Supporting Information

Improved model for the refractive index: application to potential components of ambient aerosol

Rémi Bouteloup and Didier Mathieu*
CEA, DAM, Le Ripault, 37260 Monts (France)

E-mail:didier.mathieu@cea.fr
Phone: +33-02-47-34-41-85

Present data were collected from two sources: the popular CRC Handbook of Chemistry and Physics (http://hbcponline.com) and the online TCI catalog of chemical compounds from Tokyo Chemical Industry (http://www.tcichemicals.com/en/us/index.html). We considered only data for molecular liquids, excluding ionic liquids, and made no attempt to account for the role of temperature. Many compounds (2295) are common to both sets. They were used to assess experimental uncertainties from the value of the root mean square difference (RMSD) between CRC and TCI refractive indexes, which is 0.0067. As mentioned in the article, significant discrepancies between the two were observed in a number of cases. They are detailed here in descending order. The structural formulas of the compounds are freely available from the TCI website.

The most severe discrepancy is for 1,2-dibromobutane (533-98-2). The CRC value of 1.4025 is much below the TCI range of values: 1.513-1.516. the latter is closer to the present prediction (1.4894). Various independent measurements carried out in the 1950s (Zhurnal Obshchei Khimii 1950, 20, 677-83; ibid 1952, 22, 852-9; J. Org. Chem. 1955, 20, 1430-42; Optika i Spektros. 1957, 3, 211-20) agree closely with the TCI data, and an earlier measurement reported in 1926 reported an even larger value of 1.558 (Compt. Rend. 182, 788-90).

The second most severe deviation is for 3-methoxy-1-propyne (627-41-8). This time, the CRC value of 1.5035 is very large compared to the TCI range (1.394-1.398) and to the predicted value of 1.3969. Experimental values from literature are 1.3945 (J. Org. Chem. 1957, 22, 780-3), 1.3948 (J. Org. Chem. 1960, 25, 687-90) and 1.3945 (Bull. Soc. Chim. France 1963, 10, 2258-60).

The next most severe error is for α-butyl-benzenemethanol (583-03-9). The CRC value of 1.4086 is significantly lower than TCI data (1.508-1.510) and than the predicted value of 1.5142. The lowest value found in literature is 1.4806 (J. Am. Chem. Soc. 1950, 72, 4368-73), still significantly higher than the CRC value. Although a value as high as 1.546 has been measured (Compt. Rend. 1962, 254, 887-8), other experimental data reported in literature (J. Am. Chem. Soc. 1942, 64, 520-1 and 1955, 77, 1590-4, Arzneimittel Forschung 1962, 12, 347-52) are all between 1.5078 and 1.5112 in very close agreement with TCI data.

For butylchlorodimethylsilane (1000-50-6), the CRC value of 1.5145 is very high compared to TCI data (1.417-1.421) and to the present prediction (1.4159). A value of 1.4208 was reported in 1957 (Doklady Akademii Nauk SSSR, 102, 1131-4). Furthermore, a value of 1.4205 is reported in the catalog of Gelest, Inc. (https://www.gelest.com/product/n-butyldimethylchlorosilane/).

For 1-methoxynaphthalene (2216-69-5), the CRC value of 1.6940 is again much larger than the TCI range (1.621-1.624) and than the predicted value of 1.5860. The TCI data are in excellent

agreement with literature values of 1.6221 (Chemische Berichte 1956, 89, 2105-14) and 1.6214 (J. Am. Chem. Soc. 1958, 409-21).

For cycloheptanol (502-41-0), the CRC value of 1.4071 is lower than TCI data (1.475-1.479) and than the GF prediction of 1.4745. For this compound, there are plenty of experimental data in literature, ranging from 1.4406 (Helv. Chim. Acta 1949, 32, 256-65) to 1.485 (J. Am. Chem. Soc. 1938, 1323-38).

For 3-hydroxyisovaleric acid (625-08-1), the CRC value of 1.5081 is large compared to TCI data (1.439-1.444) and to the predicted value of 1.4548. The CRC value is consistent with measurements reported in 1947 (Zhurnal Obshchei Khimii 1947, 17, 1327-31). Another paper (Soobshcheniya Akademii Nauk Gruzinskoi 1949, 10, 193-6) reports a much lower value of 1.1581.

Extensive data are available as well for benzenemethanethiol (100-53-8). The CRC value of 1.5151 is specially low compared to TCI data (1.575-1.577) and to GF predictions (1.5793). All literature values are very consistent, ranging from 1.5715 (J. Am. Chem. Soc. 1956, 78, 6414) to 1.5757 (Izvest. Vostoch. Filialov. Akad. Nauk SSSR 1957, 60-7).

In contrast, we could not find any original research paper for the RI of 1-fluoro-4-iodobenzene (352-34-1). The CRC value of 1.5270 is low compared to the TCI range (1.582-1.584) and to the GF value of 1.6057. The Yaws Handbook of Physical Properties for Hydrocarbons and Chemicals, (Second Edition) recommends the CRC value, while many online catalogs report the value of 1.583 (for instance: https://www.alfa.com or https://www.sigmaaldrich.com).

Similarly, although the CRC value of 1.5969 for thiazole (288-47-1) is very large compared to the TCI range (1.534-1.543) and to the GF value (1.5369), we could not find alternative data in original research papers. However, values reported in online catalogs clearly support the predicted value and TCI data. For instance, https://www.chemblink.com/products/288-47-1.htm reports a range of 1.5365-1.5395, and https://www.sigmaaldrich.com/ recommends a value of 1.538. The Yaws Handbook reports the same value of 1.5969 as the CRC Handbook. However, this value is assigned a code of 2 in the Yaws Handbook, indicating that this is an estimated value, rather than an experimental one.

Similarly, The CRC value of 1.4932 reported in the CRC Handbook for isopropyl methyl sulfide (1551-21-9) appears especially high compared to TCI (1.438-1.440) and GF (1.4477) data. This time, the value of 1.4363 from the Yaws Handbook (with of code value of 1 indicating that this is a genuine experimental one) supports the TCI data.

For methyl 2-bromobutanoate (3196-15-4), the CRC value of 1.4029 appears too low. The TCI range of 1.451-1.453 is again in fair agreement with the GF prediction (1.4487). Measured values reported in literature range from 1.4497 (J. Am. Chem. Soc. 1952, 74, 3582-6) to 1.4508 (Chemische Berichte 1960, 93, 2222-9).

The same situation is observed for *m*-toluoyl chloride (1711-06-4), with a low CRC value of 1.5050, quite different from TCI data (1.548-1.550) and GF (1.5455) prediction, which are in excellent agreement with the only experimental value of 1.5466 found in literature (Chemicker Zeitung 1979, 103, 9-17).

The CRC value of 1.5161 is also especially low for 2-methoxynitrobenzene (91-23-6). The TCI range (1.559-1.563) is in good agreement with data from literature that range from 1.5590 (Chemistry & Industry 1982, 12, 412) to 1.5597 (J. Chem. Soc. 1933, 1257-9). The GF value (1.5351) lies between the CRC value and all alternative ones.

For pentamethylsilanamine (2083-91-2), TCI data (1.396-1.398) and the GF prediction (1.4074) are consistent, while the CRC value is larger (1.4379). Experimental values in literature range from 1.3950 to 1.4380 (Russ. J. Org. Chem. 2010, 46, 791-793).

For 2-bromohexane, the CRC value of 1.4832 is relatively large, while all other values are very consistent: 1.444-1.446 (TCI), 1.4411 (GF), 1.4421 (J. Chem. Soc. 1953, 1538-41) and 1.4457 (Izvestiya Akademii Nauk SSSR, 1963, 1368-73).

For allyl cinnamate (1866-31-5), the CRC value of 1.530 is the same as the only value we could find in literature (Nauk Beloruss. SSR 1948, 114-24). It is fairly consistent with the predicted GF value of 1.5395, whereas TCI values are higher (1.567-1.569). The latter are consistent with the value of 1.5661 reported in the Fenaroli's Handbook of Flavor Ingredients, Fourth Edition, while both values of 1.530 and 1.5661 are listed in the Encyclopedia of Food and Color Additives, Volume 1.

For 1-pentanamine (110-58-7), the CRC value of 1.448 is larger than all others: 1.409-1.413 for TCI, 1.4196 for the GF prediction, and literature values ranging from 1.41 (Nefte-khimiya 1961, V1, 555-63) to 1.4224 (J. Am. Chem. Soc. 1956, 78, 1705-9).

For 2-thienylmethanol (636-72-6), the CRC value of 1.528 is found only in a patent (US 2577191) from 1951. A wealth of other sources agree on larger values: 1.563-1.566 for TCI, 1.5728 for GF, and at least six alternative sources reporting values ranging from 1.5593 (J. Org. Chem. 1950, 15, 790-4) to 1.5692 (Khimiya Atsetilena, 1968, 256-9).

For triethylsilane (617-86-7), the CRC value of 1.447 is larger than TCI values (1.411-1.413), the GF prediction (1.3959) and most data from literature which range from 1.4088 (Metalloorganicheskaya Khimiya 1989, 2, 1305-7) to 1.4123 (Geslest catalog). It is consistent with only one article (Doklady Akademii Nauk SSSR 1964, 158, 660-3).

Similarly, the CRC value of 1.5125 for phenoxytrimethylsilane (1529-17-5) is larger than all others we could find: 1.476-1.480 for TCI, 1.4805 for GF, and 10 experimental values ranging from 1.475 (Bull. Soc. Chim. France 1967, 4, 1221-4) to 1.4837 (Zhurnal Obshchei Khimii 1960, 30, 3352-8).

The CRC value of 1.4621 for tetramethylputrescine (111-51-3) is also larger than all alternative ones: 1.427-1.430 for TCI, 1.4412 for GF, and others ranging from 1.4261 (J. Org. Chem. 1957, 22, 1225-8) to 1.4316 (J. Chem. Soc. Transactions 1913, 103, 1689-1704).

For clarity, the table below summarizes the above mentioned data regarding the five most severe discrepancies between CRC and TCI refractive indexes.

CASRN	Chemical name	RI values	references
533-98-2	1,2-dibromobutane	1.4025	CRC
		1.513-1.516	TCI
		1.4894	GF
		1.5128-1.517	lit.
627-41-8	3-methoxy-1-propyne	1.5035	CRC
		1.394-1.398	TCI
		1.3969	GF
		1.3945	J. Org. Chem. 1957, 22, 780-3
		1.3948	J. Org. Chem. 1960, 25, 687-90
		1.3945	Bull. Soc. Chim. France 1963, 10, 2258-60
583-03-9	α-butyl-benzenemethanol	1.4086	CRC
		1.508-1.510	TCI
		1.5142	GF
		1.4806	J. Am. Chem. Soc. 1950, 72, 4368-73
		1.546	Compt. Rend. 1962, 254, 887-8
		1.5078-1.5112	lit.
1000-50-6	butylchlorodimethylsilane	1.5145	CRC
		1.417-1.421	TCI
		1.4159	GF
		1.4208	Doklady Akademii Nauk SSSR, 102, 1131-4
		1.4205	https://www.gelest.com
2216-69-5	1-methoxynaphthalene	1.6940	CRC
		1.621-1.624	TCI
		1.5860	GF
		1.6221	Chemische Berichte 1956, 89, 2105-14
		1.6214	J. Am. Chem. Soc. 1958, 409-21

Quite remarkably, whenever data from the CRC Handbook and from the TCI catalog significantly differ from each other, the TCI value appears to be more consistent with literature. The reader should not conclude that refractive indexes compiled in the CRC Handbook are of low quality, as it should be kept in mind that the above mentioned compounds represent a tiny fraction of the whole CRC database. However, present results demonstrate the value of cross-comparisons between various sources of data. They also point to the database of Tokyo Chemical Industry, Inc. as a valuable source of data, in addition to more academic resources.