The comparative analysis of some uniformity tests

P. Yu. Blinov
Department of Theoretical and Applied Informatics
Novosibirsk State Technical University
Novosibirsk, Russian Federation
blindizer@yandex.ru

B.Yu. Lemeshko
Department of Theoretical and Applied Informatics
Novosibirsk State Technical University
Novosibirsk, Russia
Lemeshko@ami.nstu.ru

Abstract — **In the paper some statistical tests intended for testing of uniformity have been considered. Distributions of test statistics, the power of tests under different competing hypotheses have been studied. Considered tests have been ranked by the test power. Advantages and disadvantages of individual tests have been shown. Also, it has been shown that the large part of the tests traditionally used for testing uniformity has the bias under some kind of competing hypotheses. It is underlines that special uniformity tests haven't clear advantage over nonparametric goodness-of-fit tests used for testing uniformity in general.**

Keywords — uniform distribution, hypothesis testing, test statistic, test power

I. Introduction

The uniform distribution is one of common distributions in applied mathematics statistics and probability theory. It is often used to describe the measurement error of some instruments or measuring systems. Simulation of pseudorandom values according to different parametric laws relies on sensors of uniform pseudorandom values. Parametric laws are urgently needed in the systems of statistical simulation. Testing the uniformity actually represents goodness-of-fit testing the hypothesis of uniform distribution of the observed sample $x_1, ..., x_n$. In some papers, the authors states that testing composite hypothesis can be reduced to test simple hypothesis of uniformity on the interval $[0,1]$ because if $x_1, ..., x_n$ belong law with probability distribution function $F(x)$, then random variable $y_i = F(x_i)$ is uniformly distributed on unit interval. All of these factors explain the increasing interest in the choice of simple and computationally efficient procedures for testing hypotheses about the uniform law of analyzed samples.

The various statistical tests used for testing hypothesis of uniformity can be divided into two subsets. These are common goodness-of-fit tests applicable for testing of uniformity and special tests oriented on testing hypothesis that sample $x_1, ..., x_n$ is uniform distributed.

The presence of numerous tests put not simple problem of choosing for specialists, because available information in papers doesn't allows to give preference to certain test, while every specialist is interested not only in correctness of using of tests, but else in reliability of statistical inferences.

In this paper, a lot of considered tests are studied by the method of statistical simulations. The number of experiments carried out for statistical modeling is usually assumed equal to 1 660 000 in the study of the distributions of test statistics. On the one hand, such number of experiments allows tracing the qualitative picture of test statistic distributions in depend on various factors. In the other hand, this number of experiments provides acceptable accuracy of the power estimates and unknown probabilities.

II. The statement of testing uniformity

In the most of uniformity tests, ordered statistics of quantity X are used ($x_{(1)} < x_{(2)} < ... < x_{(n)}$ are elements $x_{(i)}$ of variation series of the sample). Further designation $U_i = x_{(i)}, i = \overline{1,n}$, will be used in expressions of statistical tests.

As usually tests are oriented on testing of simple hypothesis H_0 on interval $[0,1]$. However, if hypothesis of uniformity is tested on interval $[a,b]$ then elements $x_{(i)}$ of variation series $x_{(1)} < x_{(2)} < ... < x_{(n)}$ are modified to corresponding (required in the tests) ordered statistics as: $U_i = \dfrac{x_{(i)} - a}{b - a}$, $i = \overline{1,n}$, $U_0 = 0$, $U_{n+1} = 1$.

To test composite hypothesis of uniformity H_0: $F(x) = \dfrac{x - a}{b - a}$, $x \in [a,b]$, where a and b are non-known, we proceed as follows. Using the variation series $a < x_{(1)} < x_{(2)} < ... < x_{(n)} < b$ of sample $X_1, X_2, ..., X_n$, the parameter estimates are obtained as follows:

$$\hat{a} = x_{(1)} - \frac{x_{(n)} - x_{(1)}}{n-1}, \quad \hat{b} = x_{(n)} + \frac{x_{(n)} - x_{(1)}}{n-1}.$$

It is obviously that testing of composite hypothesis of uniformity for sample $X_1, X_2, ..., X_n$ on interval $[\hat{a}, \hat{b}]$ equal to testing of simple hypothesis of uniformity for sample with sample size $n-2$ on interval $\left[x_{(1)}, x_{(n)}\right]$. The required values of order statistics for testing such hypothesis obtained by expressions: $U_{i-1} = \dfrac{x_{(i)} - x_{(1)}}{x_{(n)} - x_{(1)}}$, $i = \overline{2,(n-1)}$.

A number of considered tests can be divided into three groups. The first group has statistics based on interval between elements, in most of cases differences between neighbor elements denoted as

$$D_i = U_i - U_{i-1},$$

where $U_0 = 0$, $U_{n+1} = 1$, n is the size of the sample. In the second group test statistics used difference between theoretical (expected) and empirical data. These tests also called as tests based on the empirical distribution function (EDF tests), and goodness-of-fit tests are contained in this group. The third group has statistics based on entropy estimator. The third group includes the tests based on the entropy estimator.

III. Alternative hypotheses

We compared the power of tests for relatively sample size *n*=10, 20, 30, 40, 50, 100, 150, 200, 300. The hypothesis under test H_0 was chosen as uniform law. Alternative hypothesis H_i was chosen as beta distribution with the density

$$f(x) = \frac{1}{\theta_2 \mathrm{B}(\theta_0, \theta_1)} \left(\frac{x-\theta_3}{\theta_2} \right)^{\theta_0 - 1} \left(1 - \frac{x-\theta_3}{\theta_2} \right)^{\theta_1 - 1}.$$

where $\mathrm{B}(\theta_0, \theta_1) = \Gamma(\theta_0)\Gamma(\theta_1)/\Gamma(\theta_0+\theta_1)$ is beta-function, $\theta_0, \theta_1 \in (0, \infty)$ are parameters of the form, $\theta_2 \in (0, \infty)$ is shape parameter, $\theta_3 \in (-\infty, \infty)$ is bias parameter, $x \in [0, \theta_2]$. This distribution was chosen because the fact that the standard uniform distribution is a special case of the beta distribution with the parameters of form $\theta_0 = 1$ and $\theta_1 = 1$. We denote the function of beta distribution with values of parameters $B_I(\theta_0, \theta_1, \theta_2, \theta_3)$. So, three alternative hypotheses H_1, H_2, H_3, which are quite close to H_0, can be written by

$$H_1 \; : \; F(x) = B_I(1.5,\, 1.5,\, 1,\, 0), \; x \in [0,1];$$

$$H_2 \; : \; F(x) = B_I(0.8,\, 1,\, 1,\, 0), \; x \in [0,1];$$

$$H_3 \; : \; F(x) = B_I(1.1,\, 0.9,\, 1,\, 0), \; x \in [0,1].$$

The distribution functions and the density functions of these hypotheses are presented in Figure 1 and 2, respectively.

It is worth noting that the distribution function of alternative H_1 crossed the function of the uniform distribution, while the distribution functions of alternatives H_2 and H_3 are located above and below the function of uniform distribution, respectively. And abilities to distinguish hypothesis H_0 from H_1 and from H_2 and H_3 in tests are different. The comparative analysis shows that most of the considered tests have inability to distinguish hypothesis H_0 from H_1 under small sample size n, in other words these tests are biased in such cases.

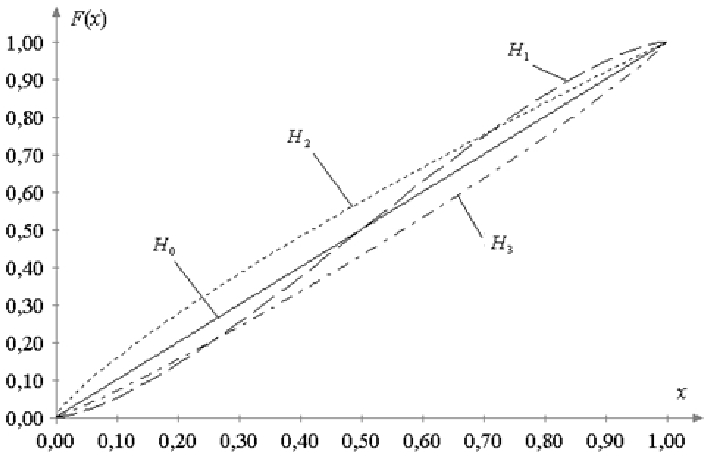

Figure 1. The distribution functions corresponding to the hypotheses

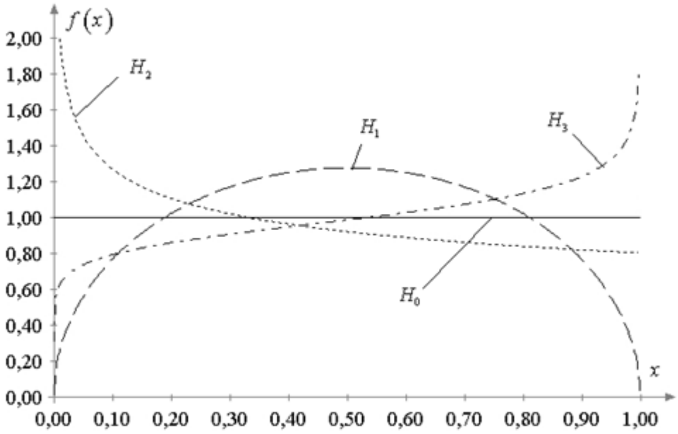

Figure 2. The density functions corresponding to the hypotheses

The distributions $G(\omega_n | H_i)$ of statistic of the Sherman test corresponding to truth of H_0 and H_1 are shown in figure 3 for illustration of fact of bias under sample sizes $n = 10$ and $n = 50$.

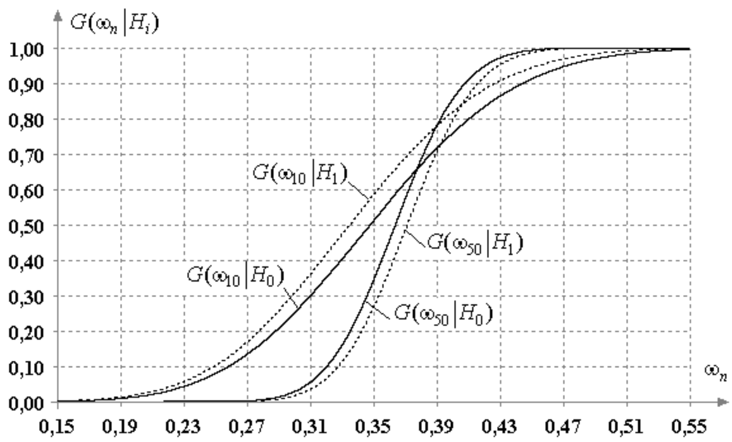

Figure 3. The distributions $G(\omega_n | H_i)$ statistic of the Sherman`s test

The test is right sided, and tested hypothesis is rejected under large values of test statistic. As we can see, the distribution of statistic $G(\omega_{10}|H_1)$ shifted relatively $G(\omega_{10}|H_0)$ not to right, but to left under truth of H_1. Therefore, power $1-\beta$ is less than corresponding α. The bias disappear with increasing of n (see $G(\omega_{50}|H_0)$ and $G(\omega_{50}|H_1)$ in figure 3).

IV. Simulation result

The expressions for statistics of special uniformity tests are presented in Table 1. The description of the most considered tests, their test statistics and power of tests are shown in [1, 2]. In this case, presented in [1, 2] results of the comparative analysis of uniformity tests are supplemented by the research of Correa test [3], modified Anderson-Darling test [4], Hermans-Rasson test [5,6].

The Table 2 contains considered tests ordered by decreasing of power (quantity $1-\beta$) under alternatives H_1, H_2 and H_3 ($n=100$, $\alpha=0.1$). The dark mark means that the test is biased under small sample size n, in other words that quantity α larger than $1-\beta$. This bias take a place to a lesser extent in Neyman-Barton tests N_2 and N_3. This disadvantage isn't observed only for some tests: Kuper test, Watson test, entropy tests, Cheng-Spiring test, Swartz test, second Cressie test, chi-squared Pearson test, Hermans-Rasson test, Pardo test, Correa test and modification of Anderson-Darling test.

Entropy tests used different entropy estimator gives high power under alternative hypothesis H_1. Whereas their power is relatively worst under alternatives H_2 and H_3. It should be noted that only modifications of entropy test have bias under alternative H_2 for small sample size n. It is recognized that power of these tests and also Cressie tests and Pardo test depends from choosing of parameter m called as window size also.

The Neyman-Barton test N_2 shows good power under H_1 and relatively good power under H_2 and H_3. The Hegazy-Green tests and Frosini test demonstrate consistently good ability to distinguish alternative hypotheses from uniformity distribution. The low powers are shown by tests, the statistics of which use the differences of successive values of order sample $U_i - U_{i-1}$ (Sherman test, Kimball test, Moran tests, Greenwood tests, Greenwood-Quesenberry-Miller test). The Cheng-Spiring test, demonstrated quite high power under H_1, shows low power under H_2 and H_3. The lowest power is demonstrated by Yang test under all considered alternative hypotheses. Among the non-parametric goodness-of-fit tests, the good powers are obtained by Zhang tests Z_A and Z_C and Anderson-Darling tests.

Only about 20 tests, concentrated at the top of the columns of Table 2, can be confidently recommended to use in statistical analysis by the reason of the received power estimators presented in this table.

V. Conclusions

Unfortunately, the distributions of most special uniformity tests depend on the sample size, therefore the researchers must rely on the tables of percent points. The similar issue occurs in using nonparametric goodness-of-fit Zhang tests.

The decision-making about results of testing of hypothesis based on p-value is more reasonably than one based on comparison of statistic values with percent points. In [1] the approach permitting to find p-value estimates using statistical methods of modeling, for the tests, the distributions of which are non-known under truth of tested hypothesis, are proposed. This approach is implemented in the interactive mode in the process of the hypothesis testing.

It is found from comparative analysis of tests, which can be used for testing the hypothesis of uniformity, that using of single certain test can be incorrect in forming the reliable statistical inference. The applying more than one test based on different measure of deviation of empirical distribution from theoretical distribution improves the quality of statistical inference. It is better to use some series of tests, which have certain advantages for more objective inferences.

VI. Acknowledgment

This work is supported by the Russian Ministry of Education and Science (project 2.541.2014K).

References

[1] B.Yu. Lemeshko, P.Yu. Blinov. Tests for checking the deviation from uniform distribution law. Guide on the application. – Moscow: INFRA-M, 2015. – 183 p. DOI: 10.12737/11304 (in Russian)

[2] *Blinov P.Yu., Lemeshko B.Yu.* The Comparative Analysis of Tests in the Problem of Testing the Hypothesis of Uniformity // Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Nonparametric Approach" – AMSA'2015, Novosibirsk–Belokuricha, Russia, 14-19 September, 2015. – P. 92-100.

[3] Correa J. C., A new estimator of etropy // Communications in Statistics —Theory and Methods —1995. — V.24. Issue 10. — C. 2436-2446.

[4] Rahman Mezbahur, Larry M. Pearson, Herbert C Heien. A Modifed Anderson-Darling Test for Uniformity// Bulletin of the Malaysian Mathematical Sciences Society. Second Series, Vol 29, P. 11-16.

[5] M. Hermans, and J. P. Rasson (1985), "A new Sobolev test for uniformity on the circle," Biometrika, Vol. 72, No. 3, P. 698-702.

[6] M. Borgan, K. Borgan and A. Futschik (2002) "A data driven smooth for circular uniformity," Annals of the Institute of Statistical Mathematics, Vol. 54, No. 1, P. 29-44.

TABLE 1. STATISTICS OF CONSIDERED TESTS FOR UNIFORMITY

№	Test	Test statitstic	№	Test	Test statitstic	№	Test	Test statitstic
1	Sherman	$\omega_n = \frac{1}{2}\sum_{i=1}^{n+1}\left\|D_i - \frac{1}{n+1}\right\|$	6	Hegazy-Green T_1	$T_1 = \frac{1}{n}\sum_{i=1}^{n}\left\|U_i - \frac{i}{n+1}\right\|$	11	Greenwood	$G = (n+1)\sum_{i=1}^{n+1}(D_i)^2$
2	Kimball	$A = \sum_{i=1}^{n+1}\left(D_i - \frac{1}{n+1}\right)^2$	7	Hegazy-Green T_1^*	$T_1^* = \frac{1}{n}\sum_{i=1}^{n}\left\|U_i - \frac{i-1}{n-1}\right\|$	12	Greenwood-Qesenberry-Miller	$Q = \sum_{i=1}^{n+1}(D_i)^2 + \sum_{i=1}^{n}(D_{i+1} * D_i)$
3	Moran 1	$B = \sum_{i=1}^{n+1}(D_i)^2$	8	Hegazy-Green T_2	$T_2 = \frac{1}{n}\sum_{i=1}^{n}\left(U_i - \frac{i}{n+1}\right)^2$	13	Swartz	$A_n^* = \frac{n}{2}\sum_{i=1}^{n}\left(\frac{U_{i+1} - U_{i-1}}{2} - \frac{1}{n}\right)^2$, where $U_0 = -U_1$, $U_{n+1} = 2 - U_n$
4	Moran 2	$M_n = -\sum_{i=1}^{n+1}\ln\left[(n+1)D_i\right]$	9	Hegazy-Green T_2^*	$T_2^* = \frac{1}{n}\sum_{i=1}^{n}\left(U_i - \frac{i-1}{n-1}\right)^2$	14	Cressie 1	$S_n^{(m)} = \sum_{i=0}^{n+1-m}\left(U_{i+m} - U_i - \frac{m}{n+1}\right)^2$, $m < \frac{n}{2}$
5	Yang	$M = \frac{1}{l}\sum_{i=1}^{n}\min(D_i, D_{i+1})$	10	Frosini	$B_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\|U_i - \frac{i-0.5}{n}\right\|$	15	Cressie 2	$L_n^{(m)} = -\sum_{i=0}^{n+1-m}\ln\left[\frac{n+1}{m}(U_{i+m} - U_i)\right]$, $m < \frac{n}{2}$

№	Test	Test statitstic	№	Test	Test statitstic
16	Cheng-Spiring	$W_p = \left[(U_n - U_1)\frac{n+1}{n-1}\right]^2 \Big/ \sum_{i=1}^{n}(U_i - \bar{U})^2$	17	Pardo	$E_{m,n} = \frac{1}{n}\sum_{i=1}^{n}\frac{2m}{n(U_{i+m} - U_{i-m})}$

№	Test	Test statitstic
18	Neyman-Barton N_k; $k = 2,3,4$	$N_2 = \sum_{j=1}^{2}V_j^2$, where $V_j = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\pi_j(U_i - 0.5)$, $\pi_1(y) = 2\sqrt{3}y$; $\pi_2(y) = \sqrt{5}(6y^2 - 0,5)$; $\pi_3(y) = \sqrt{7}(20y^3 - 3y)$; $\pi_4(y) = 3(70y^4 - 15y^2 + 0,375)$
19	Dudewicz-Van Der Mulen	$H(m,n) = -\frac{1}{n}\sum_{i=1}^{n}\ln\left\{\frac{n}{2m}(U_{i+m} - U_{i-m})\right\}$, where $m -$ integer and $m \le \frac{n}{2}$; if $i+m \ge n$, then $U_{i+m} = U_n$, if $i-m \le 1$, then $U_{i-m} = U_1$
20	The first modification of entropy test	$HY_1 = -\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{U_{i+m} - U_{i-m}}{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}\right)$, where $\hat{F}(U_i) = \frac{n-1}{n(n+1)}\left(i + \frac{1}{n-1} + \frac{U_i - U_{i-1}}{U_{i+1} - U_{i-1}}\right)$, $i = \overline{2,(n-1)}$, $\hat{F}(U_1) = 1 - \hat{F}(U_n) = \frac{1}{(n+1)}$
21	The second modification of entropy test	$HY_2 = -\sum_{i=1}^{n}\ln\left(\frac{U_{i+m} - U_{i-m}}{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}\right) \times \left(\frac{\hat{F}(U_{i+m}) - \hat{F}(U_{i-m})}{\sum_{j=1}^{n}\left(\hat{F}(U_{j+m}) - \hat{F}(U_{j-m})\right)}\right)$

Continue Table 1

№	Test	Test statitstic
22	Correa	$C(m,n)=\frac{1}{n}\sum_{i=1}^{n}\ln(b_i)$, where $b_i=\frac{\sum_{j=i-m}^{i+m}\left(X_{(j)}-\bar{X}_{(i)}\right)(j/n-i/n)}{\sum_{j=i-m}^{i+m}\left(X_{(j)}-\bar{X}_{(i)}\right)^2}$; $\bar{X}_{(i)}=\sum_{j=i-m}^{i+m}\frac{X_{(j)}}{2m+1}$.
23	Hermans-Rasson	$T_{n,\infty}=\frac{n}{\pi}-\frac{1}{2n}\sum_{i,j=1}^{n}\left\|\sin\left(U_i-U_j\right)\right\|$
24	The modification of Anderson-Darling	$V^2=n\left[\frac{(U_1-c_1)^2}{(1-c_1)}+\sum_{i=1}^{n}\frac{(U_i-c_i)^2}{c_i(1-c_i)}(c_{i+1}-c_i)+\frac{(U_n-c_n)^2}{c_n}\right]$, where $c_i=(i-0.375)/(n+0.25)$

TABLE 2. THE TESTS RANKED BY POWER (FOR $\alpha=0.1; n=100$)

№	hypothesis H_1	$1-\beta$	hypothesis H_2	$1-\beta$	hypothesis H_3	$1-\beta$
1	Hermans-Rasson	0.903	Anderson-Darling	0.648	Anderson-Darling	0.526
2	The second modification of entropy test	0.883	Hegazy-Green T_1	0.610	Hegazy-Green T_1	0.522
3	Zhang Z_A	0.850	Zhang Z_C	0.606	Frosini	0.522
4	Neyman-Barton N_2	0.837	Frosini	0.603	Hegazy-Green T_1^*	0.520
5	Cressie 2	0.820	Hegazy-Green T_2	0.602	The modification of Anderson-Darling	0.519
6	Zhang Z_C	0.819	Neyman-Barton N_2	0.597	Hegazy-Green T_2	0.508
7	Dudewicz-Van Der Mulen	0.790	Kramer-von-Misses-Smirnov	0.595	Kramer-von-Misses-Smirnov	0.507
8	The first modification of entropy test	0.789	Hegazy-Green T_1^*	0.595	Hegazy-Green T_2^*	0.506
9	Correa	0.782	Zhang Z_K	0.590	Zhang Z_C	0.463
10	Watson	0.779	The modification of Anderson-Darling	0.585	Zhang Z_A	0.459
11	Neyman-Barton N_3	0.766	Hegazy-Green T_2^*	0.585	Kolmogorov	0.450
12	Neyman-Barton N_4	0.739	Neyman-Barton N_3	0.577	Neyman-Barton N_2	0.447
13	Kuper	0.732	Zhang Z_A	0.574	Zhang Z_K	0.438
14	The modification of Anderson-Darling	0.730	Neyman-Barton N_4	0.557	Neyman-Barton N_3	0.416
15	Cheng-Spiring	0.722	Kolmogorov	0.542	Neyman-Barton N_4	0.381
16	Zhang Z_K	0.617	Pardo	0.463	χ^2 Pearson	0.374
17	χ^2 Pearson	0.593	χ^2 Pearson	0.448	Pardo	0.291
18	Swartz	0.583	Kuper	0.364	Dudewicz-Van Der Mulen	0.275
19	Anderson-Darling	0.505	Watson	0.356	The first modification of entropy test	0.275
20	Hegazy-Green T_1^*	0.443	The first modification of entropy test	0.328	The second modification of entropy test	0.267
21	Hegazy-Green T_2^*	0.409	Dudewicz-Van Der Mulen	0.327	Corea	0.267
22	Pardo	0.408	Cressie 1	0.314	Watson	0.257
23	Frosini	0.384	Correa	0.313	Kuper	0.254
24	Kramer-von-Misses-Smirnov	0.358	The second modification of entropy test	0.266	Cressie 2	0.226
25	Hegazy-Green T_1	0.322	Greenwood-Qesenberry-Miller	0.244	Cressie 1	0.218
26	Kolmogorov	0.322	Swartz	0.226	Swartz	0.206
27	Hegazy-Green T_2	0.308	Cressie 2	0.217	Greenwood-Qesenberry-Miller	0.186
28	Greenwood-Qesenberry-Miller	0.290	Sherman	0.204	Kimball	0.165
29	Kimball	0.279	Kimball	0.201	Moran 1	0.165
30	Moran 1	0.279	Moran 1	0.201	Greenwood	0.165
31	Greenwood	0.279	Greenwood	0.201	Sherman	0.154
32	Sherman	0.215	Moran 2	0.193	Moran 2	0.143
33	Cressie 1	0.187	Hermans-Rasson	0.169	Hermans-Rasson	0.110
34	Moran 2	0.187	Cheng-Spiring	0.168	Cheng-Spiring	0.106
35	Yang	0.115	Yang	0.108	Yang	0.104