

October 1, 2019 Theory Working Group Call

Summary of the Call for the Newsletter:

The EFI RCN will be hosting a forecasting challenge using NEON data. More details will be forthcoming over the next year. In preparation for the forecasting challenge the Theory Working Group has spent their last two calls discussing forecasting output, metadata, and standards that will allow comparisons from the forecasting challenge and with additional forecasts. The data and metadata from forecasts will also allow for analyses about the patterns of predictability in ecological systems and thinking about how scale and complexity affect prediction/predictability. The group has begun to look at other metadata standards and EML and will be working with folks from other Working Groups and anyone else interested in participating in developing forecasting standards moving forward. The group's goal is to distill the discussion to date about standards into an outline that identifies points of consensus and points still under discussion, and then to look at a few existing forecasts to see if the proposed approach could easily be applied. If you want to participate in discussions about forecasting standards, there will be calls on Oct 29 and Dec 3 at 1:30pm US Eastern time.

Participants: Will Pearse, John Foster, Mike Dietze, Jason McLachlan, Jody Peters, Amanda Gallinat, Carl Boettiger, Peter Adler

Agenda:

- Forecasting standards - will have a joint call with CI folks on Oct 29 and Dec 3 starting at 1:30pm Eastern
- Things to follow up with from August 20 Call
 - Write down mock guidelines. What would the simplest data table look like? Level 1 metadata. What do the metadata fields look like?
 - Take a stab at this and use PEcAn and Portal as test cases
 - Ask PIs to fill out table as a test case
 - Update from Peter - Minimal data table for forecast obs & pred
 - Google sheet - LINK REMOVED
 - If we ask people to contribute forecasts to central database - this is what we want to ask for.
 - Project identifier - identify from others forecasts
 - Subproject - if forecasting multiple things/multiple sites
 - Response variable - what are you forecasting

- Observed and predicted values and prediction date - on a particular date what is observed value, what is predicted value, and what is the date the prediction was forecast. Observation date and forecast date = forecast horizon
- Uncertainty - lowest value of CI, highest value of CI, CI_quantile is the type/magnitude/quantile (95%, 50%??), and maybe something about the type of CI (bayesian, frequentist?)
 - Mike: CI quantile/type - those will be constant across a lot of values. More metadata rather than data because it describes how the interval was collected. But definitely want this in the metadata.
 - But what if someone wants to add 95 and 50? Maybe embed it in the name so people can report any interval they want. Also wouldn't need to have low or high - people could enter what they want. Would just need to specify the quantile as part of the variable name
 - Have quantiles as one option. Alternative option - specify some ensemble identifier if you were reporting ensemble numbers. Make clear one is reporting some set of interval widths or reporting ensemble members. Have a way of naming that
 - Will: would be helpful to know what kind of analyses we want to do. Everything is reasonable, but not sure if the main reason we want to have the information is to know something specific. What are we going to be using the data for?
 - Multiple goals - broader EFI effort is the RCN is kicking off next year with model forecast comparison challenge from NEON catalog. Want to have this information from those forecasts. Need protocol for how those forecasts are being reported.
 - First application - multiple models predicting the same quantity from the RCN forecasting challenge
 - Theory group has posed questions about patterns about predictability in ecological systems. How things like scale and complexity affect prediction/predictability. Want to have a comparative approach to see what affects predictability. Predictors of predictability.
 - Signals of when predictions are about to break down
 - Want to be rigid with Peter's table - don't give people options for what to report especially in regards to quantiles
 - Also think about what is the predicted value? The 50% quantile? Is the response variable continuous?
 - Come up with a variety of metrics that are ways that model will fit well. Otherwise the winning model may narrowly match it.

- Do we need to specify the distribution of the response? If binomial do we need to ask about the scale? Why not the predictive distribution (in whatever distribution you are drawing from) then see where the data fall relative to that.
 - Distributions can be anything. Bimodal distribution is different than the others. 95% isn't going to work very well
 - If binomial and are on 0 to 1, but would want to be on the continuous predictor.
- Carl gave this update on EML on Slack on 9-16-19
 - The ecological metadata language, EML 2.2.0 has just been released last week! <https://eml.ecoinformatics.org/whats-new-in-eml-2-2-0.html>. EML began in the 90's as a community metadata standard, even predating XML-schema as a concept, so that's a lot of change in both technology and ecological practice to have survived. EML may not have gotten a lot of uptake among everyday ecologists and today XML can feel a bit clunky, but as many of us know it's widely used by LTER, NEON, Arctic data science center. This update introduces markdown support and the notion of a data paper written in EML. May have some utility & lessons to learn as we talk about standards for forecasting in [#cyberinfrastructure](#) and [#theory](#)
 - Ecological metadata EML is a good template to start with.
 - Describe taxonomic coverage - species and taxa IDs
 - Spatial and temporal coverage
 - Other more messy options is putting in info about the methods. This isn't machine readable information, but lets people go back to see why things may not be matching up between two entries.
 -
- Update from Will - propose model meta data/descriptors
 - Was looking for a consistent stat ontology - lots of people have tried but failed. See STATO - been dead for a year, but is up online.
 - Will's advice - not worth making a general one or using one that already exists. Is a huge task. Make one going off what Will has started on Slack link above.
 - Looking for info about complexity
 - Different Level of complexity for reporting
 - Just output (predictions) - Base level. Standard comparisons
 - Sandboxed enough can make a new prediction. Can we break your model and what does it look like
 - Sandboxed enough model we can refit to new data - Max level. How generalizable is your model?
 - Spatial and temporal info overlaps with EML

- Nrows - what does this mean? For inputs/explanatory variables that are fitting the model (calibration) as opposed to the predictions.
 - Would be great to capture this, but in cases such as with process based models it will be hard to get the calibration information
 - With EML you have both options - my data is tabular there is nrows. Or my data is not tabular and we can't enter that information. Both options are available
 - Uncertainty is another way to get at this
 - Metadata correlated with types of uncertainty - there is good guidance in Tom & Mevin's book on regularization. Ask Mevin about this.
 - Have something in the metadata about whether the model estimates observation error. Report observation if it exists. Observation error and initial conditions error are separate
- One worry about EML - will it be a barrier for folks to enter their model data. Carl thinks this can be addressed using tools in R that will generate the metadata that we are looking for
- Can we use a tiering system for metadata that is similar to Will's tiering set for complexity?
 - 1st tier - Get absolutely critical variables we need,
 - 2nd tier - what would be nice variables to have
 - 3rd tier - what are the variables that we would like to have
- Leave everything open so anyone working on any kind of forecasting model that can be added to the comparison
- Mike: there's overlap between Will's proposal and an idea I previously floated for recording what variable, uncertain, and assimilated in a model

UNCERTAINTY LEVELS

| | N | C | D | P | A |
|-----------------|---|---|---|---|---|
| Param | | | | X | |
| Init Cond | | | | | X |
| Driver | | | X | | |
| Process Error | | | | X | |
| Random Effects | X | | | | |
| Model Structure | X | | | | |

N = No
C = Contains
D = Derived from data
P = Propagates
A = Assimilates

- This has the advantage of being easy to fill out if people know what each thing means
- Metrics for predictive success focuses on prediction horizon. But there are things that are on different scales (weather vs climate; succession vs carbon). Important to include this as well. Steady states vs prediction horizon.
 - What information do we want that isn't in the forecast horizon?
 - If forecast horizon is temporal, then need to define it carefully
 - Grain and extent - want to know how independent is the forecast from the training data in space. In time it is easy to quantify. In space, it might be harder. Peter's table - ask for spatial extent and grain in metadata. Ask for center point. Then get coordinates for forecast and get a bounding box
 - Can we boil it down to Y/N - are you making a prediction for a place where you have training data?
- Things in the notes from the last CI call:

- Google Dataset description:
LINK REMOVED
 - Asked the question of whether current forecaster hit hurdles in using archives like Zenodo and OSF
 - Discussion otherwise focused on reproducibility
- Next steps
 - Prep for joint CI call on standards - Oct 29 at 1:30
 - Distill into a draft - outline points of consensus and points still under discussion
 - Look at some existing forecasts in light of the emerging vocabulary from Peter's Google sheet, Will's metadata, and Carl's EML update
 - Where are sure things vs where we will run into trouble
 - Do this as a group ourselves - or split up among the group
 - Blurb for Newsletter
 - Jody will draft
 - Working group scope / "Terms of Reference"? Chairs?. Jody will send example
 - 10-2-19 - Here are examples from Decision Science and Partners. Decision Science is just getting started and haven't had a chance to fill out the details and add their own thoughts. The draft that is in this link is a template that Jody created going off the Partners terms of Reference
 - Examples:
 - LINK REMOVED
 - Other Theory goals?
 -