The Study of Spatio-temporal Correlation Structure of Groundwater Data in Ilam Using Bees and Genetic Algorithms

Mehdi Omidi[1*], Iradj Yadegari[2]
[1]Department of Mathematics, Ilam University, Ilam, Iran.
[2] Department of Mathematics and Statistics, University of Ottawa, ON, Canada.

Abstract:
Spatio-temporal covariance functions are main parts of geostatistical analysis of environmental data. Many classes of covariance functions along with a large number of parameters have been proposed to estimate correlation structure of the spatial-temporal data. This paper surveys using and comparing of two important heuristic algorithms Bees Algorithm (BA) and Genetic Algorithm (GA) for finding optimal values of spatio-temporal covariance parameters based on full likelihood, Composite and Weighted Composite likelihood function. Next, the accuracy and computational time of these methods are compared based on a simulation study. Finally, the combined method is used to explore the spatio-temporal correlation structure of the monthly level of groundwater data in Ilam province, Iran.
Keywords: Spatio-temporal covariance, Bees algorithm, Genetic algorithm.
Mathematics Subject Classification (2010): 62H11, 91B72.

---

1 Introduction

Dependency structure of the geostatistical spatio-temporal models are often based on covariance functions. Many researchers have attempted to introduce and develop general and flexible classes of parametric spatio-temporal covariance functions (Stein , 2005; Mateu *et al.* , 2008; Omidi and Mohammadzadeh , 2016). Although the Maximum Likelihood (ML) technique is the most popular method for parameter estimation of spatio-

*Speaker: m.omdi@ilam.ac.ir

temporal covariance functions, a major problem with this method is computing the inverse of spatio-temporal covariance matrix in the structure of likelihood function for the Gaussian random field. To decrease the computational time, two alternative estimation methods including Composite Likelihood ($C\ell$) and Weighted Composite Likelihood ($Wc\ell$) have been proposed (Bevilacqua *et al.* (2012)). However, all likelihood methods need to use robust optimization algorithms to find the maximum likelihood estimates.

Many optimization algorithms have inspired by studies of social animals and social insects (Pham and Castellani , 2015). In the BA, the new generations are created by combining random and neighborhood searching. The BA has been applied to many problems in the literature. For example, Omidi and Mohammadzadeh (2018) used it for non-Gaussian modeling of rainfall data, and Ayyad *et al.* (2019) put it into use for optimization problems related to spatial behavior of rat sightings.

Another popular heuristic algorithm is Genetic Algorithm (GA). In the GA, the stronger generation has more chance to be the winner based on a set of mechanisms such as selection, crossover and mutation. The GA was introduced by Holland (1975) and developed and explained by Goldberg (1989).

In this article, we have implemented the Bees and Genetic algorithms to optimize the full likelihood, $C\ell$, and $Wc\ell$ function to estimate the parameters of spatio-temporal covariance function. In this line, the efficiency of the algorithms was evaluated with Mean Square Errors (MSE) by generating 100 independent zero-mean Gaussian random fields. In addition, real data were used to assess the efficiency of algorithms to estimate the spatio-temporal dependency structure of data. The accuracy of the fitted models was assessed by the Mean of Squared Prediction Errors (MSPE). Our data were the level of groundwater in two cities, Dehloran and Abdanan, in the west of Iran.

2 Spatio-temporal Covariance Functions

Let $\{Z(s;t) : s \in \mathcal{S} \subseteq R^d, t \in\subseteq R^+\}$ denotes a spatio-temporal random field at the location s and time t with the assumption of existing finite variance $Var(Z(s;t))$. The mean and spatio-temporal covariance functions for this random field are defined by $E(Z(s;t)) = \mu(s;t)$ and $C(s, s', t, t') = Cov(Z(s;t), Z(s';t'))$, respectively. Moreover, $2\gamma(s, s'; t, t') = Var(Z(s;t) - Z(s';t'))$ is defined as spatio-temporal variogram for $Z(s;t)$. For intrinsic space-time stationary random fields, the variogram reduces to $Var(Z(s;t) - Z(s';t')) = 2, \gamma(h, u)$ and based on weakly stationary, the covariance function reduces to

$$Cov(Z(s;t), Z(s';t')) = C(h, u),$$

where $h = ||s - s'||$ and $u = |t - t'|$ are spatial and temporal lags, respectively. Under the assumption of weakly stationarity, the variogram will be obtained by $2\gamma(s, s'; t, t') = 2\,(C(0,0) - C(h,u))$.

Using a valid spatio-temporal covariance function is one of the important problems in the spatio-temporal data analysis. Many classes of these functions have been constructed which the simplest one is called a separable and introduced as $C(h,u) = C_1(||h||)\,C_2(|t|)$, where $C_1(\cdot)$ and $C_2(\cdot)$ are positive definite functions and they called purely spatial and purely temporal covariance function, respectively. For example, by choosing $C_1(u) = \sigma^2/(a|u|^{2\gamma_1} + 1)$ and $C_2(h) = \exp\{-b||h||^{2\gamma_2}\}$, the model

$$\text{Model } 1 : C(h,u) = \frac{\sigma^2}{(a|u|^{2\gamma_1} + 1)} \exp\{-c||h||^{2\gamma}\}, \tag{2.1}$$

is a separable spatio-temporal covariance function in $R^d \times R$, where $\gamma_1, \gamma_2 \in [0,1]$, and σ, a, and b are positive values.

Although, ease of application and saving computational time are two major advantages of the separable covariance class, in some applications this assumption is inappropriate. In this setting nonseparable covariance functions must be used. Gneiting (2002) proposed a valid and flexible class of nonseparable spatio-temporal covariance models using completely monotone and Bernstein functions. This class is given by $C(h,u) = \frac{\sigma^2}{\psi(u^2)^{d/2}} \exp\{-\frac{\phi(||h||)^2}{\psi(u^2)}\}$, where $\phi(\cdot)$ and $\psi(\cdot)$ are completely monotone and Bernstein functions, respectively. For example, by choosing $\phi(t) = \exp\{-b\,t^{\gamma}\}$, $\psi(t) = (1 + at^{\gamma_1})^{\gamma_2}$, the model

$$\text{Model } 2 : C(h,u) = \frac{\sigma^2}{(a|u|^{\gamma_1} + 1)^{\gamma_2 d/2}} \exp\{-\frac{b||h||^{2\gamma}}{(a|u|^{2\gamma_1} + 1)^{\gamma_2\gamma}}\}, \tag{2.2}$$

is a non-separable spatio-temporal covariance function in $R^d \times R$, where γ_1, γ_2 and γ are in $[0,1]$ and σ, a, and b are positive values.

Suppose that σ, a, b and θ are positive values and γ_1, γ_2 and $\gamma \in (0,1]$. Then, the models

$$\text{Model } 3 : C(h,u) = \sigma^2 \exp\{-(a||h||^{2\gamma_1} + b|u|^{2\gamma_2})^{\gamma}\} \tag{2.3}$$

$$\text{Model } 4 : C(h,u) = \sigma^2\{(1 + a||h||)^{\gamma_1} + (1 + b|u|)^{\gamma_2}) - 1\}^{-\theta} \tag{2.4}$$

and

$$
\begin{aligned}
\text{Model } 5: C(h,u) &= \sigma^2 \exp\{-((1+a||h||)^{\gamma_1} + (1+b|u|)^{\gamma_2}) - 2)\}^{\gamma} && (2.5)\\
\text{Model } 6: C(h,u) &= \sigma^2 \exp\{-(a||h||^{2\gamma_1} + (1+b|u|)^{\gamma_2} - 1)^{\gamma}\} && (2.6)\\
\text{Model } 7: C(h,u) &= \sigma^2 \{(1+a||h||)^{\gamma_1} + b|u|^{2\gamma_2}\}^{-\theta} && (2.7)\\
\text{Model } 8: C(h,u) &= \sigma^2 \{1+a||h||^{2\gamma_1} + b|u|^{2\gamma_2}\}^{-\theta} && (2.8)
\end{aligned}
$$

are stationary non-separable spatio-temporal covariance functions in $R^d \times R$.

2.1 Maximum Likelihood Methods

The ML method is the most common technique for estimating the parameters of the spatio-temporal covariance. Some asymptotic properties like consistency, efficiency, and unbiasedness of ML estimators makes it a popular technique in applications. Let $Z(s;t)$, with $s \in \mathcal{S} \subseteq R^d$ and $t \in \{t_1, \ldots, t_T\} \subseteq R^+$, be a real-valued stationary space-time from a zero-mean Gaussian random field with realizations $\mathbf{z} = (z(s_1, t_1), \ldots, z(s_M, t_T))$ at M spatial sites and T time points. The log-likelihood function for these observations can be reduced as

$$\ell(\theta) = -\log|\Sigma(\theta)| - \mathbf{z}'\Sigma^{-1}(\theta)\mathbf{z} \tag{2.9}$$

where $\Sigma(\theta)$ denotes the parametric spatio-temporal covariance function, and θ is a set of all unknown parameters. Without loss of generality, the ML estimates were obtained by minimizing $-\ell(\theta)$ in the parameter space.

Finding the ML estimates requires to calculate the inverse of the spatio-temporal covariance matrix. To avoid these drawbacks, $C\ell$ can be helpful to estimate the parameters instead of the full likelihood function (Varin *et al.* , 2019). In $C\ell$, the parameters of covariance function are estimated using the distribution of all paired observation differences and also its corresponding spatio-temporal variogram function. Bevilacqua *et al.* (2012) showed that the -logarithm of composite likelihood function (Cℓ) can be written as

$$C\ell(\theta) = \sum_{i=1}^{M\times T}\sum_{j>i}^{M\times T} \log\gamma(s_i - s_j, t_i - t_j;\theta) + \frac{u^2(s_i, s_j, t_i, t_j)}{2\gamma(s_i - s_j, t_i - t_j;\theta)}. \tag{2.10}$$

For reducing the computational time, when the of M and T have large values, the Eq. (2.10) can be approximated by weighted composite likelihood ($Wc\ell$) function given by

$$Wc\ell(\theta) = \sum_{i=1}^{M\times T}\sum_{j>i}^{M\times T} w_{ij}(\log\gamma(s_i - s_j, t_i - t_j;\theta) + \frac{u^2(s_i, s_j, t_i, t_j)}{2\gamma(s_i - s_j, t_i - t_j;\theta)}), \tag{2.11}$$

where w_{ij} is an indicator function, and, $w_{ij} = 1$ whenever $|s_i - s_j| \leq h_w$ and $|t_i - t_j| \leq u_w$, and, $w_{ij} = 0$ otherwise.

3 Simulation Study

In order to apply the BA for minimizing the full likelihood, $C\ell$, and $Wc\ell$ presented in Eq.s (2.9), (2.10) and (2.11) we use abbreviations $B\ell$, $Bc\ell$, and $Bwc\ell$, respectively. In the BA, the parameters were assumed as $n = 100$, $m = 40$, $e = 10$, $n_{ep} = 6$, $n_{es} = 3$. In the GA, the mutation and crossover rate were chosen 0.005 and 0.85, respectively. Moreover, the new values of the GA were produced based on 2% of the best values from previous values and 8% were selected randomly and the other 90% produced by the mechanism of crossover and mutation. Dimension of 7500, for $N^2 = 10^2$ and $T = 75$, was selected for simulation data. Each one of the algorithms along with likelihood approaches were applied for 100 independent simulations from zero-mean Gaussian random fields. Studying all combinations of eight models, two algorithms, and three estimation methods needs to generate 48 independent simulations. However, we selected only the first four models presented in Eqs. (2.1), (2.2), (2.3), and (2.4). The first model is separable and the others are non-separable models. Other models in (2.5) to (2.8) have similar analyses. Data were simulated based on the following values of the parameters:

- Model 1: $\sigma^2 = 1.7$, $a = 0.8$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$, $b = 2$.
- Model 2: $\sigma^2 = 1.7$, $a = 0.8$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$, $b = 2$, $\gamma = 0.5$.
- Model 3: $\sigma^2 = 1.7$, $a = 2$, $\gamma_1 = 0.5$, $b = 0.8$, $\gamma_2 = 0.5$, $\gamma = 0.5$.
- Model 4: $\sigma^2 = 1.7$, $a = 2$, $\gamma_1 = 0.5$, $b = 0.8$, $\gamma_2 = 0.5$, $\theta = 1.5$.

Table 1 contains MSE and mean computational time in the estimation of spatio-temporal covariance parameters producing 50 new generations (iterations) for 100 independent simulations based on the BA and GA. As it is shown in Table 1, MSE of the full likelihood is less than MSE of $C\ell$ and $Wc\ell$ in combining with Genetic and Bees algorithms. If we define total MSE (TMSE) as sum of MSEs for all estimated parameters of a model, we can sort the algorithms. Therefore, for all four covariance functions we can say that

$$TMSE_{Wc\ell} > TMSE_{C\ell} > TMSE_{\ell}, \tag{3.1}$$

It must be noted that sorting GA and BA according to MSE of single estimators or TMSE is not possible.

Table 1: MSE of the estimated covariance parameters using BA and GA algorithms in $Wc\ell$, $C\ell$, and full likelihood estimation methods for simulated data.

Model	Method	Model Parameters							TMSE	Time
		σ^2	a	γ_1	b	γ_2	γ	θ		
Model 1	Bwcℓ	0.8495	2.6692	0.2315	1.7016	0.2542	-	-	5.706	4.3×10^3
	Bcℓ	0.8484	2.7372	0.2452	1.2725	0.2555	-	-	5.3588	11.4×10^3
	Bℓ	0.3345	0.2683	0.1356	0.8526	0.2216	-	-	1.8126	83.8×10^3
	Gwcℓ	1.2001	2.3198	0.2400	2.8124	0.2496	-	-	6.8219	1.6×10^3
	Gcℓ	0.7910	2.2643	0.2621	1.0139	0.2746	-	-	4.6059	7.3×10^3
	Gℓ	0.4585	0.3795	0.1118	0.7492	0.2138	-	-	1.9128	32.7×10^3
Model 2	Bwcℓ	0.8442	2.2662	0.2195	1.4238	0.1972	0.2265	-	5.1772	6.1×10^3
	Bcℓ	0.8408	1.7296	0.2098	1.2311	0.2172	0.2076	-	4.4361	26.5×10^3
	Bℓ	0.2447	1.1375	0.1722	1.1423	0.2032	0.1676	-	3.0675	104.0×10^3
	Gwcℓ	0.9507	2.4901	0.2606	2.0948	0.2412	0.2433	-	6.2807	2.5×10^3
	Gcℓ	0.8104	2.0423	0.2204	1.7614	0.2170	0.2408	-	5.2923	11.2×10^3
	Gℓ	0.2207	1.0782	0.1883	1.1356	0.1945	0.1846	-	3.0019	41.1×10^3
Model 3	Bwcℓ	0.8380	2.3366	0.2274	2.1492	0.2472	0.2040	-	6.0024	5.6×10^3
	Bcℓ	0.8116	1.8371	0.2387	1.5678	0.2252	0.2285	-	4.9089	28.1×10^3
	Bℓ	0.1916	1.1912	0.1924	0.3630	0.1626	0.1114	-	2.2122	95.7×10^3
	Gwcℓ	1.4951	2.1035	0.2565	2.0066	0.2575	0.2345	-	6.3537	2.2×10^3
	Gcℓ	1.2243	1.5227	0.2423	1.7755	0.2102	0.2236	-	5.1986	10.4×10^3
	Gℓ	0.2429	1.1015	0.2313	0.3860	0.1709	0.1356	-	2.2682	36.6×10^3
Model 4	Bwcℓ	0.8040	2.8392	0.2287	2.8107	0.2317	-	1.4889	8.4032	6.3×10^3
	Bcℓ	0.8286	2.4996	0.2258	2.9127	0.2567	-	1.1791	7.9025	26.1×10^3
	Bℓ	0.4839	1.9700	0.1967	0.5387	0.1814	-	1.3896	4.7603	78.3×10^3
	Gwcℓ	1.2416	2.5301	0.2653	1.6632	0.2522	-	1.3297	7.2821	2.7×10^3
	Gcℓ	0.9575	2.4733	0.2445	1.4637	0.2452	-	1.1535	6.5377	12.2×10^3
	Gℓ	0.5060	1.1311	0.2322	0.5015	0.1542	-	0.9399	3.4649	38.7×10^3

4 Real Data Analysis

This section contains application of the BA and GA algorithms to estimate of the spatio-temporal covariance parameters for a monthly average of groundwater in two cities, Dehloran and Abdanan, in Ilam province at the west of Iran. Data were monthly records of the groundwater that include 7200 records over 15 years from 2003 to 2017, and, observed at 40 different stations. There were few missing values that were replaced by linear interpolation. Furthermore, $156 \times 40 = 6240$ records were used for fitting the model and $24 \times 40 = 960$ records for assessing the accuracy of the fitted model.

Exploratory data analysis showed departure from normality in the data. However, the Kolmogorov-Smirnov test for transformed data by Box-Cox transformation with $\lambda = 0.4241$ and p-value=0.072, was not significant. The data was detrended by $\mu(s,t) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 t$, where the p-values for longitude, latitude, and time were attained 0.000, 0.001 and 0.001 , respectively. So, the residuals along with zero-mean Gaussian random filed were regarded to fit the models from 1 to 8 to the data.

In order to estimate the spatio-temporal covariance parameters, both algorithms were applied to minimize functions in Eq.s (2.9), (2.10), and (2.11). To assess the accuracy, the models were fitted to the data from 2003 to 2015. Then, using spatio-temporal Kriging the monthly level of the groundwater in all stations during 24 months from 2016 to

Table 2: MSPE based on the estimated models using spatio-temporal kriging

Moldel	Bwcℓ	Bcℓ	Bℓ	Gwcℓ	Gcℓ	Gℓ
Model 1	0.2011	0.1544	0.1419	0.2825	0.2461	0.1513
Model 2	0.2016	0.1773	0.1417	0.1560	0.2176	0.0982
Model 3	0.1572	0.2143	0.1173	0.1418	0.1441	0.1353
Model 4	0.1861	0.1685	0.1293	0.1058	0.1134	0.0586
Model 5	0.1552	0.2109	0.1461	0.1369	0.1168	0.0772
Model 6	0.1946	0.1292	0.1304	0.1488	0.1370	0.0705
Model 7	0.0963	0.1512	0.0910	0.1360	0.1323	0.0580
Model 8	0.2218	0.1740	0.1103	0.1063	0.1349	0.0597

2017 (960 records) were predicted. Next, using Mean Square Prediction Error (MSPE) the accuracy of the methods and algorithms were compared. Considering MSPE, the estimation based on likelihood function in both algorithms had minimum value among the functions. Also, comparing MSPE between BA and GA in corresponding functions suggested that the better algorithm depends on the covariance model. Finally, model 7, which has the smallest values of MSPE based on $G\ell$, is relatively better among the others for spatio-temporal prediction of groundwater data in the south of Ilam province. This fact shows the estimation based on likelihood function using GA has better fit to the data.

Conclusion

In this paper, to estimate spatio-temporal covariance parameters, we combined heuristic numerical methods BA and GA with three likelihood-based estimation methods, including full, composite and weighted composite likelihood methods. Simulation shows that, there is a trade-off between computational time and MSE of the estimated parameters. The heuristic algorithms were applied to minimize -log-likelihood functions and the results were compared by MSE, and, the computational time. For a given algorithm and a covariance model, the computational time of the full likelihood is considerably larger than computational time of $C\ell$ and $Wc\ell$. Also, given an estimation method, the GA has less computational time than BA.

Moreover, $Wc\ell$ is faster than other methods, which is reasonable because its corresponding likelihood function is shorter than others. In another hand, as presented in Eq. (3.1), total MSE of estimation methods can be sorted such that the full likelihood and weighted composite likelihood have the smallest and largest total MSE, respectively. Finaly, We used groundwater data to evaluate and assess combining these two heuristic algorithms with likelihood-based estimation methods and the level of groundwater in all stations were predicted using spatio-temporal kriging and the efficiency of the model was assessed by MSPE.

References

Ayyad, C., Mateu, J., Mehdi Omidi, Tamayo-Uria, I., and Mohammadzadeh, M., (2019), Trivariate non-Gaussian copulas to analyse the spatial behaviour of rat sightings. *Statistica Neerlandica*, **73** (2), 256-273.

Bevilacqua, M., Gaetan, C., Mateu, J., and Porcu, E. (2012), Estimating Space and Spacetime Covariance Functions for Large Data Sets: a Weighted Composite Likelihood Approach, *Journal of the American Statistical Association*, **107**, 268-280.

Gneiting, T., (2002), Nonseparable, Stationary Covariance Functions for Space-Time Data, *Journal of the American Statistical Association*, **97**, 590-600.

Goldberg, D. E., (1989), Genetic algorithms in search, optimization, and machine learning, addison wesley, reading, ma, 1989. NN Schraudolph and J, **3**(1).

Holland, J. (1975), *Adaptation in natural and artificial systems: an introductory analysis with application to biology*, Control and artificial intelligence.

Mateu, J., Porcu, E. and Gregori, P. (2008), Recent Advances to Model a Anisotropic Space-time Data, *Statistical Methods and Applications*, **17**, 209-223.

Omidi, M. and Mohammadzadeh, M. (2016), A New Method to Build spatio-temporal Covariance Functions: Analysis of Ozone Data, *Statistical Papers*. **57**, 689-703.

Omidi, M. and Mohammadzadeh, M. (2018). Spatial Interpolation Using Copula for non-gaussian Modeling of Rainfall Data, *Journal of The Iranian Statistical Society*, **17**(2):165179, 2018.

Pham, D. T. and Castellani, M. (2015), A comparative study of the Bees Algorithm as a tool for function optimisation, *Cogent Engineering*, **2**(1):1091540.

Stein, M. (2005), Space-time covariance functions, *Journal of the American Statistical Association*, **100**, 310-321.

Varin, C., Reid, N., and Firth, D. (2011), An overview of composite likelihood methods,*Statistica Sinica*, **21**, 5-42.