Nowcasting Burn Probability for the Contiguous United States

t.r.keeping@pgr.reading.ac.uk

Theodore Keeping[1] | Prof. Sandy Harrison[1] | Prof. Iain C. Prentice[2]

1) University of Reading
2) Imperial College London

University of Reading

LEVERHULME Centre for Wildfires, Environment and Society

LEMONTREE Land Ecosystem Models based On New Theory, obseRvations and ExperimEnts

An Adapted Generalised Linear Modelling (GLM) Method to Predict the Daily Probability of Wildfire Occurrence

GLMs are widely used for modelling wildfire properties. We resolve three key issues with the method (predictor selection, appropriate predictor range, compression) to create a daily fire occurrence model for the contiguous US.

Candidate Predictors	Variable Selection	Domain Optimisation	Power-law Rescaling
50 candidate predictors selected from the literature. Spanning meteorological, vegetation, and human-impact effects.	To identify the key effects (and their predictors) to include. A robust algorithm was needed. The forwards-backwards algorithm was selected, adding new predictors then testing if existing ones should be switched out.	Some predictors are primarily influential in a certain range, for instance, the first few mm of rainfall most reduce fire probability. Thus, predictors were truncated to optimise model performance.	To address known issues of compression in GLM outputs (Nelder and Wederburn, 1972), a power-law transform was applied to the output
Constant	100 runs	100 runs per unique set	Constant

2100 models (of 10,000 potential models)

Figure 1: overview of the model for fire occurrence and its application as an ensemble

1. A forwards-backwards algorithm was used to identify which predictors to use, by adding new predictors then testing if existing ones should be switched out.
2. Some predictors are primarily influential in a certain range, so predictors were truncated to optimise model performance.
3. A power-law transform was applied to the final output to reduce compression.

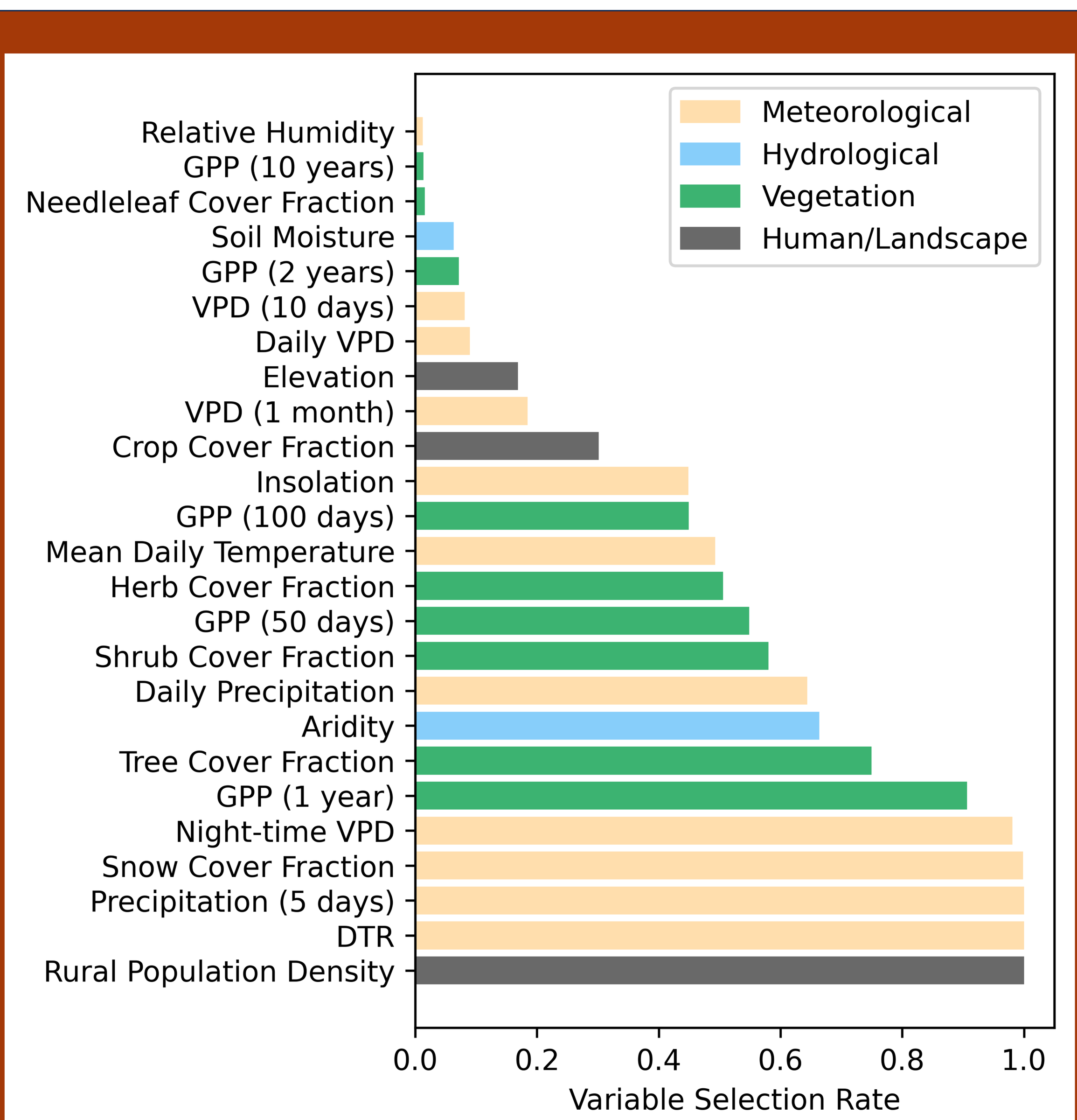

Figure 2: The rate at which predictors were selected for the final model. The top three predictors are virtually always selected. The graph only show predictors selected in more than 1% of 2000 runs.

The Geospatial Average

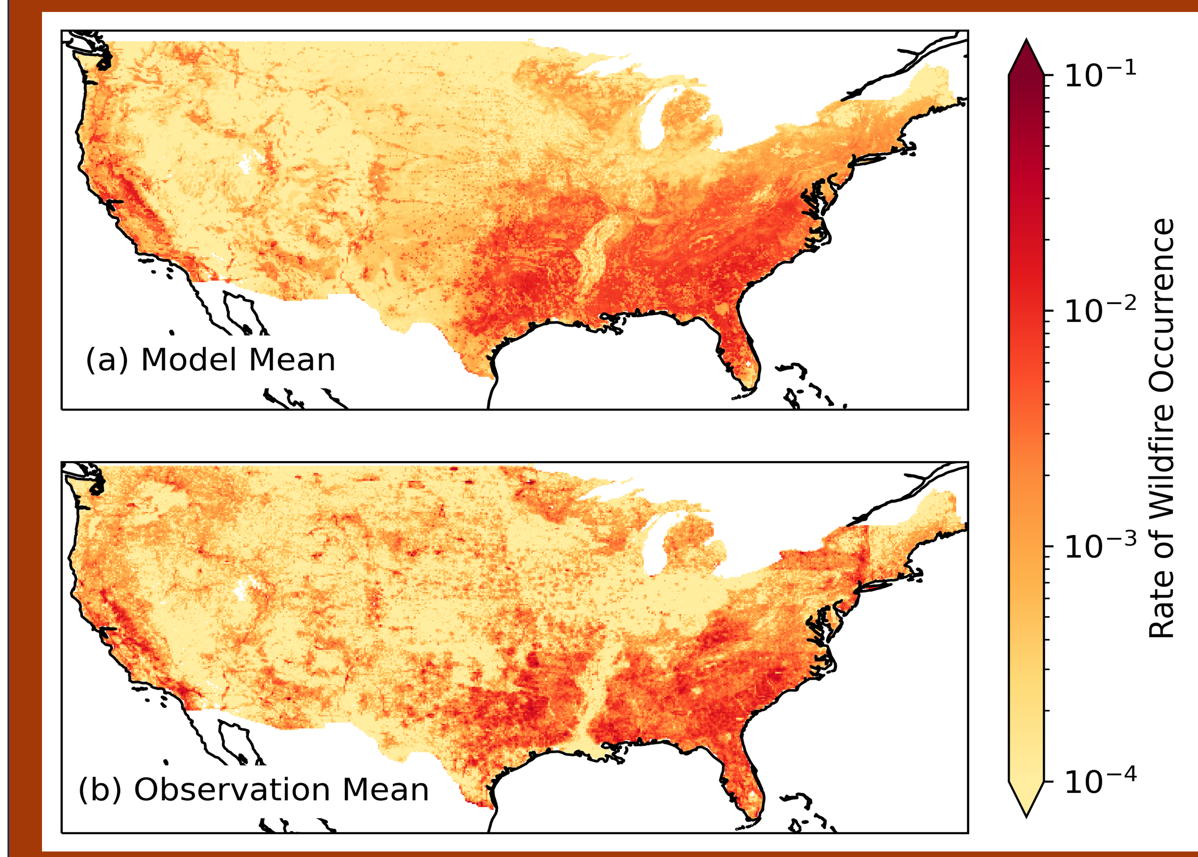

Figure 3: top, mean of daily modelled probability of wildfire occurrence. Bottom, mean observed rate of wildfire occurrence. Study period of 2002 -2018

The geospatial average of modelled daily fire occurrence probability is in good agreement with the observational record.

Interannual Variability

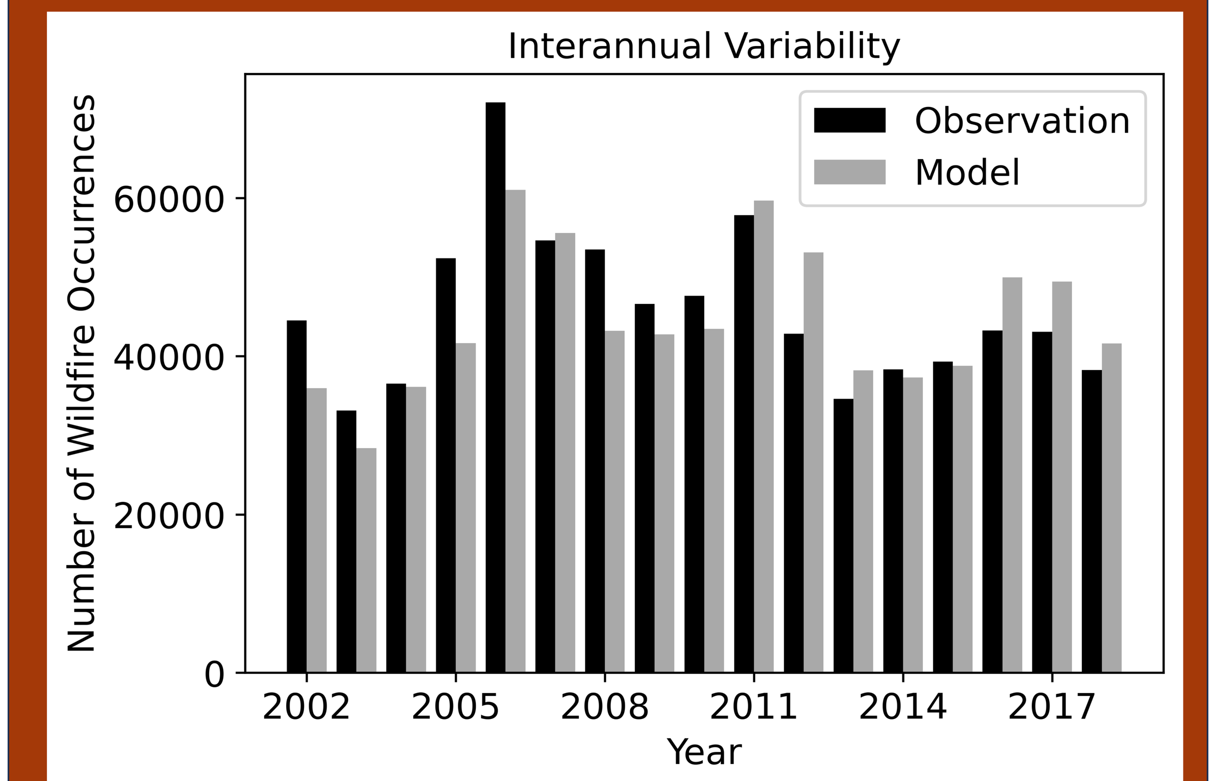

Figure 4: the total modelled and observed fire occurrences in the contiguous United States from 2002-2018.

The total annual number of wildfire occurrences simulated by the model corresponds well to observed high and low fire years.

The Seasonal Cycle

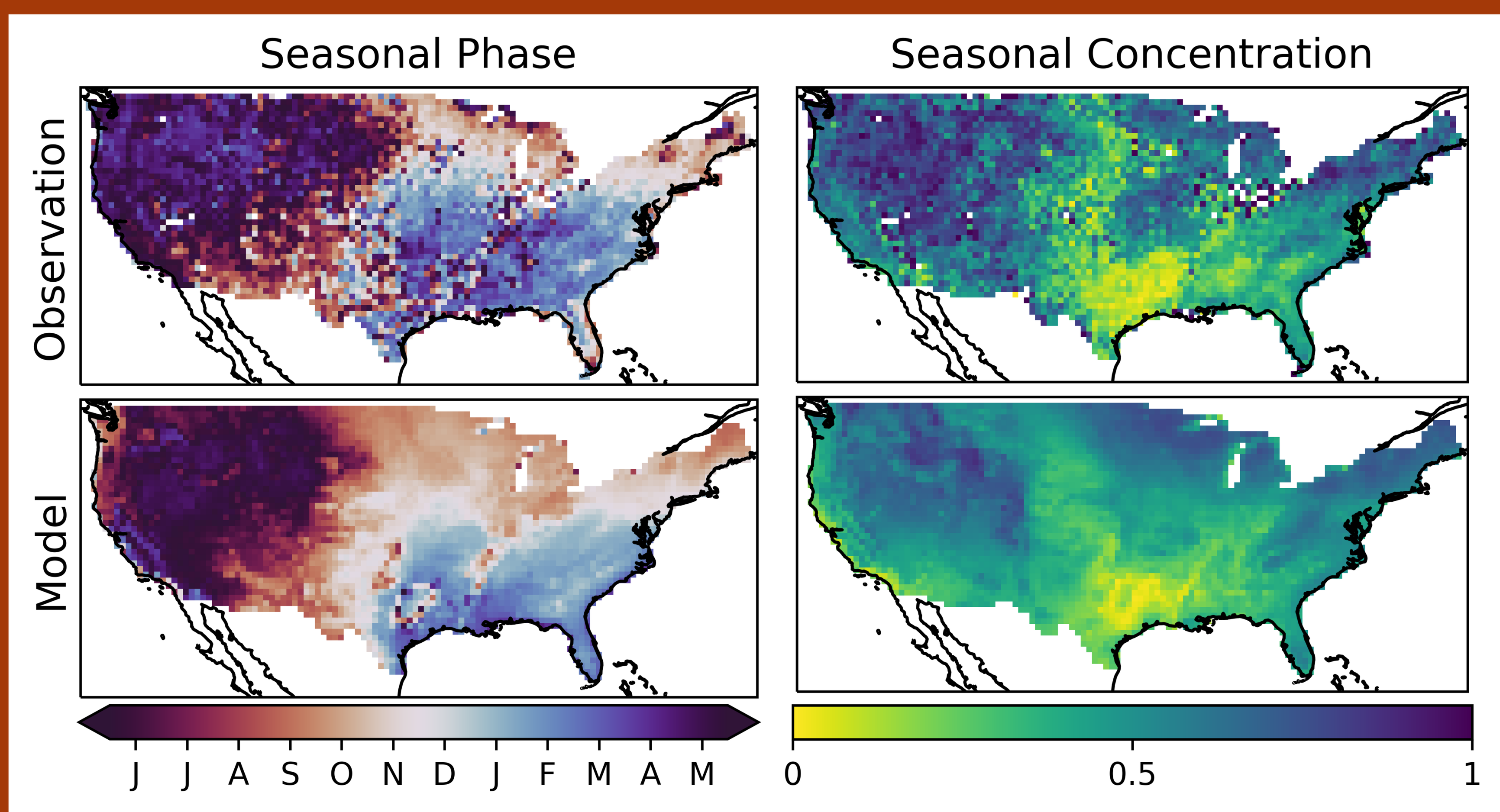

Figure 5 (above): comparative plots of the modelled and observed seasonal concentration and phase (defined below). Good correspondence is broadly shown with the exception of the NW coast.

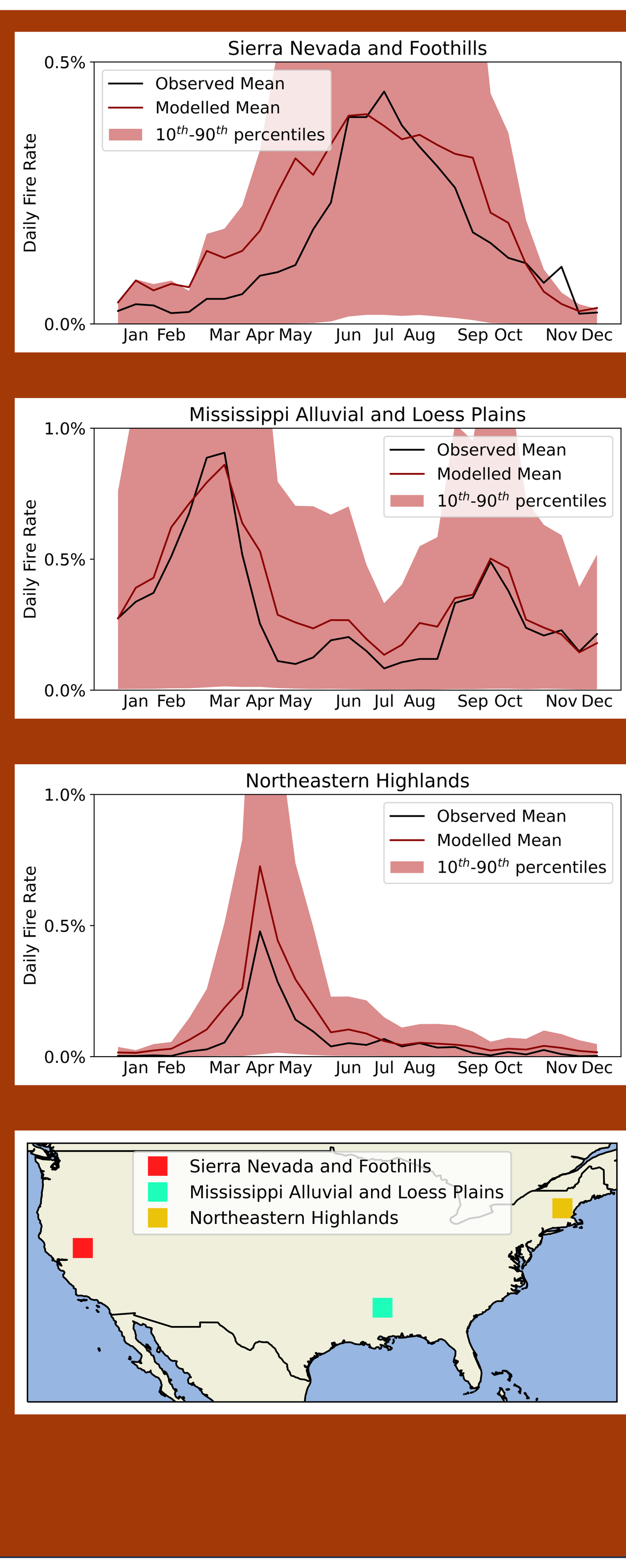

Figure 6 (right): the modelled (red) and observed (black) seasonal cycles for wildfire occurrence in three characteristic regions. Exhibiting the central summer-peak associated with the arid West; the bimodal cycle associated with regions having wet or humid summers; and, the spring peak and long tail observed in the North East.

SEASONAL PHASE:

The average centre of the fire season. Most meaningful when the season is characterised by a single, symmetrical peak.

SEASONAL CONCENTRATION:

The extent to which fire occurrence is clustered in the year, where 0 indicates that there are the same number of fires each month and 1 indicates fires occur in a single month.

Table 1: the minimum, mean and maximum of key model benchmarks. The geospatial, seasonal and interannual benchmarks are those used in FireMIP (Kelley et al., 2013). AUC is the area under the receiver operating curve – a metric for the ranking accuracy of the model

	Min	Mean	Max
AUC	0.853	0.865	0.875
NME Interannual	0.660	0.669	0.682
NME Seasonal Concentration	0.702	0.833	1.024
NME Geospatial	0.425	0.453	0.486
Seasonal MPD	0.127	0.153	0.174

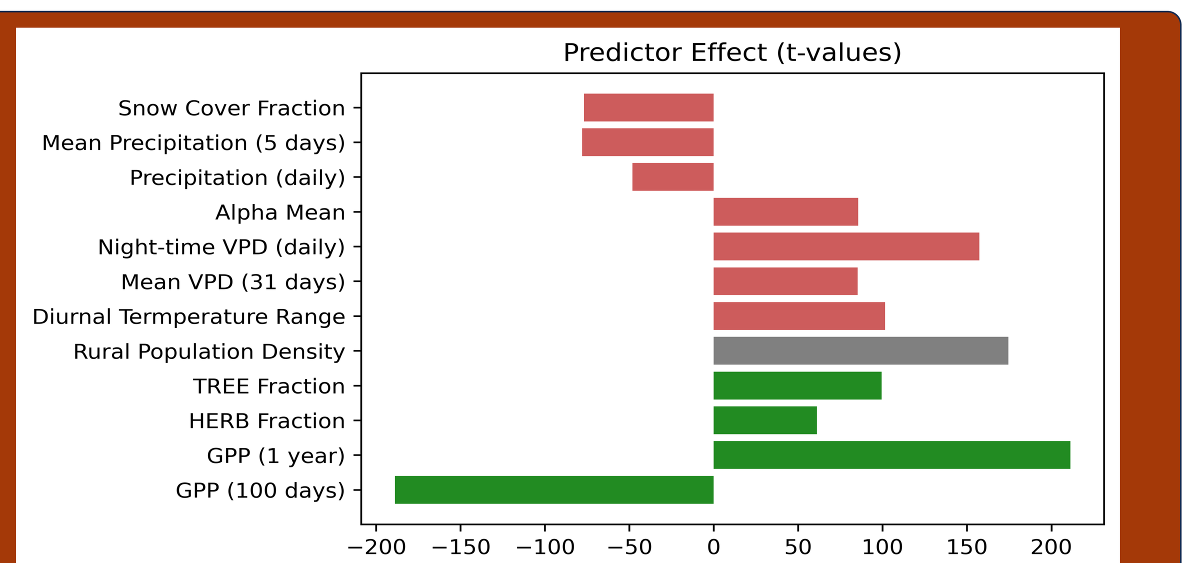

Figure 7: the t-values for each of predictors in the best performing model (defined as the model with the highest AUC of the pareto-superior set).

Citations:

1. Nelder, J.A. and Wedderburn, R.W., 1972. Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *135*(3), pp.370-384.
2. Short, K.C., 2021. Spatial wildfire occurrence data for the United States, 1992-2018 [FPA_FOD_20210617].
3. Kelley, D.I., Prentice, I.C., Harrison, S.P., Wang, H., Simard, M., Fisher, J.B. and Willis, K.O., 2013. A comprehensive benchmarking system for evaluating global vegetation models. *Biogeosciences*, *10*(5), pp.3313-3340.

@KeepingTheo

Theodore Keeping

Email: t.r.keeping@pgr.reading.ac.uk