Supplement: Understanding Dominant Factors for Precipitation over the Great Lakes Region

1 Covariates for Experiments

We consider three types of covariates for experiments: local and regional atmospheric variables, and large scale climate indices. For each climate index, we consider all 12 months' values previous to a particular season as covariates. For local and regional variables, we consider the winter and autumnal averages. The following table lists the covariates we considered for the experiments.

Table 1: Covariates for Precipitation prediction

Type	Variables
Atmospheric (Station level)	Winter Minimum Temperature (DJF_Tmin), Winter Maximum Temperature (DJF_Tmax), Autumn Minimum Temperature (SON_Tmin), Autumn Maximum Temperature (SON_Tmax), Sea Level Pressure (SLP), Convective Available Potential Energy (CAPE), Air Temperature at 500mb (AIR_500)
Atmospheric (Regional averages)	Regional Average Winter Minimum Temperature (DJF_TRegmin), Regional Average Winter Maximum Temperature (DJF_TRegmax), Regional Average Autumn Minimum Temperature (SON_TRegmin), Regional Average Autumn Maximum Temperature (SON_TRegmax), Regional Average Sea Level Pressure (SLPReg), Regional Average Convective Available Potential Energy (CAPEReg), Regional Average Air Temperature at 500mb (Reg_AIR_500)
Large-Scale Climate Indices	North Atlantic Oscillation (NAO), East Atlantic Pattern (EA), West Pacific Pattern (WP), East Pacific/North Pacific Pattern (EPNP), Pacific/North American Pattern (PNA), East Atlantic/West Russia Pattern (EAWR),Scandinavia Pattern (SCA), Tropical/Northern Hemisphere Pattern (TNH), Polar/Eurasia Pattern (POL), Pacific Transition Pattern (PT), Nino 1+2, Nino 3, Nino 3.4, Nino 4, Southern Oscillation Index (SOI), Pacific Decadal Oscillation (PDO), Northern Pacific Oscillation (NP), Tropical/Northern Atlantic Index (TNA), Tropical/Southern Atlantic Index (TSA), Western Hemisphere Warm Pool (WHWP)

2 Permutation Test & Stable Factors

As discussed in the main paper, the random permutation test enables testing the significance of each non-zero coefficient obtained from LASSO. Here we provide more details on how the permutation test behaves, and how the stable coefficients are selected.

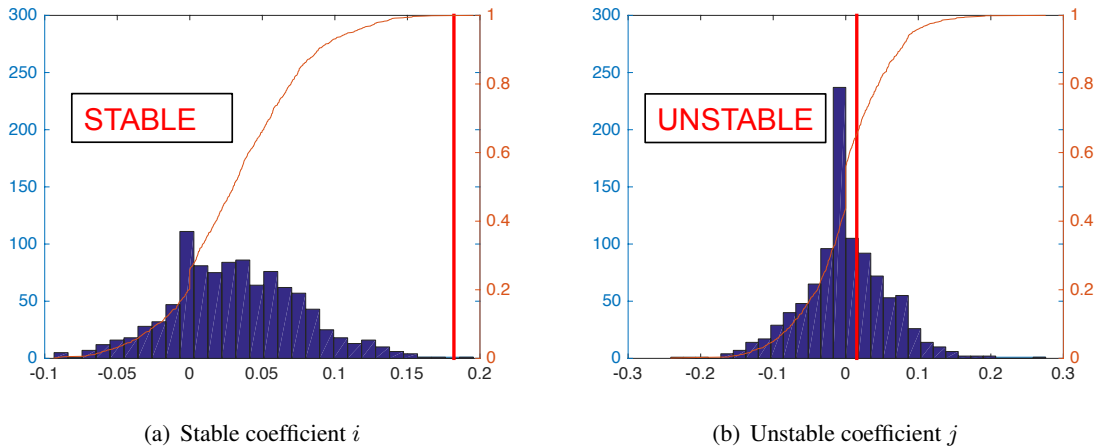(a) Stable coefficient i (b) Unstable coefficient j

Figure 1: Behavior of Stable and Unstable variables during random permutation test. The histogram represents an empirical approximation of the distribution of the coefficient value under the null hypothesis that **y** is exchangeable. A low p-value (a) shows that the estimated value lies to the tail of the distribution.

Fig. 1 illustrates the results of permutation test on two coefficients i and j. the coefficient i is stable since it lies at the tail of the empirical distribution and thus has low p-value (< 0.05) so that we can reject the null hypothesis that the estimated value occurred due to random chance. However, for coefficient j, the estimated value lies near the mode, and hence obtains a high p-value.

In Fig. 2, we plot the stable features that are selected using LASSO and permutation test in the training set. Evidently, increasing the regularization parameter λ in LASSO leads to pruning and we obtain a smaller set of stable parameters. However, note that the permutation test for each value of λ is independent, and therefore the pruning exhibited in Fig. 2 is a sign of stability of the selected features, rather than an artifact of the LASSO solution.
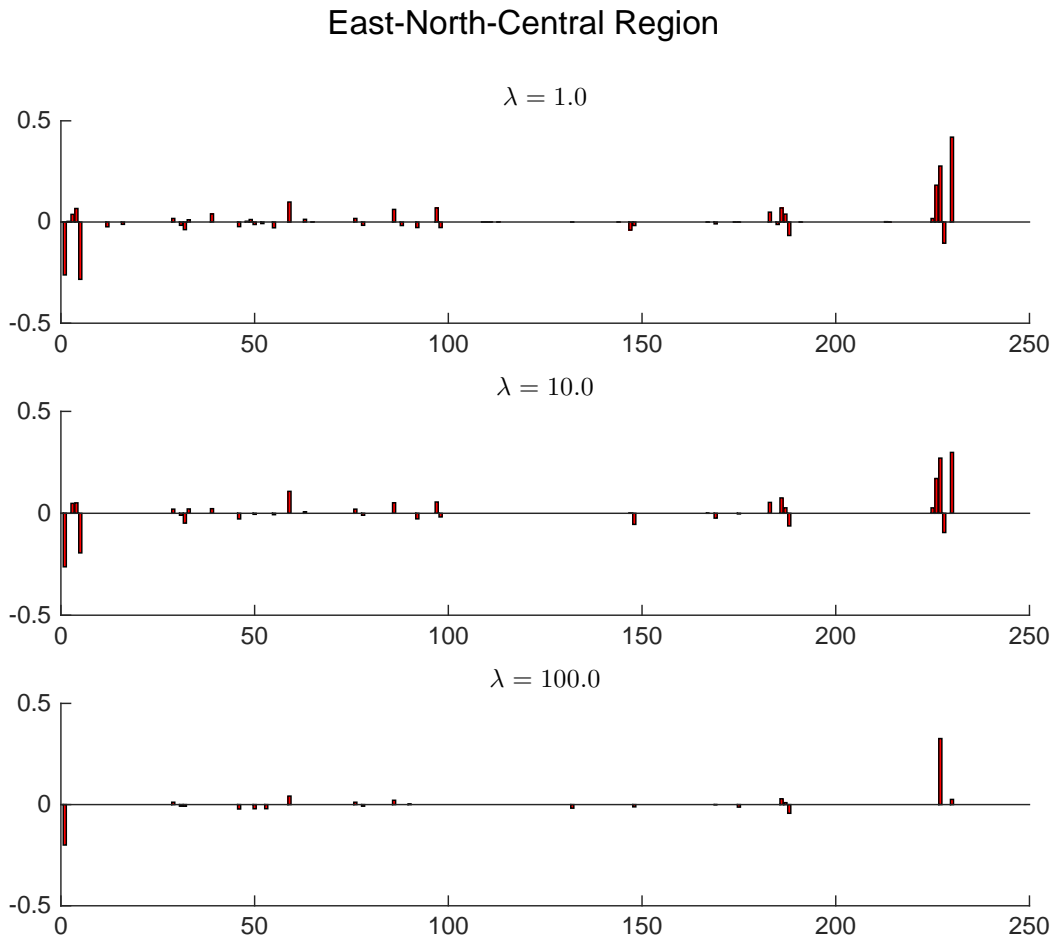
Figure 2: Stability of dominant factors at different penalization values. At higher penalization values, the set of coefficients is pruned, but no additional coefficients are introduced into the set.

3 Geopotential Height Anomalies

In Fig. 3, we plot the anomalies in geopotential height, similar to Fig. 8 in the main paper, but for averaged over the 10 lowest precipitation years in the ENC region. Note the high pressure system which moves from Siberia in September to North America across the Pacific Ocean. The high pressure system obstructs moisture flow into the Great Lakes, and also causes downdraft from upper atmosphere, thus reducing convection.

September: Low precip

October: Low precip

November: Low precip
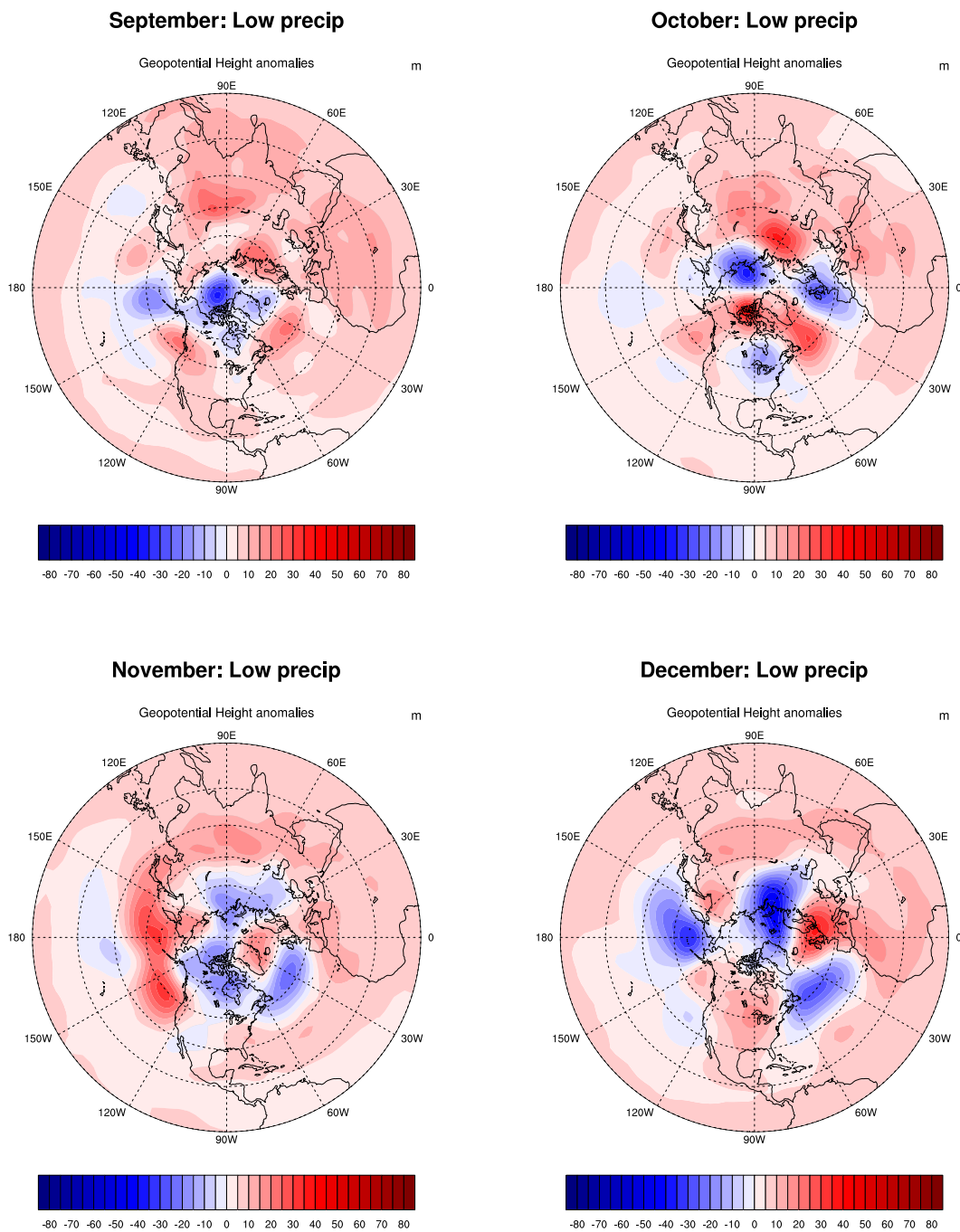

December: Low precip

Figure 3: Average Geopotential height anomalies over the 10 lowest precipitation years in ENC region for the months leading upto winter. Note the high pressure system over Siberia moves across the Pacific into North America. The Polar Pattern (POL) is a dominant factor for the ENC region and is closely related to this pressure system.