February 20, 2020 Theory Working Group Call

Attendees: Mike Dietze, Christy Rollinson, Adam Clark, Amanda Gallinat, Peter Adler, Jody Peters

Agenda:

1. Manuscript Ideas:
    - **a.** Hypotheses for predictability in ecological forecasting
        - **i. Homework from Dec 2019 call: Revisit the Google doc and think of analyses/methods to test the hypotheses now**
            1. **Self-assign sections on the Google doc that folks are most interested in. If there is something that everyone is interested in we can focus on that, or if it is more distributed than can do something broader.**

    b. This may be more than 1 paper. Each category could be a different kind of paper
    c. Not expected to answer hypotheses immediately, but could give the state of the science on where we are on the hypotheses and what needs to be done to answer the hypotheses.
    d. Think about the roles different individuals are willing to put into this
    e. Define how the paper will move forward and the interests/commitments
    f. Could do lit review or do quantitative analyses/simulations
    g. Simulations could let us get going now. And could help with distributing the work
    h. What is the road map for what we want to achieve? If Owen sorted everything out in his paper 5 years ago and we are done, then great. But not sure it is all sorted out. What is left to do? How do we get there?
    i. This paper could set out the conceptual figures that then later get tested with real forecasts
    j. 1st step - what do we mean by predictability and how do we test that.
    k. Would it be possible to start testing hypotheses with the datasets we have in hand? Compadre/Comadre? Preliminary analyses? Limitations of those analyses? Something concrete we can get our teeth into?
2. Phone call Notes
    a. 2 goals
        i. 1) Self-assign
        ii. 2) Are we pitching as a conceptual paper. Here are ideas in the theory of forecasting. OR are there analyses that can be done with existing datasets to test these hypotheses or simulations?
            1. If we go the analyses/simulations route, we'll need to have multiple manuscripts
        iii. Who is the audience?
            1. Forecasters who are also thinking about theory? OR
            2. Non-forecasters/early-stage forecasters to show how it fits into work already being done? Mike leans towards this one

        a. Forecasting community is tiny
        b. Predictability in ecology is the way forecasting is cross-cutting in ecology
        c. Pitch to ecologists broadly. Don't get down in the weeds of mathematical theory that would exclude the broader community. But don't want to produce the "50 questions in aquatic ecology (or other discipline) papers" that comes out every so often
    3. Christy - excited about. The list of hypotheses helps to answer fundamental ecological questions. This is the unique niche for this paper. Forecasting is fundamentally about theory and classic questions that we have yet to resolve.
    4. Potential title: The Role of Forecasting in Addressing Fundamental Questions in Ecology
    5. Shoot for high impact ecology journal: Frontiers, Ecological Letters
b. Adam Clark intro (this is his first call). He was invited to join by Peter after they had a discussion around the Detto et al paper. Adam is trying to separate signal variability from noise variability in time series data. Make sense to talk about randomness in forecast in terms of true randomness or model error. Adam's goal is to sit in for a couple of calls to figure out what is happening in the group.
c. Larger effort - EFI is about building community of practice with focus on iterative forecasts where forecasts happen on a timescale where you can see if your predictions are correct or not
d. Forecasting provides a win-win. We can tackle applied problems useful to society, but can also answer fundamental questions about the predictability of systems.
e. Previous/on-going work to develop standards for forecasts is so that we can compare predictability across forecasts/ecosystems
f. From Hypotheses List which topics resonated with the group on the call
    i. Which hypotheses/topics are folks on this call interested in and then find out the interest of those not on the call.
    ii. For next call -
        1. get a rough outline for each subsection and
        2. a bit more depth about questions that we want to present and
        3. think about how would we test them and
        4. is there low hanging fruit that we could test now with existing data without having to write a new proposal?
    iii. Amanda: This Figure 3 at the top of the Hypotheses list - provides an opportunity to make this figure clear and relevant for the hypotheses outlined

iv. Christy: Common thread between hypotheses - connecting big or small scale phenomena that can be measured/managed. A lot of the hypotheses are about scaling in different ways

v. Hypotheses are oriented around the figure each in some way

vi. Define predictability as the scales over space and time where we are doing better than chance. As we move out in space/time/phylogeny - things that are less similar make the uncertainty increase. If that limit is a key concept (this idea is in the Petchey 2015 paper as well), then we get a lot of traction understanding the growth rate. The slower the growth rate the longer the forecast limit

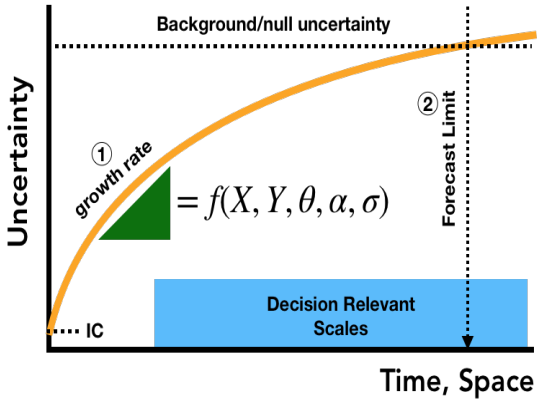

Figure 3: The predictability of a forecast is measured by the rate at which forecast uncertainty grows, in space or time, and the limit at which the forecast performs no better than chance. IC = initial conditions, X = exogenous drivers, Y = internal system state, θ = parameters, □ = random effect variability,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

vii. In ecology we don't know which uncertainty dominates in which ecosystems. What uncertainty dominates in which situations and what type of conditions contribute to those uncertainties (phylogeny, life history, etc)

viii. When we bring in phylogeny it is easy to get population focused. But there are also a nice representation of biogeochemical hypotheses as well. Perhaps it all comes down to not understanding microbes. But even there the pools and fluxes would have some kind of hypotheses about what those.

ix. Questions about Scale, Complexity. How predictability changes with model complexity. Where can you get by with a simple model, where do you need a complex model?

x. Questions about Transferability. If you have 6 congeneric species and have good data on 5 of them, then can parameterize demographic models for those 5 can you predict what covariance will matter for species #6, and what the parameters will be? Frame the transferability in hierarchical models. How variable are model structures hierarchically across systems.

    1. Christy is very interested in this. Common vs rare species. When we can't get data for the rare species, how can we borrow the strength from the information we have for the common species to make predictions for the rare species

2. Can we predict the predictability of different species or processes from those they are related to and at what scales. Mike has started to do some of this with their microbe work.

xi. Amanda - set up the questions with Figure 3 as the structure/framework. E.g. for each section -Which parameters tend to drive uncertainty? Which are the hardest to estimate? -What are the primary limits to model transferability? -How does predictability vary with scale? With system complexity?

xii. Initial paper - lay out hypotheses that are pitched to very broad audience. **The Role of Forecasting in Addressing Fundamental Questions in Ecology**. Make the point that there is a theoretical side to forecasting that tackles grand challenge questions in our discipline.

xiii. Adam: Uncertainty is propagated. For the time-series work he has done, break into 3 categories.
   a. 1) Not getting the starting point right (chaotic system).
   b. 2) Getting the model wrong
   c. 3) Something related to stochasticity in the system
2. A distinction like this could be helpful for a theoretical paper. But it moves away really quickly from something that is testable with real-world data
3. In Figure 3 #1(starting point is incorrect) is X, and #2 and 3 (model is wrong/stochasticity in the system) is sigma

xiv. For example, if we nail down 3 hypotheses that we want to test. Then there would be 3 yellow curves that could be compared. At what rate does uncertainty increase for these different systems/forecasts

xv. Then what are the factors that contribute to the shape of that curve. Some are more reducible then others. Model misspecification is reducible and stochasticity is not

xvi. Peter: What if we constrain ourselves by thinking about case studies that we can make figure 3 and make them to compare the curves?
1. Mike: Right now it is pretty limited. There are a small, but limited papers. The RCN's goal is run forecasting challenge to generate uncertainty partitioning so we can do this, but that challenge won't happen for a couple of years. So before we get to that point it may be useful to lay out these questions and perspectives for a broader audience even if we can't answer the questions yet.

xvii. Christy: case studies can limit how people think about connecting to other systems (e.g., fish case study is hard to apply to savanna system)
1. Last section of paper- lay out the theory/hypotheses at the beginning/middle of paper, then in the last section we show the next step. Here is how we think about comparing. This will lead to getting into the operations, but won't be the full focus of the paper

xviii. Mike: not opposed to testing hypotheses, but one way to move forward is to go through the hypotheses adn ask how do we test this hypothesis and if there are low-hanging fruit that is easy to test, then definitely do it

xix. Peter: Come up with categories that we can lump hypotheses. Find some way to organize the different kind of hypotheses. The challenge is to do the work. Easy to list hypotheses off the top of our head. But need to dig into the literature and see what else relates to this or play around with toy models.

xx. Next step?
    1. Dig into subsections and do lit review. And define conceptual figures to test hypotheses. Perhaps between lit review and how to test it can come to quick conclusion on what we can write a paper on based on meta-analyses or simulation. And show here are the hard questions and here are the easy questions

xxi. If someone was teaching a grad seminar - divide up these sections for grad students to tackle. There are a bunch of folks who teach forecasting classes. Can divvy up work. Peters's is a year from fall. Mike's is next spring

xxii. Grad student working group may have students interested in participating on this

xxiii. Adam: has seen papers that test questions with 6 methods and they found that 1 method works best. For this paper focus on methods that are well understood in each group and make the case that the methods were applied with best practices to the data/field

xxiv. Taxonomy/Phylogeny - can Amanda/Will dive into this more to flesh this section out?
    1. They can dig into hypotheses that are there. Then if there is specific framework that will contribute to drawing all sections together will save time in the end

xxv. Amanda is worried that folks will fill in the hypotheses that are most relevant and then the group will come to the realization that to synthesize across categories, we need to figure out how those hypotheses relate to scaling/transferability, etc
    1. If we can take some clear guidelines when looking for the hypotheses that are being filled in, that can help direct their search.

xxvi. Another way to look at it is to say: We have identified a few key concepts: Uncertainty, Uncertainty partitioning, Scale, Complexity, Transferability. See as you dive into the theoretical questions for each category - where do the existing questions of predictability fit into those key concepts.

xxvii. Scale/Geography/Temporal, patterns in variability - Christy

xxviii. Peter - work on Google doc to organize and identify people who commented on to reach out to see if they would help flesh out specific sections