Land Use Back-Cast Methods

Sarah M. McDonald

USGS, Lower Mississippi Gulf Water Science Center

Land Use Work Group Office Hours

July 10, 2025

U.S. Department of the Interior
U.S. Geological Survey

Decision

Approval of the back-cast methods described in the workplan for representing historical land use conditions from 1985-2012 in the Phase 7 suite of models.

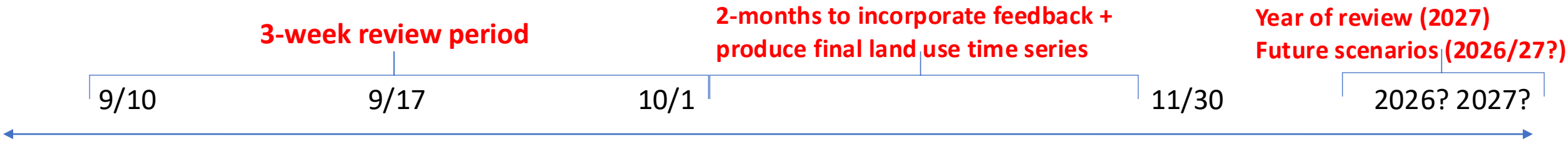
- Since the methods and data are under development, the specific rulesets need to be tested. At this stage, my view on the vote is this:
  - We are voting to approve using annual remotely sensed data to deconstruct the high-resolution land use in place of the Phase 6 method which deconstructed decadal residential development from the census.
- The specific rulesets (e.g. translation from 5-classes to the 16 Phase 7 land use classes) need to be tested. I don't think the group needs to approve the detailed methods, if all are comfortable with the datasets being used (NLCD, ag census, FIA, etc.).
- **Do folks agree with this?**

Do we want to spend a few minutes reviewing the back-cast methods slides from the June LUWG?

Data Review

Office hours for the products, or at least for the back-cast data, should reflect what was listed out in the Methods. There are a variety of pivot tables mentioned that should be addressed in office hours. We may not need office hours for present land use. As for future, when will that need to be incorporated? Is that more for scenario development? Can we wait on that? Back-cast data should have a cursory review for fatal flaws as mentioned in the above bullet prior to delivery to the modelers. Then the remaining review time can be done during the year of review.

- Below is a suggested timeline for data review… let's discuss!



Start Review Period

- Post draft annual land use time series table by county (1985-2040)
- Post draft validation results
- Jurisdictions begin reviewing county-level trends

LUWG Meeting

- Presentation on draft annual land use time series table by state (1985-2040)
- Discussion / questions on data and trends

End Review Period

- Jurisdictions provide feedback on land use trends, including fatal flaws that need to be corrected before incorporation into CAST

Incorporate feedback

Deliver final land use

- Annual land use time series by summary unit
- Future land use (2022-2040) for historic trends scenario only

Develop forecast scenarios with jurisdictions and Year of Review

- Jurisdictions work with land data team to develop alternate land use scenarios
- Model year of review

Deliverables

Can CBP provide county-level summary statistics on annual acres of Phase 7 land use change (important given that counties have a significant role in land use planning and land use decisions)? Could you clarify if working at smaller scales (catchments, HUCs? or Hexes?) watershed-wide for annual land use change, or only for pilot areas for comparing aggregate CB-LULC and NLCD (smaller scales also available in the database to support state and local policy-making decisions)? Would the Chesapeake Bay Program consider publishing its standard 1 sq km GIS hex grid?

- Annual land use (Phase 7 classes) from 1985-2013 by summary unit.
- Providing these summaries by county and catchments is lowing hanging fruit and can be done.
- Given the shortening timeline, the hexagons may not be watershed-wide. If this is something useful for the jurisdictions, I could provide this after the fact.
- The hexagon grid is used for a variety of projects, including our forecasts. We can post these on Chesapeake Data for folks.

Is this method similar to other published methods for doing back-casts?

- There are two primary methods of back-casts:
  1. Change detection algorithms applied to historical imagery (many different algorithms do this)
  2. Modeling some amount of change back through time based on some information (i.e. population change)
- This back-cast uses method 1 – detection of land cover change features from remotely sensed imagery.
- What's unique about this project we are combining the high-res where change doesn't occur with coarser change detection data where change did occur. The classification (Phase 7 land use classes) is designed for water quality modeling, which is also unique.

Segmentation

How will the new segmentation that's being decided by the WQGIT be used for the back-cast? Assuming we're staying consistent and the back-cast will have the same segmentation as the present and future? We will not use the Phase 6 segmentation?

- The segmentation (Land River Segments or LRSEGs) being discussed at the water quality GIT are the units/geographies that will be used in CAST for all dates of land use (e.g. back-cast, present/high-res, forecasts).
- We will not be using Phase 6 segmentation. The only case in which this would happen would be if the WQGIT votes to do so.
- The final land use deliverable will be a table of land use acres by year (1985-2040) per each of these geographies (i.e., county x HUC12 x MS4 x CSO x orographic region).

Validation

Is MD the only state w/ tax parcel data? How does this skew the data validation for other states? Will we be able to see any of the results of the validation?

- To my knowledge, MD is the only state with tax parcel data that is attributed with the most recent “year_built” field. In the cases of redevelopment, year_built represents the new development.
- It limits the fine-scale validation of development trends to MD only.
- SILVIS data (decadal census block groups attributed with housing density) could be useful for high-level trend validation. The primary challenge to this is that SILVIS only represents residential growth. If block groups contain commercial/industrial growth, there is no way to distinguish if the data are “wrong” or tracking a different type of development.
  - Potential solution is masking block groups with tax-parcel scale residential/commercial masks developed for 2021/22. This would restrict the data to residential areas, assuming there is no change in the residential versus commercial footprint over 40 years. This seems unlikely, particularly in densely developed areas and urban centers.
- **If any jurisdictions have data to validate development trends, please share with me ASAP**
- Validation results will be made available with the final product and documented in the Phase 7 documentation. Whatever draft validation results are complete by the September LUWG will be provided at that time.

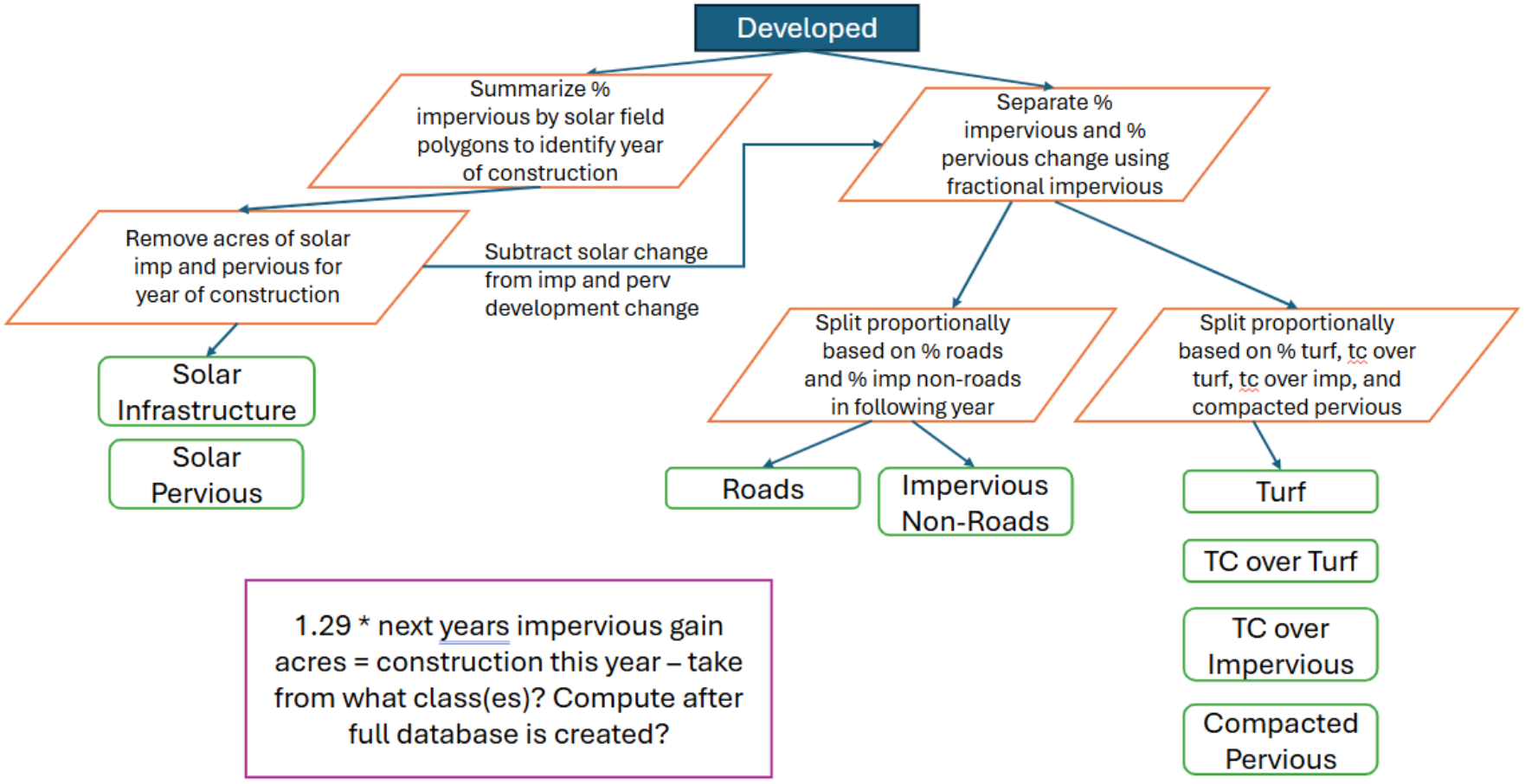
How will commercial/industrial development be identified?

- All development will be driven by the detection of impervious surfaces. The model will not distinguish commercial, industrial, and residential development.
- Generally, we expect to see this footprint shrink in all developed sectors. As we move back through time, the 2013 impervious footprint may:
  - Stay the same (e.g. was developed prior to 1985) OR
  - Transition back to a non-developed state (e.g. farm, forest/wetlands)
- The case of redevelopment may be detected but needs to be tested. Generally, barren lands and impervious lands have similar spectral properties making them difficult to distinguish. Transitions from impervious to barren to impervious may not be detectable.

Does non-development change also include solar?

- For the sake of this project, I am including solar in the developed footprint.
- What is possible:
  - Deconstructing the current (2017/18) solar impervious footprint to its pre-solar state (e.g. farm, forest/wetlands)
- What is not possible:
  - Mapping solar that existed prior to 2017/18 and that is not present in 2017/18.
- Why?
  - We have AI-derived polygons of the solar footprint for 2017/18 and 2021/22. We have annual % impervious layers for 1985-2023. The combination of these products allows us to identify when solar panels were first installed, but not classify solar that wasn’t present in 2017/18.

Can we walk through Figure 7?



- This is an example of what the rules to expand the developed class from the 5-class schema to the Phase 7 land uses.
- These rules have not been tested and are subject to change, but are what I anticipate doing.

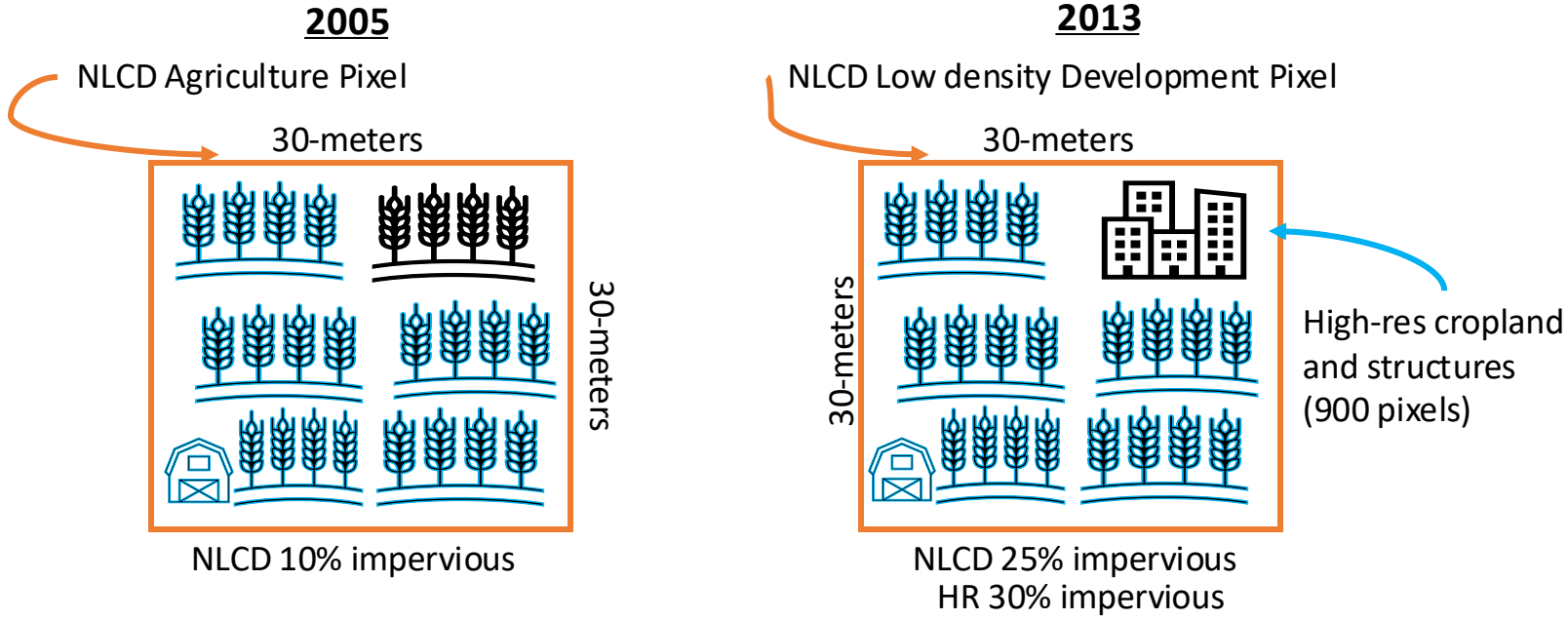
Decision Rules

"In some cases, a land use that is not the majority but meets a certain threshold may take precedence. For example, a rural house on a farm field may be majority agriculture, but could be classified as development as to not lose the impervious surfaces. In cases of a tie, the class will be assigned in a hierarchical order that is to be determined."

We would be interested in the factors CBP would consider when making these decisions, especially if the numbers are sensitive to the approach used. On one hand, as MDP noted in previous comments, reclassifying based on cell majority will likely exclude very low density residential subdivisions (e.g. development on 5-20 acre lots) from the aggregate 2013 landscape. If I recall correctly, a significant amount of impervious surface found in the high-resolution land use data was not found in the 30-meter data.

However, others have since noted that a total reclassification as developed could understate the extent to which the surrounding landscape absorbs runoff associated with these structures, particularly in rural areas. We wondered how CBP is thinking through the impact of these decisions one way or another, and if any sort of validation or tabular interpolation has been considered using a combination of historical Census data and the high-resolution data. I realize we are bumping into some limitations of the available historical data here.

- Context: This step in the workplan is to identify the common 5-class schema between NLCD and the Phase 7 land uses. The 5-classes will be used for NLCD to identify areas of change and their aggregate class.
- The high-res 5-class is only used for cross-walking NLCD, meaning what we mapped in the high-res in 2013 is what is used as our "starting point". NLCD also has majority rulesets, where development is priority over all other classes.



2013 HR is already 70% cropland – in this case, only the 15% reduction in impervious would be reclassed to agriculture

Was the calibration period/back-cast updated between Phase 5 and Phase 6

- The data (acres of land use) were updated.
- The methods did not change, but Phase 6 used the first high-res 2013/14 land use product as the starting point.

Why do we need to review forecasts at this stage? Are they needed in CAST at the same time as the back-cast? Will there be documentation with the forecast data for review?

- The baseline conditions (historic trends scenario) was requested to be delivered to the modeling team by Nov. 30.
- I need to follow-up with the modelers on how and when this data is used in the models. If this product review can be pushed to a later date, I suggest we do so to allow more time and resources on the back-cast.
- The documentation will not be ready in September. It will be ready prior to the year of review.

Will the procedure for the "figure 7" expansion from the 5-class to the Phase 7 land uses be documented in the methods? And can you share any code/scripts created to implement the methods?

- The methods will be detailed and documented in the Phase 7 documentation. Once complete, they can be shared with the group for their awareness.
- The code can be shared upon completion.

Other Questions?

What is a summary unit?

The analysis units used in CAST (Chesapeake Assessment Scenario Tool)

- A combination of layers (has not yet been approved for Phase 7)

- Land-River segments (LRSEGs)
  - Land segments = county boundaries
  - River segments = 12-digit Hydrologic Unit Codes (HUCs)
  - Shoreline = tidal shoreline boundary
  - Climate boundaries = orographic regions
    - Mostly apparent in mountainous areas

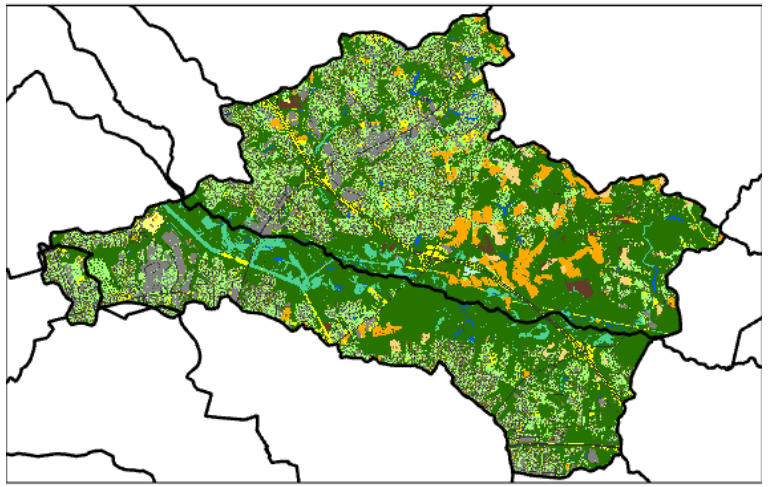
- Summary Units
  - LRSEGs
  - Municipal Separate Storm Sewer System (MS4)
  - Federal Lands
  - Combined Sewer Overflows (CSO)

What Does “Land Use” Mean?

High-resolution LULC
56 classes

Mapped from aerial imagery, Light Detecting and Ranging (LiDAR), and ancillary data sources.

Phase 7 Aggregate Land Use
*16 classes



Reclassifies mapped LULC and summarizes as acres per class per summary unit to form the base land use for CAST.

CAST Land Use
*49 classes (or load sources)

County	LRSEG	Land Use	Acres
ABC	1	Grass	100
ABC	2	CSS Roads	200
ABC	3	Pasture	50
DEF	4	Septic	35
DEF	5	True Forest	500

Incorporates reported data (census of agriculture, state annually-reported forest harvest and construction acres, CSO separations) with the base land use acres. Produces acres of land use by summary unit in which BMPs are applied and loads are calculated.

LULC = Land Use/Land Cover
LRSEG = Land River Segment
CAST = Chesapeake Assessment Scenario Tool
BMP = Best Management Practice
CSO = Combined Sewer Overflows
* Final classification schema for Phase 7 has not yet been approved.

What Does “Land Use” Mean?

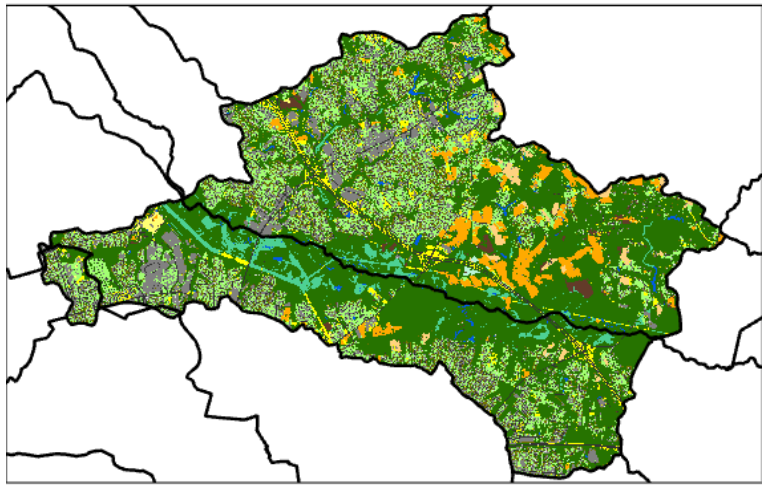
High-resolution LULC

56 classes

Mapped from aerial imagery, Light Detecting and Ranging (LiDAR), and ancillary data sources.

Phase 7 Rollup Land Use

*16 classes



Reclassifies mapped LULC and summarizes as acres per class per summary unit to form the base land use for CAST.

CAST Land Use

49 classes (or load sources)

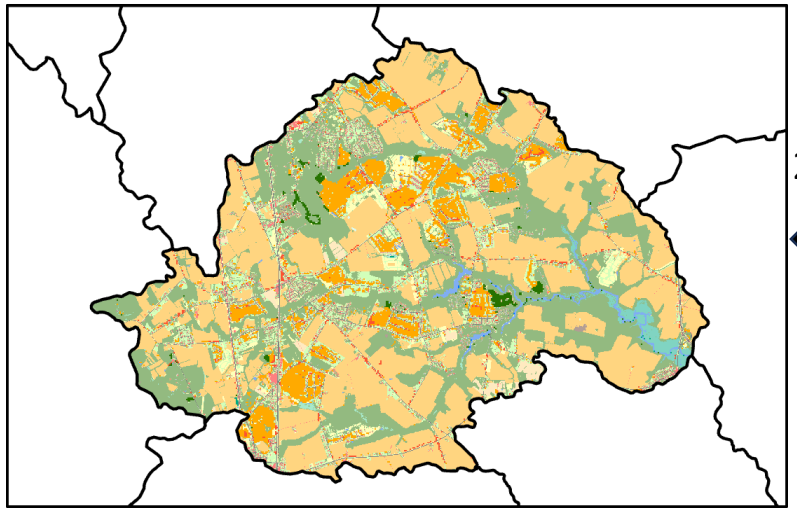
County	LRSEG	Land Use	Acres
ABC	1	Grass	100
ABC	2	CSS Roads	200
ABC	3	Pasture	50
DEF	4	Septic	35
DEF	5	True Forest	500

Incorporates reported data (census of agriculture, state annually-reported forest harvest and construction acres, CSO separations) with the base land use acres. Produces acres of land use by summary unit in which BMPs are applied and loads are calculated.

LULC = Land Use/Land Cover
LRSEG = Land River Segment
CAST = Chesapeake Assessment Scenario Tool
BMP = Best Management Practice
CSO = Combined Sewer Overflows
* Final classification schema for Phase 7 has not yet been approved.

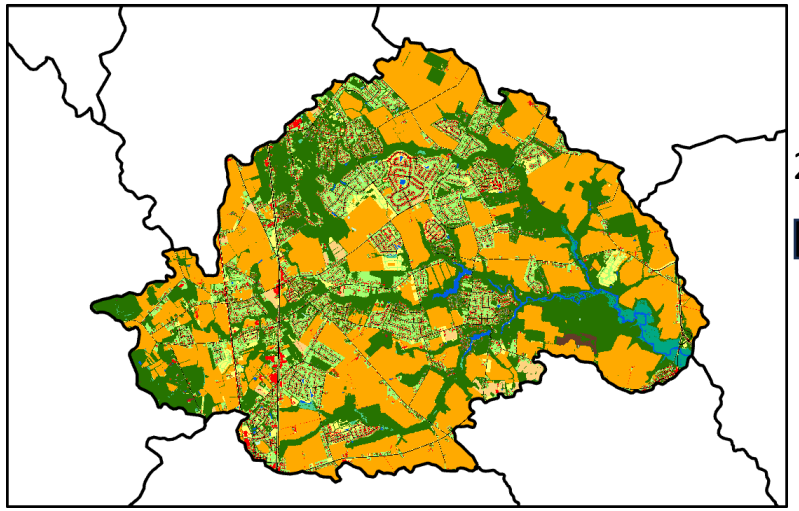
How do we assess land use over time?

Past Land Use (30m)
1985-2012



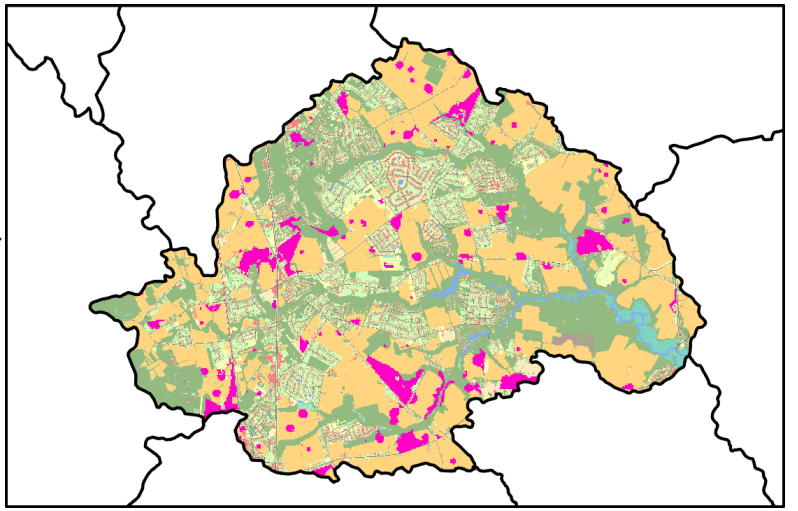
2013

Present Land Use (1m)
2013-2022



2022

Future Land Use (30m)
2023-2100



Annual historic land use condition and trends by summary unit. NLCD detects and classifies change back through time with historical satellite imagery from Landsat at 30-meter resolution. The present is deconstructed where change is detected and summarized by summary unit.

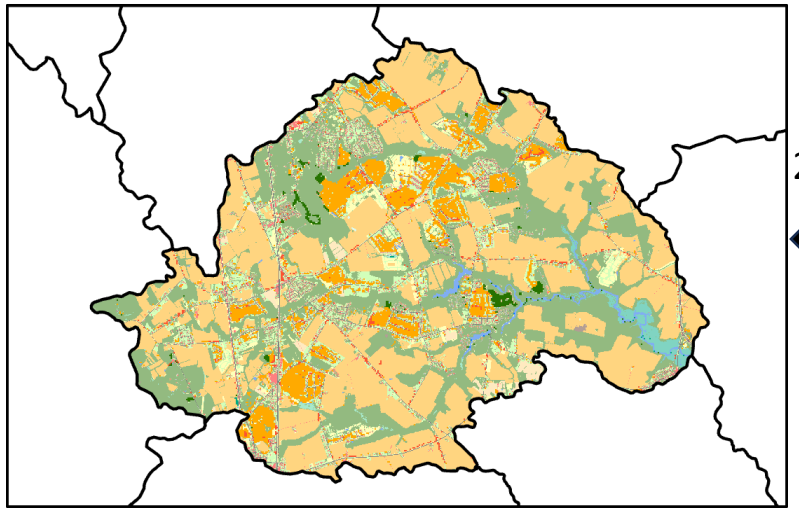
LULC = Land Use/Land Cover
LRSEG = Land River Segment
NLCD = National Land Cover Database

The land use conditions in the present, derived from the LULC at 1-meter resolution and by summary unit. Annualized from the mapped dates 2013/14, 2017/18, and 2021/22. Serves as the starting point for the back-cast and forecasts.

Future land use trends modeled with the Chesapeake Bay Land Change Model (CBLCM). Urban growth model that converts forest and farmland to development to allocate for population growth. Usually predicted in 5- or 10-year increments. Modeled at 30-meter resolution and summarizes land use by summary unit for modeled years.

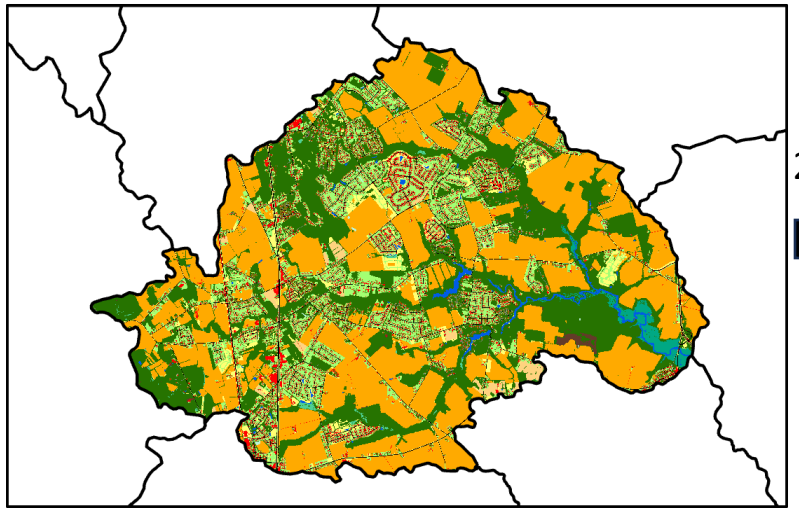
How do we assess land use over time?

Past Land Use (30m)
1985-2012



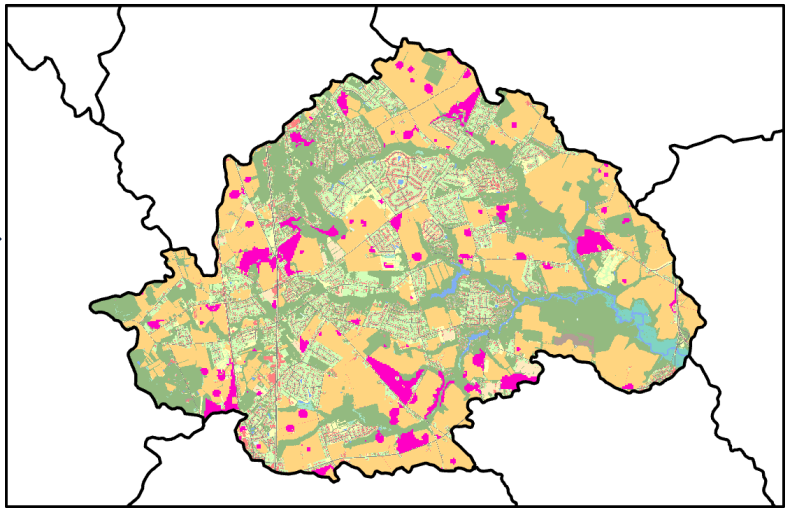
2013

Present Land Use (1m)
2013-2022



2022

Future Land Use (30m)
2023-2100



Annual historic land use condition and trends by summa
and classifies cha
with historical sa
Landsat at 30-me
present is decon
is detected and summarized by summary unit.

LULC = Land Use/Land Cover
LRSEG = Land River Segment
NLCD = National Land Cover Database

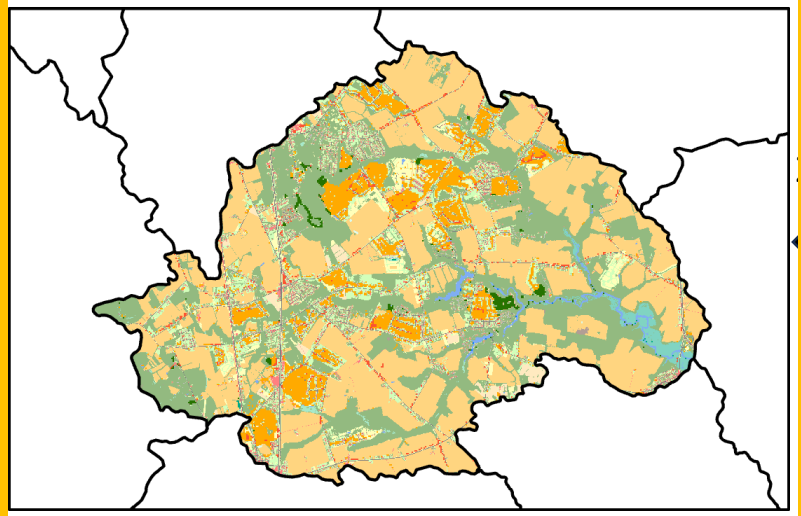
The land use conditions in the
for the back-cast and forecasts.

Three distinct land use products that work together to provide a consistent classification of land use trends for the past, present and future

Future land use trends modeled with
the Chesapeake Bay Land Change
n growth
orest and
ment to allocate
. Usually
year
increments. Modeled at 30-meter resolution and summarizes land use by summary unit for modeled years.

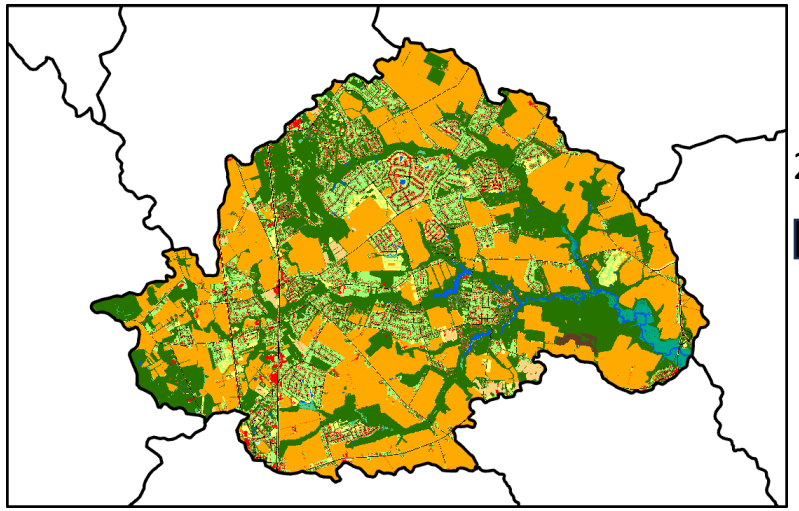
How do we assess land use over time?

Past Land Use
1985-2012



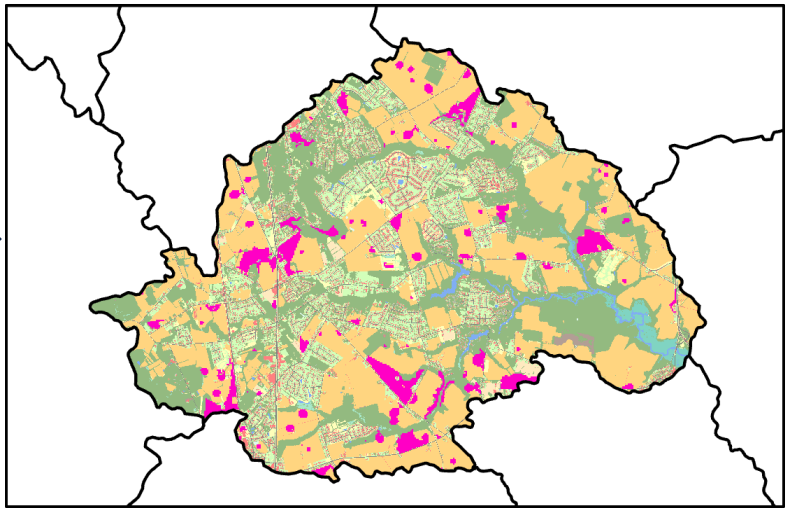
2013

Present Land Use
2013-2022



2021

Future Land Use
To 2025 and beyond



Annual historic land use condition and trends by summary unit. NLCD detects and classifies change back through time with historical satellite imagery from Landsat at 30-meter resolution. The present is deconstructed where change is detected and summarized by summary unit.

The land use conditions in the present, derived from the LULC at 1-meter resolution and by summary unit. Annualized from the mapped dates 2013/14, 2017/18, and 2021/22. Serves as the starting point for the back-cast and forecasts.

Future land use trends modeled with the Chesapeake Bay Land Change Model (CBLCM). Urban growth model that converts forest and farmland to development to allocate for population growth. Usually predicted in 5- or 10-year increments. Modeled at 30-meter resolution and summarizes land use by summary unit for modeled years.

LULC = Land Use/Land Cover
LRSEG = Land River Segment
NLCD = National Land Cover Database

Phase 7 Back-cast Goal

This project will utilize the spatial and categorical accuracy of 1-meter resolution Land Use/Land Cover (LULC) to represent the present and the temporal accuracy of Landsat derived products to deconstruct the landscape back through time.

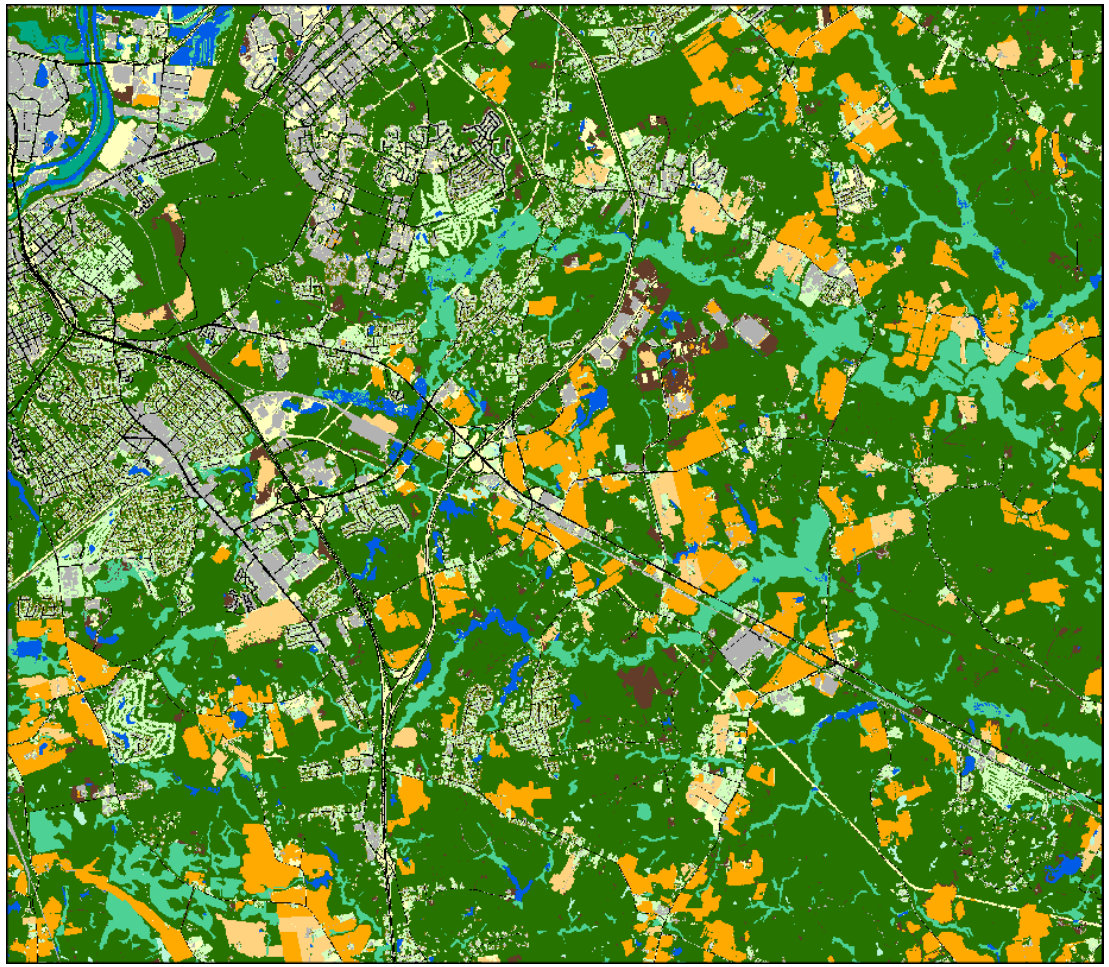
How is the Land Use Back-cast Used?

1. To set appropriate planning targets and Total Maximum Daily Load (TMDL) allocations
    a) Phase III WIPS were based on necessary change in anthropogenic sources of nitrogen, phosphorus, and sediment from 1995. If the partnership follows past precedent, an accurate land use trend will be important for appropriate goal setting.
2. To estimate spatially explicit loading rates from monitored loads.
    a) The historical land use is used as an input to CalCAST, a statistical version of Chesapeake Assessment Scenario Tool (CAST) used to estimate optimal parameters for use in CAST during the development period (Berger et al., 2024), to estimate spatially explicit loading rates from monitored loads. The resulting loading rates are input to CAST to apply Best Management Practices (BMPs) and other information to produce Nitrogen (N), Phosphorus (P), and Sediment (S) loads. An accurate land use status and trend is important for producing a model that best matches observed data.
3. The historical land use is used as inputs to scenarios run in CAST and the Dynamic Watershed Model representing those land use years. The scenarios are evaluated against monitored loads to evaluate model performance and load reduction progress in the TMDL indicator and Monitored and Expected Total Reduction Indicator for the Chesapeake (METRIC).

Phase 7 Back-cast Methods

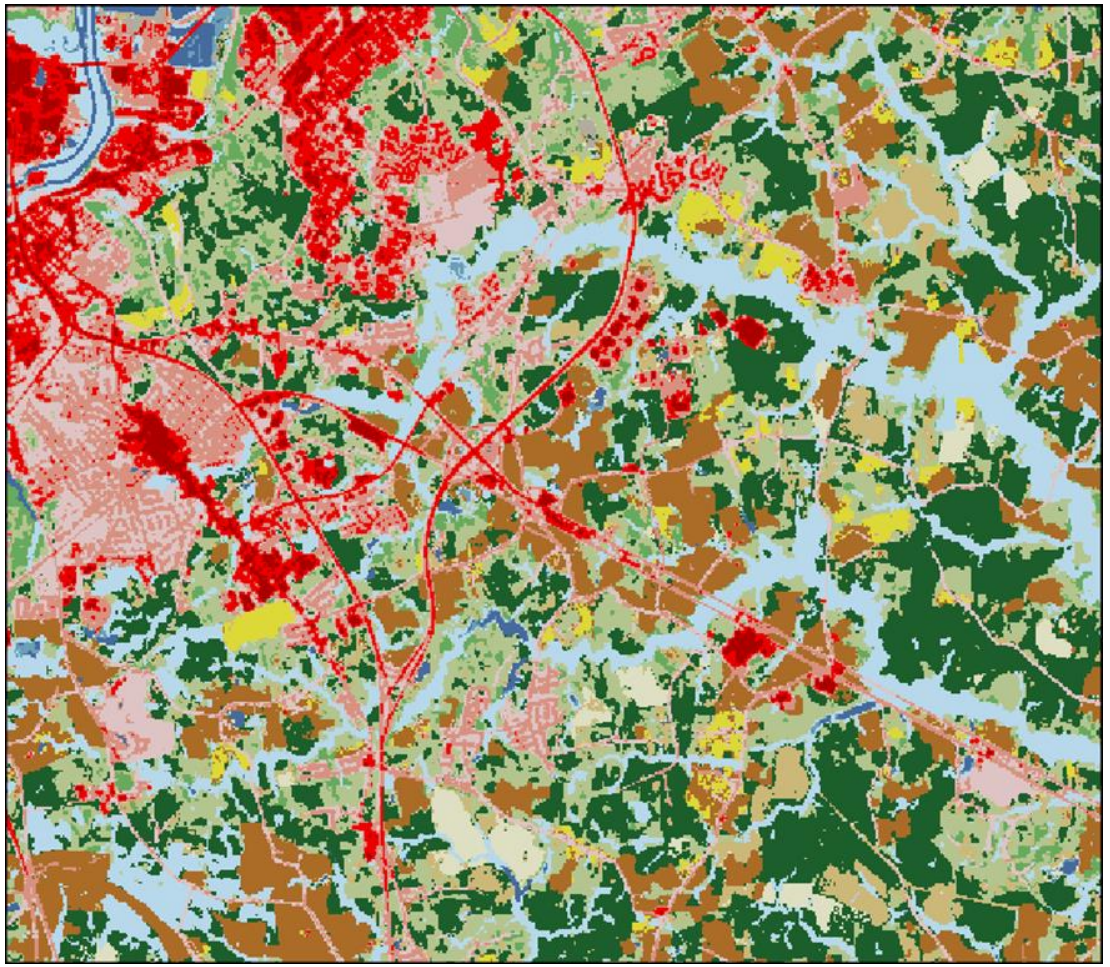
Step 1: Compare 1-meter LULC and 30-meter NLCD

1-meter LULC: 2014



56-classes, displaying 16 phase 7 classes

30-meter NLCD: 2014

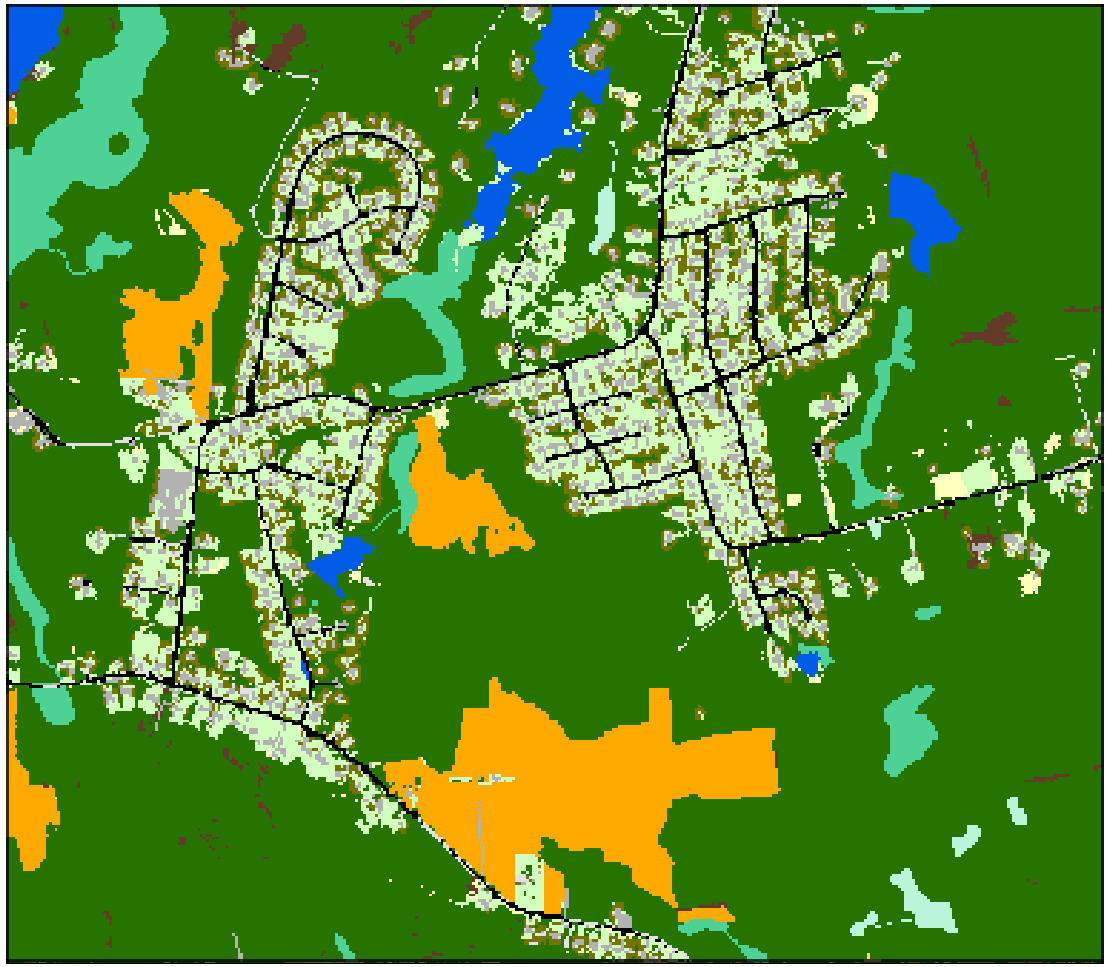


16 classes

Phase 7 Back-cast Methods

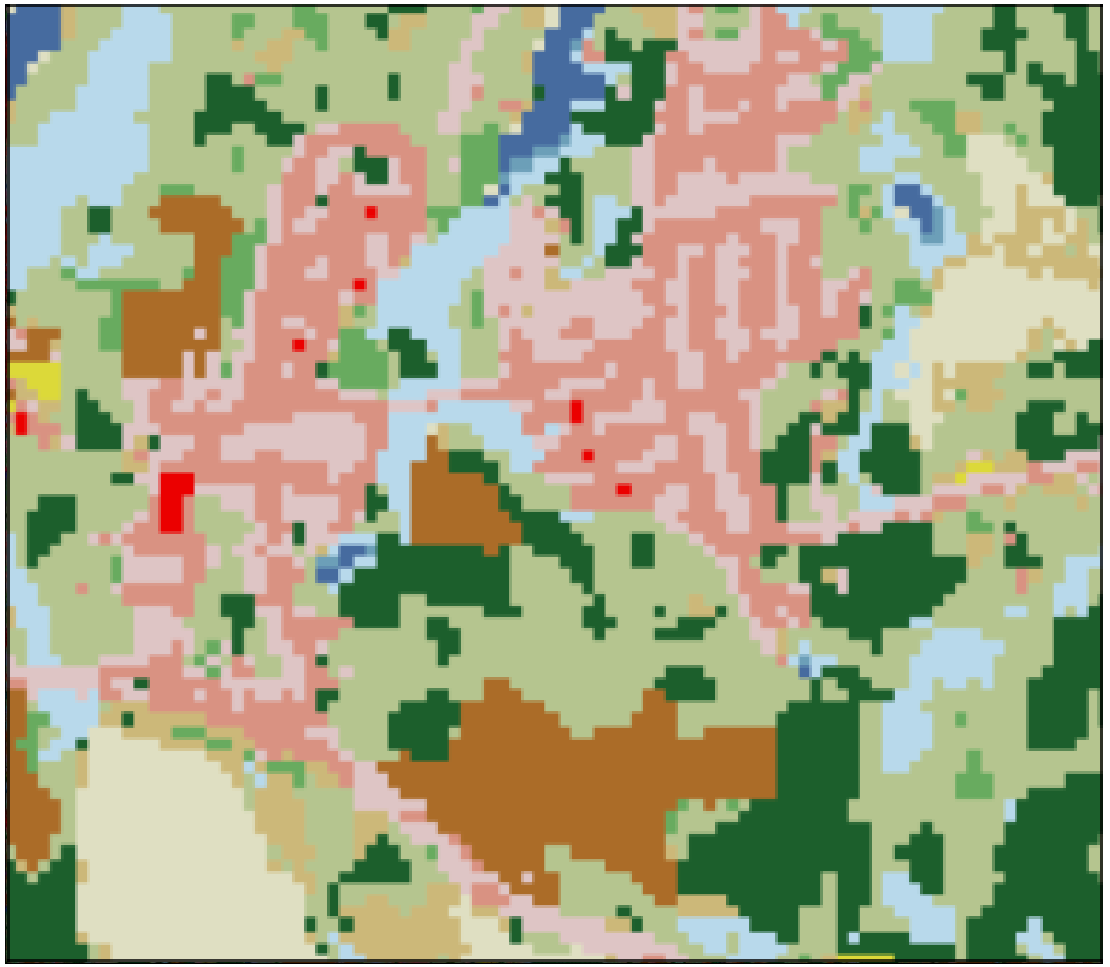
Step 1: Compare 1-meter LULC and 30-meter NLCD

1-meter LULC: 2014



56-classes, displaying 16 phase 7 classes

30-meter NLCD: 2014

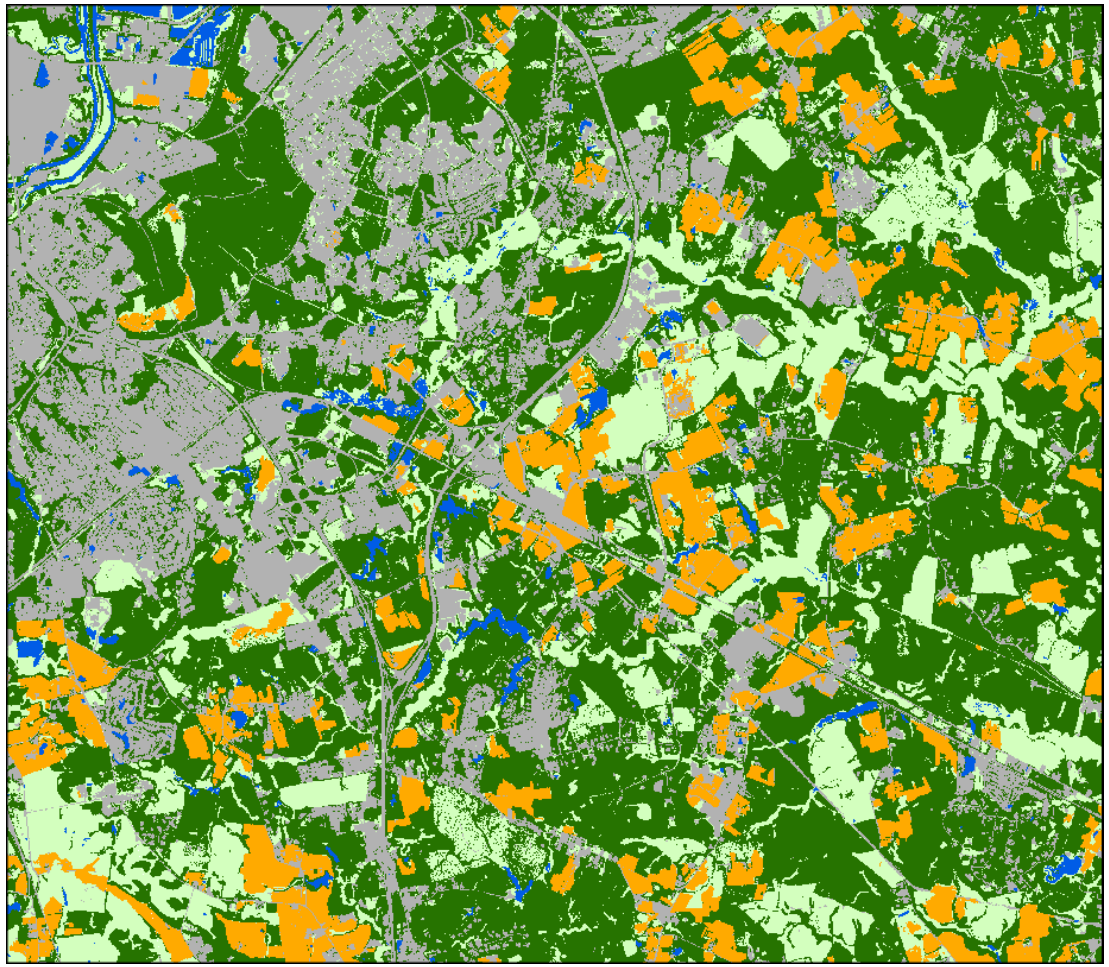


16 classes

Phase 7 Back-cast Methods

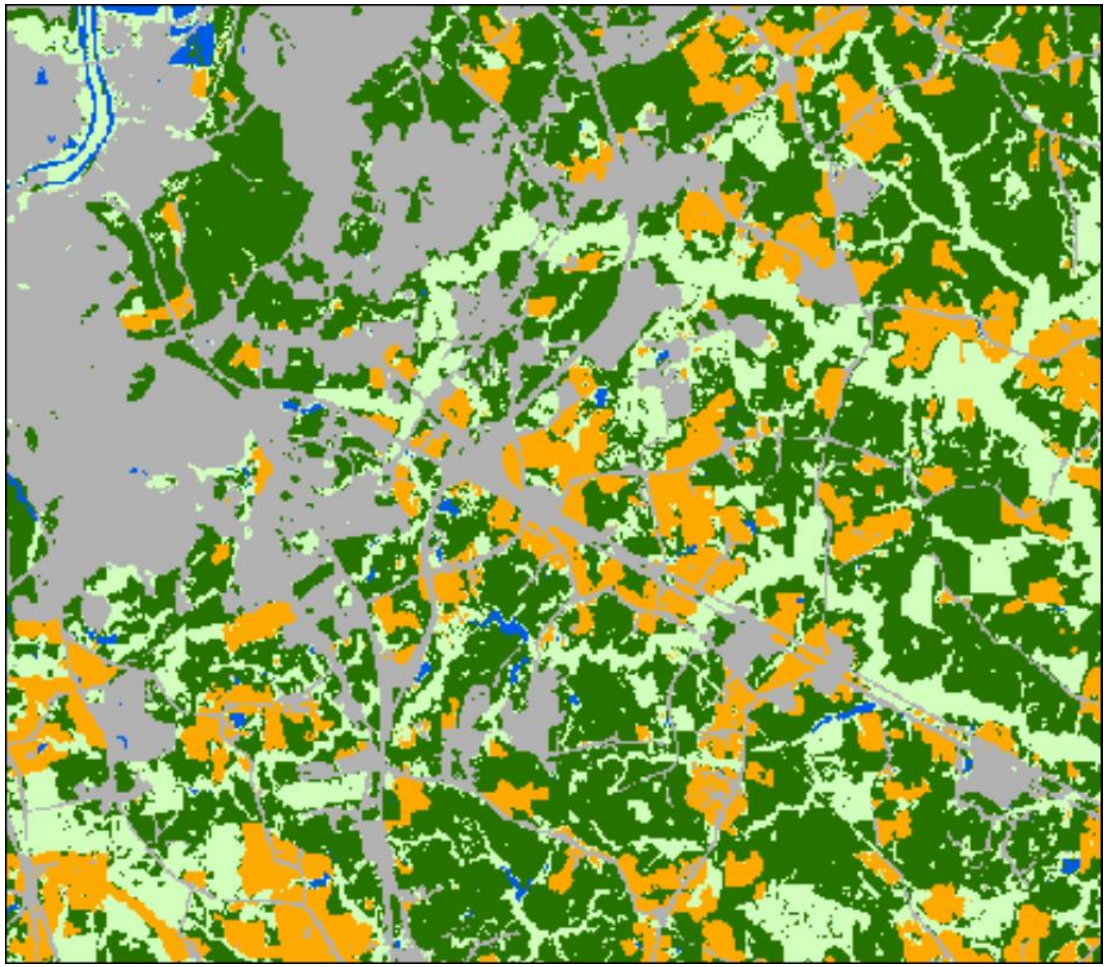
Step 2: Create shared schema for apples-to-apples analysis

1-meter LULC: 2021



56-classes, displaying 5-class schema

30-meter NLCD: 2021

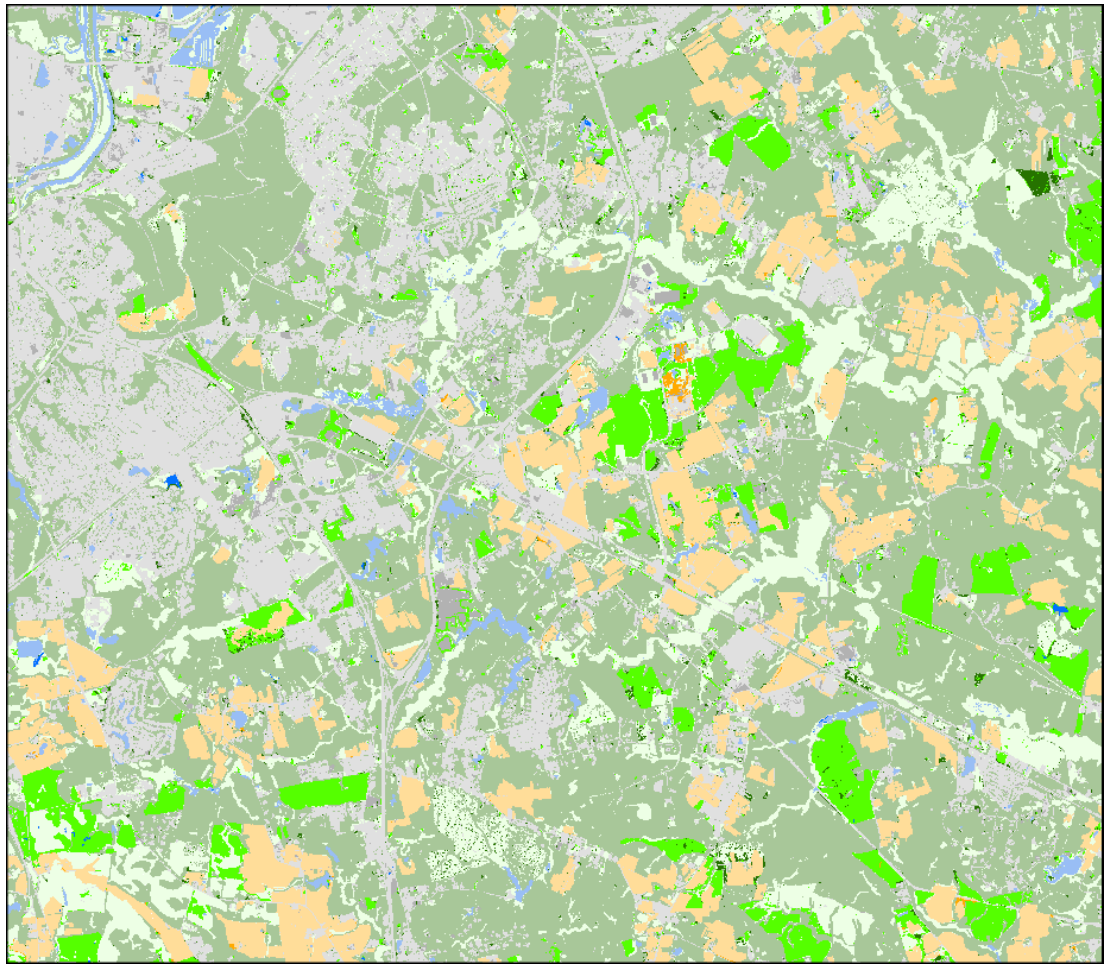


16 classes, displayed 5-class schema

Phase 7 Back-cast Methods

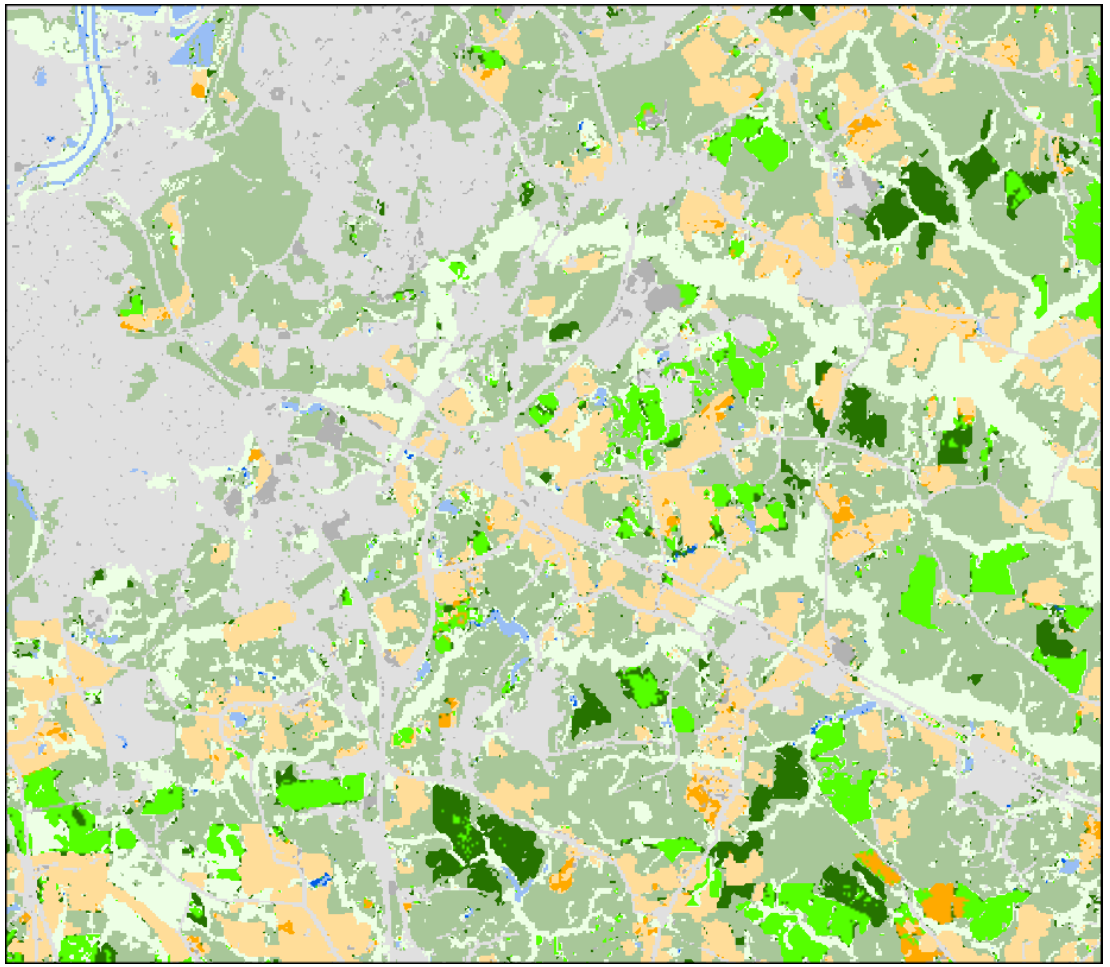
Step 2: Create shared schema for apples-to-apples analysis

1-meter LULC Change: 2014-2021 (2021 shown)



56-classes, displaying 5-class schema

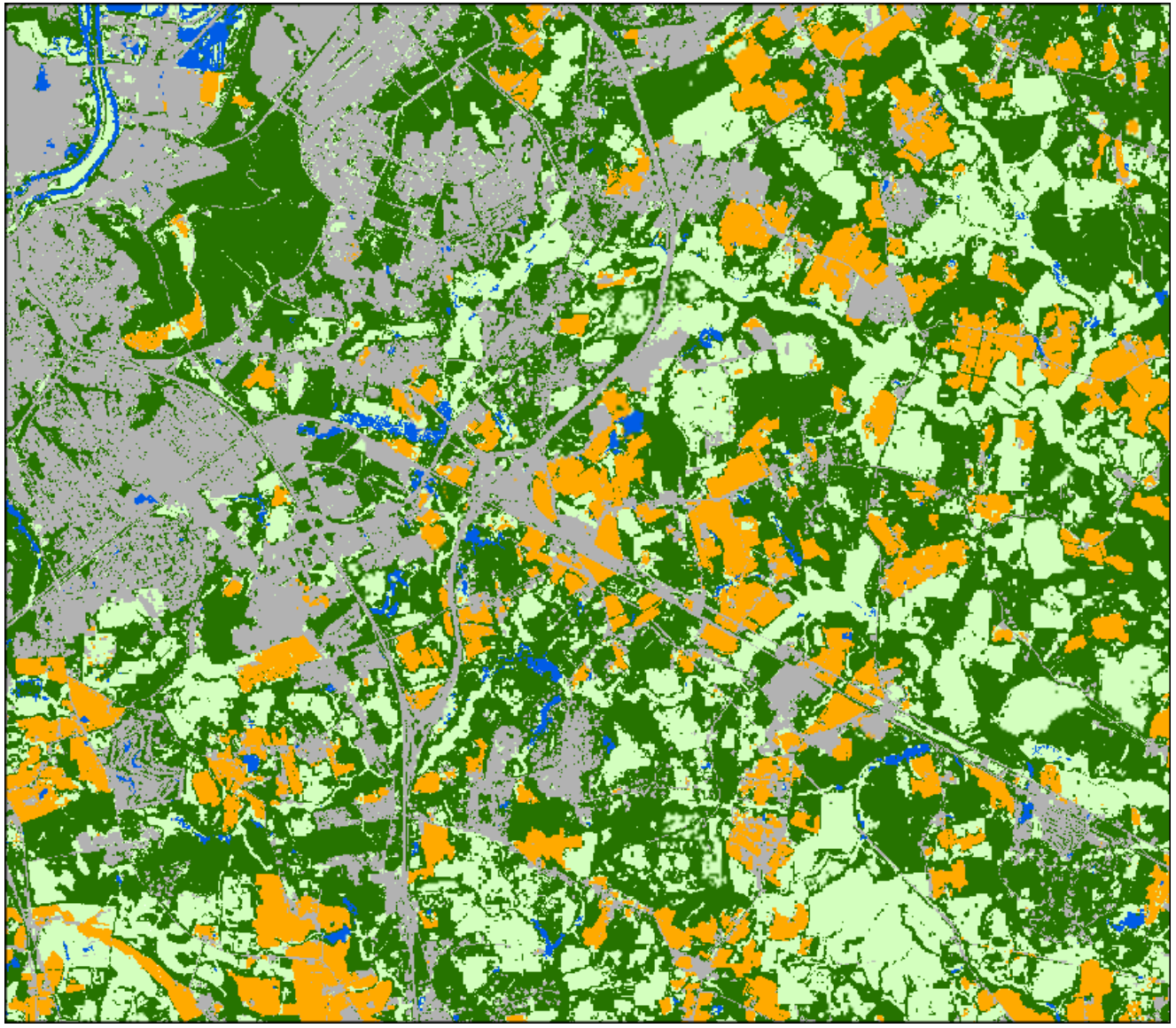
30-meter NLCD: Change 2014-2021 (2021 shown)



16 classes, displayed 5-class schema

Phase 7 Back-cast Methods

Step 3: Deconstruct 2012 LULC with historical NLCD change



Starting from the 1-meter resolution in 2014 to…

2012

2010

2005

2000

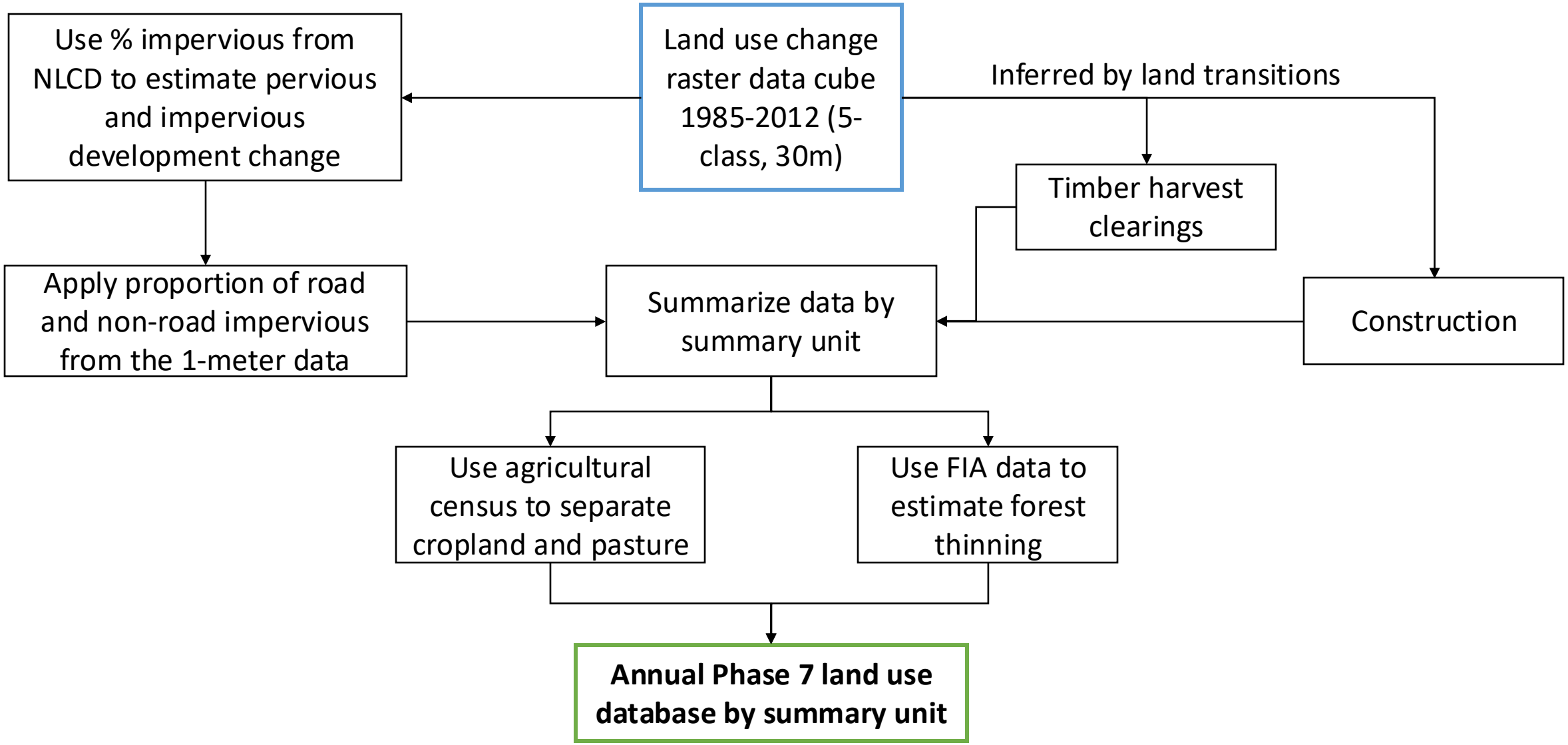
1995

1990

This animation is every 5 years for visualization purposes, the method will do this annually.

Phase 7 Back-cast Methods

Step 4: Expand the 5-classes to the Phase 7 classes



Phase 7 Back-cast Methods

Step 5: Validate trends

1. Development
    a) Compare mapped development change with Maryland and Delaware parcel data that contains “year built” attribute.

2. Agriculture
    a) Census of agriculture to validate trends in the agricultural footprint

3. Forestry
    a) Forest Inventory and Analysis (FIA) data to validate forestry trends

How does the Phase 7 Process Differ with Phase 6 Methods?

1. Temporal resolution is finer in the Phase 7 method
    a) Phase 6 method deconstructed residential development to be farm or forest using the Decennial Census (every 10 years) and the Chesapeake Bay Land Cover Data Series (CBLCD) (every 4-6 years). Phase 7 method utilizes annual data that did not exist during Phase 6 development.

2. Spatial Resolution is finer in the Phase 7 method
    a) Although the change detected from CBLCD and NLCD are 30-meter resolution products, the Phase 6 method aggregated the 30-meter products to the summary unit to identify the proportions of natural and agricultural lands to convert to development from the census. The Phase 7 method maps the land use change transitions explicitly at 30-meter resolution.

What are the potential effects of these changes?

- More temporal granularity (annual Landsat derivatives versus 4-6 year timesteps)
- More development change than Phase 6 because it is not restricted to change in housing units
- More change overall (e.g. timber harvest, ag expansion and contraction)

- The effects on loads have not yet been quantified

Approval of Back-cast Methods

- The back-cast methods need formal approval by LUWG.

- To ensure we meet the September 30th, 2025 deadline, we should plan to make this decision prior to the September LUWG meeting.

- What is the group preference for making this decision?
  - Decision request via email
  - Add a shortened meeting in July or August

Questions?

Sarah McDonald
smcdonald@chesapeakebay.net