August 9, 2022 Theory Working Group Call

Attendees: Christy Rollinson, Glenda Wardle, Abby Lewis, Quinn Thomas, Caleb Robbins, Jody Peters, Cole Brookson, Shubhi Sharma, Steph Brodie, Mike Dietze
Regrets: Noel Juvigny-Khenafou, Amanda Gallinat

Agenda:

1. Poll for calls in September to December

2. Introductions

3. NEON Forecasting Challenge background info and data - see if Mike/Quinn can provide input for this

4. Hypotheses from MEE manuscript to test with NEON Forecasting Challenge forecasts

5. What predictions could be documented before Round 3 of the challenge to test at the end of the round?

6. Any changes to the Round 3 rules that are needed for analyzing predictions?

7. 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

Notes
1. Look at the predictability of ecological variables using the NEON Challenge and look for trends.
2. Overview of the NEON Challenge
   a. The Challenge is funded by the Rsearch Coordination Network funded by NSF.
   b. Sets protocols and rules and develops tools for creating forecasts from NEON data that haven't been collected yet.
   c. Challenge spans 5 ecosystem/population themes
       i. Phenology - greenness/redness of plants
       ii. Terrestrial carbon fluxes - measured by eddy covariance
       iii. Aquatic ecosystems - temp, do, chl-a
       iv. Tick population
       v. Beetle communities

d. Status of the Challenge - round 1 went all in where we tried to do everything. With design teams for each teams, Quinn/Carl Boettiger working on CI, and teams that submitted provided forecasts.
e. Round 1 mostly has phenoloyg forecasts
f. Have ~5 forecasts for terrestrial
g. 0 for beetles
h. Water has 1-2 teams
i. First thing to set up is that we don't have a catalog. We can't create an analysis.
j. The project envisioned this type of analysis happening in Round 5
k. Round 2 has had very few forecasts submitted. But has been an opportty to update the standards to follow the EFI developed standards
   i. Also been able to have almost real time forecasts for phenology (0 day latency), water (1-2 day latency), terrestrial flux (~5 day latency)
   ii. Now have capacity to have a running score of the forecasts
l. The critical thing is that you can't do synthesis without submissions. There isn't a huge catalog to analyze yet.
m. When the group does synthesis, Quinn encourages the group to reach out to the groups that submitted forecasts to include them on the manuscript development/synthesis
n. Most useful thing right now is - 1) submit forecasts, and recruit people to submit forecasts, and 2) think about revisions for round 3
o. Can this group provide feedback to make changes for round 3.
p. Are there sites that can be added? Don't want to add a new theme since that takes a heavy list
q. Are there key metrics that are missing from the automated scoring
r. Want to be able to
s. The idea of having rounds for the Challenge will be going away. Classes and courses like to use the Challenge for class projects. Want to make the Challenge accessible for anytime so people can submit every single day.
t. Want submissions to be done at anytime.
u. Want to have folks get their forecast automated.
v. Have called it a "challenge" so it doesn't feel competitive. The challenge is to forecast NEON data before it is collected. We have built the CI to support that.
w. Perhaps a Theory or WG focus efforts on getting people to submit forecasts for 1-2 months to motivate people and make it so people don't have to submit all the time.
   i. "Forecast blitz"
x. Phenology was popular because it was real-time and the phenology modeling field is active. People building those models may be more familiar with automation and CI in general.
   i. Think that the fact that you can submit every day let people automate their forecasts
   ii. Getting the right people involved was useful - Andrew Richardson who invented the phenocam network was involved in the design team

        iii.   Phenology models are fairly straightforward to implement

    y. Think the land and water themes are ripe for getting more submissions

        i.   The comparison of gases in water and gases in air is really interesting

        ii.   Could have Theory Group members develop forecasts for this and have a theory group blitz and could have similar forecast structure

        iii.   Perfect example of leveraging by CI and filter by who is submitting forecasts so you can compare the curated set of forecasts submitted by this group and can drop the noise of other groups that have submit

    z. From Christy: I would love to compare phenology & carbon... in theory those should be very similar, but does that work in practice

3. With the idea of Characterizing predictability of different - would it work for the forecasts/themes and the limited set of models we are working with?
   a. This is the Grand Challenge question for the whole field.
   b. May be worth to narrow the focus. What is a bite sized bit that could be accomplished
   c. Could have everyone submit models that apply across themes
   d. Once we get a workflow the group agrees on and gets up and running should allow for making changes and adding complexity
   e. Abby's experience had file called "model" that people would change, but the rest of the workflow for submitting forecast were the same which allowed people to change the the model

4. There isn't any

5. There is a template repository that is available - this is nice because it sets up GitHub actions to make the forecast automated

6. 2 required steps
   a. 1) read in targets (call to a URL)
   b. 2) submitting (this is a function that is available in an r package)
   c. When you get covariates it is a little more trickier because you have to get them from the right time. Quinn is working on the new tools for that
   d. The other thing is the metadata. This would be a good test to see if the metadata standards is useful for the synthesis that this group would want to do

7. We are in the early days of the synthesis. We are trying to make sure to manage expectations. This will not be the definitive synthesis. It is a pathway to synthesis and we are tyring to figure out how to best support that by site gap analysis, forecast blitz for a specific time, and covariate assessment. This is a measured and good direction, but be patient with where everyone is at the broader level

8. Christy likes the idea of focusing on standardized covariates, the NOAA met driver availabiltiy freaks out for a week

9. How did the RCN foresee answering the synthesis question when writing the RCN grant. How did the RCN they see answering the question?
   a. Proposed to build the catalog and then proposed to have a meeting to work on the synthesis. Kept it flexible

    b. The Phenology Challenge group is working on a manuscript doing a cross model synthesis approach. It is also a first step and working with the submissions as they are

    c. When you work with a genuine set of forecasts you have things that go awry (e.g., someone forgets to submit one day)

    d. Core question of predictability led up to the core planning leading up to the launch of EFI and moving forward. Theory has always been at the heart of EFI.

10. Compare to the weather forecasting - 50 years on they are still improving. This is something that we will be working on long term

11. Think we can make useful progress if we focus on specific sub-questions. And be honest with how far we can get in answering those questions

12. One question at the interface of predictability and transferability - does predictability increase/decrease along traditional environmental gradients

    a. In phenology example - preliminary analysis looks like predictability decreases later in the season. But is that because the forecasts submitted later were not so good

    b. This highlights the benefit of a blitz that this group would be in control of

13. We hope getting up a simple model is not difficult. If it is, then that is feedback the RCN/Challenge wants

14. It'd be interesting to see if the "best" models were developed by ecologists, statisticians, or collaborations

    a. From Mike: I don't know if we'll do "better" but I really like the idea of throwing the "same" model at different challenges so that we know, on the backend, that we'll be able to say something about transferability

    b. And site

    c. From Abby: I think many folks assume process-based models will have greater transferability than statistical models. It would be cool to do that comparison using established process-based models. But of course one limitation is data availability to run those models as forecasts

15. All of EFI should be gearing up to increase participation in the next year to get papers out from the different Challenge areas.

    a. If Theory is interested, then Theory is motivated to submit forecasts or encourage others to submit forecasts

16. Think there is room to put Theory to work in a different way

    a. From the outside if there is a Theory group they must have more foundation, so their model should hit it in a different way compared to other models coming in with fancy statistical method

    b. Have this group wear the Theory hat more - what would be a better way to build a model based on principles so Theory can inform practice. Use the wiseness of not overparameterize.

    c. Let's not just blitz this one. Let's put something in the mix that does better than the random mix.

    d. Thinking about metadata - Christy's experience from judging cherry blossom forecast. Judged based on stats vs theory

      i.   When groups submit forecasts - is there anything that documents where the models come from, the covariates, or the model structure

      ii.   Can the metadata unpack the motivation?

          1.   Currently it does not. For phenology had teams write up details about their models

          2.   EFI standards started from the Theory group, the merged with CI/methods before splitting off on its own.

          3.   Things that are there that are important are

               a.   Measure of how uncertainty is being handled for key factors (initial condition, drivers, parameters)

               b.   And the dimensionality - how many parameters, how many covariates, how many state variables

                    i.   This was a compromise of not documenting anything vs documentation that was so onerous that people wouldn't want to fill it out

                    ii.   Haven't had enough teams put the metadata standards through a test.

                    iii.   Having teams work on this now would be awesome and set up for years 4 and 5

                    iv.   In the existing phenology analysis boiled it down to boolean analysis - did the model of uncertainty or not, did the model include covariates?

      iii.   Thinking in figures

          1.   Christy would look at the standards/metadata and think about the hypothetical figures to develop from those.

          2.   Everything comes down to variable, horizon, metric, and model - those are the things to play around with and then on top of that are all the ways to describe the environment

      iv.   Another big gap - getting the right metadata gotten efficiently. Can you make it easy for people to choose from dropdown options

      v.   Did the metadata development think about it from a survey perspective?

          1.   No purely thought about what is the information and how to encode it into EML?

          2.   It wasn't envisioned to be interview based

          3.   Would be worth think about whether some kind of GUI or web portal would make it easy for people to enter/validate their metadata

      vi.   Take home from Quinn before he dropped off the call - the ideas that are being discussed here are perfectly in align with what the RCN was set up for. No longer thinking about where the data is coming from. Now thinking about the models and the forecast.

e.   Wrapping up

      i.   Have touched on a number of options. Here are a few action items

1. Look at current standards - do they meet the goals of current figures we want to make
2. Get started on a shared framework for submitting theory group forecasts

ii. Is anyone excited about pushing forward on these two?
   1. Mike has been looking on the standards so need a fresh set of eyes
   2. In September when Mike has more time, wants to go back to FluxCourse example with C model running and submitting to the Challenge. It is process based, but is very simple. In one day, 2-3 teams were able to add to the model and extend it. For C it is a good base model for exploring additional complexity. Could provide another example of a front to back workflow
      a. This would be a good jumping off point for other themes
      b. Is it at a point where Abby could look at it and think about
      c. Yes - you can find it here: https://github.com/mdietze/FluxCourseForecast
   3. Also the Forecasting Standards draft v0.4

iii. Would something like a lit search to see what the state of published models for each of the challenges be useful for framing the models/complexity we're getting in Challenge submissions?
   1. Could send around a Google doc to compile this
   2. Start looking at what range of complexity we are getting
   3. Will also help to carve out a more precise questions
   4. Phenology is around change in greeness
      a. Would decisions made in deciduous (or temperate) systems in N. America be applicable to non-deciduous systems like in Australia
      b. Round 2 of phenology challenge inclused grasslands and shurblands. But haven't moved into the boreal or tropical systems/ evergreen systems. They have phenology but it is harder

iv. Getting back to sketching out what kind of graphs we want to make and then looking at what forecasts are available or where are their gaps

Action items
- Lit search
- Sketch graphs
- Set up framework and code for submitting forecasts