



A Review of Experimental Solubilities and a General Correlation between the Temperature-Dependent Solubility and Solute and Solvent Molar Masses for Binary n-Alkane Mixtures

Sverre Gullikstad Johnsen*

SINTEF, TRONDHEIM, NORWAY

ARTICLE INFO

Article history:

Received 18 June 2021

Revised 3 January 2022

Accepted 5 January 2022

Available online 8 January 2022

Keywords:

Alkane

Paraffin

Precipitation

Solubility

Solid-liquid equilibrium,

ABSTRACT

The solubility of a "heavy" alkane (solute) in a "light" alkane (solvent) is generally temperature dependent. Moreover, it is determined by the molar masses of the solute and solvent. In the current paper, published solubility data for binary normal-alkane mixtures is reviewed (solid-liquid equilibrium). A total of 43 unique solute-solvent data-sets, obtained from a total of 24 published experimental studies, are collected and presented in a systematic manner. Based on thermodynamic considerations and the experimental data, it is demonstrated that there is a log-linear relationship between the solubility and the temperature in the dilute range. Linear regression is employed to 1) obtain data-set-specific solubility-temperature best-fit parameters and 2) obtain a general correlation between the solubility and the solvent and solute molar masses and the temperature. Finally, it is demonstrated that the developed correlation carries predictive power even for multi-component mixtures by utilizing solvent and solute average molar masses.

© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Petroleum waxes are mainly associated with the aliphatic fraction of crude oil, and normal-alkanes, forming needle-like macro-crystals, are recognized as the main contributor to solid deposits forming during e.g. production and transportation of oil and gas [1]. Significant costs are associated with the prevention and mitigation of wax precipitation and deposition since accumulation of solid wax in oil pipe-lines may lead to increased operational expenses (e.g. compressors, heating, chemical inhibitors, and man-hours) and reduced production (e.g. diminished flow capacity and periods of shut-in), or in the worst-case abandonment of the entire field [2]. Being able to predict solid-liquid equilibrium in petroleum systems, is paramount in developing and designing transporting and processing solutions. Naturally, the oil industry's need for accurate predictions of wax deposition has led to immense research activity on e.g. thermodynamic modelling and experimental characterisation of waxy oils. State-of-art modelling allows accurate predictions of phase-behaviour in complex fluids, see e.g. Coutinho et al. [3], Heidariyan et al. [4], Shahdi

and Panacharoensawad [5], Wang and Chen [6], and Ghasemi and Whitson [7].

Simplified model oils are commonly utilized in wax deposition studies and for validation of mathematical models, e.g. by Singh et al. [8], Paso and Fogler [9], Wu et al. [10], and Johnsen et al. [11]. Although advanced, highly accurate thermodynamic models exist, for wax formation in hydrocarbon systems, the author's own experience as well as that of colleagues show that it is sometimes useful to have a quick and easy way of calculating rough estimates, without the need for extensive input data. This could be to e.g. evaluate suggested model oils, to initialize computational models, just produce semi-realistic data for student assignments, or whenever it is impractical to perform a rigorous thermodynamic calculation.

In this paper, published solubility data for binary normal-alkane mixtures is reviewed. A total of 43 unique solute-solvent data-sets, obtained from a total of 24 published experimental studies, are collected and presented in a systematic manner. It is demonstrated that there is a log-linear relationship between the solubility and the temperature in the dilute range. Linear regression is employed to 1) obtain data-set-specific solubility-temperature best-fit parameters and 2) obtain a general correlation between the solubility and the solvent and solute molar masses and the temperature. Finally, it is demonstrated that the developed correlation carries predictive

* Corresponding author.

E-mail address: sverre.g.johnsen@sintef.no

Nomenclature

A, B	Linear regression parameters.
ΔC_{pm}	Molar specific heat capacity difference between solid and liquid states of the solute, at the melting point.
\mathbf{e}	Column vector of errors.
$f(\mathbf{e})$	Best-fit error functional.
f_1, f_2, f_3	Solute- and solvent-dependent solubility parameters.
ΔH_m	Solute molar enthalpy of fusion.
ΔH_{tr}	Solute molar enthalpy of solid-solid transition.
I	Number of data-points.
J	Number of regression parameters.
k	Huber loss function parameter.
M	Molar mass.
MAD	Median of absolute deviations.
N	Number of carbon atoms in alkane-chain.
R	Universal gas constant.
T	Temperature.
T_m	Melting temperature
T_{tr}	Solid-solid transition temperature.
\mathbf{X}	$I \times J$ matrix containing experimental parameters.
x_s	Solute solubility (mol-fraction).
\mathbf{Y}	Column vector of calculated data.
\mathbf{y}	Column vector of experimental data.
<i>Greek Symbols</i>	
$\boldsymbol{\beta}$	Column vector of model parameters.
$\boldsymbol{\rho}$	Error vector function.
γ	Activity coefficient.
σ	Standard deviation.
<i>Sub/super-scripts</i>	
1	Solvent.
2	Solute.
$\hat{}$	Best-fit parameter.
i, j	Matrix row and column indices.

power even for multi-component mixtures by utilizing solvent and solute average molar masses.

2. Definitions

Normal alkanes, also known as normal paraffins, are straight carbon-chain molecules saturated with hydrogen atoms such that no branches or double-bonds exist. The different alkanes are denoted by CN, where N indicates the number of carbon atoms in the chain, the *carbon-number*. It is common to add the prefix *n* to identify straight-chain alkanes. In this paper, however, the *n* prefix is left out. The number of hydrogen atoms in the alkane molecule is given by $2(N + 1)$, and the *molar mass* of an alkane is given by $M(N) = (14.026N + 2.016)g/mol$. A *binary system* is a mixture consisting of two alkane species, only. The *solute* refers to the heavier of the two alkane species, and the *solvent* refers to the lighter. The *solubility* (saturation mol-fraction) is defined as the maximum amount of solute that can be dissolved in a specific quantity of solvent, at a specified temperature. The current study is limited to solid-liquid phase equilibrium. A *data-point* is a measured solubility-temperature pair. The complete set of data-points, from all the literature references, for a specific binary system, will be referred to as a *data-set*.

Experimental parameters may vary between different experiments, e.g. the system temperature, the solvent or solute properties, or a function of these. The *model parameters* for a specific

model, however, are constant. For a set of I experimental solubilities expressed as a column vector, \mathbf{y} , a linear model in J model parameters, can be expressed as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}$ where \mathbf{Y} is the column vector of I calculated approximate data, and the experimental and model parameters are represented by the $I \times J$ matrix \mathbf{X} , where $X_{i,1} \equiv 1 \forall i$, and the column vector $\boldsymbol{\beta}$, respectively. It is required that the number of model parameters exceeds the number of experimental parameters by one.

The aim of *linear regression* is to find a linear model that can be employed to predict or estimate the experimental data with acceptable accuracy. If a model with only one experimental parameter is chosen, $Y_i = \beta_1 + X_{i,2}\beta_2$ and the procedure is referred to as *simple regression*. If, on the other hand, a model of two or more input parameters is chosen, $Y_i = \beta_1 + \sum_{j=2}^J X_{i,j}\beta_j$ and the procedure is referred to as *multiple regression*. The *error vector* is defined as $\mathbf{e} = \mathbf{y} - \mathbf{Y}$ and the *best-fit model parameters* are given by the vector $\hat{\boldsymbol{\beta}}$, that minimizes the *error functional*, $f(\mathbf{e}) = \sum_i^N \rho_i(e_i)$ where each component of the vector function $\boldsymbol{\rho}$ is a function of the corresponding component of the error vector. The choice of statistical method determines the representation of the error functional.

3. Published Solubility Data

Experimental solubility data utilized in correlation development in the current paper, have been obtained from 24 publications, for a total of 43 binary systems with solute carbon-numbers ranging from 8 to 36 and solvent carbon-numbers ranging from 3 to 14. Most data were reported for atmospheric pressure, but propane and butane are gaseous at atmospheric pressure, so the works by Seyer and Fordyce [12] and Godard [13] were performed at the vapour pressure of the solvent. The solubility data references are summarized in Table 1, and the complete set of experimental data is plotted in Fig. 1.

Campestrini and Stringari [14] and Tassin et al. [15] provide comprehensive lists of references to published solid-fluid equilibrium data for n-alkanes in methane (C1), ethane (C2), and propane (C3). These data were not employed for correlation development.

Fig. 1 suggests that solubilities can be approximated by a log-linear relationship with the system temperature in the dilute range. The deviation from this behaviour increases for increasing solubility, and somewhat arbitrarily it was decided to focus on data below a mol-fraction of 0.1. This disqualifies, from the current study, eleven data-sets due to lack of data-points below the limit (C8-C6, C18-C6, C18-C7, C18-C10, C19-C7, C20-C6, C20-C10, C22-C6, C22-C7, C25-C7, C25-C14).

4. Theory and Method

The solubility of solids in liquids, in terms of the mol-fraction, can be modelled by [38,39]

$$\ln x_s = \ln f_1 + f_2 \ln T + f_3/T, \quad (1)$$

where $f_1 \equiv f_1(M_1, M_2) = \frac{T_m^{-f_2}}{\gamma_2} \exp \left[f_2 + \left(\frac{\Delta H_m}{T_m} + \sum_{tr} \Delta H_{tr} T_{tr} \right) / R \right]$, $f_2 \equiv f_2(M_2) = -\Delta C_{pm} / R$, and $f_3 \equiv f_3(M_2) = [f_2 T_m - (\Delta H_m + \sum_{tr} \Delta H_{tr}) / R]$, M_1 and M_2 are the solvent and solute molar masses, respectively, γ is the activity coefficient, ΔH_m is the molar solute enthalpy of fusion, R is the universal gas constant, T is the system temperature, T_m is the solute melting-point temperature, ΔC_{pm} is the molar heat capacity difference between solid and liquid states of the solute at T_m , and ΔH_{tr} and T_{tr} are the molar enthalpy and temperature of solid-solid phase transitions, respectively.

The 1st order Taylor expansion of Eq. (1), around some temperature $T^* < T_m$, is

$$\ln x_s \approx A + B \cdot T, \quad (2)$$

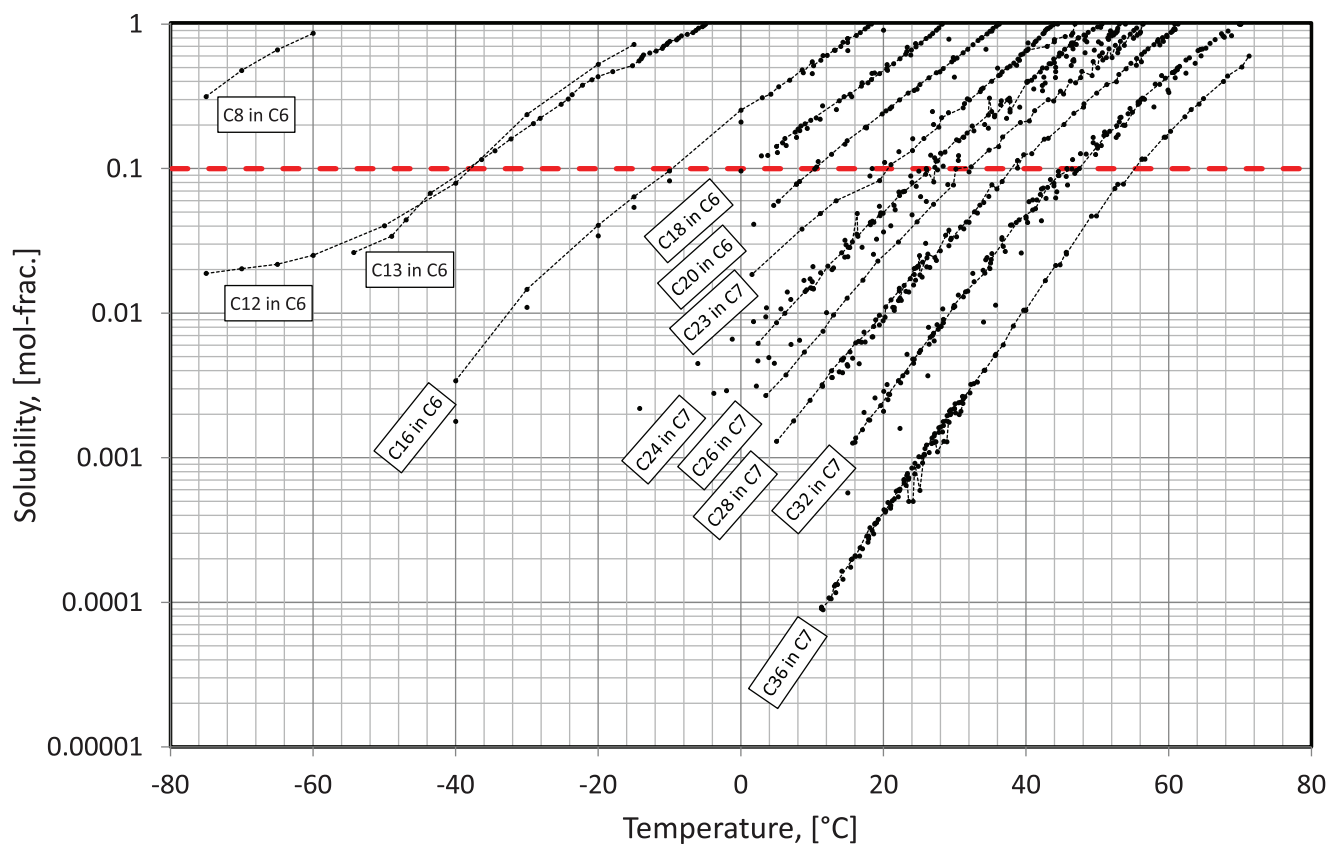


Fig. 1. The complete set of experimental solubility-temperature data. Selected data-sets have been emphasized with labels and black dotted lines, and the 0.1 mol-fraction line is emphasized as the red dashed line.

where $A \equiv A(M_1, M_2, T^*) = \ln f_1 + f_2 \ln T^* + f_3/T^* - BT^*$, and $B \equiv B(M_2, T^*) = f_2/T^* - f_3/T^{*2}$. Due to the observed log-linearity of the solubility, below the mol-fraction of 0.1, A and B are expected to be independent of T^* . It has been assumed that the activity coefficient, melting and transition temperatures, heat capacities and enthalpies can be expressed as functions of the molar masses, only. That is, effects of e.g. pressure or molecular structure/nature have not been considered.

Common methods of linear regression include Least-squares regression (LS) and Least absolute deviations regression (LAD). The current paper employs the Huber loss function [40,41], which combines the strengths of the LAD (robustness) and the LS (accuracy) methods. For the Huber method, the vector function ρ may be expressed as

$$\rho_i(e_i) = \begin{cases} e_i^2 & \text{if } -k \leq e_i \leq k \\ 2k|e_i| - k^2 & \text{if } e_i < -k \text{ or } k < e_i \end{cases} \quad (3)$$

where, $k = 1.5\sigma$. σ is an estimate of the standard deviation of the population random errors, and for normally distributed errors, $\sigma = 1.483MAD$ gives a good estimate. MAD is the median of the absolute deviations, $|e_i|$. This is a robust method that performs reasonably well even when the basic assumptions of the statistics are false [41].

5. Results

To investigate the A and B (Eq. (2)) dependency on the solvent and solute molar masses, simple regression was performed on each data-set to obtain data-set-specific best-fit parameters. These are plotted against the molar masses in Figs. 2 and 3 and are cited in Table 2. In Figs. 2a and 3a linear trends are added for the four series with more than two best-fit points (C24, C28, C32 and C36).

It is clear that there are linear relationships between A and B and the molar masses. Although some of the A and B outliers seem to deviate significantly from the linear trends, no data-sets were disqualified for this reason. Some of the scatter can be explained by the dependency on the other molar mass, and in fact, the final regression expression was not very sensitive to elimination of the outliers. This owes to the robustness of the Huber method.

Having revealed a linear relationship between the regression parameters, A and B , and the molar masses, multiple regression was performed on the entire set of data-points, from all the data-sets, to obtain the best-fit model

$$\ln x_s \approx (6.435 - 6.627 \cdot 10^{-4}M_1 - 3.446 \cdot 10^{-2}M_2) + (1.499 - 2.989 \cdot 10^{-2}M_2)T/100, \quad (4)$$

where the temperature is in $^{\circ}\text{C}$, and any M_1 dependency of B is neglected in accordance with the discussion in Section 4 (Eq. (2)).

In the appendix (Fig. A.1), all the solubility data-sets are presented. In addition, data-set-specific best-fit trend-lines (Eq. (2))(red line) and the general correlation (Eq. (4))(blue line) are shown. The Eq. (4) predictions alone, are drawn for the data-sets with no data-points below the 0.1 mol-fraction.

6. Discussion

Several authors, e.g. Jennings and Weispfennig [37], have pointed out that solubility data is correlated with the solute melting temperature, so that the solubility curves collapse onto each other if plotted against $(T - T_m)$. The end-point of the solubility curve x_s , at pure solute, is of course at the solute melting point. Thus, all solubility curves should terminate in the same point, in an x_s vs. $(T - T_m)$ plot. The starting point of the curve, however, at pure solvent, will be at the melting point of the solvent, which

Table 1

Literature references to experimental solubility data for binary alkane mixtures. The references in the hatched table cells did not provide data points below the mole-fraction threshold of 0.1.

Solute/Solvent	Propane (C3)	Butane (C4)	Pentane (C5)	Hexane (C6)	Heptane (C7)	Octane (C8)	Decane (C10)	Dodecane (C12)	Tetradecane (C14)
Octane (C8)				[16]					
Dodecane (C12)				[16]					
Tridecane (C13)				[17]					
Hexadecane (C16)				[16], [18]	[19]				
Heptadecane (C17)				[16]					
Octadecane (C18)				[19]	[20], [21]		[22]		
Nonadecane (C19)					[21]				
Eicosane (C20)				[19]	[22], [23]		[22]		
Docosane (C22)				[24]	[25]				
Tricosane (C23)					[26]				
Tetracosane (C24)	[13]	[13]	[13]	[18]	[27], [28], [25], [29]		[30], [31]	[27]	
Pentacosane (C25)					[26]				[32]
Hexacosane (C26)					[29], [26]				
Octacosane (C28)			[33]		[33], [29], [26]		[34], [30]	[34]	
Dotriacontane (C32)	[12]	[12]	[33]	[35],[16]	[36], [33], [20], [28]	[35]	[35], [30]	[35]	
Hexatriacontane (C36)				[34]	[33], [34], [28],[37]	[34]	[34], [30]	[34]	

4

Table 2

Data-set specific solubility best-fit parameters, A and B (Eq. (2)), obtained for solubilities below a mol-fraction of 0.1.

Solute/Solvent	Propane (C3)	Butane (C4)	Pentane (C5)	Hexane (C6)	Heptane (C7)	Octane (C8)	Decane (C10)	Dodecane (C12)	Tetradecane (C14)
Octane (C8)									
Dodecane (C12)				-1.062 0.0410					
Tridecane (C13)				0.993 0.0867					
Hexadecane (C16)				-1.100 0.1105					
Heptadecane (C17)				-1.025 0.1261					
Octadecane (C18)									
Nonadecane (C19)									
Eicosane (C20)				-3.340 0.100	-3.379 0.1061				
Docosane (C22)					-4.112 0.098				
Tricosane (C23)					-4.048 0.0847				
Tetracosane (C24)	-6.055 0.1294	-5.554 0.1267	-5.380 0.1232	-4.955 0.1062	-5.327 0.1165		-4.913 0.0805	-5.002 0.0959	
Pentacosane (C25)									
Hexacosane (C26)					-6.375 0.1318				
Octacosane (C28)			-7.345 0.1392		-7.233 0.1322		-7.329 0.1325	-7.459 0.1348	
Dotriacontane (C32)	-10.335 0.1704	-8.728 0.1447	-9.106 0.1539	-9.824 0.1715	-8.852 0.1441	-8.381 0.1346	-8.910 0.1450	-8.767 0.1398	
Hexatriacontane (C36)			-11.307 0.1731	-11.200 0.1699	-11.060 0.1619	-11.081 0.1677	-11.171 0.1693	-11.182 0.1654	

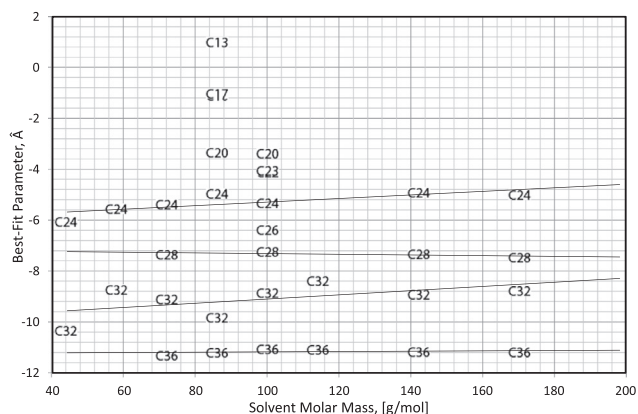
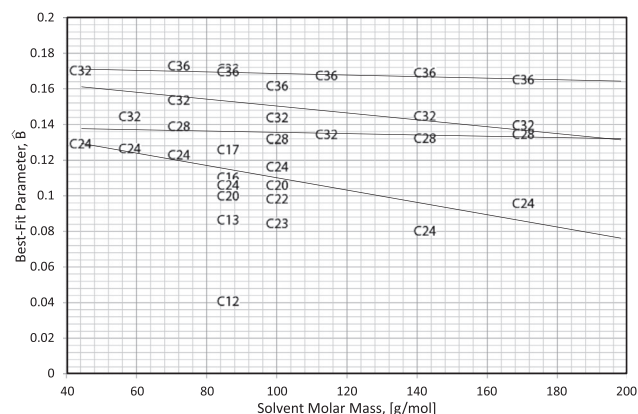
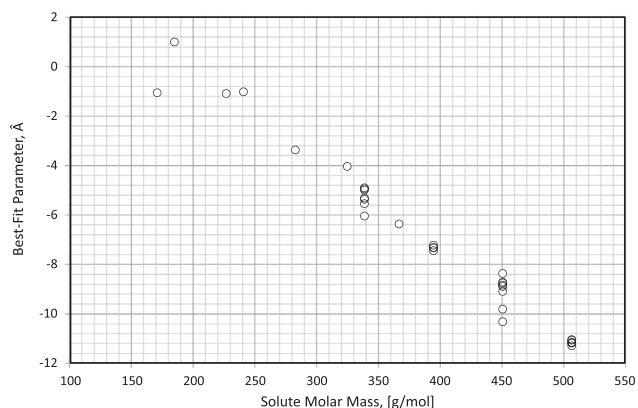
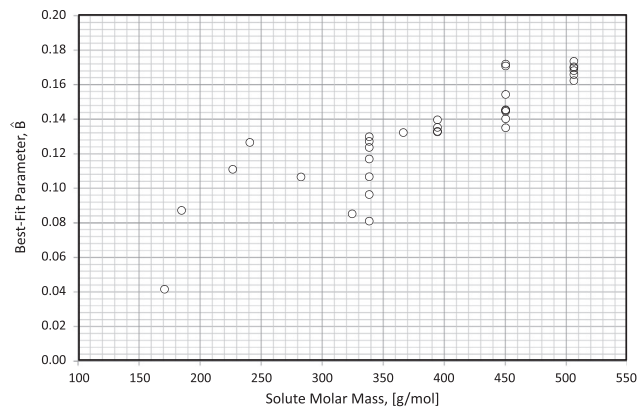
(a) \hat{A} vs. M_1 (a) \hat{B} vs. M_1 (b) \hat{A} vs. M_2 (b) \hat{B} vs. M_2

Fig. 2. Data-set specific best-fit parameter \hat{A} relation to the solvent and solute molar masses.

varies depending on the solvent molar mass. The melting temperature of n-alkanes ($^{\circ}\text{C}$) can be approximated by e.g. the Dollhopf correlation [42,43],

$$T_m(N) = \frac{414.6}{1 + 6.86/N} - 273.15, \quad (5)$$

where N is the number of carbon atoms in the alkane-chain. In Fig. 4, the complete set of reviewed solubility data are plotted against $(T - T_m)$, using melting temperatures obtained from Eq. (5). It can indeed be seen that most of the solubility curves gather in a narrow band. The C13 in C6 data-set (Morawski et al. [17]), identified by red circles in Fig. 4, stands out, however. It is noted that while Eq. (5) produces $T_{m,C13} = -1.76^{\circ}\text{C}$, Morawski et al. [17] reported a C13 melting temperature of -5.55°C . Using the latter improves the C13 in C6 solubility curve by shifting the $x_s = 1.0$ point to $(T - T_m) \approx 0$ (see black circles in Fig. 4), but the general impression is that the measured C13 in C6 solubilities might be unnaturally high. Fig. A.1c illustrates how the general correlation (Eq. (4)) generally under-predicts the C13 in C6 data.

It is observed that melting temperatures are prone to errors depending on the method of measurement and the purity of the substance; e.g. isomerization may affect the melting temperature significantly. Presuming that the melting temperature is chiefly depending on the molar mass, the solubility's dependency on the solute melting temperature is warranted by the molar mass dependency of the A and the B in Eq. (2). Thus, the current paper did not take the solute melting temperature as input for the developed correlation.

Fig. 3. Data-set specific best-fit parameter \hat{B} relation to the solvent and solute molar masses.

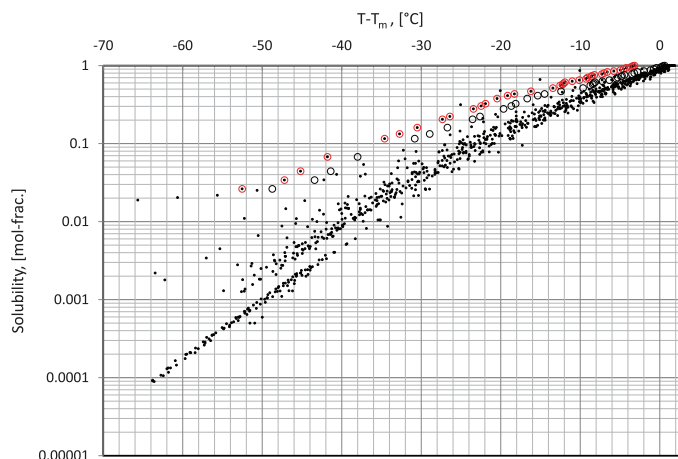


Fig. 4. Black dots: The complete set of experimental solubility data plotted against $(T - T_m)$, using Dollhopf et al. [42] melting temperatures (Eq. (5)). Red circles: The Morawski et al. [17] C13 in C6 data-set. Black circles: The Morawski et al. data-set using their own wax melting temperature instead of the Dollhopf et al. melting temperature.

Several of the referenced authors fail to state the purity or nature of the solutes and solvents used in their studies. It is suspected, however, that impurities only introduce minor errors in the experimental solubility data. Provost et al. [26] stated: "It is shown that the nature of the solvent has no major influence on the solubility...", and Rakotsoana et al. [32] concluded that the solu-

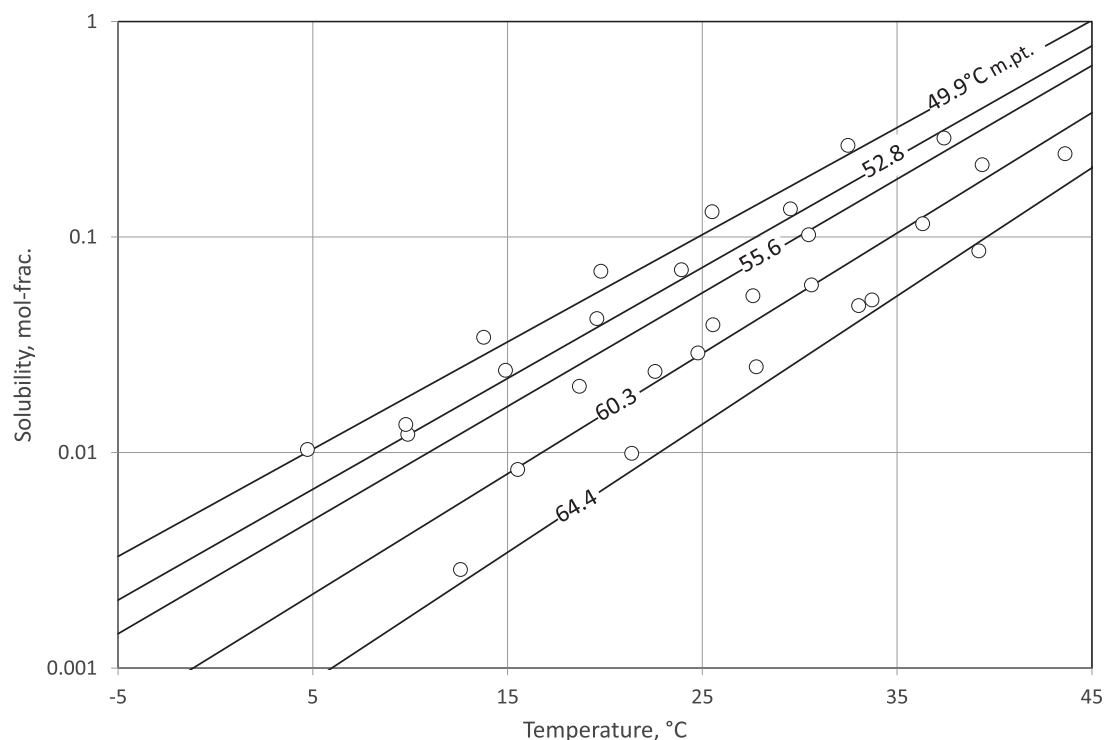


Fig. 5. Comparison of calculated solubilities (Eq. (4)) with experimental solubilities for petroleum waxes in a multi-component petroleum distillate (Solvent 1)[44]. Waxes are identified by their melting points, and molar masses, as input for Eq. (4), were taken from [44].

Table 3
Properties of solvent and waxes employed by Berne-Allen and Work [44].

		B.Pt. /M.Pt., [°C]	Molar mass, [g/mol]	Sp.grav.
Solvent	1	105	105	0.722
Wax	1	49.9	333	
	2	52.8	346	
	3	55.6	356	
	4	60.3	380	
	5	64.4	408	

bility of a multi-component wax is similar to that of the single-component wax whose carbon-number is equal to the mixture average carbon-number. These statements indicate that Eq. (4) may be utilized or adapted to more complex systems than were studied in this paper. This conjecture is supported by the good agreement between Eq. (4) and the experimental solubilities for paraffin waxes in petroleum distillates obtained by Berne-Allen and Work [44], as demonstrated in Fig. 5. Berne-Allen and Work did not elaborate on the purity of the waxes employed in their study, but specified: “the solvents were selected with the point in view of obtaining a wide spread over all the lighter fractions from petroleum.” The properties of the Berne-Allen-Work Solvent 1 and waxes 1-5 are cited in Table 3.

In the referenced literature, various methods of obtaining the solubilities were exercised. The two main strategies were 1) reducing the temperature and 2) heating the sample, looking for the first crystal to precipitate out or the last crystal to melt, respectively. Seyer and Fordyce [12], Sadeghzad et al. [22], Ashbaugh et al. [30], Johnsen [31] and Seyer [35] stated explicitly that the solubilities were found by observing the first crystals precipitate out. Dernini and De Santis [18] and Madsen and Boistelle [33], [34] did not clearly state what method they employed. The remaining authors established the saturation point by observing the last crystals dissolve. Dernini and De Santis [18], Seyer and Fordyce [12] and Seyer [35] stated that there was good agreement

between saturation temperatures obtained by heating and cooling (less than 0.1°C difference). Good agreement may not always be the case, however. Seyer [35] commented that there typically may be a significant difference between the dissolution and precipitation temperatures recorded for heavier alkanes. In the case of super-saturation, the solubility may be severely over-predicted. No obvious signs of super-saturation were identified in the experimental data, but it may be suspected that an effect of super-saturation is to reduce the data-set-specific solubility slope, B . It is observed, in Fig. A.1ap, that the Ashbaugh et al. [30] data ascends more slowly than the Madsen and Boistelle [34] data. Furthermore, the C24-C10 data-set, comprised from the Ashbaugh et al. [30] and Johnsen [31] data, produced one of the most severe outliers in Fig. 3b although there is good agreement between the two experiments. The C32-C10 data-set, comprised from the Ashbaugh et al. [30] and Seyer [35] data, did not give evidence of such an effect, however.

The current regression analysis was based on a log-linear relation between solubility and temperature, but the deviation from such behaviour increases for increasing solubility as seen in Figs. 1 and A.1. Therefore, an upper validity limit is required, for the regression analysis. The analysis was thus performed on the subset of the referenced experimental data, with reported solubilities below 0.1 (dilute solution). This upper solubility limit was chosen somewhat arbitrarily with the aim of including as much data as possible at the same time as getting as good a curve fit as possible. On one hand, for a higher limit, more data would be included in the regression analysis, but the increasing deviation from log-linear behaviour would potentially cause a generally less accurate curve-fit. On the other hand, for a lower limit, less data would be included, which might also result in a less accurate curve fit. Consequently, it was accepted that the resulting curve fit would be inaccurate above the upper solubility limit. It is observed, however, that for some data series, the general correlation (Eq. (4)) reproduces data above the the upper solubility limit accurately (see e.g. figs. A.1d(C16 in C6), A.1f-h(C18 in C6, C7, and C10), and A.1j(C20

in C6), in the appendix). Above the upper solubility limit, the tendency is that the general correlation under/over-predicts the solubility for solute carbon-numbers below/above 18, respectively. The possibility to find a more advanced correlation that will reproduce the experimental data more accurately, for a wider range of temperatures, has not been ruled out by the current study.

The current study used mol-fractions in presenting data and establishing the general correlation in Eq. (4). It is noted that the presentation would look somewhat different if e.g. mass-fractions were used, due to the non-linear relationship between mol- and mass-fractions and the large span in solute-solvent molar mass ratios considered. Although mol-fractions were preferred in the present study, the conclusions would be similar if mass-fractions were used instead.

Performing the best-fit procedure in the manner described in Section 5 means that all data-points are given the same weight. Furthermore, this means that the procedure gives more total weight to the data-sets with many data-points than to the data-sets with fewer data-points. Since a majority of the referenced data is obtained for the heaviest alkanes (C28-36) it is expected that the general correlation fit these data best. By utilizing a more complete specter of experimental data, the correlation will adapt to fit also the lighter alkanes better, possibly at the cost of the heavier alkanes fit. To give each data-set equal influence, multiple regression was performed to find the M_1 and M_2 dependency of A and B , respectively. The resulting correlation did not differ significantly from Eq. (4), but generally gave less accurate predictions of the experimental data.

7. Conclusions

A review of published solubility data for binary n-alkane mixtures is presented. Analysis of a total of 43 binary systems, from a total of 24 publications, revealed that there is a log-linear relationship between the solubility and the temperature, in the dilute range. Data-set-specific linear regression was performed to obtain data-set-specific best-fit parameters for the solubility-temperature data, for solubilities below a mol-fraction of 0.1, and it was seen that there is a clear linear relationship between the best-fit parameters and the solvent and solute molar masses. Linear regres-

sion was thus employed to establish a general correlation between the solubility and the solvent and solute molar masses and the temperature. Qualitative assessment shows that the developed correlation is successful at predicting the solubility trends seen in the experimental data, and reasonable predictions are obtained for the data-sets where no or only a few data-points exist below the mol-fraction of 0.1. Furthermore, evidence is given that the correlation provides predictive power for multi-component mixtures by utilizing average solute and solvent molar masses as input. More experiments are needed, however, for light solutes and a wider range of solvents in particular, to establish a more reliable correlation. The developed general correlation should be used with caution outside its validity range and for mixtures for which it has not been tested.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Sverre Gullikstad Johnsen: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Acknowledgements

This study was financed by SINTEF. The author is grateful for valuable feedback from colleagues and in particular Markus Hays Whitson (Whitson AS, Trondheim, Norway).

Appendix A. Detailed Comparison of Calculated and Experimental Data

In Fig. A.1, experimental solubility data-sets along with data-set-specific best-fit curves (for data below 0.1 solute mol-fraction) and the general correlation given in Eq. (4) are shown for each data-set.

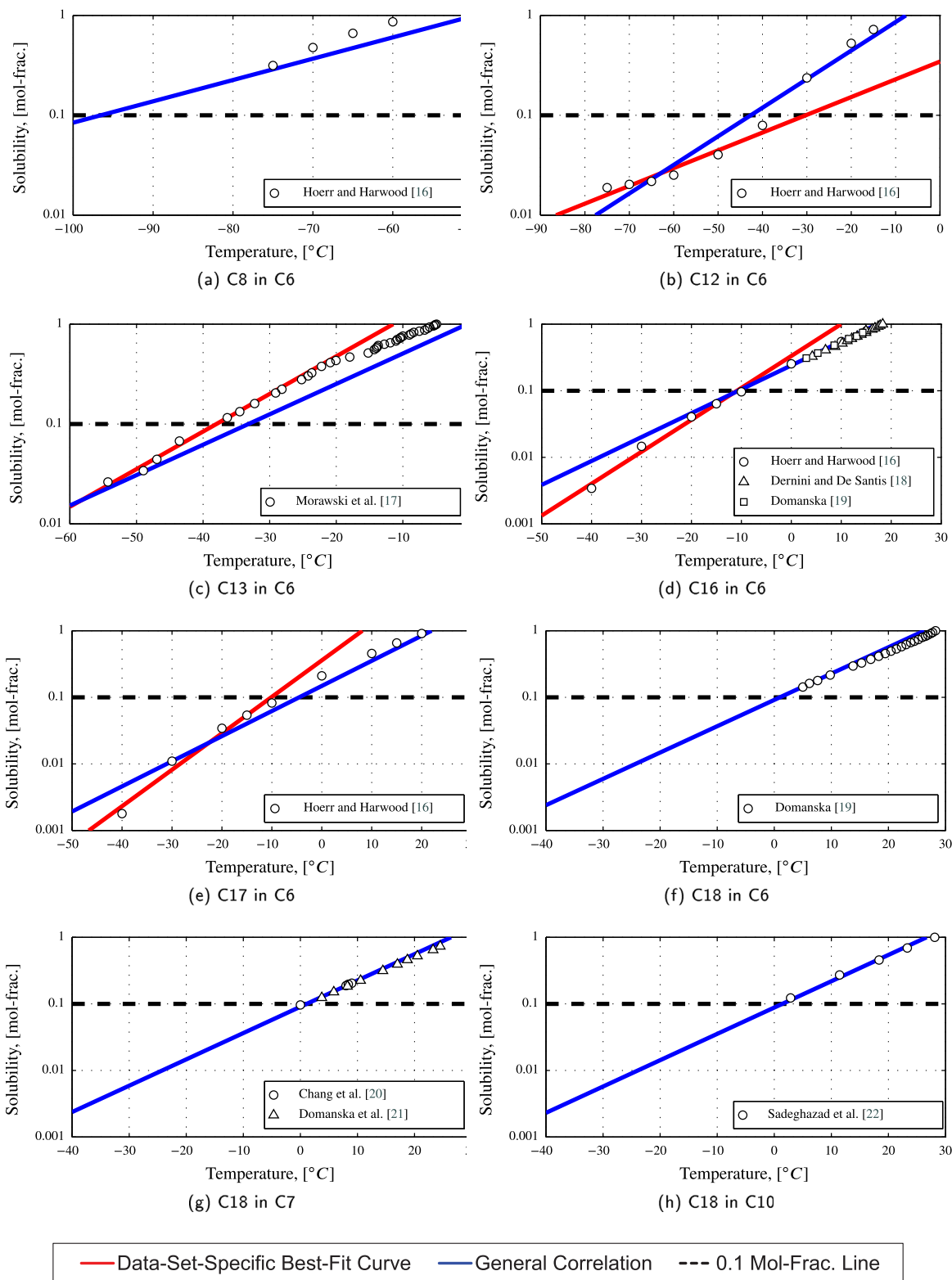


Fig. A.1. Experimental solubility data-sets along with data-set-specific best-fit curves (for data below 0.1 solute mol-fraction) and the general correlation given in Eq. (4).

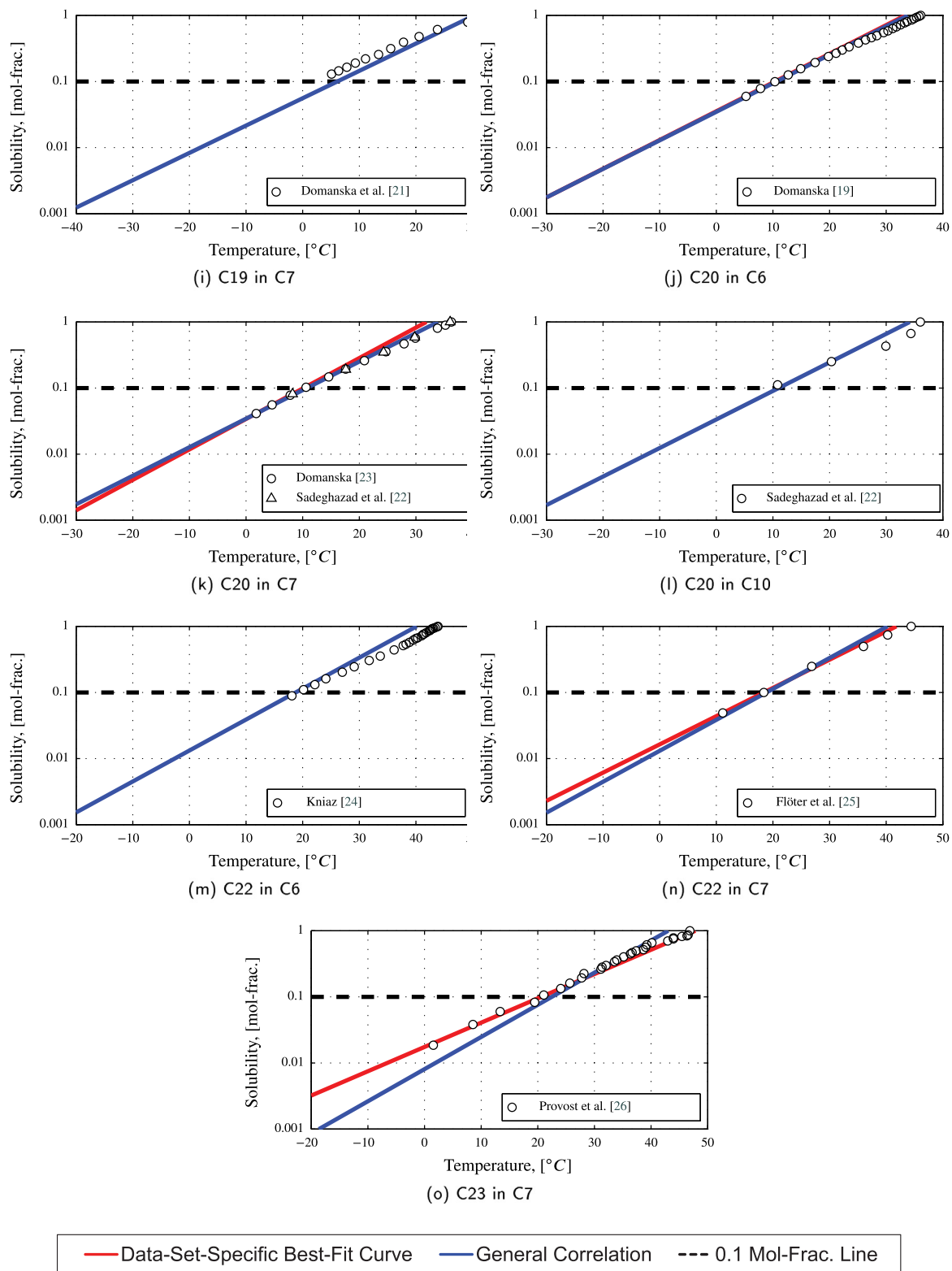


Fig. A.1. Continued

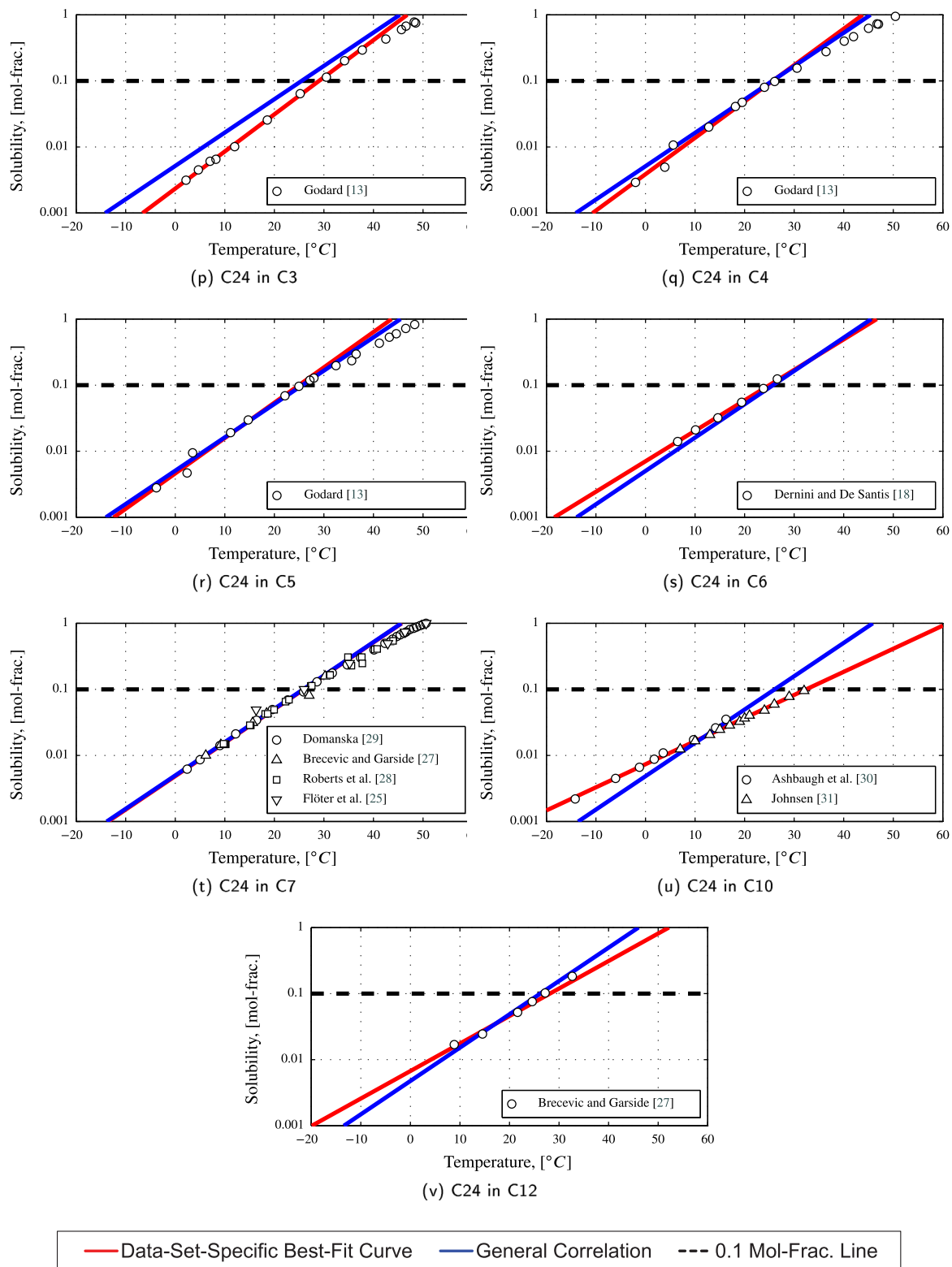


Fig. A.1. Continued

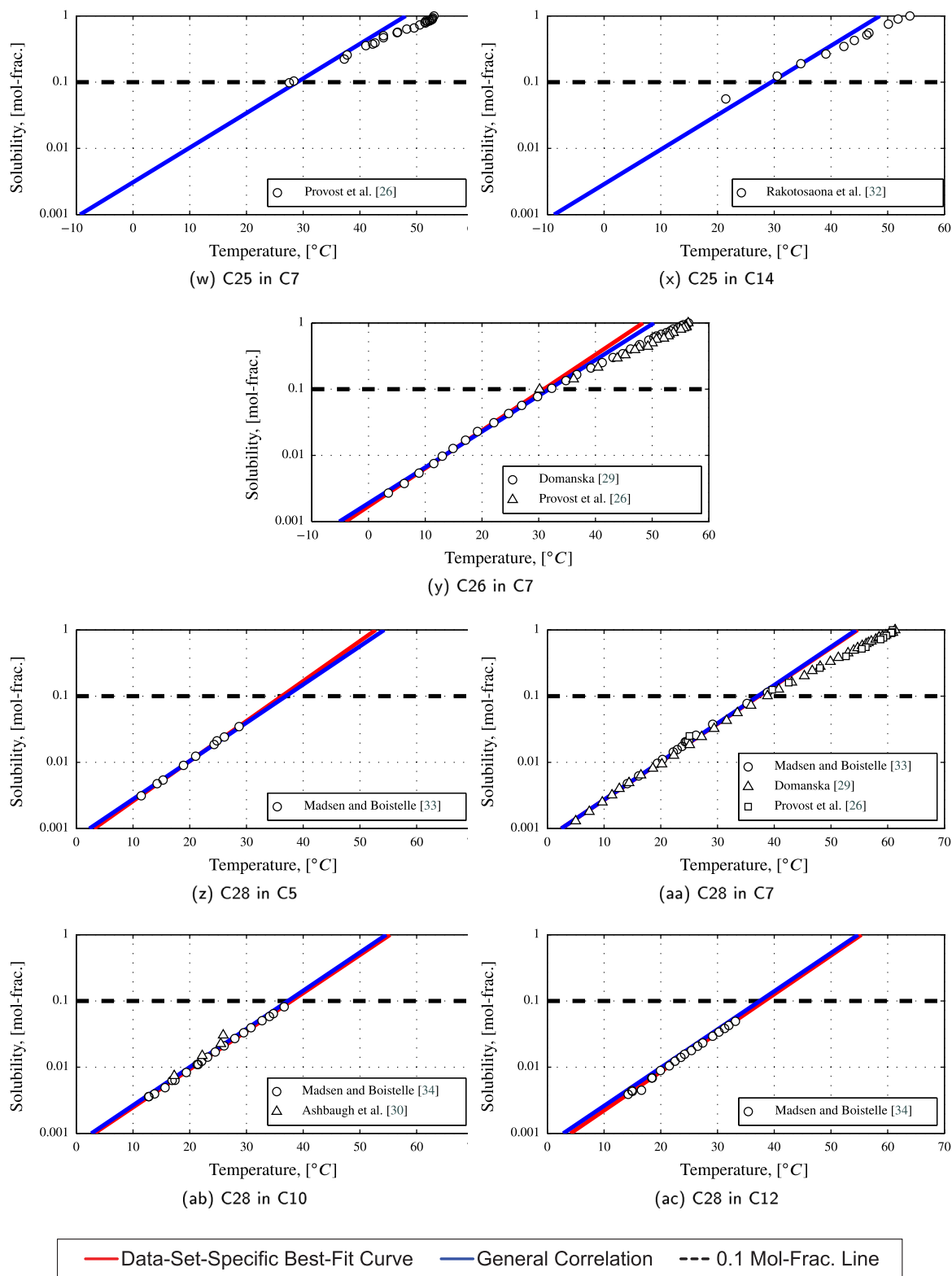


Fig. A.1. Continued

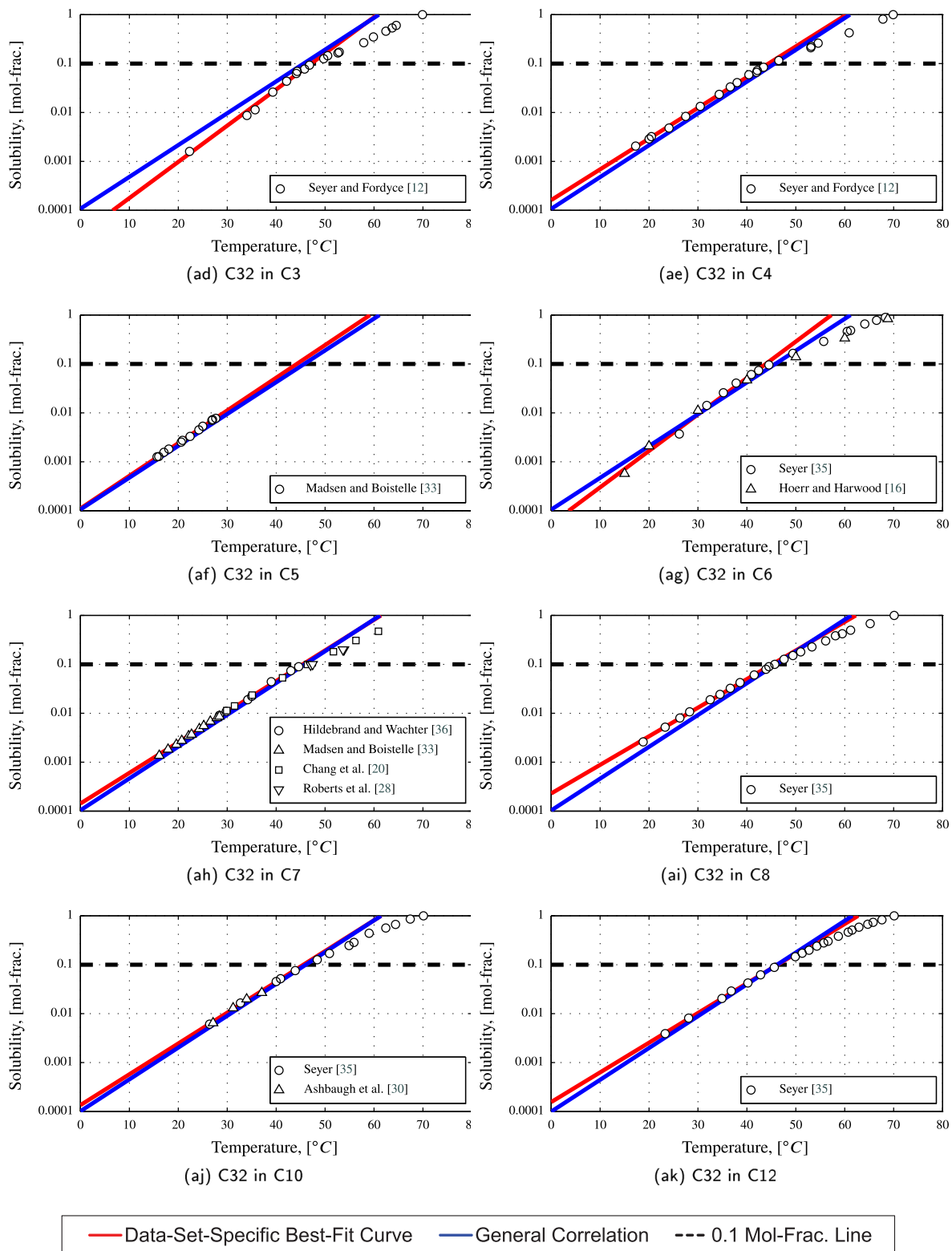


Fig. A.1. Continued

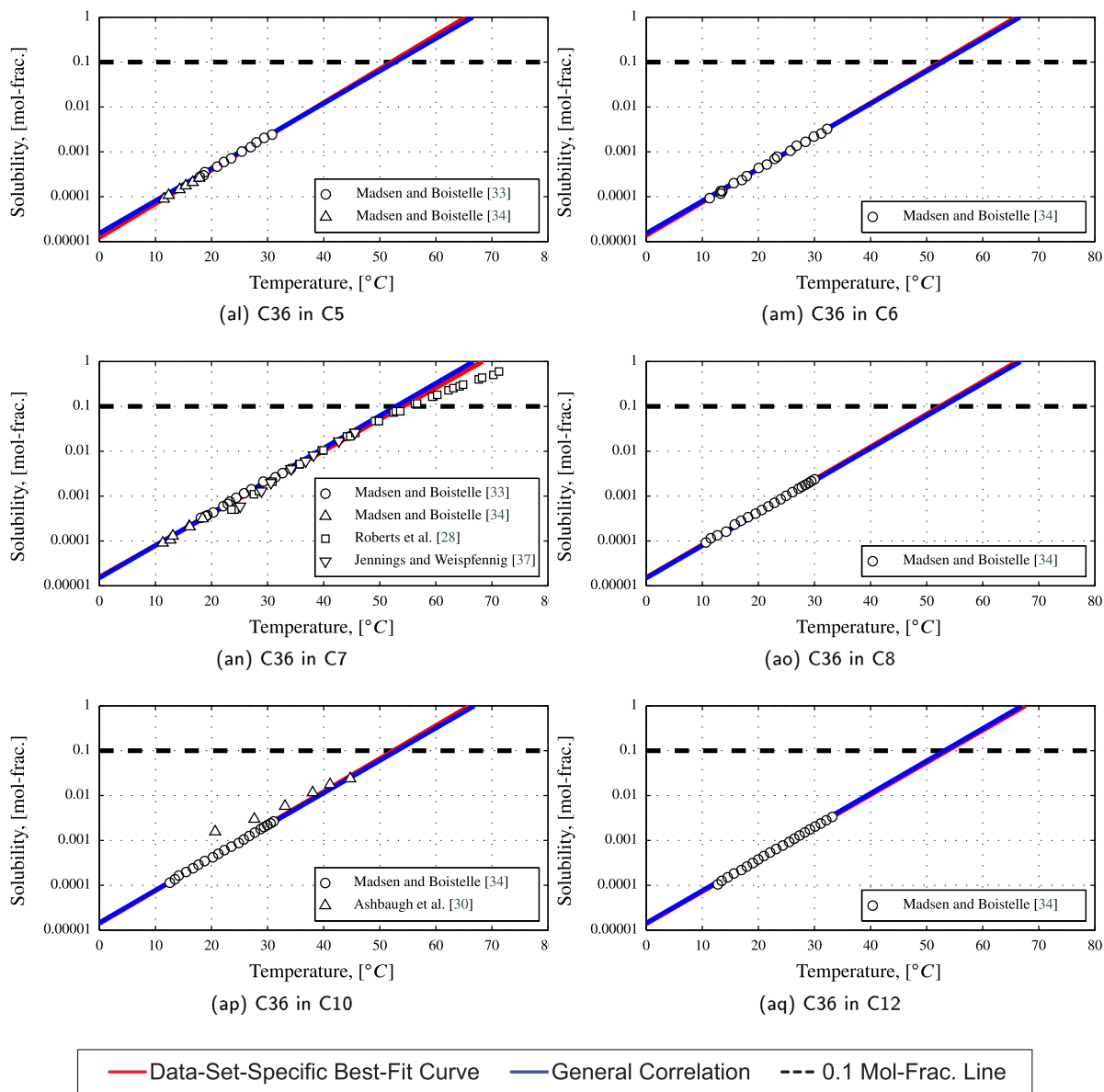


Fig. A.1. Continued

References

- [1] S. Misra, S. Baruah, K. Singh, Paraffin problems in crude oil production and transportation: A review, *SPE Production & Facilities* 10 (1) (1995), doi:10.2118/28181-PA.
- [2] J.C. Gluyas, J.R. Underhill, The staffa field, block 3/8b, UK north sea, *Geological Society, London, Memoirs* 20 (1) (2003) 327–333, doi:10.1144/GSL.MEM.2003.020.01.28.
- [3] J.A.P. Coutinho, B. Edmonds, T. Moorwood, R. Szczepanski, X. Zhang, Reliable wax predictions for flow assurance, *Energy & Fuels* 20 (3) (2006) 1081–1088, doi:10.1021/ef050082i.
- [4] H. Heidariyan, M. Ehsani, T.J. Behbahani, M. Mohammadi, Experimental investigation and thermodynamic modeling of wax precipitation in crude oil using the multi-solid model and PC-SAFT EOS, *Energy & Fuels* 33 (10) (2019) 9466–9479, doi:10.1021/acs.energyfuels.9b01445.
- [5] A. Shahdi, E. Panacharoensawad, et al., SP-Wax: Solid-liquid equilibrium thermodynamic modeling software for paraffinic systems, *Software X* (2019) 145–153, doi:10.1016/j.softx.2019.01.015.
- [6] M. Wang, C.C. Chen, Predicting wax appearance temperature and precipitation profile of normal alkane systems: An explicit co-crystal model, *Fluid Phase Equilibria* 509 (2020) 112466, doi:10.1016/j.fluid.2020.112466.
- [7] M. Ghasemi, C.H. Whitson, PVT modeling of complex heavy oil mixtures, *Journal of Petroleum Science and Engineering* 205 (2021) 108510, doi:10.1016/j.petrol.2021.108510.
- [8] P. Singh, R. Venkatesan, H.S. Fogler, N. Nagarajan, Formation and aging of incipient thin film wax-oil gels, *AIChE Journal* 46 (5) (2000) 1059–1074, doi:10.1002/aic.690460517.
- [9] K.G. Paso, H.S. Fogler, Bulk stabilization in wax deposition systems, *Energy and Fuels* 18 (4) (2004) 1005–1013, doi:10.1021/ef034105+.
- [10] C.H. Wu, K. Wang, P.J. Shuler, Y. Tang, J.L. Creek, R.M. Carlson, S. Cheung, Measurement of wax deposition in paraffin solutions, *AIChE Journal* 48 (9) (2002) 2107–2110, doi:10.1002/aic.690480923.
- [11] S.G. Johnsen, S.M. Hanetho, P. Tetlie, S.T. Johansen, M.A. Einarsrud, I. Kaus, S.C. Rone, Paraffin wax deposition studies on coated and non-coated steel surfaces, in: *Proceedings of Heat Exchanger Fouling and Cleaning IX - 2011*, 2011, doi:10.13140/2.1.5085.6968.
- [12] W.F. Seyer, R. Fordyce, The mutual solubilities of hydrocarbons. I. the freezing point curves of dotriacontane (dicetyl) in propane and butane, *Journal of the American Chemical Society* 58 (10) (1936) 2029–2031, doi:10.1021/ja01301a060.
- [13] H.P. Godard, *The solubility of tetracosane in propane, butane, and pentane*, University of British Columbia, 1937.
- [14] M. Campestrini, P. Stringari, Solubilities of solid n-alkanes in methane: Data analysis and models assessment, *AIChE Journal* 64 (6) (2018) 2219–2239, doi:10.1002/aic.16071.
- [15] N.G. Tassin, S.B. Rodríguez Reartes, M.S. Zabaloy, M. Cismondi, Modeling of solid-fluid equilibria of pure n-alkanes and binary methane + n-alkane systems through predictive correlations, *The Journal of Supercritical Fluids* 166 (2020) 105028, doi:10.1016/j.supflu.2020.105028.
- [16] C.W. Hoerr, H.J. Harwood, Solubilities of high molecular weight aliphatic compounds in n-hexane, *The Journal of Organic Chemistry* 16 (5) (1951) 779–791, doi:10.1021/jo01145a020.
- [17] P. Morawski, J.A.P. Coutinho, U. Domanska, High pressure (solid+liquid) equilibria of n-alkane mixtures: experimental results, correlation and prediction, *Fluid Phase Equilibria* 230 (2005) 72–80, doi:10.1016/j.fluid.2004.11.020.
- [18] S. Dernini, R. De Santis, Solubility of solid hexadecane and tetracosane in hexane, *The Canadian Journal of Chemical Engineering* 54 (4) (1976) 369–370, doi:10.1002/cjce.5450540421.
- [19] U. Domanska, *Selected Data on Mixtures: Series A*, in: *International Data Series, Thermodynamics Research Center*, 1990, pp. 83–93.
- [20] S.S. Chang, J.R. Maurey, W.J. Pummer, Solubilities of two n-alkanes in various solvents, *Journal of Chemical and Engineering Data* 28 (2) (1983) 187–189, doi:10.1021/je00032a017.
- [21] U. Domanska, T. Hofman, J. Rolinska, Solubility and vapour pressures in saturated solutions of high-molecular-weight hydrocarbons, *Fluid Phase Equilibria* 32 (3) (1987) 273–293, doi:10.1016/0378-3812(87)85059-8.
- [22] A. Sadeghzad, R.L. Christiansen, G.A. Sobhi, M. Edalat, The prediction of cloud point temperature: In wax deposition, *SPE*, 2000, doi:10.2118/64519-MS. Number SPE64519 in *SPE Asia Pacific Oil and Gas Conference and Exhibition*.
- [23] U. Domanska, *Selected Data on Mixtures: Series A*, *International Data Series, Thermodynamics Research Center*, 1987.
- [24] K. Kniaz, Solubility of n-docosane in n-hexane and cyclohexane, *Journal of Chemical and Engineering Data* 36 (4) (1991) 471–472, doi:10.1021/je00004a035.
- [25] E. Flöter, B. Hollanders, T.W. de Loos, J. de Swaan Arons, The ternary system (n-heptane + docosane + tetracosane): The solubility of mixtures of docosane and tetracosane in heptane and data on solid-liquid and solid-solid equilibria in the binary subsystem (docosane + tetracosane), *Journal of Chemical and Engineering Data* 42 (3) (1997) 562–565, doi:10.1021/je960332c.
- [26] E. Provost, V. Chevallier, M. Bouroukba, D. Petitjean, M. Dirand, Solubility of some n-alkanes (C23, C25, C26, C28) in heptane, methylcyclohexane, and toluene, *Journal of Chemical and Engineering Data* 43 (5) (1998) 745–749, doi:10.1021/je980027m.
- [27] L. Brecevic, J. Garside, Solubilities of tetracosane in hydrocarbon solvents, *Journal of Chemical and Engineering Data* 38 (4) (1993) 598–601, doi:10.1021/je00012a032.
- [28] K.L. Roberts, R.W. Rousseau, A.S. Teja, Solubility of long-chain n-alkanes in heptane between 280 and 350 K, *Journal of Chemical and Engineering Data* 39 (4) (1994) 793–795, doi:10.1021/je00016a035.
- [29] U. Domanska, *Selected Data on Mixtures: Series A*, in: *International Data Series, Thermodynamics Research Center*, 1987, pp. 271–285.
- [30] H.S. Ashbaugh, A. Radulescu, R.K. Prud'homme, D. Schwahn, D. Richter, L.J. Fetters, Interaction of paraffin wax gels with random crystalline/amorphous hydrocarbon copolymers, *Macromolecules* 35 (18) (2002) 7044–7053, doi:10.1021/ma0204047.
- [31] S.G. Johnsen, et al., The solubility of n-tetracosane (nC24) in n-heptane (nC7) and n-decane (nC10), *Fuel* 112 (2013) 594–595, doi:10.1016/j.fuel.2012.03.015.
- [32] R. Rakotosaona, M. Bouroukba, D. Petitjean, N. Hubert, J.C. Moïse, M. Dirand, Binary phase diagram (C14+C25) and isopleth (C14+wax): comparison of solubility of n-pentacosane and multiparaffinic wax in n-tetradecane, *Fuel* 83 (7–8) (2004) 851–857, doi:10.1016/S0016-2361(03)00249-7.
- [33] H.E.L. Madsen, R. Boistelle, Solubility of long-chain n-paraffins in pentane and heptane, *J. Chem. Soc., Faraday Trans. 1* 72 (1976) 1078–1081, doi:10.1039/F19767201078.
- [34] H.E.L. Madsen, R. Boistelle, Solubility of octacosane and hexatriacontane in different n-alkane solvents, *J. Chem. Soc., Faraday Trans. 1* 75 (1979) 1254–1258, doi:10.1039/F19797501254.
- [35] W.F. Seyer, Mutual solubilities of hydrocarbons. II. the freezing point curves of dotriacontane (dicetyl) in dodecane, decane, octane, hexane, cyclohexane and benzene, *Journal of the American Chemical Society* 60 (4) (1938) 827–830, doi:10.1021/ja01271a019.
- [36] J.H. Hildebrand, A. Wächter, The solubility of n-dotriacontane (dicetyl), *The Journal of Physical and Colloid Chemistry* 53 (6) (1949) 886–890, doi:10.1021/j150471a012.
- [37] D.W. Jennings, K. Weispenfennig, Experimental solubility data of various n-alkane waxes: effects of alkane chain length, alkane odd versus even carbon number structures, and solvent chemistry on solubility, *Fluid Phase Equilibria* 227 (2005) 27–35, doi:10.1016/j.fluid.2004.10.021.
- [38] R.F. Weimer, J.M. Prausnitz, Complex formation between carbon tetrachloride and aromatic hydrocarbons, *The Journal of Chemical Physics* 42 (10) (1965) 3643–3644, doi:10.1063/1.1695773.
- [39] P.B. Choi, E. Mclaughlin, Effect of a phase transition on the solubility of a solid, *AIChE Journal* 29 (1) (1983) 150–153, doi:10.1002/aic.690290121.
- [40] P.J. Huber, Robust estimation of a location parameter, *The Annals of Mathematical Statistics* 35 (1) (1964) 73–101, doi:10.1214/aoms/1177703732.
- [41] D. Birkes, Y. Dodge, *Alternative Methods of Regression*, Wiley-Interscience, 1993.
- [42] W. Dollhopf, H.P. Grossmann, U. Leute, Some thermodynamic quantities of n-alkanes as a function of chain length, *Colloid & Polymer Science* 259 (1981) 267–278, doi:10.1007/BF01381772.
- [43] H. Laux, T. Butz, G. Meyer, M. Matthäi, G. Hildebrand, Thermodynamic functions of the solid-liquid transition of hydrocarbon mixtures, *Petroleum Science and Technology* 17 (9–10) (1999) 897–913, doi:10.1080/10916469908949755.
- [44] A. Berne-Allen, L.T. Work, Solubility of refined paraffin waxes in petroleum fractions, *Industrial and Engineering Chemistry* 30 (7) (1938) 806–812, doi:10.1021/ie50343a019.