ROBUST NONLINEAR REGRESSION FOR PARAMETER ESTIMATION IN PRESSURE TRANSIENT ANALYSIS

A REPORT SUBMITTED TO THE DEPARTMENT OF ENERGY RESOURCES ENGINEERING OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

By
Parag Bandyopadhyay
June 2014
I certify that I have read this report and that in my opinion it is fully adequate, in scope and in quality, as partial fulfillment of the degree of Master of Science in Petroleum Engineering.

Prof. Roland Horne
(Principal Advisor)
Abstract

Parameter estimation in pressure transient analysis is used to match analytical transient flow models to measured field data. This matching of nonlinear models to observed data is also referred to as regression. Ordinary least squares (OLS) is the most commonly used regression method. However, the assumptions inherent to OLS that; a) errors are present only in the dependent variable (pressure) and b) these errors follow a Gaussian distribution, may make it unsuitable for certain data sets.

In this research report, the development of methods that address the possibility of errors in both pressure and time variables is discussed first. These methods were tested and compared to OLS and found to provide more accurate estimates in cases where random time errors are present in the data. These methods were then modified to consider errors in breakpoint flow rate measurement.

OLS parameter estimates for datasets with non-Gaussian error distributions are shown to be biased. A general method was developed based on maximum likelihood estimation theory that estimates the error distribution iteratively and uses this information to estimate parameters. This method was compared to OLS and found to be more accurate for cases with non-Gaussian error distributions.

In the final chapter, we discuss issues relating to computational performance such as hybrid methods for efficient and robust parameter estimation and scaling of methods with increasing problem size. Stochastic iteration methods, which are used commonly in machine learning problems, were adapted for use with the methods developed in the report. These methods were shown to be computationally efficient for larger problems while maintaining accuracy.
Acknowledgments

I would like to thank my advisor, Prof. Roland Horne for his patience, encouragement and advice. I also thank the Stanford University Petroleum Research Institute: Innovation in Well Testing (SUPRI-D) affiliates for funding this research and the students for providing a regular forum for brainstorming and intellectual debugging.

I thank my friends and classmates in the Energy Resource Engineering department, Cyril, Wenyue, Kevin, Steven, Joseph, Spencer, Chris, Dylan, Scott and Jack; I will cherish their friendship for the years to come.

I also acknowledge my parents for their encouragement and support throughout my education.
Contents

Abstract .............................................................................................................................................. v
Acknowledgments ........................................................................................................................... vii
Contents ............................................................................................................................................. ix
List of Tables ....................................................................................................................................... xi
List of Figures ...................................................................................................................................... xiii
1. Introduction .................................................................................................................................. 1
   1.1. Problem Statement .................................................................................................................. 1
   1.2. Prior Work .............................................................................................................................. 3
   1.3. Outline of the Report .............................................................................................................. 4
2. Nonlinear Regression: Formulation and Solution ............................................................................. 7
   2.1. Pressure Transient Analysis: Background ............................................................................. 7
   2.2. Least Squares Formulation ................................................................................................. 9
   2.3. Solving the Nonlinear Optimization Problem ....................................................................... 10
       2.3.1. The Gauss-Marquardt Method ..................................................................................... 11
       2.3.2. Diagonal Scaling as Preconditioner ............................................................................ 13
       2.3.3. Line Search .................................................................................................................. 13
   2.4. Inherent Assumptions of OLS ............................................................................................... 14
3. Orthogonal Distance Regression .................................................................................................... 15
   3.1. The General ODR Problem ................................................................................................... 15
   3.2. Dastan’s Orthogonal Distance Regression .......................................................................... 17
   3.3. Boggs Regression ............................................................................................................... 18
       3.3.1. Efficient Solution of the Boggs System ........................................................................ 20
   3.4. Iterative Ratio of Variance Estimation ................................................................................. 22
   3.5. Performance Comparison ...................................................................................................... 24
       3.5.1. Random Time Errors ................................................................................................... 24
       3.5.2. Breakpoint Error ........................................................................................................... 30
4. Breakpoint Regression ..................................................................................................................... 32
   4.1. Breakpoint Regression .......................................................................................................... 32
   4.2. Flow Rate Regression (FRR) .............................................................................................. 34
   4.3. Hybrid ODR .......................................................................................................................... 35
   4.4. Performance Comparison ...................................................................................................... 37
       4.4.1. Breakpoint Error ......................................................................................................... 37
       4.4.2. Flow Rate Error .......................................................................................................... 41
5. General Likelihood Maximization .................................................................................................... 45
   5.1. GLM for Errors in Pressure ................................................................................................... 45
   5.2. GLM for Errors in Pressure and Time .................................................................................. 50
   5.3. Performance Comparison ...................................................................................................... 52
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.1. Examples</td>
<td>53</td>
</tr>
<tr>
<td>5.3.2. Statistics of Starting Guesses</td>
<td>55</td>
</tr>
<tr>
<td>6. Performance and Scaling</td>
<td>58</td>
</tr>
<tr>
<td>6.1. Hybridization with OLS</td>
<td>58</td>
</tr>
<tr>
<td>6.2. Scaling with Problem Size</td>
<td>60</td>
</tr>
<tr>
<td>6.2.1. Stochastic Methods</td>
<td>62</td>
</tr>
<tr>
<td>7. Conclusion</td>
<td>66</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>69</td>
</tr>
<tr>
<td>References</td>
<td>71</td>
</tr>
<tr>
<td>A. Test Matrix</td>
<td>73</td>
</tr>
</tbody>
</table>
### List of Tables

Table 3.1: Results of Monte Carlo estimation of parameters for Datasets 1-8 with only pressure noise up to 5% standard deviation of the maximum pressure change.  

Table 3.2: Results of Monte Carlo estimation of parameters of Datasets 1-8 with only time error with 1 hour standard deviation.  

Table 3.3: Results of Monte Carlo estimation of parameters of Datasets 1-8 with pressure and time error.  

Table 4.1: Median error of parameters estimated during 30 Monte Carlo estimation runs for Datasets 4-6 from the Test Matrix, using BPR and ODR+BPR, with up to three hours of breakpoint error.  

Table 4.2: Results of 30 Monte Carlo runs with up to 3 hours of breakpoint error, optionally with 1 psi pressure and 30 minutes of time noise.  

Table 4.3: Median error of parameters estimated during 30 Monte Carlo estimation runs for Datasets 4-6 from the Test Matrix, using BP+FRR and BP+FRR+ODR, with up to three hours of breakpoint error and 10% flow rate measurement error.  

Table 4.4: Results of Monte Carlo runs with up to 3 hours of breakpoint error and 10% flow rate measurement error, optionally with 1 psi pressure and 30 minutes of time noise.  

Table A-1: Known model parameters for all data sets in the test matrix.  

Table A-2: Description of data sets in the test matrix.
List of Figures

Figure 1.1: a) Data from PDG indicates only two usable sections for conventional buildup analysis. b) A closer look at a section of the data indicates that the measurements are fluctuating continuously. Reproduced from Liu (2013). .................................................. 2

Figure 2.1: a) The unit flow rate response for an infinite-acting reservoir, b) the pressure response for the corresponding variable flow rate calculated using the superposition principle. ........................................................................................................... 9

Figure 3.1: Example of breakpoint error resulting in inadequate match to pressure data using ordinary least squares (OLS). ................................................................................................................. 15

Figure 3.2: Schematic of the sparsity pattern of the Boggs Hessian matrix for a problem with three model parameters and ten data points. ........................................................................................................... 21

Figure 3.3: Error in ODR parameter estimate (for Dataset 1) for assumed values of the ratio of variance of errors. ......................................................................................................................... 22

Figure 3.4: Example of ODR fit with residuals reflecting the true statistics for noise in pressure and time. (Dataset 8 with synthetic noise) ................................................................................................. 23

Figure 3.5: Results of 100 Monte Carlo runs using OLS, ODR and Boggs on Dataset 1 with variations of pressure and time noise. .......................................................................................................... 25

Figure 3.6: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, ODR and Boggs methods, for Dataset 1 with no noise. ........ 29

Figure 3.7: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, ODR and Boggs methods, for Dataset 1 with maximum pressure and time noise. ................................................................. 29

Figure 3.8: Effect of adding breakpoint error to Dataset 1 on the permeability estimate using OLS, ODR and Boggs methods. .................................................................................................................. 30

Figure 3.9: Maps of percentage permeability estimation error for a range of starting guesses using OLS, ODR and Boggs methods, for Dataset 1 with three hours of breakpoint noise. ........................................................................................................................................ 31
Figure 4.1: Effect of flow rate measurement error on OLS match to the data. ............. 34
Figure 4.2: Effect of increasing breakpoint error (Dataset 1) on the estimated permeability using OLS, ODR and BPR. .................................................................................................................. 37
Figure 4.3: Statistics of starting guesses for OLS, ODR and BPR for breakpoint error of three hours in Dataset 1 .............................................................................................................. 38
Figure 4.4: Statistics of starting guesses for OLS, ODR and BPR for breakpoint error of 1 hour and 10% flow rate error in Dataset 1 .......................................................................................... 41
Figure 5.1: Dataset 1 with lognormal noise, the actual model and the OLS match........ 46
Figure 5.2: Histogram of a) pressure residuals of OLS fit b) original pressure error from the example in Figure 5.1 .................................................................................................................. 47
Figure 5.3: Example of arbitrary distribution of errors derived from model match to data and the corresponding pdf defined by piecewise polynomial splines.............................. 48
Figure 5.4: Lognormal noise in Dataset 1, the histogram of synthetic errors and the final spline pdf function is shown in the right. The noisy data, actual model, OLS and GLM match are compared on the left .......................................................... 53
Figure 5.5: Noise and outliers centered around 5 psi in Dataset 1, the histogram of synthetic errors and the final spline pdf function is shown on the right. The noisy data, actual model, OLS and GLM match are compared on the left .......................................................... 54
Figure 5.6: Outlier-only noise centered around 3 and -3 psi in Dataset 1, the histogram of synthetic errors and the final spline pdf function is shown on the right. The noisy data, actual model, OLS and GLM match are compared on the left .......................................................... 55
Figure 5.7: Outlier-only noise centered around 3 and -3 psi in Dataset 8, the histogram of synthetic errors and the final spline pdf function is shown on the right. The noisy data, actual model, OLS and GLM match are compared on the left .......................................................... 55
Figure 5.8: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, GLM and OLS+GLM methods, for Dataset 1 with outliers averaging 5 psi .................................................................................................................. 56
Figure 5.9: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, GLM and OLS+GLM methods, for Dataset 1 with multi-modal outliers around 5 and -5 psi .................................................................................................................. 57
Figure 6.1: Example of the benefits of hybridization of OLS with GLM (repeated from Figure 5.9).

Figure 6.2: Map of permeability error for Dataset 1 with three hours of breakpoint error for varying starting guesses, computed using OLS, ODR and hybrid OLS+ODR.

Figure 6.3: Map of computation time for Dataset 1 with three hours of breakpoint error for varying starting guesses, computed using OLS, ODR and hybrid OLS+ODR.

Figure 6.4: Plot of computing time for OLS, Boggs and Boggs with Schur complement (SC) solver with increasing number of data points using parameters of Dataset 1.

Figure 6.5: Comparison of computing time with increasing problem size using OLS, ODR and BoggsSC. The parameters from Dataset 1 were used to generate the data.

Figure 6.6: Comparison of scaling with problem size and parameter errors using full OLS and stochastic OLS for Dataset 1 with no additional noise.

Figure 6.7: Comparison of scaling with problem size and parameter errors using full OLS and stochastic OLS for a Dataset 1 with pressure noise of 1 psi standard deviation.

Figure 6.8: Comparing the mean and standard deviation of the permeability estimate (from 50 runs and noise realizations) using stochastic OLS for varying number of stochastic data points.

Figure 6.9: Plot of standard deviation of the permeability estimate using stochastic OLS (from 50 runs) on Dataset 1 with varying levels of noise in the data and varying number of stochastic points per iteration.

Figure A-1: Flow rate history used for the datasets in the test matrix.
Chapter 1

1. Introduction

1.1. Problem Statement

Pressure is one of the most important data streams for reservoir diagnostics and monitoring. Traditionally, well testing is carried out as a controlled-condition downhole flow experiment, with periods of flow and no flow enforced by surface or downhole valves. The resulting pressure response ‘discovers’ reservoir features such as faults and boundaries as the signal travels further away from the well into the reservoir. Interpretation of these transient data yields information about the length scale of these reservoir features and their relative configuration. The most common well and reservoir properties interpreted from pressure transient data are wellbore storage, skin, reservoir permeability, distance to and nature of reservoir boundaries.

The introduction of permanent downhole gauge (PDG) technology has allowed continuous monitoring of downhole pressure, temperature and flow-rates. This continuous surveillance allows monitoring of changes to well and reservoir conditions and enables proactive control and optimization of recovery. However, the data from these PDGs are a surveillance of current operations rather than a controlled flow test. The flow rate changes may not be perfect step-changes as prescribed by traditional well test procedure. Constant fluctuation in operating conditions causes transient responses in both pressure and flow rate. This complicates the interpretation of these data. The interpretation is complicated further by the presence of other nonidealities such as noise in pressure and flow rate measurements and sections with missing data due to malfunctioning equipment. The presence of these nonideal artifacts is more likely for PDG data due to the large timescales over which data are collected; an example is indicated in Figure 1.1.
Figure 1.1: a) Data from PDG indicates only two usable sections for conventional buildup analysis. b) A closer look at a section of the data indicates that the measurements are fluctuating continuously. Reproduced from Liu (2013).
Pressure transient analysis involves two steps: a) model identification and b) parameter estimation. Based on inspection of the data, diagnostic plots and prior knowledge of the reservoir flow system, a transient flow model is identified. In the parameter estimation step, estimates of reservoir and well parameters such as permeability, skin and wellbore storage are made by matching the measured data to the reservoir model chosen in the previous step. The focus of this research has been the parameter estimation step.

Nonlinear regression techniques are used for parameter estimation in many petroleum engineering applications where models must be calibrated to data. The most commonly used regression method is the ordinary least squares (OLS). This method is mathematically convenient, but incorporates certain assumptions about the nature of the errors in the data. OLS assumes errors only in the dependent variables and assumes them to follow a Gaussian distribution. For large, imperfect data sets such as measurements made using a permanent downhole gauge (PDG), neither assumption may necessarily be valid. There may be measurement errors in multiple measured variables, measurement biases over time and intervals with missing data.

The focus of this research was to develop regression methods that are robust to the nonidealities expected from PDG data. The assumptions of the ordinary least squares method were examined and alternatives that bypass these assumptions were developed.

1.2. Prior Work

Nonlinear parameter estimation in well test analysis dates back to the use of semilog graph paper and straight line extrapolations to estimate permeability and skin were common. Specific model identification was dependent on analysis limited to straight lines on a semilog graph. Introduced in the 1980s, the Bourdet derivative plot (1989) allowed identification of specific flow regimes and the estimation of parameters from the pertinent sections of the data. Rosa (1983) described a method to calculate parameter gradients in the Laplace space for use in computer-aided nonlinear regression based on the ordinary
least squares method. The choice of solution methods such as the Gauss-Marquardt method and line search in this research were inspired by Rosa’s work.

Dastan (2009, 2010) introduced Orthogonal Distance Regression (ODR) for robust parameter estimation in pressure transient analysis. Itthisawatpan (2012) discussed the Boggs et. al. (1987) method as an alternative to ODR, and demonstrated a gain in efficiency for smaller problems. This research follows Dastan’s and Itthisawatpan’s work to investigate the advantages of using ODR methods for well testing problems with errors in the time variable.

This investigation considered errors in the flow rate history, which is often assumed to be error-free in other studies. Previous work on errors in the flow rate record consists of aberrant section identification, Athichanagorn (2000), and breakpoint identification, Nomura (2006). A method that parameterizes the flow rate history and adjusts these parameters to better match the measured pressure data is demonstrated and compared with OLS and ODR methods.

1.3. Outline of the Report

This report is organized as follows; Chapter 2 begins with a brief introduction to the analytical pressure transient function for a multi rate flow rate signal. This analytical function is the prototype for the functions that are to be fitted to data. The ordinary least squares formulation for the regression/history matching problem is developed from its origins in maximum likelihood estimation theory. The assumptions of the OLS method are singled out for focus in the following chapters.

Chapter 3 introduces and discusses methods that assume random error in the independent time variable. The accuracy and sensitivity to starting guesses are examined for these methods.

Chapter 4 focuses on methods that consider error in the flow rate history, a variable that is often assumed to be error-free. Breakpoint and flow rate regression methods are
introduced and their accuracy is briefly compared to OLS for cases with error in the flow rate record.

Chapter 5 describes the formulation of a general maximum likelihood estimator for arbitrary distribution of errors in the dependent variable. The formulation was extended to consider errors in the independent variable (time) as well. Chapter 6 focuses on the convergence behavior and computational performance of the methods discussed in the previous chapters. Hybrid strategies for robust convergence and efficient computation were tested. The report is concluded in Chapter 7 and possible future work is discussed.
Chapter 2

2. Nonlinear Regression: Formulation and Solution

We begin this chapter with a short background on the origin and form of the analytical solutions of pressure transient problems. The superposition principle is introduced as the way of expressing the analytical unit-rate solution for a variable flow rate history. Then, the least-squares objective function for matching these models to observed field data is derived from its basis in the Gaussian likelihood function. Techniques for minimization of the objective function by optimization are discussed, as are auxiliary strategies for efficient and accurate solution.

The chapter is concluded with a look at two of the most significant assumptions behind OLS, the following chapters indicate our efforts to work around these two assumptions.

2.1. Pressure Transient Analysis: Background

The analytical solutions for a pressure transient under a constant, unit-rate change are derived from the diffusion equation for slightly compressible, single-phase fluids in a homogeneous, isotropic porous medium (Equation 2.1).

\[ \nabla p = \frac{\phi \mu c_t}{k} \frac{\partial p}{\partial t} \] (2.1)

This equation is solved for boundary conditions representing specific well and reservoir configurations using a suitable coordinate system. The self-symmetry of the problem usually leads to a reduced problem in dimensionless pressure and time variables which are solved for a unit step-change flow rate impulse.

This solution is often obtained by solving Equation 2.1 in Laplace space for the appropriate boundary conditions. The Laplace space solution is inverted to real variables using
numerical algorithms such as the Stehfest (1970) algorithm. The methods discussed in this research report used the Stehfest transform with eight coefficients to obtain pressure solutions in the time domain.

This unit-rate pressure solution can be expressed as a general function of time and certain reservoir and fluid parameters. An example of one such function with known and unknown variables is indicated in Equation 2.2.

\[
P_u(t, \alpha) = f(t, k, S, R_e, \phi, \mu, c_t, r_w)
\]  

(2.2)

This unit-rate pressure function varies in form and input variables depending on the type of model i.e. flow media and boundary conditions. For dual-porosity reservoirs, the variables for storativity (ω) and transmissivity ratios (λ) are introduced.

The partial differential equation for single-phase, slightly compressible flow is linear. Therefore, the solution for a variable flow rate can be found by convolution of the unit flow rate solution with the variable flow rate history, as indicated in Equation 2.3. Figure 2.1 shows an example of the application of superposition to obtain the analytical pressure solution for a variable flow rate history. Here, \( \tau \) indicates a breakpoint, i.e. the time at which a change in flow rate occurs.

\[
\Delta p(t, q, \alpha) = \int_0^t q(t - \tau) \frac{dP_u(\tau, \alpha)}{d\tau} d\tau = \sum_{i=0}^n (q_{\tau+1} - q_\tau)P_u(t - \tau_i, \alpha)
\]  

(2.3)
Figure 2.1: a) The unit flow rate response for an infinite-acting reservoir, b) the pressure response for the corresponding variable flow rate calculated using the superposition principle.

This discussion on the superposition theorem will be referenced and expanded on in upcoming chapters on the discussion of breakpoint error and relevant regression methods.

2.2. Least Squares Formulation

The data collected during a well test include pressure and flow rate measured at a corresponding time. The pressure data collected from the field are designated $p_{di}$, and the analytical pressure function is defined as a function of time, flow rate history and reservoir parameters. The time and flow rate history are assumed to be independent variables. If some pressure data ($p_{di}$) is observed at corresponding time instant ($t_i$), we express the residual pressure error in measurement as shown in Equation 2.4. Assuming a Gaussian distribution, the probability density function for these residuals is indicated in Equation 2.5.

$$\epsilon_{pi} = p_{di} - p(t_i, \bar{q}, \bar{a})$$

$$f(\epsilon_{pi}) = -\frac{1}{\sigma_p \sqrt{2\pi}} \exp\left(-\frac{\epsilon_{pi}^2}{2 \sigma_p^2}\right)$$

(2.4)   

(2.5)
Assuming the errors to be independent and identically distributed for each observation in the data, the likelihood of the parameter set is expressed as the product of the probability of each individual residual, as indicated in Equations 2.6-2.7.

\[
L(\bar{\alpha} | \bar{p_d}) = \prod_{i=1}^{n} f(\epsilon_{pi})
\]  

\[
L(\bar{\alpha} | \bar{p_d}) = \prod_{i=1}^{n} \frac{1}{\sigma_p \sqrt{2\pi}} \exp\left(-\frac{(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))^2}{2 \sigma_p^2}\right)
\]  

(2.6)  

(2.7)

This maximum likelihood estimation of parameters is posed as an optimization problem. The problem can be linearized in part by considering the log likelihood, which yields the objective function indicated in Equation 2.8, which must be minimized.

\[
\min_{\alpha} E(\alpha) = \sum_{i=1}^{n} (p_{di} - p(t_i, \bar{q}, \bar{\alpha}))^2
\]

(2.8)

For the given data, the parameter set that maximizes the likelihood function is the least squares ‘best estimate’. The objective function takes the form of the sum of squares of the residuals in the data at each step. This objective function is to be minimized, leading to the more common name for this method, ‘Least Squares’, which will be referred to here as Ordinary Least Squares (OLS)

2.3. Solving the Nonlinear Optimization Problem

For a linear model, the least squares optimization problem is linear and can be solved in a single step. Nonlinear functions such as the analytical pressure solutions must be solved iteratively. The solutions for nonlinear optimization problems fall into two major categories: local and global methods. Local search methods require a starting guess and are informed by the first and/or second derivatives of the function at that point until a local solution is found.
Global search methods employ heuristics that combine direct sampling and local search to find solutions that are likely to be unbiased by the starting guess, if one is needed at all. ‘Global’ search is a misnomer because the global optimum is not guaranteed in nonlinear problems, simply more likely to be discovered using these methods.

For parameter estimation in well test analysis, local search methods are combined with a reasonably good starting guess from diagnostic plots. We used Newton-type local methods for solving the optimization problem for the least squares formulation as well as the techniques described later in this report. The solution methods and supporting algorithms are discussed next.

2.3.1. The Gauss-Marquardt Method

All Newton-type, nonlinear local optimization methods require a starting guess of the unknown parameters. The gradient vector and Hessian matrix at this starting point are calculated as indicated in Equations 2.9 and 2.10.

\[
g_j = \frac{\partial E}{\partial \alpha_j} = -2 \sum_{i=1}^{n} (p_{di} - p(t_i, \bar{q}, \bar{\alpha})) \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_i}
\]

\[
H_{jk} = \frac{\partial^2 E}{\partial \alpha_j \partial \alpha_k} = -2 \sum_{i=1}^{n} \left[ (p_{di} - p(t_i, \bar{q}, \bar{\alpha})) \frac{\partial^2 p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j \partial \alpha_k} - \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_k} \right]
\]

The update is calculated as indicated in Equations 2.11 and 2.12.

\[
H \Delta \bar{\alpha}^k = -g
\]

\[
\bar{\alpha}^{k+1} = \bar{\alpha}^k + \rho \Delta \bar{\alpha}^k
\]

This is the Newton method. The calculations in Equations 2.9 -2.12 are iterated and the parameters are updated until the objective function is sufficiently minimized or some convergence criterion is met. The convergence criteria used for the methods in this research
were minimum change of $10^{-5}$ in the objective function and maximum 5% change in the parameter values between iterations, up to a maximum of 40 iterations.

Calculation of the parameter update is one of the key steps in this process. The Newton method is not guaranteed to converge unless the Hessian is strictly positive-definite. To ensure positive-definiteness the second-order derivatives are removed from the Hessian calculation in Equation 2.10. This is known as the Gauss-Newton method (Equation 2.13). Other methods such as the Newton-Greenstadt and the Newton-Barua (1988) modify the eigenvalues of the original Hessian during solution to ensure convergence. The methods described in this research used the Gauss-Newton approximation for the Hessian with pressure derivatives calculated numerically using a central difference of $10^{-8}$.

$$ H_{jk} \approx +2 \sum_{i=1}^{n} \frac{\partial p(t_i, \bar{q}, \bar{a})}{\partial \alpha_j} \frac{\partial p(t_i, \bar{q}, \bar{a})}{\partial \alpha_k} $$

(2.13)

The Newton method has the property of quadratic convergence, however, it is slow farther away from the solution and faster as the solution is approached. In contrast, the steepest descent method uses the update rule indicated in Equation 2.14, where $\rho$ is an optimum step size.

$$ \Delta \bar{\alpha}^k = -\rho g $$

(2.14)

By definition, the steepest descent direction always reduces the objective function. However, this method converges faster farther away from the solution and gets slower as it approaches the solution.

The Levenberg-Marquardt method combines the steepest descent method and the Newton method to gain from their relative strengths. The Gauss-Marquardt update rule is indicated in Equation 2.15, where the Gauss-Newton approximation for the Hessian is used.

$$ (H + \lambda_{LM} \text{diag}(H))\Delta \bar{\alpha}^k = g $$

(2.15)

The Levenberg-Marquardt parameter ($\lambda_{LM}$) is used for interpolation between Newton and steepest descent directions. Small values of $\lambda_{LM}$ result in a search direction similar to the
Newton step and larger $\lambda_{LM}$ to the steepest descent step. For the first iteration, $\lambda_{LM}$ is set to 0.1 and is multiplied by a factor of ten for every successful iteration and divided by ten for every unsuccessful iteration, resulting in faster Newton-like convergence as we approach the solution.

### 2.3.2. Diagonal Scaling as Preconditioner

Rosa (1983) indicated a method for scaling of the Hessian matrix to improve the condition number, which helps in more accurate computation of the search direction. The scaled Hessian and gradient are computed as shown in Equations 2.16-2.17., and the update is calculated as shown in Equations 2.18-2.19

\[
HS_{jk} = \frac{H_{jk}}{\sqrt{H_{jj} H_{kk}}}
\]  
\text{(2.16)}

\[
GS_j = \frac{g_j}{\sqrt{H_{jj}}}
\]  
\text{(2.17)}

\[
(HS + \lambda I) \Delta \alpha_s = -GS
\]  
\text{(2.18)}

\[
\Delta \alpha_{j} = \frac{\Delta \alpha_{sj}}{\sqrt{H_{jj}}}
\]  
\text{(2.19)}

By definition, the scaled Hessian is the covariance matrix for the parameters and hence may also be used for calculation of confidence intervals for parameter estimates.

### 2.3.3. Line Search

Once the system of equations is solved to obtain a search direction, the next step is to determine a step length. This is a one-dimensional search problem that may be solved by
approximating the objective in the search direction as a quadratic function as indicated in Equation 2.20. The method used in this work was introduced originally by Bard (1970). The minimum for the quadratic approximation of the objective function can be found as indicated in Equation 2.21.

\[
\min_{ρ} E(ρ) \cong aρ^2 + bρ + c
\]  (2.20)

\[
ρ_{min} = -\frac{b}{2a} = -\frac{E(0)}{2(E(1) - E(0) - g ∆α^k)}
\]  (2.21)

If the resultant step succeeds in reducing the objective function, the optimum step size is recalculated based on the value of the objective function at the previous step size. This method was referred to by Rosa (1983) as the ‘interpolation-extrapolation’ method.

2.4. Inherent Assumptions of OLS

In the derivation of the ordinary least squares objective function, we assume errors to be present only in the dependent variable, i.e. pressure, and assumed the time variable and the flow rate history to be free of errors. Obviously, this assumption is not valid in cases where instrument or operation errors are introduced in these ‘independent’ variables.

Another assumption inherent to the least squares likelihood and objective function is the implication of Gaussian distribution of errors. This assumption makes the subsequent likelihood maximization mathematically convenient, but may not necessarily be valid especially if the pressure error distribution exhibits special behavior due to bias or instrument malfunction.

In the following chapters, we will discuss the effect of systematic or biased errors in multiple variables on parameter estimation. We will then discuss the development and testing of methods that are robust to such errors.
Chapter 3

3. Orthogonal Distance Regression

One of the limitations of ordinary least squares (OLS) is that it assumes errors to be present only in the dependent variable (pressure). However, there may be noise in time measurement or a lack of synchronization between the pressure and flow rate measurement. This chapter discusses methods that consider errors in both the pressure and time variable.

3.1. The General ODR Problem

Figure 3.1 shows an example where OLS is inadequate to match the model to data; this is because of a constant time shift between the pressure and flow rate history. Such errors are more generally referred to as breakpoint error. OLS assumes errors only in the pressure and hence is unable to account for breakpoint errors in the data. The same applies to random time measurement errors in data.

Figure 3.1: Example of breakpoint error resulting in inadequate match to pressure data using ordinary least squares (OLS).
The presence of errors in the independent variable (time) is accounted for in the general class of Errors in Variables (EIV) regression methods. Deming (1943) regression refers to the more specific problem with two variables with assumed independent errors; the pressure and time errors are independent and drawn from their respective error distributions. These errors and variables are represented as shown in Equations 3.1 and 3.2.

Because we assume these errors to have Gaussian distributions, the likelihood function for a given set of unknown model parameters is formulated as shown in Equation 3.3, and the equivalent log likelihood objective function to be minimized is Equation 3.4. This is the general orthogonal distance regression (ODR) objective function for pressure transient analysis problems.

\[ \epsilon_{pi} = p_{di} - p(t_i^*, q, \bar{\alpha}) \]
\[ \epsilon_{ti} = t_{di} - t_i^* \]

\[ L(\bar{\alpha} | \bar{p_d}, \bar{t_d}) = \prod_{i=1}^{n} \frac{1}{\sigma_p \sqrt{2\pi}} \exp \left( -\frac{(p_{di} - p(t_i^*, q, \bar{\alpha}))^2}{2 \sigma_p^2} \right) \]
\[ \times \frac{1}{\sigma_t \sqrt{2\pi}} \exp \left( -\frac{(t_{di} - t_i^*)^2}{2 \sigma_t^2} \right) \]

\[ \min \alpha \ E(\alpha) = \sum_{i=1}^{n} \left( p_{di} - p(t_i^*, q, \bar{\alpha}) \right)^2 + \frac{\sigma_p^2}{\sigma_t^2} \left( t_{di} - t_i^* \right)^2 \]

One key distinction between the ODR and OLS objective function is the need to specify the variance in pressure and time errors, or at least their ratio. This introduces a complication since the ratio of pressure and time variance is usually unknown in advance. Assuming the ratio to be unity essentially makes the objective function a sum of the Euclidian distance between the observed data point and the model. The minimum Euclidean distance is the length of the orthogonal line connecting the data observation to its corresponding point on the model fit. This is why this method is often referred to as ‘orthogonal distance’ regression.
In this method, we consider the time errors to be independent. In the case of systematic error due to mismatch in breakpoints in the pressure and time data, this assumption does not hold because the individual time errors are correlated. This is discussed in later sections.

3.2. Dastan’s Orthogonal Distance Regression

Dastan (2010) introduced the first implementation of ODR for application to pressure transient analysis. In that study, combinations of linear and logarithmic difference in pressure and time error were investigated. The conclusion was that using the linear distance performed best for problems with errors in both variables. Dastan defined the ratio of variance as the ratio of maximum change in pressure and time, essentially normalizing the data (Equation 3.5).

$$\beta = \frac{\max \Delta p}{\max \Delta t}$$

(3.5)

$$\min_{\alpha} E = \sum_{i=1}^{n} (p_{di} - p(t_i^*(\bar{\alpha}, \bar{q}, \bar{\alpha}))^2 + \beta^2 (t_{di} - t_i^*(\bar{\alpha}))^2$$

(3.6)

Dastan’s implementation of ODR treats the true time variable ($t^*$) as a function of the model unknowns ($\alpha$) at the current iteration. Therefore, before every iteration for updating the model unknowns, the orthogonal time point ($t^*$) must be found by minimizing the objective function with respect to $t^*$ (Equation 3.7), which is equivalent to solving the nonlinear Equation 3.8 iteratively. In practice, the number of iterations required to find within $10^{-8}$ tolerance range from one to five.

$$t_i^*(\alpha) \rightarrow \min_{t_i^*(\alpha)} E = (p_{di} - p(t_i^*(\bar{\alpha}, \bar{q}, \bar{\alpha}))^2 + \beta^2 (t_{di} - t_i^*(\bar{\alpha}))^2$$

(3.7)

$$\frac{\partial E}{\partial t_i^*(\alpha)} = (p_{di} - p(t_i^*(\bar{\alpha}, \bar{q}, \bar{\alpha})) \frac{\partial p(t_i^*, \bar{q}, \bar{\alpha})}{\partial t_i^*} + \beta^2 (t_{di} - t_i^*(\bar{\alpha})) = 0$$

(3.8)
Once these orthogonal points have been found, they can be used to calculate the model parameter update. The expressions for the ODR gradient and the Hessian are indicated in Equation 3.9-3.10. The second-order derivatives are neglected according to the Gauss-Newton approximation.

\[
g_j = \frac{\partial E}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{2\beta^2}{\beta^2 + \left(\frac{\partial p}{\partial t_i^*}\right)^2} \left[ (t_{di} - t_{i^*}^*(\alpha)) \left. \frac{\partial p}{\partial t_i^*} \right|_{\alpha_j} - (p_{di} - p(t_{i^*}^*(\alpha), \bar{q}, \bar{\alpha})) \right] \left. \frac{\partial p}{\partial \alpha_j} \right|_{t_i^*}
\]

(3.9)

\[
H_{jk} = \frac{\partial^2 E}{\partial \alpha_j \partial \alpha_k} = \sum_{i=1}^{n} \frac{2\beta^2}{\beta^2 + \left(\frac{\partial p}{\partial t_i^*}\right)^2} \left. \frac{\partial p}{\partial \alpha_j} \right|_{t_i^*} \left. \frac{\partial p}{\partial \alpha_k} \right|_{t_i^*}
\]

(3.10)

Dastan’s ODR is essentially a two-step process; the \( t^* \) estimation step effectively minimizes the objective function (sum of orthogonal distances) by adjusting the model \( t^*(\alpha) \) for a given set of model parameters. The model parameters \( (\alpha) \) are then updated keeping the \( t^* \) constant. This process is iterated until the convergence criteria are met. Any further references to orthogonal distance regression (ODR) in this report refer to Dastan’s implementation.

### 3.3. Boggs Regression

Boggs et al. (1987) introduced a method for solution of the general ODR problem and compared its computational efficiency to OLS. Itthisawatpan (2012) applied the Boggs method to parameter estimation in pressure transient analysis and compared it to ODR in terms of accuracy and computational efficiency. Itthisawatpan concluded that both ODR and Boggs regression are more accurate and robust than OLS in cases with breakpoint error in the data.
Unlike ODR, the model time variable \((t^B)^1\) is treated independently of the choice of the model unknowns \((\bar{a})\). These model time variables are estimated simultaneously with the model parameters at every iteration. This eliminates the iterative estimation of \(t^*\) that is required for ODR. Itthisawatpan (2012) indicated that this is computationally efficient with minimal loss in accuracy or robustness.

The theory and derivation of the objective function follows that of the general ODR discussed in Section 3.1. The objective function is indicated in Equation 3.11. In Boggs’ original paper, a time shift is used as the model parameter instead of model time \((t^B)\), but for the sake of consistency with the general ODR formulation, we shall treat the model time variable \((t^B)\) explicitly. Mathematically and algorithmically, this changes nothing. The gradient and the Hessian (according to the Gauss-Newton approximation) are indicated in Equations 3.12-3.17.

\[
\min_{\alpha} E = \sum_{i=1}^{n} (p_{dl} - p(t_{Bi}, \bar{q}, \bar{a}))^2 + \frac{\sigma_p^2}{\sigma_t^2} (t_{di} - t_{Bi})^2
\]

(3.11)

\[
g_{\alpha j} = \frac{\partial E}{\partial \alpha_j} = \sum_{i=1}^{n} -2 (p_{dl} - p(t_{Bi}, \bar{q}, \bar{a})) \left. \frac{\partial p}{\partial \alpha_j} \right|_{t_{Bi}}
\]

(3.12)

\[
g_{t_{Bj}} = \frac{\partial E}{\partial t_{Bj}} = -2 \left[ (p_{dj} - p(t_{Bj}, \bar{q}, \bar{a})) \left. \frac{\partial p}{\partial t_{Bj}} \right|_{\bar{a}} + \frac{\sigma_p^2}{\sigma_t^2} (t_{dj} - t_{Bj}) \right]
\]

(3.13)

\[
H = \begin{bmatrix}
\frac{\partial^2 E}{\partial \alpha_j \partial \alpha_k} & \frac{\partial^2 E}{\partial \alpha_j \partial t_{B_k}} \\
\frac{\partial^2 E}{\partial t_{B_j} \partial \alpha_k} & \frac{\partial^2 E}{\partial t_{B_j} \partial t_{B_k}}
\end{bmatrix} = \begin{bmatrix}
H_{\alpha \alpha} & H_{\alpha t_{B}} \\
H_{t_{B} \alpha}^T & H_{t_{B} t_{B}}
\end{bmatrix}
\]

(3.14)

---

1 For Boggs formulation, the model time variable is referred to as \(t^B\) to avoid ambiguity with \(t^*\) from ODR.
\[ H_{\alpha_j\alpha_k} = \sum_{i=1}^{n} 2 \left[ \frac{\partial p}{\partial \alpha_j} \bigg|_{t_{Bi}} \frac{\partial p}{\partial \alpha_k} \bigg|_{t_{Bi}} \right] \]

\begin{align*}
H_{\alpha_jt_{Bk}} &= 2 \left[ \frac{\partial p}{\partial \alpha_j} \bigg|_{t_{Bk}} \frac{\partial p}{\partial t_{Bk}} \bigg|_{\bar{\alpha}} \right] \\
H_{t_{Bj}t_{Bk}} &= \frac{\partial^2 E}{\partial t_{Bj} \partial t_{Bk}} = \begin{cases} 
2 \left( \frac{\partial p}{\partial t_{Bj}} \bigg|_{\bar{\alpha}} \right)^2 + 2 \frac{\sigma_p^2}{\sigma_t^2} & \text{if } j = k \\
0 & \text{if } j \neq k
\end{cases}
\end{align*}

(3.15) \hspace{1cm} (3.16) \hspace{1cm} (3.17)

Ithisawatpan (2012) suggested resetting \( t^B \) to the ODR \( t^* \) every five iterations to ensure similar results to ODR because his primary objective was to improve the computational efficiency of ODR. In this study, we did not do so, as our objective was to study the native accuracy and robustness of the Boggs method.

### 3.3.1. Efficient Solution of the Boggs System

Because Boggs method treats each individual data point as a model parameter, the system of coupled equations to be solved each iteration can be very large. For large data sets, such as those measured with a permanent downhole gauge (PDG), the computational requirements of Boggs method could become impractical. Fortunately, given the independence of time errors, a portion of the Hessian matrix is diagonal (Equation 3.17). The sparsity pattern of the Boggs Hessian matrix for a problem with ten data points and three model parameters is indicated in Figure 3.2. Usually, there are many more data points than model parameters and therefore a large portion of the Hessian is diagonal. Solving such a system with large systematic sparsity creates an opportunity for a specialized solution method.
Figure 3.2: Schematic of the sparsity pattern of the Boggs Hessian matrix for a problem with three model parameters and ten data points.

The system of equations for computing a Boggs update is indicated in Equation 3.17 in block matrix form. Performing a Schur complement on this system of equations, we obtain the decoupled equations for model parameter and Boggs time ($t^B$) update indicated in Equations 3.18-19. The decoupled equation for the $t^B$ update is diagonal and hence can be calculated in $O(n)$ time versus $O(n^3)$ time for the model parameter update.

\[
\begin{bmatrix}
  H_{\alpha\alpha} & H_{\alpha t_B} \\
  H_{t_B t_B}^T & H_{t_B t_B}
\end{bmatrix}
\begin{bmatrix}
  \Delta \alpha \\
  \Delta t_B
\end{bmatrix}
= -
\begin{bmatrix}
  g_{t_B}
\end{bmatrix}
\]

\[
\begin{bmatrix}
  H_{\alpha\alpha} - H_{\alpha t_B} H_{t_B t_B}^{-1} H_{t_B t_B}^T
\end{bmatrix}
\Delta \alpha
= -g_{\alpha} - H_{\alpha t_B} H_{t_B t_B}^{-1} g_{t_B}
\]

\[
H_{t_B t_B} \Delta t_B
= -g_{t_B} - H_{t_B t_B}^T \Delta \alpha
\]

A different derivation of this algorithm was introduced in the original Boggs et al. (1987) paper and compared to OLS.
3.4. Iterative Ratio of Variance Estimation

In the previous sections, we discussed the effect of assuming the ratio of variance to be either unity or assigned the ratio of variance to be equal to the ratio of scale of variation of the signal. Equating the ratio of variance to the ratio of scale of signal variation is equivalent to the assumption that the signal to noise ratio is the same for both pressure and time measurements.

In practice however, the magnitude of pressure and time errors are caused by the instrument errors, while the magnitude of pressure change is controlled by the formation properties. The magnitude of the time signal is simply the test duration. Also, there is significantly more pressure measurement error than arbitrary time measurement errors in a typical pressure transient data set. So the signal scaling estimate for ratio of variance is not valid.

![Figure 3.3: Error in ODR parameter estimate (for Dataset 1) for assumed values of the ratio of variance of errors.](image)

The true ratio of variance is seldom known a-priori, the general subject literature suggests estimating this from secondary variables or assuming from prior. Figure 3.3 shows the effect of assumed ratio of variance on the ODR estimate of parameters for Dataset 1 with synthetic noise of equal variance in psi and hours. The effect of the assumed ratio of variance on the final solution is found to be significant.

However, if the model parameters are estimated accurately, we can recover the pressure and time errors from the residuals of the model match to the data. Figure 3.4 shows an
example. The pressure and time residuals resemble the histogram of the original errors added to the data and the ratio of variance calculated from these residuals is likely to be accurate.

![Figure 3.4: Example of ODR fit with residuals reflecting the true statistics for noise in pressure and time. (Dataset 8 with synthetic noise)](image)

When we converge to a solution iteratively, the ratio of variance estimated from the residuals also approaches the true ratio of variance. So, once the mean of the residuals begins to approach zero (assumption of unbiased Gaussian error), we can estimate the ratio of variance of the residuals for calculation of the objective function for subsequent iterations. Using this strategy, we are effectively estimating the ratio of variance along with the model parameters as we iterate towards a solution. We implemented this technique for both ODR and Boggs methods and observed that the ratio of variance is most often estimated correctly within an order of magnitude of its true value.
3.5. Performance Comparison

In this section, we describe the performance of OLS, ODR and Boggs methods for parameter estimation of synthetic data sets from the test matrix (described in Appendix 1). We tested these methods on synthetic data sets so that their accuracy could be measured in absolute terms as the true model parameter values were known. The performance of these methods was compared in terms of accuracy and convergence. To test accuracy, we tested these methods with varying levels of noise added to the data sets. Convergence was tested by varying the starting guess within a range of parameter values and comparing the accuracy of the estimated parameters. The discussion on computational performance is saved for Chapter 6, which is dedicated to the issues of computational efficiency and scaling reasonably to problem size.

3.5.1. Random Time Errors

For data with random pressure and time errors, 100 random realizations of noise were generated for each data set at each noise level. The level of noise was increased to test the robustness of the methods to increasing noise. The accuracy is quantified in terms of mean and standard deviation of the estimation error. The ideal performance is zero mean and standard deviation for all noise levels. First, we tested with noise only in pressure, then with only time noise and later with combined pressure and time noise. The results of this test on Dataset 1 are shown in Figure 3.5. The maximum pressure noise has standard deviation of 5% of the maximum pressure change in the dataset. The maximum time noise has standard deviation of one hour. For combined pressure and time noise cases, the amount of noise in both pressure and time was increased proportionally to these maximums.

Tables 3.1-3.3 indicate the mean and standard deviation of the estimated parameters calculated from 30 random noise realizations for each case and each dataset at the maximum noise level mentioned earlier. The results indicate that OLS is the most inaccurate estimator for cases with any time errors. ODR is the most accurate and invariant estimator. Because this is a test of accuracy, the true solution was provided as a starting guess, robustness to the starting guess was tested next.
Figure 3.5: Results of 100 Monte Carlo runs using OLS, ODR and Boggs on Dataset 1 with variations of pressure and time noise.
Table 3.1: Results of Monte Carlo estimation of parameters for Datasets 1-8 with only pressure noise up to 5% standard deviation of the maximum pressure change.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>ODR</td>
</tr>
<tr>
<td>1</td>
<td>k</td>
<td>0.27%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.20%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.52%</td>
</tr>
<tr>
<td>2</td>
<td>k</td>
<td>0.75%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>3.42%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.57%</td>
</tr>
<tr>
<td>3</td>
<td>k</td>
<td>4.48%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>13.88%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3.39%</td>
</tr>
<tr>
<td>4</td>
<td>k</td>
<td>0.04%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.06%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.62%</td>
</tr>
<tr>
<td>5</td>
<td>k</td>
<td>8.79%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>32.62%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.72%</td>
</tr>
<tr>
<td>6</td>
<td>k</td>
<td>1.09%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>3.66%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.51%</td>
</tr>
<tr>
<td>7</td>
<td>k</td>
<td>10.87%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>37.91%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.50%</td>
</tr>
<tr>
<td>8</td>
<td>k</td>
<td>1.10%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>2.23%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.50%</td>
</tr>
</tbody>
</table>
Table 3.2: Results of Monte Carlo estimation of parameters of Datasets 1-8 with only time error with 1 hour standard deviation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>OLS Mean</th>
<th>OLS Standard Deviation</th>
<th>ODR Mean</th>
<th>ODR Standard Deviation</th>
<th>Boggs Mean</th>
<th>Boggs Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>k</td>
<td>0.43%</td>
<td></td>
<td></td>
<td>10.09%</td>
<td>0.91%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.79%</td>
<td></td>
<td></td>
<td>35.25%</td>
<td>3.11%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>8.02%</td>
<td></td>
<td></td>
<td>28.67%</td>
<td>6.48%</td>
</tr>
<tr>
<td>2</td>
<td>k</td>
<td>0.33%</td>
<td></td>
<td></td>
<td>13.62%</td>
<td>0.38%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.24%</td>
<td></td>
<td></td>
<td>47.54%</td>
<td>1.28%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.84%</td>
<td></td>
<td></td>
<td>32.16%</td>
<td>7.09%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.04%</td>
<td></td>
<td></td>
<td>0.84%</td>
<td>0.08%</td>
</tr>
<tr>
<td>3</td>
<td>k</td>
<td>0.19%</td>
<td></td>
<td></td>
<td>14.53%</td>
<td>0.18%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.60%</td>
<td></td>
<td></td>
<td>49.68%</td>
<td>0.54%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>7.41%</td>
<td></td>
<td></td>
<td>37.31%</td>
<td>6.47%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.26%</td>
<td></td>
<td></td>
<td>0.80%</td>
<td>0.05%</td>
</tr>
<tr>
<td>4</td>
<td>k</td>
<td>7.66%</td>
<td></td>
<td></td>
<td>17.64%</td>
<td>0.13%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>28.26%</td>
<td></td>
<td></td>
<td>63.18%</td>
<td>0.26%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>12.65%</td>
<td></td>
<td></td>
<td>36.91%</td>
<td>1.59%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>23048.32%</td>
<td></td>
<td></td>
<td>71720.70%</td>
<td>679.14%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>143015924.19%</td>
<td></td>
<td></td>
<td>39551849.91%</td>
<td>9.39%</td>
</tr>
<tr>
<td>5</td>
<td>k</td>
<td>13.94%</td>
<td></td>
<td></td>
<td>27.19%</td>
<td>0.11%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>50.08%</td>
<td></td>
<td></td>
<td>98.81%</td>
<td>0.77%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.24%</td>
<td></td>
<td></td>
<td>32.00%</td>
<td>6.68%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.30%</td>
<td></td>
<td></td>
<td>0.50%</td>
<td>0.07%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>18830.05%</td>
<td></td>
<td></td>
<td>47261.53%</td>
<td>2286.46%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>4422.51%</td>
<td></td>
<td></td>
<td>11661.50%</td>
<td>33.11%</td>
</tr>
<tr>
<td>6</td>
<td>k</td>
<td>12.76%</td>
<td></td>
<td></td>
<td>23.56%</td>
<td>1.94%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>46.08%</td>
<td></td>
<td></td>
<td>84.60%</td>
<td>7.04%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.01%</td>
<td></td>
<td></td>
<td>35.16%</td>
<td>4.21%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.27%</td>
<td></td>
<td></td>
<td>0.57%</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>8487.99%</td>
<td></td>
<td></td>
<td>34371.72%</td>
<td>1112.07%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>1331.97%</td>
<td></td>
<td></td>
<td>5792.47%</td>
<td>16.29%</td>
</tr>
<tr>
<td>7</td>
<td>k</td>
<td>1.71%</td>
<td></td>
<td></td>
<td>16.83%</td>
<td>8.25%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>5.69%</td>
<td></td>
<td></td>
<td>58.43%</td>
<td>28.38%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.99%</td>
<td></td>
<td></td>
<td>27.88%</td>
<td>8.38%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>86.67%</td>
<td></td>
<td></td>
<td>29.61%</td>
<td>1.78%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>11.61%</td>
<td></td>
<td></td>
<td>41.74%</td>
<td>11.91%</td>
</tr>
<tr>
<td>8</td>
<td>k</td>
<td>24.71%</td>
<td></td>
<td></td>
<td>11.30%</td>
<td>0.85%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>86.67%</td>
<td></td>
<td></td>
<td>29.61%</td>
<td>1.78%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>11.61%</td>
<td></td>
<td></td>
<td>41.74%</td>
<td>11.91%</td>
</tr>
</tbody>
</table>
Table 3.3: Results of Monte Carlo estimation of parameters of Datasets 1-8 with pressure and time error.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pressure &amp; Time Noise</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>ODR</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>k</td>
<td>6.82%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>20.96%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>18.29%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dataset 2</td>
<td>k</td>
<td>14.34%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>50.08%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.24%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.05%</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>k</td>
<td>5.05%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>15.94%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.79%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.54%</td>
</tr>
<tr>
<td>Dataset 4</td>
<td>k</td>
<td>12.64%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>48.10%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.44%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>18686.60%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>6277.89%</td>
</tr>
<tr>
<td>Dataset 5</td>
<td>k</td>
<td>1.97%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>9.33%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.47%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.45%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>29433.39%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>33750633.51%</td>
</tr>
<tr>
<td>Dataset 6</td>
<td>k</td>
<td>7.78%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>28.16%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.29%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.18%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>33622.83%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>83784507.96%</td>
</tr>
<tr>
<td>Dataset 7</td>
<td>k</td>
<td>5.98%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>20.92%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>11.33%</td>
</tr>
<tr>
<td>Dataset 8</td>
<td>k</td>
<td>19.98%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>73.11%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>33.85%</td>
</tr>
</tbody>
</table>
Accurate estimation for a wide range of starting guesses is a test of convergence for the estimation method. Figure 3.6 maps the percentage error in permeability estimate for Dataset 1 for a range of input permeability and skin as starting guess. No noise was added to the data for this test. The figure reveals that OLS converges to the true solution for a wider range of starting guesses (blue is a close estimate, red is far from the true estimate). The true solution is at the center of the map (log k=1, log C=-2). ODR is the least robust to starting guesses further from the true solution.

![Figure 3.6](image)

Figure 3.6: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, ODR and Boggs methods, for Dataset 1 with no noise.

![Figure 3.7](image)

Figure 3.7: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, ODR and Boggs methods, for Dataset 1 with maximum pressure and time noise.

Figure 3.7 indicates the result of this test for maximum pressure and time error (standard deviation of 5% max pressure change and 1 hour) in Dataset 1. As we discovered earlier, OLS is least robust to errors in time. This is evident from Figure 3.7 where OLS converges
to an inaccurate solution for all starting guesses, but does so consistently for a wide range of starting guesses. On the other hand, ODR and Boggs can estimate the parameters accurately in spite of time error in the data, as long as the starting guess is close to the true solution. For more distant starting guesses, ODR and Boggs are unable to converge to an accurate solution.

### 3.5.2. Breakpoint Error

In this section, we look at the effect of breakpoint error on parameter estimation using the methods discussed in this chapter. Figure 3.8 indicates the result of increasing breakpoint error in Dataset 1 on the estimated permeability using OLS, ODR and Boggs methods. ODR and Boggs seem to be much less sensitive to breakpoint error than OLS. However, beyond three hours of breakpoint error, the ODR and Boggs estimates are affected significantly. This is an example with a single breakpoint, for examples with many breakpoints, the threshold at which Boggs and ODR estimates to deteriorate may be earlier.

![Figure 3.8: Effect of adding breakpoint error to Dataset 1 on the permeability estimate using OLS, ODR and Boggs methods.](image)

Figure 3.9 indicates the map of error in permeability estimate for a range of starting guesses for Dataset 1 with one hour of breakpoint error. The true solution is at the center of the map.
(log k=1, log C=-2). Similar to the case with random time errors, ODR and Boggs can estimate permeability accurately for starting guesses close to the true solution, but do not converge successfully for guesses farther away. OLS is unable to converge to a suitable accurate solution for any starting guess, far or close to the true solution.

Figure 3.9: Maps of percentage permeability estimation error for a range of starting guesses using OLS, ODR and Boggs methods, for Dataset 1 with three hours of breakpoint noise.

In conclusion, OLS is robust and convergent for a wide range of starting guesses, however in cases with time noise or breakpoint error it fails to estimate an accurate solution. ODR and Boggs regression are alternate approaches that account for the presence of errors in the time variable. This makes them more accurate even in the presence of time noise or breakpoint error. However, these methods are not robust for starting guesses further away from the true solution.
Chapter 4

4. Breakpoint Regression

In the previous chapter, we discussed formulations that consider errors in both pressure and time variables. ODR and Boggs method were found to be more accurate than OLS for problems with time errors, but had weaker convergence than OLS for a wider range of starting guesses. In cases with breakpoint error, ODR and Boggs were found to be more accurate than OLS, but only up to a certain amount of breakpoint error. ODR and Boggs methods are based on a formulation that considers the time errors to be independent (uncorrelated), which is valid for random time measurement errors. In the case of breakpoint errors, however, individual time errors are correlated because all the time measurements are shifted by a constant amount of error.

4.1. Breakpoint Regression

In this section, we describe a formulation which, instead of treating individual time errors independently, treats the shift of the breakpoint itself as an unknown parameter alongside other unknown model parameters. The currently known breakpoint is used as a starting guess for this unknown parameter. This method is referred to as Breakpoint regression (BPR). Because the multirate pressure is a strong function of the breakpoint, if there is no breakpoint error, the breakpoint need not be adjusted to reduce pressure error. If a component of the pressure error is caused by breakpoint error, then the error is reduced by directly adjusting breakpoints.

To treat the breakpoint as an unknown model parameter alongside reservoir properties such as permeability and skin, the corresponding derivative terms in the gradient and Hessian need to be calculated. The superposition equation (Equation 4.1) indicates the multirate
pressure function as a function of the step changes in flow rate and breakpoint variables. \( \tau_i \) denotes the \( i^{th} \) breakpoint and \( q_{\tau_i} \) is the flow rate at times before the breakpoint. The derivative of the superposition pressure with respect to a certain breakpoint \( \tau_i \) is expressed in Equation 4.2.

\[
p(t, q, \bar{\alpha}) = P_i - \sum_{i=0}^{n} (q_{\tau_{i+1}} - q_{\tau_i}) P_u(t - \tau_i, \bar{\alpha})
\]  

(4.1)

\[
\frac{\partial p(t, q, \bar{\alpha})}{\partial \tau_i} = (q_{\tau_{i+1}} - q_{\tau_i}) \frac{\partial P_u(t - \tau_i, \bar{\alpha})}{\partial \tau_i} = -(q_{\tau_{i+1}} - q_{\tau_i}) \frac{\partial P_u}{\partial t} \bigg|_{t=\tau_i}
\]  

(4.2)

The Jacobian matrix is defined as shown in Equation 4.3. The Gauss-Newton Hessian and gradient for OLS can be written in terms of the Jacobian matrix as shown in Equations 4.4-4.5.

\[
J_{ij} = \frac{\partial p(t_i, q, \bar{\alpha})}{\partial \alpha_j}
\]  

(4.3)

\[
H = J^T J
\]  

(4.4)

\[
g = -J^T E; \quad E = (p_{di} - p(t_i, q, \bar{\alpha}))
\]  

(4.5)

If we are to treat breakpoints (\( \tau_i \)) as unknowns, the corresponding derivative terms are added to the Jacobian as indicated in Equation 4.6.

\[
J = \begin{bmatrix}
\frac{\partial p(t_i, q, \bar{\alpha})}{\partial \alpha_j} & \frac{\partial p(t_i, q, \bar{\alpha})}{\partial \tau_j}
\end{bmatrix}
\]  

(4.6)
This modified Jacobian is used in Equations 4.4-4.5 to form the Gauss-Newton gradient and Hessian for BPR. Because breakpoints are treated as unknowns, they must be updated alongside the model parameters after every iteration.

4.2. Flow Rate Regression (FRR)

Apart from breakpoint error, there is another kind of error that may affect measurements that are part of the flow rate history. There may be noise in the measurement of the flow rate measurement itself due to inaccurate measurement at surface or improper closure of valves.

Figure 4.1 shows a synthetic example of a drawdown and buildup sequence where the actual flow rate during the buildup is nonzero due to incomplete closing of valves. Parameter estimation using OLS in such cases causes large errors as evident in the example.

Figure 4.1: Effect of flow rate measurement error on OLS match to the data.
In a manner similar to BPR, Flow Rate Regression (FRR) treats the flow rates as unknowns. The expression for derivative of pressure with respect to flow rate magnitude is derived from the superposition equation (Equation 4.1) and is indicated in Equation 4.7. The corresponding derivative terms are added to the Jacobian as shown in Equation 4.8.

\[
\frac{\partial p(t, q, \bar{\alpha})}{\partial q_i} = P_u(t - \tau_i, \bar{\alpha}) - P_u(t - \tau_{i-1}, \bar{\alpha})
\]  

(4.7)

\[
J = \begin{bmatrix}
\frac{\partial p(t_i, q, \bar{\alpha})}{\partial \alpha_j} & \frac{\partial p(t, q, \bar{\alpha})}{\partial q_j}
\end{bmatrix}
\]

(4.8)

Examination of the superposition equation reveals that the multi rate pressure function is linear with respect to flow rate magnitudes. The unit rate pressure function itself is strongly linear with respect to permeability, due to the proportionality of permeability and dimensionless pressure. Therefore, matching permeability and flow rate simultaneously is an over-determined problem. To constrain the flow rates to the correct scale we need to fix at least one of the nonzero flow rate measurements, effectively estimating the ratio of the flow rates with the chosen fixed rate rather than the flow rate itself. In our implementation we fixed the first flow rate to be exact and did not add noise to it during testing with errors.

### 4.3. Hybrid ODR

In the previous chapter, we described the ODR method, used to accommodate random time measurement errors. In this section, we will describe modifications to ODR that allow hybridization with BPR, FRR or both. This hybrid method will be able to account for errors in random errors in time measurement along with errors in breakpoint and/or flow rate measurement.

The ODR gradient and Hessian (derived in the previous chapter) are repeated in Equations 4.9 and 4.10. They can be expressed in terms of the Jacobian as derived in Equations 4.11-14.
\[ g_j = \sum_{i=1}^{n} \frac{2\beta^2}{\beta^2 + \left( \frac{\partial p}{\partial t_i^*} \right)^2} \left[ (t_{di} - t_i^*(\bar{\alpha})) \left. \frac{\partial p}{\partial t_i^*} \right|_{\alpha_j} - (p_{di} - p(t_i^*(\bar{\alpha}), \bar{q}, \bar{\alpha})) \right] \left. \frac{\partial p}{\partial \alpha_j} \right|_{t_i^*} \]  

(4.9)

\[ H_{jk} = \frac{\partial^2 E}{\partial \alpha_j \partial \alpha_j} = \sum_{i=1}^{n} \frac{2\beta^2}{\beta^2 + \left( \frac{\partial p}{\partial t_i^*} \right)^2} \left[ \left. \frac{\partial p}{\partial \alpha_j} \right|_{t_i^*} \left. \frac{\partial p}{\partial \alpha_k} \right|_{t_i^*} \right] \]  

(4.10)

\[ B_{ii} = \frac{2\beta^2}{\beta^2 + \left( \frac{\partial p}{\partial t_i^*} \right)^2} \]  

(4.11)

\[ E_{i} = (t_{di} - t_i^*(\bar{\alpha})) \left. \frac{\partial p}{\partial t_i^*} \right|_{\alpha} - (p_{di} - p(t_i^*(\bar{\alpha}), \bar{q}, \bar{\alpha})) \]  

(4.12)

\[ g = j^T B E \]  

(4.13)

\[ H = j^T B j \]  

(4.14)

The objective function for the hybrid ODR is the same as the original ODR objective function (Equation 3.6). The hybrid ODR algorithm treats the flow rate variables (breakpoints, flow rates) as unknowns alongside the model parameters. The current flow rate variable measurements are used as a starting guess. The Jacobian is modified by adding the necessary derivative terms for the additional variables as derived for BPR and FRR. This updated Jacobian is then used to calculate the gradient and Gauss-Newton Hessian and calculate the parameter updates. The flow rate variables are updated alongside the model parameters after every iteration.
4.4. Performance Comparison

In this section, we compare the performance of BPR, FRR and their hybrid ODR versions for parameter estimation. As in the previous chapter, the performance of these methods was compared in terms of accuracy and convergence. To test accuracy, we tested these methods with varying levels of noise added to the data sets. Convergence was tested by varying the starting guess within a range of parameter values and comparing the accuracy of the estimated parameters.

4.4.1. Breakpoint Error

In Figure 4.2, we repeat the test from the previous chapter where the accuracy of OLS, ODR and now BPR are compared for increasing breakpoint error. As observed before in Figure 3.8, OLS is the most sensitive to breakpoint error, ODR less sensitive because it accounts for errors in time but for large amounts of breakpoint error it begins to make estimation errors. In our test, BPR was the least sensitive to breakpoint error, it estimated the parameters accurately for all the breakpoint errors tested.

Figure 4.2: Effect of increasing breakpoint error (Dataset 1) on the estimated permeability using OLS, ODR and BPR.
Figure 4.3 maps the error in permeability estimated using OLS, ODR and BPR for Dataset 1 with three hours of breakpoint error, for varying starting guesses of permeability and wellbore storage on a log scale. While OLS is robustly convergent for a wide range of starting guesses, it is unable to find an accurate solution because it cannot address breakpoint error. As noted earlier, ODR is accurate for a limited range of starting guesses close to the true solution (center of the map, k=10 md, C=0.01 STB/psi). BPR, however, combines accuracy and robustness to starting guesses and is able to estimate the permeability accurately for a wide range of starting guesses, even for problems with breakpoint error.

Table 4.2 indicates mean and standard deviation of the estimated parameters calculated from 30 random realizations of breakpoint error up to three hours per breakpoint, for each dataset in the test matrix. BPR was tested with and without additional pressure and time noise of standard deviation of 1 psi and 30 minutes respectively. The results indicate that in the absence of pressure and time noise, BPR estimates the all parameters accurately with a standard deviation of less than 1-2%, for Datasets 1-3 and 7. For Dataset 8, with higher permeability and wellbore storage, the deviation is higher but still reasonable. For the datasets with dual-porosity models (Dataset 4-6), there is much higher standard deviation in parameter values, especially the dual-porosity parameters (λ, ω). This is due to convergence failure in some cases. In such cases, the dual-porosity parameter estimates are incorrect by a few orders of magnitude, leading to meaningless mean and standard
deviation estimates. The median is a better estimator of centrality in those cases and is indicated in Table 4.1.

In the presence of pressure and time noise, there is much more variance and bias in the parameter estimates using BPR. These datasets with similar noise realizations yield better results with the hybrid ODR+BPR, due to its robustness to arbitrary time errors by design.

Table 4.1: Median error of parameters estimated during 30 Monte Carlo estimation runs for Datasets 4-6 from the Test Matrix, using BPR and ODR+BPR, with up to three hours of breakpoint error.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Parameter</th>
<th>BPR-No p and t noise</th>
<th>BPR</th>
<th>BPR+ODR</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>k</td>
<td>0.11%</td>
<td>0.55%</td>
<td>0.04%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.09%</td>
<td>0.17%</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
<td>6.01%</td>
<td>5.39%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>1.92%</td>
<td>100.00%</td>
<td>0.57%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.02%</td>
<td>43.96%</td>
<td>1.41%</td>
</tr>
<tr>
<td>5</td>
<td>k</td>
<td>6.39%</td>
<td>3.49%</td>
<td>0.02%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>6.76%</td>
<td>1.72%</td>
<td>0.06%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.76%</td>
<td>3.27%</td>
<td>10.75%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.04%</td>
<td>0.19%</td>
<td>0.03%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>0.05%</td>
<td>1.37%</td>
<td>224.96%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.10%</td>
<td>40.49%</td>
<td>6.96%</td>
</tr>
<tr>
<td>6</td>
<td>k</td>
<td>0.26%</td>
<td>1.47%</td>
<td>0.03%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.01%</td>
<td>1.16%</td>
<td>0.11%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
<td>6.32%</td>
<td>0.26%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.00%</td>
<td>0.31%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>54.70%</td>
<td>63.54%</td>
<td>26.21%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.00%</td>
<td>24.03%</td>
<td>5.24%</td>
</tr>
</tbody>
</table>
Table 4.2: Results of 30 Monte Carlo runs with up to 3 hours of breakpoint error, optionally with 1 psi pressure and 30 minutes of time noise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BPR-No p and t noise</td>
<td>BPR</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>k</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>k</td>
<td>0.06%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.01%</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>k</td>
<td>14.52%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>45.27%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.70%</td>
</tr>
<tr>
<td>Dataset 4</td>
<td>k</td>
<td>27.66%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>89.92%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.60%</td>
</tr>
<tr>
<td>Dataset 5</td>
<td>k</td>
<td>26.66%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>89.92%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.60%</td>
</tr>
<tr>
<td>Dataset 6</td>
<td>k</td>
<td>9.07%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>28.86%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3.02%</td>
</tr>
<tr>
<td>Dataset 7</td>
<td>k</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
</tr>
<tr>
<td>Dataset 8</td>
<td>k</td>
<td>0.96%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>2.72%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.43%</td>
</tr>
</tbody>
</table>
4.4.2. Flow Rate Error

Breakpoint regression (BPR) and flow rate regression (FRR) can be combined to treat both breakpoints and flow rate measurements as unknown variables to account for noise in these measurements. This is made possible by modifying the Jacobian to include the necessary derivative terms for both kinds of variables. This method is referred to as BP+FRR.

In this section, we show the test of the datasets for cases with errors in both the breakpoints and flow rate measurements. Figure 4.4 maps the error in permeability estimated using BPR, FRR and BP+FRR for Dataset 1 with three hours of breakpoint error and 10% error in flow rate measurement. The correct parameter values are at the center of the maps (log k=1, log C= -2). All three methods show similar robustness in convergence, as they converge to similar solutions for almost all starting guesses attempted.

This example indicates that BPR on its own is unable to correct for the presence of flow rate measurement error resulting in over 50% estimation error for permeability. The permeability estimated using FRR is more accurate, indicating that the flow rate error has a higher impact on the estimate than the breakpoint error in this example. However, BP+FRR, being able to correct both errors accurately estimates the permeability for most starting guesses attempted.

Figure 4.4: Statistics of starting guesses for OLS, ODR and BPR for breakpoint error of 1 hour and 10% flow rate error in Dataset 1
Table 4.4 lists the mean and standard deviation of estimated model parameters calculated from 30 random realizations of breakpoint error up to 3 hours and flow rate error of up to 10% for each dataset in the test matrix.

Similar to the previous test for breakpoint errors, FRR was tested with and without additional pressure and time noise of standard deviation of 1 psi and 30 minutes respectively. The results are similar and indicate that in the absence of pressure and time noise, FRR estimates all parameters reasonably accurately and consistently for non-dual-porosity models (Datasets 1-3, 7, 8).

For the datasets with dual-porosity models (Dataset 4-6), there is much higher standard deviation in parameter values, especially the dual-porosity parameters ($\lambda, \omega$), due to convergence failure in some cases, the median parameter estimation error is indicated in Table 4.3 for these datasets.

In the presence of pressure and time noise, there is much more variance and bias in the parameter estimates using BP+FRR. The same datasets with similar noise realizations yielded better results with the hybrid ODR+BP+FRR, due to its robustness to arbitrary time errors.
Table 4.3: Median error of parameters estimated during 30 Monte Carlo estimation runs for Datasets 4-6 from the Test Matrix, using BP+FRR and BP+FRR+ODR, with up to three hours of breakpoint error and 10% flow rate measurement error.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Median BP+FRR- No p and t noise</th>
<th>BP+FRR</th>
<th>BP+FRR+ODR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 4</td>
<td>k</td>
<td>0.07%</td>
<td>0.29%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.26%</td>
<td>1.62%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.02%</td>
<td>2.31%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>53.56%</td>
<td>64.19%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.08%</td>
<td>35.80%</td>
</tr>
<tr>
<td>Dataset 5</td>
<td>k</td>
<td>0.33%</td>
<td>0.38%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.00%</td>
<td>0.60%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.02%</td>
<td>8.29%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.00%</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>66.12%</td>
<td>100.00%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.00%</td>
<td>48.53%</td>
</tr>
<tr>
<td>Dataset 6</td>
<td>k</td>
<td>0.03%</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.00%</td>
<td>0.97%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
<td>2.89%</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.00%</td>
<td>0.12%</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>1.87%</td>
<td>667.67%</td>
</tr>
<tr>
<td></td>
<td>λ</td>
<td>0.03%</td>
<td>35.40%</td>
</tr>
</tbody>
</table>
Table 4.4: Results of Monte Carlo runs with up to 3 hours of breakpoint error and 10% flow rate measurement error, optionally with 1 psi pressure and 30 minutes of time noise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BP+FRR-No p and t noise</td>
<td>BP+FRR</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>k</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>k</td>
<td>0.09%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.30%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.02%</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>k</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.00%</td>
</tr>
<tr>
<td>Dataset 4</td>
<td>k</td>
<td>6.29%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>20.57%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.32%</td>
</tr>
<tr>
<td>Dataset 5</td>
<td>k</td>
<td>14.19%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>40.29%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.54%</td>
</tr>
<tr>
<td>Dataset 6</td>
<td>k</td>
<td>10.88%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>36.88%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.01%</td>
</tr>
<tr>
<td>Dataset 7</td>
<td>k</td>
<td>0.20%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.75%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.04%</td>
</tr>
<tr>
<td>Dataset 8</td>
<td>k</td>
<td>0.83%</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>2.35%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.42%</td>
</tr>
</tbody>
</table>
Chapter 5

5. General Likelihood Maximization

The ordinary least squares method (OLS) is based on the assumption of Gaussian probability distribution of pressure errors. In the previous sections, we investigated ODR and Boggs regression which address the possibility of errors in time as well as pressure, while still assuming a Gaussian distribution of these errors. In this chapter, we discuss the development of a parameter estimation method based on a generalized framework of likelihood maximization (GLM) for arbitrary distributions of error.

5.1. GLM for Errors in Pressure

Conventional estimation methods assume certain analytical distributions of error and use corresponding analytical likelihood functions to estimate the maximum likelihood parameters. The most commonly used probability distribution function is the Gaussian distribution function because the log-likelihood objective function is the sum of squares of the residuals (difference between model and data point). This objective function has derivatives linear to the model derivatives and is easy to compute. Robustness and mathematical convenience have made the ordinary least squares method (OLS) a very commonly used estimator.

Other available estimators are based on error distributions such as uniform, lognormal or exponential. The use of these estimators is contingent upon the assumption that the errors in the data follow the specific distribution. This is unlikely to be known in advance and the use of an inappropriate estimator will cause biased estimates. An example is indicated in Figure 5.1, OLS was used for parameter estimation on Dataset 1 with added lognormal
error. The actual model and the OLS fit to the data demonstrate the estimation error caused due to the use of an incorrect estimator.

Figure 5.1: Dataset 1 with lognormal noise, the actual model and the OLS match.

As mentioned before, the correct distribution of errors is unlikely to be known in advance. However, the true distribution of errors can be inferred accurately once the true parameters have been found; the residuals in the data represent the errors. As we converge iteratively towards a maximum likelihood parameter estimate, the residuals converge to the errors. Even with an incorrect estimator, these residuals can be used to infer the underlying distribution of the errors. For instance, lognormal pressure error was added to Dataset 1. Figure 5.1 indicates the OLS fit to this noisy dataset. The distribution of the residuals was compared and found to resemble the original lognormal distribution of the errors, as shown in Figure 5.2.
As we converge towards the solution, the residuals converge towards the errors in the data. The general likelihood maximization (GLM) method infers the distribution of errors iteratively from the residuals. Based on this distribution, a likelihood function is formulated and maximized using Gauss-Marquardt optimization.

Given a histogram of the residuals at any iteration of model parameters, we assume this to approximate the probability density function of the pressure errors in the data. Smoothing splines are fitted to the histogram data and are used to describe the continuous probability density function (pdf). Cubic piecewise polynomials are used to describe this pdf, denoted by the function $f_{p}$ in Equation 5.2. An example is shown in Figure 5.3.

\[
\epsilon_{pi} = p_{dt} - p(t_i, \bar{q}, \bar{\alpha})
\]

\[
f(\epsilon_{pi}) = f_{p}(p_{dt} - p(t_i, \bar{q}, \bar{\alpha}))
\]

The spline function must be extrapolated to describe extreme values of residuals that are not available in the data histogram, and have to be nonnegative over the domain of calculation. One may normalize the pdf to a scale of 0 to 1 to represent probabilities. However, this is not essential because the maximum likelihood point remains unaffected by scaling factors.
The shape of the probability distribution function (pdf) is obtained from the residuals of the model, however they need to be centered. The pdf may be set to have a mean or mode at a specified value if that information is available. The implementation developed in this research defaults to a mean of zero in the absence of a specified center to the pdf.

Assuming that the errors are independent results in the likelihood function shown in Equation 5.4, and the log likelihood objective function for the current iteration is formulated in Equation 5.5.

\[
L(\bar{\alpha}|\bar{p}_d) = \prod_{i=1}^{n} f_p(\epsilon_{pi})
\]  
\[\text{(5.3)}\]

\[
L(\bar{\alpha}|\bar{p}_d) = \prod_{i=1}^{n} f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))
\]  
\[\text{(5.4)}\]

\[
\max \log(L(\bar{\alpha}|\bar{p}_d)) = \min - \sum_{i=1}^{n} \log \left( f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha})) \right)
\]  
\[\text{(5.5)}\]
The derivations of the gradient and Gauss-Newton Hessian for this objective function are shown in Equations 5.6-14.

\[
g_j = \frac{\partial E}{\partial \alpha_j} = - \sum_{i=1}^{n} \frac{1}{f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))} \frac{\partial f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))}{\partial \alpha_j}
\]

\[
g_j = - \sum_{i=1}^{n} \frac{1}{f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))} \frac{\partial f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))}{\partial (p_{di} - p(t_i, \bar{q}, \bar{\alpha}))} \frac{\partial (p_{di} - p(t_i, \bar{q}, \bar{\alpha}))}{\partial \alpha_j}
\]  

Equation (5.6)

Equation (5.7)

The derivative of the spline \(f_p\) function is denoted by the shorthand \(f_p'\), computed at the current set of parameters \((\bar{\alpha})\).

\[
\frac{\partial f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))}{\partial (p_{di} - p(t_i, \bar{q}, \bar{\alpha}))} = f_p'(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))
\]  

Equation (5.8)

\[
g_j = \sum_{i=1}^{n} \frac{f_p'(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))}{f_p(p_{di} - p(t_i, \bar{q}, \bar{\alpha}))} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{f_p'}{f_p} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j}
\]  

Equation (5.9)

\[
H_{jk} = \frac{\partial^2 E}{\partial \alpha_j \partial \alpha_k} = \frac{\partial}{\partial \alpha_k} \left( \sum_{i=1}^{n} \frac{f_p'}{f_p} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j} \right)
\]  

Equation (5.10)

\[
H_{jk} = \sum_{i=1}^{n} \left( \frac{f_p'}{f_p} \frac{\partial^2 p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j \partial \alpha_k} \right) + \frac{\partial}{\partial \alpha_k} \left( \frac{f_p'}{f_p} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j} \right)
\]  

Equation (5.11)

\[
\frac{\partial}{\partial \alpha_k} \left( \frac{f_p'}{f_p} \right) = \frac{f_p \frac{\partial f_p'}{\partial \alpha_k} - \frac{\partial f_p}{\partial \alpha_k} \frac{f_p'}{f_p^2}}{f_p^2} = -\frac{f_p f_p'' + f_p' f_p'}{f_p^2} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_k}
\]  

Equation (5.12)

\[
H_{jk} = \sum_{i=1}^{n} \left( \frac{f_p'}{f_p} \frac{\partial^2 p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j \partial \alpha_k} \right) + \frac{f_p' f_p' - f_p f_p''}{f_p^2} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_k} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j}
\]  

Equation (5.13)

\[
H_{jk} \approx \sum_{i=1}^{n} \frac{f_p' f_p' - f_p f_p''}{f_p^2} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_k} \frac{\partial p(t_i, \bar{q}, \bar{\alpha})}{\partial \alpha_j}
\]  

Equation (5.14)
These expressions for gradient and Hessian are used iteratively to calculate parameter updates to minimize the objective function. For the next iteration, the residuals are recalculated using the updated parameters. These residuals are used to define a new error probability distribution function using splines, the value of the objective function is recalculated for the new distribution and minimized by calculating the parameter update.

This process is iterated until the convergence criteria are met. The convergence criterion for this method is the reduction in value of the objective function within any iteration (before and after parameter update). The value of this convergence criterion was set at $10^{-5}$ for the implementation of this method shown in this report.

### 5.2. GLM for Errors in Pressure and Time

In the previous section, we developed the GLM method for errors only in pressure measurements. Now, we extend the concept to a formulation that considers general error distributions for both pressure and time. The errors in the pressure and time measurement are expressed as indicated in Equations 5.15-5.16 and the respective spline probability distribution functions are defined as indicated in Equations 5.17-5.18.

\[
\epsilon_{pi} = p_{di} - p(t_i^*, \bar{q}, \bar{\alpha}) \tag{5.15}
\]

\[
\epsilon_{ti} = t_{di} - t_i^* \tag{5.16}
\]

\[
f(\epsilon_{pi}) = f_p(p_{di} - p(t_i^*, \bar{q}, \bar{\alpha})) \tag{5.17}
\]

\[
f(\epsilon_{ti}) = f_t(t_{di} - t_i^*) \tag{5.18}
\]

Assuming independent and identically distributed pressure and time errors, the likelihood function is derived as Equation 5.19.

\[
L(\alpha|p_d, t_d) = \prod_{i=1}^{n} f_p(p_{di} - p(t_i^*, \bar{q}, \bar{\alpha})) \times \prod_{i=1}^{n} f_t(t_{di} - t_i^*) \tag{5.19}
\]
Maximizing the log likelihood yields the minimization problem indicated in Equation 5.20. The gradient and Gauss-Newton Hessian for this problem are derived in Equations 5.21-5.29. The system of equations to be solved for computing the parameter updates is indicated in Equation 5.30.

\[
\min E = - \sum_{i=1}^{n} \left[ \log \left( f_p(p_{di} - p(t_i^*, \bar{q}, \bar{a})) \right) + \log \left( f_i(t_{di} - t_i^*) \right) \right]
\]  
(5.20)

\[
g_{\alpha j} = \frac{\partial E}{\partial \alpha_j} = - \sum_{i=1}^{n} \frac{1}{f_p(p_{di} - p(t_i^*, \bar{q}, \bar{a}))} \frac{\partial f_p(p_{di} - p(t_i^*, \bar{q}, \bar{a}))}{\partial \alpha_j}
\]  
(5.21)

\[
g_{\alpha j} = \sum_{i=1}^{n} \frac{f_p'(p_{di} - p(t_i^*, \bar{q}, \bar{a}))}{f_p(p_{di} - p(t_i^*, \bar{q}, \bar{a}))} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j}
\]  
(5.22)

\[
g_{t_j} = \frac{\partial E}{\partial t_j^*} = \frac{f_p'(p_{di} - p(t_i^*, \bar{q}, \bar{a}))}{f_p(p_{di} - p(t_i^*, \bar{q}, \bar{a}))} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial t_j^*} + \frac{f_i'(t_{di} - t_i^*)}{f_i(t_{di} - t_i^*)}
\]  
(5.23)

We use the notation \( f_p' \) and \( f_p \) to denote the values of the pressure pdf evaluated at the current model and time parameter values.

\[
H_{\alpha j \alpha k} = \sum_{i=1}^{n} \left( \frac{f_p'}{f_p} \frac{\partial^2 p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j \alpha_k} \right) + \frac{f_p f_p' - f_p f_p''}{f_p^2} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_k}
\]  
(5.24)

\[
H_{\alpha j t_k^*} = \left( \frac{f_p'}{f_p} \frac{\partial^2 p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j \partial t_k^*} \right) + \frac{f_p f_p' - f_p f_p''}{f_p^2} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j} \frac{\partial p(t_k^*, \bar{q}, \bar{a})}{\partial t_k^*}
\]  
(5.25)

\[
H_{t_k^* j t_k^*} = \left( \frac{f_p'}{f_p} \frac{\partial^2 p(t_k^*, \bar{q}, \bar{a})}{\partial t_k^*} \right) + \frac{f_p f_p' - f_p f_p''}{f_p^2} \left( \frac{\partial p(t_k^*, \bar{q}, \bar{a})}{\partial t_k^*} \right)^2 + \frac{f_i' f_i' - f_i f_i''}{f_i^2}
\]  
(5.26)

The Gauss-Newton approximation of the Hessian neglects second-order derivatives to ensure a positive-definite Hessian.

\[
H_{\alpha j \alpha k} \approx \sum_{i=1}^{n} \frac{f_p f_p' - f_p f_p''}{f_p^2} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_j} \frac{\partial p(t_i^*, \bar{q}, \bar{a})}{\partial \alpha_k}
\]  
(5.27)
\[ H_{a_j t^* k} \equiv \frac{f_p^l f_p^l - f_p f_p''}{f_p^2} \frac{\partial p(t^*_k, \bar{q}, \bar{\alpha})}{\partial \alpha_j} \frac{\partial p(t^*_k, \bar{q}, \bar{\alpha})}{\partial t^*_k} \]

(5.28)

\[ H_{t^* j t^*} \equiv \frac{f_p^l f_p^l - f_p f_p''}{f_p^2} \left( \frac{\partial p(t^*_k, \bar{q}, \bar{\alpha})}{\partial t^*_k} \right)^2 + \frac{f_t^l f_t^l - f_t f_t''}{f_t^2} \]

(5.29)

\[
\begin{bmatrix}
H_{a a} & H_{a t^*} \\
H_{a t^*}^T & H_{t^* t^*}
\end{bmatrix} \begin{bmatrix}
\Delta \alpha \\
\Delta t^*
\end{bmatrix} = - \begin{bmatrix}
ger \alpha \\
ger_{t^*}
\end{bmatrix}
\]

(5.30)

The final system of equations has a similar sparsity pattern to the example in Section 3.1.1, due to the independence of the time errors. A similar reduced system of equations can be derived from Equation 5.30 for improved computational efficiency.

The implementation of this formulation has met challenges. For \( n \) unknown model parameters to be estimated from \( z \) data points, this method estimates \( n + z \) parameters, which may lead to an over-determined problem. The current implementation of this method adjusts the objective function in the vicinity of the starting guess by adjusting the time errors and terminates prematurely. The program terminates the estimation procedure near the supplied starting guess.

Alternatively, an approach similar to Dastan’s treatment of the time errors as a function of the model parameters could be used. This is likely to constrain the time errors and emphasize adjustment of model parameters over time errors, which seems to be the problem in the current implementation.

### 5.3. Performance Comparison

In this section, we describe the performance of the general likelihood maximization (GLM) method for errors in pressure only. The performance of this methods is compared to that of OLS in terms of accuracy and convergence. To test accuracy, we generated synthetic examples with outliers and non-Gaussian error distributions and attempted to estimate the
model parameters. Convergence was tested by varying the starting guess within a range of parameter values and comparing the accuracy of the estimated parameters.

5.3.1. Examples

The first example, indicated in Figure 5.4, is that of lognormal pressure error in Dataset 1. The error histogram and the spline fit to the histogram are indicated in Figure 5.4. For lognormal error, all errors are positive, so the OLS estimate is biased because it tries to balance the positive and negative residuals according to the assumed symmetric Gaussian distribution. The GLM estimate, with the spline probability distribution function (pdf) is able to match the true data accurately by adaptively estimating not only the parameters but also the distribution of errors in the data.

![Error Histogram and Spline pdf](image1)

Figure 5.4: Lognormal noise in Dataset 1, the histogram of synthetic errors and the final spline pdf function is shown in the right. The noisy data, actual model, OLS and GLM match are compared on the left.

The second example, indicated in Figure 5.5, is that of Gaussian error and outliers centered around 5 psi added to Dataset 1. The error histogram and the spline fit to the histogram are indicated in Figure 5.5. The positive outliers cause the OLS fit to be biased, evident from the example. The GLM estimate, with the spline probability distribution function (pdf) centered at a zero mode (expecting outliers) is able to match the true data accurately.

![Error Histogram and Spline pdf](image2)
Figure 5.5: Noise and outliers centered around 5 psi in Dataset 1, the histogram of synthetic errors and the final spline pdf function is shown on the right. The noisy data, actual model, OLS and GLM match are compared on the left.

The next examples (Figure 5.6 and 5.7), are of pure outlier noise around 3 and -3 psi in an unequal multimodal distribution. There are more outliers around the positive mode than the negative. Figure 5.6 is the example for this noise added to Dataset 1 and Figure 5.7 for Dataset 8.

In both cases, the higher number of positive outliers caused the OLS fit to be biased towards the positive outliers. The GLM estimate, with the spline probability distribution function (pdf) fixed at an error mode of 3 psi was able to estimate the model parameters accurately for both examples.
5.3.2. Statistics of Starting Guesses

Accurate estimation for a wide range of starting guesses is a test of convergence for the estimation method. Figure 5.8 maps the percentage error in permeability estimate for the
example in Figure 5.5 with outliers of around 5 psi error. OLS and GLM were used to estimate model parameters for a range of input permeability and skin as starting guess. The true solution is at the center of the map (log k=1, log C=-2).

The figure indicates that the OLS estimate of permeability had more than 50% error for all starting guesses. The presence of positive outliers diminishes the ability of OLS to estimate permeability accurately. GLM, however, converged to the true solution for a wider range of starting guesses because it does not assume a certain error distribution and is able to accommodate the presence of outliers. For starting guesses further away from the true solution, GLM was not as accurate. We also indicate the results for a hybrid OLS+GLM method, where OLS is used to estimate a starting guess for GLM. The results show that this hybrid method is more robust and accurate than GLM with an arbitrary starting guess. This is likely because the OLS result provides a good starting estimate of the distribution of residuals, which is important for convergence to a better solution because GLM is highly sensitive to the initial guess. This hybrid method is discussed further in the next chapter.

![Figure 5.8: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, GLM and OLS+GLM methods, for Dataset 1 with outliers averaging 5 psi.](image)

Results of a similar test for multimodal noise (example in Figure 5.6) added to Dataset 1 are shown in Figure 5.9. The true solution is the same as before, at the center of each map. The results are similar to that of the previous example, OLS was least accurate due to its intolerance of outliers. OLS was reasonably robust for a large range of starting guesses,
but consistently estimated an inaccurate solution. GLM was more accurate, but not as robust to the starting guess. OLS+GLM successfully combines robustness and accuracy.

Figure 5.9: Maps of percentage permeability estimation error for a range of starting guesses (in log scale) using OLS, GLM and OLS+GLM methods, for Dataset 1 with multi-modal outliers around 5 and -5 psi.
Chapter 6

6. Performance and Scaling

This chapter discusses computational efficiency issues, including hybrid schemes for robust and efficient solution and the scaling of various methods with respect to the problem size.

6.1. Hybridization with OLS

One of the key observations regarding OLS from the previous chapters was its robustness to a wider range of starting guesses compared to ODR, Boggs and GLM. These methods are more accurate for a wider range of errors and error distributions by design, but are not as robust for starting guesses further away from the true solution.

Hybrid solution methods are a good way to combine the robustness of OLS with the accuracy of these other methods. OLS is used to provide a preliminary estimate of the solution, as it is more likely to approach the solution for a wider range of starting guesses. The OLS estimate is then used as a starting guess for the more accurate method chosen (ODR, Boggs, GLM). Figure 6.1 indicates the benefits of hybrid OLS+GLM used for an example with outliers discussed in the previous chapter. OLS is not accurate due to the presence of outliers in the data, consistently estimating permeability with an error of ~35%. However, this estimate, when used as a starting guess for GLM provides an accurate parameter estimate.
Figure 6.1: Example of the benefits of hybridization of OLS with GLM (repeated from Figure 5.9).

OLS usually requires fewer evaluations of the model function per iteration than ODR, Boggs and GLM. ODR and Boggs require additional evaluations of the model function to calculate the pressure derivative with respect to time for calculation of the time variables \((t^*, t^B)\). GLM needs calculation of the objective function twice as often per iteration than OLS. Therefore, OLS is usually faster to compute per iteration than the other methods.

Hybrid methods require more iterations of OLS to get close to the solution and fewer iterations of the more accurate, computationally intensive algorithm than such an algorithm would require from a further starting guess. Most often, the computation time required for OLS to provide a preliminary estimate is offset by the time saved due to fewer iterations of the more computationally intensive method. Figures 6.2 and 6.3 indicate one such example.

Figure 6.2 maps the error in permeability estimate using OLS, ODR and hybrid OLS+ODR for the case from Section 3.5.2 for Dataset 1 with three hours of breakpoint error. The hybrid OLS+ODR combines the robustness of OLS and accuracy of ODR as evident in the figure. The figure maps the time taken (in seconds) for solution for each starting guess using the three methods. OLS+ODR includes the time taken for OLS to provide the starting guess for ODR. Because less computation is required to approach the solution using OLS and fewer ODR iterations are required to converged to the accurate solution, the overall
time required for OLS+ODR is less than that of plain ODR. In this way, hybridization with OLS results in a fast, robust and accurate solution.

Figure 6.2: Map of permeability error for Dataset 1 with three hours of breakpoint error for varying starting guesses, computed using OLS, ODR and hybrid OLS+ODR.

Figure 6.3: Map of computation time for Dataset 1 with three hours of breakpoint error for varying starting guesses, computed using OLS, ODR and hybrid OLS+ODR.

6.2. Scaling with Problem Size

With the advent of surveillance instruments such as the permanent downhole gauge (PDG), the amount of data available for analysis is becoming increasingly larger. Instead of a few hours of buildup data, the data may contain years of pressure surveillance. These large data sets need methods that are not only accurate but also scale efficiently. For instance, the Boggs system of equations has more variables than the number of pressure measurements. This can become immensely time consuming for data from PDGs, and will scale unfavorably for larger data sets.
Figure 6.4 indicates the scaling of computing time with increasing number of data points using OLS, Boggs and the Boggs system using the Schur complement (BoggsSC) solution described in Section 3.1.1. While OLS scales linearly with problem size, Boggs method seems to scale in polynomial time, BoggsSC is more efficient and has almost-linear scaling.

Figure 6.4: Plot of computing time for OLS, Boggs and Boggs with Schur complement (SC) solver with increasing number of data points using parameters of Dataset 1.

Figure 6.5 indicates the results of a similar test for OLS, ODR and BoggsSC, OLS scales the most efficiently. While BoggsSC takes less time than ODR for relatively smaller problems, for problems with more than ~4500 data points it takes more time than ODR. Ithisawatpan’s (2012) results indicated that Boggs was more efficient than ODR for the datasets tested, however, such large datasets were not investigated.
6.2.1. Stochastic Methods

In the previous section, we compared the scaling performance of the OLS, ODR and BoggsSC and found that there is near-linear scaling of computation time with problem size. However, the more accurate ODR and BoggsSC may take three to five times more time than OLS and the computing time may be very high for the extremely large problems that are usually posed by PDG data.

Stochastic gradient descent is a method that is used routinely in the field of machine learning for training models to large data sets. The idea is to randomly sample a subsection of the data at the start of every iteration, calculate the objective and gradient for this subset of the data and update the model parameters to reduce the objective function for that subset. For the next iteration, a different subset of points is selected randomly. Instead of using gradient descent, the concept of stochastic iterations was adapted to the Gauss-Marquardt method discussed earlier. The objective function is calculated at least twice per iteration, at the start and end, and the convergence criterion applies to the minimum reduction in
objective function within an iteration. In the implementation developed in this research, the change in objective function had to be greater than 0.01 times the number of stochastic points to continue to next iteration.

Figure 6.6: Comparison of scaling with problem size and parameter errors using full OLS and stochastic OLS for Dataset 1 with no additional noise.

Figure 6.7: Comparison of scaling with problem size and parameter errors using full OLS and stochastic OLS for a Dataset 1 with pressure noise of 1 psi standard deviation.

Figure 6.6 compares the time required and mean parameter error for full OLS and OLS with stochastic iterations set to a fixed number of data points every iteration for a dataset with no noise. Stochastic OLS takes the same amount of time regardless of problem size and has mean parameter error less than a fraction of a percent. Figure 6.7 makes the same comparison for a similar dataset, but with pressure noise of standard deviation 0.5 psi added.
to the data. While stochastic OLS still converges in constant time regardless of the size of the problem, the mean parameter estimate error is more than that of OLS, because the number of stochastic points per iteration is fixed.

Since data points are chosen randomly each iteration, the final parameter estimate obtained using stochastic methods will vary every time it is used. Figure 6.8 indicates the mean and standard deviation calculated from 50 Monte Carlo runs for varying number of stochastic points in a 1000 point data set with pressure error with standard deviation of 0.5 psi. For a smaller fraction of stochastic points, the mean and variance is higher than for a larger fraction of the data, where the mean and standard deviation converge to the mean error and variance caused by the random realizations of noise.

![Figure 6.8: Comparing the mean and standard deviation of the permeability estimate (from 50 runs and noise realizations) using stochastic OLS for varying number of stochastic data points.](image)

We then investigated the effect of noise level and chosen number of stochastic points on the parameter estimates. Figure 6.9 indicate the results of this study on a 100 point data set with varying level of pressure noise. For no noise in the dataset, the same parameters are estimated every time regardless of the number of points. This is due to the fact that for
perfect data, the correct parameters can be estimated correctly from very few data points. With increasing noise, the variance in the results increased, especially for lower fractions of stochastic points from the full dataset.

![Figure 6.9: Plot of standard deviation of the permeability estimate using stochastic OLS (from 50 runs) on Dataset 1 with varying levels of noise in the data and varying number of stochastic points per iteration.](image)

This stochastic method can be used with all of the methods discussed in this report, barring Boggs method and GLM. For Boggs, because every data point has an associated variable, it is not possible to choose a different subset every iteration. Subsampling is still possible, but the statistics of the overall data set are not represented.

We did not test GLM with stochastic methods because the concept behind GLM is to derive the likelihood function from the residuals. If some information is left out in each iteration, it is likely to reduce the accuracy of the method. Theoretically, for large data sets we may assume stationarity of the errors over the subsample. In practice, this was not observed for the data sets that were tested.
Chapter 7

7. Conclusion

Parameter estimation is an important part of the pressure transient analysis workflow. Ordinary least squares (OLS) is the most commonly used method due to its generality and mathematical simplicity. OLS makes two major inherent assumptions; a) the presence of errors in the dependent (pressure) variable only and b) that these pressures follow a Gaussian distribution.

In Chapter 3, we discussed the development of Orthogonal Distance Regression (ODR) and Boggs methods, which consider the possibility of errors in the time measurements, which are usually treated as independent variables. OLS estimates of model parameters were found to be inaccurate and biased in the presence of random time errors, while ODR and Boggs method were found to be more accurate in such cases. OLS, however, was found to be robustly convergent for a larger range of starting guesses than ODR or Boggs.

For cases with breakpoint errors, ODR and Boggs were more accurate than OLS, but only up to a certain point beyond which estimates became inaccurate. Breakpoint regression (BPR) was developed to correct for breakpoint errors directly. Flow rate regression (FRR) was developed as an extension of BPR to address errors in flow rate measurement. Both BPR and FRR were found to be accurate for larger amounts of breakpoint error that ODR and Boggs were incapable of handling accurately. Additionally, BPR and FRR are robust to as wide a range of starting guesses as OLS, while being able to handle errors in the flow rate history. A hybrid variant of ODR (ODR+BPR/FRR) was developed for datasets that have random time errors in addition to errors in breakpoint and flow rate measurements and found to be more accurate for such cases.
In Chapter 5, we discussed the development of a method that is based on the concept of General Likelihood Maximization (GLM) for an arbitrary distribution of errors in the pressure measurement. This method bypasses the assumption of Gaussian distribution of errors that is inherent to OLS and hence is more widely applicable. Tests on datasets from the test matrix with non-Gaussian error distributions indicate that, while OLS estimates tend to be biased for such cases, GLM estimates are accurate. The concept was extended to a formulation for errors in both the pressure and time variable, but the implementation indicated problems with convergence that is likely due to the over-determined nature of the problem.

In Chapter 6, we discussed computational scaling and efficiency issues. Hybridization with OLS was shown as a way to lend robustness for a wide range of starting guesses. These hybrid methods are also faster to compute than the original methods while retaining accuracy. The scaling of computation time with problem size for these methods was also explored. Stochastic iteration methods were investigated as a means to manage the computing time needed for extremely large problems. The investigation revealed a relationship between the number of stochastic points per iteration, the amount of noise in the data and the variance of the estimated parameters. For large problems, this can be used to select the number of stochastic points per iteration for achieving estimates within a certain confidence interval.
Nomenclature

\( t_{dl} \) = \( i \)th time measurement in data set

\( p_{di} \) = \( i \)th pressure measurement in data set

\( \phi \) = porosity (fraction)

\( \mu \) = fluid viscosity (cP)

\( c_t \) = fluid compressibility (psi\(^{-1}\))

\( k \) = permeability (darcy)

\( S \) = skin factor (dimensionless)

\( C \) = wellbore storage coefficient (STB/psi)

\( R_e \) = distance to reservoir boundary (feet)

\( \omega \) = storativity ratio for dual porosity medium

\( \lambda \) = interporosity flow coefficient for dual porosity medium

\( \alpha \) = vector of unknown model parameters

\( \tau_i \) = \( i \)th breakpoint time (hours)

\( q_i \) = \( i \)th flow rate measurement (STB/day)

\( r_w \) = radius of wellbore (feet)

\( P_u \) = Unit rate pressure response function

\( \epsilon_{pi} \) = error in \( i \)th pressure measurement

\( \epsilon_{ti} \) = error in \( i \)th time measurement

\( \sigma_p \) = standard deviation of pressure errors

\( \sigma_t \) = standard deviation of time errors

\( \rho \) = line search step size

\( \lambda_{LM} \) = Levenberg Marquardt parameter
\( t_i^* \) = model variable in ODR corresponding to the \( i^{th} \) time measurement

\( t_i^B \) = model variable in Boggs method, corresponding to the \( i^{th} \) time measurement

\( \beta \) = ratio of variance estimate in ODR

\( f_p \) = piecewise cubic polynomial function to represent pressure error pdf

\( f_t \) = piecewise cubic polynomial function to represent time error pdf
References


Nomura, M., (2006), ‘Processing and Interpretation of Pressure Transient Data from Permanent Downhole Gauges’, Stanford University PhD thesis

Appendix A

A. Test Matrix

A test matrix was developed to test the methods discussed and developed in this research. Six types of reservoir models were implemented. Table A-1 lists the known model parameters. The models, flow rate history and the model unknown parameters are shown in Table A-2. Figure A-1 shows the flow history profiles used for generating the data sets.

Table A-1: Known model parameters for all data sets in the test matrix

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial pressure (psi)</td>
<td>5000</td>
</tr>
<tr>
<td>Porosity</td>
<td>0.1</td>
</tr>
<tr>
<td>Fluid Viscosity (cP)</td>
<td>1</td>
</tr>
<tr>
<td>Fluid Compressibility (1/psi)</td>
<td>3e-6</td>
</tr>
<tr>
<td>Wellbore radius (feet)</td>
<td>0.3</td>
</tr>
<tr>
<td>Formation thickness (feet)</td>
<td>30</td>
</tr>
<tr>
<td>Fluid Formation Volume Factor (bbl/STB)</td>
<td>1</td>
</tr>
</tbody>
</table>
Table A-2: Description of data sets in the test matrix

<table>
<thead>
<tr>
<th>#</th>
<th>Model</th>
<th>Flow Rate History</th>
<th>$k$ (md)</th>
<th>$S$</th>
<th>$C$ (STB/psi)</th>
<th>$L$ (feet)</th>
<th>$\omega$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Infinite line source</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>2</td>
<td>Circular no-flow boundary</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>3</td>
<td>Circular constant pressure boundary</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>4</td>
<td>Dual porosity infinite line source</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>5</td>
<td>Dual porosity circular no-flow boundary</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>6</td>
<td>Dual porosity circular constant pressure boundary</td>
<td>Flow history 1</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>7</td>
<td>Infinite line source</td>
<td>Flow history 2</td>
<td>10</td>
<td>3</td>
<td>0.01</td>
<td>1500</td>
<td>0.01</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>8</td>
<td>Infinite line source</td>
<td>Flow history 3</td>
<td>200</td>
<td>-2</td>
<td>0.1</td>
<td>3000</td>
<td>0.1</td>
<td>$10^{-7}$</td>
</tr>
</tbody>
</table>
Figure A-1: Flow rate history used for the datasets in the test matrix.