Interpretable Models Capture the Complex Relationship Between Climate Indices and Fire Season Intensity in Maritime Southeast Asia

William S Daniels[1], Rebecca R Buchholz[2], Helen M Worden[2], Fatimah Ahamad[3], and Dorit Hammerling[1]

[1]Colorado School of Mines
[2]National Center for Atmospheric Research (UCAR)
[3]AQ Expert Solutions

November 22, 2022

Abstract

There have been many extreme fire seasons in Maritime Southeast Asia (MSEA) over the last two decades, a trend which will likely continue or accelerate due to climate change. Fires, in turn, are a major driver of atmospheric carbon monoxide (CO) variability, especially in the Southern Hemisphere. Here we attempt to maximize the amount of CO variability that can be explained via human-interpretable statistical models that use only climate mode indices as predictor variables. We expand upon previous work through the complexity at which we study the connections between climate mode indices and atmospheric CO (a proxy for fire intensity). Specifically, we present three modeling advancements. First, we analyze five different climate modes at a weekly timescale, which increases explained variability by 15% over models on a monthly timescale. Second, we accommodate multiple lead times for each climate mode index, finding that some indices have very different effects on CO at different lead times. Finally, we model the interactions between climate mode indices at weekly timescales, which provides a framework for studying these interactions at a higher level of complexity than previous work. Furthermore, we perform a stability analysis and show that our model for the MSEA region is robust, which adds weight to the scientific interpretation of the selected model terms. We believe that the complex relationships quantified here will be useful for scientists studying modes of variability in MSEA and for forecasters looking to maximize the information they glean from climate modes.

Interpretable Models Capture the Complex Relationship Between Climate Indices and Fire Season Intensity in Maritime Southeast Asia

William S. Daniels[1], Rebecca R. Buchholz[2], Helen M. Worden[2], Fatimah Ahamad[3], Dorit 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

Key Points:

- We quantify the connections between climate and carbon monoxide (as a proxy for fire intensity) in more detail than previous work.
- Our model explains 70% of the variability in atmospheric carbon monoxide on a weekly timescale using only climate mode indices.
- The impact of certain indices on carbon monoxide variability changes as their lead time in the model increases.

Corresponding author: William Daniels, `wdaniels@mines.edu`

Abstract

There have been many extreme fire seasons in Maritime Southeast Asia (MSEA) over the last two decades, a trend which will likely continue or accelerate due to climate change. Fires, in turn, are a major driver of atmospheric carbon monoxide (CO) variability, especially in the Southern Hemisphere. Here we attempt to maximize the amount of CO variability that can be explained via human-interpretable statistical models that use only climate mode indices as predictor variables. We expand upon previous work through the complexity at which we study the connections between climate mode indices and atmospheric CO (a proxy for fire intensity). Specifically, we present three modeling advancements. First, we analyze five different climate modes at a weekly timescale, which increases explained variability by 15% over models on a monthly timescale. Second, we accommodate multiple lead times for each climate mode index, finding that some indices have very different effects on CO at different lead times. Finally, we model the interactions between climate mode indices at weekly timescales, which provides a framework for studying these interactions at a higher level of complexity than previous work. Furthermore, we perform a stability analysis and show that our model for the MSEA region is robust, which adds weight to the scientific interpretation of the selected model terms. We believe that the complex relationships quantified here will be useful for scientists studying modes of variability in MSEA and for forecasters looking to maximize the information they glean from climate modes.

1 Introduction

The relationship between fire and climate has been extensively studied. Fire intensity and burned area are related to the amount, type, and dryness of available fuel, all of which respond closely to water conditions driven by climate variability (van der Werf et al., 2008). This relationship is complex and varies across the different regions of the globe. For instance, drought conditions were found to increase fire potential in Southern Africa, but decrease fire potential in Northern Africa (Andela & van der Werf, 2014).

Climate modes, such as the El Niño Southern Oscillation (ENSO), capture variability in the global climate system. Studies have used these climate modes to help explain the complex relationship between climate and fire, often via regression models. ENSO has been found to influence fires in North America (Mason et al., 2017; Shabbar et al., 2011), Maritime Southeast Asia (Chen et al., 2017; Fuller & Murphy, 2006; Reid et al., 2012), the Amazon (Alencar et al., 2011; Fonseca et al., 2017), and Africa (Andela & van der Werf, 2014; N'Datchoh et al., 2015). Furthermore, studies have found that fire behavior can respond to several distinct climate modes (Andreoli & Kayano, 2006; Chen et al., 2016; Saji & Yamagata, 2003), with Cleverly et al. (2016) showing that the interactions between these climate modes are particularly important for explaining drought and rainfall in Australia (which in turn are major drivers of fire activity). This indicates that fire behavior is affected not only by the isolated influence of multiple modes, but also by their interactions (e.g., whether or not the modes are in phase).

In addition to identifying the climate modes that most influence fire behavior in a given region, studies such as Chen et al. (2016) and Wooster et al. (2012) identify lead times that correspond to the maximum predictive performance of the climate modes being studied. Similarly, Shawki et al. (2017) examines how far in advance the 2015 fire event in Indonesia can be predicted using climate based models, finding that lead times of up to 25 weeks can still provide useful predictions.

These fire-climate connections have been previously studied using satellite observations of fire properties (e.g., Ceccato et al. (2010), Chen et al. (2016), and Wooster et al. (2012)). The Moderate Resolution Imaging Spectroradiometer (MODIS) instruments onboard the Terra and Aqua satellites provide fire count data for each overpass as well as a burned area data product (Giglio et al., 2006, 2016, 2018). However, using fire counts or burned area directly presents a number of challenges. Fire count products ignore differences in fire size and intensity, burned area products can miss underground peat fires, and both products can miss fires obscured by smoke (Giglio et al., 2006, 2018; Shawki et al., 2017).

One alternative is to model atmospheric carbon monoxide (CO) instead of fire counts, burned area, or aerosol optical depth (AOD) directly. CO is produced by incomplete combustion from biomass burning, fossil fuel use, and indirectly by photochemistry (Buchholz et al., 2018; Holloway et al., 2000), and its link to fires is well established (Edwards, Emmons, et al., 2006). In fact, biomass burning is the primary source of atmospheric CO variability in the Southern Hemisphere, making CO anomalies a useful proxy for fire intensity (Bloom et al., 2015; Buchholz, Worden, Park, et al., 2021; Voulgarakis et al., 2015). Buchholz, Worden, Park, et al. (2021) show that MODIS AOD and CO observations from the Measurement of Pollution in the Troposphere (MOPITT) instrument over the Maritime Southeast Asia (MSEA) region are highly correlated, further justifying the use of CO as an alternative to fire products or AOD. Since CO variability in the Southern Hemisphere is closely linked to biomass burning (and biomass burning responds to variability in the climate), we expect that CO also responds to climate variability. Compared to the study of fire counts, burned area, or AOD, less research has gone into the connection between atmospheric CO and climate variability. Furthermore, modeling atmospheric CO concentrations provides information on co-emitted atmospheric pollutants in addition to being a proxy for fire intensity.

Edwards, Pétron, et al. (2006) found that CO observations from MOPITT are correlated with ENSO. Buchholz et al. (2018) expanded on Edwards, Pétron, et al. (2006) by showing that atmospheric CO anomalies in a number of Southern Hemisphere regions are related to four different climate modes (including ENSO) and that the interactions between these climate modes are important for explaining atmospheric CO anomalies. In this study, we examine the relationship between atmospheric CO and climate variability, further focusing on the MSEA region because of its extremely large CO anomalies (Buchholz, Worden, Park, et al., 2021). While we focus on a single region in this paper, the modeling framework we have developed can easily be applied to other parts of the globe.

In this paper, we propose a framework for studying the connections between climate and atmospheric CO (as a proxy for fire intensity) in more detail than previous work. To do this, we extend the models from Buchholz et al. (2018) via the following advancements. First, we use week-averaged data rather than month-averaged data, significantly increasing predictive performance. Second, we include the Madden-Julian Oscillation (MJO) via a proxy index, resulting in models that are better able to capture extreme CO anomalies in MSEA. Third, we develop a regularization-based model fitting framework that allows for models with multiple lags of a single climate mode. Fourth, we assess the stability of the selected model terms, which adds weight to their scientific interpretation and increases overall model interpretability. Finally, we explore the use of our model in a forecasting setting to assess how much variability can be explained using climate mode indices alone. Note that we do not attempt to outperform or even match current forecasting tools that utilize additional modes of variability beyond climate modes (e.g., Groot et al. (2006), Shawki et al. (2017)), as we are only interested in the connections between climate modes and CO. These advancements result in models that capture more complex relationships and have better predictive performance than those presented in Buchholz et al. (2018) while remaining human-interpretable. We believe that these models will be useful for scientists studying modes of variability in MSEA and forecasters looking to maximize the information they glean from climate modes.

The rest of this paper is laid out as follows. In Sections 2 and 3, we describe the data and our statistical model, respectively. In Section 4, we discuss our model fitting

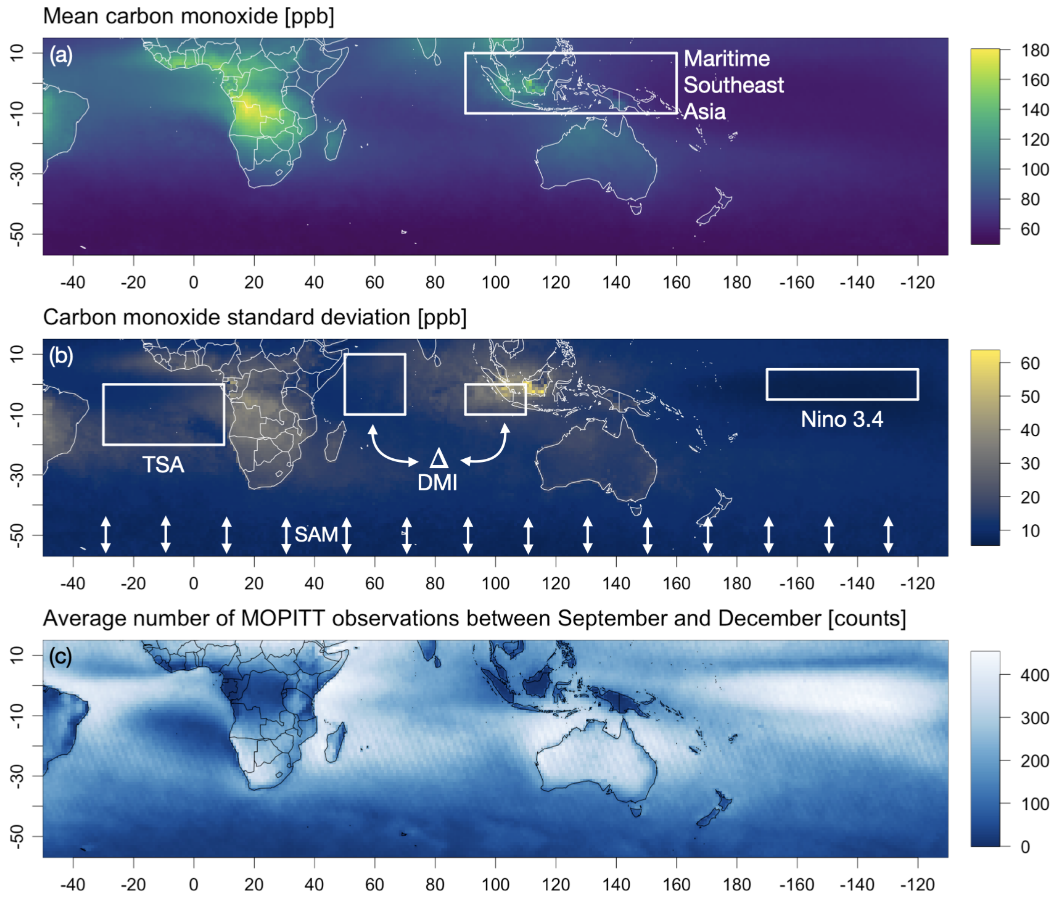

Figure 1. MOPITT CO data during the Southern Hemisphere fire season (defined here as September through December) from 2001 to 2019. Data are filtered as described in Section 2.1. (a) Average of all MOPITT CO observations (n = 217,995,648) with the Maritime Southeast Asia (MSEA) region shown in white (n = 12,985,456). (b) CO standard deviation with the spatial range of influence of the four climate mode indices discussed in Section 2.2 shown in white. (c) Average number of MOPITT observations falling within each grid cell during fire season. Note that the landmasses in MSEA have fewer observations than other regions, which could be influencing the high CO standard deviations in this region. All three subfigures are plotted on the same 1°×1° grid.

framework. In Sections 5 and 6, we present results and assess improvements in model interpretability and predictive performance, respectively, over the models presented in Buchholz et al. (2018). Finally, we summarize our work in Section 7.

2 Observational Data Sets

We model atmospheric CO using a linear regression framework in which the response variable (CO) is modeled as a linear combination of predictor variables (climate mode indices and their proxies). The following subsections describe the data used as our response and predictor variables. Note that "covariate" is synonymous with "predictor variable" and is used throughout for brevity.

2.1 Response Variable

For the response, we use carbon monoxide column-averaged volume mixing ratios (referred to as simply CO) from the MOPITT instrument onboard the Terra satellite (Drummond et al., 2010). The units of column-averaged volume mixing ratios (VMR) are parts per billion by volume (ppb). Using column-averaged volume mixing ratios instead of total column CO removes dependence on surface topography and pressure changes (Buchholz, Worden, Park, et al., 2021).

MOPITT has complete Earth coverage about every three days with a footprint size of 22×22 km^2. We use the V8 retrieval algorithm with validation results described in Deeter et al. (2019). To reduce systematic and random error, we select daytime, land-only retrievals from the joint near infrared (NIR) and thermal infrared (TIR) product. Daytime retrievals over land have a higher sensitivity to CO than nighttime or ocean retrievals due to higher thermal contrast. We use the joint product because it includes additional information from reflected solar radiation over land (Worden et al., 2010). See Buchholz et al. (2018), Deeter et al. (2007), and Deeter et al. (2014) for details.

Because MOPITT retrievals are dependent on clear sky conditions, we expect sampling error to both bias our CO time series lower and increase its variability. This is because MOPITT observations might not be available nearest to fire source regions and cloud patterns can significantly reduce the amount of data available over the region. This issue is also present in other satellite-observed data sets, such as fire counts or aerosol optical depth (Reid et al., 2012). However, we do not expect these features to significantly impact our results for two reasons. First, the magnitude of the response will only impact the magnitude of the fitted coefficients, not their relationship relative to each other. Therefore, interpretation of selected model terms is still valid in a relative sense. Second, linear models fit via regularization (which we employ and discuss in Sections 3 and 4) are well suited for handling noisy or variable data and will not overfit to the noise when tuned correctly. An analysis of how much variability in our response is attributed to cloud sampling is the focus of another study.

We aggregate CO observations into a single biomass burning region in the Southern Hemisphere: Maritime Southeast Asia (MSEA), defined here as -10° to 10° latitude and 90° to 160° longitude (see Figure 1(a)). We focus on MSEA because it is a biomass burning region that experiences significant CO anomalies, or concentrations well above average (Buchholz, Worden, Park, et al., 2021). Note that there are fewer MOPITT observations over land within the MSEA region on average (see Figure 1(c)). This is likely a result of higher cloud fractions and geophysical noise over land scenes compared to water scenes in MSEA. The Supporting Information file contains a plot and discussion of the cloud fraction from the Terra-MODIS cloud mask over MSEA. We create a weekly time series for MSEA by averaging all of the observations falling within the region boundaries for each week. This time series ranges from 2001 to 2019, resulting in 19 years of data and 991 weekly averages. Despite the relatively lower number of observations falling over MSEA landmasses, there are still 110 observations per week on average, which we deem a suitable number for creating our response variable. We compute the seasonal cycle by taking an average over the 19 years of data for each week. We then remove this seasonal cycle from the weekly time series so that our models are better able to capture the anomalous CO observations corresponding to large burn events. Figure 2 shows the weekly CO observations, climatological average, and resulting anomalies for MSEA.

Finally, since we use CO as a proxy for fire intensity, we only model anomalies during the months that experience high CO variability due to burning. Although CO variability is highest between September and November in MSEA, we use anomalies between September and December to be consistent with Buchholz et al. (2018). This time frame results in a total of 330 weekly CO anomalies for the MSEA region.

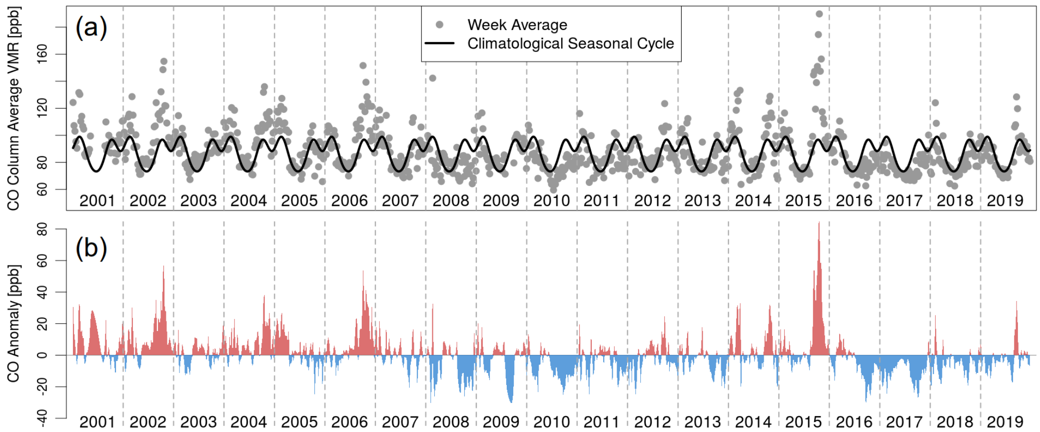

Figure 2. (a) Weekly CO observations for MSEA (grey circles) and the climatological average created by averaging each week over the 19-year time series (black line). (b) CO anomalies resulting from the difference between the weekly observations and the climatological average. Positive anomalies are shown in red and negative anomalies are shown in blue.

2.2 Predictor Variables

We are interested in connections between atmospheric CO and climate variability. Climate modes are large scale patterns that capture variation in temperature, wind, or other aspects of climate over certain spatial regions. A well known example is ENSO, which captures quasi-periodic variability in sea surface temperature and wind in the Pacific Ocean (Neelin et al., 1998; Trenberth, 2013). Climate indices are metrics that quantify the state of climate modes.

As in Buchholz et al. (2018), we consider four climate modes that represent variability in the major ocean basins of the Southern Hemisphere and tropics. The ENSO represents the Pacific Ocean, the Indian Ocean Dipole (IOD) represents the Indian Ocean, the Tropical South Atlantic (TSA) represents the southern Atlantic Ocean, and the Antarctic Oscillation (AAO) represents the Southern Ocean.

For predictor variables, we select a single climate mode index to represent each of these climate modes. To represent the ENSO, we use the Niño 3.4 index defined in Bamston et al. (1997). To represent the TSA, we use the Tropical South Atlantic Index defined in Enfield et al. (1999). These two indices are calculated using sea surface temperature (SST) anomalies in the regions shown in Figure 1(b) labeled as Nino 3.4 and TSA, respectively. To represent the IOD, we use the Dipole Mode Index (DMI) defined in Saji et al. (1999). This index is calculated from SST gradients between the two regions shown in Figure 1(b) labeled as DMI. To represent the AAO, we use the Southern Annular Mode (SAM) index defined in Thompson and Wallace (2000). This index captures Antarctic atmospheric circulation described by the poleward shift of westerly winds. This index is calculated by projecting observational height anomalies at 700 hPa and poleward of -20 degrees latitude onto the leading empirical orthogonal function of the National Centers for Environmental Prediction and National Center for Atmospheric Research reanalysis (Kalnay et al., 1996; Kistler et al., 2001). The spatial extent of this index is shown in Figure 1(b) via the arrows labeled SAM. We expect a relationship between these indices and CO, as each index is related to regional climate (e.g., rainfall), which in turn affects drought, fire, and ultimately CO concentrations.

In addition to these four indices, we also want to include variability captured by the MJO in our models. This climate mode broadly describes the eastward propagation

Figure 3. Time series of the five climate mode indices used as predictor variables in this study. Note that OLR is used as a proxy index for the MJO and that DMI is plotted using a different vertical scale.

of a convection cell that forms off the east cost of Africa and dissipates in the Pacific Ocean (Madden & Julian, 1972). The MJO is the dominant mode of intraseasonal variability in the tropics (Madden & Julian, 1994) and has been shown to increase or decrease the probability of extreme rain events by over 20% in MSEA depending on its phase (Xavier et al., 2014). The most common MJO index is described by the two primary empirical orthogonal functions (EOFs) resulting from a number of climate variables (Wheeler & Hendon, 2004). However, this index is poorly suited for use in a regression framework, as it would require a main term for both EOFs and their interaction to properly capture the phase of the MJO. This introduces multiple coefficient estimates for a single physical phenomenon, which makes it harder to model and hinders model interpretability.

Instead of using these EOFs, we use outgoing longwave radiation (OLR) anomalies to approximate the variability described by the MJO. OLR is a metric that describes how much energy is leaving the atmosphere and is one climate variable used in Wheeler and Hendon (2004) to produce the EOF index. Low OLR values indicate the presence of clouds, and hence a higher likelihood of rainfall (Birch et al., 2016). While not perfect, we believe OLR to be a decent approximation of the variability described by the MJO. Dias et al. (2017) shows that the MJO can be characterized by the variance in convection, and in Figure 3, we show that the frequency of the OLR signal captures the 30 – 90 day oscillatory movement of the MJO convection cell. This OLR proxy is better suited for a regression analysis despite losing some of the information contained in the EOF index from Wheeler and Hendon (2004).

We aggregate OLR values over the same spatial region that defines the MSEA region shown in Figure 1, and we create anomalies in the same manner as the CO anomalies described in Section 2.1. We demonstrate the benefit of including the OLR proxy in Section 6.1.

Figure 3 shows the weekly time series for each climate mode index used as a predictor variable in this study. Some of the indices have both high and low frequency components. This is most obvious in the SAM and OLR. We believe that the high frequency component of the OLR captures the oscillatory movement of the convection cell described by the MJO because both have a period of around 30 to 90 days. The climate mode index data used in this study are publicly available. The source of each index (or proxy index in the case of the MJO) is listed in Table 1.

Table 1. Climate mode indices used in this study with citations for their sources. Note that we use OLR as a proxy index for the MJO.

Climate Mode	Metric Used in Model	Source
ENSO	Niño 3.4	NOAA OOPC (2021)
IOD	Dipole Mode Index (DMI)	NOAA OOPC (2021)
TSA	Tropical South Atlantic (TSA)	NOAA OOPC (2021)
AAO	Southern Annular Mode (SAM)	NOAA CPC (2021)
MJO	Outgoing Longwave Radiation (OLR)	NOAA PSL (2021)

Note that there are other important modes of variability in the MSEA region that we do not include in our model, such as monsoons, wave phenomenon, diurnal patters, and tropical cyclones (Reid et al., 2012). These factors are excluded here because we solely aim to examine the connections between climate mode indices and atmospheric CO (as a proxy for fire intensity) in a higher level of detail than previous work, rather than build a comprehensive forecasting tool for the region.

3 Multiple Linear Regression Model

We use lagged multiple linear regression to model the relationship between CO anomalies and climate mode indices. We include first order interaction terms to capture the interconnected nature of the global climate system. Buchholz et al. (2018) found that these interaction terms were highly significant in explaining CO variability. Unlike the models in Buchholz et al. (2018), we also include squared terms to capture potential non-linear relationships between the mean CO response and the climate mode indices. For a given region, we assume that

$$CO(t) = \mu + \sum_k a_k \, \chi_k(t - \tau_k) + \sum_{i,j} b_{ij} \, \chi_i(t - \tau_i) \, \chi_j(t - \tau_j) + \sum_l c_l \, \chi_l(t - \tau_l)^2 + \epsilon(t), \quad (1)$$

where $CO(t)$ is the CO anomaly at time t, μ is a constant mean offset, a_k, b_{ij}, and c_l are coefficients, χ are the climate indices, τ is the lag value for each index in weeks, $\epsilon(t)$ is a random error component, and k, i, j, and l iterate over the number of climate indices used in the analysis. Note that we standardize the climate indices, χ, before fitting the model so that coefficient estimates can be directly compared. We consider lags between one and 52 weeks for each index. We also enforce strong hierarchy, meaning that any covariate that appears in an interaction or squared term must also appear as a main effect. Strong hierarchy has long been recommended for models with interactions, as it helps avoid misinterpretation of the included covariates (Nelder, 1977). See the Supporting Information file for more details on strong hierarchy.

Although the high frequency variability present in the weekly climate index data has important near-term effects, we do not expect it to have a large impact on the amount, type, and dryness of available fuel far into the future. This is because we believe that short anomalies do not last long enough to drastically alter large scale fuel reserves. Therefore, we want covariates with longer lags to capture progressively lower frequency components of the climate indices.

To accomplish this, we apply more smoothing to the climate mode indices as the length of their lag in the statistical model increases. In brief, we do not smooth indices for lags below four weeks to capture as much high frequency signal as possible in these short term relationships. For lags between four and 52 weeks, we use Gaussian kernels to linearly increase the amount of smoothing applied to the indices. More information on our smoothing scheme can be found in the Supporting Information file.

4 Variable Selection and Model Fitting

We consider 52 lags of each climate mode index, quadratic terms, and all pairwise interactions, which results in far more covariates than observations. In this regime, there is not a unique least squares solution, so another model fitting method is needed to compute coefficient estimates. Furthermore, we want to perform variable and lag selection to obtain human-interpretable models. Buchholz et al. (2018) broke this process up into two parts. First, they iterated through all possible lag combinations. At a given combination of lag values, stepwise selection was used for variable selection. This resulted in a list of optimally performing models, with one model for each combination of lag values. Adjusted R^2 was then used to select a single model from this list. By iterating through the lag values in this manner, Buchholz et al. (2018) was able to use stepwise selection without large computational resources. However, this strategy allowed for only a single lag of each index in the models.

To capture more complex relationships involving multiple lags of a given index, we instead consider all possible lags for each index simultaneously. This makes the search space too large for stepwise selection, so we instead employ regularization for both variable and lag selection. In the linear regression setting, regularization is a method of computing coefficient estimates that balances model fit and the overall magnitude of the coefficients with the goal of finding models that generalize well to new data. Furthermore, regularization is well suited for problems with more covariates than observations, making it feasible to consider all lag values for each index simultaneously.

We use a flexible regularization penalty called the Minimax Concave Penalty (MCP) (Zhang, 2010). Similar to the Least Absolute Shrinkage and Selection Operator (LASSO) penalty (Tibshirani, 1996), the MCP shrinks insignificant coefficient estimates to exactly zero, which leads to interpretable models with relatively few terms. Additionally, the MCP results in less biased estimates for the remaining non-zero coefficients by allowing for larger coefficients on the significant terms (Zhang, 2010). We found that using the MCP instead of the LASSO improved model performance. The MCP introduces a second parameter, η, that controls the MCP penalty in addition to the tuning parameter, λ, which is present in all regularization methods. The λ parameter balances how well the model fits to data and the overall magnitude of the coefficients (with a smaller overall magnitude leading to models with less terms). Compared to the LASSO, the MCP relaxes as the coefficients get larger and plateaus after they reach a certain magnitude. The η parameter controls when this plateau occurs, with smaller η values enabling larger coefficient estimates on the significant terms. Optimal λ and η values need to be learned from data.

To select parameter values, we perform a simple grid search over a range of η and λ values. We use the MCP to fit a model at each combination of η and λ values (implemented in `R` via the `RAMP` package from Hao et al. (2018)). We then choose between the resulting models via the Extended Bayesian Information Criterion (EBIC). The EBIC applies a much stronger penalty to large models (i.e., models with many selected terms) than other information criteria through a third parameter, γ, which is defined on the range

$[0, 1]$. When $\gamma = 0$, the EBIC is identical to the Bayesian Information Criterion (BIC), but when $\gamma = 1$, the EBIC is much harsher than the BIC. This is well suited for applications in which the number of possible covariates is large, but the optimal model might in fact be quite small. Since the number of potential covariates in this application is vast (recall that each lag value represents a different covariate), we use the EBIC rather than the BIC to select the final model. After finalizing the model terms in this manner, we refit their coefficient estimates via maximum likelihood.

More details on regularization, the MCP, the EBIC, and how we select parameter values can be found in the Supporting Information file. In the remaining sections, we discuss how this modeling framework and the choice of γ can be used to address our two goals of model interpretability and predictive performance.

5 Interpreting Fitted Models

Here we examine the physical implications of the models fit using the procedure described in Section 4. We focus on connections between climate and CO in MSEA through an analysis of selected indices and lag values.

5.1 A Framework for Identifying Optimally Performing Models at Various Complexities

We can create a list of "optimally performing" models at decreasing complexities (i.e., number of terms) by increasing the EBIC parameter, γ, on the range $[0, 1]$, as larger γ values increase the penalty on large models. Optimal here refers to the fact that these models are the result of a grid search over the other two free parameters, λ and η. For MSEA, this procedure results in the models listed in Figure 4. The color of each box corresponds to the γ value that was used to generate the model contained within it. Note that multiple γ values can produce the same model. Within each box, the name of the index and the corresponding lag is listed (in the format "`name_lag`"), along with the coefficient estimates and standard errors.

Moving from left to right in Figure 4, we see that the models decrease in size (from 17 terms to nine), while their performance drops only slightly (from adjusted R^2 of 0.68 to 0.60). By examining the terms that remain in the model as it becomes more parsimonious, we can determine which indices and lags are most influential in explaining variability in the response.

For MSEA, we can see that the Niño 3.4 index lagged at four weeks remains in the model with a positive coefficient estimate. This makes sense, as ENSO is a major climate driver in the tropics, with positive anomalies resulting in warmer, drier conditions (Nur'utami & Hidayat, 2016). The lag of four weeks indicates that it takes about four weeks for the effect of a Niño 3.4 anomaly to impact CO anomalies. Additionally, the Niño 3.4 lag of four weeks appears as a squared term in the most parsimonious model, indicating that there is a nonlinear relationship between Niño 3.4 and CO. This is confirmed by examining the residuals of a model fit to solely the Niño 3.4 lag of four weeks (not shown).

The selected DMI lags also suggest an interesting relationship. Note that positive DMI anomalies are associated with reduced rainfall in parts of MSEA, while negative DMI anomalies are associated with increased rainfall (Nur'utami & Hidayat, 2016). A DMI lag of 12 weeks remains in the model as it become more parsimonious, as well as a shorter lag that switches from one to four weeks between the smallest two models. The coefficient on the longer lag is negative, while the coefficient on the shorter lag is positive. The coefficient on the shorter lag implies that reduced rainfall (i.e., positive DMI anomalies) results in more CO on average, and vise versa. This is likely the result of an

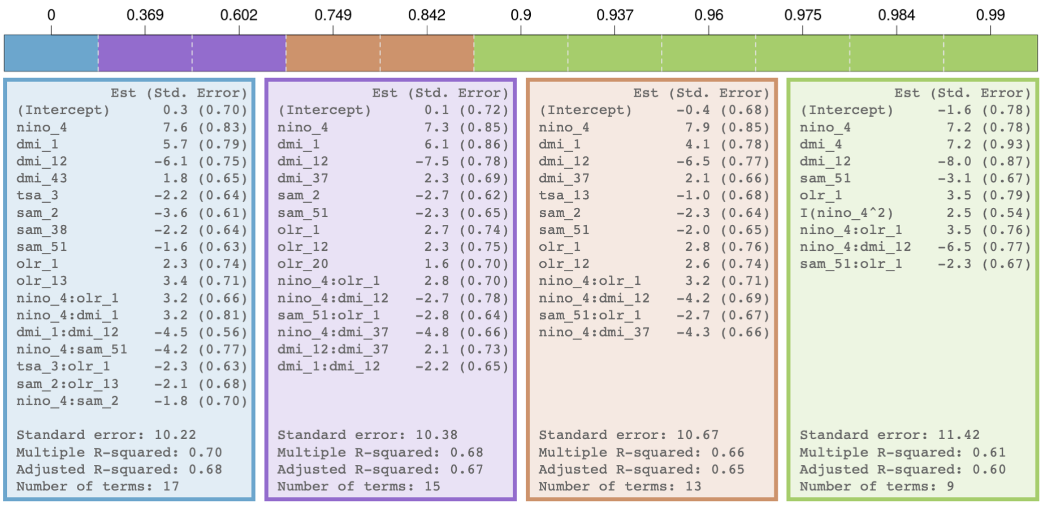

Figure 4. Optimal models for the MSEA region for a logarithmic sequence of γ values. Note that multiple γ values can produce the same model. The color of each box corresponds to the γ value that was used to generate the model contained within it. The model terms are listed in the format "**name_lag**," where lags are in weeks. Interaction terms are listed in the format "**name1_lag1:name2_lag2**." Coefficient estimates and standard errors are listed for each term, and summary statistics are listed below each model. Note that "nino" refers to the Niño 3.4 index.

intuitive relationship: reduced rainfall leads to drier conditions that are more prone to burning (and hence more CO). Similar to the ENSO relationship, these dry conditions take one to four weeks to impact CO. The coefficient on the longer lag, however, implies the opposite: reduced rainfall (i.e., positive DMI anomalies) results in less CO on average, and conversely, increased rainfall results in more CO on average. This could be because rainfall leads to vegetation growth, which ultimately provides more fuel for fires. The length of this lag is longer, implying that it takes around 12 weeks for the increased vegetation growth to impact CO concentrations.

The effect of these two DMI lags is compounding. That is, more vegetation from DMI-driven rainfall at a 12 week lead time results in more fuel for burning when a subsequent positive DMI anomaly drives drier conditions. This is supported by the negative coefficient on the interaction between the DMI lag of 12 weeks and one week present in the largest model in Figure 4. Because the coefficient is negative, there is less CO on average when the DMI has the same phase (i.e., either a positive or negative anomaly) at both a 12 and one week lag.

An OLR term lagged at one week remains in the MSEA model as it becomes more parsimonious with a positive coefficient estimate. This again makes sense, as positive OLR anomalies are associated with less cloud cover and hence less rain. The one week lag suggests that an OLR-driven decrease in rain leads to more CO in the short term, likely as a result of increased burning. The TSA index, on the other hand, is only included in the largest model. This could be because the TSA describes sea surface temperatures in the southern Atlantic Ocean, which is very far from MSEA. Therefore, it makes sense that the TSA is less important than the other indices in explaining CO variability in MSEA, as the other indices are based on aspects of the global climate system located closer to MSEA.

Finally, two Niño 3.4 interaction terms remain in the model as it becomes more parsimonious. One interaction is with the OLR at a one week lag and the other is with the DMI at a 12 week lag. The sign of these interaction terms is the same as the non-Niño 3.4 component. This indicates that the effects of these indices are amplified when they are in phase, a result that has been previously identified in the literature (Cleverly et al., 2016; Nur'utami & Hidayat, 2016). Note that studies like Islam et al. (2018) have shown that there is increased fire potential when Niño 3.4 and DMI are both positive. Our model agrees with this finding (see the Niño - DMI interaction in the largest model), but also expands on this finding by showing that Niño also amplifies the effect of DMI at longer lead times (see the Niño - DMI interaciton in the smallest model). Our results are also consistent with Reid et al. (2012), who show that an increase in fire activity occurs during the ENSO warm phase and positive IOD phase. Reid et al. (2012) also found evidence of a relationship between ENSO and IOD. We expand on this work by specifying the Niño 3.4 and DMI lead times that most significantly influence CO and by showing how the Niño - DMI interaction changes at different lead times.

These findings largely agree and expand upon the results in Buchholz et al. (2018). For MSEA, Buchholz et al. (2018) found that a Niño 3.4 lag of one month, DMI lag of eight months, TSA lag of five months, and SAM lag of one month were important predictors. The largest model presented in this study contains a Niño 3.4 lag of four weeks, DMI lag of 43 weeks, TSA lag of three weeks, and SAM lag of two weeks. All but the TSA term (which we will show to be less important for MSEA in Section 5.2) agree closely on their selected lag. However, the models we present here are capable of including multiple lags of a single index, which expands on the work in Buchholz et al. (2018) and highlights more complex relationships between climate and CO.

5.2 Assessing Stability of Selected Model Terms

While the scientific conclusions drawn in the previous section seem to agree with and expand upon current literature, we want to ensure that the selected covariates are in fact meaningful. That is, we want to avoid over-interpreting the role of covariates if slight changes in data result in drastically different models, as these models would not be capturing a meaningful physically-based relationship but would rather be artifacts of the specific training data.

Therefore, we perform one-year-out resampling to assess the stability of selected covariates. We perform the resampling on the largest model from Figure 4 because it contains most of the terms present in the smaller models. Specifically, we perform the following resampling procedure. We first iterate through the years present in the data. For each year, we create a testing set containing all data falling within that year and a training set containing the remainder of the data. We then train two models using only data from the training set. We force the first model (called the "constant structure model") to retain the same covariates as the model trained on all of the data but allow for different coefficient estimates. We let the second model (called the "varying structure model") to completely change based on the particular training set, meaning that it can have different covariates and coefficient estimates than the model trained on all of the data. We then test these two models on the corresponding test set and compute the root mean square error (RMSE) for both.

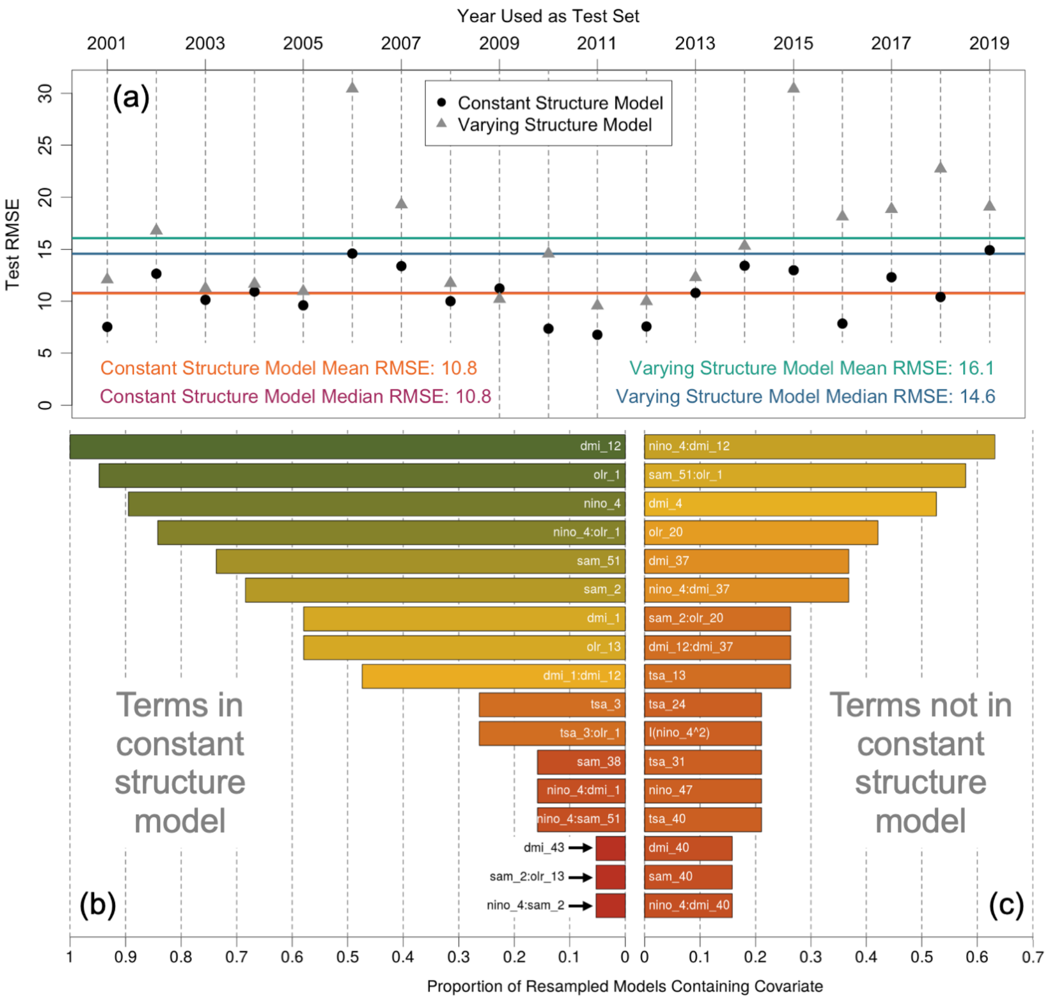

Figure 5. Results from the one-year-out resampling. Constant structure model refers to the model forced to retain the structure of the model trained on all of the data, but with refit coefficient estimates. New model refers to the model allowed to completely change according to the particular training set. (a) shows the out-of-sample prediction error for each testing set. The year on the horizontal axis indicates which year was used to test the models. The constant structure model almost always outperforms the varying structure model. (b) shows the frequency with which constant structure model terms appear in the varying structure models. Similarly (c) shows the frequency with which terms not present in the constant structure model appear in the varying structure model. The most significant covariates from Figure 4 appear in many of the retrained models. The color in (b) and (c) corresponds to the proportion on the horizontal axis and is included for visual clarity. Note that "nino" refers to the Niño 3.4 index.

Figure 5 shows the results of this resampling and is divided into three sections. Figure 5(a) shows the out-of-sample prediction error (RMSE) from both models for each training set. The year on the horizontal axis corresponds to the year reserved for the testing set. The RMSE of the constant structure model tends to perform as well or better than the varying structure model. This provides justification for using the form of the model trained on all data as the representative model for MSEA and further interpreting its covariates, as the relationships captured by this model do a better job at explaining the data than those in the varying structure models. Note that the RMSE of the varying structure model is largest when 2006 and 2015 are left out of the training set. These years contained some of the largest CO anomalies of the 19 year time series (see Figure 2). This indicates that: 1) these extreme fire years are important in driving the form of the model trained on all data, and 2) this framework should be used with caution in a forecasting setting.

Figure 5(b) and Figure 5(c) show how often certain terms appear in the varying structure models (that is, the models allowed to completely change according to the new training data). This gives some indication of the stability of the various model terms. If a term is present in many of the retrained models, then the modeling framework is likely picking up a physically-based relationship. Terms that are absent from many of the retrained models are more likely artifacts of the specific training set, rather than a true physical relationship.

Figure 5(b) shows how often the constant structure model terms reappear in the varying structure models. Notably, the terms present in the most parsimonious model from Figure 4 are most likely to appear in the retrained models. This indicates that these terms are explaining the most stable aspect of the physical relationship. Other terms, such as the 43 week DMI lag, rarely appear in the retrained models. This indicates that less consideration should be given to these terms when attempting to explain the physical relationship between climate and CO.

Figure 5(c) shows how often terms not present in the constant structure model appear in the retrained models. Note the different scales on the horizontal axis between subfigures 5(b) and 5(c). In Figure 5(c) we see that a selection of terms not in the constant structure model appear relatively frequently in the retrained models. Recall that when moving from the second smallest to the smallest model in Figure 4, the shorter DMI lag switches from one week to four weeks. In Figures 5(b) and (c), we see that both the one and four week DMI lags show up in about half of the retrained models. This indicates that these terms are interchangeable, and determining which is included likely depends on the other selected covariates.

Figures 5(b) and (c) further confirm that the terms present in the most parsimonious model for the region (see Figure 4) are capturing meaningful signal and are not simply artifacts of the specific training set. This is because these terms remain in a large majority of the retrained models, each of which is trained on a different subsample of the data. Furthermore, Figure 5(c) illustrates that the interaction between Niño 3.4 lagged at four weeks and DMI lagged at 12 weeks, although not present in the constant structure model, is still a significant interaction in explaining CO variability in MSEA. This also holds for the interaction between SAM lagged at 51 weeks and OLR lagged at one week. The terms that are included less often in the retrained models are likely more data dependent and help the model capture subtleties in the response. As a result, it is more likely that these terms would change with small changes in the data. An example is the TSA term lagged at three weeks present in the constant structure model. This term appears in less than 30% of the retrained models, which confirms the analysis in Section 5.1 that finds that TSA is less important in explaining CO variability in MSEA.

The stability analysis presented here provides further justification for assigning scientific weight to selected model terms, as it shows that certain stable terms are not simply artifacts of the particular training set used to fit the model. In particular, we confirm that a number of terms from the smallest model presented in Figure 4 are very stable: DMI lagged at 12 weeks, OLR lagged at one week, Niño 3.4 lagged at four weeks, a short DMI lag (of either one or four weeks depending on the remaining model terms), SAM lagged at 51 weeks, the interaction between Niño 3.4 lagged at four weeks and OLR lagged at one week, and the interaction between Niño 3.4 lagged at four weeks and DMI lagged at 12 weeks. This provides further evidence that these terms specify the most significant relationships between climate and atmospheric CO in MSEA.

6 Assessing Model Predictions

We now turn our attention to the predictive performance of selected models. We again focus on the largest model from Figure 4, as this model has the best predictive capabilities. Strong predictive performance indicates that there is indeed a connection between climate mode indices and CO variability and that our model is able to capture part of this connection. Therefore, strong predictive performance gives additional weight to the scientific interpretation of the selected model terms. Note that the performance metrics discussed in this section (e.g., percent of variability explained) are not meant to be an assessment of our model's forecasting ability, but rather an assessment of how well we can explain the response (CO variability) using only our predictors (climate mode indices).

6.1 Model Predictions with No Minimum-Lag-Threshold

In this subsection we impose no requirements on the minimum lag value allowed in the models, meaning that we allow lags of one to 52 weeks as in Figure 4. In Figures 6 and 7 we demonstrate the predictive capabilities of our model and highlight two interesting results.

Figure 6 shows weekly observations and predictions from two model variants. Note that these predictions are in-sample, meaning that they are predictions of the observations used to train the model. The top plot of Figure 6(a) shows predictions from a model completely refit to a data set excluding the OLR, and the bottom plot shows predictions from the full model (i.e., the model presented in Figure 4). We can see that including the OLR results in a slight decrease in RMSE and increase in both R^2 and adjusted R^2. Note that adjusted R^2 is a better metric for comparing the two models, as it accounts for the number of terms in each model. Similar to R^2, higher adjusted R^2 values indicate a better fit. Furthermore, in Figure 6(b) and (c), we highlight two of the most anomalous years, which shows that the OLR helps capture the extreme CO anomalies. This makes sense for 2015 in particular, as the MJO and our OLR proxy experienced an extreme anomaly during this year.

Figure 7 shows month-averaged observations and predictions from two different model variants. The top plot of Figure 7(a) shows predictions from a month-based model. To create this model, we took month-averages of the predictor variables and then trained the model on only these month-averaged covariates using the framework presented in Section 4. We imposed no restrictions on the terms included in this model, as we do not want to introduce information from the weekly data that would not otherwise be available in the monthly data. The bottom plot shows month-averaged predictions from the model trained on weekly data (i.e., the model shown in Figure 4). We see a noticeable increase in model performance when using the weekly data, suggesting that the weekly data is able to capture meaningful signal beyond the month-averages. This is an interesting result, as it suggests that the higher frequency signals present in the climate indices are in fact meaningful signal and not simply noise. This is perhaps most important for OLR (the proxy for localized MJO), which has a higher frequency component than the other included climate indices. This increase in performance can be seen clearly during the 2015 CO anomaly.

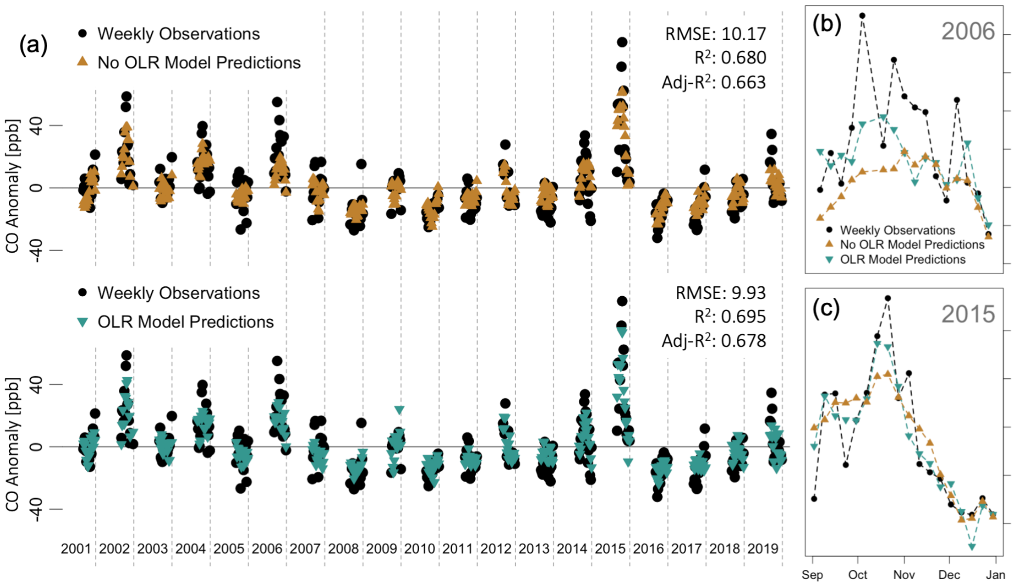

Figure 6. In-sample predictions from two model variants. In (a), the top plot shows predictions from the optimal model without the OLR, and the bottom plot shows predictions from the optimal model with the OLR. Adding the OLR appears to increase predictive performance during the extreme CO anomalies shown in (b) and (c).

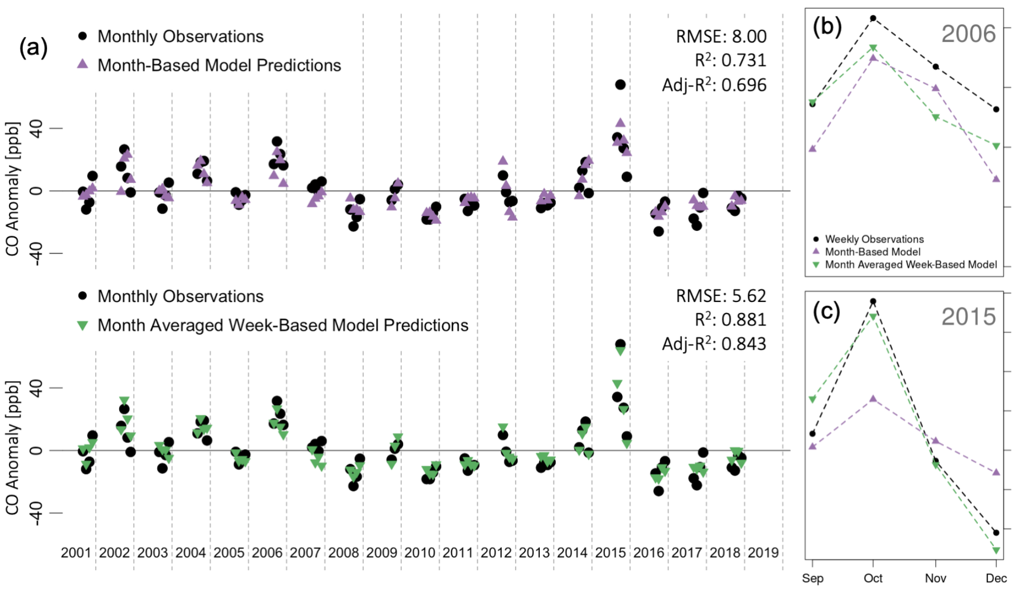

Figure 7. In-sample predictions from two additional model variants. In (a), the top plot shows predictions from a model trained on month-averaged covariates, and the bottom plot shows month-averaged predictions from a model trained on week-averaged covariates. The increase in model performance indicates that there is meaningful signal in the higher frequency climate index data, which is clearly seen in the anomalous years shown in (b) and (c).

Note that the predictions from these models are an improvement over the models in Buchholz et al. (2018). When using week-averaged data to train the model, we are able to explain 88% of the variability in the month-averaged CO observations. The model in Buchholz et al. (2018) explains 75% of the month-averaged CO. This increase in predictive performance is likely a result of: 1) the ability to include multiple lags of a single climate mode index, 2) the additional signal contained in the week-averaged data, and 3) the inclusion of the OLR proxy index.

6.2 Increasing Minimum-Lag-Threshold

The predictions shown in Subsection 6.1 are useful for demonstrating model performance and the comparative benefit of using the OLR and week-averaged data. However, these models include an OLR term lagged at one week (see Figure 4), which means that they can only be used to forecast one week ahead. In this section, we explore the capabilities of our model in a more practical forecasting environment. Note that we are not attempting to outperform or even match state-of-the-art forecasting tools that utilize modes of variability beyond just climate modes. Instead, we are interested in exploring the forecasting performance of our statistical model trained solely on climate mode indices, which will potentially help forecasters attempting to build more sophisticated tools.

To increase the prediction horizon, we implement a minimum-lag-threshold that only allows lags greater than the threshold value to be included in the model. Because increasing this threshold reduces the number of possible covariates, we also extend the maximum lag value as the minimum-lag-threshold is increased. Specifically, we consider lags between the minimum-lag-threshold and 52 weeks plus this threshold. This ensures that all models are based on one year of climate data, making it easier to compare their predictive performance.

Figure 8 shows a selection of model performance metrics as this minimum-lag-threshold is increased. We focus on the largest model generated from the range of EBIC γ values, as this model has the best predictive performance. The top plot in Figure 8 shows the number of terms in the selected model for each minimum-lag-threshold. The second plot shows the adjusted R^2 value of the selected models. As expected, the model performance drops off as the minimum lag is increased. However, this decline is not very rapid. That is, models with a high minimum-lag-threshold still explain a large percent of the variability in atmospheric CO anomalies. This is promising, as it means that predictions can be made farther in advance without losing too much predictive performance. The third plot shows another performance metric: the average out-of-sample prediction error from one-year-out resampling. Here we successively leave one year out, train the model on the remaining data, and test it on the left out year. The average RMSE is then taken for each different training and testing set pair and plotted as a function of minimum-lag-threshold. We see that performance falls off, although gradually.

We think that the gradual nature of the decline in model performance is a result of the climate indices exhibiting high auto-correlation (not shown). Since many of the short lags are highly correlated to longer lags of the same index, we think that these longer lags are able to explain much of the same CO variability when the shorter lags are excluded. This is again promising, as it means that predictions can be made decently far in advance (on the order of a half year) without dramatically compromising performance.

To further visualize model performance at increasingly large minimum-lag-thresholds, we consider predictions for the 2015 CO event in MSEA. Figure 9 shows predictions from the models corresponding to the minimum-lag-thresholds from Figure 8. The predictions largely capture the structure of the CO observations for minimum-lag-thresholds below 25 weeks (about six months). After this point, the predictions begin to flatten out (i.e., not capture the extremes in the response) and the predicted spike starts earlier in the

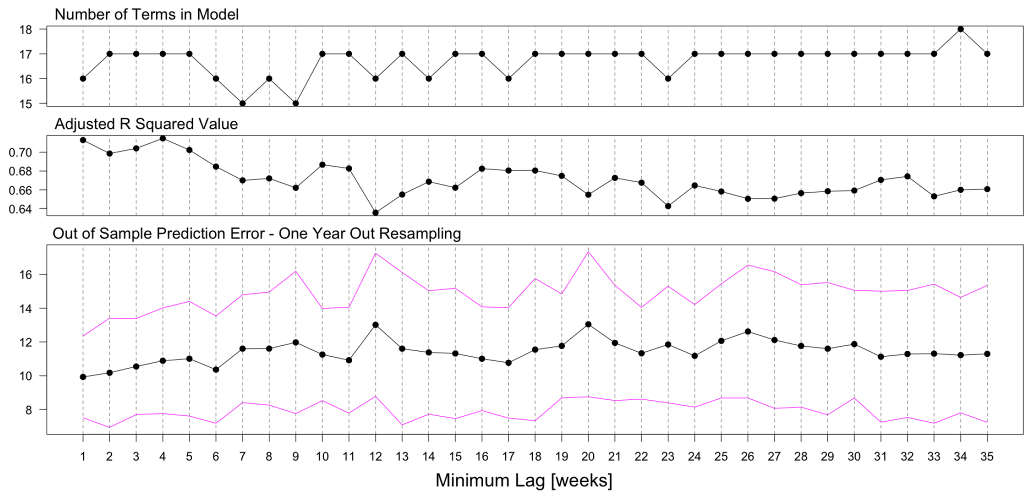

Figure 8. Model performance for MSEA at increasing minimum-lag-thresholds. Top plot shows the number of terms in the selected model. Middle plot shows the adjusted R^2 value of the selected model. Bottom plot shows an average out-of-sample prediction error for each model with magenta lines showing ± one standard deviation. Here we iteratively leave one year out, train the model on the remaining data, and test it on the left out year. Plotted is the average RMSE with ± one standard deviation lines in magenta from this procedure as a function of minimum lag. We can see that model performance drops off with an increasing minimum-lag-threshold, although at a fairly gradual pace.

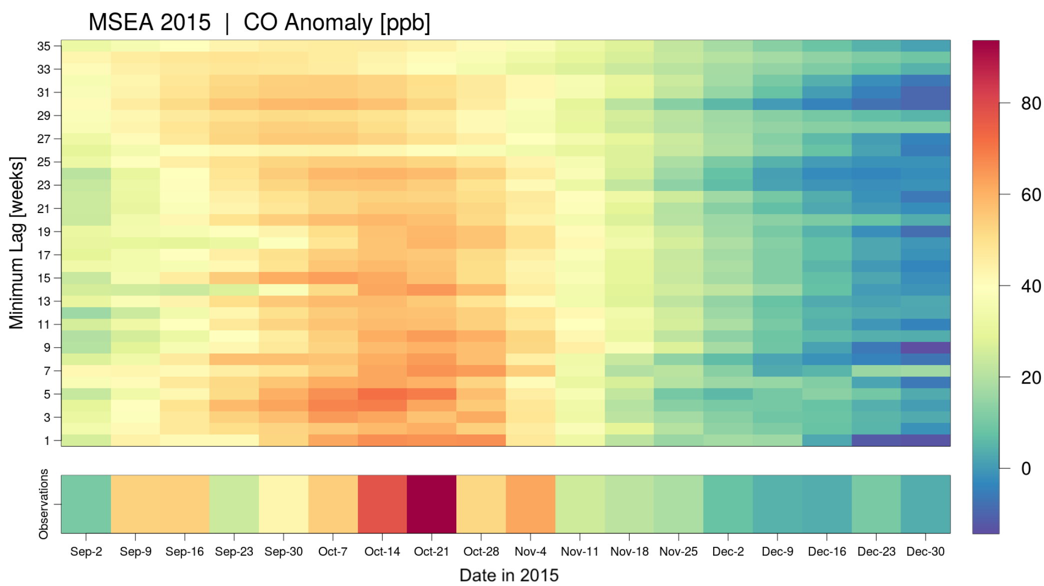

Figure 9. Predictions of the 2015 CO anomalies [ppb] in MSEA for a range of minimum-lag-thresholds. Color represents CO anomalies, and the horizontal axis represents time. MOPITT observations are shown as a horizontal bar along the bottom of the figure. The remaining vertical axis corresponds to the minimum-lag-threshold used to fit the model, and hence each row of the figure contains predictions from a different model. The minimum-lag-threshold can be interpreted as the prediction horizon of the model. We see that the general structure of the observed CO anomalies is preserved for minimum lags under 25 weeks (about half a year).

year (i.e., in early September instead of early October). This result largely agrees with Shawki et al. (2017), who found that a drought metric could be reasonably predicted 180 days (about 25 weeks) in advance. However, unlike Shawki et al. (2017), our predictions rely solely on past climate mode index anomalies, rather than forecasts from a global climate model.

7 Summary

We build on previous work aimed at explaining the relationship between climate and atmospheric CO variability. Atmospheric CO is a useful proxy for fire intensity, as fires are the main source of CO variability in the Southern Hemisphere and CO is remotely sensed on a global scale.

Our proposed regularization framework highlights a variety of optimally performing models at decreasing complexities, isolating the most important indices and lag values as the models become more parsimonious. For MSEA, we identify the Niño 3.4 index lagged at four weeks as a primary driver of atmospheric CO. Other important climate indices are the DMI and OLR (as a proxy for the MJO). We further identify that Niño 3.4 interactions with the OLR and DMI are significant predictors, suggesting that the effect of these indices is amplified when they are in phase. Finally, we show that including multiple lags of the DMI is important for explaining CO variability in MSEA. While these results broadly agree with current literature, we go beyond the usual treatment of climate mode indices on a seasonal time scale by identifying the specific weekly lead times for each index that have the most influence on CO variability.

We also perform a resampling-based sensitivity analysis to quantify the robustness of the model fit to all data. We find that the model forced to retain the covariates from the model trained on all data performs as well or better than the model allowed to completely change based on the training set. This provides justification for using the models from Figure 4 as the representative models for MSEA. Additionally, we determine which covariates are most likely to remain in the model when trained on slightly different data, finding that the terms in the most parsimonious model from Figure 4 are also the most robust. This justifies assigning scientific weight to the selection of these terms, as it suggests that they are capturing a physically-based relationship and are not simply artifacts of the specific training set used.

We show that our model for the MSEA region can explain around 70% of the variability in the weekly CO anomalies solely using climate indices as predictor variables. We further use model predictions to highlight the importance of the OLR (as a proxy for the MJO) in overall model performance and in explaining the most extreme CO anomalies. Similarly, we show that month-averaged predictions from a model trained on week-averaged data outperform predictions from a model trained on month-averaged data. This suggests that there is meaningful signal in the week-averaged data and justifies its use over month-averaged data. Note that the predictions from these models are an improvement over those in Buchholz et al. (2018), as they explain 88% of the variability in month-averaged CO observations compared to 75%.

Finally, we perform a minimum-lag-threshold study to assess the performance of our model in a forecasting setting. We find that models for MSEA are still able to explain around 65% of the weekly atmospheric CO variability when forced to only use lags greater than 35 weeks. While we do not attempt to outperform or even match state-of-the-art forecasting tools, we believe that this information is useful to forecasters hoping to maximize the information they glean from climate modes when developing more sophisticated tools.

Overall, we believe that our modeling framework quantifies the relationship between climate mode indices and atmospheric CO (as a proxy for fire intensity and as a measure of air quality) at a level of complexity not previously studied. We do this by utilizing climate mode indices on a weekly timescale, accommodating multiple lead times of each climate mode, and including complex interactions between climate mode indices at a weekly timescale. We believe that this work will be useful for scientists studying modes of variability in MSEA.

Acknowledgments

The NCAR MOPITT project is supported by the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Program. The MOPITT team also acknowledges support from the Canadian Space Agency (CSA), the Natural Sciences and Engineering Research Council (NSERC) and Environment Canada, and the contributions of COMDEV and ABB BOMEM. The National Center for Atmospheric Research (NCAR) is sponsored by the National Science Foundation.

Open Research

MOPITT carbon monoxide data are publicly available through NASA. See `https://doi.org/10.5067/TERRA/MOPITT/MOP02J_L2.008`. Climate index data are produced and maintained by NOAA. See `https://stateoftheocean.osmc.noaa.gov` and `http://www.cpc.ncep.noaa.gov`. Only a subset of the MOPITT V8 Level 2 carbon monoxide data is used in this work. The processed carbon monoxide and climate mode index data used in this work are publicly available through NCAR. See `https://doi.org/10.5065/s6rv-rc57` (Buchholz, Worden, Ahamad, et al., 2021). The `R` code used to implement the model fitting framework proposed in this work can be accessed through GitHub. See `https://github.com/wsdaniels/COmodeling`.

References

Alencar, A., Asner, G. P., Knapp, D., & Zarin, D. (2011). Temporal variability of forest fires in eastern Amazonia. *Ecological Applications*, *21*(7), 2397–2412. doi: 10.1890/10-1168.1

Andela, N., & van der Werf, G. R. (2014). Recent trends in African fires driven by cropland expansion and El Niño to la Niña transition. *Nature Climate Change*, *4*(9), 791–795. doi: 10.1038/nclimate2313

Andreoli, R. V., & Kayano, M. T. (2006). Tropical Pacific and South Atlantic effects on rainfall variability over Northeast Brazil. *International Journal of Climatology*, *26*(13), 1895–1912. doi: 10.1002/joc.1341

Bamston, A. G., Chelliah, M., & Goldenberg, S. B. (1997). Documentation of a highly ENSO-related sst region in the equatorial pacific: Research note. *Atmosphere - Ocean*, *35*(3), 367–383. doi: 10.1080/07055900.1997.9649597

Bien, J., Taylor, J., & Tibshirani, R. (2013). A lasso for hierarchical interactions. *The Annals of Statistics*, *41*(3), 1111–1141. doi: 10.1214/13-AOS1096

Birch, C. E., Webster, S., Peatman, S. C., Parker, D. J., Matthews, A. J., Li, Y., & Hassim, M. E. E. (2016). Scale interactions between the MJO and the Western Maritime Continent. *Journal of Climate*, *29*, 2471-2492. doi: 10.1175/JCLI-D-15-0557.1

Bloom, A. A., Worden, J., Jiang, Z., Worden, H., Kurosu, T., Frankenberg, C., & Schimel, D. (2015). Remote-sensing constraints on South America fire traits by Bayesian fusion of atmospheric and surface data. *Geophysical Research Letters*, *42*(4), 1268–1274. doi: 10.1002/2014GL062584

Breheny, P., & Huang, J. (2011). Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *The Annals of Applied Statistics*, *5*(1), 232–253. doi: 10.1214/10-AOAS388

Buchholz, R. R., Hammerling, D., Worden, H. M., Deeter, M. N., Emmons, L. K., Edwards, D. P., & Monks, S. A. (2018). Links between carbon monoxide and climate indices for the Southern Hemisphere and tropical fire regions. *Journal of Geophysical Research: Atmospheres*, *123*(17), 9786–9800. doi: 10.1029/2018JD028438

Buchholz, R. R., Worden, H. M., Ahamad, F., Daniels, W. S., & Hammerling, D. M. (2021). *Weekly carbon monoxide anomalies over maritime southeast asia and weekly climate indices.* (Version 1.0. UCAR/NCAR - GDEX) doi: 10.5065/s6rv-rc57

Buchholz, R. R., Worden, H. M., Park, M., Francis, G., Deeter, M. N., Edwards, D. P., ... Kulawik, S. S. (2021). Air pollution trends measured from Terra: CO and AOD over industrial, fire-prone, and background regions. *Remote Sensing of Environment*, *256*, 112275. doi: https://doi.org/10.1016/j.rse.2020.112275

Ceccato, P., Nengah Surati Jaya, I., Qian, J. H., Tippett, M. K., Robertson, A. W., & Someshwar, S. (2010). *Early warning and response to fires in Kalimantan, Indonesia* (Tech. Rep.). International Research Institute for Climate and Society.

Chen, Y., Morton, D. C., Andela, N., Giglio, L., & Randerson, J. T. (2016). How much global burned area can be forecast on seasonal time scales using sea surface temperatures? *Environmental Research Letters*, *11*(4), 45001. doi: 10.1088/1748-9326/11/4/045001

Chen, Y., Morton, D. C., Andela, N., van der Werf, G. R., Giglio, L., & Randerson, J. T. (2017). A pan-tropical cascade of fire driven by El Niño/Southern Oscillation. *Nature Climate Change 2017 7:12*, *7*, 906-911. doi: 10.1038/s41558-017-0014-8

Cleverly, J., Eamus, D., Luo, Q., Coupe, N. R., Kljun, N., Ma, X., ... Huete, A. (2016). The importance of interacting climate modes on Australia's contribution to global carbon cycle extremes. *Scientific Reports*, *6*(1), 1–10. doi: 10.1038/srep23113

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., ... Andreae, M. O. (2019). Radiance-based retrieval bias mitigation for the MOPITT instrument: The version 8 product. *Atmospheric Measurement Techniques*, *12*(8), 4561–4580. doi: 10.5194/amt-12-4561-2019

Deeter, M. N., Edwards, D. P., Gille, J. C., & Drummond, J. R. (2007). Sensitivity of MOPITT observations to carbon monoxide in the lower troposphere. *Journal of Geophysical Research*, *112*(D24), D24306. doi: 10.1029/2007JD008929

Deeter, M. N., Martínez-Alonso, S., Edwards, D. P., Emmons, L. K., Gille, J. C., Worden, H. M., ... Wofsy, S. C. (2014). The MOPITT Version 6 product: Algorithm enhancements and validation. *Atmospheric Measurement Techniques*, *7*(11), 3623–3632. doi: 10.5194/amt-7-3623-2014

Dias, J., Sakaeda, N., Kiladis, G. N., & Kikuchi, K. (2017). Influences of the MJO on the space-time organization of tropical convection. *Journal of Geophysical Research: Atmospheres*, *122*, 8012-8032. doi: 10.1002/2017JD026526

Drummond, J. R., Zou, J., Nichitiu, F., Kar, J., Deschambaut, R., & Hackett, J. (2010). A review of 9-year performance and operation of the MOPITT instrument. *Advances in Space Research*, *45*, 760-774. doi: 10.1016/J.ASR.2009.11.019

Edwards, D. P., Emmons, L. K., Gille, J. C., Chu, A., Attié, J.-L., Giglio, L., ... Drummond, J. R. (2006). Satellite-observed pollution from Southern Hemisphere biomass burning. *Journal of Geophysical Research*, *111*(D14), D14312. doi: 10.1029/2005JD006655

Edwards, D. P., Pétron, G., Novelli, P. C., Emmons, L. K., Gille, J. C., & Drummond, J. R. (2006). Southern Hemisphere carbon monoxide interannual variability observed by Terra/Measurement of Pollution in the Troposphere (MOPITT). *Journal of Geophysical Research*, *111*, D16303. doi:

10.1029/2006JD007079
Enfield, D. B., Mestas-Nuñez, A. M., Mayer, D. A., & Cid-Serrano, L. (1999). How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures? *Journal of Geophysical Research: Oceans*, *104*(C4), 7841–7848. doi: 10.1029/1998jc900109

Fonseca, M. G., Anderson, L. O., Arai, E., Shimabukuro, Y. E., Xaud, H. A. M., Xaud, M. R., … Aragão, L. E. O. C. (2017). Climatic and anthropogenic drivers of northern Amazon fires during the 2015–2016 El Niño event. *Ecological Applications*, *27*, 2514-2527. doi: 10.1002/EAP.1628

Fuller, D. O., & Murphy, K. (2006). The ENSO-fire dynamic in insular Southeast Asia. *Climatic Change*, *74*(4), 435–455. doi: 10.1007/s10584-006-0432-5

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, *217*, 72–85. doi: 10.1016/j.rse.2018.08.005

Giglio, L., Csiszar, I., & Justice, C. O. (2006). Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *Journal of Geophysical Research: Biogeosciences*, *111*, 2016. doi: 10.1029/2005JG000142

Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, *178*, 31–41. doi: 10.1016/j.rse.2016.02.054

Groot, W. J. d., Field, R. D., Brady, M. A., Roswintiarti, O., & Mohamad, M. (2006, Sep 30). Development of the indonesian and malaysian fire danger rating systems. *Mitigation and Adaptation Strategies for Global Change*, *12*(1), 165. Retrieved from `https://doi.org/10.1007/s11027-006-9043-8` doi: 10.1007/s11027-006-9043-8

Hao, N., Feng, Y., & Zhang, H. H. (2018). Model selection for high-dimensional quadratic regression via regularization. *Journal of the American Statistical Association*, *113*(522), 615–625. doi: 10.1080/01621459.2016.1264956

Holloway, T., Levy, H., & Kasibhatla, P. (2000). Global distribution of carbon monoxide. *Journal of Geophysical Research: Atmospheres*, *105*(D10), 12123–12147. doi: 10.1029/1999JD901173

Islam, M. A., Chan, A., Ashfold, M. J., Ooi, C. G., & Azari, M. (2018, 9). Effects of el-niño, indian ocean dipole, and madden-julian oscillation on surface air temperature and rainfall anomalies over southeast asia in 2015. *Atmosphere 2018, Vol. 9, Page 352*, *9*, 352. Retrieved from `https://www.mdpi.com/2073-4433/9/9/352/htmhttps://www.mdpi.com/2073-4433/9/9/352` doi: 10.3390/ATMOS9090352

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., … Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437 - 472. doi: 10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., … Fiorino, M. (2001). The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, *82*(2), 247–268.

Madden, R. A., & Julian, P. R. (1972). Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of Atmospheric Sciences*, *29*(6), 1109 - 1123. doi: 10.1175/1520-0469(1972)029⟨1109:DOGSCC⟩2.0.CO;2

Madden, R. A., & Julian, P. R. (1994). Observations of the 40–50-day tropical oscillation—a review. *Monthly Weather Review*, *122*(5), 814 - 837. doi: 10.1175/1520-0493(1994)122⟨0814:OOTDTO⟩2.0.CO;2

Mason, S. A., Hamlington, P. E., Hamlington, B. D., Jolly, W. M., & Hoffman, C. M. (2017). Effects of climate oscillations on wildland fire potential in the continental United States. *Geophysical Research Letters*, *44*, 7002-7010. doi: 10.1002/2017GL074111

N'Datchoh, E. T., Konaré, A., Diedhiou, A., Diawara, A., Quansah, E., & Assamoi, P. (2015). Effects of climate variability on savannah fire regimes in West Africa. *Earth System Dynamics*, *6*, 161-174. doi: 10.5194/ESD-6-161-2015

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F. F., Wakata, Y., Yamagata, T., & Zebiak, S. E. (1998). ENSO theory. *Journal of Geophysical Research: Oceans*, *103*(C7), 14261–14290. doi: 10.1029/97jc03424

Nelder, J. A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society. Series A (General)*, *140*(1), 48. doi: 10.2307/2344517

NOAA CPC. (2021). *Climate prediction center - teleconnections: Antarctic oscillation.* `https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/aao/aao.shtml`.

NOAA OOPC. (2021). *Ocean observations panel for climate - state of the ocean climate.* `https://stateoftheocean.osmc.noaa.gov/`.

NOAA PSL. (2021). *Physical sciences laboratory - interpolated olr.* `https://psl.noaa.gov/data/gridded/data.interp_OLR.html`.

Nur'utami, M. N., & Hidayat, R. (2016). Influences of IOD and ENSO to Indonesian rainfall variability: Role of atmosphere-ocean interaction in the Indo-pacific sector. *Procedia Environmental Sciences*, *33*, 196–203. doi: 10.1016/j.proenv.2016.03.070

Reid, J. S., Xian, P., Hyer, E. J., Flatau, M. K., Ramirez, E. M., Turk, F. J., ... Maloney, E. D. (2012). Multi-scale meteorological conceptual analysis of observed active fire hotspot activity and smoke optical depth in the Maritime Continent. *Atmospheric Chemistry and Physics*, *12*(4), 2117–2147. doi: 10.5194/acp-12-2117-2012

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian ocean. *Nature*, *401*(6751), 360–363. doi: 10.1038/43854

Saji, N. H., & Yamagata, T. (2003). Possible impacts of Indian Ocean Dipole mode events on global climate. *Climate Research*, *25*(2), 151–169. doi: 10.3354/cr025151

Shabbar, A., Skinner, W., & Flannigan, M. D. (2011). Prediction of seasonal forest fire severity in Canada from large-scale climate patterns. *Journal of Applied Meteorology and Climatology*, *50*(4), 785–799. doi: 10.1175/2010JAMC2547.1

Shawki, D., Field, R. D., Tippett, M. K., Saharjo, B. H., Albar, I., Atmoko, D., & Voulgarakis, A. (2017). Long-lead prediction of the 2015 fire and haze episode in Indonesia. *Geophysical Research Letters*, *44*(19), 9996. doi: 10.1002/2017GL073660

Thompson, D. W. J., & Wallace, J. M. (2000). Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate*, *13*(5), 1000 - 1016. doi: 10.1175/1520-0442(2000)013⟨1000:AMITEC⟩2.0.CO;2

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*, 267-288. doi: 10.1111/J.2517-6161.1996.TB02080.X

Trenberth, K. (2013). *El Nino Southern Oscillation (ENSO)* (Tech. Rep.). National Center for Atmospheric Research (NCAR).

van der Werf, G. R., Randerson, J. T., Giglio, L., Gobron, N., & Dolman, A. J. (2008). Climate controls on the variability of fires in the tropics and subtropics. *Global Biogeochemical Cycles*, *22*(3). doi: 10.1029/2007GB003122

Voulgarakis, A., Marlier, M. E., Faluvegi, G., Shindell, D. T., Tsigaridis, K., & Mangeon, S. (2015). Interannual variability of tropospheric trace gases and aerosols: The role of biomass burning emissions. *Journal of Geophysical Research: Atmospheres*, *120*(14), 7157–7173. doi: 10.1002/2014JD022926

Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivariate

MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, *132*(8), 1917 - 1932. doi: 10.1175/1520-0493(2004)132⟨1917:AARMMI⟩2.0.CO;2

Wooster, M. J., Perry, G. L., & Zoumas, A. (2012). Fire, drought and El Niño relationships on Borneo (Southeast Asia) in the pre-MODIS era (1980-2000). *Biogeosciences*, *9*(1), 317–340. doi: 10.5194/bg-9-317-2012

Worden, H. M., Deeter, M. N., Edwards, D. P., Gille, J. C., Drummond, J. R., & Nédélec, P. (2010). Observations of near-surface carbon monoxide from space using mopitt multispectral retrievals. *Journal of Geophysical Research: Atmospheres*, *115*, 18314. doi: 10.1029/2010JD014242

Xavier, P., Rahmat, R., Cheong, W. K., & Wallace, E. (2014). Influence of Madden-Julian Oscillation on Southeast Asia rainfall extremes: Observations and predictability. *Geophysical Research Letters*, *41*(12), 4406–4412. doi: 10.1002/2014GL060241

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, *38*(2), 894–942. doi: 10.1214/09-AOS729

Supporting Information for "Interpretable Models Capture the Complex Relationship Between Climate Indices and Fire Season Intensity in Maritime Southeast Asia"

William S. Daniels[1], Rebecca R. Buchholz[2], Helen M. Worden[2],

Fatimah Ahamad[3], Dorit 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

Contents of this file

1. Text S1 to S4
2. Figure S1 to S2

1. Introduction

This Supporting Information file contains additional text and figures to help interpret the main text of "Interpretable Models Capture the Complex Relationship Between Climate Indices and Fire Season Intensity in Maritime Southeast Asia." Specifically, it contains additional details about:

1. Cloud cover over the study region, which affects sampling bias.

---

2. The smoothing that we apply to the climate mode indices before using them in our model.

3. The mathematical details of the regularization-based model fitting framework we propose in the main text.

2. Cloud cover in Maritime Southeast Asia

There are noticeably fewer observations over landmasses in the Maritime Southeast Asia (MSEA) region than over water scenes (see Figure 1(c) in main text). As we mention in the main text, cloud masking over the region is likely a large contributor to this effect. Figure S2 shows the cloud fraction from the Terra-MODIS cloud mask averaged between 2002 and 2019. On average, there are clearly more clouds over MSEA landmasses than water scenes. Despite the relatively lower number of observations falling over MSEA landmasses, there are still 110 observations per week on average, which we deem a suitable number for creating our response variable.

3. Additional information on climate mode index smoothing

We employ the following smoothing strategy on the climate mode indices used as predictor variables in our models. We do not smooth the indices for lags below four weeks, as we want to capture as much high frequency signal as possible from these very short term relationships. For lags between four and 52 weeks, we use a Gaussian kernel to smooth the indices, with the bandwidth value increasing every four weeks. To select bandwidth values, we first found the bandwidth that seemed to best capture the long term trend in the climate indices. This was then set as the maximum bandwidth and a continuous sequence of bandwidth values was created between no smoothing and this maximum value.

Figure S1 shows every other level of smoothing applied to the climate indices over two years of data. The black curve is the original weekly climate index time series, which is used for lags one through three. The colored curves show every other level of smoothing up to the maximum smoothing applied to lags of one year and greater. Note that the vertical axis has been omitted from Figure S1 for visual clarity since its purpose is solely to show the relative levels of smoothing applied to each climate index.

4. Mathematical details of regularization-based model fitting framework

A general expression for the coefficient estimates generated by regularization is given by

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

where X is the data matrix containing covariates as column vectors, $\beta = (\beta_0, \beta_1, ..., \beta_q)$ is a vector containing an intercept (β_0) and the coefficients corresponding to the covariates in X, y is the response vector, and $p(\beta)$ is some penalty applied to the coefficients. In Equation 1, i iterates through the number of observations (n) and j iterates through the number of covariates (q). The first term is the sum of squared residuals and can be thought of as a measure of fit. The LASSO penalty, given by

$$p(\beta) = \lambda \sum_{j=1}^{q} |\beta_j| \tag{2}$$

has the added benefit of shrinking coefficient estimates to exactly zero, hence performing variable selection (and lag selection for our application). The tuning parameter, $\lambda \geq 0$, is a free parameter that balances the fit term and the penalty term. We discuss our method for selecting λ values shortly.

Instead of the traditional 1-norm used in the LASSO, we apply a slightly more flexible penalty: the minimax concave penalty (MCP). The MCP penalty is given by

$$p(\beta) = \sum_{j=1}^{q} f(\beta_j), \tag{3}$$

where

$$f(\beta_j) = \begin{cases} \lambda|\beta_j| - \frac{\beta_j^2}{2\eta} & \text{if } |\beta_j| \leq \eta\lambda \\ \frac{\eta\lambda^2}{2} & \text{otherwise.} \end{cases} \tag{4}$$

While the LASSO penalty increases linearly with $|\beta_j|$, the MCP penalty gradually levels off until eventually applying a constant penalty after $|\beta_j|$ surpasses a threshold defined by the free parameter $\eta \geq 1$. We discuss our method for selecting η values shortly. The MCP results in less biased estimates for non-zero regression coefficients (Zhang, 2010). Essentially, it allows for larger coefficient estimates on the significant terms (which might be closer to the "true" relationship we are attempting to model). We found that using the MCP penalty over the 1-norm penalty from the LASSO increased model performance. The price we pay for this generality is the introduction of a second parameter, η, in additional to the traditional tuning parameter, λ, that weights the penalty term.

The typical procedure for selecting parameter values (e.g., η and λ) involves minimizing the loss function (i.e., Equation 1) for a sequence of λ values, called a solution path. A single model is then selected from the solution path using an information criterion (e.g., AIC or BIC) or cross-validation test error. Here we use a more general form of the BIC, called the Extended Bayesian Information Criterion (EBIC), given by

$$BIC_\gamma(s) = BIC(s) + 2\gamma \log \tau(s), \tag{5}$$

where s is the model being evaluated, BIC is the standard form of the BIC, τ is the number of possible models with equation dimension (i.e., number of terms) as s, and $\gamma \in [0, 1]$ controls the extra penalty contained in the second term.

The EBIC can apply a much stronger penalty to large models (i.e., models with many selected terms) than the BIC. This is well suited for applications in which the number of possible covariates is large, but the true model might in fact be quite small. Since we believe this to be the case for the atmospheric CO application, we use the EBIC rather than the BIC or cross-validation test error to select λ.

With these more flexible adaptations to the traditional LASSO, we are left with a number of free parameters: λ, the tuning parameter, η, which controls the MCP penalty, and γ, which controls the EBIC. For a given combination of these parameters, we fit the coefficients using the `RAMP` package in `R` (Hao et al., 2018). `RAMP` is a recent regularization method that efficiently computes a hierarchy-preserving solution path for quadratic regression (i.e., models including squared and interaction terms). Enforcing hierarchy, or more specifically strong hierarchy, requires that terms present in an interaction are also present as main effects. Strong hierarchy (also known as the marginality principle) has long been recommended for models with interactions, as it helps avoid misinterpretation of the included covariates (Nelder, 1977). Another benefit of the `RAMP` algorithm is its remarkable efficiency. `RAMP` is able to compute full solution paths much faster than similar hierarchy-preserving algorithms available in `R`, such as `hierNet` (Bien et al., 2013) or `ncvreg` (Breheny & Huang, 2011).

We select parameter values with a simple grid search broken into two steps:

1. Select a γ value on $[0, 1]$. Values closer to 0 will result in larger models and values closer to 1 will result in smaller models.

2. For the given γ value, vary λ and η simultaneously. For each combination of λ and η, fit regression coefficients using the `RAMP` package. Select the model that minimizes the EBIC computed with the selected γ value.

(i) The `RAMP` algorithm automatically computes a data-driven sequence of λ values, so no user input is required.

(ii) We vary η on a logarithmic sequence from 1.001 to 6. This range was selected manually by trial-and-error and tuned specifically for this application. We tested this range on a number of different covariate combinations and response regions (including MSEA), and the selected η value always fell well within this range. Note that the optimal η value is completely data dependent and this sequence will need to be adjusted for different applications or data.

References

Bien, J., Taylor, J., & Tibshirani, R. (2013). A lasso for hierarchical interactions. *The Annals of Statistics*, *41*(3), 1111–1141. doi: 10.1214/13-AOS1096

Breheny, P., & Huang, J. (2011). Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *The Annals of Applied Statistics*, *5*(1), 232–253. doi: 10.1214/10-AOAS388

Hao, N., Feng, Y., & Zhang, H. H. (2018). Model selection for high-dimensional quadratic regression via regularization. *Journal of the American Statistical Association*, *113*(522), 615–625. doi: 10.1080/01621459.2016.1264956

Nelder, J. A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society. Series A (General)*, *140*(1), 48. doi: 10.2307/2344517

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, *38*(2), 894–942. doi: 10.1214/09-AOS729

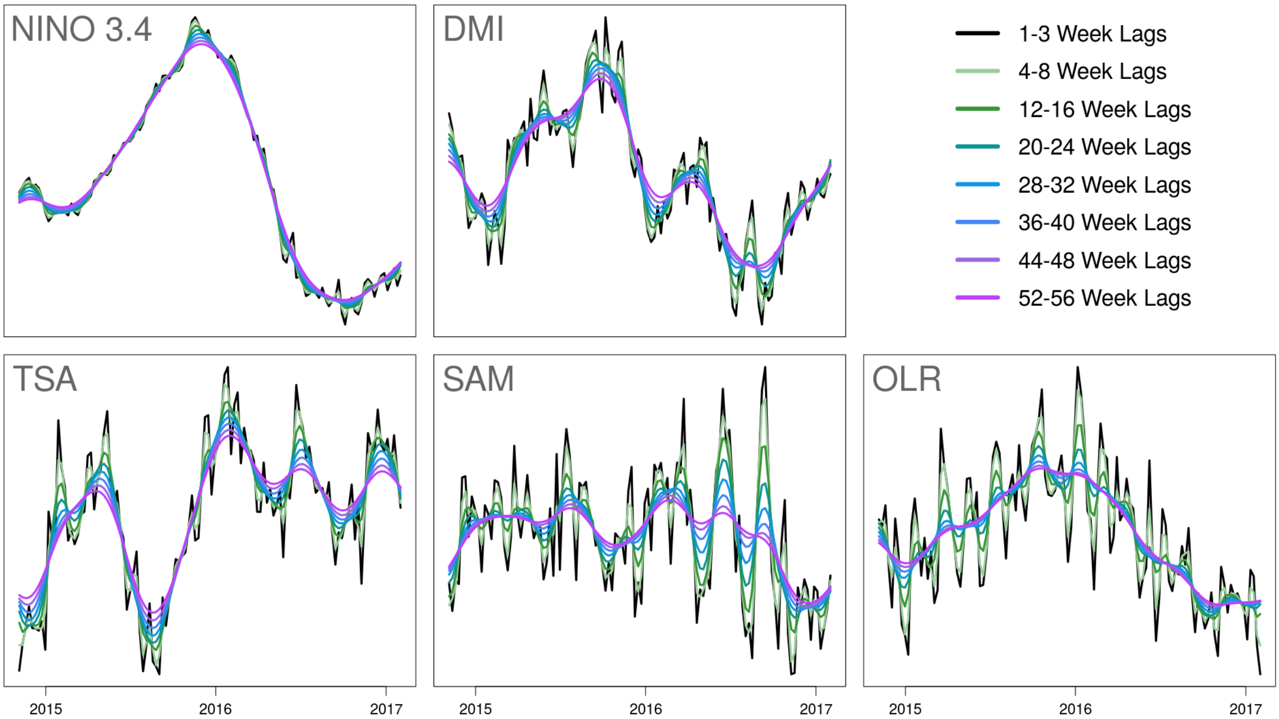

Figure S1. Black curve shows the original climate index data, which is used for lags of one through three weeks. Colored curves show every other level of smoothing applied to the climate index data, which is used for lags of four through 52 weeks. Vertical axis has been omitted for visual clarity.

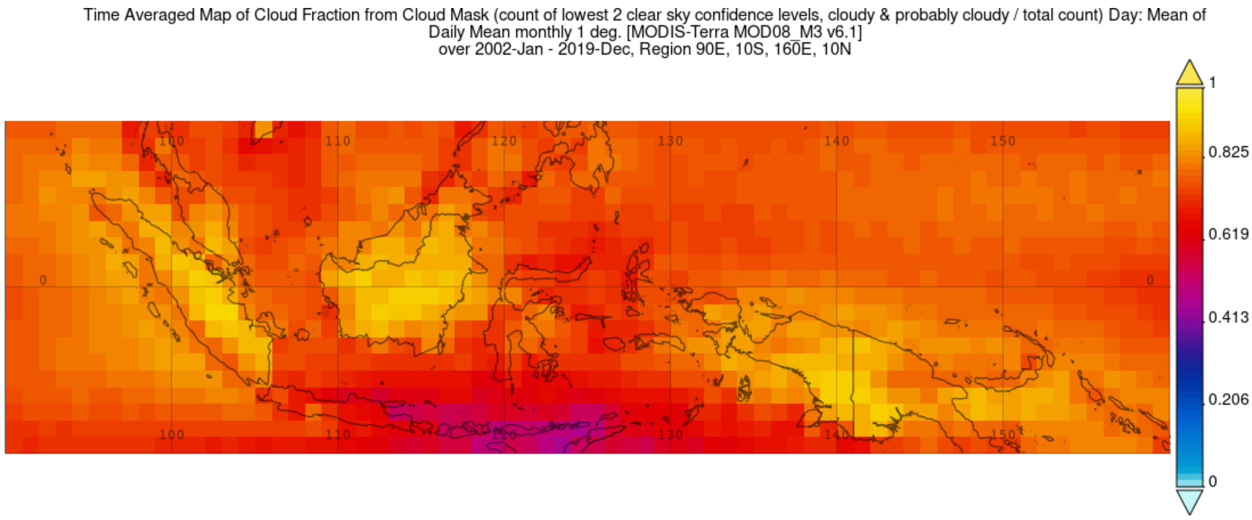

Figure S2. Cloud fraction from Terra-MODIS, averaged between 2002 and 2019, processed with NASA EarthData Giovanni (`https://giovanni.gsfc.nasa.gov/giovanni/`). The scale represents 0 (no cloud) to 1 (all cloud).