January 22, 2024 Theory Working Group Call

Attendees: Marcus Lapeyrolerie, Cole Brookson, Caleb Robbins, Abby Lewis, Bilgecan Sen, Shubhi Sharma, Alyssa Willson

Agenda:

1. Announcement: Abstracts due for the EFI 2024 conference in ~2 weeks on Feb 1
    a. Details about the conference and links to abstract submissions and registration are on the conference webpage: https://bit.ly/efi2024
    b. RMetS is accepting submissions for the joint Special Issue: **"For a future informed by science at the climate-ecology interface"** in *Meteorological Applications* and *Climate Resilience and Sustainability*. The deadline has been extended to August 29 so anyone can submit manuscripts, but also to allow people attending the conference to consider submitting manuscripts.
    c. Details about the Special Issue: https://rmets.onlinelibrary.wiley.com/hub/journal/14698080/call-for-papers/si-2022-011060

2. Blog post idea for code review materials (Jody)
    a. Jody will draft the blog post and run it by the group. There is no definitely timeline for this, but hopefully within the next month or two

3. Manuscript - using the NEON Forecasting Challenge to explore predictability across variables and scales (Caleb)
    a. On the November call we talked about looking at plot with a summary of R2 vs Forecast Horizon overall (vs by month like Caleb showed in November)
    b. On November call also talked about looking at the dashboard to see the forecast and raw data patterns to help give direction for analyses
    c. Repo where Caleb is working on this: https://github.com/robbinscalebj/NeonPredictability
    d. NEON Forecasting Challenge registration updates
        i. New registration form
        ii. Neon4cast R package update from the newsletter: If you are using the EFI docker image in your automated workflow, it will update automatically. Otherwise, we recommend manually updating the neon4cast package with the following code. remotes::install_github("eco4cast/neon4cast"). While the updates are backwards compatible (you will not need to change your code), you will be required to register your model if you have not done so already.
    e. Caleb showed greeness forecasts
    f. Y-axis is normed NSE - goes 0 to 1
    g. Models don't incorporate initial conditions

h. GAMs have been hard to run because of the massive amount of data. Getting there by splitting up by variable
i. Fits can be good by site
j. How much to care about autocorrelation and the oscillations of points
k. Peaks through time
l. Every point is a mean of a forecast at the forecast horizon for a week
m. In the plots: each facet is week of year.
n. How do the met forecasts do over these horizons? If initial conditions are not included then the met forecast may be useful
o. Need a model to account for initial conditions
p. Expect models that include initial conditions would have boost of performance in short horizons, but when initial conditions are not included, then think that performance may improve at the later horizons.
q. GAMs - expect GAMs to be good at getting oscillations, but that appear to be smoothed over here.
    i. Think if each forecast was fit with a GAM then the oscillations would be picked up
    ii. Analyzing cross-time affects, think may be reducing the oscillations
r. Looked at Caleb's GAM code
    i. Forecast week and horizon
    ii. No drivers - not looking at effect of drivers on the models
    iii. Forecast week potentially matters because of some other driver due to seasonality - winter conditions are different from summer conditions due to air temp (or other relevant driver)
    iv. Could analyze with mean temp or classes of air temp or classes of some other predictor.
    v. Is week categorical because it is grouping by week number?
    vi. Thinking about it environmentally - looking at it by week, examining rate fo decay of forecast performance is different from first week fo the year vs in the summer.
        1. From Abby's experience with water forecasts are a function of air temp when the water is stratified vs mixed due to weather
        2. Caleb's work lets us see how forecast performance changes over time
    vii. Start with forecast performance declining and then it increases - wonder why that is
    viii. If you add a driver like temp, it can help with interpretation in terms of seasonality and would be interesting to look at the differences between sites where temps are different at southern vs northern sites
    ix. There is an interaction between forecast and site ID. Forecast is not the same at all the sites.
    x. Forecast week might be a better way to reflect the changing in light availability

s. Don't think the model will capture the wiggliness at the fine scale if it is fit on the whole dataset
    i. What Caleb wants to have happen is want the GAMs to fit the specific trends and then see the overall thing that is happening at a bigger spatial level
        1. Not sure if the model is doing what Caleb wants it to do
    ii. Q for Marcus: How long would it take to train an LSTM to the forecasts?
        1. Depending on the type of machine available, but not as long as you might think
            a. This could be a way to get at the initial conditions
            b. Marcus trains all ML models on his lab server in 4 minutes on a per site basis
            c. This is something Caleb will connect with Marcus on and play around with

4. Manuscript - Uncertainty analysis that decomposes different uncertainties and ties that to intrinsic predictability which would have some analyses from the Challenge forecasts. Simulations confronted with some data. (Shubhi and Cole)
    a. Update, Cole will be leading the analysis on the predictability
    b. Previously had looked at weight permutation entropy (WPE)
    c. Suprised at how unpredictable the sensor based measurements were, e.g., the carbon measurements
    d. Cole trying to wrap his head around data gaps. Bilgecan had used a model to interpolate data
    e. 3 ways for handling data gaps
        i. With a model
        ii. Average up
        iii. And using NAs
    f. Cole has been playing around with ways to average up
        i. WPE does not do well with data gaps - so can interpolate to fill the gaps
        ii. Depending on how big the data gap is there may be places where it is harder to fill in the gaps
        iii. Predictability is contingent on data.
        iv. Variance with gaps will explode. You can leave that and let the variance decomposition help you out.
        v. But Cole is trying to average up to eliminate some data gaps. Tested this by averaging for 2-7 days and viewed how that has reduced gaps
        vi. When you average your timescale you increase predictability.
    g. What are the tradeoffs in different ways to fill the gap?
        i. Leave behind a model free measure or
        ii. Pair variance decomposition with the metrics will let us pinpoint what it will improve in the predictability of systems

1. For beetle, the observation process is so noisy that the predictability will be low and uncertainty decomp will point to noisy data so the recommendation is that we need better data
2. For time series with lots of data, we might see the recommendations be to scale up or average. As you scale up or average then the gaps will be filled

iii. Bilgecan - penguin paper there was an application need to fill the gaps. People were building models that way already. It was not related to permutation entropy but was being done separately for other reasons such as for forecasting or looking at population dynamics for conservation and management decision.
1. For that purpose it made sense to fill in the mising data since it was already being used that way.
2. But don't think this should be the default. When you do this it is no longer a model free metric and the method used to fill the gaps will affect predictability

iv. For Bilgecan's other paper - used NAs. Didn't fill the gaps. Lets see what the data tells us and see what the data tells about predictability.
1. Can use NAs to calculate permutation entropy
2. This will affect sample size

v. Bilgecan - recommends unless there is an obvious reason to fill in gaps, then should go with NAs

vi. When looking at it from an uncertainty perspective, it will be more uncertain, but can see a pattern

h. How do you treat missing points and how your uncertainty decreases and how does your permutation entropy decrease

i. Since this is Challenge data - interesting to see how people deal with the missing models in their data
   i. Could be a helpful sub-analysis of the paper to talk about how data gaps have affected predictability

j. Another option - what about calculate WPE for a shortened part of the time series?
   i. For the Challenge, people will truncate the data to a more advantageous section of data

5. GitHub repos
   a. GitHub repo: eco4cast/predictability - comparative analysis of predictability
   b. GitHub repo: Forecast_submissions - forecasts submitted to the Challenge

6. Model Development for the NEON Challenge
   a. GitHub repo: eco4cast/Forecast_submissions
   b. New model descriptions document

Previous Notes and Links for Reference

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

8. Resources the group pulled together to test hypotheses
    a. Google sheet with a summary of drivers, data availability, number of sites, etc for the Challenge themes
    b. Lit review of models typically used for the NEON Forecasting Challenge themes
        i. Here is a google doc to compile the models
    c. Figures of hypotheses that can be examined using the forecast challenge output
        i. Google slides with images
    d. GitHub repo with code that lets people drop in models to create forecasts for the challenge: https://github.com/abbylewis/EFI_Theory