

August 21, 2019 Methods & Tools Working Group Call

Summary of this call that Mike posted on Slack:

This team's goal is to generate an "task view" of useful ecological forecasting software tools for the EFI website, with the goal of having a first draft live by the end of the calendar year.

At a high level this would focus on broad categories that emerged out of group discussions at the EFI2019 meeting:

Visualization

Uncertainty quantification & propagation

Data ingest

Data cleaning

Data management (including databases)

User interfaces for the end user (best practices)

Reproducibility

Modeling & Statistics

Workflows / Pipeline Automation

Decision Support / Decision Science tools

Beyond giving high-level structure to the page (e.g. conceptual figure), Methods and Tools will focus most of their efforts on 1, 2, and 8 but thinks it would be great to work with CI on 3-7 and 9, and to work with Decisions on 10. They also want to poll some of the current forecasting teams to find out more about what tools they're actually using. Besides being useful unto itself, we think that it sets us up well for thinking more deeply about what tools are missing and how to fill them (spring 2020)

Agenda

From last meeting: Poll/Survey for what folks want out of the working group -- goals?

- Interest in short term goals: tool awareness and best practices, etc?
- Willingness to contribute blogs posts and similar
- Top 5 long term challenges/wish list
- Identify a tool they can share/want to see (or a case study)

Notes From the August 21 Call

Attendees: Rich Fulford, Leah, Jake Zwart, Alexey Shiklomanov, Mike Dietze, Elise Zipkin, Mevin Hooten

- Recap from last meeting
- Short term: Want to start narrowing down short term goals
 - Tool awareness
 - This might be an easier thing to tackle before moving towards best practices
 - ID the strengths and weaknesses will help us think about best practices and needs

- Vignettes showing how to use for forecasting
 - We could outsource vignette development to tool developers and fold into EFI community
- Also useful to have a **high** level summary of broad class of tools before digging into vignettes
- Pipeline diagram of a generic forecast
 - Offshoots of that pointing to tools and their vignettes
 - Reproducibility
 - DA techniques
- Could be synergistic effort with Cyberinfrastructure
- Could also link to blog posts
- Should this be interactive with links to blog posts, etc..?
 - Could start with static slide first and make more interactive if time allows
- Already have some noted/discussion from May meeting about the major categories of tools (Ethans' group). ~8 essential components of ecological forecasting
 - Data ingest, UQ, etc.
- **Next steps:**
 - Going through May notes
 - List from May Notes (with some additions during call) - bolded is likely under the purview of this group:
 - **Visualization**
 - **Uncertainty quantification & propagation**
 - Data ingest
 - Data cleaning
 - Data management (including databases)
 - User interfaces for the end user (best practices)
 - Reproducibility
 - **Modeling & Statistics**
 - Generic tools
 - Classic LM, GLM, GLMM, DLM (standard models, possibly multiple implementations)
 - Time series analysis: ARIMA, **etc**
 - Spatial
 - GP
 - Jags, Stan, Nimble (general purpose tools)
 - Data Assimilation
 - Methods: KF, EnKF, 4DVar
 - Tools: DART
 - ML
 -

- Workflows/PipelinesAutomation?
 - Decision Support / Decision Science **tools**
 - [Alexey] what's the difference between CI and methods/tools? There seems to be a lot of grey area...
 - [Mike] CI is more workflow side of things whereas methods/tools is more stats side of things
 - [Alexey] math == tool
 - Start new google doc to organize **[[Mike]]**
 - Homework for folks to fill in some of the details of known packages/tools and do some research on what else is out there?
 - Share that we're doing this with CI group to contribute to collaboratively
 - Poll some of the forecast teams about what they're actively using right now
 - Draft figure
 - Timeline
 - Now is crazy with start of semester
 - High level by next meeting (late Sept)
 - Would like to have a first version online by end of year (official RCN launch)
 - Best practices
- Medium term: What are the things that we'd like to achieve by RCN meeting summer 2020?
 - Best practices?
 - Definitions / semantics (e.g. what's the difference between forecasting and projection?)
 - [Jake] USGS is currently struggling with this - dividing line was arbitrarily made at a ~10 year future prediction, >10year = projection, <10 year = medium to short-term forecast
 - What is calibration vs. data assimilation vs. inference.....?
 - Consistency with other communities?
 - Identification of gaps in tools and methods
 - What is an achievable REU forecasting project (10 week, undergrad)?
 - Bottlenecks in ingest, tools, etc that someone would have to build themselves
 - Could tap into Cayelan's group - they had an undergrad present a poster on a lake methane emissions forecast that was developed over the summer, 2019
 - This would identify pain points to forecasting
 - Also would overlap with methods
 - Could identify a case study or examples to make this more clear / discrete
 - Build out more of the tool awareness materials
- Long term: Are there long-term things we'd like to be working towards?
 - Homework

- Discuss on Slack, summarize on next Agenda Doc
- Poll for recurring meeting
- Leah: JSM meeting section