Modeling

Mathematical Framing of Real-World Problems

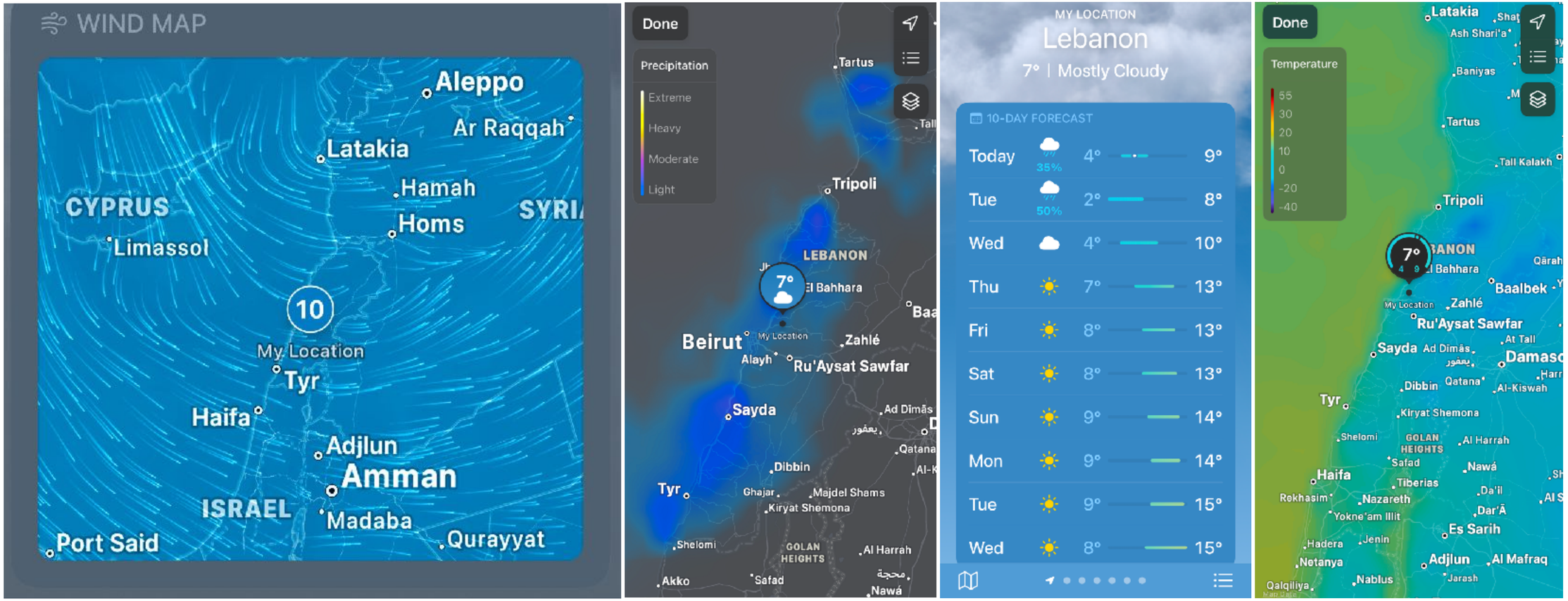
Joseph Bakarji

The ‘model’ machine learning scientist

- **Curiosity** to find learn about models and explore datasets.
- **Initiative** to take part in projects with an unfamiliar area of expertise.
- **Ability to identify** whether a problem requires machine learning or not.
- **Ability to quickly learn** enough background about the subject matter required for you to develop a sound hypothesis. e.g. how do you deal with weather data?
- **Ability to distill** complex real-world problems into mathematical representations, answering questions such as: what is the hypothesis, the are fitting parameters etc.
- **Coding maturity** to be able to test existing models, and develop new solutions. And having found a solution, the ability to package it into a product others will use.

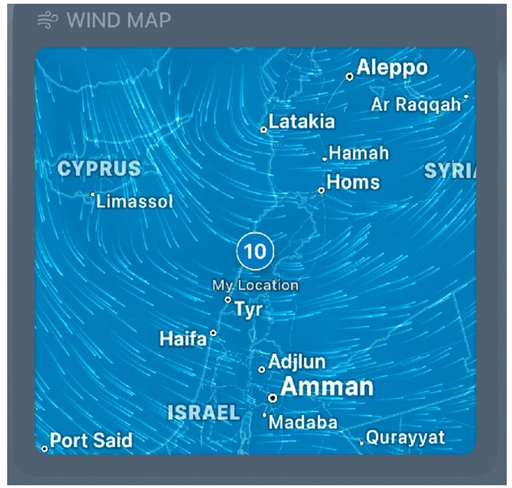
Here's an example

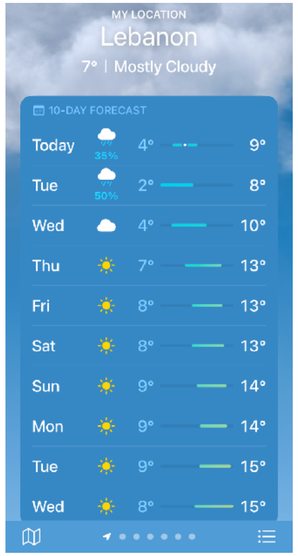
Weather data

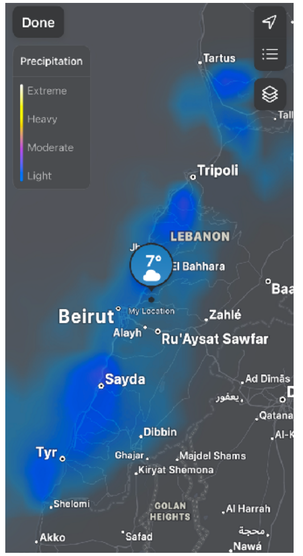


Good questions to ask about the weather

Weather data



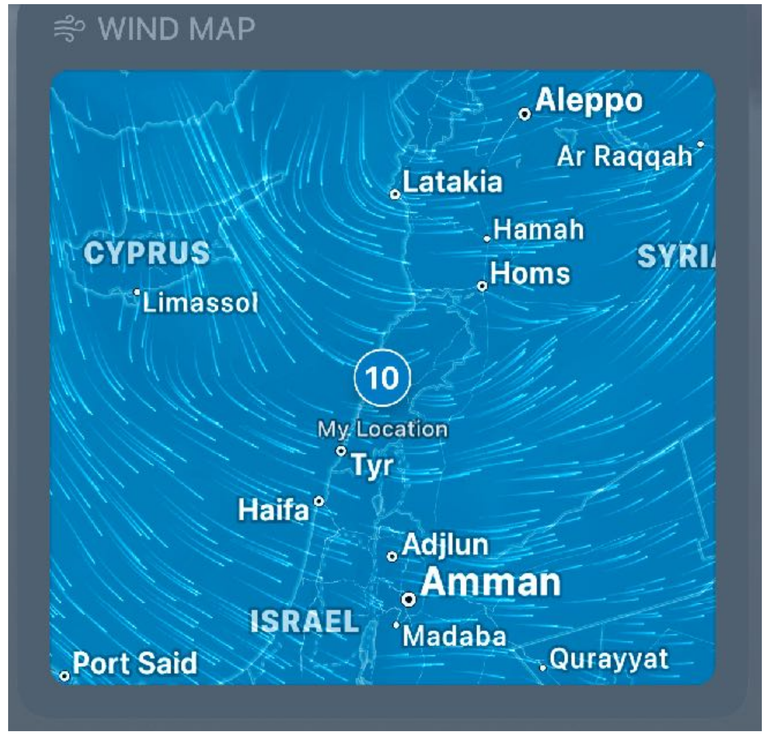




Wind

How is the data collected?

What does it look like?



What are the measurements of interest?

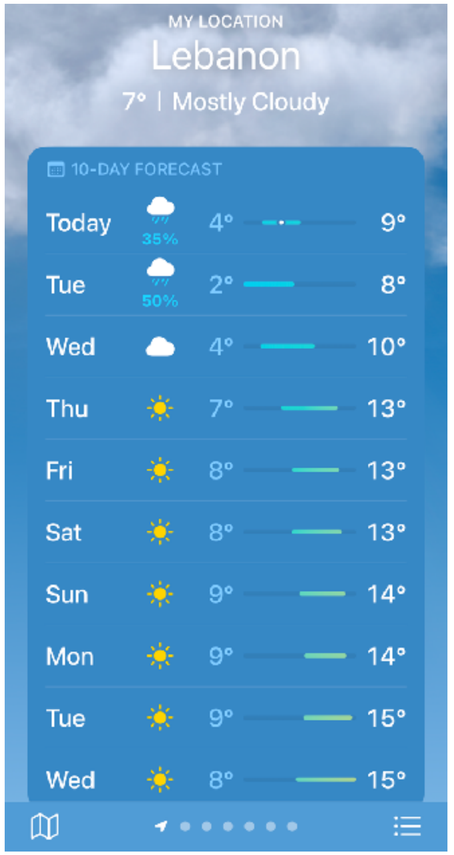
What are existing models?

What are existing models?

Temperature

How is the data collected?

What does it look like?



What are the measurements of interest?

What are existing models?

What are existing models?

Temperature

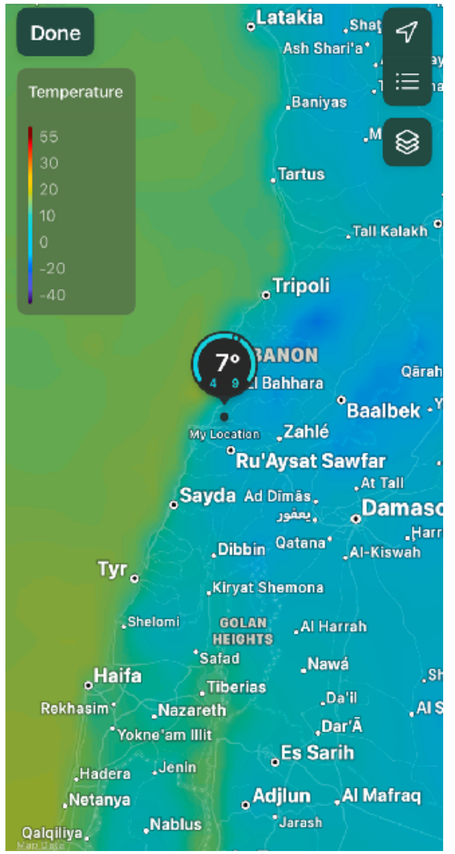
How is the data collected?

What does it look like?

What are the measurements of interest?

What are existing models?

What are existing models?



Precipitation

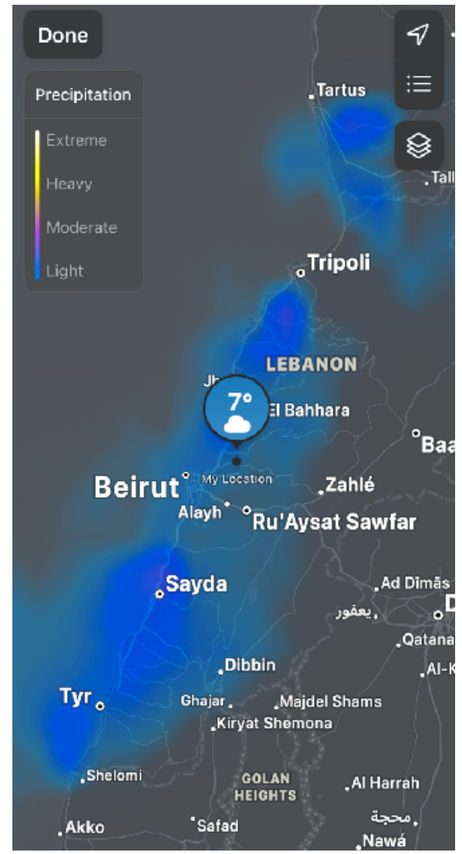
How is the data collected?

What does it look like?

What are the measurements of interest?

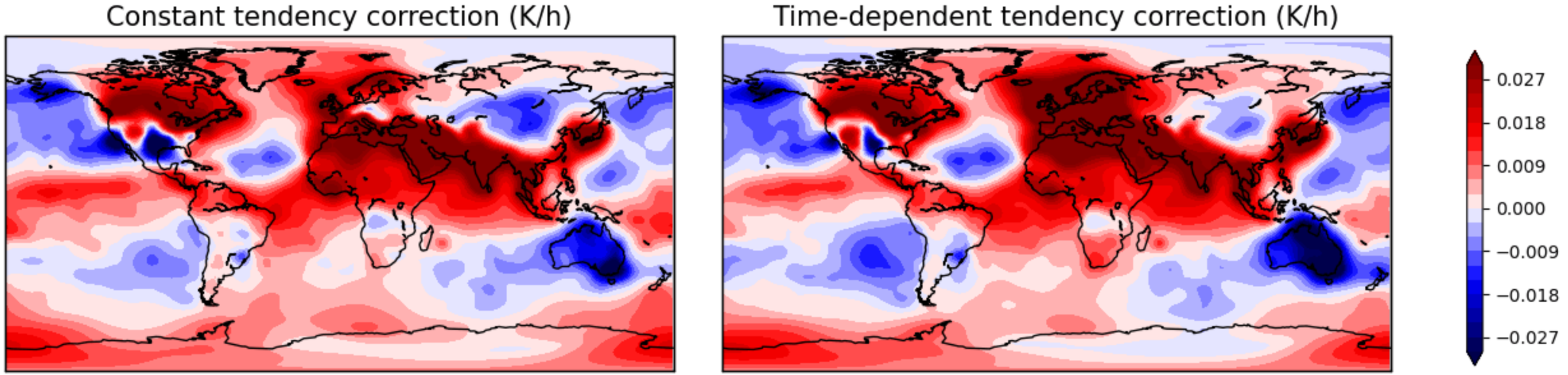
What are existing models?

What are existing models?



Who models weather? And do they use ML?

Google It: machine learning for weather prediction



Machine learning is set to play a growing role in numerical weather prediction, but physics-based forecasting techniques will continue to be important

European Centre for Medium-Range Weather Forecasts is currently pursuing a three-pronged approach:

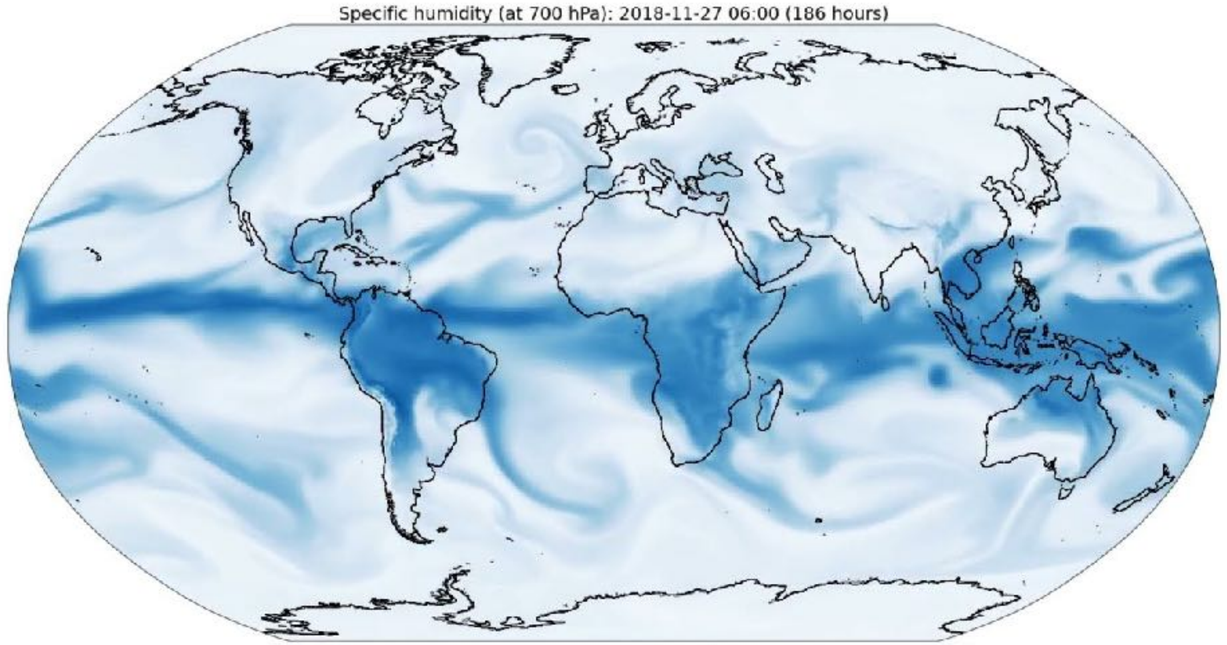
- The use of machine learning to boost traditional techniques,
- The development of a pure machine-learning forecasting model,
- An experimental use of machine learning on weather observations to build a forecasting system.

For the longer term, we will start investigating the design of a **foundation model** for weather and climate.

https://www.ecmwf.int/en/about/media-centre/news/2024/machine-learning-play-growing-role-weather-forecasting-says-dg#:~:text=Machine%20learning%20is%20well%20suited,and%20to%20produce%20global%20forecasts.

Big Tech tackling non-tech problems

GraphCast: AI model for faster and more accurate global weather forecasting - Deep Mind



- Forecasts typically rely on Numerical Weather Prediction (NWP), which begins with carefully defined **physics equations**, which are then translated into computer algorithms run on supercomputers. While this traditional approach has been a triumph of science and engineering, designing the equations and algorithms is time-consuming and requires deep expertise, as well as costly compute resources to make accurate predictions.

- GraphCast is a weather forecasting system based on machine learning and Graph Neural Networks (GNNs), which are a particularly useful architecture for processing spatially structured data.

https://deepmind.google/discover/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/

Research papers

Open Access Review

Machine Learning Methods in Weather and Climate Applications: A Survey

by Liuyi Chen [1], Bocheng Han [1], Xuesong Wang [2], Jiazhen Zhao [3], Wenke Yang [1] and Zhengyi Yang [1,*]

[1] School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia
[2] Commonwealth Scientific and Industrial Research Organisation, Data 61, Sydney, NSW 2015, Australia
[3] Key Laboratory of Meteorological Disaster, Nanjing University of Information Science and Technology, Nanjing 210044, China
* Author to whom correspondence should be addressed.

Appl. Sci. **2023**, *13*(21), 12019; https://doi.org/10.3390/app132112019

Submission received: 9 September 2023 / Revised: 22 October 2023 / Accepted: 28 October 2023 / Published: 3 November 2023

(This article belongs to the Special Issue **Methods and Applications of Data Management and Analytics**)

Download Browse Figures Versions Notes

Abstract

https://www.mdpi.com/2076-3417/13/21/12019

Research papers

machine learning for weather prediction

About 589,000 results (**0.10** sec)

Machine learning in **weather prediction** and climate analyses—applications and perspectives [PDF] mdpi.com

B Bochenek, Z Ustrnul - Atmosphere, 2022 - mdpi.com

… The main goal of this study is to present a review of the **machine learning** … of **machine learning** techniques as a new method that helps to solve important and complex issues in **weather** …

☆ Save 99 Cite Cited by 202 Related articles All 9 versions »

Machine learning for applied **weather prediction** [PDF] ieee.org

SE Haupt, J Cowie, S Linden… - 2018 IEEE 14th …, 2018 - ieeexplore.ieee.org

… of applying **machine learning** to **weather** forecasting challenges. The Dynamic Integrated foreCasting (DICast®) System was one of the first automated **weather** forecasting engines. It is …

☆ Save 99 Cite Cited by 71 Related articles All 4 versions

[HTML] Comparison of **machine learning** methods for photovoltaic power forecasting based on numerical **weather prediction** [HTML] sciencedirect.com

D Markovics, MJ Mayer - Renewable and Sustainable Energy Reviews, 2022 - Elsevier

… This study compares 24 **machine learning** models for deterministic day-ahead power forecasting based on numerical **weather predictions** (NWP), tested for two-year-long 15-min …

☆ Save 99 Cite Cited by 235 Related articles All 6 versions

[PDF] **Machine learning** applied to **weather** forecasting [PDF] stanford.edu

M Holmstrom, D Liu, C Vo - Meteorol. Appl, 2016 - cs229.stanford.edu

… intelligence used mainly **machine learning** techniques, mostly … three papers on **machine learning** for **weather prediction** we … to be the popular **machine** learning model choice for **weather** …

☆ Save 99 Cite Cited by 173 Related articles All 2 versions »

Ask GPT

how is machine learning used for weather modeling and prediction

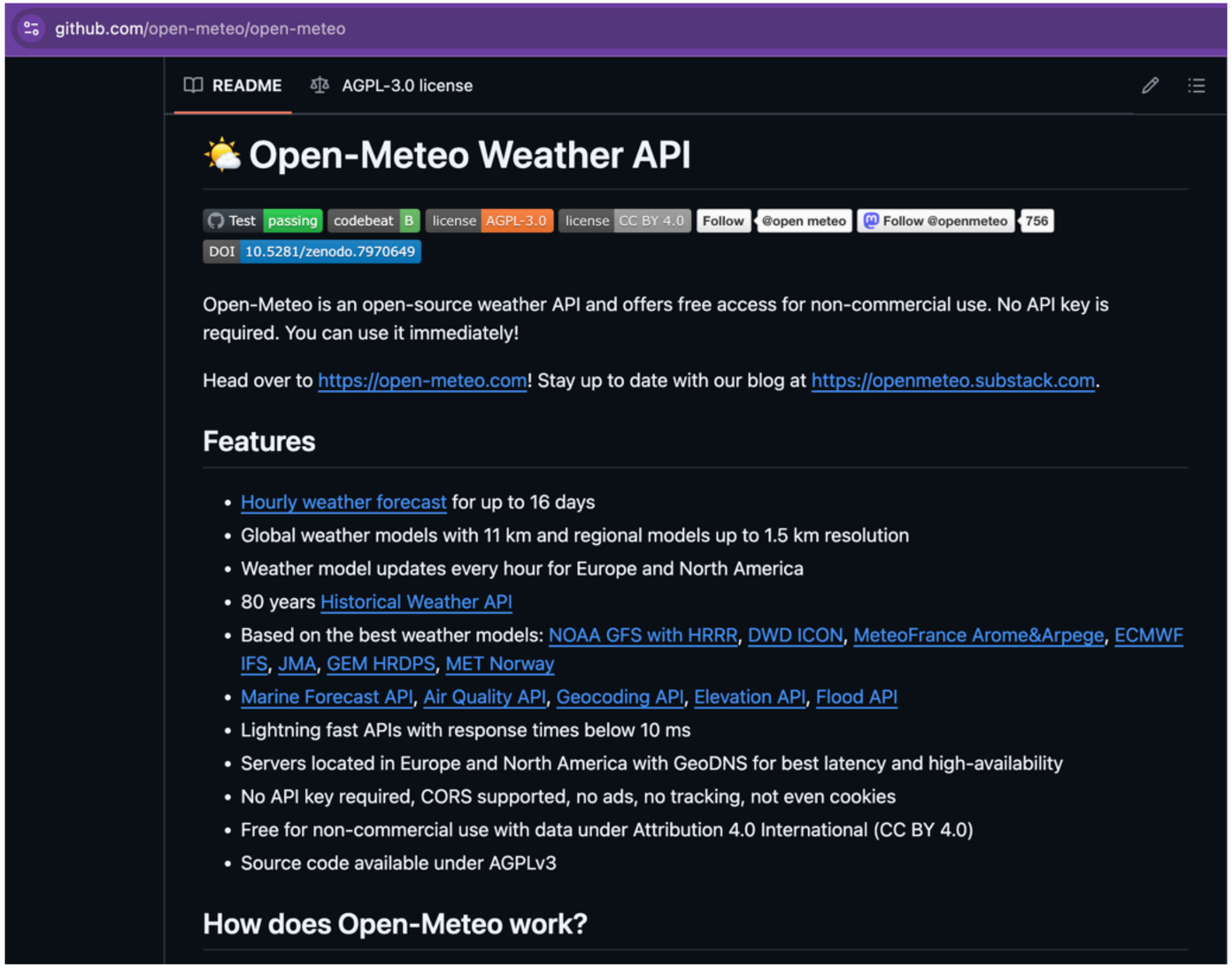
Machine learning (ML) has become an integral part of weather modeling and prediction, enhancing traditional methods by leveraging large datasets and providing more accurate, granular, and timely forecasts. Here's how ML is used in this field:

---

1. Data Integration and Preprocessing

- **Data Sources:** Weather models use diverse datasets, including satellite imagery, radar data, ground-based sensors, and historical climate records.
- **Feature Engineering:** ML models preprocess raw data to extract meaningful features, such as temperature gradients, wind patterns, or atmospheric pressure variations.

Start implementing!



Start implementing!

github.com/pangeo-data/WeatherBench

README | Code of conduct | MIT license

WeatherBench: A benchmark dataset for data-driven weather forecasting

🧯🧯🧯 WeatherBench 2 has been released. It provides an updated and much improved benchmark including more comprehensive and more easily accessible datasets.🧯🧯🧯

Binder

If you are using this dataset please cite

> Stephan Rasp, Peter D. Dueben, Sebastian Scher, Jonathan A. Weyn, Soukayna Mouatadid, and Nils Thuerey, 2020. WeatherBench: A benchmark dataset for data-driven weather forecasting. arXiv: https://arxiv.org/abs/2002.00469

This repository contains all the code for downloding and processing the data as well as code for the baseline models in the paper.

Reach out and ask

- After you do all your research, if you know someone who has expertise in the field, **ask them!** Your mechanical engineering friends, a professor at your university, an author of a paper, etc.
- Go back to pen and paper, and make sure you understand the modeling framework and assumptions.
- Implement a simple algorithm on a simple dataset if you can: it will improve your intuition about the data; even if it's fake data.
- ...

We did it all “wrong”