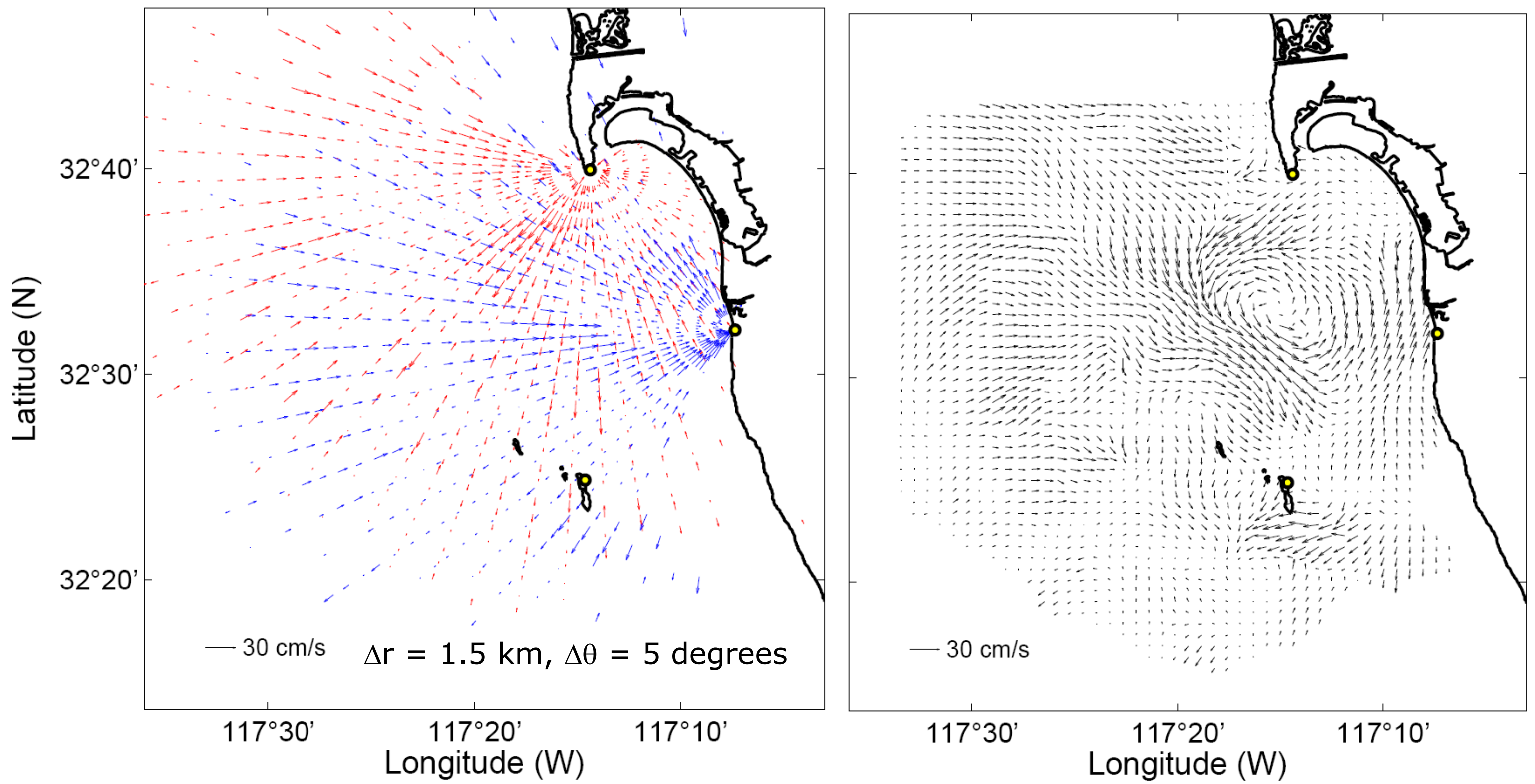
A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities



Jang Gon Yoo and Sung Yong Kim, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

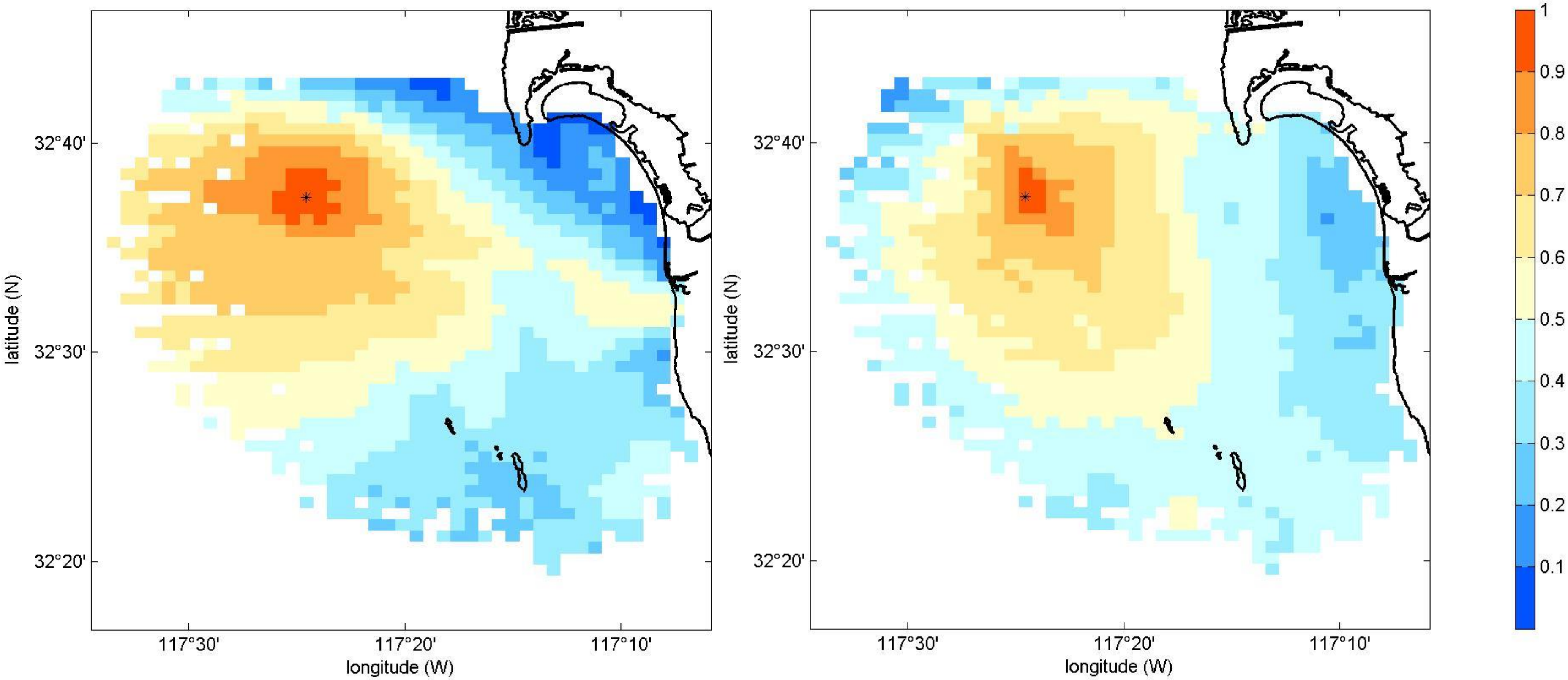
Bruce Cornuelle, Scripps Institution of Oceanography, La Jolla, USA

P. M. Kosro and Alex L. Kurapov, Oregon State University, Corvallis, USA

KAIST

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

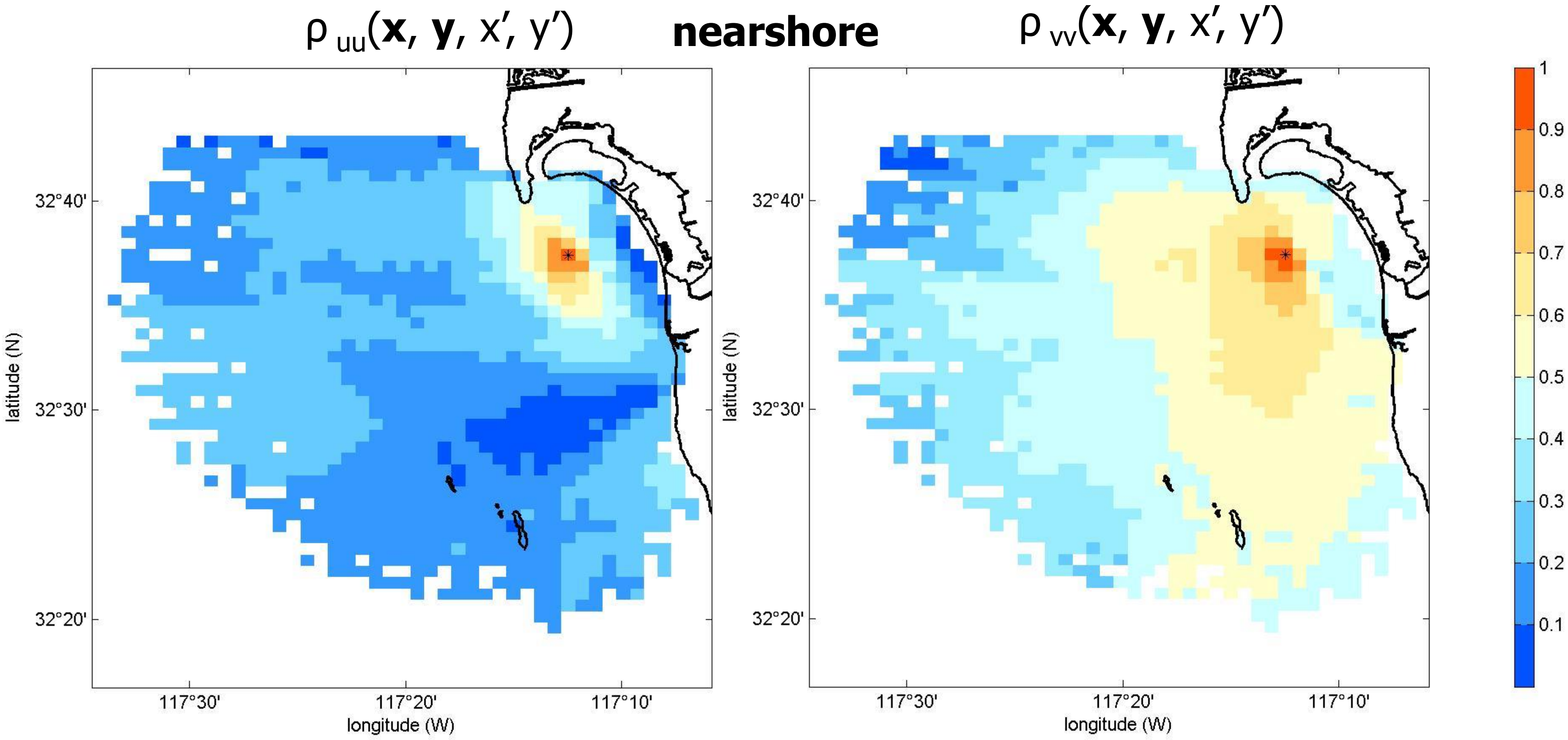
$\rho_{uu}(\mathbf{x}, \mathbf{y}, x', y')$ **offshore** $\rho_{vv}(\mathbf{x}, \mathbf{y}, x', y')$



- Correlations of the 2-year vector current data

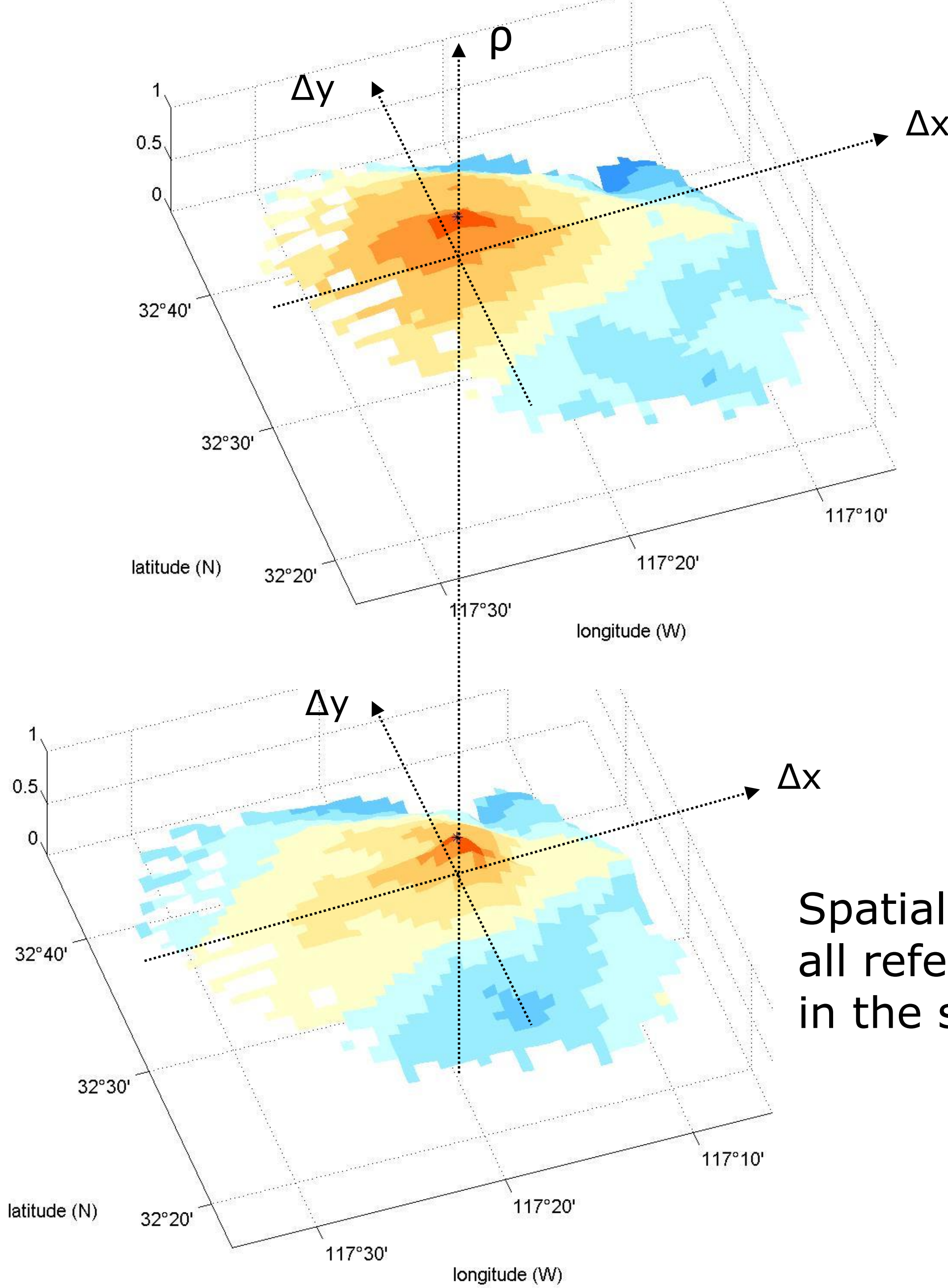
(Kim *et al*. JGR, 2007)

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities



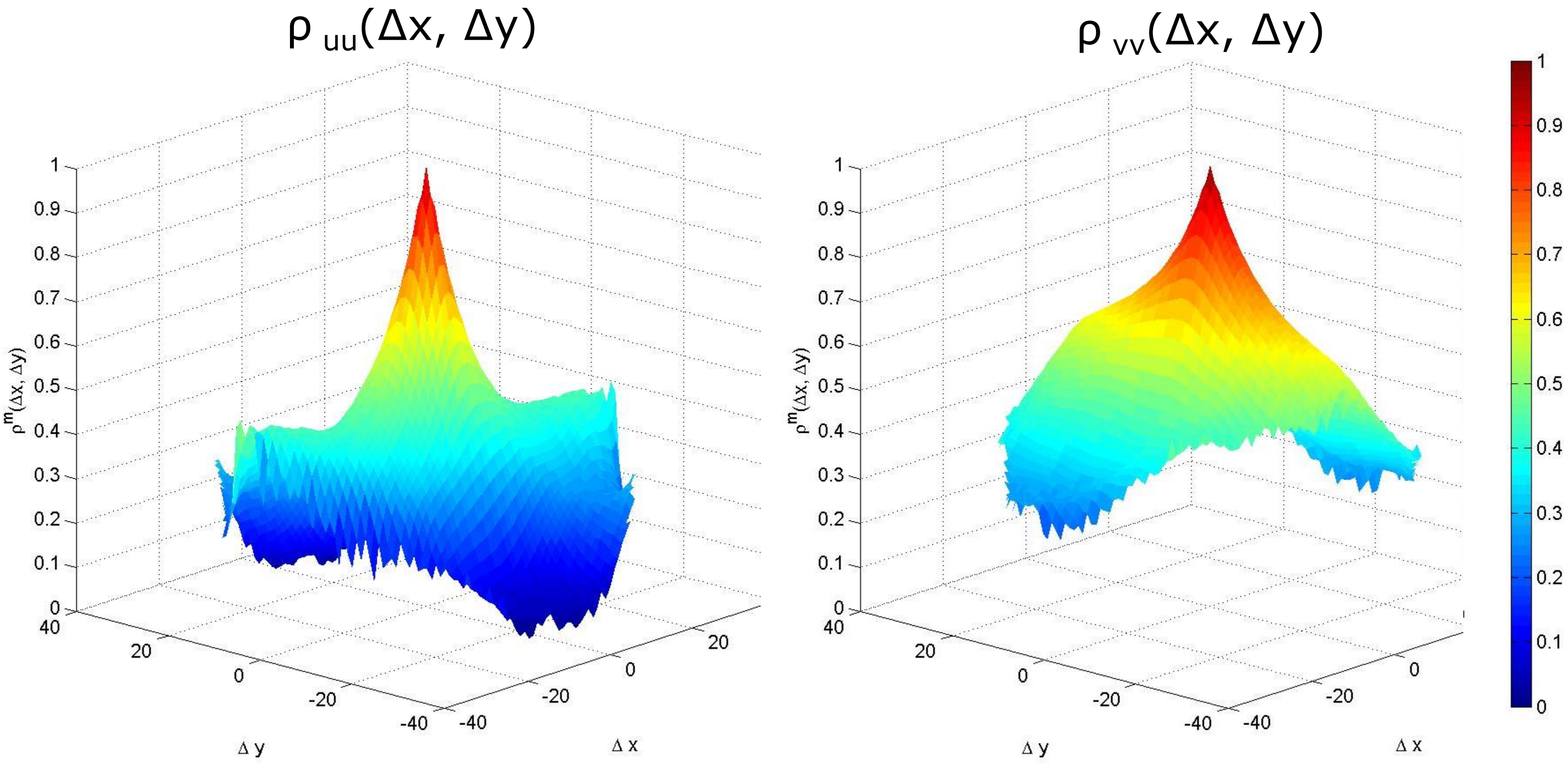
- De-correlation scale is the function of space (x, y).

(Kim *et al*. JGR, 2007)



Spatial lag average over all reference grid points in the study domain

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

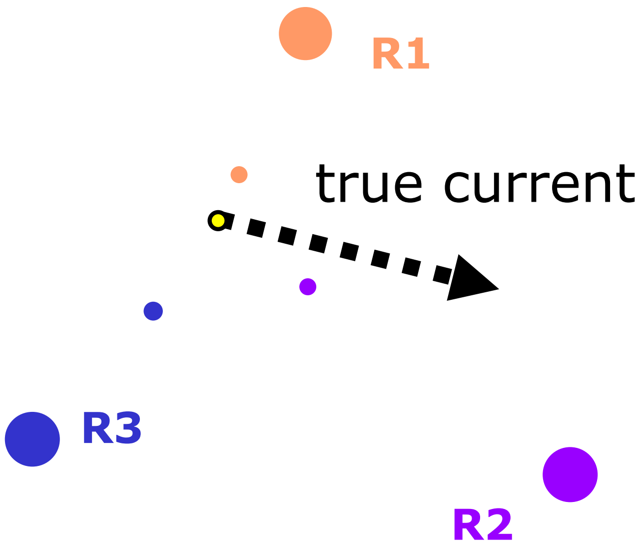


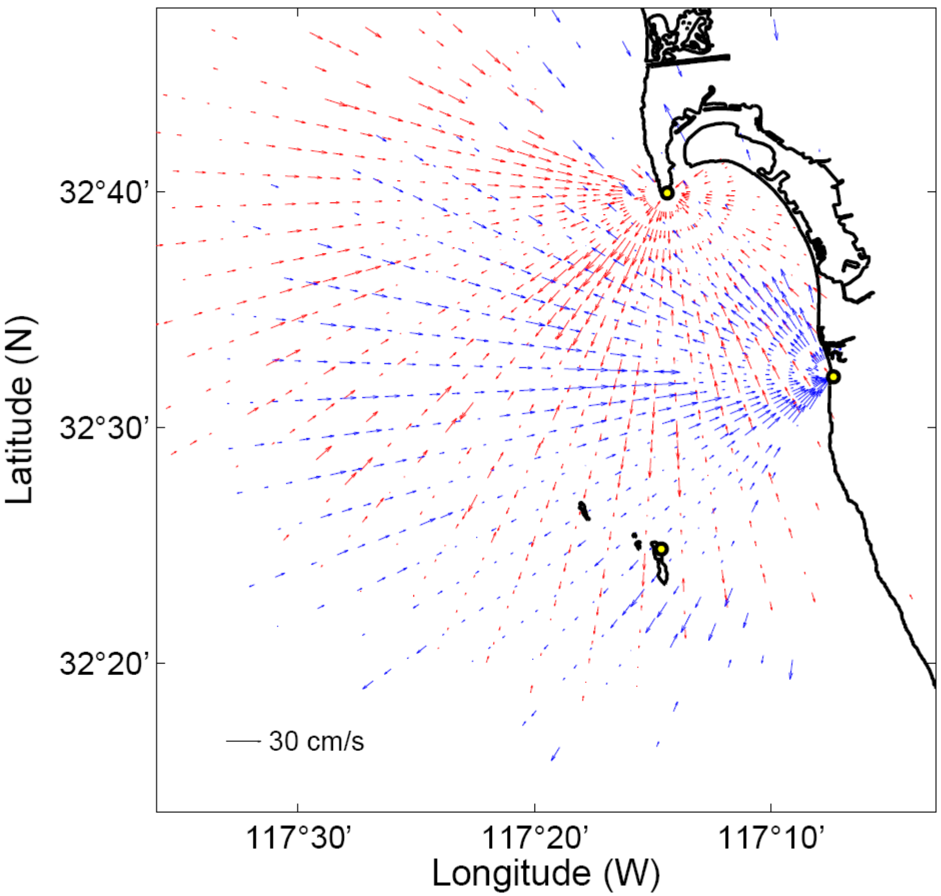
- Spatially composite correlation over the study domain
- Exponential shape (not Gaussian)

(Kim *et al*. JGR, 2007)

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

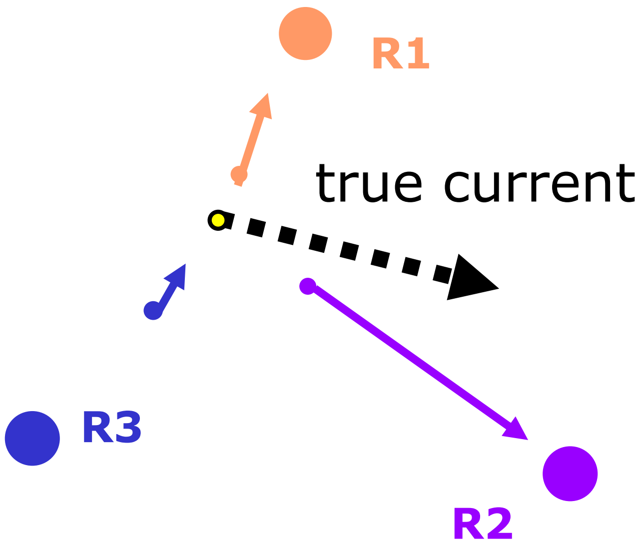
$$r(t) = \mathbf{g}^{\dagger}\mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$

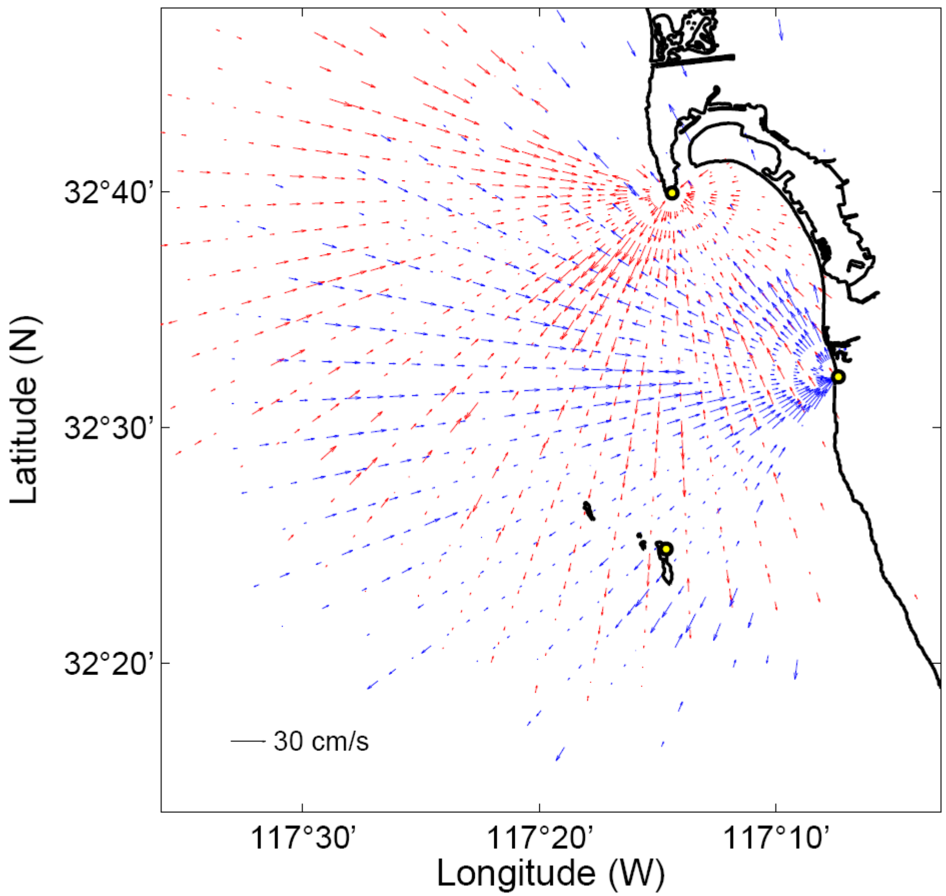




A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

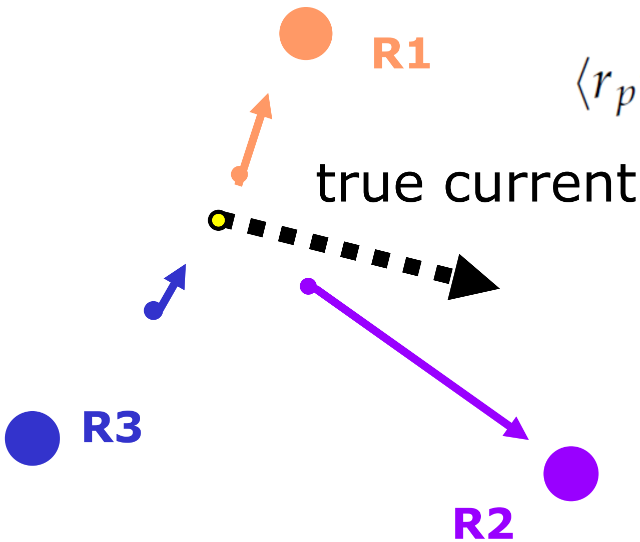
$$r(t) = \mathbf{g}^{\dagger}\mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$





A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

$$r(t) = \mathbf{g}^{\dagger}\mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$



$$\langle r_p r_q^{\dagger}\rangle = \mathbf{g}_p^{\dagger}\langle \mathbf{w}_p \mathbf{w}_q^{\dagger}\rangle \mathbf{g}_q, \quad (11)$$

$$= \cos\theta_p \cos\theta_q \langle u_p u_q^{\dagger}\rangle + \cos\theta_p \sin\theta_q \langle u_p v_q^{\dagger}\rangle$$

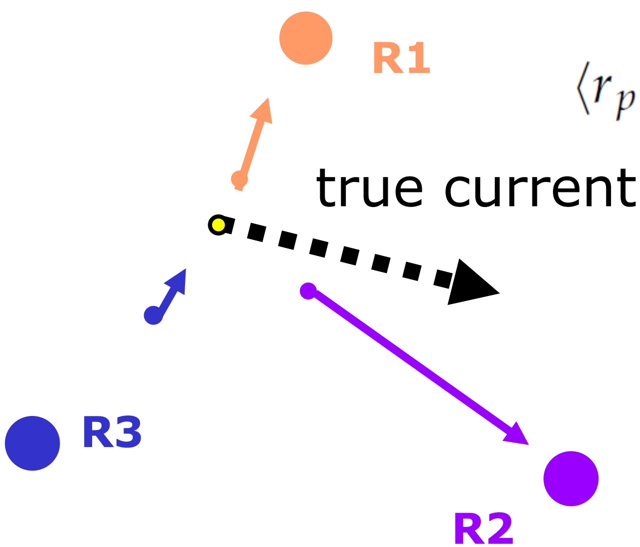
$$+ \sin\theta_p \cos\theta_q \langle v_p u_q^{\dagger}\rangle + \sin\theta_p \sin\theta_q \langle v_p v_q^{\dagger}\rangle + \langle \epsilon_p \epsilon_q^{\dagger}\rangle.$$

$$\mathrm{vec}\left(\langle \mathbf{r}\mathbf{r}^{\dagger}\rangle\right) = \alpha \mathrm{vec}\left(\langle \mathbf{u}\mathbf{u}^{\dagger}\rangle\right) + \beta \mathrm{vec}\left(\langle \mathbf{u}\mathbf{v}^{\dagger}\rangle\right) + \gamma \mathrm{vec}\left(\langle \mathbf{v}\mathbf{u}^{\dagger}\rangle\right) + \delta \mathrm{vec}\left(\langle \mathbf{v}\mathbf{v}^{\dagger}\rangle\right) + \mathrm{vec}\left(\langle \epsilon\epsilon^{\dagger}\rangle\right), \quad (13)$$

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}, \quad (14)$$

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

$$r(t) = \mathbf{g}^\dagger \mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$



$$\langle r_p r_q^\dagger \rangle = \mathbf{g}_p^\dagger \langle \mathbf{w}_p \mathbf{w}_q^\dagger \rangle \mathbf{g}_q, \tag{11}$$

$$= \cos\theta_p \cos\theta_q \langle u_p u_q^\dagger \rangle + \cos\theta_p \sin\theta_q \langle u_p v_q^\dagger \rangle$$

$$+ \sin\theta_p \cos\theta_q \langle v_p u_q^\dagger \rangle + \sin\theta_p \sin\theta_q \langle v_p v_q^\dagger \rangle + \langle \epsilon_p \epsilon_q^\dagger \rangle.$$

$$\mathrm{vec}\left(\langle \mathbf{r}\mathbf{r}^\dagger \rangle\right) = \alpha \mathrm{vec}\left(\langle \mathbf{u}\mathbf{u}^\dagger \rangle\right) + \beta \mathrm{vec}\left(\langle \mathbf{u}\mathbf{v}^\dagger \rangle\right)$$

$$+ \gamma \mathrm{vec}\left(\langle \mathbf{v}\mathbf{u}^\dagger \rangle\right) + \delta \mathrm{vec}\left(\langle \mathbf{v}\mathbf{v}^\dagger \rangle\right) + \mathrm{vec}\left(\langle \epsilon\epsilon^\dagger \rangle\right), \tag{13}$$

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}, \tag{14}$$

$$\langle u(\mathbf{x}_p) u(\mathbf{x}_q)^\dagger \rangle = \sum_{m=-M}^{M} \sum_{n=-N}^{N} A_{mn} e^{i\mathbf{k}_{mn}\Delta\mathbf{x}_{pq}},$$

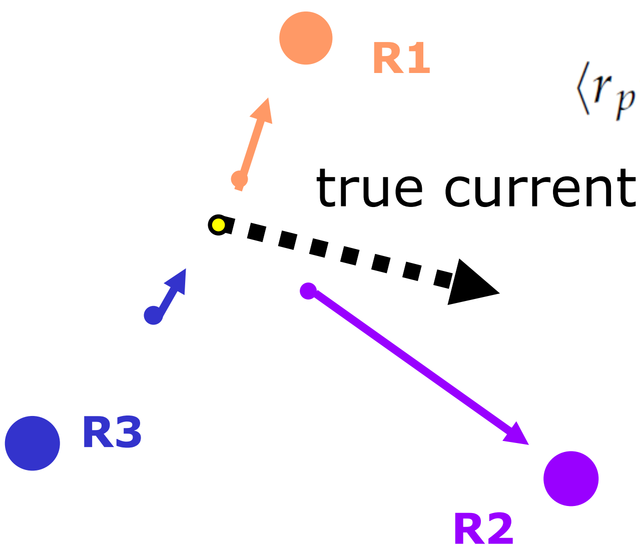
where

$$\mathbf{k}_{mn} = (k_m, l_n),$$

$$\Delta\mathbf{x}_{pq} = (x_p - x_q, y_p - y_q).$$

A direct estimate of horizontal spatial covariance from non-orthogonally sampled surface velocities

$$r(t) = \mathbf{g}^{\dagger}\mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$



$$\begin{aligned}\langle r_p r_q^{\dagger}\rangle &= \mathbf{g}_p^{\dagger}\langle \mathbf{w}_p \mathbf{w}_q^{\dagger}\rangle \mathbf{g}_q, \\ &= \cos\theta_p\cos\theta_q\langle u_p u_q^{\dagger}\rangle + \cos\theta_p\sin\theta_q\langle u_p v_q^{\dagger}\rangle \\ &+ \sin\theta_p\cos\theta_q\langle v_p u_q^{\dagger}\rangle + \sin\theta_p\sin\theta_q\langle v_p v_q^{\dagger}\rangle + \langle \epsilon_p \epsilon_q^{\dagger}\rangle.\end{aligned} \quad (11)$$

$$\begin{aligned}\mathrm{vec}\left(\langle \mathbf{r}\mathbf{r}^{\dagger}\rangle\right) &= \alpha\,\mathrm{vec}\left(\langle \mathbf{u}\mathbf{u}^{\dagger}\rangle\right) + \beta\,\mathrm{vec}\left(\langle \mathbf{u}\mathbf{v}^{\dagger}\rangle\right) \\ &+ \gamma\,\mathrm{vec}\left(\langle \mathbf{v}\mathbf{u}^{\dagger}\rangle\right) + \delta\,\mathrm{vec}\left(\langle \mathbf{v}\mathbf{v}^{\dagger}\rangle\right) + \mathrm{vec}\left(\langle \epsilon\epsilon^{\dagger}\rangle\right),\end{aligned} \quad (13)$$

L=134; d = 9045x1; M = 100; N = 100;

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}, \quad (14)$$

The number of radial grid points — The number of basis functions in the x and y directions

TABLE 1. Dimensions of the vectors and matrices in the direct estimate of variance and covariance.

Notation	Variance	Covariance
$\mathbf{d}$	$L \times 1$	$L(L+1)/2 \times 1$
$\mathbf{m}$	$3(2M+1)(2N+1) \times 1$	$4(2M+1)(2N+1) \times 1$
$\mathbf{G}$	$L \times 3(2M+1)(2N+1)$	$L(L+1)/2 \times 4(2M+1)(2N+1)$

Motivation

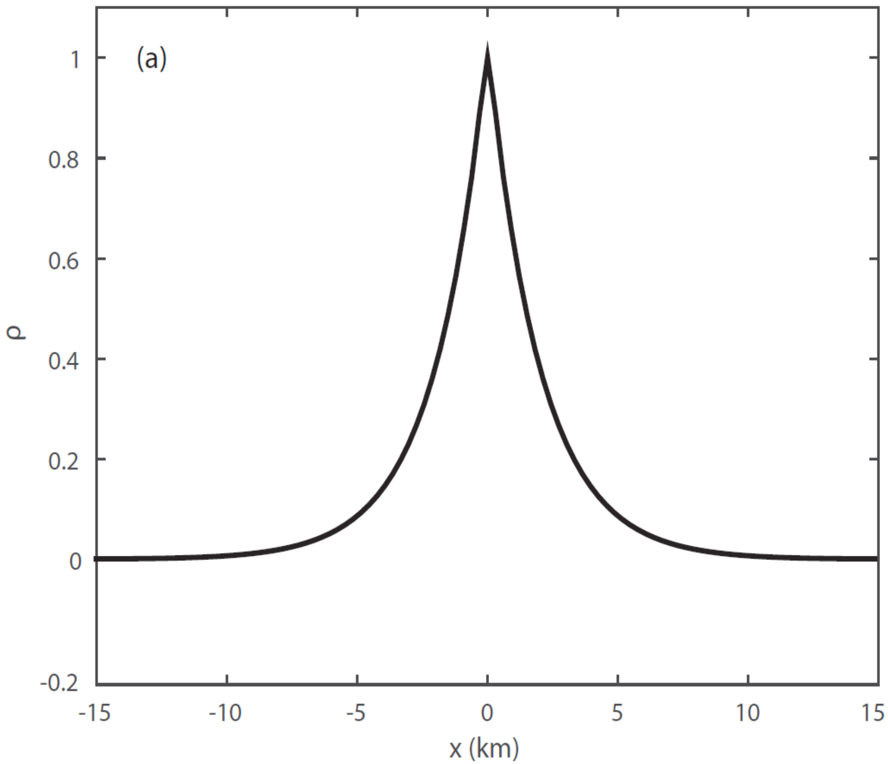
- When we describe the variability of the spatial-temporal data (fields or system) and characterize them, we may examine their covariance structure and decorrelation scales.
- However, the data may not be evenly sampled in space.
- Mapping of data on regularly spaced grid may be required, so the spatial covariance/correlation of the mapped fields contains a bias associated with given assumptions.
  - Correlation: a normalized structure of covariance
- So, how can we directly estimate covariance of unevenly sampled data in space?

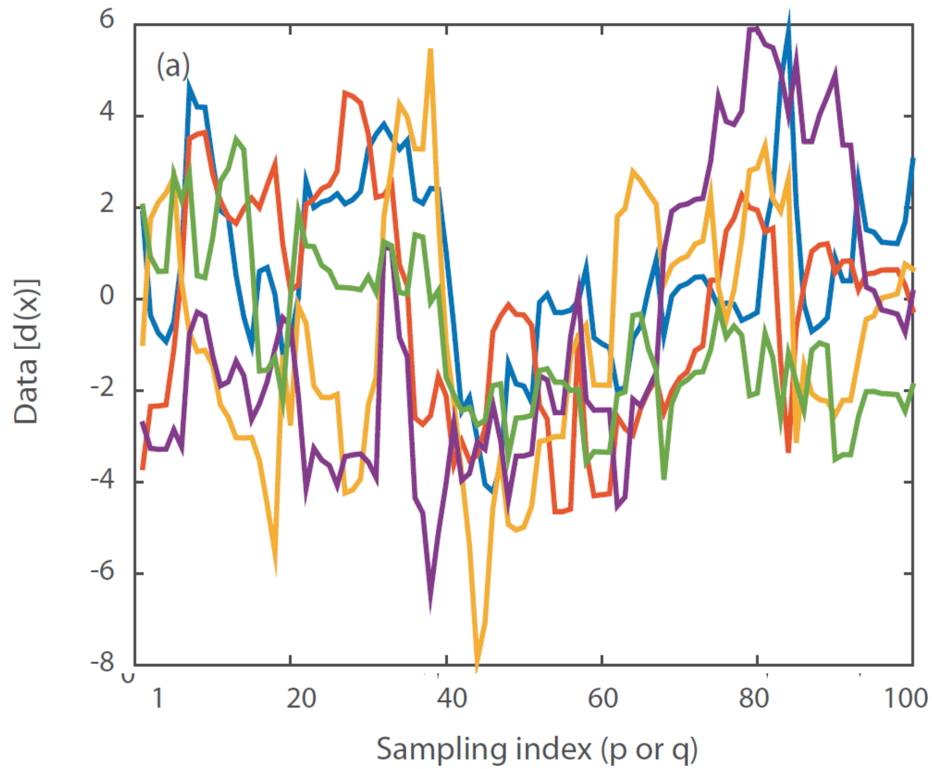
Covariance estimates (1D, scalar)

- Covariance vs. Energy spectra
  - Exponential function with a decorrelation length scale of 2km
  - Unevenly sampled data (plotted on the sampling index, not the physical domain

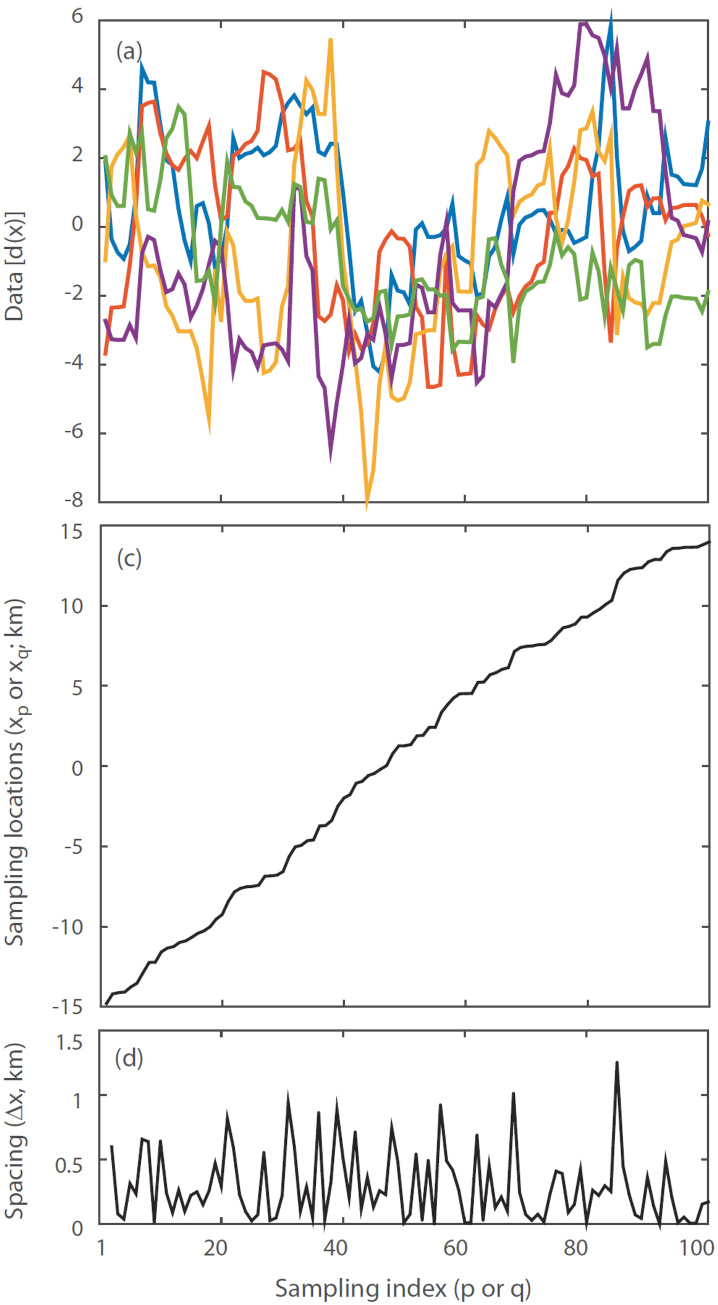
$$\int_{-\infty}^{\infty} e^{-\frac{|x|}{\lambda}} e^{-ikx} \mathrm{d}x = \frac{2\lambda}{1+\lambda^2 k^2}$$

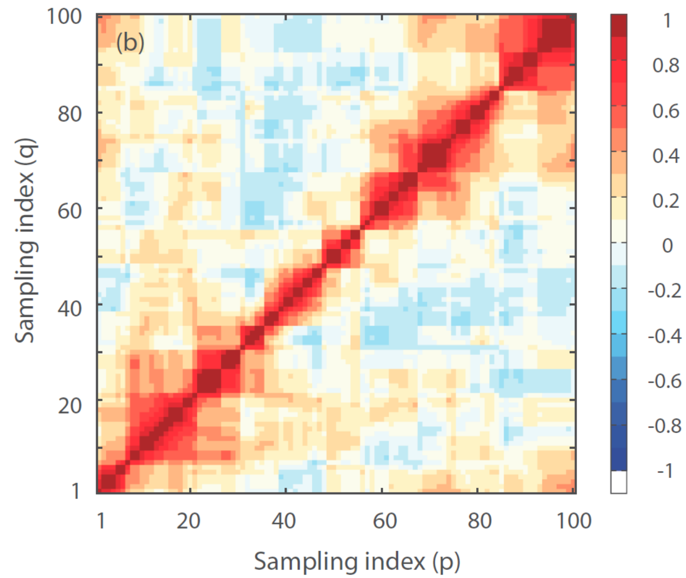
$$d(x,t) = \sum_m a_m(t)\cos(k_m x) + b_m(t)\sin(k_m x),$$
$$= \sum_m \chi_m(t)\phi_m(x),$$





Covariance estimates (1D, scalar)





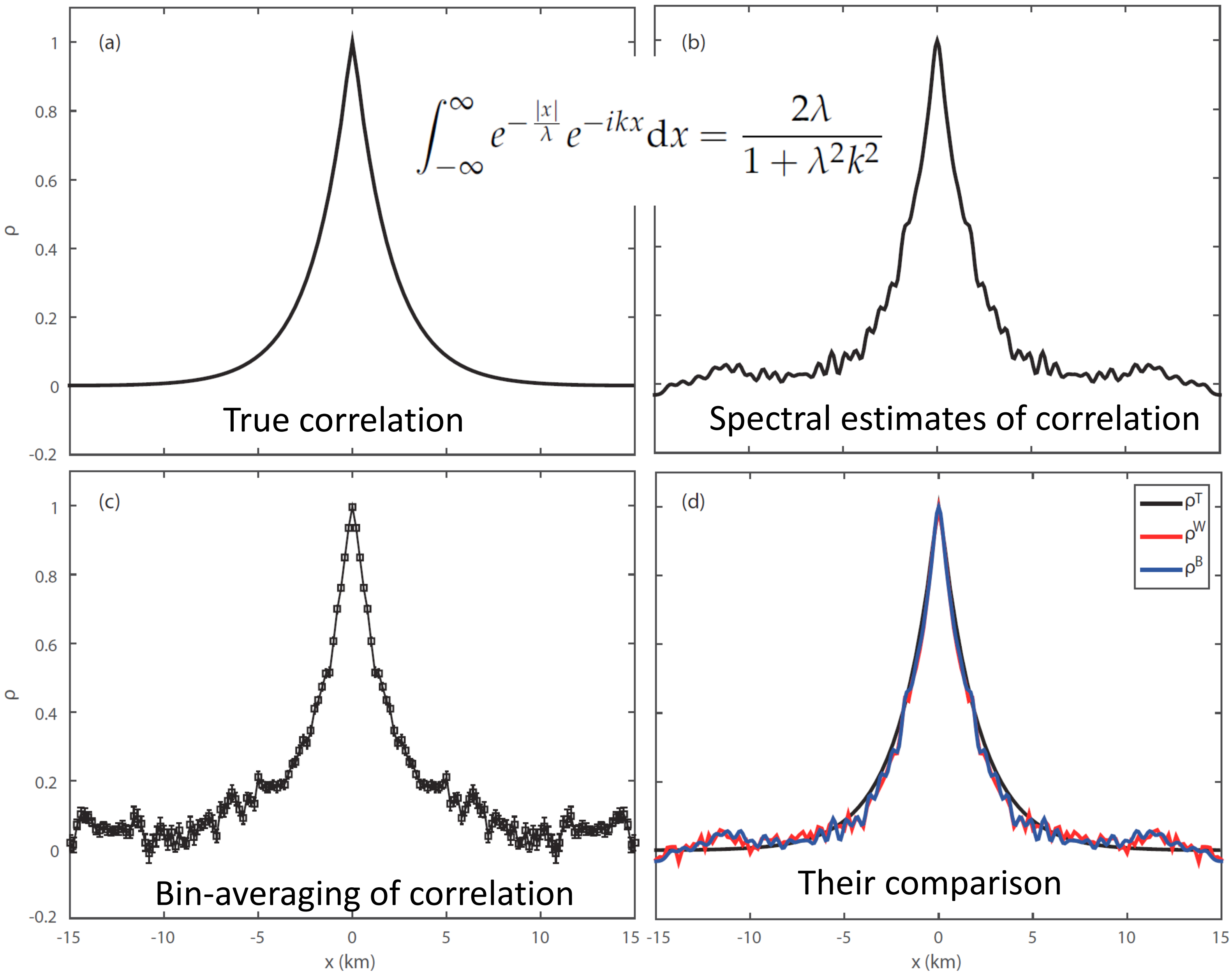
$$\xi_m^2(k) = \xi^2(k_m) = \frac{2\lambda}{1 + k_m^2\lambda^2},$$

$$\begin{aligned} d(x,t) &= \sum_m a_m(t)\cos(k_m x) + b_m(t)\sin(k_m x), \\ &= \sum_m \chi_m(t)\phi_m(x), \end{aligned}$$

$$\begin{aligned} c(x_p, x_q) &= \langle d(x_p,t)\, d(x_q,t)^\dagger \rangle, \\ &= \sum_m \langle \chi_m \chi_m^\dagger \rangle\, \phi_m(x_p)\phi_m(x_q), \\ &= \hat{c}(\Delta x), \qquad \hat{c}(\Delta x) = \sum_n A_n \cos(k_n \Delta x), \end{aligned}$$

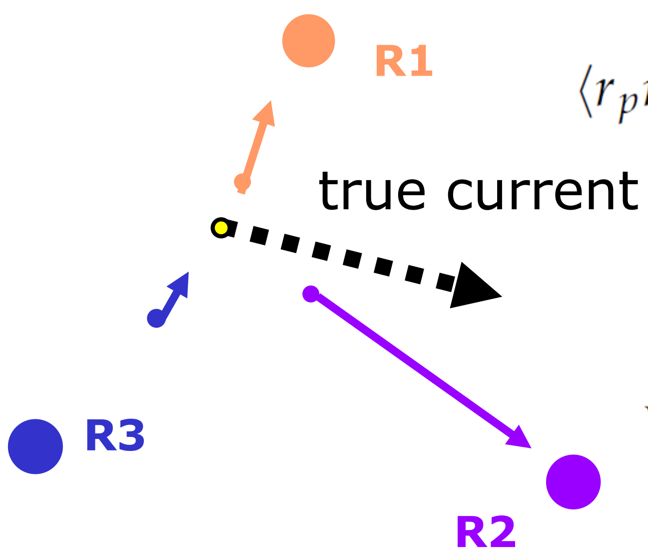
Assumptions: Stationary and homogeneous fields

Covariance estimates (1D, scalar)



$$\int_{-\infty}^{\infty} e^{-\frac{|x|}{\lambda}} e^{-ikx} \mathrm{d}x = \frac{2\lambda}{1+\lambda^2 k^2}$$

Covariance estimate (2D, vector)



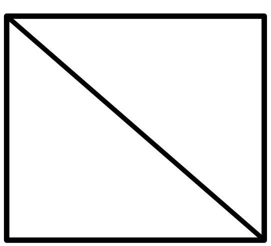
$$r(t) = \mathbf{g}^{\dagger}\mathbf{w}(t) + \epsilon(t) = u(t)\cos\theta + v(t)\sin\theta + \epsilon(t),$$

$$\langle r_p r_q^{\dagger}\rangle = \mathbf{g}_p^{\dagger}\langle \mathbf{w}_p \mathbf{w}_q^{\dagger}\rangle \mathbf{g}_q, \tag{11}$$

$$= \cos\theta_p \cos\theta_q \langle u_p u_q^{\dagger}\rangle + \cos\theta_p \sin\theta_q \langle u_p v_q^{\dagger}\rangle$$
$$+ \sin\theta_p \cos\theta_q \langle v_p u_q^{\dagger}\rangle + \sin\theta_p \sin\theta_q \langle v_p v_q^{\dagger}\rangle + \langle \epsilon_p \epsilon_q^{\dagger}\rangle.$$

$$\mathrm{vec}\left(\langle \mathbf{r}\mathbf{r}^{\dagger}\rangle\right) = \alpha\mathrm{vec}\left(\langle \mathbf{u}\mathbf{u}^{\dagger}\rangle\right) + \beta\mathrm{vec}\left(\langle \mathbf{u}\mathbf{v}^{\dagger}\rangle\right) + \gamma\mathrm{vec}\left(\langle \mathbf{v}\mathbf{u}^{\dagger}\rangle\right) + \delta\mathrm{vec}\left(\langle \mathbf{v}\mathbf{v}^{\dagger}\rangle\right) + \mathrm{vec}\left(\langle \epsilon\epsilon^{\dagger}\rangle\right), \tag{13}$$

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}, \tag{14}$$

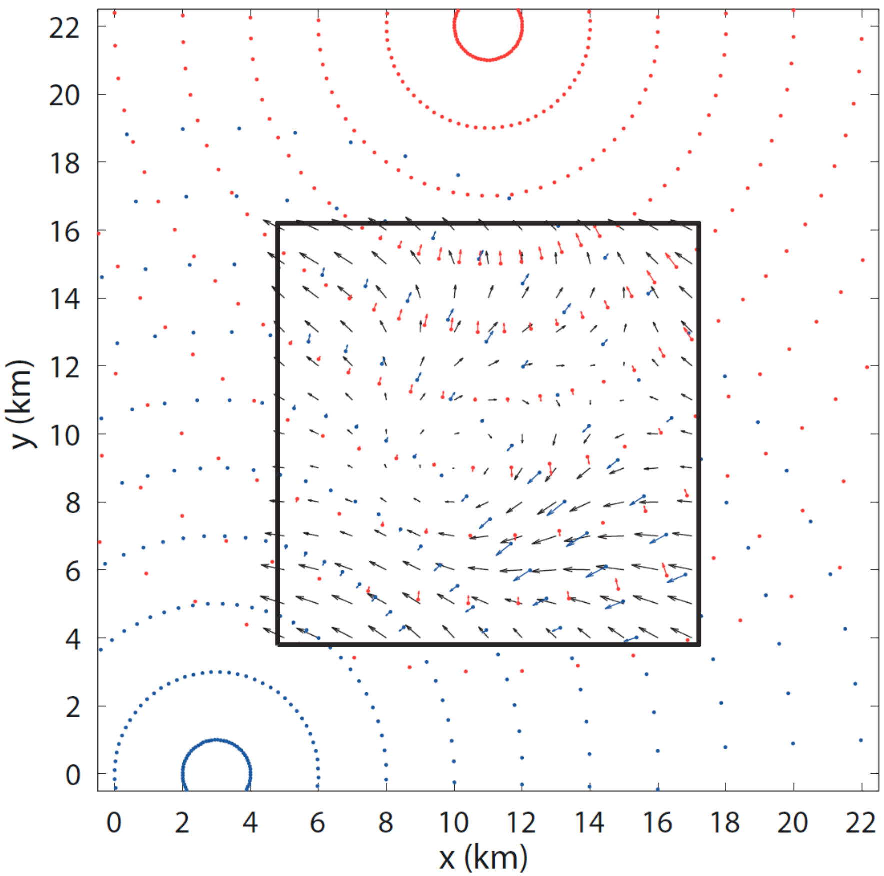


L=134; d = 9045x1; M = 100;

TABLE 1. Dimensions of the vectors and matrices in the direct estimate of variance and covariance.

Notation	Variance	Covariance
$\mathbf{d}$	$L \times 1$	$L(L+1)/2 \times 1$
$\mathbf{m}$	$3(2M+1)(2N+1) \times 1$	$4(2M+1)(2N+1) \times 1$
$\mathbf{G}$	$L \times 3(2M+1)(2N+1)$	$L(L+1)/2 \times 4(2M+1)(2N+1)$

An idealized and spectral model



$$u(x,y,t) = \sum_{m=-M^*}^{M^*} \sum_{n=-N^*}^{N^*} \sum_{s=-S^*}^{S^*} \hat{A}_{mns} \cos \vartheta_{mns} + \hat{B}_{mns} \sin \vartheta_{mns}, \tag{A9}$$

$$v(x,y,t) = \sum_{m=-M^*}^{M^*} \sum_{n=-N^*}^{N^*} \sum_{s=-S^*}^{S^*} \hat{C}_{mns} \cos \vartheta_{mns} + \hat{D}_{mns} \sin \vartheta_{mns}, \tag{A10}$$

where

$$\vartheta_{mns} = k_m x + l_n y - \sigma_s t = 2\pi \left(\frac{m}{L_x} x + \frac{n}{L_y} y \right) - \sigma_s t, \tag{A11}$$

$$\hat{A}_{mns} = (\xi_{mn})^{\frac{1}{2}} \xi_s \mathrm{N}(0,1), \tag{A12}$$

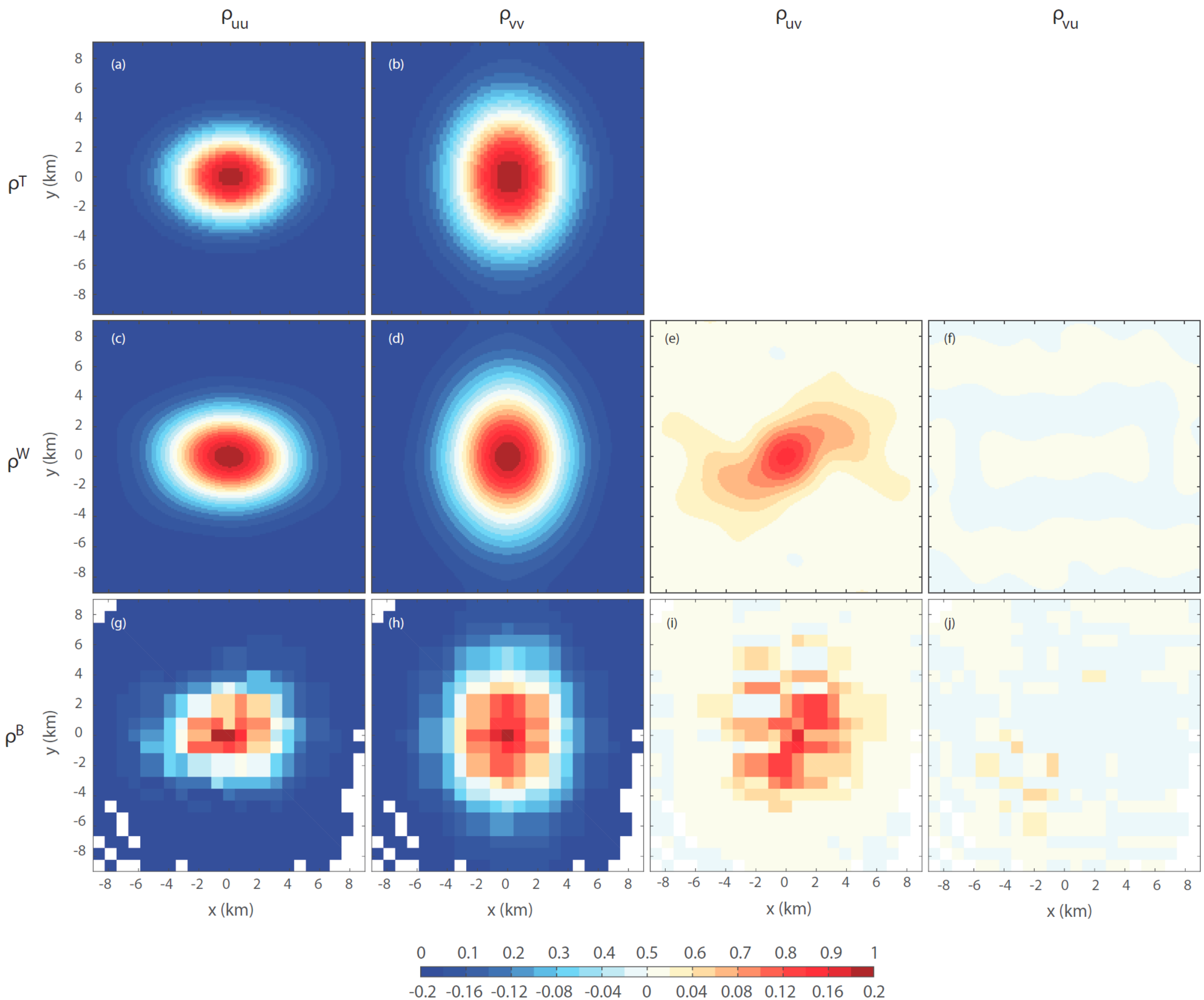
with the power spectrum of the current field in the wavenumber $[\xi_{mn} = \xi(k_m, l_n)]$ and frequency $[\xi_s = \xi(\sigma_s)]$ domains. The wavenumber spectrum can be approximated with a two-dimensional Gaussian or exponential function:

$$\xi(k_m, l_n) = \pi \lambda_x \lambda_y \exp\left(-k_m^2 \lambda_x^2 - l_n^2 \lambda_y^2\right), \tag{A13}$$

$$\xi(k_m, l_n) = \frac{4\lambda_x \lambda_y}{(1 + k_m^2 \lambda_x^2 + l_n^2 \lambda_y^2)^{3/2}}, \tag{A14}$$

and λ_x and λ_y denote the de-correlation length scales in the x and y directions, respectively.

Covariance estimates (2D, vector)



Numerical model domain

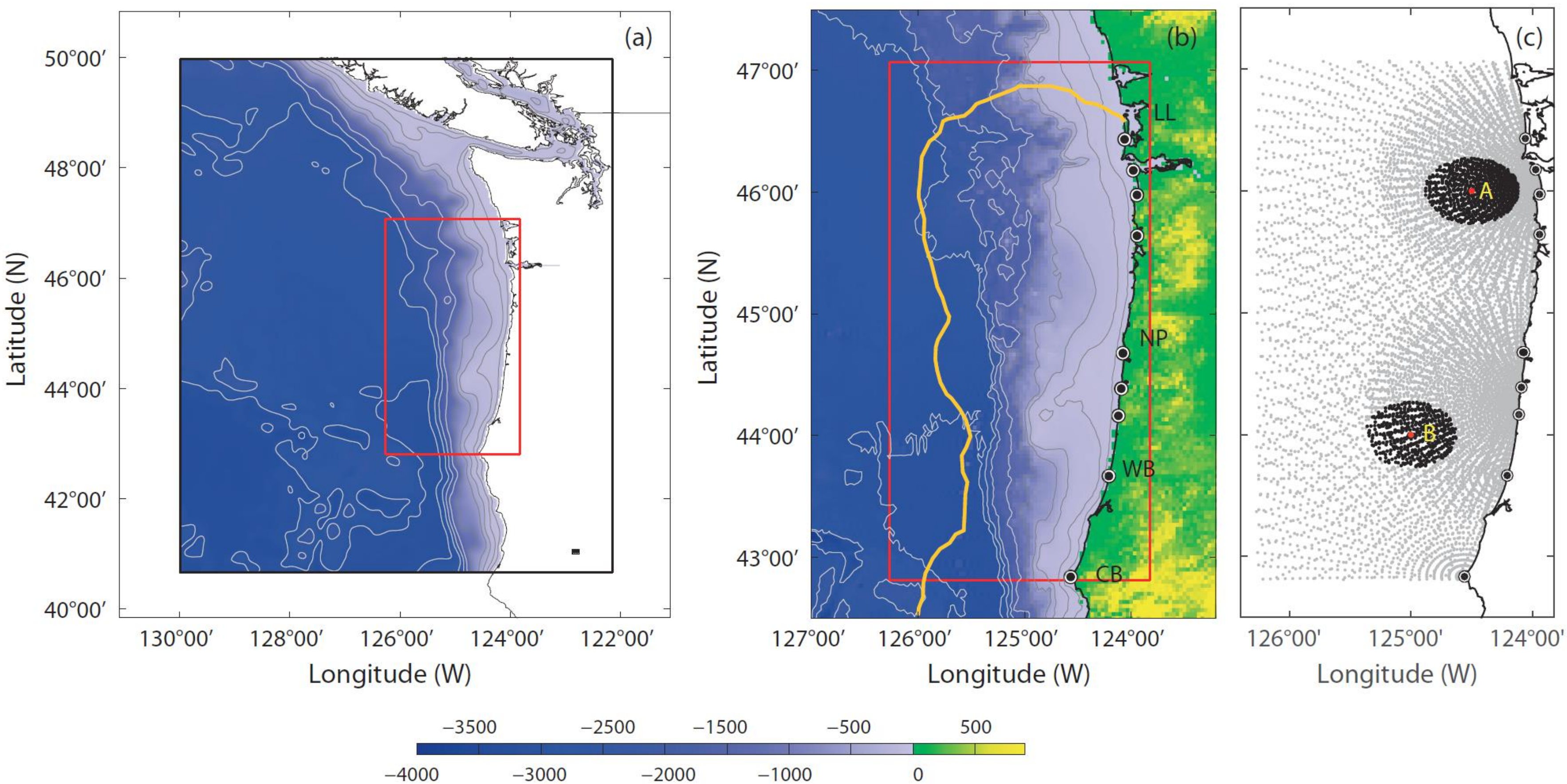
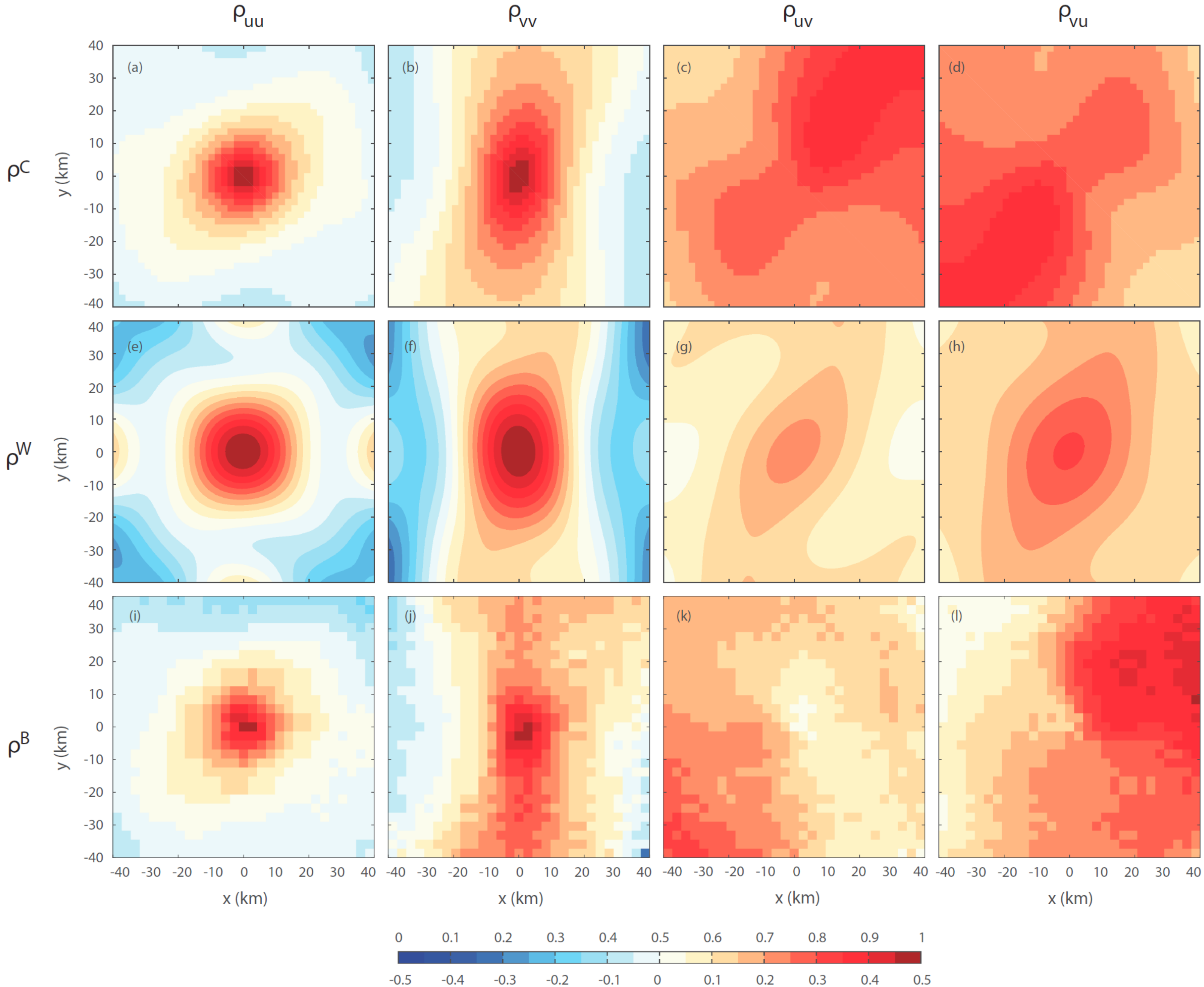


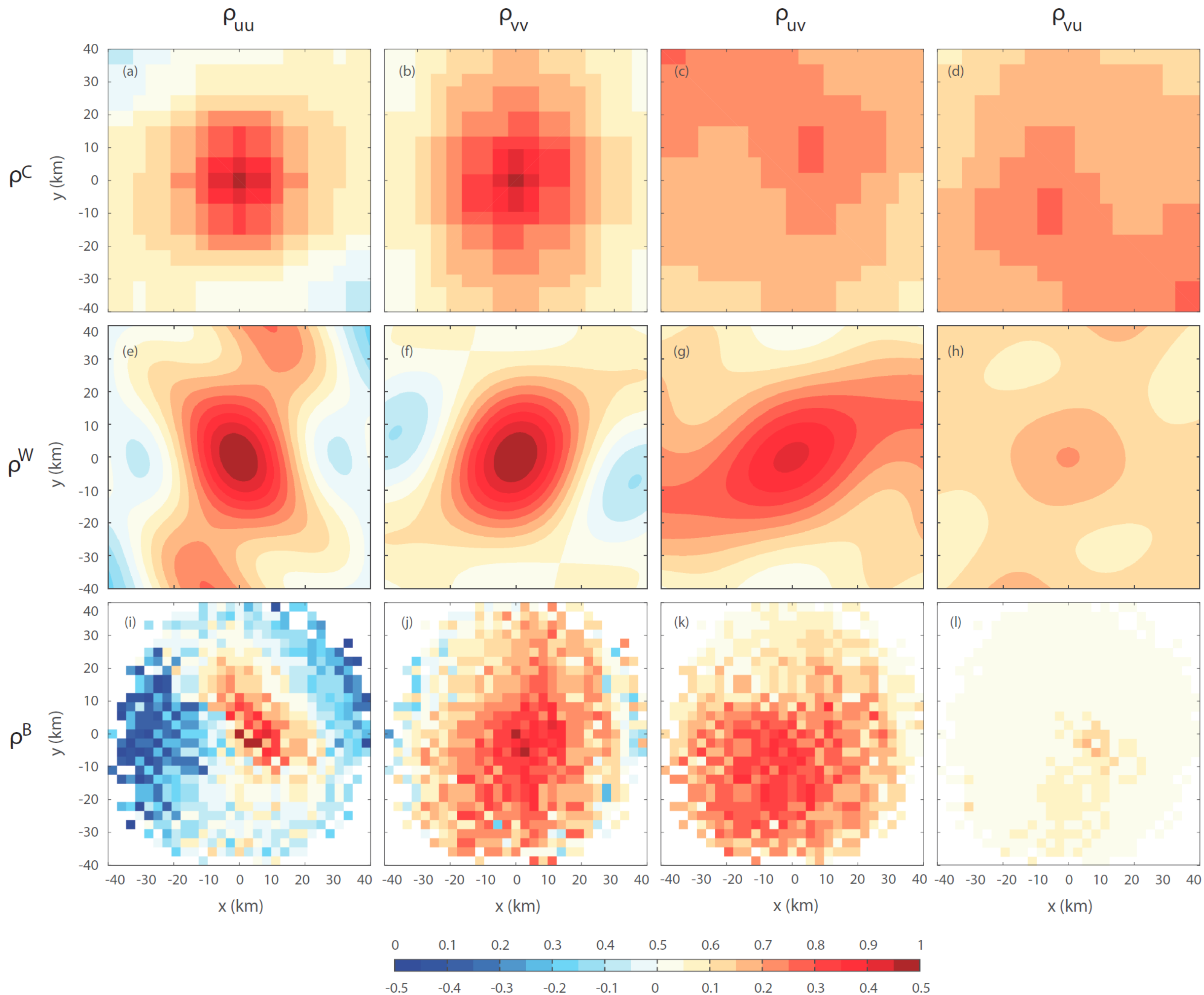
TABLE 1. Dimensions of the vectors and matrices in the direct estimate of variance and covariance.

Notation	Variance	Covariance
$\mathbf{d}$	$L \times 1$	$L(L+1)/2 \times 1$
$\mathbf{m}$	$3(2M+1)(2N+1) \times 1$	$4(2M+1)(2N+1) \times 1$
$\mathbf{G}$	$L \times 3(2M+1)(2N+1)$	$L(L+1)/2 \times 4(2M+1)(2N+1)$

Covariance estimates (2D, vector) - model



Covariance estimates (2D, vector) - observations



Summary

- A direct method of spatial covariance estimates on the scalar and vector data can minimize the bias due to an intermediate step of gridding.
- Solving the inverse problem can be computationally expensive, but we may use advanced computational resources.
- The number of realizations and the density of unevenly samplings (sparse vs. dense) can affect the performance of the proposed analysis, which should be examined with 1D/simple examples.