EOF Analysis Preliminary Results

October 19, 2012

Overview

We analyze the empirical orthogonal function (EOF) of mean sea surface temperatures (SSTs) anomalies from August-October in the Atlantic (Atl), Pacific (Pac), and joint Atlantic Pacific (joint) basins. The goal is to determine if there is significant information in the EOF's principal components (PCs) that would capture SST's impact on seasonal tropical cyclone (TC) activity on interannual timescales.

EOF and First Principle Component

Atlantic Basin

We begin by averaging the SST over August to October for each year and compute the EOF and PCs of the resulting seasonal SST. **Figure 1** (top panel) shows both the Aug-Oct Atlantic SST first PC and Aug-Oct Atlantic TC counts. The quantities have a 0.6 linear correlation coefficient. However, there is a a noticeable upward trend in both time-series starting around 1995 that contributes significantly to the positive correlation. If we de-trend the two time-series by removing the decadal mean, the linear correlation drops from 0.6 to 0.33 (see **Figure 1** bottom panel), which confirms that a large portion of the positive correlation is due to the upward trend in both time-series. Although has to be mindful of the significance of the trend given the short record.

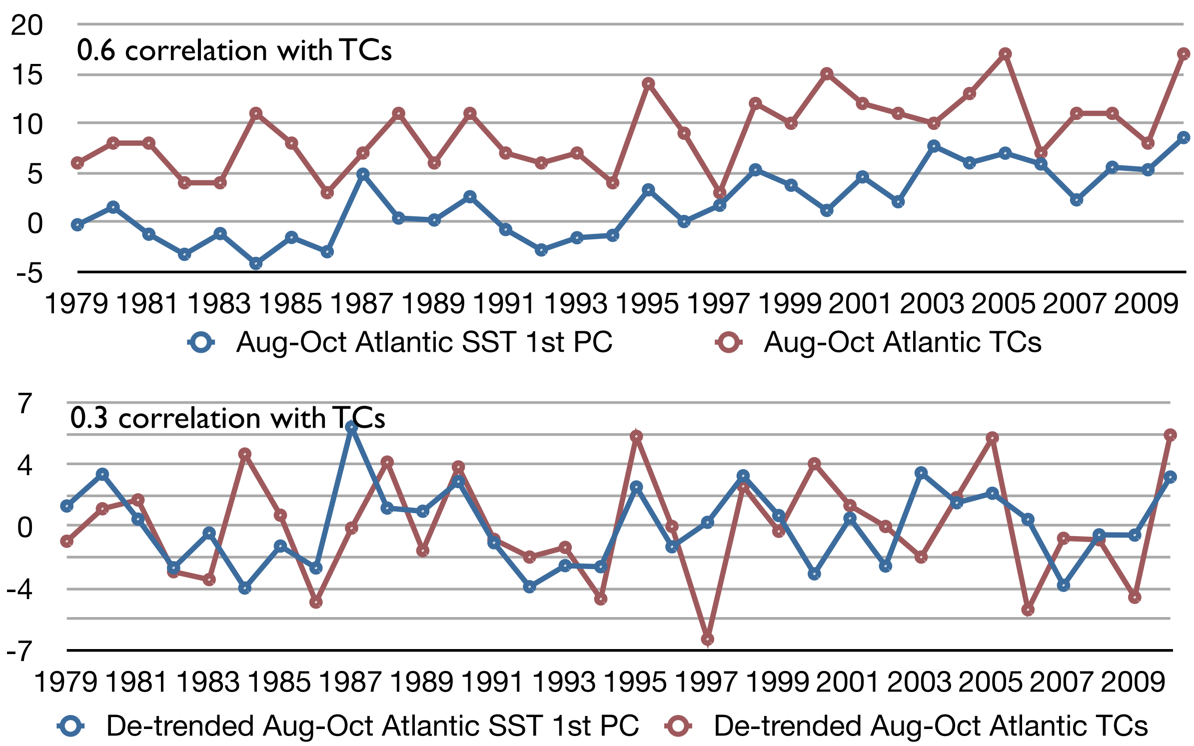


Figure 1 - *Top*: Correlation between Atlantic Aug-Oct SST 1st PC (blue) and Aug-Oct TCs.(red). *Bottom*: Correlation between Atlantic Aug-Oct SST 1st PC (blue) and Aug-Oct TCs.(red). The majority of the linear correlation between the Atlantic SST's 1st PC and TCs is the shared upward linear trend between the two time-series

Pacific Basin

We repeat the same analysis using the Aug-Oct Pacific SST EOF and first principal component. Once we compute the first PC, we correlate it with Aug-Oct Atlantic TCs. Unlike, the Atlantic 1st PC, its Pacific counterpart does not exhibit the significant upward linear trend and as it can be seen in **Figure 2** (bottom panel) the correlation remains virtually unchanged once we de-trend the data. This suggests that on interannual timescales, Pacific SSTs provide robust information about Atlantic TC frequency,

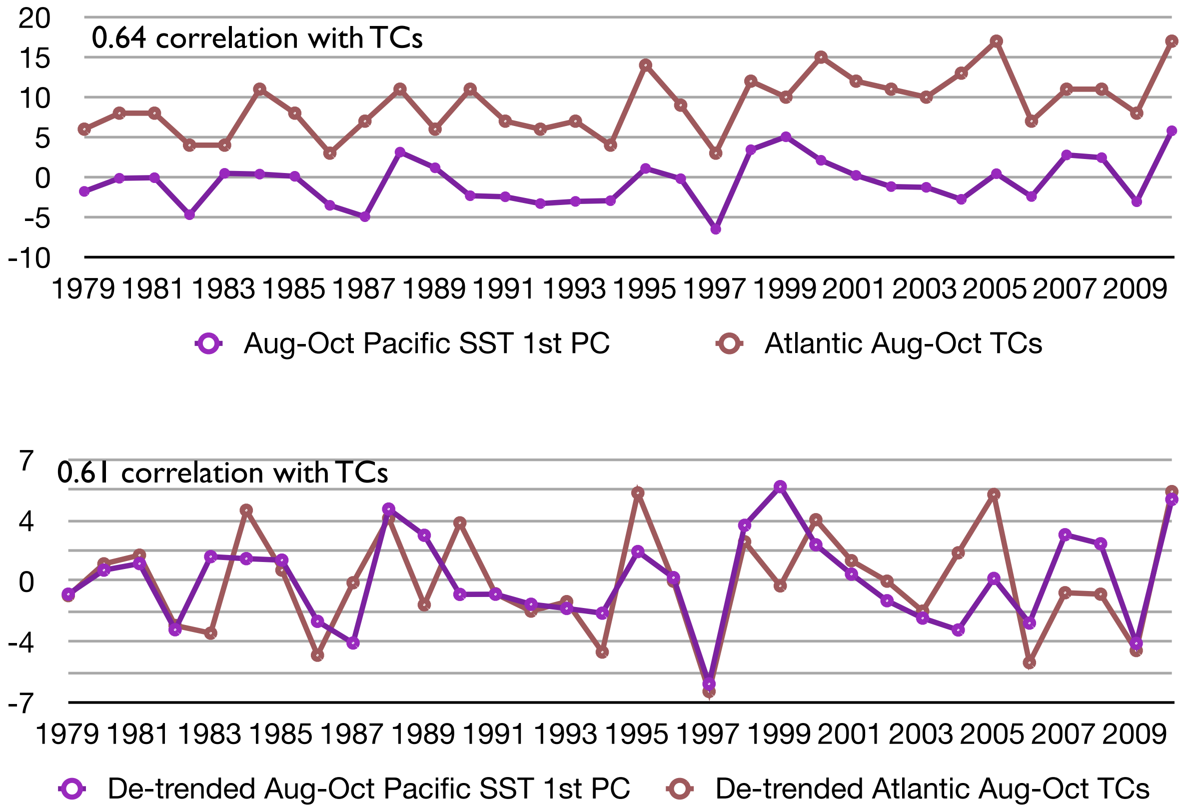


Figure - *Top*: Correlation between Pacific Aug-Oct SST 1st PC (purple) and Aug-Oct TCs (red). *Bottom*: Correlation between Pacific Aug-Oct SST 1st PC (purple) and Aug-Oct TCs (red). The linear correlation between the two time-series is not due a linear trend.

Joint Basin

In this analysis, we join the Atlantic and Pacific ocean basins (excluding land) and compute the joint basin's mean Aug-Oct SST. Subsequently we compute its EOF and PCs.

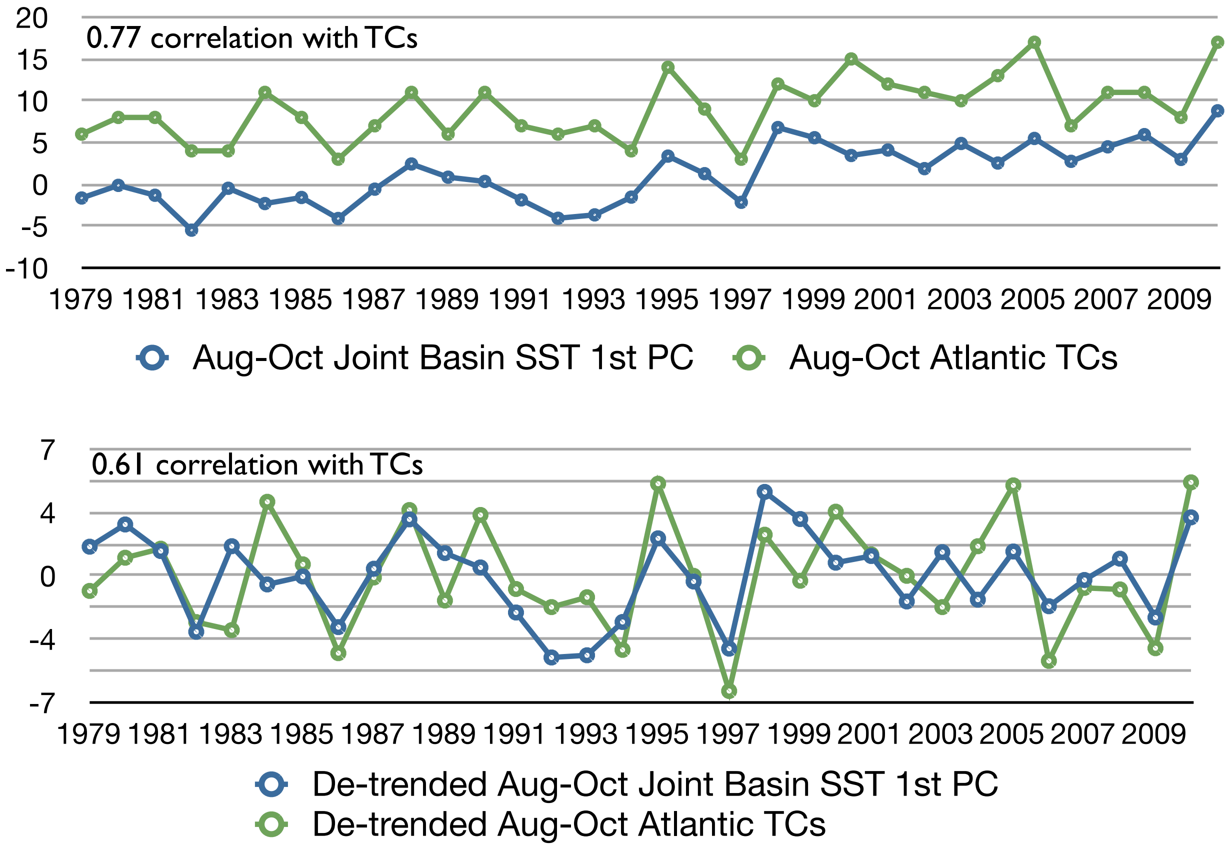


Figure 3 - *Top*: Correlation between Joint Basin Aug-Oct SST 1st PC (blue) and Aug-Oct TCs (green). *Bottom*: Correlation between Pacific Aug-Oct SST 1st PC (purple) and Aug-Oct TCs (green).

Given the previous analysis, one would think that the majority of TC variability as explained by the Joint Basin 1st PC comes from the Pacific Basin. However, a partial correlation analysis reveals that it is the Atlantic basin that explains ~30% of the variability in TC counts on an interannual time-scale. Table 1 shows the partial correlation of between the Joint Basin's EOF 1st PC and Atlantic TCs controlling for the Atlantic EOF's 1st PC (top left cell) and controlling for the Pacific EOF's 1st PC (top right cell). The results seem to indicate that both PCs explain the same amount of variability in annual Atlantic TC counts. However, when we de-trend the 1st PC and TC counts, we find that the Atlantic EOF's 1st PC explains roughly ~40% of the variability versus ~10% for the Pacific EOF's 1st PC.

	Corr(JointPC, TCs | AtlPC)	Corr(JointPC, TCs | PacPC)
Raw PCs	0.57	0.61
De-trended PCs	0.59	0.32

Table 1 - Partial linear correlation coefficients between the Joint Basin EOF's 1st PC and Atlantic TCs while controlling for the Atlantic Basin EOF's 1st PC (left) and the Pacific Basin EOF's 1st PC (right). Both PCs explain the same amount of variability in annual TC counts. However, when we de-trend the PCs and TC counts before running the analysis, the Atlantic EOF's1st PC explains roughly 40% of the variability versus 10% for the Pacific EOF's 1st PC.

Multivariate Regression Analysis

We ran a multivariate regression analysis using the principle components from the SST anomaly EOF as predictors and the Atlantic seasonal TC counts as target variables (N=32). For each basin, we select the first five PCs and feed them incrementally into an L1-regularized linear regression model (LASSO) where the weights of the coefficients are adjusted based on how well each predictor performs. To test each model, we perform a leave-k out cross-validation where k=1,2,4, and 8. Under such approach, a regression model is trained using N-k observations. The model is subsequently tested on the k hidden observations. We construct as many models as needed to test (*i.e.* hide and predict) all observations. We then report the linear correlation between observed and predicted Aug-Oct TC counts.

Tables 2-4 show the linear correlation of the linear regression models built for the Atlantic, Pacific, and Joint basins. Each row shows the linear correlation between observed and prediction Aug-Oct TCs as predicted by a linear regression model trained using the first PC (1 PC), first and second PCs (2 PC), first, second, and third PCs (3 PC) ... *etc*. The results show that the Pacific SST's first and second PCs have a strong correlation with Aug-Oct TCs (0.78). Unlike the other basins, no PC other than first contribute to much of the signal. In the Joint Basin, it is likely that the Pacific basin is dominating the signal. We also repeat the same experiment using de-trended PCs and TC counts.

	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.56	0.59	0.60	0.61
2 PCs	0.55	0.56	0.60	0.60
3 PCs	0.57	0.60	0.63	0.62
4 PCs	0.58	0.60	0.62	0.63
5 PCs	0.57	0.59	0.61	0.63

Table 2 - The linear correlation between observed and predicted Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five PCs of Aug-Oct Atlantic SST. The first PC explains most of the signal's variability, with additional PCs adding little information

	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.28	0.28	0.29	0.34
2 PCs	0.26	0.24	0.31	0.33
3 PCs	0.32	0.32	0.37	0.38
4 PCs	0.37	0.37	0.43	0.45
5 PCs	0.36	0.36	0.43	0.44

Table 2a - The linear correlation between observed and predicted de-trended Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five de-trended PCs of Aug-Oct Atlantic SST. The first PC explains most of the signal's variability, with additional PCs adding little information

	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.59	0.60	0.59	0.61
2 PCs	0.75	0.75	0.74	0.78
3 PCs	0.74	0.74	0.74	0.78
4 PCs	0.74	0.73	0.71	0.73
5 PCs	0.72	0.72	0.70	0.72

Table 3 - The linear correlation between observed and predicted Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five PCs of Aug-Oct Pacific SST. The first PC explains most of the signal's variability, with the second PC adding some information.

ASO Joint	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.59	0.59	0.59	0.61
2 PCs	0.60	0.60	0.61	0.64
3 PCs	0.61	0.60	0.62	0.65
4 PCs	0.60	0.59	0.58	0.59
5 PCs	0.57	0.56	0.51	0.54

Table 3a - The linear correlation between observed and predicted de-trended Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five de-trended PCs of Aug-Oct Pacific Basin SST. The first PC explains most of the signal's variability, with additional PCs adding little information.

ASO Joint	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.72	0.73	0.74	0.74
2 PCs	0.73	0.75	0.75	0.77
3 PCs	0.77	0.78	0.78	0.79
4 PCs	0.76	0.78	0.78	0.78
5 PCs	0.74	0.76	0.76	0.75

Table 4 - The linear correlation between observed and predicted Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five PCs of Aug-Oct Joint Basin SST. The first PC explains most of the signal's variability, with additional PCs adding little information.

ASO Joint	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
1 PC	0.63	0.64	0.64	0.64
2 PCs	0.63	0.64	0.65	0.66
3 PCs	0.66	0.67	0.68	0.69
4 PCs	0.65	0.66	0.65	0.66
5 PCs	0.63	0.64	0.61	0.64

Table 4a - The linear correlation between observed and predicted de-trended Aug-Oct TCs as predicted from a linear regression model trained on (up to) the first five PCs of de-trended Aug-Oct Joint Basin SST. The first PC explains most of the signal's variability, with additional PCs adding little information.

If we include all 96 PCs (32 from each basin) and let the model choose the best 10 and 5 predictors respectively we get the weights shown in Tables 5 and 6. If we force the model to select 1 or 2 predictors it always selects the Joint Basin 1st PC (even for 2 predictors).

Predictor	Atl 21st	Joint 29th	Atl 10th	Atl 8th	Joint 18th	Joint 1st	Atl 4th	Pac 28th	Pac 3rd	Pac 1st
Weight	-0.1194	0.0169	-0.0083	-0.0024	-0.0006	0.0006	-0.0004	-0.0002	-0.0001	0.0001

Table 4 - Predictors and associated weights for a model that selects the best 10 PCs to predict annual Atlantic TC counts. The majority of coefficients beyond the first one have marginal impact.

Predictor	Atl 21st	Joint 29th	Atl 10th	Joint 1st	Pac 1st
Weight	-0.0763	0.0083	-0.0011	0.0005	0.0001

Table 5 - Predictors and associated weights for a model that selects the best 5 PCs to predict annual Atlantic TC counts. The majority of coefficients beyond the first two have marginal impact.

When we include all PCs from the three basins and look at which combination of PCs provide the best performance we get no significant improvement over analyzing each basin's EOFs individually.

Based on the models returned from the selecting the best 10, 5, and 2 PCs we predict annual TC activity and correlate the predicted counts with the observed ones. The results summarized in Table 6 show that there isn't much information in including more than 1 or 2 PCs into the analysis. If we de-trend all PCs as well as annual TC counts, the model's performance drops significantly with the weights of the predictors highlighted above remaining virtually unchanged.

4

ALL PCs	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
Best 10 PCs	0.71	0.74	0.76	0.75
Best 5 PCs	0.72	0.75	0.76	0.77
Best 2 PCs	0.72	0.74	0.76	0.76

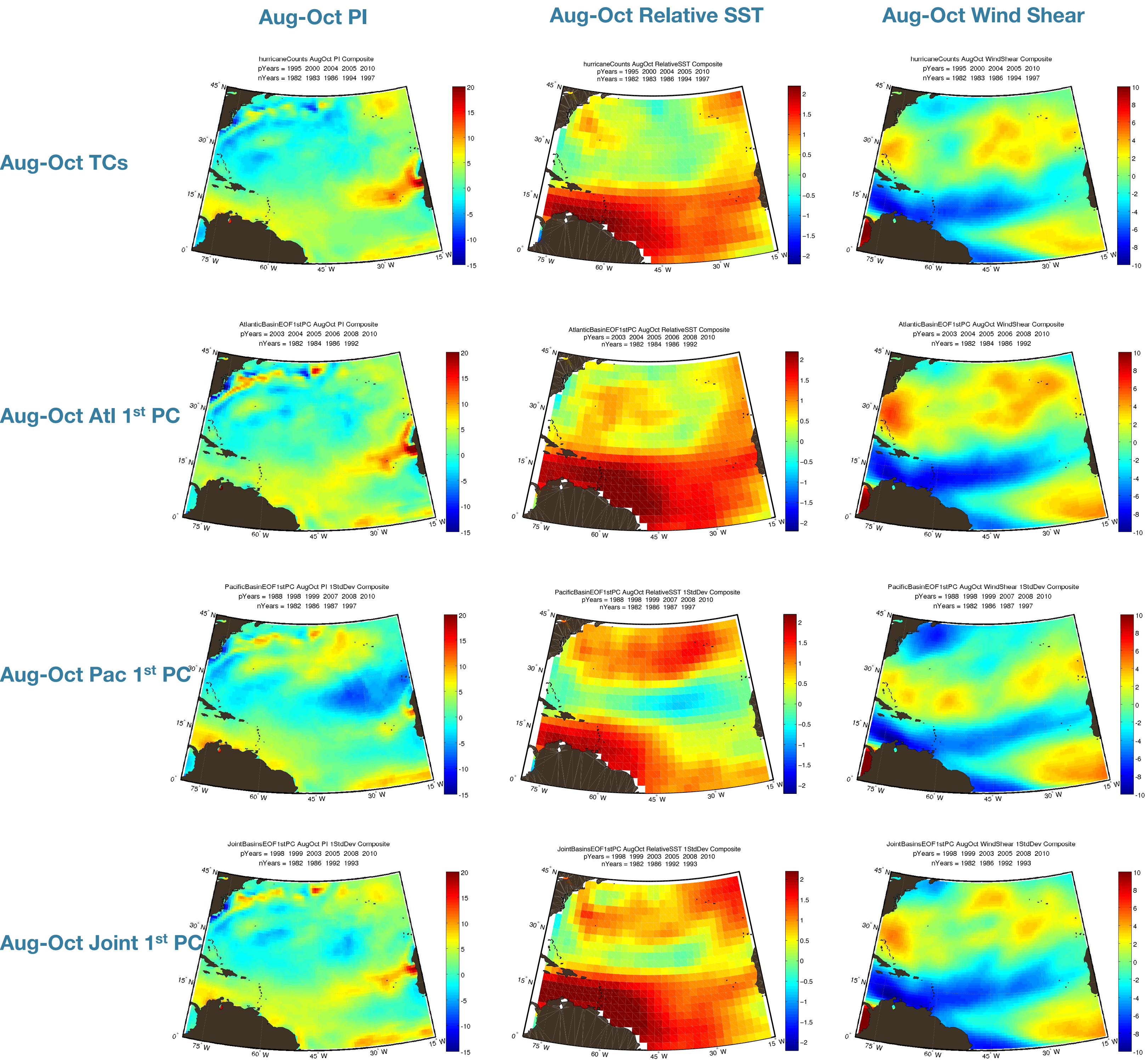
Table 6 - The linear correlation between observed and predicted Aug-Oct TCs as predicted from a linear regression model trained on the best 10, 5, and 2 PCs from all 96 PCs of Aug-Oct SST from the three basins analyzed.

ALL PCs	Leave 1 Out	Leave 2 Out	Leave 4 Out	Leave 8 Out
Best 10 PCs	0.55	0.53	0.62	0.53
Best 5 PCs	0.54	0.57	0.60	0.55
Best 2 PCs	0.54	0.55	0.59	0.57

Table 7 - The linear correlation between observed and predicted de-trended Aug-Oct TCs as predicted from a linear regression model trained on the best 10, 5, and 2 de-trended PCs from all 96 PCs of Aug-Oct SST from the three basins analyzed.

Composites Based on SST Anomalies EOF

We built the composites based on the Atlantic, Pacific, and Joint Basin EOF's 1st PC. Notice that both the Atlantic and Joint basin EOF's have composites closely resembling those of the TC count composites. The Pacific 1st PC however, doesn't capture the large scale conditions over the Atlantic as well as the two other PCs, this further confirms that solely monitoring the SST anomalies in the Pacific does not provide enough information on the Pacific's impact on the large scale conditions over the Atlantic that affect TC activity



Composites Based on Atlantic Aug-Oct SST Anomalies

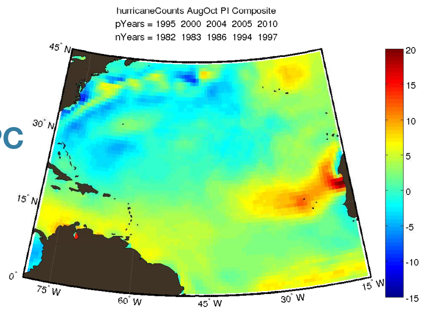
We built composites for a variety of SST regions in the Atlantic: (1) Aug-Oct Atlantic 1PC (same as previous page), (2) North Atlantic basin mean SST anomaly (0-40N, 18-80W); (3) Atlantic main development region (MDR) mean SST anomaly (5-25N 18-80W); and (4) The most dominate region from the Atlantic 1st PC (see appendix). From this analysis it is clear that the majority of the Aug-Oct Atlantic SST signal is related to the large scale conditions that affect TCs over the Atlantic.

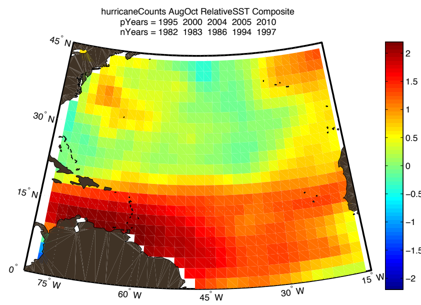
Aug-Oct PI

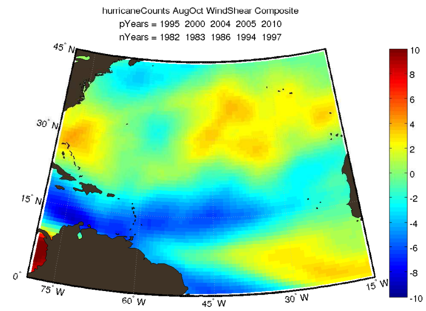
Aug-Oct Relative SST

Aug-Oct Wind Shear

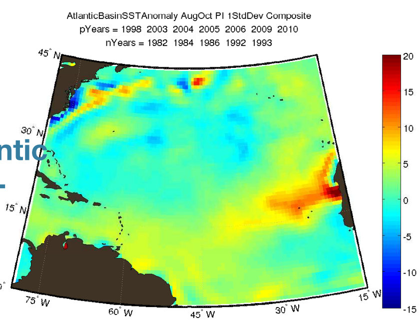
Aug-Oct Atl 1st PC

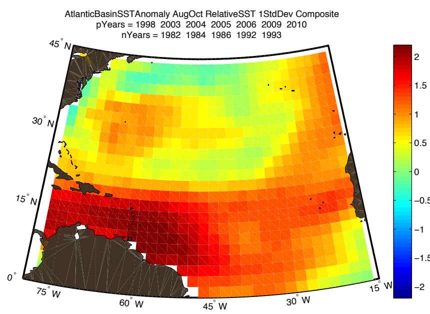


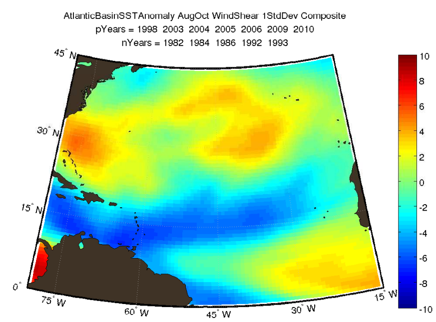




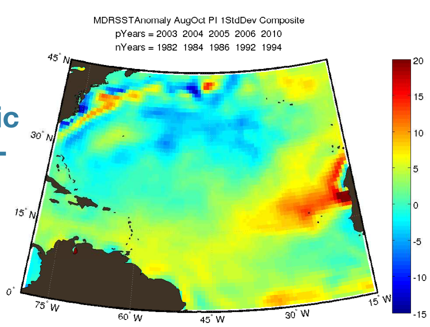
Aug-Oct N. Atlantic Basin mean SST Anomalies

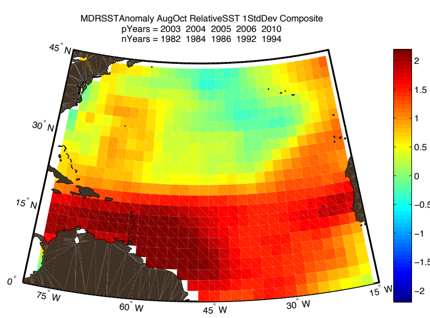


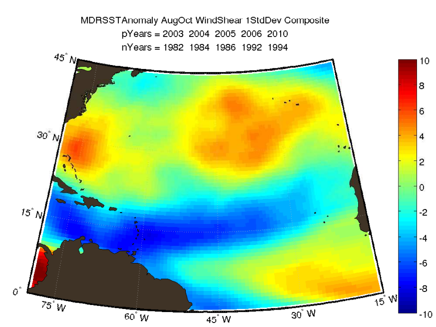




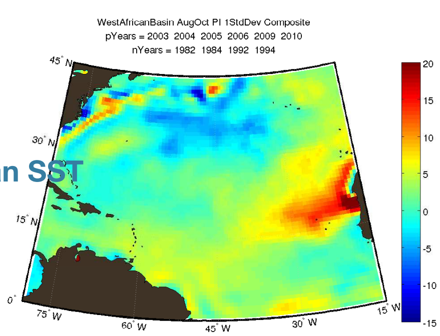
Aug-Oct Atlantic MDR mean SST Anomalies

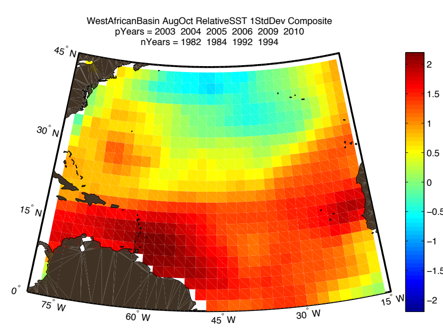


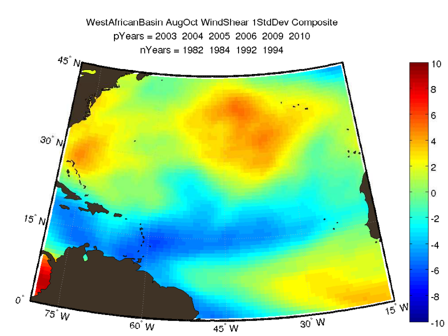




Aug-Oct W. Atlantic mean SST Anomalies



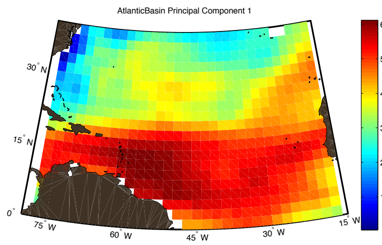




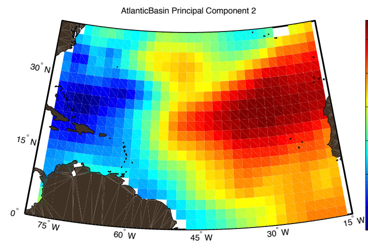
EOF Analysis

Appendix

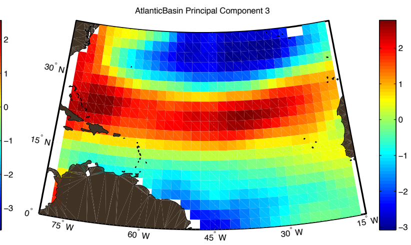
Aug-Oct Atlantic SST EOF Spatial Distribution



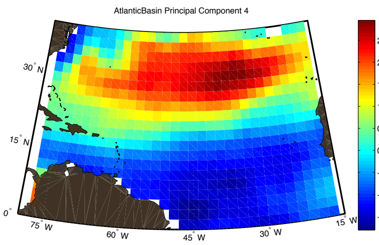
1st PC



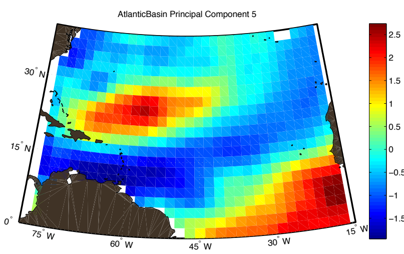
2nd PC



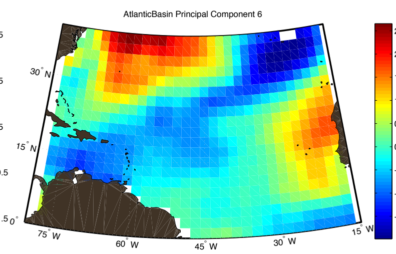
3rd PC



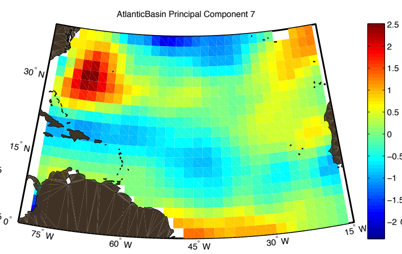
4th PC



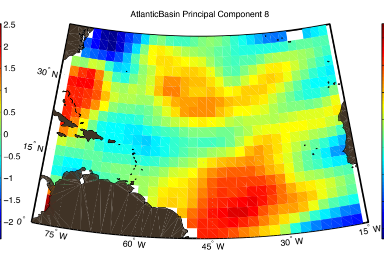
5th PC



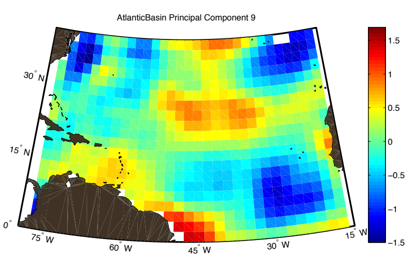
6th PC



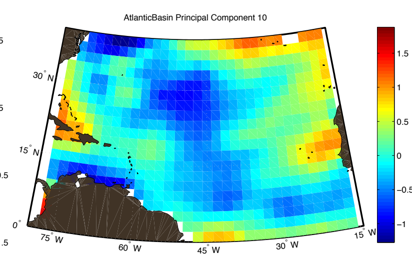
7th PC



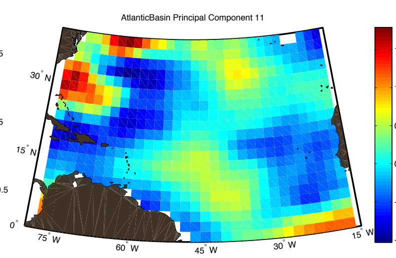
8th PC



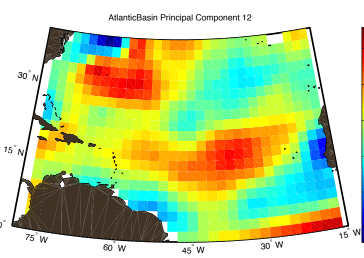
9th PC



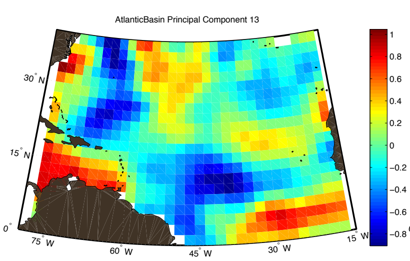
10th PC



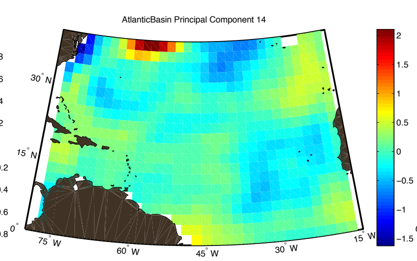
11th PC



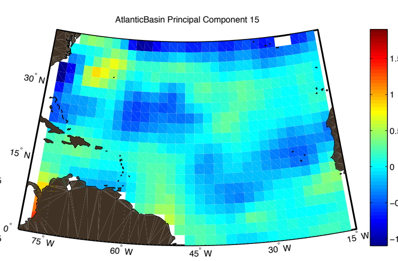
12th PC



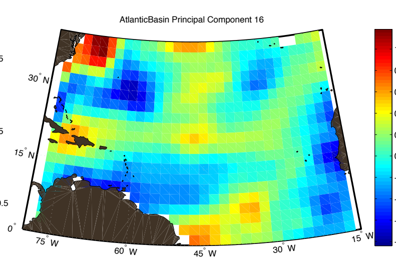
13th PC



14th PC

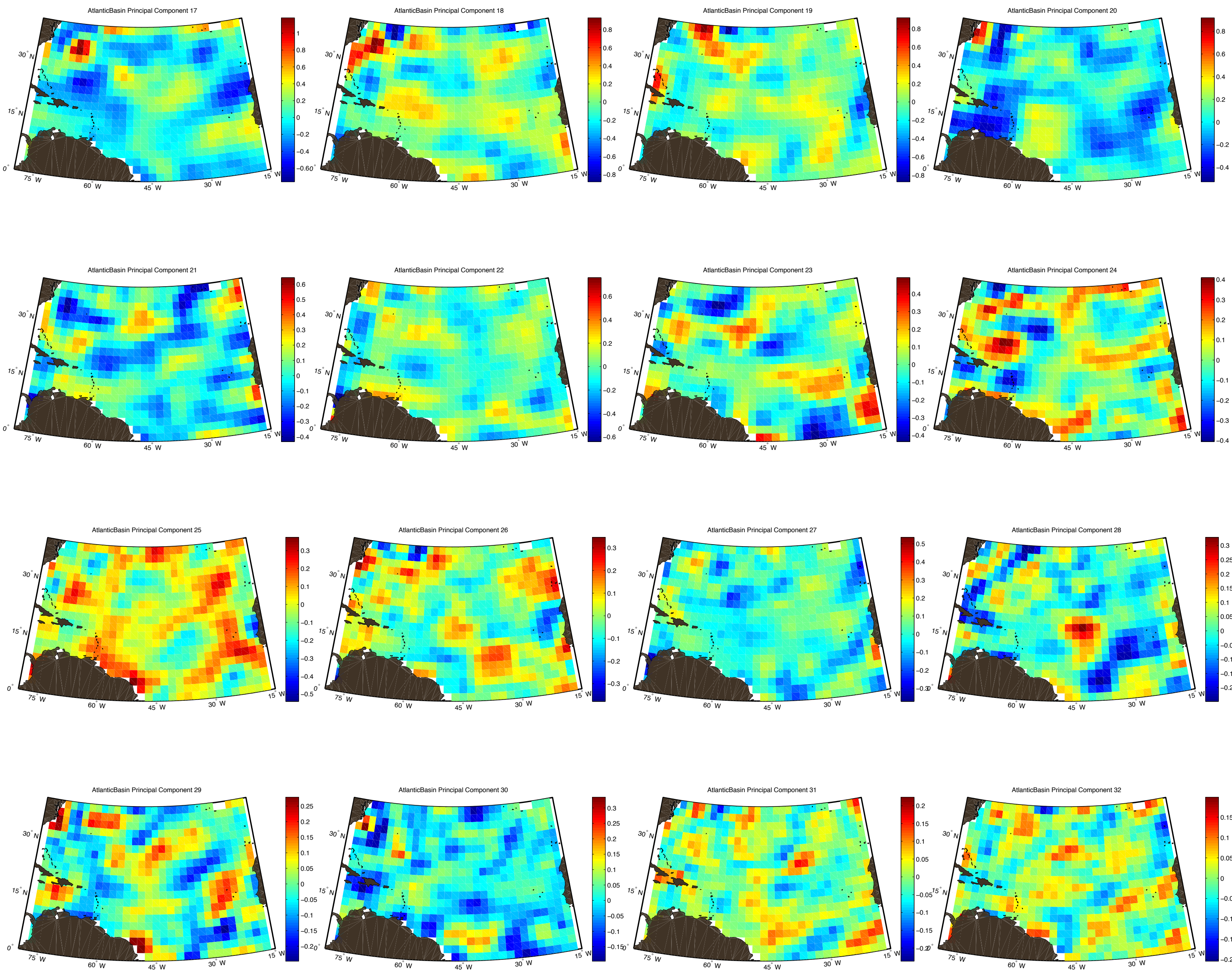


15th PC

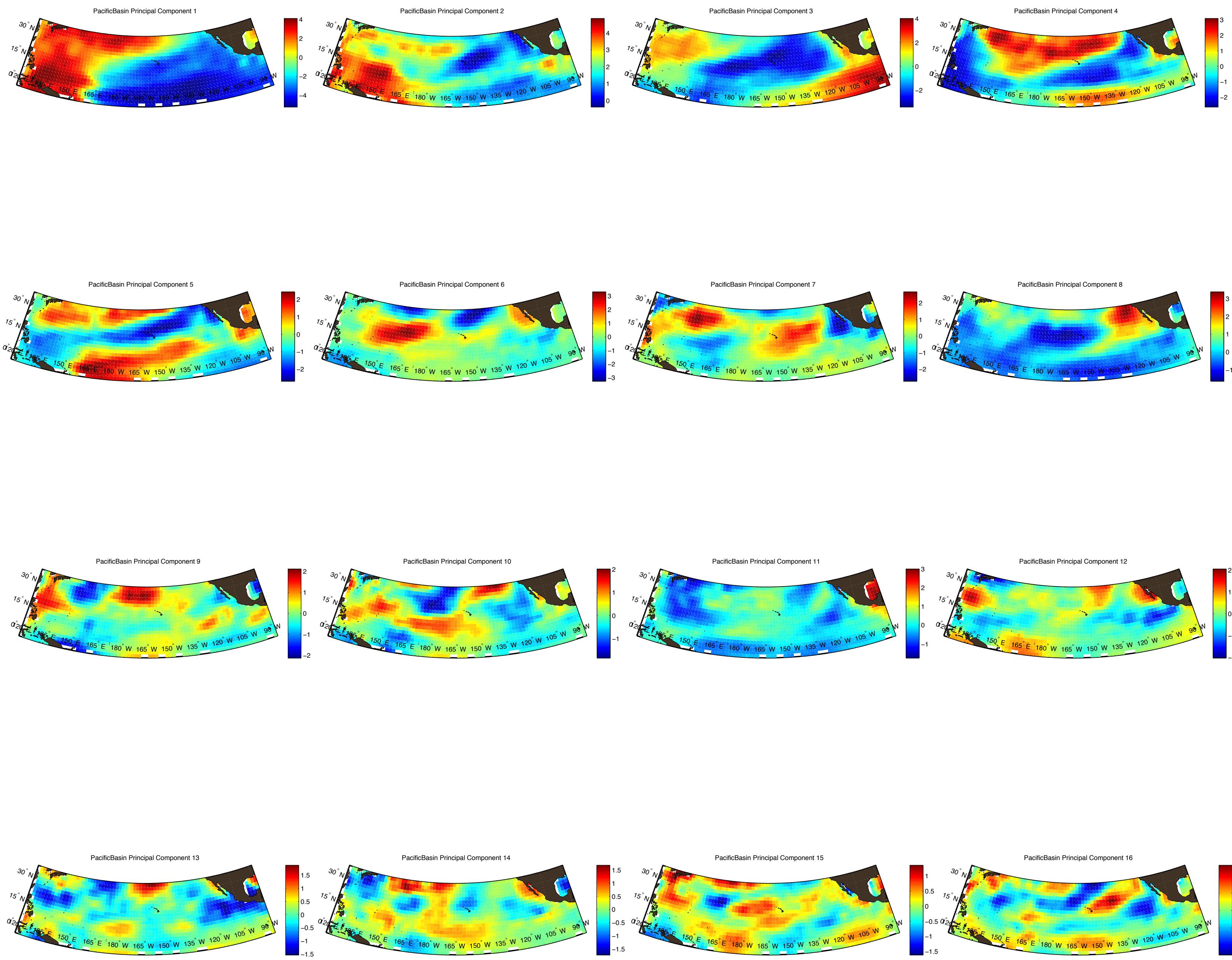


16th PC

Aug-Oct Atlantic SST EOF Spatial Distribution (Continued)



Aug-Oct Pacific SST EOF Spatial Distribution



Aug-Oct Pacific SST EOF Spatial Distribution (Continued)