Land Use Policy 41 (2014) 438–452

Contents lists available at ScienceDirect

Land Use Policy

journal homepage: www.elsevier.com/locate/landusepol

# Modeling residential development in California from 2000 to 2050: Integrating wildfire risk, wildland and agricultural encroachment

Michael L. Mann[a,b,*], Peter Berck[b], Max A. Moritz[c], Enric Batllori[c], James G. Baldwin[d], Conor K. Gately[d], D. Richard Cameron[e]

[a] The George Washington University, Geography Department, Washington, District of Columbia, United States
[b] UC Berkeley, Agricultural & Resource Economics, Berkeley, CA, United States
[c] UC Berkeley, Environmental Science, Policy, and Management, Berkeley, CA, United States
[d] Boston University, Earth & Environment, Boston, MA, United States
[e] The Nature Conservancy, San Francisco, CA, United States

ARTICLE INFO

*Article history:*
Received 29 October 2013
Received in revised form 13 March 2014
Accepted 25 June 2014

*Keywords:*
Housing
Spatial econometrics
Conservation
Wildfire
California
Wildland Urban Interface

ABSTRACT

Between 1940 and 2000, nearly 10 million housing units were constructed throughout California. This increased interaction between human and natural communities creates a number of significant socio-ecological challenges. Here we present a novel spatially explicit model that allows better characterization of the extent and intensity of future housing settlements using three development scenarios between 2000 and 2050. We estimate that California's exurban land classes will replace nearly 12 million acres of wild and agricultural lands. This will increase threats to ecosystems and those presented by wildfire, as the number of houses in 'very high' wildfire severity zones increases by nearly 1 million.



## Introduction

Between 1940 and 2000 nearly 10 million housing units were constructed throughout California (US Census Bureau, 1990, 2000a,b). Although urban growth is pronounced in most of California's urban centers, its impact is far outweighed by the acreage disturbed by low-density exurban and rural development. Almost 80% of the acreage used in recent development over the US has been outside of urban areas (Heimlich and Anderson, 2001; Newburn and Berck, 2006), as individuals seek low cost housing and more rural living amenities (Crump, 2003). These low density settlements affect increasingly large swaths of land, with nearly 57% of recent development occurring on lots of 10 acres or larger (Heimlich and Anderson, 2001; Newburn and Berck, 2006). We estimate the loss of sparsely settled and agricultural land through the expansion of exurban and rural communities between 2000 and 2050. These future exurban and rural developments will be encumbered by the complex and consequential interactions among settlements, climates, and the ecosystems. Here we examine the interaction between the these developments and the fire driven ecology where they are located.

The interaction of human and natural communities creates a number of significant environmental challenges. These challenges include climate change, loss of wildlife habitat and ecosystem fragmentation, introduction of invasive species, threats to endangered and sensitive species, as well as water and air pollution issues (Alavalapati et al., 2005; Hammer et al., 2004; Radeloff et al., 2005). All this, which is further exacerbated by the appeal of development in areas with high ecological value (McGranahan, 1999), can have significant consequences for ecosystems services. Therefore, the persistence and growth of exurban settlements creates complex patterns under which species and habitats, depending on their capacity for adaptation and resilience, ebb and flow with the course of human development (Hansen et al., 2005). Along with threats to the natural environment, the increasing proximity to wildlands (i.e., the expansion of Wildland–Urban Interface, WUI) brings risks to human communities as well. Low housing densities and increased exposure to natural lands can make exurban communities both more likely to experience natural disasters and makes many of their effects more costly (Calkin et al., 2005; Gebert et al., 2007a,b; Gude et al., 2008a,b; Liang et al., 2008).

After a respite induced by the December 2007 to June 2009 credit crunch and recession, housing development in California

* Corresponding author at: The George Washington University, Geography Department, Washington, District of Columbia, United States. Tel.: +1 202 994 4567.
*E-mail address:* mmann1123@gwu.edu (M.L. Mann).

http://dx.doi.org/10.1016/j.landusepol.2014.06.020
0264-8377/

has resumed along with its complex mix of economic, social and ecological impacts. In 2013, housing starts in the state expanded to 2.3 times their 2009 levels, nearly 70% of the 1990 to 2012 average (CBIA, 2013a, 2013b). The expansion of human settlements is of primary concern as it shapes a series of irreversible spatial and temporal patterns on the landscape. These patterns determine both the direct and indirect costs and benefits associated with development.

The development of new housing is primarily a response to the demographic pressures of population, economic growth, and other incentive structures that underpin the process of household formation (Heimlich and Anderson, 2001). The specific location of new housing, however, is driven by more complex factors including autoregressive factors, spatial spillovers, terrain, climate, access to employment and services, transportation costs, aesthetics, weather and cultural and environmental amenities, amongst others. Researchers have made efforts to forecast housing development using a variety of methods.

Cellular automata are capable of replicating the complex yet highly structured spatial patterns based on a set of deterministic or probabilistic rules that determine the state of a cell based on the states of its neighbors. Models like SLEUTH have been successfully employed to model land-use change and urban growth for discrete land-use or urban development classes (see Dietzel and Clarke, 2007; Irwin and Geoghegan, 2001). Although their usefulness is widely acknowledged, these models have been criticized for the complex and arbitrary nature of calibration, and the inability to attribute simulated patterns to particular drivers such as changes in population, employment etc. (Dietzel and Clarke, 2007; Irwin and Geoghegan, 2001; Jantz et al., 2010). Another set of models, dynamic simulations, model the interactions between the drivers of a land-use system. This is accomplished by creating a set of differential equations that portrays *a priori* a simplified representation of the complex states and interactions between system components (Lambin et al., 2004).

Empirical or statistical models of land-use and land-use change focused on modeling deforestation (Mann et al., 2010, 2014; Nelson et al., 2004; Pfaff, 1999) have also been applied to urban land-uses type and housing density (Landis and Zhang, 1998; Newburn and Berck, 2006). Broadly, this class of model determines the likelihood of conversion based on exogenous information on initial land-use, site characteristics, accessibility, community characteristics, and policy factors. The majority of these models are implemented as discrete choice models where land-use is classified as residential, commercial, or industrial, or they classify density in broad or narrow density categories. Many of these applications have endogeniety problems. The problems are caused by inclusion of accessibility measures based on transportation networks that are jointly determined with land-use choices, especially over longer time periods (Chomitz and Gray, 1996; Irwin and Geoghegan, 2001; Jacoby, 2000). Due to the complexity of implementation, these models are also typically limited to local or semi-regional case studies (Theobald, 2005). Additionally, spatially explicit discrete-choice regression models face difficulties in estimation due to the complex likelihood functions and other numerical challenges (Holloway et al., 2007). However recent advances in Bayesian techniques have rendered these models more computationally tractable (Holloway et al., 2007).

Hybrid models use a handful methodologies allowing each to interact or drive the behavior of another module (Berry et al., 1996; Pijanowski et al., 2002; Veldkamp and Fresco, 1996; Walker et al., 2007). Due to the flexibility and modular nature of these models, they are often adapted to represent complex interactions between systems, such as policy tools, socio-economic drivers, or impacts on biodiversity, and real estate values. The Spatially Explicit Regional Growth Model (SERGoM) utilizes two core modules to forecast housing density classes (Theobald, 2005). The first module estimates the demand for new housing based on county-level population and a county specific housing to population ratio. The second module allocates housing based on a set of weights developed from the local growth rates over two periods and a measure of travel time to the nearest urban core. Weights are then adjusted to improve accuracy for the observed density classes in 1990 and 2000. This approach has been adopted by the EPA for the Integrated Climate and Land-Use Scenarios (ICLUS) model (EPA, 2010).

Here we present a novel spatially explicit model that allows us to better characterize the extent and intensity of housing settlements for California up to 2050. Our spatial panel econometric approach stands out from existing models due to the ease of implementation and attribution, estimation over a long historic record, the lack of reliance on transportation networks and other endogenous variables as a basis for land-use change, implementation of spatial spillovers, and explicit consideration of housing density.

## Materials and methods

The following section lays out the methodology used to predict housing densities for the state of California from the year 2000 to 2050. A spatial panel regression, with robust standard errors, is used to estimate the effect of spatial and temporal lags as well as exogenous variables such as climate on the spatial distribution of housing density in each period. County-level demographic forecasts drive the total supply of housing for future periods. To provide a range of estimates depicting potential patterns of housing development and therefore of associated interactions with climate and ecosystems, three development scenarios are used: business as usual, greater urban development, and further rural development.

### *Response variable*

Our model estimates housing density measured as housing units per acre. Historical housing density from 1940 to 2000 is derived from the Census Bureau's split census block group data (US Census U.S. C. Bureau, 2000a,b). Block groups represent the aggregation of a cluster of census blocks. Block groups typically represent between 600 and 3000 people with a target size of 1500 people (US Census U.S.C. Bureau, 2012b). Split block groups (SBG) add additional accuracy by breaking groups by the boundary of other tabulation entities including Native American areas, voting districts, or urban boundaries. SBGs therefore provide a much more accurate representation of the housing stock. After the removal of undevelopable land (see below) all SBGs have a median size of 115 acres, with a first and third quartile of 58.8 and 280.1, respectively. Urban and rural classified SBGs have a median size of 96.2 and 524 acres, respectively, whereas in very sparsely populated or unpopulated areas SBGs can be as large as 593,000 acres.

Retrospective estimates of housing counts are provided by data from the census long form, which includes tabulations of 'year housing structure built' (US Census U.C. Bureau, 2007). A housing unit may include houses, apartments, mobile homes either occupied or vacant (Radeloff et al., 2005), and year housing structure built "refer[s] to when the building was first constructed, not when it was remodeled, added to, or converted" (US Census U.C. Bureau, 2012a). This data provides the retrospective data on housing counts at the SBG level. Following the approach of Hammer et al. (2007), the houses built in each successive decade are added to create an estimate of the number of houses present in each decade from 1940 to the year 2000, where year 2000 SBG level estimates match actual housing counts.

Although census data is currently available for 2010, this part of the census was reassigned to collection under the American Community Survey which samples only 1 in 40 households versus 1

**Table 1**
Housing net-density and wildland–urban classifications.

| Density class | Definition | Lower bound houses/Acre | Urban Status |
|---|---|---|---|
| Very-high | ≥4 units per acre | ≥4 | Urban |
| High | 1–4 units per acre | 1 | |
| Medium | <1 per acre to 1 per 2.4 acres | 0.41 | Rural |
| Low | <1 per 2.4 acres to 1 per 40 acres | 0.025 | |
| Sparse | <1 per 40 acres | <0.025 | |

in 6 for the 2000 census long form (Gardner et al., 2012). Therefore we considered the 2010 census to have a margin of error too high for rural housing counts. Two factors lead to underreporting of housing stocks in previous decades. Like all survey questions, the census long form suffers from response errors and additionally, retrospective estimates do not include houses that have been demolished, destroyed or that are uninhabited. Underestimation of housing counts were estimated to be zero for year 2000 and 6%, 11%, 14%, 28%, 46% for decades 1990, 1980, 1970, 1960, 1950, respectively. To correct for this source of error SBG housing counts were adjusted to match county level 100% sampling housing counts (US Census Bureau, 1990, 2000a,b). This adjustment is known as allotment, as the fraction of total estimated housing counts by county (Table H34) are used to allocate the actual county housing count between SBGs. To further refine the accuracy of the data, all undevelopable lands were removed and housing units distributed to the developable portions of the original SBGs. Density classifications are considered 'net-density' as all undevelopable lands are excluded from the dataset. A description of those excluded areas and its implications are described later in the Modeling Framework section. Our dependent variable is net density and it is calculated as the number of housing units per acre in a SBG.

To ease the discussion of results housing density classifications have been adapted from those from Newburn and Berck (2006) and Theobald (2001) and are presented in Table 1. One housing unit per acre is defined as the cutoff for urban areas because development above this threshold typically requires the installation of city water and sewer systems (Newburn and Berck, 2006).

### *Predictor variables*

The exogenous predictors of housing density tested in this study include a set of environmental and geographic factors. These include: (a) a one-period temporally lagged density variable (*Den(t − 1)*), (b) climate norms (1971–2000) of mean annual high (*Maxtemp*) and low temperatures in Celsius degrees (*Mintemp*), and mean annual rainfall in millimeters (*Aveppt*) (Flint and Flint, 2012), (c) linear distance to Native Indian lands and National Parks (Cal Fire, 2011) and (d) county and decade fixed effects (*CountyFE* and *TimeFE*, respectively). We include climatic temperature and precipitation norms to account for the highly variable and localized climatic conditions of California. For California, we believe climate norms make a more effective amenity variable than proximity to the ocean or freshwater bodies. Although proximity to the Pacific is a draw for much of Southern California, it is equally undesirable for many places in Northern California, due to low year-round temperatures and fog. Similarly, fresh water bodies of California run the gamut from the clear waters of Lake Tahoe, to the floodplains around South Bay, to the toxic Salton Sea. Climatic norms therefore provide one of the few ubiquitous natural amenities. For rural areas, we test how proximity to national parks influences development patterns. We expect that results will be mixed. Although proximity to parks may indicate high levels of natural amenities, for most of the state's parks, it generally also entails low levels of accessibility. We also test the influence of proximity to Native Indian settlements to account for settlement patterns more easily explained by historical events and politics, than by site desirability.

### *Modeling framework*

We use panel data to model housing densities for individual SBGs over time. The use of spatial panel data in this study helps to alleviate two key problems, unobserved spatial and temporal dynamics, and homogeneity (lack of variance). If pooled together, the integration of a statewide set of SBGs ($N$ = 29,379) over the 1950–2000 period ($T$ = 6) allows for nearly 176,256 observations, which amounts to unprecedented degree of observed variance over both space and time. This, however, is limited by the use of county-group specific estimation. County-group regressions are desirable for two reasons: (1) it reduces computational complexity of estimation and (2) it better represents the regional heterogeneity of explanatory variables. Spatial dependence is controlled for using a spatially weighted dependent variable (with coefficient *Rho*). Temporal autocorrelation is modeled using temporal lags (*Den(t − 1)*), time fixed effects with time dummies for each period (*Time_FE*) and county fixed effects (*County_FE*).

Spatial dependence is a special case of cross-sectional dependence that occurs due to similarities between neighboring regions, and creates a situation whereby data can no longer be considered independently generated (Anselin, 1999; Anselin et al., 2008). For this reason a spatially lagged fixed effect panel model is developed in the following form (Elhorst, 2010):

$$y_{it} = \rho \sum_{j=1}^{N} W_{ij} y_{jt} + \alpha_c + x_{it}\beta + y_{t-p}\phi + \gamma_t + \varepsilon_{it} \tag{1}$$

where ρ is the coefficient describing spatial dependence as modeled by $\sum_{j=1}^{N} W_{ij} y_{jt}$, a spatially lagged measure of $y$ for each individual $i$ and its 'neighbors' $j$ as defined by $W_{ij}$ in time-period $t$ (weighted average of neighboring SBGs net density). The specification of the spatial weights matrix $W_{ij}$ is row-standardized and is described in more detail below. $W_{ij}$ is assumed to remain constant over time (Anselin et al., 2008). $\alpha_c$ is the county-level fixed effect constants, and controls for unobserved characteristics of each county $c$, $\gamma_t$ is a time fixed effect that controls for unobserved characteristics of each decade, $x_{it}\beta$ is a $K \times 1$ vector of regression coefficients for descriptive variables $x_i$ at period $t$, where $K$ is the number of descriptive variables. $x_{it}$ include possible exogenous determinants of housing density as described in the *Predictor Variables* section. $y_{t-p}\phi$ is a $P \times 1$ vector of regression coefficients for temporally lagged values of $y$ for period $t - p$, where the temporal lag length ($p$) for choice for $y$ is determined by minimizing the regression Akaike Information Criterion (AIC). $\varepsilon_{it}$ is a $N \times T$ matrix of disturbances.

Controlling for both time and county fixed effects allows us to disentangle the effect of spatial dependence from that of spatial heterogeneity and of omitted variables. Neighborhood specifications of four general types are tested. First, simple distance band measured in meters (*Dist*). Second, $K$-nearest neighbors where each polygon has exactly $K$ neighbors (*Kneigh*). Third, polygon adjacency and its higher orders, where $P$ = 1 is a polygon and all adjacent neighbors, and $P$ = 2 is union of neighbors and their adjacent neighbors (*Poly*). Finally, neighbor distance plus polygon adjacency, which corresponds to the union of the polygon adjacency $P$ = 1 and a neighborhood distances band of m (*PolyDist*). To avoid problems with estimation, all SBGs with zero neighbors use a $K$-nearest neighbor to avoid zero neighbors ($K$ = 1). The neighborhood specification is found through an iterative search mechanism to reduce median squared regression error. The complete process of selecting the final neighborhood definition of *Dist* at a distance band of 20,000 meters is outlined in Appendix A.

*Estimation technique*

Considering the empirical complexity of spatial panel regression and forecasting, regressions are run on county-groups for the 1950 to 2000 decades. Groupings are largely determined by Jepson ecoregions (Hickman, 1993), where the region of largest inhabitable area assigns membership. In order to provide greater local heterogeneity, larger groups are split to obtain a final set of groups with 4.5 counties on average. Due to the complexity of estimation, the sample size used in Eq. (1) is limited to 1500 random SBGs. To avoid sampling bias, the probability of any draw is weighted by the acreage of the individual SBG. Eq. (1) is estimated separately for rural and urban communities, with an urban/rural cutoff of 1 unit per acre (Table 1). Separated estimation allows for regional heterogeneity as well as differing rates of urban and rural growth and spatial spillover. However, $W_{ij}$ accounts for all neighbors regardless of region or urban/rural membership.

*Spatial panel forecasting*

Panel forecasts were estimated as follows (Elhorst, 2010):

$$\hat{y}_{i,t+\tau} = \alpha_c + (I - \hat{\rho}W)^{-1}\hat{\beta}X_{i,t+\tau} + \hat{\gamma} + (I - \hat{\rho}W)^{-1}\hat{\mu}_i \qquad (2)$$

The $t+\tau$ forecast consists of four terms. The first term is the county fixed effect ($\alpha_c$) and the second is the effect of the forecasted independent variables $[(I - \hat{\rho}W)^{-1}\hat{\beta}X_{i,t+\tau}]$, when spatial autocorrelation is accounted for. The third term is our estimate of the future year fixed effect ($\hat{\gamma}$) that is estimated based on predicted population, see below. Finally, the last term is the spatially multiplied mean error for an individual SBG$_i$ $[(I - \hat{\rho}W)^{-1}\hat{\mu}_i]$. It is added to $\hat{y}_{i,t+\tau}$ to center each individual SBG's error around zero (Smirnov, 2010). Essentially, this introduces an individual, spatially weighted, fixed effect *i* for the prediction.

*Model assumptions*

Out-of-sample forecasts (2010–2050) are adjusted by demographic population forecasts provided by the CA Bureau of Finance (California, 2012). It is assumed that these demographic forecasts will more accurately represent population changes compared to extrapolating historic trends in housing stock. These population figures can be translated to the expected number of houses by applying a person to house ratio. Because high rural amenity counties may have high rates of second houses (Theobald, 2001), it is assumed that the ratio of persons to housing units remains constant at the observed level in year 2000 *for each county*. This allows for county-level heterogeneity in occupancy rates which should more accurately distribute future population driven by housing growth. Historical ratios are presented in Fig. 1 with each color representing a unique county (U.C. Bureau, 1995; California, 2012). This also controls for the effects of excess housing stock as might be observed during the recent housing boom. Holding the ratio of persons to homes constant slows the rate of housings starts (in overbuilt areas) in initial periods by allocating new residents first to existing houses. New houses therefore will not be allocated until the demand for new housing exceeds current stocks, as described by the expected total population and person to housing ratio.

Adjusting time fixed effects to match demographic forecasts does not significantly deteriorate prediction accuracy relative to an out-of-sample panel forecast that omits time fixed effects. To evaluate the effectiveness of adjusting time fixed effects to match demographic forecasts, we estimate the root mean squared error (RMSE) for within-sample (2000) estimates both with and without 'population-correction'. Mean RMSEs are comparable for both uncorrected and population corrected estimates, as shown in Fig. 2 violin plots, with mean values of 17.01 and 17.08, respectively.

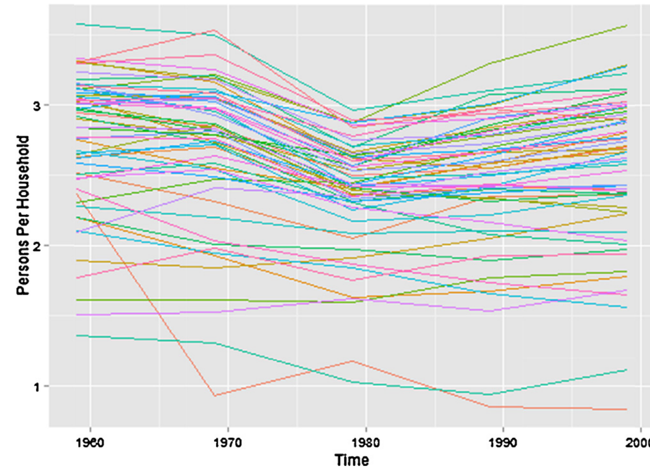


**Fig. 1.** Persons per household by county (1960–2000). Number of persons per household, where each line/color represents a county.

Although median and mean RMSE are slightly higher for population corrected estimates, these differences are reasonable considering the importance of population for optimizing future housing forecasts.

*Forecasting scenarios*

Initial housing density forecasts are assumed to be business as usual (*BAU Growth*), while redistributing houses from urban to rural or vice versa builds the *Urban Growth* and *Rural Growth* scenarios, respectively. The total number of new houses remains constant under the three scenarios. By default, in each period the number of new houses in a county follows the *BAU* partition between urban and rural areas. Increasing or reducing the initial rural partition by 25% in each decade accomplishes the *Rural Growth* and *Urban*

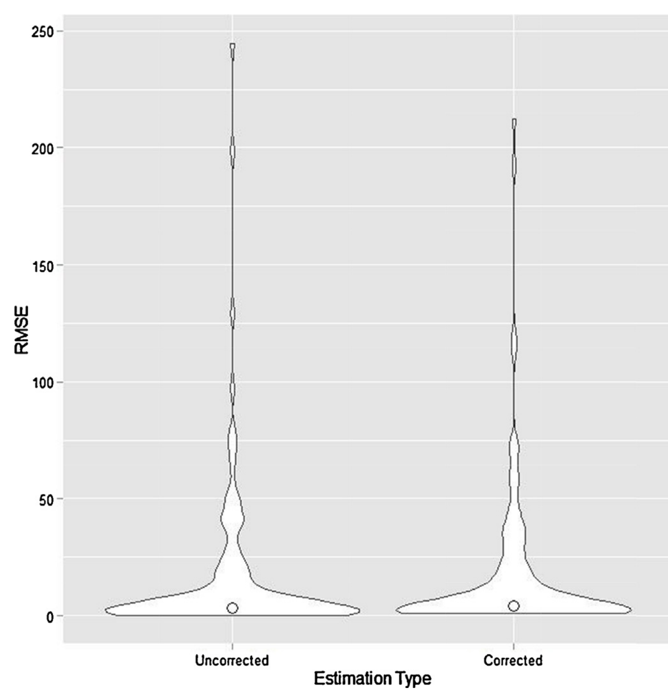


**Fig. 2.** Distribution of RMSE of (un)corrected in-sample estimates. Distribution of root mean square errors for year 2000 across all regressions, using panel regression results only (Uncorrected), and adjusting time FE to match population estimates (Corrected). Mean values are shown as white dots.

**Table 2**
Regression results from spatial panel estimation.

| Variable | Estimate (median) | *P*-value (median) | % Significant (90% level) | Urban status |
|---|---|---|---|---|
| Rho | 5.906E−2 | 0.025 | 54.545 | Urban |
| Constant | 1.424E−01 | 0.455 | 18.182 | |
| Mintemp | −1.819E−03 | 0.318 | 9.091 | |
| $\text{Mintemp}^2$ | 2.636E−04 | 0.295 | 27.273 | |
| Maxtemp | −9.403E−03 | 0.454 | 18.182 | |
| $\text{Maxtemp}^2$ | 1.252E−04 | 0.351 | 18.182 | |
| Aveppt | −2.676E−04 | 0.188 | 36.364 | |
| $\text{Den}(t-1)$ | 1.386E+00 | 0.000 | 100.000 | |
| $\text{Den}(t-1)^2$ | −4.241E−01 | 0.000 | 90.909 | |
| $\text{Den}(t-1)^3$ | 5.170E−02 | 0.022 | 72.727 | |
| TimeFE | – | – | – | |
| CountyFE Median $R^2$ | −0.891 | – | – | |
| Rho | 2.452E−02 | 0.133 | 0.500 | Rural |
| Constant | −1.593E+00 | 0.081 | 54.545 | |
| Mintemp | 2.034E−02 | 0.362 | 27.273 | |
| $\text{Mintemp}^2$ | −2.706E−03 | 0.350 | 27.273 | |
| Maxtemp | 1.858E−01 | 0.309 | 36.364 | |
| $\text{Maxtemp}^2$ | −3.841E−03 | 0.482 | 27.273 | |
| Aveppt | −8.227E−03 | 0.003 | 54.545 | |
| $\text{Den}(t-1)$ | 1.018E+00 | 0.000 | 100.000 | |
| $\text{Den}(t-1)^2$ | −2.611E−03 | 0.001 | 90.909 | |
| TimeFE | – | – | – | |
| CountyFE | – | – | – | |
| Median $R^2$ | 0.77 | | | |
| Median $R^2$ | 0.881 | | | |

*Variable abbreviations:* Rho – spatial lag coefficient. Mintemp, Maxtemp–mean annual minimum and maximum temperatures from 1971 to 2000. Aveppt – average annual precipitation for 1971–2000. $\text{Den}(t-1)$ housing density with a one-period time lag. TimeFE – vector of decadal time dummies.

*Growth* housing redistributions, respectively. For example, if county A allocated 100 rural homes and 1000 urban homes in the *BAU* scenario, rural areas would receive an additional 25 homes and urban areas 25 homes fewer in the *Rural Growth* scenario. The total number of houses needed for 'relocation' is estimated for each period and county. Urban and rural housing densities are adjusted upward or downward by an optimization algorithm with an objective function that multiplies densities by a scalar in order to reflect the number of new houses expected in urban or rural communities in each period. Housing densities are therefore adjusted downwards or upwards to create *Urban* and *Rural Growth* scenarios bounding the business as usual scenario.

### *Impact indicators*

To assess wildfire risks posed to housing over time we estimate the number of houses in each wildfire hazard severity zone (FHSZ) as described in 2007 by the California Fire and Resource Assessment Program (FRAP, 2007). This dataset designates five hazard classes based on each area's fuel rank and probability of wildfire event. The median acreage of each FHSZ polygon is close to 7 acres compared to the 121 acres for each split block group. This however overstates the variance of the FHSZ, which is largely composed of patches, on the order of hundreds of square miles (Fig. A2), broadly following the map of functional plant type. In order to account for the discrepancy in polygon extent we use a spatial union algorithm to calculate the housing density for each FHSZ polygon. For the relatively few split block groups that cross a FHSZ type boundary, we would expected our model, which assumes homogeneity within a SBG, to slightly overstate the number of houses in the higher severity zone. It likewise underestimates the risk associated with spatial spillovers, whereby proximity to 'at risk' areas also increases risk for any given house.

To assess the potential impacts to agriculture, we estimate the total percentage of SBGs with pasture and/or agricultural land that are expected to experience growth between 2000 and 2050. To identify agricultural and pasture lands we utilize the National Land Cover Database (Fry et al., 2011). Here we identify all SBGs containing agricultural land and report the percentage of which that are expected to face development pressure.

### *Regression results*

There are 22 regressions, two each (urban, rural) for the 11 county-groups. Regression results are reported as median coefficient estimates and p-values are reported Table 2. To better demonstrate the localized importance of some of these variables, we also report the percentage of regressions in which each variable was significant at the 90% level. Coefficients indicate that temporally lagged housing densities increase density at low levels, level out at medium densities, and increase slightly or plateau at higher densities and are almost universally significant. Spatial spillovers were positive at the median in both urban and rural areas. At the median for urban areas, a one-unit increase in average neighboring densities implies a 0.059 unit increase for the SBG of interest, and ranges from 0.0061 to 0.20 units. For rural areas, the same increase in neighboring densities implies between a 0.12 unit decrease to a 0.083 unit increase. This speaks to the localized nature of development and spillovers. For some regions growth begets neighboring growth, while in others it precludes it. We interpret the interaction of temporal and spatial density lags as follows for rural communities: controlling for the fact that a location tend to become more dense over time, increased housing density in neighboring rural communities does not necessarily spill over into adjacent rural communities. In fact, in a small number of instances the regression indicates that growth in neighboring farm communities may actually decrease the rate of density growth in neighboring communities, in essence relieving the pressure for development in extensive agricultural areas. In the *BAU* scenario for the 2010–2050 period, a total of 270 SBGs transitioned from uninhabited to inhabited through the mechanism of spatial spillovers.

In relation to environmental variables, all significant values of minimum observed temperature Mintemp and $\text{Mintemp}^2$ were positive and negative respectively for urban areas, with the

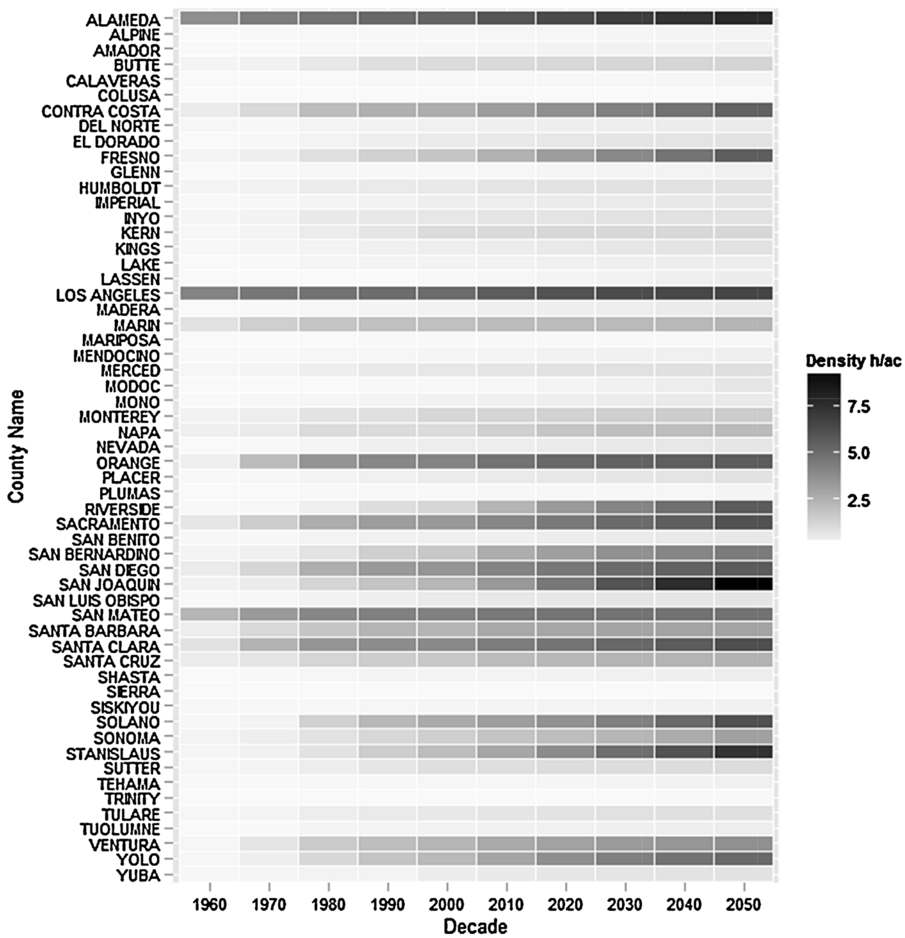


**Fig. 3.** Median BAU county-level housing density by decade (omitting San Francisco). Housing density (homes/acre) for California counties. Darker shades of grey indicate higher housing densities. Omits San Francisco because its disproportionately high density obscures the trends in other counties.

exception of the county group that includes the San Francisco Bay and South Bay counties. Where significant, Maxtemp and Maxtemp$^2$ are positive and negative respectively for the Bay Area, Lake Tahoe region, and the upper South Coast county groups. The south Central Valley county group showed a significant opposite effect. Both maximum and minimum temperatures had little significant effect on the location of rural development. Increased levels of precipitation (Aveppt) tended to decrease development density although with a small marginal effect. Due to broad insignificance, the variables distance to national parks and tribal land have been omitted from the final model.

For the median regression, the model describes 88% [range of 79–97%] of historical variance in housing density. To see how the model performs across growth classes, we divided the sample into 12 groups based on the total change in housing density between 1950 and 2000 (Fig. A3). We then computed the in-sample root mean squared prediction error for each class. We find that all middle growth class RMEs are on the order of 0.5 houses/acre and are relatively constant, with lower RMSE at both tails.

*Forecast and scenario results*

Figs. 3 and 4 (and Table A2) characterize county-level development for the *BAU Growth* scenario forecasts. Driven by demographic population forecasts (California, 2012), the highest levels of expected housing growth are for San Francisco, Sacramento, Contra Costa, Fresno, Riverside, San Joaquin, Solano, San Bernardino, Stanislaus and Yolo counties. These counties contain or abut the rapidly expanding communities of Modesto, Fresno, Fremont, Hayward, Stockton, San Bernardino, Sacramento or San Francisco. The first six of these communities have doubled or nearly doubled their population since the 1990 census, and the last two have added 125–200 thousand new residents (CSDC, 2011). The slowest predicted growth occurs in the geographically isolated, North Central Valley and Sierra Nevada counties of Alpine, Colusa, Mariposa, Plumas, Sierra, Sutter, and Trinity.

Fig. 5 represents the distribution of SBG new housing growth under *BAU*, *Urban* and *Rural Growth* scenarios. The distribution of densities are shown by both the explicit distribution in the violin plot, and by the 25th and 75th percentiles bounded by the boxplots and the median as indicated by the white dot. In all scenarios, the bottom of the distribution as seen in the violin plot, especially in rural areas (<1 house/acre), narrows over time indicating the ongoing development of sparsely inhabited areas. The range of 25th and 75th percentiles narrow noticably in the *Rural Growth* scenario as new houses are built in areas previously at the bottom of the distribution. The opposite is true to a lesser degree with the *Urban Growth* scenario as more areas remain sparsely inhabited while densely populated areas continue to grow, skewing the distribution upwards. The change observed in the *Urban Growth* scenario is however limited by the fact that development under the *BAU* scenario already occurs predominantly in urban areas.

The historical and forecast changes in the distribution of the housing development classes can be seen clearly in the stacked bar plots in Fig. 6, which depict both the acres affected and total house

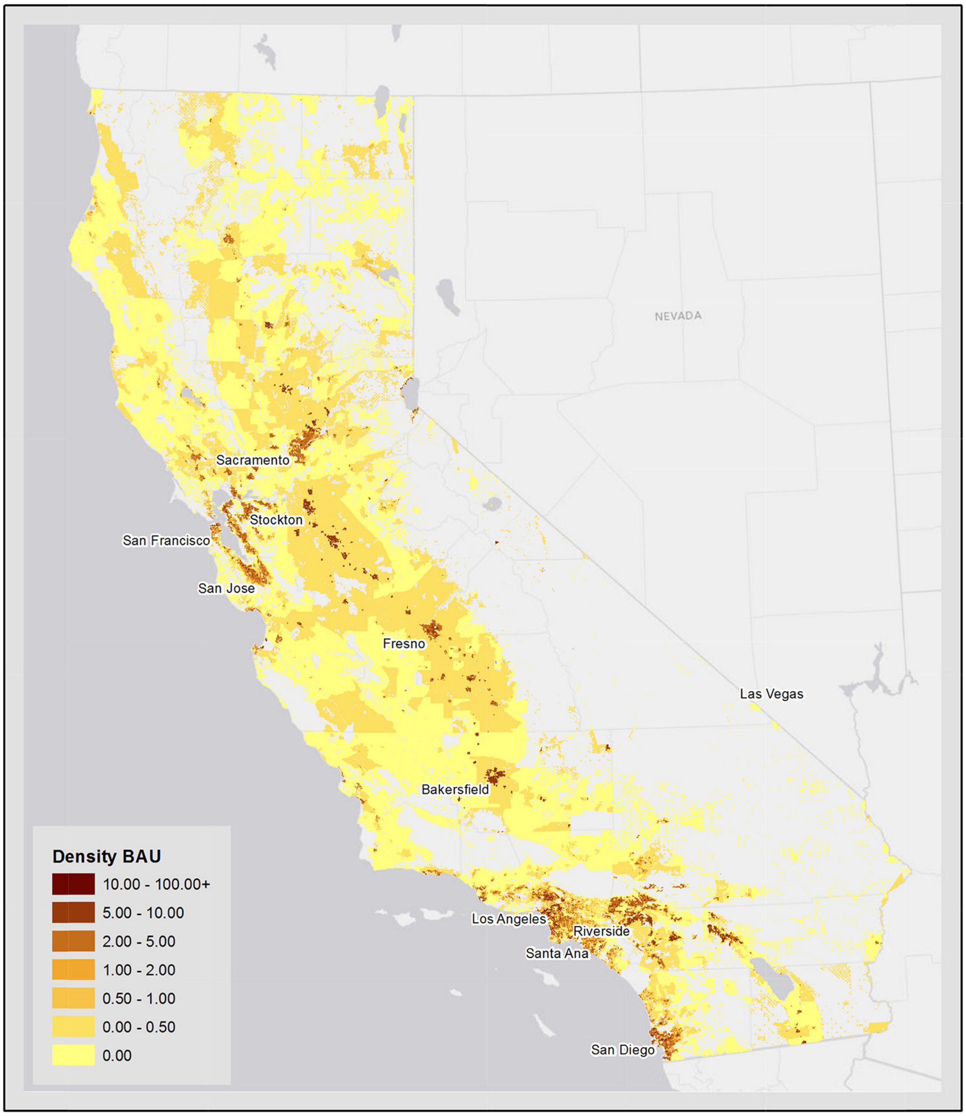


**Fig. 4.** Change in density (houses/acre) 2000–2050 in business as usual scenario. Map of forecast change in housing density (2000–2050) for the business as usual scenario. Darker shades of maroon indicate a greater increase in the housing density for this period. Grey areas represent protect or otherwise undevelopable areas.

counts for each density class outlined in Table 1. Rows indicates the evolution of total developable acres (top) and total number of houses (bottom) by housing density class. Columns represent the three forecast scenarios. These results are also presented in tabular format in Table 3.

Fig. 7 predicts the number of non-urban houses from 2010 to 2050 in each fire hazard severity zone (FHSZ). It presents housing counts over the forecast period for all non-urban FHSZ categories for the three forecast scenarios. Excluding urban, both moderate and very-high severity classes have the highest housing counts and growth rates. This high level and rate of development in the highest risk areas will likely increase the human and economic costs of wildfire.

We also estimate the impact of forecast housing development on agricultural and pasture lands. We estimate that approximately 36,000 $km^2$ of agricultural land, or 90% of split block groups containing agricultural lands, will experience positive housing growth pressures by mid-century. The same can be said for pasturelands with approximately 10,000 $km^2$, or 58% of SBGs with pasture, experiencing growth pressure. However our assumption of homogeneous settlement patterns likely overstates the influence of the lowest density areas on their environment.

**Table 3**
Acreage by density classification in 2050.

| Density class | Business as usual | Rural scenario | Urban scenario | 2000 level | Urban status |
|---|---|---|---|---|---|
| Very-high | 1,994,425 | 1,853,012 | 2,056,776 | 879,018 | Urban |
| High | 1,071,926 | 2,305,730 | 803,543 | 1,975,942 | |
| Medium | 1,674,039 | 1,553,045 | 973,923 | 901,134 | Rural |
| Low | 12,867,409 | 16,281,818 | 9,895,336 | 8,055,132 | |
| Sparse | 29,089,052 | 24,703,246 | 32,967,273 | 34,885,624 | |

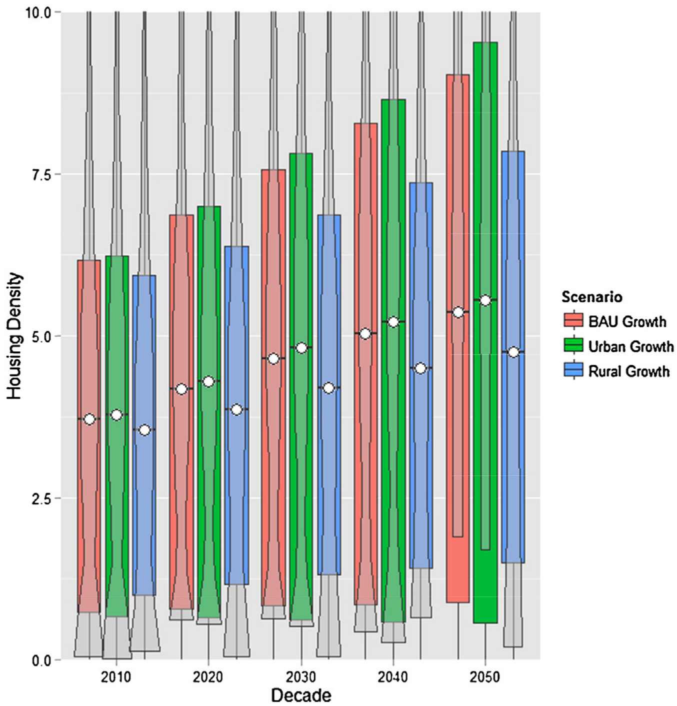


**Fig. 5.** Split block group (SBG) housing density by scenario and year. Mixed violin and box plot of housing density forecasts for three development scenarios. The distribution of densities are shown by both the explicit distribution in the violin plot, and by the 25th and 75th percentiles bounded by the box plots and the median as indicated by the white dot.

*Modeling framework considerations*

While the use of the complete block groups would assume homogeneity across the geography, split block groups allow for far greater heterogeneity (Hammer et al., 2004). Furthermore, all undevelopable lands are removed and all remaining housing units distributed to the developable portions of the SBGs. As such, housing densities presented in this paper can be considered 'net' rather than 'gross' densities. We defined undevelopable lands as water (U.C. Bureau, 2010; Library, 2008), conservation easements (NCED, 2012), military bases (US Military, 2011) and private conservation (non-easement) and public lands (Cal Fire, 2011) which includes all State and Federal lands, as well as the CPAD database (CPAD, 2012).

As modeled by the spatial panel, development innovations are limited to a temporal lag, climate, and spatial spillovers from neighboring SBGs. As such, random seeding of new population centers are not represented within these models. Although random seeding occurs in the real world, we assume that the likelihood of seeding follows the non-random drivers of spatial spillovers from neighboring blocks, and other drivers controlled for in this model.

## Discussion

Our results evidence a rapid loss of sparsely developed and agricultural areas in California by 2050 under the *BAU* scenario. The most striking feature of Fig. 5 is the marked decline in sparsely developed lands in the *BAU* scenario. These areas, which account for all SBGs with less than 1 house per 40 acres, constitute many privately-owned lands in extensive (livestock grazing) or intensive (irrigated crops) agricultural uses, and undeveloped areas outside of protected or undevelopable areas. In respect to agriculture, we forecast that 90% and 58% of agricultural and pasturelands, respectively, will experience positive housing growth pressure. The strong and uniform decline in sparsely settled land speaks to the rapid loss of wildlands and "working" landscapes; a trend which seems unlikely to flag in the coming decades without intervention.

In the *BAU* scenario, the sparse density housing class is largely replaced by low density lands, with increases of 60% to 12.9 million acres in the period from 2000 to 2050 (Table 3). In turn, medium density land is expected to increase by 86% to 1.7 million acres. Both of these land classes are of particular interest because they roughly correspond to the wildland urban interface (WUI) classifications (Register, 2001; Theobald and Romme, 2007). Because WUI lands lie at the intersection of human and relatively wild populations, the development of these lands will have disproportionate effects on both society and ecosystems. The spread of low housing density will likely lead to widespread landscape fragmentation and other important changes to California's unique and valuable ecosystems. Land fragmentation not only results in the direct loss of wildlife habitat and biodiversity, but also increases the rates of invasion by non-native species, alters water quality and availability, increases fire risk and recreation pressures, among others (e.g., Hammer et al., 2004; Alavalapati et al., 2005; Newburn and Berck, 2006). All these effects may be exacerbated in the context of climate change, posing serious challenges for landscape management and conservation. Furthermore, increased WUI acreage will have important implications for land management as well as increased costs and risks to society (Calkin et al., 2005; Gebert et al., 2007a,b; Gude et al., 2008a,b; Liang et al., 2008).

Although substantial portions of California are dedicated as military, multiple-use, and wild public spaces (seen in grey in Fig. 4) and laudably contain most of the state's greatest natural treasures, many of the largest portions appear to have been selected due to their inaccessibility and unsuitability for settlement or reclamation, rather than their merits as wild spaces of interest. The remaining and developable portions of the map will be shaped by the choices made about the location and extent of human settlements. Two potential scenarios, *Urban* and *Rural Growth* can be compared with *Business as Usual* outcomes. With significant intervention, pushing an additional one quarter of all expected rural housing growth into urban areas, the low-density class is still expected to increase by 22% from 8.1 to 9.9 million acres between 2000 and 2050 (Table 3). If those houses were diverted into rural areas instead, the low-density class would be expected to nearly double, covering over 16 million acres in the final period. As such, policies encouraging urban development can limit low-density expansion to nearly 61% of its possible extent. Translated into acreage, the *Urban Growth* scenario maintains an additional 8.3 million acres in the most sparsely habituated land class, compared to the *Rural Growth* scenario.

More densely populated rural areas (medium density lands) follow a similar pattern with one interesting distinction (Table 3): The *Business as Usual* scenario sees the largest expansion of medium density communities, increasing by 86% relative to 72% for *Rural Growth* and 8.3% for *Urban Growth*, with a higher proportion of houses moving from low to medium density in *BAU*. *Rural Growth* instead sees most of its rural development as transition from sparse to low density, with much of the initial low and middle density areas transitioning to high-density communities.

During the 1955–1985 period, wildland fires in California were responsible for the destruction of 3533 structures and resulted in 25 deaths (Hammer et al., 2007). This statistic is representative of the extent of the interplay between anthropogenic and natural systems throughout the state. At the Federal level, 92 people were killed and 10,159 houses were lost due to fires during the 2002–2006 period, resulting in a combined cost of $6.3 billion (Gude et al., 2008a,b). Despite these losses, housing development and thus the WUI have continued to expand in fire prone areas supported by state and federal subsidies for wildfire suppression and mitigation (Olmstead et al., 2012). To offset these risks, the California State Building Standards Commission adopted building codes to provide

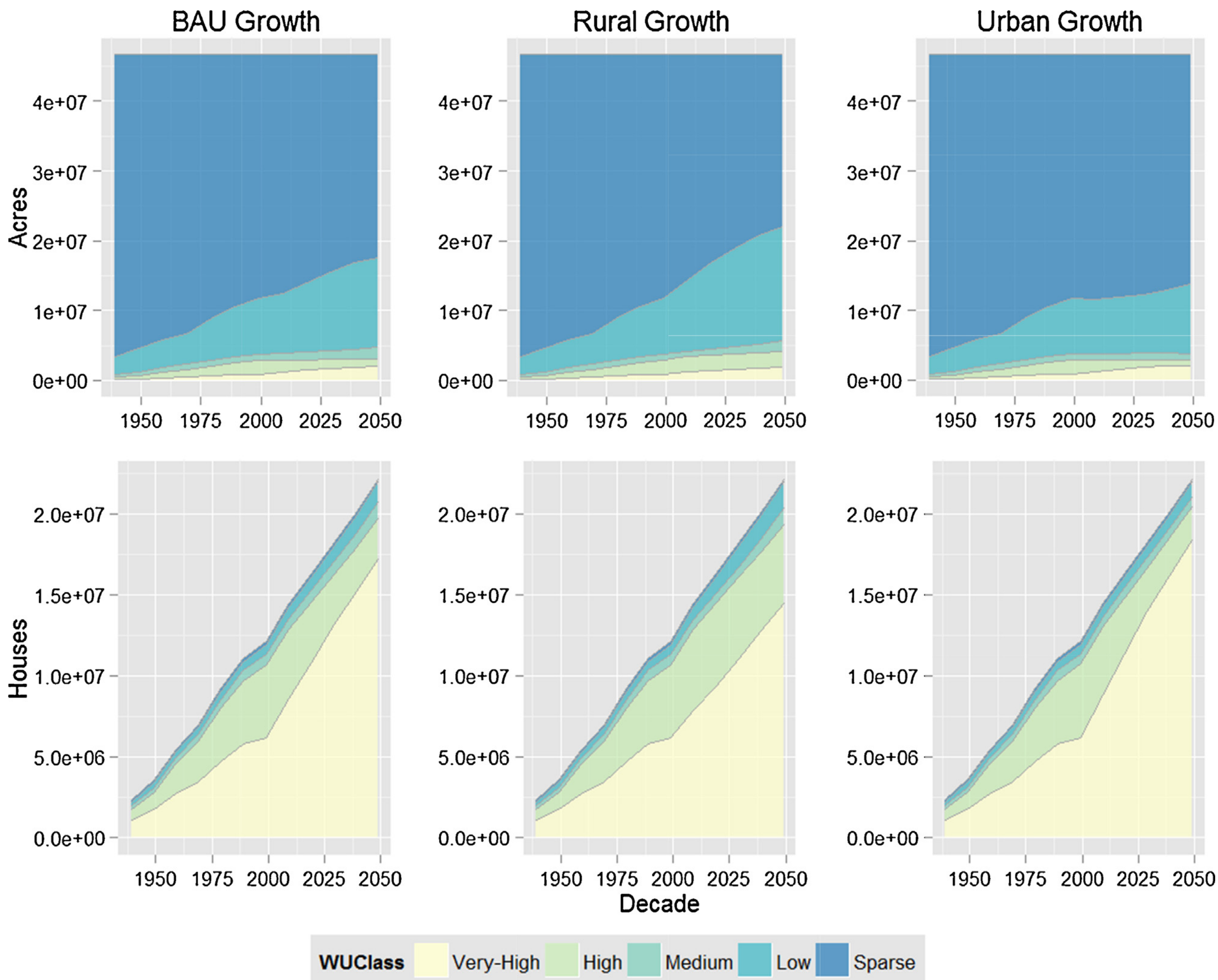


**Fig. 6.** Acres and house counts by wildland urban class and year. Rows indicate the evolution of total developable acres (top) and total number of houses (bottom), by housing density class over time. Graphs in the first column are BAU results, second column rural growth, and third column urban growth scenarios. Within a graph, the colored areas shows proportion of developable acres (resp. total house count) by WUC class, by year.

strict regulation on new building within fire-prone WUI areas. For instance, Code 701A is a comprehensive reassessment that required the use of fire-resistant materials, systems, and assembly in new construction, as well as the use of fire-safe landscaping (OSFM, 2009). Other locales, such as communities in San Diego, added additional requirements such as the use of 'shelter in place' (SIP) standards in new construction (Fu, 2012).

Future risks to settlements are indicated by housing counts in 'Very High' wildfire severity zones. Housing counts in these 'Very High' risk zones are expected to reach 1.7 million in the BAU scenario by 2050 with as many as 2.2 million or as few as 1.6 million houses in the *Rural and Urban Growth* scenarios, respectively (Fig. 7). This implies at least 640 thousand to 1.2 million new homes in the highest wildfire risk areas between 2000 and 2050. Much of this high-risk development is likely to occur on the outskirts of cities on the southern coasts including Los Angeles, San Diego and Santa Barbara, but also outlying communities of the Bay area (Fig. 4). Although WUI building codes and fire prevention fees aim to reduce risks and mitigate subsidies, it is unclear that these efforts will be sufficient to discourage significant high-risk rural development.

New developments will have both short- and long-term effects on risk management for human communities. In the short-term, building enhancements should reduce the likelihood that any *new* structure will accrue losses. Although cases like 'The Trails' development in San Diego, where over 1/3 of fire hardened homes were destroyed, indicate that many basic questions about wildfire risk management remain unanswered (Maranghides and Mell, 2009). In the long-term, these changes will likely engender greater "anthropogenic hubris" as *fire resistant* communities push further into wildlands (Fu, 2012). These new developments also have clear consequences for *existing* housing stock. In high-density areas, the addition of fire resistant developments should reduce fuel loads, and increase and consolidate available resources. In low-density lands, common in high fire risk areas, additional development may leave older stock exposed to higher ignition risks and greater dispersion of the limited suppression and mitigation resources.

Continued residential development in and around major metropolitan areas will also contribute to increased roadway congestion. Many urban transportation networks in California already experience heavy congestion at peak periods, and the ability to expand road capacity is constrained by cost and geography. Areas likely to see increased congestion include Los Angeles, San Diego, Sacramento, Fresno and the whole of the Bay area. While it has been shown that increases in residential population density are associated with a reduction in per-capita vehicle miles traveled (Cervero and Murakami, 2009; Ewing and Cervero, 2010), estimates of the size of this effect are generally modest, ranging from a 5 to

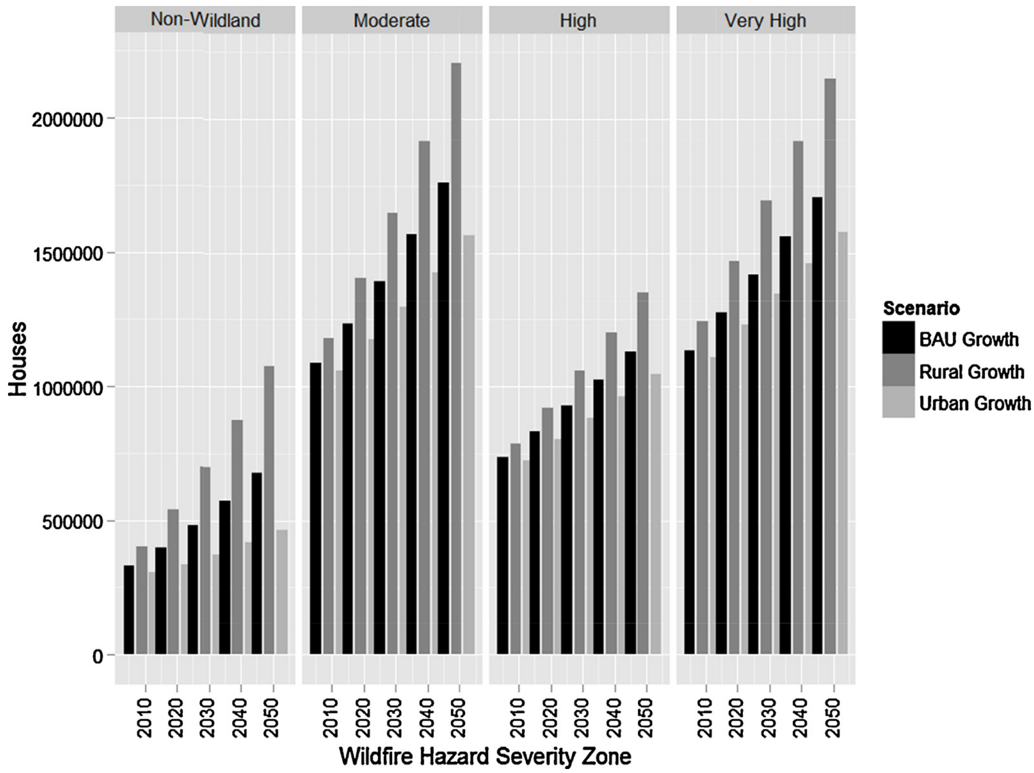


**Fig. 7.** House counts by wildfire hazard severity zone and growth scenario. Number of houses in each wildfire hazard severity zone, by forecast scenario. The urban land class has been omitted in order to better depict relevant fire risks.

25% reduction for a doubling of a density (NRC, 2010). The consequence of this is that any expansion to metro area populations will tend to increase vehicle travel. This will be more pronounced where the development occurs in lower density suburban and exurban parcels. The consequences of increased traffic congestion can be severe, as lower vehicle speeds both reduce fuel economy and increase the emissions of harmful air pollutants. Levy et al. (2010) estimated that the costs of increased mortality risk due to congestion-related $PM_{2.5}$, $SO_2$ and $NO_x$ emissions in Los Angeles exceeded $3 billion in 2010. As well, the additional fuel burned due to traffic congestion in U.S. urban areas in 2011 generated carbon dioxide emissions in excess of 25 million metric tons (Shrank and Lomax, 2012), nearly 2% of total on-road $CO_2$ emissions in the U.S. that year (EPA, 2013). The lack of extensive public transit systems in many of California's major metropolitan areas will continue to limit the ability of municipal and regional planners to mitigate traffic congestion from increased residential development over the next several decades.

## Conclusions

The Californian landscape is defined as much by its natural hazards as it is by its beauty. Households here are formed under a unique set of low-probability threats, and the houses themselves may be 'hardened' to make them less vulnerable. However, each of the development scenarios set out in this paper has its own risks for both human and natural communities. To minimize risks from fire, planners may aim to encourage smart urban development and to consolidate existing rural communities into more robust configurations. The risks that these changes present for local ecosystems and human communities will vary according to both the initial land class, its composition, and the resultant one. Although relatively untested, strategies that include locating irrigated agriculture, golf courses, or other land uses, may act as buffers to fire effects along the WUI would appear to be steps in the right direction. While urban infill will minimize anthropogenic effects in wildlands, this should not be done at expense of doing nothing to improve the robustness of rural communities. Planners, communities and policymakers must therefore find a balance between the short- and long-term risks and benefits of different development patterns.

Progress has been made to encourage technology and market signals into rural development, through the use of building codes and cost sharing fire protection fees (Fu, 2012). Although welcome, these requirements and signals seem inadequate to restructure the path and patterns of rural development (Olmstead et al., 2012). Rural planning and its associated infrastructure should account for agriculture, biodiversity, conservation, as well as risk-management priorities. Meanwhile, new protected and limited-use lands could increase the effectiveness of existing wildlands through greater extent and connectivity, enhancing landscape and species diversity, ecosystem resilience, and function. Conservation of these areas would more clearly restrict the extent of human development, promote denser settlements, and thereby minimize risks and cost to society.

## Acknowledgments

Funding for this study was generously provided by The Nature Conservancy in California. The authors wish to thank Jing Lu and Emmalee Dolfi for their work developing the historical populated places dataset.

## Appendix A. Neighborhood definition

The specification of the neighborhood matrix is outlined in the methods section. Table A1 outlines the distribution of rho which describes the direction of spatial spillovers between SBGs in each regression. Estimates of rho are both positive and negative

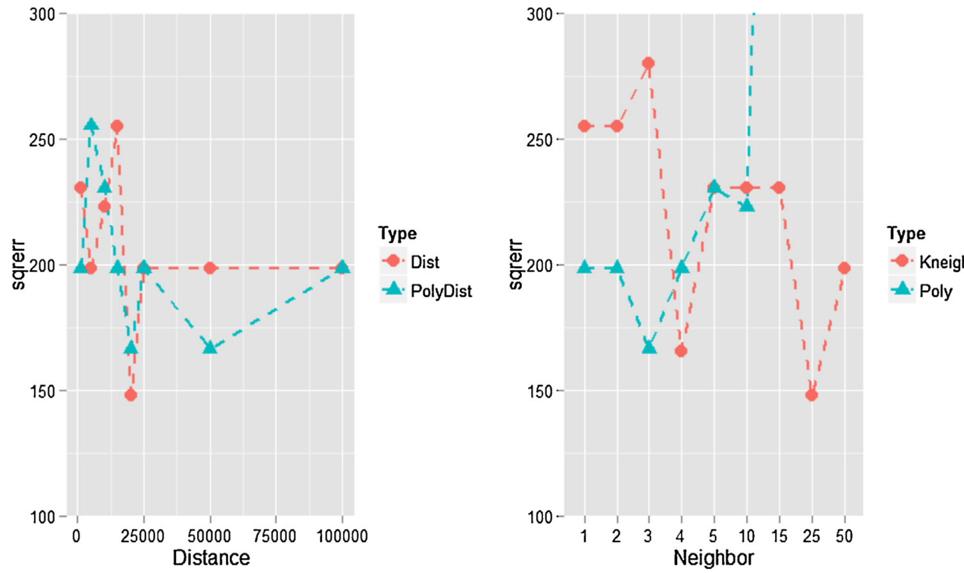


**Fig. A1.** Median squared error by distance and neighbor bands.

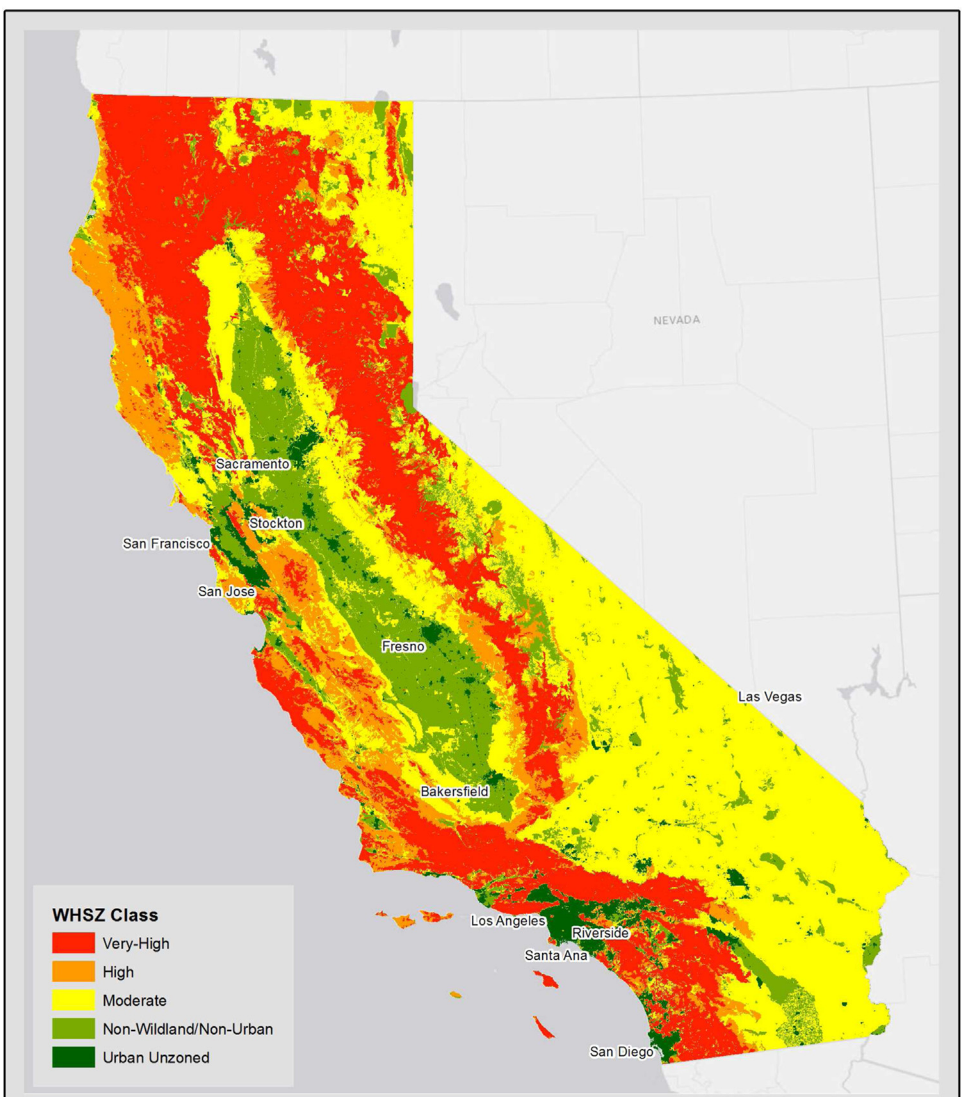


**Fig. A2.** Fire Hazard Severity Zones of California. FHSZ represent the risk to housing due to the areas fuel rank and probability of a wildfire event.

**Table A1**
Examples spatial spillover under possible neighborhood assumptions.

| Neighbor specification | Rho | | | Rho *P* |
|---|---|---|---|---|
| | Lower | Median | Upper | Median |
| *Dist (m)* | | | | |
| 1000 | −0.068 | −0.016 | 0.036 | 0.445 |
| 5000 | −0.303 | −0.036 | 0.097 | 0.200 |
| 10,000 | −0.862 | −0.033 | 0.231 | 0.170 |
| 15,000 | −0.988 | −0.066 | 0.243 | 0.153 |
| 20,000 | −0.561 | −0.064 | 0.322 | 0.064 |
| 25,000 | −0.415 | −0.102 | 0.383 | 0.151 |
| 50,000 | −0.901 | −0.162 | 0.581 | 0.131 |
| 100,000 | −0.866 | −0.105 | 0.757 | 0.130 |
| *PolyDist (m)* | | | | |
| 1000 | −0.134 | −0.051 | 0.039 | 0.176 |
| 5000 | −0.307 | −0.072 | 0.062 | 0.156 |
| 10,000 | −0.862 | −0.078 | 0.128 | 0.205 |
| 15,000 | −0.988 | −0.081 | 0.243 | 0.226 |
| 20,000 | −0.561 | −0.114 | 0.322 | 0.142 |
| 25,000 | −0.415 | −0.144 | 0.383 | 0.112 |
| 50,000 | −0.901 | −0.167 | 0.581 | 0.146 |
| 100,000 | −0.866 | −0.146 | 0.757 | 0.130 |
| *Poly (P)* | | | | |
| 1 | −0.137 | −0.049 | 0.035 | 0.203 |
| 2 | −0.629 | −0.074 | 0.036 | 0.101 |
| 3 | −0.367 | −0.079 | 0.031 | 0.178 |
| 4 | −0.831 | −0.079 | 0.033 | 0.164 |
| 5 | −0.720 | −0.072 | 0.106 | 0.209 |
| 10 | −0.368 | −0.065 | 0.159 | 0.254 |
| 15 | −0.374 | −0.061 | 0.147 | 0.227 |
| *Kneigh (K)* | | | | |
| 1 | −0.050 | −0.013 | 0.036 | 0.279 |
| 2 | −0.103 | −0.017 | 0.020 | 0.215 |
| 3 | −0.158 | −0.022 | 0.019 | 0.351 |
| 4 | −0.236 | −0.033 | 0.011 | 0.148 |
| 5 | −0.213 | −0.051 | 0.019 | 0.191 |
| 10 | −0.493 | −0.141 | 0.063 | 0.051 |
| 15 | −0.492 | −0.132 | 0.072 | 0.051 |
| 25 | −0.544 | −0.166 | −0.015 | 0.062 |
| 50 | −0.719 | −0.186 | 0.175 | 0.090 |

**Table A2**
Median housing density by county and decade.

| Name | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 | 2030 | 2040 | 2050 |
|---|---|---|---|---|---|---|---|---|---|---|
| YUBA | 0.08 | 0.15 | 0.21 | 0.21 | 0.43 | 0.54 | 0.63 | 0.71 | 0.77 | 0.80 |
| YOLO | 0.14 | 0.40 | 1.13 | 1.81 | 2.18 | 2.93 | 3.69 | 4.26 | 4.75 | 5.15 |
| VENTURA | 0.13 | 0.71 | 1.59 | 2.04 | 2.30 | 2.73 | 3.01 | 3.27 | 3.46 | 3.67 |
| TUOLUMNE | 0.02 | 0.07 | 0.17 | 0.21 | 0.24 | 0.28 | 0.33 | 0.35 | 0.40 | 0.43 |
| TULARE | 0.13 | 0.21 | 0.41 | 0.51 | 0.59 | 0.69 | 0.77 | 0.82 | 0.86 | 0.87 |
| TRINITY | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 |
| TEHAMA | 0.04 | 0.05 | 0.09 | 0.11 | 0.13 | 0.14 | 0.16 | 0.21 | 0.26 | 0.32 |
| SUTTER | 0.13 | 0.22 | 0.44 | 0.66 | 0.89 | 0.92 | 0.94 | 0.96 | 0.97 | 0.98 |
| STANISLAUS | 0.15 | 0.29 | 0.75 | 1.49 | 2.07 | 2.86 | 3.83 | 4.96 | 6.07 | 7.34 |
| SONOMA | 0.21 | 0.37 | 0.78 | 1.15 | 1.44 | 1.81 | 2.04 | 2.32 | 2.67 | 3.05 |
| SOLANO | 0.10 | 0.21 | 1.38 | 2.24 | 2.70 | 3.14 | 3.60 | 4.28 | 5.15 | 6.13 |
| SISKIYOU | 0.07 | 0.07 | 0.09 | 0.13 | 0.15 | 0.18 | 0.20 | 0.23 | 0.25 | 0.28 |
| SIERRA | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 |
| SHASTA | 0.03 | 0.05 | 0.12 | 0.19 | 0.19 | 0.22 | 0.27 | 0.31 | 0.36 | 0.42 |
| SANTA CRUZ | 0.46 | 0.70 | 1.25 | 1.50 | 1.69 | 2.10 | 2.23 | 2.33 | 2.39 | 2.42 |
| SANTA CLARA | 0.77 | 2.37 | 3.46 | 3.80 | 3.91 | 4.33 | 4.70 | 5.12 | 5.65 | 6.17 |
| SANTA BARBARA | 0.41 | 1.10 | 1.74 | 2.36 | 2.33 | 2.77 | 2.84 | 2.90 | 2.95 | 2.97 |
| SAN MATEO | 2.32 | 3.29 | 4.02 | 4.24 | 4.23 | 4.52 | 4.68 | 4.74 | 4.75 | 4.75 |
| SAN LUIS OBISPO | 0.06 | 0.12 | 0.24 | 0.38 | 0.51 | 0.59 | 0.62 | 0.67 | 0.71 | 0.74 |
| SAN JOAQUIN | 0.25 | 0.46 | 1.23 | 1.81 | 2.26 | 3.35 | 4.58 | 5.97 | 7.61 | 9.42 |
| SAN FRANCISCO | 14.43 | 14.44 | 14.59 | 15.08 | 15.51 | 17.45 | 18.07 | 18.61 | 18.77 | 18.77 |
| SAN DIEGO | 0.48 | 1.19 | 2.56 | 3.34 | 3.50 | 4.13 | 4.57 | 5.08 | 5.42 | 5.68 |
| SAN BERNARDINO | 0.18 | 0.33 | 0.75 | 1.45 | 1.68 | 2.61 | 3.14 | 3.63 | 4.06 | 4.47 |
| SAN BENITO | 0.08 | 0.08 | 0.22 | 0.29 | 0.31 | 0.37 | 0.43 | 0.49 | 0.55 | 0.58 |
| SACRAMENTO | 0.63 | 1.49 | 2.62 | 3.19 | 3.36 | 4.02 | 4.53 | 5.05 | 5.54 | 6.07 |
| RIVERSIDE | 0.04 | 0.11 | 0.34 | 0.92 | 1.21 | 2.35 | 3.27 | 4.08 | 4.86 | 5.67 |
| PLUMAS | 0.02 | 0.03 | 0.04 | 0.05 | 0.07 | 0.08 | 0.10 | 0.12 | 0.14 | 0.17 |
| PLACER | 0.04 | 0.09 | 0.18 | 0.33 | 0.45 | 0.53 | 0.59 | 0.66 | 0.72 | 0.77 |
| ORANGE | 0.37 | 2.11 | 3.52 | 4.00 | 4.12 | 4.77 | 5.16 | 5.40 | 5.53 | 5.61 |
| NEVADA | 0.02 | 0.05 | 0.16 | 0.34 | 0.42 | 0.43 | 0.49 | 0.55 | 0.60 | 0.64 |

Table A2 (*Continued*)

| Name | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 | 2030 | 2040 | 2050 |
|---|---|---|---|---|---|---|---|---|---|---|
| NAPA | 0.34 | 0.47 | 1.04 | 1.04 | 1.05 | 1.40 | 1.79 | 1.98 | 2.03 | 2.14 |
| MONTEREY | 0.26 | 0.43 | 0.75 | 0.87 | 1.15 | 1.23 | 1.33 | 1.46 | 1.52 | 1.54 |
| MONO | 0.02 | 0.09 | 0.15 | 0.19 | 0.27 | 0.33 | 0.39 | 0.45 | 0.51 | 0.57 |
| MODOC | 0.01 | 0.01 | 0.03 | 0.05 | 0.05 | 0.11 | 0.19 | 0.30 | 0.44 | 0.65 |
| MERCED | 0.11 | 0.19 | 0.37 | 0.54 | 0.62 | 0.69 | 0.77 | 0.82 | 0.88 | 0.91 |
| MENDOCINO | 0.04 | 0.06 | 0.11 | 0.15 | 0.17 | 0.20 | 0.24 | 0.28 | 0.31 | 0.36 |
| MARIPOSA | 0.01 | 0.01 | 0.03 | 0.04 | 0.04 | 0.05 | 0.05 | 0.06 | 0.06 | 0.07 |
| MARIN | 0.78 | 1.43 | 1.78 | 1.92 | 1.95 | 2.11 | 2.11 | 2.16 | 2.20 | 2.36 |
| MADERA | 0.03 | 0.04 | 0.12 | 0.17 | 0.28 | 0.34 | 0.41 | 0.43 | 0.47 | 0.55 |
| LOS ANGELES | 4.12 | 4.62 | 4.77 | 4.99 | 5.01 | 5.64 | 6.02 | 6.31 | 6.47 | 6.55 |
| LASSEN | 0.01 | 0.01 | 0.01 | 0.03 | 0.05 | 0.11 | 0.19 | 0.26 | 0.36 | 0.44 |
| LAKE | 0.03 | 0.06 | 0.15 | 0.17 | 0.22 | 0.27 | 0.32 | 0.37 | 0.42 | 0.45 |
| KINGS | 0.09 | 0.16 | 0.24 | 0.33 | 0.38 | 0.46 | 0.56 | 0.65 | 0.72 | 0.77 |
| KERN | 0.09 | 0.20 | 0.45 | 0.71 | 0.98 | 1.08 | 1.12 | 1.14 | 1.14 | 1.15 |
| INYO | 0.16 | 0.21 | 0.52 | 0.59 | 0.62 | 0.68 | 0.73 | 0.76 | 0.79 | 0.81 |
| IMPERIAL | 0.06 | 0.08 | 0.19 | 0.26 | 0.37 | 0.43 | 0.48 | 0.53 | 0.60 | 0.64 |
| HUMBOLDT | 0.19 | 0.27 | 0.45 | 0.53 | 0.57 | 0.68 | 0.75 | 0.78 | 0.79 | 0.80 |
| GLENN | 0.03 | 0.04 | 0.08 | 0.10 | 0.10 | 0.13 | 0.16 | 0.19 | 0.24 | 0.31 |
| FRESNO | 0.17 | 0.38 | 0.86 | 1.35 | 1.75 | 2.49 | 3.18 | 3.92 | 4.71 | 5.58 |
| EL DORADO | 0.03 | 0.08 | 0.22 | 0.39 | 0.49 | 0.55 | 0.63 | 0.69 | 0.74 | 0.75 |
| DEL NORTE | 0.11 | 0.09 | 0.24 | 0.26 | 0.31 | 0.35 | 0.38 | 0.42 | 0.45 | 0.49 |
| CONTRA COSTA | 0.49 | 1.10 | 2.08 | 2.52 | 2.59 | 3.18 | 3.65 | 4.22 | 4.80 | 5.44 |
| COLUSA | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.03 |
| CALAVERAS | 0.01 | 0.01 | 0.03 | 0.05 | 0.07 | 0.08 | 0.09 | 0.11 | 0.14 | 0.18 |
| BUTTE | 0.19 | 0.28 | 0.59 | 0.88 | 0.99 | 1.08 | 1.14 | 1.18 | 1.21 | 1.23 |
| AMADOR | 0.01 | 0.03 | 0.07 | 0.10 | 0.13 | 0.16 | 0.20 | 0.25 | 0.28 | 0.32 |
| ALPINE | 0.01 | 0.04 | 0.07 | 0.09 | 0.10 | 0.12 | 0.13 | 0.15 | 0.17 | 0.19 |
| ALAMEDA | 3.69 | 4.38 | 4.80 | 5.20 | 5.33 | 5.91 | 6.35 | 6.87 | 7.37 | 7.76 |

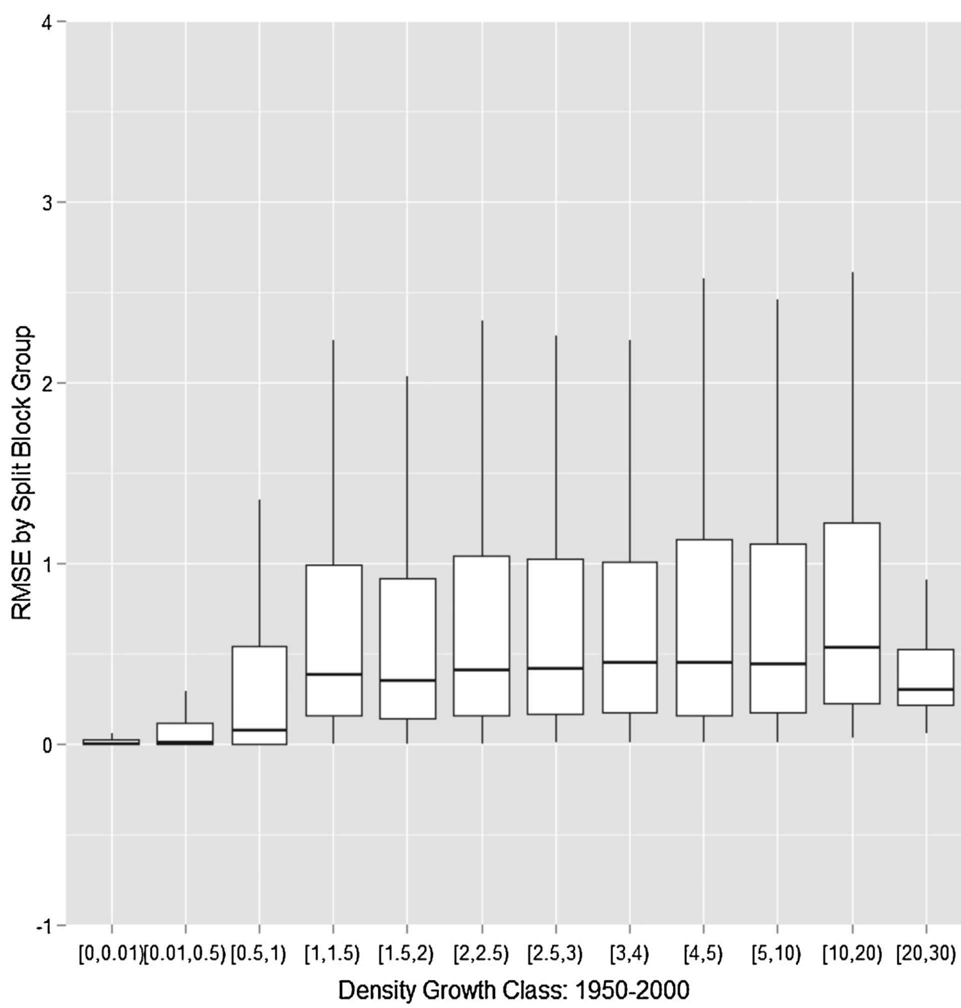


**Fig. A3.** Root Mean Square Error for Density Growth Classes. In-sample SBG RMSE boxplots portray model accuracy across a variety of housing density growth classes. Where classes are defined by the amount of housing density growth experienced between 1950 and 2000. Boxplots represent the 25th and 75th percentiles, with the median indicated by a black dash.

regardless of neighbor specification. Positive values indicate a positive spillover whereby a SBG's housing density increases by some percentage of its neighbor's average current density. With the opposite holding true for negative values of rho. The heterogeneity of rho as described at different spatial scales points to the complexity and localized nature of housing development. For some regions and levels of development, growth begets neighboring growth, while in others it precludes it.

Fig. A1 demonstrates the relative insensitivity of distance band on squared regression error and the improvement of fit with higher order *K* and *P* neighbors. Two specifications were identified for as possible final specifications. *PolyDist* ($m$ = 20,000) is chosen because it provides the minimal value of median squared regression error (*sqrerr*) and provides a consistent methodological foundation whereby the effects of localized spillovers are limited to a reasonable distance. *Kneigh* ($K$ = 25) while also minimizing squared error is harder to defend. Twenty-fifth order neighbors are conceivable in urban areas with small SBGs but abstract for large rural SBGs.

## References


Alavalapati, J.R.R., Carter, D.R., Newman, D.H., 2005. Wildland–urban interface: challenges and opportunities. Forest Policy Econ. 7 (5), 705–708.
Anselin, L., 1999. Spatial Econometrics. School of Social Sciences–University of Texas Dallas, Richardson, TX.
Anselin, L., Le Gallo, J., Jayet, H., 2008. Spatial panel econometrics. Adv. Stud Theor Appl. Econometr. 46, 625–660, http://dx.doi.org/10.1007/978-3-540-75892-1_19.
Bureau, U.S.C., 2000a. Census 2000 Summary File 1 100-Percent Data.
Berry, M., Hazen, B., MacIntyre, R., Flamm, R., 1996. Lucas: a system for modeling land-use change. IEEE Comput. Sci. Eng. 3 (1), 24–35.
Bureau, U.C., 1995. California Population of Counties by Decennial Census: 1900 to 1990. Washington, DC.
Bureau, U.C., 2007. Summary File 3 – Technical Documentation 2000 Census of Population and Housing. US Census Bureau, Washington, DC.
Bureau, U.C., 2010. Census 2010 TIGER/Line Files. US Census Bureau, Washington, DC.
Bureau, U.C., 2012a. Housing Characteristics from STF3 APPENDIX B, "Definitions of Subject Characteristics", Retrieved from http://www.census.gov/geo/lv4help/apen_bhous.html#YEAR (08.02.12).
Bureau, U.S.C., 1990. Population and Housing Unit Counts: California.
Bureau, U.S.C., 2000b. Census 2000 Summary File 3. Washington, DC.
Bureau, U.S.C., 2012b. Census 2000 Geographic Definitions (retrieved 08.02.12).
California, S.o., 2012. Population Projections for California and Its Counties 2010–2050. Department of Finance, Sacremento, CA.
Calkin, D.E., Gebert, K.M., Jones, G., Neilson, R.P., 2005. Forest service large fire area burned and suppression expenditure trends, 1970–2002. J. Forest. 103 (4), 179–183.
CBIA, 2013a. 2013 Residential New Housing Units & Valuations, Retrieved from: http://www.cbia.org/tasks/sites/cbia/assets/File/Residential%202013.pdf
CBIA, 2013b. California Housing Starts 1954–2012, Retrieved from: http://www.cbia.org/tasks/sites/cbia/assets/File/Historical%20Housing%20Starts%201954-2012.pdf
Cervero, R., Murakami, J., 2009. Effects of built environments on vehicle miles traveled: evidence from 370 US urbanized areas. Environ. Plan. A 42, 400–418.
Chomitz, K., Gray, D., 1996. Roads, land use, and deforestation: a spatial model applied to belinze. World Bank Econ. Rev. 10 (3), 487–512.
CPAD, 2012. California Protected Areas Database Version 1.7, Retrieved from: www.calands.org
Crump, J., 2003. A place in the country: exurban and suburban development in Sonoma county, California. Environ. Behav. 35 (2), 187–202.
CSDC, 2011. Historical Census Populations of California, Counties, and Incorporated Cities, 1850–2010, Retrieved from: http://www.dof.ca.gov/research/demographic/
Dietzel, C., Clarke, K., 2007. Toward optimal calibration of the SLEUTH land use change model. Trans. GIS 11 (1), 29–45.
Elhorst, P., 2010. Spatial panel data models. In: Fischer, M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer, New York.
EPA, 2010. In: O. o. R. a. D. U. S. E. P. Agency (Ed.), Integrated Climate and Land-Use Scenarios (ICLUS) V1.3 user's manual: arcgis tools and datasets for modeling US housing density growth. Environmental Protection Agency, Washington, DC.
EPA, 2013. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2011. United States Environmental Protection Agency, Washington, DC.
Ewing, R., Cervero, R., 2010. Travel and the built environment. J. Am. Plan. Assoc. 76 (3), 265–294.
Fire, C., 2011. Multisource Land Ownership Database. Cal Fire, Sacremento, CA.
Flint, A., Flint, L., 2012. Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. Ecol. Process. 1 (2.).
FRAP, 2007. Fire Hazard Severity Zone Re-Mapping Project, Retrieved from: http://frap.fire.ca.gov/projects/hazard/fhz.php
Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Wickham, J., 2011. Completion of the 2006 national land cover database for the conterminous United States. PE & RS 77 (9), 858–864.
Fu, A., 2012. The facade of safety in California's shelter-in-place homes: history, wildfire, and social consequence. Crit. Sociol., 1–17.
Gardner, E., Kimpel, T., Zhao, Y., 2012. In: O. o. F. Management (Ed.), American Community Survey – User Guide. Office of Financial Management, State of Washington.
Gebert, K., Calkin, D., Yoder, J., 2007a. Estimating suppression expenditures for individual large wildland fires. Western J. Appl. Forest. 22 (3), 188–196.
Gebert, K.M., Calkin, D.E., Yoder, J., 2007b. Estimating suppression expenditures for individual large wildland fires. Western J. Appl. Forest. 22 (3), 188–196.
Gude, P., Rasker, R., Noort, V.D., 2008a. Potential for future development on fire-prone lands. J. Forest. 106 (4), 198–205.
Gude, P., Rasker, R., Van Den, N., 2008b. Potential for future development on fire-prone lands. J. Forest. 106 (4), 198–205.
Hammer, R.B., Radeloff, V.C., Fried, J.S., Stewart, S.I., 2007. Wildland–urban interface housing growth during the 1990s in California, Oregon, and Washington. IJWF 16 (3), 255–265.
Hammer, R.B., Stewart, S.I., Winkler, R.L., Radeloff, V.C., Voss, P.R., 2004. Characterizing dynamic spatial and temporal residential density patterns from 1940–1990 across the North Central United States. Landsc. Urban Plan. 69 (2–3), 183–199.
Hansen, A.J., Knight, R.L., Marzluff, J.M., Powell, S., Brown, K., Gude, P.H., Jones, K., 2005. Effects of exurban development on biodiversity: patterns, mechanisms, and research needs. Ecol. Appl. 15 (6), 1893–1905.
Heimlich, R.E., Anderson, W.D., 2001. Development at the Urban Fringe and Beyond: Impacts on Agriculture and Rural Land Agricultural Economic Report No. 803. U.S. Department of Agriculture, Economic Research Service, Washington, DC.
Hickman, J., 1993. The Jepson Manual: Higher Plants of California. University of California Press, Berkeley, CA.
Holloway, G., Lacombe, D., LeSage, J., 2007. Spatial econometric issues for bio-economic and land-use modelling. J. Agric. Econ. 58 (3), 549–588.
Irwin, E.G., Geoghegan, J., 2001. Theory, data, methods: developing spatially explicit economic models of land use change. Agric. Ecosyst. Environ. 85 (1–3), 7–24.
Jacoby, H., 2000. Access to markets and the benefits of rural roads. Econ. J. 110 (465), 713–737.
Jantz, C., Goetz, S., Donato, D., Claggett, P., 2010. Designing and implementing a regional urban modeling system using the SLEUTH cellular urban model. Comput. Environ. Urban Syst. 34 (1), 1–16.
Lambin, E., Wainwright, J., Mulligan, M., 2004. Modelling Land-Use Change. In: Anonymous (Ed.), Environmental Modeling: Finding Simplicity in Complexity. John Wiley & Sons, New York.
Landis, J.D., Zhang, M., 1998. The Second Generation of the California Urban Futures Model. Part 2: Specification and Calibration Results of the Land-use Change Submodel.
Levy, J.I., Buonocore, J.J., Von Stackelberg, K., 2010. Evaluation of the public health impacts of traffic congestion: a health risk assessment. Environ. Health 9 (65.).
Liang, J., Calkin, D., Gebert, K., Venn, T., Silverstein, R., 2008. Factors influencing large wildland fire suppression expenditures. Int. J. Wildland Fire 17 (5), 650–659.
Library, C.S.I., 2008. Linear Hydrologic Features. State of California, Sacramento, CA.
Mann, M., Kaufmann, R., Bauer, D., Gopal, S., Nomack, M., Womack, J., Soares-Filho, B., 2014. Pasture conversion and competitive cattle rents in the Amazon. Ecol. Econ. 97, 182–190.
Mann, M.L., Kaufmann, R.K., Bauer, D., Gopal, S., Vera-Diaz, M.D.C., Nepstad, D., Amacher, G., 2010. The economics of cropland conversion in Amazonia: the importance of agricultural rent. Ecol. Econ. 69, 1503–1509.
Maranghides, A., Mell, W., 2009. A Case Study of a Community Affected by the Witch and Guejito Fires. US Department of Commerce, Gaithersburg, MD.
McGranahan, D.A., 1999. Natural Amenities Drive Population Change. Agricultural Economics Report No. 781. US Department of Agriculture, Economic Research Service, Washington, DC.
Military, U., 2011. United States Military Installations Database Scott Air Force Base. Military Surface Deployment and Distribution Command Transportation Engineering Agency, IL.
NCED, 2012. National Conservation Easement Database. Multiple: Multiple.
Nelson, G., De Pinto, A., Harris, V., Stone, S., 2004. Land use and road improvements: a spatial perspective. Int. Reg. Sci. Rev. 27 (3), 297–325.
Newburn, D.A., Berck, P., 2006. Modeling suburban and rural-residential development beyond the urban fringe. Land Econ. 82 (4), 481.
NRC, 2010. In: C. o. M. f. E. G. G. Emissions (Ed.), Verifying Greenhouse Gas Emissions: Method to Support International Climate Agreements. National Research Council, Washington, DC.
Olmstead, S., Kousky, C., Sedjo, R., 2012. Wildland Fire Suppression and Land Development in the Wildland/Urban Interface. Joint Fire Science Program, pp. 1–14.
OSFM, 2009. In: O. o. t. S. F. Marshal (Ed.), Task Force Recommendations on Revisions to Building Code Chapter 7a – Materials and Construction Methods for Exterior Wildfire Exposure.
Pfaff, A., 1999. What drives deforestation in the Brazilian Amazon. J. Environ. Econ. Manage. 37, 26–43.
Pijanowski, B., Brown, D., Shellito, B., Manik, G., 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. Comput. Environ. Urban Syst. 26 (6), 553–575.

Radeloff, V.C., Hammer, R.B., Stewart, S.I., Fried, J.S., Holcomb, S.S., McKeefry, J.F., 2005. The wildland–urban interface in the United States. Ecol. Appl. 15 (3), 799–805.
Register, F., 2001. Federal Register Notes.
Shrank, D., Lomax, T., 2012. In: T. T. Institute (Ed.), The 2012 Urban Mobility Report. Texas A&M University, College Station, TX.
Smirnov, O.A., 2010. Modeling spatial discrete choice. Reg. Sci. Urban Econ. 40 (5), 292–298.
Theobald, D.M., 2001. Land-use dynamics beyond the american urban fringe. Geogr. Rev. 91 (3), 544–564.
Theobald, D.M., 2005. Landscape patterns of exurban growth in the USA from 1980 to 2020. Ecol. Soc. 10 (1), 32.
Theobald, D.M., Romme, W.H., 2007. Expansion of the US wildland–urban interface. Landsc. Urban Plan. 83, 340–354.
Veldkamp, A., Fresco, L., 1996. A conceptual model to study the conversion of land use and its effects. Ecol. Model. 85, 253–270.
Walker, W., Gao, S., Johnston, R., 2007. UPlan: geographic information system as framework for integrated land use planning model. J. Transport. Res. Board 1994 (1), 117–127.