Interpretable Model Captures Complex Relationship between Climate Variability and Fire Season Intensity in Maritime Southeast Asia

William Daniels[1],

R. R. Buchholz[2], H. M. Worden[2], F. Ahamad[3], D. M. Hammerling[1]

[1]Colorado School of Mines, Golden, CO, USA

[2]National Center for Atmospheric Research, Boulder, CO, USA

[3]AQ Expert Solutions, Jalan Dato Muda Linggi, Negeri Sembilan, Malaysia

IAWF Fire and Climate 2022

wdaniels@mines.edu

May 26, 2022

COLORADO SCHOOL OF MINES

NSF

NCAR UCAR

A great team!

Dr. Rebecca Buchholz

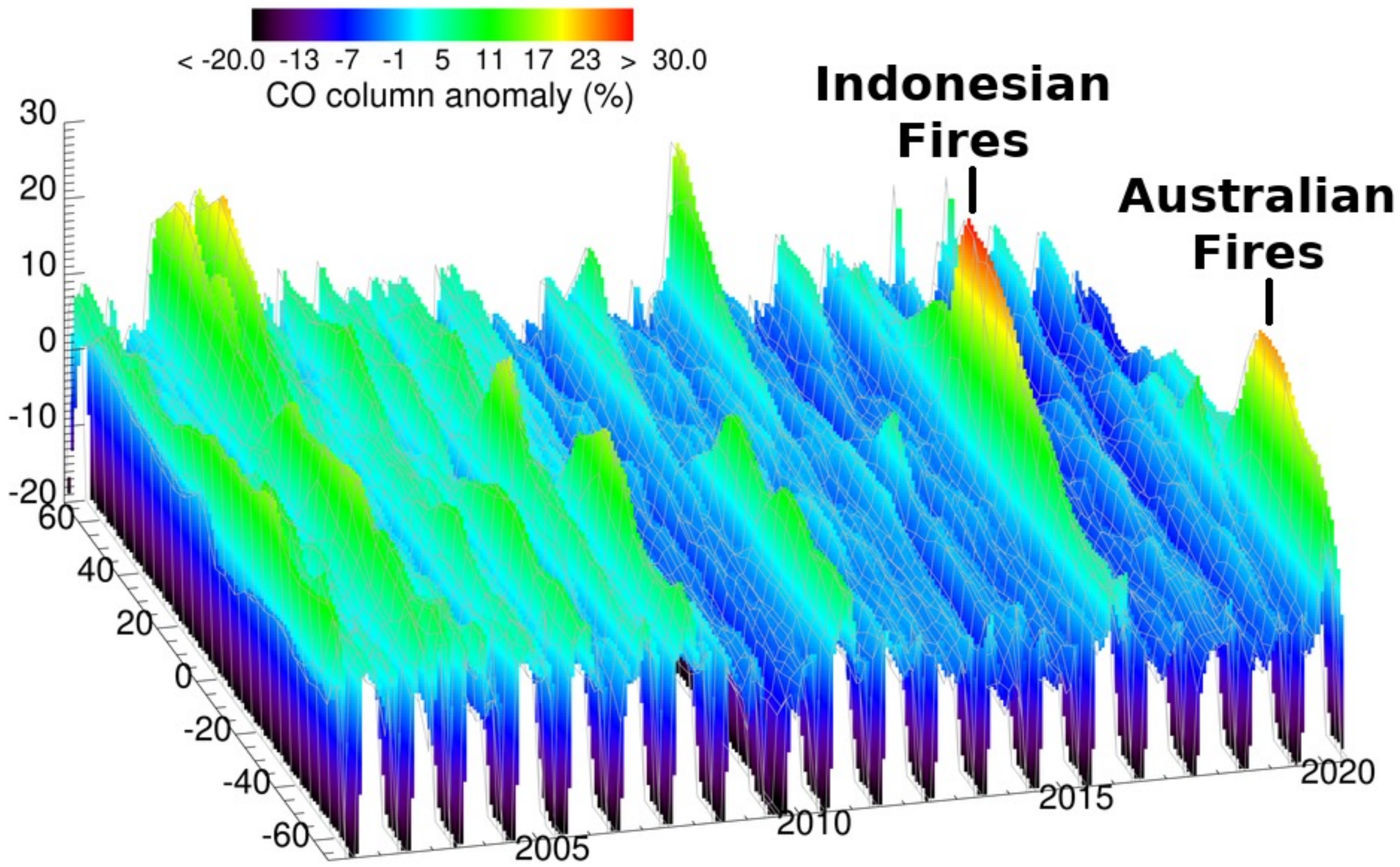
Dr. Helen Worden

Dr. Fatimah Ahamad

Dr. Dorit Hammerling

Motivation

Certain Southern Hemisphere regions experience extreme carbon monoxide (CO) anomalies as a result of biomass burning.



October 2015

Palangkaraya, Indonesia

January 2020

Canberra, Australia

Motivation

Certain Southern Hemisphere regions experience extreme carbon monoxide (CO) anomalies as a result of biomass burning.

October 2015

Palangkaraya, Indonesia

Our goals:

1. Predict CO at useful lead times
2. Build interpretable models for scientific conclusions

January 2020

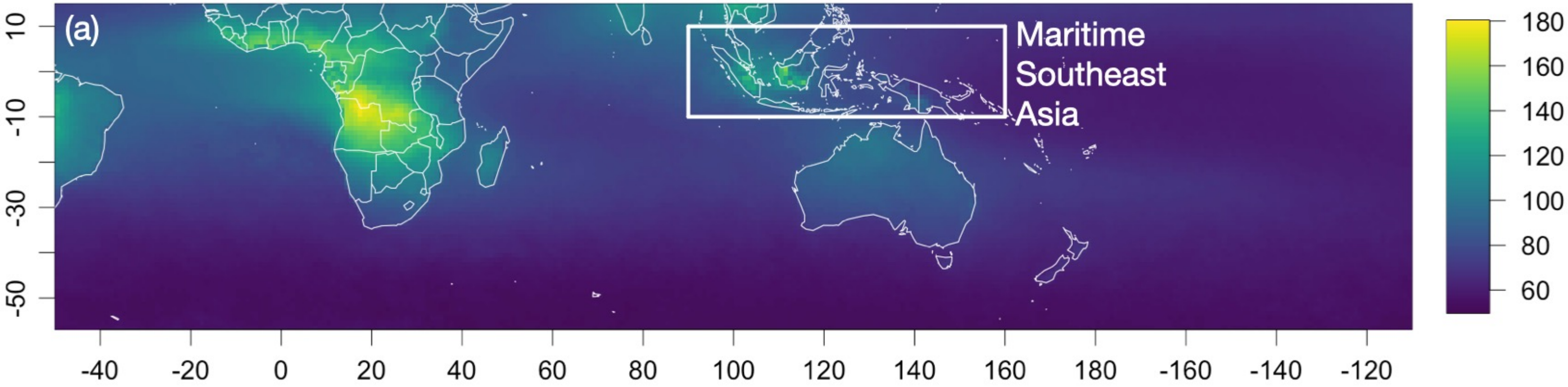
Canberra, Australia

Response variable: carbon monoxide

Use multiple linear regression to model atmospheric CO.

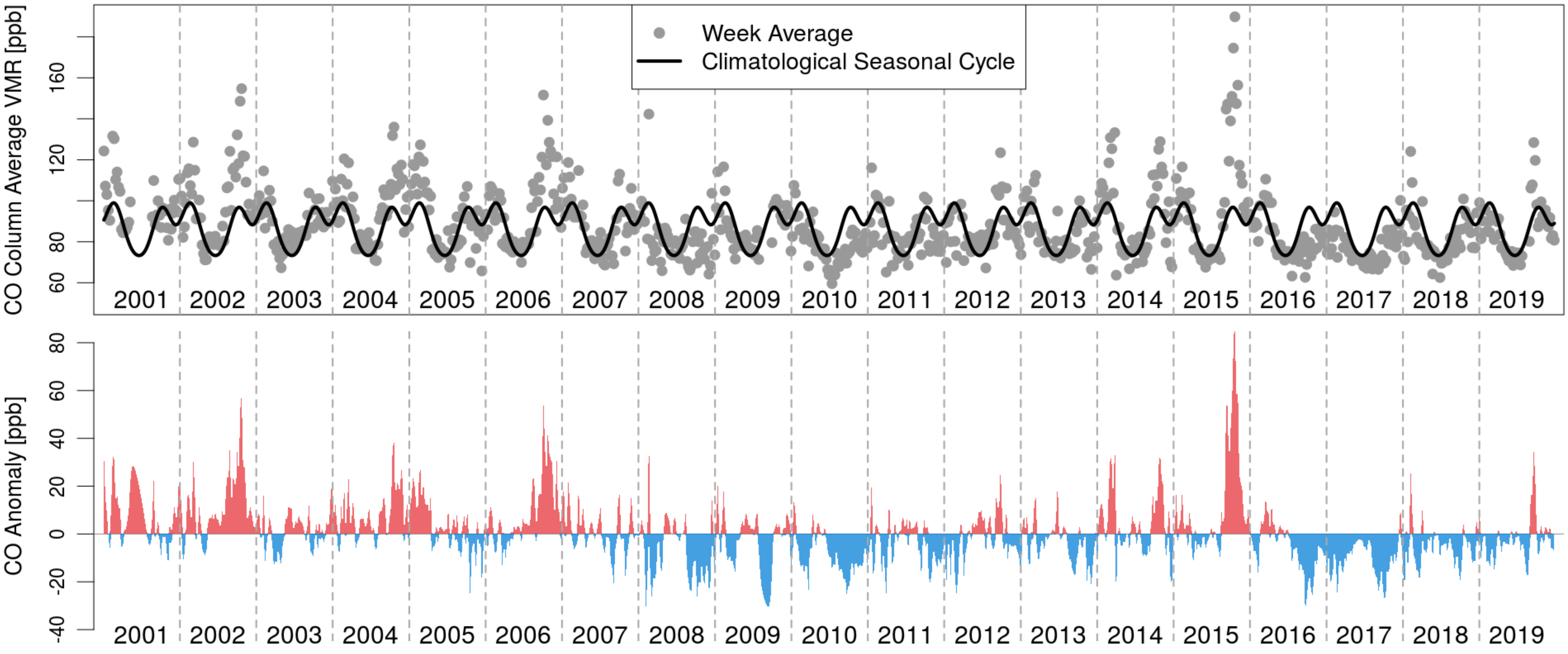
CO aggregated within the MSEA biomass burning region via spatial and temporal averages.

Mean carbon monoxide [ppb]



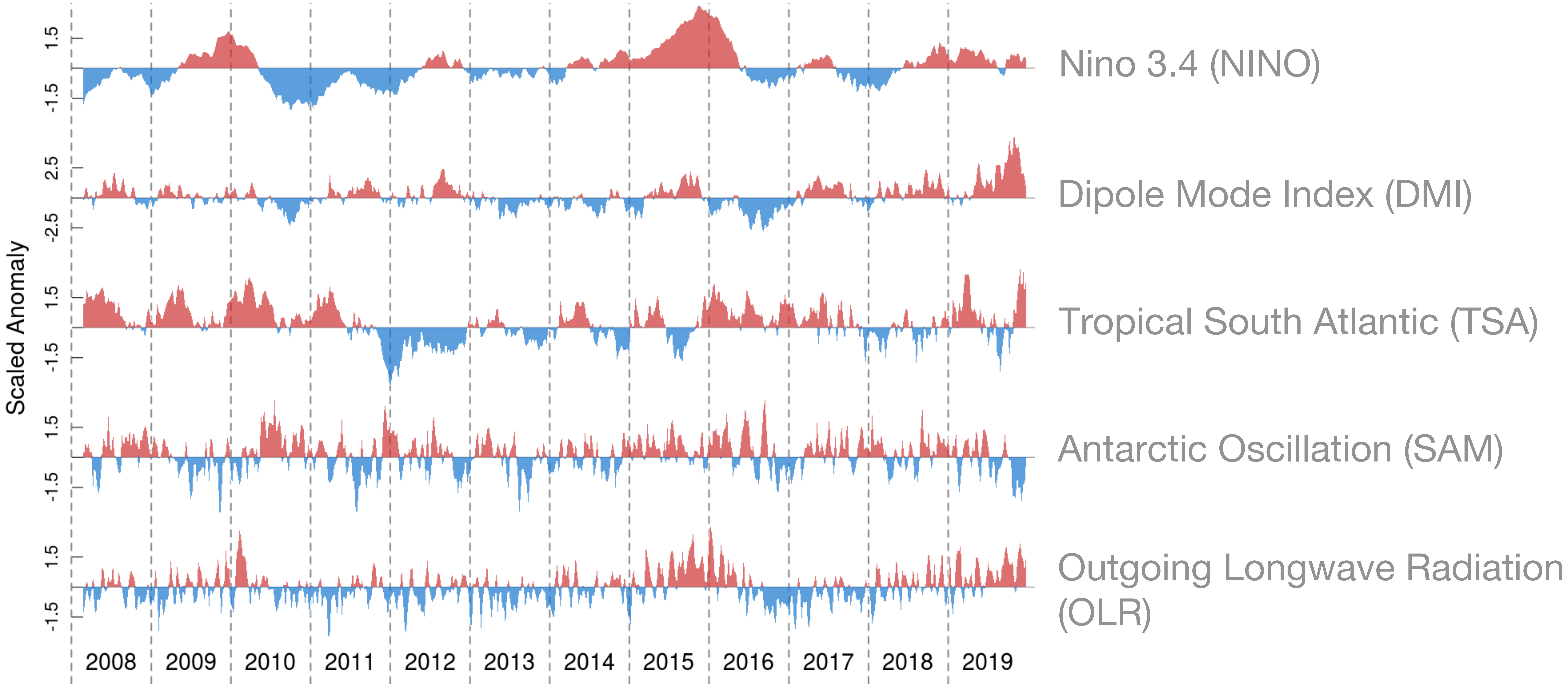
Response variable: carbon monoxide

Response variable: Deseasonalized, week-averaged CO anomalies at time t



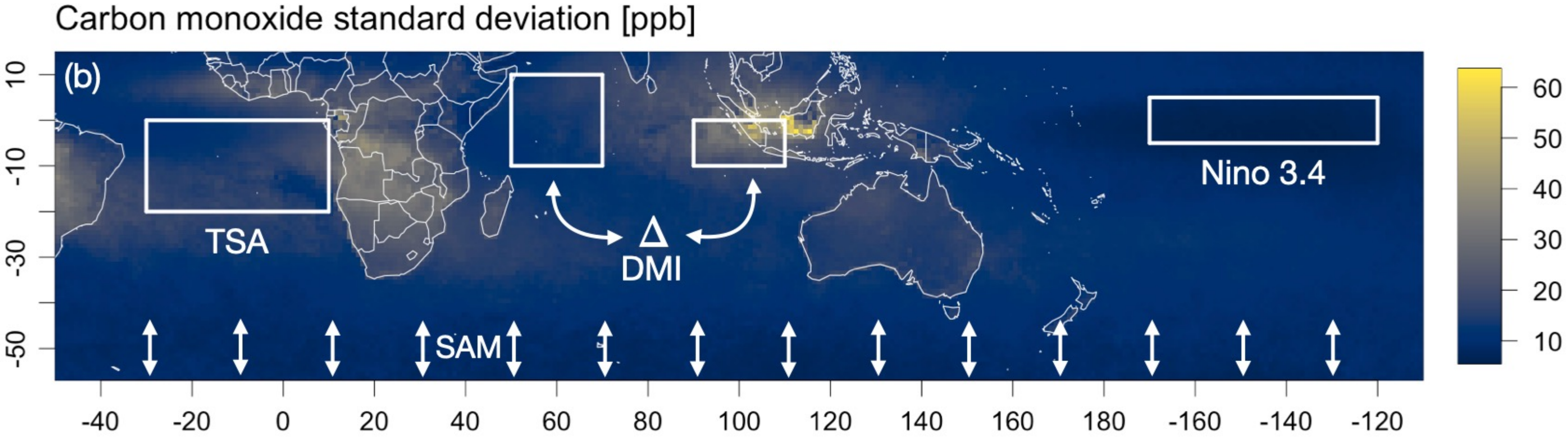
Covariates: climate mode indices

Climate mode indices are metrics that describe aperiodic variability in climate



Covariates: climate mode indices

Covariates: Week-averaged climate mode indices lagged at time t - τ



Statistical model

We use lagged multiple linear regression model with first order interactions and squared terms

$$CO(t) = \mu + \underbrace{\boxed{\sum_k a_k \chi_k(t-\tau_k)}}_{\text{Main effects}} + \underbrace{\boxed{\sum_{i,j} b_{ij} \chi_i(t-\tau_i)\chi_j(t-\tau_j)}}_{\text{Interaction terms}} + \underbrace{\boxed{\sum_l c_l \chi_l(t-\tau_l)^2}}_{\text{Squared terms}} + \epsilon(t)$$

$CO(t)$ - CO anomaly in a given response region at time t

μ - constant mean displacement

χ - climate indices

τ - lag value for each index in weeks

$\epsilon(t)$ - error term

Regularization framework for variable and lag selection

We consider lags between 1 and 52 weeks for each index

- Results in far more covariates than observations
- Regularization well suited for this regime

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{q} \beta_j X_{ij} \right)^2 + p(\beta)$$

Regularization framework for variable and lag selection

We consider lags between 1 and 52 weeks for each index

- Results in far more covariates than observations
- Regularization well suited for this regime

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{q} \beta_j X_{ij} \right)^2 + p(\beta)$$

We use the minimax concave penalty (MCP)

MCP $p(\beta) = \sum_{j=1}^{q} f(\beta_j)$ where $f(\beta_j) = \begin{cases} \lambda|\beta_j| - \frac{\beta_j^2}{2\eta} & \text{if } |\beta_j| \leq \eta\lambda \\ \frac{\eta\lambda^2}{2} & \text{otherwise.} \end{cases}$

Regularization framework for variable and lag selection

	η_1	η_2	η_3	…
λ_1	$\text{Model}_{1,1}$	$\text{Model}_{1,2}$	$\text{Model}_{1,3}$	
λ_2	$\text{Model}_{2,1}$	$\text{Model}_{2,2}$	$\text{Model}_{2,3}$	
λ_3	$\text{Model}_{3,1}$	$\text{Model}_{3,2}$	$\text{Model}_{3,3}$	
…				

Pick best model using the Extended Bayesian Information Criterion (EBIC)

- Balances model fit and complexity
- Control penalty with free parameter $\gamma \in [0,1]$
- $\gamma \to 1$ results in smaller models
- $\gamma \to 0$ results larger models

Parameter summary

Regularization → λ

MCP → η

EBIC → γ

Interpretable models lead to scientific conclusions

$$\gamma = 1$$

```
             Est (Std. Error)
(Intercept)    -1.6 (0.78)
nino_4          7.2 (0.78)
dmi_4           7.2 (0.93)
dmi_12         -8.0 (0.87)
aao_51         -3.1 (0.67)
olr_1           3.5 (0.79)
I(nino_4^2)     2.5 (0.54)
nino_4:olr_1    3.5 (0.76)
nino_4:dmi_12  -6.5 (0.77)
aao_51:olr_1   -2.3 (0.67)

Adjusted R-squared:  0.60
```

Smallest model highlights important climate-chemistry connections:

1. NINO has strong influence on CO at a four week lead time

Interpretable models lead to scientific conclusions

$$\gamma = 1$$

```
              Est (Std. Error)
(Intercept)    -1.6 (0.78)
nino_4          7.2 (0.78)
dmi_4           7.2 (0.93)
dmi_12         -8.0 (0.87)
aao_51         -3.1 (0.67)
olr_1           3.5 (0.79)
I(nino_4^2)     2.5 (0.54)
nino_4:olr_1    3.5 (0.76)
nino_4:dmi_12  -6.5 (0.77)
aao_51:olr_1   -2.3 (0.67)

Adjusted R-squared:  0.60
```

Smallest model highlights important climate-chemistry connections:

1. NINO has strong influence on CO at a four week lead time
2. Effect of DMI depends on length of lag

Interpretable models lead to scientific conclusions

$$\gamma = 1$$

	Est	(Std. Error)
(Intercept)	-1.6	(0.78)
nino_4	7.2	(0.78)
dmi_4	7.2	(0.93)
dmi_12	-8.0	(0.87)
aao_51	-3.1	(0.67)
olr_1	3.5	(0.79)
I(nino_4^2)	2.5	(0.54)
nino_4:olr_1	3.5	(0.76)
nino_4:dmi_12	-6.5	(0.77)
aao_51:olr_1	-2.3	(0.67)

Adjusted R-squared: 0.60

Smallest model highlights important climate-chemistry connections:

1. NINO has strong influence on CO at a four week lead time
2. Effect of DMI depends on length of lag
3. NINO interactions suggest that NINO amplifies effect of other indices

Model has good predictive skill at useful lead time

$$\gamma = 0$$

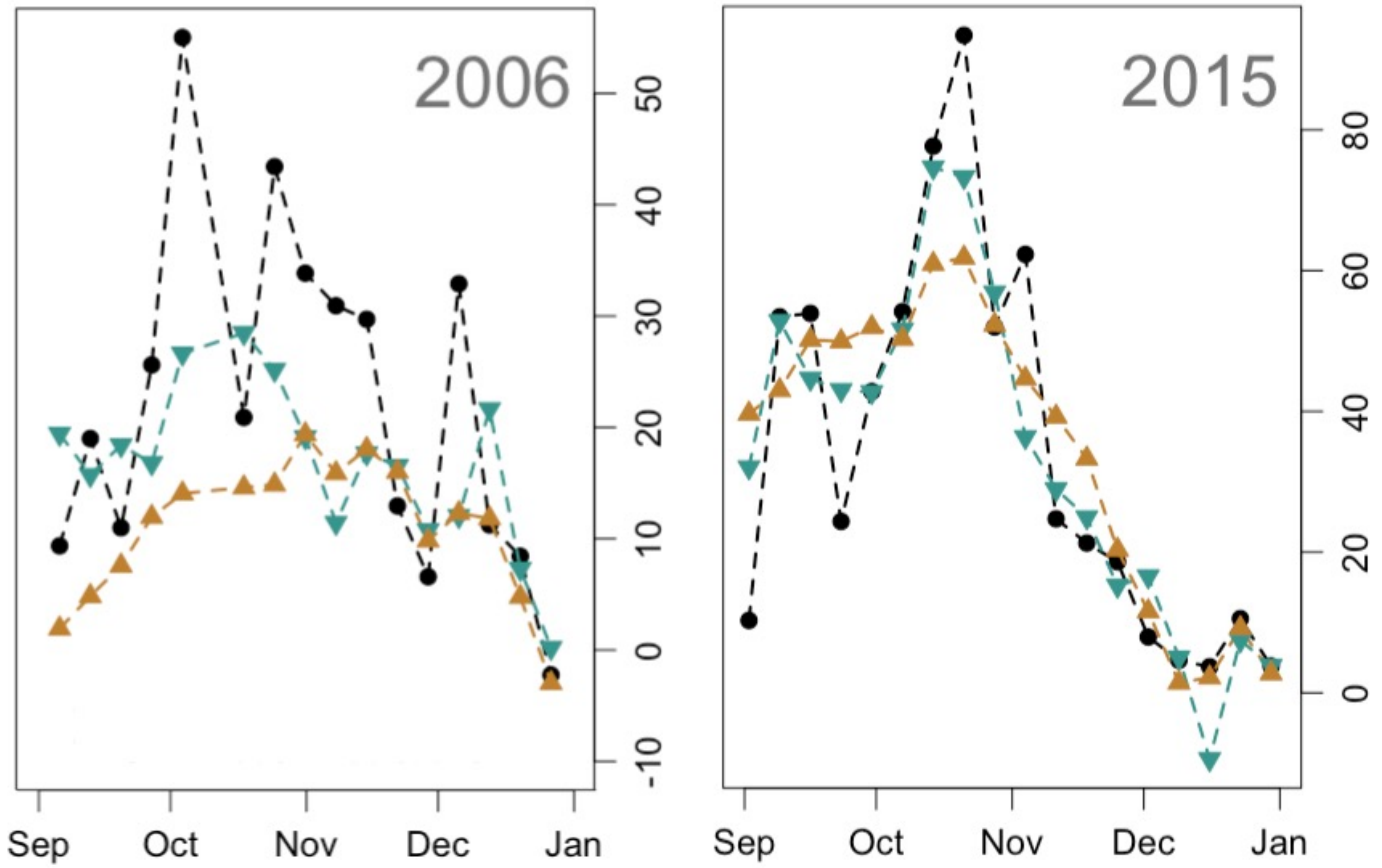
```
                 Est (Std. Error)
(Intercept)          0.3 (0.70)
nino_4               7.6 (0.83)
dmi_1                5.7 (0.79)
dmi_12              -6.1 (0.75)
dmi_43               1.8 (0.65)
tsa_3               -2.2 (0.64)
aao_2               -3.6 (0.61)
aao_38              -2.2 (0.64)
aao_51              -1.6 (0.63)
olr_1                2.3 (0.74)
olr_13               3.4 (0.71)
nino_4:olr_1         3.2 (0.66)
nino_4:dmi_1         3.2 (0.81)
dmi_1:dmi_12        -4.5 (0.56)
nino_4:aao_51       -4.2 (0.77)
tsa_3:olr_1         -2.3 (0.63)
aao_2:olr_13        -2.1 (0.68)
nino_4:aao_2        -1.8 (0.70)

Adjusted R-squared:  0.68
```

OLR helps capture the most extreme CO anomalies

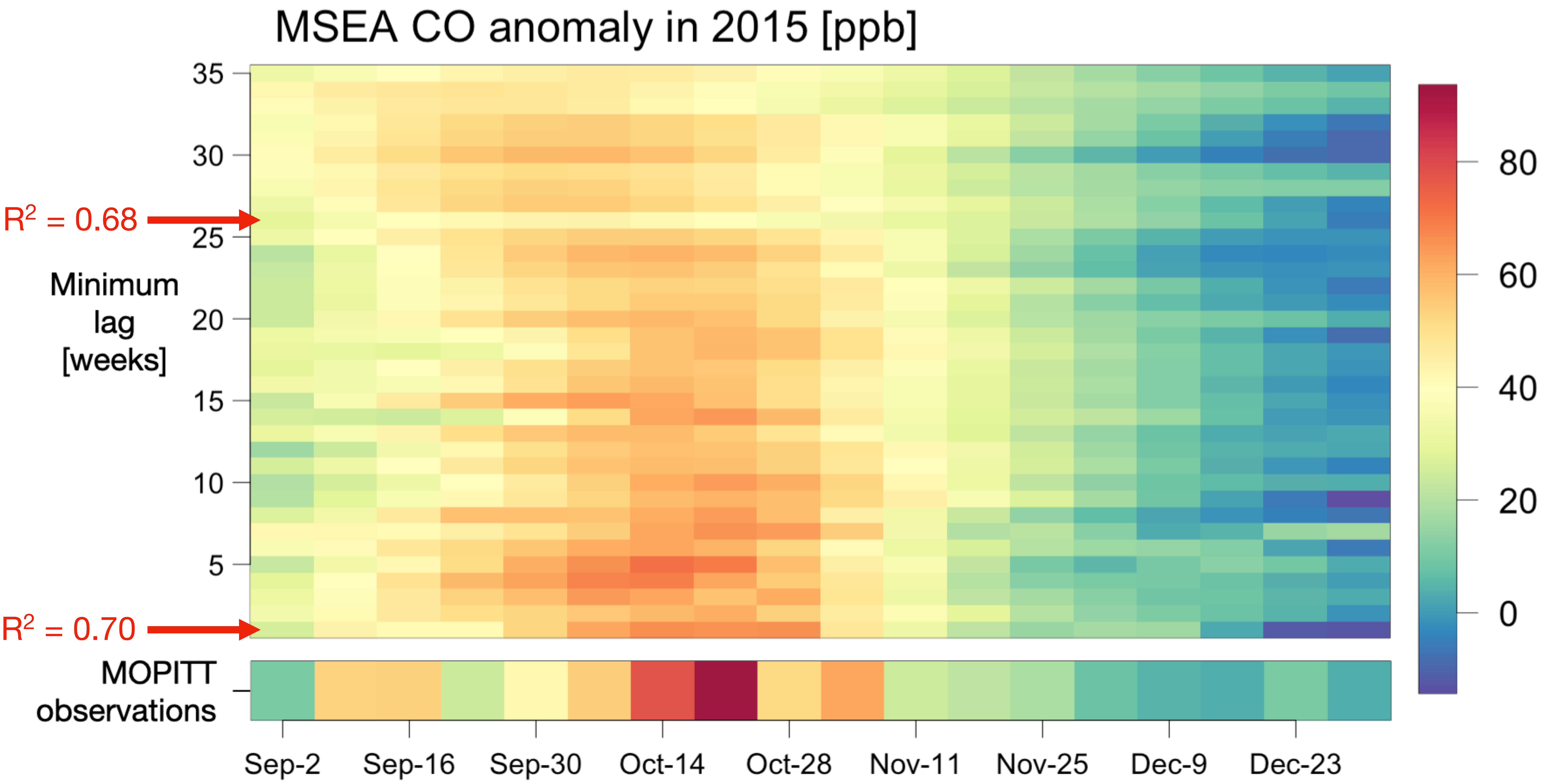
● Weekly Observations
▲ No OLR Model Predictions
▼ OLR Model Predictions

Adjusted R^2

No OLR Model	OLR Model
0.66	0.68



Model has good predictive skill at useful lead time



Conclusions

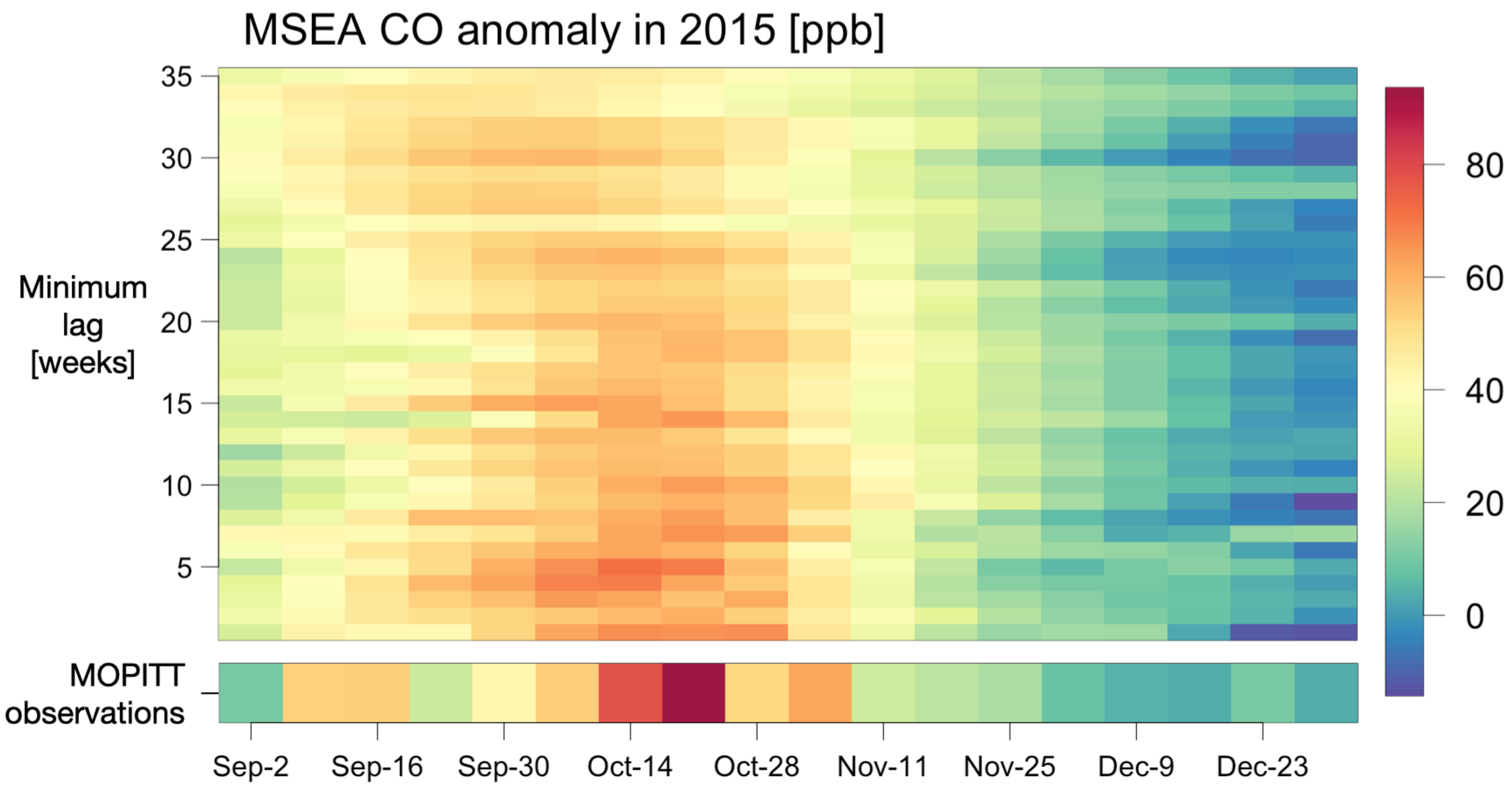
We are using natural variability in the climate to model atmospheric CO (a proxy for fire intensity)

Interpretable models

	Est	(Std. Error)
(Intercept)	-1.6	(0.78)
nino_4	7.2	(0.78)
dmi_4	7.2	(0.93)
dmi_12	-8.0	(0.87)
aao_51	-3.1	(0.67)
olr_1	3.5	(0.79)
I(nino_4^2)	2.5	(0.54)
nino_4:olr_1	3.5	(0.76)
nino_4:dmi_12	-6.5	(0.77)
aao_51:olr_1	-2.3	(0.67)

Adjusted R-squared: 0.60

Good predictive skill



Thank you! Questions?

See manuscript on EarthArXiv for details:

William Daniels

wdaniels@mines.edu

IAWF Fire and Climate 2022

COLORADO SCHOOL OF MINES

NCAR UCAR