ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2019-66181

Adaptive Models for Nonstationary Spatial Covariance Structures

Durağan Olmayan Uzamsal Kovaryans Yapılar İçin Uyarlanabilir Modeller

Joseph Thomas EGHWERIDO[a], Julian Ibezimako MBEGBU[b]

[a]Federal University of Petroleum Resources Effurun, Department of Mathematics and Computer Science, Delta State, NIGERIA
[b]University of Benin, Department of Mathematics, Edo State, NIGERIA

ABSTRACT Objective: In modeling 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 spatial domains, the spatial covariance of interest are always heterogeneous. However, this article proposed locally adaptive covariance for the spatial domain whose covariance is nonstationary in their spatial domain. The objectives of the study are to propose parametric, non-parametric and semi-parametric models for nonstationary spatial structure, continuous model for nonstationary spatial processes whose distance is far apart and to propose the adaptive weighting scheme approach that generates the optimal value for the nonparametric and semi-parametric models. **Material and Methods:** The spatial covariances are derived by applying the concept of adaptive weighting scheme approach on the covariance proposed in Nott and Dunsmuir (2002). Consequently, the local adaptive bandwidth for the nonstationary covariance was obtained for both the nonparametric and semi-parametric models. Simulations are conducted on the proposed model to examine the proposed model. **Results and Conclusion**: The results obtained are compared with existing models. The results indicate proposed spatial covariance are driven by the local bandwidths, penalty, weighted scheme, and tuning parameters. The adaptive models performed better in relation to existing covariances in terms of their mean square prediction errors (MSPE). The proposed models were further applied to real life Sulphate spatial data.

Keywords: Adaptive; locally, nonstationary; spatial covariance; variability

ÖZET Amaç: Çevresel süreçleri modellerken, dünyanın zaman içinde doğal insan kaynaklı çevresel değişikliklere nasıl tepki verdiğini açıklamak, araştırmak ve öngörmek için multi-disipliner yöntemler kullanılır. Sonuç olarak, uzamsal süreçleri incelerken uzamsal alanlar, ilgilenilen uzamsal kovaryans her zaman heterojendir. Bununla birlikte, bu makalede kovaryansı uzamsal alanlarında durağan olmayan uzamsal alan için yerel olarak uyarlanabilir kovaryans önerilmiştir. Makalenin amaçları durağan olmayan uzamsal yapılar, mesafesi çok uzak olan durağan olmayan uzamsal süreçler için sürekli model için parametrik, non-parametrik ve yarı-parametrik modeller önermek ve, non-parametrik ve yarı-parametrik modeller için optimal değeri yaratan adaptif uyarlamalı ağırlıklandırma şeması önermektir. **Gereç ve Yöntemler:** Uzamsal kovaryanslar, Nott ve Dunsmuir (2002)'de önerilen uyarlamalı ağırlıklandırma şeması yaklaşımı kovaryans üzerine uygulanarak türetilmiştir. Sonuç olarak, hem non-parametrik hem de yarı parametrik modeller için durağan olmayan kovaryans için yerel adaptif band genişliği elde edildi. Önerilen modeli değerlendirmek için önerilen model üzerine simulasyonlar yapıldı. **Bulgular ve Sonuç:** Elde edilen sonuçlar mevcut modellerle karşılaştırıldı. Bulgular önerilen uzamsal kovaryansın, yerel bant genişlikleri, ceza, ağırlıklı şema ve ayar parametreleri tarafından yönlendirildiğini göstermektedir. Uyarlanabilir modeller, ortalama karesel öngörü hataları (MSPE) açısından mevcut kovaryanslara göre daha iyi performans göstermiştir. Önerilen modeller ayrıca gerçek hayattaki sülfat uzamsal verilerine de uygulanmıştır.

Anahtar sözcükler: Uyarlanabilir; yerel, durağan olmayan; uzamsal kovaryans; değişkenlik

Interests in spatial data analysis are increasing as a result of increasing in spatial data that are present in atmospheric, hydrological, environmental, agricultural and meteorological processes. For example, the amount

Correspondence: Joseph Thomas EGHWERIDOA
Federal University of Petroleum Resources Effurun, Department of Mathematics and Computer Science, Delta State, NIGERIA/NİJERYA
E-mail: eghwerido.joseph@fupre.edu.ng

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 23 Mar 2019 ***Received in revised form:*** 25 Nov 2019 ***Accepted:*** 27 Noc 2019 ***Available online:*** 11 Dec 2019

of contamination at any given location depends on the amount of contamination in the surrounding locations prior to that period and varies. Like the meteorological processes such as; synoptic wind patterns and orographic effects exhibit inherently nonstationary processes that vary.[1-3] Thus, the understanding of the underlying spatial structures of the processes relies on variogram and covariance. But most methods of obtaining such spatial variogram and covariance assumed stationarity.[4] This often negates the nonstationary spatial variogram and covariance obtained in environmental, geological, agricultural and ecological processes. Consequently, there is need to study the nonstationarity that exist in these spatial processes.

The nonstationary processes in spatio-temporal processes can be observed in space and time framework. Let $x_1, x_2,... x_n$ be spatial locations and $Y(x)$ be a spatial process in the Euclidean space, where $x \in \Re^d$. Then, a spatial process say, $Y(x)$ is nonstationary if either the spatial covariance $\mathrm{Cov}\left[Y(x_i),\, Y(x_j)\right] i,\, j = 1,2,3,...n$ depends on the locations x_i and $\mathrm{x}_j \in \Re^d$ or $E\left[\mathrm{Y}(x)\right]$ varies over the spatial field. Our focus shall be on nonstationarity that rises from the heterogeneous correlation between observations that are separated in the same space lags, with a constant mean and variance. This assumption of nonstationarity is reasonable as large scale spatially varying means and variances are often removed before analyzing the dependence structure that exist in the spatial process $Y(x)$. Nonstationary spatial process can be diagnosed by examining the variability that exists in the empirical covariances at the same space lags. Large variability indicates nonstationarity.

Recently, modelling nonstationarity has received more attention over the years due to higher demands in practice. For example, nonstationary modelling framework was proposed through a deformation technique.[1,5-7] A covariance for a nonparametric model for variogram using the approach of convolution by representing the random field as a linear piecewise structure was proposed.[8] A more flexible nonstationary model based on convolution of stationary spatial processes using spatially varying kernels was proposed in literature.[9-11] A model based on the convolution method that allowed the latent process to be dependent was also proposed.[2,12] The generalized kernel convolution approach to a class of nonstationary covariance functions also adopted in literature.[13] 13 A nonstationary model in space using the concept of dimension expansion method was also examined.[14] This was further examined using a thin-plate spline method to obtain the nonstationary covariance that exits both in space and time.[4] The nonstationary covariance by using latent space approach by projecting the C dimension into 2D correlation structure using the covariate in the covariance.[5,15] Some examples of latent space methods for generating spatial covariance are found in literature.[16-19] A nonstationary covariance through stochastic partial differential method that allows the explanatory variables to be added to the structure was also proposed by some researchers.[20] A nonstationary covariance for nonparametric variogram through spectral representation and using a quadratic programming to solve the regularized inverse problem by employing generalized spline approach to estimate the spectrum was also proposed.[21] An optimal discretization nonstationary covariance for nonparametric variogram and covariogram with a Fourier-Bessel matrix us used in the literature.[22] A covariance for a semi-parametric model using B-spline approach to obtain the linear combination of the threshold was also observed in some literature researched.[23]

The advent of big data, led many to many to proposed other predictive models that make the computation feasible while carrying out nonstationarity modelling.[24-29]

This study proposed a locally adaptive nonstationarity in space for parametric, non-parametric and semi-parametric models, whose variability changes with location along with some asymptotic properties. Our model demonstrates advantages over the nonstationary spatial covariance in some existing literature researched as well as a combination of nonstationary spatial covariance with a stationary temporal effect.[3,4,14]

The article is organized as follows: Section 2 is a detail review of some existing nonstationary spatial process. Section 3 studies the parametric, nonparametric and semi-parametric nonstationary models, while Section 4 applies the nonstationary model to gas pipeline data and Section 5 provides the concluding remarks.

MATERIAL AND METHODS

REVIEW OF MODELS FOR NONSTATIONARY PROCESSES IN SPACE

Let $x_1, x_2,... x_n$ be the spatial locations in domain $\Re^d$ and $Y(x_1), Y(x_2),... Y(x_n)$ the spatial processes in the spatial locations. Then, the space domain $Y(.)$ can be expressed in terms of covariogram as

$$Var\,(Y(x_i), Y(x_j)) = Var\,(Y(x_i)) + Var\,(Y(x_j)) - 2Cov\,(Y(x_i), Y(x_j)) \tag{1}$$

with the variogram defined as

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

A space process $Y\left(\left[x, z\right]\right)$ for z latent dimensions with a dimension p using the lasso-penalized least squares given as

$$\hat{\theta},Z = \underset{\theta,z}{\text{argmin}} \sum_{i<j}\left(v_{i,j} - y_\theta\left[d_{i,j}(X,Z')\right]\right)^2 + \lambda \sum_{x,1\,k=1}^{p}\left\|Z'_k\right\|_1 \tag{3}$$

where $\left(d_{i,j}\left[X,Z'\right]\right)$ is the i, jth elements of the distance matrix of the augmented locations $\left[X,Z\right]_{i,j}$ are the moment estimates of $y_\theta(.)$, Z_k is the kth column of Z, $\|\,.\,\|_1$ is the L_1 norm.[14] The time is treated as replicates.

A spatial covariance for a process as a complete monotone function is given as

$$Cov\,(h) = \sum_{j=1}^{m+p} \beta_j f_j^{[q]}(h^2) \quad for\ q = p\text{-}1 \tag{4}$$

where $f_j^{[q]}$ is obtained as a B-spline of Cox-de Boor recursion formula $f_j^{\iota}(x)$ such that for $\iota \geq 1$.[30]

$$f_j^{\iota}(x) = \frac{m+1}{\iota}\left\{f_j^{\iota\text{-}1}\left(x+1\right) - \tau_{j\text{-}p}\, f_j^{[\iota\text{-}1]}\left(x\right) + \tau_{j\text{-}p+\iota+1} f_{j+1}^{[\iota\text{-}1]}\left(x\right) - f_{j+1}^{[\iota\text{-}1]}\left(x+1\right)\right\} \tag{5}$$

with

$$f_j^{[0]}(x) = \frac{m+!}{x+1}\left(\tau_{j\text{-}p+1}^{x+1} - \tau_{j\text{-}p}^{x+1}\right), \quad if\ 0 \leq \tau_{j\text{-}p}, \tau_{j\text{-}p+1} \leq 1 \tag{6}$$

otherwise 0

The $\beta = (\beta_1,..., \beta_{m+p})^T$ are the parameters that are obtained as a weighted least squares (WLS)

$$\beta_{WLS} = \text{argmin}_{\beta_{j \geq 0}} \sum_{i=1}^{L} w_j \left\{\hat{C}_E(h_i) - \sum_{j=1}^{m+p}\beta_j f_j^{[q]}(h^2)\right\}^2 \tag{7}$$

with w_i is a weighted scheme given as

$$w_i = \frac{\left|N(h_i)\right|}{\left\{1\text{-}\hat{C}_E(h_i)\right\}^2}, \quad i = 1,2,3,...n \tag{8}$$

where $N(h_i)$ is the set of data that are pair with $|N(h_i)|$ as the cardinality of $N(h_i)$ for the distance lag h_i is the distance lag.[31] The drawback of their method is the sensitivity of the spline fitting to the number of knots and thus required the Akaike information criterion and Bayesian information criterion respectively to determine the location and number of knots.

A nonparametric variogram through spectral representation approach using quadratic regularized inverse.[21] Their covariance was defined as

$$2y\,(h) = \int_0^\infty \{1\text{-}J_0\,(wh)\}r(w)f\,(w)dw, \quad hw\geq 0 \tag{9}$$

where $J_0\,(.)$ is the Bessel function of first kind of order zero, $r(w) = \frac{1\text{-}w^2}{w^2}$, $w\in(0,\infty)$ is the regularization function and $f\,(w)\geq 0$; $\int_0^\infty f\,(w)dw<\infty$ a non-decreasing function on $(0,\infty)$.

The problem with the equation (9) is that it requires a numerical solution approach to perform its computations with a weighted sum of $w = 2N_0(N_0\text{-}1)$, where N_0 are the nodes.

Furthermore, the nonstationary covariance for the spatial process with unknown distribution shall be emphasized in the next section. It is evident that the spatial process in a single unit is same, but varies from one location to another. The variable or locally adaptive bandwidths for each of these spatial locations shall also be obtained in next section.

PARAMETRIC NONSTATIONARY SPATIAL MODEL

Statistical modeling of spatial processes is often based on spatial covariogram of the spatial processes.

Let $x_1, x_2,\ldots x_n$ be the spatial locations in domain $\Re^d$ and $Y(x_1)$, $Y(x_2)$,... $Y(x_n)$ the spatial processes in the spatial locations. Minimizing the objective of as

$$Y_i(x)\text{-}m^T\,(x_1)\,B_i^{-1}\,Y_i \geq 0, \quad for\ i = 1,2,3,\ldots n. \tag{10}$$

Subject to the constraint

$$B_i^{-1} \geq 0, \quad i = 1,2,3,\ldots n. \tag{11}$$

where $Y_i(x) = n$ x 1 column vector of independent nonstationary realizations; $m^T\,(x_1) = \Im_i(x - x_i)$a n x 1 column vector of cross-covariance between the sample at the observed and sample at the unobserved locations; $Y_i = (Y_i\,(x_1), Y_i\,(x_2),\ldots,Y_i\,(x_n))^T$; $\beta_i = \Im_i(x_j - x_k)$, n x n covariance matrix between the process at locations x_j and x_k for $j,k = 1,2,3,\ldots n$. $\delta_i(x)$ is a column vector that has a zero mean nonstationary error term.

Thus, by Karush-Kuhn-Tucker method we defined the spatial minimization problem above as

$$L(B_i^{-1},\lambda) = \left|\left|Y_i\,(x) - m^T\,(x_i)\,B_i^{-1}Y_i\right|\right|^2 + \lambda_i\left|\left|B_i^{-1}\right|\right|_2^2 \quad \lambda_i\in[0,1] \tag{12}$$

where, λ_i are n x 1 vectors of adaptive turning parameters such that

$$\lambda_i = \frac{\sum_{p=1}^{n}\sum_{j=1}^{n} B_{pj} - B_i}{(n - 1)\sum_{p=1}^{n}\sum_{j=1}^{n} B_{pj}}.$$

The solution to the optimization problem in equation (12) is given as

$$\delta_{\lambda i}^2\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T\,(x_i)B_i^{-1}\,m\left(x_j\right) - 2\lambda_i\left(B_i^{-1}\right)^T B_i^{-1} \quad i, j = 1,2,3,\ldots n \tag{13}$$

where $\Im(x_i - x_j)$ is the variance of $Y_i(x)$ called the sill.

For some possible values of λ_i equation (13) reduces to model one.[3] When λ_0=0. When $\lambda_i = Y_i Y_i^T m^T(x_i)\, m^T(x_j)$. Equation (13) yields

$$\delta^2_{\lambda i}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - 3m^T(x_i)B_i^{-1} m\left(x_j\right) \quad i, j = 1,2,3,...n. \tag{14}$$

Otherwise, for $\lambda_i = 1$, equation (13) becomes

$$\delta^2_{\lambda i}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i)B_i^{-1} m\left(x_j\right) - 2\left(B_i^{-1}\right)^T B_i^{-1} \quad i, j = 1,2,3,...n. \tag{15}$$

It thus follows that equation (13) depends on the adaptive turning parameters. Hence, we have been able to propose an optimized adaptive spatial covariance that considered the lags as an important factor for nonstationary.

NONPARAMETRIC NONSTATIONARY SPATIAL MODEL

Statistical modelling of spatial processes is often based on Gaussian processes. This facilitates prediction, but normality is not necessarily an adequate modelling assumption for process whose distribution is uncertain. Hence, a spatial process can be expressed in the nonparametric form if the spatial covariance is positive definite. Consequently, the nonparametric form of the adaptive nonstationary process is obtained by minimizing

$$L_{np}(B_i^{-1}, \lambda) = \left\|Y_i(x) - m^T(x_i)\, B_i^{-1} W_0 Y_i\right\|_2^2 + \lambda_i \left\|B_i^{-1}\right\|_2^2 \quad \lambda_i \in [0,1] \tag{16}$$

where $W_0 = diag\left(d_{01}, d_{02}, d_{03}, ..., d_{0n}\right)$ an n x n matrix such that d_{0i} are the simplified Gaussian kernel weights for the locations $x_1, x_2, ... x_n$, given as

$$d_{0i}^{\ker} = \frac{K_i\left(\frac{|x_0 - x_i|}{b_i^2}\right)}{\sum_i^m K_i\left(\frac{|x_0 - x_i|}{b_i^2}\right)}$$ where b_i are the adaptive local bandwidth. The K is a kernel function given as

$$K_i\left(\frac{|x_0 - x_i|}{b_i^2}\right) = \exp\left(-\left(\frac{|x_0 - x_i|}{b_i^2}\right)\right)$$. Following same process, we can express equation (16) as

$$Q\left(B_i^{-1}\right) = \left(\hat{C}ov\,(h) - m^T(x_i)B_i^{-1} Y_i\right)^T W_0 \left(\hat{C}ov\,(h)\, m^T(x_i)B_i^{-1} Y_i\right) \tag{17}$$

Thus, the resulting nonparametric spatial covariance has explicit representation as

$$Cov_{np}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i)B_i^{-1} W_0\, m\,(x_j) - 2\lambda_i\left(B_i^{-1}\right)^T B_i^{-1} \quad i, j = 1,2,3,...n. \tag{18}$$

Provided $m^T(x_i)B_i^{-1} W_0\, m\,(x_j)$ is positive definite. Furthermore, investigating the behaviour of the adaptive vectors of tuning parameters λ_i at 0,1, and $Y_i Y_i^T W_0 m^T(x_i)$ for $0 \le i \le n$. If $\lambda_i = 0$ and $W_0 = 1$, equation (18) reduces to model one.[3] Otherwise, for $\lambda_i = 0$

$$Cov\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i)B_i^{-1} W_0\, m\,(x_j) \quad i, j = 1,2,3,...n. \tag{19}$$

For $\lambda_i = 1$, equation (18) becomes

$$Cov_{np1}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i)B_i^{-1} W_0 m\,(x_j) - 2\left(B_i^{-1}\right)^T B_i^{-1} \quad i, j = 1,2,3,...n. \tag{20}$$

and for $\lambda_i = Y_i Y_i^T W_0 m^T(x_i) m\,(x_i)$ *and* $\lambda_i \geq 0$. equation (18) becomes

$$Cov_{np2}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - 3m^T(x_i)B_i^{-1} W_0 m\,(x_j) \quad i, j = 1,2,3,...n. \tag{21}$$

The above models are the adaptive nonparametric models when stationarity and distributions are violated.

In the previous application of the adaptive nonparametric models, the spatial processes in the neighbourhood of the locations were used to predict the spatial variable at the unsampled location. Suppose the processes are not within the neighbourhood, then, we shall propose a model for such processes through kriging.

NONPARAMETRIC CONTINUOUS KRIGING FOR NONSTATIONARY SPATIAL MODEL

Prediction of a spatial variable at an unobserved location often depends on the available variable in the neighbourhood of such sample. Now, suppose the locations x_i and x_j are far apart from the unobserved location, then, the covariance $\Im\left(x_i, x_j\right)$ approaches zero as the lag h tends to infinity. Hence, a quantity η is introduced to penalize the parameter B_i^{-1} such that such observations within such neighbourhood are used to obtain the predicted variables. Thus, minimize the estimated variance subject to the penalized quantity using lasso regression approach

$$L\left(B_i^{-1}, \lambda, \tau\right) = Y_i^T\left(x\right) Y_i\left(x\right) - 2Y_i^T W_0 m^T\left(x_i\right) B_i^{-1} Y_i + Y_i^T W_0 \left(B_i^{-1}\right)^T m\left(x_i\right) m^T\left(x_i\right) B_i^{-1} Y_i - 2\lambda_i B_i^{-1} - 2\eta\, B_i^{-1}\left(B_i^{-1}\right)^T \tag{22}$$

For some $\tau \in \left[0,1\right]$, $\lambda_i \in \left[0,\infty\right]$, $\eta \in \left[0,1\right]$. The partial derivatives of equation (22) with respect to the penalty and the mixing parameters, and the spatial weights give the nonstationary spatial covariance model for the continuous processes as

$$Cov_{npc}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i)B_i^{-1} W_0 m\,(x_j) - 2\lambda_i\left(B_i^{-1}\right)^T - 2\eta\left(B_i^{-1}\right)^T B_i^{-1} \quad i, j = 1,2,3,...n. \tag{23}$$

Theorem 3.1

Let $f(x - x_i) = \exp\left(-\left\| x - x_i\right\|^2\right)$ *be the distribution function and* $\upsilon_i(x) = \dfrac{f(x - x_i)}{\sum_{j=1}^{n} f(x - x_i)}$ *be the weight function, then* $\upsilon_i(x)$ *approaches zero as the* $\left\| x - x_i\right\| \to \infty$ *for i = 1,2,3,...n.*

Proof of Theorem 3.1

Since $\upsilon_i(x)$ *are weight functions, then, it must sums up to one. Thus,*
$f(x - x_1) = \exp\left(-\left\| x - x_1\right\|^2\right), f(x - x_2) = \exp\left(-\left\| x - x_2\right\|^2\right),..., f(x - x_n) = \exp\left(-\left\| x - x_n\right\|^2\right)$. *Let*

$$A = \sum_{j=1}^{n} f(x - x_i) = \exp\left(-\left\| x - x_1\right\|^2\right) + , \exp\left(-\left\| x - x_2\right\|^2\right) + \exp\left(-\left\| x - x_3\right\|^2\right) + ... + \exp\left(-\left\| x - x_n\right\|^2\right),$$

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

Hence, $\sum_{i=1}^{n} \upsilon_i(x) = \sum_{i=1}^{n} \dfrac{\exp\left(-\left\| x - x_i\right\|^2\right)}{A} = 1.$

Next, we shall show that $\dfrac{\exp\left(-\left\| x - x_i\right\|^2\right)}{A} \to 0$. *We observed that* $\upsilon_i(x) = \dfrac{\exp\left(-\left\| x - x_i\right\|^2\right)}{\sum_{j=1}^{n} \exp\left(-\left\| x - x_i\right\|^2\right)}$. *But* $\exp\left(-\left\| x - x_i\right\|^2\right) = 0$, *as* $\left\| x - x_i\right\|^2 \to \infty$. *Clearly,* $\upsilon_1(x) \to 0$

SEMI-PARAMETRIC ADAPTIVE MODEL FOR NONSTATIONARY SPATIAL STRUCTURE

Semi-parametric nonstationary spatial covariance involves fitting spatial processes that is robust to mis-specification through an adaptive mixing parameter λ_i.The adaptive mixing parameters λ_i satisfies the condition

$0 \leq \lambda_i \leq 1$. Let $x_1, x_2, \ldots x_n$ be the spatial locations in domain $\Re^d$ and $Y(x_1), Y(x_2), \ldots Y(x_n)$ the spatial processes in the spatial locations. Hence, a linear combination of the parametric and nonparametric models will yield

$$L\left(B_i^{-1}, \lambda\right) = \| Y_i\left(x\right) - m^T B_i^{-1} Y_i \|_2^2 + \lambda_i \left(\| Y_i\left(x\right) - m^T\left(x_i\right) B_i^{-1} W_0 Y_i \|_2^2 \right)\left(1 - \| Y_i\left(x\right) - m^T\left(x_i\right) B_i^{-1} W_0 Y_i \|_2^2\right) \quad (24)$$

where 1 *is n x* 1 column vector and W_0 is a diagonal matrix of kernel weights. However, for convenience, the computational form of equation (24) for the locations *i, j* is given as

$$Cov_{sp}\left(x_i, x_j\right) = \Im\left(x_i - x_j\right) - m^T(x_i) B_i^{-1} m\,(x_j) + 2\lambda_i \left(m^T(x_i)\, B_i^{-1} W_0\, m\,(x_j)\right)\left(1 - m^T(x_i) B_i^{-1} m(x_j)\right) \quad i, j = 1,2,3,\ldots n. \quad (25)$$

It is clear from the theorem (3.1) that we do not need numerous sampled values to predict a variable at the unsampled location. Thus, for our proposed nonstationary covariance $Cov\left(x_i, x_j\right)$, simple kriging variance for nonstationary spatial process can be easily obtained as

$$\sum_{i=1}^{n} \upsilon_i\,(x) Cov\left(x_i, x_j\right) \quad (26)$$

The simple kriging predictor is given as[32]

$$\sum_{i=1}^{n} \upsilon_i\,(x)\; m^T(x_i) B_i^{-1} Y_i(x) \quad (27)$$

The nonparametric and semi-parametric nonstationarity depend on the bandwidths as the smoothing parameter. Hence, we shall derive the bandwidths such that they sum up to one.

ADAPTIVE LOCALLY BANDWIDTH

The performance of Local Linear Estimator depends on how the bandwidth b_i are chosen. Hence, we choose the bandwidth $b_i^{opt} = \eta^{opt} b_i$ where η^{opt} is an optimal real number that minimizes the mean square prediction error (MSPE).

Theorem 3.2

Let y_i *for* $i = 1,2,3,\ldots n$ *be the set of individual realization from a spatial process from the locations* x_i *such that* $\gamma = \sum_{i=1}^{n} y_i$. *Then,* $\theta_i = \frac{y_i}{\gamma}$ *it implies that* $\sum_{i=1}^{n} \theta_i = 1$.

Proof of Theorem 3.2

Since $\gamma = \sum_{i=1}^{n} y_i$, then, $\sum_{i=1}^{n} \theta_i = \sum_{i=1}^{n} \frac{y_i}{\gamma} = \frac{\sum_{i=1}^{n} y_i}{\gamma}$

It is easy to see that

$$\sum_{i=1}^{n} \gamma = 1$$

Thus, we generate the locally adaptive bandwidths in the following manner:

Let y_i for $i = 1,2,3,\ldots n$ are the spatial processes observed at the locations x_i with $\gamma = \sum y_i$. Let the weight be denoted by

$$\theta_i = \frac{y_i}{\gamma} \quad i = 1,2,3,\ldots n \quad (28)$$

Clearly,

$$\sum_{i=1}^{n} \theta_i = 1 \quad (29)$$

Suppose we denote the local bandwidths by b_i such that $\sum_{i=1}^{n} b_i = 1$. Then, we can represent the local bandwidths as

$$b_i = (1 - \theta_i) = \left(1 - \frac{y_i}{\gamma}\right) = \left(\frac{\gamma - y_i}{\gamma}\right) \quad i = 1,2,3,...n \tag{30}$$

Taking sum of both sides of the equation (30), we have:

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} \left(\frac{\gamma - y_i}{\gamma}\right) \quad i = 1,2,3,...n \tag{31}$$

When equation (31) is simplified, we have

$$\sum_{i=1}^{n} b_i = \frac{n\gamma - \sum y_i}{\gamma} = \frac{n\gamma - \gamma}{\gamma} = (n - 1) \tag{32}$$

Obviously, $\sum_{i=1}^{n} b_i \neq 1$. However for $\sum_{i=1}^{n} b_i = 1$, we multiply the right-hand side of equation (32) by $\frac{1}{n-1}$. Thus,

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} \left(\frac{\gamma - y_i}{\gamma}\right)\left(\frac{1}{n - 1}\right) = \sum_{i=1}^{n} \frac{\gamma - y_i}{\gamma(n - 1)} \tag{33}$$

Now, the local bandwidth is denoted as

$$b_i = \frac{\gamma - y_i}{\gamma(n - 1)} \quad i = 1,2,3,...n \tag{34}$$

provided,

$$0 \leq \frac{\gamma - y_i}{(n - 1)} \leq 1 \quad i = 1,2,3,...n \tag{35}$$

RESULTS

SIMULATION STUDY

The behavior of the nonstationary covariance of the adaptive covariance is investigated by conducting simulation studies with the aid of Matlab and R software. Datasets were generated from uniform distribution with replication number n = 1000 random samples. The Mean Square Prediction Error (MSPE) at different locations was used to evaluate the prediction performance of the different models with

$$MSPE = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{Y(x_{ij}) - \hat{Y}(x_{ij})\right\}^2 \quad i = 1,2,3,...n$$

Table 1 reports simulated results for the adaptive covariance proposed with fixed and locally adaptive parameters. The adaptive covariance is fixed if the values of the parameters are kept constant in all spatial locations.

TABLE 1: Mean and standard error (in parentheses) of the Mean Squared prediction Error (MSPE) performance for adaptive covariance.

	Fixed					Adaptive			
Simulated Models	b	λ	η	τ	MSPE	b^{opt}	η^{opt}	τ^{opt}	$MSPE^{opt}$
APEx 1					609.5511(9.5858e+03)				608.1271(9.5858e+03)
APEx 2					609.6286(9.5858e+03)				608.1271(9.5858e+03)
APEx3					611.0231(9.5858e+03)				608.1287(9.5858e+03)
APSp1					6.9321(5.6509e-09)				6.9321(5.6509e-09)
APSp 2					6.9321 (5.6509e-09)				6.9321(5.6509e-09)
APSp 3					6.9321 (5.6509e-09)				6.9321(5.6509e-09)
APGa 1					427.4226(1.0474e+04)				426.0120(1.0474e+04)
APGa 2					427.6278(1.0474e+04)				426.0120(1.0474e+04)
APGa 3					428.9605(1.0474e+04)				26.0136(1.0474e+04)
NP 1	0.7345		41470		0.0085(2.0413e-09)	0.0010	28000		0.0085(2.5237e-11)
NP 2	0.6391		20392		0.0086(0.0019e-00)	0.0080	17000		0.0085(1.6491e-10)
NP 3	0.7832	0.2988	10000		0.0085(2.3004e-04)	0.0070	18000		0.0085(1.7901e-11)
NPC	0.4565	0.9784	30500	0.9978	0.0085(5.3283e-05)	0.0050	22000	0.0764	0.0085(2.5237e-11)
SAPEx	0.7888	0.5841	17649		3.5251e+03(1.047e+04)	0.0020	15000		0.5181(1.0474e+04)
SAPSp	0.7888	0.8544	23040		3.7577e-04(5.6509e-09)	0.0100	16800		0.0085(5.6509e-09)
SAPGa	0.7888	0.5988	24000		4.4802e+03(9.586e+03)	0.0100	23700		0.7396(9.5859e+03)

AP: Adaptive Parametric, NP: Nonparametric, SAP: Semi-parametric Adaptive Parametric, Ex: Exponential, Sp: Spherical, Ga: Gaussian. 1, 2, 3 indicates $\lambda_i=1, \lambda_i\geq 0, \lambda\neq 0$.

We shall apply real life spatial data to examine the performances of the proposed adaptive models and further compare the results with existing models in literature.

REAL LIFE DATA APPLICATION

To demonstrate the potentials of the adaptive nonstationary covariance, we applied the nonstationary models to data set and make comparison with existing models. The datasets sampled for sulphate in the event of oil spills arising from the operation of the construction of “6 x 40” Km Tuomo To Ogbainbiri Oil and Gas Pipeline Project in South-South of Nigeria. The datasets are obtained as secondary data from the environmental impact assessment conducted by Department of Environmental and Toxicology, Federal University of Petroleum Resources, Effurun, Delta State, Nigeria, FUPRE (2018). The sulphate data are measured in mg/l. The exponential continuity function was used to evaluate the nugget effect, sill and the range parameters. Figure 1(a)-1(c) through Figure 2(a)-2(c) are the plots of the real life data for semivariograms and covariograms for exponential, spherical and Gaussian models.

DISCUSSION

In this article, an adaptive model for nonstationary spatial covariance has been explicitly derived. However, the adaptive smoothing parameter was also presented. Performances of the smoothing parameter based on MSPE selection criteria are compared with the help of a simulation study and real life data.

However, based on the simulation in Table 1, the adaptive parameters in NP 3 have the smallest standard error in MSPE with same mean. The adaptive covariance has the lowest bandwidth compare to the fixed. More so, the fixed tuning parameter increases with the increase in parameters. The semi-parametric models have the lowest standard error in spherical model in all replicates for both adaptive and fixed models and its standard error increases as the replicate increases.

Firstly, the MSPE of parametric models in Table 2 shows that Nd and PM 2 have same standard error in spherical models in all replicates and increases as the replicate increases. The true spherical standard error is the largest across all replicates and decreases as replicate increases. The mean and standard error of Gaussian PM 1 increases as the replicate increases.

In addition, in Table 3, the parameters of the CSB 3 have the lowest standard error. The standard error of CSB 1 and HHC models increases as the replicate increases.

In Figure 1(a)-1(c), the nugget effect C0 equal zero with a range parameter, α = 70000, σ^2 = 8 and sill of 108.

Furthermore, the results of the real life application show that the mean of the locally adaptive and the fixed parameters are same in all cases for adaptive and fixed exponential models in Table 4. The fixed NP 2 has the least bandwidth. We quickly observed that the more the tuning parameter, the less the bandwidth; thus, the correlation. The bias of the variable bandwidth and tuning parameters are smaller than the fixed bandwidth. In Table 5, the exponential model has the smallest standard error in PM 1 and Nd models. The Gaussian model has the largest standard error in all the models. Spherical PM 1 has the smallest mean.

However, in Table 6, the mean of the HHC nonparametric model is the smallest but the standard error of the CSB models is the smallest.

TABLE 2: Mean and standard error (in parentheses) of the Mean Squared Prediction Error (MSPE) for parametric model.

Model	Spherical	Gaussian	Exponential
True	191.14(9.3939e+00)	0.9760(8.9778e-04)	20.867(0.0109e+00)
Nd	6.9321(5.6509e-09)	426.0103(1.0474e+04)	608.1255(9.5858e+03)
PM 1	6.9324(5.0858e+08)	1.2919e+03(9.43e+04)	1.8382e+03(8.63e+04)
PM 2	6.9321(5.6509e-09)	426.0120(1.0475e+04)	608.1271(9.5863e+03)

Nd: Nott and Dunsmuir, PM: Parametric Model, 1, 2 indicates $\lambda_i=1, \lambda_i\geq 0$.

TABLE 3: Mean Squared Prediction Error (MSPE) performance for nonparametric model.

Nonparametric Models	Mean	Standard Error
CSB 1	16.9598	1.5123e+05
CSB 3	16.9913	4.8023e-05
HHC	8.4917	356.1304

(HHC), (CBS) Nonparametric Model 1 and 2. [21,33-34]

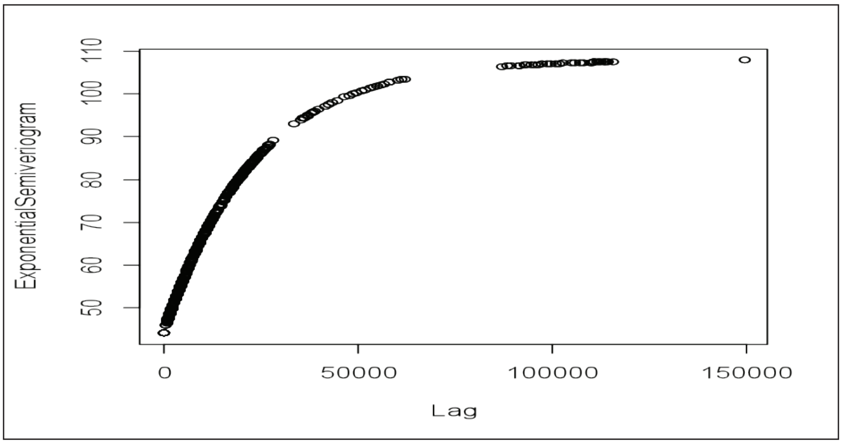

FIGURE 1(a): Exponential semivariogram oil and gas pipeline data.

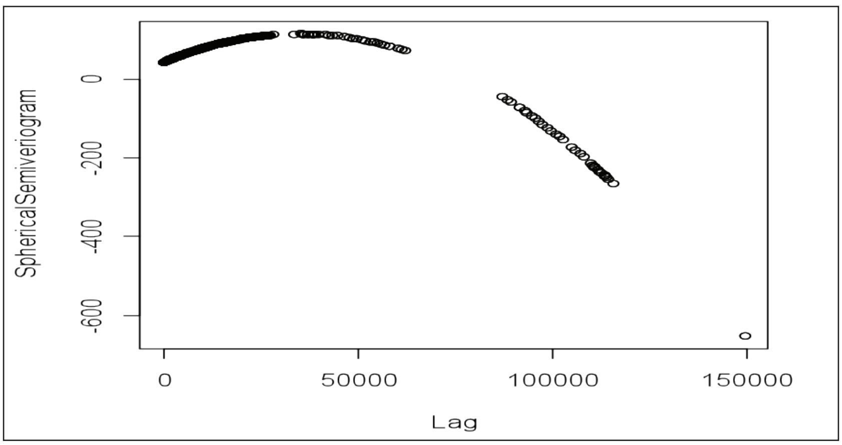

FIGURE 1(b): Spherical semivariogram oil and gas pipeline data.

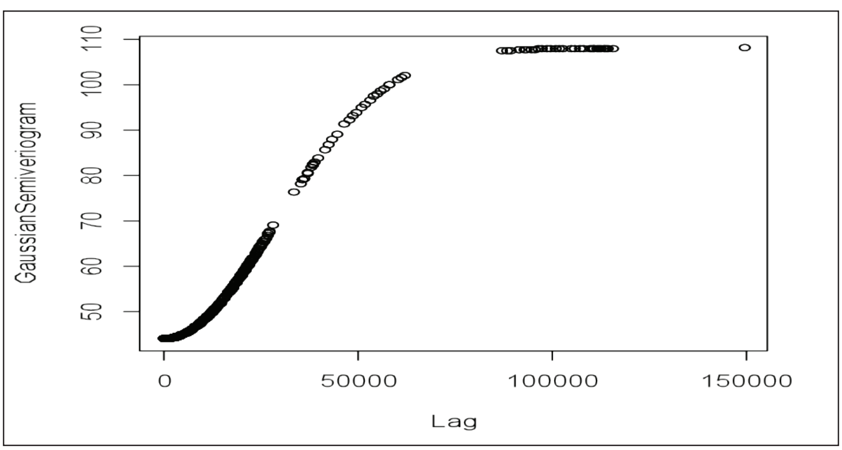

FIGURE 1(c): Gaussian semivariogram oil and gas pipeline data.

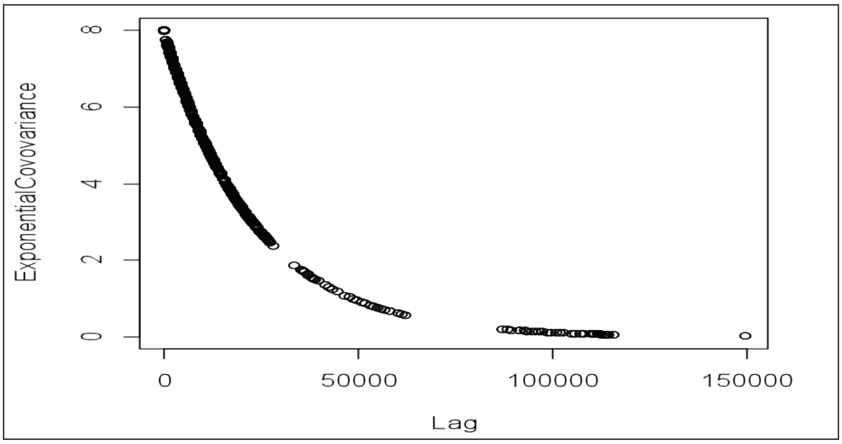

FIGURE 2(a): Exponential covariogram oil and gas pipeline data.

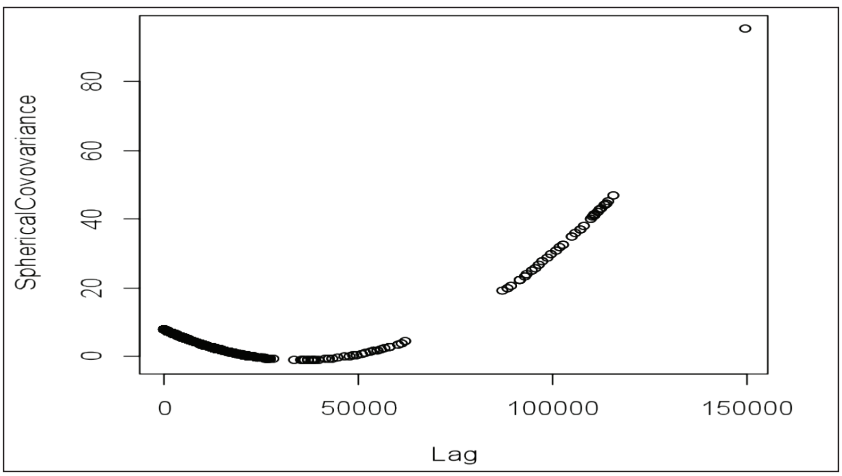

FIGURE 2(b): Spherical covariogram oil and gas pipeline data.

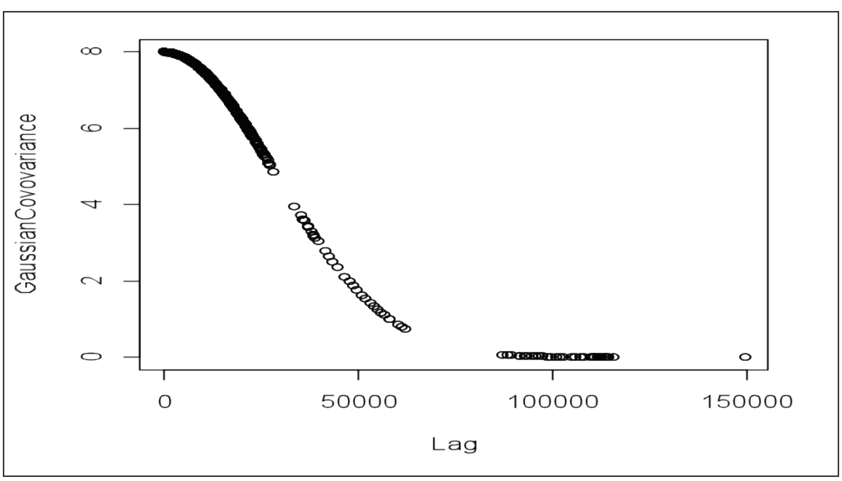

FIGURE 2(c): Gaussian covariogram oil and gas pipeline data.

TABLE 4: Mean and standard error (in parentheses) of the Mean Squared Prediction Error (MSPE) of Tuomo and Ogbainbiri oil and gas pipeline data for adaptive models and parameters.

	Fixed					Adaptive			
Models	b	λ	η	τ	MSPE	b^{opt}	η^{opt}	τ^{opt}	$MSPE^{opt}$
APEx 1					1.1161(8.7483e-16)				1.1161(7.6339e-16)
APEx 2					1.1161(8.7483e-16)				1.1161(6.9973e-16)
APEx 3					1.1162(8.7483e-16)				1.1161(7.0210e-16)
APSp 1					0.6191(2.3405e-08)				0.6191(2.3429e-08)
APSp 2					0.6191(2.3405e-08)				0.6191(2.7969e-08)
APSp 3					0.6191(2.3405e-08)				0.6191(2.7971e-08)
APGa 1					1.2426e+07(1.6764e+06)				3.6644e+06(1.6498e+06)
APGa 2					4.5358e+06(1.6764e+06)				4.5357e+06(2.1778e+06)
APGa 3					5.6240e+06(1.6764e+06)				3.6643e+06(2.4633e+06)
NP 1	0.2204		68		0.9764(2.4036e-13)	0.0279	50		0.9764(9.5086e-14)
NP 2	0.0765		56		0.9764(1.1621e-10)	0.0057	67		0.9764(3.9603e-13)
NP 3	0.1519	0.2988	93		0.9764(1.4544e-11)	0.0296	81		0.9764(2.6647e-14)
NPC	0.1271	0.9784	67	0.9978	0.9764(2.4924e-11)	0.0385	70	0.5764	0.9764(3.8887e-14)
SAPEx	0.5588	0.9711	50		1.1154 (3.5684e-07)	0.0286	38		1.1161 (3.3728e-07)
SAPSp	0.5588	0.7734	54		0.6155 (2.3405e-08)	0.0290	45		0.6191 (2.2547e-08)
SAPGa	0.5588	0.8788	51		4.5351e+06 (136.6298)	0.0277	47		4.5356e+06 (1.2249e+03)

AP: Adaptive Parametric, NP: Nonparametric, SAP: Semi-parametric Adaptive Parametric, Ex: Exponential, Sp: Spherical, Ga: Gaussian. 1, 2, 3 indicates $\lambda_i=1, \lambda_i\geq 0, \lambda\neq 0$.

TABLE 5: Mean and standard error (in parentheses) of the Mean Squared Prediction Error (MSPE) of Tuomo and Ogbainbiri oil and gas pipeline data for parametric models.

Models	**Spherical**	**Gaussian**	**Exponential**
True	191.14(939.39e-00)	2.9582e+17(3.8126e+32)	20.8665(1.089e-00)
Nd	0.6191(2.3405e-08)	4.5356e+06(1.6764e+06)	1.1161(8.7483e-16)
PM 1	0.0953(2.1064e-08)	1.3607e+07(1.5108e+07)	5.3010(7.8734e-16)
PM 2	0.6191(5.0899e-05)	1.9232e+08(5.9590e+36)	1.1161(4.2670e-05)

Nd: Nott and Dunsmuir, PM: Parametric Model, 1, 2 indicates $\lambda_i=1, \lambda_i\geq 0$.

TABLE 6: Mean and standard error of the Mean Squared Prediction Error (MSPE) of Tuomo and Ogbainbiri oil and gas pipeline data for Huang et la. (2011), Cherry et al. (1996) and Shapiro Botha (1991) nonparametric models.

Nonparametric Models	**Mean**	**Standard Error**
CBS 1	70.0285	7.3037e+04
CBS 2	68.5761	3.8341e-08
HHC	34.1647	1.9117e+04

(HHC), (CBS) Nonparametric Model 1 and 2. [21,33-34]

CONCLUSION

We have derived the concept of locally adaptive models for nonstationary covariance spatial processes. The idea allows each location to be fitted with its own tuning parameters instead of adopting a unified tuning parameter across all locations. Furthermore, this concept produces a simple way to obtain a valid nonnegative definite covariance function irrespective of a given covariance matrix. On comparing the results with existing covariance, the adaptive covariances has smaller value for bias and produce an estimate for nonstationary spatial covariance that is better than the covariance and the other classical existing covariance.[3]

Acknowledgement

The first author will like to acknowledge Prof. A. Rim for the immerse discussion and Dr. Emmanuel Esemedafe for editing this work.

Author contributions

The study proposed a new family of nonstationary covariance for parametric, nonparametric and semi-parametric models. It also develops a function that generates the optimal value for the bandwidth selection.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Supporting information

No supporting information is available as part of the online article.

Authorship Contributions

Idea/Concept: *Joseph Thomas Eghwerido;* ***Design:*** *Joseph Thomas Eghwerido;* ***Control/Supervision:*** *Julian Ibezimako Mbegbu;* ***Data Collection and/or Processing:*** *Joseph Thomas Eghwerido;* ***Analysis and/or Interpretation:*** *Joseph Thomas Eghwerido;* ***Literature Review:*** *Joseph Thomas Eghwerido;* ***Writing The Article:*** *Joseph Thomas Eghwerido;* ***Critical Review:*** *Julian Ibezimako Mbegbu;* ***References and Fundings:*** *Joseph Thomas Eghwerido.*

REFERENCES

1. Guttorp P, Sampson P. Methods for estimating heterogeneous spatial covariance functions with environmental applications, Hand-book of statistics Wiley, 1994; volume 12. [Crossref]
2. Fuentes M. A new high frequency kriging approach for nonstationary environmental processes. Environmetrics, 2001; 12:469-483. [Crossref]
3. Nott DJ, Dunsmuir WT M. Estimation of nonstationary spatial covariance structure. Biometrika Trust, 2002; 89(4):819-829. [Crossref]
4. Shand L, Li B. Modeling nonstationarity in space and time. Biometrics, 2017; 1-10. [Crossref] [PubMed] [PMC]
5. Sampson PD, Guttorp P. Nonparametric estimation of nonstationary spatial covariance structure. Journal of the American Statistical Association, 1992; 87:108-119. [Crossref]
6. Schmidt AM, O'Hagan, A. Bayesian inference for non- stationary spatial covariance structure via spatial deformations. JRSSB, 2003; 65:743-758. [Crossref]
7. Aberg S, Lindgren F, Malmberg A, Holst J, Holst U. An image warping approach to spatio-temporal modeling. Environmetrics, 2005; 16:833-848. [Crossref]
8. Barry RP, Ver-Hoef J. Blackbox kriging: Spatial prediction without specifying variogram models. Journal of Agricultural, Biological, and Environmental Statistics, 1996; 1:297-322. [Crossref]
9. Higdon D. A process convolution approach to modelling temperatures in the North Atlantic. Journal of Environmental Engineering and Science, 1998; 5:173-190. [Crossref]
10. Higdon DM, Swall J, Kern J. Nonstationary spatial modeling. Bayesian Statistics, 1999; 6:761-768.
11. Risser DM, Calder AC. Regression based covariance functions for nonstationary spatial modeling. Environmetrics, 2015; 26:284-297. [Crossref]
12. Fuentes M. Spectral methods for nonstationary spatial processes. Biometrika Trust, 2002; 89:197-210. [Crossref]
13. Paciorek CJ, Schervish MJ. Nonstationary covariance functions for Gaussian process regression. Advances in Neural Information Processing Systems, 2004; 16(16):273-280.
14. Bornn L, Shaddick G, Zidek JV. Modeling nonstationary processes through dimension expansion. Journal of the American Statistical Association, 2012; 107:281-289. [Crossref]
15. Schmidt MA, Guttorp P, O'Hagan A. Considering covariates in the covariance structure of spatial processes. Environmetrics, 2011; 22:487-500. [Crossref]

16. Meiring W, Guttorp P, Sampson P. Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. Environmental and Ecological Statistics, 1998; 5:197-222. [Crossref]

17. Le N, Sun L, Zidek J. Spatial prediction and temporal backcasting for environmental fields having monotone data patterns. The Canadian Journal of Statistics, 2001; 29:529-554. [Crossref]

18. Damian D, Sampson P, Guttorp P. Bayesian estimation of semi-parametric nonstationary spatial covariance structures. Environmetrics, 2001; 12:161-178. [Crossref]

19. Guttorp P, Meiring W, Sampson P. A space-time analysis of ground level ozone data. Environmetrics, 1994; 5(3):241-254. [Crossref]

20. Ingebrigtsen R, Finn L, Ingelin S. Spatial models with explanatory variables in the dependence structure. Spatial Statistics, 2014; 8:20-38. [Crossref]

21. Huang C, Hsing T, Cressie N. Nonparametric estimation of the variogram and its spectrum. Biometrika, 2011; 98:775-789. [Crossref]

22. Genton GM, David JG. Non-parametric variogram and covariogram estimation with fourier-bassel matrices. Journal of Computational Statistics and Data Analysis, 2002; 41:47-57. [Crossref]

23. Im H, Stein ML, Zhu Z. Semi-parametric estimation of spectral density with irregular observations. Journal of the American Statistical Association, 2007; 102:726-735. [Crossref]

24. Cressie N, Johannesson G. Fixed rank kriging for very large spatial data sets. Journal of the Royal Statistical Society, Series B (Statistical Methodology), 2008; 70:209-226. [Crossref]

25. Banerjee S, Gelfand AE, Finley AO, Sang H. Gaussian predictive models for large spatial datasets. Journal of the Royal Statsitical Society, Series B, 2008; 70:825-848. [Crossref] [PubMed] [PMC]

26. Eidsvik J, Finley AO, Banerjee S, Rue H. Approximate bayesian inference for large spatial datasets using predictive process models. Computational Statistics and Data Analysis, 2012; 56:1362-1380. [Crossref]

27. Ren Q, Banerjee S. Hierarchical factor models for large spatially misaligned datasets: a low-rank predictive process approach. Biometrics, 2013; 69:19-30. [Crossref] [PubMed] [PMC]

28. Katzfuss, M. Bayesian nonstationary spatial modeling for very large datasets. Environmetrics, 2013; 24:189-200. [Crossref]

29. Nychka D, Bandyopadhyay S, Hammerling D, Lindgren F, Sain S. A multiresolution gaussian process model for the analysis of large spatial datasets. Journal of Computational and Graphical Statistics, 2015; 24:579-599. [Crossref]

30. Choi I, Li B, Wang X. Nonparametric estimation of spatial and space-time covariance function. Journal of Agricultural, Biological, and Environmental Statistics, 2013; 18(4):1-20. [Crossref]

31. Cressie N. Fitting variogram models by weighted least squares. Mathematical Geology, 1985; 17:563-586. [Crossref]

32. Gilmour A, Cullis B, Welham S, Gogel B, Thompson R. An efficient computing strategy for prediction in mixed linear models. Computational Statistics and Data Analysis, 2004; 44:571-586. [Crossref]

33. Cherry S, Banfield J, Quimby W.F. An evaluation of a nonparametric method of estimating semi-variograms of isotropic spatial process. Journal of Applied Statistics, 1996, 23(4):435-449. [Crossref]

34. Shapiro A, Botha JD. Variogram fitting with a general class of conditionally nonnegative definite functions. Computational Statistics and Data Analysis, 1991; 11:87-96. [Crossref]