REVIEW OF LAND-USE AND LAND COVER DATASET AND METHODOLOGY

SCIENTIFIC AND TECHNICAL ADVISORY COMMITTEE

The US EPA Chesapeake Bay Program (CBP) has requested an urgent, short-turn around peer review of certain critical land use and land cover inputs to the Phase 5 Bay-wide watershed model. In responding to these questions, the Scientific and Technical Advisory Committee (STAC) assembled a diverse team of experts to triage aspects of these questions most relevant to the establishment of Bay-wide water quality regulation and provide rapid feedback to the CBP.

CHARGE QUESTIONS

The Chesapeake Bay Program requested assistance with the following questions:

1. How important are low-density residential development and rural and suburban roads to estimating nutrient and sediment loads to the Bay?
2. Are the assumptions and methodology used to estimate the extent of developed lands scientifically sound?
3. Are the assumptions and methodology used to forecast changes in developed lands, agriculture, and populations on sewer and septic systems scientifically sound?
4. Are the assumptions and methodology used to estimate and spatially represent extractive lands scientifically sound?
5. What improvements to the methodologies are needed?

The reviewers often combined responses to Charge Questions 1-3. A synopsis of key findings is prevented below. Verbatim reports from individual reviewers are presented in the following sections.

PANEL

- Chris Pyke, STAC/US Green Building Council (review coordinator)
- Kurt Gottschalk, USDA Forest Service
- David Theobald, Natural Resource Ecology Lab, Colorado State University
- Andrew Lister, USDA Forest Service
- Tonya Lister, USDA Forest Service
- Tanya T. Spano, Department of Environment Programs, Water Resources Program, Metropolitan Washington Council of Governments
- Thomas Johnson, Office of Research and Development, US Environmental Protection Agency
- *Administrative support from Liz Van Dolah, Chesapeake Research Consortium*

SUMMARY[1]

Overall, the reviewers were impressed by the technical quality the work. They recognized that the Chesapeake Bay Program has confronted challenging conceptual and technical issues and appropriately applied state-of-the-art approaches.

Charge Questions: How important are low-density residential development and rural and suburban roads to estimating nutrient and sediment loads to the Bay? Are the assumptions and methodology used to estimate the extent of developed lands scientifically sound?

Reviewers recognized that low-density development is a prominent feature of the watershed, but they believed that there are significant barriers to understanding the implications of this type of land use for the Chesapeake Bay.

Reviewer #1 indicated, "*I think it's fair to say this question is pushing the limits of current understanding of watershed hydrology…the issue is connectivity.*" Connectivity at scales relevant to low-density development is not explicitly represented in the Phase 5 watershed model, "*HSP is a lumped parameter model. The pattern of land-use within model segments is not explicitly reflected in simulations of flow and nutrient loading…As a calibrated model, however, calibration of the Phase 5 model with improved land-use data should capture to some extent the influence of low-density land-use including pattern and extent, albeit in an opaque way.*"

Reviewer #1 summarized, *"Overall, I think the methods employed to estimate the extent of developed lands are well conceived, transparent, and while based on a number of assumptions, scientifically sound."*

Reviewer #1 continued, "*Formal error analysis isn't necessary or, I suspect, even possible, but I think it's important philosophically to understand when using this data (referring here to future projections) that the data represents a scenario or plausible future condition rather than a forecast.*" Later in the comments, Reviewer #1 continues, *"…you refer in many places to future land-use dataset as a "forecast." I would argue/urge you to present this data as a scenario, i.e., a plausible future condition, rather than a forecast. The term forecast implies a single most likely future…While your methods are reasonable and scientifically sound, many other potential scenarios of future land-use could also be produced that are equally reasonable and sound."*

Reviewers #2 and 3 had similar observations: *"We think that given what they are trying to do, they have done an excellent job and used logical and sensible analyses in an attempt to aggregate factors that can contribute to Chesapeake Bay pollution. Our main concern is that there are SO many modeled variables in this dataset that there is a large potential for dramatically inflated or deflated numbers. What I'd like to see is a large sensitivity analysis."*

Reviewer #4 echoed many of the same issues while also raising the issue of other land-use and land cover types, *"Commercial/industrial lands are commingled with residential land uses in NLCD[National Land Cover Data], 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."*

[1] The Summary was written by Chris Pyke, Vice Chair, Scientific and Technical Advisory Committee

Reviewer #4 made the following observation about the overall level of land-use and land cover change expected, *“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?”*

Reviewer #5 tried to estimate the direct 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).

Reviewer #5 observed that, *“Fundamental changes in the real estate market dynamics make it grossly inadequate to use historic parameterization for these estimates [residential growth].”*

Reviewer #6 commented, *“The rationale used in 5.3.1 to refine the estimated acres of impervious surface generally appears to be based on a rational method; but it is difficult to determine if the resulting increase in impervious surface (i.e., presumed increased accuracy) necessarily results in an overall increase in the amount of nutrient/sediment loads associated with those additional acres of impervious surface…”*

Reviewer #6 continued, *“Bottom line, the increased accuracy of estimating this acreage should not automatically be presumed to result in increased accuracy in loading rates and/or reallocation of loads to these new identified acres… It is not good enough to assume that corrections can/will be made in 2017…It is critical therefore, that detailed reviews of various segments need to be done to ensure that watershed model estimates that would reallocate loads given these additional impervious acres bear a reasonably close relationship to actual load/calibration data for those land uses.”*

Charge Questions: Are the assumption and methodology used to forecast changes in developed lands and populations on sewer and septic systems scientifically sound?

It was substantially more difficult for the reviewers to comment on the specifics of the sewer and septic system forecasting. This may reflect (1) the challenge of finding individuals with depth of experience in these issues and (2) the difficulty to creating a framework of reference to evaluate staff work in this area. In the latter case, the issue is that the staff analysis makes a number of important assumptions and generalizations. These may be practical and reasonable, but, unfortunately, there a few ways to independently assess their validity. Consequently, there may be no objective reason to question the assumptions, but there is also no basis to affirm them either.

Reviewers #2 and 3 were generally supportive of the approach, writing, “*We think this was a clever approach, although they might want to factor in some Census socio economic data when predicting if sewers are going to go in and/or be used…they could calibrate a “likelihood of sewer hookup” vs. socioeconomic data index for each pixel…”*

Reviewer #6 had some significant reservations about the back-casting approach used to relate population with sewer and septic utilization: *“Back-casting Population on Sewer & Septic – The graph on slide #33, and the curve to the right of the 2006 data point, has some serious flaws as presented because the curve might well be bounded by the upper limit/maximum, but it cannot be presumed to predict the* *future trend and/or rate of change with the timeline that is shown (i.e., to 2030) given current/severe economic constraints; and the fact that vertical growth is really not accounted for in the future densification model (i.e., New Jersey example). It was agreed that the CBP's planned workshop this fall could help address this issue and help determine how to better relate demographic projections to increased sewer/septic needs."*

Reviewer #6 also elaborated on concerns regarding the use of projections: *"I recommended that because the future cannot be predicted with any accuracy, that various 'futures scenarios' be used rather than implying that any one line/curve/set of modeling assumptions can be relied up to predict future loads."*

Summary

A synthesis of the reviews suggests several high-level issues for consideration:

1. **Conduct systematic sensitivity analysis** to understand connections between regional land use and land cover and the Phase 5 watershed model.

2. **Develop multiple plausible land use and land cover scenarios** and explicitly consider uncertainty among plausible scenarios.

3. **Incorporate more mechanistic understanding of land market dynamics into land use and land cover modeling**. Current efforts are dominated by tools based on spatial and temporal patterns. These approaches are inadequate to develop policy-relevant future scenarios given fundamental changes in regional economics and structural social and demographic trends (e.g., aging populations and changes in consumer preferences). This will require new types of expertise and input to the CBP, such as from land use economists.

STAC has strongly recommended several related analyses in previous peer reviews of CBP land use and land cover modeling effort. The CBP did not specifically respond to these earlier comments, and it is clear that these prior recommendations were not addressed in the work reviewed here.

Closely related issues were raised again by this set of reviewers, including:

- The need for systematic sensitivity analysis;
- The utility of multiple plausible scenarios; and
- The need to communicate uncertainties associated with land-use and land cover futures.

The current review raises these issues again. The absence of work on these issues makes it difficult to respond to some of the charge questions and undermines CBP decision making.

In the future, it would be more productive for the CBP to **engage STAC periodically** in the development of land use and land cover modeling and representation approaches. This would allow for a more sustained and constructive engagement. This review was requested very late in the development and application of these methods under very tight deadlines. **The timing of the request offered few, if any, opportunities for feedback and iteration in these recommendations.**

In many cases, the justification for approaches used by the CBP is ultimately that the technique used was the best, practical option available with the time and resources available. These are legitimate rationales, but the results are challenging to peer-review. As noted above, **the CBP has not been responsive to important issues raised in previous reviews**, and this suggests a pattern of requesting peer review on nearly completed work products with few opportunities for improvement. This is an unacceptable pattern and inconsistent with the US EPA Peer Review Handbook (http://www.epa.gov/peerreview/pdfs/Peer%20Review%20HandbookMay06.pdf).

REVIEWER #1

1) How important are low-density residential development and rural and suburban roads to estimating nutrient and sediment loads to the Bay?

While it's clear that low-density development is a significant percentage of watershed area, the impact of such development on flow, N, P, and sediment is not well understood. I think it's fair to say this question is pushing the limits of current understanding of watershed hydrology. The key issues here are hydrologic connectivity, flow pathways, and juxtaposition with pollutant sources within the landscape. That said, there is relatively good documentation of the impacts of roads and road-stream crossings on watershed hydrology and water quality parameters including nutrient and sediment loadings. Generally speaking, as linear features in the landscape roads can intercept and disrupt natural drainage patterns. Roadside ditches and gullies also concentrate and channel water and associated pollutants over distance through undeveloped areas which can then discharge to rivers, streams and other water bodies. For example, the impacts of forest silvacultural roads on sediment loading to nearby water bodies are well documented (note this is an extreme example as much of the sediment impact is due to logging activities, but the basic principle of connectivity should apply to other roads). I would thus suggest that even at low densities the impacts of roads do have a significant impact on flow, and if juxtaposed with a sediment and/or nutrient source do contribute in a significant way to Bay nutrient loading.

The impacts of isolated, low-density homes (rooftops, driveways, etc.) are less clear. Again the issue is connectivity. If these areas drain to roadside ditches they could directly contribute pollutants to local water bodies. If draining to undisturbed woodlots where all runoff is infiltrated these areas may have no significant impact on local streams. An important caveat, however, is for N (and to a lesser extent possibly P) loading associated with groundwater. If low-density development involves an N source, such as lawn fertilizer or garbage, even if runoff is infiltrated, groundwater N concentrations could be increased resulting in increased N loading to local streams. The cumulative effects of increased N in groundwater from widely dispersed low-density development could thus potentially be significant. In summary, this is a difficult question to answer given current understanding of how different land-use pattern and processes interact with hydrologic, biogeochemical, and pollutant sources to influence flow, N, P, and sediment loading to streams. Current understanding does suggest the cumulative impacts of low-density roads are likely to be significant, and that the impacts of low density homes, while less clear, could also be a significant source of pollutants to the Bay.

2) Are the assumptions and methodology used to estimate the extent of developed lands scientifically sound?

Any effort such as this involves numbers of assumptions and simplified approaches for representing a highly complex state/process. Overall I think the methods employed to estimate the extent of developed lands are well conceived, transparent, and while based on a number of assumptions, scientifically sound. I commend the authors for their careful thought in assembling and extracting relevant information from a wide range of data sources to better represent developed lands.

My only comment is that when referring to the land-use dataset it would be good to acknowledge that characterizing and projecting future land-use is a complex and difficult task, and that this effort, however good, is just one attempt. Results are sensitive to the parameters and assumptions used to characterize existing developed land-use, methods to hindcast land-use, and methods to project future land-use. Formal error analysis isn't necessary or, I suspect, even possible, but I think it's important philosophically to understand when using this data (referring here to future projections) that the data represents a scenario or plausible future condition rather than a forecast. Other scenarios could also be developed.

3*) Are the assumptions and methodology used to estimate and spatially represent extractive lands scientifically sound?*

Same comment as for #2 above.

4) Are the assumptions and methodology used to forecast changes in developed lands, agriculture, and populations on sewer and septic systems scientifically sound?

Same comment as for #2 above. In addition, the representation of sewersheds will be extremely important as this connotes direct hydrologic connectivity of impervious areas and associated pollutant sources with rivers and streams (per my comments in #1). The approach used to identify and project sewered areas is reasonable and appropriate. The response rate on your survey is about what I would expect. Water utilities vary tremendously in levels of staffing, work loads and sophistication. I would encourage you to keep after those utilities that did not respond to your survey to try and get as much information as you can to improve current and future representation of sewered areas. Also, if the model will accept this type of information, septic maintenance is very important. If there is a way to incorporate or, if already represented, to improve measures of septic maintenance, failure rates, and/or variability in performance this could potentially have a significant influence on estimates of loading to the Bay from low-density developed areas.

5) What improvements to the methodologies are needed?

See comments below.

<u>General Comments:</u>

The following are a few additional general comments. Most refer to how the data will be used rather than the specific methods used to delineate and project land-use change.

a) HSPF is a lumped parameter model. The pattern of land-use within model segments is not explicitly reflected in simulations of flow and nutrient loadings. Thus, even with detailed land-use data developed here, the influence of low-density land-use on flow and water quality may not be well represented by the model, and hence useful to decision making. As a calibrated model, however, calibration of the Phase 5 model with improved land-use data should capture to some extent the influence of low-density land-use including pattern and extent, albeit in an opaque way. Thus I think the effort to develop this finer detailed land-use data is well justified. In addition, data such as developed here will be critical for improving our scientific understanding of how differences in land-use pattern and extent influence watershed hydrology. There are likely long term benefits from this work that will emerge as this data will enable hydrologic studies and model improvements that will ultimately feedback to improved management decision making in the Bay.

b) Are all “developed pervious” areas treated the same in the model? This may be necessary simplification, but it should be noted that infiltration rates in developed open space (parks, lawns, urban woodlots, etc.) tends to vary along a continuum. In certain areas where disturbance is minimal the degree of perviousness (e.g., as represented by infiltration rates) can be similar to undeveloped areas, whereas in areas where ground has been compacted or built on fill developed open space can have close to zero infiltration and essentially function like impervious cover. In terms of classifying developed lands, you may want to consider adjusting the pervious/impervious percent values to reflect the fact that some seemingly pervious open space actually behaves hydrologically closer to an impervious surface.

c) In the attached written document (Claggett, Irani, and Thompson; Methods for estimating past, present, and future developed land uses the Chesapeake Bay watershed), you refer in many places to the future land-use dataset as a “forecast.” I would argue/urge you to present this data as a scenario, i.e., a plausible future condition, rather than a forecast. The term forecast implies a single most likely future, with probabilities attached. While your methods are reasonable and scientifically sound, many other potential scenarios of future land-use could also be produced that are equally reasonable and sound. Land-use change is a complex process that depends upon a large number of factors we can’t accurately predict decades out. It is important to acknowledge this uncertainty in some way. Ultimately, to protect the Bay the TMDL and WIPs should be robust across a wide range of plausible future land-use change and other factors. This is a philosophical point perhaps beyond the scope of this review but extremely important.

d) Just an interesting aside FYI, a modeling study by EPA ORD in the Monocacy watershed (looking at potential impacts of climate and land-use change) suggests that future increases in developed land would decrease N, P, and sediment loads due to conversion of current agriculture lands to developed. It’s common to think of development as all bad. This presents an interesting twist. The salient point being that the impacts of changes in developed lands can vary depending on what is being replaced. Perhaps common sense, but interesting to note.

REVIEWERS #2-3[2]

As we understand, state and federal laws require that TMDL estimates be made in the Chesapeake Bay watershed. Since 1985, TMDL estimates were made using the CBWM, with the latest version being v.5.3 Results of this are used by states and counties to set TMDL and BMP guidelines. Apparently, this "lumped parameter model" uses several inputs, among them, GIS data that are aggregated to small watershed-political boundary polygons.

Description of inputs required: "To be compatible with the Phase 5.3 CBWM, annual data on the extent of high intensity and low intensity impervious and pervious developed lands, extractive lands, population on sewer, number of septic systems, and population on septic are required for each modeling segment spanning the period years 1984 to 2005 for model calibration and for 2010, 2017, and 2025 for informing state watershed implementation plans. "

We understand that the CBLCD was developed to support this effort, among other things. The classes of the landcover product are:
11 Open Water OW
21 Developed Open Space DOS
22 Low Intensity Developed LID
23 Moderate Intensity Developed MID
24 High Intensity Developed HID
31 Barren BN
32 Unconsolidated Shore US
41 Deciduous Forest DF
42 Evergreen Forest EF
43 Mixed Forest MF
52 Shrub Scrub SS
71 Grassland/Herbaceous GH
81 Pasture/Hay PH
82 Cultivated Crops CC
90 Woody Wetlands WW
95 Emergent Wetlands EW

As we understand the methods, they refined the 2001 NLCD to match the above classes, and they then used the MDA Federal Cross Correlation Analysis, which they don't explain, but has something to do with spectral trajectories and comparing spectral data from one year with those from another year and inferring change. So they took 2001 as the base and did CCA to make the 1992 and 2006 datasets, and then they used the 1992 to make the 1984 dataset. They didn't have the "extractive" class, so they got mining data from the states and made several assumptions to burn "extractive" into the raster datasets. *"Surface mines for states where only point locations were provided were represented spatially as circular polygons equal in size to the reported or modeled active acreage at each site. The extractive polygons throughout the watershed were merged together to form an "extractive" land use mask. Using this dataset, the acreage of extractive land use was tabulated within each modeling segment."*

This is a challenge, but might be the least of all evils. However, when there are "point" datasets, why not simply intersect the point with the appropriate year's data, and if it's sitting in a "barren" or "developed"

[2] Reviewer 2 and 3 provided a single consolidated set of comments.

NLCD polygon (created by vectorizing the NLCD(CBLCD) with appropriate filters and labeling each resulting vector polygon with CBLCD LC class), label it as such? It seems better than using a circle, because, for West Virginia, for example, we have seen that the disturbance polygon refers to a very long and strange-shaped polygon of disturbance. The other thing they could do is do a very quick PI – look at several thousand points and label them as yes/no extractive, or yes/no sitting in a barren polygon AND extractive. See final notes at bottom or review r.e. sensitivity analysis.

"To avoid double counting of land cover pixels, any increase in the difference of extent of extractive land use and barren land cover in a modeling segment was assumed to be due to developed land cover classes beneath the extractive mask and was therefore subtracted from the extent of developed land cover classes in the mask."--I don't quite understand the wording of this. Don't they simply recode the "extractive-developed and extractive-barren" pixels as "extractive" BEFORE the summarization to the reporting polygons?

"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). "

****This seems like a lot – it's hard to believe that's the rule, especially with residential housing, which I think is probably the largest source of land cover change in that area. I suspect bare soil might get mislabeled as impervious surface by CBLCD, so might already include some of the added footprint.

Also, suddenly "bare construction" is listed. This document should list ALL of the classes required by the 5.3 model, not just table 1, which is the cb dataset that was constructed. It would be nice to see all of the classes required by v5.3 in a table, or in a matrix of "cbld vs needed", so as we read the document we know up front what classes are required.

They also created a new set of classes: high intensity developed pervious/impervious, middle... and low… by intersecting the NLCD impervious layer with the 4 developed classes -- this seems reasonable.

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? I'm not sure, but it would be worth looking into for comparing your growth estimates with what Nowak and Stein came up with.

On page 8: *"In place of SLEUTH, the CBLCD was analyzed in each modeling segment to estimate the proportions of forest or farmland converted to development and infill occurring between 1984 and 2006. These proportions of conversion and development are assumed to continue through the year 2025. "*

--This seems unrealistic – could that rate really be consistent for 15 years? Especially with the economy and the decline in new construction? I would probably revise that down. See note at end of review r.e. sensitivity analysis.

"While using 2005 secondary road density to distribute 1990 and 2000 data does introduce some error, 1990 and 2000 road datasets of comparable quality to the 2006 NAVTEQ data do not exist for the region. "

--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.

Basically, we are assuming 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 New Jersey. If not, a more intensive review of the methods might be required.

They used Sewer data in various ways, and tried to find sewer data from localities: *"In areas where data were not provided by a state or local government or from a waste water treatment plant, the CBPO simulated the extent of existing sewer service areas using a thresholded and log-transformed raster dataset of year 2000 population density (produced using similar methods as were used to rasterize the housing unit data). The logarithmic transformation was used to normalize the population density data in the surface raster. The standard deviations in the data range were examined to find the optimal threshold for representing sewer service areas in Maryland. A threshold of 1.5 standard deviations from the mean (> -0.4177) was chosen and used to reclassify the surface raster into a binary grid. A low pass filter (ignoring no data) was then used to smooth the data and the output was converted from a floating point to an integer grid. The resulting integer grid was used to represent potential sewer service areas and serve as a mask for summarizing the original population surface data by county."*

-- We think this was a clever approach, although they might want to factor in some census socio economic data when predicting if sewers are going to go in and/or be used. Many people might have access to sewers but keep their septics because they don't have thousands of dollars to connect. I suspect there's some existing data on areas with sewers, or recent sewers, with which they could calibrate a "likelihood of sewer hookup" vs. socioeconomic data index for each pixel, and use that to weight your averages. This seems like an important point, because if you're drastically over or underestimating sewage amounts, then it might affect models a lot. See note at end of review r.e. sensitivity.

They needed to fix the "low intensity developed" data: *"The CBLCD was used almost exclusively to produce the Phase 5.3.0 land use dataset because it meets the required spatial, temporal, and categorical resolution and consistency requirements of the Phase 5.3 model. Landsat satellite derived land cover maps, however, do not fully represent the extent of low-density residential development, roads, and impervious surfaces associated with rural residential lots. This omission can have significant impacts on both the extent of developed lands and the rate of change in developed lands. "*

Again: *"To avoid double-counting impervious surfaces detected in the satellite imagery with impervious surfaces estimated based on ancillary data and coefficients, high-density residential and dense urban areas must be separated from dispersed developed pixels commonly found in low-density residential and rural areas. "*

This seems odd, again – don't they just recode the areas PRIOR to summing up the per reporting polygon data? This is an interesting approach, one that we in NRS (NE) FIA did to deal with the same issues. We basically calculated a road 1/0 grid from the Dynamap 2000 road data, did a moving window filter that summed up around road pixels in 7 pixel radius circle around each input pixel, and then iteratively parameterized X in the function (if(NLCD = one of the natural classes) and if(road density grid = X), then output pixel = "human impacted", which is the equivalent of your low density residential. See http://treesearch/pubs/14257 Again, see note on sensitivity.

On page 19: this is very confusing, and it's not clear exactly where they're going with it. For example, they appear to be stratifying by urban and rural, and used census “urban areas” as the thing that did the stratifying – this seemed somewhat arbitrary. It seems like it might be better to use block group level housing or pop'n density. With the roads – this in particular is very sensitive to the choice of a heuristic for road width-MANY miles of a 10-m wide strip adds up to a lot of area, and I just think that small systematic biases in this conversion process can lead to very large areas. This would be a great opportunity for a sensitivity analysis using a “Google Earth” PI where you simply zoom randomly into a whole bunch of secondary roads and measure their “true” widths, and compare that to the average value you applied. What do the results imply about choices of the road-based impervious surfaces?

On page 19, why did they do the logarithmic transformation of the focalsum operation? It wasn’t clear to us.

Again, at the top of page 20, there are a lot of assumptions on the amount of impervious surface in these newly generated low density residential areas – they cite a paper, and they do a “random sample” within the area. We believe they should do a more rigorous assessment of this using a PI method. 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.

On page 20 they touch on the above concern about knowing road widths for different types of roads. This should maybe be mentioned up above in the document, not here.

They are trying to calculate acreages of different datasets, including lawns. *“For the purposes of analyzing nutrient loads to the Bay, it is important to distinguish lawns from all other types of vegetation in developed areas because many lawns receive fertilizer applications. In areas identified as developed in the satellite land cover datasets, all pervious lands are assumed to be lawns. For example, we assume that 94% of areas classed as Developed Open Space in the satellite land cover dataset are estimated to be lawns. “*

****Is every lawn equally likely to contribute to the model? Is it true that all pervious high density residential area is lawn? We would argue not. 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.” We understand that there is a parameter in the model that deals with “attenuation” or something that shows how far from the Chesapeake Bay 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. They could do some sort of a simple weighting procedure that incorporates slope, stream density or flow accumulation from a DEM. We recognize that it’s hard to find research on this for large areas, but believe it is a somewhat important point that should be addressed – how do you in some way weight the attributes by likelihood of contributing material to the bay.

Overall comments: We think that given what they are trying to do, they have done an excellent job and used logical and sensible analyses in an attempt to aggregate factors that can contribute to Chesapeake Bay pollution. Our main concern is that there are SO many modeled variables in this dataset that there is a large potential for dramatically inflated or deflated numbers. What I'd like to see is a large sensitivity analysis. You could do this in a few ways:

a) some discussion of the “likely range of reasonable values” for each of the GIS values you summarize for each reporting polygon. For example, where you have a “rule of thumb” value for lot size, you might provide a range of reasonable lot size values and recalculate the GIS layers based on that (perhaps the low bound, the middle bound, and the high bound. You could then provide the v5.3 model with 3 scenarios – the “low” values, the middle (or chosen) values, and the high values. You could do this with most of your other layers.

b) If the above is difficult, you could do something as simple as the “regiongrow” command to “grow” and the “shrink” command to shrink the area of certain classes – e.g., impervious, extractive, road cover, etc. and do several runs of v5.3 with a range of areas of each class above and below those that you supply the model. That way, you can quickly assess the variability of model results in response to changes in the areas of different pollution-causing attributes

c) You could try a PI to get ideas on “overall accuracy” of each of your attributes’ classes of interest. You could then simply factor overall accuracy per factor into some sort of a sensitivity analysis – I think when you combine many layers, each of which has an error, there’s some sort of a multiplicative effect – you do something crazy like multiply all of the accuracies of the inputs together (I’m not sure of this, but I think it’s in the literature) to get an estimate of total accuracy of the combined layer.

We really think that it’s important to have some sort of a sensitivity analysis to give the model output consumers some idea of the confidence in the numbers they are using, because decisions made have a LARGE upstream influence, like in the anecdote Andy recounted on how the township engineer in his Pennsylvania town said how they have an impossible requirement to reduce TMDL because of the results of some “model” which told our county how much they need to reduce TMDL to be in compliance.

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

REVIEWER #4

Per your request, I have provided a review of the technical methods regarding the Chesapeake Bay Land Change Model. I organized my review around the primary questions you sought feedback on, followed by some miscellaneous suggestions.

1. At a regional scale, how effective are **impervious** surface at conveying sediment and nutrients to streams or generating sediment & nutrients to streams?
   a. Is there a need to differentiate "effective" impervious from total impervious?
   b. As land cover becomes more accurate, then what loads generated from these … how sensitive is the model? What are the important, non-linear relationships?

2. Are assumptions and methodology used to estimate and spatially represent **developed** lands sound?
   a. I have seen underestimation of developed lands from NLCD, but it would be helpful to quantify the number of housing units (and/or %) that are not represented in the NLCD development classes – how much underestimation is there?
   b. 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.
   c. 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?
   d. 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.
   e. Page 10: It would be helpful to provide summary tables that show that the density computations are correct and converge to the expected numbers (see sentence below). *"There may be some rounding errors because of the very low densities associated with 30 m cells. Adding up all the cell values within a single block group would produce the exact number of total housing units reported in that block group."*
3. Are assumptions and methodology used to estimate and spatially represent **extractive** lands sound?
   a. It would be useful to have a table that describes how many mines were permitted in each state, along with average and standard deviation of the area occupied by each mine. Does this include gravel mining operations in riparian/stream areas? Are the data from US National Atlas (actually from USGS http://minerals.usgs.gov/) useful to round these data out?

4. Are assumptions and methodology used to **forecast** changes sound?
   a. 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?

b. In the SLEUTH model, what is the basis for exclusion to occur on lands of steep slopes > 21% (why 21%? Is there some empirical study or local regulations that this is based on?)? Why are rurally zoned lands assumed to be off-limits to growth? Zoning typically doesn't preclude development from happening, just typically a different or lower density. Are there special cases that occur in the CBW where this is appropriate?
c. 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). . .

5. What **improvements** to data and methodology could be made?
    a. It seems that the modeling segments is very coarse … if spatial juxtaposition of land uses and proximity to streams is important (the literature suggests it is), then are boundaries derived from 1:500,000 scale streams adequate? It seems like at a minimum USGS 1:100,000 (medium resolution) should be used to link to existing datasets and EPA regulations.
    b. 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.
    c. 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. At minimum, it would be helpful to describe the accuracy assessment from the original paper (Yang et al. 2003) … a better method would be to use an independent aerial photo dataset based on a random sample design. There is a hint of this on page 20: "We estimated the percentage to be around 26% based on a random sample of aerial imagery across the watershed " – more details about the response methods (interpretation of ISA, area of the "sample"), etc. would be useful. Also, since the random points were constrained to fall along roads, these are not simple random sample. That is, the probability of sampling is much higher with more road miles (typically higher impervious surface)… there is some stratification between suburban and rural, but the points that would be selected there would have a higher chance to fall on areas with higher road density and therefore (presumably) higher impervious area. This likely leads to a bias in the estimates of impervious surface area.
    d. It also would be interesting to compare the ISA estimates for the Chesapeake Bay Watershed for those reported in Theobald et al. (2008) that also specifically incorporate exurban and rural housing development.
    e. 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 permited feedlots, chicken houses, etc.
    f. Need to have a representation or at least conceptual model of the sensitivity of TMDL to % impervious… is this a linear or non-linear relationship?

REVIEWER #5

CHARGE QUESTION: How important are low-density residential development and rural and suburban roads to estimating nutrient and sediment loads to the Bay?

This is an important question; however, unfortunately, I do not think it can be answered at this time. The most direct and objective answer would be developed from a sensitivity analysis of the Phase 5 watershed model and a range of plausible land use and land cover scenarios. Unfortunately, staff indicated that such an analysis has not been conducted and, based on their best intuition; it should not make a big difference in model outputs.

CHARGE QUESTION: Are the assumptions and methodology for development the urban and extractive land use datasets scientifically sound?

CHARGE: Are the assumptions and methodology used to forecast changes in developed lands and populations on sewer and septic systems scientifically sound?

Ratio of previous to impervious is likely to underestimate "pervious" areas with "impervious" hydrologic characteristics. The issue is not the appearance of the surface, but the hydrologic characteristics of the subsurface. Stu Schwartz at University of Maryland Baltimore County has demonstrated that compaction during construction means that "pervious" areas, such as lawns, around homes and commercial buildings generate runoff similar to convention impervious surfaces.

These empirical findings undermine the report's simple statement that, "*Previous developed lands are by definition, not impervious."* (page 21). The text indicates that it is important to identify these areas because of high nutrient loads from fertilizer. This is true, but the work of Dr. Schwartz and colleagues suggests that these "previous areas" actually function more like impervious areas, and the combination of high runoff coefficients and high nutrient input could be particularly problematic.

This would seem to be most significant for suburban residential lots (e.g., those with a median size of 0.344 acres and a presumed impervious cover of 26%).

FINDING: Assumptions about the hydrologic characteristics of residential land use and cover may underestimate runoff and pollutant loadings.

Consideration for housing growth in the absence of commercial and institutional land uses is unrealistic. This reflects prevailing practice in some parts of the land use and land cover community, but it is inadequate.

The 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. For example, see http://law.du.edu/images/uploads/rmlui/conferencematerials/2007/Thursday/DrNelsonLunchPresentation/NelsonJAPA2006.pdf or a contrary dialog http://www.calculatedriskblog.com/2009/08/research-on-homeownership-rate-through.html

The CBP's modeling program must be cognizant of these debates and issues and explicitly consider them in development land use and land cover scenarios. As STAC has indicated before, there is no single, deterministic answer and policy needs to be informed by a plausible range of future conditions. That's not the case here.

Assumptions underlying the current, single-scenario, deterministic approach are likely to significantly overestimate the extent of residential land use, while significantly underestimating non-residential land use.

The use of single, deterministic projections permeates the analysis. For example, it underpins the analysis of future sewer system expansion and the number of homes on septic systems. The analysis is entirely dependent on deterministic projections and historic ratios of development.

Assumptions underlying the sewer and septic analysis seem likely to overestimate the number of large-lot single family homes that will require septic systems and underestimate the potential of new technologies for small-scale, distributed waste water treatment to further reduce septic system growth.

CHARGE: Are the assumptions and methodology used to estimate and spatially represent extractive lands scientifically sound?

No comment.

CHARGE: What improvements to the methodologies are needed?

CBP staff are working near the state-of-the-art for the land use and land cover classification and modeling community. They are applying best practices to practical challenges and working around very challenging conceptual and technical issues.

I believe that three of the most fundamental areas for improvement include:

4. Conducting systematic sensitivity analysis to understand connections between regional land use and land cover and the Phase 5 watershed model.
    a. This was a recommendation from a previous STAC review of land use and land cover-related activities and it is reiterated here.

5. Explicitly incorporating uncertainty and multiple plausible scenarios into land use and land cover activities.
    a. This was a recommendation from a previous STAC review of land use and land cover-related activities and it is reiterated here.

6. Incorporate more mechanistic understanding of land market dynamics into land use and land cover modeling. Current efforts are dominated by tools based on spatial and temporal patterns. These approaches are inadequate to develop policy-relevant future scenarios given fundamental changes in regional economics and structural social and demographic trends (e.g., aging populations and changes in consumer preferences).
    a. This will require new types of expertise and input to the CBP, such as from economists familiar with land use and real estate dynamics.

In the future, it would be more productive to engage STAC periodically in the development of these methods. This review was requested very late in the development and application of these methods under very tight deadlines. The timing of the request offered no opportunities for feedback and iteration in these recommendations.

In many cases, the justification for approaches used by the CBP is ultimately that the technique used was the best, practical option available with the time and resources available. These are legitimate issues, but the results are challenging to peer review. As noted above, the CBP has not been responsive to important issues raised in previous reviews, and this suggests a pattern of requesting peer review only to "rubber stamp" nearly completed work products.

REVIEWER #6

<u>Comments on Presentation & Discussion</u>

1. **Impervious Surface Estimates** – The rationale used in 5.3.1 to refine the estimated acres of impervious surface generally appears to be based on a rational method; but **it is difficult to determine if the resulting increase in impervious surfaces (i.e., presumed increased accuracy) necessarily results in an overall increase in the amount of nutrient/sediment loads associated with those additional acres of impervious surface (i.e., increased accuracy in loadings from these land uses)** because the characteristics of the rural/suburban land uses associated with these additional impervious surface acres may not have the same per acre pollutant loads as traditional urban impervious surface acres. It appeared that in several cases **ground-truthing was done to verify these types of acres to other databases. That effort should be expanded to a wider range of sites,** as the current effort seems to assume that if the watershed model changes match Maryland's databases, that that confirms the methodology throughout the entire watershed. **That Bay watershed-wide assumption is not supported in my view.**

   Bottom line, the increased accuracy of estimating this acreage should <u>not</u> automatically be presumed to result in increased accuracy in loading rates and/or reallocation of loads to these newly identified acres. Because there are major management implications and cost impacts that would result if these urban loads were to be doubled, and such potential impacts cannot be minimized. **It is not good enough to assume that corrections can/will be made in 2017, and that such changes in loading assumptions will be 'minor' - and therefore can be ignored until then.**

   It is critical therefore, that that detailed reviews of various segments need to be done to ensure that watershed model estimates that would reallocate loads given these additional impervious acres bear a reasonably close relationship to actual loading/calibration data for those land uses. This should be done not only for various land use types but also in the various jurisdictions to ensure that a one-size-fits-all/assumed loading do not overestimate these loads. **There should also be a concerted effort to provide some sort of accuracy/error-bars around both the resulting impervious acres and the estimated loads. Especially since it is known that the watershed model does not account for differences in pattern or proximity of pervious surfaces to streams – which would affect loads in real world applications. I agree with the general recommendation that some sort of weighting function be considered for these assumptions.**

2. **Urban Impervious and Pervious Load Coefficients** – The tables on slides #20 & 21, show summary coefficients for various land cover classes, and the various coefficients for each states that drives those Bay-wide averages. As I recall, **the states had not vetted these figures yet**, and that the only detailed comparisons was for two Virginia counties (August and Rockingham). It appears that those counties' figures defined Virginia's state-wide numbers, but I don't recall how the other states/DC's figures were developed. And without additional vetting and understanding of how the figures relate to actual land cover; **I'm not sure of the overall state accuracy, or if using 'average' for Bay-wide**

figures makes any sense, or if the differences between the states/DC can rationally be explained. It was also recommended **that comparisons should be made between land cover types versus by state/DC (and to include the number of sites that the coefficients were developed from)** - because that type of comparison may be far more important and instructive regarding coefficients and what figures seem more reasonable. **It was also recommended that data from Montgomery County be evaluated as** apparently they have high resolution percent impervious data that can be used to 'ground-truth' the watershed model data. In any case, **empirical data/evidence should be provided whenever possible to support the various assumptions.**

Estimates of land cover coefficients should be derived to reflect actual soil/land types and practices, not the artifacts of state boundaries – unless those boundaries actually reflect physical differences. Perhaps something other than state figures or Bay-wide averages would be more appropriate to use. It was also recommended that **a sensitivity analysis be done to determine how much of change in those coefficients is statistically significant** (i.e., displaying two decimal places does not automatically reflect greater accuracy); **and that analysis should be done for nitrogen and phosphorus and sediment.**

3. **Efforts to limit assumptions/limit double-counting** – Several techniques were used and some described in the text, but those efforts were not were not always well documented. **All such assumptions and constraints should be documented and quantified; because defining an upper limit to avoid double-counting may be appropriate, but if techniques require that this be done it can also mask a problem with the methodology itself.**

4. **Back-casting Population on Sewer & Septic** – **The graph on slide #33, and the curve to the right of the 2006 data point, has some serious flaws as presented because the curve to might well be bounded by the upper limit/maximum, but it cannot be presumed to predict the future trend and/or rate of change with the timeline that is shown (i.e., to 2030) given current/severe economic constraints; and the fact that vertical growth is really not accounted for in the future densification model** (i.e., New Jersey example). **It was agreed that the CBP's planned workshop this fall could help address this issue and help determine how to better relate demographic projections to increased sewer/septic needs.** This workshop could help to better determine how growth and hence future sewer needs can reasonably be predicted give: a) the constraints of TMDLs and reaching limits of technology – balanced against higher influent sewage concentrations (due to eliminating I/I flows) and those climate change impacts that must be accounted for; and b) the challenges of reuse, and other trade-offs between stormwater and wastewater load allocations in the urban sector.

Comments on Methods Paper

1. **Methodology & Rationale** - Overall the paper provided a good overview of the methodology and rationale. However, as noted in the review of the presentation, **there was not much discussion**

regarding accuracy, sensitivity, or relationship to actual loads generated – critical issues that the paper should address.

2. **Limitations of watershed model and Use to Project Future Land Use /Resultant Loads –** Overall the points made reflect limitations to the watershed model that are not always openly acknowledged, and that are further exacerbated when additional assumptions are used to predict future land use/loads. **I recommended that because the future cannot be predicted with any accuracy, that various 'futures scenarios' be used rather than implying that any one line/curve/set of modeling assumptions can be relied up to predict future loads.** Note: Washington Metropolitan Council of Governments used this approach when developing our own 2050 futures/Region Forward work.