symmetry

MDPI

Article

Fuzzy Divisive Hierarchical Clustering of Solvents According to Their Experimentally and Theoretically Predicted Descriptors

Miroslava Nedyalkova [1], Costel Sarbu [2], Marek Tobiszewski [3] and Vasil Simeonov [4,*]

[1] Department of Inorganic Chemistry, Faculty of Chemistry and Pharmacy, University of Sofia, 1 James Bourchier Blvd., Sofia 1164, Bulgaria; nhmn@chem.uni-sofia.bg
[2] Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 400084 Cluj-Napoca, Romania; csarbu@chem.ubbcluj.ro
[3] Department of Analytical Chemistry, Chemical Faculty, Gdańsk University of Technology (GUT), 11/12 G. Narutowicza St., 80-233 Gdańsk, Poland; marektobiszewski@wp.pl
[4] Department of Analytical Chemistry, Faculty of Chemistry and Pharmacy, University of Sofia, 1 James Bourchier Blvd., Sofia 1164, Bulgaria
* Correspondence: vsimeonov@chem.uni-sofia.bg

Received: 5 September 2020; Accepted: 22 October 2020; Published: 24 October 2020

Abstract: The present study describes a simple procedure to separate into patterns of similarity a large group of solvents, 259 in total, presented by 15 specific descriptors (experimentally found and theoretically predicted physicochemical parameters). Solvent data is usually characterized by its high variability, different molecular symmetry, and spatial orientation. Methods of chemometrics can usefully be used to extract and explore accurately the information contained in such data. In this order, advanced fuzzy divisive hierarchical-clustering methods were efficiently applied in the present study of a large group of solvents using specific descriptors. The fuzzy divisive hierarchical associative-clustering algorithm provides not only a fuzzy partition of the solvents investigated, but also a fuzzy partition of descriptors considered. In this way, it is possible to identify the most specific descriptors (in terms of higher, smallest, or intermediate values) to each fuzzy partition (group) of solvents. Additionally, the partitioning performed could be interpreted with respect to the molecular symmetry. The chemometric approach used for this goal is fuzzy c-means method being a semi-supervised clustering procedure. The advantage of such a clustering process is the opportunity to achieve separation of the solvents into similarity patterns with a certain degree of membership of each solvent to a certain pattern, as well as to consider possible membership of the same object (solvent) in another cluster. Partitioning based on a hybrid approach of the theoretical molecular descriptors and experimentally obtained ones permits a more straightforward separation into groups of similarity and acceptable interpretation. It was shown that an important link between objects' groups of similarity and similarity groups of variables is achieved. Ten classes of solvents are interpreted depending on their specific descriptors, as one of the classes includes a single object and could be interpreted as an outlier. Setting the results of this research into broader perspective, it has been shown that the fuzzy clustering approach provides a useful tool for partitioning by the variables related to the main physicochemical properties of the solvents. It gets possible to offer a simple guide for solvents recognition based on theoretically calculated or experimentally found descriptors related to the physicochemical properties of the solvents.

Keywords: solvents; fuzzy hierarchical clustering; fuzzy associative-clustering; physicochemical descriptors

1. Introduction

The large number of different solvents used for many important chemical processes and technologies need special attention since their properties depend on a range of specific chemical and physical parameters such as melting and boiling point, water solubility, polarity, vapor pressure, density, viscosity, and even toxicity and many others.

Solvents can be separated by one of four basic methods: by solvent power (solubility polarity, acidity/basicity, properties/parameters), evaporation rate/boiling point, chemical structure, and hazard classification. Within the latter, this evaluation identifies both physical hazards (e.g., flash point, flammability, or reactivity) and toxicity, etc. The partitioning based on chemical structure groups used three groups: hydrocarbons, and oxygenated and chlorinated solvents [1–3].

Parker [1] divides them into: protic, aprotic, and inert according to the dipolarity of the solvent molecules and their ability to act as hydrogen bond donors. One disadvantage of a classification scheme such as this is that the groups are not restraining.

Partitioning of solvents based on physicochemical properties proved to be a significant and challenging problem [2–6]. Special interest provides a new study [7] where a new solvent similarity index is introduced, aiding in discovering the most suitable solvent for specific purposes. The solvent similarity index was calculated based on 261 pure solvents at 298 K, and classification was done for the solvents according to their solvation properties. Pushkarova et al. [8] used, as empirical characteristics of solvent-solute interactions via Taft-Kamlet-Abboud, polarity functions to determine the solvatochromic polarity. The practice of solvatochromic probing is growing rapidly but classification of media based on these values can be difficult. The paper focuses on the artificial neural networks (ANN) for the classification of solvent on the basis of their solvatochromic characteristics. Also, the influence of data variation on the stability of classification has been studied.

In the study of Gramatica et al. [9] a neuron nets approach was used for solvent separation. In general, many other chemometric methods contributed to proper solvent selection for practical needs like regression analysis, factor analysis, or partial least square regression [3–5].

Bradley et al. [10] used the Abraham general solvation model to predict the solvent coefficients for all organic solvents. The models were used to propose sustainable solvent replacements for commonly used solvents.

Recent efforts are concentrated on the application of chemometric strategies as suitable tools for classification of solvents (as objects of the analysis) characterized by many properly selected variables (chemical, structural, and physicochemical descriptors [11–15]. The majority of the methodologies are well developed and widely used for classification, interpretation, and modeling purposes like cluster analysis, principal components and factor analysis, artificial neural networks, partial least square regression, and discriminant analysis. A limited number of applications are related to fuzzy analysis [16,17].

Fuzzy clustering and partitioning also finds application in solvents characterization [18].

Fuzzy clustering analysis offers unique opportunities for decomposition of a large data set into a fixed number of similarity groups or clusters. Indeed, the classical cluster analysis (hierarchical or non-hierarchical) could achieve similar results but the strong advantage of the fuzzy partitioning strategy is the opportunity to locate a certain object (or variable) not to a single group of similarity but to calculate a function of membership for each object. Thus, a single object could be attributed to more than one cluster. This makes the interpretation efforts more loosely allowing considering specific distribution of objects into clusters with respective degree of membership. It eliminates ambiguity in interpretation or often unavoidable overlapping of clusters.

The major goal of the present study is to achieve a reliable partitioning of a large number of solvents with broad practical use by application of fuzzy partitioning methodology.

In this study, the fuzzy divisive hierarchical clustering and the powerful fuzzy divisive hierarchical associative-clustering method, which offer an excellent possibility to associate each fuzzy partition of samples to a fuzzy set of characteristics (descriptors), were successfully applied for the characterization

of 259 solvents, according to their 15 specific descriptors (experimentally found and theoretically predicted). What is quite new is the partitioning of solvents and their association with different descriptors with high, moderate, and low values. The obtained results clearly demonstrated the efficiency and information power of the advanced fuzzy clustering method in solvents characterization and clustering.

2. Materials and Methods

2.1. Fuzzy Clustering Methods

The application of fuzzy logic for various scientific and technical goals has been commented on for decades [19]. This approach differs from the classical hard clustering where each object of the data set finds its own cluster. Thus, an object either belongs to a defined cluster or is out of it. The application of Fuzzy theory to the problem of finding similarity between objects of interest leads to the conclusion that a particular object can belong simultaneously to more than one cluster, but with different degrees of membership (DOMs) between 0 and 1 [20,21]. In one of the possible approaches to so-called fuzzy c-means clustering (FCM), each cluster is replaced by a cluster prototype [22,23] with a respective center, which contains information about the size and the shape of the cluster. The degrees of membership are computed from the distances of the data point to the cluster centers. These distances are responsible for the value of DOM and determine the cluster properties and shape (point, line, etc.) [24].

There are different algorithms in fuzzy clustering applications, the most used being the binary divisive algorithm and the generalized fuzzy c-means algorithm (GFCM). The fuzzy methods briefly described above and the corresponding software were clearly described and efficiently applied in previous papers [25–30].

2.2. Data Set

The dataset consists of 269 solvents. Each solvent was described by 15 variables (molecular descriptors and experimentally obtain properties) shown below in Table 1.

Table 1. Molecular descriptors and experimentally obtained properties.

No.	Variable Name	Code
1	Melting point (experimental)	MPe
2	Melting point (calculated)	MPc
3	Boiling point (experimental)	BPe
4	Boiling point (calculated)	BPc
5	Density	Dens
6	Water solubility (experimental)	WSe
7	Water solubility (calculated)	WSc
8	Vapor pressure (experimental)	VPe
9	Vapor pressure (calculated)	VPc
10	Henry Law constant (experimental)	HLe
11	Henry Law constant (calculated)	HLc
12	Octanol/water coefficient (experimental)	logKOWe
13	Octanol/water (calculated)	logKOWc
14	Octanol/Air (calculated)	logKOAc
15	Bioconcentration Factor	log BCF

In the present study, the following set of subprograms implemented in the EPI Suite™ version 4.10 were used: MPBPWIN™, WATERNT™, HENRYWIN™, KOAWIN™, KOWWIN™, and BCFBAF™.

The melting point (MP), boiling point (BP), and vapor pressure (VP) within the MPBPWIN™ module in EPI Suite™ were applied to predict the properties of our interests. The MPBPWIN™ estimates melting point by the two methods: (1) the Joback Method (a group contribution method); (2) the Gold and Ogle method MP = 0.5839 * BP (in °K). Boiling point is valued by an adaptation of the Stein and Brown (1994) method, which is also a group contribution method. Vapor pressure is predictable as well by the methods: (1) Antoine, (2) Modified Grain method, and (3) the Mackay method. WATERNT™ estimates water solubility directly using a "fragment constant" method similar to that used in the KOWWIN™ program.

The Henry's law constant is estimated by the subprogram HENRYWIN™, which calculates (air/water partition coefficient) using both the group contribution and the bond contribution methods.

This KOAWIN™ program evaluates the logarithm of the octanol-air partition coefficient (KOA) of an organic compound with the compound's octanol–water partition coefficient (Kow) and Henry's law constant (HLC). For the KOAWIN only a chemical structure was needed for estimation of KOA. In the KOAWIN structures are implemented by the SMILES codes (Simplified Molecular Input Line Entry System). The KOA is possible to be predicted from the octanol–water partition coefficient (KOW) and Henry's law constant (H) by the subsequent equation:

$$KOA = KOW(RT)/H$$

where R is the ideal gas constant and T is the absolute temperature. KOA and KOW are unitless values. H/RT is the unit less Henry's law constant, also known as the air–water partition coefficient (KAW).

Therefore, the equation to estimate KOA is:

$$KOA = KOW/KAW$$

The KOWWIN™ program is for the octanol–water partition coefficient prediction. The basis of prediction in KOWWIN is a "fragment constant" methodology. In this "fragment constant" method, the starting structure is divided and then evaluated.

The comparison with the available experimental data shows a high level of correlation. In such a way, missing data in the large data set could be replaced.

3. Results and Discussion

3.1. Fuzzy Divisive Hierarchical Clustering of Descriptors

The fuzzy clustering of the variables (15 in total) aims to check the following:

- If the experimental values of the respective variables conform with the calculated one (i.e. if they fall within a fuzzy cluster with high membership function);
- If the partitioning procedure could determine stable groups of similarity between the variables with high DOM;
- The procedure is important for revealing information about possible descriptors for classification of the solvents in interest.

In the supplemental information section (Supplement T1) the fuzzy partitioning results for 15 variables are presented. In total, 28 groups are considered. The summary of the final partitioning is shown below:

A1—only HLc is included (a typical outlier)
A2—MPe MPc BPe BPc Dens WSe WSc VPe VPc HLe logKOWe LogKOWc logKOAc logBCF (the rest of the variables show a high level of similarity with a distinct difference from HLc).

In the next steps of fuzzy partitioning respective groups of similarity based on DOM will be sought.

A21—MPe MPc BPe BPc Dens VPe VPc HLe logKOWe LogKOWc logKOAc logBCF
A22—WSe WSc

In this partitioning stage, the experimentally found and theoretically calculated values of water solubility are extracted as a group of similarity different from the rest of variables in subgroup A21.

A211—BPe BPc
A212—MPe MPc Dens VPe VPc HLe logKOWe LogKOWc logKOAc logBCF
A2111—BPe
A2112—BPc
A2121—Dens

At this level of fuzzy partitioning, one finds separation between the experimentally and theoretically found values of boiling points and density. According to the DOM values the differences are small and the similarity between these three variables is significant,

A2122—MPe MPc VPe VPc HLe logKOWe LogKOWc logKOAc logBCF
A21221—VPe VPc HLe logKOWe LogKOWc logKOAc logBCF
A21222—MPe MPc
A212211—VPe VPc

The separation of two other groups of similarity is indicated melting point (experimental and theoretical values) and vapor pressure (experimental and theoretical values).

A212212—HLe logKOWe LogKOWc logKOAc logBCF
A2122111—VPc
A2122112—VPe
A2122121—logBCF

This stage of fuzzy partitioning reveals a slight difference between vapor pressure (theoretical and experimental values), and a more specific role of logBCF as compared to the stable group of logKOW (experimental and theoretical values) and logKOAc (calculated values).

A2122122—HLe logKOWe LogKOWc logKOAc
A21221221—HLe logKOWe LogKOWc
A21221222—logKOAc
A212212211—LogKOWc
A212212212—HLe logKOWe
A2122122121—logKOWe
A21221221212—HLe
A212221—MPe
A212222—MPc
A221—WSc
A222—Wse

The fuzzy partitioning carried out for 15 variables characterizing a set of solvents revealed the following fuzzy linkage of the variables:

- Very good coincidence between experimentally determined and theoretically calculated values of the variables characterizing the solvents; this means that if experimental values of some solvents are missing, calculation substitutes could be successfully used for classification and interpretation goals;

- HLc was defined as a typical outlier;
- The group of variables characterizing the distribution between different media (important for toxicity properties determination) is very compact;
- The parameters characterizing physicochemical properties (MP, BP, WS, and VP) indicate various type of similarity with the other parameters—water solubility is the most distant to the rest of parameters, followed by BP and MP; density is closest to BP; logBCF is slightly different as compared to the rest of "toxicity esteems."

Additional material could be found in Supplement

3.2. Fuzzy Divisive Hierarchical Clustering of Solvents

To compare the partitions, and the similarity and differences of the investigated solvents, we have to analyze both the characteristics of the prototypes corresponding to the partitions hierarchy obtained by applying fuzzy divisive hierarchical clustering and DOMs of solvents corresponding to all fuzzy partitions. The results presented in Table 2 clearly illustrate the most specific characteristics of each fuzzy partition and their similarity and differences.

The initial two clusters A1 and A2 indicate that one typical outlier is present in the list of solvents—perfluorooctane, whose properties are completely different from those of the other 268 solvents. The further divisive fuzzy clustering indicates the level of the membership function of each solvent into each of the next groups included (22 in total).

Next, Table 2 shows the final fuzzy partitioning with the prototypes of the partitions, ranked solvents for each group and the range of DOM.

Table 2. Final fuzzy partitioning.

Final Fuzzy Partitions	Prototypes of Fuzzy Partitions	Solvents (Ranked in Decreasing Order)	DOMs Range
	Parameters of Prototype		
A111111	−24.06; −54.87; 233.98; 203.55; 21.61; 8685.76; 1293.07; 1.39; 0.83; 0.00; 0.03 1.20; 1.94; 6.05; 0.98	19	0.6348
A111112	−14.24; 188.97; 186.54; 2.09; 18.27; 9376.14; 42.79; 2.64; 0.31; 0.00; 1.53; 0.75; 5.65; 0.58; −77.87	202, 259, 217, 162,	0.7181–0.4695
A11112	−65.36; −50.74; 131.11; 132.12; 43.49; 7923.34; 3573.41; 21.71; 20.30; 0.00; 0.02; 1.75; 1.88; 3.84; 0.39	102, 117	0.6730–0.2889
A11121	−60.21; −53.00; 142.95; 135.46; 42.30; 5004.94; 2542.52; 83.02; 78.23; 0.00; 0.03; 1.84; 1.84; 4.05; 0.59	54, 38, 52, 53, 9, 134	0.7262–0.2825
A11122	−71.27; −76.34; 107.00; 147.22; 122.42; 3417.29; 4425.60; 198.31; 70.84; 0.01; 0.03; 1.78; 1.53; 4.90; 0.33	118, 104, 57, 96	0.8062–0.8062
A11211	80.41; −23.42; 319.82; 112.29; 573.17; 4808.29; 3389.89; 6.06; 92.03; 0.00; 0.31; 1.33; 9.59; 0.80; −114.19	157, 223	0.9574–0.9491
A11212	−24.89; 76.99; 154.75; 8.68; 2427.15; 2994.14; 647.18; 53.07; 7.58; 0.00; 1.52; 1.23; 4.64; 0.93; −35.57	260, 188, 231	0.7190–0.2987
A11221	−58.62; 121.52; 135.12; 1.73; 7497.75; 6004.92; 136.19; 103.68; 0.21; 0.00; 1.44; 1.69; 3.73; 0.63; −78.32	233, 145, 179, 144	0.5978–0.3376
A11222	−47.51; 136.10; 138.25; 2.50; 6432.46; 3613.26; 122.45; 59.46; 0.38; 0.00; 1.70; 1.71; 3.92; 0.62; −42.74	200, 218, 234, 227	0.5407–0.2093

Table 2. *Cont.*

Final Fuzzy Partitions	Prototypes of Fuzzy Partitions Parameters of Prototype	Solvents (Ranked in Decreasing Order)	DOMs Range
A12111	−18.28; −10.51; 184.30; 150.65; 12.53; 31.27; 28.06; 20.28; 13.51; 0.19; 0.83; 1.56; 2.13; 4.38; −3.25	136, 129, 115, 55, 17, 110, 6, 67, 80, 49, 97, 112, 26, 109, 124, 86, 44, 23, 193, 68, 87, 135, 58, 45, 5, 75, 50, 37, 46, 33, 114, 133, 65, 29, 18, 149, 197, 254, 11, 238, 95, 84, 81, 88, 22, 100, 28, 77, 94, 56, 31, 121, 122, 32, 208, 41, 30, 1, 90, 240, 36, 111, 47, 89, 185, 216	0.9956–0.2976
A12112111	16.28; 89.97; 337.65; 355.25; 3.90; 97.01; 4.08; 1.86; 5.41; 0.02; 0.03; 6.00; 7.21; 8.30; 2.05	99 103 35 43 116 27 92 78 101 64 119 268	0.6005–0.0600
A12112112	10.77; 55.33; 305.37; 309.65; 2.75; 95.45; 9.19; 1.95; 1.25; 0.02; 0.03; 5.99; 6.11; 7.37; 2.08	123, 8, 24, 25, 71, 66, 105, 130	0.6368–0.1414
A121121211	18.65; 29.24; 273.01; 219.70; 8.12; 125.43; 5.96; 5.76; 0.99; 0.03; 0.02; 1.72; −0.08; 6.95; −0.20	48, 21, 152	0.6432–0.0937
A121121212	2.70; 30.76; 269.06; 258.53; 3.06; 123.06; 4.16; 2.48; 0.42; 0.03; 0.01; 5.05; 2.10; 7.25; 0.81	73, 120, 113, 2	0.7989–0.3787
A12112122	4.40; −7.63; 243.97; 239.15; 4.24; 134.37; 21.91; 3.45; 1.32; 0.03; 0.03; 3.57; 3.47; 5.96; 1.12	98, 126, 108, 61, 60, 93, 76	0.7381–0.1055
A121122	−14.32; 16.36; 184.44; 149.07; 117.55; 465.95; 353.89; 34.18; 24.34; 0.02; 0.40 2.88; 4.28; 3.77; −5.09	253, 13, 106, 170, 34, 125, 62, 10, 42, 132, 15, 40, 69	0.5034–0.0892
A1212111	18.33; 218.26; 232.00; 5.49; 20.77; 18.76; 16.13; 3.73; 4.80; 0.02; 1.99; 1.41; 6.44; 1.22; −48.90	248, 177, 261, 137, 148, 163, 225, 138, 158, 181, 265, 147, 207, 269, 205, 150, 256, 161, 171, 155, 143, 186, 235, 250, 239	0.5992–0.1036
A1212112	−19.64; 169.10; 172.41; 7.39; 15.19; 17.95; 15.96; 8.17; 3.94; 0.05; 1.51; 1.24; 5.20; 1.14; −52.63	201, 262, 212, 198, 139, 190, 183, 222, 173, 266, 214, 246	0.5090–0.1133
A1212121	−61.90; 89.25; 97.95; 11.00; 8.68; 14.72; 46.64; 47.82; 5.89; 0.07; 0.48; 0.66; 3.55; 0.87; −25.63	156, 264, 169, 224, 182, 154, 153, 219, 141, 164, 180, 210	0.5958–0.1890
A1212122	−45.95; 126.62; 125.31; 4.78; 6.59; 9.59; 14.37; 13.52; 1.99; 0.06; 0.66; 0.77; 4.30; 0.78; −29.46	213, 142, 196, 242, 236, 151, 192,175, 176, 178, 165, 191, 252, 251, 160, 263, 203, 189 226, 204, 241, 194, 249	0.7261–0.1019
A12122	−43.23; 50.28; 135.60; 63.69; 61.31; 128.10; 130.95; 86.49; 35.51; 0.11; 0.90; 1.75; 3.55; 2.46; −17.62	232, 159, 211, 195, 168, 187, 245, 172, 174, 237, 199, 228, 74,167, 85, 267, 51, 14, 230, 91, 39, 12, 3, 255, 63, 131, 79, 140, 257, 244, 127	0.9979–0.1005
A1221	−45.95; 126.62; 125.31; 4.78; 6.59; 9.59; 14.37; 13.52; 1.99; 0.06; 0.66; 0.77; 4.30; 0.78; −29.46	215, 258, 221, 243, 166, 146, 247, 206, 209, 220	0.6506–0.3753
A1222	−45.48; 37.33; 145.26; 205.85; 747.83; 1979.16; 1053.58; 52.32; 18.56; 0.01; 0.39; 2.23; 2.41; 6.05; −4.47	4, 82, 16, 70, 229, 7, 59, 184, 128, 83, 107, 72	0.9676–0.1448
A2	42.00; −58.65; 105.90; 102.10; 1.73; 3371.80; 0.00; 57.84; 33.90; 0.04; 2,450,000.00; 1.62; 6.17; −1.15; 3.74	20	1.0000

3.3. Fuzzy Divisive Hierarchical Associative-Clustering of Solvents and Descriptors

To compare the partitions, and the similarity and differences of solvents, we have to analyze the DOMs corresponding to all fuzzy partitions for both the samples and characteristics (descriptors). The results obtained by applying the fuzzy divisive hierarchical associative-clustering method using the descriptor data are presented in Table 3. By carefully analyzing the fuzzy partitions at each level

(partition history/hierarchy) in parallel with the descriptor considered data, the following remarks may be taken. The fuzzy partitioning of the solvents with indication of the descriptors related to each fuzzy partition (cluster) is depicted in Table 3.

Table 3. The fuzzy partitioning of the solvents and variables (descriptors).

Solvents	Variables	DOM Solvents	DOM Variables
A1 *244 15 10 257 125 132 106 42 140 4 62 220 40 82 34 267 79 131 268 209 92 27 255 230 74 239 167 155 63 170 90 71 76 205 269 130 8 199 25 111 185 105 250 89 216 13 207 159 228 150 101 119 2 69 214 120 152 95 246 11 78 81 240 249 21 137 93 18 116 221 22 181 98 258 253 208 172 174 235 126 245 99 51 148 85 265 114 103 36 87 35 261 24 47 161 43 64 229 45 66 68 147 49 194 41 211 123 138 238 256 186 97 188 225 254 158 171 67 109 26 237 124 143 5 180 23 112 195 44 37 160 183 222 113 80 262 210 139 198 73 169 33 166 135 252 251 197 29 65 176 168 133 88 50 46 86 48 75 164 190 189 60 212 175 108 187 248 61 58 151 213 193 192 149 156 236 153 17 201 178 263 163 196 84 14 100 191 16 142 129 154 6 94 39 173 3 177 264 242 55 224 115 241 182 266 110 165 91 121 122 1 56 32 219 141 30 31 77 28 204 226 203 12 59 136 232 260 70 243 127 215 7 247 206 118 134 157 184 223 146 9 96 231 128 83 52 104 227 53 54 38 72 107 162 202 217 218 259 19 234 200 102 57 233 179 144 145 117*	1–10 12–15	1.000	1.000–0.9957
A2 *20*	11	1.000	1.000
A11 *218 162 259 234 38 217 202 72 107 54 19 200 53 102 57 233 179 144 145 227 117 104 52 83 9 231 128 96 223 157 184*	5–7	0.9991–0.5167	0.8819–0.7767
A12 *111 13 199 152 214 76 246 89 150 228 18 174 172 245 239 253 159 170 114 47 255 211 41 256 194 137 74 195 237 180 66 23 210 238 230 181 186 49 68 161 126 254 171 45 97 98 143 169 124 87 138 235 109 26 67 261 158 112 5 208 225 147 37 265 44 93 22 85 240 160 222 148 80 164 21 123 139 198 262 36 183 33 64 197 252 251 135 43 65 81 24 187 35 60 103 29 113 46 176 73 133 75 50 86 189 248 88 11 167 190 51 168 212 249 95 99 175 108 48 213 61 78 58 120 151 156 193 149 192 2 116 153 267 236 201 178 163 263 90 17 196 84 100 216 14 191 207 154 142 129 94 6 264 119 177 39 173 3 242 224 101 185 55 130 241 115 182 105 266 165 110 91 121 25 122 250 1 56 32 219 30 141 31 8 77 28 204 226 82 203 12 71 269 205 136 155 232 4 27 92 63 34 268 106 131 62 79 140 257 125 42 10 244 132 15 209 40 220 258 69 221 229 166 16 188 70 59 243 260 127 215 206 247 7 118 146 134*	8 9 2 13 14 12 10 15 1 4	0.9999–0.5020	0.9996–0.8953
A111 *19 162 217 202 259 53 102 54 179 38 117 223 52 157 9*	6	0.9607–0.3961	0.8819
A112 *218 234 200 72 107 227 144 145 57 233 83 104 128 231 96 184*	5 7	0.9967–0.3022	0.9–0.3022
A1121 *72 107 57 83 128 104 96*	7	0.8–0.2718	0.7987
A1122 *218 234 200 227 144 145 233 231 184*	5	0.9–0.2123	0.7709
A121 *106 82 119 105 8 90 71 25 120 2 11 95 101 78 229 81 92 216 221 268 130 125 27 4 257 166 21 116 132 185 69 93 260 22 258 244 10 98 155 220 269 205 59 209 126 16 240 70 243 36 99 250 215 207 7 40 255 127 247 103 239 206 79 131 87 63 118 134 146*	13 14 12 10 9 15 1 8 2	0.9–0.3475	0.9–0.6356

Table 3. *Cont.*

Solvents	Variables	DOM Solvents	DOM Variables
A122 *253 88 212 48 228 190 198 61 262 47 252 251 159 197 256 17 248 222 175 263 151 201 213 139 6 108 29 73 236 193 129 60 210 176 100 94 178 113 55 183 196 142 173 58 149 242 115 156 254 163 133 191 195 110 3 192 241 264 39 14 169 224 122,121 164 266 194 1 33 165 160 238 158 167 177 32 37 30 56 154 170 211 189 31 225 219 182 153 168 204 187 5 180 141 91 84 50 65 226 203 12 138 152 171 80 214 28 77 237 86 230 135 41 246 46 174 124 111 89 136 44 232 143 245 42 109 172 75 147 140 66 26 114 23 112 186 123 150 161 97 18 261 15 265 62 13 67 34 199 148 85 68 137 51 45 24 249 49 235 43 188 74 208 76 64 35 267 181*	4 3	0.9–0.5186	0.9–0.8369
A1211 *258 269 205 155 257 215 209 250 207 247 243 166 40 7 255 206 63 131 79 146*	8 9 2	0.7318–0.2235	0.8666–0.5420
A1212 *8 105 90 71 25 120 11 82 95 119 2 106 4 130 125 21 78 116 69 132 81 92 27 98 101 93 10 260 59 22 126 99 70 16 216 127 240 36 103 185 87 268 244 229 221 118 220 134 239*	12 10 13 14 1 15	0.9382–0.2880	0.9–0.9379
A12121 *106 81 240 216 185 260 101 268*	1		
A12122 *90 120 25 8 82 2 119 105 4 130 78 95 116 71 125 11 27 69 132 98 93 10 126 59 21 99 22 127 16 70 103 36 229 118 221 92 87 134 244 220 239*	12 13 10 14 15	0.8–0.3352	0.9408
A121221 *119 11 95 93 22 21 98 10 126 87 221 244 220*	13 14 12	0.9–0.2867	0.9–0.8168
A121222 *25 8 82 105 2 120 130 78 71 90 116 27 132 4 69 99 59 103 16 70 36 127 125 92 118 239 134 229*	10 15	0.8–0.2226	0.9–0.7336
A1212221 *1 119 11 95 22 21 87*	14	0.7–0.3422	0.9207
A1212222 *2 98 10 126 93 221 220 244*	13 12	0.5–0.1733	0.9–0.5802
A1212221 *8 25 132 69 105 71 27 59 127 16 70 118 92 239 134*	15	0.8–0.1720	0.7326
A1212222 *2 120 78 4 82 130 90 116 99 36 103 125 229*	10	0.7–0.1169	0.8960
A1221 *48 61 17 47 73 60 108 113 6 129 55 115 42 140 110 111 89 66 15 62 123 34 37 13 114 88 18 5 267 26 24 188 50 109 80 97 124 112 43 135 35 64 76 49 68 74*	4	0.9–0.3307	0.9159
A1222 *252 251 263 175 212 151 213 190 139 176 173 183 222 228 198 236 242 256 266 241 160 178 196 142 62 253 210 201 156 192,191 264 169 194 226 248 165 219 224 189 163 246 177 225 193 197 149 138 203 158 214 180 170 254 154 141 164 153 182 238 147 159 33 195 204 211 30 94 84 152 237 121 167 91 122 65 171 12 32 3 39 31 58 14 100 28 1 168 265 56 187 143 150 261 77 249 230 161 186 29 148 136 232 137 41 174 245 75 172 86 235 85 208 181 51 67 199 45*	3	0.9–0.3509	0.8366

For the final goal of fuzzy partitioning of the objects (solvents) was performed by the use of 10 variables (only the experimentally found ones) (Table 4).

Table 4. Solvents included in each class and the respective class descriptor.

Class	*Solvents*	*Variables*
1	*9 19 38 52 53 54 102 117 157 162 179 202 217 223 259 57 72 83 96 104 107 128*	WSe
2	*144 145 184 200 218 227 231 233 234*	Dens
3	*7 40 63 79 131 146 155 166 205 206 207 209 215 243 247 250 255 257 258 269*	VPe
4	*81 101 106 185 216 240 260 268*	MPe
5	*11 21 22 87 95 119*	logKOA
6	*10 93 98 126 220 221 244*	logKOWe
7	*8 16 25 27 59 69 70 71 92 105 118 127 132 134 239*	logBCF
8	*2 4 36 78 82 90 99 103 116 120 125 130 229*	HLe
9	*5 6 13 15 17 18 23 24 26 34 35 37 42 43 44 46 47 48 49 50 55 60 61 62 64 66 68 73 74 76 80 88 89 97 108 109 110 111 112 113 114 115 123 124 129 133 135 140 188 267 1 3 12 14 28 29 30 31 32 33 39 41 45 51 56 58 65 67 75 77 84 85 86 91 94 100 121 122 136 137 138 139 141 142 143 147 148 149 150 151 152 153 154 156 158 159 160 161 163 164 165 167 168 169 170 171 172 173 174 175 176 177 178 180 181 182 183 186 187 189 190 191 192 193 194 195 196 197 198 199 201 203 204 208 210 211 212 213 214 219 222 224 225 226 228 230 232 235 236 237 238 241 242 245 246 248 249 251 252 253 254 256 261 262 263 264 265 266*	BPe
10	*20*	HLc

The fuzzy partitioning performed reveals the following classes of solvents:

Class 1 (WSe): *9 19 38 52 53 54 102 117 157 162 179 202 217 223 259 57 72 83 96 104 107 128*

Iodoethane Diethyl glutarate 1,1-dichloroethane Dimethyl phthalate quinoline 2,4-dimethyl-3-pentanone n-butyl acetate 1,2-dichloroethane Glycerol-1,3-Dibutyl ether m-cresol Dimethyl adipate Glycerol-1,2,3-triethyl ether diethyl carbonate Caprylic acid diethanolamide DMEU 1-hexanol 4-methyl-2-pentanone Butylacetate 1-chloropropane 2-chloropropane 2-Methyltetrahydrofuran Diethyl succinate.

Mainly chlorinated solvents and similar ones except for diethyl carbonate and buthylacetate (CHLORINATED SOLVENTS CLASS) with major descriptor WSe (experimental water solubility value; calculated water solubility gives the same separation).

Class 2 (Dens): *144 145 184 200 218 227 231 233 234*

Bromoethane Di-isopropyl ether Benzonitrile Acetophenone Isobutyl acetate Di-n-propyl ether Carbon disulfide Benzaldehyde Chloroform.

Nonpolar and volatile solvents except for isobutyl acetate (NON-POLAR AND VOLATILE SOLVENTS CLASS) major descriptor DENS (density).

Class 3 (VPe): *7 40 63 79 131 146 155 166 205 206 207 209 215 243 247 250 255 257 258 269*

Benzene 1,1,1-trichloroethane Dichloromethane Methyl formate Diethyl ether 1,1-dichloroethylene Oleyl alcohol Fluorobenzene Tetraethylene glycol Ricinoleic acid Triethylene glycol Butyl stearate Methyl benzoate 1,1,2,2-tetrachloroethane 1,8-Cineole gamma-Valerolactone Nopol alpha-Terpineol beta-Terpineol PolyEthyleneGlycol 200.

This is a mixture of polar solvents—acids and esters with non-polar ones such as benzene or 1,1,1-TCA (POLAR AND NON-POLAR SOLVENTS MIXED CLASS separated mainly by descriptor vapor pressure experimental values; the calculated value gives the same results).

Class 4 (MPe): *81 101 106 185 216 240 260 268*

Glycerol triacetate Oleic acid Menthanol Diisooctylsuccinat Dioctylsuccinate Isosorbide dioctanoate DPMU PolyEthyleneGlycol 600.

POLAR SOLVENTS CLASS I defined by descriptor melting point I.

Class 5 (logKOA): *11 21 22 87 95 119*

2-Pyrrolidone Sulfolane Propylene carbonate N-methylacetamide Glycerol Water

POLAR SOLVENTS CLASS II defined by descriptor logKOA.

Class 6 (logKOWe): *10 93 98 126 220 221 244*

Phenetole Diisobutyl adipate Geranyl acetate Menthanyl acetate Trichloroethylene Pentyl acetate 1-Octanol.

The grouping was defined properly as a polar solvent except trichloroethylene POLAR SOLVENTS CLASS III defined by logKOW (experimental or theoretical).

Class 7 (logBCF): *8 16 25 27 59 69 70 71 92 105 118 127 132 134 239*

Ethyl myristate Isoamyle acetate Butyl laurate Methyl abietate 2,6-dimethyl-4-heptanone N,N-dimethylaniline Nitrobenzene Benzyl benzoate Methyl stearate Methyl myristate Dimethyl 2-methylglutarate 1,1,3,3-tetramethyl urea Diethyl phthalate Diethyl adipate Glycerol-1,2,3-tributyl ether.

The group of polar ones, except for nitrobenzene—POLAR SOLVENTS CLASS IV defined by logBCF.

Class 8 (HLe): *2 4 36 78 82 90 99 103 116 120 125 130 229*

Ethyl laurate Glycerol-1,2-dibutyl ether Acetyltributyl citrate Diisoamylsuccinate N,N-Diethylolcapramide Dibenzyl ether Butyl palmitate Methyl linolenate Methyl ricinoleate Methyl laurate Ethyl benzoate Dibutyl sebacate Anisole.

The group defined by a HIGH MOLECULAR WEIGHT POLAR SOLVENTS defined by HLe (experimental).

Class 9 (BPe): *5 6 13 15 17 18 23 24 26 34 35 37 42 43 44 46 47 48 49 50 55 60 61 62 64 66 68 73 74 76 80 88 89 97 108 109 110 111 112 113 114 115 123 124 129 133 135 140 188 267*

1 3 12 14 28 29 30 31 32 33 39 41 45 51 56 58 65 67 75 77 84 85 86 91 94 100 121 122 136 137 138 139 141 142 143 147 148 149 150 151 152 153 154 156 158 159 160 161 163 164 165 167 168 169 170 171 172 173 174 175 176 177 178 180 181 182 183 186 187 189 190 191 192 193 194 195 196 197 198 199 201 203 204 208 210 211 212 213 214 219 222

224 225 226 228 230 232 235 236 237 238 241 242 245 246 248 249 251 252 253 254 256 261 262 263 264 265 266

This group is quite large. Most of the solvents are polar except for: carbon tetrachloride, xylenes, and bromobenzene.

Cyclohexanol Isododecane Di-n-butyl acetate 1-chlorobutane Glycerol-2-methyl monoether m-dichlorobenzene p-Cymene Methyl palmitate Isopropylacetate chlorobenzene Isopropyl palmitate 2,6-dimethylpyridine 1-bromobutane Butyl myristate Furfurylic alcohol 1-2,4-dimethylpyridine Dihydromyrcenol 3-Hydroxypropionic acid Benzyl alcohol Cyclohexanone 1,3-Dioxan-5-ol Diisobutyl succinate Glycerol-2-ethyl monoether Toluen Methyl Linoleate.

N,N-Dimethyldecanamide N-methylformamide Cyclopentane Propylene glycol Iodobenzene Glycerol-1,3-dimethyl ether Piperidine o-xylene Aniline Diisobutyl glutarate Tetrahydrofurfurylic alcohol 3-Methoxy-3-methyl-1-butanol p-xylene cis-decaline Dimethylisosorbide mesitylene Glycerol-1,2-dimethyl ether Isopropyl myristate d-Limonene 1,3-Dioxolane-4-methanol Propionic acid

N-decane Carbon tetrachloride Cyclopentyl methyl ether N-pentane Triethylamine Propyl formate Ethanol Ethyl acetate 1-Butanol 4-picoline 3-methyl-2-butanone n-Propyl acetate propionitrile Dimethyl sulfoxide 1,3-Dioxolane Cyclohexane Formamide Diethylamine Iso-octane Glycerol-1,2,3-trimethyl ether Dimethyl succinate 1,3-Propanediol Propylene glycol Butyronitrile N,N-dimethylformamide Ethyl formate β-Pinene 2,2,2-trifluoroethanol 3-pentanone Pyridine 2-pentanone n-heptane 3-Butyl-1-methylimidazolium tetrafluoroborate 1-Decanol N-methyl-pyrrolidin-2-one α-Pinene 1,2-dimethoxyethane 2-methoxyethanol Methyl oleate Decamethylcyclopentasiloxane Diethylene glycol Glycerol-2-butyl monoether Tributylamine 1-pentanol EthylHexyllactate Nitromethane.

Tert-butyl alcohol 1,4-dioxane Glycerol-1-ethyl monoether Cyclohexene N,N-dimethylacetamide Ethyl palmitate 5-(Hydroxymethyl)furfural 2-butanone 2-methyl-2-butanol styrene Methyl acetate Pyrrolidine N,N-Dimethyloctanamide Glycerol carbonate Acetone 2-aminoethanol tert-butyl methyl ether Acetylacetone 3-picoline Dipropyleneglycol.

2-pentanol n-butylamine Diphenyl ether 2-propanol Ethylene glycol Ethyl linolenate Methanol Cyclopentyl methyl ether Nitroethane Phenol Isobutyl alcohol Ethylenediamine.

β-Farnesen Tetrachloroethylene Tetrahydrofuran 3-pentanol Methyl 5-(dimethylamino) 2-methyl-oxopentanoate 2,4,6-trimethylpyridine Glycerol-1,3-diethyl ether 2-butanol Acetic anhydride Ethyl linoleate trifluoroacetic acid n-hexane Ethyl lactate Cyclopentanon o-dichlorobenzene 3,3-dimethyl-2-butanone Dimethyl glutarate 1-propanol Glycerol-1-methyl monoether n-octane m-xylene Bromobenzene Choline acetate Ethyl oleate Acetic acid.

Acetonitrile Glycerol-1,2,3-tributyl ether morpholine 3-methyl-1-butanol Acetone 1,4-Cineole Terpineol acetate 2-Furfuraldehyde beta-Myrcene Terpinolene Cyclademol Glycofurol (n = 2) Solketal HMPTA DEGDEE DEGDME Ethyl propionate TEGDME Dimetylsulfoxide.

Class 10 (HLc): *Outlier Perfluorooctane 20*

The solvents underlined above do not strictly belong to logical formation of similarity classes and seem more to be rather odd than reasonable as members of the respective class (polar, non-polar, or volatile solvents determined by specific variables). A careful check of the position of these 12 solvents into the fuzzy partitioning groups indicates that all of them have quite low maximal value of DOM as determined by fuzzy analysis (this values is shown next to the name of the solvent).

The few exceptions found (only 9 out of 259 solvents), namely:

(diethyl carbonate, benzene, nitrobenzene);(o-, m-, p-xylene) and (carbon tetrachloride, bromobenzene, trichloroethylene), are resultant to their low maximal DOM, so their position into one group of similarity is not stable and they could be considered either as members of the group with low probability, or members of a different class.

In Table 5 summarized results according to obtained classes are presented.

The table could be used as a practical guide for selection of type of solvents based on their physicochemical properties.

Table 5. Defined classes of solvents with the descriptors.

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
Iodoethane 9	1			x		**Wsol**
Diethyl glutarate 19	1	x				**Wsol**
1,1-dichloroethane 38	1			x		**Wsol**
2,4-dimethyl-3-pentanone 54	1			x		**Wsol**
1-hexanol 57	1		x			**Wsol**
4-methyl-2-pentanone 72	1			x		**Wsol**
1-chloropropane 96	1			x		**Wsol**
n-butyl acetate 102	1				x	**Wsol**
2-chloropropane 104	1			x		**Wsol**
2-Methyltetrahydrofuran 107	1				x	**Wsol**
1,2-dichloroethane 117	1			x		**Wsol**
Diethyl succinate 128	1	x				**Wsol**
Glycerol-1,3-Dibutyl ether 157	1	x				**Wsol**
m-cresol 162	1		x			**Wsol**
Dimethyl adipate 179	1	x				**Wsol**
Glycerol-1,2,3-triethyl ether 202	1	x				**Wsol**
diethyl carbonate 217	1				x	**Wsol**
Caprylic acid diethanolamide 223	1	x				**Wsol**
Benzaldehyde 233	1		x			**Wsol**
Chloroform 234	1			x		**Wsol**
DMEU 259	1	x				**Wsol**
Dimethyl phthalate 52	1		x			**Wsol**
Quinoline 53	1		x			**Wsol**
Butylacetate 83	1		x			**Wsol**
Bromoethane 144	2			x		**Dens**
di-isopropyl ether 145	2			x		**Dens**
Benzonitrile 184	2		x			**Dens**
Isobutyl acetate 218	2				x	**Dens**
di-n-propyl ether 227	2			x		**Dens**
carbon disulfide 231	2			x		**Dens**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
Benzene 7	3			x		**VPe**
1,1,1-trichloroethane 40	3			x		**VPe**
Dichloromethane 63	3			x		**VPe**
methyl formate 79	3				x	**VPe**
diethyl ether 131	3			x		**VPe**
1,1-dichloroethylene 146	3			x		**VPe**
Oleyl alcohol 155	3		x			**VPe**
Fluorobenzene 166	3			x		**VPe**
tetraethylene glycol 205	3	x				**VPe**
Ricinoleic acid 206	3		x			**VPe**
triethylene glycol 207	3	x				**VPe**
Butyl stearate 209	3		x			**VPe**
methyl benzoate 215	3		x			**VPe**
1,8-Cineole 247	3		x			**VPe**
gamma-Valerolactone 250	3		x			**VPe**
alpha-Terpineol 257	3		x			**VPe**
beta-Terpineol 258	3		x			**VPe**
1,1,2,2-tetrachloroethane 243	3			x		**VPe**
Nopol 255	3		x			**VPe**
PolyEthyleneGlycol 200 269	3	x				**VPe**
Dioctylsuccinate 216	4		x			**MPe**
Isosorbide dioctanoate 240	4		x			**MPe**
Glycerol triacetate 81	4	x				**MPe**
Oleic acid 101	4		x			**MPe**
Menthanol 106	4		x			**MPe**
Diisooctylsuccinate 185	4		x			**MPe**
DPMU 260	4	x				**MPe**
PolyEthyleneGlycol 600 268	4	x				**MPe**
2-Pyrrolidone 11	5		x			**logKOA**
Sulfolane 21	5	x				**logKOA**
Propylene carbonate 22	5	x				**logKOA**
N-methylacetamide 87	5	x				**logKOA**
Glycerol 95	5	x				**logKOA**
Water 119	5				x	**logKOA**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
Phenetole 10	6		x			**logKOW**
Diisobutyl adipate 93	6		x			**logKOW**
Geranyl acetate 98	6		x			**logKOW**
1-Octanol 244	6		x			**logKOW**
Menthanyl acetate 126	6		x			**logKOW**
Trichloroethylene 220	6			x		**logKOW**
pentyl acetate 221	6		x			**logKOW**
Ethyl myristate 8	7	x				**logBCF**
Isoamyle acetate 16	7				x	**logBCF**
Methyl abietate 27	7		x			**logBCF**
Methyl myristate 105	7		x			**logBCF**
Butyl laurate 25	7		x			**logBCF**
2,6-dimethyl-4-heptanone 59	7		x			**logBCF**
N,N-dimethylaniline 69	7		x			**logBCF**
Nitrobenzene 70	7		x			**logBCF**
Benzyl benzoate 71	7		x			**logBCF**
Methyl stearate 92	7		x			**logBCF**
Geraniol 239	7		x			**logBCF**
Dimethyl 2-methylglutarate 118	7	x				**logBCF**
1,1,3,3-tetramethyl urea 127	7	x				**logBCF**
Diethyl phthalate 132	7		x			**logBCF**
Diethyl adipate 134	7		x			**logBCF**
Ethyl laurate 2	8	x				**HLe**
Glycerol-1,2-dibutyl ether 4	8	x				**HLe**
dibenzyl ether 90	8		x			**HLe**
Methyl linolenate 103	8		x			**HLe**
Methyl ricinoleate 116	8		x			**HLe**
Methyl laurate 120	8		x			**HLe**
ethyl benzoate 125	8		x			**HLe**
Dibutyl sebacate 130	8		x			**HLe**
Acetyltributyl citrate 36	8		x			**HLe**
Diisoamylsuccinate 78	8		x			**HLe**
N,N-Diethylolcapramid 82	8		x			**HLe**
Butyl palmitate 99	8		x			**HLe**
Anisole 229	8		x			**HLe**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
Triethylamine 1	9				x	**BPe**
propyl formate 3	9				x	**BPe**
Cyclohexanol 5	9	x				**BPe**
Ethanol 12	9				x	**BPe**
di-n-butyl acetate 13	9				x	**BPe**
Ethyl acetate 14	9				x	**BPe**
1-chlorobutane 15	9			x		**BPe**
Glycerol-2-methyl monoether 17	9	x				**BPe**
m-dichlorobenzene 18	9		x			**BPe**
p-Cymene 23	9		x			**BPe**
Methyl palmitate 24	9		x			**BPe**
Isopropylacetate 26	9	x				**BPe**
1-Butanol 28	9				x	**BPe**
4-picoline 29	9				x	**BPe**
3-methyl-2-butanone 30	9				x	**BPe**
n-Propyl acetate 31	9				x	**BPe**
Propionitrile 32	9				x	**BPe**
Dimethyl sulfoxide 33	9	x				**BPe**
Chlorobenzene 34	9			x		**BPe**
Isopropyl palmitate 35	9		x			**BPe**
2,6-dimethylpyridine 37	9	x				**BPe**
1,3-Dioxolane 39	9				x	**BPe**
Cyclohexane 41	9			x		**BPe**
1-bromobutane 42	9			x		**BPe**
Butyl myristate 43	9		x			**BPe**
Furfurylic alcohol 44	9	x				**BPe**
Formamide 45	9	x				**BPe**
2,4-dimethylpyridine 46	9		x			**BPe**
Dihydromyrcenol 47	9		x			**BPe**
3-Hydroxypropionic acid 48	9	x				**BPe**
Benzyl alcohol 49	9		x			**BPe**
Cyclohexanone 50	9		x			**BPe**
Diethylamine 51	9				x	**BPe**
1,3-Dioxan-5-ol 55	9		x			**BPe**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
iso-octane 56	9			x		**BPe**
Glycerol-1,2,3-trimethyl ether 58	9	x				**BPe**
Glycerol-2-ethyl monoether 61	9	x				**BPe**
Toluene 62	9			x		**BPe**
1,3-Propanediol 67	9	x				**BPe**
N-methylformamide 68	9	x				**BPe**
Glycerol-1-butyl monoether 73	9	x				**BPe**
Cyclopentane 74	9			x		**BPe**
Propylene glycol 75	9	x				**BPe**
Iodobenzene 76	9		x			**BPe**
Butyronitrile 77	9				x	**BPe**
Glycerol-1,3-dimethyl ether 80	9	x				**BPe**
N,N-dimethylformamide 84	9	x				**BPe**
ethyl formate 85	9				x	**BPe**
β-Pinene 86	9		x			**BPe**
Piperidine 88	9				x	**BPe**
o-xylene 89	9			x		**BPe**
2,2,2-trifluoroethanol 91	9				x	**BPe**
3-pentanone 94	9				x	**BPe**
Aniline 97	9	x				**BPe**
Pyridine 100	9				x	**BPe**
Diisobutyl glutarate 108	9		x			**BPe**
Tetrahydrofurfurylic alcohol 109	9	x				**BPe**
3-Methoxy-3-methyl-1-butanol 110	9	x				**BPe**
p-xylene 111	9			x		**BPe**
cis-decaline 112	9		x			**BPe**
Dimethylisosorbide113	9	x				**BPe**
Mesitylene 114	9		x			**BPe**
Glycerol-1,2-dimethyl ether 115	9	x				**BPe**
2-pentanone 121	9				x	**BPe**
n-heptane 122	9			x		**BPe**
Isopropyl myristate 123	9		x			**BPe**
d-Limonene 124	9		x			**BPe**
1,3-Dioxolane-4-methanol 129	9	x				**BPe**
Propionic acid 133	9				x	**BPe**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
n-decane 135	9		x			**BPe**
3-Butyl-1-methylimidazolium tetrafluoroborate 136	9		x			**BPe**
1-Decanol 137	9		x			**BPe**
N-methyl-pyrrolidin-2-one 138	9	x				**BPe**
α-Pinene 139	9		x			**BPe**
carbon tetrachloride 140	9			x		**BPe**
1,2-dimethoxyethane 141	9				x	**BPe**
2-methoxyethanol 142	9				x	**BPe**
Methyl oleate 143	9		x			**BPe**
Decamethylcyclopentasiloxane 147	9		x			**BPe**
diethylene glycol 148	9	x				**BPe**
Glycerol-2-butyl monoether 149	9	x				**BPe**
Tributylamine 150	9		x			**BPe**
1-pentanol 151	9	x				**BPe**
EthylHexyllactate 152	9		x			**BPe**
Nitromethane 153	9				x	**BPe**
tert-butyl alcohol 154	9				x	**BPe**
1,4-dioxane 156	9				x	**BPe**
Glycerol-1-ethyl monoether 158	9	x				**BPe**
Cyclohexene 159	9			x		**BPe**
N,N-dimethylacetamide 160	9	x				**BPe**
Ethyl palmitate 161	9		x			**BPe**
5-(Hydroxymethyl)furfural 163	9		x			**BPe**
2-butanone 164	9			x		**BPe**
2-methyl-2-butanol 165	9				x	**BPe**
Styrene 167	9		x			**BPe**
Methyl acetate 168	9				x	**BPe**
Pyrrolidine 169	9				x	**BPe**
N,N-Dimethyloctanamide 170	9		x			**BPe**
Glycerol carbonate 171	9	x				**BPe**
Acetone 172	9				x	**BPe**
2-aminoethanol 173	9	x				**BPe**
tert-butyl methyl ether 174	9				x	**BPe**
Acetylacetone 175	9	x				**BPe**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
3-picoline 176	9	x				**BPe**
Dipropyleneglycol 177	9	x				**BPe**
2-pentanol 178	9				x	**BPe**
n-butylamine 180	9				x	**BPe**
diphenyl ether 181	9		x			**BPe**
2-propanol 182	9				x	**BPe**
Ethylene glycol 183	9	x				**BPe**
Ethyl linolenate 186	9		x			**BPe**
Methanol 187	9				x	**BPe**
Cyclopentyl methyl ether 188	9				x	**BPe**
Nitroethane 189	9				x	**BPe**
Phenol 190	9	x				**BPe**
isobutyl alcohol 191	9				x	**BPe**
Ethylenediamine 192	9				x	**BPe**
β-Farnesen 193	9		x			**BPe**
Tetrachloroethylene 194	9			x		**BPe**
Tetrahydrofuran 195	9				x	**BPe**
3-pentanol 196	9				x	**BPe**
Methyl 5-(dimethylamino) 2-methyl-oxopentanoate 197	9	x				**BPe**
2,4,6-trimethylpyridine 198	9		x			**BPe**
n-propylamine 199	9				x	**BPe**
Acetophenone 200	9				x	**BPe**
Glycerol-1,3-diethyl ether 201	9	x				**BPe**
2-butanol 203	9				x	**BPe**
acetic anhydride 204	9	x				**BPe**
Ethyl linoleate 208	9		x			**BPe**
trifluoroacetic acid 210	9				x	**BPe**
n-hexane 211	9			x		**BPe**
Ethyl lactate 212	9	x				**BPe**
Cyclopentanone 213	9	x				**BPe**
o-dichlorobenzene 214	9		x			**BPe**
3,3-dimethyl-2-butanone 219	9				x	**BPe**
Dimethyl glutarate 222	9	x				**BPe**
1-propanol 224	9				x	**BPe**

Table 5. *Cont.*

Name	Class	Polar	Non-Polar	Volatile	Polar-Volatile	Descriptor
Glycerol-1-methyl monoether 225	9	x				**BPe**
n-octane 226	9			x		**BPe**
m-xylene 228	9			x		**BPe**
Bromobenzene 230	9		x			**BPe**
Choline acetate 232	9		x			**BPe**
Ethyl oleate 235	9		x			**BPe**
Acetic acid 236	9				x	**BPe**
Acetonitrile 237	9				x	**BPe**
Morpholine 241	9				x	**BPe**
3-methyl-1-butanol 242	9				x	**BPe**
Acetone 245	9				x	**BPe**
2-Furfuraldehyde 249	9	x				**BPe**
Glycofurol (n = 2) 254	9	x				**BPe**
Solketal 256	9	x				**BPe**
HMPTA 261	9	x				**BPe**
DEGDEE 262	9	x				**BPe**
DEGDME 263	9	x				**BPe**
ethyl propionate 264	9				x	**BPe**
TEGDME 265	9	x				**BPe**
Dimetylsulfoxide 266	9	x				**BPe**
n-pentane 267	9			x		**BPe**
Isododecane 6	9		x			**BPe**
Diisobutyl succinate 60	9		x			**BPe**
Methyl Linoleate 64	9		x			**BPe**
Dimethyl succinate 65	9		x			**BPe**
N,N-Dimethyldecanamide 66	9		x			**BPe**
Glycerol-1,2,3-tributyl ether 238	9		x			**BPe**
1,4-Cineole 246	9		x			**BPe**
Terpineol acetate 248	9		x			**BPe**
beta-Myrcene 251	9		x			**BPe**
Terpinolene 252	9		x			**BPe**
Cyclademol 253	9		x			**BPe**

4. Conclusions

The fuzzy hierarchical clustering of a large group of solvents into 10 classes of similarity made it possible to find patterns of the chemicals with specific properties divided by important descriptors. The fuzzy partitioning method applied helped in finding relationships between solvents of various nature (polar, non-polar, volatile etc.) and the physicochemical variables used. Additionally, the chemometric analysis has proven that if there are missing data of specific descriptors the theoretical calculation of them is possible with very high level of approximate to the experimentally observed and established physicochemical indicators.

Thus, the present study offers a simple methodological approach to the complex problem of solvent partitioning.

In order to understand the similarity and differences of various solvents, fuzzy divisive hierarchical clustering and fuzzy divisive hierarchical associative-clustering were successfully applied. The fuzzy partition hierarchy of solvents and descriptors associated allowed identifying partitions (groups) of solvents with more or less similar characteristics in terms of higher, smallest, or intermediate values of considered descriptors.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-8994/12/11/1763/s1.

Author Contributions: Conceptualization, M.N.; methodology, C.S.; M.N.; software C.S.; M.N.; validation, M.N., V.S. formal analysis, M.N., M.T., C.S., V.S.; investigation, M.N.; resources, M.N.; data curation, C.S., M.N.; writing—original draft preparation, M.N., C.S., M.T., V.S.; writing—review and editing, M.N., V.S.; visualization, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Information and Communication Technologies for a Single Digital Market in Science, Education and Security" of the Scientific Research Center, grant number NIS-3317 and National roadmaps for research infrastructures (RIs) grant number [NIS-3318]. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Acknowledgments: The author M.N. is grateful for the additional support by the project "Information and Communication Technologies for a Single Digital Market in Science, Education and Security" of the Scientific Research Center, NIS-3317 and National roadmaps for research infrastructures (RIs) grant number [NIS-3318].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Parker, A. Protic-dipolar aprotic solvent effects on rates of bimolecular reactions. *Chem. Rev.* **1967**, *69*, 1–32. [CrossRef]
2. Tobiszewski, M.; Nedyalkova, M.; Madurga, M.; Pena-Pereira, F.; Namieśnik, J.; Simeonov, V. Pre-selection and assessment of green organic solvents by clustering chemometric tools. *Ecotoxicol. Environ. Saf.* **2018**, *147*, 292–298. [CrossRef] [PubMed]
3. Katritzky, A.; Fara, D.; Kuanar, M.; Hur, E.; Karelson, M. The Classification of Solvents by Combining Classical QSPR Methodology with Principal Component Analysis. *J. Phys. Chem. A.* **2005**, *109*, 10323–10341. [CrossRef] [PubMed]
4. Molnar, M.; Komar, M.; Brahmbhatt, H.; Babić, J.; Jokić, S.; Rastija, V. Deep Eutectic Solvents as Convenient Media for Synthesis of Novel Coumarinyl Schiff Bases and Their QSAR Studies. *Molecules* **2017**, *22*, 1482. [CrossRef]
5. Chastrette, M.; Rajzmann, M.; Chanon, M.; Purcell, K. Approach to a general classification of solvents using a multivariate statistical treatment of quantitative solvent parameters. *J. Am. Chem. Soc.* **1985**, *107*, 1–11. [CrossRef]
6. Laurence, C.; Legros, J.; Chantzis, A.; Planchat, A.; Jacquemin, D. A Database of Dispersion-Induction DI, Electrostatic ES, and Hydrogen Bonding α1 and β1 Solvent Parameters and Some Applications to the Multiparameter Correlation Analysis of Solvent Effects. *J. Phys. Chem. B* **2015**, *119*, 3174–3184. [CrossRef]
7. Driver, M.; Hunter, C. Solvent similarity index. *PCCP* **2020**, *22*, 11967–11975. [CrossRef]
8. Pushkarova, Y.; Kholin, Y. The classification of solvents based on solvatochromic characteristics: The choice of optimal parameters for artificial neural networks. *Cent. Eur. J. Chem.* **2010**, *10*, 1318–1327. [CrossRef]

9. Sahigara, F.; Ballabio, D.; Todeschini, R.; Consonni, V. Defining a novel k-nearest neighbors approach to assess the applicability domain of a QSAR model for reliable predictions. *J. Cheminform.* **2013**, *5*, 27. [CrossRef]
10. Bradley, J.-C.; Abraham, M.H.; Acree, W.E.; Lang, A.S.I.D. Predicting Abraham model solvent coefficients. *Chem. Cent. J.* **2015**, *12*, 2–10. [CrossRef]
11. Johnson, A.R.; Vitha, M.F. Chromatographic SelectivityTriangles. *J. Chromatogr. A* **2011**, *1218*, 556–586. [CrossRef] [PubMed]
12. Katritzky, A.R.; Tamm, T.; Wang, Y.; Sild, S.; Karelson, M. A Unified Treatment of Solvent Properties. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 692–698. [CrossRef]
13. Poole, C.F.; Karunasekara, T. Solvent Classification for Chromatography and Extraction. *J. Planar Chromatogr.* **2012**, *25*, 190–199. [CrossRef]
14. Lesellier, E. Rpider Diagram: A Universal and Versatile Approach for System Comparison and Classification: Application to Solvent Properties. *J. Chromatogr. A* **2015**, *1389*, 49–64. [CrossRef]
15. Wypych, G. *Handbook of Solvents*, 2nd ed.; Chem Tec Publishing: Toronto, ON, Canada, 2001.
16. Sarbu, C.; Pop, H.F. Fuzzy Soft-Computing Methods and Their Applications in Chemistry. In *Reviews in Computational Chemistry*; Lipkowitz, K.B., Larter, R., Cundari, T.R., Eds.; Wiley-VCH: Hoboken, NJ, USA, 2004; Chapter 5, pp. 249–332.
17. Halgamuge, S.K.; Wang, L. (Eds.) *Classification and Clustering for Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2005.
18. Guidea, A.; Sarbu, C. Fuzzy characterization and classification of solvents according to their polarity and selectivity. A comparison with the Snyder approach. *J. Liq. Chromatogr. Relat. Technol.* **2020**, *43*, 336–343. [CrossRef]
19. Kaufman, L.; Rousseeuw, P.J. *Finding Groups in Data: An Introduction to Cluster Analysis*; John Wiley & Sons: New York, NY, USA, 2009.
20. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]
21. Rouvray, D. *Fuzzy Logic in Chemistry*; Academic Press: San Diego, CA, USA, 1997; p. 364.
22. Sarbu, C.; Pop, H.F. Fuzzy Soft-Computing Methods and Their Applications in Chemistry. *Rev. Comput. Chem.* **2004**, *20*, 249–332.
23. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Press: New York, NY, USA, 1987; p. 272.
24. Hoppner, F.; Klawonn, R.K.; Kruse, R.; Runkler, T. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*; John Wiley &Sons: Chichester, UK, 1999; p. 300.
25. Sârbu, C.; Zehl, K.; Einax, J.W. Fuzzy divisive hierarchical clustering of soil data using gustafson-kessel algorithm. *Chemom. Intell. Lab. Syst.* **2007**, *87*, 121–129. [CrossRef]
26. Sârbu, C.; Moţ, A.C. Ecosystem discrimination and fingerprinting of Romanian propolis by hierarchical fuzzy clustering and image analysis of TLC patterns. *Talanta* **2011**, *85*, 1112–1117. [CrossRef]
27. Pop, H.; Dumitrescu, D.; Sârbu, C. A Study of Roman Pottery (terra sigillata) Using Hierarchical Fuzzy Clustering. *Anal. Chim. Acta* **1995**, *310*, 269–279. [CrossRef]
28. Pop, H.; Sârbu, C. The fuzzy hierarchical cross-clustering algorithm. Improvements and comparative study. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 510–516. [CrossRef]
29. Dumitrescu, D.; Pop, H.; Sârbu, C. Fuzzy Hierarchical Cross-Classification of Greek Muds. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 851–857. [CrossRef]
30. Sârbu, C.; Horovitz, O.; Pop, H. A Fuzzy Cross-Classification of Chemical Elements, Based on Their Physical, Chemical and Structural Features. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1098–1108. [CrossRef]