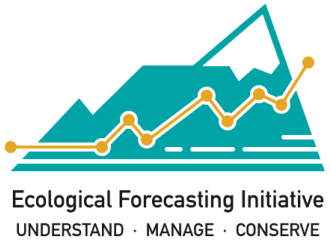


Can you predict the future? Introduction to the NEON Forecasting Challenge

ESA workshop 3

Monday 7th August 2023

Workshop lead: Freya Olsson
Virginia Tech, USA (freyao@vt.edu)

Workshop overview

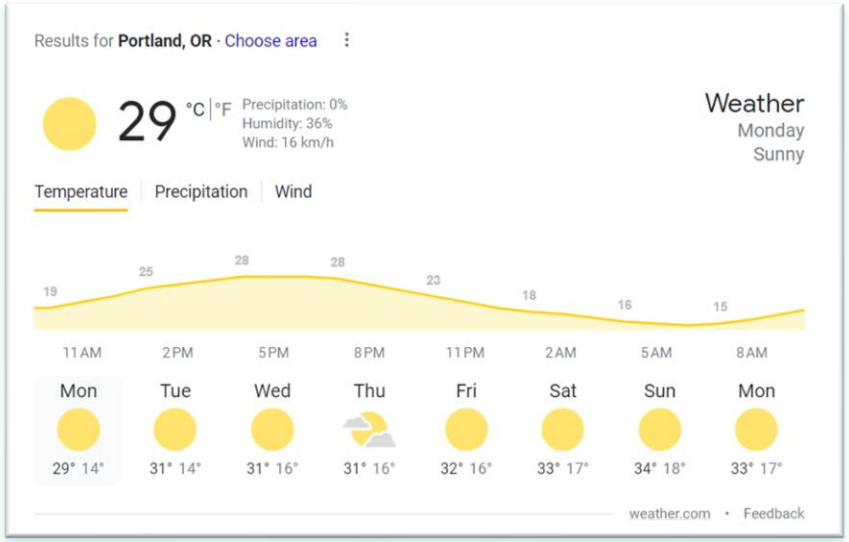
Objectives:

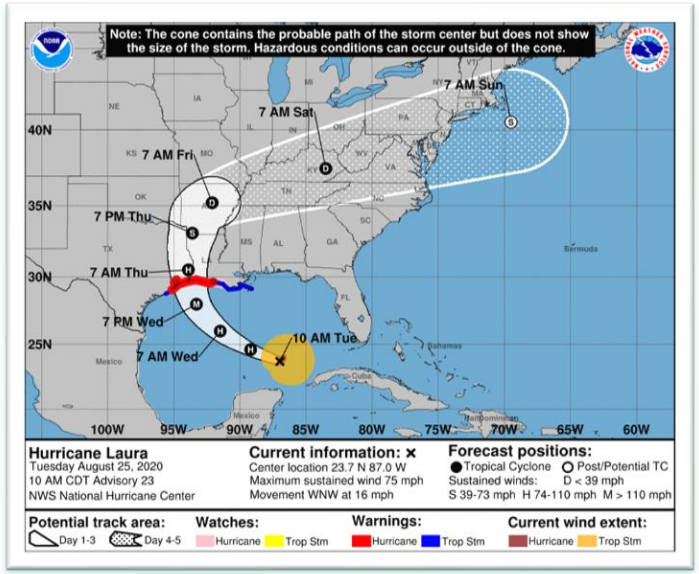
1. Highlight some key concepts of ecological forecasting
2. Introduce NEON and the Forecasting Challenge
3. Walk through a simple forecast workflow
4. Point to resources to get involved and find more information

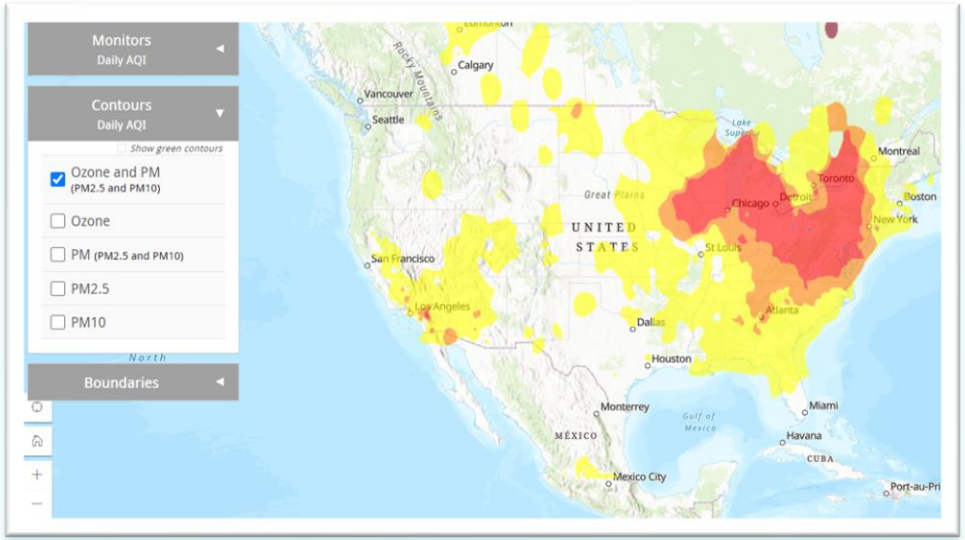
Workshop Overview

11:30-11.55	Introductory presentation
11.55-12.10	**Break and R set-up**
12:10-13:00	Hands-on coding/follow-along

Why forecast?







Near-term, iterative, ecological forecasts

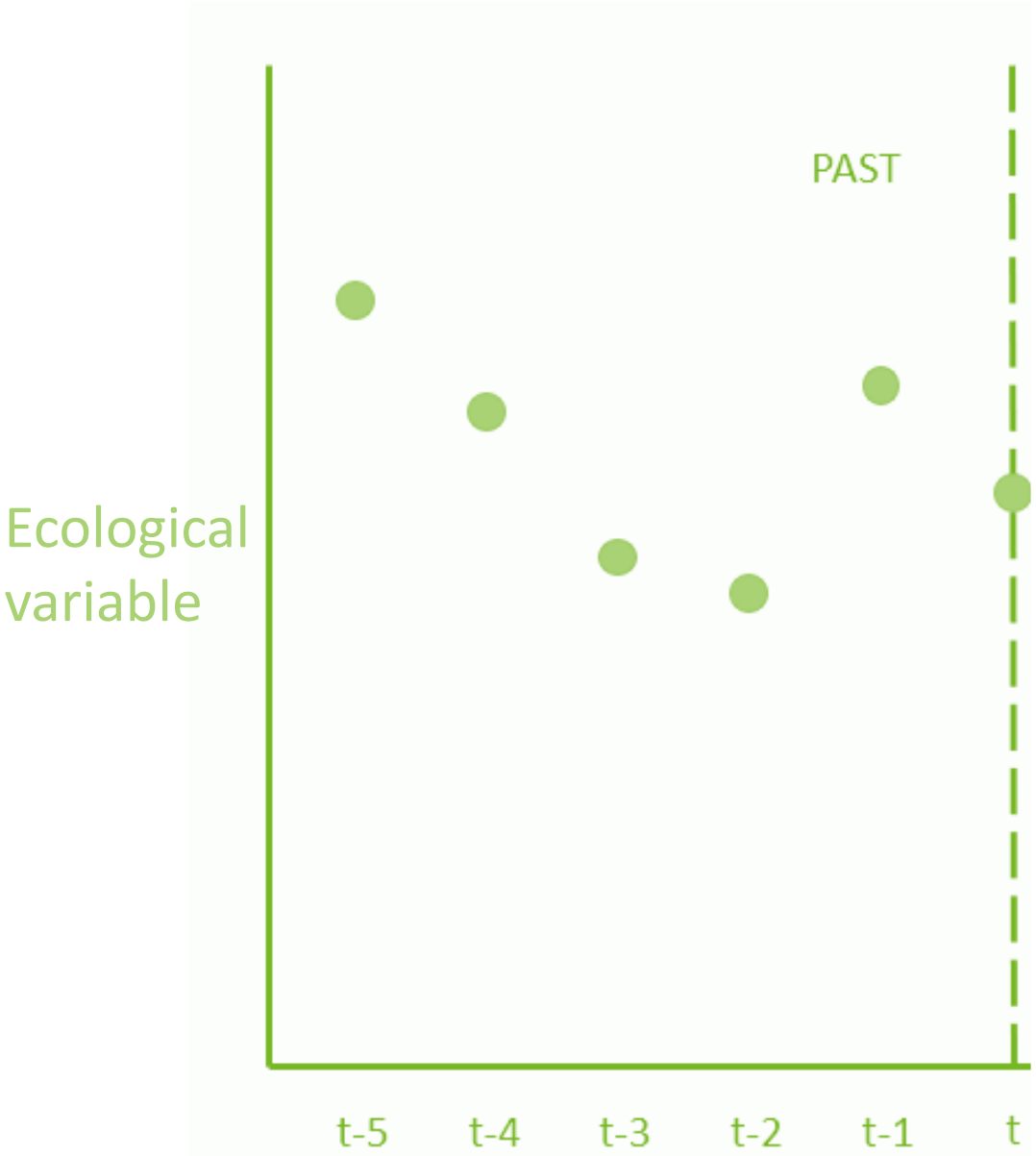
- **Near-term** = sub-daily to decadal timescales
- **Iterative** = process of repeatedly validating forecasts, updating model initial conditions and parameters, and issuing new forecasts as new data become available
- **Ecological forecast** = future predictions of physical, chemical, or biological variables *with quantified uncertainty*

Examples:

1. Forecast of river dissolved oxygen concentration for the next 1-48 hours for fish stocking

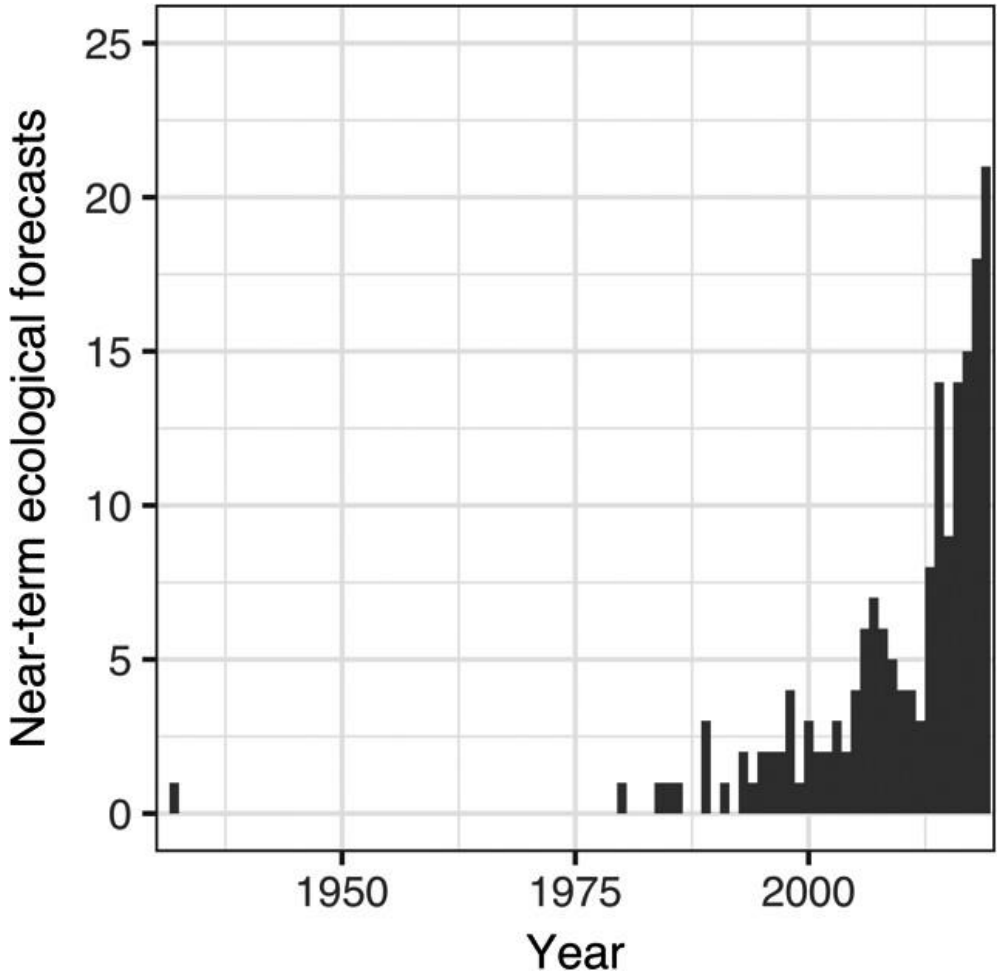
2. 1-3 month ahead predictions of the % chance of leaf fall to estimate peak leaf-peeping

3. Forecasts of tick abundance for the next 1-30 days in a popular hiking area


PAST
Ecological variable
t-5
t-4
t-3
t-2
t-1
t

Ecological forecasting

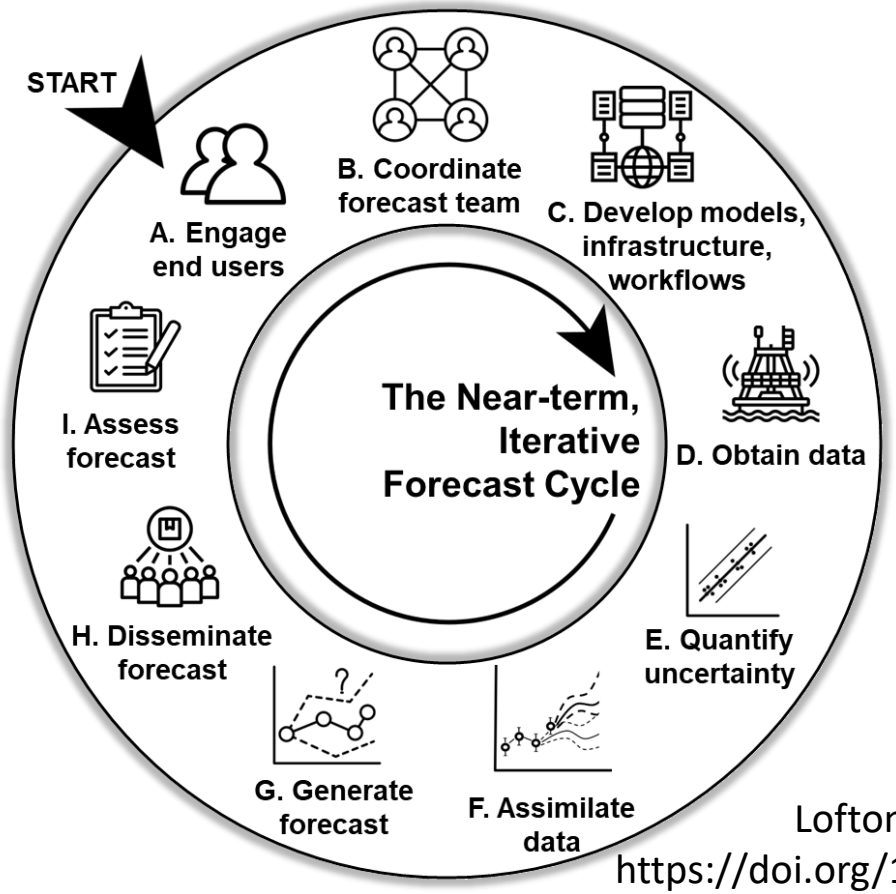
- An emerging and growing field



Lewis et al., 2021 (https://doi.org/10.1002/eap.2500)

Forecasting Challenges

- Forecasting is challenging!



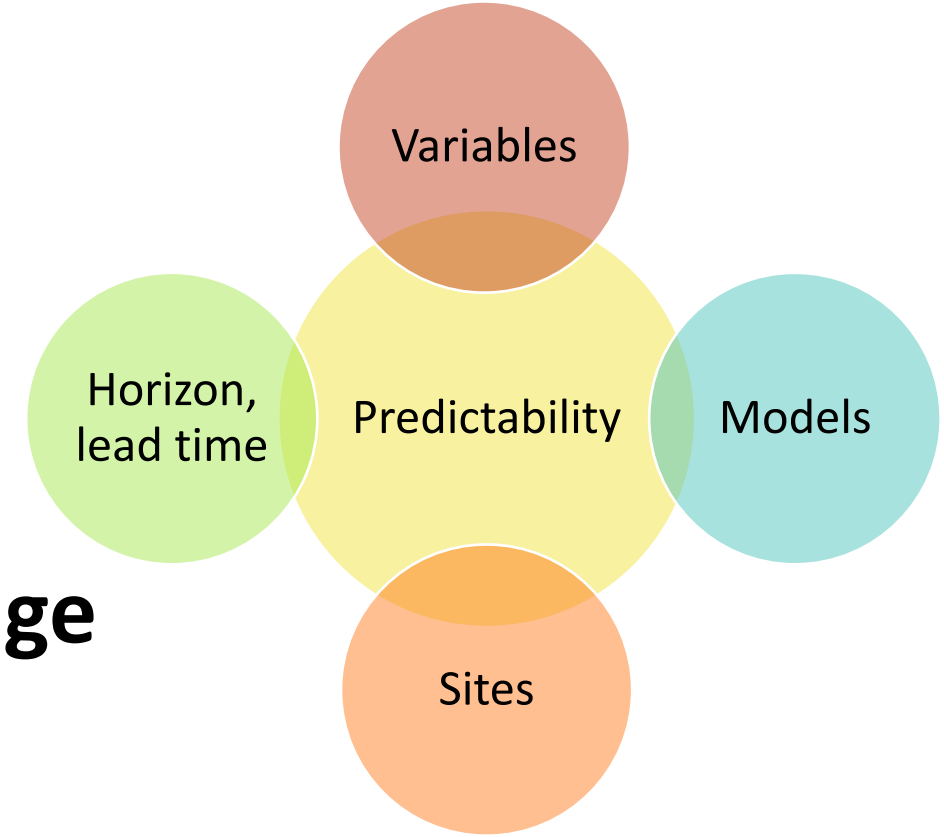
Lofton, M. E *et al*. (2023)
https://doi.org/10.1111/gcb.16590

- A Challenge to catalyze progress

Why a forecasting challenge? – the power of many forecasts!

1. A community of forecasting
   - Standards
   - Development of tools and infrastructure
   - A forecasting **platform**

2. Answer questions of predictability



The **EFI-NEON Forecasting Challenge** was born (2021)!

Ecological Forecasting Initiative Research Co-ordination Network

- EFI RCN Goals
  - lower barriers
  - community building
  - infrastructure
  - platform development

NSF-sponsored

Ecological Forecasting Initiative
Research Coordination Network
5-year project

Create a community of practice that builds capacity for ecological forecasting by leveraging NEON data products.

https://ecoforecast.org/rcn/

Funded by the National Science Foundation (DEB-1926388)

EFI RCN supported by NSF grant DEB-1926388

What is neon?

- The National Ecological Observatory Network (NEON) is a **continental-scale observation facility**
- To collect **long-term open access ecological data**
- **47 terrestrial and 34 aquatic sites**

2.1. NEON Mission

NEON is a National Science Foundation-sponsored facility for research and education on long-term, large-scale ecological change. NEON's goals are derived from the Integrated Science and Education Plan.

The goals of NEON are to:

- Enable understanding and forecasting of the impacts of climate change, land use change, and invasive species on aspects of continental-scale ecology such as biodiversity, biogeochemistry, infectious diseases, and ecohydrology
- Enable society and the scientific community to use ecological information and forecasts to understand and effectively address critical ecological questions and issues
- Provide physical and information infrastructure to support research, education, and land management.

From: https://www.neonscience.org/sites/default/files/NEON_Strategy_2011u2_0.pdf

What is the EFI-NEON Challenge?

"A platform for the community to make predictions of conditions at NEON sites before the data are collected"

- All 81 sites
- 6 themes

What is the EFI-NEON Challenge?

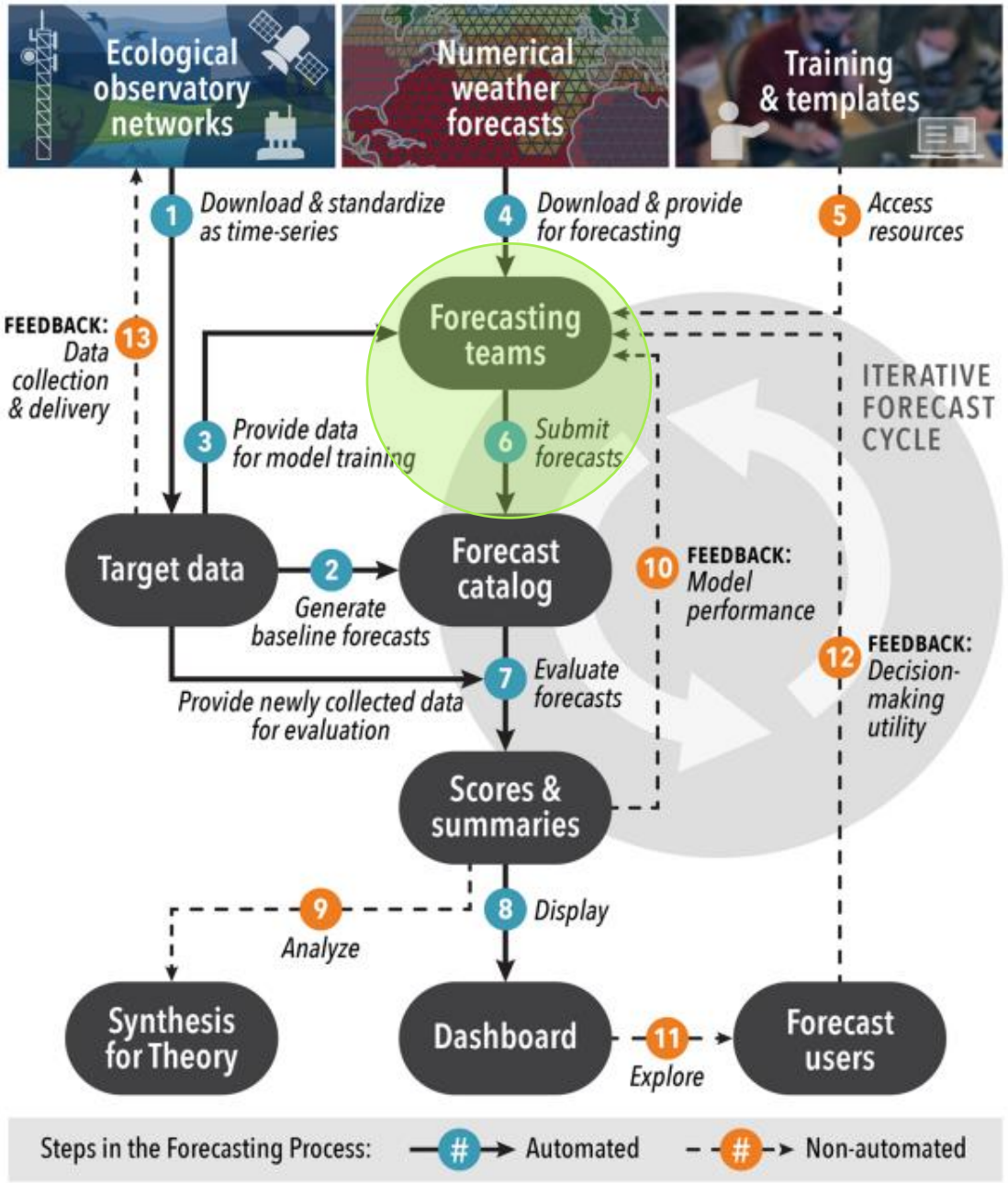
"A platform for the community to make predictions of conditions at NEON sites before the data are collected"

- All 81 sites
- 6 themes
- > 15,000 forecast submitted!

Forecast =
A prediction of future environmental conditions that includes quantified uncertainty

Thomas, et al. (2023). The NEON Ecological Forecasting Challenge. *Frontiers in Ecology and the Environment*, *21*(3), 112–113. https://doi.org/10.1002/fee.2616

Challenge overview



Thomas, et al. (2023). The NEON Ecological Forecasting Challenge. *Frontiers in Ecology and the Environment*, *21*(3), 112–113. https://doi.org/10.1002/fee.2616

Workshop overview

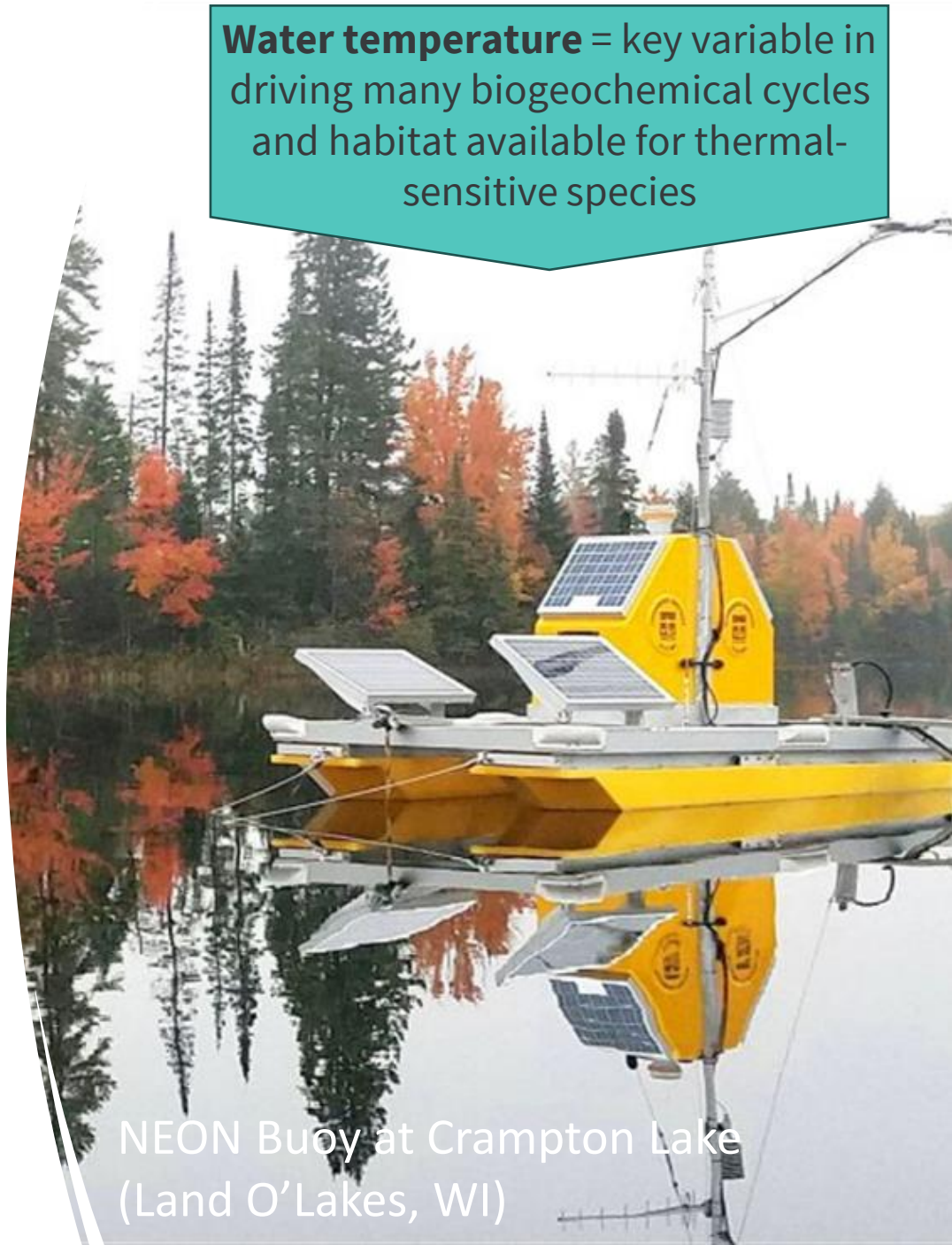
Objectives:

1. Highlight some key concepts of forecasting
2. Introduce NEON and the Forecasting Challenge
3. Walk through a simple forecast workflow
4. Point to resources to get involved and find more information

Hands-on workshop:

- **Aquatics theme – Can we predict how water temperature will change over the next month?**
  - Water temperature in lakes
  - NEONs water temperature data product (DP1.20264.001)
  - 30 day forecast horizon
  - Data latency of 2-3 days
- **Simple baseline** model to build on

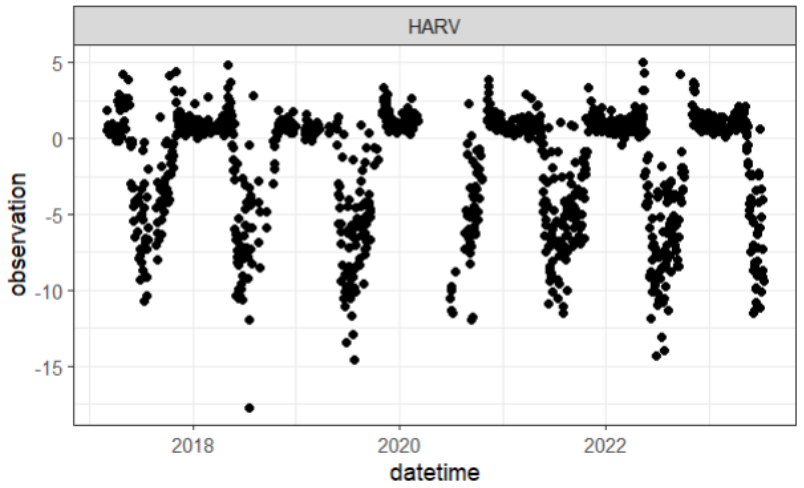
Water temperature = key variable in driving many biogeochemical cycles and habitat available for thermal-sensitive species



NEON Buoy at Crampton Lake (Land O'Lakes, WI)

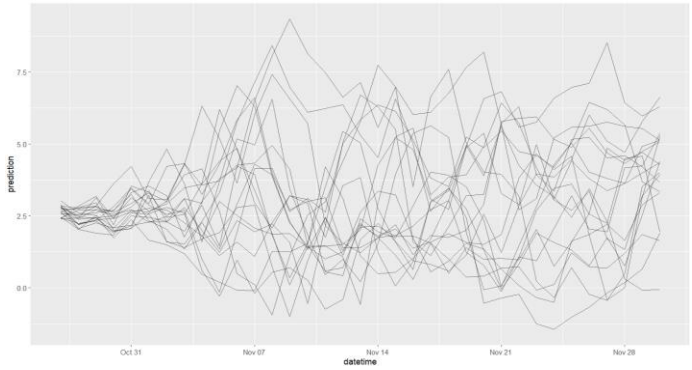
Diversity of NEON aquatic sites

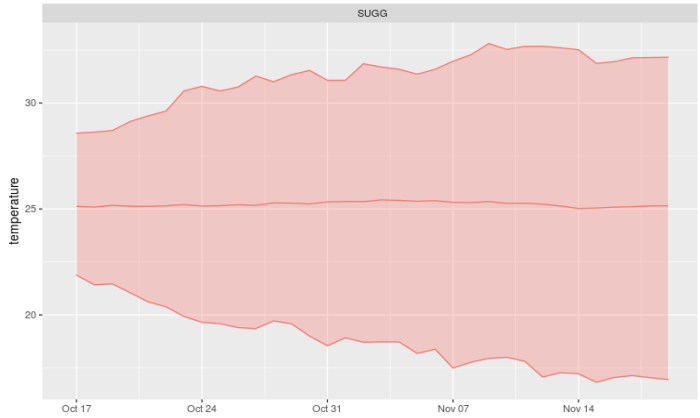
Some forecast terminology:



- Targets – water temperature

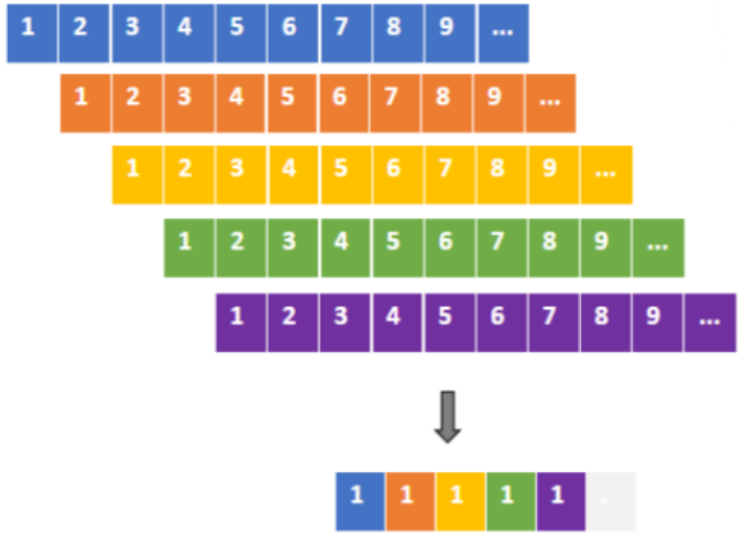
- Uncertainty – forecasts must include an estimate of uncertainty.
  The uncertainty can be represented using different model runs (**ensemble members**) or the statistics of the forecast (**mean and standard deviation**).





Some more forecast terminology:

- NOAA data – National Oceanic and Atmospheric Administration weather forecasts
- 3 NOAA forecast data products available in neon4cast:
  - Stage_1: **raw forecasts** from NOAA.
    30 member ensemble forecast
  - Stage_2: processed from stage_1 Recommended for **future forecasts. Hourly inputs**
  - Stage_3: the **historic data product**. A 'stacked' data set taking every 1 day ahead forecast.
    Useful for model training/calibration.

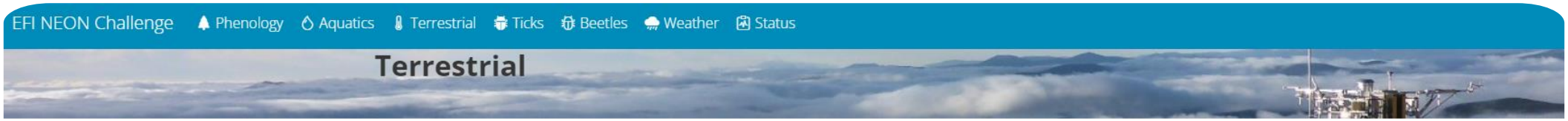


A little more forecast terminology:

Scores – a means to assess forecast skill. The Challenge uses the Continuous Rank Probability Score (crps). Uses both the **accuracy** (mean) and the **precision** (sd) of the forecast.

Scores → Dashboard → Users!

Read more: https://projects.ecoforecast.org/neon4cast-docs/Evaluation.html



On this page

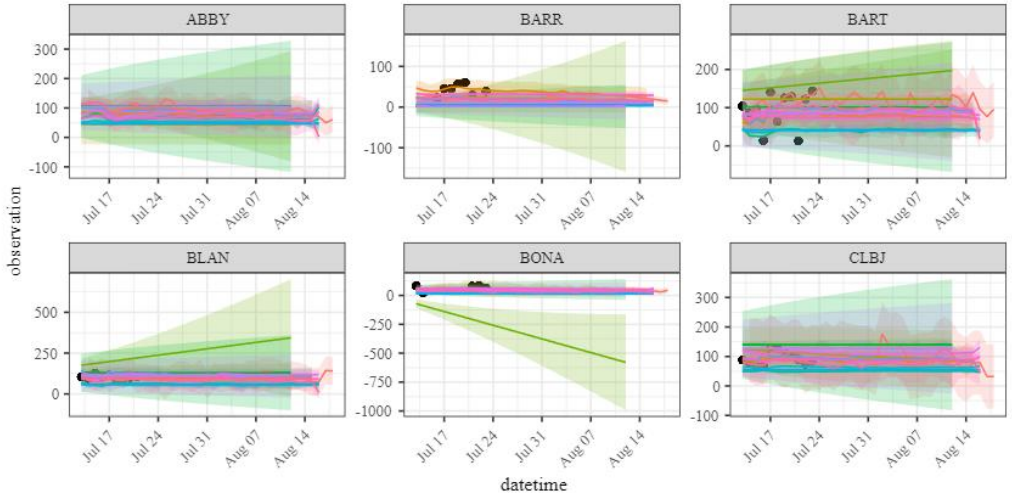
- Daily averages
- Forecasts
- CO2 Flux
- Leaderboard
- Submission statistics

Edit this page
Report an issue

Daily averages

Forecasts

These plots show the most recently submitted forecast (a single `reference_datetime`) for which we have at least 10 observations.
Models which did not submit a forecast on the given reference date are not shown. ::: panel-tabset ## Latent evaporation



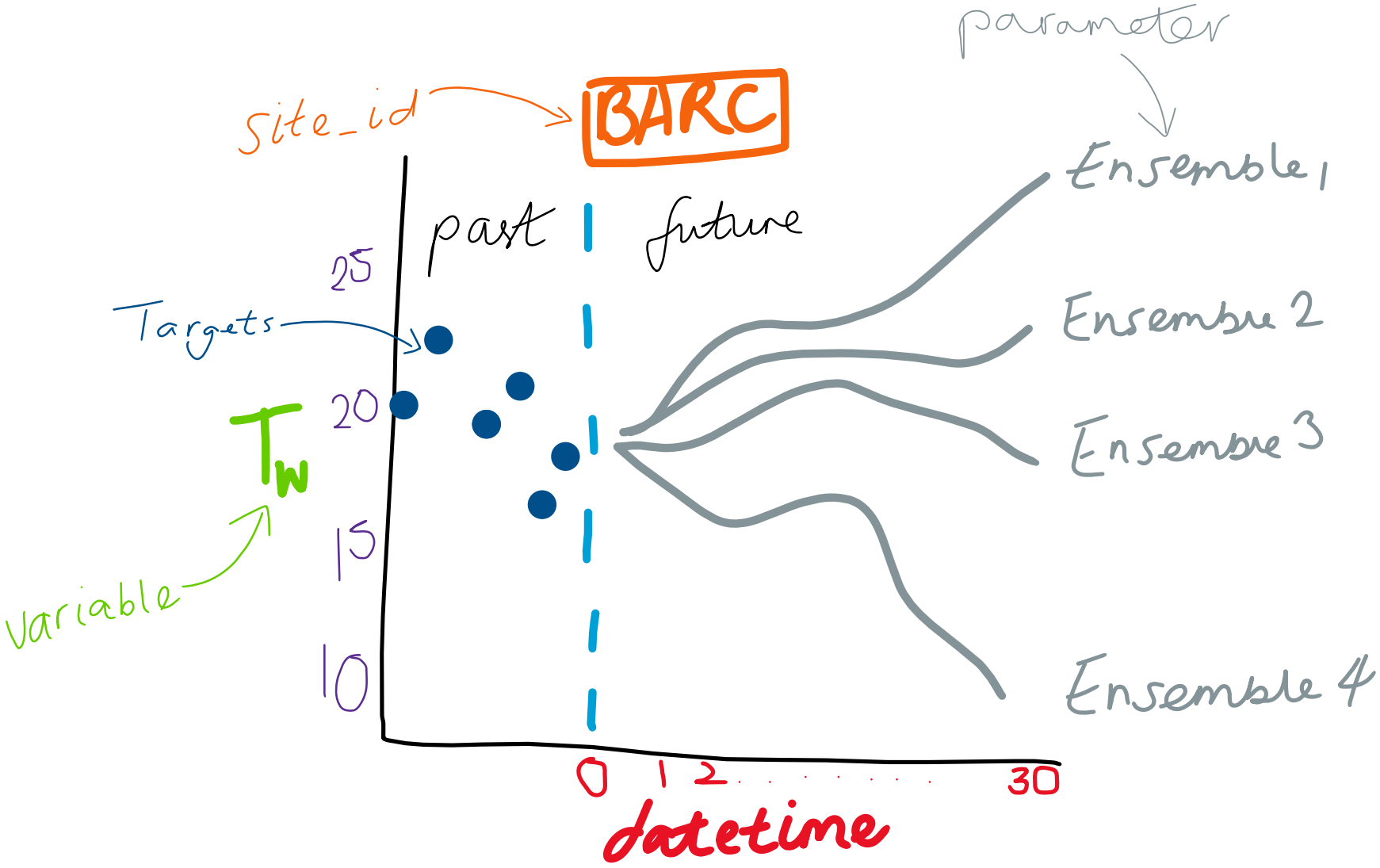
https://projects.ecoforecast.org/neon4cast-dashboard/

A little more forecast terminology:

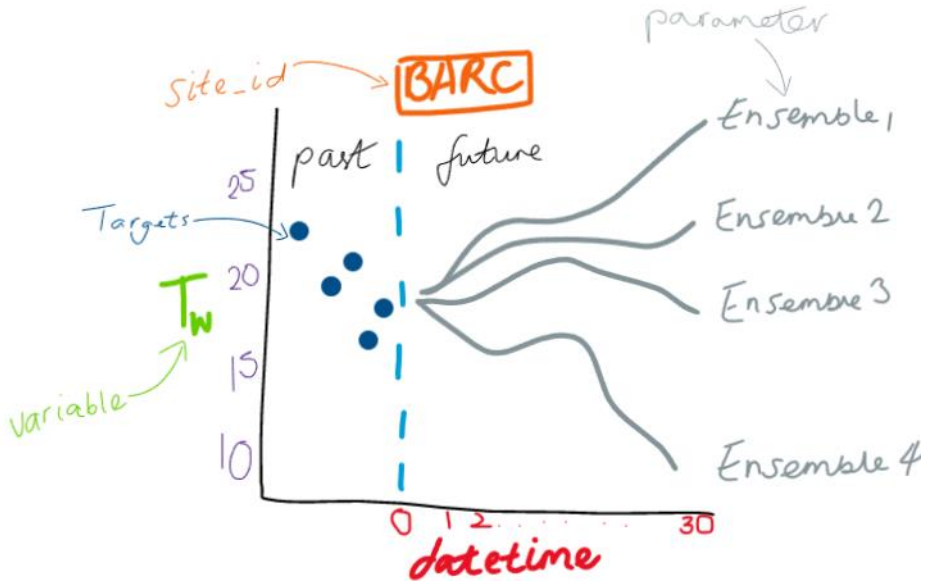
Standards - Help maintain consistency in forecast generation, submissions and scoring

Need to submit a forecast in a standardized format

- file format (csv or NetCDF)
- file name format ([theme]-[reference_datetime]-[team_name].csv)
- specific column names
- column format (datetime/character/integer/etc.)

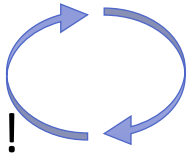

parameter
site_id
BARC
Ensemble 1
past
future
25
Targets
Ensemble 2
20
T_W
Ensemble 3
15
variable
10
Ensemble 4
0
1
2
30
datetime

datetime	reference_datetime	site_id	family	parameter	variable	prediction	model_id
2023-01-12	2023-01-11	BARC	ensemble	1	temperature	22.63563	test_mod
2023-01-12	2023-01-11	BARC	ensemble	2	temperature	26.75148	test_mod
2023-01-12	2023-01-11	BARC	ensemble	3	temperature	24.65157	test_mod
2023-01-12	2023-01-11	BARC	ensemble	4	temperature	25.1389	test_mod
…	…	…	…	…	…	…	test_mod
2023-02-10	2023-01-11	BARC	ensemble	1	temperature	19.40379	test_mod
2023-02-10	2023-01-11	BARC	ensemble	2	temperature	24.89667	test_mod
2023-02-10	2023-01-11	BARC	ensemble	3	temperature	25.98961	test_mod
2023-02-10	2023-01-11	BARC	ensemble	4	temperature	26.40593	test_mod



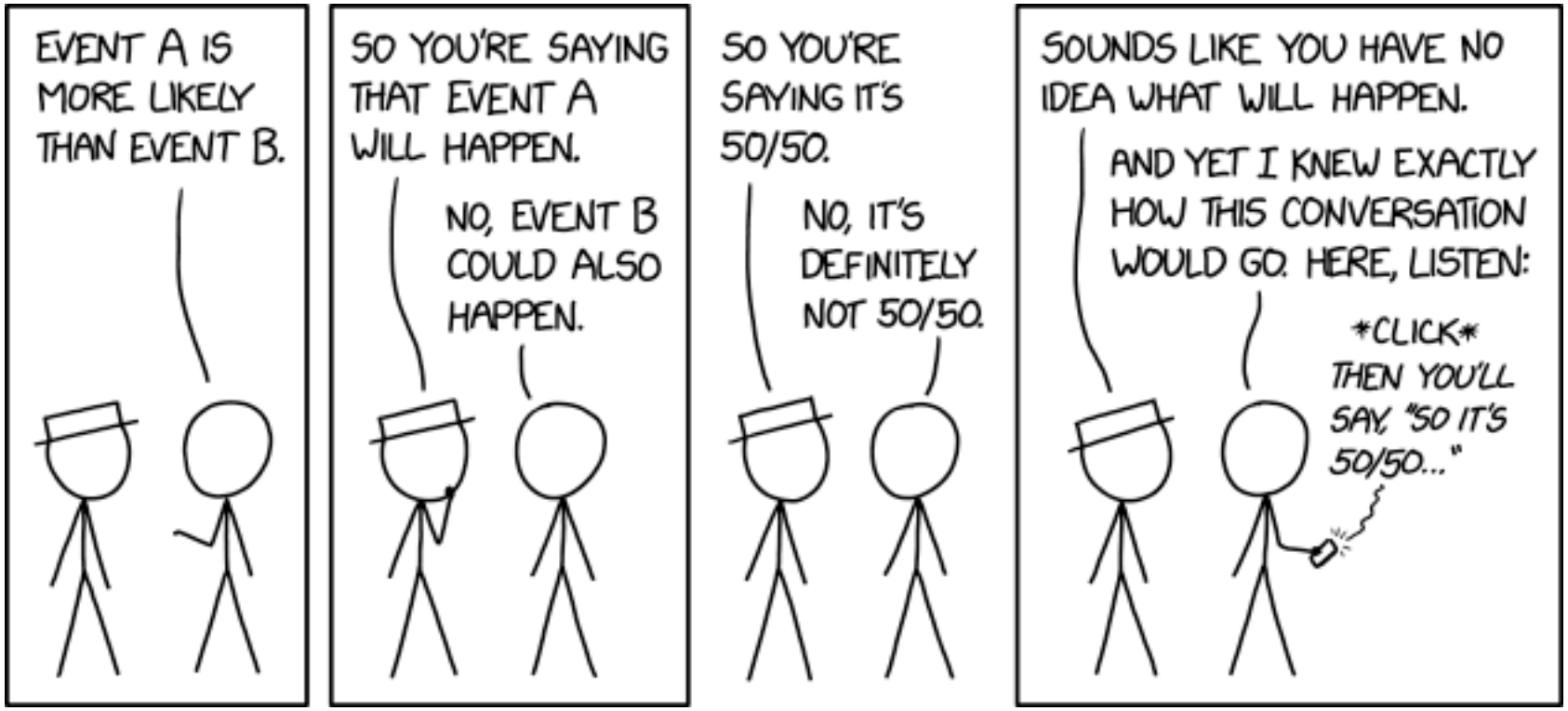
Basic workflow to submit a forecast

1. Read EFI-NEON Challenge documentation (neon4cast.org)
2. Investigate the forecast target variables
3. Build/apply your model!
4. Produce forecast of future conditions – SUBMIT TO THE CHALLENGE!
5. Register, complete model description, and submit your forecasts
6. Wait for the scores to come in and revel in the glory of predicting the future (~5-day before first evaluation)
7. Use new data to update the model
8. Submit another forecast! And another…!



Workflow automation

Questions?

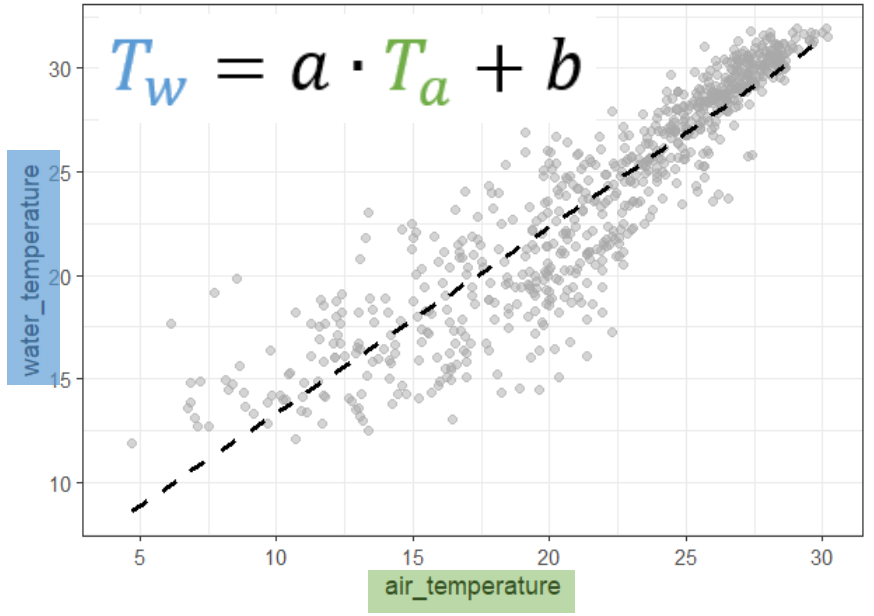


https://xkcd.com/2370

Let’s forecast!

1. Follow-along R markdown
   - Forecasting water temperature using a Linear model with air temperature

2. Modify the model and submit your forecast!
   - More/other covariates
   - Different model structures
   - Other variables



Thank you for attending!

- Big thanks to the EFI-NEON team - especially **Quinn Thomas** (Virginia Tech) and **Carl Boettiger** (UC-Berkley)
- Tech assistants

If you would like some help getting your forecasts up and running during ESA – email me to set up a time!

freyao@vt.edu

Visit ecoforecast.org & neon4cast.org