

Global sensitivity analysis of Alkali–Surfactant–Polymer enhanced oil recovery processes

Enrique Carrero, Nestor V. Queipo ^{*}, Salvador Pintos, Luis E. Zerpa

Applied Computing Institute, Faculty of Engineering, University of Zulia, Venezuela

Received 13 April 2006; received in revised form 10 November 2006; accepted 14 November 2006

Abstract

After conventional waterflooding processes the residual oil in the reservoir remains as a discontinuous phase in the form of oil drops trapped by capillary forces and is likely to be around 70% of the original oil in place (OOIP). The EOR method so-called Alkaline–Surfactant–Polymer (ASP) flooding has been proved to be effective in reducing the oil residual saturation in laboratory experiments and field projects through reduction of interfacial tension and mobility ratio between oil and water phases.

A critical step for the optimal design and control of ASP recovery processes is to find the relative contributions of design variables such as, slug size and chemical concentrations, in the variability of given performance measures (e.g., net present value, cumulative oil recovery), considering a heterogeneous and multiphase petroleum reservoir (sensitivity analysis).

Previously reported works using reservoir numerical simulation have been limited to local sensitivity analyses because a global sensitivity analysis may require hundreds or even thousands of computationally expensive evaluations (field scale numerical simulations). To overcome this issue, a surrogate-based approach is suggested.

Surrogate-based analysis/optimization makes reference to the idea of constructing an alternative fast model (surrogate) from numerical simulation data and using it for analysis/optimization purposes. This paper presents an efficient global sensitivity approach based on Sobol's method and multiple surrogates (i.e., Polynomial Regression, Kriging, Radial Base Functions and a Weighed Adaptive Model), with the multiple surrogates used to address the uncertainty in the analysis derived from plausible alternative surrogate-modeling schemes.

The proposed approach was evaluated in the context of the global sensitivity analysis of a field scale Alkali–Surfactant–Polymer flooding process. The design variables and the performance measure in the ASP process were selected as slug size/concentration of chemical agents, and cumulative oil recovery, respectively. The results show the effectiveness and efficiency of the proposed approach since it allows establishing the relative contribution of the design variables (main factors and interactions) to the performance measure variability using a limited number of computationally expensive reservoir simulations.

© 2006 Elsevier B.V. All rights reserved.

Keywords: ASP flooding; Global sensitivity analysis; Surrogate-based modeling; Ensemble of surrogates

1. Introduction

After conventional waterflooding processes the residual oil in the reservoir remains as a discontinuous phase in the form of oil drops trapped by capillary forces and is likely to be around 70% of the original oil in place —

^{*} Corresponding author. Tel.: +58 261 7598630; fax: +58 261 7598411.

E-mail addresses: ecarrero@ica.luz.ve (E. Carrero), nqueipo@ica.luz.ve (N.V. Queipo), spintos@ica.luz.ve (S. Pintos), lzerpa@ica.luz.ve (L.E. Zerpa).

OOIP (Doshier and Wise, 1976). The EOR method so-called Alkaline–Surfactant–Polymer (ASP) flooding has been proved to be effective in reducing the oil residual saturation in laboratory experiments and field projects through reduction of interfacial tension and mobility ratio between oil and water phases (Clark et al., 1988; Meyers et al., 1992; Vargo et al., 1999; Demin et al., 1999).

A critical step for the optimal design and control of ASP recovery processes is to find the relative contributions of design variables (global sensitivity analysis) such as, slug size and chemical concentrations, in the variability of given performance measures (e.g., net present value, cumulative oil recovery), considering a heterogeneous and multiphase petroleum reservoir. Global sensitivity analysis (Saltelli and Tarantola, 2002 and references therein) allows addressing settings such as:

- What are the main effects (model output vs. input variables)?
- Can we safely fix one or more of the input variables without significantly affecting the output variability (screening)?
- How can we rank a set of input variables according to their contribution to the output variability (variables prioritization)?
- If we could eliminate the uncertainty of one or more of the input variables which ones should be chosen (variable selection for maximum uncertainty reduction)?
- If and which (group of) parameters interact with each other (parameter interactions)?
- What are the main regions of interest in the parameter space if additional samples become available?
- Does the model reproduce well known behavior of the process of interest (model validation)?

Previously reported works using reservoir numerical simulation have been limited to local sensitivity analyses (Wu, 1996; Zhijian et al., 1998; Manrique et al., 2000; Qi et al., 2000; Hernández et al., 2001) and global but discrete analyses —factorial designs (Delshad et al., 2005). The latter approach given its discrete nature may hide the true nature of the process behavior or may require an unaffordable amount of experiments. While the ideal approach is a continuous global sensitivity analysis, it would require hundreds or even thousands of computationally expensive evaluations (field scale numerical simulations). To overcome this issue, a surrogate-based approach is suggested. Surrogate-based analysis/optimization makes reference to the idea of constructing an alternative fast model

(surrogate) from numerical simulation data and using it for analysis/optimization purposes. This approach has been successful in the analysis/optimization of computationally expensive simulation-based models in the aerospace (Giunta et al., 1997; Balabanov et al., 1998), automotive (Craig et al., 2002; Kurtaran et al., 2002) and oil industries (Queipo et al., 2002a,b). Li and Padula (2004) and Queipo et al. (2005) recently reviewed different surrogate models used in the aerospace industry.

On the other hand, comparative studies have shown that depending on the problem under consideration a particular modeling scheme may outperform the others, and in general, it is not known a priori which one should be selected (Simpson et al., 2001; Jin et al., 2001). This paper presents an efficient global sensitivity approach based on Sobol's method (Sobol, 1993) and multiple surrogates (i.e., Polynomial Regression, Kriging, Radial Base Functions and a Weighed Adaptive Model) with the multiple surrogates used to address the uncertainty in the analysis derived from plausible alternative surrogate-modeling schemes (Zerpa et al., 2005; Sanchez et al., 2006; Goel et al., in press). The proposed approach is evaluated using a field scale Alkali–Surfactant–Polymer flooding process (modeled with the UTCHEM reservoir numerical simulator) with design variables and performance measure selected as slug size/concentration of chemical agents, and cumulative oil recovery, respectively.

The UTCHEM is a three-dimensional, multiphase, multicomponent reservoir simulator of chemical flooding processes developed at the University of Texas at Austin (Pope and Nelson, 1978; Camilleri et al., 1987; Bhuyan et al., 1990). The basic governing differential equations consist of: a mass conservation equation for each component, an overall mass conservation equation that determines the pressure (the pressure equation), an energy balance, and Darcy's Law generalized for multiphase flow. The resulting flow equations are solved using a block-centered finite-difference scheme. The solution method is implicit in pressure and explicit in concentration, similar to the well known IMPES method used in blackoil reservoir simulators. A Jacobi conjugate gradient method is used to solve the system of finite-difference equations resulted from the discretization of the pressure equation.

2. Problem of interest

While a global sensitivity analysis allows addressing the issues posed in the introduction, in this paper we

focus on answering the following questions in the context of oil recovery processes:

- Can we safely fix one or more of the input variables without significantly affecting the output variability (screening)?
- How can we rank a set of input variables according to their contribution to the output variability (variables prioritization)?
- If we could eliminate the uncertainty of one or more of the input variables which ones should be chosen (variable selection for maximum uncertainty reduction)?
- If and which (group of) parameters interact with each other (parameter interactions)?

More precisely, given an ASP flooding oil recovery process, having design variables such as, chemical concentrations of alkali, surfactant, and polymer, and slug size, what are the most important variables and interactions?

It will be shown (Section 3.1) that all these questions can be answered by solving the following problem: Given a function $f(\mathbf{x})$ of vector of design variables x of size N_{dv} , find its decomposition as the sum of terms of increasing dimensionality as:

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i < j} f_{ij}(x_i, x_j) + \dots + f_{12\dots K}(x_1, x_2, \dots, x_K) \quad (1)$$

with,

$$\int_0^1 f_{i_1 \dots i_s} dx_k = 0 \quad (2)$$

$k = i_1, \dots, i_s$, where $1 \leq i_1 < \dots < i_s \leq N_{dv}$,

3. Solution methodology

With reference to Fig. 1, the proposed methodology involves the following steps:

1. Generate a sample of the design variables space using a modified Latin hypercube experimental design. This sampling procedure has been shown to be very effective for selecting values of input variables for the analysis of the output of a computer code (McKay et al., 1979).
2. Conduct numerical simulations (via UTCHEM) using the sample (input) from the previous step and

obtain the corresponding objective function values (output).

3. Using the input/output pairs obtained in the previous step construct multiple surrogate models based on Polynomial Regression, Kriging, Radial Basis Functions and a Weighted Adaptive Model. This surrogate models will be discussed later in this section.
4. With the surrogate models constructed, conduct a global sensitivity analysis using Sobol's method to calculate the global and total sensitivity indices for each of the design variables. These sensitivity indices will allow the ranking of design variables in order of importance.

3.1. Global sensitivity analysis

To understand the concept, assume a surrogate model of a square integrable function, $f(\mathbf{x})$, as a function of a vector of design variables, x , whose values have been scaled between zero and one (this assumes that the design domain is box-like). This surrogate model can be decomposed as the sum of functions of increasing dimensionality as:

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i < j} f_{ij}(x_i, x_j) + \dots + f_{12\dots K}(x_1, x_2, \dots, x_K) \quad (3)$$

If the following condition

$$\int_0^1 f_{i_1 \dots i_s} dx_k = 0 \quad (4)$$

is imposed for $k = i_1, \dots, i_s$, where $1 \leq i_1 < \dots < i_s \leq N_{dv}$, the decomposition described in Eq. (3) is unique and each term in the sum can be obtained by computing the following integrals:

$$\int f(\mathbf{x}) \Pi dx_k = f_0 \quad (5)$$

$$\int f(\mathbf{x}) \prod_{k \neq i} dx_k = f_0 + f_i(x_i) \quad (6)$$

from which $f_i(x_i)$ can be found, and

$$\int f(\mathbf{x}) \prod_{k \neq i, j} dx_k = f_0 + f_i(x_i) + f_j(x_j) + f_{i,j}(x_i, x_j) \quad (7)$$

from which $f_{i,j}(x_i, x_j)$ can be obtained. The higher dimensional summands are similarly found except for the last one that is calculated using Eq. (3). Furthermore, ensured by the condition expressed in Eq. (4), the summands are orthogonal.

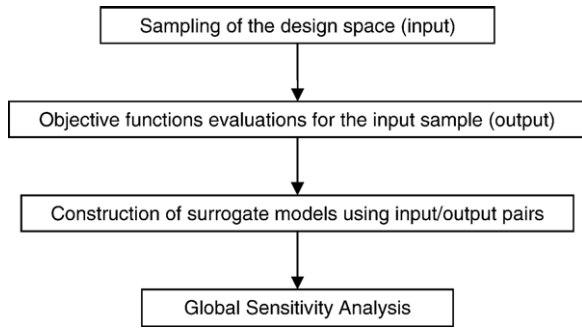


Fig. 1. Architecture of the proposed methodology.

Therefore the partial variances, that is, the contribution of each of the summands to the total variance observed in the response, can be shown to be:

$$D_{i_1 \dots i_s} = \int f_{i_1 \dots i_s}^2 dx_{i_1} \dots dx_{i_s} \quad (8)$$

with the total variance being equal to:

$$D = \int f^2(\mathbf{x}) d\mathbf{x} - f_0^2 \quad (9)$$

which can also be expressed as,

$$D = \sum_{s=1}^p \sum_{i_1 < \dots < i_s}^p D_{i_1 \dots i_s} \quad (10)$$

since the functions have been shown to be orthogonal.

Each partial variance gives a measure of the contribution of each independent variable or set of variables to the total variance, and provides an indication of their relative importance. Note that all the required integrations are conducted on the surrogate (fast) model and can in principle be calculated accurately provided an integration numerical procedure is available (e.g., Gaussian quadrature).

The relative importance of a design variable is quantified by a set of indices, namely, individual (S_i) and total (S_i^{total}) sensitivity indices. The former refer to the fraction of the total variance contributed by a particular variable in isolation, while the latter represents the contribution (expressed as a fraction) of all the partial variances in which the variable of interest is involved. The individual sensitivity index S_i represents the influence of a design variable x_i to a function variability without accounting for any of its interactions with other variables and is given as:

$$S_i = D_i / D \quad (11)$$

To calculate the total sensitivity index (S_i^{total}) of a design variable x_i , the design variable vector x is divided into two complementary subsets, x_i and \mathbf{Z} where \mathbf{Z} is a vector containing $x_1, x_2, x_3, \dots, x_n$ ($n \neq i$). The purpose of using these subsets is to isolate the influence of x_i on the $f(\mathbf{x})$ variability from the influence of the remaining design variables included in \mathbf{Z} . The total sensitivity index for x_i can be calculated as:

$$S_i^{\text{total}} = D_i^{\text{total}} / D \quad (12)$$

and

$$D = D_i^{\text{total}} + D_Z \quad (13)$$

where D_Z is defined as the sum of partial variances of all possible subsets variables in \mathbf{Z} . The D_i^{total} term is calculated as $D - D_Z$ since D_Z can be straightforwardly calculated (Sobol, 1993).

Formulations of the Sobol's method that account for non-rectangular domains and correlated inputs are available; see, for example, Jacque et al. (2004), and Mack et al. (2005) for a recent application. A detailed discussion of global sensitivity methods and applications can be found in Saltelli and Tarantola (2002), Frey and Patil (2002) and the references therein.

3.2. Surrogate-modeling

It is an inverse problem where due to the limited amount of available data: i) alternative surrogates can provide reasonable approximations to function f , and ii) each surrogate may offer the best fit to f depending on the region of the design space. Since the contributions of design variables to the variability of the objective function are unknown we suggest to use multiple surrogates considering they can be constructed at no significant additional computational cost. Three alternative surrogate models, namely, Polynomial Regression (PRG), Kriging (KRG), and Radial Basis Functions (RBF) will be considered. In addition, a Weighted Adaptive Model (WAM) of these surrogates will also be included in the

Table 1
Design variable restrictions

Design variable	Range		Units
	Min	Max	
Alkaline concentration (Na_2CO_3)	0	0.5898	(meq/ml)
Surfactant concentration	0.001815	0.005	(Vol. fract.)
Polymer concentration	0.0487	0.12	(wt.%)
Slug size	111	326	(days)

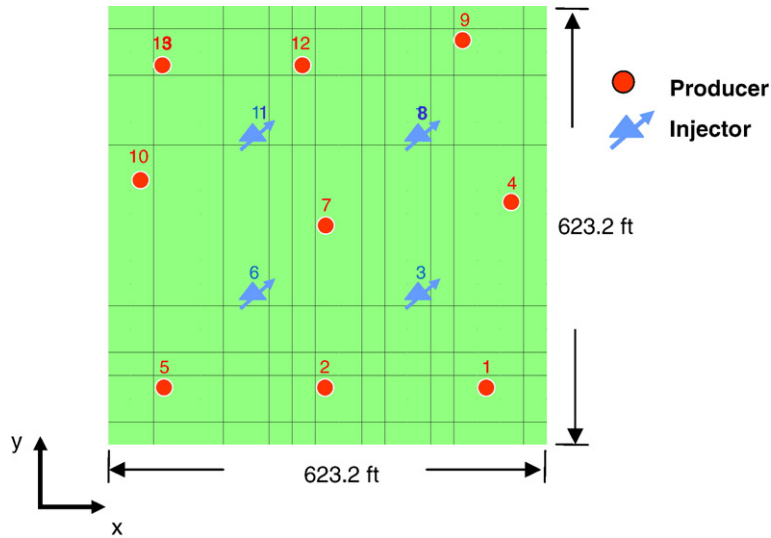


Fig. 2. Well pattern illustration.

study. The WAM can be shown to reduce the variance estimation with respect to that of the individual surrogates (Bishop, 1995). Throughout this section, given the stochastic nature of the surrogates, the available data is considered a sample of a population.

3.2.1. Polynomial Regression model (PRG)

The regression analysis is a methodology that studies the quantitative association between a function of interest y , and m prediction variables z_i^j , where there are n values of the function of interest y_i , for a set of prediction variables z_i^j (Draper and Smith, 1966). For each observation i a linear equation is formulated:

$$y_i = \sum_{j=1}^m \beta_j z_i^j + \varepsilon_i \quad E(\varepsilon_i) = 0 \quad V(\varepsilon_i) = \sigma^2 \quad (14)$$

where the errors ε_i are independents with expected value equal to zero and variance σ^2 . The estimated parameters $\hat{\beta}_j$ (by least squares) are unbiased and have minimum variance.

Eq. (1) is expressed in matrix form as:

$$y = Z\beta + \varepsilon \quad E(\varepsilon) = 0 \quad V(\varepsilon) = \sigma^2 I \quad (15)$$

where Z is a $n \times m$ matrix with the prediction variable values. The vector of the estimated parameters is:

$$\hat{\beta} = (Z^T Z)^{-1} Z^T y \quad (16)$$

Considering a new set of design values z , the variance of the predicted response $z^T \hat{\beta}$ is:

$$V(y(z)) = \sigma^2 (z^T (Z^T Z)^{-1} z + 1) \quad (17)$$

In this work the regression model considered is a second-order polynomial model of the form:

$$y = \beta_0 + \sum_{i=1}^p \beta_i x_i + \sum_{i=1}^p \sum_{j=1}^p \beta_{ij} x_i x_j \quad (18)$$

3.2.2. Kriging model (KRG)

These models suggest estimating deterministic functions as:

$$y(x_j) = \mu + \varepsilon(x_j) \quad (19)$$

where, y is the function to be modeled, μ is the mean of the population, and ε is the error with zero expected value, and with a correlation structure that is a function of a generalized distance between the sets of design values. In this work we use a correlation structure (Sacks et al., 1989) given by:

$$\text{cov}(\varepsilon(x_i), \varepsilon(x_j)) = \sigma^2 \exp \left(- \sum_{h=1}^p \theta_h (x_i^h - x_j^h)^2 \right) \quad (20)$$

where, p denotes the number of dimensions in the set of design variables x ; σ , identifies the standard deviation of the population, and, θ_h is a parameter which is a measure of the degree of correlation among the data along the h direction.

Specifically, given a set of n input/output pairs (x, y) , the parameters, μ , σ , and θ are estimated such that the likelihood function is maximized (Sacks et al.,

1989). The model estimates for a new set of design values is given by:

$$\bar{y}(x) = \bar{\mu} + r^T R^{-1} (y - L\bar{\mu}) \quad (21)$$

where the line above the letters denotes estimates, r identifies the correlation vector between the new set of design values and the points used to construct the model, R is the correlation matrix among the n sample points, and L denotes an n -vector of ones.

The estimation variance is given by:

$$V(\bar{y}(x)) = \sigma^2 \left[1 - r^T R^{-1} r + \frac{(1 - L^T R^{-1} r)}{L^T R^{-1} L} \right] \quad (22)$$

3.2.3. Radial Basis Functions (RBF)

The radial basis functions have been developed for the interpolation of scattered multivariate data. The method uses linear combinations of m radially symmetric functions, $h(x)$, based on Euclidean distance or other such metric, to approximate response functions as,

$$y_i = \sum_{i=1}^m w_i h_i(x) + \varepsilon_i \quad (23)$$

where w are the coefficients of the linear combinations, h the radial basis functions and ε_i independent errors with variance σ^2 .

The flexibility of the model, its ability to fit many different functions, derives from the freedom to choose different values for the weights. The radial basis functions and any other parameter they might contain are fixed.

Radial basis functions are a special class of functions. Their main feature is that their response decreases (or increases) monotonically with distance from a central point. The centre, the distance scale, and the precise shape of the radial function are parameters of the model.

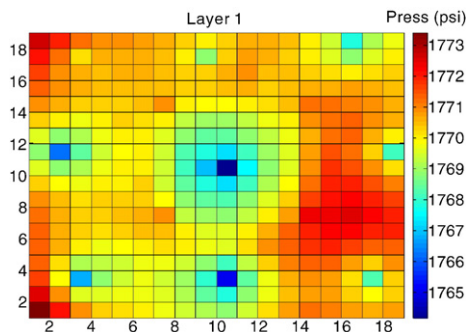


Fig. 3. Initial reservoir pressure distribution.

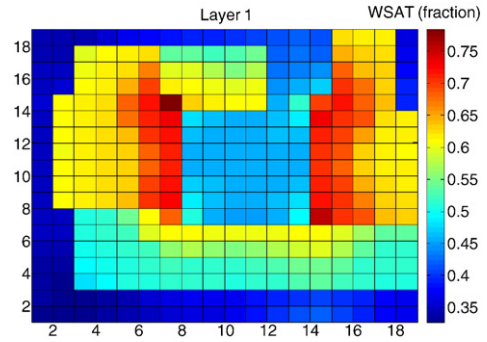


Fig. 4. Initial water saturation field.

A typical radial function is the Gaussian which, in the case of a scalar input, is

$$h(x) = \exp\left(-\frac{(x-c)^2}{r^2}\right) \quad (24)$$

its parameters are its centre c and its radius r . The response of the Gaussian RBF decreases monotonically with the distance from the centre, giving a significant response only in the center neighborhood.

A radial basis functions model can be expressed as,

$$y = Hw + \varepsilon \quad V(\varepsilon) = \sigma^2 \quad (25)$$

where H is the design matrix given by,

$$H = \begin{bmatrix} h_1(x_1) & h_2(x_1) & \cdots & h_m(x_1) \\ h_1(x_2) & h_2(x_2) & \cdots & h_m(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ h_1(x_n) & h_2(x_n) & \cdots & h_m(x_n) \end{bmatrix} \quad (26)$$

Similarly to the polynomial regression method, the Eq. (25) is solved by least squares to obtain the optimal weights,

$$\hat{w} = A^{-1} H^T y \quad (27)$$

where A^{-1} is the variance matrix and is given by,

$$A^{-1} = (H^T H)^{-1} \quad (28)$$

The variance estimator σ^2 of the error is given by,

$$\hat{\sigma}^2 = \frac{y^T P^2 Y}{\text{trace}(P)} \quad (29)$$

where P is the projection matrix,

$$P = I - HA^{-1}H^T \quad (30)$$

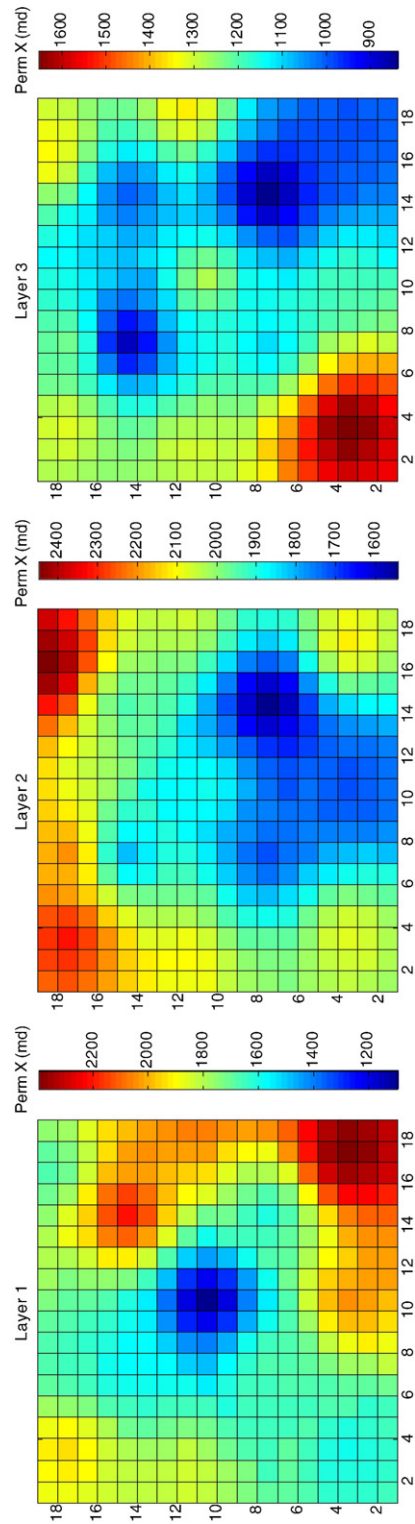


Fig. 5. Horizontal permeability field.

Table 2
Reservoir and fluid properties

Property	Value	Units
Reservoir depth	4150 (1265)	ft (m)
OOIP	395,427 (62,868)	bbls (m ³)
Oil viscosity	40	(cp)
Porosity	0.3	(fraction)
Average initial pressure	1770	(psi)
Well radius	0.49 (0.15)	ft (m)
Skin factor	0.0	(adim)
Water salinity	C _{Na}	0.0583 (meq/ml)
	C _{Ca}	0.0025 (meq/ml)

The model estimates for a new set of design values is given by,

$$y(x) = \mathbf{Z}^T \hat{\mathbf{w}} \quad (31)$$

where, \mathbf{Z} is a column vector with the radial basis functions evaluations,

$$\mathbf{Z} = \begin{Bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_c(x) \end{Bmatrix} \quad (32)$$

and the estimation variance is the variance of the prediction $\mathbf{z}^T \hat{\mathbf{w}}$ plus the error variance:

$$V(y) = V(\mathbf{Z}^T \hat{\mathbf{w}}) + V(\varepsilon) = (\mathbf{Z}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{Z} + 1) \frac{y^T \mathbf{P} y}{p-m} \quad (33)$$

3.2.4. Weighted adaptive model

This model suggests to estimate deterministic functions as:

$$y_{\text{WAM}}(x) = \sum_{i=1}^k \alpha_i(x) y_{\text{sur}} i(x) \quad (34)$$

where y_{WAM} is the weighted adaptive model, $y_{\text{sur}} i$ is the prediction of the surrogate model i , and α_i the

weight of the surrogate i , and k the number of surrogates. Note the adaptive nature of the model since the weights are a function of x .

Assuming unbiased and independent predictions, the unbiased weighted adaptive model has minimal variance when the weights are determined as follows (Bishop, 1995),

$$\alpha_i = \frac{\frac{1}{V(i)}}{\sum_{j=1}^k \frac{1}{V(j)}} \quad (35)$$

where $V(i)$ is the prediction variance of the i surrogates. In this case the individual surrogates are PRG, KRG and RBF ($k=3$).

4. Case study: Alkaline–Surfactant–Polymer flooding process

As previously stated the problem of interest is to find the contribution of design variables such as alkali, surfactant and polymer concentrations, and time of injection, to the total variability of the response of an oil recovery process by ASP injection. The ranges of the design variables are presented in Table 1. The cumulative oil production is calculated at 487 days.

As illustrated in Fig. 2, the ASP flooding pilot has an inverted five-spot pattern and a total of 13 vertical wells, 9 producers and 4 injectors. The reservoir is at a depth of 4150 ft, has an average initial pressure of 1770 psi, and the porosity is assumed to be constant throughout the reservoir and equal to 0.3. The numerical grid is composed of $19 \times 19 \times 3$ blocks in the x , y and z directions. Figs. 3, 4, and 5 show the initial reservoir pressure, initial water saturation and horizontal permeability fields, respectively. The OOIP is 395,427 bbls, the crude oil viscosity is 40 cp, the initial brine salinity is 0.0583 meq/ml and the initial brine divalent cation

Table 3
Injection scheme

Slug	Injection time (days)	PV	Component concentration								
			C _w ¹	C _{surf} ¹	C _{pol} ²	C _{Cl} ³	C _{Ca} ³	C _{Mg} ³	C _{CO₃} ³	C _{Na} ³	C _{H+} ³
Polymer preflush	26	0.05	1.0	0	0.0974	0.015667	0.0019	0.004774	0.009122	0.01461	111.0034
AS preflush	25	0.1	0.99574	0.00426	0	0.07168	0.0034	0.0067	0.3339	0.52517	111.0767
ASP slug	715	0.41	0.99637	0.00363	0.0974	0.04948	0.0067	0.00831	0.3351	0.3929	111.839
Polymer drive	50	0.5	1.0	0	0.05	0.03586	0.00665	0.00132	0.0164	0.09	111.0034
Postflush	275	1.0	1.0	0	0	0.0135	0.00185	0.004774	0.008	0.0146	111.0034

1. Conc. unit = Vol. fract.

2. Conc. unit = wt.%.

3. Conc. unit = meq/ml.

concentration is 0.0025 meq/ml. A summary of the reservoir and fluid properties is presented in the Table 2. The injection scheme is described in Table 3. This is the reference configuration whose details can be found in the sample data files of the UTCHEM program.

Three flowing phases and eleven components are considered in the numerical simulations. The phases are water, oil and microemulsion, while the components are water, oil, surfactant, polymer, chloride anions, divalent cations (Ca^{2+} , Mg^{2+}), carbonate, sodium, hydrogen ion, and oil acid. The ASP interactions are modeled using the reactions: *in situ* generated surfactant, precipitation and dissolution of minerals, cation exchange with clay and micelle, and chemical adsorption.

Note the detailed chemical reaction modeling, and the heterogeneous and multiphase petroleum reservoir

under consideration. A typical run time for the numerical simulations associated with the reference configuration using UTCHEM is 853 s of CPU time in a PC with an AMD Athlon 64 2.0 GHz Processor and 960 MB of RAM.

5. Results and discussion

After validating the implementation of the solution methodology using the well known Lagendre polynomials benchmark case (Archer et al., 1997), it is used to solve the problem of interest as described in the case study.

Fig. 6 shows the results of the global sensitivity analysis for each of the surrogate models considered in this work with the design variables denoted as x_1 :

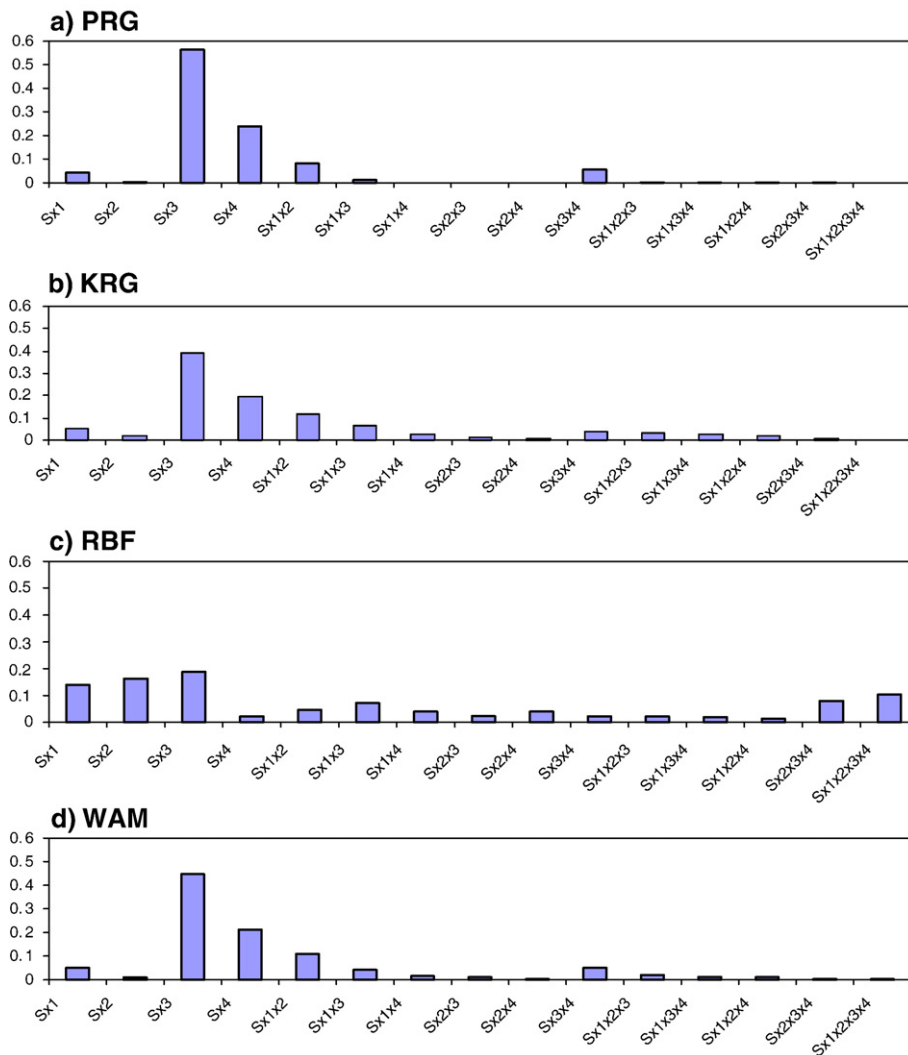


Fig. 6. Global sensitivity indices, a) PRG, b) KRG, c) RBF, d) WAM (ASP flooding case study).

alkaline concentration, x_2 : surfactant concentration, x_3 : polymer concentration and x_4 : slug size. The rank of the individual indices ($S_3 > S_4 > S_1 > S_2$) (for most surrogates but RBF) indicates that polymer concentration (x_3) has the greatest individual sensitivity index, follow by the slug size (x_4) and alkaline concentration (x_1), while the surfactant concentration (x_2) has a nearly insignificant individual sensitivity index.

The RBF model was obtained using the Matlab functions provided by Orr (1999), with general cross validation for model selection. The RBF model obtained has a number of centers almost equal to the number of points and provided an overfitted approximation to the function of interest. As a result, the RBF model exhibited a high prediction variance, did not provide a good approximation of the ASP flooding process, and had little influence on the results corresponding to the WAM model so it was not included in the rest of the analyses. RBF, in general, though, can avoid overfitting problems by using regularization methods in the modeling process. Note that the multiple surrogates approach and weighted average model for global sensitivity analyses is not significantly influenced by modeling problems of particular surrogates.

The most important interaction is that of alkaline and surfactant concentrations (x_1x_2), with a much greater contribution to the total variability than those of the individual variables x_1 and x_2 . While the polymer

concentration and slug size (x_3x_4) interaction can be seen as significant its contribution to the total variability is small compared to that of the individual variables (see Fig. 7).

Fig. 8 shows the relative value of the total sensitivity indices. The polymer concentration (x_3) has the greatest percentage contribution to the variability of the response of all the surrogates models considered (PRG: 55%, KRG: 39%, WAM: 44%). The slug size (x_4) of the ASP solution also has an important contribution to the variability of the response (PRG: 27%, KRG: 22%, WAM: 23%), followed by the alkali and surfactant concentration. In summary, the rank of the total sensitivity indices provided for most of the surrogates was $S_3^{\text{total}} > S_4^{\text{total}} > S_1^{\text{total}} > S_2^{\text{total}}$. While a similar rank was found when only the individual contributions were considered, the latter case did not account for variable interactions. For screening purposes total sensitivity indices should be considered and the design variable candidates to be fixed are those with lowest total sensitivity index (i.e., surfactant concentration, x_2).

While in general it is not possible to anticipate in oil recovery processes what are the most important design variables and interactions, having found polymer concentration (x_3) and alkaline and surfactant concentration (x_1x_2) as the most important design variable and interaction, respectively, is consistent with the facts that: i) polymer concentration is critical to attain a proper

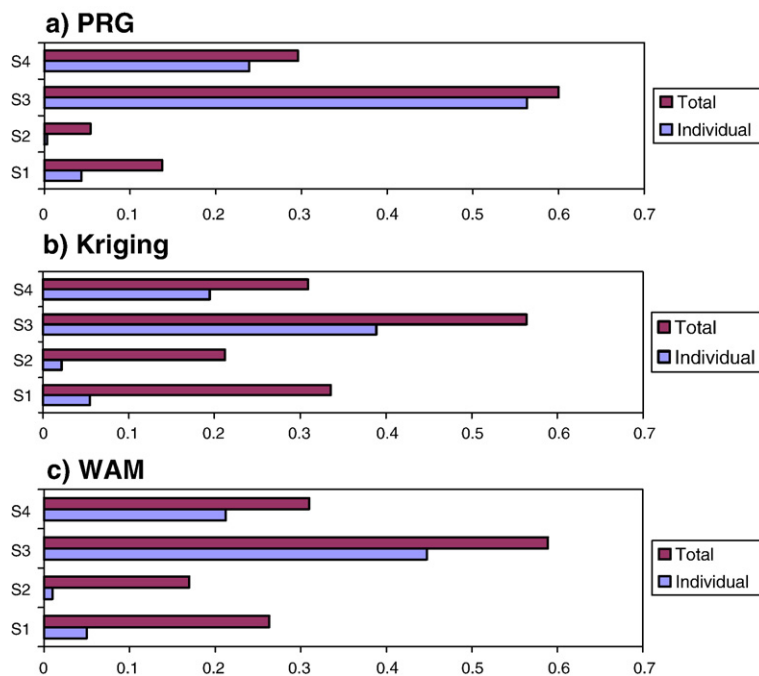


Fig. 7. Individual and total sensitivity indices. a) PRG, b) KRG, c) WAM (ASP flooding case study).

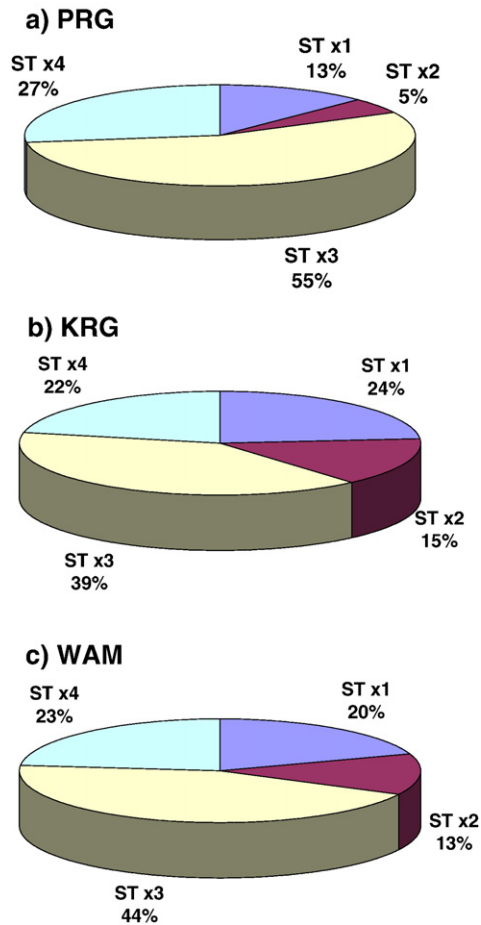


Fig. 8. Relative value of the total sensitivity indices. a) PRG, b) KRG, c) WAM (ASP flooding case study).

volumetric swept efficiency (Lake, 1989) and ii) there is a synergy between alkaline and surfactant to lower the interfacial tension between oil and water and consequently enhance the displacement of oil (Wang and Gu, 2005).

6. Conclusions

- A methodology for global sensitivity analysis, a critical step for the optimal design and control of oil recovery processes (e.g., ASP flooding), has been proposed. It involves the coupled execution of fast surrogate models constructed from field scale numerical simulation data with a global sensitivity analysis algorithm (Sobol's method). The reservoir numerical simulations are conducted using UTCHEM from the University of Texas at Austin. Furthermore, the surrogate-modeling uncertainty (different models can

provide reasonable approximations to the available data) is addressed through the use of multiple surrogates, namely, Polynomial Regression, Kriging, Radial Basis Functions, and a Weighted Average Model.

- After proper validation, the proposed approach was evaluated using an ASP flooding pilot in a heterogeneous and multiphase petroleum reservoir with an inverted five-spot pattern and a total of 13 vertical wells, 9 producers and 4 injectors, three flowing phases and eleven components, using a detailed chemical reaction modeling. The contribution of each of the design variables (alkaline, surfactant and polymer concentrations and slug size) and their interactions to the total variability were assessed finding that: the polymer concentration was the most important variable explaining more than half of the total variability, and that the interactions were not significant except for that of alkaline and surfactant concentrations. These results are consistent with the need to achieve a proper volumetric sweep efficiency and the well known synergy between alkaline and surfactant to enhance oil recovery.
- The complex heterogeneous multiphase and multi-component nature of most oil recovery processes (e.g., thermal, chemical and miscible gas flooding) does not allow anticipating the most important variables and interactions. A global sensitivity analysis based on surrogate models can provide an answer to this issue and be useful in other related problems such as identifying main factors, variable selection for maximum uncertainty reduction, main regions of interest in the context of optimization, and model validation.

Nomenclature

α	Weights of the WAV model
\hat{B}	Estimated parameters
ε	Error
μ	Mean of the population
θ	Correlation parameter
σ	Standard deviation
D	Total variance
$D_{i,...,is}$	Partial variance
f	Objective function
I	Identity matrix
k	Number of surrogates
KRG	Kriging model
L	n -vector of ones
p	Number of variables
r	Correlation vector
R	Correlation matrix

PRG	Polynomial Regression model
S_i	Individual sensitivity index
S_i^{total}	Total sensitivity index
V	Variance
x	The design variables
X	Constraints set
WAV	Weighted Average Model
y	Function of interest
z	Prediction variables
Z	Matrix of the prediction variables

Subscripts

surr Surrogate model

Acknowledgements

The authors gratefully acknowledge the financial support provided to this project by Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT) under Grant F-2005000210, Venezuela. We also thank the Center for Petroleum and Geosystems Engineering of The University of Texas at Austin for providing the UTCHEM compositional simulator.

References

- Archer, G., Saltelli, A., Sobol, I., 1997. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation* 58, 99–120.
- Balabanov, V.O., Haftka, R.T., Grossman, B., Mason, W.H., Watson, L.T., 1998. Multidisciplinary response model for HSCT wing bending material weight. AIAA Paper 98-4804 Presented at the 1998 7th AIAA/USAF/NASA/ISSMO Symp. on Multidisciplinary Anal. and Optim., St. Louis, MO.
- Bishop, C.M., 1995. *Neural Networks for Pattern Recognition*, vol. 364–371. Oxford University Press, p. 368.
- Bhuyan, D., Lake, L.W., Pope, G.A., 1990. Mathematical Modelling of High-pH Chemical Flooding. SPE paper 17398. SPE Reservoir Engineering 213–220.
- Camilleri, D., Engelsens, S., Lake, L.W., Lin, E.C., Ohno, T., Pope, G.A., Sepehrmoori, K., 1987. Description of an improved compositional micellar/polymer simulator. Paper SPE 13967. SPE Reservoir Engineering 427–432.
- Clark, S.R., Pitts, M.J., Smith, S.M., 1988. Design and application of an alkaline-surfactant-polymer recovery system to the West Kiehl. Paper SPE 17538 Presented at the SPE Rocky Mountain Regional Meeting, Casper, WY.
- Craig, K.J., Stander, N., Dooge, D.A., Varadappa, S., 2002. MDO of automotive vehicles for crashworthiness using response surface methods. AIAA paper 2002-5607 Presented at the 9th AIAA/ISSMO Symp. on Multidisciplinary Anal. and Optim., Atlanta, GA.
- Delshad, M., Pope, G.A., Sepehrmoori, K., 2005. A Framework to Design and Optimize Chemical Flooding Processes. DOE Report DE-FC-26-03NT15412.
- Demin, W., Jiecheng, Ch., Junzheng, W., Zhenyu, Y., Hongfu, L., 1999. Summary of asp pilots in Daqing oil field. Paper SPE 57288 presented at the Asia Pacific Improved Oil Recovery Conference, Kuala Lumpur, Malaysia.
- Dosher, T.M., Wise, F.A., 1976. Enhanced oil recovery potential. An estimate. Paper SPE 5800. *Journal of Petroleum Technology* 575.
- Draper, N.R., Smith, H., 1966. *Applied Regression Analysis*. John Wiley and Sons, Inc., p. 128.
- Frey, H., Patil, S., 2002. Identification and review of sensitivity analysis methods. *Risk Analysis* 22 (3), 553–578.
- Giunta, A.A., Balabanov, V., Burgee, S., Grossman, B., Haftka, R.T., Mason, W.H., Watson, L.T., 1997. Multidisciplinary optimisation of a supersonic transport using design of experiments theory and response surface modelling. *Aeronautical Journal* 101 (1008), 347–356.
- Goel, T., Haftka, R.T., Shyy, W., Queipo, N.V., in press. Ensemble of surrogates. *Structural and Multidisciplinary Optimization Journal*.
- Hernández, C., Chacón, L., Anselmi, L., Baldonado, A., Qi, J., Phillip, C., Pitts, M.J., 2001. ASP system design for an offshore application in the La Salina Field, Lake Maracaibo. Paper SPE 69544 presented at the SPE Latin American and Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina.
- Jacque, J., Laverne, C., Devictor, N., 2004. Sensitivity analysis in presence of model uncertainty and correlated inputs. *Special Issue of Reliability Engineering and System Safety, SAMO*.
- Jin, R., Chen, W., Simpson, T.W., 2001. Comparative studies of meta-modeling techniques under multiple modeling criteria. *Structural and Multi-Disciplinary Optimization* 23, 1–13.
- Kurtaran, H., Eskandarian, A., Marzougui, D., Bedewi, N.E., 2002. Crashworthiness design optimization using successive response surface approximations. *Computational Mechanics* 29, 409–421.
- Lake, L.W., 1989. *Enhanced Oil Recovery*. Prentice Hall, Englewood Cliffs, NJ, pp. 410–413.
- Li, W., Padula, S., 2004. Approximation methods for conceptual design of complex systems. In: Chui, C., Neaumtu, M., Schumaker, L. (Eds.), *Eleventh International Conference on Approximation Theory*. May 18–24, 2004.
- Mack, Y., Goel, T., Shyy, W., Haftka, R.T., Queipo, N.V., 2005. Multiple surrogates for shape optimization for bluff-body facilitated mixing. 43rd AIAA Aerospace Sciences Meeting and Exhibit, 10 Jan–13 Jan 2005. (Reno, Nevada) AIAA-2005-0333.
- Manrique, E., De Carvajal, G., Anselmi, L., Romero, C., Chacón, L., 2000. Alkali/surfactant/polymer at VLA 6/9/21 Field in Maracaibo Lake: experimental results and pilot project design. Paper SPE 59363 presented at the SPE/DOE Improved Oil Recovery Symposium, Tulsa, Oklahoma.
- McKay, M., Beckman, R., Conover, W., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239–245.
- Meyers, J.J., Pitts, M.J., Wyatt, K., 1992. Alkaline-surfactant-polymer flood of the West Kiehl, Minnelusa Unit. Paper SPE/DOE 24144 presented at SPE/DOE Eighth Symposium on Enhanced Oil Recovery, Tulsa, Oklahoma.
- Orr, M., 1999. *Matlab Functions for Radial Basis Functions Networks*. Institute for Adaptive and Neural Computation, Division of Informatics, Edinburgh University, Edinburgh EH8 9LW, Scotland, UK. www.anc.ed.ac.uk/rbf/.
- Pope, G.A., Nelson, R.C., 1978. A chemical flooding compositional simulator. Paper SPE 6725, Society of Petroleum Engineers J., p. 18.
- Qi, Q., Hongjun, G., Dongwen, L., Ling, D., 2000. The pilot test of ASP combination flooding in Karamay oil field. Paper SPE 64726 presented at the SPE International Oil and Gas Conference and Exhibition, Beijing, China.

- Queipo, N.V., Goicochea, J., Pintos, S., 2002a. Surrogate modeling-based optimization of SAGD processes. *Journal of Petroleum Science and Engineering* 35/1–2, 83–93.
- Queipo, N.V., Verde, A., Canelón, J., Pintos, S., 2002b. Efficient global optimization of hydraulic fracturing designs. *Journal of Petroleum Science and Engineering* 35/3–4, 151–166.
- Queipo, N.V., Hafka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K., 2005. Surrogate-based analysis and optimization. *Progress in Aerospace Sciences* 41, 1–28.
- Sacks, J., Welch, W., Mitchell, T., Wynn, H., 1989. Design and analysis of computer experiments. *Statistical Science* 4, 409–435.
- Saltelli, A., Tarantola, S., 2002. On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. *Journal of American Statistical Association* 97 (459), 702–709.
- Sanchez, E., Pintos, S., Queipo, N.V., 2006. Toward an optimal ensemble of kernel-based approximations with engineering applications. *Proceedings of the IEEE World Congress on Computational Intelligence*, Vancouver, BC, Canada, July 16–21, 2006.
- Simpson, T.W., Peplinski, J.D., Koch, P.N., Allen, J.K., 2001. Meta-models for computer based engineering design: survey and recommendations. *Engineering with Computers* 17, 129–150.
- Sobol, I., 1993. Sensitivity estimates for non-linear mathematical models. *Mathematical Modeling and Computational Experiments* 1 (4), 407–414.
- Vargo, J., Turner, J., Vergnani, B., Pitts, M., Wyatt, K., Surkalo, H., Patterson, D., 1999. Alkaline–surfactant/polymer flooding of the Cambridge Field. Paper SPE 55633 presented at the SPE Rocky Mountain Regional Meeting, Gillette, Wyoming.
- Wang, W., Gu, Y., 2005. Experimental studies of the detection and use of produced chemicals in alkaline/surfactant/polymer floods. *SPE Reservoir Evaluation and Engineering* 8 (5), 362–371.
- Wu, W., 1996. Optimum design of field-scale chemical flooding using reservoir simulation. PhD Thesis, The University of Texas at Austin.
- Zhijian, Q., Yigen, Z., Xiansong, Z., Jialin, D., 1998. A successful ASP flooding in Gudong oil field. Paper SPE 39613 presented at the SPE/DOE Improved Oil Recovery Symposium, Tulsa, Oklahoma.
- Zerpa, L.E., Queipo, N.V., Pintos, S., Salager, J.L., 2005. An optimization methodology of alkaline–surfactant–polymer flooding processes using field scale numerical simulation and multiple surrogates. *Journal of Petroleum Science and Engineering* 47, 197–208.