Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science

David John Gagne

National Center for Atmospheric Research

Amy McGovern[1], Imme Ebert-Uphoff[2], and Ann Bostrom[3]

1. University of Oklahoma, 2. Colorado State University/CIRA,
3. University of Washington

July 5, 2022

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Motivation

Environmental Data Science (2022), 1: e6, 1–15
doi:10.1017/eds.2022.5

CAMBRIDGE UNIVERSITY PRESS

POSITION PAPER

Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science

Amy McGovern[1,6,*], Imme Ebert-Uphoff[2,3,6], David John Gagne II[4,6] and Ann Bostrom[5,6]

[1]School of Computer Science and School of Meteorology, University of Oklahoma, Norman, OK 73019, USA
[2]Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA
[3]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO 80523, USA
[4]Computational and Information Systems Laboratory, National Center for Atmospheric Research, Boulder, CO 80301, USA
[5]Evans School of Public Policy & Governance, University of Washington, Seattle, WA 98195, USA
[6]NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography
*Corresponding author. E-mail: amcgovern@ou.edu

Received: 14 December 2021; **Revised:** 22 February 2022; **Accepted:** 23 February 2022

Keywords: Artificial intelligence; climate; ethics; weather

Abstract

Given the growing use of Artificial intelligence (AI) and machine learning (ML) methods across all aspects of environmental sciences, it is imperative that we initiate a discussion about the ethical and responsible use of AI. In fact, much can be learned from other domains where AI was introduced, often with the best of intentions, yet often led to unintended societal consequences, such as hard coding racial bias in the criminal justice system or increasing economic inequality through the financial system. A common misconception is that the environmental sciences are immune to such unintended consequences when AI is being used, as most data come from observations, and AI algorithms are based on mathematical formulas, which are often seen as objective. In this article, we argue the opposite can be the case. Using specific examples, we demonstrate many ways in which the use of AI can introduce similar consequences in the environmental sciences. This article will stimulate discussion and research efforts in this direction. As a community, we should avoid repeating any foreseeable mistakes made in other domains through the introduction of AI. In fact, with proper precautions, AI can be a great tool to help *reduce* climate and environmental injustice. We primarily focus on weather and climate examples but the conclusions apply broadly across the environmental sciences.

Impact Statement

This position paper discusses the need for the environmental sciences community to ensure that they are developing and using artificial intelligence (AI) methods in an ethical and responsible manner. This paper is written at a general level, meant for the broad environmental sciences and earth sciences community, as the use of AI methods continues to grow rapidly within this community.

Paper: https://doi.org/10.1017/eds.2022.5

- Artificial Intelligence and Machine Learning (AI/ML) algorithms are being rapidly adopted by the environmental science community for a wide range of applications
- Ethical issues with broad AI/ML adoption have become more prominent recently, especially in areas like facial recognition, criminal sentencing, autonomous vehicles, and deep fakes
- Many environmental scientists are not concerned about ethics in AI because they believe their applications are being done for social good
- **Goal:** Discuss ethical issues around the use of AI/ML in the environmental scientists and why they need to be addressed before we deploy AI/ML more broadly

Ways AI can go wrong for Environmental Sciences

Ways in which AI can go wrong for environmental sciences

Issues related to training data:

1. Non-representative training data, including lack of geo-diversity
2. Training labels are biased or faulty
3. Data is affected by adversaries

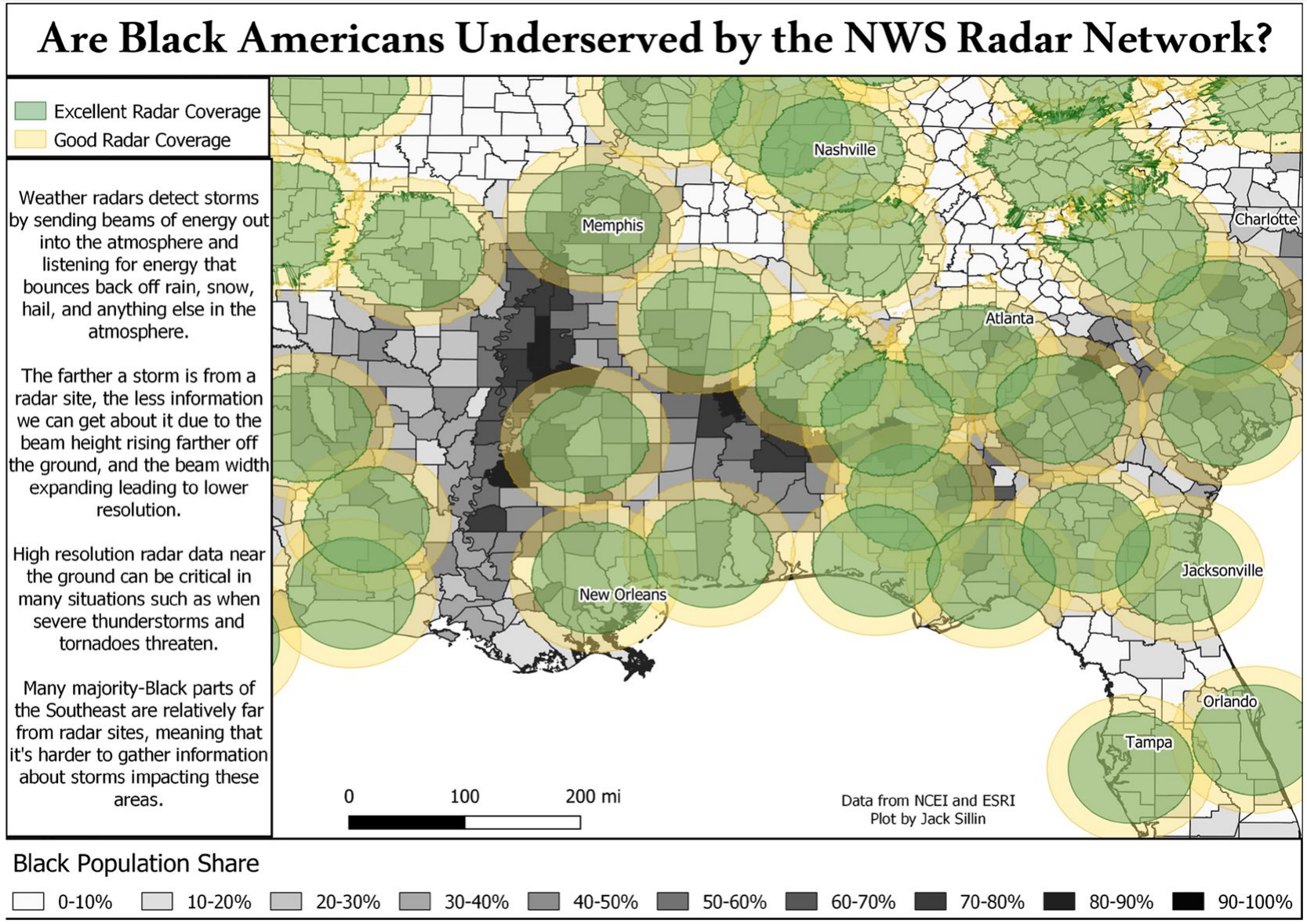
Issues related to AI models:

1. Model training choices
2. Algorithm learns faulty strategies
3. AI learns to fake something plausible
4. AI model used in inappropriate situations
5. Non-trustworthy AI model deployed
6. Lack of robustness in the AI model

Other issues related to workforce and society:

1. Globally applicable AI approaches may stymie burgeoning efforts in developing countries.
2. Lack of input or consent on data collection and model training
3. Scientists might feel disenfranchised.
4. Increase of CO_2 emissions due to computing

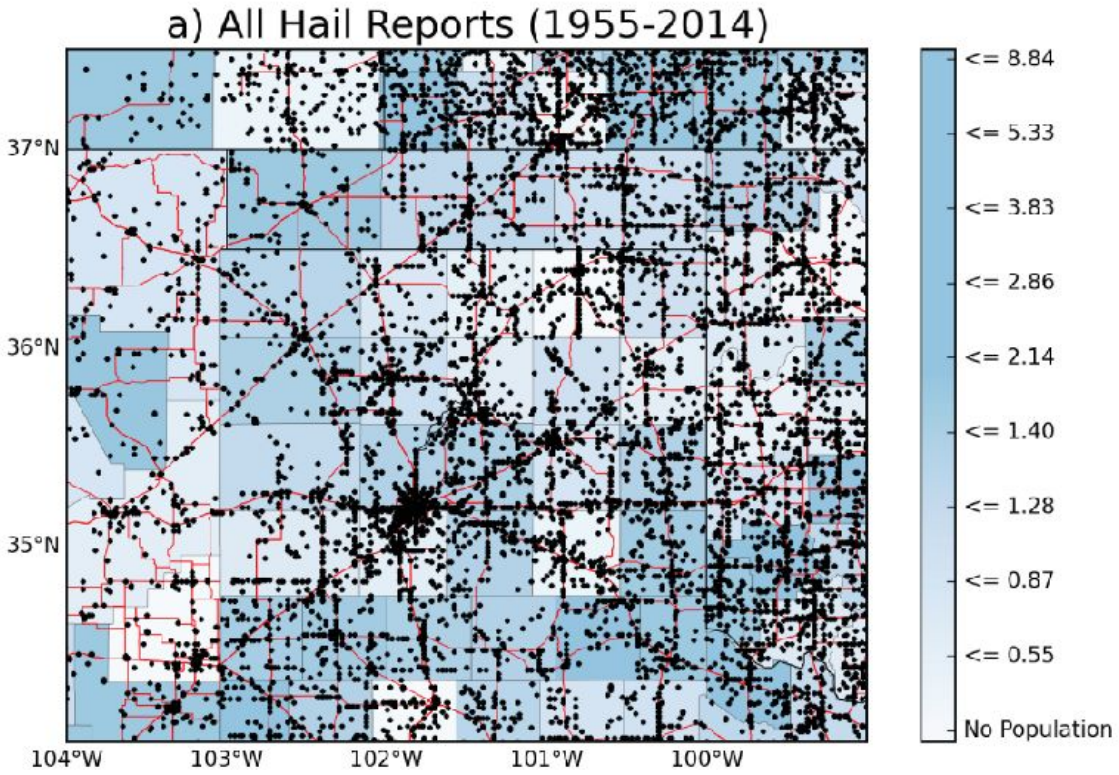
Non-representative training data



From Jack Sillin @JackSillin:

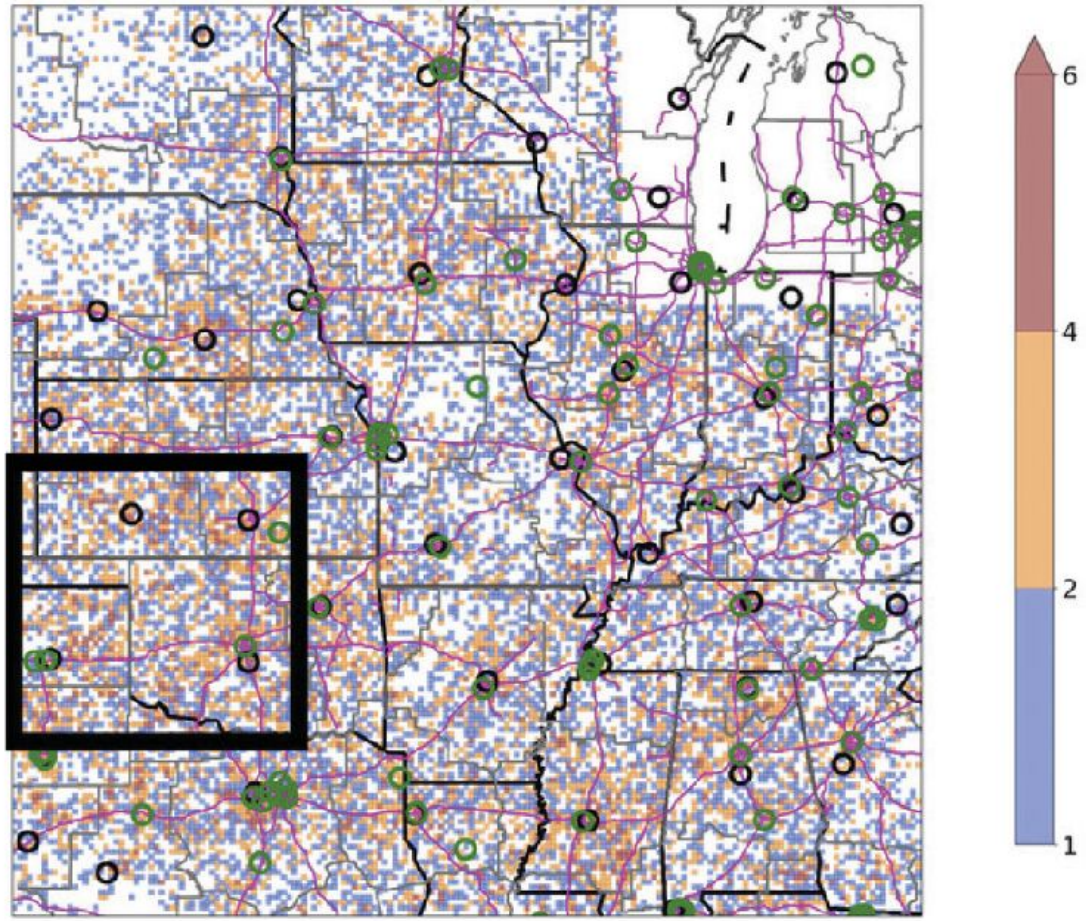
https://twitter.com/JackSillin/status/1372957704138981378?s=20

Biased or faulty training labels



(a) Hail reports follow population

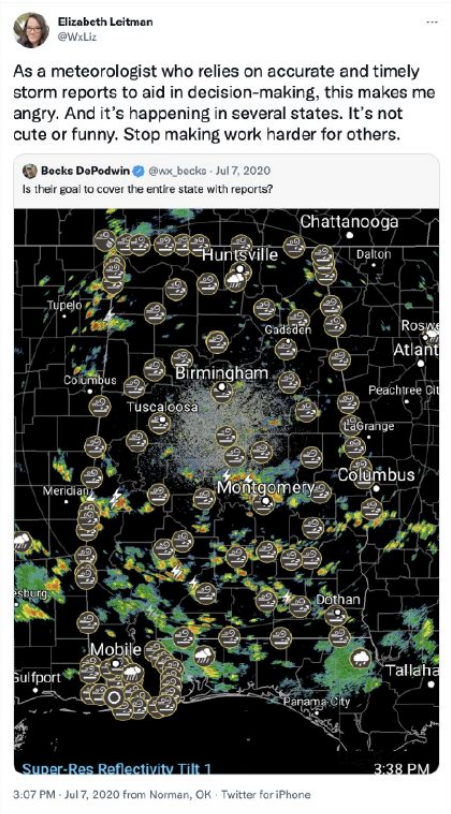
Allen, J. T., and M. K. Tippett, 2015: The characteristics of United States hail reports: 1955–2014. Electronic J. Severe Storms Meteor., 10 (3), 1–31.



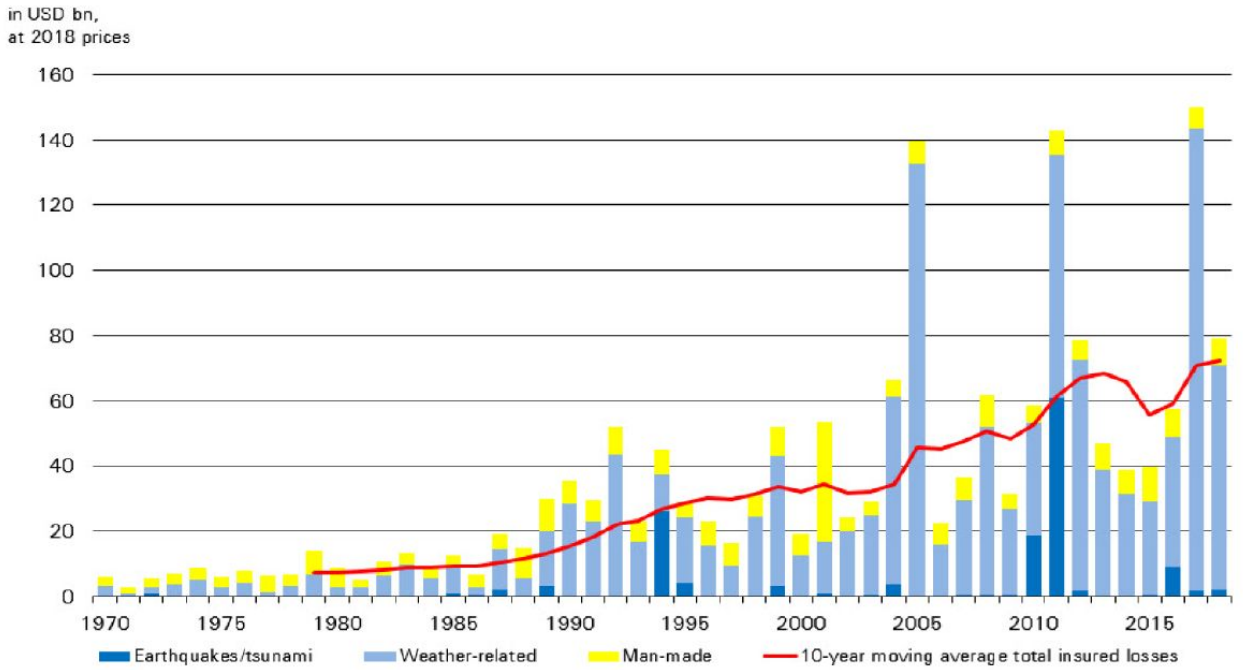
(b) Tornado reports follow population

Potvin, C. K., Broyles, C., Skinner, P. S., Brooks, H. E., & Rasmussen, E. (2019). A Bayesian Hierarchical Modeling Framework for Correcting Reporting Bias in the U.S. Tornado Database, *Weather and Forecasting*, *34*(1), 15-30. Retrieved Jul 24, 2021

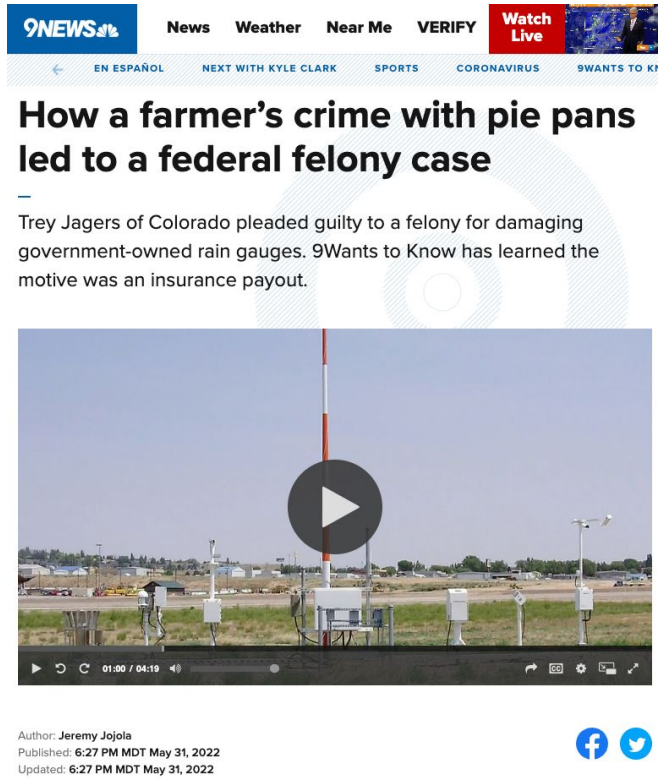
Adversarial Data



(a) Screenshot of hacked data from mPing from Cappucci (2020)



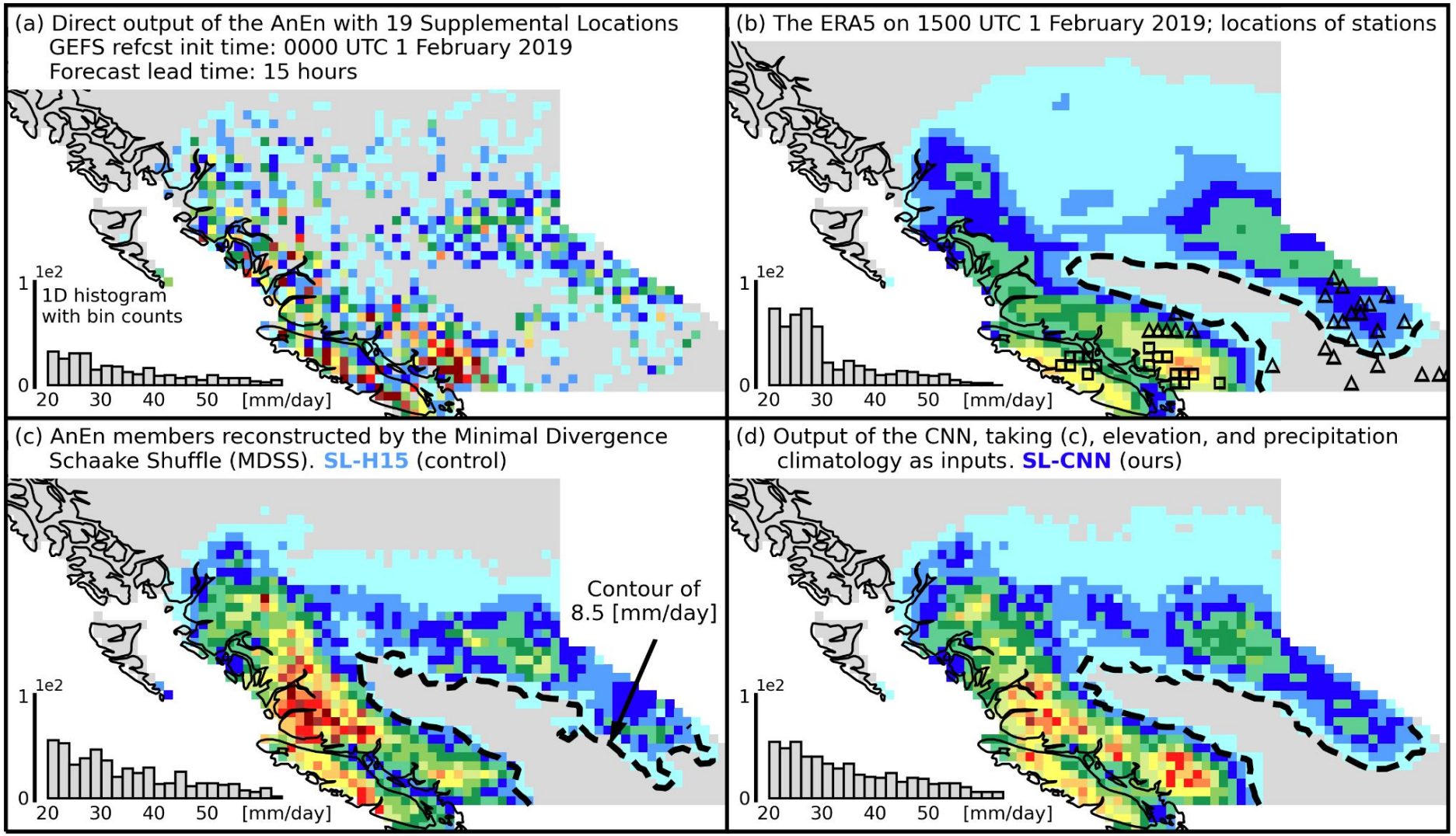
(b) Catastrophe related insured losses from 1970 to 2018 from (RE, 2018; Scotti, 2019). With at least 10% of weather-based insurance claims estimated as fraudulent (Sorge, 2018), there is increasing motivation for humans to report incorrect severe weather data



https://www.9news.com/article/news/crime/colorado-farmer-pie-tins-rain-gauges-federal-felony/73-1638f8a4-967a-4f12-af11-d6903e8b5d0d

Model Training Targets

Spatial Consistency: ensuring forecast variation matches spatial variation and properties



Sha, Y., D. J. Gagne II, G. West, and R. Stull, 2022: A hybrid analog-ensemble, convolutional-neural-network method for post-processing precipitation forecasts. *Mon. Weather Rev.*, **-1**, https://doi.org/10.1175/MWR-D-21-0154.1.

Untrustworthy AI models deployed

Figure 7. Three vastly different damage predictions for the same hypothetical 7.0-magnitude Seattle area earthquake delivered by different versions of the same AI system. Figure from Fink (2019), crediting the Seattle Office of Emergency Management.

Other Issues

- **Globally applicable AI approaches may stymie local forecasting products**
  E.g., private weather app missing warnings for local hazards
- **Lack of input or consent on data collection and model training**
  Data choices can perpetuate biases and power dynamics
- **Scientists feeling disenfranchised**
  Over focus on AI research can cause neglect of other important topics/methods and disenfranchises scientists without AI skills
- **Increase of CO2 emissions due to computing**
  Large AI models and data storage/transmission require a lot of compute and power resources

TAI4ES Summer School Trustathon: Seeing Value through Roleplaying

User role description: Create personas about end-users of a given challenge problem algorithm

Department of Transportation Official

Background: Wind speeds are an important factor for transportation officials when making decisions about closing and (reopening) bridges. Strong winds can cause driving conditions to deteriorate, making driving on them incredibly dangerous. Crashes are not only dangerous to those driving, but they could also block important routes for emergency response vehicles. The decision to close the bridges is also very important because they are needed for people to evacuate, especially from barrier islands, so closing the bridge means potentially taking away some people's ability to evacuate from a storm.

Key user needs: The transportation official needs extremely precise data on wind speeds so they can effectively walk the line between keeping pathways open for evacuation and making sure driving conditions are safe.

User role developed by Chris Wirz

Daily Value Reflection

How does each kind of ML analysis provide insight (if any) about the value the ML provides for the user persona?

Trustworthy AI Approach Evaluation:

Participants perform different kinds of ML evaluations on pre-trained models each day

- Verification statistics
- XAI (shallow and deep)
- Case studies and failure modes
- Uncertainty Quantification

Event (June 27-30) website: https://www2.cisl.ucar.edu/events/tai4es-2022-summer-school

Slides now available. Videos will be released soon

Summary

Ways in which AI can go wrong for environmental sciences

Issues related to training data:

1. Non-representative training data, including lack of geo-diversity
2. Training labels are biased or faulty
3. Data is affected by adversaries

Issues related to AI models:

1. Model training choices
2. Algorithm learns faulty strategies
3. AI learns to fake something plausible
4. AI model used in inappropriate situations
5. Non-trustworthy AI model deployed
6. Lack of robustness in the AI model

Other issues related to workforce and society:

1. Globally applicable AI approaches may stymie burgeoning efforts in developing countries.
2. Lack of input or consent on data collection and model training
3. Scientists might feel disenfranchised.
4. Increase of CO_2 emissions due to computing