Improving Atmospheric Carbon Monoxide Models

William Daniels[1], Dorit Hammerling[1], Rebecca Buchholz[2]

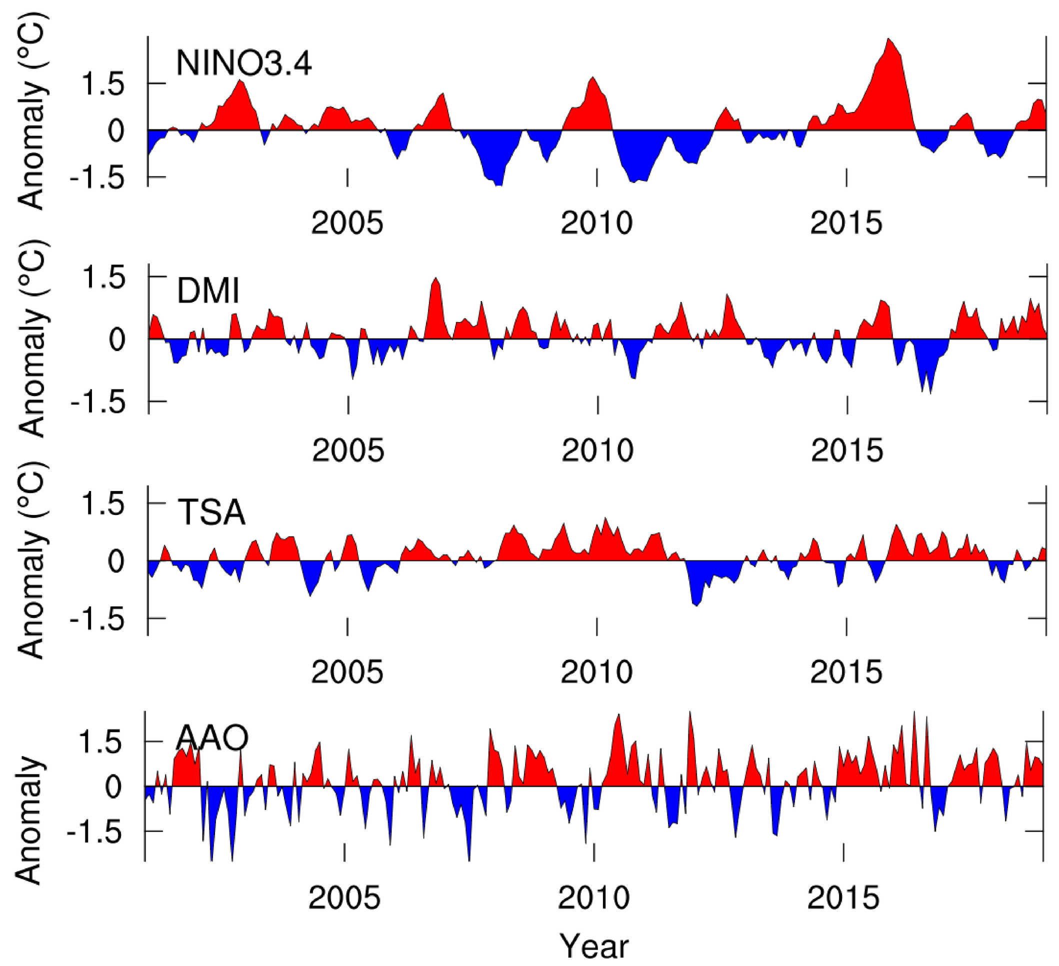
Colorado School of Mines[1], National Center for Atmospheric Research[2]

Motivation

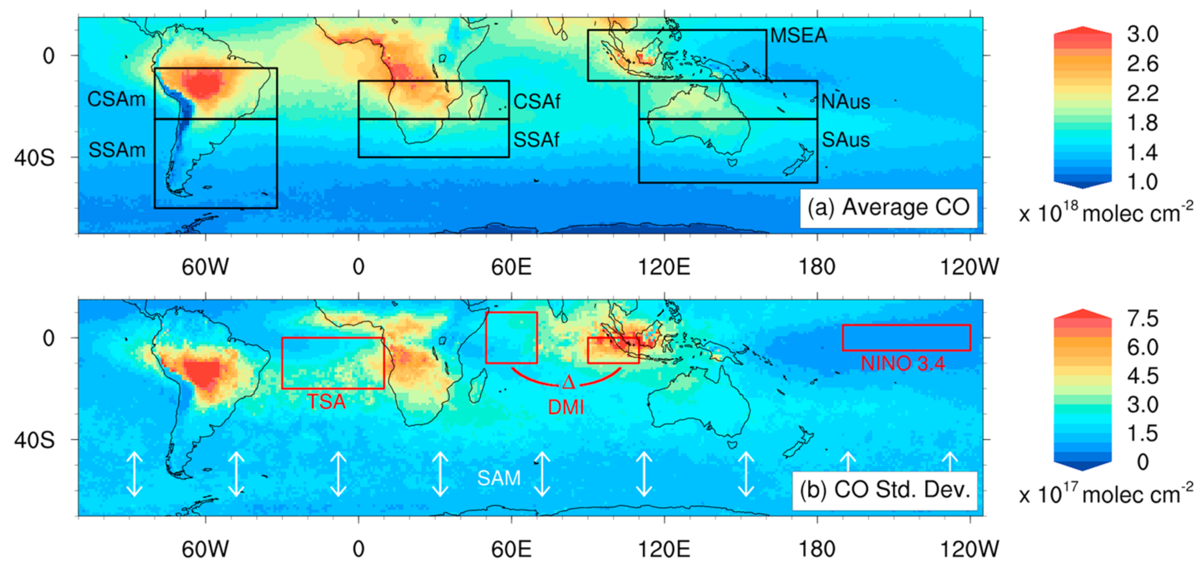
- Large burn events (like the 2019/2020 Australia fires) are the primary source of atmospheric carbon monoxide (CO) in the Southern Hemisphere.
- Atmospheric CO contributes to the greenhouse effect and has negative health impacts for both humans and vegetation.
- Predictive CO models can help countries or cities prepare for large burn events.

Observational Data Sets

- Large burn events are closely tied to the climate through wind and sea surface temperatures, as hot weather dries out vegetation.
- Climate indices provide useful metrics for summarizing the aperiodic changes in climate.



- CO measurements are taken from the MOPITT instrument on board the Terra satellite. They are aggregated into seven biomass burning regions [1].



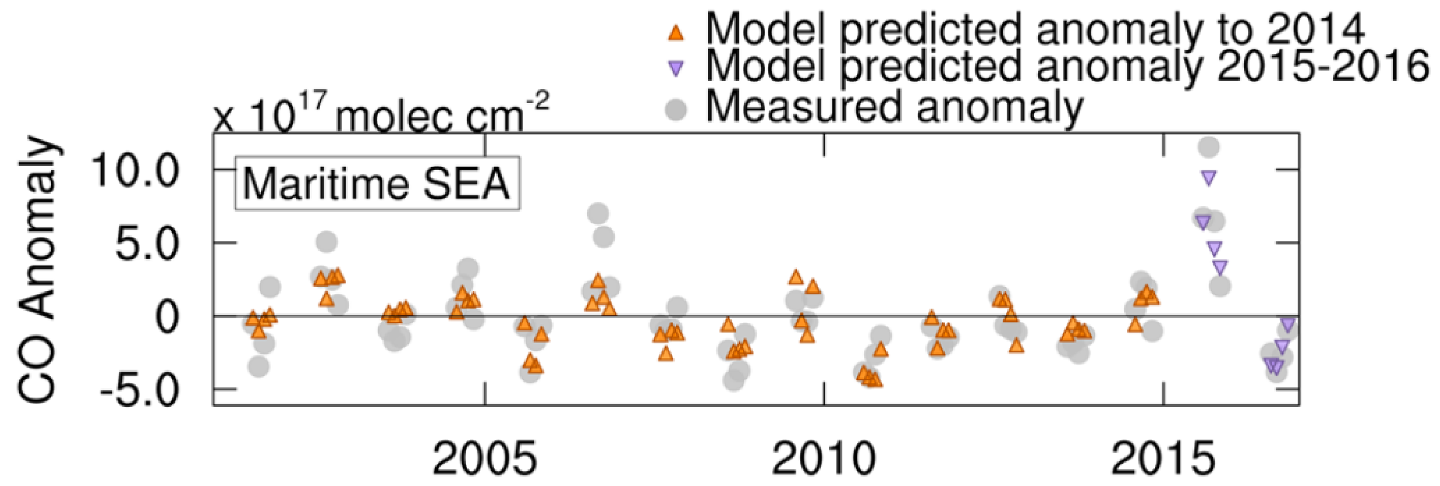
Statistical Model

We use a multiple linear regression model with first order interaction terms to explain the relationship between atmospheric CO and month-averaged climate indices [1, 2].

$$CO(t) = \mu + \sum_k a_k \cdot \chi_k(t - \tau_k) + \sum_{i,j} b_{ij} \cdot \chi_i(t - \tau_i) \cdot \chi_j(t - \tau_j)$$

- $\boldsymbol{CO(t)}$ is the CO anomaly in a given response region at time t
- χ are the climate indices
- τ is the lag value for each index in months

Models are able to explain 50-75% of atmospheric CO variability, depending on the response region [1].

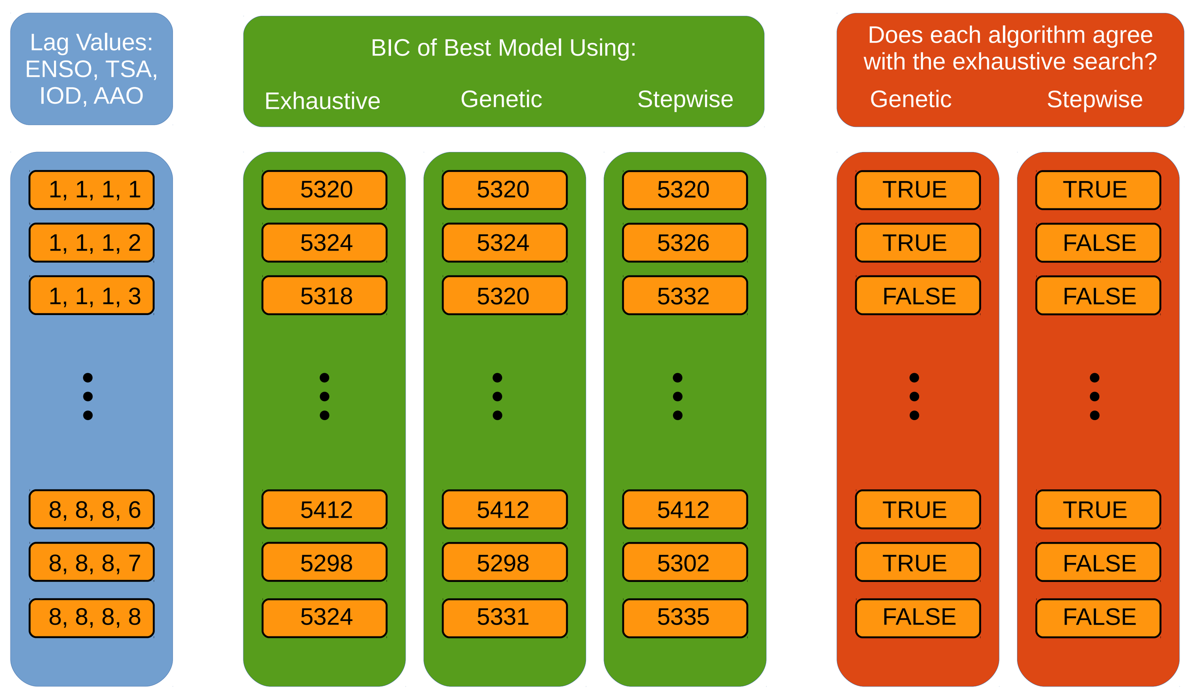


Improvements to Model Selection Codebase

Updates to the model selection codebase have been implemented in the R package `southernHemisphereCO`.

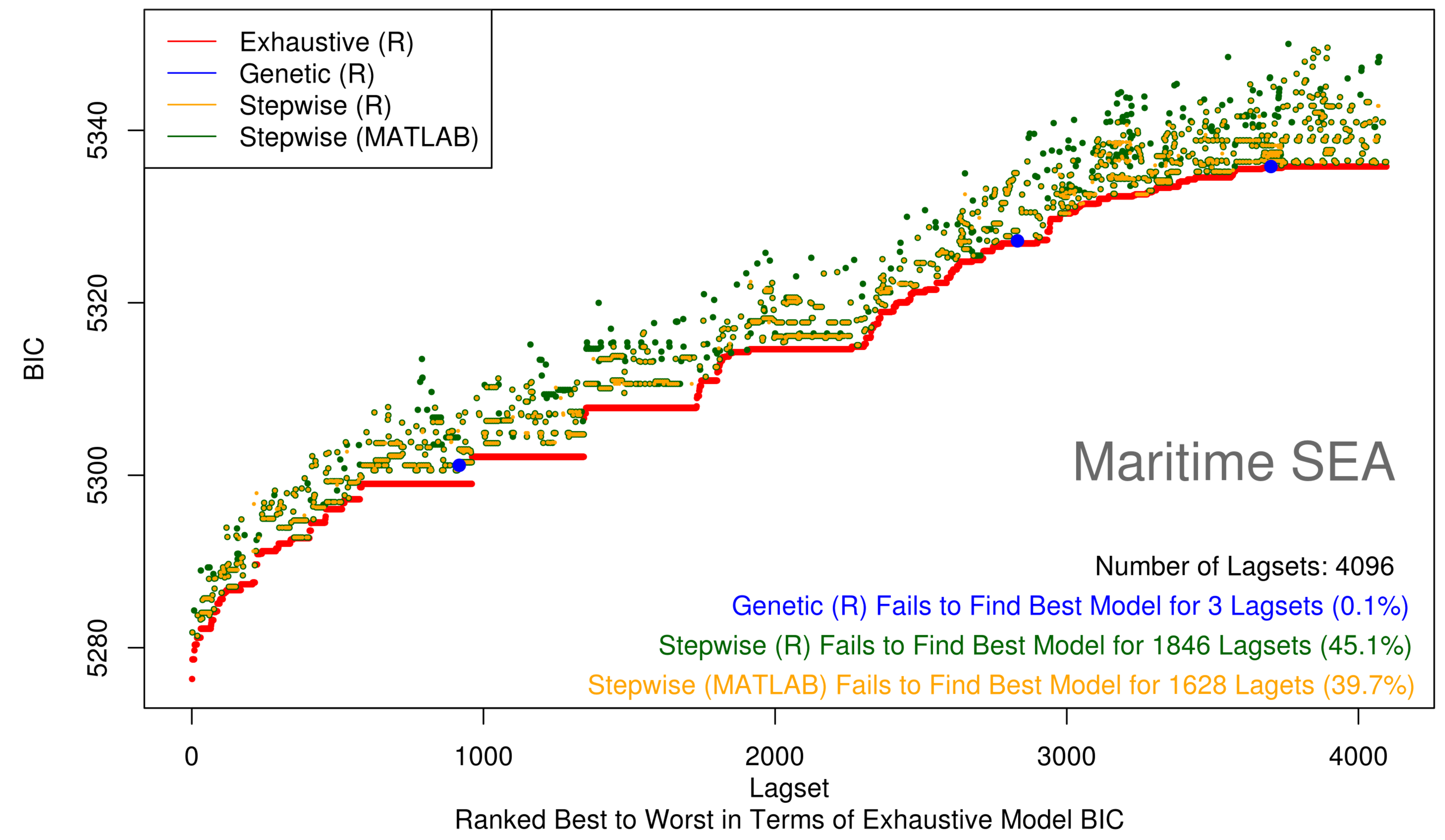
- Multiple lags of a single climate mode index can now be included.
- Exhaustive search added.
- Genetic search added in addition to stepwise selection [3].

Exhaustive search always finds the "best" model. Performance of the other model selection techniques is assessed as follows:



Comparison of Model Selection Capabilities

For each possible lagset, the "best" model found by stepwise selection (MATLAB), stepwise selection (R), the genetic algorithm (R), and an exhaustive search (R) are compared. Note that a lower BIC value corresponds to a better model.



Summary of runtimes in the Maritime SEA response region:

Search Algorithm	Mean Run Time ±1 standard deviation [minutes]
Stepwise (R)	1.9 ±0.04
Exhaustive (R)	7.1 ±0.11
Genetic (R)	7.5 ±0.01
Stepwise (MATLAB)	10.4 ±0.01

Future Work

1. Add MJO climate mode index and compare different ENSO indices.
2. Add an anthropogenic index related to burning.
3. Explore multiple lags of a single index (be careful of correlation).
4. Test models on 2019/2020 Australia fires.

References

[1] R. R. Buchholz, D. Hammerling, H. M. Worden, M. N. Deeter, L. K. Emmons, D. P. Edwards, and S. A. Monks. Links between carbon monoxide and climate indices for the southern hemisphere and tropical fire regions. *Journal of Geophysical Research: Atmospheres*, 123, 2018.

[2] Peter Simonson and Dorit Hammerling. Refactoring data-driven model selection code for improvements in interpretability, generality, and computational expense. *NCAR Technical Notes*, 2018.

[3] Vincent Calcagno and Claire de Mazancourt. glmulti: An R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, 34(12):29, 2010.