May 3, 2022 Theory Working Group Call

Attendees: Amanda Gallinat, Gerbrand Koren, Noel Juvigny-Khenafou, Glenda Wardle, Jonathan Tonkin, Cole Brookson, Jody Peters, Jaime Ashander, Elyssa Collins, Mike Dietze, Steph Brodie
Regrets: Abby Lewis, Andrew Allyn

Agenda:
1. Poll for Calls in June-August

2. Manuscript submitted - yay!

3. Theory WG recap for EFI Conference
    a. <5 minute pre-recorded presentation
    b. Recap of manuscript and short review of ideas for future Theory WG calls (outcome from point 3 below)
    c. Amanda will do the pre-recorded update

4. Recap from icebreaker from the April call about ideas from the group about what to work on next
    a. Test some of the hypotheses put forth in the manuscript
    b. Several versions of figure 1 that lays out forecasting - produce something in a non-manuscript format, but through EFI that lays out underlying theory of forecasting and expand and provide through EFI (not necessarily through traditional manuscripts). Graphical abstract of what forecasting is about
    c. Test transferability
    d. Work on/test the toy model that Abby and Elyssa started a while ago
    e. Use existing experiments to reanalyze with a forecasting approach. Start to understand how to do science differently. Put into practice the idea of how forecasting can change what we are currently doing. Especially some of these experiments that have been done across multiple sites
    f. Journal club - dive into and discuss papers that cover existing theory. (student group has had fun and productive discussions on complexity and simplicity - could possibly have cross-over call)
    g. Opportunity for synthesis and to learn from each other. Thinking about different scales of ecological organization and where predictability scales across that. Synthesis across domains (marine, terrestrial, freshwater). Think about different end users - resource managers, ag management, water management. Sounds like some of the ideas in the very first manuscript outline-- comparing predictability across levels of biological organization, geography, life history, etc
    h. Model validation/model selection using large datasets or past datasets. Transferability

      i.   Use manuscript as starting point to put things into practice to show benefit of forecasts. Like having the manuscript to motivate to read new papers and organize thoughts more than just discussing a topic.

      j.   Working on the uncertainty component of the manuscript - is uncertainty the key part of forecasting or required? Can work to explain the different sources and show how to do it and visualize.

5.   Dicussion during the call

    a.   Elyssa's gave an update about the toy model she and Abby had worked on in the manuscript writing process

        i.   Think this is a great idea at the high level

        ii.   Something to think about - what are the early steps that need to happen to move us towards these types of analyses?

        iii.   From 2017 uncertainties paper - next step was forecasting a whole bunch of things to understand the dominant sources of uncertainty in different systems.

        iv.   So think part of it is pulling togeht erh forecasts out there to do these types of analyses to go beyond the lit survey paper that Abby led recently in 2020. How do we go beyond the lit survey? What is the next step?

        v.   We now have Forecasting Challenge for 5 themes across multiple sites. No one is analyzing the output from a Theory perspective. Currently thinking about it within a theme and what we can learn within a theme vs across themes.

        vi.   Could go beyond to the EFI community. If you could convert yoru existing forecast to the EFI standard and make it availabe for synthesis and be a co-author, then can leverage a big synthesis and can do gap filling

        vii.   Where are existing foreacsts not there yet? Transition from modeling to forecasting - what would it take to get people who do modeling analyses to get into an analysis of predictability

        viii.   From VT Lit Review - there are few papers that have done enough uncertainty partitioning to do these sort of analyses.

        ix.   Forecast Challenge - has forecasts by theme by sites by team that could be analyzed. There are thousands of forecasts that have been submitted.

        x.   Year 2 of Challenge is a building year for the CI with a big push in Year 3.

        xi.   If the Theory group has ideas, now would be a great time to think through ideas

        xii.   Have EFI [Forecast Project page](#)

        xiii.   For Forecast Challenge - do people need to use a standard to submit their forecasts?

             1.   People have to submit their forecasts to a standard portal in a standard format.

             2.   In theory teams are supposed to provide metadata, but in practicie it hasn't always done. But tracking down that data isn't too much work

        3. If we wanted to take a subset of the forecasts to partition uncertainty, this is what would take more work

        4. We know the total uncertainty used in the forecast. They will say - do I have drivers (Y/N) and how many drivers if they had them

        5. Do models with temp perform better than models with precip

        6. People not required to submit the raw models or a workflow that is easily executable. (Has been discussed, but is a higher bar than we are willing require).

        7. From Glenda: What I am hearing from MIke's ideas is that it is not clear what is needed to submit along with a forecast that makes synthesis easier to do.

    xiv. If we had a clear set of analyses and a clear protocol and if authorship is offered, then think there will be teams willing to participate

    xv. We could start thinking through this by thinking of 1 theme. Think about what it would take to do the analyses within 1 theme and then expand on from there.

    xvi. Would need to start actioning things that would go into the manuscript

b. Reanalyzing experimental data is another suggestion that Jono is interested in. If there is long-term data (like Cedar Creek) and could

c. Think about what it takes for good synthesis in the new forecasting era. When you try to bring together all the population matrix models, what is lacking is in the older literature, authors would add geographic location. We have foresight along with hindsight - what could be a gap like this that we would want to

    i. We have the option now to get into forecasters head as they put together their forecast. Are there things we can suggest to guide forecasters

    ii. What makes a study that gets published better suited for a synthesis?

    iii. Thinking about the forecast challenge - what more do we need?

d. The Phenology Challenge has the most teams submitting and the most forecasts submitted. In Round 1 it was the only one with forecasts submitting daily (as opposed to monthly for the rrest of the themes). Now in Round 2 additional themes have more forecasts being submitted on more the daily time period (aquatics, terrestrial water and carbon fluxes)

e. Phenlogy is also a good test case because it uses fairly simple models. In contrast to the carbon flux. Carbon is now running daily, but models are using 3-4 state variables and 5 fluxes to keep track of.

f. Latitudinal gradient for the 1st round was eastern temperate deciduous.

g. Thinking about how predictable is nature - would you expect there to be a pattern based on latitude? Is it clearer in the north that it goes bang vs in the south?

    i. From Mike's current analyses: Model skill improves as you go south. The most skillful forecast leaf out earliest.

h. Going back to experiments -

i. There is momentum to think about what specific hypotheses we can test with the Forecasting Challenge

    i. 2 hypotheses from the manuscript

              1. Predictability decrease with forecast horizon
              2. Predictabiltiy increases with aggregation (by biological system or scale)

    ii. What is the possibility for the phenology challenge vs the others?

    iii. Most challenges are making predictions at the site scale

    iv. No challenge is being run on a wall to wall basis, but phenology if you gave them a grid to run on, most of them could do it. If gave them the driver data and pulled remotely sensed greenness think most teams could do gridded runs to answer a specific question

    v. It will be a long term project in that this year would need to figure out what it would require for a protocol for the next round

    vi. Advantage of phenology - they are currently running continuous (vs spring and fall runs). So if we came up with a protocol, could have phenology test it out this fall

    vii. Now it is predicting the composite greeness index from phenocams (cameras pointed at forests). So it is stand scale aggregate index, but the thing people are predicting is the greeness index. Every day get 35 new numbers ofr tomorrow, the next day out to day 35 (not every teams does this, but in theory have this rolling forecast).

    viii. Have been looking at lead time in the phenology analyses

    ix. Have not looked at uncertainty or thought about transferability (only have looked at the intercept for site and seen that it is different).

    x. Don't have a draft manuscript yet for Phenology analysis, but think will have a draft manuscript by the next Theory call.

j. Enthusiasm for testing questions in manuscript - predictability and maybe aggregation, uncertainty partitioning, and transferability

    i. Think aggregation overlaps with opportunities for synthesis with different scales of ecological organization.

k. Have good opportunities with existing forecasts from the challenge and from the EFI community (reach out to people who create forecasts)

    i. Would be good to tap into these resources before producing new forecasts

    ii. Is there anything the group can look at before the next call to think about what we need to ask for to make synthesis more feasible.

    iii. https://ecoforecast.org/efi-rcn-forecast-challenges/

    iv. Details about the specific challenge themes here: https://projects.ecoforecast.org/neon4cast-docs/

    v. Put these links in the Slack with a description of the homework
        1. Look at the resources and identify are there hypotheses in the mansucript that you feel motivated to test using these resources
        2. Map the challenge to the hypotheses we have in the manuscript
            a. Beetles suitable to aggregation
            b. Phenology suitable to exploring uncertainty and transferability

vi.   Use networking at the May Conference to see what the EFI community is excited about