

Sensitivity analysis, calibration and uncertainty quantification for heavy oil viscosity estimates with a corresponding states model

Diego T. Volpato,^{*,†,‡} Lucas V. A. Oliveira,[†] Sofia P. Bittencourt,[†] Danilo M. D. D. da Silva,[†] Edson T. M. Manoel,[†] and Vinicius Girardi[†]

[†]*Software Development section, Engineering Simulation and Scientific Software (ESSS), St. Orlando Phillippi, 100 - 1^o Floor, Florianópolis, Brazil.*

[‡]*Graduate Program, LNCC/MCTI, St. Getúlio Vargas, 333, Petrópolis/RJ, Brazil.*

E-mail: diego@esss.co

Abstract

The main goal of this work is to assess heavy oil viscosity estimates by a Corresponding States Principle (CSP) model using a Bayesian approach in an efficient way. To determine and select relevant parameters for model calibration, an enhanced Elementary Effects method is used to evaluate sensitivity measures of CSP tuning parameters. With the combination of sensitivity analysis and Bayesian calibration, a unified procedure to automatically tune CSP viscosity model while reducing the number of tuning parameters is devised. Moreover, the Bayesian approach provides additional information on CSP model uncertainties and credible regions inherited from experimental data. To evaluate such uncertainties in CSP viscosity model, it was used five heavy oil samples available in the literature. The viscosity curves constructed by 50th-percentile from Monte Carlo realizations for the CSP calibration show good agreement when compared with classical Least-Squares regression (deterministic), demonstrating the potential of the sensitivity assessment for both Bayesian and deterministic approaches. However, when Bayesian calibration is used, limitations of CSP viscosity estimates are detected through violation of credible regions, suggesting that heavy oil viscosity estimates for rel-

atively low pressure conditions can be insufficiently accurate for the CSP model considered in this study.

Introduction

Viscosity experiments are one of the most commonly reported results from PVT laboratories due to its importance to transport calculations in engineering. Viscosity estimation for oil samples is an important stage to properly determine a complete set of thermodynamic and transport properties for oil samples. In general, such experimental data are provided in a dedicated viscosity data report^{1,2}. However, mixture viscosity model fitting is a challenging task^{1,3}, usually performed with high associated errors when compared with experimental data. Given the importance of viscosity as a transport property, improving viscosity estimates is a relevant topic for PVT analysis.

Several models for viscosity have been proposed in the literature. Among them, the most prominent are Lohrenz-Bray-Clark (LBC)⁴ and the models there are based on Corresponding States Principle (CSP), such the one proposed by Pedersen *et al.*. Another corresponding state principle model is SUPERTRAPP, provided by NIST (US National Institute of Standards and Technology), which is actually a computer pro-

gram based on the method proposed by [Ely and Hanley](#). While LBC correlation is known for its simplicity, CSP stands out for its predictive capacity, being one of the most adopted models in the industry. Despite this fact, an additional tuning step is still necessary from experimental data to obtain an optimized CSP viscosity model. Even in reservoir simulations, where little impact caused by fluid properties are generally assumed to be true, improper viscosity estimates can lead to large deviations⁶.

Considering the impact of viscosity estimation on the fluid model, our goal is to propose an approach for tuning some parameters that can improve the outcomes of a CSP based method. Besides that, we also demonstrate a scheme for performing sensitivity analysis and provide uncertainties information for the tuned parameters. In particular, we are interested in a even more difficult problem, which is the estimation of heavy oils viscosities. Due little information about the heavier fractions during the compositional characterization procedure, this kind of mixture has intrinsic associated uncertainties. More specifically, we provide a reliable method for tuning the CSP model proposed by [Lindeloff et al.](#), which is an adaptation of the [Pedersen et al.](#) method to this kind of fluids. To calibrate the model, a Bayesian approach is employed using a Transitional Markov Chain Monte Carlo method called Cascading Adaptive Transition Metropolis in Parallel⁸ (CAT-MIP), which results in probability distributions for tuning parameter as output, in contrast with classical fitting methods based on deterministic continuous optimization problems (e.g., Least-Squares regressions), where only parameter values are known after the calibration procedure. The probability distributions provide information about the effect of uncertainties incorporated from data on the model predictive capabilities^{9,10}. Moreover, the model outcomes are given as credible regions, which can shed light on conditions where predictions are poor or inappropriate. Since that the Bayesian approach is computationally intensive, we use an enhanced Elementary Effects method¹¹ to evaluate overall sensitivity measures for CSP tuning parameters and determine the most influ-

ential parameters to be tuned, resulting in an efficient Bayesian calibration using Sensitivity Analysis information. Such Bayesian calibration approach, as far as the authors are aware, is unreported in the CSP viscosity model literature.

The manuscript is outlined as follows: we begin with a Materials and Methods section, in which we describe experimental data and the methods we used in this work. In the next section, we show the Results. After that, a Discussion section is presented. At last, we draw some Conclusions.

Materials and Methods

In the following section, we describe the experimental data and methods used to model the viscosity of heavy oil with estimates of associated uncertainties. The section is divided as:

- Experimental data: viscosity measurement data are described alongside data source. A brief description of which techniques for heavy oil characterization is provided;
- A corresponding states model for viscosity estimation;
- The Elementary Effects method^{11,12}: a screening sensitivity analysis method to determine which fitting parameters are relevant for tuning and uncertainty quantification;
- Bayesian CSP model calibration and uncertainty quantification;
- A unified approach for heavy oil viscosity estimation: a strategy considering sensitivity analysis and Bayesian CSP model calibration.

Experimental data

Five oil samples were considered in order to demonstrate the proposed approach. Information such as oil compositions and viscosity measurements were gathered from Table 1 and 2 in

a paper by Lindeloff et al. To be consistent with this previous work, the same oil labels will be kept, but only samples 1, 2, 5, 7 and 8 are considered in the present study. We summarize oil compositions and viscosity data in Table 1 and 2, respectively. All samples have predominant molar fractions of C_7^+ , characterizing them as heavy oils.

Table 1: Oil samples overall compositions⁷ given as %mol. Molar mass is provided as g/mol and density of C_7^+ as g/cm³.

Component	Oil 1	Oil 2	Oil 5	Oil 7	Oil 8
N ₂	0.900	0.310	0.040	0.340	0.310
CO ₂	0.140	0.080	1.210	1.480	0.060
C ₁	38.780	19.430	18.920	25.610	14.990
C ₂	2.030	1.470	0.040	0.040	1.130
C ₃	0.060	0.350	0.040	0.020	0.280
i-C ₄	0.010	0.610	0.030	0.020	0.590
n-C ₄	0.050	0.290	0.050	0.030	0.150
i-C ₅	0.000	0.450	0.050	0.030	0.330
n-C ₅	0.000	0.260	0.050	0.010	0.090
C ₆	0.040	0.900	0.230	0.330	0.480
C ₇ ⁺	57.990	75.860	79.340	72.090	81.590
Molar mass C ₇ ⁺	296.000	337.500	530.200	533.600	343.000
Density of C ₇ ⁺	0.955	0.945	1.009	1.001	0.948

In order to apply the methods to estimate viscosities, a proper fluid characterization and lumping procedure is required. Due experimental limitations, the heavier components characterization procedure is necessary to circumvent the lack of information about properties of components heavier than heptane fractions, i.e., critical and saturation properties. Besides that, lumping is particularly mandatory to reduce the number of pseudo-fractions (up to C_{200} with the adopted method for heavy oils) into fewer representative pseudo-components to avoid performance and phase equilibria convergence issues. We closely follow Pedersen et al. and related original research articles^{13–17} to estimate pseudo-components properties and to perform lumping on mixture components.

A corresponding states model for viscosity estimates

The viscosity model based on corresponding states used in this work was originally published by Pedersen et al. with further improvements

published by Pedersen and Fredenslund. Adaptations to include better estimates for heavy oils were originally published by Lindeloff et al.. The corresponding states principle was firstly introduced in the thesis of Johannes Diderik van der Waals in 1873 and this subject was published later by him again in another article¹⁹. In the scope of viscosity calculations, it can be used to provide a scale-factor that allows the calculation of a multi-component mixture viscosity from an accurate correlation for a single component fluid. This is done by assuming that the reduced viscosities of any two fluids at the same reduced must be equal^{3,7}. Thus,

$$\eta_r^o(P_r, T_r) = \eta_r^{\text{mix}}(P_r, T_r) \quad (1)$$

where the symbols η_r^o and η_r^{mix} denote, respectively, the reference fluid and mixture viscosities. Because of the limitation of viscosity data at the critical point (η_c), the following correlation is used^{3,7,18}:

$$\eta_r = \frac{\eta}{\eta_c} = \frac{\eta}{T_c^{-1/6} P_c^{2/3} M^{1/2}} \quad (2)$$

where P_c and T_c are the critical pressure and temperature and M is the molar mass.

By combining these assumptions, the resulting expression for the viscosity of the mixture can be estimated through

$$\begin{aligned} \eta_{\text{mix}}(P, T) = & \left(\frac{T_{c,\text{mix}}}{T_{co}} \right)^{-1/6} \\ & \times \left(\frac{P_{c,\text{mix}}}{P_{co}} \right)^{2/3} \\ & \times \left(\frac{M_{\text{mix}}}{M_o} \right)^{1/2} \\ & \times \frac{\alpha_{\text{mix}}}{\alpha_o} \eta_o(P_o, T_o) \end{aligned} \quad (3)$$

where P_o and T_o are the reference pressure and temperature given by

$$P_o = P \frac{P_{co}}{P_{c,\text{mix}}} \frac{\alpha_o}{\alpha_{\text{mix}}} \quad (4)$$

$$T_o = T \frac{T_{co}}{T_{c,\text{mix}}} \frac{\alpha_o}{\alpha_{\text{mix}}} \quad (5)$$

and the parameter α is a correction factor given

Table 2: Viscosity data for each oil sample⁷.

Oil 1 ($T = 55$ °C)		Oil 2 ($T = 77$ °C)		Oil 5 ($T = 49$ °C)		Oil 7 ($T = 49$ °C)		Oil 8 ($T = 77$ °C)	
P (bar)	Viscosity (cP)	P (bar)	Viscosity (cP)	P (bar)	Viscosity (cP)	P (bar)	Viscosity (cP)	P (bar)	Viscosity (cP)
345.7	8.1	200	11.1	137.9	2268	207.9	2597	213	18.2
311.3	7.7	170	10.4	110.3	2085	173.4	2405	176	17.2
276.8	7.4	140	9.8	82.7	1898	138.9	2219	173.5	17.1
242.3	7.0	110	9.3	55.2	1735	104.4	2042	143	16.2
231.0	6.9	70	8.9	41.4	1760	90.6	1975	101	15
221.6	6.8	-	-	27.6	2100	83.7	1943	62	13.6
214.7	6.7	-	-	13.8	2404	76.9	1912	52	13.3
207.9	6.6	-	-	-	-	70	1882	47	13.3
202.6	6.5	-	-	-	-	63.1	1857	35.2	13.5
173.4	6.9	-	-	-	-	42.4	2723	25	14.4
138.9	7.5	-	-	-	-	28.6	3406	12.2	16
104.4	8.2	-	-	-	-	14.8	4250	1	19.1
70.0	9.0	-	-	-	-	7.9	4803	-	-
35.5	10.6	-	-	-	-	1	5676	-	-
9.8	13.4	-	-	-	-	-	-	-	-
1.0	22.7	-	-	-	-	-	-	-	-

by¹⁸

$$\alpha = 1 + 7.378 \times 10^{-3} \rho_r^{1.847} M^{0.5173} \quad (6)$$

with

$$\rho_r = \frac{\rho_o \left(P \frac{P_{co}}{P_{c,mix}}, T \frac{T_{co}}{T_{c,mix}} \right)}{\rho_{co}}. \quad (7)$$

In this work, methane is used as the reference component, except at conditions where methane assumes the solid form.

The mixing rules applied to evaluate the critical properties are from Murad and Gubbins²⁰. The mixture critical temperature is computed with

$$T_{c,mix} = \frac{\sum_{i=1}^{N_c} \sum_{j=1}^{N_c} z_i z_j \tilde{V}_{cij} \sqrt{T_{ci} T_{cj}}}{\sum_{i=1}^{N_c} \sum_{j=1}^{N_c} z_i z_j \tilde{V}_{cij}} \quad (8)$$

and the mixture critical pressure with

$$P_{c,mix} = \frac{8 \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} z_i z_j \tilde{V}_{cij} \sqrt{T_{ci} T_{cj}}}{\left(\sum_{i=1}^{N_c} \sum_{j=1}^{N_c} z_i z_j \tilde{V}_{cij} \right)^2} \quad (9)$$

where

$$\tilde{V}_{cij} = \left[\left(\frac{T_{ci}}{P_{ci}} \right)^{1/3} + \left(\frac{T_{cj}}{P_{cj}} \right)^{1/3} \right]^3 \quad (10)$$

The mixture molecular weight is evaluated

through

$$\begin{aligned} M_{mix} &= 1.304 \times 10^{-4} \times (1stCSP) \\ &\times \left(\overline{M}_w^{2.303(2ndCSP)} - \overline{M}_n^{2.303(2ndCSP)} \right) \\ &+ \overline{M}_n \end{aligned} \quad (11)$$

where $(1stCSP)$ and $(2ndCSP)$ are tuning coefficients, which are equal to 1 by default. Also, \overline{M}_w and \overline{M}_n are the weight average and number average molecular weights, respectively:

$$\overline{M}_w = \frac{\sum_{i=1}^{N_c} z_i M_i^2}{\sum_{i=1}^{N_c} z_i M_i} \quad (12)$$

$$\overline{M}_n = \sum_{i=1}^{N_c} z_i M_i \quad (13)$$

The methane density, ρ_o , is calculated using the mBWR (modified Benedict-Webb-Rubin) equation of state²¹. Furthermore, the reference methane viscosity is obtained with the model suggested by Hanley and modified by Pedersen and Fredenslund as following:

$$\eta'(\rho, T) = \eta_0(T) + \eta_1(T)\rho + \Delta\eta(\rho, T) \quad (14)$$

where ρ is the methane density in mol/L and

$$\Delta\eta(\rho, T) = F_1 \Delta\eta'(\rho, T) + F_2 \Delta\eta''(\rho, T) \quad (15)$$

Parameters F_1 and F_2 are weights that de-

fine a smooth transition between functions $\Delta\eta'(\rho, T)$ and $\Delta\eta''(\rho, T)$. For a given temperature T in kelvin, the weighting parameters are calculated as

$$F_1 = \frac{\tanh(s\Delta T) + 1}{2} \quad (16)$$

$$F_2 = \frac{1 - \tanh(s\Delta T)}{2} \quad (17)$$

$$\Delta T = T - T_F \quad (18)$$

where T_F is the freezing point of methane and s ($= 1/5$ as default) is a smoothing factor. Due to the smooth transition, continuity is preserved and when $\Delta T \ll 0$, $\Delta\eta(\rho, T) \approx \Delta\eta''(\rho, T)$. Likewise, when $\Delta T \gg 0$, $\Delta\eta(\rho, T) \approx \Delta\eta'(\rho, T)$. When ΔT is close to zero, a linear combination $\Delta\eta(\rho, T) = F_1\Delta\eta'(\rho, T) + F_2\Delta\eta''(\rho, T)$ is used.

The term $\eta_0(T)$ represents the dilute gas viscosity, which is given as:

$$\begin{aligned} \eta_0(T) = & \frac{GV_1}{T} + \frac{GV_2}{T^{2/3}} + \frac{GV_3}{T^{1/3}} \\ & + GV_4 + GV_5T^{1/3} + GV_6T^{2/3} \\ & + GV_7T + GV_8T^{4/3} + GV_9T^{5/3} \end{aligned} \quad (19)$$

with constants GV_1 to GV_9 listed in Table 3.

Table 3: Constants for equation 19.

GV_1	$2.090\,975 \times 10^5$
GV_2	$2.647\,269 \times 10^5$
GV_3	$-1.472\,818 \times 10^5$
GV_4	$4.716\,74 \times 10^4$
GV_5	$-9.491\,872 \times 10^3$
GV_6	$1.219\,979 \times 10^3$
GV_7	$-9.627\,993 \times 10^1$
GV_8	$4.274\,152$
GV_9	$-8.141\,531 \times 10^{-2}$

The term $\eta_1(T)$ is defined by the following empirical expression:

$$\eta_1(T) = A + B \left(C - \ln \frac{T}{F} \right)^2 \quad (20)$$

in which $A = 1.696985927$, $B = -0.133372346$, $C = 1.4$ and $F = 168.0$.

The term $\Delta\eta'$, which plays an important role

in the dense liquid region, is computed with

$$\begin{aligned} \Delta\eta'(\rho, T) = & \exp \left(j_1 + \frac{j_4}{T} \right) \\ & \times \left[\exp \left[\rho^{0.1} \left(j_2 + \frac{j_3}{T^{3/2}} \right) \right. \right. \\ & \left. \left. + \theta \rho^{1/2} \left(j_5 + \frac{j_6}{T} + \frac{j_7}{T^2} \right) \right] - 1.0 \right] \end{aligned} \quad (21)$$

where the parameter θ is denoted by

$$\theta = \frac{\rho - \rho_c}{\rho_c} \quad (22)$$

and the constants j_1 to j_7 , for viscosities in μP , are given in Table 4.

Table 4: Constants for equation 21.

j_1	10.3506
j_2	17.5716
j_3	-3019.39
j_4	188.730
j_5	0.0429036
j_6	145.290
j_7	6127.68

The term $\Delta\eta''$ is a modification¹⁸ to include the regions where methane is solid at its reference conditions, i.e., when temperature is below 91 K. In this way, at reference temperatures below the freezing point of methane, this term is the one that mostly influences the viscosity and is calculated as

$$\begin{aligned} \Delta\eta''(\rho, T) = & \exp \left(k_1 + \frac{k_4}{T} \right) \\ & \times \left[\exp \left[\rho^{0.1} \left(k_2 + \frac{k_3}{T^{3/2}} \right) \right. \right. \\ & \left. \left. + \theta \rho^{1/2} \left(k_5 + \frac{k_6}{T} + \frac{k_7}{T^2} \right) \right] - 1.0 \right] \end{aligned} \quad (23)$$

where the constants k_1 to k_7 , for viscosities in μP are given in Table 5.

Table 5: Constants for equation 23.

k_1	-9.74602
k_2	18.0834
k_3	-4126.66
k_4	44.6055
k_5	0.9676544
k_6	81.8134
k_7	15649.9

In order to improve heavy oil viscosity estimates, Lindeloff et al. proposed to use – instead of methane viscosity as reference at some T_o values – the viscosity of stable crude oils for the corresponding state, which is given by

$$\log_{10} \eta'_{oil} = -0.07995 - 0.01101M - \frac{371.8}{T} + \frac{6.215M}{T} \quad (24)$$

where the viscosity is evaluated in cP.

In the equation above, when temperature is above 564.49 K, the sign of -0.01101 is changed from negative to positive. We remark that, in equation 24, the real mixture temperature T is used. Also, M is the average molecular weight, which is calculated to make the equation 24 suitable to live oils. Thus, we have that

$$M = \overline{M}_n \left(\frac{1.5}{Viscfac_3 \times (3rdCSP)} \right)^{Viscfac_4 \times (4thCSP)} \quad (25)$$

for $\left(\frac{\overline{M}_w}{\overline{M}_n} \right) \leq 1.5$. Otherwise, it is calculated as

$$M = \overline{M}_n \left(\frac{1}{Viscfac_3 \times (3rdCSP)} \right)^{Viscfac_4 \times (4thCSP)} \times \left(\frac{\overline{M}_w}{\overline{M}_n} \right)^{Viscfac_4 \times (4thCSP)} \quad (26)$$

where $(3rdCSP)$ and $(4thCSP)$ are tuning parameters with default values equal to 1.0, \overline{M}_n is the number average molecular weight computed with equation 12, \overline{M}_w is the weight average molecular weight calculated with equation

13 and

$$Viscfac_3 = 0.2252 \frac{T}{\overline{M}_n} + 0.9738 \quad (27)$$

$$Viscfac_4 = 0.5354 \times Viscfac_3 - 0.1170 \quad (28)$$

as terms that take into account a temperature dependence for the oil average molecular weight. We highlight that this term is not present in Lindeloff et al. work, but it is considered in Pedersen et al..

The equation 24 is applied only in atmosphere pressure conditions. To capture the pressure effects, the following correction is employed

$$\eta_{heavy} = \eta'_{oil} \exp \left(0.00384 \frac{P^{0.8226} - 1}{0.8226} \right) \quad (29)$$

for viscosities in cP and pressures in atm.

In summary, to estimate mixture viscosity, the classical CSP model is applied for methane reference temperatures greater than 75 K. For reference temperatures < 50 K, the modified heavy oil model described above is used. And, for temperatures between 50 K and 75 K, the reference viscosity is calculated as a smooth transition with respect to each model as follows

$$\eta = F_1 \eta_{CSP} + F_2 \eta_{heavy} \quad (30)$$

with F_1 and F_2 computed as in equations 16 and 17, respectively, but now

$$\Delta T = T - T_{mean} \quad (31)$$

$$T_{mean} = \frac{50.0 + 75.0}{2} \quad (32)$$

The Elementary Effects method

The Elementary Effects^{11,12} (EE) method is a screening method in sensitivity analysis. This method evaluates sensitivity measures in a qualitative way, indicating the most influential parameters (or factors) of a computational model. The EE main advantage is that the method is not computationally expensive, allowing to estimate sensitivity measures for computationally intensive models with good accuracy even when compared to variance-based methods¹¹.

To evaluate the model factors sensibility measures, the EE method calculates incremental ratios called Elementary Effects. These ratios are calculated performing randomly One-factor-At-a-Time (OAT) perturbations in each model factor. In the following, we define the elementary effect related to the i -th input factor for a scalar model $y(\mathbf{X})$, with $\mathbf{X} = (X_1, X_2, \dots, X_k)$ as the k input factors re-scaled to $[0, 1]$. Thus, we have that:

$$d_i(\mathbf{X}) = \frac{y(X_1, \dots, X_i + \Delta, \dots, X_k) - y(\mathbf{X})}{\Delta} \quad (33)$$

where \mathbf{X} is uniformly distributed and sampled from the region of experimentation Ω , which is composed by a p -level grid for each X_i input factor^{11,12}. Here, the perturbation Δ is given as $p / [2(p - 1)]$, and p is an even natural number.

As sensitivity measures, based on [Campolongo et al.](#) work, we use the mean of the distribution of the absolute values of the elementary effects (s^*), which is sufficient to provide a reliable ranking of overall importance for the input factors. However, to ease the ranking procedure and to enhance its interpretability, we employ a normalized s^* instead:

$$s_i^* = \frac{1}{r} \sum_{j=1}^r |d_i(\mathbf{X}^{(j)})| \quad (34)$$

$$\bar{s}_i^* = \frac{s_i^*}{\sum_{l=1}^k s_l^*} \quad (35)$$

where r denotes the number of sampling (or trajectories) for each \mathbf{X} sampled from Ω . For further details about EE calculations, we refer to [Campolongo et al.](#), [Morris](#).

In this work, the sensitivity measures are used to determine which parameter are the most important in the CSP viscosity estimates up to an importance threshold. This way, we automatically select which CSP fitting parameters must be tuned to improve results with a computationally efficient strategy.

Bayesian CSP model calibration and uncertainty quantification

To tune the CSP viscosity model, we use a Bayesian approach to take into account uncertainties on experimental data. With tuning parameters calibrated through this approach, uncertainties on the CSP predicted viscosities can be estimated and we can assess the model limitations based on violations of credible regions. In contrast with classical deterministic fitting (e.g., Least-Squares regressions), the calibrated parameters are given as probability distributions, providing insights about the region of occurrence, the range of values and etc. Moreover, when tuning parameters' distributions are obtained by the calibration, an uncertainty propagation can be performed for predictive scenarios. Such additional information is not provided by deterministic fitting methods. As far as we know at the moment this manuscript is written, the aforementioned Bayesian analysis is unreported in the literature for CSP viscosity estimates.

The calibration procedure adopted here is based on the Bayes' theorem:

$$P_{\text{posterior}}(\mathbf{X}|\mathbf{y}) = \frac{P_{\text{likelihood}}(\mathbf{y}|\mathbf{X})P_{\text{prior}}(\mathbf{X})}{P_{\text{evidence}}(\mathbf{y})} \quad (36)$$

where $P_{\text{prior}}(\mathbf{X})$ is the probability distribution (or probability density function) of input parameters given *a priori*, the conditional probability $P_{\text{likelihood}}(\mathbf{y}|\mathbf{X})$ represents the likelihood of observations \mathbf{y} occur subject to input parameters data \mathbf{X} , and $P_{\text{evidence}}(\mathbf{y})$ denotes the evidence from observed data \mathbf{y} . The conditional probability $P_{\text{posterior}}(\mathbf{X}|\mathbf{y})$ is called posterior distribution of input parameters \mathbf{X} given that \mathbf{y} is true. The application of equation 36 combined with Markov Chain Monte Carlo (MCMC) methods results in a procedure that fits the posterior distribution in order to obtain the best parameters' distributions \mathbf{X} given the observed data \mathbf{y} .

To perform the Bayesian calibration, instead of initial guesses to input parameters, prior distributions ($P_{\text{prior}}(\mathbf{X})$) are required as initial hypothesis (e.g., input parameters are described

by uniform distributions within given ranges). These distributions are estimates that must be the most informative as possible to have outcomes closer to observations^{23,24}. The better is the initial hypothesis, the better will be the posterior distribution. Another required information beforehand is the likelihood function ($P_{\text{likelihood}}(\mathbf{y}|\mathbf{X})$). Here we assume that observable data follow a normal distribution with mean value (μ) equal to the simulation result (CSP viscosity estimate) with a standard deviation σ . Thus, we have that:

$$P_{\text{likelihood}}(\mathbf{y}|\mathbf{X}) = \prod_{i=1}^{N_{\text{exp}}} \frac{1}{\sigma\sqrt{2\pi}} \times \exp \left[-\frac{1}{2} \left(\frac{y_i^{\text{obs}} - y_i^{\text{model}}}{\sigma} \right)^2 \right] \quad (37)$$

in which y_i^{obs} denotes the experimental viscosity data, y_i^{model} are the CSP viscosity estimates, and N_{exp} is the number of experimental data. Since σ is unknown, we calibrate it as a hyperparameter.

In this work, we consider as input parameters (\mathbf{X}) the tuning CSP coefficients: (*1stCSP*), (*2ndCSP*), (*3rdCSP*), and (*4thCSP*). To perform Monte Carlo simulations, we employ a Transitional Markov Chain Monte Carlo method called Cascading Adaptive Transition Metropolis in Parallel⁸ (CATMIP). For further details about Bayesian model calibration, uncertainty quantification and MCMC methods, we refer to [Oden et al.](#), [McClarren](#), [Sivia and Skilling](#), [Box and Tiao](#), [Tarantola](#).

A unified approach for heavy oil viscosity estimates

In this sub-section, we propose a strategy to automatically tune CSP viscosity model based on the sensitivity analysis and Bayesian calibration described in the previous sub-sections. Such strategy is motivated by the fact that CSP viscosity model has four tuning parameters but, depending on the oil characteristics, just a subset of tuning parameters is required. Analyzing the CSP model, it is clear that (*1stCSP*)

and (*2ndCSP*) are related to light oils (small amount of heavier components fractions) or gases such that the reference state is calculated using methane. Likewise, when the reference state is obtained using a crude oil based fluid, parameters (*3rdCSP*) and (*4thCSP*) play an important role in model tuning. However, how we could determine which parameters have to be tuned in a generic way, i.e., for any given mixture and regardless of CSP configuration? This is the fundamental question that motivated us to elaborate a simple and practical solution – that we describe in the following – using the method of Elementary Effects.

The main idea is to use EE method to construct a rank of importance for the CSP tuning parameters. With the importance rank, we apply an importance threshold to filter which parameters are relevant to the model. Here we select parameters according to \bar{s}_i^* measures. The EE calculations have low computational cost (in time) and this strategy is a non-intrusive way to assess the most influential parameters, thus modifications in the model are not required and the application is straightforward. Below we outline the steps for automatically perform CSP viscosity model calibration:

1. Provide prior distributions for all CSP tuning parameters, set the percentage of these parameters' value to be used as perturbation in EE calculations (e.g., 25% of the default values) and set the importance threshold to filter tuning parameters;
2. Set thermodynamic data for CSP viscosity model (pressure, temperature, composition, and etc). This step requires classical procedures as mixture characterization, equation of state setup and flash calculation to determine the oil composition (liquid phase);
3. Perform sensitivity analysis on CSP tuning parameters with EE method. At the end of this step, an importance ranking for tuning parameters is obtained;
4. Based on the importance threshold, filter the most influential CSP tuning parameters;

5. Calibrate CSP viscosity model with Bayesian approach using only the filtered tuning parameters.

Some remarks about the strategy above: the last step is performed by deterministic fitting methods in the literature^{3,7,18}. Thus, if desired due to frequent use, deterministic tuning can be employed instead of the Bayesian approach. Another relevant point is the way that sensitivity analysis is applied here is not new, since that such analysis is commonly used to rank and determine the most relevant factors that affect model outcomes^{26,27}. However, in the context of CSP viscosity estimation, sensitivity analysis and clarification on which tuning parameters must be calibrated based only on model response to perturbations are unreported in the literature. Moreover, the determination of which CSP tuning parameters have to be subject to calibration reduces the size of parameter vector and improves the calibration procedure, decreasing computational time cost and increasing parameters' identifiability^{28,29}.

Results

In this section, we present results of the proposed unified approach on each oil sample described in Experimental data section. For comparison purpose only, calibration is also performed with a Least-Squares formulation (deterministic) solved by a Dual Annealing³⁰ method, which is a global optimization algorithm available at SciPy³¹ (version 1.3.2). All thermodynamics models, equilibria calculations and CSP viscosity estimations are implemented in a proprietary software called PVTlib³². For Sensitivity Analysis, the Elementary Effects method from SALib³³ (version 1.3.11) is employed. Probabilistic programming and Monte Carlo simulations (CATMIP) are implemented with PyMC3³⁴ (version 3.8).

In the following, for each oil sample we provide three different results:

- Sensitivity measure \bar{s}_i^* by Elementary Effects method. We use this result to apply importance threshold as $\bar{s}_i^* \geq 0.05$. If this

criterion is satisfied for *any* experimental condition in the simulation, the related parameter is calibrated;

- A comparison of Bayesian approach for model Calibration and Uncertainty Quantification (UQ) with and without the deterministic fitting (Least-Squares), showing the CSP result when the fitting is not performed;
- A table summarizing statistics of calibrated parameters (and hyper-parameters): most probable value (MPV), highest density interval (HDI) of 3%-97%, and standard deviation.

Table 6 summarizes the simulation setup for each oil case. For all cases, we draw 2500 samples from the posterior distributions and used 25 Monte Carlo steps in each CATMIP stage, which are `draws` and `n_steps` CATMIP arguments in PyMC3, respectively. We used two heavier components characterization for heavy oils: based on Soave-Redlich-Kwong¹⁷ (SRK) and Peng-Robinson³⁵ (PR) cubic equations of state. The characterized and lumped mixture for each oil case is provided at the Supporting Information file (see A.1). The posterior distributions for CSP tuning parameters and standard deviation (hyper-parameter) are available at A.2. For the Elementary Effects method, we used $p = 4$ grid levels and 40 trajectories. We considered a region of experimentation construct by a range of $\pm 25\%$ for each CSP tuning parameter, i.e., 0.75 to 1.25. The results for each case is provided below.

Table 6: The prior probability distributions used in the Bayesian calibration for model parameters and hyper-parameter. Heavier components characterization is also shown for each case. \mathcal{U} denotes uniform distribution.

Case	Characterization	1stCSP	2ndCSP	3rdCSP	4thCSP	Standard Deviation (σ)
Oil 1	SRK	-	-	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 10])$
Oil 2	SRK	-	-	$\mathcal{U}([0, 3])$	-	$\mathcal{U}([0, 3])$
Oil 5	PR	-	-	$\mathcal{U}([0, 8])$	$\mathcal{U}([0, 8])$	$\mathcal{U}([0, 2000])$
Oil 7	SRK	-	-	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 4000])$
Oil 8	SRK	-	-	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 3])$	$\mathcal{U}([0, 200])$

Oil 1

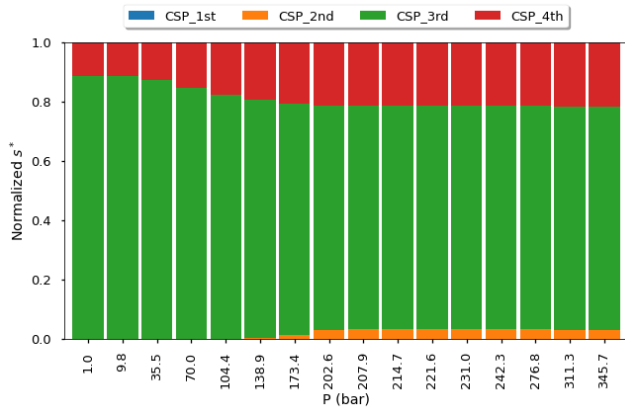


Figure 1: Normalized absolute mean elementary effect for each Oil 1 viscosity measurement.

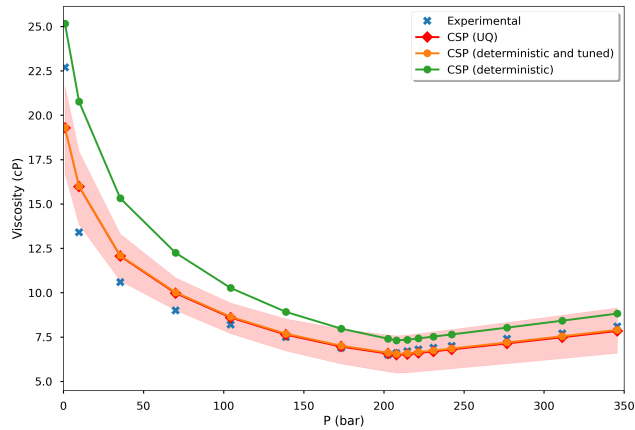


Figure 2: CSP calibration and uncertainty quantification for Oil 1. The credible region is drawn from 1st percentile to 99th percentile.

Table 7: 94% Highest Density Interval (HDI) and Most Probable Value (MPV) obtained for Oil 1.

	HDI 3%	HDI 97%	MPV
Std deviation	0.912914	1.990789	1.301332
3rdCSP	1.054233	1.149737	1.108259
4thCSP	0.933108	1.352186	1.149314

Oil 2

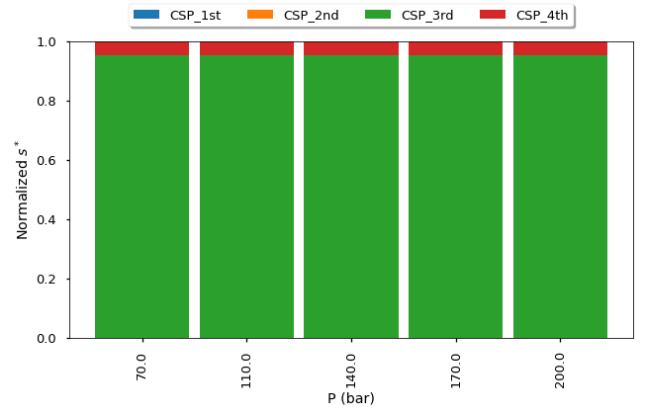


Figure 3: Normalized absolute mean elementary effect for each Oil 2 viscosity measurement.

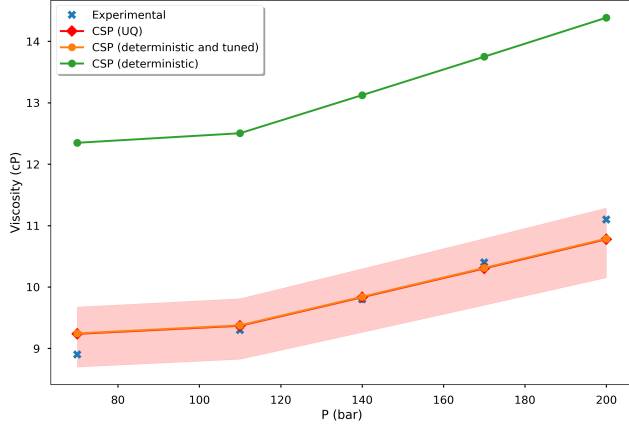


Figure 4: CSP calibration and uncertainty quantification for Oil 2. The credible region is draw from 1st percentile to 99th percentile.

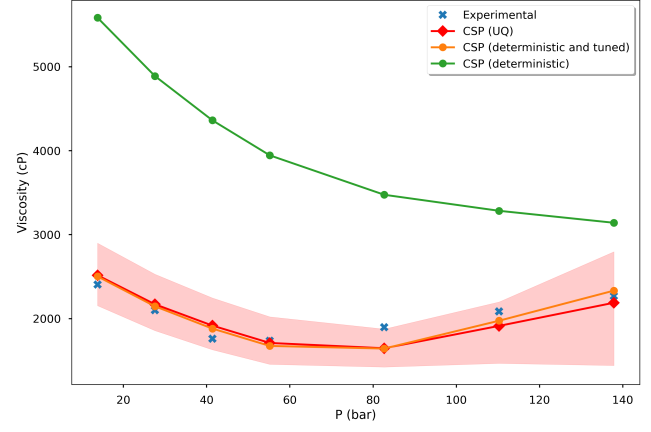


Figure 6: CSP calibration and uncertainty quantification for Oil 5. The credible region is draw from 1st percentile to 99th percentile.

Table 8: 94% Highest Density Interval (HDI) and Most Probable Value (MPV) obtained for Oil 2.

	HDI 3%	HDI 97%	MPV
Std deviation	0.115992	0.795249	0.259651
3rdCSP	1.097432	1.126872	1.112312

Table 9: 94% Highest Density Interval (HDI) and Most Probable Value (MPV) obtained for Oil 5.

	HDI 3%	HDI 97%	MPV
Std deviation	80.099300	443.740566	161.381002
3rdCSP	1.261799	1.318719	1.298014
4thCSP	2.004492	5.466199	4.188029

Oil 5

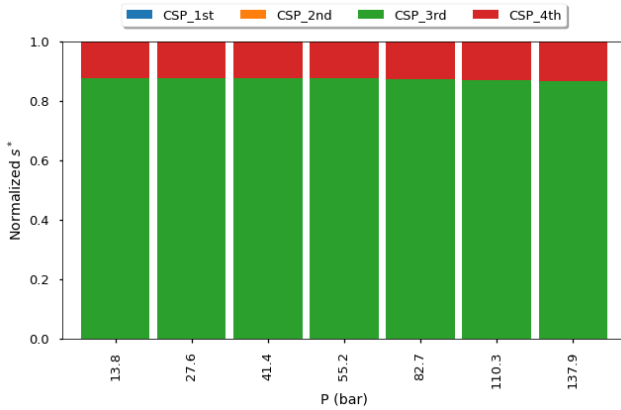


Figure 5: Normalized absolute mean elementary effect for each Oil 5 viscosity measurement.

Oil 7

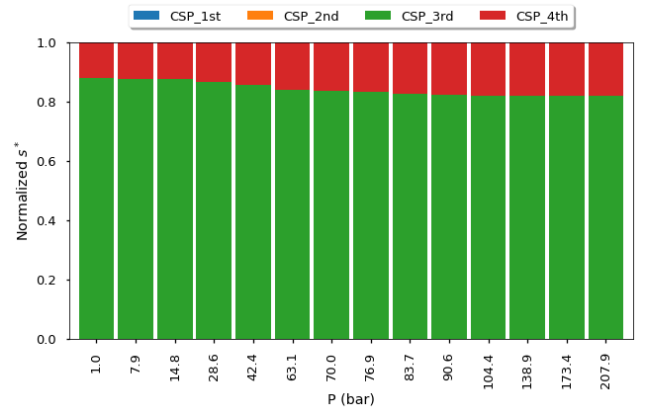


Figure 7: Normalized absolute mean elementary effect for each Oil 7 viscosity measurement.

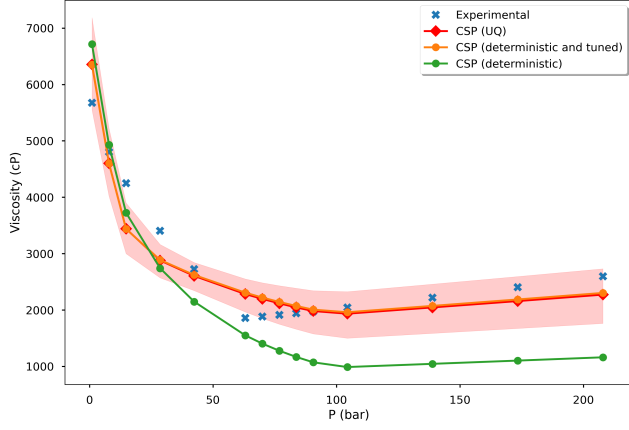


Figure 8: CSP calibration and uncertainty quantification for Oil 7. The credible region is draw from 1st percentile to 99th percentile.

Table 10: 94% Highest Density Interval (HDI) and Most Probable Value (MPV) obtained for Oil 7.

	HDI 3%	HDI 97%	MPV
Std deviation	294.597272	666.682146	421.596259
3rdCSP	1.111932	1.172000	1.145146
4thCSP	1.535499	1.931302	1.758063

Oil 8

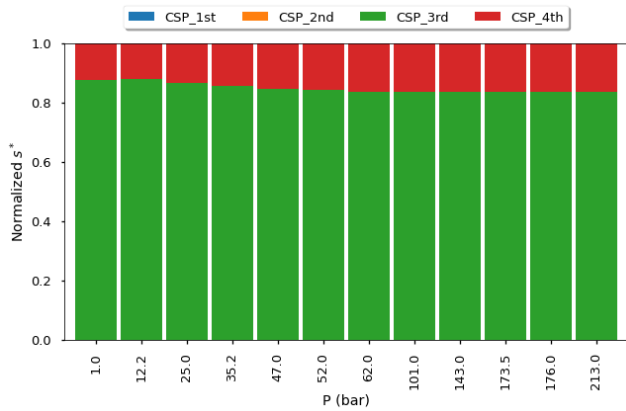


Figure 9: Normalized absolute mean elementary effect for each Oil 8 viscosity measurement.

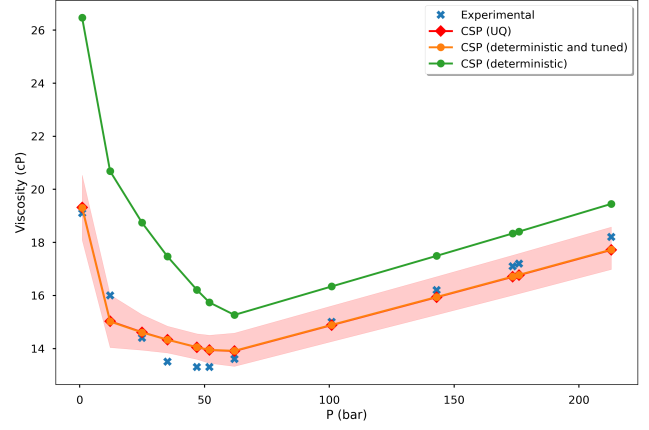


Figure 10: CSP calibration and uncertainty quantification for Oil 8. The credible region is draw from 1st percentile to 99th percentile.

Table 11: 94% Highest Density Interval (HDI) and Most Probable Value (MPV) obtained for Oil 8.

	HDI 3%	HDI 97%	MPV
Std deviation	0.392378	1.054078	0.615806
3rdCSP	1.146479	1.185162	1.169052
4thCSP	1.575960	2.032369	1.800634

Discussion

In this section, we discuss and analyze the results. As a starting point, we note that filtered tuning parameters results show good agreement between deterministic and Bayesian for all cases. The best deterministic curve almost matches the 50th-percentile curve from Bayesian approach, indicating a proper verification (and consistency) of the methods. More specifically for the study cases used in this work, only (*3rdCSP*) and (*4thCSP*) are identified as influential parameters for CSP viscosity model calibration for all oil samples. This result is expected, since we are considering only heavy oil samples.

About the sensitivity measures and how they were applied in the cases, as the viscosity experiments were performed for fixed temperature while pressure is define at a wide range, the normalized overall sensitivity measures showed just slightly changes within all evaluated pressure

conditions, suggesting that adopting $\bar{s}_i^* \geq 0.05$ as calibration criterion is satisfied for *any* experimental condition is reasonable for the cases considered in the present work.

From the tables summarizing the Highest Density Interval (HDI), it is clear that parameters with higher sensitivity measures have smaller HDI range. This is expected because changes in the most influential parameters generate larger changes on model outcomes, while less influential parameters produces smaller changes on outcomes due to perturbations.

Relevant points have to be mentioned for CSP model viscosity. For Oils 1, 7 and 8 the employed model exhibits statistical violations (credible regions do not cover all experimental data). Such violations occur mainly in the relatively low pressure regions of the experiments (below 100 bar). These regions also show a large uncertainty range and curve slope variations. A possible explanation is that when pressure decreases, light hydrocarbon components leave liquid phase increasing the fraction of heavier components in the liquid phase. Thus, liquid phase becomes more dense and possibly exhibits a behavior closer to Non-Newtonian, in which heavy oil CSP viscosity model is not appropriate^{36,37}. This interpretation is clarified due to the Credible Region provided by the Bayesian approach, which is not present in classical deterministic regressions.

Finally, it is worth note that CSP viscosity model considered here demonstrated that it has to be tuned. Without tuning, it shows high discrepancy/error when compared to experimental data. Hence, results suggests that CSP viscosity model has to be tuned at least when dealing with heavy oils.

Conclusions

We presented a probabilistic framework considering sensitivity analysis, calibration and uncertainty quantification applied on heavy oil estimations with a corresponding states model (CSP). This approach provides new information when compared with classic deterministic methods (for instance, Least-Squares meth-

ods), such as regions where the prediction is not reliable as well as which parts the range of uncertainties are wider. This way, additional knowledge of the analyzed CSP viscosity model is obtained concerning the applicability context in the light of statistical violations. Moreover, the non-intrusive sensitivity analysis method of Elementary Effects was employed to determine the most influential fitting parameters. With such information, parameters which do not effectively contribute to model calibration are neglected in this stage, decreasing time computational costs and enhancing calibration performance. The results clarified the limitations of CSP viscosity model applied to heavy oil samples, indicating that for low pressure and mid temperature conditions, the model can be inappropriate. The presented unified approach has potential to be automatically applied on data from viscosity experiment reports, avoiding the necessity to perform trials-and-errors strategy to fit the model and choose which parameters must be calibrated.

Supporting Information Available

Additional results for oil characterization and Bayesian CSP model calibration.

References

- (1) Whitson, C. H.; Brulé, M. R., et al. *Phase behavior*; Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers . . . , 2000; Vol. 20.
- (2) Pedersen, K. S.; Christensen, P. L.; Shaikh, J. A. *Phase Behaviour of Petroleum Reservoir Fluids*; CRC press, 2014; Vol. 2; p 443.
- (3) Pedersen, K. S.; Fredenslund, A.; Christensen, P. L.; Thomassen, P. Viscosity of crude oils. *Chemical Engineering Science* **1984**, *39*, 1011–1016.
- (4) Lohrenz, J.; Bray, B. G.; Clark, C. R., et al. Calculating viscosities of reservoir

- fluids from their compositions. *Journal of Petroleum Technology* **1964**, *16*, 1–171.
- (5) Ely, J. F.; Hanley, H. Prediction of transport properties. 1. Viscosity of fluids and mixtures. *Industrial & Engineering Chemistry Fundamentals* **1981**, *20*, 323–332.
 - (6) Hernandez, J.; Vesovic, V.; Carter, J.; Lopez, E. Sensitivity of reservoir simulations to uncertainties in viscosity. SPE/DOE Improved Oil Recovery Symposium. 2002.
 - (7) Lindeloff, N.; Pedersen, K. S.; Ronningsen, H. P.; Milner, J., et al. The corresponding states viscosity model applied to heavy oil systems. Canadian International Petroleum Conference. 2003.
 - (8) Minson, S. E.; Simons, M.; Beck, J. L. Bayesian inversion for finite fault earthquake source models I – theory and algorithm. *Geophysical Journal International* **2013**, *194*, 1701–1726.
 - (9) Oden, J. T.; Babuška, I.; Faghihi, D. Predictive computational science: Computer predictions in the presence of uncertainty. *Encyclopedia of Computational Mechanics Second Edition* **2017**, 1–26.
 - (10) McClarren, R. G. *Uncertainty Quantification and Predictive Computational Science*; Springer, 2018.
 - (11) Campolongo, F.; Cariboni, J.; Saltelli, A. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software* **2007**, *22*, 1509–1518.
 - (12) Morris, M. D. Factorial sampling plans for preliminary computational experiments. *Technometrics* **1991**, *33*, 161–174.
 - (13) Pedersen, K.; Thomassen, P.; Fredenslund, A. SRK-EOS calculation for crude oils. *Fluid Phase Equilibria* **1983**, *14*, 209–218.
 - (14) Pedersen, K. S.; Thomassen, P.; Fredenslund, A. Thermodynamics of petroleum mixtures containing heavy hydrocarbons. 1. Phase envelope calculations by use of the Soave-Redlich-Kwong equation of state. *Industrial & Engineering Chemistry Process Design and Development* **1984**, *23*, 163–170.
 - (15) Pedersen, K. S.; Thomassen, P.; Fredenslund, A. Thermodynamics of Petroleum Mixtures Containing Heavy Hydrocarbons. 3. Efficient Flash Calculation Procedures Using the SRK Equation of State. *Industrial and Engineering Chemistry Process Design and Development* **1985**, *24*, 948–954.
 - (16) Pedersen, K. S.; Blilie, A. L.; Meisingset, K. K. PVT Calculations on Petroleum Reservoir Fluids Using Measured and Estimated Compositional Data for the Plus Fraction. *Industrial and Engineering Chemistry Research* **1992**, *31*, 1378–1384.
 - (17) Kreibjerg, K.; Pedersen, K., et al. Controlling VLE equilibrium with a cubic EOS in heavy oil modeling. Canadian International Petroleum Conference. 2006.
 - (18) Pedersen, K. S.; Fredenslund, A. An improved corresponding states model for the prediction of oil and gas viscosities and thermal conductivities. *Chemical Engineering Science* **1987**, *42*, 182–186.
 - (19) van der Waals, J. D. The law of corresponding states for different substances. KNAW, Proceedings. 1913; pp 1912–1913.
 - (20) Murad, S.; Gubbins, K. Corresponding states correlation for thermal conductivity of dense fluids. *Chemical Engineering Science* **1977**, *32*, 499–505.
 - (21) McCarty, R. A modified Benedict-Webb-Rubin equation of state for methane using recent experimental data. *Cryogenics* **1974**, *14*, 276–280.

- (22) Hanley, H. Prediction of the viscosity and thermal conductivity coefficients of mixtures. *Cryogenics* **1976**, *16*, 643–651.
- (23) Sivia, D.; Skilling, J. *Data analysis: a Bayesian tutorial*; OUP Oxford, 2006.
- (24) Box, G. E.; Tiao, G. C. *Bayesian inference in statistical analysis*; John Wiley & Sons, 2011; Vol. 40.
- (25) Tarantola, A. *Inverse problem theory and methods for model parameter estimation*; SIAM, 2005.
- (26) Saltelli, A.; Tarantola, S.; Campolongo, F.; Ratto, M. *Sensitivity analysis in practice: a guide to assessing scientific models*; Wiley Online Library, 2004; Vol. 1.
- (27) Trucano, T. G.; Swiler, L. P.; Igusa, T.; Oberkampf, W. L.; Pilch, M. Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering & System Safety* **2006**, *91*, 1331–1357.
- (28) Loeppky, J.; Bingham, D.; Welch, W. Computer model calibration or tuning in practice. *University of British Columbia, Vancouver, BC, Canada* **2006**,
- (29) Arendt, P. D.; Apley, D. W.; Chen, W. Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. *Journal of Mechanical Design* **2012**, *134*.
- (30) Xiang, Y.; Sun, D.; Fan, W.; Gong, X. Generalized simulated annealing algorithm and its application to the Thomson model. *Physics Letters A* **1997**, *233*, 216–220.
- (31) Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272.
- (32) Silva, D. M. D. D.; Soprana, A. B.; Manoel, E. T. M.; Volpato, D. T.; Bittencourt, S. P. Development of a software library for equilibria simulations of multiphase multicomponent hydrocarbon systems. XI Iberoamerican Conference on Phase Equilibria and Fluid Properties for Process Design: Equifase 2018. 2018.
- (33) Herman, J.; Usher, W. SALib: an open-source Python library for sensitivity analysis. *Journal of Open Source Software* **2017**, *2*, 97.
- (34) Salvatier, J.; Wiecki, T. V.; Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2016**, *2*, e55.
- (35) Pedersen, K.; Milter, J.; Sørensen, H., et al. Cubic equations of state applied to HT/HP and highly aromatic fluids. SPE Annual Technical Conference and Exhibition. 2002.
- (36) Roenningsen, H. P.; Bjoerndal, B.; Baltzer Hansen, A.; Batsberg Pedersen, W. Wax precipitation from North Sea crude oils: 1. Crystallization and dissolution temperatures, and Newtonian and non-Newtonian flow properties. *Energy & Fuels* **1991**, *5*, 895–908.
- (37) Roenningsen, H. P. Prediction of viscosity and surface tension of North Sea petroleum fluids by using the average molecular weight. *Energy & fuels* **1993**, *7*, 565–573.