Journal of Mathematics and Statistics

Original Research Paper

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

[1]Wayan Somayasa, [2]Jamhir Safani, [3]La Ode Ngkoimani and [4]Muhamad Zamrun F.

[1]*Department of Mathematics, Halu Oleo University, 93232 Kendari, Indonesia*
[2]*Department of Geophysics Engineering, Halu Oleo University, 93232 Kendari, Indonesia*
[3]*Department of Physics, Halu Oleo University, 93232 Kendari, Indonesia*
[4]*Department of Geology Engineering, Halu Oleo University, 93232 Kendari, Indonesia*

Article history
Received: 13-03-2017
Revised: 21-03-2017
Accepted: 29-04-2017

Corresponding Author:
Wayan Somayasa
Department of Mathematics,
Halu Oleo University, 93232
Kendari, Indonesia
Email: wayan.somayasa@uho.ac.id

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.

Science Publications

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_1h_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_1h_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 using arbitrary probability measure. We note that the ordinary (h_1h_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 Somaysa, 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:

$$\Xi_{n_1 \times n_2} := \left\{ \left(t_{n_1 \ell}, s_{n_2 k} \right) : 1 \le \ell \le n_1, 1 \le k \le n_2 \right\} \subset D, n_1 \ge 1, n_2 \ge 1$$

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 $t_{n_1\ell} = a_1 + \frac{(a_2 - a_1)}{n_1}\ell$ and $s_{n_2k} := b_1 + \frac{(b_2 - b_1)}{n_2}k$, for $1 \le \ell \le n_1$, $1 \le k \le n_2$, (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 $\lambda_D^{2*} := \frac{1}{|D|}\lambda_D^2$ in the sense the corresponding discrete measure converges to λ_D^{2*}. That is, let $P_{n_1n_2}$ be a discrete probability measure on the Borel σ-algebra $B(D)$, defined by:

$$P_{n_1n_2}(B) := \frac{1}{n_1n_2}\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2} 1_B\left(t_{n_1\ell}, s_{n_2k}\right),\ B \in B(D)$$

where, 1_B is a 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 $P_{n_1n_2} \Rightarrow \lambda_D^{2*}$, 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 σ-algebra $B(D)$ with the corresponding distribution function F, can we construct a design $\Xi_{n_1 \times n_2}$, such that $P_{n_1n_2} \Rightarrow P$. The answer is "yes". We can determine the design points $\left(t_{n_1\ell}, s_{n_2k}\right)$ in natural way by firstly partitioning the interval $[a_1, a_2]$ as $\left\{t_{n_11}, t_{n_12}, \ldots, t_{n_1n}\right\}$ using the equation $F\left(t_{n_1\ell}, b_2\right) = \ell/n_1$, for $1 \le \ell \le n_1$. Next, for a fixed ℓ and k, with $1 \le k \le n_2$, the point $\left(t_{n_1\ell}, s_{n_2k}\right)$ is generated by solving the equation $F\left(t_{n_1\ell}, s_{n_2k}\right) = \ell k/(n_1n_2)$. The solution is unique provided F is continuous and strictly increasing on D. Hence, by this sampling method $\Xi_{n_1 \times n_2}$ 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 $\Xi_{n_1 \times n_2}$ 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 $u(t, s) := 12/(t_2s_2)$ and the distribution function $F(t, s) := 12(1-1/t)(1/2-1/s)$, for $1 \le t \le 2$ and $2 \le s \le 3$, 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 $n_1 \ge 1$ and $n_2 \ge 1$, the design point is computed by the formula:

$$t_{n_1\ell} = \frac{2n_1}{2n_1 - \ell}\ and\ s_{n_2k} = \frac{6n_2}{3n_2 - k}, 1 \le \ell \le n_1, 1 \le k \le n_2$$

We notice that $P_{n_1n_2}$ can also be written as:

$$P_{n_1n_2}(B) := \frac{1}{n_1n_2}\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2} 1_B\left(F_1^{-1}(\ell/n_1), F_2^{-1}(k/n_2)\right), B \in B(D)$$

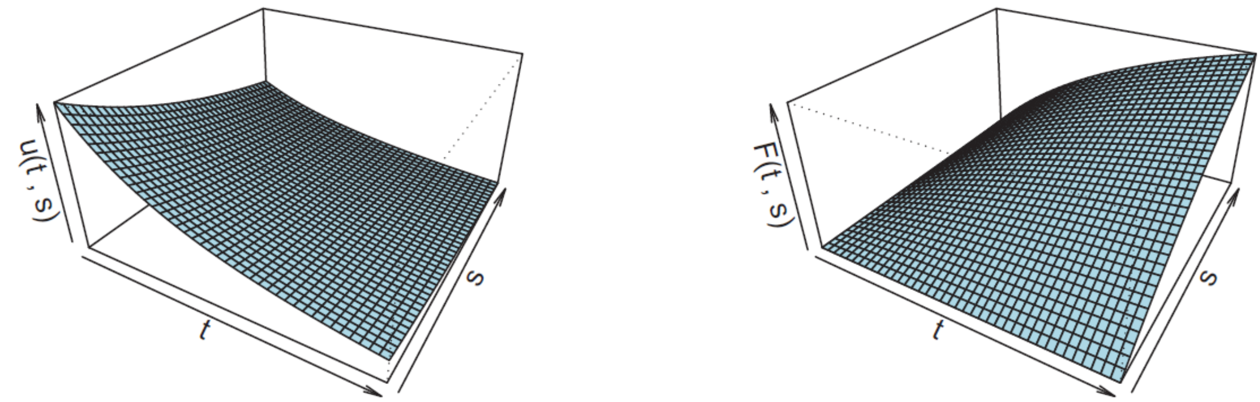


Fig. 1. Left side: The three dimensional perspective plot of the density function of P. Right side: The cumulative distribution functions of P on the compact rectangle $D = [1, 2] \times [2, 3]$

Let a nonparametric regression model $Y(x) = g(x) + \varepsilon(x)$, for $x \in D$ be observed on $\Xi_{n_1\times n_2}$, where g is an unknown function of bounded variation on D and ε is the unobserved random error defined on a common probability measure (Ω,F,P), say, with $E(\varepsilon(x)) = 0$ and $Var(\varepsilon(x)) = \sigma^2>0$ for every $x \in D$. Then the matrix of independent observations $Y_{n_2\times n_1} := \left(Y\left(t_{n_1\ell},s_{n_2k}\right)\right)_{\ell=1,k=1}^{n_1,n_2}$ satisfies the following decomposition:

$$Y_{n_2\times n_1} g_{n_2\times n_1} + \varepsilon_{n_2\times n_1} \tag{1}$$

where, $g_{n_2\times n_1} := \left(g\left(t_{n_1\ell},s_{n_2k}\right)\right)_{\ell=1,k=1}^{n_1,n_2}$ and $\varepsilon_{n_2\times n_1} := \left(\varepsilon\left(t_{n_1\ell},s_{n_2k}\right)\right)_{\ell=1,k=1}^{n_1,n_2}$ is an $n_2 \times n_1$ matrix of random errors having independent and identically distributed components with $E\left(\varepsilon\left(t_{n_1\ell},s_{n_2k}\right)\right)=0$ and $Var\left(\varepsilon\left(t_{n_1\ell},s_{n_2k}\right)\right) =\sigma^2>0$, for $1 \le \ell \le n_1$; $1 \le k \le n_2$. For the sake of brevity we write $Y\left(t_{n_1\ell},s_{n_2k}\right)$, $g\left(t_{n_1\ell},s_{n_2k}\right)$ and $\varepsilon\left(t_{n_1\ell},s_{n_2k}\right)$ throughout this paper as $Y_{k\ell}$, $g_{k\ell}$ and $\varepsilon_{k\ell}$, 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, 19993; Stute, 1997) a common feature of the GoF test for regression falls into the following framework. Let $V := [z_0,..., z_p, z_{p+1},..., z_m]$ and $W := [z_0,..., z_p]$, $p \le m$, be linear subspaces of $L_2(D, P)$, where $z_1,..., z_p, z_{p+1},..., z_m$ 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 $g\in V$. Since g is observed as a function that lies in V and on the other hand $W\subseteq V$, then g admits an orthogonal decomposition $g \equiv g_1 \oplus g_2$, where $g_1 \in W$ and $g_2 \in V \cap W^\perp$, with $\langle g_1, g_2\rangle_P = 0$. Hence the problem of testing H_0: $g\in W$ while observing $g\in V$ can be handled by testing the hypotheses:

$$H_0 : g_2 \equiv 0 \; against \; H_1 : g_2 \equiv f_1 \text{, } for\, some\, f_1 \in V \cap W^\perp \tag{2}$$

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 $n_1\geq1$ and $n_2\geq1$. 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 $n_1\geq1$ and $n_2\geq1$. 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:

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

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 $0\leq h_1\leq(a_2-a_1)$ and $0\leq h_2\leq(b_2-b_1)$. The moving sums process of the matrix of observations $Y_{n_2\times n_1}$ consisting of:

$$\left(\lfloor n_1F_1(t+h_1)\rfloor - \lfloor n_1F_1(t)\rfloor\right)\left(\lfloor n_2F_2(s+h_2)\rfloor - \lfloor n_2F_2(s)\rfloor\right)$$

components, is denoted by $MSh_1h_2\left(Y_{n_2\times n_1}\right)(t,s)$, defined as:

$$MSh_1h_2\left(Y_{n_2\times n_1}\right)(t,s) := \frac{1}{\sqrt{n_1 n_2}} \sum_{k=\lfloor n_2F_2(s)\rfloor+1}^{\lfloor n_2F_2(s+h_2)\rfloor} \sum_{\ell=\lfloor n_1F_1(t)\rfloor+1}^{\lfloor n_1F_1(t+h_1)\rfloor} Y_{k\ell} \tag{3}$$

for $(t, s)\in D_{h_1h_2} := [a_1, a_2-h_1] \times [b_1, b_2-h_2]$, where $\lfloor x\rfloor := \max\{z\in\mathbb{Z} : z\leq x\}$. The MOSUM test for the hypothesis H_0: $g_2 \equiv 0$ against H_1: $g_2 \equiv f_1$, for some $f_1\in V\cap W^\perp$, is realized by using the statistics:

$$MI\left(Y^*_{n_2\times n_1}\right) = \frac{1}{\hat{\sigma}_n}\int_D f_1(t,s)\, dMS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right)(t,s)$$

rejecting H_0 for large value of $MI\left(Y^*_{n_2\times n_1}\right)$, where $\hat{\sigma}_n$ a consistent estimator of σ given e.g., in Arnold (1981), pp. 148-149. We note that for every $(t,s)\in D_{h_1h_2}$, $MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right)(t,s)$ can be expressed in the following manner which is useful for the derivation of the main result. Indeed:

$$MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right)(t,s) = \frac{1}{\sqrt{n_1 n_2}} \sum_{k=1}^{n_2}\sum_{\ell=1}^{n_1} 1_{[t,t+h_1]\times[s,s+h_2]}\left(t_{n_1\ell},s_{n_2k}\right) Y_{k\ell}$$

$(t,s)\in D_{h_1h_2}$, meaning that for fixed h_1 and h_2, $MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right)$ can be viewed as a partial sums process indexed by $\{[t, t + h_1] \times [s, s + h_2]: (t,s) \in D_{h_1 h_2}$ of rectangles in D, cf. (Bischoff and Somayasa, 2009; Somayasa *et al*., 2015a). This end can be verified based on the assumption that $F(t, s) = F_1(t, s)F_2(t, s)$ and the continuity of F_1 and F_2 on $[a_1, a_2]$ and $[b_1, b_2]$, respectively. Indeed we have:

$$F_1\left(t_{n_1\lfloor n_1 F_1(t)\rfloor+1}\right)=\frac{\lfloor n_1 F_1(t)\rfloor+1}{n_1}\geq\frac{n_1 F_1(t)}{n_1}=F_1(t)$$

implying $t \leq t_{n_1\lfloor n_1 F_1(t)\rfloor+1}$ by the monotonicity of F_1. Further, since:

$$F_1\left(t_{n_1\lfloor n_1 F_1(t+h_1)\rfloor}\right)=\frac{\lfloor n_1 F_1(t+h_1)\rfloor}{n_1}\leq F_1(t+h_1)$$

we get $t_{n_1\lfloor n_1 F_1(t+h_1)\rfloor} \leq t + h_1$. By the similar argument we can also show that $s \leq s_{n_2\lfloor n_2 F_2(s)\rfloor+1}$ and $s_{n_2\lfloor n_2 F_2(s+h_2)\rfloor} \leq s + h_2$. Hence for fixed $(t,s) \in D_{h_1 h_2}$, the sum in $MS_{h_1 h_2}\left(Y^*_{n_2\times n_1}\right)(t,s)$ involves all $Y_{k\ell}$ with the associated design point $\left(t_{n_1\ell}, s_{n_2 k}\right)$ lies in the rectangle $[t, t]_1 \times [s, s]_2 := [t, t + h_1] \times [s, s + h_2]$.

If the uniform probability measure λ_D^{2*} on D is considered as the design, we have the following formula for every $(t,s) \in D_{h_1 h_2}$:

$$\begin{aligned} MSh_1 h_2\left(Y_{n_2\times n_1}\right)(t,s) &= \frac{1}{\sqrt{n_1 n_2}} \sum_{k=\left\lfloor \frac{n_2(s-b_1)}{(b_2-b_1)}\right\rfloor+1}^{\left\lfloor \frac{n_2(s+h_2-b_1)}{(b_2-b_1)}\right\rfloor} \sum_{\ell=\left\lfloor \frac{n_1(t-a_1)}{(a_2-a_1)}\right\rfloor+1}^{\left\lfloor \frac{n_1(t+h_1-a_1)}{(a_2-a_1)}\right\rfloor} Y_{k\ell} \\ &= \frac{1}{\sqrt{n_1 n_2}} \sum_{k=\left\lfloor \frac{n_2(s-b_1)}{(b_2-b_1)}\right\rfloor+1}^{\left\lfloor \frac{n_2(s-b_1)}{(b_2-b_1)}\right\rfloor+\left\lfloor \frac{n_2 h_2}{(b_2-b_1)}\right\rfloor} \sum_{\ell=\left\lfloor \frac{n_1(t-a_1)}{(a_2-a_1)}\right\rfloor+1}^{\left\lfloor \frac{n_1(t-a_1)}{(a_2-a_1)}\right\rfloor+\left\lfloor \frac{n_1 h_1}{(a_2-a_1)}\right\rfloor} Y_{k\ell} \end{aligned} \tag{4}$$

which consists of $\left(\left\lfloor \frac{n_1 h_1}{(a_2-a_1)}\right\rfloor\right)\left(\left\lfloor \frac{n_2 h_2}{(b_2-b_1)}\right\rfloor\right)$ components of the matrix $Y_{n_2\times n_1}$ of the observations. Hence, by following the terminology introduced in (Bischoff and Gegg, 2014; Chu *et al*., 1995) we call the term h_1h_2 in this study the window size of the process.

By the definition of $MI\left(Y_{n_2\times n_1}\right)$ it can be seen that when the sample supports H_1, the value of $MI\left(Y_{n_2\times n_1}\right)$ will be large as it is contributed by the quantity that asymptotically rational to a positive constant given by $\int_{D_{h_1 h_2}} f_1(t,s)d\zeta_{f_1}(t,s)$, see Theorem 3.1 and Remark 3.2 below. Conversely if the sample comes from H_0, the value of $MI\left(Y_{n_2\times n_1}\right)$ will small as $\int_{D_{h_1 h_2}} f_1(t,s)d\zeta_{f_1}(t,s)$ vanishes uniformly. Hence, rejecting H_0 for large $MI\left(Y_{n_2\times n_1}\right)$ leads us to a conclusion that MOSUM test so defined is a reasonable test.

The next result presents asymptotic size α test for testing (2) based on the statistic $MI\left(Y^*_{n_2\times n_1}\right)$. Let $S_{l;P}$ be the generalized (h_1h_2)-Slepian filed with the parameter space $D_{h_1 h_2}$, see the appendix. It is not too difficult to proof the following result.

Theorem 3.1

Let the matrix of the observations $Y_{n_2\times n_1}$ satisfies Model (1) such that the unknown regression function $g \in L_2(D, P)$ *have an orthogonal decomposition* $g \equiv g_1 \oplus g_2$ *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 α test for H_0: $g_2 \equiv 0$ against H_1: $g_2 \equiv f_1$, for some continuous functions $f_1 \in V \cap W^{\perp}$, will reject H_0 if and only if $MI\left(Y^*_{n_2\times n_1}\right) \geq m_{1-\alpha}$, where $m_{1-\alpha}$ is a constant determined from the distribution of the statistic:

$$I\left(S_{l;P}\right) := \frac{1}{\sigma}\int_{D_{h_1 h_2}} f_1\, d\zeta_{g_1} + \frac{1}{\sigma}\int_{D_{h_1 h_2}} f_1\, d\zeta_{g_2} + \int_{D_{h_1 h_2}} f_1\, dS_{l;P}$$

by the equation $P\{I(S_{l;P}) \geq m_{1-\alpha} | H_0\} = \alpha$. Thereby when H_0 is true, the limit of $MI\left(Y^*_{n_2\times n_1}\right)$ is given by $\frac{1}{\sigma}\int_{D_{h_1 h_2}} f_1\, d\zeta_{g_1} + \int_{D_{h_1 h_2}} f_1\, dS_{l;P}$, where for i = 1, 2, ζ_{g_i} is a function on $D_{h_1 h_2}$, defined as:

$$\zeta_{g_i}(t,s) := \int_{[t,t+h_1]\times[s,s+h_2]} g_i(x,y)P(dx,dy)$$

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

$$G_{\Upsilon_{P;n_1 n_2}}(f) := P\left\{MI\left(Y^*_{n_2 n_1}\right) \geq m_{1-\alpha} \mid g_2 \equiv f\right\}, f \in V \cap W^{\perp}$$

which converges point-wise to G_{Υ_P}, defined by:

$$G_{\Upsilon_P}(f) := P\left\{\int_{D_{h_1 h_2}} f_1\, dS_{\ell;P} \geq m_{1-\alpha} - \frac{1}{\sigma}\int_{D_{h_1 h_2}} f_1\, d\zeta_{g_1+f}\right\}, f \in V \cap W^{\perp}$$

We note that all integrals involved in the definition of $I(S_{l;P})$ are in the sense of Riemann-Stieljes type defined e.g., in Strook (1994).

Remark 3.2

Unfortunately, the probability distribution model of $I(S_{l;P})$ is not tractable for arbitrary probability measure P, by the reason the increments of $S_{l;P}$ is not stochastically independent and the limit $I(S_{l;P})$ still depends on the choice of g_1, unless g_1 and f_1 are orthogonal on $D_{h_1 h_2}$. Therefore, the test is implemented in practice by approximating $m_{1-\alpha}$ by generating Monte Carlo simulation. However, when the design is constructed under the λ_D^{2*} (regular lattice), it can be shown that $I\left(S_{l;\lambda_D^{2*}}\right)$ has independent increment, see Proposition A.6. As pointed out in Proposition A.7, it can be shown that:

$$\int_{D_{h_1 h_2}} f_1 \, dS_{l;\lambda_D^{2*}} \sim N\left(0, 4\int_{D_{h_1 h_2}} f_1^2(t,s)\lambda_D^{2*}(dt,ds)\right)$$

Hence, under H_0, it holds:

$$I\left(S_{l;\lambda_D^{2*}}\right) \sim N\left(\int_{D_{h_1 h_2}} f_1 \, d\zeta_{g_1}, 4\int_{D_{h_1 h_2}} f_1^2(t,s)\lambda_D^{2*}(dt,ds)\right)$$

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]×[2, 3]

			Left Tile Probabilities					
Models	h_1	h_2	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	0.1784	0.1454
	0.05	0.10	0.3961	0.3317	0.2780	0.2163	0.1717	0.1394
	0.05	0.15	0.3995	0.3285	0.2741	0.2136	0.1712	0.1394
	0.05	0.25	0.3859	0.3288	0.2780	0.2136	0.1749	0.1441
	0.15	0.10	0.7076	0.6111	0.5111	0.3944	0.3185	0.2583
	0.15	0.15	0.6994	0.5939	0.5005	0.3876	0.3147	0.2549
	0.15	0.25	0.7053	0.5878	0.4933	0.3826	0.3106	0.2513
	0.20	0.10	0.7914	0.6729	0.5569	0.4433	0.3632	0.2961
	0.20	0.15	0.8028	0.6699	0.5570	0.4320	0.3461	0.2801
	0.20	0.25	0.7606	0.6456	0.5391	0.4230	0.3401	0.2715
	0.25	0.15	0.8627	0.7092	0.6046	0.4731	0.3769	0.3038
	0.25	0.20	0.8469	0.7219	0.6014	0.4718	0.3868	0.3170
	0.25	0.25	0.8430	0.7218	0.6022	0.4620	0.3734	0.3054
1st order	0.05	0.05	1.3817	1.1644	0.9910	0.7621	0.6107	0.4957
	0.05	0.10	1.3281	1.1220	0.9528	0.7399	0.5919	0.4763
	0.05	0.15	1.3093	1.0707	0.8769	0.6862	0.5513	0.4489
	0.05	0.25	1.2394	1.0434	0.8706	0.6861	0.5510	0.4540
	0.15	0.10	2.3136	1.9527	1.6371	1.2710	1.0402	0.8258
	0.15	0.15	2.2480	1.8734	1.5791	1.2190	0.9905	0.7913
	0.15	0.25	2.2690	1.8688	1.5619	1.2087	0.9669	0.7828
	0.20	0.10	2.6455	2.2453	1.8779	1.4240	1.1408	0.9297
	0.20	0.15	2.5909	2.1499	1.8101	1.4051	1.1334	0.9156
	0.20	0.25	2.4980	2.1154	1.7507	1.3345	1.0797	0.8581
	0.25	0.15	2.7584	2.3539	1.9714	1.5372	1.2017	0.9837
	0.25	0.20	2.6930	2.2771	1.9429	1.5069	1.2122	0.9461
	0.25	0.25	2.7318	2.3243	1.9429	1.4797	1.1786	0.9427
2nd order	0.05	0.05	1.5798	1.2735	1.0398	0.7529	0.5651	0.4170
	0.05	0.10	1.5010	1.2493	1.0375	0.7546	0.5587	0.4083
	0.05	0.15	1.4574	1.1982	0.9903	0.7459	0.5597	0.4280
	0.05	0.25	1.4599	1.2139	1.0052	0.7507	0.5727	0.4498
	0.15	0.10	2.4139	1.9833	1.5187	1.0345	0.7012	0.4632
	0.15	0.15	2.3644	1.8953	1.4916	1.0549	0.7296	0.4740
	0.15	0.25	2.3958	1.9212	1.5290	1.0459	0.7556	0.5330
	0.20	0.10	2.7522	2.1992	1.6782	1.0582	0.6975	0.4036
	0.20	0.15	2.5483	2.0170	1.5789	1.0728	0.6967	0.4155
	0.20	0.25	2.6376	2.0509	1.5926	1.0819	0.7365	0.4702
	0.25	0.15	2.8682	2.2099	1.6637	1.0857	0.6623	0.3250
	0.25	0.20	2.8535	2.2716	1.7374	1.1395	0.7370	0.4260
	0.25	0.25	2.8014	2.2198	1.7144	1.1171	0.7506	0.4563

DOI: 10.3844/jmssp.2017.106.126

Thus, it can be stated that an asymptotically size α test will reject H_0, if:

$$MI\left(Y_{n_2\times n_1}\right)\geq \Phi^{-1}(1-\alpha)\sqrt{4\int_{D_{h_1h_2}} f_1^2(t,s)\lambda_D^{2*}(dt,ds)}$$

provided f_1 and g_1 are also orthogonal on $D_{h_1h_2}$ as functions in $L_2\left(D_{h_1h_2},\lambda_D^{2*}\right)$. Under the same condition as before we further get the corresponding power function of the test as given by:

$$G_{\Upsilon_{\lambda_D^{2*};n_1n_2}}(f):=P\left\{MI\left(Y_{n_2\times n_1}\right)\geq \Phi^{-1}(1-\alpha)\sqrt{4\int_{D_{h_1h_2}} f_1^2 d\lambda_D^2}\mid g_2\equiv f\right\}$$

which converges point-wise to the following power function:

$$G_{\Upsilon_{\lambda_D^{2*}}}(f):=1-\Phi\left(\Phi^{-1}(1-\alpha)-\frac{\frac{1}{\sigma}\int_{Dh_1h_2} f_1\,d\zeta_f}{\sqrt{4\int_{Dh_1h_2} f_1^2(t,s)\lambda_D^{2*}(dt,ds)}}\right),f\in V\cap W^{\perp}$$

Although the $m_{1-\alpha}$ 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 $MI\left(Y_{n_2\times n_1}\right)$, then the *p-value* can be obtained by approximating the following probability:

$$p-value=P\left\{MI\left(Y_{n_2\times n_1}\right)\geq m^*\mid H_0\right\}$$

For the design λ_D^{2*} we have:

$$p-value=P\left\{MI\left(Y_{n_2\times n_1}\right)\geq m^*\mid H_0\right\}\approx 1-\Phi\left(\frac{m^*}{\sqrt{4\int_{D_{h_1h_2}} f_1^2(t,s)\lambda_D^{2*}(dt,ds)}}\right)$$

provided $f1$ and $g1$ are orthogonal on $D_{h_1h_2}$. 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 $MI\left(Y_{n_2\times n_1}\right)$.
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 $CU\left(Y_{n_2\times n_1}\right)$, defined by:

$$CU\left(Y_{n_2\times n_1}\right):=\int_D f_1\,(t,s)\,d\frac{1}{\hat{\sigma}_n}S_{n_1n_2}\left(Y_{n_2\times n_1}\right)(t,s)$$

where:

$$S_{n_1n_2}\left(Y_{n_2\times n_1}\right)(t,s):=\frac{1}{\sqrt{n_1n_2}}\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2}1_{[a_1,t]\times[b_1,s]}\left(t_{n_1\ell},s_{n_2k}\right)Y_{k\ell},(t,s)\in D$$

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:

$$S_{n_1n_2}\left(Y_{n_2\times n_1}\right)(B):=\frac{1}{\sqrt{n_1n_2}}\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2}1_B\left(t_{n_1\ell},s_{n_2k}\right)Y_{k\ell},B\in A:=B(D)$$

which is commonly called set-indexed partial sums process, (Alexander and Pyke, 1986; Gaenssler, 1993; Pyke, 1983; Xie and MacNeill, 2006). For each $\omega\in\Omega,S_{n_1n_2}\left(Y_{n_2\times n_1}\right)(\omega)$ constitutes a signed measure on A, therefore the integral involved in the statistic $CU\left(Y_{n_2\times n_1}\right)$ 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 $S_{n_1n_2}$ coincides with $MS_{h_1h_2}$ when $S_{n_1n_2}$ is restricted to the Vapnick-Chervonenkis class (VCC) $\{[t_1, t_2]\times[s_1, s_2]: a_1<t_1, t_2<a_2, b_1<s_1, s_2<b_2\}$ 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 α test for testing (2). The proof is devoted to the appendix.

Theorem 4.1

Let g have an orthogonal decomposition $g \equiv g_1 \oplus g_2$ 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 α test for the hypothesis H_0: $g_2 \equiv 0$ against H_1: $g_2 \equiv f_1$, for some continuous functions $f_1 \in V \cap W^{\perp}$ will reject H_0 if and only if $CU\left(Y_{n_2 \times n_1}\right) \geq \Phi^{-1}(1-\alpha) \| f \|_P$.

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 $\Psi_{P;n_1 n_2}$: $V \cap W^{\perp} \rightarrow (0, 1)$ defined by:

$$\Psi_{P;n_1 n_2}(f) := P\left\{ \begin{array}{l} CU\left(Y^*_{n_2 \times n_1}\right) \\ \geq \Phi^{-1}(1-\alpha) \| f_1 \|_P | g_2 \equiv f \end{array} \right\}, f \in V \cap W^{\perp}$$

be the power function of the test of size α based on the statistic $CU\left(Y^*_{n_2 \times n_1}\right)$ derived above. Then under the assumption of Theorem 4.1 it holds:

$$\lim_{n_1, n_2 \rightarrow \infty} \Psi_{P;n_1 n_2}(f) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{\langle f_1, f \rangle_P}{\| f_1 \|_P} \right) =: \Psi_P(f)$$

where, Φ 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 $CU\left(Y^*_{n_2 \times n_1}\right)$ for arbitrary design P can be computed analytically. In particular the *p-value* of the test can be approximated by the formula:

$$p - value = P\left\{ \frac{1}{\hat{\sigma}_n} CU\left(Y_{n_2 \times n_1}\right) \geq t^* | H_0 \right\} \approx 1 - \Phi\left(\frac{t^*}{\| f_1 \|_P} \right)$$

where, t^* is the observed value of $CU\left(Y^*_{n_2 \times n_1}\right)$. The decision making process is mainly based on the *p-value* instead of computing the value of $\Phi^{-1}(1-\alpha)\|f_1\|_P$.

Simulation Study

In this section we present simulation study to approximate the quantiles of the statistic $MI\left(Y^*_{n_2 \times n_1}\right)$ and to demonstrate the finite sample behavior of the test based on $MI\left(Y^*_{n_2 \times n_1}\right)$ as well as $CU\left(Y_{n_2 \times n_1}\right)$. We visualize the simulation result by scattering the graphs of the power functions of both tests.

Simulating the Quntiles of $MI\left(Y^*_{n_2 \times n_1}\right)$

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 α varies in the interval (0, 1). In this study we approximate the $m_{1-\alpha}$ for α = 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 60×70 observations to approximate the generalized ($h_1 h_2$)-Slepian field $S_{l;P}$. 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 $Y(t, s) = \beta_0 z_0(t, s) + \varepsilon(t, s)$ while we are observing a first order model given by $Y(t, s) = \beta_0 z_0(t, s) + \beta_1 z_1(t, s) + \beta_2 z_2(t, s) + \varepsilon(t, s)$, where $z_0(t, s) = 1$, $z_1(t, s) = t$, $z_2(t, s) = s$, for $(t, s) \in D$, β_0, β_1 and β_2 are unknown parameters. By a re-parametrization, the model can be represented as:

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

where, $\tilde{z}_0(t,s) := 1$, $\tilde{z}_1(t,s) := t - 2\ln(2)$ and $\tilde{z}_2(t,s) = s - 6\ln(3/2)$ are orthogonal in $L_2(D, P)$ and γ_0, γ_1, γ_3 are unknown parameters. The observations are generated from a constant model defined as:

$$Y\left(t_{n_1 \ell}, s_{n_2 k}\right) = \tilde{z}_0\left(t_{n_1 \ell}, s_{n_2 k}\right) + \varepsilon\left(t_{n_1 \ell}, s_{n_2 k}\right), 1 \leq \ell \leq 60, 1 \leq k \leq 70$$

In this simulation g_2 is hypothesized as zero under H_0 and $g_2 \equiv f_1 \equiv \tilde{z}_3 + \tilde{z}_4$ under H_1. Hence, the critical region is constructed using the test statistic:

$$MI\left(Y^*_{n_2 \times n_1}\right) = \int_{D_{h_1 h_2}} \left(\tilde{z}_1 + \tilde{z}_2\right)(t,s)\, d \frac{1}{\hat{\sigma}_n} MS_{h_1 h_2}\left(Y^*_{n_2 \times n_1}\right)(t,s)$$

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

First Order Model

In the second case we give approximation to the (1-α)-quantiles of the statistic $MI\left(Y^*_{n_2 \times n_1}\right)$ for testing first order model, that is we test:

$$H_0 : Y(t,s) = \beta_0 z_0(t,s) + \beta_1 z_1(t,s) + \beta_2 z_2(t,s) + \varepsilon(t,s)$$

while we are observing a second order model:

$$Y(t,s) = \beta_0 z_0(t,s) + \beta_1 z_1(t,s) + \beta_2 z_2(t,s) + \beta_3 z_3(t,s) + \beta_4 z_4(t,s) + \varepsilon(t,s)$$

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 Grahm-Schmidt orthogonalization procedure, the model can be written by introducing new parameter system as:

$$Y(t,s) = \gamma_0 \tilde{z}_0(t,s) + \gamma_1 \tilde{z}_1(t,s) + \gamma_2 \tilde{z}_2(t,s) + \gamma_3 \tilde{z}_3(t,s) + \gamma_4 \tilde{z}_4(t,s) + \varepsilon(t,s)$$

where, γ_0, γ_1, γ_2, γ_3, γ_0 are unknown parameters and:

$$\tilde{z}_0(t,s) = 1, \tilde{z}_1(t,s) = t - 2\ln(2), \tilde{z}_2(t,s) = s - 6\ln(3/2)$$
$$\tilde{z}_3(t,s) = t^2 + [4\ln(2) - 3]t + (6\ln(2) - 8\ln^2(2) - 2)$$
$$\tilde{z}_4(t,s) = s^2 + [36\ln(3/2) - 15]s - (6\ln(3/2)[36\ln(3/2) - 15] + 6)$$

The hypothesis is tested by proposing a statistic:

$$MI\left(Y^*_{n_2 \times n_1}\right) = \int_{D_{h_1 h_2}} (\tilde{z}_3 + \tilde{z}_4)(t,s)\, d\frac{1}{\hat{\sigma}_n} MS_{h_1 h_2}\left(Y^*_{n_2 \times n_1}\right)(t,s)$$

By generating the observations under H_0 using the model:

$$Y\left(t_{n_1 \ell}, s_{n_2 k}\right) = \tilde{z}_0\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_1\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_2\left(t_{n_1 \ell}, s_{n_2 k}\right) + \varepsilon\left(t_{n_1 \ell}, s_{n_2 k}\right)$$

for $1 \le \ell \le 60$, $1 \le k \le 70$, we present the quantiles of $MI\left(Y^*_{n_2 \times n_1}\right)$ under H_0 in Table 2 for several combinations of h_1 and h_2.

Second Order Model

For the last case we simulate the (1-α)-quantiles of the statistic $MI\left(Y^*_{n_2 \times n_1}\right)$ for testing the hypothesis H_0 that states a second order model is adequate versus H_1 that assumes a third order model:

$$Y(t,s) = \gamma_0 \tilde{z}_0(t,s) + \gamma_1 \tilde{z}_1(t,s) + \gamma_2 \tilde{z}_2(t,s) + \gamma_3 \tilde{z}_3(t,s) + \gamma_4 \tilde{z}_4(t,s) + \gamma_5 \tilde{z}_5(t,s) + \varepsilon(t,s)$$

where, γ_0, γ_1, γ_2, γ_3, γ_4, γ_5 are unknown constants with:

$$\tilde{z}_0(t,s) = 1, \tilde{z}_1(t,s) = t - 2\ln(2), \tilde{z}_2(t,s) = s - 6\ln(3/2),$$
$$\tilde{z}_3(t,s) = t^2 + [4\ln(2) - 3]t + (6\ln(2) - 8\ln^2(2) - 2),$$
$$\tilde{z}_4(t,s) = s^2 + [36\ln(3/2) - 15]s - (6\ln(3/2)[36\ln(3/2) - 15] + 6),$$
$$\tilde{z}_5(t,s) = t^3 - 1.38452t^2 - 0.19293t + 1.03650$$

We present the simulation results for a choice $f_1 \equiv z_5$ in Table 2 by generating the observations under $H0$ from the model:

$$Y\left(t_{n_1 \ell}, s_{n_2 k}\right) = \tilde{z}_0\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_1\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_2\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_3\left(t_{n_1 \ell}, s_{n_2 k}\right) + \tilde{z}_4\left(t_{n_1 \ell}, s_{n_2 k}\right) + \varepsilon\left(t_{n_1 \ell}, s_{n_2 k}\right), 1 \le \ell \le 60, 1 \le k \le 70$$

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 λ_D^{2*}. Our purpose is to demonstrate that the finite sample power function $\Psi_{P;n_1 n_2}(f)$ lies closed enough to its limit $\Psi_P(f)$ as f varies in $V \cap W^{\perp}$ 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 $\|f_1\|_P = 0.37045$ for $f_1 \equiv \tilde{z}_1 + \tilde{z}_2$, so that $\Phi^{-1}(1-\alpha)\|f_1\|_P = 0.60934$, when $\alpha = 0.05$. In order to make f varies in $V \cap W^{\perp}$, we define $f \equiv \lambda f_1$, for $\lambda \in \mathbb{R}$ and generate the samples independently from the normal distribution with mean $\frac{1}{\sqrt{n_1 n_2}}\left(g_1\left(\Xi_{n_1 \times n_2}\right) + \lambda f_1\left(\Xi_{n_1 \times n_2}\right)\right)$ and the variance σ^2 which is assumed to be unknown. In this case we estimate σ^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: $g \in V = [\tilde{z}_0, \tilde{z}_1, \tilde{z}_2]$ holds true. In this simulation we develop the graphs of $\Psi_{P;n_1 n_2}(\lambda f_1)$ for $\alpha = 0.05$, where:

$$\Psi_{P;n_1 n_2}(\lambda f_1) = P\left\{\frac{1}{\hat{\sigma}_n}\int_D f_1\, dS_{n_1 n_2}\left(Y^*_{n_1 \times n_2}\right) \ge 0.60934\right\}$$

together with its point-wise limit $\Psi_P(\lambda f_1) = 1-\Phi(1.64485-0.37045\lambda)$, where we chose the VCC $\{[1, t] \times [2, s]: (t, s) \in D\}$ instead of the much larger family A_0 as the index sets. The simulation results are exhibited in

Fig. 2 for λ varies in the closed interval [0, 15] presented in two panels (*a*) and (*b*) associated with the sample sizes 60×65 and 70×75, respectively. It can be seen that for constant model the curves of $\Psi_{P;n_1 n_2}$ approximates well those of the limit Ψ_P achieving the size $\alpha = 0.05$ at $\lambda = 0$ as they should be.

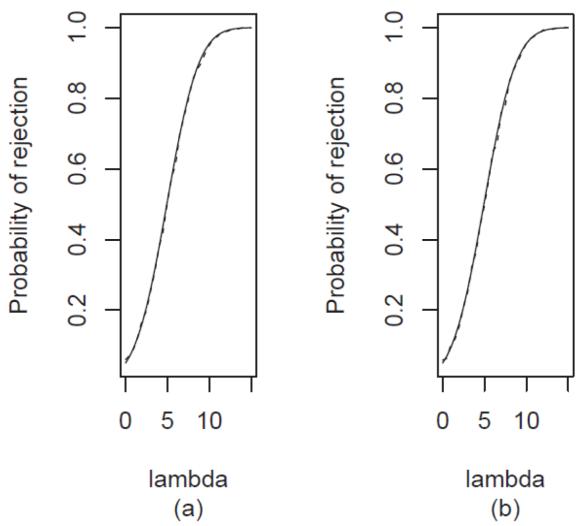


Fig. 2. The graphs of the empirical power functions of the asymptotic size $\alpha = 0.05$ CUSUM test for constant model represented by using dotted line approximated by the limit power function (Ψ_P (λf_1)) scattered using smooth lines. The design points are generated using the CDF $F(t, s) = 12(1\text{-}1/t)(1/2\text{-}1/s)$ on D

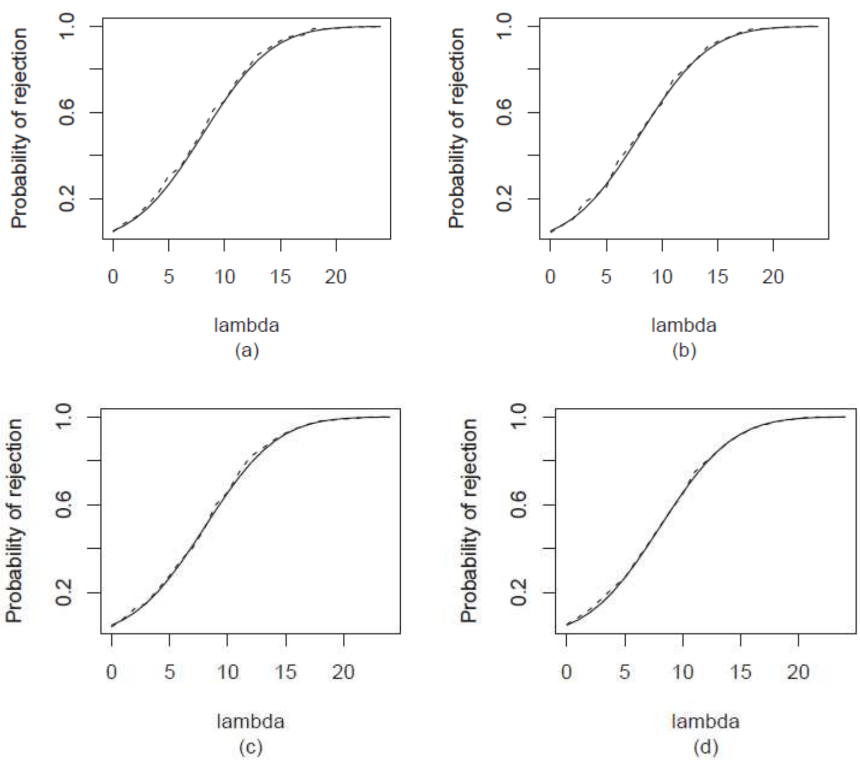


Fig. 3. The graphs of the empirical power function of the asymptotic size $\alpha = 0.05$ CUSUM test for constant model under the design generated using the scaled Lebesgue measure λ_D^{2*} (regular lattice) scattered by using dotted lines approximated by the limit power function $\Psi_{\lambda_D^{2*}}\left(\lambda f_1\right)$ 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	$\tilde{\sigma}_n$
Constant	MOSUM			0.20780
	$h_1 = 0.20$; $h_2 = 0.25$	3.56799	6.2157e-6	
	$h_1 = 0.25$; $h_2 = 0.25$	3.10043	0.00007	
	CUSUM	1.71874	1.2766e-5	
	KS	1.66759	0.03210	
	CvM	0.48443	0.01240	
First order	MOSUM			0.20163
	$h_1 = 0.20$; $h_2 = 0.25$	1.65455	0.01668	
	$h_1 = 0.25$; $h_2 = 0.25$	1.46001	0.03022	
	CUSUM	1.06940	0.00297	
	KS	0.89523	0.42660	
	CvM	0.09341	0.20460	
Second order	MOSUM			0.19759
	$h_1 = 0.20$; $h_2 = 0.25$	1.94889	0.16492	
	$h_1 = 0.25$; $h_2 = 0.25$	1.28138	0.26086	
	CUSUM	2.60534	0.00459	
	KS	0.92080	0.10380	
	CvM	0.09080	0.01240	

Figure 3 illustrates the empirical power function of the CUSUM test for constant model when the design points are constructed using the uniform measure λ_D^{2*}. The observations are generated from the model:

$$Y^*\left(t_{n_1\ell}, s_{n_2k}\right) = \frac{1}{\sqrt{n_1 n_2}}\left(\tilde{\omega}_0\left(t_{n_1\ell}, s_{n_2k}\right) + \lambda f_1\left(t_{n_1\ell}, s_{n_2k}\right)\right) + \varepsilon\left(t_{n_1\ell}, s_{n_2k}\right)$$

for testing constant model using the statistic $\frac{1}{\hat{\sigma}_n}\int_D f_1\, dS_{n_1 n_2}\left(Y^*_{n_2\times n_1}\right)$, where for $(t, s)\in D$, $\tilde{\omega}_0(t,s) = 1$, $f_1(t, s) = \tilde{\omega}_1(t,s) + \tilde{\omega}_2(t,s)$, with $\tilde{\omega}_1(t,s) = t\text{-}3/2$ and $\tilde{\omega}_2(t,s) = s\text{-}5/2$. The set $\{\tilde{\omega}_0, \tilde{\omega}_1, \tilde{\omega}_2\}$ builds orthogonal basis of $V\subset L_2\left(D,\lambda_D^{2*}\right)$. In this case $t_{n_1\ell} = \ell/n_1 + 1$ and $s_{n_2k} = k/n_2 + 2$, for $1 \le \ell \le n_1$ and $1\le k \le n_2$. The graphs of the finite sample power function $\Psi_{\lambda_D^{2*};n_1 n_2}\left(\lambda f_1\right)$ together with the point-wise limit $\Psi_{\lambda_D^{2*}}\left(\lambda f_1\right)$, for $\alpha = 0.05$ and λ varies in [0, 14] are exhibited in Fig. 3 in four panels: (a), (b), (c) and (d) associated with the sample sizes 40×50, 60×65, 70×75 and 80×85, respectively, where:

$$\Psi_{\lambda_D^{2*};n_1 n_2}\left(\lambda f_1\right) = P\left\{\frac{1}{\hat{\sigma}_n}\int_D f_1\, dS_{n_1 n_2}\left(Y^*_{n_2\times n_1}\right) \ge 0.67151\right\}$$
$$\Psi_{\lambda_D^{2*}}\left(\lambda f_1\right) = 1 - \Phi(1.64485 - 0.40825\lambda)$$

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 Ψ_P as well as $\Psi_{\lambda_D^{2*}}$ give very good approximation to $\Psi_{P;n_1 n_2}$ and $\Psi_{\lambda_D^{2*};n_1 n_2}$, 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 $\frac{1}{\hat{\sigma}_n}\int_D f_1\, dS_{n_1 n_2}\left(Y^*_{n_2\times n_1}\right)$, where the observations are generated independently from the normal distribution with mean given by:

$$\frac{1}{\sqrt{n_1 n_2}}\left(\tilde{g}_1\left(t_{n_1\ell}, s_{n_2k}\right) + \lambda f_1\left(t_{n_1\ell}, s_{n_2k}\right)\right)$$

and unknown variance σ^2. In this case $\tilde{g}_1 \equiv \tilde{z}_0 + \tilde{z}_1 + \tilde{z}_2 \in W$ and $f_1 \equiv \tilde{z}_3 + \tilde{z}_4 \in V\cap W^\perp$. That is under H_0 we assume that a first order model is adequate, while we are observing a second order model. Since $\|f_1\|_P = 1.50177$, then for $\alpha = 10\%$, we have:

$$\Psi_{P;n_1 n_2}\left(\lambda f_1\right) = P\left\{\frac{1}{\hat{\sigma}_n}\int_D f_1\, dS_{n_1 n_2}\left(Y^*_{n_2\times n_1}\right) \ge 1.92459\right\}$$
$$\Psi_{\lambda_D^{2*}}\left(\lambda f_1\right) = 1 - \Phi(1.28155 - 1.50177\lambda)$$

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 $I^2 := [0, 1] \times [0, 1]$. Then the corresponding orthogonal version of the regression functions $\{z_0, z_1, z_2, z_3, z_4, z_5\}$ are given by:

$$\tilde{w}_0(t,s)=1, \tilde{w}_1(t,s)=t-1/2, \tilde{w}_2(t,s)=s-1/2,$$
$$\tilde{w}_3(t,s)=t^2-t/12-7/24, \tilde{w}_4(t,s)=s^2-s/12-7/24$$

We test $H_0: g\in W:=[\tilde{w}_0,\tilde{w}_1,\tilde{w}_2]$ against $H_1: g\in V:=[\tilde{w}_0,\tilde{w}_1,\tilde{w}_2,\tilde{w}_3,\tilde{w}_4]$ by the statistic $\frac{1}{\hat{\sigma}_n}\int_{I^2} f_1\, dS_{n_1 n_2}\left(Y^*_{n_2\times n_1}\right)$, where the observations are generated independently based on the normal distribution with mean:

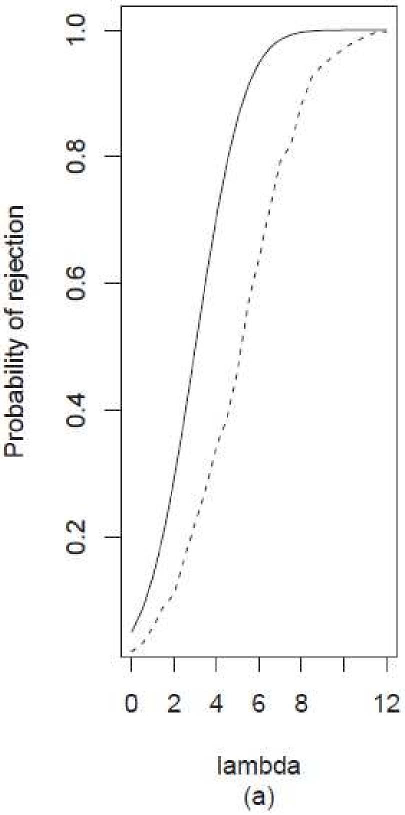
$$\frac{1}{\sqrt{n_1 n_2}}\left(g_1\left(\ell/n_1, k/n_2\right)+\lambda f_1\left(\ell/n_1, k/n_2\right)\right)$$

and unknown variance σ^2, thereby $g_1\equiv\tilde{w}_0+\tilde{w}_1+\tilde{w}_2\in W$ and $f_1\equiv\tilde{w}_3+\tilde{w}_4\in V\cap W^{\perp}$. In this case we get after little computation $\|f_1\|_{\lambda^2_{I^2}}=0.38879$ and there for $\Phi^{-1}(0.90)\|f_1\|_{\lambda^2_{I^2}}=0.49825$, giving the powers:

$$\Psi_{\lambda^2_{I^2};n_1n_2}(\lambda f_1)=P\left\{\frac{1}{\hat{\sigma}_n}\int_{I^2} dS_{n_1n_2}\left(Y^*_{n_2\times n_1}\right)\geq 0.49825\right\}$$
$$\Psi_{\lambda^2_{I^2}}(\lambda f_1)=1-\Phi(1.28155-0.38879\lambda)$$

evaluated at $\lambda f_1\in V\cap W^{\perp}$, for $\lambda\in R$.

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 λ 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 $CU\left(Y_{n_2\times n_1}\right)$ under H_0 which is given by $\Phi^{-1}(1-\alpha)\|f_1\|_P$, for a given $\alpha\in(0, 1)$ and $f_1\in V\cap W^{\perp}$.



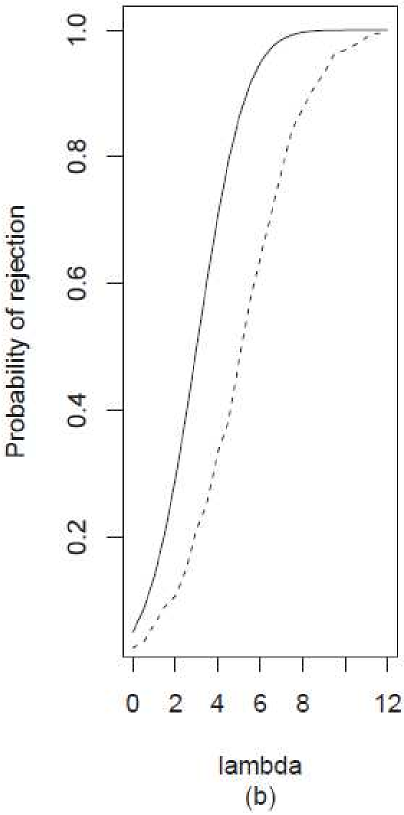


Fig. 4. The graphs of the empirical power functions of the asymptotic size α = 10% CUSUM test for first order model under the design generated using the CDF $F(t, s) = 12(1-1/t)(1/2-1/s)$ for $(t, s)\in D$ represented by using dotted line approximated by the limiting power function $\Psi_P(\lambda f_1)$ scattered using smooth lines. (a) Sample size = 60×65, (b) Sample size = 70×75

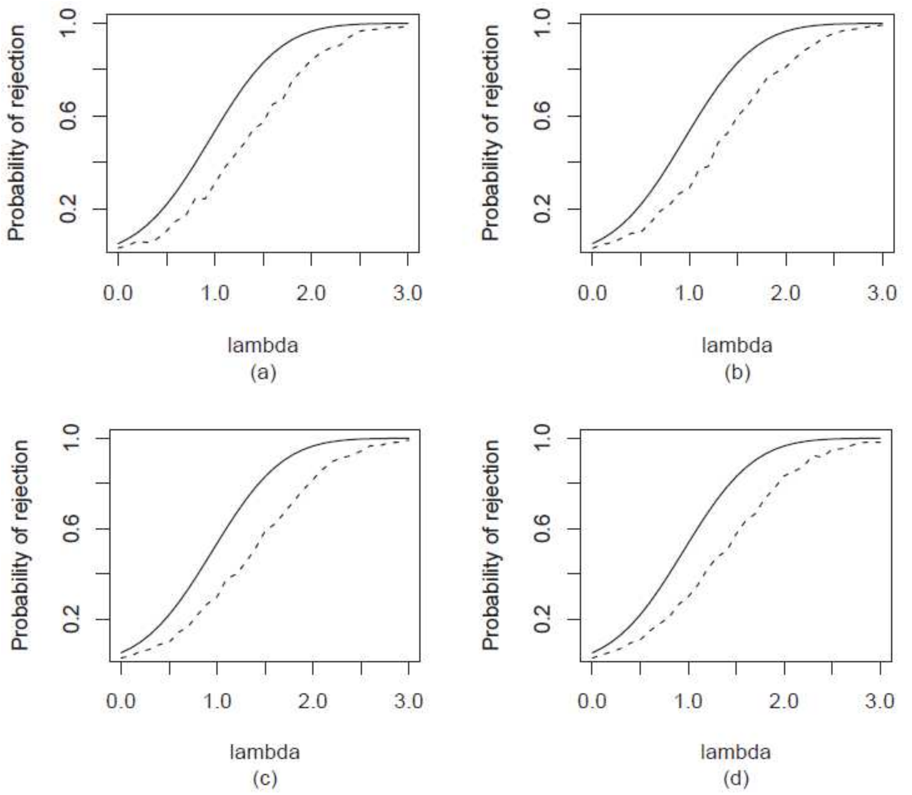


Fig. 5. The graphs of the empirical power functions of the asymptotic size $\alpha = 10\%$ CUSUM test for first order model represented by using dotted lines approximated by the limit power function $\Psi \lambda_{I^2}^2 \left(\lambda f_1\right)$ scattered using smooth lines. The design is $n_1 \times n_1$ regular lattice: (a) 50×55, (b) 60×65, (c) 70×75 and (d) 80×85

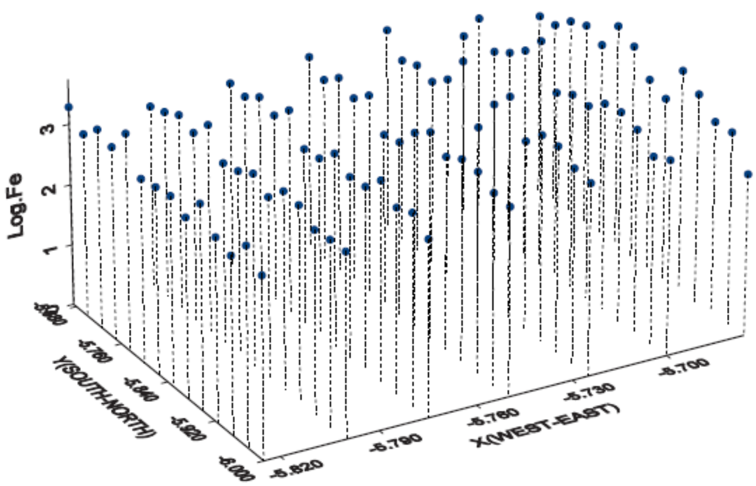


Fig. 6. The three dimensional scatter plot of the trend surface of the logarithm of the percentage of Fe. Source of data: Tahir (2010)

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 $\{[0, t] \times [0,s]: 0 \le t, s \le 1\}$ 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 $\tilde{\sigma}_n$ 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 $\alpha \le 1.668\%$ and $\alpha \le 0.297\%$, 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 α, 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 α set less then or equal 0.459 and 1.24%, respectively, we recommend that first order model is the most appropriate for the LogFe.

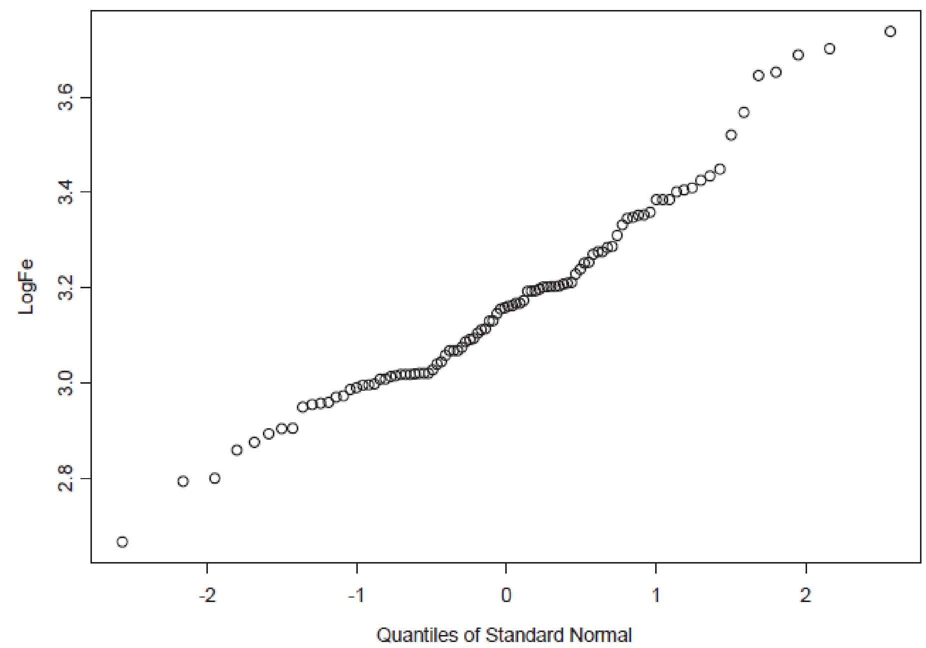


Fig. 7. The qqnorm plot for LogFe. Two sided Kolmogorov-Smirnov for normality assign the critical value 0:0933 and the *p*-value 0.0352. Source of data: Tahir (2010)

Concluding Remark

We established an asymptotically size α 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_1h_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 regrading 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: 0471197459, 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-sum 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. Engkoimani, 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. Ngkoimani, 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., BirkhÄauser, 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 based on the results of drilling bores on several points (case study: South mining region of PT. Aneka Tambang Tbk., Pomalaa, Southeast Sulawesi). Research Report, Halu Oleo University, Kendari.

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_1h_2)-Slepian Field

Definition A.1

For i = 1, 2, let h_i with $0 < h_1 < (a_2\text{-}a_1)$ and $0 < h_2 < (b_2\text{-}b_1)$ be positive real numbers. A process $S_{l;P} := \{S_{l;P}(t, s): (t, s) \in D_{h_1h_2}\}$ is called a generalized (h_1h_2-)Slepian filed, if and only if $S_{l;P}$ is a centered Gaussian process with the covariance function:

$$\gamma_{S_{l;P}}\left((t_1, s_1), (t_2, s_2)\right) = P\left(\left[t_1, t_1\right]_1 \times \left[s_1, s_1\right]_2 \cap \left[t_2, t_2\right]_1 \times \left[s_2, s_2\right]_2\right)$$

for $(t_1, s_1), (t_2, s_2) \in D_{h_1h_2} := [a_1, a_2\text{-}h_1] \times [b_1, b_2\text{-}h_2]$, where $[x, x]_1$ and $[y, y]_2$ stand for $[x, x + h_1]$ and $[y, y+h_2]$, respectively. In the case where P is the Lebesgue measure λ^2 on D, we call $S_{l;P}$ a generalized (h_1h_2-)Slepian filed, denoted by $S_{l;\lambda^2}$, with the covariance function:

$$\gamma_{S_{l;\lambda^2}}\left((t_1, s_1), (t_2, s_2)\right) = \left(h_1 - |t_1 - t_2|\right)^+ \left(h_2 - |s_1 - s_2|\right)^+$$

where, $x^+ := \max\{x, 0\}$, 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 $S_{l;P}$ can be summarized as follows:

- For every $(t, s) \in D_{h_1h_2}$, $S_{l;P}(t, s) \sim N(0, P([t, t]_1 \times [s, s]_2)$
- Let $W_P(t, s) := W_P([a_1, t] \times [b_1, s])$. For fixed h_1, h_2 defined above, let:

$$\Delta W_P(x, y) := W_P(x + h_1, y + h_2) - W_P(x, y + h_2) - W_P(x + h_1, y) + W_P(x, y)$$

Since $\Delta W_P(x, y)$ is also distributed as $N(0, P([t, t]_1 \times [s, s]_2)$, then $S_{l:P}$ is said to be equivalent in distribution to ΔW_P, *denoted by* $S_{l;P} \overset{D}{=} \Delta W_P$.

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

Let $\{\varepsilon_{n_2 \times n_1}\}_{n_1 \geq 1; n_2 \geq 1}$ be a sequence of independent and identically distributed random matrices, where $\varepsilon_{n_2 \times n_1} := (\varepsilon_{k\ell})_{\ell,k=1}^{n_1,n_2}$, $\varepsilon_{k\ell} := \varepsilon(t_{n_1\ell}, s_{n_2k})$, with $E(\varepsilon_{k\ell}) = 0$ and $Var(\varepsilon_{k\ell}) = \sigma^2 < \infty$, for $1 \leq \ell \leq n_1$ and $1 \leq k \leq n_2$. Let the experimental design $\Xi_{n_1 \times n_2}$ is constructed using a probability measure P on $D = [a_1, a_2] \times [b_1, b_2]$, such that $P = P_1 \times P_2$ 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_2-1, we have:

$$\frac{1}{\sigma} MS_{h_1h_2}\left(\varepsilon_{n_2 \times n_1}\right) \Rightarrow S_{l;P}$$

Proof

By applying the well-known Prohorov's theorem (cf. Billingsley (1999), pp. 35-40), we show first that the finite dimensional distribution of $\frac{1}{\sigma}MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)$ converges to that of $S_{l;P}$. That is for any points $(t_1, s_1), (t_2, s_2),..., (t_m, s_m)$ in D and any constants $b_1,...,b_m$, we show that the sequence $\left\{C_{n_1n_2}\right\}$, for $n_1\geq1$, $n_2\geq1$, where $C_{n_1n_2} := \sum_{i=1}^{m} b_i \frac{1}{\sigma}MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)(t_i, s_i)$ converges in distribution to $\sum_{i=1}^{m} b_i S_{l;P}(t_i, s_i)$. The last is normally distributed with mean zero and variance $\sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j\ P([t_i, t_i]_1 \times [s_i, s_i]_2 \cap [t_j, t_j]_1 \times [s_j, s_j]_2)$. By the definition of $MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)$, we get:

$$
\begin{aligned}
&Var\left(C_{n_1n_2}\right)\\
&=\sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j \frac{1}{\sigma^2} E\left(MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)(t_i, s_i) MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)(t_j, s_j)\right)\\
&=\sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j \sum_{k=\lfloor n_2F_2(s_i)\rfloor+1}^{\lfloor n_2F_2(s_i+h_2)\rfloor}\ \sum_{\ell=\lfloor n_1F_1(t_i)\rfloor+1}^{\lfloor n_1F_1(t_i+h_1)\rfloor}\ \sum_{k'=\lfloor n_2F_2(s_j)\rfloor+1}^{\lfloor n_2F_2(s_j+h_2)\rfloor}\ \sum_{\ell'=\lfloor n_1F_1(t_j)\rfloor+1}^{\lfloor n_1F_1(t_j+h_1)\rfloor} \frac{E\left(\varepsilon_{k\ell}\varepsilon_{k'\ell'}\right)}{\sigma n_1 n_2}
\end{aligned}
$$

Furthermore, by the independence of $\varepsilon_{k\ell}$, we have:

$$
\begin{aligned}
&\sum_{k=\lfloor n_2F_2(s_i)\rfloor+1}^{\lfloor n_2F_2(s_i+h_2)\rfloor}\ \sum_{\ell=\lfloor n_1F_1(t_i)\rfloor+1}^{\lfloor n_1F_1(t_i+h_1)\rfloor}\ \sum_{k'=\lfloor n_2F_2(s_j)\rfloor+1}^{\lfloor n_2F_2(s_j+h_2)\rfloor}\ \sum_{\ell'=\lfloor n_1F_1(t_j)\rfloor+1}^{\lfloor n_1F_1(t_j+h_1)\rfloor} \frac{E\left(\varepsilon_{k\ell}\varepsilon_{k'\ell'}\right)}{\sigma n_1 n_2}\\
&=\left(\begin{array}{l}\min\left\{\frac{\lfloor n_2F_2(s_i+h_2)\rfloor}{n_2}, \frac{\lfloor n_2F_2(s_j+h_2)\rfloor}{n_2}\right\}\\ -\max\left\{\frac{\lfloor n_2F_2(s_i)\rfloor}{n_2}, \frac{\lfloor n_2F_2(s_j)\rfloor}{n_2}\right\}\end{array}\right)\\
&\times\left(\begin{array}{l}\min\left\{\frac{\lfloor n_1F_1(t_i+h_1)\rfloor}{n_1}, \frac{\lfloor n_1F_1(t_j+h_1)\rfloor}{n_1}\right\}\\ -\max\left\{\frac{\lfloor n_1F_1(t_i)\rfloor}{n_1}, \frac{\lfloor n_1F_1(t_j)\rfloor}{n_1}\right\}\end{array}\right)
\end{aligned}
$$

which converges as $n_1, n_2\to\infty$, to:

$$
\begin{aligned}
&\left(\min\left\{F_2(s_i+h_2), F_2(s_j+h_2)\right\}-\max\left\{F_2(s_i), F_2(s_j)\right\}\right)\\
&\times\left(\min\left\{F_1(t_i+h_1), F_1(t_j+h_1)\right\}-\max\left\{F_1(t_i), F_1(t_j)\right\}\right)\\
&=P_2\left(\left[s_i, s_i\right]_2\cap\left[s_j, s_j\right]_2\right)P_1\left(\left[t_i, t_i\right]_1\cap\left[t_j, t_j\right]_1\right)\\
&\times P\left(\left[t_i, t_i\right]_1\times\left[s_i, s_i\right]_2\cap\left[t_j, t_j\right]_1\times\left[s_j, s_j\right]_2\right)\\
&=\left(F_2\left(\min\left\{s_i+h_2, s_j+h_2\right\}\right)-F_2\left(\max\left\{s_i, s_j\right\}\right)\right)\\
&\times\left(F_1\left(\min\left\{t_i+h_1, t_j+h_1\right\}\right)-F_1\left(\max\left\{t_i, t_j\right\}\right)\right)
\end{aligned}
$$

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 $n_1, n_2\to\infty$:

$$
\begin{aligned}
&Var\left(C_{n_1n_2}\right)\\
&\to\sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j P\left(\left[t_i, t_i\right]_1\times\left[s_i, s_i\right]_2\cap\left[t_j, t_j\right]_1\times\left[s_j, s_j\right]_2\right)
\end{aligned}
$$

which is the variance of $\sum_{i=1}^{m} b_i S_{l;P}(t_i, s_i)$. Next we show that for every $\in>0$, it holds $\lim_{n_1, n_2\to\infty} L(\in) = 0$, where:

$$
\begin{aligned}
&L(\in)\\
&:=\sum_{i=1}^{m} b_i \sum_{k=\lfloor n_2F_2(s_i)\rfloor+1}^{\lfloor n_2F_2(s_i+h_2)\rfloor}\ \sum_{\ell=\lfloor n_1F_1(t_i)\rfloor+1}^{\lfloor n_1F_1(t_i+h_1)\rfloor} E\left(\left|\frac{1}{\sigma\sqrt{n_1n_2}}\varepsilon_{k\ell}\right|^2\times 1_{\left\{\left|\frac{1}{\sigma\sqrt{n_1n_2}}\varepsilon_{k\ell}\right|>\in\right\}}\right)
\end{aligned}
$$

Let $K:= \max_{1\leq i\leq m} |b_i|$, since $\varepsilon_{k\ell}$ is independent and identically distributed, then we get:

$$
\begin{aligned}
&0<\lim_{n_1, n_2\to\infty} L(\in)\leq \lim_{n_1, n_2\to\infty}\frac{mK}{\sigma^2}\left(\frac{\lfloor n_2F_2(s_i+h_2)\rfloor}{n_2}-\frac{\lfloor n_2F_2(s_i)\rfloor}{n_2}\right)\\
&\times\left(\frac{\lfloor n_1F_1(t_i+h_1)\rfloor}{n_1}-\frac{\lfloor n_1F_1(t_i)\rfloor}{n_1}\right)E\left(|\varepsilon_{11}|^2\times 1_{\left\{|\varepsilon_{11}|>\in\sigma\sqrt{n_1n_2}\right\}}\right)\\
&=\frac{mK}{\sigma^2}P\left(\left[t_i, t_i\right]_1\times\left[s_i, s_i\right]_2\right)\lim_{n_1, n_2\to\infty}E\left(|\varepsilon_{11}|^2\times 1_{\left\{|\varepsilon_{11}|>\in\sigma\sqrt{n_1n_2}\right\}}\right)=0
\end{aligned}
$$

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 $C_{n_1n_2}$ converges in distribution to $\sum_{i=1}^{m} b_i S_{l;P}(t_i, s_i)$. 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:

$$
MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)(t, s) = S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)\left([t, t]_1\times[s, s]_2\right)
$$

that is the MOSUM process coincides with the CUSUM process when the index sets are restricted to the VCC $\{[t, t]_1 \times [s, s]_2: (t, s)\in D\}$ of subsets of D. Due to the results of Alexander and Pyke (1986) and Pyke (1983), the tightness of $MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)$

immediately follows from the tightness of $S_{n_1 n_2}\left(\varepsilon_{n_2\times n_1}\right)$, establishing the proof.

Definition A.4

Let $\Gamma_1 := \left\{[t_0,t_1],[t_1,t_2],\cdots,\left[t_{(m_1-1)},t_{m_1}\right]\right\}$ be a set of m1 closed intervals on [a_1, a_2], such that $a_1 = t_0 < t_1 < t_2 < \cdots < t_{m_1} = a_2$. Let Γ_2 := $\left\{[s_0,s_1],[s_1,s_2],\cdots,\left[s_{(m_2-1)},s_{m_2}\right]\right\}$ be a set of m_2 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_1m_2 closed rectangles which builds a partition on D. For $1\le w_i \le m_i$, with i = 1, 2, let $J_{w_1 w_2}$ be the element of K defined by $J_{w_1 w_2} := \left[t_{w_i-1},t_{w_i}\right]\times\left[s_{w_i-1},s_{w_i}\right]$. The increment of $S_{l;P}$ on $J_{w_1 w_2}$ is denoted by $\Delta_{w_1 w_2} S_{l;P}$, given by:

$$\begin{aligned}\Delta_{w_1 w_2} S_{l;P} &:= S_{l;P}\left(t_{w_1},s_{w_2}\right) - S_{l;P}\left(t_{w_1-1},s_{w_2}\right)\\ &-S_{l;P}\left(t_{w_1},s_{w_2-1}\right) + S_{l;P}\left(t_{w_1-1},s_{w_2-1}\right)\end{aligned}$$

Remark A.5

Let $\left[t_{w_i},t_{w_i+1}\right]$ × $\left[s_{w_i},s_{w_i+1}\right]$ and $\left[t_{w_i'},t_{w_i'+1}\right]$ $\times\left[s_{w_i'},s_{w_i'+1}\right]$ be arbitrary disjoint rectangles on D. Then $Cov\left(\Delta_{w_1+1w_2+1}S_{l;P},\Delta_{w_1'+1w_2'+1}S_{l;P}\right)$ is computed by the following formula:

$$\begin{aligned}
&Cov\left(\Delta_{w_1+1w_2+1}S_{l;P},\Delta_{w_1'+1w_2'+1}S_{l;P}\right)\\
&= P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2+1},s_{w_2+1}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2}),s_{w_2})\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2}),s_{w_2})\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2}),s_{w_2})\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1+1},t_{w_1+1}\right]_1\times\left[s_{w_2}),s_{w_2})\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2},s_{w_2}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2},s_{w_2}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2},s_{w_2}\right]_2\cap\left[t_{w_1'+1},t_{w_1'+1}\right]_1\times\left[s_{w_2'},s_{w_2'}\right]_2\right)\\
&-P\left(\left[t_{w_1},t_{w_1}\right]_1\times\left[s_{w_2},s_{w_2}\right]_2\cap\left[t_{w_1'},t_{w_1'}\right]_1\times\left[s_{w_2'+1},s_{w_2'+1}\right]_2\right)
\end{aligned}$$

Proposition A.6

The ordinary (h_1h_2)-Slepian field $S_{l;\lambda^2}$ is a centered Gaussian process with independent increments.

Proof

Let $\left[t_{w_i},t_{w_i+1}\right]$ × $\left[s_{w_i},s_{w_i+1}\right]$ and $\left[t_{w_i'},t_{w_i'+1}\right]$ $\times\left[s_{w_i'},s_{w_i'+1}\right]$ be arbitrary disjoint rectangles on D. We assume without loss of generality that $t_{w_1+1}\le t_{w_1'}$, $s_{w_2+1}\le s_{w_2'}$ and $0\le w_i, w_i'\ m_i - 1$. In addition we need to set $\left|t_{w_1+1}\le t_{w_1'}\right| < h_1$ and $\left|s_{w_2+1}\le s_{w_2'}\right| < h_2$. By the property of Gaussian distribution, it is sufficient to show that $Cov\left(\Delta_{w_1+1w_2+1}S_{l;\lambda^2},\Delta_{w_1'+1w_2'+1}S_{l;\lambda^2}\right) = 0$. By recalling the definition of the increments of $S_{l;\lambda^2}$, we get the following result when P is substituted by λ^2:

$$\begin{aligned}
&Cov\left(\Delta_{w_1+1w_2+1}S_{l;\lambda^2},\Delta_{w_1'+1w_2'+1}S_{l;\lambda^2}\right)\\
&=\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(h_2 - s_{w_2'} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(h_2 - s_{w_2'} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(h_2 - s_{w_2'} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(h_2 - s_{w_2'} + s_{w_2+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2}\right)
\end{aligned}$$

DOI: 10.3844/jmssp.2017.106.126

$$\begin{aligned}
&-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(h_2 - s_{w_2'} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(h_2 - s_{w_2'} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(h_2 - s_{w_2'+1} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(h_2 - s_{w_2'} + s_{w_2}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(h_2 - s_{w_2'} + s_{w_2}\right)\\
&=\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(-s_{w_2'+1}\right)-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(-s_{w_2'+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(-s_{w_2'}\right)+\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(-s_{w_2'}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(-s_{w_2'+1}\right)+\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(-s_{w_2'+1}\right)\\
&+\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(-s_{w_2'}\right)-\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(-s_{w_2'}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(-s_{w_2'+1}\right)+\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(-s_{w_2'+1}\right)\\
&+\left(h_1 - t_{w_1'+1} + t_{w_1+1}\right)\left(-s_{w_2'}\right)-\left(h_1 - t_{w_1'} + t_{w_1+1}\right)\left(-s_{w_2'}\right)\\
&+\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(-s_{w_2'+1}\right)-\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(-s_{w_2'+1}\right)\\
&-\left(h_1 - t_{w_1'+1} + t_{w_1}\right)\left(-s_{w_2'}\right)+\left(h_1 - t_{w_1'} + t_{w_1}\right)\left(-s_{w_2'}\right)=0
\end{aligned}$$

Since both rectangles are arbitrary and disjoint, then by the normality of the increments of $S_{l;\lambda^2}$, $\Delta_{w_1+1w_2+1}S_{l;\lambda^2}$ and $\Delta_{w_1'+1w_2'+1}S_{l;\lambda^2}$ are independent which is completing the proof.

Proposition A.7

Let $I_{\ell k} := \left[t_{n_1\ell}, t_{n_1\ell+1}\right] \times \left[s_{n_2k}, s_{n_2k+1}\right]$ be any rectangle in $D_{h_1h_2}$. For any h_1, h_2 with $0<h_1<(a_2-a_1)$ and $0<h_2<(b_2-b_1)$, it holds $Var\left(\Delta_{\ell+1,k+1}S_{l;\lambda^2}\right) = 4\lambda^2(I_{\ell k})$.

Proof

By a result in multivariate analysis, $Var\left(\Delta_{\ell+1,k+1}S_{l;\lambda^2}\right)$ = aE(hh^T)a^T, where the vector h is defined as:

$$h = \begin{pmatrix} S_{l;\lambda^2}\left(t_{n_1\ell+1}, s_{n_2k+1}\right), S_{l;\lambda^2}\left(t_{n_1\ell}, s_{n_2k+1}\right), \\ S_{l;\lambda^2}\left(t_{n_1\ell+1}, s_{n_2k}\right), S_{l;\lambda^2}\left(t_{n_1\ell}, s_{n_2k}\right) \end{pmatrix}^{\mathrm{T}}$$

and a := (1,-1,-1, 1). The definition of the covariance function $\gamma_{S_{l;\lambda^2}}$ leads us to the following variance-covariance matrix of h:

$$E\left(hh^T\right) = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{12} & c_{22} & c_{13} & c_{14} \\ c_{13} & c_{23} & c_{33} & c_{34} \\ c_{14} & c_{24} & c_{34} & c_{44} \end{pmatrix}$$

where:

$$\begin{aligned}
&c_{11} = c_{22} = c_{33} = c_{44} = h_1h_2, c_{12} = h_1\left(h_2 - s_{n_2k+1} + s_{n_2k}\right)\\
&c_{13} = h_2\left(h_1 - t_{n_1\ell+1} + t_{n_1\ell}\right), c_{14} = \left(h_1 - t_{n_1\ell+1} + t_{n_1\ell}\right)\left(h_2 - s_{n_2k+1} + s_{n_2k}\right),\\
&c_{23} = \left(h_1 = t_{n_1\ell+1} + t_{n_1\ell}\right)\left(h_2 - s_{n_2k+1} + s_{n_2k}\right),\\
&c_{24} = \left(h_1 - t_{n_1\ell+1} + t_{n_1\ell}\right)h_2, c_{34} = h_1\left(h_2 - s_{n_2k+1} + s_{n_2k}\right)
\end{aligned}$$

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

$$\begin{aligned}
&aE\left(hh^T\right)a^T = 4h_1h_2 + 4\left(h_1 - t_{n_1\ell+1} + t_{n_1\ell}\right)\left(h_2 - s_{n_2k+1} + s_{n_2k}\right)\\
&-4\left(h_1 - t_{n_1\ell+1} + t_{n_1\ell}\right)h_2 - 4\left(h_2 - s_{n_2k+1} + s_{n_2k}\right)h_1\\
&= 4h_1h_2 + 4h_1h_2 - 4h_1\left(s_{n_2k+1} + s_{n_2k}\right) - 4h_2\left(t_{n_2\ell+1} + t_{n_2\ell}\right)\\
&+4\left(t_{n_2\ell+1} + t_{n_2\ell}\right)\left(s_{n_2k+1} + s_{n_2k}\right) - 4h_1h_2 + 4h_1\left(s_{n_2k+1} + s_{n_2k}\right)\\
&-4h_1h_2 + 4h_2\left(t_{n_2\ell+1} + t_{n_2\ell}\right)\\
&= 4\left(t_{n_2\ell+1} + t_{n_2\ell}\right)\left(s_{n_2k+1} + s_{n_2k}\right) = 4\lambda^2\left(I_{\ell k}\right)
\end{aligned}$$

which is completing the proof.

Appendix B. Proof of Theorem 3.1

By the linearity of $MS_{h_1h_2}$ on $\mathbb{R}^{n_2\times n_1}$ and the definition of the cumulative sums operator, we get:

$$\int_{D_{h_1h_2}} f_1\, MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right) = \int_{D_{h_1h_2}} f_1\left(MS_{h_1h_2}\left(\frac{g_{n_2\times n_1}}{\sqrt{n_1n_2}}\right) + MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)\right)$$

By Equation 3, for every $(t, s)\in D_{h_1h_2}$, it holds:

$$\begin{aligned}
&\frac{1}{\hat{\sigma}_n}MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right)(t,s)\\
&= \frac{1}{\hat{\sigma}_n}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_1}1_{[t,t]_1\times[s,s]_2}\left(t_{n_1\ell}, s_{n_2k}\right)\frac{g\left(t_{n_1\ell}, s_{n_2k}\right)}{n_1n_2}\\
&+\frac{1}{\hat{\sigma}_n}MS_{h_1h_2}\left(\varepsilon_{n_2\times n_1}\right)(t,s)\\
&\frac{1}{\hat{\sigma}_n}\int_{[t,t]_1\times[s,s]_2} g\,dP_n + \frac{1}{\hat{\sigma}_n}MS_{h_1h_2}\left(\varepsilon_{n_2n_1}\right)(t,s)
\end{aligned}$$

The right-hand side of the last equation converges in distribution to:

$$\frac{1}{\sigma}\zeta_g(t,s) + S_{l;P}(t,s)$$

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

$$\frac{1}{\hat{\sigma}_n}\int_{D_{h_1h_2}} f_1\, MS_{h_1h_2}\left(Y^*_{n_2\times n_1}\right) \Rightarrow \frac{1}{\sigma}\int_{D_{h_1h_2}} f_1\,\zeta_g(t,s) + \int_{D_{h_1h_2}} f_1\, S_{l;P}(t,s)$$

which is establishing the proof.

Appendix C. Set-Indexed Gaussin white Noise

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

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

$$E_P\left(W_P(A)W_P(B)\right) = P(A\cap B), \forall A,B \in A_0$$

The sample path (trajectory) of W_P is concentrated in $C(A_0)$ (*cf.* (Lifshits, 2012), pp. 13-15).

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

Let the experimental design $\Xi_{n_1\times n_2}$ be constructed by using the probability measure P on $B(D)$ and let the sequence of the matrix of random errors $\varepsilon_{n_2\times n_1} := (\varepsilon_{k\ell})_{k=1,\ell=1}^{n_2,n_1}$ consists of i.i.d. random variables with $E(\varepsilon_{k\ell}) = 0$ *and* $Var(\varepsilon_{k\ell}) = \sigma^2 < 1$, for $n_1 \geq 1$ *and* $n_2 \geq 1$, where $\varepsilon_{k\ell} := \varepsilon\left(t_{n_1\ell}, s_{n_2k}\right)$. Then we have:

$$\frac{1}{\sigma}S_{n_1n_2}\left(\varepsilon_{n_2n_1}\right) \Rightarrow W_P, as\, n_1 \text{ and } n_2 \to \infty$$

Proof

Let $c_1,\ldots, c_m$ be any constant and $B_1,\ldots,B_m$ be any subset in A_0, $m \geq 1$. We define linear combination $F_{n_1n_2} := \sum_{i=1}^m c_i S_{n_1n_2}\left(\varepsilon\left(\Xi_{n_1\times n_2}\right)\right)(B_i)$. Then we get:

$$Var\left(F_{n_1n_2}\right) = E\left(\sum_{i=1}^m\sum_{j=1}^m c_ic_jS_{n_1n_2}\left(\varepsilon_{n_1\times n_2}\right)(B_i)S_{n_1n_2}\left(\varepsilon_{n_1\times n_2}\right)(B_j)\right)$$

$$= \sum_{i=1}^m\sum_{j=1}^m c_ic_j\frac{1}{n_1n_2}\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2}1_{\{B_i\cap B_j\}}\left(t_{n_1\ell}, s_{n_2k}\right)$$

$$= \sum_{i=1}^m\sum_{j=1}^m c_ic_jP_n\left(B_i\cap B_j\right) \to \sum_{i=1}^m\sum_{j=1}^m c_ic_jP\left(B_i\cap B_j\right)$$

where, $\sum_{i=1}^m\sum_{j=1}^m c_ic_jP(B_i\cap B_j)$ is actually the variance of $\sum_{i=1}^m c_iW_P(B_i)$ which is normally distributed with mean zero and variance $\sum_{i=1}^m\sum_{j=1}^m c_ic_jP(B_i\cap B_j)$. Next we show that the Lindeberg condition is fulfilled. For every $\in > 0$, let:

$$L(\varepsilon) := \sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2}E\left(\left|\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^m c_i 1_{B_i}\left(t_{n_1\ell}, s_{n_2k}\right)\varepsilon_{k\ell}\right|^2 \times 1_{\left\{\left|\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^m c_i 1_{B_i}\left(t_{n_1\ell}, s_{n_2k}\right)\varepsilon_{k\ell}\right| > \in\right\}}\right)$$

and let $M := \max_{1\leq i\leq m}|c_i|$. Then by the i.i.d. property of $\varepsilon_{k\ell}$ and by the well-known bounded convergence theorem it holds:

$$0 \leq \lim_{n_1,n_2\to\infty} L(\varepsilon) \leq \lim_{n_1,n_2\to\infty} M^2m^2E\left(|\varepsilon_{11}|^2\, 1_{\left\{|\varepsilon_{11}| > \frac{\in\sqrt{n_1n_2}}{Mm}\right\}}\right) = 0$$

Thus, by the Lindeberg-Levy central limit theorem, the finite dimensional distribution of $S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)$ converges weakly to that of W_P. Since $S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)$ is d_P-continuous, we define the modulus of continuity of $S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)$ as:

$$W\left(S_{n_1n_2};\delta\right) := \sup_{\{P(A\Delta B)<\delta\}}\left|S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)(A) - S_{n_1n_2}\left(\varepsilon_{n_2\times n_1}\right)(B)\right|$$

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

$$\lim_{\delta\to 0}\limsup_{n_1,n_2\to\infty} P\left\{\omega\in\Omega : W\left(S_{n_1n_2}(\omega);\delta\right) > \in\right\} = 0, \forall \in > 0$$

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

Proposition C.3

The process W_P constitutes a finite signed measure P-almost surely on A_0. That is there exists a set $\Omega' \subset \Omega$ with $P(\Omega'^C) = 0$ such that $\forall\omega\in\Omega'$, $W_P(\omega)$ is a finite signed measure on A_0.

Proof

Since $Var(W_P(A)) = P(A)$, $\forall A \in A_0$, then $Var\left(W_P(\emptyset)\right) = P(\emptyset) = 0$. This implies $\exists\Omega'$ with $P(\Omega'^C) = 0$, such that $W_P(\emptyset;\omega) = 0$, $\forall\omega\in\Omega'$. Next we show countable additivity. Let $\{A_n: n\geq 1\}$ be a sequence of disjoint sets in A_0, then $W_P\left(\bigcup_{j=1}^\infty A_j\right) \sim N\left(0, \sum_{j=1}^\infty P(A_j)\right)$. Since on the other hand $\sum_{j=1}^\infty W_P(A_j)$ is also distributed as $N\left(0,\sum_{j=1}^{\infty}P\left(A_j\right)\right)$ 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:

$$W_P\left(\bigcup_{j=1}^{\infty}A_j\right)\overset{D}{=}\sum_{j=1}^{\infty}W_P\left(A_j\right)$$

This means that W_P is countably additive *P-a.s.*, finishing the proof.

Proposition C.4

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

$$Cov\left(\int_D f dW_P, \int_D g dW_P\right) = \langle f,g\rangle_P$$
$$\int_D (f-g)dW_P \sim N\left(0, \| f-g\|_P^2\right)$$

Proof

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

$$Cov\left(\int_D f dW_P, \int_D g dW_P\right) = E_P\left(\lim_{n\to\infty}\sum_{j=1}^{n}\sum_{j=1}^{n}a_j b_j W_P\left(A_j\right)W_P\left(B_j\right)\right)$$
$$= \int_\Omega \lim_{n\to\infty}\sum_{j=1}^{n}\sum_{j=1}^{n}a_j b_j W_P\left(A_j;\omega\right)W_P\left(B_j;\omega\right)P(d\omega)$$
$$= \lim_{n\to\infty}\sum_{j=1}^{n}\sum_{j=1}^{n}a_j b_j \int_\Omega W_P\left(A_j;\omega\right)W_P\left(B_j;\omega\right)P(d\omega)$$
$$= \lim_{n\to\infty}\sum_{j=1}^{n}\sum_{j=1}^{n}a_j b_j P\left(A_j\cap B_j\right) = \lim_{n\to\infty}\int_D f_n(t,s)g_n(t,s)P(dt,ds)$$
$$\int_D f(t,s)g(t,s)P(dt,ds) = \langle f,g\rangle_P$$

Hence by this equation we further get $\int_D f\, dW_P$ is normally distributed with mean 0 and variance $\| f\|_P^2$. This result immediately implies $\int_D(f\text{-}g)\ dW_P$ 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 $n_2\geq 1$:

$$CU\left(Y_{n_2\times n_1}\right) = \int_D f_1\ (t,s)dS_{n_1 n_2}\left(Y_{n_2\times n_1}\right)(t,s)$$
$$= \sqrt{n_1 n_2}\int_D f_1\ (t,s)dS_{n_1 n_2}\left(\frac{g_1\left(\Xi_{n_1\times n_2}\right)}{\sqrt{n_1 n_2}}\right)(t,s)$$
$$+\int_D f_1\ dS_{n_1 n_2}\left(\varepsilon_{n_2\times n_1}\right)(t,s)$$

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:

$$CU\left(Y_{n_2\times n_1}\right)\overset{D}{\to}\sqrt{n_1 n_2}\int_D f_1\ (t,s)d\varphi g_1(t,s) + \int_D f_1\ (t,s)dW_P(t,s)$$

However, since $f1$ and $g1$ are orthogonal in $L_2(D, P)$ when H_0 is true, we get $\int_D f_1\ (t,s)d\varphi g_1(t,s)\langle f_1, g_1\rangle_P = 0$. Hence, under H_0, $CU\left(Y_{n_2\times n_1}\right)$ 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 α test as;

$$CU_\alpha\left\{Y_{n_2\times n_1}\in\mathbb{R}^{n_2\times n_1} : CU\left(Y_{n_2\times n_1}\right)\geq\Phi^{-1}(1-\alpha)\| f_1\|_P\right\}$$

Appendix E. Proof of Corollary 4.2

If $g_2 \equiv f$ for any $f\in V\cap W^{\perp}$, by applying the similar argument as in the proof of Theorem 4.1, we get;

$$CU\left(Y^*_{n_2\times n_1}\right)\overset{D}{\to}\langle f_1, f\rangle_P\int_D f_1\ (t,s)dW_P(t,s), as\ n_1, n_2\to\infty$$

Based on this convergence we further get:

$$\lim_{n_1, n_2\to\infty}\Psi_{P;n_1 n_2}(f)$$
$$= \lim_{n_1, n_2\to\infty}P\left\{CU\left(Y^*_{n_2\times n_1}\right)\geq\Phi^{-1}(1-\alpha)\| f_1\|_P | g_2\equiv f\right\}$$
$$= P\left\{\langle f_1, f\rangle_P + \int_D f_1\ (t,s)dW_P(t,s)\geq\Phi^{-1}(1-\alpha)\| f_1\|_P\right\}$$
$$= 1-\Phi\left(\Phi^{-1}(1-\alpha)-\frac{\langle f_1, f\rangle_P}{\| f_1\|_P}\right)$$