Conditional Generative Adversarial Networks for Wildfire Spread Prediction

Daniel Lopez

VANDERBILT UNIVERSITY

Wildfires are a devastating force

01

775,000 RESIDENCES FLAGGED AS “EXTREME” RISK

02

WILDFIRE AEROSOLS ESTIMATED 300,000 PREMATURE DEATHS PER YEAR

03

AN AVERAGE OF 1.2 MILLION ACRES OF US WOODLAND BURN EVERY YEAR

04

A LARGE WILDFIRE — OR CONFLAGRATION — IS CAPABLE OF MODIFYING THE LOCAL WEATHER CONDITIONS

The Future of Wildfires

According to the **National Oceanic and Atmospheric Administration**: “Climate change, including increased heat, extended drought, and a thirsty atmosphere, has been a key driver in increasing the risk and extent of wildfires in the western United States during the last two decades. Wildfires require the alignment of a number of factors, including temperature, humidity, and the lack of moisture in fuels, such as trees, shrubs, grasses, and forest debris. All these factors have strong direct or indirect ties to climate variability and climate change.”

Building Blocks

- Next Day Wildfire Spread: A Machine Learning Data Set to Predict Wildfire Spreading from Remote-Sensing Data by Huot et al. 2022
    1. Developed primary dataset for this study
    2. Machine learning application of a convolutional autoencoder for wildfire spread prediction

Next Day Wildfire Spread

A Data Set to Predict Wildfire Spreading from Remote-Sensing Data

Data Card   Code (5)   Discussion (0)   Suggestions (0)

My Contributions

Develop a conditional generative adversarial network (cGAN) for fire spread prediction

- cGANs have shown success in radar remote sensing meteorology
- cGANs can be optimized for both spatial accuracy and realism

Compare a cGAN performance to the original authors autoencoders

- Motivating question: can a cGAN reliable predict wildfire spread?

Generative Adversarial Networks

- Train two networks with opposing objectives
  - Generator: learns to generate samples
  - Discriminator: learns to distinguish between generated and real samples

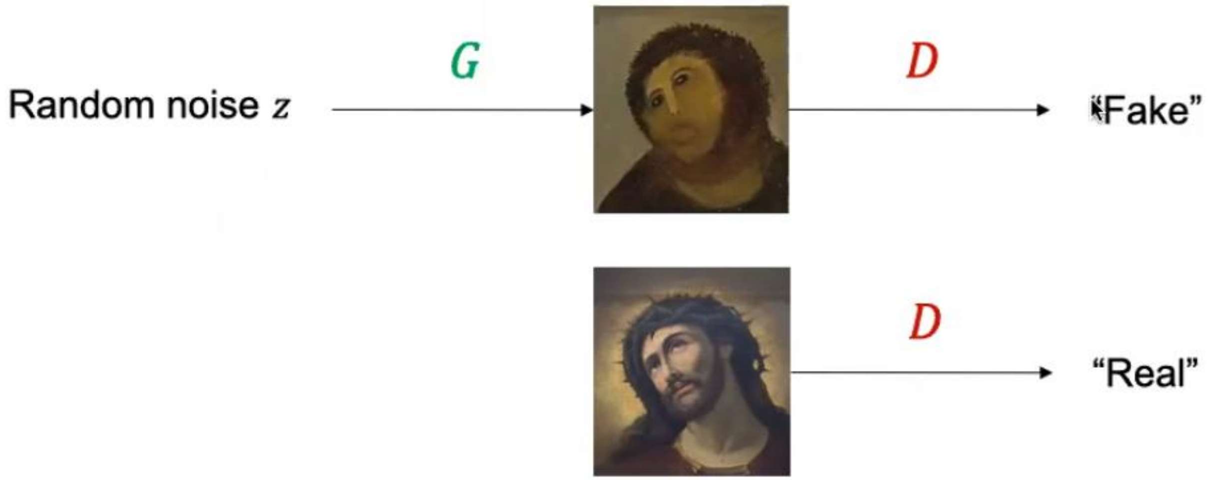


Figure adapted from F. Fleuret

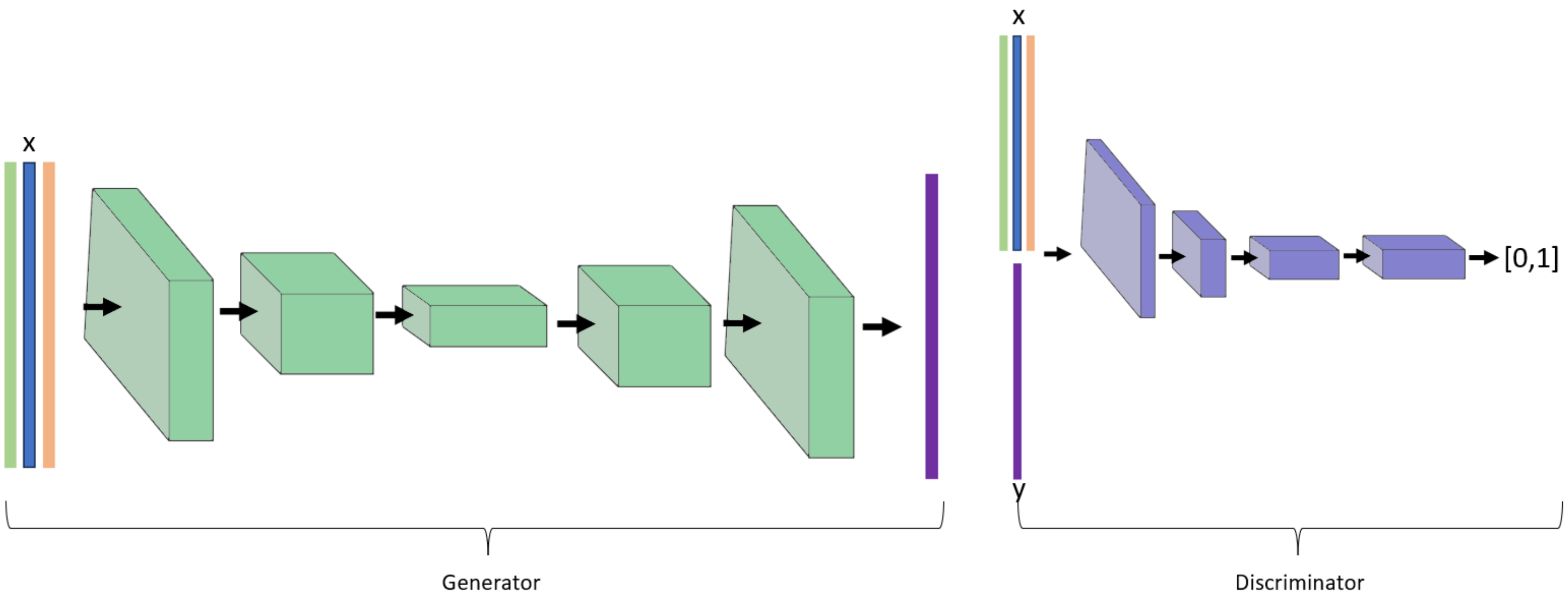
$$G : \mathbb{R}^n \times \mathcal{Y} \rightarrow \mathbb{R}^m,$$

$$D : \mathbb{R}^m \times \mathcal{Y} \rightarrow [0, 1]$$

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, NIPS 2014

GAN Objective

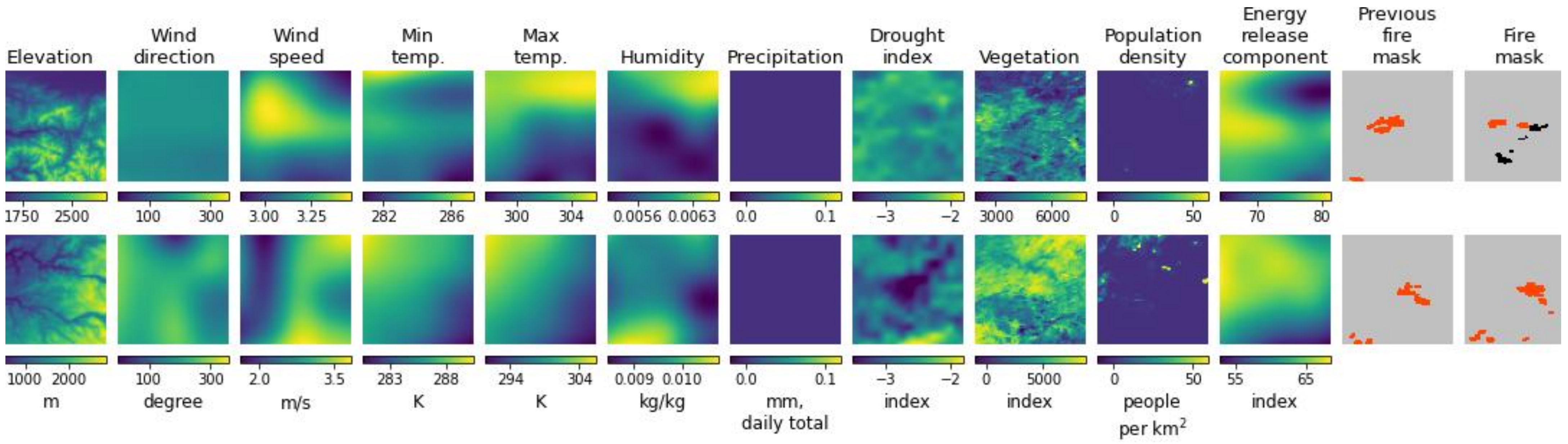
- $\min_G \max_D \mathscr{L}(G, D) = \mathbb{E}_{x,y\sim p_{\text{data}}}[\log D(x|y)] + \mathbb{E}_{z\sim p_z, y\sim p_y}[\log(1 - D(G(z|y)|y))]$
  - Attempt to minimize both the loss of the generator and discriminator
- Here, we implement a variation of pix2pix
- Incorporate "strong reliable" methods for GANs



Final cGAN Objective

- $$\min_G \max_D \mathscr{L}(G, D) = \mathbb{E}_{x,y\sim p_{\text{data}}}\left[\log D(x|y)\right] + \mathbb{E}_{z\sim p_z, y\sim p_y}\left[\log(1 - D(G(y)|y))\right] + \gamma \cdot \mathbb{E}_{z\sim p_z, y\sim p_y}\left[-\alpha y \log(G(y)) - \beta(1-y)\log(1-G(y))\right]$$
  - Binary cross entropy for the discriminator
  - Weighted binary cross entropy for generator segmentation loss
  - Binary cross entropy for generator adversarial loss
  - Gamma, alpha, beta are hyperparameters

Wildfire Spread Dataset

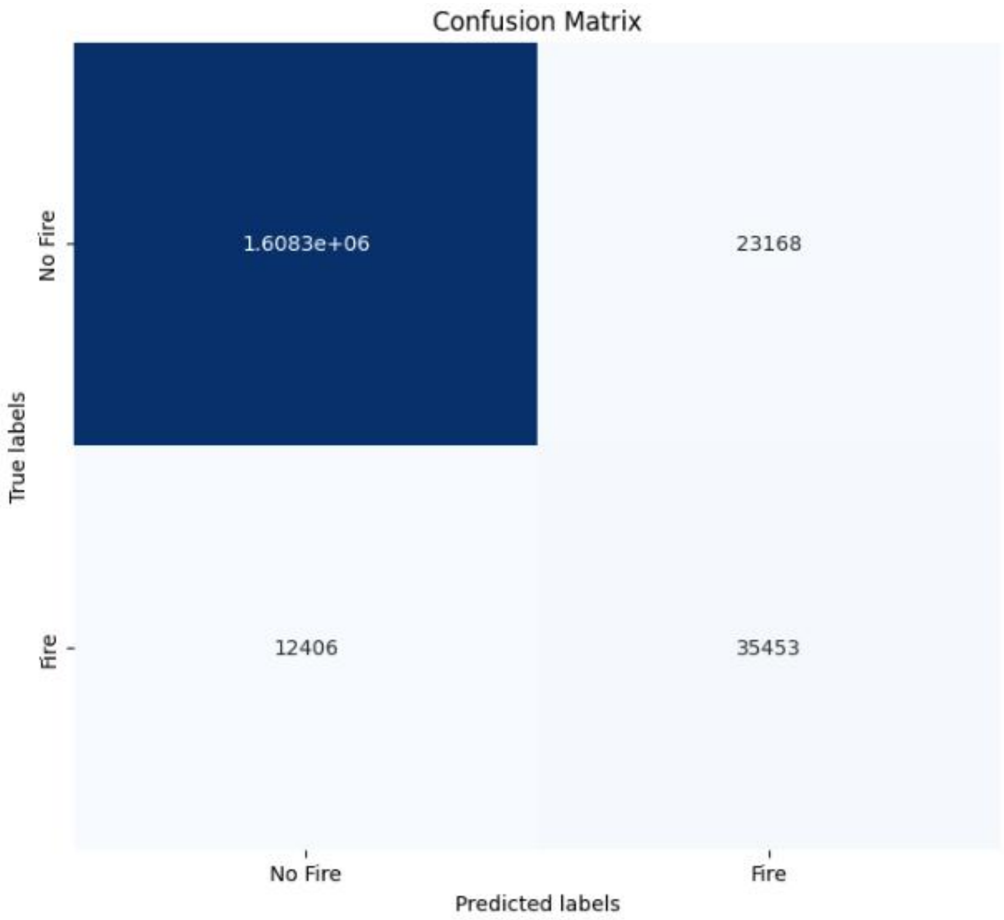


- From 2012 to 2020 contiguous US
- 18445 samples
- Each sample if 64x64km region with 1km resolution

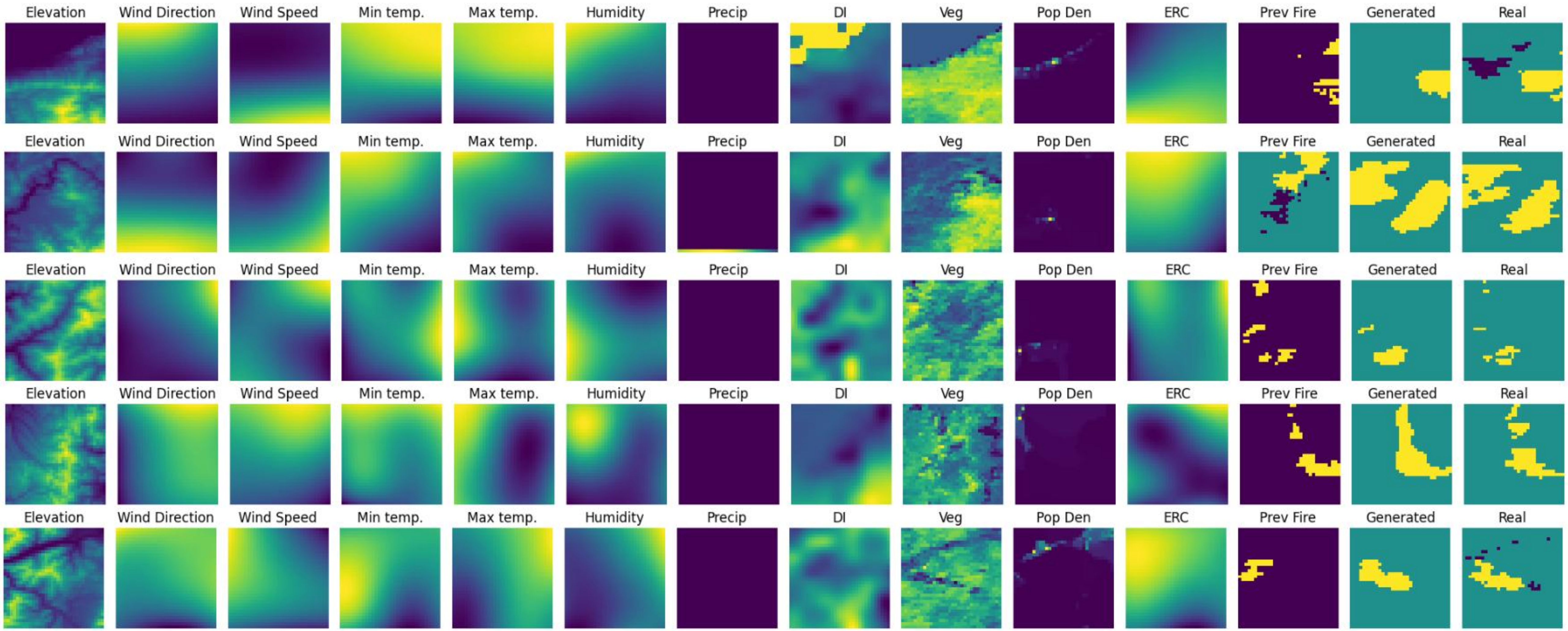
Model Performance

Table 4.1: WILDFIRE SPREADING PREDICTION METRICS

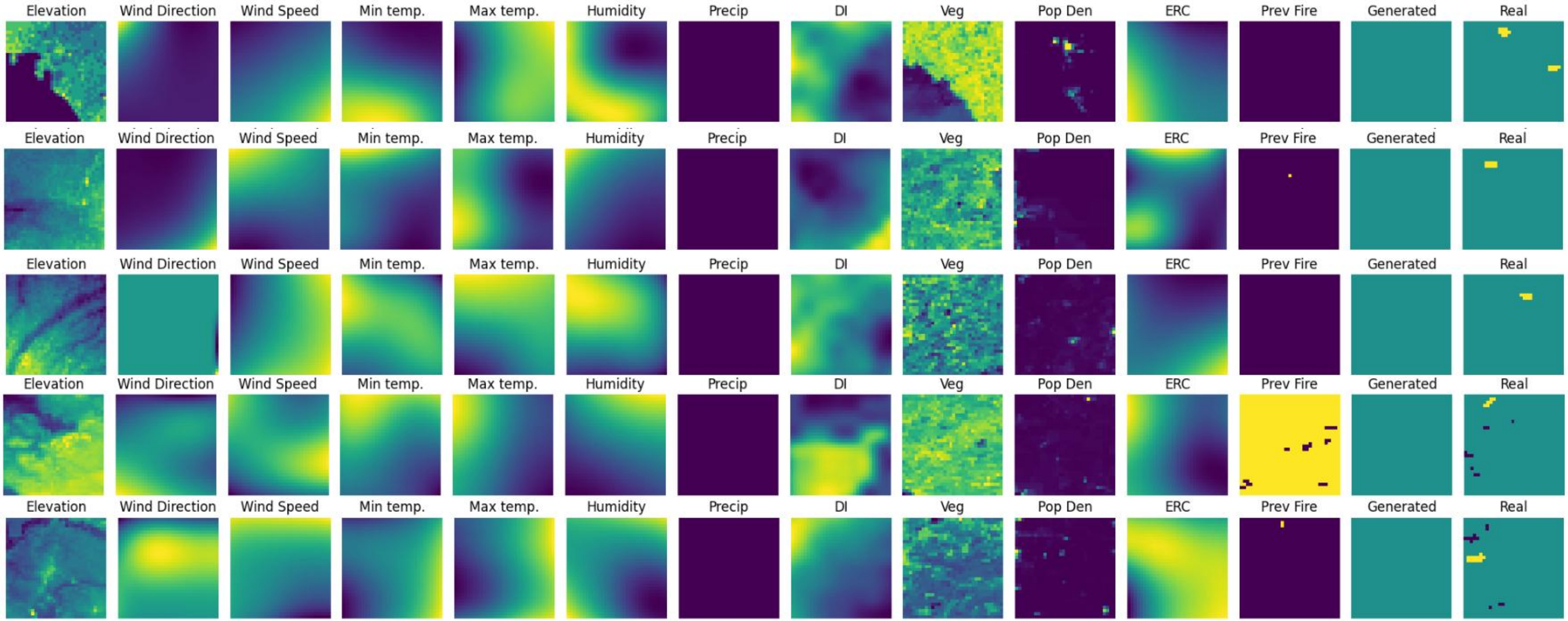
Model	AUC (PR)	Precision	Recall
cGAN	74.0	61.2	73.8
Encoder Decoder	28.4	33.6	43.1
Random Forest	22.5	26.3	46.9
Logistic Regression	19.8	32.5	35.3
Persistence	11.5	35.7	27.3



Better Predictions



Worse Predictions



Ablation Study

- Removing a features each time we train the model
- Determine which variables are the most important for wildfire spread prediction

Table 4.2: Ablation Test Results for Removing One Feature

Removed Feature	AUC PR (%)	Precision (%)	Recall (%)
Previous Fire Mask	43.1	51.6	51.6
Humidity	39.7	50.2	50.2
Max Temperature	37.8	46.1	46.1
Wind Speed	26.0	43.3	43.3
Vegetation	21.1	35.2	35.2
Precipitation	18.9	30.7	30.7
Population	20.1	36.2	36.2
Wind Direction	29.6	42.2	42.2
Min Temperature	12.8	28.8	28.8
Elevation	18.6	42.0	42.0
Drought	22.5	43.5	43.5
ERC	22.4	51.3	51.3

Discussion

- Remarkably strong improvements against an autoencoder
- AI here is not actually necessary
  - Physical models are complicated, slower, not benchmarked in this dataset
- More work on model for determining trustworthiness
  - Discussion on deterministic vs probabilistic predictions

Questions?