November 2, 2022 Theory Working Group Call

Attendees: Shubhi Sharma, Christy Rollinson, Jody Peters
Regrets: Noel Juvigny-Khenafou, Carl Boettiger

Agenda:

1. What hypotheses from the manuscript could be explored within one or across the NEON Forecasting Challenge themes or with forecasts listed on the EFI forecasting profiles webpage or from the EFI community
    a. Hypothesis 1: The rate of decline in predictability over increasing forecast horizons differs across variables and scales
    b. Hypothesis 2: Predictability increases with biological and ecological aggregation
2. Resources
    a. Lit review of models typically used for the NEON Forecasting Challenge themes
        i. Here is a google doc to compile the models
    b. Figures of hypotheses that can be examined using the forecast challenge output
        i. Google slides with images
    c. GitHub repo with code that lets people drop in models to create forecasts for the challenge: https://github.com/abbylewis/EFI_Theory
3. Next Steps Discussed on the October Call to Follow up with on This Call
    a. Code for models for the Forecast Challenge
    b. Recheck expectations for questions we want to answer
        i. In the past the group has come up with expectations for the Challenge themes
            1. This Google sheet has the previous expectations, some of which is still useful. The group can update and add expectations for each theme relative to the hypotheses listed above from the Theory paper.
    c. Summary for what has been submitted for the Challenge - Shubhi was going to work on pulling this together
4. Note from the November Call - Discussion jumping off from the the comments Noel/Carl posted (see a copy of their comments at the bottom of the notes)
    a. Think voices dominated in EFI have focused on the stats which seems like it is easier to compare than the biological aspect. So thinking about how to dive into the biology is good
    b. Wondering about selecting a Challenge theme to focus on
    c. Thinking about phenology - don't want step on the toes of the group writing the phenology paper
        i. The group is getting close to having a draft that they might be able to share with the group in order to know what else to explore (e.g., now that we have all the NEON sites available,

   ii. But that paper focused on spring green up and was not for all the NEON sites so there may be opportunities to look at things like autumn redness or look at something like if daylength of climate influences leaf out.

d. Is there a table that summarizes what is involved with the NEON CHallenge - e.g., frequencies of the drivers, data available for the different themes?
   i. No, but the info is in the Challenge and could be brought into a table
   ii. The group started this table to bring in that information
   iii. Thinking about partitioning/comparing similar forecasts/data
      1. Phenology vs Terrestrial vs Aquatic
         a. Have similar data latencies
         b. May expect to see similarities between chl-a and plant greenness and may have similarities between aquatic DO/Temp and terrestrial fluxes
      2. In terms of aggregation
         a. Think this could be examined by looking at ticks vs beetles
   iv. We are worried about having enough forecasts to compare across the themes and if there aren't many submissions needing to rely on the Theory group to submit forecasts is tough. Think the Theory group can provide a few forecasts for a few themes, but not lots of forecasts for all the themes
      1. Want to think further about what are the hurdles to doing the meta-analysis?
      2. Is the problem a biology challenge or a technical challenge? Do we have enough understanding of the biology but need help creating the forecasts or do we need more understanding of the biology of the different themes?
      3. Thinking about the literature. Do we have knowledge of robust modeling frameworks for the different themes?
         a. We have the Google doc list with the different models used for the different themes, so this is a place to start.
      4. For example, the papers available about predicting phenology have exploded, but don't think that is the case for some of the other themes such as the beetles - Shubhi found one paper and knows of another that has been submitted
      5. Is there a way to describe the biology and then think about the technical hurdle?
      6. Is there a lack of information or a lack of skills?
   v. Can we approach it like Abby's paper? Utilize what we know about the models that are used in the different systems. E.g. If the biological data is super noisy, then this is what we expect.
   vi. Could we think about the first principles of uncertainty and applying that framework to different ecological systems.
   Then set up hypotheses or predictions for the themes to look at forecasts

or do one or two models for a system to see if theory is predicting them correctly.

vii. What about taking the hypotheses we have and go back to the existing lit in the domains for the Challenge themes and looking at what information is available and what supports the current hypotheses?
    1. Everyone is coming from it from their own personal exerpience
    2. If we can classify info into the hypotheses
    3. Then we can say these 12 papers say it might be like this vs these 2 papers
    4. Have a table for people to fill out with lit that they already know and then shift to be more systemic to expand the lit review
    5. If we can give people a form to fill out then it will give them direction and could be used to work with undergrads to train them on how to read literature

viii. For next meeting
    1. Shubhi will get the summary forecast information put together
    2. Jody to share the table Christy created and share the idea of having a form that is more concrete so we have a collective understand - pitch the idea on the next call and brainstorm what would go into the form

Slack Notes from Noel regarding point 1 above:

Is there a way to combine both hypotheses? Somehow I keep thinking that they are linked and should be addressed simultaneously, even more after re-reading Abby's paper.

Currently, I am under the impression that we are getting sucked into the technical aspect of modelling a little bit too much and forgetting the biology behind our datasets. We somehow also made that dichotomy in Abby's paper slightly with the arguments exposed to support hypothesis 1 being more technicals and the ones for hypothesis 2 more biological. However Abby also stated a potential link between the two aspects when she mentions that related variables seem to have similar rate of decline. I would say that this observation is probably due to linked ecological processes meaning that variables respond on a similar spatiotemporal scale and thus behave similarly to different modelling frameworks (it is my wild guess though). This could be our link between hypothesis 1 and 2.

Here, for the first hypothesis we say "differ" which is quite general but also quite obvious when we think about it. Better would be to state in which direction it "differs" and "why". Is it a technical artefact? A property of the model? An accumulation of noise in the data? The result of

a biological process? Also, similarly to our argument for hypothesis 2, different organisms/response variables means different life history/life span, therefore a different perception of time and space, thus finding the appropriate time frame and grain for the particular process we try to forecast is the central element that ensures the accuracy of the prediction over different horizons.

This makes me think this hypothesis should be reframed/reformulated slightly for the purpose of our modelling exercise and concentrate on the aggregation method used (like hypothesis 2 states). I am under the impression that our question should be about the influence of aggregation/how data are pooled or collected on the forecasting horizon.

I think that if we think that way our first selection criteria for datasets would be to find at least one that records biological processes at the right spatial and temporal scale. Do we think that weekly tick measurements over a few plots is an accurate scale to observe the dynamics and test hypotheses about aggregation and forecasting horizon across scales? This has to be related to the lifespan and traits of the organism I guess. If we have datasets that are accurate enough, then we can select the best ones to answer questions about forecasting horizon and aggregation. Maybe I am rambling around a bit and this was already discussed in a past meeting. Looking at the NEON challenge datasets I am incline to say the phenology dataset is the best one to address both hypotheses. If I am not mistaken it has the finer temporal resolution but also it might be easier at this stage to work with sessile organisms to avoid a layer of complexity linked to dispersal.

Slack Notes From Carl:

I also can't join today (teaching) but would love to re-egnage with the group. (had to step away since I was the handling editor for the MEE submission I felt I needed to keep isolated from the internal processes of the paper).

Super interesting points @Noel Juvigny-Khenafou. I agree with your points on both hypotheses. (Minor remark on hypothesis 1: I think predictivity doesn't always decline as horizon increases -- e.g. as we enter winter all of our forecasts are becoming pretty predictable at most temperate sites).

Regarding which theme to start with for hypothesis 2: I think any could be great, but we might not want to restrict the analysis to the level of aggregation pre-determined by the challenge? Consider beetles for instance, where there is lots of ecological aggregation to explore -- predicting individual plots, or individual species counts is super noisy, but aggregations across

taxa or up to the site level are a bit easier. aggregations to the eco-domain become even more predictable, dominated by an obvious seasonal signal. (the beetles team discussed a lot what would be the right level of aggregation with the assumption that predicting individual trap-level counts would be dominated by observation noise, while aggregating across all NEON sites nation-wide would be dominated by seasonal patterns, and so the most ecologically driven variation would appear a the site level of aggregation, but that was mostly just a hypothesis)