Predicting Climate Change with Carbon Emissions

By: Joshua Ko

Acceleration of climate change

- Human activity
  - Deforestation
  - Fossil fuels
- An abundant amount of greenhouse gases
  - Carbon dioxide
  - Methane
  - CFCs
  - Water
  - Sulfur hexafluoride

Tackle climate change with machine learning

- Observe how carbon dioxide changes the climate
  - Temperature
  - Precipitation
- Determine which areas are more susceptible
- Limit carbon emissions in those areas

About the data

Carbon emission and climate data are provided by NOAA

- Carbon dioxide data
  - Collected by flasks several meters above surface
  - Measured in ppm
- Climate data
  - Average monthly temperature
  - Total monthly precipitation
  - Other variables are sparse

Correlation between target and CO_2

Temperature

Site	Correlation
Boulder, CO	-0.1832
Hilo International, HI	0.3147
Hohenpeißenberg, GM	-0.1658
Cape Florida, FL	0.4674
Los Angeles, CA	0.129
San Francisco, CA	0.2284
Vestmannaeyjar, IC	0.4409
Temple, TX	0.1077

Precipitation

Site	Correlation
Boulder, CO	0.3693
Hilo International, HI	-0.027
Hohenpeißenberg, GM	-0.143
Cape Florida, FL	0.0535
Los Angeles, CA	-0.052
San Francisco, CA	-0.1724
Vestmannaeyjar, IC	0.4842
Temple, TX	0.0993

Linear Regression

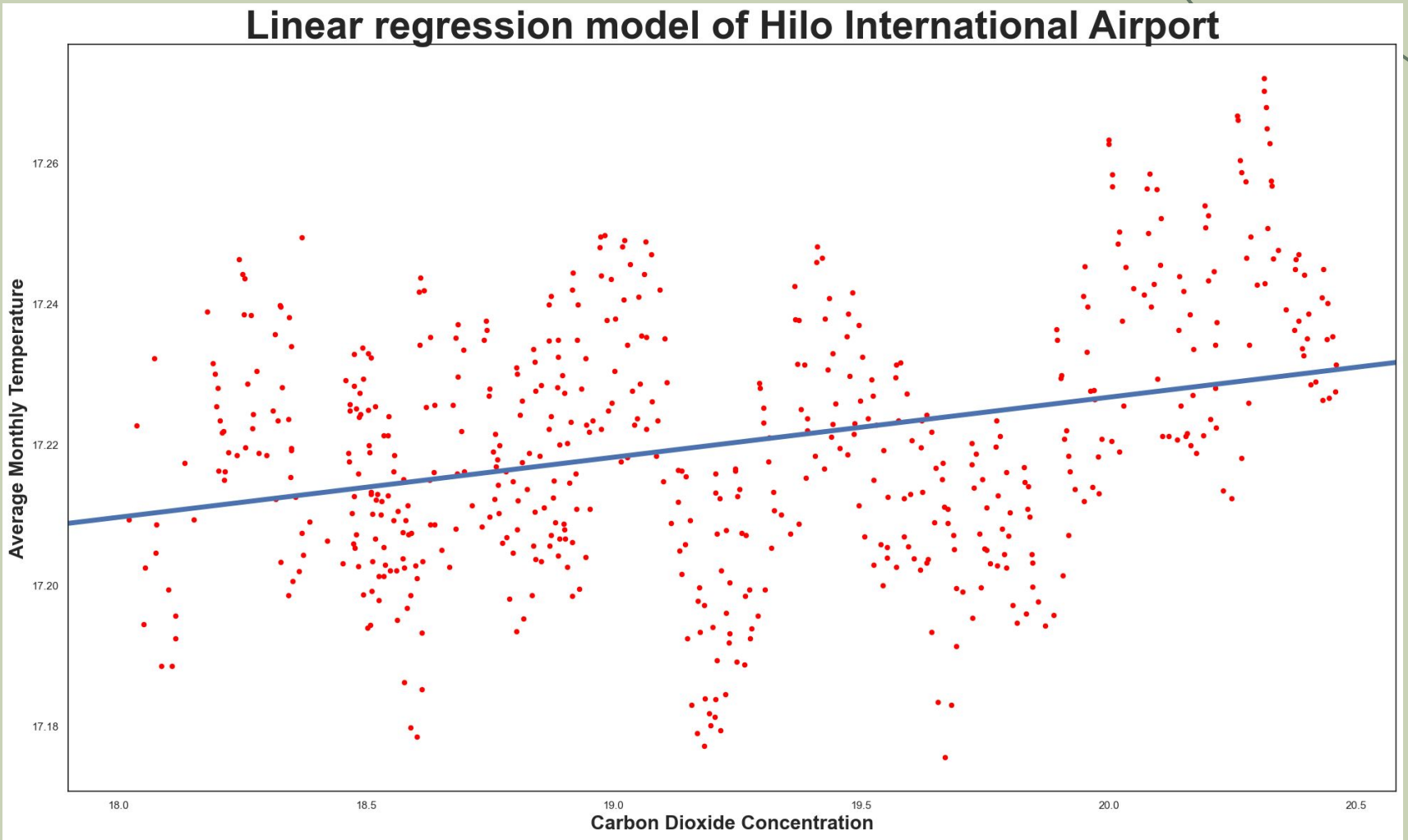
Temperature

Site	Adj. R-squared	Constant (P-value)	Slope (P-value)	Mean Squared Error
Boulder, CO	0.034	19.3995 (0.000)	-0.1278 (0.058)	0.0884
Hilo International, HI	0.099	17.0561 (0.000)	0.0085 (0.000)	0.0168
Hohenpeissenberg, GM	0.023	17.5397 (0.000)	-0.0380 (0.019)	0.0642
Cape Florida, FL	0.218	16.3450 (0.000)	0.0468 (0.000)	0.0316
Los Angeles, CA	0.010	16.6399 (0.000)	0.0226 (0.113)	0.0364
San Francisco, CA	0.047	16.4865 (0.000)	0.0232 (0.002)	0.0249
Vestmannaeyjar, IC	0.191	15.4149 (0.000)	0.0660 (0.000)	0.0359
Temple, TX	0.004	16.2759 (0.000)	0.0404 (0.232)	0.0742

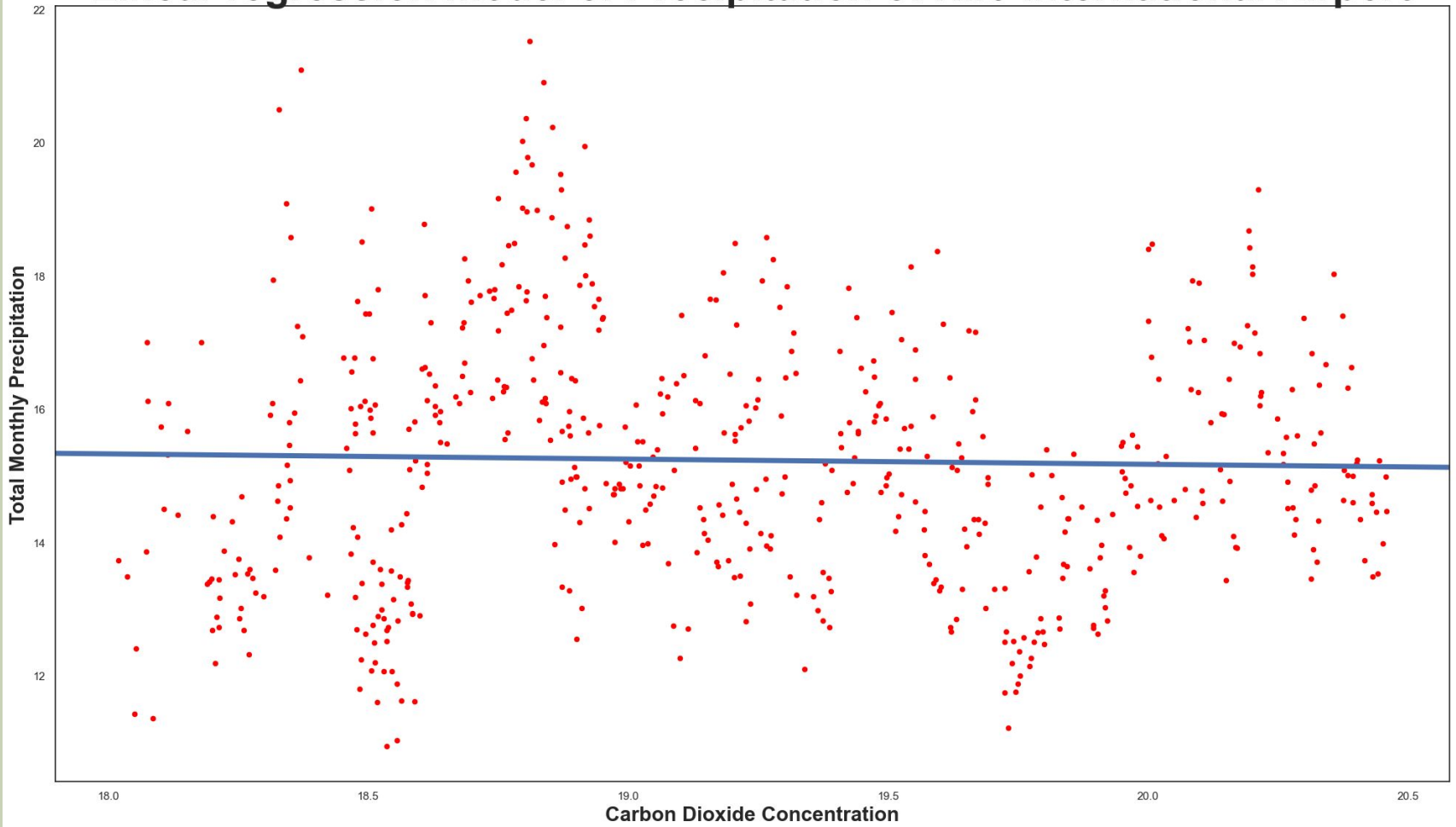
Precipitation

Site	Adj. R-squared	Constant (P-value)	Slope (P-value)	Mean Squared Error
Boulder, CO	0.128	-37.4274 (0.001)	2.1898 (0.000)	0.7101
Hilo International, HI	-0.001	16.7257 (0.000)	-0.0782 (0.000)	1.8941
Hohenpeissenberg, GM	0.016	18.5583 (0.000)	-0.4554 (0.019)	0.8980
Cape Florida, FL	-0.001	5.6583 (0.220)	0.2008 (0.386)	1.3351
Los Angeles, CA	-0.004	9.0495 (0.327)	-0.2912 (0.524)	1.173
San Francisco, CA	0.024	25.2646 (0.004)	-1.0209 (0.020)	1.4654
Vestmannaeyjar, IC	0.231	-15.8312 (0.000)	1.4020 (0.000)	0.6780
Temple, TX	0.002	-4.0718 (0.676)	0.5325 (0.270)	1.0595

** The low r-squared score indicates that the model does a poor job explaining the variance of the data


Linear regression model of Hilo International Airport
Average Monthly Temperature
17.26
17.24
17.22
17.20
17.18
18.0
18.5
19.0
19.5
20.0
20.5
Carbon Dioxide Concentration

Linear regression model of Precipitation of Hilo International Airport



XGBoost Algorithm

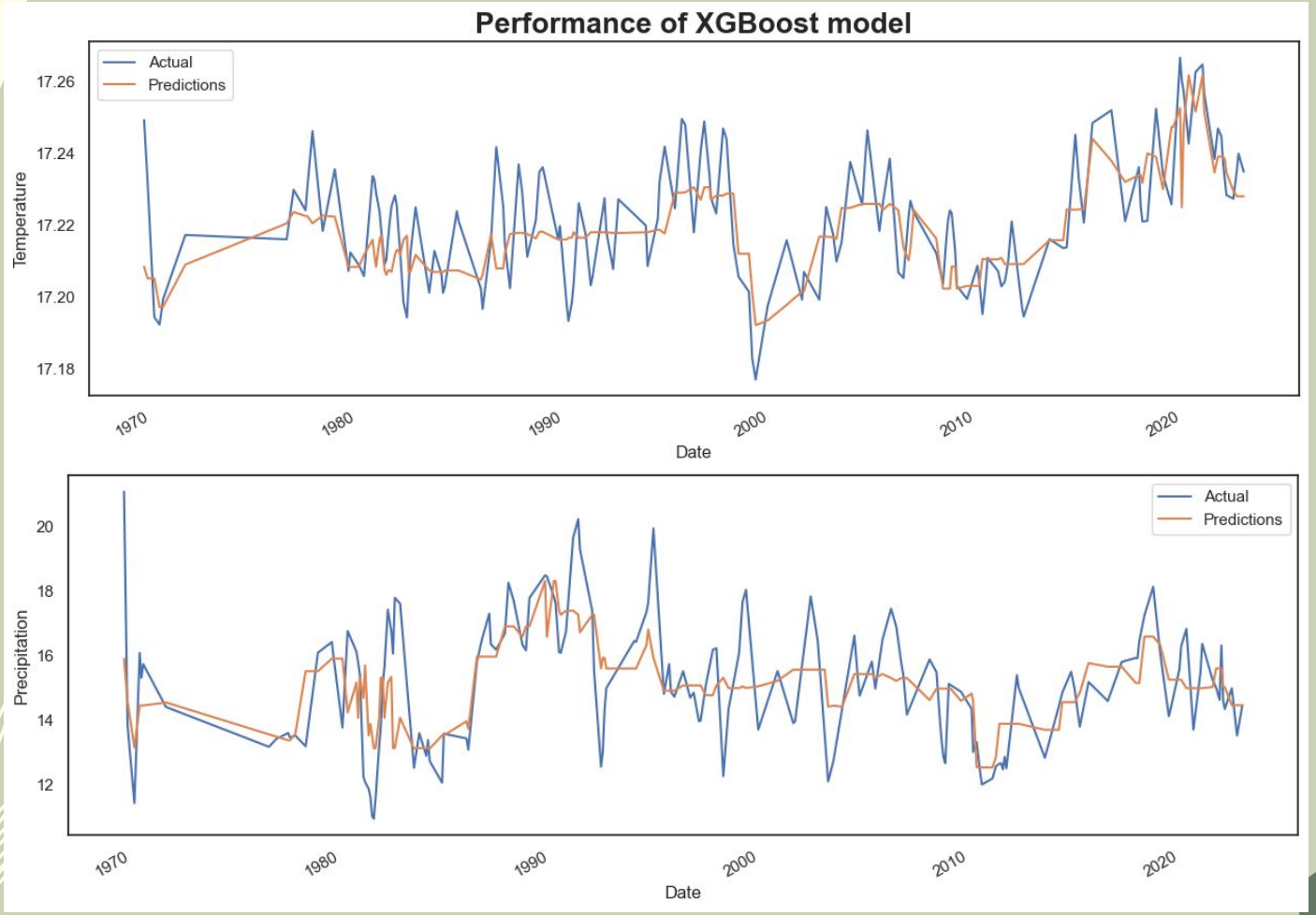
Temperature

Site	Alpha	Learning Rate	Max Depth	Number of estimators	Score	MSE of test set
Boulder, CO	1	0.05	5	500	-0.2777	0.0939
Hilo International, HI	0	0.01	5	1000	0.4683	0.0124
Hohenpeissenberg, GM	0	0.01	5	1000	0.07849	0.0624
Cape Florida, FL	1	0.1	5	500	0.13622	0.0297
Los Angeles, CA	0	0.01	5	1000	0.10141	0.0424
San Francisco, CA	0	0.1	5	500	0.02760	0.0254
Vestmannaeyjar, IC	1	0.5	5	500	0.04004	0.0400
Temple, TX	0	0.01	6	1000	0.08986	0.0894

Precipitation

Site	Alpha	Learning Rate	Max Depth	Number of estimators	Score	MSE of test set
Boulder, CO	2	0.01	6	1000	-0.1690	0.6511
Hilo International, HI	1	0.1	5	500	0.3088	1.4465
Hohenpeissenberg, GM	2	0.05	7	500	0.0710	0.8750
Cape Florida, FL	2	0.005	5	1000	0.02428	1.2637
Los Angeles, CA	2	0.05	5	500	-0.1548	1.1681
San Francisco, CA	2	0.05	6	500	-0.1926	1.3976
Vestmannaeyjar, IC	2	0.05	5	500	0.2856	0.6798
Temple, TX	2	0.1	5	500	0.2027	0.8553

** The r-squared (score column) relatively higher than the linear model


Performance of XGBoost model
Actual
Predictions
Temperature
17.26
17.24
17.22
17.20
17.18
1970
1980
1990
2000
2010
2020
Date
Actual
Predictions
Precipitation
20
18
16
14
12
1970
1980
1990
2000
2010
2020
Date

Linear model vs. XGBoost

Temperature

Site	Linear MSE	XGBoost MSE
Boulder, CO	0.0884	0.0939
Hilo International, HI	0.0168	0.0124
Hohenpeißenberg, GM	0.0642	0.0624
Cape Florida, FL	0.0316	0.0297
Los Angeles, CA	0.0364	0.0424
San Francisco, CA	0.0249	0.0254
Vestmannaeyjar, IC	0.0359	0.04
Temple, TX	0.0742	0.0894

Precipitation

Site	Linear MSE	XGBoost MSE
Boulder, CO	0.7101	0.6511
Hilo International, HI	1.8941	1.4465
Hohenpeißenberg, GM	0.898	0.875
Cape Florida, FL	1.3351	1.2637
Los Angeles, CA	1.173	1.1681
San Francisco, CA	1.4654	1.3976
Vestmannaeyjar, IC	0.678	0.6798
Temple, TX	1.0595	0.8553

Summary

- According to the correlation table
  - Temperature changes affects Cape Florida, Iceland, Hawaii
  - Precipitation changes affects Iceland and Boulder, CO
- Using the XGBoost is recommended
  - Slight overfitting
  - Underfitting when predicting temperature of Cape Florida
  - Prepare for changes in the environment
  - Acknowledge the dangers of and limit production of CO_2
- Both models performed worse when predicting precipitation
  - Other factors should be considered
  - Effect of transpiration and tradewinds

Next Steps

- Experiment with other greenhouse gases
  - XGBoost has a unique property of feature importances
- Include more aspects of climate
  - Wind speed
  - Cloud coverage
  - Relative Humidity
  - Tradewinds
- Include more alternative parameters for XGBoost algorithm

Thank you

Email: joshuawko@gmail.com
Github: Jko0425