Received: 4 November 2021 | Accepted: 9 January 2023

DOI: 10.1111/1752-1688.13131

AMERICAN WATER RESOURCES ASSOCIATION JAWRA JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

RESEARCH ARTICLE

The Chesapeake Bay Land Change Model: Simulating future land use scenarios and potential impacts on water quality

Peter R. Claggett | **Labeeb Ahmed** | **Frederick M. Irani** | **Sarah McDonald** | **Renee L. Thompson**

U.S. Geological Survey, Lower Mississippi-Gulf Water Science Center, Annapolis, Maryland, USA

Correspondence
Peter R. Claggett, U.S. Geological Survey, Lower Mississippi-Gulf Water Science Center, 1750 Forest Drive, Suite 130, Annapolis, Maryland 21401, USA.
Email: pclaggett@usgs.gov

Funding information
USEPA National Center for Environmental Research, Grant/Award Number: DW-14-95813701

Abstract

The Chesapeake Bay Land Change Model (CBLCM) is an open-source pseudo-cellular automata land change model tailored for loose coupling with watershed models. The CBLCM simulates infill development, residential and commercial development, natural land and agricultural land conversion, and growth served by sewer or septic wastewater treatment. The CBLCM is unique among land change models by simulating multiple types of development and explicitly accounting for infill development and the spatial patterns of development densities. The CBLCM was used to simulate five future land use scenarios, holding population constant, for all counties within and adjacent to the Chesapeake Bay watershed from 2013 to 2055. Results are presented here for the state of Maryland over the period 2013–2025 to illustrate model functionality and validation. The growth management (GM) scenario achieved the least development and potential impacts to natural and agricultural lands while accommodating the same amount of population growth as the other four scenarios. Scenarios focusing exclusively on natural or agricultural land protection shifted development to unprotected areas resulting in unforeseen water quality consequences. Simultaneously achieving more compact development while protecting the most valued natural and agricultural lands requires a combination of GM and land conservation policies and actions.

KEYWORDS
urban growth, land use change, Chesapeake Bay, water quality, future scenarios

1 | INTRODUCTION

"All models are wrong, but some are useful" (Box, 1976). This aphorism is frequently referenced by modelers, but usually in private. The idea that all models are wrong particularly applies to land change models (Pontius et al., 2008) and belies the challenge of accurately predicting future landscape conditions based primarily on satellite images of the past. Land change modelers have increasingly focused attention on improving the accuracy of their models (Pontius et al., 2018), but they have paid less attention to making models more useful for which accuracy is but one aspect and not necessarily the most important one (Brown et al., 2013; Sohl & Claggett, 2013). The Chesapeake Bay Land Change Model (CBLCM) was developed with usefulness as the primary objective and therefore simulates changes in urban land use (e.g. residential

Paper No. JAWR-21-0151-P of the *Journal of the American Water Resources Association* (JAWR).

Discussions are open until six months from publication.

Research Impact Statement

Land use planning policies and land conservation actions have an important role to play in the long-term restoration of large aquatic ecosystems. Planning and conservation are not usually considered to be best management practices for restoring water quality because they do not result in immediate pollutant reductions. Viewed from a future perspective, however, they are arguably the least expensive and most effective means of reducing pollution. The CBLCM was designed to produce the future land use and fine-scale population distribution data needed by watershed models to quantify the potential water quality impacts of alternative future scenarios. This has been achieved for the Chesapeake Bay watershed, enabling several states to take credit for the contributions of their planning and conservation actions toward reducing future pollution.

and commercial development) in addition to urban land cover (e.g. impervious surfaces). This is the most fundamental and much overlooked difference between it and other regional-scale land change models. It is a critical difference because the nutrient inputs from herbaceous vegetation can vary by orders of magnitude depending on whether the land is cropland, pasture, turf grass, or forest clearings undergoing natural succession.

The Chesapeake Bay watershed encompasses 156,000 square kilometers (km^2) and has the highest land area to water volume ratio of any coastal estuary in the world (Van der Leeden, 1975). This attribute makes the Chesapeake Bay (the Bay) watershed particularly vulnerable to pollution originating from land surface activities because pollutant concentrations in runoff are not as diluted compared to other coastal estuaries. In 1998, the US Environmental Protection Agency (USEPA) listed the Bay and 90% of its tidal tributaries as impaired due to excessive nitrogen, phosphorus, and sediment pollution, most of which results from historic and ongoing land use and management practices (USEPA, 2010). In 2003, scientists estimated that if recent trends continue, "the area of developed land in the watershed will increase by more than 60% by 2030" (Boesch & Greer, 2003). In 2007, the USEPA's Inspector General found that "new development is increasing nutrient and sediment loads at rates faster than restoration efforts are reducing them" (USEPA, 2007). In 2009, the USEPA established total maximum daily load (TMDL) regulatory limits on the quantity of pollutants and sediment entering the Bay to meet water quality standards (USEPA, 2010). According to the TMDL, federal, state, and local government agencies are required to reduce nitrogen, phosphorus, and sediment pollution entering the Bay by 25%, 24%, and 20%, respectively, by the year 2025 and maintain those reductions in perpetuity. In developing their strategies for reducing pollution, state and local governments are required to account for anticipated growth in pollution sources. All land uses yield some level of nutrient and sediment pollution following a rain event. Cropland, specifically corn, is one of the highest polluting land uses for nutrients whereas forest is the lowest. Lands undergoing construction contribute the most sediment per unit area.

Anticipating the information needed to account for growth, the U.S. Geological Survey and University of Maryland began testing and employing land change models in the Bay watershed to inform restoration strategies (Claggett & Jantz, 2004). Following extensive reviews of available regional-scale land change models (USEPA, 2000), the Chesapeake Bay Program (CBP) Partners selected the Slope, Land Cover, Exclusion, Urbanization, Transportation, and Hillshade (SLEUTH) urban growth model (Clarke et al., 1997) as the most feasible and applicable for implementation in the watershed. The calibration and validation processes in SLEUTH were customized for use in the Bay watershed and the model was calibrated separately in 15 unique regions in the watershed to account for the diversity of urban patterns and rates of change (Jantz et al., 2010).

In the process of implementing the SLEUTH model and reviewing similar models such as CLUE-S (Verburg et al., 2002) and DINAMICA (Soares-Filho et al., 2002), four limitations of these cellular automata models became apparent. First, urban land is simulated as a single "urban" land cover class. Urban environments are diverse consisting of roads, residential, commercial, industrial, institutional, and recreational land uses which are only partially represented by urban land cover. Datasets such as the National Land Cover Database (NLCD) (Dewitz & U.S. Geological Survey, 2021; Homer et al., 2004, 2015) have four developed classes defined by their level of imperviousness. The *Developed Open Space* class has the lowest levels of impervious surfaces, less than 20% within a 30-m cell, but even this class fails to capture the majority of low-density residential development which is dominated by turf grass and trees over turf grass (Claggett et al., 2013). Moreover, aggregating the diverse components of urban land use into a single class limits a modeler's ability to interpret, understand, and communicate observed patterns of change, as well as density. For example, large infrastructure projects are readily detectable as urban development in 30-m resolution satellite imagery but are rare and episodic events, extrapolating them into the future as regularly recurring events would be erroneous. Second, due to the moderate spatial resolution of Landsat, Indian Remote Sensing, Sentinel, and other satellites used for producing regional-scale land cover products and the necessity to standardize classification rules across broad regions, such products only capture a fraction of low-density residential areas and narrow features such as two-lane rural roads (McCauley & Goetz, 2004). Third, sole reliance on moderate-resolution remotely sensed imagery for model calibration and validation ignores infill and redevelopment,[1] hereafter referred to as "infill,"

[1]For the purposes of this paper, "infill" refers to the development of unused or under-utilized lots within developed areas and includes new construction on previously developed lots which is often referred to as "redevelopment" and the occupation of vacant buildings and the densification of residents and employees within existing buildings.

which can accommodate a large portion of population growth, one of the main drivers of urbanization. The development of small vacant lots within urbanized areas and new construction in already developed areas are not easily detectable in spectral image products unless such changes correspond to changes in surface spectral properties. Fourth, the allocation of growth in SLEUTH is optimized based on a set of parameters calibrated on recent observations of urban growth. Optimization reflects a high level of confidence that the assumptions, parameters, and policies influencing recent trends apply to future conditions. While optimization may reflect reality in regions where zoning laws and ordinances are strictly enforced and do not change over time, most states and local governments in mid-Atlantic region do not have such controls in place and even those that do, such as Maryland, may lack effective implementation (Lewis et al., 2009) or grant frequent variances.

The relationship between population growth and urban development is complex and varies spatially and temporally (Meyer & Turner, 1992). While development may occur at disproportionate and asynchronous rates compared to changes in population and employment, utilizing population and employment as drivers of urbanization is common practice in regional-scale land change models. A combination of population and jobs at place of work were recently used to represent drivers of urban growth in the greater Baltimore area (Jantz et al., 2014). The Integrated Climate and Land Use Scenarios model relates changes in county-level population to changes in various population density classes corresponding to different urban land uses (USEPA, 2009; USEPA, 2017), and the FUTURES (Meentemeyer et al., 2013) model uses the ratio of change in population to change in urban land cover to develop an aerial demand for growth. The CBLCM uses exogenous projections of county-level growth in both population and employment to determine demand for future residential and commercial development, respectively. In contrast to these other models, the CBLCM accounts for the portion of population and employment growth accommodated by infill and redevelopment and estimates maximum expected development densities, both of which contribute to a more direct translation of demand to an aerial developed footprint.

The USGS developed the CBLCM to address some of the limitations of existing models and to enable stakeholders to account for growth in their watershed implementation plans, as required by the TMDL, through the simulation of plausible future development scenarios that reflect trends but also account for the effects of land use and land conservation policies and actions. More specifically, the CBLCM was designed to:

- Utilize exogenous population and employment forecasts as the drivers of development.
- Explicitly account for infill and redevelopment.
- Simulate multiple types of developed land use (e.g., residential, commercial, and mixed use).
- Simulate a range of development densities.
- Simulate the effects of a broad variety of land use and land conservation policies and actions.
- Simulate development anywhere in the conterminous United States based on nationally available datasets.

2 | METHODOLOGY

The CBLCM is coded in Python and R (R Core Team, 2013) as an independent pseudo-cellular automata model that translates exogenous projections of population and employment into patches of residential and commercial development on the landscape after accounting for infill and redevelopment (Figure 1). The CBLCM produces 101 stochastic Monte Carlo simulations of future growth for any given set of assumptions (e.g., a future scenario) and utilizes county-level projections of population and employment as the sole drivers of urban growth. The CBLCM simulates three types of growth (residential, commercial, and mixed use); translates these land classes into estimates of impervious surface, turf grass, and other urban subclasses based on local development densities; accounts for infill/redevelopment; and relies on fine-scale data on changes in housing, employment, and land cover as spatial predictors of development. Each CBLCM simulation represents an equally plausible independent representation of future development patterns. The Monte Carlo simulations can be averaged within overlay zones such as the segmentation used in watershed models to represent the most plausible outcome for each scenario while also enabling the quantification of spatial uncertainty. The amount of simulated growth is controlled by the exogenous population and employment projections, estimated infill/redevelopment rates, and estimated housing and employment densities. By explicitly representing housing and job densities for every simulated residential and commercial patch of growth, future development can be reclassified consistent with local land use and zoning definitions (e.g., low-density residential with 0.2–2.0 dwelling units per hectare) and consistent with the specific requirements of TMDL watershed models (e.g., impervious cover, turf grass, tree cover).

The CBLCM modeling domain consists of an area of interest (AOI) and region of interest (ROI), both representing sets of counties with the AOI nesting within the ROI. The purpose of the ROI is to capture all major metropolitan areas potentially influencing the AOI. For example, New York City is well outside the Bay watershed but influences land development in the northeastern portion of the watershed. To evaluate the utility of the CBLCM for informing land conservation and planning decisions in the Bay watershed, the ROI includes all states and the District of Columbia intersecting the watershed and the AOI includes all 206 counties within, intersecting, and adjacent to the Bay watershed (Figure 2). This paper focuses on the state of Maryland for model illustration and validation purposes. Maryland represents a longitudinal cross

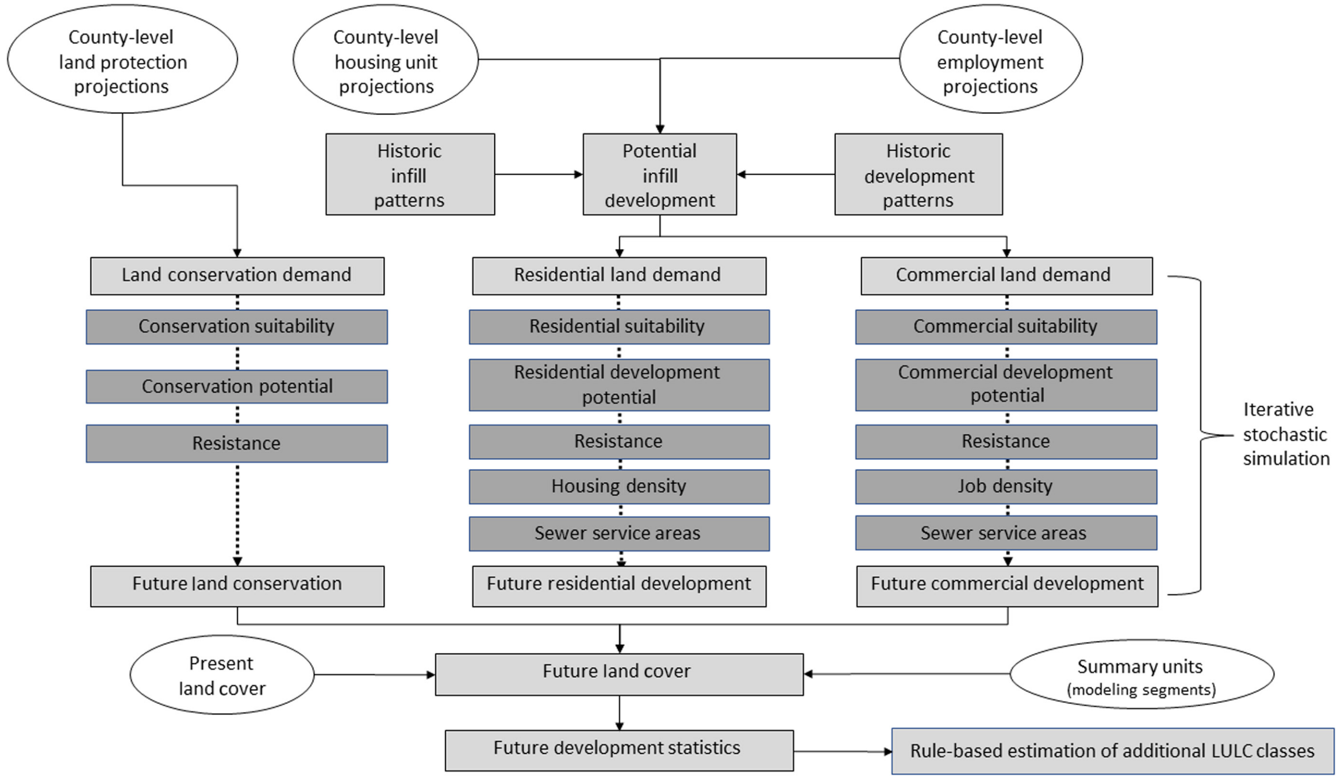

FIGURE 1 Overview diagram of the Chesapeake Bay Land Change Model (CBLCM). Non-shaded circles represent exogenous data. Light-shaded rectangles represent model functions. Dark-shaded rectangles represent the simulation aspects of the model.

section of all major physiographic provinces (i.e., Coastal Plain, Piedmont, Ridge and Valley, and Appalachian Plateau) in the Bay watershed, and it is the only state in the region with timeThe citationseries land use data for use in validating the CBLCM. For each of the five scenarios simulated, the areas of impervious surfaces, pervious developed surfaces, natural land loss, agricultural land loss, and population growth on sewer and septic systems were assessed for every land-river modeling segment used in the CBP Watershed Model (Shenk & Linker, 2013) and Chesapeake Assessment and Scenario Tool (CAST; CBP, 2020). CAST is essentially a static spreadsheet representation of the fully dynamic CBP Watershed Model. CAST was developed to rapidly simulate water quality scenarios without requiring time-consuming runs of the dynamic watershed model. Land-river segments represent the combination of small watersheds (e.g., 12-digit Hydrologic Unit Codes), county boundaries, and orographic rainfall pattern boundaries. Average annual nitrogen, phosphorus, and sediment loads delivered to streams were assessed and compared for each of the five scenarios in Maryland for the years 2013–2025.

Because the CBLCM was originally designed to provide future land use conditions as inputs to CAST, this objective dictated the required land use class outputs from the CBLCM (Table 1), the scale of aggregation units (roughly equivalent to 12-digit Hydrologic Unit Codes), and the temporal resolution (10- to 12-year increments) of the outputs. The CBLCM requires significantly more raw input data than regional land change models such as SLEUTH, but all required data are available for the conterminous United States (Data S1). All spatial data inputs were resampled to 30 m raster images aligned to the NLCD. The translation of all raw input data for the ROI, such as Census Blocks (US Census Bureau, 2010a, 2010b), into data structures with attributes required for running the CBLCM simulation module has been automated using a set of scripts collectively referred to as the preprocessor module. The simulation module executes the Monte Carlo runs and includes a set of scripts that summarize the outputs for user-defined overlay geographies such as watersheds, Hydrologic Unit Codes, Census Tracts, and the modeling segments used in CAST. The variety of outputs produced by the CBLCM is outlined and discussed in Data S2.

To describe how the CBLCM simulates future greenfield development (i.e., the conversion of natural and agricultural lands to development, hereafter referred to as "development"), a botanical metaphor is used that includes the concepts of "planting," "sprouting," and "maturation." Given exogenous demands[2] for new jobs and housing per county after accounting for infill, planting locations for new residential and commercial development seeds are selected at random from the universe of relative residential or commercial development potentials represented by 30 m resolution cells suitable for development. Suitable cells include those that are undeveloped, unprotected, gently sloped, and zoned for development. These cells are weighted by their relative development potentials derived from separate residential and commercial logistic regression analyses. For these analyses, all Census Blocks within Maryland were classed as either residential or commercial by their relative

[2]Exogenous demands for jobs and housing were derived from state or state-sponsored population and/or employment projections.

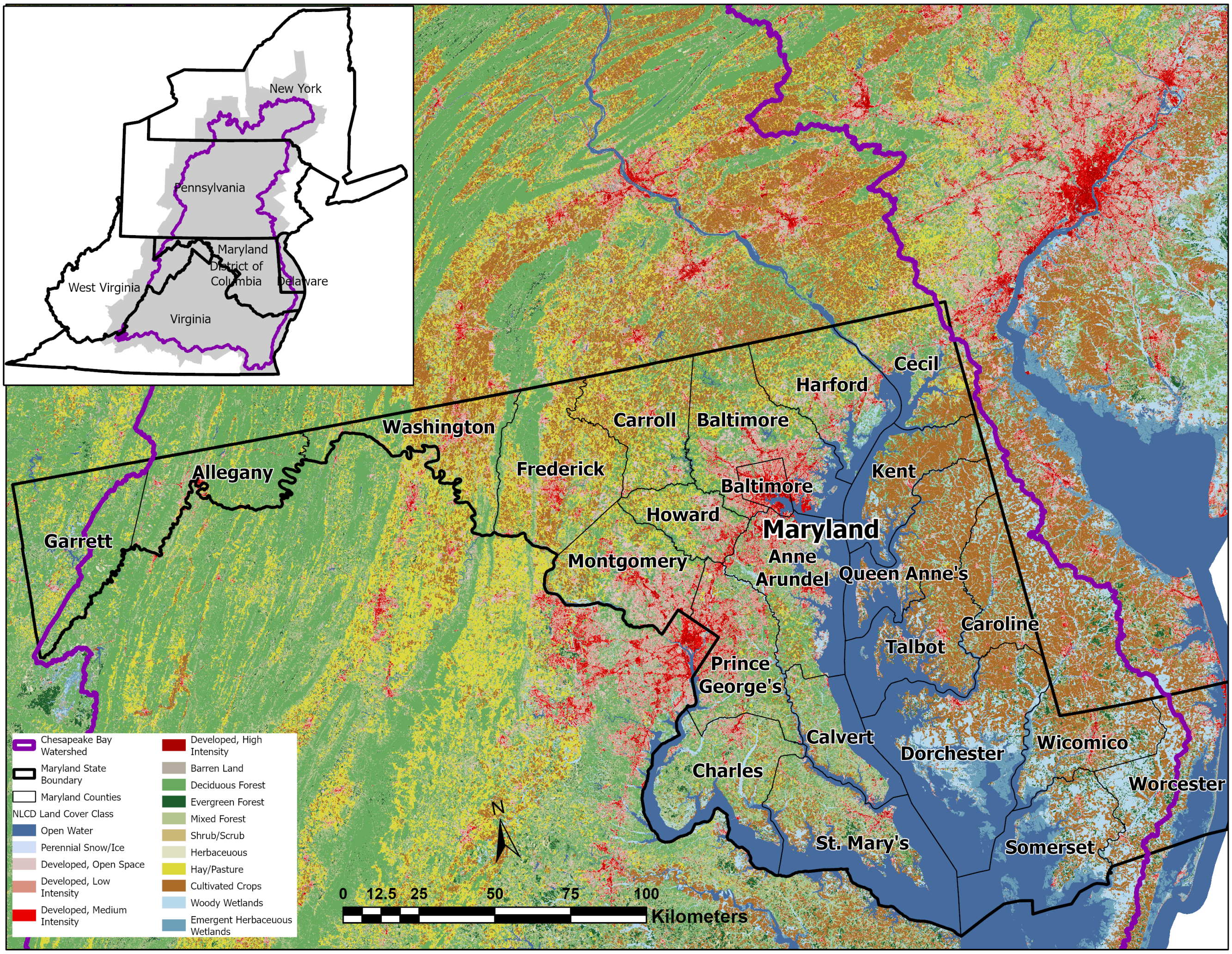

FIGURE 2 Maryland counties overlaid on the 2016 National Land Cover Database (Dewitz, 2019). The region of interest (mid-Atlantic states) and area of interest (shaded) are shown in the inset map.

amounts of housing units and jobs. Blocks with housing units greater than or equal to the number of jobs were classed as residential, all others were classed as commercial. Within these residential and commercial areas, impervious surface change cells derived from differencing the 2001 and 2011 NLCD were randomly sampled to generate independent point locations of observed development. The same number of points were randomly allocated to suitable undeveloped areas as inputs to the regressions. In each state, the same eight potential explanatory variables (Table 2) were evaluated and only significant variables, with *p*-values over 0.05, were retained for each analysis.

After a suitable seed location is identified, it is assigned a potential patch size selected at random from the cumulative frequency distribution of observed residential and commercial patch sizes in each state associated with growth over the previous decade. The seed then sprouts and matures into a patch of cells. Patch growth proceeds on a cell-by-cell basis, starting with the cell adjacent to the initial seed cell that has the lowest relative resistance to growth. Resistance is a function of proximity to roads and distance from the seed cell for each patch (i.e., closer = lower resistance). Resistance values control the shape of new patches which will be circular in rural areas far from roads and barriers to growth (e.g., slope, marshes, and open water) and more irregular or elongated shaped patches near roads and existing development. As residential and commercial patches are grown on the landscape, housing units or jobs are accumulated from the underlying housing and employment density raster datasets.

Residential and commercial development densities were estimated for all census blocks in Maryland for the year 2010. For the smallest blocks, less than 10 acres, developed area was equated to the area of each block. For all other blocks, developed area was based on the 2011 NLCD (combined area of the four developed classes: 21–24) or equated with the suitable land area in the block if it was completely undeveloped. Initial residential densities by block were estimated as the total number of housing units reported divided by developed area. Similarly, initial commercial densities by block were estimated as the total number of jobs (US Census Bureau, 2016), averaged over the period 2009–2011, divided by the total developed area. After estimating initial densities, each block was assigned the maximum density of

TABLE 1 CBLCM simulated land uses cross-walked to the required Chesapeake Assessment Scenario Tool (CAST) mapped land uses.

CBLCM interpreted land uses	CAST mapped land uses
Impervious surfaces	Impervious roads
	Impervious, non-roads
	Tree Canopy over impervious
Turf Grass	Turf grass
Tree Canopy over Turf Grass	Tree canopy over turf grass
Natural land conversion	Forest
	Non-tidal floodplain wetlands
	Non-tidal other wetlands
Herbaceous land conversion	Open space
	Cropland (agriculture)
	Pasture (agriculture)
Not simulated	Water
Not simulated	Wetlands tidal

TABLE 2 Spatial explanatory variables used to estimate development potential for all five scenarios. All distances were estimated based on travel time over the existing road network at posted speeds (HERE, 2013). Off-road travel cells were assigned a travel speed of 8 km per hour. All continuous travel time variables were log transformed.

Variable	Explanation
Distance to urban areas and clusters	Travel time to designated urban areas and urban clusters (US Census, 2010b)
Distance to residential growth hot spots	Travel time to residential growth hot spots, represented by census blocks exceeding the 75th percentile of housing unit increases throughout Maryland from 2000 to 2010 (US Census Bureau, 2010a).
Distance to commercial growth hot spots	Travel time to commercial growth hot spots, represented by census blocks exceeding the 75th percentile of job increases throughout Maryland from 2003 to 2010 (US Census, 2016).
Slope	Percent slope derived from a 30-m digital elevation model (US Geological Survey, 2019)
Pre-development land use (natural or agriculture)	2011 National Land Cover Dataset reclassed as a binary dataset with natural lands defined as classes 41–43, 52, 71, 90, and 95 and agriculture defined as classes 81 and 82 (Dewitz, et al., 2021).
Distance to regional employment centers	Travel time to regional employment centers, represented by census blocks with greater than 3.2 jobs per square kilometer.
Distance to regional waterbodies and amenities	Travel time to large waterbodies (e.g., lakes, reservoirs, estuaries, and ocean) and to public parks and open space (Moore et al., 2019; US Geological Survey, 2018)
Distance to regional sewer service areas	Travel time to areas served by municipal wastewater treatment plants (Chesapeake Bay Program (CBP), 2020)

itself and its adjacent neighbors to ensure that future simulated growth would reflect the densest development in any given neighborhood. Minimum densities of 1 housing unit per 10 hectares and 1 job per 0.4 hectare were enforced for all blocks and were assigned to blocks lacking any housing unit or job data. All densities were converted to housing units or jobs per 30 m cell to produce the final residential and commercial density raster datasets. The average density for each patch was coupled with impervious surface land use coefficients developed by Cappiella and Brown (2001) to translate the areas of residential and commercial growth into impervious and pervious surface estimates required by CAST.

New residential and commercial patches stop growing when one of the following three conditions is met: (1) it reaches its assigned patch size; (2) it reaches its attainable size (e.g., the assigned size may be unattainable because cells surrounding the seed are unsuitable for growth); or (3) housing demand for the county is reached. At the end of each time step, the housing and employment density rasters can be magnified by a density scalar to represent trends toward densification or other scenario storyline features. When used, the density scalar is multiplied by the development potential to weigh the increase in densities near urban areas more relative to rural areas. Magnifying the densities results in less overall growth in future decades given a constant demand for growth. The number of stochastic iterations per time step and length of time step are defined by the user. For each future scenario in Maryland, 101 Monte Carlo iterations were run from 2013 to 2025. This number of iterations proved sufficient to yield consistent average results and variability.

Not all projected increases in households and/or jobs, however, will result in new development. A portion of new households and jobs will be accommodated within existing urbanized areas as infill and therefore not result in the conversion of natural or agricultural lands to development. The percentages of housing and employment accommodated through infill were estimated for census blocks within urban areas and places (US Census Bureau, 2010a). Maximum expected neighborhood densities were used to estimate the minimum amount of impervious cover associated with observed increases in housing or jobs over the 2000s. These estimates were then compared to observed changes in NLCD impervious cover from 2001 to 2011 for residential infill and from 2006 to 2016 for commercial infill to estimate the portion of each county's increases in housing and/or jobs that are unaccounted for by observed growth and therefore assumed to be accommodated for by infill. As a result, infill as estimated for this study represents a direct measure of NLCD under-detection of areal changes in impervious cover in developed areas. Relying on the NLCD for this purpose required defining areas eligible for infill and redevelopment because NLCD under-detection of areal changes in impervious cover in rural areas may be associated with the dispersed, low-density nature of rural development. After assessing infill proportions for all counties, demands for new residential and commercial development were estimated by subtracting the portion of household and employment growth accommodated by infill from the total projected growth in households and jobs.

To be accepted for use by the CBP partners, the CBLCM must be transparent, validated, peer-reviewed, explicit in its quantification of uncertainty, uniquely parameterized for each state, and rely on state-sanctioned population and/or employment projections. Model transparency may be the most important criterion for decision-maker acceptance (Waddell, 2011). Achieving transparency required frequent communication with stakeholders to ensure that the model is understood and to involve them in the process of model development. It also required simulating land use change at a fine-enough resolution (30 m cells) that the forecasted patterns of growth could be visually critiqued. Ensuring that the model be easily understood led to the decision to use logistic regression as the means for producing gridded development potential surfaces, as opposed to using artificial neural networks or Bayesian methods, due to widespread familiarity with correlation analysis. Stakeholders were briefed on the model development process and provided feedback and suggestions several times per year over the 8-year model development period. Once initial future simulations of urbanization were produced, a workshop was held in 2016 to review and critique the results with representatives from each of the six states in the Chesapeake Bay watershed and the District of Columbia. Prior to this, meetings were held with the planning staff from a rapidly urbanizing county, Howard County, Maryland, and a rural county, Talbot County, Maryland in 2012 and 2015. For both Talbot and Howard Counties, local data for characterizing current urban growth conditions, drivers of growth, and constraints on growth were requested. In addition, several meetings were held with staff from Maryland's Department of Natural Resources, Department of the Environment, and Department of Planning. The data collected from these agencies included zoning, priority funding areas, transit stations, sewer service areas (existing and proposed), land use comprehensive plans (e.g., proposed future densities or zoning), special protection areas (e.g., critical areas, riparian buffers, erosion prone soils, flood zones), protected lands, agricultural districts, conservation priorities (i.e., farms, natural lands, and wetlands), and planned transportation improvements. Consulting with these agencies provided perspective on how growth patterns have occurred and might change in rural and urban settings.

3 | SCENARIOS

A future scenario represents a set of model inputs and plausible assumptions consistent with a logically coherent storyline (National Research Council, 2014). Five alternative future scenarios were developed to evaluate model performance, explore the effects of exclusions such as zoning and land protection, and to provide a meaningful set of benchmarks for comparison against stakeholder-generated scenarios. The five scenarios include: “Historic trends” (HT), “Current zoning” (CZ), “Growth management” (GM), “Forest conservation” (FC), and “Agricultural conservation” (AC). The storylines, input data, and modeling assumptions for each scenario are described below. Each scenario is based on the same set of exogenous population and employment projections to facilitate comparison based on the impacts of land use policy and land conservation assumptions associated with each scenario. As discussed below, the HT scenario forms the basis for the CZ scenario which in turn serves as the basis for the GM, FC, and AC scenarios.

3.1 | Historic trends

The HT scenario assumes that the forces, policies, and regulations that influenced development patterns over the 2000s will continue unabated into the future. Under this scenario, historical rates of infill and new development estimated by the CBLCM are perpetuated into the future. Areas developed through 2013, emergent wetlands, open water, steep slopes, and public and protected lands are excluded from future development. Because the early 2000s was a period of rapid, dispersed development, this scenario is the one most reflective of sprawling growth patterns.

3.2 | Current zoning

The CZ scenario builds on the HT scenario by restricting residential and commercial growth to areas explicitly zoned for these uses or mixed-use development where such zoning exists. Development is less restricted in counties with broad or vague zoning categories. The cooperative population and employment forecasts produced by the Washington Metropolitan Council of Governments (COG)[3] are used to exclude future growth within areas projected to grow by 10 or fewer households or jobs between 2010 and 2045. The CBLCM-estimated infill rates are replaced with ones estimated based on the proportion of the Maryland Department of Planning's modeled estimates of the percent of the future population served by sewer from 2010 to 2025. The CBP Partners determined the CZ scenario to be the most plausible and most representative of a "business as usual" condition.

3.3 | Growth management

The GM scenario builds on the CZ scenario and represents a move toward the implementation of smart growth policies that serve to densify and concentrate growth in areas with sufficient infrastructure and services to support it. Infill rates, the proportion of growth in urban areas, and urban housing and job densities are increased by up to 10% at the beginning of the initial time step. These percentage increases are then weighted by the residential and commercial development potentials. Sewer service areas are expanded by 1.6 kilometers and future growth is prohibited on soils unsuitable for septic systems.[4]

3.4 | Forest conservation

The FC scenario also builds on the CZ scenario and represents immediate implementation of aggressive natural land conservation zoning, ordinances, easements, and acquisition. Under this scenario, growth is excluded from all riparian zones, wetlands, areas subject to a 1-m rise in sea level, lands within 1.6 km of National Wildlife Refuges, large natural land tracts (> 101 hectares), lands within 305 m of estuarine and coastal shorelines, and from other natural lands with high conservation values (Chesapeake Conservation Partnership (CCP), 2020).

3.5 | Agricultural conservation

The AC scenario builds on the CZ scenario and represents immediate implementation of aggressive agricultural land conservation zoning, ordinances, easements, and acquisition. Under this scenario, growth is excluded from designated agricultural districts, 100-year floodplains, prime agricultural lands, farmland of state importance, prior converted wetlands, and farms with other high conservation values (Chesapeake Conservation Partnership (CCP), 2020).

4 | WATER QUALITY AND POLLUTANT LOADING

Total nitrogen (TN), total phosphorus (TP), and total suspended sediment (TSS) pollutant loading rates produced by CAST are specific to individual land uses and vary spatially and temporally due to differences in precipitation, slope, soil characteristics, proximity to streams, groundwater nutrient concentrations, fertilizer and manure applications, plant uptake rates, atmospheric deposition, and the number and type of BMP's implemented in an area among other factors (Shenk & Linker, 2013). On average, natural lands and mixed open (unmanaged herbaceous and shrub lands) land uses have the lowest pollutant loading rates (Chesapeake Bay Program (CBP), 2020; Figure 3). Cropland, a component of agricultural land, typically has the highest TN loading rates, but the rates vary significantly depending on crop type and manure and fertilizer application rates. Construction has the highest pollutant loading rate for TP and TSS, but these yields are ephemeral, lasting for just a single year rather than persisting into the future which must be accounted for when examining the net changes in pollutant loads over the simulation period. Pollutant loads from streambank erosion and bed scour are also simulated in CAST and attributed mostly to impervious surfaces due to their direct contribution to runoff volumes and velocities (Chesapeake Bay Program (CBP), 2020). Rather than report these loads separately, they were apportioned back to the major land uses to better understand the water quality implications of land use change.

[3]The COG region includes all of Charles, Montgomery, Prince George's, Anne Arundel, Howard, Carroll, and Frederick counties in Maryland.

[4]Unsuitable soils based on depth to bedrock, drainage class, saturated hydraulic conductivity, and flood frequency.

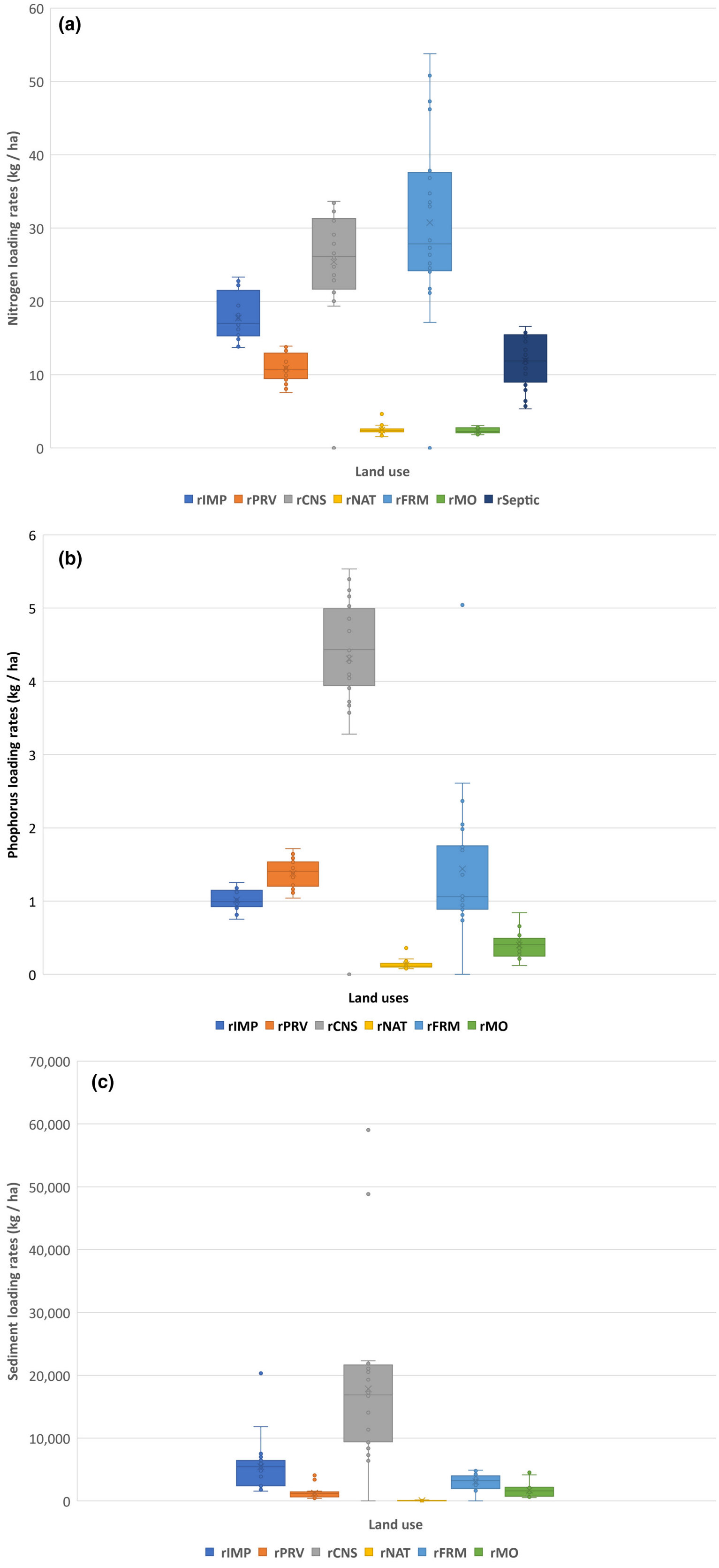

FIGURE 3 Land use and septic system county-level edge-of-stream pollutant loading rates and variability for total nitrogen (a), total phosphorus (b), and total suspended sediment (c; kilograms per hectare) estimated by CAST for the year 2025. Land uses are abbreviated as: rCNS, construction; rFRM, agriculture; rIMP, impervious; rMO, mixed open space; rNAT, natural; rPRV, pervious; rSeptic, individual onsite wastewater treatment systems. Boxes represent the interquartile range of values and whiskers represent the maximum and minimum values for Maryland counties.

To assess differences in pollutant loads associated with each land use scenario, CAST was run on the 2025 CZ scenario land use simulated by the CBLCM. For this CAST run, changes in wastewater and BMPs were excluded to isolate the effects of land use change on pollutant loads from the effects of BMP implementation and treatment plant technologies. The edge-of-stream,[5] county-level average annual land use loads in kilograms per hectare per year (kg/ha/year) for TN, TP, and TSS produced by CAST for the CZ scenario were used as land use pollutant loading coefficients and multiplied by the area of each land use simulated by the CBLCM for each of the other four scenarios.

5 | VALIDATION

The CBLCM is designed to simulate plausible levels of residential and commercial development under alternative future scenarios informed by decadal-scale trends and patterns of urban development. To accurately assess the consequences of urban development on water quality, the amount and density of growth should be valid in addition to the antecedent land use conditions (e.g., natural or agricultural land) and accessibility to public sewer. Validation of an urban land use model based solely on moderate-resolution satellite-derived land use/cover data (e.g., NLCD) is not possible because Landsat-derived land use/cover data do not fully capture low-density residential development (Claggett et al., 2013) nor do they distinguish between residential and commercial growth. The Maryland Department of Planning (MDP) maintains a statewide tax parcel database and statewide land use datasets for 2002 and 2010 which were used in combination with the NLCD to validate the CZ scenario. Maryland is the only state in the Bay watershed that maintains such a database. Information on recent development was extracted from the tax parcel database for the years 2000–2019. Recently developed residential parcels include residential homes, residential condominiums, town houses, or apartments built since the year 2000 with an improvement value greater than or equal to $100,000, lot sizes less than or equal to 10 hectares, and with less than 500 m^2 of impervious surfaces in 2001. Recently developed commercial parcels include commercial, industrial, commercial condominium, exempt commercial, commercial/residential, residential/commercial, and country clubs built since the year 2000 with an improvement value greater than or equal to $100,000 and lot sizes less than or equal to 202 hectares.

Differences between the amount of simulated and observed development may result from differences between projected and observed population and employment, infill rates, and development densities. While projected population and employment are exogenous drivers of growth in the CBLCM, actual population and employment change drive observed changes in land use and differences between them will impact the accuracy of modeled results regardless of infill and other factors influencing land consumption. MDP's population projections used to inform the simulations of future land use under all scenarios were higher for most Maryland counties in 2020 compared to actual changes in population reflected in the population estimates produced by the U.S. Census Bureau for the same year (Table 3).

Infill rates directly impact the amount of simulated residential and commercial development on the landscape. Accommodating 50% of projected growth in housing as infill will reduce the amount of land converted to residential development by roughly half. Data on housing (US Census Bureau, 2010b) and MDP's parcel data were used to validate county-level residential infill rates from 2000 to 2010 and data on new jobs (US Census Bureau, 2016) were used to validate county-level commercial infill development rates from 2006 to 2016. Increases in housing or jobs within a Census Block with no corresponding increase in new residential or commercial parcel development over the same period were categorized as observed infill and divided by the total increases in housing or jobs in each county to produce observed county-level infill rates.

Residential land consumption rates represent the area of residential land use change per year or per new housing unit. Commercial land consumption rates represent the area of commercial land use change per year or per new job. Normalization of the land consumption rates controls for variability in the number of housing units, jobs, and time periods. Simulated residential land consumption rates from 2013 to 2025 were compared to observed residential rates for the periods 2001–2011 and 2011–2019. Simulated commercial land consumption rates from 2013 to 2025 were compared to observed commercial rates for 2006–2016. Observed rates of residential and commercial land consumption were assessed using tax parcel data and data on housing units and jobs (US Census Bureau, 2010b, 2016). The difference in residential and commercial observation periods is related to the limited temporal availability of fine-scale employment data.

To validate residential and commercial development densities, the ratios of modeled to observed parcel-based densities were calculated and compared across all census blocks. In addition, the spatial patterns of modeled and observed development were compared because density varies spatially as does land suitability and development potential. Development densities and potential are typically highest closer to urban areas and transportation corridors. To validate the simulated patterns of new residential and commercial development for the CZ scenario over 101 Monte Carlo iterations, major roads and urban areas were buffered in 1.6-kilometer (km) increments (Figure 4) and the amount of observed and modeled development throughout Maryland was aggregated within these buffer zones.

[5]Edge-of-stream loads represent the amount of pollution entering streams with an average annual flow equal to or greater than 100 cubic feet per second, typical of larger second and most third order streams.

TABLE 3 Difference between projected and estimated 2020 population (persons) by county.

County name	MDP_Projected_2020	Census_Estimated_2020	Difference
Allegany	74,150	70,057	−5.52%
Anne Arundel	580,006	582,777	0.48%
Baltimore	827,758	826,017	−0.21%
Calvert	94,600	93,072	−1.62%
Caroline	34,050	33,492	−1.64%
Carroll	175,900	169,092	−3.87%
Cecil	104,600	103,419	−1.13%
Charles	167,036	164,436	−1.56%
Dorchester	34,300	31,853	−7.13%
Frederick	267,782	265,161	−0.98%
Garrett	30,300	28,852	−4.78%
Harford	250,025	256,805	2.71%
Howard	332,273	328,200	−1.23%
Kent	20,900	19,192	−8.17%
Montgomery	1051,989	1051,816	−0.02%
Prince George's	923,144	909,612	−1.47%
Queen Anne's	48,477	51,167	5.55%
St. Mary's	120,149	114,687	−4.55%
Somerset	26,750	25,453	−4.85%
Talbot	38,850	36,972	−4.83%
Washington	156,800	151,146	−3.61%
Wicomico	106,200	103,990	−2.08%
Worcester	53,100	52,403	−1.31%
Baltimore City	615,813	586,131	−4.82%

6 | RESULTS AND DISCUSSION

6.1 | Forecasted changes in land use

The five land use scenarios simulated for the state of Maryland resulted in a broad range of outcomes (Table 4; Figure 5). The HT scenario produced over twice as much development and associated natural land and agricultural land conversion compared to the remaining four scenarios due to its significantly lower infill/redevelopment rates compared to the other scenarios (Table 5), reinforcing the importance of infill and redevelopment compared to other factors influencing future growth. This is particularly true in counties that have high infill rates such as Montgomery County, Maryland. For example, decreasing the infill rate from 80% to 60% results in a doubling of land demand, whereas decreasing the infill rate from 40% to 20% increases land demand by only 25%. Counties with high infill rates often have relatively high population and employment projections which further amplify the differences in land demand associated with changes in the rates. The additional exclusion of some rural lands in the other four scenarios also influences the quantity of new development by deflecting growth toward higher density areas, but this is a minor influence in counties with ample undeveloped space zoned for low-density development which is common in most counties.

The GM scenario produced the least amount of new development and land conversion because it has the highest infill rates, highest residential and commercial densities, and highest proportion of growth in urban areas relative to the other scenarios. These results support the importance of comprehensive plans, zoning, and ordinances as effective means of minimizing impacts to both natural and agricultural lands if they are followed and enforced. This assumes though that such policies and regulations do not divert demand to neighboring or outlying counties, neither of which are simulated by the CBLCM. Such shifts in development patterns could be avoided through the regional adoption of a common set of land use policies.

Dispersed patterns of growth reflected in the HT scenario, and to a lesser degree in the CZ scenario, have similar proportions of natural to agricultural land converted to development compared to the more concentrated growth patterns in the GM scenario. At the statewide level, the proportions of natural and agricultural land within developable areas are similar among these three scenarios across an urban to rural

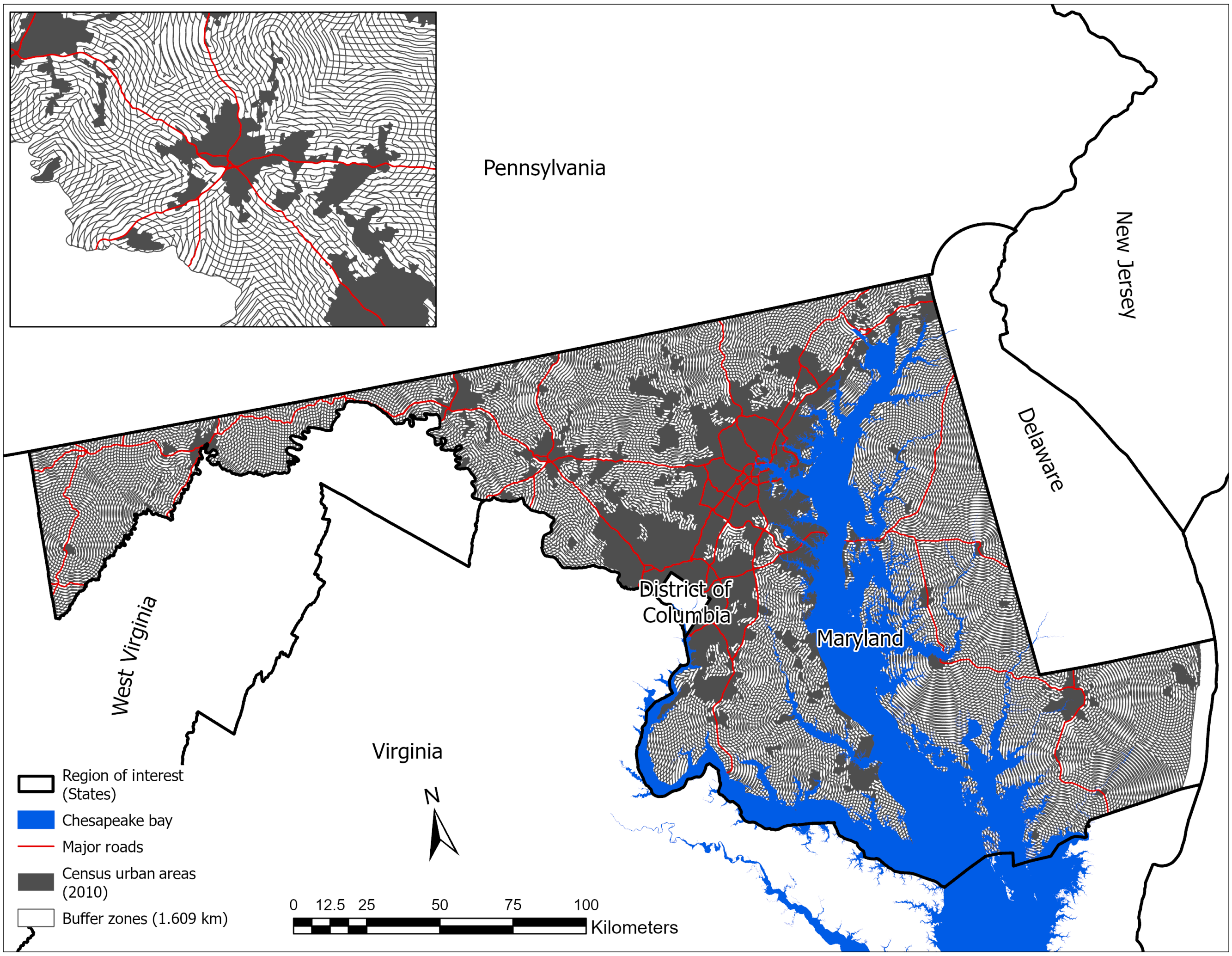

FIGURE 4 Buffer zones (1.609 km) from major roads and Census urban areas used for analyzing development patterns.

TABLE 4 Land use change by scenario (area units in hectares; septic systems represent the number of individual onsite wastewater treatment systems).

Scenario	Impervious	Pervious[a]	Natural	Agriculture	OpenSpace	Septic	Conversion ratio[b]
Historic trends (HT)	8145	20,173	(14,667)	(14,760)	1108	33,339	0.99
Current zoning (CZ)	3468	7842	(6428)	(5360)	477	12,469	1.20
Growth management (GM)	2009	4648	(3620)	(3328)	290	677	1.09
Forest conservation (FC)	3307	8125	(4685)	(7336)	589	12,766	0.64
Agricultural conservation (AC)	3323	7019	(8869)	(1617)	143	11,026	5.49

[a]Includes turf grass and tree canopy over turf grass.

[b]Ratio of natural to agricultural lands converted to development.

gradient. At the county scale, this is not always the case, leading to unanticipated policy impacts. For example, if developable natural lands within a county are primarily located adjacent to urban areas served by sewer, then expanding sewer services could disproportionately impact natural lands. Coupling aggressive natural land protection policies with sewer service expansion would reduce impacts to natural lands but would also deflect growth into lower density, agricultural areas served by septic resulting in higher amounts of impervious cover and turf grass per household without additional zoning constraints.

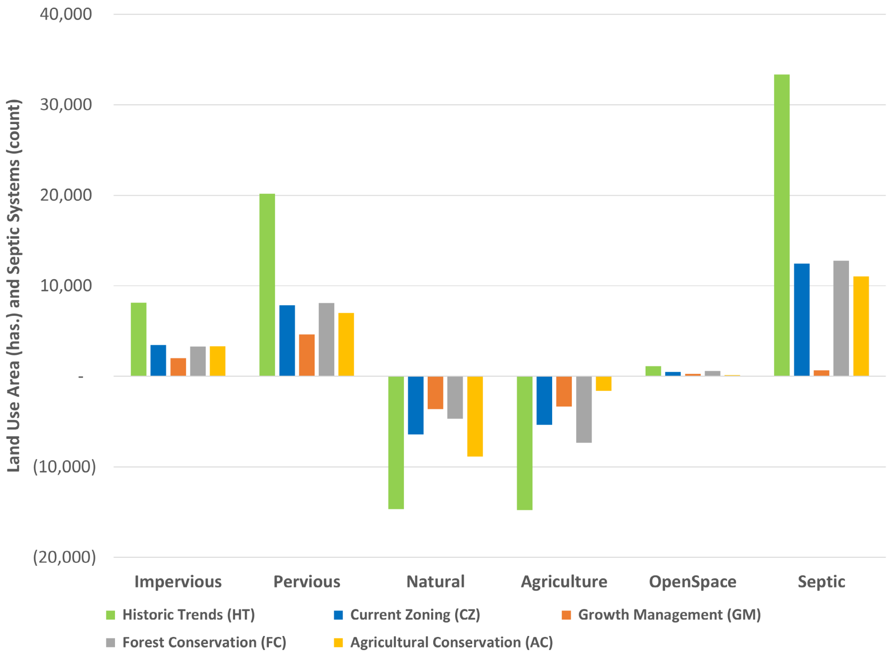

FIGURE 5 Land use change and septic growth in Maryland for the simulation period 2013–2025 in hectares for five scenarios: HT, CZ, GM, FC, and AC. Pervious surfaces include turf grass and trees over turf grass.

The conversion of natural land was highest under the AC scenario while the conversion of agricultural land was highest under the FC scenario. Both the AC and FC scenarios exclude large areas from development, deflecting growth to remaining developable lands which were mostly natural in the AC scenario and mostly farmed in the FC scenario. Land conservation mostly deflects growth rather than diminishes it because the CBLCM allocates growth until the overall demand is met or the land available for development is exhausted and this latter condition only occurs in the most highly developed jurisdictions such as Baltimore City.

The GM scenario minimized the amount of growth served by septic systems compared to all other scenarios because it expands sewer service areas and directs more growth to urban areas which are mostly served by sewer. Urban areas also have higher development densities and lower amounts of impervious surface per household compared to rural areas. For example, residential lot sizes larger than 0.81 hectares (2 acres) are estimated to have 0.10 hectares (0.25 acres) of impervious surface compared to 0.01 hectares (0.028 acres) of impervious surface per household in dense urban neighborhoods.

6.2 | Forecasted changes in water quality

The HT scenario produced the greatest increases in total nitrogen, total phosphorus, and total suspended sediment loads compared to all other scenarios (Table 6) due to its lower infill rates. This is expected because the HT scenario had over twice the amount of development (impervious and pervious) compared to the other four scenarios. The GM scenario produced the lowest increases in pollution because it produced the lowest amount of growth. Across all scenarios, pollutant load increases are highest in counties expected to experience the most overall increase in developed land area. This relationship is moderated by the proportion of natural to agricultural land in high growth areas. Increases in loads due to development can be countered by the conversion of agricultural land to development as modeled in Frederick County. In contrast, Anne Arundel and Charles counties have higher amounts of natural land in areas susceptible to development which amplifies their net change in pollutant loads.

Differences in forest and agriculture conversion are highlighted by the AC and FC scenarios and manifest in large differences in nitrogen loads. The FC scenario produced less than half of the nitrogen pollution compared to the AC scenario and more than 30% less phosphorus and sediment mainly because it deflected growth onto agricultural lands. Differences in pollutant loads among scenarios are similarly noticeable at the county level despite variability in land use loading rates at that scale (Figures 6–8). The only stark exceptions to the state-wide scenario trends were for nitrogen in the agriculturally dominated counties east of the Bay (e.g., Wicomico, Worcester, Queen Anne's, Talbot, Caroline, Dorchester, Somerset, and Kent) where the AC scenario produced higher TN loads than all other scenarios. While the absolute change in TN loads was very low in these counties across all scenarios, deflecting the small amount of expected growth onto forest and wetland areas leads

TABLE 5 Infill rates by scenario. The CZ, FC, and AC scenarios share the same infill rates. Land Demand Ratios represent proportional differences in population and jobs that must be accommodated by new development, as opposed to infill and redevelopment.

	HT infill rates		CZ, FC, AC infill rates		GM infill rates		Land demand ratios (HT/CZ)		Land demand ratios (CZ/GM)	
County names	Residential	Commercial	Residential	Commercial	Residential	Commercial	Residential	Commercial	Residential	Commercial
Allegany	49%	67%	n/a	78%	n/a	86%	n/a	147%	n/a	154%
Anne Arundel	54%	56%	64%	64%	71%	71%	129%	124%	122%	122%
Baltimore	58%	73%	78%	78%	85%	85%	188%	120%	154%	153%
Baltimore City	92%	91%	n/a	100%	n/a	100%	n/a	n/a	n/a	n/a
Calvert	38%	36%	84%	84%	93%	93%	397%	410%	220%	220%
Caroline	17%	45%	71%	71%	79%	79%	289%	191%	133%	133%
Carroll	37%	40%	85%	85%	94%	94%	428%	409%	239%	239%
Cecil	27%	42%	58%	58%	64%	64%	173%	136%	116%	116%
Charles	27%	44%	54%	54%	60%	60%	160%	122%	113%	113%
Dorchester	20%	42%	35%	42%	39%	47%	122%	100%	106%	108%
Frederick	34%	46%	82%	82%	90%	90%	360%	293%	179%	179%
Garrett	7%	27%	7%	27%	8%	30%	100%	100%	101%	104%
Harford	38%	55%	68%	68%	74%	74%	192%	138%	126%	126%
Howard	36%	46%	84%	84%	92%	92%	391%	332%	206%	206%
Kent	25%	36%	76%	75%	83%	83%	305%	259%	145%	144%
Montgomery	61%	72%	93%	93%	100%	100%	547%	392%	n/a	n/a
Prince George's	56%	72%	82%	82%	91%	91%	247%	160%	187%	187%
Queen Anne's	25%	34%	84%	84%	92%	92%	457%	406%	205%	204%
Somerset	44%	67%	46%	67%	51%	73%	105%	100%	109%	125%
St. Mary's	23%	24%	65%	65%	72%	72%	224%	219%	123%	123%
Talbot	14%	56%	64%	64%	70%	70%	235%	119%	122%	121%
Washington	35%	64%	86%	86%	94%	94%	457%	252%	251%	251%
Wicomico	26%	54%	78%	78%	85%	85%	330%	207%	154%	154%
Worcester	54%	49%	96%	96%	100%	100%	1299%	1442%	n/a	n/a

TABLE 6 Pollutant loads by scenario, 2013–2025 (kilograms).

Scenario	Nitrogen	Phosphorus	Sediment
HT	445,190	35,132	104,674,328
CZ	178,664	15,061	46,223,641
GM	6508	8427	26,017,014
FC	127,119	12,361	37,030,104
AC	256,755	17,783	53,741,402

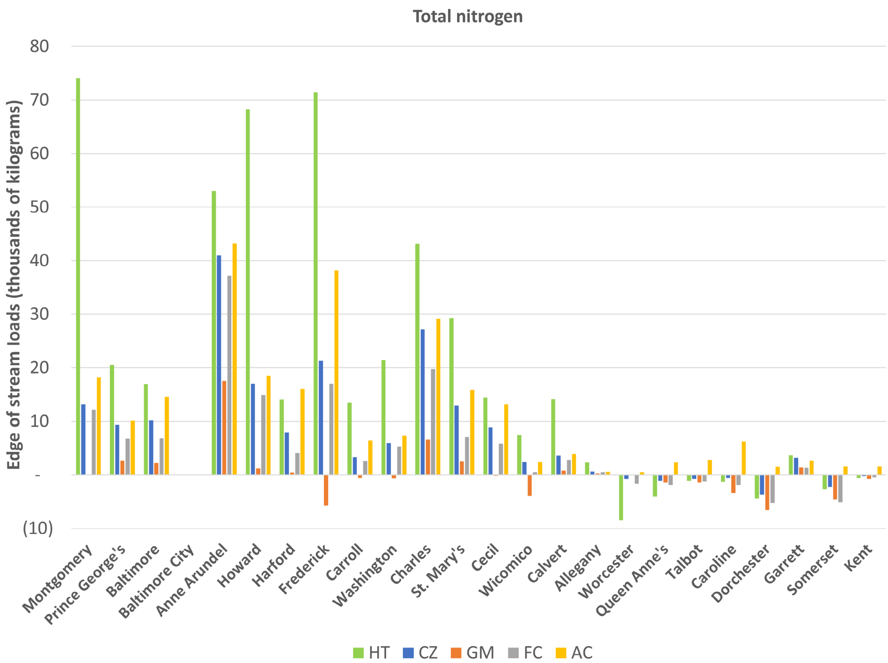

FIGURE 6 Change in total nitrogen loads by county per scenario, expressed in thousands of kilograms across the simulation period, 2013–2025. Counties are ordered from most populated (left) to least populated (right).

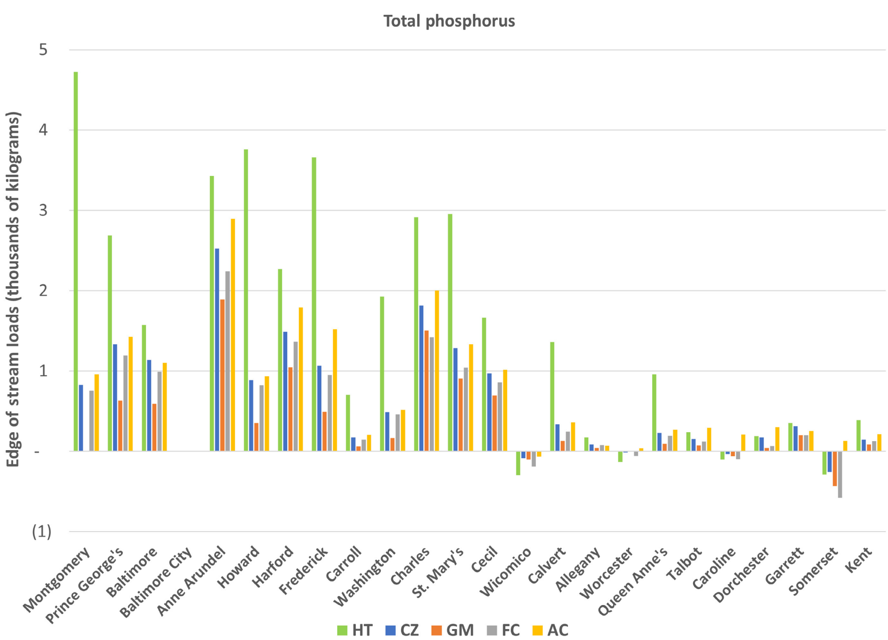

FIGURE 7 Change in total phosphorus loads by county per scenario, expressed in thousands of kilograms across the simulation period, 2013–2025. Counties are ordered from most populated (left) to least populated (right).

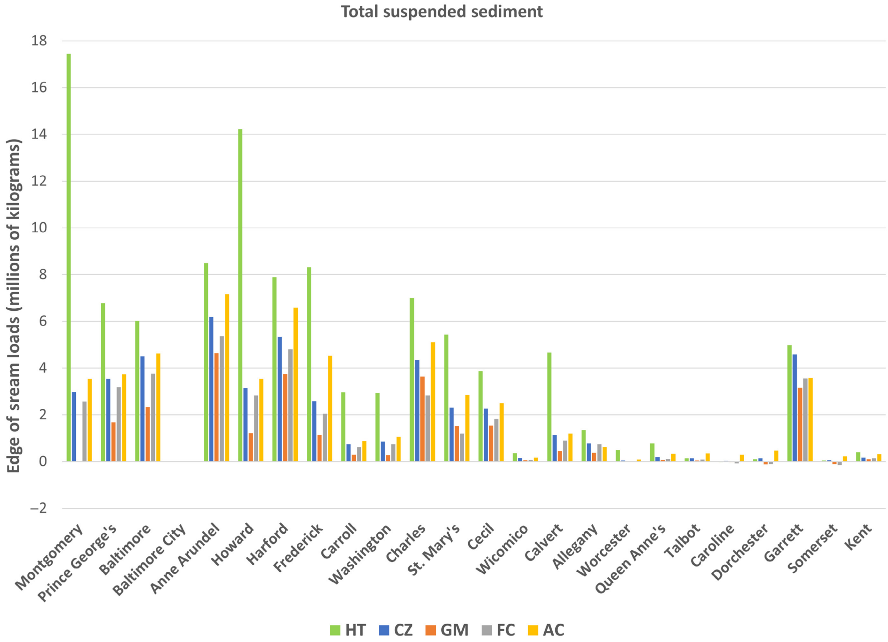

FIGURE 8 Change in total sediment loads by county per scenario, expressed in millions of kilograms across the simulation period, 2013–2025. Counties are ordered from most populated (left) to least populated (right).

to a greater load increase than concentrating growth or diminishing demand for new development. This phenomenon reflects the relatively high agricultural TN loading rates in the eastern counties where corn and soy are dominant crops and poultry manure is plentiful.

Higher demand for new development translates into higher magnitudes of new development and greater increases in pollutant loads from the developed sector. Per-capita pollutant loads vary based on the density of development and type of wastewater service. High-density development served by sewer service yields the lowest per-capita pollutant loads whereas low-density development served by onsite septic systems yields the highest per-capita pollutant loads. The differences in pollutant loads associated with development densities, however, are minor compared to the differences in pollutant loads associated with predevelopment land use conditions which vary spatially within each county. Agricultural lands in Maryland can contribute up to 20 times more TN per unit area compared to natural lands and therefore when agriculture is the dominant land use, particularly along the urban fringe where new development is more probable, net reductions in TN loads from large lot (>=0.4 ha) residential development are likely (Table 7). At higher densities (<0.4 ha), additional pollution from more septic systems and impervious cover exceeds the pre-development pollution from agriculture in 16 of the 23 counties in Maryland. Seven of the eight eastern counties are the exceptions, with agriculture still polluting more on a per-unit area basis compared to 0.2-hectare lot residential development on septic due to the high loading rates for agriculture in these counties relative to development loading rates. However, this analysis does not account for the potential increased application of manure on remaining agricultural land, the increased rate of nutrient leaching associated with heavier applications, or any other types of agricultural intensification. Moreover, this analysis does not account for variability in the nutrient export rates as a function of land cover composition (Wickham et al., 2008).

6.3 | Validation of results

The average annual amount of development simulated across the state of Maryland was significantly lower than observed from the NLCD in four of the five scenarios simulated including the CZ scenario (Table 8). This was unexpected given that the NLCD underestimates the extent of low-density residential growth. The HT scenario simulated 81% more average annual growth than the NLCD and over twice as much growth estimated from parcel data. However, the HT scenario resulted in almost equal proportions of natural and agricultural land conversion as did the parcel analysis (see Table 4).

Simulated rates of residential and commercial infill under the CZ scenario are both weakly correlated with observed rates of infill based on parcel data (Figure 9). The higher residuals at lower levels of infill are mostly associated with less populous rural counties. More populous urban and suburban counties have denser development and correspondingly smaller Census blocks compared to rural counties, enabling more accurate estimation of infill rates because within smaller, denser populated blocks, residential and commercial infill is less confused with other forms of observed development such as new or expanded roads, quarries, and dispersed growth. Surprisingly, the CBLCM-estimated infill rates used for the HT scenario were strongly correlated with observed rates from parcel data indicating that the rates used in the other

TABLE 7 Ten-year net change in total nitrogen pollutant loads (kilograms) associated with the hypothetical development of 20-hectares of natural or agricultural land at three different development densities (2-ha., 0.4-ha., and 0.2-ha. lot sizes) in Maryland counties.

	Net change (10, 2-hectare lots)		Net change (50, 0.4-hectare lots)		Net change (100, 0.2-hectare lots)	
County	Natural land	Agricultural land	Natural land	Agricultural land	Natural land	Agricultural land
Allegany	1549	(2022)	6830	1975	13,163	7900
Anne Arundel	1873	(4199)	8335	1107	16,108	8634
Baltimore	1968	(12,236)	8871	(6526)	17,212	1666
Calvert	1272	(7285)	5515	(4063)	10,571	794
Caroline	2157	(15,407)	9593	(9401)	18,535	(822)
Carroll	2241	(10,991)	10,010	(4680)	19,370	4372
Cecil	1921	(7309)	8673	(1653)	16,839	6306
Charles	1554	(6898)	7022	(2362)	13,637	4023
Dorchester	1051	(12,665)	4525	(10,004)	8653	(5869)
Frederick	2272	(9508)	10,104	(3153)	19,524	5926
Garrett	1264	(5434)	5127	(2977)	9604	1280
Harford	2004	(10,106)	9102	(4111)	17,705	4251
Howard	2117	(7020)	9645	(633)	18,777	8312
Kent	1409	(8259)	6347	(4150)	12,315	1703
Montgomery	1971	(8352)	9043	(2193)	17,644	6362
Prince Georges	1528	(5841)	6889	(1313)	13,369	5064
Queen Annes	1524	(11,730)	6799	(7460)	13,151	(1287)
St. Marys	1531	(7730)	6858	(3410)	13,281	2777
Somerset	1031	(18,364)	4363	(15,928)	8295	(12,046)
Talbot	1325	(11,050)	5748	(7683)	11,018	(2658)
Washington	1940	(5043)	8903	980	17,371	9364
Wicomico	1943	(17,814)	8506	(12,686)	16,353	(5158)
Worcester	1549	(19,521)	6582	(15,912)	12,532	(10,312)

TABLE 8 Modeled average annual development (hectares) from 2013 to 2025 compared to annual average observed development, 2013–2019, from the NLCD and the Maryland Department of Planning.

Scenario	Modeled	NLCD	Parcels
HT	2360	1303	1075
CZ	943	1303	1075
GM	555	1303	1075
FC	953	1303	1075
AC	862	1303	1075

scenarios may be more aspirational than realistic. Observed infill rates were only available for the 2001–2011 period and could change substantially by the year 2025.

The stochastic allocations of residential and commercial growth for all iterations were guided by two development potential raster datasets produced using logistic regression with coefficients of determination (R^2) of 0.81 for residential growth and 0.92 for commercial growth. Observed commercial development occurred closer to major roads and urban areas compared to simulated commercial development (Figure 10). Observed rural residential development occurred closer to urban areas but slightly farther from major roads compared to simulated rural residential development. Because development densities tend to be higher near urban areas, these differences in simulated versus observed patterns may contribute to an underestimation of development densities. Simulated residential development patterns relative to distance to major roads are closely aligned to observed patterns. For commercial development, the simulated patterns align with observed ones at distances greater than 17–18 km from major roads. At shorter distances, commercial development occurred closer to major roads compared to simulated growth. Approximately 75% of observed commercial development occurred within 5 km of urban areas compared to only 50%

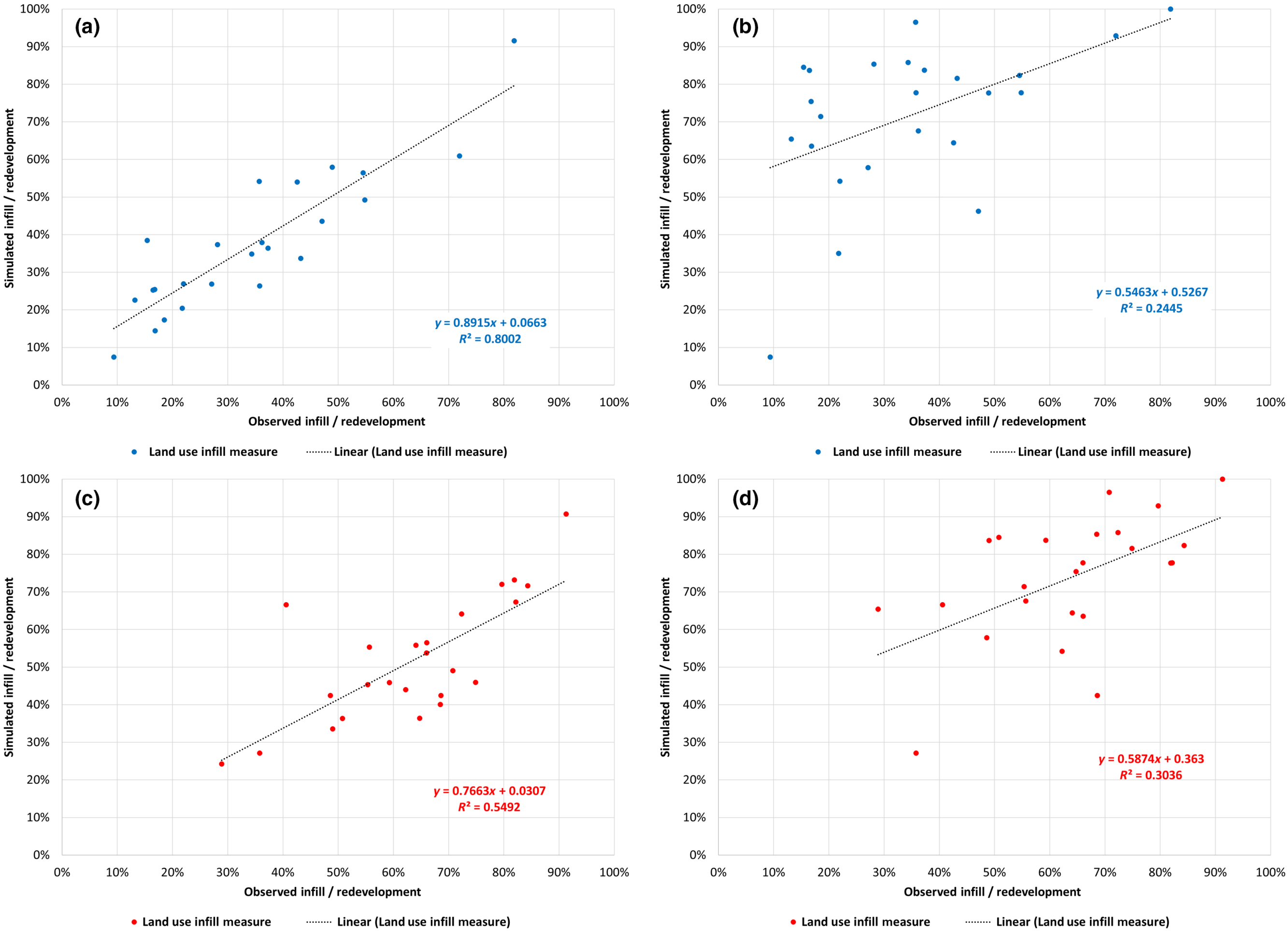

FIGURE 9 Validation of county-level estimated commercial (red) and residential (blue) infill/redevelopment rates based on a comparison of observed parcel development and NLCD development within qualifying infill blocks. Comparisons are presented for the HT scenario (a, c) and CZ scenario (b, d).

of simulated commercial growth. The patterns of residential development relative to urban areas yielded similar results with 70% of observed growth and 50% of simulated growth occurring within 5 km.

There is a weak relationship between modeled residential and commercial densities even at the larger Census Tract level (aggregations of Census Blocks) with regression coefficients of 0.23 and 0.40, respectively. Because the numerators in the density calculations, that is, the number of houses or number of jobs, are the same in the modeled and parcel data, differences in densities are explained by differences in the denominators: residential or commercial areas. In highly urbanized Census Blocks with mixed use development and large numbers of housing units and jobs, modeled densities should be lower than those derived from parcel data because the portion of census blocks that may be purely commercial is combined with the portion that may be purely residential in the numerators for both residential and commercial modeled densities whereas residential and commercial parcels within each census block are clearly differentiated. In areas with separately zoned residential and commercial areas, modeled densities should be higher than parcel-based densities because the NLCD-developed classes are based on the intensity of development and besides from golf courses, the NLCD data seldom class large lawns and landscaped areas as developed. These expectations were partially realized. Examining differences between the modeled and parcel-based patterns of residential densities reveals areas in central and southern Maryland where modeled densities were consistently higher than observed (Figure 11). These areas often have large lot sizes leading to low parcel-based densities but high modeled densities because only the impervious portions of each lot identified in the NLCD are considered as part of the modeled residential or commercial area.

Commercial and residential land consumption rates (growth per housing unit, growth per job, and growth per year) are highly variable by county and by time interval (Figures 12 and 13). Expectedly, the more populous counties tend to have the highest land consumption rates per year and lowest rates per housing unit or job. Errors in the estimation of infill are evident in the annual land consumption rates which are indicative of the overall magnitude of growth: Higher infill rates result in less development. Errors in the estimation of densities are evident in the

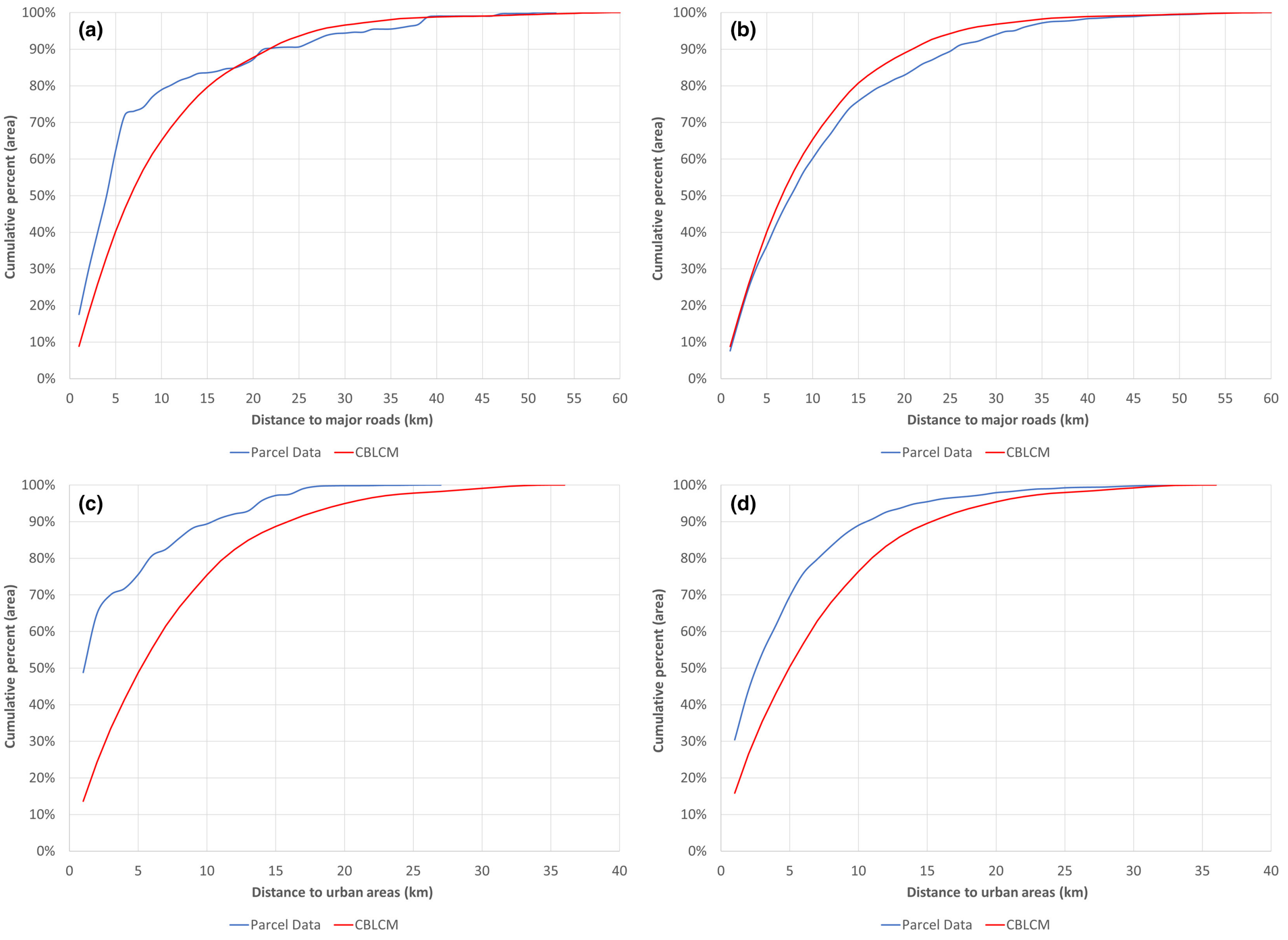

FIGURE 10 Validation of observed (parcel data) versus modeled (CBLCM) spatial patterns of commercial development relative to (a) major roads and (c) urban areas and of residential development relative to (b) major roads and (d) urban areas.

per-job and per-housing unit land consumption rates: Higher densities result in less new development per house and per job. For commercial development, simulated annual land consumption rates were lower than observed rates in most counties and higher than observed rates in only two of the more populous counties (Anne Arundel and Harford) and in five of the six least populous counties. Simulated land consumption rates per job were higher than observed rates for all counties. These comparisons reveal that the simulated infill rates should be lower and the simulated densities higher to better align with observations of change based on parcel data. Because not a lot of commercial development typically occurs within a county over just a 10-year period, especially in rural counties, a few very large or very small parcels can easily skew the overall land consumption rates making them particularly challenging to validate. Fluctuations in occupancy rates can also create conditions where one might observe large fluctuations in employment coupled with minimal new construction. Fluctuations in commercial occupancy and sector-specific floor area ratio requirements are not explicitly accounted for in the CBLCM.

Similar to commercial development, the high infill and redevelopment rates used in the CZ scenario result in a general underestimation of annual residential land consumption rates for all but one county compared to the 2001–2011 observed period and in 18 of the 24 counties over the more recent 2011–2019 observed period. The differences in annual land conservation rates between the two observed periods indicate that the past is not always prologue to the future. The accuracy of most land change models, including this one, at least partially depends on this assumption being true. The simulated land consumption rates per housing unit were like those observed over both time periods for most counties revealing that fine-scale differences between modeled and observed residential densities at the Census Block level might cancel out at the county level.

Development trends are highly episodic and influenced by fluctuations in market forces and public and private investments, and therefore, stationarity may be another false assumption. Lacking prior knowledge about episodic and uncertain events, it is unlikely that any regional urban growth model calibrated with data from the 1990s and used to forecast urbanization through the 2000s could accurately simulate the rapid rate of growth during the housing boom (2000–2005) or the collapse of the housing market beginning in 2006 (Figure 14). Land change

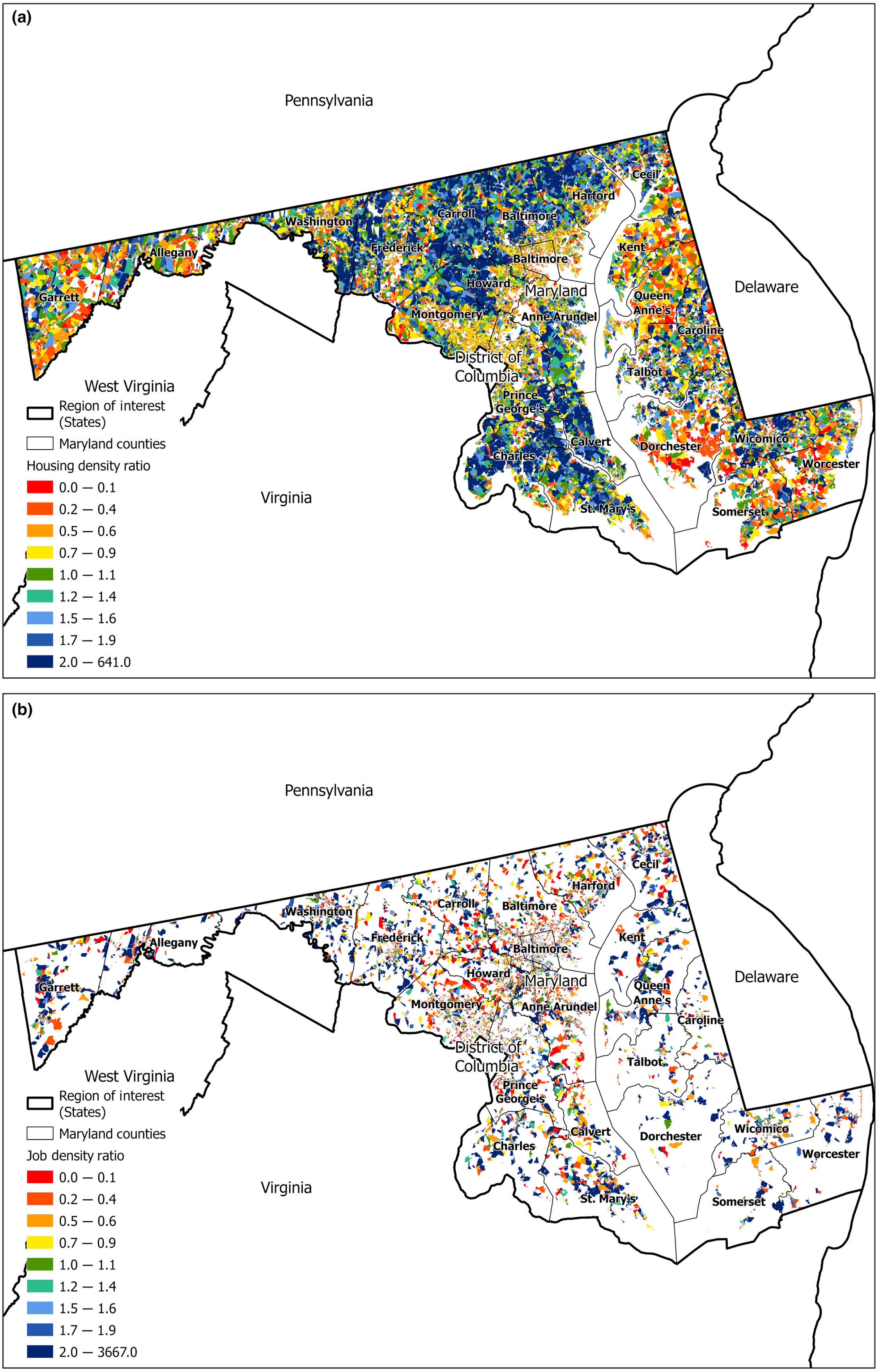

FIGURE 11 Ratios of modeled to parcel-based development densities by Census Block for (a) residential development (housing units per 30 m cell) and for (b) commercial development (jobs per 30 m cell).

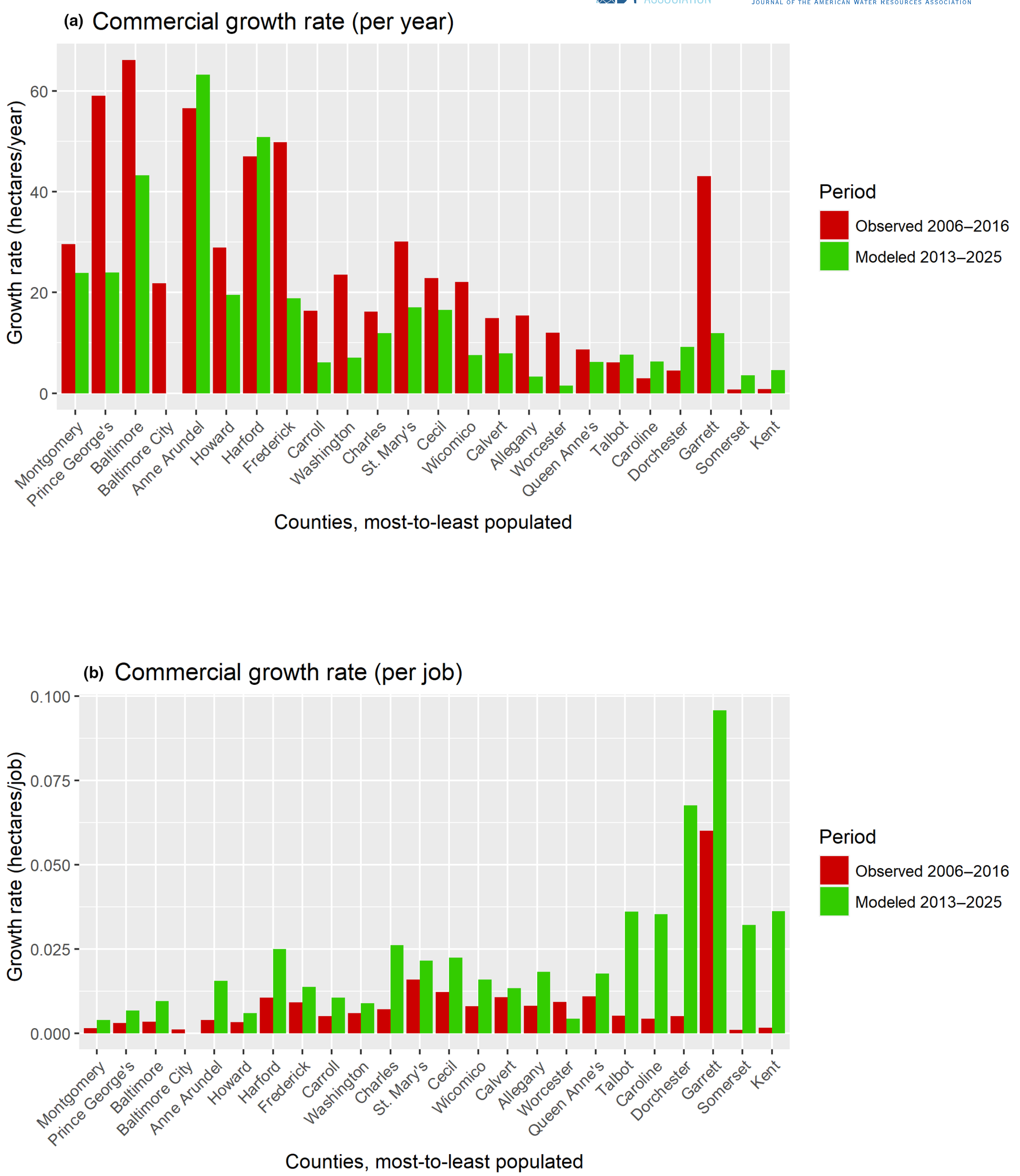

FIGURE 12 Observed and modeled commercial land consumption rates by county in units of hectares per year (a) and hectares per job (b) are shown for the CZ scenario with observed rates of commercial land use change from 2006 to2016 (red) and modeled rates of commercial land use change from 2013 to 2025 (green). Counties are ordered from most populated (left) to least populated (right).

modelers attempt to consider such possibilities by simulating alternative future scenarios. Scenarios can even be crafted to simulate rare episodic events such the effects of a global pandemic. Whether such scenarios can inform policies and decisions however rests as much with their plausibility, transparency, and stakeholder involvement in their design as it does in statistical accuracy assessments.

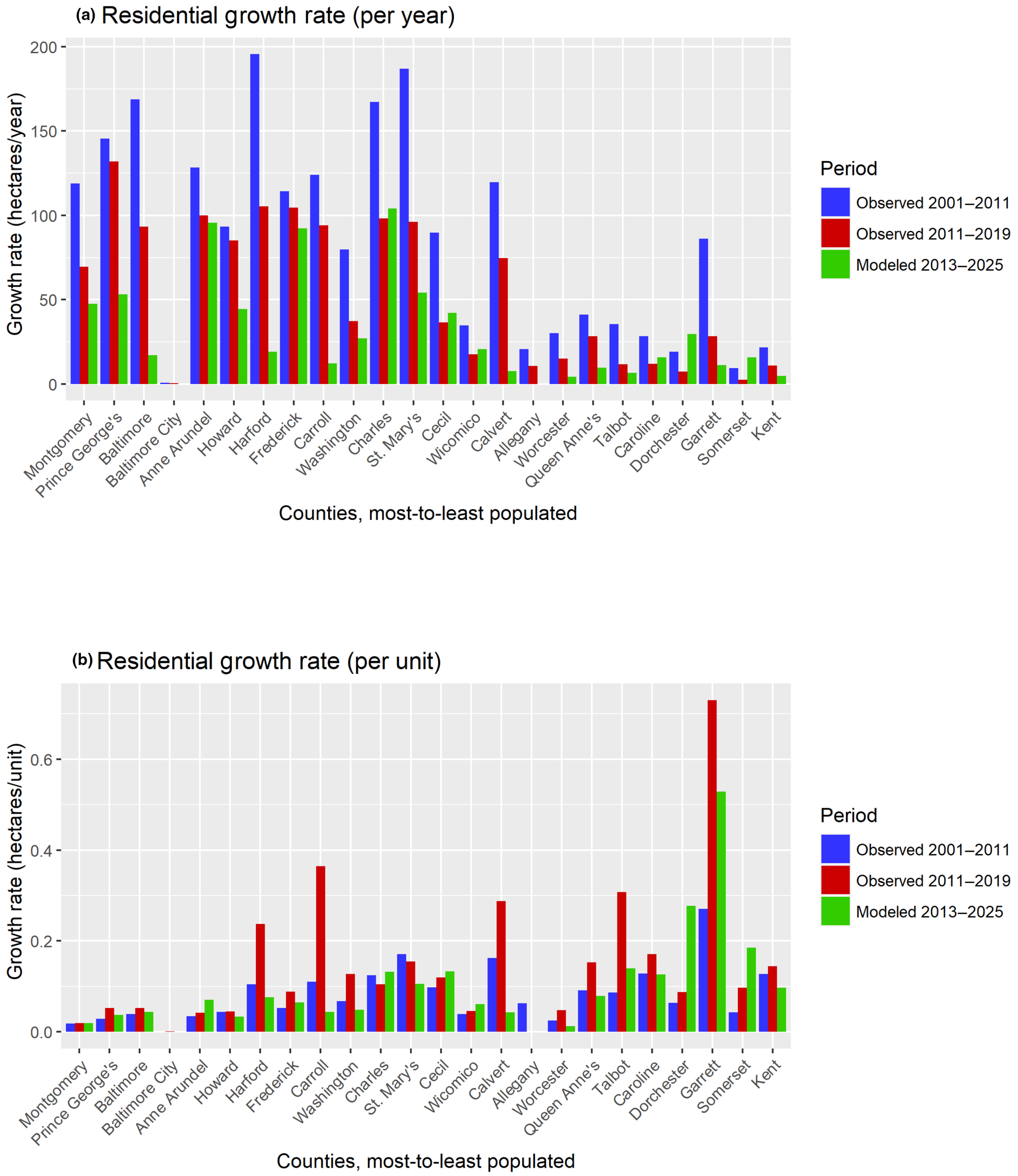

FIGURE 13 Observed and modeled residential land consumption rates by county in units of hectares per year (a) and hectares per housing unit (b) are shown for the CZ scenario with observed rates of residential land use change from 2001 to 2011 (blue), observed rates of residential land use change from 2011 to 2019 (red), and modeled rates of residential land use change from 2013 to 2025 (green). Counties are ordered from most populated (left) to least populated (right).

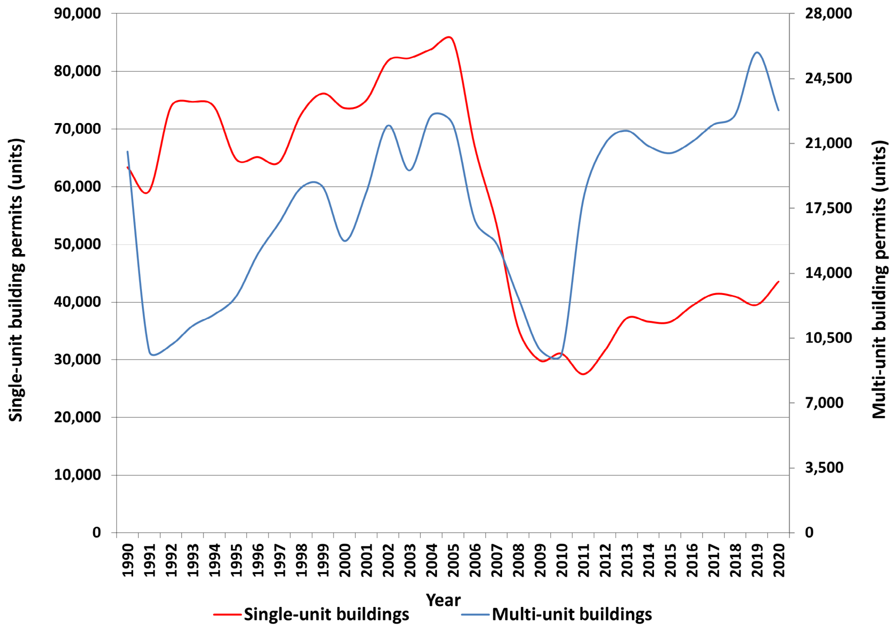

FIGURE 14 Trends in single-unit and multi-unit residential building permits in Maryland from 1990 to 2020 (US Census Bureau, 2021).

7 | CONCLUSIONS

Simulating the impacts of future alternative land use scenarios on water quality is not novel. In the context of informing state restoration plans to comply with a TMDL, however, the CBLCM simulation of future scenarios is unique and establishes a precedent for other long-term, regional-scale restoration efforts (e.g., Delaware River Basin, Great Lakes, Gulf of Mexico, Puget Sound, and the Everglades). The CBLCM has been adopted by the CBP Partners as an operational model that is maintained and updated every 2 years to inform restoration plans. The simulation of alternatives to the baseline "CZ" scenario that reflects public and private policies, programs, and land protection efforts enables the quantification and formal crediting of those efforts for reducing impacts to water quality. Five alternative scenarios are presented in this paper to illustrate the potential magnitude and range of water quality improvements achievable through growth management, agricultural conservation, and forest conservation. Alternative future scenarios help raise awareness of the importance of understanding landscape patterns and composition across the urban to rural gradient for developing policies that minimize impacts to natural lands and water quality. Scenario planning is particularly essential for assessing the complex impacts of climate change on the Chesapeake Bay because changes in climate will likely influence human migration, development patterns, and agricultural land use practices.

Land use planning is arguably the first "Best Management Practice" that should be considered to minimize environmental impacts from future development (Hirshman & Kosco, 2008). By concentrating development and locating it in areas with sufficient infrastructure to support it (i.e., roads, schools, and wastewater treatment plants), land use planners can minimize increases in impervious cover and the conversion of natural and agricultural lands resulting from population growth. Supporting this hypothesis, the GM scenario achieved the greatest reductions in pollution compared to all other scenarios. This scenario, however, is generalized by design and does not guarantee protection of the highest valued forests or farmland nor does it necessarily achieve other desired social and economic outcomes. Land use planning is a process involving tradeoffs between social, economic, and environmental objectives. Future land use scenarios should be designed accordingly to reflect a combination of growth management, land protection, social and economic equity, and other objectives. To make informed decisions, local government officials need to understand the range of plausible future land use conditions and associated impacts, including impacts to water quality. The CBLCM provides a plausible and objective way of quantifying impacts to water quality from different urban growth trajectories and land use policies. It is well suited for this purpose because it simulates both residential and commercial development and quantifies the conversion of natural and agricultural lands along with future populations served by public sewer or private septic systems. These forecasts are driven by exogenous projections of population and employment which are already used to inform state and county planning and investment decisions and therefore contribute to stakeholder acceptance of model results. The CBLCM also accounts for infill and development densities, making it particularly useful for simulating the impacts of common land use policies. However, validating regional-scale forecasts of urban land use, as opposed to land cover, will continue to be challenging because of the limited availability of multi-temporal land use datasets.

While the CBLCM produces forecasts of residential and commercial growth that align with observed trends in a few counties, accurately simulating development at the county scale is challenging, especially in areas with low levels of population and employment growth, large amounts of land suitable for development, and less restrictive zoning regulations. In rural counties, the densities of new residential and commercial growth may vary significantly compared to long-term average densities. A few warehouses or a single golf course may represent most new commercial growth in a rural county over a decade and their very-low development densities may mark a sharp departure from long-term average densities. Predicting such developments requires local information on developer interests, speculation, and incentives and information on the employment sectors (e.g., warehouses, financial, services, etc.) served by new development. Moreover, in all counties transportation and energy infrastructure projects including road construction, lane expansions, and the installation of new utility transmission lines and pipelines are often completely decoupled from county-level changes in population and employment yet visible as new development in the NLCD and possibly even in parcel data. Simulations of residential growth may be improved by considering trends in seasonal housing separately from total housing and by considering household characteristics such as size and income as potential new variables for explaining patterns in residential development. Simulations of commercial growth may be improved with data on office space vacancies, floor area requirements, and statistical estimates of future commercial office space demand by employment sector.

AUTHOR CONTRIBUTIONS

Peter R. Claggett: conceptualization, writing, reviewing, editing. **Labeeb Ahmed:** conceptualization, coding, writing. **Frederick M. Irani:** conceptualization, coding. **Sarah McDonald**: conceptualization, coding. **Renee L. Thompson:** conceptualization, writing.

ACKNOWLEDGMENTS

This project was supported by the USGS Land Change Science program, USGS Priority Ecosystem Studies program, USEPA Chesapeake Bay Program Office, and by a grant from the USEPA National Center for Environmental Research (DW-14-95813701). External project collaborators include Gerrit Knapp and Uri Avin of the National Center for Smart Growth, Research, and Education and the planning departments of Talbot and Howard Counties in Maryland. We would especially like to thank Naomi Detenbeck of USEPA's Office of Research and Development for her guidance, patience, and support.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

Data used in this project can be accessed by contacting the authors.

ORCID

Peter R. Claggett https://orcid.org/0000-0002-5335-2857
Labeeb Ahmed https://orcid.org/0000-0003-4524-9611
Frederick M. Irani https://orcid.org/0000-0002-2424-0135
Sarah McDonald https://orcid.org/0000-0003-3534-325X
Renee L. Thompson https://orcid.org/0000-0003-1463-5173

REFERENCES

Boesch, D.F., and J. Greer. 2003. *Chesapeake Futures: Choices for the 21st Century*. Edgewater, Maryland: Chesapeake Research Consortium, Inc.

Box, G.E. 1976. “Science and Statistics.” *Journal of the American Statistical Association* 71(356): 791–9.

Brown, D.G., P.H. Verburg, R.G. Pontius, Jr., and M.D. Lange. 2013. “Opportunities to Improve Impact, Integration, and Evaluation of Land Change Models.” *Current Opinion in Environmental Sustainability* 5(5): 452–7.

Cappiella, K., and K. Brown. 2001. *Impervious Cover and Land Use in the Chesapeake Bay Watershed*. Ellicott City, MD: Center for Watershed Protection.

Chesapeake Bay Program (CBP). 2020. *Chesapeake Assessment and Scenario Tool (CAST) Version 2019*. Annapolis, MD: Chesapeake Bay Program Office.

Chesapeake Conservation Partnership (CCP). 2020. Chesapeake Conservation Atlas. Annapolis, MD. http://www.chesapeakeconservation.org/our-work/goal-mapping/our-valued-lands/.

Claggett, P.R., F.M. Irani, and R.L. Thompson. 2013. “Estimating the Extent of Impervious Surfaces and Turf Grass across Large Regions.” *JAWRA Journal of the American Water Resources Association* 49(5): 1057–77.

Claggett, Peter R., and C.A. Jantz. 2004. “Assessing Development Pressure in the Chesapeake Bay Watershed: An Evaluation of Two Land-Use Change Models.” *Environmental Monitoring and Assessment* 94(1–3): 129–46.

Clarke, K.C., S. Hoppen, and L.J. Gaydos. 1997. “A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area.” *Environment and Planning B: Planning and Design* 24: 247–61.

Dewitz, J. 2019. National Land Cover Database (NLCD) 2016 Products: U.S. Geological Survey data release, https://doi.org/10.5066/P96HHBIE

Dewitz, J., and U.S. Geological Survey. 2021. *National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021): U.S. Geological Survey Data Release*. Sioux Falls, SD. https://doi.org/10.5066/P9KZCM54.

HERE, 2013. NAVSTREETS Street Data v5.0, Chicago, IL.

Hirshman, D., and J. Kosco. 2008. *Managing Stormwater in Your Community: A Guide for Building an Effective Post-Construction Program*. EPA Publication No: 833-R-08-001. Ellicott City, MD: Center for Watershed Protection.

Homer, C., J. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown. 2015. “Completion of the 2011 National Land Cover Database for the Conterminous United States–Representing a Decade of Land Cover Change Information.” *Photogrammetric Engineering & Remote Sensing* 81(5): 345–54.

Homer, C., Huang, C., Yang, L., Wylie, B. K., and Coan, M. 2004. "Development of a 2001 National Land-Cover Database for the United States". *Photogrammetric Engineering & Remote Sensing* 70(7): 829–40.

Jantz, C., S. Drzyzga, and M. Maret. 2014. “Calibrating and Validating a Simulation Model to Identify Drivers of Urban Land Cover Change in the Baltimore, MD Metropolitan Region.” *Land* 3(3): 1158–79.

Jantz, C.A., S.J. Goetz, D.I. Donato, and P.R. Claggett. 2010. “Designing and Implementing a Regional Urban Modeling System Using the SLEUTH Cellular Urban Model.” *Computers, Environment and Urban Systems* 34(1): 1–16.

Lewis, Rebecca, Gerrit-Jan Knaap, and Jungyul Sohn. 2009. “Managing Growth with Priority Funding Areas: A Good Idea whose Time Has Yet to Come.” *Journal of the American Planning Association* 75(4): 457–78.

McCauley, S., and S.J. Goetz. 2004. “Mapping Residential Density Patterns Using Multi-Temporal Landsat Data and a Decision-Tree Classifier.” *International Journal of Remote Sensing* 25(6): 1077–94.

Meentemeyer, R.K., W. Tang, M.A. Dorning, J.B. Vogler, N.J. Cunniffe, and D.A. Shoemaker. 2013. “FUTURES: Multilevel Simulations of Emerging Urban–Rural Landscape Structure Using a Stochastic Patch-Growing Algorithm.” *Annals of the Association of American Geographers* 103(4): 785–807.

Meyer, W.B., and B.L. Turner. 1992. “Human Population Growth and Global Land-Use/Cover Change.” *Annual Review of Ecology and Systematics* 23(1): 39–61.

Moore, R.B., McKay, L.D., Rea, A.H., Bondelid, T.R., Price, C.V., Dewald, T.G., and Johnston, C.M. 2019. User's Guide for the National Hydrography Dataset Plus (NHDPlus) High Resolution: U.S. Geological Survey Open-File Report 2019–1096. Reston, VA. https://doi.org/10.3133/ofr20191096.

National Research Council (NRC). 2014. *Advancing Land Change Modeling: Opportunities and Research Requirements*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18385.

Pontius, R.G., W. Boersma, J.-C. Castella, K. Clarke, T. de Nijs, C. Dietzel, Z. Duan, et al. 2008. “Comparing the Input, Output, and Validation Maps for Several Models of Land Change.” *The Annals of Regional Science* 42(1): 11–37.

Pontius, R.G., J.-C. Castella, T. De Nijs, Z. Duan, E. Fotsing, N. Goldstein, K. Kok, et al. 2018. “Lessons and Challenges in Land Change Modeling Derived from Synthesis of Cross-Case Comparisons.” In *Trends in Spatial Analysis and Modelling* 143–64. Cham, Switzerland: Springer.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for statistical Computing. ISBN 3-900051-07-0. http://www.R-project.org/.

Shenk, G.W., and L.C. Linker. 2013. “Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model.” *JAWRA Journal of the American Water Resources Association* 49(5): 1042–56.

Soares-Filho, B.S., G.C. Cerqueira, and C.L. Pennachin. 2002. “DINAMICA—A Stochastic Cellular Automata Model Designed to Simulate the Landscape Dynamics in an Amazonian Colonization Frontier.” *Ecological Modelling* 154(3): 217–35.

Sohl, T.L., and P.R. Claggett. 2013. “Clarity Versus Complexity: Land-Use Modeling as a Practical Tool for Decision-Makers.” *Journal of environmental management* 129: 235–43.

U.S. Geological Survey (USGS) Gap Analysis Project (GAP). 2018. Protected Areas Database of the United States (PAD-US): U.S. Geological Survey Data Release. Denver, CO. https://doi.org/10.5066/P955KPLE.

US Census Bureau. 2010a. “TIGER/Line Shapefiles (Machine Readable Data Files)/Prepared by the U.S. Census Bureau, 2010.” https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

US Census Bureau. 2010b. “Census Summary File 1; Tables H3 – Occupancy Status; Generated Using American FactFinder.” http://factfinder.census.gov.

US Census Bureau. 2016. *LEHD Origin–Destination Employment Statistics (LODES) Dataset Structure Format Version 7.2, Revised 03-03-2016*. Washington, DC: US Census Bureau.

US Census Bureau. 2021. *Building Permits Survey*. Washington, DC: US Census Bureau. https://www.census.gov/construction/bps/.

US Environmental Protection Agency (USEPA). 2000. *Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns, Office of Research and Development*. Washington, DC: U.S. Environmental Protection Agency. EPA/600/R-00/098.

US Geological Survey. 2019. “The National Map.” https://www.usgs.gov/the-national-map-data-delivery.

USEPA. 2007. *Development Growth Outpacing Progress in Watershed Efforts to Restore the Chesapeake Bay*. Washington, DC: Office of Inspector General, U.S. Environmental Protection Agency Report No. 2007-P-00031.

USEPA. 2009. Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines. Global Change Research Program, National Center for Environmental Assessment, Washington, DC; EPA/600/R-08/076F

USEPA. 2010. “Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus, and Sediment.” U.S. Environmental Protection Agency, Region 3, Philadelphia, Pennsylvania.

USEPA. 2017. “Updates to the Demographic and Spatial Allocation Models to Produce Integrated Climate and Land Use Scenarios (ICLUS) (Version 2).”

Van der Leeden, F. 1975. *Water Resources of the World*. Port Washington, New York: Water Information Center Publication.

Verburg, P.H., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon, and S.S. Mastura. 2002. “Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model.” *Environmental Management* 30(3): 391–405.

Waddell, P. 2011. “Integrated Land Use and Transportation Planning and Modelling: Addressing Challenges in Research and Practice.” *Transport Reviews* 31(2): 209–29.

Wickham, J.D., T.G. Wade, and K.H. Riitters. 2008. “Detecting Temporal Change in Watershed Nutrient Yields.” *Environmental Management* 42(2): 223–31.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Claggett, Peter R., Labeeb Ahmed, Frederick M. Irani, Sarah McDonald and Renee L. Thompson 2023. "The Chesapeake Bay Land Change Model: Simulating Future Land Use Scenarios and Potential Impacts on Water Quality." *JAWRA Journal of the American Water Resources Association* 00 (0): 1–26. https://doi.org/10.1111/1752-1688.13131.