

February 19, 2025 Joint Methods & CI Working Group Call

Attendees: Chris Jones, Chris Brown, Jody Peters, Hassan Moustahfid, Dave Durden, John Smith, Brittany Barker, Will Hammond, Saeed Shafiei Sabet

Regrets: Jake Zwart

1. EFI2025 Conference Update; <https://bit.ly/efi2025>
 - a. Registration is open. Early bird deadline is March 14, Full registration and late-breaking poster deadline is April 14
 - b. The Spatial Forecast Workshop was accepted!
 - c. Development needed for the workshop
 - i. Some of the null models (linear regression or single grid analyses) are good places to start
 - ii. Need to make sure that everything is scored properly and make sure GitHub actions is working so people can submit
 - d. Next steps?
 - i. John and Will met this morning and have a working parametric forecast!
 1. This is grid wise
 2. Had questions about target file generation
 3. Where does targets get stored? Forecasts are archived - right?
 - a. Everything should be archived to the bucket provided by Carl. John couldn't find the bucket
 - b. Here it is: <https://data.ecoforecast.org/spat4cast-targets/>. Or maybe not
 - c. Dave has an old link in his browser for the targets, but it is not accessible
 - d. Need to write the targets to a folder that exists
 - e. Ping Quinn and Carl about this
 - ii. John has been able to run everything locally. The issues have been writing to the write folders to get GitHub Actions working.
 - iii. The goal of this workshop is to introduce the spatial forecasting challenge and get participants to run an example model. We will provide example R code for generating a spatial forecast and discuss differences in submitting to this forecasting challenge vs the other NEON-EFI forecasting challenges. In this example we will focus on locations of wildfire burns and recovery using MODIS leaf area index (LAI) data. Time will be devoted to outlining grabbing training data, formatting model output for submission, how models will be scored, and reviewing visualizations of model scoring.
 1. Parametric forecasts will be submitted differently
 - a. What do we want the submissions
 - b. Fit auto-ARIMA for each of the grid cells
 - c. Thinking want to submit 2 geotif files with the mean and SD for each grid cell forecasting in

- d. Could put metadata in for type of submission (parametric vs ensemble) then have 1
 - e. Do it with the option that saves space
 - 2. Ensembles submitted as collections of geotif files and have a specific scoring routine for scoring ensembles
- iv. Want to grab some other bounding boxes to have a few examples that people can think about submitting to. Right now have August Complex
- v. Will devote most of the time to outline how to grab training data using ingest planetary data function, format for submission, and explain how models will be scored with CRPS and log score
 - 1. Have visualizations where model is performing well or is not performing well
 - a. Could do something similar to the NEON challenge visualizations <https://projects.ecoforecast.org/neon4cast-ci/performance.html>
 - b. John thinks we can do something similar to the visualization so people can see the null models and the parametric model example
 - c. **Ping Quinn and Carl about this**
 - d. On the webpage there is a view source code, but there is an error with that
- vi. **Main priority**
 - 1. Dave, Chris - can you set aside time with John and Will to get the GitHub Actions set up? Once that is set up then everything else should be easy.
 - a. Need to get this done now so we can make progress on the materials of the workshop
 - 2. Think the main issue is the bucket issue, so get Carl involved as well
 - 3. Goal is for the subgroup to meet in the next 2 weeks
- vii. Small things
 - 1. Have data stored locally to make it easier to run the model, but want to walk through how to pull data
 - 2. Start of data - 2002 to April 2025
- viii. Brittany has R code to download and process Soil moisture (SMAP) data for all the years. Extracts data for the root zone.
 - 1. This can be shared with the group as a covariate option
- ix. Other sites
 - 1. Have focal dataset with the August Complex
 - 2. Could use Brittany's list of other sites
 - 3. How hard is it to get the bounding boxes?
 - a. It isn't too hard. Just need shapefiles
 - b. Have East Troublesome fire as well just don't run it as the default

- c. **Brittany will get all the bounding boxes - the lat, longs and include info about the size of the fire.**
 - d. **Then once all the infrastructure is in place then can slip in the other bounding boxes for the sites**
 - 4. Would be cool to add NEON sites since people could run other covariates
 - 5. Brittany can get the bounding boxes. Brittany recommends not using places with small sized fires e.g., Konza
 - a. Perhaps start with the larger fire sizes
 - b. But small size burns aren't a problem and they may be a little easier to work with to start to make it easier for people to test out during the workshop
 - 6. Konza is a prescribed burn. Do we want to exclude or include prescribed burns?
 - a. Data is at 0.1 degrees which is 11km. So the prescribed burn will be hard to justify as an example
 - b. LAI is 500 m resolution. The 0.1 degree is a buffer when we grab data. We subset that with the bounding box once the data is grabbed
 - c. Have a 20x21 grid - so could get smaller fires than the August Complex, but 500m is still coarse
 - d. So maybe something intermediate. The Bayes temporal spatial model is fit as a regular grid because it has neighbor information and then mask everything by the bounding box at the end. It is more annoying to fit something on an irregular grid. So it is a little wasteful computationally, but does include other information.
 - e. Consider surveying the participants on what they think would be useful in the future. Use this to identify future directions
 - i. Other spatial forecasts of interest
 - ii. Other fire related information that would be useful
 - iii. Put it in a Google form!
 - iv. Side note: one future direction - a system with more dynamic responses, e.g., predict invasive species in coastal system
 - 1. When the time comes, could get data from Chris, Hassan, NOAA
2. Notes from December call: Any updates about incorporating other LAI products to replace MODIS
- a. Landsat harmonized with Sentinel-2 option - Dave was going to work on getting the data into the function already in the repo - this is as raw bands. Still need to calculate LAI from Landsat
 - i. <https://planetarycomputer.microsoft.com/dataset/storage/hls>
 - ii. Dave - how did this go as you were prepping for AGU?

1. Hasn't incorporated it into the function yet.
 2. Previously tried to add Sentinel-2 in hdfm format, but the other hdf5 format should be easier
 3. Brittany is also working with the SMAP in the hdf5 format. Was an issue with the intallation of gdal not being recognized.
 4. Dave will give it a shot to resolve the issue in the new year.
3. Additional fire sites with long term fire histories
- a. Brittany did a GIS analyses to look at fire histories at NEON sites and other long term sites (see Table on next page)
 - b. Did spatial join of Justin's wildfire and NEON's sites - tallied the sites per year and by size and earliest
 - c. Talladega National Forest had most fires, other top sites are Disney Wilderness, Santa Rita, LBJ Grasslands
 - d. All the locations have recent fires
 - e. Additional places to consider adding: Several locations in Colorado, Wyoming, and New Mexico have recently had very large fires, such as Pike National Forest (e.g., Hayman and Waldo Canyon Fires), near Grand Junction (Pine Gulch Fire), Medicine Bow NF in WY (Mullen Fire), and the Gila NF in NM (e.g., Silver, Black, Whitewater-Baldy). These are also relatively dry forest ecosystems.

siteID	siteName	state	n_fire_yrs	av_fire_km2	earliest_yr	latest_yr	field_dominant_nlcd_classes
TALL	Talladega National Forest	AL	18	52.342968	1999	2020	Deciduous Forest Evergreen Forest Mixed Forest
DSNY	Disney Wilderness Preserve	FL	14	48.504342	1984	2012	Pasture/Hay Woody Wetlands
SRER	Santa Rita Experimental Range	AZ	12	214.897271	1960	2018	Shrub/Scrub
CLBJ	LBJ National Grasslands	TX	11	41.883456	1996	2019	Deciduous Forest Grassland/Herbaceous
GRSM	Great Smoky Mountain National Park, Twin Creeks	TN	11	32.72719	1920	2016	Deciduous Forest Evergreen Forest
SJER	San Joaquin	CA	11	18.222906	1950	2020	Evergreen Forest Grassland/Herbaceous Shrub/Scrub
SOAP	Soaproot Saddle	CA	9	5.822705	1932	2020	Evergreen Forest Shrub/Scrub
ONAQ	Onaqui	UT	8	67.774086	1986	2020	Evergreen Forest Shrub/Scrub
KONZ	Konza Prairie Biological Station	KS	6	34.868609	2013	2018	Deciduous Forest Grassland/Herbaceous
MLBS	Mountain Lake Biological Station Additional TOS Boundary	VA	5	8.898438	2007	2017	Deciduous Forest
KONA	Konza Prairie Biological Station	KS	3	2.749077	2014	2018	Cultivated Crops
BONA	Caribou-Poker Creeks Research Watershed	AK	2	24.0199769	2004	2016	Deciduous Forest Evergreen Forest Mixed Forest Woody Wetlands
DEJU	Delta Junction	AK	2	29.903834	1999	2006	Evergreen Forest Shrub/Scrub Woody Wetlands