June 23, 2020 Theory Working Group Call

Attendees: Mike Dietze, Christy Rollinson, Lauren Buckley, Glenda Wardle, Winslow Hansen, Peter Adler, Jaime Ashander, Amanda Gallinat, John Foster, Kathryn Wheeler, Abby Lewis, Lynda Bradley, Hassan Moustahfid, Christa Torrens

Homework for next time: Continue vocab discussion. Complete homework RCN topics - fill out spreadsheets

Agenda/Notes:

1. Anna and Gretchen to give an overview of their manuscript to look at the difference between forecasting and prediction vocabulary
    a. Different fields of forecasting define forecasting. This project was spurred on by the 2019 EFI meeting in DC where there were a number of definitions given.
    b. Input from Diana about the diversity of the forecasting and being able to explain what ecological forecasting is to the group
    c. Worked with grad students to look at subfields of forecasting - weather, earthquakes, tsunamis, hurricanes, finance & business, economics, criminology, politics. Conducted lit search to see how the other fields define forecasting.
    d. Sending out to authors who have published on forecasting.
    e. Have 100 responses so far
    f. Working on analyzing data to write up
    g. One of the goals of the paper will be to provide a framework for tiers/components of forecasting. Ultimate goal is decision support, but even if you aren't at that point, you can still contribute to forecasting.
    h. Found that people have used the 3 terms in different ways. Some fields don't use "forecasting", but only use "prediction", etc.
    i. This gives the ecological forecasting community flexibility in what works for us to use the terms as we want. Want to find out what works for the group as a whole.
    j. From looking at the results here is a summary of what Anna sees for the definitions.
        i. Prediction is the most nebulous. It was often described as having a guess that doesn't use data.
        ii. Forecasting is making a guess and then you quantify the uncertainty around that guess. Prediction does not include uncertainty (so may not include data, or may be used data to make prediction, but don't quantify the uncertainty).
        iii. Projection - given a certain scenario what will happen (this is the most universally accepted definition that projection gives a scenario).
    k. Would forecast give an extrapolation? If a forecast is making a guess where we don't have data, then we are trying to extrapolate outside what we have trained the model.
    l. Quantification of uncertainty. Extrapolation means integration of a present condition.

m. Was there a split between people who thought of quantitative outcomes vs qualitative outcomes? There hasn't been too much of that in the survey. It wasn't an overwhelming trend, but did come up a few times.
    i. Considering across economics or weather - will there be a recession or will there be rain.
    ii. Examples/outcomes are also given by binning quantities so seems more qualitative
    iii. User or a provider of the forecast will change what people think of the definitions
n. Peter: Forecast - subset of prediction. Fit stats model and you predict the response. But that is part of your training response. Mike: does the opposite. Prediction and projection are subsets within forecasts.
    i. Fit standard regression model with predictions for response variable. Prediction because it isn't scenario driven and you can't interpret in a non-conditional way. They are predictions, they don't need to be predictions into the future, but can be predictions.
    ii. Nowcast vs hindcast
        1. Hindcast would also be a good one to define. Some people treat hindcasts as forecasts. Others use hindcasts, but it is fitting the model to the data
        2. In NOAA a weather prediction - taking the current observations and throwing them into numerical/computer models to forecast the models to get predictions. Forecast and prediction are linked which is opposite of what Mike said. Forecast is to obtain the prediction. Is a projection a forecast? It is another term for forecasting.
o. Once Anna/Gretchen have figures to share it would be good to get input from the group.
p. Glenda: Is the common use the best outcome? Prediction is the bigger set, subset of forecasts, and subset of projection. If this body of practice set up a consistent way to use these terms, then we could use what the survey says or not. But it will be helpful for thinking about communicating. Glenda's example: Nissan → Nissan-Datsun → Datsun. There was a game plan on how to change that terminology. Thinking about our EFI communication plan could be good to think about.
q. Thinking about criteria for how we define the terms may be more important than the label/term used. E.g., probabilistic statement, prediction to the future, using data not used in the prediction

2. Getting the whole group caught up. SEE NOTES INTERSPERSED BELOW IN THE HOMEWORK SECTION FOR THE FOLLOWING TWO AGENDA ITEMS.
    a. Go over the homework
    b. Does anyone who looked over the background material have any questions/comments?

Homework:

1. **People new to the Theory group:** Go through the background material previously created by the group. (a) is the newest document, (c) is the oldest document, (d) is Peter's slides for the RCN presentation. In this material, are there hypotheses that are not showing up? Things you agree or disagree with?
    a. Brainstorm of the top 3 things people want to communicate about the forecasting hypotheses we have been discussing in a manuscript: Theory Group: Top 3 Things You Learned & Top 3 Things to Communicate
    b. Look at the Common Framework Slides
    c. Hypotheses
    d. Peter's Slides from his RCN Presentation: https://drive.google.com/open?id=1gHcT2e8VEY5eVm_KW6UAgKKOPbAo4aFB
    e. Winslow: Lots of material and good ideas for thinking through how processes operate through temporal and spatial scales. This is a new angle to think about where we can make predictions across those scales.
        i. Because there is alot of literature on the temporal and spatial scales, we have circled around to finding where we are making a new contribution instead of rehashing what has been done and adding the "forecasting" terminology on it.
        ii. What are the steps to actually doing the forecasts is valuable.
    f. Hassan - problem we are facing is how to focus on the framework that can be standardized across ecological forecasting. This hasn't been addressed by many communities. To be able to do forecasts - here are the blocks/steps to get there. Dealing with different environments/systems - e.g., weather forecasts were standardized across environments to be able to make comparisons. In ecological forecasting - push to do more forecasts so we can have multiple forecasts to allow for standardization across systems
    g. Mike participated in the National Academy Earth Prediction workshop - ecological forecasting is thinking more broadly about forecasts than that community which is focused on the boundary and initial condition problem - there are other limits than just those two.
        i. Mechanistic modeling is needed. Data access and basic ecology.
            1. For the mechanistic models there is more going on with the predictability and classes of forecasting than just the boundary and initial condition problems. Many systems are influenced by their heterogeneity and variability. This is the challenge.
                a. Example of boundary condition problem. Climate forecasts the initial conditions no longer matter. If we don't have anthropogenic emissions it would be orbital influences or volcanic emissions. Uncoupled (outside of the feedbacks

within earth system model). Boundary conditions are drivers. Drivers are feeding back to the prediction because the ecology changes. But most predictions are offline. Most of the drivers don't change given what we predict with the forecast.

b. Fisheries example - Including harvest into models. It isn't influenced by climate, but is responding much more to harvest.
c. There was an RFF (Resources for the Future) talk at the EFI 2019 meeting that talked about a 3 tiered response to incorporating human behavior: predicting ecosystem, predicting ecoystem and how people respond, predicting ecosystem and how policy responds. This was Jaime's talk which you can see here:
    i. https://www.ashander.info/posts/2019/05/ecological-forecasting-integrated-ses/ has a brief abstract and link to youtube: https://www.youtube.com/watch?v=SvzjnB18sAk

h. From the applied work of NOAA - Take observations and model them and see what comes next. At the Agency level - take what has been developed in academia and transition them to be operational. Readiness level (NOAA uses the same system that NASA uses) indicates when forecasts are ready to be operational.
    i. A lot of the pipeline for basic science to application is built within the discipline interested in the specific topic (e.g., weather, fisheries, etc). But it isn't necessarily broadly applicable across systems. Ecological forecasting is in between the numerical and statistical models.
i. Temporal and spatial processes - what makes sense to put in forecasts. These theoretical concepts are promising. And will advance how we do forecasts. If we can understand the type of prediction problems then we can learn from other forecasts and apply to new systems.
    i. Theoretical insight from Lorenz put the weather forecasts in place to put lots of money into monitoring to be able to monitor for that type of prediction problem.
    ii. Ecology has a bunch of traditions of modeling. They are not all well-suited for forecasting.

2. Vocabulary

a. Take Anna Sjodin and Gretchen Stokes 5-10 minutes survey about forecasting definitions: https://uidaho.co1.qualtrics.com/jfe/form/SV_4PIltUH8Ne9xRcx
b. Add definitions to the Vocab Google Doc
    i. The following are notes from the June 23 call:
    ii. Null models - the thing we need to beat in RCN competition to

1. Null model = basis for comparison. Question for the NEON Forecasting Challenge is what is the specific null model for each of the Challenges.
2. May not be vocab discussion of the definition of what a null model is, but it may instead be about agreeing on what the appropriate null models are across the 5 Challenges
3. Stationary vs non-stationary. From the Vocab Doc - B1 (persistence) and B2 (historical means) are the null models used in practice. The persistence model is where uncertainty = random walk. Historical means - this is what is shown in the forecast limit figure. There is an upper bound that you are not doing better than the stats you have seen in the past.
4. There are interesting questions on short times scales about what people should be using. If you have a forecast on a diurnal cycle the persistence model does bad quickly. But if you are predicting noon tomorrow using info from noon today rather than using 5 pm today, the prediction will do well. A persistence forecast is different from historical means which would say what do I see on June 24 that I see for every other year on June 24 I have data for?
5. How simple does a model have to get before it is null vs a simple process model?
6. Broader definition for null model - what lets us detect change in the system. If a forecast is about describing or predicting a system, the null is the no-change state for the hypothesis of the thing that causes change.
    a. What is randomness or chance?
        i. In one sense the null is what is the random expectation accounting for the processes
    b. The null lets us know the skill of the forecast. Do you need to predict change in the system? At what temporal scale do we need to detect change?
    c. The type of null could be guided by the type of forecasting problem. Nice to have guidance to know what kind of null is needed for what kind of forecast.
    d. Would a null model have to describe the historical variability or uncertainty in the system. Random observation error. The historical means would capture the past uncertainty within that state.
        i. In the weather example, the forecast captures the diurnal cycle better than the seasonal cycle. But historical means looks at the weights.
    e. How many time scales are you allowed to look at? If just doing persistence and historical means - one goes up square root of time/distance the other is flat they hit each

other at some point. If you have multiple tiem scales you might have multiple change points at different scales.

i. Flux forecast example that Mike has been looking at recently - has 3 nulls. Persistence (last half hour is same as the next half hour; 1 day lag - 5 pm today is same as 5pm yesterday; historical mean) Found that knowing yesterday doesn't do much better than the historical mean for very long
ii. RCN Challenge - NEON fully operational for only a year. If you have the simplest possible model and you feed historical drivers as alternatives when you don't have historical time series, what would the model predict if given historical mean climate rather than the actual climate?
iii. For practical purposes imagine having decision tree to choose null for questions on spreadsheet (se the spreadsheets linked below for each forecasting challenge)
iv. Null model is chosen as a way to test skill. Depending on how mature your model is, you will choose different benchmarks. If you don't have enough data to evaluate diurnal or seasonal trends then whatever you have for the mean is your null. Once you can make those diurnal or seasonal trends then you have that. Then add on top of that. Peter gave an example of his forecasts looking at density dependence.

7. For the NEON Forecasting Challenge - will people generate the null model so everyone will use the same null model for evaluating? That is the intent - so is important to agree on this so we are equitable across different challenges. But won't be totally equable because of the amount of data (aquatics = 5 minute readings, beetles = 6 sampling points per year).
    a. Beauty of standard null will be able to compare across challenge topics. But may not be able to do that across all forecast challenge topics.
    b. Can quantify skill relative to observations. Calibrate as we go

3. Practice forecast comparisons

a. This will be homework for the next call

i. We may want to add a row: "appropriate null"

b. RCN NEON Forecast Challenge Forecasts. We have created Google sheets linked to each of the following forecast topics with questions 3b-3j. Go through the topics you are most familiar with to fill out these questions. For question 3j the format in the Google sheet is slightly different. For each type of uncertainty, enter whether that uncertainty dominates in the first 1/3, middle 1/3, or last 1/3 of the forecast horizon. The common frameworks slides are listed below question 3j for you to reference.

i. Leaf or mosquito phenology <- We (organizing committee) decided on starting with leaf phenology (specifically forecasting PhenoCam observations) for phase 1
ii. Carbon and water fluxes (eddy flux) between land and atmosphere
iii. Aquatic - chlorophyll a
iv. Ticks - abundance. Timing of the peak or abundance through time as observations come in
v. Total (Beetle or macroinvert) abundance (all species)
    1. Still being sorted out

c. For each of the RCN forecasts, what would the units on the x axis be in the following figure?

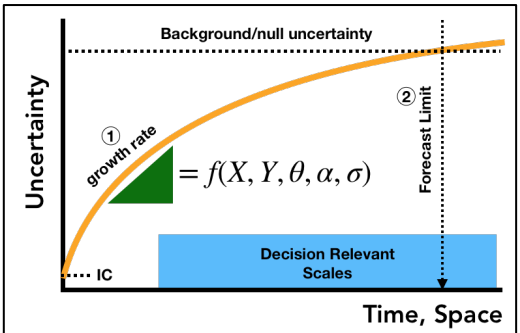


Figure 3: The predictability of a forecast is measured by the rate at which forecast uncertainty grows, in space or time, and the limit at which the forecast performs

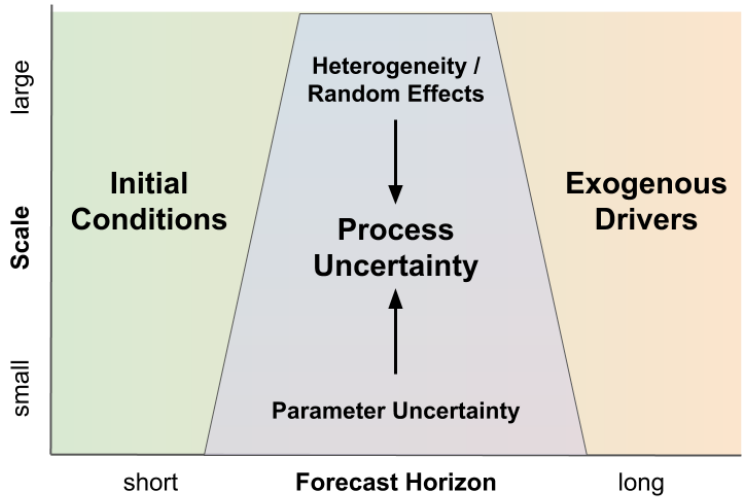
d. What is the level of organization being forecast (organ/physiology, individual, population, community, ecosystem)?
e. What is the phylogenetic scale of the forecast (if applicable)?
f. What is the trophic scale of the forecast (if applicable)?
g. What would you say is/are the relevant timescale(s) of the process itself?
h. What do you think is the forecast horizon (time till the prediction is doing no better than chance) for this system at the NEON plot/sensor spatial scale?
i. What would you say is the relevant spatial scale of the process itself?
j. How would you describe the spatial scale of the forecast (relative to the process itself)

k. If we divide the time between t=0 and the forecast horizon into 1/3s, what input uncertainty do you think dominates the forecast uncertainty at each point in time?

uncertainty	First 1/3	Middle 1/3	Last 1/3
Initial conditions			
drivers			
parameters			
Random effects			
Process error			

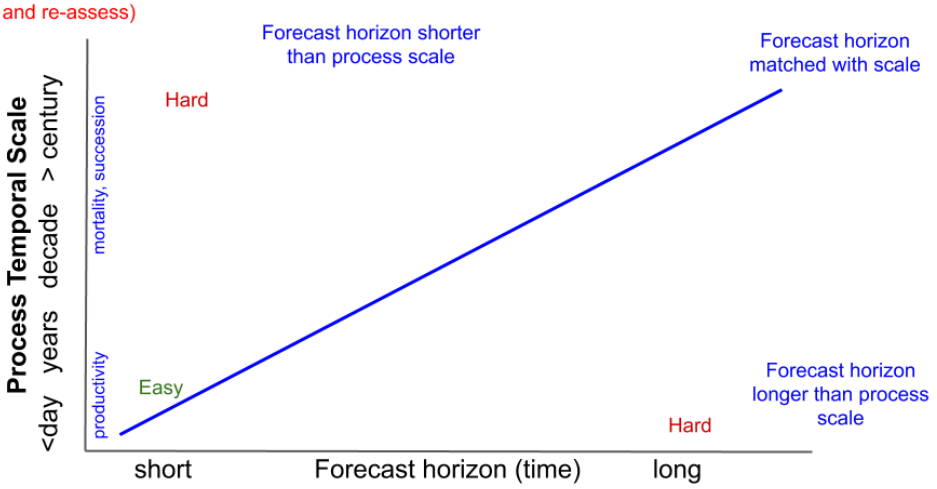
Common Framework slides for reference:

l. Slide 8

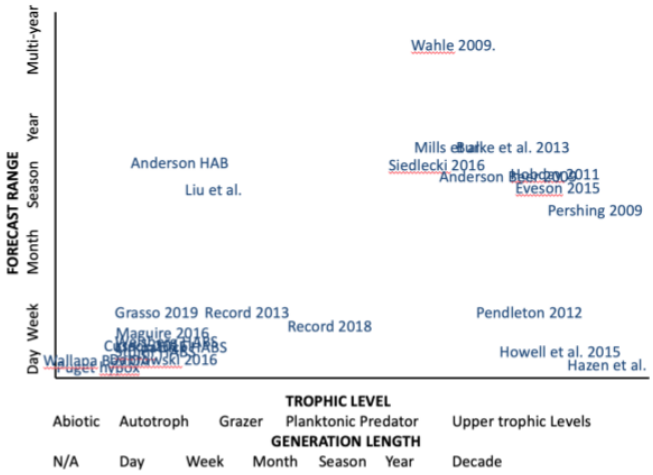


m. Slide 9

Rollinson Note: Are temporal scaling issues in forest modeling are separate from levels of organization or not. Can the temporal scale be independent of the forecast horizon? (I did this before seeing the Adler pre-print; will read and re-assess)



n. Slide 10



This is from a literature review I've been working on for marine ecological forecasts. Just from my notes-- I've just eyeballed the positions on this graph. Placing it here as food for thought. -Nick