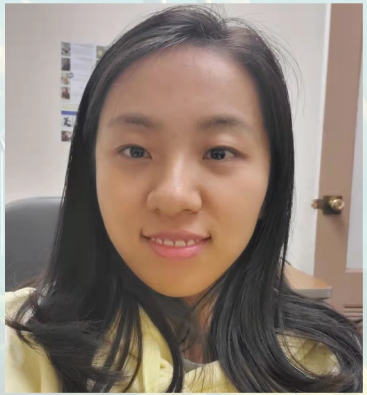
Image generated by DALL-E

CS663 Project Final Presentation

Wind Energy Prediction

Breezy Brains

Breezy Brains - Team Member



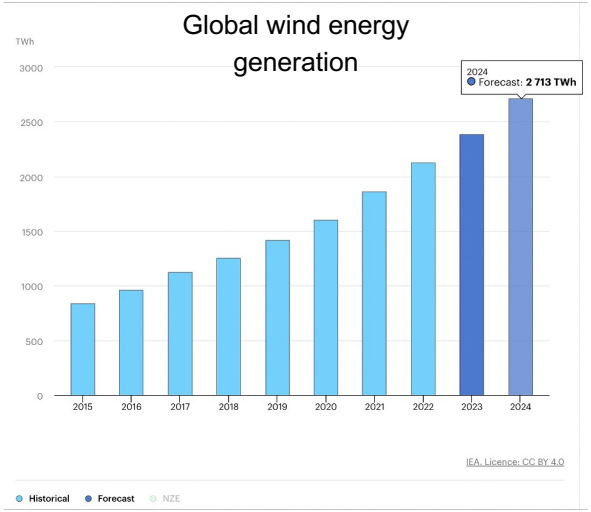
Qianru Wei

Lawrence Ng

Image generated by DALL-E

The Importance of Wind Energy Prediction

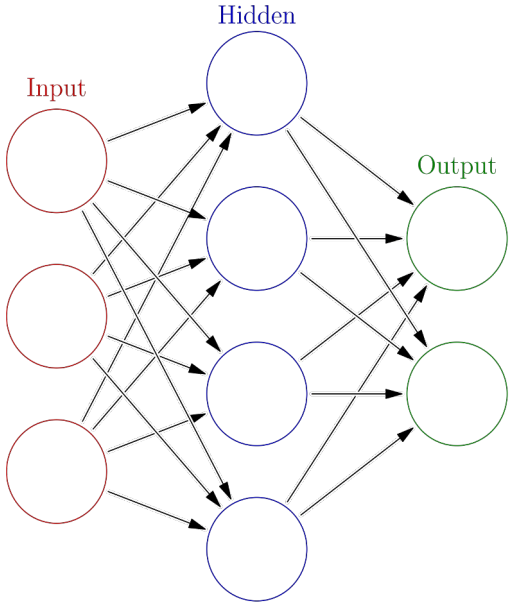
- Growing Demand for Renewable Energy
- Wind energy provided 10% of total electricity nationwide in the U.S. in 2022. (https://www.energy.gov/articles/us-department-energy-projects-steepe-growth-as-wind-power-sector#:~:text=Wind%20energy%20provided%2010%25%20of,most%20capacity%2C%20with%204%2C028%20MW.)
- Energy shortage problems



IEA, Wind power generation in the Net Zero Scenario, 2015-2030, IEA, Paris https://www.iea.org/data-and-statistics/charts/wind-power-generation-in-the-net-zero-scenario-2015-2030, IEA. Licence: CC BY 4.0

Scope and Potential Applications

- Develop ML models to predict **Wind Speed** based on historical data.
- Benchmark our model forecasts against **Pyrecast (NBM)** wind predictions.
- Convert the wind speed prediction to energy production prediction based on the wind turbine locations and types.
- Applications of Improved Predictions
  - Optimize energy allocation/usage
  - Reduce the cost of wind energy generation



Artificial Neural Network (ANN)

Previous Similar Work

Table 1. Neural networks and input features used in previous recent studies.

Reference, (Year)	Training Strategy	Input Features
Zameer et al [43], **(2017)**	ANN and Genetic Programming	Wind Speed, Wind D
Liu et al [48], **(2018)**	Elman neural network & LSTM	Wind Speed, Wavelet
Lin and Liu [42], **(2020)**	Isolation Forest Filtering & DNN	Wind Speed, Wind D
Devi et al [44], **(2020)**	Cuckoo Search Optimization, LSTM-EFG, EEMD	Wind Speed, Wind P
Acikgoz et al [46], **(2020)**	Extreme Machine Learning, ANN, K-folds	Wind Speed, Wind D Power
Niu et al [47], **(2020)**	Sequence-to-Sequence Modeling, Attention-based GRU	Wind Speed, Wind D
Yildiz et al [45], **(2021)**	Variational Mode Decomposition, CNN	Wind Speed, Wind P Images

https://doi.org/10.1371/journal.pone.0256381.t001

Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks

Henrique do Nascimento Camelo [a,*], Paulo Sérgio Lucio [a], João Bosco Verçosa Leal Junior [b], Paulo Cesar Marques de Carvalho [c], Daniel von Glehn dos Santos [b]

Forecasting with hybrid model

DOI: 10.1049/rpg2.12085

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Ultra-short-term multi-step wind power forecasting based on CNN-LSTM

Qianyu Wu[1] | Fei Guan[1] | Chen Lv[2] | Yongzhang Huang[1]

Ultra-short-term forecasting with CNN-LSTM Model

Project Workflow

ERA5

Historical Wind data (Data)

ERA5

PYRECAST

Verify the quality of the wind predictions

Prediction: Wind Speed

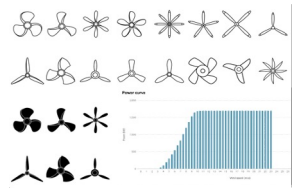
Wind Predictions at Turbine locations

Energy generated by Wind Turbines at certain location/Region

Applications

E.g.Optimize energy allocation/usage; Suggest locations to build new Wind Turbines

Wind Turbine locations (Data)



Power curve of different Wind Turbine models (Data)

ERA5 hourly data on single levels from 1940 to present

Dataset | Atmosphere (surface) | Atmosphere (upper air) | Global | Reanalysis

DATA DESCRIPTION	
Data type	Gridded
Projection	Regular latitude-longitude grid
Horizontal coverage	Global
Horizontal resolution	Reanalysis: 0.25° x 0.25° (atmosphere), 0.5° x 0.5° (ocean waves) Mean, spread and members: 0.5° x 0.5° (atmosphere), 1° x 1° (ocean waves)
Temporal coverage	1940 to present
Temporal resolution	Hourly
File format	GRIB
Update frequency	Daily

ERA5 Subsets:

- Chicago (The windy city)
  - Latitude x 5 (42.0 ~ 41.0)
  - Longitude x 5 (-88.0 ~ -87.0)
  - Timepoints x 8760 (Year 2023, hourly)
- Bay area
  - Latitude x 17 (40.0 ~ 36.0)
  - Longitude x 17 (-124.0 ~ -120.0)
  - Timepoints x 8760 (Year 2023, hourly)

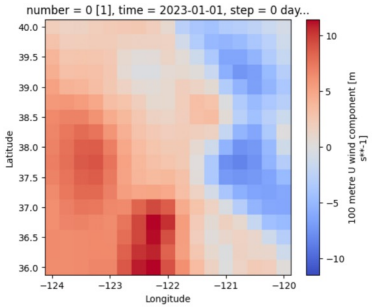
- No missing values
- u: wind toward east
- v: wind toward north
- 10 or 100: meters above ground level

https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

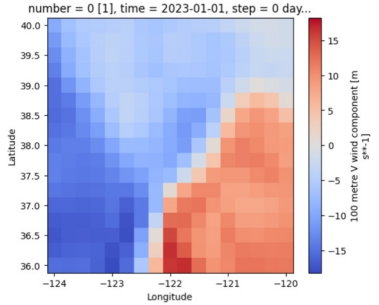
EDA - ERA5 Feature Engineering

- Convert Eastward Wind Speed, Northward Wind Speed into Actual Wind Speed and Direction(sin, cos)

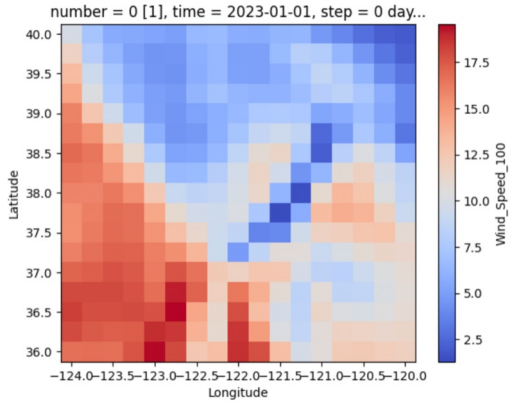
Eastward Wind Speed



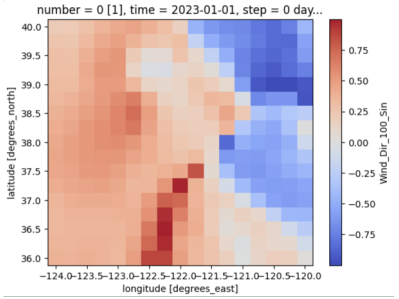
Northward Wind Speed



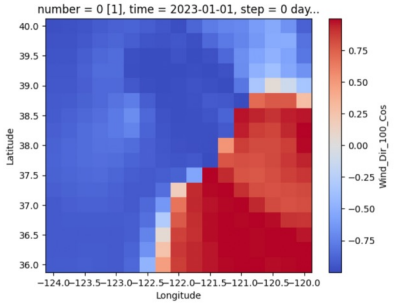
Actual Wind Speed



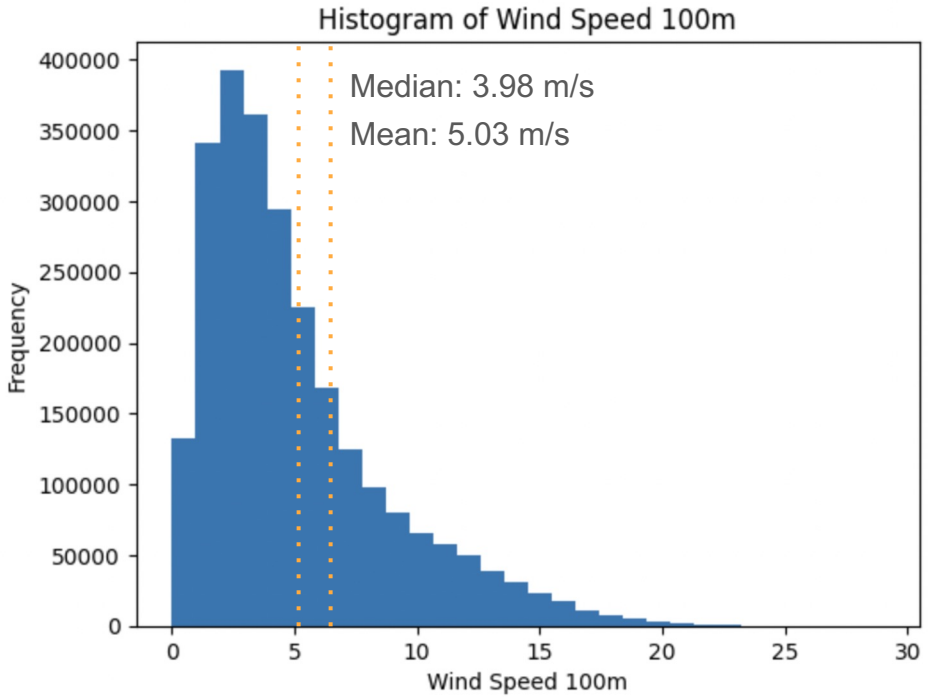
Actual Wind Direction Sin



Actual Wind Direction Cos



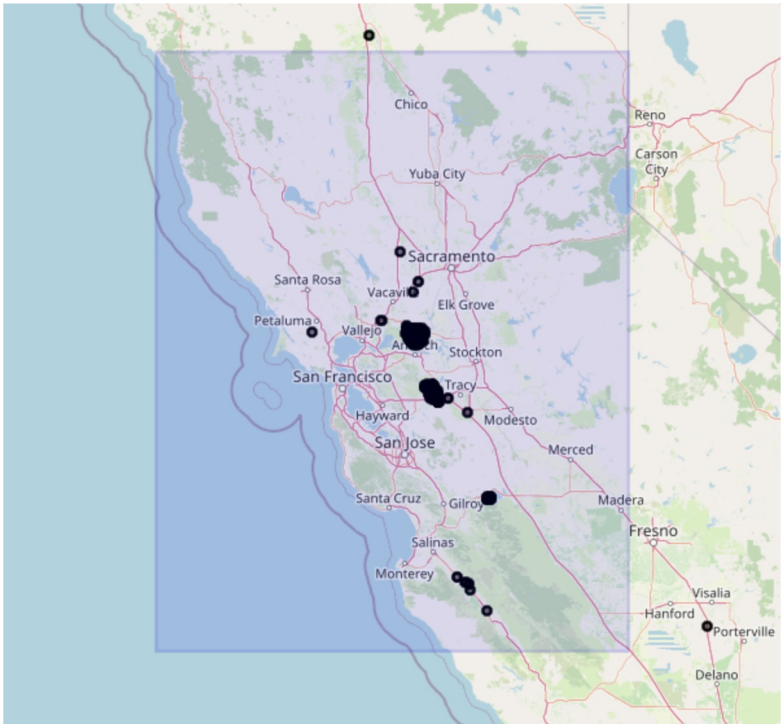
EDA - ERA5 Feature Engineering



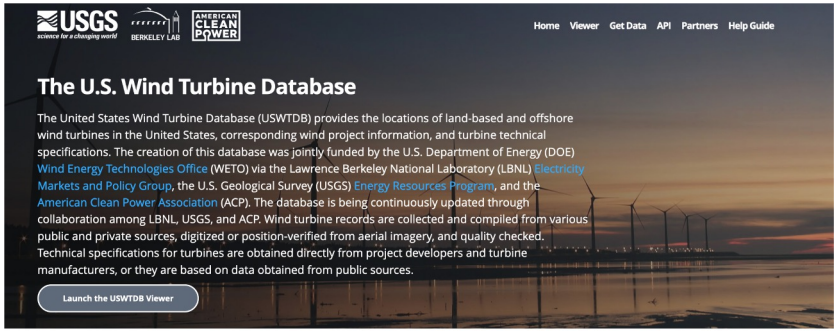

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
MultiIndex: 2531640 entries, (Timestamp('2023-01-01
Data columns (total 23 columns):
 #   Column                  Dtype
---  ------                  -----
 0   number                  int64
 1   step                    timedelta64[ns]
 2   surface                 float64
 3   valid_time              datetime64[ns]
 4   East_Wind_10m           float32
 5   North_Wind_10m          float32
 6   Temperature_2m          float32
 7   East_Neutral_Wind_10m   float32
 8   North_Neutral_Wind_10m  float32
 9   East_Wind_100m          float32
 10  North_Wind_100m         float32
 11  Wind_Speed_100          float32
 12  Wind_Speed_10           float32
 13  Neutral_Wind_Speed_10   float32
 14  Wind_Dir_100            float32
 15  Wind_Dir_10             float32
 16  Neutral_Wind_Dir_10     float32
 17  Wind_Dir_100_Sin        float32
 18  Wind_Dir_100_Cos        float32
 19  Wind_Dir_10_Sin         float32
 20  Wind_Dir_10_Cos         float32
 21  Neutral_Wind_Dir_10_Sin float32
 22  Neutral_Wind_Dir_10_Cos float32
dtypes: datetime64[ns](1), float32(19), float64(1),
memory usage: 270.7 MB
```

EDA - Turbine Location Map



San Francisco Bay Area



About the Database

In 2016, USGS, LBNL, and the American Wind Energy Association (AWEA, the predecessor of ACP) began collaborating on development of the USWTDB. Their goal was to create a joint product that would be more comprehensive and accurate than their individual wind turbine data sets. Federal agencies began using these combined data in April 2017, and in April 2018 the data were released to the public via this portal.

Chicago Area

EDA - Turbine Location Map

- Geopandas dataframe
- 73352 rows, 28 columns
- **Active** and inactive turbines

Missing Values

```
SF_area.t_model.value_counts().sort_values()

29-STALL-225      1
108               1
GE 1.5-77         1
GE 2.7-116        2
GE2.82-127        2
GE1.85-82.5       4
GE2.3-116        19
GE2.7-116        21
V47-0.66         31
V90-1.8          31
SWT-2.3-101      34
MWT62/1.0        41
SWT-2.3-93       50
GE1.7-100        51
V90-3.0          53
NTK65            73
V80-1.8          90
GE1.5-77        108
MM92            175
missing         314
Name: t_model, dtype: int64
```

```
SF_area.t_manu.value_counts()

GE Wind       208
Vestas        205
missing       182
REpower       175
Bonus         132
Siemens        84
Nordtank       73
Mitsubishi     41
Wincon          1
Norwin          1
Name: t_manu, dtype: int64
```

Duplicate Values

```
SF_area_active.t_model.value_counts()

t_model
MM92            175
108             162
GE1.5-77        108
V80-1.8          90
NTK65            73
V90-3.0          53
GE1.7-100        51
SWT-2.3-93       50
MWT62/1.0        41
SWT-2.3-101      34
V47-0.66         31
V90-1.8          31
GE2.7-116        21
GE2.3-116        19
GE1.85-82.5       4
GE2.82-127        2
GE 2.7-116        2
GE 1.5-77         1
29-STALL-225      1
Name: count, dtype: int64
```

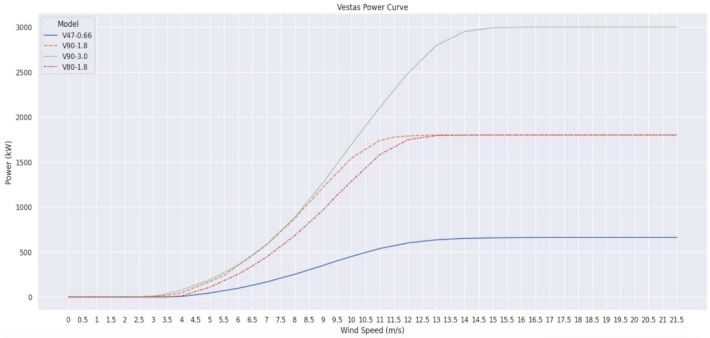
```
[80] SF_area_active.t_model.value_counts()

t_model
MM92            175
108             162
GE1.5-77        109
V80-1.8          90
NTK65            73
V90-3.0          53
GE1.7-100        51
SWT-2.3-93       50
MWT62/1.0        41
SWT-2.3-101      34
V90-1.8          31
V47-0.66         31
GE2.7-116        23
GE2.3-116        19
GE1.85-82.5       4
GE2.82-127        2
29-STALL-225      1
Name: count, dtype: int64
```

EDA - Turbine Energy Production

wind-turbine-models

https://www.thewindpower.net/turbines_manufacturers_en.php

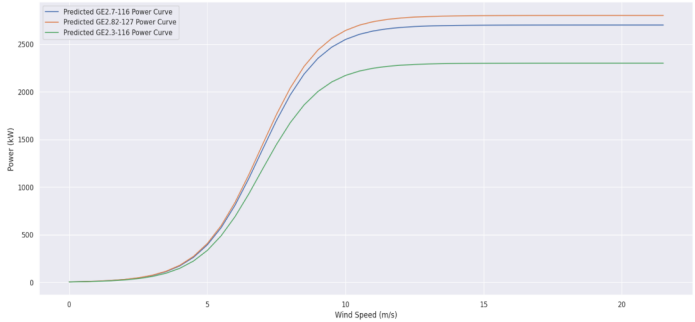



```
ge_df = df[df.Manufacturer == 'GE Wind']
ge_df
```

	Manufacturer	Model	Rated Power (kW)	Cut-in Wind Speed (m/s)	Rated Wind Speed (m/s)	Cut-out Wind Speed (m/s)	Rotor Diameter (m)	Swept Area (m^2)	Power Density 1 (W/m^2)	Power Density 2 (m^2/kW)	...	17	17.5	18	18.5	19	19.5	20	20.5	21	21.5
0	GE Wind	GE1.7-100	1700.0	3.0	10.0	23.0	100.0	7854.0	216.5	4.6	...	1700.0	1700.0	1700.0	1700.0	1700.0	1700.0	1700.0	1700.0	1700.0	1700.0
6	GE Wind	GE1.5-77	1500.0	3.5	12.0	25.0	77.0	4657.0	322.1	3.1	...	1500.0	1500.0	1500.0	1500.0	1500.0	1500.0	1500.0	1500.0	1500.0	1500.0
7	GE Wind	GE2.7-116	2700.0	3.0	NaN	NaN	116.0	10660.0	253.3	3.9	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	GE Wind	GE2.82-127	2800.0	3.0	8.0	25.0	127.0	12667.0	221.0	4.5	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	GE Wind	GE1.85-82.5	1850.0	3.0	13.0	25.0	82.5	5346.0	346.1	2.9	...	1850.0	1850.0	1850.0	1850.0	1850.0	1850.0	1850.0	1850.0	1850.0	1850.0
10	GE Wind	GE2.3-116	2300.0	3.0	10.4	22.0	116.0	10569.0	217.6	4.6	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

6 rows × 55 columns

https://en.wind-turbine-models.com/turbines



Model	GE1.7-100	GE1.5-77	GE2.7-116	GE2.82-127	GE1.85-82.5	GE2.3-116
0	0.0	0.0	0.0	0.0	0.0	0.0
0.5	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0
1.5	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0
2.5	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	71.312873	73.95409	6.0	60.748003
3.5	40.0	20.0	111.181456	115.299287	38.0	94.710129
4	100.0	46.33	171.881777	178.247769	75.0	146.41781
4.5	195.0	83.0	262.363335	272.080495	130.0	223.494693

EDA Summary

ERA5 Subsets

- Chicago and Bay Area subsets
- No missing values
- Eastward and Northward Wind info at 10, 100 meters
- Feature Engineering:
  - Actual Wind Speed
  - Wind Direction - sin, cos

Turbine Location map

- Active and inactive turbines
  - 73352 rows, 28 columns
- Reduce dataset to active turbines at Chicago & Bay Area
- Impute missing values from other turbines in the same wind farm

Turbine Energy Production

- Combine information from two sources
- Some inconsistency between two sources
- Missing turbine information:
  - Impute from similar turbines
  - Manufacture information

EDA Chicago Data Colab Notebook
EDA Bay Area Data Colab Notebook

Turbine Mapping Colab Notebook

Turbine Energy Production Notebook

Project Workflow

ERA5

Historical Wind data (Data)

Prediction: Wind Speed

ERA5

PYRECAST™

Verify the quality of the wind predictions

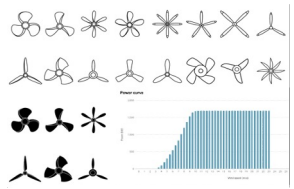
Wind Predictions at Turbine locations

Energy generated by Wind Turbines at certain location/Region

Applications

E.g.Optimize energy allocation/usage; Suggest locations to build new Wind Turbines

Wind Turbine locations (Data)



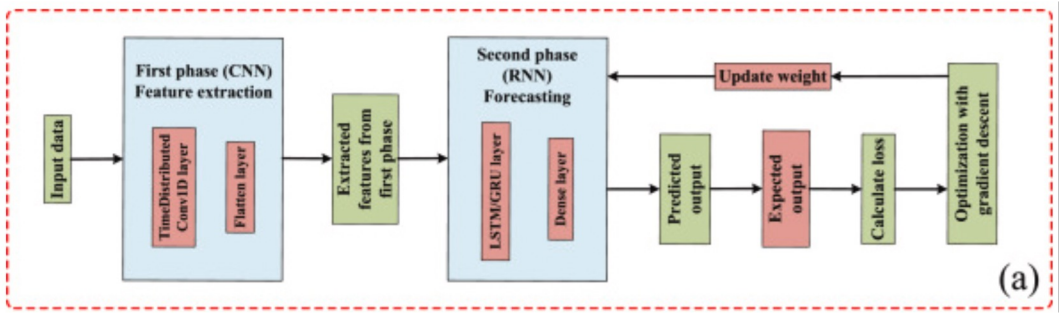
Power curve of different Wind Turbine models (Data)

CNN - LSTM model

Wind Speed
Wind Direction
(past 6 hours)

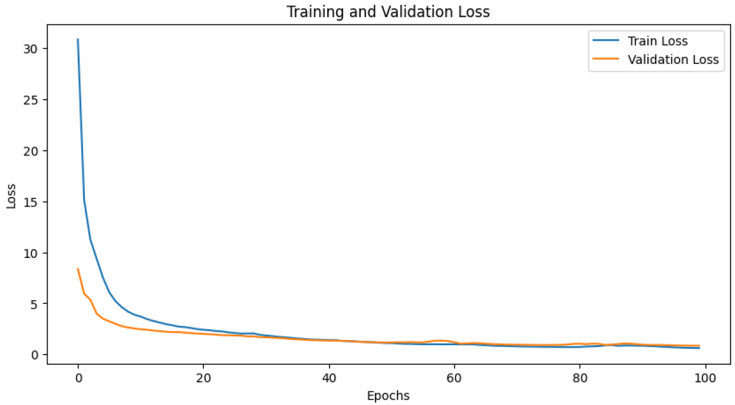
→ CNN - LSTM →

Wind Speed
(next 1 hour)



Duan et al.

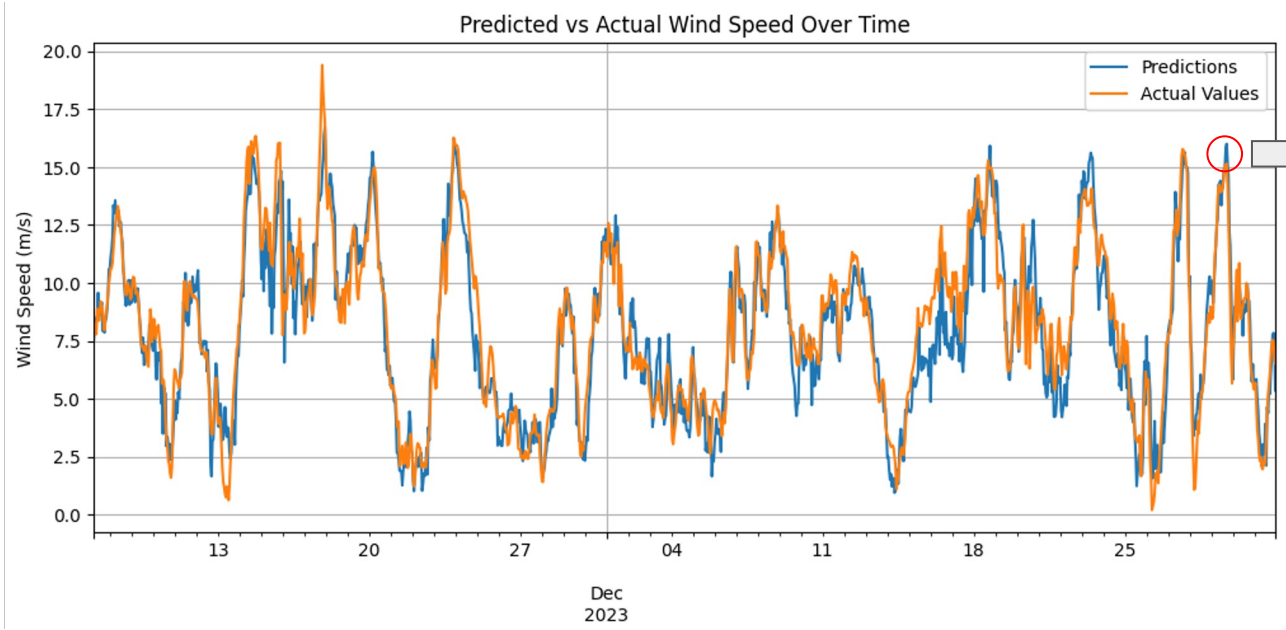
Hyper-parameters:
- Activation function: ReLU
- MSE Loss
- Adam Optimizer
- Learning rate 0.001

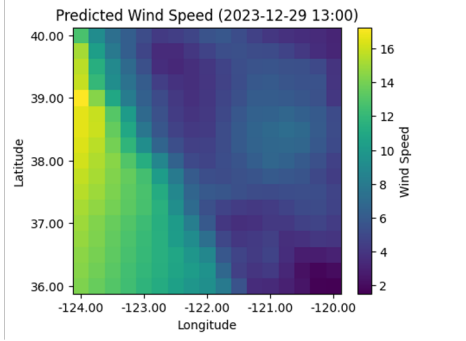


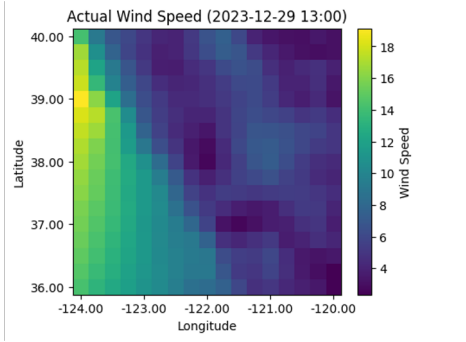

```
ConvLSTM_Model(
  (conv1): Conv2d(3, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (conv2): Conv2d(16, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (lstm): LSTM(2048, 50, batch_first=True)
  (linear): Linear(in_features=50, out_features=289, bias=True)
)
```

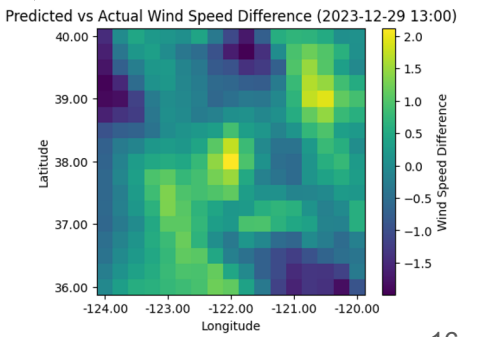
Test set MSE: 1.13

Prediction vs Actual Wind Speed









Latitude 40, Longitude -124

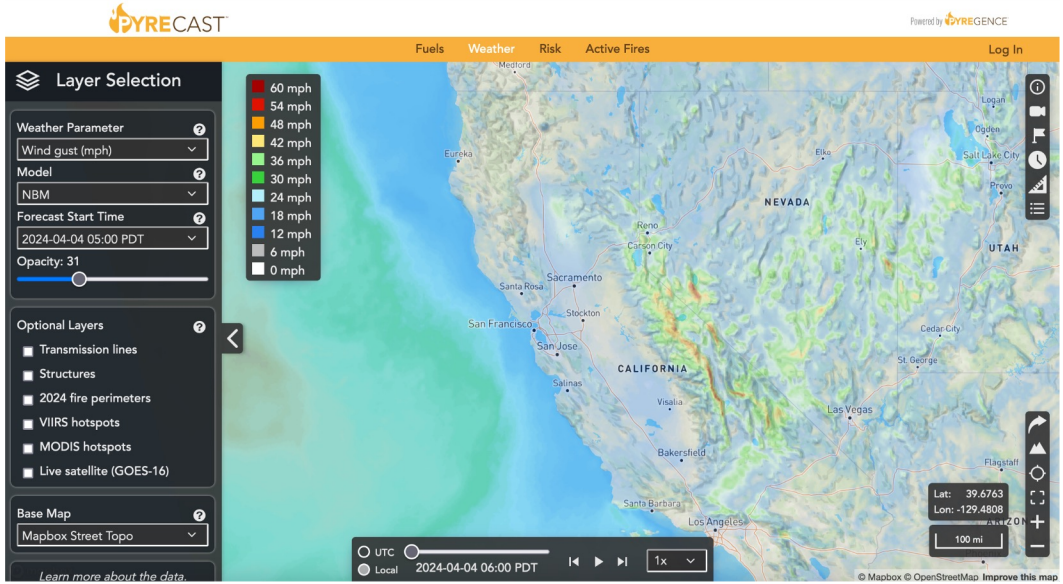
MSE: 0.62

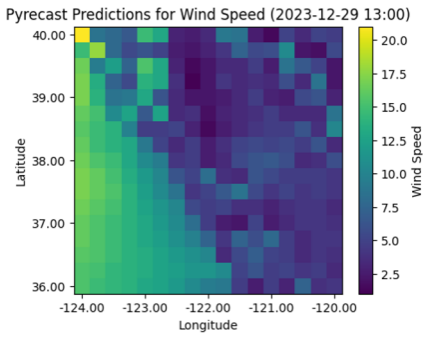
Pyrecast (NBM) Prediction vs Actual Wind Speed

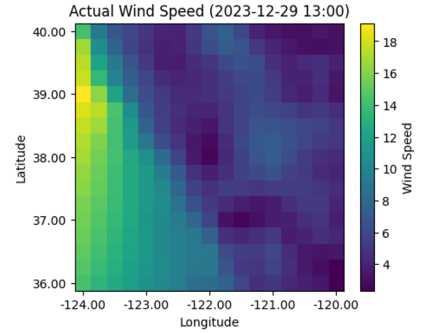
NBM - National Blend of Models, developed by the National Weather Service (NWS)

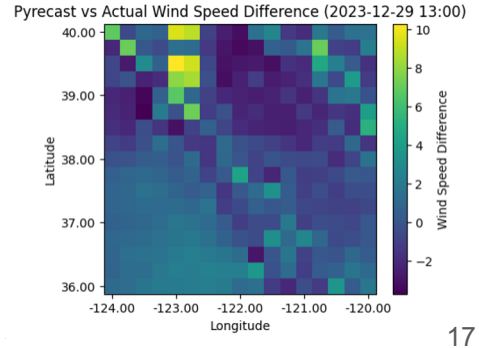
Ensemble forecasting: GFS, ECMWF, NAM, HRRR, MOS

Primarily use laws of physics and fluid dynamics to simulate atmospheric phenomena









MSE: 5.27

Project Workflow

ERA5

PYRECAST™

Verify the quality of the wind predictions

ERA5

Historical Wind data (Data)

Prediction: Wind Speed

Wind Predictions at Turbine locations

Energy generated by Wind Turbines at certain location/Region

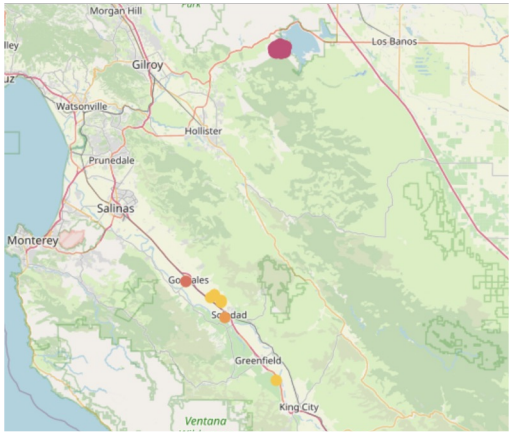
Applications

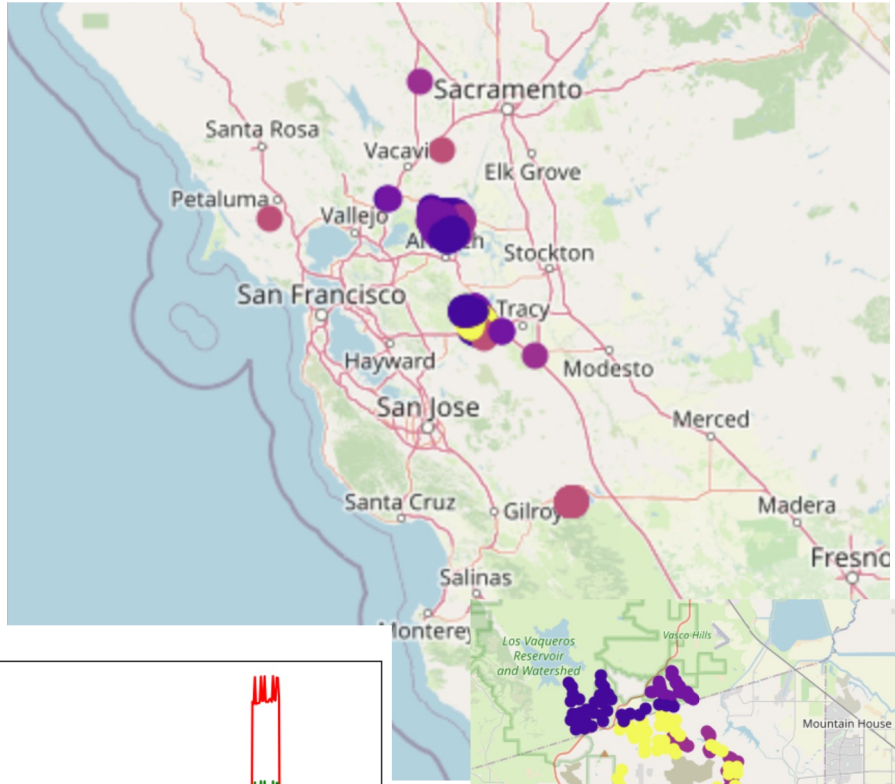
E.g.Optimize energy allocation/usage; Suggest locations to build new Wind Turbines

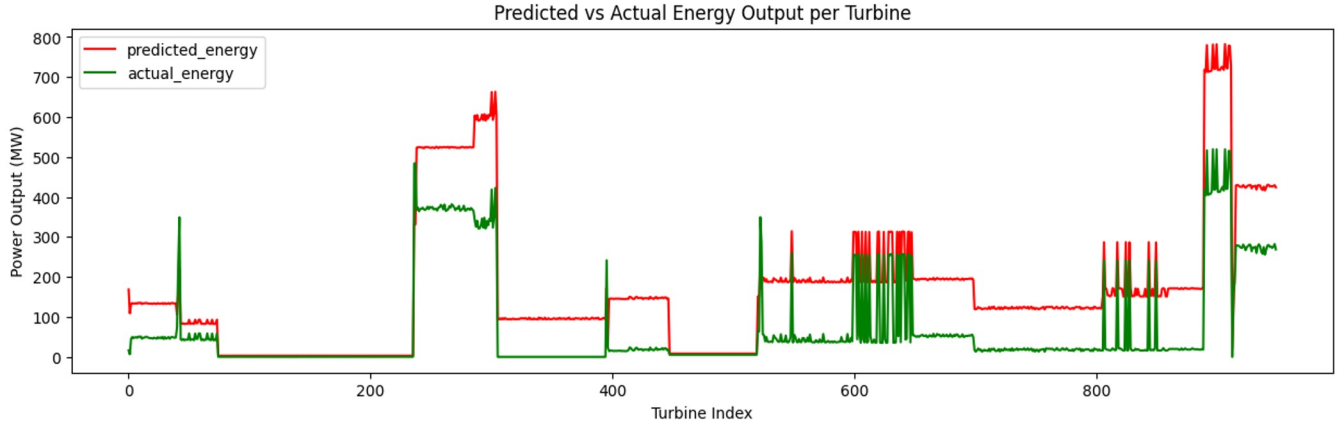
Wind Turbine locations (Data)

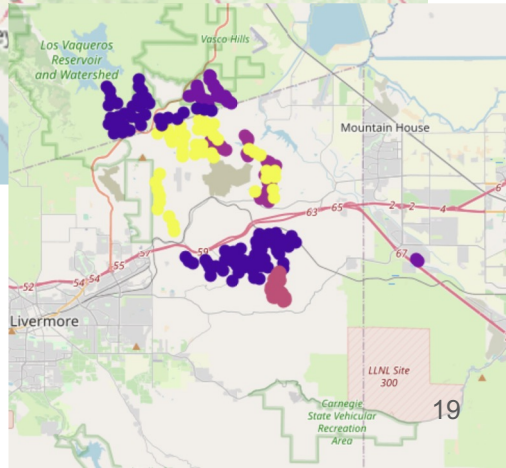
Power curve of different Wind Turbine models (Data)

Wind Energy Prediction

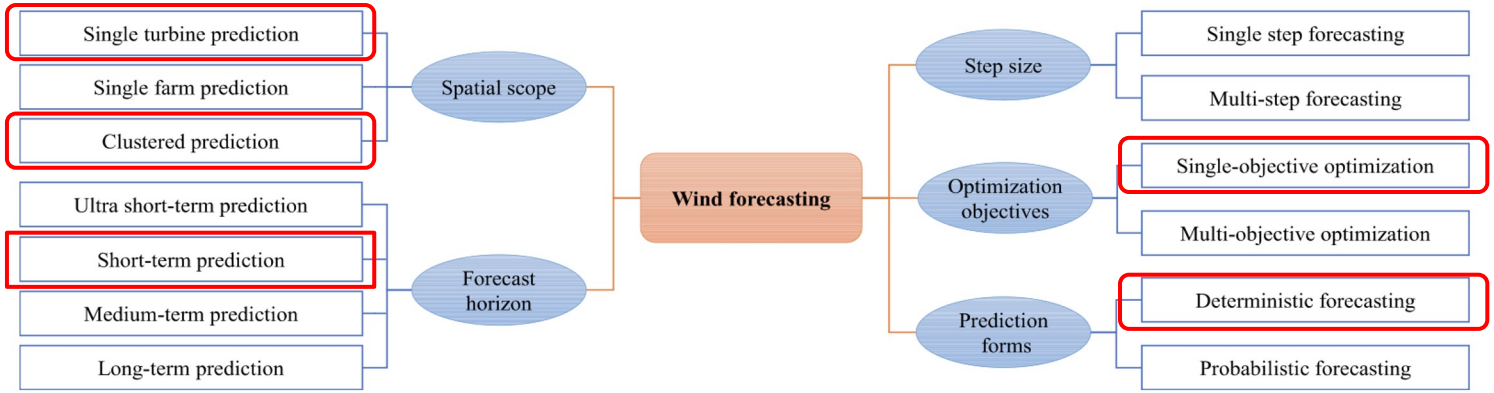








Summary



Data Processing | **Machine Learning** | **Visualization**

Xie Y, Li C, Li M, Liu F, Taukenova M. An overview of deterministic and probabilistic forecasting methods of wind energy. iScience. 2022 Dec 14;26(1):105804. doi: 10.1016/j.isci.2022.105804. PMID: 36624842; PMCID: PMC9823194.

Future Work

- Use global data or US data
- Different ML models
- More granular training data with more features
- Test accuracy of model in different regions
- Perform prediction comparison against models that use ERA5 data
- Explore applications where our predictions can be used

Image generated by DALL-E

Acknowledgements

Professor David Guy Brizan

Professor David Saah

Professor Fernanda Lopez Ornelas

Aditya Kunatharaju
Parisa Abbasi

Shane Romsos

PyTorch

UNIVERSITY OF SAN FRANCISCO
CHANGE THE WORLD FROM HERE

jupyter

ERA5

PYRECAST

Google Earth Engine

wind-turbine-models

THE WINDPOWER
Wind Energy Market Intelligence

USGS science for a changing world
BERKELEY LAB
AMERICAN CLEAN POWER
The U.S. Wind Turbine Database

ArcGIS Online

UNIVERSITY OF SAN FRANCISCO
CHANGE THE WORLD FROM HERE

Image generated by DALL-E