

March 29, 2021 Joint Methods & CI Working Group Call

Attendees: Mike Dietze, Ethan White, Abby Lewis, Jody Peters, Quinn Thomas, Hassan Moustahfid, Jake Zwart, Rob Kooper, Leah Johnson, Denis Valle

Agenda/Notes:

1. Uncertainty Quantification & Propagation, Modeling & Stats Task View
 - i. Use the [Task View 1 on Reproducible Workflows](#) as a guide
 - ii. How is the Uncertainty section coming?
 - iii. Do we want to reference the figure Quinn added before the March 1 call throughout the Task View (currently at the top of the Task View)? Or is there another figure to use?
 - iv. Kelly - any thoughts after reading through it? Kelly not available for today's call. Here are here thoughts
 1. Hi All, I am going through the Uncertainty Quantification & Propagation, Modeling & Stats Task View, and I have a few main high level suggestions (also included as comments in the doc itself):
 1. In the top of the document, an organization scheme suggests a structure, but the following sections don't follow that structure. Each section often outlines 1). the principles of the concept, and 3). the tools available, but is largely lacking "2). How EFI envisions meeting those principles. It might be hard to outline how EFI envisions meeting the principles for each subsection (& I am not totally sure this is the goal). However, I would suggest putting a short section (3-5 sentences) after the overview that outlines how EFI envisions meeting the principles of this topic, specifically, EFI's best practices for a "successful forecasting system is an effective model with properly specified uncertainty."
 2. It might be good to add a sentence to frame the focus on ARIMA models in the statistical model tools section.
 3. I added some text to explain different types of uncertainties in the Uncertainty section, and I will probably keep editing that section tomorrow, and can help trim down where needed.
 - v. One of the main thing to work on is the conceptual figure as the text is coming together.
 - vi. Mostly done with the DA section. It has a different tone then the other sections.
 - vii. Will need another read through to make the language consistent
 - viii. Need someone committed to creating a file to stash things that are extraneous

- ix. For the conceptual figure. The key things that are in Figure 1 looks like there are 3 key things: Data, Model, Decision
 - 1. The current fig is from Mike's book, but can take this and make it more simple.
 - 2. Hassan has offered to work on drafting a more simple figure
- x. Ben/Abby can work on making the text consistent across the sections then the group can take another read through.
- xi. Alexey and Jake started a table for DA. If there is something that is missing, then please add it
- xii. Do we want to put things in that we have actually used?
 - 1. Nimble particle filter, enKF
 - 2. Lot of these things don't do a hundred percent of what we want for all applications (e.g. weather ensembles and the particle filter in Nimble)
 - 3. Do we want the list to be shorter with things we can use or make the list longer with
 - 4. Jake has only used custom code for DA. This is an important point that a lot of people use custom code.
 - 5. One option - add a column that says, "Example of Ecological Applications"
 - 6. Know that all the resources work for DA, but to use the same DA to do the forecasting is where it gets tricky. Are the DA tools built for setting initial conditions, but you have to switch functions so the DA is now the input to your model doing forecast. Or are you using the resources to propagate different types of uncertainty to extend the filter out 30 days for the forecast?
 - a. Could use DA for forecasting but it may not be specifically designed for that.
 - 7. DART will have the one with the most examples that we know people are using it.
 - 8. Main argument of using DART over your own code is that a lot of effort was put into what to do when the spatial scale gets really big and painful. But if someone wants to assimilate one variable at one site, DART is a pain - you wouldn't start with DART for this example. DART - big models, big data. It is worth the effort when it is big.
 - 9. Easy to say generally the large vs small data, large vs small models. And adding another column for ecological examples.
 - 10. 2 new DA tools to add:
 - a. EMPIRE is similar to DART - designed for Big models, but less developed.

- b. Lavendar is brand new. It is unique in that it will do ensemble variational DA.
 - 11. Bayesian tools MCMC algorithms? Has anyone worked with them? Mike has used it, but only for calibration, not for DA. Can't think of an example where someone used it in ecological context
 - 12. Anyone familiar with Open DA? They now have a GitHub. Quinn is skeptical because it says that it is linked to the GLM but there has been a number of edits to the code that probably hasn't been checked. This may not be developed far enough.
 - 13. For people participating in the forecasting challenge expect they will be more interested in the packages for the small model uses.
 - xiii. General structure - propose to drop how EFI will fill missing sections. But don't want to put EFI on the hook for this. But would be good to have a paragraph about gaps in some sections.
 - 1. DA is obvious place this would be good
 - 2. Have a number of tools for small models.
 - 3. Limitations to classic Bayes tools
 - 4. Know the resources designed for big spatial/data is tough for ecologists to use.
 - 5. Had discussion with Nimble crew about the EFI wish list of what we would like to see in Nimble
 - 6. Not a good library of how to take a time series model and convert to a forecast model.
 - 7. Mike - Has been encouraging people to think past enKF. Not enough examples for people to emulate. Don't have to rely on the assumptions of enKF to solve problems.
 - xiv. This month - Abby will work on cleaning up the text that the group can look through the next month.
- 2. [NEON Ecological Forecast Challenge](#) CI Update
 - a. Started thinking about a book down repository to move the text from the website in order to have a GitHub repo that reflects the rules. Want to have the website linked to the bookdown site.
 - b. Carl started a repository for an R package to do some checking - validate and check a submission. Still need to work on this and a lot of the pieces are there.
 - c. People submit their forecast the first time, have issues, and then with the next submission they have no problem because those issues are worked out.
- 3. Visualization/Decision Support Tools, User Interface Task View
 - i. Check in with Denis Valle during the second half of the call about this Task View
 - 1. Web-based interactive tools

2. Working with models and try to make the results of the models more amenable to decision making and trying to come up with an interactive decision support tools to help with this
3. Hope more modelers will create decision support tools. Can have complicated model that few people have the background to use it, but have a tool that lets people interact with the results from the models.
4. Go beyond thinning about the decision you want to influence. As you are doing this it can affect the development of the model.
5. Denis published on malaria epidemiology
6. Had a recent workshop with SESYNC. In the end showed the decision support tools. After the workshop it became evident that there is a real need for creating a community of practice. But don't have a place to go to interact with others to see what others have learned and to share with each other.
7. Have google group with lectures and resources.
8. There is the technical aspect of programming the resources, but there are also the need to interact with stakeholders and decision makers about what way to visualize the results best.
9. Seem like a lot of good overlap with EFI.
10. Ideas for moving forward - trying to see if there is interest in coming up with additional tutorials to create tools. Special edition for journal. ESA session.
11. What EFI has been doing:
 - a. NEON Forecast Challenge
 - b. STC Proposal - was unsuccessful but some of the ideas around decision support was around is there some basic working templates for different types of tools that could be customized to give you a place to start so you aren't starting at 0
 - c. Task Views
12. Idea for moving forward - work together on the Visualization Task View
13. Coming from a modeling perspective, you might not have contact with decision makers, but you might have some ideas of x,y, and z of how the tool can be used. This is different from having decision makers coming to ask for what is needed
 - a. Service driven vs demand driven
14. Cost-benefit of different outcomes and analyses. Benefit of different support tools - people can put in their own costs.
15. Some cost-benefits can change from stakeholder to stakeholder or from year to year. If you put in the option to change the costs, that will give the tool more flexibility and less need to recreate when there are updates that are made.

16. Denis has some publications along the line of web-based interactive tools, then had a SESYNC workshop where people shared their applications. There is no grad course that focus on these types of tools. There is a basic need for building a community.
17. Denis can send the publications and the website developed
18. The idea of a decision support competition was discussed for the NEON Challenge. I think there would be interest in there being an interest.
 - a. Much easier for us to know how to score the forecasts (stats question). Less obvious to know how to score the decisions tools.
 - b. How to turn the submission of lots of forecasts into a decision support. This could be something interesting to think about moving forward.
 - c. In the Challenge - people are mostly submitting probabilistic forecasts.
 - i. Run time - could be incorporated. If it takes forever for a model to run. It could be a good forecasting model, but bad for a decision support tool.
 - ii. Close integration between the person creating the model and the people using the data is needed
 - iii. Understanding trade offs
19. Discrete first steps
 - a. Denis will go over the Visualization Task View and he will will share SESYNC workshop materials and publications that the group can look through.
 - i. Here is what Denis shared after the meeting:
 - ii. Here is the website with the material that we developed for the SESYNC workshop “Interactive Web-Based Visualizations and Decision Support Tools in Shiny/R for Quantitative Scientists”: https://lsw5077.github.io/shiny_workshop/
 - iii.
 - iv. Here are some of the articles we have written that either include decision support tools or that talk about them:
 - v.
 - vi. Millar, J.; Toh, K.; **Valle, D.** 2020. [To screen or not to screen: an interactive framework for minimizing costs of mass malaria treatment interventions](#). BMC Medicine, 149.
 - vii.

- viii. **Valle, D.;** Toh, K.; Millar, J. 2019. [Rapid prototyping of decision support tools for conservation](#). Conservation Biology, 33:6, 1448-1450
- ix.
- x. **Valle, D.;** Millar, J.; Amratia, P. 2016. [Spatial heterogeneity can undermine the effectiveness of country-wide test and treat policy for malaria: a case study from Burkina Faso](#). Malaria Journal, 15:513.
- xi.

4. Data Ingest, Cleaning, Management

- i. Placeholder until we are further along with the other Task Views or have an identified leader for this

5. Forecasting Workflow Updates

- a. Look through the final steps of the Workflow to identify what is covered in the Forecast Challenge and what will be good to keep in mind as a need for the future