Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

https://www.czasopisma.uni.lodz.pl/foe/
1(362) 2023
DOI: https://doi.org/10.18778/0208-6018.362.01

Małgorzata Graczyk https://orcid.org/0000-0002-5550-7007

Poznań University of Life Sciences, Faculty of Agronomy Horticulture and Bioengineering, Department of Mathematical and Statistical Methods, Poznań, Poland
malgorzata.graczyk@up.poznan.pl

Bronisław Ceranka https://orcid.org/0000-0002-5395-1593

Poznań University of Life Sciences, Faculty of Agronomy Horticulture and Bioengineering, Department of Mathematical and Statistical Methods, Poznań, Poland
bronislaw.ceranka@up.poznan.pl

Notes on the Efficiency of Spring Balance Weighing Designs with Correlated Errors for An Even Number of Objects

Abstract: In this paper, some issues regarding the efficiency of spring balance weighing designs for a selected class are presented. We give some conditions determining the relations between the parameters of such designs and construction examples.

Keywords: D-efficient design, spring balance weighing design

JEL: C02, C18, C90

1. Introduction

We study the linear model $\mathbf{y} = \mathbf{X}\mathbf{w} + \mathbf{e}$, where:

1) $\mathbf{y}$ is an $n \times 1$ random vector of observed measurements;
2) $\mathbf{X} \in \Psi_{n \times p}(0,1)$, the class of $n \times p$ matrices $\mathbf{X} = (x_{ij})$, having elements equal to 0 or 1, $i = 1,2,\ldots,n$, $j = 1,2,..,p$;
3) $\mathbf{w}$ is a $p \times 1$ vector representing unknown measurements of objects;
4) $\mathbf{e}$ is an $n \times 1$ vector of random errors; moreover, we assume that there are no systematic errors, the errors are uncorrelated and they have different variances, i.e. $\mathrm{E}(\mathbf{e}) = \mathbf{0}_n$, $\mathrm{Var}(\mathbf{e}) = \sigma^2\mathbf{G}$, where $\sigma > 0$ is a known parameter, $\mathbf{G} = (1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}$ is the $n \times n$ positive definite matrix of known elements, $-(n-1)^{-1} < \rho < 1$.

The applications of spring balance weighing designs are unlimited and involve economic surveys, see Banerjee (1975), Ceranka and Graczyk (2014), with agricultural experiments, see Ceranka and Katulska (1987a; 1987b; 1989), Graczyk (2013), as well as bioengineering, see Gawande and Patkar (1999). Some problems presented in the literature concerning weighing designs are connected with optimality criteria, see Jacroux and Notz (1983), Koukouvinos (1996), and with new construction methods, see Gail and Kiefer (1982), Ceranka and Graczyk (2010; 2012), Katulska and Smaga (2010).

For the estimation of the vector of unknown measurements of objects $\mathbf{w}$, we use the normal equation $\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\hat{\mathbf{w}} = \mathbf{X}'\mathbf{G}^{-1}\mathbf{y}$. On the understanding that $\mathbf{G}$ is a known positive definite matrix, $\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}$ is nonsingular if and only if $\mathbf{X}$ is of full column rank. In that case, the generalised least squares estimator of $\mathbf{w}$ is given by $\hat{\mathbf{w}} = \left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{G}^{-1}\mathbf{y}$ and $\mathrm{Var}(\hat{\mathbf{w}}) = \sigma^2(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}$.

2. The main result

We choose D-optimality from among all the possible optimisation criteria. The concept of D-optimal design was considered in the following papers: Raghavarao (1971), Banerjee (1975), Shah and Sinha (1989). We present the definition of D-optimal design below.

Definition 2.1. The design $\mathbf{X}$ is regular D-optimal in the class $\Psi_{n \times p}(0,1)$ for variance matrix of errors $\sigma^2\mathbf{G}$ if $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)^{-1} = \min\left\{\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)^{-1} : \mathbf{X} \in \Psi_{n \times p}(0,1)\right\}$.

For any combination of the number of objects p and the number of measurements n, we are not able to determine a regular D-optimal design. In such a case, a highly D-efficient design is considered, see Bulutoglu and Ryan (2009).

Definition 2.2 (Bulutoglu, Ryan, 2009). The design $\mathbf{X}$ is highly D-efficient in the class $\Psi_{n\times p}(0,1)$ if $\mathrm{D}_{\text{eff}}=\left[\frac{\det(\mathbf{X}'\mathbf{X})}{\det(\mathbf{Y}'\mathbf{Y})}\right]^{1/p}\geq 0.95$, where $\mathbf{Y}$ is the matrix of D-optimal spring balance weighing design.

Ceranka and Graczyk (2018) extended the above-presented definition for the case of D-efficient design in the class $\Psi_{n\times p}(0,1)$ for any variance matrix of errors $\sigma^2\mathbf{G}$.

Definition 2.3 (Ceranka, Graczyk, 2018). The design $\mathbf{X}$ is highly D-efficient in the class $\Psi_{n\times p}(0,1)$ for a given variance matrix of errors $\sigma\ \mathbf{G}$, if $\mathrm{D}_{\text{eff}}=\left[\frac{\det(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})}{\det(\mathbf{Y}'\mathbf{G}^{-1}\mathbf{Y})}\right]^{1/p}\geq 0.95$, where $\mathbf{Y}$ is the matrix of D-optimal spring balance weighing design for the same variance matrix of errors $\sigma^2\mathbf{G}$.

The aim of this paper is to introduce new results that pertain to the methods of construction and set the existence conditions determining the D-optimal or highly D-efficient spring balance weighing designs under the assumption that the random errors are correlated and they have the same variances. For a given class of design matrices $\Psi_{n\times p}(0,1)$, the scientific problem concerns the determination of the design matrix of the optimal spring balance weighing design. If in any class mentioned, a design matrix does not exist, highly D-efficient spring balance weighing designs with given efficiency coefficients are determined.

We remind the theorem determining the parameters of the highly D-efficient design given in Graczyk and Ceranka (2022a).

Theorem 2.1 (Graczyk, Ceranka, 2022a). If $p = 2k$ and $\mathbf{X}\in\Psi_{n\times p}(0,1)$ with the variance matrix of errors $\sigma^2\mathbf{G}=\sigma^2\left((1-\rho)\mathbf{I}_n+\rho\mathbf{1}_n\mathbf{1}_n'\right)$, then:

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)\leq\frac{(p-1)(1-\rho)}{1+\rho(n-1)}\left(\frac{np}{4(p-1)(1-\rho)}\right)^p.$$

The equality is attained if and only if:

$$\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}=\frac{n}{4(p-1)(1-\rho)}\left(p\mathbf{I}_p+\frac{p-2-\rho(n+p-2)}{1+\rho(n-1)}\mathbf{1}_p\mathbf{1}_p'\right).\tag{2.1}$$

In the design given in Theorem 2.1, the number of elements equal to 1 in each row of the design matrix $\mathbf{X}\in\Psi_{n\times p}(0,1)$ is the same. The design satisfying condition (2.1) is not regular D-optimal because $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)$ did not attain the lowest bound in the class $\Psi_{n\times p}(0,1)$. That is the reason why we consider the design matrix $\mathbf{X}\in\Psi_{n\times p}(0,1)$ with two different numbers of elements equal to 1 in rows. Let $p = 2k$ and

$$\mathbf{k} = \begin{bmatrix} k\mathbf{1}_{n_1} \\ (k+1)\mathbf{1}_{n-n_1} \end{bmatrix}. \tag{2.2}$$

Theorem 2.2 (Graczyk, Ceranka, 2022a). If $p = 2k$, $\mathbf{k}$ is of the form (2.2) and $\mathbf{X} \in \Psi_{n\times p}(0,1)$ with the variance matrix of errors $\sigma^2\mathbf{G} = \sigma^2\left((1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}\right)$, then $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le \left(p+1-\frac{\rho np(p+2)}{(p+1)(1+\rho(n-1))}\right)\left(\frac{n(p+2)}{4(p+1)(1-\rho)}\right)^p$.

The equality is attained if and only if:

$$\mathbf{X}'\mathbf{G}^{-1}\mathbf{X} = \frac{n(p+2)}{4(p+1)(1-\rho)}\left(p\mathbf{I}_p + \left(1-\frac{\rho n(p+2)}{(p+1)(1+\rho(n-1))}\right)\mathbf{1}_p\mathbf{1}_p^{'}\right). \tag{2.3}$$

The design satisfying condition (2.3) is regular D-optimal in the class $\Psi_{n\times p}(0,1)$ with the variance matrix of errors $\sigma^2\mathbf{G} = \sigma^2\left((1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}\right)$.

In order to determine a highly D-efficient design according to Definition 2.3, let us consider $\mathrm{D}_{\mathrm{eff}} = \left[\frac{\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)}{\det\left(\mathbf{Y}'\mathbf{G}^{-1}\mathbf{Y}\right)}\right]^{1/p}$, where $\mathbf{X}$ is the design matrix satisfying condition (2.1) and $\mathbf{Y}$ is the design matrix satisfying condition (2.3). Therefore, we have:

$$\mathrm{D}_{\mathrm{eff}} = \left[\frac{\frac{(p-1)(1-\rho)}{1+\rho(n-1)}\left(\frac{np}{4(p-1)(1-\rho)}\right)^p}{\left(p+1-\frac{\rho np(p+2)}{(p+1)(1+\rho(n-1))}\right)\left(\frac{n(p+2)}{4(p+1)(1-\rho)}\right)^p}\right]^{1/p} =$$

$$= \frac{p(p+1)}{(p+2)(p-1)}\left[\frac{(p-1)(p+1)(1-\rho)(1+\rho(n-1))}{(1+\rho(n-1))\left((p+1)^2(1+\rho(n-1))-\rho np(p+2)\right)}\right]^{1/p} =$$

$$= \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(1-\rho)(p+1)}{(p+1)^2(1-\rho)+\rho n}\right)^{1/p}$$

Result 2.1. For any $\mathbf{X} \in \Psi_{n\times p}(0,1)$ with the variance matrix of errors $\sigma^2\mathbf{G} = \sigma^2\left((1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}\right)$, the efficiency of the design is defined as:

$$\mathrm{D}_{\mathrm{eff}} = \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(1-\rho)(p+1)}{(p+1)^2(1-\rho)+\rho n}\right)^{1/p}. \tag{2.4}$$

If $\rho = 0$, then we obtain the condition given in Ceranka and Graczyk (2019), and in such case $\mathrm{D}_{\mathrm{eff}} = \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{p-1}{p+1}\right)^{1/p}$.

If $\rho \to 1$, then $\mathrm{D}_{\mathrm{eff}} \to 0$, when $\rho \to \frac{-1}{n-1}$, then $\mathrm{D}_{\mathrm{eff}} \to \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(p+1)}{p(p+2)}\right)^{1/p}$. Moreover, $\lim_{p\to\infty} \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(p+1)}{p(p+2)}\right)^{1/p} = 1$. So, if $\rho \to \frac{-1}{n-1}$, then $\mathrm{D}_{\mathrm{eff}} \to 1$.

Next, let us consider the efficiency as a function of ρ. $\mathrm{D}_{\mathrm{eff}}^{'}(\rho) = \zeta(\rho)\cdot\xi(\rho)$, where:

$$\zeta(\rho) = \frac{(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(1-\rho)(p+1)}{(p+1)^2(1-\rho)+\rho n}\right)^{(1-p)/p},$$

$$\xi(\rho) = \frac{(p-1)(p+1)\left((p+1)^2(1-\rho)+\rho n\right)-(p-1)(1-\rho)(p+1)\left(n-(p+1)^2\right)}{\left((p+1)^2(1-\rho)+\rho n\right)^2}.$$

After calculating, we get $\mathrm{D}_{\mathrm{eff}}^{'} = \zeta(\rho)\cdot\frac{-n\left(p^2-1\right)}{\left((p+1)^2(1-\rho)+\rho n\right)^2}$. Because $\mathrm{D}_{\mathrm{eff}}^{'} < 0$, then $\mathrm{D}_{\mathrm{eff}}$ is a decreasing function under condition $\rho \to \frac{-1}{n-1}$.

Based on Graczyk and Ceranka (2022b), we have:

Result 2.2. Let $\mathbf{X} \in \Phi_{n\times p}\{0,1\}$ and the condition $\mathbf{X}\mathbf{1}_p = k\mathbf{1}_n$ be satisfied. If $p \equiv 0 mod 2$, then $n = 2(p-1)$ and $\det(\mathbf{X}'\mathbf{X}) \le (p-1)(0.5p)^p$. The equality is fulfilled if and only if $\mathbf{X}'\mathbf{X} = 0.5\left(p\mathbf{I}_p + (p-2)\mathbf{1}_p\mathbf{1}'_p\right)$.

From the above-presented result, we obtain:

Corollary 2.1. Let $p \equiv 0(mod\,2)$. The design $\mathbf{X} \in \Phi_{n\times p}\{0,1\}$ with the variance matrix of errors $\sigma^2\mathbf{G} = \sigma^2\left((1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}\right)$ is highly D-efficient if and only if $n = 2(p-1)$ and under this condition $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le \frac{(p-1)(1-\rho)}{1+\rho(2p-3)}\left(\frac{p}{2(1-\rho)}\right)^{p}$.

Thus:

Corollary 2.2. The efficiency of any $\mathbf{X} \in \Psi_{n\times p}(0,1)$ with the variance matrix of errors $\sigma^2\mathbf{G} = \sigma^2\left((1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^{'}\right)$ equals:

$$\mathrm{D}_{\mathrm{eff}} = \frac{p(p+1)}{(p-1)(p+2)}\left(\frac{(p-1)(1-\rho)(p+1)}{(p+1)^2(1-\rho)+2\rho(p-1)}\right)^{1/p}. \qquad (2.5)$$

Based on D-efficiency given in (2.5), we give the efficiency coefficients of spring balance weighing design $\mathbf{X} \in \Psi_{n \times p}(0,1)$ under condition $p \leq 20$ for different ρ.

Table 1. The efficiency of spring balance weighing design $\mathbf{X} \in \Psi_{n \times p}(0,1)$ under condition $p \leq 20$ for different ρ values

	ρ								
P	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
4	0.97149	0.96376	0.95426	0.94228	0.92670	0.90555	0.87501	0.82648	0.73346
6	0.98904	0.98454	0.97896	0.97186	0.96248	0.94952	0.93035	0.89874	0.83438
8	0.99440	0.99151	0.97790	0.98327	0.97710	0.96845	0.95543	0.93338	0.88645
10	0.99665	0.99465	0.99213	0.98889	0.98454	0.97838	0.96898	0.95275	0.91697
12	0.99779	0.99632	0.99447	0.99208	0.98889	0.98444	0.97714	0.96469	0.93647
14	0.99844	0.99732	0.99590	0.99407	0.99157	0.98800	0.98245	0.97259	0.94918
16	0.99884	0.99796	0.99684	0.99539	0.99341	0.99056	0.98610	0.97810	0.95918
18	0.99911	0.99839	0.99749	0.99631	0.99470	0.99237	0.98871	0.98209	0.96617
20	0.99929	0.99870	0.99796	0.99698	0.99565	0.99371	0.99065	0.98507	0.97148

Source: own calculation

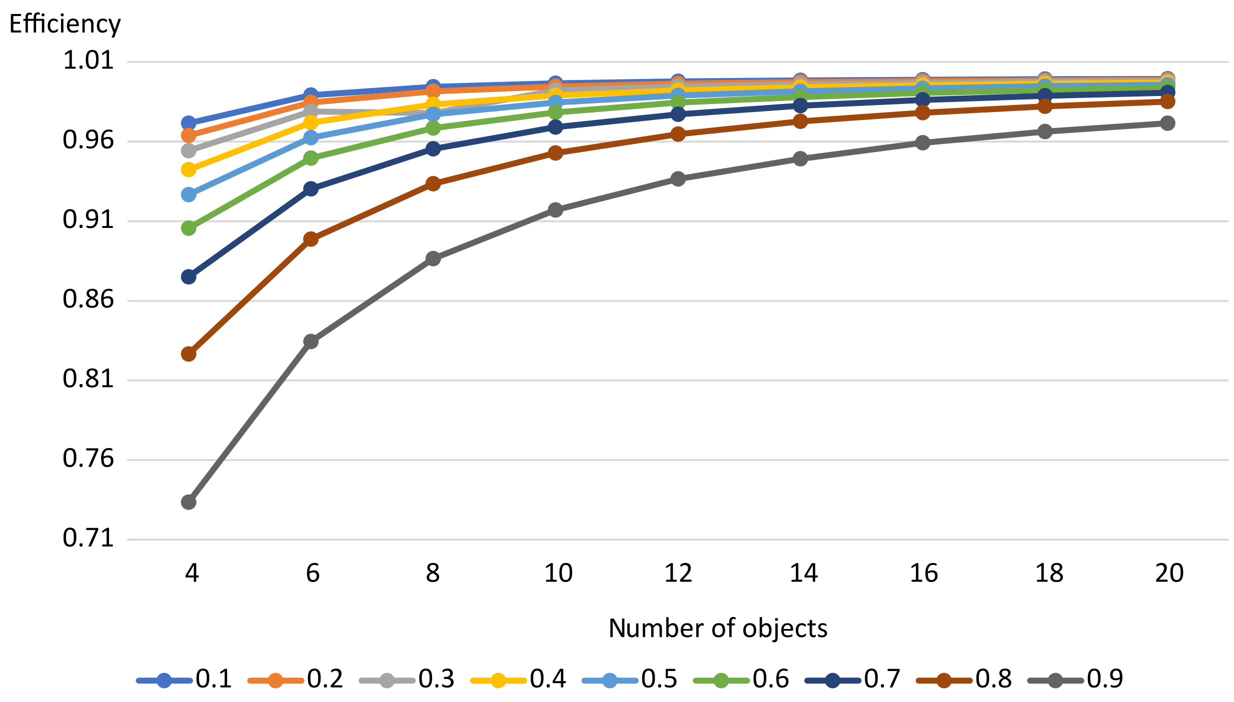

Figure 1. The efficiency coefficients of spring balance weighing design $\mathbf{X} \in \Psi_{n \times p}(0,1)$ under condition $p \leq 20$ for different ρ values: 0.1–0.9

3. Conclusions

Here, some issues related to the determining of the D-optimal spring balance weighing design are presented. For any combination of the numbers of objects p and measurements n, we are not able determine a D-optimal design. In such a case, highly D-efficient designs are proposed as the best designs for determining unknown measurements of p objects. In this case, the efficiency coefficient of spring balance weighing design was determined. Its properties were investigated. Moreover, for some ρ values, the efficiency coefficients were calculated and the trend of changes was presented.

References

Banerjee K.S. (1975), *Weighing Designs for Chemistry, Medicine, Economics, Operation Research, Statistics*, Marcell Dekker Inc, New York.

Bulutoglu D.A., Ryan K.J. (2009), *D-optimal and near D-optimal 2k fractional factorial designs of Resolution V*, "Journal of Statistical Planning and Inference", no. 139, pp. 16–22.

Ceranka B., Graczyk M. (2010), *Some construction of optimum weighing designs*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 235, pp. 235–239.

Ceranka B., Graczyk M. (2012), *Notes on the optimum chemical balance weighing designs*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 269, pp. 91–101.

Ceranka B., Graczyk M. (2014), *On certain A-optimal biased spring balance weighing designs*, "Statistics in Transition, New Series", vol. 15(2), pp. 317–326.

Ceranka B., Graczyk M. (2018), *Highly D-efficient designs for even number of objects*, "Revstat-Statistical Journal", no. 16, pp. 475–486.

Ceranka B., Graczyk M. (2019), *Recent developments in D-optimal designs*, "Communications in Statistics – Theory and Methods", vol. 48, no. 6, pp. 1470–1480.

Ceranka B., Katulska K. (1987a), *Zastosowanie optymalnych sprężynowych układów wagowych*, „Siedemnaste Colloquium Metodologiczne z Agro-Biometrii", PAN, pp. 98–108.

Ceranka B., Katulska K. (1987b), *Zastosowanie teorii sprężynowych układów wagowych do analizy doświadczeń z mieszankami*, "Listy Biometryczne", no. XXIV, pp. 17–26.

Ceranka B., Katulska K. (1989), *Application of the biased spring balance weighing theory to estimation of differences of line effects for legume content*, "Biometrical Journal", no. 31, pp. 103–110.

Gail Z., Kiefer J. (1982), *Construction methods for D-optimum weighing designs when* $n \equiv 3 \, mod \, 4$, "The Annals of Statistics", no. 10, pp. 502–510.

Gawande B.N., Patkar A.Y. (1999), *Application of factorial design for optimization of Cyclodextrin Glycosyltransferase production from Klebsiella Pneumoniae AS–22*, "Biotechnology and Bioengineering", vol. 64, no. 2, pp. 168–173.

Graczyk M. (2013), *Some applications on weighing designs*, "Biometrical Letters", vol. 50(1), pp. 15–26.

Graczyk M., Ceranka B. (2022a), *Regular D-optimal spring balance weighing designs with correlated errors*, "Communication in Statistics: Theory and Methods" [to appear].

Graczyk M., Ceranka B. (2022b), *Contribution to spring balance weighing designs*, "Biometrical Letters", vol. 59(1).

Jacroux M., Notz W. (1983), *On the Optimality of Spring Balance Weighing Designs*, "The Annals of Statistics", no. 11, pp. 970–978.

Katulska K., Smaga Ł. (2010), *On some construction of D-optimal chemical balance weighing designs*, "Colloquium Biometricum", no. 40, pp. 37–45.

Koukouvinos Ch. (1996), *Linear models and D-optimal designs for* $n \equiv 2 \, mod \, 4$, "Statistics and Probability Letters", no. 26, pp. 329–332.

Raghavarao D. (1971), *Constructions and Combinatorial Problems in Designs of Experiments*, John Wiley Inc., New York.

Shah K.R., Sinha B.K. (1989), *Theory of Optimal Designs*, Springer-Verlag, Berlin.

Uwagi o efektywnych sprężynowych układach wagowych o skorelowanych błędach dla parzystej liczby obiektów

Streszczenie: W artykule przeanalizowano problematykę związaną z efektywnością wybranych klas sprężynowych układów wagowych, przy założeniu, że błędy pomiarów są skorelowane. Podano zależności pomiędzy parametrami tych układów.

Słowa kluczowe: układ D-efektywny, sprężynowy układ wagowy

JEL: C02, C18, C90

Received: 2022-07-26; revised: 2022-11-17. Accepted: 2023-02-09