Proceedings of the 7th International Conference on Statistics: Theory and Applications (ICSTA'25)
Paris, France - August, 2025
Paper No. 150
DOI: 10.11159/icsta25.150

Asymmetric Cross-correlation of Multivariate Spatial Stochastic Processes: A Primer

Xiaoqing Chen[1]
[1]Department of Mathematics and Statistics, University of Exeter
Stocker Rd, Exeter, EX4 4PY, Exeter, United Kingdom
xiaoqing.a.chen@gmail.com

Abstract **-** Multivariate spatial phenomena are ubiquitous, spanning domains such as climate, pandemics, air quality, and social economy. Cross-correlation between different quantities of interest at different locations is asymmetric in general. Such asymmetric cross-correlation is a fundamental yet often overlooked phenomenon in multivariate spatial statistics. This paper provides the visualisation, structure, and properties of asymmetric cross-correlation. It reviews mainstream multivariate spatial models and analyzes their capability to accommodate asymmetric cross-correlation. Finally, it demonstrates the impact of accounting for asymmetric cross-correlation on model accuracy using a one-dimensional simulated example.

Keywords**:** asymmetric cross-correlation, symmetric autocorrelation, multivariate spatial stochastic processes, visualization, structure, properties

1. Introduction

Multivariate spatial phenomena are ubiquitous across disciplines. Examples include climate variables (temperature, precipitation, wind speed), pandemic indicators (protein mutation rate, UV radiation intensity, vaccination coverage), air quality measures (PM2.5, NO_2, O_3), and socioeconomic factors (crime rates, housing prices, income levels). The common characteristic of these phenomena is that each quantity at a location s_i interacts with: (1) itself at nearby locations (same-variate auto-correlation); (2) other quantities at the same location (same-location auto-correlation); (3) other quantities at nearby locations (cross-correlation).

Auto-correlations (types (1) and (2)) are symmetric by definition. For example, the correlation between PM2.5 at locations s_i and s_j is identical to the correlation between PM2.5 at s_j and s_i. Similarly, the correlation between PM2.5 and O_3 at a single location s_i is the same as the correlation between O_3 and PM2.5 at s_i.

In contrast, cross-correlation (type (3)) is generally asymmetric. For instance, the cross-correlation between PM2.5 in Saudi Arabia and Sea Salt (SS) in Egypt differs from the cross-correlation between PM2.5 in Egypt and SS in Saudi Arabia. Mathematically, $corr(X(s_i), Y(s_j)) \neq corr(X(s_j), Y(s_i))$, where X represents PM2.5 and Y represents SS.

Figure 1 shows the empirical same-variate auto-correlation matrix plot of PM2.5, displaying the $corr(PM2.5(s_i), PM2.5(s_j)) \equiv corr(PM2.5(s_j), PM2.5(s_i))$ across four equal-width longitude strips: [−180°,−90°), [−90°, 0°), [0°, 90°), [90°, 180°]. The symmetry is evident.

Figure 2 is the empirical cross-correlation plots. It displays $corr(PM2.5(s_i), PM2.5(s_j)) \neq corr(PM2.5(s_j), PM2.5(s_i))$ across four equal-width longitude strips. The asymmetry is prominent.

Lon: [−180°, −90°) Lon: [−90°, 0°) Lon: [0°, 90°) Lon: [90°, 180°]

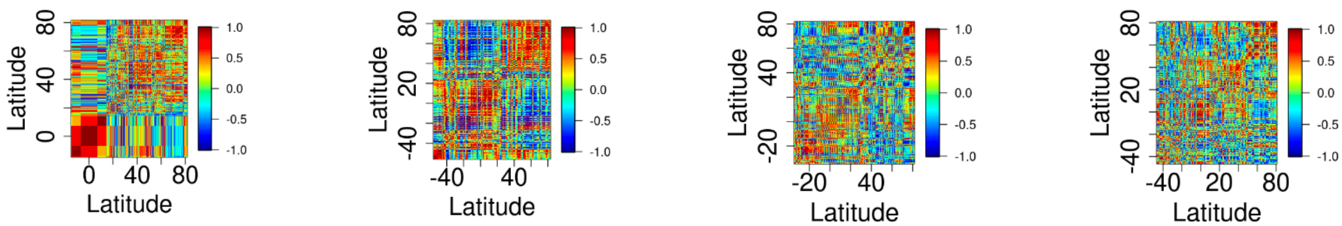

Fig. 1: Empirical same-variate auto-correlation matrix plots for PM2.5 across four longitude strips. All symmetric about y = x.

Lon: [−180°, −90°) Lon: [−90°, 0°) Lon: [0°, 90°) Lon: [90°, 180°]

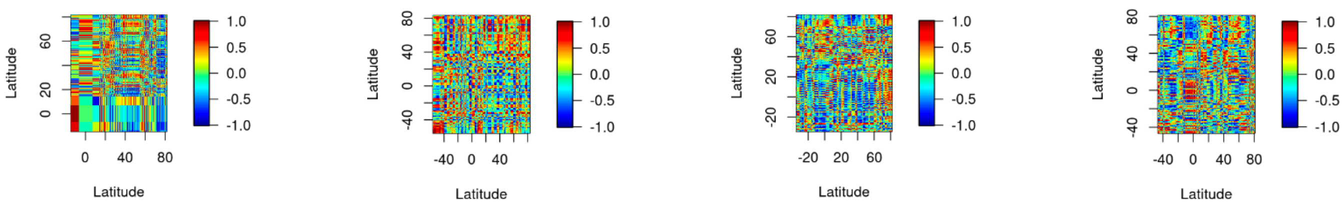

Fig. 2: Empirical cross-correlation matrix plots for PM2.5 and Sea Salt across four longitude strips. All asymmetric.

2. Structure of Joint Covariance Matrix

Consider data from a multivariate spatial stochastic process $\{(Y_1(s_i), \ldots, Y_p(s_i)) : i = 1, \ldots, n\}$.

Let $\mathbf{Y} = [\mathbf{Y}_1^T(\cdot), \ldots, Y_p^T(\cdot)]^T = [\mathbf{Y}_1^T, \ldots, Y_n^T]^T$ represent a vector of np random variables, where $\mathbf{Y}_1(\cdot) \in \mathbb{R}^n$, while $\mathbf{Y}_1 \in \mathbb{R}^p$. The joint covariance matrix $\Sigma_{np \times np} = \left[cov\left(Y_l(s_i), Y_k(s_j)\right)\right] = \left[C^{lk}(s_i, s_j)\right]$, where $C^{lk}(s_i, s_j)$ is

- a cross-covariance if $l \neq k$ and $i \neq j$
- a same-variate auto-covariance if l = k but $i \neq j$
- a same-location auto-covariance if $l \neq k$ but i = j.

The joint covariance matrix can be organised in two ways: one is location-based, and the other is variate-based. For the former, the structure of the joint covariance matrix is shown in Figure 3, and for the latter, the structure of the joint covariance matrix is shown in Figure 4.

$$\Sigma_{np\times np} = \begin{bmatrix} [cov(\boldsymbol{Y}(s_1),\boldsymbol{Y}(s_1))]_{p\times p} & [cov(\boldsymbol{Y}(s_1),\boldsymbol{Y}(s_2))]_{p\times p} & \cdots & [cov(\boldsymbol{Y}(s_1),\boldsymbol{Y}(s_n))]_{p\times p} \\ [cov(\boldsymbol{Y}(s_2),\boldsymbol{Y}(s_1))]_{p\times p} & [cov(\boldsymbol{Y}(s_2),\boldsymbol{Y}(s_2))]_{p\times p} & \cdots & [cov(\boldsymbol{Y}(s_2),\boldsymbol{Y}(s_n))]_{p\times p} \\ \vdots & \vdots & \ddots & \vdots \\ [cov(\boldsymbol{Y}(s_n),\boldsymbol{Y}(s_1))]_{p\times p} & [cov(\boldsymbol{Y}(s_n),\boldsymbol{Y}(s_2))]_{p\times p} & \cdots & [cov(\boldsymbol{Y}(s_n),\boldsymbol{Y}(s_n))]_{p\times p} \end{bmatrix}_{np\times np},$$

Fig. 3: Structure of the joint covariance matrix, with block matrices grouped by locations.

$$\Sigma_{np\times np} = \begin{bmatrix} [cov(\boldsymbol{Y}_1(\boldsymbol{s}),\boldsymbol{Y}_1(\boldsymbol{s}))]_{n\times n} & [cov(\boldsymbol{Y}_1(\boldsymbol{s}),\boldsymbol{Y}_2(\boldsymbol{s}))]_{n\times n} & \cdots & [cov(\boldsymbol{Y}_1(\boldsymbol{s}),\boldsymbol{Y}_p(\boldsymbol{s}))]_{n\times n} \\ [cov(\boldsymbol{Y}_2(\boldsymbol{s}),\boldsymbol{Y}_1(\boldsymbol{s}))]_{n\times n} & [cov(\boldsymbol{Y}_2(\boldsymbol{s}),\boldsymbol{Y}_2(\boldsymbol{s}))]_{n\times n} & \cdots & [cov(\boldsymbol{Y}_2(\boldsymbol{s}),\boldsymbol{Y}_p(\boldsymbol{s}))]_{n\times n} \\ \vdots & \vdots & \ddots & \vdots \\ [cov(\boldsymbol{Y}_p(\boldsymbol{s}),\boldsymbol{Y}_1(\boldsymbol{s}))]_{n\times n} & [cov(\boldsymbol{Y}_p(\boldsymbol{s}),\boldsymbol{Y}_2(\boldsymbol{s}))]_{n\times n} & \cdots & [cov(\boldsymbol{Y}_p(\boldsymbol{s}),\boldsymbol{Y}_p(\boldsymbol{s}))]_{n\times n} \end{bmatrix}_{np\times np},$$

Fig. 4: Structure of the joint covariance matrix, with block matrices grouped by variates.

For a clear illustration, we expand one of the off-diagonal blocks $[cov(Y(s_1), Y(s_2))]_{p\times p}$ in Fig. 3 and $[cov(Y_1(s), Y_2(s))]_{n\times n}$ in Fig. 4 to expose the asymmetry. Specifically, the off-diagonal block $[cov(Y(s_1), Y(s_2))]_{p\times p}$ is shown in Fig. 5.

$$[cov(\boldsymbol{Y}(s_1), \boldsymbol{Y}(s_2))]_{p\times p} = \begin{bmatrix} C^{11}(s_1, s_2) & C^{12}(s_1, s_2) & \cdots & C^{1p}(s_1, s_2) \\ C^{21}(s_1, s_2) & C^{22}(s_1, s_2) & \cdots & C^{2p}(s_1, s_2) \\ \vdots & \vdots & \ddots & \vdots \\ C^{p1}(s_1, s_2) & C^{p2}(s_1, s_2) & \cdots & C^{pp}(s_1, s_2) \end{bmatrix}_{p\times p},$$

Fig. 5: Structure of off-diagonal block $[cov(Y(s_1), Y(s_2))]_{p\times p}$.

Here, $cov(Y_l(s_1), Y_p(s_2)) \triangleq C^{lp}(s_1, s_2) \neq C^{pl}(s_1, s_2) \triangleq cov(Y_p(s_1), Y_l(s_2))$ in general.
And the off-diagonal block $[cov(Y_1(s), Y_2(s))]_{n\times n}$ is shown below in Fig. 6.

$$[cov(\boldsymbol{Y}_1(\boldsymbol{s}), \boldsymbol{Y}_2(\boldsymbol{s}))]_{n\times n} = \begin{bmatrix} C^{12}(s_1, s_1) & C^{12}(s_1, s_2) & \cdots & C^{12}(s_1, s_n) \\ C^{12}(s_2, s_1) & C^{12}(s_2, s_2) & \cdots & C^{12}(s_2, s_n) \\ \vdots & \vdots & \ddots & \vdots \\ C^{12}(s_n, s_1) & C^{12}(s_n, s_2) & \cdots & C^{12}(s_n, s_n) \end{bmatrix}_{n\times n},$$

Fig. 6: Structure of off-diagonal block $[cov(Y_1(s), Y_2(s))]_{n\times n}$

Here, $cov(Y_1(s_1), Y_2(s_n)) \triangleq C^{12}(s_1, s_n) \neq C^{12}(s_n, s_1) \triangleq cov(Y_1(s_n), Y_2(s_1))$ in general.

In contrast, diagonal blocks in Fig. 3 and 4 representing auto-covariance are always symmetric. Specifically, for the same-location auto-covariance: $cov(Y_l(s_1), Y_p(s_1)) \triangleq C^{lp}(s_1, s_1) = C^{pl}(s_1, s_1) \triangleq cov(Y_p(s_1), Y_1(s_1))$. And for the same-variate auto-covariance: $cov(Y_1(s_1), Y_1(s_n)) \triangleq C^{11}(s_1, s_n) = C^{11}(s_n, s_1) \triangleq cov(Y_1(s_n), Y_1(s_1))$.

It is worth noting that while off-diagonal blocks representing cross-covariance are generally asymmetric, the joint covariance matrix $\Sigma_{np\times np}$ itself must be positive definite and therefore symmetric as a whole. The asymmetry in the off-diagonal blocks must be symmetric across the main diagonal of $\Sigma_{np\times np}$.

3. Properties of Auto-correlation and Cross-correlation

The auto-/cross-correlation is defined as $Corr_{lk}(s_i, s_j) = \frac{C_{lk}(s_i, s_j)}{\sqrt{(C_{ll}(s_i, s_i))}\sqrt{(C_{kk}(s_j, s_j))}}$.

3.1. Auto-correlation Properties

- Symmetric: auto-correlation matrices are symmetric about the main diagonal. See Fig. 1.
- Maximum at diagonal: the main diagonal values $Corr^{ll}(s_i, s_i)$ of the auto-correlation matrix have the largest magnitude of 1, where l = 1, . . . p, and i = 1, . . . , n. See Fig. 1.
- Bounded magnitudes: the magnitude of the off-diagonal values in the auto-correlation matrix must be less than or equal to the main diagonal value of 1. Specifically,
    - $|Corr_{lk}(s_i, s_i)| \leq |Corr_{ll}(s_i, s_i)|$ for same-location auto-correlation.
    - $|Corr_{ll}(s_i, s_j)| \leq |Corr_{ll}(s_i, s_i)|$ for same-variate auto-correlation.
- Sign flexibility: auto-correlation can be both positive and negative. See Fig. 1.

3.2. Cross-correlation Properties

- Asymmetric: cross-correlation matrices are generally not symmetric. See Fig. 2, 5, 6.
- Maximum location flexibility: the largest magnitude (i.e., 1) does not necessarily occur on the main diagonal. See Fig. 2.
- Unbounded by diagonal: magnitudes of off-diagonal values are not necessarily less than main diagonal values. See Fig. 2.
- Overall bound: $\left|Corr_{lk}(s_i, s_j)\right| \leq \left|Corr_{ll}(s_i, s_i)\right|\left|Corr_{kk}(s_j, s_j)\right| \equiv 1$. See [1].
- Sign flexibility: cross-correlation can be both positive and negative. See Fig. 2.

4. A Criterion for Assessing Model Quality

Since the off-diagonal blocks representing cross-covariance in $\Sigma_{np\times np}$ are generally asymmetric, the capability to capture this asymmetry serves as one criterion for assessing multivariate spatial models. Below, we briefly review mainstream modelling methods for constructing the joint covariance matrix $\Sigma_{np\times np}$ and analyze their capability of incorporating the asymmetric cross-covariance.

4.1 Intrinsic Correlation Model

The intrinsic correlation model [2] (also called separable model [3]) decomposes a covariance matrix block at a given pair of locations as:

$$C(s_i, s_j) = \rho(s_i, s_j)\, V_{p\times p},$$

where $\rho(s_i, s_j)$ is the pure spatial correlation between locations and $V_{p\times p}$ is the variance-covariance matrix among p variates.

The joint covariance matrix then becomes a Kronecker product:

$$\Sigma_{np\times np} = H_{n\times n} \otimes V_{p\times p},$$

Since $V_{p\times p}$ is symmetric by definition, each off-diagonal block in $\Sigma_{np\times np}$ is a spatial correlation scalar times a symmetric matrix, making the off-diagonal block symmetric. Therefore, this model lacks a mechanism to accommodate asymmetric cross-covariance.

4.2 Kernel Convolution Approach

The kernel convolution approach [4] generates each variate field individually from a common underlying hidden process:

$$Y_l(s_i) = \sigma_l \int k_l(s_i - t)\, g(t)\, dt$$

where g(·) is a standard Gaussian process and $k_l(\cdot)$ is a square integrable kernel function.

Then the $(l, r)^{th}$ cross-covariance block has elements:

$$C^{lr}(s_i, s_j) = \sigma_l \sigma_r \int\int k_l(s_i - t) k_r(s_j - t') \rho(t - t')\, dt\, dt'$$

By introducing a shift parameter Δ to the location separation lag $s - t \triangleq h$, asymmetry can be accommodated since $h_{ij} - \Delta \neq h_{ji} - \Delta$.

4.3 Multivariate Matérn Approach

The multivariate Matérn method [5] models both the auto-correlation and cross-correlation using Matérn correlations. Specifically,

$$Corr_{ll}(h) = M(h; \nu_l, \kappa_l)$$
$$Corr_{lk}(h) = \beta_{lk} M(h; \nu_{lk}, \kappa_{lk}),$$

where M(·) represents the Matérn correlation function, with parameter ν controlling the small-scale smoothness near the origin, and κ controlling the rate of the correlation decay at a large spatial scale [6]. β_{lk} is the cross-correlation between two variates l, k regardless of location. Separation lag $h \in R^d$.

By introducing a shift parameter Δ to the spatial separation lag *h*, asymmetric cross-correlation can be accommodated.

4.4 Conditional Modelling Approaches

Mardia [7] proposed modelling each Y_i conditionally as:

$$E[Y_i | Y_{-i}] = \mu_i + \Sigma_{j \in N(i)} \beta_{ij} (Y_j - \mu_j); \quad Var[Y_i | Y_{-i}] = \Gamma_i$$

This method constructs the joint precision matrix $\Sigma^{-1}_{np \times np}$ rather than the covariance matrix, making the incorporation of asymmetric cross-covariance less straightforward and challenging.

Cressie [8] constructs the joint covariance matrix through conditional means and covariances:

$$E\left(Y_q(s_i) | Y_r(\cdot): r = 1,2,\ldots,(q-1)\right) = \sum_{r=1}^{q-1} \int_D b_{qr}(s_i, v) Y_r(v) dv$$

$$cov\left(Y_q(s_i), Y_q(s_j) | Y_r(\cdot): r = 1,2,\ldots,(q-1)\right) = C_{q|(r<q)}(s_i, s_j)$$

Similar to kernel convolution, the $b_{qr}(s_i, v)$ function can accommodate asymmetric cross-covariance by introducing a shift parameter Δ.

4.5 Remark

The mainstream multivariate models can be broadly categorized into two types. One is the unconditional type, which directly models the off-diagonal cross-covariance matrices. Examples include the *intrinsic correlation model* (Section 4.1), the *kernel convolution model* (Section 4.2) and the *multivariate Matérn model* (Section 4.3). The other type consists of conditional modelling methods discussed in Section 4.4.

The unconditional types can be further divided into two kinds: one is modelling the cross-covariance matrices in a $p \times p$ dimension (e.g., *intrinsic correlation model*), which lacks the mechanism to accommodate the asymmetry. The other is modelling the cross-covariance matrices in an $n \times n$ dimension (e.g., kernel convolution, multivariate Matérn), which can accommodate asymmetry through a spatial shift parameter Δ.

5. An Example

A 1D simulation was conducted using Cressie's conditional approach [8]. The spatial domain D = [-10, 10] was discretized with grid size 0.1, resulting in 200 locations. A tri-variate process was modelled across the domain.

The goal was to predict true process values at the first 50 locations of the first field, given the remaining noisy observations in the same field and 400 observations from the other two fields.

Prediction results were compared between models with and without a shift parameter Δ. Figure 7 shows that the model with Δ, which accounts for the asymmetry, produces more accurate prediction results (solid line) than the one without it (the dashed line).

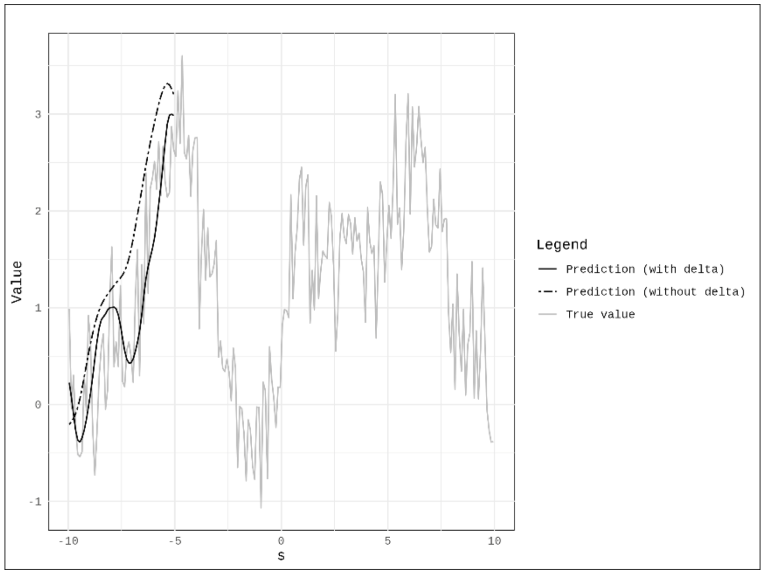

Fig. 7 Prediction results for the first 50 locations of the first field, given the noisy observations from the remaining locations in the first field and the 400 observations from the other two fields.

Table 1 compares model accuracy using two metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The model with the shift parameter Δ has lower MAE and RMSE than the one without.

Table 1: Compare MAE and RMSE between the Model with Δ and the One without Δ.

Model	MAE	RMSE
With Δ	0.245	0.289
Without Δ	0.706	0.740

6. Conclusion

Multivariate spatial phenomena are ubiquitous, spanning multiple disciplines. Quantities of interest not only interact with themselves at nearby regions but also interact with other quantities at nearby locations, as reflected by cross-correlation (or cross-covariance). The cross-correlation is generally asymmetric, and the capability to capture this asymmetry is one of the criteria for assessing model quality.

Admittedly, other factors like computational efficiency also matter in model assessment. Methods like the intrinsic correlation model (Section 4.1) and Mardia's conditional method (Section 4.4) may not readily capture asymmetric cross-covariance but offer computational efficiency. For example, Mardia's conditional approach produces a sparse precision matrix that speed up overall computation.

Asymmetric cross-covariance matrix blocks reside in the joint covariance matrix $\Sigma_{np \times np}$while the sparsity is present in the joint precision matrix $\Sigma_{np \times np}^{-1}$. Methods that can simultaneously address asymmetric cross-covariance and maintain computational efficiency (e.g., a sparse precision matrix) as in [9] are increasingly needed.

Acknowledgements

The author was supported by the Alan Turing Institute doctoral studentship *EP/N510129/1*. The author is grateful to Professor Peter Diggle for his kind discussion.

References

[1] H. Wackernagel, *Multivariate Geostatistics: An Introduction with Applications*, Springer Science & Business Media, 2013, p. 150.

[2] H. Wackernagel, *Multivariate Geostatistics: An Introduction with Applications*, Springer Science & Business Media, 2013, p. 155.
[3] S. Banerjee, B. P. Carlin, and A. E. Gelfand, *Hierarchical Modeling and Analysis for Spatial Data*, CRC Press, 2014, pp. 263.
[4] J. M. Ver Hoef and R. P. Barry, "Constructing and fitting models for cokriging and multivariable spatial prediction," *J. Stat. Plann. Inference*, vol. 69, no. 2, pp. 275–294, 1998.
[5] T. Gneiting, W. Kleiber, and M. Schlather, "Matérn cross-covariance functions for multivariate random fields," *J. Amer. Stat. Assoc.*, vol. 105, no. 491, pp. 1167–1177, 2010.
[6] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, Springer Science & Business Media, 1999, p. 32.
[7] K. V. Mardia, "Multi-dimensional multivariate Gaussian Markov random fields with application to image processing," *J. Multivariate Anal.*, vol. 24, no. 2, pp. 265–284, 1988.
[8] N. Cressie and A. Zammit-Mangion, "Multivariate spatial covariance models: a conditional approach," *Biometrika*, vol. 103, no. 4, pp. 915–935, 2016.
[9] X. Chen, P. Diggle, J. V. Zidek, and G. Shaddick, *Highly Multivariate Large-scale Spatial Stochastic Processes -- A Cross-Markov Random Field Approach*, arXiv:2408.10396, 2025. [Online]. Available: https://arxiv.org/abs/2408.10396