A COMPREHENSIVE EVALUATION OF EMPIRICAL CORRELATIONS FOR COMPUTING
THE SOLUBILITY OF CO_2 IN WATER

P. Sánchez Upton¹,² and E. Santoyo¹

¹Centro de Investigación en Energía, UNAM, Posgrado en Ingeniería (Energía), Temixco, Mor., México
²Gerencia de Proyectos Geotermoelectricos, CFE, Morelia, Mich., México
e-mail: pedro.sanchez@cfe.gob.mx | esg@mazatl.cie.unam.mx

ABSTRACT

A new correlation to estimate the solubility of carbon dioxide in water for the temperature interval of geothermal interest (from 25°C to 350°C) is proposed. This equation was derived from a comprehensive statistical analysis applied to an updated thermodynamical database containing experimental solubility data reported in the literature. In addition, it is also presented a useful equation to quantify the propagated error generated through the use of the new polynomial correlation developed. These equations will be implemented into a new geothermal wellbore simulator (under construction) for modelling the vertical-ascendant two-phase flow in geothermal wells.

INTRODUCTION

The characterization of geothermal fluids is fundamental to properly assess both reservoir and wellbore studies (Battistelli et al., 1997). Natural hydrothermal fluids are complex mixtures composed of water (H_2O), total dissolved solids (TDS), and non-condensible gases, NCG (Fournier, 1981; Henley et al., 1984; Santoyo et al., 1991; Nicholson, 1993; Sánchez-Upton, 2002). Because of H_2O is the main constituent of these mixtures, and also for simplicity, it is still a common practice to use thermodynamic and transport properties of pure water for predicting the behaviour of such fluids. However, if the mass fractions of TDS and/or NCG are significant, the pure water approach should not be applied.

The geothermal fluids are usually represented by mixtures containing the three major components found in any geothermal system, i.e., H_2O, carbon dioxide (CO_2) and sodium chloride (NaCl). The lack of experimental data dealing with the thermodynamic study of these mixtures have limited to have an equation of state (EOS) for a reliable prediction of their thermodynamic parameters, e.g., the solubility of CO_2 in H_2O (e.g., Pritchett et al., 1982; Battistelli et al., 1997). This is mainly due to the huge discrepancies that exist in the calculation of the Henry law constants (K_H) when different available equations are used.

The goal of this paper comes out, precisely, because of the significant discrepancies found during the calculation of the solubility of CO_2 in H_2O, which has motivated to proposed a new reliable equation for computing it and for estimating its associated error. This equation will be coupled to a geothermal wellbore simulator for studying the vertical-ascendant two-phase flow in geothermal wells.

PREVIOUS WORKS

The behaviour of the CO_2 solubility in H_2O for the temperature interval 0°-350°C, computed through the use of five empirical equations, is shown in Fig. 1. Most of these correlations were based on the experimental thermodynamic works carried out by Malinin (1959), Ellis and Golding (1963), Tödheide and Franck (1963), Takenouchi and Kennedy (1964, 1965), Zawisza and Malesińska (1981) and Zheng et al. (1997).

As can be observed, the majority of these correlations show an acceptable agreement in the estimation of the K_H at temperatures lower than 75 °C. However, at higher temperatures, the values of the CO_2 solubility show a remarked uncertainty at a given temperature. Malinin (1959) was apparently the first researcher that reported a comprehensive methodology to determine the solubility of CO_2 in H_2O, using his own experimental results. He used a modified version of the Henry’s law that includes the exponential correction factor (better known as poynting factor) for taking into account the effect due to high pressures, according to the following equation:

\[
\log \left( \frac{f_{CO_2} \cdot y_{CO_2}}{x_{CO_2}} \right) = \log (K_H) + \frac{V_{CO_2} (P - P_S)}{2.3026 RT_K} \tag{1}
\]
where $f_{\text{CO}_2}^0$ is the fugacity of gaseous CO$_2$ [MPa]; $y_{\text{CO}_2}$ and $x_{\text{CO}_2}$ are the molar fractions of CO$_2$ in the gaseous and liquid phases, respectively; $K_H$ is the Henry’s constant [MPa$^{-1}$]; $V_{\text{CO}_2}$ is the partial molar volume (PMV) of CO$_2$ at infinite dilution [cm$^3$/g mol]; $P$ is the total pressure [MPa]; $P_s$ is the saturation pressure of the solvent [MPa]; $R$ is the universal gas constant [8.3144 J/g mol K]; and $T_K$ is the absolute temperature [K].

All the previous experimental works have been used as a fundamental basis to generate several correlations to compute the solubility of CO$_2$. These equations have been mainly correlated as a function of temperature (e.g., Malinin, 1974; Pritchett et al., 1982; Fournier, 1989; Crovetto, 1991; Moya and Iglesias, 1992; Battistelli et al., 1997; McKibbin and McNabb, 1999). However, there are some other equations where the $K_H$ value is computed as a function of temperature and the water saturation pressure (e.g., Harvey, 1996: Fig. 1). From this brief analysis, it is clearly observed the inconsistency to predict the Henry’s constant in the temperature interval 75°C- 350°C using this group of five empirical equations (Fig. 1).

**NEW CORRELATION**

The new correlation was based on the thermodynamic principle that study the phase equilibrium as a fundamental basis to determine the solubility of CO$_2$ in H$_2$O. Regarding this, Prausnitz et al. (1999) suggest to employ an extended equation of the Henry’s law, which is given by:

$$\ln \left( \frac{f_{\text{CO}_2}^0}{x_{\text{CO}_2}} \right) = \ln (K_H) + \frac{\int_0^P V_{\text{CO}_2} dP}{RT_K}$$

(2)

where $f_{\text{CO}_2}^0$ is the CO$_2$ fugacity in H$_2$O.

It was also necessary to use the thermodynamic definition of $K_H$, which is given by:

$$K_H = \lim_{x_{\text{CO}_2} \to 0} \frac{f_{\text{CO}_2}^0}{x_{\text{CO}_2}}$$

(3)

The fugacity of CO$_2$ in H$_2$O was determined using the following equation:

$$f_{\text{CO}_2}^0 = \phi_{\text{CO}_2-H_2O} x_{\text{CO}_2} P$$

(4)

where $\phi_{\text{CO}_2-H_2O}$ is the fugacity coefficient of CO$_2$ in H$_2$O. The PMV of CO$_2$ in H$_2$O at infinite dilution was computed through the equation proposed by Moya e Iglesias (1992):
\[ \dot{V}_{CO_2} = C_1 + C_2 T_K + C_3 T_K^2 + C_4 T_K^3 + C_5 T_K^4 \]  

(5)

The constants of the Eq. (5) are compiled in Table 1. This equation was proposed to be applied in the temperature interval from 273.15 K (0°C) to 523.15 K (250°C), at any pressure. Although, this equation could be valid for temperatures up to 623.15 K (350°C), if the working pressure is lower or equal to 29.4 MPa. For temperatures ranging from 523.15 K to 623.15 K, and a pressure interval from 29.4 MPa to 98.1 MPa, the PMV should be corrected by multiplying Eq. (5) with a correction factor given by the following expression:

\[ \lambda_{PMV} = \alpha + \beta T_K + \gamma P + \delta T_K P \]  

(6)

Table 1. Coefficients of the PMV equation (5).

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_1</td>
<td>603.4837594</td>
</tr>
<tr>
<td>C_2</td>
<td>-5.813094063</td>
</tr>
<tr>
<td>C_3</td>
<td>2.229467211 x 10^{-2}</td>
</tr>
<tr>
<td>C_4</td>
<td>-3.828517025 x 10^{-5}</td>
</tr>
<tr>
<td>C_5</td>
<td>2.497055956 x 10^{-8}</td>
</tr>
</tbody>
</table>

The coefficients used in Eq. (6) are shown in Table 2.

Table 2. Coefficients of the Eq. (6).

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>0.18780892</td>
</tr>
<tr>
<td>(\beta)</td>
<td>1.5293077x10^{-3}</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>3.3965628x10^{-2}</td>
</tr>
<tr>
<td>(\delta)</td>
<td>-6.4434299x10^{-5}</td>
</tr>
</tbody>
</table>

The correction factor will be equal to 1.0 at 29.4 MPa, and it normally decreases as the pressure increases. These two equations (5 and 6) were derived from the results reported by Malinin (1974). The fugacity coefficient of CO_2 in H_2O can be computed through the use of a virial equation proposed by Spycher and Reed (1988). From these group of equations, the fundamental data required for estimating the Henry’s constant (\(K_H\)) include the molar fractions of CO_2 in both phases as well as the pressure and temperature data. The PMV of CO_2 at infinite dilution depends only on temperature from 0°C to 250°C. At higher temperatures (from 250°C to 350°C), PMV depends on temperature and pressure. According these assumptions, a poynting factor for these two temperature regions needs to be quantified. In the first case, after the integration of the poynting factor term and making the proper substitutions, the Eq. (2) adopts the following form:

\[
\ln \left( \frac{f_{CO_2-H_2O}}{x_{CO_2}} \right) = \ln (K_H) + \frac{\dot{V}_{CO_2}(T)(P - P_S)}{RT_K} 
\]

(7)

This expression is also known as the Krichevsky-Kasarnovsky equation. For the second temperature interval (T>250 °C), the integration of the poynting factor modifies the Eq. (2) in a more complex form, which is given by:

\[
\ln \left( \frac{f_{CO_2-H_2O}}{x_{CO_2}} \right) = \ln (K_H) + \frac{\dot{V}_{CO_2}(T)}{RT_K} \left[ (\alpha + \beta T_K)(P - P_S) + (\gamma + \delta T_K)(p^2 - p_S^2) \right] 
\]

(8)

Note that the second terms of the right of the Eqs. (7) and (8) are dimensionless. From Eq. (7), it follows that a graph of \( \ln \left( \frac{f_{CO_2-H_2O}y_{CO_2}P}{x_{CO_2}} \right) \) versus the total pressure \( P \) using the solubility data at constant temperature must lie on a straight line with a slope equal to \( \dot{V}_{CO_2}(T) \). The intercept of this straight line with the y-axis will define the value of the Henry’s constant (\(K_H\)) for the isotherm data (or temperature) considered, Fig. 2.

Figure 2. Estimation of the Henry’s constant using the thermodynamic method suggested by the equation (7).

Eq. (8) can be also analysed in the same form for inferring the values of the Henry’s constant (\(K_H\)). In this case, a plot of \( \ln \left( \frac{f_{CO_2-H_2O}y_{CO_2}P}{x_{CO_2}} \right) - \frac{\dot{V}_{CO_2}(T)(\gamma + \delta T_K)p^2}{2RT_K} \) versus the total pressure \( P \) using the solubility data at
temperatures between 250°C and 350°C will lie on a straight line. The intercept of this straight line with the ordinate axis will also give the value of the Henry’s constant.

RESULTS

A comprehensive thermodynamic database was created containing sixteen data groups of the CO₂ solubility at temperatures that range from 25 °C to 350 °C. These solubility data were mainly taken from the experimental works performed by Malinin (1959); Takenouchi and Kennedy (1964); Zawisza and Małeńska (1981); and Zheng et al. (1997). These data were subjected to a thorough statistical analysis in order to perform a weighted linear regression procedure of the data for computing the values of the Henry constants. For these purposes, all the solubility data points (at constant temperature) were initially plotted, as those shown in Fig. 2. Each data set was then submitted to a weighted linear regression (WLR) method according to the statistical theory suggested by Bevington (1969). The method considered the errors in both \( y \) and \( x \) variables for obtaining the best straight line.

An outlier detection algorithm was also coupled to the WLR method using the statistical theory on this subject proposed by Barnett and Lewis (1987). The application of this statistical methodology was required for determining, with accuracy, the value of the intercept (or Henry’s constant) for each solubility data set analysed at a given temperature. The slope of the straight line, as well as the errors of the intercept and the slope associated with the fitting statistical process were also computed. All these numerical analyses were performed using a Visual Fortran computer code, which was developed for these purposes. All the results obtained through this statistical study are summarised in Table 3.

In order to extend the temperature interval of the Henry’s constant from 0°C to 350°C, it was necessary to include the solubility data reported by Zheng et al. (1997). In this way, the Henry’s constant values were calculated with Eq. (7), using the experimental information presented by Zheng et al. (1997). The calculated values obtained from applying Eq. (7) were almost identical to those calculated by Zheng et al. (1997). This analysis confirmed the behavior of the solubility of CO₂ in H₂O for temperatures below 100 °C, where actually the available solubility data show a good agreement.

After concluding the calculation of the \( K_H \) values at each experimental temperature, it was possible to represent in a plot all the Henry’s constant values as a function of temperature (Fig. 3). A polynomial regression of third-degree applied to these data (\( K_H \) and temperature) enabled that they were appropriately fitted. To verify the accuracy of the regression, it was necessary to carry out an estimation of residuals with the polynomial function fitted (Bevington, 1969).

From the residual analysis, it was found that the data point at 250°C (Fig. 3) was far enough from the general tendency of the solubility curve. On the basis of this criterion, it was decided to reject it for improving the final fitted equation.

The polynomial model so obtained to describe the behavior of the Henry’s constant is finally represented by the following equation:

\[
\text{Ln} (K_H) = a + b T_c + c T_c^2 + d T_c^3
\]  
(9)
where $K_H$ is the Henry’s constant and $T_c$ is the fluid temperature (°C). The coefficients of this polynomial equation and their respective errors are reported in Table 4. The solubility curve that represents the Henry’s constant (MPa⁻¹) in the temperature interval 0°C-350°C is shown in Fig. 4. This curve was generated using the third-degree polynomial function given by Eq. (9).

Table 4. Coefficients of the Eq. (9) and their associated errors due to the polynomial regression.

<table>
<thead>
<tr>
<th>Polynomial regression coefficients</th>
<th>Errors of the coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=4.517428673$</td>
<td>$s_a=9.6930869 \times 10^{-2}$</td>
</tr>
<tr>
<td>$b=2.5554534 \times 10^{-2}$</td>
<td>$s_b=2.335664 \times 10^{-3}$</td>
</tr>
<tr>
<td>$c=-1.02213 \times 10^{-4}$</td>
<td>$s_c=1.5366 \times 10^{-5}$</td>
</tr>
<tr>
<td>$d=9.30689 \times 10^{-8}$</td>
<td>$s_d=2.88233 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

where $s_a, s_b, s_c, s_d, s_b, s_c, s_d, s_d,$ and $s_f$ are the coefficient errors due to the polynomial regression and the temperature measurement, respectively. This equation enables the associated errors with the calculation of the Henry’s constant values to be reliably computed.

**DISCUSSION OF RESULTS**

The values computed with the new improved Henry’s law equation derived here, differs from those calculated on the basis of the previous correlations. Three identified reasons can explain such differences. First, the earlier experimental works (Malinin, 1959 and 1974; Ellis and Golding, 1963; Takenouchi and Kennedy, 1964) used the fugacity coefficient for pure CO₂ (Majumdar and Roy, 1956), instead to consider the fugacity coefficient for the binary system CO₂-H₂O (Spycher and Reed, 1988). It has been demonstrated that these two fugacity coefficients can produce quite different values of the $K_H$ depending on the CO₂ concentration level, at a given pressure and temperature. Second, the PMV of CO₂ at infinite dilution was considered constant in the first studies, however, Malinin (1974) showed that it depends not only on the temperature but also on the experimental pressure. Third, none of the experimental works previously carried out have proposed a supplementary error equation for predicting the uncertainty associated with the calculation of the CO₂ solubility. Such an equation should be an important criterion for selecting the most suitable and accurate equation for estimating the CO₂ solubility for any subsequent application.

**ERROR PROPAGATION EQUATION**

The intrinsic error involved with the use of Eq. (9) can be estimated using the error propagation theory suggested by Bevington (1969) and Verma (2002). Thus, the overall error of the Eq. (9) can be computed by means of the following equation:

$$s_Z^2 = \left( \frac{\partial Z}{\partial a} \right)^2 s_a^2 + \left( \frac{\partial Z}{\partial b} \right)^2 s_b^2 + \left( \frac{\partial Z}{\partial c} \right)^2 s_c^2 + \left( \frac{\partial Z}{\partial d} \right)^2 s_d^2 + \left( \frac{\partial Z}{\partial T} \right)^2 s_T^2$$

(10)

where $Z$ is the dependent variable to be evaluated, that is, $Z=\ln(K_H)$. After reducing terms and solving the partial derivative terms of Eq. (10), a simplified equation was obtained:

$$s_Z^2 = s_a^2 + 2C^2 s_b^2 + 4C^4 s_c^2 + 6C^6 s_d^2 + 2C^2 s_b^2 + 6C^4 s_c^2 + 12C^6 s_d^2$$

(11)

**CONCLUSIONS**

A comprehensive thermodynamic/numerical methodology has been developed for computing the solubility of CO₂ in H₂O. From this methodology, an new improved correlation for estimating the Henry’s constant of the binary system H₂O-CO₂, as a function of temperature, has been proposed. This new correlation is based on a third-degree polynomial function, which has presented a low propagated error or uncertainty. The new solubility equation will be codified in the EOS subroutine for coupling it into the wellbore numerical simulator to be developed for modelling the vertical-ascendant two-phase flow in geothermal wells. Finally, it is strongly convenient to remark the necessity to perform more experimental works, specially at lower pressures, within the temperature interval of geothermal interest.

**ACKNOWLEDGEMENTS**

The authors would like to thank the financial support provided by the Engineering Post-graduate Programme of the UNAM (DGEP) and the Geothermal Department of CFE. Thanks are also due to Dr. Surendra P. Verma for his valuable critical comments on this investigation.
REFERENCES


