

THE UNIVERSITY OF CALGARY

PARAMETER ESTIMATION IN RESERVOIR SIMULATION

by

THOMAS BENG SWEE TAN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMICAL AND PETROLEUM ENGINEERING

CALGARY, ALBERTA

AUGUST, 1991

© THOMAS BENG SWEE TAN



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Voire référence*

Our file *Notre référence*

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-315-99539-4

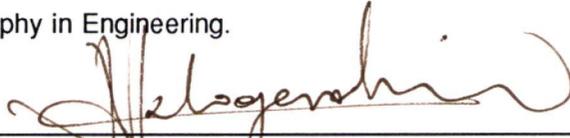
Canada

THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

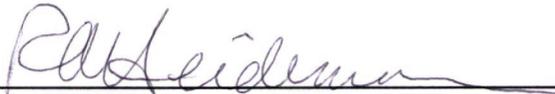
The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for Acceptance, a thesis entitled,

"PARAMETER ESTIMATION IN RESERVOIR SIMULATION"

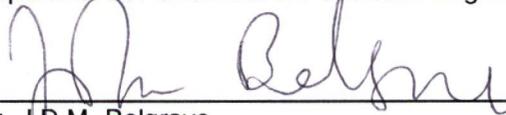
submitted by Thomas Beng Swee Tan in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering.



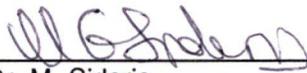
Dr. N.E. Kalogerakis (Supervisor)
Department of Chemical & Petroleum Engineering



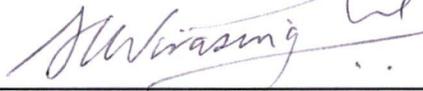
Dr. R.A. Heidemann
Department of Chemical & Petroleum Engineering



Dr. J.D.M. Belgrave
Department of Chemical & Petroleum Engineering



Dr. M. Sideris
Department of Surveying Engineering



Dr. S.C. Wirasinghe
Department of Civil Engineering



Dr. D.Y. Peng
University of Saskatchewan (External)

Date Sept. 13, 1991

ABSTRACT

This thesis addresses the problem of history matching hydrocarbon reservoirs using numerical simulation models. In current practice, estimated reservoir parameters such as porosity and permeability are modified by trial and error until an acceptable fit is obtained. There is currently no available industrial three-dimensional three-phase simulator that has automatic history matching capability.

A computationally efficient Gauss-Newton method for parameter estimation was developed in this work. The stability and convergence characteristics of the proposed method were verified by using complex chemical kinetic problems and a single-phase compressible flow model.

Subsequently, the proposed method was incorporated into a fully implicit multi-component three-dimensional three-phase simulator. The resulting model was tested on the second SPE Comparative Problem which is a numerically difficult three-phase coning problem. The model was able to simultaneously match observed pressures, water-oil ratios, gas-oil ratios and flowing bottom hole pressures. The parameters that were varied were the permeability and the porosity. A zonation approach was used. The proposed algorithm was able to recover the original values of permeability and porosity and the computational requirement was an order of magnitude less than previously published methods.

An additional important benefit of using an automatic history matching simulator is that it is possible to analyze the quality of the match obtained. As the simulator is based on the Gauss-Newton method, the variance-covariance matrix of the unknown reservoir parameters can be calculated using least squares theory. It is shown that by analyzing this information, as well as the eigenvectors of the normal equation matrix, it is possible to identify zones which are highly correlated. Illustrative examples were used to show how the theory could be applied to quantify the reliability of the estimated reservoir parameters and to identify the reservoir parameters that do not have a significant effect on the match of the measured field variables.

In actual application, a priori estimates of the most probable values of the parameters are usually available. Bayesian estimation was used to incorporate this prior information into the simulator. In addition, the penalty function approach was used to impose constraints on the parameter values. Test cases were used where the postulated models were quite different from the actual models. With prior estimates and constraints, reasonable parameter estimates were obtained. It was shown that important practical questions such as the presence of impermeable barriers in the reservoir, estimates of in-place fluid volumes and probable reservoir extensions could be better answered by the use of an automatic history matching simulator.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to Dr. Nicolas Kalogerakis for his guidance, advice and encouragement during the course of this work. Truly it can be said that if I have been able to see further, it is because I have stood on the shoulders of a giant.

My special thanks to Dr. Heidemann, Head of Department, to Mr. DenHartog, Chairman and Mr. Fisher, President of the D&S Group for their kind permission and support of my pursuit of a doctoral degree.

It was also a pleasure to work with the professors, graduate students and staff of the department while completing the requirements of the Ph.D program.

Finally, I would like to dedicate this work to my wife, Peggy, and my son, Anthony, in return for all the years that I have missed with them while I was engaged in my research work. Their love, support and understanding has made possible this thesis.

Dedicated to Peggy and Anthony

TABLE OF CONTENTS

	Page
APPROVAL PAGE	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
LIST OF SYMBOLS	xix
1. INTRODUCTION	1
1.1 Reservoir Simulation	1
1.2 Incentives for Reservoir Simulation	2
1.3 Steps in a Typical Simulation Study	4
1.4 History Matching	5
1.4.1 Current Approach	5
1.4.2 Motivation for Automatic History Matching	6
1.5 Problem Statement	7
1.6 Research Objectives	10
1.7 Outline of Thesis	12
2. REVIEW OF LITERATURE	15
2.1 Methods used for Automatic History Matching	16
2.1.1 Nonlinear Regression Methods	16
2.1.2 Optimal Control Methods	18
2.1.3 Regularization Methods	19

2.1.4	Other Methods	20
2.2	Identifiability of Parameters	21
2.3	Incorporation of Prior Information and Constraints	21
2.4	Summary of Review - The State of the Art	24
3.	MATHEMATICAL FORMULATION OF RESERVOIR SIMULATION MODELS	26
3.1	Classification of Models based on Phase Behaviour	26
3.2	Grid Systems	29
3.3	Differential Equations	30
3.4	Auxiliary Equations	33
3.5	Master Mole Fraction Concept	36
3.6	Boundary Conditions	37
3.7	Initial Conditions	37
3.8	Discretization	38
3.9	Well Model	39
3.10	Methods of Solution	40
3.11	Disadvantages and Advantages of Fully Implicit Solution	42
3.12	Jacobian Matrix	42
3.13	Solution of Jacobian	45
3.14	Pseudocode Flow Chart of the Simulator	46
4.	PARAMETER ESTIMATION TECHNIQUES	49
4.1	Development of Least Squares from Maximum Likelihood Estimation	51

4.2	Bayesian Estimation	54
4.3	Methods of Estimation	56
5.	PROPOSED COMPUTATIONALLY EFFICIENT GAUSS-NEWTON METHOD	59
5.1	Problem Statement	59
5.2	The Gauss-Newton Method	61
5.3	Integration Method - The Implicit (Euler) Method	63
5.4	Normal Implementation of the Gauss-Newton Method	65
5.5	Efficient Implementation of the Gauss-Newton Method	67
5.5.1	Matrix inversion	67
5.5.2	Sequential Integration of Sensitivity Equations	69
5.6	Convergence Testing using Chemical Kinetic Models	72
5.6.1	Example 1	73
5.6.2	Example 2	76
5.6.3	Example 3	79
5.6.4	Example 4	82
5.7	Discussion	84
6.	APPLICATION TO A SINGLE-PHASE MODEL	85
6.1	Single-Phase Compressible Flow Model	85
6.2	Performance Evaluation	86
6.2.1	Single-Phase Gas Reservoir	86
6.2.2	Five-Spot Pilot Example	93
6.3	Discussion	98

7.	APPLICATION TO A THREE-DIMENSIONAL THREE-PHASE MODEL	100
7.1	Model Equations	100
7.2	Gauss-Newton Method	102
7.3	Model Features and Operation	104
7.3.1	Cautious Step Size Policy	105
7.3.2	Pseudo-Inverse Option	107
7.3.3	Implicit Formulation	108
7.4	Application of the Automatic History Matching Model	108
7.4.1	Matching Reservoir Pressure	109
7.4.2	Matching Water-Oil Ratio	118
7.4.3	Matching Gas-Oil Ratio	121
7.4.4	Matching Bottom Hole Pressure	124
7.4.5	Matching Combinations of Observed Data	126
7.5	Correlation of Computational Requirements with Number of Parameters	130
8.	RELIABILITY OF PARAMETER ESTIMATES	133
8.1	The Covariance Matrix of the Parameter Estimates	134
8.2	Correlation between Pairs of Parameters	136
8.3	Multiple Correlation	137
8.4	Numerical Results and Discussion	140
8.4.1	Case Study I :	
	A 3 x 3 Symmetrical Model	142
8.4.2	Case Study II :	
	A 3 x 3 Asymmetrical Model	154
8.4.3	Case Study III : The SPE Second Comparative	

Solution Problem (Chappelear and Nolen,1986)	159
8.5 Practical Guidelines	166
9. INCORPORATION OF PRIOR KNOWLEDGE AND CONSTRAINTS	167
9.1 Incorporating Prior Information using Bayesian Methodology	168
9.2 Incorporating Constraints using the Penalty Function Approach	171
9.3 Model Application	175
9.4 Case Study IV	176
9.4.1 Match of Actual Model	182
9.4.2 Postulated Model A	185
9.4.3 Postulated Model B	199
9.5 Case Study V	212
9.5.1 Vertical Cross-section Model	212
9.5.2 Three-dimensional Model	219
9.6 Practical Application of the Automatic History Matching Model	223
10. CONCLUSIONS AND RECOMMENDATIONS	225
10.1 Conclusions	225
10.2 Recommendations	227
10.2.1 Increasing the Parameter Types	227
10.2.2 Improving the Cautious Step-size Policy	228
10.2.3 Improving the Postulated Model and Discrimination of Postulated Models	228

REFERENCES

229

APPENDIX A

243

LIST OF TABLES

Table	Title	Page
5.1	Pseudo-Code Flow Chart for Gauss-Newton Method	66
5.2	Number of Iterations for Example 1	75
5.3	Number of Iterations for Example 2	77
5.4	Number of Iterations for Example 3	81
5.5	Number of Iterations for Example 4	82
6.1	Single-Phase Gas Reservoir Problem Description	87
6.2	Estimates of Parameter Values at each Iteration for Single-Phase Gas Reservoir	91
6.3	Five-Spot Pilot Example Problem Description	94
6.4	Five-Spot Pilot Example - Parameter Convergence	96
6.5	Comparison of Equivalent Model Runs for Each Gauss-Newton Iteration	98
7.1	Match of Original Permeabilities	111
7.2	Match of Original Porosities	112
7.3	Comparison of Model Equivalent Runs with Number of Parameters	131
8.1	Correlation Matrix - Full Data - 3 x 3 Symmetrical Problem	144
8.2	Eigenvector Matrix - Full Data - 3 x 3 Symmetrical Problem	145
8.3	Standard Deviations - 3x3 Symmetrical Problem	149
8.4	Correlation matrix - Limited Data - 3 x 3 Symmetrical Problem	150

8.5	Eigenvector Matrix - Limited Data - 3 x 3 Symmetrical Problem	151
8.6	Standard Deviations and Eigenvector of Smallest Eigenvalue - 3x3 Asymmetrical Problem - Limited Data	156
8.7	Estimated Permeabilities, Standard Deviations and Eigenvector for SPE Problem 2 - Full Data	162
8.8	Estimated Permeabilities and Standard Deviations for SPE Problem 2 - Limited Data	163
9.1	Description of Actual Model for Case Study IV	179
9.2	Relative Permeability Table for Case Study IV	180
9.3	Parameter Match of Actual Model for Case Study IV	184
9.4	Case IV - Postulated Model A - Parameter Estimates	187
9.5	Case IV - Postulated Model B - Parameter Estimates	201
9.6	Description of Models Used for Case Study V	215
9.7	Case V - Vertical Cross-Section Model - Parameter Estimates	217
9.8	Case V - Vertical Cross-Section Model - Fluids In-place Estimates	218
9.9	Case V - Three Dimensional Model - Parameter Estimates	221
9.10	Case V - Three Dimensional Model - Fluids In-place Estimates	222

LIST OF FIGURES

Figure	Title	Page
3.1	Schematic of Model Jacobian for 1 Cell with Two Neighbours	44
5.1	Schematic of Jacobian of Model and Sensitivity Equations	68
5.2	Schematic of Sequential Solution of Model and Sensitivity Equations	71
5.3	Given Measurements for Example 1	74
5.4	Simulated Measurements for Example 2	78
5.5	Simulated Measurements for Example 3	80
5.6	Simulated Measurements for Example 4	83
6.1	Single-Phase Gas Reservoir Example	88
6.2	Estimates of Parameter Values at each Iteration in Single-Phase Reservoir Run	92
6.3	Five-Spot Pilot Example	95
6.4	Estimates of Parameter Values at each Iteration in Five-Spot Pilot Example	97
6.5	Equivalent Model Runs to Compute Sensitivity Coefficients	99
7.1	Reduction of Objective Function with Iteration in Run No. 1 (Match of Reservoir Pressures)	113
7.2	Match of Pressure Profile of Cell (1,1,8) in Run No. 1	114

7.3	Estimates of Parameter Values at each Iteration in Run No. 1	115
7.4	Match of Water-Oil Ratio in Run No. 6	120
7.5	Match of Gas-Oil Ratio in Run No. 9	122
7.6	Match of Bottom Hole Pressure in Run No. 11	125
7.7	Match of Pressure Data in Run No. 13 (Combined Data)	127
7.8	Match of WOR and GOR in Run No. 13 (Combined Data)	128
7.9	Correlation of Computational Requirements for Runs 1 to 13	132
8.1	Schematic of 3 x 3 Grid Cell Model	141
8.2	3 x 3 Symmetrical Problem with Limited Data Match of Pressures at Cell 1	152
8.3	3 x 3 Symmetrical Problem with Limited Data Match of Pressures at Cell 9	153
8.4	3 x 3 Asymmetrical Problem with Limited Data Match of Pressures at Cell 1	157
8.5	3 x 3 Asymmetrical Problem with Limited Data Match of Pressures at Cell 9	158
8.6	SPE Problem 2 Limited Data Match of Pressure Measurements	164
8.7	SPE Problem 2 Limited Data Match of WOR and GOR	165
9.1	Case Study IV - Actual and Postulated Models	177
9.2	Case Study IV Porosity and Permeability Distributions	178

9.3	Case Study IV Model A	
	Well #1 - Match of Pressure	189
9.4	Case Study IV Model A	
	Well #2 - Match of Water-Oil Ratio	190
9.5	Case Study IV Model A	
	Well #2 - Match of Gas-Oil Ratio	191
9.6	Case Study IV Model A	
	Well #2 - Match of Oil Rate	192
9.7	Case Study IV Model A	
	Well #2 - Match of Pressure	193
9.8	Case Study IV Model A	
	Well #3 - Match of Water-Oil Ratio	194
9.9	Case Study IV Model A	
	Well #3 - Match of Gas-Oil Ratio	195
9.10	Case Study IV Model A	
	Well #3 - Match of Oil Rate	196
9.11	Case Study IV Model A	
	Well #3 - Match of Pressure	197
9.12	Case Study IV Model B	
	Well #1 - Match of Pressure	203
9.13	Case Study IV Model B	
	Well #2 - Match of Water-Oil Ratio	204
9.14	Case Study IV Model B	
	Well #2 - Match of Gas-Oil Ratio	205
9.15	Case Study IV Model B	
	Well #2 - Match of Oil Rate	206

9.16 Case Study IV Model B	
Well #2 - Match of Pressure	207
9.17 Case Study IV Model B	
Well #3 - Match of Water-Oil Ratio	208
9.18 Case Study IV Model B	
Well #3 - Match of Gas-Oil Ratio	209
9.19 Case Study IV Model B	
Well #3 - Match of Oil Rate	210
9.20 Case Study IV Model B	
Well #3 - Match of Pressure	211
9.21 Case Study V - Schematic of Vertical Cross-Section	213

LIST OF SYMBOLS

b	constraints
C	observation matrix
D	depth of cell
f_I	fugacity of component I
G	sensitivity matrix
H	Gauss-Newton matrix
I	component index
k	parameter vector
k_{rp}	relative permeability of phase p
K	absolute permeability
K_{Ip}	equilibrium ratio of Component I in phase p
L	likelihood function
m	number of observations at each time
n_o	number of observation times
n_b	number of constraints
n_c	number of components
n_p	number of phases
p_o	prior distribution
p	number of estimated parameters
P_p	pressure of phase p in grid cell
P_c	capillary pressure
P_{wb}	flowing well bottom hole pressure
q_I	production/injection rate of component I
q_p	rate of flow of phase p

Q	weighting matrix
r_w	wellbore radius
R	correlation matrix
S	objective function
S_p	saturation of phase p
t	time
u	user specified vector
V	grid cell volume
V	covariance matrix of observation errors
V	eigenvector matrix of H
W	covariance matrix of prior distribution of parameters
WI	well index defined in Eq. 3.26
x	state vector
x_{Ip}	mole fraction of Component I in phase p
X_I	master mole fraction of Component I
\hat{y}	measured variable
y	calculated variable
z_{ij}	random numbers sampled from the standard normal distribution

Greek Letters

α	parameter indicating normalized distance of initial guess from optimum
α_1	specified constant for 1 th penalty function

γ_p	specific weight of phase p
Δ	difference operator
Λ	eigenvalue matrix
λ	step length
μ_p	viscosity of phase p
ξ	penalty function
ρ_p	density of phase p
σ_e	normalized standard deviation of measurement errors
T	physical transmissibility
ϕ	porosity
ψ	parameters of normal distribution function

Subscripts

B	Bayes
g	gas
i	observation time index
l	constraint index
LS	least squares
m	mean
max	maximum value
min	minimum value
ML	maximum likelihood
o	oil
p	phase (e.g. oil, gas, water)
x	direction corresponding to x axis

y direction corresponding to y axis
w water
z direction corresponding to z axis

Superscripts

j iteration level for Gauss-Newton iteration
l iteration level for model iteration
n last time step taken
n+1 next timestep
* value of the variable at the optimum
T transpose

CHAPTER 1

INTRODUCTION

1.1 Reservoir Simulation

The mathematical modelling of the dynamics of fluid flow, heat and mass transfer in underground hydrocarbon reservoirs is termed reservoir simulation. The degree to which the model can be used successfully to predict the behaviour of the actual reservoir is a function of the input data used and the adequacy of the mathematical model.

The mathematical model consists of the set of partial differential equations, together with appropriate boundary conditions, which describes the flow of oil, gas and water in the reservoir. The model equations must account for mass transfer between the phases, consider gravity, capillary and viscous forces, and allow for heterogeneous reservoirs of arbitrary geometry. The solution of these equations using classical analytical methods of mathematical physics is not practicable, and numerical methods implemented on high speed computers are necessary. The explosive growth in computational speed and memory capacity of computers has made possible the routine use of reservoir simulators and contributed in part to the rationale for the research work described in this thesis.

1.2 Incentives for Reservoir Simulation

The main incentive for reservoir simulation is the potential increased recovery of oil and gas from the reservoir by better reservoir management. A simulator that has accurately matched prior production history can be used to evaluate plans for optimum development of the reservoir. The major benefits of a simulation study are summarised below.

(a) The history matching process improves understanding of the reservoir behaviour. The reservoir engineer conducting the study increases his knowledge of the important variables affecting reservoir behaviour in the course of developing and modifying the grid cell model and its parameter distributions to match the prior history.

(b) It assists in reservoir definition, such as the size and extent of the reservoir. A properly matched model can point out deficiencies in reservoir description based on insufficient geological information.

(c) We can obtain more accurate estimates of fluid in-place and the extent of aquifer support. Very simplified models of the reservoir can be quickly built for material balance purposes.

(d) With a simulator, we can quickly compare different operating strategies without expensive pilot tests. The physical reservoir can be developed and produced only once. However the numerical model can be run many times in order to examine the various alternatives in production scenarios.

(e) The simulator can be used for prediction of reservoir performance under primary or secondary/tertiary recovery using various injection fluids.

(f) We can compare various flooding patterns and the effects of well spacing on sweep efficiency and ultimate recovery.

(g) In addition, we can compare different injection and production rates to find the optimum ones.

(h) Quite often the results of the simulation study identify unswept oil regions suitable for locating new wells.

(i) We can determine critical well rates to avoid coning and locate completion intervals. Single well models with greater definition near the well bore can be used to investigate gas and water coning.

(j) Laboratory experiments can be modelled to obtain fluid and rock properties that can be used in full field modelling.

In order to realise these benefits, reservoir description data of high quality, accurate production history data and pressure observations are necessary. The more accurate the data, the more confidence can be placed on the resulting match. The reliability of the predictions depends on the acceptability of the match. The current industry guideline is that the prediction period should only be as long as the historical period.

1.3 Steps in a Typical Simulation Study

A properly conducted reservoir study will take time to complete. Most studies involve essentially the same kind of activities but the time necessary to complete each phase of the study will differ from reservoir to reservoir. In general, however, the usual steps and typical percentage of time spent on each of these are:

(a) Definition of the problem (5%).

It is necessary that the scope of the study and the objectives to be achieved be well defined prior to beginning the study.

(b) Data Compilation and Review (30%).

Reservoir description data, seismic data, core analysis, well log analysis, pressure transient analysis, PVT laboratory tests, special fluid displacement tests, production history data, all these have to be analyzed and reviewed for missing data or inconsistencies.

(c) Model Building (10%).

The data have to be assembled in a form demanded by the input format of the simulator being used. Decisions have to be made on the type of models to be run, the number of individual wells to be analyzed, what cross-section and areal models should be run and the number of cells for the full field model.

(d) History Matching (40%).

Here the task is to duplicate the reservoir behaviour

with the simulation model by modifying the input data as necessary to obtain a match.

(e) Performance predictions and sensitivities (10%).

With a satisfactorily matched model, different operating strategies and sensitivity to well production / injection rates can be studied to maximize recovery or to achieve the objectives of the study.

(f) Reporting the results and conclusions (5%).

From the above, we can see that history matching is one of the most time-consuming phases of a typical simulation study.

1.4 History Matching

1.4.1 Current Approach

In a typical reservoir study, a grid cell model that could have thousands of grid cells is built to represent the underground geological structure. The reservoir parameters such as porosity, permeability and depth which are attributed to these grid cells have to be estimated from analysis of data from a few wells drilled into the structure. There is therefore considerable uncertainty in the estimated values of many of the reservoir parameters.

In order to verify the values of the estimated parameters, the postulated model is used to simulate past reservoir performance and the results are compared with

historical data. Very often, significant differences will be obtained. History matching is the process of estimation of reservoir parameters, such as the porosity and permeability distributions so that the observed field data is replicated by the numerical model. The current state of the art is the reservoir engineer performing the study will use his experience and judgement to vary the parameters by trial and error until a suitable match is obtained. The search for these parameters is often the most time consuming and frustrating part of a reservoir study.

1.4.2 Motivation for Automatic History Matching

Extensive experience is often required to perform a simulation study. The engineer should have a good knowledge of the important parameters that affect reservoir behaviour. As well, manpower costs have been increasing while advances in computer technology have led to a rapid decline in computer costs.

The aim therefore is to reduce the tedium of making trial and error runs and instead to let the computer do the work of estimation of the parameters using optimization techniques. This will allow the reservoir engineer to spend more time on the interpretation of the results, and less time on the process of making trial and error runs to estimate the magnitude of the parameters.

The resulting efficiencies will reduce the time required

to perform history matching and therefore the total length of time for the study.

1.5 Problem Statement

The available field data may consist of observed pressures, production rates, gas-oil ratios, water-oil ratios, and bottom hole pressures measured at some wells at a few times during the history of the reservoir.

Using the least squares (LS) technique, an objective function is defined as

$$S_{LS}(\mathbf{k}) = \sum_{i=1}^{n_o} [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)]^T \mathbf{Q}_i [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)] \quad (1.1)$$

where $\hat{\mathbf{y}}$ represents a vector of m measurements at observation times t_i , n_o is the number of observation times, \mathbf{y} represents the vector of corresponding calculated values, \mathbf{Q} is an $m \times m$ matrix of weighting factors which may vary with time and \mathbf{k} is a vector of p parameters to be estimated.

The weighting matrix \mathbf{Q} is often chosen as a diagonal matrix which normalizes the data, so that all measurements are of the same order of magnitude. On statistical grounds, this is the correct choice if the error in the measurements is proportional to the magnitude of the variable, that is, we have a constant percentage error.

The history matching problem consists of estimating the unknown reservoir parameters in the numerical model so that

the objective function, S , is minimized. The numerical model involves one or more partial differential equations depending on the scope of the model. These are reduced by grid discretization to a set of ordinary differential equations which can be written in general as

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{k}; \mathbf{u}) \quad (1.2)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (1.3)$$

where \mathbf{x} is the n dimensional vector of state variables, \mathbf{x}_0 is the given initial state, \mathbf{k} is the p dimensional vector of parameters to be estimated, \mathbf{u} is a vector of user specified variables or parameters.

For example, the vector \mathbf{x} , in a black oil model could be the pressure and saturation variables. The vector \mathbf{k} could be unknown permeabilities and porosities. The vector \mathbf{u} could be user specified production rates.

The objective is to find the unknown parameters, \mathbf{k} , by matching the model predictions to the actual field data taken over several observation times, t_i , $i = 1, 2, \dots, n_0$. These measurements, represented by the m dimensional vector, $\hat{\mathbf{y}}(t_i)$, could be the pressure at wells, water-oil ratios, gas-oil ratios and so on. The corresponding model predictions, $\mathbf{y}(t_i)$, are related to the model state variables through a relationship of the form,

$$y(t) = h(x(t)) \quad (1.4)$$

The solution of the model equations for a given set of reservoir parameters would take minutes to hours of computer time depending on the number of grid cells in the model and the computer used. It is therefore desirable to use a parameter estimation method that does not require a large number of function evaluations (model runs).

Standard methods for parameter estimation using nonlinear regression include steepest descent and the Gauss-Newton method. These require calculation of sensitivity coefficients. These are the partial derivatives of the reservoir state variables with respect to the reservoir parameters, that is,

$$\frac{\partial(\text{state variable})}{\partial(\text{estimated parameter})} \quad (1.5)$$

These are normally obtained by numerical differentiation, that is, each parameter is perturbed slightly and a full model run is made to find the change in the reservoir variables. If there are p parameters, we would require at least $(p+1)$ runs for each iteration of the regression method. For example, if there are 10 parameters, and we need 10 iterations for the regression method, and each run took 1 hour, we would need 110 runs or 110 cpu hours. The extensive computation time is the reason why no practical method for automatic history matching

is in use today. Mattax and Dalton (1990) writing in the prestigious Society of Petroleum Engineers Reservoir Simulation Monograph stated that automatic history matching is not widely used and probably should not be the method of choice for most problems.

1.6 Research Objectives

The main goal of this research project was to develop an efficient algorithm for automatic history matching, and to incorporate it into a three-dimensional three-phase reservoir simulator. This would result in the first practical model with automatic history matching capability. The development and research objectives that led to the realization of this goal follow.

In fundamental research work comparing various parameter estimation methods, a computationally efficient form of the quadratically convergent Gauss-Newton method was proposed. To test whether the modifications suggested would still result in a convergent method, the proposed method was used to estimate the reaction rate constants in several typical chemical kinetic problems described using ordinary differential equations. The results of this numerical experimentation demonstrated that the proposed method was convergent, and required significantly less computational time for these kinetic models.

The next objective was to incorporate the proposed method

in a single-phase reservoir model. The resulting model was tested with single-phase problems described in previously published papers on automatic history matching. Again encouraging results were obtained in estimating parameters with significant reduction in computational time.

Building on this success, the method was next incorporated in a three-dimensional three-phase reservoir simulator that had been developed by this author. Several conceptual problems had to be overcome before a working model could be put together. Development testing was performed using a difficult standard industrial problem. The model was able to match observations such as pressures, water-oil ratios, gas-oil ratios as well as combinations of these observations. The computational speed was also an order of magnitude faster than other published methods.

The questions of identifiability of parameters and accuracy of estimates were then addressed. The objective was to identify correlations between pairs of parameters as well as between groups of parameters. This information could then be used to identify the reservoir parameters that do not have a significant effect on the match of the observed variables.

In actual application of the automatic history matching model in a reservoir study, it is extremely unlikely that the grid cell model used for the reservoir would be a true representation of the underground reservoir. Thus it is very probable that the parameter estimation would result in values

for the physical parameters that are either unrealistically high or low. Furthermore, a priori estimates of the most probable values of the parameters are usually available. To complete the model, the final objective was to incorporate the prior information via Bayesian estimation, and to provide maximum and minimum constraints on the parameter values.

1.7 Outline of Thesis

Chapter 1 provides an introduction to reservoir simulation, its uses and the steps involved in a typical reservoir study. The problem of history matching is defined and the difficulties of automatic history matching outlined. The research objectives are then stated.

In Chapter 2, the literature on automatic history matching is reviewed. The various attempts at developing reservoir models using nonlinear regression methods for parameter estimation are surveyed. Optimal control methods have also been used in the literature. Papers on identifiability of parameters and incorporation of prior information and constraints are discussed. Finally, the state of the art is presented.

Chapter 3 introduces the mathematical formulation of reservoir simulation models, how they are classified, the methods of solution, Jacobian building, the concept of variable substitution and iterative methods of solution of the Jacobian. The operation of a typical simulator is described.

Parameter estimation techniques are the subject of Chapter 4. The least squares objective function is developed from the theory of maximum likelihood estimation. The composite objective function when Bayesian estimation is used to incorporate prior information is also developed. The methods of estimation of the unknown parameters are reviewed and discussed.

In Chapter 5, the normal Gauss-Newton method is detailed, together with a flow chart. The sensitivity equations that have to be solved when the model consists of differential equations are described. The proposed computationally efficient Gauss-Newton method is then described. The chapter then presents the convergence tests using typical chemical kinetic problems.

Chapter 6 covers the application of the proposed method to a single-phase model. The model is used to solve problems reported in the literature and its efficiency compared with other methods.

The incorporation of the proposed method into a three-dimensional three-phase model is the subject of Chapter 7. The model operation and features are described. Numerical tests to match common field observation data are described using a standard industrial problem.

Chapter 8 describes how at convergence, the covariance matrix may be obtained. Identification of correlation between pairs or groups of parameters is discussed. Numerical

examples are given.

In Chapter 9, the incorporation of prior information via Bayesian estimation and use of penalty functions for constrained estimation in the model are discussed. Case studies with incorrect postulated grid models of the true reservoir are presented to show the stabilizing effect of the Bayesian modification and the need for constraints.

Finally, the conclusions and recommendations are presented in Chapter 10.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter the previously published algorithms that have been applied to the problem of automatic history matching in reservoir simulation are reviewed. Nonlinear regression techniques are widely applied in many areas of scientific research, and today, commercial codes are generally available in many statistical packages. However, the typical problems that are solved with these multi-purpose programs involve fairly simple models. The computational time in evaluating the models for each estimate of the regression parameters is usually not of much concern.

Unfortunately this is not the case for reservoir simulation where each model evaluation involves the integration of several thousands of ordinary differential equations over long time intervals. In addition, because the field measurements used as observation variables are available only at a few discrete well locations and at a few times, and because of the large number of combinations of possible parameters, the chosen set of parameters used to obtain a solution may not be unique. This is characterized as an ill-posed problem, where the identified parameters may change significantly but without changing significantly the value of the objective function.

2.1 Methods used for Automatic History Matching

Due to the complexity of the numerical model and the large number of unknown parameters, the literature on automatic history matching in reservoir simulation is relatively sparse. In addition, the standard techniques for parameter estimation cannot be used without modifications. The techniques used in automatic history matching fall into several major categories.

2.1.1 Nonlinear Regression Methods

These methods include the method of steepest descent, the Gauss-Newton method, quasi-linearization and Newton's method.

Jacquard and Jain (1965) were among the early workers in this field. They presented a modified steepest descent method for use with a two-dimensional single-phase model. Jahns (1966), Slater and Durrer (1971), Thomas et al. (1972) also developed variants of the gradient and Gauss-Newton method.

These nonlinear regression methods require in general the calculation of all sensitivity coefficients. These are the partial derivatives of the reservoir variables (such as pressure, saturation and temperature) with respect to the reservoir parameters (typically porosity and permeability).

The usual method used to obtain these coefficients is numerical differentiation. Each parameter is perturbed independently and a full simulation run is made to evaluate the sensitivity of the reservoir variables to the perturbed

parameter as a function of time. This must be repeated at each iteration of the regression for each estimate of the parameters. In general, if there are p parameters to be estimated, it is necessary to perform $(p+1)$ equivalent simulator runs to calculate the sensitivity coefficients. The main advantage of developing the sensitivity coefficients using numerical differentiation is that it can be easily applied to any simulator without significant modification of the simulator.

Another method for calculating the sensitivity coefficients is to derive the governing differential equations by differentiating the original model differential equations with respect to each of the p parameters. The resulting p sensitivity equations are also differential equations that have to be solved in conjunction with the model equations. The work involved to solve each of these p equations is similar to the work to solve the model equations. The required work is therefore equivalent to $(p+1)$ model runs. Thus there is no apparent difference in computational work between the two methods to obtain the sensitivity coefficients.

Since there are normally many unknown parameters and the time required for each simulation run is quite significant, the regression methods as presented in the literature require excessive amounts of computing time which renders them practically unusable. Thus our literature review has found

only applications of these regression methods to simple single-phase problems or small two-phase problems.

2.1.2 Optimal Control Methods

Optimal control methods have also been applied to the history matching problem. Chen et al. (1974) and Chauvent et al. (1975) presented single-phase models based on a first-order optimal control method. The method involves the solution of a set of adjoined ordinary differential equations together with the ordinary differential equations of the model. It therefore requires the equivalent of two simulation runs per iteration of the parameter search. This is a significant improvement over the other published methods. However, the method has only linear convergence properties and a very large number of iterations are required for highly non-linear multiphase problems. Watson et al. (1980) published a two-dimensional two-phase (oil and water) algorithm based on this approach. Yang et al. (1987) used variable metric methods to obtain a better rate of convergence with the optimal control method. The two-phase (oil and water) test problems used were simple one-dimensional ten cell models and two-dimensional ten by ten models. Wasserman et al. (1975) presented a multiphase model based on optimal control theory, however, it matches only pressure data. Furthermore, he claimed that a rigorous optimal control method for a three-dimensional three-phase model would be prohibitive in cost for large scale

simulations.

If a second-order optimal control method is desired, then the sensitivity coefficients have to be calculated too, much like the nonlinear regression methods. Dogru et al. (1981) calculated the computational effort required to develop the sensitivity coefficients using the second-order optimal control method and found that it was more economical to use the nonlinear regression methods if the number of parameters was less than the number of ordinary differential equations of the model. This is usually the case in a reservoir model with several thousand grid cells as the number of parameters that can be reasonably estimated with the observed data is usually much less.

Furthermore, as Makhoul (1990) found, the coefficient matrix of the adjoint equations cannot be easily inverted using the iterative methods developed for the solution of the Jacobian of the original model, and more expensive direct methods must be utilized.

2.1.3 Regularization Methods

The regularization of an ill-posed inverse problem involves the modification of the original problem to one that is well-posed and whose solution approximates that of the original problem. In particular, the usual least squares objective function is augmented with a weighted function which measures the degree of non-smoothness with respect to the

parameter estimates. The augmenting term exerts a penalty action against anomalous oscillations in the parameter estimates.

Lee et al. (1986) applied regularization with bicubic spline approximation to single-phase two-dimensional problems. Chung and Kravaris (1990) incorporated a priori information with the method. Makhlouf et al. (1990) extended the method to a three-dimensional three-phase simulator. The only parameter matched was permeability. However excessive computational times were experienced and algorithmic difficulties were probably also indicated by their conclusion that estimation of absolute permeabilities in three-phase reservoirs is not recommended.

2.1.4 Other Methods

Veatch and Thomas (1971) presented a simple "direct" method that computes porosity and permeability given pressure and saturation observations at every grid location at several times. In practice, this amount of information is never available.

Coats et al.(1970) proposed a method based on linear least squares and linear programming which involved making a number of simulation runs with reservoir properties supplied by a random number generator.

Chen (1988) introduced a generalized pulse spectrum

technique (GPST) for two-dimensional two-phase models. This involves solving the history matching problem using a multigrid technique, in which the grid is successively refined until parameter estimates converge to a solution. The method requires pressure and saturation data at each observation point.

2.2 Identifiability of Parameters

In addition to the problem of obtaining estimates of the parameters, the questions of identifiability of parameters and the accuracy of the estimates need to be addressed.

For single-phase problems, Shah et al. (1978), using covariance analysis, investigated the accuracy of estimates and the optimal level of zonation for one-dimensional situations. Dogru et al. (1977) studied well test situations and developed estimates for the reliability of predictions. Padmanabhan and Woo (1976) used a recursive estimation technique for estimating parameters for well tests. These are valid only for single phase situations. Watson et al. (1984) examined the identifiability and accuracy of parameter estimates for two-phase flow in one-dimensional situations.

2.3 Incorporation of Prior Information and Constraints

With the use of an incorrect model, the automatic history matching problem is usually ill-conditioned. Estimates of parameter values could be unrealistically high or low. The

presence of several local minima in the objective function could result in premature termination of the parameter estimation algorithm. In order to partially alleviate some of these difficulties in the application of the automatic history matching model, it is necessary to incorporate constraints on the minimum and maximum values of the parameter estimates. In addition, when prior estimates of the most probable values of the parameters are available, these should be incorporated into the algorithm.

Gavalas et al. (1976) also incorporated prior geological information in a one-dimensional one-phase reservoir model using a classical Bayesian approach. However, they separated the Bayesian method from the zonation method. For their Bayesian development, they used empirical correlations for the porosity and permeability distributions. The objective function they minimised consisted of the least squares function and a Bayesian contribution. They did not incorporate any constraints in their model. Chung and Kravaris (1990) combined prior information with the method of regularization in a two-dimensional gas model. Relative weighting of the various terms of the objective function was accomplished using rules of thumb in a stepwise procedure. No constraints were employed to limit parameter values. Yang and Watson (1991) used the Bayesian method to estimate relative permeability curves in a two-dimensional two-phase model. In their work, the optimum weighting factor for the Bayesian term

was estimated in an algorithm based on observation of slope changes of the minimum composite objective function obtained from many regression runs using various weighting factors. In the field of groundwater hydrology, Neuman and Yakowitz (1982) also introduced a composite objective function to include a priori estimates of the parameter values. The weighting factor for the a priori information was determined by analyzing residual errors of the estimation. Cooley (1982) extended this approach to include prior information that consisted of nonlinear combinations of several types of parameters.

When the simulation model used is not a true representation of the reservoir, the parameter estimates will often attain unrealistic values. This may be prevented by including parameter inequality constraints. Coats et al.(1970) in a linear programming approach incorporated inequality constraints in single-phase and two-phase models. They noted that the final parameter values in several case studies were usually at their upper or lower limits. Farouq Ali (1988) pointed out that with this algorithm, extremal values were likely to be obtained when the bounds on the parameters are not symmetrical about the true values. Using a Gauss-Newton method, Thomas et al. (1972) introduced box type constraints on the parameters in a single-phase model. Yang et al. (1987) used a sequential quadratic programming approach to incorporate inequality constraints with variable

metric methods in a two-dimensional two-phase model.

2.4 Summary of Review - The State of the Art

As evidenced by the above survey, the problem of automatic history matching in reservoir simulation is an extremely difficult one, compounded by the number of possible parameters, the complexity of the underlying numerical model requiring extensive computational time for each estimate of the parameter values, and the reality that the grid cell model would never be an exact representation of the reservoir.

Due to the requirement of computing the sensitivity coefficients for the nonlinear regression methods, these methods have largely fallen into disfavour for application to automatic history matching. All of the published models in the last decade have utilized the first-order optimal control method because it apparently requires only two equivalent simulation runs per parameter estimation iteration. The disadvantages of the first order optimal control method however are:

(a) its linear convergence characteristics which necessitates hundreds of iterations for highly nonlinear multiphase problems,

(b) the complexity of programming of the method,

(c) the current experience that iterative methods for matrix inversion cannot be used for the adjoined equations,

(d) the unavailability of sensitivity coefficients to compute the variance of the estimated parameters.

The only other published three-dimensional three-phase model, by Makhlouf (1990), uses the method of regularization, requires hours of supercomputing time, and is unable to estimate permeability from usually available field data. There is no currently available three-dimensional three-phase reservoir simulator in the world today that has automatic history matching capability. Reservoir engineers today rely on accumulated experience and perform manual trial and error runs in order to perform history matching.

CHAPTER 3

MATHEMATICAL FORMULATION OF RESERVOIR SIMULATION MODELS

3.1 Classification of Models based on Phase Behaviour

Reservoir models may be classified based on their representation of the phase behaviour of the reservoir. Usually only three immiscible phases are identified in the reservoir. These are the water, oil and gas phases. The various kinds of simulators, in increasing order of complexity are:

(a) Black-Oil Models. These assume that the gas phase contains only gas with no oil or water vaporized in it, the oil phase has oil with gas dissolved in it and the water phase has only water with no oil or gas dissolved in it. The formulation is based on volume balances of the equivalent surface volumes of the three phases. The fluid properties are functions of pressure only. The black-oil model was the earliest form of simulator developed and it is most commonly used due to its simplicity. However, considerable simplifying assumptions have been made in its treatment of phase behaviour which may lead to inaccuracies.

(b) Compositional Models. Compositional models must be used for reservoirs where the phase behaviour depends on composition as well as pressure. These include reservoirs with volatile oil, and gas condensate reservoirs. These

models maintain a mass (molar) balance for each separately identifiable component or chemical species, each of which may exist in any or all of the three phases (water, oil and gas). The phases are mixtures of these components in varying proportions. The compositions of the phases are calculated using flash calculations with equilibrium K values (functions of pressure and composition) or with equations of state (EOS) such as the Peng-Robinson EOS (1976) or the Redlich-Kwong EOS (1949). Viscosities and densities can then be calculated from phase compositions.

Compositional models have been presented in the literature by Nolen (1973), Kazemi et al. (1978), Coats (1980), Young and Stephenson (1983), Acs et al. (1985) and Watts (1986). These models provide the most rigorous treatment of phase behaviour but are more complex to program and time consuming to run. However, they can correctly solve problems such as mixing of fluids with significantly different properties, variable bubble point problems and miscible displacement problems. In order to handle these problems in the framework of a black-oil model, special artifices must be devised, such as that by Ridings (1971) for fluid mixing, and by Todd and Longstaff (1972) for pseudo-miscible displacement.

Limited compositional models have also been introduced. These assume the equilibrium ratios (K values) are functions of pressure only. This assumption is normally quite reasonable for reservoir fluid systems and it also reduces the

need to build multi-dimensional tables of K values. The model used in this thesis belongs to this class of models.

(c) Thermal models. In addition to being fully compositional, these models consider temperature effects on fluid properties. The energy balance is solved in conjunction with the component mass balances. These models are primarily used for simulation of heavy oil reservoirs where steam is injected to provide thermal energy to the oil in order to reduce its viscosity. In such models, steam tables must be built into the simulator to provide the properties of water in the gaseous phase. Thermal models without reaction are called steam models. The first such models were published by Coats (1974, 1976, 1978).

The recovery process of in-situ combustion involves the injection of oxygen (air) into the reservoir. Combustion of the oil in place provides thermal energy that reduces oil viscosity. Carbon dioxide, inert gas and steam resulting from combustion displace oil ahead of the burning front. Simulation programs that model these reaction mechanisms are extremely complex and are called combustion models. Crookston et al. (1979), Coats (1980) and Youngren (1980) have presented combustion models.

(d) Chemical and Polymer Flooding Models. In the chemical/polymer flooding process, surfactant/polymer slugs are injected into the reservoir to reduce oil/water interfacial tension. Phase behaviour of the

oil/water/surfactant-polymer system is extremely complex and multiple liquid phases may be formed. Adsorption of surfactant and polymer on to reservoir rock surfaces reduces the concentration of the injected fluids and must be accounted for in the model. Todd and Chase (1979), Bondor et al. (1972), among others, have published papers on their models.

3.2 Grid Systems

Most simulation models are capable of handling flow in multiple dimensions. In the cartesian coordinate system, the grid cell models may be one-dimensional (linear), two-dimensional areal (x,y) or vertical (x,z) cross-sections, or three-dimensional (x,y,z) full field grid systems. For detailed study of single well behaviour, radial models are used. These can be two-dimensional (r,θ) or three-dimensional (r,θ,z) radial models.

In addition, some models have the capability whereby selected grid cells, such as those containing well locations, may be refined into sub-cartesian grids (local grid refinement), such as that published by Wasserman (1987), or into radial grids (hybrid grid) as presented by Pedrosa and Aziz (1985). The purpose of this is to increase accuracy in the description of displacement fronts, especially near wells without greatly increasing the number of grid cells in the reservoir model. One drawback of these schemes is the probability of obtaining unrealistic flood fronts due to

numerical dispersion. Heineman (1983) has introduced dynamic grid refinement whereby the grid cells may be refined or grossed in time with the passage of a displacement front.

To simulate naturally fractured reservoirs, one approach is to create two domains of grid cells, one for the matrix and one for the fracture. The physical dimensions of these two domains are equivalent so that they occupy the same physical space. Then each cell in the matrix domain is connected by a special connection to the corresponding cell in the fracture domain. The special connections represent transfer functions that approximate the movement of fluids from the matrix cells to the fracture cells. In such grid systems, fluid flow occurs between adjacent cells and between non-neighbour cells as well due to the special connections. Such models have been discussed by Gilman (1987), and Lee and Tan (1987).

3.3 Differential Equations

As discussed in section 3.1, the compositional model has the most rigorous theoretical background and so the development of the simulation equations will be based on the general compositional model. It can be shown that the black-oil model is simply a subset of the compositional model.

We will consider the most general case where there are n_c components, each of which may exist in any or all of the three phases (oil, gas and water). Let x_{I_o} be the mole fraction of the I^{th} component in the oil phase, x_{I_g} the mole fraction of

the I^{th} component in the gas phase and x_{Iw} the mole fraction of the I^{th} component in the water phase. The fundamental equation is the law of conservation of mass, simply stated,

$$(\text{Input}) - (\text{Output}) = (\text{Rate of Accumulation}) \quad (3.1)$$

The rate of fluid flow through porous media is given by Darcy's Law. For each of the three phases, we have

$$Q_g = -K \frac{k_{rg}}{\mu_g} (\nabla P_g - \gamma_g \nabla D) \quad (3.2)$$

$$Q_o = -K \frac{k_{ro}}{\mu_o} (\nabla P_o - \gamma_o \nabla D) \quad (3.3)$$

$$Q_w = -K \frac{k_{rw}}{\mu_w} (\nabla P_w - \gamma_w \nabla D) \quad (3.4)$$

where K is the absolute permeability, k_{rg} , k_{ro} , k_{rw} are the relative permeabilities of the phases, μ_g , μ_o , μ_w are the phase viscosities, γ_g , γ_o , γ_w are the specific weights of the phases, P_g , P_o , P_w are the phase pressures, D is the depth.

As component I can exist in all three phases, the molar flux density for component I is,

$$x_{Ig} \rho_g Q_g + x_{Io} \rho_o Q_o + x_{Iw} \rho_w Q_w \quad (3.5)$$

where ρ_g, ρ_o, ρ_w are the molar densities of the phases.

The rate of accumulation of each phase per unit bulk of porous medium is,

$$\begin{aligned} & \frac{\partial}{\partial t} (\phi S_g) \\ & \frac{\partial}{\partial t} (\phi S_o) \\ & \frac{\partial}{\partial t} (\phi S_w) \end{aligned} \quad (3.6)$$

where S_g, S_o, S_w are the saturations of the phases, ϕ is the porosity, and t is time. The rate of accumulation of component I is therefore,

$$\frac{\partial}{\partial t} [\phi (x_{Ig} \rho_g S_g + x_{Io} \rho_o S_o + x_{Iw} \rho_w S_w)] \quad (3.7)$$

Further we can include a source or sink term, q_I , which is the molar rate of production/injection of component I per unit volume. Then for component I, the conservation equation can be written as,

$$\begin{aligned} & -\nabla \cdot [x_{Ig} \rho_g q_g + x_{Io} \rho_o q_o + x_{Iw} \rho_w q_w] + q_I \\ & - \frac{\partial}{\partial t} [\phi (x_{Ig} \rho_g S_g + x_{Io} \rho_o S_o + x_{Iw} \rho_w S_w)] \end{aligned} \quad (3.8)$$

Substituting Darcy's Law as in Eqs. 3.2 to 3.4 , we get the differential equation,

$$\begin{aligned} \nabla \cdot [x_{Ig}\rho_g K \frac{k_{Ig}}{\mu_g} (\nabla P_g - \gamma_g \nabla D) + x_{Io}\rho_o K \frac{k_{Io}}{\mu_o} (\nabla P_o - \gamma_o \nabla D) \\ + x_{Iw}\rho_w K \frac{k_{Iw}}{\mu_w} (\nabla P_w - \gamma_w \nabla D)] + q_I \\ = \frac{\partial}{\partial t} [\phi (x_{Ig}\rho_g S_g + x_{Io}\rho_o S_o + x_{Iw}\rho_w S_w)] \end{aligned} \quad (3.9)$$

If there are n_c components, we would have n_c differential equations.

3.4 Auxiliary Equations

Additional equations are required to complete the formulation. These are the conditions that the sum of the mole fractions in any phase must be 1 and the sum of the phase saturations must be 1.

$$\sum_{I=1}^{n_c} x_{Ig} = 1 \quad (3.10)$$

$$\sum_{I=1}^{n_c} x_{Io} = 1 \quad (3.11)$$

$$\sum_{I=1}^{n_c} x_{Iw} = 1 \quad (3.12)$$

$$S_o + S_g + S_w = 1 \quad (3.13)$$

If we select the independent variables as $x_{Ig}, (I=1, \dots, n_c), x_{Io}, (I=1, \dots, n_c), x_{Iw}, (I=1, \dots, n_c), P_o, S_o, S_w, S_g$, then all the other variables can be defined as functions of the independent variables.

Densities and viscosities are functions of the phase pressures and compositions.

$$\begin{aligned} \rho_g &= f(P_g, x_{Ig}) \\ \rho_o &= f(P_o, x_{Io}) \\ \rho_w &= f(P_w, x_{Iw}) \\ \mu_g &= f(P_g, x_{Ig}) \\ \mu_o &= f(P_o, x_{Io}) \\ \mu_w &= f(P_w, x_{Iw}) \end{aligned} \quad (3.14)$$

The phase relative permeabilities are expressed as functions of saturation. Stone's (1973) probability model for three-phase relative permeability is commonly used.

$$\begin{aligned} k_{Ig} &= f(S_g) \\ k_{Io} &= f(S_w, S_g) \\ k_{Iw} &= f(S_g) \end{aligned} \quad (3.15)$$

Capillary pressure relationships couple the phase pressures.

$$\begin{aligned} P_g - P_o &= P_{cgo}(S_g) \\ P_o - P_w &= P_{cwo}(S_w) \end{aligned} \quad (3.16)$$

The permeability K is a time independent function of position and is normally directional as well. Thus,

$$\begin{aligned} K_x &= K_x(x, y, z) \\ K_y &= K_y(x, y, z) \\ K_z &= K_z(x, y, z) \end{aligned} \quad (3.17)$$

The depth, D , is a function of position.

$$D = D(x, y, z) \quad (3.18)$$

The porosity is a function of pressure and position.

$$\phi = \phi(P_o, x, y, z) \quad (3.19)$$

Equilibrium values (K -values) relate the mole fractions of component I between each pair of phases. These K -values are functions of pressure and composition, assuming the temperature is fixed in an isothermal model. Thus,

$$\begin{aligned} \frac{X_{Ig}}{X_{Io}} &= K_{Igo}(P_o, X_{jg}, X_{jo}) \\ \frac{X_{Ig}}{X_{Iw}} &= K_{Igw}(P_o, X_{jg}, X_{jw}) \\ \frac{X_{Io}}{X_{Iw}} &= K_{Iow} = \frac{K_{Igw}}{K_{Igo}} \end{aligned} \quad (3.20)$$

There are $2 n_c$ relationships for Eq. 3.20 as the last equation is not independent of the first two.

In an equation-of-state model, the equality of component fugacities in each phase provides the $2 n_c$ relationships.

$$\begin{aligned} f_{Io} &= f_{Ig} \\ f_{Io} &= f_{Iw} \end{aligned} \quad (3.21)$$

3.5 Master Mole Fraction Concept

As there are $3 n_c + 4$ independent variables to be solved for, it is desirable to reduce the number of independent variables. Coats (1980) introduced the master mole fraction concept whereby the mole fractions of component I in each phase are defined in terms of a master mole fraction X_I which is equivalent to the mole fraction of component I in a selected (base) phase.

For each component, a base phase is selected. Then all K-values of that component will be defined with respect to the base phase. For example, if the oil phase is selected as the base phase, the master mole fraction of component I, X_I , is

equal to x_{Io} . Then,

$$\begin{aligned}x_{Io} &= K_{Ioo}x_{Io} = K_{Ioo}x_I = x_I \\x_{Ig} &= K_{Igo}x_{Io} = K_{Igo}x_I \\x_{Iw} &= K_{Iwo}x_{Io} = K_{Iwo}x_I\end{aligned}\tag{3.22}$$

The three variables, x_{Ig} , x_{Io} , x_{Iw} are now expressed in terms of a single variable x_I .

As there are n_c master mole fractions, the independent variables are $X_{I,I=1,nc}$, P_o , S_o , S_g , S_w , that is $n_c + 4$ in number. The number of equations we have consist of the n_c differential equations (Eq. 3.9) and the 4 phase and saturation constraint equations (Eqs. 3.10 to 3.13).

3.6 Boundary Conditions

It is usually assumed that the reservoir lies within a closed space across which there is no flow and fluid injection and production takes place at wells which are represented by point source or sink terms. At each well, either the pressure or the flow rate for each phase is specified.

3.7 Initial Conditions

The usual initial condition of the reservoir is the state of static equilibrium at which velocities of all phases are zero. This is accomplished by specifying phase pressures and initial fluid compositions at a reference depth and using the capillary pressure and density relationships to compute

pressure and saturations using the ordinary differential equations,

$$\frac{dP_p}{dD} = \rho_p(P) g \quad \dots p=0, w, g \quad (3.23)$$

3.8 Discretization

The numerical solution of partial differential equations by the method of finite differences simply involves replacing each of the partial derivatives by equivalent finite difference quotients. The spatial solution region is divided into a discrete set of grid blocks, and the time domain is divided into discrete time intervals. Then, instead of calculating a continuous solution to the differential equations, we obtain an approximate solution at a discrete set of grid cells in the x-y-z space at discrete times. The derivation of the finite difference form of the differential equations is simple but lengthy and is available in a number of numerical analysis texts (Burden and Faires (1981), Smith (1969)). The application to differential equations of reservoir flow has been presented by Peaceman (1977), Crichlow (1977), Settari and Aziz (1979).

The resulting form of the difference equation is, for component I,

$$\begin{aligned}
& \Delta [T X_{Ig} \rho_g \frac{k_{rg}}{\mu_g} (\Delta P_o + \Delta P_{cg} - \gamma_g \Delta D) + \\
& T X_{Io} \rho_o \frac{k_{ro}}{\mu_o} (\Delta P_o - \gamma_o \Delta D) + \\
& T X_{Iw} \rho_w \frac{k_{rw}}{\mu_w} (\Delta P_o - \Delta P_{cw} - \gamma_w \Delta D)] + q_I \\
& - \frac{V}{\Delta t} \delta [\phi (X_{Ig} \rho_g S_g + X_{Io} \rho_o S_o + X_{Iw} \rho_w S_w)]
\end{aligned} \tag{3.24}$$

where T is the physical transmissibility.

3.9 Well Model

The rate of flow of component I, from a well completion located in a grid cell is given by the usual productivity index equation,

$$\begin{aligned}
q_I = (WI) [& \frac{k_{rg}}{\mu_g} \rho_g X_{Ig} (P_b - P_{wb}) \\
& + \frac{k_{ro}}{\mu_o} \rho_o X_{Io} (P_b - P_{wb}) \\
& + \frac{k_{rw}}{\mu_w} \rho_w X_{Iw} (P_b - P_{wb})]
\end{aligned} \tag{3.25}$$

where (WI) is the phase independent part of the productivity index, P_b is the cell pressure and P_{wb} is the flowing bottom hole pressure opposite the centre of the grid cell.

For square grid cells, and in the absence of any near wellbore restrictions, Peaceman (1978) has shown (WI) to be,

$$(WI) = \frac{2\pi kh}{\ln(0.2\Delta x/r_w)} \quad (3.26)$$

More complicated expressions for (WI) are available for non-square grid cells with anisotropic permeability. These have been given by Peaceman (1983).

3.10 Methods of Solution

We now have all the equations and state variables to be determined for each grid cell. Next we have to solve these equations. The aim is to obtain values for the state variables of all the grid cells at each new time level. Commonly used methods of solution include the following.

(a) IMPES solution. In this method, the component equations are combined into a single equation based on pressure only. The values of all the grid cell pressures are then solved for simultaneously. The values of the saturation and composition variables are calculated explicitly after the pressure solution. This procedure was introduced by Stone and Garder (1961) and extended by Macdonald and Coats (1970). The method is fast but has severe restrictions on timestep sizes due to instability.

(b) Semi-Implicit solution. To increase stability, instead of calculating saturations explicitly, the mobility terms and capillary pressures at each new time level are estimated by extrapolation based on the gradient at the

beginning of the timestep. For example, for water relative permeability,

$$k_{rw}^{n+1} = k_{rw}^n + \left(\frac{\partial k_{rw}}{\partial S_w} \right)^n [S_w^{n+1} - S_w^n] \quad (3.27)$$

Letskeman and Ridings (1970) implemented this approach for coning models. It can be shown that this formulation is equivalent to taking just one step of a Newtonian iteration. Thus the problem with this approach is that the gradients used at the beginning of the timestep may not be correct.

(c) Fully Implicit solution. This is the Newtonian iteration solution of the set of simultaneous equations Eqs. 3.24 and 3.10 to 3.13. All state variables in the equations are at the $n+1$ time level. To obtain estimates of the changes in the state variables, a Jacobian matrix is set up and solved for at each iteration. The solution requires a number of iterations until the changes in the state variables become acceptably small.

(d) Adaptive Implicit solution. This is a variant of the fully implicit approach. Depending on the magnitude of the changes of each state variable, each variable is treated implicitly or explicitly, and the treatment could change from iteration to iteration, timestep to timestep. Adaptive implicit methods have been presented by Thomas and Thurnau (1982), Tan (1987).

3.11 Disadvantages and Advantages of Fully Implicit Solution

Fully implicit models require much more extensive coding, significant computational work in matrix operations, and storage for coefficient matrices. The work required for each timestep is considerably more than other solution methods. However, this is more than compensated for by the stability of the model, its ability to take much larger timesteps in problems that exhibit large pore volume throughputs, well coning, gas percolation and high transmissibility variation.

For use in automatic history matching, we would prefer to use a fully implicit model in order to obtain maximum stability and convergence for a wide range of parameter estimates.

3.12 Jacobian Matrix

Eqs. 3.24 and 3.10 to 3.13 are differentiated with respect to the set of unknown state variables of each cell using the chain rule. The resulting Jacobian is shown schematically in Figure 3.1 for a cell containing 3 components and with two neighbours. It will be noted that only the first n_c equations, which are the component mass balances, have flow terms which introduce unknown state variables of neighbouring blocks. The constraint equations for a given cell involve only state variables of that cell. Since the work required to invert a matrix whose elements are submatrices of size n is proportional to n^3 , it is extremely

desirable to further reduce the value of n . We can take advantage of the fact that the constraint equations appear only on the main diagonal of the Jacobian matrix. We can diagonalize the constraint submatrix by Gaussian elimination and use it to eliminate 4 unknowns (columns) in the n_c component balance equations. This results in an n_c by n_c submatrix for each element of the Jacobian matrix and thus significantly reduced computational work to invert the Jacobian.

The procedure to eliminate the constraint submatrix is coded very generally to account for a variable number of phases and components in each cell. Row and column pivoting is used to ensure diagonal dominance.

	Left Cell Upstream	Current Cell	Right cell Downstream
unknowns	P X ₁ X ₂ X ₃ S _o S _w S _g	P X ₁ X ₂ X ₃ S _o S _w S _g	P X ₁ X ₂ X ₃ S _o S _w S _g
Comp 1 Bal.	x x x x x x x	x x x x x x x	x
Comp 2 Bal.	x x x x x x x	x x x x x x x	x
Comp 3 Bal.	x x x x x x x	x x x x x x x	x
$\Sigma x_{I_o} = 1$		x x x x	
$\Sigma x_{I_g} = 1$		x x x x	
$\Sigma x_{I_w} = 1$		x x x x	
$\Sigma S_p = 1$			x x x

Fig 3.1 Schematic of Model Jacobian for 1 Cell with Two Neighbours

If the cells are numbered in normal ordering, the resulting Jacobian will have a banded structure. The Jacobian of a one-dimensional model will have three bands, a two-dimensional model Jacobian will have five bands, and a three-dimensional Jacobian will have seven bands. However, if communicating faults are present, or the model represents dual-porosity, dual-permeability or if special connections joining non-neighbouring cells are set up, then arbitrary sparse matrices will be generated. These need special treatment for iterative matrix inverse methods.

3.13 Solution of Jacobian

The above development reduces to a matrix equation of the form,

$$Ax^{n+1} = b \quad (3.28)$$

where x is the vector of state variables for all the cells. The inverse of the matrix A can be found by direct methods such as the D4 method by Price and Coats (1974) or the nested dissection method of George (1973). When the matrix A is large, however, its inverse is more economically found by an iterative method. Many iterative methods have been published, such as successive over-relaxation (SOR) and its variants (BSOR, LSOR), the strongly implicit procedure (SIP) suggested by Stone (1968), Orthomin and its variants proposed by Vinsome (1976), nested factorization by Appleyard and Chesire (1983),

and iterative D4 with minimization by Tan and Letkeman (1982). The details of these methods are best obtained from the indicated references. The preferred iterative methods are those which do not require any iteration parameters to be supplied by the user and thus are completely transparent to the user. In the automatic history matching model, a sparse matrix iterative solution method was employed.

3.14 Pseudocode Flow Chart of the Simulator

A simplified flow chart of the simulator is presented.

Initialization -

Read in model dimensions, grid system, control parameters, start and stop times, tolerances, unit system.

Read in fluid properties, rock properties, relative permeability relationships, capillary pressures. Read in the reservoir description.

Read in initial reservoir pressures, datum depths, fluid contacts, saturation pressures.

Calculate intercell transmissibilities.

Calculate for each grid cell, the initial pressure and saturations, fluid

compositions and properties such as viscosity and density.

Time Dependent Data -

Read in timestep size, timestep controls, well locations, well rates and constraints, output options etc., at the current time. Increment time step if current time is less than stop time.

Solve -

Calculate well rates, intercell flow rates, accumulation rates. Calculate residuals for the cells and check against tolerances.

If residuals less than tolerances then
 compute material balances;
 update current time;
 print output report if required;
 return to Time Dependent Data;

Else

 set up Jacobian;
 solve linear system using
 iterative method to get Δx ;
 update properties e.g.
 pressures, saturations, fluid

```
properties;  
return to Solve;  
End if
```

CHAPTER 4

PARAMETER ESTIMATION TECHNIQUES

The problem of trying to match a numerical model consisting of algebraic or differential equations to experimental data by variation of parameters appearing in the equations is known as parameter estimation. When the model equations involve linear expressions in the unknown parameters, the solution is usually very simple - the estimates can be obtained directly by solving a set of simultaneous linear equations. However, if the model is nonlinear, the solution involves the finding of the minimum value of a nonlinear function in multi-dimensional space. This requires an iterative solution which may or may not converge to the global optimum, due to the presence of local minima or extremely shallow valleys near the optimum.

In general, there are two classes of models, algebraic models and differential equation models. In the algebraic form,

$$y = f(x, k) \tag{4.1}$$

we have a vector of measured or observed variables y , a vector x of known state variables, or which have been set by the experimentalist, and a vector of unknown parameters k .

More difficult to solve are models which are formulated

in terms of differential equations. A special subclass are the standard dynamic models which are characterized by a set of state variables which change with time according to first order nonlinear differential equations.

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{k}; \mathbf{u}) \quad (4.2)$$

Here \mathbf{x} is the vector of state variables, \mathbf{k} is the vector of unknown parameters, and \mathbf{u} is a vector of user specified variables. The initial conditions are usually known and some of the states of the system are observed at various points in time. The vector of calculated variables, \mathbf{y} , is related to the state vector \mathbf{x} by the nonlinear relationship,

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)) \quad (4.3)$$

Examples of dynamic systems include chemical reaction systems where the kinetic parameters have to be estimated.

The automatic history matching problem in reservoir simulation is a classical nonlinear parameter estimation problem. The numerical model consists of a set of mixed differential and algebraic equations as discussed in Chapter 3, and is equivalent to those considered in the chemical engineering literature by Petzold (1982), and Caracotsios and Stewart (1985). The parameters \mathbf{k} , to be estimated usually consist of the porosity and permeability distributions, though other physical properties such as rock compressibility,

relative permeability, fluid properties may also be considered as parameters.

In parameter estimation, we try to find a set of \mathbf{k} such that some scalar function of the errors is minimised. The differences between the observed variables $\hat{\mathbf{y}}$ and the calculated variables \mathbf{y} is the vector of residuals, $\boldsymbol{\varepsilon}$,

$$\boldsymbol{\varepsilon} = \hat{\mathbf{y}} - \mathbf{y} \quad (4.4)$$

In least squares estimation, the objective function $S_{LS}(\mathbf{k})$ is defined as the sum of the squares of the residuals.

$$S_{LS}(\mathbf{k}) = \sum_{i=1}^{n_o} [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)]^T \mathbf{Q}_i [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)] \quad (4.5)$$

In general, the values of \mathbf{Q} , the weighting matrix, are chosen to scale the relative magnitude of the measured variables. On statistical grounds, if the error between the measured and the predicted variables have zero mean and a known covariance matrix \mathbf{V} , then \mathbf{Q} should be the inverse of this covariance. This can be proved using the method of maximum likelihood (ML) estimation.

4.1 Development of Least Squares from Maximum Likelihood Estimation

In order to provide a mathematical basis for parameter estimation, it is necessary to assume that the errors in the

observations have a probability distribution, usually the normal distribution. Then the parameters in the probability model of the errors may be estimated together with the unknown parameters of the model. The joint probability density function of all the errors ε in the observed variables is therefore $p(\varepsilon|\psi)$. This is also called the likelihood function.

$$L(k, \psi) = p(e|\psi) = p(\hat{y}-y|\psi) \quad (4.6)$$

Here ψ represents the parameters in the normal distribution such as the mean and the variance. The variables y are calculated using the numerical model given any values for the model parameters k . The maximum likelihood estimate is obtained by finding the values of k and ψ which maximize L .

Since the logarithm is a monotonic increasing function of its argument, the values of k, ψ that maximize L also maximize $\log L$. Taking the logarithms often simplifies the final objective function to be maximized, and this is the case here. Thus,

$$L^*(k, \psi) = \log L(k, \psi) = \log p(\hat{y}-y|\psi) \quad (4.7)$$

If the errors in different experiments, or in groups of observations at each observation time, are independent, the joint probability density p is the product of the individual experiments' probability densities, and,

$$L^*(k, \Psi) = \sum_{i=1}^{n_o} \log p(\hat{y}(t_i) - y(t_i) | \Psi) \quad (4.8)$$

The normal distribution, p , is given by,

$$p(\mathbf{e}_i, \mathbf{V}_i) = \frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{V}_i|^{\frac{1}{2}}} \exp(-\frac{1}{2} \mathbf{e}_i^T \mathbf{V}_i^{-1} \mathbf{e}_i) \quad i=1, 2, \dots, n_o \quad (4.9)$$

where it is assumed the errors in each experiment are independent with zero means and covariance matrix \mathbf{V}_i . \mathbf{V}_i has the dimensions of $m \times m$ where there are m observations in each experiment at observation time t_i .

Substituting Eq. 4.9 into Eq. 4.8, we obtain,

$$L^*(k, \mathbf{V}_i) = -\frac{mn_o}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n_o} \log (|\mathbf{V}_i|) - \frac{1}{2} \sum_{i=1}^{n_o} \mathbf{e}_i^T \mathbf{V}_i^{-1} \mathbf{e}_i \quad (4.10)$$

Next, if we assume that the value of the covariance matrices $\mathbf{V}_i, i=1, \dots, n_o$ are known, then the only non-constant term in Eq. 4.10 is

$$-\frac{1}{2} \sum_{i=1}^{n_o} \mathbf{e}_i^T \mathbf{V}_i^{-1} \mathbf{e}_i \quad (4.11)$$

and maximizing $L^*(\mathbf{k})$, is equivalent to minimizing,

$$S_{ML}(\mathbf{k}) = \sum_{i=1}^{n_o} \mathbf{e}_i^T \mathbf{V}_i^{-1} \mathbf{e}_i \quad (4.12)$$

Comparing equations 4.5 and 4.12, we see that maximum likelihood estimation is equivalent to least squares when $\mathbf{Q}_1 = \mathbf{V}_1^{-1}$. It should be noted that \mathbf{V} is the covariance matrix of the errors of the observations and is not the covariance matrix of the probability distribution of the unknown parameters \mathbf{k} , which we will consider next in Bayesian estimation.

4.2 Bayesian Estimation

There is usually some prior information about the permissible values, or the most probable values of the unknown parameters \mathbf{k} . For example, it is obvious that the value of the porosity parameter must be a positive number between 0 and 1. The most probable value of a permeability parameter may be obtained from well test analysis of a well located in the parameter zone. This prior information can be put on a mathematical basis by assigning a probability distribution to it and then incorporating it into the objective function. This prior distribution $p_o(\mathbf{k})$ biases the parameter search in favour of parameter values for which the prior density function is relatively large. Naturally the bias decreases as the number of observations increases.

In practice, our prior information on a parameter k often takes the form of a mean value $k_m \pm \sigma$, where σ is the standard deviation. A normal prior distribution is often used for simplicity as suggested by Bard (1974), namely

$$p_o(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(k-k_m)^2}{2\sigma^2}\right) \quad (4.13)$$

The prior information may be taken into account in the estimation process by combining the likelihood function L , and the prior density. The posterior density is proportional to their product,

$$L(k, \psi) p_o(k) \quad (4.14)$$

which follows from Bayes' Theorem (1763).

Following the same discussion as in the previous section, instead of maximizing Eq. 4.7, we maximize,

$$L_B(k, \psi) = \log p(\hat{y}-y|\psi) + \log p_o(k) \quad (4.15)$$

Writing the prior distribution of the parameters, $p_o(k)$ for p parameters with a known covariance matrix, W , of dimension $p \times p$, and with mean k_m , we obtain,

$$p_o(k) = (2\pi)^{-\frac{p}{2}} |W|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (k-k_m)^T W^{-1} (k-k_m)\right] \quad (4.16)$$

Then the logarithm of $p_o(\mathbf{k})$ is

$$\log p_o(\mathbf{k}) = -\frac{D}{2} \log(2\pi) - \frac{1}{2} \log |W| - \frac{1}{2} (\mathbf{k} - \mathbf{k}_m)^T W^{-1} (\mathbf{k} - \mathbf{k}_m) \quad (4.17)$$

The Bayesian objective function to be minimized is obtained from Eq. 4.15 by substituting Eq. 4.9 and Eq. 4.17, ignoring constant terms and changing signs.

$$S_B(\mathbf{k}) = \sum_{i=1}^{n_o} [\hat{y}(t_i) - \mathbf{y}(t_i)]^T \mathbf{Q}_i [\hat{y}(t_i) - \mathbf{y}(t_i)] + (\mathbf{k} - \mathbf{k}_m)^T W^{-1} (\mathbf{k} - \mathbf{k}_m) \quad (4.18)$$

This composite objective function will be used in Chapter 9, when we incorporate prior information into the automatic history matching model.

4.3 Methods of Estimation

The optimum value of \mathbf{k} has to be found by an iterative procedure. Gradient methods of solution are usually employed. At the j^{th} iteration, we have a current estimate of $\mathbf{k}^{(j)}$, and we seek a new value $\mathbf{k}^{(j+1)}$ using the equation,

$$\mathbf{k}^{(j+1)} = \mathbf{k}^{(j)} - \lambda \mathbf{R} \mathbf{g} \quad (4.19)$$

where λ is a scalar, R is a matrix to be defined, and g the gradient vector of S , i.e. $g^{(j)} = \partial S / \partial k^{(j)}$. Various gradient methods differ from each other in the choice of R and λ . R determines the new step direction from $k^{(j)}$ and λ defines the step length.

In the method of steepest descent, $R = I$. This method converges very slowly for ill-conditioned problems, and thus is not suitable for reservoir parameter estimation.

When R is the inverse of the Hessian matrix, i.e. the matrix of second derivatives, then the algorithm becomes the Newton-Raphson method. This method has quadratic convergence properties, but it requires computation of the second derivatives of S , a very costly process, especially for differential equations.

The Gauss-Newton method computes an approximation to the Hessian matrix without actually computing any second derivatives of the model equations. Only the first derivatives, $\partial f / \partial k$ are required. One difficulty with this method is the possibility of overshoot when the linear approximation is not quite valid. Then the step length λ should be reduced, and the optimum step size is often found by interval halving or subsidiary optimization along the step direction.

A class of methods termed variable metric, or quasi-Newton, exist that approximate the Hessian matrix but use information from only first derivatives. These methods employ

conjugate directions. In essence, estimates of the Hessian are successively improved with each iteration.

In a large number of test problems, Bard (1970) found the Gauss-Newton method to be somewhat more efficient than the other methods. We will use the Gauss-Newton method in this work because it possesses quadratic convergence properties, does not require second derivatives, and because we can compute statistical information from the Gauss-Newton matrix at convergence to provide reliability estimates of the final parameter values.

CHAPTER 5

PROPOSED COMPUTATIONALLY EFFICIENT GAUSS-NEWTON METHOD

In this chapter, the proposed modifications to the Gauss-Newton method are developed. First, we will restate the problem, present the derivation of the Gauss-Newton method, outline the normal implementation of the method, and then introduce the modifications that would significantly reduce the computational effort.

5.1 Problem Statement

Any reservoir model, whether it is a simple black-oil model, a compositional model or a thermal model, and regardless of its dimensionality, can be represented generally by the set of ordinary differential equations,

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{k}; \mathbf{u}) \quad (5.1)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (5.2)$$

where \mathbf{x} is an n dimensional vector of state variables (such as pressure and saturations), \mathbf{x}_0 is the initial state, \mathbf{k} is the p dimensional vector of parameters to be estimated (usually porosity, permeability) and \mathbf{u} represents the user specified variables (e.g. flowrates).

The vector of calculated variables, \mathbf{y} , is related to the

state vector \mathbf{x} by the nonlinear relationship,

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)) \quad (5.3)$$

As an example, the calculated water-oil ratio of a well is a function of the pressure, saturations and fluid composition of the grid cells the well is completed in.

In many problems, Eq. 5.3 assumes the linear form,

$$\mathbf{y}(t) = \mathbf{C} \mathbf{x}(t) \quad (5.4)$$

where \mathbf{C} is an $m \times m$ matrix. This occurs when the calculated variable is also a state variable, for example, pressure. In particular, for reservoir simulation problems, this matrix is usually very sparse with zero entries everywhere except where the corresponding state variable is involved in a functional relationship with the calculated variable. This implies the work required to solve the numerical model is extremely disproportionate to the number of measured variables. The function evaluations are therefore time consuming and as a result the number of function evaluations should be minimized.

In least squares (LS) estimation, the parameters are obtained by minimizing the LS objective function.

$$S_{LS}(\mathbf{k}) = \sum_{i=1}^{n_o} [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)]^T \mathbf{Q}_i [\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)] \quad (5.5)$$

where \mathbf{Q} is an $m \times m$ user supplied weighting matrix.

5.2 The Gauss-Newton Method

The method used to obtain successive estimates of the parameter vector \mathbf{k} is the Gauss-Newton method. This is a quadratically convergent method requiring only first derivatives. A brief description of the method as presented by Kalogerakis and Luus (1983) follows.

Suppose at the j^{th} iteration, an estimate $\mathbf{k}^{(j)}$ of the unknown parameter vector is available. To obtain the next estimate $\mathbf{k}^{(j+1)}$, the output vector \mathbf{y} is expanded by Taylor series to yield,

$$\mathbf{y}^{(j+1)}(t) = \mathbf{y}^{(j)}(t) + \left(\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} \right)^T \left(\frac{\partial \mathbf{x}^T}{\partial \mathbf{k}} \right)^T (\mathbf{k}^{(j+1)} - \mathbf{k}^{(j)}) \quad (5.6)$$

When the measured variables are linearly related to the state variables, substitution of Eq. 5.4 into Eq. 5.6 yields,

$$\mathbf{y}^{(j+1)}(t) = \mathbf{C} \mathbf{x}^{(j)}(t) + \mathbf{C} \mathbf{G}(t) \Delta \mathbf{k}^{(j+1)} \quad (5.7)$$

where $\mathbf{G}(t)$ is the $n \times p$ sensitivity matrix $(\partial \mathbf{x}^T / \partial \mathbf{k})^T$. Substituting $\mathbf{y}^{(j+1)}(t)$ from Eq. 5.7 into the objective function Eq. 5.5, and setting $\partial S / \partial \mathbf{k}^{(j+1)} = 0$, we obtain a set of linear algebraic equations,

$$\begin{aligned} & \left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} \mathbf{C}(t_i) \mathbf{G}(t_i) \right] \Delta \mathbf{k}^{(j+1)} \\ & - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} (\mathcal{Y}(t_i) - \mathbf{y}(t_i)) \end{aligned} \quad (5.8)$$

which is of the form $\mathbf{A} \Delta \mathbf{k}^{(j+1)} = \mathbf{b}$. This set of linear equations can be solved to yield $\Delta \mathbf{k}^{(j+1)}$ and thus the next estimate $\mathbf{k}^{(j+1)}$ of the parameter vector is obtained.

Similarly for the nonlinear case of Eq. 5.3, linearization of the output vector around $\mathbf{k}^{(j)}$ yields,

$$\mathbf{y}^{(j+1)}(t) = \mathbf{y}^{(j)}(t) + \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right)^T \left(\frac{\partial \mathbf{x}^T}{\partial \mathbf{k}} \right)^T (\mathbf{k}^{(j+1)} - \mathbf{k}^{(j)}) \quad (5.9)$$

Following the same development that led to Eq. 5.8, we obtain,

$$\begin{aligned} & \left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right) \mathbf{Q} \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right)^T \mathbf{G}(t_i) \right] \Delta \mathbf{k}^{(j+1)} \\ & - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right) \mathbf{Q} (\mathcal{Y}(t_i) - \mathbf{y}(t_i)) \end{aligned} \quad (5.10)$$

Thus when the observation relationship is nonlinear of the form of Eq. 5.3, we use Eq. 5.10 for the Gauss-Newton method.

The sensitivity matrix \mathbf{G} is obtained by differentiating both sides of the model equation Eq. 5.1 with respect to \mathbf{k} to yield

$$\frac{d\mathbf{G}(t)}{dt} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^T \mathbf{G}(t) + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{k}}\right)^T \quad (5.11)$$

To obtain the initial conditions, it is observed that the initial state vector \mathbf{x}_0 is independent of the parameter vector \mathbf{k} , and therefore,

$$\mathbf{G}(t_0) = 0 \quad (5.12)$$

5.3 Integration Method - The Implicit (Euler) Method

In chemical engineering, the integration methods used for ordinary differential equations are usually higher order methods such as the Runge-Kutta or Gear's method. For reservoir modelling, the equations involved are partial differential equations, and these are reduced by grid discretization into a set of ordinary differential equations. The methods used for integration of these parabolic equations are either explicit (forward) difference, or implicit (backward) difference methods. The implicit difference method is equivalent to the implicit Euler method. More advanced reservoir simulators use the implicit method because of its unconditional stability in spite of the additional work in setting up the Jacobian and solving the simultaneous sets of equations. Much larger timesteps can be used with the implicit method. The implicit method will therefore be used in this work.

To integrate the model equations, Eq. 5.1 is expressed in

difference form,

$$\frac{\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}}{\Delta t} = \mathbf{f}(\mathbf{x}^{(n+1)}, \mathbf{k}; \mathbf{u}) \quad (5.13)$$

and $\mathbf{x}^{(n+1)}$ is obtained iteratively. Suppose at the 1th iteration, we have an estimate $\mathbf{x}^{(1)}$ of the state vector. To obtain the next estimate $\mathbf{x}^{(1+1)}$, we have the expansions,

$$\mathbf{x}^{(1+1)} = \mathbf{x}^{(1)} + \Delta \mathbf{x} \quad (5.14)$$

and

$$\mathbf{f}(\mathbf{x}^{(1+1)}, \mathbf{k}; \mathbf{u}) = \mathbf{f}(\mathbf{x}^{(1)}, \mathbf{k}; \mathbf{u}) + \left[\left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \right]^{(1)} \Delta \mathbf{x} \quad (5.15)$$

Substituting these into Eq. 5.13, and rearranging,

$$\left[\left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T - \frac{1}{\Delta t} \right]^{(1)} \Delta \mathbf{x} = \frac{\mathbf{x}^{(1)} - \mathbf{x}^{(n)}}{\Delta t} - \mathbf{f}(\mathbf{x}^{(1)}, \mathbf{k}; \mathbf{u}) \quad (5.16)$$

This is of the form $\mathbf{A}\Delta\mathbf{x} = \mathbf{b}$, and can be solved for $\Delta\mathbf{x}$ and thus to obtain $\mathbf{x}^{(1+1)}$. Thus the model equation can be integrated to obtain values of the state vector \mathbf{x} at the observation times. The step size, Δt , is usually adjusted to control local truncation error.

The sensitivity equations can be integrated using the implicit method in a similar way. The resulting iterative equation is

$$\begin{aligned} & \left[\left(\frac{\partial \mathbf{F}^T}{\partial \mathbf{x}} \right)^T - \frac{1}{\Delta t} \right]^{(l)} \Delta \mathbf{G} \\ & - \frac{\mathbf{G}^{(l)} - \mathbf{G}^{(n)}}{\Delta t} - \left[\left(\frac{\partial \mathbf{F}^T}{\partial \mathbf{x}} \right)^T \right]^{(l)} \mathbf{G}^{(l)} - \left[\left(\frac{\partial \mathbf{F}^T}{\partial \mathbf{k}} \right)^T \right]^{(l)} \end{aligned} \quad (5.17)$$

The procedure for the Gauss-Newton method thus consists of the solution of equations 5.16 and 5.17 to obtain values of the state vector \mathbf{x} and the sensitivity matrix \mathbf{G} . At the observation times, these values of \mathbf{x} and \mathbf{G} are used in Eq.5.8 or 5.10 to generate the Gauss-Newton matrix.

5.4 Normal Implementation of the Gauss-Newton method

The model equation, Eq.5.1, represents a set of n ODEs and the sensitivity equation, Eq. 5.11, represents a set of $n \times p$ ODEs. Most authors such as Bard (1970), Thomas (1972), Dogru (1981) in describing the implementation of the Gauss-Newton method have recommended that all $n(p+1)$ equations be integrated simultaneously using the same integration step sizes. This integration must be performed at each iteration of the Gauss-Newton method in order to supply the values of \mathbf{x} and \mathbf{G} required in Eq.5.8. While this is practical for small values of n and p , the simultaneous integration would not be feasible for reservoir simulation models. A typical reservoir model may have a value of n of 10000, and the number of parameters, p , may be 20. The normal Gauss-Newton implementation would require the simultaneous integration of

Table 5.1: Pseudo-Code Flow Chart for Gauss-Newton Method.

Flow Chart (Pseudo-Code)

1. Read initial model parameters, k .
2. Integrate model and sensitivity equations over time domain. Compute LS objective function. Invert matrix size $n(p+1)$ each Newtonian iteration in each timestep.
3. Generate Gauss-Newton matrix and solve for Δk .
4. Using new estimate of k , integrate model equations, compute LS objective function.
5. If objective function value is $>$ previous value, adjust Δk , and goto 4.
6. If $\Delta k >$ tolerance goto 2, else converged

200000 ODEs over the history match time period for each Gauss-Newton iteration.

The pseudo-code for a typical flow chart for the Gauss-Newton method is shown in Table 5.1.

5.5 Efficient Implementation of the Gauss-Newton Method

We will now introduce two important considerations that will significantly reduce the computational effort for the generation of the sensitivity coefficients and thus lead to an efficient and practical implementation of the Gauss-Newton method.

5.5.1 Matrix Inversion

If the model equations Eq. 5.16 and the sensitivity equations, Eq. 5.17, are compared, it will be observed that they have the same coefficient matrix. If the unknown vectors $\Delta \mathbf{x}$ and $\Delta \mathbf{G}$ are linearized into a single vector, $\Delta \mathbf{x}$, $\Delta(\partial \mathbf{x} / \partial k_1)$, $\Delta(\partial \mathbf{x} / \partial k_2), \dots \Delta(\partial \mathbf{x} / \partial k_p)$, of length $n(p+1)$, then the Jacobian matrix is of the form shown in Figure 5.1. Now if the equations are simultaneously integrated by a library routine, it is quite possible that the entire $n(p+1)$ matrix will be inverted without recognizing the fact that the coefficient matrix is identical for each successive set of n equations.

F					$\Delta \mathbf{x}$
	F				$\Delta \left(\frac{\partial \mathbf{x}}{\partial k_1} \right)$
		F			$\Delta \left(\frac{\partial \mathbf{x}}{\partial k_2} \right)$
		
				F	$\Delta \left(\frac{\partial \mathbf{x}}{\partial k_p} \right)$

where **F** is

$$\left[\left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T - \frac{1}{\Delta t} \right]$$

Fig. 5.1 Schematic of Jacobian of Model and Sensitivity Equations

Very substantial savings in computation time can be obtained by inverting the coefficient matrix of the first n equations and then performing matrix multiplication to get the remaining p sets of n unknowns. Typical reservoir models utilize several thousand grid cells for reservoir description and the computational effort in inverting a matrix of this size usually takes up to 70% of the computing time for a single timestep even with an efficient iterative method. Therefore in utilizing the inverted Jacobian from the model equations at each iteration, the time it takes to integrate the $n \times p$ sensitivity equations is much less than $n \times p$ equivalent model runs.

5.5.2 Sequential Integration of Sensitivity Equations

During the integration of the model equations, timestep failures may frequently occur. These failures may occur if the current timestep size is too large initially, as result of which the model iteration equation, Eq. 5.16, may not converge in a reasonable number of iterations. In other cases, to control truncation error, maximum allowable changes are set for each timestep and if these changes are exceeded, the timestep is repeated. The timestep size is reduced successively until convergence occurs or allowable changes over the timestep occur.

If the model and sensitivity equations are simultaneously integrated, then for timestep cuts, the work performed in

integrating the sensitivity equations would have had been performed in vain. It is therefore suggested that the sensitivity equations should be integrated only after the model equations have converged for any timestep, i.e. the solutions of the model and sensitivity equations are decoupled. We would first solve the model equations for a given timestep, and then the sensitivity equations for the same timestep, and repeat the sequence for the next timestep. This is shown schematically in Figure 5.2.

This sequential solution has another significant advantage. Referring to the iteration equation for the sensitivity equations, Eq. 5.17, it can be seen that the coefficients for the sensitivity equations are iteration dependent. As the values of the state variables x change during the solution of the model equations, the coefficient matrix of the sensitivity equations changes, but these changes are not a result of the changes in the sensitivity variables $\partial x / \partial k$. This implies a different set of sensitivity equations is created by each model iteration and this may cause additional iterations in the convergence of the sensitivity equations in the course of a timestep. The solution of the sensitivity equations after the model equations in a given timestep implies the coefficients of the sensitivity equations will be constant, effectively linearizing the sensitivity equations.

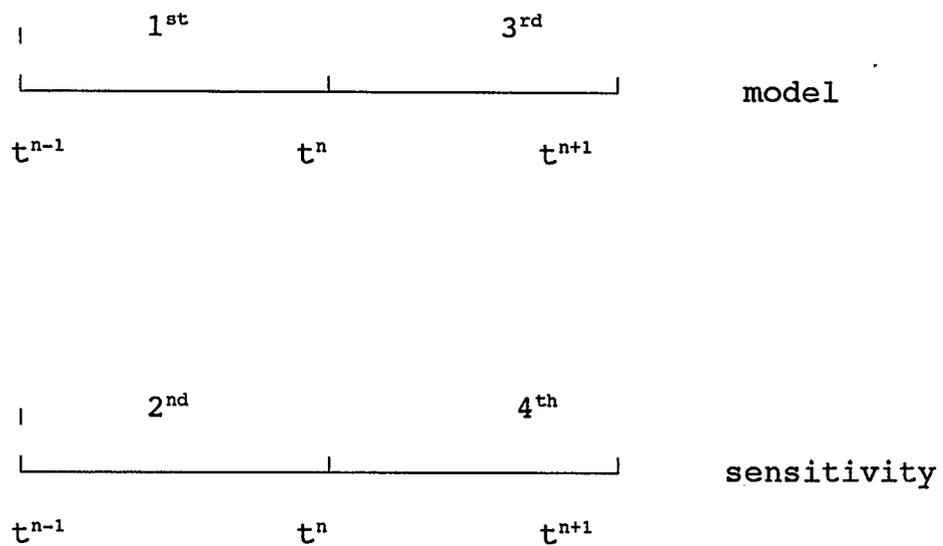


Fig. 5.2 Schematic of Sequential Solution of Model and Sensitivity Equations.

The combination of sequential integration and the utilization of the inverted Jacobian from the model equations to solve the sensitivity equations greatly reduces the work effort to obtain the sensitivity coefficients. The sensitivity equations are solved only once at the last iteration of each successful timestep of the implicit procedure for the evaluation of the state variables of the model equations, rather than simultaneously at every iteration of every successful or unsuccessful timestep. Furthermore, the sensitivity coefficients are simply obtained by matrix multiplication using a previously inverted Jacobian available from the model equations. It is therefore expected that the work to solve the sensitivity equations for p parameters will be very much less than the p equivalent model runs required in the normal implementation of the Gauss-Newton method.

5.6 Convergence Testing using Chemical Kinetic Models

It is of interest to investigate whether or not the proposed method is convergent. Numerical experimentation was performed with the four typical chemical engineering kinetic problems used by Kalogerakis and Luus (1983) in their work on using the information index to extend the region of convergence. The initial values of the parameter estimates were set at the limits of the regions of convergence reported in the paper.

5.6.1 Example 1

This is the pyrolytic dehydrogenation of benzene to diphenyl described using the two parameter model,

$$\begin{aligned} \frac{dx_1}{dt} &= -r_1 - r_2 & x_1(0) &= 1 \\ \frac{dx_2}{dt} &= r_1/2 - r_2 & x_2(0) &= 0 \end{aligned} \quad (5.18)$$

$$r_1 = k_1(x_1^2 - x_2(2 - 2x_1 - x_2))/0.726$$

$$r_2 = k_2(x_1x_2 - (1 - x_1 - 2x_2)(2 - 2x_1 - x_2))/3.852$$

where x_1 and x_2 are the moles of benzene and diphenyl per mole of benzene feed. This system was studied by Seinfeld and Gavalas (1970) for parameter estimation. The parameters k_1 and k_2 are to be estimated from the measured values of the state variables x_1 and x_2 obtained as a function of time shown in Figure 5.3. The optimum values (k_1^* , k_2^*) of the parameters have previously been determined to be 355.4 and 403.3. Table 5.2 shows the number of iterations required by the proposed Gauss-Newton method to converge to the optimum values from different starting points. The table column heading of (100,100) means that the initial guess of k_1 is 100 x k_1^* , and of k_2 is 100 x k_2^* . A listing of the program is provided in Appendix A.

Fig. 5.3 - Given Measurements for Example 1

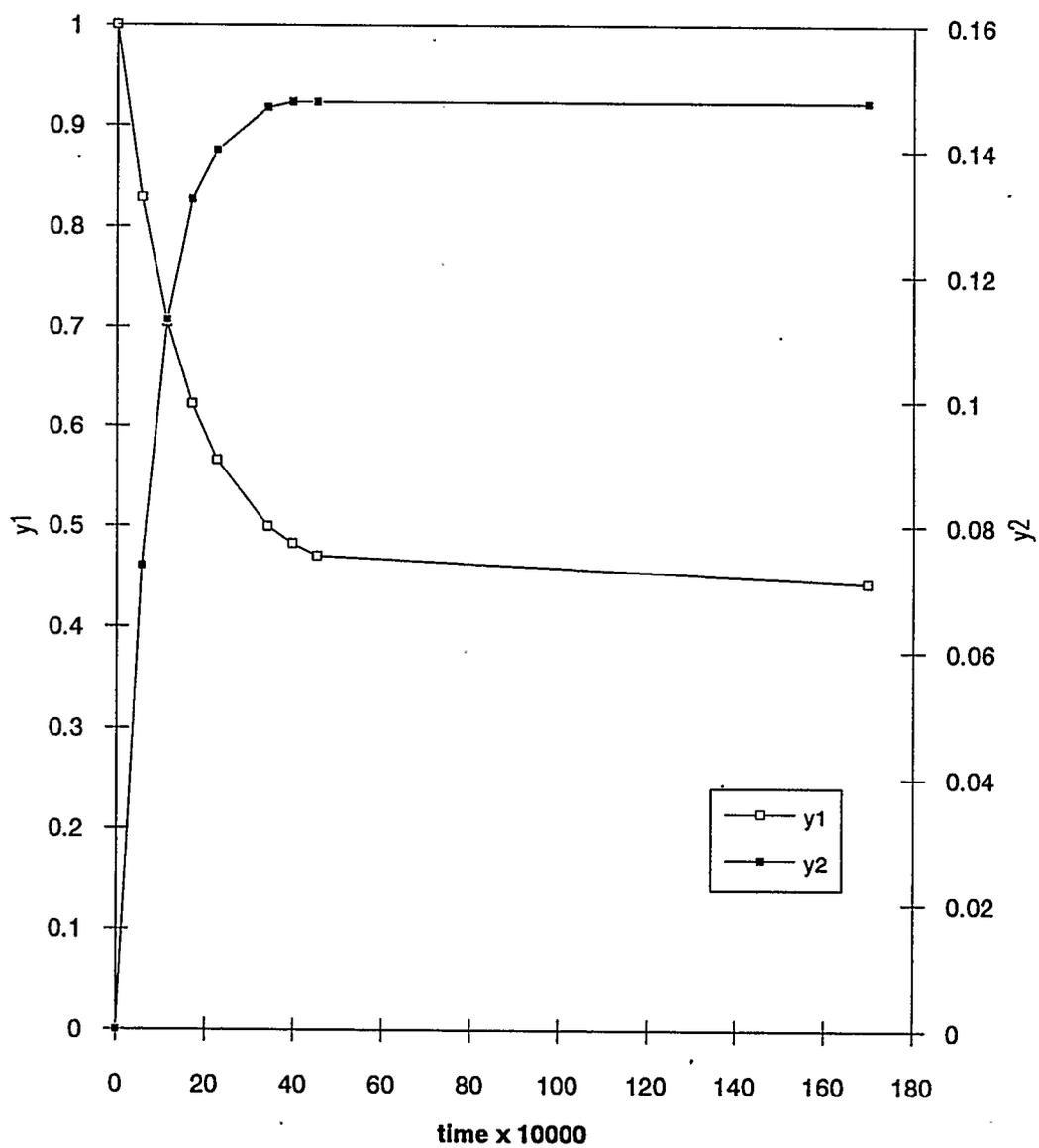


Table 5.2: Number of Iterations for Example 1				
$(k_1^0/k_1^*, k_2^0/k_2^*)$	(100,100)	(1,100)	(100,1)	(10000,10000)
iterations	9	10	9	15

5.6.2 Example 2

This is a model of an isothermal CSTR with complex reactions, used by Lapidus and Luus (1967) for optimal control studies. There are seven state variables, x_1 to x_7 , and seven differential equations.

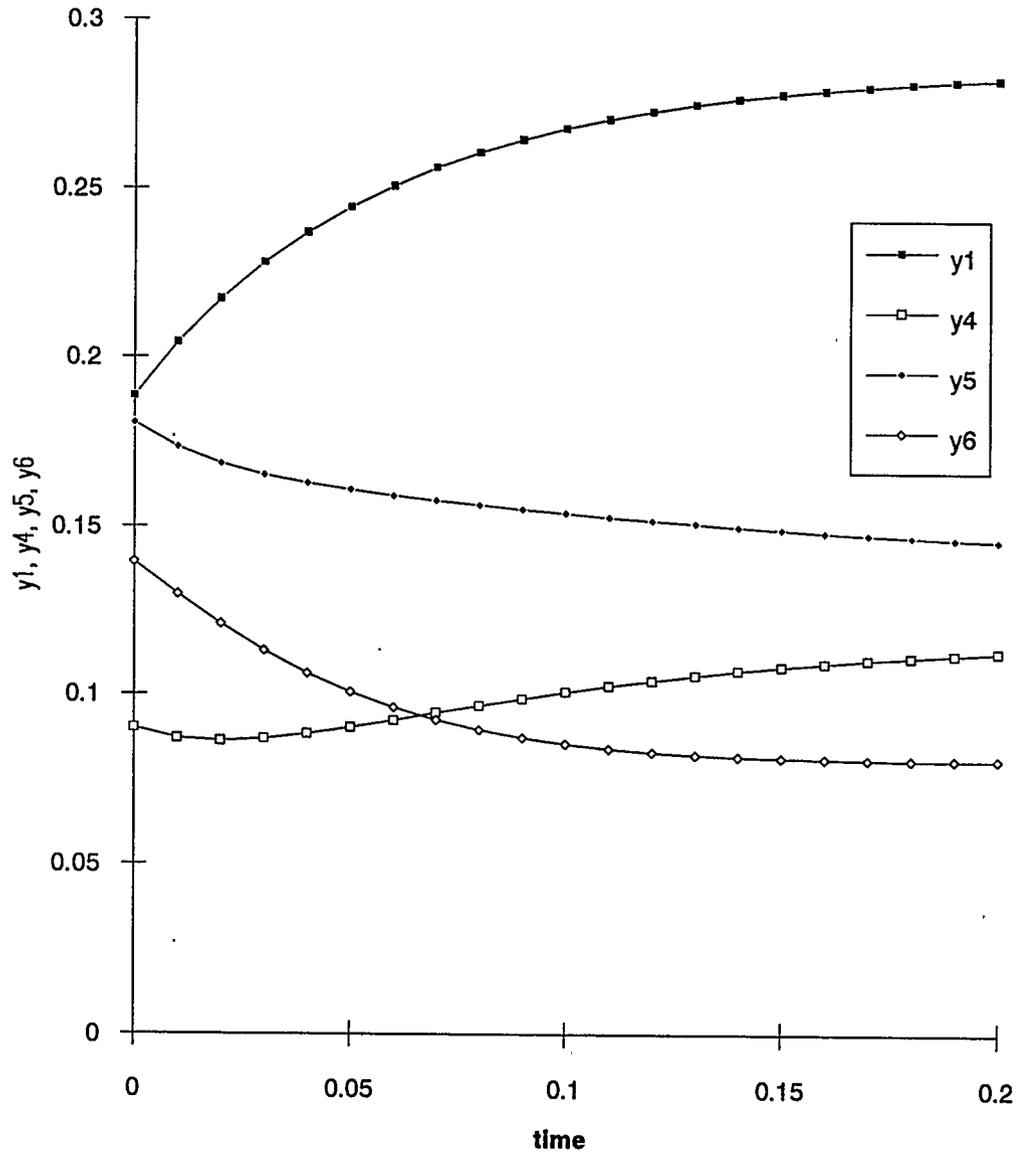
$$\begin{aligned}
 dx_1/dt &= k_5 - qx_1 - k_1 x_1 x_2 - k_4 x_1 x_6 \sqrt{0.9}; & x_1(0) &= 0.1883 \\
 dx_2/dt &= 7.0 - qx_2 - k_1 x_1 x_2 - 2k_2 x_2 x_3; & x_2(0) &= 0.2507 \\
 dx_3/dt &= 1.75 - qx_3 - k_2 x_2 x_3; & x_3(0) &= 0.0467 \\
 dx_4/dt &= -qx_4 + 2k_1 x_1 x_2 - k_3 x_4 x_5; & x_4(0) &= 0.0899 \\
 dx_5/dt &= -qx_5 + 3k_2 x_2 x_3 - k_3 x_4 x_5; & x_5(0) &= 0.1804 \\
 dx_6/dt &= -qx_6 + 2k_3 x_4 x_5 - k_4 x_1 x_6 \sqrt{0.9}; & x_6(0) &= 0.1394 \\
 dx_7/dt &= -qx_7 + 2k_4 x_1 x_6 \sqrt{0.9}; & x_7(0) &= 0.1046 \\
 q &= 8.75 + k_5
 \end{aligned}$$

(5.19)

There are five unknown parameters, k_1 to k_5 and four observed state variables x_1 , x_4 , x_5 , x_6 . The values of the parameters are assumed to be $k_1 = 17.6$, $k_2 = 73.0$, $k_3 = 51.3$, $k_4 = 23.$, $k_5 = 6.0$. Using these values, simulated measurements of the observed variables were obtained by integrating the model equations. This data is shown in Figure 5.4, and is used as the observed data to match by varying the initial values of the parameters. The values of k were varied in sequence, and α is the normalized distance of the initial guess from the optimum. Thus an initial guess of $(0.176 \times 10^3, 0.731 \times 10^2, 0.513 \times 10^2, 0.23 \times 10^2, 0.601 \times 10^1)$ corresponds to $\alpha = 10$ in the direction $(1, .1, .1, .1, .1)$. Table 5.3 shows the number of iterations required by the modified Gauss-Newton method.

Table 5.3: Number of Iterations for Example 2						
direction	(1,1,1,1,1)	(1,.1,.1,.1,.1)	(.1,1,.1,.1,.1)	(.1,.1,1,.1,.1)	(.1,.1,.1,1,.1)	(.1,.1,.1,.1,1)
α	15	160	290	310	70	10
iteration	9	10	15	17	8	8

Fig. 5.4 Simulated Measurements for Example 2



5.6.3 Example 3

This describes a reversible one-enzyme system. It was used by Garfinkel et al. (1966) and Enright and Hull (1976) as a test problem for stiff differential equation solvers. The model equations are,

$$\begin{aligned}
 dx_1/dt &= -k_1x_1x_2 + (k_2+k_3)x_3 - k_4x_1x_4; & x_1(0) &= 3.365 \times 10^{-7} \\
 dx_2/dt &= -k_1x_1x_2 + k_2x_3; & x_2(0) &= 8.261 \times 10^{-3} \\
 dx_3/dt &= k_1x_1x_2 - (k_2+k_3)x_3 + k_4x_1x_4; & x_3(0) &= 9.38 \times 10^{-6} \\
 dx_4/dt &= -k_4x_1x_2 + k_3x_3; & x_4(0) &= 1.642 \times 10^{-3}
 \end{aligned}
 \tag{5.20}$$

There are four state variables, of which only x_1 , x_2 are observed. The parameter values are assumed to be $k_1 = 3.0 \times 10^5$, $k_2 = 2.0 \times 10^1$, $k_3 = 1.0 \times 10^2$, and $k_4 = 9.0 \times 10^5$. Using these k values, simulated data was generated by solving the differential equations. The data is shown in Figure 5.5 and used as the observed data to match. Again, the values of k were varied in sequence and α is the maximum step length in each of the directions. The iterations needed to converge are shown in Table 5.4.

Fig. 5.5 - Simulated Measurements for Example 3

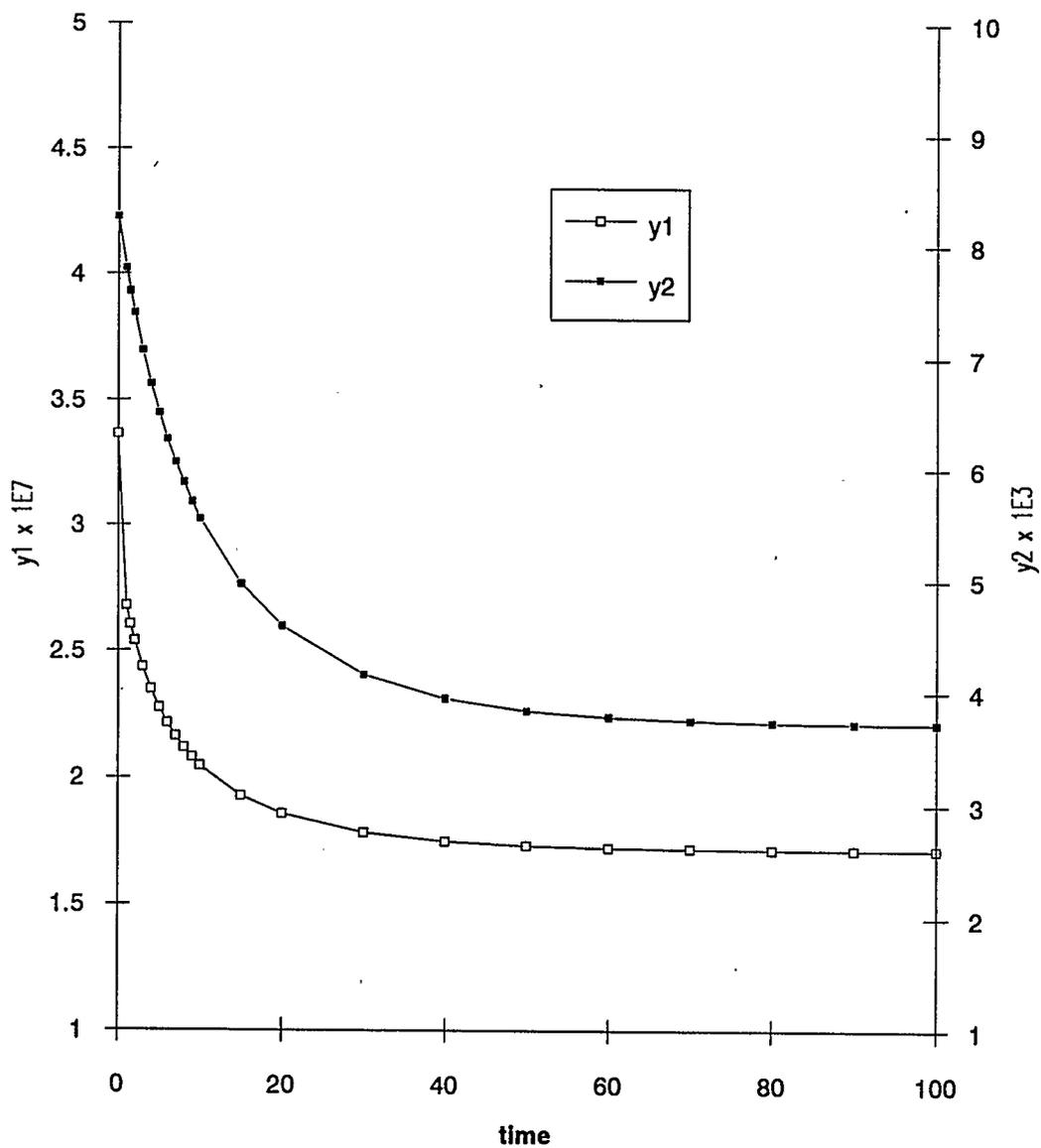


Table 5.4: Number of Iterations for Example 3					
direction	(1,1,1,1)	(1,.1,.1,.1)	(.1,1,.1,.1)	(.1,.1,1,.1)	(.1,.1,.1,1)
α	90	110	40	190	30
iterations	10	13	12	11	8

5.6.4 Example 4

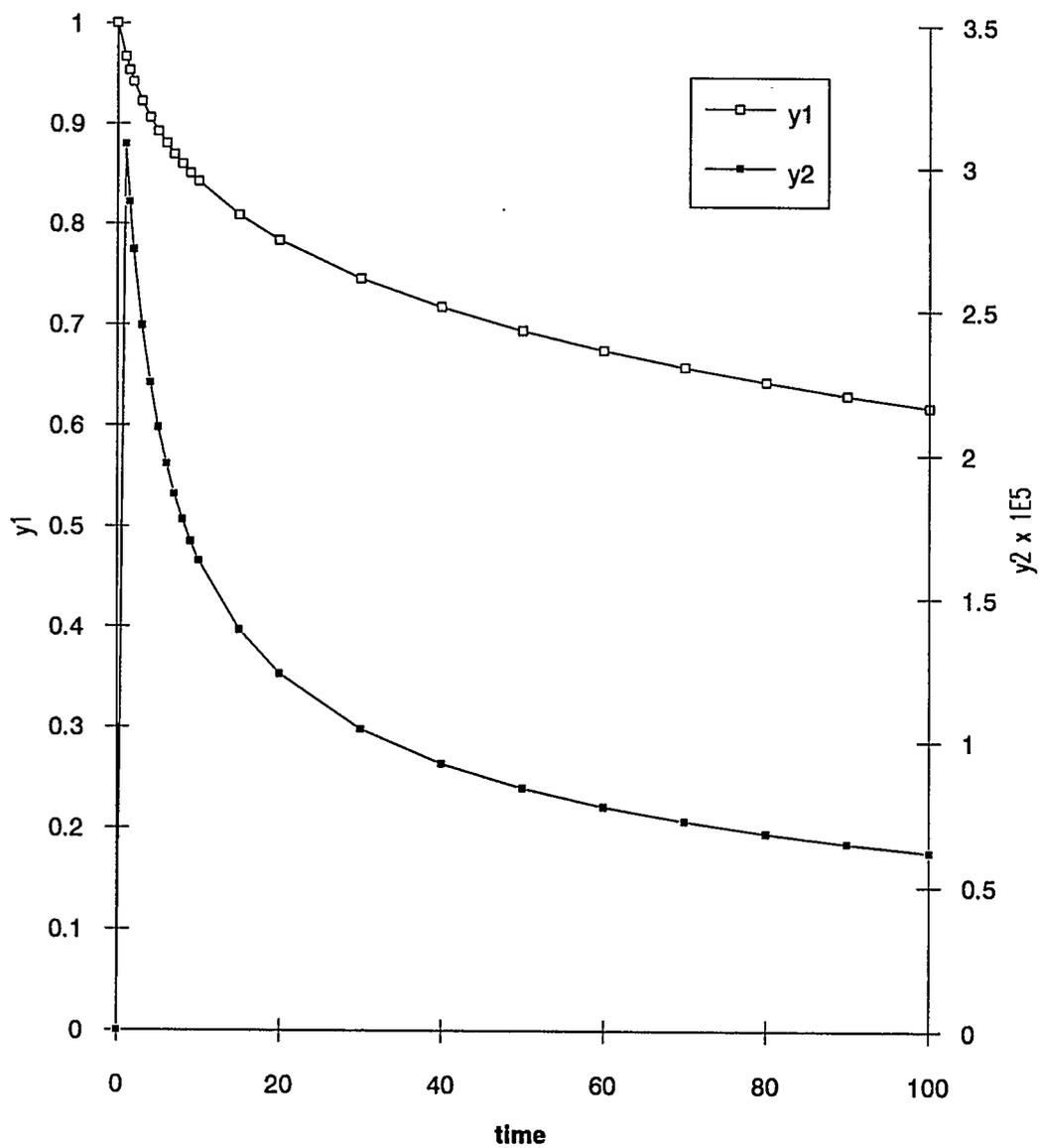
As a final test problem, the reaction rate expressions introduced by Robertson (1967) are considered. These have also been used to evaluate stiff differential equation solvers.

$$\begin{aligned} dx_1/dt &= -k_1x_1+k_2x_2(1-x_1-x_2); & x_1(0) &= 1. \\ dx_2/dt &= k_1x_1-k_2x_2(1-x_1-x_2)-k_3x_2^2; & x_2(0) &= 0. \end{aligned} \tag{5.21}$$

There are two state variables x_1 and x_2 , both of which are measured. The values of the rate constants, $k_1 = 0.04$, $k_2 = 1.0 \times 10^4$, $k_3 = 3.0 \times 10^7$, were used to generate simulated data as shown in Figure 5.6. Note that the initial value of y_2 , the simulated value of x_2 , is 0 at $t = 0$. The number of Gauss-Newton iterations required for convergence are shown in Table 5.5 for a number of different initial parameter values.

Table 5.5: Number of Iterations for Example 4				
direction	(1,1,1)	(1,.1,.1)	(.1,1,.1)	(.1,.1,1)
α	110	20	5	850
iterations	14	13	5	19

Fig. 5.6 - Simulated Measurements for Example 4



5.7 Discussion

Since convergence was obtained for the same α as tested by Kalogerakis and Luus (1983), the numerical experimentation has shown that the proposed modifications to the Gauss-Newton method have not reduced the region of convergence of the method. As the number of iterations required to converge to the optimum is very reasonable it can be inferred that the convergence characteristics of the Gauss-Newton method has not been impaired. In the next chapter, we will apply the method to a single-phase reservoir model.

CHAPTER 6

APPLICATION TO A SINGLE-PHASE MODEL

The next research objective was to incorporate the computationally efficient Gauss-Newton algorithm into a single-phase compressible flow model in order to investigate its applicability to the models typical of reservoir simulation. This chapter describes the testing of the resulting model with two single-phase problems reported by Coats et al. (1970) and Thomas et al. (1972) in their work on automatic history matching models.

6.1 Single-Phase Compressible Flow Model

The general compositional model developed in Chapter 3 can be considerably simplified by making the following assumptions:

- (a) there is only a single phase in the reservoir,
- (b) the single phase consists only of one component,
- (c) there is no gravity effect on flow,
- (d) the source/sink term is constant,
- (e) the viscosity is constant.

With these assumptions, the resulting partial differential equation is,

$$\nabla \cdot \frac{K}{\mu} \nabla P + q = \frac{\partial}{\partial t} (\phi \rho)$$

(6.1)

This can be reduced by grid discretization, for a grid of n cells, to a set of n ordinary differential equations, written as

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{k}; \mathbf{u}) \quad (6.2)$$

with initial conditions,

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (6.3)$$

where \mathbf{x} is an n dimensional vector of grid cell pressures - the state variables, \mathbf{x}_0 is the given initial state, \mathbf{k} is the p dimensional vector of porosities or permeabilities to be estimated, and \mathbf{u} is a vector of user specified parameters, such as flowrates.

This is analogous to the model equation described in Chapter 5 in the development of the Gauss-Newton method. Thus the theory presented in Chapter 5 is readily applicable to the single-phase compressible flow model.

6.2 Performance Evaluation

6.2.1 Single-Phase Gas Reservoir

This example was taken from Coats et al. (1970). It represents a gas reservoir with an 8 x 8 computing grid. The reservoir description is given in Table 6.1. The reservoir is divided into 6 zones of constant permeability as shown in

Table 6.1: Single-Phase Gas Reservoir Problem Description

Reservoir Dimensions	20000 x 20000 feet
Grid Cells	8 x 8
Grid Dimensions	2500 x 2500 feet
Thickness	25 feet
Porosity	0.15
Initial Pressure	2000 psi

Figure 6.1. The observed data consists of the pressures at eight observation wells with locations shown in Figure 6.1, measured at 360 days using the true zone permeabilities. The timestep used in the run was a constant 30 days.

The parameters to be estimated were the permeabilities of the six zones. As initial estimates of the true zone permeabilities, a constant value of 0.5 md was used in all zones. The iteration history of the six zone permeabilities are shown in Table 6.2 and in Figure 6.2. In seven iterations of the Gauss-Newton method, the actual zone permeabilities were recovered. The convergence was rapid, indicating the stability of the proposed Gauss-Newton modifications, that is, sequential solution of the model and sensitivity equations.

The cputime required for the evaluation of the model equations was 3.74 secs on the computer utilized. The cputime required for the evaluation of the model and sensitivity equations for one Gauss-Newton iteration was 5.83 secs. This is the equivalent of 1.56 model runs.

Previous implementations of the Gauss-Newton method would have required 7 model runs for each Gauss-Newton iteration for this 6 parameter problem. Thus there is a significant decrease in computational effort compared with previous implementations of the Gauss-Newton method.

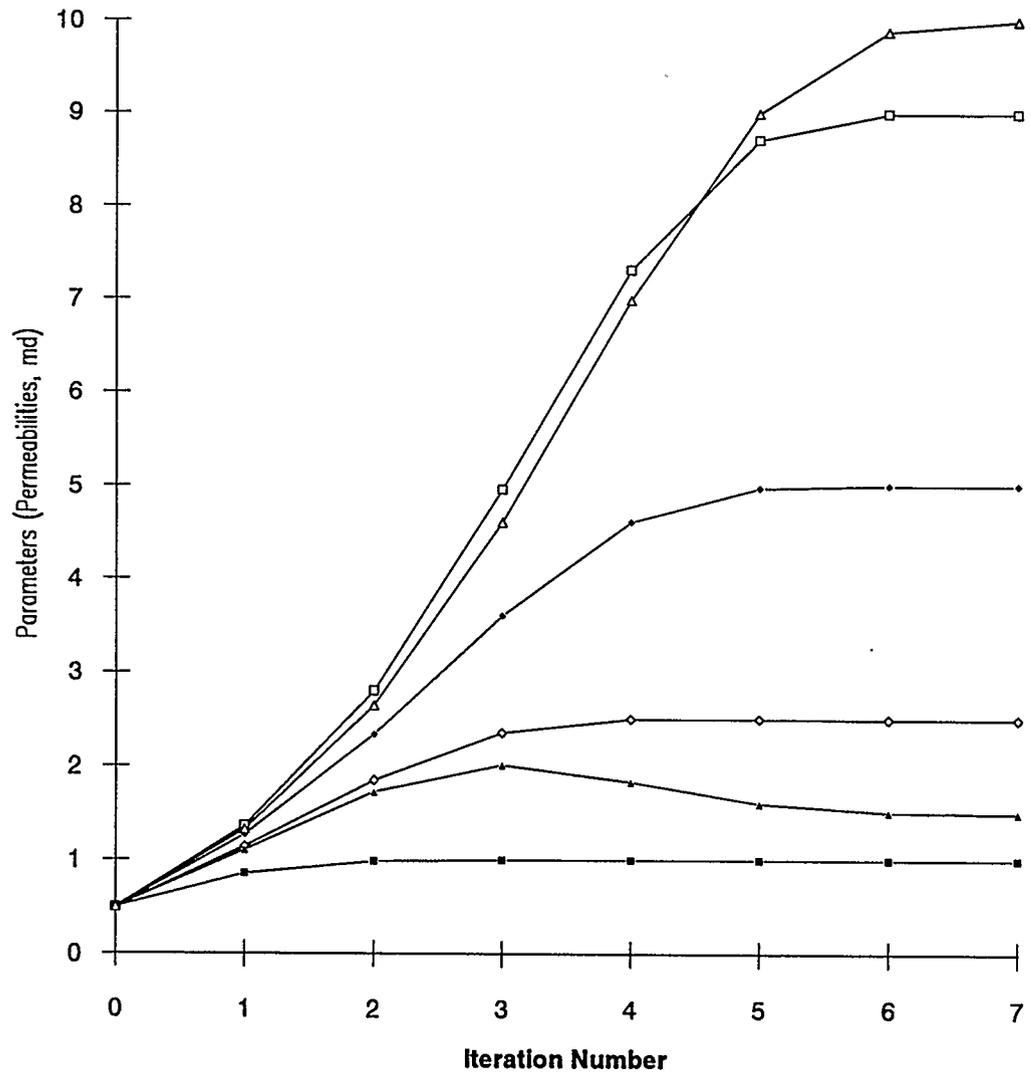
Coats' linear programming algorithm required 30 runs made with parameters generated using a random number generator.

Thomas' algorithm required 25 runs to estimate the parameters to the same accuracy. A first-order optimal control method would have required a minimum of 2 equivalent model runs per iteration. Since the Gauss-Newton is a second-order convergent method, its convergence behaviour is much better than the first-order optimal control method. This implies the proposed implementation of the Gauss-Newton method will take fewer iterations to obtain a parameter estimate than the first-order optimal control method and also require less computer time for each iteration for this particular example.

The case was repeated with 8 zones instead of 6 by subdividing zones 1 and 6. The cputime required for the evaluation of the model and sensitivity equations was 6.51 secs or the equivalent of 1.74 model runs. This compares with the 9 equivalent model runs that would have been required using previous implementations of the Gauss-Newton method.

Iter	LS Value	k1	k2	k3	k4	k5	k6
0	313324	0.5000	0.500	0.500	0.500	0.500	0.500
1	58449	0.8559	1.364	1.271	1.146	1.108	1.329
2	10107	0.9870	2.800	2.339	1.848	1.726	2.645
3	1384.5	0.9994	4.954	3.599	2.355	2.013	4.601
4	115.23	1.0000	7.315	4.602	2.505	1.836	6.988
5	3.4723	1.0000	8.715	4.970	2.505	1.605	8.996
6	0.0230	1.0000	8.992	5.001	2.501	1.512	9.881
7	0.000002	1.0000	9.000	5.000	2.500	1.500	9.999

Fig. 6.2 - Estimates of Parameter Values at Each Iteration in Single-Phase Reservoir Run



6.2.2 Five-Spot Pilot Example

This example is similar to the problem described by Jahns (1966) and later by Thomas et al. (1972). A 20 x 20 grid is used to describe an isolated five-spot in a large reservoir. There are eight zones of constant permeability as shown in Figure 6.3. The full reservoir description is given in Table 6.3. A constant production rate of 3000 bbl/d is equally divided among the four central wells. The observation data consists of the drawdown pressures at each of the four observation wells measured every hour over a twenty hour period using the true zone permeabilities. The timestep size used for the model integration was 1 hour.

The Gauss-Newton iteration process was started with an initial estimate of 500 md. for all eight zones. The values of the individual zone permeabilities at each iteration are shown in Table 6.4 and graphed in Figure 6.4. The true permeabilities of the eight zones were recovered in seven iterations. No stability problems were observed.

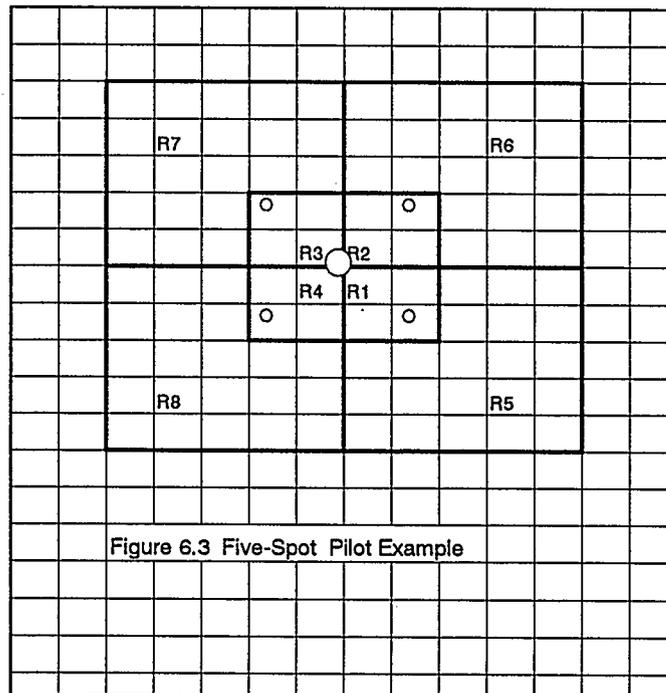
The cputime required for one model run was 195.22 secs. To evaluate the model and sensitivity equations for one iteration of the Gauss-Newton method, the cputime required was 243.47 secs or 1.25 equivalent model runs. This compares very favourably with the nine equivalent model runs that would be required by the normal implementation of the Gauss-Newton method.

The results also show that the larger the reservoir

model, the more significant will be the cost savings of the proposed implementation in terms of equivalent model runs. This is because the proportion of time spent in inverting the Jacobian of the model will increase as the number of equations increase.

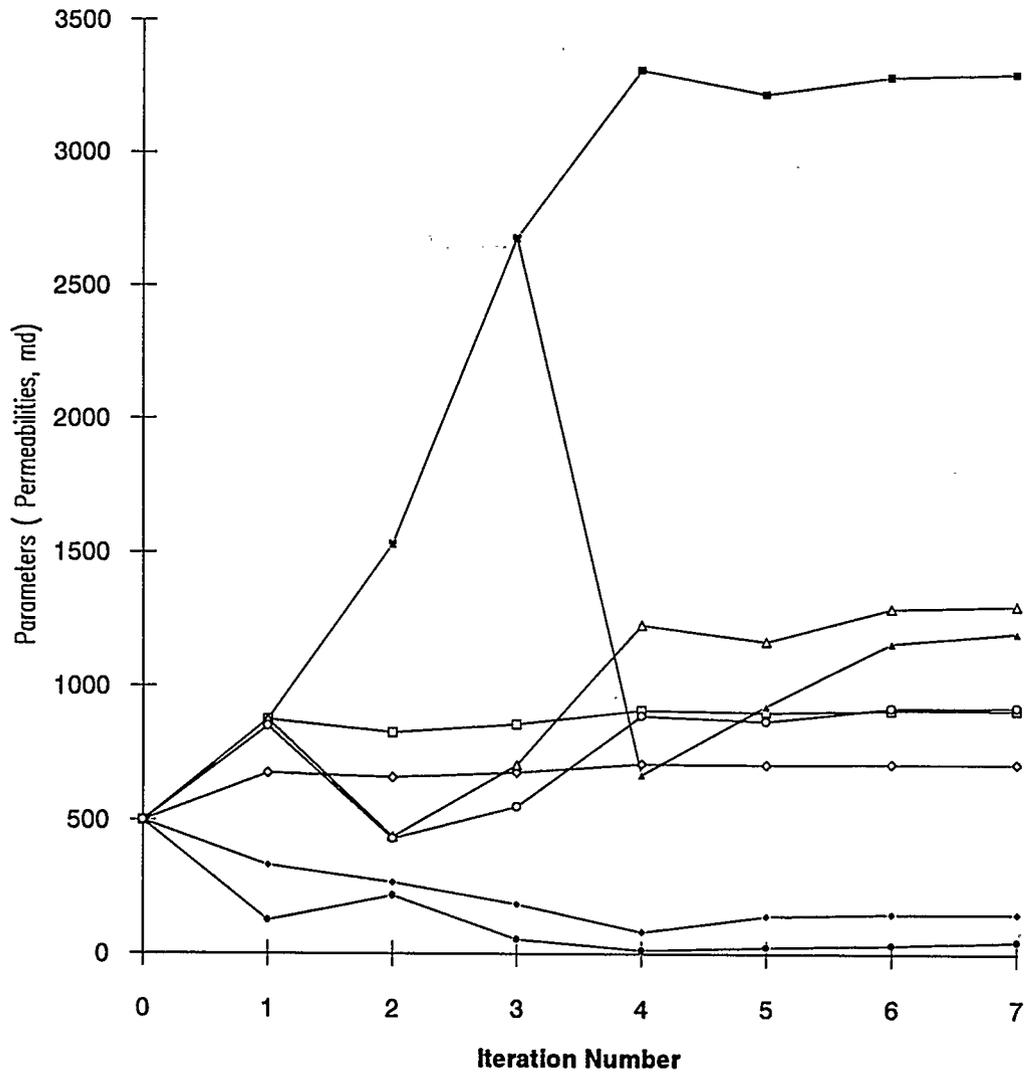
Table: 6.3 Five-Spot Pilot Example Problem Description

Reservoir Dimensions	1800 x 1800 feet
Grid Cells	20 x 20
Grid Dimensions	90 x 90 feet
Thickness	Region 1 Region 2,3,6,7 Region 4,8
	6 ft 8.5 ft 7.5 ft
Porosity	0.16
Compressibility	4.66×10^{-4} v/v/psi
Initial Pressure	2000 psi
Producing Well Locations	(10,10), (11,10), (10,11), (11,11)
Producing Well Rates	750 bbl/d
Observation Well Locations	(6,6), (15,6), (6,15), (15,15)



Iter	LS	k1	k2	k3	k4	k5	k6	k7	k8
0	1121	500	500	500	500	500	500	500	500
1	734.9	875	875	332.9	676.1	875	875	125	850.2
2	366.8	1531	825.9	267	660.3	1531	439	218.8	431.4
3	113.9	2680	857.9	186.1	678.2	2680	705.2	54.7	551.4
4	28	3313	910.1	81.1	709.7	669.9	1228	13.67	887.6
5	2.712	3222	901.5	141.9	706.9	924.9	1167	23.93	868.7
6	0.0321	3287	909.8	149.5	709.9	1160	1290	32.15	916.8
7	0.00001	3300	910	150	710	1199	1300	45.55	919.7

**Fig. 6.4 - Estimates of Parameter Values at each Iteration
in Five-Spot Pilot Example**

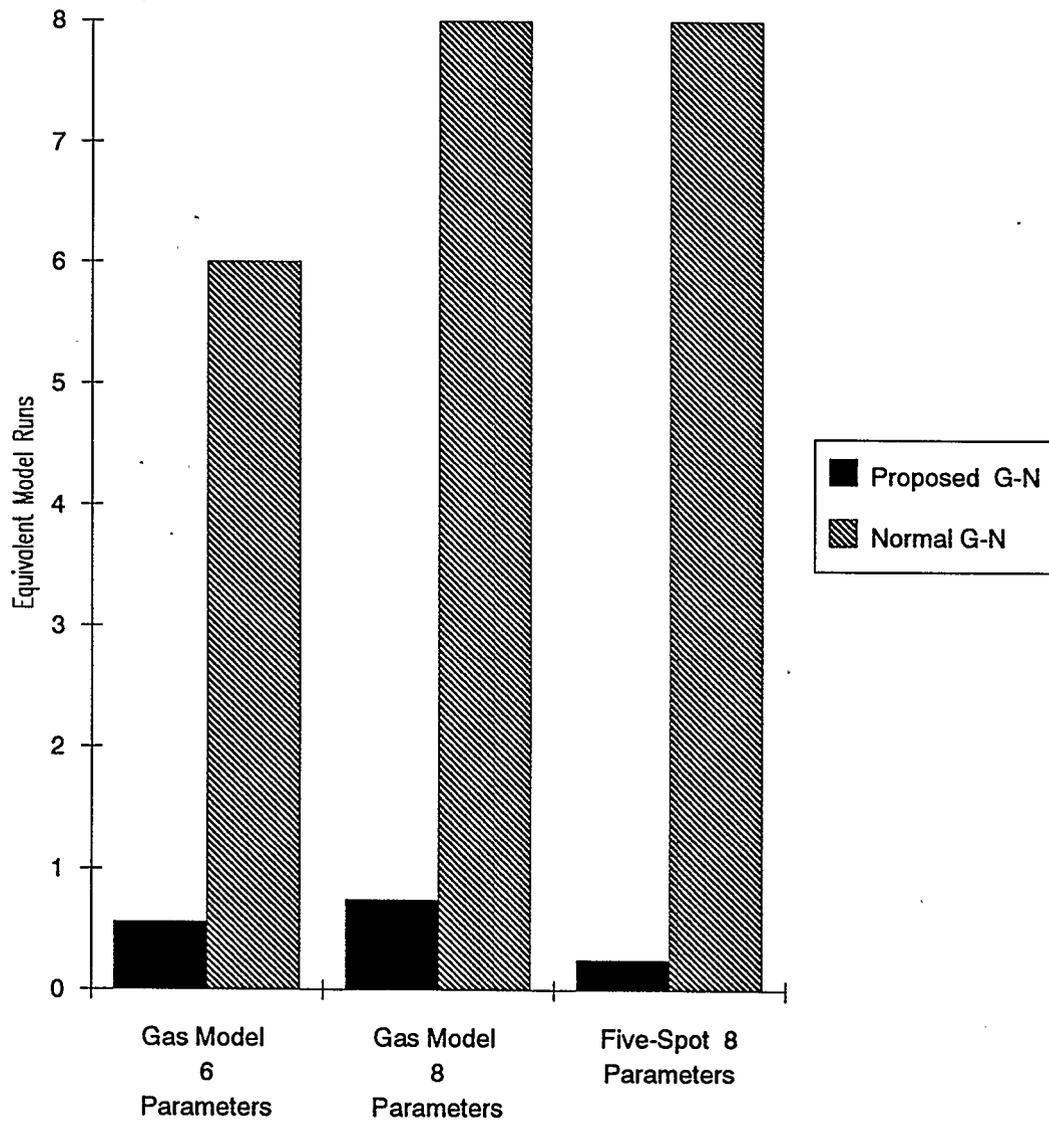


6.3 Discussion

Table 6.5 summarizes the results of the two numerical examples. The magnitude of the differences are graphically illustrated in Figure 6.5. It is clear that the proposed modifications to the Gauss-Newton method has reduced significantly the computational effort for each iteration of the Gauss-Newton method even for this relatively simple single-phase reservoir model. With this milestone reached, the next objective is the incorporation of the method in a three-dimensional three-phase reservoir model.

Problem	No of Parameters	Proposed Gauss-Newton	Standard Gauss-Newton
Gas Model	6	0.56	6
Gas Model	8	0.74	8
Five-Spot	8	0.25	8

Fig. 6.5 - Equivalent Model Runs to Compute Sensitivity Coefficients



CHAPTER 7

APPLICATION TO A THREE-DIMENSIONAL THREE-PHASE MODEL

In this chapter, the modified Gauss-Newton method is applied to a fully implicit three-dimensional three-phase limited compositional simulator. The model with its automatic history matching capabilities is described. A numerically difficult three-phase coning problem was used to test out the model.

7.1 Model Equations

The model consists of a set of ordinary differential equations which describe the molar balances of each component. The phase behaviour is represented by K values and any component can exist in any phase. The model is presented succinctly in difference notation.

Assuming there are n_p phases and n_c components, the equation component balance, for $I = 1$ to n_c , is

$$\sum_{p=1}^{n_p} \Delta \left[T \frac{k_{IP}}{\mu_p} \rho_p x_{IP} (\Delta P + \Delta P_{cp} - \gamma_p \Delta z) \right] - \frac{V}{\Delta t} \delta \left(\phi \sum_{p=1}^{n_p} \rho_p S_p x_{IP} \right) + q_I \quad (7.1)$$

A set of algebraic constraints completes the model. These are:

Saturation constraint equation

$$\sum_{p=1}^{n_p} S_p = 1 \quad (7.2)$$

n_p mole fraction constraint equations, $p = 1, n_p$

$$\sum_{I=1}^{n_c} X_{Ip} = 1 \quad (7.3)$$

Master mole fraction variables as introduced by Coats (1980) are used to reduce the number of mole fraction state variables from $n_p \times n_c$ in number to only n_c variables. The component mole fraction variables x_{Ip} in Eqs. 7.1 and 7.3 are replaced by

$$X_{Ip} = K_{Ip} X_I \quad (7.4)$$

where K_{Ip} is the equilibrium K value of component I in phase p, and X_I is the master mole fraction for component I.

The set of unknown state variables for each grid cell for n_c components in a three phase system would then be $P, X_1, X_2, \dots, X_{n_c}, S_o, S_w, S_g$.

The numerical integration method used to solve the model equations is the implicit method. Eqs. 7.1 to 7.4 are differentiated in terms of the set of unknown variables using

the chain rule. The resulting Jacobian is used to calculate the unknown variables using Newton's iteration scheme.

7.2 Gauss-Newton Method

The Gauss-Newton method requires the evaluation of the sensitivity coefficients. The matrix of sensitivity coefficients, \mathbf{G} , is obtained by differentiating the model equations with respect to the parameter vector \mathbf{k} . For the mole balance equation for component I ,

$$\sum_{p=1}^{n_p} \frac{\partial}{\partial k} \left\{ \Delta \left[T \frac{k_{Ip}}{\mu_p} \rho_p x_{Ip} (\Delta p + \Delta p_{cp} - \gamma_p \Delta z) \right] \right\} - \frac{V}{\Delta t} \frac{\partial}{\partial k} \left\{ \delta \left(\phi \sum_{p=1}^{n_p} \rho_p S_p x_{Ip} \right) \right\} + \frac{\partial q_I}{\partial k} \quad (7.5)$$

Differentiation of the saturation constraint equation yields,

$$\sum_{p=1}^{n_p} \frac{\partial S_p}{\partial k} = 0 \quad (7.6)$$

Differentiation of the n_p mole fraction constraint equations yields,

$$\sum_{I=1}^{n_c} \frac{\partial x_{Ip}}{\partial k} = 0 \dots \dots \dots; \quad p=1, n_p \quad (7.7)$$

The set of unknown sensitivity coefficients for each grid cell would then be

$$\frac{\partial P}{\partial \mathbf{k}}, \frac{\partial X_1}{\partial \mathbf{k}}, \frac{\partial X_2}{\partial \mathbf{k}}, \dots, \frac{\partial X_{nc}}{\partial \mathbf{k}}, \frac{\partial S_o}{\partial \mathbf{k}}, \frac{\partial S_w}{\partial \mathbf{k}}, \frac{\partial S_g}{\partial \mathbf{k}}$$

where \mathbf{k} is the vector of parameters to be estimated. The differentiation is carried out using the chain rule. The initial conditions are given by Eq. 5.12.

The Gauss-Newton method therefore requires the solution of the set of sensitivity equations given by Eqs. 7.5 to 7.7 for the matrix of sensitivity coefficients. These equations are solved in conjunction with the model equations. During each time step, the model equations are solved to obtain the values of the state variables $P, X_1, X_2, \dots, X_{nc}, S_o, S_w, S_g$. These values are used in the sensitivity equations to solve for the sensitivity coefficients.

$$\frac{\partial P}{\partial \mathbf{k}}, \frac{\partial X_1}{\partial \mathbf{k}}, \frac{\partial X_2}{\partial \mathbf{k}}, \dots, \frac{\partial X_{nc}}{\partial \mathbf{k}}, \frac{\partial S_o}{\partial \mathbf{k}}, \frac{\partial S_w}{\partial \mathbf{k}}, \frac{\partial S_g}{\partial \mathbf{k}}$$

At the observation times, the sensitivity coefficients are substituted into,

$$\left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} \mathbf{C}(t_i) \mathbf{G}(t_i) \right] \Delta \mathbf{k}^{(j+1)} - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} (\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)) \quad (7.8)$$

if the measured variables are linearly related to the state variables,

$$\mathbf{y}(t) = \mathbf{C} \mathbf{x}(t) \quad (7.9)$$

or into,

$$\begin{aligned} & \left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right) \mathbf{Q} \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right)^T \mathbf{G}(t_i) \right] \Delta \mathbf{k}^{(j+1)} \\ & - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \left(\frac{\partial \mathbf{h}^T}{\partial \mathbf{x}} \right) \mathbf{Q} (\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)) \end{aligned} \quad (7.10)$$

if the observation relationship is nonlinear of the form of,

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)) \quad (7.11)$$

Finally at the end of the history period, Eq. 7.8 or 7.10 is solved to yield the next estimate of the parameter vector.

The Gauss-Newton method has been implemented in an efficient manner, approximating the sensitivity equations and utilizing information from the solution of the model equations at each timestep in order to reduce the computational work to obtain the values of the sensitivity coefficients.

7.3 Model Features and Operation

The parameters that can be estimated by the automatic history matching model are permeability and porosity. The procedure is to first establish regions encompassing one or

more grid cells for which parameter values are to be determined. Each region has a constant parameter value. An initial estimate of the parameter value must be provided. The observed variables that can be matched are the pressures of the grid cells, producing water-oil ratios and gas-oil ratios of the individual wells, and flowing bottom hole pressures. The observed data can be weighted individually in the objective function and can be sparse or frequently available throughout the reservoir history. The weighting factors for any variable can change with time. An automatic timestep selector adjusts the timesteps ensuring that the simulator calculates values for the observed variables at the corresponding observation times.

7.3.1 Cautious Step Size Policy

The normal procedure for the Gauss-Newton method is, given an estimate of the unknown parameters, to make a complete pass through the history with the model to obtain the value of the objective function. Then the sensitivity equations are evaluated and this usually requires solution of the model and sensitivity equations simultaneously albeit sequentially for each timestep over the history period. This is because the integration of the sensitivity equations over the history requires that values of the model state variables be available and it is not practical to store these at each timestep. The Gauss-Newton matrix is generated from the

sensitivity coefficients and it is solved to provide a new estimate of the parameter vector. The new estimate is usually not accepted immediately and a model run is performed with the new estimated parameters to obtain the new value of the objective function. If the new value is greater than the previous objective function value, the step length is adjusted either using quadratic interpolation down the step direction as suggested by Thomas et al. (1972) or using the bisection rule to find a step length that provides an improved estimate of the objective function. In either case, further model evaluations are required before the new estimate of the parameter vector is accepted. The sensitivity coefficients for the next iteration of the Gauss-Newton method are obtained by solving the model and sensitivity equations again simultaneously. It can therefore be noted that each iteration of the Gauss-Newton method requires the minimum of two identical model evaluations together with the sensitivity equation evaluation if the Gauss-Newton step is immediately successful.

In this implementation of the Gauss-Newton method, it was found that the work required to evaluate the sensitivity coefficients was not significantly more than the work required for each model evaluation. As a result the two or more model evaluations required to evaluate each step change for each Gauss-Newton iteration was a major portion of the computational effort. It was therefore decided to accept each

new estimate of the parameter vector without requiring that the new value of the objective function be less than the previous value. In order to avoid but not prevent overstepping, a cautious stepsize policy was adopted with respect to the maximum change of each parameter during each iteration. The step changes of each parameter are normalized and maximum step change limitations can be imposed by the user so that the new estimate of the parameter vector is obtained along the direction found by the Gauss-Newton method. In summary, in this implementation, each iteration of the Gauss-Newton method consists of the simultaneous solution of the model and sensitivity equations followed by the solution of the Gauss-Newton matrix.

7.3.2 Pseudo-Inverse Option

An option available in the program is the use of the pseudo-inverse method as described by Lawson and Hanson (1974). This method computes the solution vector by a procedure that avoids the instability that can occur when a very ill-conditioned problem is treated as full rank. This may happen when the parameters being sought for are not identifiable from the observed data. For example, a parameter zone may be outside the drainage radius of a well and hence is not observable from the data measurements at the well. The technique replaces the Gauss-Newton matrix by a rank deficient matrix and then computes the minimal length solution to the

problem. In this implementation, the rank of the pseudo-inverse is increased whenever the estimated change in the parameters is less than a certain tolerance.

7.3.3 Implicit Formulation

The choice of model formulation is also important for the success of the automatic history matching algorithm. Kalogerakis and Luus (1983) have shown the importance of using a robust integration routine for parameter estimation in complex chemical systems in order to increase the region of convergence. Similarly in reservoir simulation, a fully implicit formulation for the reservoir model ensures stability and convergence for a wide range of parameter estimates. It is essential that the model itself be well constructed, robust and stable with sufficient provisions for timestep cuts and repeats and possess an effective iterative matrix solution algorithm, prior to development of the automatic history matching algorithm.

7.4 Application of the Automatic History Matching Model

The resulting model was applied to the Second SPE Comparative Solution Problem (Chappelear and Nolen, 1986). This is a radial coning problem. The reservoir grid consists of ten concentric rings and fifteen layers. A gas cap, an oil zone and a water zone are all present. Both water and gas coning occurs. Over a short time frame, severe changes in

production rate were imposed causing rapid changes in pressure and phase saturations as well as phase appearance and phase disappearance. Full details of the model parameters, fluid properties and other necessary data can be found in the reference and will not be repeated here. The problem itself is a difficult one and combining its solution with the parameter estimation problem is a good test of the stability and effectiveness of the automatic history matching model. The model was used to generate artificial observation data using the original values of the reservoir parameters. Then starting with an arbitrary initial guess of the reservoir parameters, it was attempted to recover the original data by matching the observed data.

7.4.1 Matching Reservoir Pressure

The reservoir pressure is a state variable in the model equations and thus the relationship between the observed variables and the state variables is a linear one of the form of Eq. 7.9.

The matrix C consists of a sparse matrix with $C_{ij} = 1$ for the grid cell locations where the pressure measurements are taken.

The SPE Comparison Problem 2 has fifteen layers of constant permeability and porosity. These layers were taken as the reservoir zones for which parameters are to be estimated. The original reservoir description was used to

generate the observed data. With automatic timestep selection, the model required 16 timesteps to complete the 900 day problem. The cputime required for this base model run was 166 secs. This value is used to calculate equivalent model runs. At the end of each timestep, the pressures of the first column of cells were recorded. This data was used as the observed data. As weighting matrix, Q , the identity matrix was used.

In the first test, Run 1, layers 3 to 12 were selected as the zones whose parameters were to be estimated. The horizontal permeabilities of these ten zones were the unknown parameters. The initial guess of the permeabilities was a uniform value of 200 md. The Gauss-Newton method recovered the original permeability values of the ten zones in 9 iterations. The objective function value decreased from 0.911×10^6 to 0.306×10^{-2} . Figure 7.1 shows the change in the objective function value with iteration number. The estimated values of the ten zone permeabilities are given in Table 7.1. The deviations of these values from the original values, also given in Table 7.1, are very small. Figure 7.2 shows the initial pressure profile of cell (1,1,8) and the final match of the observed data. The pressure profiles for the second, third and fourth iterations are also plotted. The iteration history of the ten zone permeabilities is shown in Figure 7.3.

The cputime required for the 9 iterations was 3238 secs, equivalent to 19.51 model runs. The average cputime for each

Table 7.1 : Match of Original Permeabilities						
Layer	Permeability(md)	Run 1	Run 2	Run 6	Run 12	Run 13
1	35.0		34.98			
2	47.5		47.54			
3	148.	148.0	148.0			
4	202.	202.1	202.0			
5	90.	90.0	90.0			89.64
6	418.5	418.5	418.5	433.7	418.5	418.3
7	775.	775.0	775.0	761.4	774.8	775.3
8	60.	60.0	60.0	65.0	60.0	60.0
9	682.	681.7	682.0	627.6	681.7	686.6
10	472.	472.3	471.8			476.8
11	125.	125.1	125.2			
12	300.	299.8	299.7			
13	137.5		137.4			
14	191.		191.4			
15	350.		349.3			

Table 7.2: Match of Original Porosities					
Layer	Porosity	Run 3	Run 4	Run 7	Run 14
1	0.087		0.0867		
2	0.097		0.0982		
3	0.111	0.1109	0.1105		
4	0.160	0.1601	0.1610		
5	0.130	0.1299	0.1298		0.1297
6	0.170	0.1704	0.1700	0.1660	0.1699
7	0.170	0.1700	0.1697	0.1800	0.1718
8	0.080	0.0800	0.0794	0.0747	0.0788
9	0.140	0.1402	0.1421	0.1406	0.1420
10	0.130	0.1297	0.1283		0.1273
11	0.120	0.1212	0.1202		
12	0.105	0.1031	0.1003		
13	0.120		0.1313		
14	0.116		0.1554		
15	0.157		0.1335		

**Fig. 7.1 - Reduction of Objective Function with Iteration
in Run No. 1 (Match of Reservoir Pressures)**

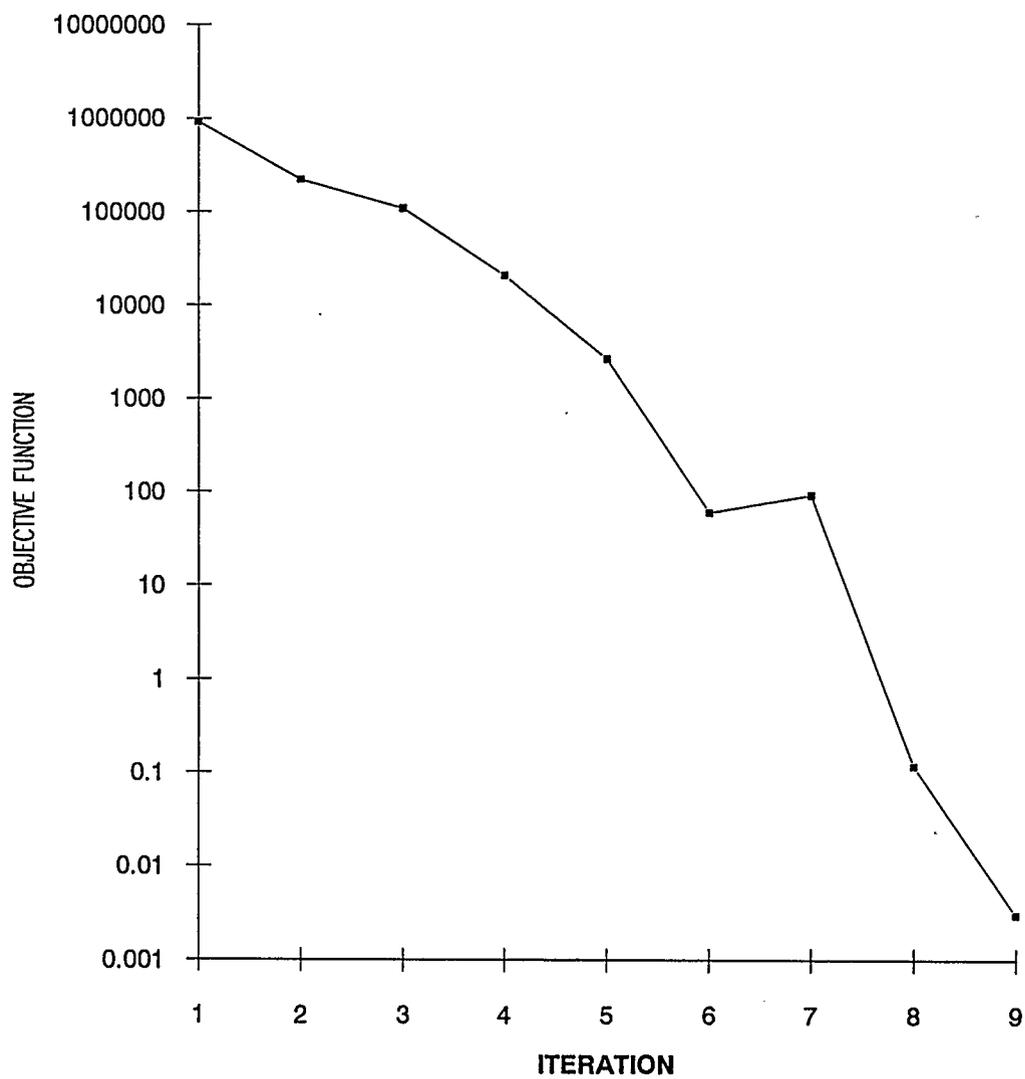


Fig. 7.2 - Match of Pressure Profile of Cell (1,1,8) in Run No 1.

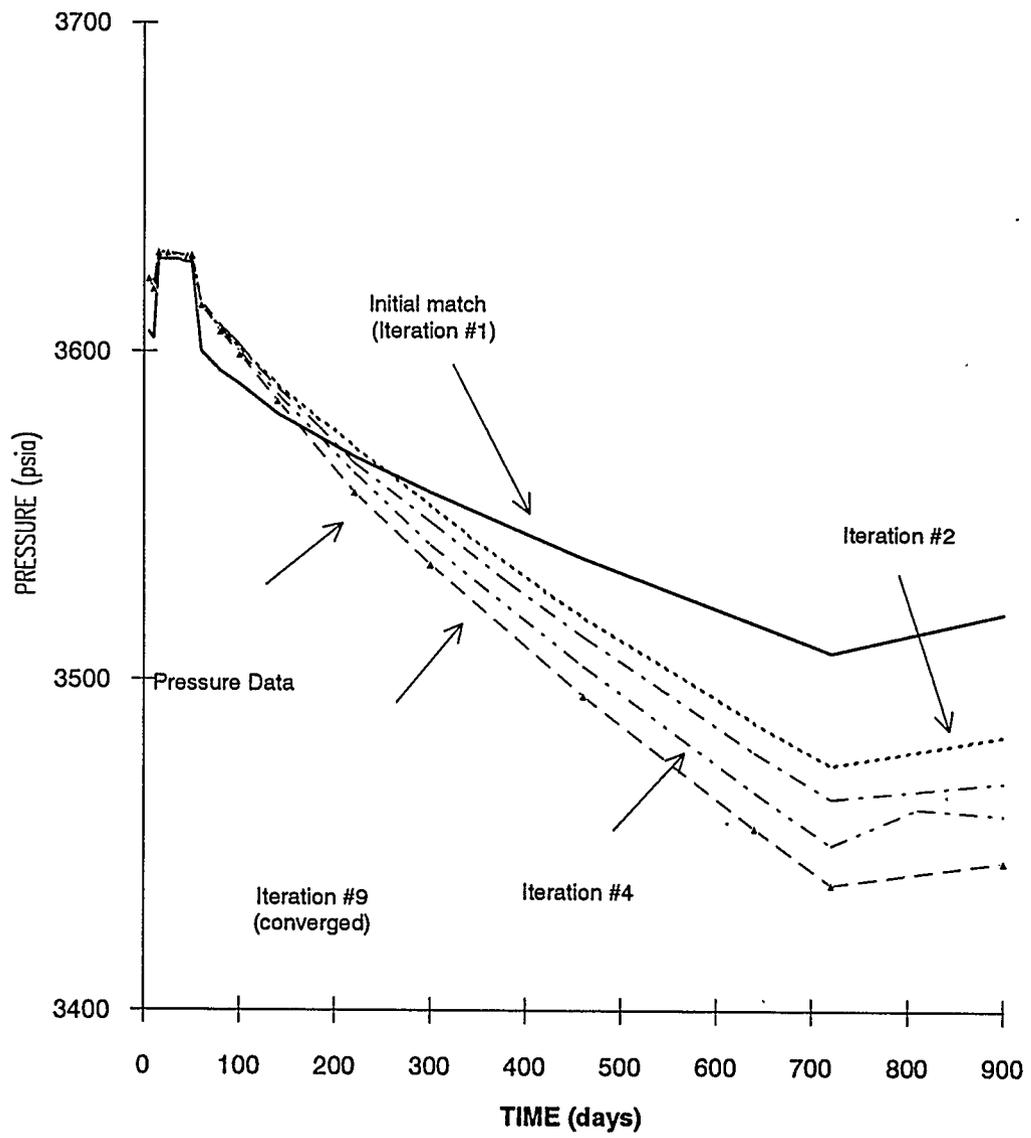
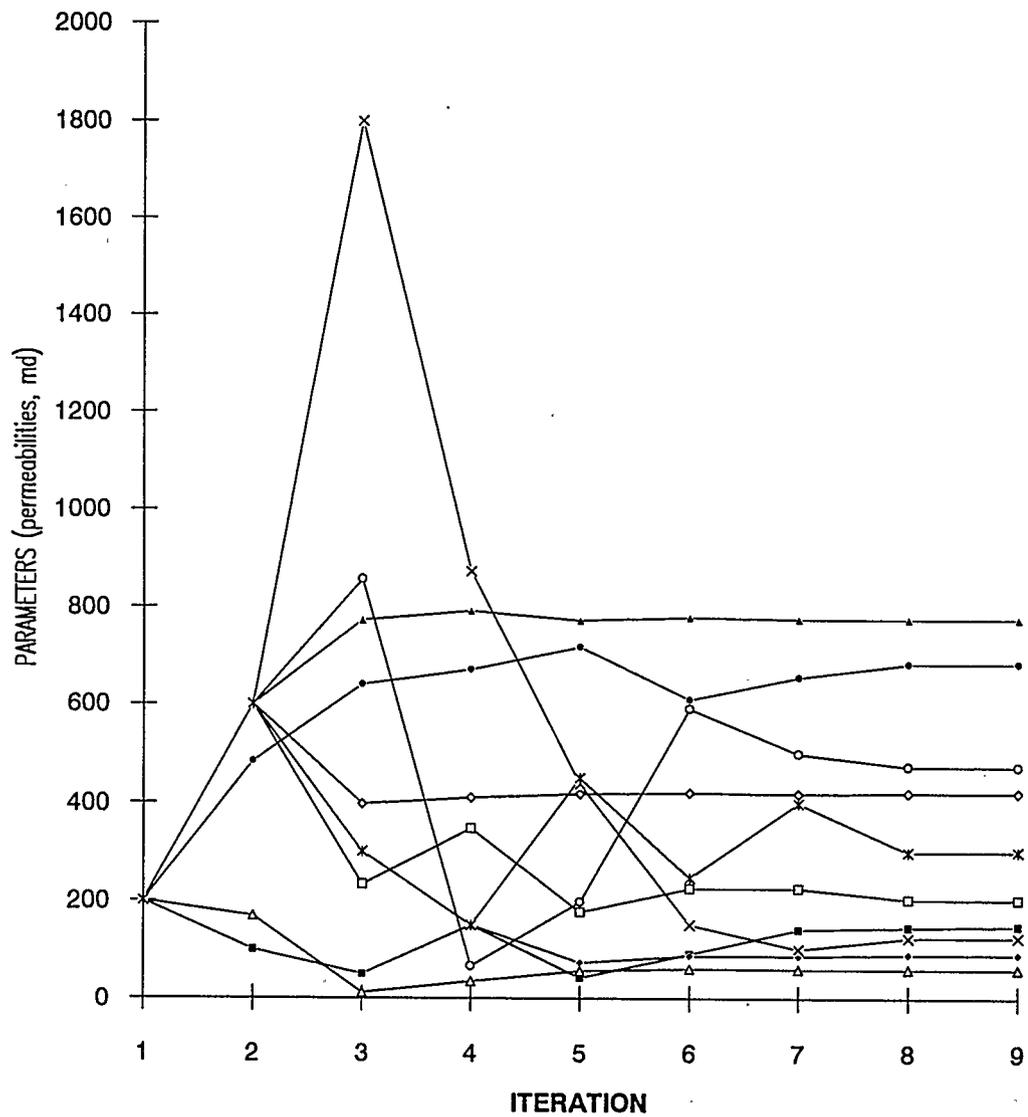


Fig 7.3 - Estimates of Parameter Values at Each Iteration in Run No. 1



Gauss-Newton iteration was therefore 2.17 model runs. This is the time required for one model run and the evaluation of the sensitivity coefficients for each estimate of the parameters. The time required for the evaluation of the sensitivity coefficients for the ten parameters was therefore 1.17 model runs. This compares very favourably to the 10 equivalent model runs, one for each parameter, that would have been required by normal numerical differentiation for each Gauss-Newton iteration. This result represents an order of magnitude reduction in computational requirement and makes automatic history matching through nonlinear regression practical and economically feasible for the first time in the history of reservoir simulation.

It should also be pointed out that the Gauss-Newton method is a quadratically convergent method and as such it converges in very few iterations compared to the first order optimal control method introduced by Chen et al. (1974). This is especially true for the highly nonlinear multiphase reservoir problems. The cputime requirement for this implementation of the Gauss-Newton method is extremely competitive with the optimal control method, even on a single iteration basis. Of course, with the advantage of possessing quadratic convergence, the proposed modified Gauss-Newton method outperforms the latter method.

For the next test, Run 2, the horizontal permeability of

all fifteen layers of the model were initially set at a uniform value of 200 md. The Gauss-Newton method converged to the optimum in 12 iterations reducing the objective function value from 0.833×10^6 to 0.692×10^{-2} . The estimated values of the fifteen zone permeabilities are given in Table 7.1. Again these values are very similar to the original values. The cputime taken for the 12 iterations was 4536 secs, equivalent to 27.33 model runs, or an average of 2.28 model runs for each iteration. Thus the time required for the sensitivity coefficient evaluation for the fifteen parameters was 1.28 model runs. Deriving the sensitivity coefficients by previously published implementations of the Gauss-Newton method would have required 15 equivalent model runs. The comparison is again extremely favourable. This case also shows that for an increase of 5 parameters, the equivalent number of runs per iteration increased from 1.17 to 1.28 only.

Run 3 is similar to Run 1 except that the porosity values of the ten zones were to be estimated. An estimate of 0.1 was used for the initial porosity of all the zones. The objective function was reduced from 0.391×10^5 to 0.874×10^{-2} in thirteen iterations. Table 7.2 shows the values of the estimated porosities as well as the original porosities for comparison. The cputime for the thirteen iterations was equivalent to 26.62 model runs or an average of 2.05 model runs for each Gauss-Newton iteration for this ten parameter

case.

For Run 4, the porosity values of all 15 layers were initially set at 0.1. The initial objective function value was 0.376×10^5 . Using the full rank matrix, the Gauss-Newton method was unable to converge, indicating the sensitivity coefficient matrix was extremely ill-conditioned. Using the pseudo-inverse option previously described, with an initial rank of 8, the objective function was reduced to 4.59. Further iteration reduced the objective function to 1.5635 with a rank of 11, and to 0.15874 with a rank of 13. The estimated porosities are shown in Table 7.2. Additional iterations with a higher rank of the pseudo-inverse of the Gauss-Newton matrix resulted in divergence indicating near collinearity of the eigenvectors and very poor conditioning. An average cputime of 2.41 model runs was taken for each Gauss-Newton iteration. The results show that unlike permeability, the observed data was insufficient to distinguish all fifteen values for the porosity parameter. It should be pointed out that the ill-conditioning of the problem due to limited observability can only be overcome by supplying more information to the model such as additional observed data or by reparameterization of the reservoir model.

7.4.2 Matching Water-Oil Ratio

A well can produce from several grid cells at the same

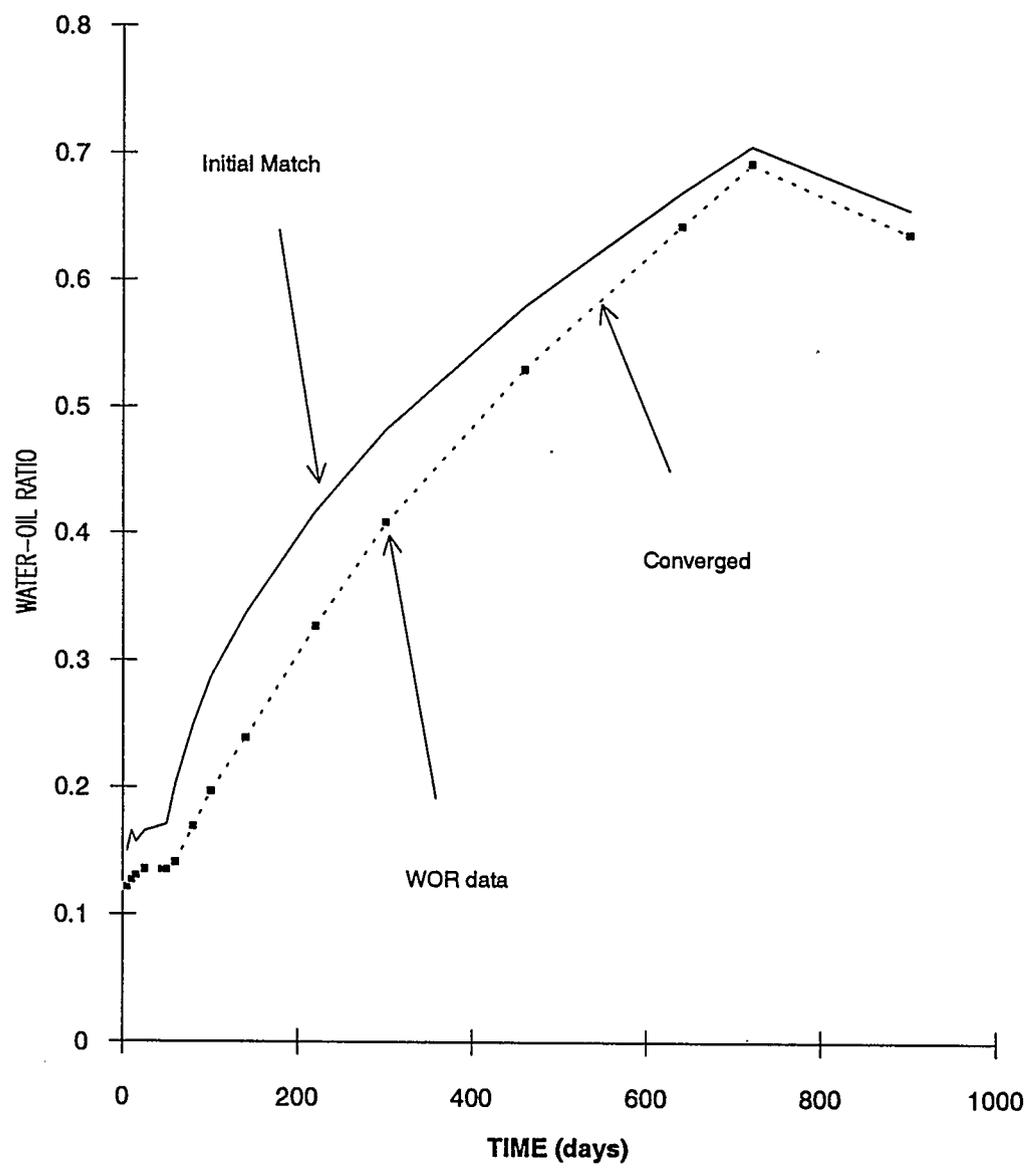
time. Thus the water-oil ratio is a complex time-dependent function of the state variables of several grid cells. The relationship between the output vector and the state vector is of the form given by Eq. 7.11.

In the SPE Problem 2, the centrally located well is completed in layers 7 and 8. From the base run utilizing the original description, the water-oil ratio was obtained at each of the 16 timesteps. This set of observations was used as the observed data to be matched in the next set of runs. Figure 7.4 shows the observed data. A diagonal matrix with elements of 1000 was used as the weighting matrix Q .

In Run 5, the horizontal permeabilities of two zones, layers 7 and 8 were estimated starting from an initial value of 200 md. With five iterations of the Gauss-Newton method, the objective function was reduced from 0.446×10^5 to 2.6. The final parameter values were 773.7 and 58.99 compared with the original values of 775 and 60. The equivalent number of model runs for each Gauss-Newton iteration was 1.36 for this two parameter case.

For Run 6, using the same observed data, it was attempted to estimate the horizontal permeabilities of four zones, layers 6 to 9 with a starting guess of 200 md. The objective function was reduced from 0.517×10^5 to 2.05 in thirteen iterations. The initial and final water-oil ratio profiles

Fig. 7.4 - Match of Water-Oil Ratio in Run No. 6



are shown in Figure 7.4. Table 7.1 compares the estimated permeabilities with the true values. The match with the observed data is actually very good even though the actual permeabilities were not recovered very closely. Note that the calculated permeability-thickness of the four zones is within four percent of the actual permeability-thickness. Thus a good estimate of the permeability-thickness is obtained.

The porosity of layers 6 to 9 were estimated in Run 7 using an initial guess of 0.1. After five Gauss-Newton iterations, the objective function was reduced from 0.158×10^5 to 2.55. Table 7.2 compares the estimated porosities with the original values. Further iterations did not reduce the objective function any further. The equivalent number of model runs for each iteration was 1.49.

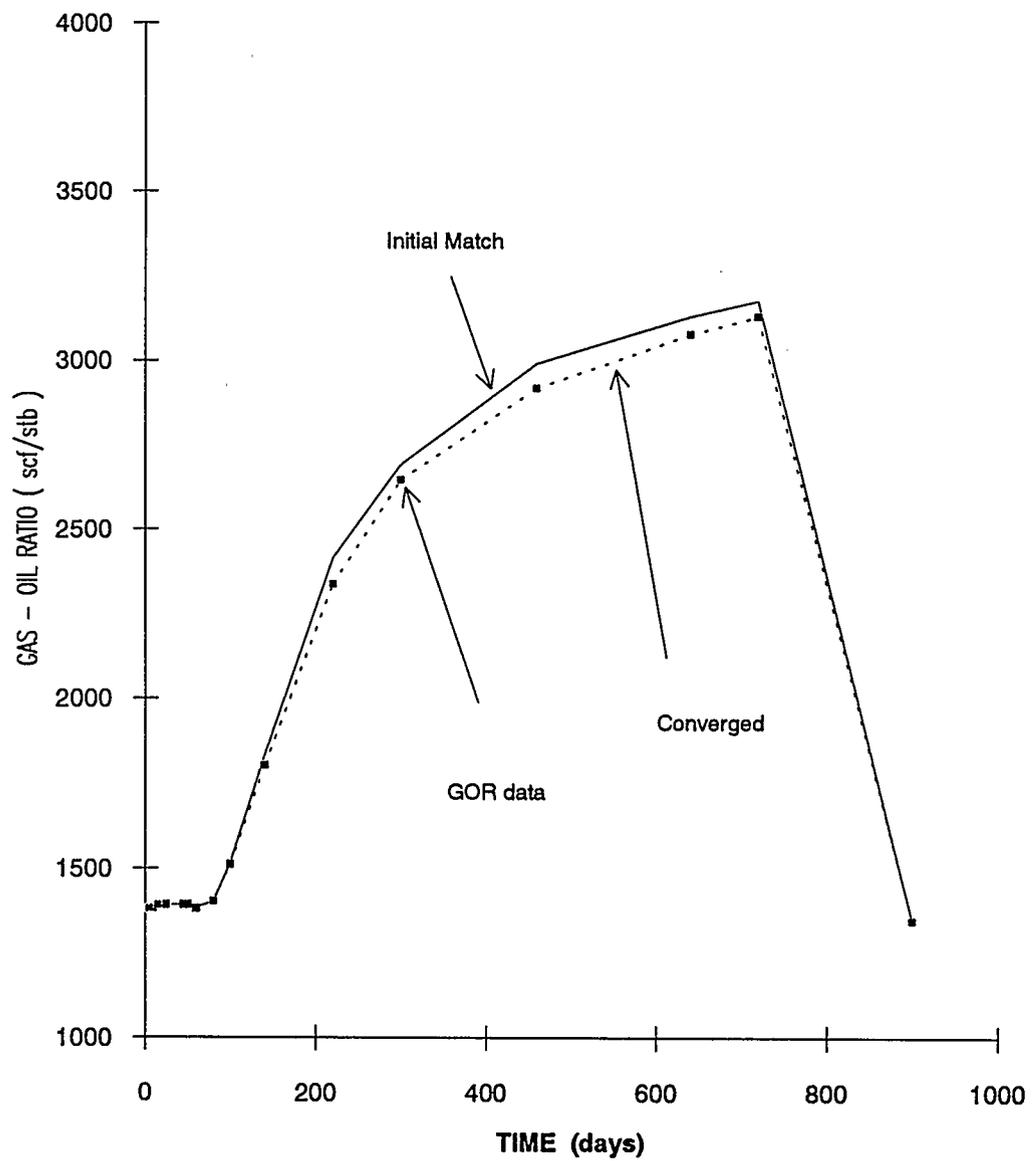
7.4.3 Matching Gas-Oil Ratio

The gas-oil ratio, like the water-oil ratio, is a derived quantity from a complex function of the state variables. Eq. 7.11 is solved numerically to obtain new estimates of the parameter vector.

The observed data consists of the gas-oil ratio reported at each of the 16 timesteps in the base model run using the original data. This data is shown in Figure 7.5. As weighting matrix, the identity matrix is used.

Using two zones, layers 7 and 8, and the horizontal permeabilities as the unknown parameters, the program was

Fig. 7.5 - Match of Gas-Oil Ratio in Run No. 9



unable to recover the original values using an initial guess of 200 md. Starting with guesses much closer to the original values, it was able to converge to the correct solution. For example, in Run 8 using initial estimates of 770 and 70 md, the final estimated values were 774.5 and 60.51 in three iterations. The cputime required was an equivalent of 1.25 model runs for each iteration.

Using porosity as the parameter and layers 7 and 8 as the two zones in Run 9, the program was able to recover the porosities from an initial estimate of 0.1. Six iterations were required to obtain a final estimate of 0.1727 and 0.0789. The initial and final match values are shown in Figure 7.5. The average time for each iteration was 1.30 equivalent model runs.

Unlike the runs to match water-oil ratio, it was not possible to estimate the values of either the permeability or porosity for four zones using the observed gas-oil ratio data. This indicates the gas-oil ratio is a much more nonlinear function than the water-oil ratio. The combined effect of the linearization of the gas-oil ratio function together with the linearization of the output vector with respect to the parameter vector as in Eq. 7.10 is to reduce the region of convergence. Initial estimates that are closer to the correct values must be provided as in Run 8.

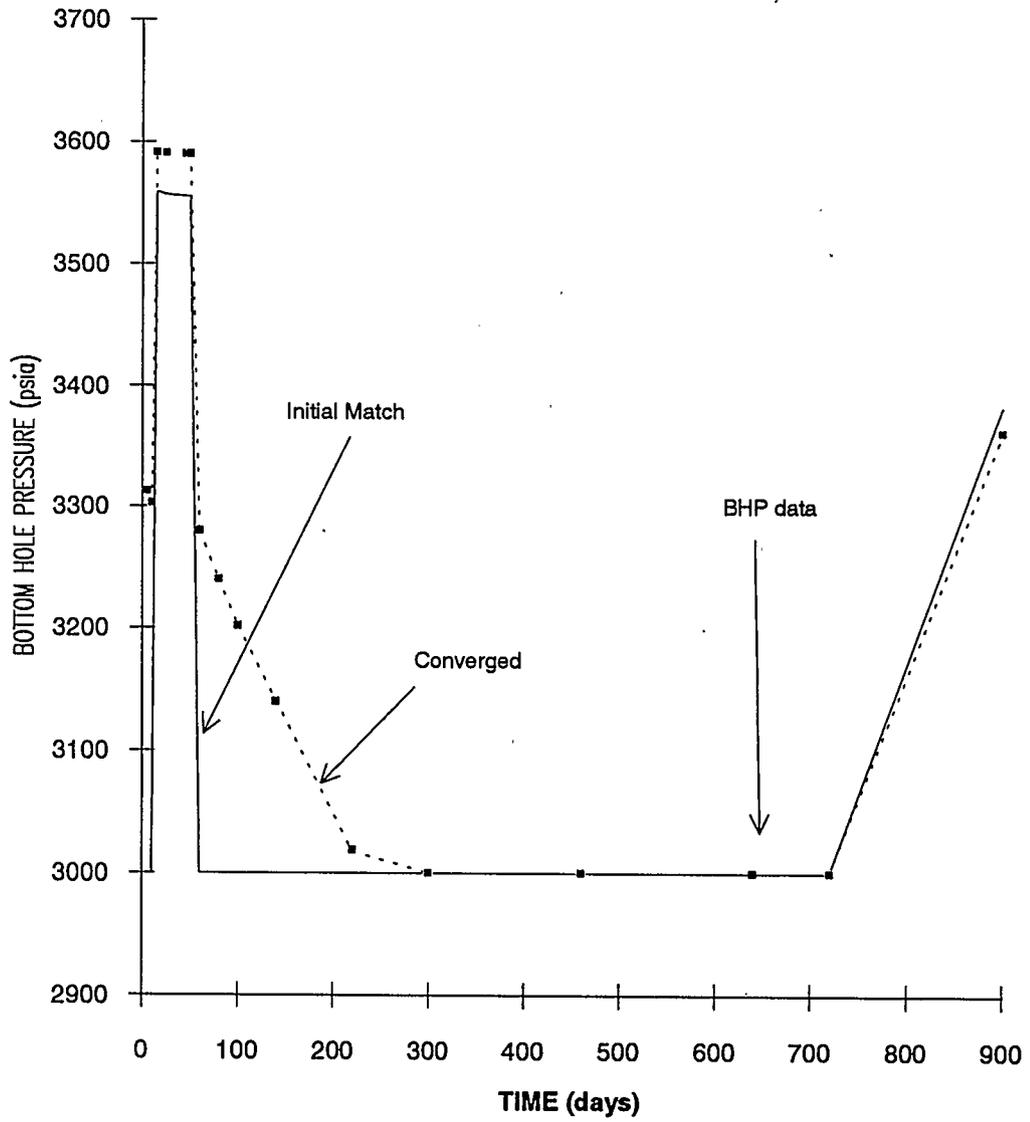
7.4.4 Matching Bottom Hole Pressure

In a fully implicit model, the bottom hole pressure of each well is also a state variable when the well is capacity restricted. At other times, the well is bottom hole pressure restricted and the bottom hole pressure is fixed at the minimum value allowed. The functional relationship between the observed vector and the state vector is that of Eq. 7.9.

The observed data for this case consists of the bottom hole pressure measurements for the central well at each timestep. This data is displayed in Figure 7.6. Since the functional relationship only exists when the measured bottom hole pressure is greater than the minimum bottom hole pressure, the actual data region is much shorter. Using only one zone, layer 7, and an initial estimate of 200 md, Run 10 predicted the true permeability to be 761.3 md in 7 iterations. The actual permeability was 775 md. The average number of equivalent model runs was 1.41 per Gauss-Newton iteration.

With the same data, Run 11 tried to estimate the permeabilities of layers 7 and 8 starting from an initial guess of 200 md. After 17 iterations, the estimated values were 743.7 and 107.8 md. Each iteration required an average of 1.51 equivalent model runs. The initial and final match values of the bottom hole pressure are shown in Figure 7.6. Again the match is very good.

Fig 7.6 - Match of Bottom Hole Pressure in Run No.
11



7.4.5 Matching Combinations of Observed Data

The observed data from producing wells in a reservoir would normally consist of all the data previously matched individually. As the Gauss-Newton equations 7.8 and 7.10 are fundamentally similar, it is an easy matter to construct the Gauss-Newton matrix elements regardless of whether the relationship Eq. 7.9 or Eq. 7.11 is used. The program can therefore match reservoir pressures, water-oil ratios, gas-oil ratios and flowing bottom hole pressures simultaneously.

The observed data for the remaining runs consists of the WOR, GOR, and flowing bottom hole pressure measurements from the previous runs, and the reservoir pressures at the two locations of the well, layers 7 and 8. This data is shown in Figures 7.7 and 7.8.

For Run 12, the horizontal permeabilities of layers 6 to 9 were estimated starting from an initial guess of 200 md. The objective function was reduced from 0.1997×10^7 to 0.97×10^{-1} in 12 iterations. Table 7.1 shows the final estimates of the permeabilities. In comparison with Run 6, which used only water-oil ratio data, it can be seen that the estimates are closer to the actual values. This is to be expected as more data is available in this run. The equivalent number of model runs for each Gauss-Newton iteration was 1.52.

Fig 7.7 - Match of Pressure Data in Run No 13
(Combined Data)

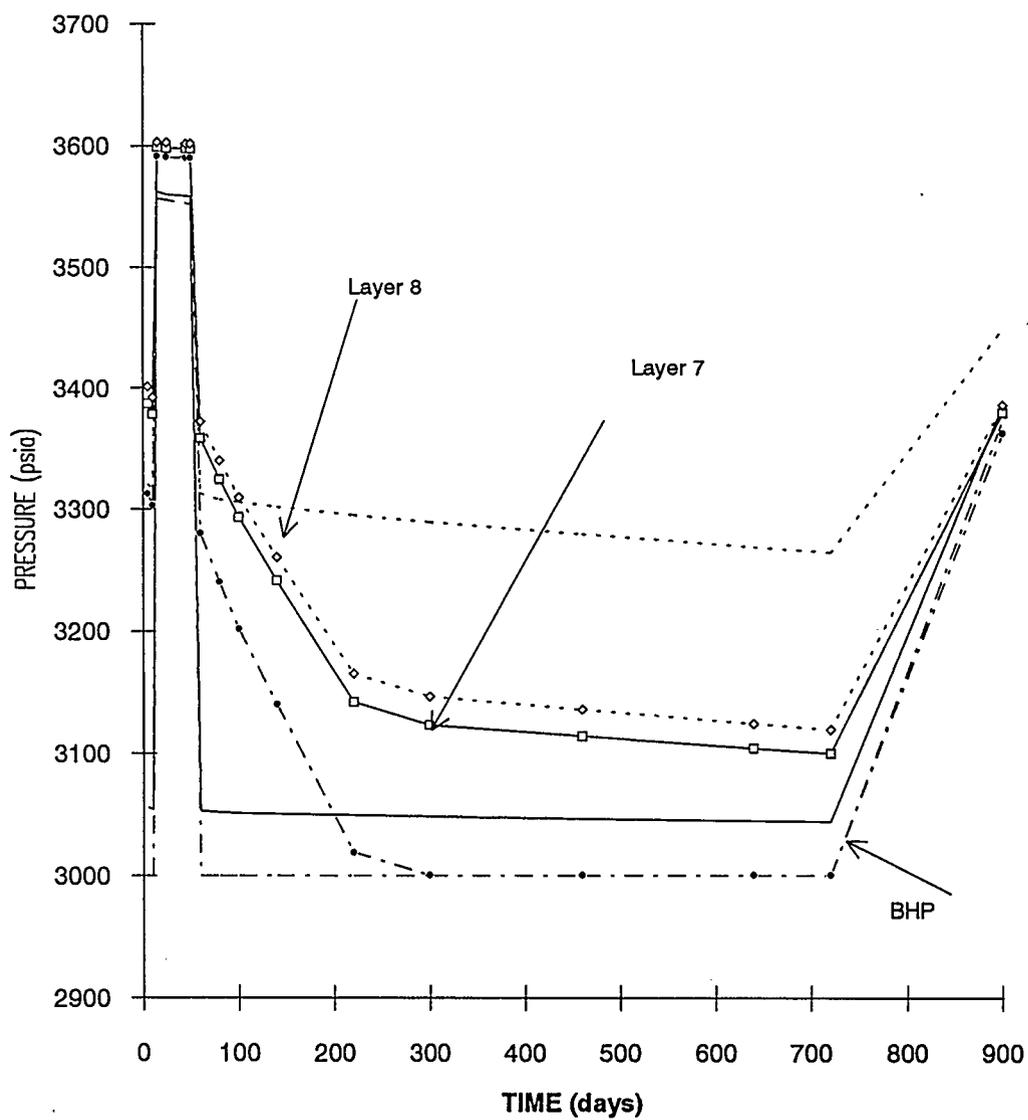
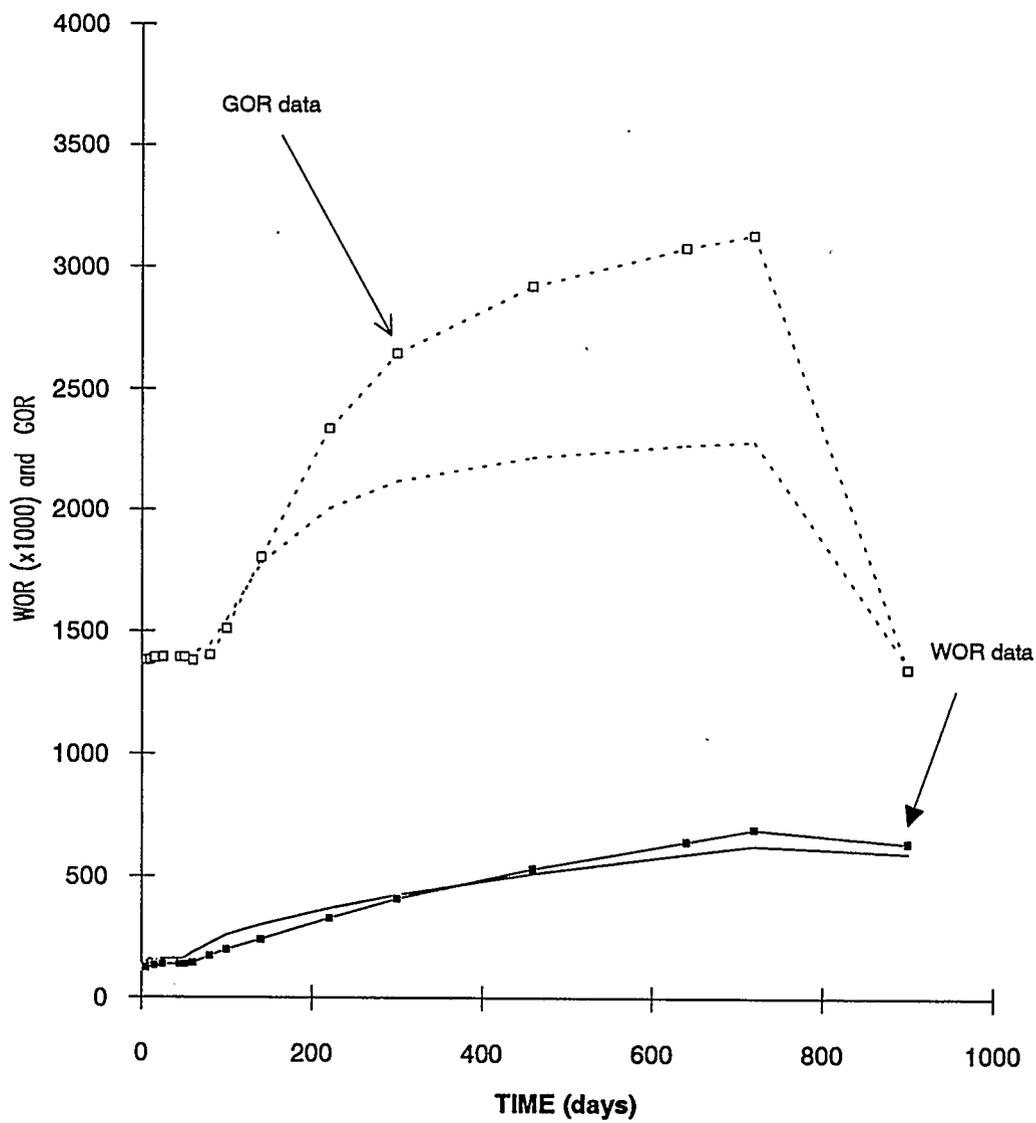


Fig. 7.8 - Match of WOR and GOR in Run No. 13
(Combined Data)



It is of interest to investigate the extent of reservoir that is identifiable from measurements made only at the well locations. Run 13 estimated the horizontal permeabilities of 6 zones, layers 5 to 10, starting from the initial guess of 200 md. The pseudo-inverse option was necessary to ensure initial stability. As the rank of the pseudo-inverse was increased, it was noted that the parameters that had the most significant changes were the ones furthest away from the well locations. The objective function was reduced from 0.33×10^7 to 0.91 in 22 iterations. Table 7.1 shows the final estimates of the 6 zone permeabilities. The initial and final profiles of each of the calculated variables are displayed in Figures 7.7 and 7.8. Each Gauss-Newton iteration required an average of 1.76 model runs. An attempt to estimate permeability values for 8 zones from the available data failed. It would be necessary to have observation data over a longer history in order to identify the parameters of more zones.

The porosities of the 6 zones, layers 5 to 10, were the parameters for Run 14. After 10 iterations, the objective function was reduced from 0.1938×10^6 to 0.456. Table 7.2 shows the estimated porosities. The equivalent number of model runs per iteration was 1.67. Again it was not possible to estimate porosity values for 8 layers.

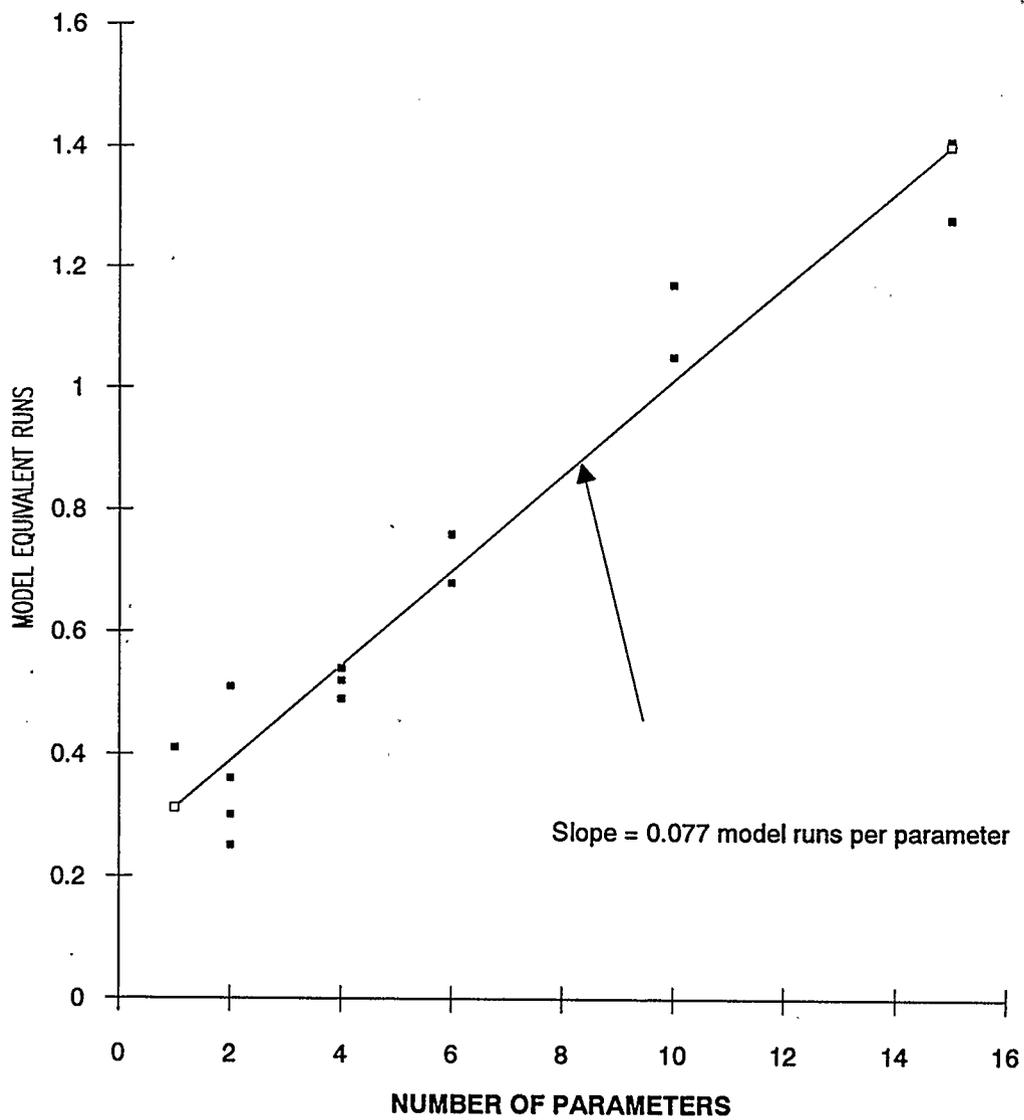
7.5 Correlation of Computational Requirements with Number of Parameters

The runs described above demonstrate the capability of the automatic history matching model to match the types of production data commonly obtained. Even though the observed data used in these runs were synthetic, the runs were extremely useful in the development stage of this very complex model to establish its validity and to investigate the convergence behaviour of the algorithm.

Computational times in terms of equivalent model runs for each Gauss-Newton iteration were presented for each case. In order to quantify the time required to develop the sensitivity coefficients only, these times were sorted in order of increasing number of parameters and presented in Table 7.3. The values are graphed in Figure 7.9 and a simple linear least squares fit was made of the data. The equation of the correlation line was $0.233 + 0.077x$. This indicates that after an initial overhead of 0.233 model runs to set up the algorithm, an additional 0.077 equivalent model run was required to develop the sensitivity coefficients for each additional parameter. This is much less than the additional run per parameter required by previously published Gauss-Newton methods. These numbers should serve only as a guideline as each model run using a different set of parameters would take a different amount of time and a simple linear correlation was used.

Table 7.3: Comparison of Model Equivalent Runs with Number of Parameters		
No. of Parameters	Model Equivalent Runs	Run No
1	0.41	10
2	0.36	5
2	0.25	8
2	0.30	9
2	0.51	11
4	0.54	6
4	0.49	7
4	0.52	12
6	0.76	13
6	0.68	14
10	1.17	1
10	1.05	3
15	1.28	2
15	1.41	4

Fig. 7.9 - Correlation of Computational Requirements for Runs 1 to 13



CHAPTER 8

RELIABILITY OF PARAMETER ESTIMATES

In a history matching exercise, the engineer would like to know how accurately the parameters have been estimated, whether or not a particular parameter has any influence on the history match and to what degree. The traditional procedure of history matching by varying parameters based on engineering judgement may provide a match, but certainly does not provide the engineer with any estimate of the degree of confidence one may place on the reliability of these parameters. Often the engineer is only too pleased to find a set of parameters that results in a minimally acceptable match.

Least Squares (LS) estimation is commonly used to identify the objective function to be minimized and is the basis of many parameter estimation algorithms. Drawing from the extensive body of knowledge that has been developed in the fields of statistical inference and multivariate analysis, it is possible to develop relationships for the variances of estimated parameters, establish confidence intervals and detect highly correlated parameters from the results of a linear least squares fit. One of the benefits of using a simulator with an automatic history matching capability is the further information provided to the engineer about the quality of the match and the reliability of the estimates of the

parameters.

As discussed in the literature review, the techniques used in automatic history matching fall into two major categories, optimal-control based methods and nonlinear regression methods. The first-order optimal-control method proposed by Chen et al. (1974) and Chauvent et al. (1975) involves the solution of a set of adjointed ordinary differential equations together with the ordinary differential equations of the reservoir model. However, this provides only the gradient of the least squares objective function with respect to the parameters, and this information is not sufficient to generate the desirable covariance matrix that quantifies the reliability of the estimates.

Nonlinear regression methods such as the Gauss-Newton method have also been proposed by Thomas et al. (1972) and Tan and Kalogerakis (1991). These methods require the computation of the sensitivity coefficient matrix (the partial derivatives of the reservoir variables with respect to the reservoir parameters) at each iteration of the parameter search. The sensitivity coefficient matrix can be used to determine the covariance matrix of the parameters which quantifies the accuracy of the estimates. This is an additional advantage of the nonlinear regression methods over optimal control methods.

8.1 The Covariance Matrix of the Parameter Estimates

Nonlinear regression parameter estimates obtained by the

Gauss-Newton method possess characteristics of linear LS estimates due to the nature of the Gauss-Newton method.

The Gauss-Newton equation (also known as the "normal equation") was developed in full in Chapter 5.

$$\left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} \mathbf{C}(t_i) \mathbf{G}(t_i) \right] \Delta \mathbf{k}^{(j+1)} - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} (\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)) \quad (8.1)$$

This can be represented by

$$\mathbf{H} \Delta \mathbf{k}^{(j+1)} = \mathbf{b} \quad (8.2)$$

This set of linear equations can be solved to yield $\Delta \mathbf{k}^{(j+1)}$ and therefore the next estimate $\mathbf{k}^{(j+1)}$ of the parameter vector.

When convergence is achieved, the matrix \mathbf{H} by analogy to linear least squares can be used to develop measures of the variance of the parameters.

Under the null hypothesis that the errors in the measurements, \mathbf{e}_i are independently and normally distributed with zero mean and variance-covariance matrix $\text{COV}(\mathbf{e}_i) = \sigma_e^2 \mathbf{Q}_i^{-1}$, the covariance matrix of the parameters is given by

$$[\text{COV}(\mathbf{k})] = \sigma_e^2 \mathbf{H}^{-1} \quad (8.3)$$

where H is calculated with the converged parameter values, k^* , and $\hat{\sigma}_e^2$ is an unbiased estimate of σ_e^2 that can be obtained from

$$\hat{\sigma}_e^2 = \frac{S(k^*)}{(d.f.)} \quad (8.4)$$

where (d.f.) are the degrees of freedom which are equal to the total number of measurements minus the number of parameters, that is, $(d.f.) = n_0 m - p$.

8.2 Correlation between Pairs of Parameters

The variances of the parameters appear as the diagonal elements of $COV(k)$ and the covariances between the parameters appear as the off-diagonal elements. In general, the larger the variance of a parameter, the less confidence one would place on the estimated parameter value. The covariances between any two parameters is an indication of how closely these two parameters are correlated. The higher the covariance, the closer is the correlation, or the near dependence between the pair of parameters. If any two parameters are closely correlated, this implies that these parameters have the same influence on the observations and it may not be possible to identify the values of these two parameters separately. If this is the case, the two parameters may be combined into a single parameter.

In order to show this relationship clearer, the

covariance matrix can be transformed into a correlation matrix, R , using the simple matrix operations, namely,

$$R = D^{-1} [COV(k)] D^{-1} \quad (8.5)$$

where D is a diagonal matrix whose entries are the reciprocals of the square roots of the diagonal elements of matrix $Cov(k)$, i.e. the reciprocals of the standard deviations of the parameters. The diagonal elements of R will be 1 as a result of these operations, and the off-diagonal elements will have values whose absolute magnitude will range from 0 to 1. If an off-diagonal element has an absolute value close to 1, the corresponding pair of parameters is highly correlated. Thus observation of the elements of the correlation matrix R will assist in identifying the degree of correlation between pairs of zones.

8.3 Multiple Correlation

Unfortunately, cases do occur where a set of three or more parameters have a mutual near dependence whereas no two of the parameters are nearly dependent. In such a case, the off-diagonal elements of the correlation matrix R , corresponding to these parameters taken as pairs, will not be close to 1 or -1. Correlations between three or more parameters are very difficult to detect using the covariance or the correlation matrix. However, in this case, the eigenvalue decomposition of the matrix H can be used to

identify groups of parameters that are highly correlated and which may be combined and hence provide guidelines on how to reduce the number of parameters.

A geometrical approach may be taken to illustrate this property. Consider a p dimensional ellipsoid around the minimum of the LS objective function S , where p is the number of parameters to be estimated. Let this ellipsoid represent the region of indifference, i.e. the region around the minimum where S does not vary significantly. If this ellipsoid is highly elongated in one direction, then there is considerable uncertainty in the value of the parameters in that direction. We wish to find the points of the p dimensional ellipsoid which are furthest away from the origin and also those which are closest. These points determine respectively, the least determined and best determined linear combinations of the parameters. It can be shown that an eigenvalue decomposition of the matrix H will provide us with the information to describe this ellipsoid. The eigenvalue decomposition of H , a $(p \times p)$ matrix, yields the p eigenvalues and corresponding eigenvectors of H . The p eigenvectors form the direction of the p principal axes of the ellipsoid. The eigenvalues corresponding to the eigenvectors represent inversely the magnitude of the length of each principal axis. The longest axis, corresponding to the smallest eigenvalue defines the worst determined direction, and the shortest axis corresponding to the largest eigenvalue defines the best

determined direction.

Since H is symmetric, the eigenvalue decomposition of H is also an orthogonal decomposition, namely,

$$H = V \Lambda V^T \quad (8.6)$$

where Λ is a diagonal matrix of positive eigenvalues which can be arranged in decreasing order. V is an orthogonal matrix whose columns are the normalized eigenvectors of H and hence,

$$V^T V = I \quad (8.7)$$

The eigenvalue decomposition of H can be used to solve Eq. 8.2 which can now be written as,

$$V \Lambda V^T \Delta k^{(j+1)} = b \quad (8.8)$$

Since V is an orthogonal matrix, $\Delta k^{(j+1)}$ is readily obtained as

$$\Delta k^{(j+1)} = V \Lambda^{-1} V^T b \quad (8.9)$$

Usually the eigenvalues of H will have a wide range of values. If some of the k parameters are highly correlated, then the ratio of the largest eigenvalue to the smallest eigenvalue (usually referred to as the condition number of matrix H) will be very large. Note that the eigenvalues are not in a one-to-one correspondence with the parameters. That

is, for example if the fifth and sixth parameters are highly correlated, this does not mean that the fifth and sixth eigenvalues will be small. However, inspection of the eigenvectors corresponding to the smallest eigenvalues will allow the engineer to deduce which parameters are highly correlated. The elements of each eigenvector are the cosines of the angles the eigenvector makes with the axes corresponding to the p parameters. If r of these parameters are highly correlated, then at least r of the parameter axes are closely collinear, i.e. lie in the same direction, and the angle between them is very small. Thus the eigenvector corresponding to the smallest eigenvalue will have significant contributions from r of these parameters, and the cosines of the angles will tend to 1 as the angles tend to zero. Thus the larger elements of the eigenvector will identify a particular combination of parameters which are nearly dependent. The same logic applies to any other eigenvalues / eigenvectors which are also very small.

8.4 Numerical Results and Discussion

In order to illustrate how to use such information in analyzing reservoir history matching results, let us begin with a very simple problem so that we can observe the relationships between the zones as predicted by the theory without getting lost in the geometry of the problem. In each case the eigenvalues, matrix of eigenvectors, covariance

	1	2	3
1	1 (1,1)	2 (1,2)	3 (1,3)
2	4 (2,1)	5 (2,2)	6 (2,3)
3	7 (3,1)	8 (3,2)	9 (3,3)

Fig. 8.1 - Schematic of 3 x 3 Grid Cell Model

matrix and the correlation matrix are calculated by the automatic history matching program. The reported standard deviations of the parameters are normalized by dividing with the individual parameter.

8.4.1 Case Study I : A 3 x 3 Symmetrical Model

Consider a uniform areal 3x3 model with a producing well situated in grid cell (1,1). The model is shown schematically in Figure 8.1. The fluid properties used are those of the SPE Second Comparative Solution Problem (Chappelear and Nolen, 1986). The model is initialized with oil and a connate water saturation. As depletion starts, the pressure of the cell with the producing well drops below the saturation pressure and solution gas is liberated. This is a three-phase problem. The physical details of the problem are quite irrelevant as the purpose is to show the correlation between the zones.

We assume that each cell has the same properties, such as grid cell dimensions, depth and porosity. The permeability in either direction is the same. The problem therefore is symmetrical about the diagonal (1,1), (2,2) and (3,3). The parameters to be estimated are the porosities of each grid cell. The observed data is generated by running the model for a period of 100 days and observing the pressures of the grid cells at intervals of 10 days. Artificial data are generated by adding noise to the noise-free observations by using random numbers z_{ij} , distributed normally with mean zero and variance

equal to 1 and then using the equation

$$\hat{y}_j(t_1) = y_j(t_1) + z_{1j} * |y_j(t_1)| * \sigma_e \quad (8.10)$$

Here, $\hat{y}_j(t_1)$ are the "noisy" observations of variable j at time t_1 , $y_j(t_1)$ are the exact observations of variable j at time t_1 (calculated by the model using the exact parameter values) and σ_e is the standard deviation of the measurement errors, taken as a fraction of the magnitude of the errors.

The reservoir simulation model is then run to match the noisy observations with a weighting matrix Q that is a diagonal matrix with elements $|y_j(t_1)|^{-2}$. At convergence, the eigenvectors, eigenvalues and correlation matrix are inspected. Initially assume the reservoir is divided into 9 zones, each cell comprising one zone, and numbered in natural ordering. The pressures of each grid cell are observed. For the first case a value for σ_e of 0.001 was used. The correlation matrix is shown in Table 8.1 and the eigenvector matrix in Table 8.2.

The correlation matrix can be used to identify quickly any correlation among the parameters. For example, if we concentrate on parameter 5, we can readily see that the elements 2,4,6 and 8 are large. Therefore parameter 5 is highly correlated with parameters 2,4,6 and 8. Not surprisingly, these are the zones (cells) which are directly adjacent to zone (cell) 5. Similarly we see that column 1 has largest elements corresponding to parameters 2 and 4. These

TABLE 8.1 : Correlation Matrix - Full Data - 3 x 3 Symmetrical Problem									
	1	2	3	4	5	6	7	8	9
1	1.000	-0.1514	0.03726	-0.1514	0.05341	-0.01194	0.03727	-0.01193	-0.00588
2	-0.1514	1.000	-0.5262	0.03526	-0.4195	0.1650	0.01469	0.07712	-0.02157
3	0.03726	-0.5262	1.000	0.01471	0.1007	-0.5490	0.00359	0.00914	0.1471
4	-0.1514	0.03526	0.01471	1.000	-0.4195	0.07709	-0.5262	0.1650	-0.02159
5	0.05341	-0.4195	0.1007	0.4195	1.000	-0.4257	0.1008	-0.4257	0.1735
6	-0.01194	0.1650	-0.5490	0.07709	-0.4257	1.000	0.00916	0.1237	-0.5347
7	0.03727	0.01469	0.00359	-0.5262	0.1008	0.00916	1.000	-0.5491	0.1472
8	-0.01193	0.07712	0.00914	0.1650	-0.4257	0.1237	-0.5491	1.000	-0.5349
9	-0.00588	-0.02157	0.1471	-0.02159	0.1735	-0.5347	0.1472	-0.5349	1.000

Table 8.2: Eigenvector Matrix - Full Data - 3 x 3 Symmetrical Problem									
	1	2	3	4	5	6	7	8	9
1	-0.0465	0.00003	-0.1057	-0.0001	-0.1207	0.9059	-0.3880	-0.00009	0.03107
2	-0.4238	0.4617	-0.4602	0.5277	-0.2052	0.00299	0.2390	0.09116	-0.09761
3	-0.3645	0.5046	0.04238	-0.3880	0.5858	0.05658	-0.00670	-0.3076	0.1383
4	-0.4237	-0.4615	-0.4600	-0.5281	-0.2049	0.00306	0.2392	-0.09097	-0.09766
5	-0.3782	0.000003	0.05546	-0.00023	-0.3213	-0.3490	-0.6480	-0.00037	0.4571
6	-0.2959	0.1795	0.3669	-0.2657	-0.09028	-0.01148	-0.1040	0.6303	-0.5086
7	-0.3645	-0.5048	0.04193	0.3885	0.5854	0.05652	-0.00698	0.3075	0.1385
8	-0.2959	-0.1796	0.3666	0.2658	-0.09055	-0.01145	-0.1037	-0.6301	-0.5090
9	-0.2342	-0.00005	0.5384	0.00015	-0.3096	0.2257	0.5418	0.00003	0.4636

cells 2 and 4 are adjacent to cell 1. Similar relationships can be observed for the other parameters. The correlation matrix can therefore be used to assess which zones (parameters) are highly correlated. In this example, adjacency leads to a degree of correlation, and which is clearly demonstrated by the correlation matrix.

Each column of the eigenvector matrix is an eigenvector corresponding to the eigenvalues of the matrix H . As the eigenvalues are arranged in decreasing order, the last column corresponds to the smallest eigenvalue. The elements of each eigenvector have a direct correspondence with the parameters. This is borne out by the observation that rows 2 and 4, 3 and 7, 6 and 8 are nearly equal. Recall that the reservoir model is symmetric about a diagonal and thus rows 2 and 4 represent the equal contributions of cells 2 and 4 to the eigenvector, and similarly for the other pairs.

The standard deviation of the parameters is shown in column 1 of Table 8.3. In this particular case, all the grid cell pressures are observed and thus there is enough information to estimate reliably the values of all the parameters. The standard deviation of the parameters is therefore quite small. Column 2 of Table 8.3 shows the standard deviation of the same nine parameters when the data noise is increased by changing σ_e to 0.005. As expected, the standard deviation of the parameters increases with noise. If the model is correct and the only errors are due to the

measurement errors, then an estimate of the variance of the measurement error can be obtained from Eq. 8.4. These requirements are met here. The estimated variance is 0.0000299 corresponding to a standard deviation of 0.0055. This agrees well with the value of 0.005 used to generate the noise. A χ^2 test could be readily performed to confirm that this is the case. Dumez et al. (1977) used similar tests for chemical kinetic models.

Next we demonstrate the effect of having fewer data (limited observability) on the identifiability and reliability of the estimated parameters. Let us assume that the observations are obtained only at cells 1 (1,1) and 9 (3,3). In this case, we are not able to estimate the values of all nine cells. The standard deviation of the 9 parameters are shown in column 3 of Table 8.3. The standard deviations are smallest for the cell where the well is located (cell 1), the cell with the observation well (cell 9) and the diagonal cell (cell 5). The standard deviation is smaller for cell 1 than for cell 9 because the well provides more excitation, dropping the pressure rapidly in cell 1, and the pressure of the observation cell drops in response to the pressure drop of cell 1. The pressure drop is more significant in cell 1 than in cell 9. The cells on either side of the diagonal have huge standard deviations indicating that it is not possible to identify and reliably estimate the parameter values of these cells with the information available. In addition, due to the

symmetry of the problem, the parameters are highly correlated. In this situation, one could say that all six cells 2,3,4,6,7 and 8 have similar influence on the observations. The correlation matrix, reproduced in Table 8.4, cannot show the effect of correlations between more than two parameters as discussed in the mathematical background. However, the eigenvector corresponding to the smallest eigenvalue as shown in Table 8.5 shows the correlation between these six zones. The eigenvector elements are the cosines of the eigenvector axis with the parameter axes. Parameters which are nearly collinear will have smaller angles between them and the eigenvector axis and this will be indicated by larger cosine values. The largest entries are indeed those of the parameters 2,3,4,6,7 and 8.

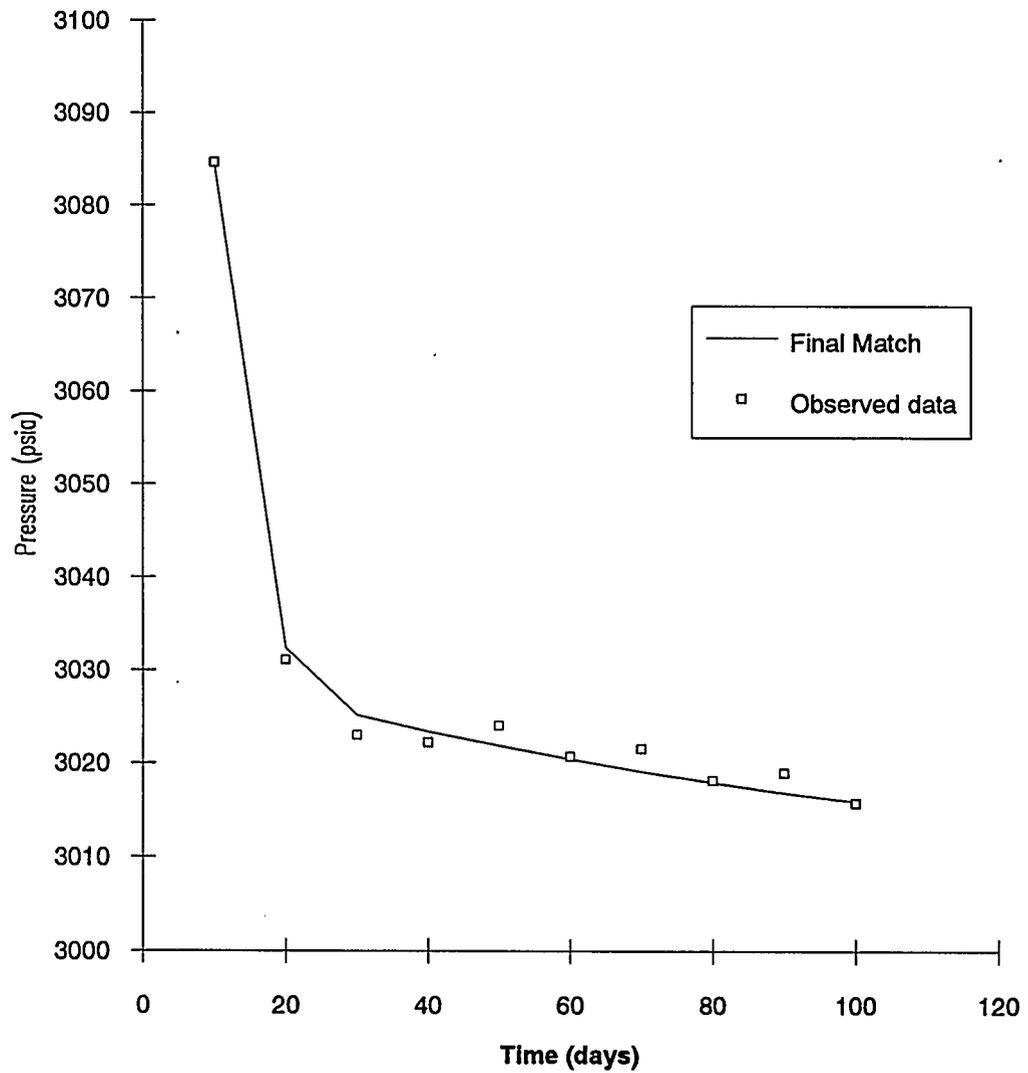
To reduce the number of parameters then, one would lump adjacent zones of highest variance. The suggested zonation would be zone 1 consisting of cell 1, zone 2 consisting of cell 5, zone 3 consisting of cell 9, and zone 4 consisting of the rest of the cells (2,3,4,6,7,8). With the suggested zonation using 4 parameters, the value of the objective function was 0.164×10^{-4} . Figures 8.2 and 8.3 show the match of the calculated pressures to the noisy data. The match using 9 or 4 parameters was virtually indistinguishable showing the insensitivity of the predictions to parameters with high variance.

Zone	Std. Dev. $\sigma_\epsilon=0.001$ Full Data	Std. Dev. $\sigma_\epsilon=0.005$ Full Data	Std. Dev. $\sigma_\epsilon=0.001$ Limited Data
1	0.06745	0.54280	0.18
2	0.04077	0.18713	2016.2
3	0.04840	0.26567	1022.7
4	0.04078	0.16417	2013.6
5	0.07820	0.41325	40.63
6	0.08368	0.39927	809.1
7	0.04840	0.32481	1056.8
8	0.08370	0.33589	752.2
9	0.07433	0.37696	8.32

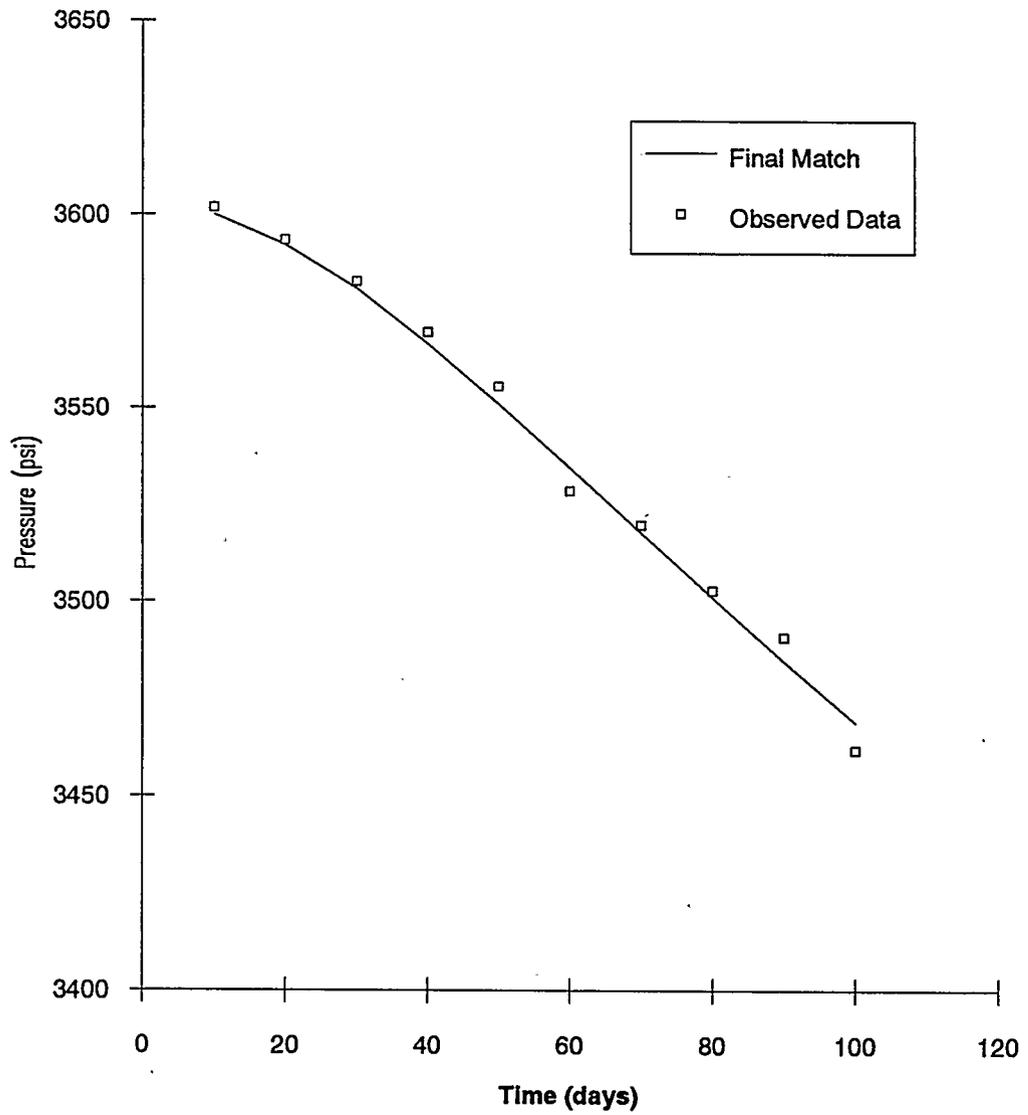
Table 8.4 : Correlation matrix - Limited Data - 3 x 3 Symmetrical Problem									
	1	2	3	4	5	6	7	8	9
1	1.000	-0.6787	-0.2694	0.6780	0.1110	0.7340	0.2483	-0.7216	-0.0914
2	-0.6787	1.000	0.3795	-1.000	-0.4541	-0.7781	-0.3128	0.7148	0.5354
3	-0.2694	0.3795	1.000	-0.3794	0.3985	-0.3522	-0.9957	0.2945	0.5417
4	0.6780	-1.000	-0.3794	1.000	0.4546	0.7776	0.3126	-0.7142	-0.5361
5	0.1110	-0.4541	0.3985	0.4546	1.000	0.3784	-0.4805	-0.3647	-0.0623
6	0.7340	-0.7781	-0.3522	0.7776	0.3784	1.000	0.3509	-0.9943	0.01520
7	0.2483	-0.3128	-0.9957	0.3126	-0.4805	0.3059	1.000	-0.2535	-0.4982
8	-0.7216	0.7148	0.2945	-0.7142	-0.3647	-0.9943	-0.2535	1.000	-0.1210
9	-0.0914	0.5354	0.5417	-0.5361	-0.0623	0.01520	-0.4982	-0.1210	1.000

Table 8.5 : Eigenvector Matrix - Limited Data - 3 x 3 Symmetrical Problem									
	1	2	3	4	5	6	7	8	9
1	-0.0311	0.9970	-0.06308	0.03239	-0.00812	0.00118	-0.00007	0.00000	-0.00004
2	-0.1756	0.04337	0.5821	-0.3493	0.07393	-0.03725	-0.2027	-0.1498	0.6603
3	-0.1946	-0.00782	0.1356	0.2511	-0.1787	0.6148	-0.0044	0.6668	0.1587
4	-0.1758	0.04337	0.5834	-0.3427	0.1007	0.00336	0.2049	0.1498	-0.6594
5	-0.4114	-0.01738	0.2760	0.7182	0.09156	-0.4791	-0.00314	0.02049	-0.00527
6	-0.1405	-0.00679	-0.01384	-0.1161	-0.6726	-0.1822	-0.6577	0.01543	-0.2194
7	-0.1925	-0.00774	0.1341	0.2580	-0.1445	0.5729	0.01254	-0.7134	-0.1419
8	-0.1497	-0.00730	-0.01648	-0.1050	-0.6463	-0.1694	0.6951	-0.03106	0.1888
9	-0.8068	-0.04213	-0.4514	-0.2993	0.2302	0.02985	-0.00900	0.00193	0.00142

**Fig. 8.2 - 3 x 3 Symmetrical Problem with Limited Data
Match of Pressures at Cell 1**



**Fig. 8.3 - 3x3 Symmetrical Problem with Limited Data
Match of Pressures at Cell 9**



8.4.2 Case Study II : A 3 x 3 Asymmetrical Model

Having used the symmetrical model to show how the parameters interact in a well defined case, we now turn to the more general situation of the asymmetrical model. Keeping everything else unchanged, the porosities of the cells were varied. Specifically, in natural ordering, the values used were 0.1, 0.15, 0.05, 0.2, 0.3, 0.1, 0.15, 0.1 and 0.2. Observations were taken only at the well cell 1 (1,1) and at cell 9 (3,3).

With the limited number of observations, it is not possible to estimate the individual porosities of all nine cells. Proceeding as in the previous case, we use a preliminary zonation taking each cell as a separate zone. Column 1 of Table 8.6 shows the standard deviations of each of the nine parameters as generated by the automatic history matching program. The cells sorted in ascending order of standard deviation are 1,9,2,4,5,7,3,6 and 8. If we locate these cells on the grid, we see that the standard deviations increase radially from the producing well. Thus apart from the well cell 1, and the observation cell 9, two rings are apparent. One consists of cells 2,4 and 5, and the other consists of cells 3,6,7 and 8. This is to be expected on physical grounds and differs from the symmetrical case where the highest variances are on both sides of the diagonal joining cells 1 and 9. The eigenvector corresponding to the smallest eigenvalue also confirms that the parameters that are

most closely correlated are those of cells 3,6,7 and 8. The elements of this eigenvector are reproduced in column 2 of Table 8.6.

The zonation suggested by these standard deviation values would be zone 1 consisting of cell 1; zone 2 consisting of the ring of cells 2,4,5; zone 3 consisting of the next ring of cells, 3,6,7 and 8; and zone 4 consisting of cell 9. The minimum value of the least squares objective function using these four parameters is 0.1584×10^{-4} . Figures 8.4 and 8.5 show the match of the observed pressures. The quality of the match obviously has not decreased with the fewer parameters. The standard deviation of the four parameters is now much smaller with fractional magnitudes of 0.15, 0.29, 0.56 and 0.53. The predicted porosity values are 0.086, 0.16, 0.134 and 0.2048 which compares with the volumetrically averaged values of 0.1, 0.217, 0.1 and 0.2. The conclusion that can be drawn from this is that the user of an automatic history matching model should use a zonation that reflects the quantity of data available. The advantages of using fewer parameters are that the model would use less computer time, the standard deviations of the parameters are smaller giving more confidence in the estimates of the values.

Table 8.6 : Standard Deviations and Eigenvector of Smallest Eigenvalue - 3x3 Asymmetrical Problem - Limited Data		
Zone	Std. Dev	Eigenvector of Smallest Eigenvalue
1	0.3484	-0.00018
2	28.2466	0.00802
3	217.149	-0.2087
4	28.7422	-0.00092
5	82.2182	-0.04818
6	520.476	0.6262
7	160.398	0.1617
8	608.605	-0.7319
9	8.0241	0.00472

Fig. 8.4- 3x3 Asymmetrical problem with Limited Data
Match of Pressures at Cell 1

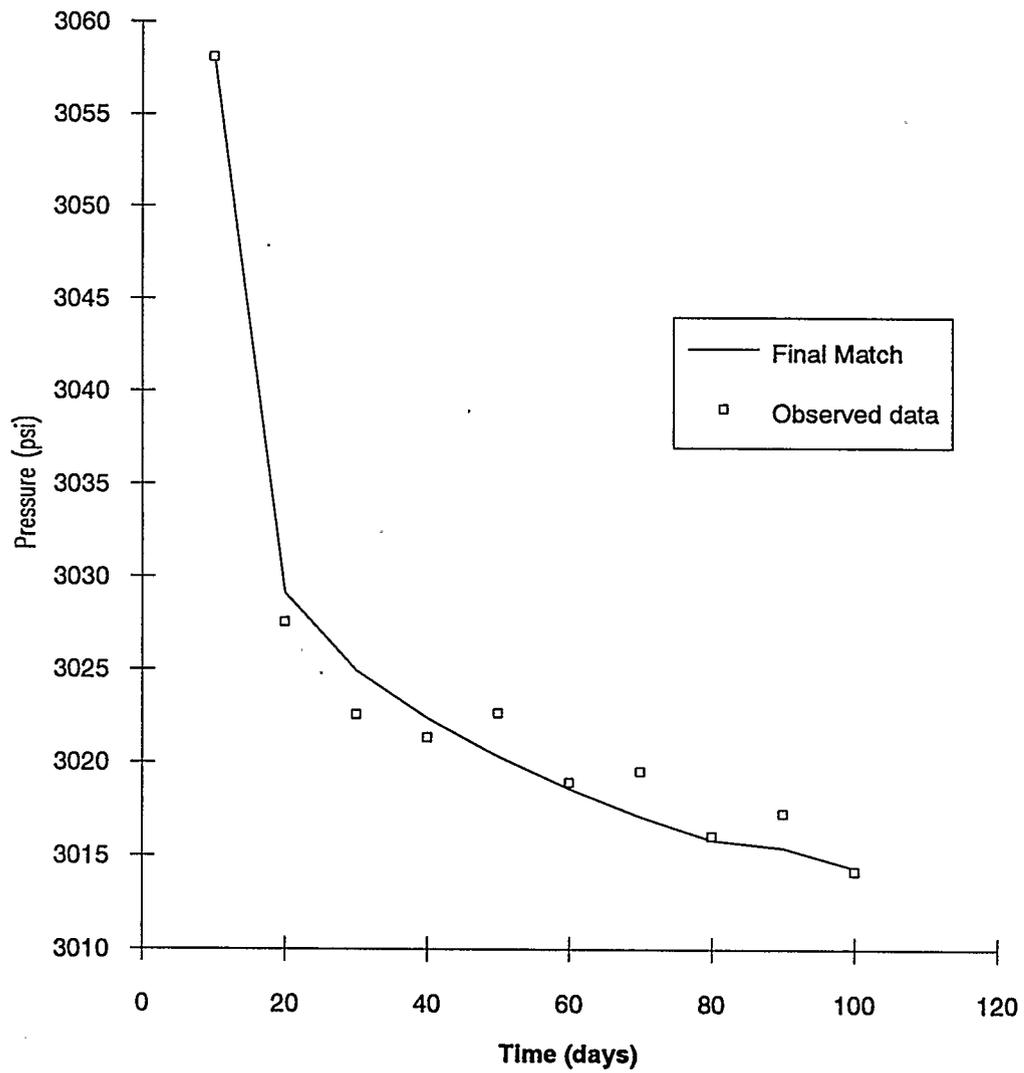
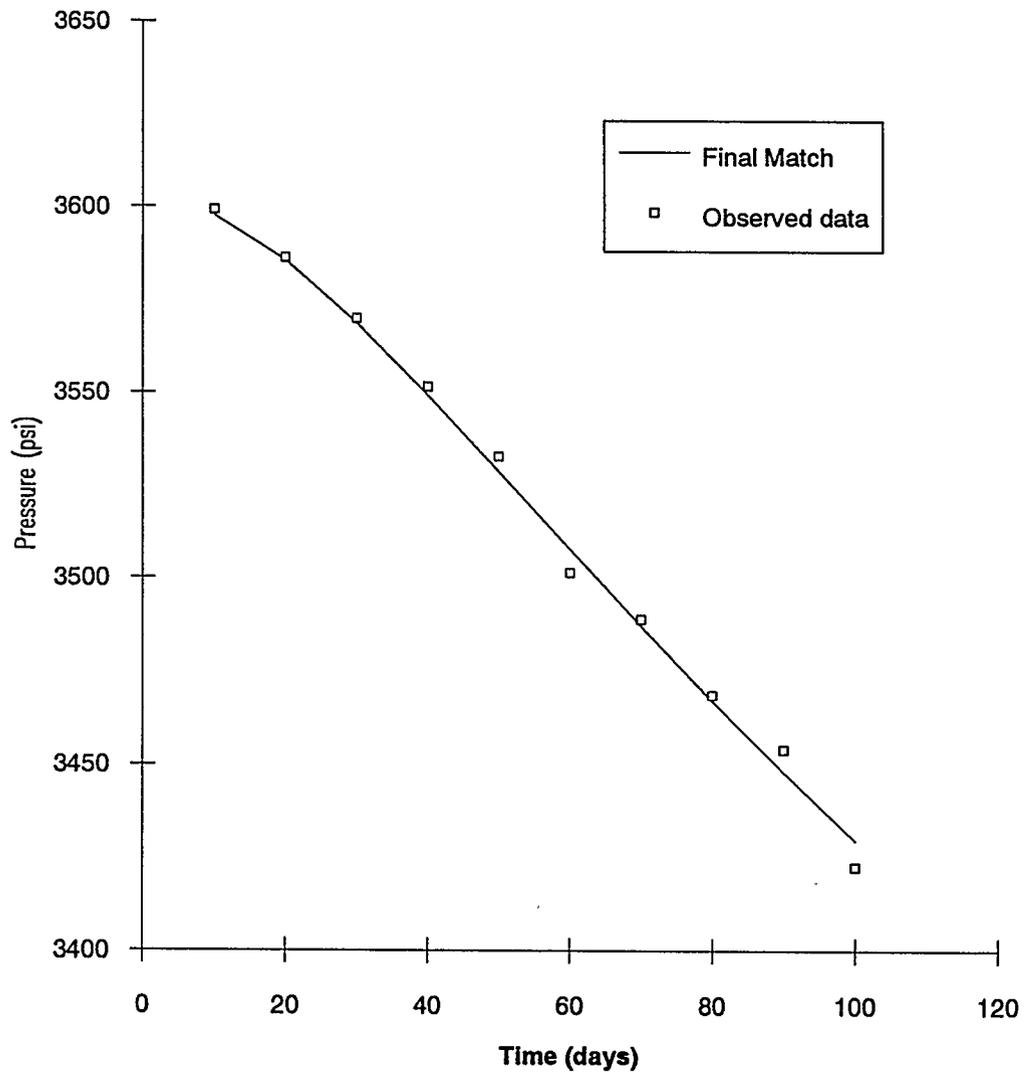


Fig. 8.5 - 3x3 Asymmetrical Problem with Limited Data
Match of Pressure at Cell 9



8.4.3 Case Study III : The SPE Second Comparative Solution Problem (Chappelear and Nolen, 1986)

As mentioned earlier, the model in this problem has 15 layers, each of constant horizontal permeability. We will try to estimate the values of these permeabilities using the observations made at the well. The well is completed in layers 7 and 8 and with permeability thicknesses of 6200 and 480 md ft. respectively. The model represents a three phase coning problem and the measurements made are the producing gas-oil ratio, water-oil ratio, flowing bottom-hole pressure, and the reservoir pressures of the well locations. The observations are generated running the model with the original description and at the end of each of the 16 timesteps required to complete the run, recording the above variables. Noise was then added to the data to reflect measurement errors as described previously. The value of σ_e used was 0.001.

In the preliminary zonation, all fifteen layers were treated as separate zones. An initial guess of 300 md was used for the permeabilities of all the 15 zones. After ten Gauss-Newton iterations, the objective function was 0.147×10^{-1} . The values of the estimated permeabilities, the true permeabilities, the calculated standard deviations and the eigenvector corresponding to the smallest eigenvalue are shown in Table 8.7. The largest elements of the eigenvector correspond to zones 1,2,3,12,13,14,15 indicating these zones are highly correlated. The standard deviation of the

parameters are smallest for zones 4,5,6,7 and 8. Parameters 9,10,11 have standard deviations of relative magnitude greater than 1. Because of this, zones 9,10,11 are combined with zones 12 to 15 in the revised zonation. Zones 1,2 and 3 are lumped into one zone. The revised zonation consists of lumped zone 1, zones 2,3,4,5,6 corresponding to layers 4,5,6,7,8 and lumped zone 7, making a total of 7 parameters.

With this revised zonation, the objective function was reduced to 0.392×10^{-2} . The final match is shown in Figures 8.6 and 8.7. The values of the estimated permeabilities and their standard deviations are shown in Table 8.7. It is of interest to note that parameter 5 corresponding to layer 7 has the lowest error. The well is completed in layers 7 and 8 but has a higher permeability thickness in layer 7, and thus has more production from this layer. Parameter 3, corresponding to layer 5, has a higher error than zones further away than it from the well locations even though this is contrary to expectation. This layer has a lower permeability thickness and its effect may have been missed in the timing of the observations.

The effect of less observed data on the reliability of the estimates is now considered. Assuming that only the pressures of the two well locations are measured, the calculated values of the permeabilities and the standard deviations of the 15 parameters, taking each layer as a zone, are shown in Table 8.8. The values of the permeabilities of

layers 7 and 8 can be accepted, but all the other layers have standard deviations of relative magnitude much greater than 1. We can safely combine these layers without affecting the match of the observed data. Combining adjacent zones of high variance, the suggested zonation is zone 1 representing layers 1 to 6, zone 2 consisting of layer 7, zone 3 consisting of layer 8 and zone 4 representing layers 9 to 15. When the automatic history matching model is run with this zonation, we obtain estimated values for the permeabilities of layers 7 and 8 that are much closer to the original values as shown in Table 8.8. The standard deviation of the estimates has also decreased, increasing their reliability.

Table 8.7 : Estimated Permeabilities, Standard Deviations and Eigenvector for SPE Problem 2 - Full Data						
Layer	True Permeability (md)	Estimated Permeability	Std. Dev. Relative magnitude	Eigenvector of smallest eigenvalue	Estimated Permeability Reduced Zones	Std. Dev. Relative Magnitude
1	35.0	237.5	8.6771	0.1280	246.2	0.1436
2	47.5	329.2	14.1603	-0.2501		
3	148.0	22.4	7.2519	0.1037		
4	202.0	289.2	0.4075	-0.00405	248.0	0.2135
5	90.0	240.8	0.6912	0.0108	128.1	0.3807
6	418.5	324.6	0.4029	-0.00557	478.6	0.1013
7	775.0	676.2	0.0533	0.00027	696.8	0.0194
8	60.0	52.62	0.2091	0.00007	67.48	0.0873
9	682.0	636.0	1.1684	-0.01937	480.48	0.1207
10	472.0	411.4	1.2098	-0.00737		
11	125.0	821.1	2.6144	-0.04280		
12	300.0	323.7	12.8429	-0.2848		
13	137.5	159.9	33.9166	0.8613		
14	191.0	246.1	12.4082	0.1693		
15	350.0	256.4	12.5646	-0.2373		

Layer	Estimated Permeability md	Std. Dev. Relative Magnitude	Estimated Permeability md	Std. Dev. Relative Magnitude
1	226.9	180.37	263.4	0.0168
2	261.4	332.44		
3	1229.	11.09		
4	111.8	61.14		
5	5.899	252.47		
6	50.55	13.05		
7	985.8	0.27	788.3	0.0086
8	68.12	0.26	61.64	0.0165
9	283.6	11.01	392.5	0.1910
10	333.6	111.95		
11	325.4	405.86		
12	316.9	1206.45		
13	315.3	2824.31		
14	315.5	2791.76		
15	315.6	1611.55		

**Fig 8.6 - SPE Problem 2 - Limited Data
Match of Pressure Measurements**

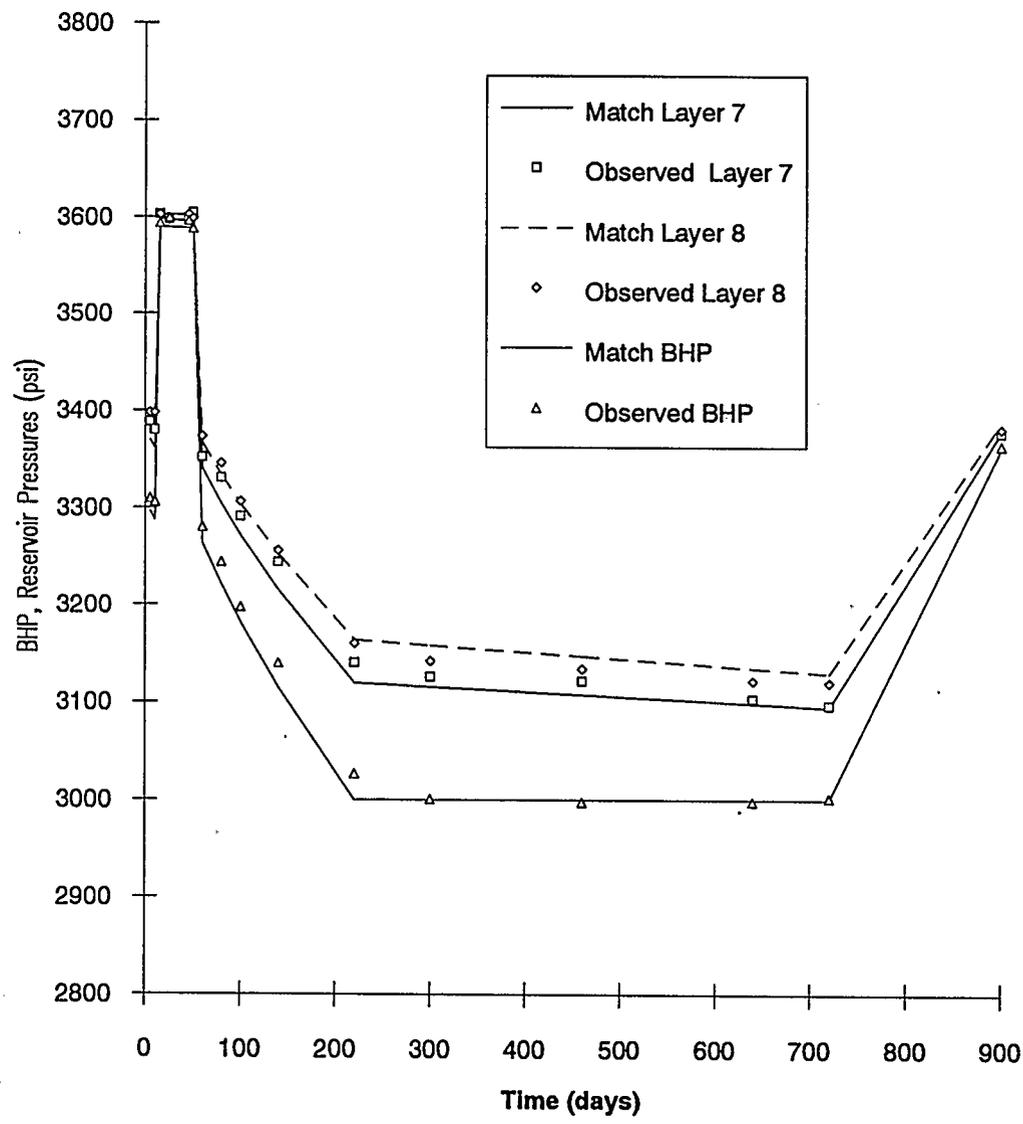
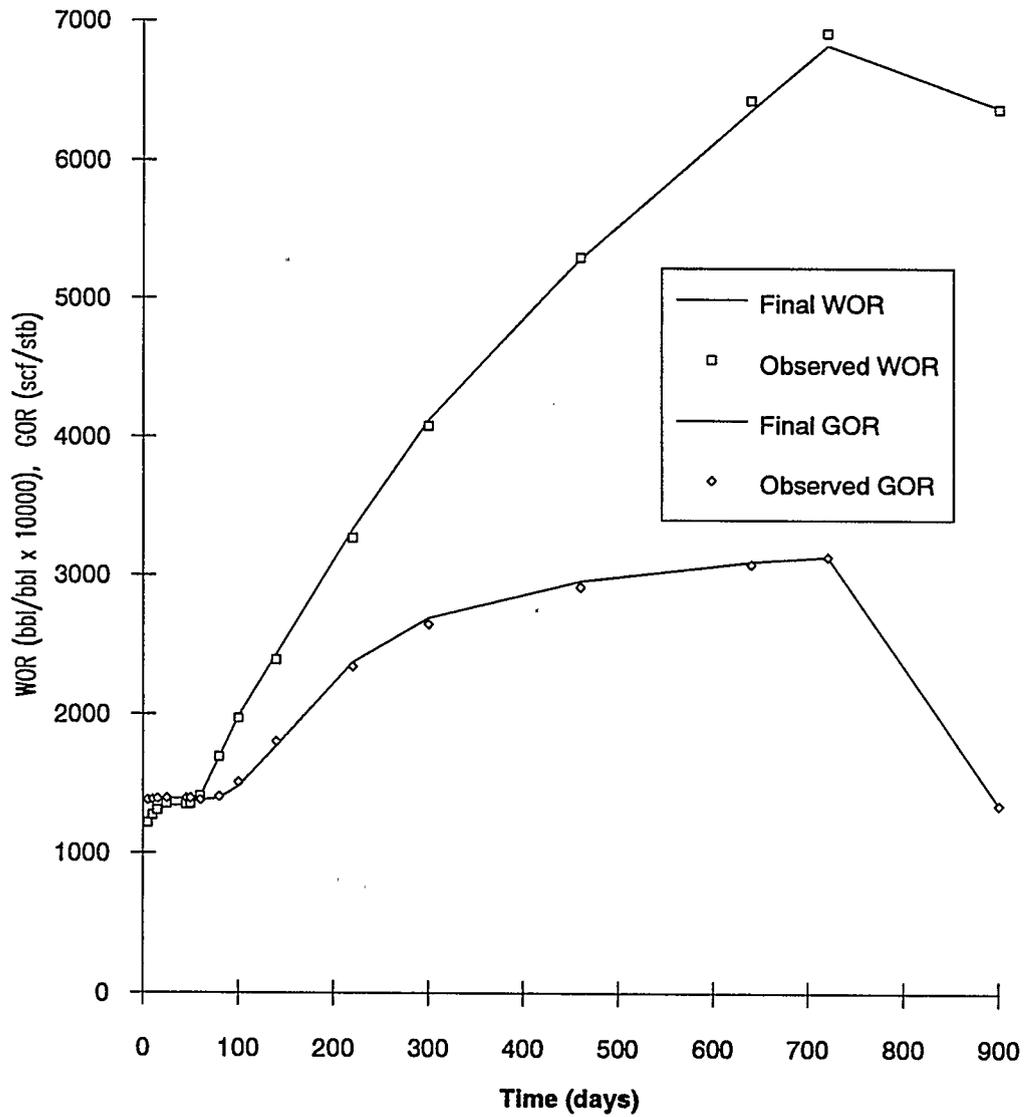


Fig. 8.7 - SPE Problem 2 - Limited Data
Match of WOR and GOR



8.5 Practical Guidelines

The practical application of the techniques demonstrated in this chapter can now be summarized. Using an initial parameterization of the reservoir, the automatic history matching model is run. If groups of parameters are highly correlated, i.e. the reservoir has been over-parameterized for the data available, then it will be noticed that the model will converge to a minimum value of the LS objective function before beginning to diverge as successively higher ranks of the H^{-1} are used. This will be an indication that a converged solution will be extremely difficult to obtain using all the parameters. At this point, the correlation matrix, eigenvalues and eigenvectors as well as the variances of the parameters computed at the current minimum of the LS objective function should be inspected. Adjacent zones of high variance should be combined. With the reduced zonation, the automatic history matching model can be rerun, often with the effect of further decreasing the LS objective function and reduction of the variance of the parameters.

CHAPTER 9

INCORPORATION OF PRIOR KNOWLEDGE AND CONSTRAINTS

In the numerical examples described so far, the postulated grid cell model was identical to the actual model used to generate the observation data.

However, when the automatic history matching simulator is utilized for actual field studies, the postulated grid cell model may not accurately represent the reservoir description since the geological interpretation of seismic data and well samples may be in error. In addition, the variation in rock properties may be such that the postulated grid cell model may not have enough detail to represent it. Several researchers such as Christie (1987) and White (1987) have attempted to use very fine grids with hundreds of thousands of cells to model reservoir heterogeneities. Unfortunately, with current computer hardware and numerical techniques, it is still not possible to use many thousands of parameters to describe the variation in porosity and permeability in reservoir rock both areally and vertically. We have to resort to a zonation approach whereby the reservoir is divided into a relatively small number of zones in each of which the parameters have constant values that have to be estimated. The shape and arrangement of these zones in the postulated model will be quite different than the true reservoir distribution. The

problem of optimally defining the zoning pattern is still unresolved and most published work involves a trial and error approach.

When the postulated model is very much different than the actual model, the resulting estimated parameter values could also be very different from the true values such that they become physically unrealistic. In such circumstances, it is helpful to incorporate into the algorithm prior knowledge of the most probable values of the parameters, and also to be able to impose constraints on the magnitudes of the values. This helps to influence the parameter search towards physically acceptable values.

9.1 Incorporating Prior Information using Bayesian Methodology

In many practical situations, even before we start estimating the unknown parameters from the field data, some information about the values of the parameters is often available from other sources or from general physical considerations. For example, from core data analysis and well test analysis, the estimated porosity and permeability values may be obtained as ϕ_m and K_m . It is reasonable to assume that the probable values of these parameters have normal distributions with means equal to the estimated values k_m and standard deviation σ_k . It was shown in Chapter 4, that using maximum likelihood estimation, the objective function could be written as,

$$S = S_{LS} + S_{prior} \quad (9.1)$$

where S_{LS} is defined by Eq. 4.5 and,

$$S_{prior} = (\mathbf{k} - \mathbf{k}_m)^T \mathbf{W}_k^{-1} (\mathbf{k} - \mathbf{k}_m) \quad (9.2)$$

Here \mathbf{W}_k is the covariance matrix of the probability distribution of the unknown parameters \mathbf{k} , and is a diagonal matrix with its elements being the reciprocals of σ_k^2 .

The Gauss-Newton method can be readily modified to include the prior term. At the j^{th} iteration,

$$\mathbf{k}^{(j+1)} - \mathbf{k}_m = \Delta \mathbf{k}^{(j+1)} + \mathbf{k}^{(j)} - \mathbf{k}_m \quad (9.3)$$

Therefore substituting in Eq. 9.2,

$$S_{prior} = (\Delta \mathbf{k}^{(j+1)} + \mathbf{k}^{(j)} - \mathbf{k}_m)^T \mathbf{W}_k^{-1} (\Delta \mathbf{k}^{(j+1)} + \mathbf{k}^{(j)} - \mathbf{k}_m) \quad (9.4)$$

Setting $\partial S_{prior} / \partial \Delta \mathbf{k}^{(j+1)}$ to zero, rearranging and combining with Eq. 5.8, the resulting Gauss-Newton equation is

$$\left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} \mathbf{C}(t_i) \mathbf{G}(t_i) + \mathbf{W}_k^{-1} \right] \Delta \mathbf{k}^{(j+1)} - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} (\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)) - \mathbf{W}_k^{-1} (\mathbf{k}^{(j)} - \mathbf{k}_m) \quad (9.5)$$

Note that since W_k^{-1} is simply a diagonal matrix, only the main diagonal of the Gauss-Newton matrix is modified. The effect of prior information can be deduced easily. If the user specified standard deviations σ_k are large, implying the user has little confidence in the prior estimate of k_m , then W_k^{-1} is small and it makes an insignificant contribution to the Gauss-Newton matrix. On the other hand, if the user has a strong belief that the prior estimates of k_m are representative of the actual reservoir values, then the user specified σ_k will be small, W_k^{-1} will be large and dominant in the Gauss-Newton matrix, and hence, $k^{(j+1)}$ will stay close to k_m .

Other authors such as Yang and Watson (1991), Chung and Kravaris (1990) have incorporated a weighting factor for the Bayesian term, and attempted to find rules to determine the optimum weighting factor.

In our work, we have not considered the problem of optimizing the weighting factor for the a priori estimates of the parameter values as there is really no rigorous approach for doing this. The selection of the weighting factor is very much a subjective process in that the weight is a function of how strongly the user believes in the a priori information he has supplied. If there is a significant amount of historical observation data available, the objective function will be dominated by the least squares contribution. The contribution of the a priori information to the objective function increases as the amount of data decreases. If the estimated

parameter values after regression are significantly different from the prior estimates, the user then either has to question his pre-conceptions or to support them by increasing the weighting factor for the prior information. The approach we have taken is to provide the user with this flexibility.

9.2 Incorporating Constraints using the Penalty Function

Approach

While prior information may be used to influence parameter estimates towards physically realistic values, there is no guarantee that the final estimates will not reach extremal values especially when the postulated model is not correct and there is a large amount of data available. The only way then to ensure that this does not happen is to restrict the feasible region by the use of parameter inequality constraints. For example, the range of values for porosity must be greater than 0 and less than 1, and the value of permeability must be greater than or equal 0.

Various methods of incorporating boundary constraints exist, such as the gradient projection method. In this work, we have chosen the penalty function method because it can be incorporated into the Gauss-Newton algorithm readily and it has a negligible computational overhead. The penalty function approach works well when the solution is expected to be in the interior of the feasible region, which is true in this application. The objective function is modified in such a way

that it remains almost unchanged in the interior of the feasible region, but increases dramatically as the solution approaches the constraints.

For each constraint, l , of the form,

$$b_l(\mathbf{k}) \geq 0 \quad (9.6)$$

a penalty function is assigned,

$$\xi(\mathbf{k}) = \frac{\alpha_l}{b_l(\mathbf{k})} \quad (9.7)$$

where α_l is a small specified positive constant that determines how close the parameter vector approaches the boundary. The objective function to be minimised is expanded to include the sum of the penalty functions thus,

$$S = S_{LS} + S_{prior} + S_{penalty} \quad (9.8)$$

where,

$$S_{penalty}(\mathbf{k}) = \sum_{l=1}^{n_b} \frac{\alpha_l}{b_l(\mathbf{k})} \quad (9.9)$$

if there are n_b constraints.

The Gauss-Newton method is further modified to include the penalty function. At the j^{th} iteration,

$$b_i^{(j+1)} = b_i^{(j)} + \frac{\partial b_i}{\partial k} \Delta k^{(j+1)} \quad (9.10)$$

Substituting into Eq. 9.9, and differentiating S_{penalty} with respect to $\Delta k^{(j+1)}$,

$$\frac{\partial S_{\text{penalty}}}{\partial \Delta k^{(j+1)}} = \sum_{i=1}^{n_b} -\alpha_i (b_i^{-2} - 2b_i^{-3} \frac{\partial b_i}{\partial k} \Delta k^{(j+1)}) \frac{\partial b_i}{\partial k} \quad (9.11)$$

Setting this to zero, rearranging and adding to Eq. 9.5, we obtain the final form of the modified Gauss-Newton iterative equation,

$$\left[\sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} \mathbf{C}(t_i) \mathbf{G}(t_i) + \mathbf{W}_k^{-1} + \sum_{i=1}^{n_b} \frac{\alpha_i}{b_i^3} \left(\frac{\partial b_i}{\partial k} \right)^2 \right] \Delta \mathbf{k}^{(j+1)} - \sum_{i=1}^{n_o} \mathbf{G}^T(t_i) \mathbf{C}^T(t_i) \mathbf{Q} (\hat{\mathbf{y}}(t_i) - \mathbf{y}(t_i)) - \mathbf{W}_k^{-1} (\mathbf{k}^{(j)} - \mathbf{k}_m) + \frac{1}{2} \sum_{i=1}^{n_b} \frac{\alpha_i}{b_i^2} \frac{\partial b_i}{\partial k} \quad (9.12)$$

This is of the form $\mathbf{A} \Delta \mathbf{k}^{(j+1)} = \mathbf{b}$ and can be solved for $\Delta \mathbf{k}^{(j+1)}$.

With constraints in the form of,

$$\begin{aligned} b_i(k_i) - k_{(\max, i)} - k_i &\geq 0 \\ b_{(i+1)}(k_i) - k_i - k_{(\min, i)} &\geq 0 \end{aligned} \quad (9.13)$$

the penalty function terms appear only in the main diagonal of the \mathbf{A} matrix.

The general effect of the penalty terms can easily be identified. In the interior of the feasible region, α_1/b_1 is small resulting in a small contribution to the main diagonal of the \mathbf{A} matrix. The values of $\Delta \mathbf{k}^{(j+1)}$ will not be affected. Near the l^{th} constraint, α_1/b_1 is large, dominating the diagonal element of the parameter to which the constraint is applied. Then Δk_1 will be very small.

The penalty function method requires little modification of the Gauss-Newton algorithm and is superior to the box type constraint method. As discussed by Yang et al. (1987), the step direction in the box type method is determined independently of the constraints, and the step length is limited if the bounds are violated. This could result in early termination of the algorithm when a boundary is encountered. With the penalty function approach, the step direction is confined to the feasible region only.

In addition, as the magnitude of the main diagonal of the Gauss-Newton matrix is increased by contributions from the Bayes function and the penalty function, the eigenvalues of the matrix become larger, and the ratio of the largest to the smallest eigenvalue (the condition number) is decreased. This increases the stability of the nonlinear regression algorithm.

9.3 Model Application

As explained earlier, in an actual application, it is extremely unlikely that the grid cell model used to describe the reservoir will correctly represent the underground reservoir structure with all its variation in areal extent and parameter distribution. Furthermore, the equations used to calculate fluid flow and phase behaviour in the numerical model are also approximations to the true physical processes. Thus the postulated simulation model can be quite different from the actual physical model. Regression analysis to find parameter values that cause the postulated model to match the actual field data may result in parameter values that are unrealistically large or small. In order to influence the parameter search towards a priori beliefs of the most probable values of the parameters, the incorporation of prior information via a method such as Bayesian estimation is very useful. In order to restrict the estimated parameter values within feasible limits, the incorporation of constraints via a method such as the use of penalty functions is needed.

There are thus two major problems associated with automatic history matching in reservoir simulation. The first is the correct representation of the reservoir with a grid cell model with a limited number of zones, and the second is the regression analysis necessary to find the parameter values that result in a "best" fit. The automatic history matching model presented by Tan and Kalogerakis (1991a, 1991b) and in

this thesis was designed to solve the second problem. Unfortunately as yet, there is no rigorous analysis of the first problem of accurate model description, and the design of postulated models to represent the reservoir is at best a trial and error process, utilizing as much prior information as possible from other sources such as seismic mapping, and geological analysis. In spite of this, the automatic history matching model is still very useful as a tool to help predict the existence of impermeable boundaries, suggest the possibility of reservoir extensions, provide estimates of the volumes of oil, gas and water in place, and at the same time limit estimates of the parameters to within realistic values. The two hypothetical case studies presented here are intended to demonstrate these.

9.4 Case Study IV

The actual model used to generate the observation data and the postulated models used to match the data are shown in Figure 9.1. This case study is a two-dimensional three-phase problem. In the actual model, an impermeable boundary extends halfway across the reservoir partially separating Well #1 in grid cell (6,3) and Well #3 in cell (10,4) from Well #2 in cell (7,5). Well #1 is a water injector while the other two wells are producing wells. The effect of the impermeable

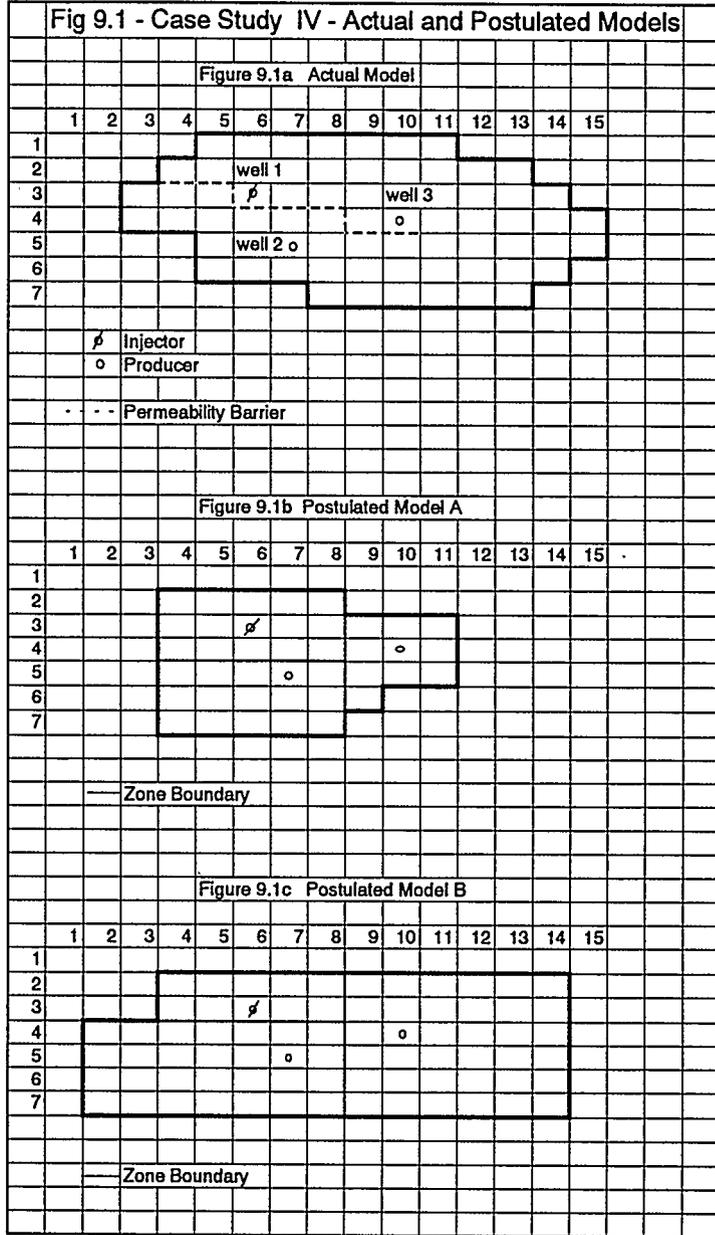


Fig. 9.2 - Case Study 9.1 - Porosity, Permeability Distributions

Figure 9.2a Porosity Distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1					0.14	0.14	0.14	0.14	0.14	0.14	0.14				
2				0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13		
3			0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	
4			0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
5					0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
6					0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
7								0.08	0.08	0.08	0.08	0.08	0.08		

Figure 9.2b Permeability Distribution

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1					1000	1000	1000	1000	1000	1000	1000				
2				1000	1000	1000	1000	1000	1000	1000	1000	50	50		
3			500	500	500	1000	1000	1000	1000	1000	1000	50	50	50	
4			500	500	500	500	500	500	1000	1000	1000	50	50	50	50
5					500	500	500	500	500	1000	1000	50	50	50	50
6					500	500	500	500	500	1000	1000	50	50	50	
7								500	500	1000	1000	50	50		

X direction permeability same as Y direction
 Permeability Barrier overrides values

Table 9.1: Description of Actual Model for Case Study IV

Grid dimension in X and Y directions	300 ft
Thickness of cells	50 ft
Top depth of cells	9035 ft
Initial Pressure at Depth 9035 ft is	3600 psia
Depth of Gas-Oil, Water-Oil Contacts	9035, 9300 ft
Well #1 - Injector II = 20	
Water injection rate = 1500 bbls/d	
Injection pressure = 50000 psi	
Well #2 - Producer Wellbore Index = 10	
Maximum Oil Rate = 1000 bbl/d	
Minimum Bottom Hole Pressure = 2000 psia	
Well #3 - Producer Wellbore Index = 5	
Maximum Oil Rate = 500 bbl/d	
Minimum Bottom Hole Pressure = 2000 psia	

Table 9.2: Relative Permeability Table for Case Study IV			
SL	KRG	KROG	PCG
0.22	1.00	0.00	3.9
0.30	0.8125	0.00	3.5
0.40	0.5000	0.00	3.0
0.50	0.4200	0.00	2.5
0.60	0.3400	0.00	2.0
0.70	0.2400	0.02	1.5
0.80	0.1000	0.10	1.0
0.90	0.0220	0.33	0.5
0.96	0.005	0.60	0.2
1.00	0.000	1.0	0.0
SW	KRW	KROW	PCW
0.22	0.00	1.0000	7.0
0.30	0.066	0.4000	4.0
0.40	0.15	0.1250	3.0
0.50	0.24	0.0649	2.5
0.60	0.33	0.0048	2.0
0.70	0.49	0.0024	1.5
0.80	0.66	0.0000	1.0
0.90	0.83	0.0000	0.5
0.95	0.915	0.0000	0.25
1.00	1.00	0.0000	0.0

boundary is to delay the response of Well #2, which is located closer, to the water injection in Well #1. The water injection rate is not sufficient to maintain voidage replacement and the producing wells soon drop below saturation pressure. The gas-oil ratio increases as free gas is formed in the reservoir. The water-oil ratio of Well #3 increases first, until it reaches a value of 3.05 when the well is shut in at 480 days. The water front then sweeps around the no-flow boundary to Well #2. The pressure of Well #2 begins to increase when Well #3 is shut in as then there is over water injection. The producing gas-oil ratio of this well reaches a maximum of 4482 scf/stb and then decreases as the free gas goes back into solution. The water-oil ratio of Well #2 just begins to increase when the base run is terminated at 720 days.

As the actual reservoir is partially divided by the impermeable boundary, and the water front has not yet broken through at Well #2 at the termination of the base run, the estimation problem is thus more difficult. This is typical of actual reservoir situations where the reservoir parameters have to be estimated even though the reservoir has not been completely produced.

The reservoir dimensions are described in Table 9.1. The reservoir is initially undersaturated with a connate water saturation. No aquifer is present. The permeability and porosity distributions are also detailed in Figure 9.2. These

distributions were set up with irregular regions so that parameter zones would encompass several permeability or porosity values. The reservoir fluid properties are identical to those used in the Second SPE Comparative Solution Problem (1986). The relative permeability values are shown in Table 9.2. The well constraints are also detailed in Table 9.1. The transmissibility between any pair of cells (i_1, j_1) , (i_2, j_2) may be calculated in two ways, using the series average of the two cell permeabilities or using the permeability of cell (i_1, j_1) . The second option is used here. This allows no-flow boundaries to be specified. The pressures of the cells in which the wells are located, the water-oil ratios, gas-oil ratios and the oil rates of the producing wells are recorded every thirty days in the base run and used as observation data to be matched. A small amount of noise was added using random numbers distributed normally with a mean of zero and a standard deviation of the measurement errors of 0.001 as discussed previously in Chapter 8.

9.4.1 Match of Actual Model

The actual grid model was first used to see if there was sufficient observation information to identify the original distribution of porosity and permeability. The reservoir was divided into fourteen zones, and initial guesses for porosity and permeability values were 0.1 and 200 md respectively. These initial guesses were used for all subsequent runs. The

automatic history matching model was first run without any prior estimates of the porosity and permeability available. Column 4 of Table 9.3 shows the final estimates of the fourteen parameter values. The objective function was reduced from 0.4409×10^9 to 0.119×10^5 in 26 Gauss-Newton iterations. Comparing with the original values, it can be seen that the estimates of the porosity values are not acceptable. Thus a local minimum must have been reached.

The simulator was then rerun using Bayesian estimation. The prior estimate of the porosity values was set at 0.1 with a standard deviation of 0.001. The prior estimate of the permeability values was set at 200 md with a standard deviation of 50 md. The LS objective function was reduced from 0.4409×10^9 to 0.276×10^3 in 30 Gauss-Newton iterations. The Bayes objective function was 0.356×10^4 . The revised estimates are shown in column 5 of Table 9.3. It can be observed that these are much closer to the actual values. This shows the beneficial effect of Bayesian estimation in stabilizing the parameter search. The Bayesian contribution to the composite objective function influenced the parameter search to the correct values.

The run terminated at a minimum of the composite objective function where the Bayes contribution was significantly higher than the LS contribution. In order to recover the correct values, the weighting factor for the Bayes objective function was set to zero, effectively removing its

Table 9.3: Parameter Match of Actual Model for Case Study IV					
Zones	Parameter	Actual	LS	Bayes	Final
1	Porosity	0.14	0.1898	0.1321	0.1398
2	Porosity	0.13	0.0453	0.1297	0.1300
3	Porosity	0.12	0.1785	0.1234	0.1201
4	Porosity	0.11	0.0873	0.1095	0.1099
5	Porosity	0.10	0.1532	0.1005	0.1000
6	Porosity	0.09	0.0176	0.0878	0.0900
7	Porosity	0.08	0.1127	0.0837	0.0799
8	X Permeability	1000.	1037.	1026.	1001.
9	X Permeability	500.	494.2	485.1	499.
10	X Permeability	50.	66.93	52.1	50.07
11	Impermeable Boundary	0.	0.0664	0.0088	0.001
12	Y permeability	1000.	1024.	934.5	999.2
13	Y permeability	500.	313.4	503.8	498.7
14	Y permeability	50.	46.55	49.39	49.98

contribution, and the run restarted from the previously converged values. The final estimates are shown in column 6 of Table 9.3 and are very close to the actual values. The LS objective function was reduced to 6.4 in 4 Gauss-Newton iterations.

9.4.2 Postulated Model A

The postulated model used to match the observation data is purposely made different from the actual model. The relative position of the wells remain the same, but the areal extent of the reservoir is modified reflecting the fact that in practice, the true reservoir boundaries are rarely known. In this case, the postulated model is shifted slightly sideways and made smaller. The intent here is to see if the results of the parameter estimation using an incorrect postulated model could be used as a guide to revising the postulated model towards a more accurate representation of the reservoir. In particular, we want to see if the presence of an impermeable boundary can be detected and whether the correct in-place volumes of the fluids can be estimated. The zonation used reflects the normal approach towards zonation of a reservoir such as this with only three wells. The postulated model as shown in Figure 9.1b is divided into three porosity zones and nine permeability zones as detailed in Table 9.4. Each well is allocated a zone for the porosity parameter. The permeability zones are set up to find the average permeability

in the x and y directions for each of the porosity zones as well as to find the inter-zonal permeabilities as an impermeable barrier is suspected.

Initial guesses for the porosity and permeability values were once again 0.1 and 200 md respectively. The automatic history matching model was run initially without any prior estimates of porosity and permeability, and without any constraints. The objective function was reduced from 0.352×10^9 to 0.287×10^6 in 38 Gauss-Newton iterations. The estimated parameter values are given in column 4 of Table 9.4. It is difficult to provide correct estimates to compare against as the postulated zones encompass varying proportions of the original model zones. Nevertheless, we would expect porosity values about 0.1 and permeability values of approximately 1000, 500 and 50 md. Inspecting the calculated values, we see porosity values varying from a low of 0.045 to a high of 0.378, some permeability values higher than 2000 md, one as high as 18430 md.

Figures 9.3 to 9.11 show the match of water-oil ratios, gas-oil ratios, oil rates for Wells #2 and #3, and the match of pressures for all three wells. The profiles using the initial guesses for porosity and permeabilities are also shown. It was at first thought that the matches were not satisfactory based on the magnitude of the objective function at the end of the regression and the differences between the actual and calculated parameter values. However the figures

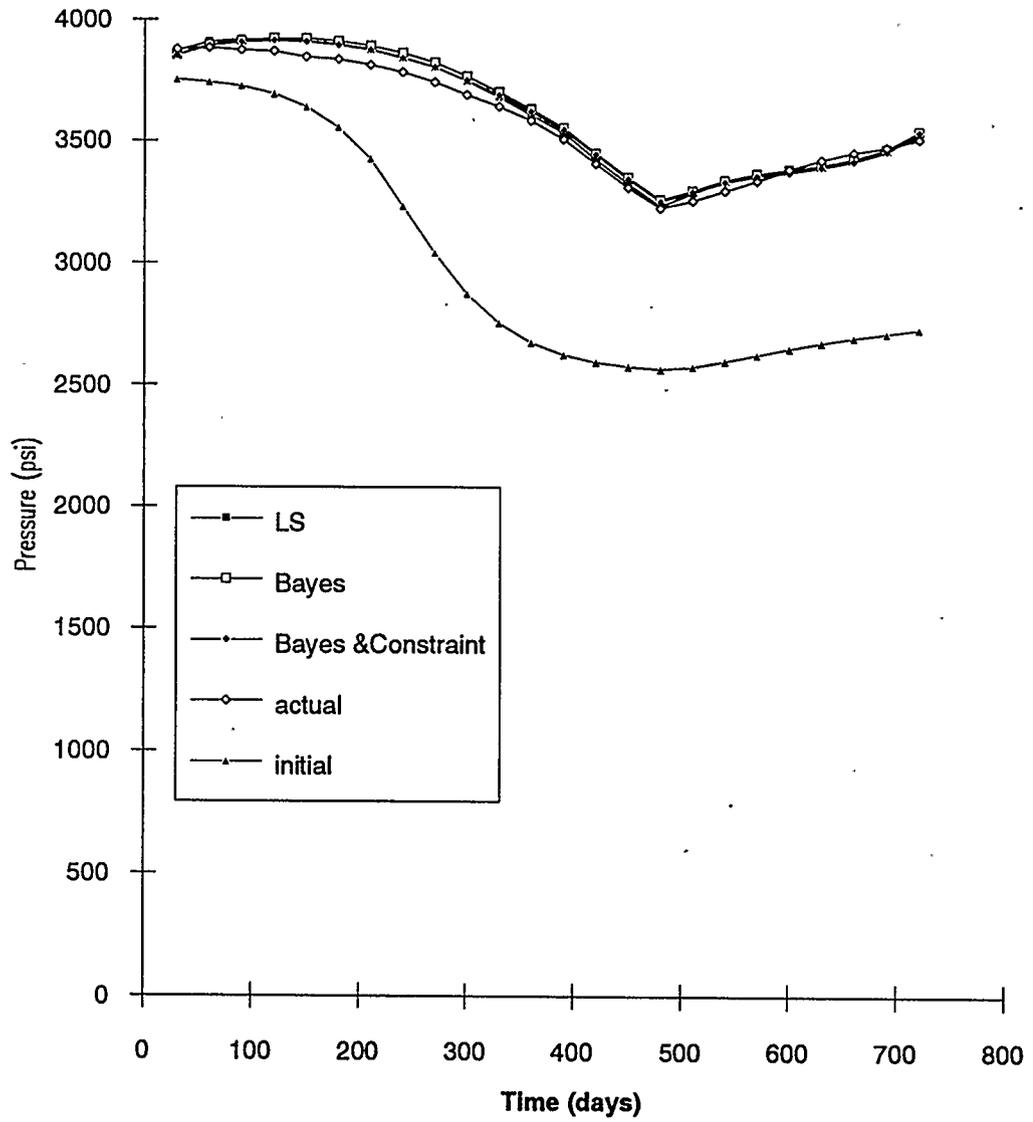
Table 9.4: Case Study IV - Postulated Model A - Parameter Estimates					
Zones	Parameter	Region	LS	Bayes	Bayes & Constraint
1	Porosity	(4-8,2-4)	0.045	0.078	0.078
2	Porosity	(4-8,5-7)	0.235	0.2487	0.2473
3	Porosity	(9-11,3-6)	0.378	0.3014	0.3094
4	X Permeability	(4-7,2-4)	2131.	744.8	717.2
5	X Permeability	(4-7,5-7)	74.	77.	90.83
6	X Permeability	(8,3-4)	1722.	783.	1092.
7	X Permeability	(8,5-7)	71.	66.	54.6
8	X Permeability	(9-11,3-6)	4132.	2513.	2557.
9	Y Permeability	(4-8,2-3)	185.	1340.	1732.
10	Y Permeability	(4-8,4)	1.6	2.311	2.106
11	Y Permeability	(4-8,5-7)	18430.	5013.	3000.
12	Y Permeability	(9-11,3-6)	67.	73.4	87.56

show that the predicted profiles compare very closely to the actual profiles and the match is indeed quite acceptable.

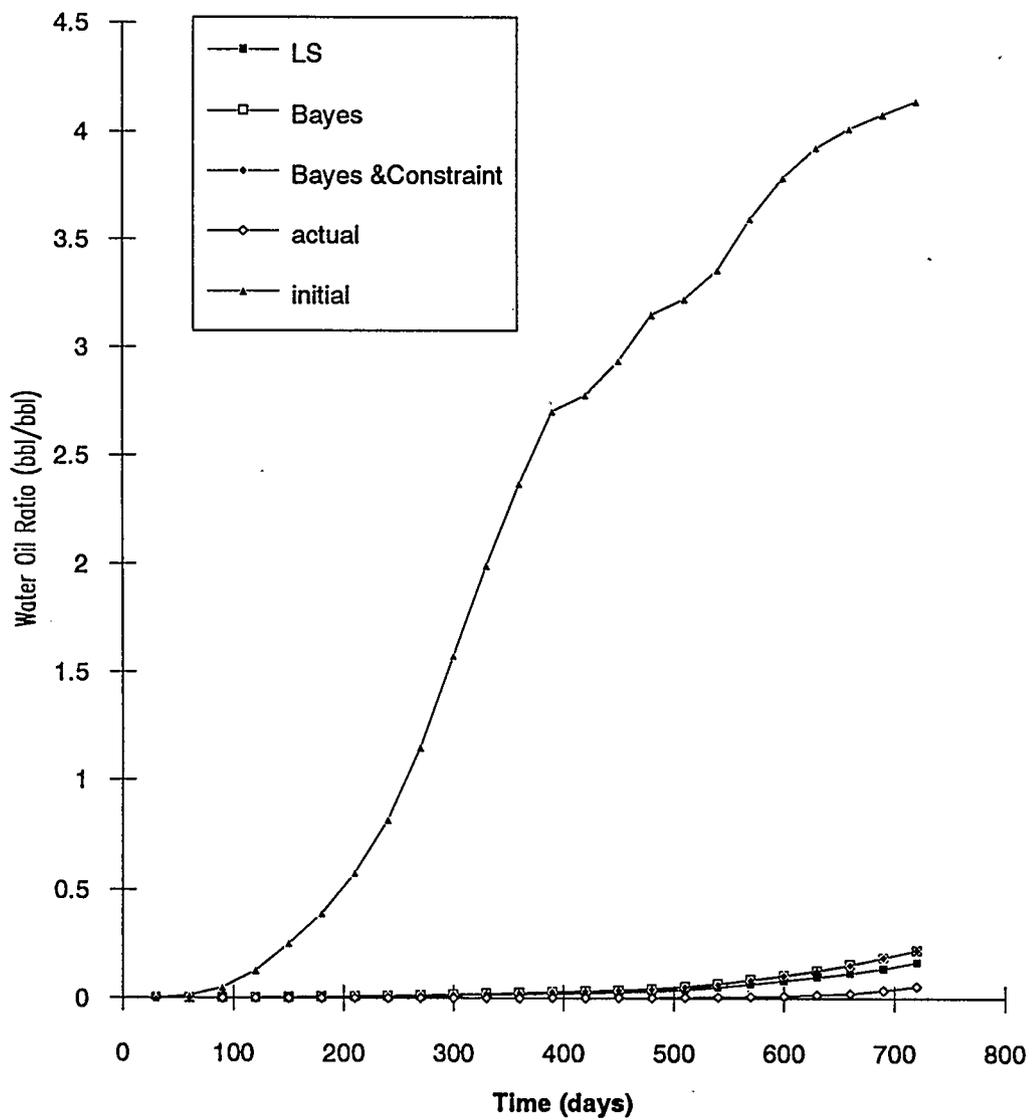
If we compare the pore volumes of the zones to the approximate drainage volumes of the wells in the original model, we can see that the volume of the zone allocated to Well #1 should be smaller, the volumes allocated to Wells #2 and #3 should be larger. The estimated values therefore have provided some indication that the reservoir volumes should be re-allocated and in particular the reservoir extent should be increased for Well #3. The initial oil in-place for the original model was 4,297,353 bbls. With the initial guess of porosity, the postulated model had an oil in-place of 2,252,872 bbls. With the final parameter estimates, the oil in-place was 4,494,478 bbls which is close to the true value. The inter-zonal permeability between Well #1 and Well #2 is also small, indicating restriction in flow between the two wells.

The unrealistic high values of some of the permeabilities provide an indication that the postulated model is not an accurate representation of the true reservoir. In order to influence the parameter search towards more realistic parameter values, prior estimates of the porosity values were supplied with a mean of 0.1 and a standard deviation of 0.01, and prior estimates of permeabilities were given with a mean of 200 md and a standard deviation of 50 md. No constraints were imposed. The results of the regression run using

**Fig 9.3 - Case Study IV - Model A
Well #1 - Match of Pressure**



**Fig 9.4 - Case Study IV - Model A
Well #2 - Match of Water Oil Ratio**



**Fig 9.5 - Case Study IV - Model A
Well #2 - Match of Gas-Oil Ratio**

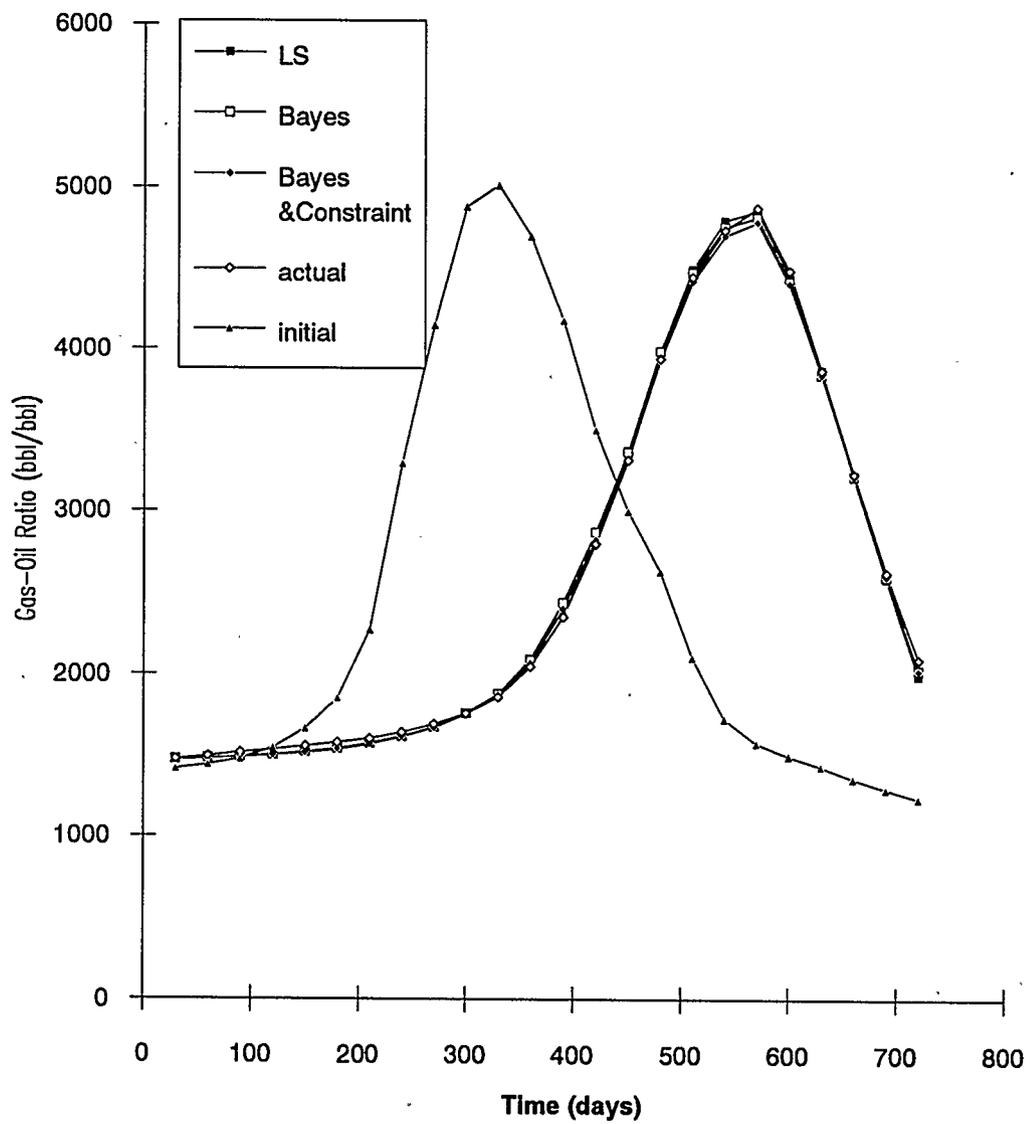
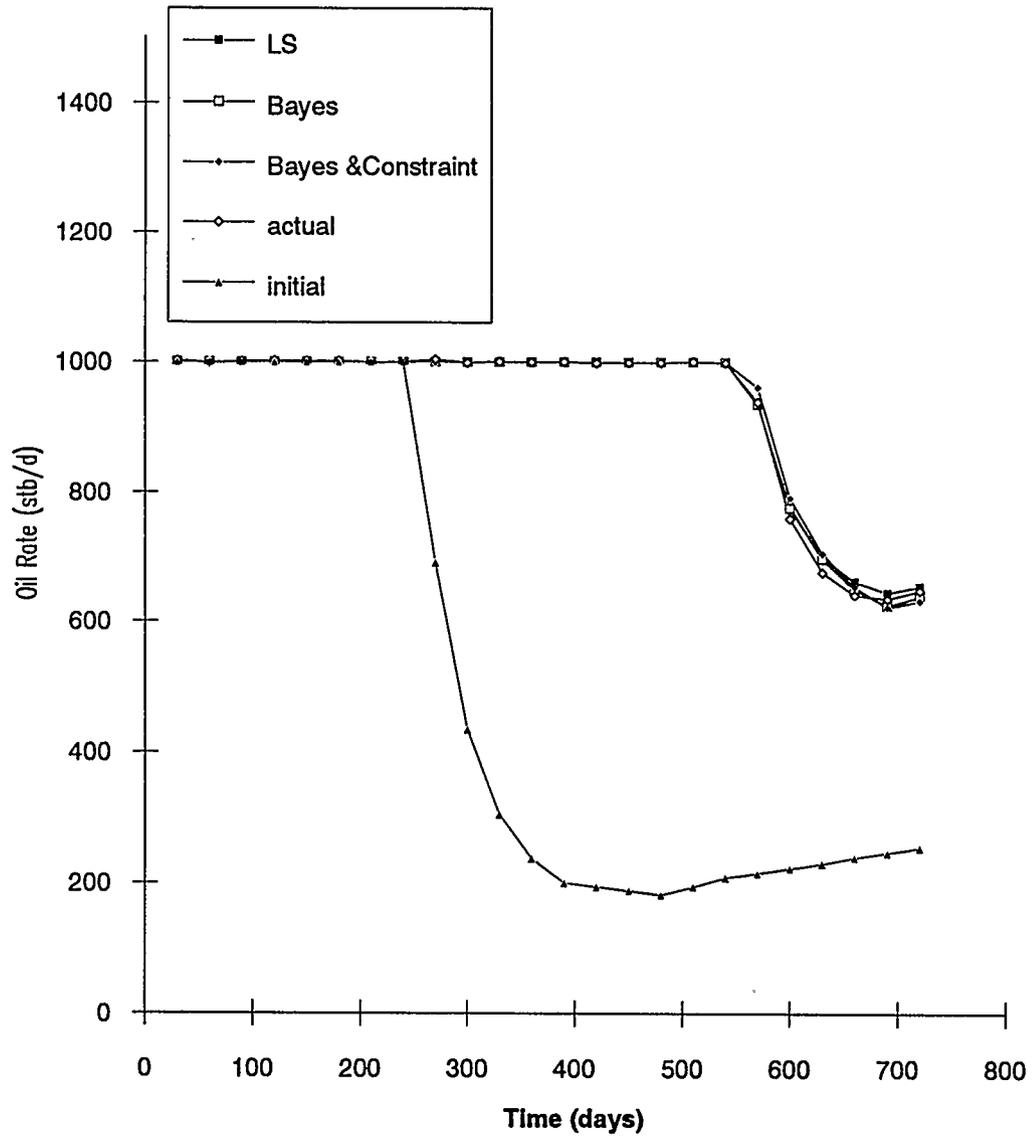
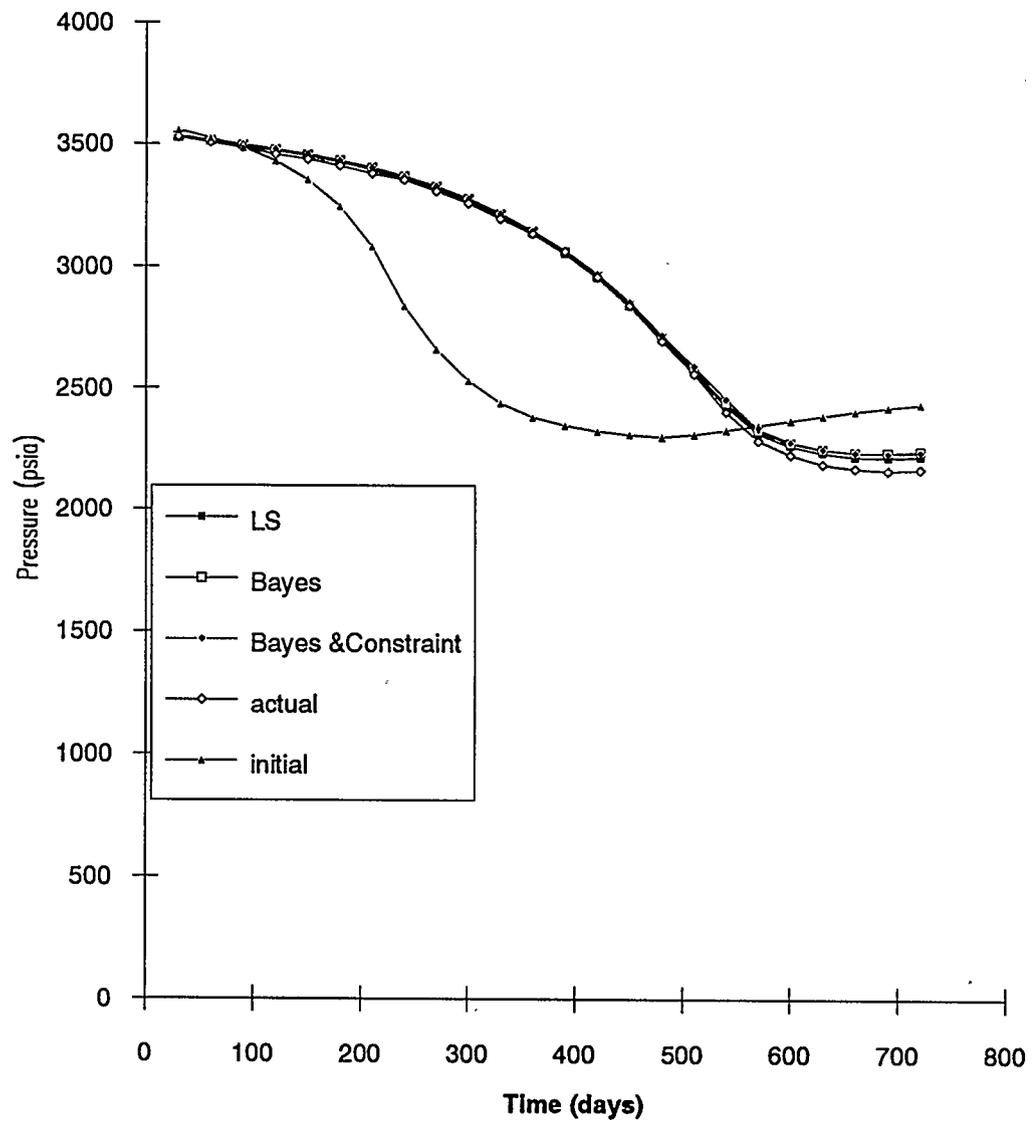


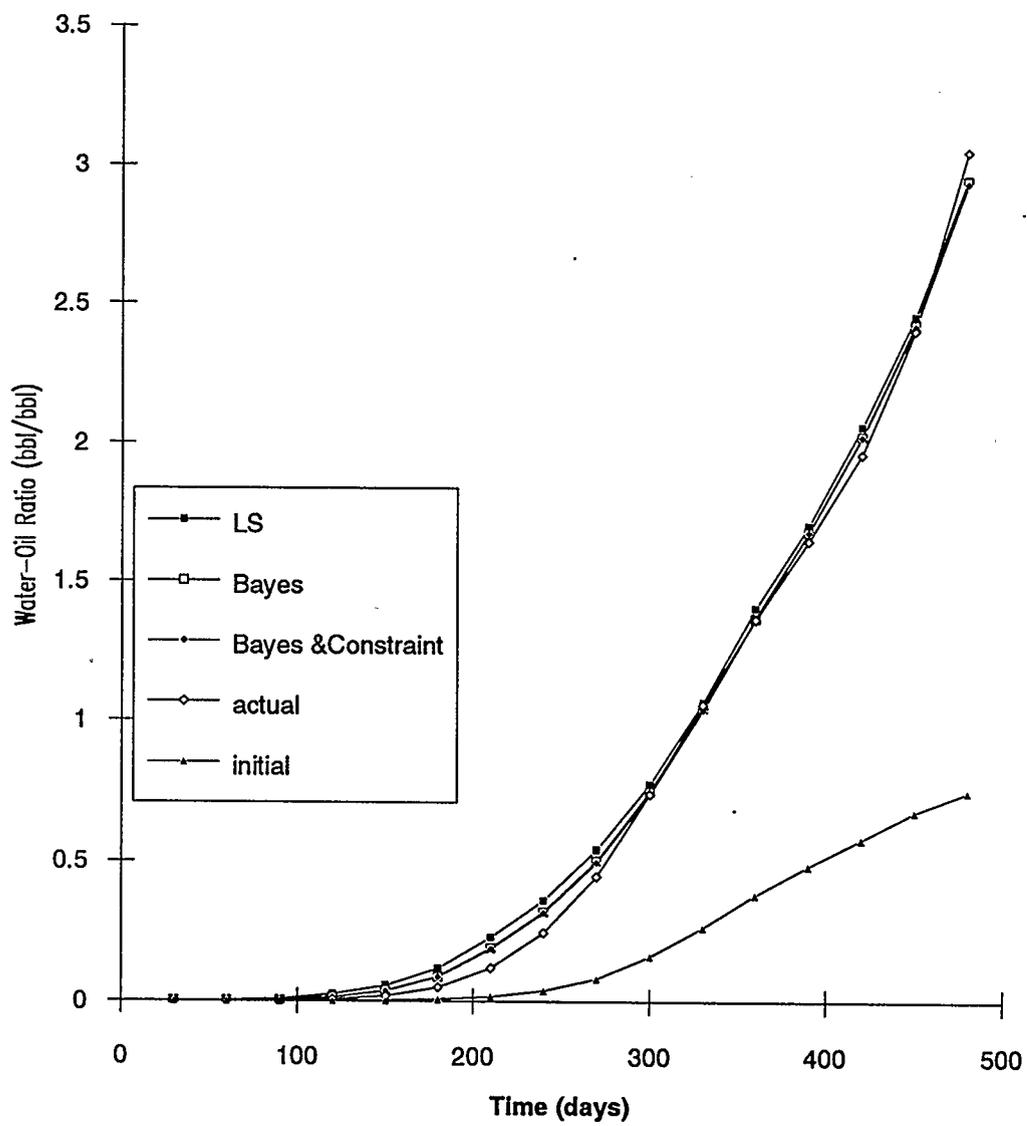
Fig 9.6 - Case Study IV - Model A
Well #2 - Match of Oil Rate



**Fig 9.7 - Case Study IV - Model A
Well #2 - Match of Pressure**



**Fig 9.8 - Case Study IV - Model A
Well #3 - Match of Water-Oil Ratio**



**Fig 9.9 - Case Study IV - Model A
Well #3 - Match of Gas-Oil Ratio**

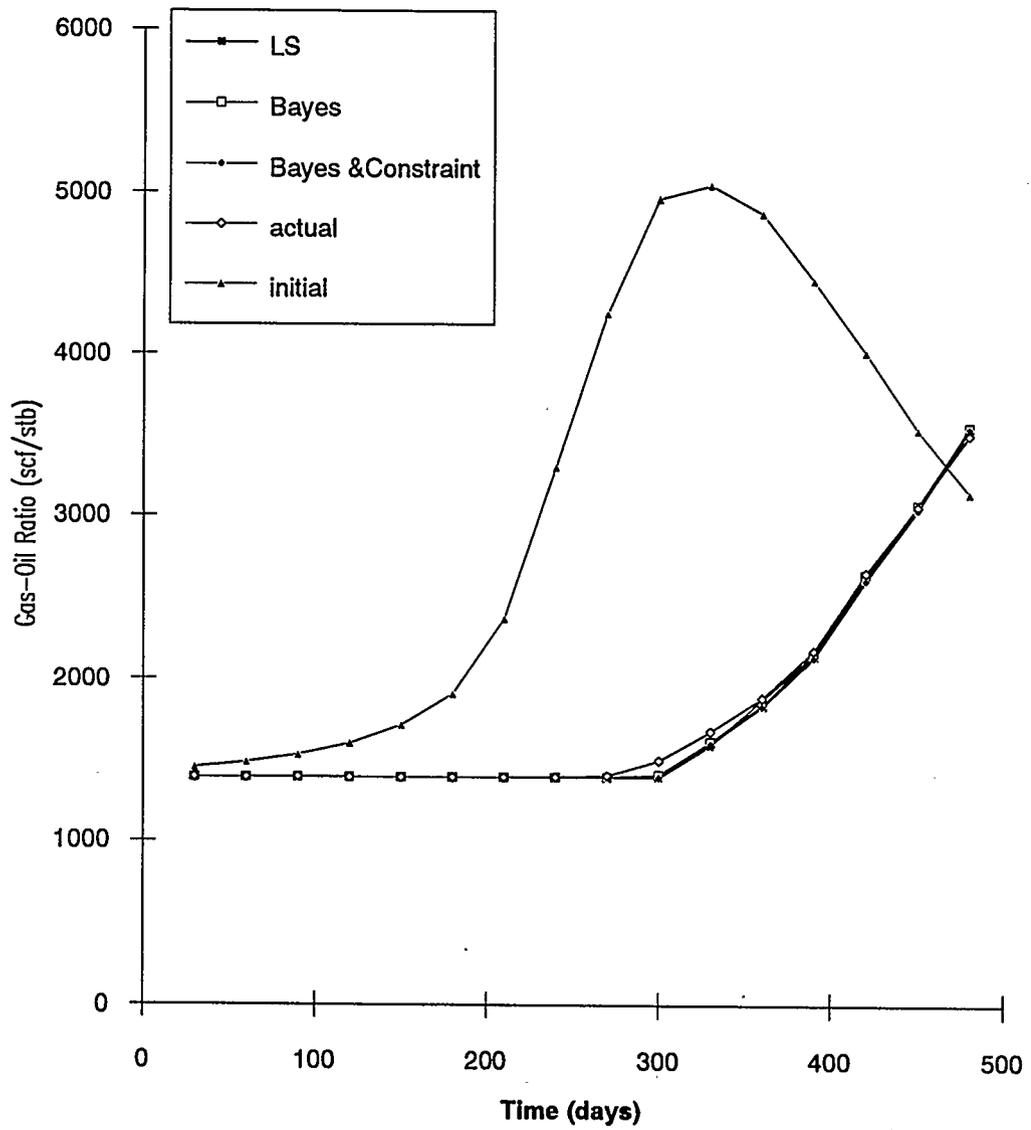
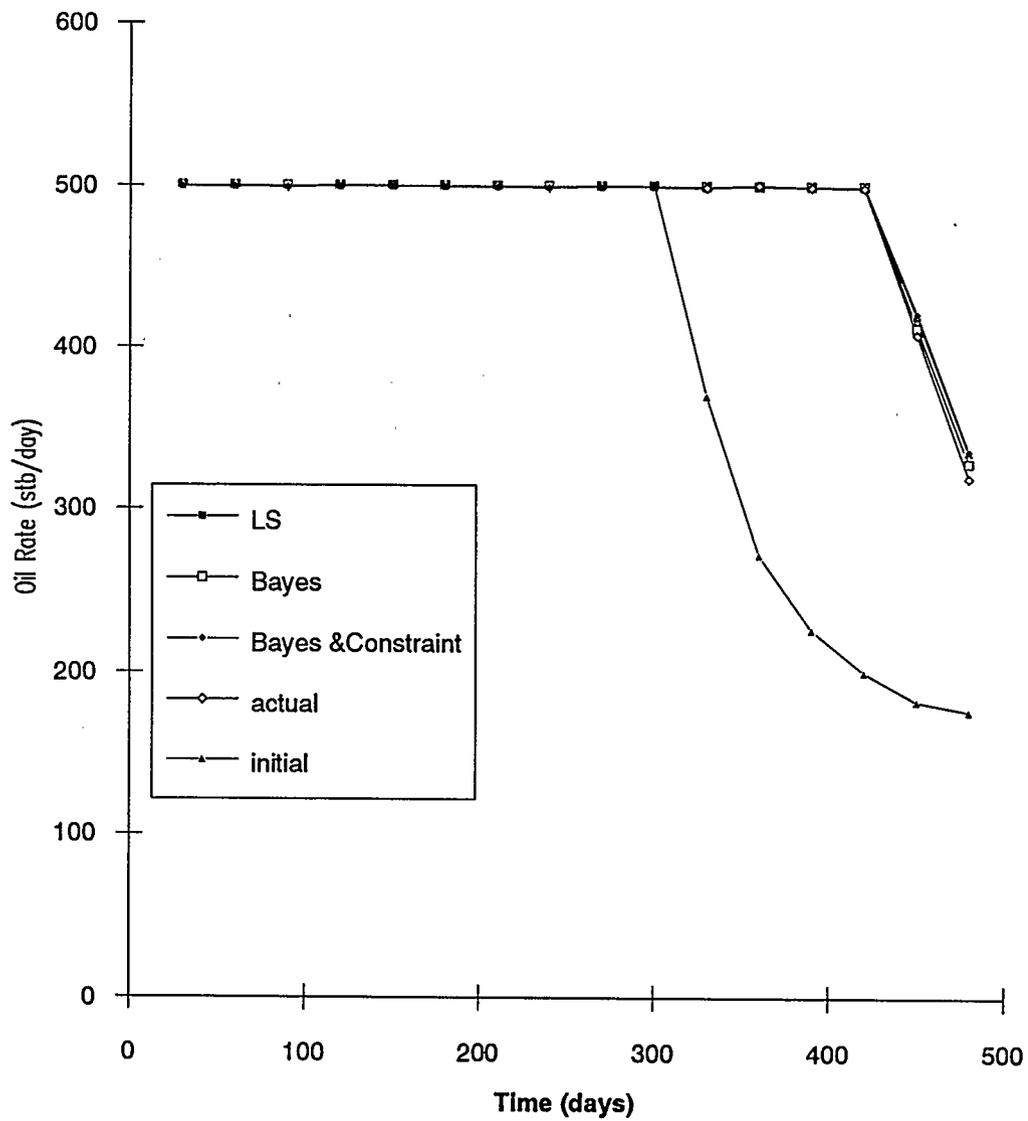
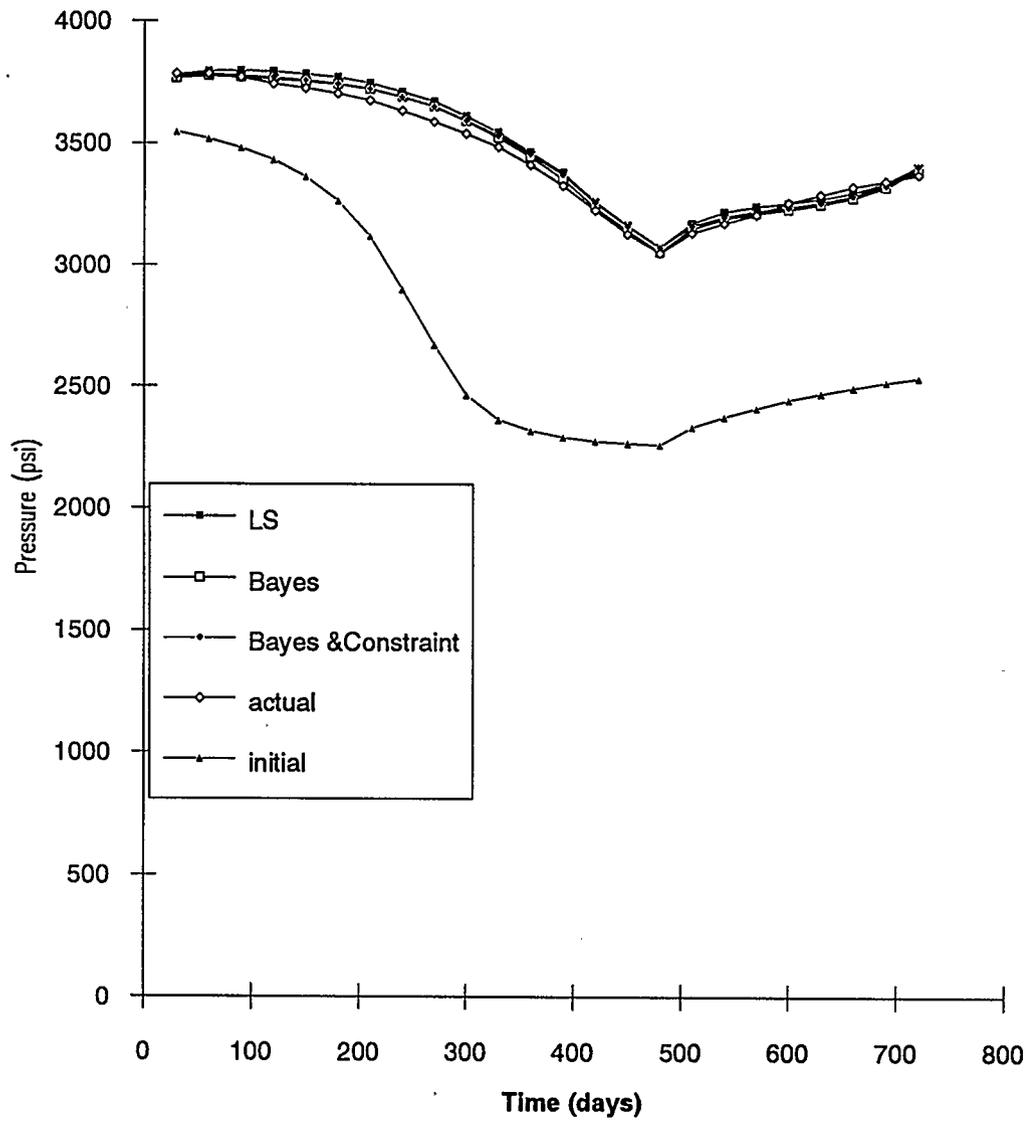


Fig 9.10 - Case Study IV - Model A
Well #3 - Match of Oil Rate



**Fig 9.11 - Case Study IV - Model A
Well #3 - Match of Pressure**



Bayesian estimation and the same initial guesses are given in column 5 of Table 9.4. The LS objective function was reduced from 0.352×10^9 after 25 iterations to 0.3118×10^6 with a Bayes function value of 0.12849×10^5 . Considering the relative magnitude of the LS and Bayes function values, it is clear that the observation data has more effect on the final parameter estimates than our prior estimates. The use of prior information has not changed significantly the porosity estimates. They still suggest the same reservoir volume modifications. The estimated oil in-place is 4,460,968 bbls. The maximum value of the permeability estimates however has been reduced to 5013 md from 18430 md. This shows that adding prior information using Bayesian estimation helps to reduce the variation in estimated values, providing values that are more realistic. However it can be termed as a "soft" constraint in that it does not impose an absolute limit on the values that can be attained during the parameter search.

When it is required that the parameter estimates do not exceed specified minimum and maximum values, it is necessary to use a method such as the penalty function method to limit the parameter values within the constraint values. In the next run, in addition to supplying prior estimates, we set the constraints for porosity such that the minimum value is 0.001 and the maximum value is 0.5, the minimum value for permeability to be 0.001 md and the maximum value as 3000 md. After 22 iterations, the objective function was reduced from

0.352×10^9 to 0.3058×10^6 with a Bayes function value of 0.74×10^4 . The estimated parameter values supplied in column 6 of Table 9.4. The highest permeability value for the zones has been further reduced to the limit of 3000 md. The estimated value for the oil in-place is 4,493,352 bbls. The matches of the observation data for the runs with Bayesian estimation, with Bayesian and constraints, are also shown in Figures 9.3 to 9.11. It is obvious that the matches are similar to the run without Bayesian estimation and constraints and yet with more realistic permeability values.

9.4.3 Postulated Model B

The results of the previous postulated model indicated the boundaries of the three porosity zones should be changed in order to obtain more reasonable values for the porosity. The second postulated model was set up as shown in Figure 9.1c, with the reservoir extent enlarged and the porosity zones assigned to each well modified based on the prior results such that the average porosity is about 0.1. Thus the area around Well #1 is reduced, and the areas around Wells #2 and #3 enlarged. Comparing the resulting reservoir description with the actual reservoir in Figure 9.1a, we can see that we are getting closer to the actual model. However there is no analysis that can be performed on the results of the prior run that would indicate the porosity zonation and permeability arrangement that exists in the actual model.

Using this postulated model and a porosity and permeability zonation similar to postulated model A, the automatic history matching model was run initially without any prior information and constraints. The final parameter estimates are shown in column 4 of Table 9.5. The LS objective function was reduced from 0.124×10^9 to 0.8115×10^6 . The porosity values are reasonable, but one of the permeability values reaches an unrealistic value of 27960 md. The inter-zonal permeability between Wells #1 and #2 is very small, with a value of 0.0067 md. It could be argued that this too is an unrealistic value, but the analysis of reservoir behaviour indicates that this permeability barrier is the only explanation for the delayed response of pressure and water-oil ratio of Well #2 to the water injection in Well #1. The reduced permeability of 13.4 md between Wells #2 and #3 could be a clue that though there is no complete barrier between these two wells, the permeability barrier between Wells #1 and #2 could extend further between Wells #2 and #3. The initial oil in-place estimate was 4,414,501 bbls which is close to the true value of 4,297,353 bbls.

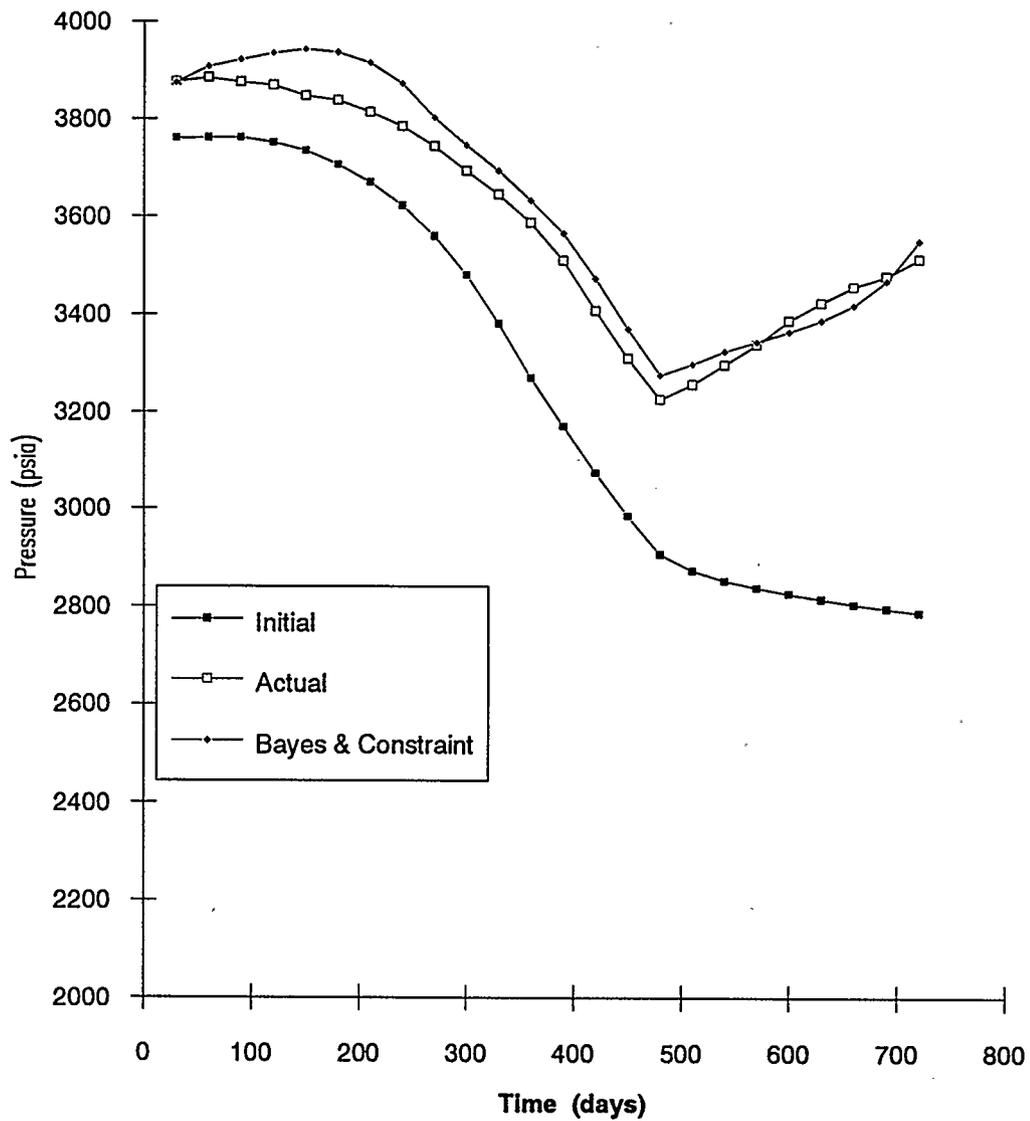
Next, prior estimates of the porosity, given as 0.1 ± 0.01 , and of permeability given as 200 ± 50 md are supplied to the model. Starting again from the same initial guess, the revised estimates are given in column 5 of Table 9.5. The LS objective function value is reduced from 0.124×10^9 to 0.785×10^6 with a Bayes function value of 0.539×10^5 . The initial

Table 9.5: Case Study IV - Postulated Model B - Parameter Estimates					
Zones	Parameter	Region	LS	Bayes	Bayes & Constraint
1	Porosity	(4-8,2-3)	0.075	0.1016	0.128
2	Porosity	(2-8,4-7)	0.131	0.1308	0.1301
3	Porosity	(9-14,2-7)	0.095	0.086	0.08
4	X Permeability	(4-7,2-3)	2314	1922	3000.
5	X Permeability	(2-7,4-7)	4516	3242	3000.
6	X Permeability	(8,2-3)	440	669	792.7
7	X Permeability	(8,4-7)	13.4	13.7	14.3
8	X Permeability	(9-14,2-7)	27960	11220	3000
9	Y Permeability	(4-8,2)	1489	1062	1120
10	Y Permeability	(4-8,3)	0.0067	0.0094	0.0425
11	Y Permeability	(2-8,4-7)	609	407	367
12	Y Permeability	(9-14,2-7)	731	435	243.6

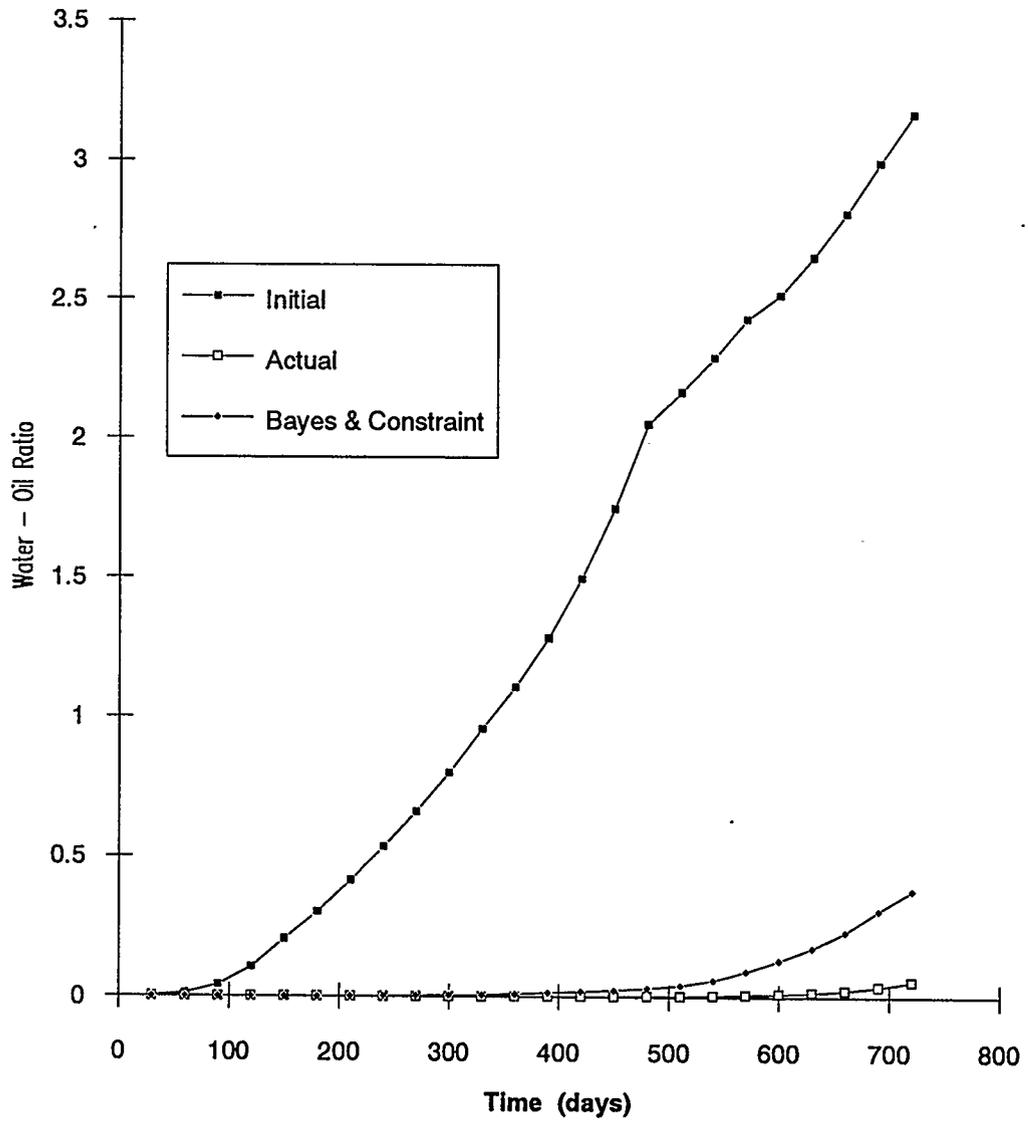
oil in-place estimate is 4,378,679 bbls. The maximum permeability value is now 11220 md, and this relatively high value shows that the incorporation of prior information using Bayesian estimation does not prevent extremal values when the LS objective function is significantly larger than the value of the prior function.

In order to restrict the parameters to realistic values, the feasible region must be constrained. The next run sets the constraint values for the permeability to be ≥ 0.001 and ≤ 3000 md in addition to the supplied prior estimates. The results are shown in column 6 of Table 9.5. The initial oil in-place estimate is 4,394,679 bbls. The match of the observation values is shown in Figures 9.12 to 9.20. A visually acceptable match is obtained for this second postulated model. Several parameter zones are restricted to the maximum permeability value. This indicates the postulated model is different from the actual model but there is no theoretical analysis available that will assist us in rezoning the model. Several feasible models must be created and compared to select the most realistic model. It will be impossible to create a model that represents exactly the underground reservoir. Nevertheless, it is encouraging that on a macroscopic basis, the postulated models provide us with reasonable estimates of oil in-place, and will suggest the presence of impermeable barriers as shown in this case study.

**Fig 9.12 - Case Study IV - Model B
Well #1 - Match of Pressure**



**Fig 9.13 - Case Study IV - Model B
Well #2 - Match of Water-Oil Ratio**



**Fig 9.14 - Case Study IV - Model B
Well #2 - Match of Gas-Oil Ratio**

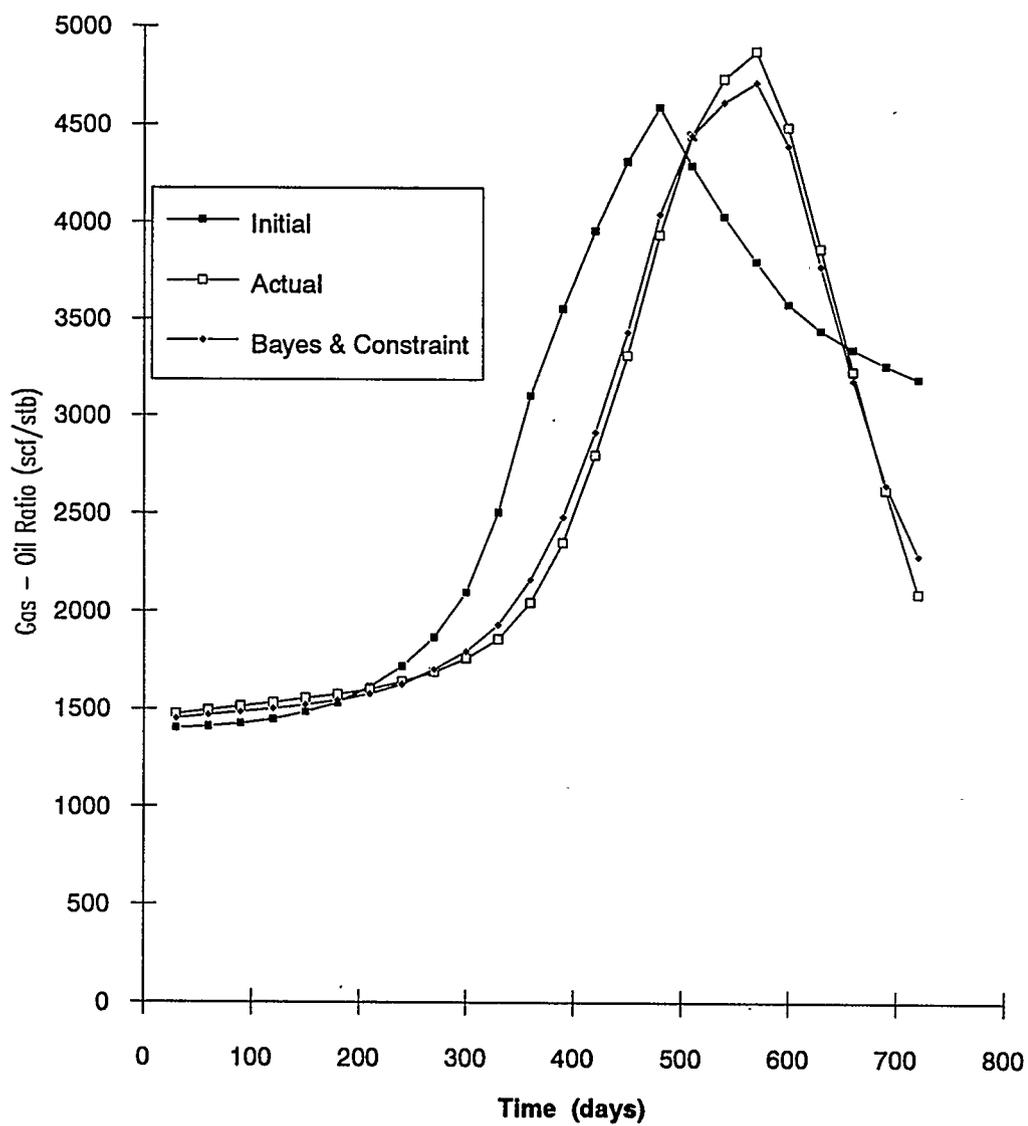
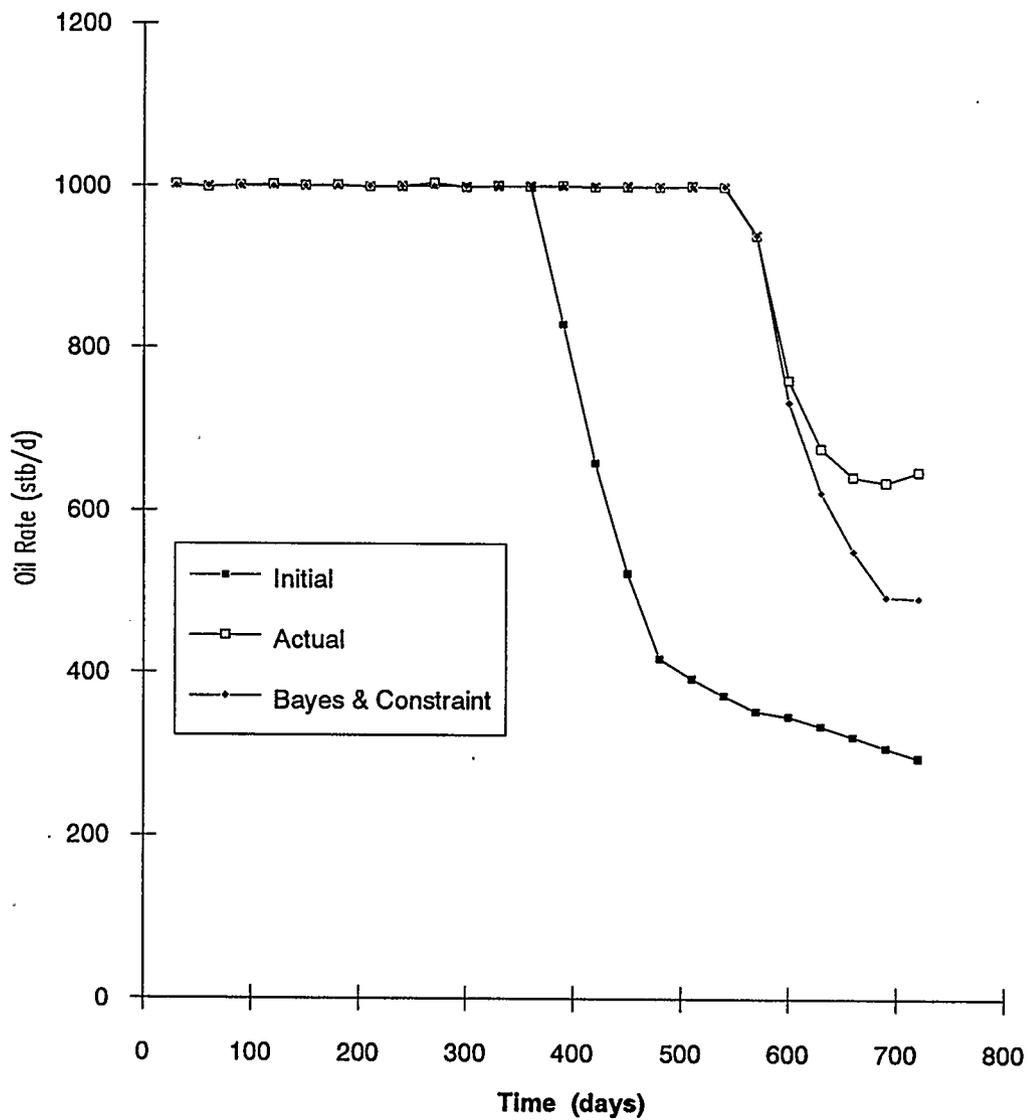
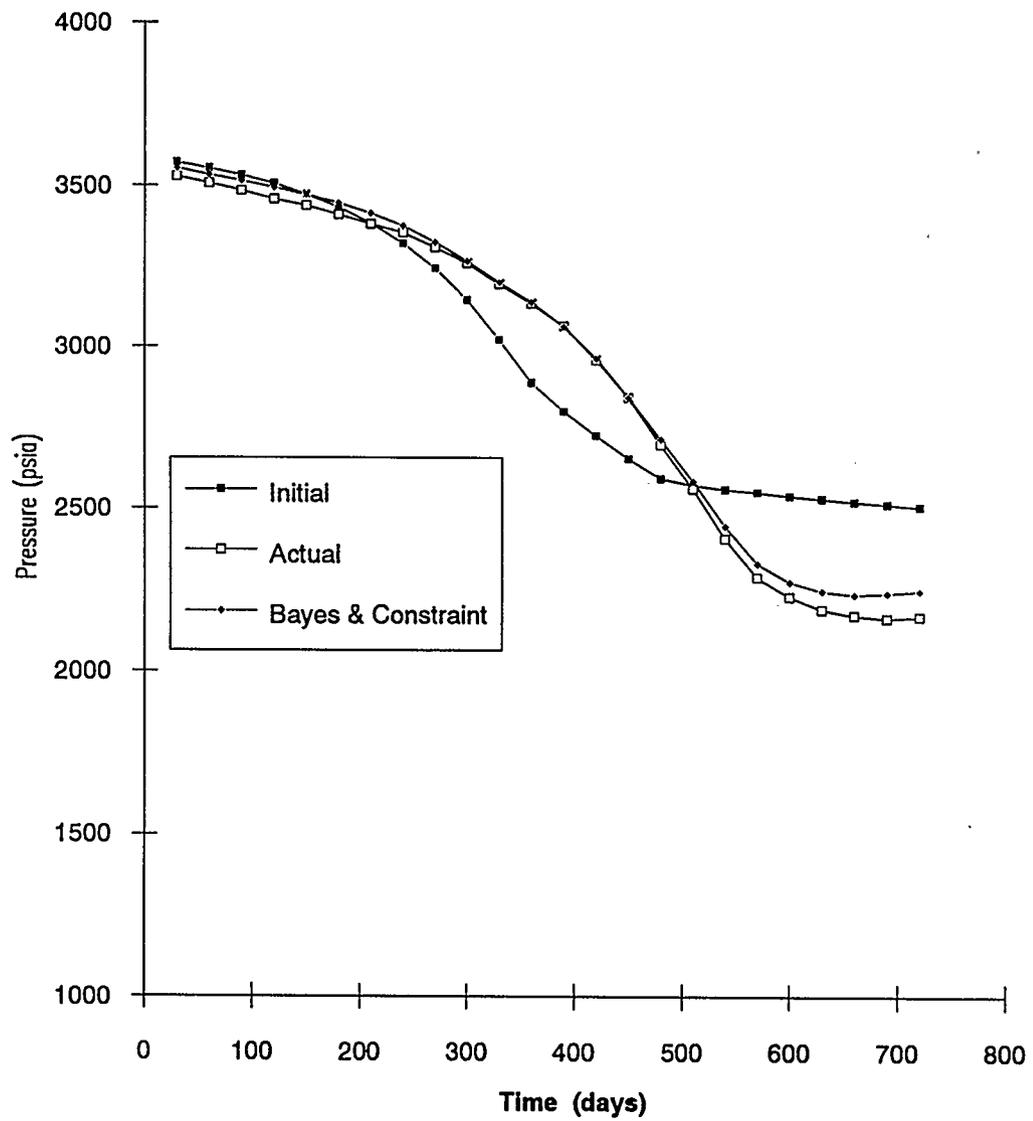


Fig 9.15 - Case Study IV - Model B
Well #2 - Match of Oil Rate



**Fig 9.16 - Case Study IV - Model B
Well #2 - Match of Pressure**



**Fig 9.17 - Case Study IV - Model B
Well #3 - Match of Water-Oil Ratio**

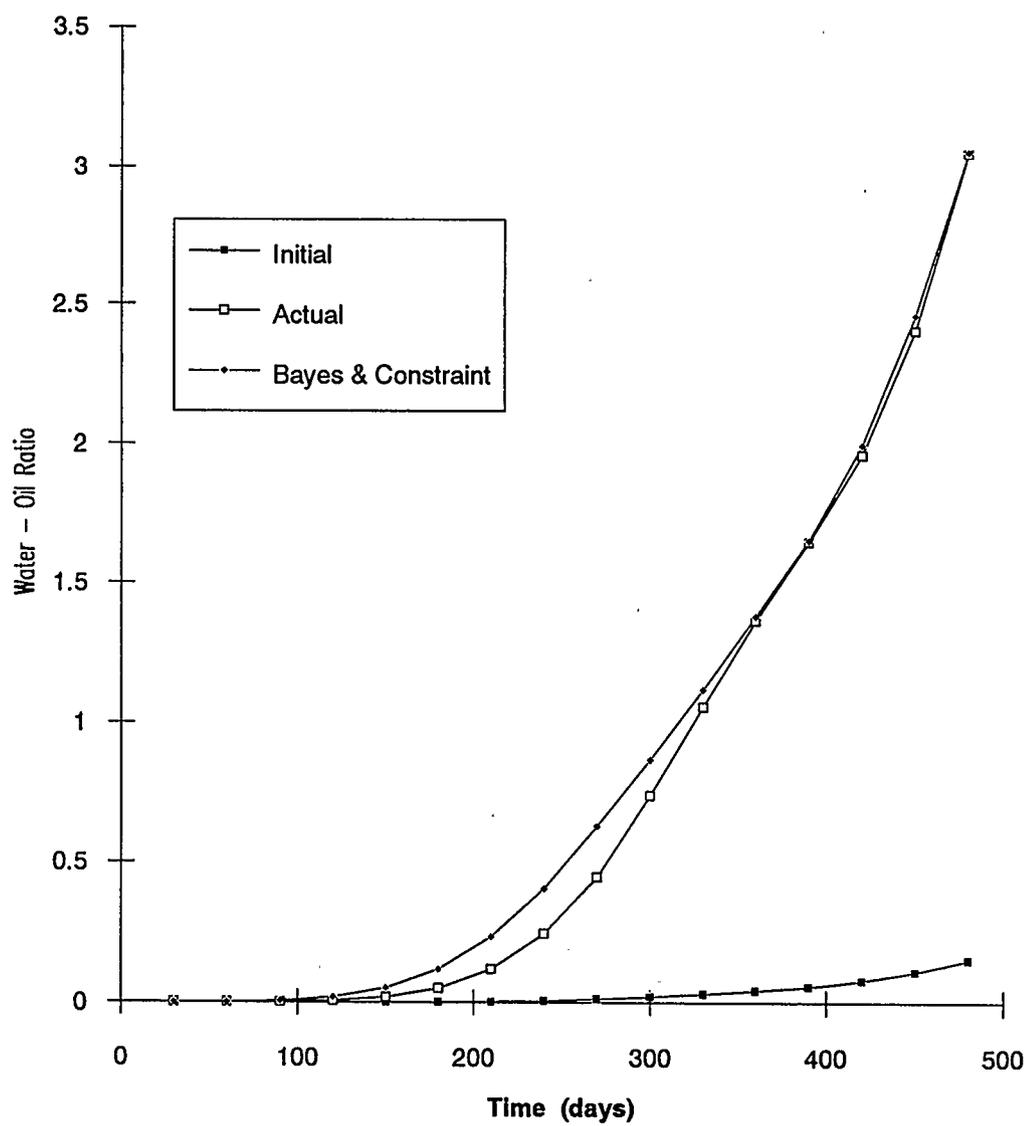


Fig 9.18 - Case Study IV - Model B
Well #3 - Match of Gas-Oil Ratio

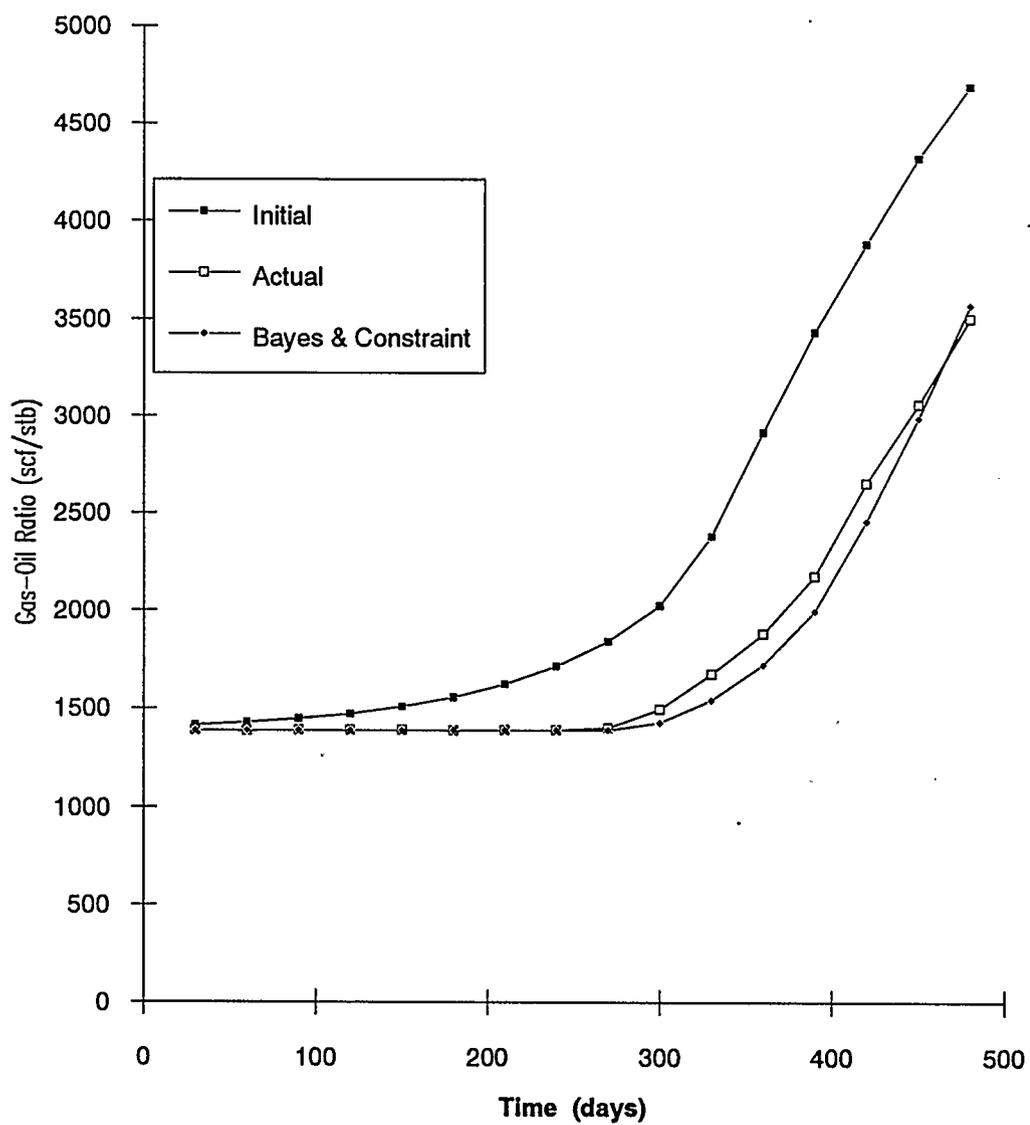
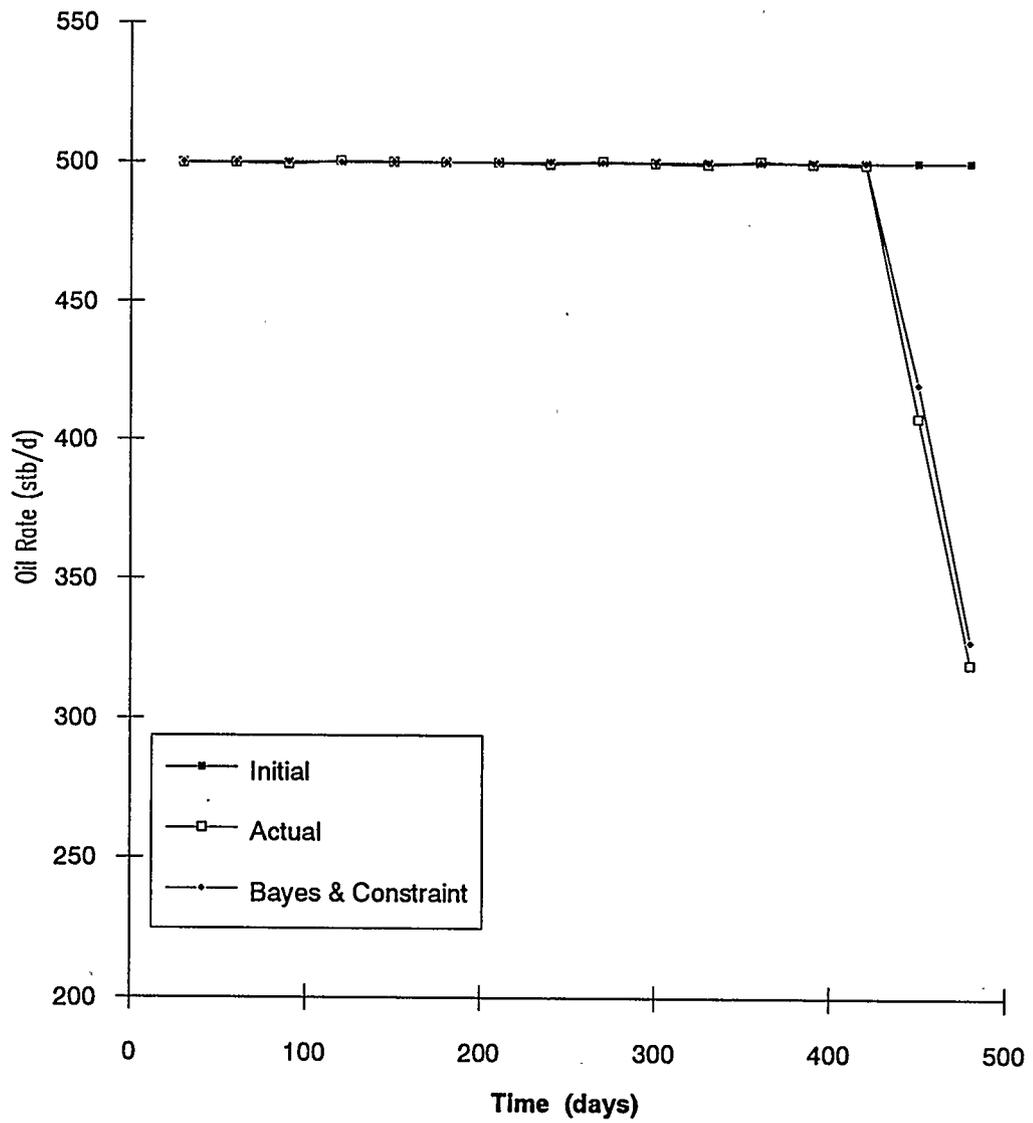
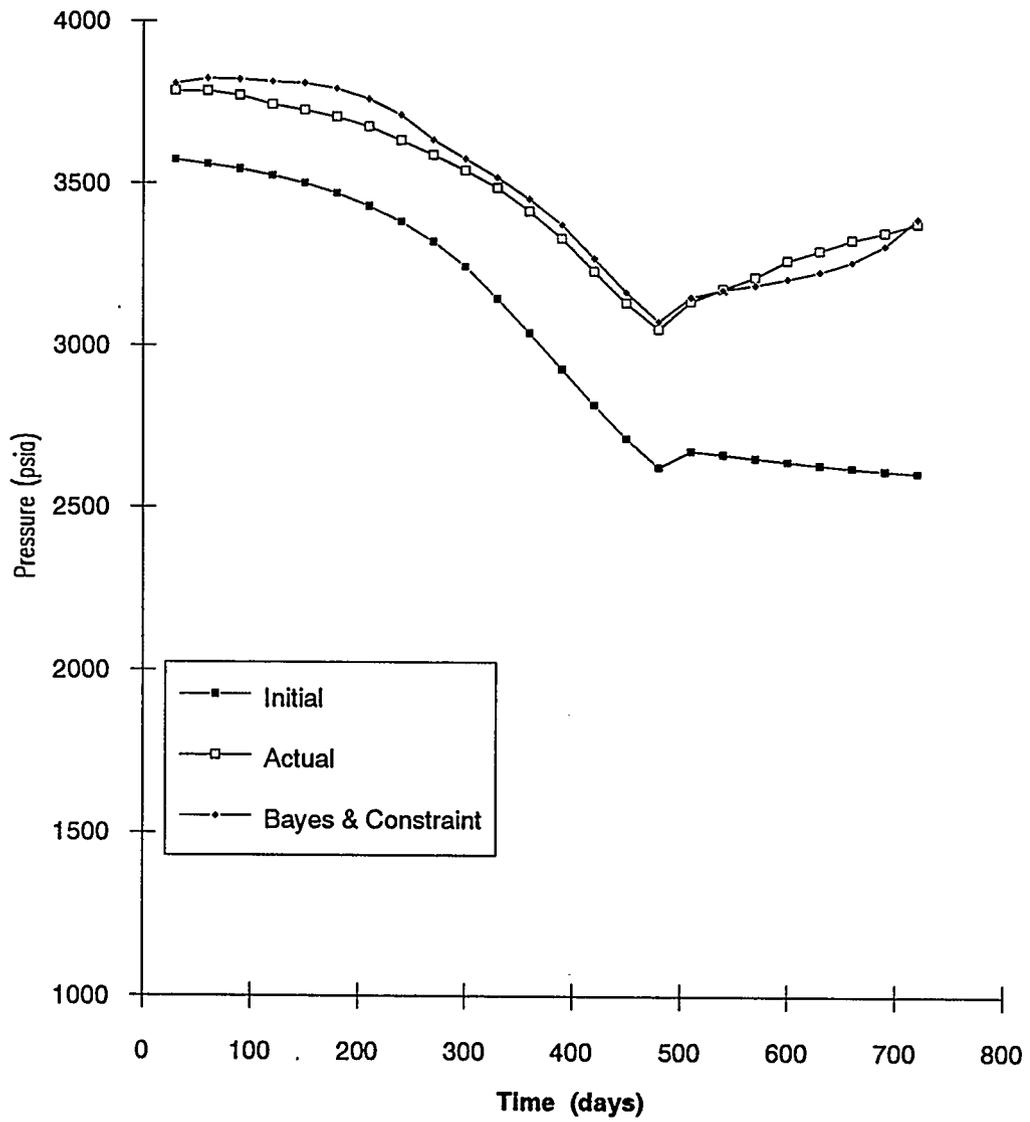


Fig 9.19 - Case Study IV - Model B
Well #3 Match of Oil Rate



**Fig 9.20 - Case Study IV - Model B
Well #3 - Match of Pressure**



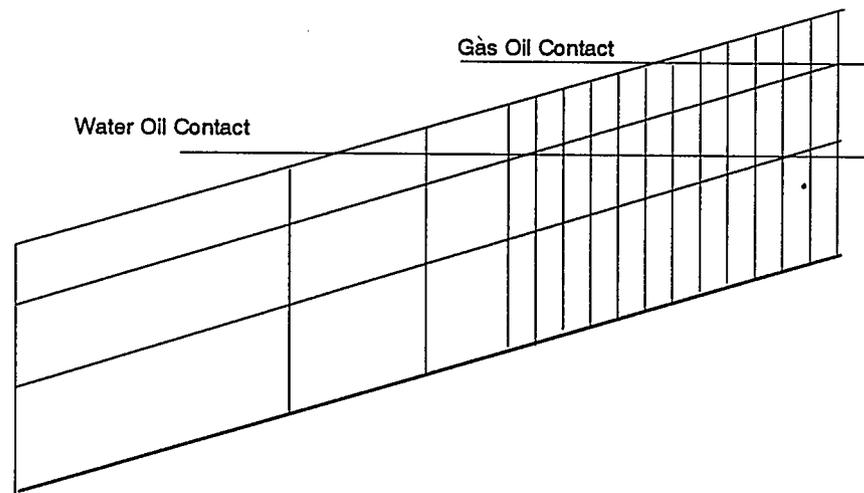
9.5 Case Study V

The purpose of this case study is to further investigate the reliability of the estimates of in-place fluids especially in situations where an aquifer and a gas cap are present. In development drilling of reservoirs, the target of the wells is usually the productive pay, normally the oil zone. Wells are seldom drilled into an aquifer or a gas cap to verify its existence and measure its extent. As a result, the size and extent of the aquifer or gas cap must be estimated by material balance techniques. Here we will use the automatic history matching model to predict the size of the aquifer or gas cap, and at the same time incorporate prior values using Bayesian estimation while constraining the parameters to within realistic values. We will also use different lengths of history to see the effect on the estimates of in-place fluids.

9.5.1 Vertical Cross-section Model

The model is based on a tilted monoclinial reservoir limited by an unconformity and connected to an aquifer as shown schematically in Figure 9.21. It is represented in the base model by an XZ cross-section consisting of 15 cells in the horizontal (X) direction, and 3 layers in the vertical (Z) direction. There is a small gas cap, an oil zone, and an extensive aquifer. These are explicitly modelled by the grid cell model. The fluid properties and relative permeability

Fig 9.21 Case Study V - Schematic of Vertical Cross-Section



tables are the same as in Case Study 9.1. The porosity and permeability distributions are detailed in Table 9.6. Each layer has a constant but different porosity and permeability in the X and Z directions.

The reservoir is produced from three wells in the oil zone, one near the aquifer, one in the middle and one near the gas cap. Well specifications are included in Table 9.6. The base run is carried out to 1800 days. As the oil zone is depleted, the gas cap expands resulting in high gas-oil ratios for the producing well near the gas cap. At 990 days, this well is shut in with a gas-oil ratio of 26,700 scf/stb. The aquifer provides pressure support and at the same time encroaches into the oil zone. The water-oil ratios of the two remaining wells increase with time, with the well nearer the water-oil contact showing the largest increase. At the end of the run, its water-oil ratio was 8.2 bbl/bbl. The producing gas-oil ratios, water-oil ratios, oil rates of the wells and pressures of the well cells are recorded every ninety days and used as observation data to be matched.

We will now assume in the postulated model that we do not know the true extent of the aquifer or the gas cap. The two columns of cells at either end of the cross-section are removed. The aquifer is reduced to less than half its original size, and the gas cap is reduced to one-third. The zonation used for the parameters is also different from the original distribution of porosity and permeability. For the

Table 9.6 : Description of Models Used for Case Study V

Vertical Cross-Section

Grid dimensions in X direction -
 20000, 5000, 1000, 300, 300, 300, 300, 300, 300, 300, 300,
 300,
 300, 300, 300, 300 ft
 Grid dimensions in Y direction - 3000 ft
 Thickness of layers 1 to 3 - 50, 75, 100 ft
 Top depth of first layer -
 9600, 9525, 9450, 9375, 9300, 9225, 9150, 9075, 9000,
 8925, 8850, 8775, 8700, 8625, 8550 ft
 Porosity of layers 1 to 3 - 0.4, 0.08, 0.10
 X direction Permeability layers 1 to 3
 - 300, 100, 300 md
 Z direction Permeability of layers 1 to 3
 - 40, 20, 0 md
 Initial Pressure at Depth 8700 ft is 3600 psia
 Depth of Gas-Oil, Water-Oil Contacts 8700, 9300 ft
 Well #1 - Producer Location (12,1,2) (12,1,3)
 Well Bore Index = 20
 Maximum Oil Rate = 1000 bbls/d
 Minimum Bottom Hole Pressure = 1000 psia
 Well #2 - Producer Location (10,1,2) (10,1,3)
 Well Bore Index = 30
 Maximum Oil Rate = 2000 bbls/d
 Minimum Bottom Hole Pressure = 1000 psia
 Well #3 - Producer Location (8,1,1) (8,1,2)
 Well Bore Index = 30
 Maximum Oil Rate = 2000 bbls/d
 Minimum Bottom Hole Pressure = 1000 psia

Three Dimensional Model

Grid dimensions in Y direction - 1000 ft
 Well #1 - Producer Location (12,2,2) (12,2,3)
 Well #2 - Producer Location (10,5,2) (10,5,3)
 Well #3 - Producer Location (8,1,1) (8,1,2)

porosity parameter, five zones consisting of columns of cells were used. These are detailed in Table 9.7. Only one zone, consisting of the entire grid model, was used for the permeability in the x direction. Similarly, only one zone was used for the permeability in the z direction. The initial guess of porosity was 0.15 for all the porosity parameters, 500 md for the x direction permeability, and 50 md for the z direction permeability. The prior estimates for porosity and permeability were the same as the initial guesses. The standard deviations for porosity and permeability were 0.01 and 50 md respectively. Constraints for the porosity were ≥ 0.001 and ≤ 0.5 and for permeability were ≥ 0.001 md and ≤ 3000 md.

Runs were made using varying lengths of history. Specifically these were 360, 720, 1080, 1440 and 1800 days. The estimated parameters for each of these runs are shown in Table 9.7. The calculated in-place values of oil, gas and water are shown in Table 9.8 and compared with the actual values. It can be seen that the calculated values are within 17% of the actual values using a data length of 360 days, and the error in general decreases as the data length increases. However it can be noted that the estimates at 720 days are as good as the estimates at 1800 days. It appears that even using a postulated model that is different from the actual model, the estimates of aquifer size and gas cap size are quite satisfactory. Early time estimates are also quite

Table 9.7 : Case Study V - Vertical Cross-Section Model - Parameter Estimates						
History Period						
Zone	Parameter	0 - 360 days	0 - 720 days	0-1080 days	0-1440 days	0-1800 days
1	Porosity (3,1,1-3)	0.2198	0.2354	0.2394	0.2394	0.2543
2	Porosity (4-6,1,1-3)	0.0658	0.3088	0.0645	0.1671	0.1630
3	Porosity (7-9,1,1-3)	0.1152	0.1079	0.0964	0.0994	0.1025
4	Porosity (10-12,1,1-3)	0.1454	0.0652	0.0829	0.1055	0.1053
5	Porosity (13,1,1-3)	0.2076	0.3783	0.4064	0.3253	0.3146
6	X Permeability (3-13,1,1-3)	294.5	352.1	276.6	252.3	266.7
7	Z Permeability (3-13,1,1-3)	36.99	15.59	14.31	28.95	24.03

Table 9.8: Case Study V - Vertical Cross-Section Model - Fluids In-place Estimates						
	oil (10 ⁷ stb)	% diff	gas (10 ⁷ mcf)	% diff	water (10 ⁹ bbls)	% diff
actual	1.968		3.152		1.902	
0-360 days	2.195	11.54%	3.268	3.68%	1.576	-17.1%
0-720 days	1.961	-0.3%	3.119	-1.04%	1.819	-4.36%
0-1080 days	2.004	1.83%	3.209	1.82%	1.704	-10.4%
0-1440 days	2.067	5.04%	3.212	1.93%	1.764	-7.25%
0-1800 days	2.061	4.71%	3.192	1.29%	1.862	-2.09%

reasonable. This implies that a history match of a reservoir early in its producing life can provide estimates of its in-place volumes, and as more data becomes available, this data will help to refine estimates of porosity and permeability distributions.

9.5.2 Three-dimensional Model

The previous vertical cross-section model was extended to a three-dimensional model for this test case. The resulting model is a 15 x 5 x 3 model with 5 cells in the Y direction. The width of the reservoir is extended to 5000 feet from 3000 feet. The initial in-place volumes of oil, gas and water are larger than in the cross-sectional model. The reservoir description, fluid properties are identical to the cross-sectional case. The Y direction permeability is made the same as the X direction permeability. The wells were also redistributed along the width (Y direction) of the reservoir. In all other respects, the reservoir model is the same as the cross-sectional model. The base run is also carried out to 1800 days recording observation data every 90 days.

The postulated model is again created by deleting two planes of cells at each end of the model, making the model an 11 x 5 x 3 model effectively reducing the size of the postulated aquifer and gas cap. One more parameter zone was added for the permeability in the Y direction. The parameters in this model thus represent all possible parameters that

currently can be used in the automatic history matching simulator, i.e porosity and the permeability in each of the co-ordinate directions. Prior estimates for porosity and permeability, constraints as detailed in the cross-section case were also added.

Runs were again made with varying lengths of historical observation data. The estimated values for the parameters are shown in Table 9.9. Some porosity estimates are at the constraint limits, particularly that of the second porosity parameter. There is no well in this zone, so most likely there is insufficient information to characterize this zone. The utility of the penalty function approach in limiting parameter values to within the constraint limits is again demonstrated.

The calculated fluids in-place values for each of the runs are shown in Table 9.10, and compared with the original values. The conclusions that were drawn for the cross-section case are substantiated by the three dimensional case. Reasonable estimates of aquifer size, oil in-place and gas cap size can be obtained even with fairly short periods of historical data, and in general, the accuracy of the estimates increases with the length of the historical period. Oil and gas in-place estimates are within 10% of the actual values and this is more than acceptable in field practice.

Table 9.9 : Case Study V - Three Dimensional Model - Parameter Estimates						
History Period						
Zone	Parameter	0 - 360 days	0 - 720 days	0-1080 days	0-1440 days	0-1800 days
1	Porosity (3,1-5,1-3)	0.2605	0.2607	0.1677	0.1938	0.2420
2	Porosity (4-6,1-5,1-3)	0.1462	0.1568	0.5000	0.5000	0.0010
3	Porosity (7-9,1-5,1-3)	0.1264	0.1241	0.0560	0.1061	0.0982
4	Porosity (10-12,1-5,1-3)	0.1318	0.0967	0.0458	0.0772	0.1229
5	Porosity (13,1-5,1-3)	0.1920	0.2814	0.5000	0.2963	0.2580
6	X Permeability (3-13,1-5,1-3)	285.4	297.1	192.4	317.1	337.4
7	Y Permeability (3-13,1-5,1-3)	121.1	111.6	144.6	452.9	218.3
8	Z Permeability (3-13,1-5,1-3)	16.44	16.67	4.451	3.367	6.163

Table 9.10: Case Study V - Three Dimensional Model - Fluids In-place Estimates						
	oil (10 ⁷ stb)	% diff	gas (10 ⁷ mcf)	% diff	water (10 ⁹ bbls)	% diff
actual	3.281		5.253		3.169	
0-360 days	3.587	9.32%	5.318	1.25%	3.164	-0.19%
0-720 days	3.419	4.21%	5.240	-0.24%	3.174	-0.13%
0-1080 days	3.034	-7.53%	5.084	-3.2%	2.441	-22.9%
0-1440 days	3.239	-1.26%	5.016	-4.5%	2.747	-13.3%
0-1800 days	3.342	1.86%	5.092	-3.05%	2.805	-11.5%

9.6 Practical Application of the Automatic History Matching Model

The steps involved in a reservoir study using the model can be summarised as follows:

(a) Using all available information, the engineer will construct a postulated grid cell model of the reservoir.

(b) The model will then be subdivided into zones following as closely as possible any geological zonation. The zones for porosity and permeability need not be identical. The cells allocated to a zone do not need to be contiguous.

(c) The engineer will then optionally provide estimates of the most probable values of the parameters as well as constraints on the minimum and maximum values.

(d) The model is then run to provide estimates of the parameter values.

(e) At convergence or at the minimum of the LS objective function, the variances of the parameters as well as the eigenvectors of the smallest eigenvalues should be inspected to identify any highly correlated zones. Adjacent zones of high variance should be combined.

(f) Any zones with values close to the constraint limits should be carefully analyzed with a view to modifying the postulated grid cell representation of the reservoir.

(g) If changes are made to the model repeat from step (c), otherwise carry on with the study. It should be noted that the nature of the problem is such that it is impossible

to obtain a postulated model that uniquely represents the reservoir and the engineer should update the match when additional information becomes available.

CHAPTER 10

CONCLUSIONS AND RECOMMENDATIONS

10.1 Conclusions

In this work, several modifications were proposed to reduce the computational effort of parameter estimation using the Gauss-Newton method. Numerical experiments were carried out with chemical kinetic problems involving ordinary differential equations to show the convergence of the method was not affected. Then the method was implemented in a single-phase compressible flow simulator. Comparisons with published test cases showed the method was significantly faster.

A fully implicit three-dimensional three-phase multi-component simulator with an automatic history matching capability was developed. The model can simultaneously match observed pressures, water-oil ratios, gas-oil ratios and flowing bottom hole pressures. A zonation approach is used. Parameters that currently can be estimated are the permeability in all directions as well as porosity. The model was used to match observation data generated using the Second SPE Comparative Solution Problem.

The model is extremely efficient in computational requirements. Based on the run times obtained from the test cases, the program required only $1.233 + 0.077(\text{number of$

parameters) equivalent model runs for each Gauss-Newton iteration. This is significantly less than the (number of parameters + 1) equivalent model runs required by previously published Gauss-Newton methods.

Further, by using several examples, it was demonstrated that one can obtain significant information about the reliability of the estimates of the parameters when an automatic history matching model based on the Gauss-Newton method is used. This is an important advantage over manual matching. It was shown that the correlation matrix developed at convergence will suggest which pairs of zones are highly correlated. For correlation of multiple zones, the eigenvectors corresponding to the smallest eigenvalues should be inspected to locate the parameters that correspond to the largest values. Zones which are highly correlated usually have the largest variances. The variances of the parameters also depend on the measurement errors. In general, the match can often be improved by removing or combining adjacent zones whose parameters have high variance.

In actual application, the postulated model will be quite different than the actual reservoir. It was demonstrated that under these circumstances, the regression analysis could result in parameter values that are highly unrealistic. However this problem can be alleviated by introducing prior information using Bayesian estimation to influence the parameter search towards more realistic values. Inviolable

limits on the feasible parameter values can be imposed using constraints via the penalty function method.

Test cases were presented using postulated models that were quite different than the actual models that were used to generate the artificial observation data. It was concluded from the test cases that the automatic history matching model can be used to detect the presence of impermeable barriers, estimate in-place volumes of oil, gas and water and suggest reservoir extensions that would make the postulated model conform more closely to the actual reservoir.

With the combination of an effective parameter estimation method - the Gauss-Newton method, utilization of prior information using Bayesian estimation, and limiting constraints on parameter values, the three-dimensional three-phase simulator with automatic history matching capability is a powerful tool that can be used effectively for parameter estimation in reservoir modelling.

10.2 Recommendations

10.2.1 Increasing the Parameter Types

The parameters that can be estimated should be extended to include fluid properties as well as saturation functions. These may be represented by spline functions and then the parameters of the functions will be estimated. The zonation approach assumes a constant parameter value over a range of

cells. This can be extended such that the parameter value to be estimated is a multiplying factor of the values in the range of cells. This will allow a distribution to be raised or lowered in the range of cells. Another approach would be to use a spline approximation for the parameter distribution and then to estimate the coefficients of the spline.

10.2.2 Improving the Cautious Step-size Policy

In the current implementation, each iteration consists of the solution of the model equations and sensitivity equations to generate the Gauss-Newton matrix. The step size and direction predicted by the Gauss-Newton method is accepted regardless whether the objective function is larger at the new estimates. It is desirable that if this occurs, the new estimate not be accepted, but the Gauss-Newton matrix generated at the new estimate be used in combination with the previous estimate's Gauss-Newton matrix to provide a new step size and direction that hopefully results in a lower objective function value.

10.2.3 Improving the Postulated Model and Discrimination of Postulated Models

Methods should be devised to redesign the postulated model if it results in a poor fit of the observation data. Statistical tests to distinguish between competing models should be investigated.

REFERENCES

Acs, G., Doleschall, S., and Farkas, E.: "General Compositional Model," *SPEJ* (Aug. 1985) 543-53.

Appleyard, J.R. and Chesire, L.M.: "Nested Factorization," paper SPE 12264 presented at the 1983 SPE Symposium on Reservoir Simulation, San Francisco, Nov 16-18.

Aziz, K. and Settari, A.: *Petroleum Reservoir Simulation*, Applied Science Publishers Ltd., London, (1979).

Bard, Y.: "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems," *SIAM J. Numer. Anal.* (1970), **7**, 157.

Bard, Y. : *Nonlinear Parameter Estimation*, Academic Press Inc., Florida, 1974.

Bayes, T.: "Essay towards solving a problem in the doctrine of chances," *Philos. Trans. Roy. Soc.* **53**, 370-418. [Reprinted in *Biometrika*, **45**, 293-315 (1958)]

Bondor, P.L., Hirasaki, G.J. and Tham, M.J.: "Mathematical Simulation of

Polymer Flooding in Complex Reservoirs," *SPEJ* (Oct 1972) 369-82.

Burden, R.L. and Faires, J.D.: *Numerical Analysis*, Prindle, Weber and Schmidt, Boston, (1981)

Caracotsios, M. and Stewart, W.E.: "Sensitivity Analysis of Initial Value Problems with Mixed ODES and Algebraic Equations", *Comp. & Chem. Eng* (1985) Vol 9, 359-365.

Chappelear, J.E. and Nolen, J.S.: "Second Comparative Project: A Three Phase Coning Study", *JPT* (Mar. 1986) 345-353.

Chavent, G., Dupuy, M. and Lemonnier, P.: "History Matching by Use of Optimal Control Theory," *SPEJ* (Feb. 1975) 74-86.

Chen, W.H., Gavalas, G.R., Seinfeld, J.H. and Wasserman, M.L.: "A New Algorithm for Automatic History Matching," *SPEJ* (Dec. 1974) 593-608.

Chen, Y.M. et al.: "GPST Inversion Algorithm for History Matching in 2-D Two-Phase Simulator Models," *Applied Numer. Math.*, 4, 83 (1988).

Christie, M.A.: "Numerical Techniques for High-Resolution Simulation of

Instabilities", paper SPE 16005 presented at the Ninth SPE Symposium on Reservoir Simulation, San Antonio, Texas, Feb 1987.

Chung, C.B. and Kravaris, C.: "Incorporation of A Priori Information in Reservoir History Matching by Regularization," unsolicited paper SPE 21615 received in 1990.

Coats, K.H., Dempsey, J.R. and Henderson, J.H.: " A New Technique for Determining Reservoir Description from Field Performance Data," *SPEJ*. (March 1970) 66-74.

Coats, K.H.: " An Equation of State Compositional Model," *SPEJ* (Oct 1980) 363-76.

Coats, K.H.:"Three-Dimensional Simulation of Steam Flooding," *SPEJ* (Dec 1974) 574-92; Trans., AIME, **257**.

Coats, K.H.:"Simulation of Steam Flooding with Distillation and Solution Gas," *SPEJ* (Oct 1976) 235-47.

Coats, K.H.:"A Highly Implicit Steam Flood Model," *SPEJ* (Oct 1978) 369-83.

Coats, K.H.: "In-Situ Combustion Model," *SPEJ* (Dec 1980) 533-54; *Trans., AIME*. **269**.

Cooley, R.L.: "Incorporation of prior information on parameters into nonlinear regression groundwater flow models: 1. Theory," *Water Resour. Res.* **18** 965-76, 1982.

Crichlow, H.B.: *Modern Reservoir Engineering - A Simulation Approach*, Prentice-Hall Inc., New Jersey, (1977).

Crookston, R.B., Culham, W.E. and Chen, W.H.: " A Numerical Simulation Model for Thermal Recovery Processes," *SPEJ* (Feb 1979) 37-58; *Trans., AIME*, **267**.

Dogru, A.H., Dixon, T.N., and Edgar, T.F.: " Confidence Limits on the Parameters and Predictions of Slightly Compressible Single-Phase Reservoirs," *SPEJ* (Feb 1977) 42-56.

Dogru, A.H. and Seinfeld, J.H.: " Comparison of Sensitivity Coefficient Calculation Methods in Automatic History Matching," *SPEJ* (Oct 1981) 551-557.

Dumez, F.J., Hosten, L.H., and Froment, G.F.: " The Use of Sequential Discrimination in the Kinetic Study of 1-Butene Dehydrogenation," *Ind. Eng. Chem. Fundam.*, **16**, (1977) 298-301.

Enright, A.H. and Hull, T.H.: In "*Numerical Methods for Differential Equations*;" Lapidus, L. and Schiesser, W.E., Eds.; Academic Press: New York, 1976, p.45

Farouq Ali, S.M. and Berkowitz, B.S.: "Limiting Aspects of Automatic History Matching," paper 88-39-03 presented at the CIM 39th Annual Technical Meeting, Calgary, June 1988.

Garfinkel, D., Ching, D.W., Adelman, M. and Clark, P.: "Techniques and Problems in Construction of Computer Models of Biochemical Systems Including Real Enzymes," *Ann. N. Y. Acad. Sci.* (1966), 1054

Gavalas, G.R., Shah, P.C., and Seinfeld, J.H.: "Reservoir History Matching by Bayesian Estimation," *SPEJ* (Dec 1976) 337-50.

George, J.A.: "Nested Dissection of a Regular Finite Element Mesh," *SIAM J. Numerical Analysis* (April 1973) **10** 345-63.

Gilman, J.R., and Kazemi, H.: "Improved Calculations for Viscous and Gravity Displacement in Matrix Blocks in Dual Porosity Simulators," paper SPE 16010 presented at the 1987 SPE Symposium on Reservoir Simulation, San Antonio, Feb 1-4.

Heinemann, Z.E., Gerken, G. and von Hantelmann, G.: "Using Local Grid Refinement in a Multiple Application Reservoir Simulator," paper SPE 12255 presented at the 1983 SPE Symposium on Reservoir Simulation, San Francisco, Nov 16-18.

Jacquard, P. and Jain, C.: "Permeability Distribution from Field Pressure Data," *SPEJ* (Dec. 1965) 281-294.

Jahns, H.O.: "A Rapid Method for obtaining a Two Dimensional Reservoir Description from Well Response Data," *SPEJ* (Dec. 1966) 315-327.

Kalogerakis, N. and Luus, R.: "Improvement of Gauss-Newton Method for Parameter Estimation through the Use of Information Index", *I & EC Fundamentals* (Nov. 1983) Vol 22, 436-445.

Kazemi, H., Vestal, C.R., and Shank, G.D.: "An Efficient Multicomponent Numerical Simulator," *SPEJ* (Oct 1978) 355-68.

Lapidus, L. and Luus, R.: *Optimal Control of Engineering Processes*, Blaisdell, MA (1967)

Lawson, C.L. and Hanson, R.J.: *Solving Least Squares Problems*, Prentice Hall Inc., Englewood Cliffs (1974)

Lee, B.Y. and Tan, T.B.: "Application of a Multiple Porosity/Permeability Simulator in Fractured Reservoir Simulation," paper 16009 presented at the 1987 SPE Symposium on Reservoir Simulation, San Antonio, Feb 1-4.

Lee, T., Kravaris, C. and Seinfeld, J.H.: " History Matching by Spline Approximation and Regularization in Single-Phase Areal Reservoirs," *SPEE*, **1**, 521, 1986.

Letkeman, J.P. and Ridings, R.L.: "A Numerical Coning Model," *SPEJ* (Dec 1970) 418-24, Trans., AIME **249**.

Macdonald, R.C. and Coats, K.H.: "Methods for Numerical Simulation of Water and Gas Coning," *SPEJ* (Dec 1970) 425-35, Trans., AIME, **249**.

Makhlouf, E.M. et al.: "A General History Matching Algorithm for Three-Phase Three-Dimensional Petroleum Reservoirs," unsolicited paper SPE 20383

received in 1990.

Mattax, C.C. and Dalton, R.L. (Eds): *Reservoir Simulation*, Monograph Series, SPE. Richardson, TX (1990) **13**.

Neuman, S.P. and Yakowitz, S.: "A Statistical Approach to the Inverse Problem of aquifer hydrology: 1. Theory," *Water Resources Res.*, **18** 965-76, 1982.

Nolen, J.S.: "Numerical Simulation of Compositional Phenomena in Petroleum Reservoirs," paper SPE 4274 presented at the 1973 SPE Symposium on Numerical simulation of Reservoir Performance, Houston Jan. 11-12.

Padmanabhan, L. and Woo, P.T.: "A New Approach to Parameter Estimation in Well Testing," paper SPE 5741 presented at the 1976 SPE Symposium on Numerical simulation of Reservoir Performance, Los Angeles, Feb 19-20.

Peaceman, D.W. *Fundamentals of Numerical Reservoir Simulation*, Elsevier Scientific Publishing Company, Amsterdam, (1977).

Peaceman, D.W.: "Interpretation of Well-Block Pressures in Numerical Reservoir Simulation," *SPEJ* (Jun 1978) 183-94, *Trans., AIME*, **253**.

Peaceman, D.W.: "Interpretation of Well-Block Pressures in Numerical Reservoir Simulation with Nonsquare Grid Blocks and Anisotropic Permeability," *SPEJ* (June 1983) 531-43.

Pedrosa, O.A. and Aziz, K.: "Use of Hybrid Grid in Reservoir Simulation," paper SPE 13507 presented at the 1985 SPE Symposium on Reservoir Simulation, Dallas, Feb 10-13.

Peng, D.Y. and Robinson, D.B.: "A New Two-Constant Equation of State," *Ind. Eng. Chem. Fund.* (1976) **15**, 59-64.

Petzold, L.R.: "A description of DASSL: a differential /algebraic system solver", Sandia Tech. Rep. 82-8637 (1982).

Price, H.S. and Coats, K.H.: "Direct Methods in Reservoir Simulation," *SPEJ* (June 1974) 295-308, *Trans., AIME*, **257**.

Redlich, O. and Kwong, J.N.S.: "On the Thermodynamics of Solutions. V. An Equation of State. Fugacities of Gaseous Solutions," *Chem. Reviews* (Feb. 1949) **44**, 233-44.

Ridings, R.L.: "Handling Fluid Properties in Reservoir Simulation Studies,"

paper SPE 3703 presented at the 1971 SPE California Regional Meeting, Los Angeles, Nov 4-5.

Robertson, H.H.: In "*Numerical Methods*;; Walsh, J., Ed.; Thompson: Washington, 1967; p. 178.

Seinfeld, J.H. and Gavalas, G.R.: " Analysis of Kinetic Parameters from Batch and Integral Reactor Experiments," *AICHE J.* (1970) **16**, 644.

Shah, P.C., Gavalas, G.R. and Seinfeld, J.H.: " Error Analysis in History Matching: The Optimum Level of Parameterization," *SPEJ* (June 1978) 219-228.

Slater, G.E. and Durrer, E.J.: "Adjustment of Reservoir Simulation Models to Match Field Performance," *SPEJ* (Sep. 1971) 295-305.

Smith, G.D.: *Numerical Solution of Partial Differential Equations*, Oxford University Press, London, 1969.

Stone, H.L., and Garder, A.O.: "Analysis of Gas-Cap or Dissolved-Gas Drive Reservoirs," *SPEJ* (June 1961) 91-102.

Stone, H.L.:" Iterative Solution of Implicit Approximations of Multi-Dimensional Partial Differential Equations," *SIAM J. Numerical Analysis* (1968) 5, 530-58.

Stone, H.L.:" Estimation of Three-Phase Relative Permeability and Residual Oil Data," *J. Cdn. Pet. Tech.* (Oct-Dec 1973) 53-61.

Tan, T.B. and Letkeman, J.P.:" Application of D4 Ordering and Minimization in an Effective Partial Matrix Inverse Iterative Method," paper SPE 10493 presented at the 1982 SPE Symposium on Reservoir Simulation, New Orleans, Feb 1-3.

Tan, T.B.:"Implementation of an Improved Adaptive Implicit Method in a Thermal Compositional Simulator," paper SPE 16028 presented at the 1987 SPE Symposium on Reservoir Simulation, San Antonio, Feb 1-4.

Tan, T.B. and Kalogerakis, N.: "An Efficient Automatic History Matching Algorithm for Reservoir Modelling", to appear in *J. Comp. Phys.*

Tan, T.B. and Kalogerakis, N.: "A Fully Implicit Three-Dimensional Three-Phase Simulator with Automatic History Matching Capability," paper SPE 21205 presented at the Eleventh SPE Symposium on Reservoir Simulation,

Anaheim, Feb 17-20, 1991.

Tan, T.B. and Kalogerakis, N.: "A Three-Dimensional Three-Phase Automatic History Matching Model: Reliability of Parameter Estimates," paper CIM/AOSTRA 91-84 presented at the CIM/AOSTRA 1991 Technical Conference, Banff, April 21-24, 1991.

Thomas, G.W. and Thurnau, D.H.: "The Mathematical Basis of the Adaptive Implicit Method," paper SPE 10495 presented at the 1982 SPE Symposium on Reservoir Simulation, New Orleans, Feb 1-3.

Thomas, L.K., Hellums, L.J. and Reheis, G.M.: "A Nonlinear Automatic History Matching Technique for Reservoir Simulation Models," *SPEJ* (Dec. 1972) 508-514.

Todd, M.R. and Longstaff, W.J.: "The Development, Testing and Application of a Numerical Simulator for Predicting Miscible Flood Performance," *JPT* (July 1972) 874-82; *Trans., AIME*, **253**.

Todd, M.R. and Chase, C.A.: "A Numerical Simulator for Predicting Chemical Flood Performance," paper SPE 7689 presented at the 1979 SPE Symposium on Reservoir Simulation, Denver, Feb 1-2.

Veatch, R.W. and Thomas, G.W.: "A Direct Method for History Matching," paper SPE 3515 presented at 46th Annual SPE Fall Meeting, New Orleans, Oct. 3-6, 1971.

Vinsome, P.K.: "Orthomin, an Iterative Method for Solving Sparse Sets of Simultaneous Linear Equations," paper SPE 5729 presented at the 1976 SPE Symposium on Numerical simulation of Reservoir Performance, Los Angeles, Feb 19-20.

Wasserman, M.L., Emanuel, A.S. and Seinfeld, J.H.: "Practical Applications of Optimal Control Theory to History Matching Multiphase Simulator Models", *SPEJ* (Aug. 1975) 347-355.

Wasserman, M.L.: "Local Grid Refinement for Three-Dimensional Simulators," paper SPE 16013 presented at the 1987 SPE Symposium on Reservoir Simulation, San Antonio, Feb 1-4.

Watson, A.T., Seinfeld, J.H., Gavalas, G.R. and Woo, P.T.: "History Matching in Two-Phase Petroleum Reservoirs", *SPEJ* (Dec. 1980) 521-532.

Watson, A.T., Gavalas, G.R. and Seinfeld, J.H.: "Identifiability of Estimates of Two-Phase Reservoir Properties in History Matching," *SPEJ* (Dec 1984) 697-

705.

Watts, J.W.: "A Compositional Formulation of the Pressure and Saturation Equations," *SPEE* (March 1986), 243-52.

White, C.D. and Horne, R.N.: "Computing Absolute Transmissibility in the Presence of Fine-Scale Heterogeneity," paper 16011 presented at the Ninth SPE Symposium on Reservoir Simulation, San Antonio, Texas, Feb 1987.

Yang, P., Armasu, R.V. and Watson, A.T.: "Automatic History Matching with Variable-Metric Methods", paper SPE 16977 presented at the 62nd Annual Technical Conference, Dallas, Sep. 27-30, 1987.

Yang, P-H. and Watson, A.T.: "A Bayesian Methodology for Estimating Relative Permeability Curves," *SPEE*, **6**, 259-65, 1991.

Young, L.C. and Stephenson, R.E.: "A Generalized Compositional Approach for Reservoir Simulation," *SPEJ* (Oct. 1983) 727-42.

Youngren, G.K.: "Development and Application of an In-Situ Combustion Reservoir Simulator," *SPEJ* (Feb 1980) 39-51; Trans., AIME, **269**.

APPENDIX A

```

C      ODE      GAUSS NEWTON METHOD FOR LEAST SQUARES
C
C      EXAMPLE PROBLEM 1 OF DR KALOGERAKIS'S PAPER
C
PROGRAM GN
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1      FS(10,10),RS(10,10),BS(10,10),
1      XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2      L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3      F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
DIMENSION DK(10),TK(10),VK(10),XK(10)
CHARACTER*16 FILE
LOGICAL SCALE
Q(1) = 1.E10
Q(2) = 10.E10
WRITE(*,700)
700  FORMAT(' OUTPUT FILE ? ', '$)
READ(*,701) FILE
701  FORMAT(A)
OPEN(UNIT=6,FILE=FILE,STATUS='UNKNOWN',FORM='FORMATTED')
NEXPT = 9
NN = 2
NP = 2
VK(1) = 355.4 *10000
VK(2) = 403.3 *10000
TOL = 0.0001
CALL FUNCT(VK,VALK)
WRITE(6,*) ' INITIAL P.I,K1,K2 = ',VALK,(VK(I),I=1,NP)
WRITE(6,100)
100  FORMAT(1X,6X,'ITER',6X,'PERF.INDX',8X,' NEW K1',8X,'NEW K2'//)
DO 20 ITTR = 1,100
SCALE = .TRUE.
IPASS = 0
3    DO 5 K = 1,NP
B(K) = 0.
DO 4 KK = 1,NP
4    A(K,KK) =0.
5    CONTINUE
CALL COEF(VK)
IF(SCALE) THEN
DO 6 I = 1,NP
B(I) = B(I) * VK(I)
DO 6 J = 1,NP
A(I,J) = A(I,J) * VK(I) * VK(J)
6    CONTINUE
END IF
L = 0
DO 7 I = 1,NP
DO 7 J = 1,NP
L = L + 1
7    AX(L) = A(I,J)
CALL MATINV(AX,NP,L1,L2)
L = 0
DO 9 I = 1,NP
DXT = 0
DO 8 J = 1,NP
L = L + 1

```

```

8   DXT = DXT + AX(L) * B(J)
   XK(I) = DXT
9   CONTINUE
   DKMAX = 0
   WRITE(*,*) 'XK',(XK(I),I=1, NP)
   DO 14 K = 1, NP
   IF(DABS(XK(K)).GT.DKMAX) DKMAX = DABS(XK(K))
   DK(K) = XK(K)
   IF(DABS(DK(K)).GT.0.5) DK(K)=DSIGN(0.5D0,DK(K))
14  CONTINUE
   IF(DKMAX.LT.0.001) GO TO 21
   SIGNX = 1.0
15  DO 16 K = 1, NP
   TK(K) = VK(K)*(1. + DK(K))
16  CONTINUE
   CALL FUNCT(TK,VAL)
   WRITE(*,*) 'TK,VAL',(TK(I),I=1, NP),VAL
   IF(VAL.GT.VALK) THEN
   IF(IPASS.EQ.0) THEN
   IPASS = 1
   DO 17 K = 1, NP
17  DK(K) = -DK(K)
   GO TO 15
   ELSE IF(IPASS.EQ.1) THEN
   IPASS = 2
   IF(DKMAX.GT.0.5) THEN
   DO 13 K = 1, NP
13  XK(K) = XK(K)/(2.*DKMAX)
   END IF
   DO 12 K = 1, NP
12  DK(K) = XK(K)
   GOTO 15
   END IF
   END IF
   DO 22 K = 1, NP
22  VK(K) = TK(K)
   VALK = VAL
   WRITE(6,101) ITTR,VAL,(VK(I),I=1, NP)
101 FORMAT(1X,I5,F12.5,5G13.5)
20  CONTINUE
21  CONTINUE
99  STOP
   END
   SUBROUTINE COEF(VK)
   IMPLICIT REAL*8(A-H,O-Z)
   COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1     FS(10,10),RS(10,10),BS(10,10),
1     XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2     L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3     F(10),R(10),A(10,10),B(10),Q(10)
   LOGICAL ERROR
   DIMENSION Y(2,9),T(9),VK(10)
   DATA T/ 0., 5.63E-4, 11.32E-4, 16.97E-4, 22.62E-4, 34.E-4,
1  39.7E-4, 45.2E-4, 169.7E-4/
   DATA (Y(1,I),I=1,9)
1  /1.,0.828, 0.704, 0.622, 0.565, 0.499, 0.482, 0.470, 0.443/
   DATA (Y(2,I),I=1,9)
1  /0.,0.0737,0.113, 0.1322,0.1400,0.1468,0.1477,0.1477,0.1476/
10  DO 10 K = 1, NP
   BK(K) = VK(K)
   DO 20 K = 1, NN
20  X(K) = Y(K,1)

```

```

DO 30 I = 1,NN
DO 30 J = 1,NP
30 DXDK(I,J) = 0.
DT = 0
IFLAG = 1
TIMSTR = T(1)
DO 100 IEXPT = 2,NEXPT
TIMEND = T(IEXPT)
CALL IEULER(TIMSTR,TIMEND)
DO 9 K = 1,NP
DO 6 I = 1,NN
6 B(K) = B(K) + DXDK(I,K) *Q(I)* (Y(I,IEXPT)-X(I))
DO 7 KK = 1,NP
DO 7 I = 1,NN
7 A(K,KK) = A(K,KK) + DXDK(I,K) *Q(I)* DXDK(I,KK)
9 CONTINUE
TIMSTR = TIMEND
100 CONTINUE
RETURN
END
SUBROUTINE FUNCT(VK,VALUE)
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1 FS(10,10),RS(10,10),BS(10,10),
1 XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2 L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3 F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
DIMENSION Y(2,9),T(9),VK(10)
DATA T/ 0., 5.63E-4, 11.32E-4, 16.97E-4, 22.62E-4, 34.E-4,
1 39.7E-4, 45.2E-4, 169.7E-4/
DATA (Y(1,I),I=1,9)
1 /1.,0.828, 0.704, 0.622, 0.565, 0.499, 0.482, 0.470, 0.443/
DATA (Y(2,I),I=1,9)
1 /0.,0.0737,0.113, 0.1322,0.1400,0.1468,0.1477,0.1477,0.1476/
VALUE= 0.
DO 1 K = 1,NP
1 BK(K) = VK(K)
DO 2 K = 1,NN
2 X(K) = Y(K,1)
TIMSTR = T(1)
IFLAG = 0
DT = 0.
DO 10 I = 2,NEXPT
TIMEND = T(I)
CALL IEULER(TIMSTR,TIMEND)
DO 7 K = 1,NN
VALUE = VALUE + (Y(K,I)-X(K))**2 * Q(K)
7 CONTINUE
TIMSTR = TIMEND
10 CONTINUE
RETURN
END
SUBROUTINE IEULER(TA,TB)
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1 FS(10,10),RS(10,10),BS(10,10),
1 XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2 L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3 F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
IF(DT.EQ.0) DT = (TB-TA)/10.

```

```

T = TA
IF((T+DT).GT.TB) DT = TB - TA
DO 100 IT = 1,10000
DO 10 I = 1,NN
10 XS(I) = X(I)
15 CALL INTGRT
DTF = DT
IF(ERROR) THEN
DT = DT * 0.5
DO 16 I = 1,NN
16 X(I) = XS(I)
GOTO 15
END IF
DO 20 I = 1,NN
XG(I) = X(I)
20 X(I) = XS(I)
DT = DT * 0.5
DO 30 II = 1,2
CALL INTGRT
IF(ERROR) THEN
DO 25 I = 1,NN
25 X(I) = XS(I)
GOTO 15
END IF
30 CONTINUE
ERR = 0
DO 40 I = 1,NN
40 IF(DABS((XG(I)-X(I))/X(I)).GT.ERR) ERR = DABS((XG(I)-X(I))/X(I))
IF(ERR.LE.TOL) THEN
T = T + DTF
DO 45 I = 1,NN
45 X(I) = XG(I)
DT = DTF
IF(IFLAG.EQ.1) CALL INTSEN
IF(T.EQ.TB) GOTO 110
IF(ERR.EQ.0) THEN
DT = 5.* DTF
ELSE
DT = TOL/ERR * DTF
END IF
IF(DT.GT.5.*DTF) DT = 5.*DTF
IF((T+DT).GT.TB) DT = TB - T
ELSE IF(ERR.GT.TOL) THEN
DT = DTF * 0.5
DO 50 I = 1,NN
50 X(I) = XS(I)
END IF
100 CONTINUE
110 CONTINUE
RETURN
END
SUBROUTINE INTSEN
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1 FS(10,10),RS(10,10),BS(10,10),
1 XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2 L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3 F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
RDT = 1./DT
CALL EVDFDK
DO 10 J = 1,NP

```

```

DO 10 I = 1,NN
10  RS(I,J) = 0.0
    CALL DIFFDX
    L = 0
    DO 22 I = 1,NN
    DO 22 J = 1,NN
    L = L + 1
    AX(L) = DFDX(I,J)
22  IF(I.EQ.J) AX(L) = AX(L) - RDT
    CALL MATINV(AX,NN,L1,L2)
    DO 200 IT = 1,100
        CALL EVALFD
        BMAX = 0
        DO 110 J = 1,NP
        DO 110 I = 1,NN
        BS(I,J) = RS(I,J) - FS(I,J)
        IF(IT.GT.1) THEN
            ERR = 0.
            IF(RS(I,J).NE.0) ERR = DABS(BS(I,J)/RS(I,J))
            IF(ERR.GT.BMAX) BMAX = ERR
        END IF
110  CONTINUE
        IF(IT.GT.1.AND.BMAX.LT.0.0001) GOTO 210
        DXMAX = 0
        DO 150 J = 1,NP
        DO 120 I = 1,NN
120  BP(I) = BS(I,J)
        L = 0
        DO 125 II= 1,NN
        DXT = 0
        DO 124 JJ= 1,NN
        L= L + 1
124  DXT = DXT + AX(L) * BP(JJ)
        DX(II) = DXT
125  CONTINUE
        IF(IT.GT.1) THEN
            DO 126 I = 1,NN
            ERR = 0
            IF(DXDK(I,J).NE.0.) ERR = DABS(DX(I)/DXDK(I,J))
            IF(ERR.GT.DXMAX) DXMAX = ERR
126  CONTINUE
        END IF
        DO 127 I = 1,NN
        RS(I,J) = RS(I,J) + DX(I) * RDT
127  DXDK(I,J) = DXDK(I,J) + DX(I)
150  CONTINUE
        IF(IT.GT.1.AND.DXMAX.LT.0.001) GOTO 210
200  CONTINUE
        WRITE(*,*) ' FAILED TO INTEGRATE SENSITIVITY EQUATIONS '
210  CONTINUE
    RETURN
    END
    SUBROUTINE INTGRT
    IMPLICIT REAL*8(A-H,O-Z)
    COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1      FS(10,10),RS(10,10),BS(10,10),
1      XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2      L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3      F(10),R(10),A(10,10),B(10),Q(10)
    LOGICAL ERROR
    ERROR = .FALSE.
    RDT = 1./DT

```

```

DO 10 I = 1,NN
R(I) = 0.
DO 50 IT = 1,100
CALL EVALF
BMAX = 0
DO 15 I = 1,NN
BP(I) = R(I) - F(I)
IF(IT.GT.1) THEN
ERR = DABS(BP(I)/R(I))
IF(ERR.GT.BMAX) BMAX = ERR
END IF
15 CONTINUE
IF(IT.GT.1.AND.BMAX.LT.0.0001) GO TO 60
CALL DIFFDX
DO 20 I = 1,NN
20 DFDX(I,I) = DFDX(I,I) - RDT
L = 0
DO 23 I = 1,NN
DO 23 J = 1,NN
L = L + 1
23 AX(L) = DFDX(I,J)
CALL MATINV(AX,NN,L1,L2)
L = 0
DO 25 I = 1,NN
DXT = 0
DO 24 J = 1,NN
L = L + 1
24 DXT = DXT + AX(L) * BP(J)
DX(I) = DXT
25 CONTINUE
DXMAX = 0
DO 26 I = 1,NN
IF(X(I).NE.0.) THEN
ERR = DABS(DX(I)/X(I))
ELSE
ERR = 0.
END IF
IF(ERR.GT.DXMAX) DXMAX = ERR
26 CONTINUE
DO 30 I = 1,NN
R(I) = R(I) + DX(I) * RDT
30 X(I) = X(I) + DX(I)
IF(IT.GT.1.AND.DXMAX.LT.0.001) GOTO 60
50 CONTINUE
ERROR = .TRUE.
60 CONTINUE
RETURN
END
SUBROUTINE EVALF
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1 FS(10,10),RS(10,10),BS(10,10),
1 XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2 L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3 F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
R1= BK(1)*(X(1)*X(1)-X(2)*(2.-2.*X(1)-X(2)))/0.726)
R2= BK(2)*(X(1)*X(2)-(1.-X(1)-2*X(2))*(2.-2.*X(1)-X(2)))/3.852)
F(1) = -R1 - R2
F(2) = R1*0.5 - R2
RETURN
END

```

```

SUBROUTINE EVALFD
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1      FS(10,10),RS(10,10),BS(10,10),
1      XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2      L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3      F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
DO 30 I = 1,NN
DO 30 J = 1,NP
FD = DFDK(I,J)
DO 20 K = 1,NN
20  FD = FD + DFDX(I,K) * DXDK(K,J)
FS(I,J) = FD
30  CONTINUE
RETURN
END
SUBROUTINE DIFFDX
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1      FS(10,10),RS(10,10),BS(10,10),
1      XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2      L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3      F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
DR1DX1 = BK(1)*(2.*X(1)+2.*X(2)/0.726)
DR1DX2 = BK(1)/0.726*(-(2.-2.*X(1)-X(2)) + X(2))
DR2DX1 = BK(2)*(X(2)+(2.-2.*X(1)-X(2))/3.852
1          +2.*(1.-X(1)-2.*X(2))/3.852)
DR2DX2 = BK(2)*(X(1)+2.*(2.-2.*X(1)-X(2))/3.852
1          +(1.-X(1)-2.*X(2))/3.852)
DFDX(1,1) = - DR1DX1 - DR2DX1
DFDX(1,2) = - DR1DX2 - DR2DX2
DFDX(2,1) = 0.5 * DR1DX1 - DR2DX1
DFDX(2,2) = 0.5 * DR1DX2 - DR2DX2
RETURN
END
SUBROUTINE EVDFDK
IMPLICIT REAL*8(A-H,O-Z)
COMMON NEXPT,NN,NP,DT,RDT,ERROR,TOL,IFLAG,
1      FS(10,10),RS(10,10),BS(10,10),
1      XS(10),XG(10),X(10),DX(10),DFDX(10,10),BP(10),L1(10),
2      L2(10),BK(10),AX(100),DFDK(10,10),DXDK(10,10),
3      F(10),R(10),A(10,10),B(10),Q(10)
LOGICAL ERROR
DR1DK1 = X(1)*X(1) - X(2)*(2.-2.*X(1)-X(2))/0.726
DR2DK2 = X(1)*X(2) - (1.-X(1)-2.*X(2))*
1      (2.-2.*X(1)-X(2))/3.852
DFDK(1,1) = -DR1DK1
DFDK(1,2) = -DR2DK2
DFDK(2,1) = 0.5 * DR1DK1
DFDK(2,2) = - DR2DK2
RETURN
END

```