Chapter 5: Exploratory Methods for Spatio-Temporal Data

Jim Faulkner

Space-Time Reading Group
University of Washington

Feb 5, 2018

Outline

1. Spectral Analysis
2. Empirical Covariance Functions
3. Empirical Orthogonal Function Analysis
4. Principal Oscillation Patterns
5. Canonical Correlation Analysis

Spectral Analysis

Why use spectral analysis?

- Partition overall variation into different scales, e.g., seasonal, annual, decadal
- Determine which components account for the most variation
- Simplifies correlation structure – spectral transformation decorrelates
- Can allow for reduced-rank approximations

3.5.1 Spectral Representations via Orthogonal Series

Consider a process that can be written as a function of time, $f(t)$, over the interval (a, b). Define a sequence of spectral basis functions $\phi_k(t)$ for $k = 0, 1, \ldots,$ to be *orthogonal* over the interval (a, b) if

$$\int_a^b \phi_k(t)\phi_l(t)dt = \begin{cases} 0, & k \neq l, \\ \delta, & k = l, \end{cases} \tag{3.118}$$

where $\delta > 0$. These functions are *orthonormal* if $\delta = 1$.

Now we expand $f(t)$ in terms of these basis functions:

$$f(t) = \sum_{k=0}^{\infty} \alpha_k \phi_k(t), \tag{3.119}$$

where α_k are the *weights* or *spectral coefficients* of the expansion.

$$\alpha_k = \int_a^b f(t)\phi_k(t)dt \quad k = 1, 2, \ldots, \tag{3.120}$$

The spectral coefficients are the projection of $f(t)$ onto the orthogonal basis functions.

Trigonometric Series Expansion

Consider $f(t)$ defined on interval $(-1/2, 1/2)$. Let $\phi_k(t)$ correspond to the orthonormal trigonometric basis functions $\sin(2\pi kt)$ and $\cos(2\pi kt)$. We can write

$$f(t) = \frac{a_0}{2} + \sum_{k=0}^{\infty} \{a_k \cos(2\pi kt) + b_k \sin(2\pi kt)\} \tag{3.122}$$

The *Fourier coefficients* are:

$$a_k = 2 \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t) \cos(2\pi kt) dt, \quad k = 0, 1, \ldots \tag{3.123}$$

$$b_k = 2 \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t) \sin(2\pi kt) dt, \quad k = 0, 1, \ldots \tag{3.124}$$

Using Euler's relationship, $e^{\pm i2\pi kt} \equiv \cos(2\pi kt) \pm i\sin(2\pi kt)$, we can simplify :

$$f(t) = \sum_{k=0}^{\infty} \alpha_k e^{\pm i2\pi kt} \quad -1/2 \leq t \leq 1/2, \tag{3.125}$$

where $\phi_k(t) \equiv e^{\pm i2\pi kt}$ and $\alpha_k \equiv a_k + ib_k$ are both complex functions. This result is the *Fourier Transform* or the *spectral representation theorem*.

3.5.2 Discrete-Time Spectral Expansion

Let $\{Y_t : t = 1, \ldots, T\}$ be a times series and define $\{\phi_k(t) : t = 1, \ldots, T; k = 1, \ldots, p_\alpha\}$ to be a complete set of basis functions. Define $\boldsymbol{Y} \equiv (Y_1, \ldots, Y_T)'$, $\boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_{p_\alpha})$, and $\boldsymbol{\Phi} \equiv (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{p_\alpha})$, where $\boldsymbol{\phi}_k \equiv (\phi_k(1), \ldots, \phi_k(T))$. Then the spectral expansion of $\boldsymbol{Y}$

$$Y_t = \sum_{k=1}^{p_\alpha} \alpha_k \phi_k(t) \tag{3.126}$$

can be written in matrix notation as

$$\boldsymbol{Y} = \boldsymbol{\Phi}\boldsymbol{\alpha} \tag{3.127}$$

Multiply both sides of (3.127) by $\mathbf{\Phi}$ to obtain the spectral coefficient vector,

$$\boldsymbol{\alpha} = \left(\mathbf{\Phi}'\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}'\boldsymbol{Y} \tag{3.128}$$

which is in the form of a least-squares estimator. When the basis functions are orthonormal, $\mathbf{\Phi}'\mathbf{\Phi} = \boldsymbol{I}$, then (3.128) becomes

$$\boldsymbol{\alpha} = \mathbf{\Phi}'\boldsymbol{Y} \tag{3.129}$$

Note that the operation $\mathbf{\Phi}'\boldsymbol{Y}$ can be carried out using the *Fast Fourier Transform* (FFT), and the operation $\mathbf{\Phi}\alpha$ can be carried out using the inverse FFT.

Figure 5.1. SST anomalies one location through time.

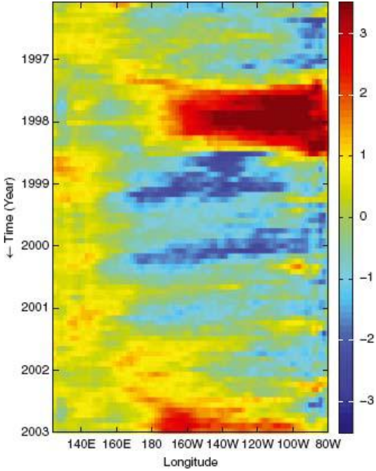
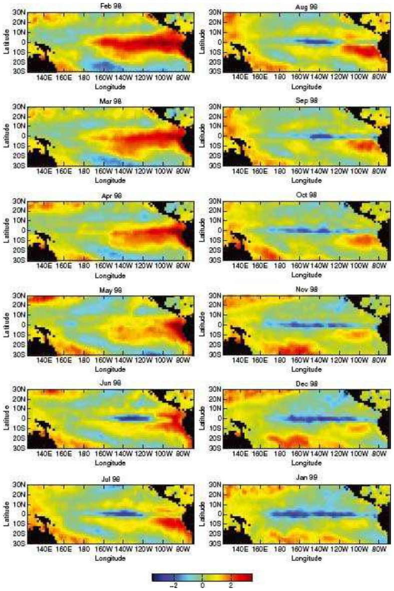


Figure 5.4. SST anomalies in tropical Pacific through time.



Empirical Covariance Functions

Assume we have observations $\mathbf{Z}_t \equiv (Z(\mathbf{s}_1;t),\ldots,Z(\mathbf{s}_m;t)')$ for $t = 1,\ldots,T$. An $m \times m$ empirical (averaged over time) lag-τ spatial covariance matrix is given by

$$\hat{\mathbf{C}}_Z^{(\tau)} \equiv \frac{1}{t-\tau}\sum_{t=\tau+1}^{T}(\mathbf{Z}_t - \hat{\boldsymbol{\mu}}_Z)(\mathbf{Z}_{t-\tau} - \hat{\boldsymbol{\mu}}_Z)', \quad \tau = 0,1,\ldots,T \quad (5.1)$$

where the empirical spatial mean is given by

$$\hat{\boldsymbol{\mu}}_Z \equiv \frac{1}{T}\sum_{t=1}^{T}\mathbf{Z}_t$$

Figure 5.6. Lag-0 covariance and correlation along equator for SST.

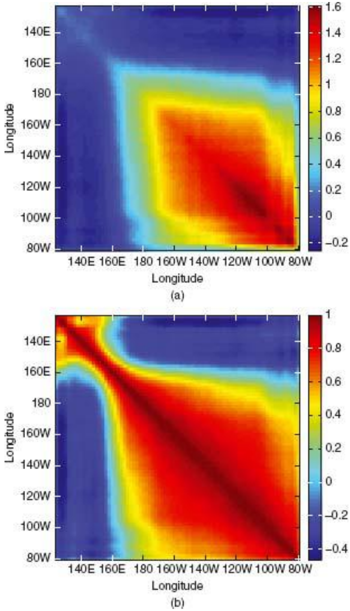
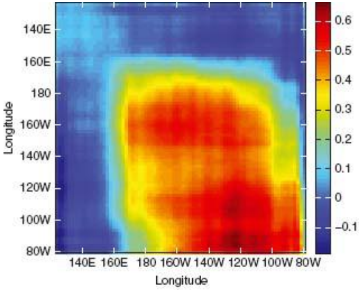
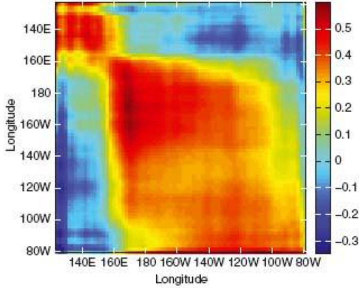


Figure 5.7. Lag-6 covariance and correlation along equator for SST.



(a)



(b)

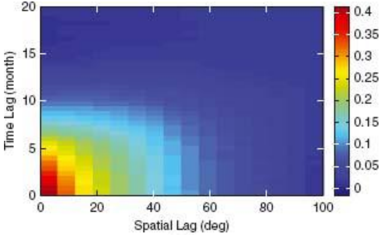
The estimated spatio-temporal covariance at spatial lag $\mathbf{h}$ and time lag τ is given by

$$\hat{C}_Z(\mathbf{h};\tau) \equiv \frac{1}{|N_s(\mathbf{h})|}\frac{1}{|N_t(\tau)|}\times \sum_{\mathbf{s}_i,\mathbf{s}_j \in N_s(\mathbf{h})} \sum_{t,r\in N_t(\tau)} (Z(\mathbf{s}_i;t)-\hat{\mu}(\mathbf{s}_i))(Z(\mathbf{s}_j;r)-\hat{\mu}(\mathbf{s}_j))$$

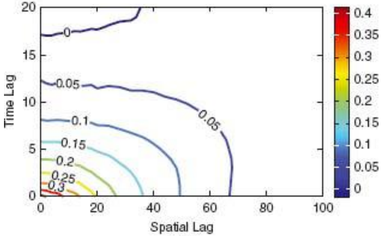
where $N_s(\mathbf{h})$ refers to pairs of spatial locations with spatial lag within some tolerance of $\mathbf{h}$, $N_t(\tau)$ refers to pairs of time points with time lag within some tolerance of τ, and $|N(\cdot)|$ refers to the cardinality (number of elements) in the set $N(\cdot)$. Also,

$$\hat{\mu}_Z(\mathbf{s}_i) \equiv \frac{1}{T}\sum_{t=1}^{T} Z(\mathbf{s}_i;t)$$

Figure 5.9. SST spatio-temporal covariance.



(a)



(b)

5.3 Empirical Orthogonal Function (EOF) Analysis

- EOFs are eigenvectors from eigen (spectral) decomposition of covariance matrix
- In discrete formulation is PCA, in continuous is Karhunen-Loève expansion.
- Typically used
  1. Diagnostically to find principal spatial structures and how those vary
  2. To reduce dimensionality in spatio-temporal data sets and to reduce noise

5.3.1 Spatially Continuous Formulation

Consider data $\{Z_t(\mathbf{s}) : \mathbf{s} \in D_t, t = 1, 2, \dots\}$ from a continuous spatial process measured at discrete time intervals. We want to find an optimal separable orthogonal decomposition:

$$Z_t(\mathbf{s}) = \sum_{k=1}^{\infty} \alpha_t(k)\phi_k(\mathbf{s}), \tag{5.17}$$

such that $\text{var}(\alpha_t(1)) > \text{var}(\alpha_t(1)) > \dots$, and $\text{cov}(\alpha_t(i), \alpha_t(k)) = 0$ for all $i \neq k$.

The solution is the Karhunen-Loève expansion, which allows the decomposition:

$$C_Z^{(0)}(\mathbf{s}, \mathbf{r}) = \sum_{k=1}^{\infty} \lambda_k \phi_k(\mathbf{s}) \phi_k(\mathbf{r}), \tag{5.18}$$

where $\{\phi_k(\cdot)\}$ are the eigenfunctions and $\{\lambda_k\}$ are the eigenvalues of the Fredholm integral equation:

$$\int_{D_s} C_Z^{(0)}(\mathbf{s}, \mathbf{r}) \phi_k(\mathbf{s}) d\mathbf{s} = \lambda_k \phi_k(\mathbf{r}) \tag{5.19}$$

This can be solved numerically, but is difficult and usually not done in practice.
The kth "amplitude" times series $\{\alpha_t(k) : t = 1, 2, \dots\}$ is

$$\alpha_t(k) = \int_{D_s} Z_t(\mathbf{s}) \phi_k(\mathbf{s}) d\mathbf{s}, \quad t = 1, 2, \dots \tag{5.22}$$

5.3.2 Spatially Discrete Formulation

- Let $\mathbf{Z}_t \equiv (Z_1(\mathbf{s}_1), \ldots, Z_t(\mathbf{s}_m))'$ and define the kth discrete EOF to be $\boldsymbol{\psi}_k \equiv (\psi_k(\mathbf{s}_1), \ldots, \psi_k(\mathbf{s}_m))$, where $\boldsymbol{\psi}_k$ is the vector in the linear combination $a_t(k) = \boldsymbol{\psi}_k' \mathbf{Z}_t$, for $k = 1, \ldots, m$.
- In general, $\boldsymbol{\psi}_k$ is the vector that maximizes $\text{var}(a_t(k))$ subject to the constraints $\boldsymbol{\psi}_k' \boldsymbol{\psi}_k = 1$ and $\text{cov}(a_t(k), a_t(j)) = 0$ for all $j \neq k$. This is equivalent to solving the eigen decomposition of $C_Z^{(0)}$:
$$C_Z^{(0)} = \boldsymbol{\Psi} \boldsymbol{\Lambda} \boldsymbol{\Psi}'$$
where $\boldsymbol{\Psi} \equiv (\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_m)$ is the $m \times m$ orthonormal matrix of eigenvectors ($\boldsymbol{\Psi}' \boldsymbol{\Psi} = I$), $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_m)$ is the $m \times m$ diagonal matrix of eigenvalues decreasing down the diagonal, $\text{var}(a_t(k) = \lambda_k)$, and Z_t is centered with mean $\mathbf{0}$.

Figure 5.17. SST first and second EOFs.

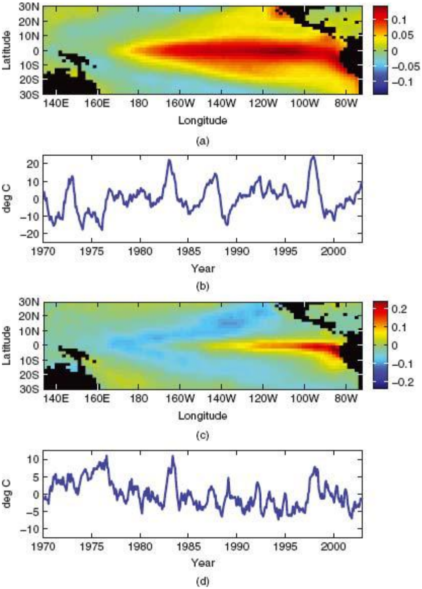
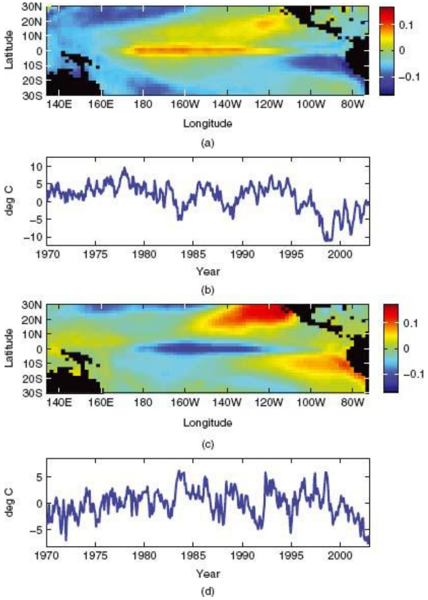


Figure 5.18. SST third and fourth EOFs.



5.5 Principal Oscillation Patterns (POPs)

- Spectral decomposition of the propogator matrix for the first-order dynamical system provides information about underlying dynamics
- POP analysis based on spectral decomposition of propogator matrix of a spatio-temporal process expressed as a first-order dynamical system
- Consider the first-order discrete linear system

$$\mathbf{Z}_t = \mathbf{M}\mathbf{Z}_{t-1} \tag{5.26}$$

where $\mathbf{Z}_t \equiv (Z_t(\mathbf{s}_1), \ldots, Z_t(\mathbf{s}_1))'$, and $\mathbf{M}$ is the $m \times m$, full-rank but non-symmetric propogator matrix.

- In short, if we calculate the normalized SVD $\mathbf{M} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{V}'$ then

$$\mathbf{Z}_t = \mathbf{W}\mathbf{V}'\mathbf{Z}_t = \mathbf{W}\mathbf{a}_t \tag{5.28}$$

  where $\mathbf{a}_t \equiv \mathbf{V}'\mathbf{Z}_t$ are the *POP coefficients*.
- The columns of $\mathbf{W}$, $\{\mathbf{w}_k\}$ are the *principal oscillation patterns* and the columns of $\mathbf{V}$, $\{\mathbf{v}_k\}$ are called the *adjoint bases*.
- We can form the recursion $\mathbf{a}_t = \boldsymbol{\Lambda}\mathbf{a}_{t-1}$ or $a_t(k) = \lambda_k a_{t-1}$, which results in solutions $a_t(k) = (\lambda_k)^t$.
- These coefficients evolve according to $a_t(k) = \lambda_k^t e^{i\phi_k t}$, where $\gamma_k = |\lambda_k|$.
- Components of interest are $a_t(k)$, γ_k, ϕ_k, $a_0(k)/e$ and $\tau_k = -1/\ln(\gamma_k)$.

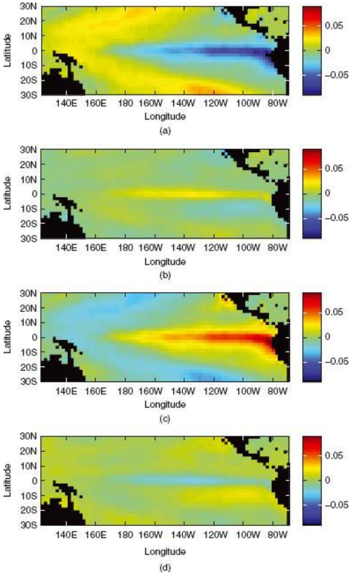
5.5.1 Calculation of POPs

- To account for uncertainty in the data, we write the process as a first-order vector-autoregressive process:

$$\mathbf{Z}_t = \mathbf{M}\mathbf{Z}_{t-1} + \boldsymbol{\eta}_t \tag{5.33}$$

  where $\boldsymbol{\eta}_t$ has mean $\mathbf{0}$ with uncorrelated elements.
- Under second-order stationarity, $\mathbf{M} = \mathbf{C}_Z^{(1)}(\mathbf{C}_Z^{(0)})^{-1}$
- A method-of-moments estimator is $\hat{\mathbf{M}} = \hat{\mathbf{C}}_Z^{(1)}(\hat{\mathbf{C}}_Z^{(0)})^{-1}$
- Calculate eigenvalue decomposition of $\hat{\mathbf{M}}$, which results in $\hat{\mathbf{M}} = \hat{\mathbf{W}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{W}}^{-1}$
- Set $\hat{\mathbf{V}}' = \hat{\mathbf{W}}^{-1}$, then the POP-coefficient estimates can be obtained from $\hat{\mathbf{a}}_t = \hat{\mathbf{V}}'\mathbf{Z}_t$

Figure 5.23. SST POP for 10th eigenvalue. a) real part, b) negative imaginary part, c) negative real part, c) imaginary part.



5.6 Spatio-Temporal Canonical Correlation Analysis (CCA)

- CCA obtains linear combinations of two sets of random variables whose correlations are maximal
- Can apply to two random variables indexed by space and time
- Suppose we have two data sets $\{\mathbf{Z}_t \equiv (Z_t(\mathbf{s}_1), \ldots, Z_t(\mathbf{s}_m))'\}$ and $\{\mathbf{X}_t \equiv (X_t(\mathbf{x}_1), \ldots, X_t(\mathbf{x}_\ell))'\}$ with a possibly different spatial domain but the same temporal domain $(t = 1, \ldots, T)$
- The *kth canonical correlation* is defined as

$$r_k \equiv \text{corr}(\boldsymbol{\xi}_k' \mathbf{Z}_t, \boldsymbol{\psi}_k' \mathbf{X}_t) = \frac{\boldsymbol{\xi}_k' \mathbf{C}_{Z,X}^{(0)} \boldsymbol{\psi}_k'}{(\boldsymbol{\xi}_k' \mathbf{C}_Z^{(0)} \boldsymbol{\xi}_k)^{1/2} (\boldsymbol{\psi}_k' \mathbf{C}_X^{(0)} \boldsymbol{\psi}_k)^{1/2}} \quad (5.36)$$

- For $k = 1$ we can write:

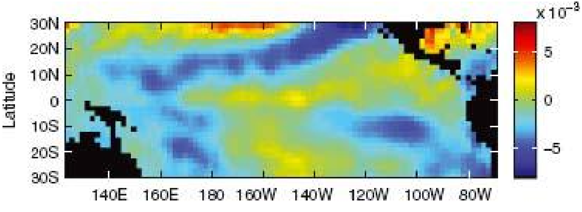
$$r_1^2 = \frac{[\tilde{\xi}_1'(\mathbf{C}_Z^{(0)})^{-1/2}\mathbf{C}_{Z,X}^{(0)}(\mathbf{C}_X^{(0)})^{-1/2}\tilde{\psi}_1]^2}{(\tilde{\xi}_1'\tilde{\xi}_1)(\tilde{\psi}_1'\tilde{\psi}_1)}$$

  where $\tilde{\xi}_1 \equiv (\mathbf{C}_Z^{(0)})^{1/2}\xi_1$ and $\tilde{\psi}_1 = (\mathbf{C}_X^{(0)})^{1/2}\psi_1$.
- Turns out that r_1^2 is the largest singular value in the singular value decomposition of

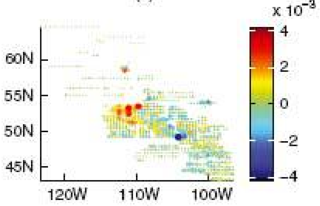
$$(\mathbf{C}_Z^{(0)})^{-1/2}\mathbf{C}_{Z,X}^{(0)}(\mathbf{C}_X^{(0)})^{-1/2}$$

  where $\tilde{\xi}_1$ and $\tilde{\psi}_1$ are the left and right singular vectors.
- The vectors ψ_1 and ξ_1 can then be calculated and the time series of canonical variables $a_t(1) \equiv \xi_1'\mathbf{Z}_t$ and $b_t(1) \equiv \psi_1'\mathbf{X}_t$ can be obtained.

Figure 5.26. First CCA patterns for SST and Mallard counts

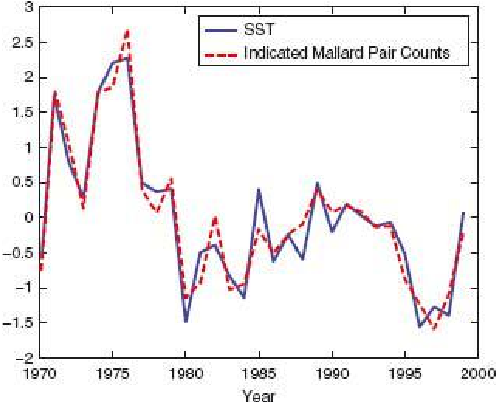


(a)



(b)

Figure 5.27. First canonical variables for SST and Mallard counts



References

- Cressie, N., and C. K. Wikle. 2011. Statistics for spatio-temporal data. John Wiley & Sons, New Jersey.