ADAPTIVE PARAMETRIC MODEL FOR NONSTATIONARY SPATIAL COVARIANCE

Julian Ibezimako Mbegbu[1a] and Joseph Thomas Eghwerido[2b*]

[a]Department of Statistics, University of Benin, Benin City, Edo State, NIGERIA. E-mail: julian.mbegbu@uniben.edu[1]
[b]Department of Mathematics, Federal University of Petroleum Resources, P. M. B. 1221 Effurun, Delta State, NIGERIA. E-mail: eghwerido.joseph@fupre.edu.ng[2]
Corresponding author: eghwerido.joseph@fupre.edu.ng
Received: 12th Jun 2019 Accepted: 5th Mar 2020 Published: 30th Jun 2020
DOI: https://doi.org/10.22452/mjs.vol39no2.4

ABSTRACT In modelling environment processes, multi-disciplinary methods are used to explain, explore and predict how the earth responds to natural human-induced environmental changes over time. Consequently, when analyzing spatial processes in environmental and ecological studies, the spatial parameters of interest are always heterogeneous. This often negates the stationarity assumption. In this article, we proposed the adaptive parametric nonstationary covariance structure for spatial processes. The adaptive tuning parameter for this model was also proposed for nonstationary processes. The flexibility and efficiency of the proposed model were examined through simulation. A real life data was used to examine the efficiency of the proposed model. The results show that the proposed models perform competitively with existing models.

Keywords: Adaptive, Locally, Nonstationary, Spatial Covariance, Variability.

1. INTRODUCTION

In modelling spatial processes in environmental, agricultural and meteorological processes, the parameters of interest are not always known. However, most methods for obtaining such measurements assumed stationarity (Cressie 1993) which often negates the measurement obtained in environmental and ecological processes. For examples, the dispersion of the transportation of atmospheric pollutants, topographic effect, and weather patterns are good cases of nonstationary spatial processes where stationarity is assumed. Also, most environmental processes exhibit spatially nonstationary covariance structure over sufficiently large spatial scales (Fuentes, 2001). However, several studies like Fuentes (2002), Haas (1990a) and Haas (1990b), Guttorp and Sampson (1994)and many others have shown that meteorological processes such as; synoptic wind patterns and orographic effects often exhibit nonstationary covariance. Consequently, however, the understanding of the observed values is needed to predict values at unobserved location. Thus, the nonstationary processes observed in these spatial processes rely on the covariance and variogram techniques.

Let $x_1, x_2, \cdots x_n$ be spatial locations and $Y(x_i)$ the space domain, where $x \in \Re^d$ .Then, the realization say, $Y(x_i)$is nonstationary if either

$Cov[Y(x_i),Y(x_j)]$ $i<j,\ \ i,j=1,2,\ldots n$ for x_i and $x_j \in \Re^d$ depends on the locations x_i and x_j or $E[Y(x_i)]$varies over the random field. However, but the covariance function in these locations changes with locations.

In spatial statistics, spatial processes were considered in Sampson and uttorp (1992) to be nonstationary. They represented the covariance of the processes as a latent space. Higdon et al. (1999) proposed a nonstationary covariance function using a convolution of two kernel functions. This was further expanded by Paciorek and Schervish (2006) by using the square root of the quadratic form of the spatial covariance. A moving window approach was proposed in David and Genton (2000). Hyoung-Moon et al. (2005) and Parker et al. (2016) regrouped the heterogeneous spatial processes in the same location into sub-regions whose structures were homogeneous in both the mean and covariance. Bornn et al. (2012) proposed a nonstationary covariance in space using the concept of dimension expansion method. This method was further enhanced in Shand and Li (2017) using a thin-plate spline method to obtain the nonstationary covariance in space and time. Jaehong et al. (2017) proposed isotropic and nonstationary covariance using a differential operator approach. Fuentes (2002) proposed a nonstationary spatial process using the spectral density convolution approach. Higdon (1998) proposed a convolution based approach nonstationarity covariance. Ingebrigtsen et al. (2014) proposed covariance structure in nonstationary processes through stochastic partial differential method. This method allows the explanatory variables to be added to the structure. A linear mixed model approach was proposed in Haskard et al. (2010) to determine the mean and covariance of the nonstationary process of soil potassium on gamma radiometry. Lasso regression approach was proposed in Hsu et al. (2012) to select the basis function in modelling the nonstationary covariance. Huser and Genton (2016) proposed a nonstationary max stable dependence model and obtained the covariance using the pairwise likelihood method. Schmidt et al. (2011) proposed nonstationary covariance by using latent space model by projecting the C dimension in Sampson and Guttorp (1992) to 2D correlation structure using the covariate in the covariance. Some examples of latent space models are found in Meiring et al. (1998), Le et al. (2001), Sampson et al. (2001), Damian et al. (2003) and Guttorp et al. (2007). More so, nonstationarity as a sum of the stationary process and basis function with its coefficients as a departure from nonstationary were proposed in (Nychka and Saltzman 1998, Nychka et al. 2002). Several Bayesian methods for solving nonstationarity problem have been developed over the years in Katzfuss (2013), Katzfuss and Cressie (2012), but Risser and Calder (2015) proposed MCMC model for posterior distribution.

Motivated by the articles researched and based on the results obtained from existing spatial literature research such as the nonparametric estimation of spatial and space-time covariance function, non-parametric method of estimating semi-variograms of isotropic spatial processes, and the estimation of nonstationary spatial covariance structure, we proposed adaptive parametric nonstationary covariance for spatial processes whose variability depends on the lags between the spatial processes. Its major characteristic was that more parameters were added to make it more flexible. This model aims to attract wider range of application in agriculture, environmental sciences, hydrology and other related areas.

This article is organized as follows: Section 2 discusses the review of spatial covariogram and semi-variogram. In Section 3, we discuss the adaptive parametric model and the adaptive parameter for optimizing the adaptive model. Simulation of the formulated models was examined in section 4 together with a real life application. Section 5 contains the conclusions.

2. REVIEW OF SPATIAL DOMAIN

A fundamental notion underlying most of the current modeling approaches were that, the spatial covariance of the environmental processes can be regarded as approximately stationary over small spatial regions. This notion describes spatially varying isotropic covariance structure. Thus, for a spatial realization $Y(x_i)$, the semi-variogram is defined as

$$\gamma(x_i, x_j) = \frac{1}{2} E\left[Y(x_i) - Y(x_j)\right]^2, i < j, \quad i = 1, 2, 3, \cdots, n. j = 1, 2, 3, \cdots, n \tag{1}$$

However, in Matheron (1963), Equation (1) is said to be second order stationary process if the spatial covariance function is given as

$$Cov(h) = Cov(Y(x), Y(x+h)), \tag{2}$$

where h is the lag and x the location. However, the covariance between any two locations says x_1 and x_2 in Equation (1) depends on the spatial lag vector connecting them. Shand and Li (2017) consider the exponential covariance function for the space domain as $Cov(Y(x_i), Y(x_j)) = \sigma^2 \exp(-\phi_s h)$, for

$$i < j, \; i = 1, 2, 3, \cdots, n. j = 1, 2, 3, \cdots, n \tag{3}$$

where ϕ_s is the range parameter. Then $h = \lfloor x_i - x_j \rfloor$ and $h = \|h\|$ where $\|b\|$ is a Euclidean norm of vector $\|b\|$ and σ^2 is the variance of the process.

3. METHODOLOGY

In this section, a new adaptive parametric nonstationary spatial covariance is proposed.

3.1 *Proposed adaptive parametric nonstationary spatial covariance*

Let $x_1, x_2, \cdots, x_n$ be the spatial locations in domain $\Re^d$ and $Y(x_1), Y(x_2), \cdots, Y(x_n)$ the spatial processes in the spatial locations. Then, the nonstationary spatial process is expressed as

$$Y_i(x) = m^T(x_i) B_i^{-1} Y_i + \delta(x_i) \quad for \quad i = 1, 2, 3, \cdots, n \tag{4}$$

where $m^T(x_i)$ is the cross covariance between observed and unobserved locations, B_i is the covariance vector between the processes at locations $x_i^{'s}$, Y_i is the column vector of the random process, and $\delta(x_i)$ is the error term.

The spatial residual from Equation (4) can be expressed as

$$\begin{aligned} E[\delta^2] &= E[Y^T(x_i)Y(x_i)] \\ &- 2E\left[\left(B_i^{-1}\right)^T Y_i m(x_i) Y(x_i)\right] \\ &+ E\left[\left(B_i^{-1}\right)^T Y_i m(x_i) m^T(x_i) Y_i^T B_i^{-1}\right] \end{aligned} \tag{5}$$

Spatial processes vary from one location to another depending on the distance between the processes. However, to obtain optimal value of spatial process at unobserved location, the distance is penalized. Thus, an optimization problem is set up by minimizing the objective of Equation (4) as

$$Y_i(x) - m^T(x_i) B_i^{-1} Y_i \geq 0, \qquad for \quad i = 1, 2, 3, \cdots, n. \tag{6}$$

Subject to the constraint

$$B_i^{-1} \geq 0 \qquad i = 1, 2, 3, \cdots, n. \tag{7}$$

However, due to over fitting and high dimensional of the data analysis, using Karush-Kuhn-Tucker technique, spatial minimization problem of Equations (6) and (7) is given as

$$\begin{aligned} L\left(B_i^{-1}, \lambda\right) &= \left\| Y_i(x) - m^T(x_i) B_i^{-1} Y_i \right\|_2^2 + \lambda_i \left\| B_i^{-1} \right\|_2^2 \\ \lambda_i &\in [0,1] \end{aligned} \tag{8}$$

where, λ_i are $n \times 1$ vectors of adaptive tuning parameters that are data dependent, such that $\lambda_i \geq 0$. The values of the adaptive parameters shrink the spatial regression equation coefficient towards zero and add some spatial bias that reduces the nonstationary covariance of the estimator. While ℓ_2-norm is used to keep the spatial equation rotationally invariant.

Thus, for computational purpose, Equation (8) can be expressed as

$$\begin{aligned} L\left(B_i^{-1}, \lambda\right) &= Y^T(x_i) Y(x_i) \\ &- 2\left[\left(B_i^{-1}\right)^T Y_i m(x_i) Y(x_i)\right] \\ &+ \left(B_i^{-1}\right)^T Y_i m(x_i) m^T(x_i) Y_i^T B_i^{-1} \\ &+ \lambda_i \left(\left(B_i^{-1}\right)^T B_i^{-1} - 2\right) \end{aligned} \tag{9}$$

Taking partial derivative of Equation (9) with respect to B_i^{-1} and equating to zero, we have

$$\begin{aligned}&[Y_i m(x_i) Y(x_i)]\\&= Y_i m(x_i) m^T(x_i) Y_i^T B_i^{-1}\\&+\lambda_i B_i^{-1}\end{aligned} \tag{10}$$

Substituting Equation (10) into Equation (5) we have the adaptive parametric model as

$$\begin{aligned}&E[\delta^2] = E[Y^T(x_i) Y(x_i)]\\&-2E\left[\left(B_i^{-1}\right)^T\left(Y_i m(x_i) m^T(x_i) Y_i^T B_i^{-1} + \lambda_i B_i^{-1}\right)\right]\\&+E\left[\left(B_i^{-1}\right)^T Y_i m(x_i) m^T(x_i) Y_i^T B_i^{-1}\right]\end{aligned} \tag{11}$$

where $E[Y_i, Y_i^T] = \left(B_i^{-1}\right)^T, E[Y(x_i), Y(x_i)] = \Im_i(x, x)$ and by symmetric property of a matrix, we have the proposed adaptive parametric model for nonstationary spatial covariance (AP 1) for location x as

$$\begin{aligned}&\delta^2_{iAP1_i}(x, x) = \Im_i(x, x) - m^T(x_i) B_i^{-1} m(x_i)\\&-2\lambda_i \left(B_i^{-1}\right)^T B_i^{-1} \qquad i = 1, 2, 3, \cdots, n\end{aligned} \tag{12}$$

where $\Im_i$ is the variance of $Y_i(x)$ or possibly the sill.

3.2 Proposed adaptive parameter for generating optimal model

In this section, we shall propose the adaptive parameter for generating the optimal model. Let

$$\rho = \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij} \text{ and } \lambda_i \neq 0 \text{ such that } \phi = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}}{\rho}$$

Let the adaptive parameter at location zero be

$$\lambda_0 = 1 - \phi = 1 - \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}}{\rho} = \frac{\rho - \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}}{\rho} = 0 \tag{13}$$

Now, at some other locations other than zero, we multiply Equation (13) by $\frac{1}{n^2-1}$. Thus, we have

$$\lambda_i = \frac{n\rho - B_{i\bullet}}{\rho}\frac{1}{(n^2-1)} \tag{14}$$

However,

$$\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}\frac{n\rho - B_{i\bullet}}{\rho}\frac{1}{(n^2-1)} = 1 \tag{15}$$

Thus,

$$\lambda_i = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} nB_{ij} - B_{i\bullet}}{(n^2-1)\sum_{i=1}^{n}\sum_{=1}^{n} B_{ij}} \quad i = 1,2,3,\ldots,n \tag{16}$$

Next, we shall consider some possible fixed values of λ_i.

For $\lambda = 0$ Equation (12) reduces to David and William (2002) model.

Now, taking second partial derivative of Equation (9) with respect to B^{-1}_i and equating to zero, we have:

$$\lambda_i = -Y_i Y_i^T m^T(x_i) m(x_i); \tag{17}$$

but λ_i cannot be negative. Thus, taking the norm of both sides we have

$$\|\lambda_i\| = \|Y_i Y_i^T m^T(x_i) m(x_i)\|, \tag{18}$$

Hence, substituting the value of λ_i into Equation (12) we have

$$\begin{aligned} \delta_{i_i}^2(x,x) &= \Im_i(x,x) - m^T(x_i) B_i^{-1} m(x_i) \\ &- 2E\left[Y_i Y_i^T m^T(x_i) m(x_i) (B^{-1}_i)^T B^{-1}_i\right] \\ & i, j = 1,2,3,\cdots,n \end{aligned} \tag{19}$$

On simplifying Equation (19), where $E\left[Y_i, Y_i^T\right] = \left(B_i^{-1}\right)^T$, $E\left[Y(x_i), Y(x_i)\right] = \Im_i(x,x)$ and by symmetric property of matrix, we have the parametric spatial covariance model (PM 1) for the location x as

$$\delta^2_{iPM1_i}(x,x) = \Im_i(x,x) - 3m^T(x_i)B_i^{-1}m(x_i) \qquad i = 1,2,3,\cdots,n \tag{20}$$

Otherwise, for $\lambda_i = 1$ Equation (12), the parametric spatial covariance model (PM 2) for the location x as

$$\delta^2_{iAP1_i}(x,x) = \Im_i(x,x) - m^T(x_i)B_i^{-1}m(x_i) - 2\lambda_i\left(B_i^{-1}\right)^T B_i^{-1} \qquad i = 1,2,3,\cdots,n \tag{21}$$

In the above cases, spatial processes in the neighborhood were used to predict the variable at unsampled location. Now, suppose the processes available are not within the neighborhood, we shall formulate a model on how such variables could be krig.

3.3 *Proposed parametric continuous kriging for nonstationary spatial covariance model*

If the locations x and s are far away from the unobserved site, then, the covariance $\Im_i(x,s)$ approaches zero as the lag h tends to infinity. Hence, a quantity $\tau_i = \frac{y_i}{\psi}$; $\left(\psi = \sum_{i=0}^{n} y_i\right)$ (where y_i are spatial processes) is introduced to penalize the parameter B_I^{-1} such that the observations within the neighborhood of the target point are used to obtain the predicted variables.

Proposition 3.1

Let $f(x - x_i) = \exp\left(-\|x - x_i\|^2\right)$ *be the distribution function of a spatial process and* $\upsilon_i(x) = \frac{f(x-x_i)}{\sum_{j=1}^{n} f(x - x_j)}$ *be the weight function, then* $\upsilon_i(x)$ *approaches zero as the*

$$\|x - x_i\| \to \infty \quad for \quad i = 1,2,3,\cdots,n$$

Proof

Since the $\upsilon_i(x)$ *are weight functions, we show that their sums equal one.*

$$f(x-x_1) = \exp\left(-\|x-x_1\|^2\right), f(x-x_2) = \exp\left(-\|x-x_2\|^2\right), \cdots, f(x-x_n) = \exp\left(-\|x-x_n\|^2\right) = A,$$

Let,

$$\upsilon_1(x)=\frac{\exp\left(-\|x-x_1\|^2\right)}{A},\upsilon_2(x)=\frac{\exp\left(-\|x-x_2\|^2\right)}{A},$$
$$\cdots,\upsilon_n(x)=\frac{\exp\left(-\|x-x_n\|^2\right)}{A}$$

Hence,

$$\sum_{i=1}^{n}\upsilon_i(x)=\frac{\exp\left(-\|x-x_1\|^2\right)}{A}+\frac{\exp\left(-\|x-x_2\|^2\right)}{A}+ \cdots+\frac{\exp\left(-\|x-x_n\|^2\right)}{A} \tag{22}$$

Clearly,

$$\sum_{i=1}^{n}\upsilon_i(x)=1. \tag{23}$$

It is clear from the proposition (3.1) that numerous sampled values are not needed to predict a variable at unsampled location. Thus sample kriging variance, K_{var} for the spatial process can easily be obtained in Gilmour et al. (2004) by:

$$K_{\text{var}}=\sum_{i=1}^{n}\upsilon_i(x)\delta^2_i(x_i,x_i) \tag{24}$$

Next, we show that $\upsilon_i(x)\to 0$ as $\|x-x_i\to\infty\|$

Observe that $\upsilon_i(x)=\dfrac{\exp\left(-\|x-x_i\|^2\right)}{\sum_{j=1}^{n}\exp\left(-\|x-x_i\|^2\right)}$ But $\exp\left(-\|x-x_i\|^2\right)=0\ as\|x-x_i\|\to\infty$. Clearly, $\upsilon_i(x)\to 0.$

The lasso regression equation of Equations (6) and (7) can be expressed as

$$\begin{aligned}&L\left(B_i^{-1},\lambda,\tau\right)=Y^T(x_i)Y(x_i)\\&+Y^T_i\left(B^{-1}_i\right)^T m(x_i)m^T(x_i)Y^T_iB^{-1}_iY_i\\&-2Y^T_i(x_i)m^T(x_i)B^{-1}_iY_i-2\lambda_iB^{-1}_i-\tau_iB^{-1}_i\left(B^{-1}_i\right)^T\end{aligned} \tag{25}$$

For some $\tau_i\in[0,1]$, $\lambda_i\in[0.1]$, *where* $\tau_i=1\times n$ is the vector of penalty.

The partial derivative of Equation (25) with respect to B^{-1}_i and equating to zero gives

$$\begin{aligned}&Y^T_im(x_i)m^T(x_i)B^{-1}_iY_i-2\lambda_i-\tau_iB^{-1}_i\\&=Y^T_i(x_i)m^T(x_i)Y_i\end{aligned} \tag{26}$$

However, Substituting equation (26) into (5), we have

$$\begin{aligned} E\left[\delta^2\right] &= E\left[Y^T(x_i)Y(x_i)\right] \\ &- 2E\left[\left(B^{-1}_i\right)^T\left(Y^T_i m(x_i)m^T(x_i)B^{-1}_i Y_i - 2\lambda_i - \tau_i B^{-1}_i\right)\right] \\ &+ E\left[\left(B^{-1}_i\right)^T Y_i m(x_i)m^T(x_i)Y^T_i B^{-1}_i\right] \end{aligned} \tag{27}$$

On simplifying, the proposed parametric continuous Kriging for nonstationary spatial covariance model for the location x can be expressed as

$$\begin{aligned} \delta^2_{ick_i}(x,x) &= \Im_i(x,x) - m^T(x_i)B_i^{-1}m(x_i) - 2\lambda_i B^{-1}_i \\ &- 2\tau_i B_i^{-1}\left(B_i^{-1}\right)^T \qquad i = 1,2,3,\cdots,n \end{aligned} \tag{28}$$

4. SIMULATION SET UP

The behavior of the nonstationary covariance of the adaptive covariance is investigated by conducting simulation studies with the aid of Matlab and R software (mySeed 500). Various simulations are used for the different adaptive models to examine their performance. The simulation is performed as follows:

- Datasets were generated from uniform distribution with n = 5, 10, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 350, 500, 700, 900, 950, 1000 random sample sizes. This is repeated for all the variates $m(x_i)$ and $B^{=1}_i$. More so, we generated n random variables from the uniform distribution for the following.

(a) Variates $x \in [50,1000]$, and the variates $y \in [100,2000]$ to obtain the square matrix **B**,

(b) Variates $m(x_i) \in [1,n]$ to obtain the distance between the observed and unobserved spatial processes.

(c) The assumed model for the observed spatial process is obtained as $\upsilon = 10 + 10t$, where $t \in [1,n]$ is a random variable.

- The data generated from the uniform distribution are then applied to the following models:

(i) Spherical model:

$$C_{Sph}\left(h;\sigma^2,\theta\right) = \sigma^2\left\{1 - \frac{3h}{2\theta} + \frac{1}{2}\left(\frac{h}{\theta}\right)^3\right\} \tag{29}$$

(ii) Gaussian model:

$$C_{Gau}\left(h;\sigma^2,\theta\right) = \sigma^2 \exp\left(-\frac{h^2}{\theta^2}\right) \tag{30}$$

(iii) Exponential model:

$$C_{Exp}\left(h;\sigma^2,\theta\right)=\sigma^2\exp\left(-\frac{h}{\theta}\right) \tag{31}$$

(iv) David and Williams (2002) model:

$$C_{Nd}\left(h;\sigma^2,\theta\right)=\sigma^2-m^T\left(x_i\right)B_i^T m\left(x_i\right). \tag{32}$$

(v) Proposed parametric model 1:

$$\delta_{iPM1}^2\left(x,x\right)=\Im_i\left(x,x\right)-3m^T\left(x_i\right)B_i^T m\left(x_i\right). \tag{33}$$

(vi) Proposed parametric model 2:

$$\begin{aligned}&\delta_{iPM2}^2\left(x,x\right)=\Im_i\left(x,x\right)-m^T\left(x_i\right)B_i^T m\left(x_i\right)\\&-2\left(B_i^{-1}\right)^T B_i^{-1}\end{aligned} \tag{34}$$

(vii) Proposed adaptive parametric model 1:

$$\begin{aligned}&\delta_{iAP1}^2\left(x,x\right)=\Im_i\left(x,x\right)-m^T\left(x_i\right)B_i^T m\left(x_i\right)\\&-2\lambda_i\left(B_i^{-1}\right)^T B_i^{-1}.\end{aligned} \tag{35}$$

(viii) Proposed adaptive parametric model 2:

$$\begin{aligned}&\delta_{iAP2}^2\left(x,x\right)=\Im_i\left(x,x\right)-m^T\left(x_i\right)B_i^T m\left(x_i\right)\\&-2\lambda_I B_i^{-1}\end{aligned} \tag{36}$$

(ix) Proposed adaptive parametric model 3:

$$\begin{aligned}&\delta_{ick}^2\left(x,x\right)=\Im_i\left(x,x\right)-m^T\left(x_i\right)B_i^T m\left(x_i\right)\\&-2\tau_i\left(B_i^{-1}\right)^T B_i^{-1}-2\lambda_i B_i^{-1}.\end{aligned} \tag{37}$$

(x) Cherry et al. (1996) Nonparametric model of order 1:

$$C_{CSB1}\left(h;\sigma^2\right)=2\left(\sigma^2\left(1-Cos\left(B\right)\right)\right)+\frac{2B^2}{1+\left(\frac{B^2}{2}\right)} \tag{38}$$

(xi) Cherry et al. (1996) Nonparametric model of order 3:

$$C_{CSB3}\left(h;\sigma^2\right)=2\left(\sigma^2\left(1-\frac{Sin\left(B\right)}{B}\right)\right)+\frac{2B^2}{1+\left(\frac{B^2}{2}\right)} \tag{39}$$

(xii) Huang et al. (2011)Nonparametric model modified:

$$C_{HHC}\left(h;\sigma^2\right)=\sigma^2\left\{1-\frac{\theta}{B}Sin\left(\frac{B}{\theta}\right)\right\} \tag{40}$$

(xiii) Paciorek and Schervish (2006)

$$C^{NS}\left(h;\sigma^2\right)=\sigma^2\left|\vartheta_i\right|^{\frac{1}{4}}\left|\vartheta_j\right|^{\frac{1}{4}}\left|\frac{\vartheta_i+\vartheta_j}{2}\right|^{-\frac{1}{2}} \times \exp\left(-\sqrt{\left(x_i-x_j\right)^T\left(\frac{\vartheta_i+\vartheta_j}{2}\right)^{-1}\left(x_i-x_j\right)}\right) \tag{41}$$

Where $\vartheta_i=\vartheta\left(x_i\right)$ is the kernel matrix of the covariance matrix of the Gaussian kernel centred at x_i.

- The Mean Square Prediction Error (MSPE) was used to evaluate the flexibility and performance of the different models with

$$MSPE=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\{Y\left(x_{ij}\right)-\hat{Y}\left(x_{ij}\right)\right\}^2 \tag{42}$$

- Each of the sample size are replicated 1000 times.

The following abbreviations were used in the study: Parametric model (PM 1, PM 2), Adaptive Parametric (AP 1, AP 2, AP 3), David and Williams (2002) parametric (Nd), Ordinary Kriging model Choi et al. (2013) (OK), Exponential (Ex), Spherical (Sp) and Gaussian (Ga). Huang et la. (2011) (HHC), Cherry et al. (1996) and Shapiro and Botha (1991) (CSB), Higdon et al. (1999) (HIG) and Paciorek and Schervish (2006) (PS).

Figure 1 displays 1000 sampled spatial distributions at different locations. Figures 2 and 3 are the plots for the simulated models showing the adaptive and global parameters for covariograms and semi-variograms for Exponential, Gaussian, Spherical and the David and Williams (2002) models.

In Figures 2 through 3, the spatial covariance of the adaptive model is smaller, has a bi-covariance with spherical model and with skewed covariance. Tables 1, 2 and 3 showed the results of the simulation.

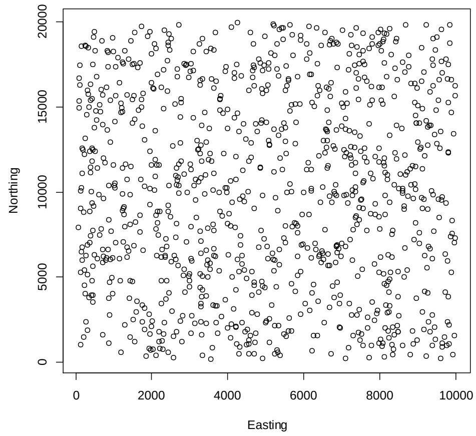


Figure 1. Spatial distributions of spatial processes in the various locations

Firstly, we consider the performance of the models developed when the parameters are fixed, then compare it to when it is adaptive. A model is fixed if the value of the parameters is kept fixed in all spatial locations.

In all cases, the adaptive parameters in APGa has the smallest standard error in MSPE; although the mean seems same for all models except for APGa 3 adaptive models. The adaptive penalized parameter of the adaptive model is smaller than the fixed model in all cases.

In Table 2, Nd and PM 2 have same spherical standard error in all and the smallest spherical standard error. The spherical standard error of Nd and PM 2 increases. The true spherical standard error is the largest across all. The spherical mean of Nd, PM 1 and PM 2 are same.

In Tables 3, CSB 3 having the lowest standard error at order 3. The standard error of HIG has the largest standard error and the smallest mean. The HIG and PS tend to zero as the lag increases.

4. DATA ANALYSIS

In this section, we prove the flexibility and efficiency of the new model by using real life data. The data were distribution of 35 Sulphate spatial data in mg/l at the construction of Tuomo/Ogbainbiri oil and gas pipeline project in South-South Nigeria. Figures 4 and 5 show the stochastic semi-variogram and variogram of the exponential, spherical and Gaussian adaptive parametric models. In Figure 4, the nugget effect $C_0 = 0$; range parameter, $\theta = 70000$ and sill of 108.

Table 4 is the summary results of the exponential, Gaussian and spherical adaptive parametric model.

The adaptive models in Table 4 give the lowest values for the standard error in

MSPE among all fitted models in Gaussian and exponential models except for exponential PM 1. Thus, the adaptive model is chosen as a better model for the data.

Table 1. Mean and Standard Error (in Parentheses $\times 10^{-05}$) of the Mean Squared Prediction Error (MSPE) Comparison of Performance for Fixed and Adaptive Models

	FIXED			Adaptive		
Model	λ	η	*MSPE*	λ^{opt}	η^{opt}	$MSPE^{opt}$
APEx 1	0.7294		9.955110(958.58)	0.01000		6.0812710(6.6758)
APEx 2	0.6633		7.096286(958.58)	0.0201		6.0812710(6.4358)
APEx 3	0.3523	0.9678	7.110231(958.58)	0.0067	0.0190	6.0812870(6.8758)
APSp 1	0.2887		8.733500(56.509)	0.0050		6.9321000(56.509)
APSp 2	0.1760		7.437500(56.509)	0.0040		6.9321000(46.329)
APSp 3	0.1569	0.9978	9.998700(56.509)	0.0029	0.0057	6.9321000(59.329)
APGa 1	0.3350		6.784226(19.874)	0.0406		4.260120(0.10474)
APGa 2	0.2883		6.416578(16.474)	0.0337		4.260120(0.10474)
APGa 3	0.1763	0.5764	4.289605(30.564)	0.0252	0.0106	2.601360(0.10474)

Table 2. Mean and Standard Error (In Parentheses $\times 10^{-05}$) of the Mean Squared Prediction Error (MSPE) Simulated Data Comparison for Parametric

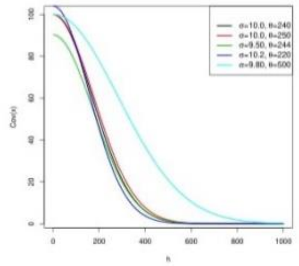
Model	Spherical	Gaussian	Exponential
OK	1.6e+07(2460)	5.950000000(910)	12.68960000(1220)
True	191.14(93939)	0.9760000(89.778)	20.8670000(10900)
Nd	6.9321(5.6509)	426.0103(0.10474)	608.12550(0.95858)
PM 1	6.9324(50.858)	1.2919e+03(0.943)	1.8382e+030(0.863)
PM 2	.93210(5.6509)	426.0120(0.10475)	608.1271(0.095863)

Table 3. The Mean and Standard Error ($\times 10^{-05}$) of the MSPE Performance for Some Models

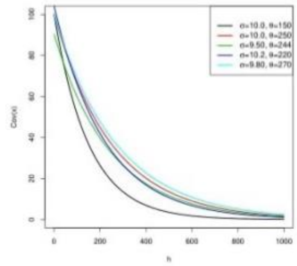
Model	Mean	Standard error
CSB 1	16.9598	0.15123
CBS 3	16.9913	4.80230
HHC	8.49170	0.00356
HIG	6.23220	48198
PS	7.42400	171400

Table 4. Mean and Standard Error (in parentheses $\times 10^{-05}$) of the Mean Squared Prediction Error (MSPE) of Adaptive Models for the Tuomo and Ogbainbiri Oil and Gas Pipeline Data

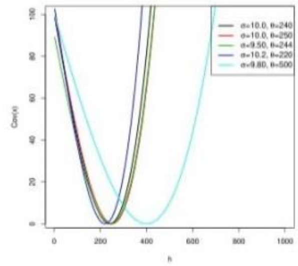
Model	Gaussian	Spherical	Exponential
True	1.92(939000)	2.96(381000)	2.01(208000)
Nd	0.62(234)	2.53(167000)	1.11(274000)
PM 1	0.10(0.00212)	1.36(151000)	0.30(287000)
PM 2	0.62(5.09)	1.92(596000)	1.12(27000)
AP 1	0.03(4.88)	1.93(592000)	0.11(272000)
AP 2	0.03(4.88)	1.93(592000)	0.11(272000)
AP 3	0.01(4.88)	1.93(592000)	0.08(271000)



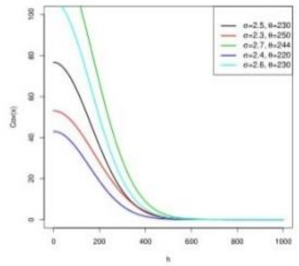
(a) Gaussian Covariogram



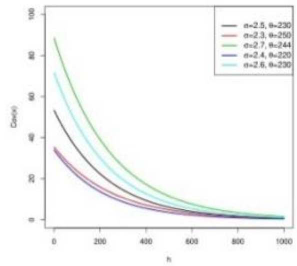
(b) Exponential Covariogram



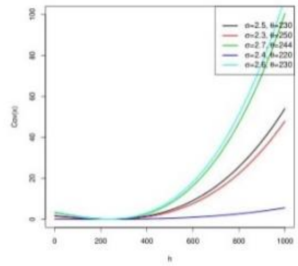
(c) Spherical Covariogram



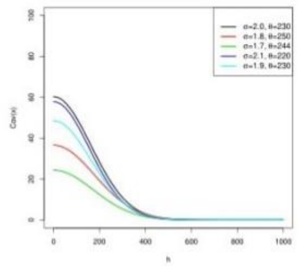
(d) Nott and Willian (2002) Gaussian Covariogram



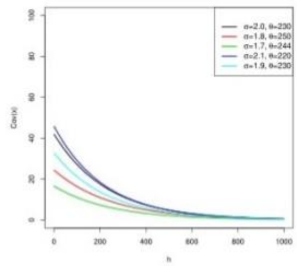
(e) Nott and Willian (2002) Exponential Covariogram



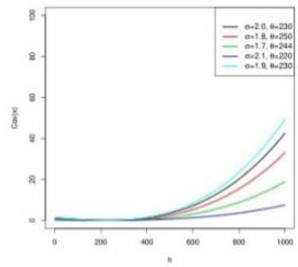
(f) Nott and Willian (2002) Spherical Covariogram



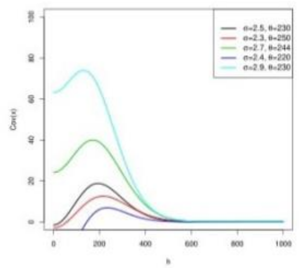
(g) Parametric Gaussian Covariogram Distribution for Model 1



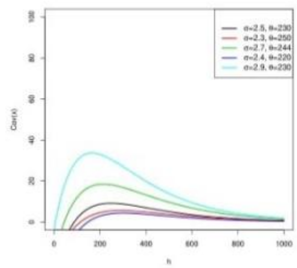
(h) Parametric Exponential Covariogram Distribution for Model 1



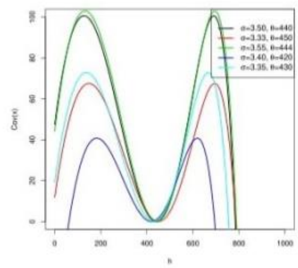
(i) Parametric Spherical Covariogram Distribution for Model 1



(j) Parametric Gaussian Covariogram Distribution for Model 2

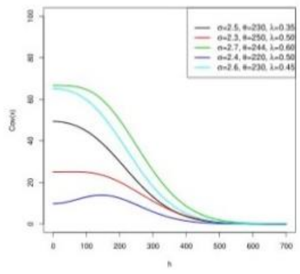


(k) Parametric Exponential Covariogram Distribution for Model 2

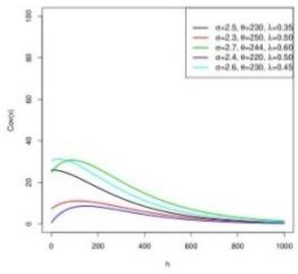


(l) Parametric Spherical Covariogram Distribution for Model 2

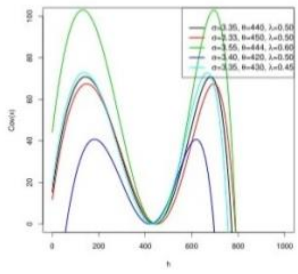
Figure 2. Covariogram Plots for Different Values of Parameters with Various Models



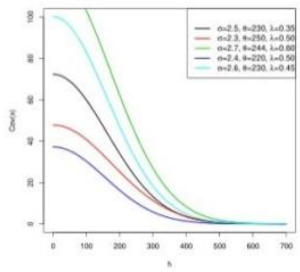
(a) Adaptive Parametric Model 1 for Gaussian Distribution



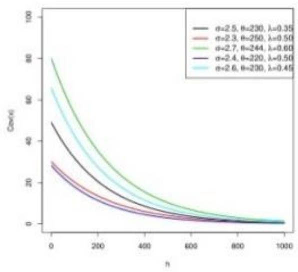
(b) Adaptive Parametric Model 1 for Exponential Distribution



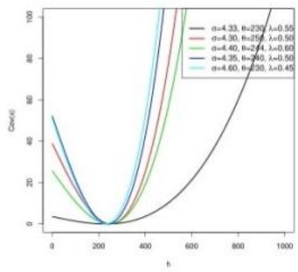
(c) Adaptive Parametric Model 1 for Spherical Distribution



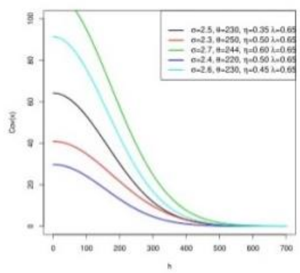
(d) Adaptive Parametric Model 2 for Gaussian Distribution



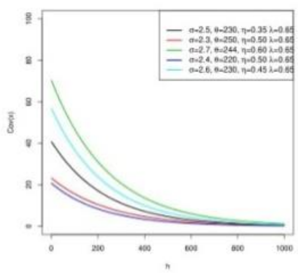
(e) Adaptive Parametric Model 2 for Exponential Distribution



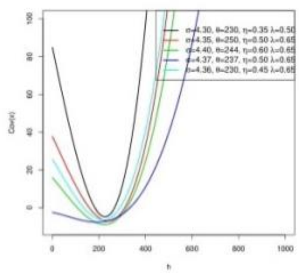
(f) Adaptive Parametric Model 2 for Spherical Distribution



(g) Adaptive Parametric Model 3 for Gaussian Distribution



(h) Adaptive Parametric Model 3 for Exponential Distribution



(i) Adaptive Parametric Model 3 for Spherical Distribution

Figure 3. Covariogram Plots for Different Values of Parameters with Various Models

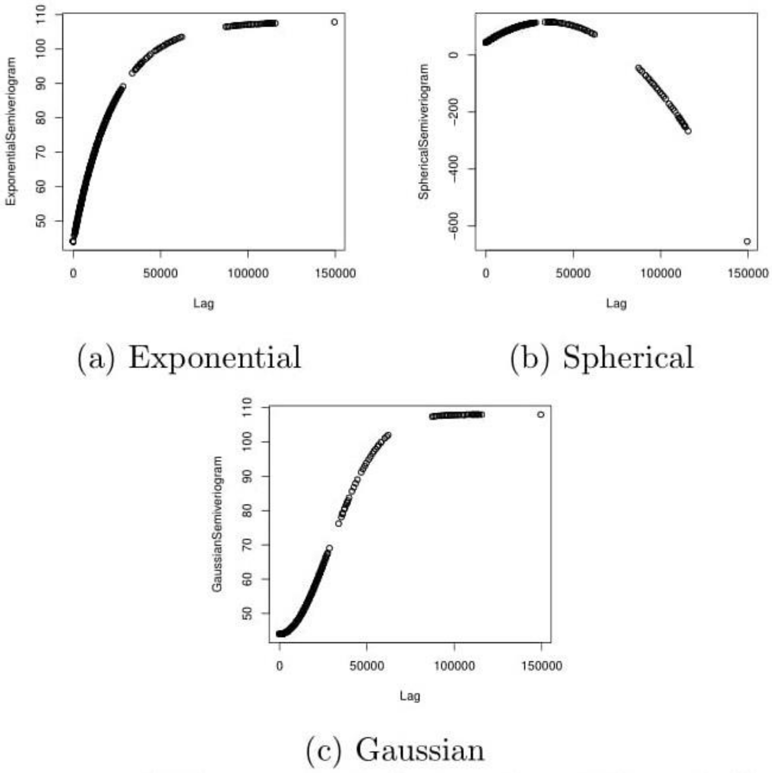


Figure 4. Semivariogram of Tuoma and Ogbainbiri Oil and Gasline Pipeline Data
(a) Exponential (b) Spherical (c) Gaussian

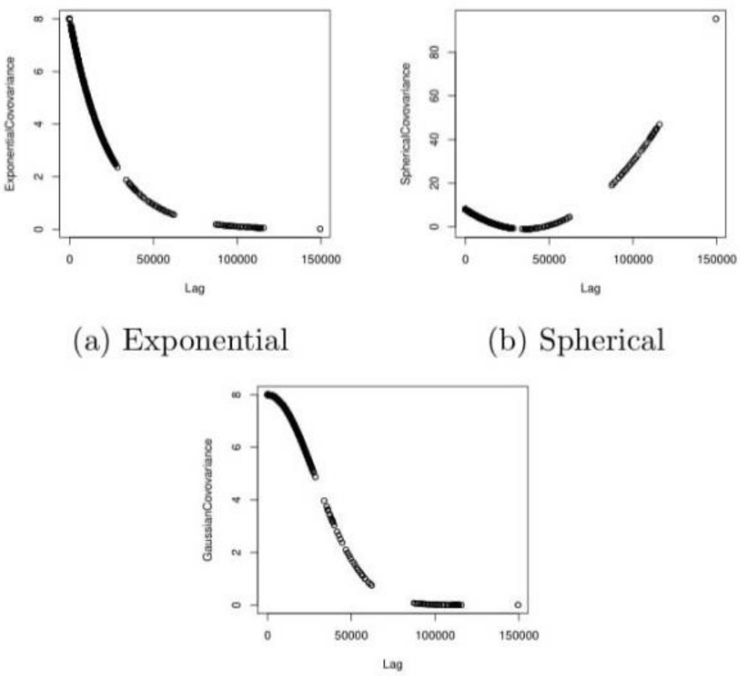


Figure 5. Covariogram of Tuoma and Ogbainbiri Oil and Gasline Pipeline Data
(a) Exponential (b) Spherical (c) Gaussian
model is chosen as a better model for the data

5. CONCLUSION

We have derived the concept of locally adaptive model for nonstationary covariance spatial processes. The idea allows each location to be fitted with its own tuning parameters instead of adopting a unified turning parameter across all locations. Furthermore, this concept produces a simple way to obtaining a valid nonnegative definite covariance function irrespective of a given covariance matrix. On comparing the results with existing models, the adaptive models have the smallest standard error. The proposed models produced an estimate for nonstationary spatial covariance that are better than David and Williams (2002) parametric model and other classical existing models.

The study developed a new family of parametric models for spatial covariance function. A closed form solution to the family of continuous model for nonstationary spatial processes is also developed. An adaptive parameter that generate the optimal value of the propose model was also developed.

The adaptive parametric models was implemented in the genetic algorithm in Matlab 2017 and R 3.5.1 programs.

6. ACKNOWLEDGEMENTS

The authors wish to acknowledge Chief Thomas Boyi Aviokirhie for his immersed effort to make this study a reality.

7. REFERENCES

Bornn, L., Shaddick, G., and Zidek, J. V. (2012). Modeling Nonstationary Processes through Dimension Expansion. Journal of the American Statistical Association, 107, 281-289.

Cherry, S., Banfield, J. and Quimby,W. F. (1996). An Evaluation of a Nonparametric Method of Estimating Semi-variograms of Isotropic Spatial Process. Journal of Applied Statistics, 23(4), 435-449.

Choi, I., Li, B., and Wang, X. (2013). Nonparametric estimation of spatial and spacetime covariance function. Journal of Agricultural, Biological, and Environmental Statistics, 18(4), 1-20.

Cressie, N. (1993). Statistics for Spatial Data. New York. Wiley.

Damian, D., Sampson, P., and Guttorp, P. (2003). Variance Modeling for Nonstationary Spatial Processes with Temporal Replication.Journal of Geophysical Research Atmospheres, 108, (D24) A rt. No. 8778.

David, J. G. and Genton, M. G. (2000). Variogram Model Selection via Nonparametric Derivative Estimation. Mathematical Geology, 32 (3), 249-270.

David, N. J. and William, D. T. M. (2002). Estimation of Nonstationary Spatial Covariance Structure. Biometrika Trust, 89(4), 819-829.

Fuentes, M. (2001). A High Frequency Kriging Approach for Nonstationary Environmental Processes. Environmetrics, 12, 469-483.

Fuentes, M. (2002). Spectral Methods for Nonstationary Spatial Processes. Biometrika, 89, 197-210.

Gilmoura, A., Cullis, B., Welham, S., Gogel, B., and Thompson, R. (2004). An Efficient Computing Strategy for Prediction in Mixed Linear Models. Journal of Computational Statistics and Data Analysis, 44, 571-586.

Guttorp, P. and Sampson, P. D. (1994). Methods for Estimating Heterogeneous Spatial Covariance Functions with Environmental Applications. In Handbook of Statistics X II: Environmental Statistics, E d. G. P. Patil and C. R. Rao, pp. 663-90. New York: Elsevier/North Holland.

Guttorp, P. Fuentes, and M. Sampson, P. (2007). Using Transforms to Analyze Space-time Processes. In Statistics Methods of Spatio-Temporal Systems, V. Isham, B. Finkelstadt, L. Held (eds). Chapman and Hall/CRC: Boca Raton, 77-150.

Haas, T. (1990a). Kriging and Automated Variogram Modeling within a Moving Window, Atmospheric Environment. Part A. General Topics, 24(7), 1759-1769.

Haas, T. (1990b). Lognormal and Moving Window Methods of Estimating Acid Deposition. Journal of the American Statistical Association, 85(412), 950--963.

Haskard, K. A., Rawlins, B. G., and Lark, R. M. (2010). A Linear Mixed Model with Nonstationary mean and Covariance for Soil potassium based on Gamma Radiometry. Journal of Biogeosciences, 7, 2081-2089.

Higdon, D. (1998). A Process Convolution Approach to Modelling Temperatures in the North-Atlantic. Journal of Environmental Engineering and Science, 5,173-190.

Higdon, D. Swall, J. and Kern, J. (1999). Non-stationary spatial modeling. In Bayesian Statistics 6, Bernardo J, Berger J, Dawid A, Smith A (eds). Oxford University Press: Oxford, UK; 761–768.

Hsu, N., Chang, Y. and Huang, H. (2012). A group Lasso Approach for Nonstationary Spatial Temporal Covariance Estimation. Environmetrics, 23, 12-23.

Huang, C., Hsing, T., and Cressie, N. (2011). Nonparametric Estimation of the Variogram and Its Spectrum. Biometrika, 98, 775-789.

Huser, R. and Genton, M. G. (2016). Nonstationary Dependence Structures for Spatial Extremes. \textit{ Journal of Agricultural, Biological, and Environmental Statistics, 12(3), doi: 10.1007/s13253-016-0247-4.

Hyoung-Moon, K., Mallick, B. K., and Holmes, C. C. (2005). Analyzing Nonstationary Spatial Data Using Piecewise Gaussian Processes. Journal of the American Statistical Association, 100(470), 653-668.

Ingebrigtsen, R., Finn, L. and Ingelin, S. (2014). Spatial Models with Explanatory Variables in the Dependence Structure. Spatial Statistics, 8, 20-38.

Jaehong, J., Mikyoung, J. and Genton, M. G. (2017). Spherical Process Models for Global Spatial Statistics. Journal of Statistical Science, 32(4), 501-513.

Katzfuss, M. (2013). Bayesian Nonstationary Spatial Modeling for very Large Datasets, Environmetrics, 24, 189-200.

Katzfuss, M., and Cressie, N. (2012). Bayesian Hierarchical Spatio-temporal Smoothing for very Large Datasets, Environmetrics, 23(1), 94-107.

Le, N., Sun, L., and Zidek, J. (2001). Spatial Prediction and Temporal

Backcasting for Environmental Fields having Monotone Data Patterns. The Canadian Journal of Statistics, 29, 529--554.

Meiring, W., Guttorp, P., and Sampson, P. (1998). Space-time Estimation of Grid-cell hourly Ozone Levels for Assessment of a Deterministic Model. Environmental and Ecological Statistics, 5, 197-222.

Matheron, G. (1963). Principles of Geostatistics, Econom. Geol. 58, 1246-1266.

Nychka, D. and Saltzman, N. (1998). Design of Air Quality Monitoring Networks. Case Studies in Environmental Statistics, 132, 51-76.

Nychka D., Wikle C., and Royle, J. (2002). Multiresolution Models for Nonstationary Spatial Sovariance Functions, Statistical Modelling, 2(4), 315.

Paciorek, C. and Schervish, M. (2006). Spatial modelling using a new class of nonstationary covariance functions. Environmetrics Journal, 17, 483-506.

Parker, R. J., Reich, B. J. and Eidsvik, J. (2016). A Fused Lasso Approach to Nonstationary Spatial Covariance Estimation. Journal of Agricultural, Biological, and Environmental Statistics, DOI: 10.1007/s13253-016-0251-8

Pickle, S. M., Robinson, T. J., Birich, J. B. and Anderson-Cook, C. M. (2008). Semi-parametric Approach to Robust Parameter Design, Journal of Statistical Planning and Inference, 138, 114-131.

Risser, D. M. and Calder, A. C. (2015). Regression Based Covariance Functions for Nonstationary Spatial Modeling. Environmetrics, 26, 284--297.

Sampson, P., Damian, D., Guttorp, P., and Holland, D. M. (2001). Deformation Based Nonstationary Spatial Covariance Modelling and Network Design. In Spatio-temporal Modelling of Environmental processes, Colecion Treballs D Informatica I Technologia, Num. 10, J Mateu,M Fuentes (eds). Universitat Jaume I:Castellon, Spain; 125-132.

Sampson, P. and Guttorp, P. (1992). Nonparametric Estimation of Nonstationary Spatial Covariance Structure. Journal of the American Statistical Association, 87,108-119.

Schmidt, M. A., Guttorp, P. and O Hagan, A. (2011). Considering Covariates in the Covariance Structure of Spatial Processes. Environmetrics Journal, 22, 487-500.

Shand, L. and Li, B. (2017). Modeling Nonstationarity in Space and Time. Biometrics, 1-10.

Shapiro, A. and Botha, J. D. (1991) Variogram fitting with a general class of conditionally nonnegative definite functions. Comput. Statist. Data Anal., 11, 87 -96.