Application of Data-Driven Techniques to SAGD and Solvent-Aided Methods

by

Seyide Olumuyiwa Hunyinbo

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

PETROLEUM ENGINEERING

Department of Civil and Environmental Engineering
University of Alberta

© Seyide Olumuyiwa Hunyinbo, 2021

Abstract

The warm vaporized solvent injection process has been proposed as a more environmentally friendly alternative to steam-based technologies for bitumen recovery. The process typically involves injecting heated solvent vapor into a horizontal injector; the solvent condenses and dissolves into bitumen, while the diluted oleic phase would flow towards a horizontal producer. An optimization process is important because of its potential reduction of solvent loss to the reservoir and energy requirements while maximizing bitumen recovery. Hence, this research proposes a workflow for optimizing the multiple conflicting performance objectives associated with the warm vaporized solvent injection process. Specific considerations phase behavior constraints, multiple realizations of reservoir heterogeneity, and computational efficiency are considered. It is expected that this workflow can be readily integrated into the design and decision-making processes in reservoir management, especially where multiple geostatistical realizations are involved.

Apart from performing automated optimization and quantification of geological uncertainties and requiring lower computational effort compared to reservoir simulation, data-driven models offer better accuracy than semi-analytical or proxy models based on Butler's equation. Hence, this thesis also presents another workflow for real-time forecasting, uncertainty assessment of SAGD profiles, and optimization of steam allocation using a real SAGD dataset which includes operational data, geological, and well design parameters. The workflow includes the development of a predictive model using the random forest algorithm, and clustering, Bayesian updating, Monte Carlo sampling, and

genetic algorithm for the real-time prediction of SAGD injection and production data. This workflow can update predictions in real-time, perform uncertainty quantification of the forecasts, and optimize steam allocation, making it a practical tool for development planning and field-wide optimization.

Preface

This thesis is the original work of Seyide Olumuyiwa Hunyinbo. No part of this thesis has been previously published.

Chapter 4 has been published as Seyide Hunyinbo, Zhiwei Ma, and Juliana Y. Leung, "Incorporating phase behavior constraints in the multi-objective optimization of a warm vaporized solvent injection process," *Journal of Petroleum Science and Engineering*, vol. 205. I was responsible for the conceptualization and methodology, formal analysis, and writing. Zhiwei Ma assisted with the conceptualization and methodology. Professor Juliana Y. Leung was responsible for the supervision, project administration, funding acquisition, conceptualization, data curation, reviewing and editing.

Chapter 5 of this thesis was submitted for publication as Seyide Hunyinbo, Zhiwei Ma, and Juliana Y. Leung, "Determination of the Optimal Operational Parameters for a Warm Vaporized Solvent Injection Process, Using an Efficient Multi-Objective Optimization Workflow that Considers Reservoir Heterogeneity Uncertainty" in the proceedings of the *SPE Annual Technical Conference and Exhibition (ATCE)* to be held on 21–23 September, 2021 in Dubai, UAE. I was responsible for the conceptualization, methodology, and writing. Zhiwei Ma assisted with the conceptualization and methodology. Professor Juliana Y. Leung was responsible for the supervision, funding acquisition, conceptualization, and editing.

Chapter 6 of this thesis was submitted for publication as Seyide Hunyinbo, Prince Azom, Amos Ben-Zvi, and Juliana Y. Leung, "A Machine Learning Approach to Real-time

Uncertainty Assessment of SAGD Forecasts and the Optimization of Steam Allocation" in the proceedings of the *SPE Canadian Energy Technology Conference and Exhibition* to be held on 16–17 March, 2022 in Calgary, Alberta, Canada. I was responsible for the conceptualization, methodology, and writing. Prince Azom, Amos Ben-Zvi, and Juliana Y. Leung were responsible for the supervision, funding acquisition, conceptualization, methodology, review, and editing.

For all the Chapters, I was responsible for literature review, code development, data analysis and writing.

This work is dedicated to God, to Oluwatoyin, my beloved mother and Seji, my late father.

Acknowledgments

First, I would like to thank my awesome supervisor, Dr. Juliana Leung, who has been a huge inspiration for me in terms of her professionalism and attitude to work. She always motivated me and encouraged me to never loose sight of my goal. Her endless support was very critical to the successful completion of my thesis. I would also like to show gratitude to Prince Azom and Amos Ben-Zvi for believing in my abilities and allowing me to collaborate with them at Cenovus Energy Inc.

My appreciation also goes to my defense committee members; Dr. Hassan Dehghanpour, Dr. Yong Li, and Dr. Yashar Pourrahimian, for their guidance and advice. I am also grateful to my course mates, family, and friends; Semande Hunyinbo, Longtong Dafyak Olanrewaju Adebola, Adejuwon Kehinde, Fortune Amadi, Adeshina Aina, David Nwankwo, Taiwo Dada, Temitope Okesanya, Francis Nwabia, Chuanyao Zhong, Luis Coimbra, Najmudeen Sibaweihi, Phuong Nguyen, Jose Leonardo Guevara, and my research teammates for their moral support.

Finally, I would like to express my appreciation to my country Nigeria, who supported my program financially through the Petroleum Technology Development Fund (PTDF), the University of Alberta Future Energy Systems (FES), the Mitacs Accelerate program, Cenovus Energy, the Alberta Graduate Excellence, and the Donald Lougheed Engineering Graduate scholarships. I thank the Computer Modeling Group (CMG) and MathWorks™ for providing the academic licenses for STARS™ and WINPROP™, and MATLAB™, respectively.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANN	=	Artificial Neural Network
AHTMS-VAPEX	=	Azeotropic High Temperature Multicomponent Solvent Vapor Extraction
ASP	=	Alkaline-Surfactant-Polymer
cSOR	=	Cumulative Steam-to-Oil Ratio
cSolOR	=	Solvent Retained-Oil Ratio
DAS	=	Distributed Acoustic Sensing
DTS	=	Distributed Temperature Sensing
EOR	=	Enhanced Oil Recovery
ESP	=	Electric Submersible Pump
ES-SAGD	=	Expanding Solvent Steam Assisted Gravity Drainage
FCD	=	Flow Control Device
GA	=	Genetic Algorithm
GHG	=	Greenhouse Gas
GSLIB	=	Geostatistical Software Library

HTMS-VAPEX	=	High Temperature Multicomponent Solvent Vapor Extraction
MCS	=	Monte Marlo Simulation
MOEA	=	Multi-Objective Evolutionary Algorithms
MOGA	=	Multi-Objective Genetic Algorithm
MOO	=	Multi-Objective Optimization
ML	=	Machine Learning
MSE	=	Mean Squared Error
NMSE	=	Normalized Mean Squared Error
NPGA	=	Niched Pareto Genetic Algorithm
NPV	=	Net Present Value
NSGA	=	Non-Dominated Sorting Genetic Algorithm
OOIP	=	Original Oil in Place
PAES	=	Pareto-Archived Evolution Strategy
PCA	=	Principal Component Analysis
P_R	=	P10, P50 or P90
PSO	=	Particle Swamp Optimization

RF	=	Recovery Factor
RSM	=	Response Surface Methodology
RVEA	=	Reference Vector Evolutionary Algorithm
SISIM	=	Sequential Indicator Simulation
SAGD	=	Steam Assisted Gravity Drainage
SAP	=	Solvent Aided Process
SA-SAGD	=	Solvent Assisted Steam Assisted Gravity Drainage
SED	=	Squared Euclidean Distance
SolOR	=	Solvent Retained-Oil Ratio
SPEA	=	Strength Pareto Evolutionary Algorithm
SVD	=	Singular Value Decomposition
VAPEX	=	Vapor Extraction
VEGA	=	Vector Evaluated Genetic Algorithm

List of Symbols

a_j^{k+1} = Activation of node j in layer k+1

b = Bias node for neural network

BHP_{inj} = Bottom-hole injection pressure

BHT_{inj} = Bottom-hole injection temperature

$cSOR_{avg}$ = Average cumulative steam-to-oil ratio across SAGD well-pairs

d = Annual compound discount rate

e_{P_R} = Deviation of actual production data from P10, P50 or P90 predictions from the random forest model predictions

$E\{f(x)\}$ = Weighted average value of the objective function

$f(\vec{v})$ = Objective function

$f(x, P_{low-case})$ = Objective function for low-case realization

$f(x, P_{mid-case})$ = Objective function for mid-case realization

$f(x, P_{high-case})$ = Objective function for high-case realization

f = Objective function

f_m^{i-1} = m^{th} objective function of the neighboring individual $i-1$

f_m^{i+1} = m^{th} objective function of the neighboring individual $i + 1$

f_m^{min} = Population-minimum value of the m^{th} objective function

f_m^{max} = Population-maximum value of the m^{th} objective function

$f(x)$ = Objective function

$h_k(x)$ = Equality constraint

H_{inj} = Cumulative enthalpy of the injected solvent

H_{inj}^{max} = Maximum cumulative enthalpy of the injected solvent

$H_{inj}^{normalized}$ = Normalized enthalpy

g = Activation or transfer function

$g_j(x)$ = Inequality constraint

i_{rank} = Rank of individual i

$i_{distance}$ = Crowding distance of individual i

J = mean squared Euclidean distance

j_{rank} = Rank of individual j

$j_{distance}$ = Crowding distance of individual j

M_o = Molecular weight of oil

M_{sol} = Molecular weight of solvent

N = Population size

N_{P_R} = Number of P_R type well-pairs

N_{wells_total} = Total number of well-pairs used for training

N_{wells} = Total number of well-pairs for steam optimization

P_{WCS} = Price per barrel for Western Canadian Select oil grade

$P_{inj_{m,n}}$ = Injector bottom-hole pressure for well n at month m

$[P_{inj}]$ = Decision matrix for GA minimization problem

$P(P_{R_0}|P_{R_t})$ = Posterior probability that a well is a P_R type well given that it as a P_R type well at day t

$P(P_{R_t}|P_{R_0})$ = Probability that a well is a P_R type well after production up till day, t

$P(P_{10_t}|P_{10_0})$ = Probability that a well is a P_{10} type well after production up till day, t

$P(P_{50_t}|P_{50_0})$ = Probability that a well is a P_{50} type well after production up till day, t

$P(P_{90_t}|P_{90_0})$ = Probability that a well is a P_{90} type well after production up till day, t

$P(P_{10_0})$ = Probability of that a well is a P10 type well at day, 0

$P(P_{50_0})$ = Probability of that a well is a P50 type well at day, 0

$P(P_{90_0})$ = Probability of that a well is a P90 type well at day, 0

P_{min} = Dew point pressure at 30°C (minimum injection temperature for this study).

P_{max} = Upper pressure limit

P = Random number which represents bottom-hole injection pressure

P_{val} = Injection pressure value corresponding to P

P_{inj} = Bottom-hole injection pressure

P_t = Parent population

$q_{o_{P_R}}^{predicted,i}$ = P10, P50 or P90 oil production prediction at timestep i

$q_o^{actual,i}$ = Actual oil production data at timestep i

Q_{o_i} = Total oil production (in m^3) for month i

Q_{s_i} = Total steam injection (in m^3) for month i

Q_t = Offspring population

R_t = Combined population

r = Monthly discount rate

S = Entropy

S_o = Oil saturation

T = Random number which represents injection temperature

T_{val} = Injection temperature value corresponding to T

T_{sat}= Saturation temperature at a pressure P_{val}

T_{inj} = Bottom-hole injection temperature

u = Decision vector

U_{inj} = Internal energy of the injected solvent

v or $\vec{v}$ = Decision vector

$V_{solvent}^{inj}$ or $V_{C_3}^{inj}$ = Total volume of gaseous C_3 is injected

$V_{solvent}^{prod}$ or $V_{C_3}^{prod}$= Total volume of C_3 recovered at the surface

V_{oil}^{prod} = Total volume of bitumen produced

V_{inj} = Volume of fluid injected

$w_{j,i}^{k}$ = Weight of node i in the neural network layer k

x= Decision vector

$x^{(i)}$ = *i-th* observation

X= Decision vector or Input data

X_{steam} = Cost of steam (per barrel of oil)

x_i^{lower} = Lower bound for decision variable x_i

x_i^{upper} = Upper bound for decision variable x_i

x_i^k = Value of node i in layer k

x_{C3} = Propane mole fraction

x_{oil} = Oil mole fraction

$\bar{x}_{sol}$ = Solvent volume fraction

Z = Mean adjusted data

ΔT = Temperature range

σ_Z = Covariance matrix of mean adjusted data, Z

α = Normalization constant for the distribution $P\left(P_{R_0}|P_{R_t}\right)$

$\mu^{(j)}$ = *j-th* cluster centroid

$\vec{u}$ and $\vec{v}$ = Decision vectors

$\prec_c$ = Crowded comparison operator

μ_{c1} = Chemical potential for C_1

μ_{c3} = Chemical potential for C_3

n_{c1} = Number of moles of C_1

n_{c3} = Number of moles of C_3

ρ_{oil} = Oil density

ρ_{sol} = Solvent density

β = Coefficients for the factorial experimental design regression model

ε = Random error term

Introduction

Most of Canada's oil sands are located in Western Canada, notably in the provinces of Alberta and Saskatchewan, and the heavy oil/bitumen resources can be produced through in-situ thermal and/or solvent-based methods. Due to the high viscosity of bitumen, thermal and solvent based methods are implemented to lighten bitumen viscosity through heat transfer (in the form of steam or hot water injection) and/or mass transfer (injection of light hydrocarbon solvents or carbon-dioxide). The most common thermal extraction technique is steam-assisted gravity drainage (SAGD) method, which involves the injection of high-quality steam into the reservoir to thermally mobilize viscous heavy oil or bitumen. The recent challenges that are associated with SAGD include high operating costs, technical constraints, water consumption, and greenhouse gas (GHG) emissions, hence non-thermal methods (e.g., solvent-aided recovery processes) have been proposed as promising alternatives.

Solvent-based and steam-solvent hybrid methods often utilize a similar well-pair configuration as the SAGD process. For example, the warm Vapor Extraction (VAPEX) process, which involves the injection of vaporized solvent between the dew point to superheated conditions in-situ, such that it can dissolve into the bitumen. Compared to the SAGD, solvent-based techniques usually cost more to operate, hence the need for optimization of the pertinent operational constraints using techniques such as the genetic algorithm (GA)., however, one of its limitations is slow computation time.

This research focuses on three major workflows which consists of: (1) multi-objective optimization of the warm VAPEX process assuming homogeneous reservoir while considering

phase behavior constraints; (2) is an expansion of (1) to consider a heterogenous reservoir; (3) real-time forecasting and optimization of real-world SAGD injection and production data. (1) and (2) are performed using the Non-dominated Sorting Genetic Algorithm II (NSGA-II) proposed by Deb et al. (2000). Chapter 3 elaborates the problem statement and objectives of this research.

A practical workflow for optimizing various design variables associated with solvent-based bitumen recovery processes is important. Therefore, a workflow, which optimizes multiple conflicting objective functions for the warm VAPEX process assuming homogeneous reservoir, is proposed in Chapter 4. The addition of non-condensable gas (methane) into the solvent (propane) is examined. The design variables are propane mole fraction, injection pressure and temperature, and the objective functions (i.e., oil recovery factor, solvent retained-to-oil ratio, and energy consumption), are defined. The workflow combines experimental design, proxy modeling through the artificial neural network (ANN), and NSGA-II to estimate the optimal decision variables. The ANN technique is incorporated to reduce the computational costs associated with reservoir simulations. Specific considerations including phase behavior constraints and computational efficiency are examined and incorporated. It is anticipated that this workflow can be readily integrated into the design and decision-making processes in reservoir management.

Chapter 5 elaborates on a robust multi-objective optimization workflow, similar to the one in Chapter 4, but for a more practical scenario of the warm VAPEX process involving geostatistical realizations of reservoir heterogeneities. To account for geologic uncertainty, a separate proxy model is constructed for each realization, and for each objective function, the minimum, the weighted mean, and the maximum over these realizations are considered. The developed approach can identify a set of optimal design variables in a computationally efficient manner. The presence

of shale baffles or barriers can affect the performance of the warm vaporized solvent injection process.

In Chapter 6, a novel algorithm for real-time forecasting and optimization (based on the Net Present Value (NPV)) of SAGD process is proposed. This machine learning framework can update predictions in real-time, be applied for the quantification of the uncertainties associated with the forecasts, and optimize steam allocation, making it a practical tool for development planning and field-wide optimization.

Since this is a paper-based thesis, Chapters 4, 5, and 6 consist of literature review and summary. Chapter 7 highlights the general conclusions and recommendations from this work.

Chapter 2 Problem Statement and Research Objectives

For solvent-based bitumen extraction processes, trade-offs exist between its performance objectives such (e.g., solvent loss, recovery factor, energy requirement), therefore requiring the multi-objective optimization (MOO) framework. Also, it is imperative to incorporate the phase behavior/thermodynamic properties of non-condensable gases in a MOO problem since these gases are usually present in solvents. Usually, since it is impossible to determine the actual distribution of reservoir heterogeneities, reservoir models are represented with multiple realizations of properties (e.g., low-case, mid-case, and high-case) using stochastic methods. The uncertain distribution of the reservoir properties can make optimization to be challenging, therefore the optimal solutions may be identified by considering a worst-case scenario, best case scenario or simply by finding the average of a performance objective over the realizations. However, the process of determining the optimal decision variables using reservoir modeling and sensitivity analysis is computationally inefficient and less effective, hence the need for the integration of data-driven or machine learning (ML) technique.

Also, a ML-based approach for real-time forecasting, uncertainty quantification and optimization of the SAGD method, while considering operational and reservoir uncertainty is uncommon, and the use of real field data for this data-driven approach is not in public domain. Although, semi-analytical models (Dehdari and Dong, 2017) have been proposed for SAGD production forecasting, they have not been widely adopted because of the difficulty in integrating operational data into the models.

Therefore, the objective of this research is to create a ML-based workflow that includes numerical reservoir modeling, proxy modeling and a Pareto-based MOO algorithm for the optimization of a solvent-based process. Another data-driven workflow which uses actual field dataset for dynamic forecasting, uncertainty quantification, and optimization of steam allocation during SAGD operation is also proposed.

Chapter 3 Methodology

Several techniques exist in searching for the optimal decision variable for a constrained optimization problem, which may be expresed in terms of minimizing or maximizing an objective function *f(x)* of decision vector *x = (*x_1*,* x_2*, ...,* x_N*)* whose values limited to the bounds $x_i^{lower} \leq x_i \leq x_i^{upper}$, and are constrained to meet a set of constraints including equality $h_k(x) = 0$ and inequality constraints $g_j(x) \geq 0$. However, an optimization problem may be expressed such that there are no constraints and no bounds for *x,* and this is referred to an unconstrained optimization problem.

Optimization problems can also be categorized based on *x, f(x),* $h_k(x)$ and $g_j(x)$. For example, single-variable and multi-variable problems are unconstrained problems with one-dimensional *x* and multi-dimensional *x* respectively. An optimization problem comprising of a linear $h_k(x)$ and $g_j(x)$, and non-linear $h_k(x)$ and $g_j(x)$ are linearly constrained and non-linearly constrained optimization problems respectively. For the linearly constrained problems, if *f(x)* are linear and *x* are continous variables, such problems are called linear programs. A non-linearly constrained optimization problem with a non-linear objective function may be referred to as a non-linearly constrained non-linear program. The proposed workflows in Chapters 4 and 5 consist of a multi-variable, multi-objective, non-linearly constrained non-linear program.

To find the solutions for an optimization problem, several algorithms exist. These algorithms or heuristics can be gradient-based (e.g. gradient descent, stochastic gradient descent), Hessian-based (e.g. Newton's method, Quasi-Newton's method) Marquardt's method etc.), or algorithms with non-differentiable objective functions (e.g. simulated annealing, particle swarm optimization,

differential evolution, genetic algorithm, etc.). The genetic algorithm is implemented in this research due to its advantages such as; ability to do a global search, resistance to being trapped in local optima (Sivanandam and Deepa, 2008), capacity to handle noisy objective functions, use of function evaluations instead of derivatives (Bittencourt and Horne, 1997), etc. The algorithm is explained below.

3.1. Genetic Algorithm

The genetic algorithm (GA) is an optimization technique inspired by natural evolution, which operates based on a populution of artificial chromosomes or solutions, where each chromosome has a fitness value or objection function, which is a measure of the quality of the solutions to an optimization problem. The GA begins by randomly creating a population, then the process of selection which is based on the fitness function, and recombination is performed to produce a new generation of population. Usually selection is done with replacement with highty fit chromosomes having the possibility of being selected more than once. Examples of selection methods include Roulette Wheel, tournament, random stochastic and truncation selection.

Recombination consists of two probabilistic genetic operators which are crossover and mutation. The crossover operation involves the genetic mixing of two parent chromosomes selected to produce offspring chromosomes. The execution of the crossover operator is dependent on a random number between 0 and 1, which is generated based on a uniform probability distribution. This number is compared to a pre-determined crossover rate, and if this is greater, no crossover occurs. A crossover operator is applied when the random number is less than or equal to the

crossover rate. One-point, multi-point and uniform crossover are common crossover operations. The mutation is performed after crossover and this operation is applied to one parent chromosome by flipping one or more allele (the value occuring at a position in a chromosome). The decision on whether to perform mutation is similar to crossover (i.e. based on a comparision between a pre-determined rate to a random number). Since mutation rates are usually small, it is less likely for it to occur.

The process of selection and recombination is performed iteratively and a sequence of successor generations of a population are created with a decrease or increase in the fitness values of the chromosomes, depending on whether the optimization is a minimization ot maximization problem, until a stopping criterion is met.

For multi-objective optimization problems, variant of the GA, such as the Non-Dominated Sorting Genetic Algorithm (NSGA) can be used.

3.2. Data-Driven Modeling Concepts

As previously discussed, the determination of optimal decision variables using GA is based on the minimization or maximization of a fitness or objective function. For petroleum engineering problems, objective functions can be reservoir models, however, since these functions are evaluated for each chromosome in a population, and reservoir simulations may take several hours or days to complete, the solution of such problems may require huge computional footprint and lot

of time. Hence, making data-driven or machine learning (ML) techniques to be effective. The ML-based objective functions are usually referred to as proxy models.

Machine learning is the use of statistical, mathematical tools or heuristics, coupled with domain knowledge to model a physical system of interest. ML algorithms are categorized as supervised (e.g., linear regression, random forest algorithm, artificial neural network (ANN), support vector machines, etc.), unsupervised (K-means, hierarchical clustering, principal component analysis (PCA), etc.) and reinforcement learning. To develop a data-driven model, three sets of data are required, which are: training set, test set and validation set. The training set is employed for model training, and the validation set provides an unbiased assessment of the model performance during training, which sometimes triggers the termination of the training process. The test set is required to perform an unbiased evaluation of the model at the end of the training process. Usually, the training set take a larger proportion of the entire dataset or observations compared to the test and validation set. A subset of a particular observation fed into the model is referred to as input variables or predictors, while the expected output (in the case of supervised learning) is called a target variable. Therefore, in supervised learning, there is a functional mapping from the predictors to the target variables. Unsupervised learning algorithms do not require target variables. In

reinforcement learning, an agent learns from interaction with its environment through different trial and error techniques in order to maximize a reward function.

The data-driven methods that are employed in this research include ANN, random forest algorithm, PCA and K-means clustering.

3.2.1. Artificial Neural Network

The artificial neural network (ANN) is created to mimic a biological neuron by accepting signals (inputs) from neighbouring neurons and processing them. Based on the results of the processing, a neuron can decide whether to fire an output signal or not. If the output signal is triggered, it may either be 0 or 1, or real value between 0 and 1. An ANN architecture typically consists of an input vector/layer, one or more hidden layers, and an output layer. The processing process is implemented by feeding an input vector X into each neuron, and the output is calculated based on the function $g(f(x))$. $f(x)$ is a linear combination of the elements in X (i.e. x_i; $i = 1, 2...., m$) with the weights $w_{j,i}$ and/or bias b, and g is a non-linear transfer function for $f(x)$ to the output value a in the hidden layer. To calculate the output for a successor hidden layer, a is fed and $g(f(a))$ is implemented until the final output value y (i.e., the output from the output layer) is obtained.

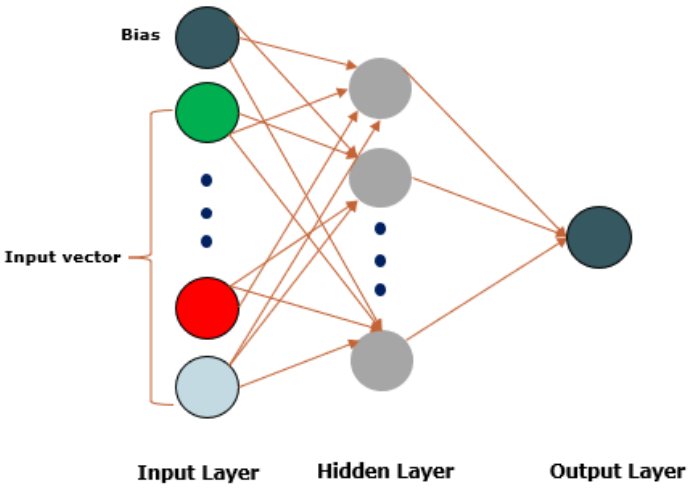


Figure 3-1: ANN Architecture

The weights are updated through a backpropagation algorithm which includes an optimization framework where the error between the actual target values and model predictions (e.g., the Mean Squared error, MSE) is minimized after several batches of training.

3.2.2. Random Forest Algorithm

The random forest algorithm is an ensemble learning method which combines predictions from decision trees for better predictive accuracy. The algorithm uses bootstrap aggregation or bagging, which is a random sampling with replacement that reduce the variance of decision trees. The bootstrapped dataset is the same size as the original dataset, and depending on the number of decision trees, multiple trees are trained, and the average of outputs from the trees is the predicted output. Compared to ANN, the random modeling technique has less computational footprint, with fewer tuning parameters (Muhammad et. al., 2017).

3.2.3. Principal Component Analysis

Principal component analysis (PCA) performs dimensionality reduction for a dataset by projecting it to a lower dimension space. For the input data X of dimension $m \times n$ (n = number of variables; m = number of samples), a mean-adjusted data Z (of dimension $m \times n$) is calculated by subtracting the mean of variable X^j ($j = 1, 2..., n$) from each of its data point X_i^j ($i = 1, 2..., m$) to eliminate bias. Next, using a singular value decomposition (SVD) technique, the eigenvectors (or principal components) of a $n \times n$ covariance matrix of Z. The covariance matrix σ_Z is computed thus:

$$\sigma_Z = \frac{1}{m}\sum_{i=1}^{m} Z \times Z^T, \tag{1}$$

3.2.4. K-means Clustering

K-means is a common cluster analysis method for recommender systems, anomaly detection, reservoir characterization, etc. Other clustering algorithms include density-based scan and hierarchical clustering. K-means identifies internal groupings within data by grouping observations or samples into k-clusters based on similarities in data, and a measure of this similarity is the squared Euclidean distance. Cluster assignment of the observations is performed by minimizing the mean squared Euclidean distance, J from each observation to its nearest cluster centroid.

$$J = \frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} w^{(i,j)} \left\| x^{(i)} - \mu^{(j)} \right\|_2^2\right], \tag{2}$$

Where $\mu^{(j)}$ is the *j-th* cluster centroid, $w^{(i,j)}$ is 1 if the observation $x^{(i)}$ is assigned to cluster j, otherwise, $w^{(i,j)}$ is 0.

A limitation of K-means is its sensitivity to the initialization of the cluster centroids, hence, a common practice is to repeat random initialization of centroids and clustering until optimal grouping is obtained. To measure the performance of a clustering process, the silhouette value, which measures the similarity between an observation and its cluster compared to other clusters, is used. The silhouette value has an interval of [-1, 1]; a value close to -1 means that an observation is misclassified while a value close to 1 means that an observation is closer to other observations within its cluster compared to the remaining samples in the dataset. A value of 0 suggests that a sample can belong to more than cluster.

3.3. Uncertainty Quantification

In the absence of real-world data, uncertainty assessment is critical to the analysis, prediction, and optimization of physical systems, and a common source of this uncertainty is data uncertainty. Data uncertainty may occur in form of a random variable in which the accuracy of a distribution is dependent on the amount of available data, or from measurement error, or when data exists as a range of values (Mahadevan and Sarkar, 2009). Uncertainty in data can also come from subjective interpretation of geologic properties. Bootstrapping and the Monte Carlo method are some of the widely adopted techniques for uncertainty analysis. For this research, the Monte Carlo sampling method is employed as it is used to predict the possible outcomes of an uncertain event through the random sampling of a finite number of realizations, which is based on a probability distribution.

The application of the Monte Carlo sampling can be found in literatures such as Bieker et. al., (2006), Al-Mudhafar and Rao (2016) and Mehana et al. (2019). In this work, the probability distribution for sampling is determined using the Bayesian approach (Russell and Norvig, 2016). The Bayesian theorem is a posterior probability estimation method for estimating the likelihood of an event based on apriori data. Given events A and B, the Bayesian theorem can be expressed as:

$$P(B|A) = \frac{P(B|A)\,P(A)}{P(B)}, \tag{3}$$

where *P(A|B)* is the conditional of A given the occurrence of B, *P(A)* and P(B) are the probabilities of A and B respectively irrespective of any other event, and *P(B|A)* is the conditional of B given the occurrence of A. *P(A|B)* is posterior probability and *P(A)* is prior probability.

Chapter 4 Incorporating Phase Behavior Constraints in Multi-Objective Optimization of Warm Vaporized Solvent Injection Process

Abstract

The warm vaporized solvent injection process has been proposed as a more environmentally friendly alternative to steam-based technologies for bitumen recovery. The process typically involves injecting heated solvent vapor into a horizontal injector; the solvent condenses and dissolves into bitumen, while the diluted oleic phase would flow towards a horizontal producer. Despite the promising results reported from several pilot projects near Fort McKay, Alberta, successful commercial-scale extraction is costly and would require a detailed optimization of the pertinent design variables. The main challenge is that this is a multi-objective optimization (MOO) problem, which aims to balance the trade-offs between conflicting performance objectives while honoring the various operational constraints. In this study, a systematic workflow is formulated to optimize these multiple conflicting performance objectives considering phase behavior constraints.

A 2D synthetic model based on typical Athabasca oil sands properties is constructed to simulate the warm vaporized solvent process. The addition of non-condensable gas (methane) into the solvent (propane) is examined. The resultant changes in thermodynamic properties and equilibrium phase behavior are considered in determining the practical limits of the decision variables (e.g., bottom-hole injection pressure and temperature). The objective functions,

including oil recovery factor, solvent retained-to-oil ratio, and energy consumption, are defined, and a factorial experimental design is employed to identify a subset of decision variables that exhibit minimal redundancy internally and create the dataset for proxy model development. To reduce the computational costs associated with reservoir simulations, proxy models, e.g., the artificial neural network (ANN), is developed and applied. Finally, a Pareto-based MOO scheme is implemented to estimate the optimal decision variables.

Despite the higher front-end loading requirement of the ANN proxy modeling, the MOO with proxy modeling still requires significantly less execution/running time as compared to a MOO with traditional flow simulation (e. g., a 97% reduction in CPU time). This reduced running time is important for alleviating the computational load when evaluating the objective functions during the optimization process. More importantly, this optimization scheme is capable of identifying a set of optimal decision variables.

This work presents a practical workflow for optimizing various design variables associated with many solvent-based bitumen recovery processes. Specific considerations including the practical limits for operating constraints and computational efficiency are examined and incorporated. It is anticipated that this workflow can be readily integrated into the design and decision-making processes in reservoir management.

4.1. Introduction

A vast majority of Canada's oil sands are located in Western Canada, notably in the provinces of Alberta and Saskatchewan. According to the Alberta Energy Regulator (2019), close to 80% of Alberta's heavy oil reserves can be produced through in-situ thermal and/or solvent-based methods. Due to their high viscosity and specific gravity, heavy oil and bitumen enhanced oil recovery schemes are employed to reduce the oil viscosity through heat (in the form of steam or hot water injection) and/or mass transfer (injection of light hydrocarbon solvents).

Popular thermal extraction techniques are steam-assisted gravity drainage (SAGD) and cyclic steam stimulation (CSS). Current challenges associated with these techniques include high operating costs, technical constraints, operational safety concerns, water consumption, and greenhouse gas (GHG) emissions. Non-thermal methods (e.g., solvent-based processes) have been proposed as promising alternatives. These solvent-based (with or without steam) techniques can potentially offer the following benefits (Zhang et al., 2019a):

- Injecting pure solvent is suitable for thin reservoirs where heat loss from steam is substantial.
- Water treatment cost is lower for solvent-based methods (less or no steam is needed).
- Limited solvent solubility in water renders these solvent-based processes to be more effective in reservoirs underlain by a bottom water zone.

- Heat requirement is lower in the pure solvent injection method in contrast to the steam-based method since for most light hydrocarbon solvents, the dew point temperature at reservoir pressure is usually much lower than the steam temperature.

Therefore, many simulations, laboratory, and field studies involving the solvent-based and steam-solvent hybrid methods have been presented. Butler and Mokrys (1991) proposed the vapor extraction (VAPEX) method, which utilizes a similar well configuration as in SAGD. This process involves the injection of a pure vaporized solvent to reduce the bitumen viscosity. Many previous simulation studies have concluded that the production rate from VAPEX is usually too low because the solvent diffusion rate is too low (Shi and Leung, 2014a, b). A modification of VAPEX is the warm vaporized solvent (warm VAPEX) method. Warm VAPEX or the patented N-Solv® technique (Nenniger and Nenniger, 2001) involves the injection of a heated solvent vapor close to dew point conditions, allowing the solvent to condense and dissolve once in contact with the in-situ cold bitumen. The solvent is produced with the bitumen, then separated, and re-injected into the reservoir. The mechanisms for viscosity reduction in warm VAPEX are solvent diffusion, solvent dispersion, and heat transfer. Additional heat transfer to the bitumen is achieved through the release of latent heat after condensation (Nenniger and Dunn, 2008). The N-Solv® process, which is a variant of the warm VAPEX method, involves the injection of a pure heated vaporized solvent, such as propane (C_3), and pilot test results conducted near Fort McKay, Alberta, have proven the effectiveness of this technique for commercial-scale bitumen extraction (Nenniger and Dunn, 2008). Other simulation studies, such as the high-temperature multicomponent solvent vapor extraction (HTMS-VAPEX) and Azeotropic HTMS-VAPEX (AHTMS-VAPEX) processes

developed by Khaledi et al. (2018), also demonstrated that it is possible to leverage the benefits of solvent dilution with effective heat transfer for achieving optimal recovery.

Despite the successful implementation of N-Solv® at the pilot scale and, to some extent, the field scale (Emissions Reduction Alberta, 2016), commercial-scale extraction still requires detailed optimization of the pertinent decision variables. The first challenge is that this is a multi-objective optimization (MOO) problem, which aims to balance the trade-offs between multiple conflicting performance objectives while considering operational constraints. Secondly, the commercial supply of light hydrocarbon solvents usually has about 5–10% hydrocarbon impurities and non-condensable gases, such as methane (C_1); non-condensable gases may also be added to the solvent mixture to delay liquid condensation in the solvent chamber (Das, 2008). However, solvent mixtures usually partition into separate phases (oleic and vapor) inside the solvent chamber due to variation in solubility for different components (Das, 2008; Zhang et al., 2019a). Lighter components (e.g. C_1) may accumulate near the top of the reservoir, providing a thermal barrier to the overburden heat loss, but it may also be inhibiting the chamber from propagating vertically (Das, 2008; Ma and Leung (2020a); heavier components (e.g. C_3) tend to stay in the oleic phase in the extracted chamber. Apart from reducing the gravity drainage potential (due to density differences between the vapor and liquid phases) for the mobilization of bitumen (Das, 2008), the accumulation of liquid C_3 also leads to solvent retention in the reservoir. Hence, the performance of the warm VAPEX process is contingent on the amount of non-condensable gas. Thirdly, for a given solvent concentration and pressure, solvent solubility reduces with increasing temperature; injection at superheated conditions may hinder liquid condensation at the solvent-bitumen interface, with a negative impact on solvent diffusion and viscosity reduction. However,

superheated solvents also provide more thermal energy for viscosity reduction. The net effect of solvent composition and injection conditions on bitumen recovery and solvent retention is a complex function of the dominant rate-controlling step (diffusion or heat transfer) and the extent of solvent partitioning (condensation) inside the vapor chamber behind the chamber interface. In the end, conducting a simple sensitivity analysis, where one or two variables are varied at a time, to identify the optimal solvent concentration, bottom-hole injection temperature, and pressure is not feasible. A MOO strategy, which involves all possible combinations of the decision parameters, is necessary to assess the trade-offs between multiple conflicting objectives and the complex interplay between these decision parameters.

Three distinct objective functions, including oil recovery, solvent loss, and energy requirement, are considered in this study. Solvent type, solvent concentration, bottom-hole injection temperature, bottom-hole injection pressure, reservoir heterogeneity, well configuration, and pre-heating period are factors that can influence the performance of the process. In most optimization studies, a single objective function is defined; for example, Al-Gosayir et al. (2013) applied a hybrid genetic algorithm technique in the optimization of the steam-over-solvent injection process in fractured reservoirs (SOS-FR), where the Money Recovery Factor (MRF) combining the cumulative steam/solvent injection and production and oil production, was defined. In other cases, a MOO problem can be formulated as a single objective optimization (SOO) problem by aggregating all the objectives into one weighted objective function or by changing all but one of the objectives into constraints. The limitation, however, is that the trade-offs between objectives cannot be easily evaluated when aggregated functions are used. Also, it may be impossible to find an optimal solution if the search space is non-convex (Ngatchou et al., 2005). For instances in

which there are more than three objective functions, several objective functions can be grouped considering the trade-offs among them, reducing the total objective function count (Hutahaean et al., 2017).

Solving a MOO problem entails searching for an optimal set of solutions along the Pareto-optimal front. Common MOO techniques include Vector Evaluated Genetic Algorithm (VEGA) (Schaffer, 1985), Niched Pareto Genetic Algorithm (NPGA) (Horn et al., 1994), Reference Vector Evolutionary Algorithm (RVEA) (Cheng et al., 2016), and Multi-Objective Genetic Algorithm (MOGA) (Fonseca and Fleming, 2011). Other widely adopted MOOs are elitist Multi-Objective Evolutionary Algorithms (MOEAs) such as Non-dominated Sorting Genetic Algorithm-II (NSGA-II) (Deb et al., 2000), Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler and Thiele, 1998), and the Pareto-Archived Evolution Strategy (PAES) (Knowles and Corne, 1999). Many studies have utilized MOO workflows in various enhanced oil recovery (EOR), reservoir modeling, and well placement applications. For example, Hamdi et al. (2018) proposed a MOO approach which is based on the sequential Gaussian process to history match pressure, gas, oil, and water production rates from a tight reservoir, while RVEA was implemented to history match the production data for a real-field reservoir model considering more than three objective functions in Hutahaean et al. (2016). Also, an ensemble-based MOO was applied to the optimization of long- and short-term water flooding in Fonseca et al. (2014), and Min et al. (2014) compared a number of MOO algorithms for production history matching. Proposed by Deb et al. (2000), the NSGA-II, which is a variant of the genetic algorithm (GA), finds a diverse set of optimal solutions along a Pareto front by using a non-dominated sorting approach and an elitist-based technique. In contrast to PAES and SPEA, the NSGA-II is more efficient with a computational complexity of O

(MN^2), where *N* is the population size and *M* is the number of objectives (Deb et al., 2000). Zhang et al. (2019b) developed a hybrid NSGA-II workflow to history match and optimize an alkaline-surfactant-polymer (ASP) flooding process, where two objectives, including oil recovery and chemical usage, were considered. The NSGA-II with a similarity-based selection method was used to optimize the placement of a non-vertical well (Rostamian et al., 2019), where the net present value (NPV) and oil recovery factor are incorporated as the objective functions. (Ma and Leung (2020a) applied a Pareto-based multi-objective optimization (i.e., NSGA-II) for the design of a warm solvent injection process. In that work, only pure C_3 was injected, and the co-injection with other non-condensable gases was not explored.

The objectives of this study include the following: (1) developing a base simulation model to examine the influence of C_1 co-injection with C_3, bottom-hole injection pressure, and temperature on the performance of the warm VAPEX process; (2) performing an experimental design to identify relevant decision variables and sampling a set of training/ testing data that can be used for constructing various proxy models of the objective functions; (3) integrate an artificial neural network (ANN) proxy modeling technique with MOO (NSGA-II) to reduce the computational costs of physics-based (simulation) models when the evaluating objective function values; (4) identifying a set of Pareto-optimal decision variables for a MOO problem including C_3 loss, oil recovery factor, and injected enthalpy. Although incorporating an economic objective function (e.g., Net Present Value, NPV), into the optimization framework could be useful; however, due to the uncertainties in the NPV calculations (market conditions and company policies), a single aggregated economic function is not considered here. Besides, the three individual objective functions are the key economic drivers for most solvent projects. The MOO framework would

facilitate the examination of the trade-offs that may exist among these objectives. However, the inclusion of an economic objective function may be considered in future work. GHG emissions are often quantified by measuring the amount of GHGs produced at the surface: e.g., combustion of fuel gas to generate steam, thermal energy and power consumption, or while flaring gas. However, one of the limitations of this study is that those surface facilities are not incorporated in workflow, and only subsurface conditions and recovery performance are captured. Therefore, explicit quantification of GHG emission in the MOO workflow is not considered. Although there are some potential sources of GHG emissions for solvent-based technologies, including flaring of solution gas and solvent, and consumption of fuel gas for solvent heating, it is widely expected that the GHG emissions associated with solvent-based techniques would be lower than those from traditional steam-based methods, and that is due to the overall lower operating temperatures (usually between 7 and 90 ^{o}C), in comparison to typical steam injection at approximately 230 ^{o}C (Emissions Reduction Alberta, 2016).

The NSGA-II MOO algorithm is employed not only because of its common applicability but for its computational efficiency and elitist approach for selecting the optimal solutions. There are three particular contributions from this work. First, a novel parameterization scheme is devised to represent the interdependency among the three decision variables, ensuring that the injection conditions are either at the dew point or within a particular window of superheating. This scheme facilitates crucial phase behavior constraints to be directly incorporated into the MOO framework, which is a particularly significant consideration in most EOR applications. Secondly, the results offer important insights about the optimal operating strategies for the warm vaporized solvent injection process, where complex physical mechanisms including mass and heat transfer are

involved. The MOO workflow can reveal subtle patterns in the decision variables that are not necessarily detectable based on traditional analyses. Thirdly, although some commercial simulation packages have a built-in optimization module that can perform a variety of sensitivity analysis, proxy modeling, and multi-objective function optimization, usually, a user is offered limited flexibility in terms of the problem formulation; for example, only Particle Swarm Optimization (PSO) algorithm can be used, and the number of objective functions may be limited. Besides, the interdependency of the three decision variables cannot be explicitly incorporated. Therefore, there is a need to formulate a more flexible framework using other widely adopted and robust MOO algorithms, such as NSGA-II.

4.2. Methodology

4.2.1. Reservoir Model Description

A synthetic 2D, IK Cartesian, single porosity-permeability homogenous reservoir model, with reservoir dimensions of 35 × 20 × 32 m is built using a fully implicit thermal-compositional reservoir simulator, STARS™ (CMG, 2019b). Reservoir model properties representative of the Athabasca oil sands are assigned (Ma and Leung (2020a, b, c). The simulation is run for 15 years (5448 days) and a 4-month preheating period is imposed. A summary of the model inputs is shown in Table 4-1.

Table 4-1: Base model properties for warm VAPEX process

Description	Parameters	Input
Grid properties	Dimension of reservoir (m)	35 x 20 x 32

	Permeability for the I direction (mD)	2500
	Permeability for the J direction (mD)	2500
	Permeability for the K direction (mD)	1500
	Porosity (%)	32
Initial conditions	Reference depth (m)	200
	Reservoir pressure (kPa)	500
	Reservoir temperature (°C)	10
	Average initial water saturation (%)	13
	Molar concentration of dissolved C_1 (mole %)	5
Components	Names	Bitumen, Propane (C_3H_8), Methane (CH_4), Water (H_2O)
Rock/Fluid properties	Bitumen viscosity at 15°C and 101.325 kPa (cP)	562204
	Rock wettability	Water wet (capillary pressure ignored)
	Model for evaluating 3-phase k_{ro}	Stone's second model
	Relative permeability end points	k_{rw} = 0.79, k_{row} = 0.95, k_{rg} = 0.50, k_{rw} = 0.95
Well-pair constraints	Injector bottom-hole pressure (kPa)	1719

	Producer bottom-hole pressure (kPa)	1519
	Injection temperature (°C)	50

Fig. 4-1 shows the model configuration. Only one-half of the reservoir domain is constructed, assuming symmetric propagation of the solvent chamber. Moreover, a simple model is created to improve simulation speed so that multiple case scenarios can be developed efficiently for sensitivity studies and proxy modeling.

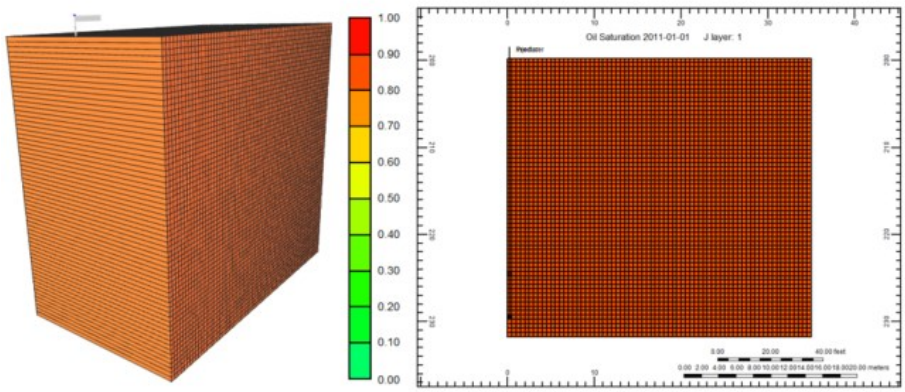


Figure 4-1: Illustration of the simulation domain.

4.2.2. Grid Size, Molecular Diffusion, and Mechanical Dispersivity Sensitivity

To investigate the impact of numerical dispersion only (without molecular diffusion and mechanical dispersivity) on model performance, three grid sizes were examined. For the 1 m × 20 m × 1 m (base case), 0.667 m × 20 m × 0.667 m, and 0.5 m × 20 m × 0.5 m, the total computing

time are approximately 5 minutes, 20 minutes, and 30 minutes, respectively using a personal computer with 16.0 GB installed RAM and Intel(R) Core i7 processor.

Figs. 4-2 to 4-4 show that model response is sensitive to grid size. In Fig. 4-2, the solvent saturation, oil saturation, gaseous phase C_1 mole fraction, and gas saturation for each grid size are presented. Solvent saturation is given as:

$$\text{Solvent saturation} = S_o \times \bar{x}_{sol}, \tag{1}$$

where S_o is oil saturation. $\bar{x}_{sol}$ is the solvent volume fraction (assuming additivity), which is expressed as:

$$\bar{x}_{sol} = \frac{\frac{(1 - x_{oil}) \times M_{sol}}{\rho_{sol}}}{\frac{x_{oil} \times M_{oil}}{\rho_{oil}} + \frac{(1 - x_{oil}) \times M_{sol}}{\rho_{sol}}}. \tag{2}$$

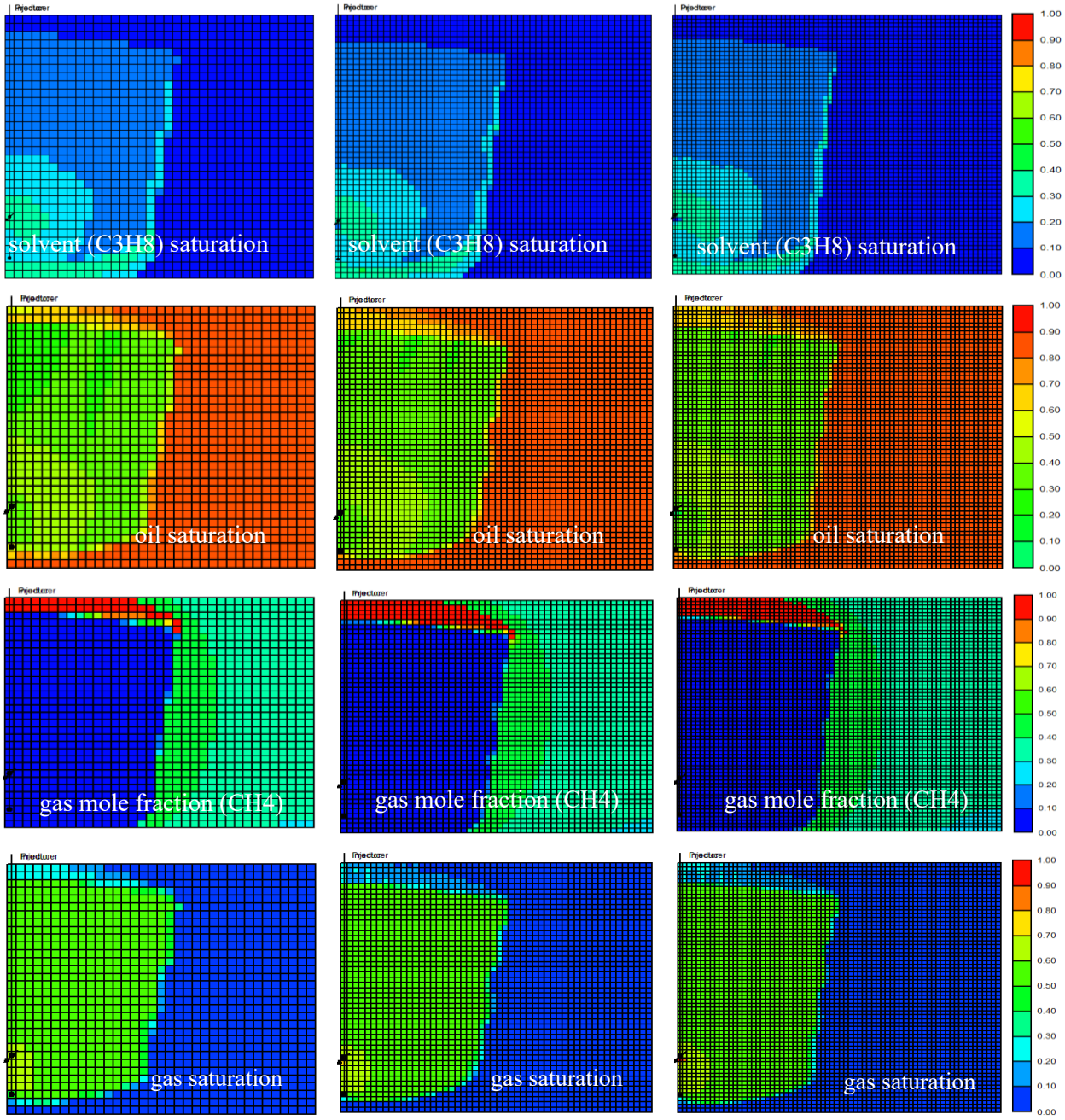
M_{oil} is the molecular weight of oil, M_{sol} is the molecular weight of solvent, ρ_{oil} and ρ_{sol} are the oil and solvent densities, respectively, and x_{oil} is the oil mole fraction.

From Figs. 4-3 and 4-4, it is noted that as the grid size is reduced beyond 0.667 m, any changes in oil recovery factor, enthalpy, cumulative C_3 injected, and cumulative C_3 produced are not overly significant. The grid size of 0.5 m × 20 m × 0.5 m could not be reduced any further due to wellbore instability error (wellbore diameter is comparable to the grid size).

To assess the influence of grid sizes, molecular diffusion, and mechanical dispersivity on numerical and physical dispersion, a sensitivity analysis is performed, and the results are summarized in Table 4-2. Three different grid sizes, as well as various combinations of molecular

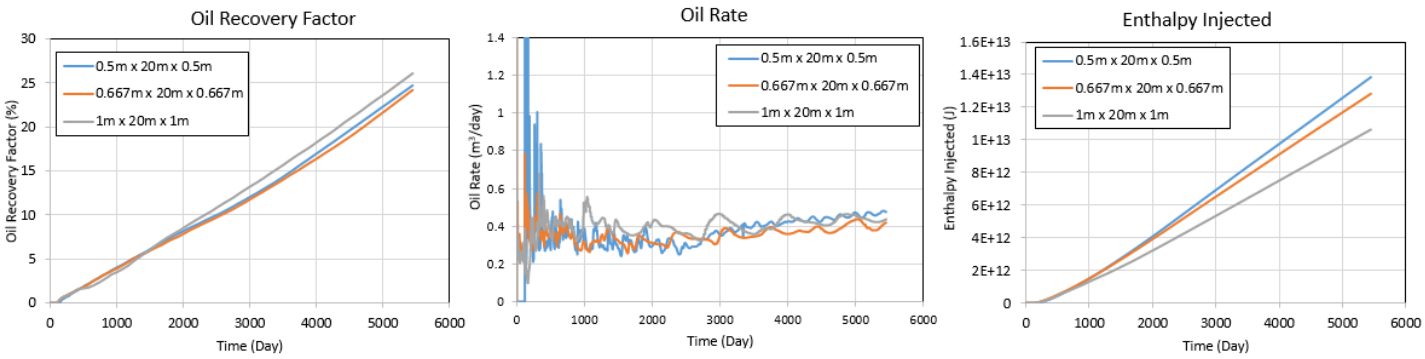
diffusion and mechanical dispersivity, are tested. All the values tested are within the ranges that were reported in Das and Butler (1996) and Perkins and Johnston (1963). The results in Table 4-2 also indicate that oil recovery may vary by 15% for the ranges of molecular diffusion and mechanical dispersivity tested. Therefore, in the end, values of 2.00×10^{-5} m^2/day and 4.32×10^{-5} m are assigned for diffusion coefficient and dispersion, respectively; this choice is consistent with other values reported in the literature (Ji, 2014). As for the grid size, in order to minimize numerical dispersion, the finest mesh (0.5 m × 20 m × 0.5 m) is selected.

Time – 5448 days



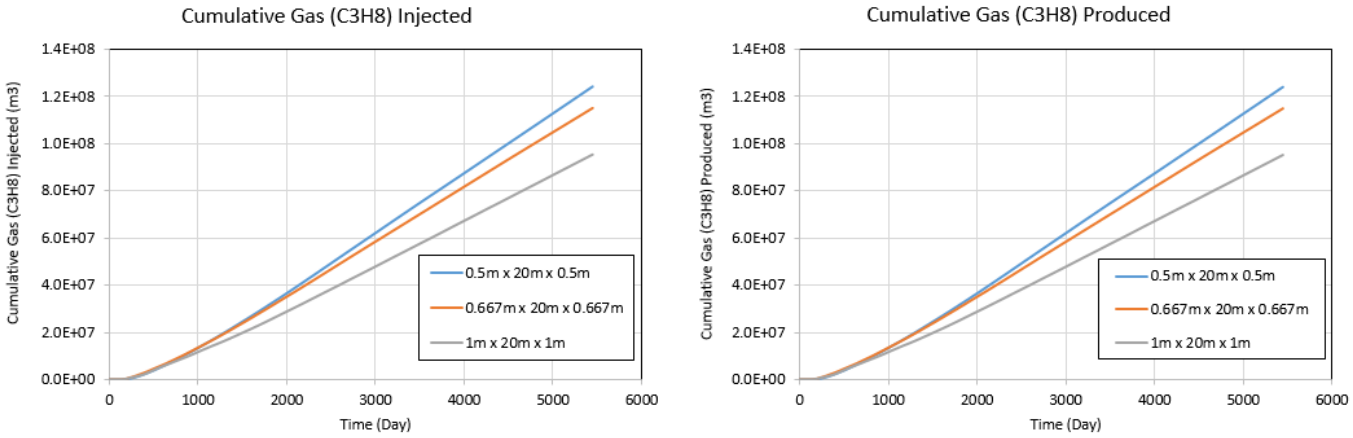
$1m \times 20m \times 1m$ $0.67m \times 20m \times 0.67m$ $0.5m \times 20m \times 0.5m$

Figure 4-2: Grid size sensitivity.



(a) (b) (c)

Figure 4-3: Grid size sensitivity. (a) – Oil recovery factor; (b) – Oil rate; (c) – Enthalpy Injected.



(a) (b)

Figure 4-4: Grid size sensitivity. (a) – Cumulative solvent injected; (b) – Cumulative solvent produced.

Table 4-2: Grid size, molecular diffusion, and mechanical dispersivity sensitivity results

Case	Grid Size (m)	Molecular Diffusion (m^2/day)	Mechanical Dispersivity (m)	Oil Recovery Factor (%)
1	0.5	4.32×10^{-7}	2.00×10^{-4}	32
2	0.5	4.32×10^{-6}	2.00×10^{-4}	32
3	0.5	4.32×10^{-5}	2.00×10^{-6}	46

4	0.5	4.32×10^{-5}	2.00×10^{-5}	47
5	0.67	4.32×10^{-7}	2.00×10^{-4}	30
6	0.67	4.32×10^{-6}	2.00×10^{-4}	34
7	0.67	4.32×10^{-5}	2.00×10^{-6}	47
8	0.67	4.32×10^{-5}	2.00×10^{-5}	47
9	1	4.32×10^{-7}	2.00×10^{-4}	32
10	1	4.32×10^{-6}	2.00×10^{-4}	33
11	1	4.32×10^{-5}	2.00×10^{-6}	47
12	1	4.32×10^{-5}	2.00×10^{-5}	47

4.2.3. Solvent Phase Behaviour and Fluid Model

Understanding the phase behavior of solvent mixtures is critical to optimizing the warm VAPEX process. Methane, a carrier gas responsible for solvent propagation towards the chamber-bitumen interface, may impede overburden heat loss and help to increase the dew point pressure of the C_1-C_3 mixture at a particular temperature. However, it also acts as an insulative layer, reducing the transfer of thermal energy between the solvent chamber and bitumen, which is detrimental to bitumen recovery. It is desirable to inject at a condition close to the dew point to maximize mutual diffusivity between the bulk bitumen and condensed solvent, as well as the transfer of latent heat of vaporization to the colder bitumen (Nenninger and Dunn, 2008). On the other hand, injecting at superheated conditions may be beneficial in terms of providing additional thermal energy to mobilize the bitumen. To illustrate the impacts of bottom-hole injection temperature and pressure, the P-T relationship is examined for several C_1 mole fractions using WINPROP^{TM} (CMG, 2019a).

Fig. 4-5 shows the P-T relationship for a C_1- C_3 binary mixture. K-values for the reservoir fluids are shown in Table 4-3. From Fig. 4-5, it is evident that an increase in C_1 mole fraction leads to an increase in dew point pressure at a particular temperature. Bitumen viscosity plot is shown in Fig. 4-6.

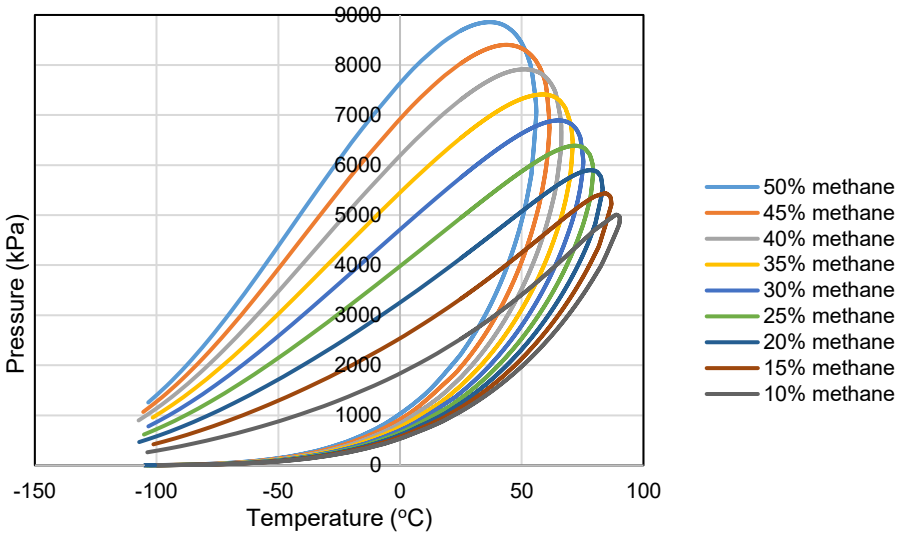


Figure 4-5: P-T Diagram of a methane-propane binary mixture.

Table 4-3: Fluid model

	Propane (C_3H_8)	**Methane (CH_4)**	**Water (H_2O)**
KV1 (kPa)	9.0085×10^5	5.45475×10^5	1.1860×10^7
KV4 (°C)	-1872.46	-879.84	-3816.44
KV5 (°C)	-247.99	-265.99	-227.02

Physically, transport across the gaseous (vaporized solvent) phase and oleic (bitumen) phases is due to rapid kinetic processes (condensation and evaporation), while transport within each phase is due to convection and diffusion. The bitumen phase close to the interface becomes mobilized

where the C_3 concentration has increased sufficiently. Convection (mainly driven by gravity) would subsequently remove the mobilized bitumen. A boundary layer is formed, where the bitumen is mobilized due to solvent diffusion, and the mobilized bitumen is removed, controlling the equilibration process. In particular, the rates at which these two mechanisms are occurring (diffusion and convection) would depend on the boundary layer thickness. In the numerical simulation, both these mechanisms are modeled under the following assumptions: (1) equilibrium is attained at each grid block; (2) the smallest resolution is that of a grid block (i.e., it is not possible to model a boundary layer thickness less than the grid block size); (3) total dispersion is the combination of molecular diffusion, mechanical dispersion, and numerical dispersion.

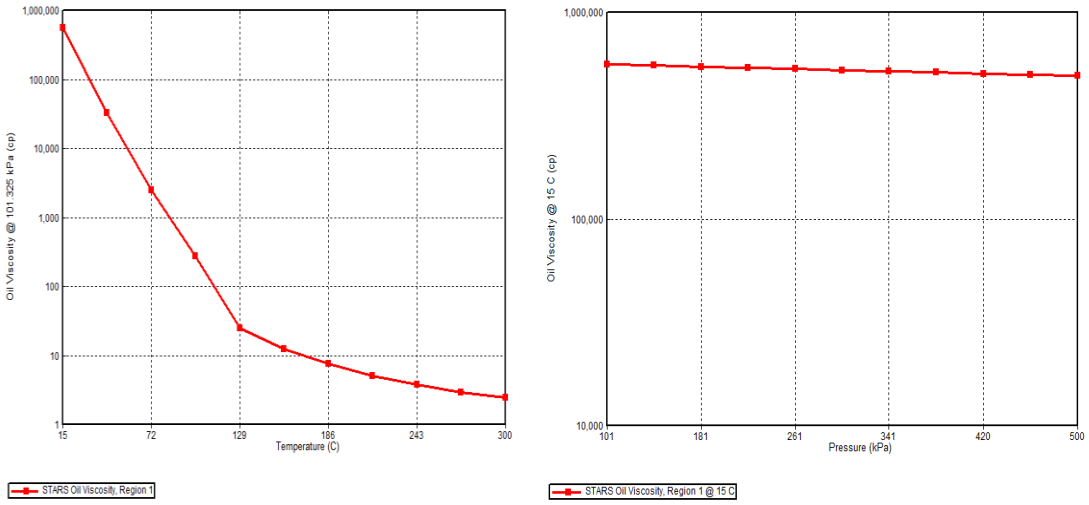


Figure 4-6: Bitumen viscosity model

It should be emphasized that various assumptions have been made (sections 4.2.1 - 4.2.3), and the primary ones are summarized here: (1) 3D effects are not incorporated; the models used here are 2D along the x-z plane. (2) Blowdown physics are not considered; it is assumed that any solvent lost in the reservoir is retained and not recoverable. (3) The temperature dependency of relative

permeability functions is neglected. (4) Reservoir heterogeneity is not examined. It is expected that reservoir heterogeneity would introduce additional mixing, and the diffusion and dispersion represented here in the homogeneous model are less than what would have been observed in a heterogeneous reservoir. However, despite the simplifications, the simulation model does capture many of the key physical mechanisms relevant to the heat and mass transfer processes. The model is also able to provide both qualitative and quantitative information about the effect of C_1 co-injection with C_3 on bitumen recovery factor and solvent retention in the reservoir. Most importantly, the focus of this work is to illustrate how a MOO workflow can take into account the phase envelope constraint and be used to gain insights about optimal ranges for several key operational parameters.

4.2.4. Selection of Design Variables and Objective Functions

The identification of the design variables that strongly influence the warm VAPEX process is crucial to any MOO scheme; hence a 2-level full factorial experimental design is employed to assess qualitatively the influence of various design variables on the objective functions, examine any correlation among them, and eliminating potential redundant variables. The two objective functions are solvent retained-oil ratio (SolOR) and recovery factor (RF):

$$SolOR = \frac{V_{solvent}^{inj} - V_{solvent}^{prod}}{V_{oil}^{prod}}, \tag{3}$$

where $V_{solvent}^{inj}$ is the total volume of gaseous C_3 is injected (in m^3), $V_{solvent}^{prod}$ is the total volume of C_3 recovered at the surface (in m^3) and V_{oil}^{prod} is the total volume of bitumen extracted (in m^3).

All the parameters are at surface conditions. Tables 4-4 and 4-5 present the experimental design set-up and the results. Table 4-4 is sorted in order of decreasing *RF*.

Table 4-4: Design input

Design variables		**Low (-1)**	**High (+1)**
C3 mole fraction (fraction)	A	0.5	0.9
Bottom-hole injection Pressure (kPa)	B	1078	1976
Bottom-hole injection Temperature (°C)	C	25	50
Preheating period (months)	D	4	12

Table 4-5: Standard order table

#	**A**	**B**	**C**	**D**	**SOR (m^3/m^3)**	**RF (%)**
8	1	1	1	-1	80.322	42.532
16	1	1	1	1	80.456	42.056
4	1	1	-1	-1	83.582	41.690
12	1	1	-1	1	83.959	40.611
11	-1	1	-1	1	108.706	9.383
10	1	-1	-1	1	97.159	8.707
15	-1	1	1	1	66.544	8.458
3	-1	1	-1	-1	110.986	8.340
2	1	-1	-1	-1	98.266	7.483
14	1	-1	1	1	60.775	7.388
7	-1	1	1	-1	66.781	7.259
6	1	-1	1	-1	59.299	6.063

13	-1	-1	1	1	34.511	3.017
9	-1	-1	-1	1	68.306	2.929
5	-1	-1	1	-1	30.532	2.185
1	-1	-1	-1	-1	65.450	1.988

The regression model for a 2-level factorial experiment with four factors may be expressed as.

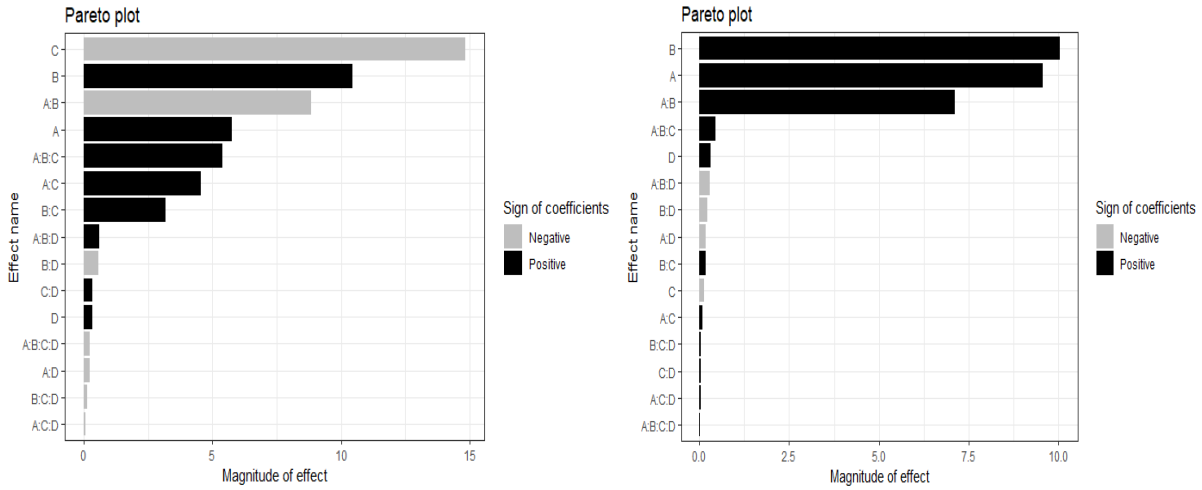
$$\begin{aligned} Y = \beta + \beta_A.x_A + \beta_B.x_B + \beta_C.x_C + \beta_D.x_D + \beta_{AB}.x_Ax_B \\ + \beta_{AC}.x_Ax_C + \beta_{BC}.x_Bx_C + \beta_{BD}.x_Bx_D + \beta_{CD}.x_Cx_D \\ + \beta_{AD}.x_Ax_D + \beta_{ABC}.x_Ax_Bx_C + \beta_{ABD}.x_Ax_Bx_D \\ + \beta_{BCD}.x_Bx_Cx_D + \beta_{ACD}.x_Ax_Cx_D \\ + \beta_{ABCD}.x_Ax_Bx_Cx_D + \varepsilon \end{aligned} \quad . \qquad (4)$$

β is the intercept, β_A, β_B, β_C, and β_D represent the effects of Factors, A, B, C, and D respectively, β_{AB} denotes the effect due to the interaction between Factors A and B, while β_{AC} denotes the effect due to the interaction between Factors A and C, and so on. ε is the random error term.

Fig. 4-7 is a Pareto plot showing the absolute values of all model coefficients in decreasing order as horizontal bars. The sign of each coefficient is denoted by the bar colors; black for positive coefficients (or effects) and grey for negative coefficients (or effects). The Pareto plot is created using the PID package (Dunn, 2021) in R software. Similar Pareto plots can be found in several references (Okafor, 2020; Jiju, 2014; Anirban, et. al., 2016). Among the four design variables, the preheating period (Factor D) exhibited significant redundancy internally– effects of Factor D, or combinations of D and other factors, are minimal in comparison to Factors A (C_3 mole fraction), B (bottom-hole injection pressure), and C (bottom-hole injection temperature). This led to the

choice of C_3 mole fraction, bottom-hole injection pressure, and temperature as the primary design variables for this study. It should be mentioned that the maximum operating pressure (MOP) for all cases is around 5000 kPa, and the maximum threshold is not exceeded.

Fig. 4-7 also show that complex interactions between Factors A, B, and C may have an impact on the two objective functions: for instance, Factors A and B exhibit a positive effect on *SolOR* and *RF*, while the combined interaction between several factors (e.g., Factors A and B) has a negative and positive effect on both *SolOR* and *RF*, respectively. Besides, the magnitude of influence on each objective function varies for different combinations of factors. Although these relationships cannot be exhaustively studied using a 2-level factorial design, which neglects interactions at the intermediate levels, the results offer a preliminary assessment of the complicated trade-offs that may exist between these objectives.



(a) (b)

Figure 4-7: 2-level factorial experimental design results. (a) - effect on SolOR; (b) - effect on RF.

4.2.5. Sensitivity Analysis

In this section, different injection conditions and C_1 co-injection concentrations are tested. First, for a given C_1 mole fraction, the mixture is injected at different dew point pressure and temperature for C_1 co-injection ranging from 0% to 50%. Next, the solvent is injected at different superheated conditions by varying the temperature at a given pressure.

4.2.5.1. Injection at Dew Point Conditions

Fig. 4-8 compares the oil viscosity for different C_1 concentrations. As expected, *RF* is proportional to the solvent chamber size. It may also be anticipated that *RF* would decrease with an increase in C_1 co-injection (or decrease in C_3 concentration); this is true except when C_1 mole fraction exceeds 20%. One of the plausible explanations for this trend reversal is that the dew point pressure of the C_1-C_3 mixture, at a fixed temperature, increases with C_1 concentration; therefore, an increase in bottom-hole injection pressure is required at higher C_1 concentration, and that increase in injection pressure leads to a higher *RF*. The thermal insulation effect of C_1 is observed in Fig. 4-9, where C_1 accumulates at the top of the reservoir and acts as a barrier to overburden heat loss.

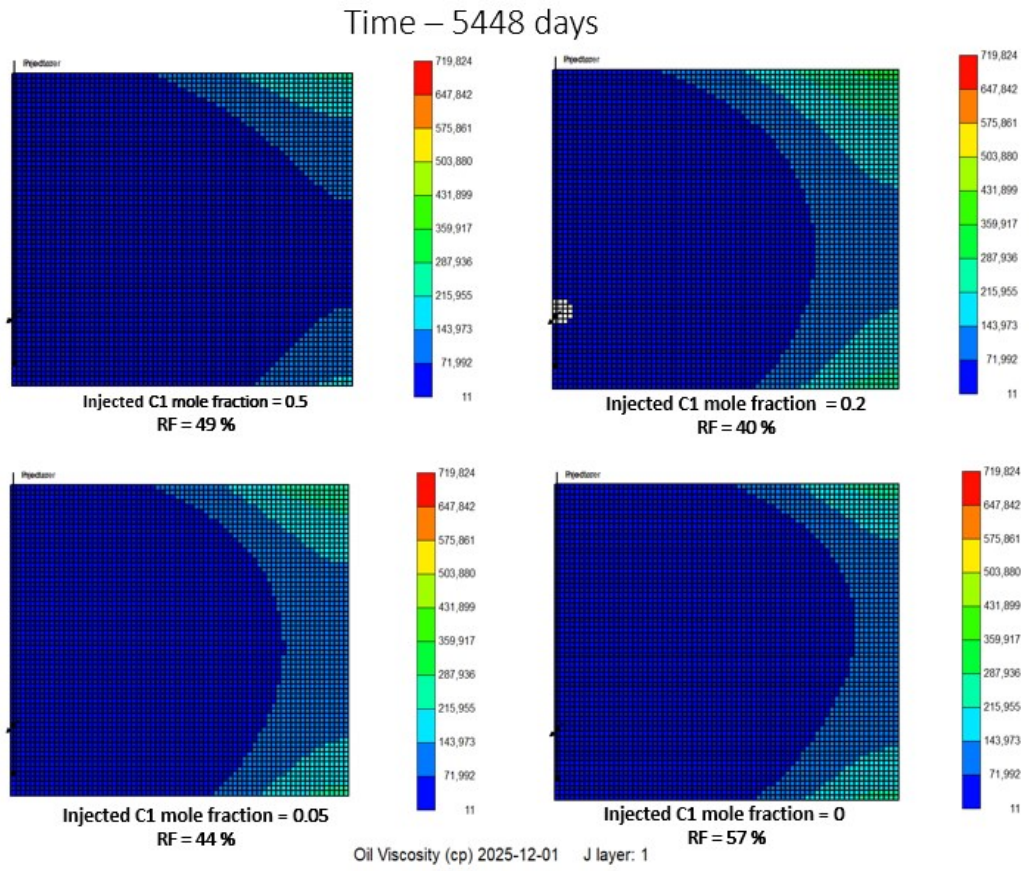


Figure 4-8: Effect of co-injecting C_1 on oil viscosity. Bottom-hole injection pressure is the dew point pressure at 50°C.

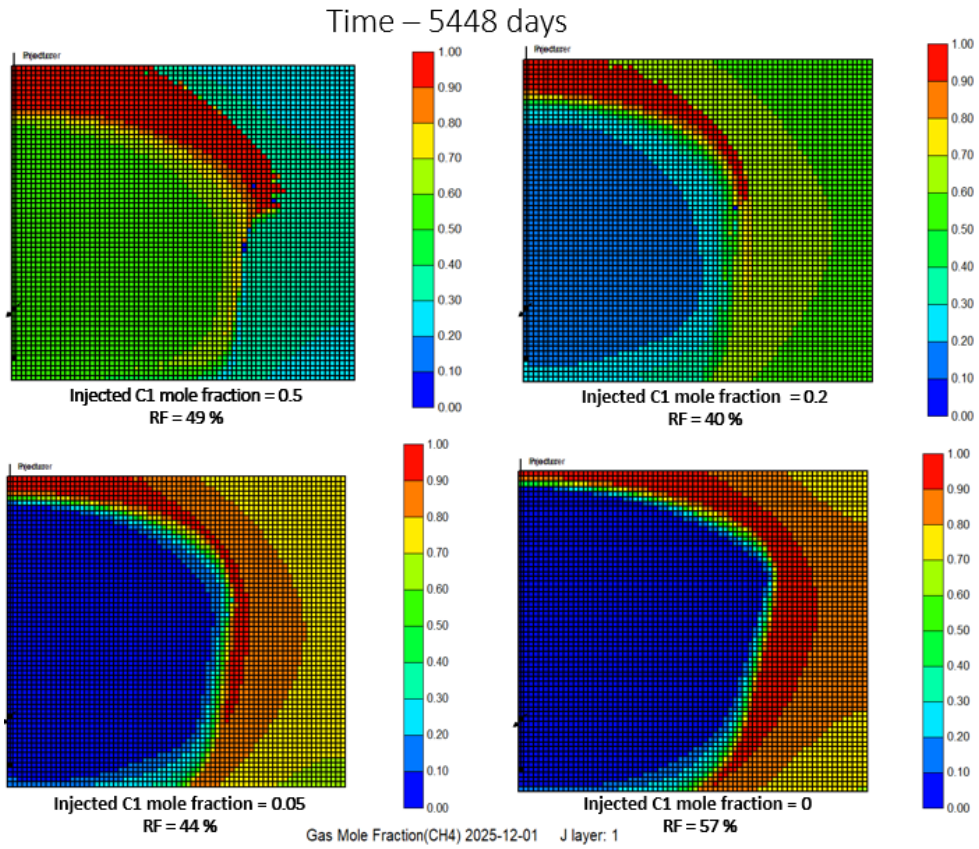


Figure 4-9: Effect of C_1 co-injection on gas-phase C_1 mole fraction. Bottom-hole injection pressure is the dew point pressure at 50°C.

The effect of C_1 co-injection on C_3 concentration in the oleic phase is also examined in Fig. 4-10. In addition to its impact on bottom-hole injection pressure and RF, there is a potential trade-off in the retention of C_3 in the oleic phase. C_1 is needed to keep the solvent in the gaseous phase; reducing C_1 concentration may cause more liquid C_3 to accumulate in the near-wellbore region and be retained. As shown in Fig. 4-10, more C_3 is accumulated in the oleic phase at the base of the reservoir because of gravity segregation. According to Fig. 4-11, for low dew point temperatures (< 35°C), the temperature is inversely related to *SolOR*, and the opposite trend is observed for high dew point temperatures (> 40°C). This may be due to increased oil production

as a result of the increase in injection (dew point) temperature. The results presented in Figs. 4-8 to 4-11 serve to illustrate the trade-offs among different objective functions of a given decision variable.

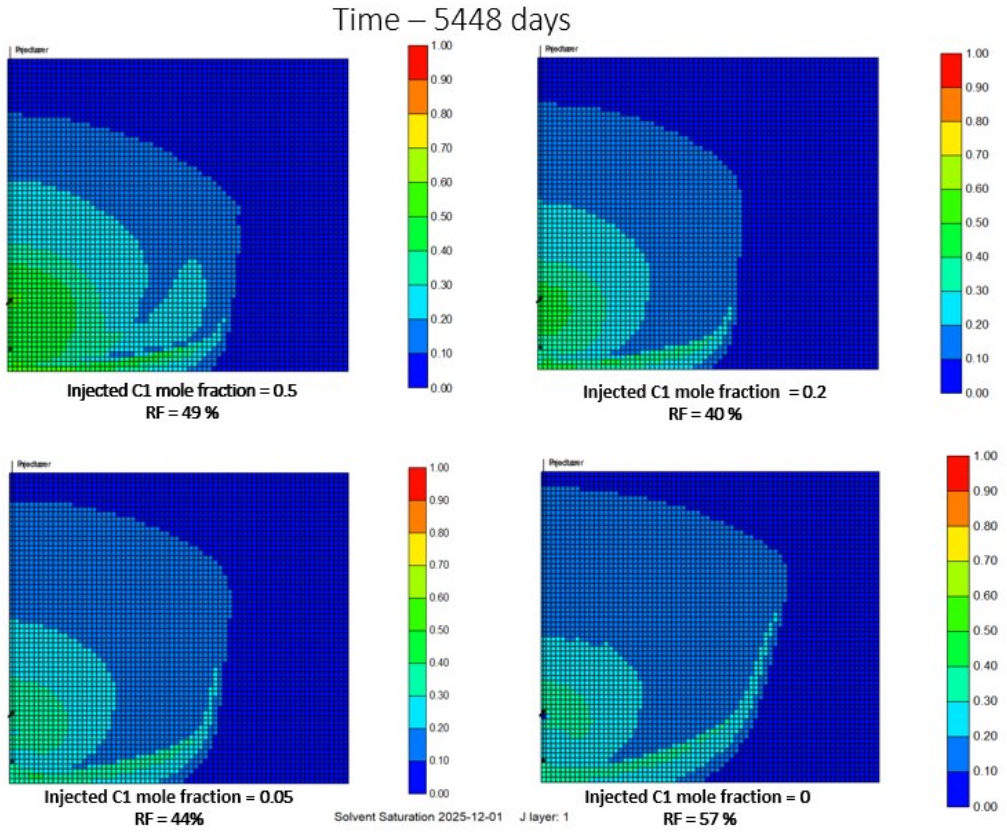
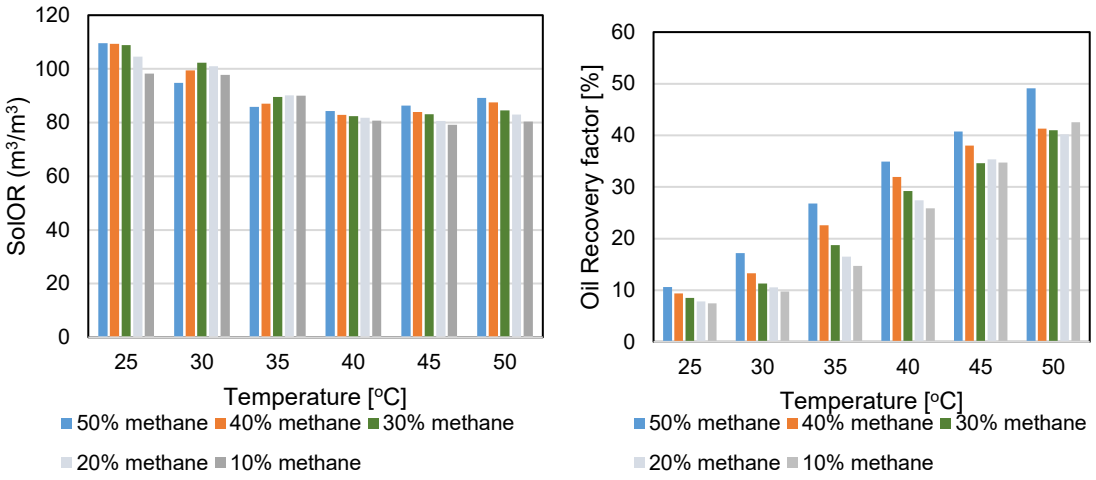


Figure 4-10: Effect of C_1 co-injection on C_3 concentration in the oleic phase. Bottom-hole injection pressure is the dew point pressure at 50°C.



(a) (b)

Figure 4-11: Effect of C_1 co-injection at different dew point conditions. (a) – SolOR; (b) – Oil Recovery Factor.

4.2.5.2. Injection at Superheated Conditions.

Generally speaking, more heat is added when injecting at superheated conditions for a given pressure, so a higher *RF* is expected. The results in Fig. 4-12 show that more C_3 remains in the extracted chamber at 30°C (i.e., more C_3 stays in the oleic phase and more likely to be trapped) than at 45°C, suggesting a decrease in solvent retention with increasing temperature. There is a reduction in solvent solubility at temperatures much higher than the dew point conditions of the C_1-C_3 mixture. The drawback of reduced solvent solubility may negate any positive impact of enhanced heat transfer at high temperatures. The trade-off between adding more heat and decreasing solvent solubility at higher injection temperatures may also lead to a reduction in *RF* (Fig. 4-13). However, this trend is also dependent on the injection (dew) pressure, as shown in Fig. 4-14.

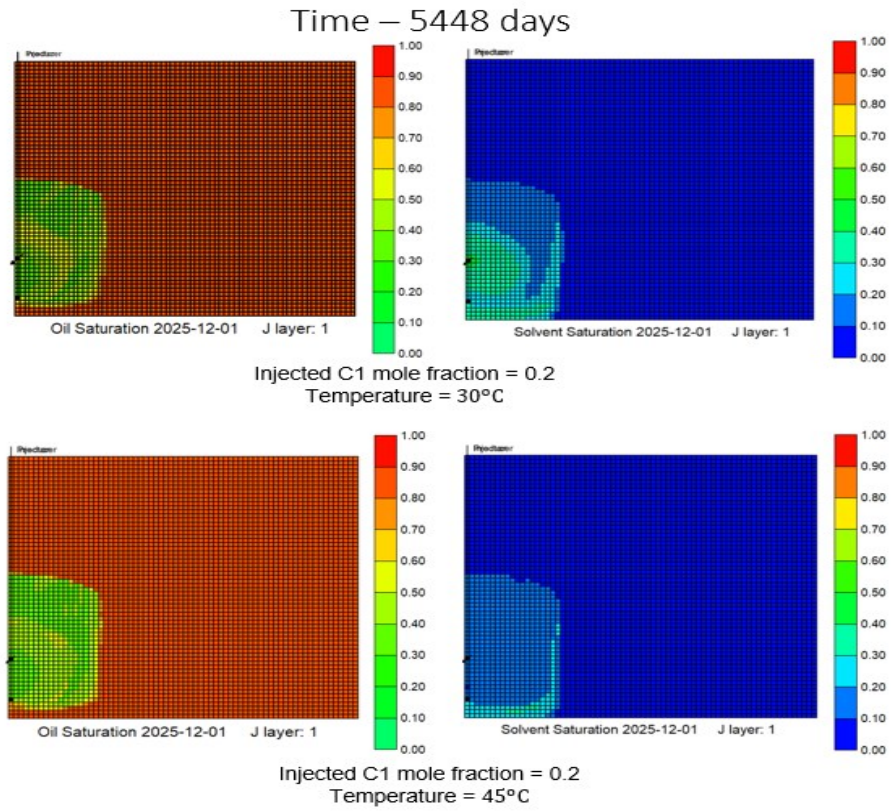


Figure 4-12: Effect of injecting at superheated conditions - BHP_{inj} = 1240 kPa.

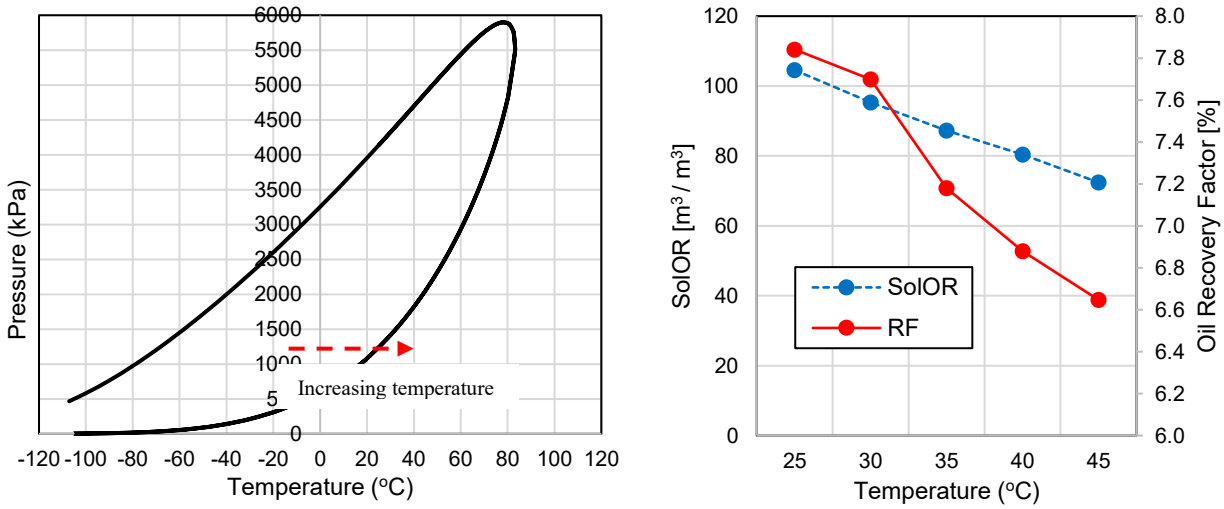


Figure 4-13: Injecting at superheated conditions - 20% methane, BHP_{inj} = 1240 kPa.

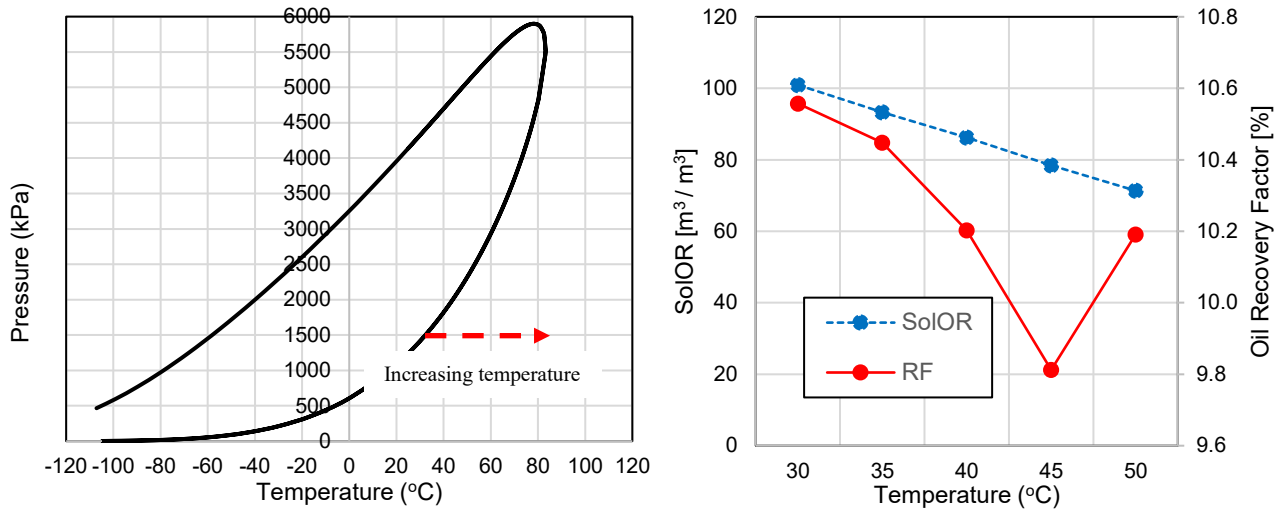


Figure 4-14: Injecting at superheated conditions - 20% methane, BHPinj = 1414 kPa.

In the end, the sensitivity analysis reveals that there are no simple (e.g., monotonic) relationships between the design variables and objective functions. A MOO analysis framework (e.g., multi-objective evolutionary algorithm or MOEA) is needed to infer a set of optimal operating parameters for the warm VAPEX process.

4.2.6. Pareto-Based Multi-Objective Optimization

The Non-Dominated Sorting Genetic Algorithm II (NSGA-II) is employed to search for a set of Pareto-optimal solutions that are uniformly spaced along the Pareto-optimal front (Deb et al., 2002). Pareto dominance and Pareto optimality are commonly used concepts in MOO problems. Consider a MOO problem with three decision variables, a decision vector $u = (u_1, u_2, u_3)$ is said to dominate (denoted by $\prec$) another vector $v = (v_1, v_2, v_3)$, if and only if

$$\forall_i \in \{1, 2, .. m\}, \qquad f_i\,(u) \leq f_i\,(v)$$
$$\text{and} \quad j \in \{1, 2, .. m\}: \quad f_j\,(u) < f_j\,(v), \tag{5}$$

where m is the total number of objectives, $f_i\,(u)$ or $f_j\,(u)$ and $f_i\,(v)$ or $f_j\,(v)$ are the objective functions for vectors u and v, respectively. It is implied that $f(u)$ is better than $f(v)$ for all objectives and there is at least one objective function (j) for which $f_j(u)$ is strictly better than $f_j(v)$ (Ngatchou et al. 2005).

4.2.6.1. NSGA-II

An initial parent population P_t is created randomly and sorted into sets of non-dominated solutions (fronts): best solutions with rank = 1; second-best solutions with rank = 2, etc. The offspring population Q_t with size N is generated from P_t through the operations of tournament selection, crossover, and mutation. Elitism is ensured by combining the parent population P_t of size N and the offspring population Q_t to form a combined population R_t of size 2N. Non-dominated sorting is also applied to the group R_t, and the new parent population P_{t+1} is filled with solutions from fronts with decreasing rank until a total size of N is reached. It is likely that when the last allowable front is considered, there are more solutions in that front than the remaining slots within P_{t+1}; in such instances, solutions in the least crowded region are selected to fill the slots. All remaining individuals in the combined population are discarded.

The new population P_{t+1} of size N is then used for subjected to another round of selection, recombination, and mutation to create a new offspring population Q_{t+1} , also of size N. The

selection of parents among P_t and P_{t+1} is based on the crowded comparison procedure, which is also used in the tournament selection step for selecting individuals from the last allowable front, as mentioned earlier. The crowded comparison operator, $\prec_c$ is defined as follows:

$$\begin{aligned} i \prec_c j \quad &\text{if } i_{rank} < j_{rank} \\ \text{or if } (i_{rank} = j_{rank}) &\text{ and } (i_{distance} > j_{distance}) \end{aligned} \tag{6}$$

where i_{rank} and j_{rank} are the ranks of individuals i and j, respectively, within the population; $i_{distance}$ and $j_{distance}$ are their crowding distances. The crowding distance, $i_{distance}$, of an individual i is the total side length of the cuboid formed by the neighboring solutions to the left and right of i. A 2D example (with 2 objectives) is illustrated in Fig. 4-15.

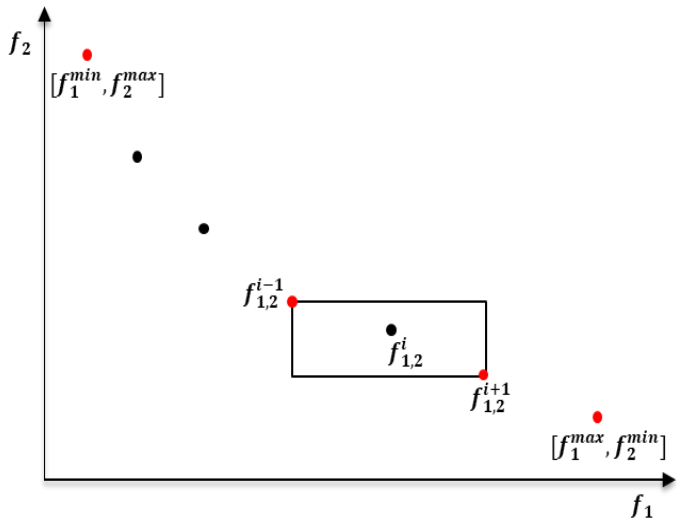


Figure 4-15: Schematic of crowding distance calculation.

Mathematically, for *m* objectives,

$$i_{distance} = \frac{\left|f_1^{i+1} - f_1^{i-1}\right|}{f_1^{max} - f_1^{min}} + \frac{\left|f_2^{i+1} - f_2^{i-1}\right|}{f_2^{max} - f_2^{min}} + \cdots \frac{\left|f_m^{i+1} - f_m^{i-1}\right|}{f_m^{max} - f_m^{min}}, \quad (7)$$

where f_m^{i-1} and f_m^{i+1} are the m^{th} objective function of the neighboring individuals $i-1$ and $i+1$, respectively, f_m^{min} and f_m^{max} are the population minimum and maximum values of the m^{th} objective function. Every individual within the population is composed of three decision variables, which are C_3 mole fraction, bottom-hole injection temperature, and bottom-hole injection pressure. In this study, the population size and the number of iterations are set to 200 and 50, respectively. It should be noted that the NSGA-II optimization is repeated several times with different initial guesses to ensure that a global minimum is reached.

4.2.6.2. Parameterization Scheme for MOO Implementation

There is an interdependency among the decision variables for this study (i.e., C_3 mole fraction, bottom-hole injection temperature, and bottom-hole injection pressure) since the solvent mixtures are to be injected at either the dew point or superheated conditions. This criterion implies that for a certain C_3 mole fraction, there is a particular "feasibility window" in the P-T diagram for the corresponding bottom-hole injection pressure and temperature. This constraint renders the MOO problem to be rather peculiar. Hence, a search space is identified, specified by the shaded portion in Fig. 4-16, to represent this feasibility window for a specific C_3 mole fraction. A novel parameterization scheme is formulated to represent any combinations of pressure and temperature within the search space for a particular C_3 mole fraction.

The temperature and pressure values are represented with two separate indicators (*T* and *P*), with values ranging between 1 and 10; the values are then mapped back to the search space to obtain the actual temperature and pressure for a certain C_3 mole fraction. Here are the steps:

1. Formulate each decision variable vector as C_3 mole fraction (x_{C3}), *T,* and *P*.

2. For the value of x_{C3}, determine P_{max} = upper limit of pressure along the dew point curve and P_{min} = dew point pressure at 30°C (the minimum injection temperature for the MOO). The mapped pressure P_{val} corresponding to indicator *P* is:

$$P_{val} = P_{min} + \left(\frac{P-1}{10-1}\right) \times (P_{max} - P_{min}) \tag{8}$$

3. The mapped temperature value T_{val} corresponding to indicator *T* is:

$$T_{val} = T_{sat} + \left(\frac{T-1}{10-1}\right) \times \Delta T \tag{9}$$

T_{sat} = saturation temperature at P_{val}. ΔT represents the temperature range of superheated conditions and is set to 20°C.

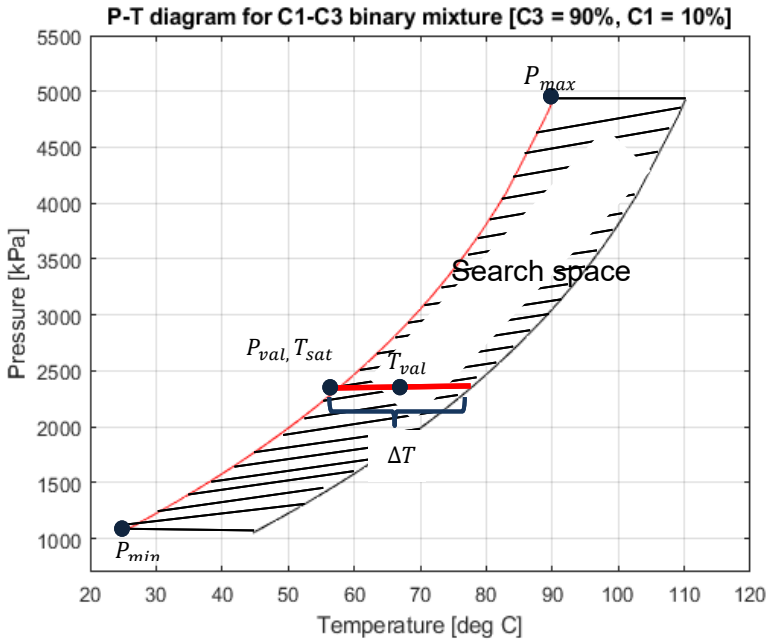


Figure 4-16: Search space for the implementation of NSGA-II (C_3 mole fraction = 0.9). Lower dew point line in red.

4.2.7. Proxy Modeling for Objective Function Evaluation

Evaluating the objective functions using the numerical simulation can be computationally expensive and inefficient; the total CPU run time for a single simulation execution could range from a few minutes to several days, depending on the model size and process complexity. Hence, a proxy model, which is calibrated using a limited number of simulation runs, can be used to compute the objective functions directly. In this study, the artificial neural network (ANN) technique is employed due to its ability to analyze complex relational behaviors between a set of input variables (e.g. the thermodynamic properties of solvent mixtures) and their outputs (e.g., objective functions). ANN has been applied to a wide variety of reservoir engineering problems such as production forecast (Kubota and Reinert, 2019; Cao et al., 2016), optimization of warm

solvent injection processes (Ma and Leung, 2020a), investigation of reservoir heterogeneities in SAGD (Wang et al., 2018; Zheng et al., 2018a; Wang and Leung, 2015), and the analysis of SAGD production performance (Ma et al., 2015, Ma et al., 2017). Zheng et al. (2018a, b) developed an ANN-based proxy model to predict both oil production rate and steam-oil-ratio for heterogeneous reservoirs with different shale barrier configurations, and the same proxy was later combined with a genetic algorithm-based optimization scheme for inferring the unknown shale barrier configurations from the time-series production data.

An ANN architecture is usually composed of an input layer, one or more hidden layers, and an output layer. Each layer usually consists of a certain number of neurons or nodes and a bias unit. For the input layer, the neurons consist of the input parameters or features (i.e., decision variables in this study), while the output neurons include the target variables (i.e., the objective functions in this study). The output of node j in layer $k+1$ is given by a_j^{k+1}:

$$z_j^{k+1} = b + \sum_{i=1}^{n} w_{j,i}^k . x_i^k, \tag{10}$$

$$a_j^{k+1} = g\left(z_j^{k+1}\right) \tag{11}$$

where b is the bias; $w_{j,i}^k$ is the weight of node i in layer k ; x_i^k refers to the input value of node i in layer k; z_j^{k+1} is the weighted sum of the input values for node j in layer $k+1$, g is the activation function. The weights can be estimated through a backpropagation update algorithm based on a variety of techniques such as gradient descent, Adaptive Moment Estimation (Adam), and Levenberg-Marquardt, where the mismatch in model predictions (e.g., the mean square error, MSE, or other cost functions) is minimized after several batches of training. In this study, the

ANN-based proxy model is constructed using the Neural Network Toolbox in Matlab™ (Beale, et al., 2018), and the network properties are shown in Table 4-6. To generate a set of flow simulation cases for training the proxy model, the following factorial experimental design is performed: there are 9 equally-spaced levels for C_3 mole fraction, and for each level, 6 and 5 equally-spaced levels of bottom-hole injection temperature and bottom hole injection pressure are used, respectively. Therefore, a total of 270 cases are used for training and testing.

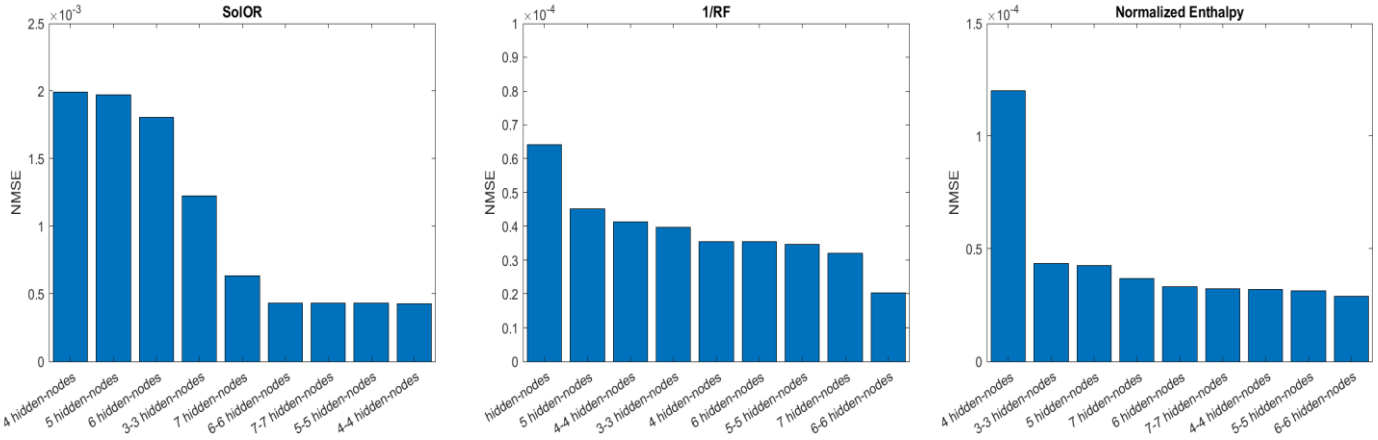
Table 4-6: Network properties

Property	
Activation function (hidden layer)	Tan-Sigmoid
Activation function (target)	Linear
Training function	Levenberg-Marquardt
Performance function	Mean Square Error
Maximum epochs	1000

Three objective functions are defined: *SolOR*, 1/*RF*, and normalized enthalpy injected (representing the total energy requirement). A separate ANN model is constructed for each objective function. For the objective function of 1/*RF*, the network structure has 2 hidden layers with 6 nodes; for the objective function of *SolOR*, the network consists of 2 hidden layers (each with 4 nodes); finally, the network for normalized enthalpy has 2 hidden layers (each with 6 nodes). The ANN architecture is obtained after a 10-fold cross-validation process, and the network structure with the lowest normalized mean squared error (NMSE) is selected. The results of the cross-validation process are shown in Table 4-7 and Fig. 4-17.

Table 4-7: 10-fold cross validation

Number of nodes per hidden layer (s)	4	5	6	7	3-3	4-4	5-5	6-6	7-7
NMSE (SolOR)	1.99×10^{-3}	1.97×10^{-3}	1.81×10^{-3}	6.36×10^{-4}	1.23×10^{-3}	4.27×10^{-4}	4.30×10^{-4}	4.31×10^{-4}	4.31×10^{-4}
NMSE (1/RF)	3.55×10^{-5}	4.51×10^{-5}	3.55×10^{-5}	3.21×10^{-5}	3.96×10^{-5}	4.14×10^{-5}	3.47×10^{-5}	2.03×10^{-5}	6.41×10^{-5}
NMSE (Normalized Enthalpy Injected)	1.20×10^{-4}	4.27×10^{-5}	3.32×10^{-5}	3.68×10^{-5}	4.35×10^{-5}	3.20×10^{-5}	3.15×10^{-5}	2.89×10^{-5}	3.23×10^{-5}



(a) (b) (c)

Figure 4-17: Cross validation results. (a) – SolOR; (b) – 1/RF; (c) – Normalized Enthalpy Injected.

Typically, the test split ratio for machine learning algorithms ranges between 10 – 50 % (Afendras and Markatou, 2019; Racz, Bajusz and Heberger, 2021; Ma and Leung, 2020a). To determine the optimal train-test split, 3 different ratios (85/15%, 75/25%, and 70/30%) are examined, and for all the ratios, the ANN model predictions of the objective functions are highly consistent with those generated using reservoir simulation – no significant change in the coefficient of determination (R^2) is observed (i.e., R^2 = 0.989, 0.993, and 0.992 respectively). Therefore, it is concluded that

any of the tested train-test split ratios could be used. Out of the 270 cases, 203 cases (75%) are used for training and validation, and 67 cases (25%) are employed for testing purposes.

4.3. Results and Discussion

4.3.1. Proxy Modeling and MOO Results – 2 Objective Functions

The performance of the proxy models is assessed by comparing the objective function values computed from reservoir simulation with the proxy's predicted values for the 67 test cases, as shown in Fig. 4-18. For both the *SolOR* and *1/RF* prediction, the normalized mean squared error is low (i.e. approximately 0.384 and 0.011, respectively), while their coefficients of determination (R^2) are close to 1, indicating that the network predictions closely match the target values from flow simulation.

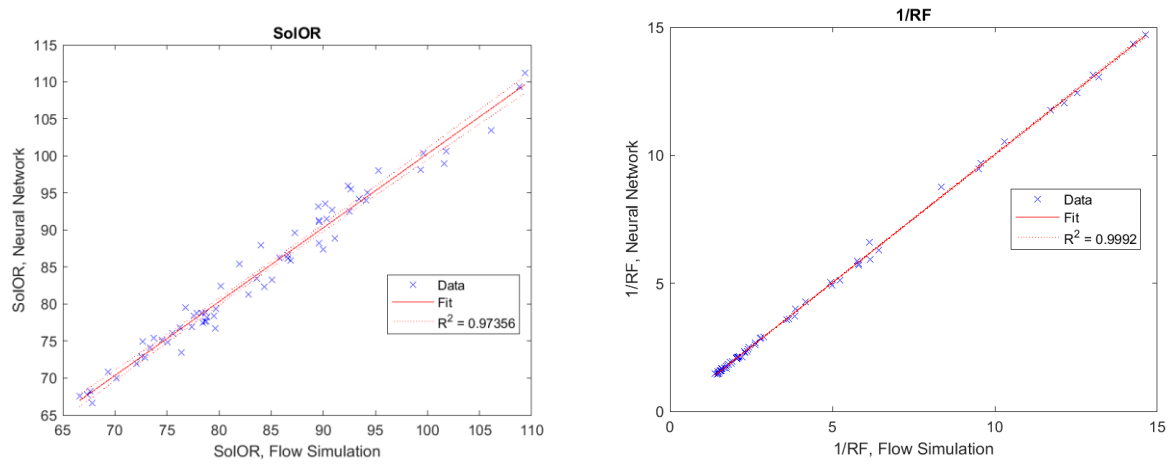


Figure 4-18: Comparison of proxy model predicted values and flow simulation output. (a) – SolOR; (b) – 1/RF.

The results of the 2-objective MOO approach are shown in Figs. 4-19 to 4-21. *N* is 200 and a front that represents the trade-off between solvent retained-oil ratio and oil recovery is obtained after the maximum number of iterations (i.e., 50 iterations) is reached, the total computing time is approximately 7 minutes. The initial population has yielded extremely high values for at least one of the objectives (1/*RF*), and as the number of iterations increases, a set of Pareto-optimal solutions emerge (Fig. 4-19). Fig. 4-20 compares the distributions of decision parameters for the initial population with those corresponding to the most optimal solutions (inside the red circle in Fig. 4-19 – *SolOR*: 70.4 to 79.1 and $1/RF$: 1.68 to 2.28). All solutions along the front are deemed to be Pareto optimum. Among the solutions, it is possible to select a subset (highlighted by a red circle in Fig. 4-19) for which 1/*RF* and *SolOR* are the lowest. A couple of remarks should be noted: (1) The selection of this optimal subset of solution can be subjective. (2) Such a selection is more obvious if the Pareto front resembles an "*L*-shaped" function. For the initial population, the decision variables are spanning over wide ranges, while the optimal design entails injecting solvent mixtures with high C_3 mole fraction (~0.9) at relatively lower pressures and high temperatures. To check the accuracy of the proxy model, three optimal solutions (i.e. [(x_{C3}), *T,* and *P*] = [0.85, 2203 kPa, 71°C], [0.87, 2366 kPa, 76°C] and [0.9, 2671 kPa, 83°C]), with corresponding objective functions ([*1/RF, SolOR*]) [2.25, 70.93], [2.04, 73.11], and [1.75, 77.35], respectively, are randomly selected. The objective function calculations are then repeated using the more accurate flow simulation, and the corresponding values are [2.35, 71.13], [2.01, 73.09], and [1.71, 78.34], reflecting errors of 4%, 1.5%, and 2% respectively for *1/RF* and 0.3%, 0.03% and 1.3% for *SolOR*. This check shows that the proxy model can closely reproduce the flow simulation predictions, confirming that it can be used to replace the more expensive flow simulation calculations for evaluating the objective functions.

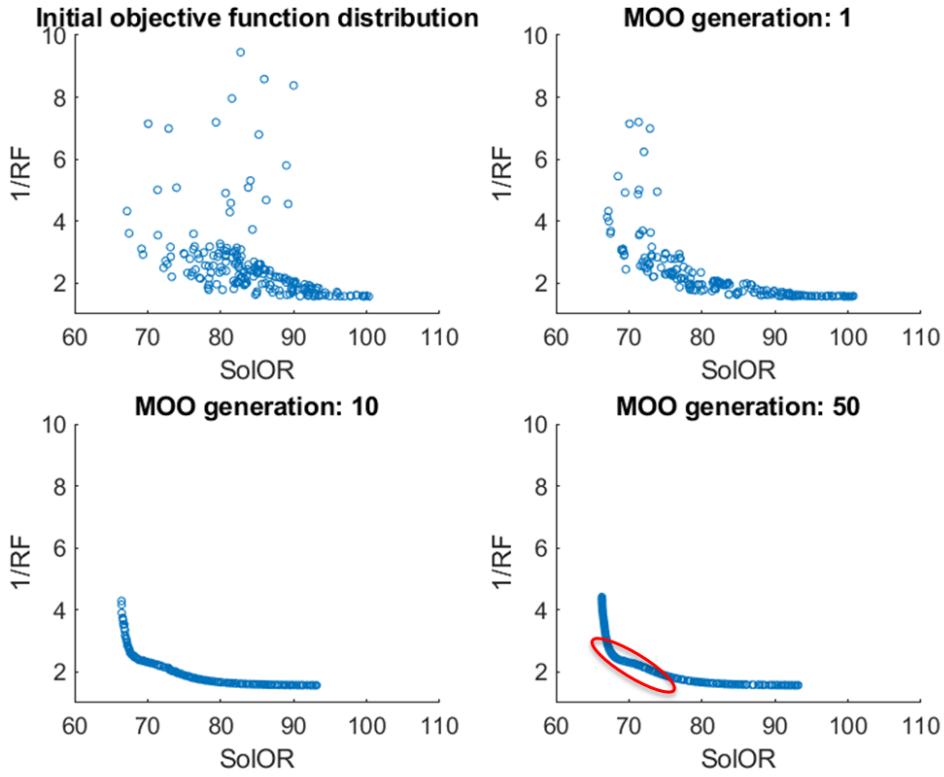


Figure 4-19: Movement of the Pareto front. Optimal trade-off solutions are enclosed by the red circle.

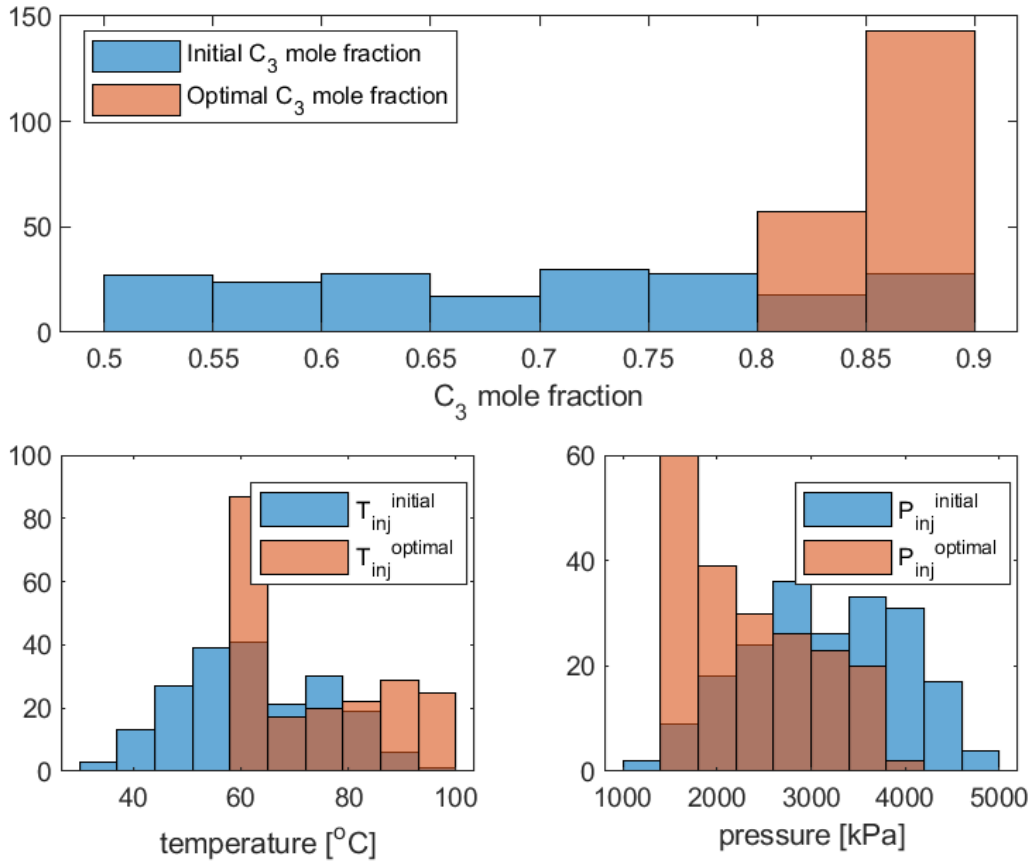
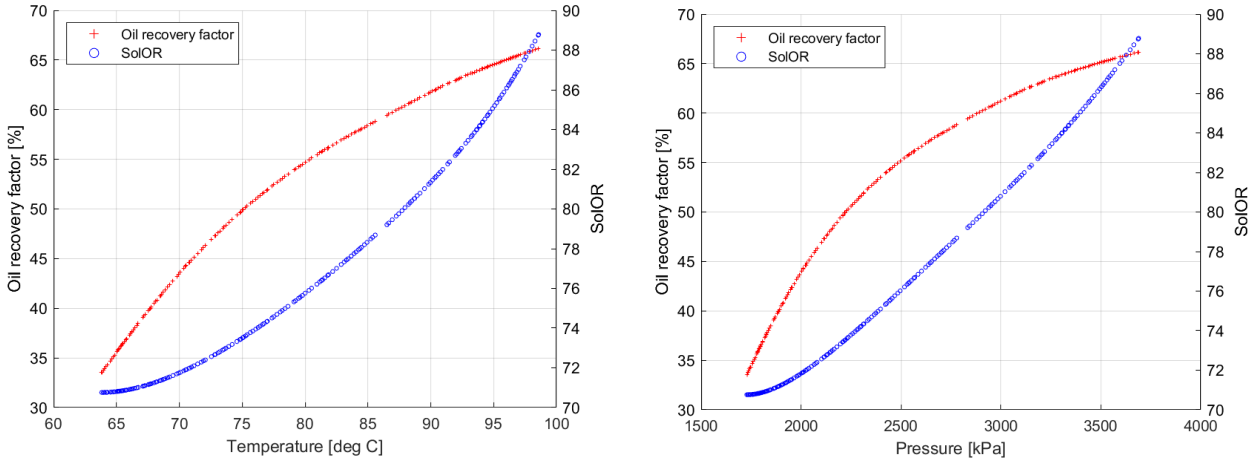


Figure 4-20: Distribution of decision variables.

With 2 objective functions and 3 design variables, a multi-dimensional space is needed to represent the interactions among all the objectives and variables. Fig. 4-21 shows the projection of that space along the temperature and pressure axes. Both oil recovery and *SolOR* increase with bottom-hole injection pressure and temperature. For the optimum C_3 mole fraction, any improvement in oil recovery due to an increase in temperature and pressure also leads to increased C_3 retention. Therefore, to achieve a balance between these two objectives, the optimal solutions depict a combination of low to medium bottom-hole injection pressure and high bottom-hole injection

temperature. A summary of the Pareto-optimal solutions for the 2-objective MOO is shown in Table 4-8.



(a) (b)

Figure 4-21: Optimal decision variables and their corresponding objective functions. (a) - Temperature; (b) – Pressure.

Table 4-8: Optimal trade-off solutions

$SolOR\ \left(m^3/m^3\right)$	$1/RF$	RF (%)	C_3 mole fraction	Pressure (kPa)	Temperature (°C)
[70.4 79.1]	[1.68 2.28]	[43.8 59.4]	[~0.90]	[2204 2827]	[69.9 85.6]

One of the drawbacks of this 2-objective MOO analysis is the omission of energy requirement as an objective: high bottom-hole injection temperature and pressure implies an undesirable increase in energy consumption; hence another MOO is performed considering an additional objective function of energy requirement, which can be minimized along with *SolOR* and $1/RF$.

4.3.2. Proxy Modeling and MOO Results – 3 Objective Functions (Considering Energy Requirement)

The third objective, called Normalized Enthalpy is defined as:

$$H_{inj}^{normalized} = \frac{H_{inj}}{H_{inj}^{max}} \tag{12}$$

where H_{inj} is the cumulative enthalpy of the injected solvent (in Joules), as shown in the following equations; H_{inj}^{max} is the maximum H_{inj} among the 270 cases used for proxy modeling.

$$H_{inj} = U_{inj} + \mathrm{P}_{inj}\mathrm{V}_{inj}, \tag{13}$$

where U_{inj} denotes the internal energy of injected fluid, P_{inj} is the injection pressure, and V_{inj} refers to the volume of fluid injected. U_{inj} for a C_1-C_3 binary mixture is expressed as follows:

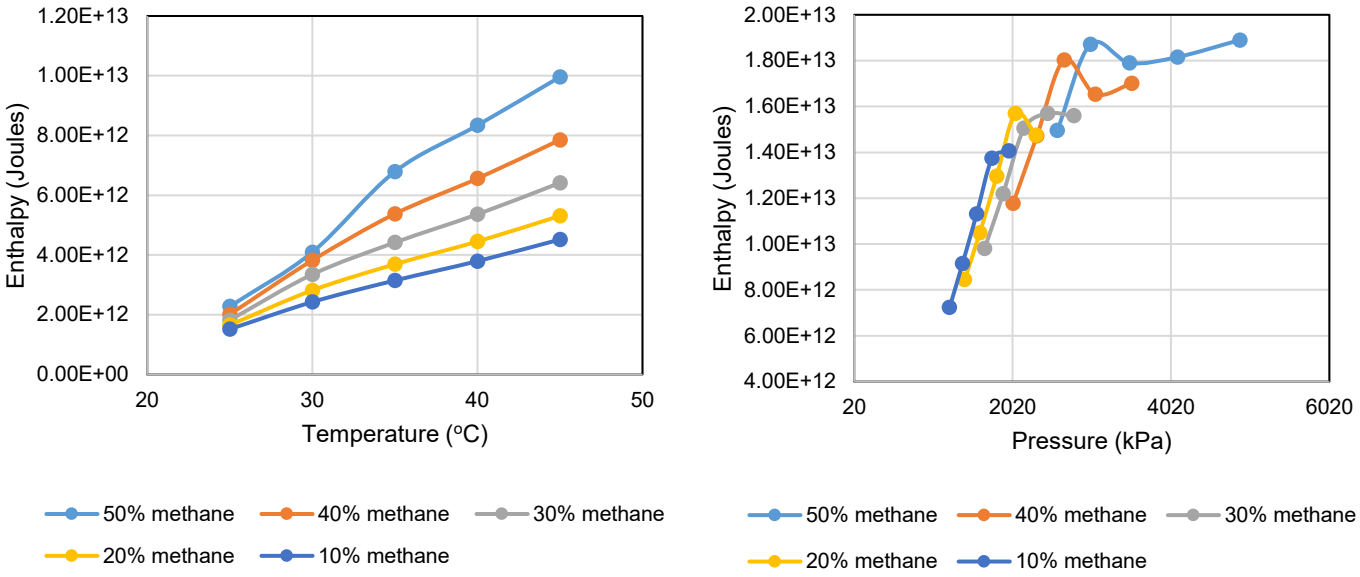
$$U_{inj} = T_{inj}S - \mathrm{P}_{inj}\mathrm{V}_{inj} + \mu_{c1}n_{c1} + \mu_{c3}n_{c3}, \tag{14}$$

where T_{inj} refers to injection temperature, S denotes entropy, μ_{c1} and μ_{c3} are the chemical potentials for C_1 and C_3, respectively, and n_{c1} and n_{c3} are the number of moles for C_1 and C_3 respectively. Therefore,

$$H_{inj} = T_{inj}S - \mathrm{P}_{inj}\mathrm{V}_{inj} + \mathrm{P}_{inj}\mathrm{V}_{inj} + (n_{c1} + n_{c3})\left[\mu_{c1}\frac{n_{c1}}{n_{c1}+n_{c3}} + \mu_{c3}\frac{n_{c3}}{n_{c1}+n_{c3}}\right], \tag{15}$$

where $\frac{n_{c1}}{n_{c1}+n_{c3}}$ and $\frac{n_{c3}}{n_{c1}+n_{c3}}$ are the mole fractions of C_1 and C_3 respectively. Therefore, there is a direct relationship between enthalpy and the three decision variables (i.e.., C_3 mole fraction,

bottom-hole injection temperature, and bottom-hole injection pressure), which are illustrated in Fig. 4-22.



(a) (b)

Figure 4-22: Enthalpy as a function of temperature and pressure. (a) – Temperature; (b) - Pressure.

For the 67 test cases, a comparison of the proxy model predictions of $H_{inj}^{normalized}$ values with reservoir simulation values is presented in Fig. 4-23. Compared to reservoir simulation, the mean squared error is 1.61×10^{-4}, while R^2 is approximately equal to 1. The final Pareto-optimal solutions after 50 iterations are shown in Figs. 4-24 to 4-25. The entire Pareto-optimal set can be divided into three distinct groups. For the entire solution set, the optimal C_3 mole fraction is about 0.9.

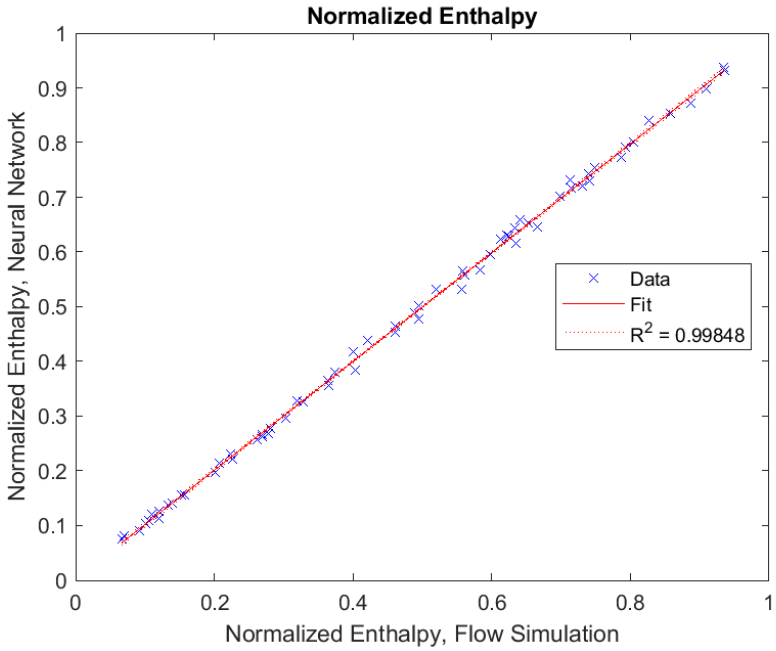


Figure 4-23: Comparison of proxy model predicted values and flow simulation output.

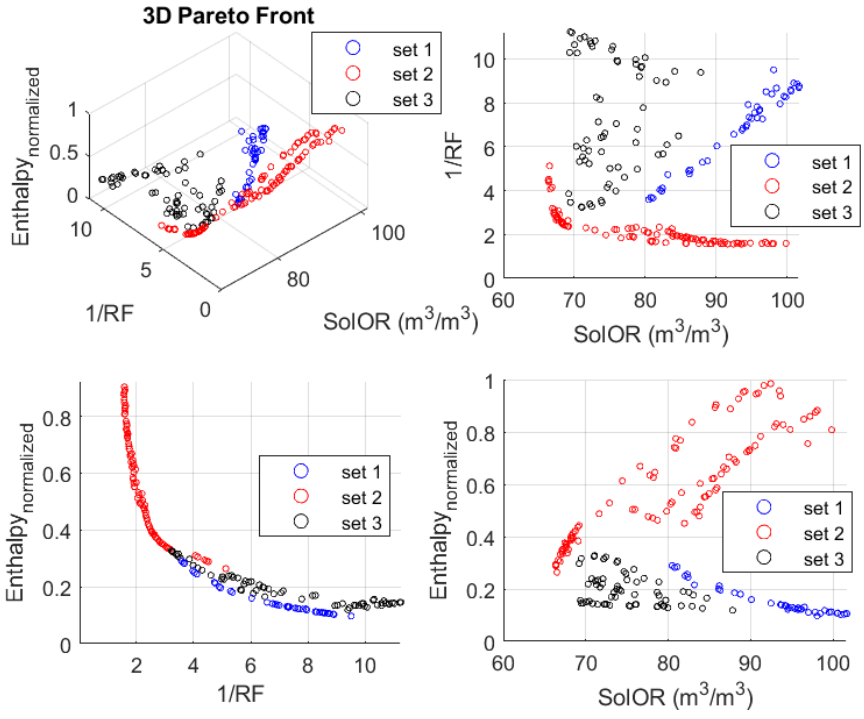


Figure 4-24: Pareto front for the 3-objective MOO.

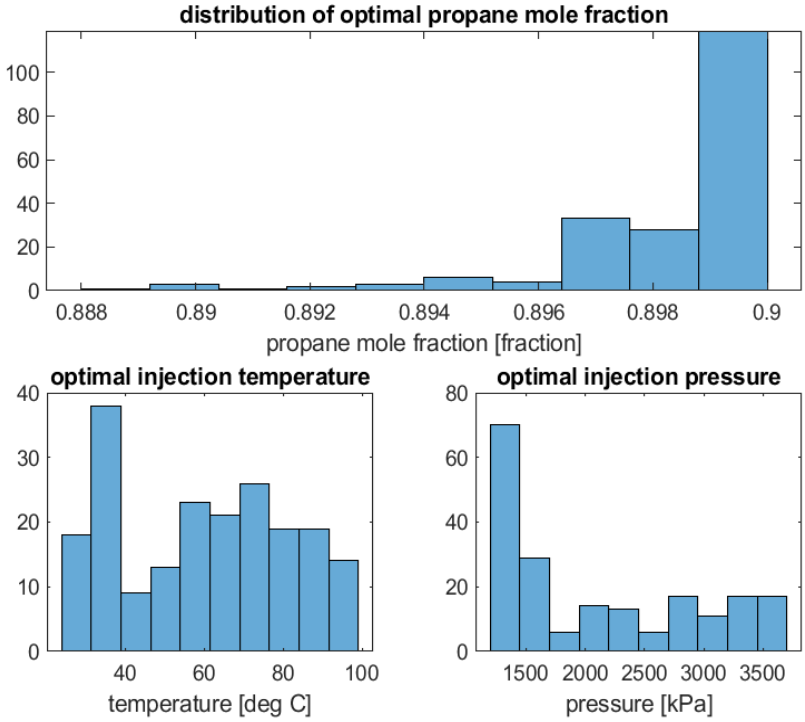


Figure 4-25: Distribution of optimal decision variables for all 3 clusters.

Individuals in solution set 1 exhibit low enthalpy due to low bottom-hole injection temperature and pressure. A positive correlation exists between 1/*RF* and *SolOR*; an increase in bottom-hole injection pressure and temperature would result in higher *RF* and less *SolOR* (although more C_3 stays in the liquid phase inside the chamber, as shown in Fig. 4-26). It is also plotted in Fig. 4-27, the dew point pressures corresponding to all the optimal temperatures when C_3 mole fraction = 90%. The conclusion is that these solutions correspond to scenarios where the solvent is injected near its dew point conditions at low temperatures and low pressures; the enthalpy requirement is low, and *RF* may not be high. There is significant solvent retention in the reservoir as liquid condensation occurs quickly in the extracted chamber before much C_3 can propagate towards the solvent-bitumen interface (Fig. 4-26).

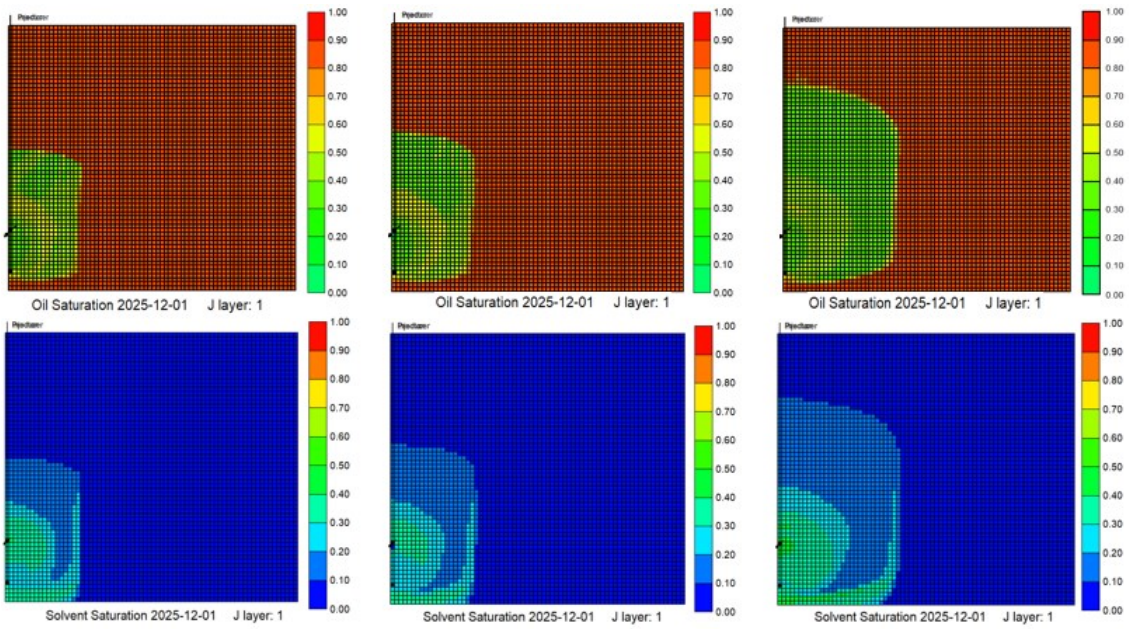


Figure 4-26: The impact of optimal decision variables for the solution set 1 on propane loss.

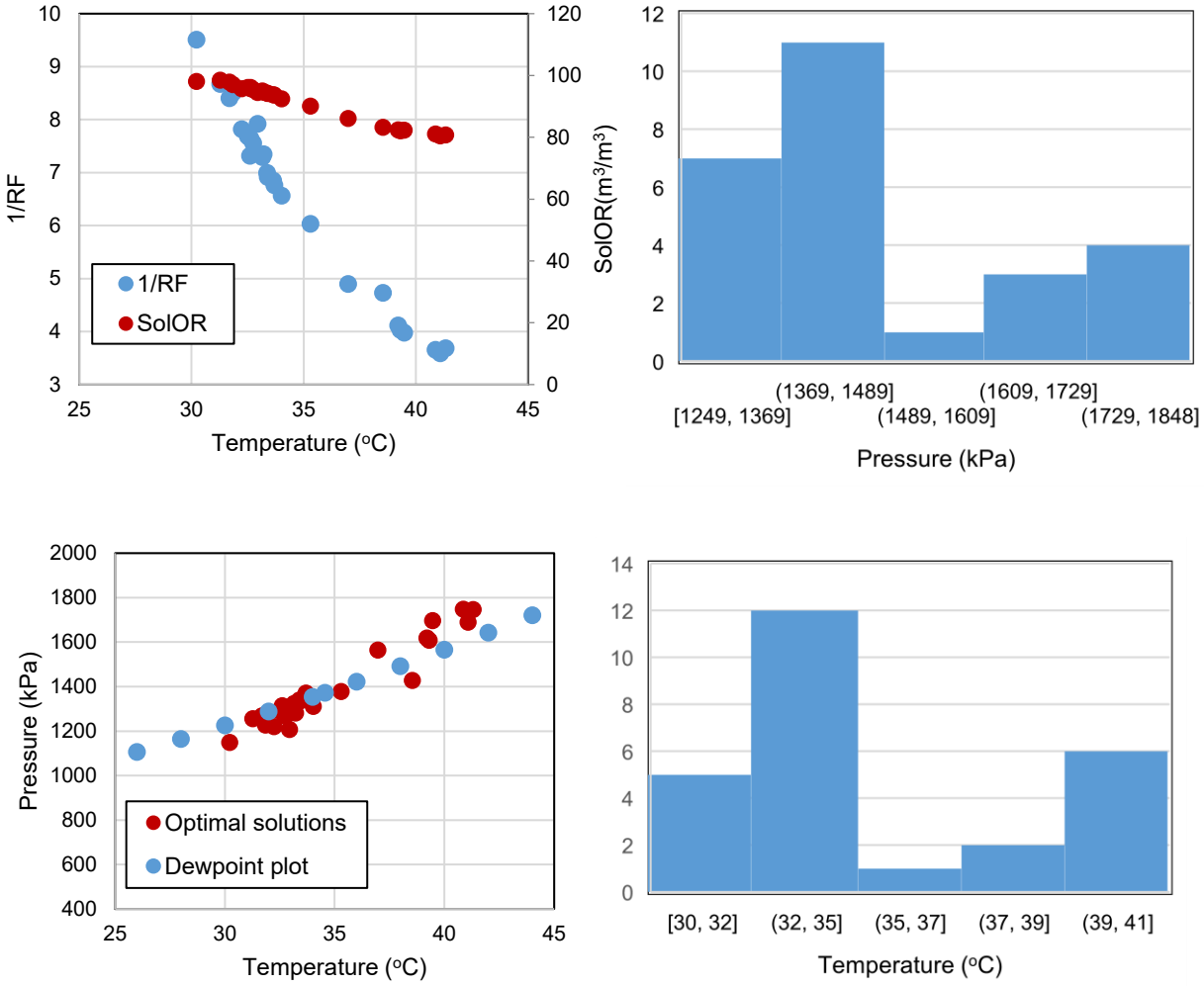


Figure 4-27: Solution set 1. (a) – optimal bottom-hole injection temperature and pressure and their corresponding objectives; (b) - distributions of optimal decision variables.

In contrast to solution set 1, solution set 2 exhibits a tradeoff between *SolOR* and the 1/*RF*, as well as a direct relationship between *SolOR* and enthalpy. An increase in bottom-hole injection temperature and pressure would result in higher *RF* but more *SolOR* (i.e. more C_3 loss as demonstrated in Fig. 4-28. The conclusion is that these solutions correspond to scenarios where

the solvent is injected at superheated conditions and high pressures; *RF* is high, while solvent retention and enthalpy requirement may also be high (see Fig. 4-29).

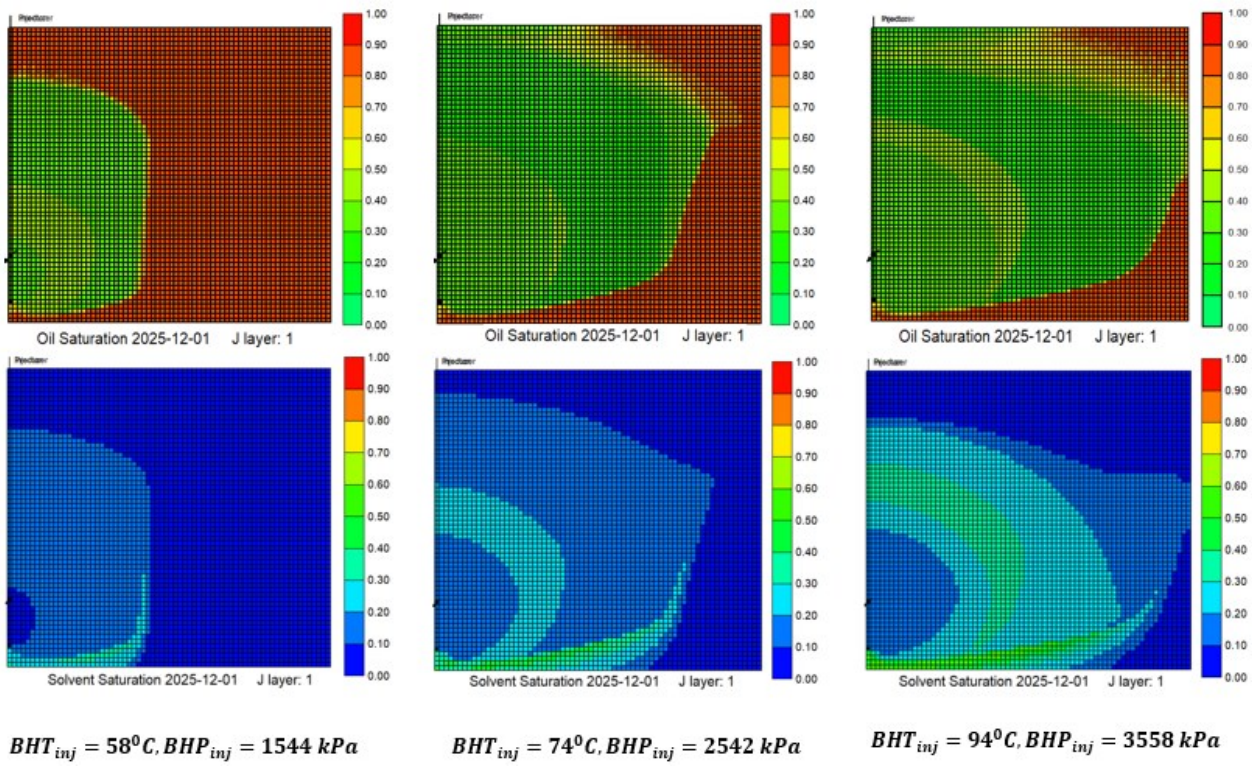


Figure 4-28: The impact of optimal decision variables for the solution set 2 on propane loss.

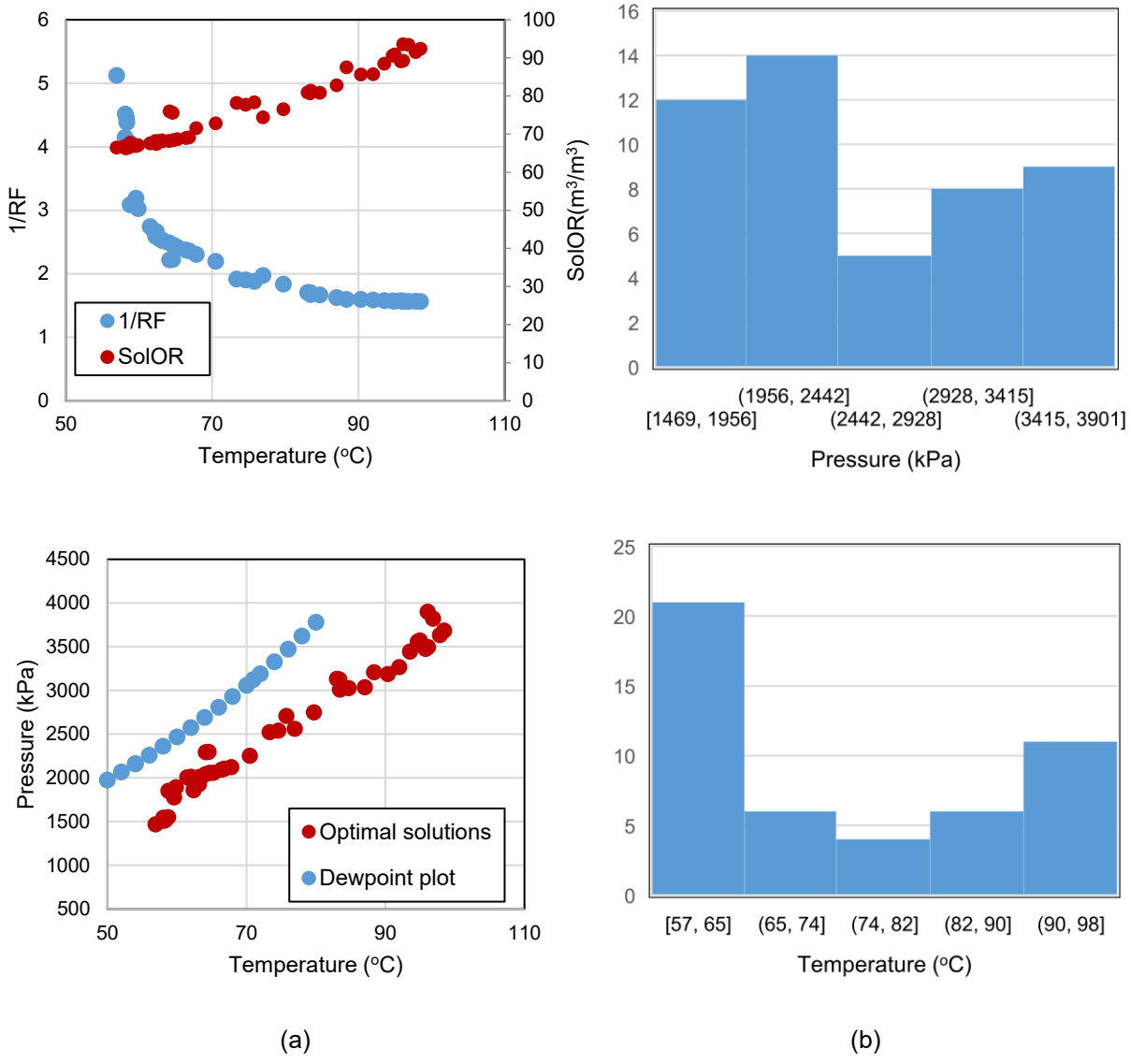


Figure 4-29: Solution set 2: (a) – optimal bottom-hole injection temperature and pressure and their corresponding objectives; (b) - distributions of optimal decision variables.

Individuals in solution set 3 comprise cases with lower bottom-hole injection pressures and temperatures compared to set 2. The trends among different objectives are similar to those of solution set 1, except that the values for *SolOR* are lower; however, this reduction in *SolOR* is

accompanied by an increase in enthalpy requirement and a reduction in $1/RF$ (see Figs. 4-30 to 4-32).

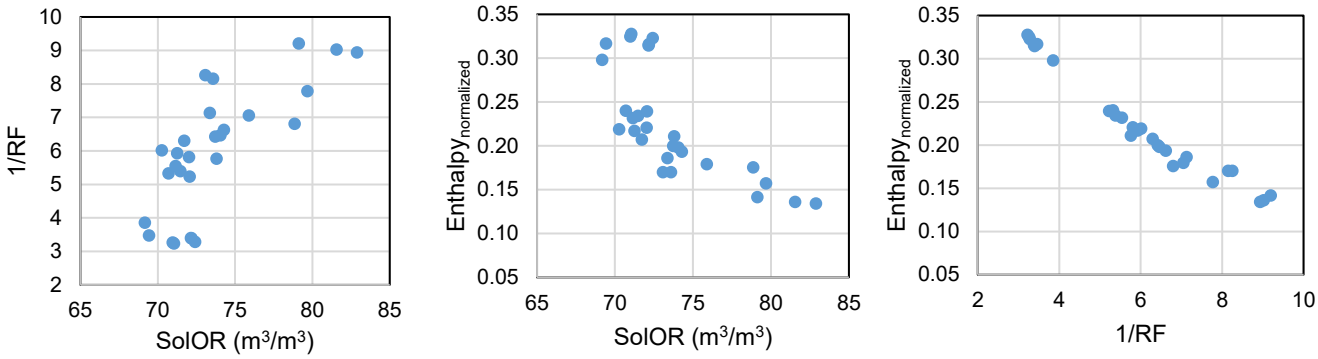


Figure 4-30: Solution set 3.

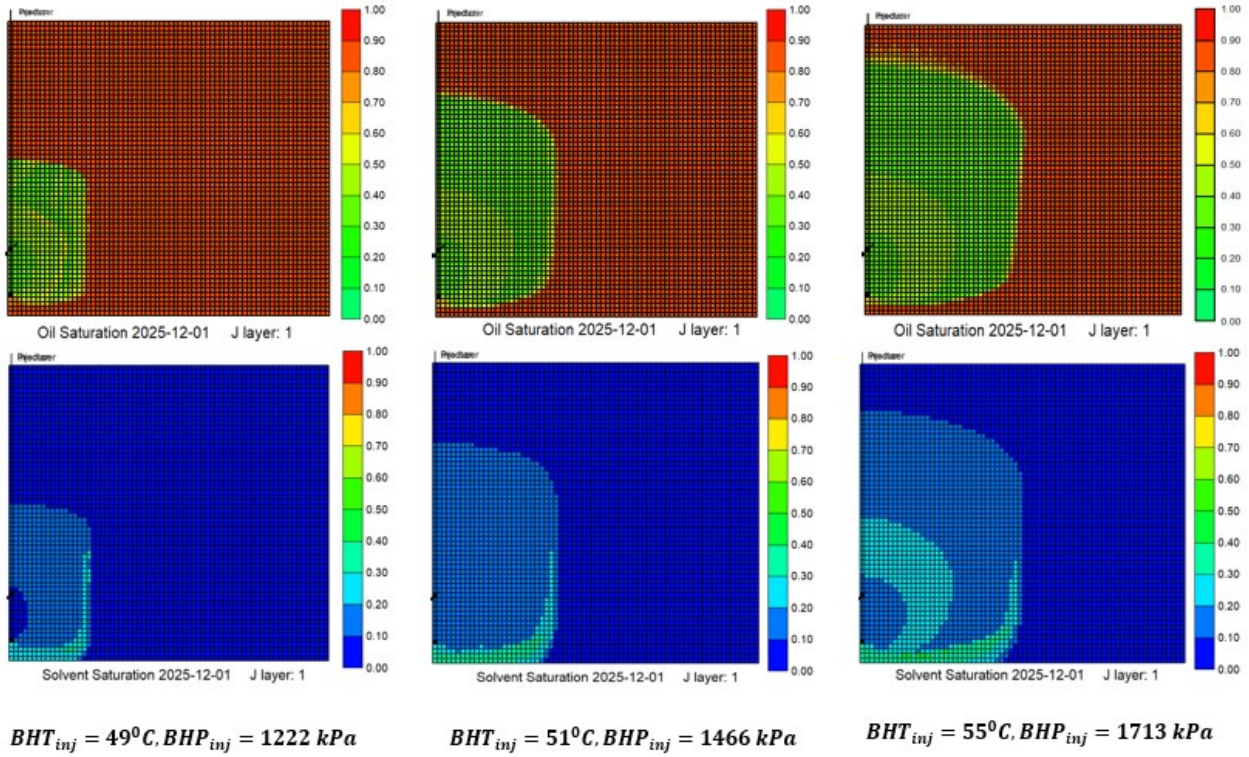


Figure 4-31: The impact of optimal decision variables for the solution set 3 on propane loss.

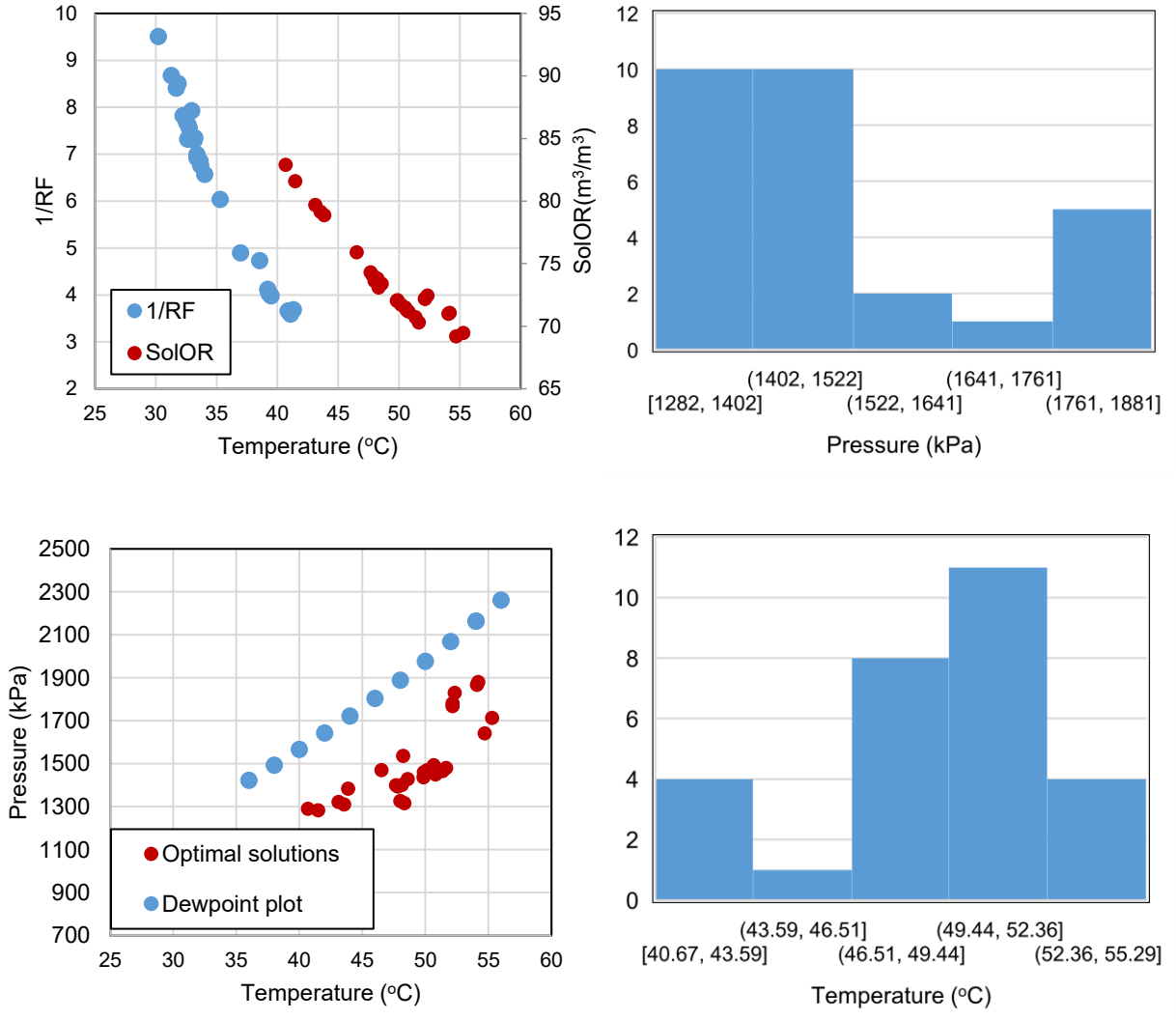


Figure 4-32: Solution set 3. (a) – optimal bottom-hole injection temperature and pressure and their corresponding objectives; (b) - distributions of optimal decision variables.

The computational requirements for the MOO workflow with proxy modeling and without proxy modeling (i.e., running flow simulation for every objective function evaluation) are compared in Table 4-9. There is substantial front-end loading for the proxy modeling (> 8,000 minutes); however, this training needs to be performed only once. The subsequent NGSA-II calculation takes only 20 minutes for a run with 2 objective functions or 30 minutes for a run with 3 objective

functions, while if the flow simulation is used to evaluate each objective function, 300,005 minutes are needed for each NGSA-II run. As mentioned in section 4.2.6.1, the entire NGSA-II step should be repeated multiple times with different initial guesses when searching for the global minimum. It is clear that when multiple NGSA-II runs are required, the extra costs associated with the proxy modeling can be further justified.

Table 4-9: Computational requirement comparison

Steps	**MOO with Proxy Model**	**MOO with Flow Simulation (no Proxy Model)**
Proxy Modeling	Constructing Training and Testing Dataset: 270 flow simulation cases × 30 min/case = 8,100 minutes Training of Proxy (ANN) Model = 1.5 minutes	0 (N/A)
NGSA-II (2-objective functions)	0.001 min per objective function evaluation × 2 objective functions × 200 individuals in a population × 50 iterations = 20 minutes	30 min per simulation run (objective function evaluations) × 200 individuals × 50 iterations = 300,000 minutes
	Rest of NGSA-II code = 5 minutes	Rest of NGSA-II code = 5 minutes
NGSA-II (3-objective functions)	0.001 min per objective function evaluation × 3 objective functions × 200 individuals in a population × 50 iterations = 30 minutes	30 min per simulation run (objective function evaluations) × 200 individuals × 50 iterations = 300,000 minutes
	Rest of NGSA-II code = 5 minutes	Rest of NGSA-II code = 5 minutes

Total Time	~8,125 minutes (for MOO with 2 objective functions) or ~8,135 minutes (for MOO with 3 objective functions)	~300,005 minutes (for MOO with either 2 or 3 objective functions)

The optimal operating parameters would be sensitive to reservoir heterogeneities; in other words, the final Pareto-optimal solutions may differ if geologic heterogeneities are considered. The presence of shale baffles and barriers may impede solvent chamber propagation and/or lead to the accumulation of liquid solvent in the reservoir, thereby causing a reduction in bitumen recovery and an increase in the amount of trapped solvent in the reservoir. Future work will extend the current workflow to include multiple realizations or models of heterogeneous reservoir properties; the goal is to identify a set of decision variables that are most robust/optimal against all realizations. Compared to a previous study where only pure C_3 was injected (Ma and Leung, 2020a), the optimal bottom-hole injection pressure here exhibits more variability, ranging between 2204 – 2827 kPa for the case with two objective functions, while a slightly lower pressure of 2000 – 2500 kPa was reported by Ma and Leung (2020a). This is because co-injecting C_1 has increased the overall dew point pressure of the C_1-C_3 solvent mixture. The optimal injection pressure is higher when a non-condensable gas (e.g. C_1) is co-injected with the solvent (e.g. C_3). Also, it is reported that the N-Solv® pilot involves injecting highly purified butane (C_4) at low pressures and near-saturated conditions (i.e. 40 – 60^{O}C) to achieve maximum energy efficiency (Emissions Reduction Alberta, 2016). These trends resemble those of solution sets 1 and 3 for the case with three objective functions. There are some differences in the absolute values for several reasons: (1) The models presented in this study are based on a 2D homogeneous reservoir; (2) The optimal injection pressure in this study is higher since a C_1-C_3 mixture with higher dew point pressure,

instead of C_4, is injected. It is reasonable to conclude that the proposed MOO workflow is useful for identifying different combinations of optimal operational parameters in a relatively fast and computationally efficient manner.

4.4. Summary

- A homogenous reservoir model for the warm VAPEX technique is constructed, and a sensitivity analysis is performed to examine the influences of various design variables on oil recovery and C_3 retention. The results demonstrate that C_1 provides a thermal barrier for overburden heat loss and hinders the accumulation of a C_3-concentrated liquid phase in the vapor chamber that leads to an ineffective propagation of the solvent towards the chamber edge. Also, injecting near the dew point conditions can improve oil recovery, but it also causes an increase in C_3 retention in the reservoir. Among the factors considered, a 2-level factorial experimental design is employed to identify decision variables that are most relevant to the process.

- A MOO workflow, which includes an ANN proxy model for objective function evaluation, and the NSGA-II MOO algorithm for determining a set of Pareto-optimal solutions is employed. The decision variables are C_3 mole fraction, bottom-hole injection temperature, and bottom-hole injection pressure. A novel parameterization scheme is proposed to incorporate the phase behavior constraints.

- For the 2-objective MOO problem, the optimal solutions are those that correspond to high C_3 fraction (~0.9), along with high bottom-hole injection temperatures and low pressures. For the

3-objective MOO which also considers the minimization of injected enthalpy, three sets of Pareto-optimal solutions corresponding to low, high, and low to medium bottom-hole injection temperature and pressure conditions are identified, while the optimal C_3 mole fraction remains as 0.9. Also, the solution sets illustrate the different trade-offs that may exist between the objective functions for those varying operating scenarios.

- This study demonstrates the ability of a MOO workflow to identify combinations of optimal design parameters for a solvent-based bitumen recovery process. Specific considerations regarding the solvent phase behavior, computational cost, and multiple conflicting objectives have been incorporated. It is anticipated that this workflow can be readily integrated into the decision-making processes in heavy oil reservoir management. Future work will modify the developed workflow to account for multiple geostatistical realizations of uncertain heterogeneous reservoir properties. The coupling of the simulation model with a surface network (i.e., wells and production facilities) should be explored.

Chapter 5 Determination of the Optimal Operational Parameters for a Warm Vaporized Solvent Injection Process, Using an Efficient Multi-Objective Optimization Workflow that Considers Reservoir Heterogeneity Uncertainty

Abstract

The warm vaporized solvent injection process is a bitumen recovery technique that involves the injection of heated solvent vapor (through a horizontal injector) into the reservoir. As the solvent contacts the colder bitumen, it condenses and dissolves into the bitumen, while the diluted oleic phase would flow towards the horizontal producer. Despite the promising results reported from reservoir simulations, laboratory experiments and pilot studies, commercial-scale extraction has not been reported. Since solvent-based technologies may be more expensive to operate, in comparison to existing steam-based processes, a robust optimization of the pertinent design variables is needed. The trade-offs between multiple conflicting performance objectives must be balanced, while honoring operational constraints and effectively incorporating reservoir heterogeneity uncertainties into the optimization process. This study presents a new multi-objective optimization (MOO) workflow that incorporates the uncertainty introduced by multiple realizations of reservoir properties in the identification of optimal operational parameters (or design variables).

Geostatistical realizations of reservoir heterogeneities are constructed by superimposing stochastically distributed shale barriers into a 2D base (homogeneous) Athabasca oil sands model. To achieve more optimal mixing and mass transfer, an innovative scheme is implemented to incorporate the thermodynamic behavior due to the addition of non-condensable gas (methane) into the solvent (propane) as constraints. Several design variables (i.e., injection compositions, temperature, and pressure) are optimized. Also, a proxy model (developed using the artificial neural network method) is used to determine the correlation between these design variables and two objective functions: oil recovery factor and propane loss. To account for geologic uncertainty, a separate proxy model is constructed for each realization, and for each objective function, the minimum, the weighted mean, and the maximum over these realizations are considered. Finally, a Pareto-based MOO algorithm, which is based on the Non-dominated Sorting Genetic Algorithm II (NSGA-II), is implemented.

The developed approach can identify a set of optimal design variables in a computationally efficient manner. The results are analyzed to formulate different options of optimal design parameters. The trade-offs between these different options are examined and quantified. Also, the presence of shale baffles or barriers can affect the performance of the warm vaporized solvent injection process.

Since reservoir models are usually presented with multiple realizations and due to the presence of non-condensable gases in many solvent-assisted methods, this workflow is robust and has the potential to be applied to several solvent-based bitumen recovery processes. Real field data can also be used to train the models, such that the workflow can be extended to model pilot and real-field operations.

5.1. Introduction

Due to the high viscosity of bitumen, thermal and solvent-based methods are implemented to lighten bitumen viscosity through heat transfer (in the form of steam or hot water injection) and/or mass transfer (injection of light hydrocarbon solvents or CO_2). The recent challenges that are associated with many thermal extraction techniques include technical constraints, operational safety concerns, water consumption, and greenhouse gas (GHG) emissions, hence non-thermal methods (e.g., solvent-aided recovery processes) have been proposed as promising alternatives.

Studies concerning solvent-based and steam-solvent hybrid methods were discussed in the literature, and they often utilize a similar well configuration as the widely adopted steam assisted gravity drainage (SAGD) process. For example, Butler and Mokrys (1991) proposed the vapor extraction (VAPEX) method which involves the injection of a solvent vapour at dew point conditions for the reduction of bitumen viscosity through mass transfer. However, according to Shi and Leung (2014a, b), the production rate from VAPEX is usually low, which may be due to the slow nature of mass transfer. Therefore, the warm VAPEX process, which leverages both heat and mass transfer, was introduced. This technique involves the injection of vaporized solvent between the dew point to superheated conditions in-situ, such that it can dissolve into the bitumen. A variant of warm VAPEX is the NSolv® process (Nenniger and Nenniger, 2001) which involved the injection of pure warm vaporized solvent close to dew point conditions. Additional heat is transferred to the bitumen from the release of latent heat after condensation (Nenniger and Dunn, 2008). The solvent is produced with the bitumen, then separated, and re-injected into the reservoir. Other solvent assisted methods, which include the co-injection of steam and solvent at low concentrations are the solvent aided process (SAP) (Gupta and Gittins, 2005), expanding solvent

SAGD (ES-SAGD) (Nasr et al., 2002) and solvent assisted SAGD (SA-SAGD) (Dickson et al., 2013).

Although, some of these solvent-based methods have been piloted, commercial-scale extraction may pose certain challenges. Compared to the SAGD, solvent-assisted techniques usually cost more to operate, hence the need for optimization of the pertinent operational constraints (design variables) that control the process. Also, the uncertain distribution of reservoir heterogeneity renders the decision-making process to be more challenging. Therefore, the trade-offs that exist between multiple conflicting performance objectives must be balanced while considering operational constraints and incorporating the uncertainty introduced by multiple realizations of reservoir heterogeneity into the optimization process.

To integrate the uncertainty associated with multiple realizations (e.g., low-case, mid-case, and high-case) of reservoir heterogeneity within an optimization workflow, one approach may be to give the highest weight to the realization with the most likely chance of occurrence (mid-case), and the weighted average of the objective function using the low-case, mid-case, and high-case realization is defined for the optimization task. For equiprobable realizations, the average of the objective functions over all realizations can be used (van Essen et al., 2009). Also, in Kathrada and Azri (2019), the optimization of well controls is performed across multiple realizations of geologic uncertainties, in which the Swanson's rule was employed to create a pseudo-objective function. According to Hurst et al. (2000), the Swanson’s 30-40-30 rule is a good representation of the average values (e.g., in reserve estimation) for most distributions apart from the highly skewed distributions.

Conducting a simple analysis where one or two variables are varied at a time, to identify the optimal parameters is not feasible. Hence, an optimization framework, which considers all the possible combinations of the decision variables, is necessary. In some optimization problems, a single objective function is defined. For example, in Al-Gosayir et al. (2013), where a hybrid genetic algorithm (GA) technique is implemented to optimize a steam-over-solvent injection process in fractured reservoirs, an economic objective function, which is the Money Recovery Factor (MRF) is defined. The MRF combines cumulative steam/solvent injection and recovery, as well as oil production. However, the limitation of a single objective function is that the trade-offs between objective functions cannot be easily evaluated (Ngatchou et al., 2005). Therefore, in this study, a multi-objective optimization (MOO) approach is employed, in which two objective functions (i.e., cumulative solvent lost-to-oil ratio (cSolOR) and oil/bitumen recovery factor (RF) are defined.

MOO involves the identification of an optimal set of decision variables along the Pareto-optimal front. Vector Evaluated Genetic Algorithm (VEGA) (Schaffer, 1985), Niched Pareto Genetic Algorithm (NPGA) (Horn et al., 1994), and Multi-Objective Genetic Algorithm (MOGA) (Fonseca and Fleming, 2011) are popular MOO techniques. Other widely adopted MOO algorithms are the elitist multi-objective evolutionary algorithms (MOEA) such as the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002), Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler and Thiele, 1998), and the Pareto-Archived Evolution Strategy (PAES) (Knowles and Corne, 1999). Recent studies have applied some of these MOEA in enhanced oil recovery (EOR) and well placement. For example, Zhang et al. (2019b) proposed a hybrid NSGA-II workflow to history match and optimize an alkaline-surfactant-polymer (ASP)

flooding process, where two objectives, including oil recovery and chemical usage, are considered. Other applications of the NSGA-II in petroleum engineering include the optimization of ES-SAGD, where the decision variables are the solvent mole fraction, injector bottom-hole pressure and temperature (Min et al., 2017), identification of the optimal operational parameters (i.e. solvent mole fraction and duration of solvent injection in a cycle) for the Steam Alternating Solvent (SAS) process (Coimbra et al., 2019), and the optimization of the warm VAPEX process for heterogeneous reservoirs (Ma and Leung, 2020a). According to Ma and Leung (2020a), the NSGA-II was adopted to determine the optimal injector and producer bottom-hole pressures of this process. In their work, which did not examine the co-injection of methane or incorporate phase behaviour constraints, they considered only 2 heterogeneous cases, and developed proxy models each case using a response surface methodology (RSM) and artificial neural network (ANN), then incorporated these proxy models into an NSGA-II MOO workflow. The optimal solutions for the heterogeneous cases were compared to a homogenous case and they concluded that shale barriers can have a negative effect on the warm VAPEX process.

For this study, a NSGA-II MOO workflow that incorporates the uncertainty introduced by multiple realizations of a heterogeneous reservoir is proposed to identify the optimal operational constraints of the warm VAPEX process, involving the co-injection of propane (C_3) and methane (C_1). In order to achieve these objectives, the following are performed in this paper: (1) develop three separate realizations (low-case, mid-case and high-case) of a heterogeneous reservoir model; (2) For each realization, create separate ANN-based proxy models to correlate the decision variables to each of the objective functions; (3) Find the Pareto-optimal decision variables for the MOO problem while considering the minimum, weighted average or Swanson's mean, and maximum

values of each objective function, among all the realizations. Since multiple realizations of heterogeneous properties and the presence of a non-condensable gas (i.e., C_1) are incorporated, this workflow is robust and can easily be applied to several solvent-based bitumen recovery processes. The workflow can also be extended to identify optimal operating strategies in pilot and real-field operations by using field data to train the proxy models.

5.2. Methodology

5.2.1. Reservoir Modeling

A base homogenous two-dimensional single porosity-permeability reservoir model, with a dimension of 35 m × 20 m × 32 m and a cell size of 0.5 m × 20 m × 0.5 m in the x, y and z directions respectively, is constructed using a fully implicit thermal-compositional reservoir simulator, STARSTM (CMG, 2019b). The reservoir model is populated with reservoir properties that are representative of the Athabasca oil sands (Ma and Leung, 2019a, b, c) as shown in Table 5-1. The symmetric propagation of the solvent chamber in the reservoir is assumed, hence one-half of the reservoir is modeled. The horizontal well trajectory is parallel to the y-axis.

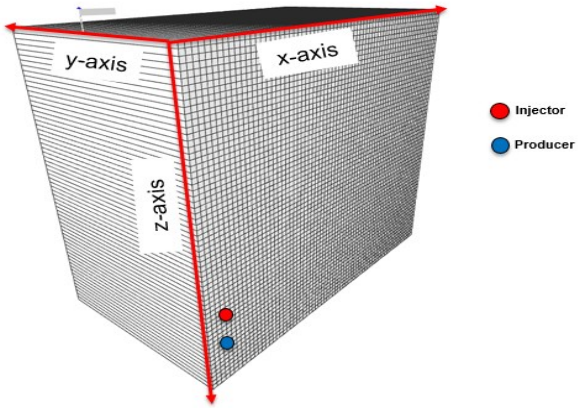


Figure 5-1: Reservoir model description.

Each model incorporates a different realization of the shale barrier distribution. As applied in the geostatistical software library (GSLIB), the sequential indicator simulation (SISIM) is implemented to create these realizations (Deutsch and Journel, 1998). The realizations have a 10% shale volume and are constructed with a spherical variogram model. The major and minor directions of continuity both have a range of 10 meters, while a 1 meter range is imposed in the vertical direction.

Table 5-1: Reservoir model properties

Description	Parameters	Value
Reservoir properties	Dimension of reservoir (ijk)	35m x 20m x 32m
	Number of grid blocks (ijk)	70 x 1 x 64
	Grid block size (ijk)	0.5m x 20m x 0.5m
	Sand permeability in the I direction (mD)	2500

	Sand permeability in the J direction (mD)	2500
	Sand permeability in the K direction (mD)	1500
	Shale permeability in the I direction (mD)	0.001
	Shale permeability in the J direction (mD)	0.001
	Shale permeability in the K direction (mD)	0.001
	Sand porosity (%)	32
	Shale porosity (%)	3
Initial conditions	Reference depth (m)	200
	Reservoir pressure (kPa)	1100
	Reservoir temperature (^{o}C)	10
	Sand initial water saturation (%)	13
	Shale initial water saturation (%)	25
	Molar concentration of dissolved C_1 (mole %)	5
Solvent type	Names	Propane (C_3) – Methane (C_1) mixture
Reservoir fluids	Names	Bitumen, Methane (C_1), Water (H_2O)
Rock/Fluid properties	Rock compressibility (kPa^{-1})	2.0×10^{-6}
	Rock heat capacity ($J/m^3.^{o}C$)	2.35×10^{6}
	Rock thermal conductivity ($J/m.^{o}C$)	1.468×10^{5}

	Water thermal conductivity (J/m.^{o}C)	5.35 x 10^{4}
	Oil thermal conductivity (J/m.^{o}C)	1.15 x 10^{4}
	Gas thermal conductivity (J/m.oC)	1.39 x 10^{2}
	Bitumen viscosity at 15^{o}C and 101.325 kPa (cP)	562204
	Propane effective molecular diffusion coefficients for the I,J,K directions (m^2/day)	4.32 x 10^{-5}
	Mechanical (convective) dispersivity of oil and phases for the I,J,K directions (m)	2.00 x 10^{-5}
	Rock wettability	Water wet (capillary pressure ignored)
	Model for evaluating 3-phase k_{ro}	Stone's second model
	Relative permeability endpoints	k_{rw} = 0.79, k_{row} = 0.95, k_{rg} = 0.50, k_{rw} = 0.95
Well constraints	Pressure differential between injector and producer (kPa)	200

To represent multiple geological realizations of a heterogeneous reservoir model, three separate realizations are constructed by superimposing stochastically, distributed shale barriers into the base model. In creating these realizations, the sequential indicator simulation (SISIM), which is based on the geostatistical software library (GSLIB) (Deutsch and Journel, 1998), is implemented. For each of the realizations, the facies are sand and shale, and it is assumed that the well-pair is completed in clean sand. Also, the realizations have a 10% shale volume, and are constructed using

a spherical variogram model. Both the major and minor directions of continuity have a range of 10 meters, while a 1-meter range is imposed in the vertical direction. As shown in Fig. 5-2, the shale barriers or baffles are denoted with a blue color, while the sand distribution is represented in yellow. The simulation is run for approximately 15 years, starting with a 4-month preheating period.

Fig. 5-3 shows the propagation of the solvent chamber at different snapshots of time (year 1, year 4, year 9, and year 15), in which the solvent front bypasses the lowly- permeable shale baffles, while the shale barriers impede solvent propagation in the vertical direction.

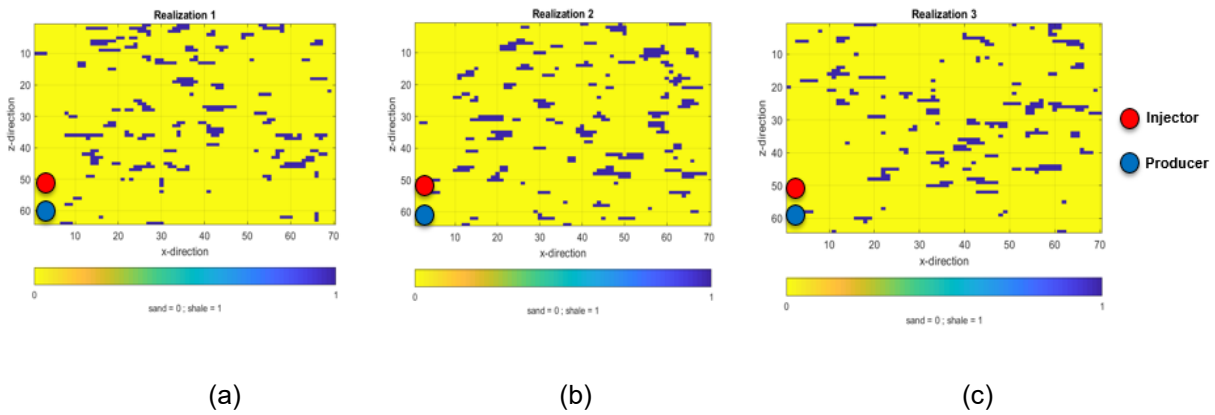


Figure 5-2: The shale and sand distribution: a – Realization 1 (low-case); b – Realization 2 (mid-case); c – Realization 3 (high-case).

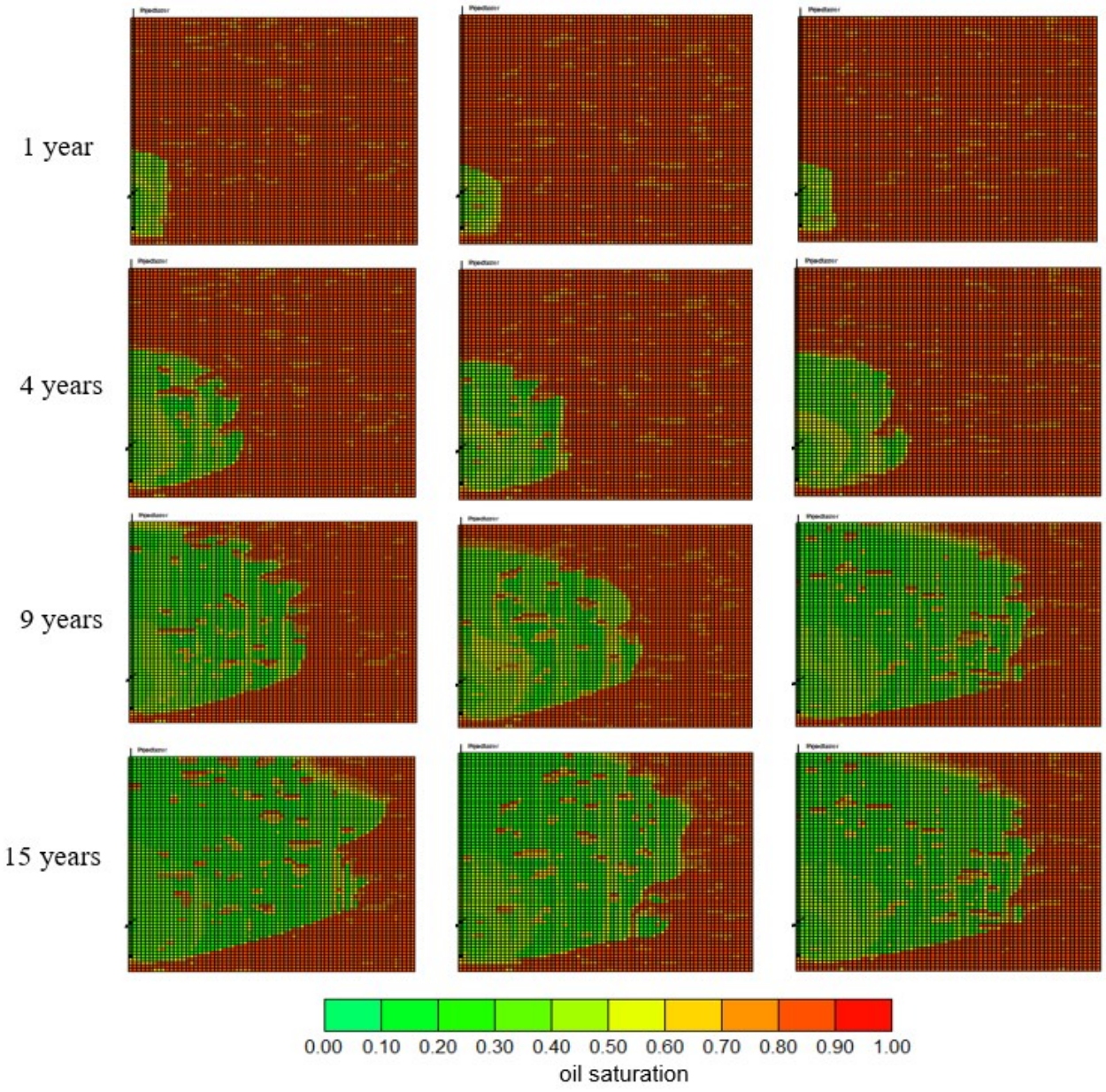


Figure 5-3: Oil saturations: a – Realization 1 (low-case); b – Realization 2 (mid-case); c – Realization 3 (high-case).

The construction of the geologic realizations is such that the sand units in Realization 1 (low-case) is poorly distributed when compared to Realization 2 (mid-case), while the sand distribution in Realization 3 (high-case) is better than Realization 2, hence as shown in Fig. 5-4, the *RF* for Realization 3 is greater than for Realization 2, while the *RF* for Realization 2 is greater than the *RF* for Realization 1.

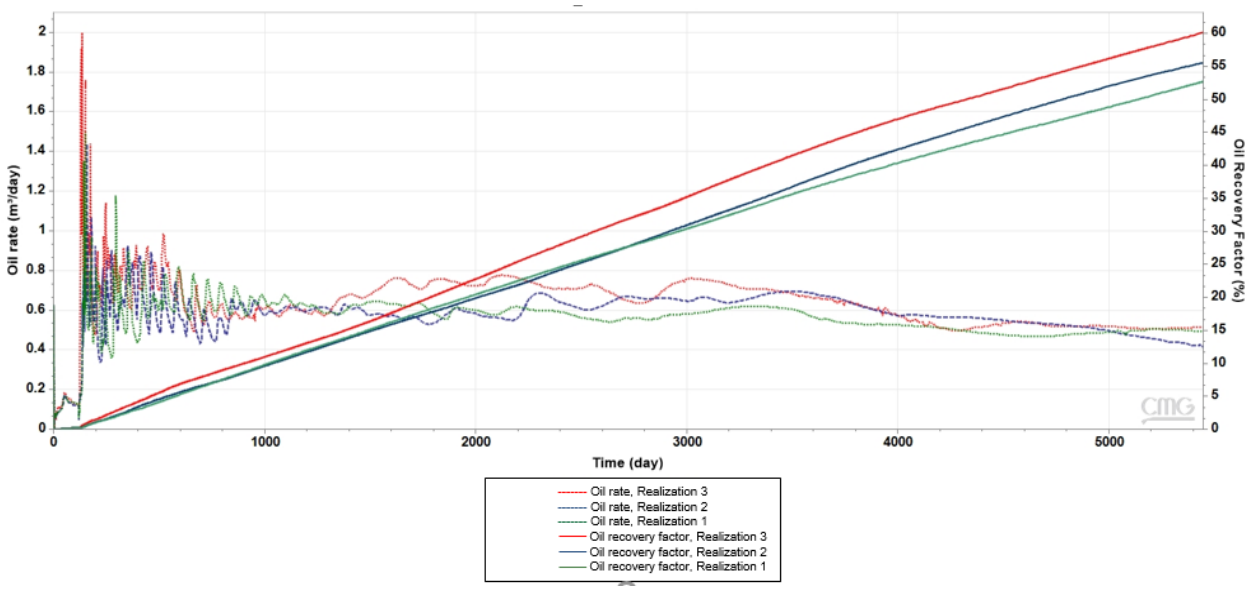


Figure 5-4: Oil recovery factor and production profiles: a – Realization 1 (low-case); b – Realization 2 (mid-case); c – Realization 3 (high-case).

5.2.2. Decision Variables and Objective Functions

The identification of the decision variables that would affect the objective functions and their trade-offs is crucial to any MOO. Similar to previous studies described earlier, the decision variables for our work include propane mole fraction in the injected solvent, injection temperature, and injection bottom-hole pressure. Since the MOO is a minimization problem, and because it is desirable to

minimize *cSolOR* and maximize *RF*, the objective functions are *cSolOR* and the inverse of *RF* (*1/RF*), which are defined as follows:

$$cSolOR = \frac{V_{C_3}^{inj} - V_{C_3}^{prod}}{V_{oil}^{prod}}; \quad (1)$$

$$1/RF = \frac{OOIP}{V_{oil}^{prod}}, \quad (2)$$

Where $V_{C_3}^{inj}$ is the total volume of gaseous propane injected (in m^3), $V_{C_3}^{prod}$ is the total volume of propane recovered at the surface (in m^3), V_{oil}^{prod} is the total volume of bitumen extracted (in m^3), and OOIP denotes the original oil in place (in m^3). All the parameters are specified at surface conditions.

For the MOO task, in order to capture the variety exhibited by the different realizations, three approaches are considered: using either the minimum, weighted average, or maximum values of each objective function over the realizations. The weighted average values are obtained by computing a Swanson's mean as shown below.

$$E\{f(x)\} = 0.3\, f(x, P_{low-case}) + 0.4\, f(x, P_{mid-case}) + 0.3\, f(x, P_{high-case}), \quad (3)$$

Where $E\{f(x\}$ is the weighted average value of the objective function over the 3 realizations (i.e. low-case, mid-case, and high-case) being considered, x is the decision vector and $f(x, P_{low-case})$, $f(x, P_{mid-case})$, and $f(x, P_{high-case})$ are the objective functions using the low-case, mid-case, and high-case realizations respectively.

5.2.3. Multi-Objective Optimization Problem

For a MOO minimization problem with three decision variables (or vector), a decision vector, $u = (u_1, u_2, u_3)$ will dominate (denoted by $\prec$) another decision vector $v = (v_1, v_2, v_3)$, if and only if

$$f_i(u) \leq f_i(v), \ \forall_i \in \{1, 2, .. m\};$$
$$\text{and } f_j(u) < f_j(v), \qquad \forall_j \in \{1, 2, .. m\}, \tag{4}$$

Where m denotes the total number of objectives, $f(u)'s$ are the objective functions for vector u, and $f(v)'s$ are the objective functions for vector v. The idea is that $f(u)$ is lower than or as best as $f(v)$ for all objectives and there is at least one objective function for which $f_j(u)$ is strictly less than $f_j(v)$.

Since the MOO problem for this study is formulated with three decision variable and two objective functions, the decision vector u (or v) is given by:

$$u = [u_1\ u_2\ u_3\ ] = \left[x_{c_3}\ T_{inj}\ P_{inj}\ \right], \tag{5}$$

Where x_{c_3}, T_{inj}, and P_{inj} are the decision variables, which are propane mole fraction in the injected solvent, injection temperature, and injection pressure respectively. The objective functions vector is shown in Eqn. 6.

$$[f_1(u,),\ f_2(u)\ ] = [cSolOR,\ 1/RF], \tag{6}$$

The aim of this study is to identify the optimal sets of x_{c_3}, T_{inj}, and P_{inj} vectors that would minimize $cSolOR$ and $1/RF$. Typically, algorithms that perform a non-dominated sorting of the

solution space are used to search for a set of optimal solutions that are uniformly spaced along a Pareto-optimal front; for this work, a widely-adopted technique called NSGA-II (Deb et al., 2002) is applied.

5.2.3.1. The Non-Dominated Sorting Genetic Algorithm II (NSGA-II)

The procedure for an iteration of non-dominated sorting using the NSGA-II is described below:

1. An initial parent population P_t of size N is created randomly and sorted into sets of non-dominated solutions: the best solutions are assigned a rank of 1; the second-best solutions are assigned a rank of 2, and so on.
2. The offspring population Q_t of size N is created from P_t through the selection, crossover, and mutation operations.
3. Elitism is ensured by combining P_t and Q_t to form a combined population R_t of size $2N$.
4. Non-dominated sorting is applied to R_t, and the new parent population P_{t+1} is filled with solutions from fronts of R_t with decreasing rank until the total size of N is reached. Whenever the last allowable front is considered, there may be are more solutions than the remaining slots within P_{t+1}. In such instances, solutions within the least crowded region are selected to fill these slots and the remaining individuals in R_t are discarded.
5. P_{t+1} is subjected to another round of selection, recombination, and mutation to create a new offspring population Q_{t+1} of size N.

A previously stated, the selection of individuals from the last allowable front of R_t to fill P_{t+1} can be based on crowding distance comparison. The crowded comparison operator, $\prec_c$ (Deb et al., 2002) is defined as follows.

$$i \prec_c j \quad \text{if } i_{rank} < j_{rank};$$
$$\text{or } i \prec_c j \quad \text{if } i_{rank} = j_{rank} \text{ and } i_{distance} > j_{distance}, \tag{7}$$

Where i_{rank} and j_{rank} denotes the ranks of individuals i and j, respectively, within R_t, while $i_{distance}$ and $j_{distance}$ are their crowding distances. The crowding distances, $i_{distance}$ and $j_{distance}$ are the total side length of the cuboid formed by the neighboring solutions to the left and right of i and j respectively. Fig. 5-5 and Eqn. 8 describe the crowding distance for a 2-objective MOO.

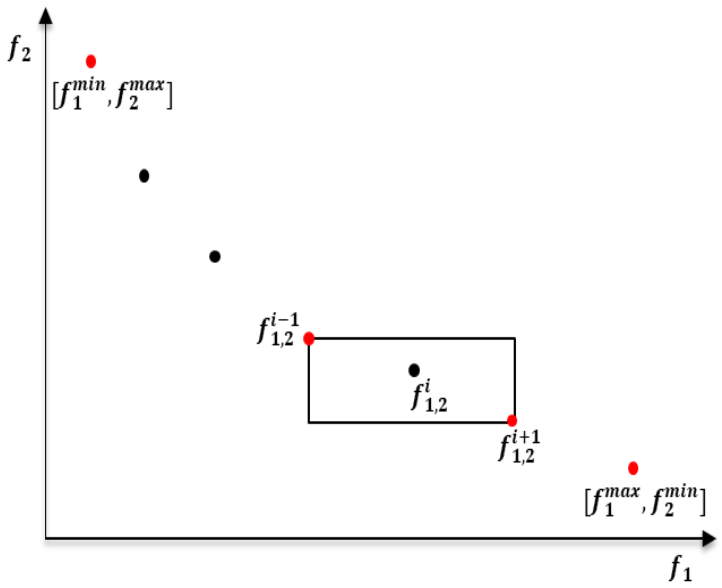


Figure 5-5: Schematic for crowding distance calculation.

$$i_{distance} = \frac{\left|f_1^{i+1} - f_1^{i-1}\right|}{f_1^{max} - f_1^{min}} + \frac{\left|f_2^{i+1} - f_2^{i-1}\right|}{f_2^{max} - f_2^{min}} + \cdots \frac{\left|f_m^{i+1} - f_m^{i-1}\right|}{f_m^{max} - f_m^{min}}, \quad (8)$$

Where f_m^{i-1} and f_m^{i+1} are the m^{th} objective function of the neighboring individuals $i-1$ and $i+1$, respectively; f_m^{min} and f_m^{max} are the minimum and maximum values of the m^{th} objective function. The NSGA-II parameters for this study are: population size = 200, number of iterations = 50; mutation and cross-over probability is 0.3. The optimal solutions are obtained once the maximum number of iterations is reached.

5.2.3.2. Parameterization Scheme – Incorporating Thermodynamic Constraints

Since the solvent mixture (e.g., propane-methane mixture) for the warm VAPEX process are injected at either the dew point or superheated conditions for optimal solvent-oil interactions, for a given propane (C_3) mole fraction, a "feasibility window" for injection conditions of the solvent mixture is defined within the propane-methane phase diagram. This window renders the MOO problem to be rather peculiar. Therefore, a novel parameterization scheme is incorporated within the NSGA-II to constrain the pressure and temperature conditions for a given C_3 composition within the window (or search space). As shown in Fig. 5-6, the search space for the case of the injection of a C_1-C_3 solvent mixture with a 90% C_3 mole fraction is specified by a shaded portion.

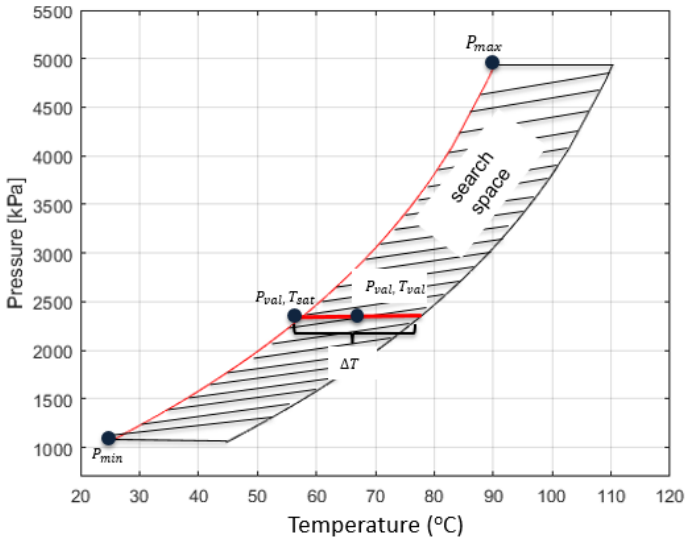


Figure 5-6: Search space for the implementation of NSGA-II (P-T diagram for C1-C3 mixture; C3 mole fraction is 0.9). Lower dew point curve shown in red.

Two separate indicators, P and T with values ranging between 1 and 10 are randomly generated within the genetic algorithm framework of the NSGA-II to denote pressure and temperature values respectively. The indicator values are then mapped back to the search space to obtain the actual temperature and pressure values for a particular C_3 mole fraction. The parameterization procedure is described below:

1. Randomly create T and P, with values ranging from 1 to 10.
2. Formulate the decision vector of the population as C_3 mole fraction (x_{C3}), P, and T.
3. For the value of x_{C3} (lower bound = 0.5; upper bound = 0.9), determine P_{max} and P_{min} from the dew point curve. P_{max} is the upper pressure limit along the dew point curve and P_{min} is the dew point pressure at 30°C (the minimum injection temperature for this study).

4. Estimate the mapped pressure and temperature values (P_{val} and T_{val}) corresponding to the indicators *P* and *T* using Eqns. 9 and 10 respectively.

$$P_{val} = P_{min} + \left(\frac{P-1}{10-1}\right) \times (P_{max} - P_{min}); \quad (9)$$

$$T_{val} = T_{sat} + \left(\frac{T-1}{10-1}\right) \times \Delta T, \quad (10)$$

Where T_{sat} is the saturation temperature at P_{val} and ΔT is the temperature range of superheated conditions. For this work, ΔT is set to 20°C.

5.2.4. Objective Function Evaluation Using Proxy Modeling

The evaluation of the objective functions with reservoir simulation can be computationally inefficient. For example, in this work, the population size is 200 and the number of iterations is 50, the required number of reservoir simulations will be 10,000 runs per geologic realization, which is a huge computational load. This is because the computing time for a single simulation run could range from a few minutes to several days, depending on the model size and process complexity. Therefore, a proxy model, which is calibrated using a limited number of simulation runs, can be employed to evaluate the objective functions. The proxy modeling technique employed in this study is the ANN. The ANN modeling approach aims to approximate the complex non-linear relationships between a set of input variables and output attributes. It has been applied in the optimization of warm solvent injection processes (Ma and Leung, 2020a), investigation of reservoir heterogeneities in SAGD (Wang et al., 2018; Zheng et al., 2018a; Wang and Leung, 2015), and the analysis of SAGD production performance (Ma et al., 2015). Typically, an ANN

design consists of an input layer, one or more hidden layers, and an output layer. Each layer usually consists of a certain number of nodes and a bias unit, and for the input layer, the nodes are the features (i.e., decision variables for this study), while the output nodes are the target variables (i.e., the objective functions in this study). The mathematical relationship between the nodes is given by Eqn. 11.

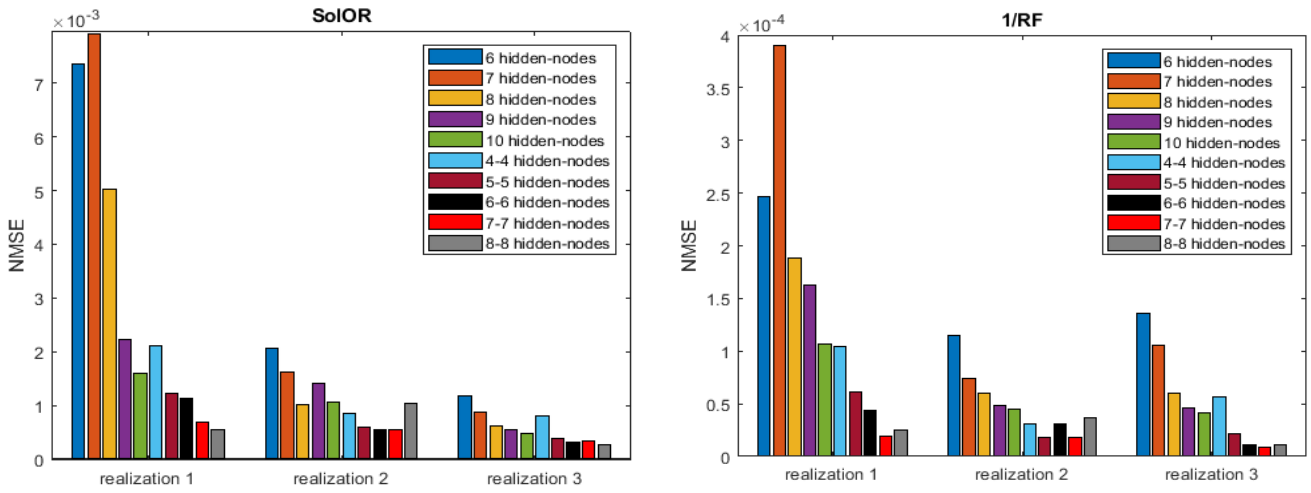
The output of node j in layer $k+1$, a_j^{k+1} is given by:

$$a_j^{k+1} = g\left(\sum_{i=1}^{n} w_{j,i}^{k} . x_i^k + b\right), \tag{11}$$

Where b denotes the bias, $w_{j,i}^{k}$ is the weight corresponding to the connection between node i in layer k and node j in layer $k+1$, x_i^k refers to the input value of node i in layer k and g is the activation function. The weights are estimated through a backpropagation update algorithm that consists of an optimization framework, such as gradient descent, Adaptive Moment Estimation (Adam), Levenberg-Marquardt, etc., where the deviation between the actual target values and model predictions (which is defined in terms of a cost function, such as mean square error, MSE) is minimized after several batches of training.

In this study, a separate proxy model is created for each of the 3 geologic realizations, and since there are 2 objective functions (*cSolOR* and 1/*R*F), a total of 6 proxy models are constructed. A 10-fold cross-validation technique is employed to determine the optimal number of hidden layers and nodes per layer, using the normalized mean squared error (NMSE) performance metric as the selection criteria. The cross-validation approach is performed for the 6 proxy models and the

results are shown in Fig. 5-7. The optimal architecture with a low value of NMSE over the 3 geologic realizations for both *cSolOR* and *1/RF* consists of 2 hidden layers, with 6 nodes per layer. This optimal network configuration is then implemented using the neural network toolbox in Matlab™ (Beale et al., 2018). A total of 270 cases are used for training and testing [230 samples (85%) for training and 40 samples (15%) for testing].



(a) (b)

Figure 5-7: Cross validation results: a - cSolOR; b – 1/RF.

5.3. Results and Discussion

5.3.1. Proxy Modeling

The predictions of the ANN based proxy-models for 40 test cases are validated with the reservoir simulations, as shown in Figs. 5-8 and 5-9. For all the results, the coefficient of determination (R^2) is close to 1, indicating that the predictions from the ANN model match closely to the outputs from reservoir simulation.

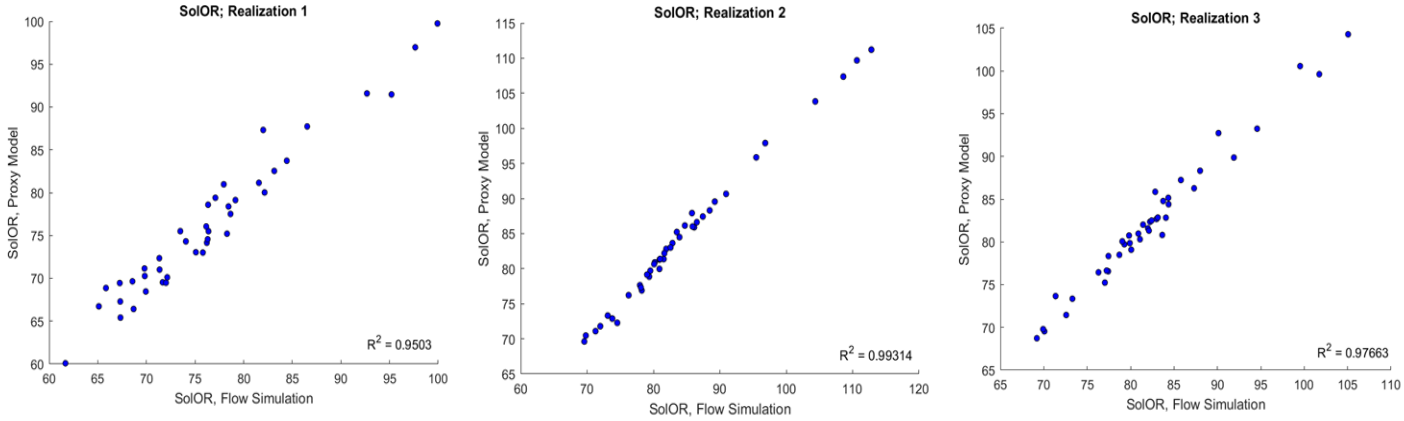


Figure 5-8: Comparison of proxy model predicted values and flow simulation output of cSolOR.

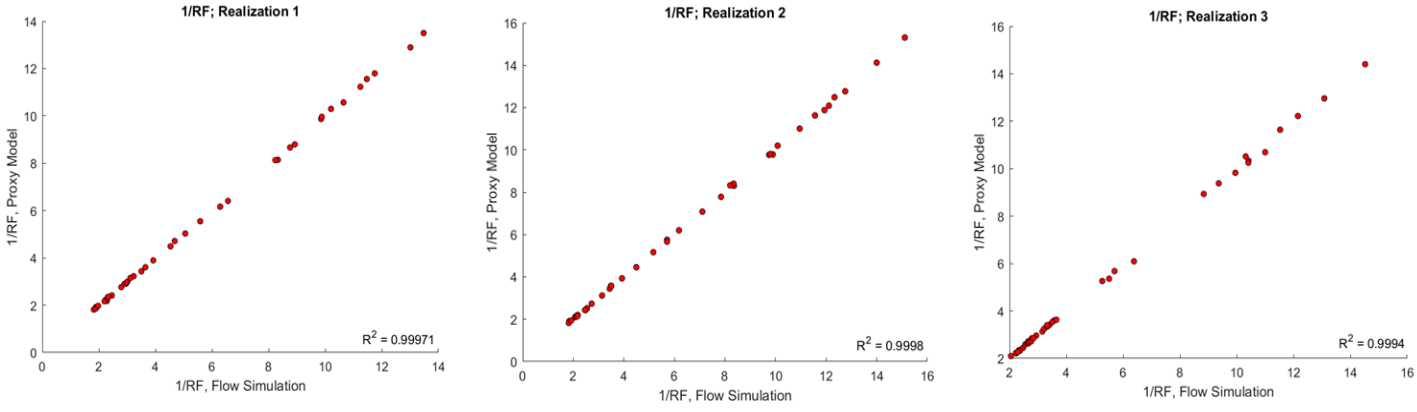


Figure 5-9: Comparison of proxy model predicted values and flow simulation output of 1/RF.

5.3.2. MOO

The final Pareto-optimal solutions after 50 iterations are shown in Fig. 5-10. Three different Pareto fronts corresponding to the minimum, weighted average, and maximum objective function values over all the realizations are attained. If the Pareto front is derived from the minimum objective function value among all realizations, it is generally expected that the realizations contributing to the lowest objective functions are those exhibiting minimal heterogeneity (e.g., very few shale barriers). Similarly, if the Pareto front is attained based on the maximum objective function value, the realizations contributing to this case would likely be those exhibiting high degree of heterogeneity (e.g., many shale barriers in the near-well region).

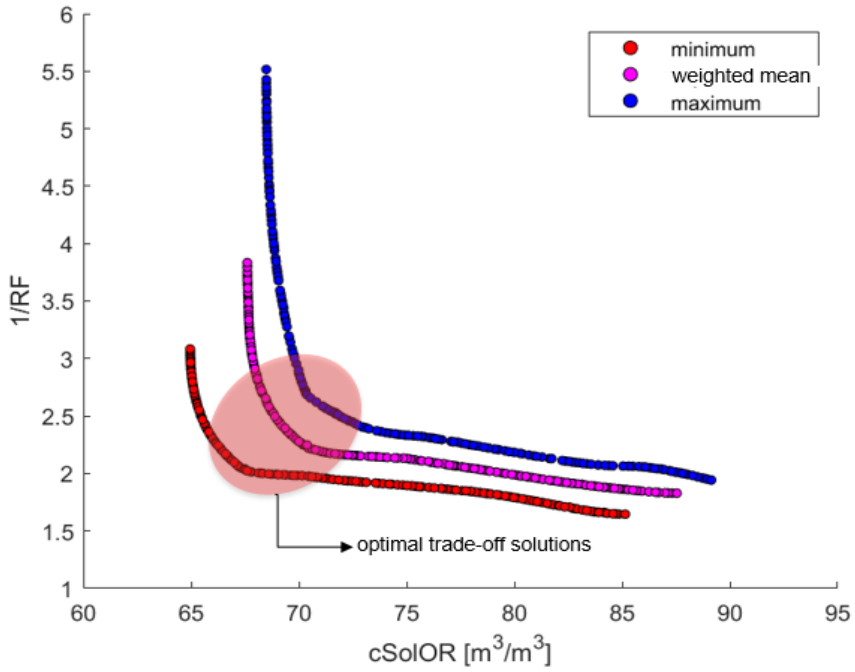


Figure 5-10: Pareto-optimal fronts.

If the minimum values of *cSolOR* and 1*/RF* over the realizations (best case scenario) are considered, the optimal propane mole fraction, injection pressure, and temperature ranges would be 0.52 - 0.90, 1648 - 4489 kPa, and 60 - 70°C, respectively. As shown in Fig. 5-11, out of the 200 optimal solutions, the most frequent scenario is the injection of solvent with a high propane mole fraction (modal propane mole fraction = 0.9) at high temperature and low injection pressure. When referring to the phase diagram of the binary methane-propane mixture, this translates to solvent injection at superheated conditions to capture the trade-offs between solvent retention and bitumen recovery.

A plausible explanation can be conjectured to understand why injecting at superheated condition is preferred for heterogeneous reservoirs. When solvent is injected into a heterogeneous reservoir,

there is a risk of early condensation before establishing sufficient contact with the in-situ bitumen. The early condensation may be due to velocity variation along the chamber front as it encounters shale baffles and preferentially fingers through highly-permeability zones, thereby leading to lower temperature and solvent condensation behind the chamber edge. Solvent condensation before contact with bitumen can lead to the accumulation of solvent droplets in the chamber which may be trapped if its saturation is less than the critical saturation. Therefore, injection at superheated conditions lowers the likelihood of early solvent condensation, hence limiting the amount of solvent lost in the reservoir while the extra heat from the superheated solvent may be transferred to the bitumen for its viscosity reduction. The addition of methane or the increment of methane proportion within the methane-propane solvent mix may also delay liquid condensation in the solvent chamber (Das, 2008); however, methane addition may also reduce the recovery factor, as it may act as an insulative barrier for the transfer of thermal energy to the bitumen. The trade-offs between these different considerations may have led to the identification of a few solutions along the Pareto front, which involve the co-injection of higher methane content (i.e., moderate propane mole fraction).

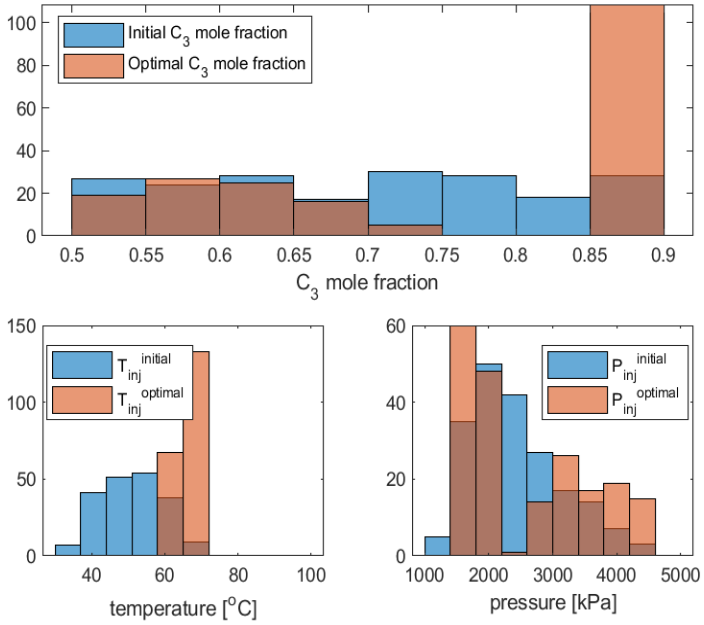


Figure 5-11: Distribution of decision variables considering the minimum objective function value over all realizations.

Considering the weighted average (or Swanson's mean) of *cSolOR* and 1/*RF* over the realizations, the optimal propane mole fraction, injection pressure, and temperature ranges would be 0.51 - 0.90, 1573 - 4680 kPa, and 58 - 70ºC, respectively. From Fig. 5-12, conclusions similar to what is described earlier can be made. However, since a weighted average value is usually higher than the minimum value for a set of data, the position of the Pareto front is shifted towards the right of the '*minimum values'* case.

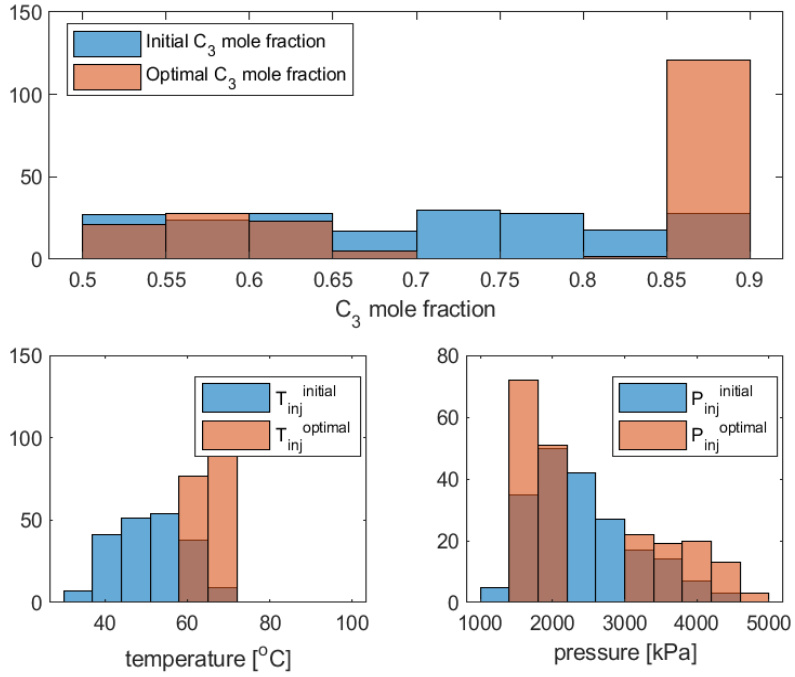


Figure 5-12: Distribution of decision variables considering weighted average objective function value over all realizations.

Finally, considering the maximum value of *cSolOR* and 1/*RF* over the realizations (these are deemed worst-case scenarios because they generally consist of many shale barriers in the near-well region) are considered. The optimal propane mole fraction, injection pressure, and temperature ranges between 0.50 - 0.90, 1413 - 4838 kPa, and 55 - 70°C, respectively. Once again, the pareto front is shifted to the right of both of the *'Swanson's mean'* and *'minimum values'* cases. Similar conclusions can be drawn.

An interesting conclusion can be drawn after analyzing all three ways of assessing the objective functions: there is no significant change in the ranges of optimal solutions of all 3 decision

variables. This demonstrates the robustness of this workflow in identifying optimal parameters against possible scenarios of geologic heterogeneity and phase behavior constraints.

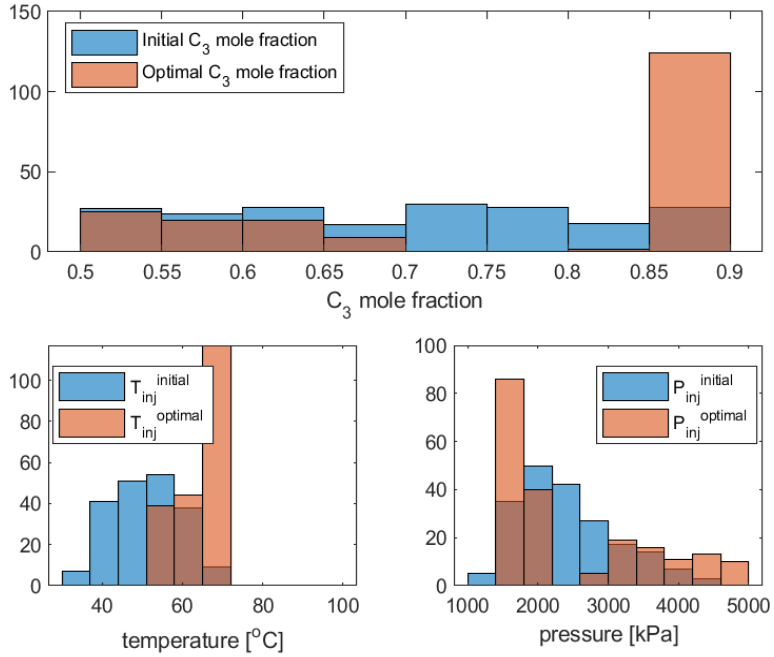


Figure 5-13: Distribution of decision variables considering maximum objective function value over all realizations.

The ranges of the optimal decision parameters and the corresponding objective function for each of the scenarios are summarized in Table 5-2.

Table 5-2: Optimal decision variables and objective functions

Approach	x_{c3} (fraction)	T_{inj} (°C)	P_{inj} (kPa)	cSolOR (m^3/m^3)	1/RF (1/fraction)
Minimum	[0.52 0.90]	[60 70]	[1648 4489]	[65 85]	[1.7 3.1]
Weighted average	[0.51 0.90]	[58 70]	[1573 4680]	[67 87]	[1.8 3.8]
Maximum	[0.50 0.90]	[55 70]	[1413 4838]	[68 89]	[1.9 5.5]

Although flow simulation of solvent injection processes in heterogeneous reservoirs can be computationally intensive; this workflow is able to optimize the warm VAPEX process with significant improvement in terms of computational efficiency. For example, if a reservoir simulator was used for all the objective function evaluation, it will take 900,000 minutes or approximately 625 days (i.e., 50 iterations $\times$ 200 simulation runs per iteration $\times$ 30 minutes/simulation run $\times$ 3 geologic realizations) without parallel computation. However, using a personal computer with 16.0 GB installed RAM and Intel(R) Core i7 processor, the total computing time is approximately 30 minutes.

5.4. Summary

A MOO workflow that considers phase behavior constraints and geologic uncertainties for the optimization of the warm VAPEX process is presented. The decision variables are propane mole fraction in the injected solvent, injection pressure, and injection temperature, while the objective functions are *cSolOR* and *1/RF*. Several steps are implemented in this workflow: First, three separate reservoir models, representing a low-case, mid-case, and high-case realization, are considered; for each realization, an ANN-based proxy model is constructed to correlate one objective function to the three decision variables. Next, three ways of formulating the objective function in the MOO framework are examined: minimum, weighted average, and maximum values over all the realizations. The NSGA-II, together with the developed proxy models and a novel

parameterization scheme which incorporates phase behavior constraints, is employed to search for the optimal parameters along a Pareto front.

This study demonstrates the ability of a MOO workflow to identify optimal decision parameters for a solvent-based bitumen recovery process. The optimization results also show that there is no significant change in the final optimal solutions, regardless of how the objective functions are evaluated. The results suggest that our workflow is robust in identifying optimal parameters considering a range of different heterogeneity realizations. It is anticipated that this workflow can be readily integrated into the decision-making processes in heavy oil reservoir management.

Chapter 6 A Machine Learning Approach to Real-time Uncertainty Assessment of SAGD Forecasts and the Optimization of Steam Allocation

Abstract

Field development planning and economic analysis require reliable forecasting of bitumen production. Forecasting at the field level may be done using reservoir simulations, type-curve analysis, and other (semi-)analytical techniques. Performing reservoir simulation is usually computationally expensive and the non-uniqueness of a history-matched solution leads to uncertainty in the model predictions and production forecasts. Analytical proxies, such as Butler's model and its various improvements, allow for sensitivity studies on input parameters and forecasting under multiple operational scenarios and geostatistical realizations to be conducted rather quickly, despite being less accurate than reservoir simulation. Similar to their reservoir simulation counterparts, proxy models can also be tuned or updated as more data are obtained. Type curves also facilitate efficient reservoir performance prediction; however, in practice, the performance of many SAGD well-pairs tends to deviate from a set of pre-defined type curves.

Historical well data is a digital asset that can be utilized to develop machine learning or data-driven models for the purpose of production forecasting. These models involve lower computational effort compared to numerical simulators and offer better accuracy compared to proxy models based on Butler's equation. Furthermore, these data-driven models can be used for automated optimization, quantification of geological uncertainties, and "What If" scenario analysis.

This paper presents a novel machine learning workflow that includes a predictive model development using the random forest algorithm, clustering, Bayesian updating, Monte Carlo sampling, and genetic algorithm for accurate forecasting of real-world SAGD injection and production data, and optimization. The training dataset involves field data that is typically available for a SAGD well (e.g., operational data, geological, and well design parameters). Just as importantly, this machine learning workflow can update predictions in real-time, be applied for the quantification of the uncertainties associated with the forecasts, and optimize steam allocation, making it a practical tool for development planning and field-wide optimization. To the best of the author's knowledge, this is the first time that machine learning algorithms have been applied to a SAGD data set of this size.

6.1. Introduction

Steam Assisted Gravity Drainage (SAGD) is a thermal in-situ bitumen recovery technology proposed by Butler et al. (1981). Production of bitumen via SAGD from the Western Canada sedimentary basin has been prolific, reaching over a million barrels per day by 2017 (Canada Energy Regulator, 2020). SAGD involves the injection of high-quality steam into the reservoir to thermally mobilize viscous heavy oil or bitumen. The well configuration for this mechanism is designed in such a way that bitumen can be effectively drained under gravity. A horizontal producer well is located at the base of the reservoir while another horizontal well (steam injector) is drilled approximately 5 meters above the producer. Steam is injected into the injector well and condenses. The energy given off by the condensing steam is transferred to the rock, water, and hydrocarbon present in situ. As the hydrocarbons are heated, the oleic phase viscosity is reduced,

and it flows to the producer well where it is pumped to the surface along with connate water, gas, and the condensed steam.

The efficacy of the SAGD process is sensitive to operational parameters and reservoir heterogeneity. Typical current practice is to assign engineers to monitor well operation and make inferences about subsurface conditions. These inferences can be used for optimizing well and field operations, as well as for long-term development planning. An automated approach to effective real-time field optimization and development planning has not been widely adopted in the industry. Semi-analytical or proxy models (Dehdari and Dong, 2017) have been proposed for field-wide planning and optimization. The governing equation in these models is typically based on Butler's model and, hence, they are computationally efficient. However, such proxy models have not been widely adopted for several reasons: firstly, it is difficult to incorporate changes in operational constraints into these models; secondly, the model parameters are generally unknown and must be estimated and tuned according to field data; it is not uncommon to even modify the form of the model to better fit the actual data. As a result, such models are hard to be incorporated for real-time forecasting and uncertainty assessment when dealing with real-time dynamic data.

Data-driven modelling is based on the analysis of data characterizing the system of interest and often involves the application of machine learning methods to build models that describe the behavior of the corresponding physical process. There are three main machine learning approaches: supervised, unsupervised, and reinforcement learning. In supervised learning, there is a functional mapping from the inputs to a set of target variables (e.g., random forest, artificial neural network, etc.). In unsupervised learning, there are no target variables (e.g., K-means clustering and principal component analysis (PCA). In reinforcement learning, an agent learns

from interaction with its environment through different trial and error techniques in order to maximize a reward function.

Supervised and unsupervised learning methods have been adopted in many petroleum engineering applications. For instance, supervised learning was implemented in the optimization of a solvent-assisted process (Ma and Leung, 2020a), the production forecast for mature fields that are simultaneously driven by steam and water flooding (Kubota and Reinert, 2019), investigation of SAGD production performance (Ma et al., 2015), production forecasting (Zheng et al., 2018) and analysis of reservoir heterogeneities for SAGD (Wang et al., 2018; Zheng et al., 2018; Wang and Leung, 2015). The K-means clustering algorithm is a common technique for classifying datasets based on similarities among data, and PCA can be used for feature selection and reduce the dimension of data. Both methods have been integrated into a couple of reservoir engineering applications, which include prediction of the fluid properties of crude oil systems (Almashan et al., 2020), analysis of SAGD production data (Ma et al., 2015), and reservoir characterization (Gilbert et al., 2004). Ma. et. al. (2015) applied K-means clustering to separate a dataset of Canadian SAGD fields (gathered from the public domain) into different groups, and separate ANN models were constructed to predict several production indicators (e.g., total oil production) from petrophysical data. A similar approach was employed in Amirian et al. (2014) using a synthetic dataset from flow simulation results. The major drawbacks of these works are: (1) the datasets used have missing and limited information; key operational parameters are generally missing; (2) time-series outputs (e.g., oil production or steam injection vs. time profiles) were not considered.

In this work, instead of ANN, the random forest algorithm, in which input data are fed into multiple decision trees that are trained independently and the output is the average of predictions from each

tree, is used. This modeling technique is less expensive computationally, with fewer tuning parameters compared to ANN (Muhammad et. al., 2017). Therefore, it has often been applied to a variety of reservoir-engineering problems, such as history matching sensitivity analysis (Aulia et al., 2017) and production forecasting for tight hydrocarbon reservoirs (Liao et. al., 2020). To quantify the uncertainty associated with the production forecasts given a set of uncertain geologic variables, the Monte Carlo simulation (MCS) is used. The application of MCS can be found in Bieker et. al., (2006), Al-Mudhafar and Rao (2016) and Mehana et al. (2019).

The development of a data-driven model for SAGD forecasting facilitates the optimization of pertinent variables (e.g. operational constraints). In this work, the genetic algorithm (GA) is used to optimize steam allocation. Compared to other local, gradient-based, optimization algorithms (gradient descent, Levenberg–Marquardt algorithm, etc.), GA offers several advantages including its ability to perform a global search, the capacity to handle noisy objective functions, and its resistance to being trapped in local optima (Sivanandam and Deepa, 2008). A number of recent works have applied GA and other elitist multi-objective GA, such as the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002), to design heavy oil/bitumen recovery processes: optimizing solvent-assisted SAGD processes (Al-Gosayir et al., 2011), determining optimal conditions for high-temperature solvent injection techniques, in which a money recovery factor is defined as the objective function (Leyva-Gomez and Babadagli, 2017), designing steam alternating solvent process (Coimbra et al. 2019), where the objective functions are recovery factor and cumulative propane injection. Optimization problems can also be formulated to maximize the net present value (NPV) of the SAGD process using a reinforcement learning technique (Guevara et al., 2018) and for steam allocation optimization in real-time (Sibaweihi et al., 2020).

Although the aforementioned studies have successfully implemented various data-driven workflows for SAGD production forecasting and optimizing steam allocation, the impacts of geologic and operational uncertainties were not considered. Furthermore, synthetic data sets were used to train the data-driven models in these studies, rendering it challenging to extrapolate their conclusions to real field dataset. Therefore, a major contribution of this work is the use of a real SAGD field dataset for dynamic forecasting, uncertainty quantification, and optimization of steam allocation. The proposed workflow is shown to provide reliable prediction update (with its associated uncertainty) in real-time and steam allocation optimization solutions that can be used for decision-making under real-field conditions. To achieve this goal, the following are performed: (1) construct a random forest model to perform real-time multi-year forecasting (2) integrate the random forest predictive model and K-means clustering into a real-time uncertainty assessment framework that would update the forecast in real-time by incorporating new production data dynamically while quantifying both the geologic and operational uncertainties (3) predict the optimal steam allocation over a multi-year time horizon considering the uncertainties estimated in (2).

6.2. Proposed Approach

6.2.1. Case Study 1: Real-time Uncertainty Assessment of Forecasts

To perform real-time uncertainty analysis along with production forecasting, a workflow that integrates geologic uncertainty with the variability in operating strategies is proposed. Hence, our

workflow (Fig. 6-1) consists of two main parts: predictive modeling and real-time uncertainty analysis of model predictions via integration of continuous production data measurements.

All the models in this work are implemented in Matlab (MathWorks, 2020). The modeling starts with selecting the input features. Feature selection involves identifying and formulating attributes that are most relevant to the problem. Typically, irrelevant, redundant, and highly correlated variables are removed. This exercise can be performed using statistical techniques or domain knowledge. The features are then used to create two models:

1. A random forest model that predicts oil production and steam injection time-series from static and dynamic (well operational) variables.

2. PCA with K-means clustering that classifies all well-pairs based on their static variables only.

The random forest algorithm is an ensemble method which aggregates predictions from multiple decision trees to yield a more accurate prediction. Because decision trees are weak learners that are prone to overfitting, they are usually not the ideal predictive learning tool. Hence, the random forest algorithm combines the simplicity of these weak learners with flexibility, resulting in better predictive accuracy (Breiman L, 2001). In this algorithm, a bootstrapped dataset of the same size as the original dataset is created through sampling with replacement, and depending on the number of decision trees, several bootstrapped samples are used to train the trees. The average of the output variables over all trees is the predicted output.

K-means clustering identifies internal structures within data by dividing a set of items (e.g. well-pairs) into clusters or sub-categories based on similarities in the data. In this algorithm,

observations are grouped into k-clusters based on similarities between them, and a measure of this similarity is the squared Euclidean distance (SED).Cluster assignment of the observations is performed by minimizing the mean SED from each observation to its nearest cluster centroid . Ding and He (2014) established a connection between the K-means clustering algorithm and principal component analysis (PCA), and according to them, principal components are continuous solutions of cluster membership indicators in K-means clustering, allowing for the determination of the optimal number of clusters using PCA. PCA is used for the dimensionality reduction of a dataset by projecting it to a lower-dimensional space. Firstly, a mean-adjusted dataset is calculated by subtracting the mean of each variable from every data point in a dataset to eliminate bias, then the singular value decomposition (SVD) method is used to estimate the principal components of the covariance matrix of the data. A detailed explanation of the application of PCA for analyzing SAGD data can be found in Ma et. al. (2015).

The two models created are then integrated into a real-time uncertainty analysis module that performs two functions: (1) updates the forecast in real-time by incorporating the new production data and (2) incorporates both geologic uncertainty and operational uncertainty (i.e., uncertainty in the dynamic variables) into the forecasting. Geologic uncertainty arises due to the incomplete ability to adequately sample the subsurface (very few wells are drilled to map an entire field or reservoir). Operational uncertainty, on the other hand, arises due to not always operating a well optimally (if it can be assumed there is only one optimal way to operate a well). Geologic uncertainty can be assessed using multiple realizations of the uncertain reservoir properties. For example, three different realizations of the geologic data representing the low-, mid-, and high-cases can be selected (e.g., P90, P50, and P10, respectively). a given realization of geologic data

can be deemed most probable and classified as a P50 realization. Subsequently, other realizations can be deemed less probable and classified as either P10 (high case) or P90 (low case). Other realizations (e.g. P25) can be obtained but only the P10, P50, and P90 were used for this work. It is a trivial matter to see how the algorithm can be extended to handle more realizations. In the context of this paper, a P90 realization of geologic data, for instance, is a set of static variables, including porosity, permeability, thickness, saturation, etc., representative of a "low-case". It is defined using a combination of well measurements and domain expertise. In quantifying geologic uncertainty, a Monte Carlo sampling technique is used to draw a finite number of samples.

In this work, production data is used to update the posterior probability distribution of the occurrence of a given well type (e.g., P10, P50, P90) in a Bayesian framework. Multiple realizations of the static variables are then sampled according to the updated posterior distributions. Each realization is then subjected to a cluster assignment (using the K-means clustering results in the modeling step). Depending on the assigned cluster label, values of the dynamic variables are randomly selected from other well-pairs in that same cluster. The key assumption here is that well-pairs that belong to a given geologic group or cluster are expected to be operated similarly. Any variance in well operations for a given cluster then informs operational uncertainty. The probability of inter-cluster membership is not considered in this study.

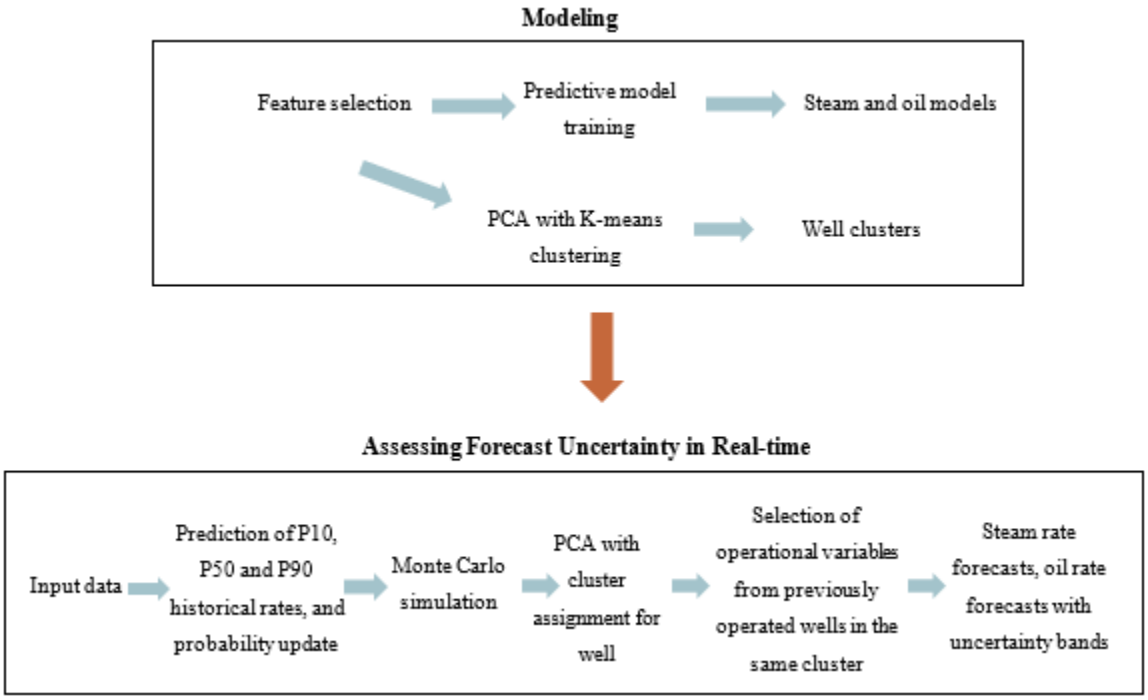


Figure 6-1: Summary of the entire workflow showing the two main parts: modeling and forecast uncertainty assessment.

6.2.1.1. Feature Selection

Here, the variable selection process is done using our domain knowledge of SAGD process dynamics. The input parameters, as well as their units and nominal values, are shown in Table 6-1. The well efficiency proxy, which is one of the input variables, is introduced to capture how the age of a well-pair or pad can affect its performance. Although other formulations of this proxy variable are possible, for the sake of convenience, it is defined as the production start date corresponding to the well-pair being considered. In order to aid the feature selection, parameters are further classified as static or dynamic. Static variables are expected to stay constant for a large portion (or all) of a well-pair's operating life (e.g. well and geologic data). Dynamic variables are

expected to change during day-to-day operations (e.g. operational data). 19 different parameters (4 dynamic and 15 static variables) are available for each well.

Table 6-1: Input variables

Variable	Units	Nominal Value	Type
Elapsed Time	day	1000	Dynamic
Injector Bottomhole Pressure	kPa	3000	Dynamic
ESP Pump Speed	Hz	40	Dynamic
Steam Injection Rate	m^3/day	350	Dynamic
Effective Length	m	1000	Static
Well Efficiency Proxy	dimensionless	2020	Static
Spacing	m	100	Static
Rich Pay Thickness	m	20	Static
Non-Rich Pay Thickness	m	5	Static
Bottom Water Thickness	m	1	Static
Stand-Off	m	3	Static
Rich Vertical Permeability	Darcy	3.5	Static
Non-Rich Vertical Permeability	milliDarcy	50	Static
Rich Porosity	fraction	0.35	Static
Non-Rich Porosity	fraction	0.30	Static
Bottom Water Porosity	fraction	0.30	Static
Rich Oil Saturation	fraction	0.85	Static
Non-Rich Oil Saturation	fraction	0.75	Static
Bottom Water Oil Saturation	fraction	0.45	Static

It should be noted that the elapsed time is considered as an input attribute, and this variable is important since it controls the stages in a well-pair's lifecycle, with each stage having a distinctive

governing process. The addition of the well efficiency proxy variable is based on the idea that as time progresses, operational strategies become more efficient; for example, a well that started operations in the year 2020 is expected to be more efficient, with everything else being equal, than another well that started operations in 2000. Steam injection rate is sensitive to injection pressure. An increase in injection pressure increases the steam chamber temperature (since it is required to maintain steam at saturated conditions) which reduces bitumen viscosity; since a decrease in viscosity drives bitumen drainage, there would be a higher steam injection rate for a given voidage replacement. The effect of injection pressure on latent heat and ultimately heat loss to overburden and underburden may affect steam efficiency. From a well performance perspective, instantaneous changes in pump speed influence oil production rates locally (in time). For the predictive model training process, well, geologic, and operational data as considered as input features for the prediction of steam injection and oil production rates (output features). For the K-means clustering, the inputs are the principal components from PCA on the geologic data only.

6.2.1.2. Predictive Model Training

Two separate random forest models are trained (steam and oil model) using the selected dynamic and static input features. The dataset consists of well (dynamic) data, geologic data, daily steam, and oil rates. The training data set consists of data from 152 Christiana Lake and Foster Creek well-pairs over a 3-year (or 1000 days) period. The well-pairs in this training dataset represent well-pairs that have been operated in the past. Model training is based on 3 years of time-series data to avoid ramp down and blowdown physics (which is not implemented or modeled in this work). Elapsed time and injection bottom-hole pressure are the dynamic parameters for the steam

model, while the dynamic inputs for the oil model are elapsed time, ESP pump speed, and steam injection rate. Hence, the steam injection rate is the output from the steam model and input into the oil model, considering that the operator is almost always steam constrained. The same static parameters are also used as input features in both models.

Before training these models, all the well-pairs are classified as either a P10, P50, or P90 type well. A well is assigned as a certain well type by comparing its rate profiles to each of the P10, P50, or P90 well forecast profiles previously calculated using the semi-analytical model. For example, if a well's actual data matches a P10 forecast more closely than a P50 or P90 forecast, it is classified as a P10 type well. The classification of well-pairs is based on 3 years of production data. Next, for each well type, the geologic and dynamic data of every well corresponding to that well type are assembled to develop the random forest models. For example, to construct the P90 models, the static data (well design parameters, well efficiency proxy, and P90 geologic data), dynamic data (elapsed time and operational data), and time-series steam injection and oil production of every P90 well-pair are included to train the steam injection and oil production random forest models.

Table 6-2: Random forest model properties

Property	Value (Steam Model)	Value (Oil Model)
Ensemble aggregation method	Bagging	Bagging
Number of trees	20	20
Number of predictors	17	18
Number of predictors to select at random for each split	17	18
Number of observations	138,096	138,096
Minimum observations per leaf	5	5

Minimum observations per branch node	10	10
Pruning criterion	Mean Squared Error	Mean Squared Error
Split criterion	Mean Squared Error	Mean Squared Error
Proportion of well-pairs used for training	90%	90%
Proportion of well-pairs used for testing	10%	10%

The coefficients of determination (R2) for the steam and oil models are 0.6 and 0.5 respectively, and these low R2 values are due to noise which is common in most SAGD field data. Results of the feature importance analysis for the two models are shown below. As expected, both models are more sensitive to dynamic data inputs because the response variables (steam and oil rates) also vary with time.

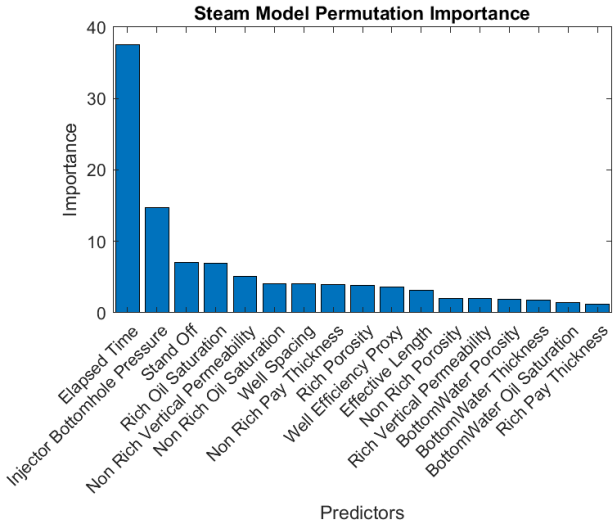


Figure 6-2: Feature importance for the steam model.

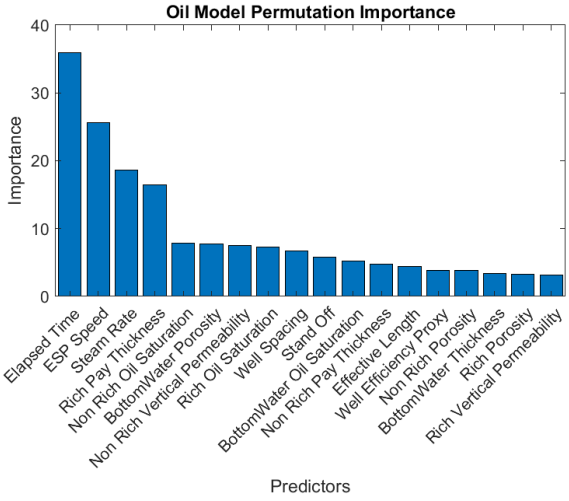


Figure 6-3: Feature importance for the oil model.

6.2.1.3. PCA and K-Means Clustering

Here, principal component analysis (PCA) on static data for dimensionality reduction is implemented. The 15 static variables are reduced to 6 principal components in such a way that a significant amount of its variance is explained. Results from the PCA are shown in Fig. 6-4. Since the 6 principal components explain 80% of the variance of the data, the optimal number of clusters is also set to 6 in accordance with the findings in Ding and He (2014). Although this approach is used in this study, clustering can also be performed directly on the static data.

Fig. 6-5 shows the 3D and 2D clustering results in the principal components space. Each plot represents a well and the clusters are highlighted with different colors. The model training, data compression (through PCA), and K-means clustering processes are shown in Fig. 6-6. The outcome of this step is to identify a set of clusters based on static (geologic) data alone. The results are used to classify any new well based on its static data. Besides, the appropriate/historical

operational strategies for a given cluster can be obtained by assembling the corresponding dynamic (operational) variables from all well-pairs associated with that particular cluster.

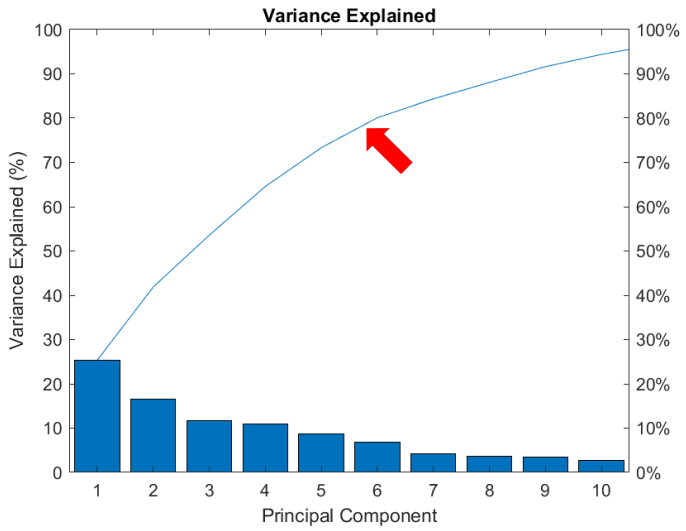
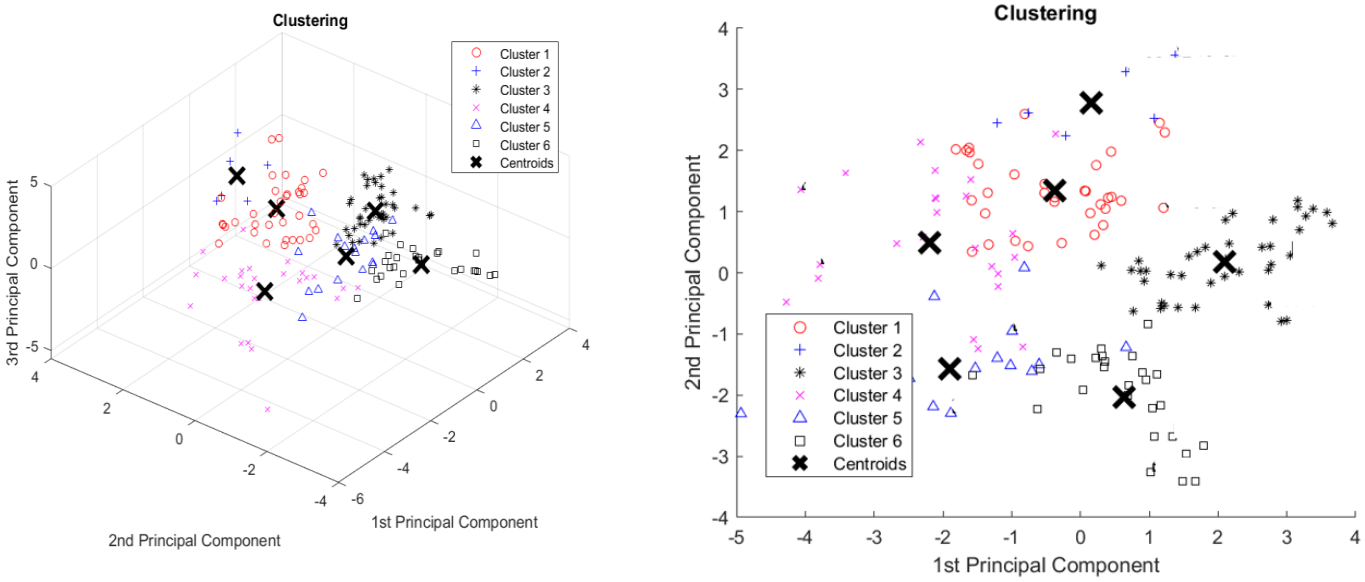


Figure 6-4: Percentage of variance explained from PCA.



(a) (b)

Figure 6-5: K-Means Clustering. (a) – cluster assignments-3D, (b) – cluster assignments-2D

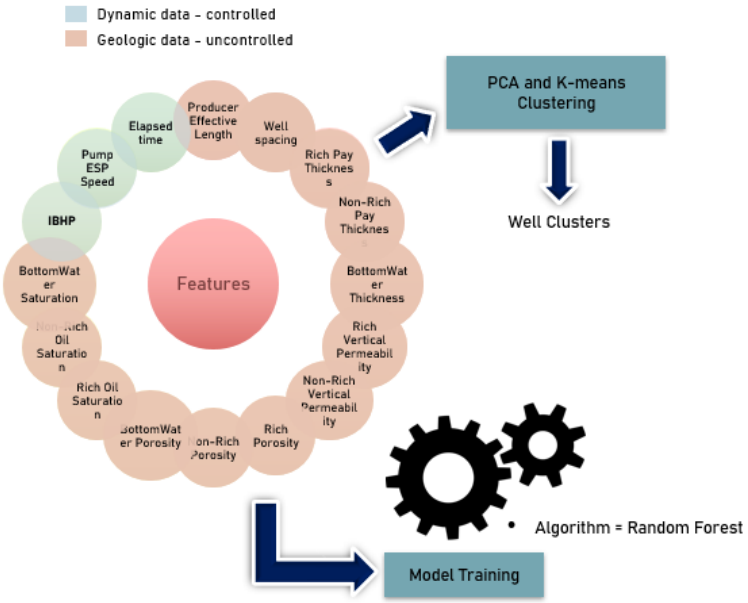


Figure 6-6: Modeling process.

6.2.1.4. Assessing Forecast Uncertainty in Real-time

In this section, we demonstrate how the predictive models developed in section 6.2.1.3 can be used to forecast production for a new well based on its static data and subsequently update this forecast and the associated uncertainty as new production data is collected. As discussed in section 6.2.1.2, all the well-pairs in the training dataset have already been classified as either a P10, P50, or P90 type well. The prior probability distribution of the occurrence of a P10, P50, or P90 type well, $P(P_{R_0})$, can be estimated using Eqn. 1.

$$P(P_{R_0}) = \frac{N_{P_R}}{N_{wells\_total}} \quad (1)$$

subscript R denotes 10, 50, or 90, while P_R represents P10, P50, or P90. N_{P_R} is the number of P_R type well-pairs and N_{wells_total} is the total number of well-pairs used in the training dataset. For a new well with no production data, the probability distribution for the well type of this well is that of the prior distribution $P(P_{R_0})$. Therefore, $P(P_{R_0})$ signifies 3 probabilities: that a well is a P10 type well, a P50 type well, or a P90 type well, at day zero, respectively.

As more production data are obtained, $P(P_{R_0})$ is updated to posterior probability $P(P_{R_0}|P_{R_t})$. $P(P_{R_0}|P_{R_t})$ represents the probability of a well being classified as a well type P_R at day zero conditional to its classification based on production data collected up to day t. The updating procedure is further explained as follows:

i. After x production timesteps of Δt (i.e., $t = x\Delta t$), a measure of closeness (or deviation) of actual production data to the model prediction is defined as:

$$e_{P_R} = \frac{\left|\sum_{i=1}^{x}\left(q_{o_{P_R}}^{predicted,i} - q_o^{actual,i}\right)\right|}{x}, \quad (2)$$

where e_{P_R} is the deviation of actual production data from the P10, P50, or P90 predictions from the random forest models: $q_{o_{P_R}}^{predicted,i}$ is the model's P10, P50, or P90 oil production prediction at timestep i, and $q_o^{actual,i}$ is the actual oil production data at timestep i.

ii. $P(P_{R_0}|P_{R_t})$ is estimated using the Bayes theorem as shown in Eqns. 3.

$$P\left(P_{R_0}|P_{R_t}\right) = \frac{P\left(P_{R_t}|P_{R_0}\right) P\left(P_{R_0}\right)}{P\left(P_{R_t}\right)}, \tag{3}$$

where $P\left(P_{R_t}|P_{R_0}\right)$ is the likelihood function or the conditional probability that a well is a P_R type well after producing up until day t, given P_{R0}. $P\left(P_{R_t}|P_{R_0}\right)$ is calculated using Eqn. 4 below.

$$P\left(P_{R_t}|P_{R_0}\right) = \frac{\frac{1}{e_{P_R}}}{\frac{1}{e_{P_{10}}} + \frac{1}{e_{P_{50}}} + \frac{1}{e_{P_{90}}}} \tag{4}$$

Eqn. 3 can also be expressed as:

$$P\left(P_{R_0}|P_{R_t}\right) = \frac{P\left(P_{R_t}|P_{R_0}\right)P\left(P_{R_0}\right)}{P\left(P_{R_t}|P_{10_0}\right)P\left(P_{10_0}\right) + P\left(P_{R_t}|P_{50_0}\right)P\left(P_{50_0}\right) + P\left(P_{R_t}|P_{90_0}\right)P\left(P_{90_0}\right)} \tag{5}$$

Since the denominator in Eqn. 5 cannot be assessed easily, a normalization constant, α is used (Russell and Norvig, 2016). Hence, $P\left(P_{R_0}|P_{R_t}\right)$ is calculated by computing the terms within the angle brackets, followed by normalization as shown in Eqn. 6.

$$P\left(P_{R_0}|P_{R_t}\right) = \alpha\langle P\left(P_{R_t}|P_{R_0}\right) P\left(P_{R_0}\right)\rangle \tag{6}$$

Step (ii) is repeated as data is obtained, hence, there is a resulting change in $P\left(P_{R_0}|P_{R_t}\right)$ with time.

At time *t*, geologic uncertainty is assessed by sampling the P10, P50, and P90 geologic data multiple times with replacement. The probability distribution for each sampling is $P\left(P_{R_0}\right)$ when determining forecast uncertainty at day zero, while the probability distribution is $P\left(P_{R_0}|P_{R_t}\right)$ for forecasts after production up till day, t.

For each sampled set of static variables, which is drawn from either a P10, P50, or P90 geologic data is converted to principal components, which are fed into the K-means clustering for cluster assignment and the dynamic variables are extracted from all well-pairs corresponding to that particular cluster (section 6.2.1.3), capturing the operational uncertainty. Next, both the static and dynamic variables are used to predict/forecast SAGD injection/production using the trained random forest models. Hence, the coupling of geologic and operational uncertainty is achieved. An example is presented in Fig. 6-7. Assuming we sample either the P10, P50, or P90 geologic data of Well-X, Well-X is assigned to a cluster based on the sampled geologic data, then *n* sets of operational constraints are extracted (where n is the number of well-pairs belonging to that particular cluster). Next, the random forest models are executed n times to forecast n Well-X's injection and production profiles from day t. Finally, Well-X's geologic data is sampled k times, resulting in $k \times n$ predicted injection and production profiles. The $k \times n$ predictions represent the uncertainty in forecast due to both geologic and operational uncertainties, and this uncertainty band or interval varies as $P\left(P_{R_0}|P_{R_t}\right)$ changes with time. It should be noted that the actual number of unique rate profiles is $3 \times n$ since there are just 3 sets of geologic data.

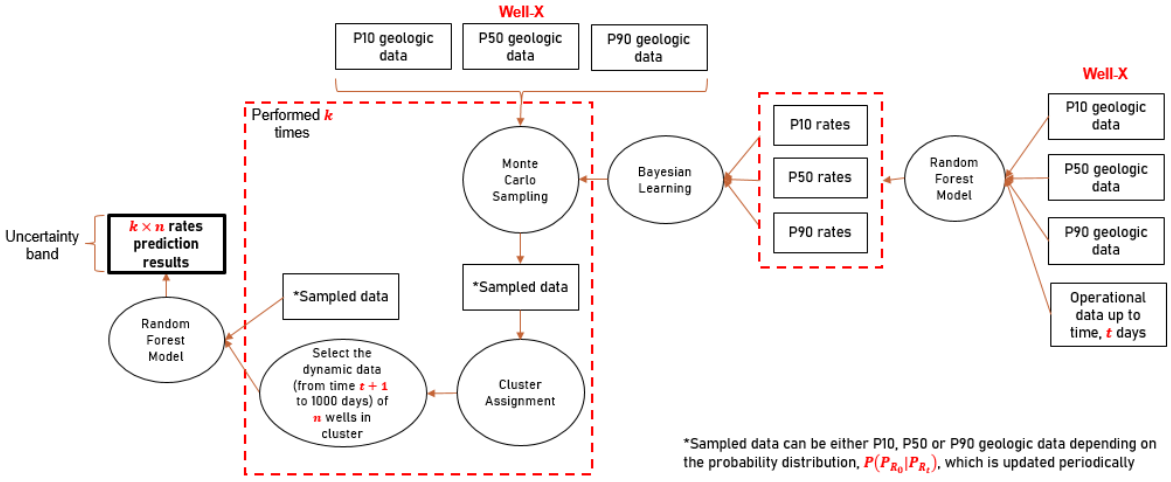


Figure 6-7: A forecasting and uncertainty analysis case for a hypothetical well, Well-X

6.2.2. Case Study 2: Optimization of Steam Allocation

Optimization of steam allocation is often challenging due to the uncertainty (or confidence) interval being high for some well-pairs. A probabilistic approach is formulated, such that the greater the uncertainty, the lesser the weight that should be applied to the steam allocation determination for that well-pair. A simple rescaling to available steam capacity gives the absolute steam allocation for each well-pair. As time progresses, and due to the Bayesian updating, the confidence interval converges and becomes fairly constant (or stable) for each well-pair. At that point, the probability of steam allocation weighting, based on the confidence interval becomes unnecessary and the allocation process can proceed as determined by any optimization algorithm. A common optimization algorithm is the genetic algorithm (GA) in which solutions are represented as chromosomes; a population of solutions is initialized randomly. Each chromosome is evaluated according to a fitness or objective function, which is a measure of how good a solution

or chromosome is when compared to other solutions in the same population. A process of fitness-based selection of parent chromosomes and their recombination is performed iteratively, and successive generations of the population are created with a decrease or increase in the fitness values of the chromosomes, depending on whether the optimization is a minimization or maximization problem, until a stopping criterion is met.

6.2.2.1. Optimization Framework

In this sub-section, a steam allocation optimization framework that proceeds after all the well-pairs have had their uncertainty band stabilized is presented. In essence, this is equivalent to using a history-matched model for steam allocation optimization. The GA-based optimization framework is performed for 10 well-pairs. An economic objective function is defined based on NPV. Usually, GA is used for minimization problems, and since the aim is to maximize NPV, the objective function is defined as $\frac{1}{NPV}$. Note that due to the correlation between steam injection rate and injection bottom-hole pressure (the independent variable), the optimization of injection bottom-hole pressure implies that steam injection rate is optimized. Hence, the decision variables are the injection bottom-hole pressures. To reduce the dimension of the decision variable space and complexity of the optimization problem, monthly injection bottom-hole pressures are used.

A salient assumption here is that all the well-pairs used in demonstrating this optimization process have their uncertainty interval stabilized after 12 months of operation (which may not necessarily be so in reality). Hence, for each well-pair, we select, out of the P10, P50, and P90 predictions, the one that gives the least deviation between actual and predicted data after 12 months starting

from day zero. The geologic data associated with that prediction are extracted for the optimization process. Since we will be making forecasts for the next 22 months, optimization is based on a 22-month forecast period (between months 13 and 34). To estimate the oil rate for a given well-pair, we used the producer operational data of another well-pair in the training dataset, whose geologic data is the closest to the extracted geologic data of that given well-pair (in terms of minimum Euclidean distance in the PCA space).

The decision variables are given by;

$$\left[P_{inj}\right] = \begin{bmatrix} P_{inj_{1,1}} & P_{inj_{1,2}} & \cdots & P_{inj_{1,z}} \\ \cdot & \cdot & \cdots & \cdots \\ P_{inj_{m,1}} & P_{inj_{m,2}} & \cdots & P_{inj_{m,z}} \end{bmatrix} \tag{7}$$

$\left[P_{inj}\right]$ denotes the decision matrix and $P_{inj_{m,z}}$ is the injector bottom-hole pressure for well z at month m. In this study, z = 10 and m = 22., hence the number of decision variables is 220 (10 well-pairs × 22 months forecast period/well pair). Eqn. 7 can be expressed in the vectorized form as;

$$\vec{v} = [v_1\ v_2 \ldots \ldots\ v_{z\times m}] = \left[P_{inj_{1,1}}\ \ldots\ P_{inj_{m,1}}\ P_{inj_{1,2}} \ldots\ P_{inj_{m,2}} \ldots\ P_{inj_{1,z}} \ldots\ P_{inj_{m,z}}\right] \tag{8}$$

where $\vec{v}$ is the decision vector. The objective function, $f(\vec{v})$ is defined as;

$$f(\vec{v}) = \left[\frac{1}{\text{NPV}}\right] \tag{9}$$

NPV (assuming there are no initial investments) is calculated using Eqn. 10.

$$\text{NPV} = \sum_{i=1}^{m} \frac{6.2898 \cdot \left[\left(\sum_k^{N_{wells}} Q_{o_i}\right) \cdot P_{WCS} - \left(\sum_k^{N_{wells}} Q_{s_i}\right) \cdot \frac{X_{steam}}{cSOR_{avg}}\right]}{(1+r)^i} \tag{10}$$

where Q_{o_i} and Q_{s_i} are the total oil production and steam injection volumes (in m^3) for month i (assuming there are 30 days in a month), P_{WCS} is the price per barrel for Western Canadian Select oil grade, X_{steam} is the cost of steam (per barrel of oil), $cSOR_{avg}$ is the average cumulative steam-to-oil ratio across SAGD well-pairs, N_{wells} denotes the total number of well-pairs and r is the monthly discount rate. 6.2898 is the conversion rate from barrel to m^3, and P_{WCS}, X_{steam}, $cSOR_{avg}$ and r are assumed to be equal to \$25/bbl., \$2.5/bbl., 2.5 bbl./bbl. and 9% respectively (Cenovus Energy Inc, 2018, 2019; Millington and Murillo, 2015; National Energy Board, 2006). Usually, discount rates are per annum; hence the conversion of an annual discount rate to a monthly rate is shown in Eqn. 11 below.

$$r = (1 + d)^{\frac{1}{12}} - 1, \tag{11}$$

where d denotes the annual compound discount rate.

The GA optimization parameters are shown in Table 6-3. The initial population is sampled from a set of uniform probability distributions. Since the typical operator of a SAGD field is usually steam constrained, a good step is to add this constraint to the loss function of the predictive model used to resolve an optimization problem. However, the random forest method (the modeling technique for this case study) does not allow easy access to the prediction loss function, as opposed to other techniques such as a Long Short Term Memory network (LSTM). An LSTM network can be trained using various gradient-based techniques, where constraints can be explicitly defined as part of the loss function using techniques such as the Lagrangian multipliers. On the other hand, such formulation is not common in the naïve form of the random forest modeling. To account for steam capacity constraints during both model prediction and steam allocation optimization in this work,

we, therefore, explicitly set the upper bound (a proxy for steam constraint) using the operational data from a subset of the training dataset. To define the lower and upper bounds of the decision vector for each of the well-pairs, we select, out of the clusters they are assigned to, n injector bottom-hole pressure measurements (between months 13 and 34). The lower bound is the 10th percentile of the n injector bottom-hole pressures values while the upper bound is the 90th percentile. To establish a baseline for comparison, a base case is constructed based on the median injector bottom-hole pressure.

Table 6-3: Parameters for GA optimization

Parameter	Value
Population size	50
Number of generations	50
Mutation operator	Gaussian
Crossover operator	Scattered
Number of decision variables	220

6.3. Results and Discussions

6.3.1. Real-time Uncertainty Assessment of Forecasts

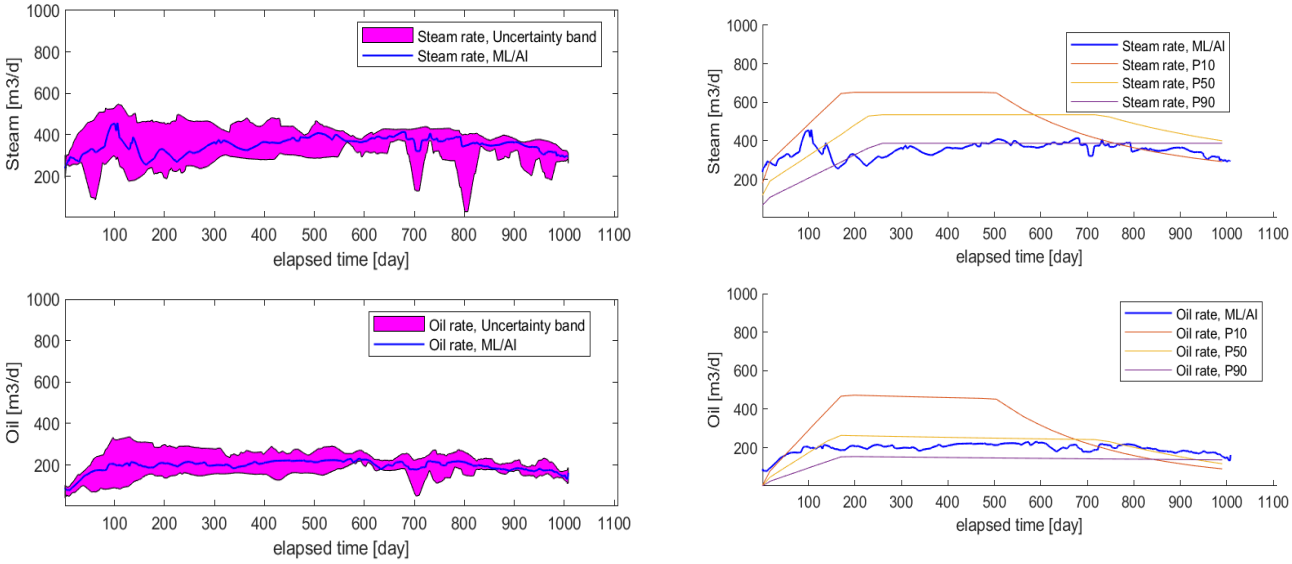
In this section, we present the results for real-time uncertainty assessment using the proposed workflow on Well-1 and compared the predictions against the P10, P50, and P90 forecast profiles from a semi-analytical SAGD model which is based on Butler's theory as shown from Fig. 6-8 to

Fig. 6-16. The figures are presented for every 100 days of live data obtained, with the uncertainty assessment plots to the left and the comparison plots to the right. The dips in the uncertainty band (or interval) are the shut-ins or other operational events being transferred from the well-pairs in the training dataset belonging to the same cluster as Well-1. The median value of the forecasts (or predictions) which make up the uncertainty band is represented with a blue line.

The proposed approach can assess prediction uncertainty in real-time as more data are obtained. As shown in Fig. 6-8a, at day zero, the prediction uncertainty interval at the early SAGD stage is large, and this is due to limited data (as there are no prior production data) and the high variability (or uncertainty) in the start-up and ramp-up phase of SAGD well-pairs. A similar result can also be seen after production for 100 days (Fig. 6-9a) while the model prediction data initially follow the P10 profile (Fig. 6-8b and Fig. 6-9b). The forecasts (and its associated uncertainty) are updated as more data are obtained, leading to the reduction in the uncertainty interval (Fig. 6-10a and Fig. 6-11a), and better injection/production forecasts compared to the semi-analytical SAGD model (Fig. 6-10b and Fig. 6-11b), up till after 300 days of production in which there are no significant changes in the interval (Fig. 6-12). This is because $P\left(P_{R_0}|P_{R_t}\right)$ does not change significantly after 300 days, leading to the sampling of the same geological data and subsequent convergence to a particular cluster (cluster 6). Hence, similar prediction uncertainty interval onwards as shown in Figs. 6-12 to 6-16.

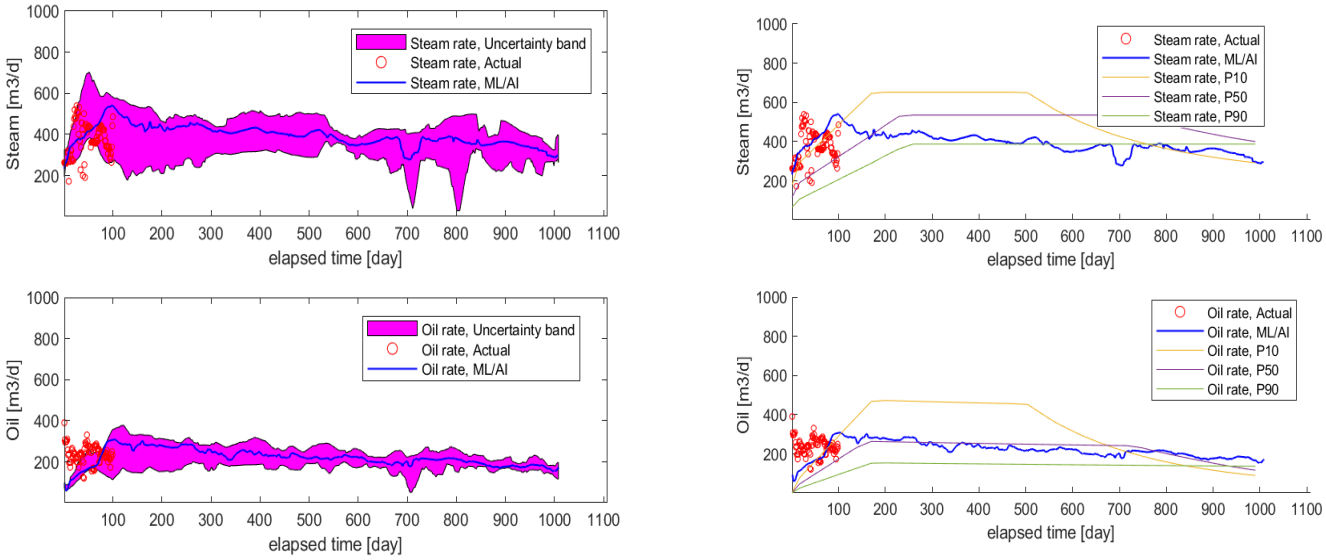
In addition to predicting more accurately than the semi-analytical SAGD model, the implementation of this workflow is faster than numerical reservoir simulation as it took approximately 1 hour of computing time in total using a workstation with 16.0 GB installed RAM

and Intel(R) Core (TM) i5 processor while a 10 well-pair simulation using available commercial reservoir simulators is expected to take several days.



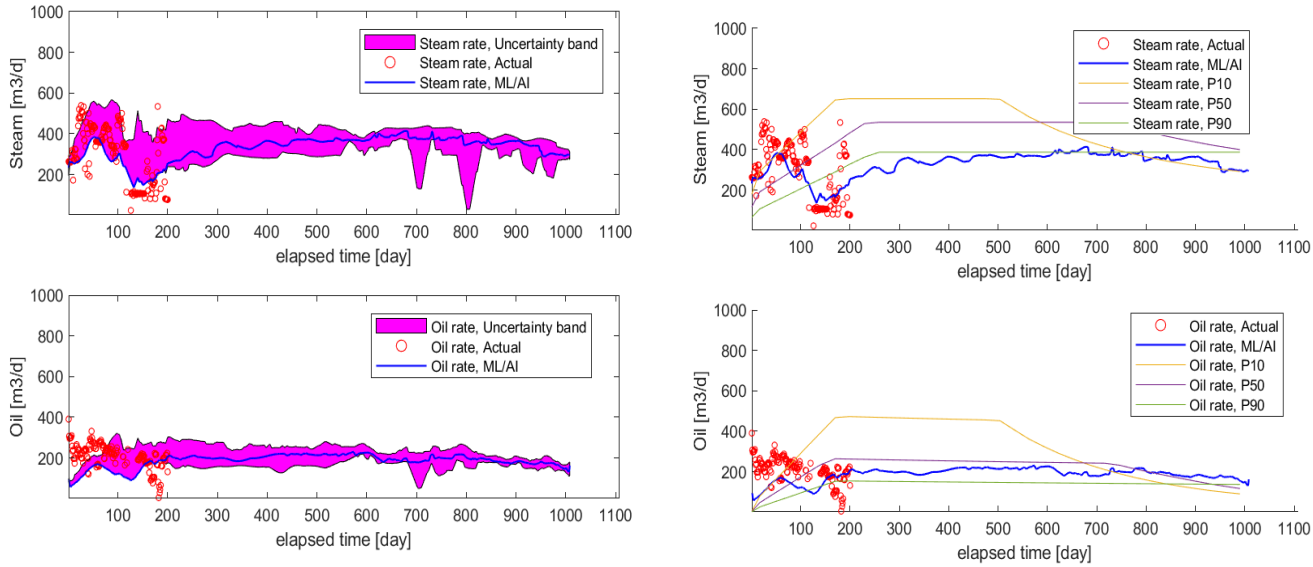
(a) (b)

Figure 6-8: No prior data. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



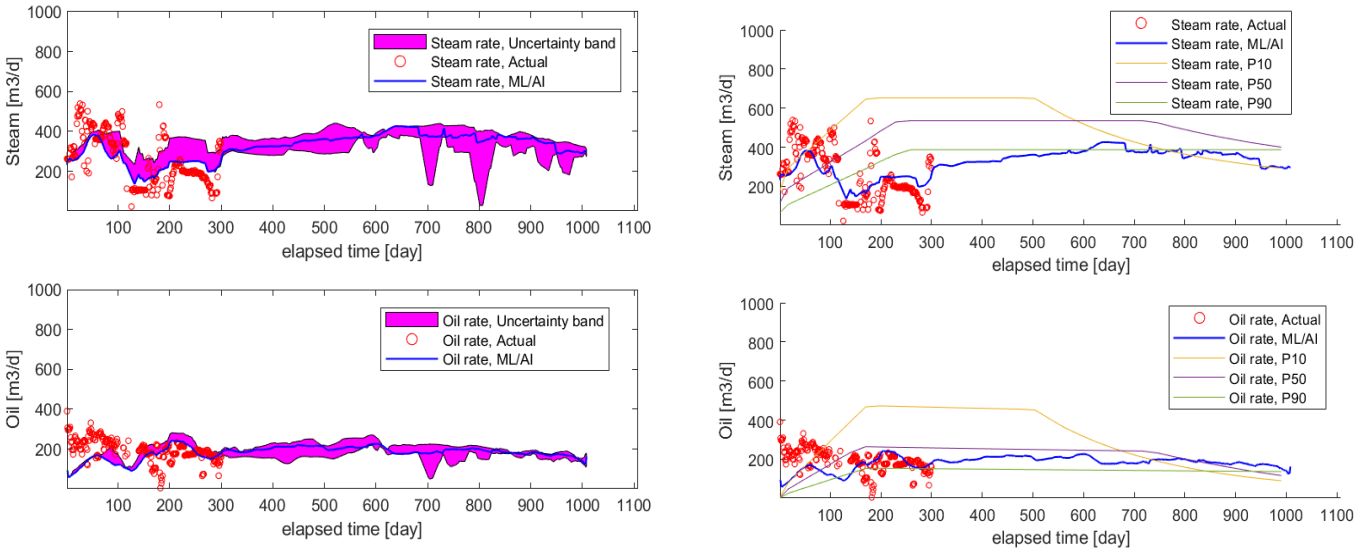
(a) (b)

Figure 6-9: After 100 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



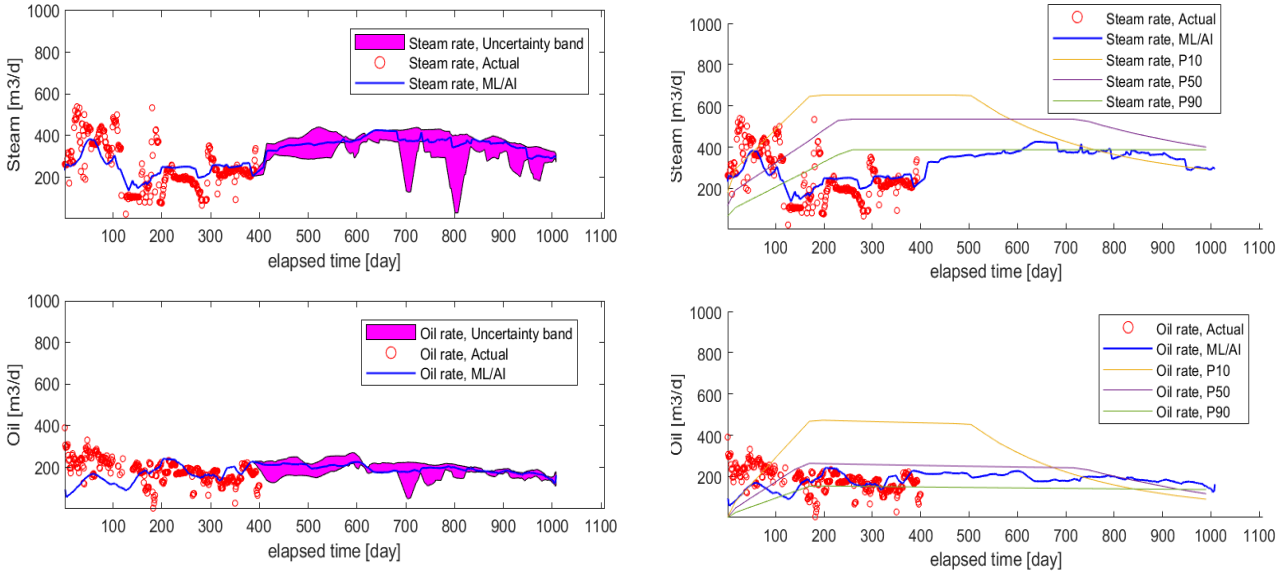
(a) (b)

Figure 6-10: After 200 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



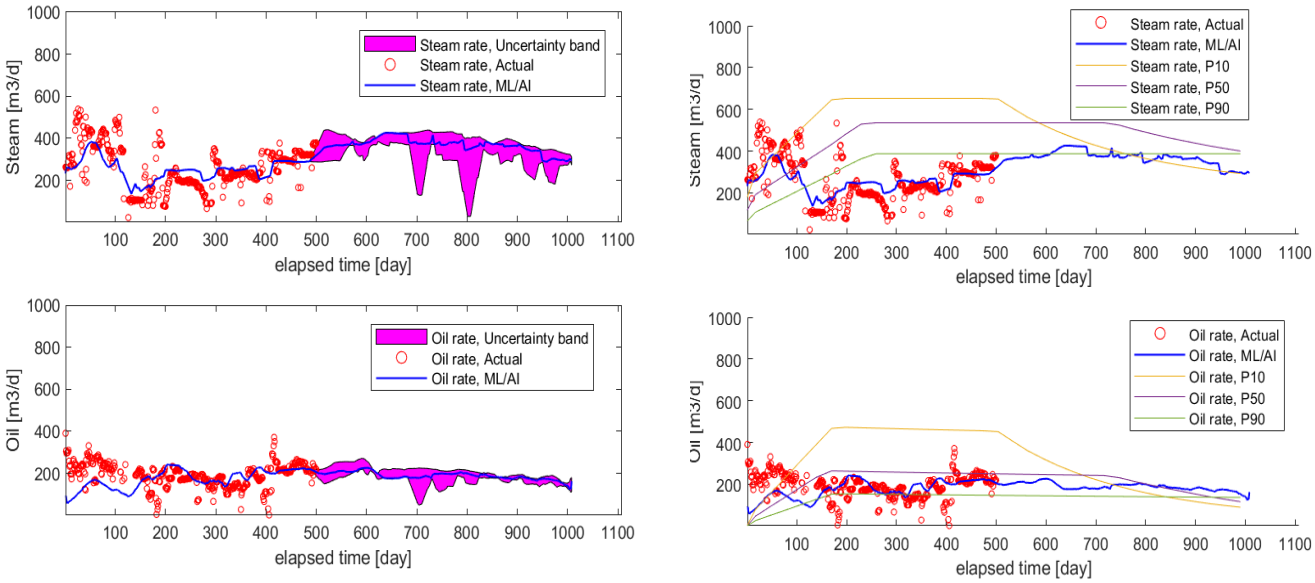
(a) (b)

Figure 6-11: After 300 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



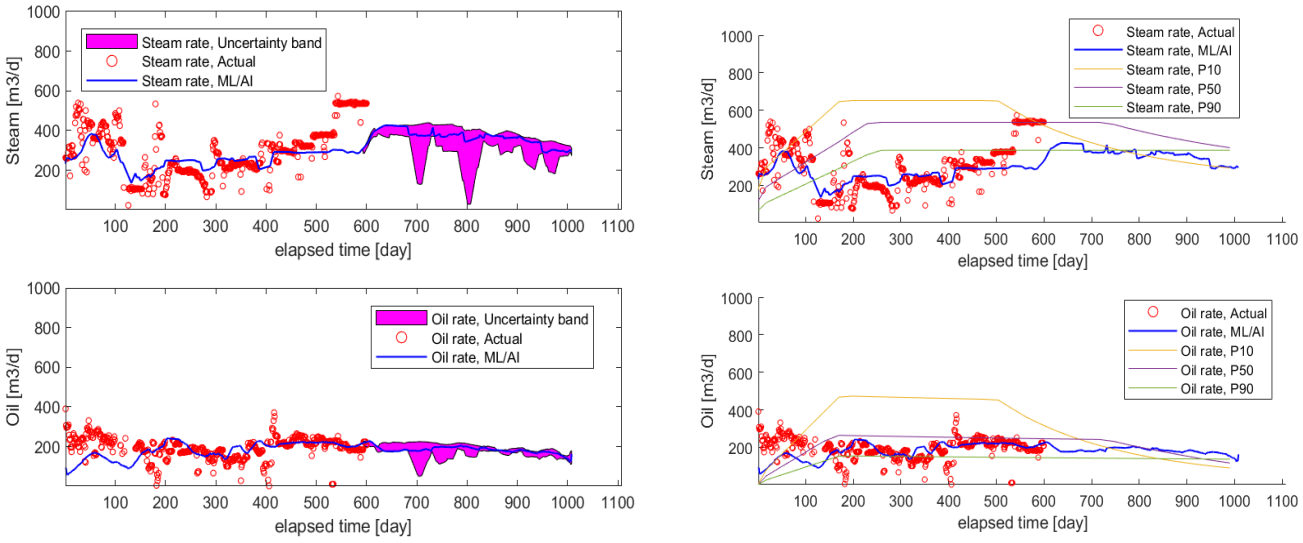
(a) (b)

Figure 6-12: After 400 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



(a) (b)

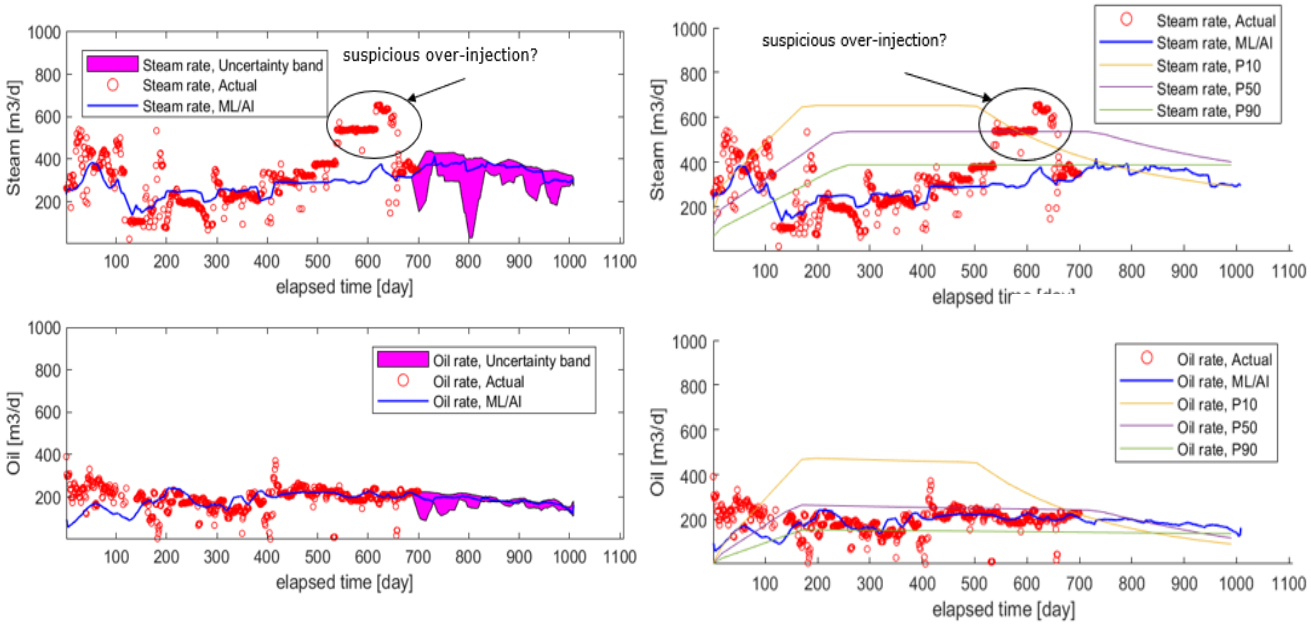
Figure 6-13: After 500 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



(a) (b)

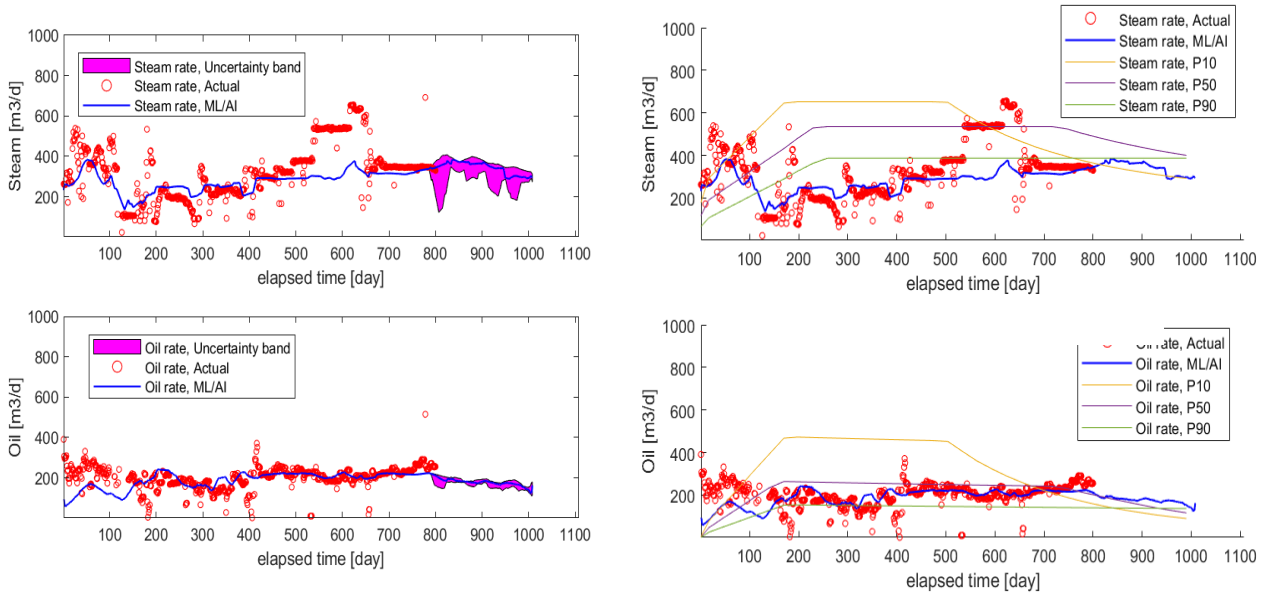
Figure 6-14: After 600 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model

Furthermore, since operational uncertainty is based on the multiple probable sets of operational constraints within cluster 6, abnormal well events can easily be detected, and this is highlighted in Fig. 6-15. There is a significant difference between the actual and predicted data between the 500th and 700th day, indicating a suspicious case of over-injection (as there is no corresponding increment in oil production); therefore, such detections may allow for faster implementation of corrective measures.



(a) (b)

Figure 6-15: After 700 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model



(a) (b)

Figure 6-16: After 800 days of production. (a) – prediction confidence interval, (b) - comparison with forecast profiles from semi-analytical model

6.3.2. Optimization of Steam Allocation

The results of steam allocation optimization for 10 well-pairs, using the approach described in section 6.2.2 is discussed here. The optimization plot is shown in Fig. 6-17. By operating at the optimal steam injection rates, there is a 5% increase in NPV, from an initial value of $220MM to $230MM.

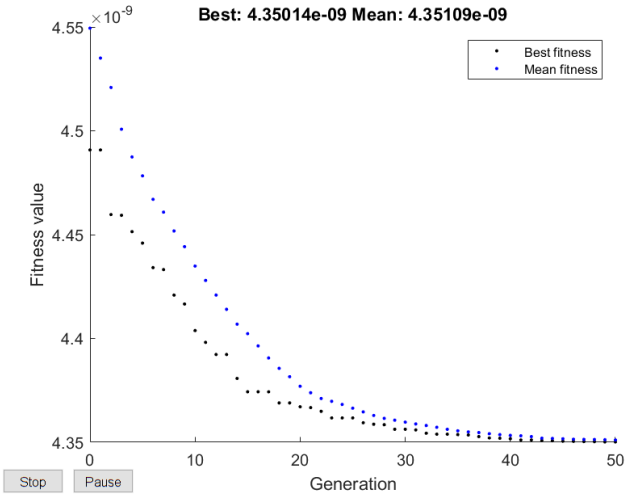
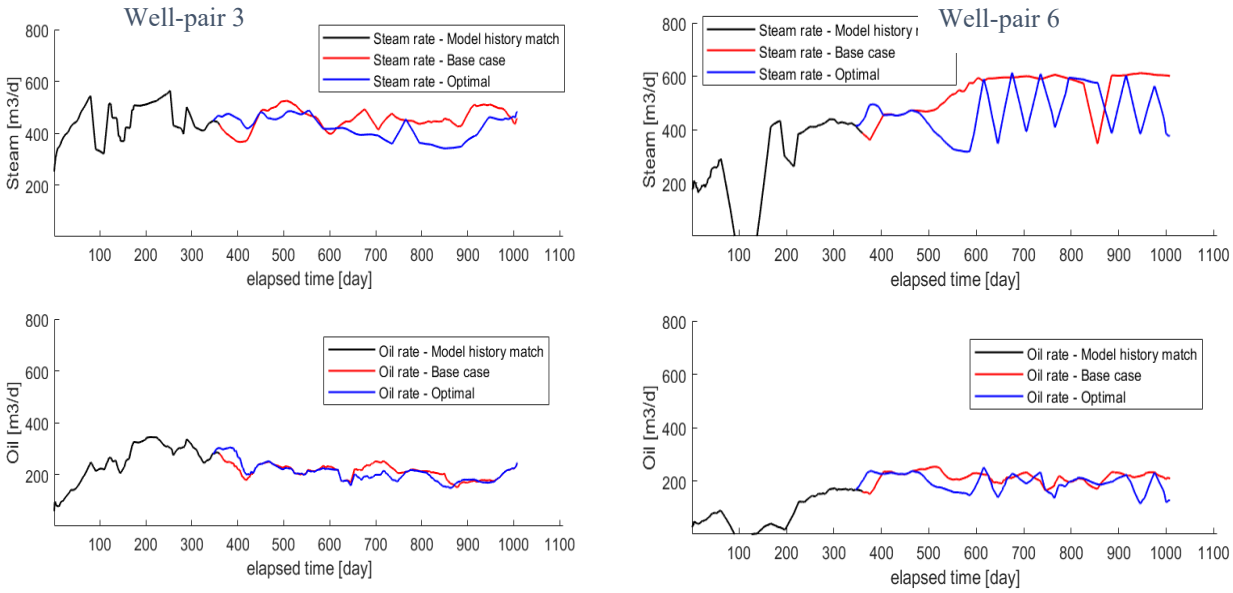


Figure 6-17: Minimization of. fitness value (1/NPV)

The optimal steam injection and their corresponding oil production rates for some of the well-pairs being considered are presented in Figs. 6-18 and 6-19. (The results for the other well-pairs can be found in the appendix section). The model's match to the 12-month historical data is represented with a black line. The red line represents the base case, and it is interesting to note that for most of these well-pairs, the optimal steam injection and oil production rates are similar to the base case, indicating that the well-pairs are being operated near optimal levels.

However, the optimal steam injection and oil production rates are greater than the base case for Well 9, whereas, on average, the optimal steam injection rate is lower than the base case in well-pairs 3 and 6. The implication of these results is that it is optimal to starve well-pairs 3 and 6 of some steam and re-allocate it to Well 9 to produce more oil. Usually, well-pairs that require lower steam injection rates for similar oil production rates as the base case are indicative of good geology. Also, it is evident that Well 9 had more potential to produce oil but was starved of steam in the

non-optimized scenario. This is a typical case of how our optimization framework can be used in real-time steam allocation decisions.



(a) (b)

Figure 6-18: Comparison between base case and optimal rates. (a) – Well-pair 3, (b) – Well-pair 6

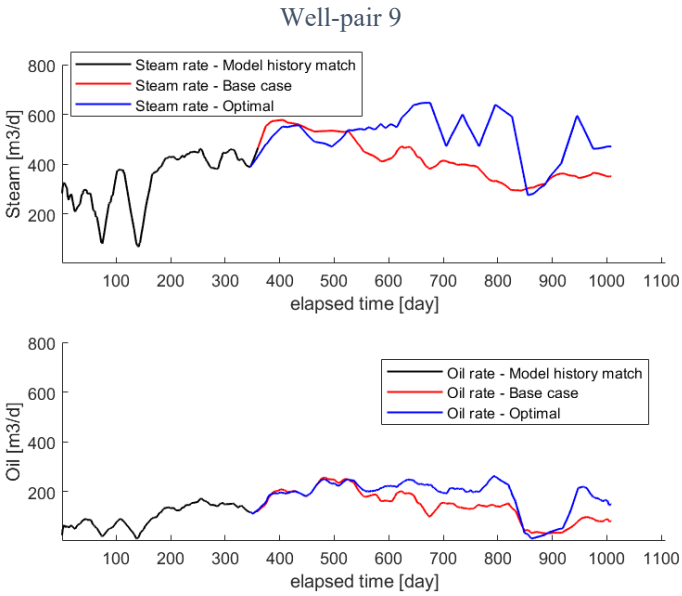


Figure 6-19: Comparison between base case and optimal rates: Well-pair 9

6.4. Summary

A machine learning workflow for real-time SAGD production forecasts and uncertainty quantification of model predictions is presented. We began by introducing the machine learning concepts used in this work, which include various supervised learning (random forest algorithm) and unsupervised learning (PCA and K-means clustering) techniques. Then, utilized the relationship between PCA and K-means clustering analyses as the basis for selecting the optimal number of clusters. Next, we demonstrated a Bayesian-based workflow to determine the prediction uncertainty due to geology through Monte Carlo sampling (the posterior probability distributions are updated by incorporating the new production data) and operational uncertainty. Finally, a steam allocation optimization scheme is adopted: using the steam model, the oil model, and genetic algorithm, the steam injection rates are optimized based on an economic objective function (NPV). Therefore, the following conclusions can be drawn from our work.

1. It is possible to perform real-time forecasting and uncertainty quantification simultaneously.

2. By customizing a series of machine learning tools, we coupled geological and operational uncertainty, allowing for appropriate diagnosis of anomalous well events.

3. The predictive model is faster than numerical reservoir simulation and less expensive, which can lead to significant cost savings. The ability to integrate real-time production data and geologic uncertainties to update our understanding of the subsurface conditions and generate reliable forecasts can also lead to increased oil production, reduced steam usage, increased operational efficiency, and reduced GHG emissions intensity.

4. The machine learning workflow and examples (based on actual SAGD data) explored in this study will find immediate application in field-wide steam allocation optimization and SAGD well/field surveillance.

5. The use of a comprehensive real field dataset is a key novelty of this work. Previous studies often utilize incomplete data from public data or synthetic data, rendering extrapolating their conclusions and models to field applications challenging.

Chapter 7 Conclusions and Recommendations

7.1. Conclusions

The conclusions from this research are summarized as follows.

A MOO workflow that incorporates reservoir simulation, proxy modeling and the NSGA-II algorithm is created to identify the optimal operational parameters for the warm VAPEX process while considering a homogenous and multiple realizations of a heterogenous reservoir case. The workflow also consists of a novel parameterization scheme which considers the interdependency between the operational parameters, ensuring that the injection conditions of solvent mixtures are either at the dew point or within a particular window of superheated conditions. The MOO workflow can reveal subtle patterns in the decision variables that are not easily identifiable through sensitivity analysis. Since reservoir models are usually in form of multiple realizations, and due to the presence of non-condensable gases in injected solvents for many solvent-based techniques, this workflow is robust and has the potential to be applied to real field cases of solvent-based processes.

Also, this research demonstrates the ability of a data-driven workflow to perform real-time SAGD forecasting, uncertainty quantification and steam allocation optimization. The incorporation of operational uncertainty within this workflow facilitates appropriate detection of anomalous well events. Similar to many machine learning framework, the data-driven model is computational efficient and less expensive compared to numerical reservoir modeling.

7.2. Recommendations

A top-down model which couples subsurface physics with a surface network (i.e., wells and surface production system) within an optimization framework should be explored. It is also recommended to use real data to train the proxy models that are required for the generation of the objective function values.

The integration of distributed temperature sensing (DTS) data, distributed acoustic sensing (DAS) data and consideration of completion information (e.g. the impact of splitters, flow control devices (FCDs) and their configurations) is a good opportunity for future work as it could easily be incorporated into the proposed workflow. It is expected that the learnings from this work can also be applied to the understanding of solvent-aided technology. In particular, the SAGD data-driven model should be modified with the knowledge of solvent physics, for faster forecast, uncertainty analysis, optimization and decision making.

References

Afendras, G., and Markatou, M. (2019). Optimality of training/test size and resampling effectiveness in cross-validation. *Journal of Statistical Planning and Inference. Vol. 199, pp 286 – 301, https://doi.org/10.1016/j.jspi.2018.07.005.*

Alberta Energy Regulator (2019). Oil sands. *https://www.aer.ca/providing-information/by-topic/oil-sands, Accessed in November, 2019.*

Al-Gosayir, M., Leung, J., Babadagli, T., and Al-Bahlani, A.M. (2013). Optimization of Steam-over-Solvent Injection in Fractured Reservoirs (Sos-Fr) Method Using Hybrid Techniques: Testing Cyclic Injection Case. *Journal of Petroleum Science and Engineering 110: 74–84. https://doi.org/10.1016/j.petrol.2013.08.036.*

Al-Mudhafar, W.J., and Rao, D. (2016). Integrating Design of Experiments, Proxy Modeling, and Monte-Carlo Simulation for Combined Uncertainty Quantifications of Geological and Production Data in the Cyclic GAGD Process. *https://doi-org.login.ezproxy.library.ualberta.ca/10.2118/182287-MS*

Almashan, M., Narusue, Y., and Morikawa, H. (2020). Estimating PVT Properties of Crude Oil Systems Based on a Boosted Decision Tree Regression Modelling Scheme with K-Means Clustering. *Paper presented at the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Bali, Indonesia, October 2019. https://doi-org.login.ezproxy.library.ualberta.ca/10.2118/196453-MS*

Amirian, E., Leung, J. Y. W., Zanon, S. D. J., and Dzurman, P. J. (2014). Integrated cluster analysis and artificial neural network modeling for steam-assisted gravity drainage performance prediction in heterogeneous reservoirs, *Expert Systems with Applications*

Anirban C. M., Tanushri, S., and Kiranchand, G. R. (2016). Optimization of Automotive Suspension System by Design of Experiments: A Nonderivative Method. *Hindawi Publishing Corporation, Advances in Acoustics and Vibration, Volume 2016, Article ID 3259026, 10 pages, http://dx.doi.org/10.1155/2016/3259026*

Aulia, A., Jeong, D., Saaid. I.M., Shuker, M.T., and El-Khatib, N.A. (2017). A New History Matching Sensitivity Analysis Framework with Random Forests and Plackett-Burman Design. *SPE Symposium: Production Enhancement and Cost Optimisation, Kuala Lumpur, Malaysia*

Beale, M. H., Hagan, M.T., and Demuth, H.B. (2018). Neural Network Toolbox User's Guide. *MathWorks, no. March. https://doi.org/10.1002/0471221546*

Bieker, H.P., Slupphaug, O., and Johansen, T.A. (2006). Optimal Well-Testing Strategy for Production Optimization: A Monte Carlo Simulation Approach. *Paper presented at the SPE Eastern Regional Meeting, Canton, Ohio, USA, October 2006 https://doi-org.login.ezproxy.library.ualberta.ca/10.2118/104535-MS*

Bittencourt, A.C., and Horne, R.N. (1997). Reservoir Development and Design Optimization. *Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 5-8 October. DOI: 10.2118/38895*Mahadevan, S., and Sarkar, S. (2009). Uncertainty Analysis Methods. *US Department of Energy, Washington, DC, USA, November, 2009*

Breiman, L. (2001). Random Forests. Machine Learning 45, 5–32 (2001). *https://doi.org/10.1023/A:1010933404324*

Butler, R.M., and Stephens, D.J. (1981). The Gravity Drainage of Steam-heated Heavy Oil to Parallel Horizontal Wells. *Journal of Canadian Petroleum Technology, 20 (2) (1981), pp. 90-96*

Butler, R.M., and Mokrys, I.J. (1991). A New Process (VAPEX) For Recovering Heavy Oils Using Hot Water And Hydrocarbon Vapour. *Journal of Canadian Petroleum Technology 30 (01). https://doi.org/10.2118/91-01-09.*

Canada Energy Regulator, (2020). Canada's Energy Future 2017 Supplement: Oil Sands Production. *https://www.cer-rec.gc.ca/en/data-analysis/canada-energy-future/2017-oilsands/index.html*

Cenovus Energy Inc, (2018). Christiana lake in-situ Progress Resport. *https://www.aer.ca/providing-information/data-and-reports/activity-and-data/in-situ-performance-presentations*

Cenovus Energy Inc, (2019). 2019 Investors Day presentation. October 2, 2019

Cao, Q., Banerjee, R., Gupta, S., Li, J., Zhou, W., and Jeyachandra. B.. (2016). Data Driven Production Forecasting Using Machine Learning . *SPE Argentina Exploration and Production of Unconventional Resources Symposium*

Cheng, R., Jin, Y., Olhofer, M., and Sendhoff, B.. (2016). A Reference Vector Guided Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation, pp. 1–19, 2016.*

Coimbra, L., Ma, Z., and Leung, J.Y. (2019). Practical Application of Pareto-Based Multi-Objective Optimization and Proxy Modeling for Steam Alternating Solvent Process Design. *Paper presented at the SPE Western Regional Meeting, San Jose, California, USA, April 2019. doi: https://doi-org.login.ezproxy.library.ualberta.ca/10.2118/195247-MS*

Computer Modeling Group. (2019a). WINPROP User Guide.

Computer Modeling Group. (2019b). STARS User Guide. *vol. 1, pp. 1–18.*

Das, S.K. (2008). Distribution of Multi-Component Solvents in Solvent Vapour Extraction Chamber. *1–10. https://doi.org/10.2118/117694-ms*.

Das, S.K., and Butler, R.M. (1996). Diffusion Coefficients of Propane and Butane in Peace River Bitumen. *The Canadian Journal of Chemical Engineering, Vol. 74, pp. 985-992, December 1996.*

Deb K., Agrawal S., Pratap A., Meyarivan T. (2000) A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II. *In: Schoenauer M. et al. (eds) Parallel Problem Solving from Nature PPSN VI. PPSN 2000. Lecture Notes in Computer Science, vol 1917. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45356-3_83*

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA -II. *IEEE Transactions on Evolutionary Computation 6 (0): 182–97.*

Dehdari, V., and Dong, C. (2017). Calibrating a Semi-Analytic SAGD Forecasting Model to 3D Hetetogenous Reservoir Simulations. *SPE Canada Heavy Oil Technical Conference Calgary, Alberta, Canada*

Dickson J. L., Dittaro L. M., Boone T. J. (2013). Integrating the Key Learnings from Laboratory, Simulation, and Field Tests to Assess the Potential for Solvent Assisted – Steam Assisted Gravity Drainage. S*PE Heavy Oil Conference, Calgary, Alberta, Canada*

Deutsch, C., and Journel, A.G. (1998). GSLIB: Geostatistical Software Library and User's Guide. *Second Edition. https://doi.org/10.1080/00401706.1995.10485913.*Fonseca, C. M., and Fleming. P J. (2011). Genetic Algorithms and Multiobjective Optimization. *Multiobjective Genetic Algorithms for Clustering, no. July: 25–50. https://doi.org/10.1007/978-3-642-16615-0_2.*

Dunn, K. (2021). Process Improvement Using Data. *Release 60da13, https://learnche.org/pid/PID.pdf?60da13*

Emissions Reduction Alberta. (2016). CCEMC Project ID: H110031 - BEST Pilot Project Non-Confidential Final Report. *February, 2016*

Fonseca, C. M., and Fleming. P J. (2011). Genetic Algorithms and Multiobjective Optimization. Multiobjective Genetic Algorithms for Clustering, no. July: 25–50. *https://doi.org/10.1007/978-3-642-16615-0_2.*

Fonseca, R.M., Stordal, A.S., Leeuwenburgh, O., Van Den Hof, P.M.J., and Jansen, J.D. (2014). Robust Ensemble-based Multi-objective Optimization. *ECMOR XIV - 14th European Conference on the Mathematics of Oil Recovery, Sep 2014, Volume 2014, p.1 - 14. https://doi.org/10.3997/2214-4609.20141895.*

Fonseca, R.M., Stordal, A.S., Leeuwenburgh, O., Van Den Hof, P.M.J., and Jansen, J.D. (2014). Robust Ensemble-based Multi-objective Optimization. *ECMOR XIV - 14th European Conference on the Mathematics of Oil Recovery, Sep 2014, Volume 2014, p.1 - 14. https://doi.org/10.3997/2214-4609.20141895.*

Gilbert, R.J., Liu, Y., William, A., and Preece, R. (2004). Reservoir Modeling: Integrating Various Data at Appropriate Scales. *The Leading Edge, 23 (8) (2004), pp. 784-788*

Guevara, J.L., Patel, R.G., Trivedi, J.J., (2018). Optimization of steam injection for heavy oil reservoirs using reinforcement learning. *In: In SPE International Heavy Oil Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/193769-MS*

Gupta S. and Gittins S. (2002). Field Implementation of Solvent Aided Process. *Canadian International Petroleum Conference, Petroleum Society*

Hamdi, H., Couckuyt, I., Dhaene, T., and Costa Sousa, M. (2018). Efficient Multi-Objective History-Matching Using Gaussian Processes. *ECMOR XVI - 16th European Conference on*

the Mathematics of Oil Recovery, Sep 2018, Volume 2018, p.1 - 18. https://doi.org/10.3997/2214-4609.201802146.

Horn, J., Nafpliotis, N., and Goldberg, D.E. (1994). Niched Pareto Genetic Algorithm for Multiobjective Optimization. *IEEE Conference on Evolutionary Computation - Proceedings 1: 82–87. https://doi.org/10.1109/icec.1994.350037.*

Hurst, A, Brown, G. C, and Swanson, R. I. (2000). Swanson's 30-40-30 rule. *AAPG Bulletin, Vol. 84, No. 12 (December 2000), pp. 1883 - 1891.*

Hutahaean, J., Demyanov, V., and Christie, M. (2017). On Optimal Selection of Objective Grouping for Multiobjective History Matching. *SPE Journal. 22. 10.2118/185957-PA.*

Hutahaean, J., Demyanov, V., and Christie, M. (2016). Many-objective optimization algorithm applied to history matching. *IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 2016, pp. 1-8, doi: 10.1109/SSCI.2016.7850215.*

Ji, D. (2014). Simulation Study of Steam-Solvent Phase Behaviour in Solvent Aided SAGD Process and Its Effect on Oil Recovery. *Master's Degree Thesis.*

Jiju, A. (2014). Design of Experiments for Engineers and Scientists (Second Edition). *Elsevier, 2014, Pages 7-17, ISBN 9780080994178, https://doi.org/10.1016/B978-0-08-099417-8.00002-X.*

Kathrada, M., and Azri, K. (2019). A Pragmatic Approach to Reservoir Simulation Optimisation Under Uncertainty. *Paper presented at the SPE Reservoir Characterisation and Simulation*

Conference and Exhibition, Abu Dhabi, UAE, September 2019. doi: https://doi.org/10.2118/196659-MS

Khaledi, R., Motahhari, H.R., Boone, T.J., Fang, C., and Coutee, A.S. (2018). Azeotropic Heated Vapour Extraction- A New Thermal-Solvent Assisted Gravity Drainage Recovery Process. *https://doi.org/10.2118/189755-ms.*

Knowles, J., and Corne, D. (1999). The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Pareto Multiobjective Optimisation. *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999 1: 98–105. https://doi.org/10.1109/CEC.1999.781913.*

Kubota, L., and Reinert, D. (2019). Machine Learning Forecasts Oil Rate in Mature Onshore Field Jointly Driven by Water and Steam Injection The Mature Field under Study : Oil Reservoir under Simultaneous Water and Steam Injection. *SPE Annual Technical Conference and Exhibition, Calgary, Alberta, Canada.*

Liao, L., Zeng, Y., Liang, Y., and Zhang, H. (2020). Data Mining: A Novel Strategy for Production Forecast in Tight Hydrocarbon Resource in Canada by Random Forest Analysis. *Institute of Petroleum Technology Conference, Dhahran, Saudi Arabia*

Leyva-Gomez, H., and Babadagli, T. (2017). High-Temperature Solvent Injection for Heavy-Oil Recovery From Oil Sands: Determination of Optimal Application Conditions Through Genetic Algorithm. *SPE Res Eval & Eng 20 (2017): 372–382. doi: https://doi-org.login.ezproxy.library.ualberta.ca/10.2118/183638-PA*

Ma, Z. and Leung, J. Y. (2020a). Design of warm solvent injection processes for heterogeneous heavy oil reservoirs: A hybrid workflow of multi-objective optimization and proxy models. *Journal of Petroleum Science and Engineering, 107186.*

Ma, Z. and Leung, J. Y. (2020b). Integration of deep learning and data analytics for SAGD temperature and production analysis. *Computational Geosciences, 1-17.*

Ma, Z. Leung, J. Y. (2020c). A knowledge-based heterogeneity characterization framework for 3D steam-assisted gravity drainage reservoirs. *Knowledge-Based Systems, 192, 105327.*

Ma, Z., Leung, J.Y., Zanon, S., and Dzurman, P. (2015). Practical implementation of knowledge-based approaches for steam-assisted gravity drainage production analysis. *Expert Systems with Applications 42(21): 7326-7343.*

Ma, Z., Leung, J.Y., and Zanon, S.D. (2017). Practical data mining and artificial neural network modeling for SAGD production analysis. *Journal of Energy Resources Technology, Transactions of the ASME 139(3): 03290.*

MathWorks, (2020). Statistics and Machine Learning Toolbox™ User's Guide. *https://www.mathworks.com/help/releases/R2020b/pdf_doc/stats/stats.pdf*

MathWorks, (2020). Global Optimization Toolbox: User's Guide (r2011b). *https://www.mathworks.com/help/releases/R2020b/pdf_doc/gads/gads.pdf*

Mehana, M., Callard, J., Mansi, M. and Gong, Y. (2019). Integrating Production Analysis With Monte Carlo Simulation for Estimated Ultimate Recovery EUR Prediction. *Paper presented at the SPE Eastern Regional Meeting, October 15–17, 2019*

Millington, D., and Murillo, C.A. (2015). Canadian Oil Sands Supply Costs and Development Projects (2015 – 2035). *Canadian Energy Research Institute (CERI). Study No. 152, August 2015*

Min, B., Kannan, K., and Srinivasan, S. (2017). Quick screening of Pareto-optimal operating conditions for expanding Solvent–Steam assisted gravity drainage using hybrid multi-objective optimization approach. *Energies 10 (7), 966.*

Min, B., Kang, J. M., Chung, S., Park, C., & Jang, I. (2014). Pareto-based multi-objective history matching with respect to individual production performance in a heterogeneous reservoir. *Journal of Petroleum Science and Engineering, 122, 551-566.*

Muhammad, W.A., Monjur, M., and Yacine, R. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy and Buildings. *Volume 147, 15 July 2017, Pages 77-89*

Nasr, T., Beaulieu, G., Golbeck, H., Heck, G., (2002). Novel expanding solvent-SAGD process "ES-SAGD". *In: Paper Presented at the Canadian International Petroleum Conference, Calgary, Alberta, Canada.*

National Energy Board (2006). Canada's Oil Sands: Opportunities And Challengesto 2015: An Update, An Energy Market Assessment. *Calgary, Alberta; 2006*

Nenniger, J.E., and Dunn, S.G. (2008). How Fast Is Solvent Based Gravity Drainage? *https://doi.org/10.2118/2008-139.*

Nenniger, J., and Nenniger, E. (2001). Method and Apparatus for Stimulating Heavy Oil Production. *Canadian Patent CA2351148A1.*

Ngatchou, P., Zarei, A., and El-Sharkawi, M. A. (2005). Pareto Multi Objective Optimization. *84–91.*

Okafor, A (2020). Cooling and machining strategies for high speed milling of titanium and nickel super alloys (Chapter 5). *High Speed Machining, Academic Press, 2020, Pages 127-161, ISBN 9780128150207, https://doi.org/10.1016/B978-0-12-815020-7.00005-9.*

Perkins T. K., and Johnston O.C. (1963) A review of Diffusion and Dispersion in Porous Media. *SPE Journal,Vol. 3, pp. 70-84, March 1963.*

Racz, A., Bajusz, D., and Heberger, K. (2021). Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification. Molecules 2021, 26, 1111. *https://doi.org/10.3390/molecules26041111*

Rostamian, A., Jamshidi, S., and Zirbes, E. (2019). The Development of a Novel Multi-Objective Optimization Framework for Non-Vertical Well Placement Based on a Modified Non-Dominated Sorting Genetic Algorithm-II.

Russel, S.J., and Norvig P. (2016). Artificial Intelligence – A Modern Approach. *Third Edition-Person Education (2016)*

Schaffer, J. D. (1985). Multiple Objective Optimization with Vector Evaluated Genetic Algorithms. *In Proceedings of First International Conference on Genetic Algorithms, 93–100. https://doi.org/10.1017/CBO9781107415324.004.*

Sibaweihi, N., Patel, R.G., Guevara, J.L., Ian D.G., and Trivedi, J.J., (2020). Real-time steam allocation workflow using machine learning for digital heavy oil reservoirs. *Journal of Petroleum Science and Engineering. Volume 199, April 2021, 108168. https://doi.org/10.1016/j.petrol.2020.108168*

Sivanandam, S. N., and Deepa, S. N. (2008). Introduction to Genetic Algorithms. *Springer-Verlag Berlin Heidelberg. https://doi-org.login.ezproxy.library.ualberta.ca/10.1007/978-3-540-73190-0*

Shi, J., and Leung, J.Y. (2014a). Physics-Based Proxy Modelling of Solvent Transport in VAPEX Process. *Canadian Journal of Chemical Engineering 92 (8): 1467–80. https://doi.org/10.1002/cjce.21992.*

Shi, J., and Leung, J.Y. (2014b). Semi-Analytical Proxy for Vapex Process Modeling in Heterogeneous Reservoirs. *Journal of Energy Resources Technology 136(3): 032904.*

van Essen, G., Zandvliet, M., Van den Hof, P., Bosgra, O., Jansen, J.D. (2009). Robust waterflooding optimization of multiple geological scenarios. *SPE J. 14 (1), 202–210.*

Wang, C., and Leung, J.Y. (2015). Characterizing the Effects of Lean Zones and Shale Distribution in Steam-Assisted-Gravity-Drainage Recovery Performance. *SPE Reservoir Evaluation and Engineering 18 (3): 329–45. https://doi.org/10.2118/170101-PA.*

Wang, C., Ma, Z., Leung, J.Y., and Zanon, S. (2018). Correlating Stochastically Distributed Reservoir Heterogeneities with Steam-Assisted Gravity Drainage Production. *Oil and Gas Science and Technology 73. https://doi.org/10.2516/ogst/2017042.*

Zhang, K., Zhou, X., Peng, X., and Zeng, F. (2019a). A Comparison Study between N-Solv Method and Cyclic Hot Solvent Injection (CHSI) Method. *Journal of Petroleum Science and Engineering 173 (June 2018): 258–68. https://doi.org/10.1016/j.petrol.2018.09.061.*

Zhang, Z., Jung, H.Y., Datta-Gupta, A. and Delshad, M. (2019b). History Matching and Optimal Design of Chemically Enhanced Oil Recovery Using Multi-Objective Optimization. *https://doi.org/10.2118/193860-ms.*

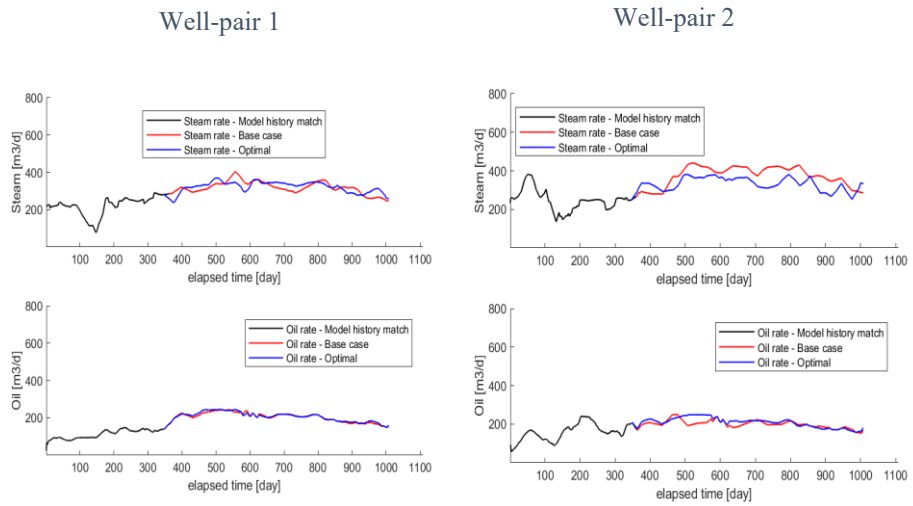
Zheng, J., Leung, J.Y., Sawatzky, R.P., and Alvarez, J.M. (2018a). A proxy model for predicting SAGD production from reservoirs containing shale barriers. *Journal of Energy Resources Technology, Transactions of the ASME 140(12): 122903.*

Zheng, J., Leung, J.Y., Sawatzky, R.P., and Alvarez, J.M. (2018b). An AI-based workflow for estimating shale barrier configurations from SAGD production histories. *Neural Computing and Applications. https://doi.org/10.1007/s00521-018-3365-9.*

Zitzler, E., and Thiele, L. (1998). Multiobjective Optimization Using Evolutionary Algorithm. *Lecture Notes in Computer Science 1498: 292–301.*

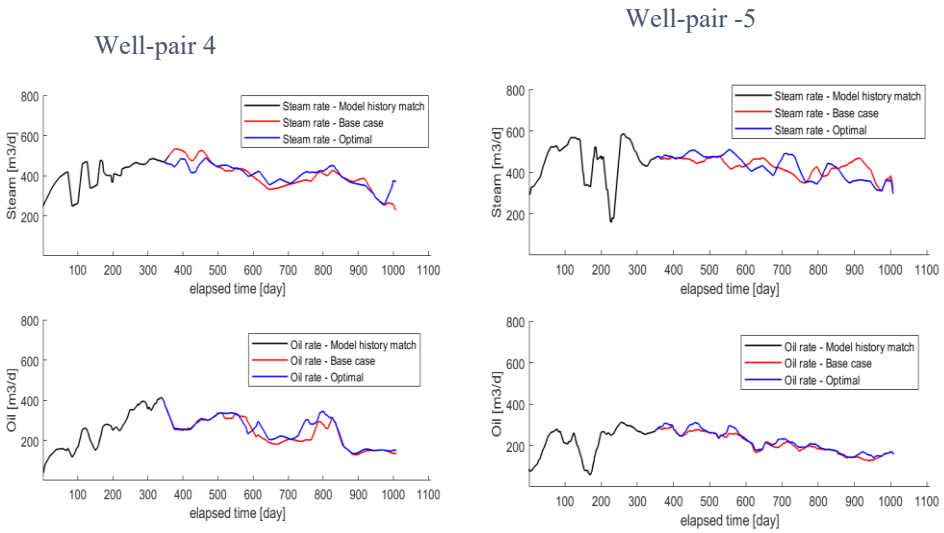
Appendices

Appendix 1: Supplementary steam allocation optimization results from Chapter 6.



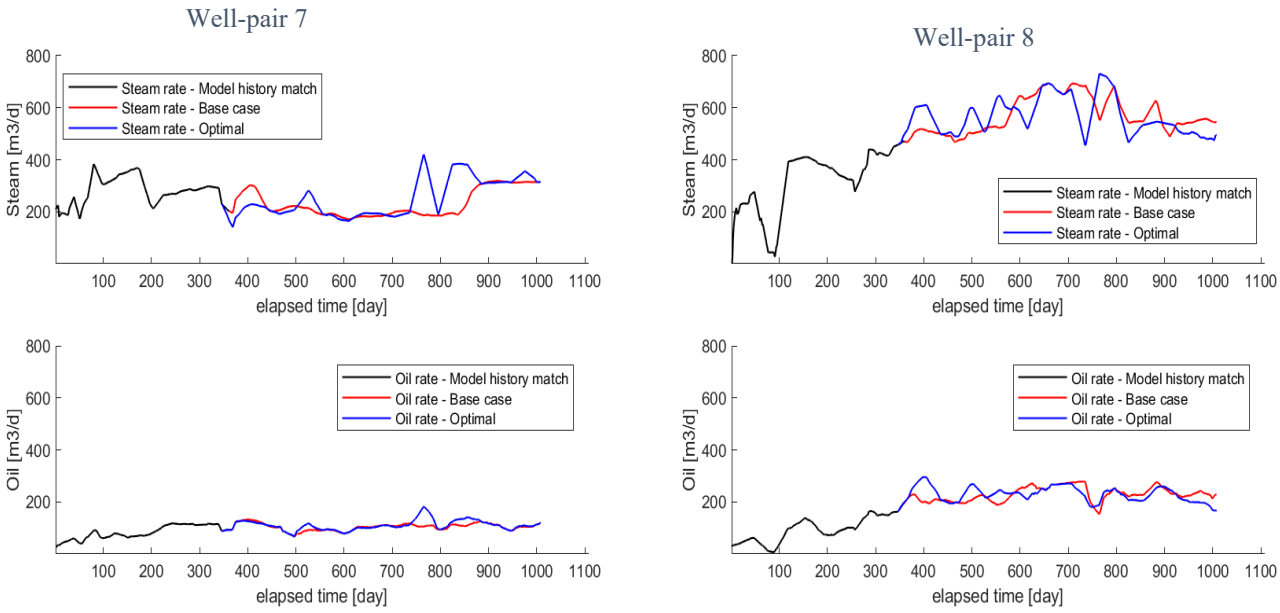
(a) *(b)*

Figure A1: Comparison between base case and optimal rates. (a) – Well-pair 1, (b) – Well-pair 2



(a) *(b)*

Figure A2: Comparison between base case and optimal rates. (a) – Well-pair 4, (b) – Well-pair 5



(a) *(b)*

Figure A3: Comparison between base case and optimal rates. (a) – Well-pair 7, (b) – Well-pair 8

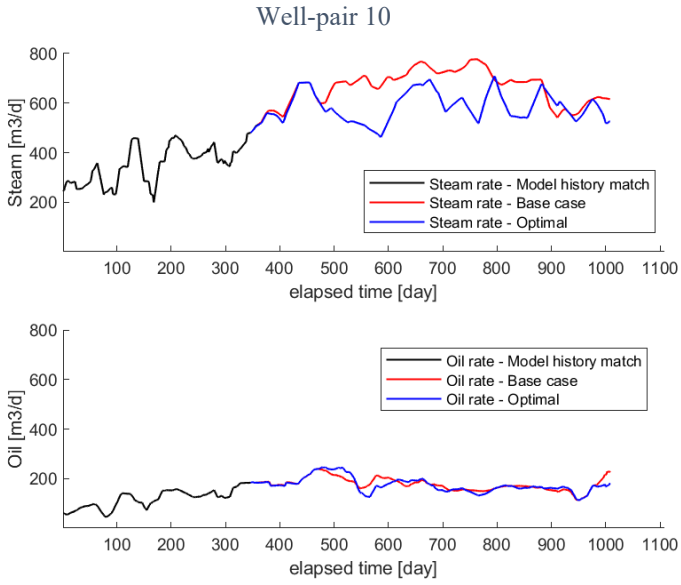


Figure A4: Comparison between base case and optimal rates: Well-pair 10

Appendix 2: Dataset used in proxy model development in Chapter 4.

Case #	Temperature (°C)	Pressure (kPa)	C_3 mole fraction	SolOR (m^3/m^3)	RF (%)	1/RF	Normalized Enthalpy
1	25	2226	0.5	109.545	10.65	9.39	0.093
2	30	2585	0.5	94.840	17.20	5.81	0.171
3	35	3003	0.5	85.894	26.78	3.73	0.278
4	40	3499	0.5	84.357	34.89	2.87	0.383
5	45	4104	0.5	86.290	40.72	2.46	0.480
6	50	4890	0.5	89.142	49.09	2.04	0.616
7	30	2226	0.5	104.999	10.61	9.42	0.100
8	35	2585	0.5	91.234	17.85	5.60	0.180
9	40	3003	0.5	84.106	27.35	3.66	0.287
10	45	3499	0.5	83.670	34.74	2.88	0.385
11	50	4104	0.5	86.499	39.49	2.53	0.475
12	55	4890	0.5	88.938	49.20	2.03	0.618
13	35	2226	0.5	95.778	10.51	9.51	0.115
14	40	2585	0.5	85.982	18.35	5.45	0.196
15	45	3003	0.5	81.325	28.06	3.56	0.299
16	50	3499	0.5	81.973	35.72	2.80	0.395
17	55	4104	0.5	85.037	41.90	2.39	0.498
18	60	4890	0.5	89.144	49.09	2.04	0.622
19	40	2226	0.5	86.821	10.38	9.63	0.131
20	45	2585	0.5	78.343	19.21	5.21	0.220
21	50	3003	0.5	76.935	27.28	3.67	0.311
22	55	3499	0.5	79.498	35.81	2.79	0.411
23	60	4104	0.5	84.262	40.16	2.49	0.490
24	65	4890	0.5	88.342	49.90	2.00	0.633
25	45	2226	0.5	78.676	10.32	9.69	0.147
26	50	2585	0.5	72.555	18.03	5.55	0.231
27	55	3003	0.5	72.525	27.37	3.65	0.330
28	60	3499	0.5	75.814	33.90	2.95	0.414
29	65	4104	0.5	80.183	43.82	2.28	0.533
30	70	4890	0.5	87.754	50.24	1.99	0.640
31	25	1967	0.55	109.977	9.90	10.11	0.083

32	30	2272	0.55	96.590	15.41	6.49	0.151
33	35	2622	0.55	86.165	25.05	3.99	0.253
34	40	3027	0.55	83.429	33.28	3.00	0.355
35	45	3502	0.55	85.838	37.14	2.69	0.431
36	50	4076	0.55	87.148	46.29	2.16	0.557
37	30	1967	0.55	103.891	9.83	10.17	0.092
38	35	2272	0.55	93.343	15.25	6.56	0.157
39	40	2622	0.55	84.282	25.03	3.99	0.259
40	45	3027	0.55	83.300	31.83	3.14	0.348
41	50	3502	0.55	84.767	38.00	2.63	0.440
42	55	4076	0.55	87.927	44.03	2.27	0.543
43	35	1967	0.55	94.194	9.69	10.32	0.108
44	40	2272	0.55	86.003	15.68	6.38	0.174
45	45	2622	0.55	79.531	26.18	3.82	0.280
46	50	3027	0.55	80.882	32.18	3.11	0.358
47	55	3502	0.55	83.438	38.60	2.59	0.450
48	60	4076	0.55	87.480	43.89	2.28	0.546
49	40	1967	0.55	86.452	9.18	10.89	0.121
50	45	2272	0.55	79.020	15.56	6.43	0.191
51	50	2622	0.55	74.218	25.64	3.90	0.294
52	55	3027	0.55	76.812	33.25	3.01	0.383
53	60	3502	0.55	80.748	39.69	2.52	0.472
54	65	4076	0.55	85.537	46.02	2.17	0.569
55	45	1967	0.55	78.011	8.82	11.34	0.135
56	50	2272	0.55	71.775	16.45	6.08	0.215
57	55	2622	0.55	70.327	25.44	3.93	0.311
58	60	3027	0.55	71.976	34.35	2.91	0.408
59	65	3502	0.55	76.948	38.48	2.60	0.479
60	70	4076	0.55	82.787	44.41	2.25	0.573
61	25	1762	0.6	109.358	9.40	10.64	0.077
62	30	2027	0.6	99.437	13.30	7.52	0.131
63	35	2328	0.6	86.997	22.59	4.43	0.227
64	40	2672	0.6	82.917	31.96	3.13	0.334
65	45	3066	0.6	83.910	38.00	2.63	0.421
66	50	3527	0.6	87.470	41.29	2.42	0.496

67	30	1762	0.6	102.000	9.35	10.70	0.087
68	35	2027	0.6	94.575	13.40	7.46	0.140
69	40	2328	0.6	84.029	23.39	4.28	0.239
70	45	2672	0.6	81.802	31.46	3.18	0.334
71	50	3066	0.6	84.361	35.29	2.83	0.404
72	55	3527	0.6	85.831	43.59	2.29	0.517
73	35	1762	0.6	92.864	9.02	11.08	0.102
74	40	2027	0.6	86.235	13.56	7.37	0.158
75	45	2328	0.6	78.876	23.51	4.25	0.257
76	50	2672	0.6	78.667	31.43	3.18	0.347
77	55	3066	0.6	81.472	36.54	2.74	0.421
78	60	3527	0.6	85.419	42.28	2.37	0.512
79	40	1762	0.6	85.993	8.59	11.64	0.115
80	45	2027	0.6	79.165	13.50	7.41	0.173
81	50	2328	0.6	73.402	24.02	4.16	0.278
82	55	2672	0.6	74.939	32.34	3.09	0.372
83	60	3066	0.6	78.053	36.68	2.73	0.441
84	65	3527	0.6	82.362	43.26	2.31	0.535
85	45	1762	0.6	77.600	8.24	12.14	0.128
86	50	2027	0.6	71.761	14.19	7.05	0.195
87	55	2328	0.6	69.033	23.02	4.34	0.289
88	60	2672	0.6	70.747	31.04	3.22	0.377
89	65	3066	0.6	74.079	37.99	2.63	0.469
90	70	3527	0.6	77.817	46.41	2.15	0.575
91	25	1595	0.65	109.097	8.87	11.27	0.070
92	30	1830	0.65	100.818	12.30	8.13	0.120
93	35	2094	0.65	88.554	20.34	4.92	0.205
94	40	2392	0.65	82.663	30.31	3.30	0.312
95	45	2731	0.65	83.848	35.12	2.85	0.387
96	50	3117	0.65	84.712	43.59	2.29	0.495
97	30	1595	0.65	100.546	8.91	11.22	0.084
98	35	1830	0.65	94.319	12.35	8.10	0.131
99	40	2094	0.65	83.108	22.16	4.51	0.226
100	45	2392	0.65	80.921	29.95	3.34	0.317
101	50	2731	0.65	81.206	37.86	2.64	0.410

102	55	3117	0.65	84.658	42.47	2.35	0.493
103	35	1595	0.65	92.006	8.42	11.88	0.096
104	40	1830	0.65	86.730	12.19	8.20	0.146
105	45	2094	0.65	78.313	21.58	4.63	0.239
106	50	2392	0.65	77.327	30.29	3.30	0.335
107	55	2731	0.65	79.489	36.63	2.73	0.417
108	60	3117	0.65	82.696	42.29	2.36	0.500
109	40	1595	0.65	84.625	8.11	12.33	0.109
110	45	1830	0.65	78.740	12.38	8.08	0.163
111	50	2094	0.65	72.526	22.30	4.49	0.262
112	55	2392	0.65	73.220	29.99	3.33	0.350
113	60	2731	0.65	75.846	34.96	2.86	0.421
114	65	3117	0.65	79.427	41.99	2.38	0.513
115	45	1595	0.65	76.524	7.76	12.89	0.122
116	50	1830	0.65	71.796	13.12	7.62	0.184
117	55	2094	0.65	68.117	22.04	4.54	0.278
118	60	2392	0.65	69.324	28.86	3.46	0.358
119	65	2731	0.65	72.362	34.62	2.89	0.435
120	70	3117	0.65	76.032	40.65	2.46	0.520
121	25	1457	0.7	108.839	8.53	11.72	0.066
122	30	1667	0.7	102.339	11.30	8.85	0.110
123	35	1902	0.7	89.573	18.74	5.34	0.189
124	40	2166	0.7	82.433	29.19	3.43	0.297
125	45	2462	0.7	83.156	34.63	2.89	0.374
126	50	2795	0.7	84.506	41.00	2.44	0.463
127	30	1457	0.7	99.127	8.39	11.92	0.080
128	35	1667	0.7	93.945	11.98	8.35	0.127
129	40	1902	0.7	83.661	20.20	4.95	0.211
130	45	2166	0.7	78.977	30.07	3.33	0.312
131	50	2462	0.7	80.312	36.66	2.73	0.394
132	55	2795	0.7	82.757	42.46	2.36	0.480
133	35	1457	0.7	90.878	7.98	12.53	0.093
134	40	1667	0.7	86.567	11.28	8.87	0.138
135	45	1902	0.7	78.215	20.01	5.00	0.226
136	50	2166	0.7	75.497	29.74	3.36	0.328

137	55	2462	0.7	78.143	33.93	2.95	0.392
138	60	2795	0.7	81.172	40.88	2.45	0.483
139	40	1457	0.7	84.219	7.61	13.14	0.105
140	45	1667	0.7	79.252	11.49	8.70	0.155
141	50	1902	0.7	72.407	19.97	5.01	0.242
142	55	2166	0.7	71.537	29.33	3.41	0.341
143	60	2462	0.7	74.358	34.35	2.91	0.410
144	65	2795	0.7	76.860	42.97	2.33	0.512
145	45	1457	0.7	76.269	7.40	13.51	0.118
146	50	1667	0.7	71.928	11.72	8.54	0.172
147	55	1902	0.7	67.680	21.34	4.69	0.269
148	60	2166	0.7	68.165	28.40	3.52	0.350
149	65	2462	0.7	70.168	36.40	2.75	0.443
150	70	2795	0.7	73.629	42.42	2.36	0.527
151	25	1340	0.75	106.834	8.21	12.18	0.063
152	30	1530	0.75	102.076	10.95	9.13	0.105
153	35	1742	0.75	88.706	18.56	5.39	0.184
154	40	1978	0.75	82.429	27.85	3.59	0.282
155	45	2241	0.75	81.787	35.60	2.81	0.374
156	50	2534	0.75	83.709	39.64	2.52	0.444
157	30	1340	0.75	97.557	8.04	12.44	0.077
158	35	1530	0.75	95.980	10.85	9.22	0.119
159	40	1742	0.75	83.029	19.14	5.22	0.201
160	45	1978	0.75	78.996	28.21	3.54	0.296
161	50	2241	0.75	79.846	34.84	2.87	0.377
162	55	2534	0.75	81.762	41.78	2.39	0.466
163	35	1340	0.75	89.603	7.68	13.02	0.090
164	40	1530	0.75	86.942	10.53	9.49	0.131
165	45	1742	0.75	77.498	18.98	5.27	0.215
166	50	1978	0.75	74.079	28.77	3.48	0.316
167	55	2241	0.75	76.643	33.54	2.98	0.384
168	60	2534	0.75	78.439	41.48	2.41	0.479
169	40	1340	0.75	82.404	7.34	13.62	0.102
170	45	1530	0.75	79.663	10.68	9.36	0.147
171	50	1742	0.75	72.511	18.80	5.32	0.231

172	55	1978	0.75	70.338	28.91	3.46	0.335
173	60	2241	0.75	72.471	35.24	2.84	0.412
174	65	2534	0.75	75.502	41.15	2.43	0.493
175	45	1340	0.75	74.994	7.09	14.10	0.115
176	50	1530	0.75	72.084	10.86	9.21	0.164
177	55	1742	0.75	67.603	19.11	5.23	0.249
178	60	1978	0.75	67.329	28.34	3.53	0.346
179	65	2241	0.75	69.583	34.13	2.93	0.422
180	70	2534	0.75	72.063	39.73	2.52	0.500
181	25	1240	0.8	104.565	7.84	12.75	0.061
182	30	1414	0.8	100.928	10.56	9.47	0.100
183	35	1606	0.8	90.192	16.52	6.05	0.167
184	40	1820	0.8	81.739	27.39	3.65	0.274
185	45	2056	0.8	80.602	35.32	2.83	0.364
186	50	2317	0.8	82.940	40.21	2.49	0.439
187	30	1240	0.8	95.295	7.70	12.99	0.074
188	35	1414	0.8	93.267	10.45	9.57	0.115
189	40	1606	0.8	83.613	17.34	5.77	0.187
190	45	1820	0.8	77.421	28.58	3.50	0.295
191	50	2056	0.8	78.737	34.30	2.92	0.371
192	55	2317	0.8	80.832	40.45	2.47	0.451
193	35	1240	0.8	87.235	7.18	13.93	0.085
194	40	1414	0.8	86.263	10.20	9.80	0.128
195	45	1606	0.8	78.012	17.22	5.81	0.201
196	50	1820	0.8	73.732	27.80	3.60	0.308
197	55	2056	0.8	74.494	34.50	2.90	0.387
198	60	2317	0.8	77.460	39.69	2.52	0.461
199	40	1240	0.8	80.394	6.88	14.54	0.097
200	45	1414	0.8	78.420	9.81	10.19	0.140
201	50	1606	0.8	72.514	17.03	5.87	0.216
202	55	1820	0.8	69.671	28.18	3.55	0.324
203	60	2056	0.8	71.575	34.77	2.88	0.405
204	65	2317	0.8	73.635	41.98	2.38	0.494
205	45	1240	0.8	72.437	6.65	15.04	0.110
206	50	1414	0.8	71.351	10.19	9.81	0.158

207	55	1606	0.8	67.907	18.09	5.53	0.238
208	60	1820	0.8	66.604	28.23	3.54	0.341
209	65	2056	0.8	68.479	35.63	2.81	0.426
210	70	2317	0.8	71.197	39.07	2.56	0.488
211	25	1153	0.85	101.744	7.58	13.19	0.059
212	30	1313	0.85	98.794	10.44	9.58	0.099
213	35	1489	0.85	91.116	14.96	6.68	0.155
214	40	1684	0.85	81.745	25.99	3.85	0.260
215	45	1898	0.85	79.402	36.06	2.77	0.364
216	50	2134	0.85	82.079	40.43	2.47	0.434
217	30	1153	0.85	92.408	7.36	13.59	0.072
218	35	1313	0.85	91.170	9.90	10.10	0.110
219	40	1489	0.85	83.067	16.47	6.07	0.180
220	45	1684	0.85	76.771	27.55	3.63	0.287
221	50	1898	0.85	77.570	34.18	2.93	0.367
222	55	2134	0.85	79.120	42.56	2.35	0.461
223	35	1153	0.85	84.013	7.01	14.27	0.083
224	40	1313	0.85	84.268	9.54	10.49	0.122
225	45	1489	0.85	77.589	16.94	5.90	0.197
226	50	1684	0.85	72.660	27.85	3.59	0.304
227	55	1898	0.85	73.364	35.92	2.78	0.394
228	60	2134	0.85	76.010	41.95	2.38	0.474
229	40	1153	0.85	77.479	6.55	15.27	0.093
230	45	1313	0.85	77.521	9.25	10.81	0.134
231	50	1489	0.85	72.571	16.25	6.15	0.208
232	55	1684	0.85	69.343	27.54	3.63	0.318
233	60	1898	0.85	70.519	35.06	2.85	0.403
234	65	2134	0.85	72.891	41.06	2.44	0.483
235	45	1153	0.85	70.322	6.38	15.66	0.106
236	50	1313	0.85	69.667	9.71	10.29	0.152
237	55	1489	0.85	67.380	16.88	5.93	0.228
238	60	1684	0.85	66.344	27.12	3.69	0.330
239	65	1898	0.85	68.426	32.57	3.07	0.401
240	70	2134	0.85	70.107	39.68	2.52	0.487
241	25	1078	0.9	98.266	7.48	13.36	0.058

242	30	1225	0.9	97.825	9.73	10.28	0.093
243	35	1388	0.9	89.960	14.74	6.79	0.153
244	40	1566	0.9	80.725	25.86	3.87	0.256
245	45	1762	0.9	79.149	34.73	2.88	0.351
246	50	1976	0.9	80.322	42.53	2.35	0.442
247	30	1078	0.9	89.545	7.21	13.87	0.070
248	35	1225	0.9	90.370	9.49	10.54	0.106
249	40	1388	0.9	82.837	16.30	6.14	0.178
250	45	1566	0.9	76.808	25.52	3.92	0.270
251	50	1762	0.9	76.118	35.05	2.85	0.369
252	55	1976	0.9	77.720	42.70	2.34	0.459
253	35	1078	0.9	81.941	6.83	14.64	0.081
254	40	1225	0.9	83.189	9.60	10.41	0.121
255	45	1388	0.9	76.837	16.51	6.06	0.193
256	50	1566	0.9	72.300	26.28	3.81	0.290
257	55	1762	0.9	72.551	37.89	2.64	0.404
258	60	1976	0.9	74.955	42.76	2.34	0.473
259	40	1078	0.9	75.827	6.50	15.38	0.092
260	45	1225	0.9	76.366	8.95	11.18	0.130
261	50	1388	0.9	71.542	15.92	6.28	0.203
262	55	1566	0.9	69.066	26.12	3.83	0.304
263	60	1762	0.9	69.793	35.66	2.80	0.403
264	65	1976	0.9	72.375	41.13	2.43	0.477
265	45	1078	0.9	68.964	6.20	16.12	0.103
266	50	1225	0.9	68.837	9.24	10.82	0.146
267	55	1388	0.9	66.928	16.75	5.97	0.224
268	60	1566	0.9	65.922	25.99	3.85	0.319
269	65	1762	0.9	67.834	33.33	3.00	0.401
270	70	1976	0.9	69.791	41.26	2.42	0.494

Appendix 3: Dataset used in proxy model development in Chapter 5.

Realization #1 - Low-Case

Case #	Temperature (°C)	Pressure (kPa)	C3 mole fraction	SolOR (m^3/m^3)	RF (%)	1/RF	Normalized Enthalpy
1	25	2226	0.5	109.715	11.08	9.03	0.050
2	30	2585	0.5	94.269	17.75	5.63	0.094
3	35	3003	0.5	87.460	28.98	3.45	0.158
4	40	3499	0.5	84.497	37.10	2.70	0.234
5	45	4104	0.5	85.790	43.45	2.30	0.324
6	50	4890	0.5	91.016	49.20	2.03	0.425
7	30	2226	0.5	105.381	11.01	9.08	0.089
8	35	2585	0.5	90.986	18.17	5.50	0.144
9	40	3003	0.5	86.388	29.34	3.41	0.219
10	45	3499	0.5	83.468	37.56	2.66	0.306
11	50	4104	0.5	85.148	43.45	2.30	0.407
12	55	4890	0.5	90.576	49.82	2.01	0.516
13	35	2226	0.5	96.183	10.98	9.10	0.146
14	40	2585	0.5	85.595	18.91	5.29	0.232
15	45	3003	0.5	84.584	29.50	3.39	0.311
16	50	3499	0.5	82.146	38.31	2.61	0.401
17	55	4104	0.5	84.917	43.98	2.27	0.510
18	60	4890	0.5	89.892	50.58	1.98	0.626
19	40	2226	0.5	87.947	10.64	9.40	0.180
20	45	2585	0.5	80.610	18.85	5.30	0.282
21	50	3003	0.5	79.906	29.43	3.40	0.416
22	55	3499	0.5	79.519	38.92	2.57	0.578
23	60	4104	0.5	84.633	44.33	2.26	0.659
24	65	4890	0.5	89.148	51.40	1.95	0.764
25	45	2226	0.5	79.372	10.56	9.47	0.216
26	50	2585	0.5	73.574	19.18	5.21	0.329
27	55	3003	0.5	75.122	29.23	3.42	0.464
28	60	3499	0.5	75.806	37.86	2.64	0.637
29	65	4104	0.5	82.058	44.88	2.23	0.873

30	70	4890	0.5	88.146	52.37	1.91	0.973
31	25	1967	0.55	109.327	10.33	9.68	0.046
32	30	2272	0.55	97.971	15.48	6.46	0.087
33	35	2622	0.55	89.591	25.51	3.92	0.145
34	40	3027	0.55	84.410	34.53	2.90	0.219
35	45	3502	0.55	84.850	41.32	2.42	0.303
36	50	4076	0.55	89.387	46.81	2.14	0.403
37	30	1967	0.55	104.002	10.14	9.86	0.085
38	35	2272	0.55	94.926	15.26	6.55	0.137
39	40	2622	0.55	86.759	26.62	3.76	0.207
40	45	3027	0.55	82.984	35.09	2.85	0.290
41	50	3502	0.55	84.340	41.62	2.40	0.385
42	55	4076	0.55	88.316	47.38	2.11	0.496
43	35	1967	0.55	94.750	9.80	10.21	0.129
44	40	2272	0.55	86.209	16.23	6.16	0.204
45	45	2622	0.55	82.382	26.93	3.71	0.314
46	50	3027	0.55	81.079	36.04	2.77	0.399
47	55	3502	0.55	83.453	42.26	2.37	0.495
48	60	4076	0.55	87.869	47.62	2.10	0.612
49	40	1967	0.55	86.929	9.45	10.58	0.158
50	45	2272	0.55	79.013	16.11	6.21	0.246
51	50	2622	0.55	76.368	27.24	3.67	0.366
52	55	3027	0.55	76.904	36.20	2.76	0.504
53	60	3502	0.55	80.542	42.72	2.34	0.687
54	65	4076	0.55	86.737	48.84	2.05	0.794
55	45	1967	0.55	78.981	9.22	10.85	0.189
56	50	2272	0.55	71.752	16.65	6.01	0.288
57	55	2622	0.55	71.925	27.16	3.68	0.407
58	60	3027	0.55	72.983	35.30	2.83	0.553
59	65	3502	0.55	77.144	42.13	2.37	0.746
60	70	4076	0.55	83.751	49.13	2.04	1.000
61	25	1762	0.6	109.282	9.60	10.41	0.043
62	30	2027	0.6	100.302	13.64	7.33	0.082
63	35	2328	0.6	90.694	23.47	4.26	0.136
64	40	2672	0.6	84.370	32.81	3.05	0.206

65	45	3066	0.6	84.027	39.84	2.51	0.286
66	50	3527	0.6	88.002	44.99	2.22	0.381
67	30	1762	0.6	101.459	9.70	10.31	0.081
68	35	2027	0.6	95.234	13.73	7.29	0.132
69	40	2328	0.6	88.137	23.70	4.22	0.198
70	45	2672	0.6	82.408	33.51	2.98	0.279
71	50	3066	0.6	82.920	40.24	2.48	0.369
72	55	3527	0.6	87.215	45.40	2.20	0.475
73	35	1762	0.6	92.620	9.30	10.76	0.115
74	40	2027	0.6	86.929	14.06	7.11	0.183
75	45	2328	0.6	82.120	24.25	4.12	0.280
76	50	2672	0.6	78.750	34.47	2.90	0.400
77	55	3066	0.6	81.266	40.42	2.47	0.506
78	60	3527	0.6	86.522	46.16	2.17	0.605
79	40	1762	0.6	85.548	8.89	11.25	0.140
80	45	2027	0.6	79.909	14.24	7.02	0.219
81	50	2328	0.6	76.128	24.60	4.07	0.327
82	55	2672	0.6	75.181	34.04	2.94	0.449
83	60	3066	0.6	78.138	40.49	2.47	0.607
84	65	3527	0.6	83.870	47.00	2.13	0.812
85	45	1762	0.6	77.415	8.69	11.51	0.167
86	50	2027	0.6	72.652	15.11	6.62	0.256
87	55	2328	0.6	71.350	24.69	4.05	0.364
88	60	2672	0.6	70.957	33.38	3.00	0.491
89	65	3066	0.6	74.460	39.71	2.52	0.657
90	70	3527	0.6	80.889	46.38	2.16	0.870
91	25	1595	0.65	108.797	9.15	10.92	0.041
92	30	1830	0.65	101.455	12.61	7.93	0.078
93	35	2094	0.65	90.422	20.84	4.80	0.128
94	40	2392	0.65	85.043	31.27	3.20	0.195
95	45	2731	0.65	83.131	39.18	2.55	0.273
96	50	3117	0.65	86.607	44.32	2.26	0.364
97	30	1595	0.65	99.975	9.01	11.10	0.077
98	35	1830	0.65	94.714	12.74	7.85	0.128
99	40	2094	0.65	87.236	21.72	4.60	0.192

100	45	2392	0.65	82.633	31.64	3.16	0.271
101	50	2731	0.65	81.990	39.43	2.54	0.358
102	55	3117	0.65	85.697	44.59	2.24	0.460
103	35	1595	0.65	91.311	8.69	11.51	0.103
104	40	1830	0.65	86.950	12.83	7.80	0.165
105	45	2094	0.65	81.346	21.70	4.61	0.253
106	50	2392	0.65	77.442	32.15	3.11	0.364
107	55	2731	0.65	79.282	39.47	2.53	0.497
108	60	3117	0.65	84.621	44.95	2.22	0.645
109	40	1595	0.65	83.974	8.42	11.88	0.125
110	45	1830	0.65	79.893	12.57	7.95	0.199
111	50	2094	0.65	75.812	22.26	4.49	0.295
112	55	2392	0.65	73.598	31.71	3.15	0.406
113	60	2731	0.65	75.636	38.86	2.57	0.546
114	65	3117	0.65	82.052	45.09	2.22	0.725
115	45	1595	0.65	76.328	8.04	12.44	0.150
116	50	1830	0.65	72.938	13.68	7.31	0.232
117	55	2094	0.65	70.803	22.64	4.42	0.331
118	60	2392	0.65	70.067	31.59	3.17	0.442
119	65	2731	0.65	72.174	38.33	2.61	0.589
120	70	3117	0.65	78.581	44.01	2.27	0.775
121	25	1457	0.7	108.931	8.71	11.48	0.039
122	30	1667	0.7	102.292	11.83	8.45	0.075
123	35	1902	0.7	90.811	19.74	5.06	0.122
124	40	2166	0.7	85.841	30.01	3.33	0.186
125	45	2462	0.7	82.605	38.01	2.63	0.262
126	50	2795	0.7	84.667	43.50	2.30	0.350
127	30	1457	0.7	99.051	8.68	11.53	0.071
128	35	1667	0.7	94.224	12.06	8.29	0.122
129	40	1902	0.7	85.525	20.34	4.92	0.188
130	45	2166	0.7	82.107	30.53	3.28	0.269
131	50	2462	0.7	80.840	38.50	2.60	0.352
132	55	2795	0.7	84.354	43.66	2.29	0.451
133	35	1457	0.7	90.840	8.34	12.00	0.094
134	40	1667	0.7	87.312	11.51	8.69	0.152

135	45	1902	0.7	79.816	20.45	4.89	0.230
136	50	2166	0.7	76.724	31.10	3.22	0.335
137	55	2462	0.7	77.863	38.64	2.59	0.454
138	60	2795	0.7	82.844	44.04	2.27	0.608
139	40	1457	0.7	83.268	7.95	12.57	0.114
140	45	1667	0.7	80.067	12.11	8.25	0.180
141	50	1902	0.7	74.313	20.61	4.85	0.270
142	55	2166	0.7	73.309	30.42	3.29	0.372
143	60	2462	0.7	74.354	38.14	2.62	0.497
144	65	2795	0.7	79.562	43.46	2.30	0.657
145	45	1457	0.7	76.218	7.64	13.08	0.136
146	50	1667	0.7	72.628	12.31	8.13	0.213
147	55	1902	0.7	69.731	20.75	4.82	0.303
148	60	2166	0.7	69.972	30.05	3.33	0.405
149	65	2462	0.7	71.100	37.69	2.65	0.535
150	70	2795	0.7	76.554	42.91	2.33	0.700
151	25	1340	0.75	106.860	8.46	11.82	0.037
152	30	1530	0.75	101.986	11.43	8.75	0.072
153	35	1742	0.75	88.527	18.36	5.45	0.117
154	40	1978	0.75	86.268	28.68	3.49	0.179
155	45	2241	0.75	82.149	36.87	2.71	0.253
156	50	2534	0.75	84.064	43.19	2.32	0.338
157	30	1340	0.75	97.913	8.23	12.15	0.065
158	35	1530	0.75	94.575	11.32	8.84	0.113
159	40	1742	0.75	82.921	18.81	5.32	0.176
160	45	1978	0.75	80.919	30.13	3.32	0.266
161	50	2241	0.75	79.651	37.55	2.66	0.359
162	55	2534	0.75	82.969	43.14	2.32	0.449
163	35	1340	0.75	89.547	7.86	12.73	0.086
164	40	1530	0.75	87.069	10.84	9.22	0.139
165	45	1742	0.75	79.037	18.81	5.32	0.212
166	50	1978	0.75	77.230	29.55	3.38	0.309
167	55	2241	0.75	76.549	37.27	2.68	0.419
168	60	2534	0.75	80.973	43.35	2.31	0.558
169	40	1340	0.75	82.232	7.53	13.29	0.104

170	45	1530	0.75	80.001	10.76	9.29	0.166
171	50	1742	0.75	73.254	19.00	5.26	0.248
172	55	1978	0.75	72.919	29.26	3.42	0.344
173	60	2241	0.75	73.506	37.40	2.67	0.457
174	65	2534	0.75	78.199	42.67	2.34	0.602
175	45	1340	0.75	75.125	7.34	13.62	0.123
176	50	1530	0.75	72.617	11.24	8.90	0.197
177	55	1742	0.75	68.071	19.35	5.17	0.280
178	60	1978	0.75	69.192	29.31	3.41	0.374
179	65	2241	0.75	69.856	36.38	2.75	0.491
180	70	2534	0.75	74.995	41.96	2.38	0.640
181	25	1240	0.8	105.055	8.05	12.43	0.035
182	30	1414	0.8	101.733	10.98	9.11	0.069
183	35	1606	0.8	90.123	16.90	5.92	0.112
184	40	1820	0.8	86.061	28.03	3.57	0.173
185	45	2056	0.8	81.434	36.67	2.73	0.246
186	50	2317	0.8	83.898	43.07	2.32	0.329
187	30	1240	0.8	95.872	7.94	12.60	0.059
188	35	1414	0.8	94.062	10.68	9.36	0.106
189	40	1606	0.8	82.834	17.64	5.67	0.163
190	45	1820	0.8	81.409	28.59	3.50	0.247
191	50	2056	0.8	79.044	37.18	2.69	0.347
192	55	2317	0.8	82.086	42.97	2.33	0.460
193	35	1240	0.8	87.218	7.49	13.35	0.078
194	40	1414	0.8	86.822	10.38	9.64	0.129
195	45	1606	0.8	76.803	17.56	5.69	0.196
196	50	1820	0.8	76.612	28.35	3.53	0.288
197	55	2056	0.8	76.114	37.25	2.68	0.388
198	60	2317	0.8	79.888	43.07	2.32	0.516
199	40	1240	0.8	80.650	7.16	13.97	0.094
200	45	1414	0.8	79.463	10.10	9.91	0.154
201	50	1606	0.8	71.358	18.08	5.53	0.229
202	55	1820	0.8	72.450	28.26	3.54	0.320
203	60	2056	0.8	72.087	36.22	2.76	0.424
204	65	2317	0.8	76.966	42.23	2.37	0.556

205	45	1240	0.8	72.298	6.89	14.52	0.112
206	50	1414	0.8	71.998	10.69	9.36	0.182
207	55	1606	0.8	68.082	18.08	5.53	0.260
208	60	1820	0.8	69.052	28.17	3.55	0.349
209	65	2056	0.8	68.993	35.84	2.79	0.455
210	70	2317	0.8	73.962	41.61	2.40	0.591
211	25	1153	0.85	101.922	7.87	12.70	0.033
212	30	1313	0.85	100.433	10.29	9.71	0.066
213	35	1489	0.85	91.334	15.67	6.38	0.108
214	40	1684	0.85	85.443	26.85	3.72	0.166
215	45	1898	0.85	80.883	36.64	2.73	0.239
216	50	2134	0.85	83.486	43.49	2.30	0.322
217	30	1153	0.85	92.863	7.63	13.10	0.055
218	35	1313	0.85	91.896	10.44	9.58	0.097
219	40	1489	0.85	83.184	16.57	6.04	0.152
220	45	1684	0.85	81.384	27.72	3.61	0.230
221	50	1898	0.85	78.348	36.85	2.71	0.325
222	55	2134	0.85	82.089	43.60	2.29	0.439
223	35	1153	0.85	84.541	7.28	13.74	0.071
224	40	1313	0.85	85.124	10.05	9.95	0.119
225	45	1489	0.85	76.831	16.79	5.96	0.182
226	50	1684	0.85	75.409	28.11	3.56	0.268
227	55	1898	0.85	74.944	36.22	2.76	0.362
228	60	2134	0.85	78.701	42.84	2.33	0.481
229	40	1153	0.85	77.648	6.95	14.39	0.086
230	45	1313	0.85	78.620	9.61	10.41	0.142
231	50	1489	0.85	71.087	17.00	5.88	0.213
232	55	1684	0.85	72.576	27.37	3.65	0.299
233	60	1898	0.85	71.823	35.87	2.79	0.394
234	65	2134	0.85	76.297	42.83	2.33	0.516
235	45	1153	0.85	71.039	6.64	15.05	0.102
236	50	1313	0.85	70.853	10.03	9.97	0.168
237	55	1489	0.85	65.958	17.72	5.64	0.242
238	60	1684	0.85	68.575	27.41	3.65	0.325
239	65	1898	0.85	69.600	36.02	2.78	0.424

240	70	2134	0.85	73.491	41.64	2.40	0.549
241	25	1078	0.9	99.525	7.73	12.94	0.032
242	30	1225	0.9	98.219	10.05	9.95	0.063
243	35	1388	0.9	88.443	15.91	6.28	0.104
244	40	1566	0.9	85.321	25.88	3.86	0.161
245	45	1762	0.9	80.618	36.48	2.74	0.234
246	50	1976	0.9	82.419	44.05	2.27	0.316
247	30	1078	0.9	90.455	7.45	13.41	0.050
248	35	1225	0.9	90.378	9.81	10.19	0.090
249	40	1388	0.9	82.993	15.87	6.30	0.142
250	45	1566	0.9	81.206	26.59	3.76	0.215
251	50	1762	0.9	78.431	36.81	2.72	0.305
252	55	1976	0.9	80.517	44.24	2.26	0.410
253	35	1078	0.9	82.744	7.10	14.08	0.066
254	40	1225	0.9	83.644	9.38	10.66	0.110
255	45	1388	0.9	76.011	16.45	6.08	0.169
256	50	1566	0.9	77.032	26.51	3.77	0.252
257	55	1762	0.9	75.377	36.24	2.76	0.339
258	60	1976	0.9	77.667	43.63	2.29	0.449
259	40	1078	0.9	76.398	6.77	14.78	0.079
260	45	1225	0.9	75.884	9.09	11.00	0.131
261	50	1388	0.9	70.533	16.60	6.03	0.199
262	55	1566	0.9	72.933	26.53	3.77	0.281
263	60	1762	0.9	72.062	35.75	2.80	0.722
264	65	1976	0.9	75.281	43.18	2.32	0.943
265	45	1078	0.9	69.915	6.40	15.61	0.183
266	50	1225	0.9	69.227	9.42	10.62	0.304
267	55	1388	0.9	65.134	17.41	5.74	0.443
268	60	1566	0.9	68.122	27.16	3.68	0.597
269	65	1762	0.9	69.381	35.24	2.84	0.775
270	70	1976	0.9	73.294	42.70	2.34	1.000

Realization #2 - Mid-Case

Case #	Temperature (°C)	Pressure (kPa)	C_3 mole fraction	SolOR (m^3/m^3)	RF (%)	1/RF	Normalized Enthalpy
1	25	2226	0.5	113.488	11.15	8.97	0.050
2	30	2585	0.5	97.213	18.24	5.48	0.096
3	35	3003	0.5	83.671	31.04	3.22	0.159
4	40	3499	0.5	84.764	39.95	2.50	0.236
5	45	4104	0.5	86.523	46.96	2.13	0.328
6	50	4890	0.5	91.138	52.80	1.89	0.433
7	30	2226	0.5	108.590	11.13	8.98	0.090
8	35	2585	0.5	95.475	18.48	5.41	0.146
9	40	3003	0.5	82.506	31.44	3.18	0.221
10	45	3499	0.5	83.735	40.38	2.48	0.308
11	50	4104	0.5	85.871	47.34	2.11	0.411
12	55	4890	0.5	91.158	53.21	1.88	0.525
13	35	2226	0.5	99.065	11.00	9.09	0.147
14	40	2585	0.5	90.479	19.17	5.22	0.234
15	45	3003	0.5	80.757	31.88	3.14	0.312
16	50	3499	0.5	82.346	40.83	2.45	0.402
17	55	4104	0.5	85.295	47.80	2.09	0.514
18	60	4890	0.5	90.931	53.73	1.86	0.635
19	40	2226	0.5	90.322	10.60	9.43	0.182
20	45	2585	0.5	82.653	19.07	5.24	0.286
21	50	3003	0.5	76.998	32.07	3.12	0.415
22	55	3499	0.5	79.404	41.39	2.42	0.572
23	60	4104	0.5	84.410	48.30	2.07	0.661
24	65	4890	0.5	90.294	54.30	1.84	0.773
25	45	2226	0.5	81.258	10.60	9.44	0.218
26	50	2585	0.5	74.590	19.37	5.16	0.335
27	55	3003	0.5	72.965	31.84	3.14	0.534
28	60	3499	0.5	76.043	40.90	2.44	0.731
29	65	4104	0.5	81.422	48.56	2.06	1.000
30	70	4890	0.5	89.461	55.00	1.82	0.982
31	25	1967	0.55	112.812	10.34	9.67	0.047

32	30	2272	0.55	98.527	15.66	6.39	0.088
33	35	2622	0.55	89.610	27.22	3.67	0.147
34	40	3027	0.55	84.346	37.73	2.65	0.220
35	45	3502	0.55	84.882	44.86	2.23	0.305
36	50	4076	0.55	88.458	50.68	1.97	0.408
37	30	1967	0.55	106.819	10.27	9.74	0.085
38	35	2272	0.55	94.681	15.93	6.28	0.138
39	40	2622	0.55	86.312	27.75	3.60	0.208
40	45	3027	0.55	83.556	38.05	2.63	0.291
41	50	3502	0.55	84.167	45.22	2.21	0.387
42	55	4076	0.55	88.035	51.10	1.96	0.501
43	35	1967	0.55	97.608	10.03	9.97	0.129
44	40	2272	0.55	87.396	16.19	6.18	0.206
45	45	2622	0.55	80.648	28.63	3.49	0.313
46	50	3027	0.55	81.767	38.41	2.60	0.399
47	55	3502	0.55	83.095	45.68	2.19	0.497
48	60	4076	0.55	87.426	51.66	1.94	0.619
49	40	1967	0.55	89.562	9.61	10.40	0.159
50	45	2272	0.55	80.064	16.35	6.12	0.249
51	50	2622	0.55	75.624	28.63	3.49	0.367
52	55	3027	0.55	78.229	38.41	2.60	0.502
53	60	3502	0.55	80.893	46.06	2.17	0.682
54	65	4076	0.55	86.819	52.22	1.91	0.797
55	45	1967	0.55	81.517	9.36	10.69	0.191
56	50	2272	0.55	73.104	16.76	5.97	0.293
57	55	2622	0.55	71.417	28.67	3.49	0.409
58	60	3027	0.55	74.538	37.96	2.63	0.552
59	65	3502	0.55	77.658	45.49	2.20	0.744
60	70	4076	0.55	83.652	52.31	1.91	1.000
61	25	1762	0.6	111.864	9.74	10.27	0.044
62	30	2027	0.6	101.710	13.89	7.20	0.083
63	35	2328	0.6	93.121	24.08	4.15	0.138
64	40	2672	0.6	83.991	35.86	2.79	0.207
65	45	3066	0.6	83.585	43.23	2.31	0.288
66	50	3527	0.6	86.729	49.16	2.03	0.385

67	30	1762	0.6	104.620	9.62	10.39	0.082
68	35	2027	0.6	96.828	13.94	7.17	0.132
69	40	2328	0.6	89.029	24.56	4.07	0.200
70	45	2672	0.6	82.684	36.22	2.76	0.279
71	50	3066	0.6	82.645	43.57	2.30	0.370
72	55	3527	0.6	86.129	49.58	2.02	0.479
73	35	1762	0.6	95.633	9.26	10.80	0.115
74	40	2027	0.6	88.369	14.08	7.10	0.184
75	45	2328	0.6	82.853	25.20	3.97	0.283
76	50	2672	0.6	80.435	36.70	2.72	0.398
77	55	3066	0.6	80.969	43.84	2.28	0.503
78	60	3527	0.6	84.914	50.08	2.00	0.607
79	40	1762	0.6	87.644	8.86	11.28	0.141
80	45	2027	0.6	80.185	14.38	6.96	0.222
81	50	2328	0.6	78.451	25.31	3.95	0.331
82	55	2672	0.6	76.702	36.20	2.76	0.448
83	60	3066	0.6	77.745	43.85	2.28	0.603
84	65	3527	0.6	82.790	50.52	1.98	0.809
85	45	1762	0.6	79.807	8.65	11.56	0.169
86	50	2027	0.6	72.680	14.94	6.69	0.261
87	55	2328	0.6	74.552	25.48	3.92	0.369
88	60	2672	0.6	73.775	35.85	2.79	0.490
89	65	3066	0.6	74.893	43.41	2.30	0.654
90	70	3527	0.6	79.450	49.77	2.01	0.870
91	25	1595	0.65	111.526	9.25	10.81	0.042
92	30	1830	0.65	104.347	12.73	7.85	0.079
93	35	2094	0.65	92.262	21.58	4.63	0.129
94	40	2392	0.65	83.505	34.13	2.93	0.196
95	45	2731	0.65	82.844	41.92	2.39	0.274
96	50	3117	0.65	85.463	48.12	2.08	0.366
97	30	1595	0.65	103.159	9.13	10.95	0.078
98	35	1830	0.65	97.322	12.77	7.83	0.129
99	40	2094	0.65	90.895	21.89	4.57	0.193
100	45	2392	0.65	81.906	34.33	2.91	0.270
101	50	2731	0.65	81.612	42.27	2.37	0.358

102	55	3117	0.65	84.454	48.48	2.06	0.461
103	35	1595	0.65	94.661	8.71	11.49	0.104
104	40	1830	0.65	88.613	12.67	7.89	0.167
105	45	2094	0.65	87.159	22.26	4.49	0.256
106	50	2392	0.65	78.591	34.46	2.90	0.360
107	55	2731	0.65	79.105	42.65	2.34	0.491
108	60	3117	0.65	82.598	48.51	2.06	0.636
109	40	1595	0.65	87.380	8.33	12.01	0.127
110	45	1830	0.65	80.903	12.75	7.84	0.200
111	50	2094	0.65	79.896	22.55	4.44	0.300
112	55	2392	0.65	75.111	34.18	2.93	0.404
113	60	2731	0.65	76.232	42.19	2.37	0.543
114	65	3117	0.65	80.154	48.67	2.05	0.723
115	45	1595	0.65	79.515	8.11	12.33	0.152
116	50	1830	0.65	73.598	13.44	7.44	0.237
117	55	2094	0.65	73.004	23.06	4.34	0.336
118	60	2392	0.65	71.998	34.03	2.94	0.442
119	65	2731	0.65	72.757	41.79	2.39	0.587
120	70	3117	0.65	76.757	48.08	2.08	0.774
121	25	1457	0.7	110.635	8.80	11.36	0.040
122	30	1667	0.7	104.367	11.97	8.35	0.075
123	35	1902	0.7	92.991	19.73	5.07	0.123
124	40	2166	0.7	82.552	32.36	3.09	0.186
125	45	2462	0.7	82.182	41.01	2.44	0.263
126	50	2795	0.7	84.384	47.36	2.11	0.351
127	30	1457	0.7	101.905	8.66	11.55	0.072
128	35	1667	0.7	97.314	11.99	8.34	0.123
129	40	1902	0.7	87.089	20.30	4.93	0.189
130	45	2166	0.7	80.089	32.72	3.06	0.267
131	50	2462	0.7	80.571	41.35	2.42	0.351
132	55	2795	0.7	83.422	47.61	2.10	0.450
133	35	1457	0.7	93.852	8.26	12.11	0.095
134	40	1667	0.7	89.868	11.67	8.57	0.153
135	45	1902	0.7	81.037	20.47	4.88	0.234
136	50	2166	0.7	76.991	33.01	3.03	0.333

137	55	2462	0.7	77.703	41.39	2.42	0.450
138	60	2795	0.7	81.007	47.94	2.09	0.602
139	40	1457	0.7	87.008	7.86	12.73	0.116
140	45	1667	0.7	82.080	11.65	8.59	0.183
141	50	1902	0.7	76.671	20.63	4.85	0.274
142	55	2166	0.7	73.815	32.69	3.06	0.370
143	60	2462	0.7	74.641	40.97	2.44	0.495
144	65	2795	0.7	77.891	47.38	2.11	0.654
145	45	1457	0.7	79.534	7.56	13.23	0.138
146	50	1667	0.7	73.802	12.19	8.20	0.217
147	55	1902	0.7	75.350	20.90	4.79	0.309
148	60	2166	0.7	70.775	32.45	3.08	0.403
149	65	2462	0.7	71.959	40.54	2.47	0.533
150	70	2795	0.7	74.975	46.88	2.13	0.698
151	25	1340	0.75	110.717	8.38	11.93	0.038
152	30	1530	0.75	105.156	11.36	8.80	0.073
153	35	1742	0.75	93.103	18.25	5.48	0.118
154	40	1978	0.75	81.856	30.98	3.23	0.179
155	45	2241	0.75	81.394	40.46	2.47	0.253
156	50	2534	0.75	83.606	46.85	2.13	0.339
157	30	1340	0.75	101.491	8.17	12.24	0.066
158	35	1530	0.75	96.829	11.27	8.87	0.114
159	40	1742	0.75	88.826	18.70	5.35	0.177
160	45	1978	0.75	79.429	31.75	3.15	0.264
161	50	2241	0.75	79.358	40.53	2.47	0.352
162	55	2534	0.75	82.066	47.11	2.12	0.447
163	35	1340	0.75	93.323	7.80	12.83	0.087
164	40	1530	0.75	89.231	10.91	9.17	0.140
165	45	1742	0.75	81.811	18.75	5.33	0.214
166	50	1978	0.75	75.750	31.62	3.16	0.309
167	55	2241	0.75	76.253	40.56	2.47	0.415
168	60	2534	0.75	79.475	47.10	2.12	0.553
169	40	1340	0.75	86.222	7.42	13.48	0.105
170	45	1530	0.75	81.880	10.80	9.26	0.168
171	50	1742	0.75	74.768	19.13	5.23	0.252

172	55	1978	0.75	72.462	31.43	3.18	0.343
173	60	2241	0.75	73.557	40.14	2.49	0.455
174	65	2534	0.75	76.919	46.61	2.15	0.599
175	45	1340	0.75	79.004	7.12	14.05	0.126
176	50	1530	0.75	74.698	11.15	8.97	0.200
177	55	1742	0.75	69.469	19.63	5.09	0.285
178	60	1978	0.75	69.577	31.37	3.19	0.374
179	65	2241	0.75	71.339	39.71	2.52	0.489
180	70	2534	0.75	73.686	46.03	2.17	0.639
181	25	1240	0.8	108.634	8.07	12.39	0.036
182	30	1414	0.8	104.406	10.91	9.17	0.070
183	35	1606	0.8	91.684	17.17	5.82	0.113
184	40	1820	0.8	81.533	29.72	3.36	0.173
185	45	2056	0.8	80.489	40.17	2.49	0.245
186	50	2317	0.8	82.564	46.61	2.15	0.328
187	30	1240	0.8	99.731	7.84	12.75	0.060
188	35	1414	0.8	96.387	10.70	9.34	0.106
189	40	1606	0.8	85.794	17.53	5.70	0.165
190	45	1820	0.8	78.673	30.36	3.29	0.247
191	50	2056	0.8	77.956	40.58	2.46	0.343
192	55	2317	0.8	81.004	46.86	2.13	0.452
193	35	1240	0.8	91.533	7.47	13.39	0.079
194	40	1414	0.8	89.081	10.34	9.67	0.130
195	45	1606	0.8	80.990	17.53	5.70	0.199
196	50	1820	0.8	75.077	30.36	3.29	0.288
197	55	2056	0.8	75.360	40.11	2.49	0.385
198	60	2317	0.8	78.271	46.70	2.14	0.512
199	40	1240	0.8	84.134	7.14	14.00	0.096
200	45	1414	0.8	81.358	10.11	9.89	0.155
201	50	1606	0.8	75.939	17.87	5.60	0.235
202	55	1820	0.8	71.982	30.37	3.29	0.320
203	60	2056	0.8	72.938	39.68	2.52	0.421
204	65	2317	0.8	75.281	46.22	2.16	0.553
205	45	1240	0.8	76.300	6.89	14.52	0.114
206	50	1414	0.8	73.796	10.41	9.61	0.185

207	55	1606	0.8	69.509	18.46	5.42	0.266
208	60	1820	0.8	69.164	30.36	3.29	0.348
209	65	2056	0.8	70.573	39.29	2.54	0.452
210	70	2317	0.8	72.446	45.60	2.19	0.589
211	25	1153	0.85	103.554	7.79	12.84	0.034
212	30	1313	0.85	103.099	10.53	9.49	0.067
213	35	1489	0.85	91.205	16.03	6.24	0.109
214	40	1684	0.85	82.090	28.52	3.51	0.167
215	45	1898	0.85	79.545	39.99	2.50	0.238
216	50	2134	0.85	81.634	46.93	2.13	0.320
217	30	1153	0.85	94.854	7.54	13.26	0.055
218	35	1313	0.85	95.737	10.21	9.80	0.098
219	40	1489	0.85	84.720	16.37	6.11	0.153
220	45	1684	0.85	78.027	29.26	3.42	0.230
221	50	1898	0.85	76.476	40.44	2.47	0.321
222	55	2134	0.85	80.011	47.22	2.12	0.433
223	35	1153	0.85	86.620	7.14	14.00	0.072
224	40	1313	0.85	88.727	9.88	10.12	0.120
225	45	1489	0.85	78.770	16.44	6.08	0.185
226	50	1684	0.85	74.828	29.10	3.44	0.268
227	55	1898	0.85	74.327	39.93	2.50	0.359
228	60	2134	0.85	77.265	46.63	2.14	0.477
229	40	1153	0.85	79.027	6.85	14.60	0.087
230	45	1313	0.85	81.010	9.68	10.34	0.143
231	50	1489	0.85	73.045	16.77	5.96	0.218
232	55	1684	0.85	71.842	29.40	3.40	0.300
233	60	1898	0.85	72.046	39.54	2.53	0.391
234	65	2134	0.85	74.314	46.11	2.17	0.514
235	45	1153	0.85	71.429	6.62	15.11	0.104
236	50	1313	0.85	72.317	9.87	10.14	0.171
237	55	1489	0.85	69.809	17.24	5.80	0.248
238	60	1684	0.85	69.116	29.42	3.40	0.326
239	65	1898	0.85	69.799	39.23	2.55	0.420
240	70	2134	0.85	71.958	45.67	2.19	0.546
241	25	1078	0.9	100.628	7.58	13.19	0.032

242	30	1225	0.9	100.236	10.24	9.77	0.064
243	35	1388	0.9	90.360	15.31	6.53	0.105
244	40	1566	0.9	83.427	27.57	3.63	0.162
245	45	1762	0.9	78.229	40.41	2.47	0.232
246	50	1976	0.9	80.670	47.71	2.10	0.313
247	30	1078	0.9	91.614	7.29	13.72	0.050
248	35	1225	0.9	92.834	9.92	10.08	0.091
249	40	1388	0.9	83.910	15.67	6.38	0.143
250	45	1566	0.9	78.103	28.22	3.54	0.216
251	50	1762	0.9	75.724	40.66	2.46	0.302
252	55	1976	0.9	78.613	48.06	2.08	0.406
253	35	1078	0.9	83.601	6.94	14.42	0.066
254	40	1225	0.9	85.887	9.51	10.51	0.111
255	45	1388	0.9	77.545	15.73	6.36	0.172
256	50	1566	0.9	74.693	28.35	3.53	0.252
257	55	1762	0.9	73.297	40.33	2.48	0.336
258	60	1976	0.9	76.116	47.33	2.11	0.445
259	40	1078	0.9	76.773	6.64	15.06	0.080
260	45	1225	0.9	78.608	9.21	10.86	0.132
261	50	1388	0.9	72.050	16.05	6.23	0.202
262	55	1566	0.9	71.893	28.53	3.51	0.282
263	60	1762	0.9	71.068	40.03	2.50	0.366
264	65	1976	0.9	73.719	46.72	2.14	0.478
265	45	1078	0.9	70.014	6.38	15.69	0.095
266	50	1225	0.9	71.196	9.26	10.80	0.159
267	55	1388	0.9	67.795	16.67	6.00	0.456
268	60	1566	0.9	69.165	28.91	3.46	0.603
269	65	1762	0.9	69.189	39.57	2.53	0.772
270	70	1976	0.9	71.262	46.36	2.16	1.000

Realization #3 - High-Case

Case #	Temperature (°C)	Pressure (kPa)	C_3 mole fraction	SolOR (m^3/m^3)	RF (%)	1/RF	Normalized Enthalpy
1	25	2226	0.5	105.782	12.01	8.33	0.054
2	30	2585	0.5	86.588	19.80	5.05	0.098
3	35	3003	0.5	81.165	32.84	3.04	0.159
4	40	3499	0.5	79.149	42.70	2.34	0.233
5	45	4104	0.5	82.514	50.25	1.99	0.321
6	50	4890	0.5	86.772	57.02	1.75	0.419
7	30	2226	0.5	101.618	12.02	8.32	0.097
8	35	2585	0.5	83.564	20.37	4.91	0.151
9	40	3003	0.5	80.067	33.29	3.00	0.222
10	45	3499	0.5	78.438	43.31	2.31	0.305
11	50	4104	0.5	81.986	50.79	1.97	0.403
12	55	4890	0.5	86.671	57.69	1.73	0.510
13	35	2226	0.5	93.746	11.89	8.41	0.163
14	40	2585	0.5	80.712	21.37	4.68	0.249
15	45	3003	0.5	78.342	33.80	2.96	0.316
16	50	3499	0.5	77.063	43.86	2.28	0.402
17	55	4104	0.5	81.564	51.48	1.94	0.506
18	60	4890	0.5	86.502	58.41	1.71	0.619
19	40	2226	0.5	84.125	11.39	8.78	0.201
20	45	2585	0.5	74.695	21.46	4.66	0.297
21	50	3003	0.5	75.962	34.40	2.91	0.425
22	55	3499	0.5	74.921	44.88	2.23	0.583
23	60	4104	0.5	80.785	52.19	1.92	0.656
24	65	4890	0.5	86.528	59.00	1.70	0.756
25	45	2226	0.5	74.528	11.22	8.91	0.239
26	50	2585	0.5	68.701	22.06	4.53	0.344
27	55	3003	0.5	71.678	33.79	2.96	0.473
28	60	3499	0.5	71.659	44.25	2.26	0.641
29	65	4104	0.5	78.171	52.82	1.89	0.873
30	70	4890	0.5	86.360	59.92	1.67	0.969
31	25	1967	0.55	105.737	11.20	8.92	0.050

32	30	2272	0.55	97.068	16.41	6.10	0.093
33	35	2622	0.55	79.558	29.78	3.36	0.148
34	40	3027	0.55	78.686	40.08	2.50	0.219
35	45	3502	0.55	80.662	47.95	2.09	0.301
36	50	4076	0.55	84.188	54.64	1.83	0.397
37	30	1967	0.55	100.234	11.16	8.96	0.095
38	35	2272	0.55	93.968	16.69	5.99	0.147
39	40	2622	0.55	78.713	30.32	3.30	0.211
40	45	3027	0.55	77.536	40.67	2.46	0.291
41	50	3502	0.55	80.324	48.37	2.07	0.383
42	55	4076	0.55	83.838	55.49	1.80	0.490
43	35	1967	0.55	91.130	10.90	9.18	0.145
44	40	2272	0.55	87.533	17.32	5.77	0.224
45	45	2622	0.55	76.381	31.25	3.20	0.324
46	50	3027	0.55	76.149	41.18	2.43	0.402
47	55	3502	0.55	79.617	49.02	2.04	0.495
48	60	4076	0.55	83.487	56.14	1.78	0.607
49	40	1967	0.55	82.154	10.39	9.62	0.177
50	45	2272	0.55	78.414	17.78	5.63	0.265
51	50	2622	0.55	71.437	30.97	3.23	0.375
52	55	3027	0.55	72.657	41.96	2.38	0.510
53	60	3502	0.55	77.544	50.10	2.00	0.691
54	65	4076	0.55	83.488	56.97	1.76	0.790
55	45	1967	0.55	73.760	10.13	9.87	0.211
56	50	2272	0.55	69.597	18.52	5.40	0.307
57	55	2622	0.55	67.171	30.87	3.24	0.417
58	60	3027	0.55	69.186	41.05	2.44	0.558
59	65	3502	0.55	73.954	49.24	2.03	0.749
60	70	4076	0.55	80.391	57.00	1.75	1.000
61	25	1762	0.6	102.995	10.67	9.37	0.048
62	30	2027	0.6	99.921	14.47	6.91	0.088
63	35	2328	0.6	80.571	26.81	3.73	0.139
64	40	2672	0.6	80.230	37.79	2.65	0.207
65	45	3066	0.6	78.605	46.32	2.16	0.285
66	50	3527	0.6	83.142	53.02	1.89	0.378

67	30	1762	0.6	95.500	10.60	9.43	0.094
68	35	2027	0.6	95.917	14.53	6.88	0.145
69	40	2328	0.6	78.064	27.45	3.64	0.203
70	45	2672	0.6	78.084	38.38	2.61	0.281
71	50	3066	0.6	78.502	46.83	2.14	0.368
72	55	3527	0.6	82.434	53.90	1.86	0.472
73	35	1762	0.6	87.394	10.33	9.68	0.130
74	40	2027	0.6	88.856	14.71	6.80	0.203
75	45	2328	0.6	73.258	28.30	3.53	0.291
76	50	2672	0.6	74.626	39.46	2.53	0.408
77	55	3066	0.6	77.034	46.82	2.14	0.511
78	60	3527	0.6	81.878	54.26	1.84	0.602
79	40	1762	0.6	77.863	9.73	10.27	0.158
80	45	2027	0.6	79.792	14.85	6.73	0.241
81	50	2328	0.6	68.563	28.46	3.51	0.337
82	55	2672	0.6	71.160	38.93	2.57	0.456
83	60	3066	0.6	74.055	47.29	2.11	0.611
84	65	3527	0.6	79.519	54.99	1.82	0.814
85	45	1762	0.6	69.934	9.39	10.65	0.189
86	50	2027	0.6	73.646	15.61	6.40	0.280
87	55	2328	0.6	65.293	28.55	3.50	0.374
88	60	2672	0.6	68.296	38.52	2.60	0.498
89	65	3066	0.6	70.496	46.71	2.14	0.660
90	70	3527	0.6	75.856	54.20	1.84	0.871
91	25	1595	0.65	101.999	10.15	9.86	0.046
92	30	1830	0.65	99.950	13.34	7.50	0.085
93	35	2094	0.65	81.269	24.23	4.13	0.133
94	40	2392	0.65	81.164	35.85	2.79	0.198
95	45	2731	0.65	77.681	45.31	2.21	0.273
96	50	3117	0.65	81.684	52.04	1.92	0.362
97	30	1595	0.65	93.813	10.10	9.90	0.089
98	35	1830	0.65	93.310	13.44	7.44	0.144
99	40	2094	0.65	78.627	24.60	4.07	0.202
100	45	2392	0.65	79.412	36.42	2.75	0.275
101	50	2731	0.65	77.017	45.47	2.20	0.359

102	55	3117	0.65	81.303	52.29	1.91	0.458
103	35	1595	0.65	84.430	9.66	10.35	0.118
104	40	1830	0.65	84.461	13.20	7.57	0.185
105	45	2094	0.65	73.798	25.50	3.92	0.265
106	50	2392	0.65	75.082	37.13	2.69	0.372
107	55	2731	0.65	74.675	46.43	2.15	0.502
108	60	3117	0.65	79.812	52.61	1.90	0.653
109	40	1595	0.65	76.210	9.24	10.83	0.143
110	45	1830	0.65	75.093	13.22	7.56	0.220
111	50	2094	0.65	68.755	25.85	3.87	0.307
112	55	2392	0.65	71.666	36.71	2.72	0.414
113	60	2731	0.65	71.418	45.62	2.19	0.551
114	65	3117	0.65	76.514	52.83	1.89	0.728
115	45	1595	0.65	68.451	8.86	11.29	0.170
116	50	1830	0.65	72.445	13.71	7.29	0.256
117	55	2094	0.65	64.319	26.28	3.81	0.341
118	60	2392	0.65	67.935	36.67	2.73	0.451
119	65	2731	0.65	68.370	45.11	2.22	0.593
120	70	3117	0.65	72.880	52.12	1.92	0.776
121	25	1457	0.7	101.593	9.83	10.18	0.044
122	30	1667	0.7	100.078	12.52	7.99	0.081
123	35	1902	0.7	82.005	21.88	4.57	0.128
124	40	2166	0.7	81.010	34.41	2.91	0.190
125	45	2462	0.7	77.223	44.00	2.27	0.263
126	50	2795	0.7	80.824	50.87	1.97	0.349
127	30	1457	0.7	92.671	9.70	10.31	0.082
128	35	1667	0.7	92.563	12.61	7.93	0.138
129	40	1902	0.7	77.951	22.80	4.39	0.201
130	45	2166	0.7	79.241	34.52	2.90	0.277
131	50	2462	0.7	75.428	44.35	2.25	0.354
132	55	2795	0.7	80.311	51.41	1.94	0.450
133	35	1457	0.7	83.216	9.25	10.81	0.108
134	40	1667	0.7	84.669	12.14	8.24	0.169
135	45	1902	0.7	73.806	23.13	4.32	0.244
136	50	2166	0.7	75.282	35.41	2.82	0.343

137	55	2462	0.7	72.863	44.95	2.22	0.459
138	60	2795	0.7	78.226	52.12	1.92	0.611
139	40	1457	0.7	75.654	8.84	11.31	0.131
140	45	1667	0.7	75.541	11.99	8.34	0.202
141	50	1902	0.7	69.589	23.53	4.25	0.283
142	55	2166	0.7	72.009	34.84	2.87	0.381
143	60	2462	0.7	69.565	44.58	2.24	0.502
144	65	2795	0.7	74.249	51.62	1.94	0.661
145	45	1457	0.7	67.759	8.51	11.75	0.156
146	50	1667	0.7	67.164	12.45	8.03	0.236
147	55	1902	0.7	65.364	24.26	4.12	0.315
148	60	2166	0.7	68.748	34.47	2.90	0.415
149	65	2462	0.7	67.103	43.62	2.29	0.539
150	70	2795	0.7	71.101	50.72	1.97	0.703
151	25	1340	0.75	100.088	9.42	10.62	0.043
152	30	1530	0.75	98.401	11.93	8.38	0.079
153	35	1742	0.75	83.954	19.72	5.07	0.124
154	40	1978	0.75	79.796	33.19	3.01	0.183
155	45	2241	0.75	76.339	43.59	2.29	0.255
156	50	2534	0.75	79.906	50.37	1.99	0.338
157	30	1340	0.75	90.539	9.28	10.78	0.076
158	35	1530	0.75	92.500	11.87	8.43	0.127
159	40	1742	0.75	78.369	20.68	4.84	0.191
160	45	1978	0.75	78.284	34.27	2.92	0.276
161	50	2241	0.75	74.333	43.60	2.29	0.369
162	55	2534	0.75	78.643	50.73	1.97	0.450
163	35	1340	0.75	81.551	8.84	11.31	0.099
164	40	1530	0.75	83.908	11.42	8.76	0.156
165	45	1742	0.75	74.471	21.06	4.75	0.227
166	50	1978	0.75	74.623	33.85	2.95	0.319
167	55	2241	0.75	71.327	43.72	2.29	0.424
168	60	2534	0.75	76.295	50.92	1.96	0.562
169	40	1340	0.75	74.013	8.45	11.83	0.120
170	45	1530	0.75	75.437	11.09	9.02	0.186
171	50	1742	0.75	69.432	21.58	4.63	0.263

172	55	1978	0.75	71.666	33.55	2.98	0.353
173	60	2241	0.75	68.167	43.19	2.32	0.463
174	65	2534	0.75	72.625	50.62	1.98	0.606
175	45	1340	0.75	66.226	8.13	12.30	0.143
176	50	1530	0.75	67.307	11.39	8.78	0.218
177	55	1742	0.75	65.787	22.51	4.44	0.293
178	60	1978	0.75	68.335	33.14	3.02	0.384
179	65	2241	0.75	65.840	42.60	2.35	0.497
180	70	2534	0.75	69.827	49.65	2.01	0.644
181	25	1240	0.8	96.753	9.17	10.91	0.041
182	30	1414	0.8	97.637	11.56	8.65	0.077
183	35	1606	0.8	93.113	17.40	5.75	0.122
184	40	1820	0.8	78.011	32.16	3.11	0.177
185	45	2056	0.8	75.797	42.75	2.34	0.248
186	50	2317	0.8	79.000	50.50	1.98	0.330
187	30	1240	0.8	87.748	8.90	11.24	0.070
188	35	1414	0.8	89.969	11.43	8.75	0.119
189	40	1606	0.8	81.633	18.57	5.39	0.179
190	45	1820	0.8	76.050	32.76	3.05	0.257
191	50	2056	0.8	73.624	43.78	2.28	0.354
192	55	2317	0.8	77.240	50.98	1.96	0.464
193	35	1240	0.8	78.967	8.58	11.65	0.091
194	40	1414	0.8	81.031	10.97	9.12	0.146
195	45	1606	0.8	75.497	18.92	5.28	0.213
196	50	1820	0.8	72.935	32.73	3.06	0.297
197	55	2056	0.8	69.984	43.32	2.31	0.395
198	60	2317	0.8	74.455	50.31	1.99	0.521
199	40	1240	0.8	71.352	8.20	12.19	0.110
200	45	1414	0.8	73.936	10.56	9.47	0.174
201	50	1606	0.8	69.802	19.72	5.07	0.246
202	55	1820	0.8	69.750	32.47	3.08	0.329
203	60	2056	0.8	67.256	42.48	2.35	0.430
204	65	2317	0.8	71.306	50.09	2.00	0.561
205	45	1240	0.8	63.836	7.91	12.63	0.132
206	50	1414	0.8	66.489	10.77	9.29	0.204

207	55	1606	0.8	65.861	20.90	4.79	0.274
208	60	1820	0.8	66.899	32.29	3.10	0.358
209	65	2056	0.8	66.081	41.92	2.39	0.461
210	70	2317	0.8	68.167	49.33	2.03	0.595
211	25	1153	0.85	93.516	8.89	11.25	0.040
212	30	1313	0.85	95.616	11.15	8.97	0.075
213	35	1489	0.85	95.198	16.11	6.21	0.119
214	40	1684	0.85	77.088	30.74	3.25	0.172
215	45	1898	0.85	74.616	42.96	2.33	0.242
216	50	2134	0.85	78.198	50.58	1.98	0.323
217	30	1153	0.85	84.963	8.63	11.58	0.065
218	35	1313	0.85	88.616	10.97	9.12	0.111
219	40	1489	0.85	87.384	16.75	5.97	0.169
220	45	1684	0.85	74.773	31.69	3.16	0.240
221	50	1898	0.85	71.787	43.69	2.29	0.333
222	55	2134	0.85	76.315	51.37	1.95	0.445
223	35	1153	0.85	76.179	8.23	12.15	0.084
224	40	1313	0.85	80.848	10.52	9.51	0.135
225	45	1489	0.85	80.381	16.91	5.91	0.201
226	50	1684	0.85	71.510	31.67	3.16	0.278
227	55	1898	0.85	68.812	42.79	2.34	0.369
228	60	2134	0.85	73.189	50.47	1.98	0.486
229	40	1153	0.85	68.678	7.86	12.72	0.102
230	45	1313	0.85	72.967	10.15	9.85	0.161
231	50	1489	0.85	71.957	17.92	5.58	0.232
232	55	1684	0.85	68.842	31.64	3.16	0.309
233	60	1898	0.85	66.331	42.29	2.36	0.402
234	65	2134	0.85	70.621	49.59	2.02	0.522
235	45	1153	0.85	61.671	7.54	13.26	0.121
236	50	1313	0.85	65.230	10.18	9.83	0.190
237	55	1489	0.85	65.138	19.26	5.19	0.259
238	60	1684	0.85	66.422	31.74	3.15	0.335
239	65	1898	0.85	64.081	41.87	2.39	0.430
240	70	2134	0.85	67.700	49.41	2.02	0.554
241	25	1078	0.9	90.741	8.72	11.47	0.039

242	30	1225	0.9	91.655	10.78	9.28	0.073
243	35	1388	0.9	93.767	15.24	6.56	0.118
244	40	1566	0.9	76.530	29.82	3.35	0.169
245	45	1762	0.9	73.482	43.39	2.30	0.237
246	50	1976	0.9	77.784	51.07	1.96	0.318
247	30	1078	0.9	82.166	8.44	11.85	0.061
248	35	1225	0.9	84.557	10.55	9.48	0.104
249	40	1388	0.9	88.556	15.94	6.28	0.160
250	45	1566	0.9	73.608	30.77	3.25	0.225
251	50	1762	0.9	70.923	43.73	2.29	0.313
252	55	1976	0.9	75.804	51.71	1.93	0.418
253	35	1078	0.9	74.821	8.01	12.48	0.079
254	40	1225	0.9	76.591	10.12	9.88	0.126
255	45	1388	0.9	82.150	15.91	6.29	0.190
256	50	1566	0.9	70.850	30.74	3.25	0.261
257	55	1762	0.9	68.185	43.18	2.32	0.347
258	60	1976	0.9	72.139	50.96	1.96	0.456
259	40	1078	0.9	67.327	7.69	13.01	0.095
260	45	1225	0.9	69.175	9.80	10.21	0.150
261	50	1388	0.9	74.055	16.70	5.99	0.219
262	55	1566	0.9	68.244	30.89	3.24	0.290
263	60	1762	0.9	65.106	42.82	2.34	0.377
264	65	1976	0.9	69.696	50.50	1.98	0.489
265	45	1078	0.9	59.415	7.42	13.48	0.113
266	50	1225	0.9	61.649	9.78	10.23	0.178
267	55	1388	0.9	66.072	18.19	5.50	0.244
268	60	1566	0.9	65.859	31.22	3.20	0.315
269	65	1762	0.9	63.235	42.43	2.36	0.404
270	70	1976	0.9	67.135	49.57	2.02	0.519