EFFECT OF THE SOLUBILITY MODEL ON ANTISOLVENT CRYSTALLIZATION PREDICTED VOLUME MEAN SIZE

D. J. Widenski, A. Abbas, J. A. Romagnoli

Chemical Engineering Department, Louisiana State University, USA
School of Chemical and Biomolecular Engineering, University of Sydney, Australia

The use of predictive solubility models can be of great use for crystallization modelling, and can decrease the amount of experimental data needed to create a robust crystallization model. Predictive solubility models such as MOSCED, UNIFAC, NRTL-SAC, and the Jouyban-Acree model are compared against an empirical model for predicted solubility accuracy. The best models are subsequently compared against the empirical model for the antisolvent crystallization of acetaminophen in acetone using water as antisolvent. The effect of these solubility models on the predicted relative supersaturation and volume mean size profiles are investigated.

1. INTRODUCTION

Crystallization is a widely used unit operation. Since it can produce high purity products it is used for the production of pharmaceuticals, agrochemicals, and fine chemicals. The production of these products is a multimillion dollar industry. Any methods to improve the production of these products would be highly valued.

Optimal crystallization operation often requires a crystallization model. This crystallization model, typically based on population balances, requires a companion solubility model. The solubility prediction is an important aspect of any crystallization model because its prediction is the basis for any crystallization phenomena. Crystallization is caused by supersaturation which is defined as the difference between the solution concentration and the equilibrium concentration (absolute supersaturation), or the ratio of the solution and equilibrium concentrations (relative supersaturation).

Empirical solubility models have been extensively used in crystallization modelling (Zhou et al., 2006; Nagy et al., 2007; Nowee et al., 2008). It is of interest to understand how more predictive solubility models such as the MOSCED, NRTL-SAC, UNIFAC, and Jouyban-Acree models, can be incorporated into crystallization models and how their accuracy of predicting the solubility profiles influences the crystallization model prediction and optimal profile calculation. The outcome of combining predictive solubility modelling with the crystallization model is expected to reduce the need for solubility experimental data and consequently streamline the optimization of the crystallization process.

This paper investigates the effect of these solubility models on antisolvent crystallization predictions. Both the effect of the model on the predicted relative supersaturation profile and the effect on the volume mean size will be evaluated. The simulation software package gPROMS (Process Systems Enterprise Limited, London) will be used for our crystallization studies. The general process modelling system (gPROMS®) is an equation-oriented high-level declarative modelling, simulation, and optimization package. It allows the development of hierarchical models of arbitrary depth involving a range of process models including distributed systems and process with discontinuities. gPROMS directly supports simulation, parameter estimation, and optimization activities as well as, following the CAPE-OPEN standards, provides an open environment and can be used in

Please cite this article as: Widenski D., Abbas A. and Romagnoli J. (2009). Effect of the solubility model on antisolvent crystallization predicted volume mean size, AIDIC Conference Series, 09, 355-364 DOI: 10.3303/AOS0909041
real-time applications. The acetaminophen in acetone with water as the antisolvent is used in this study as antisolvent crystallization system.

2. SOLUBILITY MODELS

Several predictive solubility models were considered for this study namely the MOSCED, UNIFAC, NRTL-SAC, and Jouyban-Acree models. Each of these models is capable of predicting the equilibrium profiles without additional solubility data, except for the Jouyban-Acree model. The Jouyban-Acree model requires two solubility data points, the solubility of the solute in the pure solvent and the solubility of the solute in the pure antisolvent. The ternary acetaminophen-acetone-water system is used as a model system in this study at 16 °C.

The MOSCED model (Lazarzoni et al., 2005), generates infinite dilution activity coefficients. In order to obtain a non infinite dilution activity coefficient, another activity coefficient model is required. The Van Laar, Wilson, and NRTL models were each combined with the MOSCED model to evaluate which would give the best prediction to known experimental data. The MOSCED model is described by equation set 1.

\[
\ln y_{ij} = \frac{v_i^0}{RT} \left[ (\lambda_i - \lambda_j)^2 + \frac{q_i^2 q_j^2 (\tau_j - \tau_i)^2}{\psi_j} + \frac{(\alpha_j - \alpha_i) (\beta_j - \beta_i)}{\xi_j} \right] + d_{ij}
\]

\[
d_{ij} = \ln \left( \frac{v_i^0}{v_j^0} \right)^{aa} + 1 - \left( \frac{v_i^0}{v_j^0} \right)^{aa}
\]

\[
\alpha_j = 0.953 - 0.002314 [(\tau_j)^2 + \alpha_i \beta_j]
\]

\[
\psi_j = POL + 0.002629 \alpha_i \beta_j
\]

\[
\xi_j = 0.68 (POL - 1) + \left[ 3.24 - 2.4 \exp \left( -0.002687 (\alpha_i \beta_j)^{1.5} \right) \right]^{293 \over T_j}^2
\]

\[
POL = q_j^0 \left[ 1.15 - 1.15 \exp \left( -0.002337 (\tau_j)^3 \right) \right] + 1
\]

\[
\alpha_j = \alpha_j^0 \left( \frac{293}{T(K)} \right)^{0.8}, \quad \beta_j = \beta_j^0 \left( \frac{293}{T(K)} \right)^{0.8}, \quad \tau_j = \tau_j^0 \left( \frac{293}{T(K)} \right)^{0.4}
\]

(1)

This representation of the MOSCED model is used to find substance j’s infinite dilution activity coefficient in substance i. Similarly, the model can be used to find the infinite dilution activity coefficient for substance i in j by switching the subscripts i and j around. The MOSCED model contains five parameters: \( \lambda, \alpha, \beta, q, \tau, \) and \( v \) corresponding to dispersion, hydrogen bond acidity, hydrogen bond basicity, induction, polarity, and molar volume respectively. Details on these parameters and their correlated values for various compounds are given in Lazarzoni et al. (2005).

The next solubility model considered is the UNIFAC model (Anderson and Prausnitz, 1978). The UNIFAC model predicts activity coefficients based on group contributions, and is described by equation set 2.

\[
\ln y_{ij} = \ln y_{ij}^C + \ln y_{ij}^R
\]

\[
\ln y_{ij}^C = \ln \frac{\phi_{i,j}^u}{x_i} + \frac{q_i}{2} \ln \frac{\phi_{i,j}}{\phi_{i,u}} + \frac{q_i \theta_{i,j}}{x_i} \sum_j x_j l_j
\]

\[
\ln y_{ij}^R = \sum_k v_k^{(i)} \left( \ln \Gamma_k - \ln \Gamma_k^{(i)} \right)
\]

\[
\ln \Gamma_k = \ln \Gamma_k^{(i)} = Q_k \left[ 1 - \ln (\sum_m \Theta_m \Psi_{mk}) - \sum_m \frac{\Theta_m \Psi_{km}}{\sum_n \Theta_n \Psi_{nm}} \right]
\]

\[
l_j = \frac{q_i x_i}{2 (r_{l,u} - q_i)} - (r_{l,u} - 1)
\]

\[
\theta_{i,j} = \frac{q_i x_i}{\sum_j q_{j,u} x_j}, \quad \phi_{i,u} = \frac{r_{i,u} x_i}{\sum_j r_{j,u} x_j}, \quad \Theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n}, \quad \Psi_{mn} = \exp \left( \frac{\alpha_{mn}}{\tau} \right)
\]
\[ r_{i,U} = \sum_k \gamma_{i,k}^{(0)} R_k \quad , \quad q_{i,U} = \sum_k \eta_{i,k}^{(0)} Q_k \] (2)

Where \( \gamma_i \) is the activity coefficient of component \( i \), \( \gamma_i^{(0)} \) and \( \gamma_i^k \) are the combinatorial and residual parts of the activity coefficient of component \( i \), \( \Gamma_k \) is the residual activity coefficient of group \( k \), \( \Gamma_k^{(0)} \) is the residual activity coefficient of group \( k \) in a reference solution containing only groups of type \( i \), \( z \) is the coordination number, \( r_{i,U} \) is the volume structural parameter of component \( i \), \( r_{i,U} \) is the area structural parameter of component \( i \), \( \Theta_i \) is the area fraction of component \( i \), \( \Phi_i \) is the volume fraction of component \( i \), \( \Theta_m \) is the area fraction of group \( m \), \( Q_i \) is the volume structural parameter of group \( k \), \( R_k \) is the area structural parameter of group \( k \), \( y_{i,k}^{(0)} \) is the number of \( k \) groups in component \( i \), \( x_i \) is the mole fraction of component \( i \), \( X_m \) is the mole fraction of group \( m \) in the mixture. \( \Psi_{mn} \) is the group interaction parameter between \( m \) and \( n \), and \( a_{mn} \) is the measure of interaction between groups \( m \) and \( n \). The UNIFAC parameters for water, acetone, and acetaminophen molecules were calculated from the individual functional group parameters found in Poling, Prausnitz, and O’Conell (2000). Since the acetaminophen aromatic NH functional group was not listed, it was approximated as an aromatic NH$_2$ functional group in the same manner as Hojjati and Rohani (2006). The MOSCED and UNIFAC models predicted equilibrium profiles for paracetamol in acetone and water are shown in Figure 1.

![Figure 1: MOSCED & UNIFAC solubility predictions](image)

The MOSCED models all give very poor solubility predictions. They all greatly underestimate the solubility. The NRTL and Wilson models give better estimates to the shape of the solubility curve than the Van Laar model does. The UNIFAC model is the worst of the models both greatly overestimating the solubility and weakly representing the shape of the curve.

The next solubility model considered is the NRTL-SAC model (Chen et al., 2004, 2006). The NRTL-SAC model is a NRTL activity coefficient model that they modified using segment theory in a similar way as the polymer NRTL model. The NRTL-SAC model is described by equation set 3.
\[
\ln y_i = \ln y_i^C + \ln y_i^R
\]

\[
\ln y_i^C = \ln \frac{\phi_i}{x_i} + 1 - r_i \sum_j \frac{\phi_j}{r_j}
\]

\[
\ln y_i^R = \sum_m r_{m,i} (\ln \Gamma_m^C - \ln \Gamma_m^R)
\]

\[
\ln \Gamma_m^C = \frac{\Sigma_j x_{i,j} \Gamma_{j,m} \Gamma_{j,m}}{\Sigma_k x_{k,j} \Gamma_{k,m}} + \frac{\Sigma_m x_{m,j} G_{m,m}}{\Sigma_k x_{k,j} G_{k,m}} \left( \tau_{m,m} - \frac{\Sigma_n x_{n,m} G_{n,m}}{\Sigma_k x_{k,j} G_{k,m}} \right)
\]

\[
\ln \Gamma_m^R = \frac{\Sigma_n x_{n,m} \Gamma_{n,m} \Gamma_{n,m}}{\Sigma_k x_{k,j} \Gamma_{k,m}} + \frac{\Sigma_m x_{m,j} G_{m,m}}{\Sigma_k x_{k,j} G_{k,m}} \left( \tau_{m,m} - \frac{\Sigma_n x_{n,m} G_{n,m}}{\Sigma_k x_{k,j} G_{k,m}} \right)
\]

\[
x_n = \frac{\sum_i x_{i,n}}{\sum_i \sum_m x_{i,m} r_{m,i}}, \quad x_{n,i} = \frac{r_{n,i}}{\sum_m r_{n,i}}, \quad r_i = \sum_m r_{m,i}, \quad \phi_i = \frac{r_i x_i}{\sum_j r_j x_j}, \quad G_{km} = e^{-a_{km} \tau_{km}} \tag{3}
\]

Where \( y_i^C \) and \( y_i^R \) are the combinatorial and residual contributions to the activity coefficient of component \( i \). \( k, l, m, n \), and \( n \) are the segment indices, \( i \) and \( j \) are the component indices, \( x_n \) is the segment-based mole fraction of segment species \( n \), \( x_{n,i} \) is the mole fraction of component \( i \), \( x_{n,i} \) is the segment fraction of segment species \( n \) in component \( i \), \( \Gamma_m^C \) is the activity coefficient of segment species \( m \), \( \Gamma_m^R \) is the activity coefficient of segment species \( m \) in component \( i \), \( r_{n,i} \) is the number of segment species \( m \) in component \( i \), \( r_i \) is the total segment number of component \( i \), and \( \phi_i \) is the segment mole fraction of component \( i \). \( \tau_{km} \) and \( a_{km} \) are the NRTL-SAC binary interaction parameter and nonrandomness parameter between segments \( n \) and \( m \) respectively. These parameters are tabulated in Chen and Crafts (2006) for each segment pair. Further details about the development of the NRTL-SAC model can be found in Chen and Song (2004) and Chen and Crafts (2006).

Furthermore, another solubility model considered is the Jouyban-Acree model (Jouyban et al., 2006). The Jouyban-Acree model is a semi-empirical model developed to predict the solubility of pharmaceuticals in organic solutions. This model requires the solubilities of both pure components in a binary solute-solvent system, and predicts the solubility of a solute in a solvent mixture. The Jouyban-Acree model is described by equation 4.

\[
\log x_{2,mix} = f_1 \log x_{2,1} + f_3 \log x_{2,3} + f_1 f_3 \left[ \frac{C_0}{T} + \frac{C_1 (f_1 - f_2)}{T} + \frac{C_2 (f_1 - f_2)^2}{T} \right] \tag{4}
\]

Where \( f_1 \) is the solute free volume fraction of the solvent, \( f_3 \) is the solute free volume fraction of the antisolvent, \( x_{2,1} \) is the solubility of the solute in pure solvent, \( x_{2,3} \) is the solubility of the solute in pure antisolvent, \( x_{2,mix} \) is the solubility of the solute in the solvent mixture, and \( C_0, C_1, \) and \( C_2 \) are constants.

The last solubility model considered is an empirical model generated from data from Granberg and Rasmusson (2000) by Zhou et al. (2006).

\[
C^* = -5.01902 \times 10^{-12} w^6 + 1.69767 \times 10^{-9} w^5 - 2.46765 \times 10^{-7} w^4 + 2.19262 \times 10^{-5} w^3 - 1.27018 \times 10^{-3} w^2 + 3.42614 \times 10^{-2} w + 7.96086 \times 10^{-2} \tag{5}
\]

Where \( C^* \) is the equilibrium concentration (kg paracetamol/kg solvents), and \( w \) is the solute free mass percent of water. The Jouyban-Acree, NRTL-SAC, and empirical solubility predictions are shown in figure 2.
The NRTL-SAC and Jouyban-Acree solubility models both predict the equilibrium solubility much better than the MOSCED or UNIFAC models did. The empirical model fits the data very well and will be considered as the standard solubility model for benchmarking. Since the UNIFAC and MOSCED models gave such poor solubility predictions, only the NRTL-SAC and Jouyban-Acree models will be compared against the empirical model for further crystallization volume mean size sensitivity studies.

3. CRYSTALLIZATION MODEL & SIMULATION

In order to evaluate the effect of the solubility model on the predicted crystal properties a crystallization model is needed. The crystallization model is comprised of a population balance with corresponding crystallization kinetics, a mass balance, and a solubility model (Nowee et al. 2008; Widenski et al. 2009). The population balance considered was for a crystallization system with size-independent crystal growth, and with negligible attrition and agglomeration.

\[
\frac{\partial n(L, t)}{\partial t} + \frac{n(L, t) \frac{dV}{dt}}{V} + G \frac{\partial n(L, t)}{\partial L} - B = 0
\]

(6)

Where \( n(L, t) \) is the crystal density (\( \# \) of particles/m\(^3\)), \( V \) is the volume (m\(^3\)), \( G \) is the growth rate (m/s), and \( B \) is the nucleation rate (\( \# \) of particles/ s m\(^3\)).

The acetaminophen in acetone with water as the antisolvent is used in this study as anti-solvent crystallization system. The antisolvent crystallization kinetics for acetaminophen in acetone with water as the antisolvent was taken from Zhou et al. (2006). The authors developed their kinetics from previous crystallization data performed by Granberg et al. (1999, 2001).

\[
N = 8.56080 \times 10^8 \exp \left\{ -1.22850 \times 10^{-3} \frac{\ln^3 \left( \frac{\rho_c}{\rho_a} \right)}{\ln^2 \left( \frac{C}{C^*} \right)} \right\}
\]

(7)

\[
G = k_G (C - C^*)^\theta
\]

(8)

\[
k_g = 4.01067 \times 10^{-8}w^2 - 1.76198 \times 10^{-6}w + 5.78135 \times 10^{-5}
\]

(9)
\[ g = -4.22536 \times 10^{-3} \omega + 1.77428 \]  

(10)

Where \( N \) is the nucleation rate (no. of particles/m\(^3\)), \( \rho_c \) is the crystal density of acetaminophen (kg/m\(^3\)), \( C^* \) is the equilibrium concentration defined previously, \( C \) is the solution concentration (kg acetaminophen/kg solvents), \( \rho_s \) is the density of the solution (kg/m\(^3\)), \( G \) is the crystal growth rate (m/s), and \( s \) is the solute free mass percent of antisolvent (water) in the solution. Also, the growth kinetics is only valid for solute free water mass percents greater than 30%.

The simulations were executed using the gPROMS modelling package (Process Systems Enterprise, UK). The population balance was solved by backward finite difference discretization using 250 geometrically spaced intervals across the size axis from 0.5-1000 microns. To evaluate the effect of using different solubility sub-models on the predicted crystallization results two different antisolvent feed rates were used specifically 25 and 400 g/min. Both antisolvent simulations were done isothermally at 16 °C. The initial antisolvent solute-free mass fraction of water was 0.3 and the antisolvent feedrate was added until the solute free mass fraction of water reached 0.8. The initial solvent mass was 1 kg, and the amount of acetaminophen added was equal to the equilibrium value of the solubility model used. After the antisolvent feed was stopped the simulation was continued for 10 min to consume any remaining supersaturation, and allow the solution to reach equilibrium.

![Graph](image_url)

**Figure 3: Relative supersaturation profile for 25 g/min antisolvent feedrate**

Figure 3 shows the predicted relative supersaturation profiles for the 25 g/min antisolvent feed rate for the different solubility models. The empirical and Jouyban-Acree models result in similar profiles while the NRTL-SAC model deviates significantly. The variation in the NRTL-SAC supersaturation profile is due to the divergence of the NRTL-SAC equilibrium profile past 0.3 solute free mass fraction of water from the experimental data. The NRTL-SAC model predicts that the solubility of acetaminophen in the mixture does not decrease at the same composition that the empirical and Jouyban-Acree models decrease. It actually predicts a slight increase in equilibrium solubility which causes the decrease in relative supersaturation. Figure 4 shows the relative supersaturation profile for the 400 g/min feedrate. It shows the same behaviour as the lower feedrate with the exception that the profiles peak higher. This is expected because the higher feedrate should cause more nucleation due to excessive supersaturation.
The end product volume mean sizes (taken at time = 95 min) for the 400 g/min feedrate are, as expected, smaller than those for the 25 g/min feedrate. Figures 5 and 6 illustrate this situation, using the empirical model as a basis. The 400 g/min feedrate results in a volume mean size of 156 microns, but under a feedrate of 25 g/min, the volume mean size increases to 192 microns. These figures also show that the NRTL-SAC model overpredicts, while the Jouyban-Acree model underpredicts the volume mean size for both feedrates. In addition, the figures show that the prediction is more dependent to the solubility model for the 25 g/min feedrate than the 400 g/min feedrate. Intermediate feedrates between 25 and 400 g/min were then simulated with results displayed in Table 1.

Since the empirical model best matches the experimental data, its predictions are assumed to be accurate, and are used to assess the other two models. At the higher feedrate the prediction for both models improves. This is attributed to the relative difference in the nucleation rates under the two feed profiles.
The nucleation rate is lower under the lower feedrate while crystal growth becomes the dominant kinetic. The solubility model affects the size prediction over a longer period of time accumulating a larger error. In other words, when nucleation is the primary mechanism (i.e. under higher antisolvent feed rates), crystal formation occurs very rapidly. Thus, the supersaturation curve only affects crystallization for a shorter period of time. This follows that the solubility profile is in turn only influential on the size prediction for a shorter period of time. In contrast, when the crystallization process is growth dominated, the crystallization process is more affected by the longer enduring supersaturation profile. The Jouyban-Accree model results in an almost perfect size prediction in the early stages of the crystallization as opposed to the delayed nucleation of the NRTL-SAC model. This is attributed to the divergence of the NRTL-SAC equilibrium profile from the experimental data past 0.3 solute free mass fraction of water and such error will propagate into model-based optimisation calculations. A common crystallization objective function is to minimize undesired nucleation. We have shown, in a companion paper that the solubility model has a large affect on the predicted optimal antisolvent feed profiles (Widenski et al., 2009).

<table>
<thead>
<tr>
<th>Antisolvent Feedrate (g/min)</th>
<th>Solubility Model Predicted Volume Mean Size (microns)</th>
<th>% Error</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Empirical</td>
<td>Jouyban-Accree</td>
<td>NRTL-SAC</td>
</tr>
<tr>
<td>25</td>
<td>192</td>
<td>167</td>
<td>13.0</td>
</tr>
<tr>
<td>50</td>
<td>170</td>
<td>150</td>
<td>11.8</td>
</tr>
<tr>
<td>100</td>
<td>158</td>
<td>142</td>
<td>10.1</td>
</tr>
<tr>
<td>200</td>
<td>154</td>
<td>141</td>
<td>8.4</td>
</tr>
<tr>
<td>400</td>
<td>156</td>
<td>144</td>
<td>7.7</td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

This paper evaluated different predictive solubility models for use in crystallization population balance formulations. The acetaminophen in acetone with water as the antisolvent is used in this study as anti-solvent crystallization system.

Several predictive solubility models were considered for this study namely the MOSCED, UNIFAC, NRTL-SAC, and Jouyban-Accree models. The MOSCED and UNIFAC models were very poor predictors of the equilibrium solubility. However, the Jouyban-Accree and NRTL-SAC predictions showed closer agreement to experimental data.
The models predictions of relative supersaturation and volume mean size were shown to be significantly influenced by the solubility predictions errors, even for the better Jouyban-Acree and NRTL-SAC models, thus highlighting that caution is needed in selecting the right solubility formulation. As the antisolvent feedrate decreased, the solubility model error had a greater effect on the predicted volume mean size. Not only affecting the mean size, the solubility model also can also predict delayed nucleation phenomena, such as for the NRTL-SAC model. Since the solubility model is hypothesized to have a greater effect on growth dominant antisolvent crystallization experiments, this may have a large effect on both seeded antisolvent crystallization and antisolvent optimization.

5. REFERENCES


