How Emissions Will Impact Wildfire Risk

Website

Nathan Brodie Benjamin Xue Kai Morita Duncan Watson-Parris

njbrodie@ucsd.edu bxue@ucsd.edu kmorita@ucsd.edu dwatsonparris@ucsd.edu

DSC 180B B03-2, The University of California, San Diego

UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

Why Wildfires Are Important

Wildfires are occurrences of uncontrolled fires that spread rapidly across vegetation. These fires have many effects on the ecosystem, including:

- Climate Change Wildfires emit carbon dioxide into the atmosphere, which leads to an increase in temperatures globally.
- Destruction of Vegetation Many plants and soil are destroyed in the process.
- Health effects Smoke from wildfires can cause respiratory issues in humans and animals.

Vapor Pressure Deficit (VPD) represents the difference between the water vapor present in the atmosphere and the maximum amount of water vapor the atmosphere can hold. This can be used to represent how dry the plants within a surface area, which is an indicator of high wildfire chance. We plan to use deep learning models to analyze how emissions can affect VPD.

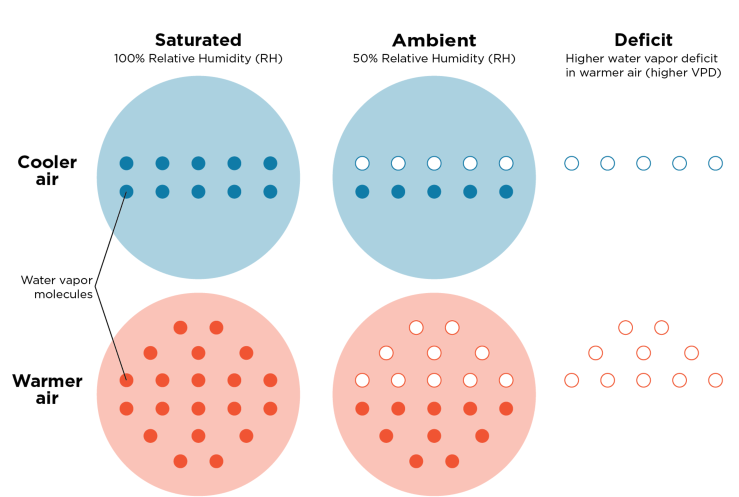
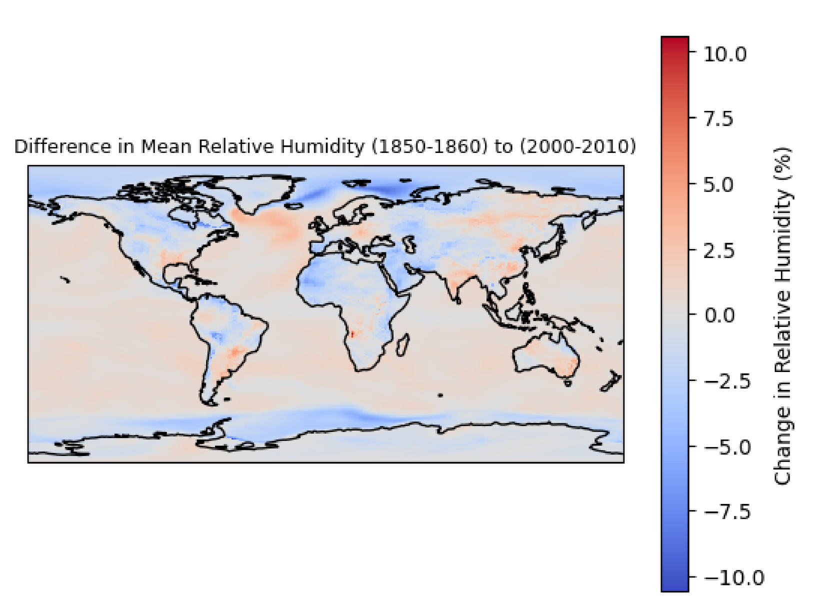


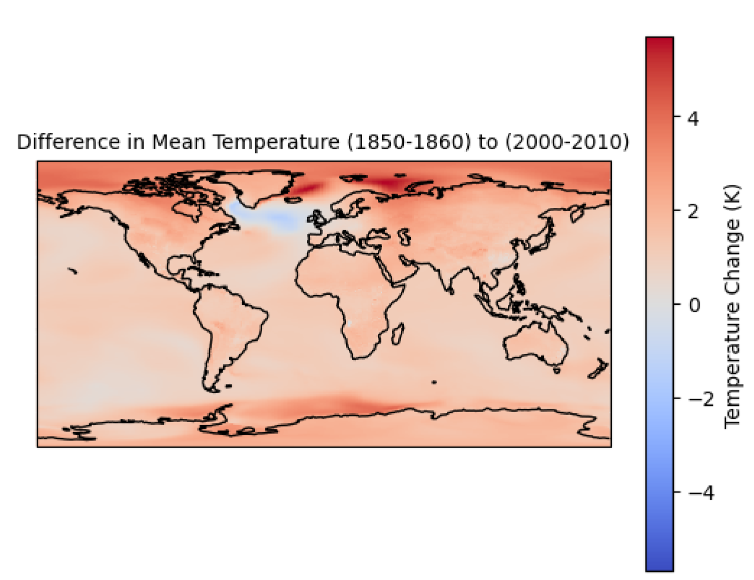
Figure 1. VPD Visualization

Data Background

- Coupled Model Intercomparison Project Phase 6 (CMIP6): A collection of many of the most advanced climate models.
- Community Earth System Model 2 (CESM2): Specific climate simulation model used.

To calculate VPD, we utilized CESM2's data regarding near-surface relative humidity and near-surface average temperature. To confirm that the data was sufficiently complete, we performed a timeseries analysis of these two variables, as shown below.





There are two steps to calculate the VPD using relative humidity and average temperature:

1. Calculate the saturation vapor pressure (SVP) using the Clausius-Clapeyron Equation.

$$SVP = 0.6112 \exp\left(\frac{17.76 \times T}{T + 243.5}\right)$$

2. Calculate the VPD as the difference between relative humidity (RH) and SVP

The Models We Used

Our models used the emissions of CO_2, CH_4, SO_2 and BC (black carbon) from several different climate scenarios to make predictions of global VPD up to the year 2100. We followed the approach of the ClimateBench paper [4] to test our predictions on the climate change scenario ssp245 (moderate climate change). Like [4], we used the following models:

1. **Linear Model**: We fit a linear model to predict VPD from the global mean temperature. This is our baseline model to compare the performance of our machine learning models.
2. **Gaussian Process**: We performed dimensionality reduction on the aerosol emissions, then fit a GP model with a Matern-1.5 kernel onto the data.
3. **Random Forest**: We used the same dimensionality reduced data as the Gaussian Process to fit a random forest.
4. **Convolutional Neural Network**: We fit a CNN-LSTM trained in 10 year chunks using ReLU activation functions.

Our Models' Predictions

Figure 2. Prediction Comparisons

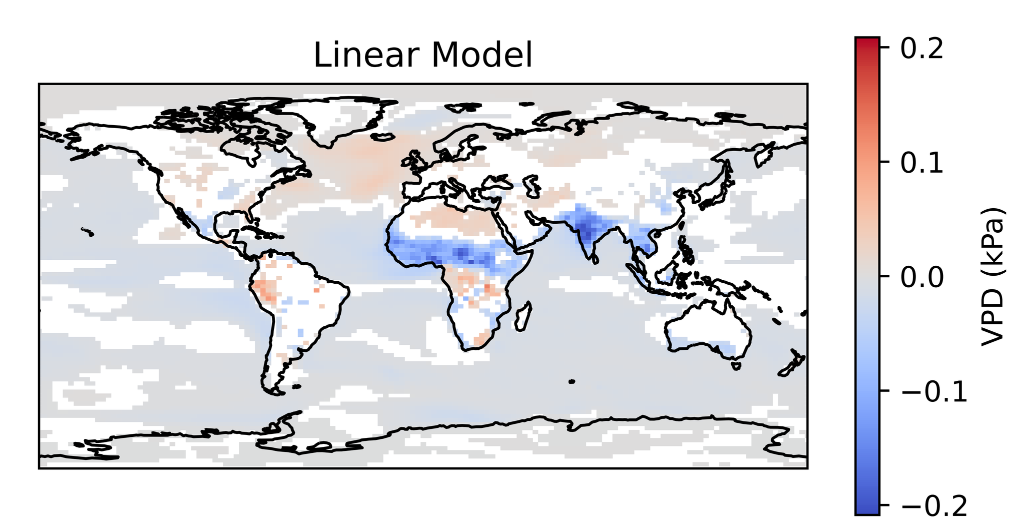


Figure 3. Linear Model Differences

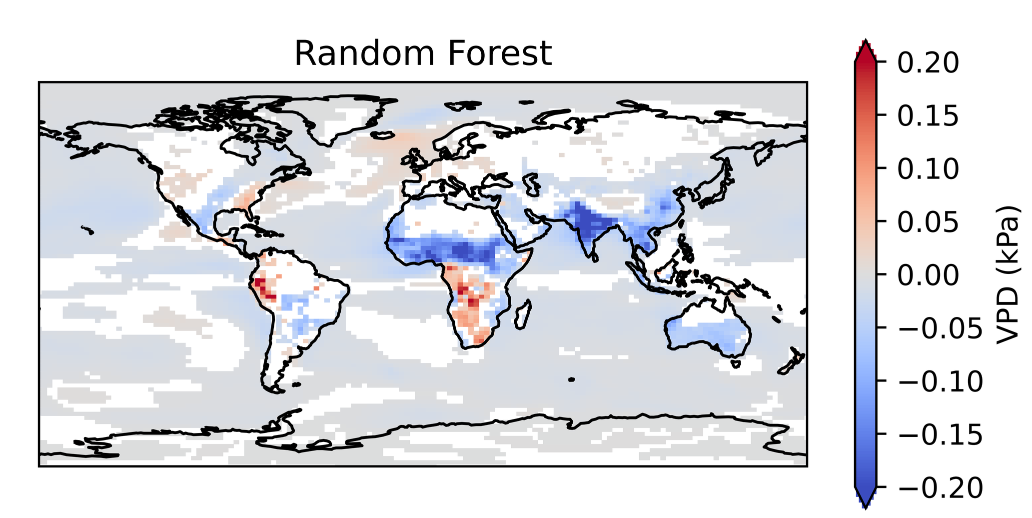


Figure 4. Random Forest Differences

Results and Discussion

To compare our emulator predictions, we calculate the root-mean square error (RMSE), normalize it spatially (NRMSE_s) and globally (NRMSE_g), then compare using a weighted sum (NRMSE). These are defined in [4] as follows:

$$\text{NRMSE}_s = \sqrt{\langle (|x_{i,j,t}|_t - |y_{i,j,t,n}|_{n,t})^2 \rangle} / |\langle y_{i,j} \rangle|_{t,n}$$

$$\text{NRMSE}_g = \sqrt{|(\langle x_{i,j,t} \rangle - \langle |y_{i,j,t,n}|_n \rangle)^2|_t} / |\langle y_{i,j} \rangle|_{t,n}$$

$$\text{NRMSE} = \text{NRMSE}_s + \alpha \cdot \text{NRMSE}_g,$$

where $\langle x_{i,j} \rangle = \frac{1}{N_{lat} N_{lon}} \sum_i^{N_{lat}} \sum_j^{N_{lon}} \cos(\text{lat}(i)) x_{i,j}$ is the global mean, and $\alpha = 5$. In the following table, we compare the NRMSE results of each of our models. Global NRMSE represents the global averages across the Earth, while spatial is concerned with differences in individual regions.

Model	Spatial	Global	Total
Linear	**0.036**	0.012	0.096
CNN	0.063	0.017	0.148
GP	0.047	**0.009**	**0.094**
RF	0.051	0.019	0.144

Table 1. NRMSE results of different climate models used.

The linear model performed the best when taking into account the regional differences spatially. Overall, the gaussian model performed the best with the lowest global average NRMSE. A linear model is generally good for making predictions on future data, however, it may not be able to capture climate trends present in VPD data. The performance of the models could be an indicator that VPD data follows a pretty consistent pattern.

Future Direction

Additional data to analyze

In order to improve upon our models we could have utilized other variables within the dataset to increase accuracy. For example, we could look at evapotranspiration, wind direction, and other variables that might not have as large of an impact as VPD but could still contribute.

Another approach we can take to improve the real-world implication of our model could be to find where trees and other possible flammable plants are. Combining our VPD data with this data will allow us to predict where wildfires will occur with greater accuracy.

Improvements to Model

Here we will talk about improvements that could have been made to the model, including different ways of hyper-parameter tuning and different Deep Learning Methods.

References

[1] Clausius-clapeyron equation.
American Meteorological Society: Glossary of Meteorology, Jan 2024.

[2] Barnard et al. Di Giuseppe, El Garroussi.
Europe faces up to tenfold increase in extreme fires in a warming climate.
npj Clim Atmos Sci, 7(30), 2024.

[3] J. T. Randerson F. Sedano.
Multi-scale influence of vapor pressure deficit on fire ignition and spread in boreal forest ecosystems.
Biogeosciences, 2014.

[4] Rao Y. Olivié D. Seland Ø. Nowack P. Camps-Valls G. et al. Watson-Parris, D.
Climatebench v1.0: A benchmark for data-driven climate projections.
Journal of Advances in Modeling Earth Systems, 14, 2022.