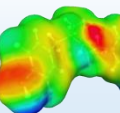


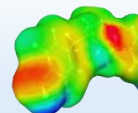
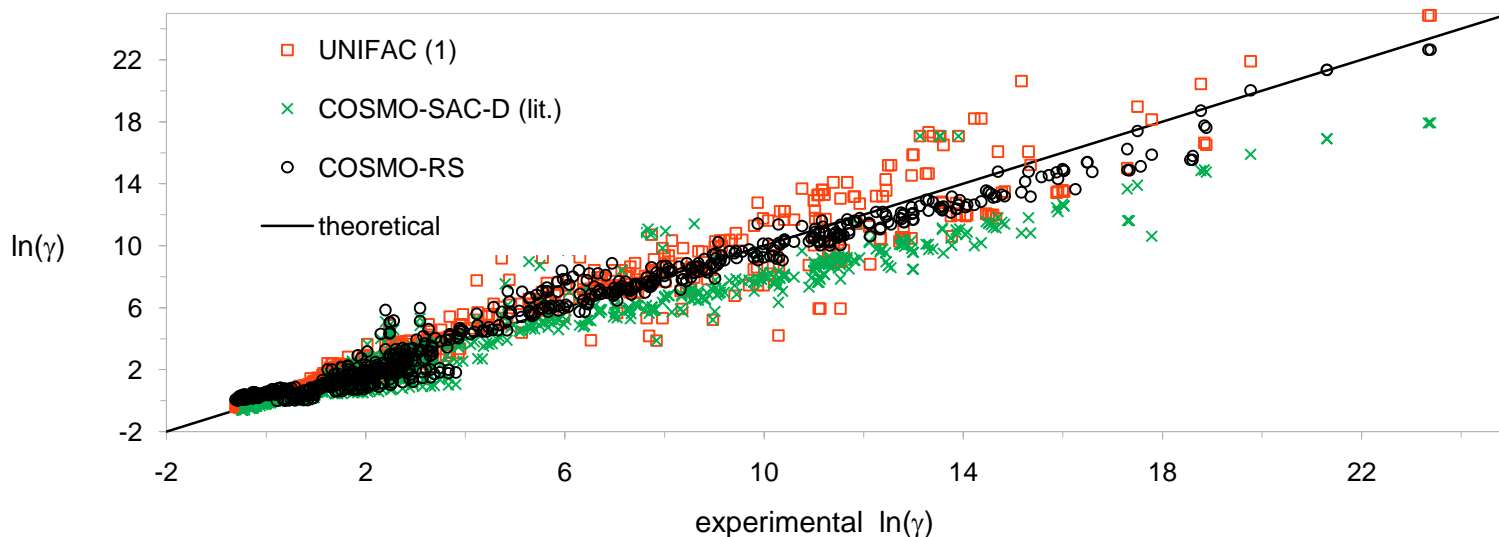
Method benchmark on recently published reference data of inf. dil. activity coefficients

COSMO*therm* versus UNIFAC and COSMO-SAC



Summary

- COSMO-RS outperforms UNIFAC and COSMO-SAC-D with respect to overall root mean square error (RMSE).
- COSMO-SAC-D comes close to COSMO-RS if especially adjusted to the present experimental data. Without special fitting COSMO-SAC-D is 0.7 ln-units worse than COSMO-RS.
- UNIFAC shows the best RMSE among the compared methods when used on organic solvents with limited chemical diversity.



Experimental Databases¹

Overview

Two different experimental databases have been used:

- A non-water DB with mixtures of organic solvents in other organic solvents
- A water DB with mixtures of organic solvents and water

Database overview

	non-water DB		water DB	
Total data points	386		359	
Temperature range / [K]	293 K	393 K	283 K	373 K
Substances	50		257	
IDAC range	0.54	76.7	2.51	9.74E+09
Chemical diversity	very low		low to medium	

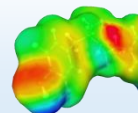
Chemical diversity overview

Groups	Subgroup	Substances in DB		Groups	Subgroup	Substances in DB	
		non-water	water			non-water	water
alkanes	linear	16	8	halogen-alkanes	chloro	2	18
	iso		11		bromo		11
	neo		5		iodo		3
cyclo alkanes	simple	2	4	carbonyls	aldehyde		8
	branched	3	4		ketone	5	18
alkenes	1,2-alkenes	5	7		carbon acids	2	9
	2,3-alkenes	1	3		ester		24
	dienes		2	amide	1		
cyclo alkenes	1,2	4	4	alcohols	primary	6	21
	diene		4		secondary		10
	triene		1		tertiary		11
alkynes			6		phenols	1	
aromatics	simple	2	22	ether			13
	halogen subst.		5	epoxids			1
	polyaromatics		23	water			1

Chemical diversity

- Important groups are missing in both DB (amine, aniline, thiol, imine, cyano, nitro, etc.)
- Bi- or poly functional substances are missing (amino-acids, diols, glycerols, halogenated acids, etc.)
- All substances where group contribution methods have weaknesses are missing (e.g. molecules with intra-molecular hydrogen bonding, halogen- or nitro substituted phenols / acids / amines/ anilines etc.)

1) Renan P. Gerber and Rafael de P. Soares, *Ind. Eng. Chem. Res.*, Article ASAP, DOI: [10.1021/ie901947m](https://doi.org/10.1021/ie901947m)



Methods

Overview of applied methods

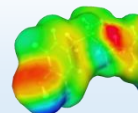
Sources

- COSMO-SAC and UNIFAC results have been taken from a recent paper of R. P. Gerber, R. de P. Soares¹. The analyzed data are to our best knowledge not excluded from the general UNIFAC or COSMO-SAC-D parameterization. COSMO-SAC-D (adj.) is fitted to the present data.
- COSMO-RS results have been calculated with COSMO $therm$ (parameterization C21_0110). The analyzed experimental data are not part of the COSMO $therm$ parameterization to the largest extend.

Methods	# data points		uses molecule conformations	adjusted to these DB	Comment
	non-water DB	water DB			
UNIFAC (1)	374	294	no		Group contribution method
UNIFAC (2)	386	114	no		Group contribution method
COSMO-SAC-D (lit.)	361	311	no		The missing 57 data points in the water-DB show a large AAD in COSMO-RS
COSMO-SAC-D (adj.)	361	311	no	yes	The missing 57 data points in the water-DB show a large AAD in COSMO-RS
COSMO-RS	386	358	yes		
COSMO-RS*	378	358	yes		Without carbon acids in non-water DB

* Acetic acid and propionic acid are known to form dimers in organic solvents, even at low concentration. The experimental extrapolation to infinite dilution may thus be flawed. If acetic acid dimers are used with COSMO-RS, the results are significantly better than for monomers.

- 1) Renan P. Gerber and Rafael de P. Soares, *Ind. Eng. Chem. Res.*, Article ASAP, DOI: [10.1021/ie901947m](https://doi.org/10.1021/ie901947m)
- 2) Gmehling, J.; Li, J.; Schiller, M. A., *Ind. Eng. Chem. Res.* **1993**, *32*, 178–193



Results

Best method

- The best non adjusted method for the organic solvent database (low chemical diversity, small IDAC range) is UNIFAC followed by COSMO-RS. If adjusted to this dataset COSMO-SAC-D is slightly better than COSMO-RS.
- If the water-DB (chemically diverse, IDAC values over 10 decades) is considered, COSMO-RS outperforms all other methods by far. Only the especially adjusted COSMO-SAC-D comes close.

Methods	Combined IDAC results for both DB				
	# points+ non-water	# points+ water	AAD(1)	RMSE+	Max dev+
UNIFAC (1)	374	294	0.80	1.26	6.08
UNIFAC (2)	386	114	0.50	1.11	5.12
COSMO-SAC-D (lit.)	386	311	1.27	1.52	7.18
COSMO-SAC-D (adj.)	361	311	0.63	0.86	3.67
COSMO-RS	361	358	0.61	0.8	3.45

The UNIFAC (2) ADD/RMSE is biased towards organic solvent, because only 114 water-DB points are included.

Methods	IDAC results for the non-water DB					IDAC results for the non-water DB				
	# points+	AAD(1)	RMSE+	Max dev+	mean dev+	# points+	AAD(1)	RMSE+	Max dev+	mean dev+
UNIFAC (1)	374	0.28	0.39	1.74	0.05	294	1.47	1.84	6.08	0.25
UNIFAC (2)	386	0.12	0.17	0.88	-0.05	114	1.8	2.31	5.12	-1.15
COSMO-SAC-D (lit.)	361	0.59	0.81	2.77	-0.36	311	2.05	2.41	7.18	-1.41
COSMO-SAC-D (adj.)	361	0.48	0.57	3.1	-0.18	311	0.81	1.11	3.67	0.39
COSMO-RS	386	0.56	0.71	3.45	-0.015	358	0.66	0.89	3.05	-0.18
COSMO-RS*	378	0.52	0.62	1.99	-0.07					
COSMO-RS**						311	0.56	0.75	2.46	-0.22

* Acetic acid and propionic acid are known to form dimers in organic solvents, even at low concentration. The experimental extrapolation to infinite dilution may thus be flawed. If acetic acid dimers are used with COSMO-RS, the results are significantly better than for monomers.

** The same points as for COSMO-SAC are used for the calculation of the water DB. The missing 57 points in COSMO-SAC have a large influence.

+ The number of points and calculated results from the supplemental material deviate from the published data. The discrepancies are significant, but do not change the picture in a qualitative fashion. The origin of the discrepancies is unknown.

