AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters®

Key Points:

- We introduce a statistical model to explain interannual variability in tropical cyclone (TC) counts directly from global sea surface temperature maps, and show it performs as well as existing models
- We present evidence that extending current statistical models by nonlinear interactions between predictors does not reduce model error
- We explain these results by arguing that if TC counts are generated by a Poisson process there is a lower limit to model error. We estimate the limit and show it is saturated by existing models

Correspondence to:

D. Wesley,
dhwesley@alumni.princeton.edu

Modeling North Atlantic Tropical Cyclone Counts Directly From Sea Surface Temperature Maps

Daniel Wesley[1] , Michael E. Mann[2], Bhuvnesh Jain[1], Colin R. Twomey[3], and Shannon Christiansen[2]

[1]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA, USA, [2]Department of Earth and Environmental Science, University of Pennsylvania, Philadelphia, PA, USA, [3]Data Driven Discovery Initiative, University of Pennsylvania, Philadelphia, PA, USA

Abstract   Annual North Atlantic tropical cyclone (TC) counts are frequently modeled as a Poisson process with a state-dependent rate. Current models based on Poisson regression can explain roughly 50% of the annual variance using three climate indices: El Niño/Southern Oscillation, average sea surface temperature (SST) in the main development region of the North Atlantic, and the North Atlantic oscillation atmospheric circulation index. We introduce a new method, based on the Elastic Net (EN) that predicts TC counts directly from global SST maps. We show it achieves performance on par with current models, without requiring manually constructed indices. To understand the performance of the EN we argue that, when TC counts are generated by independent Poisson draws, statistical models are subject to a lower limit on prediction error. We estimate this limit and show that it is saturated by both current models and our new method.

Plain Language Summary   Our work studies the relationship between the total number of North Atlantic tropical cyclones (TCs) each year and climate factors, such as sea surface temperature (SST). Many researchers have worked on this problem and have typically needed to invent climate indices by hand to use in explaining TC counts. We introduce a method that predicts TC counts directly from global SST maps, and show it performs as well as previous methods. The new method is in principle sensitive to much more information about global climate and does not require hand-crafted climate indices. We also argue that we may have reached a limit of predictive power and speculate that current models already extract all available information from climate factors to explain TC counts.

1. Introduction

Potential climate influences on the variation over time of North Atlantic Tropical Cyclone (TC) counts has been a topic of active research for some time. Numerous prior studies have examined the importance of various climate factors in influencing year-to-year variation in the seasonal number of named storms (annual TC counts). A small number of climate variables have emerged as being particularly important in modeling annual TC counts. It is well known that the El Niño/Southern Oscillation (ENSO) influences seasonal TC activity through its impact on vertical wind shear (Gray, 1984), with TC counts being enhanced during La Niña and suppressed during El Niño. Warmer ocean surface temperatures promote the formation and development of TCs (Emanuel, 1995; Gray, 1968) and numerous studies have thus incorporated the impact of sea surface temperatures (SST) over the Main Development Region (MDR, 6°–18°N, 20°–60°W) during the peak months of the hurricane season (August–October) (Emanuel, 2005; Hoyos et al., 2006; Mann et al., 2007; Sabbatelli & Mann, 2007). The North Atlantic Oscillation (NAO) is also relevant to modeling Atlantic TC activity (Elsner, 2003; Elsner & Jagger, 2006; Elsner, Jagger, & Niu, 2000; Elsner, Liu, & Kocher, 2000; Mann et al., 2007) through its impact on the tracking of storms, which determines in part whether they encounter conditions favorable for tropical cyclogenesis (Elsner, 2003; Elsner, Liu, & Kocher, 2000; Kossin et al., 2010). For recent reviews of the models and methods used for TC forecasting see (Klotzbach et al., 2019; Takaya et al., 2023) and references therein.

Previous research has demonstrated that basin-wide TC counts can be effectively modeled as a Poisson process conditioned on key climate state variables (Kozar et al., 2012; Sabbatelli & Mann, 2007). In particular, Kozar et al. (2012) considered a broad menu of features and used forward stepwise Poisson regression to show that the most skillful models for annual TC counts include ENSO, MDR SST, and NAO indices as predictors.

This prior work raises the question of whether there are additional features which can help explain interannual variability of TC counts. The predictors identified in prior work have expressions in the global temperature field. In this work we introduce a novel method, based on the Elastic Net (EN), which predicts TC counts directly from map-level SST data and does not require hand-crafted features. In principle this can capture all sources of predictive skill with expressions in global SST maps. We also explore nonlinear combinations of predictors identified in prior studies. We show the EN performs as well as existing methods and that nonlinear extensions do not significantly improve performance. We explain these results by arguing that, if TC counts are truly generated by a Poisson process, there is a natural limit to prediction performance. We estimate this limit and show it is saturated by existing methods as well as our new one.

In Section 2 we introduce our data sources and statistical methods. Section 3 describes the results of the EN and nonlinear extension studies. In Section 4 we estimate the limit on predictive skill that is saturated by these models, and we present our conclusions in Section 5.

2. Data and Methods

2.1. Data

In this paper we use climate indices, global SST data, and annual TC count series. The climate indices used are well described in previous references (e.g., Kozar et al., 2012). The key indices are the in season August-October (ASO) mean temperature in the MDR, the December-February (DJF) Nino 3.4 index and the late-to-post-season boreal winter December-March (DJFM) NAO index. For the global SST data we use the NOAA Extended Reconstructed SST v5 (ERSSTv5) (Huang et al., 2024). This data set provides coverage over nearly the entire ocean surface with mean monthly SST temperatures in a $2° \times 2°$ latitude and longitude grid. For TC counts, we use the adjusted TC counts published in (Vecchi & Knutson, 2008). This timeseries corrects for the improvement over time in the detection of TCs from technological advances such as aircraft reconnaissance and satellites.

2.2. Poisson Regression

Poisson regression has been used in many prior studies of Atlantic TC counts (Elsner, 2003; Elsner, Jagger, & Niu, 2000; Kozar et al., 2012; Mann et al., 2007; Sabbatelli & Mann, 2007). This approach assumes that the probability of observing a number of TCs y_t in year t is governed by the Poisson distribution

$$P(y_t) = \frac{\lambda_t^{y_t}}{y_t!} e^{-\lambda_t} \tag{1}$$

where λ_t parameterizes the mean counts expected in year t. Poisson regression captures variation of observed counts by assuming λ_t varies according to

$$\lambda_t = \exp(\beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_p x_{pt}) \tag{2}$$

where x_{it} is the value of "feature" or "predictor" i in year t, p is the total number of features, and the β_i are coefficients which control the influence of feature i on the expected counts. The coefficient β_0, which we sometimes refer to as the "intercept," controls the mean or unconditional count. Given a data set D with n observations we define the log-likelihood function

$$L_{\text{Poisson}}(\{\beta\}|D) = \log P(D|\{\beta\}) = \sum_{t=1}^{n} \left(y_t \beta_0 + y_t \sum_{j=1}^{p} \beta_j x_{jt} - e^{\beta_0 + \sum_{k=1}^{p} \beta_k x_{kt}} \right) \tag{3}$$

where $\{\beta\} = \beta_0, \ldots \beta_p$ and terms independent of β_i have been dropped. We then "fit" the model by choosing the set of $\hat{\beta}_i$ which maximize L_{Poisson}. In this study we quantify prediction quality using mean absolute error (MAE) defined by

$$E = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t| \tag{4}$$

where $\hat{y}_t$ is the model's prediction for the target in observation t. MAE is simply the expected offset between the prediction and the observed counts. It is less sensitive to rare observations with large residuals $y_t - \hat{y}_t$ than other error measures (such as mean square error) which promotes statistical efficiency on our relatively small data set.

2.3. Cross Validation

To compare models or feature sets, we use the standard statistical meta-algorithm of N-fold cross-validation (NFCV). NFCV ensures prediction quality is always evaluated on data that was held out of training, providing protection against overfitting and a realistic test of the skill of the model on real problems.

We first divide the data D into N partitions or "folds." For each $f = 1, 2, \ldots N$ we set aside fold f as validation data and use the remaining folds to train the model. We use the trained model to generate predictions and compute the error metric E_f on fold f. This generates a sequence of prediction errors $\{E_f\}$ on validation data not seen in training.

To compare two models A and B we use the differences in prediction error

$$\Delta_{BA,f} = E_f^{(B)} - E_f^{(A)} \tag{5}$$

to compute the mean change in prediction error Δ_{BA} and the t-statistic

$$t_{BA} = \frac{\sqrt{N} \, \mathrm{mean}_f(\Delta_{BA,f})}{\mathrm{std}_f(\Delta_{BA,f})}. \tag{6}$$

The mean change in prediction error Δ_{BA} indicates whether B is an improvement over A, and the t-statistic measures the statistical significance of any improvements. If B reduces errors then Δ_{BA} and t_{BA} will be negative, hence negative t-statistics are preferred.

For this study, we divide the 140 yearly observations 1880–2019 into $N = 5$ equally sized folds 1880–1907, 1908–1935, etc. This maximizes the chances that adjacent years (which may not be statistically independent) are assigned to the same fold.

2.4. Elastic Nets

Previous analyses (Kozar et al., 2012) have generally incorporated SST data using features inspired by an understanding of the processes involved in hurricane formation and climate. Here we investigate whether there is additional information in global SST data that is not expressed in the existing hand-crafted features. While we are only using SST data, other drivers of Atlantic TC counts (such as wind or current patterns) may be incorporated indirectly into the model through their correlation with the SST field. For example, the Nino3.4 index is based on the tropical Pacific SST field, but it is actually a metric of how the ENSO phenomenon impacts wind patterns in the tropical Atlantic that govern TC formation.

We incorporate the full SST data using an algorithm that takes global SST map-level data and uses it to predict the annual TC count directly. The technique is designed to ignore temperature observations that are not useful for predicting TCs. Each month, the ERSSTv5 data set provides temperature data on roughly 8.8×10^3 grid points covering the globe, while we have only $n = 142$ observations of annual TC counts, so we are deeply in the $p \gg n$ statistical regime.

A given temperature observation τ_{xymt} in ERSSTv5 is specified by four indices (x, y, m, t) giving the latitude and longitude indices (x, y) of the grid location, the observation month m (January–December) and the observation year t. We process the features into normalized versions by computing the mean and standard deviations over t on the training data,

$$\mu_{xym} = \mathrm{mean}_t(\tau_{xymt}), \ \sigma_{xym} = \mathrm{std}_t(\tau_{xymt}). \tag{7}$$

For clarity we replace the triplet of indices (x, y, m) with a single consolidated pixel index i. In our terminology a pixel refers to a specific geographical location together with an observation month. Using the pixel index i we define normalized features z_{it} by

$$z_{it} = \frac{\tau_{it} - \mu_i}{\sigma_i}. \tag{8}$$

On the training data z_{it} has zero mean and unit standard deviation by construction. When performing cross-validation (See Section 2.3) it is crucial that we do not leak any information about validation data into training. On validation data, we use the same Formula 8 but apply the μ_i and σ_i computed from training data.

We adapt the EN (Zou & Hastie, 2005) to the Poisson regression framework. To construct the objective function L_{EN} for training we use the SST features z_{it} in the Poisson regression log-likelihood 3 but add additional regularization terms inspired by the EN

$$L_{\text{EN}}(\beta_0, \beta_1, \dots \beta_X, \lambda_1, \lambda_2) = L_{\text{Poisson}}(\beta_0, \beta_1, \dots \beta_X) - \lambda_1 \sum_{i=1}^{X} |\beta_i| - \frac{1}{2}\lambda_2 \sum_{i=1}^{X} \beta_i^2 \tag{9}$$

where X is the total number of pixels and $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are regularization parameters. λ_1 encourages sparsity and performs variable selection by encouraging weak features to acquire zero coefficients. The λ_2 term encourages the model to assign similar weights to highly correlated features. Note we do not apply penalties to the intercept term β_0.

The EN is trained using a two-step process. In the first step we use a gradient ascent algorithm to find good values of the coefficients. Denoting by β_i^s the value of coefficient β_i in step s of this process, we initialize $\beta_i^0 = 0$ for $i > 0$ and $\beta_0^0 = \text{logmean}_t\, y_t$. In each training step we adjust the values of the β_i^s by

$$\beta_i^{s+1} = \beta_i^s + \alpha \frac{\partial L_{\text{EN}}}{\partial \beta_i} \tag{10}$$

where α is the learning rate. We have found $\alpha = 10^{-6}$ works well in practice. Once gradient ascent ceases to improve the objective function L_{EN} on training data we halt the algorithm and assign preliminary values $\beta_i' = \beta_i^{s'}$ where s' is the step with the maximum value of the objective function.

In the second training step, we adjust values of the coefficients. The EN penalties encourage coefficients to shrink toward zero and hence predictions of TC counts will be biased. To reduce bias, we construct an aggregated feature w_t using coefficients from the first stage fits

$$w_t = \sum_{i=1}^{X} \beta_i' z_{it} \tag{11}$$

then do a second Poisson regression on the training data using the two features $\{\beta_0, w_t\}$. This second Poisson regression yields coefficients γ_0 on the constant term and γ_1 on the aggregate feature. This yields fitted coefficients $\hat{\beta}_0 = \gamma_0 \beta_0'$ and $\hat{\beta}_i = \gamma_1 \beta_i'$ for $i > 0$. We then use 2 with features z_{it} to generate predictions. While we use the standard Poisson formula for prediction, the EN procedure generally produces different coefficient values than the standard Poisson regression procedure.

3. Results

We use the model of (Kozar et al., 2012) as our baseline model. To ensure a fair comparison of models, we only allow the EN model to train over data with the same month ranges that are available to the baseline model. We train the EN using SST data from current year August through next year March.

Table 1
Cross-Validated Performance Metrics for Baseline and Elastic Net Models

Model	Absolute		Relative		
	Error	Std	Error	Std	*t*-stat
Baseline	2.46	0.10			
EN (0.316, 10^4)	2.49	0.14	0.05	0.12	0.41

3.1. Elastic Net

We find that the EN can match the performance of the baseline model. The EN with $(\lambda_1, \lambda_2) = (0.316, 10^4)$ achieves cross-validated prediction error of 2.49 ± 0.14 versus 2.46 ± 0.10 for the baseline model. The performance difference is not statistically significant ($t = 0.41$). See Table 1 for details. Generally there is a broad region in the (λ_1, λ_2) plane that gives good EN performance: the values reported here were chosen through a coarse grid search.

In addition to providing a prediction, the EN also generates a map showing what oceanic regions and months contain SST information relevant to TCs. Each feature i in the EN corresponds to SST at a specific grid location during a specific month of the year. Hence, the coefficients β_i give a measure of the weight that the EN puts on that location and month in generating its predictions. The β_i for the EN described above are shown in Figure 1. Most of the pixels are close to zero, reflecting the influence of the λ_1 penalty in eliminating some features. The maps clearly show the weight placed on the area near the MDR during the TC season, with the weight falling off post-season. This reflects the importance of the MDR SST in driving TC formation. We also see features in the Pacific, notably an area near the Niño 3.4 region which acquires a prominent negative weight post season. This is consistent with the baseline model, in which higher temperatures in the Niño 3.4 region correspond to a decrease in the TC prediction.

It is also possible to train an EN using only SST data available pre-season. This model can then be used for generating true forecasts of seasonal TC counts before the season commences, which is of practical interest. We have trained a model using SST data over prior year December through current year April, resulting in β_i as shown in Figure 2. This map shows a strong positive weight put on the MDR area, increasing as the season approaches, which is consistent with the notion that MDR temperatures are important for TC formation. This EN yields an error of 2.78 ± 0.12. Unfortunately we do not have history for a baseline forecasting model available and so we cannot report a relative EN performance comparison.

3.2. Nonlinear Interactions

The Poisson regression framework allows us to explore nonlinear interactions between the predictors in the baseline model. We do this by constructing "product features" and testing whether they improve prediction errors when added to the baseline model.

Given two features *x* and *y*, we define the product feature $x \star y$ by demeaning *x* and *y* on the training data, then multiplying the resulting values together observation by observation. On validation data, we use the same means derived from the training data, so this procedure is consistent with the cross-validation procedure. We then test a series of models, one per product feature, obtained by adding the product feature to the features in the baseline model.

Table 2 summarizes the result of this test for the baseline model. Of the candidate predictors tested, only the product `nino34_djf * nao_djfm` showed a reduction in prediction error when added to the baseline model: prediction error decreases by a small amount (0.03) but the statistical significance of the effect is fairly strong ($t = -4.71$) and other checks show it is quite consistent across cross-validation folds.

In addition to the features found in the baseline model, we have also examined all product features using the full 10 predictor set studied in Kozar et al. (2012). This study revealed no interesting product features, except those trivially related to the `nino34_djf * nao_djfm` one described above. We have also explored some other forms of nonlinear interaction that did not reveal additional features of interest.

4. Limits on Explanatory Skill

In this section we argue that, if the annual TC counts are truly generated by independent Poisson draws with different rates each year, then any statistical model that predicts TC counts faces a lower limit on the MAE achievable on validation data. We estimate this limit on our data and show that it is saturated by both the baseline model (Kozar et al., 2012) and the EN.
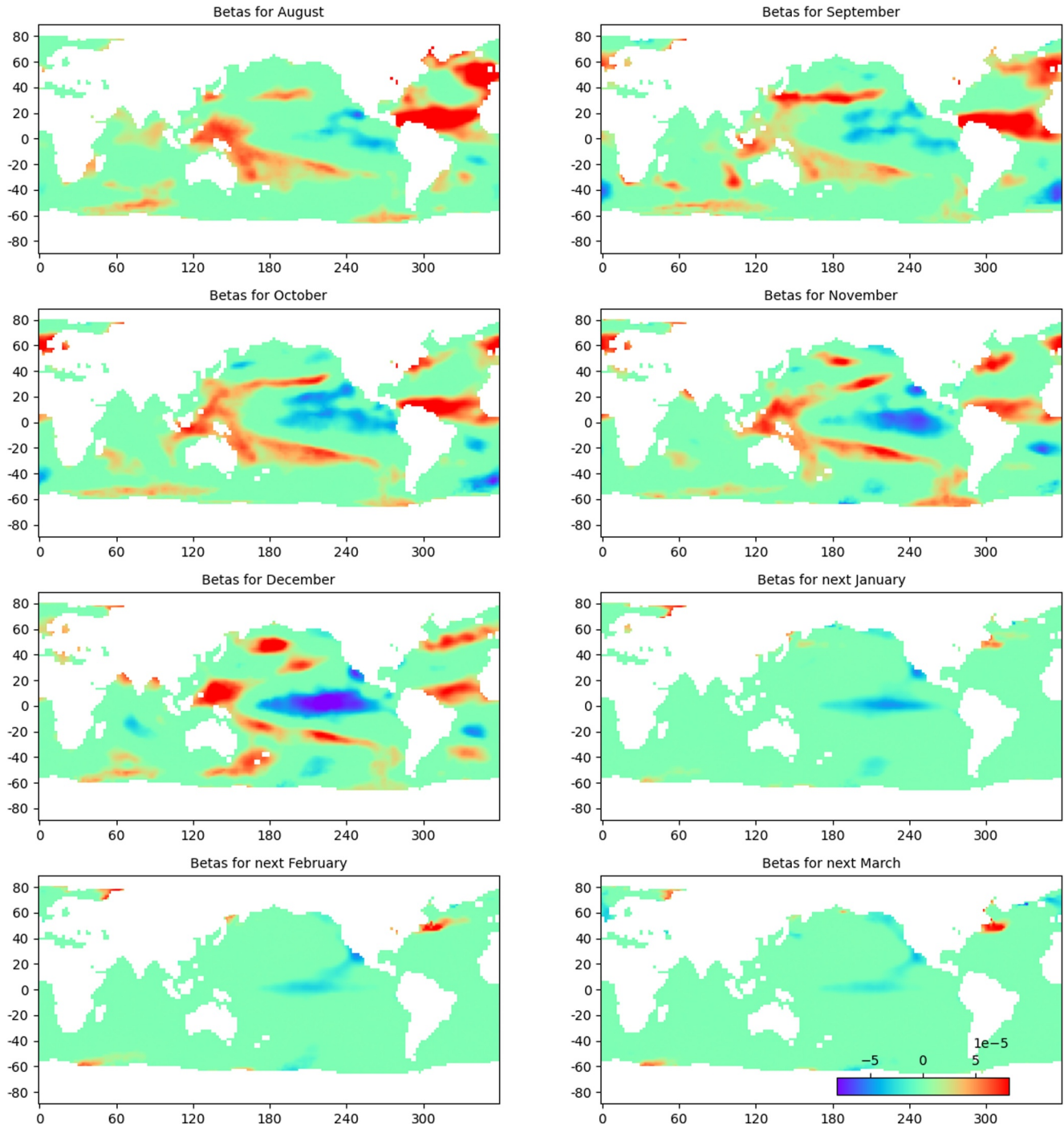
Figure 1. Coefficients $\hat{\beta}_i$ for the Elastic Net (EN) fit with $(\lambda_1, \lambda_2) = (0.316, 10^4)$ illustrating the weight the EN puts on each sea surface temperature observation.

The data passes two checks of the independent Poisson hypothesis. First, following (Sabbatelli & Mann, 2007) we compute the test statistic

$$x = \sum_{t=1}^{n} \frac{(y_t - \hat{y}_t)^2}{\hat{y}_t} \tag{12}$$
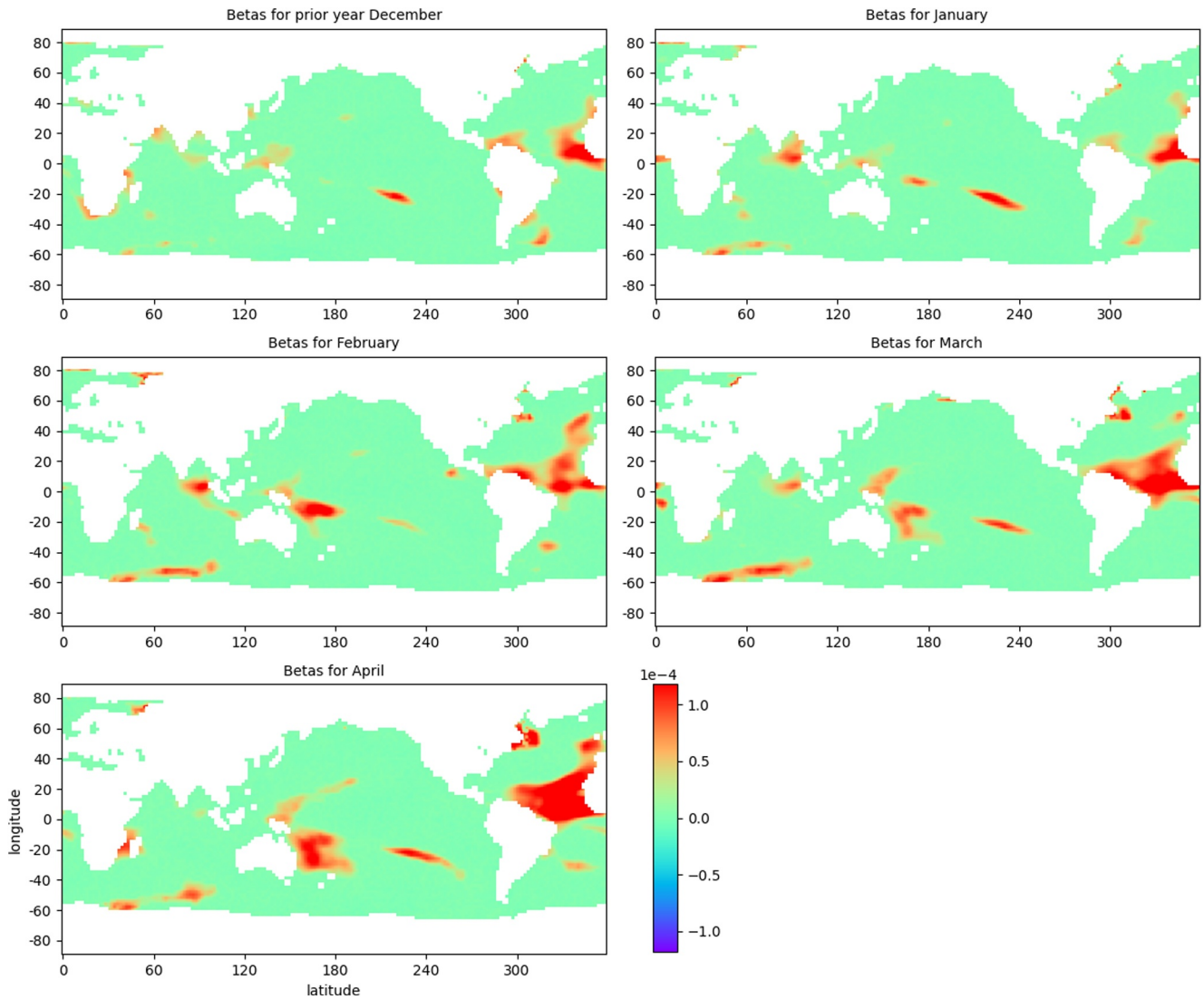
Figure 2. Coefficients $\hat{\beta}_i$ for an Elastic Net (EN) fit with $(\lambda_1, \lambda_2) = (1, 10^4)$ trained to generate pre-season forecasts, illustrating the weight the EN puts on each sea surface temperature observation.

where y_t are the observed counts and $\hat{y}_t$ are the counts predicted by the model in year t. We use the model of Kozar et al. (2012) and Sabbatelli and Mann (2007) fit over our entire sample (no cross validation) to define $\hat{y}_t$, yielding results fully consistent with Poisson distributed residuals ($p = 0.84$). Second, we find the serial autocorrelation of residuals to this model is 11%, mildly positive but consistent with zero ($z = 1.3$). There may be deviations from independent Poisson draws not captured by these tests, but these results support the independent Poisson assumption.

Focusing on a single year, the expected error for a prediction of z counts is

$$\text{err}(z, \lambda) = \sum_{y \geq 0} |y - z| P(y|\lambda) \tag{13}$$

where $P(y|\lambda)$ is the probability of observing y counts with Poisson rate λ according to 1. For fixed z this error has a lower bound

Table 2
Cross Validated Performance Metrics With Candidate Nonlinear Terms Added

Model	Absolute		Relative		
	Error	Std	Error	Std	*t*-stat
baseline	2.46	0.10			
mdr_aso * mdr_aso	2.63	0.19	0.16	0.126	1.30
mdr_aso * nino34_djf	2.50	0.09	0.04	0.009	4.36
mdr_aso * nao_djfm	2.47	0.10	0.00	0.009	0.37
nino34_djf * nino34_djf	2.47	0.11	0.01	0.013	0.84
nino34_djf * nao_djfm	2.43	0.11	−0.03	0.007	−4.71
nao_djfm * nao_djfm	2.47	0.10	0.00	0.007	0.74

$$b(z) = \min_{\lambda} \mathrm{err}(z, \lambda). \tag{14}$$

In our application we are given observed counts y_t in each year t, but the Poisson rates λ_t are unobservable. We can estimate the lower limit for MAE by noting that a very good model will make predictions z_t with a similar distribution as the actual counts y_t, so we estimate the lower bound B to MAE by averaging over observations

$$B = \frac{1}{n} \sum_{t=1}^{n} b(y_t). \tag{15}$$

The Formulas 13–15 apply to validation data only. Models can overfit to training data, potentially achieving zero prediction errors. On validation data, not used in training, these models are still subject to Poisson uncertainty and we estimate likely errors using the expectation 13. In any particular realization of the data, we may exceed B by chance. On average or for a large number of observations, we expect it to be accurate.

When we compute B for our training data we obtain $B = 2.51$. This is consistent with the baseline model performance of 2.46 ± 0.10 given in Table 1. The bound is also consistent with the EN model MAE of 2.49 ± 0.14. While the baseline model appears to be slightly lucky at the 0.5σ level, both of these models saturate the error bound.

Note the estimated lower limit 15 is independent of the model used to generate the predictions of TC rates λ_t in each year. Our argument makes no reference to how the rate predictions are generated. So long as the TC counts are truly generated by independent Poisson draws each year, the bound should apply.

5. Conclusions

Prior work modeling annual TC counts as a Poisson process with a state-dependent rate has revealed that roughly 50% of the annual variance can be predicted using three climate indices: ENSO, average SST in the MDR of the North Atlantic and NAO atmospheric circulation index (Kozar et al., 2012). The results presented here shows that nonlinear extensions of these features does not improve performance of this baseline model.

In this work we demonstrate a method, based on the EN, to predict TC counts directly from global SST temperature maps. This method includes a large potential universe of features and matches the performance of prior models without the need to create features by hand. As a byproduct, the EN generates maps which illustrate the ocean regions most useful for predicting TCs. These maps may be useful in designing new climate indices and understanding linkages between TC and global climate observables.

To understand the performance of the EN we present arguments that, if the TC counts are generated by independent Poisson draws, statistical models that predict TC counts are subject to a lower limit on prediction error. We estimate this limit and show that it is saturated by both the baseline model (Kozar et al., 2012) and the EN.

These results suggest some avenues for future research on improving statistical models of TC counts. So long as the TC counts are well described by independent Poisson draws, and given observed model performance, our limit suggests that further refinements to statistical models will not reduce prediction error on validation data. We have reported some simple tests of the independent Poisson assumption in this work, but in the future, a careful study of the TC count series could potentially identify violations. There could be a subtle relationship between counts in different years which is not captured by our tests, perhaps modulated by a conditioning variable we have not yet identified. This would violate the independence assumption and weaken or invalidate the limit, and awareness of the violation could point the way to improved models. We hope future work will illuminate this issue further.

Finally, we note that our study shares some limitations with all attempts at modeling TC counts, namely that the available historical record represents a small number of observations and that possible effects of climate change may lead to changes in the causal relationships that are difficult to discern from recent data.

Data Availability Statement

The adjusted TC counts published in Vecchi and Knutson (2008) and climate indices were originally obtained from the Penn State/IBM Nittany AI Alliance. This data, together with analysis and figure generating code, is publicly available at https://doi.org/10.5281/zenodo.14767663 (Wesley, 2025). The NOAA Extended Reconstructed SST V5 data (ERSSTv5) (Huang et al., 2024) is provided by the NOAA PSL, Boulder, Colorado, USA with access information at https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html. The figures for this manuscript were generated using the Anaconda software distribution version 24.7.1 available from https://www.anaconda.com.

References

Elsner, J. B. (2003). Tracking hurricanes. *Bulletin of the American Meteorological Society*, *84*(3), 353–356. https://doi.org/10.1175/BAMS-84-3-353

Elsner, J. B., Jagger, T., & Niu, X.-F. (2000). Changes in the rates of North Atlantic major hurricane activity during the 20th century. *Geophysical Review Letters*, *27*(12), 1743–1746. https://doi.org/10.1029/2000GL011453

Elsner, J. B., & Jagger, T. H. (2006). Prediction models for annual U.S. hurricane counts. *Journal of Climate*, *19*(12), 2935–2952. https://doi.org/10.1175/JCLI3729.1

Elsner, J. B., Liu, K.-B., & Kocher, B. (2000). Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism. *Journal of Climate*, *13*(13), 2293–2305. https://doi.org/10.1175/1520-0442(2000)013⟨2293:SVIMUS⟩2.0.CO;2

Emanuel, K. A. (2005). Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, *436*(7051), 686–688. https://doi.org/10.1038/nature03906

Emanuel, K. A. (1995). Sensitivity of tropical cyclones to surface exchange coefficients and a revised steady-state model incorporating eye dynamics. *Journal of the Atmospheric Sciences*, *52*(22), 3969–3976. https://doi.org/10.1175/1520-0469(1995)052⟨3969:SOTCTS⟩2.0.CO;2

Gray, W. M. (1968). Global view of the origin of tropical disturbances and storms. *Monthly Weather Review*, *96*(10), 669–700. https://doi.org/10.1175/1520-0493(1968)096⟨0669:GVOTOO⟩2.0.CO;2

Gray, W. M. (1984). Atlantic seasonal hurricane frequency. Part I: El Niño and 30 mb quasi-biennial oscillation influences. *Monthly Weather Review*, *112*(9), 1649–1668. https://doi.org/10.1175/1520-0493(1984)112<1649:ashfpi>2.0.co;2

Hoyos, C. D., Agudelo, P. A., Webster, P. J., & Curry, J. A. (2006). Deconvolution of the factors contributing to the increase in global hurricane intensity. *Science*, *312*(5770), 94–97. https://doi.org/10.1126/science.1123560

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2024). NOAA extended reconstructed sea surface temperature (ERSST), version 5. https://doi.org/10.7289/V5T72FNM

Klotzbach, P., Blake, E., Camp, J., Caron, L.-P., Chan, J. C. L., Kang, N.-Y., et al. (2019). Seasonal tropical cyclone forecasting. *Tropical Cyclone Research and Review*, *8*(3), 134–149. https://doi.org/10.1016/j.tcrr.2019.10.003

Kossin, J. P., Camargo, S. J., & Sitkowski, M. (2010). Climate modulation of North Atlantic hurricane tracks. *Journal of Climate*, *23*(11), 3057–3076. https://doi.org/10.1175/2010JCLI3497.1

Kozar, M. E., Mann, M. E., Camargo, S. J., Kossin, J. P., & Evans, J. L. (2012). Stratified statistical models of North Atlantic basin-wide and regional tropical cyclone counts. *Journal of Geophysical Research*, *117*(D18), D18103. https://doi.org/10.1029/2011JD017170

Mann, M. E., Sabbatelli, T. A., & Neu, U. (2007). Evidence for a modest undercount bias in early historical Atlantic tropical cyclone counts. *Geophysical Review Letters*, *34*(22), L22707. https://doi.org/10.1029/2007GL031781

Sabbatelli, T., & Mann, M. (2007). The influence of climate state variables on Atlantic tropical cyclone occurrence rates. *Journal of Geophysical Research*, *112*(D17), D17114. https://doi.org/10.1029/2007JD008385

Takaya, Y., Caron, L.-P., Blake, E., Bonnardot, F., Bruneau, N., Camp, J., et al. (2023). Recent advances in seasonal and multi-annual tropical cyclone forecasting. *Tropical Cyclone Research and Review*, *12*(3), 182–199. https://doi.org/10.1016/j.tcrr.2023.09.003

Vecchi, G. A., & Knutson, T. R. (2008). On estimates of historical North Atlantic tropical cyclone activity. *Journal of Climate*, *21*(14), 3580–3600. https://doi.org/10.1175/2008JCLI2178.1

Wesley, D. (2025). bounds_tc_counts: Version 2.1 [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.14767663

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society - Series B: Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x