Environmental Science

http://www.aimspress.com/

Volume 2, Issue 3, 668-693.
DOI: 10.3934/environsci.2015.3.668
Received date 4 February 2015,
Accepted date 18 June 2015,
Published date 24 June 2015

Research article

Methods used to parameterize the spatially-explicit components of a state-and-transition simulation model

Rachel R. Sleeter [1,] *, William Acevedo [2], Christopher E. Soulard [2] and Benjamin M. Sleeter [3]

[1] U. S. Geological Survey, Eastern Geographic Science Center, Gig Harbor, WA, USA
[2] U. S. Geological Survey, Western Geographic Science Center, Menlo Park, CA, USA
[3] U. S. Geological Survey, Western Geographic Science Center, Tacoma, WA, USA

* **Correspondence:** Email rsleeter@usgs.gov; Tel: (253) 313-3310.

Abstract: Spatially-explicit state-and-transition simulation models of land use and land cover (LULC) increase our ability to assess regional landscape characteristics and associated carbon dynamics across multiple scenarios. By characterizing appropriate spatial attributes such as forest age and land-use distribution, a state-and-transition model can more effectively simulate the pattern and spread of LULC changes. This manuscript describes the methods and input parameters of the Land Use and Carbon Scenario Simulator (LUCAS), a customized state-and-transition simulation model utilized to assess the relative impacts of LULC on carbon stocks for the conterminous U.S. The methods and input parameters are spatially explicit and describe initial conditions (strata, state classes and forest age), spatial multipliers, and carbon stock density. Initial conditions were derived from harmonization of multi-temporal data characterizing changes in land use as well as land cover. Harmonization combines numerous national-level datasets through a cell-based data fusion process to generate maps of primary LULC categories. Forest age was parameterized using data from the North American Carbon Program and spatially-explicit maps showing the locations of past disturbances (i.e. wildfire and harvest). Spatial multipliers were developed to spatially constrain the location of future LULC transitions. Based on distance-decay theory, maps were generated to guide the placement of changes related to forest harvest, agricultural intensification/extensification, and urbanization. We analyze the spatially-explicit input parameters with a sensitivity analysis, by showing how LUCAS responds to variations in the model input. This manuscript uses Mediterranean California as a regional subset to highlight local to regional aspects of land change, which demonstrates the utility of LUCAS at many scales and applications.

Keywords: state-and-transition modeling; land use and land cover; spatially-explicit modeling;

scenarios

1. Introduction

Human activities are continually transforming the land surface, emphasizing the trade-offs between the use of natural resources and preserving ecosystem functions [1,2]. Changes in land use and land cover (LULC) include human-driven processes and biophysical processes, which can be monitored and assessed with spatial data and models. Changes such as the clearing of forests for agriculture or urban development directly affect the exchange of greenhouse gases between terrestrial ecosystems and the atmosphere at local-to-global scales [3-5]. LULC change has been recognized as a necessary component for global change assessments, [6-8] triggering an increased need for developing scalable models with infrastructure that address multi-disciplinary concerns.

Spatially-explicit LULC change models are developed to describe past, present and future landscape conditions, often used to communicate theories and concepts about possible outcomes related to human and environmental interactions. These models take many forms; the National Research Council has recently conducted a review and broadly categorizes the models as: machine learning and statistical, cellular, sector-based economic, spatially disaggregated economic, agent based, and hybrid [9]. Models operate with different rule sets and parameters, often sharing components as well as limitations. For example, the machine learning and statistical approach expresses land change through a selection of statistically explanatory variables, using logistic regression, neural networks or weighted relationships [10,11]. These models assume stationarity and often lack the potential to capture variability in future patterns.

Agent based models include parameters that reflect the collective behavior of human actors (e.g. land managers, farmers, developers) and how they interact with given landscape characteristics [12]. Most case studies are done from the bottom up to address local decisions and management [13,14]. Due to the complex process involving decision rules, surveys, and public opinion, the Agent Based Model is rarely used at the national or global scales. The CLUE (Conversion of Land Use and its Effects) model is a cellular model that is used from local to global scales, taking a top-down approach to allocate change often based on scenarios [9,15]. Suitability maps and neighborhood interactions of current conditions guide the allocation of change.

Markov-chain models were one of the first approaches to modeling the continuation of LULC historical trends [16]. A transition matrix of the historic time period is used to derive transition probabilities for different conversion types. However, there are a number of limitations using traditional Markov-chains for LULC analysis. The primary drawbacks involve stationarity of the transition matrix across time and space and difficulty pulling out the causal factors driving change [17]. Many advances have been made to combine Markov-chains with cellular models to improve simulation rules and spatial patterns [18,19].

State-and-transition simulation models can be considered a Markov chain, where the probability of a state transitioning at any given time depends only on the present state. However, modern state-and-transition simulation models (e.g. VDDT, ST-Sim) have adapted beyond the technical definition of the Markov chain in order to reflect the complex spatial relationships between states on a given landscape. Our research uses the ST-Sim software package, openly available through APEX Resource Management Solutions (www.apexrms.com). ST-Sim provides a combination of stochastic, deterministic and empirical methods to predict how variables transition from one state to another within a selected time period [19]. A grid of simulation cells is classified into states that represent the landscape. Probabilities are assigned to possible transitions between states, and the landscape is then simulated through time using Monte Carlo simulation methods. In addition to tracking state classes and transitions, ST-Sim also tracks the age and time-since-transition for every simulation cell. External factors like climate change and harvest rates can have probabilities that change at different times in the simulation period; therefore, age-dependent transitions are a key parameter that makes ST-Sim non-stationary and a suitable option for capturing variability in future projections.

The U.S. Geological Survey (USGS) recently developed a custom version of ST-Sim known as the Land Use and Carbon Scenario Simulator (LUCAS), to model future projections (2000–2100) of LULC and the associated impacts on carbon dynamics in the conterminous U.S. LUCAS takes an integrated approach where changes in LULC are modeled using state-and-transition simulation and carbon stocks and flows are modeled using a stock and flow model, yet these processes are parameterized uniformly and operate synchronously. By using a consistent framework (scale, temporal period, classification scheme, and model criteria) we demonstrate how a combination of biophysical processes and anthropogenic drivers can be modeled within a single platform. The spatial scale of 1-km resolution makes it possible to incorporate downscaled global emissions scenarios at very coarse resolutions, with regional summaries of historical LULC trends at fine resolutions [8,20]. All input parameters were seamlessly created for the conterminous U.S., which allows the inputs and outputs to be subset by any zone of interest from management to ecological. Spatially-explicit results link local, regionalized change to national and global change estimates, which contribute to the transferability of data and modeling methods to various scales.

The goal of this manuscript is to outline the methods and parameters used to develop LUCAS. Special attention is given to the spatially-explicit source data incorporated into the LULC sector of the model. In particular, we use a harmonization process where over 20 national-level spatial datasets undergo a data fusion process resulting in the LULC conditions used to initialize the model. The initial conditions consist of a set of cell based rasters including a strata map of ecoregional zones [21], a state class map of LULC, and a forest age map. Spatial multipliers are additional maps that use neighborhood effects and landscape suitability to guide specific transitions. The initial carbon stock maps are introduced, while a full methodology of the stock and flow model is given in Sleeter et al. [22]. The results analyze the spatially-explicit input parameters with a sensitivity analysis, by showing how LUCAS responds to variations in the model input. Although this research has been conducted for the conterminous U.S. for multiple scenarios, we use Mediterranean California as a subset to highlight the utility of LUCAS at the regional scale and show one scenario, A1B, from the Intergovernmental Panel on Climate Change - Special Report on Emission Scenarios (IPCC-SRES) [23].

2. Methods and model parameters

2.1. Study area

The study area encompasses four Level III ecoregions covering most of California (Figure 1). The ecoregions are characterized by a mild climate of hot, dry summers and cool, moist winters with precipitation occurring from frontal storms off the Pacific Ocean [21]. Extending north to south from the Klamath Mountains to the Baja peninsula and east to west from the Sierra Nevada Mountains to the Pacific Ocean, the landscape is characterized by shrubland vegetation of chaparral, mixed with grassland and open oak woodlands, agriculturally productive valleys and high urban population agglomerations. We include the Sierra Nevada ecoregion in our analysis to model forested land cover to demonstrate how harvesting associated with forest cover affects the carbon cycle. Using all four ecoregions captures all dominant human land uses in the state of California and allows for the examination of mitigation scenarios relative to forest harvest and agriculture extensification.

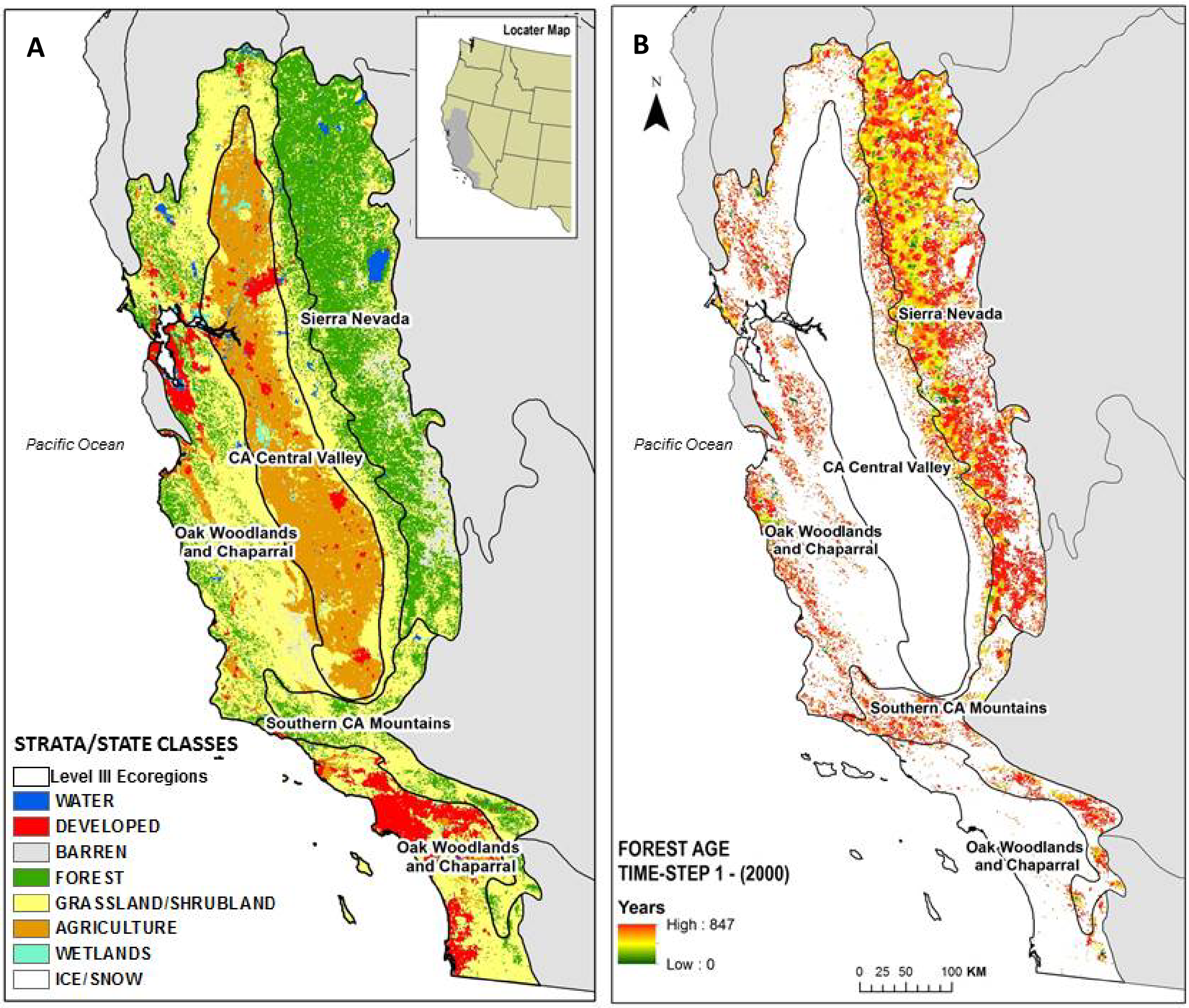

Figure 1. Initial Conditions: Strata/State Classes (A) and Forest Age (B). The strata consists of four U.S. EPA level III ecoregions (Sierra Nevada, Southern and Central California Chaparral and Oak Woodlands, Central California Valley, and Southern California Mountains) [21]. The state classes for the year 2000 are from a harmonized LULC product.

The ecoregions in the study area experienced 25,605 km^2 of LULC change between 1973 and 2000. The Central California Valley Ecoregion ranked the 6th highest changing ecoregion in the Western U.S. [24] with 7401 km^2 of change. The other ecoregions ranked 10th (Southern and Central California Chaparral and Oak Woodlands), 18th (Southern California Mountains), and 19th (Sierra

Nevada), with 13,009 km^2, 1413 km^2, and 3782 km^2 of LULC change between 1973 and 2000, respectively. The dominant LULC classes include grassland/shrubland (39.3%), forest (28.6%), agriculture (19.7%), and developed (6.3%). The overall spatial change was highest in the Central California Valley Ecoregion, estimated at 12.9% between 1973 and 2000. Overall change in the Southern and Central California Chaparral and Oak Woodlands Ecoregion was estimated at 9.7%. The Southern California Mountains and Sierra Nevada ecoregions had 5.1% and 5.0% overall spatial change, respectively. The most common types of LULC change across the study area were related to wildland fire and regrowth of vegetation after fires. The second most common types of change were linked to agricultural expansion and contraction in response to economic and climatic conditions. The third major type of change has been development, primarily conversion of agriculture to developed land in the Southern and Central California Chaparral and Oak Woodlands.

2.2. Initial conditions

The initial conditions for the LUCAS model describe the strata, LULC state classes, and forest stand age present at the start of the model simulation period (2000). The initial conditions need to be identified through a set of spatially-explicit input datasets, where the land use and land cover and forest age for each ecoregion (strata) are defined. Each component is explained in detail within this section of the manuscript.

2.2.1. Strata

The strata file is used to divide the landscape into contiguous zones such as ecoregions, management areas, vegetation types, or administrative boundaries. Each stratum has a unique pathway diagram and every cell within the entire landscape is assigned to its strata at the start of a simulation. Ecoregions originally developed by Omernik [21] and later modified by the U.S. Environmental Protection Agency [25] provide the strata used in the LUCAS model (Figure 1). Ecoregions are areas that share common characteristics relative to biophysical settings, management strategies, and socio-economic factors. Ecoregions have been demonstrated to be an effective framework for characterizing changes in U.S. land cover [26]. Stratifying the landscape allows the model to better simulate the spatial variability between ecoregions. If forest harvest rates are high in the Sierra Nevada, but nonexistent in the Central California Valley, we are able to parameterize the pathways and transition probabilities regionally to reflect that assumption.

2.2.2. State Classes

Initial state classes (Figure 1A) were set in LUCAS by using a circa-2000 LULC dataset that was created by harmonizing existing multi-temporal LULC datasets [27]. The methodology uses the principal of convergence of evidence to reclassify areas mapped by multiple national-scale LULC datasets. Datasets were combined through a cell-based data stacking process that determined the most likely cell classification based on input dataset agreement. This process allows for classification errors in some of the input datasets. Harmonization methods [28,29] were necessary because a wide range of national-scale land use/land-cover classification efforts exist, yet major differences between these data arise because of different methodologies, class definitions, and mapping objectives.

The current methodology used over 20 national LULC datasets that were readily available from various archives (Table 1). The datasets ranged in vintage from 1992 to 2011. The harmonization process (Figure 2) begins by identifying cells that are persistent land cover and do not change over the 19-year harmonization period. Based on previous USGS research [24], the amount of persistent cover should approximate 90% of the conterminous U.S. land surface between 1992 and 2011. Most differences from this amount are due to classification errors in the input datasets. In developing the harmonized dataset, only 54% of 30-m raster cells had 100% agreement between input datasets. Expanding the rules of cell inclusion to cells where a clear majority of the input datasets were in agreement led to 82% of the U.S. classified as persistent cover. Persistent cover increased to 90% of the land surface when cells were assigned a cover type even though the input dataset agreement was evenly split between two classes. For example, in the western U.S., areas changing from grassland to shrubland or classification confusion between pasture/hay and cultivated crops correspond to areas that were transitional between cover types or spectrally similar to cultivated crops. The remaining 10% of the land surface was cells that experienced a change in land cover or cells that had substantial disagreement between datasets and could not be classified based on dataset agreement. These latter cells tended to correspond to areas that were difficult to map because they contained mixed spectral responses, such as riparian areas. Data from the National Land Cover Database 2001 [30] was used to populate remaining unclassified cells. The resulting dataset was then visually inspected for quality assurance. Artificial boundaries and other processing artifacts in the input data were minimized by harmonization. The addition of three disturbance classes (forest harvesting, wildland fire, and surface mining) increased the utility of the harmonized dataset for land change analysis and modeling.

Table 1. Datasets used in the harmonization process to create a circa-2000 land use and land cover dataset for setting state classes in the LUCAS model.

Data Sources	Description
National Land Cover Database (NLCD)	Six datasets available (1992, 2001, 2006, 2011, Retro 1992, & Retro 2001) [30, 31, 32, 33]
LANDFIRE geo-spatial layers	Six datasets available (Existing Vegetation Cover (2001 refresh, 2008 refresh, & 2010 refresh) and Existing Vegetation Type (2001 refresh, 2008 refresh, & 2010 refresh)) [34]
Cropland Data Layer (CDL)	Six datasets available (2008–2013) [35]
National GAP Land Cover	One dataset available (2001) [36]
Annual datasets available on vegetation disturbance	Global Forest Cover (2000–2012) [37]; LANDFIRE Disturbance (1999–2010) [38]; WELD Land Cover Land Cover Change (2006–2010) [39]; Monitoring Trends in Burn Severity (1984–2012) [40]

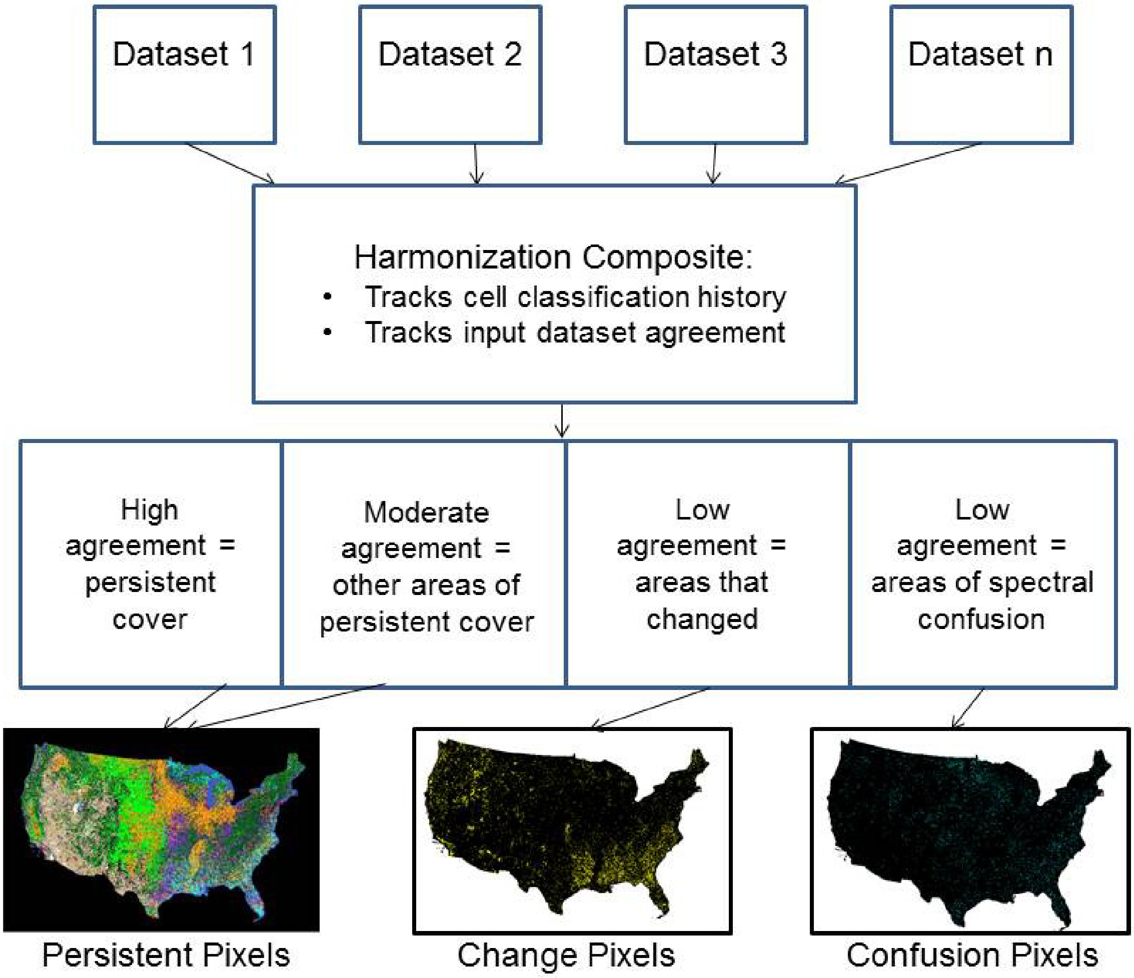

Figure 2. Conceptual diagram of the harmonization process showing the major output components.

The harmonized LULC product was subsequently recoded to match the state classes defined for the LUCAS state-and-transition simulations. LUCAS uses state classes to represent eight broad-scale LULC classes. The classification scheme used is based largely on the classification system developed by Anderson et al. [41] with some small differences (Table 2). The classification scheme represents a hybrid between use and cover, with some classes purely describing land cover, and others a mixture of cover and use. For example, development describes built-up features of the land surface and includes features such as paved surfaces, infrastructure, and housing, as well as vegetated areas predominated by a relatively high intensity anthropogenic use (e.g. golf courses). All areas of mechanical disturbance in the harmonized dataset (i.e. forest harvesting) were recoded to forest, while areas of natural disturbance in the harmonized dataset (i.e. wildfire) were recoded to grassland/shrubland or forest, based on the underlying land cover from the National Land Cover Database (NLCD). Further aggregation involved merging the cultivated crops and hay/pasture classes into a single agriculture class, merging the herbaceous and woody wetlands classes into a single wetlands class, and merging all four NLCD developed classes (open space, low intensity, medium intensity, and high intensity) into one developed class.

Table 2. Descriptions of the land-use/land-cover classes used to set the state classes in the LUCAS state-and-transition-simulation model.

LULC State Class	Description
Water	Water includes estuaries and bays, canals/aqueducts, lakes, reservoirs, rivers and streams. Cells are classified as water if vegetation and/or soils make up less than 25% of the area.
Developed	Development includes residential, industrial, commercial, transportation, and areas such as parks or other open spaces surrounded or otherwise dominated by an urban landscape.
Barren	Barren includes bare rock, gravel, sand, silt, clay, or other earthen material, with little or no "green" vegetation present. Vegetation, if present, is widely spaced and scrubby; lichen cover may be extensive.
Forest	Forests are distinguished from other vegetated surfaces based on having a tree-crown areal density greater than 20%. However, within the forest class we also include areas of recent harvest and natural disturbance, where tree cover may not meet this threshold.
Agriculture	Agricultural lands are characterized as any area used for the production of food and fiber, including cultivated cropland, pasture, orchards and vineyards, nurseries and ornamental horticulture areas, and confined livestock feeding operations.
Grassland/ Shrubland	Areas dominated by a combination of grasses (herbaceous vegetation) and shrubs (natural semi-woody vegetation, less than 6 m tall). These areas are not subject to intensive management such as tilling, but can be utilized for grazing.
Wetlands	Wetlands are those areas where the water table is at, near, or above the land surface for a significant part of most years.
Perennial Ice/Snow	All areas characterized by year-long cover of ice and/or snow.

2.2.3. Forest age

Initial conditions include a forest age parameter (Figure 1B) that is used in the simulation of forest harvesting and wildfire. In addition to tracking the state of each simulation cell, ST-Sim can also track the age of each cell. Each simulation cell is assigned an initial forest age for the year 2000, the initial year of time-step 1. By incrementing this age by one for each time-step of the simulation, the age can be tracked over time [19]. In addition, a maximum and minimum age limit can be specified for each state of forest use. Once the age for a forest simulation cell reaches the upper limit or a harvest age for its current state and strata, the forest cell becomes eligible for harvesting. For example, if the minimum age at which a forest cell can be harvested is 50 years, the forested cell would not be eligible to harvest until the cell reaches 50 years. Once a forested cell undergoes a harvest, the cell age is reset to zero. Age-based transitions are one of the major parameters in the ST-Sim software, which make the model non-stationary over time. Forest age is also one of the primary variables that allows LULC change and carbon stock and flow to be related. Carbon attributes (e.g. stocks, flux rates, automatic flows and event-based flows) all have an age factor.

An initial 2000 forest stand age map was generated from a combination of sources. The primary source came from the North American Carbon Program (NACP), where Pan et al. [42] used Forest Inventory and Analysis (FIA) data, historical fire data and the images from the National Aeronautics and Space Administration, Landsat Ecosystem Disturbance Adaptive Processing System project [43] to develop a national stand age dataset at 250-m resolution. We also used Vegetation Change Tracker [44] data to track forest disturbance. The spatially-explicit data from Pan, et al. [42] provides forest age for the year 2006 for all forested areas in the U.S. and Canada at 1-km resolution. These data were recalculated to represent forest age at the year 2000, which is the start of our scenario simulations. A general stand age distribution for each forest type (evergreen, deciduous, and mixed) was generated for each ecoregion. The ages were then reduced by enough years to get back to the 2000 start date. Cells with negative values at that point (meaning they had been harvested in the intervening years) were reassigned so that cell age at harvest generally matched the rest of the distribution (e.g. if the bulk of nearby forest in the study area was 60–80 years old in 2005, most harvested forest was randomly assigned an age somewhere in that range). For ecoregions with large extents of old growth forest, in some cases, forest was assigned to age ranges up to 350 years. USGS Land Cover Trends harvest rates [24] were used as a quality control to make sure the stand age by ecoregion had a high correlation. Vegetation Change Tracker provided the year of disturbance for the period 1984–2000. The disturbance year was converted to a separate 2000 stand age map by subtracting year of disturbance from the starting age of 2000. The disturbance age map was overlaid onto the Pan et.al. [42] stand age map to provide finer detail on known forest disturbance relative to the regionally stratified landscape.

2.3. Spatial multipliers

Spatial multipliers, also referred to as probability surfaces or suitability maps, are cell-based maps that allow external drivers such as landscape characteristics and neighborhood effects to influence the spatial patterns of LULC change. In the ST-Sim software, spatial multiplier values range from 0-1, where 0 prohibits a transition from occurring and 1 identifies areas that are the most suitable for transition. Spatial multiplier cells with the highest value (probability) are the most likely to convert to another state class, while cells with the lowest value are the least likely to transition. For example, a map of soil fertility might be incorporated in a spatial multiplier to indicate cells with a higher probability of agricultural expansion.

Within the LUCAS framework, the spatial distribution of future change across simulation cells is largely determined by spatial multipliers for four specific state class transitions (Table 3). These input parameters are critical: without them, landscape changes occur in a random pattern, and while area targets are achieved in each time-step, cell placement has no spatial significance relative to existing LULC or external drivers of change. For this research, the spatial multipliers were static over time, meaning that only one set of maps/cell values were input into LUCAS at the model initialization. ST-Sim provides users the capability to insert additional spatial multiplier values at selected time-steps; however, as yet, we have not used this feature.

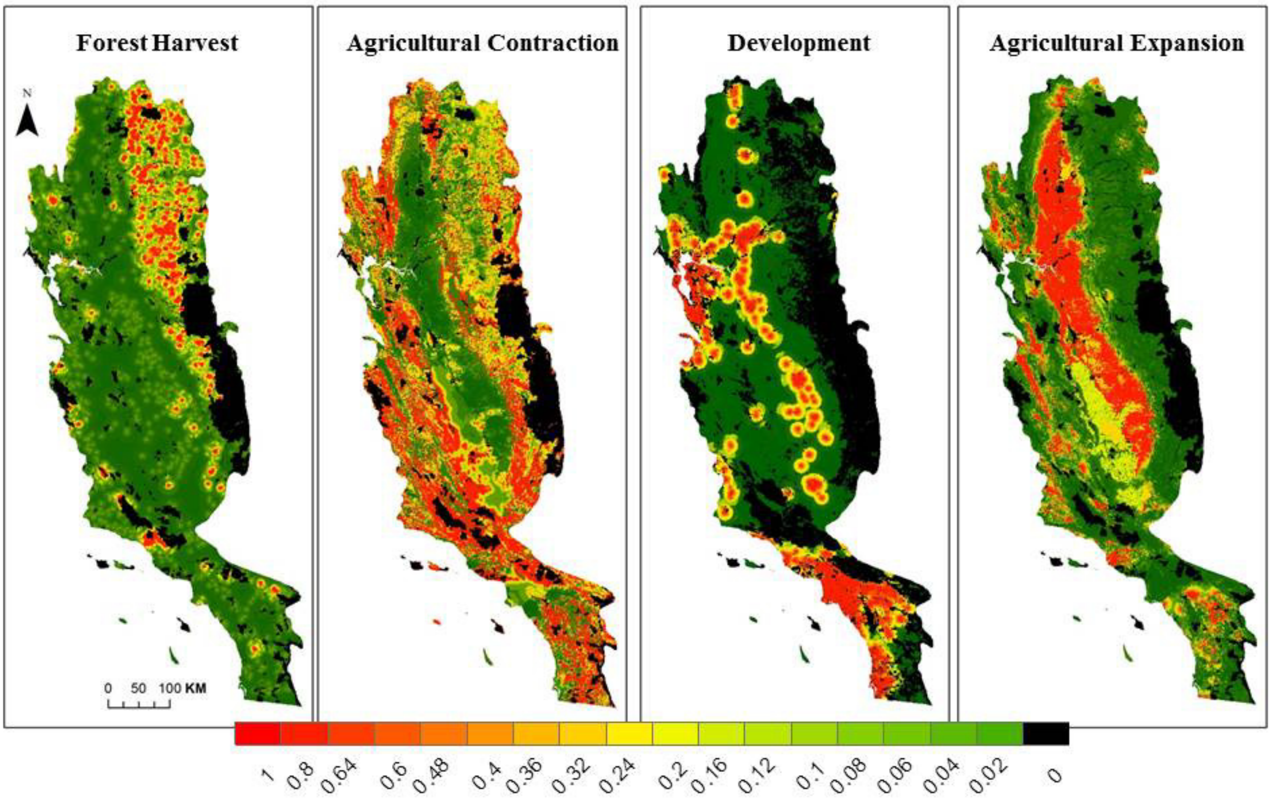

Figure 3. Spatial multipliers representing four major transition types (Forest Harvest, Agriculture Contraction, Development and Agricultural Expansion. Values range from 0–1, where 0 prohibits a transition from occurring and 1 identifies areas that are the most suitable for transitioning.

In our development of spatial multipliers (Figure 3), distance-decay theory was used as one of the ways to describe the effect that distance has on spatial interactions. The distance-decay theory states that the interaction between two locales declines as the distance between them increases [45]. For example, any gains or losses in the agricultural state class are more likely to occur in raster cells adjacent to, or within a short distance of, existing agricultural land uses. An important objective of this paper is to present source data and methods that are easily transferable to other projects. With this in mind, the primary source of the spatial multipliers is each existing state class and the distance calculated "to" the existing state class. A Euclidean distance algorithm was used measuring the straight-line distance from every cell to the nearest source cell of the four major transition types in Mediterranean California (Forest Harvesting, Agricultural Contraction, Development and Agricultural Expansion) (Table 3). The "distance-to" rasters were stratified into bins measuring 1–2 km each, displaying a distance-related gradient away from the particular state class. Distance-weighted rasters were then merged with additional contributing variables (i.e. soil condition, protected areas, population density, historical harvest and fires) to augment probabilities and further refine where changes were most suitable on the landscape. For example, the incorporation of the USGS Gap Analysis Program (GAP) Protected Areas Database [46] led to prohibited all LULC conversions (zero probabilities for these cells) on protected lands such as national parks. The spatial multipliers for each transition were developed at a 1-km resolution for the conterminous U.S., with seamless boundaries between ecoregions. All of the spatial inputs for LUCAS are developed nationally, but can be subset to any area of interest.

Table 3. Descriptions of the spatial multipliers chosen to represent spatially constrained LULC conversions. Spatial multipliers remain static over time.

Multiplier	Description	Source Data
Forest Harvesting	Sets probabilities for allowable forest harvesting transitions based on distance to historical "tree loss" (1984–2009 cumulative harvest from Vegetation Change Tracker). Conversions on protected lands were restricted (GAP 1 & 2). Small forest patches and historic fire cells from the Monitoring Trends in Burn Severity dataset were removed prior to the creation of the Euclidean distance surface."	Vegetation Change Tracker [44] PAD-US-GAP [46] Monitoring Trends in Burn Severity [40]
Agricultural Contraction	Sets probabilities of conversion to grassland/shrubland based on distance to existing grassland/shrublands and low crop capability. A majority filter of eight cells was applied to minimize the effect of small grassland/shrubland patches on modeled changes. Conversions on protected lands (GAP 1 & 2) were restricted.	Harmonized LULC [27] Crop Capability (USDA-NRCS-Soil Survey) [47] PAD-US-GAP [46]
Development	Sets probabilities of conversion to developed land with the highest probabilities occurring on land closest to existing, high density development (> 80 people/km^2). Distance to development was calculated and cells > 20 km^2 away from existing development were excluded. A majority filter of eight cells was also applied. Distance to development and distance to high population density were multiplied to produce final probability map. Conversions not allowed on protected lands (GAP Status 1 & 2 & 3).	Harmonized LULC [27] PAD-US-GAP [46] Gridded Dasymetric Population Density – USGS [48]
Agricultural Expansion	Sets probabilities of conversion to agriculture based on distance to existing agriculture and high crop capability. Emphasis placed on agricultural cells in the harmonized dataset that remained agriculture throughout the 1992–2011 time period. Restricts conversion on protected lands (GAP 1 & 2).	Harmonized LULC [27] Crop Capability(USDA-NRCS-Soil Survey) [47] PAD-US-GAP [46]

2.4. Initial carbon stocks

Initial carbon values (kgCm^2) for pools of forest biomass, deadwood, litter, and soil organic carbon are used by the LUCAS model to assess the linkages between LULC change and carbon dynamics within one modeling framework. The stock and flow model tracks the flow of carbon between pools on forest cover cells. In addition, the stock and flow model has a spatially-explicit component where the initial carbon stocks are represented in raster form. To accomplish this, two types of flows are calibrated (automatic and event-based). Automatic flows occur every year (e.g. growth rates based on age, mortality, litterfall, etc.), whereas event-based flows are triggered by LULC disturbance (e.g. wildfire, harvest, etc.). Event-based flows invoke an additional set of flow values to account for the types of LULC change. Annual simulations output from LUCAS allow the user to visually analyze carbon stock over time in relation to different amounts of disturbance (wildfire and harvest). Under a range of scenarios, conditional questions can be addressed (e.g. If X/per year is harvested in the Pacific Northwest, what are the impacts on carbon balance?).

For the initial carbon stock raster layers (Figure 4), we used a set of carbon stock density maps from the U.S. Forest Service FIA plot data [49]. These maps contain forest carbon estimates for the following pools: above ground biomass, below ground biomass, standing deadwood, down deadwood, litter, and soil. To be consistent with the carbon pools used in the stock and flow model, we combined the two biomass pools into a single living biomass map and the two deadwood pools into a single deadwood map. Because these carbon stock maps were created independently from the initial state class maps, we had to address situations where some cells were classified as forest but did not have matching spatial correspondence with the values from the FIA maps. To resolve this, we assigned the average carbon densities, summarized by ecoregion, to any forest cells not classified as

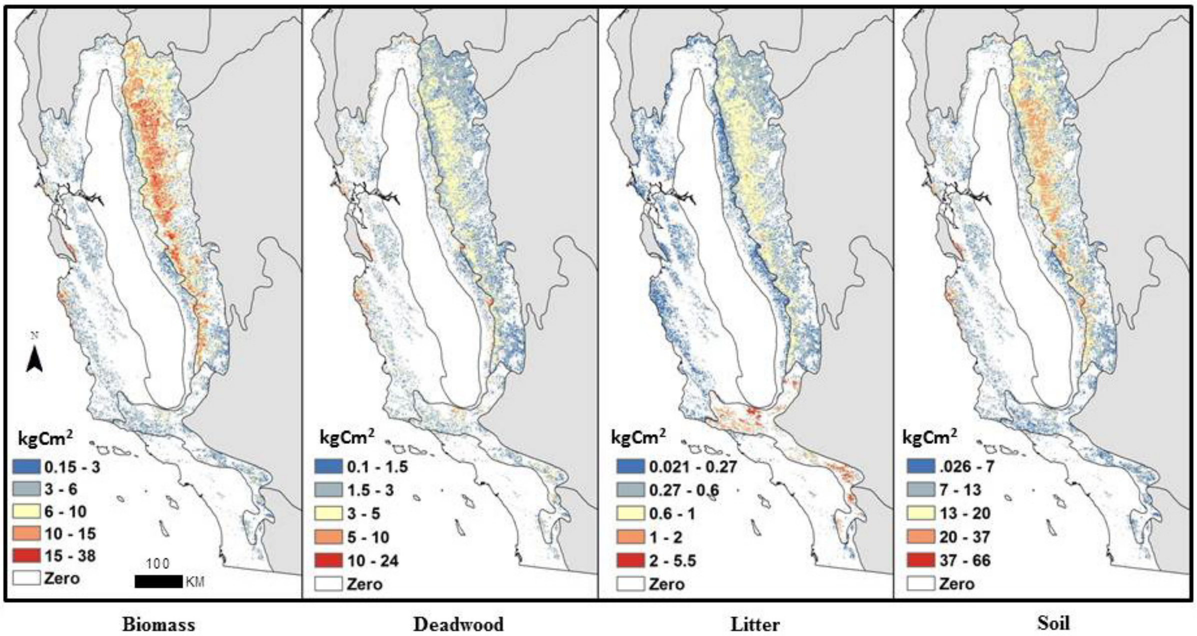

Figure 4. Initial carbon stock density maps (Biomass, Deadwood, Litter, Soil Organic) modified from U.S. Forest Service FIA plot data [49] to meet the input stock and flow parameters defined by LUCAS.

forest in the FIA data. Figure 4 shows the four carbon pools associated with forest cover. For more information regarding the derivation of carbon flow values between pools and the integration of the stock and flow model with the state-and-transition framework, see Sleeter et al. [22].

2.5. Area targets and scenarios

There are circumstances in LUCAS in which transition pathways are not represented using probabilities; rather a deterministic or fixed target is used. Transition targets, or area targets, are specified for each transition type and can vary at each time-step. The area targets for LUCAS were generated based on a downscaling method by Sleeter, et.al (2012) [20], which translates global gridded LULC projections from the Intergovernmental Panel on Climate Change—Special Report on Emission Scenarios (IPCC-SRES) [8,23,50] to regional scales (ecoregions) for the conterminous U.S. For our spatially-explicit modeling approach, we used area targets incremented at 5-year horizons until 2100 to enumerate amounts of change over time between state classes. Extensive calibration has been conducted both qualitatively by regional experts and quantitatively by assessing how much land is available by ecoregion, for each state class, to meet land-use demands over time. Conversions from one state class to another are based on the area target for the given time-step and the number of simulation cells eligible for the transition pathway during each time-step of the simulation [19]. Historic rates of change by ecoregion [24] serve as a baseline for future conditions. Future scenarios (i.e. A1B), from global assessment models were downscaled to be compatible with the spatial framework of LUCAS [20].

We developed multiple scenarios to model within LUCAS for the conterminous U.S. These included narratives and spatial data from the global community: IPCC-SRES (A1B, A2, B1, B2), Representative Concentration Pathways (8.5, 6.0, 4.5, 2.6) [51]; and extrapolations of historical trends based on findings from USGS Land Cover Trends such as: “business as usual” (1992–2000), “drought” (1986–1992), and “random” (1972–2000) [24]. The following section shows results from the A1B scenario. A1B is characterized in the United States as strong economic growth in a global market, rapid technological innovation, medium population growth, yet strong increase in urbanization. The high demand for food, biofuels and wood products leads to the intensification of agricultural and forestry land uses.

To conceptualize the input parameters discussed in the methods section relative to the full functionality of LUCAS, we categorize the major components recommended to replicate our efforts (see Figure 5).

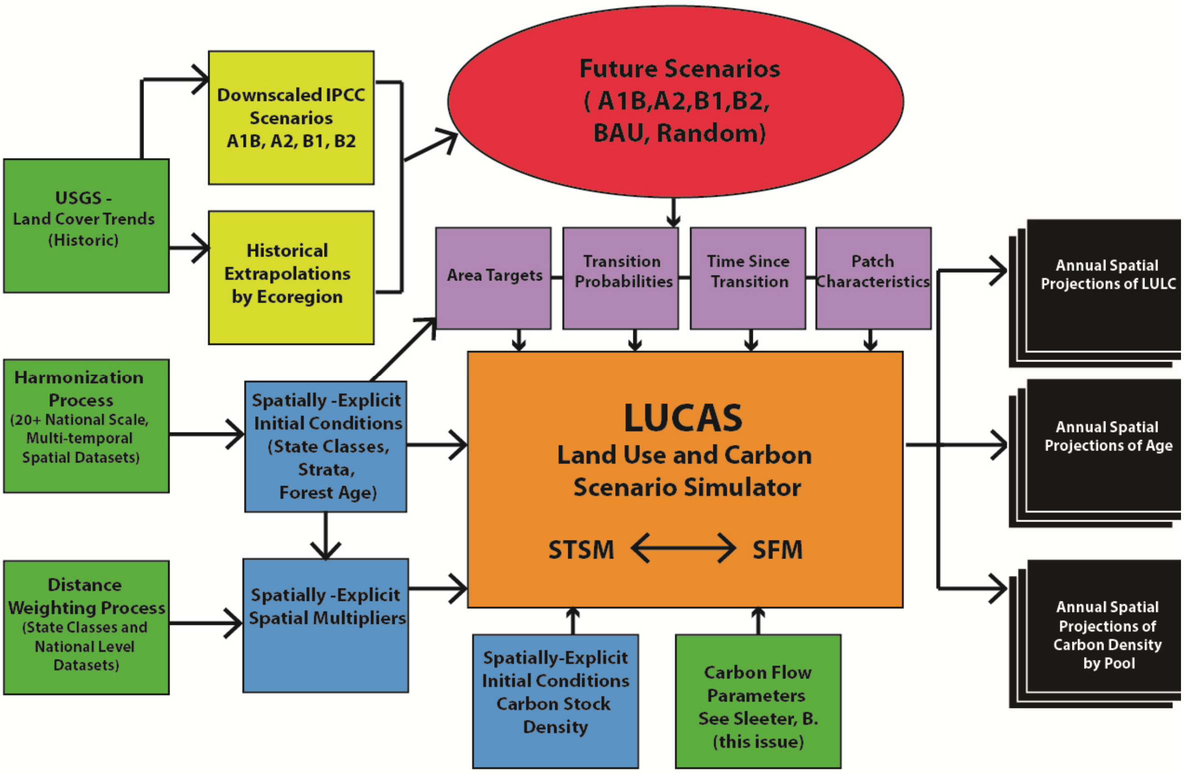

Figure 5. Conceptual diagram of LUCAS. Green = External data sources; Yellow = Scenarios based on historical trends or future global assessments; Blue = Spatially Explicit Parameters; Red = Future scenarios parameterized for LUCAS integration; Purple = State-and-transition simulation model functions; Orange = LUCAS integrated model; Black = Output options. Diagram modified from Wilson, et.al. (2014) [52].

3. Results and Discussion

3.1. LUCAS verification

There are two questions commonly addressed when verifying and validating the performance of a spatially-explicit model: (1) did the model code achieve the intention of the modeler (verification) and (2) how accurate was the spatial allocation or distribution of change across the landscape (validation). Model verification, different than model validation, tests the model code to verify whether the input commands were properly executed. Since the LUCAS model uses annual area targets for each transition group (e.g. agriculture to developed), we verified the model's performance by comparing the input area targets for scenario A1B with the modeled output at each 5-year time-step.

Three common transition groups (Grass-shrub to Agriculture, Forest Harvest, and Agriculture to Developed) are analyzed and expressed in Figure 6. The X-axis represents the input area targets (km^2) at 5-year intervals that are expected to occur during the 100-year simulation. The Y-axis represents the output results at the same 5-year interval, modeled in LUCAS. By looking at the graphs in Figure 6, it is clear that there is a very high correlation between the A1B area targets and the modeled output for each transition group tested. The R^2 values indicate a near perfect correlation, verifying that the code within the model is apportioning the exact annual increments of LULC change on the landscape. The spatial accuracy of cell placement, is not analyzed with these graphs.

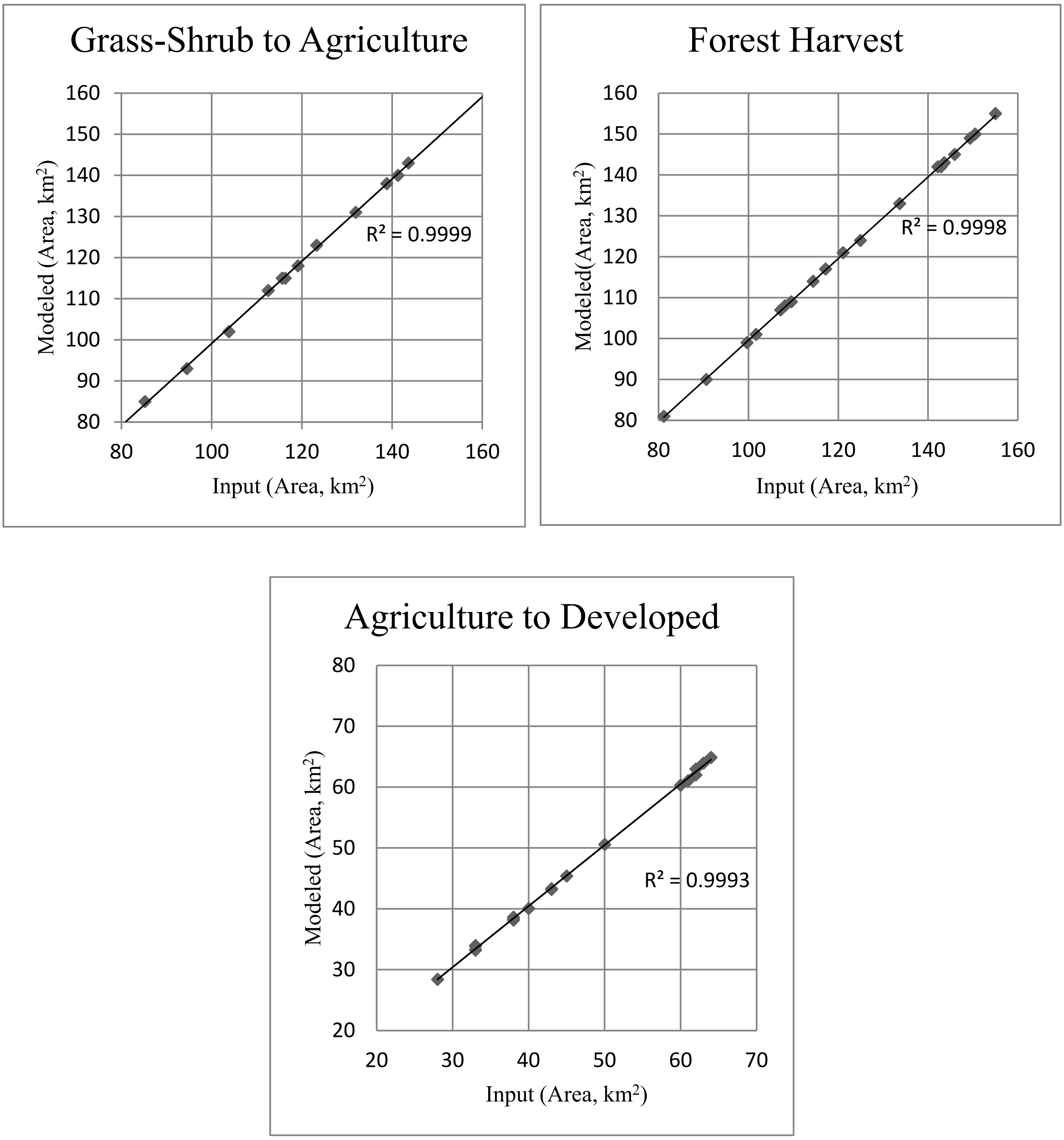

Figure 6. Model verification—Verify that the model code is executing the correct amount of change per time-step, given the area target input values. Very high correlations indicate that the model is performing the specified commands.

3.2. 100-year comparisons

One way to gauge the spatial impact of LUCAS is through observation and visualization. State class maps by strata are easy to export for selected years. Figure 7A shows an initial conditions state class map of the year 2000 for Mediterranean California. Figure 7B shows the A1B scenario modeled out to the year 2100. The A1B scenario specifies a strong, open global market and rapid technological innovation, leading to high demand for wood products and agricultural intensification [20,23]. The primary observable change between Figure 7A and B is the increase in developed/urban area (shown in red). The Oak Woodlands had the largest amount (~ 19,000 km^2) of land transition to developed land uses. From the maps, it is observed that the pattern of urban spread increases around the existing major metropolitan areas of the San Francisco Bay Area and Los Angeles/San Diego. The influence of the spatial multipliers, with protected lands set to zero, prohibits future urban development on areas not eligible for conversion. An example of the limited urban development is apparent in the East Bay hills (Figure 7), due to the network of open space/parks that are protected, yet adjacent to a sprawling urban agglomeration. Within the Central Valley ecoregion, developed lands increase (~ 7500 km^2) and are visible in Figure 7, north to south along the I-99 corridor. The Southern California Mountains and Sierra Nevada show very low net change in urban expansion.

Table 4. Zonal statistics of the LUCAS 100-yr change amounts output by state class area and stratified by ecoregion.

*Area km^2	Oak Woodlands and Chaparral			Sierra Nevada		
State Classes	2000	2100	*Net Change*	2000	2100	*Net Change*
Water	1500	1482	−18	1096	1096	0
Developed	10,204	29,251	19,047	62	68	6
Barren	1480	1480	0	3224	3224	0
Forest	16,265	16,170	−95	33404	33404	0
Grassland / Shrubland	63,618	39,746	−23,872	14563	14556	−7
Agriculture	5839	10,829	4990	162	163	1
Wetlands	863	830	−33	196	196	0
Ice/Snow	0	0	0	25	25	0
	Central Valley			Southern California Mountains		
State Classes	2000	2100	*Net Change*	2000	2100	*Net Change*
Water	473	339	−134	82	82	0
Developed	2304	9933	7629	234	1881	1647
Barren	501	501	0	168	168	0
Forest	71	71	0	4987	4925	−62
Grassland / Shrubland	9268	4647	−4621	12190	10571	−1619
Agriculture	32,067	29,618	−2449	112	146	34
Wetlands	1267	849	−418	47	47	0
Ice/Snow	0	0	0	0	0	0

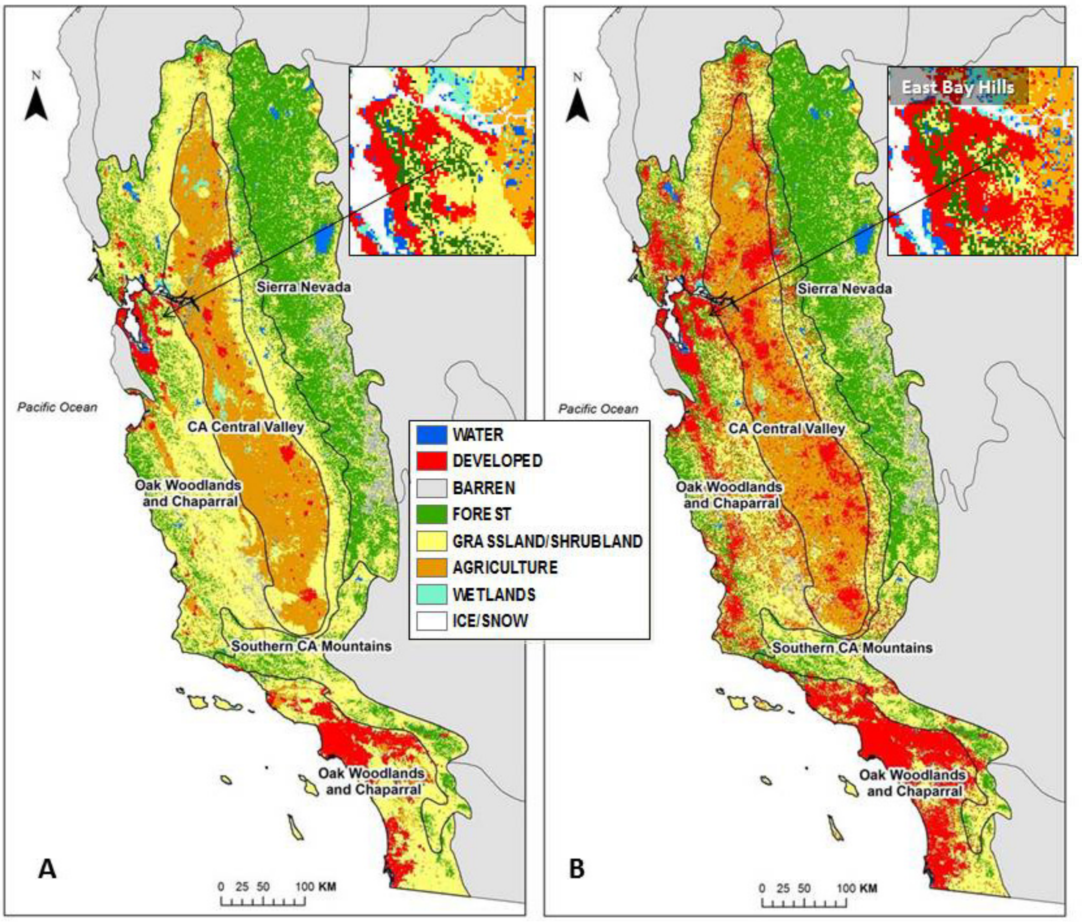

Figure 7. A is the state class map of the initial LULC conditions in the year 2000; B is the projected state class map in the year 2100 modeled with the A1B scenario.

The second highest state class change over the 100-year period is the increase in agriculture (shown in orange). The California Central Valley is one of the most important agricultural centers in the United States [53]. In order to meet the A1B demand/area target for agriculture, offset by the losses in agriculture due to urbanization, there is an expansion of the agricultural footprint along the eastern edge of the Central Valley into the foothills. Even though the soil quality to the east is not as desirable as that of the valley floor, these areas can support high value crops (e.g. vineyards and orchards). The Oak Woodlands also shows agricultural intensification (~ 5000 km^2), mainly confined to coastal valleys. These areas follow a similar pattern, where urban expansion pushes new agriculture to the periphery, resulting in new vineyards and orchards [53]. The increased demand for agricultural products, tied with urban expansion, drives competition between both land uses.

Urban expansion and agriculture intensification commonly occurs from the conversion of grassland/shrubland. As a result, the grassland/shrubland state class (shown in yellow) results in massive declines in the Oak Woodlands (−23,872 km^2) over the 100-year model period. The slight changes seen in the forest state class in Table 4 can be attributed to unidirectional change (e.g. urbanization). State class output cannot be used to visualize changes in forest use and disturbance. In order to look at significant changes related to forest cover and use, harvest and wildfire impacts must be measured over the 100-year simulation period (see Figure 8).

3.3. Sensitivity analysis

Sensitivity analysis is a technique by which the researcher evaluates the variation in the model output resulting from specific amounts of variation in model input, parameter values, or structure [9]. LUCAS uses Monte Carlo simulations to address uncertainty by processing multiple iterations with the same input parameters, showing a range of possible outcomes. The variability of transition can be measured for each transition group by using a powerful function in ST-Sim called the annual average probability distribution. This distribution measures how many times, on average, a particular cell transitioned from one class to another. The value is equal to the sum of the number of times the cell transitioned across all time-steps and iterations divided by (the number of time-steps * number of iterations). We use this distribution to summarize and visualize our output maps and test how the model responds to a changed input parameter.

Areas of forest that have been harvested or disturbed due to wildfire do not transition out of the forest class; rather, forest age is reset to zero once the disturbance occurs. This allows us to visualize changes in forest age structure and the spatial distribution of disturbance. Figure 8A shows the net change in forest age over the 100-year simulation period using 20 Monte Carlo simulations. Net change marks the difference between the forest age at the beginning of the simulation (2000) and the forest age at the end (2100). The result is a difference map showing the spatial distribution of forest age across the landscape. Positive values indicate a net increase in average age, while negative values indicate a net decrease in average age. The net difference in age over time is an indicator of disturbance intensity. Age difference values of greater than 100 (dark green on Figure 8A) reveal forest cells that were left undisturbed (did not undergo harvest or wildfire) for the entire simulation period.

Figures 8B–C show the average annual transition probability of forest wildfire and forest harvest from 2000–2100. Figure 8C illustrates the spatial variability and intensity of harvest over the 100-year period. The Sierra Nevada ecoregion had the highest historical forest cutting rates within Mediterranean California [24]. Future harvest under scenario A1B, shown in Figure 8C, follows that same trend. The region is largely protected through establishment of national parks and wilderness areas, which prohibit forest harvest activities. Therefore, much of the harvest occurs on private timberlands, which are mostly located at lower elevations. Protected areas are, however, prone to wildfire. Patch characteristics are used to influence the size distribution of the two disturbance types. Based on historic patch size and patterns, wildfires tend to be larger and more contiguous, while forest harvest is fragmented. One area that stands out with a high transition probability for wildfire is the Big Sur coast, south of Monterey Bay. Another notable wildfire simulated by LUCAS occurs west of Yosemite National Park. Interestingly, this is the same area that experienced the large “Rim Fire” of 2013.

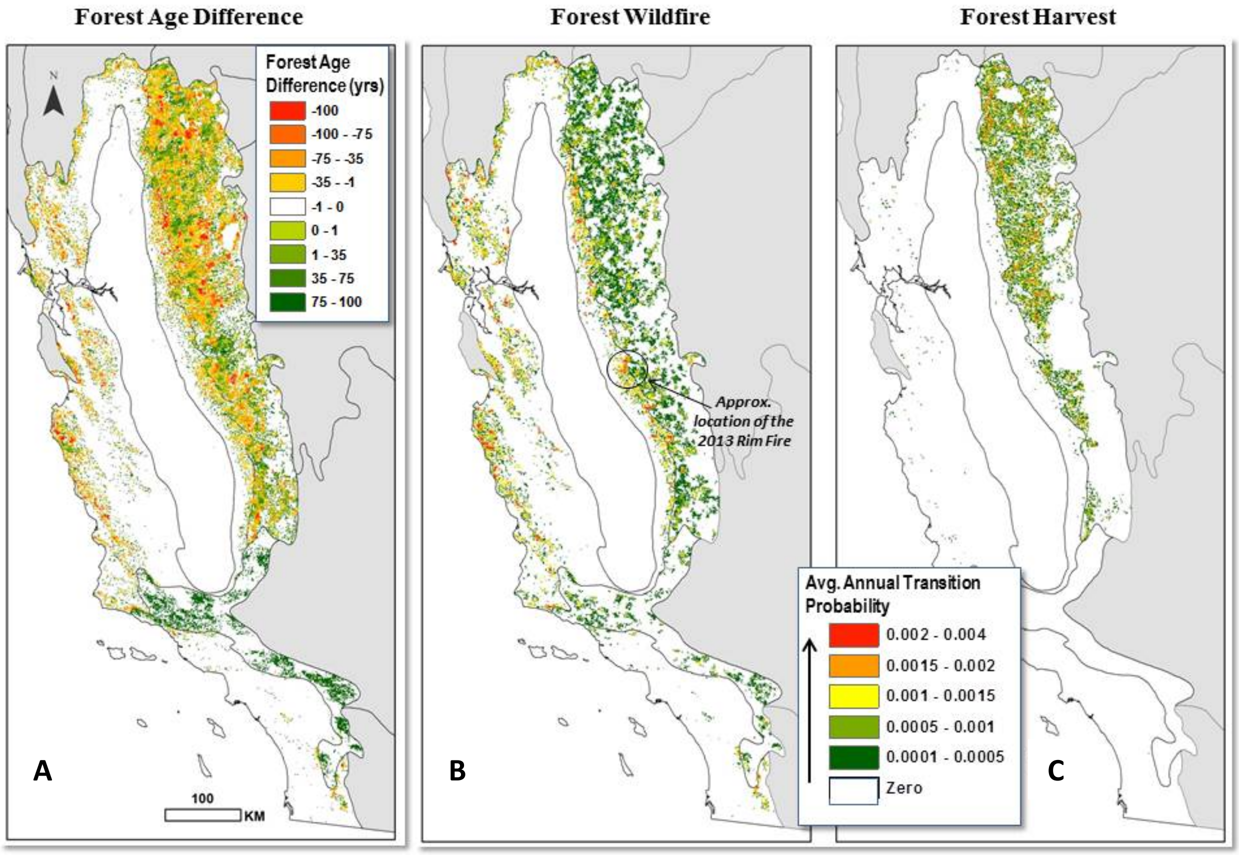

Figure 8. Forest Disturbance between the year 2000 and the year 2100. A shows the net difference in Forest Age from the initial age (2000) to the ending age (2100). Green cells represent forest cover that had a net increase in age—less disturbance; and areas in red depict forest cover that had a net decrease in age—heavy disturbance. B shows the average annual transition probability for Forest Wildfire, and C shows the average annual transition probability for Forest Harvest.

To test the significance of the spatial multipliers, we compared results from LUCAS with and without the use of spatial multipliers as input parameters. A simple sensitivity analysis was conducted to evaluate the range of uncertainty in model results by using different inputs. The three maps in Figure 9 show the evaluation of the probability of forest harvest around Lassen Volcanic National Park in northern California. Figure 9A represents the forest harvest spatial multiplier, with higher probabilities of harvest in red and lower probabilities in green. Black cells in Figure 9A represent GAP classes 1 and 2, which are protected areas prohibited from transitioning. The forest harvest spatial multiplier is described in detail in Table 3. Figures 9B–C show a 100-year simulation with 20 Monte Carlo simulations of forest harvest with and without the influence of the spatial multipliers as input parameters. Black cells in Figures 9B–C simply represent non-harvested cells. In Figure 9B, forest harvest occurs throughout the park in a random transition frequency pattern over the 100-year period. The higher frequency cells are not clustered and appear more evenly dispersed; therefore, it becomes difficult to draw conclusions about spatial and temporal harvest intensity. The random pattern reflects the high uncertainty of where forest harvest occurs. The high values in Figure

9C correspond to the general pattern of high frequency hotspots in the spatial multiplier map (Figure 9A). The influence of the spatial multiplier map on cell placement also restricts forest cells from harvest in Lassen Volcanic National Park.

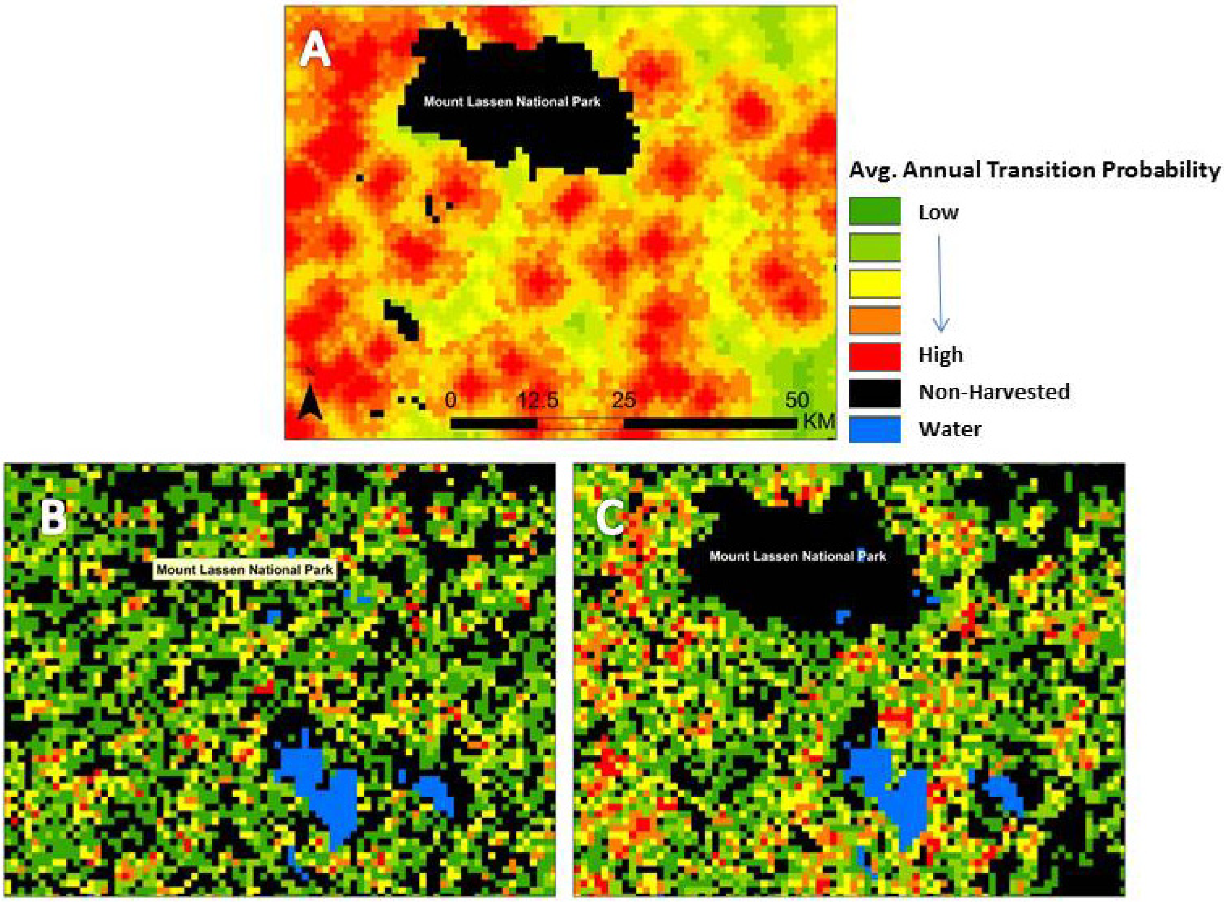

Figure 9. Annual Average Transition Probability of Forest Harvest around Mt. Lassen Volcanic National Park, CA., from 2000–2100 with 20 Monte Carlo simulations. A represents the spatial multiplier created for forest harvest transitions; B represents the 100-year simulation with spatial multipliers as input parameters; C represents the 100- year simulation without spatial multipliers as input parameters.

The transition group, Agriculture to Developed is one of the most common transitions in the Mediterranean California study area. Figures 10A–B show the average annual transition probability of agriculture transitioning to developed within the 100-year, 20 Monte Carlo simulation period for scenario A1B. Both maps reflect the same input area target (i.e. the projected amount of agriculture to developed annually); however, Figure 10A was run without spatial multipliers as input parameters, and Figure 10B was run with spatial multiplier input parameters. One of the obvious differences between the two maps is the higher uncertainty shown in Figure 10A, which illustrates a random diffusion of transition occurrence across the landscape, emphasizing the impacts of not using spatial multipliers as parameters. Figure 10B, influenced by spatial multipliers, constrains transitions to the developed-agriculture interface.

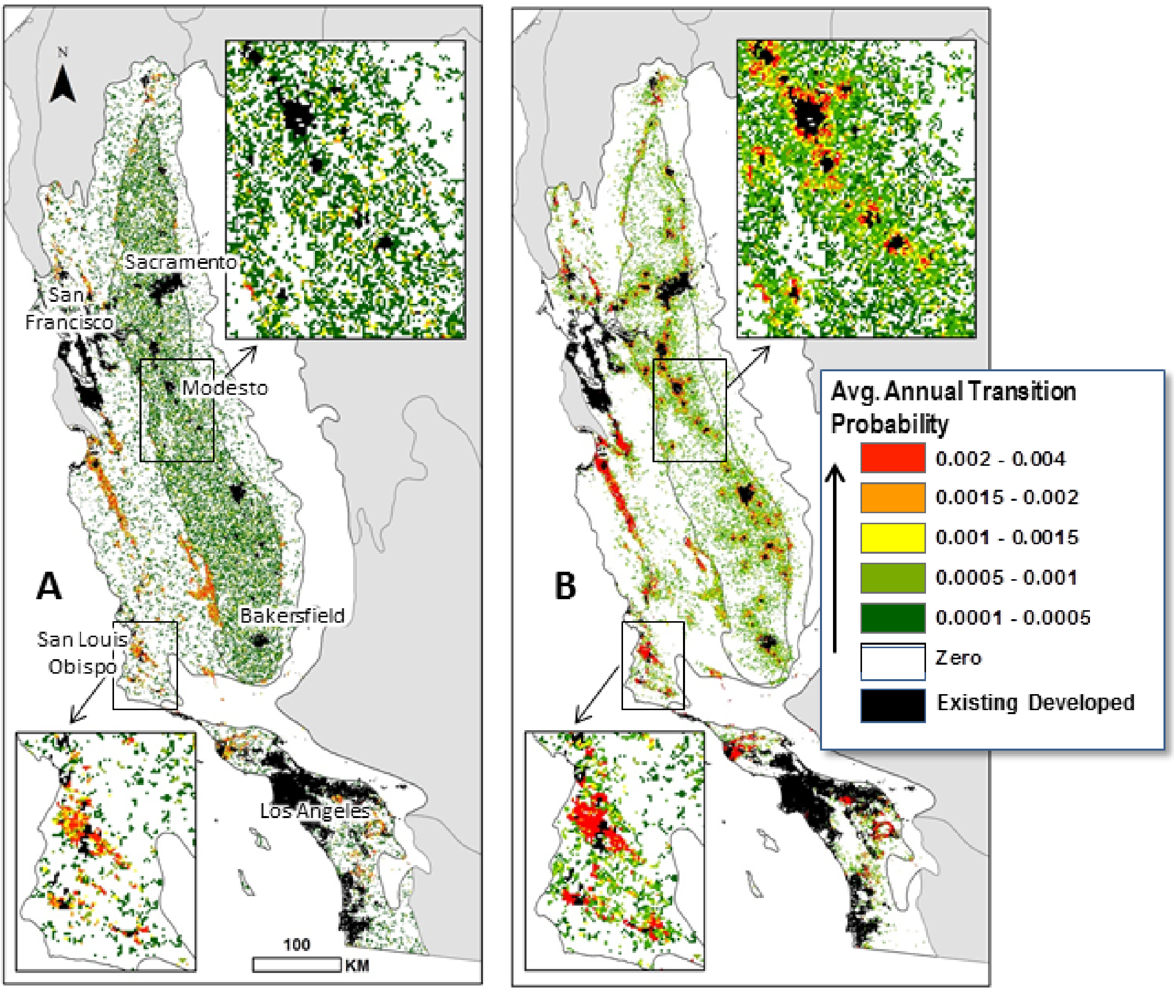

Figure 10. Annual Average Transition Probability—Agriculture to Developed from 2000–2100 with 20 Monte Carlo simulations. A represents the 100-year simulation without using spatial multipliers as spatial input parameters and B illustrates the 100-year simulation with spatial multipliers as spatial input parameters.

The Central Valley has a very large pool of available agriculture cells eligible to transition to a developed land use. In Figure 10A, there were no spatial multipliers used, resulting in a transition pattern that is highly scattered throughout the Central Valley. Most values (in green) show a very low transition frequency, and there are no clusters of high probability due to the lack of external indicators and the randomness inherent in a state-and-transition-simulation model. Figure 10A does not convey a logical pattern based on existing developed land uses and would not be a useful projection for additional spatial analysis. In contrast, Figure 10B clearly shows a logical pattern with a higher frequency of agriculture conversion occurring near existing developed cells and preferentially to areas of higher population density. Along the I-99 corridor lie established cities such as Sacramento, Stockton, Modesto and Bakersfield. Logically, the spread of development would be constrained to these areas, with probabilities declining with increased distance.

In the Oak Woodlands, the difference in pattern between Figure 10A and B is much less extreme. The overall pool of available agriculture cells is considerably smaller than in the Central Valley, and generally confined to coastal valleys such as Sonoma/Napa, Salinas, and San Luis Obispo. Due to the fact that agriculture land use has a much smaller existing footprint in this ecoregion, yet the Oak Woodlands has a high level of urbanization in the A1B scenario, the model frequently selects the same agriculture cells to transition to developed land use. Consequently, the results in Figure 10A show clusters of higher average annual change probability in the Oak Woodlands relative to the dispersed pattern in the Central Valley. In Figure 10B, given the input of spatial multipliers, and the A1B area targets, the Oak Woodlands has an even tighter pattern around existing developed. The ability of LUCAS to visually and statistically summarize 100 years of transitions within one spatial map helps the end user test the sensitivity and uncertainty of different spatially-explicit variables.

4. Conclusion

The Land Use and Carbon Scenario Simulator, referred to as LUCAS, was developed to quantify the impacts of future LULC change on carbon storage and flux in the conterminous U.S. LUCAS uses a state-and-transition simulation model fully integrated with a terrestrial carbon stock and flow model. In this manuscript we have presented methods and results to describe the spatially-explicit input parameters used to develop LUCAS. Many spatial models of LULC change are identified in the introduction highlighting the strengths and limitations of different approaches in relation to input parameters, leading to the rationale of why we chose a state-and-transition simulation model as the foundation of this research. The overarching goal is to create an adaptable modeling environment that can implement data from various disciplines, at any scale, local-to-global. Transferability and scalability is an important aspect to this research and one we continue to explore.

Defining the initial landscape conditions is a fundamental first step to spatially parameterizing a state-and-transition simulation model, and a step that should be well thought out. By synthesizing multiple national datasets through the harmonization process, we were able to address a large problem in the mapping and modeling community. There are always inconsistencies with the way LULC is defined, monitored and spatially formatted. Most of these issues stem from multiple disciplines with divergent objectives and applications. We used a unique harmonization method to converge a wide range of temporal, spatial and classification schemes, resulting in consistent definitions and rules. The original harmonization map was produced at a 30-m resolution with 21 LULC classes, which was then recoded into eight broad-scale LULC classes and resampled to a 1-km resolution. The importance of this process ensures that the initial conditions map in LUCAS can be scalable from 30-m to 1-km resolution based on a set methodology.

Strata, defined by ecoregions, allow for regional comparisons of future projections with an archive of historical trends for reference conditions. Strata provide an important linkage to connect area targets from downscaled global scenarios to ecoregions of the United States. LUCAS also uses strata to analyze region-specific drivers of change. Forest age is one of the most important parameters used to track and estimate disturbance from anthropogenic land-use change, like harvest, and natural disturbances, like wildfire. Both strata and forest age have attribute options in the model, (e.g. time-since-transition, patch characteristics) that provide non-stationarity to some of the input processes.

Spatial multipliers were developed in a simple way by applying a weighted distance between specific transition groups, reinforcing the fundamental relationship between spatial interaction and distance-decay. If a modeler had limited source data and wanted to create spatial multipliers for a state-and-transition simulation modeling project, the individual state classes and the distance relationships between them could assist the development of a simple set of spatial multipliers. Lastly, as part of the initial conditions, we introduce the spatially-explicit carbon stock density maps, and the associated source data. The unique carbon parameters (e.g. flux rates between pools, automatic and event-based flows etc.) used in the stock and flow model, are discussed in detail in Sleeter et al. [22].

When modeling future scenarios with exogenous data sources, the spatial accuracy of change projections are difficult to validate due to the fact that the modeled time period is in the future, and the set of scenario assumptions driving the change may not be grounded with current trends. We recognize that a formal validation of the spatial accuracy of LUCAS is needed to substantiate its relevance in the science community. While development of the input parameters has been our primary focus, we plan to improve accuracy measures to promote visibility. It is critical to model the uncertainty around projections of future outcomes. Monte Carlo iterations increase the model's capacity to spatially measure uncertainty around projections of future outcomes. Our results illustrate the uncertainty of specific input parameters with a sensitivity analysis by summarizing the annual average transition probabilities over the 100-year simulation period. We demonstrate the positive influence of the spatial multipliers for spatial allocation and landscape patterns. Our research provides a well-defined set of spatially-explicit input parameters combined with freely available modeling software (ST-Sim) as a methodology for producing future projections of LULC and carbon dynamics for the conterminous U.S.

Conflict of Interest

Authors declare no conflicts of interest in this paper

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

1. Vitousek PM, Mooney HA, Lubchenco J, et al. (1997) Human domination of earth's ecosystems. *Science* 277: 494-499.
2. DeFries RS, Foley JA, Asner GP (2004) Land-use choices: Balancing human needs and ecosystem function. *Front Ecol Environ* 2: 249-257.
3. IPCC (2000) In: Watson, R.T., Noble, I.R., Bolin, B., Ravindranath, N.H., Verardo, D.J., Dokken, D.J. (Eds.), Land Use, Land-Use Change, and Forestry. Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 377.
4. Foley JA, Defries R, Asner GP, et al. (2005) Global consequences of land use. *Science* 309: 570-574.
5. Pielke RA (2005) Land use and climate change. *Science* 310: 1625-1626.
6. Alcamo J (2008) Searching for the future of land: scenarios from the local to global scale. In: Alcamo J (Ed.), Environmental Futures: The Practice of Environmental Scenario Analysis. Elsevier, Amsterdam, The Netherlands.

7. USGCRP, The National Global Change Research Plan 2012-2021: A Strategic Plan for the U.S. Global Change Research Program. 132 pp. The U.S. Global Change Research Program Washington, D.C., 2012, Available from: http://downloads.globalchange.gov/strategic-plan/2012/usgcrp-strategic-plan-2012.pdf
8. Strengers B, Leemans R, Eickhout B, et al. (2004) The land-use projections and resulting emissions in the IPCC SRES scenarios as simulated by the IMAGE 2.2 model. *GeoJournal* 61: 381-393.
9. National Research Council (2014) Advancing Land Change Modeling: Opportunities and Research Requirements. Washington, DC: The National Academies Press.
10. Soares-Filho BS, Nepstad DC, Curran LM, et al. (2006) Modeling conservation in the Amazon basin. *Nature* 440: 520-523.
11. Eastman JR (2007) The Land Change Modeler, a software extension for ArcGIS. Worcester, Mass.: Clark University.
12. Matthews RB, Gilbert NG, Roach A, et al. (2007) Agent-based land-use models: A review of applications. *Landscape Ecol* 22: 1447-1459.
13. Acosta-Michlik L, Espaldon V (2008) Assessing vulnerability of selected farming communities in the Philippines based on behavioural model of agent's adaptation to global environmental change. *Glob Environ Chang* 18: 554-563.
14. Brown DG, Page SE, Riolo R, et al. (2004) Agent based and analytical modeling to evaluate the effectiveness of greenbelts. *Environ Modell Softw* 19: 1097-1109.
15. Verburg PH (2002) Land use change modelling at the regional scale: The CLUE-S model. *Environ Manage* 30: 391-405.
16. Burnham BO (1973) Markov intertemporal land use simulation model. *South J Agr Econ* 5: 253-258.
17. Baker WL (1989) A review of models of landscape change. *Landscape Ecol* 2: 111-133.
18. Turner M (1987) Spatial simulation of landscape changes in Georgia: A comparison of 3 transition models. *Landscape Ecol* 1: 29-36.
19. Daniels CJ, Frid L (2011) Predicting landscape vegetation dynamics using state-and-transition simulation models, In *Proceedings of the First Landscape State-and-Transition Simulation Modeling Conference, Portland, OR, USA, 14-16 June 2011*; Kerns, B.K., Shlisky, A.J., Daniel, C.J., Eds.; U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station: Portland, OR, USA, 2012; pp.5-22.
20. Sleeter BM, Sohl T, Bouchard MA, et al. (2012) Scenarios of land use and land cover change in the conterminous U.S.: Utilizing the special report on emission scenarios at ecoregional scales. *Glob Environ Chang* 22: 896-914.
21. Omernik JM (1987) Ecoregions of the conterminous U.S. *Ann Assoc Am Geogr* 77: 118-125.
22. Sleeter BM, Liu J, Daniel C, et al. (2015) An integrated approach to modeling changes in land use, land cover, and disturbance and their impact on ecosystem carbon dynamics: a case study in the Sierra Nevada Mountains of California. *AIMS Environ Sci* 2: 577-606.
23. Nakicenovic N, Swart R (Eds.) (2000) *IPCC Special Report on Emission Scenarios*; Cambridge University Press; Cambridge, UK; p. 570.
24. Sleeter BM, Sohl T, Loveland T, et al. (2013) Land-cover change in the conterminous United States from 1973 to 2000. *Glob Environ Chang* 23: 733-748.

25. EPA (U.S. Environmental Protection Agency), Primary distinguishing characteristics of Level III ecoregions of the continental United States. Environmental Protection Agency, 1999. Available from: http://www.epa.gov/wed/pages/ecoregions/level_iii.htm
26. Gallant AL, Loveland T, Sohl T, et al. (2004) Using an ecoregion framework to analyze land-cover and land-use dynamics. *Environ Manage* 34: S89-S110.
27. Soulard CE and Acevedo W, Multi-temporal harmonization of independent land-use/land-cover datasets for the conterminous U.S. American Geophysical Union, 2013. Available from: http://adsabs.harvard.edu/abs/2013AGUFM.B41E0448S
28. McConnell WJ, Moran EF (Eds.) (2001) Meeting in the middle: the challenge of meso-level integration. An international workshop on the harmonization of land-use and land-cover classification. LUCC Report Series No. 5. Anthropological Center for Training and Research on Global Environmental Change - Indiana University and LUCC International Project Office, Louvain-la-Neuve.
29. Jansen LJM, Groom GB, Carrai G (2008) Land-cover harmonisation and semantic similarity: some methodological issues. *J Land Use Sci* 3:131-160.
30. Homer C, Dewitz J, Fry J, et al. (2007) Completion of the 2001 national land cover database for the conterminous U.S. *Photogramm Eng Remote Sens* 73: 337-341.
31. Vogelmann JE, Howard SM, Yang L, et al. (2001) Completion of the 1990's national land cover data set for the conterminous U.S. *Photogramm Eng Remote Sens* 67: 650-662.
32. Fry J, Xian G, Jin S, et al. (2011) Completion of the 2006 national land cover database for the conterminous U.S. *Photogramm Eng Remote Sens* 77:858-864.
33. Jin S, Yang L, Danielson P (2013) A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sens Environ* 132: 159-175.
34. LANDFIRE: LANDFIRE Existing Vegetation Type layer. U.S. Department of Interior, Geological Survey, 2013 June last update. Available from: http://landfire.cr.usgs.gov/viewer/ (accessed 1 October 2012)
35. U.S Department of Agriculture, National Agricultural Statistics Service (2011) Cropland Data Layer. Available from http://nassgeodata.gmu.edu/CropScape/ (accessed on 1 October 2012).
36. U.S. Geological Survey, Gap Analysis Program (GAP). May 2011. National Land Cover, Version 2
37. Hansen MC, Potapov PV, Moore R, et al. (2013) High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342: 850-853.
38. LANDFIRE: LANDFIRE Disturbance (1999-2010). U.S. Department of Interior, Geological Survey, 2013 June last update. Available from: http://landfire.cr.usgs.gov/viewer/ (accessed on 1 October 2012).
39. Roy DP, Ju J, Kline K, et al. (2010) Web-enabled Landsat data (WELD): Landsat ETM+ composited mosaics of the conterminous U.S. *Remote Sens Environ* 114: 35-49.
40. Eidenshink J, Schwind B, Brewer K, et al. (2007) A project for monitoring trends in burn severity. *Fire Ecol Spec Issue* 3: 3-21.
41. Anderson JR, Hardy E, Roach JT, et al. (1976) A land use and land cover classification scheme for use with remote sensor data. U.S. Geological Survey, Reston, VA. USA, Professional Paper 964.
42. Pan Y, Chen JM, Birdsey R, et al. (2011) Age structure and disturbance legacy of North American forests. *Biogeosciences* 8: 715-732.

43. Masek JG, Huang R, Wolfe R, et al. (2008) North American forest disturbance mapped from a decadal Landsat record. *Remote Sens Environ* 112: 2914- 2926.
44. Huang C, Goward SN, Masek JG, et al. (2010) An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens Environ* 114: 183-198.
45. Tobler W (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46: 234-240.
46. U.S. Geological Survey GAP. Protected Areas Database of the United States (PAD-US), version 1.3 Combined Feature Class. 2012. Available from: http://gapanalysis.usgs.gov/padus/ (accessed on 13 November 2013).
47. Natural Resources Conservation Service. Soil Survey Geographic (SSURGO) Database. U.S. Department of Agriculture, 2011. Available from: http://soildatamart.nrcs.usda.gov (accessed on 22 November 2013).
48. Sleeter R, Gould M (2007) Geographic Information System Software to Remodel Population Data Using Dasymetric Mapping Methods; U.S. Geological Survey *Techniques and Methods* 11-C2; U.S. Geological Survey, Reston, VA. USA, p. 15.
49. Wilson BT, Woodall CW, DM Griffith (2013) Imputing forest carbon stock estimates from inventory plots to a nationally continuous coverage. *Carbon Balance Manag* 8: 1-15.
50. IMAGE team (2001) The IMAGE 2.2 implementation of the SRES scenarios: climate change scenarios resulting from runs with several GCMs. RIVM CD-ROM Publication 481508019, National Institute of Public Health and the Environment, Bilthoven.
51. Van Vuuren D, Edmonds JA, Kainuma M, et al. (2011) The Representative Concentration Pathways: An Overview. *Climatic Change* 109: 5-31.
52. Wilson TS, Sleeter BM, Sleeter RR, et al. (2014) Land use threats and protected areas: a scenario-based, landscape level approach. *Land* 3: 362-389.
53. Sleeter BM (2008) Late 20th century land change in the Central California Valley Ecoregion. *California Geographer* 48:27-60