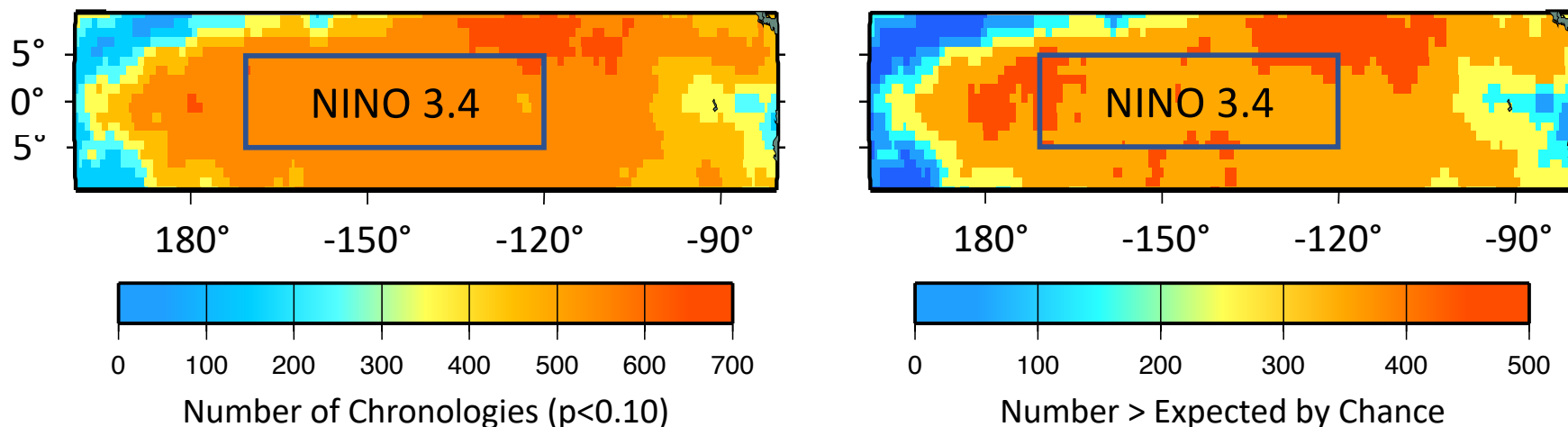


Point-by-Point Regression (PPR) was used for SST field reconstruction instead of less well resolved reduced-space space-time methods, like canonical regression, in order to preserve potentially meaningful smaller-scale spatial details of the SST field like that in the NINO 1+2 region. The calibration period in this example is 1920-1990. Autoregressive prewhitening was applied to the 1726 tree-ring candidate chronologies and SSTs at each grid point prior to regression. Only those tree-rings series correlated at the 2-tailed 90% level with SSTs were initially retained for calibration. That number ranged from 66 to 651, with a median of 547 per grid point in the NINO 3.4 region. *The El Niño rainfall teleconnection signal is very strong!* But with only 71 years for calibration, there is great potential for overfitting and including spurious predictors, with an expected number of retained chronologies by chance alone being ~ 170 for each of the 2397 grid points. The maps below illustrate this problem.



To reduce the amount of overfitting, only the m best correlated tree-ring chronologies equal to the number of calibration years ($n=71$) were used as predictors. Doing so creates a "fully saturated" regression design matrix. Choosing this "max" number is based on the fact that no unique regression solution exists if the number of variables exceeds the number of observations unless some regularization method is used. Here $m=71$ (\ll NINO 3.4 median 547) and we further reduce the degree of overfitting by using principal component regression (PCR), the regression method used in PPR.