Supporting Information

Predicting the Surface Tension of Liquids: Comparison of Four Modeling Approaches and Application to Cosmetic Oils

Valentin Goussard,[†] François Duprat,[*,‡] Vincent Gerbaud,[§] Jean-Luc Ploix,[‡] Gérard Dreyfus,[‡] Véronique Nardello-Rataj,[†] and Jean-Marie Aubry[*,†]

[†]Univ. Lille, CNRS, Centrale Lille, ENSCL, Univ. Artois, UMR 8181, UCCS - Unité de Catalyse et de Chimie du Solide, F-59000 Lille, France
[‡]Laboratoire de Chimie Organique, CNRS, ESPCI Paris, PSL Research University, 10 rue Vauquelin, 75005 Paris, France
[§]Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INP, UPS, 31432 Toulouse, France

[*]To whom correspondence should be addressed

Emails: jean-marie.aubry@univ-lille1.fr; arthur.duprat@espci.fr

Supplementary Information Contents:

Table S1. Names, SMILES notations, measured surface tensions values in $mN \cdot m^{-1}$ at 25 °C, molecular formulas, and CAS RN of the 244 molecules of the training/validation set and of the 25 molecules of the test set

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
1	Pentane	CCCCC	15.5	C_5H_{12}	109-66-0
2	Hexane	C(C)CCCC	18.0	C_6H_{14}	110-54-3
3	Heptane	CCCCCCC	19.8	C_7H_{16}	142-82-5
4	Octane	CCCCCCCC	21.1	C_8H_{18}	111-65-9
5	Nonane	CCCCCCCCC	22.4	C_9H_{20}	111-84-2
6	Decane	CCCCCCCCCC	23.4	$C_{10}H_{22}$	124-18-5
7	Undecane	CCCCCCCCCCC	24.2	$C_{11}H_{24}$	1120-21-4
8	Tridecane	CCCCCCCCCCCCC	25.6	$C_{13}H_{28}$	629-50-5
9	Pentadecane	CCCCCCCCCCCCCCC	26.7	$C_{15}H_{32}$	629-62-9
10	Hexadecane	CCCCCCCCCCCCCCCC	27.1	$C_{16}H_{34}$	544-76-3
11	Heptadecane	CCCCCCCCCCCCCCCCC	27.5	$C_{17}H_{36}$	629-78-7
12	Squalane	CC(C)CCCC(C)CCCC(C)CCCCC(C)CCCC(C)CCCC(C)C	29.0	$C_{30}H_{62}$	111-01-3
13	2,2-Dimethylbutane	CC(C)(C)CC	15.8	C_6H_{14}	75-83-2
14	2,3-Dimethylbutane	CC(C)C(C)C	16.9	C_6H_{14}	79-29-8
15	2,2,3-Trimethylbutane	CC(C)(C)C(C)C	18.3	C_7H_{16}	464-06-2
16	2,2-Dimethylpentane	CC(C)(C)CCC	17.5	C_7H_{16}	590-35-2
17	2-Methylhexane	CC(C)CCCC	18.8	C_7H_{16}	591-76-4
18-Test	2,2,3-Trimethylpentane	CC(C)(C)C(C)CC	20.2	C_8H_{18}	564-02-3
19	2,2,4-Trimethylpentane	CC(C)(C)CC(C)C	18.3	C_8H_{18}	540-84-1
20	3-Ethylhexane	CCC(CC)CCC	21.1	C_8H_{18}	619-99-8
21	2,2,4,4-Tetramethylpentane	CC(C)(CC(C)(C)C)C	19.9	C_9H_{20}	1070-87-7
22	Isooctane	CCCCCC(C)C	20.3	C_8H_{18}	592-27-8
23-Test	2-Methyloctane	CCCCCCC(C)C	21.4	C_9H_{20}	3221-61-2
24	Cyclopentane	C1CCCC1	22.0	C_5H_{10}	287-92-3
25	Cyclohexane	C1CCCCC1	24.6	C_6H_{12}	110-82-7
26	Cycloheptane	C1CCCCCC1	27.0	C_7H_{14}	291-64-5
27	Cyclooctane	C1CCCCCCC1	29.3	C_8H_{16}	292-64-8
28	Methylcyclopentane	C(CCC1)(C1)C	21.7	C_6H_{12}	96-37-7
29	Ethylcyclopentane	CCC1CCCC1	23.3	C_7H_{14}	1640-89-7
30-Test	Propylcyclopentane	CCCC1CCCC1	24.4	C_8H_{16}	2040-96-2
31	Isopropylcyclopentane	C(CC1)CC1C(C)C	23.9	C_8H_{16}	3875-51-2
32	Isobutylcyclopentane	C(CC1)CC1CC(C)C	24.3	C_9H_{18}	3788-32-7
33	Methylcyclohexane	C(CCCC1)(C1)C	23.3	C_7H_{14}	108-87-2
34	1,3-Dimethylcyclohexane	CC1CCCC(C)C1	23.3	C_8H_{16}	638-04-0
35	1,2-Dimethylcyclohexane	CC1CCCCC1C	24.4	C_8H_{16}	6876-23-9
36	Propylcyclohexane	CCCC1CCCCC1	25.8	C_9H_{18}	1678-92-8

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
37	Isopropylcyclohexane	C(CCCC1)(C1)C(C)C	26.0	C_9H_{18}	696-29-7
38-Test	Butylcyclohexane	C(CCCC1)(C1)CCCC	26.5	$C_{10}H_{20}$	1678-93-9
39	Isobutylcyclohexane	C(CCC1CC(C)C)CC1	25.3	$C_{10}H_{20}$	1678-98-4
40	*tert*-Butylcyclohexane	C(CCC1C(C)(C)C)CC1	26.2	$C_{10}H_{20}$	3178-22-1
41	Decalin	C12CCCCC1CCCC2	30.6	$C_{10}H_{18}$	493-01-6
42	Cyclopentene	C1=CCCC1	22.0	C_5H_8	142-29-0
43	Methylenecyclopentane	C(CCC1)(C1)=C	23.8	C_6H_{10}	1528-30-9
44	3-Methylcyclopentene	CC1C=CCC1	22.1	C_6H_{10}	1120-62-3
45	Cyclohexene	C1=CCCCC1	26.1	C_6H_{10}	110-83-8
46	1-Methylcyclohex-1-ene	C(=CCCC1)(C1)C	26.0	C_7H_{12}	591-49-1
47	Benzene	C1=CC=CC=C1	28.2	C_6H_6	71-43-2
48	Toluene	C1=CC=CC=C1C	28.0	C_7H_8	108-88-3
49	*o*-Xylene	C1=C(C(=CC=C1)C)C	29.6	C_8H_{10}	95-47-6
50	*m*-Xylene	C1=C(C=CC=C1C)C	28.4	C_8H_{10}	108-38-3
51	*p*-Xylene	C1=CC(=CC=C1C)C	27.8	C_8H_{10}	106-42-3
52	Ethylbenzene	C1=CC=CC=C1CC	28.6	C_8H_{10}	100-41-4
53	1,2,3-Trimethylbenzene	CC1=CC=CC(C)=C1C	28.3	C_9H_{12}	526-73-8
54	1,2,4-Trimethylbenzene	CC1=C(C)C=CC(C)=C1	29.2	C_9H_{12}	95-63-6
55	Mesitylene	CC1=CC(C)=CC(C)=C1	28.1	C_9H_{12}	108-67-8
56	Propylbenzene	CCCC1=CC=CC=C1	28.4	C_9H_{12}	103-65-1
57	*o*-Diethylbenzene	CCC1=C(CC)C=CC=C1	29.8	$C_{10}H_{14}$	135-01-3
58	*m*-Diethylbenzene	CCC1=CC(CC)=CC=C1	28.7	$C_{10}H_{14}$	141-93-5
59-Test	*p*-Diethylbenzene	CCC1=CC=C(CC)C=C1	28.5	$C_{10}H_{14}$	105-05-5
60	Butylbenzene	CCCCC1=CC=CC=C1	28.7	$C_{10}H_{14}$	104-51-8
61	Isobutylbenzene	CC(C)CC1=CC=CC=C1	27.0	$C_{10}H_{14}$	538-93-2
62	*sec*-Butylbenzene	CCC(C1=CC=CC=C1)C	28.0	$C_{10}H_{14}$	135-98-8
63-Test	*tert*-Butylbenzene	CC(C1=CC=CC=C1)(C)C	27.6	$C_{10}H_{14}$	98-06-6
64	*o*-Ethyltoluene	CC1=CC=CC=C1CC	29.7	C_9H_{12}	611-14-3
65	*m*-Ethyltoluene	CC1=CC=CC(CC)=C1	28.5	C_9H_{12}	620-14-4
66	*p*-Ethyltoluene	CC1=CC=C(CC)C=C1	28.3	C_9H_{12}	622-96-8
67	1-Phenyldecane	CCCCCCCCCCC1=CC=CC=C1	30.5	$C_{16}H_{26}$	104-72-3
68	Styrene	C1=CC=CC=C1C=C	31.0	C_8H_8	100-42-5
69	1,2,3,4-Tetrahydronaphthalene	C2=C1CCCCC1=CC=C2	33.2	$C_{10}H_{12}$	119-64-2
70	Methanol	CO	22.1	CH_4O	67-56-1
71	Ethanol	C(C)O	22.2	C_2H_6O	64-17-5
72	1-Propanol	C(CO)C	23.3	C_3H_8O	71-23-8
73	2-Propanol	CC(O)C	20.8	C_3H_8O	67-63-0

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
74	1-Butanol	C(CCO)C	24.2	$C_4H_{10}O$	71-36-3
75	2-Butanol	CC(O)CC	22.1	$C_4H_{10}O$	78-92-2
76	Isobutyl alcohol	CC(C)CO	22.4	$C_4H_{10}O$	78-83-1
77	*tert*-Butyl alcohol	CC(C)(O)C	20.0	$C_4H_{10}O$	75-65-0
78	1-Pentanol	C(CCO)CC	25.0	$C_5H_{12}O$	71-41-0
79	3-Methyl-1-butanol	CC(C)CCO	23.7	$C_5H_{12}O$	123-51-3
80-Test	1-Hexanol	CCCCCCO	25.9	$C_6H_{14}O$	111-27-3
81-Test	4-Methyl-2-pentanol	C(C(C)C)C(C)O	22.6	$C_6H_{14}O$	108-11-2
82	2-Ethyl-1-butanol	C(C(CC)CO)C	24.5	$C_6H_{14}O$	97-95-0
83	Heptanol	CCCCCCCO	26.7	$C_7H_{16}O$	111-70-6
84	1-Octanol	CCCCCCCCO	27.1	$C_8H_{18}O$	111-87-5
85	2-Ethyl-1-hexanol	C(C(CO)CC)CCC	27.1	$C_8H_{18}O$	104-76-7
86	Nonanol	CCCCCCCCCO	27.9	$C_9H_{20}O$	143-08-8
87	Decanol	CCCCCCCCCCO	28.4	$C_{10}H_{22}O$	112-30-1
88	Undecanol	CCCCCCCCCCCO	28.6	$C_{11}H_{24}O$	112-42-5
89	Dodecanol	CCCCCCCCCCCCO	29.4	$C_{12}H_{26}O$	112-53-8
90	Allyl alcohol	C=CCO	25.3	C_3H_6O	107-18-6
91	Propargyl alcohol	C#CCO	35.4	C_3H_4O	107-19-7
92	2-Methoxyethanol	C(OC)CO	30.8	$C_3H_8O_2$	109-86-4
93	Cyclohexanol	C1C(CCCC1)O	32.9	$C_6H_{12}O$	108-93-0
94	Cycloheptanol	OC1CCCCCC1	32.7	$C_7H_{14}O$	502-41-0
95	Benzyl alcohol	OCC1=CC=CC=C1	37.0	C_7H_8O	100-51-6
96	1-Phenyl-1-propanol	CCC(C1=CC=CC=C1)O	33.5	$C_9H_{12}O$	93-54-9
97	*m*-Cresol	C1=C(C)C=CC=C1O	35.7	C_7H_8O	108-39-4
98	1,2-Ethanediol	C(CO)O	47.8	$C_2H_6O_2$	107-21-1
99	1,2-Propanediol	C(C(O)C)O	36.0	$C_3H_8O_2$	57-55-6
100-Test	Trimethylene glycol	OCCCO	45.2	$C_3H_8O_2$	504-63-2
101	1,3-Butanediol	C(C(C)O)CO	37.7	$C_4H_{10}O_2$	107-88-0
102	Diethylene glycol	C(O)COCCO	44.8	$C_4H_{10}O_3$	111-46-6
103-Test	2-Ethoxyethanol	C(C)OCCO	28.6	$C_4H_{10}O_2$	110-80-5
104	Propylene glycol methyl ether	CC(O)COC	27.6	$C_4H_{10}O_2$	107-98-2
105	Ethylene glycol propyl ether	OCCOCCC	29.3	$C_5H_{12}O_2$	2807-30-9
106	2-Butoxyethanol	C(OCCO)CCC	27.4	$C_6H_{14}O_2$	111-76-2
107	Diethylene glycol ethyl ether	OCCOCCOCC	33.5	$C_6H_{14}O_3$	111-90-0
108	Triethylene glycol methyl ether	OCCOCCOCCOC	36.8	$C_7H_{16}O_4$	112-35-6
109	Ethylene glycol hexyl ether	OCCOCCCCCC	27.7	$C_8H_{18}O_2$	112-25-4
110-Test	Diethylene glycol monobutyl ether	C(C)CCOCCOCCO	29.7	$C_8H_{18}O_3$	112-34-5

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
111	Triethylene glycol	OCCOCCOCCO	45.1	$C_6H_{14}O_4$	112-27-6
112	Diethylene glycol hexyl ether	OCCOCCOCCCCCC	29.8	$C_{10}H_{22}O_3$	112-59-4
113	Triethylene glycol *n*-butyl ether	OCCOCCOCCOCCCC	32.0	$C_{10}H_{22}O_4$	143-22-6
114	Glycerol acetate	OC(CO)COC(C)=O	41.4	$C_5H_{10}O_4$	26446-35-5
115	Diethyl ether	C(OCC)C	16.5	$C_4H_{10}O$	60-29-7
116	Butyl methyl ether	COCCCC	19.5	$C_5H_{12}O$	628-28-4
117	*tert*-Butyl methyl ether	COC(C)(C)C	19.1	$C_5H_{12}O$	1634-04-4
118	Ethyl propyl ether	CCCOCC	19.3	$C_5H_{12}O$	628-32-0
119	Di-*n*-propyl ether	CCCOCCC	20.0	$C_6H_{14}O$	111-43-3
120	Diisopropyl ether	CC(OC(C)C)C	17.3	$C_6H_{14}O$	108-20-3
121	Methyl pentyl ether	O(C)CCCCC	21.4	$C_6H_{14}O$	628-80-8
122	Butyl ethyl ether	O(CCCC)CC	20.2	$C_6H_{14}O$	628-81-9
123	Di-*n*-butyl ether	CCCCOCCCC	22.4	$C_8H_{18}O$	142-96-1
124	Dipentyl ether	O(CCCCC)CCCCC	24.4	$C_{10}H_{22}O$	693-65-2
125-Test	Dihexyl ether	O(CCCCCC)CCCCCC	25.4	$C_{12}H_{26}O$	112-58-3
126	Dibenzylether	C1(COCC2=CC=CC=C2)=CC=CC=C1	39.6	$C_{14}H_{14}O$	103-50-4
127	1,2-Propylene oxide	CC1CO1	22.2	C_3H_6O	75-56-9
128	Furan	C1=COC=C1	23.8	C_4H_4O	110-00-9
129	Tetrahydrofuran	C1CCCO1	26.5	C_4H_8O	109-99-9
130	1,2-Dimethoxyethane	COCCOC	24.2	$C_4H_{10}O_2$	110-71-4
131	1,4-Dioxane	C1COCCO1	32.8	$C_4H_8O_2$	123-91-1
132	Diethylene glycol dimethyl ether	COCCOCCOC	29.4	$C_6H_{14}O_3$	111-96-6
133	Tetraethylene glycol dimethyl ether	COCCOCCOCCOCCOC	33.7	$C_{10}H_{22}O_5$	143-24-8
134	Dimethoxymethane	C(OC)OC	20.7	$C_3H_8O_2$	109-87-5
135	1,1-Dimethoxyethane	O(C(OC)C)C	21.0	$C_4H_{10}O_2$	534-15-6
136	Diethoxymethane	O(CC)COCC	20.2	$C_5H_{12}O_2$	462-95-3
137	1,1-Diethoxyethane	O(C(OCC)C)CC	20.9	$C_6H_{14}O_2$	105-57-7
138	1,1-Dibutoxyethane	CC(OCCCC)OCCCC	24.0	$C_{10}H_{22}O_2$	871-22-7
139	Paraldehyde	O(C(OC(O1)C)C)C1C	25.6	$C_6H_{12}O_3$	123-63-7
140	Ethyl orthoformate	CCOC(OCC)OCC	23.1	$C_7H_{16}O_3$	122-51-0
141-Test	Diethylene glycol diethyl ether	CCOCCOCCOCC	26.8	$C_8H_{18}O_3$	112-36-7
142	Anisole	COC1=CC=CC=C1	34.5	C_7H_8O	100-66-3
143	Phenetole	CCOC1=CC=CC=C1	32.4	$C_8H_{10}O$	103-73-1
144	*p*-Methylanisole	COC1=CC=C(C)C=C1	33.5	$C_8H_{10}O$	104-93-8
145	Allyl phenyl ether	C=CCOC1=CC=CC=C1	33.3	$C_9H_{10}O$	1746-13-0
146	Phenyl propyl ether	CCCOC1=CC=CC=C1	31.6	$C_9H_{12}O$	622-85-5
147	Acetaldehyde	CC=O	20.6	C_2H_4O	75-07-0

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
148	*n*-Butyraldehyde	CCCC=O	24.2	C_4H_8O	123-72-8
149	*trans*-Crotonaldehyde	C/C=C/C=O	25.7	C_4H_6O	123-73-9
150	Isovaleraldehyde	O=CCC(C)C	23.2	$C_5H_{10}O$	590-86-3
151	2-Furaldehyde	O=CC(OC=C1)=C1	41.3	$C_5H_4O_2$	98-01-1
152-Test	Benzaldehyde	C1=CC=CC=C1C=O	38.0	C_7H_6O	100-52-7
153	Salicylaldehyde	OC1=CC=CC=C1C=O	42.3	$C_7H_6O_2$	90-02-8
154	Acetone	CC(=O)C	22.7	C_3H_6O	67-64-1
155	2-Butanone	C(C(C)=O)C	24.0	C_4H_8O	78-93-3
156	4-Methyl-2-pentanone	C(C(C)C)C(C)=O	23.5	$C_6H_{12}O$	108-10-1
157	Mesityl oxide	CC(C)=CC(C)=O	27.9	$C_6H_{10}O$	141-79-7
158	Diacetone alcohol	C(C(C)=O)C(C)(O)C	29.6	$C_6H_{12}O_2$	123-42-2
159	5-Methyl-2-hexanone	C(CC(C)=O)C(C)C	25.3	$C_7H_{14}O$	110-12-3
160-Test	2,6-Dimethyl-4-heptanone	C(C(C)C)C(=O)CC(C)C	24.1	$C_9H_{18}O$	108-83-8
161	Cyclopentanone	O=C1CCCC1	32.8	C_5H_8O	120-92-3
162	Cyclohexanone	O=C1CCCCC1	34.3	$C_6H_{10}O$	108-94-1
163	Isophorone	CC1=CC(=O)CC(C1)(C)C	31.2	$C_9H_{14}O$	78-59-1
164	Formic acid	O=CO	38.2	CH_2O_2	64-18-6
165	Acetic acid	CC(O)=O	27.1	$C_2H_4O_2$	64-19-7
166	Propionic acid	O=C(O)CC	26.2	$C_3H_6O_2$	79-09-4
167	Acrylic acid	C=CC(O)=O	28.5	$C_3H_4O_2$	79-10-7
168	Butanoic acid	C(C(O)=O)CC	26.1	$C_4H_8O_2$	107-92-6
169	Isobutyric acid	CC(C)C(O)=O	24.6	$C_4H_8O_2$	79-31-2
170	3-Butenoic acid	OC(=O)CC=C	28.3	$C_4H_6O_2$	625-38-7
171	Methacrylic acid	CC(C(O)=O)=C	29.1	$C_4H_6O_2$	79-41-4
172	Pentanoic acid	O=C(O)CCCC	26.5	$C_5H_{10}O_2$	109-52-4
173-Test	Isovaleric acid	CC(C)CC(O)=O	25.0	$C_5H_{10}O_2$	503-74-2
174	Hexanoic acid	O=C(O)CCCCC	27.2	$C_6H_{12}O_2$	142-62-1
175	Heptanoic acid	O=C(O)CCCCCC	27.7	$C_7H_{14}O_2$	111-14-8
176	Octanoic acid	O=C(O)CCCCCCC	28.7	$C_8H_{16}O_2$	124-07-2
177	Nonanoic acid	O=C(O)CCCCCCCC	29.5	$C_9H_{18}O_2$	112-05-0
178	Oleic acid	CCCCCCCC/C=C\CCCCCCCC(O)=O	32.3	$C_{18}H_{34}O_2$	112-80-1
179	Methyl formate	O=COC	24.3	$C_2H_4O_2$	107-31-3
180	Ethyl formate	O=COCC	23.3	$C_3H_6O_2$	109-94-4
181	Propyl formate	O=COCCC	23.9	$C_4H_8O_2$	110-74-7
182	Isopropyl formate	O=COC(C)C	21.7	$C_4H_8O_2$	625-55-8
183	Butyl formate	O=COCCCC	24.5	$C_5H_{10}O_2$	592-84-7
184	Isobutyl formate	O=COCC(C)C	23.3	$C_5H_{10}O_2$	542-55-2

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
185	Pentyl formate	O=COCCCCC	25.5	$C_6H_{12}O_2$	638-49-3
186	Hexyl formate	O=COCCCCCC	26.4	$C_7H_{14}O_2$	629-33-4
187	Isopentyl formate	O=COCCC(C)C	24.4	$C_6H_{12}O_2$	110-45-2
188	Methyl acetate	CC(OC)=O	24.5	$C_3H_6O_2$	79-20-9
189	Ethyl acetate	C(OC(C)=O)C	23.5	$C_4H_8O_2$	141-78-6
190	Vinyl acetate	CC(OC=C)=O	23.6	$C_4H_6O_2$	108-05-4
191	Allyl acetate	O=C(OCC=C)C	25.8	$C_5H_8O_2$	591-87-7
192	Propyl acetate	CC(OCCC)=O	23.8	$C_5H_{10}O_2$	109-60-4
193-Test	Isopropyl acetate	CC(OC(C)C)=O	21.8	$C_5H_{10}O_2$	108-21-4
194	Isobutyl acetate	CC(OCC(C)C)=O	23.1	$C_6H_{12}O_2$	110-19-0
195	*sec*-Butyl acetate	CC(OC(C)CC)=O	23.0	$C_6H_{12}O_2$	105-46-4
196	*tert*-Butyl acetate	CC(OC(C)(C)C)=O	21.9	$C_6H_{12}O_2$	540-88-5
197	Butyl acetate	C(C)CCOC(C)=O	24.8	$C_6H_{12}O_2$	123-86-4
198	Pentyl acetate	CC(OCCCCC)=O	25.1	$C_7H_{14}O_2$	628-63-7
199	Isopentyl acetate	C(C(C)C)COC(=O)C	24.0	$C_7H_{14}O_2$	123-92-2
200	Hexyl acetate	CC(OCCCCCC)=O	26.0	$C_8H_{14}O_2$	142-92-7
201	Acetic acid octyl ester	CC(OCCCCCCCC)=O	27.3	$C_{10}H_{20}O_2$	112-14-1
202	2-Ethoxyethyl acetate	C(OC(C)=O)COCC	28.0	$C_6H_{12}O_3$	111-15-9
203	Methyl acrylate	C=CC(OC)=O	25.2	$C_4H_6O_2$	96-33-3
204	Methyl propionate	O=C(OC)CC	24.2	$C_4H_8O_2$	554-12-1
205	Methyl butyrate	O=C(OC)CCC	24.5	$C_5H_{10}O_2$	623-42-7
206-Test	Methyl isobutyrate	O=C(OC)C(C)C	23.2	$C_5H_{10}O_2$	547-63-7
207	Ethyl propionate	CCC(OCC)=O	23.8	$C_5H_{10}O_2$	105-37-3
208	Ethyl lactate	O=C(OCC)C(O)C	28.3	$C_5H_{10}O_3$	97-64-3
209	Propyl propionate	CCC(OCCC)=O	24.2	$C_6H_{12}O_2$	106-36-5
210	Ethyl crotonate	O=C(OCC)C=CC	26.6	$C_6H_{10}O_2$	623-70-1
211	Ethyl butyrate	CCCC(OCC)=O	23.9	$C_6H_{12}O_2$	105-54-4
212	Ethyl isobutyrate	CC(C)C(OCC)=O	22.7	$C_6H_{12}O_2$	97-62-1
213	Methyl isovalerate	O=C(OC)CC(C)C	23.7	$C_6H_{12}O_2$	556-24-1
214	Methyl pentanoate	O=C(OC)CCCC	25.2	$C_6H_{12}O_2$	624-24-8
215	Butyl propionate	CCC(OCCCC)=O	24.8	$C_7H_{14}O_2$	590-01-2
216	Propyl butyrate	CCCC(OCCC)=O	24.6	$C_7H_{14}O_2$	105-66-8
217	Methyl hexanoate	CCCCCC(OC)=O	25.9	$C_7H_{14}O_2$	106-70-7
218	Pentanoic acid ethyl ester	CCCCC(OCC)=O	24.7	$C_7H_{14}O_2$	539-82-2
219	Ethyl isovalerate	O=C(OCC)CC(C)C	23.2	$C_7H_{14}O_2$	108-64-5
220	Propyl isobutyrate	O=C(OCCC)C(C)C	23.3	$C_7H_{14}O_2$	644-49-5
221	Butyl butyrate	CCCC(OCCCC)=O	25.2	$C_8H_{16}O_2$	109-21-7

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
222	Ethyl hexanoate	O=C(OCC)CCCCC	25.4	$C_8H_{16}O_2$	123-66-0
223-Test	3-Methylbutanoic acid propyl ester	CC(C)CC(OCCC)=O	23.8	$C_8H_{16}O_2$	557-00-6
224	Methyl heptanoate	O=C(OC)CCCCCC	26.6	$C_8H_{16}O_2$	106-73-0
225	Ethyl heptanoate	O=C(OCC)CCCCCC	26.0	$C_9H_{18}O_2$	106-30-9
226	Methyl octanoate	O=C(OC)CCCCCCC	27.4	$C_9H_{18}O_2$	111-11-5
227	Butyl valerate	O=C(OCCCC)CCCC	25.7	$C_9H_{18}O_2$	591-68-4
228	Ethyl octanoate	O=C(OCC)CCCCCCC	26.8	$C_{10}H_{20}O_2$	106-32-1
229	Methyl nonanoate	O=C(OC)CCCCCCCC	28.7	$C_{10}H_{20}O_2$	1731-84-6
230	Pentyl pentanoate	O=C(OCCCCC)CCCC	25.7	$C_{10}H_{20}O_2$	2173-56-0
231-Test	Methyl decanoate	O=C(OC)CCCCCCCCC	28.0	$C_{11}H_{22}O_2$	110-42-9
232	Methyl dodecanoate	O=C(OC)CCCCCCCCCCC	29.1	$C_{13}H_{26}O_2$	111-82-0
233	Methyl myristate	O=C(OC)CCCCCCCCCCCCC	29.4	$C_{15}H_{30}O_2$	124-10-7
234	Octyl octanoate	O=C(OCCCCCCCC)CCCCCCC	28.3	$C_{16}H_{32}O_2$	2306-88-9
235	Methyl palmitate	O=C(OC)CCCCCCCCCCCCCCC	29.6	$C_{17}H_{34}O_2$	112-39-0
236	Methyl stearate	O=C(OC)CCCCCCCCCCCCCCCCC	30.3	$C_{19}H_{38}O_2$	112-61-8
237-Test	Methyl acetoacetate	O=C(OC)CC(=O)C	32.6	$C_5H_8O_3$	105-45-3
238	Ethyl acetoacetate	O=C(OCC)CC(=O)C	31.9	$C_6H_{10}O_3$	141-97-9
239	Dimethyl maleate	O=C(OC)C=CC(=O)OC	37.7	$C_6H_8O_4$	624-48-6
240	Diethyl maleate	O=C(OCC)C=CC(=O)OCC	32.1	$C_8H_{12}O_4$	141-05-9
241-Test	Dibutyl maleate	O=C(OCCCC)C=CC(=O)OCCCC	30.3	$C_{12}H_{20}O_4$	105-76-0
242	Decanedioic acid dibutyl ester	O=C(OCCCC)CCCCCCCCC(OCCCC)=O	31.6	$C_{18}H_{34}O_4$	109-43-3
243	Propane-1,2,3-triyl triacetate	CC(OCC(OC(C)=O)COC(C)=O)=O	35.9	$C_9H_{14}O_6$	102-76-1
244	1,2,3-tris-Butyryloxy-propane	O=C(OCC(OC(CCC)=O)COC(CCC)=O)CCC	30.5	$C_{15}H_{26}O_6$	60-01-5
245	Tricaprylin	CCCCCCCC(OCC(COC(CCCCCCC)=O)OC(CCCCCCC)=O)=O	29.3	$C_{27}H_{50}O_6$	538-23-8
246	Methyl benzoate	O=C(OC)C1=CC=CC=C1	37.3	$C_8H_8O_2$	93-58-3
247	Methyl salicylate	OC1=CC=CC=C1C(OC)=O	39.2	$C_8H_8O_3$	119-36-8
248-Test	Ethyl benzoate	O=C(OCC)C1=CC=CC=C1	34.6	$C_9H_{10}O_2$	93-89-0
249	Diethyl phthalate	O=C(OCC)C1=CC=CC=C1C(OCC)=O	36.7	$C_{12}H_{14}O_4$	84-66-2
250	Tetramethylsilane	C[Si](C)(C)C	12.3	$C_4H_{12}Si$	75-76-3
251	Triethylsilane	CC[SiH](CC)CC	20.3	$C_6H_{16}Si$	617-86-7
252	Triethoxysilane	CCO[SiH](OCC)OCC	20.8	$C_6H_{16}O_3Si$	998-30-1
253	Tetraethylsilane	CC[Si](CC)(CC)CC	22.5	$C_8H_{20}Si$	631-36-7
254-Test	Hexamethyldisiloxane	C[Si](C)(C)O[Si](C)(C)C	15.4	$C_6H_{18}OSi_2$	107-46-0
255	Hexaethyldisiloxane	CC[Si](CC)(CC)O[Si](CC)(CC)CC	22.0	$C_{12}H_{30}OSi_2$	994-49-0
256	Octamethyltrisiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	16.6	$C_8H_{24}O_2Si_3$	107-51-7
257	Octamethylcyclotetrasiloxane	C[Si]1(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O1	18.2	$C_8H_{24}O_4Si_4$	556-67-2
258	Decamethyltetrasiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	17.4	$C_{10}H_{30}O_3Si_4$	141-62-8

Table S1. (Continued)

No.	Molecule	SMILES notation	Measured ST	MF	CAS RN
259-Test	Decamethylcyclopentasiloxane	C[Si]1(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O1	18.2	$C_{10}H_{30}O_5Si_5$	541-02-6
260	Dodecamethylpentasiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	17.5	$C_{12}H_{36}O_4Si_5$	141-63-9
261	Tetradecamethylhexasiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	17.6	$C_{14}H_{42}O_5Si_6$	107-52-8
262	Hexadecamethylheptasiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	17.6	$C_{16}H_{48}O_6Si_7$	541-01-5
263	Octadecamethyloctasiloxane	C[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)C	18.9	$C_{18}H_{54}O_7Si_8$	556-69-4
264	Tetrapropoxysilane	CCCO[Si](OCCC)(OCCC)OCCC	23.1	$C_{12}H_{28}O_4Si$	682-01-9
265	Dimethyl carbonate	O=C(OC)OC	28.7	$C_3H_6O_3$	616-38-6
266	Propylene carbonate	CC1COC(O1)=O	40.9	$C_4H_6O_3$	108-32-7
267	Diethyl carbonate	O=C(OCC)OCC	26.0	$C_5H_{10}O_3$	105-58-8
268	Dipropyl carbonate	O=C(OCCC)OCCC	26.5	$C_7H_{14}O_3$	623-96-1
269	Acetic anhydride	CC(OC(C)=O)=O	32.0	$C_4H_6O_3$	108-24-7

Table S2. Names, SMILES notations, measured surface tensions in $mN \cdot m^{-1}$ at 23.5 °C, molecular formulas, and CAS RN of the 23 molecules of the cosmetic oil set.

No.	Molecule	SMILES notation	Measured ST (Std. dev.) [a]	MF	CAS RN
1	Tetradecane	CCCCCCCCCCCCCC	26.1 (0.2)	$C_{14}H_{30}$	629-59-4
2	Dicaprylyl ether	O(CCCCCCCC)CCCCCCCC	27.1 (0.2)	$C_{16}H_{34}O$	629-82-3
3	Dodecyl caprylate	CCCCCCCC(OCCCCCCCCCCCC)=O	28.3 (0.2)	$C_{20}H_{40}O_2$	20292-09-5
4	2-Propylheptyl caprylate	CCCCCCCC(OCC(CCC)CCCCC)=O	27.4 (0.2)	$C_{18}H_{36}O_2$	868839-23-0
5	Diisopropyl adipate	O=C(OC(C)C)CCCCC(OC(C)C)=O	28.7 (0.2)	$C_{12}H_{22}O_4$	6938-94-9
6	Dodecane	CCCCCCCCCCCC	24.8 (0.4)	$C_{12}H_{26}$	112-40-3
7	Neopentyl glycol di(2-ethylhexanoate)	CCCCC(CC)C(OCC(C)(C)COC(C(CC)CCCC)=O)=O	28.0 (0.3)	$C_{21}H_{40}O_4$	28510-23-8
8	Isoamyl laurate	CCCCCCCCCCCC(OCCC(C)C)=O	28.3 (0.2)	$C_{17}H_{34}O_2$	6309-51-9
9	Diisoamyl sebacate	O=C(OCCC(C)C)CCCCCCCCC(OCCC(C)C)=O	29.4 (0.4)	$C_{20}H_{38}O_4$	10340-42-8
10	Hemisqualane	CCC(CCCC(CCCC(C)C)C)C	24.9 (0.1)	$C_{15}H_{32}$	3891-98-3
11	2,2,4,6,6-Pentamethylheptane	CC(C)(C)CC(C)CC(C)(C)C	21.6 (0.1)	$C_{12}H_{26}$	13475-82-6
12	2,2,4,4,6,8,8-Heptamethylnonane	CC(C)(C)CC(C)(C)CC(C)CC(C)(C)C	24.2 (0.1)	$C_{16}H_{34}$	4390-04-9
13	Dicaprylyl carbonate	O=C(OCCCCCCCC)OCCCCCCCC	28.8 (0.2)	$C_{17}H_{34}O_3$	1680-31-5
14	Propanediol dicaprylate	CCCCCCCC(OCCCOC(CCCCCCC)=O)=O	27.8 (0.7)	$C_{19}H_{36}O_4$	56519-71-2
15	Isodecyl neopentanoate	CC(C)(C)C(OCCCCCCCCC(C)C)=O	25.4 (0.2)	$C_{15}H_{30}O_2$	60381-61-5
16	2-Ethylhexyl cocoate	CCCCCCCCCCCC(OCC(CC)CCCC)=O	28.7 (0.1)	$C_{20}H_{40}O_2$	92044-87-6
17	Neopentyl glycol diheptanoate	CC(C)(COC(CCCCCC)=O)COC(CCCCCC)=O	28.8 (0.1)	$C_{19}H_{36}O_4$	27841-04-9
18	Isononyl isononanoate	CC(C)CCCCCC(OCCCCCCC(C)C)=O	26.2 (0.1)	$C_{18}H_{36}O_2$	42131-25-9
19	Isopropyl myristate	O=C(OC(C)C)CCCCCCCCCCCCC	28.3 (0.2)	$C_{17}H_{34}O_2$	110-27-0
20	Diethylhexyl adipate	CCCCC(CC)COC(CCCCC(OCC(CC)CCCC)=O)=O	29.5 (0.2)	$C_{22}H_{42}O_4$	103-23-1
21	Isopropyl isostearate	CC(C)CCCCCCCCCCCCCCCC(OC(C)C)=O	27.6 (0.3)	$C_{21}H_{42}O_2$	31478-84-9
22	Triethylhexanoin	O=C(C(CC)CCCC)OCC(COC(C(CC)CCCC)=O)OC(C(CC)CCCC)=O	28.9 (0.3)	$C_{27}H_{50}O_6$	7360-38-5
23	D6	C[Si]1(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O[Si](C)(C)O1	18.8 [b]	$C_{12}H_{36}O_6Si_6$	540-97-6

[a] Means and standard deviations for the surface tension are computed over 3 measurements; [b] for D6 the surface tension value is taken from the literature.

Table S3. Names and sigma-moment descriptors for the 269 molecules of the complete data set.

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
1	Pentane	139.15	7.30	–0.14	0.69	–0.03	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	Hexane	159.25	8.21	–0.09	0.77	–0.02	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	Heptane	179.32	9.16	–0.03	0.87	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	Octane	199.39	10.07	0.01	0.95	0.01	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	Nonane	219.44	10.97	0.07	1.04	0.02	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	Decane	239.55	11.93	0.12	1.13	0.03	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	Undecane	259.50	12.58	0.14	1.16	0.03	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	Tridecane	299.64	14.34	0.24	1.32	0.05	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	Pentadecane	339.83	16.10	0.28	1.47	0.06	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	Hexadecane	359.74	17.34	0.41	1.64	0.09	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	Heptadecane	379.98	18.30	0.46	1.73	0.11	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	Squalane	604.79	30.89	0.26	2.91	0.04	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	2,2-Dimethylbutane	147.94	9.03	–0.17	0.91	–0.05	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	2,3-Dimethylbutane	148.74	8.76	–0.23	0.87	–0.06	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	2,2,3-Trimethylbutane	160.94	10.13	–0.22	1.04	–0.06	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	2,2-Dimethylpentane	167.89	9.93	–0.11	0.99	–0.03	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	2-Methylhexane	174.80	9.31	–0.12	0.88	–0.03	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	2,2,3-Trimethylpentane	178.67	10.90	–0.37	1.14	–0.10	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	2,2,4-Trimethylpentane	181.75	10.90	–0.14	1.09	–0.05	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	3-Ethylhexane	187.06	9.92	–0.21	0.94	–0.05	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	2,2,4,4-Tetramethylpentane	191.96	12.20	–0.07	1.27	–0.04	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	Isooctane	181.75	10.90	–0.14	1.09	–0.05	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	2-Methyloctane	215.01	11.29	–0.01	1.08	–0.01	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	Cyclopentane	121.30	6.02	–0.17	0.52	–0.03	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	Cyclohexane	135.75	5.84	–0.12	0.45	–0.02	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26	Cycloheptane	150.41	6.46	–0.02	0.50	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27	Cyclooctane	163.91	7.29	–0.03	0.58	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
28	Methylcyclopentane	139.68	7.06	–0.15	0.64	–0.03	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
29	Ethylcyclopentane	157.81	7.71	–0.13	0.68	–0.03	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	Propylcyclopentane	177.83	8.60	–0.10	0.76	–0.02	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
31	Isopropylcyclopentane	172.60	8.62	–0.11	0.77	–0.03	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	Isobutylcyclopentane	192.62	9.89	–0.09	0.91	–0.02	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
33	Methylcyclohexane	153.55	7.16	–0.05	0.61	–0.01	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
34	1,3-Dimethylcyclohexane	171.37	8.39	0.03	0.76	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	1,2-Dimethylcyclohexane	168.83	8.20	–0.06	0.73	–0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
36	Propylcyclohexane	191.60	8.79	–0.08	0.75	–0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
37	Isopropylcyclohexane	186.00	8.85	–0.18	0.76	–0.03	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	Butylcyclohexane	211.41	9.64	–0.04	0.82	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39	Isobutylcyclohexane	206.30	10.07	–0.06	0.90	–0.01	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
40	*tert*-Butylcyclohexane	197.65	10.37	–0.14	0.95	–0.03	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
41	Decalin	184.50	7.96	–0.01	0.63	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42	Cyclopentene	116.89	13.38	2.36	4.81	2.29	2.75	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
43	Methylenecyclopentane	135.87	16.16	3.97	6.26	3.78	3.94	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
44	3-Methylcyclopentene	135.52	14.43	2.46	4.98	2.45	2.91	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00
45	Cyclohexene	132.03	14.16	3.76	5.42	3.46	3.48	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
46	1-Methylcyclohex-1-ene	150.84	14.01	3.08	4.54	2.72	2.66	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
47	Benzene	122.42	22.98	–1.27	7.06	–0.95	2.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
48	Toluene	141.81	23.59	0.08	7.05	0.03	2.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
49	*o*-Xylene	157.76	24.79	1.19	7.42	0.93	2.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	*m*-Xylene	161.06	24.17	1.24	7.15	0.94	2.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
51	*p*-Xylene	161.16	23.89	1.18	6.96	0.90	2.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
52	Ethylbenzene	160.64	24.31	0.44	7.06	0.22	2.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
53	1,2,3-Trimethylbenzene	173.76	25.43	2.36	7.80	1.88	3.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
54	1,2,4-Trimethylbenzene	177.08	25.04	2.25	7.48	1.83	2.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
55	Mesitylene	180.28	24.68	2.25	7.31	1.82	2.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
56	Propylbenzene	180.59	25.29	0.67	7.13	0.34	2.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
57	*o*-Diethylbenzene	192.80	25.71	1.95	7.37	1.36	2.72	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
58	*m*-Diethylbenzene	198.62	25.41	2.00	7.12	1.37	2.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
59	*p*-Diethylbenzene	198.81	25.29	1.88	7.06	1.29	2.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
60	Butylbenzene	200.67	26.20	0.80	7.23	0.40	2.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
61	Isobutylbenzene	194.33	25.36	0.37	6.78	0.16	2.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
62	*sec*-Butylbenzene	194.29	26.27	0.56	7.26	0.34	2.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
63	*tert*-Butylbenzene	188.41	26.52	0.44	7.31	0.27	2.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
64	*o*-Ethyltoluene	175.12	25.30	1.49	7.44	1.10	2.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
65	*m*-Ethyltoluene	179.85	24.71	1.58	7.07	1.12	2.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
66	*p*-Ethyltoluene	180.01	24.67	1.55	7.06	1.11	2.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
67	1-Phenyldecane	321.14	31.56	1.09	7.66	0.45	2.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
68	Styrene	155.10	30.25	–2.65	10.22	–1.82	4.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
69	1,2,3,4-Tetrahydronaphthalene	178.00	23.92	2.54	7.01	1.56	2.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
70	Methanol	68.41	45.84	14.22	88.29	36.07	204.74	0.04	3.55	2.19	1.07	–0.02	2.18	1.44	0.72
71	Ethanol	89.37	46.71	18.15	89.73	51.06	213.54	0.04	3.84	2.46	1.29	–0.02	2.02	1.29	0.61
72	1-Propanol	109.16	46.53	17.62	87.26	48.39	206.61	0.04	3.77	2.39	1.24	–0.02	1.96	1.23	0.58
73	2-Propanol	108,19	47,02	20,95	88,20	61,79	213,04	0,04	3,88	2,51	1,36	–0,02	1,86	1,14	0,50

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
74	1-Butanol	128.95	47.02	17.89	86.39	48.99	204.31	0.04	3.66	2.31	1.19	–0.02	1.96	1.23	0.57
75	2-Butanol	126.80	45.58	19.18	82.34	55.88	196.45	0.04	3.63	2.30	1.21	–0.02	1.78	1.07	0.46
76	Isobutyl alcohol	125.95	45.23	15.17	80.20	40.35	185.36	0.03	3.48	2.16	1.08	–0.02	1.91	1.19	0.55
77	*tert*-Butyl alcohol	124.69	47.46	22.33	87.75	69.38	217.14	0.04	4.00	2.62	1.47	–0.02	1.74	1.02	0.43
78	1-Pentanol	149.15	48.04	18.06	86.71	49.06	204.75	0.04	3.62	2.28	1.18	–0.02	1.95	1.22	0.57
79	3-Methyl-1-butanol	144.62	47.82	16.55	85.13	45.83	200.04	0.04	3.71	2.32	1.19	–0.02	2.05	1.29	0.59
80	1-Hexanol	169.17	49.08	17.60	87.56	48.14	207.27	0.04	3.71	2.35	1.22	–0.02	2.00	1.27	0.60
81	4-Methyl-2-pentanol	162.30	46.98	17.44	80.61	51.28	191.25	0.03	3.51	2.21	1.17	–0.02	1.78	1.07	0.47
82	2-Ethyl-1-butanol	159.78	44.99	15.15	76.28	40.88	175.40	0.03	3.39	2.08	1.02	–0.02	1.80	1.10	0.49
83	Heptanol	189.23	49.77	17.89	86.94	48.76	205.45	0.04	3.76	2.36	1.21	–0.02	2.03	1.27	0.59
84	1-Octanol	209.34	50.69	18.10	87.25	49.16	206.10	0.04	3.67	2.32	1.20	–0.02	1.98	1.26	0.59
85	2-Ethyl-1-hexanol	201.60	47.62	13.62	77.68	36.47	177.30	0.03	3.42	2.08	1.00	–0.02	1.92	1.19	0.54
86	Nonanol	229.51	51.95	18.07	88.43	49.45	209.23	0.04	3.80	2.39	1.24	–0.02	2.06	1.29	0.59
87	Decanol	249.54	52.71	18.27	87.60	49.31	206.33	0.04	3.78	2.38	1.23	–0.02	2.04	1.28	0.59
88	Undecanol	269.56	53.69	18.11	87.98	49.12	207.12	0.04	3.82	2.39	1.23	–0.02	2.05	1.29	0.60
89	Dodecanol	289.94	54.58	17.70	88.35	47.86	208.11	0.04	3.75	2.35	1.22	–0.02	2.11	1.33	0.62
90	Allyl alcohol	104.40	53.10	10.62	84.68	23.21	185.47	0.03	3.38	2.00	0.92	–0.02	2.19	1.43	0.72
91	Propargyl alcohol	100.84	65.29	–6.88	95.93	–25.24	193.90	0.03	2.61	1.39	0.49	–0.03	3.45	2.06	1.02
92	2-Methoxyethanol	120.98	61.85	32.58	104.52	77.85	231.12	0.05	5.54	3.24	1.47	–0.02	1.77	1.09	0.50
93	Cyclohexanol	144.87	46.93	22.52	88.03	65.06	214.38	0.04	3.93	2.54	1.38	–0.02	1.81	1.10	0.48
94	Cycloheptanol	158.82	46.18	21.97	84.81	64.32	206.43	0.04	4.00	2.58	1.40	–0.02	1.79	1.05	0.44
95	Benzyl alcohol	151.51	59.65	7.31	81.41	18.00	171.70	0.03	3.14	1.86	0.84	–0.02	2.06	1.35	0.68
96	1-Phenyl-1-propanol	186.11	57.40	7.96	74.51	20.48	156.89	0.03	2.95	1.74	0.80	–0.02	1.98	1.25	0.59
97	*m*-Cresol	151.20	55.84	–19.44	78.76	–76.22	190.73	0.01	1.14	0.34	0.03	–0.03	3.06	2.26	1.50
98	1,2-Ethanediol	99.03	79.61	20.85	149.15	49.36	338.59	0.06	6.18	3.68	1.69	–0.04	3.80	2.45	1.23
99	1,2-Propanediol	117.22	75.01	22.60	136.86	55.54	310.60	0.06	5.91	3.50	1.62	–0.03	3.29	2.07	1.01
100	Trimethylene glycol	116.95	77.16	28.78	149.55	76.23	358.85	0.06	6.18	3.87	1.97	–0.03	3.25	2.14	1.10
101	1,3-Butanediol	134.55	72.73	33.00	142.08	93.12	358.07	0.06	6.03	3.88	2.09	–0.03	2.58	1.72	0.92
102	Diethylene glycol	145.13	87.98	41.44	157.99	103.88	373.92	0.07	7.55	4.57	2.23	–0.03	2.62	1.72	0.93
103	2-Ethoxyethanol	141.85	63.77	34.96	110.58	86.63	252.40	0.06	5.87	3.50	1.66	–0.02	1.69	1.06	0.51
104	Propylene glycol methyl ether	139.32	60.46	34.72	101.72	87.45	230.18	0.05	5.50	3.25	1.55	–0.02	1.57	0.93	0.41
105	Ethylene glycol propyl ether	161.38	59.12	34.15	95.09	83.26	209.45	0.05	5.35	3.13	1.45	–0.01	1.34	0.78	0.34
106	2-Butoxyethanol	181.21	59.29	35.30	94.05	87.10	208.04	0.05	5.17	3.04	1.45	–0.01	1.19	0.69	0.29
107	Diethylene glycol ethyl ether	191.81	76.92	57.69	121.90	143.85	270.86	0.08	7.71	4.54	2.14	–0.01	0.87	0.47	0.17
108	Triethylene glycol methyl ether	223.74	98.53	76.94	156.21	189.78	344.26	0.10	10.46	6.19	2.89	–0.01	0.87	0.46	0.16
109	Ethylene glycol hexyl ether	221.08	58.77	34.48	87.70	83.58	189.54	0.05	5.06	2.94	1.36	–0.01	1.07	0.59	0.22
110	Diethylene glycol monobutyl ether	233.08	80.68	57.67	128.82	144.11	290.44	0.08	7.66	4.60	2.23	–0.01	1.04	0.60	0.24

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
111	Triethylene glycol	198.90	103.92	67.40	168.80	169.38	371.01	0.10	10.08	5.98	2.80	–0.02	1.91	1.07	0.41
112	Diethylene glycol hexyl ether	274.80	82.56	56.01	127.16	137.41	280.96	0.08	7.90	4.68	2.20	–0.01	1.10	0.62	0.25
113	Triethylene glycol *n*-butyl ether	287.98	106.10	79.28	170.37	197.03	382.85	0.11	10.83	6.51	3.12	–0.01	1.26	0.73	0.30
114	Glycerol acetate	168.28	102.13	40.17	151.49	87.63	302.73	0.07	7.27	3.90	1.52	–0.02	2.51	1.59	0.80
115	Diethyl ether	132.20	29.03	26.21	44.17	65.31	103.65	0.03	3.38	2.06	1.03	0.00	0.00	0.00	0.00
116	Butyl methyl ether	151.07	28.32	23.34	38.16	53.78	82.11	0.03	2.87	1.62	0.70	0.00	0.00	0.00	0.00
117	*tert*-Butyl methyl ether	142.02	31.11	26.47	47.08	70.41	114.67	0.03	3.43	2.13	1.09	0.00	0.00	0.00	0.00
118	Ethyl propyl ether	152.06	28.10	23.85	39.44	57.36	89.78	0.03	2.98	1.73	0.83	0.00	0.00	0.00	0.00
119	Di-*n*-propyl ether	171.48	26.97	21.49	35.15	50.76	79.06	0.03	2.60	1.50	0.71	0.00	0.00	0.00	0.00
120	Diisopropyl ether	165.70	31.90	26.76	48.76	75.88	128.06	0.04	3.77	2.47	1.38	0.00	0.00	0.00	0.00
121	Methyl pentyl ether	171.02	29.25	23.56	38.62	54.62	83.75	0.03	2.94	1.68	0.74	0.00	0.00	0.00	0.00
122	Butyl ethyl ether	172.09	28.96	24.01	39.69	57.89	90.85	0.03	3.08	1.79	0.84	0.00	0.00	0.00	0.00
123	Di-*n*-butyl ether	211.61	28.83	21.72	35.13	50.32	77.80	0.03	2.59	1.48	0.68	0.00	0.00	0.00	0.00
124	Dipentyl ether	251.91	30.90	22.29	36.02	51.58	79.90	0.03	2.69	1.55	0.73	0.00	0.00	0.00	0.00
125	Dihexyl ether	291.69	32.04	21.81	35.42	50.90	79.38	0.03	2.66	1.54	0.73	0.00	0.00	0.00	0.00
126	Dibenzylether	254.59	58.08	11.83	36.76	29.70	51.57	0.02	1.94	1.00	0.39	0.00	0.00	0.00	0.00
127	1,2-Propylene oxide	100.92	34.39	23.48	40.89	48.19	69.57	0.03	2.66	1.35	0.43	0.00	0.00	0.00	0.00
128	Furan	103.30	29.34	–3.61	16.43	–3.64	11.68	0.00	0.11	0.00	0.00	0.00	0.20	0.00	0.00
129	Tetrahydrofuran	113.61	31.63	31.35	52.91	79.28	126.35	0.04	3.70	2.32	1.17	0.00	0.00	0.00	0.00
130	1,2-Dimethoxyethane	143.39	50.18	45.14	75.96	106.90	164.49	0.06	5.76	3.36	1.49	0.00	0.00	0.00	0.00
131	1,4-Dioxane	121.30	50.84	40.05	69.48	90.16	135.01	0.05	4.95	2.67	1.01	0.00	0.00	0.00	0.00
132	Diethylene glycol dimethyl ether	196.21	72.31	65.59	110.88	157.22	243.46	0.08	8.52	5.02	2.29	0.00	0.00	0.00	0.00
133	Tetraethylene glycol dimethyl ether	303.75	110.60	95.36	158.33	216.81	327.00	0.11	11.40	6.35	2.62	0.00	0.00	0.00	0.00
134	Dimethoxymethane	120.68	41.42	30.68	51.76	64.32	93.07	0.04	4.05	2.00	0.62	0.00	0.00	0.00	0.00
135	1,1-Dimethoxyethane	138.44	42.92	33.83	56.93	74.48	110.24	0.05	4.67	2.48	0.90	0.00	0.00	0.00	0.00
136	Diethoxymethane	162.45	43.14	33.77	56.24	74.53	110.54	0.05	4.77	2.55	0.96	0.00	0.00	0.00	0.00
137	1,1-Diethoxyethane	179.89	44.88	36.99	62.12	86.10	131.29	0.05	5.21	2.91	1.23	0.00	0.00	0.00	0.00
138	1,1-Dibutoxyethane	265.12	47.88	37.74	60.28	83.34	124.94	0.05	4.92	2.66	1.11	0.00	0.00	0.00	0.00
139	Paraldehyde	173.12	61.75	40.40	70.33	80.83	114.15	0.05	5.39	2.36	0.57	0.00	0.00	0.00	0.00
140	Ethyl orthoformate	213.73	58.71	45.37	75.11	97.34	141.85	0.06	6.44	3.29	1.11	0.00	0.00	0.00	0.00
141	Diethylene glycol diethyl ether	238.45	75.11	70.57	119.96	176.35	279.06	0.08	7.91	4.84	2.40	0.00	0.00	0.00	0.00
142	Anisole	151.80	36.99	5.49	19.66	9.67	17.24	0.01	0.78	0.21	0.01	0.00	0.00	0.00	0.00
143	Phenetole	172.76	37.68	6.99	20.12	11.93	19.23	0.01	0.92	0.29	0.03	0.00	0.00	0.00	0.00
144	*p*-Methylanisole	171.15	37.26	7.35	20.28	12.19	19.14	0.01	0.87	0.24	0.02	0.00	0.00	0.00	0.00
145	Allyl phenyl ether	187.42	44.44	2.82	20.75	5.93	14.85	0.00	0.50	0.12	0.01	0.00	0.00	0.00	0.00
146	Phenyl propyl ether	192.67	37.75	7.43	19.75	12.18	19.06	0.01	0.92	0.29	0.03	0.00	0.00	0.00	0.00
147	Acetaldehyde	83.43	41.75	26.72	50.69	57.13	85.47	0.03	2.94	1.58	0.55	0.00	0.00	0.00	0.00

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
148	*n*-Butyraldehyde	123.07	41.13	27.45	47.92	56.36	81.42	0.03	2.85	1.47	0.49	0.00	0.00	0.00	0.00
149	*trans*-Crotonaldehyde	119.10	49.98	35.44	68.11	84.04	130.96	0.04	3.71	2.27	1.01	0.00	0.00	0.00	0.00
150	Isovaleraldehyde	139.16	40.74	26.41	46.06	53.96	77.75	0.03	2.87	1.48	0.48	0.00	0.00	0.00	0.00
151	2-Furaldehyde	128.14	60.06	23.99	74.65	61.41	124.04	0.03	3.47	1.94	0.76	0.00	0.65	0.06	0.00
152	Benzaldehyde	145.55	48.33	20.56	49.50	47.47	75.72	0.03	2.56	1.28	0.40	0.00	0.00	0.00	0.00
153	Salicylaldehyde	150.83	48.91	12.50	41.16	26.21	48.91	0.02	1.88	0.64	0.11	0.00	0.02	0.00	0.00
154	Acetone	103.34	47.39	35.96	67.50	86.00	134.14	0.04	3.72	2.32	1.07	0.00	0.00	0.00	0.00
155	2-Butanone	122.53	45.40	34.29	62.20	79.75	122.25	0.04	3.61	2.19	0.96	0.00	0.00	0.00	0.00
156	4-Methyl-2-pentanone	157.81	45.30	32.96	59.32	76.03	116.12	0.03	3.46	2.06	0.89	0.00	0.00	0.00	0.00
157	Mesityl oxide	153.82	46.87	34.05	63.16	83.49	131.46	0.03	2.84	1.76	0.91	0.00	0.00	0.00	0.00
158	Diacetone alcohol	162.91	69.00	46.41	106.21	124.38	240.93	0.06	5.73	3.59	1.85	–0.01	0.53	0.34	0.16
159	5-Methyl-2-hexanone	178.04	48.48	35.62	63.83	82.70	126.81	0.04	3.60	2.20	0.99	0.00	0.00	0.00	0.00
160	2,6-Dimethyl-4-heptanone	212.38	42.75	28.92	49.76	63.12	93.14	0.03	3.19	1.77	0.66	0.00	0.00	0.00	0.00
161	Cyclopentanone	125.89	46.94	37.88	65.69	85.83	129.51	0.04	3.59	2.20	0.99	0.00	0.00	0.00	0.00
162	Cyclohexanone	140.59	46.59	41.01	69.07	94.93	145.12	0.04	3.92	2.50	1.21	0.00	0.00	0.00	0.00
163	Isophorone	185.92	60.33	53.11	98.03	143.69	235.69	0.05	4.65	3.24	1.91	0.00	0.00	0.00	0.00
164	Formic acid	72.47	69.63	–22.35	136.25	–119.43	369.80	0.02	2.32	0.97	0.21	–0.04	3.75	2.80	2.05
165	Acetic acid	93.19	67.11	–1.24	118.04	–46.62	283.94	0.03	2.89	1.39	0.43	–0.03	3.11	2.36	1.63
166	Propionic acid	112.52	65.46	–0.32	113.39	–44.85	272.69	0.03	2.97	1.37	0.40	–0.03	3.17	2.39	1.63
167	Acrylic acid	107.65	66.39	–8.59	116.30	–66.82	287.83	0.02	2.44	1.07	0.28	–0.03	3.22	2.43	1.68
168	Butanoic acid	132.44	66.14	–0.07	113.64	–45.00	274.24	0.03	2.93	1.36	0.40	–0.03	2.96	2.24	1.54
169	Isobutyric acid	129.79	65.02	–0.19	114.86	–48.48	284.27	0.03	2.90	1.33	0.39	–0.03	3.25	2.46	1.69
170	3-Butenoic acid	127.76	73.03	–6.40	117.19	–61.65	282.08	0.02	2.59	1.15	0.30	–0.03	3.09	2.35	1.64
171	Methacrylic acid	125.05	62.33	–5.53	106.87	–55.38	261.21	0.02	2.54	1.17	0.30	–0.03	2.97	2.25	1.55
172	Pentanoic acid	152.50	66.39	0.61	111.81	–42.36	267.56	0.03	2.99	1.37	0.40	–0.03	2.99	2.26	1.54
173	Isovaleric acid	148.12	65.02	–1.57	111.06	–48.05	269.10	0.03	2.79	1.26	0.35	–0.03	3.09	2.33	1.60
174	Hexanoic acid	172.62	67.68	0.73	113.06	–42.57	270.63	0.03	2.98	1.38	0.41	–0.03	3.01	2.27	1.55
175	Heptanoic acid	192.61	68.19	0.73	112.34	–42.30	268.69	0.03	2.98	1.39	0.41	–0.03	3.10	2.33	1.59
176	Octanoic acid	212.75	69.28	0.94	112.55	–42.08	268.59	0.03	3.01	1.39	0.41	–0.03	3.21	2.43	1.65
177	Nonanoic acid	232.86	70.19	0.86	112.93	–42.48	269.50	0.03	2.94	1.34	0.38	–0.03	2.98	2.25	1.55
178	Oleic acid	405.16	83.98	4.64	113.88	–37.40	258.97	0.03	2.88	1.29	0.36	–0.03	2.95	2.22	1.51
179	Methyl formate	94.13	46.83	20.86	48.66	40.44	66.64	0.02	2.31	0.95	0.20	0.00	0.14	0.00	0.00
180	Ethyl formate	115.18	47.58	24.22	50.19	46.38	71.88	0.02	2.51	1.07	0.25	0.00	0.11	0.00	0.00
181	Propyl formate	134.73	46.72	24.01	48.29	44.75	68.71	0.02	2.40	1.01	0.24	0.00	0.11	0.00	0.00
182	Isopropyl formate	132.56	46.59	24.90	49.08	47.24	71.53	0.02	2.56	1.09	0.28	0.00	0.10	0.00	0.00
183	Butyl formate	155.29	48.25	24.83	49.54	46.52	71.37	0.03	2.62	1.13	0.26	0.00	0.12	0.00	0.00
184	Isobutyl formate	151.83	47.03	23.81	48.45	45.61	70.41	0.02	2.54	1.10	0.26	0.00	0.12	0.00	0.00

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
185	Pentyl formate	175.33	49.07	24.88	49.32	46.20	70.59	0.03	2.62	1.10	0.25	0.00	0.11	0.00	0.00
186	Hexyl formate	194.90	49.03	24.47	48.30	45.13	68.84	0.02	2.51	1.06	0.24	0.00	0.11	0.00	0.00
187	Isopentyl formate	170.89	49.18	24.42	49.21	45.79	70.32	0.02	2.55	1.07	0.25	0.00	0.11	0.00	0.00
188	Methyl acetate	114.75	49.14	30.67	56.97	62.70	92.29	0.03	2.95	1.47	0.47	0.00	0.00	0.00	0.00
189	Ethyl acetate	135.30	49.19	32.75	57.61	66.33	95.93	0.03	3.07	1.54	0.51	0.00	0.00	0.00	0.00
190	Vinyl acetate	130.55	45.45	15.57	38.28	28.70	45.82	0.02	1.67	0.59	0.10	0.00	0.01	0.00	0.00
191	Allyl acetate	150.26	55.70	27.38	55.56	55.00	82.10	0.03	2.75	1.27	0.36	0.00	0.00	0.00	0.00
192	Propyl acetate	155.32	49.25	32.92	57.04	66.09	95.19	0.03	3.16	1.58	0.52	0.00	0.00	0.00	0.00
193	Isopropyl acetate	152.95	48.13	32.21	55.71	64.80	93.03	0.03	3.10	1.51	0.49	0.00	0.00	0.00	0.00
194	Isobutyl acetate	171.59	48.64	31.96	55.73	65.06	94.04	0.03	3.10	1.57	0.51	0.00	0.00	0.00	0.00
195	*sec*-Butyl acetate	171.01	47.59	32.12	55.11	64.97	93.31	0.03	3.03	1.52	0.51	0.00	0.00	0.00	0.00
196	*tert*-Butyl acetate	166.01	45.24	29.62	51.93	61.80	90.26	0.03	2.82	1.44	0.53	0.00	0.00	0.00	0.00
197	Butyl acetate	175.13	50.00	33.16	57.21	66.57	95.87	0.03	3.09	1.56	0.52	0.00	0.00	0.00	0.00
198	Pentyl acetate	195.22	50.84	33.14	56.97	66.10	94.97	0.03	3.07	1.54	0.51	0.00	0.00	0.00	0.00
199	Isopentyl acetate	190.76	51.38	33.26	58.09	67.80	98.19	0.03	3.13	1.60	0.54	0.00	0.00	0.00	0.00
200	Hexyl acetate	215.63	52.74	34.59	59.60	70.23	101.67	0.03	3.29	1.69	0.57	0.00	0.00	0.00	0.00
201	Acetic acid octyl ester	255.44	53.64	33.46	57.48	66.53	95.59	0.03	3.16	1.57	0.51	0.00	0.00	0.00	0.00
202	2-Ethoxyethyl acetate	188.26	70.93	48.82	86.85	103.42	152.86	0.05	5.20	2.64	0.91	0.00	0.00	0.00	0.00
203	Methyl acrylate	129.22	48.07	23.78	49.62	47.60	71.95	0.02	2.56	1.09	0.25	0.00	0.01	0.00	0.00
204	Methyl propionate	134.10	47.28	29.78	52.26	58.09	82.62	0.03	3.04	1.37	0.36	0.00	0.00	0.00	0.00
205	Methyl butyrate	154.12	47.90	29.99	52.01	58.08	82.26	0.03	3.08	1.41	0.39	0.00	0.00	0.00	0.00
206	Methyl isobutyrate	151.77	46.08	30.15	51.38	59.05	83.68	0.03	2.89	1.37	0.42	0.00	0.00	0.00	0.00
207	Ethyl propionate	154.65	47.66	32.15	54.07	63.43	89.68	0.03	3.11	1.50	0.46	0.00	0.00	0.00	0.00
208	Ethyl lactate	162.81	66.99	33.69	84.31	69.29	150.64	0.04	4.37	2.17	0.76	–0.01	0.53	0.34	0.17
209	Propyl propionate	175.11	48.60	33.31	55.18	65.68	92.67	0.03	3.34	1.63	0.50	0.00	0.00	0.00	0.00
210	Ethyl crotonate	170.71	51.93	33.00	58.80	68.07	98.48	0.03	3.25	1.60	0.49	0.00	0.00	0.00	0.00
211	Ethyl butyrate	174.90	48.65	32.80	54.50	64.37	90.63	0.03	3.28	1.57	0.47	0.00	0.00	0.00	0.00
212	Ethyl isobutyrate	172.75	47.05	33.05	54.15	65.79	92.84	0.03	3.10	1.55	0.51	0.00	0.00	0.00	0.00
213	Methyl isovalerate	169.68	47.43	29.10	50.88	56.90	80.80	0.03	2.93	1.34	0.37	0.00	0.00	0.00	0.00
214	Methyl pentanoate	174.13	48.78	30.30	52.40	58.76	83.39	0.03	3.00	1.36	0.38	0.00	0.00	0.00	0.00
215	Butyl propionate	195.13	49.13	33.12	54.46	64.54	90.58	0.03	3.26	1.58	0.47	0.00	0.00	0.00	0.00
216	Propyl butyrate	195.17	49.09	33.41	54.52	65.04	91.18	0.03	3.28	1.59	0.48	0.00	0.00	0.00	0.00
217	Methyl hexanoate	194.15	49.60	30.43	52.46	58.89	83.48	0.03	3.14	1.45	0.40	0.00	0.00	0.00	0.00
218	Pentanoic acid ethyl ester	194.96	49.68	33.22	54.96	65.02	91.45	0.03	3.28	1.57	0.47	0.00	0.00	0.00	0.00
219	Ethyl isovalerate	190.26	47.72	31.10	51.87	60.71	85.13	0.03	3.05	1.43	0.42	0.00	0.00	0.00	0.00
220	Propyl isobutyrate	192.57	47.00	33.15	53.29	65.06	91.11	0.03	3.08	1.54	0.50	0.00	0.00	0.00	0.00
221	Butyl butyrate	214.29	48.67	32.37	52.65	62.29	86.70	0.03	3.14	1.47	0.42	0.00	0.00	0.00	0.00

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
222	Ethyl hexanoate	214.75	49.89	32.86	54.26	64.14	90.13	0.03	3.16	1.51	0.45	0.00	0.00	0.00	0.00
223	3-Methylbutanoic acid propyl ester	208.34	45.84	31.49	50.91	62.36	87.48	0.03	2.84	1.43	0.47	0.00	0.00	0.00	0.00
224	Methyl heptanoate	214.32	50.60	30.72	52.86	59.52	84.47	0.03	3.00	1.34	0.36	0.00	0.00	0.00	0.00
225	Ethyl heptanoate	235.13	51.31	33.28	54.81	64.66	90.72	0.03	3.23	1.53	0.45	0.00	0.00	0.00	0.00
226	Methyl octanoate	234.28	51.44	30.62	52.79	59.24	84.11	0.03	3.14	1.46	0.42	0.00	0.00	0.00	0.00
227	Butyl valerate	234.99	50.65	33.71	54.57	65.18	91.21	0.03	3.37	1.62	0.49	0.00	0.00	0.00	0.00
228	Ethyl octanoate	254.93	51.32	32.36	53.21	62.08	86.62	0.03	3.12	1.46	0.41	0.00	0.00	0.00	0.00
229	Methyl nonanoate	254.52	52.31	30.73	52.88	59.12	83.62	0.03	3.01	1.36	0.37	0.00	0.00	0.00	0.00
230	Pentyl pentanoate	254.47	50.58	32.78	53.02	62.71	87.22	0.03	3.08	1.44	0.41	0.00	0.00	0.00	0.00
231	Methyl decanoate	274.60	53.34	30.62	52.78	58.74	83.06	0.03	3.14	1.45	0.40	0.00	0.00	0.00	0.00
232	Methyl dodecanoate	314.57	55.06	30.83	53.01	59.06	83.61	0.03	3.09	1.42	0.39	0.00	0.00	0.00	0.00
233	Methyl myristate	354.92	56.61	30.89	53.02	58.80	83.01	0.03	3.07	1.40	0.38	0.00	0.00	0.00	0.00
234	Octyl octanoate	370.10	52.89	31.73	50.86	60.53	83.81	0.03	2.96	1.36	0.38	0.00	0.00	0.00	0.00
235	Methyl palmitate	394.85	58.83	31.13	53.58	59.38	84.07	0.03	3.11	1.44	0.40	0.00	0.00	0.00	0.00
236	Methyl stearate	435.27	60.21	31.14	53.52	59.20	83.74	0.03	3.01	1.36	0.38	0.00	0.00	0.00	0.00
237	Methyl acetoacetate	157.35	80.09	39.71	87.01	79.85	126.18	0.05	4.62	2.03	0.51	0.00	0.13	0.00	0.00
238	Ethyl acetoacetate	177.94	80.48	42.28	87.92	84.17	130.11	0.05	4.64	2.10	0.55	0.00	0.12	0.00	0.00
239	Dimethyl maleate	180.68	81.29	33.35	80.03	62.22	103.81	0.04	3.98	1.42	0.23	0.00	0.29	0.01	0.00
240	Diethyl maleate	221.36	81.24	38.33	80.65	70.29	109.51	0.04	3.94	1.51	0.29	0.00	0.25	0.01	0.00
241	Dibutyl maleate	301.22	82.92	39.65	80.34	71.85	110.71	0.04	4.17	1.59	0.30	0.00	0.25	0.01	0.00
242	Decanedioic acid dibutyl ester	430.98	97.86	67.05	107.67	128.77	179.77	0.06	6.51	3.11	0.92	0.00	0.00	0.00	0.00
243	Propane-1,2,3-triyl triacetate	260.80	121.68	60.92	130.64	122.57	189.73	0.07	6.79	2.97	0.73	0.00	0.06	0.00	0.00
244	1,2,3-tris-Butyryloxy-propane	380.12	121.08	62.39	122.72	119.32	176.74	0.06	6.66	2.73	0.59	0.00	0.05	0.00	0.00
245	Tricaprylin	621.38	133.23	64.57	127.05	123.46	183.57	0.07	7.00	2.94	0.65	0.00	0.06	0.00	0.00
246	Methyl benzoate	175.33	50.96	19.53	45.89	41.81	64.40	0.02	2.27	0.98	0.21	0.00	0.00	0.00	0.00
247	Methyl salicylate	180.77	51.51	12.54	37.73	23.30	40.59	0.02	1.64	0.44	0.03	0.00	0.00	0.00	0.00
248	Ethyl benzoate	195.72	50.70	20.86	45.47	43.67	65.38	0.02	2.19	0.97	0.22	0.00	0.00	0.00	0.00
249	Diethyl phthalate	264.79	82.95	40.50	80.95	78.26	114.32	0.04	4.34	1.76	0.37	0.00	0.02	0.00	0.00
250	Tetramethylsilane	151.68	14.05	0.11	2.33	0.08	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
251	Triethylsilane	187.60	13.60	0.23	2.07	0.24	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
252	Triethoxysilane	227.09	50.65	34.62	54.96	67.23	94.24	0.04	4.38	2.04	0.62	0.00	0.00	0.00	0.00
253	Tetraethylsilane	213.22	13.08	–0.34	1.40	–0.08	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
254	Hexamethyldisiloxane	227.61	30.21	7.59	15.11	13.18	17.77	0.01	0.88	0.30	0.05	0.00	0.00	0.00	0.00
255	Hexaethyldisiloxane	313.80	21.14	1.37	4.88	3.08	4.35	0.00	0.21	0.08	0.02	0.00	0.00	0.00	0.00
256	Octamethyltrisiloxane	299.67	40.27	9.91	20.03	16.83	22.56	0.01	1.10	0.36	0.06	0.00	0.00	0.00	0.00
257	Octamethylcyclotetrasiloxane	323.26	56.24	19.15	35.56	31.34	40.82	0.02	2.19	0.61	0.07	0.00	0.00	0.00	0.00
258	Decamethyltetrasiloxane	373.51	47.78	10.20	22.31	18.03	24.77	0.01	1.23	0.42	0.08	0.00	0.00	0.00	0.00

Table S3. (Continued)

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
259	Decamethylcyclopentasiloxane	384.23	59.09	16.02	31.57	25.21	32.75	0.01	1.61	0.42	0.04	0.00	0.00	0.00	0.00
260	Dodecamethylpentasiloxane	444.93	57.07	12.36	27.03	21.91	30.14	0.01	1.48	0.50	0.09	0.00	0.00	0.00	0.00
261	Tetradecamethylhexasiloxane	525.53	68.88	14.36	31.74	24.62	33.67	0.01	1.56	0.50	0.08	0.00	0.00	0.00	0.00
262	Hexadecamethylheptasiloxane	600.63	78.31	15.47	34.55	25.19	34.04	0.01	1.48	0.45	0.07	0.00	0.00	0.00	0.00
263	Octadecamethyloctasiloxane	666.69	86.83	18.29	40.79	32.14	44.20	0.02	2.19	0.72	0.13	0.00	0.00	0.00	0.00
264	Tetrapropoxysilane	352.98	56.80	33.17	50.35	58.41	78.82	0.04	3.74	1.62	0.46	0.00	0.00	0.00	0.00
265	Dimethyl carbonate	125.68	53.77	27.00	51.41	47.63	67.48	0.02	2.55	0.87	0.14	0.00	0.00	0.00	0.00
266	Propylene carbonate	129.52	66.47	27.48	66.39	55.07	89.20	0.03	2.69	1.18	0.29	0.00	0.05	0.00	0.00
267	Diethyl carbonate	167.66	55.33	32.98	54.68	57.59	77.13	0.03	2.95	1.06	0.20	0.00	0.00	0.00	0.00
268	Dipropyl carbonate	206.65	54.14	32.61	51.96	55.62	73.32	0.03	2.98	1.05	0.19	0.00	0.00	0.00	0.00
269	Acetic anhydride	139.20	65.86	28.63	63.19	51.21	79.01	0.03	3.15	1.08	0.16	0.00	0.04	0.00	0.00

Table S4. Names and sigma-moment descriptors for the 23 molecules of the cosmetic oil set.

No.	Molecule	M_0	M_2	M_3	M_4	M_5	M_6	Hb acc1	Hb acc2	Hb acc3	Hb acc4	Hb don1	Hb don2	Hb don3	Hb don4
1	Tetradecane	319.73	15.36	0.19	1.41	0.05	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	Dicaprylyl ether	371.97	35.55	21.80	35.49	50.61	79.09	0.03	2.68	1.56	0.75	0.00	0.00	0.00	0.00
3	Dodecyl caprylate	455.16	59.20	33.25	53.62	62.82	87.29	0.03	3.07	1.43	0.40	0.00	0.00	0.00	0.00
4	2-Propylheptyl caprylate	403.68	54.86	31.89	52.32	62.95	88.90	0.03	3.04	1.53	0.48	0.00	0.00	0.00	0.00
5	Diisopropyl adipate	304.66	89.13	60.58	99.48	116.45	162.06	0.06	5.94	2.70	0.76	0.00	0.00	0.00	0.00
6	Dodecane	279.76	13.71	0.16	1.28	0.04	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	Neopentyl glycol di(2-ethylhexanoate)	453.90	85.12	56.90	93.45	114.74	162.38	0.05	5.07	2.61	0.85	0.00	0.00	0.00	0.00
8	Isoamyl laurate	389.96	56.94	32.93	54.21	64.31	90.60	0.03	3.16	1.55	0.47	0.00	0.00	0.00	0.00
9	Diisoamyl sebacate	459.07	97.28	64.29	104.98	125.31	175.96	0.06	6.06	2.94	0.87	0.00	0.00	0.00	0.00
10	Hemisqualane	321.75	16.75	–0.05	1.58	–0.02	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	2.2.4.6.6-Pentamethylheptane	242.98	14.87	–0.05	1.50	–0.04	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	2.2.4.4.6.8.8-Heptamethylnonane	303.67	18.46	–0.09	1.91	–0.06	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	Dicaprylyl carbonate	407.78	63.31	34.26	53.73	57.80	76.09	0.03	3.05	1.12	0.22	0.00	0.00	0.00	0.00
14	Propanediol dicaprylate	447.93	94.69	57.17	98.13	111.95	158.44	0.05	5.27	2.49	0.70	0.00	0.00	0.00	0.00
15	Isodecyl neopentanoate	339.86	49.58	29.83	47.98	58.30	81.33	0.03	2.90	1.44	0.43	0.00	0.00	0.00	0.00
16	2-Ethylhexyl cocoate	445.02	56.79	32.19	52.88	63.90	90.68	0.03	3.07	1.56	0.50	0.00	0.00	0.00	0.00
17	Neopentyl glycol diheptanoate	437.01	88.29	55.87	94.69	111.88	158.70	0.06	5.54	2.71	0.77	0.00	0.00	0.00	0.00
18	Isononyl isononanoate	405.00	57.64	32.86	53.56	62.92	87.90	0.03	3.22	1.52	0.45	0.00	0.00	0.00	0.00
19	Isopropyl myristate	392.97	55.83	32.31	52.45	61.55	85.48	0.03	3.15	1.47	0.43	0.00	0.00	0.00	0.00
20	Diethylhexyl adipate	491.19	94.64	59.76	99.40	118.39	167.26	0.05	5.19	2.58	0.77	0.00	0.00	0.00	0.00
21	Isopropyl isostearate	466.56	58.59	31.51	51.61	59.98	83.42	0.03	3.01	1.38	0.38	0.00	0.00	0.00	0.00
22	Triethylhexanoin	574.92	114.15	61.44	111.85	116.00	168.74	0.06	5.82	2.56	0.65	0.00	0.06	0.00	0.00
23	D6	437.35	65.67	17.66	35.46	29.14	38.58	0.02	1.85	0.54	0.08	0.00	0.00	0.00	0.00

Table S5. Measured, neural network estimated, neural network VLOO estimated, graph machine estimated, and graph machine VLOO estimated values of the surface tension in $mN \cdot m^{-1}$ for the 244 molecules of the training/validation set and the 25 molecules of the test set.

No.	Molecule	Measured ST	NN estim. ST [a]	NN VLOO estim. ST [c]	GM estim. ST [d]	GM VLOO estim. ST [f]
1	Pentane	15.5	19.8	20.4	16.2	16.4
2	Hexane	18.0	21.2	21.4	18.1	18.1
3	Heptane	19.8	22.9	23.1	19.9	19.9
4	Octane	21.1	24.6	24.9	21.2	21.2
5	Nonane	22.4	24.9	25.4	22.3	22.3
6	Decane	23.4	24.5	25.6	23.3	23.2
7	Undecane	24.2	23.5	22.6	24.1	24.1
8	Tridecane	25.6	24.8	23.8	25.6	25.6
9	Pentadecane	26.7	26.6	26.6	26.7	26.7
10	Hexadecane	27.1	27.0	27.0	27.1	27.1
11	Heptadecane	27.5	27.7	27.2	27.4	27.3
12	Squalane	29.0	29.1	29.4	28.8	29.6
13	2,2-Dimethylbutane	15.8	16.8	17.0	16.4	16.6
14	2,3-Dimethylbutane	16.9	17.3	17.3	15.9	15.2
15	2,2,3-Trimethylbutane	18.3	16.7	16.5	17.5	17.2
16	2,2-Dimethylpentane	17.5	18.7	18.8	18.3	18.5
17	2-Methylhexane	18.8	21.3	21.4	18.7	18.6
18-Test	2,2,3-Trimethylpentane	20.2	18.1 [b]	-	19.0 [e]	-
19	2,2,4-Trimethylpentane	18.3	19.4	19.5	18.6	18.7
20	3-Ethylhexane	21.1	21.9	22.0	20.8	20.7
21	2,2,4,4-Tetramethylpentane	19.9	19.8	19.8	19.5	19.1
22	Isooctane	20.3	19.4	19.3	20.1	20.1
23-Test	2-Methyloctane	21.4	23.8 [b]	-	21.1 [e]	-
24	Cyclopentane	22.0	21.0	20.6	22.2	22.3
25	Cyclohexane	24.6	24.3	24.4	24.5	24.5
26	Cycloheptane	27.0	25.6	25.4	27.0	27.1
27	Cyclooctane	29.3	25.4	25.0	28.3	27.9
28	Methylcyclopentane	21.7	20.6	20.4	22.0	22.1
29	Ethylcyclopentane	23.3	22.2	22.1	23.2	23.2
30-Test	Propylcyclopentane	24.4	23.8 [b]	-	24.2 [e]	-
31	Isopropylcyclopentane	23.9	22.7	22.6	23.6	23.6
32	Isobutylcyclopentane	24.3	23.5	23.4	24.1	24.0
33	Methylcyclohexane	23.3	23.5	23.5	24.1	24.2
34	1,3-Dimethylcyclohexane	23.3	23.7	23.7	23.6	23.7
35	1,2-Dimethylcyclohexane	24.4	23.3	23.3	24.5	24.6
36	Propylcyclohexane	25.8	26.0	26.0	26.0	26.0
37	Isopropylcyclohexane	26.0	24.4	24.3	25.5	25.4
38-Test	Butylcyclohexane	26.5	26.3 [b]	-	26.8 [e]	-
39	Isobutylcyclohexane	25.3	25.2	25.2	25.9	26.1
40	*tert*-Butylcyclohexane	26.2	23.2	23.0	26.0	26.0
41	Decalin	30.6	27.5	27.0	30.5	30.2
42	Cyclopentene	22.0	20.4	19.7	22.6	22.8
43	Methylenecyclopentane	23.8	25.9	26.6	24.3	24.5
44	3-Methylcyclopentene	22.1	21.3	21.2	22.6	22.8
45	Cyclohexene	26.1	25.6	25.4	25.2	24.9
46	1-Methylcyclohex-1-ene	26.0	25.3	25.1	26.2	26.3
47	Benzene	28.2	25.9	24.8	28.3	28.4
48	Toluene	28.0	28.0	28.0	28.2	28.2
49	*o*-Xylene	29.6	29.4	29.3	29.3	29.3
50	*m*-Xylene	28.4	28.7	28.8	28.4	28.4
51	*p*-Xylene	27.8	28.4	28.5	28.3	28.5
52	Ethylbenzene	28.6	27.8	27.7	28.3	28.2
53	1,2,3-Trimethylbenzene	28.3	30.3	30.6	28.6	29.1
54	1,2,4-Trimethylbenzene	29.2	29.8	29.9	29.1	29.1
55	Mesitylene	28.1	29.5	29.6	28.3	28.4
56	Propylbenzene	28.4	28.3	28.2	28.6	28.6

Table S5. (Continued)

No.	Molecule	Measured ST	NN estim. ST [a]	NN VLOO estim. ST [c]	GM estim. ST [d]	GM VLOO estim. ST [f]
57	*o*-Diethylbenzene	29.8	29.5	29.4	29.1	28.9
58	*m*-Diethylbenzene	28.7	29.0	29.0	28.8	28.9
59-Test	*p*-Diethylbenzene	28.5	28.8 [b]	-	28.4 [e]	-
60	Butylbenzene	28.7	28.3	28.2	29.2	29.4
61	Isobutylbenzene	27.0	27.3	27.3	28.0	28.2
62	*sec*-Butylbenzene	28.0	28.3	28.4	28.3	28.3
63-Test	*tert*-Butylbenzene	27.6	28.7 [b]	-	27.6 [e]	-
64	*o*-Ethyltoluene	29.7	29.3	29.3	29.0	28.8
65	*m*-Ethyltoluene	28.5	28.8	28.8	28.5	28.6
66	*p*-Ethyltoluene	28.3	28.7	28.7	28.1	28.0
67	1-Phenyldecane	30.5	28.6	26.2	30.3	30.7
68	Styrene	31.0	31.2	31.4	30.4	30.2
69	1,2,3,4-Tetrahydronaphthalene	33.2	29.5	28.6	33.6	34.0
70	Methanol	22.1	22.0	21.7	21.8	19.3
71	Ethanol	22.2	21.7	21.2	22.3	22.2
72	1-Propanol	23.3	22.7	22.6	23.5	23.6
73	2-Propanol	20.8	21.9	22.3	21.1	21.2
74	1-Butanol	24.2	23.7	23.6	23.8	23.7
75	2-Butanol	22.1	22.5	22.6	22.2	22.3
76	Isobutyl alcohol	22.4	23.0	23.5	22.1	21.9
77	*tert*-Butyl alcohol	20.0	23.7	25.0	20.8	21.3
78	1-Pentanol	25.0	24.4	24.0	25.1	25.1
79	3-Methyl-1-butanol	23.7	25.0	25.3	23.8	23.8
80-Test	1-Hexanol	25.9	25.6 [b]	-	25.6 [e]	-
81-Test	4-Methyl-2-pentanol	22.6	24.5 [b]	-	24.1 [e]	-
82	2-Ethyl-1-butanol	24.5	23.6	23.2	24.5	24.5
83	Heptanol	26.7	27.3	27.5	26.7	26.7
84	1-Octanol	27.1	27.3	27.4	27.3	27.4
85	2-Ethyl-1-hexanol	27.1	26.4	25.8	26.7	26.6
86	Nonanol	27.9	28.6	28.6	27.8	27.9
87	Decanol	28.4	28.8	28.9	28.4	28.4
88	Undecanol	28.6	29.0	29.1	28.6	28.6
89	Dodecanol	29.4	29.7	29.9	29.3	29.3
90	Allyl alcohol	25.3	26.2	27.0	26.2	27.2
91	Propargyl alcohol	35.4	33.3	32.0	35.2	34.3
92	2-Methoxyethanol	30.8	30.1	29.6	30.0	29.5
93	Cyclohexanol	32.9	29.3	27.6	32.4	32.3
94	Cycloheptanol	32.7	31.9	31.5	33.0	33.4
95	Benzyl alcohol	37.0	36.2	35.7	37.1	37.2
96	1-Phenyl-1-propanol	33.5	32.6	31.7	33.6	33.7
97	*m*-Cresol	35.7	35.6	34.8	35.5	34.5
98	1,2-Ethanediol	47.8	46.5	41.5	47.6	46.8
99	1,2-Propanediol	36.0	36.9	38.7	35.6	34.7
100-Test	Trimethylene glycol	45.2	37.8 [b]	-	46.9 [e]	-
101	1,3-Butanediol	37.7	37.5	37.7	37.4	35.7
102	Diethylene glycol	44.8	44.6	44.8	44.7	44.5
103-Test	2-Ethoxyethanol	28.6	30.9 [b]	-	28.5 [e]	-
104	Propylene glycol methyl ether	27.6	29.3	30.6	27.4	26.9
105	Ethylene glycol propyl ether	29.3	28.5	28.2	28.3	28.0
106	2-Butoxyethanol	27.4	28.1	28.4	28.3	28.5
107	Diethylene glycol ethyl ether	33.5	33.0	32.9	33.2	32.8
108	Triethylene glycol methyl ether	36.8	36.2	36.2	36.7	36.7
109	Ethylene glycol hexyl ether	27.7	27.0	26.5	29.0	29.4
110-Test	Diethylene glycol monobutyl ether	29.7	29.9 [b]	-	29.9 [e]	-
111	Triethylene glycol	45.1	44.9	44.3	45.1	45.2
112	Diethylene glycol hexyl ether	29.8	29.3	28.9	29.9	29.9
113	Triethylene glycol *n*-butyl ether	32.0	32.9	34.1	32.2	32.4
114	Glycerol acetate	41.4	42.0	44.4	41.3	40.7

Table S5. (Continued)

No.	Molecule	Measured ST	NN estim. ST [a]	NN VLOO estim. ST [c]	GM estim. ST [d]	GM VLOO estim. ST [f]
115	Diethyl ether	16.5	18.2	18.8	17.6	18.0
116	Butyl methyl ether	19.5	20.0	20.2	19.5	19.5
117	*tert*-Butyl methyl ether	19.1	17.9	17.6	18.9	18.8
118	Ethyl propyl ether	19.3	19.5	19.5	19.0	19.0
119	Di-*n*-propyl ether	20.0	21.1	21.6	20.4	20.5
120	Diisopropyl ether	17.3	17.4	17.6	17.4	17.4
121	Methyl pentyl ether	21.4	20.3	20.0	21.4	21.4
122	Butyl ethyl ether	20.2	19.5	19.2	20.2	20.2
123	Di-*n*-butyl ether	22.4	23.3	23.7	22.3	22.3
124	Dipentyl ether	24.4	24.4	25.1	23.5	23.2
125-Test	Dihexyl ether	25.4	25.5 [b]	-	24.3 [e]	-
126	Dibenzylether	39.6	39.2	38.7	39.6	39.6
127	1,2-Propylene oxide	22.2	20.7	19.5	22.2	22.1
128	Furan	23.8	25.5	27.8	23.8	24.1
129	Tetrahydrofuran	26.5	25.3	24.0	25.5	24.3
130	1,2-Dimethoxyethane	24.2	25.8	26.2	23.8	23.6
131	1,4-Dioxane	32.8	30.2	29.4	32.7	32.2
132	Diethylene glycol dimethyl ether	29.4	29.3	29.2	29.7	29.9
133	Tetraethylene glycol dimethyl ether	33.7	33.0	32.0	33.7	34.7
134	Dimethoxymethane	20.7	21.6	22.1	21.3	22.0
135	1,1-Dimethoxyethane	21.0	21.4	21.7	21.1	21.2
136	Diethoxymethane	20.2	19.6	19.3	20.4	20.4
137	1,1-Diethoxyethane	20.9	20.1	19.6	21.3	21.5
138	1,1-Dibutoxyethane	24.0	24.3	25.0	23.6	23.2
139	Paraldehyde	25.6	26.5	27.2	25.5	25.5
140	Ethyl orthoformate	23.1	23.9	25.6	23.1	22.9
141-Test	Diethylene glycol diethyl ether	26.8	25.1 [b]	-	27.8 [e]	-
142	Anisole	34.5	33.2	32.8	33.4	33.0
143	Phenetole	32.4	32.2	32.2	32.3	32.3
144	*p*-Methylanisole	33.5	33.3	33.5	33.0	31.9
145	Allyl phenyl ether	33.3	34.1	35.5	33.7	33.8
146	Phenyl propyl ether	31.6	31.0	30.8	32.0	32.1
147	Acetaldehyde	20.6	20.8	20.7	21.4	22.1
148	*n*-Butyraldehyde	24.2	22.6	22.4	23.9	23.8
149	*trans*-Crotonaldehyde	25.7	25.5	25.5	25.5	25.3
150	Isovaleraldehyde	23.2	22.9	22.8	23.5	23.8
151	2-Furaldehyde	41.3	38.1	35.4	41.3	41.1
152-Test	Benzaldehyde	38.0	31.6 [b]	-	36.8 [e]	-
153	Salicylaldehyde	42.3	41.3	39.6	42.3	42.8
154	Acetone	22.7	26.0	26.7	21.9	21.6
155	2-Butanone	24.0	26.0	26.2	23.2	23.0
156	4-Methyl-2-pentanone	23.5	24.9	25.1	23.8	23.9
157	Mesityl oxide	27.9	27.4	26.9	27.4	27.0
158	Diacetone alcohol	29.6	28.8	28.6	30.0	30.7
159	5-Methyl-2-hexanone	25.3	25.3	25.4	25.1	25.0
160-Test	2,6-Dimethyl-4-heptanone	24.1	22.9 [b]	-	24.8 [e]	-
161	Cyclopentanone	32.8	29.5	28.7	32.4	32.3
162	Cyclohexanone	34.3	32.0	30.9	33.6	33.3
163	Isophorone	31.2	32.1	36.0	31.2	31.1
164	Formic acid	38.2	38.4	39.9	38.0	36.4
165	Acetic acid	27.1	26.8	26.7	27.1	27.1
166	Propionic acid	26.2	26.8	26.8	26.0	25.9
167	Acrylic acid	28.5	29.1	29.7	28.5	28.5
168	Butanoic acid	26.1	26.6	26.8	25.9	25.9
169	Isobutyric acid	24.6	25.7	27.0	24.7	24.8
170	3-Butenoic acid	28.3	28.1	27.7	28.8	29.2
171	Methacrylic acid	29.1	28.8	28.5	29.6	30.3
172	Pentanoic acid	26.5	27.1	27.4	26.3	26.2

Table S5. (Continued)

No.	Molecule	Measured ST	NN estim. ST [a]	NN VLOO estim. ST [c]	GM estim. ST [d]	GM VLOO estim. ST [f]
173-Test	Isovaleric acid	25.0	25.6 [b]	-	25.0 [e]	-
174	Hexanoic acid	27.2	27.4	27.2	27.3	27.4
175	Heptanoic acid	27.7	27.8	27.9	27.7	27.8
176	Octanoic acid	28.7	27.9	26.9	28.6	28.5
177	Nonanoic acid	29.5	30.3	31.6	29.2	29.1
178	Oleic acid	32.3	31.7	29.8	32.6	34.8
179	Methyl formate	24.3	24.8	25.9	24.5	24.8
180	Ethyl formate	23.3	24.7	25.5	23.5	23.5
181	Propyl formate	23.9	23.1	22.6	24.0	24.0
182	Isopropyl formate	21.7	22.9	23.1	22.2	22.3
183	Butyl formate	24.5	24.1	24.1	24.5	24.5
184	Isobutyl formate	23.3	24.0	24.1	23.1	23.1
185	Pentyl formate	25.5	25.1	25.0	25.0	24.9
186	Hexyl formate	26.4	26.5	26.4	26.6	26.6
187	Isopentyl formate	24.4	25.2	25.3	24.1	24.1
188	Methyl acetate	24.5	23.5	23.2	23.7	23.4
189	Ethyl acetate	23.5	24.6	24.7	22.9	22.8
190	Vinyl acetate	23.6	25.1	27.0	24.3	24.5
191	Allyl acetate	25.8	26.4	26.5	26.1	26.3
192	Propyl acetate	23.8	24.4	24.4	23.7	23.7
193-Test	Isopropyl acetate	21.8	24.3 [b]	-	22.1 [e]	-
194	Isobutyl acetate	23.1	23.9	24.0	23.1	23.1
195	*sec*-Butyl acetate	23.0	24.3	24.4	23.3	23.3
196	*tert*-Butyl acetate	21.9	23.4	23.5	22.1	22.2
197	Butyl acetate	24.8	24.7	24.7	24.4	24.3
198	Pentyl acetate	25.1	24.9	24.9	24.9	24.9
199	Isopentyl acetate	24.0	24.9	25.1	24.0	24.0
200	Hexyl acetate	26.0	25.6	25.6	26.1	26.2
201	Acetic acid octyl ester	27.3	27.8	27.9	27.4	27.5
202	2-Ethoxyethyl acetate	28.0	28.0	28.3	27.6	27.4
203	Methyl acrylate	25.2	26.2	26.9	26.5	27.2
204	Methyl propionate	24.2	23.2	23.1	23.9	23.9
205	Methyl butyrate	24.5	23.3	23.2	24.5	24.5
206-Test	Methyl isobutyrate	23.2	23.7 [b]	-	23.1 [e]	-
207	Ethyl propionate	23.8	24.3	24.3	23.7	23.7
208	Ethyl lactate	28.3	29.2	29.7	28.5	28.7
209	Propyl propionate	24.2	24.5	24.5	24.2	24.2
210	Ethyl crotonate	26.6	24.5	24.4	26.0	25.7
211	Ethyl butyrate	23.9	24.3	24.3	24.2	24.2
212	Ethyl isobutyrate	22.7	25.0	25.2	23.1	23.2
213	Methyl isovalerate	23.7	23.1	23.1	23.7	23.7
214	Methyl pentanoate	25.2	23.5	23.4	25.2	25.2
215	Butyl propionate	24.8	24.3	24.3	24.8	24.7
216	Propyl butyrate	24.6	24.4	24.3	24.7	24.7
217	Methyl hexanoate	25.9	24.6	24.5	25.8	25.8
218	Pentanoic acid ethyl ester	24.7	24.4	24.4	24.9	25.0
219	Ethyl isovalerate	23.2	23.7	23.7	23.5	23.6
220	Propyl isobutyrate	23.3	24.6	24.9	23.8	24.0
221	Butyl butyrate	25.2	24.7	24.6	25.1	25.0
222	Ethyl hexanoate	25.4	25.0	24.9	25.6	25.6
223-Test	3-Methylbutanoic acid propyl ester	23.8	24.5 [b]	-	23.9 [e]	-
224	Methyl heptanoate	26.6	25.6	25.5	26.6	26.6
225	Ethyl heptanoate	26.0	26.1	26.1	26.1	26.2
226	Methyl octanoate	27.4	26.5	26.3	27.4	27.5
227	Butyl valerate	25.7	25.3	25.2	25.8	25.9
228	Ethyl octanoate	26.8	27.0	27.0	26.9	26.9
229	Methyl nonanoate	28.7	27.7	27.4	28.3	28.1
230	Pentyl pentanoate	25.7	26.6	26.8	25.8	25.8

Table S5. (Continued)

No.	Molecule	Measured ST	NN estim. ST [a]	NN VLOO estim. ST [c]	GM estim. ST [d]	GM VLOO estim. ST [f]
231-Test	Methyl decanoate	28.0	26.7 [b]	-	29.0 [e]	-
232	Methyl dodecanoate	29.1	27.8	27.5	29.2	29.1
233	Methyl myristate	29.4	28.1	27.8	29.7	29.9
234	Octyl octanoate	28.3	29.1	29.4	27.9	27.5
235	Methyl palmitate	29.6	28.3	28.0	30.1	30.5
236	Methyl stearate	30.3	29.2	28.7	30.6	31.1
237-Test	Methyl acetoacetate	32.6	30.8 [b]	-	32.8 [e]	-
238	Ethyl acetoacetate	31.9	32.2	32.1	31.6	31.2
239	Dimethyl maleate	37.7	38.9	40.5	37.5	36.0
240	Diethyl maleate	32.1	34.0	35.7	32.4	33.5
241-Test	Dibutyl maleate	30.3	30.6 [b]	-	31.8 [e]	-
242	Decanedioic acid dibutyl ester	31.6	31.8	31.7	31.5	31.8
243	Propane-1,2,3-triyl triacetate	35.9	35.2	32.8	35.9	35.3
244	1,2,3-tris-Butyryloxy-propane	30.5	30.1	30.1	30.2	29.7
245	Tricaprylin	29.3	29.4	29.9	29.6	30.4
246	Methyl benzoate	37.3	37.3	37.5	36.6	35.6
247	Methyl salicylate	39.2	39.5	39.8	39.4	40.9
248-Test	Ethyl benzoate	34.6	32.8 [b]	-	34.5 [e]	-
249	Diethyl phthalate	36.7	35.4	34.4	36.9	36.6
250	Tetramethylsilane	12.3	15.8	17.4	12.8	14.9
251	Triethylsilane	20.3	19.3	19.1	19.9	19.6
252	Triethoxysilane	20.8	22.3	22.8	20.9	20.5
253	Tetraethylsilane	22.5	20.9	20.4	22.1	21.4
254-Test	Hexamethyldisiloxane	15.4	26.3 [b]	-	15.1 [e]	-
255	Hexaethyldisiloxane	22.0	22.0	21.9	22.1	23.3
256	Octamethyltrisiloxane	16.6	18.1	21.6	16.6	16.6
257	Octamethylcyclotetrasiloxane	18.2	19.8	21.0	18.3	19.9
258	Decamethyltetrasiloxane	17.4	17.2	17.9	17.3	16.8
259-Test	Decamethylcyclopentasiloxane	18.2	17.5 [b]	-	18.1 [e]	-
260	Dodecamethylpentasiloxane	17.5	17.9	18.3	17.4	17.1
261	Tetradecamethylhexasiloxane	17.6	17.7	17.6	17.9	18.2
262	Hexadecamethylheptasiloxane	17.6	18.2	18.5	17.9	17.9
263	Octadecamethyloctasiloxane	18.9	18.8	18.3	18.7	18.1
264	Tetrapropoxysilane	23.1	26.0	27.2	23.2	23.1
265	Dimethyl carbonate	28.7	27.6	26.5	28.7	29.9
266	Propylene carbonate	40.9	37.9	35.4	40.6	38.5
267	Diethyl carbonate	26.0	25.8	25.7	26.3	26.5
268	Dipropyl carbonate	26.5	26.7	26.9	26.1	25.5
269	Acetic anhydride	32.0	34.1	35.5	32.2	32.3

All neural network estimated results were obtained by neural networks with five descriptors and 10 hidden neurons (71 parameters). All graph machine estimated results were obtained by graph machines whose node functions are neural networks with seven hidden neurons (95 parameters).

[a] Estimated values of surface tension provided by neural networks when the molecule belongs to the training/validation set, averaged over the 10 models (out of 100) that had the smallest RMSTEs.

[b] Estimated values of surface tension provided by neural networks when the molecule belongs to the test set, averaged over the 10 models (out of 100) that had the smallest VLOO scores.

[c] Virtual Leave-One-Out estimates of surface tension provided by neural networks, averaged over the 10 models (out of 100) that had the smallest VLOO scores. VLOO estimates are computed only for the molecules of the training/validation set.

[d] Estimated values of surface tension provided by graph machines when the molecule belongs to the training/validation set, averaged over the 10 models (out of 100) that had the smallest RMSTEs.

[e] Estimated values of surface tension provided by graph machines when the molecule belongs to the test set, averaged over the 10 models (out of 100) that had the smallest VLOO scores.

[f] VLOO estimates of surface tension provided by graph machines, averaged over the 10 models (out of 100) that had the smallest VLOO scores. As VLOO estimates are used for model selection, they are computed only for the molecules of the training/validation set, and not for the molecules of the test set.

Table S6. Measured, group-contribution Conte method estimated, corresponding-state Pitzer method estimated, and corresponding-state Zuo-Stenby method estimated values of the surface tension in $mN \cdot m^{-1}$ for the 269 molecules of the complete set.

No.	Molecule	Measured ST	GC estimated ST [a]	CSP1 estimated ST [b]	CSP2 estimated ST [c]
1	Pentane	15.5	17.9	16.2	15.6
2	Hexane	18.0	18.6	19.2	18.5
3	Heptane	19.8	19.2	21.1	20.3
4	Octane	21.1	19.8	22.3	21.5
5	Nonane	22.4	20.4	23.2	22.2
6	Decane	23.4	21.0	23.8	22.7
7	Undecane	24.2	21.7	24.2	23.0
8	Tridecane	25.6	22.9	24.7	23.2
9	Pentadecane	26.7	24.1	24.9	23.2
10	Hexadecane	27.1	24.8	25.0	23.2
11	Heptadecane	27.5	25.4	25.1	23.1
12	Squalane	29.0	27.3	25.6	21.4
13	2,2-Dimethylbutane	15.8	16.7	16.6	16.0
14	2,3-Dimethylbutane	16.9	17.7	17.9	17.2
15	2,2,3-Trimethylbutane	18.3	19.5	19.8	19.1
16	2,2-Dimethylpentane	17.5	17.3	18.9	18.2
17	2-Methylhexane	18.8	18.1	19.8	19.1
18-Test	2,2,3-Trimethylpentane	20.2	20.2	21.8	21
19	2,2,4-Trimethylpentane	18.3	16.8	19.3	18.6
20	3-Ethylhexane	21.1	18.8	21.9	21.1
21	2,2,4,4-Tetramethylpentane	19.9	16.7	19.7	19.0
22	Isooctane	20.3	18.7	21.2	20.4
23-Test	2-Methyloctane	21.4	19.3	22.2	21.3
24	Cyclopentane	22.0	21.1	23.2	22.4
25	Cyclohexane	24.6	25.3	24.8	24.0
26	Cycloheptane	27.0	29.6	25.9	25.0
27	Cyclooctane	29.3	33.8	26.5	25.7
28	Methylcyclopentane	21.7	9.9	23.0	22.2
29	Ethylcyclopentane	23.3	20.4	25.2	24.3
30-Test	Propylcyclopentane	24.4	21	25.7	24.8
31	Isopropylcyclopentane	23.9	12.1	24.5	23.6
32	Isobutylcyclopentane	24.3	13.3	24.7	23.7
33	Methylcyclohexane	23.3	10.6	24.1	23.2
34	1,3-Dimethylcyclohexane	23.3	20.6	24.1	23.2
35	1,2-Dimethylcyclohexane	24.4	20.6	24.1	23.2
36	Propylcyclohexane	25.8	25.2	26.3	25.3
37	Isopropylcyclohexane	26.0	10.7	24.3	23.3
38-Test	Butylcyclohexane	26.5	12.4	25.6	24.4
39	Isobutylcyclohexane	25.3	-	-	-
40	*tert*-Butylcyclohexane	26.2	-	-	-
41	Decalin	30.6	-	-	-
42	Cyclopentene	22.0	24.4	23.7	22.9
43	Methylenecyclopentane	23.8	9.5	23.4	22.6
44	3-Methylcyclopentene	22.1	22.1	23.7	22.9
45	Cyclohexene	26.1	28.6	25.6	24.7
46	1-Methylcyclohex-1-ene	26.0	10.7	24.7	23.9
47	Benzene	28.2	29.0	29.8	28.8
48	Toluene	28.0	28.6	30.5	29.4
49	*o*-Xylene	29.6	28.6	29.6	28.5
50	*m*-Xylene	28.4	28.6	29.6	28.5
51	*p*-Xylene	27.8	28.1	30.2	29.1
52	Ethylbenzene	28.6	27.9	31.5	30.4
53	1,2,3-Trimethylbenzene	28.3	-	-	-
54	1,2,4-Trimethylbenzene	29.2	-	-	-
55	Mesitylene	28.1	26.2	28.3	27.3
56	Propylbenzene	28.4	15.1	26.1	24.9

Table S6. (Continued)

No.	Molecule	Measured ST	GC estimated ST [a]	CSP1 estimated ST [b]	CSP2 estimated ST [c]
57	*o*-Diethylbenzene	29.8	27.3	30.9	29.6
58	*m*-Diethylbenzene	28.7	-	-	-
59-Test	*p*-Diethylbenzene	28.5	-	-	-
60	Butylbenzene	28.7	15.7	26.4	25.1
61	Isobutylbenzene	27.0	14.6	25.4	24.3
62	*sec*-Butylbenzene	28.0	14.7	26.3	25.0
63-Test	*tert*-Butylbenzene	27.6	13.9	24.7	23.7
64	*o*-Ethyltoluene	29.7	-	-	-
65	*m*-Ethyltoluene	28.5	28.0	30.4	29.2
66	*p*-Ethyltoluene	28.3	-	-	-
67	1-Phenyldecane	30.5	19.4	26.7	24.7
68	Styrene	31.0	24.2	32.7	31.5
69	1,2,3,4-Tetrahydronaphthalene	33.2	-	35.1	34.0
70	Methanol	22.1	24.1	27.1	26.0
71	Ethanol	22.2	24.7	32.8	31.2
72	1-Propanol	23.3	25.3	35.8	33.9
73	2-Propanol	20.8	21.8	30.9	29.3
74	1-Butanol	24.2	25.9	37.4	35.2
75	2-Butanol	22.1	22.5	34.7	32.6
76	Isobutyl alcohol	22.4	24.8	35.3	33.4
77	*tert*-Butyl alcohol	20.0	21.5	27.8	26.3
78	1-Pentanol	25.0	26.5	38.2	35.8
79	3-Methyl-1-butanol	23.7	25.4	36.4	34.2
80-Test	1-Hexanol	25.9	27.2	38.6	35.9
81-Test	4-Methyl-2-pentanol	22.6	22.7	35.4	33
82	2-Ethyl-1-butanol	24.5	26.2	38.2	35.5
83	Heptanol	26.7	27.8	38.6	35.8
84	1-Octanol	27.1	28.4	38.6	35.5
85	2-Ethyl-1-hexanol	27.1	27.4	38.4	35.2
86	Nonanol	27.9	29.0	38.4	35.1
87	Decanol	28.4	29.6	38.1	34.6
88	Undecanol	28.6	30.3	37.9	34.1
89	Dodecanol	29.4	30.9	37.6	33.7
90	Allyl alcohol	25.3	26.2	36.5	34.6
91	Propargyl alcohol	35.4	16.6	43.0	40.7
92	2-Methoxyethanol	30.8	27.6	36.6	34.7
93	Cyclohexanol	32.9	32.7	45.1	42.4
94	Cycloheptanol	32.7	37.0	44.8	41.9
95	Benzyl alcohol	37.0	34.8	52.0	48.4
96	1-Phenyl-1-propanol	33.5	34.5	45.9	42.5
97	*m*-Cresol	35.7	36.0	50.3	47.7
98	1,2-Ethanediol	47.8	48.0	69.6	63.5
99	1,2-Propanediol	36.0	47.6	66.8	60.4
100-Test	Trimethylene glycol	45.2	33.9	67.4	60.4
101	1,3-Butanediol	37.7	31.1	63.8	56.5
102	Diethylene glycol	44.8	36.8	63.9	56.8
103-Test	2-Ethoxyethanol	28.6	28.2	37.6	35.4
104	Propylene glycol methyl ether	27.6	-	32.7	30.8
105	Ethylene glycol propyl ether	29.3	28.9	38.0	35.6
106	2-Butoxyethanol	27.4	29.5	38.1	35.5
107	Diethylene glycol ethyl ether	33.5	31.3	40.8	37.5
108	Triethylene glycol methyl ether	36.8	33.6	43.8	39.5
109	Ethylene glycol hexyl ether	27.7	30.7	37.9	34.9
110-Test	Diethylene glycol monobutyl ether	29.7	32.5	40.3	36.5
111	Triethylene glycol	45.1	40.4	58.8	51.6
112	Diethylene glycol hexyl ether	29.8	33.7	39.6	35.4
113	Triethylene glycol *n*-butyl ether	32.0	35.5	41.8	36.8
114	Glycerol acetate	41.4	-	69.7	61.1

Table S6. (Continued)

No.	Molecule	Measured ST	GC estimated ST [a]	CSP1 estimated ST [b]	CSP2 estimated ST [c]
115	Diethyl ether	16.5	19.1	16.8	16.2
116	Butyl methyl ether	19.5	19.7	22.0	21.2
117	*tert*-Butyl methyl ether	19.1	17.8	19.3	18.6
118	Ethyl propyl ether	19.3	19.7	20.3	19.5
119	Di-*n*-propyl ether	20.0	20.3	22.4	21.5
120	Diisopropyl ether	17.3	17.0	15.1	14.6
121	Methyl pentyl ether	21.4	20.3	23.7	22.8
122	Butyl ethyl ether	20.2	20.3	22.4	21.5
123	Di-*n*-butyl ether	22.4	21.6	24.8	23.6
124	Dipentyl ether	24.4	22.8	25.8	24.4
125-Test	Dihexyl ether	25.4	24	26.3	24.6
126	Dibenzylether	39.6	39.8	40.2	37.6
127	1,2-Propylene oxide	22.2	29.1	22.1	21.3
128	Furan	23.8	30.3	30.9	29.8
129	Tetrahydrofuran	26.5	23.7	29.0	28.1
130	1,2-Dimethoxyethane	24.2	20.8	25.1	24.1
131	1,4-Dioxane	32.8	30.5	36.5	35.2
132	Diethylene glycol dimethyl ether	29.4	23.8	30.1	28.4
133	Tetraethylene glycol dimethyl ether	33.7	29.8	34.9	31.3
134	Dimethoxymethane	20.7	20.2	22.6	21.8
135	1,1-Dimethoxyethane	21.0	19.8	24.2	23.2
136	Diethoxymethane	20.2	21.5	23.8	22.8
137	1,1-Diethoxyethane	20.9	21.1	24.8	23.6
138	1,1-Dibutoxyethane	24.0	23.6	27.8	25.9
139	Paraldehyde	25.6	-	-	-
140	Ethyl orthoformate	23.1	23.5	28.6	26.8
141-Test	Diethylene glycol diethyl ether	26.8	25.1	29.4	27.4
142	Anisole	34.5	29.8	36.0	34.6
143	Phenetole	32.4	16.2	27.9	26.6
144	*p*-Methylanisole	33.5	-	-	-
145	Allyl phenyl ether	33.3	17.8	28.4	26.9
146	Phenyl propyl ether	31.6	16.9	28.2	26.8
147	Acetaldehyde	20.6	22.6	22.2	21.4
148	*n*-Butyraldehyde	24.2	23.8	28.1	27.0
149	*trans*-Crotonaldehyde	25.7	23.5	27.6	26.6
150	Isovaleraldehyde	23.2	23.4	27.8	26.7
151	2-Furaldehyde	41.3	-	-	-
152-Test	Benzaldehyde	38.0	40.4	42.6	40.8
153	Salicylaldehyde	42.3	25.6	51.0	46.0
154	Acetone	22.7	23.4	23.4	22.6
155	2-Butanone	24.0	23.5	25.4	24.5
156	4-Methyl-2-pentanone	23.5	23.6	26.2	25.2
157	Mesityl oxide	27.9	24.7	27.6	26.5
158	Diacetone alcohol	29.6	28.6	43.0	39.5
159	5-Methyl-2-hexanone	25.3	24.3	26.8	25.7
160-Test	2,6-Dimethyl-4-heptanone	24.1	24.2	26.4	25.2
161	Cyclopentanone	32.8	29.8	45.7	44.0
162	Cyclohexanone	34.3	34.1	44.6	42.9
163	Isophorone	31.2	55.5	37.9	36.2
164	Formic acid	38.2	30.6	54.0	50.8
165	Acetic acid	27.1	25.0	50.3	47.6
166	Propionic acid	26.2	25.6	49.0	46.2
167	Acrylic acid	28.5	26.7	49.3	46.6
168	Butanoic acid	26.1	26.2	47.8	44.8
169	Isobutyric acid	24.6	24.8	40.1	37.9
170	3-Butenoic acid	28.3	27.1	48.1	45.2
171	Methacrylic acid	29.1	25.1	46.4	43.9
172	Pentanoic acid	26.5	26.8	46.6	43.4

Table S6. (Continued)

No.	Molecule	Measured ST	GC estimated ST [a]	CSP1 estimated ST [b]	CSP2 estimated ST [c]
173-Test	Isovaleric acid	25.0	25.7	45.1	42.2
174	Hexanoic acid	27.2	27.4	45.6	42.2
175	Heptanoic acid	27.7	28.1	44.6	41.0
176	Octanoic acid	28.7	28.7	43.6	39.9
177	Nonanoic acid	29.5	29.3	42.8	38.9
178	Oleic acid	32.3	34.0	38.7	32.9
179	Methyl formate	24.3	22.7	22.2	21.4
180	Ethyl formate	23.3	23.4	25.8	24.8
181	Propyl formate	23.9	24.0	27.8	26.7
182	Isopropyl formate	21.7	22.9	26.0	25.1
183	Butyl formate	24.5	24.6	29.0	27.8
184	Isobutyl formate	23.3	23.5	27.5	26.4
185	Pentyl formate	25.5	25.2	29.7	28.3
186	Hexyl formate	26.4	25.9	30.0	28.6
187	Isopentyl formate	24.4	24.1	28.3	27.1
188	Methyl acetate	24.5	22.8	22.3	21.5
189	Ethyl acetate	23.5	23.5	25.1	24.1
190	Vinyl acetate	23.6	24.6	25.7	24.7
191	Allyl acetate	25.8	25.0	27.1	25.9
192	Propyl acetate	23.8	24.1	26.7	25.6
193-Test	Isopropyl acetate	21.8	22	22	21.1
194	Isobutyl acetate	23.1	23.6	26.4	25.3
195	*sec*-Butyl acetate	23.0	22.8	24.2	23.2
196	*tert*-Butyl acetate	21.9	21.9	22.6	21.7
197	Butyl acetate	24.8	24.7	27.8	26.5
198	Pentyl acetate	25.1	25.3	28.4	27.0
199	Isopentyl acetate	24.0	24.2	27.2	25.9
200	Hexyl acetate	26.0	26.0	28.8	27.2
201	Acetic acid octyl ester	27.3	27.2	29.1	27.3
202	2-Ethoxyethyl acetate	28.0	26.5	30.4	28.7
203	Methyl acrylate	25.2	25.2	26.2	25.2
204	Methyl propionate	24.2	23.2	25.4	24.4
205	Methyl butyrate	24.5	23.8	27.0	25.9
206-Test	Methyl isobutyrate	23.2	22.9	22.8	22
207	Ethyl propionate	23.8	23.8	27.0	25.9
208	Ethyl lactate	28.3	28.3	47.5	43.3
209	Propyl propionate	24.2	24.4	28.0	26.7
210	Ethyl crotonate	26.6	25.0	29.3	28.0
211	Ethyl butyrate	23.9	24.4	28.0	26.7
212	Ethyl isobutyrate	22.7	23.5	24.2	23.3
213	Methyl isovalerate	23.7	23.3	26.6	25.5
214	Methyl pentanoate	25.2	24.4	28.0	26.7
215	Butyl propionate	24.8	25.0	28.6	27.2
216	Propyl butyrate	24.6	25.0	28.6	27.2
217	Methyl hexanoate	25.9	25.0	28.6	27.2
218	Pentanoic acid ethyl ester	24.7	25.0	28.6	27.2
219	Ethyl isovalerate	23.2	23.9	27.4	26.1
220	Propyl isobutyrate	23.3	24.1	25.1	24.0
221	Butyl butyrate	25.2	25.7	28.9	27.4
222	Ethyl hexanoate	25.4	25.7	28.9	27.4
223-Test	3-Methylbutanoic acid propyl ester	23.8	24.6	27.8	26.4
224	Methyl heptanoate	26.6	25.7	28.9	27.4
225	Ethyl heptanoate	26.0	26.3	29.1	27.4
226	Methyl octanoate	27.4	26.3	29.1	27.4
227	Butyl valerate	25.7	26.3	29.1	27.4
228	Ethyl octanoate	26.8	26.9	29.2	27.4
229	Methyl nonanoate	28.7	26.9	29.2	27.4
230	Pentyl pentanoate	25.7	26.9	29.2	27.4

Table S6. (Continued)

No.	Molecule	Measured ST	GC estimated ST [a]	CSP1 estimated ST [b]	CSP2 estimated ST [c]
231-Test	Methyl decanoate	28.0	27.5	29.3	27.3
232	Methyl dodecanoate	29.1	28.8	29.2	26.9
233	Methyl myristate	29.4	30.0	29.1	26.5
234	Octyl octanoate	28.3	30.6	29.0	26.3
235	Methyl palmitate	29.6	31.3	29.0	26.0
236	Methyl stearate	30.3	32.5	28.9	25.6
237-Test	Methyl acetoacetate	32.6	31.9	37.6	35.6
238	Ethyl acetoacetate	31.9	32.6	37.2	35.1
239	Dimethyl maleate	37.7	31.7	39.0	36.5
240	Diethyl maleate	32.1	32.9	38.4	35.5
241-Test	Dibutyl maleate	30.3	35.4	36.9	33.3
242	Decanedioic acid dibutyl ester	31.6	37.8	32.9	28.6
243	Propane-1,2,3-triyl triacetate	35.9	37.0	40.0	36.2
244	1,2,3-tris-Butyryloxy-propane	30.5	40.8	40.1	34.5
245	Tricaprylin	29.3	48.2	37.4	29.1
246	Methyl benzoate	37.3	33.3	36.4	34.9
247	Methyl salicylate	39.2	25.7	46.6	42.1
248-Test	Ethyl benzoate	34.6	33.9	35.7	34.1
249	Diethyl phthalate	36.7	-	-	-
250	Tetramethylsilane	12.3	-	13.5	13.0
251	Triethylsilane	20.3	0.0	0.0	0.0
252	Triethoxysilane	20.8	0.0	0.0	0.0
253	Tetraethylsilane	22.5	-	20.3	19.4
254-Test	Hexamethyldisiloxane	15.4	-	-	-
255	Hexaethyldisiloxane	22.0	-	-	-
256	Octamethyltrisiloxane	16.6	-	-	-
257	Octamethylcyclotetrasiloxane	18.2	-	26.0	23.8
258	Decamethyltetrasiloxane	17.4	-	-	-
259-Test	Decamethylcyclopentasiloxane	18.2	-	25.9	23.1
260	Dodecamethylpentasiloxane	17.5	-	-	-
261	Tetradecamethylhexasiloxane	17.6	-	-	-
262	Hexadecamethylheptasiloxane	17.6	-	-	-
263	Octadecamethyloctasiloxane	18.9	-	-	-
264	Tetrapropoxysilane	23.1	-	-	-
265	Dimethyl carbonate	28.7	25.7	29.2	28.0
266	Propylene carbonate	40.9	32.6	58.0	54.9
267	Diethyl carbonate	26.0	27.0	31.2	29.7
268	Dipropyl carbonate	26.5	28.2	31.8	30.0
269	Acetic anhydride	32.0	31.9	35.2	33.4
	RMSE: [d]	-	**5.16**	**8.13**	**6.32**

[a] Estimated values of surface tension provided by Conte group-contribution method when the molecule belongs to the complete set of 269 molecules.

[b] Estimated values of surface tension provided by Pitzer corresponding-states principle method when the molecule belongs to the complete set.

[c] Estimated values of surface tension provided by Zuo-Stenby corresponding-states principle method when the molecule belongs to the complete set.

[d] Root Mean Square Error computed for the molecules of the complete set with all methods.

Table S7. Measured, group-contribution estimated, neural network estimated, and graph machine estimated values of the surface tension in $mN \cdot m^{-1}$ for the 23 molecules of the cosmetic oil set.

No.	Molecule	Measured ST	CSP Estimated ST	NN Estimated ST [a]	GM Estimated ST I [b]	GM Estimated ST II [c]
1	Tetradecane	26.1	23.2	25.8	26.1	26.3
2	Dicaprylyl ether	27.1	24.3	26.3	27.1	27.1
3	Dodecyl caprylate	28.3	25.4	30.5	28.2	28.8
4	2-Propylheptyl caprylate	27.4	25.9	29.8	28.0	28.0
5	Diisopropyl adipate	28.7	29.5	32.9	29.8	29.6
6	Dodecane	24.8	23.1	24.9	25.0	25.0
7	Neopentyl glycol di(2-ethylhexanoate)	28.0	25.3	30.2	28.0	27.2
8	Isoamyl laurate	28.3	25.5	29.9	28.3	29.1
9	Diisoamyl sebacate	29.4	27.0	31.9	31.0	31.2
10	Hemisqualane	24.9	22.6	24.9	25.7	24.6
11	2,2,4,6,6-Pentamethylheptane	21.6	20.6	21.6	23.2	22.5
12	2,2,4,4,6,8,8-Heptamethylnonane	24.2	20.8	24.1	24.7	24.2
13	Dicaprylyl carbonate	28.8	27.3	31.0	28.7	29.0
14	Propanediol dicaprylate	27.8	28.3	34.4	28.6	28.4
15	Isodecyl neopentanoate	25.4	27.3	27.9	26.3	27.1
16	2-Ethylhexyl cocoate	28.7	25.4	30.3	28.5	29.2
17	Neopentyl glycol diheptanoate	28.8	28.0	30.6	27.8	26.9
18	Isononyl isononanoate	26.2	24.8	29.9	27.5	27.1
19	Isopropyl myristate	28.3	25.5	29.6	29.0	29.2
20	Diethylhexyl adipate	29.5	27.6	32.2	30.6	31.3
21	Isopropyl isostearate	27.6	24.3	29.9	29.3	28.9
22	Triethylhexanoin	28.9	26.4	34.6	28.5	29.1
23	D6	18.8	22.4	17.3	18.4	18.2
	RMStE: [d]	-	**2.38**	**2.69**	**0.84**	**0.94**

All CSP estimated results were obtained with ZuoStenby1997 method. All neural network estimated results were obtained by neural networks with five descriptors and 10 hidden neurons (71 parameters). All graph machine estimated results were obtained by graph machines whose node functions are neural networks with seven hidden neurons (95 parameters).

[a] Values of surface tension provided by neural networks averaged over the 25 models (out of 250) that had the smallest VLOO scores.

[b] Values of surface tension provided by graph machines averaged over the 20 models (out of 1,000) that had the smallest VLOO scores; results obtained in running the Docker image "espcigm/demo:BJMA269EMOL", with the seed parameter equal to 799062158, the "hidden" parameter equal to 7, the "init" parameter equal to 1,000, and the "maxrec" parameter equal to 20, see section "graph machine results with Docker".

[c] Values of surface tension provided by graph machines averaged over the 25 models (out of 250) that had the smallest VLOO scores; results obtained in running the Docker image "espcigm/demo:BJMA269EMOL", with the seed parameter equal to 799062158, the "hidden" parameter equal to 7, the "init" parameter equal to 250, and the "maxrec" parameter equal to 25, see section "graph machine results with Docker".

[d] Root Mean Square test Error computed for the 23 molecules of the cosmetic oil set with all methods.

DOCKER CLIENT INSTALLATION

General

Docker is a lightweight virtualization tool used to launch applications inside containers on Windows, MacOS or Linux computers; interactions with the host machine are fully under control, so that, after use, the host machine can be returned exactly to its original state. The use of Docker containers requires the installation of a local client, which depends on the host machine.

More information on how to install the local client on all machines can be found on the following pages:

https://docs.docker.com/engine/getstarted/step_one/

https://docs.docker.com/engine/installation/

MacOS computer

Docker Client Installation

The Macintosh computer should run at least Yosemite OS (Mac OS X.10.3) with a minimum of 4 GB of RAM (works only for computer manufactured after 2010).

Download the Docker application by copying and pasting the following line into a browser URL bar:

https://download.docker.com/mac/stable/Docker.dmg

Open the downloaded dmg file and drag the icon Docker.app onto the Applications icon as suggested. Launch the Docker application from the Applications folder to complete the installation (the system password is required). The presence of a whale icon in the top right status bar indicates that Docker is running and ready to receive commands via a terminal window.

For more information, visit:

https://docs.docker.com/engine/installation/mac/#docker-for-mac

Note: If necessary, allocate more memory to the virtual machine using the Docker whale icon from the top right menu bar, then choosing preferences, moving the cursor to 3 GB (or more) and selecting restart. Make sure that the "Include VM in Time Machine backups" option is unselected in the General section of Docker preferences.

For more information visit the page:

https://docs.docker.com/engine/installation/mac/#docker-for-mac

Uninstalling Docker

Access the Docker menu by clicking on the whale icon located on the finder top right menu bar, select the reset option, and then click on the uninstall box. Confirm this action with the Uninstall Docker choice within the next window. Once all the elements have been removed from the computer, drag the Docker application to the Trash and empty the latter.

Linux computer

Docker Client Installation

Installation instructions can be found for many Linux distributions on those pages:

- CentOS: https://docs.docker.com/engine/installation/linux/centos/
- Debian: https://docs.docker.com/engine/installation/linux/debian/
- Fedora: https://docs.docker.com/engine/installation/linux/fedora/
- Oracle Linux: https://docs.docker.com/engine/installation/linux/oracle/
- Red Hat: https://docs.docker.com/engine/installation/linux/rhel/
- Susa: https://docs.docker.com/engine/installation/linux/suse/
- Ubuntu: https://docs.docker.com/engine/installation/linux/ubuntu/<

- Other Linux Distributions: https://docs.docker.com/engine/installation/linux/other/

Warning: depending on the chosen distribution and settings of the host computer, the requested commands might have to be entered in superuser mode, which is done with the “sudo” prefix at the beginning of the Docker command.

Docker client installation on Ubuntu

The Ubuntu Community Edition distribution is taken as an example.

Prerequisites

The ubuntu distribution must be one of the three 64-bit distributions:

- Zesty 17.4
- Xenial 16.04 (LTS)
- Trusty 14.04 (LTS)

Docker installation

- Update the apt package with the following command:

```
$ sudo apt-get update
```

- Update Install the latest version of Docker with the command:

```
$ sudo apt-get install docker-ce
```

- Check that Docker is installed correctly by downloading and running the image "hello-world":

```
$ sudo docker run hello-world
```

Uninstalling Docker and images

Follow the instructions found at the uninstall Docker section of each distribution web page.

Computer with Windows 10 or later

Docker Client Installation

The computer must run at least Windows 10 64-bit Pro, Enterprise or Education (November 1511 Build 10586 or later) and have a minimum of 4 GB of RAM. The Hyper-V package must be activated. The Docker for Windows installer will activate it if needed, which requires a reboot.

Download the Docker application with the following link:

https://download.docker.com/win/stable/InstallDocker.msi

Open the downloaded msi file for installation of the Docker application. Once installation is complete, Docker launches automatically. For more information visit the following page:

https://docs.docker.com/docker-for-windows/

Note: if needed, more memory (and more CPUs) can be allocated for the computations; access the Docker whale icon in the task bar, choose preferences, and then increase the amount of memory (and the number of CPUs) used for Docker.

Computer with Windows 7 and 8

The Windows OS must be 64-bit, and virtualization must be enabled on the machine (activation to be done in the computer BIOS).

For more information about the virtualization support for Windows, visit:

http://docs.docker.com/toolbox/toolbox_install_windows/

Installing the Docker Toolbox

The latest Docker toolbox application (about 200 MB) can be downloaded from the page:

https://github.com/docker/toolbox/releases/

Start installing the Docker toolbox with the default values; the option "View Shortcuts in File Explorer" must be unselected. The Docker toolbox and virtualbox applications are installed in the "Program Files" directory of the C: drive, whereas the shortcuts for the "Docker Quickstart", "Oracle VM VirtualBox" and "Kinematic (Alpha)" applications are installed on the desktop.

Notes: 1) the Docker toolbox installation must be done within a session with administrator privileges; the choice is then given to install the application for the administrator or for all the computer users.

2) to avoid running out of memory during the docker compilation steps, a 4 Go memory minimum is required for the docker-machine. This tuning can be done via the VM VirtualBox program. Choosing a maximum number of CPUs (or cores) will also speed up considerably the computations.

Launching a test image in a Docker container

Open a "Docker Toolbox" terminal by double-clicking the shortcut "Docker Quickstart". Wait for the prompt, which is the $ sign. At the first launch, this process can take several minutes. To test that the Docker toolbox application is working properly, enter the following command:

```
docker run hello-world
```

If everything goes right the message "Hello from Docker!" appears in the terminal window, followed by some information about the client and the Docker daemon.

Note: using the Docker toolbox application requires a working internet connection.

Uninstalling Docker and images

The protocol detailed below uninstalls Docker and all its components.

Open a Docker toolbox terminal window and issue the command:

```
$ ls docker machine
```

A list of files (xxxNamexxx) is returned by the list command.

Enter the following command to delete the matching image(s):

```
$ rm xxxNamexxx docker machine
```

Note: the kinematic application can also be used to delete images.

Once the file(s) has(ve) been deleted, use the Windows uninstaller application (located in the control panel window) to remove the Docker toolbox application from the system.

For more information, visit the page:

https://www.docker.com/toolbox

GRAPH MACHINE DEMONSTRATION WITH DOCKER CONTAINERS

Loading and launching the Docker image

Open a terminal window, type the following command and confirm it with the enter key (or copy and paste the text below):

```
docker pull espcigm/demo:BJMA269EMOL
```

The image used to create containers is then downloaded. You can check that the image is genuine by comparing its hash code that is generated at the end of the download process with the following:

sha256:ad3c0377f485c8d3419ed11013815a049a40fbf4e6d927ad4898c0f6912bb822

To open a container that will run the graph machine program in interactive mode, type the following text (or copy and paste the text below):

```
docker run -it --rm -v ~/docker:/docker espcigm/demo:BJMA269EMOL
```

Reply to questions for the available choices, for example the number of neurons or the number of trained models (see the next section for detailed examples). To accept the default values, just confirm with the enter key. It is recommended to start with the default values that enable a fairly fast execution of the demonstration. If you make a mistake at this stage, followed by a runtime error, re-enter the above command.

After completion of the demo, the container is automatically deleted. A new demo session can be started with the same command, but within a new container.

An excel file with the xlsm extension is recorded in the result subdirectory of the Docker directory created in the home directory on the host machine (`~/docker/result`, where ~ is a shortcut indicating the path to the connected user's home directory). The output filename is automatically incremented when the same command is executed.

Note: the above command is the minimal command; if for example, the "`-v ~/docker:/docker`" is omitted in the command syntax, no excel file is created on the host machine, the computed results being lost when the container is deleted. See the Advanced Options section for more information about this run command.

Explanations:

This demonstration has three parts:

- Data analysis and creation of the training model
- Training
- Test on a fresh data set, different from the training data set.

Data file analysis and creation of training model

The data file used is the excel file BJMA269EMOL.xlsx

The training data for the 269 molecules of the training set are contained in the DATA cell range. Each line of this file contains a molecule name, its SMILES code and the measured value of its property of interest (here the surface tension, TS in $mN \cdot m^{-1}$). Other data present in the file are ignored.

For each molecule, a graph machine model is automatically generated from the SMILES code. The structure of this model is derived from the 2D-structure of the molecule, and its output estimates the value of the selected property. All models are merged into a module that is trained with the desired property values.

The following monitoring messages are displayed during this phase:

```
running graphmachine
    source file: <container>BJMA269EMOL.xlsx
    data range: DATA
    testfile: <container>BJMA269EMOL.xlsx
continue and create/load training module? ([y])y
create/load training module: yes
please enter number of hidden neurons: [5] y
hidden neurons: 5
sourcing module: 100%|████████████████████████| 269/269 [00:04<00:00, 54.97it/s]
compiling module

module bjma269emol_si3_cn3_5n:
     code module/program 48/48
     type training, multiple
     mark: Model created with Chem_gm 6.0.0.11
     created 22/10/2017 09:20:00
     base: BJMA269EMOL
     property: TS
```

```
     hidden: 5
     parameter count: 59
     modelCount: 269
analysis done

 running graphmachine
      sourcefile: <container>BJMA269EMOL.xlsx
      data range: DATA
      test range: TEST
continue and create/load training module? ([y]/n)
create/load training module: yes
please enter number of hidden neurons: [5]
hidden neurons: 5
sourcing module: 100%|####################| 269/269 [00:04<00:00, 66.04it/s]
compiling module
compilation time: 7.94 s
module bjma269emol_si3_5n:
      type: training, multiple
      created: 16/03/2017 09:24:08
      mark: Model created with Chem_gm 3.5.0.4
      base: BJMA269EMOL
      property: TS
      parameter count: 59
      model count: 269
      hidden: 5
analysis done
```

Training

The training of graph machine models, just as the training of neural networks, involves the minimization of a non-convex cost function with respect to the model parameters. Therefore, a large number of models are generated with different parameter initial values, in order to capitalize on model diversity. The chosen default number of parameter initializations, hence of generated models, is 250. It is possible to change it, up to 1000, 2000 or more, depending on the speed of the machine on which the code is executed.

After training, two important quantities are computed from each model: its root mean squared error (RMSE) and its virtual leave-one-out score (VLOO score). RMSE is the value of the cost function at the end of training; it indicates the accuracy of the estimation of the property of interest for the molecules present in the training set. The VLOO score is an estimator of the ability of the model to make accurate predictions on fresh data (not present in the training set). The training outcomes are ranked in order of increasing "VLOO score" values, since the lower that score, the better the corresponding model for prediction.

The following monitoring messages are displayed during this phase:

```
continue and launch training? ([y])
launch training: yes
give the number of training run: [250]
training count: 250
give the number of epochs in each run: [150]
epochs: 150
seed: 4100967572
training: 100%|██████████████████████████████| 250/250 [00:34<00:00,  7.18it/s]
25 best training results:
     ID  RMSE  VLOO score
0    #6  1.06        1.50
1  #211  0.95        1.51
2   #58  1.06        1.56
3  #219  1.04        1.58
4  #131  1.03        1.60
5  #173  1.06        1.61
6   #19  1.03        1.64
7  #234  1.10        1.69
8  #104  1.08        1.78
```

```
9  #168  1.02        1.83
10  #92  1.22        1.86
11  #81  1.39        1.93
12 #216  0.94        1.94
13  #77  1.11        1.95
14 #182  1.37        1.98
15 #231  1.14        2.02
16  #69  1.28        2.03
17  #51  1.14        2.03
18   #9  1.10        2.03
19 #199  1.22        2.04
20  #98  1.31        2.07
21  #94  1.22        2.07
22 #189  1.27        2.07
23 #115  1.17        2.09
24 #132  1.23        2.09
training done
```

Note: RMSE and VLOO score are expressed in the same unit as the property i.e. force per unit length for the surface tension (the results displayed above are in $mN \cdot m^{-1}$); ID represents the training index considered.

At the beginning of each training, the model parameters are initialized with a pseudorandom number generator. The corresponding initialization data can then be recalled by the graph machine program with the environment variable "SEED". The displayed value for the seed variable is the random seed used (here `4100967572`).

Test

The data file used is BJMA269EMOL.xlsx, and the test data are contained in the TEST cell range. These are the 23 molecules of the cosmetic oil set.

The following monitoring messages are displayed during this phase:

```
continue and launch test? ([y])
launch test: yes
compute modules: 100%|████████████████████████| 23/23 [00:01<00:00, 11.19it/s]
results in file: /docker/result/BJMA269EMOL_5N.xlsm
test done
```

Output File

The output file has an xlsm (not xlsx) extension because it contains a macro used to close the file without getting the excel popup warning. As a result, any modification made to this file has to be saved before closing. It can be read with Excel 2010-2011 (or more recent) or with LibreOffice 4.3.4 (or more recent).

The file contains six sheets:

- training summary;
- training results;
- test results;
- training #xxx chart(yy);
- VLOO #xxx chart(yy);
- Test result chart(yy);

where xxx is the ID of the model that has the smallest VLOO score and yy is the number of models selected for the computation of the mean.

Training summary

This sheet displays the results obtained with the twenty-five models having the smallest VLOO scores. The training results are ranked in order of increasing value of the VLOO score.

For each trained model, the following quantities are displayed:

- the identification number;
- the root mean square error (RMSE);

- the Virtual Leave-One-Out score (VLOO), which is an estimator of the root mean squared estimation error that would be obtained by the model on molecules that are not present in the training set.

Training results

This sheet displays the results for the top twenty-five trainings for each molecule (selected criterion: the twenty-five smallest VLOO scores).

For each molecule and each training, the following quantities are displayed:

- the measured value of the surface tension;
- the estimated value of the surface tension;
- the estimation error (difference between the estimated value and the measured value of the surface tension);
- the leverage;
- the virtual leave-one-out estimation of the surface tension of the molecule, i.e. the estimation of the value of the surface tension that the model would have made if the molecule had not been present in the training set.

Below the results, the values returned by five Excel formulas are displayed:

- line SUM: the sum of the leverages computed after training, equal to the number of parameters of the model;
- line VLOO score: the Virtual Leave-One-Out score computed after training;
- line RMSE: the training Root Mean Squared Error;
- line Pearson: the Pearson correlation coefficient R between the measured values and the estimated values;
- line R^2: the coefficient of determination, square of the Pearson coefficient R.

Test results

The computations made on the test set are displayed in that sheet. These values are obtained by averaging the results of the twenty-five models that have the smallest VLOO training scores.

For each molecule, the following quantities are displayed:

- the measured value of the surface tension;
- the averaged estimated value;
- the estimation error.

Below the results, a line Selected models recalls the number of selected trainings used for the mean estimation.

The RMSE, Pearson coefficient, and R^2 are defined as explained in the previous subsection.

Training #xxx chart(yy)

Scatter plot (estimated values of the surface tension versus measured values of the surface tension) of the 269 molecules of the training set for the model with ID xxx that has the smallest VLOO training score. The straight line is the diagonal; this is also the case in the VLOO and test result charts described below.

VLOO #xxx chart(yy)

Scatter plot (VLOO-estimated values of the surface tension vs. measured values of the surface tension) of the 269 molecules of the training set for the model with ID xxx that has the smallest VLOO training score.

Test result chart(yy)

Scatter plot (estimated values of the surface tension vs. measured values of the surface tension) of the 23 molecules of the test set. The estimated values are averaged over the 25 models that have the smallest VLOO training score.

Advanced Options

The execution of the graph machine demonstration can be customized from the command line. The proposed default command line is:

```
docker run -it --rm -v ~/docker:/docker -e SEED=4100967572
espcigm/demo:BJMA269EMOL
```

It contains the following terms:

- "`docker`": calls the Docker daemon of the host machine;
- "`run`": launches a Docker container from the Docker image;
- "`-it`": opens and launches the interactive mode;
- "`--rm`": destroys the container at the end of the session;
- "`-v ~/docker:/docker`": creates a volume in the container, and shares it with the ~/docker tree on the host machine;
- "`-e SEED=4100967572`": creates an environment variable in the container with the value `4100967572`, which is used by graph machine as seed of the pseudorandom generator during the model parameter initialization;
- "`espcigm/demo:BJMA269EMOL`": name of the Docker image launched with the run command (with its optional security code).

Parameters can be passed to the graph machine demonstrator through the variables of the container environment. The available commands are detailed below.

Syntax

To input a new value for an environment variable, the following syntax is used:

```
-e <VariableName>=<value>
```

Several variable assignments can be used in the command line. The order of the variable assignment positions in the line is meaningless.

Available variables

Optional variables are listed below (capital letters are important):

- HELP: displays help;
- LICENSE: displays the license agreement;
- SEED: seed provided to the pseudorandom number generator; by default the seed is auto-generated;
- INIT: number of models generated; the default is 250;
- EPOCHS: number of epochs required for each training, the default is 150;
- HIDDEN: number of hidden neurons, the default is 5; this parameter may be a list in one of the formats x,y,z or <min>:<max>[:<step>]; default step is 1; in such a case, the full computation will be performed for each hidden value, and a separate result file will be written;
- MAXREC: maximum number of the selected trainings (that have the smallest VLOO scores) selected for display and computation of the predicted values for the molecules of the test set; the default is 25; this parameter may be a list in one of the formats x,y,z or <min>:<max>:[<step>]; default step is 1; in such a case, the computation will be performed for each selected trainings value, and the results will be written in the same result file;
- YES: value=1 to run the command in demo mode with all the default values; the default is value=0;
- SHAREFOLDER: variable defining a shared working directory; if value=1, the ~/docker/workdir is shared between the host machine and the container; this directory is used by graph machine to write temporary files, in particular the training module file with the so extension (.so file)); this module can thus be reused without recompilation at a new container launch; the default is value=0;
- DECOR: a comma-separated list of decorators appended to the result file name; the resulting decoration strings are separated by underscores;

Decorator code	Suffix appended
✓ t: number of training initializations;	"<value>T"
✓ e: epochs	"<value>E"
✓ n: number of hidden neurons;	"<value>N"
✓ s: number of selected models (smallest VLOOs);	"<value>S"
✓ se: seed used for pseudorandom generator;	"SEED<value>"

Example of result file name: BJMA269EMOL_250T_150E_3N_25S_SEED1513857070
By default, graph machine provides a hidden neuron decorator.

Few examples of command line

Use of an auto-generated seed:

```
docker run -it --rm -v ~/docker:/docker espcigm/demo:BJMA269EMOL
```

Use 1947 as random seed:

```
docker run -it --rm -v ~/docker:/docker -e SEED=1947 espcigm/demo:BJMA269EMOL
```

Run in automated mode with the value 1346167548 as the random seed and a “yes” answer to all questions, the decorator 250T_5N_25S_SEED1346167548 being appended to the result file name:

```
docker run -it --rm -v ~/docker:/docker -e SEED=1346167548 -e YES=1 -e
DECOR=t,n,s,se espcigm/demo:BJMA269EMOL
```

Run with the same data in automated mode as the above command, but with 5, 6, and 7 hidden neurons, 1,000 models (i.e. 1,000 different parameter initializations), each training running 200 epochs, the results using the 10, 20, 30, 40, and 50 trainings that have the smallest VLOO scores, and the training modules created being accessible to the host machine to be reused in future sessions:

```
docker run -it --rm -v ~/docker:/docker -e SEED=1346167548 -e YES=1 -e
HIDDEN=5:7 -e INIT=1000 -e EPOCHS=200 -e MAXREC=10:50:10 -e SHAREFOLDER=1
espcigm/demo:BJMA269EMOL
```

When the computations are done, three files with the names BJMA269EMOL_5N, BJMA269EMOL_6N and BJMA269EMOL_7N are written in the ~/docker/result directory. A supplementary file with the name BJMA269EMOL_RECAP_5-6-7N is also written. It contains a line chart displaying the variation of the average VLOO score with the number of hidden neurons.

GRAPH MACHINE RESULTS WITH DOCKER

All computations were made on a 3 GHz 8-Core Intel Xeon E5 Mac Pro with 32 Go of RAM running macOs Sierra 10.12.2 (Docker configuration: 16 CPUs, 16 Go allocated RAM). They were checked on a 2.8 GHz Intel Core i7 Macbook Pro Retina 13" with 8 Go of RAM running macOs Sierra 10.12.6 (3 CPUs and 4 Go of RAM for Docker), on a 1.3 GHz Intel Core i5 Macbook Air 13" with 4 Go of RAM running OS X El Capitan 10.11.6 (4 CPUs and 4 Go of RAM for Docker), on a 3.4 GHz Intel(R) Core(TM) i7-2600 PC with 4 Go of RAM running Windows 7 Pro SP1 (Docker: 8 CPUs, 3.2 Go allocated RAM), on a 3.3 GHz Intel(R) Core(TM) i5-4590 PC with 16 Go of RAM running Windows 7 Pro (Docker: 8 CPUs, 8 Go allocated RAM), and with an Ubuntu 15.04 virtual partition launched on the Mac Pro (Docker: 12 CPUs, 16 Go allocated RAM). All graph machine tasks are launched with the maximum number of available CPUs minus 2, e.g. 14 CPUs on the Mac Pro and 6 CPUs on the PCs.

Selection of the graph machine complexity with the training/validation set of 244 molecules - assessment of the accuracy of estimation with a test set of 25 molecules

The Docker image called for this task is "demo:BJMA244T25". Its hash code is the following:

sha256:c6e5432ea885a181a0e7985e72f74b4ca91d3953b0cdd5f8e5dbaa5f95c44e3e

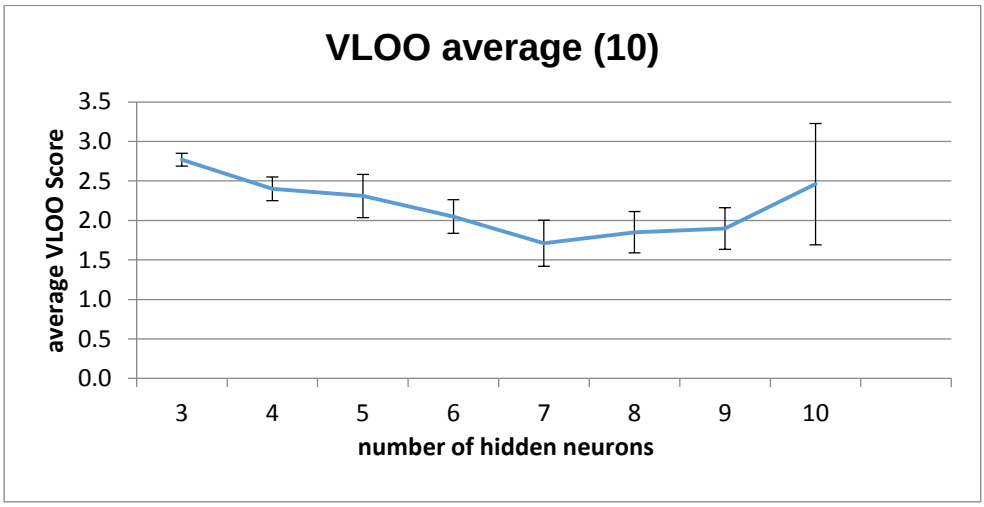
The training data, contained in the DATA cell range, are the names, the SMILES codes, and the surface tension values for the 244 molecules of the training/validation set. The test data contained in the TEST cell range are those of the 25 molecules of the test set. To select the appropriate complexity for the graph machines, the mean of the VLOO scores of the 10 models (out of 100) that have the smallest VLOO scores are compared when the number of hidden neurons increases. The Docker run command is the following, the number of hidden neurons being varied from 3 to 10:

```
docker run -it --rm -v ~/docker:/docker -e YES=1 -e SEED=1346167548 -e
HIDDEN=3:10 -e INIT=100 -e MAXREC=10 espcigm/demo:BJMA244T25
```

The main results are the following (nine files are written in the ~/docker/result directory):

Number of Hidden neurons	Best VLOO scores mean (10 models)	Smallest RMSTE	Smallest VLOO score
3	2.77	1.99	2.55
4	2.40	1.33	2.03
5	2.31	0.94	1.78
6	2.05	0.75	1.70
7	**1.71**	0.49	0,99
8	1.85	0.40	1.42
9	1.90	0.31	1.46
10	2.46	0.18	1.34

The corresponding chart given in the recap file is the following:



The seven hidden neuron model is obviously the best model for generalization.

To speed up the computation and get a coarser idea of the above results, the previous command can be launched with smaller values of the INIT and MAXREC variables:

```
docker run -it --rm -v ~/docker:/docker -e YES=1 -e SEED=1346167548 -e
HIDDEN=3:10 -e INIT=20 -e MAXREC=5 espcigm/demo:BJMA244T25
```

The surface tension estimations for the 25 molecules of the test set are the mean of the estimate values of surface tension computed with the 10 models (out of 100) that have the smallest VLOO scores. They are reported in the following table for the model with 7 hidden neurons:

Molecule	Measured ST	Estimated ST	Estimation error
2,2,3-trimethylpentane	20.2	19.0	-1.2
2-methyloctane	21.4	21.1	-0.3
Propylcyclopentane	24.4	24.2	-0.2
Butylcyclohexane	26.5	26.8	0.3
p-diethylbenzene	28.5	28.4	-0.1
tert-butylbenzene	27.6	27.6	0.0
1-hexanol	25.9	25.6	-0.3
4-Methyl-2-pentanol	22.6	24.1	1.5
Trimethylene glycol	45.2	46.9	1.7
2-Ethoxyethanol	28.6	28.5	-0.1
Diethylene glycol monobutyl ether	29.7	29.9	0.2
Dihexyl ether	25.4	24.3	-1.1
Diethylene glycol diethyl ether	26.8	27.8	1.0
Benzaldehyde	38	36.8	-1.2
2,6-Dimethyl-4-heptanone	24.1	24.8	0.7
isovaleric acid	25	25.0	0.0
isopropyl acetate	21.8	22.1	0.3
Methyl isobutyrate	23.2	23.1	-0.1
3-methylbutanoic acid propyl ester	23.8	23.9	0.1
Methyl decanoate	28	29.0	1.0
Methyl acetoacetate	32.6	32.8	0.2
Dibutylmaleate	30.3	31.8	1.5
Ethyl benzoate	34.6	34.5	-0.1
1,1,1,3,3,3-hexamethyldisiloxane	15.4	15.1	-0.3
Decamethylcyclopentasiloxane	18.2	18.1	-0.1
	Test RMSE	**0.77**	

Training with the complete set of 269 molecules and estimations of the surface tension values for the 23 cosmetic oil set

The Docker image called for this task is "demo:BJMA269EMOL". The training data, contained in the DATA cell range, are the names, the SMILES codes, and the surface tension values for the 269 molecules of the merged training/validation and test sets (previous sets used above). The test data contained in the TEST cell range are those of the 23 molecules of the cosmetic oil set. The command used is the following:

```
docker run -it --rm -v ~/docker:/docker -e YES=1 -e HIDDEN=7 -e MAXREC=25 -e
SEED=799062158 espcigm/demo:BJMA269EMOL
```

It runs a computation in automated mode with the number 799062158 as seed value for the pseudorandom number generator, the "hidden" parameter equal to 7, and uses the 25 training models (out of 250) that have the smallest VLOO scores. As the other variables have their default value, a total number of initializations of 250 and a number of epochs of 150 are used to estimate the surface tension values of the molecules of the cosmetic oil set. The training summary results, identical for the above command issued on the mentioned computers, are reported below for the 25 models that have the smallest VLOO scores:

ID	RMSE	VLOO score
#113	0.64	1.33
#75	0.68	1.49
#161	0.75	1.55
#19	0.67	1.56
#185	0.66	1.60
#122	0.61	1.62
#177	0.66	1.68
#139	0.58	1.68
#54	0.71	1.68
#99	0.66	1.71
#66	0.67	1.74
#22	0.65	1.77
#217	0.80	1.79
#80	0.62	1.81
#164	0.76	1.81
#222	0.66	1.82
#68	0.56	1.83
#170	0.68	1.85
#244	0.67	1.88
#203	0.70	1.90
#208	0.60	1.90
#125	0.82	1.93
#140	0.76	1.97
#211	0.68	1.98
#92	0.76	1.98

RMSE and VLOO score are expressed in the same unit as the surface tension ($mN \cdot m^{-1}$); ID is the index of the model (out of 250). The estimated surface tension values for the molecules of the cosmetic oil set are included in the last column of table S7.

The excel file BJMA269EMOL_SEED799062158_7N.xlsm is also written on the host machine in the ~/docker/result directory, since the sharing option is activated between the container and the host for the /docker tree.

For comparison, training times on the 6 computers are given below:

MacPro: 87 s MacBookPro: 397 s MacBook Air: 504 s PC-i5: 197 s PC-i7: 226 s MacPro under Ubuntu: 86 s

A better estimation of the surface tension values for the test molecules was obtained by increasing the number of generated models. The following command launches a similar computation with the same seed and the same hidden parameter, but with 1,000 initializations. In that case, the 20 models that have the smallest VLOO scores are used to

estimate the surface tension of the cosmetic oils. When this task is run on the Mac Pro, the training time has a duration of 338 s.

```
docker run -it --rm -v ~/docker:/docker -e YES=1 -e HIDDEN=7 –e LEVTHRES=2 -e
INIT=1000 -e MAXREC=20 -e SEED=799062158 -e SHAREFOLDER=1
espcigm/demo:BJMA269EMOL
```

The training summary results are reported below for the 20 models that have the smallest VLOO scores:

ID	RMSE	VLOO score
#291	0.53	1.06
#605	0.49	1.06
#564	0.62	1.20
#776	0.56	1.22
#762	0.55	1.25
#972	0.66	1.26
#480	0.61	1.31
#322	0.66	1.32
#113	0.64	1.33
#579	0.53	1.37
#933	0.58	1.40
#503	0.77	1.41
#886	0.64	1.43
#701	0.70	1.43
#592	0.61	1.43
#885	0.68	1.45
#983	0.70	1.45
#619	0.51	1.46
#816	0.57	1.47
#765	0.52	1.47

The estimated surface tension values for the molecules of the cosmetic oil set are included in the sixth column of table S7. The root mean square test error (0.84) is smaller than the previous one (0.94) since only the first model (one out of 25) used in the previous run to compute the root mean square test error has a smaller VLOO score compared to those of the 20 models selected when the same command is issued with 1,000 initializations.