	−10%
2	28,018,871	39,117,688	46,902,172	51,792,819	40%	20%	10%
3	28,797,147	32,682,229	34,693,103	36,115,978	13%	6%	4%
4	57,405,312	73,800,430	85,111,046	92,041,933	29%	15%	8%
5	51,257,348	54,146,181	51,539,278	47,200,686	6%	−5%	−8%
6	35,680,974	47,647,066	55,559,004	59,071,696	34%	17%	6%
7	13,269,562	13,780,804	12,800,038	11,445,472	4%	−7%	−11%
8	10,006,652	13,473,510	15,807,016	16,909,652	35%	17%	7%
9	44,519,456	56,021,194	61,616,885	62,219,659	26%	10%	1%
10	11,360,137	13,654,349	14,511,247	14,785,715	20%	6%	2%

Table D-2 (continued)

EPA region	Population				Growth rate		
	2005	2030	2060	2090	2005–2030	2030–2060	2060–2090
Storyline B2							
1	14,255,073	15,543,670	15,997,050	17,077,923	9%	3%	7%
2	28,018,871	35,422,606	42,541,498	51,216,096	26%	20%	20%
3	28,797,147	32,086,103	34,216,490	37,647,265	11%	7%	10%
4	57,405,312	69,018,096	79,590,169	92,831,550	20%	15%	17%
5	51,257,348	57,089,003	60,336,151	64,684,730	11%	6%	7%
6	35,680,974	45,225,069	54,516,450	64,720,598	27%	21%	19%
7	13,269,562	14,552,312	15,227,145	16,041,584	10%	5%	5%
8	10,006,652	12,703,387	15,085,697	17,639,212	27%	19%	17%
9	44,519,456	55,960,255	66,985,991	79,576,422	26%	20%	19%
10	11,360,137	13,326,862	14,723,525	16,493,287	17%	10%	12%

APPENDIX E

COMPONENT AND COHORT MODEL DATA

The following tables provide summary values for the entire population; actual rates used in the model vary by age, sex, and race/ethnicity.

Table E-1. Fertility rates (births per 1000 women)

Year	Low	Mid	High
1999	2,036	2,048	2,059
2000	2,032	2,055	2,080
2001	2,026	2,063	2,100
2002	2,021	2,070	2,120
2003	2,015	2,077	2,140
2004	2,009	2,083	2,161
2005	2,002	2,090	2,181
2006	1,996	2,097	2,201
2007	1,990	2,103	2,221
2008	1,984	2,110	2,241
2009	1,978	2,117	2,261
2010	1,971	2,123	2,280
2011	1,964	2,129	2,299
2012	1,958	2,135	2,318
2013	1,951	2,140	2,337
2014	1,944	2,146	2,355
2015	1,937	2,152	2,374
2016	1,931	2,157	2,392
2017	1,924	2,163	2,411
2018	1,917	2,169	2,430
2019	1,910	2,175	2,448
2020	1,903	2,180	2,467
2021	1,896	2,186	2,485
2022	1,888	2,191	2,503
2023	1,881	2,197	2,522
2024	1,873	2,202	2,540

Table E-1 (continued)

Year	Low	Mid	High
2025	1,866	2,207	2,558
2026	1,864	2,208	2,563
2027	1,862	2,210	2,567
2028	1,860	2,211	2,572
2029	1,857	2,212	2,576
2030	1,855	2,213	2,580
2031	1,852	2,213	2,584
2032	1,849	2,213	2,588
2033	1,846	2,214	2,591
2034	1,843	2,214	2,594
2035	1,840	2,214	2,597
2036	1,837	2,214	2,601
2037	1,834	2,214	2,604
2038	1,832	2,214	2,607
2039	1,829	2,214	2,610
2040	1,826	2,215	2,613
2041	1,823	2,215	2,617
2042	1,820	2,215	2,620
2043	1,818	2,216	2,624
2044	1,815	2,216	2,627
2045	1,813	2,217	2,631
2046	1,810	2,217	2,634
2047	1,807	2,218	2,637
2048	1,805	2,218	2,641
2049	1,802	2,219	2,644
2050	1,800	2,219	2,647
2051	1,797	2,219	2,650
2052	1,794	2,219	2,652
2053	1,791	2,219	2,655

Table E-1 (continued)

Year	Low	Mid	High
2054	1,789	2,219	2,658
2055	1,786	2,219	2,660
2056	1,783	2,219	2,662
2057	1,780	2,218	2,665
2058	1,776	2,218	2,667
2059	1,773	2,217	2,669
2060	1,770	2,217	2,671
2061	1,767	2,216	2,673
2062	1,764	2,216	2,675
2063	1,760	2,215	2,677
2064	1,757	2,215	2,679
2065	1,754	2,214	2,682
2066	1,751	2,214	2,684
2067	1,747	2,213	2,686
2068	1,744	2,213	2,688
2069	1,741	2,212	2,690
2070	1,738	2,212	2,692
2071	1,734	2,212	2,695
2072	1,731	2,211	2,697
2073	1,728	2,211	2,699
2074	1,725	2,210	2,701
2075	1,721	2,209	2,703
2076	1,718	2,209	2,705
2077	1,715	2,208	2,706
2078	1,711	2,207	2,708
2079	1,708	2,206	2,710
2080	1,705	2,206	2,711
2081	1,701	2,205	2,713
2082	1,698	2,204	2,714

Table E-1 (continued)

Year	Low	Mid	High
2083	1,694	2,203	2,716
2084	1,690	2,202	2,717
2085	1,687	2,201	2,719
2086	1,683	2,199	2,720
2087	1,680	2,198	2,721
2088	1,676	2,197	2,723
2089	1,672	2,196	2,724
2090	1,669	2,195	2,725
2091	1,665	2,194	2,726
2092	1,661	2,193	2,728
2093	1,658	2,191	2,729
2094	1,654	2,190	2,730
2095	1,651	2,189	2,731
2096	1,647	2,188	2,733
2097	1,643	2,187	2,734
2098	1,640	2,185	2,735
2099	1,636	2,184	2,736
2100	1,632	2,183	2,737

Table E-2. Mortality rates (lifespan-equivalent)

Year	Low	Mid	High
1999	79.74	79.74	79.68
2000	80.01	79.9	79.95
2001	80.16	80.05	80.10
2002	80.32	80.2	80.25
2003	80.47	80.36	80.40
2004	80.62	80.51	80.54
2005	80.78	80.67	80.69
2006	80.94	80.82	80.85
2007	81.09	80.97	80.99
2008	81.24	81.13	81.14
2009	81.39	81.28	81.29
2010	81.66	81.43	81.54
2015	82.41	82.19	82.26
2020	83.26	82.93	83.07
2025	83.88	83.56	83.65
2030	84.60	84.17	84.33
2035	85.20	84.78	84.88
2040	85.90	85.4	85.53
2045	86.51	86.01	86.08
2050	87.21	86.62	86.71
2055	87.80	87.22	87.22
2060	88.47	87.81	87.82
2065	89.05	88.4	88.30
2070	89.70	88.97	88.85
2075	90.24	89.53	89.28
2080	90.86	90.09	89.80
2085	91.39	90.64	90.21
2090	92.00	91.18	90.69
2095	92.51	91.71	91.06
2100	93.00	92.24	91.42

Table E-3. Projected international migration

Year	Low	Mid	High
2000	670,138	1,020,435	1,432,695
2001	593,255	1,030,425	1,570,973
2002	531,919	1,030,293	1,671,881
2003	462,200	995,679	1,704,589
2004	404,294	961,750	1,722,833
2005	355,173	928,453	1,729,993
2006	312,734	895,833	1,728,678
2007	275,252	863,540	1,720,305
2008	241,908	831,875	1,706,158
2009	211,938	800,663	1,687,319
2010	184,668	769,797	1,664,477
2011	179,040	774,125	1,699,910
2012	174,268	778,360	1,733,468
2013	170,255	782,327	1,765,441
2014	166,833	786,288	1,795,997
2015	163,809	790,010	1,825,332
2016	161,032	793,650	1,853,500
2017	158,577	797,129	1,880,682
2018	156,347	800,507	1,907,047
2019	154,457	803,815	1,932,527
2020	152,592	806,979	1,957,368
2021	166,085	840,341	2,041,510
2022	179,067	873,156	2,125,426
2023	191,267	905,230	2,208,839
2024	202,980	936,857	2,292,232
2025	214,147	967,891	2,375,342
2026	224,882	998,465	2,458,340
2027	235,272	1,028,622	2,541,223
2028	245,307	1,058,406	2,624,090

Table E-3 (continued)

Year	Low	Mid	High
2029	255,086	1,087,913	2,706,829
2030	264,891	1,117,043	2,789,455
2031	257,533	1,110,301	2,797,150
2032	250,901	1,104,092	2,804,858
2033	244,888	1,098,310	2,812,504
2034	239,382	1,092,915	2,820,177
2035	234,661	1,088,005	2,827,787
2036	230,311	1,083,384	2,835,422
2037	226,341	1,079,014	2,842,719
2038	222,757	1,074,903	2,850,066
2039	219,520	1,071,081	2,857,333
2040	216,643	1,067,573	2,864,468
2041	213,996	1,064,300	2,871,496
2042	211,696	1,061,065	2,878,370
2043	209,507	1,058,054	2,885,342
2044	207,356	1,055,215	2,892,083
2045	205,448	1,052,456	2,898,727
2046	203,620	1,049,922	2,905,317
2047	201,922	1,047,331	2,911,896
2048	200,347	1,044,964	2,918,282
2049	198,712	1,042,694	2,924,605
2050	197,203	1,040,387	2,930,848
2051	195,780	1,038,197	2,936,971
2052	194,554	1,036,208	2,942,987
2053	193,265	1,034,270	2,949,104
2054	192,017	1,032,477	2,955,004
2055	190,931	1,030,777	2,960,927
2056	189,843	1,029,070	2,966,785
2057	188,619	1,027,388	2,972,462

Table E-3 (continued)

Year	Low	Mid	High
2058	187,498	1,025,691	2,978,058
2059	186,448	1,024,149	2,983,567
2060	185,302	1,022,613	2,989,148
2061	184,102	1,021,041	2,994,495
2062	183,076	1,019,540	2,999,780
2063	181,927	1,018,071	3,004,997
2064	180,848	1,016,574	3,010,217
2065	179,608	1,015,188	3,015,235
2066	178,444	1,013,731	3,020,298
2067	177,282	1,012,362	3,025,208
2068	176,119	1,011,015	3,030,105
2069	174,912	1,009,672	3,034,978
2070	173,632	1,008,378	3,039,655
2071	172,478	1,007,139	3,044,480
2072	171,308	1,005,888	3,049,252
2073	170,122	1,004,724	3,053,972
2074	169,012	1,003,494	3,058,708
2075	167,811	1,002,407	3,063,253
2076	166,788	1,001,342	3,067,923
2077	165,628	1,000,147	3,072,497
2078	164,501	999,100	3,076,939
2079	163,329	998,085	3,081,488
2080	162,359	997,098	3,085,910
2081	161,216	996,136	3,090,373
2082	160,066	995,099	3,094,759
2083	159,026	994,178	3,099,047
2084	158,066	993,261	3,103,452
2085	156,995	992,333	3,107,792
2086	155,981	991,460	3,112,064

Table E-3 (continued)

Year	Low	Mid	High
2087	154,984	990,630	3,116,245
2088	154,078	989,894	3,120,601
2089	153,015	989,122	3,124,829
2090	152,160	988,353	3,129,030
2091	151,203	987,576	3,133,229
2092	150,290	986,934	3,137,378
2093	149,360	986,237	3,141,620
2094	148,521	985,542	3,145,634
2095	147,543	984,815	3,149,846
2096	146,699	984,223	3,153,959
2097	145,840	983,650	3,158,109
2098	145,004	983,071	3,162,141
2099	144,195	982,546	3,166,233
2100	143,407	982,038	3,170,286

APPENDIX F

SUMMARY OF MAJOR MODEL INPUTS AND ASSUMPTIONS

Table F-1. Major demographic model inputs and assumptions

Model input	Source	Major assumptions
Initial 2005 county population	Bridged-Race Vintage 2006 dataset for 7/1/2005*	Population over age 85 in each county distributed by national over-85 population
Mortality rates	Component Assumptions of the Resident Population by Age, Sex, Race, and Hispanic Origin (U.S. Census Bureau, 2000)	
Fertility rates		Birth ratio of 1046 males to 1000 females (Matthews and Hamilton, 2005)
Net international migration		
Distribution of international immigration to counties	2000 U.S. Census, SF3, Table P22 (U.S. Census Bureau, 2007)	Distribution remains consistent with 2000 pattern.
1995−2000 Domestic migration	Public Use Microdata Samples (U.S. Census Bureau, 2003), organized by the NY State Data Center	Migration patterns from this time period are used to drive future migration
Natural amenities (January temperature, January sunlight, July temperature, July humidity, water area)	USDA Natural Amenities Database (McGranahan, 1999)	Amenities calculated for 1941−1970 stay constant over time
County growth rate (1980−2000)	2000 U.S. Census (SF1, Table P12), 1980 U.S. Census (U.S. Census Bureau, 1992)	The 1980−2000 growth rate remains a constant proxy for economic growth
Functional county-to-county distance	Based on population-weighted center of each county (as of 2000 Census) and transportation infrastructure of in the 2006 National Transportation Atlas	Function distance remains static in the model, despite changes in the geographic distribution of the population and anticipated changes in infrastructure

* NCHS (National Center for Health Statistics). (2007) Bridged-race Vintage 2006 postcensal population estimates for July 1, 2000 - July 1, 2006, by year, county, single-year of age, bridged-race, Hispanic origin, and sex. Released August 16, 2007. Centers for Disease Control and Prevention (CDC), Atlanta, GA. Available online at http://www.cdc.gov/nchs/about/major/dvs/popbridge/datadoc.htm#vintage2006.

Table F-2. Major SERGoM inputs and assumptions

<table>
<tr><th>Model input</th><th>Source</th><th>Major assumptions</th></tr>
<tr><td>Housing units</td><td>2000 U.S. Census (U.S. Census Bureau, 2007)</td><td>Future growth patterns are likely to be similar to those that occurred historically 1980−2000</td></tr>
<tr><td>Undevelopable lands</td><td>UPPT (unprotected, private protected, public protected, and tribal/native lands) dataset generated with data from the Conservation Biology Institute’s PAD v4 database (CBI, 2008). Updated with current data for 21 states</td><td>Development cannot occur on these lands</td></tr>
<tr><td>Road and groundwater well density</td><td rowspan="2">Fine-grained land use/cover data (NLCD 2001)</td><td rowspan="2">Changes in transportation infrastructure will follow the footprint of existing infrastructure</td></tr>
<tr><td>Commercial/Industrial land use</td></tr>
<tr><td>County population</td><td>Developed using demographic model</td><td>See Table F-1</td></tr>
<tr><td colspan="3">Cross-cutting assumptions:
• Drivers are hierarchical (national, state, regional, county, w/in county)
• New urban areas can be emergent (like cellular automata)</td></tr>
</table>

SC
IEN
CE

Recycled/Recyclable

Printed with vegetable-based ink on paper that contains a minimum of 50% post-consumer fiber content and processed chlorine free.

EPA
United States
Environmental Protection
Agency

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT NO. G-35

Office of Research and Development
National Center for Environmental Assessment
Washington, DC 20460

Official Business
Penalty for Private Use
$300