Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
*J. Serb. Chem. Soc.* 84 (6) S173–S182 (2019) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

# Modeling of linear and nonlinear quantitative structure property relationships of the aqueous solubility of phenol derivatives

SOUMAYA KHEROUF[1], NABIL BOUARRA[1,2], AMEL BOUAKKADIA[1,3] and DJELLOUL MESSADI[1]*

[1]*Laboratory of Environmental and Food Safety, Department of Chemistry, Badji Mokhtar–Annaba University, PB12, 23000, Annaba. Algeria,* [2]*Center of Scientific and Technical Research in Physicochemical Analyzes (CRAPC), Bp 384, Siège ex-Pasna Zone Industrielle, Bou-Ismail, 42004, Tipaza, Algeria and* [3]*University Abbes Laghrour Khenchela – Algeria – BP 1252 Route de Batna Khenchela 40004, Algeria*

*J. Serb. Chem. Soc.* 84 (6) (2019) 575–590

EQUATIONS OF STATISTICAL PARAMETERS USED TO VERIFY THE PREDICTIVITY OF A QSPR MODEL

$$R^2_{\text{val}} = 1 - \left[ \frac{\sum_{i=1}^{n_{\text{ext}}} \left( y_i - \hat{y}_i \right)^2}{\sum_{n=1}^{n_{\text{ext}}} \left( y_i - \bar{y} \right)^2} \right] \tag{S-1}$$

where $y_i$ is the predicted value calculated using the regression of the predicted and experimental data of the prediction set, while in the subsequent formulas the $y_i$ value is calculated using the QSPR model.

$$Q^2_{\text{F1}} = 1 - \left[ \frac{\sum_{i=1}^{n_{\text{ext}}} \left( y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n_{\text{ext}}} \left( y_i - \bar{y}_{\text{tr}} \right)^2} \right] \tag{S-2}$$

where $\bar{y}_{\text{tr}}$ is the mean experimental value in the training set and $n_{\text{ext}}$ the number of molecules in the validation set.

* Corresponding author. E-mail: d_messadi@yahoo.fr

$$Q^2_{F2} = 1 - \left[ \frac{\sum_{i=1}^{n_{ext}} \left( y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n_{ext}} \left( y_i - \bar{y}_{ext} \right)^2} \right] \quad \text{(S-3)}$$

where $\bar{y}_{ext}$ is the mean experimental value in the prediction set.

$$Q^2_{F3} = 1 - \left[ \frac{(\sum_{i=1}^{n_{ext}} \left( y_i - \hat{y}_i \right)^2 / n_{ext})}{(\sum_{i=1}^{n_{tr}} \left( y_i - \hat{y}_{tr} \right)^2 / n_{tr})} \right] \quad \text{(S-4)}$$

$$CCC = 2 \frac{\sum_{i=1}^{n_{ext}} \left( y_i - \bar{y} \right) \left( \hat{y}_i - \bar{\hat{y}} \right)}{\sum_{i=1}^{n_{ext}} \left( y_i - \bar{y}_{ext} \right)^2 + \sum_{i=1}^{n_{ext}} \left( \hat{y}_i - \bar{\hat{y}} \right)^2 + n_{ext} \left( \bar{y}_{ext} - \bar{\hat{y}} \right)^2} \quad \text{(S-5)}$$

In this last formula, $\bar{\hat{y}}$ is the mean predicted value.

RANDOM SPLITTING

TABLE S-I. Values of endpoint log (S) and splitting (ordered response splitting)

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | $h_i$ ($h^*$ = 0.2500) | Std.Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 1 | Naphthalen-1-ol | Training | 2.6415 | 2.858 | 0.0323 | 0.7046 |
| 2 | 2,3,4,5-Tetrachlorophenol | Training | 2.2201 | 2.218 | 0.1056 | –0.007 |
| 3 | 2,3,5,6-Tetrachlorophenol | Training | 2 | 2.4854 | 0.1236 | 1.6597 |
| 4 | 2,3,5-Trimethylphenol | Prediction | 2.9031 | 3.1415 | 0.0493 | 0.7828 |
| 5 | 2,3,6-Trichlorophenol | Prediction | 2.6532 | 2.8344 | 0.048 | 0.5944 |
| 6 | 2,4,6-Trimethylphenol | Training | 3.0792 | 3.1986 | 0.0547 | 0.393 |
| 7 | Naphthalene;2,4,6-trinitrophenol | Training | 4.1383 | 3.3829 | 0.0878 | –2.5317 |
| 8 | 2,4-Dinitrophenol | Prediction | 2.525 | 3.4706 | 0.0455 | 3.0981 |
| 9 | 2,6-Dichlorophenol | Training | 3.4191 | 3.3153 | 0.0635 | –0.3435 |
| 10 | 2,6-Dimethylphenol | Prediction | 3.7945 | 3.9318 | 0.0472 | 0.4502 |
| 11 | 2-Methoxyphenol | Training | 4.3945 | 4.1961 | 0.051 | –0.6521 |
| 12 | 5-*tert*-Butyl-2-methylphenol | Training | 2.6128 | 2.4652 | 0.0688 | –0.4895 |
| 13 | 3-Nitrophenol | Training | 4.0626 | 3.9255 | 0.1227 | –0.4685 |
| 14 | 3-*tert*-Butylphenol | Training | 3.316 | 2.9756 | 0.0538 | –1.12 |
| 15 | 4,5-Dichloro-2-methoxyphenol | Training | 2.85 | 2.8653 | 0.0444 | 0.0502 |
| 16 | 2-Methyl-4,6-dinitrophenol | Training | 2.3464 | 2.4647 | 0.1442 | 0.4093 |
| 17 | 4-Butylphenol | Training | 2.7903 | 2.8705 | 0.0517 | 0.2635 |
| 18 | 4-Chloro-2-methoxyphenol | Training | 3.73 | 3.8774 | 0.07 | 0.4892 |

TABLE S-I. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | $h_i$ ($h^*$ = 0.2500) | Std.Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 19 | 4-Chlorophenol | Training | 4.4314 | 4.7052 | 0.099 | 0.9233 |
| 20 | 4-Hexylphenol | Prediction | 2.5922 | 1.9231 | 0.1216 | –2.2853 |
| 21 | 4-Propan-2-ylphenol | Prediction | 3.5136 | 3.329 | 0.04 | –0.603 |
| 22 | 4-Methoxyphenol | Training | 4.29 | 4.4224 | 0.0669 | 0.4388 |
| 23 | 4-Nitrophenol | Prediction | 4.1303 | 4.0449 | 0.093 | –0.2869 |
| 24 | 4-Nonylphenol | Training | 0.7348 | 0.4993 | 0.3403 | –0.9281 |
| 25 | 4-Octylphenol | Training | 1.1004 | 0.9775 | 0.2523 | –0.4551 |
| 26 | 4-Phenylphenol | Training | 0.9912 | 2.1335 | 0.0578 | 3.7669 |
| 27 | 4-Propylphenol | Training | 3.2375 | 3.3344 | 0.0398 | 0.3164 |
| 28 | 4-Butan-2-ylphenol | Training | 2.9823 | 2.9364 | 0.0522 | –0.1508 |
| 29 | 4-*tert*-Butylphenol | Training | 2.7634 | 2.9079 | 0.0516 | 0.475 |
| 30 | Benzene-1,2-diol | Prediction | 4.6532 | 4.6261 | 0.0796 | –0.0905 |
| 31 | Benzene-1,4-diol | Training | 4.8451 | 5.045 | 0.118 | 0.6812 |
| 32 | 2-Methylphenol | Prediction | 4.415 | 4.485 | 0.0629 | 0.2313 |
| 33 | 2-Ethylphenol | Training | 4.1474 | 4.0264 | 0.0484 | –0.397 |
| 34 | 4-methylphenol | Training | 4.301 | 4.3031 | 0.0603 | 0.0071 |
| 35 | 2,3,4,5,6-Pentachlorophenol | Prediction | 1.1461 | 1.8013 | 0.2332 | 2.3951 |
| 36 | 4-Ethylphenol | Training | 3.902 | 3.8339 | 0.042 | –0.2227 |
| 37 | Phenol | Training | 4.9463 | 5.031 | 0.1328 | 0.2912 |
| 38 | 2,3,4,5-Tetrachloro-6-methoxyphenol | Training | 1.415 | 1.4461 | 0.2508 | 0.115 |
| 39 | Naphthalen-2-ol | Training | 2.8692 | 2.8616 | 0.0326 | –0.0247 |
| 40 | 2-Nitrophenol | Training | 3.0334 | 3.6421 | 0.2303 | 2.2209 |
| 41 | 3-Ethyl-5-methylphenol | Training | 3.3644 | 3.3396 | 0.0269 | –0.0806 |
| 42 | 2-Phenylphenol | Prediction | 2.8451 | 2.3955 | 0.0879 | –1.5068 |
| 43 | 3,4,5-Trichloro-2-methoxyphenol | Training | 2.4914 | 2.1634 | 0.1087 | –1.1121 |
| 44 | 3,4-Dichlorophenol | Training | 3.9664 | 3.645 | 0.0279 | –1.0434 |
| 45 | 3,5-Dimethylphenol | Training | 3.8689 | 3.9783 | 0.0633 | 0.3619 |
| 46 | 3,5-Di-*tert*-butylphenol | Training | 3.7404 | 3.725 | 0.0296 | –0.05 |
| 47 | 3-Methoxyphenol | Prediction | 1.1461 | 1.029 | 0.2305 | –0.4273 |
| 48 | 2,3,5,6-Tetrachlorophenol | Training | 4.8312 | 4.328 | 0.0563 | –1.6582 |
| 49 | 2,3,5-Trichlorophenol | Training | 2.2625 | 2.2715 | 0.1055 | 0.0304 |
| 50 | 2,3-Dichlorophenol | Training | 2.699 | 2.9431 | 0.0438 | 0.7991 |
| 51 | 2,3-Dimethylphenol | Training | 3.9146 | 3.1638 | 0.1035 | –2.5384 |
| 52 | 2,4,5-Trichlorophenol | Training | 3.7782 | 3.7821 | 0.0317 | 0.0126 |
| 53 | 2,4,6-Trichlorophenol | Prediction | 2.9768 | 3.2497 | 0.0733 | 0.9074 |
| 54 | 2,4-Dichlorophenol | Prediction | 2.6375 | 3.1303 | 0.0546 | 1.6222 |
| 55 | 2,4-Dimethylphenol | Training | 3.6532 | 3.6201 | 0.0281 | –0.1076 |
| 56 | 2,5-Dimethylphenol | Prediction | 3.9442 | 3.7535 | 0.0304 | –0.6198 |
| 57 | 2-Chlorophenol | Training | 3.5019 | 3.8284 | 0.0349 | 1.0639 |
| 58 | 2-Propan-2-ylphenol | Prediction | 4.3918 | 4.1705 | 0.0676 | –0.7337 |
| 59 | 2,3,4-Trichlorophenol | Training | 3.6457 | 3.5536 | 0.0481 | –0.3022 |
| 60 | 2,3,4-Trichloro-6-methoxyphenol | Training | 2.699 | 2.6668 | 0.0693 | –0.1068 |
| 61 | 4,5-Dichloro-2-methoxyphenol | Training | 1.7324 | 1.7214 | 0.1543 | –0.0384 |
| 62 | 5-Chloro-2-methoxyphenol | Training | 2.7597 | 2.7566 | 0.0437 | –0.0102 |
| 63 | 3-Methylphenol | Prediction | 3.5977 | 3.6011 | 0.0275 | 0.011 |

TABLE S-I. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | $h_i$ ($h^*$ = = 0.2500) | Std.Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 64 | Benzene-1,3-diol | Prediction | 4.3424 | 4.3459 | 0.059 | 0.0116 |
| 65 | 3,4,5-Trimethylphenol | Prediction | 5.0414 | 4.6867 | 0.0782 | –1.1826 |
| 66 | 3,4-Dimethylphenol | Training | 3.1875 | 3.0489 | 0.0449 | –0.4541 |
| 67 | 3-Chlorophenol | Training | 3.7076 | 3.6555 | 0.03 | –0.1693 |
| 68 | Naphthalen-1-ol | Prediction | 4.3424 | 4.297 | 0.0583 | –0.1497 |

ORDERED RESPONSE SPLITTING

TABLE S-II. Values of endpoint log *S* and splitting (ordered response splitting)

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | *hi* ($h^*$ = = 0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 1 | Naphthalen-1-ol | Training | 2.6415 | 2.926 | 0.0295 | 0.9354 |
| 2 | 2,3,4,5-Tetrachlorophenol | Training | 2.2201 | 2.1739 | 0.1465 | –0.1622 |
| 3 | 2,3,5,6-Tetrachlorophenol | Training | 2 | 2.5017 | 0.1545 | 1.7675 |
| 4 | 2,3,5-Trimethylphenol | Training | 2.9031 | 3.1554 | 0.0613 | 0.8436 |
| 5 | 2,3,6-Trichlorophenol | Training | 2.6532 | 2.7976 | 0.0659 | 0.4841 |
| 6 | 2,4,6-Trimethylphenol | Prediction | 3.0792 | 3.2253 | 0.0664 | 0.49 |
| 7 | Naphthalene;2,4,6-trinitro-phenol | Training | 4.1383 | 3.4526 | 0.1012 | –2.3431 |
| 8 | 2,4-Dinitrophenol | Prediction | 2.525 | 3.434 | 0.0486 | 3.0189 |
| 9 | 2,6-Dichlorophenol | Training | 3.4191 | 3.2553 | 0.0711 | –0.5506 |
| 10 | 2,6-Dimethylphenol | Prediction | 3.7945 | 3.9987 | 0.0504 | 0.6789 |
| 11 | 2-Methoxyphenol | Training | 4.3945 | 4.2402 | 0.0494 | –0.5126 |
| 12 | 5-*tert*-Butyl-2-methylphenol | Training | 2.6128 | 2.5734 | 0.0577 | –0.1316 |
| 13 | 3-Nitrophenol | Prediction | 4.0626 | 3.8657 | 0.1217 | –0.6806 |
| 14 | 3-*tert*-Butylphenol | Training | 3.316 | 3.0735 | 0.0482 | –0.8052 |
| 15 | 4,5-Dichloro-2-methoxyphenol | Prediction | 2.85 | 2.8724 | 0.057 | 0.0749 |
| 16 | 2-Methyl-4,6-dinitrophenol | Training | 2.3464 | 2.3262 | 0.191 | –0.0728 |
| 17 | 4-Butylphenol | Training | 2.7903 | 2.9446 | 0.0463 | 0.5117 |
| 18 | 4-Chloro-2-methoxyphenol | Prediction | 3.73 | 3.9815 | 0.0781 | 0.8484 |
| 19 | 4-Chlorophenol | Training | 4.4314 | 4.8277 | 0.1125 | 1.3627 |
| 20 | 4-Hexylphenol | Training | 2.5922 | 2.0315 | 0.0958 | –1.9102 |
| 21 | 4-Propan-2-ylphenol | Training | 3.5136 | 3.3825 | 0.0389 | –0.4331 |
| 22 | 4-Methoxyphenol | Training | 4.29 | 4.5178 | 0.0729 | 0.7664 |
| 23 | 4-Nitrophenol | Training | 4.1303 | 4.0121 | 0.089 | –0.401 |
| 24 | 4-Nonylphenol | Training | 0.7348 | 0.6586 | 0.2547 | –0.286 |
| 25 | 4-Octylphenol | Training | 1.1004 | 1.1206 | 0.1905 | 0.0727 |
| 26 | 4-Phenylphenol | Training | 0.9912 | 2.1994 | 0.0465 | 4.0083 |
| 27 | 4-Propylphenol | Training | 3.2375 | 3.3891 | 0.0387 | 0.5008 |
| 28 | 4-Butan-2-ylphenol | Training | 2.9823 | 3.0254 | 0.0465 | 0.1431 |
| 29 | 4-*tert*-Butylphenol | Prediction | 2.7634 | 2.9905 | 0.046 | 0.753 |
| 30 | Benzene-1,2-diol | Prediction | 4.6532 | 4.6432 | 0.0712 | –0.0336 |
| 31 | Benzene-1,4-diol | Training | 4.8451 | 5.1569 | 0.1294 | 1.0824 |
| 32 | 2-Methylphenol | Training | 4.415 | 4.5513 | 0.0634 | 0.4563 |

TABLE S-II. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | *hi* (*h** = = 0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 33 | 2-Ethylphenol | Prediction | 4.1474 | 4.1134 | 0.0517 | –0.1132 |
| 34 | 4-Methylphenol | Training | 4.301 | 4.3284 | 0.056 | 0.0913 |
| 35 | 2,3,4,5,6-Pentachlorophenol | Prediction | 1.1461 | 1.7949 | 0.2938 | 2.5009 |
| 36 | 4-Ethylphenol | Training | 3.902 | 3.8773 | 0.0413 | –0.0817 |
| 37 | Phenol | Prediction | 4.9463 | 5.0952 | 0.1376 | 0.5194 |
| 38 | 2,3,4,5-Tetrachloro-6-methoxyphenol | Training | 1.415 | 1.3906 | 0.2736 | –0.0928 |
| 39 | Naphthalen-2-ol | Training | 2.8692 | 2.9303 | 0.0298 | 0.2011 |
| 40 | 2-Nitrophenol | Training | 3.0334 | 3.5182 | 0.2434 | 1.8052 |
| 41 | 3-Ethyl-5-methylphenol | Training | 3.3644 | 3.3968 | 0.0277 | 0.1065 |
| 42 | 2-Phenylphenol | Training | 2.8451 | 2.5207 | 0.0742 | –1.092 |
| 43 | 3,4,5-Trichloro-2-methoxyphenol | Training | 2.4914 | 2.1437 | 0.1444 | –1.2175 |
| 44 | 3,4-Dichlorophenol | Training | 3.9664 | 3.6597 | 0.0281 | –1.0079 |
| 45 | 3,5-Dichlorophenol | Training | 3.8689 | 4.0684 | 0.0699 | 0.6701 |
| 46 | 3,5-Dimethylphenol | Training | 3.7404 | 3.7452 | 0.0296 | 0.0157 |
| 47 | 3,5-Di-*tert*-butylphenol | Training | 1.1461 | 1.1852 | 0.1713 | 0.139 |
| 48 | 3-Methoxyphenol | Training | 4.8312 | 4.402 | 0.0583 | –1.4328 |
| 49 | 2,3,5,6-Tetrachlorophenol | Prediction | 2.2625 | 2.2394 | 0.1437 | –0.0808 |
| 50 | 2,3,5-Trichlorophenol | Prediction | 2.699 | 2.931 | 0.0581 | 0.7743 |
| 51 | 2,3-Dichlorophenol | Prediction | 3.9146 | 3.0695 | 0.119 | –2.9166 |
| 52 | 2,3-Dimethylphenol | Training | 3.7782 | 3.8151 | 0.032 | 0.1215 |
| 53 | 2,4,5-Trichlorophenol | Prediction | 2.9768 | 3.3069 | 0.0854 | 1.1182 |
| 54 | 2,4,6-Trichlorophenol | Prediction | 2.6375 | 3.1605 | 0.0662 | 1.7531 |
| 55 | 2,4-Dichlorophenol | Training | 3.6532 | 3.6291 | 0.0284 | –0.0793 |
| 56 | 2,4-Dimethylphenol | Training | 3.9442 | 3.7801 | 0.0305 | –0.5397 |
| 57 | 2,5-Dimethylphenol | Prediction | 3.5019 | 3.8719 | 0.0359 | 1.2208 |
| 58 | 2-Chlorophenol | Training | 4.3918 | 4.172 | 0.0628 | –0.7356 |
| 59 | 2-Propan-2-ylphenol | Prediction | 3.6457 | 3.6579 | 0.0496 | 0.0406 |
| 60 | 2,3,4-Trichlorophenol | Training | 2.699 | 2.5922 | 0.0957 | –0.3639 |
| 61 | 2,3,4-Trichloro-6-methoxyphenol | Prediction | 1.7324 | 1.6017 | 0.2222 | –0.4802 |
| 62 | 4,5-Dichloro-2-methoxyphenol | Training | 2.7597 | 2.7391 | 0.0593 | –0.0688 |
| 63 | 5-Chloro-2-methoxyphenol | Training | 3.5977 | 3.6427 | 0.0285 | 0.1478 |
| 64 | 3-Methylphenol | Prediction | 4.3424 | 4.3808 | 0.0555 | 0.1281 |
| 65 | Benzene-1,3-diol | Training | 5.0414 | 4.7175 | 0.0713 | –1.0887 |
| 66 | 3,4,5-Trimethylphenol | Training | 3.1875 | 3.0417 | 0.0584 | –0.4866 |
| 67 | 3,4-Dimethylphenol | Training | 3.7076 | 3.6599 | 0.03 | –0.1568 |
| 68 | 3-Chlorophenol | Training | 4.3424 | 4.3272 | 0.0546 | –0.0508 |

ORDERED BY STRUCTURES SPLITTING

TABLE S-III. Values of endpoint log *S* and splitting (structure splitting)

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | HAT i/i ($h^*$ = 0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 1 | Naphthalen-1-ol | Training | 2.6415 | 2.8388 | 0.0317 | 0.8152 |
| 2 | 2,3,4,5-Tetrachlorophenol | Prediction | 2.2201 | 2.0261 | 0.1024 | –0.8323 |
| 3 | 2,3,5,6-Tetrachlorophenol | Training | 2 | 2.2556 | 0.1032 | 1.0972 |
| 4 | 2,3,5-Trimethylphenol | Training | 2.9031 | 2.9889 | 0.0425 | 0.3563 |
| 5 | 2,3,6-Trichlorophenol | Prediction | 2.6532 | 2.7198 | 0.0578 | 0.2791 |
| 6 | 2,4,6-Trimethylphenol | Training | 3.0792 | 3.0378 | 0.0473 | –0.1723 |
| 7 | Naphthalene;2,4,6-trinitro-phenol | Prediction | 4.1383 | 3.1956 | 0.0876 | –4.012 |
| 8 | 2,4-Dinitrophenol | Training | 2.525 | 3.434 | 0.064 | 3.8196 |
| 9 | 2,6-Dichlorophenol | Prediction | 3.4191 | 3.2973 | 0.0986 | –0.5215 |
| 10 | 2,6-Dimethylphenol | Training | 3.7945 | 3.83 | 0.0492 | 0.1481 |
| 11 | 2-Methoxyphenol | Training | 4.3945 | 4.2075 | 0.0477 | –0.7792 |
| 12 | 5-*tert*-Butyl-2-methylphenol | Training | 2.6128 | 2.4648 | 0.0727 | –0.6247 |
| 13 | 3-Nitrophenol | Training | 4.0626 | 3.9875 | 0.1607 | –0.3334 |
| 14 | 3-*tert*-Butylphenol | Training | 3.316 | 3.0658 | 0.0541 | –1.0458 |
| 15 | 4,5-Dichloro-2-methoxyphenol | Training | 2.85 | 2.7359 | 0.0396 | –0.4734 |
| 16 | 2-Methyl-4,6-dinitrophenol | Prediction | 2.3464 | 2.4077 | 0.2467 | 0.2873 |
| 17 | 4-Butylphenol | Training | 2.7903 | 2.9756 | 0.0502 | 0.7728 |
| 18 | 4-Chloro-2-methoxyphenol | Training | 3.73 | 3.7692 | 0.0853 | 0.1666 |
| 19 | 4-Chlorophenol | Training | 4.4314 | 4.6549 | 0.1199 | 0.9687 |
| 20 | 4-Hexylphenol | Training | 2.5922 | 2.0914 | 0.1162 | –2.1655 |
| 21 | 4-Propan-2-ylphenol | Training | 3.5136 | 3.4046 | 0.0393 | –0.452 |
| 22 | 4-Methoxyphenol | Prediction | 4.29 | 4.4017 | 0.073 | 0.4718 |
| 23 | 4-Nitrophenol | Training | 4.1303 | 4.09 | 0.1116 | –0.1739 |
| 24 | 4-Nonylphenol | Training | 0.7348 | 0.763 | 0.3221 | 0.139 |
| 25 | 4-Octylphenol | Training | 1.1004 | 1.2088 | 0.2395 | 0.5055 |
| 26 | 4-Phenylphenol | Prediction | 0.9912 | 2.1707 | 0.0549 | 4.9325 |
| 27 | 4-Propylphenol | Training | 3.2375 | 3.4092 | 0.039 | 0.7121 |
| 28 | 4-Butan-2-ylphenol | Prediction | 2.9823 | 3.0322 | 0.0513 | 0.2081 |
| 29 | 4-*tert*-Butylphenol | Training | 2.7634 | 3.0077 | 0.0502 | 1.0191 |
| 30 | Benzene-1,2-diol | Training | 4.6532 | 4.612 | 0.0747 | –0.1739 |
| 31 | Benzene-1,4-diol | Prediction | 4.8451 | 4.9715 | 0.135 | 0.5526 |
| 32 | 2-Methylphenol | Training | 4.415 | 4.4677 | 0.0608 | 0.2212 |
| 33 | 2-Ethylphenol | Training | 4.1474 | 4.0387 | 0.0507 | –0.4537 |
| 34 | 4-Methylphenol | Prediction | 4.301 | 4.3117 | 0.0572 | 0.0448 |
| 35 | 2,3,4,5,6-Pentachlorophenol | Training | 1.1461 | 1.5037 | 0.1936 | 1.6191 |
| 36 | 4-Ethylphenol | Prediction | 3.902 | 3.8735 | 0.0403 | –0.1184 |
| 37 | Phenol | Training | 4.9463 | 5.0993 | 0.1195 | 0.6629 |
| 38 | 2,3,4,5-Tetrachloro-6-meth-oxyphenol | Training | 1.415 | 1.6981 | 0.3352 | 1.4114 |
| 39 | Naphthalen-2-ol | Training | 2.8692 | 2.8419 | 0.0321 | –0.1129 |
| 40 | 2-Nitrophenol | Prediction | 3.0334 | 3.7443 | 0.3376 | 3.5509 |

TABLE S-III. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | HAT i/i ($h^*$ = 0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 41 | 3-Ethyl-5-methylphenol | Training | 3.3644 | 3.2862 | 0.0264 | –0.322 |
| 42 | 2-Phenylphenol | Prediction | 2.8451 | 2.3956 | 0.0966 | –1.9225 |
| 43 | 3,4,5-Trichloro-2-methoxy-phenol | Training | 2.4914 | 1.9687 | 0.0958 | –2.2347 |
| 44 | 3,4-Dichlorophenol | Prediction | 3.9664 | 3.5803 | 0.0278 | –1.5919 |
| 45 | 3,5-Dichlorophenol | Training | 3.8689 | 3.8664 | 0.0731 | –0.0107 |
| 46 | 3,5-Dimethylphenol | Training | 3.7404 | 3.6526 | 0.0284 | –0.3622 |
| 47 | 3,5-Di-*tert*-butylphenol | Prediction | 1.1461 | 1.1256 | 0.2287 | –0.0947 |
| 48 | 3-Methoxyphenol | Prediction | 4.8312 | 4.3207 | 0.0558 | –2.136 |
| 49 | 2,3,5,6-Tetrachlorophenol | Training | 2.2625 | 2.072 | 0.0965 | –0.8147 |
| 50 | 2,3,5-Trichlorophenol | Training | 2.699 | 2.8132 | 0.0431 | 0.4745 |
| 51 | 2,3-Dichlorophenol | Prediction | 3.9146 | 3.1673 | 0.1697 | –3.3342 |
| 52 | 2,3-Dimethylphenol | Prediction | 3.7782 | 3.7015 | 0.0297 | –0.3164 |
| 53 | 2,4,5-Trichlorophenol | Training | 2.9768 | 3.0763 | 0.0692 | 0.4192 |
| 54 | 2,4,6-Trichlorophenol | Training | 2.6375 | 2.9738 | 0.0472 | 1.4006 |
| 55 | 2,4-Dichlorophenol | Training | 3.6532 | 3.5589 | 0.0291 | –0.3891 |
| 56 | 2,4-Dimethylphenol | Training | 3.9442 | 3.6771 | 0.0286 | –1.1019 |
| 57 | 2,5-Dimethylphenol | Training | 3.5019 | 3.7413 | 0.0332 | 0.9898 |
| 58 | 2-Chlorophenol | Prediction | 4.3918 | 4.196 | 0.0705 | –0.8256 |
| 59 | 2-Propan-2-ylphenol | Training | 3.6457 | 3.5974 | 0.0531 | –0.2019 |
| 60 | 2,3,4-Trichlorophenol | Training | 2.699 | 2.5761 | 0.1048 | –0.5283 |
| 61 | 2,3,4-Trichloro-6-methoxy-phenol | Training | 1.7324 | 1.5893 | 0.2171 | –0.6573 |
| 62 | 4,5-Dichloro-2-methoxyphenol | Training | 2.7597 | 2.6426 | 0.0462 | –0.4876 |
| 63 | 5-Chloro-2-methoxyphenol | Prediction | 3.5977 | 3.5321 | 0.026 | –0.2703 |
| 64 | 3-Methylphenol | Training | 4.3424 | 4.3484 | 0.0549 | 0.0251 |
| 65 | Benzene-1,3-diol | Training | 5.0414 | 4.6641 | 0.0719 | –1.5924 |
| 66 | 3,4,5-Trimethylphenol | Training | 3.1875 | 2.9093 | 0.0422 | –1.1555 |
| 67 | 3,4-Dimethylphenol | Training | 3.7076 | 3.5929 | 0.0316 | –0.4738 |
| 68 | 3-Chlorophenol | Training | 4.3424 | 4.3046 | 0.0549 | –0.1579 |

CADEX SPLITTING

TABLE S-IV. Values of endpoint log *S* and splitting (CADEX)

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | HAT i/i ($h^*$=0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 1 | Naphthalen-1-ol | Training | 2.6415 | 2.8593 | 0.0267 | 0.6486 |
| 2 | 2,3,4,5-Tetrachlorophenol | Training | 2.2201 | 2.0074 | 0.1244 | –0.6678 |
| 3 | 2,3,5,6-Tetrachlorophenol | Training | 2 | 2.3379 | 0.1183 | 1.057 |
| 4 | 2,3,5-Trimethylphenol | Training | 2.9031 | 3.0238 | 0.0528 | 0.3642 |
| 5 | 2,3,6-Trichlorophenol | Training | 2.6532 | 2.6659 | 0.0641 | 0.0387 |
| 6 | 2,4,6-Trimethylphenol | Training | 3.0792 | 3.0943 | 0.0555 | 0.0455 |
| 7 | Naphthalene;2,4,6-trinitrophenol | Training | 4.1383 | 3.3235 | 0.0804 | –2.4962 |
| 8 | 2,4-Dinitrophenol | Training | 2.525 | 3.3354 | 0.0516 | 2.4447 |

TABLE S-IV. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | HAT i/i ($h^*$=0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 9 | 2,6-Dichlorophenol | Training | 3.4191 | 3.157 | 0.0733 | –0.7997 |
| 10 | 2,6-Dimethylphenol | Training | 3.7945 | 3.9045 | 0.0458 | 0.3309 |
| 11 | 2-Methoxyphenol | Training | 4.3945 | 4.1851 | 0.044 | –0.6292 |
| 12 | 5-*tert*-Butyl-2-methylphenol | Training | 2.6128 | 2.523 | 0.0507 | –0.2707 |
| 13 | 3-Nitrophenol | Training | 4.0626 | 3.8012 | 0.1091 | –0.8134 |
| 14 | 3-*tert*-Butylphenol | Training | 3.316 | 3.0579 | 0.0437 | –0.7754 |
| 15 | 4,5-Dichloro-2-methoxyphenol | Training | 2.85 | 2.7468 | 0.0497 | –0.311 |
| 16 | 2-Methyl-4,6-dinitrophenol | Training | 2.3464 | 2.1881 | 0.1914 | –0.5172 |
| 17 | 4-Butylphenol | Training | 2.7903 | 2.9279 | 0.0418 | 0.4129 |
| 18 | 4-Chloro-2-methoxyphenol | Training | 3.73 | 3.8945 | 0.0676 | 0.5003 |
| 19 | 4-Chlorophenol | Training | 4.4314 | 4.7718 | 0.1013 | 1.0548 |
| 20 | 4-Hexylphenol | Training | 2.5922 | 2.0443 | 0.0915 | –1.6885 |
| 21 | 4-Propan-2-ylphenol | Training | 3.5136 | 3.351 | 0.0341 | –0.4862 |
| 22 | 4-Methoxyphenol | Training | 4.29 | 4.4649 | 0.0652 | 0.5314 |
| 23 | 4-Nitrophenol | Training | 4.1303 | 3.9489 | 0.0798 | –0.5557 |
| 24 | 4-Nonylphenol | Training | 0.7348 | 0.7157 | 0.2495 | –0.0648 |
| 25 | 4-Octylphenol | Training | 1.1004 | 1.163 | 0.1857 | 0.2038 |
| 26 | 4-Phenylphenol | Training | 0.9912 | 2.151 | 0.0441 | 3.4847 |
| 27 | 4-Propylphenol | Training | 3.2375 | 3.3576 | 0.0339 | 0.3588 |
| 28 | 4-Butan-2-ylphenol | Training | 2.9823 | 3.0094 | 0.0421 | 0.0814 |
| 29 | 4-*tert*-Butylphenol | Training | 2.7634 | 2.9742 | 0.0416 | 0.6325 |
| 30 | Benzene-1,2-diol | Training | 4.6532 | 4.5729 | 0.0652 | –0.244 |
| 31 | Benzene-1,4-diol | Training | 4.8451 | 5.0906 | 0.1176 | 0.7678 |
| 32 | 2-Methylphenol | Training | 4.415 | 4.4923 | 0.0571 | 0.2338 |
| 33 | 2-Ethylphenol | Training | 4.1474 | 4.0692 | 0.0457 | –0.2351 |
| 34 | 4-Methylphenol | Training | 4.301 | 4.2676 | 0.0504 | –0.1008 |
| 35 | 2,3,4,5,6-Pentachlorophenol | Training | 1.1461 | 1.5956 | 0.2216 | 1.4966 |
| 36 | 4-Ethylphenol | Training | 3.902 | 3.8313 | 0.0362 | –0.2116 |
| 37 | Phenol | Training | 4.9463 | 5.0712 | 0.1123 | 0.3895 |
| 38 | 2,3,4,5-Tetrachloro-6-methoxyphenol | Training | 1.415 | 1.3901 | 0.2485 | –0.0844 |
| 39 | Naphthalen-2-ol | Training | 2.8692 | 2.8637 | 0.0269 | –0.0163 |
| 40 | 2-Nitrophenol | Training | 3.0334 | 3.4509 | 0.2182 | 1.3873 |
| 41 | 3-Ethyl-5-methylphenol | Training | 3.3644 | 3.3162 | 0.0255 | –0.1433 |
| 42 | 2-Phenylphenol | Training | 2.8451 | 2.4748 | 0.0649 | –1.1248 |
| 43 | 3,4,5-trichloro-2-methoxyphenol | Training | 2.4914 | 1.9825 | 0.1172 | –1.5911 |
| 44 | 3,4-Dichlorophenol | Training | 3.9664 | 3.5646 | 0.0293 | –1.198 |
| 45 | 3,5-Dichlorophenol | Training | 3.8689 | 3.9766 | 0.0615 | 0.3267 |
| 46 | 3,5-Dimethylphenol | Training | 3.7404 | 3.649 | 0.0304 | –0.2728 |
| 47 | 3,5-Di-*tert*-butylphenol | Training | 1.1461 | 1.179 | 0.1596 | 0.1053 |
| 48 | 3-Methoxyphenol | Training | 4.8312 | 4.3481 | 0.0521 | –1.4576 |
| 49 | 2,3,5,6-Tetrachlorophenol | Prediction | 2.2625 | 2.0735 | 0.1192 | –0.5917 |
| 50 | 2,3,5-Trichlorophenol | Prediction | 2.699 | 2.8003 | 0.0536 | 0.306 |
| 51 | 2,3-Dichlorophenol | Prediction | 3.9146 | 2.9698 | 0.1188 | –2.9567 |
| 52 | 2,3-Dimethylphenol | Prediction | 3.7782 | 3.7195 | 0.0317 | –0.1754 |
| 53 | 2,4,5-Trichlorophenol | Prediction | 2.9768 | 3.1793 | 0.068 | 0.6161 |

TABLE S-IV. Continued

| ID | Name | Status | Exp. endpoint | Pred. by model eq. | HAT i/i ($h^*$=0.2500) | Std. Pred. Mod. Eq. Res. |
|---|---|---|---|---|---|---|
| 54 | 2,4,6-Trichlorophenol | Prediction | 2.6375 | 3.0317 | 0.0546 | 1.1909 |
| 55 | 2,4-Dichlorophenol | Prediction | 3.6532 | 3.5338 | 0.03 | –0.3562 |
| 56 | 2,4-Dimethylphenol | Prediction | 3.9442 | 3.6842 | 0.0308 | –0.7758 |
| 57 | 2,5-Dimethylphenol | Prediction | 3.5019 | 3.7767 | 0.0344 | 0.8217 |
| 58 | 2-Chlorophenol | Prediction | 4.3918 | 4.1108 | 0.0563 | –0.8497 |
| 59 | 2-Propan-2-ylphenol | Prediction | 3.6457 | 3.6286 | 0.0439 | –0.0515 |
| 60 | 2,3,4-Trichlorophenol | Prediction | 2.699 | 2.4588 | 0.0964 | –0.7422 |
| 61 | 2,3,4-Trichloro-6-methoxyphenol | Prediction | 1.7324 | 1.4361 | 0.2089 | –0.9786 |
| 62 | 4,5-Dichloro-2-methoxyphenol | Prediction | 2.7597 | 2.6124 | 0.0553 | –0.4452 |
| 63 | 5-Chloro-2-methoxyphenol | Prediction | 3.5977 | 3.553 | 0.0275 | –0.1333 |
| 64 | 3-Methylphenol | Prediction | 4.3424 | 4.3204 | 0.0499 | –0.0662 |
| 65 | Benzene-1,3-diol | Prediction | 5.0414 | 4.6478 | 0.0653 | –1.196 |
| 66 | 3,4,5-Trimethylphenol | Prediction | 3.1875 | 2.9092 | 0.0532 | –0.8402 |
| 67 | 3,4-Dimethylphenol | Prediction | 3.7076 | 3.563 | 0.0319 | –0.4316 |
| 68 | 3-Chlorophenol | Prediction | 4.3424 | 4.2673 | 0.049 | –0.2264 |

## ARTIFICIAL NEURAL NETWORK

TABLE S-V. Values of experimental log *S* and predicted by ANN

| ID | Name | Status | ID | Name | Status |
|---|---|---|---|---|---|
| 1 | Naphthalen-1-ol | Training | 35 | 2,3,4,5,6-Pentachlorophenol | Training |
| 2 | 2,3,4,5-Tetrachlorophenol | Training | 36 | 4-Ethylphenol | Training |
| 3 | 2,3,5,6-Tetrachlorophenol | Training | 37 | Phenol | Training |
| 4 | 2,3,5-Trimethylphenol | Training | 38 | 2,3,4,5-Tetrachloro-6-methoxy-phenol | Training |
| 5 | 2,3,6-Trichlorophenol | Training | 39 | Naphthalen-2-ol | Training |
| 6 | 2,4,6-Trimethylphenol | Training | 40 | 2-Nitrophenol | Training |
| 7 | naphthalene; 2,4,6-Trinitrophenol | Training | 41 | 3-Ethyl-5-methylphenol | Training |
| 8 | 2,4-Dinitrophenol | Training | 42 | 2-Phenylphenol | Training |
| 9 | 2,6-Dichlorophenol | Training | 43 | 3,4,5-Trichloro-2-methoxyphenol | Training |
| 10 | 2,6-Dimethylphenol | Training | 44 | 3,4-Dichlorophenol | Training |
| 11 | 2-Methoxyphenol | Training | 45 | 3,5-Dichlorophenol | Training |
| 12 | 5-*tert*-Butyl-2-methylphenol | Training | 46 | 3,5-Dimethylphenol | Training |
| 13 | 3-Nitrophenol | Training | 47 | 3,5-Di-*tert*-butylphenol | Training |
| 14 | 3-*tert*-Butylphenol | Training | 48 | 3-Methoxyphenol | Training |
| 15 | 4,5-Dichloro-2-methoxyphenol | Training | 49 | 2,3,5,6-Tetrachlorophenol | Prediction |
| 16 | 2-Methyl-4,6-dinitrophenol | Training | 50 | 2,3,5-Trichlorophenol | Prediction |
| 17 | 4-Butylphenol | Training | 51 | 2,3-Dichlorophenol | Prediction |
| 18 | 4-Chloro-2-methoxyphenol | Training | 52 | 2,3-Dimethylphenol | Prediction |
| 19 | 4-Chlorophenol | Training | 53 | 2,4,5-Trichlorophenol | Prediction |
| 20 | 4-Hexylphenol | Training | 54 | 2,4,6-Trichlorophenol | Prediction |
| 21 | 4-Propan-2-ylphenol | Training | 55 | 2,4-Dichlorophenol | Prediction |
| 22 | 4-Methoxyphenol | Training | 56 | 2,4-Dimethylphenol | Prediction |
| 23 | 4-Nitrophenol | Training | 57 | 2,5-Dimethylphenol | Prediction |
| 24 | 4-Nonylphenol | Training | 58 | 2-Chlorophenol | Prediction |

TABLE S-V. Continued

| ID | Name | Status | ID | Name | Status |
|---|---|---|---|---|---|
| 25 | 4-Octylphenol | Training | 59 | 2-Propan-2-ylphenol | Prediction |
| 26 | 4-Phenylphenol | Training | 60 | 2,3,4-Trichlorophenol | Prediction |
| 27 | 4-Propylphenol | Training | 61 | 2,3,4-Trichloro-6-methoxyphenol | Prediction |
| 28 | 4-Butan-2-ylphenol | Training | 62 | 4,5-Dichloro-2-methoxyphenol | Prediction |
| 29 | 4-*Tert*-Butylphenol | Training | 63 | 5-Chloro-2-methoxyphenol | Prediction |
| 30 | Benzene-1,2-diol | Training | 64 | 3-Methylphenol | Prediction |
| 31 | Benzene-1,4-diol | Training | 65 | Benzene-1,3-diol | Prediction |
| 32 | 2-Methylphenol | Training | 66 | 3,4,5-Trimethylphenol | Prediction |
| 33 | 2-Ethylphenol | Training | 67 | 3,4-Dimethylphenol | Prediction |
| 34 | 4-Methylphenol | Training | 68 | 3-Chlorophenol | Prediction |

TABLE S-VI. Number of compounds test and predicted value of log *S*

| Number compounds test | log ($S$ / g m$^{-3}$) cal | Number compounds test | log ($S$ / g m$^{-3}$) cal |
|---|---|---|---|
| 41 | 3.565 | 43 | 2.081 |
| 36 | 3.764 | 25 | 1.244 |
| 40 | 3.188 | 20 | 2.130 |
| 15 | 2.577 | 10 | 4.044 |
| 34 | 4.341 | 5 | 2.708 |
| 18 | 4.093 | 22 | 4.499 |
| 9 | 3.117 | 6 | 2.990 |
| 1 | 2.717 | 2 | 2.241 |
| 47 | 1.008 | 3 | 2.209 |
| 24 | 0.781 | 29 | 2.957 |