

# Introducing COSMO-based models, including F-SAC

Prof. Rafael de Pelegrini Soares, D.Sc.

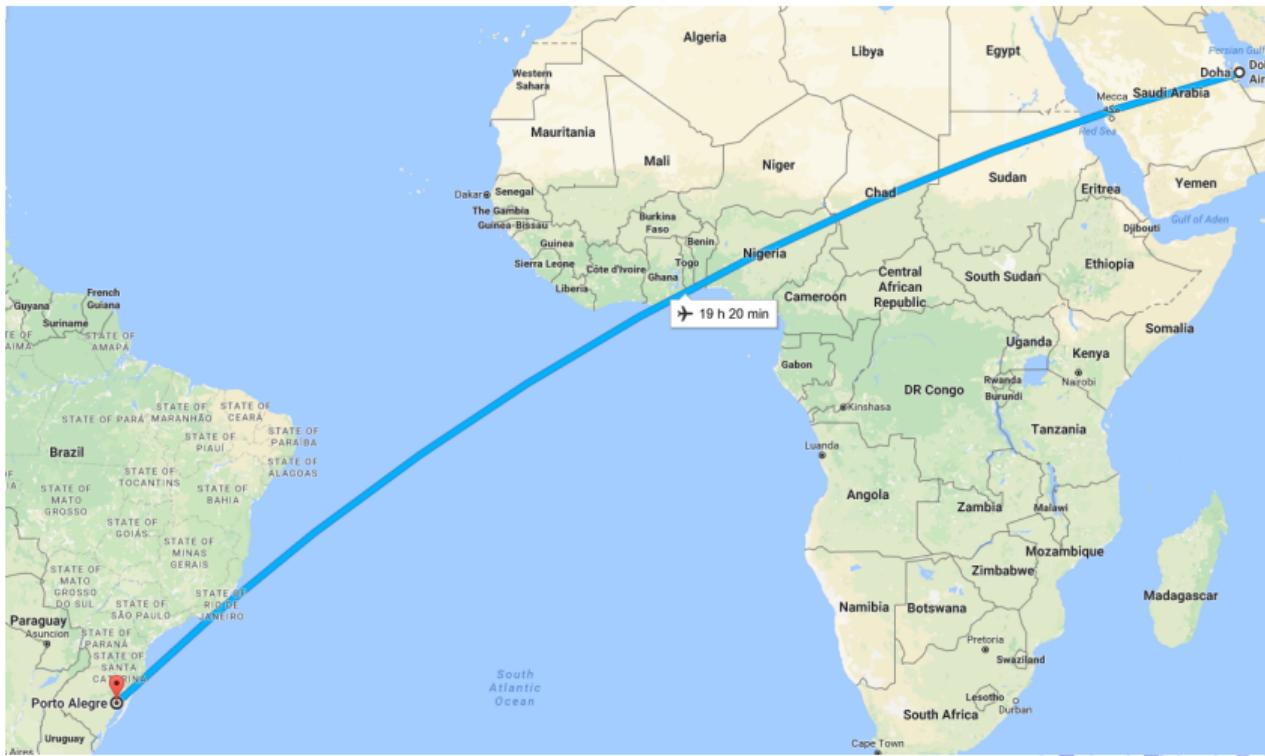


FED. UNIV. OF RIO GRANDE DO SUL - BRAZIL  
CHEMICAL ENGINEERING DEPARTMENT  
LAB. VIRTUAL DE PREDIÇÃO DE PROPRIEDADES  
<http://www.enq.ufrgs.br/labs/lvpp>



April 18, 2018

# Map

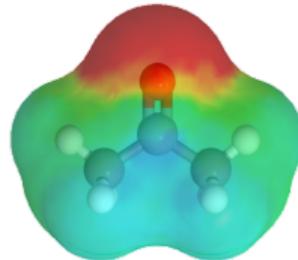
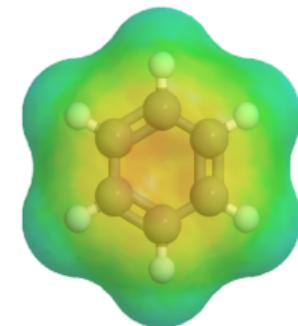


# Our Group: Virtual Laboratory for Property Predictions (LVPP)



# COSMO method (solutes alone)

- The COSMO<sup>a</sup> method was originally developed for the computation of solvation effects
- The method belongs to the class of dielectric continuum models

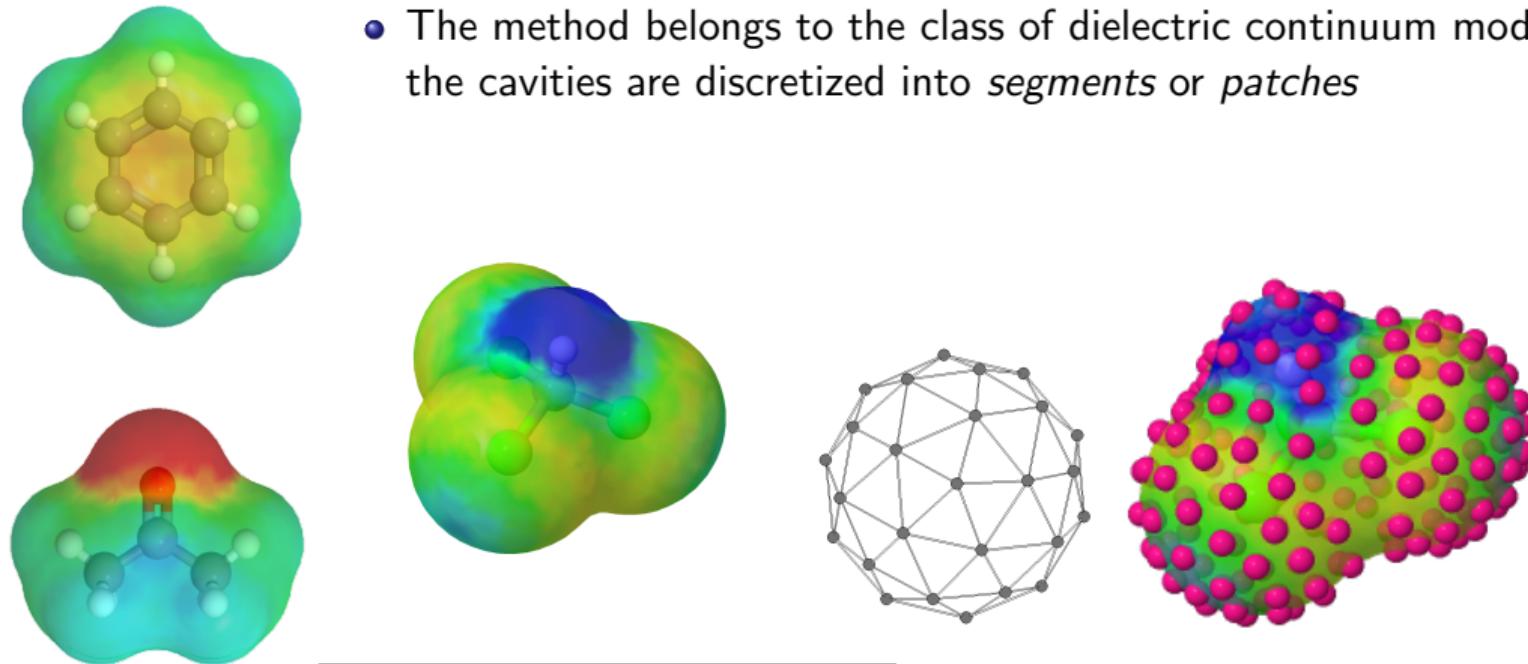


---

<sup>a</sup>A Klamt and G Schüürmann. *J. Chem. Soc., Perkin Trans. 2*, pages 799–805, 1993

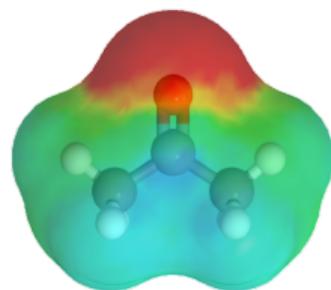
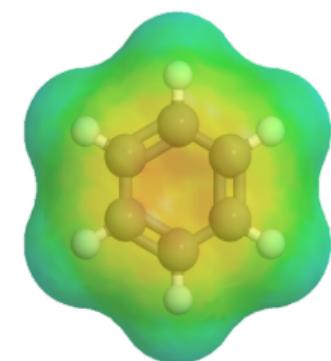
# COSMO method (solutes alone)

- The COSMO<sup>a</sup> method was originally developed for the computation of solvation effects
- The method belongs to the class of dielectric continuum models, the cavities are discretized into *segments* or *patches*



<sup>a</sup>A Klamt and G Schüürmann. *J. Chem. Soc., Perkin Trans. 2*, pages 799–805, 1993

# COSMO-RS – Surface contacting theory (mixtures)



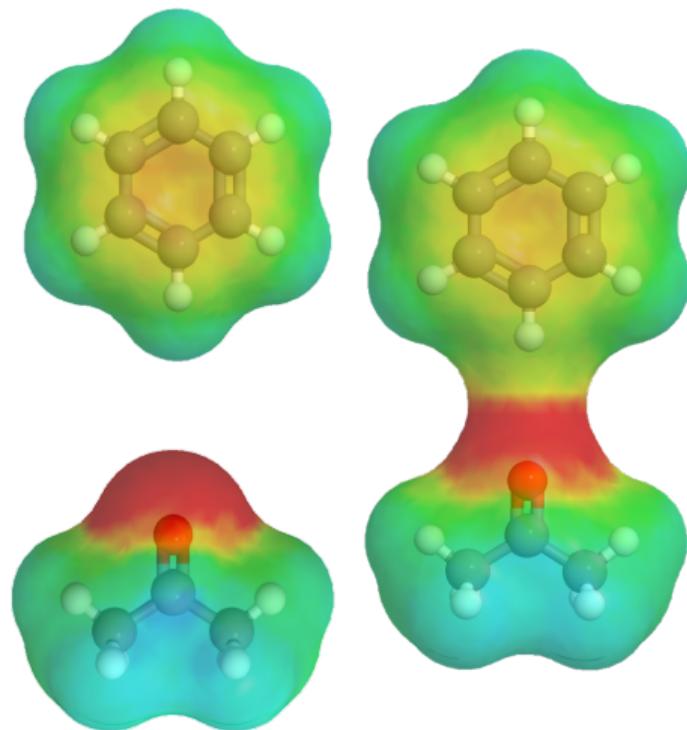
- In the COSMO-RS<sup>a</sup> methods we rely on COSMO computations
- Based on these *pure* substance computations, the mixture behavior is predicted ( $\gamma_i$ )
- The COSMO-SAC<sup>b</sup> formulation follows the same idea

---

<sup>a</sup>Andreas Klamt. *The J. of Phys. Chem.*, 99(7):2224–2235, 1995

<sup>b</sup>Shiang-Tai Lin and Stanley I. Sandler. *Ind. Eng. Chem. Res.*, 41(5):899–913, 2002

# COSMO-RS – Surface contacting theory (mixtures)



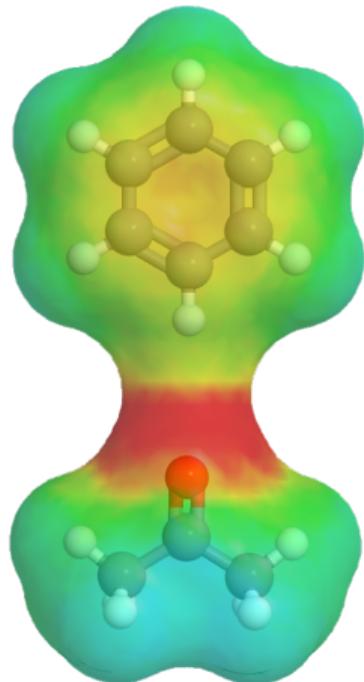
- In the COSMO-RS<sup>a</sup> methods we rely on COSMO computations
- Based on these *pure* substance computations, the mixture behavior is predicted ( $\gamma_i$ )
- The COSMO-SAC<sup>b</sup> formulation follows the same idea

---

<sup>a</sup>Andreas Klamt. *The J. of Phys. Chem.*, 99(7):2224–2235, 1995

<sup>b</sup>Shiang-Tai Lin and Stanley I. Sandler. *Ind. Eng. Chem. Res.*, 41(5):899–913, 2002

# COSMO-RS – Surface contacting theory

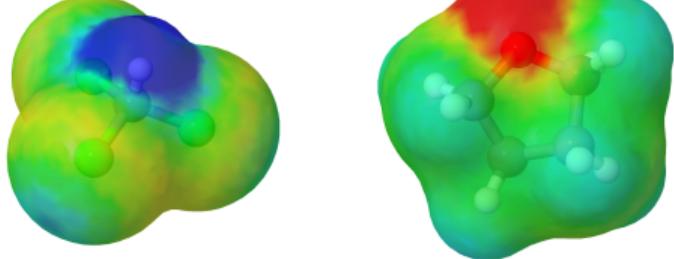


- For each surface pair contact, there is an energy change
- This results in different behavior for different substances in solution
- Clearly, there are many different possible contacting arrangements and a *statistical thermodynamics* treatment is necessary



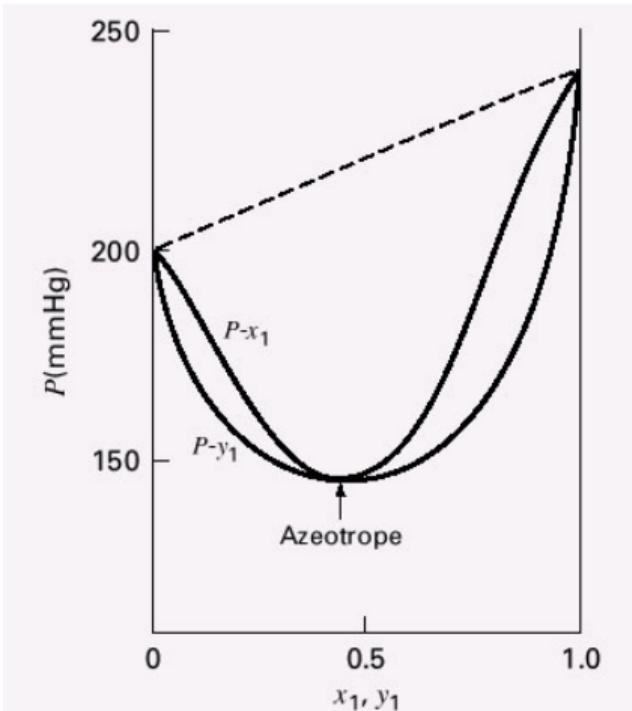
# COSMO-RS – Surface contacting theory (mixtures)

- Based on these *pure* substance computations, the mixture behavior is predicted ( $\gamma_i$ )



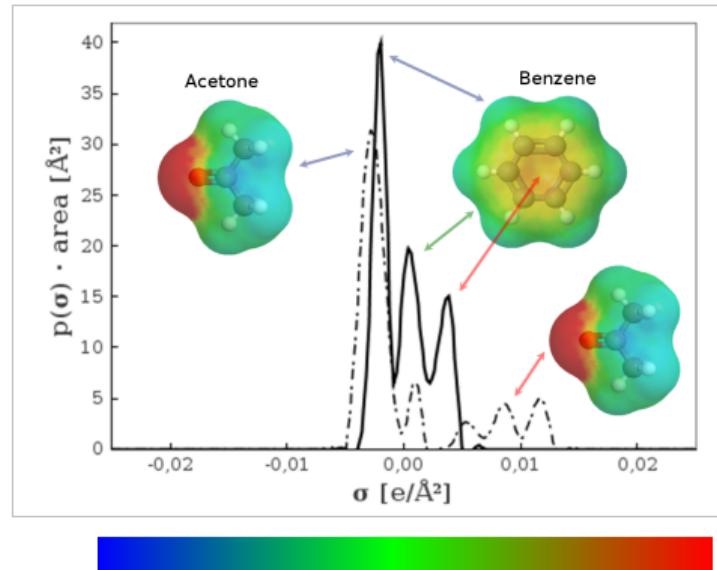
- "It is always desirable to express the properties of a solution in terms that can be calculated completely from the properties of the pure components." – J. M. Prausnitz.

*Molecular thermodynamics of fluid-phase equilibria.*  
Prentice-Hall, third edition, 1999.



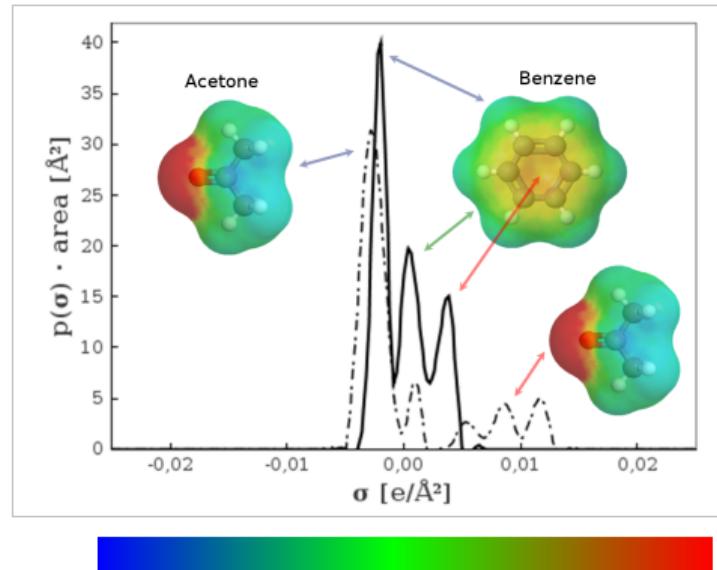
Chloroform/tetrahydrofuran at 30°C

# Sigma profile – $p(\sigma)$



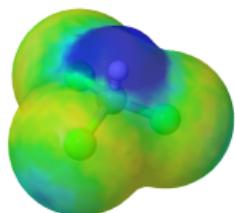
- “It is always desirable to express the properties of a solution in terms that can be calculated completely from the properties of the pure components.” – J. M. Prausnitz. *Molecular thermodynamics of fluid-phase equilibria*. Prentice-Hall, third edition, 1999.

# Sigma profile – $p(\sigma)$



- “It is always desirable to express the properties of a solution in terms that can be calculated completely from the properties of the pure components.” – J. M. Prausnitz. *Molecular thermodynamics of fluid-phase equilibria*. Prentice-Hall, third edition, 1999.
- For a statistical thermodynamics treatment (without using MD), the 3D apparent surface charges are projected into a simple histogram
- These **pure compound** distributions, known as *sigma profiles* –  $p(\sigma)$ , are the basis for computing the activity coefficients in **mixture**

# Sigma profile database



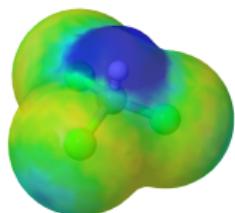
- The quantum chemistry calculations for generating the sigma profiles represent the most time-consuming aspect of COSMO-based methods
- There are several different quantum chemistry packages implementing the COSMO method, e.g.: Gaussian, Turbomole, MOPAC, DMol3 and GAMESS.
- However, different packages lead to different sigma profiles <sup>a</sup>, requiring specific model parametrizations <sup>b</sup>

---

<sup>a</sup>Eric Mullins, Richard Oldland, Y A Liu, Shu Wang, Stanley I Sandler, Chau-chyun Chen, Michael Zwolak, and Kevin C Seavey. *Ind. Eng. Chem. Res.*, 45(12):4389–4415, jun 2006

<sup>b</sup>R.P. Gerber and R. de P. Soares. *Braz. J. of Chem. Eng.*, 30:1–11, 2013

# Sigma profile database



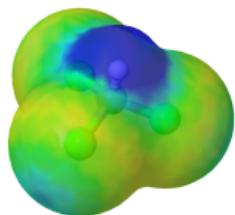
- The quantum chemistry calculations for generating the sigma profiles represent the most time-consuming aspect of COSMO-based methods
- There are several different quantum chemistry packages implementing the COSMO method, e.g.: Gaussian, Turbomole, MOPAC, DMol3 and GAMESS.
- However, different packages lead to different sigma profiles <sup>a</sup>, requiring specific model parametrizations <sup>b</sup>

---

<sup>a</sup>Eric Mullins, Richard Oldland, Y A Liu, Shu Wang, Stanley I Sandler, Chau-chyun Chen, Michael Zwolak, and Kevin C Seavey. *Ind. Eng. Chem. Res.*, 45(12):4389–4415, jun 2006

<sup>b</sup>R.P. Gerber and R. de P. Soares. *Braz. J. of Chem. Eng.*, 30:1–11, 2013

# VT-2005 database



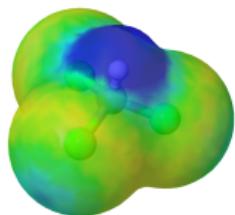
- The freely available database known as VT-2005 contains 1432 compounds<sup>a</sup>, mostly solvents and small molecules
- The VT-2006 contains 206 pharmaceutical-related molecules
- VT-databases were constructed using DMol3 (Materials Studio) which is very expensive...

---

<sup>a</sup>Eric Mullins, Richard Oldland, Y A Liu, Shu Wang, Stanley I Sandler, Chau-chyun Chen, Michael Zwolak, and Kevin C Seavey. *Ind. Eng. Chem. Res.*, 45(12):4389–4415, jun 2006



# VT-2005 database



- The freely available database known as VT-2005 contains 1432 compounds<sup>a</sup>, mostly solvents and small molecules
- The VT-2006 contains 206 pharmaceutical-related molecules
- VT-databases were constructed using DMol3 (Materials Studio) which is very expensive...

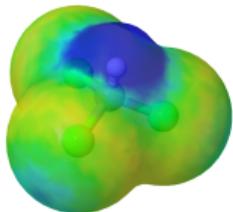
---

<sup>a</sup>Eric Mullins, Richard Oldland, Y A Liu, Shu Wang, Stanley I Sandler, Chau-chyun Chen, Michael Zwolak, and Kevin C Seavey. *Ind. Eng. Chem. Res.*, 45(12):4389–4415, jun 2006



# Alternatives available

- Alternative database with nearly 1000 compounds<sup>a</sup> using MOPAC (semi-empirical – could lead to poor results)
- GAMESS (freely available, including source code) can be used<sup>b</sup>



<sup>a</sup>Renan P. Gerber and Rafael de P. Soares. *Ind. Eng. Chem. Res.*, 49(16):7488–7496, August 2010

<sup>b</sup>Shu Wang, Shiang-Tai Lin, Suphat Watanasiri, and Chau-Chyun Chen. *Fluid Phase Equilib.*, 276(1):37–45, 2009

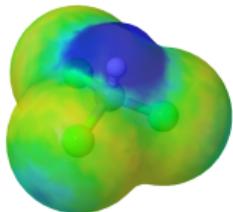
## Suggestion

Assemble an extensive and freely available database using the rigorous methods available in GAMESS



# Alternatives available

- Alternative database with nearly 1000 compounds<sup>a</sup> using MOPAC (semi-empirical – could lead to poor results)
- GAMESS (freely available, including source code) can be used<sup>b</sup>



<sup>a</sup>Renan P. Gerber and Rafael de P. Soares. *Ind. Eng. Chem. Res.*, 49(16):7488–7496, August 2010

<sup>b</sup>Shu Wang, Shiang-Tai Lin, Suphat Watanasiri, and Chau-Chyun Chen. *Fluid Phase Equilib.*, 276(1):37–45, 2009

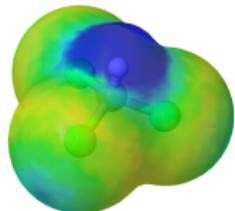
## Suggestion

Assemble an extensive and freely available database using the rigorous methods available in GAMESS



# Alternatives available

- Alternative database with nearly 1000 compounds<sup>a</sup> using MOPAC (semi-empirical – could lead to poor results)
- GAMESS (freely available, including source code) can be used<sup>b</sup>



---

<sup>a</sup>Renan P. Gerber and Rafael de P. Soares. *Ind. Eng. Chem. Res.*, 49(16):7488–7496, August 2010

<sup>b</sup>Shu Wang, Shiang-Tai Lin, Suphat Watanasiri, and Chau-Chyun Chen. *Fluid Phase Equilib.*, 276(1):37–45, 2009

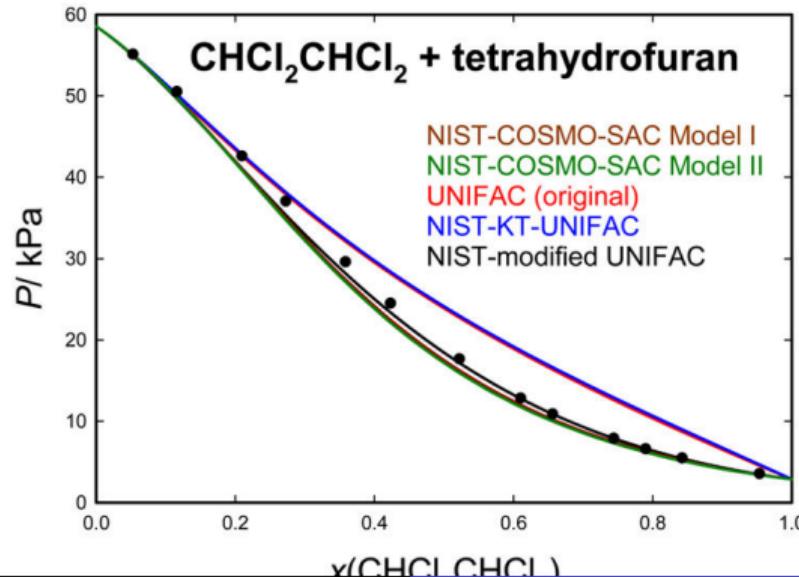
## Suggestion

Assemble an extensive and freely available database using the rigorous methods available in GAMESS



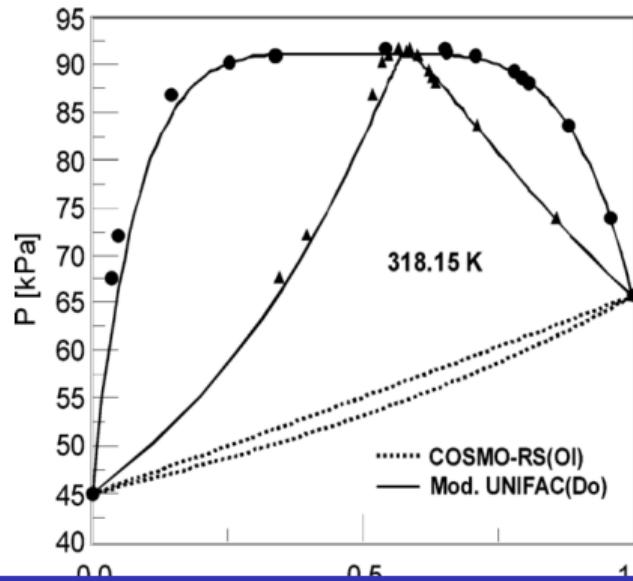
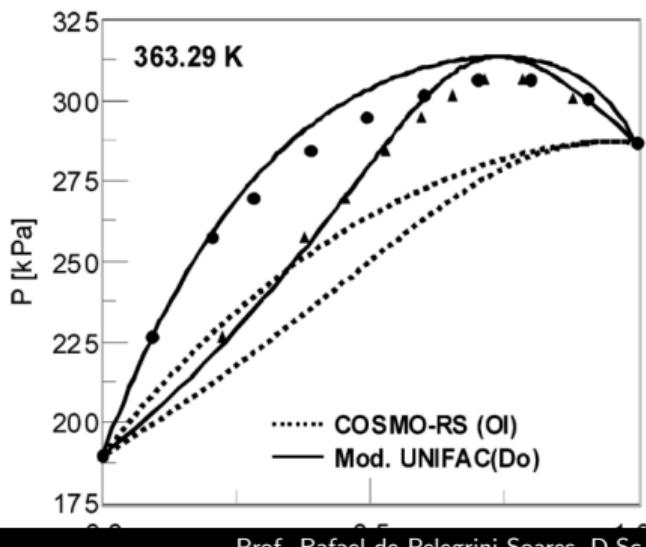
# Reliability: The good...

There are several very good results in the literature, for instance Eugene Paulechka, Vladimir Diky, Andrei Kazakov, Kenneth Kroenlein, and Michael Frenkel. *J. Chem. Eng. Data*, 2015



# The good, the bad and the ugly...

There are also bad and ugly results around, e.g. Hans Grensemann and Jürgen Gmehling. *Ind. Eng. Chem. Res.*, 44(5):1610–1624, 2005, for acetone(1)/diisopro-pylether and perfluorohexane(1)/*n*-hexane, respectively:



# Precision of COSMO-based models

- COSMO-based models have exceptional theoretical features
- However, very often empirical modifications are needed for improving the agreement with experimental data, for instance:
  - Corrections in the water apparent surface charge <sup>1</sup>
  - Empirical scaling factors <sup>2</sup> or empirically scaled surface areas <sup>3</sup>
  - Special parametrizations for different families, e.g. alcohols<sup>4</sup>
  - ...

---

<sup>1</sup>H. Grensemann and J. Gmehling. *Ind. Eng. Chem. Res.*, 44(5):1610–1624, 2005

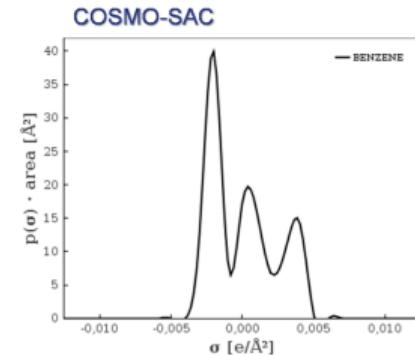
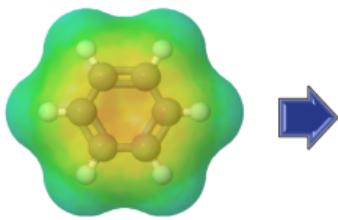
<sup>2</sup>Renan P. Gerber and Rafael de P. Soares. *Ind. Eng. Chem. Res.*, 49(16):7488–7496, August 2010

<sup>3</sup>Jens Reinisch, Andreas Klamt, Frank Eckert, and Michael Diedenhofen. *Fluid Phase Equilib.*, 310(1-2):7 – 10, 2011

<sup>4</sup>Robert Franke, Bernd Hannebauer, and Sebastian Jung. *Fluid Phase Equilib.*, 340:11–14, 2013



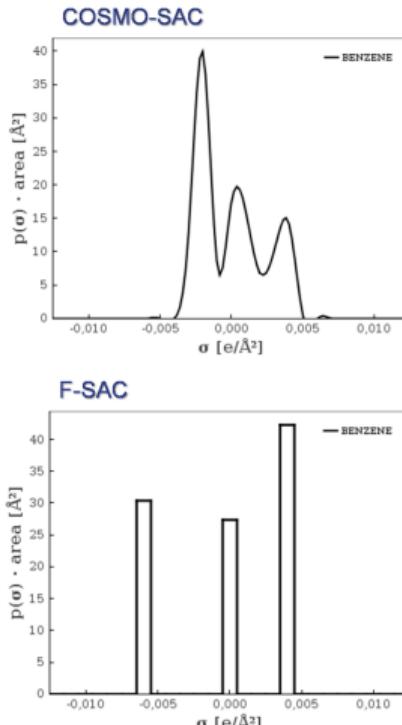
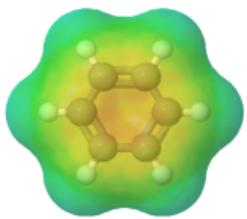
# F-SAC: Functional-Segment Activity Model



- In the F-SAC model, the quantum chemistry computation is replaced by an artificial sigma-profile<sup>a</sup>
- Reduced predictive power
- But **hopefully, with increased resolution** and with less parameters than UNIFAC-type models

<sup>a</sup>R. de P. Soares and R. P. Gerber. *Ind. Eng. Chem. Res.*, 52:11159–11171, 2013

# F-SAC: Functional-Segment Activity Model



Exp.  
Data

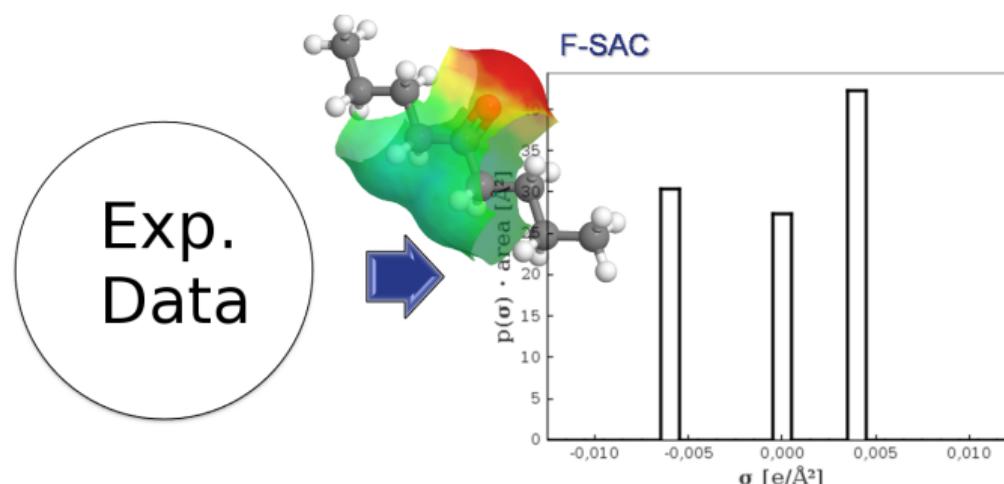


- In the F-SAC model, the quantum chemistry computation is replaced by an artificial sigma-profile<sup>a</sup>
- Reduced predictive power
- But **hopefully, with increased resolution** and with less parameters than UNIFAC-type models

<sup>a</sup>R. de P. Soares and R. P. Gerber. *Ind. Eng. Chem. Res.*, 52:11159–11171, 2013

# F-SAC: Functional-Segment Activity Model

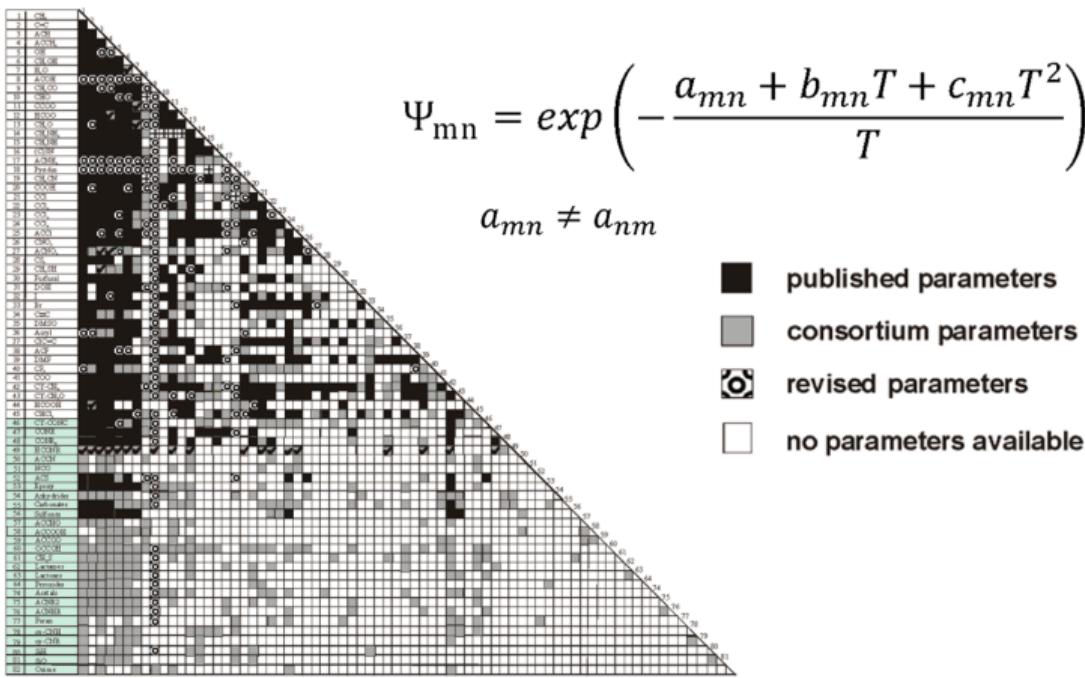
- F-SAC and COSMO-SAC model equations are identical, the difference is in the sigma profile
- In the F-SAC model there is a neutral and two charged peaks per functional group
- Simply add the sigma-profiles of functional groups to get the sigma-profile of molecules



LVPP

LABORATÓRIO VIRTUAL DE  
PREDIÇÃO DE PROPRIEDADES

# Why don't just use UNIFAC instead?



According to UNIFAC (Do) revision 5 – Antje Jakob, Hans Grensemann, Jürgen Lohmann, and Jürgen Gmehling. *Ind. Eng. Chem. Res.*, 45(23):7924–7933, 2006.

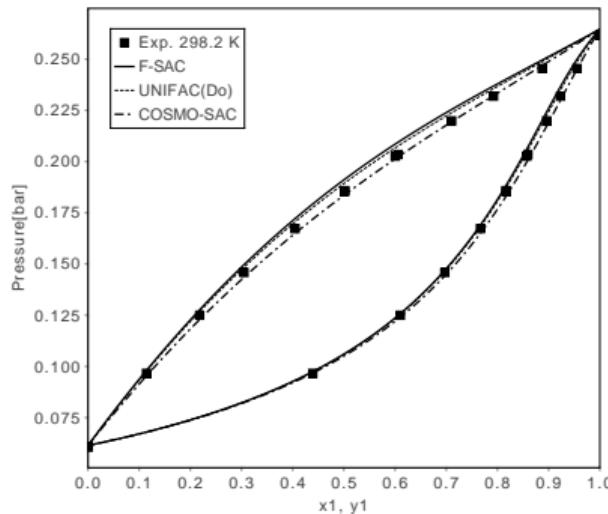
# F-SAC parameters

- For the **24 grupos** considered, only **152** parameters were calibrated.
- Only the hydrogen-bonding\* energies are **pairwise**
- All other parameters are for **pure** groups alone, reducing the total number of parameters
- In order to represent the same molecules in mixtures, **778** parameters are used in UNIFAC (Do).

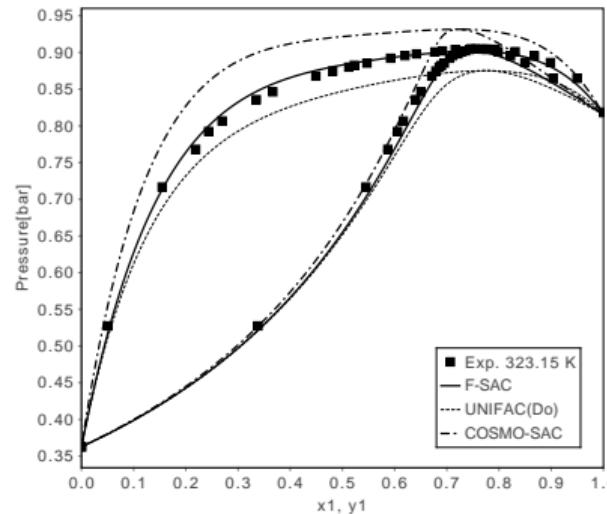
| Group                         | $Q_k^+/\text{\AA}^2$ | $Q_k^-/\text{\AA}^2$ | $\sigma_k^+/e \cdot \text{\AA}^{-2}$ |
|-------------------------------|----------------------|----------------------|--------------------------------------|
| $\text{CH}_2$                 | 0.00                 | 0.00                 | 0.0000                               |
| $\text{C}=\text{C}$           | 6.16                 | 3.70                 | 0.0050                               |
| ACH                           | 5.38                 | 6.71                 | 0.0056                               |
| $\text{CH}_3\text{COCH}_3$    | 21.97                | 80.23                | 0.0133                               |
| $\text{CH}_2\text{CHO}$       | 14.85                | 81.88                | 0.0172                               |
| $\text{CH}_3\text{COAC}_3$    | 9.56                 | 16.68                | 0.0162                               |
| c- $\text{CH}_2\text{COCH}_2$ | 13.93                | 42.78                | 0.0166                               |
| $\text{CH}_3\text{COOCH}_3$   | 28.94                | 107.69               | 0.0109                               |
| Triacetin                     | 34.68                | 62.83                | 0.0146                               |
| GBLactone                     | 18.05                | 41.30                | 0.0161                               |
| Phthalate                     | 33.60                | 69.26                | 0.0132                               |
| $\text{CH}_2\text{COOAC}_3$   | 12.59                | 64.23                | 0.0158                               |
| $\text{CH}_3\text{OCH}_2$     | 13.01                | 22.71                | 0.0091                               |
| c- $\text{CH}_2\text{OCH}_2$  | 8.12                 | 27.20                | 0.0143                               |
| $\text{N}(\text{CH}_2)_3$     | 3.60                 | 28.58                | 0.0175                               |
| DMSO                          | 11.36                | 22.84                | 0.0188                               |
| $\text{CH}_2\text{Cl}$        | 9.30                 | 9.07                 | 0.0089                               |
| $\text{CCl}_4$                | 18.40                | 26.23                | 0.0062                               |
| $\text{CH}_3\text{CCl}_3$     | 40.07                | 20.28                | 0.0039                               |
| $\text{CHCl}_3$               | 43.32                | 14.76                | 0.0035                               |
| ClAC <sub>3</sub>             | 37.41                | 12.13                | 0.0026                               |
| $\text{CH}_3\text{OH}^*$      | 8.31                 | 4.72                 | 0.0132                               |
| $\text{CH}_2\text{OH}^*$      | 7.34                 | 4.69                 | 0.0145                               |
| $\text{H}_2\text{O}^*$        | 8.84                 | 12.16                | 0.0123                               |



# VLE predictions for non-associating mixtures

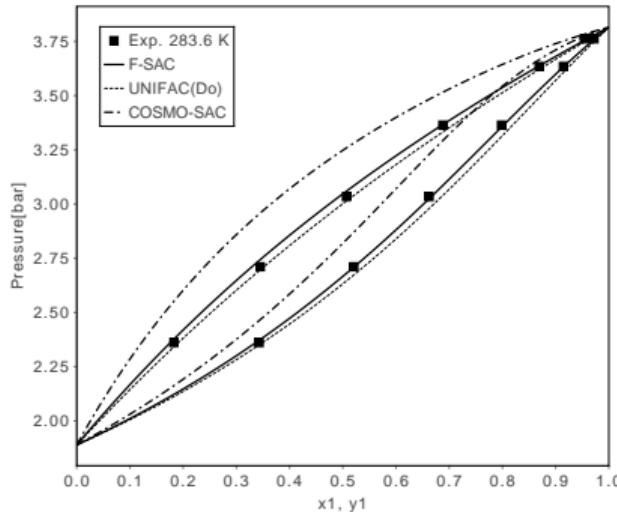


(a) Chloroform/n-heptane

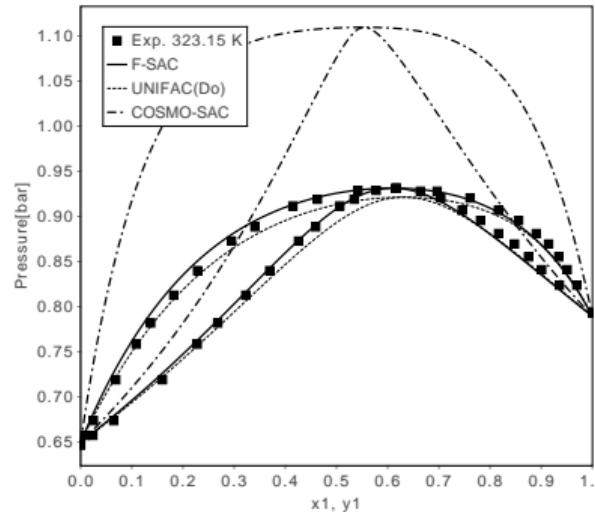


(b) Acetone/cyclohexane

# VLE predictions for non-associating mixtures



(c) Dimethyl ether/1-butene

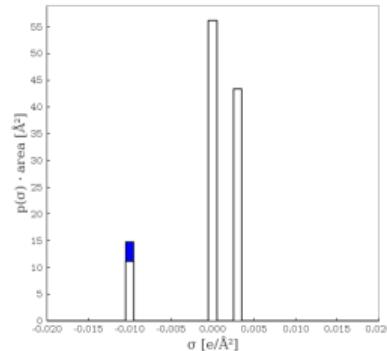
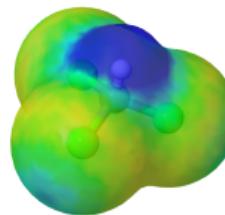
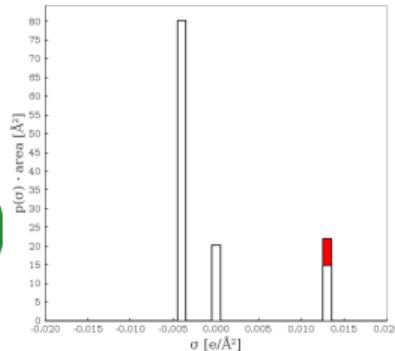
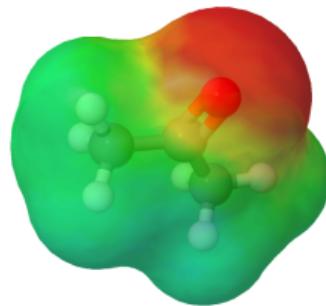


(d) Methyl acetate/1-hexene



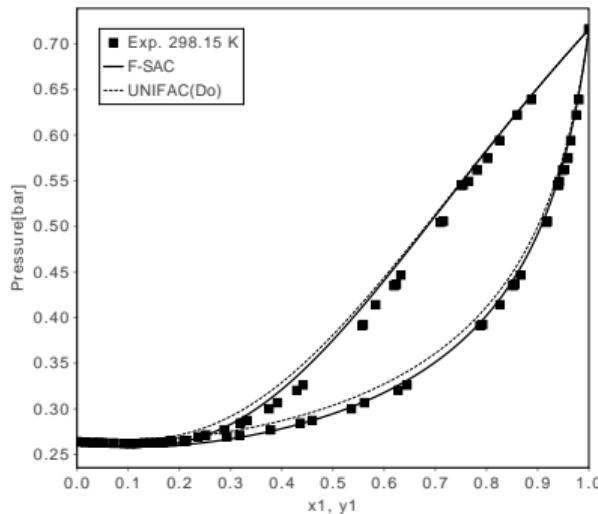
# F-SAC parameters for Hydrogen-Bonding mixtures (association)

- HB-acceptor is red (positive), HB-donor is blue (negative)<sup>5</sup>

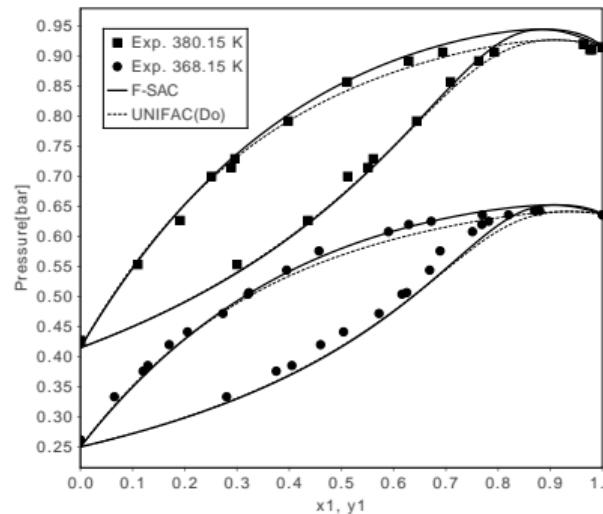


<sup>5</sup>R. de P. Soares, R.P. Gerber, L.F.K. Possani, and P.B. Staudt. *Ind. Eng. Chem. Res.*, 52:11172–11181, 2013.

# VLE predictions for associating mixtures

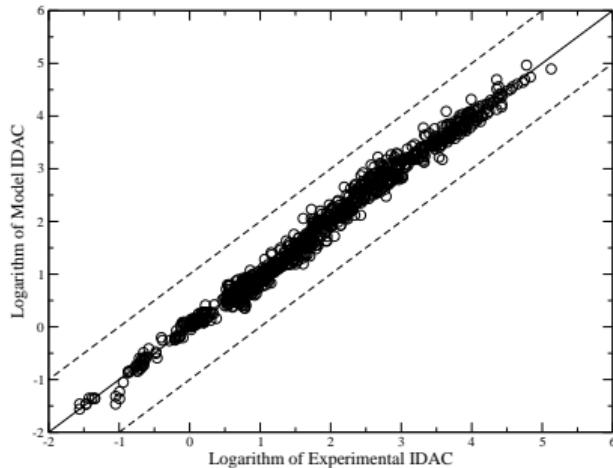


(e) diethyl-ether/chloroform

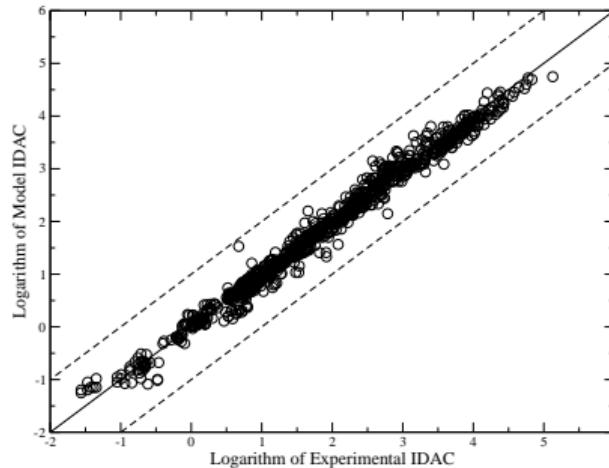


(f) toluene/3-methyl,1-butanol

# IDAC comparison for associating mixtures (water excluded)



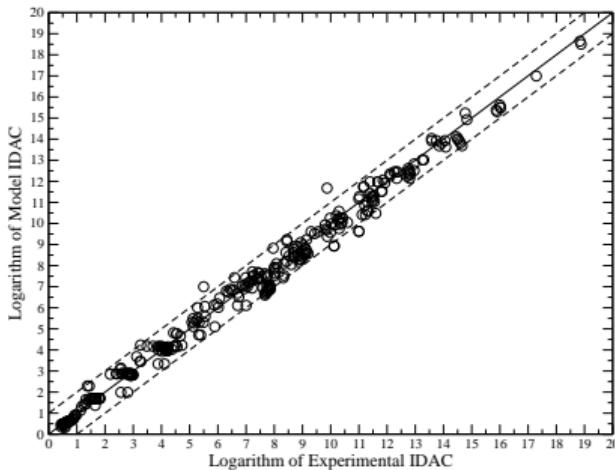
(g) F-SAC



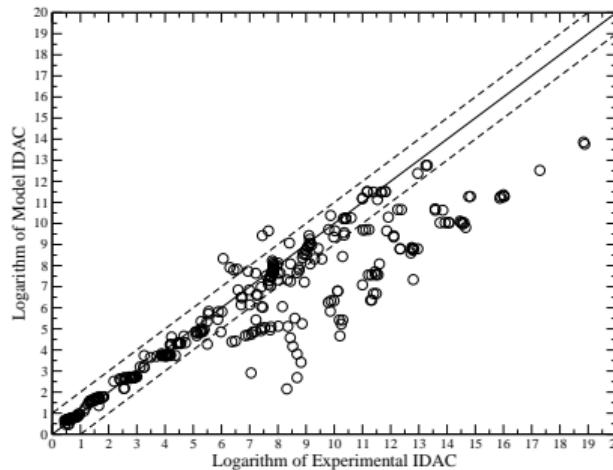
(h) UNIFAC (Do)

R. de P. Soares, R.P. Gerber, L.F.K. Possani, and P.B. Staudt. *Ind. Eng. Chem. Res.*, 52:11172–11181, 2013

# IDAC comparison for mixtures with water



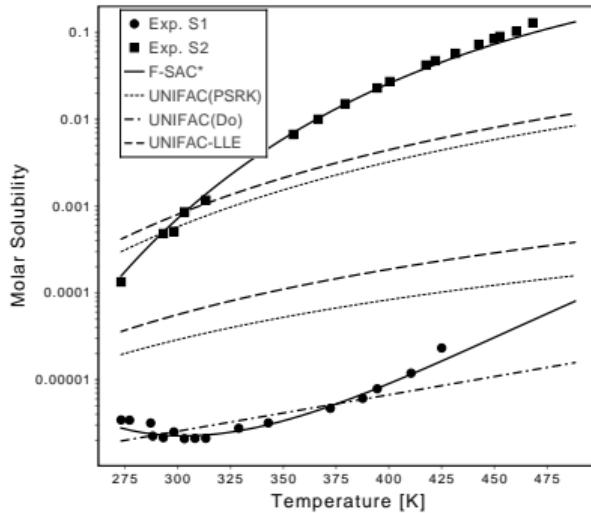
(i) F-SAC



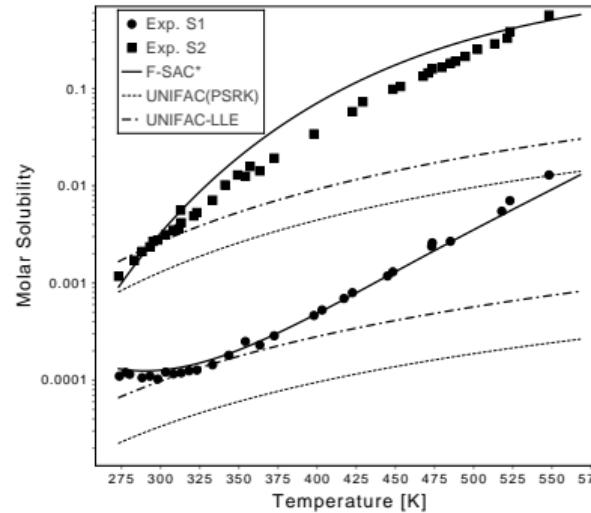
(j) UNIFAC (Do)

R. de P. Soares, R.P. Gerber, L.F.K. Possani, and P.B. Staudt. *Ind. Eng. Chem. Res.*, 52:11172–11181, 2013

# FSAC: Hydrocarbon-water mutual solubilities



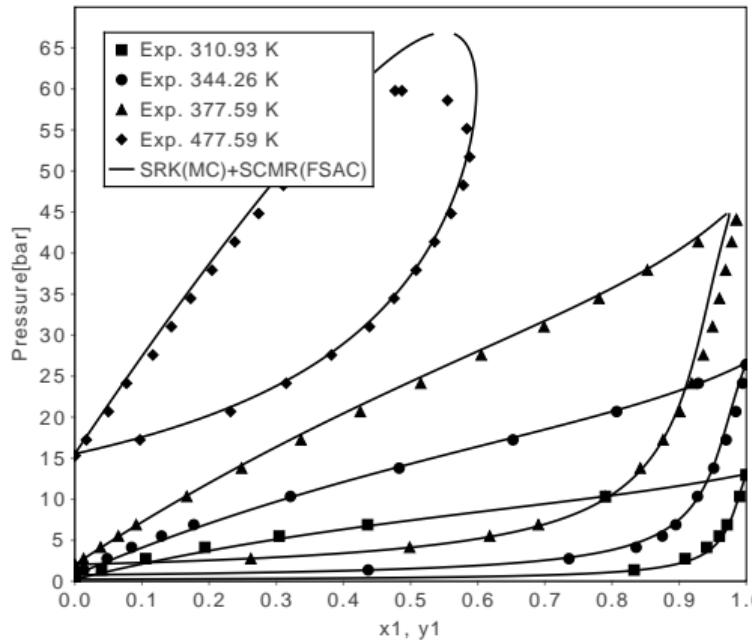
(k) n-Hexane-Water



(l) Toluene-Water

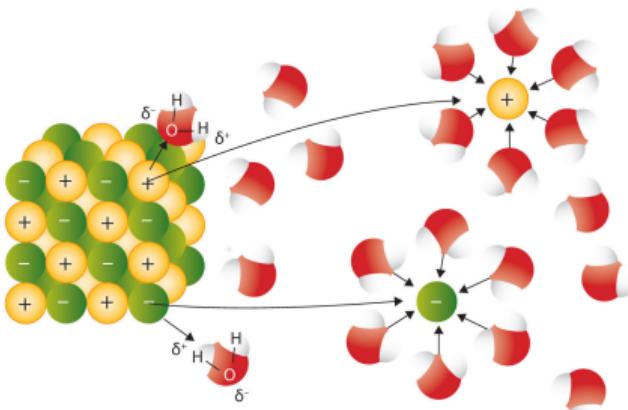
# Predicting high-pressure VLE with SCMR

- The F-SAC can be combined with SRK with Mathias-Copeman  $\alpha$  function and the Self-Consistent Mixing Rule<sup>6</sup> (SCMR) for propane-benzene, no parameter is adjusted for the mixture effects:



<sup>6</sup>Paula B. Staudt and Rafael de P. Soares. *Fluid Phase Equilibria*, 334:76–88, 2012

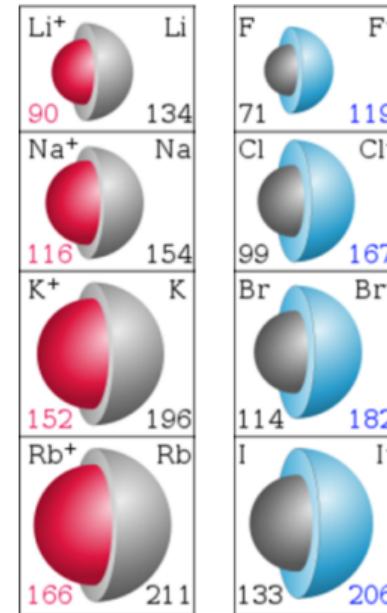
# Extension for electrolytes: eF-SAC



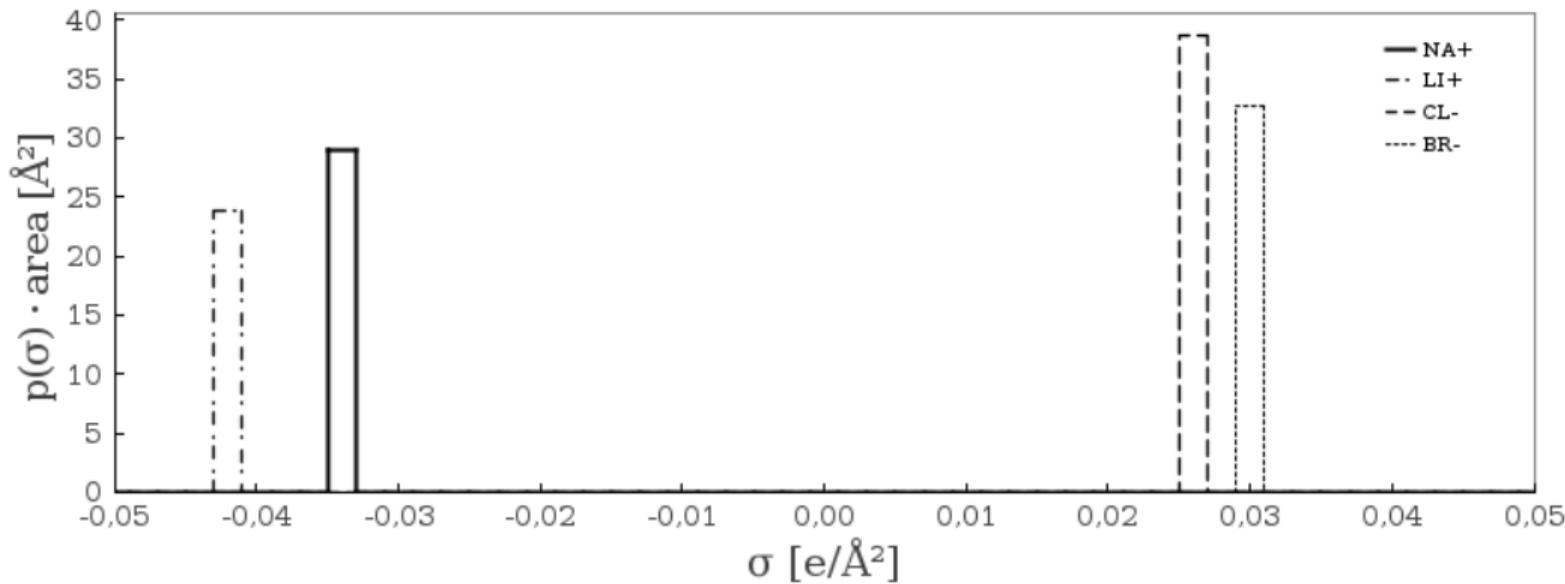
- Ions considered spherical with a given charge (+1, -1, +2, -2, etc.)
- Then the only parameter is the surface area (water parameters from previous works)
- To avoid direct contact between ions, their interaction energy is assumed high

eF-SAC  $\sigma$ -profiles

| Group         | $Q_k^+(\text{\AA}^2)$ | $Q_k^-(\text{\AA}^2)$ | $\sigma_k^+(e\text{\AA}^{-2})$ |
|---------------|-----------------------|-----------------------|--------------------------------|
| $\text{Li}^+$ | 0                     | 36,96                 | -0,027                         |
| $\text{Na}^+$ | 0                     | 29,05                 | -0,034                         |
| $\text{K}^+$  | 0                     | 47,13                 | -0,021                         |
| $\text{Rb}^+$ | 0                     | 40,17                 | -0,025                         |
| $\text{Cs}^+$ | 0                     | 41,64                 | -0,024                         |
| $\text{F}^-$  | 27,12                 | 0                     | 0,037                          |
| $\text{Cl}^-$ | 39,19                 | 0                     | 0,026                          |
| $\text{Br}^-$ | 30,11                 | 0                     | 0,033                          |
| $\text{I}^-$  | 38,69                 | 0                     | 0,026                          |

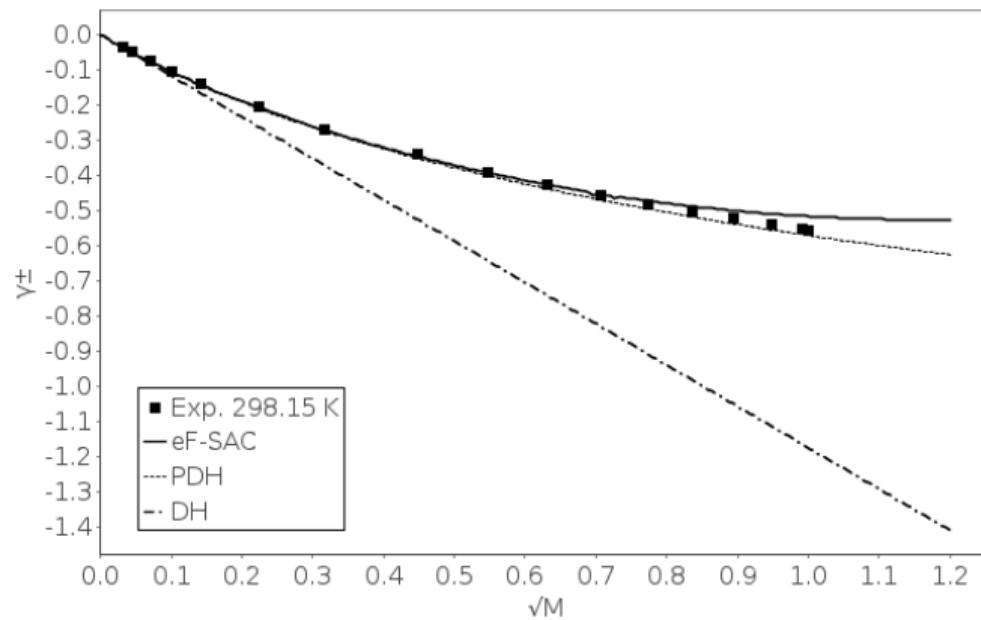


# Results: $\sigma$ -Profile for eF-SAC



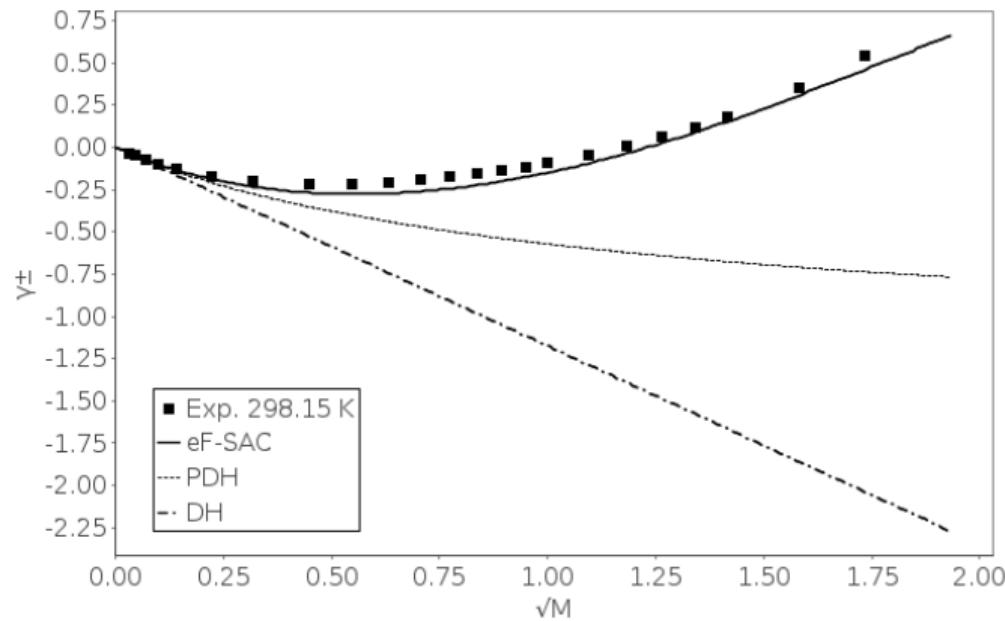
# Results: $\gamma_{\pm}$

- $\text{NaF}_{aq}$



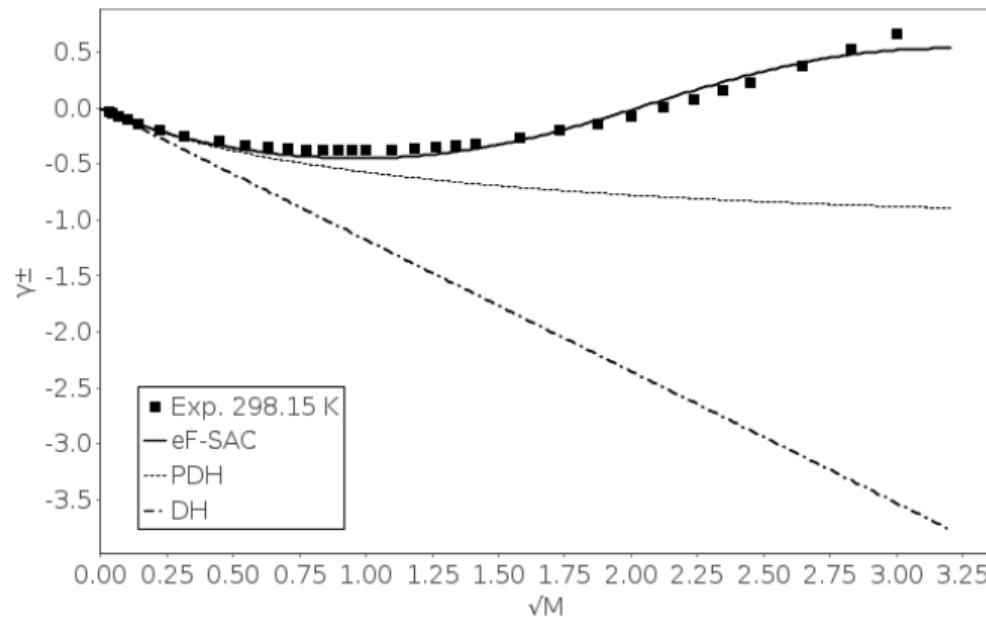
# Results: $\gamma_{\pm}$

- $Lil_{aq}$



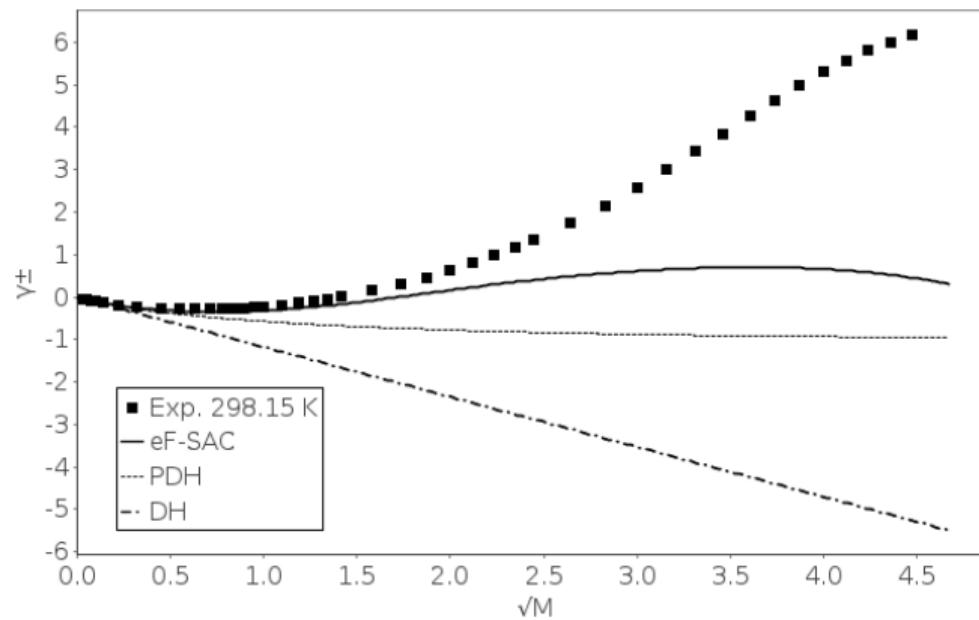
# Results: $\gamma_{\pm}$

- $\text{NaBr}_{aq}$



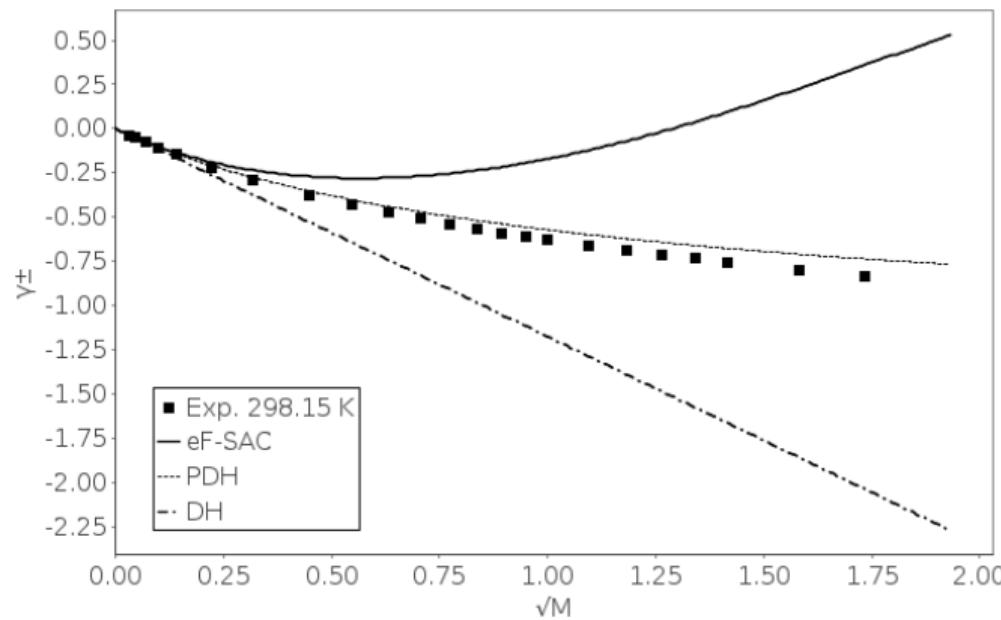
# Results: $\gamma_{\pm}$

- LiBr<sub>aq</sub>



# Results: $\gamma_{\pm}$

- $Csl_{aq}$



LABORATÓRIO VIRTUAL DE  
PREDIÇÃO DE PROPRIEDADES

# Conclusions

- COSMO-based models have exceptional theoretical features
- An extensible, open, and rigorous sigma-profile database needs to be developed
- Currently, results can be considered semi-quantitative
- Agreement with experimental data is usually obtained by means of empirical corrections
- The variation known as **F-SAC** usually produces better results but depends on fitted parameters
- There are certainly many other opportunities for improvement and new fields for application



# Conclusions

- COSMO-based models have exceptional theoretical features
- An extensible, open, and rigorous sigma-profile database needs to be developed
- Currently, results can be considered semi-quantitative
- Agreement with experimental data is usually obtained by means of empirical corrections
- The variation known as **F-SAC** usually produces better results but depends on fitted parameters
- There are certainly many other opportunities for improvement and new fields for application



# Conclusions

- COSMO-based models have exceptional theoretical features
- An extensible, open, and rigorous sigma-profile database needs to be developed
- Currently, results can be considered semi-quantitative
- Agreement with experimental data is usually obtained by means of empirical corrections
- The variation known as **F-SAC** usually produces better results but depends on fitted parameters
- There are certainly many other opportunities for improvement and new fields for application



# Thank you!



- Molecules database at <http://code.google.com/p/jcosmo/>
- Download the F-SAC demonstration code at <http://www.enq.ufrgs.br/labs/lvpp>
- F-SAC is already available in the iiSE process simulator and being integrated into DWSim (open source) process simulator
- Contact: [rafael.pelegrini@ufrgs.br](mailto:rafael.pelegrini@ufrgs.br)



-  Robert Franke, Bernd Hannebauer, and Sebastian Jung.  
*Fluid Phase Equilib.*, 340:11–14, 2013.
-  H. Gremseemann and J. Gmehling.  
*Ind. Eng. Chem. Res.*, 44(5):1610–1624, 2005.
-  Hans Gremseemann and Jürgen Gmehling.  
*Ind. Eng. Chem. Res.*, 44(5):1610–1624, 2005.
-  Renan P. Gerber and Rafael de P. Soares.  
*Ind. Eng. Chem. Res.*, 49(16):7488–7496, August 2010.
-  R.P. Gerber and R. de P. Soares.  
*Braz. J. of Chem. Eng.*, 30:1–11, 2013.
-  Antje Jakob, Hans Gremseemann, Jürgen Lohmann, and Jürgen Gmehling.  
*Ind. Eng. Chem. Res.*, 45(23):7924–7933, 2006.

-  **Andreas Klamt.**  
*The J. of Phys. Chem.*, 99(7):2224–2235, 1995.
-  **A Klamt and G Schüürmann.**  
*J. Chem. Soc., Perkin Trans. 2*, pages 799–805, 1993.
-  **Shiang-Tai Lin and Stanley I. Sandler.**  
*Ind. Eng. Chem. Res.*, 41(5):899–913, 2002.
-  **Eric Mullins, Richard Oldland, Y A Liu, Shu Wang, Stanley I Sandler, Chau-chyun Chen, Michael Zwolak, and Kevin C Seavey.**  
*Ind. Eng. Chem. Res.*, 45(12):4389–4415, jun 2006.
-  **Eugene Paulechka, Vladimir Diky, Andrei Kazakov, Kenneth Kroenlein, and Michael Frenkel.**  
*J. Chem. Eng. Data*, 2015.
-  **J. M. Prausnitz.**



LABORATÓRIO VIRTUAL DE  
PROPRIEDADES

## *Molecular thermodynamics of fluid-phase equilibria.*

Prentice-Hall, third edition, 1999.

-  L.F.K. Possani, R.L. Simoes, P.B. Staudt, and R. de P. Soares.  
*Fluid Phase Equilib.*, 384:122–133, 2014.
-  Jens Reinisch, Andreas Klamt, Frank Eckert, and Michael Diedenhofen.  
*Fluid Phase Equilib.*, 310(1-2):7 – 10, 2011.
-  R. de P. Soares and R. P. Gerber.  
*Ind. Eng. Chem. Res.*, 52:11159–11171, 2013.
-  R. de P. Soares, R.P. Gerber, L.F.K. Possani, and P.B. Staudt.  
*Ind. Eng. Chem. Res.*, 52:11172–11181, 2013.
-  Paula B. Staudt and Rafael de P. Soares.  
*Fluid Phase Equilibria*, 334:76–88, 2012.
-  Shu Wang, Shiang-Tai Lin, Suphat Watanasiri, and Chau-Chyun Chen.



LVPP

*Fluid Phase Equilib.*, 276(1):37–45, 2009.

