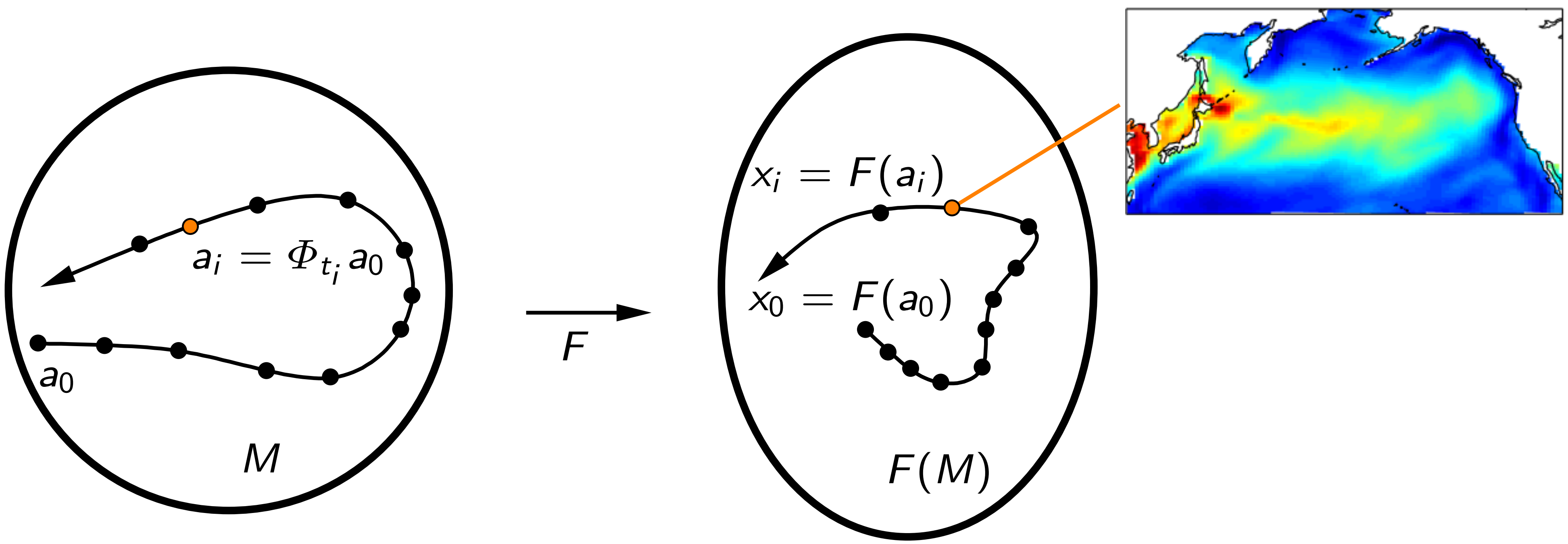
Recovering spatiotemporal modes of variability in the ocean and atmosphere using kernel and Koopman operators

Joanna Slawinska[1] & Dimitris Giannakis[2]

[1]Center for Environmental Prediction, Rutgers University

[2]Courant Institute of Mathematical Sciences, NYU

Setting



Ergodic dynamical system (M, Φ_t, μ) observed through a vector-valued function $F : M \mapsto \mathbb{R}^n$

Given time-ordered observations $\{x_0, \ldots, x_{N-1}\}$ with $x_i = F(a_i)$ and $n \gg 1$, we seek to compute approximate Koopman eigenfunctions with high smoothness

Kernel methods for Koopman eigenfunctions (Giannakis 2015, Giannakis et al. 2015)

We formulate a Galerkin method for the Koopman eigenvalue problem in a data-driven basis of $L^2(M, \mu)$ with a well-defined notion of smoothness (Dirichlet energy)

- Takens delay embeddings

$$F_s : M \to \mathbb{R}^{sd}, \quad F_s(a) = X = (F(a), F(\Phi_{-\delta t} a), \ldots, F(\Phi_{-(s-1)\,\delta t} a))$$

- Variable-bandwidth kernel

$$K_{\epsilon,s} : M \times M \to \mathbb{R}_+, \quad K_{\epsilon,s}(a_i, a_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{\epsilon \hat{\sigma}_{\epsilon,s}^{-1/m}(X_i)\hat{\sigma}_{\epsilon,s}^{-1/m}(X_j)}\right)$$

- Diffusion maps normalization

$$\tilde{K}_{\epsilon,s} = \frac{K_{\epsilon,s}(a_i, a_j)}{\sum_k K_{\epsilon,s}(a_j, a_k)}, \quad P_{\epsilon,s}(a_i, a_j) = \frac{\tilde{K}_{\epsilon,s}(a_i, a_j)}{\sum_k \tilde{K}_{\varepsilon,s}(a_i, a_k)}$$

Galerkin method for the Koopman eigenvalue problem

- We eliminate rough eigenfunctions by solving the eigenvalue problem for $L_\epsilon = v + \epsilon\Delta$

Continuous problem. Find $z \in H^1(M, h)$ and $\lambda \in \mathbb{C}$ s.t.

$$\langle \psi, v(z) \rangle + \epsilon \langle \operatorname{grad}_h \psi, \operatorname{grad}_h z \rangle = \lambda \langle \psi, z \rangle, \quad \forall \psi \in H^1(M, h)$$

Discrete approximation

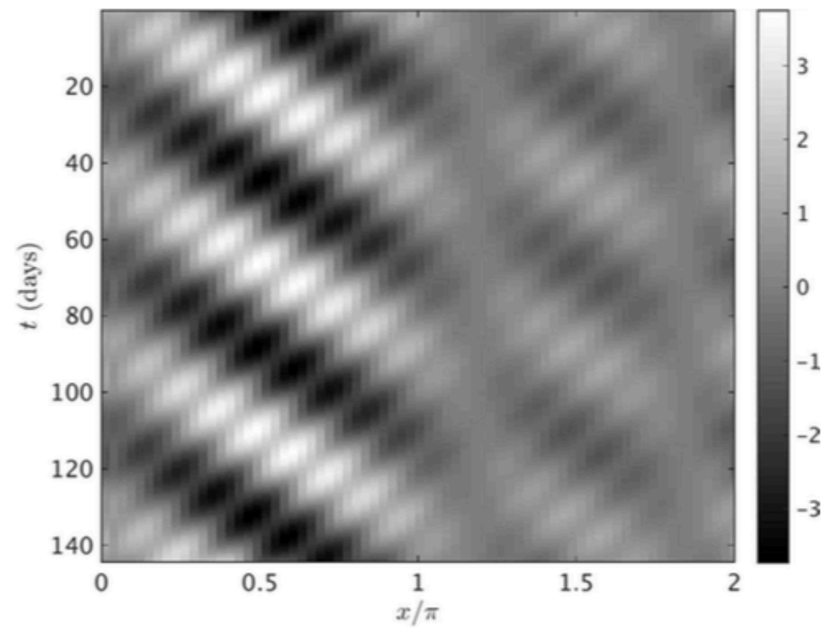
$$\sum_{j=0}^{l} (v_{ij} - \epsilon \delta_{ij}) c_j = \lambda \eta_i^{-1} c_i, \quad z = \sum_{i=0}^{l} c_i \varphi_i,$$

$$\varphi_i = \frac{\phi_i}{\eta_i^{1/2}}, \quad v_{ij} = \frac{1}{N} \sum_{k=0}^{N-1} \varphi_i(a_k) \frac{\varphi_j(a_{k+1}) - \varphi_j(a_{k-1})}{2\,\delta t}$$

- The action of v on functions is approximated via **finite differences** in time
- By construction of the $\{\varphi_i\}$ basis, $\langle \operatorname{grad}_h \varphi_i, \operatorname{grad}_h \varphi_j \rangle = \delta_{ij}$, and the scheme remains well-conditioned at large l

Traveling waves in synthetic dataset

$$u_{x,t} = (0.5 + \sin(x))[2\cos(k_1 x - \Theta_1(t)) + 0.5\cos(k_2 x - \Theta_2(t))]$$

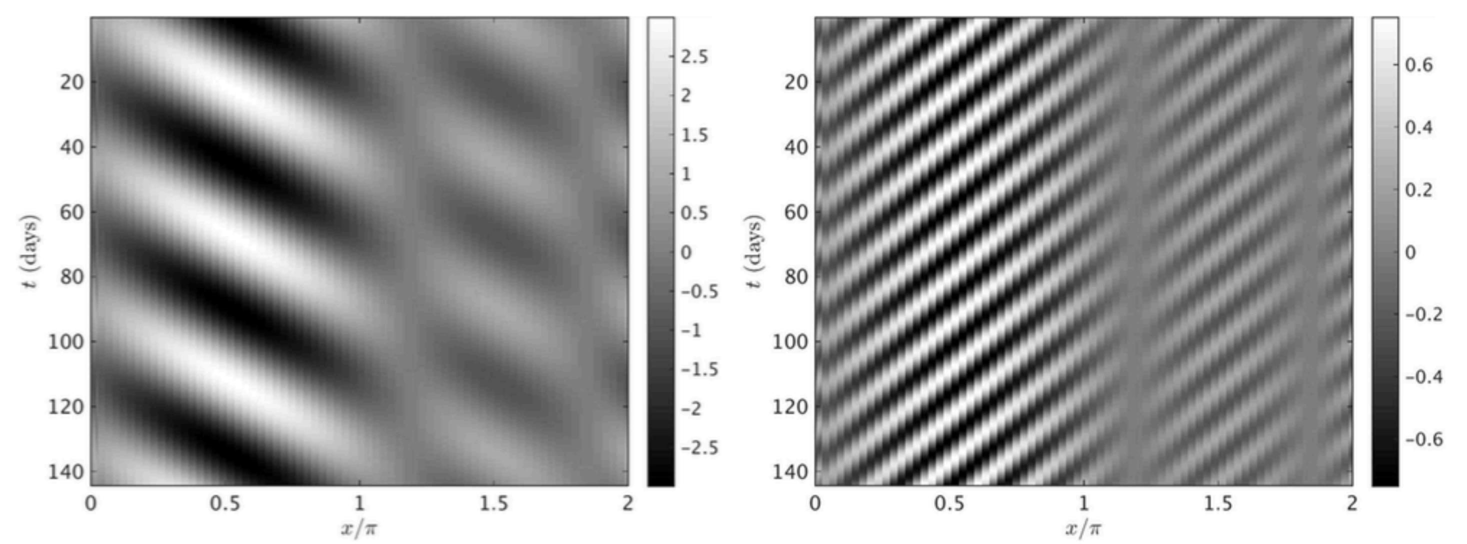


Kikuchi and Wang (2010)

$k_1 = 2, k_2 = 10$

$\Theta_i(t) = \omega_i t$

$\omega_1 = \frac{2\pi}{45}, \omega_2 = \sqrt{10}\omega_1$



Reconstructed data

Diffusion maps and Koopman eigenfunctions

In systems with pure point spectra of the Koopman operators, $\bar{\Delta} f = lim_{s \to \infty} \Delta_s f$ is well defined, and $[\bar{\Delta}, v] = 0$ (Giannakis 2015)

Eigenfunctions of $\bar{\Delta}$ provide an efficient approximation space for eigenfunctions of v

Timescale separation

Eigenfunctions recovered by diffusion maps with delays embeddings have timescale separation (Giannakis & Majda 2012, Berry et al. 2013)

Removing i.i.d. measurement noise

For data $\tilde{x}_i = x_i + \xi_i$, $x_i = F(a_i)$, corrupted by i.i.d. noise ξ_i,

$$\mathbb{E}\|\tilde{x}_i - \tilde{x}_j\|^2 = \|x_i - x_j\|^2 + 2R^2, \quad R^2 = \operatorname{var}\xi_i$$

Performing delay-coordinate maps, and taking the limit $s \to \infty$,

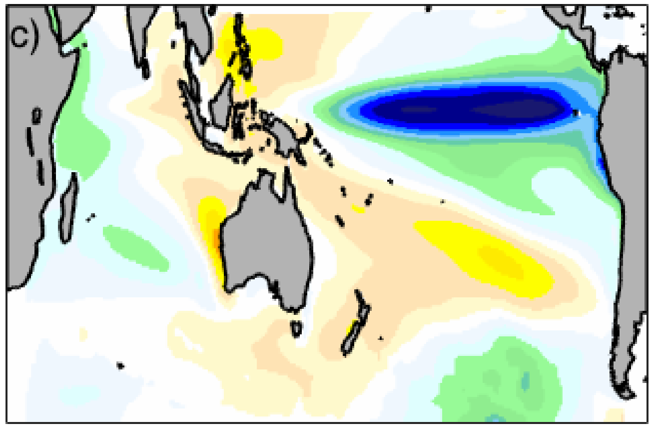
- $\|\tilde{X}_i - \tilde{X}_j\|^2 \xrightarrow{\text{a.s.}} \|X_i - X_j\|^2 + 2R^2$
- Because $\bar{g}$ is flat, the bias term in the pairwise distance produces a multiplicative bias in the kernel which cancels to $O(\epsilon^2)$ in the **diffusion maps normalization**

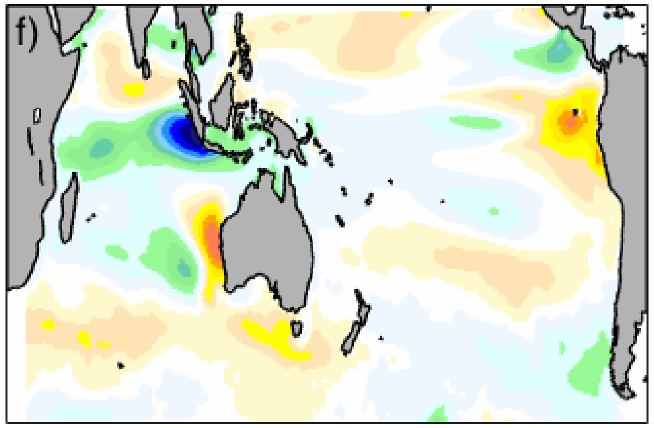
$$\tilde{K}_\epsilon(a_i, a_j) = \exp\left(-\frac{\|\tilde{X}_i - \tilde{X}_j\|^2}{\epsilon\sigma_{s,\epsilon}^{-1/m}(\tilde{X}_i)\sigma_{s,\epsilon}^{-1/m}(\tilde{X}_j)}\right) \to e^{-\frac{2R^2}{\epsilon(C^2+O(\epsilon^2))}} K_\epsilon(a_i, a_j),$$

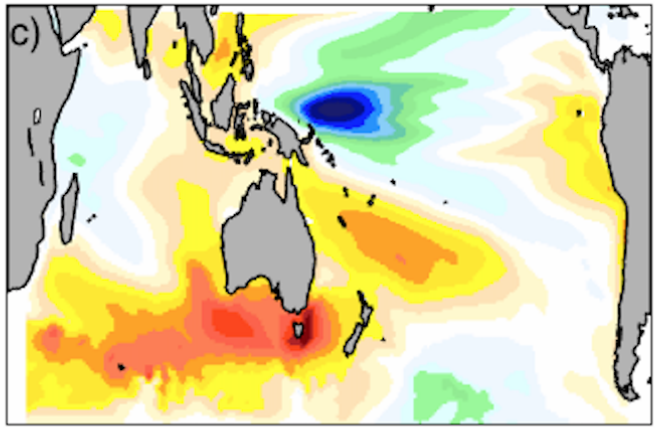
$$\sum_j \tilde{P}_\epsilon(a_i, a_j) f(a_j) = \sum_j \frac{\frac{\tilde{K}_\epsilon(a_i,a_j)}{\sum_k \tilde{K}_\epsilon(a_j,a_k)}}{\sum_l \frac{\tilde{K}_\epsilon(a_i,a_l)}{\sum_k \tilde{K}_\epsilon(a_l,a_k)}} f(a_j) \to \int_M P_\epsilon(a_i, b) f(b)\, d\mu + O(\epsilon^2)$$

- Since $\int_M P_\epsilon(a_i, b) f(b)\, d\mu = f(a_i) - \epsilon\Delta_{\bar{g}} f(a_i) + O(\epsilon^2)$, this bias does not affect the convergence of the eigenfunctions of P_ϵ to the eigenfunctions ϕ_k of $\Delta_{\bar{g}}$ as $\epsilon \to 0$, and the denoised ϕ_k can be employed in the Galerkin scheme for Koopman eigenfunctions

Indo-Pacific SST datasets







CCSM4 control
1300 y monthly-averaged SST, 1° (nominal), preindustrial forcings

GFDL CM3 control
800 y monthly-averaged SST, 1° (nominal), preindustrial forcings

HadISST
Industrial era (1870–2013) SST, 1°
Satellite era (1979–2013) SST, 1°

- No bandpass filtering or detrending performed
- Results checked for robustness against embedding windows 4–30 y, changes in spatial domain, addition of observational noise

Indo-Pacific SST modes recovered by NLSA

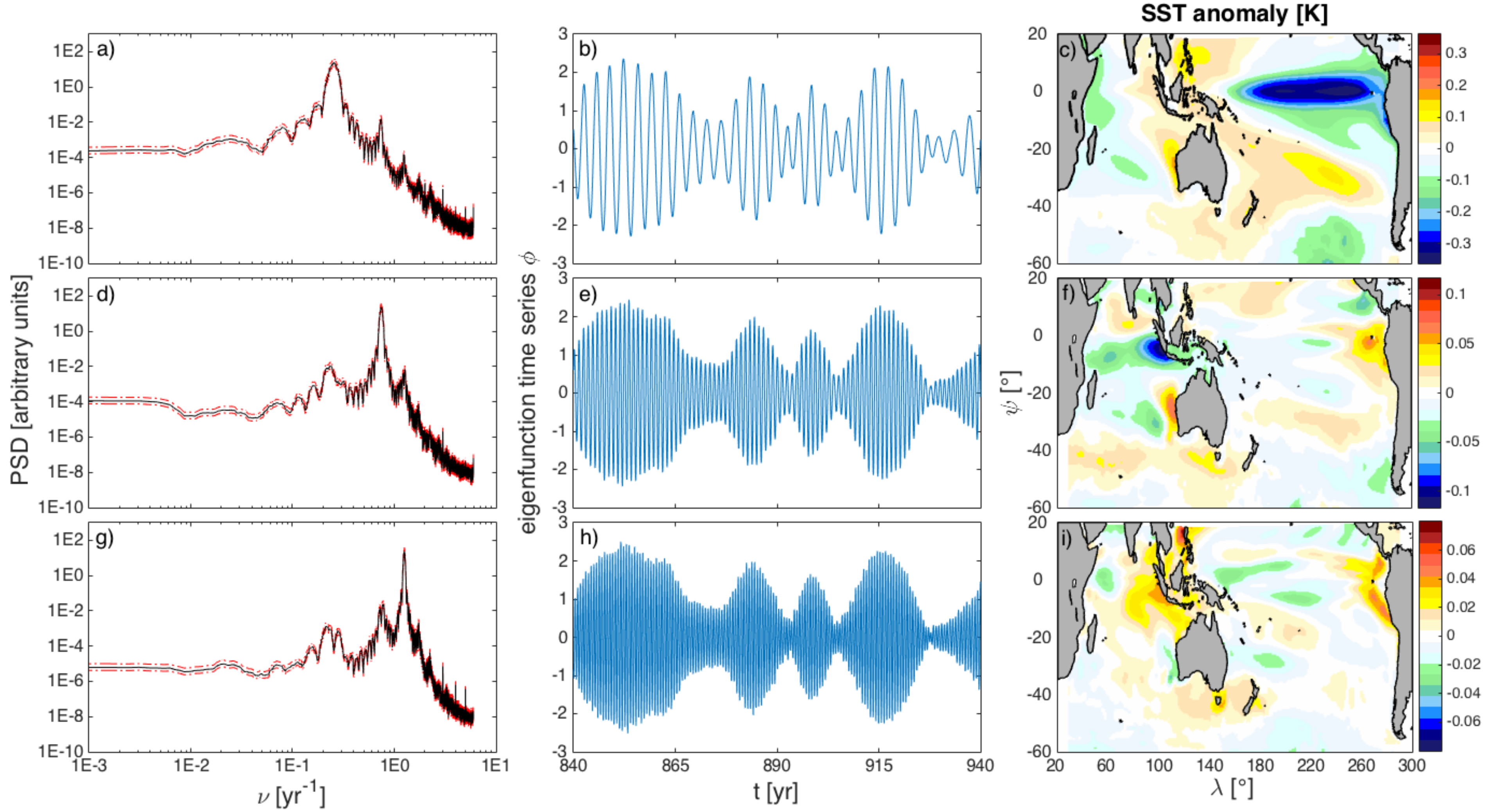
Modes from CCSM4 and CM3:

1. Annual cycle and its harmonics
2. ENSO and ENSO combination modes (McGregor et al. 2012; Stuecker et al. 2013; Ren et al. 2016)
3. Tropospheric biennial oscillation (TBO) (Meehl 1987) and associated combination modes
4. Interdecadal Pacific oscillation (IPO) (Power et al. 1999)
5. West Pacific multidecadal mode (WPMM)

Modes from HadISST:

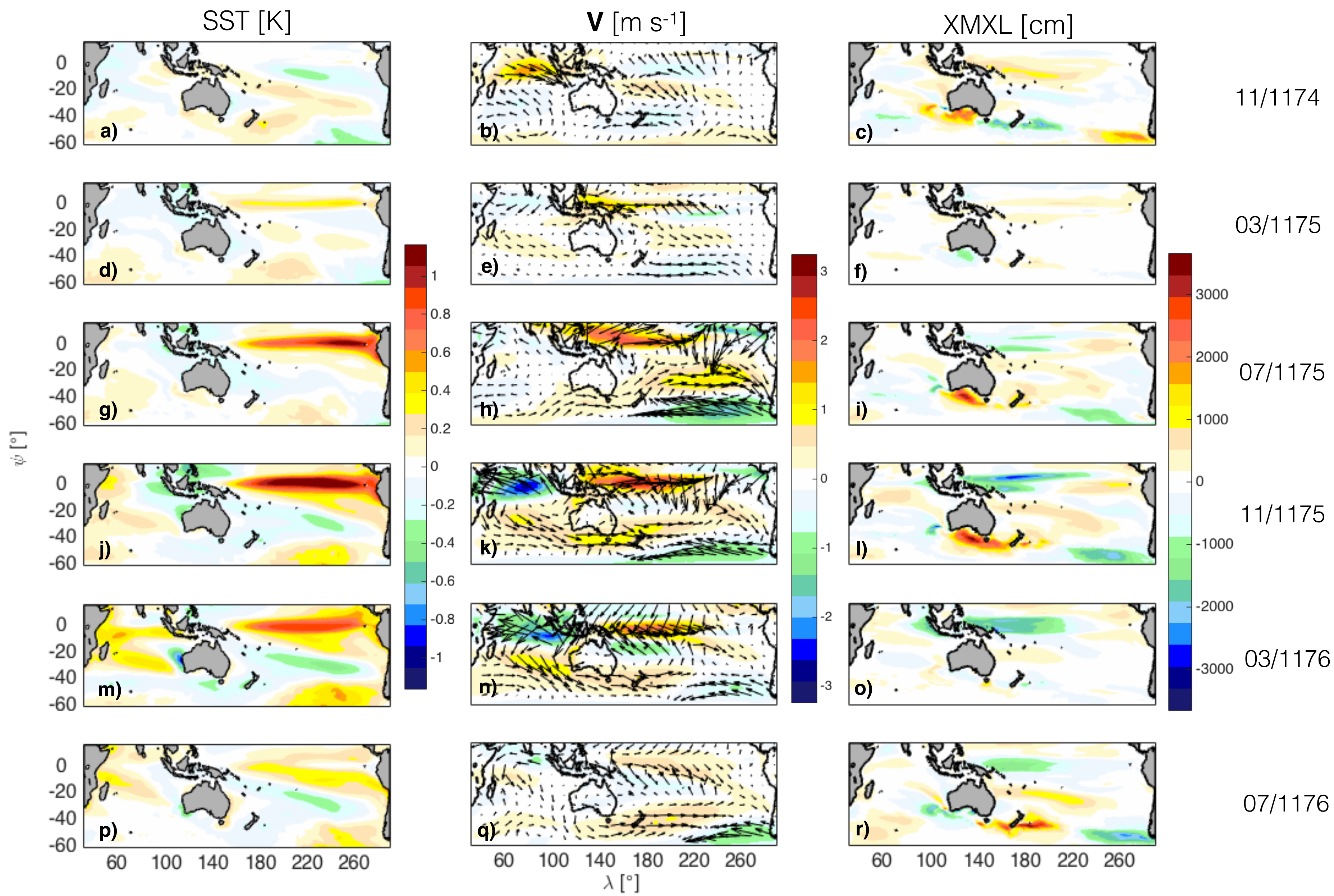
- Annual, ENSO, and TBO modes robustly recovered
- Evidence for IPO and WPMM (though of degraded quality)

ENSO and ENSO combination modes



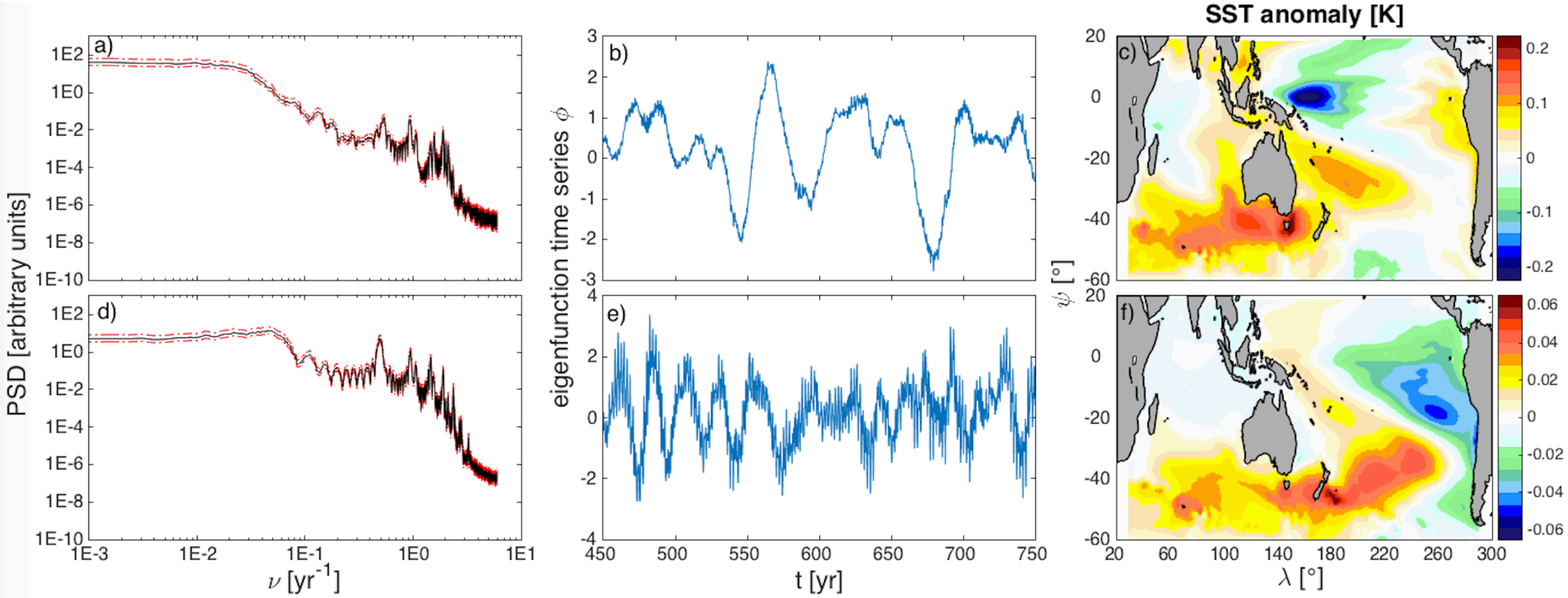
- ENSO emerges as an oscillatory pair of eigenfunctions with a $\nu_{\text{ENSO}} \approx 4$ y^{-1} frequency peak and a **decadal amplitude envelope**
- **Combination modes** predicted from quadratic nonlinearities in the coupled atmosphere-ocean system (McGregor et al. 2012; Stuecker et al. 2013) are recovered at the theoretically expected frequencies $\nu_{\text{ENSO}} \pm 1\ \text{y}^{-1}$ and degeneracies
- The ENSO and ENSO combination modes together capture the phase-locking of ENSO to the annual cycle

ENSO and ENSO combination modes

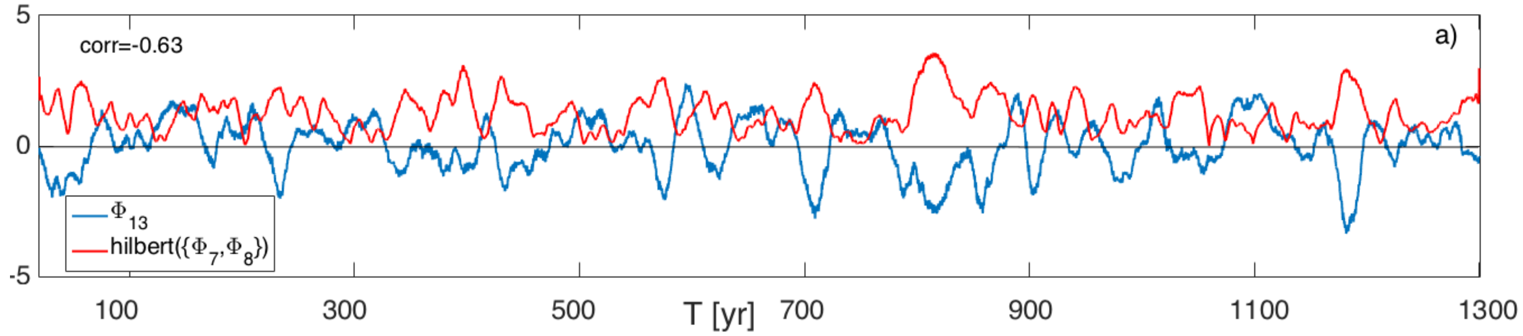


- Reconstructed surface winds exhibit **anomalous westerlies** during the development of El Niño events in boreal winter, and a **southward shift** preceding El Niño decay in boreal spring (Vecchi 2006; Stuecker et al. 2013)
- Surface circulation consistent with **Indian Ocean SST dipole** (Saji et al. 1999; Webster et al. 1999)

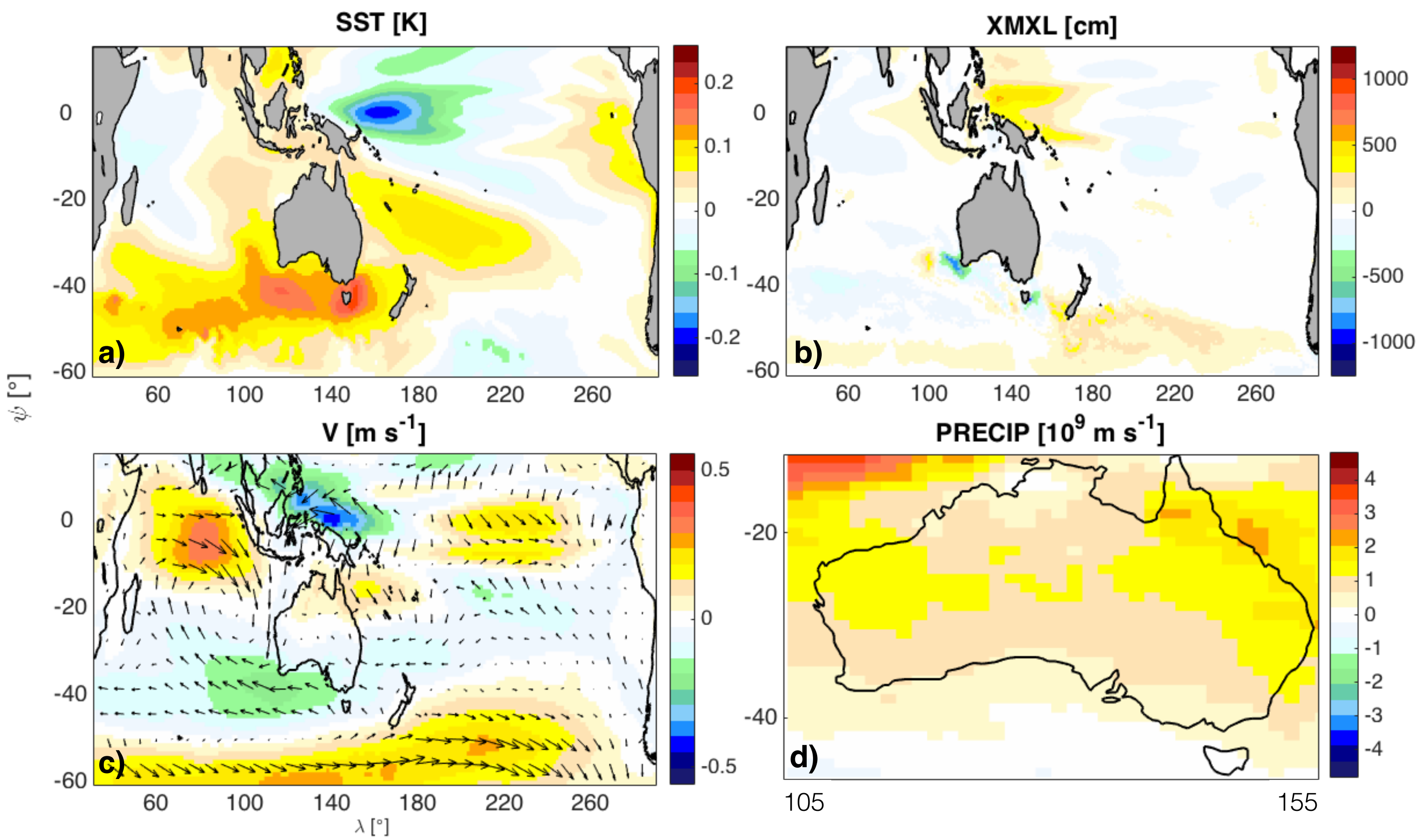
Decadal modes (WPMM & IPO)



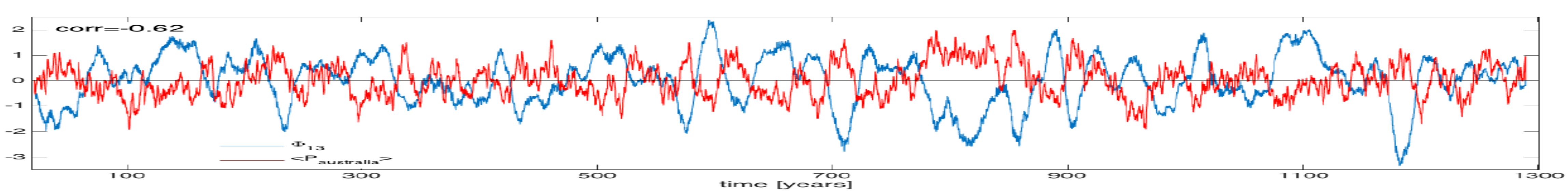
- West Pacific multidecadal mode (top) characterized by multidecadal variability and a prominent cluster of SST anomalies in the **western equatorial Pacific**
- Some similarities with 2nd EOF of decadal Pacific SST (Timmermann 2003; Ogata et al. 2013) and SST residuals (Karnauskas et al. 2009; Solomon & Newman 2012; Seager et al. 2015)
- Cold (warm) WPMM phases correlate with enhanced (suppressed) ENSO activity (corr. coeff. 0.63 in CCSM4)



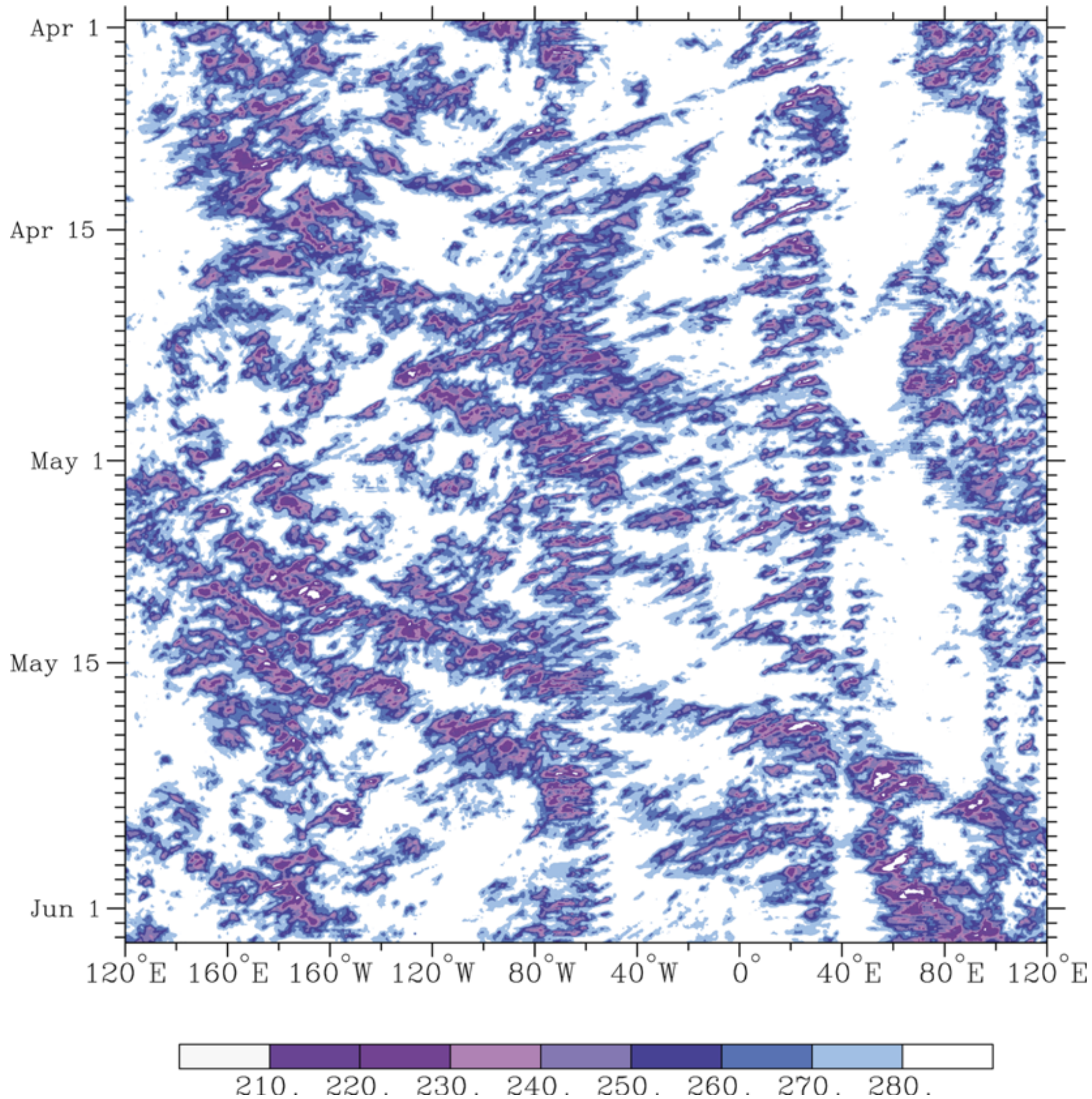
West Pacific multidecadal mode – climate impacts

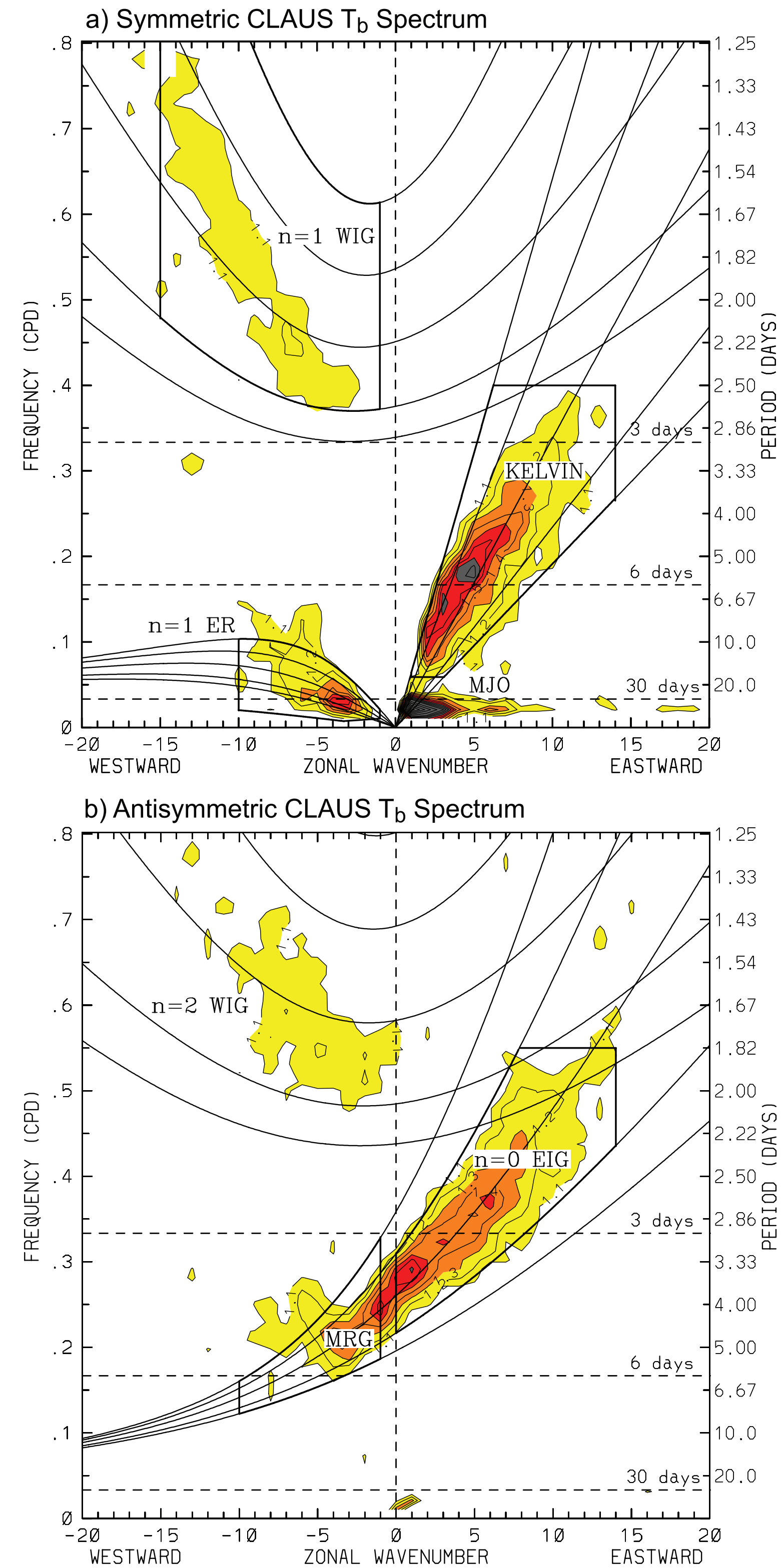


- Cold WPMM phase is characterized by **anomalous westerlies in the central Pacific** and **anomalously flat zonal thermocline profile**; such conditions are known to correlate with enhanced ENSO activity (Kirtman & Schopf 1998; Kleeman et al. 1999; Fedorov & Philander 2000)
- Circulation and SST patterns are consistent with strong impacts on Australian decadal precipitation (corr. coeff. 0.62 in CCSM4)



Analysis of organized tropical convection





Source: Kiladis et al. (2009)

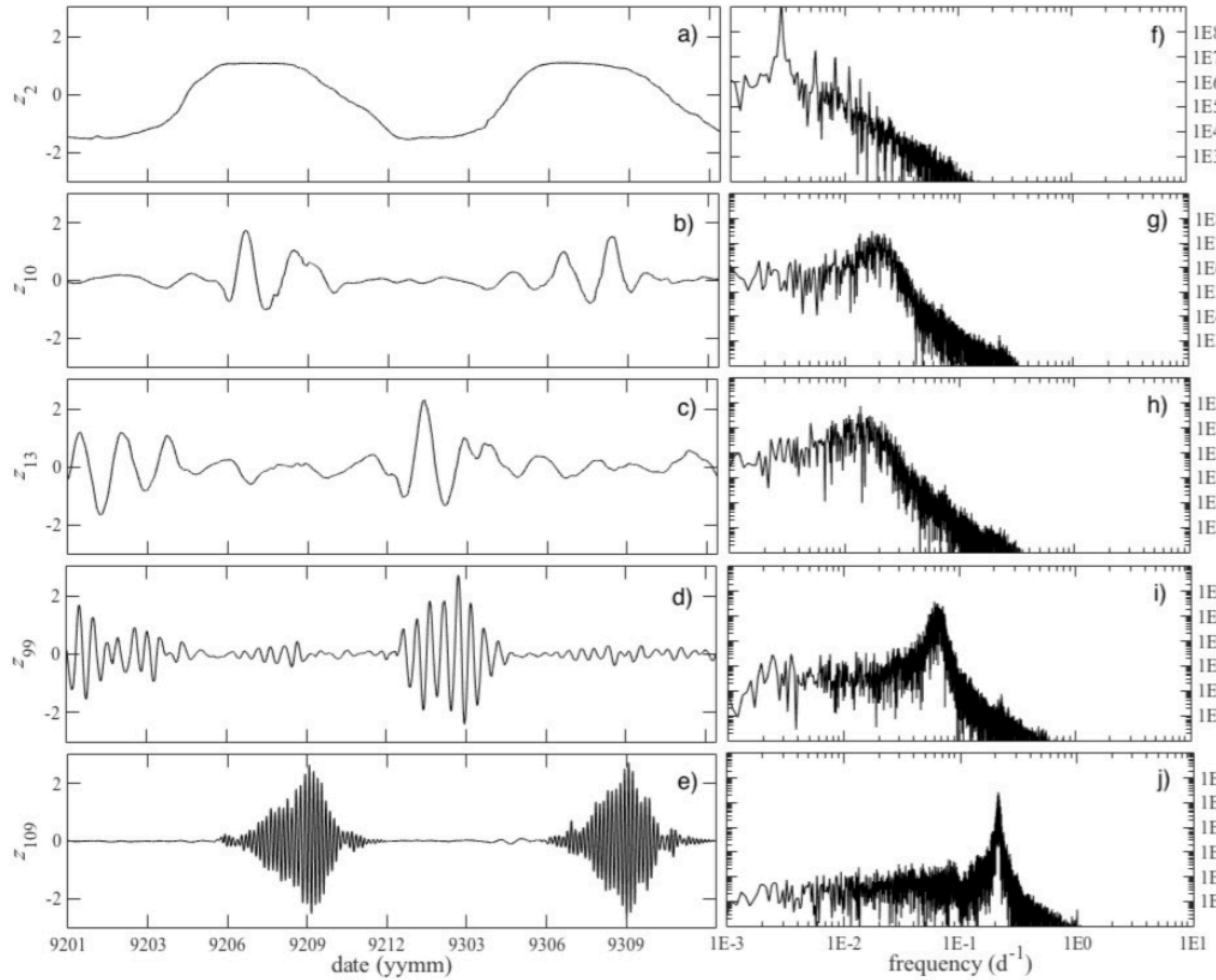
T_b from 1983-2009 CLAUS archive
s=512 delays

Koopman Eigenfunctions

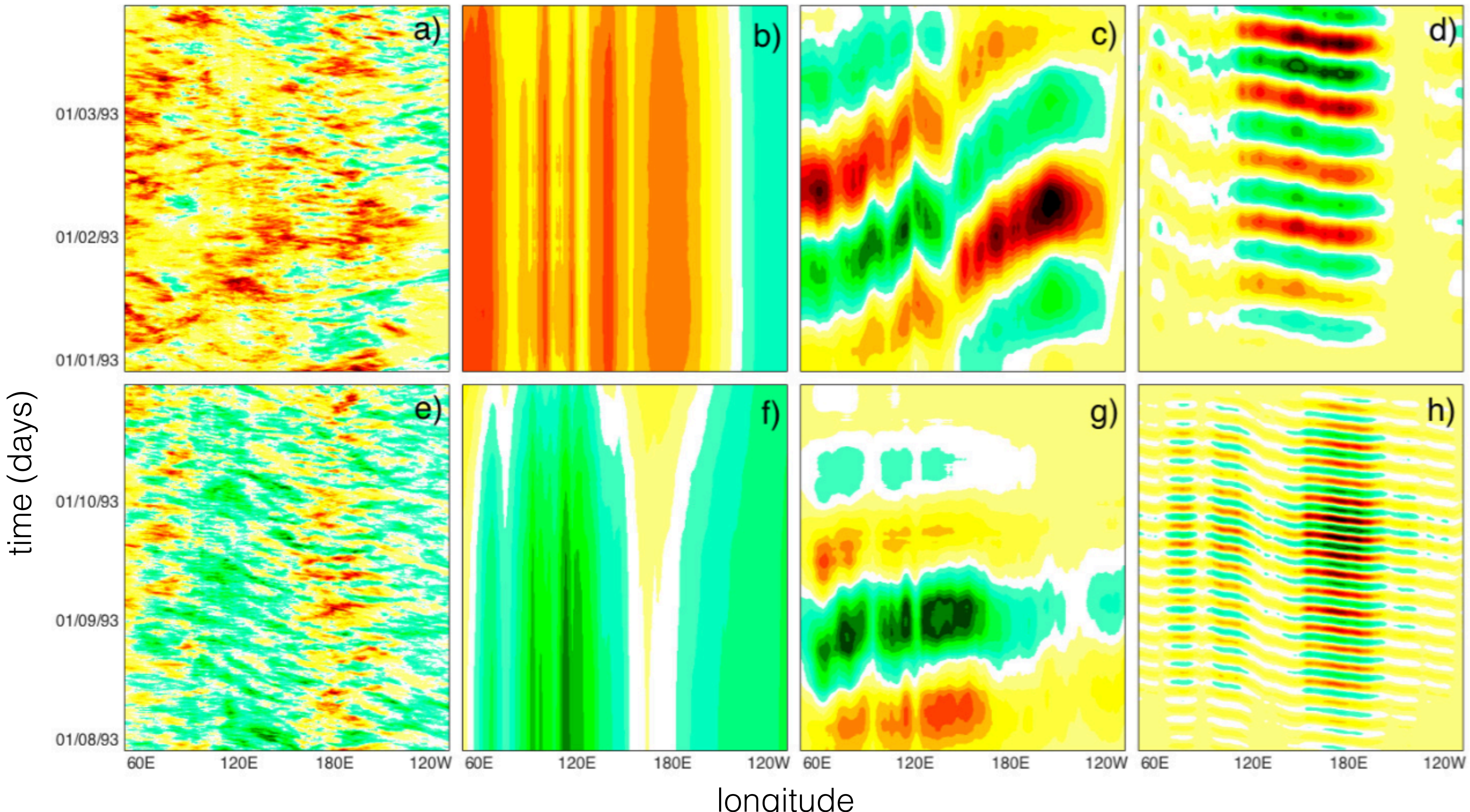
Annual cycle

Dominant intraseasonal oscillations

Convectively Coupled Equatorial Waves (CCEWs)

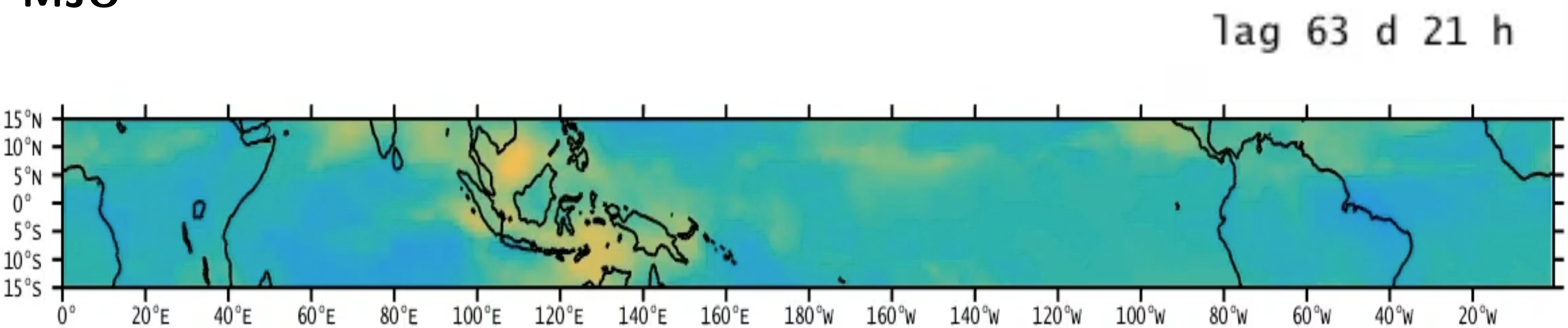


Spatiotemporal reconstruction

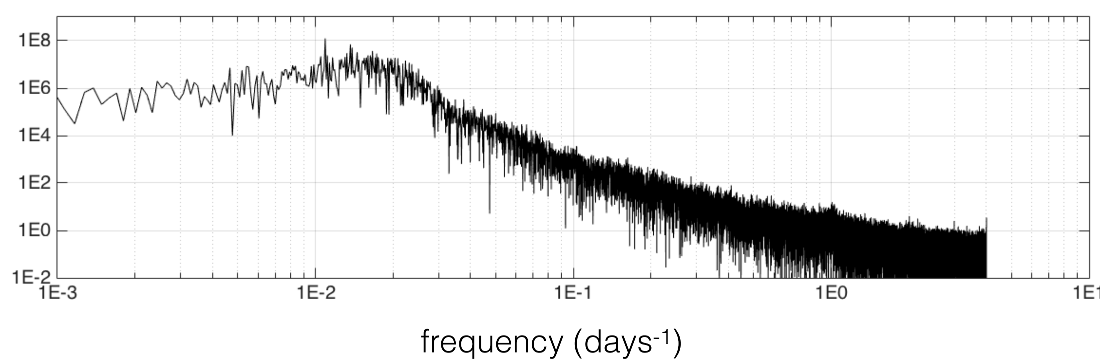


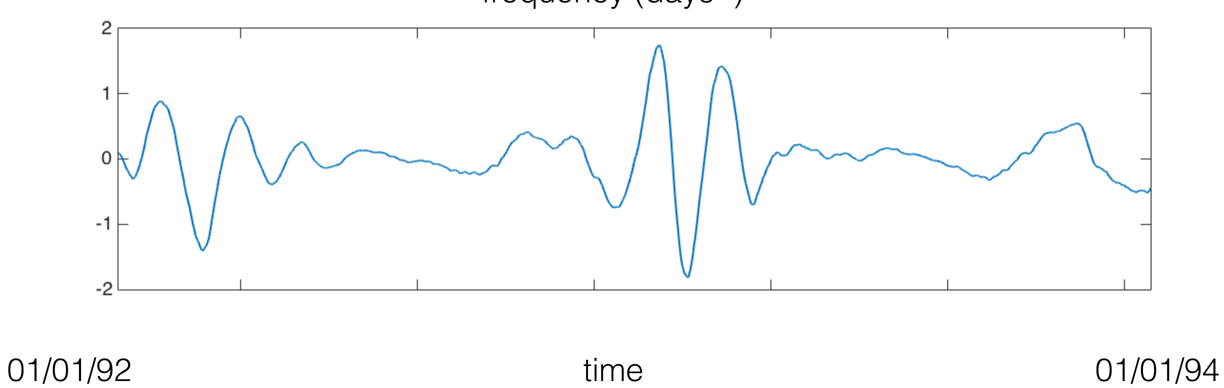
(a,e) Raw Signal
(b,f) Annual cycle
(c) Madden-Julian Oscillation (MJO)
(d,h) Westward-propagating disturbances
(g) Boreal Summer Intraseasonal Oscillation (BSISO)

MJO

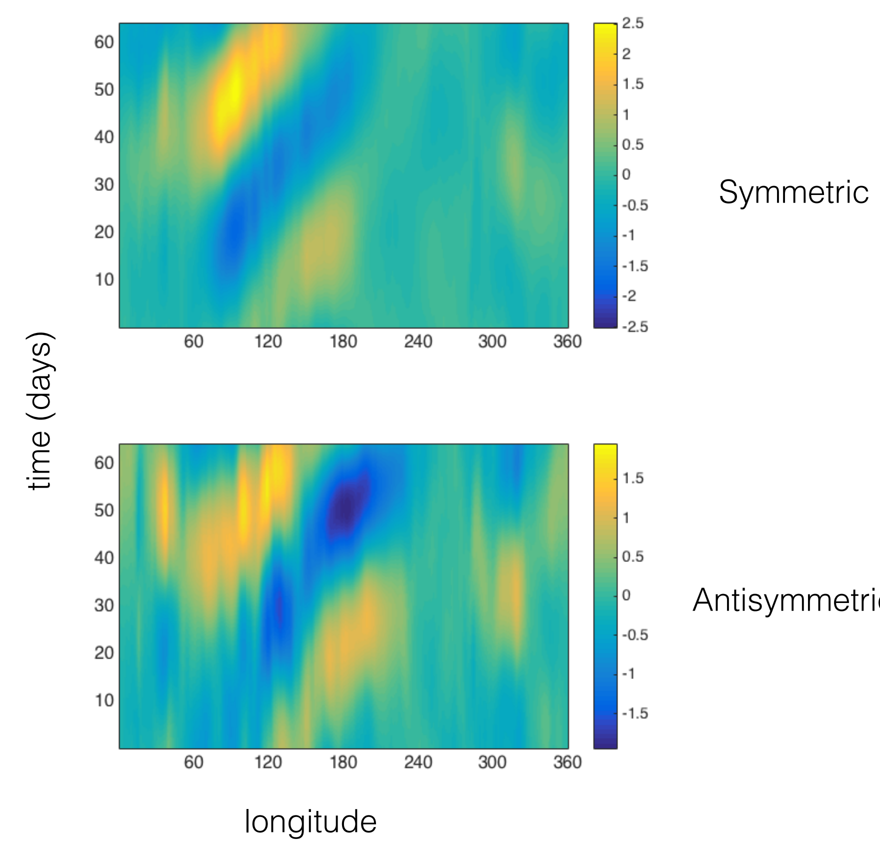


Temporal pattern and frequency spectrum

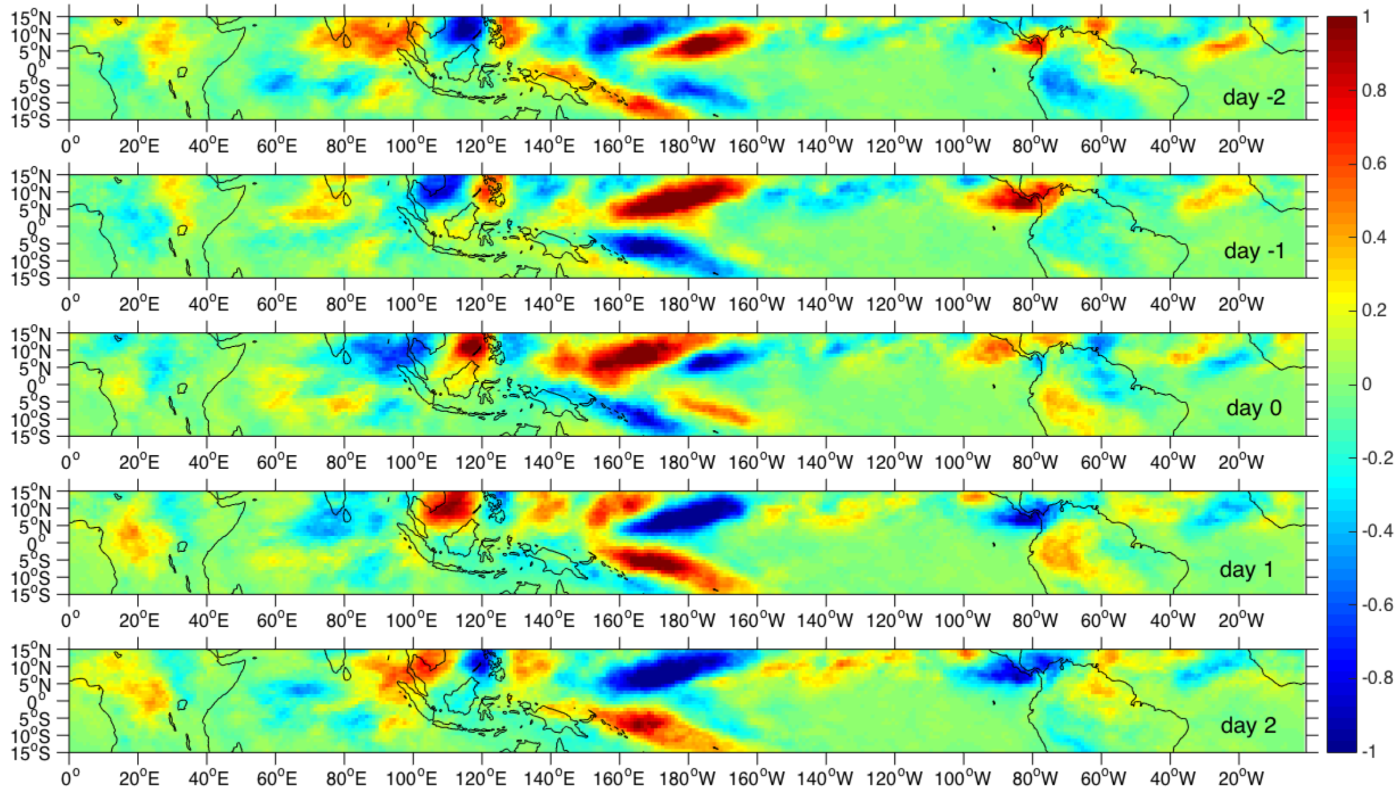




Hovmoller diagram

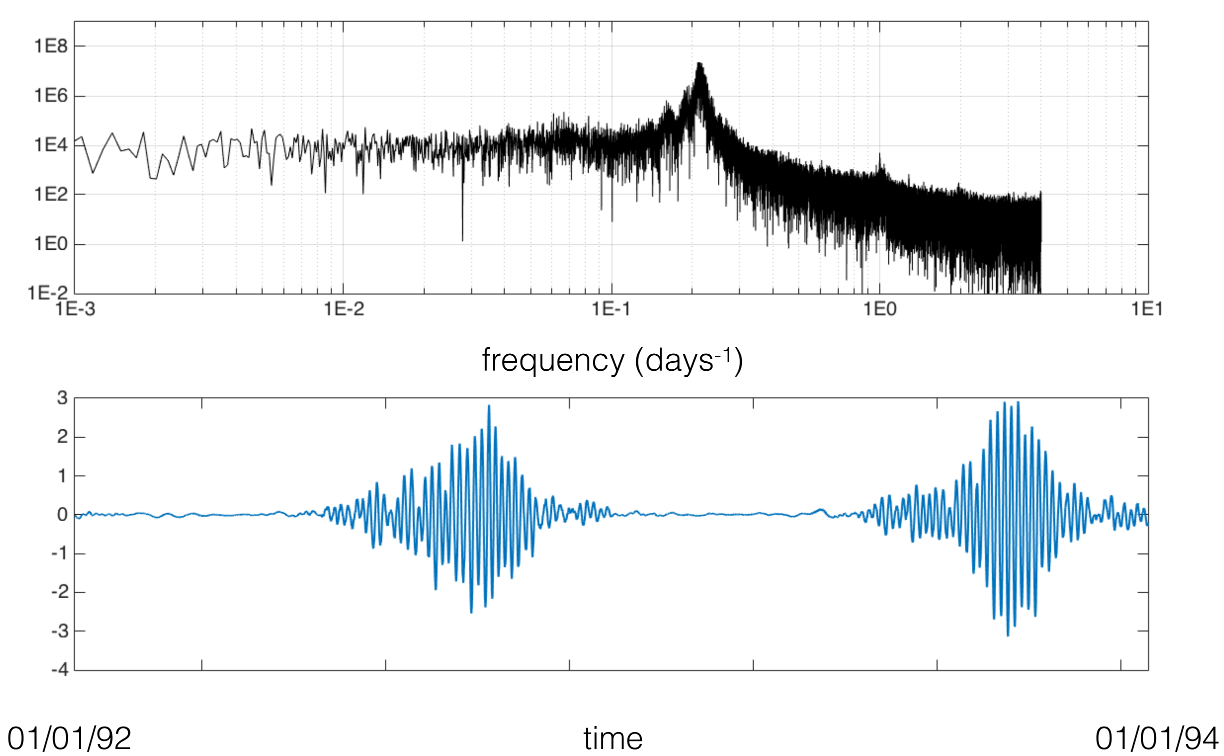


CCEW: mixed Rossby-gravity wave

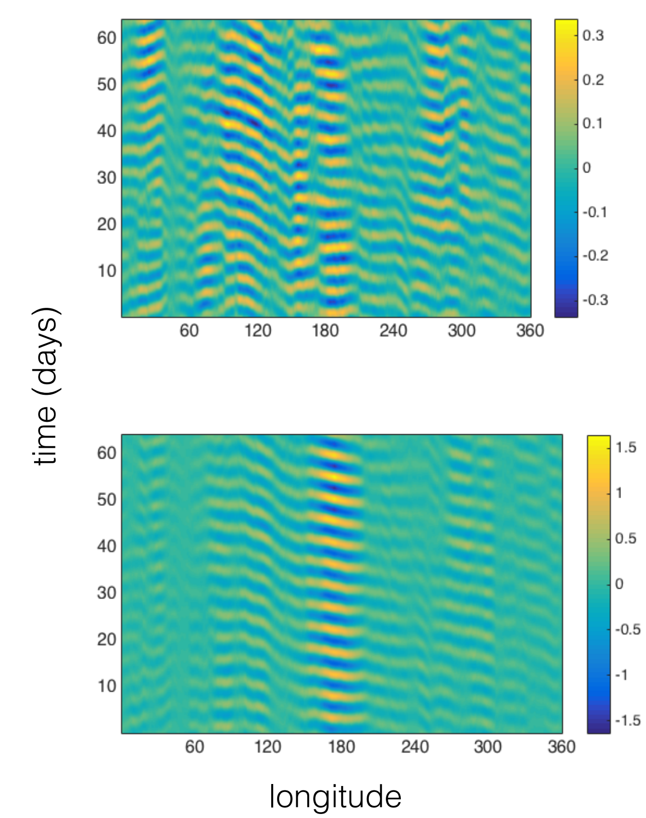


lag 63 d 21 h

Temporal pattern and frequency spectrum



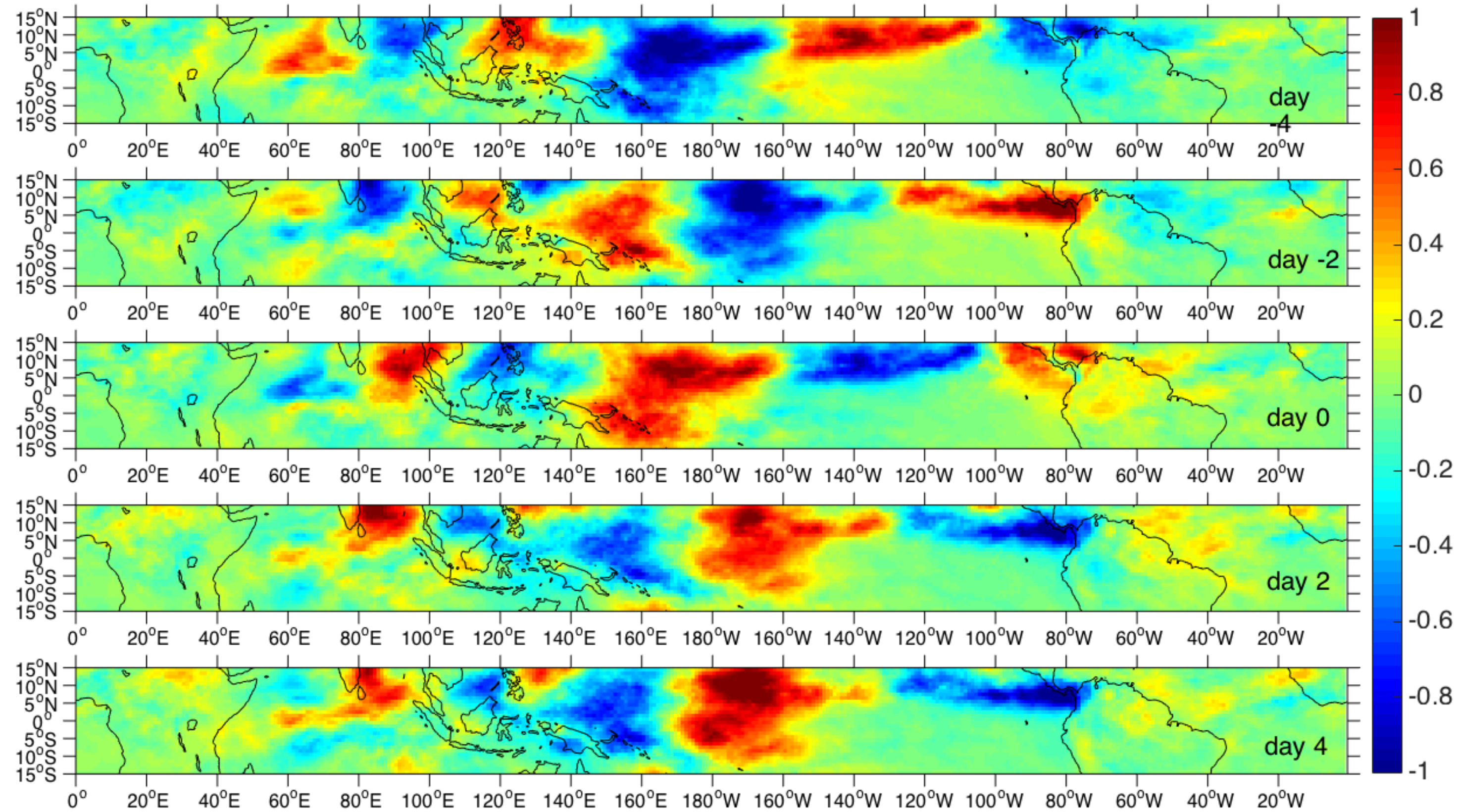
Hovmoller diagram



Symmetric

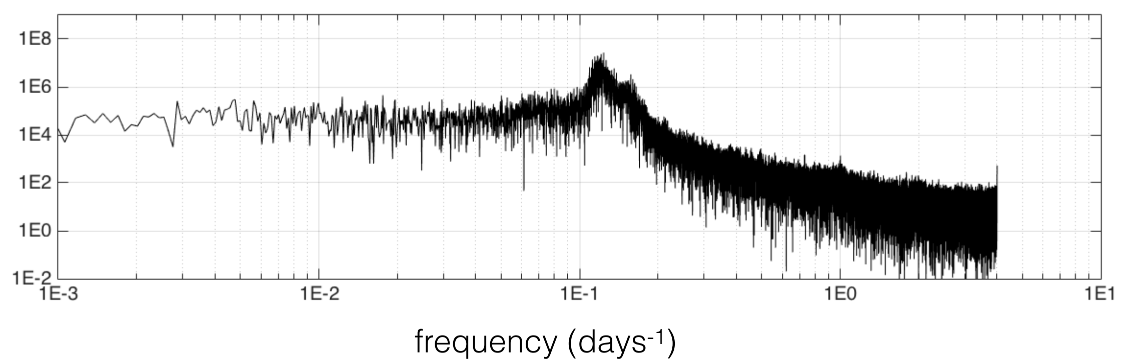
Antisymmetric

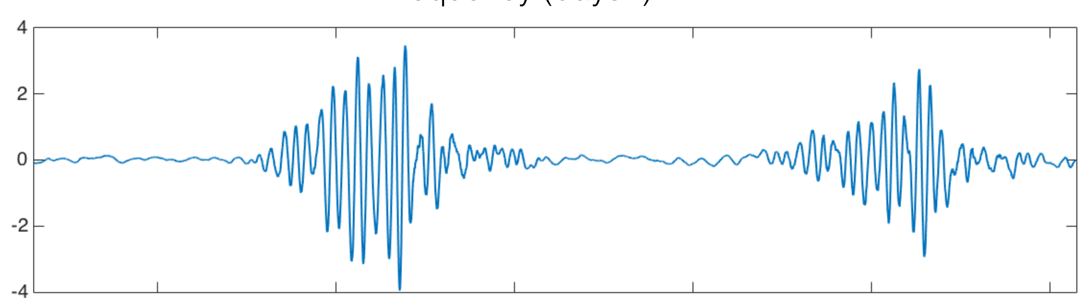
CCEW: Kelvin wave



lag 63 d 21 h

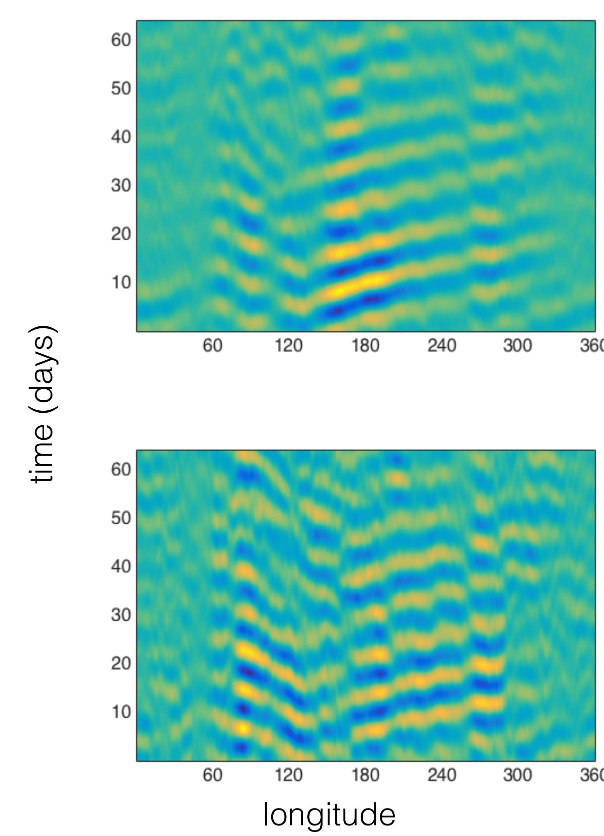
Temporal pattern and frequency spectrum





01/01/92 time 01/01/94

Hovmoller diagram

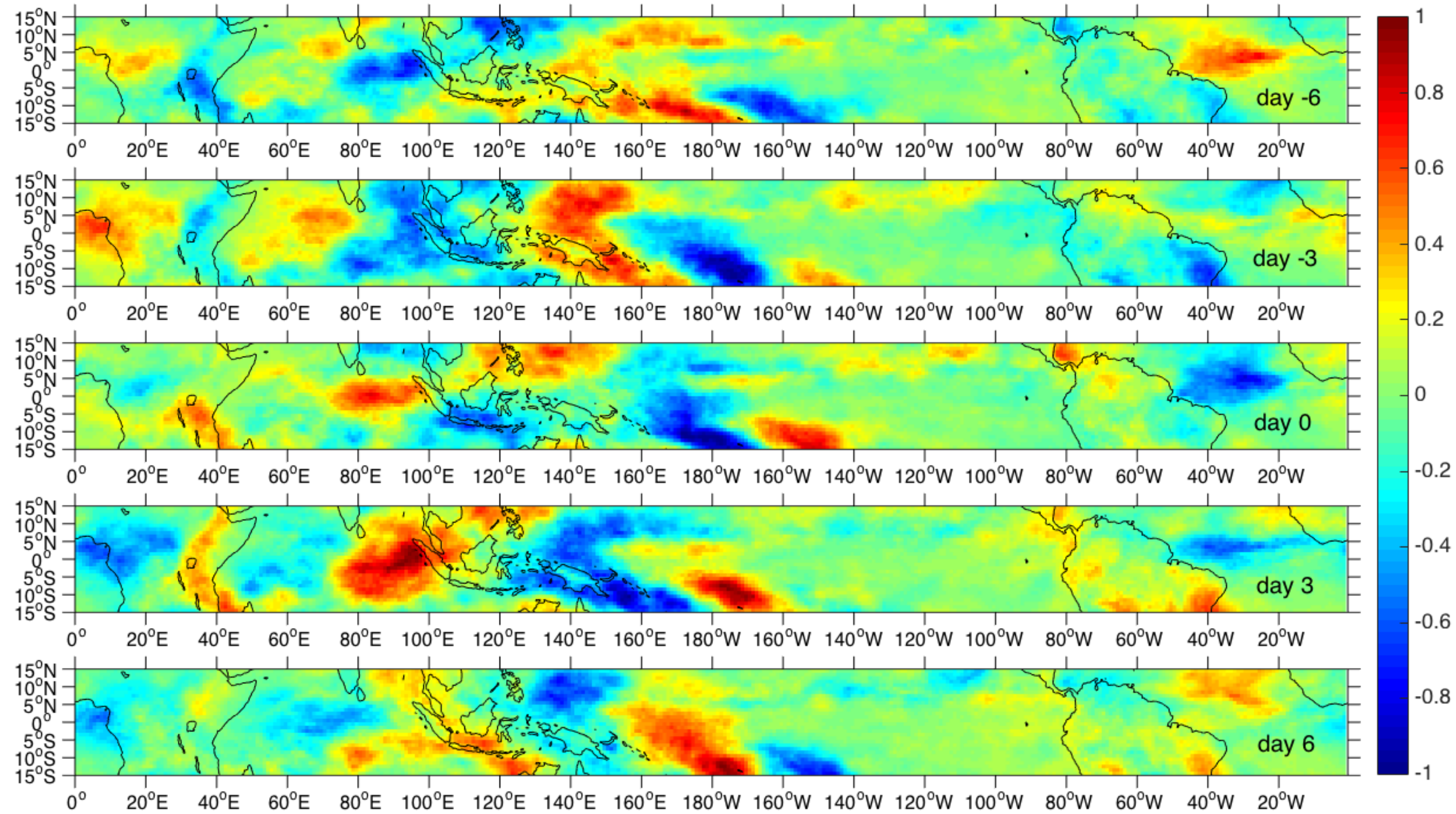


Symmetric

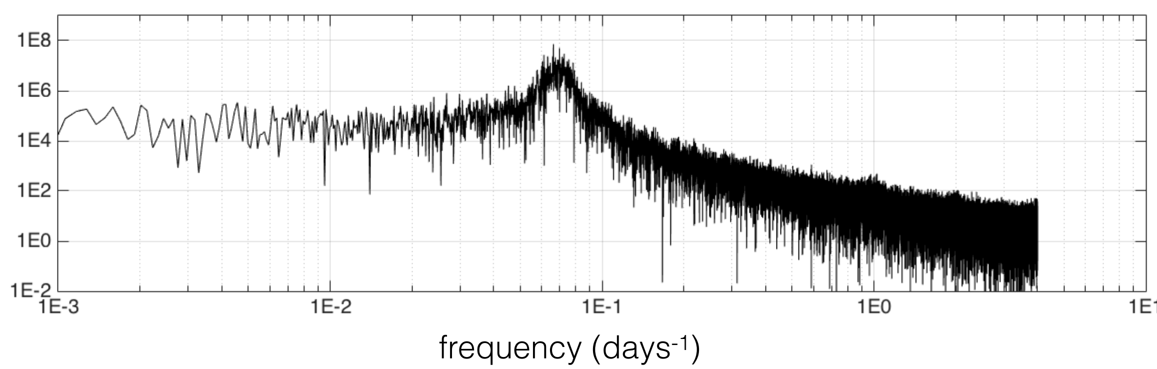
Antisymmetric

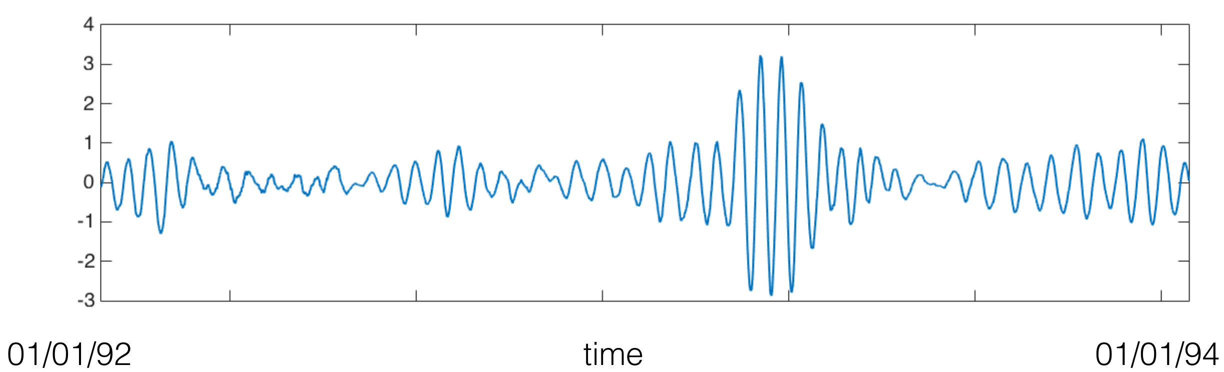
CCEW: equatorial Rossby wave

lag 63 d 21 h

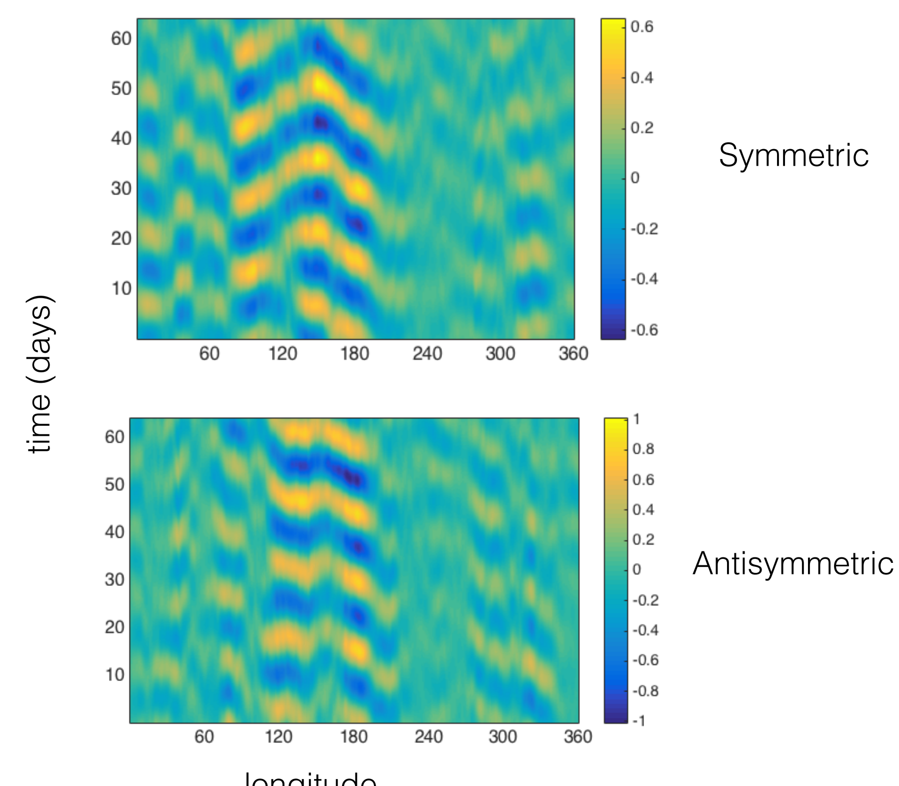


Temporal pattern and frequency spectrum





Hovmoller diagram



Summary

- We have demonstrated the potential of data-driven Koopman operator techniques for extraction of spatiotemporal patterns from high-dimensional multiscale timeseries generated by nonlinear dynamical systems.
- The method relies on constructing low-dimensional representations (feature maps) of spatiotemporal signals using eigenfunctions of the Koopman operator governing the evolution of observables in ergodic dynamical systems.
- This operator is estimated from time-ordered data through a Galerkin scheme applied to basis functions computed via the diffusion maps algorithm.
- In particular, applying this method to 2D brightness temperature data over the tropics, we identified several propagating patterns corresponding to CCEWs.
- To our knowledge, recovery of such patterns from brightness temperature data has previously not been possible via objective eigendecomposition techniques.

References

- D. Giannakis and A. J. Majda (2012), Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. Proc. Natl. Acad. Sci., 109(7),2222-2229
- D. Giannakis (2015), Data-driven Spectral Decomposition and Forecasting of Ergodic Dynamical Systems. arXiv:1507.02338
- D. Giannakis, J. Slawinska, and Z. Zhao (2015), Spatiotemporal Feature Extraction with Data-Driven Koopman Operators. J. Mach. Learn. Res. Workshop and Conference Proceedings, 44
- J. Slawinska and D. Giannakis (2016), Spatiotemporal pattern extraction with data-driven Koopman Operators for Convectively Coupled Equatorial Waves. Proceedings for 6th International Climate Informatics Workshop
- J. Slawinska and D. Giannakis (2016), Indo-Pacific Variability on Seasonal to Multidecadal Timescales. Part I: Intrinsic SST Modes in Models and Observations. J. Climate, in revision
- D. Giannakis and J. Slawinska (2016), Indo-Pacific Variability on Seasonal to Multidecadal Timescales. Part II: Multi-Scale Ocean–Atmosphere Interactions. J. Climate, in revision