May 15, 2023 Theory Working Group Call

Attendees: Shubhi Sharma, Abby Lewis, Caleb Robbins, Jonathan Borrelli, Alyssa Willson, Avery Hill, Kathryn Wheeler, Jody Peters
Regrets: Cole Brookson

Agenda:

1. Do we need to poll for a new time for the group to meet in June-August? We will send a new poll for Sept-Dec.
    a. If needed, here is a poll for scheduling calls for June to August. Make sure your time zone is selected. Focus on the general day/time that works for you. Ignore the specific dates.
    b. Stick with this same time for June 12, July 10, August 14

2. Get an Update About: Comparative analysis of predictability using an uncertainty framework to decompose how predictable something is and why or why not (Shubhi, Cole, Noel, etc)
    a. GitHub repo: eco4cast/predictability
    b. Have intrinsic predictability values for all the NEON forecasts
    c. There are a number of decisions you have to make to calculate the intrinsic predictability values
    d. Do we have the performance metrics for each of the forecasts?
        i. It would be nice to compare the forecast scores with the intrinsic predictability values
        ii. Have the forecast dashboard - but unfortunately the scoring is down right now
    e. To measuring intrinsic predictability we are using weighted permutation entropy (WPE). The more repetitions in the time series, the more predictable it is.
        i. Get a value between 0 and 1. 1 is chaos, 0 is perfectly predictable
    f. Follow up from questions from the last meeting
        i. What is embedding dimension?
        ii. Does seasonality reduce predictability?
        iii. Should you detrend the time series before analyze it?
        iv. Is there any accounting for observation error in calculating WPE?
            1. Haven't done it for NEON so far, but can talk about doing it
    g. Shubhi calculated the WPE for the NEON Challenge data streams
        i. Beetles are doing well (0.58)
        ii. Phenology and Terrestrial are not doing as well (~0.9)
            1. Explored this further - aggregated it by time scale. Daily is not predictable, but gets better at monthly and then at yearly is quite stable (low WPE)
    h. Thinking about seasonality - Phenology and Terrestrial are very seasonal/daily. Is that why these values are really high?

i. Shubhi doesn't think so. All are seasonal. Tick and beetles also have seasonality.
ii. For phenology - if we aggregate temporally will it get better?
iii. Phenocam is measured daily. Think phenology at monthly will be at a
iv. How drastic does the magnitude need to be different to know what is more predictable? If phenology-rcc-90 is 3 points more than phenology-gcc90 (0.90 vs0.87 WPE values) is it really more predictable
v. Does the index look at all the data points indiscriminate of their time or is it sensitive time of day or day of year? It is sensitive to the time. If you have data from days 1-10, that pattern will matter, it doesn't reshuffle that. So if you get a set of ancillary variables, you tell it you want it to reshuffle. It doesn't reshuffle. It conserves the data in the pattern it is recorded.
vi. Surprised that the 30 minute terrestrial data is so close to chaos. There is definitely a lot of observation error
vii. How does it distinguish observation error to process variability? Does the calculation miss that there is a diurnal cycle?
viii. Haven't done it yet, but we can discuss how to account for observation error. The random noise is driving the WPE to chaos.
ix. Should Shubhi take out the observation error?
    1. If we knew the true value then we could take it out
x. Not convinced that the patterns are due ot observation error alone. Think the diurnal or seasonal changes is why we see the patterns.
    1. Phenology averaging to an annual scale will show that there is the same amount of greenness over time unless there is an extreme drought
    2. Think the things with more predictability have interannual differences
    3. Go back to the idea of incorporating info about natural cycles would be useful
xi. Plant greenness across different sights will vary
    1. Not sure if this is the same for lakes
xii. Shubhi looked at predictability for each site
xiii. WPE is trying to capture how redundant the time series is. The more redundant or cyclical mean that it is more predictability
xiv. PE values are sensitive to Tau
    1. From Caleb: there is a 2013 paper that deals with tau and cyclical patterns. Think they have a protocol for thinking about this. Caleb applied this to dissolved oxygen.
        a. https://link.springer.com/article/10.1140/epjst/e2013-01862-7
xv. Beetles have a lot of 0 values so think that is driving the WPE down.
xvi. Most of the time aquatic chlorophyll is low and then it spikes with a bloom
xvii. Things to tackle - data availability and observation error
xviii. Does it matter that the values are not normally distributed?

1. E.g., chl-a will mostly be at low range and then will have a few very real outliers that skew the distribution.
2. This doesn't matter
3. Switched from permutation entropy to weighted permutation entropy because it weights it by the variance of your time slice

xix. Will be interesting to compare permutation entropy to performance of the null models (persistence and climatology). Think they may be related, but the difference between the metrics may be really interesting

xx. Idea was to calculate intrinsic predictability and realized predictability. The space between the two will show the model improvement space.
1. This will be the information theory metric
2. Want to have it in comparable units

xxi. What about observation error?
1. Is there a way to renormalize the index to what a pure noise would look like if we knew the observation error say for eddy covariance, could we set bounds to say even a perfect model would have this amount of noise?
2. Make it a ratio?
3. If you had a perfect model and that perfect model had real observation error what would you look for?
4. Random noise model scores 0.99. What does a perfect model look like? Completely flat.
5. What would a perfectly predictable model look like that we can then add observation error on to come up with in practice what you could actually expect to achieve
    a. Shubhi hasn't seen anything below 0.3 or 0.4 in the papers she has been looking at for ecological time series examples

xxii. Shubhi will work on developing a figure to show how the WPE is calculated

xxiii. Could recruit people to join the conversations, if someone else in the world is thinking about applying this in time series then invite them to these meetings

xxiv. If there is time of day in the raw values (thinking about terrestrial values). If you have window of 6, but have the actual window of 48, then might be good to adjut

xxv. Recommended window size is 3-7, but there is no recommendation for how to choose. Shubhi picked 5, but also calculated it for 3 and 7 and it was very variable.
1. This is using the most recent 3-7 time slices to calculate
2. It will be sensitive to observation error

xxvi. Characterizing the datasets - wavelet analysis to look at the dominant scales in the datasets will be interesting

xxvii. For the flux data, could you calculate the predictability by time of day?

3. Get an Update About: Model Development for the NEON Challenge
    a. GitHub repo: eco4cast/Forecast_submissions
    b. Goal: Submitting as many models as possible across all the Challenge themes. If we use the same models then want to see what is works well for what themes and why that might be
    c. Wishlist of models to submit
    d. Abby has been working with code Caleb had put on the repo for the all sites LASSO model and almost has that working. Hoping it will take shorter time to calibrate, but still seems intensive. Was a bit shorter for Caleb, but still takes awhile.
    e. Caleb is working up the rainforest and all sites predictors and then think it is worth modifying those again to rethink the data every time. The hyperparameters take a long time to fit, but think you can refit the models quickly. Think it will take similar amount of time to refit as what it takes for refitting the linear models
    f. Abby is hoping to get more models set up in the
    g. Caleb - for next time - want to talk next time how to incorporate more uncertainty. Caleb talked to Jake Zwart about it and it is complicated with how he is doing it. Would like to see if there are easier ways to do it.