



OPEN

## Saddlepoint p-values for a class of location-scale tests under randomized block design

Haidy N. Mohamed, Ehab F. Abd-Elfattah, Amel Abd-El-Monem & Abd El-Raheem M. Abd El-Raheem

This paper deals with a class of nonparametric two-sample location-scale tests. The purpose of this paper is to approximate the exact p-value of the considered class under a randomized block design. The exact p-value of the considered class is approximated by the saddlepoint approximation method, also by the traditional method which is the normal approximation method. The saddlepoint approximation method is more accurate than the normal approximation method in approximating the exact p-value, and does not take a lot of time like the simulation method. This accuracy is proved by applying the mentioned methods to two real data sets and a simulation study.

The tests for location-scale problem have many uses in many fields, especially when studying spatial and variations in streamflow, Zhang et al.<sup>1</sup> and Yang et al.<sup>2</sup>, also in monsoon circulation, see Kwoon et al.<sup>3</sup>. In addition, these tests arise very often in bioinformatics in the case of comparing two groups such as control and treatment in order to detect differentially expressed genes, see Neuhäuser and Senske<sup>4</sup>. There is also an importance for location-scale tests in the field of biomedical and clinical trials, when the drug or the treatment is given for the treatment group causes changes in location and scale parameters between the two groups, see Mulccioli et al.<sup>5</sup>, Rice et al.<sup>6</sup>, Lunde and Timmermann<sup>7</sup> and Marozzi<sup>8</sup>. This paper presents some location-scale linear rank tests that can form the proposed class. Lepage<sup>9</sup> test “LP1”, was the first test that combines both, the location test which is the Wilcoxon test and the scale test which is the Ansari-Bradley test. In addition, Büning and Thadewald<sup>10</sup> presented LP2, LP3 and LP4 tests which are known as modified tests for LP1. Furthermore, Rublik<sup>11</sup> proposed a test statistic consisting of a linear combination of the location test which is the Wilcoxon test and the scale test which is the Mood test. Rousson<sup>12</sup> discussed the location-scale test in case of a multivariate two-sample problem. All the aforementioned tests can be used when all observations in the population or selected samples are determined. Therefore, these tests belong to classical statistics. While if the data to be analyzed or a part of it is indeterminate, we can resort to neutrosophic statistics. For more information on this context, a number of references can be referred to, namely Smarandache<sup>13</sup>, Aslam<sup>14–16</sup>, Afzal et al.<sup>17</sup>, Albassam et al.<sup>18</sup> and Sherwani et al.<sup>19</sup>. If you interested in location shift only, you could see Hollander et al.<sup>20</sup>. They suggested many location problems and tests, with one sample and two samples.

In clinical trials, consider we have  $N$  patients will be distributed into two groups, control group and treatment group. To guarantee no selection bias in the randomization assignment for the subjects, and to achieve a certain degree of balance between the control and treatment groups, randomization designs must be used. Complete randomization design is one of the easiest ways to assign the subject into two groups using a fair coin, if the head, it goes to the control but if the tail, it goes to the treatment group, and so on until we finish all subjects. There is also a random allocation design, which depends on one of the  $\binom{N}{\frac{N}{2}}$  permutations to assign the  $N$  patients to the groups. The randomized block design is an important design that reduces selection bias and control the imbalance of group sizes. In this paper, we tend to use randomized block design, which contains  $k$  blocks of even sizes  $n_j$ , such that  $\frac{n_j}{2}$  patients are for control group and the same number for treatment group, and within each block the random allocation design is applied to assign the patients to the two groups. For more randomization designs, see Rosenberger and Lachin<sup>21</sup>.

To calculate the exact p-value of the proposed class, an accurate method is used, which is saddlepoint approximation method “SPA” that depends on the permutation distribution of the considered class. The SPA method discovered by Daniels<sup>22</sup> who investigated an approximate formula for any probability mass or density function based on its moment generating function. The theory of SPA expansions and multidimensional

Department of Mathematics, Faculty of Education, Ain Shams University, Cairo, Egypt. email: [abdelrahimmohamed@edu.asu.edu.eg](mailto:abdelrahimmohamed@edu.asu.edu.eg)

generalizations were presented in Good<sup>23,24</sup> and Barndorff-Nielsen and Cox<sup>25</sup>. Further developed by Lugannani and Rice<sup>26</sup> who proposed an approximate formula for the cumulative distribution function, CDF. Saddlepoint approximations to randomization distributions and permutation distributions were treated by Robinson<sup>27</sup> and Davison and Hinkley<sup>28</sup>. Skovgaard<sup>29</sup> introduced an approximate formula for the conditional distribution function which is a generalization of the Lugannani and Rice<sup>26</sup> formula, for more details, see Booth and Butler<sup>30</sup> and Butler<sup>31</sup>. The SPA has many advantages, first, the SPA has a correction term equal to  $O(n^{-\frac{3}{2}})$ , in contrast the central limit theorem has  $O(n^{-\frac{1}{2}})$ , Daniels<sup>32</sup>. Second, the SPA often leads to a uniformly bounded error, in contrast, the errors of the normal approximation method generally increase in the tail of the distribution. Daniels<sup>33</sup> was the first who suggested the major advantages of the SPA method and its general power and scope. Many papers were interested in approximating the exact p-value for various classes using SPA method with different randomization designs, such as, Abd-Elfattah and Butler<sup>34,35</sup>, Abd-Elfattah<sup>36–39</sup>, Abd EL-Raheem and Abd-Elfattah<sup>40,41</sup>, Kamal et al.<sup>42,43</sup> and Abd El-Raheem et al.<sup>44,45</sup>.

The exact distribution of the considered class of tests is unknown. Then, we cannot obtain the exact p-values for such tests. Therefore, we resort to the saddlepoint approximation method to approximate the exact p-values for such tests. In all previous studies related to the considered class of tests, the normal approximation method was used to approximate the exact p-values of such tests. In the current study, we use the saddlepoint approximation method to approximate the exact p-value for the considered tests. From the results of simulation study and real data analysis, it will be seen that saddlepoint approximation p-values are almost always closer to the simulated (permutation) p-values than the normal approximation. The degree of greater accuracy is readily apparent in small and intermediate size samples for which the asymptotic normality has not been attained. Thus, the saddlepoint method is a more accurate approximation method than the normal approximation method and computationally less demanding than the simulation method (permutation based, so time consuming).

This paper is partitioned as follows: “Class of location-scale tests” presents the class of non-parametric location-scale tests. The saddlepoint approximation is presented in “Saddlepoint approximation”. Section “Simulation study and real data examples” proves the accuracy of SPA in approximating the exact p-value by performing a simulation study and analyzing three real data sets. Moreover, the time consumed for calculating the SPA p-values, normal approximation p-values and simulated mid-p-values is calculated in minutes and presented in “Simulation study and real data examples”. Furthermore, the 95% and 99% confidence intervals for location and scale parameters are constructed in “Confidence intervals for location and scale parameters”. Finally, the conclusion and discussion are presented in “Conclusion”.

### Class of location-scale tests

Consider two independent samples  $X$  which is the control group and  $Y$  is the treatment group are drawn from populations with CDF  $F$  and  $G$ , with means  $\mu_1$  and  $\mu_2$  and standard deviations  $\sigma_1$  and  $\sigma_2$ , respectively. Under the randomized block design, the location-scale class is given by

$$H_B = \sum_{i=1}^k b_i \sum_{j=1}^{n_i} (a_{L_{ij}} + a_{S_{ij}}) Z_{ij}, \quad i = 1, \dots, k, j = 1, \dots, n_i, \quad (1)$$

where  $k$  is the number of blocks with even sizes  $n_i$ ,  $N = \sum_{i=1}^k n_i$ , and  $b_i = \frac{1}{m_i+1}$  is the optimum weight of block  $i$ , Elteren<sup>46</sup>. The location and scale scores of each observation  $j$  in each block  $i$  are denoted by the linear combination  $(a_{L_{ij}} + a_{S_{ij}})$ , where  $a_{L_{ij}}$  is for location score and  $a_{S_{ij}}$  is for scale score, and  $Z_{ij}$  is the group indicator takes the value 1 if the observation  $j$  in the block  $i$  is from the treatment group  $Y$  and takes the value 0 otherwise. The permutation distribution of the observations assignments within the blocks, is done under random allocation rule with  $\binom{n_i}{m_i}$  possible permutations, where  $m_i$  is the number of the treatment observations inside the block  $i$ . The asymptotic distribution of  $H_B$  is  $N(\mu_L + \mu_S, \sigma_L + \sigma_S)$ , where  $\mu_L$  and  $\mu_S$  are the means of the location and scale tests, respectively. Also,  $\sigma_L$  and  $\sigma_S$  are the standard deviations of the location and scale tests, respectively.

The considered class includes many of location-scale tests, such as, Lepage’s statistic which can be written according to the Eq. (1) as

$$LP1 = \sum_{i=1}^k b_i \sum_{j=1}^{n_i} (W_{L_{ij}} + R_{S_{ij}}) Z_{ij}, \quad (2)$$

where  $W_{L_{ij}}$  is the score of the Wilcoxon location test with mean  $\mu_L = \frac{1}{2} \sum_{i=1}^k b_i m_i (n_i + 1)$  and variance  $\sigma_L^2 = \frac{1}{12} \sum_{i=1}^k b_i^2 m_i^2 (n_i + 1)$ . Also,  $R_{S_{ij}}$  is the score of Ansari-Bradley scale test that takes the form

$$R_{S_{ij}} = 1, 2, 3, 4, \dots, \frac{n_i}{2}, \frac{n_i}{2}, \dots, 4, 3, 2, 1,$$

with mean

$$\mu_S = \frac{1}{4} \sum_{i=1}^k b_i m_i (n_i + 2),$$

and variance

$$\sigma_S^2 = \sum_{i=1}^k \frac{b_i^2 m_i^2 (n_i^2 - 4)}{48(n_i - 1)},$$

with asymptotic distribution  $N(\mu_L + \mu_S, \sigma_L + \sigma_S)$ . In addition,  $LP2$ , also called Gastwirth<sup>47</sup> test, can be written in the form of Eq. (1), as follows:

$$LP2 = \sum_{i=1}^k b_i \sum_{j=1}^{n_i} (a_{L_{ij}} + a_{S_{ij}}) U_{ij}, \quad (3)$$

where the location score is

$$a_{L_{ij}} = \begin{cases} j - \frac{n_i+1}{4} & \text{if } j \leq \frac{n_i+1}{4} \\ 0 & \text{if } \frac{n_i+1}{4} < j < \frac{3(n_i+1)}{4} \\ j - \frac{3(n_i+1)}{4} & \text{if } j \geq \frac{3(n_i+1)}{4}, \end{cases}$$

and the scale score

$$a_{S_{ij}} = \begin{cases} \frac{n_i+1}{4} - j & \text{if } j \leq \frac{n_i+1}{4} \\ 0 & \text{if } \frac{n_i+1}{4} < j < \frac{3(n_i+1)}{4} \\ j - \frac{3(n_i+1)}{4} & \text{if } j \geq \frac{3(n_i+1)}{4}, \end{cases}$$

where

$$U_{ij} = \begin{cases} 1 & Z_{i(j)} \in Y \\ 0 & \text{otherwise,} \end{cases}$$

and  $Z_{i(j)}$  is the  $j$ -th order statistic of the combined two samples  $X$  and  $Y$  in the block  $i$ .

$LP3$  test takes the form of  $LP2$  in Eq. (3) and follows the form of location-scale class in Eq. (1) with location score of Van der Waerden test

$$a_{L_{ij}} = \varphi^{-1} \left( \frac{j}{n_i + 1} \right),$$

and Klotz scale score

$$a_{S_{ij}} = \left( \varphi^{-1} \left( \frac{j}{n_i + 1} \right) \right)^2,$$

where  $\varphi^{-1}$  is the inverse cumulative distribution function of standard normal distribution.

$LP4$  test also takes the same form of  $LP2$  and  $LP3$  but with location score

$$a_{L_{ij}} = \begin{cases} -(\lfloor \frac{n_i}{4} \rfloor + 1) & \text{if } j < (\lfloor \frac{n_i}{4} \rfloor + 1) \\ i - \frac{n_i+1}{2} & \text{if } (\lfloor \frac{n_i}{4} \rfloor + 1) \leq j \leq \lfloor \frac{3(n_i+1)}{4} \rfloor \\ \lfloor \frac{n_i}{4} \rfloor + 1 & \text{if } j > \lfloor \frac{3(n_i+1)}{4} \rfloor, \end{cases}$$

and with Mood scale score

$$a_{S_{ij}} = \left( j - \frac{n_i + 1}{2} \right)^2,$$

where  $[a]$  denotes the greatest integer less than or equal to  $a$ .

All Lepage's types,  $LP2$ ,  $LP3$ , and  $LP4$  are asymptotically distributed  $N(\mu, \sigma)$ , where

$$\mu = \sum_{i=1}^k \frac{b_i m_i}{n_i} \sum_{j=1}^{n_i} (a_{L_{ij}} + a_{S_{ij}}),$$

and

$$\sigma^2 = \sum_{i=1}^k \frac{b_i^2 m_i^2}{(n_i - 1)} \sum_{j=1}^{n_i} \left( (a_{L_{ij}} + a_{S_{ij}}) - \frac{\sum_{j=1}^{n_i} (a_{L_{ij}} + a_{S_{ij}})}{n_i} \right)^2.$$

In addition, Rublik<sup>11</sup> investigated the location-scale problem with test statistic contains a combination between the Wilcoxon location score and the Mood scale score, that can take the same form of (1) as follows:

$$T = \sum_{i=1}^k b_i \sum_{j=1}^{n_i} \left( j + \left( j - \frac{n_i + 1}{2} \right)^2 \right) Z_{ij},$$

with location mean  $\mu_L = \frac{1}{2} \sum_{i=1}^k b_i m_i (n_i + 1)$ , and variance  $\sigma_L^2 = \frac{1}{12} \sum_{i=1}^k b_i^2 m_i^2 (n_i + 1)$ , and scale mean  $\mu_S = \frac{1}{12} \sum_{i=1}^k b_i m_i (n_i^2 - 1)$  and variance  $\sigma_S^2 = \frac{1}{180} \sum_{i=1}^k b_i^2 m_i^2 (n_i + 1) (n_i^2 - 4)$ . The asymptotic distribution of  $T$  is  $N(\mu_L + \mu_S, \sigma_L + \sigma_S)$ .

In this paper, we are interested in working with three test statistics only, which are *LP1* test, *LP3* test and Rublik test. For more location-scale tests, see Duran et al.<sup>48</sup> who investigated a class of location-scale nonparametric tests. Also see, Fueda and Ôhori<sup>49</sup> who designed a two-sample rank test based on the Wilcoxon test.

### Saddlepoint approximation

For simplicity, let the class (Eq. 1) be in the form

$$H_B = \sum_{i=1}^k \sum_{j=1}^{n_i} A_{ij} Z_{ij}, i = 1, \dots, k, j = 1, \dots, n_i, \tag{4}$$

where  $A_{ij} = b_i(a_{L_{ij}} + a_{S_{ij}})$ , as we noted before, that the patients within the blocks distributed under random allocation design, this means that the random variables  $Z_{ij}$  and  $Z_{ib}$  for all  $i = 1, \dots, k$  and  $j \neq b$  are dependent but independent with  $Z_{aj}$  where  $a \neq i$ .

To avoid the problem of the dependence, we constructed a conditional distribution as follow:

$$Z_{11}, \dots, Z_{knk} \sim V_{11}, \dots, V_{knk} \mid \sum_{j=1}^{n_1} V_{1j} = m_1, \dots, \sum_{j=1}^{n_k} V_{kj} = m_k,$$

where  $V_{i1}, \dots, V_{in_i}$  are independent and identically Bernoulli ( $\theta_i$ ) random variables for each  $i = 1, \dots, k$ .

This transfers the distribution of the statistic (Eq. 4) to equivalent conditional distribution as follows:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} A_{ij} Z_{ij} \sim \sum_{i=1}^k \sum_{j=1}^{n_i} A_{ij} V_{ij} \mid \sum_{j=1}^{n_1} V_{1j} = m_1, \dots, \sum_{j=1}^{n_k} V_{kj} = m_k.$$

Now, to approximate exact p-value of  $H_B$  in (Eq. 4), we need to approximate the following conditional probability

$$P \left( \sum_{i=1}^k \sum_{j=1}^{n_i} A_{ij} V_{ij} \geq h_o \mid \sum_{j=1}^{n_1} V_{1j} = m_1, \dots, \sum_{j=1}^{n_k} V_{kj} = m_k \right), \tag{5}$$

where  $h_o$  is the observed value of  $H_B$ , using the double saddlepoint approximation of Skovgaard<sup>29</sup>, the conditional probability in (Eq. 5) can be approximated as follows:

$$\text{midp}(h_o) \simeq 1 - \Phi(\hat{\omega}) - \phi(\hat{\omega}) \left( \frac{1}{\hat{\omega}} - \frac{1}{\hat{u}} \right),$$

where

$$\hat{\omega} = \text{sgn}(\hat{t}) \sqrt{2 \left[ \left\{ Q(0, \hat{S}_0) - M^T \hat{S}_0 \right\} - \left\{ Q(\hat{t}, \hat{S}) - M^T \hat{S} - h_o \hat{t} \right\} \right]},$$

$$\hat{u} = \hat{t} \sqrt{\left| Q''(\hat{t}, \hat{S}) \right| / \left| Q''_{ss}(0, \hat{S}_0) \right|},$$

where  $M = (m_1, \dots, m_k)$ , the two saddlepoints are  $(\hat{t}, \hat{S}) = (\hat{t}, \hat{s}_1, \dots, \hat{s}_k)$  and  $\hat{S}_0 = (\hat{s}_{10}, \dots, \hat{s}_{k0})$ .

The joint cumulant generating function of  $\left\{ \sum_{i=1}^k \sum_{j=1}^{n_i} A_{ij} V_{ij}, \sum_{j=1}^{n_1} V_{1j} = m_1, \dots, \sum_{j=1}^{n_k} V_{kj} = m_k \right\}$  is

$$Q(t, S) = \sum_{i=1}^k \sum_{j=1}^{n_i} \ln \left\{ (1 - \theta_i) + \theta_i \exp \left( s_i + t(a_{L_{ij}} + a_{S_{ij}}) \right) \right\},$$

where  $Q(\hat{t}, \hat{S})$  is  $(k + 1) \times (k + 1)$  Hessian matrix and  $Q''_{ss}$  is the second derivative of  $Q(0, \hat{S}_0)$  with respect to  $S$ .

To calculate  $\hat{\omega}$  and  $\hat{u}$ , we first calculate the numerator saddlepoints  $(\hat{t}, \hat{S}) = (\hat{t}, \hat{s}_1, \dots, \hat{s}_k)$ , by solving the following equations

$$Q_{S_i}'(\hat{t}, \hat{S}) = \sum_{i=1}^N \frac{\theta_i \exp(s_i + t(a_{L_{ij}} + a_{S_{ij}}))}{(1 - \theta_i) + \theta_i \exp(s_i + t(a_{L_{ij}} + a_{S_{ij}}))} = m_i, i = 1, \dots, k,$$

$$Q_t'(\hat{t}, \hat{S}) = \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\theta_i \exp(s_i + t(a_{L_{ij}} + a_{S_{ij}})) (a_{L_{ij}} + a_{S_{ij}})}{(1 - \theta_i) + \theta_i \exp(s_i + t(a_{L_{ij}} + a_{S_{ij}}))} = h_0,$$

and to find the value of  $\hat{S}_0 = (\hat{s}_{10}, \dots, \hat{s}_{k0})$ , we solve the following equation

$$Q_{S_i}'(0, \hat{S}_0) = \sum_{j=1}^{n_i} \frac{\theta_i \exp(\hat{s}_{i0})}{(1 - \theta_i) + \theta_i \exp(\hat{s}_{i0})} = m_i, i = 1, \dots, k.$$

Each of  $\hat{\omega}$  and  $\hat{u}$  dose not depend on  $\theta_i$ , so for explicitly in solving the SPA equations, we can choose  $\theta_i = \frac{m_i}{n_i}$ , then  $\hat{S}_0 = (0, \dots, 0)$  for all  $i = 1, \dots, k$ .

### Simulation study and real data examples

#### Simulation study

The main aim of using a simulation study is to prove that the saddlepoint approximation method is closer to the simulated mid-p-value than the normal approximation method. The exact p-values of Lepage test, LP3 test and Rublik test are approximated using the two methods, saddlepoint approximation and normal approximation methods. To illustrate the accuracy of SPA p-value, we compare the p-values for the two previously methods with the simulated mid-p-value, which can be calculated from simulation of one million random permutations of the treatment and control indicator  $Z_{ij}$ . The simulated mid-p-value is obtained as  $[\sum I(H > h_0) + 0.5 \sum I(H = h_0)]/10^6$ , and its donated here as “exact” p-value. The benefit of using mid-p-value instead of the p-value is that the mid-p-value is convenient in case of discrete test statistics and not conservative compared to the ordinary p-value, for more details, see Agresti<sup>50</sup>, Butler<sup>31</sup> and Delanchy et al.<sup>51</sup>. The simulated data in this section is generated from extreme value and logistic distributions with six cases which are (1):  $N = 24, k = 4, m = 3$ , case (2):  $N = 30, k = 3, m = 5$ , case (3):  $N = 40, k = 4, m = 5$ , case (4):  $N = 60, k = 6, m = 5$ , case (5):  $N = 80, k = 5, m = 8$  and case (6):  $N = 90, k = 5, m = 5$ , each with  $n_i = \frac{N}{k}$ , and  $m_i = m$ . Different location ( $\mu_1, \mu_2$ ) and scale ( $\sigma_1, \sigma_2$ ) parameters are used to generate the data according to the previous scenarios. To prove our aim this process is repeated 1000 times based on 1000 generated samples. Tables 1, 2 and 3 present the mean of the SPA p-values, normal approximation p-values, simulated mid-p-values and the percentage of approaching the SPA method to the simulated method “P.SPA”. Also, the average relative absolute error for SPA “R.E.SPA” and the average relative absolute error for normal approximation method “R.E.NA” are calculated and presented in Tables 1, 2 and 3.

In Table 1, 2 and 3, the SPA approximation is more accurate than the normal approximation method. This can be seen through the R.E.SPA, which is much smaller than R.E.NA, for all considered cases.

Lepage test						
Cases	$N = 24, k = 4, m = 3$	$N = 30, k = 3, m = 5$	$N = 40, k = 4, m = 5$	$N = 60, k = 6, m = 5$	$N = 80, k = 5, m = 8$	$N = 90, k = 9, m = 5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 1, 3, 2	3, 1, 1, 2	0, 3, 3, 3.5	0, 1, 3, 2	0, 1, 0.1, 1.5	0, 1, 3, 2
Exact p-value	0.2442	0.0202	0.0454	0.1382	0.0221	0.1085
SPA p-value	0.2426	0.0202	0.0452	0.1380	0.0221	0.1084
Normal p-value	0.2416	0.0207	0.0453	0.1378	0.0222	0.1084
P.SPA	80.1	95.6	84.1	76.0	86.1	73.4
R.E.SPA	0.0298	0.0004	0.0476	0.0216	0.1497	0.0253
R.E.NA	0.0601	0.0010	0.4318	0.0616	0.5431	0.0764
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	3, 1, 5, 10	0, 3, 1, 1.5	0, 3, 1, 2	3, 1, 5, 10	0, 3, 3, 2	3, 1, 5, 10
Exact p-value	0.2661	0.0306	0.0640	0.1586	0.0094	0.1090
SPA p-value	0.2649	0.0301	0.0635	0.1586	0.0094	0.1089
Normal p-value	0.2642	0.0304	0.0635	0.1585	0.0095	0.1089
P.SPA	81	85.3	82.0	77.5	89.6	74.6
R.E.SPA	0.0061	0.0917	0.0482	0.0019	0.1196	0.0007
R.E.NA	0.0110	1.7408	0.4445	0.0025	0.4601	0.0010

**Table 1.** Outcomes of the simulation study for the Lepage test.

LP3 test						
Cases	$N = 24, k = 4, m = 3$	$N = 30, k = 3, m = 5$	$N = 40, k = 4, m = 5$	$N = 60, k = 6, m = 5$	$N = 80, k = 5, m = 8$	$N = 90, k = 9, m = 5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	1, 0, 0.1, 1	3, 0, 5, 1	3, 0, 1, 2	1, 0, 0.1, 1	3, 0, 50, 10	1, 0, 0.1, 1
Exact p-value	0.5370	0.0025	0.1277	0.3897	0.0003	0.3677
SPA p-value	0.5366	0.0024	0.1273	0.3895	0.0003	0.3676
Normal p-value	0.5367	0.0035	0.1270	0.3893	0.0004	0.3675
P.SPA	81.7	99.4	93.8	91.6	99.3	85.1
R.E.SPA	0.0086	$6.2 \times 10^{-5}$	0.0008	0.0055	$7.6 \times 10^{-6}$	0.0091
R.E.NA	0.0205	0.0011	0.0034	0.0423	$6.6 \times 10^{-5}$	0.0390
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 2, 5, 15	3, 0, 50, 10	3, 0, 50, 10	0, 2, 5, 15	3, 0, 5, 7	0, 2, 5, 15
Exact p-value	0.2043	0.0940	0.0725	0.0539	0.4974	0.0241
SPA p-value	0.2033	0.0934	0.0721	0.0537	0.4974	0.0240
Normal p-value	0.2026	0.0934	0.0723	0.0539	0.4974	0.0243
P.SPA	83.5	89.2	94.2	92.5	89.3	91.9
R.E.SPA	0.0202	0.0009	0.0005	0.0387	0.0024	0.0765
R.E.NA	0.2401	0.0037	0.0023	0.5868	0.0097	0.5964

**Table 2.** Outcomes of the simulation study for the LP3 test.

Rublik test						
Cases	$N = 24, k = 4, m = 3$	$N = 30, k = 3, m = 5$	$N = 40, k = 4, m = 5$	$N = 60, k = 6, m = 5$	$N = 80, k = 5, m = 8$	$N = 90, k = 9, m = 5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	1, 3, 1, 2	1, 3, 0.1, 0.01	1, 3, 0.1, 0.5	1, 3, 1, 2	0, 3, 1, 2	1, 3, 1, 2
Exact p-value	0.0737	0.0478	0.0268	0.0488	0.0853	0.0210
SPA p-value	0.0733	0.0481	0.0266	0.0487	0.0853	0.0210
Normal p-value	0.0735	0.0481	0.0274	0.0488	0.0850	0.0212
P.SPA	65.7	84.6	81.9	82.7	71.7	83.8
R.E.SPA	0.0465	0.0067	0.0086	0.0155	0.0052	0.0307
R.E.NA	0.1524	0.0080	0.0220	0.0807	0.0091	0.1262
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 0.01, 1, 0.5	0, -0.1, 1, 1.5	0, 0.01, 1, 0.5	0, 0.1, 1, 1.5	0, 0.001, 1.5, 1	0, 0.01, 1, 0.5
Exact p-value	0.2644	0.2600	0.1138	0.0906	0.1147	0.0486
SPA p-value	0.2627	0.2597	0.1135	0.0904	0.1146	0.0485
Normal p-value	0.2619	0.2588	0.1133	0.0904	0.1144	0.0486
P.SPA	76	81.9	83.1	82.6	81.2	81.1
R.E.SPA	0.0064	0.0094	0.0009	0.0005	0.0004	0.0003
R.E.NA	0.0098	0.0436	0.0019	0.0011	0.0008	0.0005

**Table 3.** Outcomes of the simulation study for the Rublik test.

### Real data examples

To support the aim of this paper, three real data sets are analyzed. The first data set is from Rosenberger and Lachin<sup>21</sup>. They analyzed cholesterol rate for 50 patients. The results of the cholesterol rate for 50 patients can be found in Table 7.4 in the reference Rosenberger and Lachin<sup>21</sup>. The 50 patients were assigned randomly to control and treatment group by generating the vector of the group indicator  $Z_{ij}$ , such that 25 assign to control and 25 assign to treatment, where each block contains 5 from each group, i.e. ( $N = 50, n_i = 10, k = 5, m = 5$ ). The second data set is from a survey of household expenditure for 20 single men "treatment group" and 20 single women "control group". For this data set, ( $N = 40, n_i = 8, k = 5, m = 4$ ). The second data is presented in Büning and Thadewald<sup>10</sup>. The third data set was presented on page 39 of Hand et al.<sup>52</sup>. This data set consists of 40 measurements of cholesterol levels for 40 men were divided into two groups A and B according to two types of behaviors. The type A behavior "treatment group" is characterized by urgency and aggression. While type B behavior "control group" is relaxed. For this data set, ( $N = 40, n_i = 8, k = 5, m = 4$ ). Table 4 presents the p-values of LP1 test, LP3 test and Rublik test using simulated, SPA and normal approximation methods.

From Table 4, we can see that SPA p-value is closer to the exact p-value than the normal p-value, and this result gives more evidence that SPA method is more accurate than the normal method in approximating the

Data sets	LP1			LP3			Rublik		
	SPA	Exact	Normal	SPA	Exact	Normal	SPA	Exact	Normal
1st	0.2397	0.2403	0.2381	0.2767	0.2773	0.2733	0.2705	0.2705	0.2710
2nd	0.3655	0.3681	0.3638	0.0135	0.0138	0.0152	0.0093	0.0094	0.0104
3rd	0.0128	0.0130	0.0137	0.1216	0.1219	0.1192	0.3934	0.3911	0.3975

**Table 4.** P-values for simulated, SPA and normal approximations for the three data sets.

p-value. It remains for us to explain the reason for considering the saddlepoint approximation method as an alternative to the simulation method. The reason is that the saddlepoint approximation method requires much less computing time compared to the simulation method. To clarify this, the computing time for the different methods is calculated and this is summarized in Tables 5, 6 and 7.

From the result of the simulation study, we can see that the SPA method is more accurate than the normal approximation methods compared to the simulated exact p-value. Moreover, from Tables 5, 6 and 7 it is clear that the SPA method is faster than simulated method which needs a lot of time to approximate the exact p-value.

### Confidence intervals for location and scale parameters

The estimated confidence intervals for location parameter  $\mu_2$  and scale parameter  $\sigma_2$ , are the set of all values  $\mu_{2_0}$  and  $\sigma_{2_0}$  of the parameters  $\mu_2$  and  $\sigma_2$ , respectively, which if formulated in the claim  $H_0 : \mu_2 = \mu_{2_0}$  and  $\sigma_2 = \sigma_{2_0}$ , would not be rejected at  $\alpha$  significant level. Accordingly, if  $p(\mu_{2_0}, \sigma_{2_0})$  is the one-sided p-value for the location-scale test, then a  $(1 - \alpha)100\%$  confidence intervals of  $\mu_2$  and  $\sigma_2$  can be constructed as  $\{\mu_{2_0} : \frac{\alpha}{2} \leq p(\mu_{2_0}, \sigma_{2_0}) \leq 1 - \frac{\alpha}{2}\}$  and  $\{\sigma_{2_0} : \frac{\alpha}{2} \leq p(\mu_{2_0}, \sigma_{2_0}) \leq 1 - \frac{\alpha}{2}\}$ , respectively, see<sup>34</sup>. Assume  $D_0(\mu_{2_0}, \sigma_{2_0})$  be the observed test statistic with location parameter  $\mu_{2_0}$  and scale parameter  $\sigma_{2_0}$ , using a satisfactory grid of  $\mu_{2_0}$  and  $\sigma_{2_0}$  values with suitable increasing, the cutoff  $D_0(\cdot, \cdot)$  is a step function in  $\mu_{2_0}$  and  $\sigma_{2_0}$ , that leads to incremental increases with increasing  $\mu_{2_0}$  and  $\sigma_{2_0}$ . Here, the 3<sup>rd</sup> real data set is used to illustrate the procedure for creating

Lepage test						
Cases	$N = 24, k = 4, m = 3$	$N = 30, k = 3, m = 5$	$N = 40, k = 4, m = 5$	$N = 60, k = 6, m = 5$	$N = 80, k = 5, m = 8$	$N = 90, k = 9, m = 5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 1, 3, 2	3, 1, 1, 2	0, 3, 3, 3.5	0, 1, 3, 2	0, 1, 0.1, 1.5	0, 1, 3, 2
Exact p-value	1504 min	1615 min	1613 min	1701 min	1617 min	1713 min
SPA p-value	5.13 min	5.14 min	6.12 min	5.62 min	5.23 min	5.33 min
Normal p-value	3.41 min	3.52 min	3.44 min	4.11 min	3.21 min	3.21 min
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	3, 1, 5, 10	0, 3, 1, 1.5	0, 3, 1, 2	3, 1, 5, 10	0, 3, 3, 2	3, 1, 5, 10
Exact p-value	1515 min	1512 min	1527 min	1621 min	1629 min	1723 min
SPA p-value	5.71 min	5.21 min	5.32 min	6.22 min	6.10 min	6.30 min
Normal p-value	3.21 min	3.01 min	3.22 min	3.41 min	3.41 min	3.22 min

**Table 5.** The time consumed for calculating the SPA p-values, normal approximation p-values and simulated mid-p-values for Lepage test.

Rublik test						
Cases	$N = 24, k = 4, m = 3$	$N = 30, k = 3, m = 5$	$N = 40, k = 4, m = 5$	$N = 60, k = 6, m = 5$	$N = 80, k = 5, m = 8$	$N = 90, k = 9, m = 5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	1, 3, 1, 2	1, 3, 0.1, 0.01	1, 3, 0.1, 0.5	1, 3, 1, 2	0, 3, 1, 2	1, 3, 1, 2
Exact p-value	1603 min	1613 min	1502 min	1528 min	1619 min	1624 min
SPA p-value	5.21 min	5.66 min	5.24 min	5.44 min	5.55 min	5.36 min
Normal p-value	2.99 min	3.12 min	2.97 min	3.25 min	3.75 min	3.19 min
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 0.01, 1, 0.5	0, -0.1, 1, 1.5	0, 0.01, 1, 0.5	0, 0.1, 1, 1.5	0, 0.001, 1.5, 1	0, 0.01, 1, 0.5
Exact p-value	1533 min	1630 min	1623 min	1525 min	1522 min	1623 min
SPA p-value	6.00 min	6.25 min	6.61 min	5.43 min	5.31 min	5.29 min
Normal p-value	3.24 min	3.55 min	3.76 min	3.42 min	3.46 min	3.12 min

**Table 6.** The time consumed for calculating the SPA p-values, normal approximation p-values and simulated mid-p-values Rublik test.

LP3 test						
Cases	$N=24, k=4, m=3$	$N=30, k=3, m=5$	$N=40, k=4, m=5$	$N=60, k=6, m=5$	$N=80, k=5, m=8$	$N=90, k=9, m=5$
<b>Extreme value distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	1, 0, 0.1, 1	3, 0, 5, 1	3, 0, 1, 2	1, 0, 0.1, 1	3, 0, 50, 10	1, 0, 0.1, 1
Exact p-value	1051 min	1060 min	1072 min	1075 min	1067 min	1173 min
SPA p-value	6.20 min	6.54 min	5.61 min	6.43 min	6.21 min	6.49 min
Normal p-value	3.34 min	3.35 min	3.16 min	3.22 min	3.16 min	3.92 min
<b>Logistic distribution</b>						
$\mu_1, \mu_2, \sigma_1, \sigma_2$	0, 2, 5, 15	3, 0, 50, 10	3, 0, 50, 10	0, 2, 5, 15	3, 0, 5, 7	0, 2, 5, 15
Exact p-value	1054 min	1057 min	1162 min	1167 min	1267 min	1263 min
SPA p-value	5.00 min	5.45 min	5.21 min	5.42 min	6.31 min	6.31 min
Normal p-value	3.23 min	3.16 min	3.77 min	3.52 min	3.86 min	3.97 min

**Table 7.** The time consumed for calculating the SPA p-values, normal approximation p-values and simulated mid-p-values for LP3 test.

confidence intervals for the location and scale parameters. We use the “gofTest” R package to estimate the location  $\mu_2$  and scale  $\sigma_2$  parameters for the 3<sup>rd</sup> real data set and to test the suitability of the extreme value distribution for the considered real data set. The maximum likelihood estimations for the location and scale parameters are 227.9 and 31.17, respectively. Furthermore, the p-value of the goodness of fit test, is  $p\text{-value} = 0.887$ . We evaluate the values of  $D_o(\mu_{2_o}, \sigma_{2_o})$  using a large range of the possible values of  $\mu_{2_o}$  and  $\sigma_{2_o}$ , then for each value of  $D_o(\mu_{2_o}, \sigma_{2_o})$  the corresponding exact, saddlepoint and normal p-values are calculated. Table 8 includes the exact, saddlepoint and normal confidence intervals for LP1 test.

From Table 8, we can see that the estimated 99% confidence interval using SPA method is more accurate than the corresponding estimated confidence interval using the normal approximation method as compared to the simulated (Exact) confidence interval. For the 95% confidence interval, both methods have the same accuracy as the simulated method.

## Conclusion

In this article, various nonparametric tests for location and scale problem have been discussed and rewritten as a common linear rank class. The exact p-value of the considered class is approximated by SPA method and the normal approximation method. According to our results in the simulation study and the two real data sets, SPA performs well and achieves high accuracy in approximating the exact p-value instead of the normal method. This article can be applied in different designs, such as random allocation design, Wei’s urn design, complete design and truncated binomial design. Also, the proposed study can be extended to neutrosophic statistics, see Afzal et al.<sup>17</sup>, Albassam et al.<sup>18</sup> and Sherwani et al.<sup>19</sup>.

	Location parameter	Scale parameter
99% Confidence interval		
Exact	(151.4, 308.4)	(29.34, 111.6)
SPA	(151.4, 308.4)	(29.34, 111.6)
Normal	(146.9, 311.9)	(29.23, 115.1)
95% Confidence interval		
Exact	(158.4, 319.4)	(29.51, 122.6)
SPA	(158.4, 319.4)	(29.51, 122.6)
Normal	(158.4, 319.4)	(29.51, 122.6)

**Table 8.** 99% and 95% confidence intervals for the location and scale parameters of the 3rd data set for LP1 test.



## Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 26 April 2023; Accepted: 31 January 2024

Published online: 07 February 2024

## References

- Zhang, Q., Xu, C. Y. & Yang, T. Variability of water resource in the yellow river basin of past 50 years. *Water Resource Manag.* **23**, 1157–1170 (2009).
- Yang, T., Chen, X. & Zhang, Z. C. Spatio-temporal changes of hydrological processes and underlying driving forces in Guizhou region, Southwest China. *Stoch. Env. Res. Risk Assess.* **23**, 1071–1087 (2009).
- Kwoon, M., Jhun, J. G. & Ha, K. J. Decadal change in east Asian summer monsoon circulation in the mid-1990s. *Geophys. Res. Lett.* **34**, L21706 (2007).
- Neuhäuser, M. & Senske, R. The Baumgartner-Weiss-Schindler test for the detection of differentially expressed genes in replicated microarray experiments. *Bionformatics* **20**, 3553–3564 (2004).
- Muccioli, C. *et al.* The diagnosis of intraocular inflammation and cytomegalovirus retinitis in HIV-infected patients by laser flare photometry. *Ocul. Immunol. Inflamm.* **4**(2), 75–81 (1996).
- Rice, K. L. *et al.* Withdrawal of chronic systemic corticosteroids in patients with COPD. *Am. J. Respir. Crit. Care Med.* **162**, 174–178 (2000).
- Lunde, A. & Timmermann, A. Duration dependence in stock prices: An analysis of bull and bear markets. *J. Business Econ. Stat.* **22**, 253–273 (2004).
- Marozzi, M. Nonparametric simultaneous tests for location and scale testing: A comparison of several methods. *Commun. Stat. Simulat. Comput.* **42**(6), 1298–1317 (2013).
- Lepage, Y. A combination of Wilcoxon's and Ansari-Bradley's statistics. *Biometrika* **58**, 213–217 (1971).
- Büning, H. & Thadewald, Th. An adaptive two-sample location-scale test of Lepage-type for symmetric distributions. *J. Stat. Comput. Simul.* **65**, 287–310 (2000).
- Rublik, F. Critical values for testing location-scale hypothesis. *Meas. Sci. Rev.* **9**, 9–15 (2009).
- Rousson, V. On distribution-free tests for the multivariate two sample location scale model. *J. Multivar. Anal.* **80**, 43–57 (2002).
- Smarandache F. (2014). Introduction to neutrosophic statistics. Infinite Study.
- Aslam, M. Retracted article: Neutrosophic statistical test for counts in climatology. *Sci. Rep.* **11**, 17806 (2021).
- Aslam, M. Clinical laboratory medicine measurements correlation analysis under uncertainty. *Ann. Clin. Biochem.* **58**(4), 377–383 (2021).
- Aslam, M. Neutrosophic analysis of variance: Application to university students. *Complex Intell. Syst.* **5**, 403–407 (2019).
- Afzal, U., Alrweili, H., Ahamd, N. & Aslam, M. Neutrosophic statistical analysis of resistance depending on the temperature variance of conducting material. *Sci. Rep.* **11**, 23939 (2021).
- Albassam, M., Khan, N., Aslam, M. (2020). The W/S test for data having neutrosophic numbers: An application to USA village population. *Complexity*. 3690879.
- Sherwani, R.A.K., Shakeel, H., Saleem, M., Awan, W.B., Aslam, M. & Farooq, M. (2021). A new neutrosophic sign test: An application to COVID-19 data. *Plos One*. **16**(8).
- Hollander, M., Wolfe, D. A. & Chicken, E. *Nonparametric Statistical Methods* (Wiley, 2013).
- Rosenberger, W. F. & Lachin, J. M. *Randomization in Clinical Trials: Theory and Practice* (Wiley, 2002).
- Daniels, H. E. Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25**, 631–650 (1954).
- Good, I. J. Saddlepoint methods for the multinomial distribution. *Ann. Math. Stat.* **28**, 861–880 (1957).
- Good, I. J. The multivariate saddlepoint method and chi-squared for the multinomial distribution. *Ann. Math. Stat.* **32**, 535–548 (1957).
- Barndorff-Nielsen, O. E. Edgeworth and saddlepoint approximations with statistical applications. *J. R. Stat. Soc. B* **41**, 279–312 (1979).
- Lugannani, R. & Rice, S. Saddle point approximation for the distribution of the sum of independent random variables. *J. Appl. Probability* **12**, 475–490 (1980).
- Robinson, J. Saddlepoint approximations for permutation tests and confidence intervals. *J. R. Stat. Soc.* **44**, 91–101 (1982).
- Davison, A. C. & Hinkley, D. V. Saddlepoint approximations in resampling methods. *Biometrika* **75**, 417–431 (1988).
- Skovgaard, I. M. Saddlepoint expansions for conditional distributions. *J. Appl. Probability* **24**, 875–887 (1987).
- Booth, J. G. & Butler, R. W. Randomization distributions and saddlepoint approximations in generalized linear models. *Biometrika* **77**, 787–796 (1990).
- Butler, R. W. *Saddlepoint Approximations with Applications* (Cambridge University Press, UK, 2007).
- Daniels, H. E. The approximate distribution of serial correlation coefficients. *Biometrika* **43**, 169–185 (1956).
- Daniels, H. E. Discussion of paper by D. R. Cox. *J. R. Stat. Soc. Biometrika* **20**, 236–238 (1958).
- Abd-Elfattah, E. F. & Butler, R. W. The weighted log-rank class of permutation tests: P-values and confidence intervals using saddlepoint methods. *Biometrika* **94**, 543–551 (2007).
- Abd-Elfattah, E. F. & Butler, R. W. Log-rank permutation tests for trend: Saddlepoint p-values and survival rate confidence intervals. *Can. J. Stat.* **37**, 5–16 (2009).
- Abd-Elfattah, E. F. The weighted log-rank class under truncated binomial design: Saddlepoint p-values and confidence intervals. *Lifetime Data Anal.* **18**, 247–259 (2012).
- Abd-Elfattah, E. F. Saddlepoint p-values and confidence intervals for the class of linear rank tests for censored data under generalized randomized block design. *Comput. Stat.* **30**, 593–604 (2015).
- Abd-Elfattah, E. F. Saddlepoint p-values for two-sample bivariate tests. *J. Stat. Plan. Inference* **171**, 92–98 (2016).
- Abd-Elfattah, E. F. Saddlepoint p-values for group of linear rank tests. *Commun. Stat. Simulat. Comput.* **46**, 4274–4280 (2017).
- Abd El-Raheem, A. M. & Abd-Elfattah, E. F. Weighted log-rank tests for clustered censored data: Saddlepoint p-values and confidence intervals. *Stat. Method Med. Res.* **29**, 2629–2636 (2020).
- Abd El-Raheem, A. M. & Abd-Elfattah, E. F. Log-rank tests for censored clustered data under generalized randomized block design: Saddlepoint approximation. *J. Biopharm. Stat.* **31**, 352–361 (2021).
- Kamal, K.S., Abd El-Raheem, A.M. & Abd-Elfattah, E.F. (2021). Weighted log-rank tests for left-truncated data: Saddlepoint p-values and confidence intervals. *Commun. Stat.-Theory Methods*.
- Kamal, K. S., Abd El-Raheem, A. M. & Abd-Elfattah, E. F. Weighted log-rank tests for left-truncated data under Wei's urn design: Saddlepoint p-values and confidence intervals. *J. Biopharm. Stat.* **32**(5), 641–651 (2022).
- Abd El-Raheem, A. M., Hosny, M. & Abd-Elfattah, E. F. Statistical inference of the class of non-parametric tests for the panel count and current status data from the perspective of the saddlepoint approximation. *J. Math.* **2023**, 1–8 (2023).

45. Abd El-Raheem, A.M., Kamal, K.S., & Abd-Elfattah, E.F. (2023b). P-values and confidence intervals of linear rank tests for left-truncated data under truncated binomial design. <https://doi.org/10.1080/10543406.2023.2171431>
46. Elteren, P. H. On the combination of independent two-sample tests of Wilcoxon. *Bull. Int. Stat. Inst.* **37**, 351–361 (1960).
47. Gastwirth, J. L. Percentile modifications of two sample rank tests. *J. Am. Stat. Assoc.* **60**, 1127–1141 (1965).
48. Duran, B. S., Tsai, W. S. & Lewis, T. O. A class of location-scale nonparametric tests. *Biometrika* **63**, 173–176 (1976).
49. Fueda, K. & Ohori, K. Versatile two-sample rank tests based on Wilcoxon test. *Bull. Inform. Cybernet.* **27**, 159–164 (1995).
50. Agresti, A. A survey of exact inference for contingency tables. *Stat. Sci.* **7**(1), 131–153 (1992).
51. Delanchy, P. R., Heard, N. A. & Lawson, D. J. Meta-analysis of mid-p-values: Some new results based on the convex order. *J. Am. Stat. Assoc.* **114**(527), 1105–1112 (2018).
52. Hand, D. J., Daly, F., Lunn, A. D., McConway, K. J. & Ostrowski, E. *A Handbook of Small Data Sets* (Chapman and Hall, 1994).

### Author contributions

All authors participated equally in all tasks.

### Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to A.E.-R.M.A.-R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024