Accessing the Appropriateness of a Spatial Regression Using Generalized $(h_1h_2)$-Slepian Random Field

Wayan Somayasa, Jamhir Safani, La Ode Ngkoimani and Muhamad Zamrun F.

Department of Mathematics, Halu Oleo University, 93232 Kendari, Indonesia
Department of Geophysics Engineering, Halu Oleo University, 93232 Kendari, Indonesia
Department of Physics, Halu Oleo University, 93232 Kendari, Indonesia
Department of Geology Engineering, Halu Oleo University, 93232 Kendari, Indonesia

Abstract: In this study we derived asymptotic goodness-of-fit test (model check) for spatial regression where the critical region as well as the p-value of the tests are approximated based on the distribution of a type of the integral functional of the generalized $(h_1h_2)$-Slepian field and the set-indexed Gaussian white noise. Such random fields are obtained as the limit process of the moving and the cumulative sums processes of the sequence of random matrices consist of independent and identically distributed random variables indexed by the points of a design constructed by means of a given continuous probability measure. Although the common approach in model diagnostic for regression is based on the functional of the residuals, in this study a new different idea is proposed by directly investigating the moving and the cumulative sums of the array of the observations. It is shown that these approaches are mathematically tractable and practically more applicable. Simulation study is conducted for investigating the finite sample size behavior of the tests. An application of the procedure to a mining data is also discussed, where from the perspective of geology and geophysics, polynomial model is reasonable and suitable for the data.

Keywords: Slepian Field, Set-Indexed White Noise, Goodness-of-Fit Test, Moving Sum Process, Spatial Regression

Introduction

Modelling spatial data using spatial random field (process) approach has been increasingly studied in various scientific disciplines among them are agriculture, environmental sciences, geostatistics, geology, medicine, biology, mining industry, among others. In the statistical literatures of spatial analysis the real-valued variable observed at the space coordinate is usually regarded as a realization of a stochastic process indexed by a set of points or a family of sets which is commonly called random field, cf. Cressie (1993; Ripley, 2004; Wackernagel, 2003). The measured variable in spatial analysis might stand for percentage of either Ni, Fe or Au in mining exploration, see e.g., Tahir (2010; Somayasa et al., 2015a; 2015b; Somayasa and Wibawa, 2015; Somayasa et al., 2016) or the incidence rates for breast cancer, cf. MacNeill et al. (1994). We refer to (Cressie, 1993) for a comprehensive review and bibliography.

One important purpose of the statistical analysis for spatial data is optimal prediction of an unobserved part of the process. In the references of spatial data analysis this type of statistical inference is called kriging. As stated in the literatures mentioned above, the result of universal kriging depends heavily not only on the covariance structure of the spatial process, but also on the adequateness of the assumed model that describes the mean of the observed variable, (Cressie, 1993; Ripley, 2004; Wackernagel, 2003). This means that preliminary diagnostics involving model validity check as well as model selection must be conducted before kriging to prevent wrong conclusion. A serious spatial analysis should be accompanied with a statistical inference for checking whether or not the assumed model fit to the sample. To this end there has been many approaches and procedures proposed in the literatures how to handle a proper model check. Stute (1997; Stute et al., 2008) for a complete bibliographical information.
It is the aim of the present paper to give a significance contribution in model diagnostic for univariate spatial regression by establishing asymptotic test procedures for checking the appropriateness of an assumed model defined on a closed rectangle under an arbitrary experimental design. By combining the setup of Goodness-of-Fit (GoF) hypothesis for regression defined both in Arnold (1981; Eubank and Hart, 1993), we propose asymptotic tests method by defining a test statistic expressed as a Riemann-Stieltjes integral of a function defined on a competing alternative with respect to the so-called spatial Moving Sums (MOSUM) process of the arrays of observations. We call this test in the sequel as MOSUM test for brevity. This statistic is shown to converges in distribution (weakly) to the integral of such regression function with respect to a generalized $(h_1, h_2)$-Slepian field when the hypothesis is true. The critical region of the MOSUM test is determined based on the probability distribution of such random field. We show in the appendix that the generalized $(h_1, h_2)$-Slepian field is obtained as a limit process of the MOSUM process of the sequence of random matrices consisting independent and identically distributed random variables with finite first and second moments indexed by the design points constructed in arbitrary probability measure. We note that the ordinary $(h_1, h_2)$-Slepian field studied in Fuchang and Li (2007; Bischoff and Gegg, 2014), is a spatial process obtained as a limit process when the design points are constructed using the uniform probability measure.

The application of spatial process such as the Brownian sheet (Brownian (2) motion), Brownian pillow and the set-indexed Bronian sheet in GoF as well as Lack-of-Fit (LoF) for spatial regression has been investigated in many literatures. A common feature of most work is to test the hypothesis by investigating the continuous functional of the Cumulative Sums (CUSUM) of the residuals. The critical region was developed by studying the principal component of the corresponding functional of the Gaussian processes stated above. For example, (Stute, 1997; Stute et al., 2008) pro-posed the Kolmogorov-Smirnov functional of the so-called marked empirical process based of the residuals. MacNeill and Jandhyalla (1993; Xie and MacNeill, 2006) investigated Cramer-von Misses functional of the CUSUM process of the residuals for detecting boundary in spatial regression. Geometric approach have been proposed in the works due to Bischoff (2002; Bischoff and Somayasa, 2009; Somayasa et al., 2015a). However most of these papers have restrictive application, because the problem addressed to the computation of the quantities of the limiting distribution of the test statistic is mainly not tractable as the dimension of the experimental region gets large. We show in this study that our proposed method is more applicable.

To be able to state about the sensitivity of the MOSUM test, we discuss a comparison study by defining a similar test using the integrated regression function under alternative with respect to univariate Gaussian white noise which is a random filed defined e.g., in Alexander and Pyke (1986; Pyke, 1983; Gaensler 1993; Lifshits, 2012). This statistic is actually the limit of that defined as the integral with respect to CUSUM process of the arrays of the observations. For convenience we call such test as CUSUM test. The behavior of the tests will be studied by investigating their empirical as well as limiting power functions by simulation.

The paper is organized as follows. In section 2 we give a more precise definition of the model and the hypotheses under study to fix the idea. The detailed treatment of the MOSUM test and also its asymptotic distribution is presented in section 3. Throughout this work the test procedures are derived under more general condition incorporating the technical situation frequently encountered in mining or geological engineering in which by some practical reason the engineers sometimes cannot or will not determine the drilling bores equidistantly. We propose an experimental design by generalizing the approach introduced in Somayasa (2013) in which we construct the design points over the experimental region by means of a given probability measure defined on the experimental region. Hence our test procedure will be more applicable in practice. However, mathematically the derivation of the result seems to be more difficult. We therefore need more effort. In section 4 we study the CUSUM test. The finite sample behavior of the tests are investigated by simulation in section 5. In Section 6 we discuss the application of the methods to real data. This paper is closed with some conclusions and remarks for future works. Proofs of propositions, theorems and corollaries are postponed to the appendix.

**Model Definition**

In this section we give a brief review to the sampling scheme, the regression model and the hypotheses under study. For more detail interested reader is referred to (Somayasa, 2013; Somayasa et al., 2015a). We consider the experimental design consists of $n_1 \times n_2$ points:

$$\Xi = \{\left(t_{n_1}, s_{n_2}\right) : 1 \leq \ell \leq n_1, 1 \leq k \leq n_2\} \subset D, n_1 \geq 1, n_2 \geq 1$$

defined on a closed rectangle $D := [a_1, a_2] \times [b_1, b_2]$, say, where $a_1 < a_2$ and $b_1 < b_2$. Much of the existing
literature is concerned with the \( n_1 \times n_2 \) regular lattice, given by \( t_i = a_i + \left( \frac{a_i - a_{i-1}}{n_1} \ell \right) \) and \( s_{nk} = b_k + \left( \frac{b_k - b_{k-1}}{n_2} \right) k \), for \( 1 \leq \ell \leq n_1, 1 \leq k \leq n_2 \). (Bischoff and Somayasa, 2009; MacNeill and Jandhyala, 1993; MacNeill et al., 1994; Somayasa et al., 2015a). This kind of experimental design is from practical point of view sometimes not efficient. It is mathematically associated with the uniform probability measure defined by a scaled Lebesgue measure \( \lambda_{\ast} = \frac{1}{|D|} \lambda_{\ast} \) in the sense the corresponding discrete measure converges to \( \lambda_{\ast} \). That is, let \( P_{n_1} \) be a discrete probability measure on the Borel \( \sigma \)-algebra \( B(D) \), defined by:

\[
P_{n_1}(B) := \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{k=1}^{n_2} \mathbb{1}_{\{ t_i, s_{nk} \}}(B) \in B(D)
\]

where, \( \mathbb{1}_{B} \) is the indicator of \( B \). Then for the \( n_1 \times n_2 \) regular lattice, it can be shown by applying the well known Portmanteau theorem (Billingsley (1999), pp. 18-19), that \( P_{n_1} \Rightarrow \lambda_{\ast} \), as \( n_1, n_2 \to \infty \). Here and throughout the paper "\( \Rightarrow \)" stands for the convergence in distribution in the sense of (Billingsley, 1999).

Based on this fact we go through from the opposite direction with the question: If we are given a continuous probability measure \( P \) defined on the Borel \( \sigma \)-algebra \( B(D) \) with the corresponding distribution function \( F \), can we construct a design \( \Xi_{n_1 n_2} \), such that \( P_{n_1} \Rightarrow P \). The answer is "yes". We can determine the design points \( \{ t_{i, 1}, s_{nk} \} \) in natural way by firstly partitioning the interval \([a_1, a_2]\) as \([t_{i, 1}, t_{i, 2}, \ldots, t_{i, n_1}]\) using the equation \( F(t_{i, 1}) = \ell/n_1 \), for \( 1 \leq \ell \leq n_1 \). Next, for a fixed \( \ell \) and \( k \), with \( 1 \leq k \leq n_2 \), the point \( \{ t_{i, 1}, s_{nk} \} \) is generated by solving the equation \( F(t_{i, 1}, s_{nk}) = \ell k/(n_1 n_2) \). The solution is unique provided \( F \) is continuous and strictly increasing on \( D \). Hence, by this sampling method \( \Xi_{n_1 n_2} \) is not necessarily a regular lattice unless \( P \) is the uniform probability measure on \( B(D) \), (Somayasa, 2013). For convenience we call the probability measure \( P \) under which \( \Xi_{n_1 n_2} \) is constructed a design.

As an example, let us consider a probability measure \( P \) defined on the measurable space \( (D := [1, 2] \times [2, 3], B(D)) \), having the probability density function \( u(t, s) := 12/(t s^2) \) and the distribution function \( F(t, s) := 12(1-1/t)(1/2-1/s) \), for \( 1 \leq t \leq 2 \) and \( 2 \leq s \leq 3 \), illustrated in Fig. 1. There exist distribution functions \( F_1(t) := 2(1-1/t) \) and \( F_2(s) := 6(1/2-1/s) \) on \([1, 2]\) and \([2, 3]\), respectively, such that for \((t, s) \in D\), \( F(t, s) = F_1(t)F_2(s) \). Then for fixed \( n_1 \geq 1 \) and \( n_2 \geq 1 \), the design point is computed by the formula:

\[
t_{i, 1} = \frac{2n_1}{2n_1 - \ell} \quad \text{and} \quad s_{nk} = \frac{6n_2}{3n_2 - k}, 1 \leq \ell \leq n_1, 1 \leq k \leq n_2
\]

We notice that \( P_{n_1 n_2} \) can also be written as:

\[
P_{n_1 n_2}(B) := \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{k=1}^{n_2} \mathbb{1}_{B}(F_1^{-1}(t_{i, 1}/n_1), F_2^{-1}(k/n_2)), B \in B(D)
\]

Fig. 1. Left side: The three dimensional perspective plot of the density function of \( P \). Right side: The cumulative distribution functions of \( P \) on the compact rectangle \( D = [1, 2] \times [2, 3] \).
Let a nonparametric regression model \( Y(x) = g(x) + \varepsilon(x) \), for \( x \in D \) be observed on \( \Xi_{n \times n} \), where \( g \) is an unknown function of bounded variation on \( D \) and \( \varepsilon \) is the unobserved random error defined on a common probability measure \((\Omega,F,\mathbb{P})\), say, with \( E(\varepsilon(x)) = 0 \) and \( \text{Var}(\varepsilon(x)) = \sigma^2 > 0 \) for every \( x \in D \). Then the matrix of independent observations \( Y_{n \times n} = \left( Y \left( t_{i_1}, s_{i_2} \right) \right)_{i_1,i_2=1}^{n_1,n_2} \) satisfies the following decomposition:

\[
Y_{n \times n} = g_{n \times n} + \varepsilon_{n \times n}
\]

where,

\[
g_{n \times n} := \left( g \left( t_{i_1}, s_{i_2} \right) \right)_{i_1,i_2=1}^{n_1,n_2}
\]

\[
\varepsilon_{n \times n} := \left( \varepsilon \left( t_{i_1}, s_{i_2} \right) \right)_{i_1,i_2=1}^{n_1,n_2}
\]

is an \( n_2 \times n_1 \) matrix of random errors having independent and identically distributed components with \( E(\varepsilon(t_{i_1},s_{i_2})) = 0 \) and \( \text{Var}(\varepsilon(t_{i_1},s_{i_2})) = \sigma^2 > 0 \), for \( 1 \leq i_1 \leq n_1; 1 \leq i_2 \leq n_2 \). For the sake of brevity we write \( Y(t_{i_1},s_{i_2}) \), \( g(t_{i_1},s_{i_2}) \) and \( \varepsilon(t_{i_1},s_{i_2}) \) throughout this paper as \( Y_{t_i s_j}, g_{t_i s_j} \) and \( \varepsilon_{t_i s_j} \) respectively. It is important to note that for our result we do not need normal assumption.

As nicely noted in (Arnold, 1981; Eubank and Hart, 1993; Stute, 1997) a common feature of the GoF test for regression falls into the following framework. Let \( V := \left[ z_0, \ldots, z_p \right] \) and \( W := [z_0, \ldots, z_p] \), \( p \leq m \), be linear subspaces of \( L_2(D,P) \), where \( z_1, \ldots, z_p, \ldots, z_m \) are known regression functions which are without loss of the generality assumed to be orthogonal as functions in \( L_2(D,P) \). The space \( L_2(D,P) \) is the set of squared integrable functions on \( D \) with respect to \( P \) which is furnished with the inner product and norm denoted respectively by \( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \). We test the null hypothesis that \( g \in W \) while observing \( g \in F \). Since \( g \) is observed as a function that lies in \( V \) and on the other hand \( W \subset V \), then \( g \) admits an orthogonal decomposition \( g = g_1 \oplus g_2 \), where \( g_1 \in W \) and \( g_2 \in V \setminus W \), with \( \langle g_1, g_2 \rangle = 0 \). Hence the problem of testing \( H_0: g \in W \) while observing \( g \in V \) can be handled by testing the hypotheses:

\[
H_0 : g_2 = 0 \quad \text{against} \quad H_1 : g_2 \neq 0 \quad \text{for some} \quad f_i \in V \cap W^\perp
\]

We notice that the statement that \( g \in W \) is equivalent to that of \( \frac{1}{\sqrt{n_1 n_2}} g \in W \) for all \( n_1 \geq 1 \) and \( n_2 \geq 1 \). Hence the problem of testing \( H_0: g \in W \) can be handled by testing that of \( H_0 : \frac{1}{\sqrt{n_1 n_2}} g \in W \) for all \( n_1 \geq 1 \) and \( n_2 \geq 1 \). On the other hand by observing the localized model, the limiting distribution of the sequence of the moving as well as cumulative sums process of the observations under any competing alternative can be obtained concretely, see section 3 and section 4 below. Therefore without altering the test problem we consider in this study the localized model:

\[
Y^*_{n \times n} = \frac{1}{\sqrt{n_1 n_2}} g_{n \times n} + \varepsilon_{n \times n}
\]

In this study we restrict the consideration to \( x \in \mathbb{R}^2 \). Result for higher dimensional rectangle can be obtained immediately.

**Test based on Moving Sums of the Observations**

In this section we introduce a test based on spatial Moving Sums (MOSUM) process of the observations extending the notion of one dimensional MOSUM test defined in Chu et al. (1995). Let \( h_1 \) and \( h_2 \) be positive numbers such that \( 0 \leq h_1 \leq a_1 \) and \( 0 \leq h_2 \leq a_2 \). The moving sums process of the matrix of observations \( Y_{n \times n} \) consisting of:

\[
\left( \left\{ n_i F_1 (t + h) \right\} - \left\{ n_i F_1 (t) \right\} \right) \left( \left\{ n_i F_2 (s + h) \right\} - \left\{ n_i F_2 (s) \right\} \right)
\]

components, is denoted by \( MSh_h^* \left( Y_{n \times n} \right) (t,s) \), defined as:

\[
MSh_h^* \left( Y_{n \times n} \right) (t,s) := \frac{1}{\sqrt{n_1 n_2}} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} Y_{t+k,s+l}
\]

for \( (t,s) \in D_{h_1 h_2} := \left[ a_1, a_2 - h_1 \right] \times \left[ b_1, b_2 - h_2 \right] \), where \( \| x \| = \max \{ z \in \mathbb{Z} : z \leq x \} \). The MOSUM test for the hypothesis \( H_0^* : g_2 = 0 \) against \( H_1^* : g_2 \neq 0 \), for some \( f_i \in V \cap W^\perp \), is realized by using the statistics:

\[
M \left( Y^*_{n \times n} \right) = \frac{1}{\sigma^2} \int_{D_{h_1 h_2}} f_i (t,s) dMSh_h^* \left( Y^*_{n \times n} \right) (t,s)
\]

rejecting \( H_0^* \) for large value of \( M \left( Y^*_{n \times n} \right) \), where \( \sigma^2 \) a consistent estimator of \( \sigma \) given e.g., in Arnold (1981), pp. 148-149. We note that for every \( (t,s) \in D_{h_1 h_2} \), \( MSh_h^* \left( Y_{n \times n} \right) (t,s) \) can be expressed in the following manner which is useful for the derivation of the main result. Indeed:

\[
MSh_h^* \left( Y_{n \times n} \right) (t,s) := \frac{1}{\sqrt{n_1 n_2}} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} Y_{t+k,s+l} \left( n_i l, s_{i_2} \right) Y^*_{t,s}
\]

for \( (t,s) \in D_{h_1 h_2} \), meaning that for fixed \( h_1 \) and \( h_2 \), \( MSh_h^* \left( Y_{n \times n} \right) \) can be viewed as a partial sums process.
indexed by \( \{[t, t + h_1] \times [s, s + h_2]: (t, s) \in D_{a,b} \} \) of rectangles in \( D \), cf. (Bischoff and Somayasa, 2009; Somayasa et al., 2015a). This end can be verified based on the assumption that \( F(t, s) = F(t, s)F(t, s) \) and the continuity of \( F_1 \) and \( F_2 \) on \( [a_1, a_2] \) and \( [b_1, b_2] \), respectively. Indeed we have:

\[
F_1 \left( t_n \big| n, (o|i) \right) = \left| n, F_1 (t) \right| + \frac{1}{n} \geq \frac{n, F_1 (t) \geq 0}{n} = F_1 (t)
\]

implying \( t \leq t_n \big| n, (o|i) \big| + \) by the monotonicity of \( F_1 \). Further, since:

\[
F_1 \left( t_n \big| n, (i|o) \right) = \left| n, F_1 (t + h) \right| \leq \frac{F_1 (t + h)}{n}
\]

we get \( t_n \big| n, (o|i) \big| \leq t + h \). By the similar argument we can also show that \( s \leq s_n \big| n, (o|i) \big| \) and \( s_n \big| n, (o|i) \big| \leq s + h \).

Hence for fixed \((t, s) \in D_{a,b} \), the sum in \( MS_{a,b} \left( Y_{s,n}^* \right) \) involves all \( Y_{k+l} \) with the associated design point \( \{t_n, s_n\} \) lies in the rectangle \([t, t] \times [s, s] \supseteq [t, t + h] \times [s, s + h] \).

If the uniform probability measure \( \lambda_2^* \) on \( D \) is considered as the design, we have the following formula for every \((t, s) \in D_{a,b} \):

\[
MS_{h_2} \left( Y_{s,n}^* \right) (t, s) = \frac{1}{\sqrt{n_1 n_2}} \sum_{t_{n,1}} \sum_{s_{n,2}} Y_{k+l}
\]

\[
= \frac{1}{\sqrt{n_1 n_2}} \sum_{t_{n,1}} \sum_{s_{n,2}} Y_{k+l}
\]

which consists of \( \left( \begin{array}{c} n_1 h_1 \\ a_1 - a_i \\ b_1 - b_i \end{array} \right) \) components of the matrix \( Y_{s,n}^* \) of the observations. Hence, by following the terminology introduced in (Bischoff and Gegg, 2014; Chu et al., 1995) we call the term \( h_i h_j \) in this study the window size of the process.

By the definition of \( MI(Y_{s,n}^*) \) it can be seen that when the sample supports \( H_2 \), the value of \( MI(Y_{s,n}^*) \) will be large as it is contributed by the quantity that asymptotically rational to a positive constant given by \( \int_0^{h_2} f_i(t) d\zeta_2(t) \), see Theorem 3.1 and Remark 3.2 below. Conversely if the sample comes from \( H_0 \), the value of \( MI(Y_{s,n}^*) \) will small as \( \int_0^{h_2} f_i(t) d\zeta_2(t) \) vanishes uniformly. Hence, rejecting \( H_0 \) for large \( MI(Y_{s,n}^*) \) leads us to a conclusion that MOSUM test so defined is a reasonable test.

The next result presents asymptotic size \( \alpha \) test for testing (2) based on the statistic \( MI(Y_{s,n}^*) \). Let \( S_{a,b} \) be the generalized \((h_1,h_2)\)-Slepian filed with the parameter space \( D_{a,b} \), see the appendix. It is not too difficult to prove the following result.

**Theorem 3.1**

Let the matrix of the observations \( Y_{s,n}^* \) satisfies Model (1) such that the unknown regression function \( g \in L_2(D, P) \) have an orthogonal decomposition \( g = g_1 \otimes g_2 \) with \( g_1 \in W \) and \( g_2 \in V \cap W^\perp \). Suppose that \( g_1 \) and \( g_2 \) are continuous and have bounded variation on \( D \). Then an asymptotic size \( \alpha \) test for \( H_0: g_2 = 0 \) against \( H_1: g_2 = f_1 \), for some continuous functions \( f_1 \in V \cap W^\perp \), will reject \( H_0 \) if and only if \( MI(Y_{s,n}^*) \geq m_{1,\alpha} \), where \( m_{1,\alpha} \) is a constant determined from the distribution of the statistic:

\[
I(S_{a,b}) := \int_0^{\infty} f_1 d\zeta_2 \int_0^{\infty} f_1 d\zeta_2 + \int_0^{\infty} f_1 dS_{a,b}
\]

by the equation \( \Pr (S_{a,b} \geq m_{1,\alpha} | H_0) = \alpha \). Thereby when \( H_0 \) is true, the limit of \( MI(Y_{s,n}^*) \) is given by:

\[
\int_0^{\infty} f_1 d\zeta_2 + \int_0^{\infty} f_1 dS_{a,b}, \text{ where for } i = 1, 2, \zeta_i \text{ is a function on } D_{a,b} \text{, defined as:}
\]

\[
\zeta_i(t, s) := \int f_1 d\zeta_2 + \int f_1 dS_{a,b}
\]

Furthermore, the finite sample size power function of this test is given by:

\[
G_{I} (f, \alpha) := \Pr \left( MI(Y_{s,n}^*) \geq m_{1,\alpha} \big| g_2 = f \right), f \in V \cap W^\perp
\]

which converges point-wise to \( G_1 \), defined by:

\[
G_1 (f, \alpha) := \Pr \left( \int f_1 dS_{a,b} \geq m_{1,\alpha} - \frac{1}{\alpha} \int f_1 d\zeta_2, f \in V \cap W^\perp
\]
We note that all integrals involved in the definition of \( I(S, P) \) are in the sense of Riemann-Stieltjes type defined e.g., in Strook (1994).

**Remark 3.2**

Unfortunately, the probability distribution model of \( I(S, P) \) is not tractable for arbitrary probability measure \( P \), by the reason the increments of \( S, P \) is not stochastically independent and the limit \( I(S, P) \) still depends on the choice of \( g_1 \), unless \( g_1 \) and \( f_1 \) are orthogonal on \( D_{b_h} \).

Therefore, the test is implemented in practice by approximating \( m_{1, a} \) by generating Monte Carlo simulation. However, when the design is constructed under the \( \lambda_{a}^{2} \) (regular lattice), it can be shown that \( I(S, P) \) has independent increment, see Proposition A.6.

As pointed out in Proposition A.7, it can be shown that:

\[
\int_{D_{b_h}} f_i \, dS_{E} \sim N\left(0, 4 \int_{D_{b_h}} f_i^2(t, s) \lambda_{a}^{2}(dt, ds)\right)
\]

Hence, under \( H_0 \), it holds:

\[
I(S, P) \sim N \left( \int_{D_{b_h}} f_i \, dS_{E}, 4 \int_{D_{b_h}} f_i^2(t, s) \lambda_{a}^{2}(dt, ds) \right)
\]

Table 1. The simulated quantiles for constant, first and second order model under \( H_0 \) generated with 60×70 design points constructed using \( F(t, s) = 12(1-1=t)(1-1=s) \) on \( D = [1, 2] \times [2, 3] \)

| Models | 0.990 | 0.975 | 0.950 | 0.900 | 0.850 | 0.800 |
|--------|--------|--------|--------|--------|--------|--------|
| Const. | 0.05   | 0.05   | 0.4096 | 0.3440 | 0.2849 | 0.2223 |
|        | 0.05   | 0.10   | 0.3961 | 0.3317 | 0.2780 | 0.2163 |
|        | 0.05   | 0.15   | 0.3995 | 0.3285 | 0.2741 | 0.2136 |
|        | 0.05   | 0.25   | 0.3859 | 0.3288 | 0.2780 | 0.2136 |
|        | 0.15   | 0.10   | 0.7076 | 0.6111 | 0.5111 | 0.3944 |
|        | 0.15   | 0.15   | 0.6994 | 0.5939 | 0.5005 | 0.3876 |
|        | 0.15   | 0.25   | 0.7053 | 0.5878 | 0.4933 | 0.3826 |
|        | 0.20   | 0.10   | 0.7914 | 0.6729 | 0.5569 | 0.4433 |
|        | 0.20   | 0.15   | 0.8028 | 0.6699 | 0.5570 | 0.4320 |
|        | 0.20   | 0.25   | 0.7606 | 0.6456 | 0.5391 | 0.4230 |
|        | 0.25   | 0.15   | 0.8627 | 0.7092 | 0.6046 | 0.4731 |
|        | 0.25   | 0.20   | 0.8469 | 0.7219 | 0.6014 | 0.4718 |
|        | 0.25   | 0.25   | 0.8430 | 0.7218 | 0.6022 | 0.4620 |
| 1st order | 0.05   | 0.05   | 1.3817 | 1.1644 | 0.9910 | 0.7621 |
|         | 0.05   | 0.10   | 1.3281 | 1.1220 | 0.9528 | 0.7399 |
|         | 0.05   | 0.15   | 1.3093 | 1.0707 | 0.8769 | 0.6862 |
|         | 0.05   | 0.25   | 1.2394 | 1.0434 | 0.8706 | 0.6861 |
|         | 0.15   | 0.10   | 2.3136 | 1.9527 | 1.6371 | 1.2710 |
|         | 0.15   | 0.15   | 2.2480 | 1.8734 | 1.5791 | 1.2190 |
|         | 0.15   | 0.25   | 2.2690 | 1.8688 | 1.5619 | 1.2087 |
|         | 0.20   | 0.10   | 2.6455 | 2.2453 | 1.8779 | 1.4240 |
|         | 0.20   | 0.15   | 2.5909 | 2.1499 | 1.8101 | 1.4051 |
|         | 0.20   | 0.25   | 2.4980 | 2.1154 | 1.7507 | 1.3345 |
|         | 0.25   | 0.15   | 2.7584 | 2.3539 | 1.9714 | 1.5372 |
|         | 0.25   | 0.20   | 2.6930 | 2.2771 | 1.9429 | 1.5069 |
|         | 0.25   | 0.25   | 2.7318 | 2.3243 | 1.9429 | 1.4797 |
| 2nd order | 0.05   | 0.05   | 1.5798 | 1.2735 | 1.0398 | 0.7529 |
|          | 0.05   | 0.10   | 1.5010 | 1.2493 | 1.0375 | 0.7546 |
|          | 0.05   | 0.15   | 1.4574 | 1.1982 | 0.9903 | 0.7459 |
|          | 0.05   | 0.25   | 1.4599 | 1.2193 | 1.0052 | 0.7507 |
|          | 0.15   | 0.10   | 2.4139 | 1.9833 | 1.5187 | 1.0345 |
|          | 0.15   | 0.15   | 2.3644 | 1.8953 | 1.4916 | 1.0549 |
|          | 0.15   | 0.25   | 2.3958 | 1.9212 | 1.5209 | 1.0545 |
|          | 0.20   | 0.10   | 2.7522 | 2.1992 | 1.6782 | 1.0582 |
|          | 0.20   | 0.15   | 2.5483 | 2.0170 | 1.5789 | 1.0728 |
|          | 0.20   | 0.25   | 2.6376 | 2.0509 | 1.5926 | 1.0819 |
|          | 0.25   | 0.15   | 2.8682 | 2.2099 | 1.6637 | 1.0857 |
|          | 0.25   | 0.20   | 2.8535 | 2.2716 | 1.7374 | 1.1395 |
|          | 0.25   | 0.25   | 2.8014 | 2.2198 | 1.7144 | 1.1171 |
Thus, it can be stated that an asymptotically size $\alpha$ test will reject $H_0$ if:

$$MI\left( Y_{n,n} \right) \geq \Phi^{-1}(1-\alpha) \sqrt{\int_{D_{h,h}} f^2(t,s) \lambda_0^2(dt,ds)}$$

provided $f_1$ and $g_1$ are also orthogonal on $D_{h,h}$ as functions in $L_2\left(D_{h,h}, \lambda_0^2\right)$. Under the same condition as before we further get the corresponding power function of the test as given by:

$$G_{\lambda_0^2}(f) := P\left( MI\left( Y_{n,n} \right) \geq \Phi^{-1}(1-\alpha) \sqrt{\int_{D_{h,h}} f^2(t,s) \lambda_0^2(dt,ds)} | g_2 = f \right)$$

which converges point-wise to the following power function:

$$G_{\lambda_0^2}(f) := 1 - \Phi - \frac{1}{\sigma^2} \int_{D_{h,h}} f^2(t,s) \lambda_0^2(dt,ds)$$

Although the $m_{n,a}$ can not be calculated analytically, the application of the MOSUM test in the practice can be realized at least by simulation by approximating the $p$-value of the test. For arbitrary design $P$, let $m^*$ be the observed value of $MI\left( Y_{n,n} \right)$, then the $p$-value can be obtained by approximating the following probability:

$$p\text{-value} = P\left( MI\left( Y_{n,n} \right) \geq m^* | H_0 \right)$$

For the design $\lambda_0^2$, we have:

$$p\text{-value} = P\left( MI\left( Y_{n,n} \right) \geq m^* | H_0 \right) \approx 1 - \Phi \left( \frac{m^*}{\sqrt{\int_{D_{h,h}} f^2(t,s) \lambda_0^2(dt,ds)}} \right)$$

provided $f_1$ and $g_1$ are orthogonal on $D_{h,h}$. Hence, the realization of the MOSUM test is according to the following algorithm

1. Construct the design of experiment containing $n_1 \times n_2$ points using a given probability measure $P$.
2. Define the hypothesis of interest.
3. Compute the critical value of $MI\left( Y_{n,n} \right)$.
4. Calculate the associated $p$-value.
5. Draw decision: reject $H_0$ when $\alpha \geq p$-value.

Comparison to Set-Indexed CUSUM Method

In this section we aim to establish a different approach for testing Hypothesis 2 in that instead of considering the moving sums we define other reasonable test statistic $CU\left( Y_{n,n} \right)$, defined by:

$$CU\left( Y_{n,n} \right) := \int_{D} f_1(t,s) d\frac{1}{\sigma^2} S_{n|h}(Y_{n,n})(t,s)$$

where:

$$S_{n|h}(Y_{n,n})(t,s) := \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left( \frac{1}{\sqrt{n_1 n_2}} \left( Y_{i,n} - \mu \right) \right)(t,s)$$

is the cumulative sums process of the observations indexed by $D$. This type of stochastic process is a special case of the more general one defined by:

$$S_{n|h}(Y_{n,n})(B) := \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left( \frac{1}{\sqrt{n_1 n_2}} \left( Y_{i,n} - \mu \right) \right)(t,s)$$

which is commonly called set-indexed partial sums process, (Alexander and Pyke, 1986; Gaenssler, 1993; Pyke, 1983; Xie and MacNeill, 2006). For each $\omega \in \Omega, S_{n|h}(Y_{n,n})(\omega)$ constitutes a signed measure on $A$, therefore the integral involved in the statistic $CU\left( Y_{n,n} \right)$ can be interpreted path-wise as the integral of a function in $L_2(D, P)$ with respect to a signed measure. We refer the reader to Cohn (1980), pp. 121-153 for the definition of integral with respect to signed measure.

Based on the definition of both MOSUM and CUSUM process it can be seen that $S_{n|h}$ coincides with $MS_{h,h}$ when $S_{n|h}$ is restricted to the Vapnik-Chervonenkis class (VCC) $\left\{ \{t_1, t_2\} \times \{s_1, s_2\} : a_1 < t_1, t_2 < a_2, b_1 < s_1, s_2 < b_2 \} \}$ of subsets of $D$. By this reason the CUSUM test is viewed as a generalization of the MOSUM test. Clearly the MOSUM test differs from the CUSUM test in that each moving sum contains a fixed number of the observations, whereas the cumulative sums test incorporates more and more observations. Therefore we can conjecture that MOSUM test should be more sensitive than that of CUSUM test in detecting the change in model, see also (Chu et al., 1995). It is the purpose of this work to investigate this sensitivity property by comparing the behavior of the finite sample power functions of both tests.
The following theorem gives the asymptotic size \( \alpha \) test for testing (2). The proof is devoted to the appendix.

**Theorem 4.1**

Let \( g \) have an orthogonal decomposition \( g = g_1 \oplus g_2 \) with \( g_1 \in W \) and \( g_2 \in V \cap W^\perp \) as functions in \( L_2(D, P) \), suppose that \( g_1 \) and \( g_2 \) are continuous and have bounded variation on \( D \). Then an asymptotic size \( \alpha \) test for the hypothesis \( H_0: g_2 = 0 \) against \( H_1: g_2 = f_1 \), for some continuous functions \( f_1 \in V \cap W^\perp \) will reject \( H_0 \) if and only if

\[
\text{CU} \{ Y_{n_{\alpha/n}} \} \geq \Phi^{-1}(1-\alpha) \parallel f_1 \parallel_p.
\]

An immediate consequence of Theorem 4.1 is the asymptotic power function of the test as presented in the following corollary.

**Corollary 4.2**

Let \( \Psi_{\alpha/n} : V \cap W^\perp \rightarrow (0, 1) \) defined by:

\[
\Psi_{\alpha/n}(f) := P \left\{ \text{CU} \{ Y_{n_{\alpha/n}} \} \geq \Phi^{-1}(1-\alpha) \parallel f_1 \parallel_p | g_2 = f \right\},
\]

be the power function of the test of size \( \alpha \) based on the statistic \( \text{CU} \{ Y_{n_{\alpha/n}} \} \) derived above. Then under the assumption of Theorem 4.1 it holds:

\[
\lim_{n_{\alpha/n} \to \infty} \Psi_{\alpha/n}(f) = 1 - \Phi \left( \Phi^{-1}(1-\alpha) - \frac{f_1 \cdot f_2}{\parallel f_1 \parallel_p} \right) = \Psi_{\alpha}(f)
\]

where, \( \Phi \) is the cumulative distribution function of the standard normal distribution.

In contrast to MOSUM test, CUSUM test can be realized in the practice in relatively easier way by the reason the quantities addressed to the limiting distribution of \( \text{CU} \{ Y_{n_{\alpha/n}} \} \) for arbitrary design \( P \) can be computed analytically. In particular the \( p\)-value of the test can be approximated by the formula:

\[
p\text{-value} = P \left\{ \frac{1}{\sigma_n} \text{CU} \{ Y_{n_{\alpha/n}} \} \geq r^* \big| H_0 \right\} = 1 - \Phi \left( \frac{r^*}{\parallel f_1 \parallel_p} \right)
\]

where, \( r^* \) is the observed value of \( \text{CU} \{ Y_{n_{\alpha/n}} \} \). The decision making process is mainly based on the \( p\)-value instead of computing the value of \( \Phi^{-1}(1-\alpha) \parallel f_1 \parallel_p \).

**Simulation Study**

In this section we present simulation study to approximate the quantiles of the statistic \( \text{MI} \{ Y_{n_{\alpha/n}} \} \) and to demonstrate the finite sample behavior of the test based on \( \text{MI} \{ Y_{n_{\alpha/n}} \} \) as well as \( \text{CU} \{ Y_{n_{\alpha/n}} \} \). We visualize the simulation result by scattering the graphs of the power functions of both tests.

**Simulating the Quantiles of \( \text{MI} \{ Y_{n_{\alpha/n}} \} \)**

Not like the CUSUM test, the asymptotic critical value \( m_{1,\alpha} \) of the MOSUM test for arbitrary design \( P \) and window size \( h_1, h_2 \) can not be determined analytically as \( \alpha \) varies in the interval \((0, 1)\). In this study we approximate the \( m_{1,\alpha} \) for \( \alpha = 0.01, 0.025, 0.05, 0.10, 0.15 \) and 0.20 by Monte Carlo simulation developed according to the assumed model under \( H_0 \): constant first and second order polynomial model defined on a closed rectangle. For each model we use standard normal distribution with 60\( \times \)70 observations to approximate the generalized \( (h_1, h_2) \)-Slepian field \( S_{\alpha/\beta} \). Table 1 presents the simulation results for the experimental design is constructed using the distribution function \( F(t, s) = 12(1-1/t)(1/2-1/s) \) on \( D = [1, 2] \times [2, 3] \) defined in Section 1.

**Constant Model**

In the first case we assume under \( H_0 \) a constant model defined as \( Y(t, s) = \beta_0 z_0(t, s) + \alpha(t, s) \) while we are observing a first order model given by \( Y(t, s) = \beta_0 z_0(t, s) + \beta_1 z_1(t, s) + \beta_2 z_2(t, s) + \alpha(t, s) \), where \( z_0(t, s) = 1, z_1(t, s) = t, z_2(t, s) = s \), for \( (t, s) \in D \), \( \beta_0, \beta_1 \) and \( \beta_2 \) are unknown parameters. By a re-parametrization, the model can be represented as:

\[
Y(t, s) = \gamma_0 z_0(t, s) + \gamma_1 z_1(t, s) + \gamma_2 z_2(t, s) + \epsilon(t, s)
\]

where, \( z_0(t, s) := 1, z_1(t, s) := t-2 \ln (2) \) and \( z_2(t, s) = s-6 \ln (3/2) \) are orthogonal in \( L_2(D, P) \) and \( \gamma_0, \gamma_1, \gamma_2 \) are unknown parameters. The observations are generated from a constant model defined as:

\[
Y(t_{\delta/\ell}, s_{\delta/k}) = \tilde{z}_0(t_{\delta/\ell}, s_{\delta/k}) + \epsilon(t_{\delta/\ell}, s_{\delta/k}), 1 \leq \ell \leq 60, 1 \leq k \leq 70
\]

In this simulation \( g_2 \) is hypothesized as zero under \( H_0 \) and \( g_2 = f_1 + \tilde{z}_2 \) under \( H_1 \). Hence, the critical region is constructed using the test statistic:

\[
\text{MI} \{ Y_{n_{\alpha/n}} \} = \int_{\gamma_{\delta/\ell}} \left( z_1 + \tilde{z}_2 \right)(t, s) d \frac{1}{\sigma_n} \text{MS}_{\alpha/\beta} \{ Y_{n_{\alpha/n}} \}(t, s)
\]

whose simulated \((1-\alpha)\)-quantiles is presented in Table 2.

**First Order Model**

In the second case we give approximation to the \((1-\alpha)\)-quantiles of the statistic \( \text{MI} \{ Y_{n_{\alpha/n}} \} \) for testing first order model, that is we test:
Second Order Model

For the last case we simulate the (1-α)-quantiles of the statistic $M(\tilde{Y}_{\gamma_{n\cdot}})$ for testing the hypothesis $H_0$ that states a second order model is adequate versus $H_1$ that assumes a third order model:

$$Y(t,s) = \gamma_0 \tilde{z}_0(t,s) + \gamma_1 \tilde{z}_1(t,s) + \gamma_2 \tilde{z}_2(t,s) + \gamma_3 \tilde{z}_3(t,s) + \epsilon(t,s)$$

where, $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ are unknown constants with:

$$\tilde{z}_0(t,s) = 1, \tilde{z}_1(t,s) = t, \tilde{z}_2(t,s) = s, \tilde{z}_3(t,s) = t^2$$

Second Order Model

For the last case we simulate the (1-α)-quantiles of the statistic $M(\tilde{Y}_{\gamma_{n\cdot}})$ for testing the hypothesis $H_0$ that states a second order model is adequate versus $H_1$ that assumes a third order model:

$$Y(t,s) = \beta_0 z_0(t,s) + \beta_1 z_1(t,s) + \beta_2 z_2(t,s) + \beta_3 z_3(t,s) + \epsilon(t,s)$$

while we are observing a second order model:

$$Y(t,s) = \beta_0 z_0(t,s) + \beta_1 z_1(t,s) + \beta_2 z_2(t,s) + \beta_3 z_3(t,s) + \epsilon(t,s)$$

and generating new parameter system as:

$$Y(t,s) = \gamma_0 \tilde{z}_0(t,s) + \gamma_1 \tilde{z}_1(t,s) + \gamma_2 \tilde{z}_2(t,s) + \gamma_3 \tilde{z}_3(t,s) + \epsilon(t,s)$$

where $z_0(t,s) = 1, z_1(t,s) = t, z_2(t,s) = s, z_3(t,s) = t_2$ and $z_4(t,s) = s_2$ for $(t,s) \in D$. By the similar Gram-Schmidt orthogonalization procedure, the model can be written by introducing new parameter system as:

$$Y(t,s) = \gamma_0 \tilde{z}_0(t,s) + \gamma_1 \tilde{z}_1(t,s) + \gamma_2 \tilde{z}_2(t,s) + \gamma_3 \tilde{z}_3(t,s) + \epsilon(t,s)$$

for $1 \leq \ell \leq 60, 1 \leq k \leq 70$, we present the quantiles of $M(\tilde{Y}_{\gamma_{n\cdot}})$ under $H_0$ in Table 2 for several combinations of $h_1$ and $h_2$.

Simulating the Power

We investigate by simulation the behavior of the empirical power function of the CUSUM test for two different cases: Constant and first order model described in Subsection 5.1 under two different designs: The probability measure $P$ and the uniform measure $\mu$. Our purpose is to demonstrate that the finite sample power function $\Psi_{P_{n\cdot}}(f)$ lies closed enough to its limit $\Psi_{\mu}(f)$ as $f$ varies in $\mathcal{V} \cap \mathcal{W}^d$ when the sample size $n_1$ and $n_2$ get large. The simulation is based on 10000 runs developed using R.

Constant Model

By considering the hypothesis formulated in Subsection 5.1.1, we have $\|f\|_P = 0.37045$ for $f_1 = \tilde{z}_1 + \tilde{z}_2$, so that $\Phi^{-1}(1-\alpha)\|f\|_P = 0.60934$, when $\alpha = 0.05$. In order to make $f$ varies in $\mathcal{V} \cap \mathcal{W}^d$, we define $f = \lambda f_1$, for $\lambda \in \mathbb{R}$ and generate the samples independently from the normal distribution with mean $\frac{1}{\sqrt{n_1 n_2}} g_1(\Xi_{n_{1\cdot}}) + \lambda f_1(\Xi_{n_{1\cdot}})$ and the variance $\sigma^2$ which is assumed to be unknown. In this case we estimate $\sigma^2$ by using the consistent estimator defined in (Arnold, 1981), pp. 148-149. It is clear that the samples support the proposed model under $H_0$ if and only if $\lambda = 0$, otherwise the alternative $H_1$: $g \in \mathcal{V} = [\tilde{z}_1, \tilde{z}_2]$ holds true. In this simulation we develop the graphs of $\Psi_{P_{n\cdot}}(\lambda f_1)$ for $\alpha = 0.05$, where:

$$\Psi_{P_{n\cdot}}(\lambda f_1) = P \left( \frac{1}{\sigma_n} \int_{[0,1]} f_1 dS Y_{\gamma_{n\cdot}} \geq 0.60934 \right)$$

together with its point-wise limit $\Psi_{\mu}(\lambda f_1) = 1 - \Phi(1.64485-0.37045 \lambda)$, where we chose the VCC $[1, t] \times [2, s]; (t, s) \in D$ instead of the much larger family $A_d$ as the index sets. The simulation results are exhibited in
Fig. 2 for \( \lambda \) varies in the closed interval \([0, 15]\) presented in two panels (a) and (b) associated with the sample sizes 60\( \times \)65 and 70\( \times \)75, respectively. It can be seen that for constant model the curves of \( \Psi_{\rho,n_{n}} \) approximates well those of the limit \( \Psi_{\rho} \) achieving the size \( \alpha = 0.05 \) at \( \lambda = 0 \) as they should be.

Fig. 2. The graphs of the empirical power functions of the asymptotic size \( \alpha = 0.05 \) CUSUM test for constant model represented by using dotted line approximated by the limit power function \( (\Psi_{\rho}(\lambda f_{1})) \) scattered using smooth lines. The design points are generated using the CDF \( \tilde{F}(t, s) = 12(1-1/t)(1/2-1/s) \) on \( D \).

Fig. 3. The graphs of the empirical power function of the asymptotic size \( \alpha = 0.05 \) CUSUM test for constant model under the design generated using the scaled Lebesgue measure \( \lambda_{D}(\text{regular lattice}) \) scattered by using dotted lines approximated by the limit power function \( \Psi_{\rho_{n}}(\lambda f_{1}) \) represented by smooth lines.
Table 2. The critical and $p$-values of MOSUM and CUSUM test for Fe data with a comparison to KS and CvM tests based on the residuals proposed in (Somayasa et al., 2015a)

| Model H0       | Test   | Critical value | $p$-value | $\hat{\sigma}_o$ |
|---------------|--------|----------------|-----------|------------------|
| Constant      | MOSUM  | 3.56799        | 6.2157e-6 | 0.20780          |
|               | $h_1=0.20$; $h_2=0.25$ | 3.10043 | 0.00007 |            |
|               | $h_1=0.25$; $h_2=0.25$ | 1.71874 | 1.2766e-5 |            |
|                | CUSUM  | 1.66759        | 0.03210   |            |
|                | KS     | 0.48443        | 0.01240   |            |
| First order   | MOSUM  | 1.65455        | 0.01668   | 0.20163       |
|               | $h_1=0.20$; $h_2=0.25$ | 1.46001 | 0.03022 |            |
|                | $h_1=0.25$; $h_2=0.25$ | 1.06940 | 0.00297 |            |
|                | CUSUM  | 0.89523        | 0.42660   |            |
|                | KS     | 0.09341        | 0.20460   |            |
| Second order  | MOSUM  | 1.94889        | 0.16492   | 0.19759       |
|               | $h_1=0.20$; $h_2=0.25$ | 1.28138 | 0.26086 |            |
|                | $h_1=0.25$; $h_2=0.25$ | 2.60534 | 0.00459 |            |
|                | CUSUM  | 0.92080        | 0.10380   |            |
|                | KS     | 0.09080        | 0.01240   |            |

Figure 3 illustrates the empirical power function of the CUSUM test for constant model(20,288),(987,673) when the design points are constructed using the uniform measure $\lambda_n^{\ast}$. The observations are generated from the model:

$$
Y_n(t,s) = \frac{1}{\sqrt{n_1n_2}}(\tilde{\phi}_n(t,s) + \lambda f_1(t,s)) + \varepsilon(t,s),
$$

for testing constant model using the statistic

$$
\frac{1}{\hat{\sigma}_o}\int_D f_1 dS_{n_1n_2}(Y_{n_1n_2}),
$$

where for $(t,s) \in D$,

$$
\tilde{\phi}_n(t,s) = \tilde{\phi}_s(t) + \tilde{\phi}_t(s),
$$

with $\tilde{\phi}_s(t) = t/3$ and

$$
\tilde{\phi}_t(s) = s/2.
$$

The graphs of the finite sample power function $\Psi_{\varepsilon, \lambda_n}(\lambda f_1)$ together with the point-wise limit $\Phi(\lambda f_1)$, for $\alpha = 0.05$ and $\lambda$ varies in $[0, 14]$ are exhibited in Fig. 3 in four panels: (a), (b), (c) and (d) associated with the sample sizes $40 \times 50$, $60 \times 65$, $70 \times 75$ and $80 \times 85$, respectively, where:

$$
\Psi_{\varepsilon, \lambda_n}(\lambda f_1) = P\left(1 - \frac{1}{\hat{\sigma}_o}\int_D f_1 dS_{n_1n_2}(Y_{n_1n_2}) \geq 0.67151\right),
$$

$$
\Psi(\lambda f_1) = 1 - \Phi(1.64485 - 0.40825\lambda).
$$

The simulation shows that in the case of constant model the empirical power functions of the CUSUM test lie very close to their limits independent to the choice of the design strategy even for relatively moderate sample sizes. Thus $\Psi_\varepsilon$ as well as $\Psi_{\varepsilon, \lambda_n}$ give very good approximation to $\Psi_{\varepsilon, \lambda_n}$ and $\Psi_{\varepsilon, \lambda_n}$, respectively.

**First Order Model**

We simulate the power function of the CUSUM test for the setup considered in Subsection 5.1.2 in which we propose a test using the statistic

$$
\frac{1}{\hat{\sigma}_o}\int_D f_1 dS_{n_1n_2}(Y_{n_1n_2}),
$$

where the observations are generated independently from the normal distribution with mean given by:

$$
\frac{1}{\sqrt{n_1n_2}}(\tilde{\phi}_n(t,s) + \lambda f_1(t,s)),
$$

and unknown variance $\sigma^2$. In this case $\tilde{\phi}_n = \tilde{\phi}_s + \tilde{\phi}_t$, $\tilde{\phi}_s(t) = t/3$, $\tilde{\phi}_t(s) = s/2$ and $f_1(t,s) = f_1(t)$ and $f_1(s) = f_1(s)$, respectively. Under $H_0$ we assume that a first order model is adequate, while we are observing a second order model. Since $\|f_1\|_{L^2} = 1.50177$, then for $\alpha = 10\%$, we have:

$$
\Psi_{\varepsilon, \lambda_n}(\lambda f_1) = P\left(1 - \frac{1}{\hat{\sigma}_o}\int_D f_1 dS_{n_1n_2}(Y_{n_1n_2}) \geq 1.92459\right),
$$

$$
\Phi(\lambda f_1) = 1 - \Phi(1.28155 - 1.50177\lambda).
$$

Next, we consider the same test problem as before with a little modification in that the experimental design is now given by a regular lattice of size $n_1 \times n_2$ on the experimental region $\overline{I} = [0, 1] \times [0, 1]$. Then
the corresponding orthogonal version of the regression functions \{z_0, z_1, z_2, z_3, z_4, z_5\} are given by:

\[
\begin{align*}
\tilde{w}_0(t,s) &= t - 1/2, \\
\tilde{w}_1(t,s) &= s - 1/2,
\end{align*}
\]

We test \[H_0: g \in W := \left[\tilde{w}_0, \tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4\right]\]
against \[H_1: g \in V := \left[\tilde{w}_5, \tilde{w}_6, \tilde{w}_7, \tilde{w}_8\right]\]
by the statistic \[\frac{1}{\sigma_i^2} \int f_1 dS_{n_i} \left(Y_{n_i}^{*}\right)\]
where the observations are generated independently based on the normal distribution with mean:

\[
\frac{1}{\sqrt{n_1 n_2}} \left(g_i \left(\ell / n_1, k / n_2\right) + \lambda f_i \left(\ell / n_1, k / n_2\right)\right)
\]
and unknown variance \(\sigma^2\), thereby \(g_i = \tilde{w}_5 + \tilde{w}_6 + \tilde{w}_7 \in W\) and \(f_i = \tilde{w}_8 \in V \cap W^\perp\). In this case we get after little computation \(\|f_i\|_{2^*} = 0.38879\) and therefore \(\Phi^{-1}(0.90)\|f_i\|_{2^*} = 0.49825\), giving the powers:

\[
\left|\Psi_{j_{f_i}}(\lambda f_i)\right| = \lim_{n_i \to \infty} \frac{1}{\sigma_i^2} \int f_1 dS_{n_i} \left(Y_{n_i}^{*}\right) \geq 0.49825
\]

evaluated at \(\lambda f_i \in V \cap W^\perp\), for \(\lambda \in \mathbb{R}\).

For the two different situations we present the simulation results in Fig. 4 and 5, respectively. It can be seen therein that independent to the choice of the design strategy, the limiting power function gives relatively good approximation to that of the finite sample power function of the CUSUM test. Both quantities achieve the specified probability 10% when \(\lambda\) is set to zero even for relatively moderate sample sizes. Thus, in the practice we can realize the test by directly calculating the quantiles of the limiting distribution of \(CU \left(Y_{n_i}^{*}\right)\) under \(H_0\) which is given by \(\Phi^{-1}(1-\alpha)\|f_i\|_{2^*}\), for a given \(\alpha \in (0, 1)\) and \(f_i \in V \cap W^\perp\).

Fig. 4. The graphs of the empirical power functions of the asymptotic size \(\alpha = 10\%\) CUSUM test for first order model under the design generated using the CDF \(F(t, s) = 12(1-1/s)(1.2-1/t)\) for \((t, s) \in D\) represented by using dotted line approximated by the limiting power function \(\Psi_{j_{f_i}}(\lambda f_i)\) scattered using smooth lines. (a) Sample size = 60x65, (b) Sample size = 70x75
Fig. 5. The graphs of the empirical power functions of the asymptotic size $\alpha = 10\%$ CUSUM test for first order model represented by using dotted lines approximated by the limit power function $\Psi\lambda^2_i\left(\hat{\lambda}_{f_i}\right)$ scattered using smooth lines. The design is $n_1 \times n_1$ regular lattice: (a) 50×55, (b) 60×65, (c) 70×75 and (d) 80×85

Fig. 6. The three dimensional scatter plot of the trend surface of the logarithm of the percentage of Fe. Source of data: Tahir (2010)

Numerical Application

This section illustrates strategies for selecting appropriate model for describing the physically meaningful functional relationship between the conditional and the response variables given by the coordinate and the observed percentage of Ferum (Fe), respectively. We study the data provided in Tahir (2010) received from a mining industry, which consists of the percentage of Fe observed independently over 7×14
lattice points of drilling bores on the exploration region of the company with 7 equidistance column running from west to east and 14 equidistance rows running from south to north as scattered in Fig. 6. Here our goal is to verify by conducting both MOSUM and CUSUM tests for checking whether or not a first-order model appropriate for describing the model is. For that we regard the observation as a realization of a regression model defined on the unit rectangle $I^2$ by putting the coordinate (-5824,-6000) where the observation process was initiated as the point (0,0) and the coordinate (-5825, -5725) where the observation process was ended as the point (1, 1). To stabilize the variance we apply logarithm transformation to the percentages of Fe. We denote the transformed measurement as LogFe. Preliminary goodness of fit for the normality of the sample is presented in Fig. 7 which shows that the distribution model of LogFe is not fit to normal family.

Table 2 presents the computation results of the critical values and the corresponding $p$-values of the MOSUM and CUSUM tests compared with those proposed in (Somayasa et al., 2015a) defined by using the Kolmogorov-Smirnov (KS) and Cramer-von Mises (CvM) functionals of the partial sums process of the least squares residuals. The $p$-values of the KS and CvM tests are approximated by simulation, whereas those of the MOSUM and CUSUM tests are computed analytically. For the computation of the CUSUM, KS and CvM statistics we consider the VCC class $\{(0, t] \times [0,s]: 0 \leq t, s \leq 1\}$ instead of the family of all convex sets. See also the cumulative sums defined in (Bischoff and Somayasa, 2009; Somayasa et al., 2015a; Xie and MacNeill, 2006). For each assumed model we also calculate $\hat{\sigma}_e$ using the method proposed in (Arnold, 1981).

In the case of constant model, both the MOSUM and the CUSUM tests result in too small $p$-values. This mans that the constant model is not appropriate for LogFe under the MOSUM and CUSUM tests. But quite different result is obtained when we consider the $p$-values of the KS and CvM tests in which constant model could be adequate at level less than 3% for KS and at level less than 1% for CvM, respectively.

The MOSUM as well as the CUSUM tests do not reject a first order model for all level of significance $\alpha \leq 1.668\%$ and $\alpha \leq 0.297\%$, respectively. This means that there is a significance evidence where a first order model is fit to the sample. Under additional information obtained from the $p$-values of the KS and CvM tests in which the hypothesis is not rejected for almost all frequently used values of $\alpha$, it can be stated that first order model is an appropriate model for LogFe. This conjecture is also suitable with the scatter plot of the data. Even though second order model is also fit to the model when the test is conducted using the MOSUM and KS tests as their $p$-values show, since the CUSUM and the CvM tests show evidence that it only fit for $\alpha$ set less then or equal 0.459 and 1.24%, respectively, we recommend that first order model is the most appropriate for the LogFe.

---

Fig. 7. The qqnorm plot for LogFe. Two sided Kolmogorov-Smirnov for normality assign the critical value 0.0933 and the $p$-value 0.0352. Source of data: Tahir (2010)
Concluding Remark

We established an asymptotically size $\alpha$ test for checking the appropriateness of a spatial regression whose critical region is constructed by using the probability distribution of the statistic expressed as the integral of the competing known regression function with respect to the generalized $(h_1,h_2)$-Slepian field. This statistic appears as the limit of the integral of such regression function with respect to the MOSUM process of the matrix of the observations. Other test which is called CUSUM test is also proposed defined in the like way as in the MOSUM test. We show that our tests are more applicable in the case where the design strategy must be incorporated in the analysis. Beside that our test procedures are also tractable in the sense the quantities such as the quantiles of the limiting distribution and the $p$-values can be computed analytically. Based on the limit power functions of both tests it is shown that the MOSUM test is asymptotically more sensitive in detecting the change of the model than the CUSUM test. In the present work the result was derived under independently distributed observations. In the future we put our setup in a more general and reasonable situation in which the observations are assumed to be dependent or at least stationary. This approach will be useful for handling the modeling problem of spatial data. In a forthcoming paper of Somayasa and et al. the investigation is extended to multivariate spatial regression model defined in (Somayasa et al., 2015b; Somayasa and Wibawa, 2015; Somayasa et al., 2016) by considering the moving sum process of the multivariate recursive residuals.

Acknowledgment

The authors wish to thank Professor Wolfgang Bischoff for the kind discussion during the preparation of the manuscript. We are grateful to the Ministry of RISTEK-DIKTI for the financial support of the research through the Fundamental research award 2016.

Author’s Contributions

The corresponding author of this work has contributed in establishing the mathematical derivation of the results and developing the simulation. The second author took part in giving interpretation of the model from the perspective of geostatistics. The third author provided discussion regarding the suitability of the assumed model to the mining data. The fourth author has evaluated the first draft of the paper.

Ethics

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Alexander, K.S. and R. Pyke, 1986. A uniform central limit theorem for set-indexed partial-sum processes with finite variance. Annals Probability, 14: 582-597.

Arnold, S.F., 1981. The Theory of Linear Models and Multivariate Analysis. 1st Edn., John Wiley and Sons, Inc., New York, ISBN-10: 0471050652, pp: 475.

Athreya, K.B. and S.N. Lahiri, 2006. Measure Theory and Probability Theory. 1st Edn., Springer Science and Business Media, New York, ISBN-10: 038732903X, pp: 618.

Billingsley, P., 1999. Convergence of Probability Measures. 2nd Edn., John Wiley and Sons, Inc., New York, ISBN-10: 0471974539, pp: 296.

Bischoff, W. and A. Gegg, 2014. The Cameron-martin theorem for $(p)$-Slepian processes. Preprint, Catholic University Eichstaett-Ingolstadt, Germany.

Bischoff, W. and W. Somayasa, 2009. The limit of the partial sums process of spatial least squares residuals. J. Multivariate Anal., 100: 2167-2177. DOI: 10.1016/j.jmva.2009.04.005

Bischoff, W., 2002. The structure of residual partial sums limit processes of linear regression models. Theory Stochastic Processes, 2: 23-28.

Chu, C.S.J., K. Hornik and C.M. Kuan, 1995. MOSUM tests for parameter constancy. Biometrika, 82: 603-617. DOI: 10.2307/2337537

Cohn, D.L., 1980. Measure Theory. 1st Edn., Birkh Äauser, Inc., Boston, ISBN-10: 1489904018, pp: 373.

Cressie, N.A.C., 1993. Statistics for Spatial Data. 2nd Edn., Wiley and Sons Inc., New York, ISBN-10: 0471002550, pp: 900.

Eubank, R.L. and J.D. Hart, 1993. Commonality of cusum, von neumann and smoothing-based goodness-of-fit tests. Biometrika, 80: 89-98. DOI: 10.2307/2336759

Fuchang, G. and W.V. Li, 2007. Small ball probabilities for the Slepian Gaussian fields. Trans. Am. Math. Society, 359: 1339-1350. DOI: 10.1090/S0002-9947-06-03963-8

Gaenssler, P., 1993. On recent development in the theory of set-indexed processes (A unified approach to empirical and partial-sum processes) in Asymptotic Statistics. Springer, Berlin.

Lifshits, M., 2012. Lectures on Gaussian Processes. 1st Edn., Springer Science and Business Media, ISBN-10: 3642249388, pp: 121.

MacNeill, I.B. and V.K. Jandhyala, 1993. Change-Point methods for Spatial Data, Multivariate Environmental Statistics, Patil, G.P. and C.R. Rao (Eds.), Elsevier Science Publishers B.V., pp: 298-306.

MacNeill, I.B., Y. Mao and L. Xie, 1994. Modeling heteroscedastic age-period-cohort cancer data. Canad. J. Statist., 22: 529-539. DOI: 10.2307/3315408
Pyke, R., 1983. A uniform central limit theorem for set-indexed partial sums processes with finite variance. Ann. Probab., 79: 219-240.

Ripley, B.D., 2004. Spatial Statistics. 1st Edn., John Wiley and Sons, Inc., New Jersey, ISBN-10: 047169116X, pp: 252.

Somayasa, W. and G.N.A. Wibawa, 2015. Asymptotic model-check for multivariate spatial regression with correlated responses. Fareast J. Math. Sci., 98: 613-939. DOI: 10.17654/FJMSNov2015_613_639

Somayasa, W., 2013. The partial sums of the least squares residuals of spatial observations sampled according to a probability measure. J. Indones. Math. Soc. 19: 23-40.

Somayasa, W., E. Ruslan and L.O. Cahyono, 2015a. Engkoiimoto, Cheking adequateness of spatial regressions using set-indexed partial sums technique. Fareast J. Math. Sci., 96: 933-966.

Somayasa, W., G.N. Adhi Wibawa, L. Hamimu and L.O. Ngkoimanto, 2016. Asymptotic theory in model diagnostic for general multivariate spatial regression. Int. J. Math. Math. Sci., 2016: 2601601-2601616. DOI: 10.1155/2016/2601601

Somayasa, W., G.N.A. Wibawa and Y.B. Pasolon, 2015b. Multidimensional set-indexed partial sums method for checking the appropriateness of a multivariate spatial regression. Int. J. Math. Model Meth. Applied Sci., 9: 700-713.

Stroock, D.W., 1994. A Concise Introduction to the Theory of Integration. 2nd Edn., Birkhäuser, Berlin, ISBN-10: 0817637591, pp: 184.

Stute, W., 1997. Nonparametric model checks for regression. Ann. Stat., 25: 613-641.

Stute, W., W.L. Xu and L.X. Zhu, 2008. Model diagnosis for parametric regression in high-dimensional spaces. Biometrika, 95: 451-467.

Tahir, M., 2010. Prediction of the amount of nickel deposit processes and boundary detection. South African Stat. J., 4: 33-53.

Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications. 3rd Edn., Springer, Berlin, ISBN-10: 3540441425, pp: 388.

Xie, L. and I.B. MacNeill, 2006. Spatial residual processes and boundary detection. South African Stat. J., 4: 33-53.

Appendix A. Generalized \((h_1,h_2)\)-Slepian Field

Definition A.1

For \(i = 1, 2\), let \(h_i\) with \(0 < h_1 < (a_2-a_1)\) and \(0 < h_2 < (b_2-b_1)\) be positive real numbers. A process \(S_{h_1,h_2}(t, s): (t, s) \in D_{h_1,h_2}\) is called a generalized \((h_1,h_2)\)-Slepian field, if and only if \(S_{h_1,h_2}\) is a centered Gaussian process with the covariance function:

\[
\gamma_{S_{h_1,h_2}}((t_1,s_1),(t_2,s_2)) = P\left\{ x \in [s_1, s_2] \cap [t_1, t_2] \right\}
\]

for \((t_1, s_1), (t_2, s_2) \in D_{h_1,h_2} := [(a_1, a_2-h_1) \times [b_1, b_2-h_2]\), where \([x, y]\) and \([y, z]\) stand for \([x, \infty] \) and \([y, \infty)\), respectively. In the case where \(P\) is the Lebesgue measure \(\lambda^2\) on \(D\), we call \(S_{h_1,h_2}\) a generalized \((h_1/h_2)\)-Slepian field, denoted by \(S_{h_1/h_2}\), with the covariance function:

\[
\gamma_{S_{h_1/h_2}}((t_1,s_1),(t_2,s_2)) = \left( h_1 - |t_1 - t_2| \right) \left( h_2 - |s_1 - s_2| \right)
\]

where, \(x^\prime := \max\{x, 0\}\), see also (Bischoff and Gegg, 2014; Chu et al., 1995; Fuchang and Li, 2007).

Remark A.2

Some immediate consequences implied by the definition of \(S_{h_1,h_2}\) can be summarized as follows:

- For every \((t, s) \in D_{h_1,h_2}, S_{h_1,h_2}(t, s) \sim N(0, P([t, t_2] \times [s, s_2]))
- Let \(W_P(t, s) := W_p([a_1, t] \times [b_1, s])\). For fixed \(h_1, h_2\) defined above, let:
  \[\Delta W_P(x, y) := W_P(x+h_1+y+h_2) - W_P(x, y + h_1) - W_P(x + h_1, y) + W_P(x, y)\]

Since \(\Delta W_P(x, y)\) is also distributed as \(N(0, P([t, t_2] \times [s, s_2])\), then \(S_{h_1,h_2}\) is said to be equivalent in distribution to \(\Delta W_P\), denoted by \(S_{h_1,h_2} \overset{\Delta}{=} \Delta W_P\).

Theorem A.3. (Invariance Principle for MOSUM Process)

Let \(\{e_{n \times n}\}_{n \in \mathbb{N}, n \geq 2}\) be a sequence of independent and identically distributed random matrices, where \(e_{n \times n} := \left(e_{i,j}\right)_{i=1,j=1}^{n,n}, e_{i,j} := \epsilon_i \epsilon_j \eta_{i,j}\), with \(E(\epsilon_i) = 0\) and \(Var(\epsilon_i) = \sigma^2 < \infty\), for \(1 \leq i < \infty\) and \(1 \leq k < \infty\). Let the experimental design \(\Xi_{n_1,n_2}\) is constructed using a probability measure \(P\) on \(D = [a_1, a_2] \times [b_1, b_2]\), such that \(P = P_1 \otimes P_2\) with the associated distribution function \(F\) satisfying \(F = F_1 \times F_2\) for continuous and increasing marginals \(F_1\) and \(F_2\) on \([a_1, a_2]\) and \([b_1, b_2]\), respectively. Then as \(n_1, n_2 \to 1\), we have:

\[
\frac{1}{\sigma} MS_{h_1,h_2}(e_{n \times n}) \Rightarrow S_{h_1,h_2}
\]
Proof

By applying the well-known Prohorov’s theorem (cf. Billingsley (1999), pp. 35-40), we show first that the finite dimensional distribution of $\frac{1}{\sigma} MS_{\Delta h} (\xi_{n_{a,n}})$ converges to that of $S_{\Delta t}$. That is for any points $(t_1, s_1), (t_2, s_2), \ldots, (t_m, s_m)$ in $D$ and any constants $b_{1}, \ldots, b_{m}$ we show that the sequence $\{C_{n_{a,n}}\}$, for $n_{1} \geq 1$, $n_{2} \geq 1$, where $C_{n_{a,n}} := \sum_{i=1}^{n_{1}} b_{i} \frac{1}{\sigma} MS_{\Delta h} (\xi_{n_{a,n}})(t_{i}, s_{i})$ converges in distribution to $\sum_{i=1}^{n_{1}} b_{i} S_{\Delta t}(t_{i}, s_{i})$. The last is normally distributed with mean zero and variance $\sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}} b_{i} b_{j} P[[t_{i}, t_{j}] \times [s_{i}, s_{j}] \cap [t_{j}, t_{j}] \times [s_{j}, s_{j}]]$. By the definition of $MS_{\Delta h} (\xi_{n_{a,n}})$, we get:

$$
Var(C_{n_{a,n}}) = \sum_{i=1}^{n_{1}} b_{i}^{2} E \left[ MS_{\Delta h} (\xi_{n_{a,n}})(t_{i}, s_{i}) MS_{\Delta h} (\xi_{n_{a,n}})(t_{j}, s_{j}) \right] = \sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}} \sum_{k=1}^{n_{1}} b_{i} b_{j} b_{k} E(\xi_{i}, \xi_{j}, \xi_{k}) \frac{1}{\sigma n_{1} n_{2}}.
$$

Furthermore, by the independence of $\xi_{i}$, we have:

$$
E(\xi_{i}, \xi_{j}, \xi_{k}) = \frac{1}{\sigma n_{1} n_{2}} \left( n_{1} F_{1} (s_{i} + h_{i}) - n_{2} F_{2} (s_{i} + h_{i}) \right) \left( n_{1} F_{1} (s_{j} + h_{j}) - n_{2} F_{2} (s_{j} + h_{j}) \right) \left( n_{1} F_{1} (t_{i} + h_{i}) - n_{2} F_{2} (t_{i} + h_{i}) \right) \left( n_{1} F_{1} (t_{j} + h_{j}) - n_{2} F_{2} (t_{j} + h_{j}) \right) \left( n_{1} F_{1} (s_{i}) - n_{2} F_{2} (s_{i}) \right) \left( n_{1} F_{1} (s_{j}) - n_{2} F_{2} (s_{j}) \right) \left( n_{1} F_{1} (t_{i}) - n_{2} F_{2} (t_{i}) \right) \left( n_{1} F_{1} (t_{j}) - n_{2} F_{2} (t_{j}) \right).
$$

which converges as $n_{1}, n_{2} \rightarrow \infty$, to:

$$
= \left\{ \min\{F_{1}(s_{i} + h_{i}), F_{2}(s_{i} + h_{i})\} - \max\{F_{1}(s_{i}), F_{2}(s_{i})\} \right\} \times \left\{ \min\{F_{1}(t_{i} + h_{i}), F_{2}(t_{i} + h_{i})\} - \max\{F_{1}(t_{i}), F_{2}(t_{i})\} \right\}
$$

or

$$
P \left[ \left\{ \min\{F_{1}(s_{i} + h_{i}), F_{2}(s_{i} + h_{i})\} - \max\{F_{1}(s_{i}), F_{2}(s_{i})\} \right\} \times \left\{ \min\{F_{1}(t_{i} + h_{i}), F_{2}(t_{i} + h_{i})\} - \max\{F_{1}(t_{i}), F_{2}(t_{i})\} \right\} \right]
$$

where $\sigma = \sqrt{\frac{1}{\sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}} b_{i} b_{j} E(\xi_{i}, \xi_{j}, \xi_{k}) \frac{1}{\sigma n_{1} n_{2}}}}$.

The first equality of the preceding result follows from the assumption that $F_{1}$ and $F_{2}$ are increasing on $[a_{1}, a_{2}]$ and $[b_{1}, b_{2}]$, respectively. Thus it is shown that for $n_{1}, n_{2} \rightarrow \infty$:

$$
Var(C_{n_{a,n}}) \rightarrow \sum_{i=1}^{n_{1}} b_{i} b_{j} P[[t_{i}, t_{j}] \times [s_{i}, s_{j}] \cap [t_{j}, t_{j}] \times [s_{j}, s_{j}]].
$$

which is the variance of $\sum_{i=1}^{n_{1}} b_{i} S_{\Delta t}(t_{i}, s_{i})$. Next we show that for every $\epsilon > 0$, it holds $\lim_{n_{1}, n_{2} \rightarrow \infty} L(\epsilon) = 0$, where:

$$
L(\epsilon) := \sum_{i=1}^{n_{1}} b_{i} \sum_{k=1}^{n_{2}} \left( n_{1} F_{1}(s_{i}) - n_{2} F_{2}(s_{i}) \right) \left( n_{1} F_{1}(t_{i}) - n_{2} F_{2}(t_{i}) \right) \frac{1}{\sigma n_{1} n_{2}}\left( n_{1} F_{1}(s_{i} + h_{i}) - n_{2} F_{2}(s_{i} + h_{i}) \right) \left( n_{1} F_{1}(t_{i} + h_{i}) - n_{2} F_{2}(t_{i} + h_{i}) \right)
$$

Let $K_{\epsilon} := \max_{a_{1}, a_{2}, b_{1}, b_{2}} |b|$, since $\epsilon_{a}$ is independent and identically distributed, then we get:

$$
0 < \lim_{n_{1}, n_{2} \rightarrow \infty} L(\epsilon) \leq \frac{mk}{\sigma^{2}} \left( \frac{n_{1} F_{1}(s_{i} + h_{i})}{n_{2}} - \frac{n_{1} F_{1}(t_{i})}{n_{1}} \right) \left( \frac{n_{1} F_{1}(t_{i} + h_{i})}{n_{1}} - \frac{n_{1} F_{1}(s_{i})}{n_{2}} \right) \left( \frac{n_{1} F_{1}(s_{i})}{n_{1}} - \frac{n_{1} F_{1}(t_{i})}{n_{2}} \right) \left( \frac{n_{1} F_{1}(t_{i} + h_{i})}{n_{1}} - \frac{n_{1} F_{1}(s_{i} + h_{i})}{n_{2}} \right) = 0
$$

by the dominated convergence theorem. Hence, the Lindeberg central limit theorem (cf. Athreya and Lahiri (2006), pp. 343-345), leads us to the conclusion that $C_{n_{a,n}}$ converges in distribution to $\sum_{i=1}^{n_{1}} b_{i} S_{\Delta t}(t_{i}, s_{i})$. In the last step of the proof we have to show the tightness of the MOSUM process. It can be shown that for every $(t, s) \in D$, it holds:

$$
MS_{\Delta h} (\xi_{n_{a,n}})(t, s) = S_{n_{a,n}}(t, s) = \left[ I_{n} \times I_{m} \right] \left( t, s \right)
$$

that is the MOSUM process coincides with the CUSUM process when the index sets are restricted to the VCC $\{(t, t) \times [s, s]: (t, s) \in D\}$ of subsets of $D$. Due to the results of Alexander and Pyke (1986) and Pyke (1983), the tightness of $MS_{\Delta h} (\xi_{n_{a,n}})$
immediately follows from the tightness of $S_{\lambda,n_1}(\varepsilon_{\lambda,n_1})$, establishing the proof.

**Definition A.4**

Let $\Gamma_1 := \left\{ (t_1,t_2) \mid t_1 < t_2 < \cdots < t_m = a_2 \right\}$ be a set of m1 closed intervals on $[a_1, a_2]$, such that $a_i = t_i < t_{i+1} < \cdots < t_m = a_2$. Let $\Gamma_2 := \left\{ (s_1,s_2) \mid s_1 < s_2 < \cdots < s_{m_2} = s_2 \right\}$ be a set of m2 closed intervals on $[b_1, b_2]$, such that $b_1 = s_0 < s_1 < s_2 < \cdots < s_{m_2} = s_2$. The Cartesian product $K := \Gamma_1 \times \Gamma_2$ which are consisting of $m_1 m_2$ closed rectangles which builds a partition on $D$. Proposition A.6: The ordinary $(h_1,h_2)$-Slepian field $S_{i,j}$ is a centered Gaussian process with independent increments.

**Proof**

Let $[t_{i-1}, t_i] \times [s_{i-1}, s_i]$ and $[t_{i-1}, t_i] \times [s_{i-1}, s_i]$ be arbitrary disjoint rectangles on $D$. Then $\text{Cov}(\Delta_{i-1,i} S_{i-1,i} \times S_{i-1,i})$ is computed by the following formula:

$$\text{Cov}(\Delta_{i-1,i} S_{i-1,i} \times S_{i-1,i}) = P\left([t_{i-1}, t_i] \times [s_{i-1}, s_i] \cap [s_{i-1}, s_i] \times [s_{i-1}, s_i]\right) - P\left([t_{i-1}, t_i] \times [s_{i-1}, s_i] \cap [t_{i-1}, t_i] \times [s_{i-1}, s_i]\right)$$

where $P$ is substituted by $\varpi^2$.
\[ \left( h_i - t_{n_i} + t_{n_i+1} \right) \left( h_j - t_{n_j} + t_{n_j+1} \right) \]

\[ = \left( h_i - t_{n_i} + t_{n_i+1} \right) \left( h_j - t_{n_j} + t_{n_j+1} \right) = \left( h_i - t_{n_i} + t_{n_i+1} \right) \left( h_j - t_{n_j} + t_{n_j+1} \right) \]

Proposition A.7

Let \( I_{kl} := \left[ t_{n_{i,k}} \right] \times \left[ t_{n_{j,k}} \right] \) be any rectangle in \( D_{h,k} \). For any \( h_1, h_2 \) with \( 0 < h_1 < \lambda a_2 - a_1 \) and \( 0 < h_2 < \lambda (b_2 - b_1) \), it holds \( \text{Var}(\Delta_{i,k}, S_{j,k}) = 4 \lambda^2 (I_{kl}) \).

Proof

By a result in multivariate analysis, \( \text{Var}(\Delta_{i,k}, S_{j,k}) = a^T E(hh^T) a \), where the vector \( h \) is defined as:

\[ h = \begin{pmatrix} S_{i,k} \left( t_{n_{i,k}+1} \right) \times S_{j,k} \left( t_{n_{j,k}+1} \right) \\ S_{i,k} \left( t_{n_{i,k}+1} \right) \times S_{j,k} \left( t_{n_{j,k}+1} \right) \end{pmatrix}^T \]

and \( a := (1,-1,1,1) \). The definition of the covariance function \( \gamma_{\beta, \beta} \) leads us to the following variance-covariance matrix of \( h \):

\[ E(hh^T) = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix} \]

where:

\[ c_{11} = c_{22} = c_{33} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} + t_{n_{j,k}+1} \right) \]

\[ c_{12} = c_{13} = c_{23} = c_{14} = c_{24} = c_{34} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} - t_{n_{j,k}+1} \right) \]

\[ c_{32} = c_{33} = c_{34} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} + t_{n_{j,k}+1} \right) \]

\[ c_{33} = c_{34} = c_{43} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} - t_{n_{j,k}+1} \right) \]

\[ c_{42} = c_{43} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} + t_{n_{j,k}+1} \right) \]

\[ c_{43} = c_{44} = h_i \left( h_j - t_{n_{i,k}+1} - t_{n_{j,k}+1} \right) \]

After some computations, \( a^T E(hh^T) a \) can be further expressed as:

\[ a^T E(hh^T) a = 4h_i h_j + 4 \left( h_i - t_{n_{i,k}+1} + t_{n_{j,k}+1} \right) \left( h_j - t_{n_{i,k}+1} - t_{n_{j,k}+1} \right) \]

By Equation 3, for every \( (t, s) \in D_{h,k} \), it holds:

\[ \int_{t_{n_{i,k}}}^{t_{n_{i,k}+1}} f_t \left( \frac{\gamma_{\beta, \beta}(t, s_{n_{i,k}})}{a_{n_{i,k}}} \right) dt + MS_{h,k} \left( \epsilon_{n_{i,k}} \right) \]

By the invariance principle for MOSUM process (Theorem A.3) and the fact that \( P_\sigma \Rightarrow P \). Hence, by the continuous mapping theorem we get:

\[ 1 \sigma^2 \left( t, s \right) + MS_{h,k} \left( \epsilon_{n_{i,k}} \right) \]

\[ \Rightarrow 1 \sigma^2 \left( t, s \right) + MS_{h,k} \left( \epsilon_{n_{i,k}} \right) \]
\[
\frac{1}{d} \sum_{j=1}^{d} f_i \mathcal{M}_n \left( X_i, \eta \right) = \frac{1}{d} \sum_{j=1}^{d} f_i \mathcal{S}_n (t,s) + \int_{x_{n-1}}^{x_n} f_i \mathcal{S}_n (t,s)
\]
which is establishing the proof.

Appendix C. Set-Indexed Gaussian white Noise

**Definition C.1.** (Lifshits (2012), pp. 13-15)

Let \( P \) be a probability measure on \( D, B(D) \) and \( \mathcal{A}_0 := \{ A \in B(D) : P(A) < \infty \} \) be a class of subsets of \( D \) which has finite measure under \( P \). Let \( C(\mathcal{A}_0) \) be the space of functions on \( \mathcal{A}_0 \) which are \( d_P \) -uniformly continuous, where \( d_P \) is a pseudo metric on \( A_0 \times A_0 \) defined by \( d_P (A_1, A_2) := P(A_1, A_2) \). As usual \( C(\mathcal{A}_0) \) is furnished with the uniform norm \( \| \cdot \|_\infty := \sup_{A \in \mathcal{A}_0} | P(A) | \), for any \( w \in C(\mathcal{A}_0) \). A centered Gaussian process \( \mathcal{W}_P := \{ \mathcal{W}_P (A), A \in \mathcal{A}_0 \} \) defined on a common probability space \((\Omega, \mathcal{F}, P)\), say, is called univariate Gaussian white noise with the control measure \( P \), if and only if:

\[
E_P \left( \mathcal{W}_P (A) \mathcal{W}_P (B) \right) = P(A \cap B), \quad \forall A, B \in \mathcal{A}_0
\]

The sample path (trajectory) of \( \mathcal{W}_P \) is concentrated in \( C(\mathcal{A}_0) \) (cf. Lifshits, 2012), pp. 13-15).

**Theorem C.2.** (Construction of the Set-Indexed Gaussian white Noise)

Let the experimental design \( \mathcal{X}_{n_1, n_2} \) be constructed by using the probability measure \( P \) on \( D \) and let the sequence of the matrix of random errors \( \xi_{n_1, n_2} := (\xi_{i,j})_{i=1,\ldots,n_1; j=1,\ldots,n_2} \) consists of i.i.d. random variables with \( E(\xi_{i,j}) = 0 \) and \( \text{Var}(\xi_{i,j}) = \sigma^2 < 1 \), for \( n_1 \geq 1 \) and \( n_2 \geq 1 \), where \( \xi_{i,j} := \xi \left( x_{i,j} \right) \). Then we have:

\[
\frac{1}{\sigma^4} \mathcal{S}_{n_1, n_2} \left( \xi_{n_1, n_2} \right) := \mathcal{W}_P, \text{ as } n_1 \text{ and } n_2 \to \infty
\]

**Proof**

Let \( c_1, \ldots, c_n \) be any constant and \( B_1, \ldots, B_m \) be any subset in \( \mathcal{A}_0 \), \( m \geq 1 \). We define linear combination \( F_{n, n_2} := \sum_{i=1}^{n} c_i \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{i, n_2} \right) \right) \left( B_i \right) \). Then we get:

\[
\text{Var} \left( F_{n, n_2} \right) = E \left( \sum_{i=1}^{n} c_i \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{i, n_2} \right) \right) \left( B_i \right) \right) \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{i, n_2} \right) \right) \left( B_i \right)
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{i, n_2} \right) \right) \left( B_i \right) \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{j, n_2} \right) \right) \left( B_j \right)
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \mathcal{P}_n \left( B_i \right) \mathcal{P}_n \left( B_j \right) \to \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \mathcal{P} \left( B_i \cap B_j \right)
\]

where \( \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j P \left( B_i \cap B_j \right) \) is actually the variance of \( \sum_{i=1}^{n} c_i \mathcal{W}_P \left( B_i \right) \) which is normally distributed with mean zero and variance \( \sum_{i=1}^{n} c_i c_j P \left( B_i \cap B_j \right) \). Next we show that the Lindeberg condition is fulfilled. For every \( \epsilon > 0 \), let:

\[
L(\epsilon) := \sum_{i=1}^{n} E \left( \sum_{j=1}^{n} \left| c_i c_j \mathcal{P}_n \left( B_i \right) \mathcal{P}_n \left( B_j \right) \right| \right)^2 x_1 \left( \sum_{i=1}^{n} \sum_{j=1}^{n} \left| c_i c_j \mathcal{P}_n \left( B_i \cap B_j \right) \right| \right)
\]

and let \( M := \max_{1 \leq i \leq n} \mathcal{P}_n \left( B_i \right) \). Then by the i.i.d. property of \( \mathcal{P}_n \) and by the well-known bounded convergence theorem it holds:

\[
0 \leq \lim_{n \to \infty} L(\epsilon) \leq M^2 \mathcal{E} \left( \epsilon_1 \right)^2 \left( \mathcal{S}_{n_1, n_2} \left( \xi \left( x_{i, n_2} \right) \right) \right) = 0
\]

Thus, by the Lindeberg-Levy central limit theorem, the finite dimensional distribution of \( S_{n_1, n_2} \) converges weakly to that of \( \mathcal{W}_P \). Since \( S_{n_1, n_2} \) is \( d_P \)-continuous, we define the modulus of continuity of \( S_{n_1, n_2} \) as:

\[
W \left( S_{n_1, n_2} : \delta \right) := \sup_{\left| \chi \right| \leq \delta} \left| S_{n_1, n_2} \left( \xi \right) \right| \left( A \right) - S_{n_1, n_2} \left( \xi \right) \left( B \right)
\]

Hence in order to show tightness it suffices to show that:

\[
\lim \sup_{n \to \infty} \mathcal{P} \left( \omega \in \Omega : W \left( S_{n_1, n_2} : \delta \right) \to 0 \right) = 0, \forall \mathcal{P} > 0
\]

To this end we refer the reader to (Alexander and Pyke, 1986; Gaenssler, 1993; Pyke, 1983), establishing the proof.

**Proposition C.3**

The process \( \mathcal{W}_P \) constitutes a finite signed measure \( P \)-almost surely on \( \mathcal{A}_0 \). That is there exists a set \( \Omega \subset \Omega \) with \( P(\Omega^c) = 0 \) such that \( \forall \omega \in \Omega \), \( \mathcal{W}_P (\omega) \) is a finite signed measure on \( \mathcal{A}_0 \).

**Proof**

Since \( \text{Var}(\mathcal{W}_P (A)) = P(A), \forall A \in \mathcal{A}_0 \), then \( \text{Var}(\mathcal{W}_P (0)) = P(0) = 0 \). This implies \( \exists \Omega \) with \( P(\Omega^c) = 0 \), such that \( \mathcal{W}_P (\Omega^c) = 0, \forall \omega \in \Omega^c \). Next we show countable additivity. Let \( \left\{ A_n : n \geq 1 \right\} \) be a sequence of disjoint sets in \( \mathcal{A}_0 \), then \( \mathcal{W}_P \left( \bigcup_{n=1}^{\infty} A_n \right) \sim N(0, \sum_{n=1}^{\infty} P(A_n)) \).

Since on the other hand \( \sum_{n=1}^{\infty} \mathcal{W}_P (A_n) \) is also distributed...
as \( N(0, \sum_{j=1}^{n} P(A_j)) \) and it is well-known that a normal distribution model is determined uniquely by its mean and variance, we can conclude that both random variables are equivalent in distribution. That is:

\[
W_{\mathcal{F}}\left( \bigcup_{j=1}^{n} A_j \right) = \sum_{j=1}^{n} W_{\mathcal{F}}(A_j)
\]

This means that \( W_{\mathcal{F}} \) is countably additive \( P\)-a.s., finishing the proof.

**Proposition C.4**

For any \( f, g \in L_2(D, P) \), it holds:

\[
\text{Cov}\left( \int_{D} f dW_{\mathcal{F}} , \int_{D} g dW_{\mathcal{F}} \right) = \left\{ f, g \right\}_P
\]

\[
\int_{D} (f - g) dW_{\mathcal{F}} \sim N(0, \| f - g \|_P^2)
\]

**Proof**

First we refer the reader to Lifshits (2012), pp. 13-14 for the definition of \( \int_{D} f dW_{\mathcal{F}} \), which is defined path-wise as the integral with respect to the finite signed measure \( W_{\mathcal{F}} \). Let \( \{f_n := \sum_{j=1}^{\infty} a_j 1_{A_j} \}_{n \geq 1} \) and \( \{g_n := \sum_{j=1}^{\infty} b_j 1_{B_j} \}_{n \geq 1} \) be sequences of step functions converging uniformly to \( f \) and \( g \), respectively, with \( \bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j = D \). We notice that the existence of \( \{f_n\} \) and \( \{g_n\} \) is guaranteed by the denseness of the class of step functions in \( L_2(D, P) \). Then we have:

\[
\text{Cov}\left( \int_{D} f dW_{\mathcal{F}} , \int_{D} g dW_{\mathcal{F}} \right) = E_{\mathcal{F}} \left( \lim_{n \to \infty} \sum_{j=1}^{\infty} a_j b_j W_{\mathcal{F}}(A_j) W_{\mathcal{F}}(B_j) \right)
\]

\[
= \lim_{n \to \infty} \sum_{j=1}^{\infty} a_j b_j W_{\mathcal{F}}(A_j \cap B_j) P(d\omega)
\]

\[
= \lim_{n \to \infty} \sum_{j=1}^{\infty} a_j b_j \int_{D} W_{\mathcal{F}}(A_j \cap B_j) P(dt, ds)
\]

\[
\int_{D} f(t, s) g(t, s) P(dt, ds) = \{f, g\}_P
\]

Hence by this equation we further get \( \int_{D} f dW_{\mathcal{F}} \) is normally distributed with mean 0 and variance \( \| f \|_P^2 \). This result immediately implies \( \int_{D} (f - g) dW_{\mathcal{F}} \) is normally distributed with mean 0 and variance \( \| f - g \|_P^2 \). We are done.

**Appendix D. Proof of Theorem 4.1**

Without loss of generality we assume \( \sigma^2 = 1 \). Then by the definition of the model, if \( H_0 \) is true we have for every \( n_1 \geq 1 \) and \( n_2 \geq 1 \):

\[
CU\left( Y_{n_1} \right) = \int_{D} f_1(t, s) dS_{n_1}(Y_{n_1})(t, s)
\]

\[
= \sqrt{n_1 n_2} \int_{D} f_1(t, s) dS_{n_2}\left( \zeta_{n_1 n_2} \right)\left( t, s \right)
\]

\[
+ \int_{D} f_1 dS_{n_2}(\varepsilon_{n_1 n_2})(t, s)
\]

For large enough \( n_1 \) and \( n_2 \), it holds by recalling the invariance principle of the set-indexed process (cf. (Alexander and Pyke, 1986; Pyke, 1983; Gaenssler, 1993; Xie and MacNeill, 2006)) and the fact that \( g_1 \) is continuous and of bounded variation on \( D \).

\[
CU\left( Y_{n_1} \right) \sim \sqrt{n_1 n_2} \int_{D} f_1(t, s) d\varphi g_1(t, s) + \int_{D} f_1(t, s) dW_{\mathcal{F}}(t, s)
\]

However, since \( f_1 \) and \( g_1 \) are orthogonal in \( L_2(D, P) \) when \( H_0 \) is true, we get \( \int_{D} f_1(t, s) d\varphi g_1(t, s)\{f_1, g_1\}_P = 0 \).

Hence, under \( H_0 \), \( \text{CU}(Y_{n_1}) \) is normally distributed with mean zero and variance \( \| f_1 \|_P^2 \). Thus by applying the well known Lindeberg-Levy central limit theorem, we get the critical region of the asymptotically size \( \alpha \) test as:

\[
\text{CU}_{n}\left( Y_{n_1} \right) \in \mathbb{R}^n : \text{CU}(Y_{n_1}) \geq \Phi^{-1}(1-\alpha) \| f_1 \|_P
\]

**Appendix E. Proof of Corollary 4.2**

If \( g_2 = f \) for any \( f \in \mathcal{V} \cap W^2 \), by applying the similar argument as in the proof of Theorem 4.1, we get:

\[
\text{CU}\left( Y_{n_1}^* \right) \in \mathbb{R}^n : \text{CU}(Y_{n_1}) \geq \Phi^{-1}(1-\alpha) \| f_1 \|_P
\]

Based on this convergence we further get:

\[
\lim_{n_1, n_2 \to \infty} \Psi_{n_1 n_2}(f) = \lim_{n_1, n_2 \to \infty} P\left( \text{CU}(Y_{n_1}) \geq \Phi^{-1}(1-\alpha) \| f_1 \|_P \right)
\]

\[
= P\left( \{f_1, f_2\}_P + \int_{D} f_1(t, s) dW_{\mathcal{F}}(t, s) \geq \Phi^{-1}(1-\alpha) \| f_1 \|_P \right)
\]

\[
= 1 - \Phi\left( \Phi^{-1}(1-\alpha) - \frac{|\{f_1, f_2\}_P|}{\| f_1 \|_P} \right)
\]

126