Multidimensional Set-Indexed Partial Sums Method for Checking the Appropriateness of a Multivariate Spatial Regression

W. Somayasa, Gusti N.A. Wibawa, and Yulius. B. Pasolon

Abstract—In this paper we propose an asymptotic procedure for diagnosing the appropriateness of a multivariate spatial regression with correlated responses based on the so-called multidimensional set-indexed least squares residual partial sums processes of the observations. The limit process which is a projection of the higher dimensional set-indexed Brownian sheet is derived by applying the vectorial analogue of Prohorov's theorem. The adequacy of the assumed model is tested by using the Kolmogorov-Smirnov and Cramér-von Mises functional of the processes. Simulation based investigation are conducted in studying the finite sample size performance of the tests by comparing with that of the classical likelihood ratio test. Finally we attempt to apply the proposed method to a mining data supplied by a mining industry in Southeast Sulawesi.

Keywords—Multivariate spatial regression, set-indexed partial sums process, least squares residuals, multidimensional set-indexed Brownian sheet, model-check, Kolmogorov-Smirnov, Cramér-von Mises.

I. INTRODUCTION

Multivariate spatial regression (M.S.R.) analysis is frequently utilized as a statistical tool for empirical model building in applied sciences such as in geosciences and industry. As evidenced in the literatures, checking the appropriateness of an assumed model is important before the fitted model is further used such as in the prediction of future observations. To this stage the residual of the observation or variant of it is commonly investigated. We refer the reader to Box and Draper [7], Christensen [9], pp. 9–21 and Johnson and Wichern [11], pp. 323–328 for the case where the observations are assumed to be normally distributed and Arnold [2] for an asymptotic approach.

In this paper we aim to demonstrate the application of a technique that based on the partial sums (cumulative sum = CUSUM) of the least squares residuals to verify whether or not the assumed model is adequate. The application of CUSUM technique can also be found in the problem of quality control in industry using control chart as the tool in detecting whether a change occurs during the production process. The comparison between the CUSUM and Shewart's control chart has been established in Kovarik and Sarga [12] by a case study.

Let us consider a multivariate spatial process

$$\{\mathbf{Y}(\mathbf{t}) := (Y_i(\mathbf{t}))_{i=1}^p : \mathbf{t} := (t_j)_{j=1}^d \in [a,b]^d \subset \mathbb{R}^d\}$$

where $[a,b]^d := [a_1,b_1] \times \cdots \times [a_d,b_d] =: \mathbf{E}$ is a fixed experimental region. Throughout we write any p-column vector of real numbers or real functions $\mathbf{w} := (w_1,\ldots,w_p)$ by $(w_i)_{i=1}^p$ for convenient. We assume that $\mathbf{Y}$ follows a nonparametric spatial regression model

$$\mathbf{Y}(\mathbf{t}) = \mathbf{g}(\mathbf{t}) + \mathcal{E}(\mathbf{t}),\ \mathbf{t} \in \mathbf{E}, \tag{1}$$

where $\mathbf{g} := (g^{(i)})_{i=1}^p : \mathbf{E} \to \mathbb{R}^p$ is the true-unknown regression function, and $\mathcal{E} := (\varepsilon^{(i)})_{i=1}^p$ is the vector of random errors with $E(\mathcal{E}) = \mathbf{0} \in \mathbb{R}^p$, and $Cov(\mathcal{E}) = \mathbf{\Sigma}$ which is unknown and positive definite.

In the framework of model-check for multivariate linear regression in the first step we assume that $\mathbf{g}$ belongs to a family of polynomials or trigonometric polynomials with a given upper bound on the degree. Then based on the sample we test the hypothesis if the assumed family of the polynomials is appropriate to represent the model. More precisely, let $\mathbf{W} := [f_1,\ldots,f_m]$ be a subspace of $L_2(\lambda_{\mathbf{E}})$ generated by some given regression functions $f_1,\ldots,f_m$, where $\lambda_{\mathbf{E}}$ is the Lebesque measure on $\mathbf{E}$. We then test the hypothesis

$$H_0 : \mathbf{g} \in \times_{i=1}^p \mathbf{W} := \mathbf{W}^p \ against\ H_1 : \mathbf{g} \notin \mathbf{W}^p \tag{2}$$

This work was supported in part by the Indonesian Ministry of Research, Technology and Higher Education through the *KLN* and *Publikasi Internasional* research Grant 2015.

W. Somayasa is with the Department of Mathematic, Halu Oleo University, Jl. H.E.A. Mokodompit, Kendari 93232, Indonesia (corresponding author; e-mail: wayan.somayasa@uho.ac.id)

G.N.A. Wibawa is with the Department of Mathematic, Halu Oleo University, Jl. H.E.A. Mokodompit, Kendari 93232, Indonesia (e-mail: adhi_gst@yahoo.com)

Y. B. Pasolon is with the Department of Agrotechnology, Halu Oleo University, Jl. H.E.A. Mokodompit, Kendari 93232, Indonesia (e-mail: julpas_59@yahoo.com)

based on a sample taken according to a design. Let $\{\mathbf{Y}_{n\mathbf{j}} := \mathbf{Y}(\mathbf{t}_{n\mathbf{j}}) : \mathbf{1} \leq \mathbf{j} \leq n\mathbf{1}\}$, $n \geq 1$ be a sequence of independent observations of Model 1 on a regular lattice with n^d points

$$\Xi_n^d := \{\mathbf{t}_{n\mathbf{j}} := (t_{nj_k})_{k=1}^d \in \mathbf{E}, 1 \leq j_k \leq n\},$$

where $t_{nj_k} := a_k + \frac{j_k}{n}(b_k - a_k)$, $k = 1, \ldots, d$. Here $\mathbf{1}$ denotes $(1, \ldots, 1)^\top \in \mathcal{Z}_+^d$ and $\mathbf{j} := (j_1, \ldots, j_d)^\top \in \mathcal{Z}_+^d$. Thus the set $\mathbf{Y}(\Xi_n^d) := \{\mathbf{Y}_{n\mathbf{j}} : \mathbf{1} \leq \mathbf{j} \leq n\mathbf{1}\}$ is a "pyramidal array" of p-random vectors indexed by the lattice of d-dimensional positive integer that satisfies the model $\mathbf{Y}_{n\mathbf{j}} = \mathbf{g}(\mathbf{t}_{n\mathbf{j}}) + \mathcal{E}_{n\mathbf{j}}$, where $\mathcal{E}_{n\mathbf{j}}$ is the array of independent and identically distributed error vectors with $E(\mathcal{E}_{n\mathbf{j}}) = \mathbf{0}$ and $Cov(\mathcal{E}_{n\mathbf{j}}) = \mathbf{\Sigma}$. Without loss of generality and for notational convenience we consider in this paper $n \times n$-regular lattice $\Xi_n = \{(\ell/n, k/n) : 1 \leq \ell, k \leq n\}$ in the unit rectangle $\mathbf{I} := [0, 1] \times [0, 1]$ as the experimental condition. Our result can be immediately extended to the higher dimensional case.

In the classical multivariate analysis where $\mathcal{E}_{n\ell k}$ are assumed to be normally distributed, a more specific hypothesis than (2) is considered. That is

$$H_0 : \mathbf{g} \in \mathbf{W}^p \; against \; H_1 : \mathbf{g} \in \mathbf{V}^p, \tag{3}$$

where $\mathbf{V} := [f_1, \ldots, f_m, f_{m+1}, \ldots, f_q] \supseteq \mathbf{W}$. The set $\{f_{m+1}, \ldots, f_q\}$ are the set of additional known regression functions in $L_2(\lambda_{\mathbf{I}})$ needed to represent $\mathbf{g}$. Let $\mathbf{W}_n := [f_1(\Xi_n), \ldots, f_m(\Xi_n)]$, and $\mathbf{V}_n := [f_1(\Xi_n), \ldots, f_m(\Xi_n), f_{m+1}(\Xi_n), \ldots, f_q(\Xi_n)]$ be subspaces of $\mathbb{R}^{n\times n}$ generated respectively by the set of $n \times n$-matrixes $\{f_1(\Xi_n), \ldots, f_m(\Xi_n)\}$ and $\{f_1(\Xi_n), \ldots, f_m(\Xi_n), f_{m+1}(\Xi_n), \ldots, f_q(\Xi_n)\}$, where $f_j(\Xi_n) := (f_j(\ell/n, k/n))_{\ell,k=1}^n \in \mathbb{R}^{n\times n}$ is obtained by evaluating f_j to every point in Ξ_n, $j = 1, \ldots, q$. Furthermore, let $\mathbf{W}_n^p := \times_{i=1}^p \mathbf{W}_n$ and $\mathbf{V}_n^p := \times_{i=1}^p \mathbf{V}_n$ be the product of p copies of $\mathbf{W}_n$ and $\mathbf{V}_n$, respectively. Let $pr_{\mathbf{W}_n^p}$ be the component-wise orthogonal projector onto $\mathbf{W}_n^p$. Then by using the analogous argument as the estimation procedure defined in [9], pp. 1–5. and [11], pp. 314–328, the ordinary least squares residual of the array $\mathbf{Y}(\Xi_n)$ is given by

$$\mathbf{R}(\Xi_n) := (\mathbf{r}_{n\ell k})_{\ell,k=1}^n = \mathbf{Y}(\Xi_n) - pr_{\mathbf{W}_n^p}\mathbf{Y}(\Xi_n), \tag{4}$$

where $\mathbf{r}_{n\ell k} := (r_{n\ell k}^{(1)}, \ldots, r_{n\ell k}^{(p)})^\top$, for $1 \leq \ell, k \leq n$. Suppose $\{f_1(\Xi_n), \ldots, f_m(\Xi_n)\}$ builds an orthonormal basis (ONB) for $\mathbf{W}_n$, then $\mathbf{R}(\Xi_n)$ can be equivalently expressed as $\mathbf{R}(\Xi_n) = (R^{(i)}(\Xi_n))_{i=1}^p \in \times_{i=1}^p \mathbb{R}^{n\times n}$, where for $i = 1, \ldots, p$,

$$\begin{aligned} R^{(i)}(\Xi_n) &= Y^{(i)}(\Xi_n) \\ &- \sum_{j=1}^p \langle f_j(\Xi_n), Y^{(i)}(\Xi_n)\rangle_{\mathbb{R}^{n\times n}} f_j(\Xi_n). \end{aligned}$$

Here $\langle \cdot, \cdot \rangle_{\mathbb{R}^{n\times n}}$ denotes the inner product in $\mathbb{R}^{n\times n}$, defined by

$$\langle \mathbf{A}, \mathbf{B}\rangle_{\mathbb{R}^{n\times n}} := trace(\mathbf{A}^\top \mathbf{B}), \; \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}.$$

Several test procedures have been proposed for testing Hypothesis 3 (see [11] pp. 396–398). One of them is the classical likelihood ratio (L.R.)-test using the modified Wilk's lambda statistic defined by

$$\Lambda_n := [n^2 - q - \frac{1}{2}(p - q + m + 1)] \ln \left\{ \frac{det\widehat{\Sigma}_n^{\mathbf{V}}}{det\widehat{\Sigma}_n^{\mathbf{W}}} \right\},$$

where

$$\begin{aligned} \widehat{\Sigma}_n^{\mathbf{W}} &:= (\mathbf{Y}_{n^2} - \mathbf{X}_m(\mathbf{X}_m^\top \mathbf{X}_m)^{-1}\mathbf{X}_m^\top \mathbf{Y}_{n^2})^\top \times \\ &\quad (\mathbf{Y}_{n^2} - \mathbf{X}_m(\mathbf{X}_m^\top \mathbf{X}_m)^{-1}\mathbf{X}_m^\top \mathbf{Y}_{n^2})/n^2 \\ \widehat{\Sigma}_n^{\mathbf{V}} &:= (\mathbf{Y}_{n^2} - \mathbf{X}_q(\mathbf{X}_q^\top \mathbf{X}_q)^{-1}\mathbf{X}_q^\top \mathbf{Y}_{n^2})^\top \times \\ &\quad (\mathbf{Y}_{n^2} - \mathbf{X}_q(\mathbf{X}_q^\top \mathbf{X}_q)^{-1}\mathbf{X}_q^\top \mathbf{Y}_{n^2})/n^2. \end{aligned}$$

Thereby we define the matrices

$$\begin{aligned} \mathbf{Y}_{n^2} &:= (vec(Y^{(1)}(\Xi_n)), \ldots, vec(Y^{(p)}(\Xi_n))), \\ \mathbf{X}_m &:= (vec(f_1(\Xi_n)), \ldots, vec(f_m(\Xi_n))), \\ \mathbf{X}_q &:= (vec(f_1(\Xi_n)), \ldots, vec(f_q(\Xi_n))), \end{aligned}$$

where "vec" denotes the vec operator defined e.g. in Magnus and Neudecker [16], pp. 34–36. The test will reject H_0 at level $\alpha \in (0, 1)$, if and only if $\Lambda_n \geq \chi_{p(q-m)}^{2;(1-\alpha)}$, where $\chi_{p(q-m)}^{2;(1-\alpha)}$ is the $(1-\alpha)$-quantile of central chi-square distribution with $p(q - m)$ degrees of freedom.

The L.R.-test defined above has restriction in the implementation in that the population under study must be normally distributed. However such a distributional simplification is frequently found to be unrealistic especially in the statistical modelling of geosciences (cf. Christensen [9], pp. 263–299 and Cressie [10], pp 105–183). In the forecasting problem of times series regression, normality assumption possibly leads to spurious results. Alternative approaches with Grey and ARIMA models under modified residuals was proposed in Shu and Hsu [19].

In contrast to this classical approach, for our proposed method we do not need such normality assumption. We only consider the partial sums of the array of the residuals instead. Precisely, let $\mathcal{A}$ be the family of convex subset of $\mathbf{I}$, and d_λ be the Lebesgue pseudometric on $\mathcal{A}$. Let $\mathcal{C}(\mathcal{A})$ be the set of continuous functions on $\mathcal{A}$ with respect to d_λ. We embeds $\mathbf{R}(\Xi_n)$ into a stochastic process indexed by $\mathcal{A}$ defined by

$$\mathbf{V}_n(\mathbf{R}(\Xi_n))(B) := \sum_{\ell=1}^n \sum_{k=1}^n n\lambda(B \cap C_{\ell k})\mathbf{r}_{n\ell k}, B \in \mathcal{A}$$

where $C_{\ell k} := ((\ell-1)/n, \ell/n] \times ((k-1)/n, k/n]$. The i-th one dimensional component of $\mathbf{V}_n(\mathbf{R}(\Xi_n))$ is defined by

$$\mathbf{T}_n(R^{(i)}(\Xi_n))(B) = \sum_{\ell=1}^{n}\sum_{k=1}^{n} n\lambda(B \cap C_{\ell k}) r^{(i)}_{n\ell k}.$$

Results concerning the properties of $\mathbf{T}_n$ was studied in the work of Somayasa, Ruslan, Cahyono and Ngkoimani [22]. Interested reader is also suggested to see Bischoff and Somayasa [6] for the case of the ordinary partial sums process.

By the definition of the operator $\mathbf{V}_n$, the process $\{\mathbf{V}_n(\mathbf{R}(\Xi_n))(B) : B \in \mathcal{A}\}$ induces the sample path in the space $\mathcal{C}^p(\mathcal{A})$. As a convention we call this process a p-dimensional least squares residual partial sum process (L.S.R.P.S.P.) indexed by $\mathcal{A}$. In the present paper $\mathcal{C}^p(\mathcal{A})$ is furnished with the uniform topology induced by the metric φ defined by

$$\varphi(\boldsymbol{u}, \boldsymbol{w}) := \sum_{i=1}^{p} \|u_i - w_i\|_{\mathcal{A}},$$

for $\boldsymbol{u} := (u_i)_{i=1}^p$ and $\boldsymbol{w} := (w_i)_{i=1}^p \in \mathcal{C}^p(\mathcal{A})$, where for $i = 1, \ldots, p$,

$$\|u_i - w_i\|_{\mathcal{A}} := \sup_{A\in\mathcal{A}} |u_i(A) - w_i(A)|.$$

We propose the Kolmogorov-Smirnov (K.S.) and Cramér-von Mises (C.M.) functionals of the sequence of the p-dimensional L.S.R.P.S.P. as the test statistics for testing (2) defined by

$$KS_{n,\mathcal{A}} := \sup_{B\in\mathcal{A}} \|\boldsymbol{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))(B)\|_{\mathbb{R}^p}$$
$$CM_{n,\mathcal{A}} := \frac{1}{n^2}\sum_{B\in\mathcal{A}} \|\boldsymbol{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))(B)\|^2_{\mathbb{R}^p},$$

where $\boldsymbol{\Sigma}^{-1/2}$ a $p \times p$-symmetric matrix that satisfies $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}$. The K.S.-test (resp. C.M.-test) will reject H_0 at a level $\alpha \in (0,1)$ if and only if $KS_{n,\mathcal{A}} > q_{1-\alpha}$ (resp. $CM_{n,\mathcal{A}} > c_{1-\alpha}$), where $q_{1-\alpha}$ (resp. $c_{1-\alpha}$) is the $(1-\alpha)$-th quantile of the limiting distribution of $KS_{n,\mathcal{A}}$ (resp. $CM_{n,\mathcal{A}}$).

The properties of the one-dimensional version of the C.M.-test for spatial data was firstly investigated in MacNeill and Jandhyalla [15]. The results was generalized to one dimensional set-indexed LSRPSP by Xie and MacNeill [25]. They got the limit processes by applying the approach of MacNeill [13, 14]. In [6] the limit process of the sequence of ordinary L.S.R.P.S.P. for univariate spatial regression was derived by generalizing the geometric approach of Bischoff [5]. The extension of the method to the set-indexed L.S.R.P.S.P. has been already studied in [22]. In this paper we derive the limit process of the multidimensional set-indexed L.S.R.P.S.P. by applying the vectorial extension of Prohorov's theorem (cf. Billingsley [4], pp. 35–40).

The rest of the present paper is organized as follows. In Section II we establish a limit theorem for the p-dimensional set-indexed L.S.R.P.S.P. under H_0 as well as under H_1. For this purpose we generalize the uniform central limit theorem of Alexander and Pyke [1] to higher dimensional case, see Appendix. Examples of the limit process under H_0 and H_1 associated to polynomial models are studied in Section III. Simulation study is devoted in Section IV. Application of the method to real data is presented in Section V. The paper is closed in Section VI with a conclusion and suggestion for future work.

II. THE LIMIT OF $KS_{n,\mathcal{A}}$ AND $CM_{n,\mathcal{A}}$

A process $\mathbf{Z}_p := \{(Z^{(1)}(A), \ldots, Z^{(p)}(A))^\top : A \in \mathcal{A}\}$ is called the p-dimensional set-indexed Brownian sheet indexed by $\mathcal{A}$, if and only if it is centered Gaussian process with the covariance function

$$Cov(\mathbf{Z}_p(A_1), \mathbf{Z}_p(A_2)) = \lambda(A_1 \cap A_2)\mathbf{I}_p, A_1, A_2 \in \mathcal{A}$$

where $\mathbf{I}_p$ is the $p \times p$-identity matrix. The existence of this process can be shown by generalizing the uniform central limit theorem of Alexander and Pyke [1] and Pyke [17] to vectorial processes indexed by sets. As a result in probability theory it must be well known. However we can not find a reference where it is stated. Therefor we present the theorem together with the proof in the appendix.

Now we are ready to state the limit process of the sequence of p-dimensional set-indexed L.S.R.P.S.P. for the model specified under H_0 as well as under H_1.

Theorem 1 *Let $\{f_1, \ldots, f_m\}$ be an orthonormal bases (ONB) of $\mathbf{W}$ and h_{f_j} a function in $\mathcal{C}(\mathcal{A})$ defined by $h_{f_j}(A) := \int_A f_j d\lambda$. If f_j is continuous and have bounded variation in the sense of Hardy (cf. [8]), that is $f_j \in BV_H(\mathbf{I})$, $j = 1, \ldots, m$, then under H_0 it holds true*

$$\boldsymbol{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n)) \xrightarrow{\mathcal{D}} \mathbf{Z}^{H_0}_{p,\mathbf{f}} := \mathbf{Z}_p - pr^*_{\mathbf{W}\mathcal{H}_{\mathbf{Z}_p}}\mathbf{Z}_p.$$

where

$$pr^*_{\mathbf{W}\mathcal{H}_{\mathbf{Z}_p}}\mathbf{Z}_p = (pr^*_{\mathbf{W}\mathcal{H}_Z}Z^{(1)}, \ldots, pr^*_{\mathbf{W}\mathcal{H}_Z}Z^{(p)})^\top$$

constitutes a component-wise projection of $\mathbf{Z}_p$ in the sense of Rudin [18]. Thereby for every $u \in \mathcal{C}(\mathcal{A})$ and

$A \in \mathcal{A}$,

$$(pr^*_{\mathbf{W}_{\mathcal{H}_Z}} u)(A) := \sum_{j=1}^{m} \langle h_{f_j}, u\rangle h_{f_j}(A).$$

Here $\int^R$ denotes the Riemann-Stieltjes integral, and for $(t,s) \in I$, $u(t,s) := u([0,t]\times[0,s])$. Moreover $\mathbf{Z}_{p,\mathbf{f}}$ is a centered Gaussian process with the covariance function given by

$$K_{\mathbf{Z}^{H_0}_{p,\mathbf{Z}}}(A,B) := (\lambda(A\cap B) - \sum_{j=1}^{m} h_{f_j}(A)h_{f_j}(B))\mathbf{I}_p$$

where $h_{f_j}(A) := \int_A f_j d\lambda$.

Proof: Under H_0, Equation 4 is equivalent to

$$\mathbf{V}_n(\mathbf{R}(\Xi_n)) = \mathbf{V}_n(\mathcal{E}(\Xi_n)) - \mathbf{V}_n(pr_{\mathbf{W}^p_n}\mathcal{E}(\Xi_n)).$$

Next by Lemma 5.1 in [22], we further get

$$\begin{aligned}\mathbf{V}_n(pr_{\mathbf{W}^p_n}\mathcal{E}(\Xi_n)) &= (pr_{\mathbf{W}_{n\mathcal{H}_Z}}\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n)))^p_{i=1}\\ &= pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n)),\end{aligned}$$

where $\mathbf{W}_{n\mathcal{H}_Z} := [h_{\tilde{s}^{(n)}_1}, \ldots, h_{\tilde{s}^{(n)}_m}]$ is subspace of $\mathcal{C}(\mathcal{A})$ generated by $\{h_{\tilde{s}^{(n)}_1}, \ldots, h_{\tilde{s}^{(n)}_m}\}$, with $h_{\tilde{s}^{(n)}_w}(A) := \int_A \tilde{s}^{(n)}_w d\lambda$, for some orthonormal set of step functions $\{\tilde{s}^{(n)}_1, \ldots, \tilde{s}^{(n)}_m\} \subset BV_H(\mathbf{I})$ that satisfy $\|\tilde{s}^{(n)}_w - f_w\|_\infty \to 0$, as $n\to\infty$, $w = 1,\ldots,m$. Interested reader is referred to Lemma A.8 and Lemma A.9 in [6] for the definition of $\tilde{s}^{(n)}_w$ and their properties useful for obtaining the result.

By considering Theorem A.1 we only need to show that the process satisfies the weak convergence

$$\mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n)) \xrightarrow{\mathcal{D}} pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{Z}_p,\ n\to\infty.$$

For this we apply Prohorov's theorem that suggests to show two conditions. First, we have to show that the finite dimensional distribution of $\mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n))$ converges to that of $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{Z}_p$, where for every $A\in\mathcal{A}$, $(pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{Z}_p)(A)$ follows a centered p-variate normal distribution with the covariance matrix

$$\mathcal{C}(A,B) := \sum_{w=1}^{m} h_{f_w}(A)h_{f_w}(B)\mathbf{I}_p,\ for\ A,B\in\mathcal{A}.$$

Second, the process must be shown to be tight.

For the first objective we apply the Lindeberg-Levy multivariate central limit theorem studied e.g. in Van der Vaart [24], pp. 16. Let $A_1,\cdots,A_r$ be any convex sets in $\mathcal{A}$ and let $c_1,\ldots,c_r$ be any real numbers. We consider the asymptotic distribution of the general linear combination

$$\begin{aligned}\mathbf{U}_n &:= \sum_{k=1}^{r} c_k \mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n))(A_k)\\ &= \sum_{k=1}^{r} c_k \mathbf{\Sigma}^{-1/2}(pr_{\mathbf{W}_{n\mathcal{H}_Z}}\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n))(A_k))^p_{i=1}. \quad (5)\end{aligned}$$

Since $E(\mathbf{U}_n) = \mathbf{0}$, by the standard theory of multivariate analysis, it holds

$$\begin{aligned}&Var(\mathbf{U}_n) =\\ &\sum_{k,\ell=1}^{r} c_k c_\ell \mathbf{\Sigma}^{-1/2} E(pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n))(A_k)\times\\ &pr^\top_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}}\mathbf{V}_n(\mathcal{E}(\Xi_n))(A_\ell))\mathbf{\Sigma}^{-1/2}\\ &= \sum_{k,\ell=1}^{r} c_k c_\ell \mathbf{\Sigma}^{-1} E((B_i(A_k)B_j(A_\ell))^p_{i,j=1}),\end{aligned}$$

where $B_i(A_k) := pr_{\mathbf{W}_{n\mathcal{H}_Z}}\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n))(A_k)$, $i = 1,\ldots,p$ and $k = 1,\ldots,r$. Next, for fixed k and ℓ, we have by the definition of $pr_{\mathbf{W}_{n\mathcal{H}_Z}}$,

$$\begin{aligned}EB_i(A_k)B_j(A_\ell) &= \sum_{w,w'=1}^{m} E\int_{\mathbf{I}}^{R} \tilde{s}^{(n)}_w d\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n))\\ &\times \int_{\mathbf{I}}^{R} \tilde{s}^{(n)}_{w'} d\mathbf{T}_n(\varepsilon^{(j)}(\Xi_n)) h_{\tilde{s}^{(n)}_w}(A_k) h_{\tilde{s}^{(n)}_{w'}}(A_\ell),\end{aligned}$$

where by referring to Bischoff and Somayasa [6], for $(t,s)\in\mathbf{I}$,

$$\begin{aligned}&\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n))(t,s) :=\\ &\frac{1}{n}\sum_{k=1}^{[ns]}\sum_{\ell=1}^{[nt]}\varepsilon^{(i)}_{n\ell k} + \frac{(nt-[nt])}{n}\sum_{k=1}^{[ns]}\varepsilon^{(i)}_{n[nt]+1,k}\\ &+\frac{(ns-[ns])}{n}\sum_{\ell=1}^{[nt]}\varepsilon_{n\ell,[ns]+1}\\ &+\frac{(nt-[nt])(ns-[ns])}{n}\varepsilon^{(i)}_{n[nt]+1,[ns]+1}.\end{aligned}$$

Hence, by the definition of the Riemann-Stieljes integral (cf. Stroock [20], pp. 7–17) and by the independence of $\{\varepsilon_{n\ell k} : 1\le \ell,k\le n\}$, we further get

$$\begin{aligned}&E(\int_{\mathbf{I}}^{R}\tilde{s}^{(n)}_w d\mathbf{T}_n(\varepsilon^{(i)}(\Xi_n))\int_{\mathbf{I}}^{R}\tilde{s}^{(n)}_{w'}d\mathbf{T}_n(\varepsilon^{(j)}(\Xi_n)))\\ &= \sum_{u,v,u',v'=1}^{n}\tilde{s}^{(n)}_w(\frac{u}{n},\frac{v}{n})\frac{1}{n^2}E(\varepsilon^{(i)}_{nuv}\varepsilon^{(j)}_{nu'v'})\tilde{s}^{(n)}_{w'}(\frac{u'}{n},\frac{v'}{n})\\ &= \frac{1}{n^2}\sum_{u,v=1}^{n}\tilde{s}^{(n)}_w(\frac{u}{n},\frac{v}{n})\sigma_{ij}\tilde{s}^{(n)}_{w'}(\frac{u'}{n},\frac{v'}{n}).\end{aligned}$$

Because of the convergence $\|\tilde{s}_w^{(n)} - f_w\|_\infty \overset{n\to\infty}{\longrightarrow} 0$, for $w = 1, \ldots, m$ (see [6]), the right hand side of the last equation converges to

$$\sigma_{ij} \int_{\mathbf{I}} f_w f_{w'} d\lambda = \sigma_{ij} \langle f_w, f_{w'} \rangle_{L_2} = \sigma_{ij} \delta_{ww'},$$

where $\delta_{ww'} = 1$ if $w = w'$, and $\delta_{ww'} = 0$ if $w \neq w'$. Also we get

$$h_{\tilde{s}_w^{(n)}}(A_k) h_{\tilde{s}_{w'}^{(n)}}(A_\ell) \overset{n\to\infty}{\longrightarrow} h_{f_w}(A_k) h_{f_{w'}}(A_\ell).$$

Hence, by combining these two results we have

$$E(B_i(A_k) B_j(A_\ell)) \overset{n\to\infty}{\longrightarrow} \sigma_{ij} \sum_{w=1}^{m} h_{f_w}(A_k) h_{f_w}(A_\ell).$$

Let $\mathbf{D} := \sum_{w=1}^{m} h_{f_w}(A_k) h_{f_w}(A_\ell) \mathbf{1}\mathbf{1}^\top$ and let $\odot$ denote the Hadarmard product defined e.g. in Magnus and Neudecker [16], pp. 53–54. Then by the preceding results and by the properties of $\odot$, $Var(\mathbf{U}_n)$ satisfies

$$Var(\mathbf{U}_n) \overset{n\to\infty}{\longrightarrow} \sum_{k,\ell=1}^{r} c_k c_\ell \mathbf{\Sigma}^{-1/2} (\mathbf{\Sigma} \odot \mathbf{D}) \mathbf{\Sigma}^{-1/2} =$$

$$\sum_{k,\ell=1}^{r} c_k c_\ell \mathcal{C}(A_k, A_\ell) = Var \sum_{k=1}^{r} c_k (pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{Z}_p)(A_k).$$

Next we investigate that the Lindeberg condition is fulfilled by the central limit theorem. We observe that $\mathbf{U}_n$ can also be written as

$$\mathbf{U}_n = \sum_{u,v=1}^{n} \sum_{k,w=1}^{r,m} \frac{c_k}{n} \mathbf{\Sigma}^{-1/2} \widetilde{s}_w^{(n)}(\frac{u}{n}, \frac{v}{n}) \mathcal{E}_{nuv} h_{\widetilde{s}_w^{(n)}}(A_k).$$

Hence we get

$$\sum_{u,v=1}^{n} \| \sum_{k,w=1}^{r,m} \frac{c_k}{n} \mathbf{\Sigma}^{-1/2} \widetilde{s}_w^{(n)}(\frac{u}{n}, \frac{v}{n}) \mathcal{E}_{nuv} h_{\widetilde{s}_w^{(n)}}(A_k) \|^2_{\mathbb{R}^p}$$

$$\leq \sum_{u,v=1}^{n} \| rmM\gamma_f^2 \frac{1}{n} \mathbf{\Sigma}^{-1/2} \mathcal{E}_{nuv} \|^2_{\mathbb{R}^p}$$

$$= (rmM\gamma_f^2)^2 \|\mathbf{\Sigma}^{-1/2}\|^2 \frac{1}{n^2} \sum_{u,v=1}^{n} \|\mathcal{E}_{nuv}\|^2_{\mathbb{R}^p},$$

where

$$M := \max_{1\leq k\leq r} \{c_k\} \; and \; \gamma_f := \max_{1\leq w\leq m} \|f_w\|_\infty.$$

That why by considering the stochastically independent property of the array of the p-vector of the random errors and the well-known bounded convergence theorem, for every $\epsilon > 0$ it holds

$$0 \leq \lim_{n\to\infty} \sum_{u,v=1}^{n} E(\|rmM\gamma_f^2 \mathbf{\Sigma}^{-1/2} \frac{1}{n} \mathcal{E}_{nuv}\|^2_{\mathbb{R}^p}$$

$$\times \mathbf{1}_{\{\|rmM\gamma_f^2 \mathbf{\Sigma}^{-1/2} \frac{1}{n} \mathcal{E}_{nuv}\|_{\mathbb{R}^p} \geq \epsilon\}})$$

$$\leq (rmM\gamma_f^2)^2 \|\mathbf{\Sigma}^{-1/2}\|^2 \lim_{n\to\infty} E(\|\mathcal{E}_{n11}\|^2_{\mathbb{R}^p}$$

$$\times \mathbf{1}_{\{\|\mathcal{E}_{n11}\|_{\mathbb{R}^p} \geq \frac{\epsilon n}{rmM\gamma_f^2 \|\mathbf{\Sigma}^{-1/2}\|^2}\}}) = 0.$$

The last result implies that the Lindeberg condition is satisfied. Therefore it can be concluded that the finite dimensional distribution of $\mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{V}_n(\mathcal{E}(\Xi_n))$ converges to that of $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{Z}_p$.

To show the tightness of the process we define the modulus of continuity of the p-dimensional process $\mathcal{Z}_n := (Z_n^{(i)})_{i=1}^p := \mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}_{n\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{V}_n(\mathcal{E}(\Xi_n))$ by

$$w(\mathcal{Z}_n, \delta) := \sup_{\{A,B\in\mathcal{A}: d_\lambda(A,B)<\delta\}} d_{\mathcal{Z}_n}(A, B),$$

where $d_{\mathcal{Z}_n}(A, B) := \|\mathcal{Z}_n(A) - \mathcal{Z}_n(B)\|_{\mathbb{R}^p}$, for $A, B \in \mathcal{A}$. Since $w(\mathcal{Z}_n, \delta) \leq \sum_{i=1}^{p} w(Z_n^{(i)}, \delta)$, then two show the tightness of $\mathcal{Z}_n$ it is sufficient to investigate the tightness of $Z_n^{(i)}$ for all i. In other word, the one-dimensional component of $\mathcal{Z}_n$ must be tight in order to make the p-dimensional process to be tight. By Theorem 2.3 in [4] we need the sufficient condition in that $Z_n^{(i)}$ converges to $Z^{(i)}$, for all i. The proof is finished with Theorem 2.2 of Somayasa [22].

By Theorem 1 and the well-known continuous mapping theorem (cf. Theorem 5.1 in [4]) the sampling distribution of the statistics $KS_{n,\mathcal{A}}$ and $CM_{n,\mathcal{A}}$ can be immediately approximated by those of $\sup_{A\in\mathcal{A}} \left\| \mathbf{Z}_{p,\mathbf{f}}^{H_0}(A) \right\|_{\mathbb{R}^p}$ and $\int_{\mathbf{I}} \|\mathbf{Z}_{p,\mathbf{f}}^{H_0}(A)\|^2_{\mathbb{R}^p} dA$, respectively.

The test procedures derived above are consistent in the sense the power of the test under the competing alternative converges to 1. Consequently the behavior of the test can not be observed as the model moves away from H_0. Therefore without altering the test problem for (2) and (3) to be able to investigate how good our tests perform under the alternatives we consider a localized model defined by

$$\mathbf{Y}(\Xi_n) = \frac{1}{n} \mathbf{g}^{local}(\Xi_n) + \mathcal{E}(\Xi_n)$$

where $\mathbf{g}^{local}(\Xi_n) := \frac{1}{n} \mathbf{g}(\Xi_n) = \frac{1}{n} (g^{(i)}(\Xi_n))_{i=1}^p$.

Theorem 2 *Let $\{f_1, \ldots, f_m\}$ be an ONB of $\mathbf{W}$ and for all $i = 1, \ldots, p$, $g^{(i)}$ have bounded variation in*

the sense of Vitali on $\mathbf{I}$*. If* $f_1, \ldots, f_m$ *are continuous and have bounded variation in the sense of Hardy on* $\mathbf{I}$*. Then under the alternative we have*

$$\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n)) \xrightarrow{\mathcal{D}} \mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}^{\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{h_g} + \mathbf{Z}^{H_0}_{p,\mathbf{f}},$$

where

$$pr_{\mathbf{W}^{\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{h_g} = \mathbf{h_g} - \sum_{w=1}^{m} (\langle h_{f_w}, h_{g^{(i)}} \rangle_{\mathcal{H}_Z})_{i=1}^{p} h_{f_w}$$
$$\mathbf{h_g} := (h_{g^{(i)}})_{i=1}^{p} : \mathcal{A} \to \mathbb{R}^p,$$

with

$$h_{g^{(i)}}(B) := \int_B g^{(i)}\, d\lambda,\ B \in \mathcal{A}.$$

Proof: Under H_1 we have

$$\mathbf{R}(\Xi_n) = pr_{\mathbf{W}^{p\perp}_n} \frac{1}{n} \mathbf{g}(\Xi_n) + pr_{\mathbf{W}^{p\perp}_n} \mathcal{E}(\Xi_n).$$

Hence by applying Lemma 5.1 in [22] and by considering the definition of the p-dimensional set-indexed partial sums operator $\mathbf{V}_n$, we further get

$$\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n)) = \mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}^{\perp}_{n\mathcal{H}_{\mathbf{Z}_p}}} \frac{1}{n} \mathbf{V}_n(\mathbf{g}(\Xi_n)) + pr_{\mathbf{W}^{\perp}_{n\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{\Sigma}^{-1/2} \mathbf{V}_n(\mathcal{E}(\Xi_n).$$

Because of Theorem 1, we only need to proof $pr_{\mathbf{W}^{\perp}_{n\mathcal{H}_{\mathbf{Z}_p}}} \frac{1}{n}\mathbf{V}_n(\mathbf{g}(\Xi_n))$ converges to $pr_{\mathbf{W}^{p\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{h_g}$. By the definition of the component-wise projection and by the fact that $\langle \cdot, \cdot \rangle$ restricted to $\mathcal{H}_Z$ coincides to $\langle \cdot, \cdot \rangle_{\mathcal{H}_Z}$ (cf. Proposition 2.1 in [22]), the i-th component of $pr_{\mathbf{W}^{\perp}_{n\mathcal{H}_{\mathbf{Z}_p}}} \frac{1}{n}\mathbf{V}_n(\mathbf{g}(\Xi_n))$ has the representation

$$\frac{1}{n}\mathbf{T}_n(g^{(i)}(\Xi_n)) - \sum_{w=1}^{m} \langle h\tilde{s}_w^{(n)}, \frac{1}{n}\mathbf{T}_n(g^{(i)}(\Xi_n)) \rangle_{\mathcal{H}_Z} h\tilde{s}_w^{(n)}$$

where the terms $\frac{1}{n}\mathbf{T}_n(g^{(i)}(\Xi_n))$ and $h\tilde{s}_w^{(n)}$ converge to $h_{g^{(i)}} := \int_I g^{(i)} d\lambda$ and h_{f_w}, as $n \to \infty$, respectively, for $i = 1, \ldots, p$. The continuity of the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_Z}$ implies $\langle h\tilde{s}_w^{(n)}, \frac{1}{n}\mathbf{T}_n(g^{(i)}(\Xi_n)) \rangle_{\mathcal{H}_Z} h\tilde{s}_w^{(n)}$ converges to $\langle h_{f_w}, h_{g^{(i)}} \rangle_{\mathcal{H}_Z} h_{f_w}$, as $n \to \infty$, for all w. This convergence results finally implies

$$\mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}^{\perp}_{n\mathcal{H}_{\mathbf{Z}_p}}} \frac{1}{n} \mathbf{V}_n(\mathbf{g}(\Xi_n)) \overset{n\to\infty}{\longrightarrow} \mathbf{\Sigma}^{-1/2} pr_{\mathbf{W}^{p\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{h_g}.$$

The power function of the asymptotically size α K.S. and C.M.-tests can now be approximated respectively by the computation of the probabilities of the form

$$\mathbf{P}\{\sup_{A\in\mathcal{A}} \|pr_{\mathbf{W}^{\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} (\mathbf{\Sigma}^{-1/2}\mathbf{h_g} + \mathbf{Z}^{H_0}_{p,\mathbf{f}})(A)\|_{\mathbb{R}^p} \geq q_{1-\alpha}\}$$

and

$$\mathbf{P}\{\int_{\mathbf{I}} \|pr_{\mathbf{W}^{\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} (\mathbf{\Sigma}^{-1/2}\mathbf{h_g} + \mathbf{Z}^{H_0}_{p,\mathbf{f}})(A)\|^2_{\mathbb{R}^p} dA \geq c_{1-\alpha}\}$$

where $q_{1-\alpha}$ and $c_{1-\alpha}$ are the $1-\alpha$ quantiles of the limiting distribution of $KS_{n,\mathcal{A}}$ and $CM_{n,\mathcal{A}}$ under H_0, respectively.

Remark 3 *In the application* $\mathbf{\Sigma}$ *is sometimes unknown. In this case* $\mathbf{\Sigma}$ *can be directly replaced with a consistent estimator without altering the asymptotic results, for example with that defined in Arnold [2], i.e.,* $\widehat{\mathbf{\Sigma}}_n := \widehat{\mathbf{\Sigma}}_n^{\mathbf{V}\top} \widehat{\mathbf{\Sigma}}_n^{\mathbf{V}}$.

In the computation we dealt with the partial sums process indexed by the family of closed rectangles $\mathcal{U} := \{[0,t] \times [0,s] : 0 \leq t \leq 1,\ 0 \leq s \leq 1\} \subset \mathcal{A}$ as a spacial case. The limit process is conveniently written as $\mathbf{Z}^{H_0}_{p,\mathbf{f}})(t,s)$ instead of $\mathbf{Z}^{H_0}_{p,\mathbf{f}}([0,t]\times[0,s])$. In the practice $\mathbf{Z}^{H_0}_{p,\mathbf{f}}(t,s)$ is approximated by the p-dimensional partial sums process

$$\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))([nt]/n, [ns]/n),\ (t,s) \in \mathbf{I},$$

where $[x] := \max\{z \in \mathcal{Z} :\ z \leq x\}$. The validity of the approximation is relied on the result that

$$\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))([nt]/n, [ns]/n) \xrightarrow{\mathcal{D}} \mathbf{B}^{H_0}_{p,\mathbf{f}}(t,s)$$

for every $(t,s) \in \mathbf{I}$, where $\mathbf{B}^{H_0}_{p,\mathbf{f}}(t,s) := pr_{\mathbf{W}^{*\perp}_{\mathcal{H}_{\mathbf{Z}_p}}} \mathbf{B}_p$. Thereby B_p is the p-dimensional standard Brownian (2) motion with the covariance function

$$K_{\mathbf{B}_p}((t,s);(t',s')) = (t \wedge t')(s \wedge s')\mathbf{I}_p.$$

To this end we refer the reader to Somayasa [21]. Furthermore, by the property of the partial sums, the analogues definition of the Kolmogorov-Smirnov and the Cramer-von Mises statistics are given respectively by

$$KS_n := \max_{1\leq \ell,k\leq n} \|\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))(\ell/n, k/n)\|_{\mathbb{R}^p},$$

$$CM_n := \frac{1}{n^2} \sum_{\ell,k=1}^{n} \|\mathbf{\Sigma}^{-1/2}\mathbf{V}_n(\mathbf{R}(\Xi_n))(\ell/n, k/n)\|^2_{\mathbb{R}^p}.$$

III. EXAMPLES OF THE LIMIT PROCESSES

For illustrative purpose we present examples of the limit process associated to the model under H_0 and H_1. The assumed regression models are restricted to the polynomial relationship between the response

variables and the explanatory variables.

A. Example 1

We consider the hypothesis $H_0 : \mathbf{g} \in [f_1]^p$, against $H_1 : \mathbf{g} \in [f_1, f_2, f_3]^p$, where for $(t, s) \in I$, $f_1(t, s) = 1$, $f_2(t, s) = t$, and $f_3(t, s) = s$. Hence for any $B \in \mathcal{A}$, $h_{f_1}(B) = \lambda(B)$ and $\langle h_{f_1}, Z^{(i)} \rangle = Z(I)$. Therefore the limit process under H_0 is given by

$$\mathbf{Z}^{H_0}_{p,\mathbf{f}_1}(B) = \mathbf{Z}_p(B) - \lambda(B)\mathbf{Z}_p(\mathbf{I}),\ B \in \mathcal{A}$$

having the covariance function

$$K_{\mathbf{Z}_{\mathbf{f}_1}}(B_1, B_2) = (\lambda(B_1 \cap B_2) - \lambda(B_1)\lambda(B_2))\mathbf{I}_p.$$

The limit process under $\mathcal{U}$ is give by

$$\mathbf{Z}^{H_0}_{p,\mathbf{f}_1}(t, s) = B_p(t, s) - tsB_p(1, 1),\ (t, s) \in \mathbf{I}$$

which is the well known p-dimensional Brownian (2) bridge, having the covariance function

$$K_{\mathbf{Z}_{\mathbf{f}_1}}((t, s), (t', s')) = ((t \wedge t')(s \wedge s') - tst's')\mathbf{I}_p.$$

For this type of H_1 we get the limit process as

$$\frac{1}{2}\mathbf{\Sigma}^{-1/2}(\widetilde{\beta}_2(t^2s - ts) + \widetilde{\beta}_3(ts^2 - ts)) + \mathbf{Z}^{H_0}_{p,\mathbf{f}_1}(t, s),$$

where for $j = 2, 3$, $\widetilde{\beta}_j = (\beta_{1j}, \cdots, \beta_{pj})^\top \in \mathbb{R}^p$ is a vector of unknown constants.

B. Example 2

In the second example we suppose that a first order multivariate model is observed under H_0, whereas under H_1 a second order multivariate model is assumed. That is we consider $H_0 : \mathbf{g} \in [f_1, f_2, f_3]^p$, against $H_1 : \mathbf{g} \in [f_1, f_2, f_3, f_4]^p$, with $f_1(t, s) = 1$, $f_2(t, s) = t$, $f_3(t, s) = s$ and $f_4(t, s) = ts$. The orthonormal set obtained from $\{f_1, f_2, f_3\}$ is given by $\tilde{f}_1(t, s) = 1$, $\tilde{f}_2(t, s) = \sqrt{3}(2t - 1)$, and $\tilde{f}_3(t, s) = \sqrt{3}(2s - 1)$. Since $\mathbf{Z}_p(t, s) = \mathbf{0}$ for $t = 0$ or $s = 0$, then by the definition of $\langle \cdot \rangle$ we get

$$\begin{aligned}
\langle h_{\tilde{f}_1}, Z^{(i)} \rangle &= Z^{(i)}(I), \\
\langle h_{\tilde{f}_2}, Z^{(i)} \rangle &= \sqrt{3}Z^{(i)}(\mathbf{I}) - 2\sqrt{3}\int_{[0,1]} Z^{(i)}(t, 1)dt, \\
\langle h_{\tilde{f}_3}, Z^{(i)} \rangle &= \sqrt{3}Z^{(i)}(I) - 2\sqrt{3}\int_{[0,1]} Z^{(i)}(1, s)ds.
\end{aligned}$$

Hence the corresponding p-dimensional set-indexed PSPLSR under H_0 is given by

$$\begin{aligned}
&\mathbf{Z}^{H_0}_{p,\mathbf{f}_3})(B) = \mathbf{Z}_p(B) - \lambda(B)\mathbf{Z}_p(I) \\
&-(\sqrt{3}\mathbf{Z}_p(I) - 2\sqrt{3}\int_{[0,1]} \mathbf{Z}_p(t, 1)dt)h_{\tilde{f}_2}(B) \\
&-(\sqrt{3}\mathbf{Z}_p(I) - 2\sqrt{3}\int_{[0,1]} \mathbf{Z}_p(1, s)ds)h_{\tilde{f}_3}(B).
\end{aligned}$$

The covariance function of this process is given by

$$\begin{aligned}
&K_{\mathbf{Z}_{\mathbf{f}_3}}(B_1, B_2) = K_{\mathbf{Z}_{\mathbf{f}_1}}(B_1, B_2) \\
&-(3\int_{B_1}(2t - 1)d\lambda \int_{B_2}(2t - 1)d\lambda)\mathbf{I}_p \\
&-(3\int_{B_1}(2s - 1)d\lambda \int_{B_2}(2s - 1)d\lambda)\mathbf{I}_p.
\end{aligned}$$

The ordinary version of the limit process takes the form

$$\begin{aligned}
&\mathbf{Z}^{H_0}_{p,\mathbf{f}_3})(t, s) = \mathbf{B}_p(t, s) - tsB_p(I) \\
&-6(\mathbf{B}_p(I) - 2\int_{[0,1]} \mathbf{B}_p(t, 1)dt)ts(t - 1) \\
&-6(\mathbf{B}_p(I) - 2\int_{[0,1]} \mathbf{B}_p(1, s)ds)ts(s - 1).
\end{aligned}$$

with the covariance function

$$\begin{aligned}
&K_{\mathbf{Z}_{\mathbf{f}_3}}((t, s), (t', s')) = K_{\mathbf{Z}_{\mathbf{f}_1}}((t, s), (t', s')) \\
&-3tst's'(t - 1)(t' - 1)\mathbf{I}_p - 3tst's'(s - 1)(s' - 1)\mathbf{I}_p.
\end{aligned}$$

After some algebraic computation we get the limit process for the localized model under H_1 as

$$\frac{1}{4}\mathbf{\Sigma}^{-1/2}\widetilde{\beta}_4(t^2s^2 - t^2s - ts^2 + ts) + \mathbf{Z}^{H_0}_{p,\mathbf{f}_3})(t, s),$$

for some unknown $\widetilde{\beta}_4 = (\beta_{14}, \ldots, \beta_{p4})^\top \in \mathbb{R}^p$.

IV. SIMULATION STUDY

In this section we investigate the finite sample size behavior of the tests by conducting Monte Carlo simulation. We consider three different cases.

A. Simulation 1

In our first simulation we consider the hypothesis defined in Example with $p = 2$. The samples are generated from the localized model

$$\mathbf{Y}_{n\ell k} = \mathbf{g}_{n\ell k}/n + \mathcal{E}_{n\ell k},$$

where for $1 \leq \ell, k \leq n$, $g^{(1)}(\ell/n, k/n) = 5 + \rho\ell/n + \rho k/n$ and $g^{(2)}(\ell/n, k/n) = 3 + \gamma\ell/n + \gamma k/n$. Hence, the sample is clearly from H_0 if and only if $\rho = 0 = \gamma$. Conversely, if $\rho \neq 0$ and $\gamma \neq 0$, then the sample is from H_1. In this example the vector of random errors $\mathcal{E}_{n\ell k}$ are drawn independently from a centered bivariate normal distribution with the covariance matrix

$$\mathbf{\Sigma}_1 = \begin{pmatrix} 6.26 & -0.50 \\ -0.50 & 6.25 \end{pmatrix}.$$

Table 1: Probabilities of rejection of H_0 (Simulation 1) for $\alpha = 0.01$, where $\mathcal{E}_{n\ell k} \sim N_2(\mathbf{0}, \mathbf{\Sigma}_1)$ simulated under 10000 runs.

$n \times n$	ρ	γ	$KS^{(2)}_{n;\mathcal{U}}$	$CM^{(2)}_{n;\mathcal{U}}$	Λ_n
			$\alpha = 0.01$		
30×30	0	0	0.0096	0.0156	0.0098
	5	10	0.0792	0.1216	0.1216
	5	30	0.8538	0.9580	0.9580
	10	5	0.0850	0.1178	0.1178
	30	5	0.8520	0.9624	0.9624
50×50	0	0	0.0099	0.0098	0.0089
	5	10	0.1048	0.1592	0.1184
	5	30	0.8928	0.9668	0.9602
	10	5	0.1054	0.1574	0.1184
	30	5	0.8882	0.9694	0.9610
70×70	0	0	0.0148	0.0112	0.0104
	5	10	0.1158	0.1648	0.1194
	5	30	0.9026	0.9648	0.9580
	10	5	0.1150	0.1556	0.1112
	30	5	0.9058	0.9710	0.9640

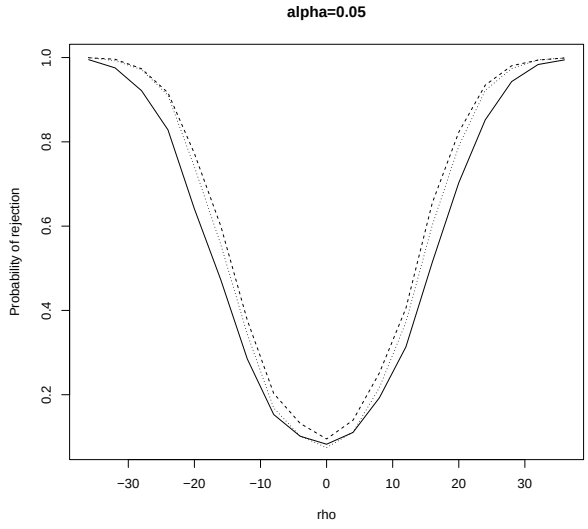

Figure 1: Approximated power functions of the K.S. (straight line), C.M. (dashed line), and L.R. (dotted line) tests for Simulation 1 with 50×50 lattice points simulated under 10000 runs

.

Table 2: Probabilities of rejection of H_0 (Simulation 1) for $\alpha = 0.05$, where $\mathcal{E}_{n\ell k} \sim N_2(\mathbf{0}, \mathbf{\Sigma}_1)$ simulated under 10000 runs.

$n \times n$	ρ	γ	$KS^{(2)}_{n;\mathcal{U}}$	$CM^{(2)}_{n;\mathcal{U}}$	Λ_n
			$\alpha = 0.05$		
30×30	0	0	0.0408	0.0544	0.0492
	5	10	0.2076	0.3276	0.2824
	5	30	0.9526	0.9914	0.9896
	10	5	0.2206	0.3316	0.2788
	30	5	0.9544	0.9888	0.9882
50×50	0	0	0.0460	0.0548	0.0508
	5	10	0.4550	0.5704	0.5334
	5	30	0.9678	0.9908	0.9886
	10	5	0.4498	0.5804	0.5338
	30	5	0.9676	0.9912	0.9910
70×70	0	0	0.0556	0.0480	0.0450
	5	10	0.2678	0.3240	0.2878
	5	30	0.9696	0.9906	0.9926
	10	5	0.2672	0.3222	0.2834
	30	5	0.9708	0.9892	0.9896

Table 1 and Table 2 present the probabilities of rejection of H_0 for $\alpha = 0.01$ and $\alpha = 0.05$, respectively, and for several varied number of lattice points, ρ, and γ. The simulation shows that for each choice of the sample size, when ρ is fixed the power of the tests increases as the values of γ get large. Similarly, for fixed γ, the power increases as the values of ρ get large. This means that the tests have a good ability to reject H_0 when the samples are actually from H_1. When H_0 is true the level of significance is well approximated by the three tests. Fig.1 exhibits the graphs of the power function of the tests for $\alpha = 0.05$. The curves are drawn under the model $g^{(1)}(\ell/n, k/n) = 5+\rho\ell/n+\rho k/n$ and $g^{(2)}(\ell/n, k/n) = 3+\rho\ell/n+\rho k/n$ by joining the points $(\rho, F(\rho))$ for 20 chosen values of ρ, where $F(\rho)$ is the percentage of time H_0 is rejected. In general there is a tendency that the L.R.-test (dotted line) and the C.M.-test (dashed line) almost have the same power. Among the three tests, the K.S. test, which is represented by a straight line, has the smallest power.

B. Simulation 2

In our next example we assume under H_0 the first-order model against a more general nonparametric alternative. We consider the case $p = 3$. Suppose the samples are generated independently from

Table 3: Probabilities of rejection of H_0 (Simulation 2) for $\alpha = 0.01$ and $\alpha = 0.05$ with 50×50 lattice points, where $\mathcal{E}_{n\ell k} \sim N_3(\mathbf{0}, \mathbf{\Sigma}_2)$ simulated under 10000 runs.

ρ	γ	δ	$KS^{(3)}_{n;\mathcal{U}}$	$CM^{(3)}_{n;\mathcal{U}}$	$KS^{(3)}_{n;\mathcal{U}}$	$CM^{(3)}_{n;\mathcal{U}}$
			$\alpha = 0.01$		$\alpha = 0.05$	
0	0	0	0.0094	0.0104	0.0378	0.0504
10	5	5	0.0124	0.0142	0.0494	0.0678
20	5	5	0.0202	0.0406	0.0778	0.1286
40	5	5	0.1092	0.3592	0.2856	0.5936
60	5	5	0.4650	0.9072	0.7478	0.9736
5	10	5	0.0106	0.0176	0.0510	0.0732
5	15	5	0.0114	0.0350	0.0664	0.1149
5	20	5	0.0190	0.0714	0.0932	0.2070
5	40	5	0.2664	0.7674	0.5378	0.9012
5	60	5	0.9054	0.9990	0.9814	0.9996
5	5	10	0.0208	0.0722	0.0934	0.2084
5	5	15	0.0788	0.3324	0.2336	0.5680
5	5	20	0.0274	0.7576	0.5540	0.9909
5	5	40	0.9996	1.0000	1.0000	1.0000

a three-variate norma random vector with the mean $\mathbf{g}(\ell/n, k/n)/n$, where for $1 \le \ell, k \le n$,

$$\mathbf{g}(\ell/n, k/n) = \begin{bmatrix} 2 + \frac{\ell}{n} + \frac{k}{n} + \rho \sin(\frac{lk}{n^2}) \\ -1 + 2\frac{\ell}{n} + 3\frac{k}{n} + \gamma \exp(\frac{lk}{n^2}) \\ 3 + \frac{\ell}{n} - 2\frac{k}{n} + \delta \exp(\frac{lk}{n^2}) \sin(\frac{kl}{n^2}) \end{bmatrix}$$

and the covariance matrix

$$\mathbf{\Sigma}_2 := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

Then H_0 holds true if and only if $\rho = \gamma = \delta = 0$.

The simulation results are presented respectively in Table 3 and Table 4 for 50×50 and 75×75 lattice points, respectively. Fig.2 exhibits the graph of the approximated power function of the tests simulated under the model

$$\mathbf{g}(\ell/n, k/n) = \begin{bmatrix} 2 + \frac{\ell}{n} + \frac{k}{n} + \rho \sin(\frac{lk}{n^2}) \\ -1 + 2\frac{\ell}{n} + 3\frac{k}{n} + \rho \exp(\frac{lk}{n^2}) \\ 3 + \frac{\ell}{n} - 2\frac{k}{n} + \rho \exp(\frac{lk}{n^2}) \sin(\frac{kl}{n^2}) \end{bmatrix}.$$

The simulation result shows that the power of the tests increases as the model moves away from H_0. More precisely, for each sample size there exist a tendency that the power increases when the values of either ρ, δ or γ get large. When ρ, γ and δ are set to zero, the power fluctuates around the nominal levels of the significance. We note that in this case the power of the K.S. and C.M.-tests can not be compared with that of the L.R.-test, because under H_1 we consider a non-parametric model. Based on Table 3, Table 4 and Fig.2 it can be concluded that the C.M.-test is more powerful than the K.S.-test independent to the sample sizes and also to the choice of α.

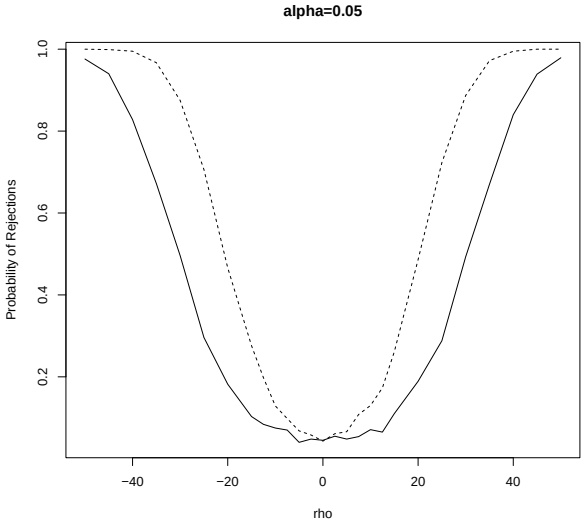

Figure 2: Approximated power functions of the K.S. (straight line) and C.M. (dashed line)-tests for Simulation 2 with 50×50 lattice points simulated under 10000 runs.

C. Simulation 3

For the second case we consider hypothesis of the form $H_0 : \mathbf{g} \in [f_1, f_2, f_3, f_4, f_5, f_6]^3$ against $H_1 : \mathbf{g} \in [f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8]^3$, where for $(t, s) \in \mathbf{I}$, $f_1(t, s) = 1$, $f_2(t, s) = t$, $f_3(t, s) = s$, $f_4(t, s) = t^2$, $f_5(t, s) = s^2$, $f_6(t, s) = ts$, $f_7(t, s) = t^3$, and $f_8(t, s) = s^3$. Suppose the data are generated from the localized model with the regression function $\mathbf{g} = (g^{(1)}, g^{(2)}, g^{(3)})^\top$ defined respectively by

$$5 - 10\frac{k}{n} + 15\frac{\ell}{n} + 6\frac{k^2}{n^2} + 8\frac{\ell^2}{n^2} + 5\frac{\ell k}{n^2} + \rho\frac{\ell^3}{n^3} + \rho\frac{k^3}{n^3}$$
$$6 + 5\frac{k}{n} + 10\frac{\ell}{n} + 6\frac{k^2}{n^2} - 5\frac{\ell^2}{n^2} + 5\frac{\ell k}{n^2} + \rho\frac{\ell^3}{n^3} + \rho\frac{k^3}{n^3}$$
$$10 - 5\frac{k}{n} + 5\frac{\ell}{n} + 6\frac{k^2}{n^2} - 10\frac{\ell^2}{n^2} - 5\frac{\ell k}{n^2} + \rho\frac{\ell^3}{n^3} + \rho\frac{k^3}{n^3}$$

Hence, the observations are clearly from H_0 if and only if $\rho = 0$, otherwise they are from H_1.

Table 4: Probabilities of rejection of H_0 (Simulation 2) for $\alpha = 0.01$ and $\alpha = 0.05$ with 75×75 lattice points, where $\mathcal{E}_{n\ell k} \sim N_3(\mathbf{0}, \mathbf{\Sigma}_2)$ simulated under 10000 runs.

ρ	γ	δ	$KS^{(3)}_{n;\mathcal{U}}$	$CM^{(3)}_{n;\mathcal{U}}$	$KS^{(3)}_{n;\mathcal{U}}$	$CM^{(3)}_{n;\mathcal{U}}$
			$\alpha = 0.01$		$\alpha = 0.05$	
0	0	0	0.0094	0.0080	0.0580	0.0458
15	5	5	0.0168	0.0222	0.0790	0.0856
20	5	5	0.0246	0.0386	0.0938	0.1316
40	5	5	0.1230	0.3512	0.3276	0.5974
60	5	5	0.5198	0.9126	0.7918	0.9750
5	10	5	0.0174	0.0172	0.0632	0.0740
5	15	5	0.0202	0.0308	0.0780	0.1170
5	20	5	0.0264	0.0668	0.1048	0.1990
5	40	5	0.3114	0.7596	0.5960	0.9006
5	60	5	0.9156	0.9996	0.9834	1.0000
5	5	10	0.0246	0.0726	0.1090	0.1950
5	5	15	0.0874	0.3240	0.2648	0.5462
5	5	20	0.2918	0.7474	0.5760	0.8920
5	5	40	1.0000	1.0000	0.9998	1.0000

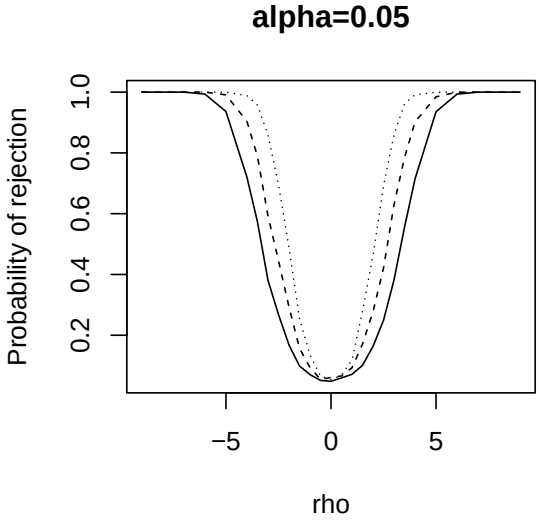

Figure 3: Approximated power functions of the K.S. (straight line), C.M. (dashed line), and L.R. (dotted line) tests for Simulation 3 with 50×50 lattice points simulated under 10000 runs.

Table 5: The Pearson's correlation matrix of the percentages of Ni, CaO and SiO_2.

	Ni	CaO	SiO2
Ni	$+1.0000$	-0.1285	-0.0004
CaO	-0.1285	$+1.0000$	$+0.3949$
SiO2	-0.0004	$+0.3949$	$+1.0000$

The graphs of the approximated power function for the three tests (K.S, C.M. and L.R. tests) are exhibited in Fig. 3. Compared to Simulation 1 and Simulation 2, there is also a similar tendency in Simulation 3 that the power of the tests increases as the values of ρ get large. The power of the tests increases as the model moves away from H_0. When $\rho = 0$, the power achieves the nominal levels of the significance. The simulation also shows that C.M.-test is more powerful than K.S.-test.

V. APPLICATION TO MINING DATA

In this section we apply the proposed method to a mining data provided by PT. ANTAM Tbk., studied in Tahir [23]. The company aims to predict the percentage of nickel (Ni) contained in the exploration region. The sample was obtained by drilling bores positioned according to a 7×14 regular lattice with 7 equidistance rows running west to east and 14 equidistance column running south to north. A laboratory analysis resulted in not only the percentage of Ni but at the same time there appeared also other substances such as calcium-monoxide (CaO) and Silicon-dioxide (SiO_2) as impurities that must be incorporated in the statistical analysis.

As preliminary investigation we present the pairs scatter plot of the percentages of Ni, CaO and SiO2 in Fig. 4. The Pearson's correlation coefficient among the three variables is exhibited in Table 5. It is shown that there exists strong correlation between Ni and CaO and between CaO and SiO_2. By this reason the correlation matrix must be taken into account in the statistical modelling of the variables. Furthermore from the individual scatter plot of the percentages of Ni, CaO and the logarithm of the percentage of SiO_2 presented respectively in Fig. 5, Fig. 6 and Fig. 7 we can infer that polynomial of lower order especially first-order model seems to be adequate to represent the data.

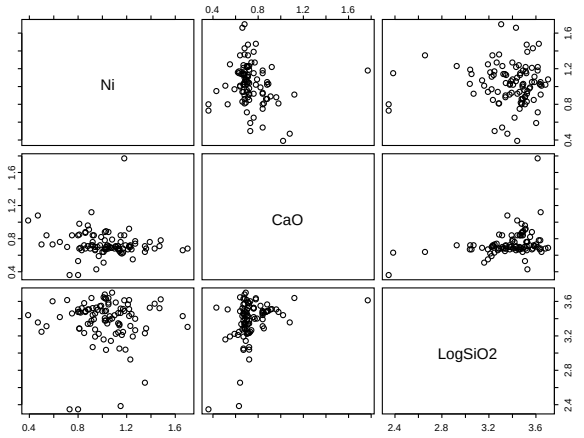

Figure 4: The pairs plot of the percentage of Ni, CaO and $LogSiO_2$. The data is compiled in the exploration region of PT. Antam Tbk in Southeast Sulawesi

.

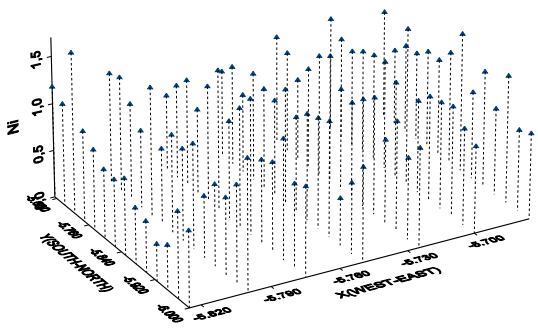

Figure 5: The scatter plot of the percentage of Ni

.

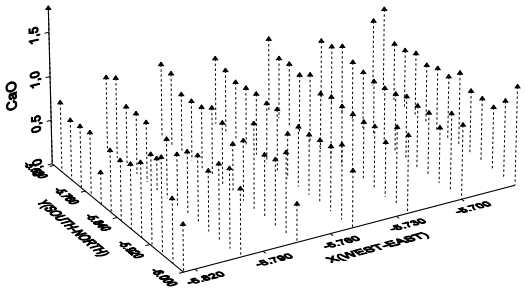

Figure 6: The scatter plot of the percentage of CaO

.

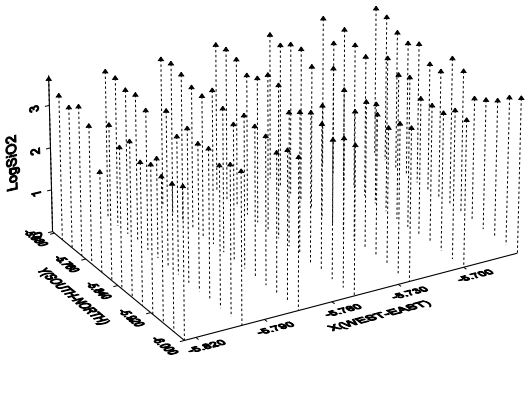

Figure 7: The scatter plot of the logarithm of the percentage of SiO_2

.

Table 6: The values of $KS^{(3)}_{7\times14}$ and $CM^{(3)}_{7\times14}$ for Ni, CaO and LogSiO2 associated to constant, first-order and second-order model.

	$KS^{(3)}_{7\times14}$		$CM^{(3)}_{7\times14}$	
Model	Data	p-value	Data	p-value
Constan	4.5244	0.0376	4.2910	0.0248
1st-order	2.0354	0.6216	0.6388	0.5913
2nd-order	2.7491	0.0174	0.7996	0.0381

The values of the statistics $KS^{(3)}_{7\times14}$ and $CM^{(3)}_{7\times14}$ for the Ni, CaO and SiO_2 data together with their approximated p-values for the assumed models are presented in Table 6 below. The constant as well as second-order polynomial models are clearly not plausible when tested using either K.S.-test or C.M.-test because both tests give relatively small p-values. The KS-test clearly rejects the constant (resp. second-order) model for $\alpha > 3.76\%$ (resp. $\alpha > 1.74\%$), whereas $C.M.$-test rejects the constant (resp. second-order) model for $\alpha > 2.48\%$ (resp. $\alpha > 3.81\%$). Next by the large p-values of the K.S. and C.M.-test associated with the first-order model, we conclude that the first-order polynomial is appropriate. Hence by the least squares method, the fitted model is

$$\begin{pmatrix} Ni(t,s) \\ CaO(t,s) \\ SiO_2(t,s) \end{pmatrix} = \begin{pmatrix} 0.8250 + 0.1507t + 0.2146s \\ 0.7533 + 0.0015t - 0.0316s \\ 3.4123 + 0.0063t - 0.0703s \end{pmatrix}$$

In the practice the company should use such a fitted model for the prediction of the unobserved points.

VI. CONCLUSION

The limit process of the sequence of p-dimensional set-indexed L.S.R.P.S.P. can be derived under a mild assumption by applying the vectorial version of Prohorov's theorem. Our technique leads to the same limit process as the geometric approach proposed in [6] for the spatial regression with single response and in [21] for MSR with correlated responses did. The limit processes can be represented and calculated in a simple way since it is a projection of the p-dimensional set-indexed Brownian sheet. The simulation study shows that the K.S. and C.M.-tests provide reasonable statistics to check the adequacy of the assumed models in the sense they achieve the level of significance as the LR-test did when the response is normally distributed. It is also shown by simulation that C.M.-test tends to have more power than the K.S.-test has. The application of the method to the Ni, CaO and SiO_2 data shows that the K.S. and C.M.-test have a good power in detecting the appropriateness of the assumed model to the sample. The MSR dealt with in this paper is that with correlated responses each of which has the same basis. In our future study we consider the model in which the responses have different bases with random experimental design rather than fixed.

APPENDIX

Theorem A.1 (Invariant Principle) For $n \geq 1$, let

$$\mathbf{X}_{n^d\times p} := \left\{ \mathbf{X}_{n\frac{\mathbf{j}}{n}} := (X^{(i)}_{n\frac{\mathbf{j}}{n}})_{i=1}^{p} : \ \mathbf{1} \leq \mathbf{j} \leq n\mathbf{1} \right\},$$

be a pyramidal arrays of independent and identically distributed random vectors with $E(\mathbf{X}_{n\frac{\mathbf{j}}{n}}) = \mathbf{0} \in \mathbb{R}^p$ and $Cov(\mathbf{X}_{n\frac{\mathbf{j}}{n}}) = \mathbf{\Sigma}$, where $\mathbf{\Sigma}$ is assumed to be positive definite, and $\mathbf{j} := (j_k)_{k=1}^{d}$. Let $\mathbf{Z}_p := (Z^{(i)})_{i=1}^{p}$ be a centered Gaussian process in $\mathcal{C}^p(\mathcal{A})$ with the covariance function

$$Cov(\mathbf{Z}_p(A), \mathbf{Z}_p(B)) = \lambda^d(A \cap B)\mathbf{I}_p, \ A, B \in \mathcal{A}.$$

Then it holds true

$$\mathbf{\Sigma}^{-1/2}\mathbf{V}^{(d)}_n(\mathbf{X}_{n^d\times p}) \xrightarrow{\mathcal{D}} \mathbf{Z}_p, \ n \to \infty,$$

where for $A \in \mathcal{A}$,

$$\mathbf{V}^{(d)}_n(\mathbf{X}_{n^d\times p})(A) := \sum_{\mathbf{1}\leq\mathbf{j}\leq n\mathbf{1}} \sqrt{n^d}\lambda^d(A \cap C_{\frac{\mathbf{j}}{n}})\mathbf{X}_{n\frac{\mathbf{j}}{n}},$$

with $C_{\frac{\mathbf{j}}{n}} := \times_{k=1}^{d}\left(\frac{j_k-1}{n}, \frac{j_k}{n}\right]$. Here $\mathbf{\Sigma}^{-1/2}$ denotes the root square of $\mathbf{\Sigma}^{-1}$, that is $\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{-1/2} = \mathbf{\Sigma}^{-1}$, $\mathcal{A}$ is the family of convex sets in $\mathbf{I}^d := \times_{j=1}^{d}[0,1] \times [0,1]$, and λ^d is the Lebesgue measure on $\mathbf{I}^d$.

Proof: (1) We show that the sequence of the finite dimensional distribution of $\mathbf{\Sigma}^{-1/2}\mathbf{V}^{(d)}_n(\mathbf{X}_{n^d\times p})$ converges to those of $\mathbf{Z}_p$. For this purpose we use the well-known multivariate Lindeberg-Feller central limit theorem (cf. [24], p. 20). Let $B_1, \ldots, B_m$ be any Borel subset in $\mathcal{A}$ and $a_1, \ldots, a_m$ be any constants. It must be shown for $n \to \infty$, that

$$\sum_{j=1}^{m} a_j\mathbf{\Sigma}^{-1/2}\mathbf{V}^{(d)}_n(\mathbf{X}_{n^d\times p})(B_j) \xrightarrow{\mathcal{D}} \sum_{j=1}^{m} a_j\mathbf{Z}_p(B_j),$$

where $\sum_{j=1}^m a_j \boldsymbol{Z}_p(B_j)$ follows a p-variate normal distribution having zero mean and the covariance matrix

$$Cov(\sum_{j=1}^m a_j \boldsymbol{Z}_p(B_j)) = \sum_{i=1}^m \sum_{j=1}^m a_i a_j \lambda^d(B_i \cap B_j)\mathbf{I}_p.$$

It suffices to show the covariance of

$$\mathbf{F}_n := \sum_{j=1}^m a_j \mathbf{\Sigma}^{-1/2} \mathbf{V}_n^{(d)}(\mathbf{X}_{n^d \times p})(B_j)$$

converges to that of $\sum_{j=1}^m a_j \boldsymbol{Z}_p(B_j)$, and $\mathbf{F}_n$ satisfies Lindeberg condition. By the definition of $\mathbf{V}_n^{(d)}$, $\mathbf{F}_n$ can be represented as

$$\begin{aligned}\mathbf{F}_n &= \sum_{j=1}^m a_j \mathbf{\Sigma}^{-1/2} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} \sqrt{n^d} \lambda^d(B_j \cap C_{\frac{\mathbf{j}}{n}}) \mathbf{X}_{n\frac{\mathbf{j}}{n}} \\ &= \frac{1}{\sqrt{n^d}} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}},\end{aligned}$$

where

$$\gamma_{n\frac{\mathbf{j}}{n}} := \sum_{i=1}^m a_i n^d \lambda^d(A_i \cap C_{\frac{\mathbf{j}}{n}}) = \sum_{i=1}^m a_i \frac{\lambda^d(A_i \cap C_{\frac{\mathbf{j}}{n}})}{\lambda^d(C_{\frac{\mathbf{j}}{n}})}.$$

Hence, by using the i.i.d. property of $\mathbf{X}_{n\frac{\mathbf{j}}{n}}$, for $\mathbf{1} \le \mathbf{j} \le n\mathbf{1}$, we have

$$Cov(\mathbf{F}_n) = \frac{1}{n^d} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} \gamma^2_{n\frac{\mathbf{j}}{n}} \mathbf{I}_p.$$

Let ψ_n be a nonnegative function on $\mathbf{I}^d$ defined by

$$\psi_n(\mathbf{t}) := \begin{cases} \gamma^2_{n\frac{\mathbf{j}}{n}};\ \mathbf{t} \in C_{\frac{\mathbf{j}}{n}} \\ 0\ ;\ otherwise \end{cases}$$

There exists an $M := \sum_{i=1}^m |a_i|$ such that $|\psi_n|$ is bounded uniformly by M^2. Also we have

$$\int_{\mathbf{I}^d} \psi_n(\mathbf{t})\ \lambda^d(d\mathbf{t})\mathbf{I}_p = Cov(\mathbf{F}_n)$$

Then by the higher dimensional Lebesgue density theorem (cf. [26], p. 148) applied to the ratio $\frac{\lambda^d(A_i \cap C_{\frac{\mathbf{j}}{n}})}{\lambda^d(C_{\frac{\mathbf{j}}{n}})}$, and the well-known Lebesgue dominated convergence theorem, we get

$$\begin{aligned}\lim_{n\to\infty} Cov(\mathbf{F}_n) &= \int_{\mathbf{I}^d} \lim_{n\to\infty} \psi_n(\mathbf{t}) \lambda^d(d\mathbf{t})\mathbf{I}_p \\ &= \int_{\mathbf{I}^d} (\sum_{j=1}^m a_j \mathbf{1}_{B_j})^2\ \lambda^d(d\mathbf{t})\mathbf{I}_p \\ &= \sum_{i=1}^m \sum_{j=1}^m a_i a_j \lambda^d(B_i \cap B_j)\mathbf{I}_p,\end{aligned}$$

where $\mathbf{1}_A$ stands for the indicator of a set A. Next we show Lindeberg condition. That is $\forall \varepsilon > 0$,

$$\begin{aligned}\lim_{n\to\infty} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} E(\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|^2_{\mathbb{R}^p} \times \\ \mathbf{1}_{\{\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|_{\mathbb{R}^p} > \varepsilon\}}) = 0.\end{aligned}$$

By the property of the Euclidean norm

$$\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|^2_{\mathbb{R}^p} \le \frac{M^2}{n^d} \| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|^2_{\mathbb{R}^p}.$$

Also $\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|_{\mathbb{R}^p} > \varepsilon$ implies

$$\| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|_{\mathbb{R}^p} \ge \frac{n^{d/2}\varepsilon}{M}.$$

Therefor we get

$$\begin{aligned}0 \le \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} E(\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|^2_{\mathbb{R}^p} \times \\ \mathbf{1}_{\{\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|_{\mathbb{R}^p} > \varepsilon\}}) \\ \le M^2 E(\| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{1}}{n}} \|^2_{\mathbb{R}^p} \mathbf{1}_{\{\| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{1}}{n}} \|_{\mathbb{R}^p} > \frac{\varepsilon n^{d/2}}{M}\}}).\end{aligned}$$

By the bounded convergence theorem (cf. Corollary 2.3.13 in [3]), we see that

$$\begin{aligned}0 \le \lim_{n\to\infty} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} E(\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|^2_{\mathbb{R}^p} \times \\ \mathbf{1}_{\{\| \frac{1}{\sqrt{n^d}} \gamma_{n\frac{\mathbf{j}}{n}} \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{j}}{n}} \|_{\mathbb{R}^p} > \varepsilon\}}) \\ \le M^2 E(\lim_{n\to\infty} \| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{1}}{n}} \|^2_{\mathbb{R}^p} \times \\ \mathbf{1}_{\{\| \mathbf{\Sigma}^{-1/2} \mathbf{X}_{n\frac{\mathbf{1}}{n}} \|_{\mathbb{R}^p} > \frac{\varepsilon n^{d/2}}{M}\}}) = 0.\end{aligned}$$

(2) We show that $\{\mathbf{V}_n(\mathbf{X}_{n^d \times p}) :\ n \ge 1\}$ is tight. Since for each $n \ge 1$, $\mathbf{V}_n(\mathbf{X}_{n^d \times p})$ has the sample path in $\mathcal{C}^p(\mathcal{A})$, the suitable definition of the modulus of continuity of $\mathbf{V}_n(\mathbf{X}_{n^d \times p})$ is

$$w(\mathbf{V}_n(\mathbf{X}_{n^d \times p}); \delta) := \sup_{\{A,B \in \mathcal{A}:\ d_{\lambda^d}(A,B) < \delta\}} \Upsilon(A, B),$$

where $\Upsilon : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_{\ge 0}$, defined by

$$\Upsilon(A, B) := \left\| \mathbf{V}_n(\mathbf{X}_{n^d \times p})(A) - \mathbf{V}_n(\mathbf{X}_{n^d \times p})(B) \right\|_{\mathbb{R}^p}$$

for any $A, B \in \mathcal{A}$. The i-th component of the p-dimensional process $\mathbf{V}_n(\mathbf{X}_{n^d \times p})$ is given by

$$\mathbf{V}_n^{(i)}(X^{(i)}_{n^d \times p}) := \frac{1}{\sqrt{n^d}} \sum_{\mathbf{1} \le \mathbf{j} \le n\mathbf{1}} \gamma^2_{n\frac{\mathbf{j}}{n}} X^{(i)}_{n\frac{\mathbf{j}}{n}}, i = 1, \ldots, p$$

that satisfies

$$w(\mathbf{V}_n(\mathbf{X}_{n^d\times p});\delta) \leq \sum_{i=1}^{p} w(\mathbf{V}_n^{(i)}(X^{(i)}_{n^d\times p});\delta).$$

Hence, in order to establish the tightness of the set-indexed process $\{\mathbf{V}_n(\mathbf{X}_{n^d\times p})(B) : B \in \mathcal{A}\}$ it suffices to show that $\mathbf{V}_n^{(i)}(X^{(i)}_{n^d\times p})$ is tight, for each $i = 1,\ldots,p$. We finish the proof by referring to the uniform central limit theorem for the one dimensional set-indexed process investigated in [1] and [17].

ACKNOWLEDGEMENT

The first author wishes to thank Prof. Bischoff the head of the Mathematical Department Catholic University Eichstaett-Ingolstadt for his constructive discussion during the preparation of the manuscript.

References:

[1] K. S. Alexander, and R. Pyke, "A uniform central limit theorem for set-indexed partial-sum processes with finite variance", *The Annals of Probability*, 14, 1986, pp. 582–597.

[2] S. F. Arnold, "The asymptotic validity of invariant procedures for the repeated measures model and multivariate linear model", *Journal of Multivariate Analysis*, 15, 1984, pp. 325–335.

[3] K. B. Athreya and S. N. Lahiri, *Measure Theory and Probability Theory*, Springer Science+Business Media, LLC., New York, 2006.

[4] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York, 1968.

[5] W. Bischoff, "structure of residual partial sums limit processes of linear regression models", *Theory of Stochastic Processes*, 2, 2002, pp. 23–28.

[6] W. Bischoff and W. Somayasa, "The limit of the partial sums process of spatial least squares residuals", *J. Multivariate Analysis*, 100, 2009, pp. 2167–2177.

[7] G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surface*, John Wiley and Sons, New York, 1987.

[8] J. A. Clarkson and C. R. Adams, "On definition of Bounded Variation for functions of two variables", *Transactions of the American Mathematical Society*, 5, 1933, pp. 824–854.

[9] R. Christensen, *Linear Models for Multivariate, Time Series, and Spatial Data*, Springer-Verlag New York, Inc., New York, 1991.

[10] Noel A. C. Cressie, *Statistics for Spatial Data*, John Wiley & Sons, New York, 1991.

[11] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis* (3rd edition), Prentice-Hall,Inc., New York, 1992.

[12] M. Kovarik and L. Sarga, "Implementing Control Charts to Corporate Financial Management", *WSEAS Trans. Math.*, 13, 2014, pp. 246–255.

[13] I. B. MacNeill, "Properties of partial sums of polynomial regression residuals with applications to test for change of regression at unknown times", *Ann. Statist.* 6, 1978, pp. 422–433.

[14] I. B. MacNeill, "Limit processes for sequences partial sums of regression residuals", *Ann. Probab.* 6, 1978, pp. 695–698.

[15] I. B. MacNeill and V. K. Jandhyala, *Change-point methods for spatial data, Multivariate Environmental Statistics eds. by G.P. Patil and C.R. Rao*, Elsevier Science Publishers B.V., 1993, pp. 298–306.

[16] J. N. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics (3rd edition)*, John Wiley & Sons Ltd., New York, 2007.

[17] W. J. Pyke, " uniform central limit theorem for partial sum processes indexed by sets", *Ann.Probab.* 79, 1983, pp. 219–240.

[18] W. Rudin, *Functional Analysis (2nd ed.)*, McGraw-Hill, Inc., New York, 1991.

[19] M. H. Shu and B. M. Hsu, "ARIMA Models Forecasting Cargo Throughput with Modified Seasonal", WSEAS Trans. Math., 13, 2014, pp. 171–181.

[20] D. W. Stroock, *A Concise Introduction to the Theory of Integration (2nd edition)*, Birkhäuser, Berlin, 1994.

[21] W. Somayasa, "On weak convergence of the partial sums processes of the least squares residuals of multivariate spatial regression", *J. Indones. Math.Soc.*, 20(2), 2014, pp. 77–94.

[22] W. Somayasa, Ruslan, E. Cahyono, L. O. Engkoimani, "Cheking adequateness of spatial regressions using set-indexed partial sums technique", *Far-east Journal of Mathematical Sciences*, 96(8), 2015, pp. 933–966.

[23] M. Tahir, "Prediction of the amount of nickel deposit based on the results of drilling bores on several points (case study: south mining region of PT. ANTAM Tbk., Pomalaa, Southeast Sulawesi", unpublished Bachelor Thesis, Halu Oleo University, Dept. of Mathematics, 2010.

[24] A. W. Van der Vaart, *Asymptotic Statistics*, Cambridge University Press, Cambridge, 1998.

[25] L. Xie and I. B. MacNeill, "Spatial residual processes and boundary detection", *South African Statist. J.* 4, 2006, pp. 33–53.

[26] A. C. Zaanen, *An Introduction to the Theory of Integration*, North Holand, Amsterdam, 1958.