

April 12, 2024 Education Working Group Call

Attendees: Antoinette Abeyta, Cazimir Kowalski, Jody Peters, John Zobitz, Jason McLachlan, Saeed Shafiei Sabet, Mary Lofton

Regrets: Rachel Torres

Agenda/Notes:

1. Look at the histories of TCUs and HBCUs
 - a. Jody started to compile info about TCUs history on [this Google doc](#)
 - b. For next month she will compile inf about HBCUs
 - c. Feel free to add information or links that you may know about
 - d. Looking at the AIHEC map, there are no TCUs in the east. But if you want to know more about the Indigenous education in the east, Ryan Emanuel's new book is a good one to read. Ryan is an EFI collaborator and Lumbee just had a book come out that talks about Lumbee experience in the east. On the Swamp: Fighting for Indigenous Environmental Justice: <https://uncpress.org/book/9781469678320/on-the-swamp/>

2. Discussion of chatGPT/AI in the classroom - articles and papers shared
 - a. I'm a student. You have no idea how much we're using chatGPT <https://www.chronicle.com/article/im-a-student-you-have-no-idea-how-much-were-using-chatgpt>
 - i. Students are using it to develop materials so they can learn from it and can learn in different ways than writing essays
 - b. Don't ban chatGPT in schools. Teach with it <https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html>
 - i. Use it as a tool to evaluate students submissions
 - ii. Teach in different ways
 - iii. Teach how to use it effectively. How to constructively think about the output that is given
 - c. This is how AI will transform how science gets done <https://www.technologyreview.com/2023/07/05/1075865/eric-schmidt-ai-will-transform-science/>
 - i. Example of FourCastNet which is an AI model that uses tens of terabytes of Earth system data and can predict the next two weeks of weather tens of thousands of times faster and more accurately than current forecasting methods. Used in conjunction with the digital twin Earth-2
 - d. A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education, <https://doi.org/10.1007/s12195-022-00754-8>
 - e. John Zobitz's Class Ethics Statement,

- f. Basham, T., Hughes-Seefeldt, R., Kovacs, J., Sitvarin, M., Sollberger, D. (2023). [AI Tools in Pedagogy](#). 2023 BIOME Institute, QUBES Educational Resources. [doi:10.25334/D5SE-TK10](https://doi.org/10.25334/D5SE-TK10)
 - i. If students are using AI, Professors can also.
 - ii. Provides syllabus statement examples starting on pg 13
 - iii. AI Lit Review tools from Texas A&M
<https://tamu.libguides.com/c.php?g=1289555>
- g. The Use of Generative Artificial Intelligence In Teaching and Learning
https://teaching.ucla.edu/resources/ai_guidance/
 - i. We will provide strategies for adopting AI technologies in a responsible, ethical manner, and innovating within each discipline, major, and course. Exploring and communicating about the opportunities and limitations to using these tools will allow instructors and students to critically think about how knowledge is created.
- h. Gradescope - <https://www.gradescope.com/> . Can scan all exams and it will group responses together so you can apply comments to each group. Doesn't do the grading for you, but groups things for you.
 - i. If it works well, could help to make grading more consistent (e.g., with 70 responses, could group them together to avoid the type of grading that starts at the beginning of grading to the end)
- i. Jason's AI example with abstracts in ML conferences - wording that comes from AI shot up after AI was available
- j. From John's experience teaching data science this semester, he can tell what students use chatGPT - the big tell is what pipe the students used. chatGPT uses the pipe, John uses the updated one
- k. With teaching R, AI can help reduce the technical side of coding, so could possibly reduce the technical side, but still want to teach model building
- l. Want to teach flexibility - how do changing the steps taken affect the output - group, order, summarize - if you change the order do you get a different answer
- m. Accessibility - it doesn't work if you don't have access to the internet. There are opportunities for reducing barriers and creating other barriers
- n. Can have it write the code, but don't always get the comments so you have mysterious biases that you don't know about. It is concerning for making decisions using black boxes where you don't know what is going on.
- o. If you ask students to get the answer - let them get the answer from chatGPT, but then ask them to interpret the code or interpret the answer. Doing the exercises in person where you can talk through it is a cool way
- p. Kann Academy - Sal Kann has teamed up with OpenAI to get algorithms that gets the math right, but asks questions that helps students think through the answers.
- q. Data Feminism book John's class is reading. Data science is expected to be objectional, but need to recognize that there is emotional input as well
 - i. Want students to author their own voice in coding (and writing)

- r. Interesting exercise - if you were to write the code, would you write it as chatGPT writes it. Thinking about it more personally and from an individual perspective
 - s. Want students to understand that chatGPT is not always true, it is replicating information that is out there.
 - t. Helping students learn the limitations to AI/chatGPT is really powerful.
 - u. Jody used chatGPT to explore the results when asking about setting up a forecast for the aquatics theme of the NEON Forecasting challenge and what process based models are recommended for forecasting water temp for the aquatics theme. If you are interested in checking out the output is listed below.
3. If we have time on the call, we could do this self-reflection in person during the call. This was suggested on the Jan call to help the group think about biases when teaching other students data science
- a. When and how did you learn about data and how to collect and analyze it?
 - b. Fill in this Google doc with your reflections

3-12-24 Jody's questions for chatGPT

How would you set up an forecast for the aquatics theme of the NEON Ecological Forecasting Challenge found here: <https://projects.ecoforecast.org/neon4cast> -ci/

Answer

Setting up a forecast for the aquatics theme of the NEON Ecological Forecasting Challenge involves several steps. Here's a general guide to get you started:

1. **Understand the Challenge**: Begin by thoroughly reading through the documentation provided on the NEON Ecological Forecasting Challenge website. Understand the objectives, evaluation criteria, data availability, and any specific requirements or guidelines for the aquatics theme.
2. **Data Exploration**: Familiarize yourself with the data provided for the aquatics theme. This may include time-series data on various aquatic parameters such as water temperature, pH, dissolved oxygen, chlorophyll concentration, etc. Understand the format of the data, any missing values, and potential data quality issues.
3. **Explore Historical Patterns**: Analyze historical data to identify any patterns, trends, or seasonal variations in the aquatic parameters. This will help you understand the underlying dynamics of the ecosystem and guide your forecasting approach.
4. **Select Forecasting Models**: Choose appropriate forecasting models based on the nature of the data and the forecasting horizon. Common models for time-series forecasting include ARIMA, SARIMA, exponential smoothing methods (e.g., Holt-Winters), and machine learning algorithms such as random forests, gradient boosting, or neural networks.

5. **Data Preprocessing**: Prepare the data for modeling by handling missing values, scaling, normalization, and feature engineering if necessary. Split the data into training and validation sets for model evaluation.
6. **Model Training**: Train the selected forecasting models using the training data. Tune model hyperparameters if applicable, and experiment with different algorithms to find the best-performing ones.
7. **Model Evaluation**: Evaluate the trained models using the validation set. Use appropriate evaluation metrics such as mean absolute error (MAE), root mean square error (RMSE), or others specified in the challenge guidelines. Compare the performance of different models and select the best-performing one.
8. **Generate Forecasts**: Once you have selected the best model, use it to generate forecasts for the specified forecasting horizon. Ensure that the forecasts are in the required format and align with the evaluation criteria specified in the challenge guidelines.
9. **Submit Forecasts**: Submit your forecasts through the designated submission portal according to the instructions provided in the challenge guidelines. Double-check that your submissions meet all the requirements and deadlines.
10. **Iterate and Improve**: Continuously refine your forecasting approach based on feedback from the challenge organizers and your own observations. Experiment with different modeling techniques, feature engineering strategies, or data preprocessing methods to improve the accuracy of your forecasts.

Remember to document your methodology, assumptions, and any challenges encountered during the forecasting process. Good luck with your forecasting efforts for the NEON Ecological Forecasting Challenge!

what process based models would you recommend for forecasting water temperature for the aquatics theme of the NEON Ecological Forecasting Challenge?

<https://projects.ecoforecast.org/neon4cast> -ci/

For forecasting water temperature in the aquatics theme of the NEON Ecological Forecasting Challenge, process-based models are particularly suitable due to their ability to simulate the underlying physical processes driving temperature dynamics in aquatic ecosystems. Here are some process-based models that you may consider:

- **CE-QUAL-W2**: CE-QUAL-W2 is a two-dimensional, laterally averaged, hydrodynamic and water quality model that simulates the complex interactions between hydrodynamics, temperature, and water quality parameters in lakes and reservoirs. It can be used to

forecast water temperature based on inputs such as meteorological data, lake geometry, and water quality parameters.

- **DYRESM-CAEDYM:** DYRESM (DYnamic REservoir Simulation Model) and CAEDYM (Computational Aquatic Ecosystem Dynamics Model) are two coupled models commonly used for simulating the dynamics of temperature, dissolved oxygen, nutrients, and other water quality variables in lakes and reservoirs. They can be applied to forecast water temperature by simulating heat exchange processes, mixing dynamics, and thermal stratification phenomena.
- **EFDC:** The Environmental Fluid Dynamics Code (EFDC) is a widely used hydrodynamic and water quality model capable of simulating temperature dynamics in aquatic systems. It considers factors such as solar radiation, wind-driven mixing, heat exchange with the atmosphere, and thermal stratification to forecast water temperature accurately.
- **SWAT:** The Soil and Water Assessment Tool (SWAT) is a comprehensive watershed model that simulates the hydrology, water quality, and temperature dynamics in river systems. It can be applied to forecast water temperature by considering various factors such as streamflow, land use, soil properties, and climate inputs.
- **HEC-RAS:** The Hydrologic Engineering Center's River Analysis System (HEC-RAS) is a widely used hydraulic modeling software capable of simulating flow and temperature dynamics in rivers and streams. It can be applied to forecast water temperature by considering factors such as streamflow, channel geometry, heat exchange processes, and thermal pollution.

These process-based models vary in complexity and computational requirements, so it's essential to choose a model that aligns with the specific objectives and requirements of the NEON Ecological Forecasting Challenge. Additionally, ensure that you have access to the necessary input data (e.g., meteorological data, hydrological data, bathymetric data) required to run these models effectively.