Title: How to compare separation selectivity of high-performance liquid chromatographic columns properly?

Authors: Filip Andrić, Károly Héberger

PII: S0021-9673(17)30151-6
DOI: http://dx.doi.org/doi:10.1016/j.chroma.2017.01.066
Reference: CHROMA 358245

To appear in: *Journal of Chromatography A*

Received date: 25-11-2016
Revised date: 23-1-2017
Accepted date: 24-1-2017

Please cite this article as: Filip Andrić, Károly Héberger, How to compare separation selectivity of high-performance liquid chromatographic columns properly?, Journal of Chromatography A http://dx.doi.org/10.1016/j.chroma.2017.01.066

Highlights

- Known pattern was disclosed by independent techniques, bunch of techniques provides same pattern
- Similarity of columns are revealed by non-parametric methods
- Sum of ranking difference and generalized pair correlation methods for fine differences
- The primary retention data are superior compared to the HS approach
- Enforcing HS model leads to significant information loss

How to compare separation selectivity of high-performance liquid chromatographic columns properly?

Filip Andrić[1], Károly Héberger[2,*]

[1] Faculty of Chemistry, University of Belgrade, Studentski trg 12-16, 11000 Belgrade, Serbia

[2] Plasma Chemistry Research Group, Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, H-1117 Budapest XI., Magyar Tudósok krt. 2, Hungary

*Corresponding author: Károly Héberger

Phone: +36 1 38 26 509

E-mail: heberger.karoly@ttk.mta.hu

Abstract

Comparison and selection of chromatographic columns is an important part of development as well as validation of analytical methods. Presently there is abundant number of methods for selection of the most similar and orthogonal columns, based on the application of limited number of test compounds as well as quantitative structure retention relationship models (QSRR), from among Snyder's hydrophobic-subtraction model (HSM) have been most extensively used.

Chromatographic data of 67 compounds were evaluated using principal component analysis (PCA), hierarchical cluster analysis (HCA), non-parametric ranking methods as sum of ranking differences (SRD) and generalized pairwise correlation method (GPCM), both applied as a consensus driven comparison, and complemented by the comparison with one variable at a time (COVAT) approach. The aim was to compare the ability of the HSM approach and the approach based on primary retention data of test solutes (log*k* values) to differentiate among ten highly similar C18 columns.

The ranking (clustering) pattern of chromatographic columns based on primary retention data and HSM parameters gave different results in all instances. Patterns based on retention coefficients were in accordance with expectations based on columns' physicochemical parameters, while HSM parameters provided a different clustering.

Similarity indices calculated from the following dissimilarity measures: SRD, GPCM Fisher's conditional exact probability weighted (CEPW) scores; Euclidian, Manhattan, Chebyshev, and cosine distances Pearson's, Spearman's, and Kendall's, correlation coefficients have been ranked by the consensus based SRD. Analysis of variance confirmed that the HSM model produced statistically significant increases of SRD values for the majority of similarity indices, *i.e*. HS transformation of original retention data yields

significant loss of information, and finally results in lower performance of HSM methodology. The best similarity measures were obtained using primary retention data, and derived from Kendal's and Spearman's correlation coefficients, as well as GPCM and SRD score values. Selectivity function, *F*s, originally proposed by Snyder, demonstrated moderate performance.

Keywords: Chromatographic column selection, Distance and orthogonality measures, Hydrophobic subtraction model, Sum of Ranking Differences, Generalized Pairwise Correlation method, Principal Component Analysis

1. Introduction

Presently there is abundant numbers of various brands of reversed-phase stationary phases available on the market, and new ones are released regularly. Since separation performance of a chromatographic column strongly depends on the properties of stationary phase, it is important to have reliable means for their comparison and adequate selection.

Comparison and selection of appropriate columns is an important part of development and validation of many chromatographic methods. During early phases of method development, a search for the most dissimilar (orthogonal) chromatographic systems is targeted in order to achieve the optimal separation of analytes and their impurities [1]. On the other hand, selection of the most similar columns is of great importance in assessing robustness and reproducibility of a method, and plays important role in method transferring processes. Information on chromatographic column similarity is of great deal in every-day laboratory practice providing valuable information to a researcher in order to choose a suitable substitute [2].

At present, numerous approaches exist to compare and select similar and dissimilar (orthogonal) chromatographic columns (systems) [3]. Some comparison methods use only a few test solutes (six or seven) in order to determine stationary phase characteristics such as: hydrophobicity, hydrophobic selectivity, shape selectivity, hydrogen bonding, ion exchange *etc*. Engelhardt's test [4,5], the extensively used Tanaka's test [6], its modifications [7-9], and the Catholic University of Leuven (Katholieke Universiteit Leuven - KUL) method [10-13] are good examples. The other approaches include establishing quantitative structure retention relationships (QSRR) models describing the chromatographic column selectivity in terms of molecular descriptors [14,15] or physicochemical properties such as linear solvation energy relationships (LSER) [16] or HSM method [17-22].

All column classification systems suffer from the problem that the inherent clustering pattern among the columns is not known.

Although, aiming to measure more or less the same column properties, some of the methods usually provide significantly different results (*e.g.* KUL method, original Tanaka's test and its modifications) [22,23]. The situation is more complicated by use of different chemometric approaches to perform comparisons. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) are most frequently employed [1,8,9]. The main drawback of unsupervised pattern recognition methods is that they provide different column grouping/ranking patterns depending on the applied amalgamation rule (*e.g.*, weighted and unweighted-average linkage, single and complete linkage, the centroid, Ward's method, *etc.*), and distance/similarity measure used [24-26]. Various dissimilarity measures such as selectivity function, *F*s [11-13,23] and orthogonality ratio (OR) based on Pearson's correlation coefficient, Spearman's rho, and generalized pairwise correlation method (GPCM) based on Fisher's conditional exact test as well as McNemar's significance test [25] are used to find the most similar or the most orthogonal systems. All of them lead to different results. Column selectivity function, *F*s, is an Euclidian distance of a target column (1) to the reference one (2) (Eq. 1). Hence it is represented by a single number, and is originally used in the framework of the hydrophobic-subtraction model.

$$F_s = [a(A_1 - A_2)^2 + b(B_1 - B_2)^2 + c(C_1 - C_2)^2 + h(H_1 - H_2)^2 + s(S_1 - S_2)^2]^{1/2} \qquad (1)$$

Where *a*, *b*, *c*, *h* and *s* are weight factors that depends on the nature and structure of the set of test compounds.

However, the same form of equation is often applied by the KUL, Euerby, Tanaka, and similar testing assays [11-13,23]. Column comparison is also done based on primary retention data (log*k* values) of testing solutes, and already mentioned techniques of multivariate data analysis [28].

The commonly used approach is to disclose the correlation matrix showing the similarities of columns pairwise. These matrices can be colored producing attractive heatmaps. As many similarity measures exist, many such heatmaps can be produced. They are rarely used for column selection as the similarity pattern depends on the definition of similarity and the color coding strongly. In this work we invent a solution to the problem of arbitrarily selected reference columns (COVAT method, see later). At the same time we apply a fair comparison method for various similarity search algorithms and for the "best" similarity metrics suitable to define the inherent clustering of chromatographic columns.

Finally, the present study aims to compare the ability of the HSM approach and the approach based on primary retention data to rank, group, and select the most similar and dissimilar (orthogonal) chromatographic columns, especially in the case of columns of highly similar properties (performance). For such purposes, we have decided to reinvestigate the emblematic data of Wilson's *et al*. [17], which is historically a starting point in developing of the HSM approach. The method evolved during years and some changes in the values of the HSM column parameters occurred in latter works [19,20]. Also, Shackman [29] recently alerted to some minor discrepancies among HSM values reported for the same columns in above mentioned sources [19,20]. Despite of these discrepancies we were motivated to use this particular data set because its intrinsic data structure is well described and supported by physicochemical properties of the columns. Therefore, it suits ideally for a method comparison study. Several chemometric approaches were used such as PCA, HCA, and novel non-parametric ranking methods - called sum of ranking differences (SRD) [30-32] and the GPCM [33,34].

2. Materials and methods (calculations)

2.1. Data selection

Wilson and co-authors [17] have studied retention behavior of 67 solutes of highly diverse molecular structure on ten very similar octadecyl silica modified chromatographic columns, under the same chromatographic conditions. They have used retention data to develop a five term equation (Eq. 1) connecting the retention of solutes and column properties, by the so called hydrophobic-subtraction model. The model allows the characterization of the column selectivity according to hydrophobicity (*H*), steric resistance to penetration into the stationary phase (*S*), hydrogen bonding of basic or acidic solutes, respectively, acidic or basic column sites (hydrogen-bond acidity *A* and basicity *B*) and ion interaction or ion-exchange capacity (*C*).

The source provides the column properties, primary retention data (log*k* values), and derived column selectivity parameters that are included in the scope of the present work as **Table 1,** and the **Tables S1** and **S2** (Supplementary material).

Table 1

The retention data are collected on a single mobile phase composition (50% acetonitrile - phosphate buffer) with a pH = 2.8, at a temperature of 35 °C. Although Wilson *et al*. [18] suggested that measured column selectivity under one set of conditions can be related to different conditions, and thus likely relative values remain constant, it is worth mentioning that any variation in the experimental conditions would impact the retention data, and thus may change the findings presented in the current work.

2.2. Data pre-processing and statistical analysis

As primary retention data are basically expressed on the same scale, considering that very similar columns have been studied under the same chromatographic conditions, basically no data pre-processing is necessary. Nevertheless, in the case of PCA and HCA, mean

centering and scaling to the unit standard deviation (standardization) was applied, as in the case of column selectivity parameters, which are expressed in different scales.

In the case of comparison of various dissimilarity measures (see section 3.3) different data pre-processing were investigated: (*i*) standardization, (*ii*) range scaling between 0 and 1, and (*iii*) rank transformation. Previous studies [35-36] suggest that variance analysis is able to reveal, whether the data preprocessing is a significant factor or not. If not, one can be sure that no artifacts are incorporated in the analysis by data pre-processing [35]. Otherwise, the best preprocessing method should be recommended [34].

All data pre-processing, descriptive statistics, PCA, HCA, and analysis of variance (ANOVA) were performed using Statistica v. 10 (Statsoft Inc. Tulsa, Oklahoma, USA).

In the case of HCA and PCA the PLS, PCA and multivariate/Batch SPC module was used, while analysis of variance was done by Factorial ANOVA tool, part of advanced models (General linear) module, Statistica v. 10.

HCA has been carried out using Ward's amalgamation rule and Euclidian distance as a distance measure. The rest of computations were performed using Microsoft Excel 2010.

2.3. Sum of ranking differences

Sum of ranking differences has been developed with a purpose to fairly rank, group and compare methods and models [30]. It has been applied to various problems so far: from checking the multi-class classification performance in the case of tobacco leaf grades [37] *via* ranking and classifying chromatographic systems [38,39], to selection of multiple tuning parameters for multivariate calibration [40].

The SRD approach is simple: It requires the input data matrix consisting of objects arranged in rows (in the present case 67 compounds, or column selectivity parameters in case of HSM data) and methods (in this case chromatographic columns, or dissimilarity measures)

arranged in columns. In order to rank, compare, and group methods, a reference must be provided. Depending on the nature of the data, a reference can be the row maximum, row minimum, row arithmetic mean (average) values, or an already known set of reference values. Rationales for the use of row maxima and minima are valid in the case of properties that are maximized or minimized under optimal conditions, such as correlation coefficients and method errors (residuals), respectively. However, in most situations the use of the average is suitable (consensus modeling). Two main points justify this: (*i*) every method is accompanied by some sort of bias, as well as random errors; hence, the usage of arithmetic means leads to errors to cancel each other (at least) to a certain extent, and (*ii*) according to maximum likelihood principle the average is the most probable value to apply instead of all studied methods individually.

After setting up the input matrix, the objects are ranked for each method including the reference, and these ranks are subtracted from the reference ranking. Absolute values of differences are then summed up and every method is associated with an SRD value. The lower the SRD value, the closer is the method to the benchmark, and *vice versa*. In that way the ranking of methods (in this case chromatographic columns, or dissimilarity measures) is obtained. The SRD values are usually expressed as the range scaled values between 0 and 100 – so called normalized SRDs.

The SRD procedure is validated in two ways. The first one, called comparison of ranks by random numbers (CRRN), either uses simulated random numbers or theoretical distribution of the random SRD values. SRD values that originate significantly differently from random distribution fall away from each side of the theoretical or fitted Gaussian-like curve at the significance level $p = 0.05$. The second validation uses sevenfold ($n > 13$) or leave-one-out cross-validation ($n < 14$) procedures to create seven, or n datasets by removing 1/7th of objects, or just 1, in each step, where n is the total number of objects. Truncated

datasets are further subjected to SRD procedure, which finally results in 7 or *n* SRD values for each variable (method), enabling the estimation of their uncertainty. Statistical difference among variables (methods) can be then tested by applying Wilcoxon's matched pair test, as well as sign test. Graphical representation of ranking and grouping may preferably be completed providing box and whisker plots.

Validation approaches are complementary. The first one provides information whether the ranking is statistically different from the random distribution of SRDs, while the second one additionally provides statistically significant difference among studied methods (variables). SRD procedure was done using several Microsoft Excel macros that can be downloaded from http://aki.ttk.mta.hu/srd/ together with input and output files.

2.4. Generalized pairwise correlation method

The method requires the same data input matrix as in the case of the SRD; however, the GPCM approach is based on completely different train of thoughts. First, two variables ($X1$ and $X2$) are selected and compared with the reference for each object pair. There are four possible outcomes: A) both $\Delta X1$ and $\Delta X2$ are positive, B) one of them is positive the other is negative, C) *vice versa*, D) both are negative. The numbers of events are counted for all possible pairs of objects. Then a suitable statistical test (*e.g.* Fisher's conditional exact, McNemar's, Chi-square and William's test) decides whether the numbers of *B* and C events are significantly different or not: *i.e.* $X1$ is superior (wins over) to $X2$, reversely (looses against $X2$), or no decision can be made (tie). In the present work, only conditional exact Fisher's test (CE) was used. Then, variables are compared pairwise with the reference, considering all possible combinations. Variables can be further ranked according to the number of wins (simple ranking), number of wins minus number of losses, or probability

weighted ranking (PW), *i.e*. based on *p*(wins)-*p*(losses) scores. The last one was used in the present work.

3. Results and discussion

3.1. Exploratory data analysis

Visual inspection of the radar plots (**Figure 1**) suggests that all studied compounds except 16, 18, 45-50, 63-65, 67, share extremely similar retention profiles (the similar shape of different size). Compounds 46-50 are strong basic solutes, while solutes 56-67 are weak acids (**Table S2**, Supplementary material). This implies that column comparison based on such a vast number of compounds should definitely contain redundant information, and that significantly smaller set of compounds could be equally suitable. Indeed, the creators of the HSM approach ultimately reduced the size of the test set to 16 solutes [19], which are still in use for determination of HSM parameters for new columns. Use of 67 solutes for testing column selectivity requires significant recourses and is simply impractical. However, in this case Wilson *et al*. [17] have done a good job in selecting ten columns which have indeed very similar behavior, as demonstrated by the retention of a large portion of the test set (n = 67).

Figure 1

Standardized primary retention data were subjected to PCA and HCA to accomplish better insight into data structure. Standardization (mean centering and scaling to the unit standard deviation) has been used as a data preprocessing step for both data sets: for log$k$ data and HSM parameters. A virtual (hypothetic) average reference column (VARC) was created based on average retention coefficients of all 67 substances on all 10 columns. Selectivity parameters (H, S, A, B, and C) of the VARC were obtained as regression coefficients of the multiple regression model using average log$k$ as dependent and selectivity parameters of solutes as independent variables (**Tables S1** and **S2**, Supplementary material).

Figure 2

The total variance (98.06%) of the overall primary retention data was described using the first two principal components (91.02% and 7.04% by PC1 and PC2 respectively). Only SB-300 and Discovery can be distinguished by PC1. However, retention data variability along PC2 reveals fine differences among the rest of the chromatographic columns. As it could be expected, based on the column property profiles (**Table 1**), SB-100 and SB-90 exhibit the greatest similarity. Close to them are all three YMC columns and the Eclipse. All of these are very close to the VARC, while Symmetry and Inertsil diverge. An outlier stationary phase (SB 300) was also observed and validated (see **Table 1**). This is also in accordance with expectations based on already known primary retention data structure. The loading plot reveals that compounds 45-50 diverge significantly from the main cluster. This is reasonable since all of them exhibit strong basic properties and are therefore retained by mechanisms, which most likely involve extensive ionic interactions with negatively charged silanols.

However, PCA based on HSM parameters leads to different disposition of chromatographic columns (**Figure 2c**). Considering the first two principal components, which account for 84.53% of the overall variance, chromatographic columns YMC 15, YMC 16, YMC 17 and Eclipse remain the closest to the VARC, but Inertsil and Symmetry, which were distant from the VARC in the case of PCA based on primary retention data, suddenly become closer. Also SB-100 and SB-90 that have been close to the VARC in the previous case and according to physicochemical parameters are very distant now.

Hierarchical cluster analysis complements the findings of PCA (**Figures S1a** and **S1b**, Supplementary material).

If one would like to substitute the VARC, the best choice would be YMC 16, and YMC 17, followed by Eclipse and YMC 15. Both, PCA and HCA may lead to the same conclusion.

3.2. Comparison of chromatographic columns by means of SRD and GPCM

Although PCA and HCA have been frequently applied to find similarities among particular chromatographic conditions [1,8,9], or specific chromatographic columns, the non-supervised character renders an ambiguity to the pattern. The main disadvantage of PCA is a loss of information with consecutive reduction of dimensionality by projecting data only to few principal components. Besides, both PCA and HCA lack in estimation of statistical significance of such comparisons.

Non-parametric ranking methods such as SRD and GPCM, have the ability to fairly rank, group and compare methods (in this particular case chromatographic columns). They are also able to find differences when other parametric tests fail. SRD and GPCM are able to fuse multiple criteria (parameters) in order to achieve the final ranking.

In the present work both, retention and HSM data have been subjected to the SRD and GPCM without any preprocessing, using VARC as a benchmark. Validation of the SRD procedure was completed by comparison with the distribution of random numbers (CRRN).

Figure 3

Both, SRD and GPCM ranking methods based on primary retention data (**Figures 3a, and 3c**) select the YMC 16 and YMC 17 as the closest to the VARC, followed by SB-100, SB-90, Inertsil and Discovery. SB-300 is distinctively separated from the rest of them (an outlier). This is again the pattern that is expected based on the well-known column property profiles (**Table 1**). All chromatographic columns are positioned far from the random number distribution curve in a narrow window of the SRD values (0.98-5.17), which indicates that all of them are powerful in ranking test solutes, mostly according to their abilities to establish hydrophobic, as well as specific interactions with the stationary phase, and not by chance. Although, the SRD and GPCM share the main ranking pattern, there are several

chromatographic columns that are separated by SRD, which are "glued together" by GPCM comparison (sharing almost the same GPCM score values). Obviously, the SRD is more suitable, when refined discrimination among similar columns is needed.

Contrary to the ranking based on primary retention data, comparisons based on HSM parameters reveals completely different patterns. Both, SRD and GPCM select Inertsil and YMC1 5 as the closest to the VARC. However, GPCM ranks the SB-300 as the worst one, *i.e.*, the farthest from the reference, and clearly separated from the rest of them, while SRD puts all three SB columns together at the very end. Also, with the exception of Inertsil and YMC1 5, the SRD puts all chromatographic columns under the 95% probability interval of the random SRD distribution curve. In this particular case this is simply a consequence of a limited number of input terms (*H*, *S*, *A*, *B* and *C*). Again, patterns provided by GPCM and SRD are slightly different; however, the inherent order is preserved (**Table S3**, Supplementary material). The ranking pattern among columns suggested by both comparison methods using HSM data is different from expectations based on column properties. This might suggest that applying the HSM methodology to retention data may result in partial distortion of column grouping (based on selectivity) information.

Figure 4

After completing the sevenfold cross-validation procedure on retention data, as described in the section 2.3, the results are depicted in a form of box and whisker plot (**Figure 4**). All three YMC columns cannot be distinguished among themselves at the predefined significance level of $p = 0.05$. The same is valid for SB-90 and SB-100. However, the group of Inertsil, Symmetry and Discovery is clearly separated from the previous ones. The same is valid for SB-300, which differs from all others. In this way identifying similar and orthogonal columns is statistically justified. This is of great importance for chromatographers since the SRD procedure can be easily implemented for "in house" column comparisons or it can be

embedded in large databases. Compared to the HSM for which the cut off condition (Fs < 3) is empirically established, the SRD approach is founded on a firm statistically basis.

Although the use of the arithmetic mean as a reference seems to be an ideal solution, sometimes the choice of reference is not unambiguous. In such cases an alternative methodology might be used – the application of pairwise variable comparisons. Correlation or distance based matrices are most often used [24-26]. However, sometimes such matrices, especially in the form of heatmaps are used merely as a convenient representation of HCA results [24-26]. We have extended the capabilities of SRD and GPCM by using each variable (column) as a reference once and only once [41]. As a final result square matrices are obtained, which are superior in finding orthogonal and similar objects, over the classical approaches based on Pearson or Spearman rank correlation coefficients [41]. We have named this approach as comparison with one variable at a time (COVAT). SRD-COVAT matrix is completely symmetrical, while GPCM produces asymmetric one by definition. This is expected since GPCM probability weighted scores differ depending on weather x or y is used as a reference. Finally, the SRD-COVAT matrix is rearranged according to ascending order of the row-wise averaged SRD scores (which is, at the same time, the ascending order of the column-wise average SRDs, as the matrix is symmetric unless ties (equal numbers) are present in the input matrix). In the case of GPCM matrix, due to the inherent asymmetry, the rearrangement is slightly different. The reference variables are arranged in columns in an ascending order of the row-wise averages score. However, column-wise averaging leads to different results; therefore, the arrangement of GPCM-COVAT matrices demands a compromise. Results for both matrices are presented in an easily perceivable way using a heatmap, with three coloring schemes: relative, absolute and Gaussian. In the present work we have decided to stick with the relative coloring, which divide the range of score values into ten sub-ranges of the same size (0-10%, 10-20%,…,90-100%) and a color is assigned to each

of them (**Figure 5**). In the case of SRD-COVAT matrix the most similar elements are easily found in the top corner of a heatmap, and along the diagonal, sharing the same or similar color (**Figures 5a** and **5b**). GPCM-COVAT matrix requires different interpretation (**Figures 5c** and **5d**). In the case of SRD we have implemented a MS Excel VBA macro for the generation of "SRD heatmaps", which is also available for download at: http://aki.ttk.mta.hu/srd/

Figure 5

Both SRD and GPCM COVAT relative heatmaps based on primary retention data (**Figures 5a** and **5c** respectively) share a common pattern. Again, columns YMC 16, YMC 17 and YMC 15 are identified as the most similar ones, and they are best correlated with the rest of the studied columns. They are clustered in the upper left square A, closely bound with Eclipse (square B). Both chromatographic columns SB-90 and SB-100 are joined together along the matrix diagonal (square C). The most dissimilar column, SB-300, is located at the very edges of the matrix (area denoted as D). The rest of them, Inertsil, Discovery and Symmetry are placed in between. This is, again, confirmation of the expected pattern based on column properties, and above described results. Based on COVAT matrices it is easy to select the most similar as well as the most dissimilar pairs of chromatographic columns. For example SB-90 can be easily replaced by SB-100. However, in the absence of SB-100, a much better choice can be found from the columns located in the upper part of the matrix, *e.g.*, Eclipse, YMC 16, or YMC 17, than those located below (Inertsil, Discovery or Symmetry).

SRD and GPCM COVAT matrices based on HSM parameters (**Figures 5b** and **5d**) demonstrate a considerably different pattern. Only the pair made of SB-90 and SB-100 columns is preserved. Also, YMC 16 and YMC 17 together with Eclipse can be found in the upper left area of both matrices. However, the upper position of Discovery, the middle

position of SB-300 column, and the marginal position of YMC 15 clearly indicate strong deviation from the expected similarity pattern among the columns, which is inherent in the column physicochemical property data as well as primary retention data. The complete list of COVAT scores is provided in the Supplementary material (**Tables S4a** - **S4d**).

3.3. Miscellaneous (dis)similarity measures in column comparison

Besides the various chemometric approaches discussed above, a simple (dis)similarity measure between two chromatographic columns is often used. Most (dis)similarity measures evaluated in this work (**Table 3**) are well known, and some of them have been already used in column comparison studies [11-13,17,23-26]. Nevertheless, a few points should be emphasized. First of all, our primary goal was to determine which type of data is the most suitable for column clustering. As different similarity measures are variously sensitive to deviations from the normal distribution and the presence of outliers, we have to use (more) robust similarity coefficients. However, we had to include absolute Pearson correlation distance, cosine and Manhattan distance because of their frequent usage. Second of all, any dissimilarity (distance) d(A,B), and similarity measure S(A,B) among two objects A and B can be interconverted by several mathematical transformations [42]. In the case of the so-called "unbound" dissimilarity measures, which lie in the range $(0, +\infty)$, transformation was done according to the Eq. 2. In the case of bound dissimilarities, ranging between 0 and 1, S(A,B) was calculated according to the Eq. 3.

$$S(A,B) = \frac{1}{1+d(A,B)} \tag{2}$$

$$S(A,B) = 1-d(A,B) \tag{3}$$

Short definitions of investigated dissimilarity measures and their corresponding similarities are given in **Table 2**. The values of distances of chromatographic columns from the VARC, based on primary retention data, as well as HSM parameters are collected in the

Table S5a, while the corresponding similarities are given in the **Table S5b** (Supplementary material).

Table 2

Figure 6

The pattern among chromatographic columns, such as possible groupings, as well as the most similar and the most distant ones from the reference, can be easily identified by fusion of similarity or dissimilarity measures. In that sense, stack plots of interval scaled (between 0 and 1) dissimilarity and similarity measures are particularly useful, providing cumulative profiles, which are sensitive to small differences among the columns. Profiles based on primary retention data (**Figures 6a** and **6c**) demonstrate high similarity of all three YMC columns, which are the closest to the VARC (the lowest cumulative distance values, and the highest values of corresponding cumulative similarities). The Eclipse follows closely the pattern, along with the pair of SB-100 and SB-90, while the last ones are Inertsil, Symmetry and Discovery. However, the sharp peak of cumulative distance and the corresponding deep valley of cumulative similarity come from SB-300. Such behavior is in accordance with the already described physical properties of columns. On the contrary, the stack plots based on HSM parameters (**Figures 6b** and **6d**) provide a completely different picture. The cumulative dissimilarity profile hardly distinguishes between all three SB columns, and the YMC 15 can be easily differentiated from the YMC 16 and YMC 17 columns. The similarity (dissimilarity) measures behave differently for Symmetry column – zigzag pattern for HSM data (**Figures 6b** and **6d**) whereas a more uniform behavior is seen in **Figures 6a** and **6c**.

Obviously, the primary retention data provide different pattern between chromatographic columns compared to Snyder's HSM selectivity parameters. In order to investigate this particular phenomenon we have decided to rank and group similarity measures by SRD in a consensus based comparison. The arithmetic mean was used as the reference. A comparison

of similarities is more convenient than comparison of corresponding dissimilarities simply because they range between 0 and 1, with 1 corresponding to identical objects and 0 to the most dissimilar ones. It should be kept in mind that different similarity measures might be on different scale, although they are located in the same range [35], especially if they are derived from similar or complementary distance metrics. According to Bajusz *et al.* [35] the relationship of such metrics and their arithmetic mean average is not linear. To get a fair comparison it is essential to scale all similarity measures to the same range using different transformation methods such as: standardization (mean centering and scaling to the unit standard deviation), interval scaling (between 0 and 1) and rank transformation.

Figure 7

Complete list of the SRD scores is given in the **Table S6**, Supplementary material. The ranking is slightly altered by the data preprocessing method. However, the main pattern remains preserved. In the case of the standardized data (**Figure 7**) the best, *i.e.*, the closest to the average (consensus) similarity measure is based on the Kendal's correlation coefficient, which is closely followed by the GPCM and SRD. All three are derived from the primary retention data. The worst similarity measures, which fall under the bell shaped distribution curve of random SRD values, and therefore are statistically insignificant, are mostly based on HSM parameters, derived from non-parametric metrics (*i.e.*, Kendal's and Spearman's rank correlation coefficients, as well as scaled SRD values, denoted by GPCM_HSM, KNT_HSM, SPR_HSM and SRD_HSM). Generally, a trend of lower SRD values of similarities based on primary retention data can be noticed. Similarity derived from the *F*s comparison function, originally proposed by Snyder, works fairly well, although not being the best ranked.

In order to estimate the effect of (*i*) a preprocessing method of similarity measures, (*ii*) the sort of similarity measures, and (*iii*) the type of chromatographic data (primary retention *vs*. HSM parameters), on the ranking of similarity measures, the uncertainty was estimated for

each of the SRD scores by the leave-one-out cross-validation procedure. In that way 594 SRD score values were obtained (11 repetitions × 3 sorts of preprocessing methods × 9 similarity measures × 2 types of chromatographic data). SRD scores were further subjected to the factorial analysis of variance with type III decomposition, resulting in the full linear model with Eq. 4.

$$Score = b_0 + b_1F_1 + b_2F_2 + b_3F_3 + b_{12}F_1F_2 + b_{13}F_1F_3 + b_{23}F_2F_3 + b_{123}F_1F_2F_3 \quad (4)$$

Where F_1 accounts for the data preprocessing methods in three levels coded as standardization (STD), range scaling (SCL), and rank transformation (RNK); F_2 represents the type of chromatographic data, coded at two levels: retention *vs*. HSM selectivity data, and finally, F_3 encodes information about similarity measures at nine levels: EUC, MNH, CHD, COR, SPR, KNT GPCM, SRD, and COS. Snyder's selectivity function *F*s, was omitted in order to make factorial design balanced (*F*s as such cannot be defined for primary retention data).

Statistical parameters of ANOVA are listed in the **Table 3**. All factors including their cross-coupling terms, with exception of F_1 and the corresponding interaction term are statistically significantly affecting the ranking of similarity measures at the predefined significance level of $p = 0.05$.

Figure 8 shows the statistical significance of factor effects in a way that is much easier to understand.

Table 3

Figure 8

Clearly, standardization and range scaling do not significantly affect the outcome of SRD ranking of similarity indices. However, in the case of primary retention data, rank transformation provides significantly lower SRD scores for all similarity coefficients, while in

the case of the HSM parameters, non-parametric based similarity measures (KNT, GPCM, SRD, and SPR) have significantly higher SRDs.

Considering the fact that the lower the SRD scores are, the better the performance of a similarity measure is, it is clear that in most cases similarities based on HSM parameters results in significantly higher SRD scores (perform worse) compared to those derived from primary retention data. The exceptions are: MNH, COR and EUC in which case the use of HS parameters leads to equal or significantly better performance (EUC in the case of rank transformed data). The best similarity measures are: KNT, SRD, GPCM, and SPR derived from primary retention data. They could be particularly useful for comparison methods based on retention coefficients such as KUL, Tanaka, Euerby *etc*. instead of currently used Euclidian distance. Also, this clearly demonstrates a significant loss of information due to reinforcement of hydrophobic-subtraction model which is based on a limited number of assumed interaction terms) to describe column selectivity. However, because the *F*s parameter is nothing else but a weighted form of Euclidian distance, it should be expected that itself, as well as its corresponding similarity measure have a moderate performance such as: CHD, COS and SPR in the case of primary retention data.

4. Conclusion

Non-parametric ranking based on Sum of Ranking Differences (SRD) as well as the Generalized Pairwise Correlation Method (GPCM) combined with conditional exact Fisher's test (CE) are able to fairly compare, rank, and cluster chromatographic columns using a consensus based approach (arithmetic mean as a reference). The methods are sensitive in exploring refined differences among ten very similar chromatographic columns. Unlike Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) the methods provide statistical significance of ranking (grouping) patterns.

The methodology of SRD and GPCM, extended by the comparison with one variable at a time (COVAT) approach, provide symmetric and asymmetric matrices, respectively. The methodology uses each variable as a reference for comparison and possible grouping of chromatographic columns is revealed without selection of one particular (dedicated) column. The relative coloring of matrices enables the most similar and orthogonal chromatographic systems to be identified in an easily perceivable way.

In all instances the ranking (clustering) pattern of chromatographic columns based on primary retention data (log*k* values), and hydrophobic-subtraction model (HSM) selectivity parameters gave different results. However, the primary retention data resulted in patterns that are consistent with differences in the columns' physicochemical parameters, while HSM provide results that are drifting away to higher or lesser degree, depending on the particular chemometric approach, plus they are inconsistent within the various methods applied.

Consensus based SRD comparison of nine similarity measures based on bound and unbound distances (dissimilarities), implied the existence of a trend in *significantly lower* SRD scores associated with similarity indices derived from *primary retention* data as compared to the ones derived from HSM parameters. The lower SRD scores indicate better performance and *vice versa*.

Analysis of variance confirmed that (*i*) the type of data preprocessing was insignificant, but (*ii*) the type of chromatographic data (primary retention *vs*. HSM parameters) significantly affects the ranking of similarity measures, *i.e*., application of the HSM to retention data results in statistically significant increase of SRD values (worsen performance) for majority of similarity indices (exception are those based on Euclidian and Manhattan distance). Therefore, significant loss of information arises during the application of the HSM. The sort of distance measure is another factor influencing the performance of similarity indices. The

best ones are obtained on primary retention data and derived from Kendall's and Spearman's correlation coefficients, as well as GPCM and SRD score values. Selectivity function, *F*s, originally proposed by Snyder, has moderate (medium) performance.

Although our investigations are limited to one (well-known) data set, the procedure and algorithm can be carried out on any data sets partially and on the whole to select the most similar and dissimilar columns.

Acknowledgment

This work was supported by the Ministry of Education, Science and Technological development of the Republic of Serbia [grant number 172017]; the National Research Development and Innovation office, OTKA, Hungary [grant number K119269]; and the Joint research project under the cooperation between the Serbian Academy of Sciences and Arts, and the Hungarian Academy of Sciences [grant numbers: HF-2016-02, and NKM 70/2016].

5. References

[1] K. M. Biswas, B. C. Castle, B. A. Olsen, D. S. Risley, M. J. Skibic, P. B. Wright, A simple and efficient approach to reversed-phase HPLC method screening, J. Pharm. Biomed. Anal. 49 (2009) 692–701.

[2] P. Dehouck, D. Visky, Y. Vander Heyden, E. Adams, Z. Kovacs, B. Noszal, D. L. Massart, J. Hoogmartens, Characterisation of reversed-phase liquid-chromatographic columns by chromatographic tests. Comparing column classification based on chromatographic parameters and column performance for the separation of acetylsalicylic acid and related compounds, J. Chromatogr. A 1025 (2004) 189–200.

[3] E. Lesellier, C. West, Description and comparison of chromatographic tests and chemometric methods for packed column classification, J. Chromatogr. A 1158 (2007) 329–360.

[4] H. Engelhardt, M. Jungheim, Comparison and characterization of reversed phases, Chromatographia 29 (1990) 59–68.

[5] U. D. Neue, B. A. Alden, Thomas H. Walter, Universal procedure for the assessment of the reproducibility and the classification of silica-based reversed-phase packings II. Classification of reversed-phase packings, J. Chromatogr. A 849 (1999) 101–116.

[6] K. Kimata, K. Iwaguchi, S. Onishi, K. Jinno, R. Eksteen, K. Hosoya, M. Araki, and N. Tanaka, Chromatographic characterization of silica C18 packing materials. Correlation between a preparation method and retention behavior of stationary phase, J Chromatogr. Sci. 27 (1989) 721–728.

[7] E. Cruz, M.R. Euerby, C.M. Johnson, C.A. Hackett, Chromatographic classification of commercially available reverse-phase HPLC columns, Chromatographia 44 (1997) 151–161.

[8] M. R. Euerby, P. Petersson, Chromatographic classification and comparison of commercially available reversed-phase liquid chromatographic columns using principal component analysis, J. Chromatogr. A 994 (2003) 13–36.

[9] M. R. Euerby, P. Petersson, W. Campbell, W. Roe, Chromatographic classification and comparison of commercially available reversed-phase liquid chromatographic columns containing phenyl moieties using principal component analysis, J. Chromatogr. A 1154 (2007) 138–151.

[10] E. Haghedooren, T. Janssens, R. Nijs, S. K. Park, E. Farkas, S. Dragovic, B. Noszal, J. Hoogmartens, E. Adams, Selecting a Suitable LC Column for Pharmaceutical Separations using a Column Characterization System, J. Liq. Chromatogr. Relat. Technol. 32 (2009) 1–25.

[11] E. Haghedooren, D. Jose, V. Noszal, J. Hoogmartens, and E. Adams, "Classification of reversed-phase columns based on their selectivity towards vancomycin compounds, Talanta 71 (2007) 31–37.

[12] P. Dehouck, D. Visky, G. Van den Bergh, E. Haghedooren, E. Adams, A. Kerner, Y.Vander Heyden, D.L. Massart, Z. Kovacs, B. Noszal, J. Hoogmartens, acilitated column ranking and selection in reversed-phase liquid chromatographic analysis, LC–GC Eur. 17 (2004) 592–601.

[13] E. Haghedooren, A. Kerner, B. Noszal, J. Hoogmartens, E. Adams, Application of an improved column characterisation system to evaluate the within and between batch variability, J. Pharm. Biomed. Anal. 44 (2007) 634–639.

[14] R. Kaliszan, M. A. van Straten, M. Markuszewski, C. A. Cramers, H. A. Claessens, Molecular mechanism of retention in reversed-phase high-performance liquid chromatography and classification of modern stationary phases by using quantitative structure–retention relationships, J. Chromatogr. A 855 (1999) 455–486.

[15] T. Baczek, R. Kaliszan, K. Novotna, P. Jandera, Comparative characteristics of HPLC columns based on quantitative structure–retention relationships (QSRR) and hydrophobic-subtraction model, J. Chromatogr. A 1075 (2005) 109–115.

[16] A. Sandi, L. Szepesy, Characterization of various reversed-phase columns using the linear free energy relationship I. Evaluation based on retention factors, J. Chromatogr. A 818 (1998) 1–17.

[17] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, R.G. Wolcott, P.W. Carr, Column selectivity in reversed-phase liquid chromatography I. A general quantitative relationship, J. Chromatogr. A 961 (2002) 171–193.

[18] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, P.W. Carr, C olumn selectivity in reversed-phase liquid chromatography II. Effect of a change in conditions, J. Chromatogr. A 961 (2002) 195–215.

[19] J.J. Gilroy, J.W. Dolan, L.R. Snyder, Column selectivity in reversed-phase liquid chromatography IV. Type-B alkyl-silica columns, J. Chromatogr. A 1000 (2003) 757–778.

[20] L.R. Snyder, A. Maule, A. Heebsh, R. Cuellar, S. Paulson, J. Carrano, L. Wrisley, C.C. Chan, N. Pearson, J.W. Dolan, J.J. Gilroy, A fast, convenient and rugged procedure for characterizing the selectivity of alkyl-silica columns, J. Chromatogr. A 1057 (2004) 49–57.

[21] L.R. Snyder, J.W. Dolan, P.W. Carr, The hydrophobic-subtraction model of reversed-phase column selectivity, J. Chromatogr. A 1060 (2004) 77–116.

[22] S. Dragovic, E. Haghedooren, T. Nemeth, I. M. Palabiyik, J. Hoogmartens, E. Adams, Evaluation of two approaches to characterize liquid chromatographic columns using pharmaceutical separations, Journal of Chromatography A 1216 (2009) 3210–3216.

[23] E. Haghedooren, T. Nemeth, S. Dragovic, B. Noszal, J. Hoogmartens, E. Adams, Comparison of two column characterisation systems based on pharmaceutical applications, J. Chromatogr. A 1189 (2008) 59–71.

[24] E. Van Gyseghem, B. Dejaegher, R. Put, P. Forlay-Frick, A. Elkihel, M. Daszykowski, K. Heberger, D.L. Massart, Y. Vander Heyden, Evaluation of chemometric techniques to select orthogonal chromatographic systems, J. Pharm. Biomed. Anal. 41 (2006) 141–151.

[25] E. Van Gyseghem, M. Jimidar, R. Sneyers, D. Redlich, E. Verhoeven, D.L. Massart, Y. Vander Heyden, Orthogonality and similarity within silica-based reversed-phased chromatographic systems, J. Chromatogr. A 1074 (2005) 117–131.

[26] M. Dumarey, R. Put, E. Van Gyseghem, Y. Vander Heyden, Dissimilar or orthogonal reversed-phase chromatographic systems: A comparison of selection techniques, Anal. Chim. Acta 609 (2008) 223–234.

[27] Peter Forlay-Frick, Elke Van Gyseghem, Karoly Heberger, Yvan Vander Heyden, Selection of orthogonal chromatographic systems based on parametric and non-parametric statistical tests, Anal. Chim. Acta 539 (2005) 1–10.

[28] M.F. Delaney, A.N. Papas, M.J. Walters, Chemometric classification of reversed-phase high-performance liquid chromatography columns, J. Chromatogr. A 410 (1987) 31–41.

[29] J.G. Shackman, Discrepancies in column parameters presented in hydrophobic subtraction model manuscripts, J. Chromatogr. A 1475 (2016) 116–118.

[30] K. Heberger, Sum of ranking differences compares methods or models fairly TRAC - Trends Anal. Chem. 29 (2010) 101–109.

[31] K. Heberger, K. Kollar-Hunek, Sum of ranking differences for method discrimination and its validation: comparison of ranks with random numbers J. Chemometr. 25 (2011) 151–158.

[32] K. Kollar-Hunek, K. Heberger, Method and model comparison by sum of ranking differences in cases of repeated observations (ties) Chemometr. Intell. Lab. Syst. 127 (2013) 139–146.

[33] K. Heberger and R. Rajko, Variable Selection using Pair-Correlation Method. Environmental Applications. SAR and QSAR Environ. Res. 13 (2002) 541–554.

[34] K. Heberger and R. Rajko, Generalization of Pair-Correlation Method (PCM) for Nonparametric Variable Selection, J. Chemometr. 16 (2002) 436–443.

[35] D. Bajusz, A. Racz, K. Heberger, Why is Tanimoto index the most appropriate choice for fingerprint-based similarity calculations? J Cheminform. 7 (2015) Art. no. 20

[36] K. Heberger, S. Kolarevic, M. Kracun-Kolarevic, K. Sunjog, Z. Gacic, Z. Kljajic, M Mitric, Branka Vukovic-Gacic, Evaluation of single cell gel electrophoresis data: Combination of variance analysis with sum of ranking differences, Mutat. Res. Genet. Toxicol. Environ. Mutagen. 771 (2014) 15–22.

[37] J. Bin, F.-F. Ai, W. Fan, J.-H. Zhou, Y.-H. Yun and Y.-Z. Liang, A modified random forest approach to improve multi-class classification performance of tobacco leaf grades coupled with NIR spectroscopy, RSC Adv. 6 (2016) 30353–30361.

[38] C. West, M. Khalikova, E. Lesellier, K. Heberger, Sum-of-ranking-differences to rank stationary phases used in packed column supercritical fluid chromatography, Journal of Chromatography A, 1409 (2015) 241–250.

[39] W. Nowik, S. Heron, M. Bonose, A. Tchapla, Separation system suitability (3S): A new criterion of chromatogram classification in HPLC based on cross-evaluation of separation capacity/peak symmetry and its application to complex mixtures of anthraquinones Analyst, 138 (2013) 5801–5810.

[40] A. J. Tencate, J. H. Kalivas, A. J. White, Fusion strategies for selecting multiple tuning parameters for multivariate calibration and other penalty based processes: A model updating application for pharmaceutical analysis, Anal. Chim. Acta. 921 (2016) 28–37.

[41] F. Andric, D. Bajusz, A. Racz, S. Segan, K. Hebeger, Multivariate assessment of lipophilicity scales – computational and reversed phase thin-layer chromatographic indices, J. Pharm. Biomed. Anal. 127 (2016) 81–93.

[42] R. Todeschini, D. Ballabio, V. Consonni, (2015) Distances and Other Dissimilarity Measures in Chemometrics, Encyclopedia of Analytical Chemistry, John Wiley & Sons, Ltd., 1–34.

Tables

Table 1 List of chromatographic columns and their properties; Data are taken from the reference [17] (Copyright Elsevier: license number 3995960369478, License date Nov 25, 2016).

No.	Column name	Abbreviation	Surface area (m^2/g)	Pore diameter (nm)	%C	$\mu mol/m^2$
1	GL Inertsil ODS-3	Inertsil	436	9.5	14.7	1.74
2	Waters Symmetry C_{18}	Symmetry	343	9	19.7	3.13
3	HP Zorbax SB C_{18}	SB-100	186	8	10.4	2.08
4	HP Zorbax SB C_{18}^{c}	SB-90	188	8	9.20	1.79
5	HP Zorbax SB-300 C_{18}	SB-300	52	30	3.25	2.09
6	HP Eclipse XDB C_{18}	Eclipse	186	8	10.7	3.00
7	YMC Pack Pro C_{18}	YMC 15	322	12.5	15.5	2.51
8	YMC Pack Pro C_{18}	YMC 16	321	12.5	16.3	2.68
9	YMC Pack Pro C_{18}	YMC 17	322	12.5	17.0	2.82
10	Supelco Discovery C_{18}	Discovery	190-220	17-20	12.5	3.12

Table 2 Dissimilarity measures and similarity measure transformation functions

Dissimilarity measure	Label	Definition	Range	Similarity measure definition or transformation function
Manhattan (City block) distance	MNH	$d(A,B)=\sum_{j=1}^{n}\lvert x_{jA}-x_{jB}\rvert$	$0<d(A,B)<+\infty$	$S(A,B)=\frac{1}{1+d(A,B)}$
Euclidean distance	EUC	$d(A,B)=\sqrt{\sum_{j=1}^{n}(x_{jA}-x_{jB})^2}$	$0<d(A,B)<+\infty$	$S(A,B)=\frac{1}{1+d(A,B)}$
Chebyshev distance	CHD	$d(A,B)=\max_{i}\left(\lvert x_{jA}-x_{jB}\rvert\right)$	$0<d(A,B)<+\infty$	$S(A,B)=\frac{1}{1+d(A,B)}$
Pearson's correlation distance defined using absolute values	COR	$d(A,B)=1-\lvert r_{AB}\rvert$ r_{AB} is the Pearson's correlation coefficient among objects A and B in *n*-dimensional Euclidean space.	$0<d(A,B)<1$	$S(A,B)=1-d(A,B)$
Spearman's rank correlation distance defined using absolute values	SPR	$d(A,B)=1-\lvert \rho_{AB}\rvert$ ρ_{AB} is the Spearman's rank correlation coefficient among objects A and B in *n*-dimensional Euclidean space.	$0<d(A,B)<1$	$S(A,B)=1-d(A,B)$
Kendall's rank correlation distance defined using absolute values	KNT	$d(A,B)=1-\lvert \tau_{AB}\rvert$ τ_{AB} is the Kendall's rank correlation coefficient among objects A and B in *n*-dimensional Euclidean space.	$0<d(A,B)<1$	$S(A,B)=1-d(A,B)$

Dissimilarity measure	Label	Definition	Range	Similarity measure definition or transformation function
Cosine distance defined using absolute values, also called Congruence coefficient	COS	$d(A,B) = 1 - \lvert\cos\theta_{AB}\rvert$	$0 < d(A,B) < 1$	$\lvert\cos\theta_{AB}\rvert = \frac{\left\lvert\sum_{j=1}^{n} x_{jA} x_{j\mathbf{B}}\right\rvert}{\sqrt{\sum_{j=1}^{n}(x_{jA})^2 \sum_{j=1}^{n}(x_{j\mathbf{B}})^2}}$
SRD	SRD	$d(A,B) = \sum_{j=1}^{n} \lvert R(x_{jA}) - R(x_{jB})\rvert$	$0 < d(A,B) < 100$[a]	$S(A,B) = \frac{1}{1+d(A,B)}$
GPCM (Conditional Fisher's exact test with probability weighted ordering, CE-PW)	GPCM	$d(A,B) = N_{winsA,B} p_{winsA,B} - N_{lossesA,B} p_{lossesA,B}$	$0 < d(A,B) < 100$[b]	$S(A,B) = \frac{1}{1+d(A,B)}$

[a] Scaled to the range between 0 and 100
[b] Commonly rescaled to fit the SRD values

Table 3 Statistical parameters of factor effects in ANOVA model (type III decomposition).

Statistically significant factor effects are marked in bold

Factor	*SS*	*D.F.*	*MS*	*F*	*p*
Intercept	**546030.5**	**1**	**546030.5**	**4891.58**	**<0.001**
F_1	223.3	2	111.6	0.086	0.920
F_2	**21570.4**	**1**	**21570.4**	**19.78**	**0.047**
F_3	**9838.3**	**8**	**1229.8**	**4.02**	**0.007**
$F_1 \times F_2$	**2179.2**	**2**	**1089.6**	**10.91**	**0.001**
$F_1 \times F_3$	**4895.3**	**16**	**306.0**	**3.06**	**0.016**
$F_2 \times F_3$	**36960.1**	**8**	**4620.0**	**46.27**	**<0.001**
$F_1 \times F_2 \times F_3$	**1597.5**	**16**	**99.8**	**4.82**	**<0.001**
Error	11173.2	540	20.7		

Figure captions:

Figure 1 Retention profiles of solutes presented in a form of radar plots of retention coefficients; Clockwise order of the chromatographic columns starting from the top point in each radar plot is: Inertsil, Symmetry, SB-100, SB-90, SB-300, Eclipse, YMC 15, YMC 16, YMC 16, and Discovery.

Figure 2 PCA of primary retention data (a, b) and hydrophobicity subtraction (HS) selectivity parameters (c, d); Score plot (a) indicates grouping of chromatographic columns, while loading diagram (b) demonstrates similarities among studied compounds; Diagram of scores (c) shows grouping of chromatographic columns according to disposition of HS selectivity parameters - loading plot (d).

Figure 3 SRD and GPCM CE-PW comparison of chromatographic columns based on primary retention data (a, b) and hydrophobicity subtraction selectivity parameters (c, d); CE-PW stands for Fisher's conditional exact test (CE), probability weighted (PW) ranking. VARC was used as the benchmark.

Figure 4 Results of the sevenfold SRD cross-validation, Box-plot visualizes the ranking of chromatographic columns (arranged in ascending order of the SRD median values), and their separation in sections according to statically significant difference (Wilcoxon-matched pair test and sign test at predefined significance level of $p = 0.05$)

Figure 5 COVAT relative heatmaps showing similarity/orthogonality relationship among chromatographic columns based on primary retention (a, c) and hydrophobicity subtraction data (b, d), using SRD (a, b) and GPCM CE-PW (c, d) comparison methods. Red color represents the lowest score value (the highest similarity), while blue marks the highest one (the lowest similarity). Color codes are provided on the right side with absolute and relative (%) values. CE-PW stands for probability weighted ranking (PW) based on Fisher's conditional exact test (CE).

Figure 6 Stacked plot: Cumulative distance (a, b) and similarity (c, d) profiles of a series of chromatographic columns, calculated from primary retention data (a, c) and hydrophobicity subtraction selectivity parameters (b, d). Similarity and distance measures were scaled between 0 and 1.

Figure 7 Ranking and comparison of similarity measures (standardized data) using average as a reference; *x* and *y* left sided axes are SRD values scaled between 0 and 100, right sided *y* axis represents relative frequencies of the theoretical distribution of ranking random numbers; Dashed XX1 line denotes statistical significance at $p = 0.05$ (left side); ret, and HS denote similarity measures calculated from the primary retention and HS data, respectively.

Figure 8 Effect of factors by analysis of variance for tenfold cross-validated SRD score values of similarity measures; the average was used for reference in ranking. Score values were plotted on the *y*-axis. Vertical bars denote 0.95 confidence intervals.

Figure 1

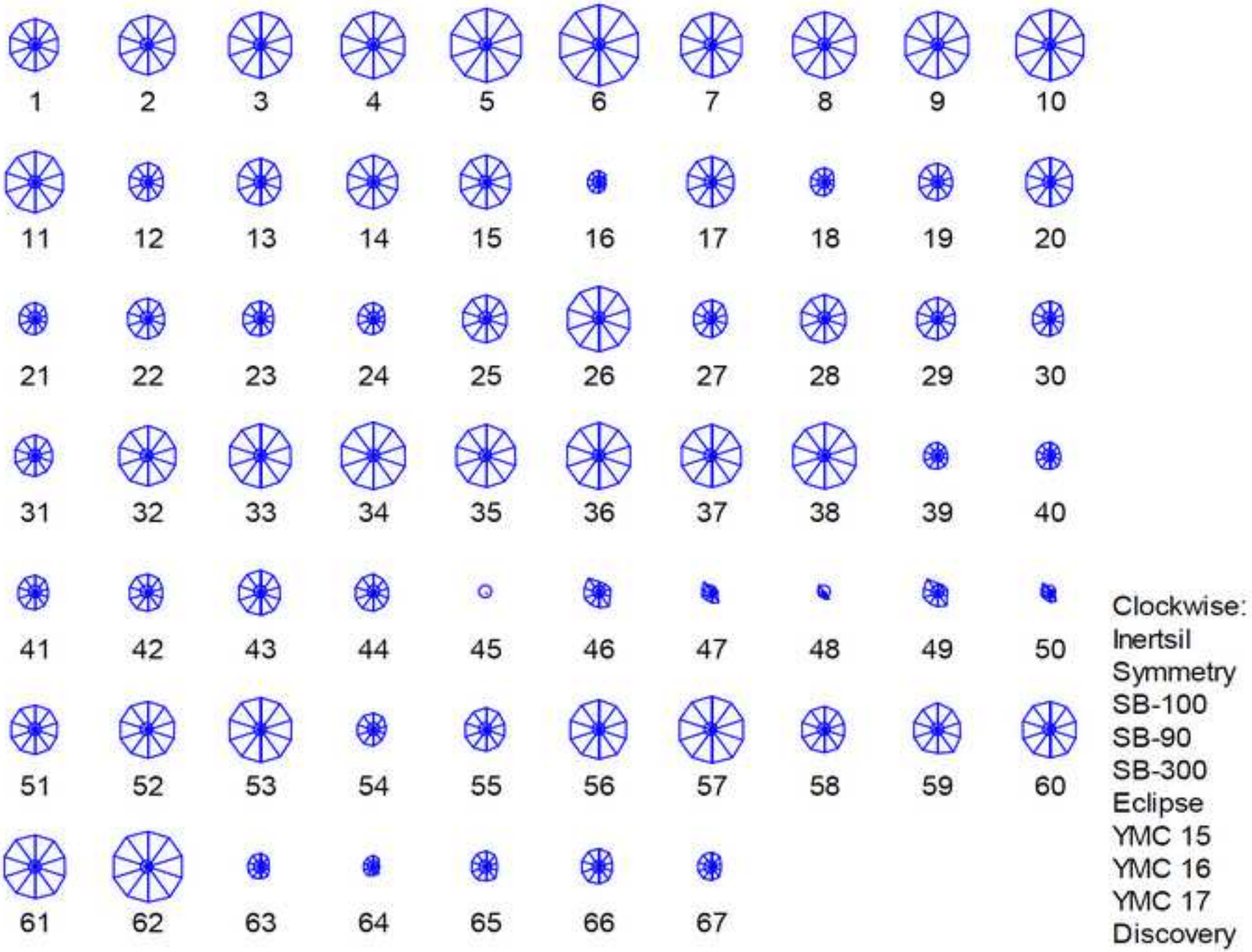

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
Clockwise:
Inertsil
Symmetry
SB-100
SB-90
SB-300
Eclipse
YMC 15
YMC 16
YMC 17
Discovery

Figure 2

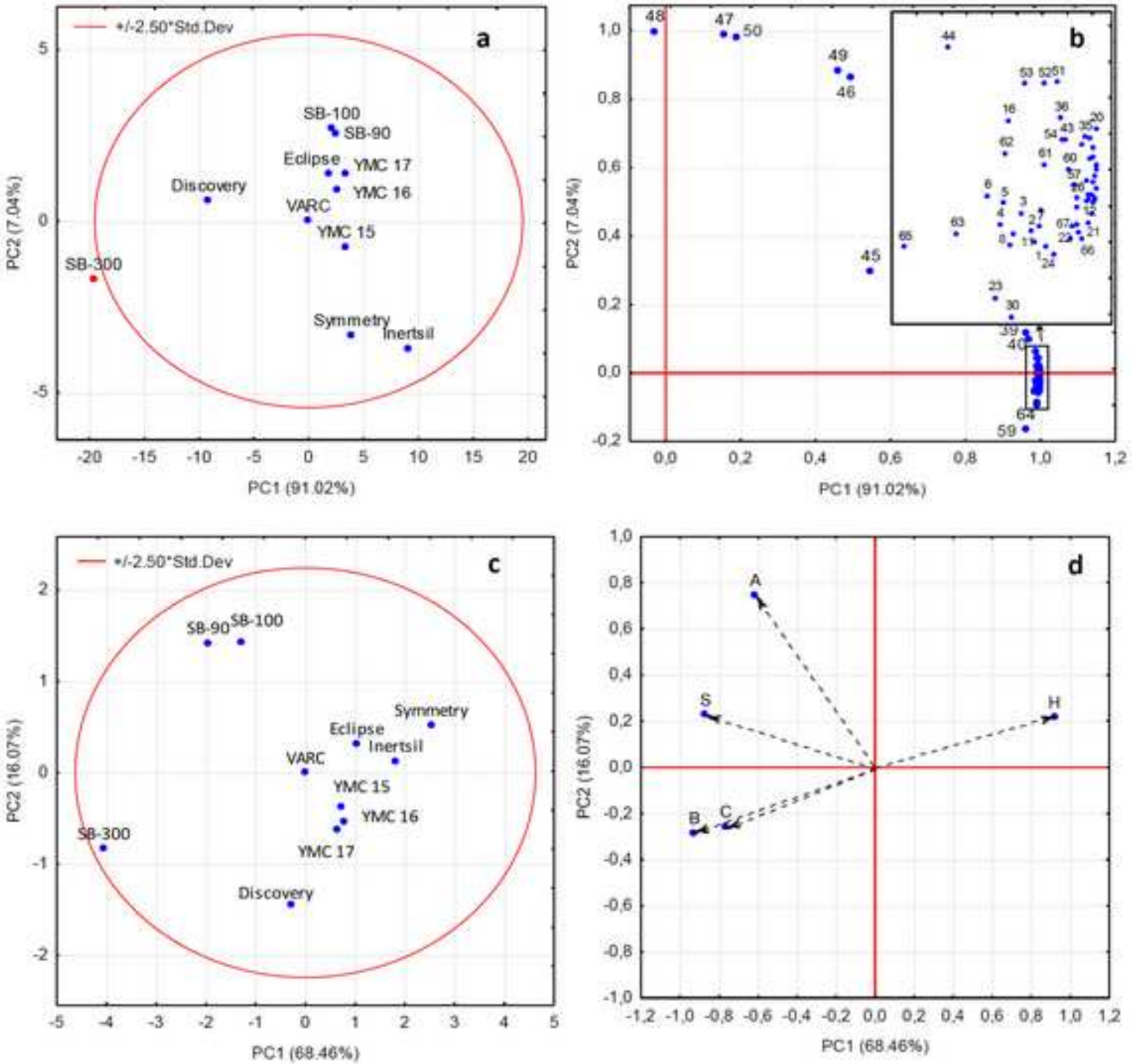

a
+/-2.50*Std.Dev
SB-100
SB-90
Eclipse
YMC 17
Discovery
YMC 16
VARC
YMC 15
SB-300
Symmetry
Inertsil
PC1 (91.02%)
PC2 (7.04%)
b
PC1 (91.02%)
PC2 (7.04%)
c
+/-2.50*Std.Dev
SB-90
SB-100
Symmetry
Eclipse
Inertsil
VARC
YMC 15
YMC 16
SB-300
YMC 17
Discovery
PC1 (68.46%)
PC2 (16.07%)
d
A
S
H
B
C
PC1 (68.46%)
PC2 (16.07%)

Figure 3

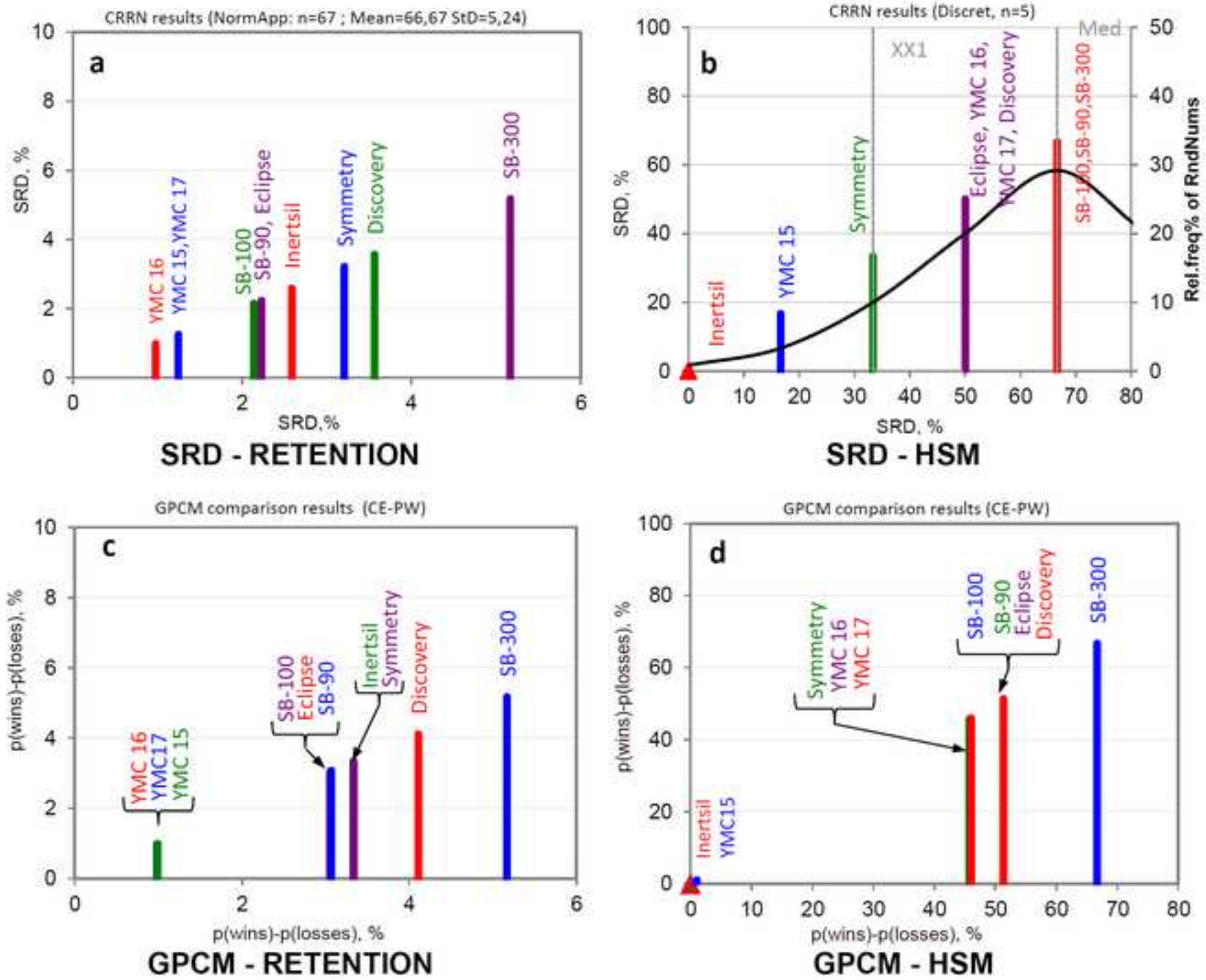

CRRN results (NormApp: n=67 ; Mean=66,67 StD=5,24)
a
SRD, %
SRD,%
YMC 16
YMC 15,YMC 17
SB-100
SB-90, Eclipse
Inertsil
Symmetry
Discovery
SB-300
SRD - RETENTION
CRRN results (Discret, n=5)
b
XX1
Med
SRD, %
Rel.freq% of RndNums
Inertsil
YMC 15
Symmetry
Eclipse, YMC 16, YMC 17, Discovery
SB-100,SB-90,SB-300
SRD - HSM
GPCM comparison results (CE-PW)
c
p(wins)-p(loses), %
YMC 16
YMC17
YMC 15
SB-100
Eclipse
SB-90
Inertsil
Symmetry
Discovery
SB-300
GPCM - RETENTION
GPCM comparison results (CE-PW)
d
Inertsil
YMC15
Symmetry
YMC 16
YMC 17
SB-100
SB-90
Eclipse
Discovery
SB-300
GPCM - HSM

Figure 4

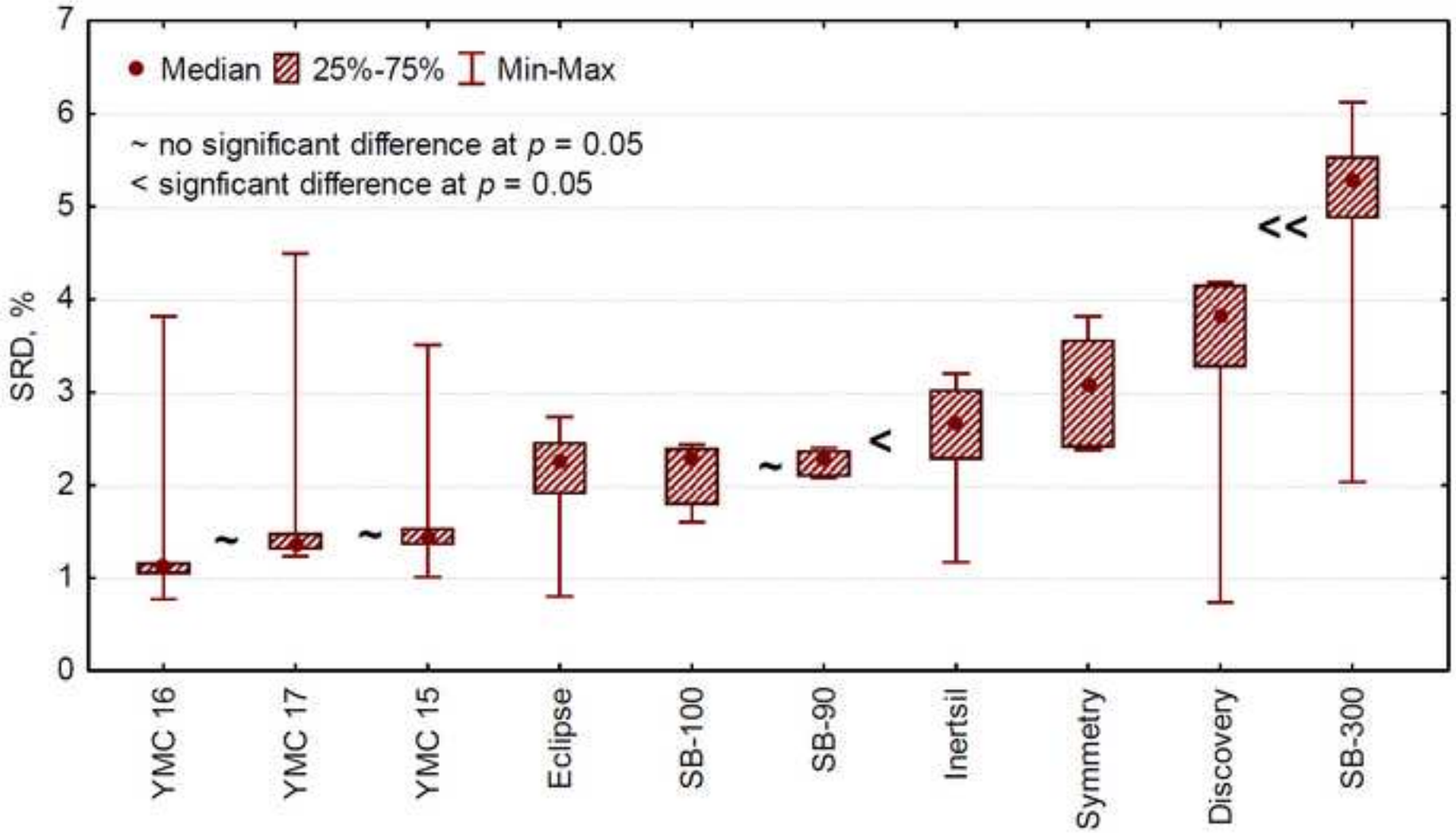

Median
25%-75%
Min-Max
~ no significant difference at p = 0.05
< signficant difference at p = 0.05
SRD, %
0
1
2
3
4
5
6
7
~
~
~
<
<<
YMC 16
YMC 17
YMC 15
Eclipse
SB-100
SB-90
Inertsil
Symmetry
Discovery
SB-300

Figure 5

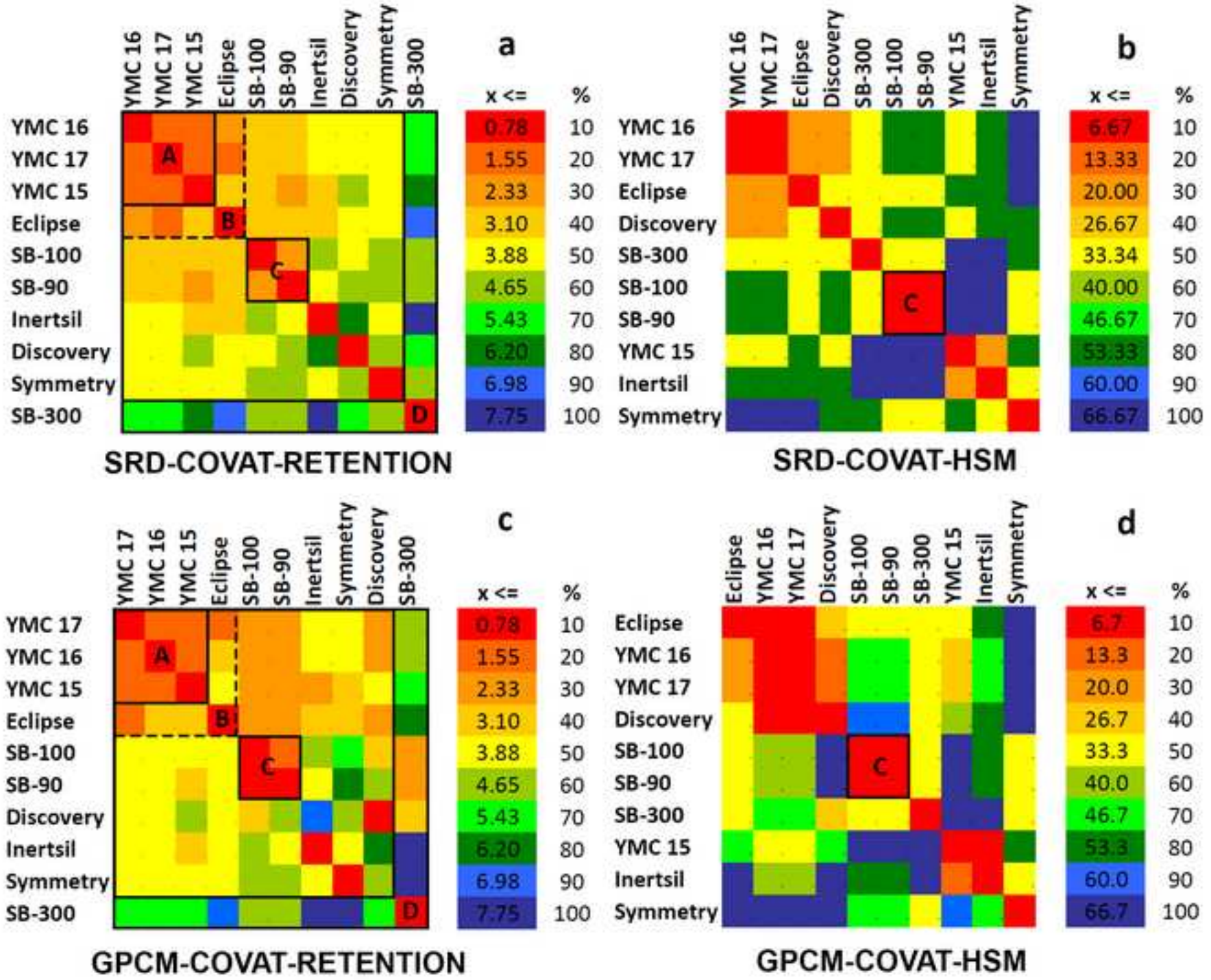

a
YMC 16
YMC 17
YMC 15
Eclipse
SB-100
SB-90
Inertsil
Discovery
Symmetry
SB-300
A
B
C
D
x <=
%
0.78 10
1.55 20
2.33 30
3.10 40
3.88 50
4.65 60
5.43 70
6.20 80
6.98 90
7.75 100
SRD-COVAT-RETENTION
b
YMC 16
YMC 17
Eclipse
Discovery
SB-300
SB-100
SB-90
YMC 15
Inertsil
Symmetry
C
x <=
%
6.67 10
13.33 20
20.00 30
26.67 40
33.34 50
40.00 60
46.67 70
53.33 80
60.00 90
66.67 100
SRD-COVAT-HSM
c
YMC 17
YMC 16
YMC 15
Eclipse
SB-100
SB-90
Inertsil
Symmetry
Discovery
SB-300
A
B
C
D
x <=
%
0.78 10
1.55 20
2.33 30
3.10 40
3.88 50
4.65 60
5.43 70
6.20 80
6.98 90
7.75 100
GPCM-COVAT-RETENTION
d
Eclipse
YMC 16
YMC 17
Discovery
SB-100
SB-90
SB-300
YMC 15
Inertsil
Symmetry
C
x <=
%
6.7 10
13.3 20
20.0 30
26.7 40
33.3 50
40.0 60
46.7 70
53.3 80
60.0 90
66.7 100
GPCM-COVAT-HSM

Figure 6

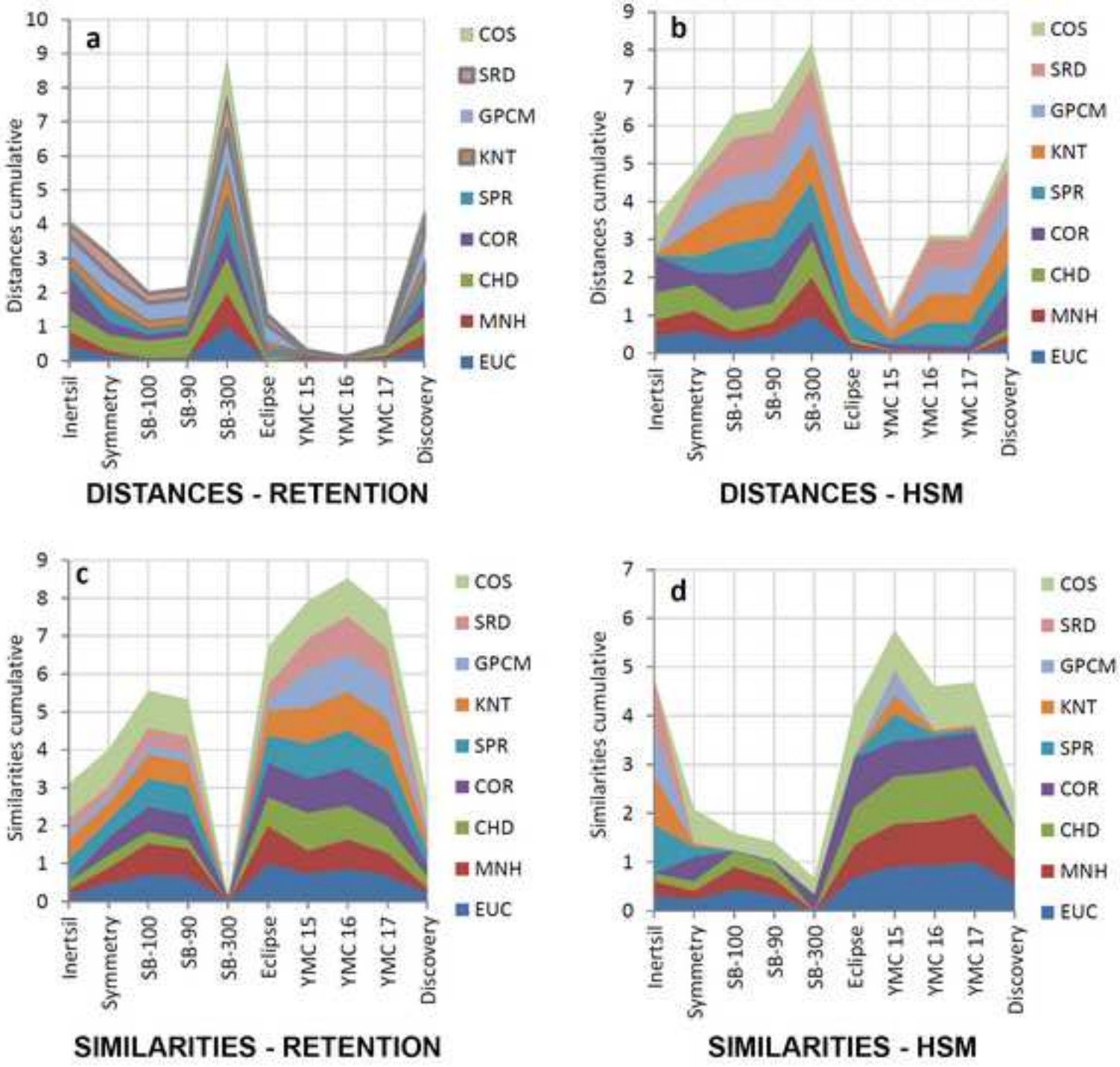

a
Distances cumulative
COS
SRD
GPCM
KNT
SPR
COR
CHD
MNH
EUC
Inertsil
Symmetry
SB-100
SB-90
SB-300
Eclipse
YMC 15
YMC 16
YMC 17
Discovery
DISTANCES - RETENTION
b
Distances cumulative
DISTANCES - HSM
c
Similarities cumulative
SIMILARITIES - RETENTION
d
Similarities cumulative
SIMILARITIES - HSM

Figure 7

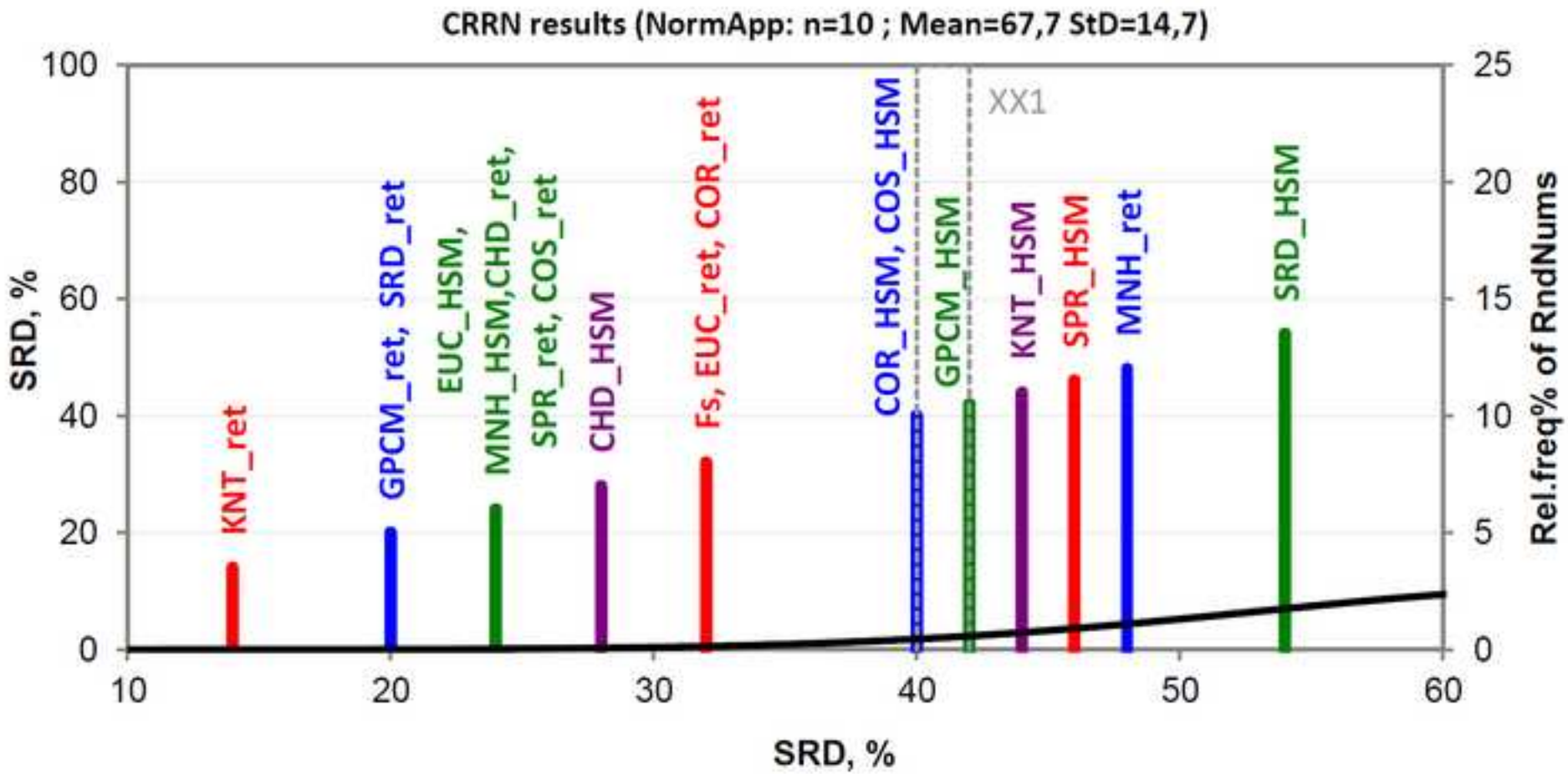
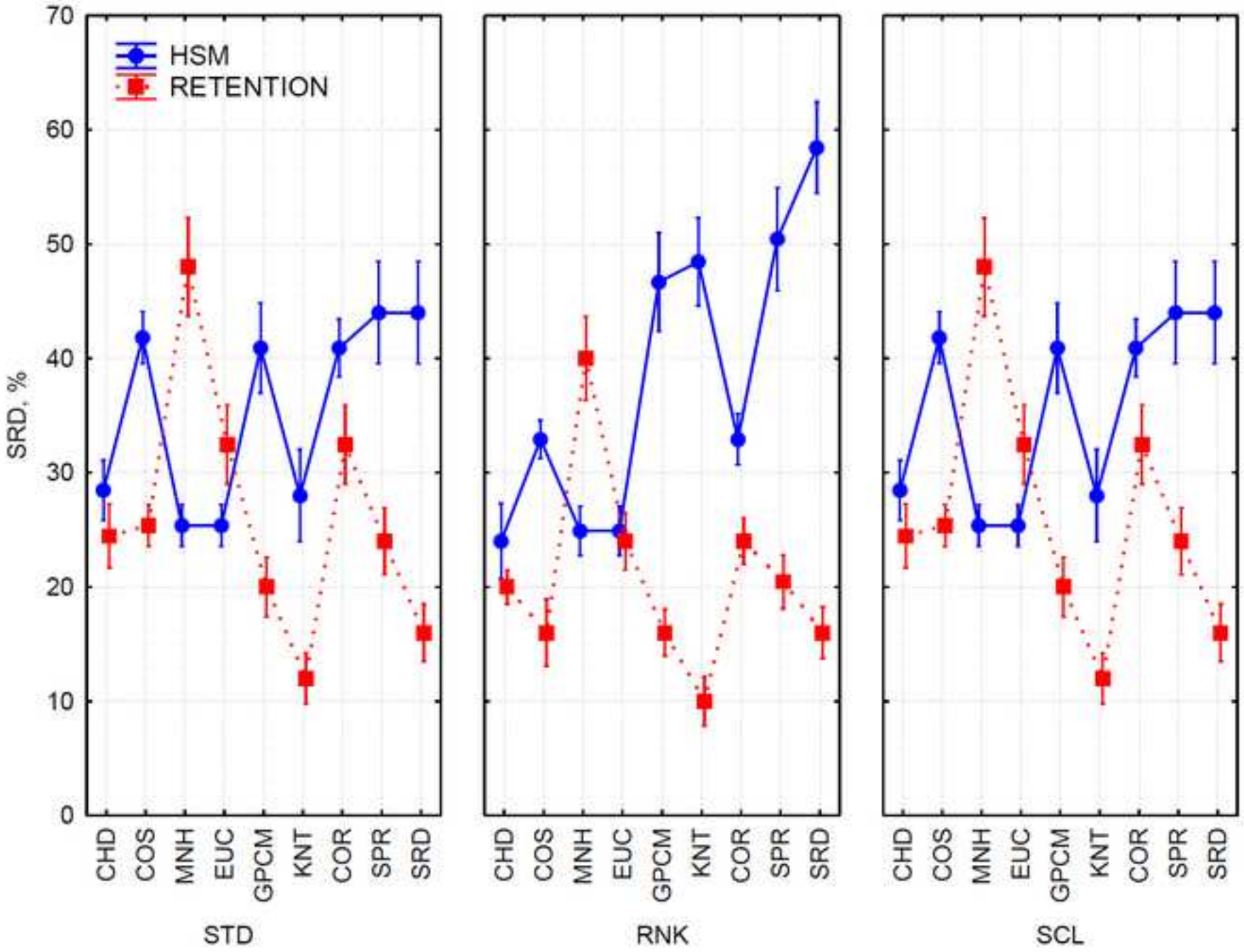

CRRN results (NormApp: n=10 ; Mean=67,7 StD=14,7)
SRD, %
Rel.freq% of RndNums
SRD, %
100
80
60
40
20
0
25
20
15
10
5
0
10
20
30
40
50
60
KNT_ret
GPCM_ret, SRD_ret
EUC_HSM,
MNH_HSM,CHD_ret,
SPR_ret, COS_ret
CHD_HSM
Fs, EUC_ret, COR_ret
COR_HSM, COS_HSM
GPCM_HSM
XX1
KNT_HSM
SPR_HSM
MNH_ret
SRD_HSM

Figure 8


HSM
RETENTION
SRD, %
70
60
50
40
30
20
10
0
CHD
COS
MNH
EUC
GPCM
KNT
COR
SPR
SRD
STD
RNK
SCL