Surface tension models for binary aqueous solutions:
A review and intercomparison
Supplemental Material

Judith Kleinheins,*[a] Nadia Shardt,[a] Manuella el Haber,[b] Corinne Ferronato,[b] Barbara
Nozière,[c] Thomas Peter,[a] and Claudia Marcolli[a]

[a] Institute for Atmospheric and Climate Science, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland.
E-mail: judith.kleinheins@env.ethz.ch
[b] Ircelyon, CNRS and Universite Lyon 1, Villeurbanne, France
[c] Royal Institute of Technology (KTH), Department of Chemistry, Stockholm, Sweden

Contents

1 Experimental surface tension data

Figure S1: Overview of the experimental surface tension data of the test substances.The data for each substance are one data set or multiple data sets compiled from the literature. For the individual data sets of each substance, see the next Figure.
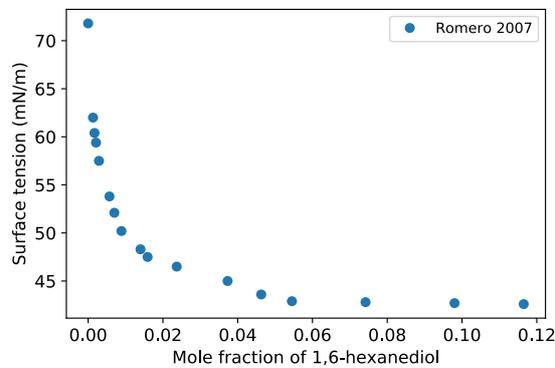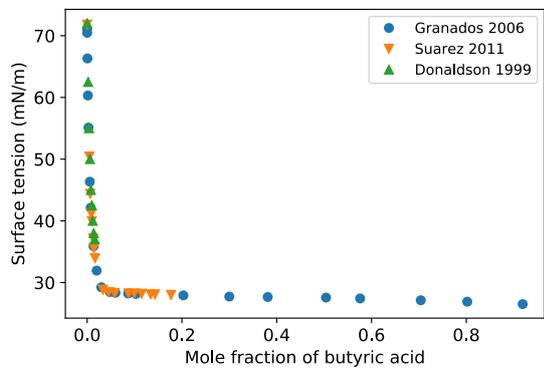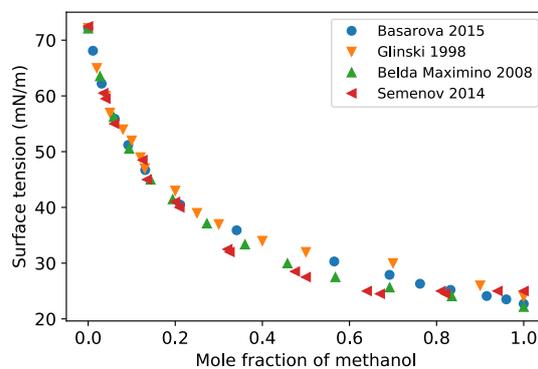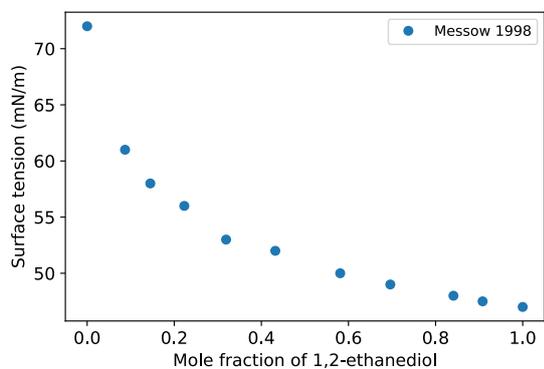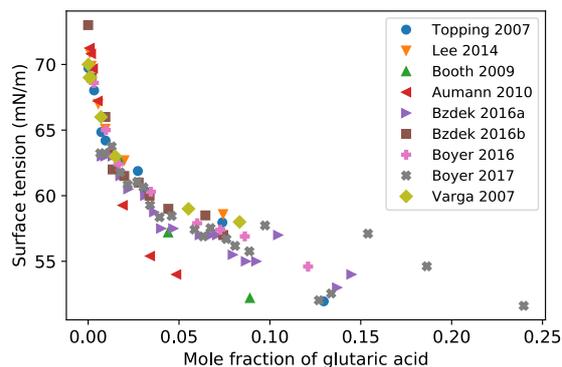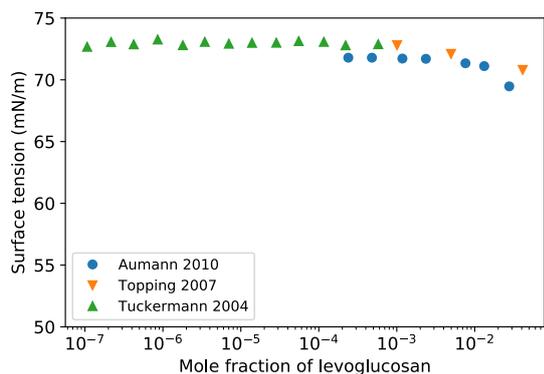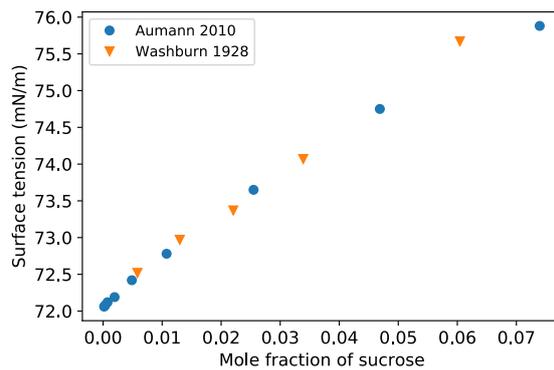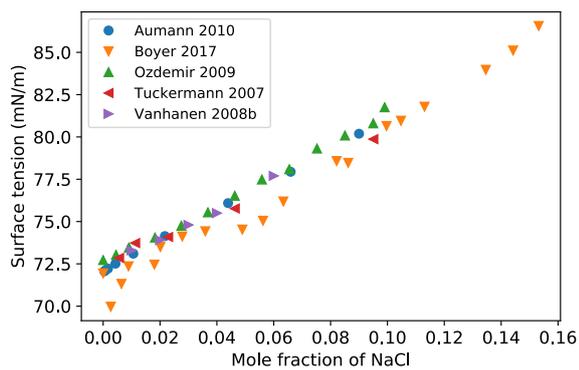
Figure S2: Surface tension of aqueous NaCl, sucrose, levoglucosan, glutaric acid, 1,2-ethanediol, methanol, butyric acid, and 1,6-hexanediol from various sources (see Table 2 in main paper for full references).
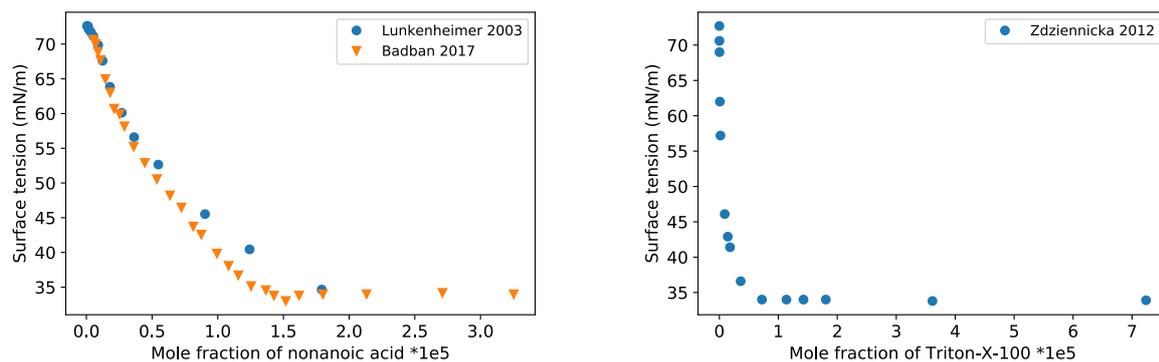
Figure S2 (continued): Surface tension of aqueous nonanoic acid and Triton X-100 from various sources (see Table 2 in main paper for full references).

2 AIOMFAC activities
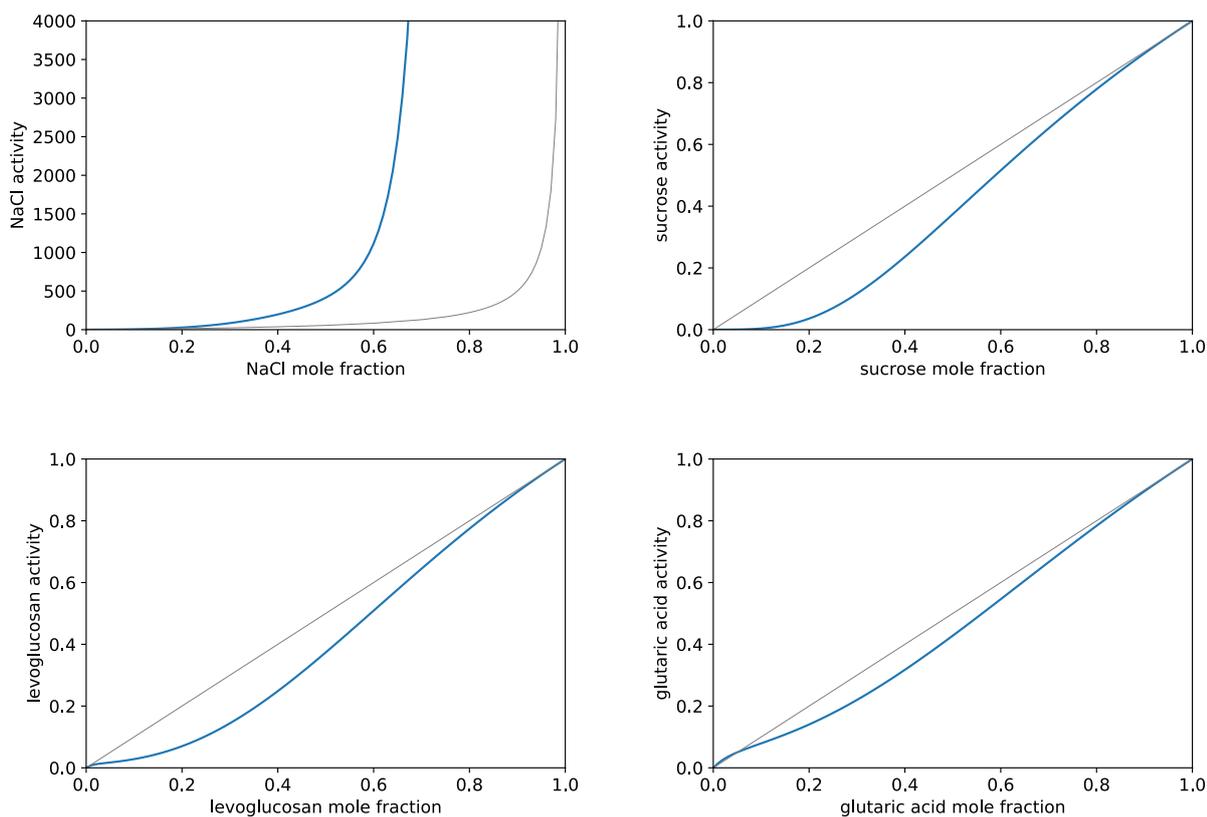
Figure S3: AIOMFAC activities (blue line) for four of the tested substances versus their mole fraction in aqueous solution. For NaCl, the activity coefficients are molality based and with a reference state of the 1–molal solution. Due to the conversion to a mole fraction basis, the activity goes to infinity (see Section 5 for more information). For all other substances, the activity coefficients are mole fraction based and with a reference state of the pure solute. The activities of the ideal solution is shown for reference (gray line).
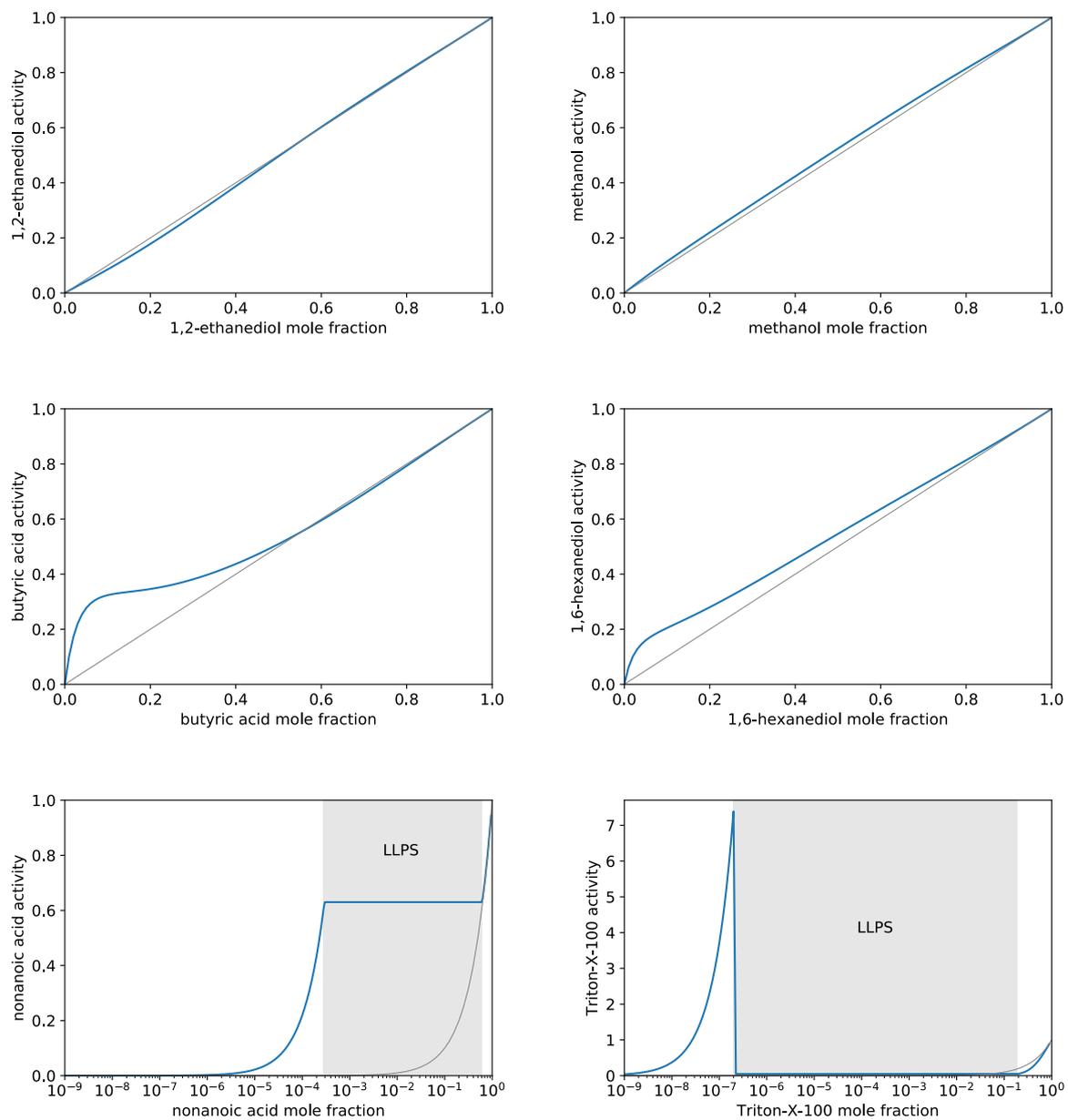
Figure S3 (continued): AIOMFAC activities (blue line) for six of the tested substances versus their mole fraction in aqueous solution. For all substances, the activity coefficients are mole fraction based and with a reference state of the pure solute. The activities of the ideal solution is shown for reference (gray line). For nonanoic acid and Triton X-100, liquid–liquid phase separation (LLPS) was predicted in the grey shaded concentration range, which is responsible for the plateau in the activity-mole fraction graphs.

4

3 Szyszkowski–Langmuir equation
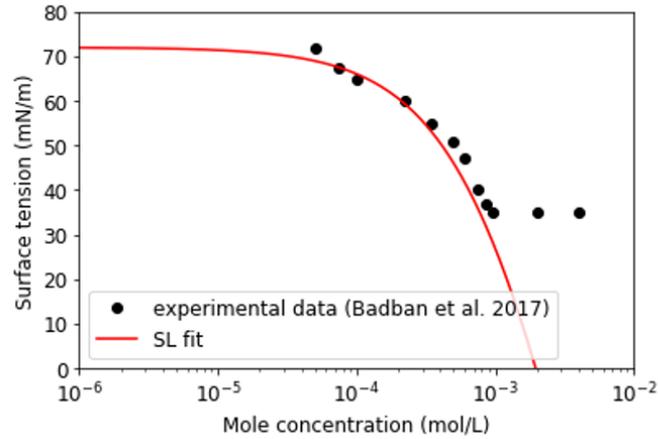
Figure S4: Szyszkowski–Langmuir (SL) fit of the surface tension of aqueous nonanoic acid.

4 Butler equation for different A_i

For the calculation of the molar surface area A_i of the substances in the mixture, the critical molar volume was calculated with UManSysProp (Topping *et al.*, 2016), using the method by "Nannoolal" for the critical molar volume and the boiling point estimation. The molar volume was calculated as $V_i = \widetilde{M}/\rho_i$ with the molecular weight $\widetilde{M}$ and the subcooled liquid density ρ_i from UManSysProp, using the method by "Girolami" and critical properties by "Nannoolal".
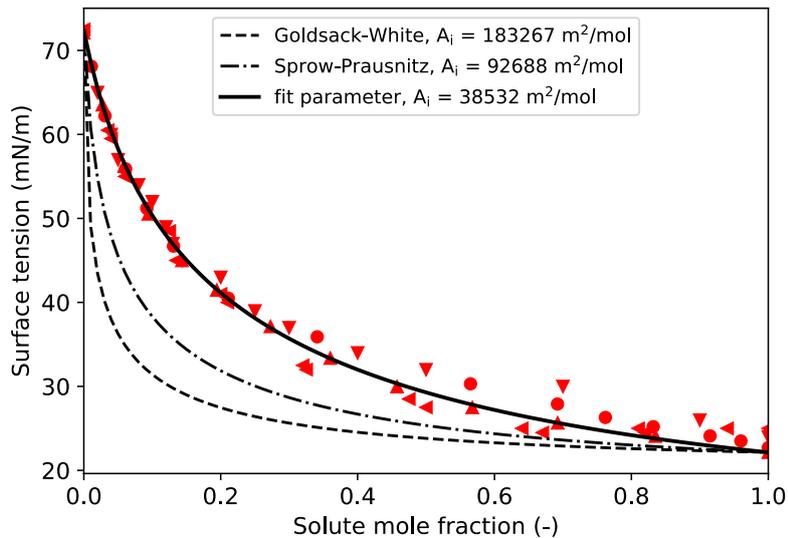
Figure S5: Sensitivity of the Butler equation to the choice of A_i: Experimental data of surface tension of aqueous methanol (red symbols) and the prediction/fit of the Butler equation (lines)

5 Activity coefficients of NaCl

Activity coefficients for NaCl calculated with AIOMFAC are Pitzer-based which means that they differ from the activity coefficients for all other tested substances in three ways:

- They are ion specific: For a solution of water and NaCl, an activity coefficient is returned for water (γ_w), one for the sodium ion Na$^+$ ($\gamma_{\mathrm{Na}+}$) and one for the chloride ion Cl$^-$ ($\gamma_{\mathrm{Cl}-}$). The mean activity coefficient of a salt is defined as

$$\gamma_\pm = \left[\gamma_+{}^{\nu^+} \cdot \gamma_-{}^{\nu^-} \right]^{1/(\nu^+ + \nu^-)}, \tag{1}$$

where ν^+ and ν^- are the stoichiometric numbers of cations (+) and anions (-) respectively (Zuend *et al.*, 2008), which for NaCl simply gives $\gamma_{NaCl} = \gamma_{\mathrm{Na}+} = \gamma_{\mathrm{Cl}-}$.

- They are molality based, which means that the activity a_{NaCl} is calculated as $a_{NaCl} = m_{NaCl} \cdot \gamma_{NaCl}$, where m_{NaCl} is the molality of NaCl in the mixture in $\mathrm{mol\,L^{-1}}$. In contrast to that, for all other tested substances the activity is calculated by multiplication with the mole fraction, i.e. as $a_i = x_i \cdot \gamma_i$.

- They are based on a reference state of the 1–molal solution. This means that $a_{NaCl} = 1$ at a molality $m_{NaCl} = 1\,\mathrm{mol\,L^{-1}}$ and therefore the activity can reach values larger than one. In contrast to that, for all other tested substances the activity coefficients are based on a reference state of the pure solute, which means that $a_i = 1$ for a mole fraction $x_i = 1$.

Due to these differences, the question arises whether these Pitzer-based activity coefficients of NaCl can be used in the tested models the same way as the mole–fraction-based activity coefficients or whether some modifications are required.

In the publication by Li and Lu (2001), Pitzer-based activity coefficients are used as well, therefore no problem is seen for the Li&Lu model. In the Butler equation, a conversion of the activities is not necessary, as long as the same standard and reference states are used for bulk and surface activities. If this is the case, any conversion factors cancel out, as only the ratio of surface and bulk activities appear in the equation. The Statistical model, however, was created for mole–fraction-based activities. The activity coefficients could be converted to a reference state of the pure solute by dividing by the limit $a_i(x_i \to 1) \coloneqq a_i^{x=1}$. However, the molality of pure NaCl is undefined (division by zero) and therefore a_{NaCl} approaches infinity as $x_{NaCl} \to 1$. For numerical purposes, $a_{NaCl}^{x=1}$ was set to be a value of $\approx 1e82$. Dividing all activities by this value leads to extremely small values and therefore to bad fit performance. A better result was achieved by not converting the activities and having the fit parameters accounting for the difference. Alternatively, the large values of the molality-based activities could be compensated by K, if negative values were allowed for this parameter. Due to $a_{NaCl}^{x=1} \neq 1$, when using the Statistical model for NaCl, Equation 6 in the main paper has to be replaced by

$$C = 1 - \left[1 - \left(1 - K a_{NaCl}^{x=1} \right) \exp\left(\frac{r S_w (\sigma_w - \sigma_i)}{kT} \right) \right] / \left(K a_{NaCl}^{x=1} \right) \tag{2}$$

and when using the Li&Lu model, Equation 18 in the main paper has to be replaced by

$$\sigma = \sigma_w + RT\Gamma^{max} \ln \frac{1}{1 + \left(\frac{1}{a_{NaCl}^{x=1}} \exp\left(\frac{\sigma_w - \sigma_i}{RT\Gamma^{max}} \right) - 1 \right) a_i}. \tag{3}$$

6 Extrapolation test

To test how well the models can extrapolate the surface tension from a limited subset of experimental data at low solute concentration to the surface tension at high solute concentration, we selected aqueous mixtures

of methanol and butyric acid as examples. We selected these because experimental data is available over the whole concentration range, and their surface tension measurements provide two different curve shape types. Figure S6 (left panel) shows the extrapolation of the models for methanol when only the data up to a solute mole fraction of $x_i = 0.2$ is used for the fit. It can be seen that the Butler equation and the Sigmoid model extrapolate the best (i.e., the differences between the extrapolated and experimental pure component surface tension values are only slight), while the extrapolation with the Connors–Wright model exhibits the largest deviation from the experimental values. In Figure S6 (right panel), the differences of the surface tension of pure methanol between measurement (σ_{real}) and extrapolation (σ_{pred}) are shown as a function of the extrapolation threshold (error in σ_i prediction). It can be seen that down to an extrapolation threshold of $x_i = 0.15$, the Butler equation and the Sigmoid model yield very accurate extrapolations with an error in the predicted σ_i below $2\,\text{mN}\,\text{m}^{-1}$. At a lower extrapolation threshold, large errors in the predicted σ_i values appear in all models, with the smallest error in the Butler equation ($4\,\text{mN}\,\text{m}^{-1}$), the Tamura model ($2\,\text{mN}\,\text{m}^{-1}$), the Statistical unconstrained model ($2\,\text{mN}\,\text{m}^{-1}$), and the Li&Lu model ($3\,\text{mN}\,\text{m}^{-1}$), depending on the specific extrapolation thresholds.
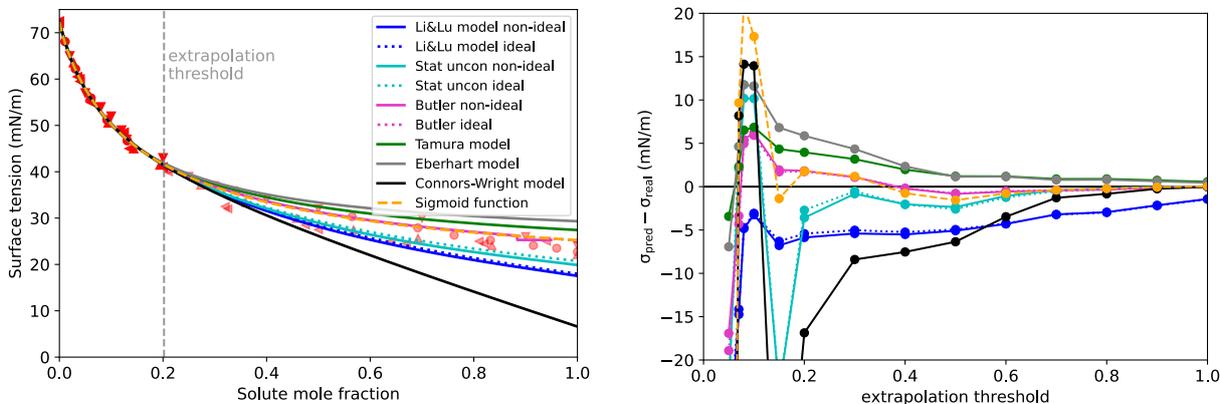
Figure S6: Extrapolation test with methanol. Left: The models were fitted to the experimental data with a solute mole fraction < 0.2 (i.e., the extrapolation threshold). The small step in the Butler non-ideal curve at high mole fraction is a result of numerical issues. Right: Difference of surface tension of pure methanol from extrapolation (σ_{pred}) and measurement (σ_{real}) for different extrapolation thresholds (markers). Connecting lines are added to guide the eye.

Analogously, Figure S7 shows the extrapolation test for butyric acid. If the given experimental data includes the part with plateauing data points, the Sigmoid, Eberhart, Stat. uncon. (ideal) and Connors–Wright models make the best extrapolations (error in $\sigma_i < 5\,\text{mN}\,\text{m}^{-1}$). However, if the plateauing part of the data is cut off (extrapolation threshold $< x_i = 0.05$), only the Butler equation (non-ideal) and the Eberhart and Connors–Wright models continue to extrapolate with moderate errors in σ_i (error $< 8\,\text{mN}\,\text{m}^{-1}$).

This extrapolation test revealed that the ability of the models to extrapolate to higher concentrations seems to depend on the substance that is being examined. However, more substances would need to be tested to confirm the results, since a small scatter in the experimental data can strongly affect the extrapolations especially for low extrapolation thresholds. Furthermore, the Butler equation (non-ideal) seems to be the most robust model in extrapolating to higher concentrations, since, except for the lowest extrapolation threshold for methanol, it never reached an error in σ_i larger then $8\,\text{mN}\,\text{m}^{-1}$. The Butler equation, however, seems to be strongly dependent on the activity coefficients. AIOMFAC, as a group contribution model, is not very accurate for larger molecules like surfactants, since the error in the contribution of a single functional group sum up to a large error if the functional group is repeated many times in one molecule. Furthermore, as a group-contribution method, AIOMFAC does not account for the position of the acid group within a molecule. This can be an explanation why the surface tension predictions for butyric acid are inaccurate at all extrapolation thresholds in this extrapolation test. AIOMFAC is also not very accurate for salts at

high concentrations (supersaturated regime) due to the lack of experimental data in that concentration range when fitting the functional groups of AIOMFAC. This might explain why it was not possible to fit σ_i for NaCl (Section 4.2 and Figure 4 in the main paper). To conclude, the Butler equation seems to be a robust model for extrapolation of limited data as long as accurate, solute specific activity coefficients are used. If the latter cannot be provided, we suggest to extrapolate the surface tension with the Sigmoid model, however, only if the experimental data covers a sufficiently large concentration range to constrain the curvature.
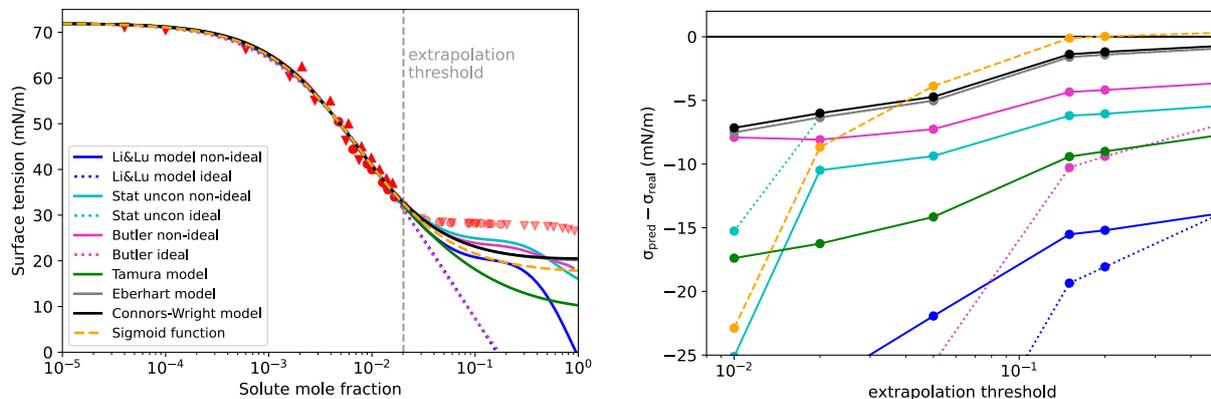
Figure S7: Extrapolation test with butyric acid. Left: The models were fitted to the experimental data with a solute mole fraction < 0.02 (= extrapolation threshold). Right: Difference of surface tension of pure butyric acid from extrapolation (σ_{pred}) and measurement (σ_{real}) for different extrapolation thresholds (markers). Connecting lines are added to guide the eye.

References

D. Topping, M. Barley, M. K. Bane, N. Higham, B. Aumont, N. Dingle and G. McFiggans, *Geoscientific Model Development*, 2016, **9**, 899–914.

A. Zuend, C. Marcolli, B. P. Luo and T. Peter, *Atmospheric Chemistry and Physics*, 2008, **8**, 4559–4593.

Z. Li and B. C.-Y. Lu, *Chemical Engineering Science*, 2001, **56**, 2879–2888.