October 7, 2024 Theory Working Group Call

Attendees: Freya Olsson, Jody Peters, Bilgecan Sen, Cole Brookson, Abby Lewis, Saeed Shafiei Sabet, Jonathan Borrelli
Regrets: Shubhi Sharma, Caleb Robbins

Agenda:

1. [ESIIL](#) working group proposal update (Cole, Shubhi, Abby, Caleb)

2. Paper discussion: synthesizing predictability
   a. https://www.nature.com/articles/s41559-019-1052-6
   b. While the paper focuses on animal populations in general, the goal of sharing this paper is not to discuss the ecological conclusions. Rather, the authors have made very explicit choices on how to measure predictability (both intrinsic and realized) which could be helpful when conceptualizing the proposal
   c. Similar to the Pennekamp paper (only published a year apart). Introduced weighted permutation entropy to an ecological audience
      i. Pennekamp paper:
         https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecm.1359
   d. This paper does what the Pennekamp paper does in a certain way
      i. Use empirical dynamical modeling - lets you look at nonlinearity.
   e. This paper doesn't look at permutation entropy and doesn't cite Pennekamp, but they are doing something very similar.
   f. Since Pennekamp was not cited, this is a good example of why there needs to be a synthesis paper to show what has been done broadly
   g. Looks like chaos is common and when do we see chaos
   h. Dimensionality of timeseries is what makes predictability interesting. We can add stochasticity to nonlinearity and dimensionality.
   i. Memory of time series - is used outside of ecology
   j. 4 characteristics
      i. Dimensionality, nonlinearity, stochasticity, and memory are all part of intrinsic predictability
   k. This paper, they used MAE instead of RMSE or CRPS
   l. Forecast horizon - they set as 1 year
   m. This is a similar analysis to what has been done with the NEON Forecast Challenge - we aren't doing life history traits so it is sort of harder to think about what axes you think predictability will be influenced by
   n. In paper animals with faster life histories were more predictable Thought that longer lived organisms would be easier to predict. But that wasn't the case.
   o. The paper is working with a different definition of predictability
   p. What does it mean to be "more predictable" - depends on the definition
   q. Chaotic dynamics are more predictable than a stochastic system.

r. For population anslyses for people interested in mean state not year to year variability then they will say the system is predictable.

s. How much is just a data problem for the life history question. Think there is an interesting question of what you are seeing is an artifact vs something inherent

t. Seen a lot of strong opinions of EDM as a form of prediction in ecology
  i. The link to other disciplines that use something similar is weaker compared to things like weighted permutation entropy. The S mapping links to other examples are few. Some in applied math. Haven't seen as many physicists or climate scientists use this approach

u. Steven Munch gave one of the EFI2024 conference keynotes on EDMs and Caleb has connected with Steve about the potential to get EDM models included add to the Forecasting Challenge submissions. Ethan Deyle at Boston University also used EDM models and may also be able to help with this.

v. Curious on how EDM holds up when you have the process worked out. The model isn't statistical - it is Newtonian dynamics where you know the process/laws

w. Future paper to consider discussing - paper Bilgecan mentioned with EDM
  i. https://www.pnas.org/doi/pdf/10.1073/pnas.1417063112

x. EDM - why hasn't someone applied it to Covid? WIth short term expect you could get some good results
  i. Think there are standard toolbox of items in disease world that are ML methods
  ii. If compared to process based models vs prediction in a model free framework - haven't seen EDM compared to the ML methods. Would it do just as well or better
    1. Haven't seen this, but think probably someone has done this
  iii. EDM is very dependent on time series. Not sure if it can make predictions across space.
    1. For Covid could make predictions for a specific county, but then can't say how it can spread across the country
    2. Not sure how it could be used in places with no prior data

y. Questions to think through
  i. How does uncertainty tie to intrinsic and realized predictability - this isn't really considered in this paper other than a few small points
  ii. Forecast horizon - most of the predictions for this paper are 1 time step out. But different metrics may deteriorate with different rates into the future
    1. Not sure how to grapple with this and how it affects our conceptualization of predictability but should include
  iii. Nonlinearity was probably detected more in longer time series - problem for their conclusions, but not much you can do with that
    1. This would be interesting to look at with higher frequency data since we have that compared to what they had with annual data

    iv.   Great to develop a list of all the ways that think about predictability and then consider how to get complimentary information from different metrics

    v.   When are you competing metrics vs having metrics that compliment each other
1. Don't think it is clear cut
2. In this context - be able to articulate what you can get from a new metric vs from the other metrics
3. RMSE tells you one thing and then you can use other metrics to tell you other related things
4. Have dialogue about how to layer metrics and get information from them
5. Creating a multi-dimensional picture of what we mean by predictability and what are the metrics that match those lenses of predictability

    vi.   Are metrics sufficiently comparable to calculate multiple metrics and see how related those responses can be?
1. Think so.
2. Not sure how to generate some of the metrics - they are emergent properties that do not have equations. The processes generate the stochasticity, dimensionality, and nonlinearity of the system.
3. Bilgecan has looked at autocorrelation and WVPE - the relationship changed shape
4. Autocorrelation and WVPE have different definitions but don't know empirically the difference between the two

    vii.   Length of time series is a limitation - with animals always has a lot of observation errors, so the NEON Challenge may be more relevant to be able to explore the metrics

    viii.   CRPS includes the uncertainty where RMSE - different metrics have different axes

    ix.   If you only calculate on mean time series it dampens all the time series - so end up with less fluctuation with lower observation error, then will have higher fluctuation

    x.   The treatment of the observation error will affect your predictability error - intrinsic and realized

    xi.   Bilgecan diagram of putting together the different ideas of predictability - in a less mathy approach to predictability than Mike's 1st principles paper
1. Use the forecast horizon as the predetermined limit - not cited on the model, but based on the model. Could be error based or uncertainty based
2. Could have forecast limit that can be error based or uncertainty based
3. Forecast horizon/forecast limit is more interesting on short term datasets

        4. If we go with a synthesis paper - would be nice to include this diagram that describes how they are measured and the types of metrics used for each of the measurements

    xii. Density dependent factors may play different role at top of food web (longer lived species) vs lower on the food web (shorter lived species) and how human influence affect the realized predictability

  z. Next call - what to focus on?

    i. Read more papers

    ii. Paper - https://www.cell.com/trends/ecology-evolution/fulltext/S0169-5347(19)30087-4?dgcid=raven_jbs_etoc_email

        1. Trying to standardize jargon on time series so could be interesting to look at if there is anything to learn for predictability

        2. Discuss this paper on the next call

    iii. Error metrics paper - the choice of error metrics can influence your overall conclusions. Ideas in this paper could feed into the synthesis and what metrics to use. Not relevant for forecasting specifically, but useful frameworks

        1. https://www.sciencedirect.com/science/article/pii/S0304380023002922?via%3Dihub

        2. Freya to lead December 2

    iv. Fishing down the food web - https://www.science.org/doi/10.1126/science.279.5352.860

3. Other papers previously shared that Jody is including here for reference

  a. The intrinsic predictability of ecological time series and its potential to guide forecasting; https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecm.1359

  b. Prediction in ecology: a first-principles framework; https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/eap.1589