Modelling of Anisotropic Spatial Random Fields Using Mixture Copulas

[Mohomed Abraj, Helen Thompson and You-Gan Wang] [Early Career and Student Statisticians Conference 2021]

Introduction

In spatial studies, one of the key assumptions is that the spatial process is isotropic. Although this assumption considerably reduces the complexity in the modelling and prediction process, ignoring the directional dependence in spatial studies can lead to an inappropriate understanding of the physical phenomenon under study as well as making incorrect inferences. We proposed a novel spatial copula method that enables to model spatial random fields while incorporating directional dependence.

Methodology

Spatial random field (SRF)

SRF is the spatial data that are continuously indexed and evaluated at a set of discrete locations in space (denote, $Z(x)$ is a SRF at location x).

Copula (Sklar,1959)

Copula is a joint distribution function (H) of two or more random variables (Y_1,Y_2), in a bivariate context, copula $C : [0,1]^2 \to [0,1]$ is given by,

$$H(y_1, y_2) = C(G_1(y_1), G_2(y_2)) \quad (1)$$

where G_1 and G_2 are cumulative distribution functions (CDFs) of Y_1 and Y_2.

Spatial copula (Bärdossy, 2006)

A spatial copula is a joint distribution function of Z at selected locations x and $x+r$, for any quantiles u and v, is given by

$$C_{\theta,k}(u,v) = C(F(Z(x)), F(Z(x+r))) \quad (2)$$

where F is the marginal CDF of Z, and r is the separation distance. Also, k is the number spatial bins and θ is the direction of interest.

Mixture Copula

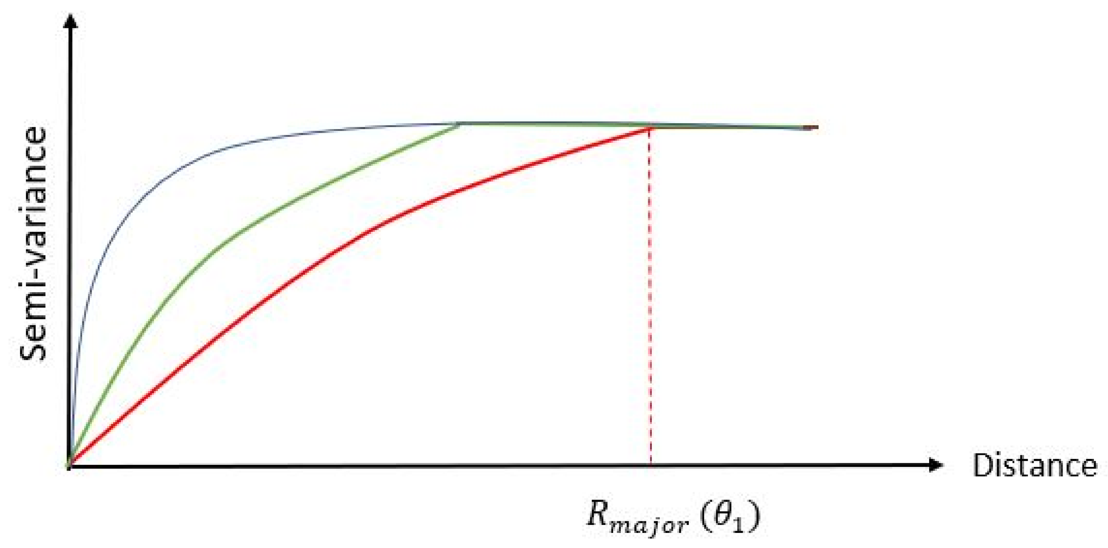
A mixture copula is a weighted combination of copulas with different directions (for $C_{\theta_1,k}$ and $C_{\theta_2,k}$), is given by C_k^m

$$C_k^m(u,v) = (1-w)C_{\theta_1,k} + wC_{\theta_2,k}$$

where w is a weight and $0 < w < 1$.

Anisotropy detection

Suppose $Z(x)$ is a univariate SRF with response Z which is collected at two-dimensional location x. Identify geometric anisotropy of $Z(x)$ using directional variograms for different directions.



The variogram having the biggest range (R_{major}) is selected as the major direction of anisotropy (say θ_1).

If (a_i, b_i) and (a_j, b_j) are coordinates of any pairs of locations at i and j. A rotation (transformation) matrix is used to select pairs of points that are orientated along or nearly to θ_1. Then, the transformed points (a'_i, b'_i) and (a'_j, b'_j) are given by

$$\begin{pmatrix} a'_i & a'_j \\ b'_i & b'_j \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} a_i & a_j \\ b_i & b_j \end{pmatrix}$$

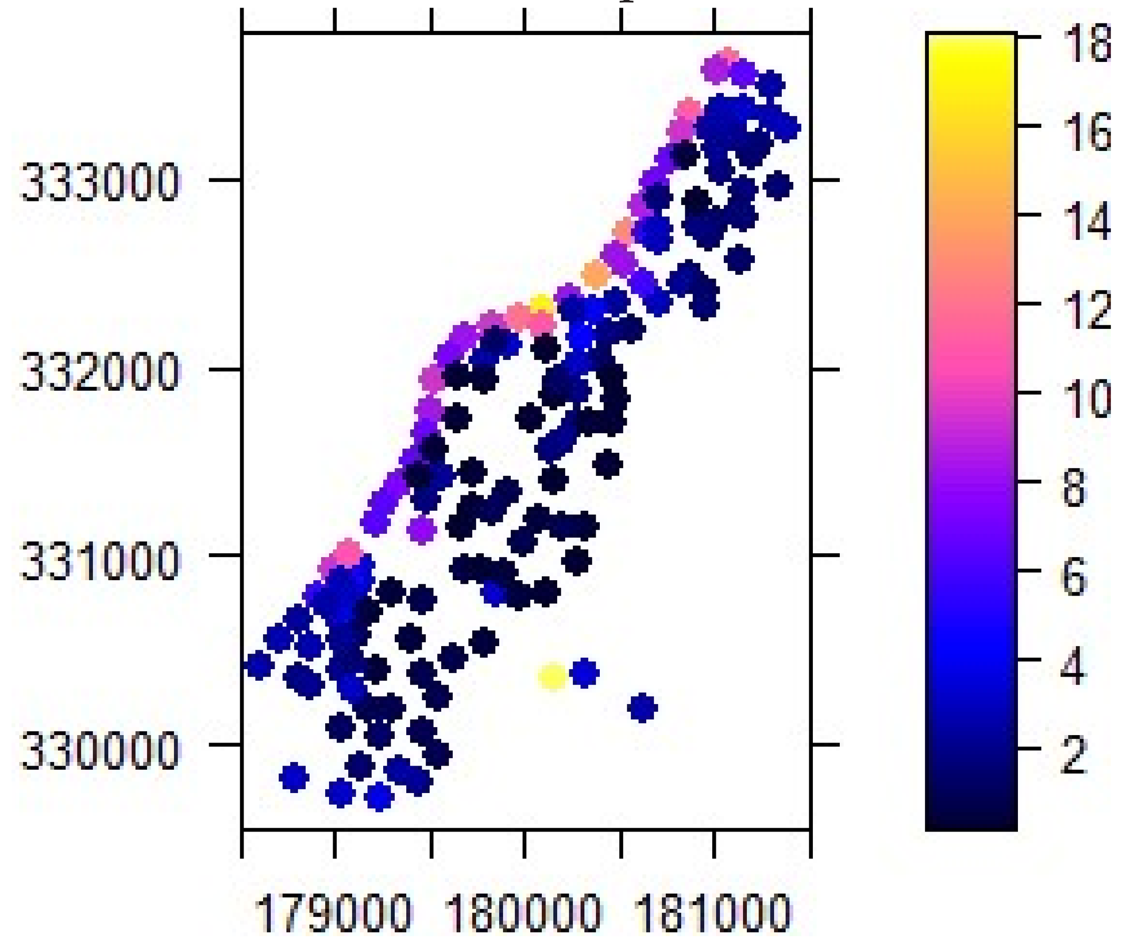
Pairs are considered to be nearly orientated along the direction θ_1 if,

$$\left| \frac{b'_j - b'_i}{a'_j - a'_i} \right| < \text{ a tolerance distance}$$

All point pairs that satisfy the above inequality are selected as the pairs for the major direction (θ_1), and the rest of the pairs are selected for the direction θ_2 (Addo et al., 2019) .

Application - Meuse river bank data

The Meuse river bank data consists of measurements of heavy metal concentrations (e.g., cadmium, zinc), collected at the observation locations in a flood plain of the river Meuse, Netherlands. The below figure shows the spatial plot of cadmium concentration samples in the Meuse data.



The cadmium concentration is used to demonstrate our new model, where the SRF of cadmium showed very strong anisotropy in the direction of 60° (θ_1). The best marginal distribution of cadmium is Weibull (shape=0.94, scale = 3.15). Based on the geometric anisotropy, the data pairs were separated into θ_1 and θ_2. Using Equation 2, spatial copulas were fitted for each direction. The mixture copulas C_k^m for five spatial bins are given in the table.

Bins(km)	$C_{\theta_1,k}$	$C_{\theta_2,k}$	C_k^m
0-150	Frank (14.18)	Gumbel (1.49)	C_1^m
150-300	Frank (5.63)	Frank (2.51)	C_2^m
300-450	Gumbel (1.48)	Clayton (0.41)	C_3^m
450-500	Frank (2.64)	Clayton (0.18)	C_4^m
500-750	Gumbel (1.27)	Joe (2)	C_5^m

Then, the spatial copula for the anisotropic SRF (cadmium) is constructed as the convex combination of mixture copulas, is given by C_h^m.

$$C_h^m = \begin{cases} C_1^m, & \text{if } 0 \le h < 150 \\ (1-\lambda_2)C_1^m + \lambda_2 C_2^m, & \text{if } 150 \le h < 300 \\ (1-\lambda_3)C_2^m + \lambda_3 C_3^m, & \text{if } 300 \le h < 450 \\ (1-\lambda_4)C_3^m + \lambda_4 C_4^m, & \text{if } 450 \le h < 600 \\ (1-\lambda_5)C_4^m + \lambda_5 C_5^m, & \text{if } 600 \le h < 750 \end{cases}$$

where, for example, $\lambda_2 = \frac{h-150}{300-150}$ and h is the mean distance of each spatial bin.

Our novel anisotropic spatial copula method was compared with an existing isotropic spatial copula method (spatial vine copula of Gräler and Pebesma, 2011). The performance in the reproduction of cadmium was assessed based on the root mean square error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE). The results showed that proposed method that considers the directional dependence in spatial copula modelling outperformed in terms of MAE, RMSE and MAPE.

Conclusion

A new spatial model is presented for anisotropic SRF using mixture copula. This model facilitates predicting the spatial response by utilizing the joint anisotropic spatial dependence of multiple directions. The new method applied to cadmium concentration data, and the ability of the prediction accuracy was compared with an existing spatial copula model, where our new method was outperformed.

Future Research

In this research, the equal weights (w=0.5) were used for mixture copula construction, however the selection of optimal weight combinations will be addressed in future research. Also, only two directions were considered, however the multiple directions could be used in mixture copula modelling. Moreover, a simulation study will be carried out in future research. Finally, this research can be extended to spatio-temporal copula modelling, where application such as wildfire data can be considered with directional dependence over time.

Contact: abrajsz20@gmail.com

References

[1] András Bárdossy. Copula-based geostatistical models for groundwater quality parameters. *Water Resources Research*, 42(11), 2006.

[2] Benedikt Gräler and Edzer Pebesma. The pair-copula construction for spatial data: a new approach to model spatial dependency. *Procedia Environmental Sciences*, 7:206–211, 2011.

[3] Emmanuel Addo, Emmanuel K Chanda, and Andrew V Metcalfe. Spatial pair-copula model of grade for an anisotropic gold deposit. *Mathematical Geosciences*, 51(5):553–578, 2019.