Agricultural Modeling Team (AMT)
Meeting Minutes

July 14th, 2023
09:00 AM – 11:00 AM
Meeting Materials

Summary of Actions and Decisions

Decision: The AMT approved the June minutes.

Action: Identify information gaps for crop yields and acreage for counties across the watershed. Explore spatial interpolation options for filling those data gaps.

Meeting Minutes

Statement of purpose:
To evaluate the crop yield and land uses in CAST and discuss potential alternatives for Phase 7.

Announcements:

- Fertilizer Expert Group Phase 7 recommendations are working through the partnership:
  - Conduct in depth searches for different sales and application data sets.
  - Evaluate differences in data processing.

Introduction: 09:00-09:05 [5 min (Zach Easton, Virginia Tech)]

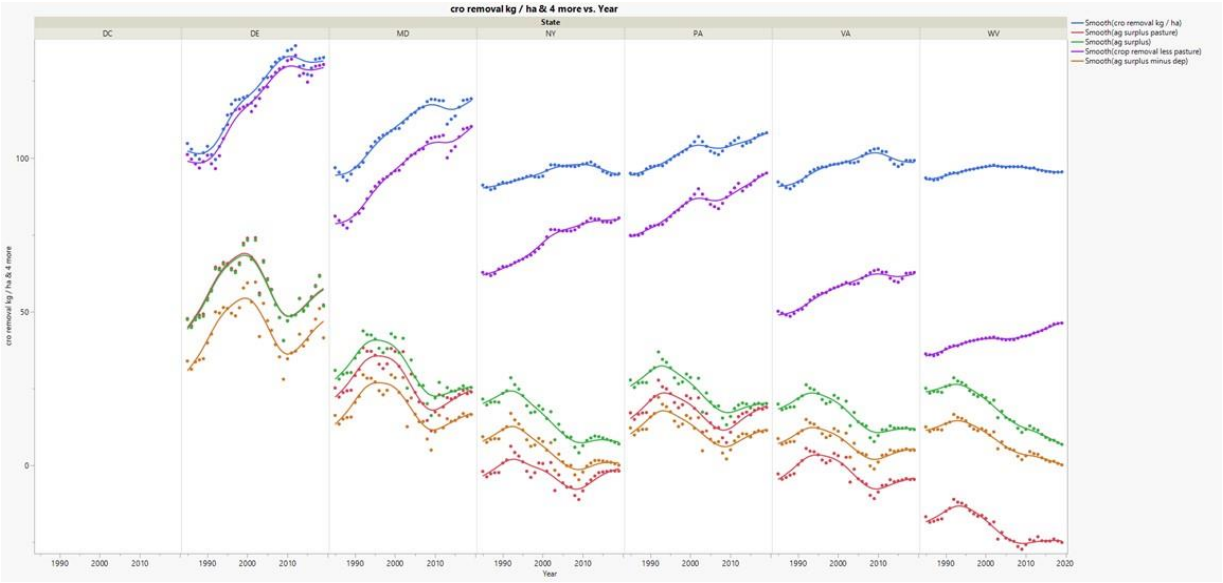
Approval of June minutes and take attendance.

Scenario Discussions: Crop Yield trends 09:05-10:55 [110 min (15 min presentation 95 min discussion) (Tom Butler, EPA; Zach Easton, Virginia Tech)]

At the June meeting there was discussion about utilizing spatial interpolation to improve the methods used to fill in crop yield data gaps at the county scale. During the past month the CBP has looked at ways to involve technical experts to achieve this improvement. Tom outlined the progress to date and sought input from the group as to how this should be implemented in CAST.

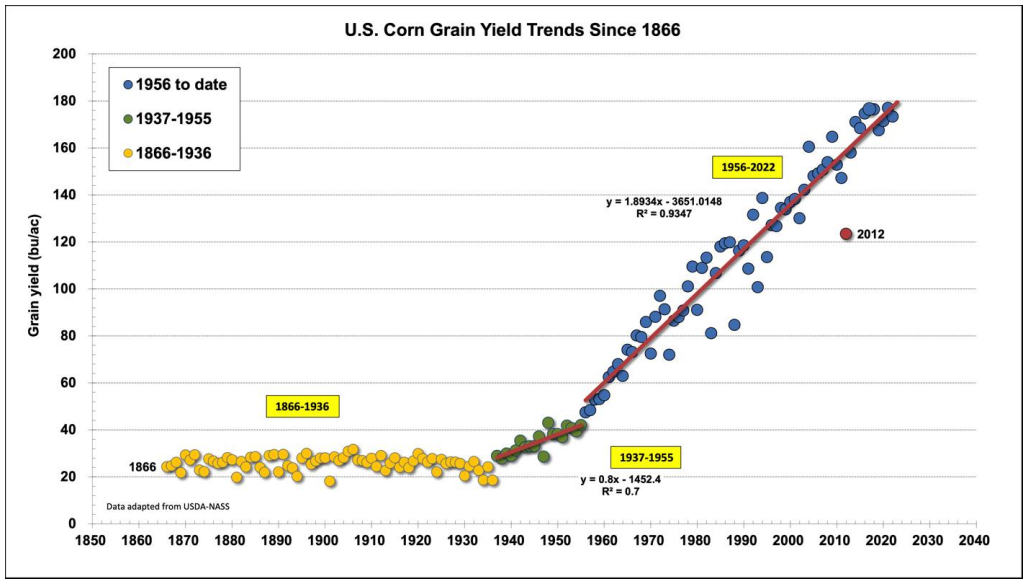
Discussion
Ken Staver: So the organic nutrients are in a county where the nutrients are, but the inorganic get distributed across counties from the whole bay bucket, right? Inorganic gets moved around more?
Tom Butler: Yes. There are still things that could influence organic though.
Ken Staver: Someone has to report organic for it to change counties though.
Tom Butler: Right.
Eric Rosenbaum (in chat): If yield increases by 2% annually, does the model increase fertilizer use by 2% or a lower increase to account for 4R nutrient stewardship practices + genetic improvements?

Tom Butler: We are not necessarily doing the 1:1. We are meeting nutrient requirements with the organic portion in the county. The rest is filled with the inorganic source. Inorganic value is fixed from the last year of data we have moving forward.
Olivia Devereux: Nutrient management is also a BMP. Number of acres are reported by states.
Robert D Sabo: How is excess fertilizer applied in the model?
Tom Butler: We only have fertilizer stock for the last year available, in this case 2016. We don't have a stock for later years that we have excess from.
Robert D Sabo: Once AAPFCO has those numbers we can use those?
Tom Butler: It's up for discussion if we want to continue using solely AAPFCO data.
Dave Montali: For 2016, I thought our stock of inorganic fertilizer was less than the amount to meet all the agronomic need we have projected. Is that right?
Olivia Devereux: Depends on where you are in the watershed and which crop you're talking about.
Jess Rigelman: In general, I'm not sure. Through the earlier years it hadn't exceeded the overall crop need, but I'm not sure about recent years. That is just overall crop need though, and timing of application also plays a part.
Mark Dubin (in chat): Supplemental NM for N and P are reportable BMPs for representing advanced NM practices. It is not assumed but reported by the partnership to the CBPO by acre by year.
Tim Larson (in chat): Farmers don't tend to stockpile excess non-organic fertilizer. They apply it as they purchase it, so the sales numbers are a useful data source!
Robert D. Sabo, (in chat): yes. @Dave



Results from Sabo et al, Chesapeake bay nutrient inventory
Chris Brosch (in chat): what is the source for the data in the graph?
Robert D. Sabo (in chat): CAST. https://iopscience.iop.org/article/10.1088/2515-7620/ac5db6. Olivia and her team also got this nutrient inventory downscaled to NHD+. https://www.usgs.gov/data/cast-data-input-disaggregation-county-and-land-river-segment-scale-national-hydrography
Chris Brosch: Don't think we can look at the pre-BMP/non-BMP conditions to answer these questions. Because the difference between expected and max yield is one element of NM. The other half is the appropriate application rate or appropriate yield.

Tom Butler: What we're trying to get at is instead of using a max yield, what is a good way to get towards a reasonable expectation for crops? Are you saying that we shouldn't be looking at a pre-BMP expected yield?
Chris Brosch: I think it's impossible to tell what a non-BMP condition looks like for management.
Tom Butler: So maybe we want to look at something other than expected yield or put a qualifier on that.
Robert D Sabo: Do you mean the N/P removed from the watershed or taken up from the plants? Or do you mean both?
Tom Butler: When talking about yield, we're referring to the mass that is produced.
Jess Rigelman: In the model we use uptake, so whole plants. Removal would be ⅔ of that.
Ken Staver: To be clear, we're dealing with reported crop yield and bushels per acre like Tom said that you get from NASS. Also, in counties that don't have yield data, do we also not have acres for those crops?
Olivia Devereux: No, NASS has acres and yields in separate tables. So we may have data for one but not the other.
Ken Staver: When I do quickstats, for the counties that don't report, you don't get either.
Jess Rigelman: For the annual NASS data, that is true. But for the 5-year ag census data we get yield and acres separately.
Alex Soroka (in chat): I thought the counties reporting far outweighed the counties which didn't report.
Robert D. Sabo (in chat): Uptake and yield is an important distinction. With yield (e.g., bushels/acres), you can apply nitrogen coefficient to get N removal directly for a mass balance. However, as I understand it, the Bay model uses the yield data to determine an uptake value. If a user wants to get crop removal, they have to apply the 2/3rds rule for all crop types.
Mark Dubin: In response to Chris's comment, the phase 6 NM expert panel looked at one study done comparing pre-NM and NM applications and found that a portion of producers that were pre-NM that were on target, some that were exceeding, but also some producers that were applying below target. So it was all across the board. There was an improvement after applying NM, but in some cases nutrients applied increased. But it was an old report so the baseline today might be different.
Dave Montali: The baseline producer would apply nutrients to the max yield whereas to temper expectations under NM would aim more towards estimated yield that is lower than max. Do we know that? I guess Mark is saying that's not universal.
Chris Brosch: Yeah, we don't know that. We could use max yield as a proxy but not sure that's appropriate. We tried that in a previous version of the model.
Mark Dubin: A lot of P6 NM panel work was on the availability of information to make decisions. If you don't know what your yields are, don't know what your nutrient value or organic component is - you're just guessing. The value of the NM was providing information to make more informed decisions. Part of that is looking at long term yields and expected yield like what Dave is saying, but it's bigger than that.
Chris Brosch: What's important to me is that we need to have identified fertilizer inputs - where that data comes from will be discussed. But that is the source of nutrients, combined with animals and biosolids. You're asking how we identify the size of the sink and equitably distribute the source, and how we do so in a way to add load source BMPs. In my mind, the relationship between max and expected yield is irrelevant. Those should be traveling together. Doesn't matter what the number is, as long as the relationship through time is reflective of reality, I'm fine with it. Might be safer to chase max yield to secure a method to apply NM, understanding that the baseline as a non-bmp condition is impossible to estimate.
Robert D. Sabo (in chat): Corn yield trends over time.



Tom Butler: If we don't find an improved way to do this, our default is what we currently do in Phase 6. So it's always possible to stay with what we have. Do other states have input on using the max yield?

Alex Soroka: How would we define maximum yield? Where do we get that kind of data?

Robert D. Sabo: Could ask seed distributors what their max yield is based on ideal conditions. I think we need to consider the impact of climate as well.

Chris Brosch (in chat): We have to be careful about theoretical and realistic max yield. Theoretical is not useful for a massive averaging model for interpreting the distribution of nutrients.

Mark Dubin: Maybe we should take a look at max yield for Phase 7 and use a study that is more current to compare to the previous value we have or to the theoretical.

Ken Staver: Are we still talking about unreported areas?

Tom Butler: Want to get the best max yield for every area. Initially, yes, we'd fill those gaps but if there's a way to have a max yield for everywhere it would be ideal.

Ken Staver: This is way too much effort for areas that we don't have reported data from NASS. Need a map of the watershed that has green for counties with reported data and red for counties that don't have reported data, because if the area is small then it seems like too much effort to spend time on this. And even if we have max yield, we still need the acres.

Eric Rosenbaum (in chat): Expected yield is what farmers will fertilize for. Farmers, in general, are not fertilizing for max yields.

Chris Brosch (in chat): A realistic max yield would be captured by the high points in a tool like NASS and interpolating between those good years. Expected yield is roughly 80% of that by design when we manage for best 5 of 7 years.

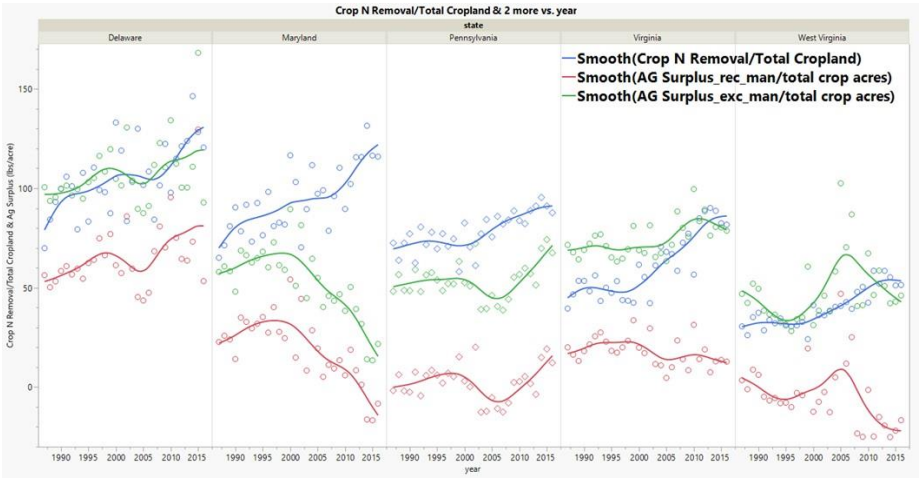
Alex Soroka (in chat): Use CDL for counties which don't report? Cropland Data Layer. That get's you acreage

Olivia Devereux (in chat): In a prior version of the model, max yield was the optimal yield in the Penn State Agronomy Guide and the similar guides from each of the other states. Actual yield is what was reported in the Ag Census/NASS annual surveys.

Robert D. Sabo: Fertilizer institute - sometimes you have censored data with NASS at the county level. You can proportionally allocate the corn yield for counties without data. I'll follow up over email about the methodology. Also, to get the expected yield - is it possible to use the reported

yields in the given year instead of the expected yield in 2001 or 2016? And use the reported data to track progress?
Tom Butler: That's kind of what we do now. There's some manipulation that happens to try and fill nulls.
Robert D. Sabo: USDA has state-level expected yield forecasts that you could use.
Tamie Veith USDA-ARS (in chat): I am not aware of access to the data source of forecasted yields mentioned.
Chris Brosch: I agree with Ken that filling in data gaps is most important, need to start there. Geospatial tools might be able to help us. The acres we have other sources to fill in that gap but the yield shouldn't be vastly different from the surrounding areas. So approaches should be different.
Robert D. Sabo: So you're saying since we already have the reported yield data, we don't have to engage these other datasets? Soil, irrigation, climate?
Chris Brosch: Yes. Issues like irrigation, genetic technology have a more recent bias on the data and a model that begins in 1985 and forecasts in 2035, those elements are minimalist in how they influence the data. We probably shouldn't capture them in our effort to model the baseline.
Tim Larson (in chat): The NASS Cropland Data Layer can provide good yearly crop acreages derived from satellite imagery. Yield could be estimated from this.
Tom Butler: How does that differ from the NASS yield data?
Alex Soroka: Cropland data layer is from a team at the USDA who is looking at reflectance data. Look at characteristic land surface using remote sensing. You get annual raster datasets, so a tabular map that tells you what is happening in a 30x30 cell. They are quite accurate for individual crop categories, but still working out the kinks when it comes to multiple crop rotations.
Olivia Devereux: We've checked with NASS and they did reply that they felt it was better we used the tabular data for attributing land use to various crops. They felt it was stronger than the cropland data layer.
Alex Soroka: I agree, would only suggest it to fill in data for counties that we don't have info for.
Mark Dubin: Looking at the cropland dataset, we've seen areas in the watershed where the data didn't hold up.
Alex Soroka (in chat): I work mostly in MD and DE, and there we've had success.
Robert D. Sabo (in chat): Here is an example of USDA weekly report for yield forecast https://www.nass.usda.gov/Publications/Todays_Reports/reports/crop0822.pdf
Robert D Sabo: There is also USGS National Resource Inventory (NRI). You can get that at county scale or cropland reporting district scale.
Tom Butler: There aren't privacy concerns with that dataset?
Robert D Sabo: The county level might be censored but regional scale would be available. You could calculate the county level though. And goes back to 1980.
Mark Dubin: We've looked at that data in the past but didn't get very far. It's through NRCS so we had trouble accessing it.
Robert D Sabo: I can reach out with a request if it's helpful.
Mark Dubin: Yes. Peter should come speak to the group about the high-resolution land use because we've been using that and comparing it to the CDL to see differences.
Olivia Devereux: Sure, let's figure out how the land use is done. But what Tom was saying is that we're trying to figure out what the nutrient applications are. We're getting off topic about data sources. We need to figure out what we're going to use to disperse nutrients to crop types in the model.
Mark Dubin: Not only is the yield important but also how many acres we're applying it to, so it's all related.
Alex Soroka: Going back to Ken's question - how much data are we missing?

Tom Butler: I think around 2%.
Chris Brosch: I would ask if this were the real number or specifically some land area?
Tom Butler: Thanks Chris for his input, this is the acreage proportion of specialty crops which happen to make up a number of crops with missing data.
Ken Staver: We only have a few counties that we don't have NASS reported data. Not every parcel reports to NASS. But when we're getting it from NASS, they've already dealt with that problem. I thought we were talking about the unreported acres in NASS?
Chris Brosch: Filling in those missing data pieces for acreage, I think we've tackled that. Filling in the data gaps for yield is much more important because the ag census inadvertently captured too many drought years. And for what it's worth I think a spatial analysis is the best way to fill that gap.
Ken Staver: Sounds reasonable to me.
Alex Soroka: Use surrounding counties for yield estimate?
Chris Brosch: Not sure, but it makes more sense to look at neighboring counties for that number than to look at the whole state and take an average or subdivide.
Robert D. Sabo (in chat): How we are allocating fertilizer is important as certain states have differing trends in surplus nutrients being left on fields if you're using mass balance data not from CAST. Axis is in lbs/acre.



Tim Larson (in chat): Yes you could interpolate spatially. You would want to interpolate by soil type from the soil survey.
Dave Montali: I think we have growth regions. Couldn't we determine the average and apply that to counties where we don't have yield data?
Chris Brosch: That could be useful when considering the surrounding counties that has missing data. But also those growth regions don't cross state lines, which I think would be a mistake to not use surrounding counties that are in a different state for these estimates.
Mark Dubin: There are adjacent counties that have completely different soil associations and therefore completely different yields.
Tim Larson: You can interpolate looking at similar soil types and regimes like that.
Robert D Sabo: Are there any proof of concepts for some of these interpolation techniques that are being discussed?
Tim Larson: I can look into it and get back to you.
Alex Soroka: I think we're making this a larger issue than it needs to be. I don't know that there's a large amount of counties that aren't reporting. Let's see how many counties there are and see

what the timeframe is that we're looking for data for. Then let's see if we can use the CDL to estimate acreage and estimate the yield from surrounding counties using spatial products available that could help us in the last 10 or 15 years.
Robert D. Sabo (in chat): the differencing procedure between state and reported counties is the most standard approach vs. some of these fancier methods being proposed.
Chris Brosch (in chat): Tim and Alex seem to know plenty more than me, but factors of emphasis for one counties' influence on another county would be useful to approximate, but the perfect can be the enemy of the good to respect comments from Mark and Robert.
Chris Brosch (in chat): Robert I'm not interested in informing the Phase 7 discussion with Phase 6 data until we've identified other datasets that we could use in the future that have been underutilized. AMT was formed to sure up some known and also unknown unexpected results in CAST, unless I am mistaken. The academic questions about nutrient uptake and mass balance interpretations are beyond the scope of our charge.
Robert D Sabo: We have country crop removal rate annual estimates across the country. A certain county might not report its apple harvest, for example, so how would we account for that? How much effort will be put into finding these out?
Ken Staver: Are the county estimates that you have the same that is in CAST?
Robert D Sabo: I don't think the numbers are exactly the same.
Cassandra Davis (in chat): Do the crops in acres belong to the small crop/specialty crop category?
Tom Butler: I don't think all of them, but not sure.
Olivia Devereux (in chat): Most are vegetables and do.
Chris Brosch: Cassie, the yield unit when it's acres does not necessarily identify the land use category, like specialty crop, that those crop types are assigned to.
Elizabeth Hoffman (in chat): I think Chris is highlighting something that is part of the larger request to identify data gaps before we go down paths -- is there documentation of some of these previous decisions available so we can review on our own time rather than rely on folks reviewing because they were involved last time?
Tom Butler (in chat): Here is the documentation: https://cast-content.chesapeakebay.net/documents/P6ModelDocumentation%2F3TerrestrialInputs.pdf
Chris Brosch: We informed the model with as much NASS information as we could and at the beginning of the simulation (blue line), the model has overestimated the yields for corn compared to the real NASS information we have. More modern times we have damped the effect of influences on that yield with however we decided to interpolate corn yield.
Robert D. Sabo (in chat): It's only gone up ~25%
Eric Rosenbaum (in chat): Is there a graph showing nitrogen fertilizer use over this time?
Kate Bresaw (in chat): Thanks for bringing this up, Chris.
Eric Rosenbaum: Curious if there is a graph of nitrogen fertilizer usage, particularly in corn, through time with yield through time.
Tom Butler: I'll look into that and circulate offline.
Elizabeth Hoffman (in chat): Corn production, the primary driver of N applications, has not increased by the same magnitude as related to fertilizer when we looked at NASS data tables for MD. We echo this concern.
Robert D Sabo: Do you want a single line for the entire watershed or could we generate these county-specific?
Tom Butler: We are looking for information in each county. We're trying to fill those gaps with the information we do have.
Ken Staver: The two steeper lines are annual NASS survey reported yields. Flat one is the method in the CB watershed model. To Chris's point, this graph is more important to focus on

than how to come up with yield estimates on a small percentage of the bay watershed that we don't have data for. This is more important in my opinion.
Chris Brosch (in chat): Can we see the line that uses only peaks in the NASS data to simulate Max?

Action: Identify information gaps for crop yields and acreage for counties across the watershed. Explore spatial interpolation options for filling those data gaps.

Closing –10:55 - 11:00 (5 minutes)

Adjourn – 11:00

Up Next: Friday, August 11th, 2023 from 09:00 - 11:00 am.

Participants

Jackie Pickford, CRC
Tom Butler, EPA-CBPO
Olivia Devereux, Devereux Consulting
Emily Dekar, USC
Becky Barlow, VA DCR
Jess Rigelman, J7 consulting, contractor to the CBPO
Scott Heidel, PA DEP
Kristen Bisom, WVCA
Cassie Davis, NYSDEC
Tyler Trostle, PA DEP
Kate Bresaw, PA DEP
Eric Rosenbaum - Agronomist, Rosetree Consulting
Helen Golimowski, Devereux Consulting, CBPO
Mark Dubin, UME/CBPO
Tad Williams Virginia Tech
Chris Brosch, Delaware
Ruth Cassilly, UMD CBP
Alex Soroka USGS
Elizabeth Hoffman, MDA 9:02 AM
Elizabeth Hoffman, MDA
Ken Staver UMD Wye
Jeff Sweeney, EPA
Karl Blankenship, Bay Journal
Dave Montali, Tetra Tech, WV, MWG
Victor Clark - Farm Freezers - DE
Robert D Sabo, EPA ORD

**Common Acronyms
AAPFCO - Association of American Plant Food Control Officials
AgWG- Agriculture Workgroup
AMT- Agricultural Modeling Team (Phase 7)
BMP- Best Management Practice
CAST- Chesapeake Assessment Scenario Tool (user interface for the CBP Watershed Model)
CBP- Chesapeake Bay Program
CBPO- Chesapeake Bay Program Office (houses EPA, federal partners, and various contractors and grantees working towards CBP goals)
CBW-Chesapeake Bay Watershed
CDL - Cropland Data Layer
CRC- Chesapeake Research Consortium
EPA- [United States] Environmental Protection Agency
NASS - National Agricultural Statistics Service
NHD - National Hydrography Dataset
NM - Nutrient Management
NRI - National Resource Inventory
PSC – Principals' Advisory Committee (CBP)
STAC- Scientific & Technical Advisory Committee
TMDL- Total Maximum Daily Load
WQGIT- Water Quality Goal Implementation Team