August 18, 2020 Theory Working Group Call

Attendees: Abby Lewis, Glenda Wardle, Amanda Gallinat, John Foster, Mike Dietze, Jody Peters

Agenda and Notes:

1. New co-chairs.
    a. Abby Lewis, Amanda Gallinat

2. Jody will send out a poll to find a new time for calls for this fall semester

3. Continue to discuss Forecasting Vocab Terms
    a. Background: theory group has been meeting for over a year, but we realized we were talking past each other because there was lack of consensus
        i. We have discussed Null Models, Forecast Horizon vs Forecast Extent, Forecast scale, Uncertainty
    b. On this call pick up with Transferability
        i. How different does a site need to be
        ii. Transferability vs generalizability
        iii. Prediction to a new location vs prediction about a different species
        iv. It seems less ambiguous when thinking about it phylogenetically. Example - you make a forecast for 4 genera. Can I take those 4 models and extrapolate to a 5th species?
        v. Another example - If you make a forecast for one lake and move it to another lake - going from one to one seems like transferability. But if you are predicting hundreds of lakes and you have a model that is calibrated across all of them is it still transferability if you have hundreds of units and unique parameters for each individual lake?
        vi. Does it apply across systems? Having site as a random effect would be one way to make a model transferable. If you want a model to apply to as many systems as possible, then maybe transferability is a really broad definition.
        vii. Does it need to be applied to a totally out of sample system?
        viii. Allometric relationship Dietze paper example - global allometry and random species effects for phylogenetic differences. Why are some taxa taller/wider/heavier than others? You can put in higher level covariates. Even though you have higher level covariates in the system - there is still the question about transferability. What can move models from one location to another?
        ix. What are we going to put this term to use/to do?
            1. Is there a tradeoff between the model that was set up for a certain system and transferring it to a new system
        x. How much do you tune your parameters? How much can you alter it? How much can your random effects change what is going on?

xi. If you re-calibrate the model, you ask about the transferability of the structure, but not the transferability of the parameters

xii. Has there been tests of the transferability of models?
   1. There are papers where people fit process-based models in one system and tested at new sites and assessed how the model is doing in general. But were any of those couched in a formal metric of the transferability?
   2. 0 to 1 ranking - 1 this model does just as well at a new site. 0 this model does no better than random at this site.
   3. Could ask this about structure and parameters? How well did the sample perform?

xiii. Physics is an example of near transferability. The laws of Physics apply everywhere.

xiv. Another context/example - camera trap dataset and machine learning. It isn't forecasting, but is becoming more relevant to how specific you want it to be in S. Africa vs when you apply the same algorithms to Australia.
   1. People report the skill of the model
   2. There are examples in species distribution

xv. Why is transferability interesting?
   1. If models are making some attempt to capture processes, it is central to the understanding of how general ecological processes are. If models transfer across systems, then we understand what is going on well.
   2. Main goal of transferability is to answer questions about how well we understand the general patterns across systems
   3. What is the value of forecasting to ecology?
      a. What keeps coming up is that the value is that forecasting gives us a sense of how predictable systems can be
      b. If the kernel of that is within transferability.
      c. Is it transferable within a certain scale and when you jump scales it is no longer transferable?
      d. Are there processes operating in different biomes?
      e. Can we know everything we need to put into a model? If we can link processes in our models that are not transferable then we can learn a lot.
      f. The expectation is that if we can understand the underlying system, even if there isn't a true set of parameters, if we understand the drivers in the uncertainty, then we will know more about the system

xvi. Keep coming back to parameters and keep coming back to structure

xvii. Transferability of the dominant uncertainties

xviii. Can we understand the patterns of uncertainties, of model parameters, and of model structures?

1. If you could put large classes of processes in large classes of models that have high levels of hierarchy to them
2. Example - a lot of population modeling is done with matrix models. There is a structure we keep coming back to. A lot of the same covariates come back over and over (temp, precip, primary productivity). The same predictors keep coming back.
3. What is the probability that certain structures are present in certain classes of models?
4. What is the temp effect in this set of models for this set of species? Use hierarchical model to predict this

xix. Glenda: Buckley paper (see citation below) on temporal and spatial variability in plant models
1. They accumulated hundreds of models, but found it hard to test the model structure
2. Need more models to pull out the generality.
3. There is a lot of small scale variation that is going on
4. Comadre/Compadre matrix database examples

xx. Want to make sure to store the right x,y, and z variables to be able to synthesize across forecasts
1. Lessons learned from demography - most people didn't provide specific enough location information where the data came from. So had to go back to each of the studies to get this information. Would have been better to get it up front.
2. RCN is beta test of the synthetic analyses. Can we synthesize across models or is there more information we need (or need in a different format) to analyze across forecasts?
3. Want to keep the metadata collection/analyses straightforward so when models change, you still have the information you need to make comparisons
4. How would you model a tortoise, a tree, a mosquito? Need the biology behind it and then you can run the models

xxi. RCN Forecasts -
1. Will transferability be something we want to examine across the NEON Forecast systems? And is there a way to compare?
2. Make comparisons across all tick forecasts, or across the phenology forecasts?
3. Need theory group to go from lofty ideas to coming up with something actionable
4. Conceptually there are things that have emerged
    a. The Near Term forecasting in Mike's lab - first step of synthesis, but has been on too small of scale to answer the transferability questions
        i. There are phenology and eruptive behavior across a number of systems. E.g., tick forecasts, algal

blooms - they have leaned on Kathryn's vegetation phenology forecast

5. Transferability and spatial forecastability - how are they different from each other and can we measure them? If there is a forecast in space vs measuring the transferability of the model - how is that different? How is the test of transferability different from an out of sample prediction? It is a gray line.
    a. Comes back to transferability of model structure and transferability of parameters
    b. In pure spatial model have assumption of no trend and isotropy
    c. Need to put process into why the mean is changing or why directionality matters or why parameters are not the same everywhere
    d. Transferability needs to be broken into different parts

xxii. Transferability in time? Do the same principles apply? Predicting out of sample vs transferability? Does it work the same way?
1. Think it does.
2. Is a model sufficiently general so that it can capture temporal variability?
3. Example when thinking about the projections out to 2100 - think about how to put that capacity for the underlying model parameter and structure to drift as you go further out in time.

xxiii. When thinking about Parameters - Is it about the values or the number of parameters?
1. Glenda's example - Colin Prentice 2011 paper looked at delta 13 of plants along aridity gradient in China. Glenda retested in Australia and did not find the same universal gradient. Good lesson of how transferability isn't as standard as you would think it would be. So Glenda is a bit nervous about thinking about transferability
2. But it is still worthwhile in that if you repeated the test/experiment on enough large gradients in the world, you might be able to tease out the variability that you see. How much is the underlying biology vs other factors?
    a. In Mike's bacteria forecasts example - differences are seen in different lab protocols. The community hasn't come up with a standard way of measuring the same thing. Looking at the lab effect
    b. Sample to lab transfer effect
    c. Glenda's Long Term Ecological Network example from Australia - long term = 10 years. But it takes about 10 years to start to get the sampling/protocols right

xxiv. Model transferability vs forecast transferability. Forecast horizon and forecast transferability.
1. Could think about spatial forecast horizon as one more way to measure spatial transferability. How far out in space does the model work?
2. May be more applicable to space than time
3. Within season - forecast 16 days out, loose predictive power, but when you reach day 16 you had to accommodate that you were accumulating uncertainties

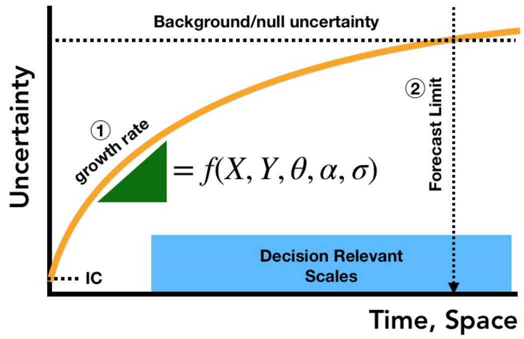
xxv. Note from Glenda with citations mentioned today
1. Here are the demography refs I referred to in our discussion today.
2. Buckley Y.M., Ramula S., Blomberg S., Burns J., Crone E., Ehrlén J., Pichancourt J.B., Quested H., and Wardle G.M. (2010) Causes and consequences of variation in plant population growth rate: a synthesis of matrix population models in a phylogenetic context. *Ecology Letters* 13(9): 1182-1197. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1461-0248.2010.01506.x
3. Burns J.H., Blomberg S.P., Crone E.E., Ehrlén J., Knight T.M., Pichancourt J.B., Ramula S., Wardle G.M. and Buckley Y.M. (2010) Empirical tests of life-history evolution theory using phylogenetic analysis of plant demography. *Journal of Ecology* 98: 334-344.
4. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.1365-2745.2009.01634.x
5. These use a database that existed before Compadre was launched.
6. Link to Compadre here
7. And the paper:
8. https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2745.12334
9. Time series data in MPMs: The mode of duration of studies in compadre 3.0 is 4 years , corresponding to the length of an average PhD project, as well as that of most funding agencies.
10. Spatial data in MPMs: The mode of number of populations studied per publication in compadre is one.

Agenda for Next Call

- Look at the Phenology example next time as a group. Do it with the consideration of the definitions from the vocab. Think about how to apply those definitions to that specific case and how we would assess them. As a Theory Group what would we want people to report on (even conceptually)? What would we want to test in the phenology example.

- **The motivation for the Phenology example was thinking through questions about scale and uncertainties that had been bouncing around in earlier discussions, things that Peter brought up in his talk, and the conceptual slides.**
    - **Can we apply this a priori to any of the NEON Forecast Challenge areas**
    - **Can we make a prediction for what we expect to see for these scale and uncertainty questions from the conceptual figures and do a community of forecasters agree on any of it**
    - **Do this exercise and then go back to the Conceptual figures to see if we are still on the same page.**
    - **And will be useful predictions for the Challenges**

4. Practice forecast comparisons using the RCN NEON Forecasting Challenge Topics
    a. Before call - work through the Phenology example
    b. On the call discuss to see if there is any confusion
    c. For future calls - do the same thing for the other systems linked below.
    d. 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 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 water temp and DO
        iv. Tick abundance. Timing of the peak or abundance through time as observations come in
        v. Beetle communities
    b. Questions (see References figures/slides starting on pg 3 for context):
        i. For each of the RCN forecasts, what would the units on the x axis be in the following figure?



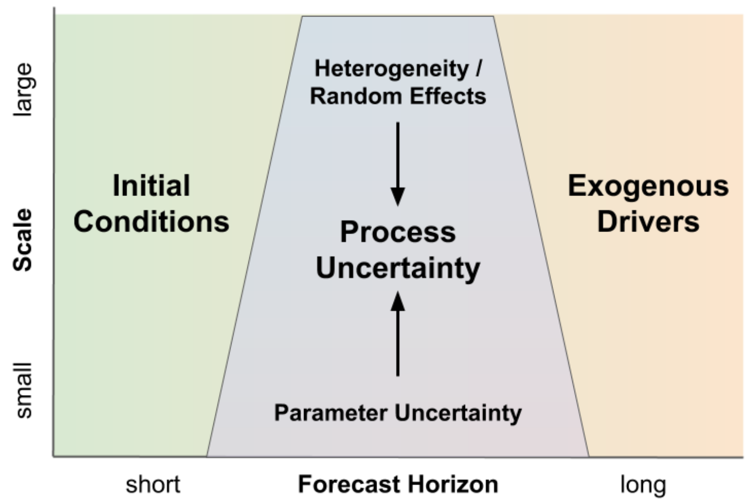
c. What is the level of organization being forecast (organ/physiology, individual, population, community, ecosystem)?
d. What is the phylogenetic scale of the forecast (if applicable)?
e. What is the trophic scale of the forecast (if applicable)?
f. What would you say is/are the relevant timescale(s) of the process itself?
g. 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?
h. What would you say is the relevant spatial scale of the process itself?
i. How would you describe the spatial scale of the forecast (relative to the process itself)

j. 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			

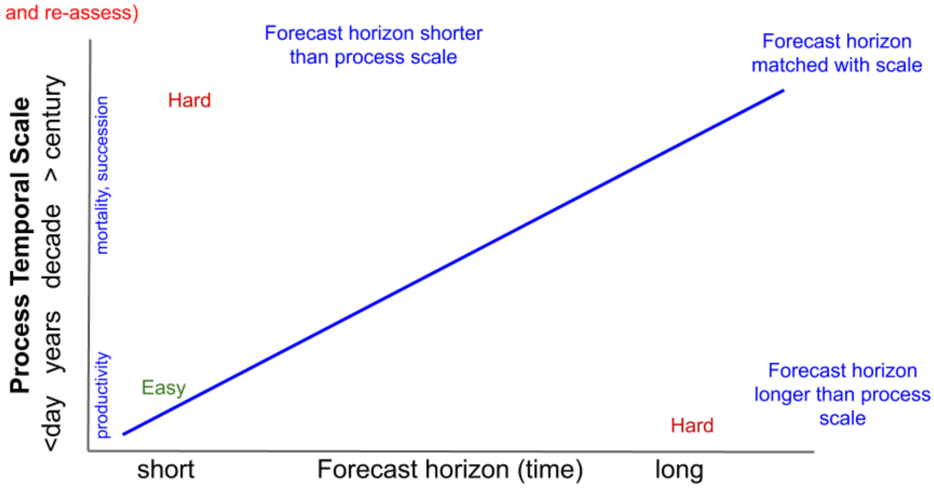
Common Framework slides for reference:

k. Slide 8

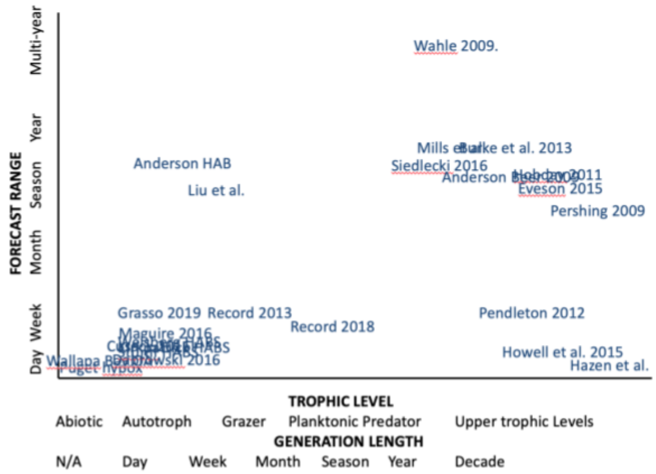


l. Slide 9

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)



m. 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

n. Uncertainty components in forecasts

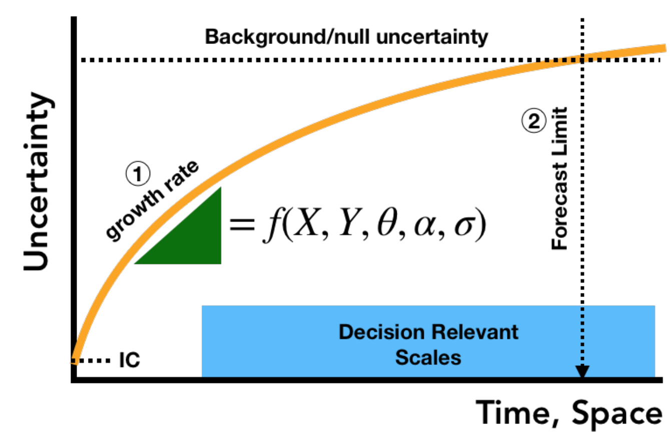


Figure 3: The predictability of a forecast is measured by the rate at which forecast uncertainty grows, in space or time, and the limit at which the forecast performs no better than chance. IC = initial conditions, X = exogenous drivers, Y = internal system state, θ = parameters, □ = random effect variability, □ = process