MEMORANDUM

TO: Christopher Pyke, US Green Building Council

FROM: Peter Claggett, US Geological Survey

DATE: February 09, 2011

RE: Response to Draft STAC Review of Land-Use and Land Cover Dataset and Methodology

Dear Review Committee Members and the Scientific and Technical Advisory Committee:

Thank you for taking the time to compile a committee, review and provide timely and insightful comments and suggestions related to the Chesapeake Bay Program's Land-Use and Land Cover Dataset and Methodology. The CBPO Land Data Team (LDT) have reviewed your comments and concerns. Over the past several months, the LDT has invested substantial effort in documenting and substantiating the many assumptions used in developing the Phase 5.3.2 developed land use dataset. To systematically address STAC comments and suggestions, the Land Data Team has organized this response under key subject headings **in bold 14-pt type** with summarized STAC comments in blue type followed by the LDT's response in *red italics*. Comments regarding the SLEUTH model are not addressed because that part of the Chesapeake Bay Land Change Model is no longer being used to simulate future scenarios associated with the Bay TMDL.

Hydrologic, Nutrient, and Sediment Load Impacts of Developed Land Uses

The reviewers posed the following questions: Is every lawn equally likely to contribute to the model? Is it true that all pervious areas in the high density residential class is lawn? We know that it's difficult to come up with reasonable numbers, but this gets to the root of an issue: for many of these analyses, "all points are created equal". They stated that they understand that there is a parameter in the model that deals with "attenuation" or something that shows how far from the CB each point is, but we would argue that there should at least be some discussion in the document of how likely each point on the surface is likely to contribute material that ends up in the bay. The reviewers suggested a simple weighting procedure that incorporates slope, stream density or flow accumulation from a DEM.

> *Spatially weighting the contribution of impervious surfaces and lawns to the Bay based on their proximity to streams, slope, probability of compaction (pervious areas only) and other factors is very reasonable. At this time, however, we have not found enough information in published literature to justify even a simple weighting procedure. The USGS is initiating a study to develop and test weights for impervious surfaces based on their connectivity to streams and impacts on flow.*

An analysis should be conducted on the Phase 5.3.2 watershed model to understand the sensitivity of nutrient and sediment load estimates to changes in developed land area and populations on sewer and septic.

> *The LDT is working closely with EPA's Watershed Modeling Team to better understand the sensitivity of the model to changes in developed land. The p532 land use has more developed land than the p530 land use yet the extent of agricultural land is the same because it is derived from the USDA Census of Agriculture and not from satellite imagery. The woody/open land use class (formerly called "forest") is whatever remains after accounting for water, development, extractive, bare-construction, and agriculture. Therefore, there will be less woody/open acreage in the p532 land use dataset compared with p530 which will compensate for some of the expected increase in loads from developed lands. The initial P532 calibration will have the same loading rates per acre per land use as the initial version p530 loading rates. The total initial load from developed land uses will increase, however, meaning that on average, the loading rates for all land uses after calibration will most likely decrease compared to the calibrated version of p530 as they are adjusted to meet in-stream monitoring data. This effect will be more pronounced in areas where the change in land use is larger. In terms of total load, the urban sector will almost certainly increase, the forest sector will almost certainly decrease, and it is likely that the agriculture and waste water sector loads will decrease slightly, all else being equal.*
>
> *The USGS will also execute nutrient and sediment SPARROW models for the Bay watershed using the new improved p532 developed land uses and impervious surface estimates as predictors of water quality. Results from the models will then be used to evaluate the effectiveness of the updated urban data in explaining observed Nitrogen, Phosphorus, and Sediment loads in streams. These analyses will test if the p532 developed land use improves the overall statistical fit of the SPARROW models compared to fitting the models with satellite derived developed land cover data.*

Future Growth and Scenarios

STAC expressed concern that the single trend, "business as usual" scenario, developed by the Land Data Team is just one of many equally plausible future scenarios and therefore multiple scenarios should be produced representing a range of possible and plausible future outcomes. Demographic, economic, and other changes may greatly alter future growth patterns compared to the recent past.

> *In the spring/summer of 2011, the LDT will host an Alternative Future Scenario workshop to solicit input from various state and local governments for constructing a range of plausible future scenarios. The LDT will amend the current land use methodology as needed to accommodate the factors and phenomena associated with any new scenarios (e.g., increased infill development, increased density, de-coupled residential and commercial development trends).*

Gompertz growth curves are inadequate guides to future development, particularly total housing demand in 2010, 2017, and 2025. Fundamental changes in real estate market dynamics make it grossly inadequate to use historic parameterization for these estimates. As recommended in an earlier STAC review, bounding the range of conditions requires scenarios based on a plausible range of regional macro- and micro-economic factors along with policy choices.

Compare Figure 5 *Illustration of Gompertz Curve Fit* with actual market dynamics, such as http://www.businessinsider.com/the-housing-chart-thats-worth-1000-words-2009-2. The Gompertz Curve does not begin to capture these dynamics. Moreover, there is evidence that many parts of the US have an oversupply of the type of single family detached housing imagined in this analysis – i.e., a sufficient supply to meet demand for decades into the future. With current rates of household formation, stagnant real income growth, and an aging population, we are likely to see a fundamental and long-term shift in real estate market dynamics.

> *The Gompertz family of growth curves has been used effectively to simulate housing growth at the sub-county scale (Reilly, 1997). Gompertz curves have an "S" shape which makes them especially suitable for modeling situations in which: a) growth is at first slow (such as when an agricultural or forested area is beginning to be developed); b) growth takes place rapidly (such as when fairly large tracts are being developed into suburban housing subdivisions); and c) growth trails off but does not come to a frank stop (such as when marginal areas within suburban areas or cities are developed or urban re-development entails marginal increases in housing density). These phenomena are evident in the Demographic Transition Model that simulates relative changes in the birth and death rates as a nation transitions from an agricultural to industrial to post-industrial economy.*
>
> *The Gompertz family of growth curves has been used effectively to simulate housing growth at the sub-county scale in New Jersey (Reilly, 1997) which is why they were adopted for use in the Chesapeake Bay watershed. At the county and municipal scales, one can observe a similar form of transition in the amount of development occurring in the region. Economic changes (e.g., the emergence of a technology corridor in northern Virginia combined with low mortgage interest rates) prompt less developed jurisdictions (e.g., Loudoun) to experience unprecedented high rates of development to accommodate new residences and businesses. Eventually, after one to three decades, the high rates of growth taper off due to further economic changes and/or limitations on the amount of land available for development.*
>
> *For short-term, ten to twenty-year, forecasts the Gompertz equation produces results that rarely diverge from a linear forecast. At the modeling segment scale, comparing the Gompertz forecasts for 2025 with a linear forecast for 2025 produced an R2 of 0.98 (p-value <0.05). In only a few cases did the Gompertz equation forecast growth divergent from linear expectations and in all cases the growth was less than expected. The most notable exceptions occurred in Loudoun, Prince William, and Fairfax Counties in Virginia. All three of these counties experienced high growth rates between 1992 and 2006 that are not expected to continue based on increased land values, zoning, and demographic projections. For Loudoun and Fairfax counties, even the dampening effect of land constraints incorporated into the Gompertz equation were insufficient*

to produce housing forecasts as low as those derived from the current local population projections.

In summary, using a linear growth equation in place of a Gompertz equation (which would improve the transparency of the model) would be inadequate for simulating even short-term growth in jurisdictions that have experienced high growth rates over the past 2-3 decades. As the reviewers have noted, however, the Gompertz curve may not be the appropriate formula for accommodating changes in development policies (e.g., infill and redevelopment) or exogenous economic and demographic changes. The LDT will investigate alternative formulae and approaches for future versions of the Chesapeake Bay Land Change Model.

One STAC reviewer tried to estimate the direction of impacts from different assumptions. Reflecting on the various modeling assumptions the Reviewer concluded that assumptions underlying the current, single-scenario, deterministic approach are likely to significantly overestimate the extent of residential land use, underestimate hydrologic impacts, and significantly underestimate the overall extent of non-residential (commercial land use).

We do not agree that the current approach overestimates the extent of residential land use. The land use methodology was segmented into three zones: urban, suburban, and rural. Satellite data were used to map urban areas, secondary road density and satellite data were used to map suburban areas, and development in rural lands was determined using rural single-detached housing units, an average lot size coefficient, state specific residential impervious surface coefficients, combined with data on all road lengths coupled with randomly sampled road width coefficients. By comparing the Phase 5.3.2 impervious surface estimates at the county scale with local data in Lancaster County, PA, all three Delaware counties, and Montgomery County, MD, it is evident that the Phase 5.3.2 land use dataset underestimates impervious surface compared with local data by about 5-20%.

Commercial and industrial land uses and agricultural structures are underestimated, particularly in rural counties because only roads and single-detached housing units were used to map development in rural areas.

We do not believe that the hydrologic impacts from impervious or pervious lands are significantly underestimated. Impervious surfaces in the Phase 5.3.2 dataset increased 93% compared with Phase 5.3.0 where they were delineated solely based on Landsat satellite derived land cover data. Much of this increase occurred in rural areas through the use of roads and housing unit information. Arguably, dispersed impervious surfaces in rural areas may have less impact on Bay water quality than concentrated impervious surfaces adjacent to streams or directly connected to streams via underground conveyance systems. Because HSPF is a lumped parameter model, all impervious surfaces in a modeling segment contribute equally to nutrient and sediment loads despite the fact that some impervious surfaces are directly connected to streams via storm drains and others are dispersed fragments of impervious surface distant from waterways.

The question of whether it is appropriate to assume that the rate of conversion from forest and farmland to development and infill that occurred between 1984 and 2006 would in fact stay constant through the year 2025 was also raised. The reviewers felt this was an unrealistic assumption given the current economy and decline in new construction and recommended the rate for conversion in the next 15 years should be reduced to account for that decline.

> *We multiplied the historic ratios (1984 – 2006) of farmland and forest conversion to developed land cover in each modeling segment by the forecasted extent of new development to determine how much farmland and forest land conversion would occur through the year 2025. We did not use historic rates of change in the CBLCM derived from the 1984-2006 land cover data series nor did we assume a constant rate of change in simulating the future trend scenario. Infill was assumed to occur in all modeling segments to a degree relative to the remaining amount of land available for development (i.e., gently sloped, undeveloped and unprotected uplands). The amount of infill was determined by a densification factor. The densification factor is based on an exponential relationship between single-detached residential parcel size and the percentage of undeveloped land within all modeling segments in Maryland. As the percentage of undeveloped land decreases below 90%, the average residential parcel size decreases exponentially. In the CBLCM, this means that forecasted development in 2010, 2017, and 2025 is multiplied by a number less than one if the amount of undeveloped land is less than 90%. The multipliers for 2010, 2017, and 2025 are based on the ratio of the exponential equation run using the extent of undeveloped land in 2006 over the extent of undeveloped land in 2010, 2017, and 2025 respectively. For the year 2025, the average densification factor was 0.97 and the lowest densification factor was 0.53.*

There are some existing urban growth models – did they look at Nowak et al.'s Journal of Forestry article on projected urban forest changes for the country paper? Doesn't "forests on the edge" make projections? It would be worth looking into for comparing your growth estimates with what Nowak and Stein came up with.

> *The Forests on the Edge paper cites the projections made in Nowak and Walton (2005). The projections in Nowak and Walton (2005) were derived by linearly extrapolating the average rate of change in the percent of a county's land defined as urban from 1990 to 2000 through the year 2050 (using the US Census defined "urban area" boundaries for 2000 and applying the definitions of urban areas in 2000 to 1990 Census Block Groups and Blocks). The authors forecast that DE will become 39.5% "urban" by the year 2050 and Maryland will become 37.5% urban. These are the only two states which fall completely within our CBLCM study area. For the year 2000, US Census Urban Area boundaries indicate that there were 1,115,469 acres of urban land in Maryland representing 18% of the total area and 194,308 acres of urban land in Delaware representing 15% of the total area. For the year 2001, the CBLCM defined urban and suburban areas compose similar acreages as the US Census Bureau's urban areas (18% in MD and 14% in DE). Using data reported in Table 1 from Nowak et al., (2005), urban area growth rates from 1990 to 2000 for the Bay states were:*

	1990 - 2000 (%)
DE	37.6
PA	25.5
WV	24.0
VA	21.5
MD	20.8
NY	20.8

The CBLCM estimates that between 1992 and 2001, urban and suburban lands increased 7% in MD and 10% in DE and total developed lands increased 11% in MD and 19% in DE. Because the CBLCM and Nowak and Walton (2005) both basically linearly extrapolated historic growth rates into the future (the Gompertz equation results in linear extrapolations for short term growth under most conditions), differences in the estimated historic rates of change lead to large discrepancies for future years. The LDT contends that the CBLCM trend scenario is more realistic than the projected change in urban area based on historic changes in Census defined urban area boundaries because the CBLCM trend scenario is constrained by the most current population projections (which account for some of the demographic factors impacting housing demand) and by the availability of suitable land for new development.

The possible reasons explaining why the rates of urban growth derived from the change in 1990 and 2000 Census Urban Area boundaries is so much larger than the CBLCM rates of change have not been fully explored. One possible reason may be the existence of many 1990 blocks and block groups with population densities just under the threshold values used by the Census Bureau to define urban areas. The Census Urban Area and Urban Cluster boundaries are primarily defined based on population density thresholds at the Census block group level (1000 persons per square mile, ppsm) combined with surrounding blocks exceeding 500 ppsm.

Nowak, D.J. and J.T. Walton, 2005. Projected Urban Growth (2000-2050) and Its Estimated Impact on the US Forest Resource. Journal of Forestry v103 (8):383-389.

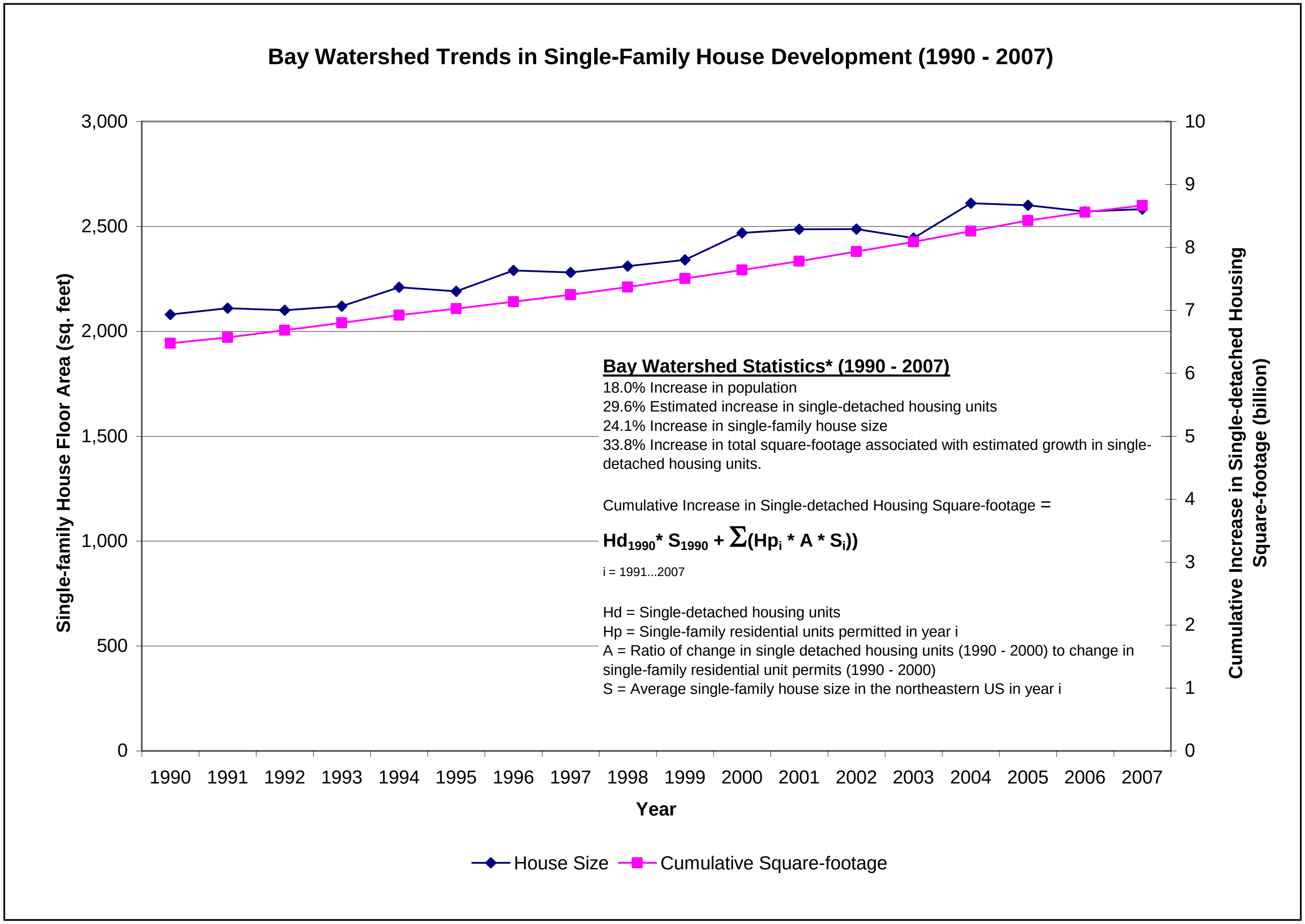
Nowak, D.J., Walton, J.T., Dwyer, J.F., Kaya, L.G., and S. Myeong, 2005. The Increasing Influence of Urban Environments on US Forest Management. Journal of Forestry v103 (8):377-382.

At the simplest level, the expectation would be a 1:1 change in population and developed lands. So, if an increase from 17 to 20 million people by 2030 (17% increase), then expectation would be roughly similar increase in developed land cover. Is there a reasonable basis for the ~60% increase cited on page 1

The estimate that "developed land in the watershed will increase by more than 60%" (STAC 2003, page 60) between 2000 and 2030 is associated with the "Recent Trends" scenario discussed in the report. The baseline extent of development appears to be derived from an early version of the Phase 4 model which likely underestimated the extent of development in the watershed. For this scenario, it was assumed that the average size of new lots would be large (0.91 to 1.45 acres per new household) with high levels of impervious cover (0.21 to 0.31 acres per household) and no infill or redevelopment. The population projections used to support the Recent Trends scenario were developed by NPA Data Services in 1999 although the LDT has been unable to locate a copy of them. Certainly differences in the population projections used in Chesapeake Futures compared to those used in the CBLCM (county-level projections downloaded in 2010 from state agency websites) may partly explain the high growth (60% increase) scenario.

An underestimate of baseline developed lands combined with large lot size and impervious surface coefficients also help to explain the high growth scenario.

Trends noted in the Chesapeake Futures report and not accounted for in the CBLCM are increases in the average size of single-family houses, residential lot size, and big-box retail development. The LDT has compiled more recent trends in the figure below.



Bay Watershed Statistics* (1990 - 2007)

18.0% Increase in population
29.6% Estimated increase in single-detached housing units
24.1% Increase in single-family house size
33.8% Increase in total square-footage associated with estimated growth in single-detached housing units.

Cumulative Increase in Single-detached Housing Square-footage =

$$\mathbf{Hd_{1990} * S_{1990} + \sum(Hp_i * A * S_i))}$$

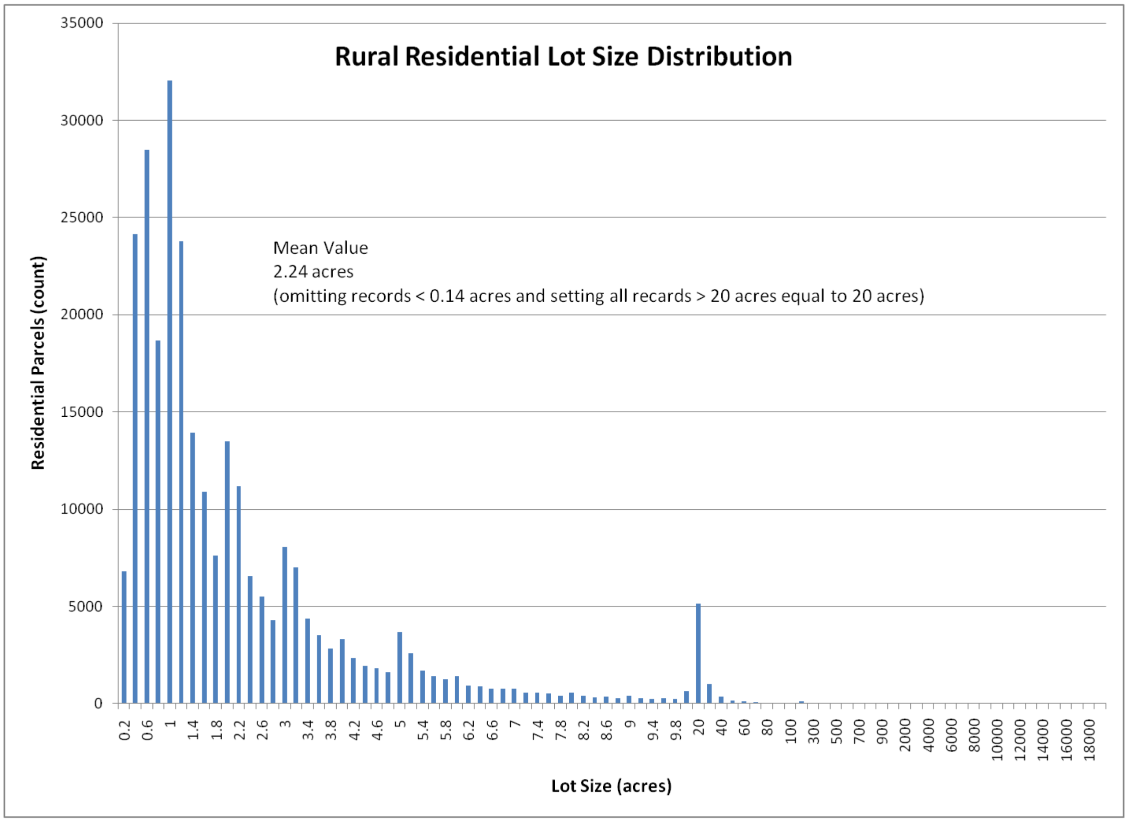
i = 1991...2007

Hd = Single-detached housing units
Hp = Single-family residential units permitted in year i
A = Ratio of change in single detached housing units (1990 - 2000) to change in single-family residential unit permits (1990 - 2000)
S = Average single-family house size in the northeastern US in year i

Model Assumptions and Sensitivity

Due to the large number of modeled variables in the Chesapeake Bay Land Cover Model (CBLCM), STAC expressed concern that there is potential for dramatically inflated or deflated numbers due to minor changes in one or several variables. To address this issue, STAC suggested that the LDT conduct a sensitivity analysis on the variables used in the CBLCM. Rather than reporting a single estimate of future growth in a region, a range of values representing low and high thresholds should be considered.

> *In response to these suggestions, the LDT evaluated several variables used in the methodology that are known to have a significant impact on the estimates of developed land. The LDT is also working closely with the CBPO Watershed Model Team to implement a series of model runs based on a range of developed land estimates. This work will be carried out following calibration of the Phase 5.3.2 model.*
>
> *Rural residential lot size, impervious surface coefficients associated with different development intensities and with rural and suburban residential development, and road width are three variables that significantly impact estimates of current, past, and future developed lands. To estimate the amount of impervious and pervious developed lands associated with residential development in rural areas, single-detached housing units reported by the Census Bureau and falling within areas mapped by the LDT as "rural" were counted and multiplied by an estimate of average rural residential lot size. To estimate average rural residential lot size, Maryland's 2007 PropertyView database was used. All "residential" parcels that were also described as 'split foyer 2 levels of living area, 'split level 3 or more levels of living area', or 'standard single family unit 1, 2 or 3 story' (to approximate the Census Bureau's definition of "single-detached") and falling within areas classed as "rural" were queried and their distribution plotted.*



Depending on the measure of central tendency used to represent this distribution, the extent of pervious developed lands in the watershed can vary by several 100,000 acres. Outliers in the distribution were eliminated based on the assumption that rural lots less than 0.14 acres would contradict our sampled average impervious coefficient associated with rural residential lots in the watershed (which was 0.14 acres). The LDT also assumed that lots over 20-acres in size would most likely not be considered typical of "residential" lots and were more likely hobby farms if they contained extensive open areas. So as to not completely omit these records from the analysis, all parcels greater than 20-acres in size were set equal to 20-acres. (Note that in every modeling segment, the percentages of land in forest and open space within 60m of secondary rural roads were calculated and the percentage of land in open space was multiplied by the extent of rural residential land to determine the amount of rural residential land in "lawn". Therefore, we do not assume that a 20-acre residential lot is all lawn). The mean rural residential lot size between 0.14 acres and 20 acres is 2.24 acres.

Because Landsat-satellite derived land cover maps were used to depict areas of dense development, termed "urban", impervious surface coefficients for urban lands were derived by overlaying the 2001 National Land Cover Dataset Impervious Surface layer on the 2001 Chesapeake Bay Land Cover Data Series (masking out only urban areas). For every county, mean values of imperviousness were derived for each of the four developed land cover classes (DOS, LID, MID, and HID). The LDT decided to use County coefficients rather than state scale or Bay-watershed coefficient because the coefficients varied significantly by county and the county *coefficients yielded the best results when comparing Phase 5.3.2 estimates of impervious surface with local high-resolution impervious surface datasets. In Montgomery County, Maryland for example, the actual extent of impervious surface based on local data is 37,600 acres. Using just satellite land cover and state level impervious coefficients yielded an estimate of 27,700 acres. County-level coefficients yielded an estimate of 29,900 acres and county-level coefficients combined with impervious surface estimates for suburban and rural roads and residences yielded an estimate of 35,362 acres. County-level coefficients were generally higher than state-level coefficients in the more developed counties. Also, in Montgomery County, impervious surface coefficients for the Phase 5.3.2 developed classes derived using local high-resolution impervious surface data were even higher than the NLCD coefficients derived for the same areas in the county. In very rural counties, county-level NLCD coefficients of imperviousness were extremely and unrealistically low (e.g., 1% for Developed Open Space). To avoid anomalies where a small number of developed pixels yield extremely low coefficients (this phenomena did not occur at the high end of the range), for every state a county-level first quartile estimate of mean imperviousness was developed for each of the four developed land cover classes. No county was allowed to have mean impervious surface values for any of the four developed classes below the first quartile minimum class value for all counties in their state.*

To estimate impervious surface in suburban and rural areas, a random sampling of impervious surfaces associated with residential lots was performed. Fifty samples were allocated per state along the secondary road network in rural areas and 50 additional points were allocated per state along the secondary road network in suburban areas. Impervious surfaces associated with the nearest residential house to each sample point were measured. Impervious surfaces were measured using heads-up digitizing from aerial imagery (typically 2005-2009 vintage). Many points were thrown out because they fell in remote or non-residential areas. The final sample sizes in each state for suburban areas ranged from 17 to 34 and from 21 to 44 in rural areas. The median values were used as the best measure of central tendency because of the relatively small sample sizes and the presence of a few high outlier values. Rural residential impervious surface coefficients are consistently higher than suburban residential coefficients due to the common presence of longer driveways, outbuildings and other structures.

Median	**Suburban**	**Rural**
DC	0.073	0.177
DE	0.116	0.149
MD	0.135	0.177
NY	0.095	0.113
PA	0.077	0.148
VA	0.085	0.150
WV	0.076	0.109
All States	0.094	0.140

In addition to estimating the extent of impervious surfaces associated with all single-detached residential houses in suburban and rural areas, the extents of impervious surface associated with all suburban and rural roads were also estimated. Suburban and rural roads constitute 31% of the impervious surfaces in the Chesapeake Bay watershed. Eighty-nine percent of all roads are 2-lane, 2-way roads. Depending on assumptions about shoulder width, a difference of 4' along 2-lane roads can make an overall difference of 70,000 acres of impervious surface in the Bay watershed. Because regional estimates of impervious surface are very sensitive to road width assumptions, the LDT conducted a literature review of design road widths and sampled road width by major road type in Pennsylvania, Maryland, and Virginia. Four major road types (2-lane/2-way suburban, 2-land/2-way rural; 4-6 lanes, 2-way; and 8+ lanes (2-way controlled access) were identified composing over 94% of all roads. The 2-lane, 2-way class was split into rural and suburban categories because suburban areas may contain parking lanes, bike lanes, hardened medians, extra-wide residential streets or other features that effectively widen the road feature.

Road Type	% of Rd Miles	Literature Range	Sample Range	Sample Mean	Selected Width
2-lanes (2-way)	88.8%	22 – 36	13 – 50	23- 25	22 (rural) 26 (suburban)
4-6 lanes (2-way)	2.4%	42 – 84	26 – 104	72	26, 36, and 72
8 + lanes (2-way) (controlled access)	2.9%	116 - 120	70 - 222	120	116

We agree that both the residential impervious and road width sampling efforts are biased samples- meaning that more samples are likely to occur in areas of dense roads. We tried to limit the degree of bias by ensuring that all samples were at least one mile apart (hence the low number of samples in DC).

One STAC reviewer recommended percent impervious associated with various land use types, namely developed open spaces should be adjusted to reflect that these types of land uses often behave effectively more like pervious surfaces.

USGS is initiating a study in the Bay watershed to quantify the impact of impervious surface connectivity to streams on stream flow. Results from the study may inform a relative weighting of impervious surface effects in future versions of the Chesapeake Bay Watershed Model.

The reviewers also questioned the assumptions surrounding impervious surface associated with low density residential development. They suggested the modelers conduct a more rigorous assessment using a PI method. Specifically they suggested the following: "one thing that would be pretty quick and easy would be to identify "change polygons" found using the change detection, and then over each of them, generate a grid of dense points. Turn this grid of points into a .kml and load it into Google Earth, and have an interpreter for, say 50 of these areas, count dots that show the disturbance footprint using the Google Earth historical imagery, which is accessible via a button on the interface. It would be quick and they could supply this information to their sensitivity analysis.

The LDT will initiate this proposed approach for ground-truthing areas of change identified in the Chesapeake Bay Land Change Data Series (1984-2006). For Chesapeake Bay counties, historical imagery previous to the early 1990's are not commonly available on Google Earth nor is it even available from the state archives. Therefore, our accuracy assessment will focus on the period 1992 to 2006. Seasonal differences and image quality will also impact the accuracy of data produced through this approach so those differences will also be noted in the analysis.

Another reviewer noted that the foundation of the methodology is the NLCD Percent Urban Impervious dataset – impervious surface is derived specifically from it from NLCD for example. So, the question is – how reliable is it? It would be useful to at least describe the accuracy assessment of PUI, errors of omission and commission, and whether these tend to occur more or less in urban vs. rural areas.

Comparing the Phase 5.3.2 land use dataset with local impervious data within the urban zone allowed us to develop impervious surface coefficients for the Chesapeake Bay Land Cover Data series urban land cover classes (e.g., Developed Open Space, Low Intensity Developed, Medium Intensity Developed, and High Intensity Developed) based on local (1m resolution) impervious surface data. We then compared those locally derived coefficients with coefficients derived from the 2001 NLCD Impervious Surface Dataset. The table below highlights results for four counties. The locally derived coefficients are almost all higher than the 2001 NLCD Impervious Surface dataset and have a significant impact on our models estimates of impervious cover. For Phase 5.3.2, we switched from using state-level impervious surface coefficients derived from the NLCD to county-level coefficients which were generally higher in the most developed counties within each state. However, as shown in the table, the county coefficients are still low. Developing a new more recent impervious surface dataset derived from Landsat satellite imagery that is trained on the current wealth of high-resolution impervious surface datasets may improve the accuracy of all county impervious surface coefficients while ensuring spatial and temporal consistency across the watershed.

SUSSEX, DE	Developed Open Space	Low Intensity Developed	Medium Intensity Developed	High Intensity Developed
P532 Land Use	11.1	25.2	56.0	72.8
Local Data	16.5	29.4	50.7	76.0
NEW CASTLE, DE	**Developed Open Space**	**Low Intensity Developed**	**Medium Intensity Developed**	**High Intensity Developed**
P532 Land Use	10.0	23.4	52.1	75.6
Local Data	16.3	35.5	58.6	81.7
KENT, DE	**Developed Open Space**	**Low Intensity Developed**	**Medium Intensity Developed**	**High Intensity Developed**
P532 Land Use	11.0	27.0	60.9	85.3
Local Data	16.6	34.6	60.2	87.1
LANCASTER, PA	**Developed Open Space**	**Low Intensity Developed**	**Medium Intensity Developed**	**High Intensity Developed**
P532 Land Use	9.7	19.8	46.9	70.5
Local Data	7.9	19.7	42.4	71.3

In addition a reviewer suggested comparing the ISA estimates for the CBW for those reported in Theobald et al. (2008) that also specifically incorporate exurban and rural housing development.

> *We will further explore the relationship between housing density and impervious surface at multiple scales. While housing density and developed lands as measured using Landsat satellite land cover data are well correlated at the County scale, at the modeling segment scale some deviations occur due to the relative dominance of commercial/industrial areas, roads, or agricultural structures at the local level. By complementing housing data with Landsat-derived land cover information in urban areas and road information in rural areas we were able to capture some of this local variation in the Phase 5.3.2 land use dataset.*

Another reviewer posed the following questions. What is the weighting function to allocate housing density based on road density? Is there an empirical basis to this? How does this compare to other dasymetric mapping techniques, such as Eicher's paper?

> *Given the extensive use of spatially distributed Census variables in the CBLCM, over the coming year the LDT will further investigate the accuracy of our dasymetric mapping technique and examine alternative published dasymetric mapping methods.*
>
> *The weighting function we used was relative road density at a 30m cell resolution. Road density was mapped using the line density function in ArcGIS with a 200m radius and parameterized so that the output would be a 30m resolution raster dataset. Public and protected lands, open water, emergent wetlands, and beach areas were excluded. The values in the road density raster were summarized for each and all census block groups. Block groups were then converted to a raster dataset using the total road density attributes as the raster value. Finally, the original road density raster was divided by the total road density raster to produce a relative road density raster layer that was used to weight the distribution of the population, total housing unit, and single-detached housing unit attributes of the block groups.*

One reviewer requested the modelers clarify the logic of allocating housing units by block-group, when blocks (a finer scale) are available? There may be other attributes that you use from the block-groups, but for those variables you could aggregate up the block-level information… it seems like you are losing a great deal of spatial variation because of this assumption.

> *There is an 80% correlation at the modeling segment scale between estimates of population using Census block data vs. Census block group attributes. We chose to use block groups because block groups contain single-detached housing unit information in addition to population and total housing unit values. Moreover, in rural areas, roads form the boundaries of Census blocks so allocating attributes from blocks to the underlying patterns of road density does not achieve a plausible result whereas the large block groups often encompass concentrated areas of roads in rural areas which provide a more intuitive allocation. Lastly, the US Census Bureau's American Community Survey has taken the place of the Decennial Census Long Form and the Block Group will be the smallest unit for which detailed housing unit information is available in the future.*

It isn't clear how the estimates within a unit of analysis (modeling segments) interact with their neighbors. That is, it's clear that the Gompertz curve is used to estimate the number of new units within a segment, but is there any assumption that as a segment begins to reach capacity that housing units "spill over" to adjacent units – it appears that each segment operates independently of what is going on in its surroundings – yet growth pressures often spread across broad units (and across stateliness).

> *Each modeling segment does operate independently. The LDT agrees with the reviewers that in actuality, growth pressures are spread across broad units. Transportation and economic relationships exist between townships, cities, and counties that result in "spill over" and other cross-over effects to adjacent units. Such relationships do not necessarily occur among modeling segment units or at the scale of those units. A growth model that represents the economic and transportation linkages among jurisdictions and feedbacks between county level drivers of growth and local factors (e.g., availability of land) would be preferable to the current approach. These are ideas that will be considered for future versions of the CBLCM.*

Land Use Characterization and Mapping

STAC reviewers expressed concern of the lack of explicit representation of Commercial/Industrial land use types in the CBLCM and therefore the model may underestimate these land use types and overestimate residential areas (page 2). Commercial/industrial lands are commingled with residential land uses in NLCD, yet there is no explicit representation of this land use type in your model. Assuming that commercial area will also expand to meet demands of more people, the residential footprint would be underestimated by your approach.

> *In Landsat satellite derived land cover datasets, commercial and industrial land uses are comingled with residential land uses. Our approach assumes that changes in commercial and industrial lands are directly proportional to changes in residential housing at the modeling segment scale. At the county scale and within older towns and cities, this assumption is well supported by the strong correlation between total housing units and the extent of developed land. We realize that the correlation between housing and all developed lands degrades at finer sub-county scales, particularly in areas developed over the past 30-40 years where commercial/industrial development was purposely separated from residential development through zoning and other means. Unfortunately, we lack access to fine scale commercial/industrial data that could be used to separate commercial/industrial from residential developed lands and translating employment projections into a demand for land is complicated by the lack of regional data on suitable vacant office space and the great variation in office space demands among and within different employment sectors (e.g., financial, government, military, retail, etc.).*
>
> *It is not clear whether assuming that changes in commercial and industrial lands are directly proportional to changes in residential housing results in current and future overestimates or underestimates of developed lands. For example, assume an industrial park and airport dominate the developed footprint in one modeling segment. In this area, the developed footprint per housing unit will be very large and small forecasted changes in housing would lead*

to disproportionately large changes in developed lands (overestimate). However, because so few housing units exist in the area, the amount of change in housing forecasted or back-casted would be likely be very small and so would any corresponding overestimate.

Considering another example, assume the developed footprint in a segment is dominated by a large new and occupied subdivision. In the near future, service sector (e.g., gas stations, "new town centers") and institutional developments (e.g., churches, schools) are likely to follow to provide for the needs of the new residents. In this area, high levels of past and future change would be simulated so long as the amount of remaining undeveloped land in the segment is sufficient to not constrain future growth. The past simulations may be accurate because prior to the subdivision, the area was largely undeveloped. Future simulations may also be accurate, but for the wrong reason. Our model would estimate significant growth due in residential development while the future growth could be largely commercial and institutional. Because we don't differentiate among the types of development, the extent of development forecast for the future may be correct but the implied cause may be incorrect.

All this said; the LDT recognizes the need to separate commercial/industrial from residential development in future versions of the CBLCM so that fine-scale simulations of growth are more plausible given the drivers and nature of change.

According to the methodology, "The 'bare-construction' land use class in the Phase 5.3 model is derived by multiplying the annual acreage of change in impervious surface by a factor of 2.5 (based on information provided by the Maryland Department of the Environment that the area disturbed for construction is ~ 2.5 times larger than the final developed area)." The reviewers felt this was an overestimate, and found it hard to believe that's the rule, especially with residential housing, which is probably the largest source of land cover change in that area. The reviewer suspected bare soil might get mislabeled as impervious surface by CBLCD, so might already include some of the added footprint.

The following information is taken from the Phase 5.3.0 documentation developed by USEPA. "Detailed records from all Maryland counties indicate that, on average, a unit area of imperviousness is generated from a construction permit covering about 10 times that impervious area but that the area cleared for construction was 2.5 times the impervious area, or on average one-quarter of the total area of the site covered by the construction permit. Accordingly, the average yearly change in impervious surface was multiplied by 2.5 to calculate the Phase 5.3 bare-construction acreage. Although that calculation is static and does not reflect year-to-year changes in construction, it provides a uniform methodology for the entire Phase 5.3 study area. Maryland permit data are available as state totals for the years 1998, 1999, 2001, 2002, 2003, and 2004. Phase 5.3 bare-construction area annual estimates for Maryland (here multiplied by a factor of four for consistency with total permitted construction acres as described above) fall within these reported values (Figure 4-12)."

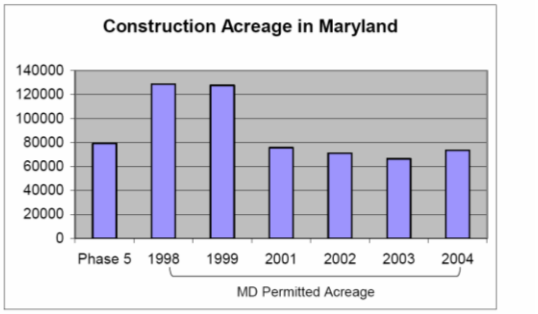


Figure 4-12. Permitted construction area in Maryland for 1998 to 2004 compared to the Phase 5.3 annua average bare-construction (bare-construction multiplied by a factor of four as described in the text) are

Attempts were made to minimize confusion between areas classed as "barren" in the 2006 CBLCD dataset and developed, unconsolidated shore and extractive lands (e.g., quarries and surface mines). Patches of barren land cover in the raster land cover dataset were overlaid on extractive polygons and overlaid on a polygon layer created from a mosaic of unconsolidated shore and open water pixels. Barren patches intersecting either extractive or unconsolidated shore pixels were reclassed as either "extractive" or "unconsolidated shore". Remaining barren patches that intersect developed land patches were reclassed as "suburban" areas.

STAC reviewers proposed an alternative suggestion for deriving the "extractive" land cover. As outlined in the Methodology that was reviewed by STAC, extractive land uses were determined using ancillary data and the land cover was adjusted using point data, expanded to the size of the disturbed area. The reviews thought the point data should be overlain with the appropriate land cover year and where the underlying land cover was barren or developed, then change it to extractive. The reviewers felt this would better represent areas that are not a circle of disturbance but rather the actual area which could be irregularly shaped (p. 9 paragraph 1).

We explored this possibility early on in our investigations. It turned out, however, that the CBLCD is not accurate enough to permit this kind of disturbed surface mine delineation. Surface mine land cover is a mix of land cover types that blend with the surrounding area. They can contain Shrub Scrub, Bare Soil, Water, different Forest and Developed land cover types as well as Pasture/Hay and Grassland/Herbaceous. Moreover, reclaimed surface mine areas, which have been documented as having a more impervious index as compared to undisturbed land, may lie in an area that is completely covered by forest according to the satellite land cover data with no way to differentiate between reclaimed surface mine and undisturbed forest land.

For West Virginia we used the exact polygon delineations of permitted surface mine disturbed areas provided to us by the Office of Abandoned Mine Lands and Reclamation (AMLR) of the West Virginia Department of Environmental Protection. This department further advised us, however, to reduce our acreage tabulations of these polygons by 1/3 to produce a more accurate estimate of actually disturbed surface mine acres. We followed this procedure in producing our final acreages for West Virginia.

> *In some states, only point data was provided and we buffered those points to create circles corresponding in size to the area reported as "disturbed" in the point attribute file. While we did overlay extractive points and polygons on the land cover dataset, we only did so to improve the accuracy of the urban classification. We also intersected the extractive circles (derived from points) and polygons on the 2006 CBLCD to assess the potential extent of confusion with developed land cover classes. We then summarized the spatial extent of extractive circles (derived from points) and irregular polygons for each modeling segment and reported those acreages irrespective of the underlying land cover. In each modeling segment, the amounts of impervious and pervious developed land underlying extractive areas were subtracted from the final extents of impervious and pervious developed lands.*

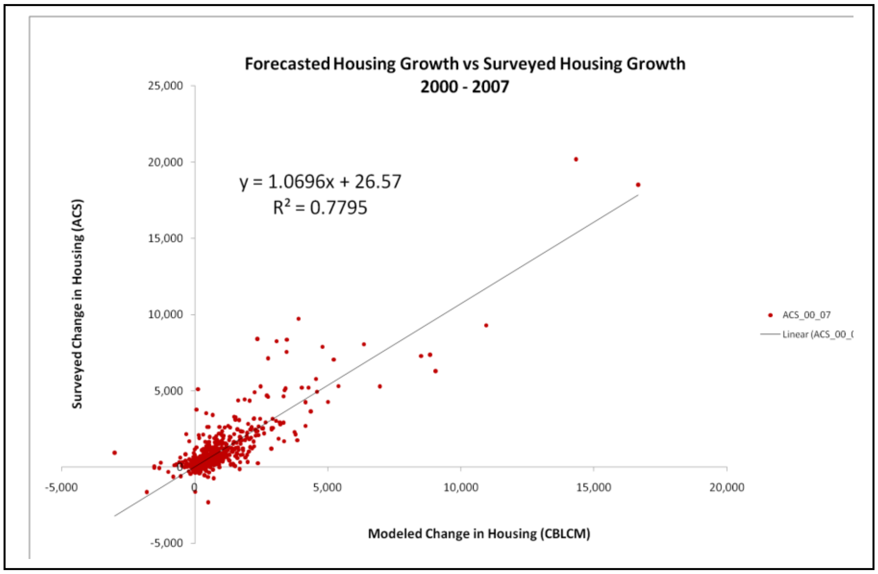
Data Considerations

Additional socio-economic data, such as income, should be incorporated in order to aid in predicting if households will convert from septic to sewer. This would enable the modelers to calibrate the likelihood of sewer hookup vs. a socio-economic data index in order to weight averages. Reviewers felt the implementation of this recommendation would help prevent over- or under-estimating sewerage amounts, which would have a large affect on the watershed model.

> *This is an interesting idea and would be applicable to areas where sewer hookups are not mandatory. While we model expansion of existing sewers along the road network based on the average distance of expansion derived from comparing current with planned sewered extents in Maryland using Maryland Department of Planning data, the modeled expansion to 2025 was fairly minor in most cases. Tetra Tech collected planned future sewered area information for some of the major Waste Water Treatment Plants although it is not clear when exactly, or if, the expansions will occur. The potential temporal and spatial uncertainty inherent in our data on planned future sewer service areas probably overwhelms any potential errors associated with a home owner's ability to pay for hookup.*

There is some discordance between the dates of data used, such as using 2000 housing density with roads from 2006, and aerial photographs (from Google and others) that are presumably of 2008 or 2009 vintage. Although this discordance is typically unavoidable, it would be useful to state what potential bias these would have when comparing different date datasets.

> *In December 2010, the US Census released the American Community Survey Data 5-year Estimates for year 2000 Block Groups. These data represent the cumulative results of surveys conducted from 2005 – 2009. We used our 2005 road dataset and dasymetric mapping technique to spatially disaggregate the ACS data to modeling segments. We then compared the ACS estimates of Total Housing Units (2005-2009) with our modeled estimate of Total Housing Units for the year 2007 (see graph below). Based on this comparison, we substituted the ACS data for our modeled estimate of Total Housing Units for the year 2006 (proportioning the ACS estimate from 2007 to 2006 based on the relative degree of modeled change in Total Housing Units from 2007 to 2006).*



Why not use the Croplands Data Layer or CLU dataset in some way? They have classes like NLCD, but have very detailed agricultural information.

> *We did use the Cropland Data Layer in previous versions of the Watershed Model (v5.0). However, decided to not use the CDL in Phase 5.3 or Phase 5.3.2 because the benefits of doing so were unclear. Using the CDL to discriminate crop types in areas classed as agriculture in the Phase 5.3.2 land cover dataset will be considered in near future to inform local implementation of Watershed Implementation Plans (local plans used to guide compliance with the Bay TMDL).*

The EPA Geospatial dataset of facility locations could be used to better identify locations of “chick houses” that caused some misclassification. That is, using the NAICS code 112 one can pull out locations of permitted feedlots, chicken houses, etc.

> *We examined using a dataset produced by NAVTEQ and license to the US EPA which contained some point data for agricultural facilities such as chicken houses. Cursory investigation however showed that the data omitted a large number of facilities observable in recent aerial photography. Due to questions about the accuracy of the data, they were not used. We would be very interested in learning if the dataset mentioned by the reviewers is the same as the one created by NAVTEQ or different.*

There is the census 1990 and 2000 roads datasets, and Dynamap 1990 and 2000 datasets, all of which are free and pretty good, though admittedly not as good as the Navteq. The reviewers assumed that the census-road density-other data merge and modeling to predict future housing development was developed by academia and already reviewed and used in NJ. If not, a more intensive review of the methods might be required.

> *The use of secondary road density to map potential subdivisions was not part of the original GAMe model used in New Jersey. The LDT used a circular line density (200m radius) function to both disaggregate and redistribute housing unit and population census data and to define and*

map suburban areas. Alternative radii distances were explored (e.g., 100m, 500m, and 1000m) by thresholding and overlaying the secondary road density rasters (created based on different radii) on top of aerial images. Increasing the radius above 200m resulted in more commission errors and decreasing the radius below 200m resulted in more omission errors. Therefore, 200m appeared to provide the best compromise for minimizing both commission and omission errors. This approach will be submitted to a scientific journal for publication and in the process, undergo a USGS and external peer review.

The LDT was not aware of the Dynamap 1990 road dataset nor were we aware about how it compares with the Dynamap 2000 dataset. We are familiar with the Census Tiger road files. The Tiger road datsets were all produced from aggregating county data reported on a rolling update basis. The US Census Bureau has said that their 1992 Tiger road files are not comparable to later files as roads added in 1998 or 2000 may have been present in 1992 but not reported. Other issues such as mapping accuracy and standards have also changed over time and will change again with the 2010 Census. The LDT will evaluate whether these issues are apparent when comparing the Dynamap 1990 and 2000 products. If the data accurately represent the extant road network in 1990 and 2000, the LDT will use them in future versions of the CBLCM.