Parsimonious patterns of sea surface temperature in the tropical Pacific Ocean

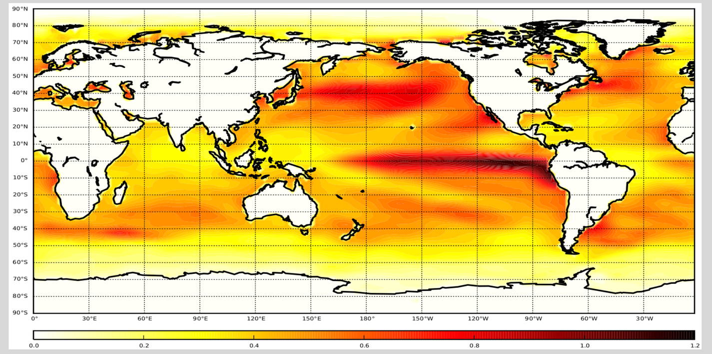
Guangoh Jheong and Gyu-Ho Lim

School of Earth and Environmental Sciences, College of Natural Science, Seoul National University, Seoul, Republic of Korea

Introduction

Motivating issue

: Principal component analysis (PCA) provides loading vectors full of nonzeros.

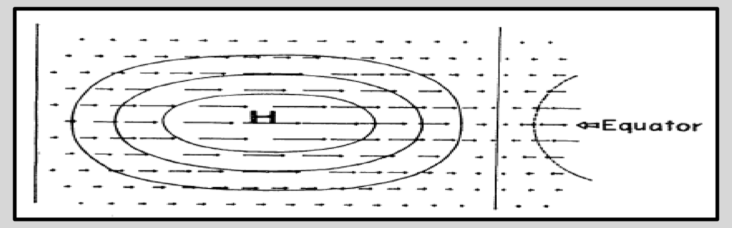


[Standard deviation of SST anomaly (1948 – 2014)]

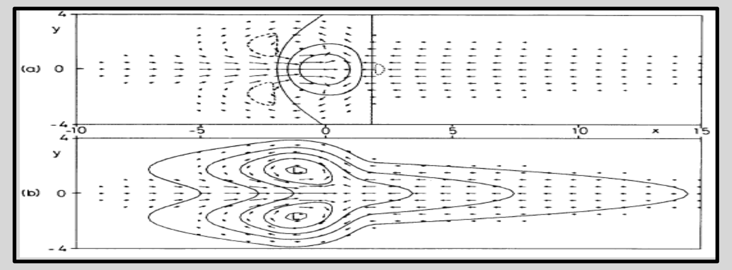
Are essentially required all the grid point values for explaining a particular climate variation?
(e.g. El Niño Southern Oscillation or North Atlantic Oscillation)

Theoretical backgrounds

Equatorial waves tend to stay close to the equator in latitudes.
[Matsuno, 1966; Gill, 1980]



Kelvin wave [Matsuno, 1966]



Atmospheric response to symmetric heating [Gill, 1980]

Conventional Methods

Generic PCA

PCA is a method of matrix decomposition to find a direction of maximal variance.

$$X = Z V^T$$

X: data matrix, Z: principal components (PCs), and
V: loading (eigen) vectors

Rotated PCA (RPCA)

Rotated PCA (RPCA) forces loading towards zero but not exact-zeros.

However, RPCA depends on several subjective factors
[Jolliffe, 2002; Hanachi et al, 2006]

- Selection of loading vectors (how many?)
- Choice of rotation criteria (e.g. varimax, quartimax, 17 other criteria)
- Decision of what to normalize (e.g PC or loading)

New Approach

Sparse PCA (SPCA)

SPCA is a PCA with a L1 regularization imposed on each eigenvector, so that the eigenvector can have as many exact-zeros as possible at the least expense of the explained variance loss.

Merits
- It can extract non-overlapping independent patterns successively.
- It manifest the center of action and a region of substantial variability.

Disadvantage
- It requires a great amount of computational resources compared with PCA and RPCA due to the determination of optimal sparsity.

Also (S)PCA mode can be seen a solution to the regression form with a regularization term as follows :

$$\tilde{v} = \underset{v}{argmin}\{\|X - Z V^T\|_F^2 + \lambda \|v\|_1\}$$

(i) generic PCA : $\lambda = 0$
(ii) SPCA : $\lambda \neq 0$

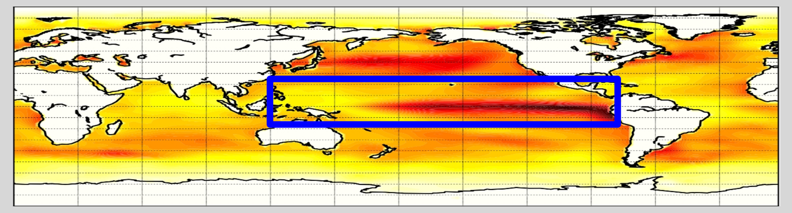
How should the sparse parameter (λ) be determined in an objective manner?

⟹ Based on Cross-validation, Bayesian Information Criteria (BIC), and Rate of loss information

Mode number	Selection criteria: RMSECV	BIC	ROIL	Average
1	40	30	60	40
2	60	50	50	50
3	60	50	50	50
4	80	70	70	70
5	90	70	80	80

Data

- Domain : 20 °S ~ 20 °N, 100 °E ~ 90 °W



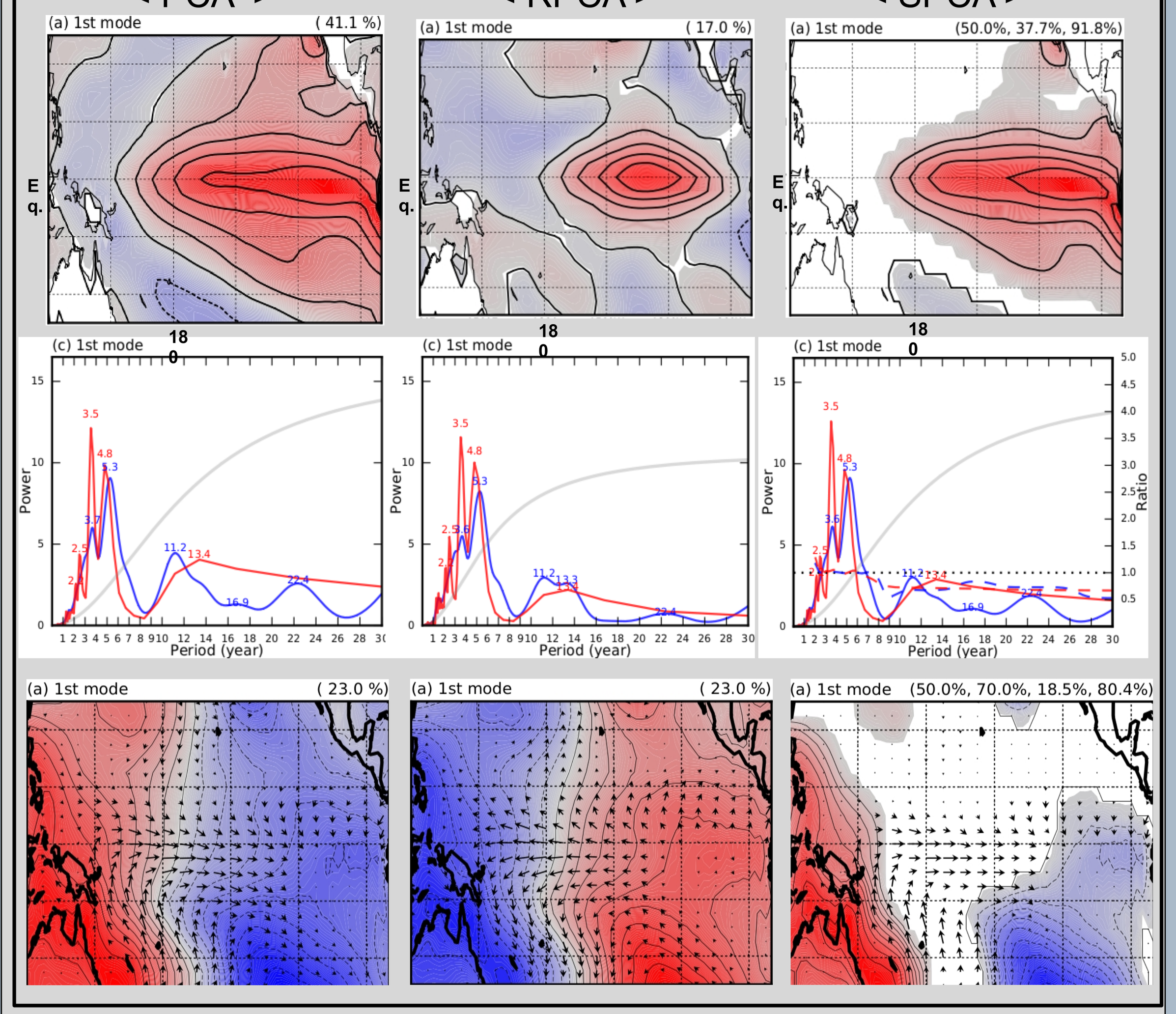
- Period : January 1948 ~ December 2014
- Variables : Monthly mean values
-- Sea surface temperature (SST)
from Centennial in situ Observation-Based Estimate (COBE)
-- Mean sea level pressure (MSLP) and surface winds
from National Center for Environmental Prediction (NCEP) reanalysis
- Preprocess : removal of calendar month mean ⟹ anomaly fields

Results

First mode of SST

(ENSO) [Rasmusson and Carpenter, 1982]
(3.5- and 5-year peaks) [White and Tourre, 2003]

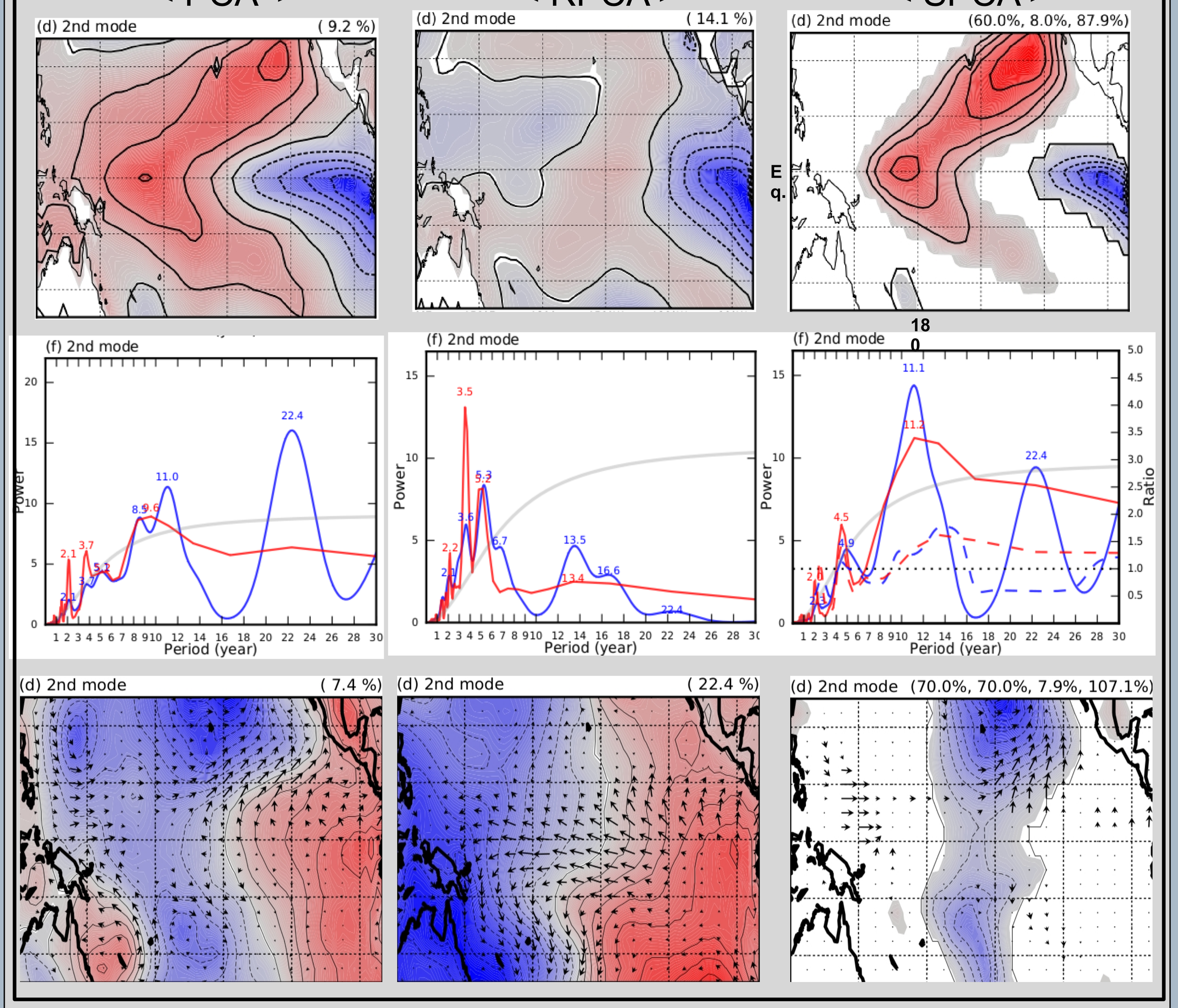
< PCA > < RPCA > < SPCA >



Second mode of SST

(ENSO Modoki / CP-type ENSO)
(4- and 12-year peaks) [Ashok et al, 2007]

< PCA > < RPCA > < SPCA >

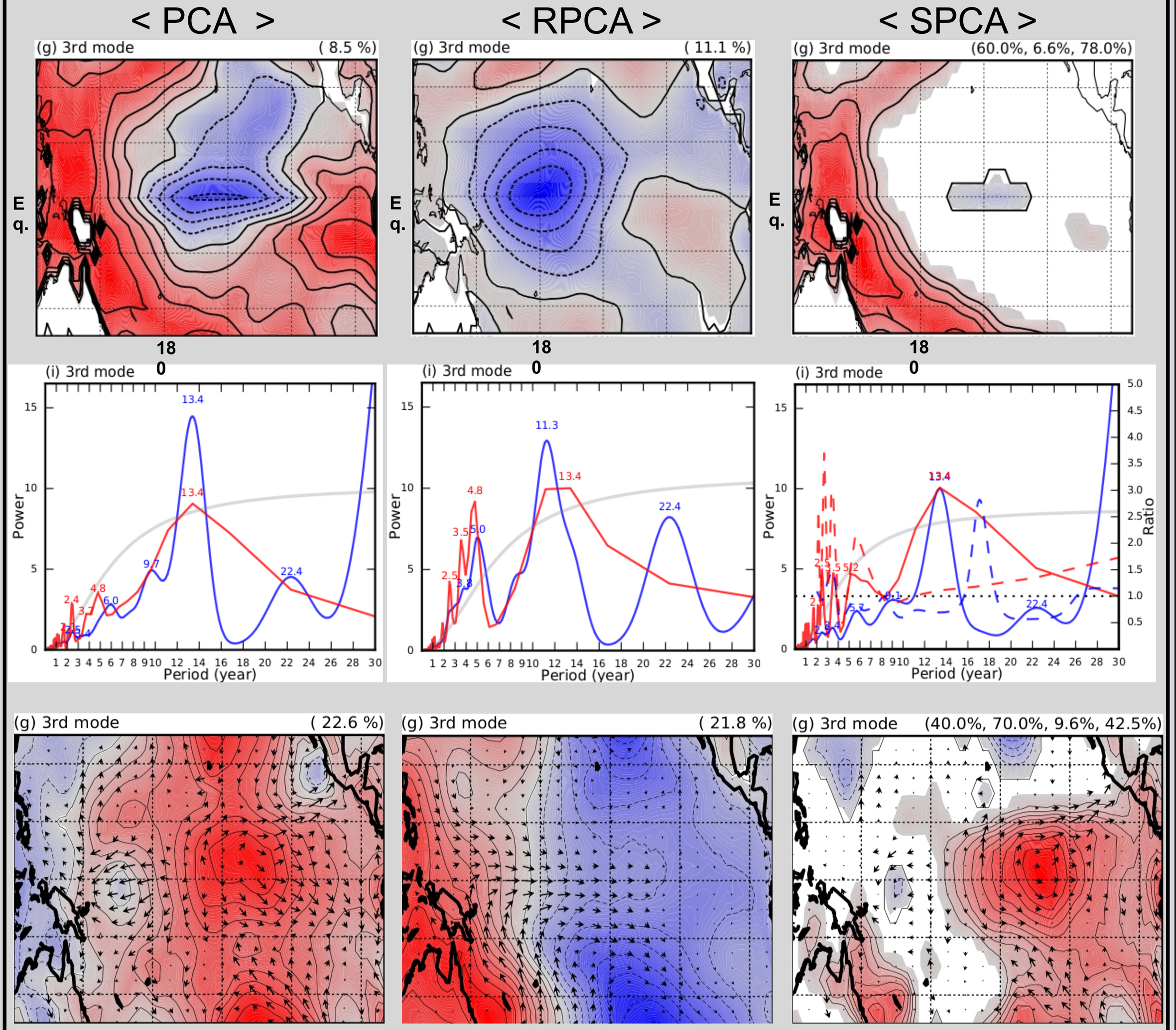


Results

Third mode of SST

(Western Pacific Warm Pool)
(10 ~ 15-year peaks) [Kug et al., 2009]

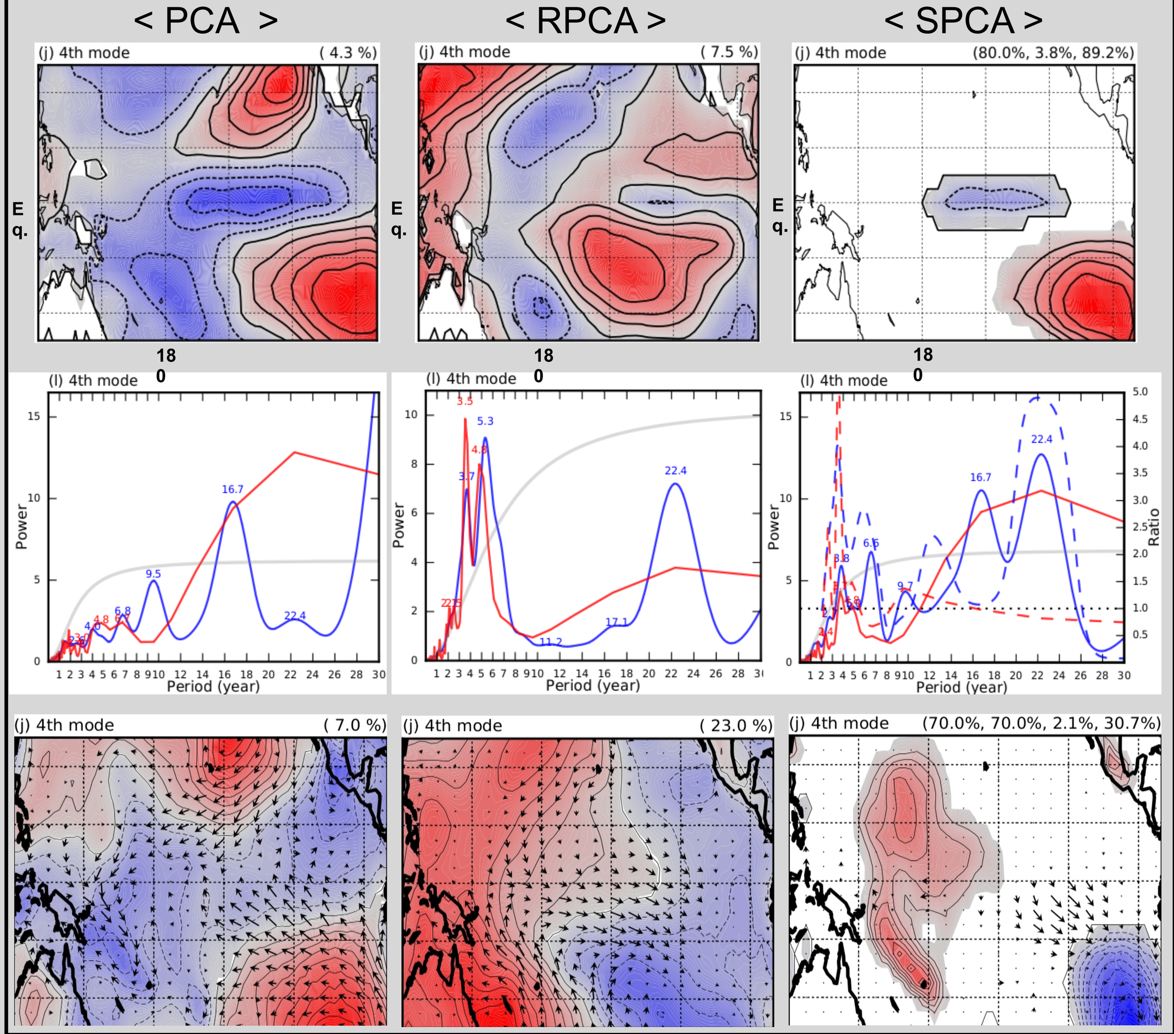
< PCA > < RPCA > < SPCA >



Fourth mode of SST

(South Pacific Meridional Mode)
(no preferred periodicity) [Chiang and Vimont, 2004]

< PCA > < RPCA > < SPCA >

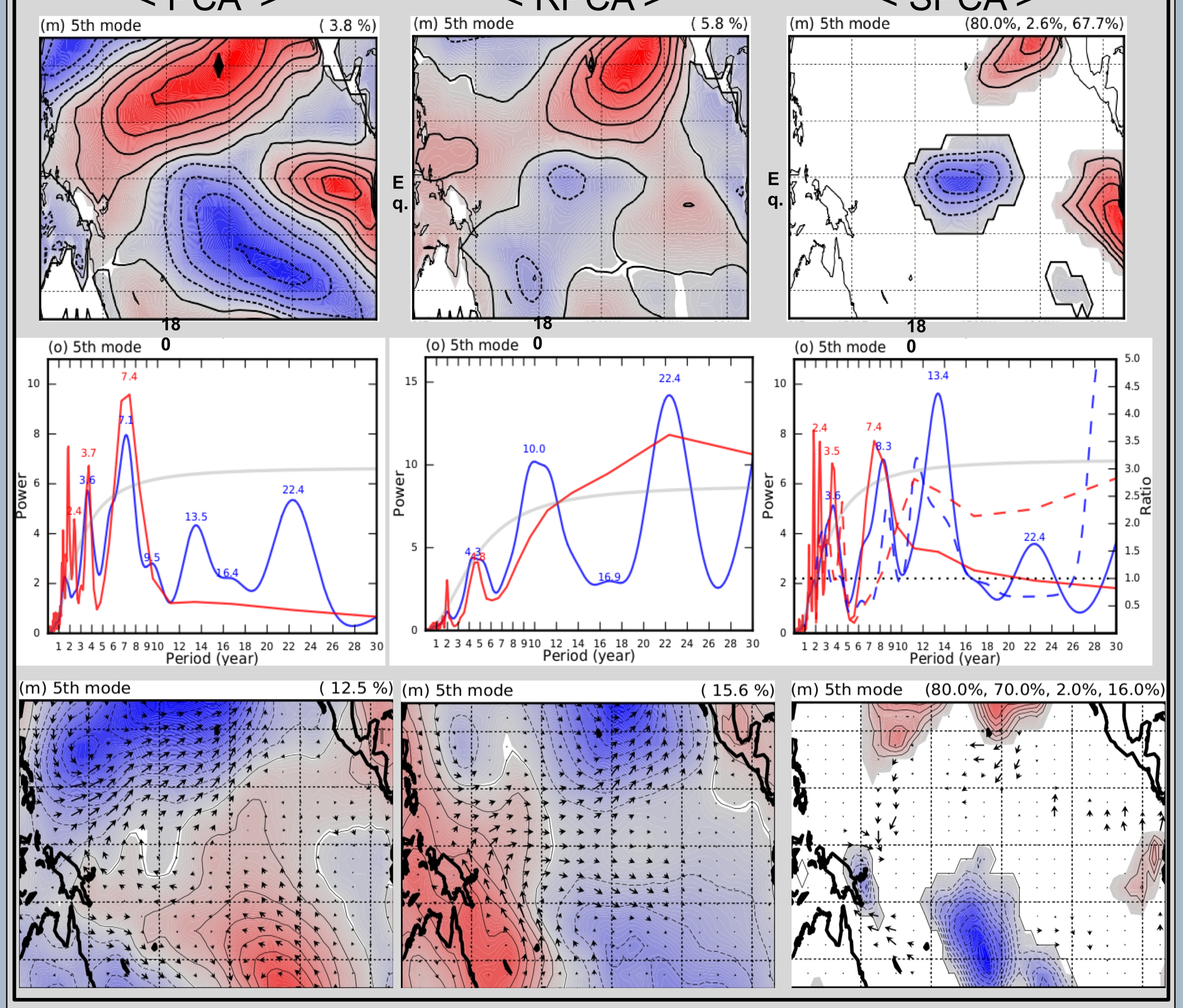


Results

Fifth mode of SST

(North Pacific Meridional Mode)
(no preferred periodicity) [Zhang et al, 2014]

< PCA > < RPCA > < SPCA >



Significant signatures and periodicities

Mode	PCA	RPCA	SPCA
First	3.5 and 4.8 ENSO	3.5 and 4.8 ENSO	3.5 and 4.8 ENSO
Second	22.4, 11.0, and 8.5 ENSO Modoki	3.5 and 5.2 ENSO	11.1 and 4.9 ENSO Modoki
Third	13.4 WPWP + CP ENSO	11.3, 5.0, and 3.8 CP ENSO	13.4 WPWP
Fourth	16.7 SPMM + NPMM	5.3 and 3.7 Unidentifiable	22.2 and 16.9 SPMM
Fifth	7.1 and 3.6 SPMM + NPMM	22.4 and 10.0 NPMM	13.4 and 8.6 NPMM

Summary and Conclusion

- The canonical ENSO has an unimodal center of action in the eastern equatorial Pacific.
- The ENSO modoki mode is a hemispherically asymmetric oscillation in SST of 11-year period.
- The western Pacific warm pool (WPWP) may be separated from the canonical ENSO and the ENSO modoki in terms of spatial patterns and temporal variations.
- The South Pacific meridional mode (SPMM) has a regionalized oscillation centered on the subtropical eastern Pacific with an inter-decadal periodicity.
- The North Pacific meridional mode (NPMM) is symmetric in space but different in periodicity, which is a different signature from the SPMM.

➡ We expect to be able to discover new spatio-temporal structures in climate variation by applying SPCA to the upper-air dataset.