Multivariate Recursive Residuals Partial Sums Processes and Its Application in Model Check for Multiresponse Regression

WAYAN SOMAYASA
Halu Oleo University
Department of Mathematics
Jalan Mokodompit, 93232 Kendari
INDONESIA
wayan.somayasa@uho.ac.id

Abstract: In this paper asymptotic test in multivariate regression based on set-indexed partial sums of the vector of recursive residuals is proposed. The limit process is derived for multivariate nonparametric regression with localized vector of regression functions under an equally spaced experimental design on a closed rectangle. Under mild condition it is shown that independent to the assumed model, the partial sums processes converges to a vector of trends plus the multivariate set-indexed Brownian sheet. The trend vanishes simultaneously when the hypothesis is true living the multivariate set-indexed Brownian sheet as the only limit process. The finite sample size behavior of the power functions of Kolmogorov-Smirnov (KS) and Cramér-von Mises (CM) type tests are investigated by simulation. It is shown that for testing multivariate polynomial model of low order the CM test seems to have larger power than the KS test has. The application of the test method in the empirical model building of corn plants data and its comparison with the classical test using Wilk's lambda statistic is also demonstrated.

Key–Words:multivariate linear regression, recursive residual, least squares residuals, partial sums process, multivariate Brownian sheet, Kolmogorov-Smirnov test, Cramér-von Mises test, Wilk's lambda.

1 Introduction

Checking the validity of multivariate linear regression using set-indexed partial sums of the vector of ordinary least squares (OLS) residuals has been intensively studied in [16, 17, 18]. The limit processes have been derived by applying various mathematical techniques. The result in [16] has been obtained by the geometric approach of [4], whereas those in [17, 18] have been established based on Prohorov's theorem (cf. [3]). However, the limit processes have been expressed as complicated functions of the multivariate set-indexed Brownian sheets which functionally depend not only on the assumed regression functions but also on the dimension of the model. Consequently, the application of the method in the practice is restricted. The quantiles of the Kolmogorov-Smirnov as well as the Cramér-von Mises type test statistics have been approximated by conducting Monte Carlo simulation. Estimation to the upper and lower bounds of the limiting power functions of the tests by appying Li-Kuelb's shift inequality have been studied in [19, 20], see also [13], pp. 53–54.

In this paper we propose an asymptotic model validity check based on the limit process of the partial sums of the vector of recursive residuals instead of the OLS residuals. Our object of study is actually an extension of that considered in [21]. The derivation of the result will be mathematically more complicated by the existence of the correlation among the components of the vector of observations. It is not like the classical likelihood ratio test studied in e.g. [6], p. 9–20 and [10], p. 395–398, for our test method we need neither normality nor any other distributional assumption given to the vector of observations, so that our proposed method is more practice.

The application of the partial sums of the recursive residuals of multivariate regression in some test problem has been investigated in [9] by extending the approaches due to [5, 7, 8, 12, 15]. However the limit process has been obtained only for vector of time series observations. For our result we need to investigate a functional central limit theorem for the sequence of high dimensional triangular arrays of vector of recursive residuals. To the best knowledge of the author, there are no documentations available for such kind of limit theorem.

To see concretely the objective of the present paper, let us consider a standard p-variate nonparametric regression defined by

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

where $\widetilde{\mathbf{Y}} = (Y^{(1)}, \ldots, Y^{(p)})^\top$ is the p-dimensional vector of observations, $\widetilde{\mathbf{g}} = (g^{(1)}, \ldots, g^{(p)})^\top$ is the true-unknown p-dimensional vector of regression functions defined on $\mathbf{D} := \Pi_{j=1}^{d}[a_j, b_j] \subset \mathcal{R}^d$, and $\widetilde{\mathcal{E}} = (\varepsilon^{(1)}, \ldots, \varepsilon^{(p)})^\top$ is the p-dimensional vector of random errors with $E(\widetilde{\mathcal{E}}) = \mathbf{0}$ and $Cov(\widetilde{\mathcal{E}}) = \mathbf{\Sigma} = (\sigma_{k\ell})_{k=1,\ell=1}^{p,p}$. Thereby $\mathbf{\Sigma}$ is a $p \times p$ dimensional positive definite matrix. Let $f_1, \ldots, f_q$ be known linearly independent regression functions in $L_2(P_0)$, where P_0 is the Lebesque measure on $\mathbf{D}$. Model validity check in multivariate linear regression analysis concerns with the problem of testing the following hypothesis

$$H_0 : Y^{(i)}(\mathbf{t}) = \sum_{j=1}^{q} \beta_{ij} f_j(\mathbf{t}) + \varepsilon^{(i)}(\mathbf{t}),\ \mathbf{t} \in \mathbf{D}, \quad (2)$$

for $i = 1, \ldots, p$, where $\beta_{i1}, \ldots, \beta_{iq}$ are unknown constants, cf [10], p. 323. Suppose Model 2 is observed independently over a triangular array of design points

$$\Xi_{n_1 \times n_2} := \{\mathbf{t}_{j_1 j_2} : 1 \le j_1 \le n_1,\ 1 \le j_2 \le n_2\}.$$

Let $\mathbf{R}_{n_1 n_2} := (\mathbf{r}_{j_1 j_2})_{j_1=1,j_2=1}^{n_1,n_2}$, be the corresponding sequence of $n_1 \times n_2$ arrays of the p-dimensional vector of the OLS residuals of Model 1, where $\mathbf{r}_{j_1 j_2} = (r_{j_1 j_2}^{(1)}, \ldots, r_{j_1 j_2}^{(p)})^\top$. The p-dimensional partial sums process of $\mathbf{R}_{n_1 n_2}$ indexed by the family of convex subsets $\mathcal{A}$ of $\mathbf{D}$ is defined by

$$\mathbf{V}_{n_1 n_2}(\mathbf{R}_{n_1 n_2})(A) = \left(\sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \mathbf{1}_A(t_{j_1 j_2}) r_{j_1 j_2}^{(i)} \right)_{i=1}^{p}.$$

By combining the univariate invariance principle of [1, 14] and the multivariate one of [11], Somayasa and et al. [16, 17] showed under an equidistance design (regular lattice), defined by

$$\mathbf{t}_{j_1 j_2} = \left(a_1 + (b_1 - a_1)\frac{j_1}{n_1}, a_2 + (b_2 - a_2)\frac{j_2}{n_2} \right),$$

that

$$\frac{1}{\sqrt{n_1 n_2}} \mathbf{\Sigma}^{-1/2} \mathbf{V}_{n_1 n_2}(\mathbf{R}_{n_1 n_2}) \Rightarrow \mathbf{Z}_{P_0}^*, \quad (3)$$

where $\mathbf{Z}_{P_0}^*$ is a p-dimensional centered Gaussian process indexed by $\mathcal{A}$, defined by

$$\mathbf{Z}_{P_0}^* = \mathbf{Z}_{P_0} - \left(\sum_{j=1}^{q} \int_{\mathbf{D}}^{R} f_j(t,s) dZ_{P_0}^{(i)}(t,s) h_{f_j} \right)_{i=1}^{p}.$$

Thereby $\mathbf{Z}_{P_0} = \left(Z_{P_0}^{(1)}, \ldots, Z_{P_0}^{(p)} \right)^\top$ is the p-dimensional set-indexed Brownian sheet, which is a centered p-dimensional Gaussian process with the covariance function given by

$$K_{\mathbf{Z}_{P_0}}(A_1, A_2) = diag(P_0(A_1 \cap A_2), \ldots, P_0(A_1 \cap A_2))$$

and $h_{f_j}(A) := \int_A f_j(x,y) P_0(dx,dy)$, for $A_1, A_2 \in \mathcal{A}$. Thereby $\int^R$ stands for the Riemann-Stiltjes integral. By the dependency of the limit process on the regression functions, the computation of the critical values of the Kolmogorov-Smirnov and Cramér-von Mises statistics defined respectively by

$$\begin{aligned} KS(\mathbf{Z}_{P_0}) &:= \sup_{(t,s)\in\mathbf{D}} \|\mathbf{Z}_{P_0}^*(t,s)\| \\ CvM(\mathbf{Z}_{P_0}) &:= \int_{\mathbf{D}} \|\mathbf{Z}_{P_0}^*(t,s)\|^2 P_0(dt,ds) \end{aligned}$$

become complicated as the order of the model gets higher. Unfortunately there are no documentations available providing mathematical techniques for computing such critical values for the case other than Brownian sheet. A good survey for the case of multivariate standard Brownian motion and Brownian bridge on the unit interval $[0,1]$ can be found in [11]. Thus, the main point of this work is to define a transformation so that the limit process in (3) does not depend on the regression models. This will be handled recursively as in the times series case investigated in [9] by extending the functional central limit theorem studied in [21].

It is worth mentioning that for testing the hypothesis H_0 against a specific alternative of the form $H_1 : g^{(i)} \in \mathbf{V} := [f_1, \ldots, f_q, f_{q+1}, \ldots, f_m]$, for all $i \in \{1, \ldots, p\}$, under normally distributed error terms the likelihood ration test leads to the test based on the well known Wilk's lambda statistic (cf. [10], p.324). However, there is no information available regarding the optimality of the test.

To give more insight on the multivariate recursive residuals we present in Section 2 its formal definition and investigate the related properties. The limit of the sequence of p-dimensional partial sums processes of recursive residuals is discussed in Section 3. Next in Section 4 we give an investigation to the performance of the KS and CM test by simulation. The application of the test method in real data is presented in Section 5. We close the paper with some conclusions and remarks for future works, see Section 6. Proofs are presented in the appendix.

2 Multivariate recursive residuals

As in univariate case, we assume throughout the paper that the design is given by a regular lattice over $\mathbf{D}$ and the observations are collected row wise initialized at the point $\mathbf{t}_{11}$ and terminated at the point $\mathbf{t}_{n_1 n_2}$.

Let (j_1^*, j_2^*) be a fixed pair of integers such that $\mathbf{Y}_{j_1^* j_2^*}$ becomes the first q-th observed vector of responses according to the preceding order. Keeping this in mind, let us define the following notations:

$$\begin{aligned} \mathbf{T}_{n_1 n_2} &:= \{(j_1, j_2) :\ 1 \le j_1 \le n_1,\ 1 \le j_2 \le n_2\} \\ \mathbf{T}_{n_1 n_2 - q} &:= \mathbf{T}_{n_1 n_2 - q + 1} \setminus \{(j_1^*, j_2^*)\}, \end{aligned}$$

where

$$\mathbf{T}_{n_1n_2-q+1}=\{(j_1^*,j_2^*),(j_1^*+1,j_2^*),\cdots,(n_1,j_2^*),\\ (1,j_2^*+1),\cdots,(n_1,j_2^*+1),\cdots,(n_1,n_2)\}.$$

Thus $\mathbf{T}_{n_1n_2},\mathbf{T}_{n_1n_2-q}$ and $\mathbf{T}_{n_1n_2-q+1}$ consist respectively of n_1n_2, n_1n_2-q and n_1n_2-q+1 ordered pairs. For every pair $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q+1}$ and every $i=1,\ldots,p$, we define the following vector of observations and random errors

$$\mathbf{Y}^{(i)}_{j_1j_2} = \left(Y^{(i)}_{11},Y^{(i)}_{21},\ldots,Y^{(i)}_{j_1j_2}\right)^\top$$

$$\mathcal{E}^{(i)}_{j_1j_2} = \left(\varepsilon^{(i)}_{11},\varepsilon^{(i)}_{21},\ldots,\varepsilon^{(i)}_{j_1j_2}\right)^\top.$$

The n_1n_2-q+1 numbers of p-variate regression models under H_0 are defined by

$$\mathbf{Y}^{(n_1,n_2)}_{j_1j_2} := \mathbf{X}^{(n_1,n_2)}_{j_1j_2}\mathbf{B}+\mathbf{E}^{(n_1,n_2)}_{j_1j_2}, \qquad (4)$$

where

$$\mathbf{Y}^{(n_1,n_2)}_{j_1j_2} := \left(\mathbf{Y}^{(1)}_{j_1j_2},\mathbf{Y}^{(2)}_{j_1j_2},\ldots,\mathbf{Y}^{(p)}_{j_1j_2}\right),$$

$$\mathbf{X}^{(n_1,n_2)}_{j_1j_2} := \left(\mathbf{f}(t_{11}),\mathbf{f}(t_{21}),\ldots,\mathbf{f}(t_{j_1j_2})\right)^\top,$$

$$\mathbf{E}^{(n_1,n_2)}_{j_1j_2} := \left(\mathcal{E}^{(1)}_{j_1j_2},\mathcal{E}^{(2)}_{j_1j_2},\ldots,\mathcal{E}^{(p)}_{j_1j_2}\right).$$

Then the OLS of $\mathbf{B}$ based on (4) is given by

$$\widehat{\mathbf{B}}^{(n_1,n_2)}_{j_1j_2}=\left(\mathbf{X}^{(n_1,n_2)\top}_{j_1j_2}\mathbf{X}^{(n_1,n_2)}_{(j_1j_2)}\right)^{-1}\mathbf{X}^{(n_1,n_2)\top}_{j_1j_2}\mathbf{Y}^{(n_1,n_2)}_{(j_1j_2)}$$

for every $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q+1}$. It is noticed that $\left(\mathbf{X}^{(n_1,n_2)\top}_{j_1j_2}\mathbf{X}^{(n_1,n_2)}_{(j_1j_2)}\right)^{-1}$ exists for large enough n_1 and n_2, provided the regression functions are linearly independent as functions in $L_2(P_0)$, cf. [22], p. 37.

Now we are ready to define the notion of p-variate recursive residuals of Model 2.

Definition 1 *Let $\widetilde{\mathbf{Y}}_{j_1j_2} := \left(Y^{(1)}_{j_1j_2},Y^{(2)}_{j_1j_2},\ldots,Y^{(p)}_{j_1j_2}\right)^\top$ be the vector of observations of Model 2 on the point $\mathbf{t}_{j_1j_2}$, for $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$. The n_1n_2-q numbers of vectors of recursive residuals based on Model 4 are defined by*

$$\widetilde{\mathbf{w}}_{j_1j_2}=\left(w^{(1)}_{j_1j_2},w^{(2)}_{j_1j_2},\ldots,w^{(p)}_{j_1j_2}\right)^\top := \frac{\widetilde{\mathbf{Y}}_{j_1j_2}-\widehat{\mathbf{B}}^{(n_1,n_2)\top}_{j_1-1j_2}\mathbf{f}(\mathbf{t}_{j_1j_2})}{\sqrt{1+\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\left(\mathbf{X}^{(n_1,n_2)\top}_{j_1-1j_2}\mathbf{X}^{(n_1,n_2)}_{j_1-1j_2}\right)^{-1}\mathbf{f}(\mathbf{t}_{j_1j_2})}},$$
$$for\ j_1\neq 1,\ j_2\in\{j_2^*,j_2^*+1,\ldots,n_2\},$$

and

$$\widetilde{\mathbf{w}}_{1j_2}=\left(w^{(1)}_{1j_2},w^{(2)}_{1j_2},\ldots,w^{(p)}_{1j_2}\right)^\top := \frac{\widetilde{\mathbf{Y}}_{1j_2}-\widehat{\mathbf{B}}^{(n_1,n_2)\top}_{n_1j_2-1}\mathbf{f}^\top(\mathbf{t}_{1j_2})}{\sqrt{1+\mathbf{f}^\top(\mathbf{t}_{1j_2})\left(\mathbf{X}^{(n_1,n_2)\top}_{n_1j_2-1}\mathbf{X}^{(n_1,n_2)}_{n_1j_2-1}\right)^{-1}\mathbf{f}(\mathbf{t}_{1j_2})}},$$
$$for\ j_1=1,\ and\ j_2\in\{j_2^*+1,\ldots,n_2\}.$$

By substituting $\widehat{\mathbf{B}}^{(n_1,n_2)}_{j_1j_2}$, for $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$, it can be shown that the i-th component of $\widetilde{\mathbf{w}}_{j_1j_2}$ coincides with that of the univariate case, see [21]. Hence, the argument that has been applied for univariate model can be adopted to get the following result.

Proposition 2 *For every $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$, there exists an $(n_1n_2-q)\times n_1n_2$ matrix $\mathbf{A}$, defined by*

$$\mathbf{A}:=\begin{pmatrix}\mathbf{a}^\top_{j_1^*+1j_2^*}\\ \vdots\\ \mathbf{a}^\top_{n_1j_2^*}\\ \mathbf{a}^\top_{1j_2^*+1}\\ \vdots\\ \mathbf{a}^\top_{n_1j_2^*+1}\\ \vdots\\ \mathbf{a}^\top_{1n_2}\\ \vdots\\ \mathbf{a}^\top_{n_1n_2}\end{pmatrix}\in\mathcal{R}^{(n_1n_2-q)\times n_1n_2},$$

such that

$$\left(\widetilde{\mathbf{w}}_{j_1^*+1j_2^*},\widetilde{\mathbf{w}}_{n_1j_2^*},\ldots,\widetilde{\mathbf{w}}_{n_1n_2}\right)^\top=\mathbf{A}\mathbf{E}_{n_1n_2},$$

where for $j_2\in\{j_2^*,j_2^*+1,\ldots,n_2\}$, and $j_1\neq 1$,

$$\mathbf{a}_{j_1j_2}\sqrt{d_{j_1j_2}} := \\ \left(-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\left(\mathbf{X}^{(n_1,n_2)\top}_{j_1-1j_2}\mathbf{X}^{(n_1,n_2)}_{j_1-1j_2}\right)^{-1}\mathbf{X}^{(n_1,n_2)\top}_{j_1-1j_2},\right.\\ \left.1,0,\ldots,0\right)^\top,$$

with

$$d_{j_1j_2}=1+\mathbf{f}^\top(\mathbf{t}_{j_1j_2})(\mathbf{X}^{(n_1,n_2)\top}_{(j_1-1j_2)}\mathbf{X}^{(n_1,n_2)}_{(j_1-1j_2)})^{-1}\mathbf{f}(\mathbf{t}_{j_1j_2}),$$

and for $j_1=1$ and $j_2\in\{j_2^*+1,\ldots,n_2\}$,

$$\mathbf{a}_{1j_2}\sqrt{d_{1j_2}} := \\ \left(-\mathbf{f}^\top(\mathbf{t}_{1j_2})\left(\mathbf{X}^{(n_1,n_2)\top}_{n_1j_2-1}\mathbf{X}^{(n_1,n_2)}_{n_1j_2-1}\right)^{-1}\mathbf{X}^{(n_1,n_2)\top}_{n_1j_2-1},\right.\\ \left.1,0,\ldots,0\right)^\top$$

with

$$d_{1j_2}:=1+\mathbf{f}^\top(\mathbf{t}_{1j_2})\left(\mathbf{X}^{(n_1,n_2)\top}_{n_1j_2-1}\mathbf{X}^{(n_1,n_2)}_{n_1j_2-1}\right)^{-1}\mathbf{f}(\mathbf{t}_{1j_2}).$$

Moreover, $\mathbf{A}$ satisfies the condition $\mathbf{A}\mathbf{A}^\top=\mathbf{I}_{n_1n_2-q}$, where $\mathbf{I}_{n_1n_2-q}$ is the $(n_1n_2-q)\times(n_1n_2-q)$ identity matrix.

By Proposition 2 every component of the vector of recursive residuals is expressible as a linear combination of the components of the vector of random errors $\mathcal{E}^{(i)}_{n_1n_2}$, for every $(j_1, j_2) \in \mathbf{T}_{n_1n_2-q}$, where the coefficients depend only on the pair $(j_1, j_2) \in \mathbf{T}_{n_1n_2-q}$, but not on $i \in \{1, \ldots, p\}$.

Let $\widetilde{\mathbf{W}}_{n_1n_2} := (\widetilde{\mathbf{w}}_{j_1j_2})^{n_1,n_2}_{j_1=1,j_2=1}$, $n_1 \geq 1$ and $n_2 \geq 1$ be the sequence of $n_1 \times n_2$ dimensional triangular arrays of the p-dimensional vectors of recursive residuals, where $\widetilde{\mathbf{w}}_{j_1j_2} := \mathbf{0}$, for $(j_1, j_2) \in \mathbf{T}_{n_1n_2} \setminus \mathbf{T}_{n_1n_2-q}$. Let $\mathcal{A}$ be the collection of convex subsets of $\mathbf{D} := [a_1, b_1] \times [a_2, b_2]$, such that $\mathcal{A}$ is totally bounded and have convergence integral entropy in the sense of [1, 14]. The set-indexed partial sums processes of the p-dimensional recursive residuals with respect to $\mathcal{A}$ is defined as

$$S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(A) := \sum_{(j_1,j_2)} \mathbf{1}_A(\mathbf{t}_{j_1j_2})\widetilde{\mathbf{w}}_{j_1j_2},$$

where the sum is defined component wise taken over all $(j_1, j_2) \in \mathbf{T}_{n_1n_2-q}$. It is worth mentioning that for $A = \emptyset$, we define $S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1\times n_2})(\emptyset) := 0$ and for any $A \in \mathcal{A}$, for which no design points $\mathbf{t}_{j_1j_2}$ with the corresponding pair $(j_1, j_2) \in \mathbf{T}_{n_1n_2-q}$ are covered by A, we define $S^{(p)}_{n_1n_2-p}(\mathbf{W}_{n_1\times n_2})(A) := 0$. To be able to sum the recursive residuals over all pairs $(j_1, j_2) \in \mathbf{T}_{n_1n_2}$, the value of $\widetilde{\mathbf{w}}_{j_1j_2}$ is set equal to zero, for $(j_1, j_2) \in \mathbf{T}_{n_1n_2} \setminus \mathbf{T}_{n_1n_2-q}$. So that we have

$$S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(A) := \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} \mathbf{1}_A(\mathbf{t}_{j_1j_2})\widetilde{\mathbf{w}}_{j_1j_2}.$$

The Kolmogorov-Smirnov and Cramér-von Mises type statistics reasonable for testing H_0 are defined respectively by

$$\mathcal{KS}^{(p)}_{n_1n_2} := \sup_{A\in\mathcal{A}} \left\| \frac{\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(A)}{\sqrt{n_1n_2-q}} \right\|$$

$$\mathcal{CM}^{(p)}_{n_1n_2} := \sum_{A\in\mathcal{A}} \left\| \frac{\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(A)}{n_1n_2-q} \right\|^2,$$

where $\mathbf{\Sigma}^{1/2}$ is the matrix that satisfies $\mathbf{\Sigma} = \mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2}$ and $\mathbf{\Sigma}^{-1/2} = (\mathbf{\Sigma}^{1/2})^{-1}$. Based on those statistics, H_0 will be rejected for large values of $\mathcal{KS}^{(p)}_{n_1n_2}$ or $\mathcal{CM}^{(p)}_{n_1n_2}$. The limit distributions of $\mathcal{KS}^{(p)}_{n_1n_2}$ and $\mathcal{CM}^{(p)}_{n_1n_2}$ will be derived in the next section. Since Proposition A1 states that the set $\{\widetilde{\mathbf{w}}_{j_1j_2} : (j_1, j_2) \in \mathbf{T}_{n_1n_2-q}\}$ consists of mutually uncorrelated p-dimensional random vectors with zero mean vector and the covariance matrix $\mathbf{\Sigma}$, by applying a well known result for multivariate normal distribution, cf. [10], p. 137, for $1 \leq j_1 \leq n_1$ and $1 \leq j_2 \leq n_2$, if the p-dimensional error vectors $\widetilde{\mathcal{E}}_{j_1j_2} = (\varepsilon^{(1)}_{j_1j_2}, \ldots, \varepsilon^{(p)}_{j_1j_2})^\top$ are assumed to be independent and identically distributed (iid) $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$, then $\widetilde{\mathbf{w}}_{j_1,j_2}$ are iid $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ too. Hence, for normally distributed error model, the limit of $\mathcal{KS}^{(p)}_{n_1n_2}$ and $\mathcal{CM}^{(p)}_{n_1n_2}$ can be straightforwardly obtained by incorporating the techniques proposed in [16, 17, 18]. However for the more general setting that methods can not be applied. In this work the limit will be established by generalizing the functional central limit theorem studied in [9] and [15].

For computational reason, in this paper we restrict the index set $\mathcal{A}$ to the family $\Re$, defined by

$$\Re := \{[a_1, x] \times [a_2, y] : \ a_1 \leq x \leq b_1, \ a_2 \leq y \leq b_2\}.$$

Hence, for every $[a_1, x] \times [a_2, y] \in \Re$, we have

$$\begin{aligned} &S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})([a_1, x] \times [a_2, y]) \\ &= \sum_{(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}} \mathbf{1}_{[a_1,x]\times[a_2,y]}(\mathbf{t}_{j_1j_2})\widetilde{\mathbf{w}}_{j_1j_2}, \end{aligned}$$

and it will be denoted as $S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(x, y)$ for brevity. Hence, the corresponding test statistics are given respectively by

$$\mathcal{KS}^{(p)}_{n_1n_2} = \sup_{(x,y)\in\mathbf{D}} \left\| \frac{\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(x, y)}{\sqrt{n_1n_2-q}} \right\|$$

$$\mathcal{CM}^{(p)}_{n_1n_2} = \int_{\mathbf{D}} \left\| \frac{\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(x, y)}{\sqrt{n_1n_2-q}} \right\|^2 dxdy.$$

In particular, for the case of a unit square experimental region $\mathcal{U} := [0, 1] \times [0, 1]$, the partial sums process indexed by such family of rectangles will take the form

$$S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(x, y) = \sum_{j_2=1}^{[n_2y]}\sum_{j_1=1}^{[n_1x]} \widetilde{\mathbf{w}}_{j_1j_2}, \ (x, y) \in \mathcal{U}.$$

So, when the experimental design is given by the regular lattice $\Xi_{n_1\times n_2}$ over $\mathcal{U}$

$$\Xi_{n_1\times n_2} = \{(\ell/n_1, k/n_2) : \ 1 \leq \ell \leq n_1, \ 1 \leq k \leq n_2\},$$

the property of the partial sums operator implies

$$\mathcal{KS}^{(p)}_{n_1n_2} = \max_{1\leq k\leq n_2: 1\leq \ell\leq n_1} \left\| \frac{\mathbf{\Sigma}^{-1/2}\sum_{j_1,j_2=1}^{\ell,k} \widetilde{\mathbf{w}}_{j_1j_2}}{\sqrt{n_1n_2-q}} \right\|$$

$$\mathcal{CM}^{(p)}_{n_1n_2} = \sum_{k,\ell=1}^{n_2,n_1} \left\| \frac{\mathbf{\Sigma}^{-1/2}\sum_{j_1,j_2=1}^{\ell,k} \widetilde{\mathbf{w}}_{j_1j_2}}{n_1n_2-q} \right\|^2.$$

In the practice, the computation of the test statistics will be based on the partial sums indexed by the family $\Re$.

3 The partial sums limit process

In this section we state the main result that gives the limit process of the sequence of the set-indexed p-dimensional partial sums processes of the recursive residuals:

$$\frac{1}{\sqrt{n_1n_2-q}}\mathbf{\Sigma}^{-1/2}S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})(A),\ A\in\mathcal{A}.$$

The proof is postponed to the appendix.

Theorem 3 *Let $\left\{\widetilde{\mathbf{W}}_{n_1\times n_2}=(\widetilde{\mathbf{w}}_{j_1j_2})^{n_1,n_2}_{j_1=1,j_2=1}\right\}$, $n_1\geq 1$ and $n_2\geq 1$, be the sequence of $n_1\times n_2$ arrays of the p-dimensional recursive residuals of Model 2 observed over a regular lattice $\Xi_{n_1\times n_2}$. Suppose that the regression functions are continuous and have bounded variation on $\mathbf{D}$. Then for $n_1,n_2\to\infty$, it holds,*

$$\frac{1}{\sqrt{n_1n_2-q}}\mathbf{\Sigma}^{-1/2}S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}_{n_1n_2})\Rightarrow\mathbf{Z}_{P_0},$$

where $\mathbf{Z}_{P_0}$ is the p-dimensional set-indexed Brownian sheet.

Theorem 3 shows that the limit process of the partial sums of the recursive residuals under H_0 is given by the p-dimensional set-indexed Brownian sheet, whatever the regression functions we have. This means that the transformation defined by the recursive residuals reduces the dependency of the limit process on the assumed model. Theoretically this will give advantage particularly in the computation of the quantiles of the test statistics.

The limit of the test statistics $\mathcal{KS}^{(p)}_{n_1n_2}$ as well as $\mathcal{CM}^{(p)}_{n_1n_2}$ can be readily obtained by applying the continuous mapping theorem (cf. Theorem 27 in [3]), as stated in the following corollary.

Corollary 4 *Under the conditions of Theorem 3, it holds for n_1 and n_2 are simultaneously large, that*

$$\mathcal{KS}^{(p)}_{n_1n_2}\Rightarrow\sup_{A\in\mathcal{A}}\|\mathbf{Z}_{P_0}(A)\|$$

and

$$\mathcal{CM}^{(p)}_{n_1n_2}\Rightarrow\int_{\mathbf{D}}\|\mathbf{Z}_{P_0}(A)\|^2dA$$

By Corollary 4, the implementation of the test in the practice can be realized by approximating the finite samples quantiles of $\mathcal{KS}^{(p)}_{n_1n_2}$ and $\mathcal{CM}^{(p)}_{n_1n_2}$ using those of $sup_{A\in\mathcal{A}}\|\mathbf{Z}_{P_0}(A)\|$ and $\int_{\mathbf{D}}\|\mathbf{Z}_{P_0}(A)\|^2dA$, respectively. More precisely, for $\alpha\in(0,1)$, let $ks_{1-\alpha}$ and $cm_{1-\alpha}$ be positive real numbers that satisfy the equations

$$\mathbf{P}\left\{\sup_{A\in\mathcal{A}}\|\mathbf{Z}_{P_0}(A)\|\geq ks_{1-\alpha}\right\}=\alpha$$
$$\mathbf{P}\left\{\int_{\mathbf{D}}\|\mathbf{Z}_{P_0}(A)\|^2dA\geq cm_{1-\alpha}\right\}=\alpha.$$

Then the rejection regions of asymptotic size α tests for testing H_0 are given respectively by

$$\mathcal{C}_{KS}:=\{\mathbf{Y}_{n_1n_2}:\ \mathcal{KS}^{(p)}_{n_1n_2}\geq ks_{1-\alpha}\},$$

when $\mathcal{KS}^{(p)}_{n_1n_2}$ is used and

$$\mathcal{C}_{CvM}:=\{\mathbf{Y}_{n_1n_2}:\ \mathcal{CM}^{(p)}_{n_1n_2}\geq cm_{1-\alpha}\},$$

when the statistic $\mathcal{CM}^{(p)}_{n_1n_2}$ is used. Thus, the problem of testing H_0 reduces to that of computing $ks_{1-\alpha}$ and $cm_{1-\alpha}$ for any pre-signed $\alpha\in(0,1)$. Moreover, without altering the test procedures, the unknown covariance function $\mathbf{\Sigma}$ can be estimated by any consistent estimator, such as by that given in [2].

To be able to investigate the limiting power of the tests, we consider the general localized version of Model 1, defined by:

$$\widetilde{\mathbf{Y}}(\mathbf{t})=\frac{1}{\sqrt{n_1n_2-q}}\widetilde{\mathbf{g}}(\mathbf{t})+\widetilde{\mathcal{E}}(\mathbf{t}),\ \mathbf{t}\in\mathbf{D},\qquad(5)$$

with $E(\widetilde{\mathcal{E}}(\mathbf{t}))=\mathbf{0}$ and $Cov(\widetilde{\mathcal{E}}(\mathbf{t}))=\mathbf{\Sigma}$. For $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q+1}$, let

$$\mathbf{G}^{(n_1,n_2)}_{j_1j_2}\ :=\ \left(\mathbf{g}^{(1)}_{j_1j_2},\mathbf{g}^{(2)}_{j_1j_2},\ldots,\mathbf{g}^{(p)}_{j_1j_2}\right).$$

Then, we have

$$\mathbf{Y}^{(n_1,n_2)}_{j_1j_2}=\frac{1}{\sqrt{n_1n_2-q}}\mathbf{G}^{(n_1,n_2)}_{j_1j_2}+\mathbf{E}^{(n_1,n_2)}_{j_1j_2}.\qquad(6)$$

Conversely, when H_0 is true, the model clearly reduces to (4). Hence, by applying Theorem 3, the asymptotic test procedure is not altered when the localized model is considered, in the sense the test leads to the same size α rejection region as that of the non localized model.

The limiting distribution of the statistics $\mathcal{KS}^{(p)}_{n_1n_2}$ and $\mathcal{CM}^{(p)}_{n_1n_2}$ for the localized model (5), when H_0 is not true is presented in the following theorem.

Theorem 5 *Suppose that the vector of regression functions $\widetilde{\mathbf{g}}=(g_1,\ldots,g_p)^\top$ is continuous and has bounded variation on $\mathbf{D}$. Let $\widetilde{\mathbf{W}}^{loc}_{n_1\times n_2}:=(\widetilde{\mathbf{w}}^{loc}_{j_1j_2})$, for $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$ be the sequence of arrays of p-dimensional vector of recursive residuals associated with the localized model (5) observed over the regular lattice $\Xi_{n_1\times n_2}$. Then, when H_0 is not true it holds*

$$\mathbf{\Sigma}^{-1/2}\frac{S^{(p)}_{n_1n_2-q}(\widetilde{\mathbf{W}}^{loc}_{n_1\times n_2})}{\sqrt{n_1n_2-q}}\Rightarrow\mathbf{\Sigma}^{-1/2}h_{\widetilde{\mathbf{g}}}+\mathbf{Z}_{P_0},$$

where

$$\begin{aligned}h_{\widetilde{\mathbf{g}}}(A)\ &:=\ \int_A\widetilde{\mathbf{g}}(x,y)P_0(dx,dy)\\ &-\ \int_A\mathbf{f}^\top(u,v)\mathbf{G}^{-1}(u,v)\mathbf{H}(u,v)P_0(du,dv),\end{aligned}$$

thereby

$$\mathbf{G}(u,v) := \int_{B_{(u,v)}} \mathbf{f}(x,y)\mathbf{f}^\top(x,y)P_0(dx,dy) \in \mathcal{R}^{q\times q}$$

$$\mathbf{H}(u,v) := \int_{B_{(u,v)}} \mathbf{f}(x,y)\widetilde{\mathbf{g}}^\top(x,y)P_0(dx,dy) \in \mathcal{R}^{q\times p}.$$

The subset $B_{(u,v)}$ is determined by the set A and the variable $(u,v) \in A$.

As a direct application of the well-known continuous mapping theorem, the asymptotic power function of the test based on the p-dimensional statistics $KS^{(p)}_{n_1n_2}$ and $CM^{(p)}_{n_1n_2}$ can be expressed as follows.

Corollary 6 *Suppose that for testing the hypothesis H_0 defined in Section 1 the localized model (5) is observed under the equidistance design $\Xi_{n_1\times n_2}$. Asymptotic power function of the size α Kolmogorov-Smirnov test is given by*

$$\lim_{n_1,n_2\to\infty} \Upsilon_{\mathcal{KS}^{(p)}_{n_1n_2}}(\widetilde{\mathbf{g}})$$
$$= \lim_{n_1,n_2\to\infty} \mathbf{P}\left\{\mathcal{KS}^{(p)}_{n_1n_2} \geq ks_{1-\alpha} \mid \widetilde{\mathbf{g}}\right\}$$
$$= \mathbf{P}\left\{\sup_{A\in\mathcal{A}} \left\|\mathbf{\Sigma}^{-1/2}h_{\widetilde{\mathbf{g}}}(A) + \mathbf{Z}_{P_0}(A)\right\| \geq \tilde{k}_{1-\alpha}\right\}.$$

Similarly, the asymptotic power function of the Cramér-von Mises test of size α is given by

$$\lim_{n_1,n_2\to\infty} \Upsilon_{\mathcal{CM}^{(p)}_{n_1n_2}}(\widetilde{\mathbf{g}})$$
$$= \lim_{n_1,n_2\to\infty} \mathbf{P}\left\{\mathcal{CM}^{(p)}_{n_1n_2} \geq cm_{1-\alpha} \mid \widetilde{\mathbf{g}}\right\}$$
$$= \mathbf{P}\left\{\int_{\mathbf{D}} \left\|\mathbf{\Sigma}^{-1/2}h_{\widetilde{\mathbf{g}}}(A) + \mathbf{Z}_{P_0}(A)\right\|^2 dA \geq cm_{1-\alpha}\right\}.$$

In this paper the finite sample size behavior of the tests will be investigated by simulation by comparing the power functions of the tests based on the $\mathcal{KS}^{(p)}_{n_1n_2}$ and $\mathcal{CM}^{(p)}_{n_1n_2}$ statistics. It can be shown easily that when H_0 is true, the term $\mathbf{\Sigma}^{-1/2}h_{\widetilde{\mathbf{g}}}$ vanishes uniformly, so that the power attains the pre signed size of the test. That is $\Upsilon_{\mathcal{KS}^{(p)}_{n_1n_2}}(\widetilde{\mathbf{g}}) = \alpha = \Upsilon_{\mathcal{CM}^{(p)}_{n_1n_2}}(\widetilde{\mathbf{g}})$, for $\widetilde{\mathbf{g}}$ varies under H_0.

4 Simulation study

We now study the power of the KS and CM type tests via Monte Carlo simulations by considering three cases. In each case the graph of the empirical power function $\Upsilon_{\mathcal{KS}^{(p)}_{n_1n_2}}$ and $\Upsilon_{\mathcal{CM}^{(p)}_{n_1n_2}}$ are developed and compared each other. The samples are generated over $\mathcal{U}$ with the experimental design given by 50×50 regular lattice

$$\Xi_{50\times 50} = \{(\ell/50, k/50) : \ 1 \leq \ell \leq 50, \ 1 \leq k \leq 50\}.$$

The vectors of random errors are generated independently from the centered multivariate normal distribution $N_p(\mathbf{0}, \mathbf{\Sigma})$ for some nonsingular covariance matrix $\mathbf{\Sigma}$. However, in the simulation it is assumed that $\mathbf{\Sigma}$ is unknown, therefore it is estimated by a consistent estimator $\widehat{\mathbf{\Sigma}}^{H_1}_{nn}/(n^2)$ defined in [2]. The number of runs is 1500.

4.1 Bivariate constant model

In the first scenario we test the hypothesis that a bivariate constant model holds true. The samples are generated based on the following localized bivariate first-order model

$$\widetilde{\mathbf{Y}} = \left(\begin{pmatrix} 5 \\ 3 \end{pmatrix} + \rho \begin{pmatrix} \frac{\ell}{n} + \frac{k}{n} \\ \frac{\ell}{n} + \frac{k}{n} \end{pmatrix}\right) / \sqrt{n^2-1} + \widetilde{\mathcal{E}}.$$

The real constant ρ varies in some interval so that the mean function $\widetilde{\mathbf{g}}$ varies in the space of vector of functions of bounded variations on $\mathcal{U}$. It is clear that the observations are from H_0, when $\rho = 0$. The two dimensional random error $\widetilde{\mathcal{E}}$ is generated independently from the bivariate normal distribution $N_2(\mathbf{0}, \mathbf{\Sigma})$, with the covariance matrix

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

Figure 1 exhibits the empirical power functions of the tests, where the graphs of $\Upsilon_{\mathcal{KS}^{(p)}_{n_1n_2}}$ and $\Upsilon_{\mathcal{CM}^{(p)}_{n_1n_2}}$ are indicated by a solid and dashed line, respectively. The left panel is for $\alpha = 0.01$, whereas the right one corresponds to $\alpha = 0.05$. The graphs show that the power increases as the model moves away from H_0. The power fluctuates around α as $\rho = 0$ as it must be. It can be seen that the test based on the Cramér-von Mises type statistic is slightly more powerful than that based on the Kolmogorov-Smirnov one.

4.2 Trivariate first-order model

In this subsection a more general hypothesis and model are simulated in which the samples are generated based on the following model

$$\widetilde{\mathbf{g}}\left(\frac{\ell}{n}, \frac{k}{n}\right) = \begin{pmatrix} 1 + \frac{2\ell}{n} + \frac{k}{n} \\ -2 + \frac{3\ell}{n} + \frac{3k}{n} \\ 3 + \frac{\ell}{n} - \frac{2k}{n} \end{pmatrix} + \rho \begin{pmatrix} \exp\{\frac{\ell}{n}\frac{k}{n}\} \\ \exp\{\frac{\ell}{n}\frac{k}{n}\} \\ \exp\{\frac{\ell}{n}\frac{k}{n}\} \end{pmatrix}$$

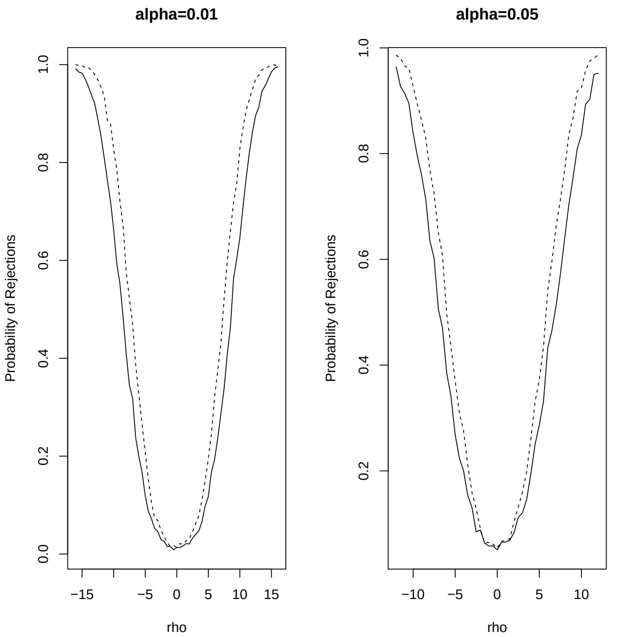

Figure 1: The graphs of $\Upsilon_{\mathcal{KS}^{(p)}_{n_1 n_2}}$ (solid line) and $\Upsilon_{\mathcal{CM}^{(p)}_{n_1 n_2}}$ (dashed line) for testing bivariate constant model.

for testing the hypothesis that a trivariate first-order model is true. Under alternative we consider a trivariate nonparametric model by adding the model specified under H_0 with an exponential terms. The error component is generated independently from the trivariate centered normal distribution $N_3(\mathbf{0}, \mathbf{\Sigma})$, where

$$\mathbf{\Sigma} = \begin{pmatrix} 0.65 & 0.31 & 1.18 \\ 0.31 & 0.58 & 0.81 \\ 1.18 & 0.81 & 2.50 \end{pmatrix}$$

The scatter plot of the empirical power functions of size $\alpha = 0.01$ and $\alpha = 0.05$ are presented in Figure 2. It can be concluded that the CM type test (dashed line) has larger power than the KS type test (solid line). The sizes of the tests are achieved when ρ is set to zero. That is as the observations are from H_0 both tests attain the pre-signed values of α.

4.3 Trivariate second-order model

The last simulation concerns with the problem of testing the hypothesis that a second-order model holds true. The vector of observations are generated based on the localized model

$$\widetilde{\mathbf{Y}} = \frac{\widetilde{\mathbf{g}}}{\sqrt{n^2 - 6}} + \widetilde{\mathcal{E}},$$

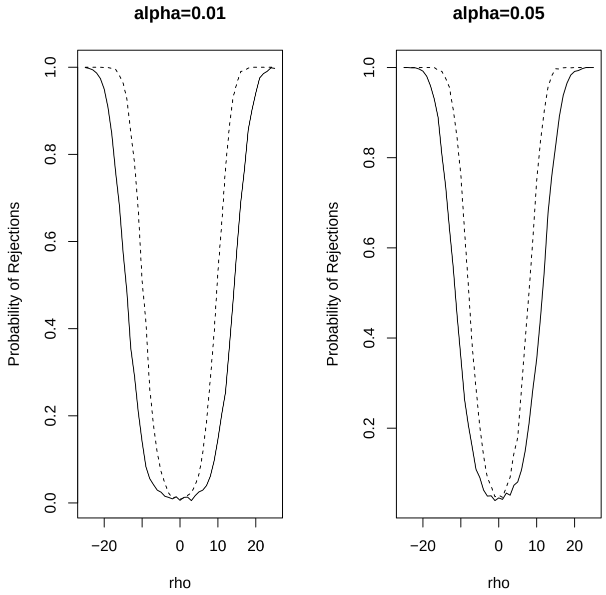

Figure 2: The graphs of $\Upsilon_{\mathcal{KS}^{(p)}_{n_1 n_2}}$ (solid line), and $\Upsilon_{\mathcal{CM}^{(p)}_{n_1 n_2}}$ (dashed line) for testing three-variate first-order model.

where

$$\widetilde{\mathbf{g}}\left(\frac{\ell}{n}, \frac{k}{n}\right) = \begin{pmatrix} 2 + \frac{\ell}{n} + \frac{k}{n} + \frac{2k^2}{n^2} + \frac{3\ell^2}{n^2} + \frac{\ell k}{n^2} \\ -1 + \frac{2\ell}{n} + \frac{3k}{n} + \frac{k^2}{n^2} - \frac{2\ell^2}{n^2} + \frac{5\ell k}{n^2}\} \\ 3 + \frac{\ell}{n} - \frac{2k}{n} + \frac{6k^2}{n^2} + \frac{8\ell^2}{n^2} + \frac{5\ell k}{n^2} \end{pmatrix} + \rho \begin{pmatrix} \sin(\frac{\ell k}{n^2}) \\ \exp\{\frac{\ell k}{n^2}\} \\ \exp\{\frac{\ell k}{n^2}\} \sin(\frac{k\ell}{n^2}) \end{pmatrix}.$$

The vector of random error $\widetilde{\mathcal{E}}$ is generated independently from the centered three-variate normal distribution $N_3(\mathbf{0}, \mathbf{\Sigma})$ with the covariance matrix as in Subsection 4.2. As in the preceding scenarios, the simulation result for testing second-order model shows that the CM type test has larger power than the KS type test has, see Figure 3. When H_0 is true, the powers of the tests attain their pre determined sizes. However, when the model is far away from H_0, the power increases gradually.

5 Application

In this section we demonstrate the application of the asymptotic test procedures to a multivariate data which is the corn plant data. The data consist of the measurements of the maximum weight of the corn yield (in gram), the maximum height of the corn plants (in

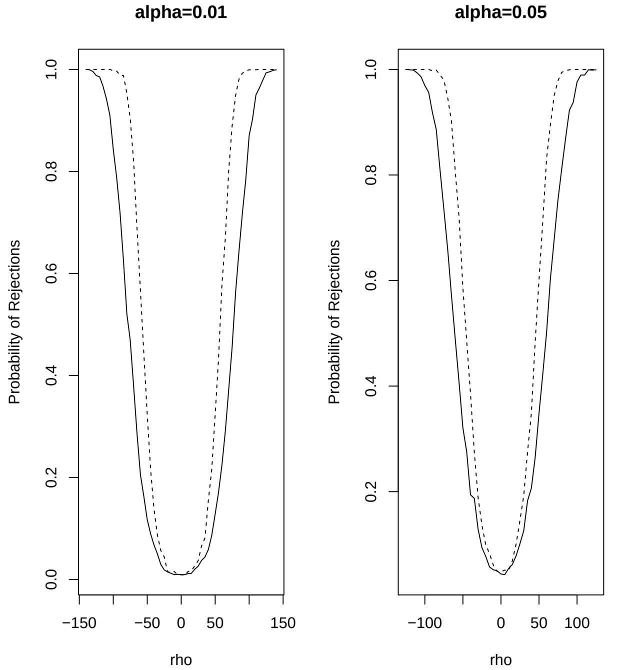

Figure 3: The graphs of $\Upsilon_{\mathcal{KS}^{(p)}_{n_1n_2}}$ (solid line), and $\Upsilon_{\mathcal{CM}^{(p)}_{n_1n_2}}$ (dashed line) for testing quadratic model.

cm) and the rate of growth (in cm/day) of 21 x 16 corn plants planted over a rectangular farm region of size [0,12 m]x[0,15.75 m] running from west to east and from north to south. The experimental design is given by a 0.75 m x 0.75 m dimensional regular lattice, see [21]. The measurements have been conducted from August 2018 to October 2018. The goal is to build a model empirically describing how the values of the three variables vary over the region as a function of the coordinate of any point on the experimental region.

Descriptive investigation using matrix scatter plot and Pearson correlation coefficient regarding the existence of the correlations among the logarithm of the maximum weight (Ln Weight), the logarithm of the maximum height (Ln Height) and the Rate of Growth shows that the three variables are positively correlated each other (Figure 4). In particular, the correlation between the Ln Weight and the Ln Height is relatively strong compared to those between the Ln Weight and the Rate of Growth and between the Ln Weight and the Rate of Growth, see also Table 1. By this preliminary diagnostic results, the empirical model building must be conducted using multivariate analysis without ignoring the inherent correlation among the variables. Furthermore, based on the normal Kolmogorov-Smirnov test presented in Table 2 there is no enough statistical evidence to say that the three variables follow a trivariate normal distribution.

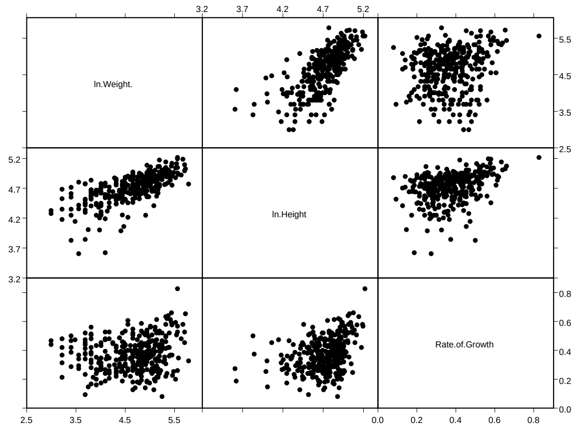

Figure 4: The pairs plot of the Ln Weight, Ln Height and the Rate of Growth of corn plants data.

Table 1: The Pearson correlation matrix for the Ln Weight, the Ln Height and the Rate of Growth.

	Ln Weight	Ln Height	Growth
Ln Weight	1.00000	0.69965	0.18816
Ln Height	0.69965	1.00000	0.34798
Growth	0.18816	0.34799	1.00000

Table 2: The Kolmogorov-Smirnov goodness of fit test using the command ”ks.gof” in R for the normality of the Ln Weight, Ln Height and Rate of Growth.

Variables	Critical Values	p-Values
Ln Weight	0.10330	0.00000
Ln Height	0.08000	0.00000
Rate of Growth	0.04390	0.50000

Three dimensional scatter plot of each variable presented respectively in Figure 5, Figure 6 and Figure 7 indicate that three dimensional polynomial of low order are reasonable for describing the regression relationship between the observed variables and the coordinate of every position on the experimental region.

The main objective is to test the validity of the assumed model based on the partial sums of the recursive residuals. The test results are presented in Table 3. When under H_0 a three dimensional constant model is assumed, both the KS and CCM type tests reject the hypothesis by the fact the corresponding p-values are very small. This conclusion is also supported by the Wilk’s lambda test when under the alternative a three dimensional first-order model is assumed by the reason the test also has a very small p-value, see the figures in the second row of Table 3. Thus it can be concluded that three dimensional constant model is not plausible for the corn plant data.

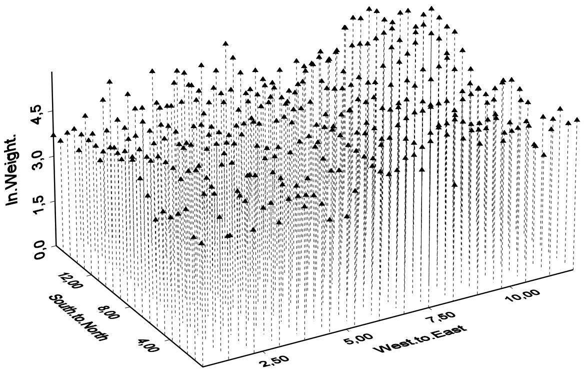

Figure 5: The three dimensional scatter plot of the Ln Weight observed over a 16×21 regular lattice.

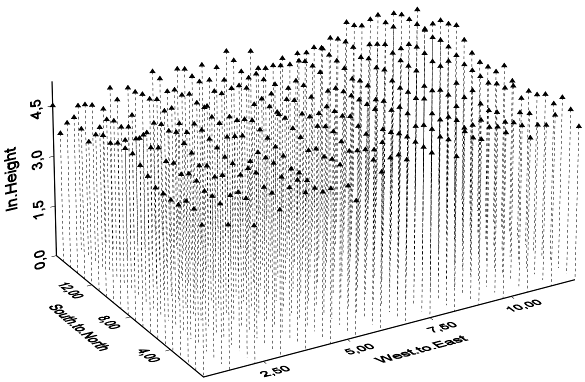

Figure 6: The three dimensional scatter plot of the Ln Height observed over a 16×21 regular lattice.

Now we test the hypothesis that three dimensional first-order model is significant. The KS as well as the CM type tests lead to the acceptance of the hypothesis. Referring to the associated p-values of the tests (see the third rows of Table 3), the hypothesis will be rejected for $\alpha \geq 21.101\%$ when the KS type test is used. Similarly, by using the CM type test the hypothesis will be rejected for $\alpha \geq 32.093\%$. Since for these numbers the probabilities of the rejection of the hypothesis when it is true is large, we decide to accept the hypothesis. The same conclusion is also obtained when under the alternative a three dimensional second-order model is considered. By employing the Wilks lambda test, the hypothesis is also not rejected since the p-value is 32.093%. We therefor conclude that first-order model is a significant model.

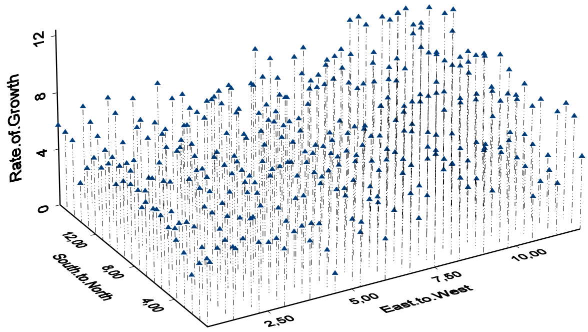

Figure 7: The three dimensional scatter plot of the Rate of Growth observed over a 16×21 regular lattice.

Table 3: The critical values and the approximated p-values of the $\mathcal{KS}^{(p)}_{21\times 16}$, $\mathcal{CM}^{(p)}_{21\times 16}$ and the Wilk's lambda tests for the corn plants data.

Model	$\mathcal{KS}^{(p)}_{21;16}$	$\mathcal{CM}^{(p)}_{21;16}$	$\Lambda_{21;16}$
Constant	24.49602	64.51159	25.21487
P-Value	0.00012	0.00035	0.00031
First Order	17.19393	16.85463	10.37576
P-Value	0.21101	0.11236	0.32093

The least squares estimate of the parameter matrix $\mathbf{B}$ is given by

$$\widehat{\mathbf{B}} = \begin{pmatrix} 0.24424 & 4.42890 & 4.49722 \\ 0.15932 & 0.40387 & 0.29157 \\ 0.07932 & 0.00587 & 0.11326 \end{pmatrix}.$$

Hence, the first-order fitted model associated with the corn plants data is as follows

$$\begin{pmatrix} \widehat{Y}_1 \\ \widehat{Y}_2 \\ \widehat{Y}_3 \end{pmatrix} = \begin{pmatrix} 0.2442 + 0.1593x + 0.0793y \\ 4.49722 + 0.40387x + 0.00587y \\ 4.49722 + 0.29157x + 0.11326y \end{pmatrix},$$

for $(x, y) \in \mathbf{D}$. By this model the values of the Ln Weight, Ln Height and Rate of Growth simultaneously increase as the coordinate of the point moves away from the origin which is put on the south-west corner of the region. In fact, when the correlation among the variables are ignored, the one dimensional partial sums method proposed in [21] can be applied to each individual variable. A routine computation gives the result that first-order model is also fit well to the corn plant data, see [21].

From the agricultural perspective, the three observed variables (Ln Weight, Ln Height and Rate of Growth) can be regarded as indicators of the fertility level of a farm land. The larger the values of these variables, the better the fertility level of the land is and viceversa, the smaller the values of these variables, the worst the fertility level of the land is. Hence, by observing the fitted model presented above, it can be concluded that the fertility level of the region gets large as the position moves away from the origin.

6 Conclusion

In this paper asymptotic procedure for testing model validity in multivariate linear regression based on the partial sums process of the recursive residuals has been established. The method is derived for the case when the probability distribution model of the vector of the observations is unknown. The limit process under H_0 is given by the multivariate set-indexed Brownian sheet independent to whatever the assumed regression models is. This result gives an advantage in that, the computation of the quantiles of the test statistic theoretically becomes easier. The application of the test method to the corn plants data give the similar result as that of the Wilk's lambda test. Three dimensional first-order polynomial model is fitted well to the corn plants data.

Appendix

Proposition A1. For the vector of recursive residuals $\widetilde{\mathbf{w}}_{j_1 j_2} = \left(w^{(1)}_{j_1 j_2}, \ldots, w^{(p)}_{j_1 j_2}\right)^\top$ defined in Definition 1, it holds $E(\widetilde{\mathbf{w}}_{j_1 j_2}) = \mathbf{0}$ and

$$Cov\left(\widetilde{\mathbf{w}}_{j_1 j_2}, \widetilde{\mathbf{w}}_{j_1' j_2'}\right) = \begin{cases} \mathbf{O} & ; j_1 \neq j_1' \; or \; j_2 \neq j_2' \\ \mathbf{\Sigma} & ; j_1 = j_1' \; and \; j_2 = j_2' \end{cases},$$

for every $(j_1, j_2),\ (j_1', j_2') \in \mathbf{T}_{n_1 n_2 - q}$.

Proof: Let $(j_1, j_2),\ (j_1', j_2') \in \mathbf{T}_{n_1 n_2 - q}$ be arbitrary. By Proposition 2, we have

$$E\left(\widetilde{\mathbf{w}}_{j_1 j_2}\right) = E\left(\mathbf{a}^\top_{j_1 j_2}\mathcal{E}^{(1)}_{n_1 n_2}, \ldots, \mathbf{a}^\top_{j_1 j_2}\mathcal{E}^{(p)}_{n_1 n_2}\right)^\top = \mathbf{0},$$

and

$$\begin{aligned} Cov\left(\widetilde{\mathbf{w}}_{j_1 j_2}, \widetilde{\mathbf{w}}_{j_1' j_2'}\right) &= E\left(\mathbf{E}^\top_{n_1 n_2}\mathbf{a}_{j_1 j_2}\right)\left(\mathbf{a}^\top_{j_1' j_2'}\mathbf{E}_{n_1 n_2}\right) \\ &= E\begin{pmatrix} \mathbf{a}^\top_{j_1 j_2}\mathcal{E}^{(1)}_{n_1 n_2} \\ \vdots \\ \mathbf{a}^\top_{j_1 j_2}\mathcal{E}^{(p)}_{n_1 n_2} \end{pmatrix}\left(\mathbf{a}^\top_{j_1' j_2'}\mathcal{E}^{(1)}_{n_1 n_2}, \ldots, \mathbf{a}^\top_{j_1' j_2'}\mathcal{E}^{(p)}_{n_1 n_2}\right) \\ &= \begin{pmatrix} \sigma_{11}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} & ,\cdots, & \sigma_{1p}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} \\ \sigma_{21}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} & ,\cdots, & \sigma_{2p}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} \\ \vdots & \vdots & \vdots \\ \sigma_{p1}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} & ,\cdots, & \sigma_{p2}\mathbf{a}^\top_{j_1 j_2}\mathbf{I}\mathbf{a}_{j_1' j_2'} \end{pmatrix}. \end{aligned}$$

Since $\mathbf{a}^\top_{j_1 j_2}\mathbf{a}_{j_1' j_2'} = 1$ for $j_1 = j_1'$ and $j_2 = j_2'$ and $\mathbf{a}^\top_{j_1 j_2}\mathbf{a}_{j_1' j_2'} = 0$, for $j_1 \neq j_1'$ or $j_2 \neq j_2'$, the result follows, establishing the proof.

Proof of Theorem 3: By the multivariate version of Donsker's theorem (cf. [3]) we need to show that the finite dimensional distribution of the sequence

$$\frac{1}{\sqrt{n_1 n_2 - q}}\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1 n_2 - q}(\widetilde{\mathbf{W}}_{n_1 n_2})$$

converges to that of $\mathbf{Z}_{P_0}$ and the sequence is tight. It is noticed that the sum is defined component-wise. Let $\gamma_1, \ldots, \gamma_m$ and $A_1, \ldots, A_m$ be arbitrary m constants and convex subsets of $\mathbf{D}$, respectively. Let

$$\mathbf{U}_{n_1 n_2 - q} := \sum_{\ell=1}^{m}\gamma_\ell \frac{\mathbf{\Sigma}^{-1/2} S^{(p)}_{n_1 n_2 - q}(\widetilde{\mathbf{W}}_{n_1 n_2})(A_\ell)}{\sqrt{n_1 n_2 - q}}.$$

Then, by Proposition A1, we get

$$\begin{aligned} Cov\left(\mathbf{U}_{n_1 n_2 - q}\right) &= E\left(\mathbf{U}_{n_1 n_2 - q}\mathbf{U}^\top_{n_1 n_2 - q}\right) \\ &= \textstyle\sum_{\ell=1}^{m}\sum_{k=1}^{m}\frac{\gamma_\ell\gamma_k}{n_1 n_2 - q} \times \\ &\textstyle\sum_{(j_1, j_2),(j_1', j_2') \in \mathbf{T}_{n_1 n_2 - q}} \mathbf{1}_{A_\ell}(\mathbf{t}_{j_1 j_2})\mathbf{1}_{A_k}(\mathbf{t}_{j_1' j_2'}) \\ &\times \mathbf{\Sigma}^{-1/2} E\left(\widetilde{\mathbf{w}}_{j_1 j_2}\widetilde{\mathbf{w}}^\top_{j_1' j_2'}\right)\mathbf{\Sigma}^{-1/2} \\ &= \textstyle\sum_{\ell=1}^{m}\sum_{k=1}^{m}\frac{\gamma_\ell\gamma_k}{n_1 n_2 - q} \\ &\times \textstyle\sum_{(j_1, j_2) \in \mathbf{T}_{n_1 n_2 - q}} \mathbf{1}_{A_\ell \cap A_k}(\mathbf{t}_{j_1 j_2})\mathbf{I}_p \\ &= \textstyle\sum_{\ell=1}^{m}\sum_{k=1}^{m}\gamma_\ell\gamma_k\frac{n_1 n_2}{n_1 n_2 - q} \\ &\times \left(\int_{A_\ell \cap A_k} P_{n_1 \times n_2}(dx, dy) + o(1)\right)\mathbf{I}_p, \end{aligned}$$

where $\mathbf{I}_p$ is the $p \times p$ dimensional identity matrix and $P_{n_1 \times n_2}$ is the discrete probability measure on $\mathcal{B}(\mathbf{D})$, such that for every $A \in \mathcal{B}(\mathbf{D})$,

$$P_{n_1 \times n_2}(A) := \frac{1}{n_1 n_2}\sum_{(j_1, j_2) \in \mathbf{T}_{n_1 n_2}} \mathbf{1}_A(\mathbf{t}_{j_1 j_2}).$$

Since for n_1 and n_2 simultaneously large, $P_{n_1\times n_2}$ converges weakly to the Lebesque measure P_0 and the ratio $\frac{n_1n_2}{n_1n_2-q}$ converges to 1, then we have

$$\lim_{n_1,n_2\to\infty} Cov\left(\mathbf{U}_{n_1n_2-q}\right)$$
$$=\sum_{\ell=1}^{m}\sum_{k=1}^{m}\gamma_\ell\gamma_k\int_{A_\ell\cap A_k}P_0(dx,dy)\mathbf{I}_p$$
$$=\sum_{\ell=1}^{m}\sum_{k=1}^{m}\gamma_\ell\gamma_k P_0(A_\ell\cap A_k)\mathbf{I}_p.$$

The right-hand side is the covariance of the linear combination $\sum_{\ell=1}^{m}\gamma_\ell\mathbf{Z}_{P_0}(A_\ell)$. Next, we show that

$$\lim_{n_1,n_2\to\infty}\sum_{(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}}\|\mathbf{b}_{j_1j_2}\|^2=0.$$

Since, for $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$, the vector $\mathbf{b}_{j_1j_2}$ has the form

$$\Big(-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top},$$
$$1,0,\ldots,0,0,0,\ldots,0\Big)^\top\in\mathcal{R}^{n_1n_2},$$

it holds

$$\|\mathbf{b}_{j_1j_2}\|^2$$
$$=\left\|\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top},1\right\|^2$$
$$\le\left\|\frac{\mathbf{f}^\top(\mathbf{t}_{j_1j_2})}{\sqrt{n_1n_2}}\left(\frac{\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}}{n_1n_2}\right)^{-1}\frac{\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}}{\sqrt{n_1n_2}}\right\|^2$$
$$+o(1).$$

The result follows by the continuity of the norm operator and the regression functions. So, following and extending the technique in [15], the first assertion follows.

Proof of Theorem 5: For the localized model, suppose that H_0 does not hold true. Then for $(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}$, we get the corresponding p-dimensional vector of recursive residuals as

$$\widetilde{\mathbf{w}}_{j_1j_2}^{loc}=\frac{\widetilde{\mathbf{Y}}_{j_1j_2}-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\widehat{\mathbf{B}}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}}}$$

interpreting $\widetilde{\mathbf{w}}_{j_1j_2}^{1}$ and $\widetilde{\mathbf{Y}}_{j_1j_2}$ as the row vectors. By substituting the vector of observation $\widetilde{\mathbf{Y}}_{j_1j_2}$ and $\widehat{\mathbf{B}}_{j_1-1j_2}^{(n_1,n_2)}$, we further get

$$\widetilde{\mathbf{w}}_{j_1j_2}^{loc}=\frac{\frac{\widetilde{\mathbf{g}}(\mathbf{t}_{j_1j_2})}{\sqrt{n_1n_2-q}}+\widetilde{\mathcal{E}}_{j_1j_2}}{\sqrt{d_{j_1j_2}}}-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\times$$
$$\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{Y}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}}}$$
$$=\frac{\widetilde{\mathbf{g}}(\mathbf{t}_{j_1j_2})}{\sqrt{d_{j_1j_2}(n_1n_2-q)}}+\frac{\widetilde{\mathcal{E}}_{j_1j_2}}{\sqrt{d_{j_1j_2}}}-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\times$$
$$\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{G}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}(n_1n_2-q)}}$$
$$-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\times$$
$$\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{E}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}}}$$

Since Proposition 2 ensures that

$$\frac{\widetilde{\mathcal{E}}_{j_1j_2}}{\sqrt{d_{j_1j_2}}}+\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{E}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}}}$$
$$=\widetilde{\mathbf{w}}_{j_1j_2},$$

which is the vector of recursive residual under H_0, then the last expression can be simplified as

$$\widetilde{\mathbf{w}}_{j_1j_2}^{loc}=\frac{\widetilde{\mathbf{g}}(\mathbf{t}_{j_1j_2})}{\sqrt{d_{j_1j_2}(n_1n_2-q)}}-\mathbf{f}^\top(\mathbf{t}_{j_1j_2})\times$$
$$\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{G}_{j_1-1j_2}^{(n_1,n_2)}}{\sqrt{d_{j_1j_2}(n_1n_2-q)}}+\widetilde{\mathbf{w}}_{j_1j_2}.$$

Let $\widetilde{\mathbf{W}}_{n_1\times n_2}^{loc}:=\left(\widetilde{\mathbf{w}}^{loc}\right)_{j_1=1,j_2=1}^{n_1,n_2}$ be the $n_1\times n_2$ array of the p-dimensional vector of the recursive residuals associated with the localized model. By considering the linearity of the partial sums operator we get for every $A\in\mathcal{A}$,

$$\frac{1}{\sqrt{n_1n_2-q}}S_{n_1n_2-q}^{(p)}(\widetilde{\mathbf{W}}_{n_1\times n_2}^{loc})(A)$$
$$=\sum_{(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}}\mathbf{1}_A(\mathbf{t}_{j_1j_2})\frac{\widetilde{\mathbf{g}}(\mathbf{t}_{j_1j_2})}{(n_1n_2-q)\sqrt{d_{j_1j_2}}}$$
$$-\sum_{(j_1,j_2)\in\mathbf{T}_{n_1n_2-q}}\mathbf{1}_A(\mathbf{t}_{j_1j_2})\mathbf{f}^\top(\mathbf{t}_{j_1j_2})$$
$$\times\frac{\left(\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)}\right)^{-1}\mathbf{X}_{j_1-1j_2}^{(n_1,n_2)\top}\mathbf{G}_{j_1-1j_2}^{(n_1,n_2)}}{(n_1n_2-q)\sqrt{d_{j_1j_2}}}$$
$$+\frac{1}{\sqrt{n_1n_2-q}}S_{n_1n_2-q}^{(p)}(\widetilde{\mathbf{W}}_{n_1\times n_2})(A).$$

The first term on the right-hand side of the last equation can be re-written as

$$\frac{n_1n_2}{(n_1n_2-q)\sqrt{d_{j_1j_2}}}\sum_{(j_1,j_2)}\frac{\mathbf{1}_A(\mathbf{t}_{j_1j_2})\widetilde{\mathbf{g}}(\mathbf{t}_{j_1j_2})}{n_1n_2}$$

$$= \frac{n_1 n_2}{(n_1 n_2 - q)\sqrt{d_{j_1 j_2}}} \int_A \widetilde{\mathbf{g}}(x,y) P_{n_1 n_2}(dx, dy)$$

where the sum is over $\mathbf{T}_{n_1 n_2 - q}$. Since the components of $\widetilde{\mathbf{g}}$ has bounded variation on $\mathbf{D}$ and $P_{n_1 n_2} \Rightarrow P_0$, then by the definition of integral component-wise and the fact that $q \ll n_1 n_2$, we get by applying the similar argument as in the univariate case (cf. [21]), that

$$\lim_{n_1, n_2 \to \infty} \int_A \frac{n_1 n_2 \widetilde{\mathbf{g}}(x,y)}{(n_1 n_2 - q)\sqrt{d_{j_1 j_2}}} P_{n_1 n_2}(dx, dy)$$
$$= \int_A \widetilde{\mathbf{g}}(x,y) P_0(dx, dy).$$

For the second term we have

$$\sum_{(j_1,j_2) \in \mathbf{T}_{n_1 n_2 - q}} \mathbf{1}_A(\mathbf{t}_{j_1 j_2}) \mathbf{f}^\top(t_{j_1 j_2})$$
$$\times \frac{\left(\mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)\top} \mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)}\right)^{-1} \mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)\top} \mathbf{G}_{j_1 - 1 j_2}^{(n_1,n_2)}}{(n_1 n_2 - q)\sqrt{d_{j_1 j_2}}}$$
$$= \frac{1}{n_1 n_2} \sum_{(j_1,j_2) \in \mathbf{T}_{n_1 n_2 - q}} \frac{(n_1 n_2) \mathbf{1}_A(\mathbf{t}_{j_1 j_2}) \mathbf{f}^\top(\mathbf{t}_{j_1 j_2})}{(n_1 n_2 - q)\sqrt{d_{j_1 j_2}}}$$
$$\times \left(\frac{\mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)\top} \mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)}}{n_1 n_2}\right)^{-1} \frac{\mathbf{X}_{j_1 - 1 j_2}^{(n_1,n_2)\top} \mathbf{G}_{j_1 - 1 j_2}^{(n_1,n_2)}}{n_1 n_2}$$
$$= \frac{n_1 n_2}{(n_1 n_2 - q)} \int_A \frac{\mathbf{f}^\top(u,v)}{\sqrt{d_{j_1 j_2}}} \times$$
$$\left(\int_{B_{u,v}} (f_k(x,y) f_\ell(x,y))_{k=1,\ell=1}^{q,q} P_{n_1 n_2}(dx, dy)\right)^{-1}$$
$$\times \left(\int_{B_{u,v}} f_k(x,y) g_\ell(x,y) P_{n_1 n_2}(dx, dy)\right)_{k=1,\ell=1}^{q,p}$$
$$P_{n_1 n_2}(du, dv).$$

Hence, by applying the similar argument as in the case of the univariate model (cf. [21]), the last term converges to

$$\int_A \mathbf{f}^\top(u,v) \mathbf{G}^{-1}(u,v) \mathbf{H}(u,v) P_0(du, dv).$$

Thus, overall we have

$$\frac{1}{\sqrt{n_1 n_2 - q}} \mathbf{\Sigma}^{-1/2} S_{n_1 n_2 - q}^{(p)}(\widetilde{\mathbf{W}}_{n_1 \times n_2}^{loc})(A)$$

converges in distribution to

$$\mathbf{\Sigma}^{-1/2} h_{\widetilde{\mathbf{g}}}(A) + \mathbf{Z}_{P_0},$$

finishing the proof.

Acknowledgements: The research was supported by the Ministry of Research and Higher Education of the Republic of Indonesia through the Fundamental Research Grant 2019.

References:

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

[2] S.F. Arnold, The asymptotic validity of invariant procedures for the repeated measures model and multivariate linear model, *J. Amer. Statist. Assoc.* 75(372), 1984, pp. 890–894.

[3] P. Billingsley, *Convergence of Probability Measures (2nd ed.)*, John Wiley & Sons, New York 1999.

[4] 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.

[5] R.L. Brown, J. Durbin and J.M. Evans, Teqhniques for testing the regression relationships over time, *J. R. Statist. Soc. B* 37(2), 1975, pp. 149–192.

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

[7] C.S.J. Chu, K. Hornik and C.M. Kuan, MOSUM test for parameter constancy, *Biometrika* 82, 1995, pp. 603–617.

[8] J.M. Dufour, Recursive stability analysis of linear regression relationships, *J. Econometrics* 19, 1982, pp. 31–76.

[9] J. Groen, G. Kapetanios and S. Price, Multivariate methods for monitoring structural change, *Working Paper No. 369*, Bank of England, 2009.

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

[11] J. Kiefer, K-sample analogous of the Kolmogorov-Smirnov and Cramér-V. Mises tests, *The Ann. of Math. Stat.*, 30(2), 1959, pp. 420–447.

[12] W. Kramér, W. Ploberger and R. Alt, Testing for structural change in dynamic models, *Econometrica* 56, 1988, pp. 1355–1369.

[13] M. Lifshits, *Lectures on Gaussian Processes*, Springer Briefs in Mathematics, Springer–Verlag, Berlin 2012.

[14] R. Pyke, A uniform central limit theorem for partial sum processes indexed by sets, *Ann. Probab.* 79, 1983, pp. 219–240.

[15] P.K. Sen, Invarianc principle for recursive residuals, *Ann. Stat.* 10(1), 1982, pp. 307–312.

[16] W. Somayasa and G.N.A. Wibawa, Asymptotic model-check for multivariate spatial regression with correlated responses *Fareast J. Math. Sci.* 98(5), 2015, pp. 613–939.

[17] W. Somayasa, G.N.A. Wibawa, and Y.B. Pasolon, Multidimensional Set-Indexed Partial Sums Method for Checking the Appropriateness of a Multivariate Spatial Regression *Int. J. Math. Models and Methods in Applied Sciences*, 9, 2015, pp. 700-713.

[18] W. Somayasa, G.N.A. Wibawa, L. Hamimu and L.O. Ngkoimani, Asymptotic theory in model diagnostic for general multivariate spatial regression, *Int. J. Mathematics and Mathematical Sciences* 2016, pp. 1–16.

[19] W. Somayasa, Estimating the power of lack-of-fit test for the mean of multivariate spatial regression, *Applied Mathematical Sciences*, **11**(52), 2017, pp. 2587–2607.

[20] W. Somayasa, Accessing the Power of Tests Based on Set-Indexed Partial Sums of Multivariate Regression Residuals, *Journal of Applied Mathematics*, 2018, pp. 1–13

[21] W. Somayasa, Recursive residuals partial sums method for testing model validity in modelling of spatial data, *WSEAS Trans. Math.*, 18, 2019, pp. 62–72.

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