applied sciences

MDPI

Conference Report

Modeling Properties with Artificial Neural Networks and Multilinear Least-Squares Regression: Advantages and Drawbacks of the Two Methods [†]

Jesus Vicente de Julián-Ortiz [1,*], Lionello Pogliani [1] and Emili Besalú [2]

[1] Molecular Topology and Drug Design Research Unit, Departament de Química Física, Facultat de Farmàcia, Universitat de València, 46100 Burjassot, Spain; liopo@uv.es

[2] Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, 17003 Girona, Spain; emili.besalu@udg.edu

[*] Correspondence: jejuor@uv.es; Tel.: +34-963-543-279

[†] A preliminary version of this article appeared in: de Julián-Ortiz, J.; Pogliani, L.; Besalú, E. Artificial Neural Networks and Multilinear Least Squares to Model Physicochemical Properties of Organic Solvents. In *Proceedings of the MOL2NET, International Conference on Multidisciplinary Sciences*, 25 December 2016–25 January 2017; Sciforum Electronic Conference Series, Vol. 2, 2016; doi:10.3390/mol2net-02-03826.

Received: 24 April 2018; Accepted: 29 June 2018; Published: 5 July 2018

Abstract: The mean molecular connectivity indices (MMCI) proposed in previous studies are used in conjunction with well-known molecular connectivity indices (MCI) to model eleven properties of organic solvents. The MMCI and MCI descriptors selected by the stepwise multilinear least-squares (MLS) procedure were used to perform artificial neural network (ANN) computations, with the aim of detecting the advantages and limits of the ANN approach. The MLS procedure can replicate the obtained results for as long as is needed, a characteristic not shared by the ANN methodology, which, on the one hand increases the quality of a description, and on the other hand also results in overfitting. The present study also reveals how ANN methods prefer MCI relatively to MMCI descriptors. Four types of ANN computations show that: (i) MMCI descriptors are preferred with properties with a small number of points, (ii) MLS is preferred over ANN when the number of ANN weights is similar to the number of regression coefficients and, (iii) in some cases, the MLS modeling quality is similar to the modeling quality of ANN computations. Both the common training set and an external randomly chosen validation set were used throughout the paper.

Keywords: physicochemical properties; QSPR; topological descriptors; MLS; artificial neural networks

1. Introduction

Recently [1], the mean molecular connectivity indices (MMCI) were introduced to model eleven properties of organic solvents. The multilinear least-squares (MLS) used to derive the quantitative structure-property relationships (QSPR) showed that three out of eleven properties, the refractive index (RI), the flash points (FP), and the ultraviolet cutoff values (UV), were modeled with the MMCI while the remaining properties were modeled with the well-known molecular connectivity indices (MCI). The MMCI indices are also centered on the basic concepts of the delta, valence delta, I- and S-indices that go back to the origins of the molecular connectivity theory [2–7]. Results from two other recent studies that used semiempirical sets of descriptors [8,9] showed that the artificial neural network (ANN) model with a variable number of hidden neurons chosen by the software improves the quality of a QSPR obtained with the aid of the multilinear least-squares (MLS) methodology, also known as multilinear regression (MLR). Nevertheless, this improvement is somewhat artificial as the ANN computations for the eleven properties employed a number of weights, due to the presence of more than one hidden neuron, much greater than the number of weights or regression coefficients in the MLS procedure. This fact can provide poor results when new data are to be predicted. This is called overfitting, and it can be avoided by guiding the training process after the predictions in a test set, by more general regularization techniques, or by dropout of the hidden neurons.

A scheme of the work is depicted in Figure 1. Data consisting of eleven physicochemical properties of solvents were randomly split into train (TR) and evaluation (EV) sets. Molecular descriptors were calculated, as explained in section Materials and Methods, for every molecule. MLS computations performed with the train set ended up choosing the best descriptors among the set of given descriptors. These best descriptors were used to perform the Multilayer Perceptron ANN (ANN-MLP) computations. To avoid overfitting, during its training process the ANN randomly selects test sets (TE) within the original TR set. Finally, the models obtained by each method are applied, for external validation, to the evaluation (EV) set. It should be underlined that the evaluation set is common to every type of computation.

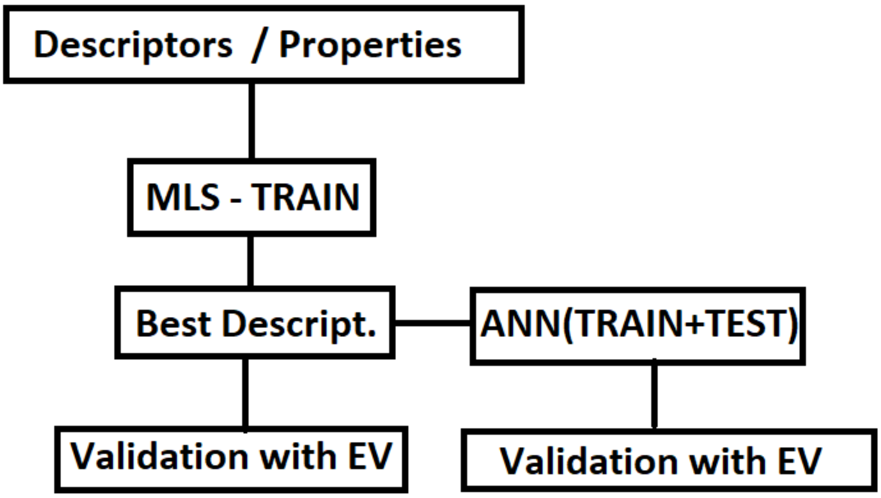


Figure 1. Flow chart of the methodology followed throughout the present work.

The aim of the present work is to pin down the real advantages and the drawbacks of the ANN methodology, and apply it to the model of the eleven properties of [1] where either MCI or MMCI are used as the descriptors. Four different types of ANN computations are here performed to detect the level of achieved improvement, if any, (a) with one hidden neuron, (b) with a pre-fixed number of hidden neurons, (c) with a variable number of hidden neurons chosen by the software, and (d) with a minor number of descriptors for the one hidden neuron case. This last case attempted to render the number of ANN weights equal to the number of MLS weights. It also monitored if ANN computations preferred either MCIs or MMCIs for modeling purposes. The descriptors for the eleven properties are those of [1]; however, whenever a property was not satisfactorily modeled by the given MCI (or MMCI) the second or third best MCI (or MMCI) was chosen. The domain of applicability of the models presented here includes substances that have been used as solvents without any other chemical restrictions.

2. Materials and Methods

2.1. The Properties

The raw material of the present study, the eleven properties of the organic solvents, is given in Table 1. The source for their values is cited in [1].

Table 1. Eleven properties of organic solvents with their molar mass M (g·mol^{-1}): T_b, boiling point (K); ε, dielectric constant; d, density (at 20 °C ± 5 °C relative to water at 4 °C, g/cc); RI, refractive index (20 °C); FP, Flashpoint (K); η, viscosity (Cpoise, 20 °C; [1] at 25 °C, [2] at 15 °C); γ, surface tension (mN/m at 25 °C); UV, Cutoff UV values (nm); μ, dipole moments in debye (1D = 10^{-18} esu cm = 3.3356 × 10^{-3} C m); MS, magnetic susceptibility (also, $-\chi \cdot 10^6$, in emu mol^{-1}, 1 emu = 1 cm^3, temperatures cover a range from 15 °C to 32 °C); and El, Elutropic value (silica).

Solvents	M	T_b	ε	d	RI	FP	η	γ	UV	μ	MS	El
(°) Acetone	58.1	329	20.7	0.791	1.359	256	0.32	23.46	330	2.88	0.46	0.43
(°) Acetonitrile	41.05	355	37.5	0.786	1.344	278	0.37	28.66	190	3.92	0.534	0.50
Benzene	78.1	353	2.3	0.84	1.501	262	**0.65**	28.22	280	0	0.699	0.27
Benzonitrile	103.1	461	25.2	1.010	1.528	344	1.24 [1]	38.79				
1-Butanol	74.1	391	17.1	0.810	1.399	308	2.95	**24.93**	215			
(°) 2-Butanone	72.1	353	18.5	0.805	1.379	270	0.40	23.97	330			0.39
Butyl Acetate	116.2	398	5.0	0.882	1.394	**295**	**0.73**	**24.88**	254			
CS_2	76.1	319	2.6	1.266	1.627	**240**	0.37	**31.58**	380	0	0.532	
CCl_4	153.8	350	2.2	1.594	1.460		**0.97**	**26.43**	263	**0**	0.691	0.14
Cl-Benzene	112.6	405	5.6	1.107	1.524	296	0.80	32.99	**287**			
1-Cl-Butane	92.6	351	7.4	0.886	1.4024	267	0.35	23.18	225			
$CHCl_3$	119.4	334	4.8	1.492	1.446		0.57	26.67	245	1.01	0.740	**0.31**
Cyclohexane	84.2	354	2.0	0.779	1.426	255	1.00	24.65	**200**	**0**	0.627	0.03
(°) Cyclopentane	70.1	323	2.0	0.751	1.400	236	0.47	21.88	200		0.629	
1,2-diCl-Benzene	147.0	453	9.9	1.306	1.551	**338**	1.32		**295**	2.50	0.748	
1,2-diCl-Ethane	98.95	356	10.4	1.256	**1.444**	288	0.79	31.86	225	1.75		
diCl-Methane	84.9	313	9.1	1.325	1.424		0.44	27.20	235	1.60	**0.733**	0.32
N,*N*-diMe-Acetamide	87.1	**438**	**37.8**	**0.937**	**1.438**	343			268	3.8		
N,*N*-diMeFormamide	73.1	**426**	**36.7**	**0.944**	**1.431**	**330**	0.92		268	3.86		
1,4-Dioxane	88.1	374	2.2	1.034	1.422	285	**1.54**	**32.75**	215	0.45	0.606	
Ether	74.1	308	4.3	0.708	1.353	**233**	0.24	16.95	**215**	1.15		0.29
Ethyl acetate	88.1	350	6.0	0.902	1.372	270	0.45	23.39	260	1.8	0.554	**0.45**
(°) Ethyl alcohol	46.1	351	24.3	0.785	1.360	281	1.20	21.97	210	1.69	0.575	
Heptane	100.2	371	1.9	0.684	1.387	272		19.65	200			0.00
Hexane	86.2	**342**	**1.9**	**0.659**	**1.375**	**250**	0.33	17.89	200			0.00
2-Methoxyethanol	76.1	398	16.0	0.965	1.402	319	**1.72**	30.84	220			
(°) Methyl alcohol	32.0	338	32.7	0.791	1.329	284	0.60	22.07	205	1.70	0.530	0.73
4-Me-2-Pentanone	100.2	391	13.1	0.800	1.396	286			334			
2-Me-1-Propanol	74.1	381	17.7	0.803	1.396	310						
2-Me-2-Propanol	74.1	356	10.9	0.786	1.387	277		19.96		**1.66**	**0.534**	
DMSO	78.1	462	46.7	1.101	1.479	368	2.24	**42.92**	268	3.96		
(°) Nitromethane	61.0	374	35.9	1.127	1.382	308	0.67	36.53	380	3.46	0.391	
1-Octanol	130.2	469	10.3	0.827	1.429	354	10.6 [2]	27.10				
(°) Pentane	72.15	309	1.8	0.626	1.358	224	0.23	15.49	200			0.00 *
3-Pentanone	86.1	375	17.0	0.853	1.392	279		24.74				
(°) 1-Propanol	60.1	370	20.1	0.804	1.384	288	2.26	23.32	210			
(°) 2-Propanol	60.1	356	18.3	0.785	1.377	295	2.30	20.93	210			0.63
Pyridine	79.1	388	12.3	0.978	1.510	293	0.94	36.56	305	2.2	0.611	0.55
Tetra Cl-Ethylene	165.8	394	2.3	1.623	1.506		0.90				**0.802**	
(°) Tetra-Hydrofuran	72.1	340	7.6	0.886	1.407	256	0.55		215	1.75		0.35 *
Toluene	92.1	**384**	**2.4**	**0.867**	**1.496**	277	0.59	27.93	285	0.36	0.618	0.22
1,1,2-triCl,triF-Ethane	187.4	**321**	**2.4**	**1.575**	**1.358**		**0.69**		**230**			**0.02**
2,2,4-triMe-Pentane	114.2	372	1.9	0.692	1.391	266	0.50		215			0.01
o-Xylene	106.2	**417**	**2.6**	**0.870**	**1.505**	**305**	0.81	29.76				
p-Xylene	106.2	411	2.3	0.866	1.495	300	0.65	28.01				
(°) Acetic acid	60.05	391	6.15	1.049	1.372			27.10		1.2	0.551	
Decalin	138.2	**465**	**2.2**	**0.879**	**1.476**						0.681	
diBr-Methane	173.8	370	7.8	1.542	2.497			**39.05**		**1.43**	0.935	
1,2-diCl-Ethylen(Z)	96.9	334	9.2	1.284	1.449					1.90	0.679	
(°) 1,2-diCl-Ethylen(E)	96.9	321	2.1	1.255	1.446					0	0.638	
1,1-diCl-Ethylen	96.9	**305**	**4.7**	**1.213**	**1.425**					1.34	0.635	

Table 1. *Cont.*

Solvents	M	T_b	ε	d	RI	FP	η	γ	UV	μ	MS	El
Dimethoxymethane	76.1	315	2.7	0.866	1.356						0.611	
(°) Dimethylether	46.1	249	5.0									
Ethylen Carbonate	88.1	511	89.6	1.321	1.425					4.91		
(°) Formamide	45.0	484	109	1.133	1.448			57.03		3.73	0.551	
(°) Methylchloride	50.5	249	12.6	0.916	1.339					1.87		
Morpholine	87.1	402	7.3	1.005	1.457						0.631	
Quinoline	129.2	510	9.0	1.098	1.629			42.59		2.2	0.729	
(°) SO_2	64.1	263	17.6	1.434						1.6		
2,2-tetraCl-Ethane	167.8	**419**	**8.2**	**1.578**	**1.487**			35.58		1.3	0.856	
tetraMe-Urea	116.2	450	23.1	0.969	1.449					**3.47**	0.634	
triCl-Ethylen	131.4	360	3.4	1.476	1.480						**0.734**	

(°) externally validated compounds; bold values: test compounds used in Artificial Neural Network Multilayer Perceptron (ANN-MLP) calculations, * for this property these two compounds ∈ {TR} (see Table 4 below) and {TR + TE} (see Table 5 below), Me = Methyl.

2.2. Descriptors

Table 2 shows the molecular connectivity χ indices, the molecular pseudoconnectivity ψ indices (pseudo-MCI), and the dual connectivity and pseudoconnectivity indices (Dual MCI, pseudo-MCI) used throughout this study. Three new indices were used: $\Delta = \Sigma_{EA} n_{EA}$, $\Sigma = \Sigma_{EA} <S_{EA}>$, and $T_{\Sigma/M} = \Sigma^3/M^{1.7}$ (M = molar mass); Δ encodes the number of electronegative atoms (n_{EA}), Σ encodes the sum of the S-State index for the electronegative atoms, N, O, F, Cl, Br ($<S_{EA}>$ is the average value for a specific atom). Table 3 shows the definitions of the MMCI (the first M stands for "mean"), which are based on averages of vertex invariants. The original Stolarsky's mean has a minus in the denominator here replaced by a plus to avoid zeroing the denominator due to equal δ_i and δ_j, although it is known that the limit of this function when δ_i tends to δ_j is finite. The present mean is a kind of pseudo-Stolarsky mean.

Table 2. Definition of the Molecular Connectivity Indices (MCI). Replacing δ with δ^v and I with S the corresponding valence, χ^v, I-State, ψ_I, and E-State, ψ_E, MCIs are obtained.

MCI	Pseudo-MCI	Dual MCI + Δ + Σ	Dual Pseudo-MCI + $T_{\Sigma/M}$
$D = \Sigma_i \delta_i$	${}^S\psi_I = \Sigma_i I$	${}^0\chi_d = (-0.5)^n \Pi_i(\delta_i)$	${}^0\psi_{Id} = (-0.5)^n \Pi_i(I_i)$
${}^0\chi = \Sigma(\delta_i)^{-0.5}$	${}^0\psi_I = \Sigma(I_i)^{-0.5}$	${}^1\chi_d = (-0.5)^{(n+\mu-1)} \Pi(\delta_i + \delta_j)$	${}^1\psi_{Id} = (-0.5)^{(n+\mu-1)} \Pi(I_i + I_j)$
${}^1\chi = \Sigma(\delta_i\delta_j)^{-0.5}$	${}^1\psi_I = \Sigma(I_iI_j)^{-0.5}$	${}^1\chi_s = \Pi(\delta_i + \delta_j)^{-0.5}$	${}^1\psi_{Is} = \Pi(I_i + I_j)^{-0.5}$
$\chi_t = (\Pi\delta_i)^{-0.5}$	${}^T\psi_I = (\Pi I_i)^{-0.5}$	$\Delta = \Sigma_{EA} n_{EA}$, $\Sigma = \Sigma_{EA} <S_{EA}>$	$T_{\Sigma/M} = \Sigma^3/M^{1.7}$

n is the number of atoms, ij means corresponds to σ bond, μ is the cyclomatic number.

Table 3. Definition of the Mean Molecular Connectivity Indices (MMCI). Replacing δ with δ^v, I, and with S the respective valence (M^v), I-State (M_I), and E-State (M_E) MMCIs are obtained.

${}^AM = \Sigma_i\ \delta_i/n$	${}^GM = \Sigma_{ij}(\delta_i\delta_j)^{1/2}$	${}^HM = 2\Sigma_{ij}\ (\delta_i{}^{-1} + \delta_j{}^{-1})^{-1}$
${}^RM = \Sigma_{ij}[(\delta_i{}^2 + \delta_j{}^2)/2]^{1/2}$	${}^SM = \Sigma_{ij}\ (\delta_i{}^2 + \delta_j{}^2)/(\delta_i + \delta_j)$	${}^UM = \Sigma_{ij}[\delta_i - \delta_j + (\delta_i{}^2 - 2\delta_i\delta_j + 5\delta_j{}^2)^{0.5}]/2$
${}^{Ho}M = \Sigma_{ij}(\delta_i{}^p + \delta_j{}^p)^{1/p}/2$	${}^LM = \Sigma_{ij}(\delta_i{}^p + \delta_j{}^p)/(\delta_i{}^{p-1} + \delta_j{}^{p-1})$	${}^{St}M = \Sigma_{ij}[(\delta_i{}^p - \delta_j{}^p)/(p\delta_i + p\delta_j)]^{1/(p-1)}$

A: arithmetic; G: geometric; H: harmonic; R: root mean square; S: symmetric; U: unsymmetric; Ho: Hölder; L: Lehmer; St: pseudo-Stolarsky.

These two tables summarize the pool of descriptors used throughout this study: n is the number of atoms in a molecule, i = 1 to n denotes the atoms of a molecule, ij denotes directly σ-bonded atoms, while p is assigned the value n in Table 3. Replacing δ with the valence delta, δ^v, in Table 2, allows the corresponding valence MCI, $\{D^v, {}^0\chi^v, {}^1\chi^v, \chi^v{}_t, {}^0\chi_d{}^v, {}^1\chi_d{}^v, {}^1\chi_s{}^v\}$, to be obtained; replacing the Intrinsic-I-State with the Electrotopological S-State index the corresponding pseudoconnectivity electrotopological indices are obtained, $\{^{S}\psi_E$, $^{0}\psi_E$, $^{1}\psi_E$, $^{T}\psi_E$, $^{0}\psi_{Ed}$, $^{1}\psi_{Ed}$, $^{1}\psi_{Es}\}$ [3–9]. This subject is further elucidated in the Appendixs A and B. Replacing in Table 3 δ, with δ^v, I and S three other subsets of MMCI: the valence, $\{^{A}M^v$, $^{G}M^v$, $^{H}M^v$, $^{R}M^v$, $^{S}M^v$, $^{U}M^v$, $^{Ho}M^v$, $^{L}M^v$, $^{St}M^v\}$, the I-State, $\{^{A}M_I$, $^{G}M_I$, $^{H}M_I$, $^{R}M_I$, $^{S}M_I$, $^{U}M_I$, $^{Ho}M_I$, $^{L}M_I$, $^{St}M_I\}$, and the E-State $\{^{A}M_E$, $^{G}M_E$, $^{H}M_E$, $^{R}M_E$, $^{S}M_E$, $^{U}M_E$, $^{Ho}M_E$, $^{L}M_E$, $^{St}M_E\}$ MMCI, respectively, are obtained. Because some S values can be negative (highly electropositive atoms) to avoid imaginary S-State MMCI values, a rescaling of the S value is undertaken as it is explained in [1]. Summing up, we have thirty-one MCI and thirty-six MMCI. Every index was obtained with a visual basic home-made program that runs on a normal PC that uses both adjacency and distance matrices [6].

2.3. Multilinear Least-Squares Regression

The stepwise multilinear least-squares (MLS) procedure of Statistica 8 that searches the whole combinatorial space built by the descriptors was used to find the best set of indices, either MCI or MMCI, for the training compounds of Table 1. They were then used to evaluate the left-out compounds (EV, those with (°) in Table 1, ~30% of all compounds, 25% for El). These best descriptors were also used for the ANN computations. To model the dipole moments, indices were multiplied by a two-valued symmetry factor, $\phi = 0, 1$, i.e., $\phi\cdot$[MCI or MMCI] = 0 or $\phi\cdot$[MCI or MMCI] = [MCI or MMCI], where zero is used for the symmetric molecules with $\mu = 0$. The choice for the number of indices of a relationship was performed bearing in mind that the ratio of data points to the number of variables should be higher or equal to five and should provide a correlation coefficient $r > 0.84$, i.e., $r^2 > 0.70$ [10]. External validation was performed for all types of model (ANN inclusive) with the set of evaluation points (EV) by adding them to check the prediction ability of the overall model. Broadly speaking, the models show robustness when 30–25% cases (the EV set) are advantageously added to complete the model.

2.4. Multi-Layer Perceptron—Artificial Neural Networks

ANN methods [11,12] that can perform regression and data validation carry out both tasks in a non-parametric way that makes no assumption regarding the relationship between y and x, where $y = f(x)$. This means that the function Property = f(indices) is not known a priori. In short, a non-parametric model is a kind of black box that tries to discover the mathematical function that can approximate the relationship between the indices and the property well enough. It uses highly flexible transfer functions with adaptable parameters that can model a wide spectrum of functional relationships. The activation functions for both hidden and output nodes used in Statistica 8 are: identity (i), logistic sigmoid (l), hyperbolic tangent (t), sine (s), and exponential (e).

ANN results were obtained with the built-in utility of Statistica 8—the multilayer perceptron neural network (MLP-ANN). This network has three-layered feedforward architecture with unidirectional full connections between successive layers (Figure 2) and error backpropagation (or backprop). The three layers are:

$$\text{input units} \rightarrow \text{hidden units} \rightarrow \text{output units}$$

Units are also known as neurons or nodes, in our case input units correspond to our variables, i.e.,

$$\text{variables (MCI or MMCI)} \rightarrow \text{hidden units} \rightarrow \text{P}$$

The only output unit, here, is the targeted property, P. In the present study the number of variables corresponds to the number of MCI or MMCI descriptors. Each neuron, or node, in a layer connects to every neuron in the next layer. The connections between neurons are the weights that determine the values assigned to the nodes. There exist additional weights assigned to the bias values that act as node value offsets; therefore, the resulting number of weights is:

$$(\text{No. input nodes} + 2)\cdot(\text{No. hidden nodes}) + 1$$

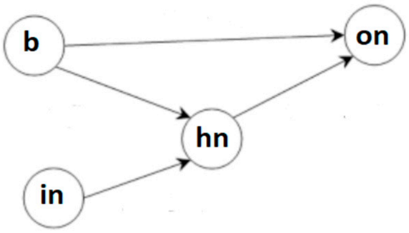


Figure 2. An ANN scheme with an input node (in), a bias node (b), a hidden node (hn), and an output node (on).

The given ANN scheme let us notice that if a weight is added to a hidden node the connections become seven. With five input nodes and seven hidden nodes [a 5-7-1 network] the weights become fifty. The weights adjusted by the training process are initially random and are handed over to all nodes of the following layer. The training process is iterative, and each iteration is called an epoch. Technically, the number of epochs is not definitive and it cannot be held as an unfailing parameter (it can exceed the given number). The weights are slightly varied in each epoch to minimize the sum-of-squares error function: SOS = $\Sigma_{i=1-N}$ $(P_{iclc} - P_i)^2$, where P_{iclc} (clc = calculated) is the ith predicted value (network outputs) of the property, and P_i is the target value. This function is the sum of differences between the prediction outputs and the target defined over the entire training set of points (compounds) N. Statistica 8 allows setting the number of networks to train and retain (Ntr/Nre). Two sets of values are here imposed: Ntr/Nre = 10^3/200 and Ntr/Nre = 10^5/200. In the corresponding tables only Ntr is shown as Nre is constant. The ANN network of Statistica 8 is optimized with the BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm to ensure a fast convergence rate [13,14].

Statistica 8, as a rule, sets by default the number of hidden nodes between 3 and 11. Nevertheless, as already told, we perform four procedures (for the 4th procedure see later on): (i) first a single hidden node, then (ii) hidden nodes from two to twelve are sequentially tried 'by hand' (i.e., the program is not allowed to change the imposed number of hidden nodes), and, finally, (iii) the program chooses the number of hidden nodes. To come as close as possible to the MLS results, it was decided (iv), to compute again the one hidden neuron case where either one or two indices with the lowest sensibility value have been deleted. In this case, for instance, the number of weights for the 4-1-1 case of T_b is 7, and it equals the number of correlation coefficients from the MLS calculations with six indices. Data required no normalization by the user, since the program performs this automatically.

Since the MLS procedure optimizes a number of regression parameters equal to the number of variables plus one (the bias parameter), a practical comparison between the two methods should only be performed when ANN uses no hidden neurons. In this case, the number of ANN weights equals the number of MLS parameters. One should expect that with a growing number of hidden neurons, the model of a property should constantly improve due to the growing number of weights for each variable (akin having a variable with many different weights). With ANN it is usually the case that the model becomes exceedingly good with a growing number of weights, and this frequently results in overfitting with exceedingly poor prediction for the external values. The choice of training (TR = 80% of the values in Table 1, excluding the externally validated compounds) and test sets (TE = 20% of the values, the bold values in this Table) usually prevents overfitting. In fact, the network is repeatedly trained for a number of cycles so long as the test error is on the decrease, otherwise the training is halted. This method, known as the 'early stopping' procedure [12], avoids the trap that the program will always choose the maximum number of hidden nodes. Each property shows an optimal number of nodes, which rarely corresponds with its maximum number.

3. Results

The results of the five procedures, one MLS and four ANN, are shown in Tables 4–9. Table 4 collects the MLS results for the eleven properties. In this table, in parenthesis the errors of the regression coefficients are given in vector form (±signs have been omitted, 2nd line of each cell, 2nd column). The training set for the elutropic value (El) includes pentane and tetrahydrofuran.

Tables 5–8 collect the different ANN-MLP results for the set of variables (descriptors, either MMCI or MCI) of Table 4. In these tables, the first column gives the δ^v type (see Appendix A), and the number of networks to train, Ntr = 10^3 or 10^5 (when both numbers gave rise to similar results Ntr = 10^3 was preferred), while the number of networks to retain is always 200. The activation functions together with the neuronal architecture are in the second column of Tables 5–8. In this column, 3rd line, the number of epochs for which the ANN-MLP calculation runs are shown for each property. In the third column is the set of variables together with their statistics. In this column, second line, are shown the sensitivities. These values come from the sensitivity analysis that quantifies the importance of the input variables of the models. The r^2 and s, statistics were obtained with the EXCEL spreadsheet plotting the observed property, P, vs. the calculated one, P_{clc}, once for the training and test compounds, N(aTR + bTE), and the second time for the training + test + evaluated compounds, N(+cEV), where a, b, and c are the number of points (i.e., compounds). We remind the reader that the MLS procedure has no test compounds, only training compounds, N(TR). No ANN weights are shown, due to their exceeding number, and because every time an ANN-MLP runs, different weights and sensitivity values are obtained.

For comparison purposes it was decided to maintain throughout the ANN calculations (see Tables 5–8) the same number of outliers excluded throughout the MLS procedure, where the exclusion was done for residuals greater than 3s. Clearly, the ANN outliers differ from the MLS ones. In Table 5, the ANN results obtained with a single hidden neuron are given. Tables 6 and 7 display the multiple neuron cases: Table 6 with an externally imposed number of hidden neurons that was cycled from 2 to 12, and Table 7 with the number of hidden neurons chosen by the program (between 3 and 11). For UV, MS, and El the program sets this number between 3 and 10. For those cases where different sets of hidden nodes achieve similar modeling, the set with the minimal number of nodes was preferred. The subset of descriptors used to model the properties showed r intercorrelation lower than 0.93. We remind the reader that in a previous study [15] it was established that indices can be considered strongly correlated if r > 0.98.

Table 4. Best set of descriptors for the properties of Table 1 with the multilinear least-squares (MLS) methodology. 1st column: δ^v type for the valence-dependent indices. 2nd column: set of descriptors and their statistical quality.

δ^v-Type	Regression Equations
$\delta^v{}_{po}(1)$	$T_b = 237.5 + 139.1\ {}^0\chi + 24.69\ D^v + 527.7\ {}^0\psi_I - 25.91\ {}^1\psi_I - 1500\ {}^0\psi_E + 41.53\ T_{\Sigma/M}$ (1) (24, 31, 3.5, 69, 21, 222, 10) N(TR) = 45, r^2 = 0.821, s = 22; N(+16EV) = 61, r^2 = 0.792, s = 25 Excluded strong outliers: Formamide & $SO_2 \in$ {EV}
$\delta^v{}_{po}(50)$	$\varepsilon = 2.804 - 12.05\ \chi^v{}_t - 5.99\cdot10^{-5}\ {}^1\chi^v{}_d + 132.7\ {}^1\chi^v{}_s + 0.021\ {}^1\psi_{Id} - 421.2\ {}^1\psi_{Es} + 38.12\ T_{\Sigma/M}$ (2) (0.9, 4.4, 10^{-5}, 28, 0.005, 124, 2.9) N(TR) = 43, r^2 = 0.858, s = 4.2; N(+16EV) = 59, r^2 = 0.896, s = 5.5 Excluded strong outliers: ethylencarbonate & quinoline ϵ {TR}, and MeOH & MeCl ∈ {EV}.

Table 4. *Cont.*

δ^{v}-Type	Regression Equations	
$\delta^{v}{}_{ppo}(-0.5)$	$d = 0.733 + 0.024\,D^{v} + 0.211\,{}^{0}\chi^{v} + 1.463\,{}^{1}\chi^{v}{}_{s} - 0.022\,{}^{S}\psi_{E} + 0.148\,\Delta$ (0.06, 0.002, 0.02, 0.3, 0.002, 0.01) N(TR) = 45, r^{2} = 0.939, s = 0.07; N(+15EV) = 60, r^{2} = 0.914, s = 0.08 Excluded outliers: MeCl & MeOH ∈ {EV}	(3)
$\delta^{v}{}_{ppo}(1)$	$RI = 1.573 - 0.156\,{}^{H}M + 0.617\,{}^{R}M + 0.067\,{}^{R}M^{v} - 0.447\,{}^{S}M - 0.086\,{}^{Ho}M^{v} - 0.012\,{}^{S}M_{E}$ (0.03, 0.01, 0.02, 0.01, 0.02, 0.01, 0.02) N(TR) = 45, r^{2} = 0.957, s = 0.04; N(+14EV) = 59, r^{2} = 0.951, s = 0.03 Excluded outliers: MeCl & MeOH ∈ {EV}	(4)
$\delta^{v}{}_{po}(-0.5)$	$\gamma = 8.683 + 0.386\,D^{v} + 397.6\,{}^{1}\chi^{v}{}_{s} + 151.9\,{}^{T}\psi_{I} - 502.4\,{}^{1}\psi_{Is} + 3.347\,\Delta$ (2.3, 0.05, 57, 36, 90, 0.7) N(TR) = 29, r^{2} = 0.835, s = 3.1; N(+10EV) = 39, r^{2} = 0.792, s = 3.1 Excluded outlier: formamide, nitromethane ∈ {EV}	(5)
$\delta^{v}{}_{ppo}(0.5)$	$FP = 387.1 + 26.99\,{}^{H}M - 94.38\,{}^{H}M_{I} + 33.03\,{}^{G}M_{E} + 114.5\,{}^{U}M_{I} - 83.10\,{}^{Ho}M_{E}$ (26, 6.2, 12, 5.2, 13, 11) N(TR) = 29, r^{2} = 0.829, s = 16; N(+11EV)= 40, r^{2} = 0.764, s = 17 Excluded outliers: Acetone ∈ {EV}	(6)
$\delta^{v}{}_{po}(-0.5)$	$\eta = -\ 0.216 + 0.001\,{}^{1}\chi_{d} + 0.486\,{}^{1}\psi_{I} + 2.20\cdot10^{-5}\,{}^{1}\psi_{Id} - 3.83\cdot10^{-6}\,{}^{0}\psi_{Ed} + 0.098\,\Sigma$ (0.2, 0.0003, 0.1, $7\cdot10^{-6}$, 10^{-7}, 0.01) N(TR) = 28, r^{2} = 0.969, s = 0.4; N(+10EV) = 38, r^{2} = 0.939, s = 0.4 Excluded outlier: MeOH ∈ {EV}	(7)
$\delta^{v}{}_{po}(5)$ $\phi = 0, 1$	$\mu = 0.038 + 0.002\,{}^{1}\chi_{d} - 0.189\,D^{v} + 0.078\,{}^{0}\chi^{v}{}_{d} + 0.077\,{}^{S}\psi_{E} + 4.039\,T_{\Sigma/M}$ (0.2, 0.0002, 0.04, 0.01, 0.01, 0.4) N(TR) = 24, r^{2} = 0.919, s = 0.4; N(+9EV) = 33, r^{2} = 0.768, s = 0.7 Excluded outlier: formamide & MeOH ∈ {EV}	(8)
$\delta^{v}{}_{po}(50)$	$UV = 299.1 + 50.54\,{}^{S}M^{v} - 37.34\,{}^{L}M^{v} - 9.048\,{}^{Ho}M_{E} + 1.310\,{}^{St}M_{E}$ (13, 4.9, 3.8, 1.1, 0.2) N(TR) = 25, r^{2} = 0.893, s = 15; N(+8EV) = 33, r^{2} = 0.826, s = 21 Excluded outlier: 4-Me-2-pentanone ϵ {TR}; 2-butanone, MeOH, acetonitrile ϵ {EV}	(9)
$\delta^{v}{}_{po}(50)$	$-\chi\cdot10^{6} = 0.617 + 0.044\,{}^{0}\chi_{d} + 2.208\,{}^{1}\chi^{v}{}_{s} - 2.212\,{}^{1}\psi_{Is} + 0.070\,\Delta - 0.016\,\Sigma$ (0.02, 0.01, 0.4, 0.5, 0.008, 0.003) N(TR) = 23, r^{2} = 0.876, s = 0.04; N(+7EV) = 30, r^{2} = 0.875, s = 0.04 Excluded outlier: nitromethane & MeOH ∈ {EV}	(10)
$\delta^{v}{}_{ppo}(1)$	$El = 0.018 + 0.181 \times 10^{-3}\,{}^{1}\chi_{d} - 0.675\cdot10^{-6}\,{}^{1}\chi^{v}{}_{d} + 0.003\,{}^{0}\psi_{Id} + 140.8\,T_{\Sigma/M}$ (0.02, 0.00006, 10^{-7}, 0.0004, 14) N(TR) = 15, r^{2} = 0.934, s = 0.06; N(+3EV) = 18, r^{2} = 0.931, s = 0.06	(11)

Table 5. ANN results with descriptors of Table 4 with one hidden neuron. 1st column: the δ^v-type and the Ntr value; 2nd col.: ANN-MLP architecture, abbreviations for the activation functions for the internal layers, the number of epochs, and training and test errors; 3rd col.: input indices, sensitivity values, and statistical parameters for the training plus test sets, a[N(aTR + bTE)], and plus the evaluation set, [N(+cEV)].

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(1)$ Ntr = 10^5	6-1-1 (e, l) * 41 0.005/0.003	$({}^0\chi, D^v, {}^0\psi_I, {}^1\psi_I, {}^0\psi_E, T_{\Sigma/M}) \rightarrow T_b$ (12) (30.67, 34.22, 41.80, 1.111, 15.76, 2.291) N(36TR + 9TE) = 45, r^2 = 0.850, s = 21; N(+16EV) = 61 r^2 = 0.820, s = 23 Excluded outlier: dMe-Ether & SO_2 ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^3	6-1-1 (e, s) 8 0.004/0.002	$(\chi_t{}^v, {}^1\chi^v{}_d, {}^1\chi^v{}_s, {}^1\psi_{Id}, {}^1\psi_{Es}, T_{\Sigma/M}) \rightarrow \varepsilon$ (13) (1.209, 1.091, 3.028, 1.108, 1.440, 6.964) N(34TR + 9TE) = 43, r^2 = 0.871, s =3.8; N(+16EV) = 59, r^2 = 0.793, s = 5.1 Excl.out.: ethylencarbonate & quinoline ϵ {TR}, formamide & acetone ∈ {EV}.
$\delta^v{}_{ppo}(-0.5)$ Ntr = 10^3	5-1-1 (t, t) 33 0.002/0.0006	$(D^v, {}^0\chi^v, {}^1\chi^v{}_s, {}^S\psi_E, \Delta) \rightarrow d$ (14) (17.99, 8.653, 2.953, 41.31, 12.37) N(36TR + 9TE) = 45, r^2 = 0.956, s = 0.1; N(+15EV) = 60, r^2 = 0.930, s = 0.1 Excluded outliers: MeCl & MeOH ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^3	6-1-1 (i, i) 20 0.001/0.0001	$({}^0\chi, D^v, {}^0\chi^v, {}^0\psi_E, \Delta, T_{\Sigma/M}) \rightarrow RI$ (15) (78.50, 212.9, 286.4, 356.0, 9.482, 1.603) N(35TR + 10TE) = 45, r^2 = 0.959, s = 0.03; N(+14EV) = 59, r^2 = 0.943, s = 0.04 Excluded outliers: formamide & MeOH ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^5	5-1-1 (e, t) 27 0.005/0.006	$(D^v, {}^1\chi^v{}_s, {}^T\psi_I, {}^1\psi_{Is}, \Delta) \rightarrow \gamma$ (16) (9.086, 34.48, 34.44, 45.45, 2.328) N(22TR + 7TE) = 29, r^2 = 0.841, s = 2.8; N(+10EV) = 39, r^2 = 0.705, s = 3.7 Excluded outlier: nitromethane & formamide ∈ {EV}
$\delta^v{}_{ppo}(0.5)$ Ntr = 10^3	5-1-1 (e, e) 39 0.009/0.009	$({}^HM, {}^HM_I, {}^GM_E, {}^UM_I, {}^{Ho}M_E) \rightarrow FP$ (17) (445.1, $1.44{\cdot}10^6$, $2.65{\cdot}10^6$, $4.22{\cdot}10^6$, $17{\cdot}10^6$) N(22TR + 7TE) = 29, r^2 = 0.801, s = 16; N(+11EV) = 40, r^2 = 0.769, s = 16 Excluded outliers: 2Me-Butane ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^3	5-1-1 (e, l) 17 0.001/0.0004	$({}^1\chi_d, {}^1\psi_I, {}^1\psi_{Id}, {}^0\psi_{Ed}, \Sigma) \rightarrow \eta$ (18) (1.982, 1.509, 1.060, 12.04, 3.824) N(22TR + 6TE) = 28, r^2 = 0.972, s = 0.3; N(+10EV) = 38, r^2 = 0.942, s = 0.4 Excluded outlier: MeOH ∈ {EV}
$\delta^v{}_{po}(5)$ [ϕ = 0, 1] Ntr = 10^3	5-1-1 (e, s) 18 0.002/0.003	$({}^1\chi_d, D^v, {}^0\chi^v{}_d, {}^S\psi_E, T_{\Sigma/M}) \rightarrow \mu$ (19) (317.2, 43.27, 17.80, 26.95, 8.546) N(19TR + 5TE) = 24, r^2 = 0.926, s = 0.4; N(+9EV) = 33, r^2 = 0.768, s = 0.7 Excluded outliers: formamide & MeOH ∈ {EV}

Table 5. *Cont.*

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(50)$ Ntr = 10^3	4-1-1 (s, i) 16 0.003/0.002	$({}^S M^v, {}^L M^v, {}^{Ho}M_E, {}^{St}M_E) \to$ UV (20) (772.2, 543.5, 28.82, 4.862) N(20TR + 5TE) = 25, r^2 = 0.892, s = 14; N(+8EV) = 33, r^2 = 0.794, s = 22 Excl. outl.: 4M2-pentanone ϵ {TR}; 2-butanone, MeOH, Acetonitrile, ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^3	5-1-1 (s, s) 15 0.008/0.001	$({}^0\chi_d, {}^1\chi^v{}_s, {}^1\psi_{Is}, \Delta, \Sigma) \to -\chi\cdot 10^6$ (=MS) (21) (1.420, 6.413, 3.061, 3.569, 1.792) N(19TR + 4TE) = 23, r^2 = 0.809, s = 0.04; N(+7EV) = 30, r^2 = 0.810, s = 0.05 Excluded outliers: nitromethane & MeOH ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^3	4-1-1 (i, i) 20 0.002/0.0003	$({}^A M^v, {}^H M_E, {}^G M_E, {}^{St}M_I) \to$ El (22) (52.93, 3072, 3020, 27.81) N(12TR + 3TE) = 15, r^2 = 0.966, s = 0.04; N(+3EV) = 18, r^2 = 0.955, s = 0.04 pentane and THF ϵ {TR}; excl. Out.: MeOH & 2-propanol ∈ {EV}

* Activation functions: e = exponential, i = identity, l = logistic, t = tanh, s = sin.

Table 6. ANN-MLP results with descriptors of Table 4 with externally imposed number of hidden neurons. The structure is similar to that in Table 5.

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(1)$ Ntr = 10^3	6-2-1 (t, t) 73 0.004/0.002	$({}^0\chi, D^v, {}^0\psi_I, {}^1\psi_I, {}^0\psi_E, T_{\Sigma/M}) \to T_b$ (23) (18.17, 50.17, 138.5, 6.414, 93.87, 4.392) N(36TR + 9TE) = 45, r^2 = 0.891, s = 17; N(+16EV) = 61 r^2 = 0.871, s = 20 Excluded outlier: SO_2 & MeOH ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^5	6-3-1 (t, e) 55 0.002/0.001	$(\chi_t{}^v, {}^1\chi^v{}_d, {}^1\chi^v{}_s, {}^1\psi_{Id}, {}^1\psi_{Es}, T_{\Sigma/M}) \to \varepsilon$ (24) (2.111, 1.902, 8.790, 3.305, 8.234, 16.43) N(34TR + 9TE) = 43, r^2 = 0.942, s =2.5; N(+16EV) = 59, r^2 = 0.830, s = 4.5 Excl. Out.: ethylencarbonate & quinoline ∈ {TR}, formamide & nitromethane ∈ {EV}
$\delta^v{}_{ppo}(-0.5)$ Ntr = 10^3	5-4-1 (t, l) 58 0.0004/0.0001	$(D^v, {}^0\chi^v, {}^1\chi^v{}_s, {}^S\psi_E, \Delta) \to$ d (25) (41.54, 29.37, 9.057, 47.73, 29.59) N(36TR + 9TE) = 45, r^2 = 0.990, s = 0.04; N(+15EV) = 60, r^2 = 0.966, s = 0.1 Excluded outliers: formamide & MeCl ∈ {EV}.
$\delta^v{}_{ppo}(1)$ Ntr = 10^3	6-2-1 (t, s) 20 0.0001/0.0004	$({}^0\chi, D^v, {}^0\chi^v, {}^0\psi_E, \Delta, T_{\Sigma/M}) \to$ RI (26) (152.0, 450.4, 1447, 596.4, 25.73, 2.743) N(35TR + 10TE) = 45, r^2 = 0.995, s = 0.03; N(+14EV) = 59, r^2 = 0.987, s = 0.05 Excluded outliers: formamide & MeOH ∈ {EV}

Table 6. *Cont.*

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(-0.5)$ Ntr = 10^5	5-4-1 (t, e) 36 0.004/0.002	$(D^v, {}^1\chi^v{}_s, {}^T\,\psi_I, {}^1\psi_{Is}, \Delta) \to \gamma$ (27) (1285, 21.98, 2093, 62687, 5.853) N(22TR + 7TE) = 29, r^2 = 0.908, s = 2.1; N(+10EV) = 39, r^2 = 0.871, s = 2.4 Excluded outlier: nitromethane & formamide ∈ {EV}
$\delta^v{}_{po}(1)$ Ntr = 10^5	5-5-1 (t, l) 35 0.003/0.009	$(D, {}^1\psi_{Is}, {}^0\psi_{Ed}, \Delta, T_{\Sigma/M}) \to FP$ (28) (8.683, 2.965, 1.212, 5.431, 5.439) N(22TR + 7TE) = 29, r^2 = 0.919, s = 10; N(+11EV) = 40, r^2 = 0.860, s = 13 Excluded outliers: nitromethane ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^5	5-3-1 (e, l) 35 0.0003/0.0003	$({}^1\chi_d, {}^1\psi_I, {}^1\psi_{Id}, {}^0\psi_{Ed}, \Sigma) \to \eta$ (29) (4.609, 5.914, 1.286, 15.86, 6.803) N(22TR + 6TE) = 28, r^2 = 0.982, s = 0.3; N(+10EV) = 38, r^2 = 0.975, s = 0.3 Excluded outlier: 2-butanone ∈ {EV}
$\delta^v{}_{po}(5)$ [ϕ = 0, 1] Ntr = 10^5	5-2-1 (t, t) 77 0.001/0.001	$({}^1\chi_d, D^v, {}^0\chi^v{}_d, {}^S\psi_E, T_{\Sigma/M}) \to \mu$ (30) (12.41, 109.7, 76.57, 90.85, 34.04) N(19TR + 5TE) = 24, r^2 = 0.970, s = 0.2; N(+9EV) = 33, r^2 = 0.874, s = 0.5 Excluded outliers: HAc, and MeOH ∈ {EV}
$\delta^v{}_{po}(0.5)$ Ntr = 10^5	4-5-1 (t, e) 142 0.002/0.0006	$(D^v, {}^0\chi^v, {}^0\psi_E, \Delta) \to UV$ (31) (604041, 1166, 22291, 18.45) N(20TR + 5TE) = 25, r^2 = 0.970, s = 7.5; N(+8EV) = 33, r^2 = 0.895, s = 13, Excl. Out.: 4M2-pentanone ϵ {TR}; nitromethane, MeOH, acetone ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^3	5-3-1 (e, s) 18 0.003/0.0008	$({}^0\chi_d, {}^1\chi^v{}_s, {}^1\psi_{Is}, \Delta, \Sigma) \to -\chi\cdot 10^6$ (=MS) (32) (3.148, 21.76, 4.090, 8.594, 3.054) N(19TR + 4TE) = 23, r^2 = 0.907, s = 0.03; N(+7EV) = 29, r^2 = 0.871, s = 0.04 Excluded outliers: nitromethane, MeOH ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^3	4-2-1 (t, s) 22 0.001/0.003	$({}^AM^v, {}^HM_E, {}^GM_E, {}^{St}M_I) \to El$ (33) (80.08, 3075, 2819, 34.79) N(12TR + 3TE) = 15, r^2 = 0.973, s = 0.03; N(+3EV) = 18, r^2 = 0.975, s = 0.03 pentane and THF ϵ {TR}; excluded outliers: acetonitrile & 2-propanol ∈ {EV}

Table 7. ANN-MLP results with the number of hidden neurons chosen by Statistica 8. Descriptors are those of Table 4. The structure is similar to that in Table 5.

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(1)$ Ntr = 10^3	6-11-1 (t, t) 39 0.005/0.005	$({}^0\chi, D^v, {}^0\psi_I, {}^1\psi_I, {}^0\psi_E, T_{\Sigma/M}) \rightarrow T_b$ (34) (17.98, 45.18, 106.2, 2.556, 72.23, 3.579) N(36TR + 9TE) = 45, r^2 = 0.846, s = 21; N(+16EV) = 61 r^2 = 0.826, s = 24 Excluded utlier: MeOH & SO_2 ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^5	6-3-1 (t, e) 66 0.002/0.001	$(\chi_t{}^v, {}^1\chi^v{}_d, {}^1\chi^v{}_s, {}^1\psi_{Id}, {}^1\psi_{Es}, T_{\Sigma/M}) \rightarrow \varepsilon$ (35) (2.598, 2.510, 10.40, 3.409, 10.99, 15.65) N(34TR + 9TE) = 43, r^2 = 0.942, s =2.5; N(+16EV) = 59, r^2 = 0.742, s = 5.7 Excl. Out.: ethylencarbonate & quinoline ϵ {TR}, formamide & acetone ∈ {EV}
$\delta^v{}_{ppo}(-0.5)$ Ntr = 10^5	5-8-1 (t, l) 18 0.001/0.001	$(D^v, {}^0\chi^v, {}^1\chi^v{}_s, {}^S\psi_E, \Delta) \rightarrow d$ (36) (20.47, 8.414, 4.606, 49.56, 19.77) N(36TR + 9TE) = 45, r^2 = 0.970, s = 0.05; N(+15EV) = 60, r^2 = 0.938, s = 0.07 Excluded outliers: MeCl & MeOH ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^5	6-4-1 (e, i) 66 0.0001/0.0004	$({}^0\chi, D^v, {}^0\chi^v, {}^0\psi_E, \Delta, T_{\Sigma/M}) \rightarrow RI$ (37) (447.3, 947.0, 1178, 1152, 39.05, 14.42) N(35TR + 10TE) = 45, r^2 = 0.990, s = 0.02; N(+14EV) = 59, r^2 = 0.984, s = 0.02 Excluded outliers: MeCl & MeOH ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^3	5-10-1 (l, s) 42 0.004/0.002	$(D^v, {}^1\chi^v{}_s, {}^T\psi_I, {}^1\psi_{Is}, \Delta) \rightarrow \gamma$ (38) (18.16, 81.96, 74.19, 173.8, 2.809) N(22TR + 7TE) = 29, r^2 = 0.890, s = 2.3; N(+10EV) = 39, r^2 = 0.851, s = 2.6 Excluded outlier: nitromethane & formamide ∈ {EV}
$\delta^v{}_{po}(1)$ Ntr = 10^5	5-4-1 (l, l) 81 0.003/0.01	$(D, {}^1\psi_{Is}, {}^0\psi_{Ed}, \Delta, T_{\Sigma/M}) \rightarrow FP$ (39) (6.663, 2.542, 1.105, 4.616, 3.220) N(22TR + 7TE) = 29, r^2 = 0.899, s = 11; N(+11EV) = 40, r^2 = 0.840, s = 14 Excluded outliers: 2Me-Butane ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^3	5-3-1 (e, l) 26 0.0003/0.0003	$({}^1\chi_d, {}^1\psi_I, {}^1\psi_{Id}, {}^0\psi_{Ed}, \Sigma) \rightarrow \eta$ (40) (6.071, 4.640, 1.164, 14.16, 7.089) N(22TR + 6TE) = 28, r^2 = 0.981, s = 0.3; N(+10EV) = 38, r^2 = 0.974, s = 0.3 Excluded outlier: 2-butanone ∈ {EV}
$\delta^v{}_{po}(5)$ $[\phi = 0, 1]$ Ntr = 10^5	5-4-1 (t, t) 49 0.001/0.0005	$({}^1\chi_d, D^v, {}^0\chi^v{}_d, {}^S\psi_E, T_{\Sigma/M}) \rightarrow \mu$ (41) (20.13, 174.3, 115.0, 202.7, 62.97) N(19TR + 5TE) = 24, r^2 = 0.977, s = 0.2; N(+9EV) = 33, r^2 = 0.835, s = 0.6 Excluded outliers: HAc, and MeOH ∈ {EV}

Table 7. *Cont.*

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(0.5)$ Ntr = 10^5	4-5-1 (t, e) 108 0.001/0.0003	$(D^v, {}^0\chi^v, {}^0\psi_E, \Delta) \rightarrow UV$ (42) (2555, 517.8, 51639, 43.21) N(20TR + 5TE) = 25, r^2 = 0.970, s = 7.3; N(+8EV) = 33, r^2 = 0.941, s = 10 Excl. Out.: 4M2-pentanone ϵ {TR}; nitromethane, 2-butanone, acetone ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^5	5-4-1 (t, i) 78 0.0004/0.0001	$({}^0\chi_d, {}^1\chi^v{}_s, {}^1\psi_{Is}, \Delta, \Sigma) \rightarrow -\chi\cdot 10^6$ (=MS) (43) (48.24, 991.4, 1672, 165.9, 112.6) N(19TR + 4TE) = 23, r^2 = 0.989, s = 0.01; N(+7EV) = 29, r^2 = 0.852, s = 0.04 Excluded outliers: nitromethane & MeOH ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^5	4-5-1 (t, t) 49 0.002/0.001	$({}^AM^v, {}^HM_E, {}^GM_E, {}^{St}M_I) \rightarrow El$ (44) (66.31, 355.9, 331.7, 27.55) N(12TR + 3TE) = 15, r^2 = 0.973, s = 0.03; N(+3EV) = 18, r^2 = 0.973, s = 0.03 pentane and THF ϵ {TR} and excluded MeOH & 2-propanol ∈ {EV}

Table 8. ANN-MLP results for the set of descriptors of Table 4 with only one hidden neuron where either one or two indices were deleted, usually, those with the lowest sensitivity values shown in Table 5. The structure is similar to that in Table 5. Only the satisfactory results are here shown.

δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{po}(1)$ Ntr = 10^5	4-1-1 (e, e) 25 0.008/0.008	$({}^0\chi, D^v, {}^0\psi_I, {}^0\psi_E) \rightarrow T_b$ (45) (816.3, 863.6, 110900, 7016972) N(36TR + 9TE) = 45, r^2 = 0.758, s = 26; N(+16EV) = 61 r^2 = 0.714, s = 29 Excluded outlier: dMe-Ether & SO_2 ∈ {EV}
$\delta^v{}_{po}(50)$ Ntr = 10^3	4-1-1 (i, s) 8 0.006/0.01	$(\chi_t{}^v, {}^1\chi^v{}_s, {}^1\psi_{Es}, T_{\Sigma/M}) \rightarrow \varepsilon$ (46) (1.033, 1.602, 1.092, 3.781) N(34TR + 9TE) = 43, r^2 = 0.761, s =5.2; N(+16EV) = 59, r^2 = 0.903, s = 5.2 Excl.out.: ethylencarbonate & quinoline ϵ {TR}, nitromethane & HAc ∈ {EV}
$\delta^v{}_{ppo}(-0.5)$ Ntr = 10^3	4-1-1 (l, t) 17 0.004/0.002	$(D^v, {}^0\chi^v, {}^S\psi_E, \Delta) \rightarrow d$ (47) (11.01, 7.934, 28.40, 4.905) N(36TR + 9TE) = 45, r^2 = 0.917, s = 0.1; N(+15EV) = 60, r^2 = 0.895, s = 0.1 Excluded outliers: SO_2 & Formamide ∈ {EV}
$\delta^v{}_{ppo}(1)$ Ntr = 10^3	4-1-1 (i, e) 14 0.0008/0.0008	$({}^0\chi, D^v, {}^0\chi^v, {}^0\psi_E) \rightarrow RI$ (48) (1220, 479.3, 31.61, 2.185) N(35TR + 10TE) = 45, r^2 = 0.926, s = 0.05; N(+14EV) = 59, r^2 = 0.914, s = 0.05 Excluded outliers: THF & MeCl ∈ {EV}

Table 8. *Cont.*

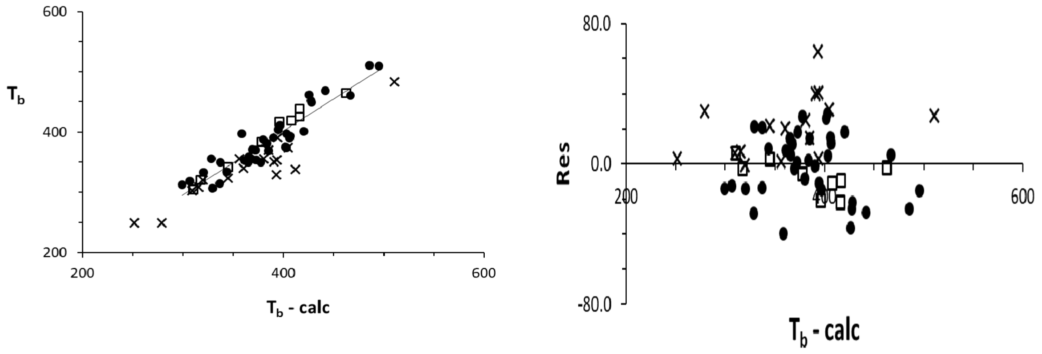
δ^v-Type	ANN-MLP	(Variables) → Property
$\delta^v{}_{ppo}(0.5)$ Ntr = 10^5	4-1-1 (i, l) 26 0.01/0.02	$({}^{H}M_I, {}^{G}M_E, {}^{U}M_I, {}^{Ho}M_E) \to FP$ (49) (10.65, 14.68, 15.90, 12.16) N(22TR + 7TE) = 29, r^2 = 0.719, s = 19; N(+11EV) = 40, r^2 = 0.702, s = 18 Excluded outliers: 2Me-Butane ∈ {EV}
$\delta^v{}_{po}(-0.5)$ Ntr = 10^3	3-1-1 (t, i) 67 0.0007/0.0003	$({}^{1}\chi_d, {}^{0}\psi_{Ed}, \Sigma) \to \eta$ (50) (1.603, 15.54, 10.70) N(22TR + 6TE) = 28, r^2 = 0.965, s = 0.4; N(+10EV) = 38, r^2 = 0.917, s = 0.5 Excluded outlier: MeOH ∈ {EV}
$\delta^v{}_{po}(5)$ [ϕ = 0, 1] Ntr = 10^5	4-1-1 (t, e) 22 0.009/0.005	$(D^v, {}^{0}\chi^v{}_d, {}^{S}\psi_E, T_{\Sigma/M}) \to \mu$ (51) (2.582, 4.178, 6.314, 3.371) N(19TR + 5TE) = 24, r^2 = 0.795, s = 0.6; N(+9EV) = 33, r^2 = 0.746, s = 0.7 Excluded outliers: HAc & MeOH ∈ {EV}

Table 9. Statistical, N/r^2 (2nd decimal figure)/s, results for the eleven properties from Tables 4–7. 2nd column: MLS, results, 3rd column: ANN with one hidden neuron (ANN 1HN) results, 4th column: ANN with externally chosen number of hidden neurons (ANN enHN) results, 5th column: ANN with software chosen number of hidden neurons (ANN snHN) results. First line shows the statistical results for the training (MLS) and train plus test (ANN) compounds, the second line shows the overall statistical results inclusive of the evaluated compounds. M stands for MMCIs (otherwise they are MCIs). The last two columns show also the number of hidden neurons (second line, underlined and bold).

P	MLS (Table 4)	ANN 1HN (Table 5)	ANN enHN (Table 6)	ANN snHN (Table 7)
T_b	45/0.82/22 61/ 0.79/25	45/0.85/21 61/0.82/23	45/0.89/17 **2**/61/0.87/20	45/0.85/21 **11**/61/0.83/24
ε	43/0.86/4.2 59/0.90/5.5	43/0.87/3.8 59/0.79/5.1	43/0.94/2.5 **3**/59/0.83/4.5	43/0.94/2.5 **5**/59/0.83/5.7
d	45/ 0.94/0.07 60/0.91/0.08	45/0.96/0.1 60/ 0.93/0.1	45/0.99/0.04 **4**/60/0.97/0.1	45/0.97/0.05 **8**/60/0.94/0.1
RI	45/ 0.96/0.04 59/0.95/0.03	45/0.96/0.03 59/0.94/0.04	45/0.995/0.03 **2**/59/0.99/0.05	45/0.99/0.02 **4**/59/0.98/0.02
γ	29/0.84/3.1 39/0.79/3.1	29/0.84/2.8 39/0.71/3.7	29/0.91/2.1 **4**/39/0.87/2.4	29/0.89/2.3 **10**/39/0.85/2.6
FP	M/29/0.83/16 40/0.76/17	M/29/0.80/16 40/0.77/16	29/0.92/10 **5**/40/0.86/13	29/0.90/11 **4**/40/0.84/14
η	28/0.97/0.4 38/0.94/0.4	28/0.97/0.3 38/0.94/0.4	28/0.98/0.3 **3**/38/0.98/0.3	28/0.98/0.3 **3**/38/0.97/0.3
μ	24/0.92/0.4 33/0.77/0.7	24/0.93/0.4 33/0.77/0.7	24/0.97/0.2 **2**/33/0.87/0.5	24/0.98/0.2 **4**/33/0.84/0.6
UV	M/25/0.89/15 33/0.83/21	M/25/0.89/14 33/0.79/22	25/0.97/7.5 **5**/33/0.90/13	25/0.97/7.3 **5**/33/0.94/10
$-\chi \cdot 10^6$	23/0.88/0.04 30/0.88/0.04	23/0.81/0.04 30/0.81/0.05	23/0.91/0.03 **3**/29/0.87/0.04	23/0.99/0.01 **4**/29/0.85/0.04
El	15/0.93/0.06 18/0.93/0.06	M/15/0.97/0.04 18/0.96/0.04	M/15/0.97/0.03 **2**/18/0.98/0.03	M/15/0.97/0.03 **5**/18/0.97/0.03

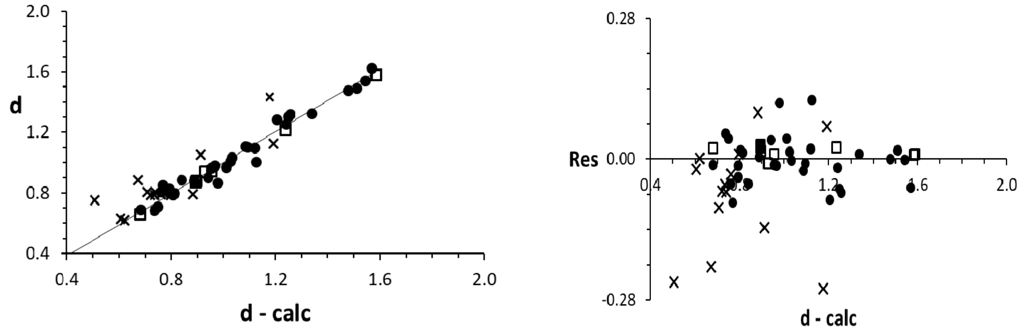
4. Plots

Figures 3–5 display the normal and residual plots of those properties that give rise to the best models and that also show optimal statistics for the evaluated points (given in the captions). All these plots follow the statistics shown in Table 9, 3rd column 2nd line. The structure and importance of this type of plots was discussed in [16,17].



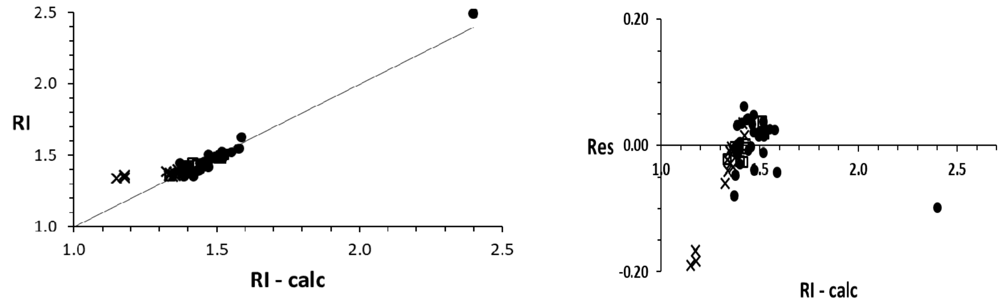
(**a**) (**b**)

Figure 3. Plot of the Experimental vs. calculated (calc) boiling points, $\mathbf{T_b}$, (**a**) and the corresponding residual plot (**b**): (•) training points, (□) test points, and (×) evaluated points. Statistics of evaluation (EV) points: N = 16, r^2 = 0.91, s = 17.



(**a**) (**b**)

Figure 4. Plot of the Experimental vs. calculated (calc) density, **d**, (**a**) and the corresponding residual plot (**b**): (•) training points, (□) test points, and (×) evaluated points. Statistics of EV points: N = 15, r^2 = 0.88, s = 0.1.



(**a**) (**b**)

Figure 5. Plot of the Experimental vs. calculated (calc) refractive index, **RI**, (**a**) and the corresponding residual plot (**b**): (•) training points, (□) test points, and (×) evaluated points. Statistics of EV points: N = 14, r^2 = 0.96, s = 0.1.

5. Discussion

For the ease of discussion and interpretation the most important and detailed statistical results collected through Tables 4–7 are summarized in Table 9. Table 8 shows a special case that will be discussed later on. While Tables 4–7 collect the detailed information about the modeling of the eleven properties, and especially about the type of indices, valence deltas, and structure of the ANN computations, Table 9 gives direct information about the different models.

Looking for MMCI indices (letter M), in MLS, they are optimal for three properties: refractive index RI, flashpoints FP, and cutoff UV.

In ANN computations with one hidden neuron, (ANN 1HN, Table 5), these are instead important descriptors for cutoff UV, flashpoints FP, and elutropic values El. It seems that properties with less training points are better modelled by MMCIs. Concerning the statistical results for the training compounds, ANN 1HN (Table 9, 1st line) improves over MLS for T_b, and El properties, while it lays behind for $-\chi\cdot10^6$, otherwise results are rather similar. With the whole set of compounds (Table 9, second line); i.e., with training (and test with ANN)—plus evaluated compound ANN 1HN calculations improve again over MLS for T_b, and El, while they stay behind with ε, γ, UV, and $-\chi\cdot10^6$.

As soon as the number of hidden neurons grows either by external choice, enHN (Table 6), or by software choice, snHN (Table 7), MMCIs are optimal descriptors only for Elutropic values (silica) El, which is the property with the lowest number of points.

The multiple hidden neuron case shows that, at the training level ANN enHN (Table 9), things improve consistently over the two previous cases (MLS and ANN 1HN) for T_b, ε, d, RI, γ, FP, μ, and UV. For $-\chi\cdot10^6$, ANN with several hidden neurons improves with respect to ANN 1HN, and for El there is an improvement only in relation to MLS (Table 4). Results for viscosity, η, are rather similar throughout the three cases. Mostly, improvement concerns both the r^2 and the s statistics. Concerning the whole set of compounds (TR, TE and EV) statistics improve in relation to the two previous cases (MLS and ANN 1HN) for T_b, ε, γ, FP, μ, and UV.

The advantage of the ANN over the MLS procedure in general is not striking in the eleven properties. In fact, with the only exception of the training plus test for the $-\chi\cdot10^6$ property it does not achieve any useful improvement.

Normally, for an optimal modeling the number of hidden neurons that are externally chosen (ANN enHN, Table 9) is smaller than the number of hidden neurons chosen by the software (ANN snHN, Table 9). In some cases, it is much smaller, like for T_b (an extreme case), d, and γ. Furthermore, ANN snHN statistics are either worse or similar to the ANN enHN ones. This means that if you intend to let the software choose the number of hidden neurons then it is better that you stick to the MLS modeling. Could that depend on the ANN initial weights considered? Probably even if it seems a general trend; i.e., it shows up with nearly all properties.

The MLS results compare rather well with the ANN 1HN results even if the ANN computations have a number of weights bigger (by two) than the number of regression coefficients of the corresponding MLS computations. Thus, we decided to perform ANN calculations by deleting the two indices with the lowest sensibility values in Table 5. In those cases where deletion of two indices gives rise to poor modeling, we deleted only one index. In this last case, the number of weights is no longer equal (actually it is bigger by one) to the number of regression coefficients or weights of the MLS case. Results are shown in Table 8, and, as the reader can notice, four properties, γ, UV, $-\chi\cdot10^6$, and El, do not show up due to poor modeling, while for properties d, FP, and μ, only one index was deleted. We also notice that the dipole moment, μ, does not obey the lowest sensibility rule (see Table 5) as following this rule we should have deleted $T_{\Sigma/M}$ index. Now, deletion of this index gives rise to a poor modeling for the dipole moment. This confirms that sensibility values change from run to run, like the weights, and they are not guidance for the absolute importance of an index, but only for its importance within a given model. The statistics here are usually not as good as in the MLS case (Table 4), with a clear and amazing exception, the modeling of the whole set of compounds for the dielectric constant, ε. Looking only at the training plus test modeling, we would simply discarded this modeling. Nevertheless, the very good modeling of the evaluated compounds helps to improve the overall model for this property. Thus, (i) before throwing away some training plus test ANN or MLS results, re-evaluate and do not forget that (ii) a very good ANN modeling may hiding somewhere.

All this comes back to the random assigning of the initial weights in ANN computations, which renders it difficult to reproduce values that seem to show up by chance. Tables 5–8 tell also that there is no fixed preferential value for the parameter Ntr (numbers of networks to train). Usually, different Ntr values give rise to rather similar statistical parameters.

Generally, the addition of the EV set does not greatly affect the overall quality of the models, showing their robustness in most cases. The differences in r^2 are not greater than 0.5, as a rule. Some exceptions are the MLS models for FP and μ, and the ANN ones for ε, γ, and μ.

Concerning the most used values for δ^v, Tables 4–7 show that the $\delta^v{}_{ppo}$ configuration is preferred, especially throughout the nHN cases (Tables 6 and 7). This choice means a strong dependence on the core electrons for higher row atoms (see Appendix A). Regarding the exponent of the fractional term in δ^v (see Appendix A), the most used values are 1, −0.5; i.e., strong hydrogen atom dependence—and 50; i.e., no hydrogen atom dependence. The strong hydrogen dependence of δ^v tells us that the hydrogen atoms should not be neglected.

Plots of Figures 3–6 exemplify the best models obtained from the given properties. These four properties, T_b, d, RI, and η (Vis) show the best statistics for the set of points evaluated. The residual plots, nevertheless, remind us that the models achieved could be further improved since the evaluated points are not placed symmetrically around the zero line, as required in a perfect model. In the graphs of Figures 4 and 5 a point appears away from the remaining points, which could anchor the regression line. However, the corresponding residual plots show that this is not the case, since their residuals are not insignificant. This is due to the large number of values concentrated in the cloud of the remaining points.

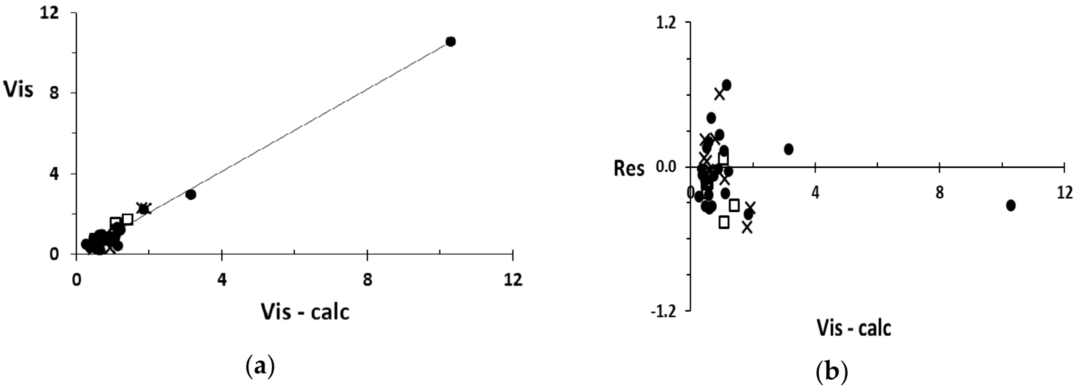


Figure 6. Plot of the Experimental vs. calculated (calc) viscosity, η, (**a**) and the corresponding residual plot (**b**): (●) training points, (□) test points, and (×) evaluated points. Statistics of EV points: N = 9, $r^2 = 0.98$, s = 0.1.

6. Conclusions

The first interesting result of the present ANN-MLP computations is that MCIs are preferred over MMCIs, especially with properties with a relatively high number of points. In fact, only El, with a minimum number of points, is usefully described with MMCI when ANN-MLP with more than one hidden neuron is performed.

The second result suggests that for the properties given it is better to impose from the outside the number of hidden neurons.

The third result shows that, with some exceptions, ANN-MLP improves on MLS calculations, even if the improvement is not dramatic.

One of the great advantages of MLS computation is that its statistical results are reproducible, no matter how many times the calculations are repeated with the same indices, the same results are obtained. The ANN-MLP results can seem, instead, as non-reproducible since the weights of the ANN-MLP calculations start with random values, and the minimization procedure usually ends up with different values from run to run. Furthermore, as a rule, different ANN-MLP computations end up in different local minima. However, it must be pointed out that repeating the training process by setting up the same procedure, by using the same seed, randomization the algorithm and precision, with the same data sets, the resulting model would be the same.

ANN-MLP results obtained with one hidden neuron either with the full set of descriptors (Table 5), or with a reduced set of descriptors (Table 8) confirm the validity of the MLS calculations. The asymmetry of the evaluated points around the zero line of the residual plots, reminds us that things might be further improved either with other types of ANN-MLP calculations or with new types of descriptors.

These results indicate that MLS models should be preferred, except when it is necessary to reach a given quality in the predictions that is only achievable with ANN-MLP models.

The present study also tells us that it is worth considering the hydrogen atoms when performing the calculations to derive the MCIs or the MMCIs, as in many cases they help to improve the quality of a model both in the MLS and ANN-MLP computations.

Author Contributions: J.V.d.J.-O., L.P. and E.B. conceived and designed the idea, analyzed the data and wrote the paper; L.P. performed the calculations.

Funding: This research received no external funding.

Acknowledgments: We would like to thank the suggestions of the referees that helped to improve the paper.

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

Appendix A. The Valence Delta

The δ^v number used in current and previous works is defined as follows [7],

$$\delta^v = \frac{(q + f_\delta^n)\delta^v(ps)}{(p \cdot r + 1)} \tag{A1}$$

$\delta^v(ps)$ is the valence of a vertex in a chemical pseudograph (or general graph) that allows multiple bonds and self-connections (or loops). Usually, in chemical graph theory simple graphs (with no multiple bonds and loops) and pseudographs are hydrogen-depleted. Parameters p is the order of a complete graph, K_p, used to encode core electrons[7], while r is its regularity ($r = p - 1$). A complete graph is a graph where every pair of its vertices is adjacent. The first order complete graph, K_1, that encodes the second row atoms, is just a vertex. Higher values for p encode higher row atoms. Parameter q in Equation (A1) is two-valued: q = 1 or p, where p = 1, 2, 3, 4, 5, 7, ...

Generally, two representations (or configurations) for δ^v are useful: a K_{po}, configuration where q = 1, and p = odd, and a K_{ppo} one where q = p and, again, p = odd.

The f_δ fractional perturbation parameter (or hydrogen perturbation) that encodes the depleted hydrogen atoms is defined in the following way,

$$f_\delta = 1 - \delta^v(ps)/\delta^v{}_m(ps) = n_{H/}\delta^v{}_m(ps) \tag{A2}$$

$\delta^v{}_m(ps)$ is the maximal $\delta^v(ps)$ value a heteroatom (a vertex) can have in a hydrogen depleted chemical pseudograph when all bonded hydrogen atoms are substituted by heteroatoms, and n_H equals the number of hydrogen atoms bonded to a heteroatom. For completely substituted heteroatoms, $f_\delta = 0$ as $\delta^v{}_m(ps) = \delta^v(ps)$ (i.e., $n_H = 0$). In hydrocarbons $\delta^v(ps) = \delta$, which is the delta number in simple chemical graphs with no multiple bonds and loops. In this case: $\delta^v = (1 + f_\delta{}^n)\delta$ (for p = 1). For quaternary carbons $f_\delta = 0$ and $\delta^v = \delta$. Exponent n in f_δ quantifies the importance of the hydrogen perturbation; i.e., the higher the n values the lower the importance of the perturbation. Different values for n give rise to different sets of indices. In this study: n = −0.5, 0.5, 1, 2, 5, 50.

Appendix B. The Intrinsic-I-State and the Electrotopological S-State Indices

The I- and E-State indices ($\psi_{E,I}$:E means electrotopological, and I intrinsic), known in the literature as I and S indices, respectively, are related to δ^v in the following way [4],

$$I = (\delta^v + 1)/\delta, S = I + \Sigma\Delta I, \text{ with } \Delta I = (I_i - I_j)/r^2{}_{ij} \quad \text{(A3)}$$

r_{ij} counts the atoms in the minimum path length separating atoms i and j, which equals the graph distance, d_{ij} + 1; ΣΔI incorporates the information about the influence of the remainder of the molecular environment, and, as it can be negative, S can also be negative for some atoms.

References

1. Pogliani, L.; de Julián-Ortiz, J.V. Artificial neural networks and multilinear least squares to model physicochemical properties of organic solvents. *Int. J. Chem. Mod.* **2014**, *6*, 241–254.
2. Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. [CrossRef]
3. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Structure-Activity Analysis*; Wiley: New York, NY, USA, 1986.
4. Kier, L.B.; Hall, L.H. The Electrotopological State. In *Molecular Structure Description*; Academic Press: New York, NY, USA, 1999.
5. Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*, 2nd ed.; Wiley-VCH: Weinheim, German, 2000.
6. Pogliani, L. From molecular connectivity indices to semiempirical connectivity terms: Recent trends in graph theoretical descriptors. *Chem. Rev.* **2000**, *100*, 3827–3858. [CrossRef] [PubMed]
7. García-Domenech, R.; Gálvez, J.; de Julián-Ortiz, J.V.; Pogliani, L. Some new trends in chemical graph theory. *Chem. Rev.* **2008**, *108*, 1127–1169. [CrossRef] [PubMed]
8. Pogliani, L.; de Julián-Ortiz, J.V. Testing selected optimal descriptors with artificial neural networks. *RSC Adv.* **2013**, *3*, 14710–14721. [CrossRef]
9. Pogliani, L.; de Julián-Ortiz, J.V. QSPR with descriptors based on averages of vertex invariants. An artificial neural network study. *RSC Adv.* **2014**, *4*, 44733–44740. [CrossRef]
10. Topliss, J.G.; Costello, R.J. Chance correlations in structure-activity studies using multiple regression analysis. *J. Med. Chem.* **1972**, *15*, 1066–1069. [CrossRef] [PubMed]
11. Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design: An Introduction*, 2nd ed.; Wiley-VCH: Weinheim, German, 1999.
12. Livingstone, D.J.; Manallack, D.T.; Tetko, I.V. Data modelling with neural networks: Advantages and limitations. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 135–142. [CrossRef]
13. Castillo, E.; Guijarro-Berdiñas, B.; Fontenla-Romero, O.; Alonso-Betanzos, A. A very fast learning method for neural networks based on sensitivity analysis. *J. Mach. Learn. Res.* **2006**, *7*, 1159–1182.
14. Broyden–Fletcher–Goldfarb–Shanno Algorithm. Available online: http://en.wikipedia.org/wiki/Broyden%E2%80%93Fletcher%E2%80%93Goldfarb%E2%80%93Shanno_algorithm (accessed on 4 July 2018).
15. Mihalic, Z.; Nikolic, S.; Trinajstic, N.J. Comparative study of molecular descriptors derived from the distance matrix. *Chem. Inf. Comput. Sci.* **1992**, *32*, 28–37. [CrossRef]
16. Besalu, E.; de Julián-Ortiz, J.V.; Pogliani, L. Trends and plot methods in MLR studies. *J. Chem. Inf. Model.* **2007**, *47*, 751–760. [CrossRef] [PubMed]
17. Besalú, E.; de Julián-Ortiz, J.V.; Pogliani, L. An overlooked property of plot methods. *J. Math. Chem.* **2006**, *39*, 475–484. [CrossRef]