THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

NONLINEAR REGRESSION: APPLICATION TO WELL TEST ANALYSIS

by
Renato de Souza Carvalho

A dissertation submitted in partial fulfillment of

the requirements for the degree of Doctor of Philosophy

in the Discipline of Petroleum Engineering

The Graduate School

The University of Tulsa

1993

THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

NONLINEAR REGRESSION: APPLICATION TO WELL TEST ANALYSIS

by

Renato de Souza Carvalho

A DISSERTATION

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Dissertation Committee

_____, Co-Chairperson

_____, Co-Chairperson

_____

_____

_____

ABSTRACT

Carvalho, R. S. (Doctor of Philosophy in Petroleum Engineering)

Nonlinear Regression: Application to Well Test Analysis (277 pp. – Chapter VII)

Co-Directed by Dr. Albert C. Reynolds, Jr. and Dr. Leslie G. Thompson

(306 words)

This work presents procedures for analyzing well-test pressure data by using nonlinear parameter estimation and statistical techniques. We present a general regression analysis procedure that combines robust least absolute value (LAV) estimation, least squares (LS) estimation, outlier detection, and statistical analysis (e.g., confidence region and confidence intervals). It is shown that our LAV implementation works effectively, and has the ability to account for outliers in the data.

We present a new method for constraining the parameters within the feasible region, which guarantees that the estimates satisfy the imposed constraints in all circumstances. The LAV implementation and the proposed constraining method are readily applied in conjunction with existing unconstrained least-squares routines.

It is demonstrated through synthetic examples that for dual porosity models, we should first regress on pressure derivative data, and then use the results of this run as initial estimates for a final regression on pressure. In this procedure, regression should be applied to the original derivative data—derivative data should not be smoothed significantly even if the data are extremely noise. This does not contradict the fact that significant smoothing is often necessary to obtain derivative data useful for model identification.

For the horizontal-well model, we show it may be important to use preliminary regressions on specific time intervals of the data where one or more parameters are not important, in order to obtain good initial estimates for a final regression. This procedure enhances the chance of obtaining convergence to the right solution. It is shown that deconvolution is important to reveal the theoretical behavior of the early time flow regimes, and that the correlation coefficients of the parameters can help to confirm the existence of specific flow regimes.

Finally, we present a new algorithm to take the Laplace transform of tabulated real data, which yields accurate results for both pressure and pressure derivative data.

ACKNOWLEDGMENTS

I wish to express my gratitude and appreciation to Dr. Albert C. Reynolds, Jr., and Dr. Leslie G. Thompson for their invaluable assistance and guidance as co-directors of this research project. I also express my gratitude to Dr. Richard A. Redner of the University of Tulsa, Dr. Ram Agarwal of Amoco Production Co., and Dr. Peyton J. Cook of the University of Tulsa, for their comments and suggestions and for serving on my dissertation committee.

I would like to also express my appreciation to all the other faculty members and to my graduate student colleagues, who contributed to my education as a TU graduate student.

I am grateful to my company Petrobrás, Petróleo Brasileiro S.A., for investing in my education and for supporting my studies at the University of Tulsa. I also wish to acknowledge TUPREP, The University of Tulsa Petroleum Reservoir Exploitation Projects, for providing copies, slides and other facilities.

This work is dedicated to my parents Washington and Alminda Carvalho, and to my daughters Renata, Carolina and Marina. I specially dedicate this work to my wife Maria Inez whose support, patience, encouragement and companionship will always be appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Well test analysis is an important tool in the petroleum industry. For many years, the interpretation of transient pressure tests has been one of the most important sources for estimating reservoir properties such as transmissibility and storativity, and for detecting the presence of boundaries and formation heterogeneities. Further, well test analysis provides important information about the well condition, i.e., stimulated or damaged. The engineer can successfully evaluate the performance of a well based on the results of transient pressure tests in conjunction with available information from geology and logs.

Interpretation of well test data consists of defining an appropriate well/reservoir model which best represents the system under study, and adjusting the reservoir-model parameters to obtain the closest match between field and model data. A careful examination of the log-log plot of the pressure and pressure derivative data versus time, and other available information, usually defines the appropriate reservoir model. Conventional type curve analysis relies on manual alignment of field pressure (and pressure derivative) responses with corresponding (idealized) model pressure responses to obtain the system parameters. This approach is inherently subjective; thus, it is possible that a single set of pressure transient data may be matched differently (by different analysts, or by the same analyst at different times), to yield inconsistent estimates of the system parameters. Depending on the model complexity and the quality of the pressure data, the interpretation of transient well tests can demand a considerable amount of work and time. Much of the work and subjectiveness of the classical type curve matching procedure can be removed by applying computer-aided, automated, type-curve matching. There are many additional advantages of automated type-curve matching over the conventional graphical techniques:

it is superior for analyzing tests with a variable flow-rate history, (since rate variations can be directly incorporated into the model pressure response[1]); it can handle complex reservoir/wellbore models (e.g., horizontal wells, bounded systems) for which type curves are inadequate or nonexistent; and it is immune to noisy pressure derivative data when applied to the analysis of pressure data only. An efficient automated method does not require subjective adjustments of parameters from the analyst and decreases the probability of numerical and procedural errors.

The objective of this work is to study the application of nonlinear regression and statistical techniques in the analysis of pressure transient tests.

1.1 Literature Review

Although nonlinear regression analysis (i.e., automatic type curve matching) is generally more applicable than classical type curve matching, it also suffers from non-uniqueness problems.[2,3,4] Different initial guesses may yield different results, and for some initial guesses, the schemes may not converge. Also, actual recorded data may contain outliers,[5,6] i.e., pressure points which (compared to the behavior of the majority of the measured data) cannot be adequately described by the model well-reservoir system pressure response. (These points are also referred to as "bad observations".) As is well known, the nonlinear LS regression analysis technique may be significantly influenced by the outliers. Thus, an important consideration in the successful application of nonlinear regression analysis to pressure transient data is the ability to account for outliers. (The robust regression analysis technique described in Chapter II addresses this issue.)

In the early application of computer-aided well test analysis, the theoretical pressures were computed by using numerical simulators, or by using simple analytical solutions such as the line-source well in an infinite homogeneous reservoir. Rapid

technical advances in computer hardware and software and the reliability of the Stehfest algorithm[7] (1970) for the inversion of Laplace transforms made automated analysis feasible. Rosa and Horne[2] (1983) successfully used the general analytical solutions for both homogeneous and non-communicating (commingled) layered reservoirs. They showed that by numerical inversion from Laplace space of both pressure and parameter gradients, it is possible to apply least squares to fit field data to a more complicated model than can be conveniently used with the traditional type curve approach. Rosa and Horne applied the first order Gauss-Newton method. They also considered the Gauss-Marquardt method (also known as Levenberg-Marquardt method) because of the ill-conditioning of the Hessian matrix observed in their work. The Marquardt modification improves the conditioning of the Hessian matrix by adding a small positive number to the diagonal elements, which increases the magnitude of the eigenvalues of the matrix.

Barua *et al.*[3] (1988) suggested a second order algorithm, the so-called Newton-Greenstadt method, and a modification of it, referred to here as the Newton-Barua method. Nevertheless, Refs. 2 and 3 reported problems in convergence when many parameters, or parameters that only marginally affect the pressure specific data set, were to be determined simultaneously. In some cases, they encountered difficulties when inaccurate initial estimates were used.

Namba and Horne[4] (1988) discussed a modified Cholesky decomposition that is applicable for symmetric positive-indefinite matrices. They claim that the modified Gauss-Cholesky method works as well as the Gauss-Newton and Gauss-Marquardt methods under favorable conditions, and it can act more reliably than the other two methods under unfavorable conditions, i.e., cases with poor initial guesses associated with ill-defined parameters such as ϕc_t.

Abbaszadeh and Kamal[1] (1988) developed an automated well test interpretation method based on the modified Levenberg-Marquardt least squares regression algorithm, where the parameter gradients are evaluated numerically. The method takes into account

any rate history by using the superposition principle of the pressure drawdown solutions of the selected well/reservoir model.

More recently, Rosa and Horne[8] (1991) employed a robust nonlinear parameter estimation that uses the least absolute value (LAV) as the criterion for the minimization. The problem is solved iteratively as a sequence of multiple linear regression problems (linearization method). However, this approach exhibits convergence problems for poor initial estimates.

Wilkinson[9] (1992) presented procedures for parameter estimation in the Laplace space for pressure transient test that were recorded beginning at $t = 0$, i.e., with the reservoir at initial condition.

Horne[10] (1992) provided a good summary of the methodology that has been used in computed-aided interpretation of pressure transient tests.

Note that many papers about computer-aided well test analysis focused on alternative algorithms to overcome the computational difficulties inherent in the classical Newton and Gauss Method. One of the objectives of our work is to delineate reliable, efficient procedures for the estimation of reservoir parameters using regression techniques based on publicly available software. The majority of public-domain codes are based on nonlinear least-squares (LS) regression analysis, and this is the focal point of our analysis strategy, (although we often employ it in a way that renders it equivalent to a robust estimator). Among the many arguments in favor of using the LS method is that reliable software (such as the package we use, LMDER[11] from the Argonne National Laboratory's MINPACK library) is readily available. We also discuss statistical analysis, in particular confidence regions and confidence intervals of the parameter estimates, which are tied directly to the least squares procedure. We demonstrate the applicability of the proposed procedures by analyzing sets of synthetic data as well as actual field tests.

1.2 Outline

Chapter II presents an overview of the nonlinear regression techniques considered meaningful for well test applications. We discuss the least squares (LS) estimation, the robust least absolute value (LAV) estimation, confidence region and confidence intervals of the parameters, the residual-based methods for outlier detection, and the handling of parameter constraints. A general regression analysis procedure that combines LAV and LS estimations, outlier detection, and statistical analysis are presented. A nice feature of our LAV estimation implementation and the novel imagining method used to constrain the parameters is that they are readily applicable in conjunction with existing unconstrained LS routines.

Chapter III focuses on the application of regression analysis to the classical wellbore storage and skin model. There, we investigate the effect of defining regression parameters in terms of correlating parametric groups instead of the individual parameters commonly used. The applicability of pressure derivative information is also investigated. Motivated by the Onur-Reynolds type curves, a two-step regression procedure that incorporates information available from pressure derivative data is presented. This procedure works better for short tests than the regression on pressure data only which is commonly used.

In Chapter IV, we apply nonlinear regression techniques to the horizontal well and dual porosity models. Regression analysis is especially helpful for these more complex models, for which conventional type curve analysis is difficult or inadequate. A general procedure for obtaining the pressure response in Laplace space for an anisotropic dual porosity system, based on the corresponding solution for the homogeneous system, is presented. For these more complex models, we demonstrate the necessity of using

pressure derivative information (in addition to the pressure data information). For the horizontal well model, we present a procedure that uses the results from preliminary regressions on specific parts of the data (where one or more of the parameters have little or no influence on the pressure response) to devise better initial estimates for regressing on all parameters using the entire span of the test data.

Chapter V reviews existing methods for verifying whether the basic assumptions of the error term are violated. Some diagnostics are based on the visual analysis of different plots of the residuals, and others are based on statistical tests. The methodology is illustrated by its application to field data.

Chapter VI discusses Laplace space convolution and deconvolution techniques. A new algorithm for computing the numerical Laplace transform of tabulated data that yields accurate Laplace transforms for both pressure and pressure derivative data is presented. We show that the proposed algorithm yields better results than the algorithm presented by Roumboutsos-Stewart[12] (1988), and slightly better results than a recent method presented by Bourgeois and Horne[13] (1993). We also briefly discuss ways of applying parameter estimation in Laplace space.

Finally, Chapter VII summarizes the results and enumerates the contributions and conclusions of this study.

CHAPTER II

NONLINEAR PARAMETER ESTIMATION TECHNIQUES

2.1 Least Squares Parameter Estimation

The basic objective of analyzing pressure transient data using nonlinear regression analysis is the determination of an optimum set of system parameters, α, such that the observed data match the ideal model data as closely as possible. The most common method of ensuring a "close" fit between measured and model data, is to require that the *sum of the squares of the residuals* attains a minimum value. Here, the uth residual, r_u, is defined as the difference between the uth measured data point, y_u, and the corresponding model response, $F_u(\alpha)$. Thus, for example, if we are regressing on measured pressure data, the appropriate uth residual would assume the form,

$$r_u \equiv y_u - F_u(\alpha) = p_{\text{meas},u} - p_{\text{model},u}. \tag{2.1.1}$$

Suppose we have n measured data points, $\{y_u\}$. Least squares (LS or L_2-norm) regression attempts to minimize the objective function, $E_2(\alpha)$, where,

$$E_2(\alpha) = \sum_{u=1}^{n} r_u^2 . \tag{2.1.2}$$

In Eq. 2.1.2, α denotes the set of parameters on which the model response depends. Thus, for example, if $\{y_u\}$ denotes measured pressures for a reservoir-well system with permeability k, skin s, and storage C, the set of model parameters assumes the form,

$$\alpha = \{k, C, s\}. \tag{2.1.3}$$

A necessary condition for the minimization of the objective function (Eq. 2.1.2) is that the partial derivative of the objective function with respect to each model parameter is zero. Thus, for the case of a well with skin and storage, the necessary condition becomes,

$$\frac{\partial E_2}{\partial k} = \frac{\partial E_2}{\partial C} = \frac{\partial E_2}{\partial s} = 0. \tag{2.1.4}$$

An outline of the method used to find the parameters which minimize the LS objective function (as implemented in **LMDER**[11]), is provided in Appendix A. Appendix B presents a summary on linear regression theory. Some important aspects of the nonlinear model theory are presented in Section 2.3.

Equation 2.1.2 indicates that the objective function for LS analysis depends on the *square* of the residuals. If our measured data set contains a few "bad" observations (i.e., outliers), the residuals associated with these points tend to be relatively large, and the ultimate minimum value of the objective function (and therefore the optimized parameter estimates) may be significantly influenced by the outliers.

2.2 Robust Parameter Estimation

One approach to reduce the impact of data outliers on the final regression result is to apply the so-called *robust* regression techniques.[5,6,8] The most well known of these methods, the LAV method, obtains parameter estimates by searching for the Least Absolute Value (i.e., the minimum L_1-norm) of the residuals. We refer to this particular robust estimator as the L^1 method, or as in Ref. 8, the LAV method. For the LAV method, the appropriate objective function, $E_1(\alpha)$, which we minimize is

$$E_1(\alpha) = \sum_{u=1}^{n} |r_u| \,. \tag{2.2.1}$$

Note that in this case, the ultimate effect of the outliers (i.e., high residual values) on the minimized objective function is reduced since the residuals are *not* squared as in the LS regression. Let us illustrate this point with the following simple linear regression model

$$y = ax + b, \tag{2.2.2}$$

where the slope a and the intercept b are to be estimated by using the data points presented in Table 2.2.1. Since the data points lie perfectly on a straight line, both LS and LAV methods yield the right slope, $a = -0.25$, and the right intercept, $b = 2.25$ (see Fig. 2.2.1). However, suppose that for some reason, the value of y_4 is recorded as 4.30 representing an outlier. Fig. 2.2.2 shows that the outlier has a large effect on the LS estimation (dashed line), which is quite different from the LS line in Fig. 2.2.1. Also, note that the outlier has essentially no effect on the LAV line.

Table 2.2.1 - Data points for the linear model example

x_u	y_u
1.0	2.00
2.0	1.75
3.0	1.50
4.0	1.25
5.0	1.00

If we seek to minimize the LAV objective function (Eq. 2.2.1) by setting the partial derivatives of $E_1(\alpha)$, with respect to the model parameters, to zero, we arrive at an ill-posed problem because the absolute value function is not differentiable at zero. Moreover, since most public domain nonlinear regression routines are based on LS analysis, (i.e., where the objective function is of the form of Eq. 2.1.2), the routines are not directly applicable to minimizing $E_1(\alpha)$.

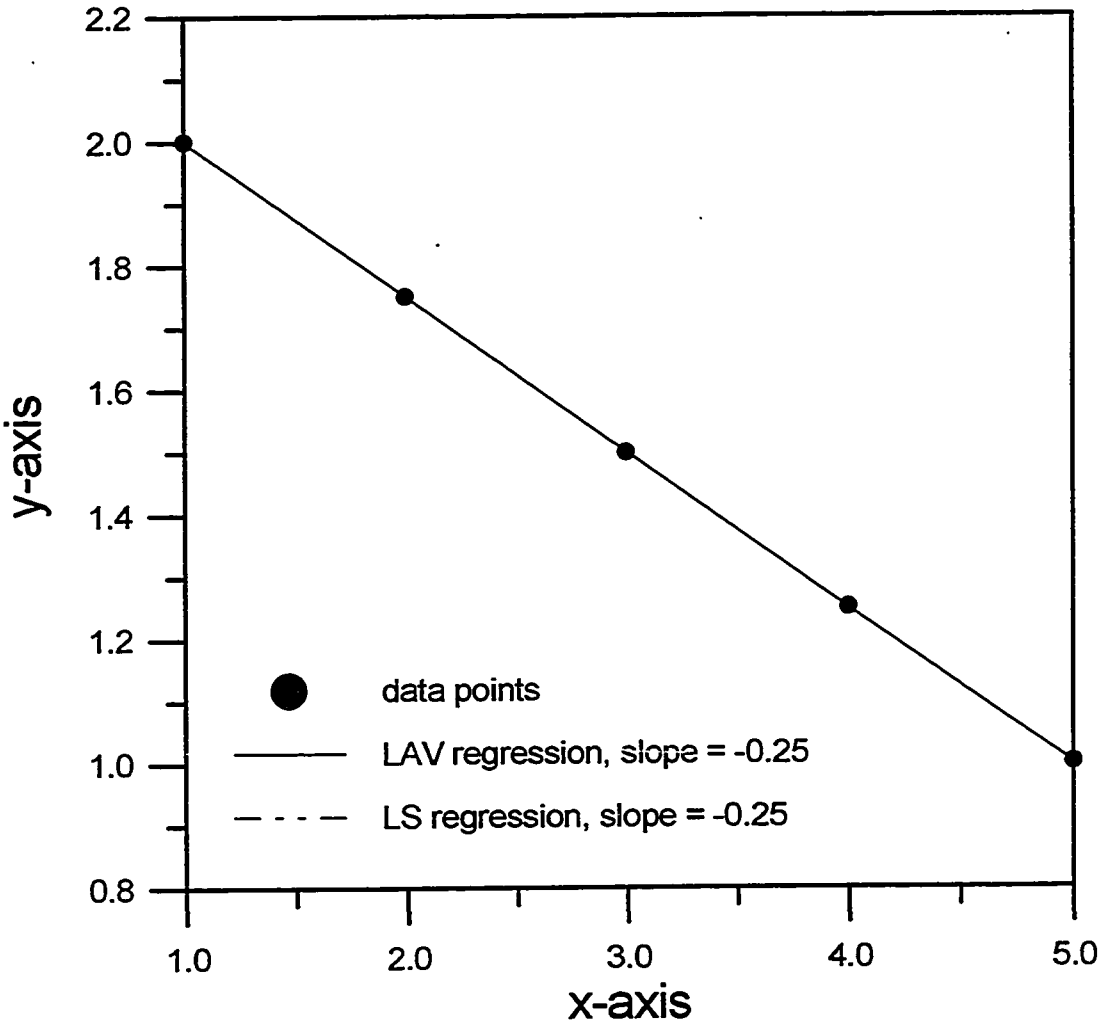

Fig. 2.2.1 - LS and LAV linear regressions with no outliers

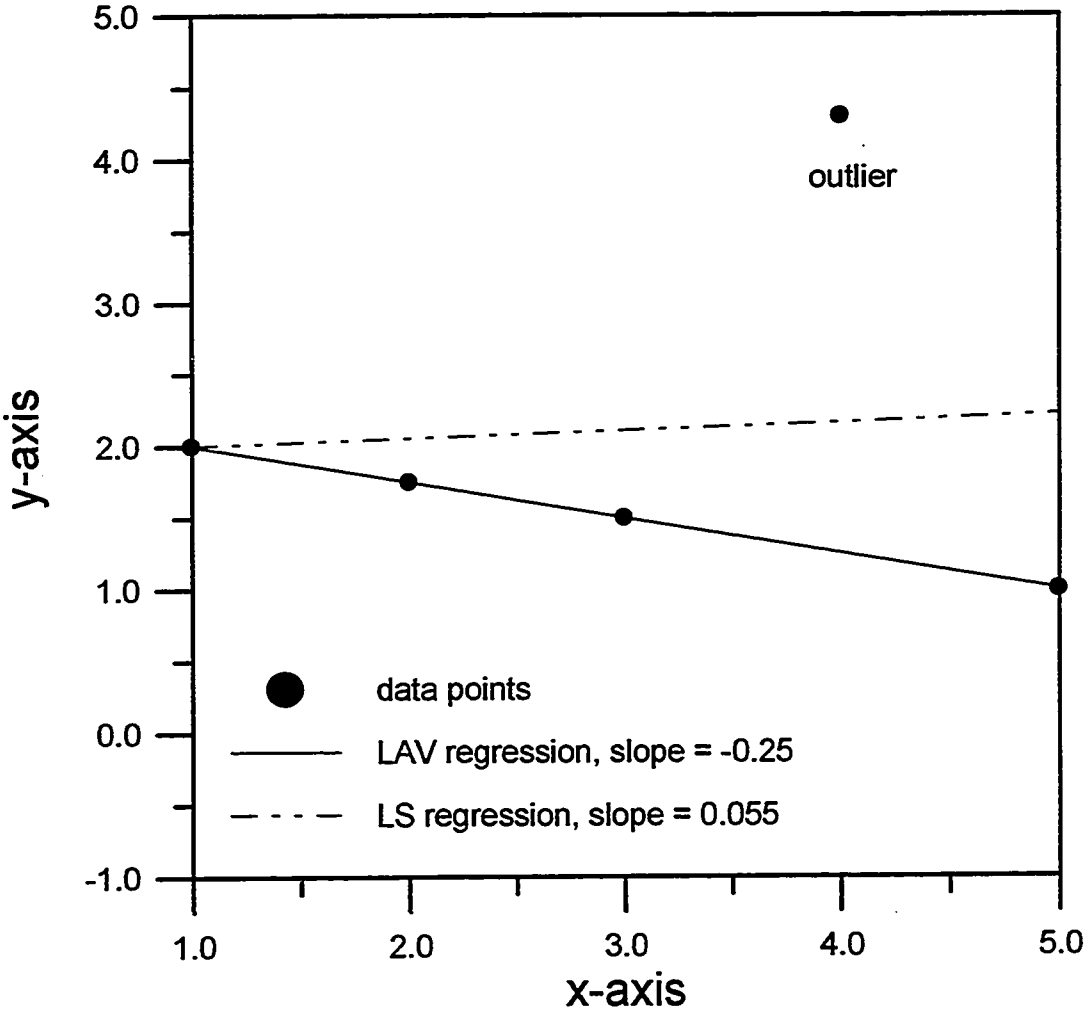

Fig. 2.2.2 - LS and LAV linear regressions with one outlier

Reference 23 proposes the following method of L_1-norm minimization. Define a new objective function, $\hat{E}_1(\alpha,\varepsilon)$, as

$$\hat{E}_1(\alpha,\varepsilon) \equiv \sum_{u=1}^{n} \sqrt{r_u^2 + \varepsilon}, \tag{2.2.3}$$

where ε is a small number. Reference 23 shows that if Eq. 2.2.3 is minimized with decreasing values of ε, and the resulting sequence of solutions converge to α^* as $\varepsilon \to 0$, then α^* minimizes $E_1(\alpha)$ defined in Eq. 2.2.1. Note that Eq. 2.2.3 can be rewritten in terms of an *equivalent* L_2-norm minimization problem as follows:

$$\hat{E}_1(\alpha,\varepsilon) \equiv \sum_{u=1}^{n} \rho_u^2, \tag{2.2.4a}$$

where,

$$\rho_u = \left(r_u^2 + \varepsilon\right)^{1/4}. \tag{2.2.4b}$$

The advantages of minimizing $\hat{E}_1(\alpha,\varepsilon)$ given by Eq. 2.2.4a instead of $E_1(\alpha)$ defined in Eq. 2.2.1 are as follows: (i) the partial derivatives of the new objective function, $\hat{E}_1(\alpha,\varepsilon)$, are defined everywhere; (this permits the use of powerful derivative based minimization methods); and (ii) public domain LS solution algorithms are directly applicable to the problem.

An iterative algorithm for L_1-norm minimization problems can be summarized as follows:[23]

Step 1: Pick an initial guess of the unknown parameters, α^0, and a small positive number ε_1; set $k = 1$. (Typically we use $\varepsilon_1 = 10^{-1}$.)

Step 2: Find α_k^* which minimizes the objective function $\hat{E}_1(\alpha,\varepsilon_k)$ given by Eq. 2.2.4a. In this step, we use the LMDER algorithm of Reference 11 (see Appendix A).

Step 3: Set $\varepsilon_{k+1} = \varepsilon_k / L$ where L is a number greater than 1. Typically $L = 10^2$.

Step 4: If the scheme converges within a pre-specified tolerance, STOP. Otherwise, set $k = k+1$ and go back to step 2.

In step 4, the convergence is verified in terms of the parameter estimates as well as in terms of the value of the objective function, i.e., for convergence we require that

$$\left|\alpha_k^* - \alpha_{k-1}^*\right| \leq \delta \tag{2.2.5}$$

and

$$\left|\hat{E}_1(\alpha_k^*, \varepsilon_k) - \hat{E}_1(\alpha_{k-1}^*, \varepsilon_{k-1})\right| \leq \varphi, \tag{2.2.6}$$

where δ and φ are pre-specified small numbers depending on the accuracy desired.

The choice of an appropriate starting value for ε_1 depends on the scaling of the specific problem. On the basis of numerical experience with this algorithm, El-Attar *et al.*[23] use ε_1 equal to one tenth of the largest absolute value of the residuals (or the average absolute value of the residuals) computed at the initial guess α^0. Typically we use ε_1 equal to 0.1.

Based on our limited experience, we have found that performing a single regression with $\varepsilon \approx 10^{-10}$ usually suffices. However, we suggest the use of the above iterative algorithm because the total number of iterations required is about the same.

The question now arises as to whether LAV regression should *always* be used to obtain the most reliable parameter estimates. In the sense that the LAV method reduces the impact of data outliers, the answer is yes. However, most of the statistical analyses used in evaluating the "goodness of fit," e.g., confidence intervals, parameter correlation, etc., are developed based on the LS approach[5] (see Appendix B and Section 2.3). In our work, we apply a combination of these two methods. In particular, we advocate the use of the LAV regression as a diagnostic tool for detecting and removing data outliers; then use LS analysis on the reduced data set for final parameter estimates and the statistics of the

match. This is in contrast to Ref. 8, where the LAV approach is used to accommodate the outliers without detecting them.

2.3 Topics on Nonlinear Model Theory

The nonlinear model is generically defined as a model of the form,

$$y = F(\alpha,\xi)+\varepsilon, \tag{2.3.1}$$

where F is the model function, $\alpha=(\alpha_1, \alpha_2,\ldots,\alpha_p)^T$ is the vector of p unknown parameters to be estimated, and $\xi=(\xi_1, \xi_2,\ldots,\xi_m)^T$ is the vector of the independent variables. The same assumptions about the error term, ε, we stated for the linear case (see Appendix B) will be assumed here; i.e., for each set of values for the independent variables, ε is independently and normally distributed with mean zero and constant variance σ^2.

In well test analysis, time is usually the only independent variable, thus, we will write our model as $F(\alpha, t)$.

We define the *error sum of squares* for the nonlinear model and the given data as

$$S(\alpha)=\sum_{u=1}^{n}[y_u - F(\alpha,t_u)]^2, \tag{2.3.2}$$

where (t_u, y_u), $u = 1, \ldots, n$, is a set of n observations. Note that since t_u and y_u are fixed observations, $S(\alpha)$ is a function of α only.

Let $\hat{\alpha}$ denote the estimate of the unknown vector of the parameters, α. For example, if we want a least-squares (LS) estimate, $\hat{\alpha}$ would be the value of α which minimizes the error sum of squares given by Eq. 2.3.2 (often called *LS objective function*). However, since F is nonlinear, the resulting p normal equations will also be nonlinear, and iterative methods must be employed in nearly all cases.

2.3.1 Gradient Methods

In well testing, and in many other engineering applications, the most common iterative procedures to minimize Eq. 2.3.2 are the gradient methods based on the classical Newton's method. Among these methods, we mention the Gauss-Newton method and the Gauss-Marquardt method.

Newton's method can be derived by using a quadratic approximation to the objective function $S(\alpha)$ obtained by expanding $S(\alpha)$ in a Taylor series up to the second order term about an initial guess for the vector of parameters. See Appendix A for more details.

Another method for obtaining the parameter estimates is the so-called *linearization method.*[16] In contrast to Newton's method, which approximates the objective function $S(\alpha)$, this method uses a linear approximation to the model function $F(\alpha, t)$. The linear approximation of $F(\alpha, t)$ about an initial guess of the parameters, α_0, is

$$F(\alpha,t_u)=F(\alpha_0,t_u)+\sum_{i=1}^{p}\left.\frac{\partial F(\alpha,t_u)}{\partial\alpha_i}\right|_{\alpha_0}(\alpha_i-\alpha_{i0}), \tag{2.3.3}$$

where α_{i0} is the ith component of the vector α_0. When α is sufficiently close to α_0 so that the approximation of Eq. 2.3.3 applies, our nonlinear model given by Eq. 2.3.1 can be approximated by the following linear equation:

$$Y_u^0=\sum_{i=1}^{p}\beta_i^0 Z_{iu}^0+\varepsilon_u, \tag{2.3.4}$$

where

$$Y_u^0=Y_u-F(\alpha_0,t_u) \tag{2.3.5a}$$

$$\beta_i^0=\alpha_i-\alpha_{i0} \tag{2.3.5b}$$

$$Z_{iu}^0 = \left.\frac{\partial F(\alpha, t_u)}{\partial \alpha_i}\right|_{\alpha=\alpha_0} . \qquad (2.3.5c)$$

The idea behind the linearization method is to solve the original nonlinear regression problem as a sequence of linear regression problems.

If we denote by b_i^0 the estimates of β_i^0 obtained by solving Eq. 2.3.4, an iterative procedure can be written, in a convenient notation as follows:

1) Assume an initial guess of the parameters α_{i0}, $i = 1, ..., p$.

2) Estimate the modified parameters ($\beta_i^0 = \alpha_i - \alpha_{i0}$) in Eq. 2.3.4 by applying linear regression theory. Denote the estimates of β_i^0 by b_i^0.

3) If the solution converges to some pre-specified small tolerance STOP. That is, if $\left|b_i^0/\alpha_{i0}\right| < \delta$, $i = 1,\ldots,p$, STOP. Otherwise, find $\alpha_{i1} = \alpha_{i0} + b_i^0$.

4) Set $\alpha_{i0} = \alpha_{i1}$, and GO TO step 2.

There is no guarantee that in step 3 the estimates are actually improved at each iteration, i.e., there is no guarantee that $S(\alpha_{i1}) < S(\alpha_{i0})$. References 8 and 16 point out that the linearization method may exhibit convergence problems in some applications.

Bard[21] discusses improvements to the above iterative procedure. He presents methods in which the step size is a function of a relaxation parameter $\lambda > 0$, i.e., the update procedure in step 3 is written as $\alpha_{i1} = \alpha_{i0} + \lambda b_i^0$, and λ is chosen such that, in each iteration, the objective function is significantly reduced.

It is important to make some observations about this method. First, observe that step 2 performs the relatively easy task of minimizing a linear model. Moreover, step 2 is not restricted to LS estimation, i.e., we can use any estimator to minimize the linear problem. For example, one may use LAV (least absolute value) estimation instead of LS estimation in step 2. Indeed, Ref. 8 uses a linear LAV routine of the commercial package IMSL to implement their LAV estimation procedure.

Secondly, note that the linearization method reduces to the standard Gauss-Newton method for LS estimation. If we denote by $\mathbf{Y}^0$, β^0 and $\mathbf{Z}_0$ the matrices (and vectors) corresponding to Eq. 2.3.5, and let $\mathbf{b}_0$ represent the least-squares estimate of β^0, Then, (in analogy with Eq. B.6) the normal equations of the linearization method become

$$\mathbf{Z}_0^T\mathbf{Z}_0\mathbf{b}_0 = \mathbf{Z}_0^T\mathbf{Y}^0. \tag{2.3.6}$$

Note that Eq. 2.3.6 is equivalent to Eq. A-2, provided that in Eq. A-2 we use the Gauss-Newton Hessian matrix given by Eq. A-8. Thus

$$\left(\mathbf{Z}^T\mathbf{Z}\right)_{jk} = \frac{1}{2}(\mathbf{H})_{jk}, \tag{2.3.7}$$

where $\mathbf{H}$ is the Hessian matrix in Eq. A-8.

Many papers in well testing applications have focused on alternative algorithms to overcome the computational difficulties inherent in the classical Newton and Gauss-Newton methods. Reference 10 presents a very good summary of the methods that have been used in well testing problems. Appendix A presents a description of the LMDER routine (a Gauss-Newton-Marquardt method implementation), which we have successfully used to date.

When the linearized form of the nonlinear model is used, all usual formulas and statistical analyses of the linear regression theory (see Appendix B) can be applied. However, care should be exercised, since all the results are valid to the extent that the linearized form provides a good approximation to the true model and that the statistical assumptions are valid.

2.3.2 Confidence Regions

An exact confidence contour is defined by taking, in Eq. 2.3.2, $S(\alpha) = a$, (a is a constant). However, for the general nonlinear model, we are unable to obtain the exact

confidence level associated with a given value a. We can still define a confidence region by the expression

$$S(\alpha) = S(\hat{\alpha})\left\{1+\frac{p}{n-p}F(p,n-p,1-\gamma)\right\}, \tag{2.3.8}$$

which provides an exact $(1-\gamma)\times 100\%$ confidence region for the linear case (see Appendix B, Eq. B.24) even though the confidence level will not be γ in the nonlinear model. We will call such a region an "approximate $(1-\gamma)\times 100\%$ confidence region" for α. In Eq. 2.3.8, $F(p,n-p,1-\gamma)$ is the value that cuts off $\gamma\times 100\%$ in the upper tail of the F-distribution with p and $n-p$ degrees of freedom; p is the number of regression parameters; n is the number of data points; and $\hat{\alpha}$ is the parameter vector which minimizes the objective function $S(\alpha)$. For the nonlinear model, Eq. 2.3.8 may define "banana-shaped" regions in contrast to the ellipsoidal regions obtained for the linear case. Clearly, the procedure of computing exact confidence contours by using Eq. 2.3.8, will require the evaluation of $S(\alpha)$ for many different values of the vector α (followed by an interpolation procedure), which normally demands an unreasonable amount of computational work. Eq. 2.3.8 is rarely applied for this reason.

The general confidence regions tend to become ellipsoidal regions if the nonlinear model *behaves linearly*, i.e., if the model can be approximated by Eq. 2.3.3 in the vicinity of the solution. Moreover, the confidence region shrinks as the significance level γ increases. Since the distances between the parameter estimates and the boundaries of the confidence region are reduced in this case, the errors introduced by the linearity assumption are reduced, and the shape of the confidence region approaches an ellipsoid.

In analogy with Eq. B.22, an approximate (linearized) confidence region can be calculated via

$$(\alpha-\hat{\alpha})^T\hat{\mathbf{Z}}^T\hat{\mathbf{Z}}(\alpha-\hat{\alpha}) = ps^2F(p,n-p,1-\gamma). \tag{2.3.9}$$

In Eq. 2.3.9, $\hat{\mathbf{Z}}$ is the $n\times p$ matrix given by Eq. 2.3.5, evaluated at the solution $\hat{\alpha}$, i.e.,

$$\hat{\mathbf{Z}} = \left\{\hat{Z}_{iu}\right\} = \left\{\left.\frac{\partial F(\alpha, t_u)}{\partial \alpha_i}\right|_{\alpha=\hat{\alpha}}\right\}, \tag{2.3.10}$$

and

$$s^2 = \frac{\sum_{u=1}^{n}\left[y_u - F(\hat{\alpha}, t_u)\right]^2}{n-p} = \frac{S(\hat{\alpha})}{n-p}. \tag{2.3.11}$$

Again, the ellipsoid given by Eq. 2.3.9 will only be a good approximation to the true confidence region if the model behaves linearly. Under the linearity assumption, the values of $S(\alpha)$ computed at any point on the boundary of the confidence region should be the same. Thus, one can compare the values of $S(\alpha)$ at points on the boundary of the confidence region and decide whether the linear assumption is valid. Draper *et al.*[16] suggest we compare values of $S(\alpha)$ at the end points of the major axes of the ellipsoidal approximated confidence region of Eq. 2.3.9. If the linear approximation is valid, $S(\alpha)$ should be constant at the boundaries of this ellipsoid and hence at the endpoints of the axes of this ellipsoid. These endpoints can be found by canonical reduction, which requires computation of the eigenvalues and the eigenvectors of the symmetric $p \times p$ matrix $\hat{\mathbf{Z}}^T\hat{\mathbf{Z}}$.

2.3.3 Confidence Intervals

In the general nonlinear model, $\hat{\alpha}$ may not have a Gaussian distribution, and the variance-covariance matrix is not necessarily of the form $(\mathbf{XX})^{-1}\sigma^2$ (see Appendix B). Thus, we do not know how to compute confidence intervals in the general case. However, we can define approximate confidence intervals for the various parameters by assuming a linearized form of the nonlinear model. Thus, in analogy with Eq. B.12, the approximate confidence limit of the parameter α_i is determined by

$$\hat{\alpha}_i - t(\gamma, n-p)\sqrt{s^2\left(\left(\hat{\mathbf{Z}}^T\hat{\mathbf{Z}}\right)^{-1}\right)_{ii}} \le \alpha_i \le \hat{\alpha}_i + t(\gamma, n-p)\sqrt{s^2\left(\left(\hat{\mathbf{Z}}^T\hat{\mathbf{Z}}\right)^{-1}\right)_{ii}}, \quad (2.3.12)$$

where $\hat{\mathbf{Z}}$ and s are given by Eqs. 2.3.10 and 2.3.11, respectively. Dogru' *et al.*[24] were the first to employ Eq. 2.3.12 in well test applications. Here (refer to Eq. B.8), the variance-covariance matrix is estimated to be

$$V(\hat{\alpha}) = (\hat{\mathbf{Z}}^T\hat{\mathbf{Z}})^{-1}s^2. \quad (2.3.13)$$

Similarly, the correlation coefficient between parameters $\hat{\alpha}_i$ and $\hat{\alpha}_j$ is defined as

$$\rho_{ij} = \frac{Cov(\hat{\alpha}_i, \hat{\alpha}_j)}{\left[Var(\hat{\alpha}_i)Var(\hat{\alpha}_j)\right]^{1/2}}. \quad (2.3.14)$$

In Eq. 2.3.14, the quantities $Var(\hat{\alpha}_i)$, $Var(\hat{\alpha}_j)$ and $Cov(\hat{\alpha}_i, \hat{\alpha}_j)$ are elements of the (linearized) variance-covariance matrix, $V(\hat{\alpha})$, defined in Eq. 2.3.13. Specifically, the element in the ith row and jth column of $V(\hat{\alpha})$ is equal to $Cov(\hat{\alpha}_i, \hat{\alpha}_j)$. Recall $Cov(\hat{\alpha}_i, \hat{\alpha}_i) = Var(\hat{\alpha}_i)$.

On the basis of the arguments presented in Appendix B, most of the equations presented here for the (unweighted) least squares nonlinear problem are also valid for the weighted least squares estimation, provided that the matrix $\hat{\mathbf{Z}}^T\hat{\mathbf{Z}}$ is replaced by $\hat{\mathbf{Z}}^T\mathbf{W}\hat{\mathbf{Z}}$ and provided that s^2 is based on the weighted least squares objective function.

2.3.3.1 Confidence Intervals for LAV Regression

In a recent paper, Rosa and Horne[8] applied a LAV robust estimation with the objective to accommodate possible outliers in the data set. They suggest computing confidence intervals for the LAV parameter estimates based on the following weighted least squares objective function:

$$E = \sum_{u=1}^{n} w_u \left[y_u - F(\alpha, t_u) \right]^2, \tag{2.3.15}$$

where

$$w_u = \frac{1}{|r_{u,LAV}|} = \frac{1}{|y_u - F(\hat{\alpha}_{LAV}, t_u)|}. \tag{2.3.16}$$

Thus, the weights given by Eq. 2.3.16 are the inverses of the residuals from the LAV regression. If the model is correct, this is equivalent to assuming that the errors at each data point have unequal variances, and that these variances are proportional to the inverse of weights computed from the data by using Eq. 2.3.16. Moreover, they claim that the confidence intervals of the weighted least squares (Eq. 2.3.15) and the confidence intervals of the LAV regression are equivalent, because the minimizer of the objective function given by Eq. 2.3.15 is exactly equal to the LAV solution. Under these assumptions, the confidence intervals for the LAV parameter estimates can be computed by using Eq. B-34. As usual, for nonlinear regression, the matrix **X** in Eq. B-34 should be replaced by the matrix **Z** defined in Eq. 2.3.5c. Since both problems have the same solution, a second regression is not required. This approach has possible problems. First, it is doubtful that the correct weights (i.e., the variance of the errors) can be estimated based on a single realization of the test. Also, it is possible that some residuals may be too small compared to the others, so that the weights assigned to those data points are unexpectedly large; This causes the confidence intervals given by Eq. B.34 to be unexpectedly small. Thus, there is a need to limit the weights to some threshold. However, our experience is that the confidence intervals of the parameters are quite different depending on the threshold. For example, when this procedure is applied to the Synthetic Example #3 (seed = –34), presented in Chapter III, the confidence intervals for skin are quite different, i.e., ±0.0040% and ±2.42% corresponding to the thresholds 10^{-10} and 10^{-5} psi, respectively. Thus, we use another approach. We advocate the detection of outliers based

on the LAV residuals. Once the outliers are removed from the original data set, a classical LS regression is applied on the remaining data. Then, all statistical analyses (including confidence regions and intervals of the parameter estimates) are performed for this LS regression.

If the constant variance assumption of the errors is true, then the reweighted least squares (RLS) regression (i.e., the LS regression after deleting the outliers) will give accurate confidence intervals. If it is the case that the errors have unequal unknown variances, then the confidence interval procedure proposed by Ref. 8 could be more accurate, provided that Eq. 2.3.16 gives reasonable estimates of the true weights, which we believe is unlikely.

2.3.4 Canonical Reduction and Verification of the Linearity Assumption

The confidence intervals given by Eq. 2.3.12 can be inaccurate because they rely heavily upon the linear approximation of the model[22] (see Eq. 2.3.3). If the linearity assumption is really questionable, one should use alternative procedures to compute confidence intervals as we will discuss in the next section.

A convenient method for testing the linearity assumption is as follows. Determine the endpoints of the p-dimensional confidence region of the parameter estimates based on the linearity assumption. If this assumption is valid, the actual values of the objective function $S(\alpha)$ calculated at these points should be the same and only slightly different from the following approximation value

$$S(\hat{\alpha}) + ps^2 F(p, n-p, 1-\gamma). \tag{2.3.17}$$

Eq. 2.3.17 comes directly from Eqs. 2.3.8 and 2.3.11.

The endpoints of the axes of the ellipsoid can be determined by *canonical reduction*. Let $\delta\alpha = \alpha - \hat{\alpha}$ be the p-dimensional column vector, and let $\mathbf{A} = \hat{\mathbf{Z}}^T\hat{\mathbf{Z}}$ be the

$p \times p$ matrix that contains the parameter gradients of the functional model. Thus the confidence region given by Eq. 2.3.9 can be written as

$$\delta\alpha^T \mathbf{A} \delta\alpha = c, \tag{2.3.18}$$

where c is a constant, $c = ps^2 F(n, n-p, 1-\gamma)$. The left-hand-side of Eq. 2.3.18 is called a *quadratic form* defined by **A**. Note that if **A** is diagonal, then the ellipsoid of Eq. 2.3.18 is already in *canonical form*, and the endpoints of the ith axis of the ellipsoid are given by $\pm(c / A_{ii})^{1/2}$. In this case, all entries of the coordinates of the endpoints are zero, except for the entry corresponding to the $\delta\alpha_i$-axis. The matrix **A** is not diagonal in general and a change of coordinates is necessary to transform Eq. 2.3.18 in canonical form (no off diagonal terms).

Since **A** is a real symmetric matrix, the eigenvalue decomposition of **A** is given by[21]

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T, \tag{2.3.19}$$

where **D** is diagonal; $D_{ii} = d_i$, the ith eigenvalue of **A**; and **P** is a real (unitary) orthogonal matrix ($\mathbf{P}^{-1} = \mathbf{P}^T$) whose columns are the normalized eigenvectors of **A**. Then Eq. 2.3.18 can be written as

$$\delta\alpha^T \mathbf{A} \delta\alpha = \delta\alpha^T \mathbf{P}\mathbf{D}\mathbf{P}^T \delta\alpha = c. \tag{2.3.20}$$

If we define $\psi = \mathbf{P}^T \delta\alpha$ we obtain

$$\delta\alpha^T \mathbf{A} \delta\alpha = \psi^T \mathbf{D} \psi = \sum_{i=1}^{p} d_i \psi_i^2 = c. \tag{2.3.21}$$

The expression $\psi^T \mathbf{D} \psi$ is the canonical form of the quadratic expression $\delta\alpha^T \mathbf{A} \delta\alpha$, and ψ is referred to as the vector of the *canonical variables*. Eq. 2.3.21 indicates that the principal axes of the ellipsoid coincide with the coordinate axes in the ψ space. Thus the endpoints of the axis of the ellipsoid along the ψ_i coordinate axis are $\pm(c / d_i)^{1/2}$, and hence the corresponding length is $2(c / d_i)^{1/2}$. Note that since **P** is unitary, the

transformation of coordinates given by $\psi = \mathbf{P}^T \delta\alpha$ is a rigid rotation of the axes, which leaves the form of the ellipsoid and distances unaffected. The endpoints of the ellipsoid (or any point in the ψ space) can be expressed in the original coordinate system via

$$\delta\alpha = \left(\mathbf{P}^T\right)^{-1}\psi = \mathbf{P}\psi. \tag{2.3.22}$$

It is important to emphasize that the canonical variables, ψ_i's, are made up of linear combinations of the original variables, such that in the new coordinate system these variables are uncorrelated.

If the linear approximation of the model is not valid within the confidence region of the parameters, such that the contours of $S(\beta)$ are markedly non-elliptical, then the confidence region described by the quadratic form given by Eq. 2.3.9 as well as the confidence intervals obtained from Eq. 2.3.12 can be misleading. In this case, the inference about the regression parameters can be based on the likelihood ratio tests.

2.3.5 Likelihood Ratio Confidence Regions

In this section, we will briefly describe a method for constructing likelihood ratio confidence regions. Since the least squares is a special case of the weighted least squares, the procedure will be introduced from the point of view of weighted least squares.

We split the vector of parameters into two components,

$$\alpha^T = \left(\alpha_1^T, \alpha_2^T\right),$$

where α is a vector of length p and α_2 is a vector of length r, $r \le p$. We are interested in testing the null hypothesis,

$$H_0: \ \alpha_2 = \alpha_{2,0},$$

where $\alpha_{2,0}$ is specified. Let $L(\alpha_1,\alpha_2)$ be the log-likelihood function based on the full model with all parameters included, and let $L(\alpha_1,\alpha_{2,0})$ be the log-likelihood function

assuming that the null hypothesis is true, i.e., $\alpha_2 = \alpha_{2,0}$. The likelihood ratio test statistic is given by[22,25]

$$\mathrm{LR}(\alpha_{2,0}) = 2\left[L(\hat{\alpha}_1, \hat{\alpha}_2) - L(\hat{\alpha}_{1,0}, \alpha_{2,0})\right], \tag{2.3.23}$$

where $(\hat{\alpha}_1, \hat{\alpha}_2)$ denotes the maximum likelihood estimator (MLE) of (α_1, α_2) for the full model, and $\hat{\alpha}_{1,0}$ is the MLE of α_1 under the constraint that $\alpha_2 = \alpha_{2,0}$. This statistic has an asymptotic chi-squared distribution with r degrees of freedom. The likelihood test rejects the null hypothesis, with a probability γ for the type I error, if

$$\mathrm{LR}(\alpha_{2,0}) \geq \chi^2(r, 1-\gamma). \tag{2.3.24}$$

The likelihood ratio test (Eq. 2.3.24) for the weighted least squares case can be written as[22,25]

$$\mathrm{LR}(\alpha_{2,0}) = n\left[\log\left(S(\hat{\alpha}_{1,0}, \alpha_{2,0})\right) - \log\left(S(\hat{\alpha}_1, \hat{\alpha}_2)\right)\right] \geq \chi^2(p-r, 1-\gamma), \tag{2.3.25}$$

where $S(\hat{\alpha})$ is the LS objective function at $\hat{\alpha}$, i.e., $S(\hat{\alpha})$ represents the weighted sum of the squared residuals given by

$$S(\hat{\alpha}) = \sum_{u=1}^{n} w_u \left[y_u - F(\hat{\alpha}, t_u)\right]^2. \tag{2.3.26}$$

An approximate $(1-\gamma)\times 100\%$ confidence region for α_2 is given by the set of all $\alpha_{2,0}^*$ satisfying

$$2\left[L(\hat{\alpha}_1, \hat{\alpha}_2) - L\left(\hat{\alpha}_{1,0}, \alpha_{2,0}^*\right)\right] \leq \chi^2(p-r, 1-\gamma). \tag{2.3.27}$$

The region defined by Eq. 2.3.24 is likelihood-based. This means that when the asymptotic distribution is a poor approximation, regions of the right shape are obtained even though the confidence levels corresponding to the regions may not be accurate. Computing such regions require repeated nonlinear minimizations even in the case that $\alpha_{2,0}$ is a scalar (i.e., a single parameter). In contrast to this procedure, the classical confidence region (or inter-

vals) requires only one minimization, but they are restricted to the quadratic approximation of the surface.

2.3.6 Handling of Parameter Constraints

In an unconstrained optimization algorithm, it is possible that non-physical values of the parameters (e.g., $k < 0$) may be selected in the process of minimizing the objective function. In order to keep the parameters physically meaningful, some type of constraint should be used. References 1 and 2 employ *penalty functions* which transform the constrained problem into an unconstrained problem by adding terms (penalty functions) to the objective function. The penalty functions (one for each constraint) are defined in such a way that the new objective function remains almost unchanged in the interior of the feasible region of the parameters, but increases rapidly as a parameter approaches the boundary. Thus, if we constrain each parameter α_i to lie between some minimum value and maximum value, i.e., $\alpha_{\min,i} < \alpha_i < \alpha_{\max,i}$, then the new unconstrained objective function, $E^+(\alpha)$, can be written as[1]

$$E^+(\alpha^k) = E(\alpha^k) + F_c \sum_{i=1}^{p} \left(\frac{\delta_{\min,i}}{\alpha_i^k - \alpha_{\min,i}} + \frac{\delta_{\max,i}}{\alpha_{\max,i} - \alpha_i^k} \right), \tag{2.3.28}$$

where α_i^k is the current value of parameter α_i, i.e., the estimate of α_i at the kth iteration, and $E(\alpha^k)$ is the original unconstrained objective function. F_c, $\delta_{\min,i}$ and $\delta_{\max,i}$ are positive values that govern the amount of contribution from the penalty functions. The superscript k refers to the current iteration. One choice for a set of values F_c, $\delta_{\min,u}$ and $\delta_{\max,u}$ is, respectively,[1]

$$F_c = 10^{-4} E(\alpha^{k-1}), \tag{2.3.29}$$

$$\delta_{\min,i} = \alpha_{0i} - \alpha_{\min,i}, \tag{2.3.30}$$

and

$$\delta_{max,i} = \alpha_{max,i} - \alpha_{0i}, \tag{2.3.31}$$

where α_{0i} is the initial estimate of parameter α_i. Note that, in this formulation, $\delta_{min,i}$ and $\delta_{max,i}$ are computed only at the beginning of the regression, but F_c changes at each iteration.

Reference 2 shows that the use of penalty functions improves the convergence of the iterative procedure; however, they also report that the penalty function method may not avoid non-physical parameter values when the initial estimates are far from the solution. Because of this defect, and also because the penalty function method option is not available in some unconstrained routines (like Subroutine **LMDER** which we use) without changing the code, we have been using two alternative methods to constrain the parameters, i.e., to ensure that we do not compute estimates which are unreasonable from a physical point of view. These two methods are presented in the next sections.

2.3.6.1 Relaxation of the Step Size

As provided, Subroutine **LMDER** is an unconstrained algorithm for optimized parameter estimation. We have used the following constraining algorithm for regression on the wellbore storage and skin model. If the new estimated value of k or C is less than zero, or the new estimate of skin factor s is less than -4, we modify the value to a physically feasible estimate by reducing the change in the size of the parameter requested by LMDER.

Thus, for example, if the step change, $\Delta\alpha_i^c$, in a parameter α_i over an iteration is such that

$$\alpha_i^+ \equiv \alpha_i^c + \Delta\alpha_i^c < \alpha_{min,i}, \tag{2.3.32}$$

where $\alpha_{min,i}$ is the minimum acceptable value of the parameter, we modify the value of $\Delta\alpha_i^c$ to $0.75\{\alpha_{min,i} - \alpha_i^c\}$. Note that in applying this procedure, the direction $\mathbf{d}^c$ computed by **LMDER** is not honored. Although we have not established a precise mathematical basis for this procedure, we have found that the method successfully expedites convergence to a physically meaningful solution for the wellbore storage and skin problem. In fact, we tried a similar method which reduces the step size, but preserves the step direction, but this procedure was not as successful as the method based on modifying individual $\Delta\alpha_i^c$.

For more complex models with many parameters (e.g., a horizontal well model), the procedure based on modifying $\Delta\alpha_i^c$ sometimes fails to converge to the correct solution; in particular, we found that one or more parameters occasionally remain fixed at a value close to the boundary. In an attempt to overcome this problem, we use another procedure which is discussed in the next section.

2.3.6.2 Imaging Extension Procedure

In order to keep the parameters physically meaningful, we devised a new constraining algorithm which is referred to as *the imaging extension procedure*. The idea behind the method is to extend the objective function in such a way that the new objective function is defined everywhere (i.e., unconstrained) and that the solution of this new unconstrained problem is the same as (or related to) the solution of the original constrained problem.

Let us illustrate the proposed method with a model of the form, $F(\alpha,t)$, where α is the single parameter of the model (say permeability, k) and t is the independent variable (say time, t). Since permeability can not assume non-positive values, the original problem should be constrained through the inequality $k > 0$. If we reflect the objective function

through the ordinate axis, the new objective function is now defined for positive values of k as well as for negative values of k (see Fig. 2.3.1). Clearly, if k^* minimizes the new objective function, then $|k^*|$ will also minimize the original problem. Thus, we choose to minimize the modified unconstrained problem. Sometimes, we also refer to the modified problem as an *equivalent problem*.

An algorithm to handle a single constraint of the form, $\alpha > \alpha_{\min}$, can be written as

Step 1. Is the current parameter value of the equivalent problem, α_{eq}^c, greater than $\alpha_{\min}$? If YES, compute the objective function and the parameter gradient in the usual way. If NO, GO TO step 2.

Step 2. Is α_{eq}^c equal to $\alpha_{\min}$? If YES, STOP. If NO, compute the corresponding parameter value of the original problem as

$$\alpha^c = \left|\alpha_{eq}^c - \alpha_{\min}\right| + \alpha_{\min}. \tag{2.3.33}$$

Step 3. Compute the current value of the unconstrained objective function, E^+, as

$$E^+\left(\alpha_{eq}^c\right) = E\left(\alpha^c\right). \tag{2.3.34}$$

Step 4. Compute the value of the gradient of the unconstrained objective function at α_{eq}^c as

$$\left.\frac{\partial E^+(\alpha)}{\partial \alpha}\right|_{\alpha_{eq}^c} = \left.\frac{-\partial E(\alpha)}{\partial \alpha}\right|_{\alpha^c}. \tag{2.3.35}$$

Once a minimizer, α_{eq}^*, of the equivalent unconstrained problem is found, the corresponding minimizer of the original problem is computed by using Eq. 2.3.33. The above algorithm can be easily extended for problems with more than one parameter subject to the same type of constraint $\alpha_i > \alpha_{\min,i}$. Many applications of the proposed method for different functional models (e.g., horizontal well and double-porosity models) indicate that the algorithm works effectively for well test applications.

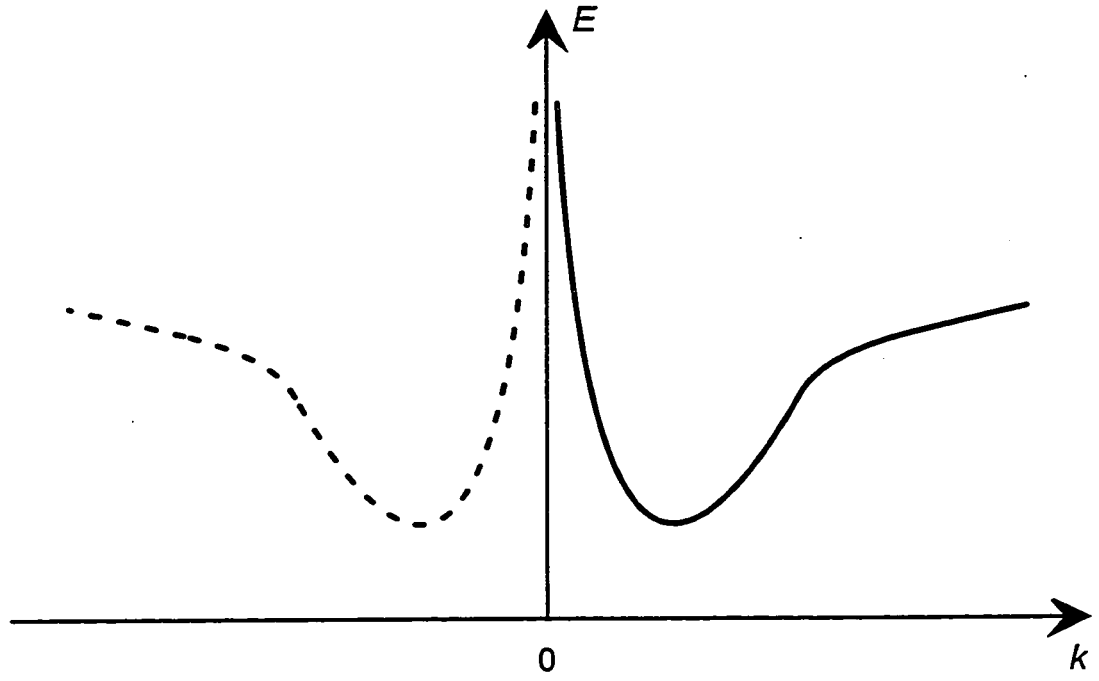

Fig. 2.3.1 - Illustration of the imaging extension of the objective function

The imaging procedure can be generalized to handle constraints of the type $\alpha_{\min,i} < \alpha_i < \alpha_{\max,i}$. In this case, the objective function is periodically repeated by a sequence of images that extend to $\pm\infty$ (see Fig. 2.3.2). The continuous curve in Fig. 2.3.2 is the original objective function, which is defined only in the interval (a,b), and the dashed curves represent the images. The minimizer of the original problem, α^*, can be computed once the minimizer α^*_{eq} of the extended objective function is found. Thus, first we apply the iterative minimization procedure to the modified (unconstrained) objective function, and then we find the corresponding solution to the original problem which is geometrically related to the solution of the modified (unconstrained) problem. Note that, by construction, at the boundaries of each image, the new (unconstrained) objective function, E^+, is continuous; however, its partial derivative with respect to the parameter that is at the boundary may not exist. In fact, its partial derivative exists (and is zero) at a "boundary" only when the corresponding derivative of the original objective function is zero, i.e., when

$$\left.\frac{\partial E(\alpha)}{\partial \alpha}\right|_{boundary},$$

at the corresponding boundary, is zero. In the event a boundary is hit during the iterative process, we assume, for that iteration, that the partial derivative with respect to the parameter at the boundary is zero.

An algorithm to handle constraints of the form, $\alpha_{\min,i} < \alpha_i < \alpha_{\max,i}$, can be written as

Step 1. Set the parameter index, $i = 1$.

Step 2. For the current value of the ith parameter of the equivalent problem, $\alpha^c_{eq,i}$, compute α^c_i via,

$$\alpha^c_i = \left|\alpha^c_{eq,i} - \alpha_{\min,i}\right| + \alpha_{\min,i}. \qquad (2.3.36)$$

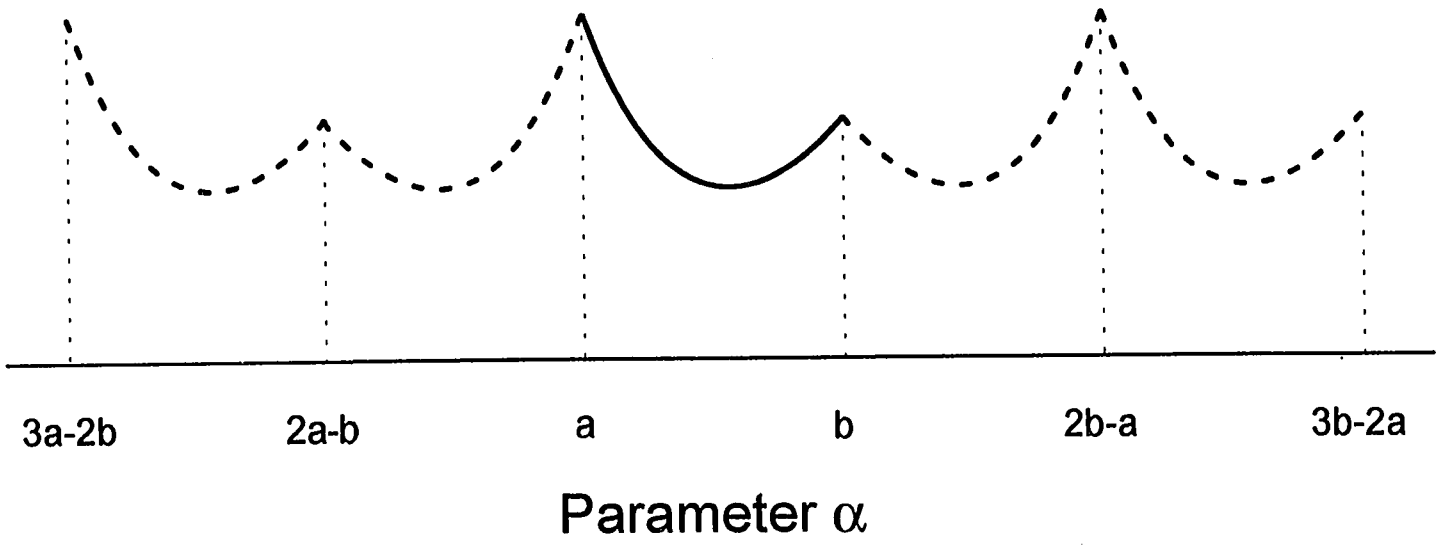

Fig. 2.3.2 - Illustration of the objective function extension for a constraint of the type $a < \alpha < b$

Step 3. Is $\alpha_{eq,i}^{c}$ greater than $\alpha_{\min,i}$? If YES, set $SIGN_i = 1$. If NO, set $SIGN_i = -1$.

Step 4. Compute

$$\alpha_i^c = \alpha_i^c - 2(\alpha_{\max,i} - \alpha_{\min,i})\,\mathrm{DINT}\left[\frac{\alpha_i^c - \alpha_{\min,i}}{2(\alpha_{\max,i} - \alpha_{\min,i})}\right], \tag{2.3.37}$$

where DINT(x) is the operation of truncating x to the whole number, i.e., DINT(x) is the greatest integer function so DINT(x) is equal to the largest integer less than or equal to x.

Step 5. Is α_i^c (given by Eq. 2.3.37) greater than $\alpha_{\max,i}$? If NO, GO TO step 6. If YES, set $SIGN_i = -SIGN_i$, and compute

$$\alpha_i^c = \alpha_{\max,i} - \left(\alpha_i^c - \alpha_{\max,i}\right). \tag{2.3.38}$$

Step 6. Is α_i^c equal to $\alpha_{\min,i}$ or $\alpha_{\max,i}$? If YES, set $SIGN_i = 0$.

Step 7. Repeat steps 2 to 6 for all parameters, i.e., for $i = 1,\ldots,p$.

Step 8. Compute the current value of the unconstrained objective function, E^+, as

$$E^+\left(\alpha_{eq}^c\right) = E\left(\alpha^c\right). \tag{2.3.39}$$

Step 9. Compute the value of the gradients of the unconstrained objective function at α_{eq}^c as

$$\left.\frac{\partial E^+\left(\alpha_{eq}\right)}{\partial \alpha_{eq,i}}\right|_{\alpha_{eq}^c} = SIGN_i \left.\frac{\partial E(\alpha)}{\partial \alpha_i}\right|_{\alpha^c}, \text{ for } i = 1,\ldots p. \tag{2.3.40}$$

In step 2, the current parameter estimate $\alpha_{eq,i}^{c}$ is reflected through the left boundary of the original objective function (i.e., the boundary $\alpha_i = \alpha_{\min,i}$) when $\alpha_{eq} < \alpha_{\min,i}$. Note that every time a reflection occurs the variable SIGN (which controls the sign of the derivative) changes sign. In step 4, the estimate is transformed to lie in the first repetitive element of the new objective function (which consists of the original

objective function and its right reflection), without changing either the value of the objective function or the value of its partial derivative. In step 5, if the estimate corresponds to the reflection of the original objective function, it is transformed (i.e., reflected through the boundary $\alpha_i = \alpha_{\max,i}$) to obtain an α_i^c value within the feasible interval. At the end, $SIGN = 1$ means that, in terms of the unconstrained problem, we are in a replica of the original objective function, and $SIGN = -1$ means that we are in a replica of the reflection.

When a minimizer, α_{eq}^*, of the equivalent unconstrained problem is found, the corresponding minimizer of the original problem is given by the coordinates of α^c (Eqs. 2.3.37 or 2.3.38) at the last iteration.

Let us illustrate the proposed method with a dual-porosity example where we regress on the parameter ω, storativity ratio. By definition, the feasible interval for the storativity ratio is $0 \le \omega \le 1$. The minimization procedure (with our constraining algorithm) was applied to synthetic pressure data which was generated with $\omega = 0.10$. Figures 2.3.3 though 2.3.7 present the changes in the parameter ω during the iterative process, for five different initial estimates, $\omega = 0.95$, 0.70, 0.50, 0.25 and 0.05, respectively.

The solid continuous lines in Figs. 2.3.3 though 2.3.7 refer to the path of iterates for the equivalent (unconstrained) problem, and the dashed lines refer to the equivalent path for the original problem. All five regressions converged to the right solution, i.e., $\omega = 0.10$. For example, Fig. 2.3.3 shows that for the initial estimate $\omega = 0.95$, the regression applied to the equivalent problem converged to $\omega_{eq} = -1.90$ (i.e., the algorithm converged to a minimum in the second image to the left), which corresponds to $\omega = 0.10$. Note that for the two cases where the initial estimates were closer to the solution, i.e., $\omega = 0.05$ and 0.25 (Figs. 2.3.7 and 2.3.6, respectively), the constraining algorithm was not necessary since ω always remained in the feasible region.

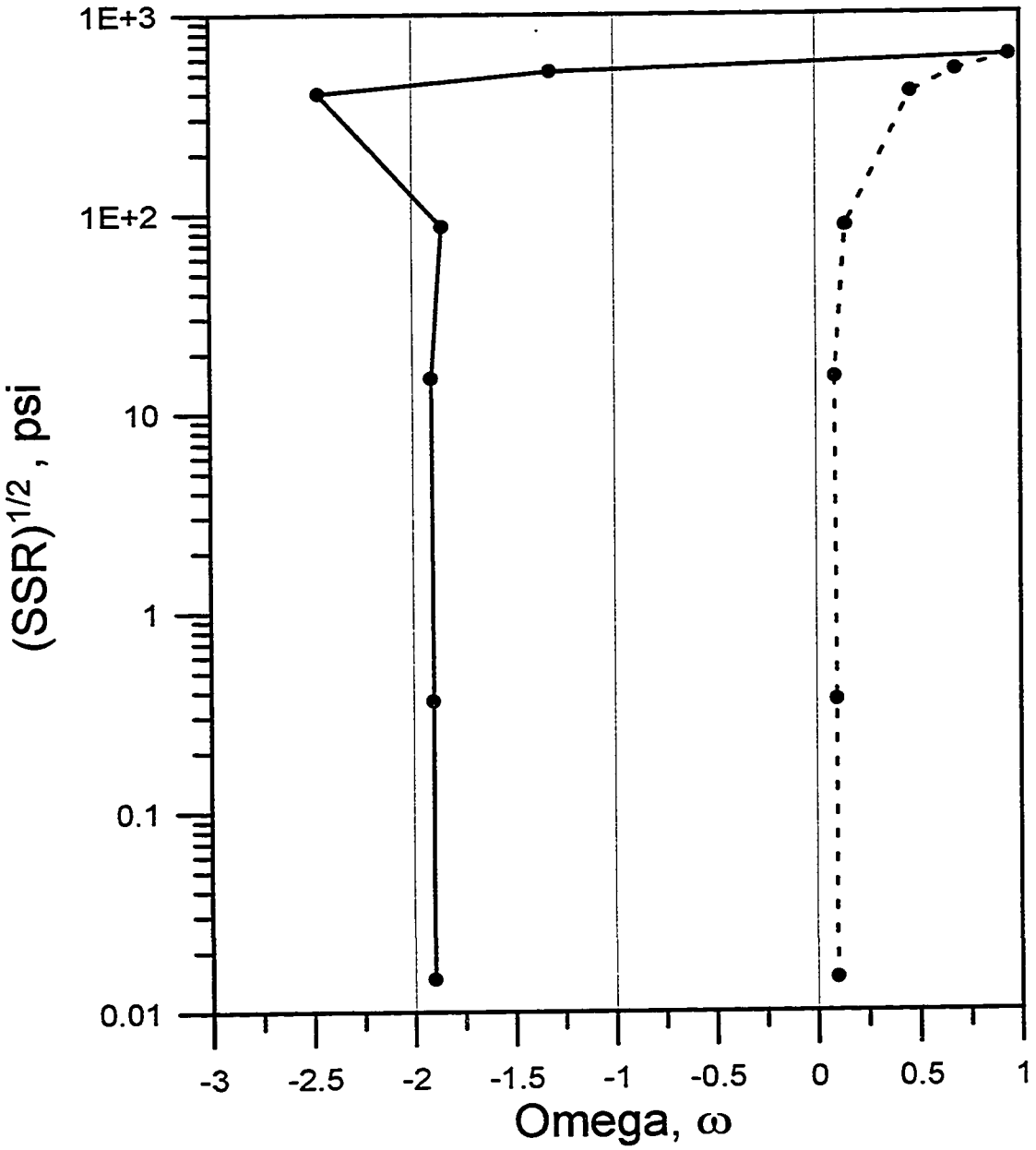

Fig. 2.3.3 - Change in the estimates of ω during the iterative process. Initial estimate ω = 0.95

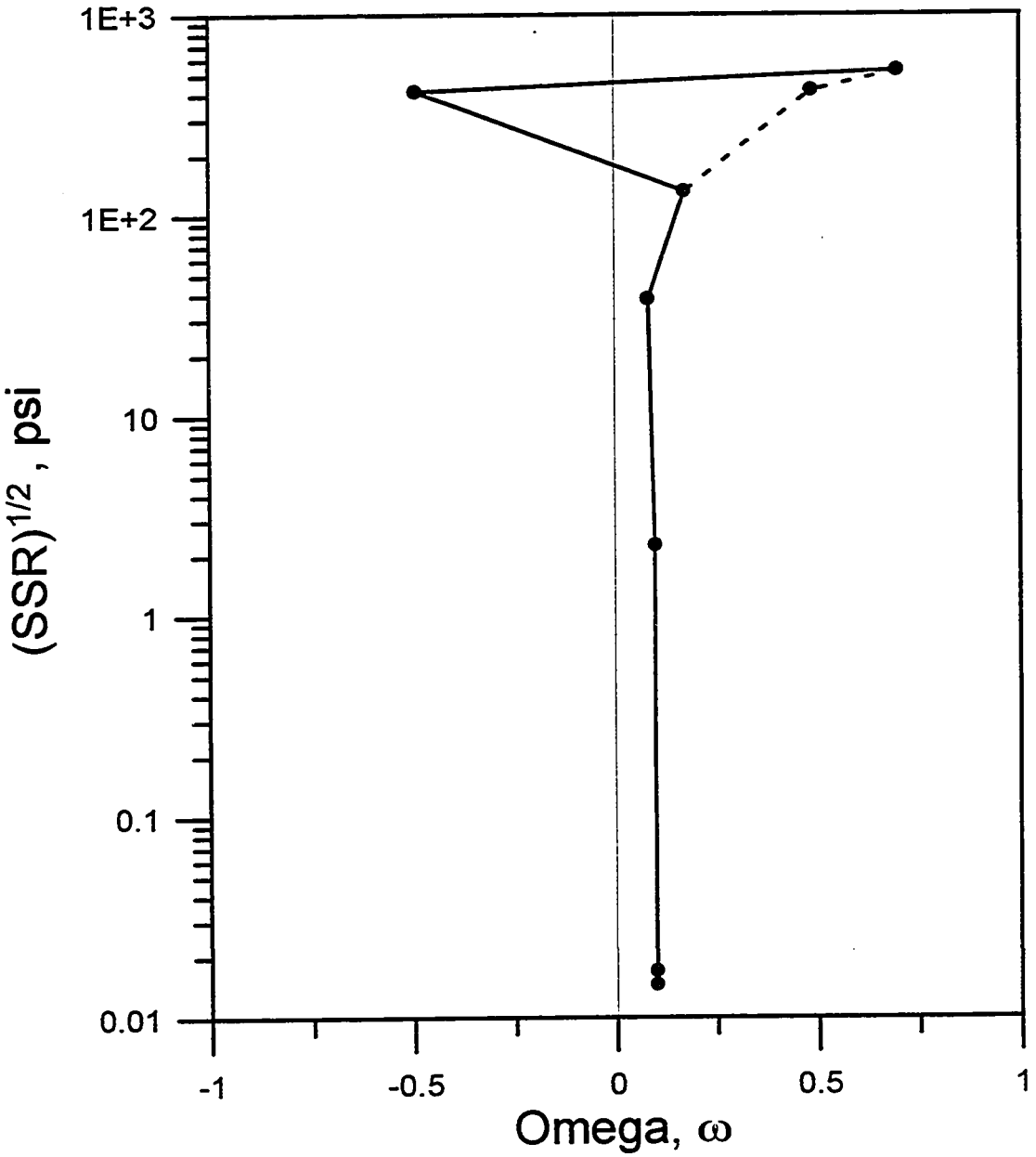

Fig. 2.3.4 - Change in the estimates of ω during the iterative process. Initial estimate ω = 0.70

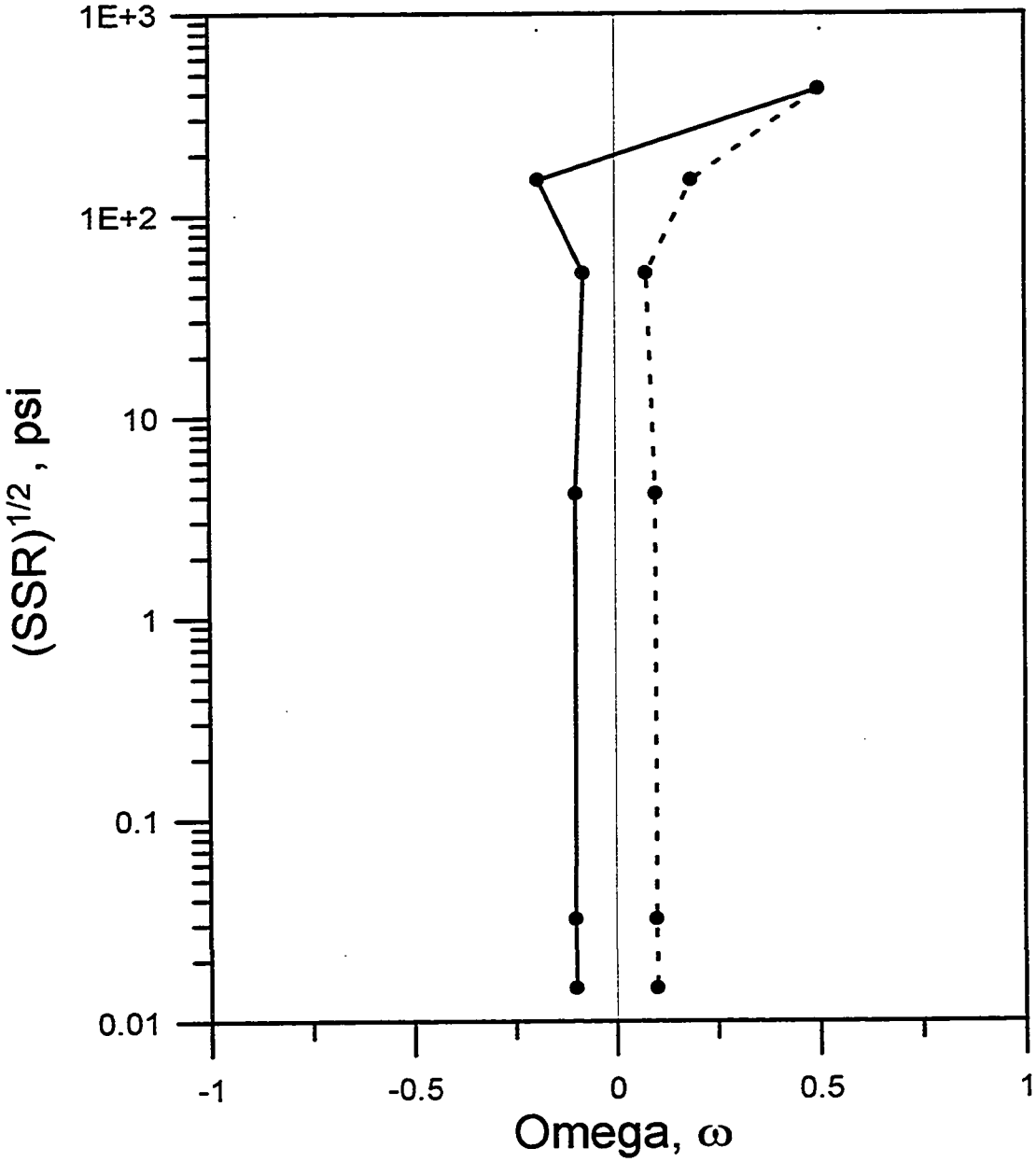

Fig. 2.3.5 - Change in the estimates of ω during the iterative process. Initial estimate ω = 0.50

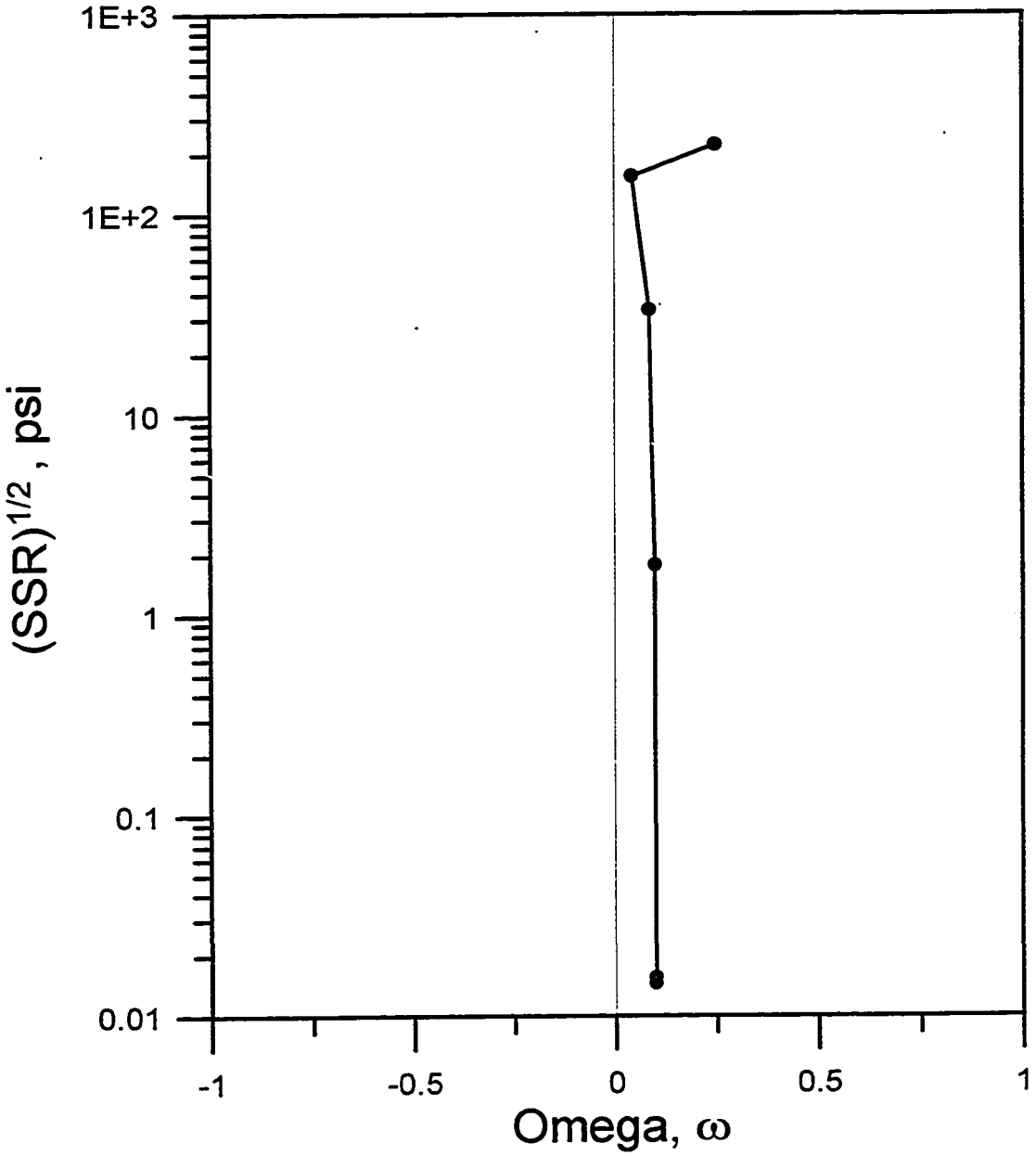

Fig. 2.3.6 - Change in the estimates of ω during the iterative process. Initial estimate ω = 0.25

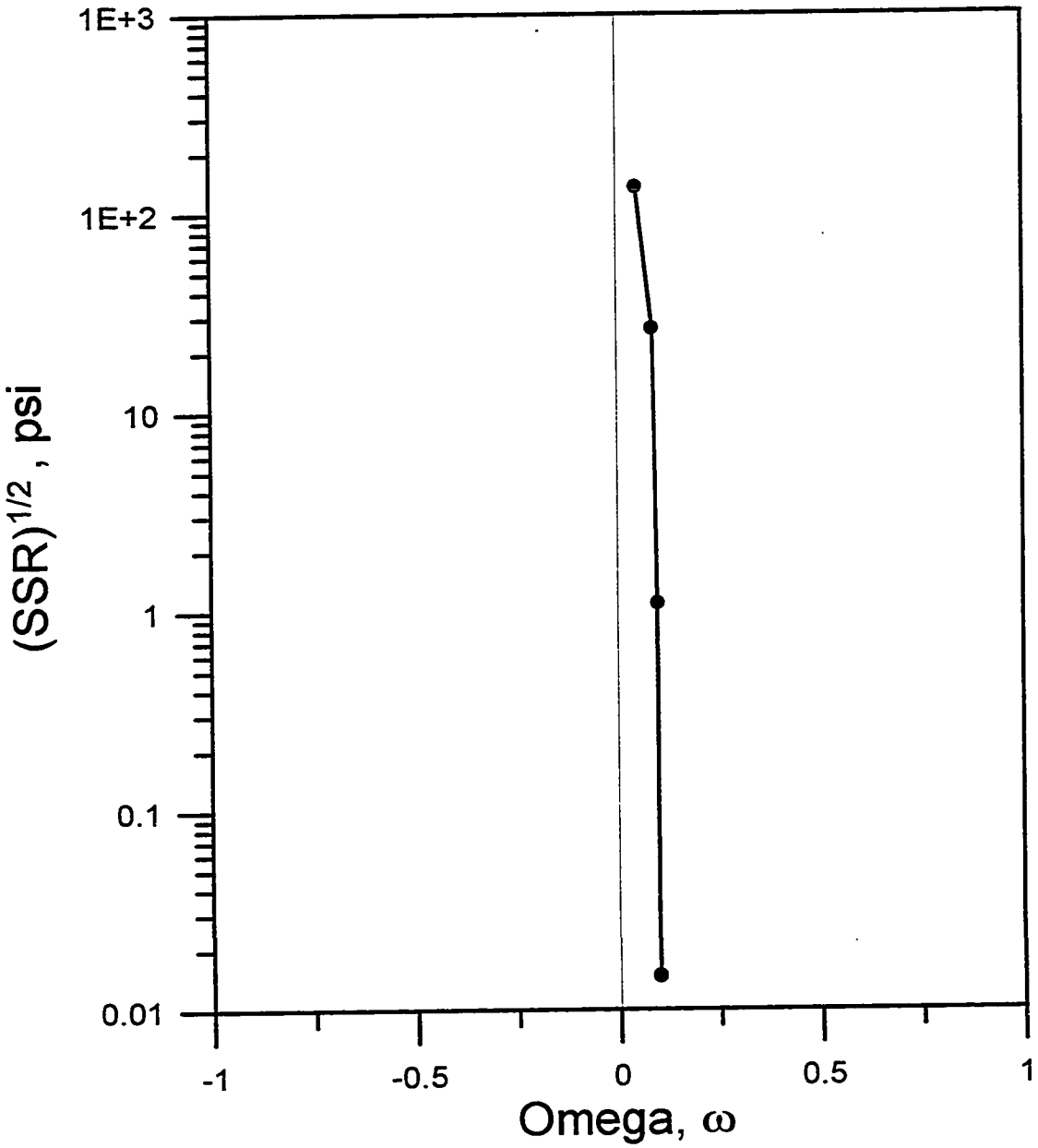

Fig. 2.3.7 - Change in the estimates of ω during the iterative process. Initial estimate ω = 0.05

From our computational experience, we make the following observations:

(i) A constraining algorithm is not necessary for cases in which we start the iterative procedure with very good initial estimates, because, in such cases, the parameter values tend to stay in the feasible region. Even with bad initial estimates, we have seen that a constraining algorithm is normally necessary only for the first few iterations.

(ii) It is expected (especially when the objective function has local minimums) that for the same problem with the same initial estimates, different constraining procedures may follow different paths during the iterative regression procedure; thus, a solution obtained with our imaging algorithm may be different from a solution obtained with some other constraining algorithm, e.g., the penalty function method.

(iii) The way the imaging algorithm was implemented, it is restricted to the simple types of constraints discussed previously.

(iv) Our imaging extension method is easily applied in conjunction with public domain unconstrained regression routines.

(v) In contrast to the penalty function method, which may not avoid a non-physical parameter value in some circumstances, the proposed method guarantees that the estimates satisfy the physical constraints imposed.

2.3.7 Goodness of Fit

In order to have an idea of the model's performance, we may use an absolute or a relative measure of goodness of fit. The standard deviation of the residuals, s, (which is an estimate of σ, the true standard deviation of the error term) is an absolute measure of goodness of fit. This value can be compared against the accuracy of the pressure measurement. One can also use the standard deviation of the residuals (or the value of the sum of the squared residuals) as the criterion for deciding among different regression

models. It seems reasonable to choose the regression model that corresponds to the smallest standard deviation.

The coefficient of determination, R^2, is a relative measure of the fit. This statistic measures the proportion of total variability in the independent variable that is explained by the regression. For LS regression, R^2 is given by[5]

$$R^2 = 1 - \sum_{u=1}^{n} \left[y_u - F(\hat{\alpha}, t_u) \right]^2 \Big/ \sum_{u=1}^{n} y_u^2 \tag{2.3.41}$$

if the model F does not have a constant term, or by

$$R^2 = 1 - \sum_{u=1}^{n} \left[y_u - F(\hat{\alpha}, t_u) \right]^2 \Big/ \sum_{u=1}^{n} (y_u - \bar{y})^2 \tag{2.3.42}$$

if the model F contains a constant term. Here, $\bar{y}$ represents the arithmetic mean of the y_u's . The largest possible value is $R^2 = 1$, which indicates perfect fit. Note that if the model contains only the intercept (i.e., the model function is given by $F(\alpha, t_u) = \alpha$), then $F(\hat{\alpha}, t_u) = \bar{y}$, and $R^2 = 0$ (see Eq. 2.3.42). Thus, R^2 is a measure of the usefulness of the parameters, other than the intercept, in the model.

2.4 Residual-Based Methods for Outlier Detection

One of the drawbacks of the least squares (LS) regression is that it is significantly affected by outliers. There are two ways to overcome this problem. The first approach is to apply robust regression, which tries to devise estimators that are not strongly affected by outliers. An example of robust regression is the least absolute value estimator (LAV) we discussed in the previous section. In this approach, the objective is to accommodate the outliers by reducing their effects on the parameter estimates without actually detecting them (see for example Ref. 8). The second approach is based on the so-called *regression diagnostics*. Here, diagnostics are certain quantities computed from the residuals of the

regression with the purpose of identifying outliers. Once the outliers are identified, they can be removed from the original data set, followed by an LS analysis on the remaining data. One of the advantages of the second approach is that statistical analysis and confidence intervals of the parameter estimates are well defined for LS regression. In the following sections, we will discuss two methods for detecting outliers in a data set.

2.4.1 Outlier Detection Based on Least-Squares Residuals

Many diagnostics for outlier detection are based on the studentized residuals resulting from LS (see Chapter V, Section 5.1). Specifically, the basic hypothesis of a normal error structure, i.e. $\varepsilon_u \approx N(0,\sigma^2)$, is rejected in favor of one which postulates a single outlier if[26]

$$\max_u \left| \frac{r_u}{\sqrt{(1-R_{uu})s^2}} \right| \geq c_\gamma, \tag{2.4.1}$$

where r_u is the uth residual; s^2 is an estimate of the true variance, σ^2; and R_{uu} is the uth diagonal element of the hat matrix,

$$\mathbf{R} = (R_{ij}) = \mathbf{Z}_0\left(\mathbf{Z}_0^T\mathbf{Z}_0\right)^{-1}\mathbf{Z}_0^T. \tag{2.4.2}$$

The definition of $\mathbf{Z}_0$ are presented in Eq. 2.3.5 (also see Chapter V, Section 5.1). The critical value c_γ is such that, under the null hypothesis, Eq. 2.4.1 holds with probability γ. A set of tables of upper bounds for c_γ is presented in Reference 26 for 5% and 1% confidence levels, respectively. The critical values c_γ can also be computed by using the following expression:[26]

$$c_\gamma = \left[\frac{(n-p)F}{n-p-1+F}\right]^{1/2}, \tag{2.4.3}$$

where F is the upper γ/n point of the F-distribution with 1 and $n-p$ degrees of freedom. Thus, if the maximum studentized residual in Eq. 2.4.1 exceeds the critical value computed, for example, by using Eq. 2.4.3, then the corresponding observation is declared an outlier. For more than one outlier, the test is repeated until all the outliers are removed. Since this is a single outlier test, the outliers are detected one at a time. Note that each time an outlier is detected, it should be removed from the data set and a new LS regression is required for computing the new studentized residuals. For example, if we have two outliers in the data we will need three LS regressions.

Rousseeuw *et al.*[5] point out that outlier detection methods based on the LS residuals may have possible problems. The outliers may mask the LS residuals resulting in the failure of the method. Even though the outlier almost always corresponds to the largest residual in the data, it may have a relatively small LS residual in some cases, especially if it is a point far away from the bulk of the observed independent variables (leverage point). For this reason, we use the robust estimates as diagnostics.

2.4.2 Outlier Detection Based on LAV Residuals

Since the residuals computed from a robust estimator is less affected by outliers, they embody powerful information for detecting all outliers in the data. In particular, Rousseeuw *et al.*[5] suggest examination of the standarized residuals r_u / s. Along with the residuals, the scale parameter s (standard deviation) should be estimated in a robust way (i.e., the scale parameter should not be affected by the outliers we want to detect).

Following Ref. 5, first we compute a preliminary estimate of the standard deviation (i.e., the scale) of the LAV residuals via

$$s^0 = 1.4826\left(1+\frac{5}{n-p}\right)\sqrt{\underset{u}{\text{med}}\ r_u^2}\,. \tag{2.4.4}$$

In Eq. 2.4.4, p represents the number of parameters in the model, and "med" denotes the median value of the squared residuals. It is known that 1.4826 $\text{med}_u|r_u|$ is a consistent estimator of the standard deviation if the residuals are normally distributed.[5,6] The term in parentheses in Eq. 2.4.4 is an empirical correction factor which is significant only when n is small. Note that Eq. 2.4.4 is defined in terms of the median so that the presence of outliers will not influence the estimate of the standard deviation of the residuals.

This preliminary scale estimate is then used to compute the final scale estimate, s^*, through the equation

$$s^* = \sqrt{\sum_{u=1}^{n} w_u r_u^2 \Big/ \left(\sum_{u=1}^{n} w_u - p \right)}, \tag{2.4.5}$$

where w_u is determined using the following criteria:

$$w_u = \begin{cases} 1 & \text{if } \left| r_u / s^0 \right| \le 3.0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.4.6}$$

Finally, the outliers (if they exist) are detected based on the values of the standarized residuals $|r_u/s^*|$ by using the following criterion. If $|r_u/s^*|$ is greater than 3.0, the uth data point is considered an outlier, and is removed from the set of observed data points. The bound 3.0 is arbitrary but reasonable because if the errors (residuals) satisfy the Gaussian distribution, the percentage probability of obtaining a residual larger than $3s^*$ (in absolute value) is only 0.27%.

This method and the previous one (based on the LS residuals) were consistent in terms of detecting the same outliers when applied to some field cases we have studied. The advantage of the test based on the LAV residuals is that all outliers (if they exist) are detected at the same time; i.e., only one LAV regression is needed to apply the LAV residual-based test.

2.5 General Regression Analysis Procedure

In the following, we outline the typical procedure we use to obtain optimum parameter estimates:

(i) We perform an LAV regression on the entire set of observed data using an initial estimate, α^0, of the model parameters. We denote the result of this regression by α^1. (In some cases, LAV regression using an arbitrary initial guess requires significantly more iterations than the corresponding LS regression. In those cases, we suggest using an initial LS run to provide good initial parameter estimates for the LAV run.)

ii) We sort the squares of the LAV residuals from the first step in ascending order, and we apply the procedure presented in Section 2.4.2 for detecting the presence of outliers in the data set.

iii) Outliers are then removed from the data set. Suppose we have found n_o outliers.

iv) We perform a final re-weighted LS regression (RLS) on the reduced set of data (i.e., $n-n_o$ values), using the parameters from step (i), i.e., α^1, as the initial estimate. Note that since α^1 is very close to the final solution, this step costs very little to run.

Statistical analyses, and in particular confidence intervals, are based on the results of the final match. In Appendix B and in Section 2.3 we discuss the application of confidence intervals and confidence region of the parameter estimates. By definition, a 95% confidence interval is a range of values within which there is a 95% probability of finding the parameter value.

The above procedure is slightly modified for applications to more complex models such as horizontal wells (see Chapter IV).

CHAPTER III

APPLICATION TO THE CLASSICAL WELLBORE STORAGE AND SKIN MODEL

3.1 Regression Analysis Procedures

Up to this point, our discussion has been general. In this section we focus on the application of regression analysis to the classical wellbore storage and skin model. Here, we restrict our discussion to LS analysis, however, our conclusions apply equally to LAV analysis.

In classical type curve matching, various tools have appeared over the years to assist the analyst in obtaining a unique match. Prior to 1984, analysis was based solely on pressure drops (or pressure rises for buildup). The subsequent use of the pressure derivative[27] reduced the non-uniqueness problems associated with the matching process. Finally, the two-step procedure (using the pressure-pressure derivative ratio) introduced by Onur and Reynolds[28] refined the matching procedure even further. Since nonlinear regression analysis is, in a sense, automated, type-curve matching, we were motivated to investigate whether the use of pressure derivative data enhances the possibility of obtaining a unique set of optimized parameters. Further, assuming that pressure derivative data does contain additional information, it is worthwhile to attempt to delineate methods of incorporating this data into the regression analysis.

Classical type curves are generated based on correlating parametric groups.[29] For example, the classical wellbore storage and skin pressure solutions are function of only three parametric groups, $\{k, k/C, C\exp(2s)/\phi\}$, or, since we assume throughout that porosity is known, $\{k, k/C, C\exp(2s)\}$. In conventional regression analysis, we optimize

the primitive parameters, i.e., $\{k, C, s\}$, but it is not clear whether we should select as our model parameters the primitive or the grouped variables.

In this chapter, we investigate the issues discussed in the preceding two paragraphs by applying various techniques to a set of synthetic well test data into which (Gaussian distributed) random noise was introduced. The data is for a well with skin and wellbore storage effects.

We introduce the following objective functions which will be used in the optimization process. The combined pressure-pressure derivative function is defined as,

$$E_2 = a\sum_{u=1}^{n}\left(\Delta p_u - \Delta p(t_u)\right)^2 + (1-a)\sum_{u=1}^{nd}\left(\Delta p'_u - \Delta p'(t_u)\right)^2, \tag{3.1.1}$$

where a is a constant, with $a = 1$ for regression on pressure only, and $a = 0.5$, for example, for simultaneous regression on pressure and pressure derivative data. Since pressure derivative data are noisier than pressure data, $a = 0.5$ actually places more emphasis to the pressure derivative data. (Note that although a can assume any value between 0 and 1, we only considered $a = 0$, 0.5, and 1 in this work). In Eq. 3.1.1, the subscripted variables denote observed data points, whereas, the unsubscripted variables refer to generated model data; *nd* denotes the number of pressure derivative values and $\Delta p'$ is the logarithmic pressure derivative, i.e., $\Delta p' = \partial \Delta p / \partial \ln t$. Model partial derivatives required by the optimization routine (in terms of primitive or grouped parameters), are provided in Appendix C.

If we analyze the pressure-pressure derivative ratio, the appropriate objective function becomes

$$E_2 = \sum_{u=1}^{nd}\left[\left(\frac{\Delta p}{2\Delta p'}\right)_u - \frac{\Delta p(t_u)}{2\Delta p'(t_u)}\right]^2, \tag{3.1.2}$$

where *nd* is the number of points at which we generate pressure derivative values. In minimizing the objective function of Eq. 3.1.2, we require partial derivatives of the model pressure-pressure derivative ratio with respect to each model parameter; i.e.,

$$\frac{\partial}{\partial \alpha_i}\left(\frac{\Delta p}{2\Delta p'}\right)=\frac{\frac{\partial \Delta p}{\partial \ln t}\frac{\partial \Delta p}{\partial \alpha_i}-\Delta p\frac{\partial^2 \Delta p}{\partial \alpha_i \partial \ln t}}{2\left(\frac{\partial \Delta p}{\partial \ln t}\right)^2}, \tag{3.1.3}$$

for each model parameter, $\alpha_i \in \alpha$. While we do not have closed expressions for the model pressure drop or its partial derivatives (with respect to time or the system parameters), we have equivalent expressions in Laplace space, and we numerically invert these (using the Stehfest algorithm[7]) to obtain desired numerical values at any time. Appendix C provides the relevant working expressions; in particular, we describe how each of the derivative terms on the right side of Eq. 3.1.3 is computed. Appendix D provides details on the FORTRAN code which implements the nonlinear parameter estimation.

Analysis based on minimizing the objective function defined in Eq. 3.1.1 is straightforward. Given a set of pressure and pressure derivative data, we provide initial estimates of the primitive variables, $\left\{k^0, C^0, s^0\right\}$, or the grouped variables, $\left\{k^0, (k/C)^0, (C\exp(2s))^0\right\}$, and iterate (using **LMDER**) until final refined parameter estimates are obtained, (or the scheme fails to converge in 100 iterations).

On the other hand, analysis based on minimizing the objective function of Eq. 3.1.2 is a two-step procedure. For manual type curve matching, use of the pressure-pressure derivative ratio increases the probability of obtaining a unique match by breaking the matching process into two parts, each of which has only one degree of freedom (i.e., the type curve is moved only horizontally or only vertically). In the first step, formation of the pressure-pressure derivative ratio removes permeability as a parameter on the type curve

ordinate. The $\Delta p/(2\Delta p')$ versus time data is then matched by sliding horizontally over the type curves to obtain the best fit for $C_D \exp(2s)$ (from the matched type curve), and k/C from the time match. In the second step, pressure and/or pressure derivative data are matched with the Bourdet *et al.*[27] type curves by fixing the time match, and sliding the data vertically over the type curves until the curve with the appropriate value of $C_D \exp(2s)$ (from the first step) is encountered. Permeability is calculated from the pressure match. By analogy with the manual procedure, the regression procedure is as follows. Select an arbitrary value of permeability, k^0, and initial guesses $\left\{(k/C)^0, (C\exp(2s))^0\right\}$. Minimize the objective function of Eq. 3.1.2 to obtain refined estimates of k/C and $C\exp(2s)$. At this point, we have several options for the second step of the analysis:

i) We can analyze **either** pressure by itself **or** both pressure and pressure derivative data (as in Eq. 3.1.1). We choose to analyze pressure by itself in the second step of the analysis, since it is not as severely affected by noise as the derivative data, and we assume that we have already extracted all usable information contained in the derivative data in the first step.

ii) In the spirit of the original Onur-Reynolds procedure, we can fix k/C and $C\exp(2s)$ at their values from the first step, and regress only on k.

iii) We can fix k/C (from the first step) and regress on $C\exp(2s)$ and k.

iv) We can use as initial estimates for the second step the refined values of k/C and $C\exp(2s)$ from the first step, along with the assumed value of permeability, k^0, and optimize all parameters, (whether primitive or grouped).

In using procedure (ii), it is important to note that even though we regress only on k, as k changes from iteration to iteration, we must also change C and s to maintain the fixed values of k/C and $C\exp(2s)$ found in step (i). Similar comments apply to

procedures (iii) and (iv). We have implemented all of the options outlined above, and have found that options (iii) and (iv) above yield the best results. Intuitively option (iv) seems best if there is difficulty in finding good initial estimates for the one-step procedures, however, all results presented here were generated using option (iii) and we encountered no difficulties.

In the first step of the two-step method, using the objective function of Eq. 3.1.2, we can regress on the primitive parameters $\{k,C,s\}$ instead of k/C and $C\exp(2s)$. With this approach, we found that different initial guesses gave very different solutions for the individual parameters k, C and s; however, all solutions gave good estimates of the correct values of k/C and $C\exp(2s)$. Although we do not use this approach, it has the advantage of eliminating the necessity to compute derivatives with respect to parameter groups.

3.2 Synthetic Examples

In this section, we apply nonlinear regression analysis to three synthetic examples. All examples involve regression using the Laplace space solution for the classical wellbore storage and skin problem.[34]

3.2.1 Synthetic Example #1

For our first example, we consider the synthetic data shown in Fig. 3.2.1, which was generated from the Agarwal *et al.*[34] solution using the data of Table 3.2.1. In analyzing this data, we consider data only in the time interval from 0.002 to 0.3 days. This eliminates any semilog data, and leaves type curve matching as the only viable classical analysis method. We tested five alternative procedures for analyzing the data. Throughout, we denote these Methods as **A**, **B**, **C**, **D** and **E**.

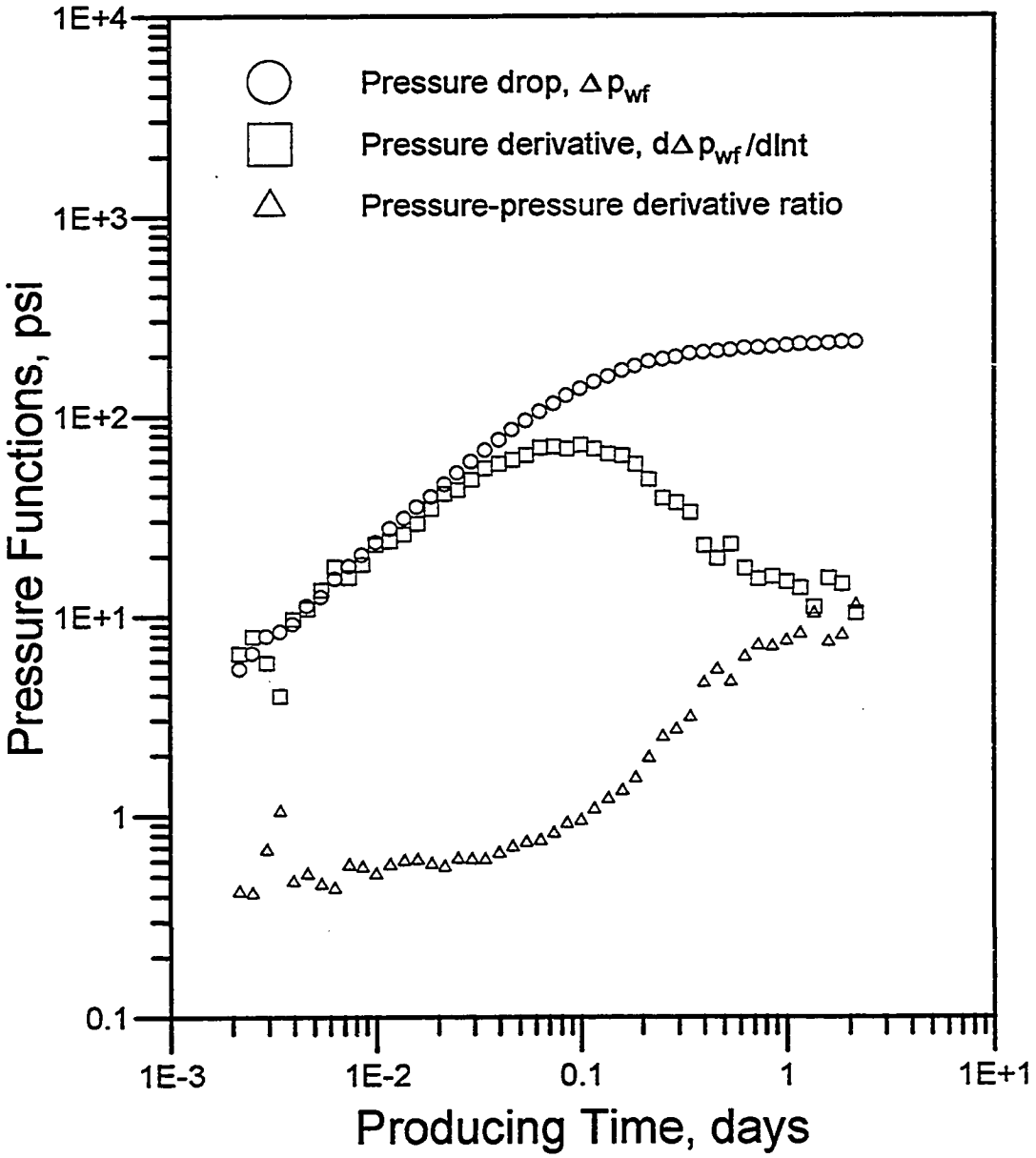

Fig. 3.2.1 - Drawdown synthetic data with noise

Methods **A** and **B** optimize grouped parameters, $\{k, k/C, C\exp(2s)\}$, by regressing, respectively, on pressure, and pressure and pressure-derivative (i.e., using Eq. 3.1.1 with $a = 0.5$). Methods **C** and **D** work with primitive variables, $\{k, C, s\}$, and are analogous to Methods **A** and **B**. Finally, Method **E** refers to the two-step procedure previously discussed.

Table 3.2.1 - Input variables to generate the synthetic data of Examples 1, 2 and 3

Permeability:	$k = 15$ md
Skin Factor:	$s = 2.6$
Wellbore Storage Coefficient:	$C = 0.02$ bbl/psi
Porosity:	$\phi = 0.25$
Thickness:	$h = 20$ ft
Viscosity:	$\mu = 1.0$ cp
Wellbore Radius:	$r_w = 0.25$ ft
Total Compressibility:	$c_t = 1\times10^{-5}$ psi^{-1}
Flow Rate:	$qB = 50$ bbl/day
Initial Pressure:	$p_i = 4000$ psi
STD Deviation of $N(0,s)$ Noise:	$\sigma = 0.35$ psi
# of Data Points per Cycle:	15

As initial estimates for the various parameters, we have used all 32 possible combinations of the following: $k = 5$, 10, 20 and 30 md; $s = 0$, 1.5, 4 and 8; and $C = 0.005$ and 0.05 bbl/psi. (Note from Table 3.2.1, that the actual values are $k = 15$ md, $s = 2.6$ and $C = 0.02$ bbl/psi.) Table 3.2.2 summarizes the results of the various runs. The following points are important to note:

Table 3.2.2 - Summary of the results for Synthetic Example #1
(time interval 0.002 to 0.3 days)

Regression type (*)	Convergence (%)	avg. # of iterations	k (md)	C (bbl/psi)	s
A	100	62	13.82	0.0199	1.96
B	88	33	14.82	0.0200	2.50
C	100	14	13.82	0.0199	1.96
D	97	12	14.82	0.0200	2.50
E	100	29	14.86	0.0200	2.52

(*) Legend:
A - Combined parameters, pressure only
B - Combined parameters, pressure and pressure derivative
C - Primitive parameters, pressure only
D - Primitive parameters, pressure and pressure derivative
E - Two-step procedure

1) When a particular method converged, it always converged to the unique final answers shown in columns 4-6 of Table 3.2.2, regardless of the initial guess.

2) Results obtained by regressing on pressure, or on pressure and pressure derivative simultaneously, were independent of whether primitive or grouped parameters were used. However, regressing on combined parameters invariably required more iterations.

3) Comparing the results of analyzing pressure only, with those obtained using both pressure and pressure derivative data, we see that introduction of the pressure derivative data (using Eq. 3.1.1) may actually decrease the chance of obtaining convergence; compare results for methods **A** and **B** and results for methods **C** and **D** in Table 3.2.2. When the calculations converged, however,

the results from the simultaneous analysis of pressure and pressure derivative data were closer to the original input values; however, this result does not generalize, i.e., we obtained different comparison for other set of synthetic data.

4) The two-step analysis (based on use of the pressure-pressure derivative ratio in the first step) always converged regardless of the initial guess. This suggests that the manner in which the pressure derivative data is incorporated in this analysis is better than simply summing the objective functions, as in Eq. 3.1.1. Note that the number of iterations in this case cannot be directly compared with the other methods, since each iteration costs less.

Figure 3.2.2 shows the contour map of the LS objective function as a function of k/C and $C\exp(2s)$. Figure 3.2.3 is a similar plot, for a fixed $C = 0.02$ bbl/psi, but here, the axes represent the primitive parameters k and s. Clearly, close to the solution, the model that uses primitive variables k and s as parameters (Fig. 3.2.3) behaves linearly, in contrast to the model of Figure 3.2.2, which utilizes the combined variables as the parameters. This fact explains why regressing on combined parameters normally requires more iterations.

3.2.2 Synthetic Example #2

We further investigated the behavior of each of the methods by analyzing a smaller range of the same set of synthetic data used in Example #1; here, we removed the wellbore storage dominated data, and analyzed data in the interval from 0.02 to 0.3 days (see Fig. 3.2.1). In this case, our initial guesses were all 24 possible combinations of the following parameter estimates: $k = 5$, and 30 md; $s = 0$, 4 and 8; and $C = 0.0015$ and 0.10 bbl/psi. The results are summarized in Table 3.2.3.

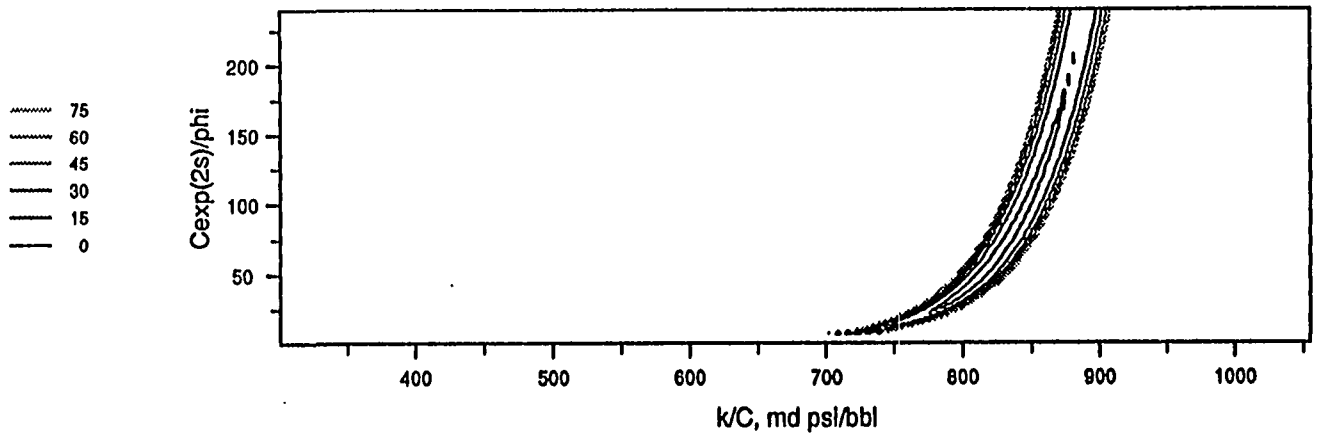

Fig. 3.2.2 - LS objective function contour map at C=0.02 bbl/psi
Synthetic Examples 1 and 2

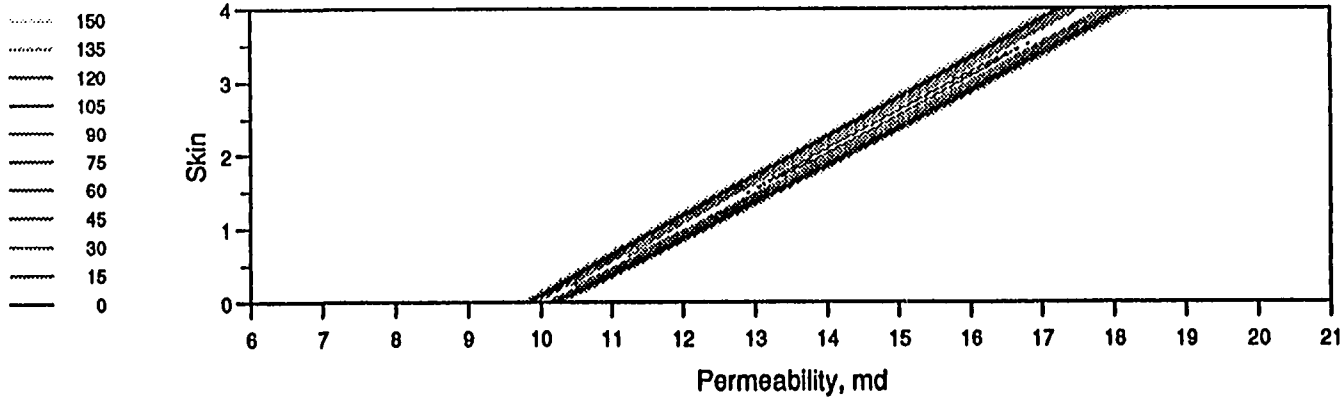

Fig. 3.2.3 - LS objective function contour map at C=0.02 bbl/psi
Synthetic Examples 1 and 2

Table 3.2.3 - Summary of the results for Synthetic Example #2
(time interval 0.02 to 0.3 days)

Regression type (*)	Convergence (%)	avg. # of iterations	k (md)	C (bbl/psi)	s
A	75	53	14.71	0.0200	2.45
B	75	35	15.07	0.0200	2.64
C	75	14	14.71	0.0200	2.45
D	75	12	15.07	0.0200	2.64
E	100	30	14.93	0.0200	2.57

(*) Legend:
A - Combined parameters, pressure only
B - Combined parameters, pressure and pressure derivative
C - Primitive parameters, pressure only
D - Primitive parameters, pressure and pressure derivative
E - Two-step procedure

Note that whereas the two-step procedure always converged to the correct solution, all the remaining methods failed to converge for 25% of the initial guesses. The individual initial guesses, for which each method failed to converge, differed from method to method. The study on this synthetic data set suggests that the two-step procedure enhances our chance of converging to good estimates of the parameters. There appears to be no advantage to using grouped parameters in preference to the primitive variables, in fact, it is apparent we should use primitive parameters when using the classical one-step method as well as in the second step of the two-step method. Finally, analysis of the sum of pressure and pressure derivative data may actually worsen the chance of converging to an acceptable solution (see Table 3.2.2, methods **B** and **D**). For the remainder of this work, we will focus our attention on analysis of pressure or pressure derivative using primitive parameters, or the two-step method.

3.2.3 Synthetic Example #3

In this example, we investigate the effects of the randomness of the measurement errors on the parameter estimates. The input data for this example are the same as shown in Table 3.2.1, except that here the standard deviation of the normally distributed error (i.e., the noise) is $\sigma = 0.5$ psi. We generated four different data sets corresponding to four different seeds in the algorithm for generating the random sequence. Thus, each set of data is a different *realization* of the same pressure transient test. For each realization, we interpret the test twice, corresponding to two different spans of time. In the first test, we use data points (15 points per log-cycle) in the time interval 2.15×10^{-3} to 0.215 days (which eliminates all semilog data); and in the second one, we use all data in the time interval 2.15×10^{-3} to 15.85 days. The initial guess for all runs are $k = 30$ md, $C = 0.05$ bbl/psi, and $s = 8$.

Table 3.2.4 summarizes the results for the various tests. In Table 3.2.4, the $\pm$ values associated with each parameter denote the 95-percent confidence intervals (see Appendix B and Section 2.3.3). For example, the first entry in column 3 gives k as 16.81 ± 4.47. This means that there is a 95% chance that the true k lies in the interval $(16.81-4.47) < k < (16.81+4.47)$. The following points should be noted:

1) The standard deviation of the residuals calculated from the regression fit (column 6 of Table 3.2.4) is, as expected, a good estimate of the true standard deviation of the errors (0.5 psi in this case).

2) No run evidenced convergence problems; in fact, the number of iterations required to obtain convergence is about the same for all cases.

3) Since all tests exhibit about one log-cycle of wellbore dominated flow period, the estimate of the wellbore storage coefficient was very accurate in all cases.

Table 3.2.4 - Summary of the results for Synthetic Example #3
(Same input data as in Table 3.2.1 except that $\sigma = 0.5$ psi)

Seed number	Time span (days)	Parameter Estimates			Std. deviation (psi)	Correlation Coefficients	Number of iterations
		k (md)	C (bbl/psi)	s			
−34	2.15×10^{-3} to 2.15×10^{-1}	16.81 ± 4.47	0.0200 ± 2.8×10^{-4}	3.548 ± 2.39	0.4810	$k, s = 0.9998$ $k, C = 0.8975$ $C, s = 0.9040$	11
−34	2.15×10^{-3} to 15.85	15.01 ± 0.185	0.0199 ± 1.1×10^{-4}	2.603 ± 0.117	0.5128	$k, s = 0.9973$ $k, C = 0.5900$ $C, s = 0.6066$	10
−48	2.15×10^{-3} to 2.15×10^{-1}	13.99 ± 3.63	0.0199 ± 3.5×10^{-4}	2.061 ± 1.92	0.5586	$k, s = 0.9997$ $k, C = 0.9027$ $C, s = 0.9105$	12
−48	2.15×10^{-3} to 15.85	15.08 ± 0.209	0.0200 ± 1.2×10^{-4}	2.650 ± 0.133	0.5754	$k, s = 0.9973$ $k, C = 0.5898$ $C, s = 0.6063$	10
−21	2.15×10^{-3} to 2.15×10^{-1}	15.33 ± 4.94	0.0200 ± 3.9×10^{-4}	2.749 ± 2.62	0.6320	$k, s = 0.9998$ $k, C = 0.9002$ $C, s = 0.9073$	12
−21	2.15×10^{-3} to 15.85	15.13 ± 0.219	0.0201 ± 1.3×10^{-4}	2.689 ± 0.139	0.5974	$k, s = 0.9974$ $k, C = 0.5898$ $C, s = 0.6062$	10
−100	2.15×10^{-3} to 2.15×10^{-1}	14.22 ± 3.70	0.0199 ± 3.4×10^{-4}	2.196 ± 1.97	0.5497	$k, s = 0.9998$ $k, C = 0.9002$ $C, s = 0.9073$	12
−100	2.15×10^{-3} to 15.85	14.99 ± 0.171	0.0199 ± 1.0×10^{-4}	2.592 ± 0.108	0.4760	$k, s = 0.9974$ $k, C = 0.5898$ $C, s = 0.6062$	10

4) The values of the correlation coefficients are independent of the seed (realization). However, they are greatly dependent on the span of the data. The higher correlation coefficients obtained in the short tests had no significant impact on the convergence of the minimization routine.

5) Practically, all four tests with the longer data span converged to the same solution, independent of the seed, and in these cases, the confidence intervals were also very small.

6) Except for the wellbore storage coefficient, the parameter estimates obtained in the short tests differ significantly from realization to realization. The estimates of skin factor varied from 2.06 to 3.55 which corresponds to more than 70% difference. The percentage difference for the permeability estimates is about 20%. The short tests give significantly larger confidence intervals for k and s.

7) For the tests with the same span in the data, all parameter estimates, regardless of the realization, lie in all confidence intervals generated for that parameter; i.e., as expected, the confidence intervals capture the uncertainty in the parameter estimates.

In the next sections, we turn our attention to the analysis of actual field data, and demonstrate the applicability of the methods to drawdown, buildup and slug-test data.

3.3 Field Examples #1

Our first field example is the drawdown data of Ref. 35; Figs. 3.3.1 and 3..3.2 show the log-log plot and the semilog plot of the data, respectively. In Fig. 3.3.1, note that the flattening derivative data towards the end of the drawdown (i.e., for $t > 20$ hr),

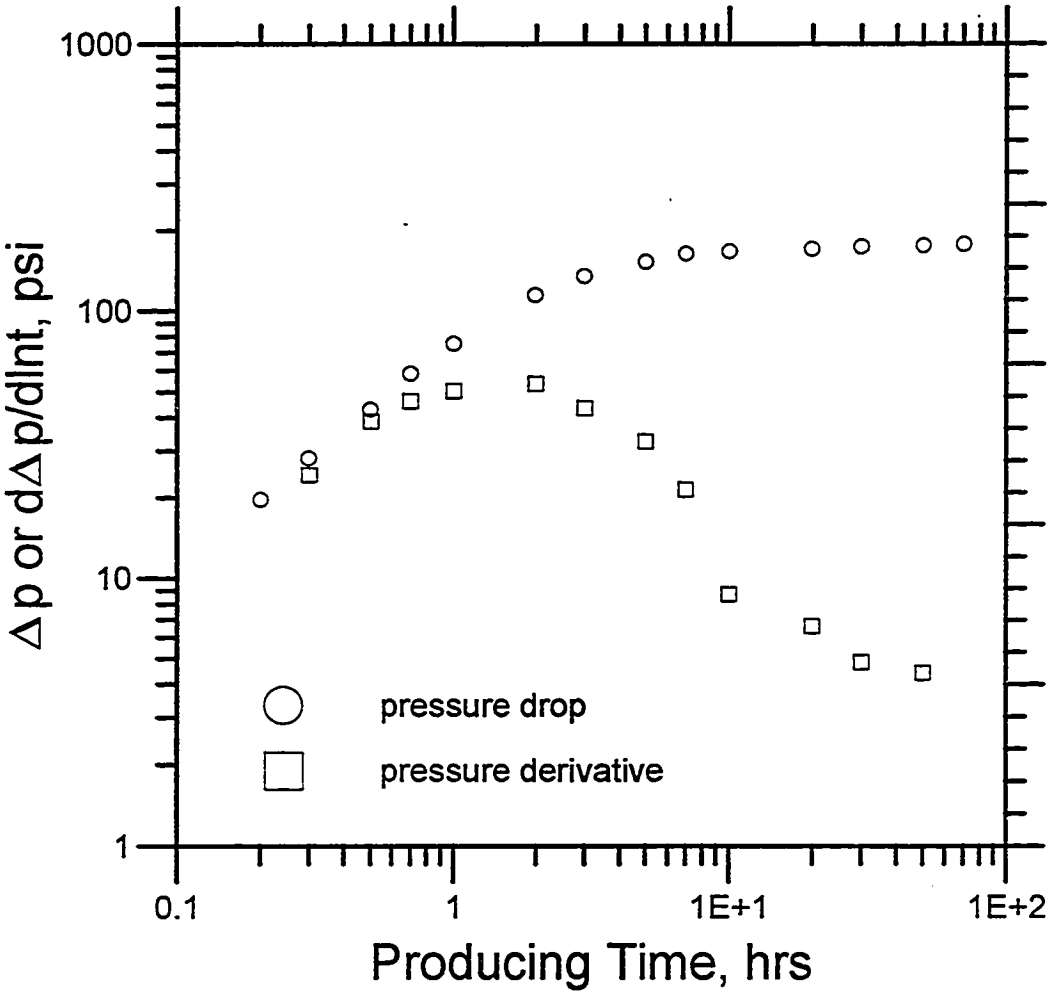

Fig. 3.3.1 - Log-log plot - Field Example #1

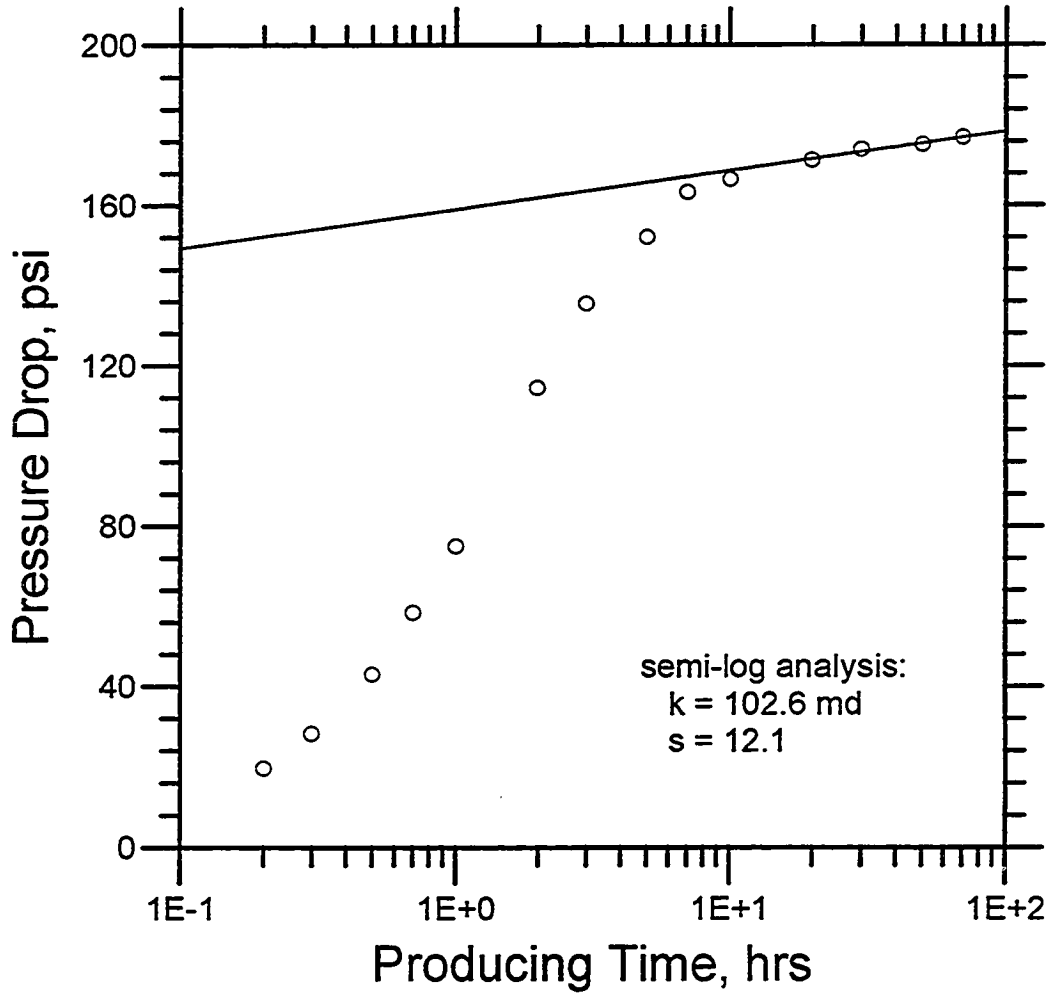

Fig. 3.3.2 - Conventional semilog analysis
Field Example #1

indicates that the last two or three data points may be semilog data. Applying standard semilog analysis to the last three data points, we obtain estimates of $s = 12.1$ and $k = 102.6$ md (see Fig. 3.3.2).

We begin by illustrating the performance of our LAV regression implementation applied to the pressure data. Table 3.3.1 shows the results of a sequence of regression runs which minimize the objective function of Eq. 2.2.4 with successively smaller values of ε; see Chapter II, Section 2.2.

Table 3.3.1 - L_1-Norm minimization applied to all data points of Field Example #1 (Tolerance = 10^{-5} and Stehfest # = 10)

Iteration l	ε_l (psi^2)	Parameter Estimates			# of Iterations	SSR (psi^2)	SAR (psi)
		k (md)	C (bbl/psi)	s			
1	10^{-2}	98.52	0.08685	11.47	28	13.337	8.3834
2	10^{-4}	99.84	0.08705	11.72	22	13.638	8.3365
3	10^{-6}	100.01	0.08710	11.75	19	13.671	8.3309
4	10^{-8}	100.04	0.08711	11.76	8	13.686	8.3274
5	10^{-10}	100.05	0.08711	11.76	5	13.691	8.3268
6	10^{-12}	100.05	0.08711	11.76	2	13.690	8.3267
single	10^{-10}	100.05	0.08711	11.76	37	13.691	8.3268

For our first run (with $\varepsilon_1 = 10^{-2}$), the initial estimates are k = 150 md, C = .001 bbl/psi and $s = 0$. We used the refined parameter estimates from this run as the initial guess for iteration 2, with $\varepsilon_2 = 10^{-4}$, and continued in this fashion until we reduced ε_l to 10^{-12}.

Note the final parameter estimates, $k = 100$ md, $C = 0.087$ bbl/psi and $s = 11.8$, are essentially independent of ε_l for $\varepsilon_l < 10^{-6}$. Further, a single run with a small value of ε_l, e.g., $\varepsilon_l = 10^{-10}$, and the initial guess used in iteration 1 also converges to the same answer in a single step. (For this reason, we sometimes perform only a single regression with a small value of ε when applying the LAV method.) Columns 7 and 8, respectively, of Table 3.3.1 show values of the sum of the squares of the residuals (SSR) and the sum of the absolute values of the residuals (SAR) at the end of each run. Note that iteration 6 (i.e., the solution) corresponds to the minimum value of SAR, not SSR.

Table 3.3.2 summarizes the results of a variety of runs on the data. The first row repeats the final results of Table 3.3.1, however, these results were obtained from a single run with $\varepsilon_l = 10^{-10}$, and the same initial guess used to obtain the results in the first row of Table 3.3.1. Based on our LAV outlier detection analysis, one point in the data set, (pressure data point #8 from the left in Fig. 3.3.1), was identified as an outlier. Fig. 3.3.3a shows the scaled residual plot (i.e., r_i/s^* versus time) for this case. It is clear from Fig. 3.3.3a that data point #8 is an outlier. The outlier was removed from the data set, and the LAV analysis was repeated. Fig. 3.3.3b shows the scaled residual plot for this second LAV regression, as expected no more outliers were found. The final parameter estimates were essentially unchanged since LAV is a robust method, (see columns 2-4 of the second row). The results for this match are shown in Fig. 3.3.4.

Results for LS analysis on pressure data with and without the detected outlier are shown in rows 3 and 4 of Table 3.3.2. Note that upon removal of the outlier, LS and LAV results are almost identical. Also, note that confidence intervals for the final parameter estimates are improved when the outlier is removed.

For comparison purposes, we also applied to this data the LS outlier detection procedure presented in Section 2.4.1. The outlier detection based on the LS residuals also indicates point #8 as the single outlier in the data set. The value of the studentized residual of point #8 is -2.820 (the largest one in absolute value), which is larger than the critical

Table 3.3.2 - Summary of the results for Field Example #1
(Tolerance = 10^{-5} and Stehfest # = 10)

Type of Regression	Parameter Estimates with Respective Confidence Intervals			# of iterations	SSR (psi^2)	SAR (psi)	Detected outlier and respective scaled residual
	k (md)	C (bbl/psi)	s				
LAV (pressure only) all points	100.1	0.0871	11.8	37	13.7	8.33	data #8 (–5.6)
LAV (pressure only) all points except #8	99.9	0.0867	11.7	29	3.21	5.05	none
LS (pressure only) all points	89.3 ± 16.2%	0.0863 ± 2.6%	9.68 ± 28.6%	11	11.2	–	data #8 (–2.820)
LS (pressure only) all points except #8	99.1 ± 11.1%	0.0868 ± 1.5%	11.6 ± 18.3%	12	3.08	–	none
LAV (press. and deriv.) all points except #1 and 14	83.0	0.0858	8.51	30	110.7	35.2	deriv. #7 (3.4) deriv. #8 (3.5)
LAV (two-step regression) all points except #1 and 14	103.0	0.0873	12.3	82 + 16	14.2	7.98	data #8 (–5.4)

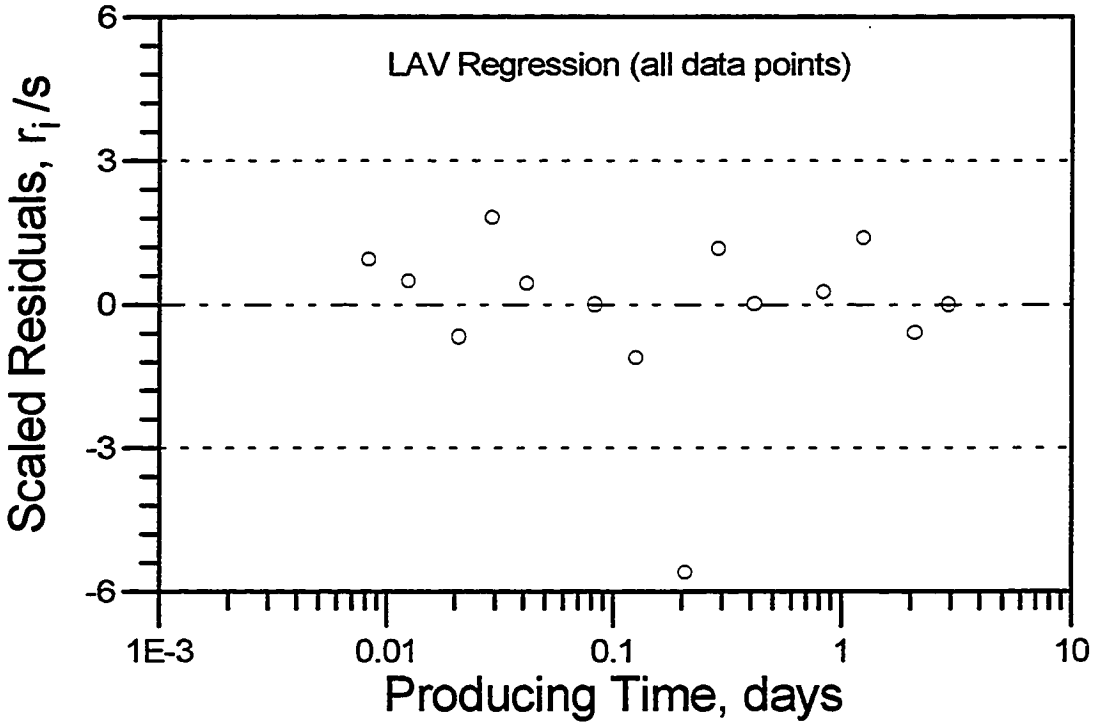

Fig. 3.3.3a - Scaled residual plot - Field Example #1

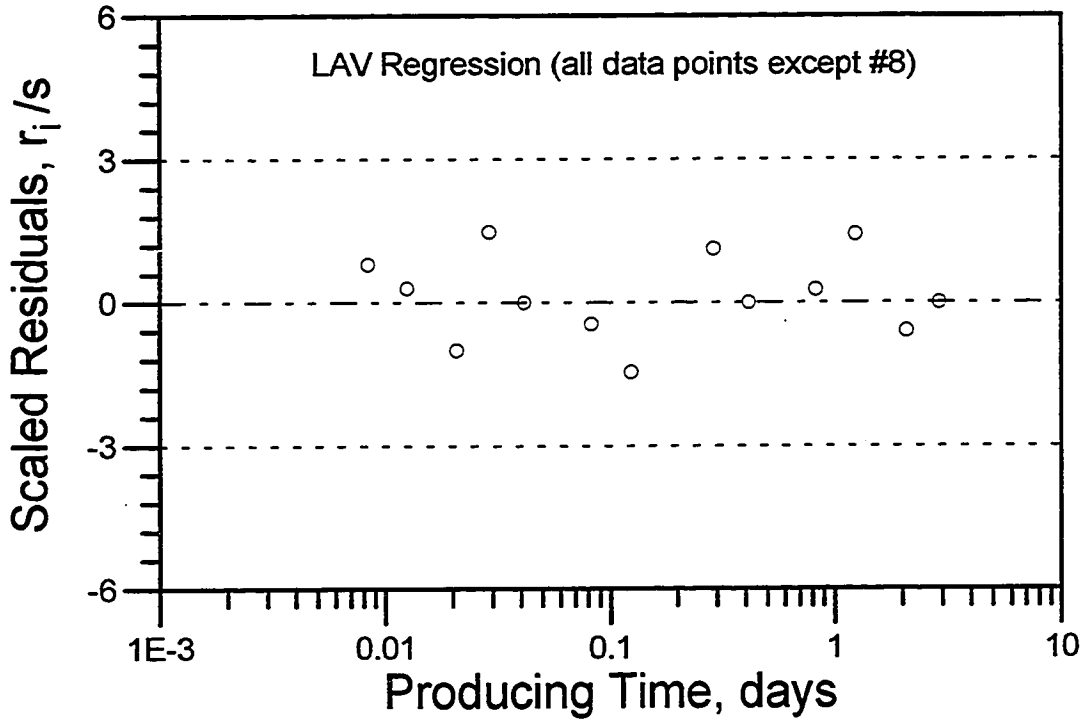

Fig. 3.3.3b - Scaled residual plot - Field Example #1

value $c_{5\%}$ given by Eq. 2.4.3, $c_{5\%} = 2.55$ in this case. The second LS regression (without data point #8) does not indicate any outliers, since the largest studentized residual in absolute value, 1.817 (data point #4), does not exceed the critical value $c_{5\%} = 2.50$.

For the other results shown in Table 3.3.2, both pressure and pressure derivative data (Eq. 3.1.1 with $a = 0.5$) were used. The first and last pressure data points (#1 and #14) were removed because our current regression routine requires equal numbers of pressure and pressure derivative points, and no pressure derivative data were available at these times. Row 5 shows the results obtained when pressure and pressure derivative data were analyzed simultaneously using the LAV approach. Note that the results for this run are significantly different from those obtained with semilog analysis. On the other hand, the final row of Table 3.3.2 shows that the two-step procedure yields results which agree closely with the results obtained by application of LAV analysis to all pressure data, as well as with results obtained from semilog analysis. The matches for the LAV regressions are shown in Figures 3.3.4 through 3.3.6.

Figures 3.3.7 through 3.3.12 show the confidence intervals (as the sides of the dashed rectangles) and the joint 2-dimensional elliptical confidence region for each pair of parameters. Note that the axes are normalized using the corresponding parameter estimates obtained from the regression. We computed the elliptical confidence region using Eq. 2.3.9, which assumes that the model is approximately linear in a neighborhood of the minimum. The correlation coefficients for each pair of parameters (Eq. 2.3.14) are presented in Table 3.3.3.

The 2-dimensional confidence regions shown in Figs. 3.3.7 through 3.3.12 were calculated by assuming that the third parameter was fixed at the estimated value obtained from the regression. For example, the confidence region for parameters C and k, assumes that there is no uncertainty (i.e., constant value) in the skin factor, s. The correct skin value is assumed to be the result obtained from the regression, in this case, $s = 9.68$.

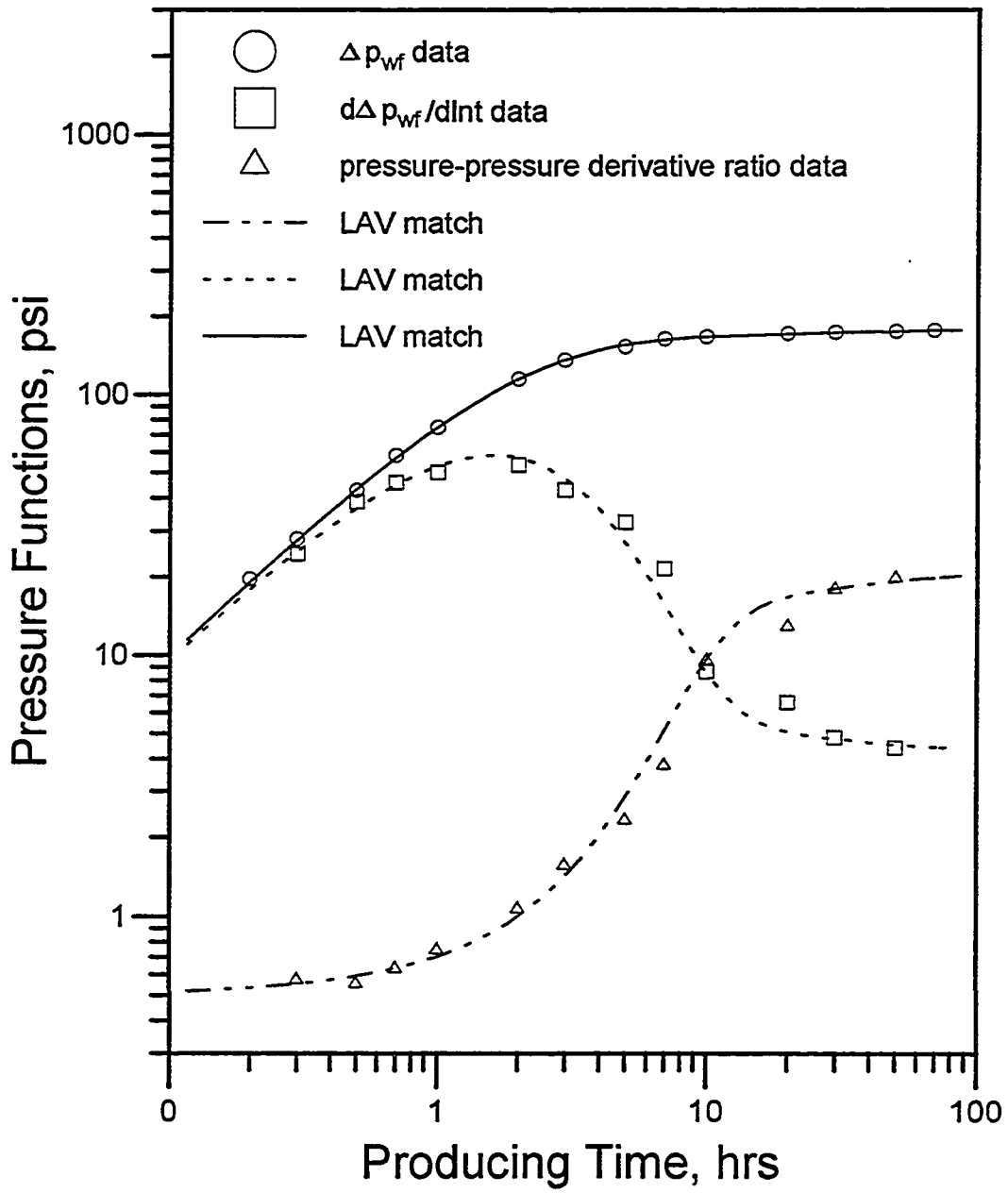

Fig. 3.3.4 - Results from the LAV method applied to the pressure-pressure derivative ratio data Field Example #1

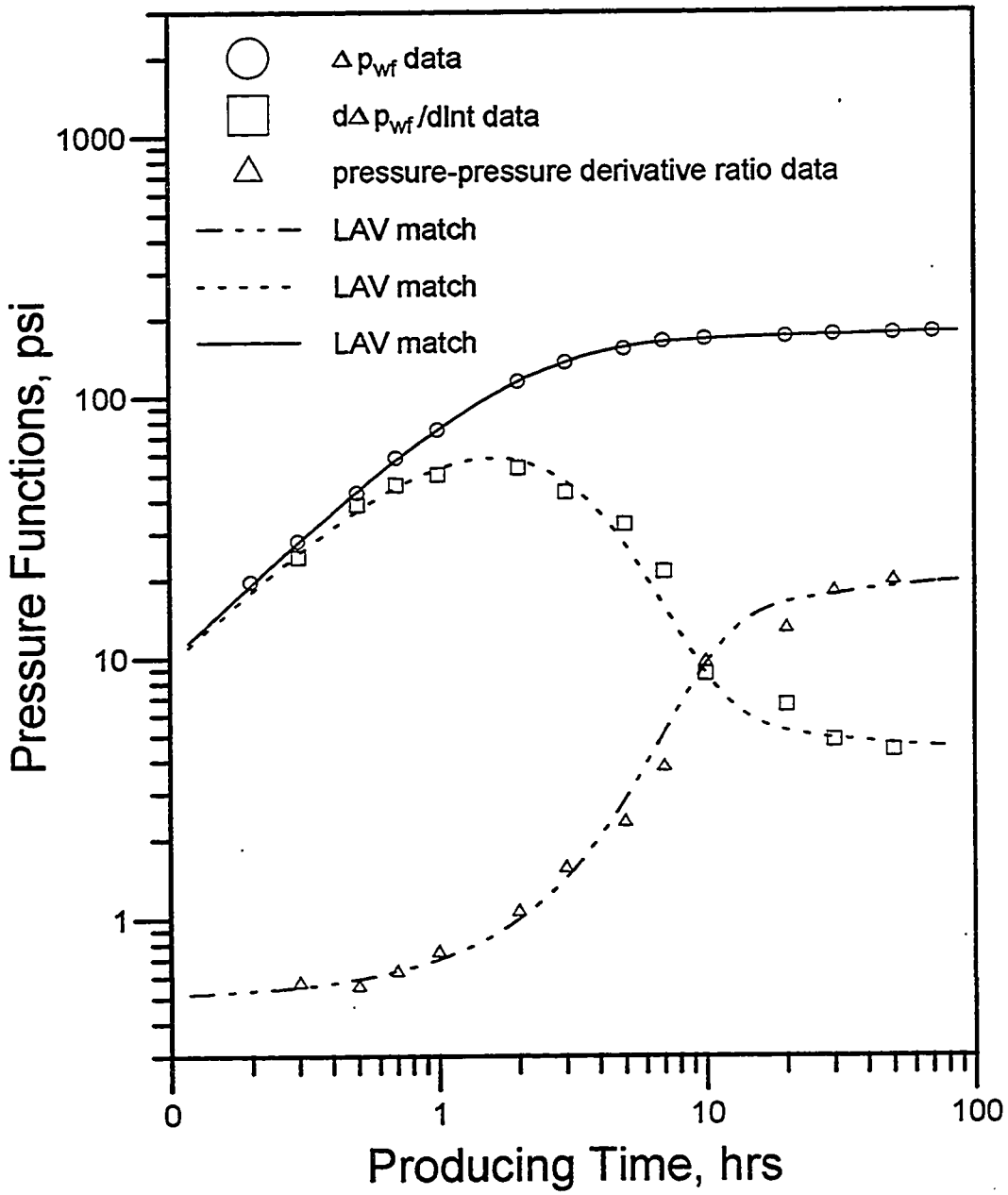

Fig. 3.3.5 - Results from the LAV method applied to the pressure data only - Field Example #1

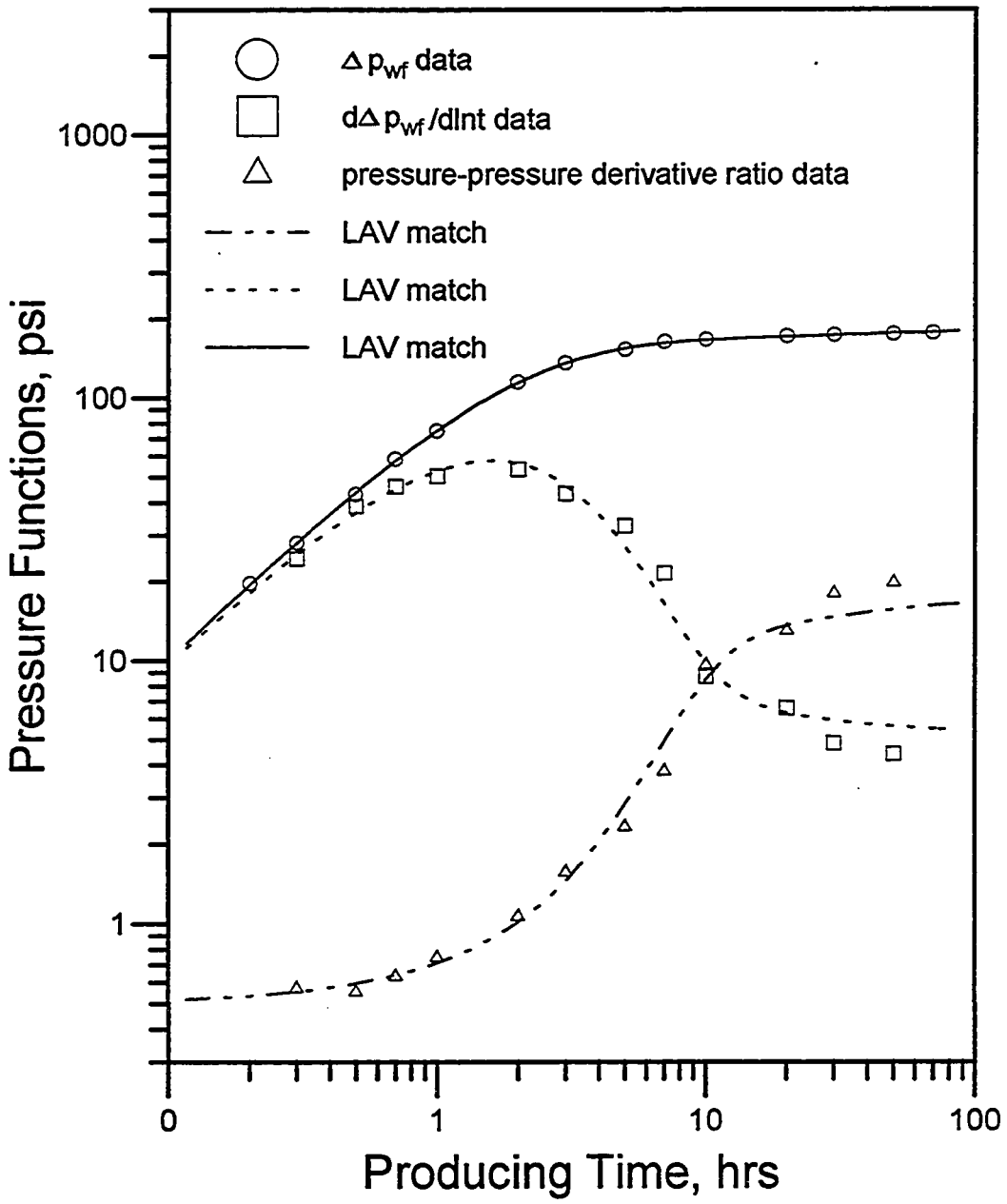

Fig. 3.3.6 - Results from the LAV method applied to the pressure and pressure derivative data Field Example #1

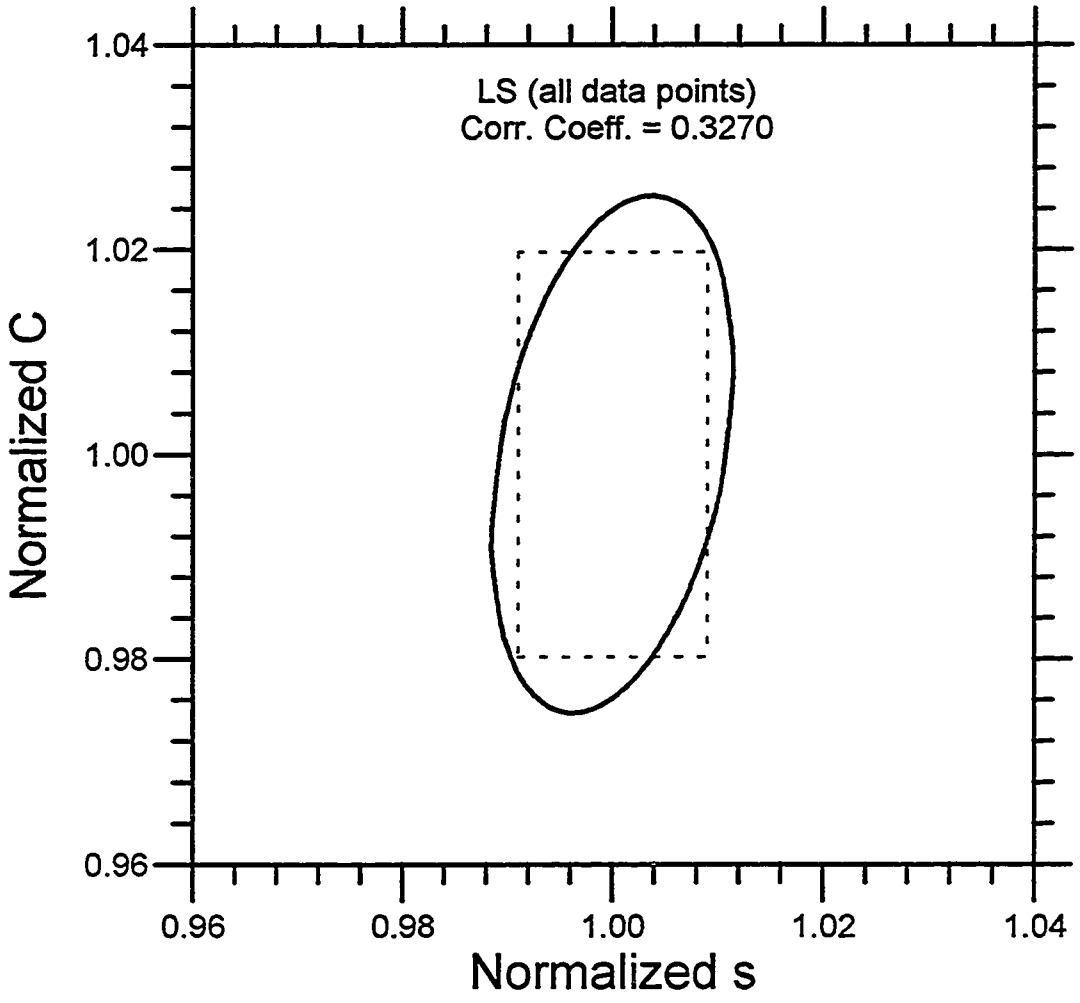

Fig. 3.3.7 - 95% confidence region versus
95% confidence intervals - Field Example #1

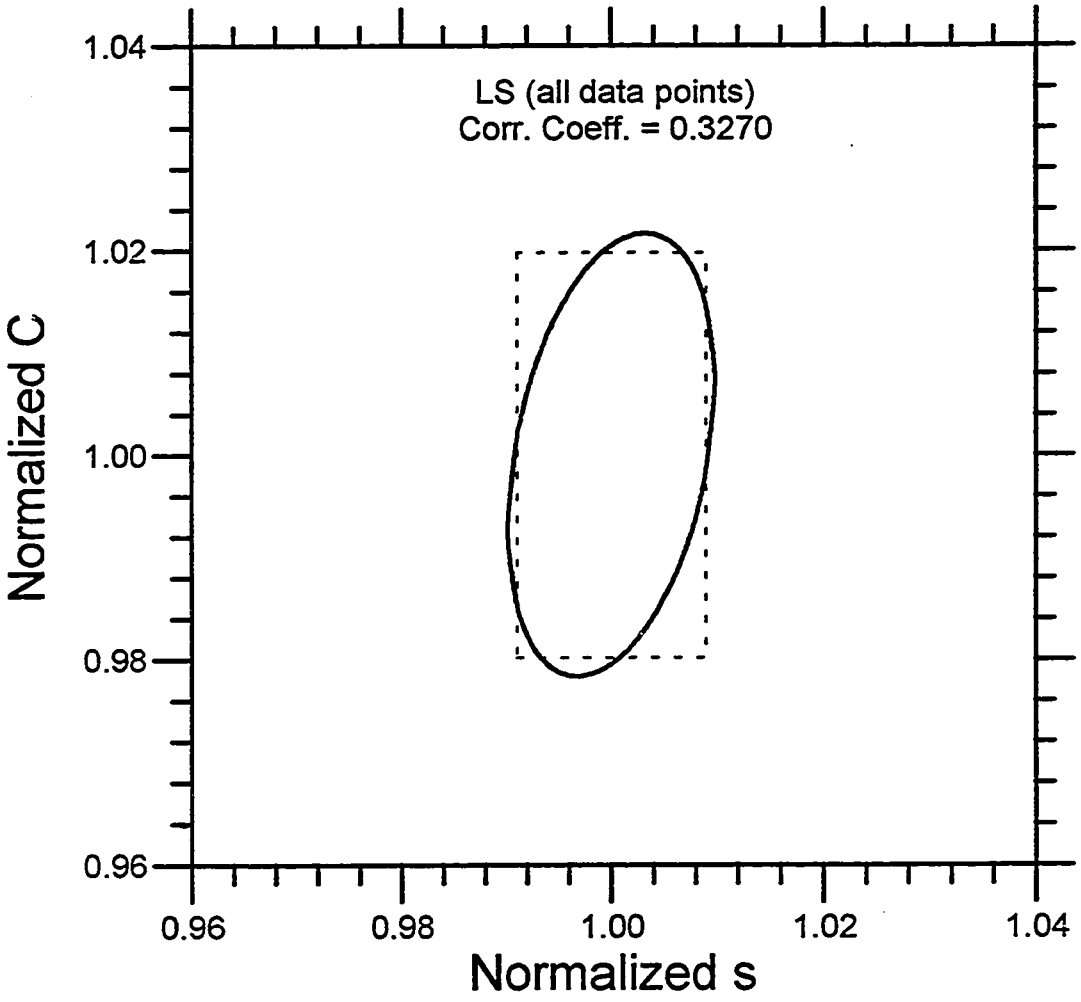

Fig. 3.3.8 - 90.25% confidence region versus 95% confidence intervals - Field Example #1

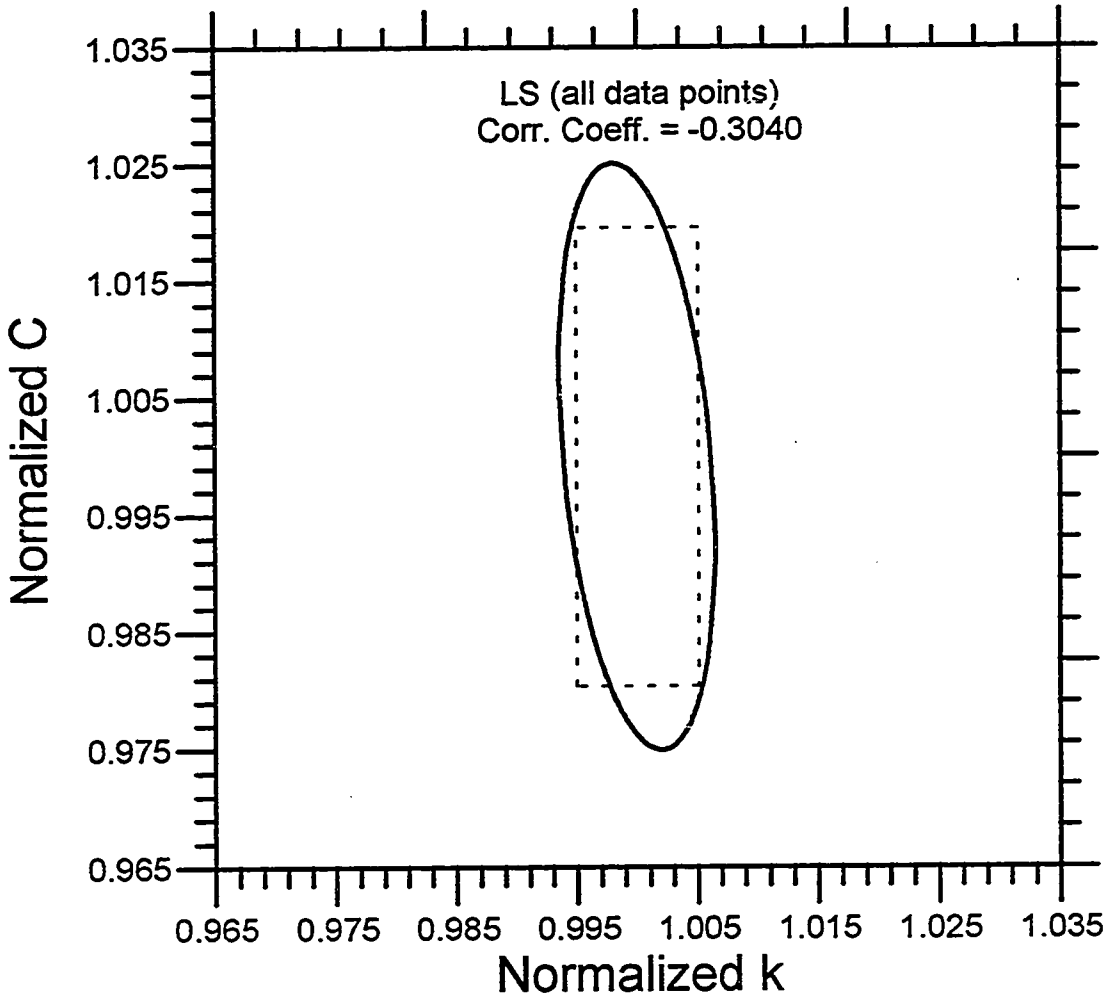

Fig. 3.3.9 - 95% confidence region versus
95% confidence intervals - Field Example #1

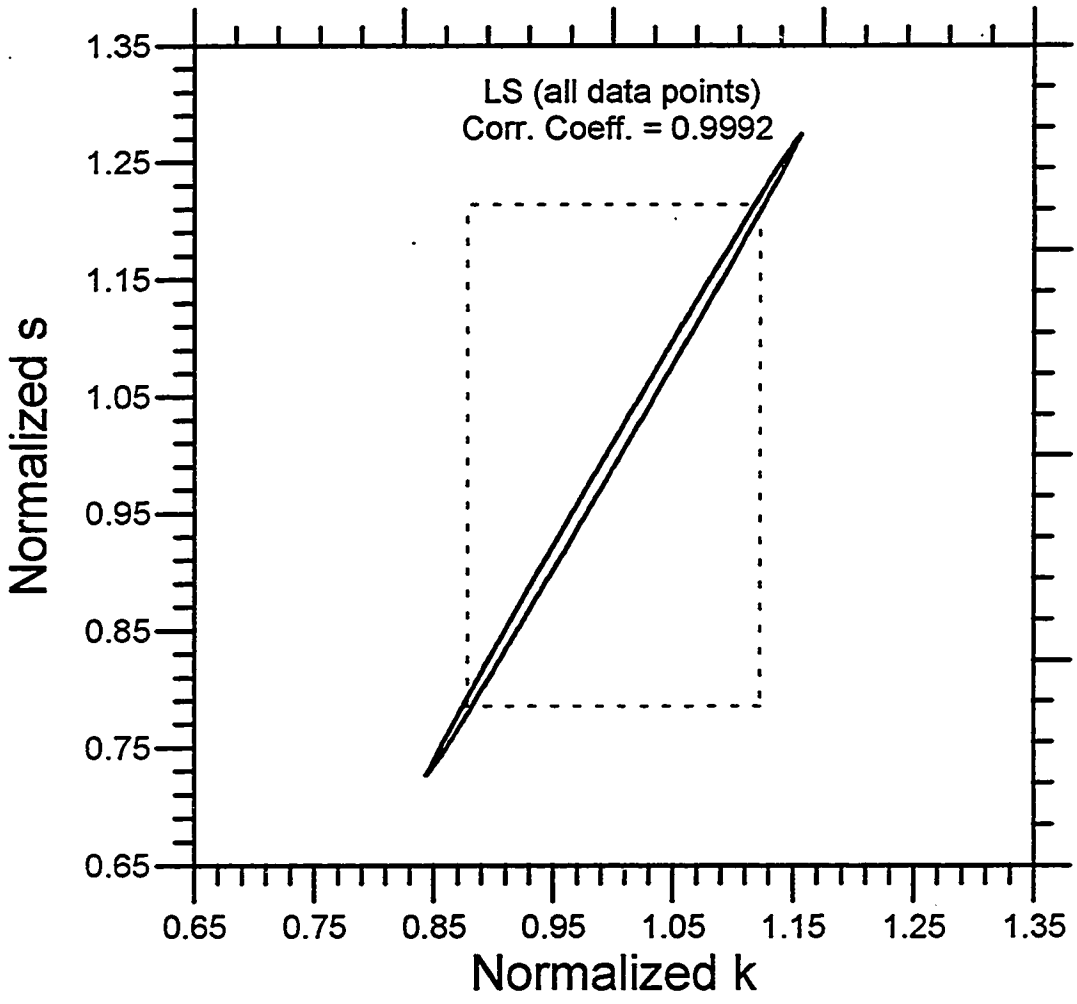

Fig. 3.3.10 - 95% confidence region versus
95% confidence intervals - Field Example #1

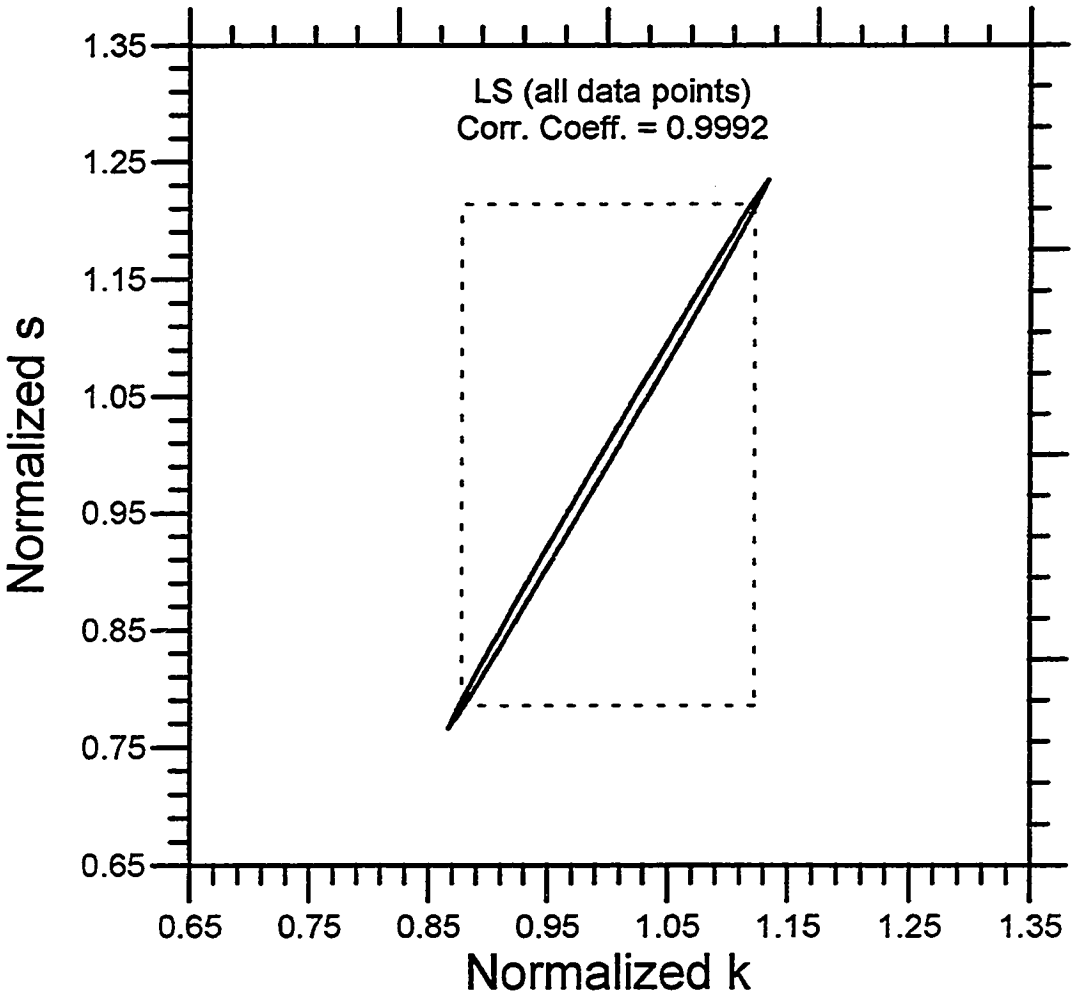

Fig. 3.3.11 - 90.25% confidence region versus 95% confidence intervals - Field Example #1

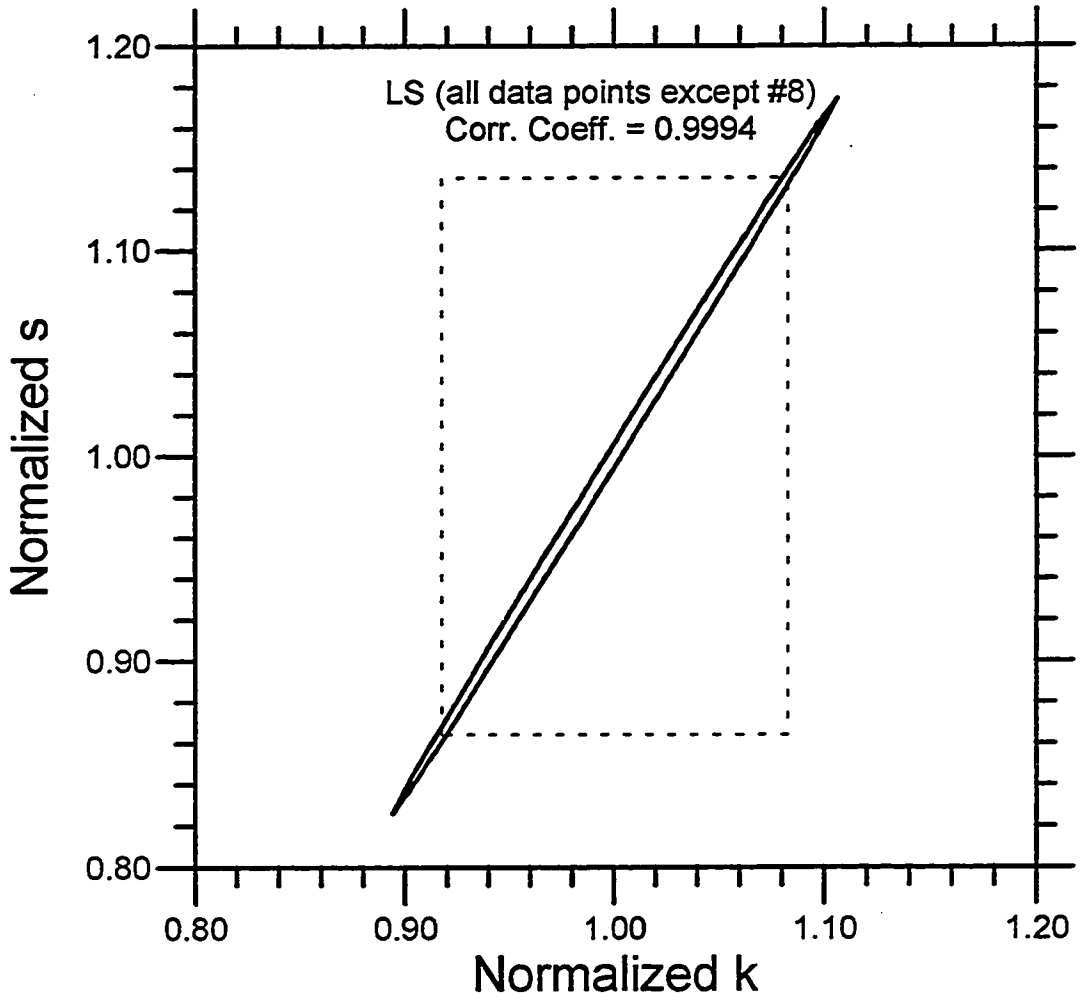

Fig. 3.3.12 - 95% Confidence region versus
95% confidence intervals - Field Example #1

Figs. 3.3.7 through 3.3.8 show that the pairs (C, s) and (C, k) are not highly correlated which implies that the rectangular regions obtained as the cross product of the two one-dimensional confidence intervals, are good approximations to the joint 2-dimensional confidence regions. Note that in these cases, the axes of the ellipses are almost parallel to the Cartesian coordinate axes. On the other hand, Figs. 3.3.10 through 3.3.12 show that k and s are strongly correlated. This implies that the angle between one of the axes of the ellipse and Cartesian axes is about 45°, and the rectangular region formed by the individual confidence limits of the parameters does not at all resemble the joint confidence region.

Table 3.3.3 - Correlation coefficients for Field Example #1

Pair of Parameters	Correlation Coefficients LS regression	
	all data points	all points except #8
(k, s)	0.99945	0.99962
(k, C)	0.60559	0.61092
(C, s)	0.61387	0.61681

Figs. 3.3.7 and 3.3.8 present the confidence regions of the pair (C, s) for 95% and 90.25% confidence levels, respectively. Note that the 90.25% confidence region better approximates the rectangular region formed by the individual confidence intervals. Similarly, Figs. 3.3.11 and 3.3.12 presents the 95% and 90.25% confidence regions for parameters C and k. It seems that, if two parameters are completely uncorrelated, the rectangular confidence region, formed by the 95% individual confidence intervals of the parameters, approximates a 90.25% (note that $0.95^2 = 0.9025$) confidence region. In other words, the one-dimensional projections of a $(1-\gamma)\times 100\%$ p-dimensional confidence region, should define corresponding $(1-\gamma)^p$ individual confidence intervals.

At this point, we verify the linearity assumption by comparing the objective function at the end points of the 3-dimensional confidence region for k, C and s (see Chapter II, Section 2.3.4). Table 3.3.4 presents the coordinates of the end points of the three axes of the ellipsoidal confidence region and the respective values of the objective function. Under the linearity assumption, the expected value of the objective function at these points (see Eq. 2.3.17) is 22.0579 psi^2.

Table 3.3.4 - Values of the objective function at the end points of the 3-dimensional confidence region of Field Example #1

k (md)	C (bbl/psi)	s	LS Objective Function (psi^2)
110.8870	0.0883231	13.8056	18.5751
67.78236	0.0842403	5.55294	31.5663
89.30944	0.0871564	9.81118	22.0028
89.36995	0.0854071	9.54736	22.0674
89.33470	0.0888154	9.67926	21.7009
89.33469	0.0837481	9.67929	22.3142

As we can in Table 3.3.4, the objective function for the first two points (associated with the same axis) is different from the expected value. Thus, the conventional confidence intervals we computed before may not be accurate. Figures 3.3.13 through 3.3.15 show the likelihood ratio confidence intervals for parameters k, s and C, respectively. The procedure we use is described in Chapter II, Section 2.3.5.

It is interesting to note that because of the non-linearity, the estimates for k ($k = 89.34$) and s ($s = 9.679$) are not centered in the likelihood ratio confidence intervals, $78.55 < k < 103.59$ and $7.616 < s < 12.416$, respectively. In this example, the conventional

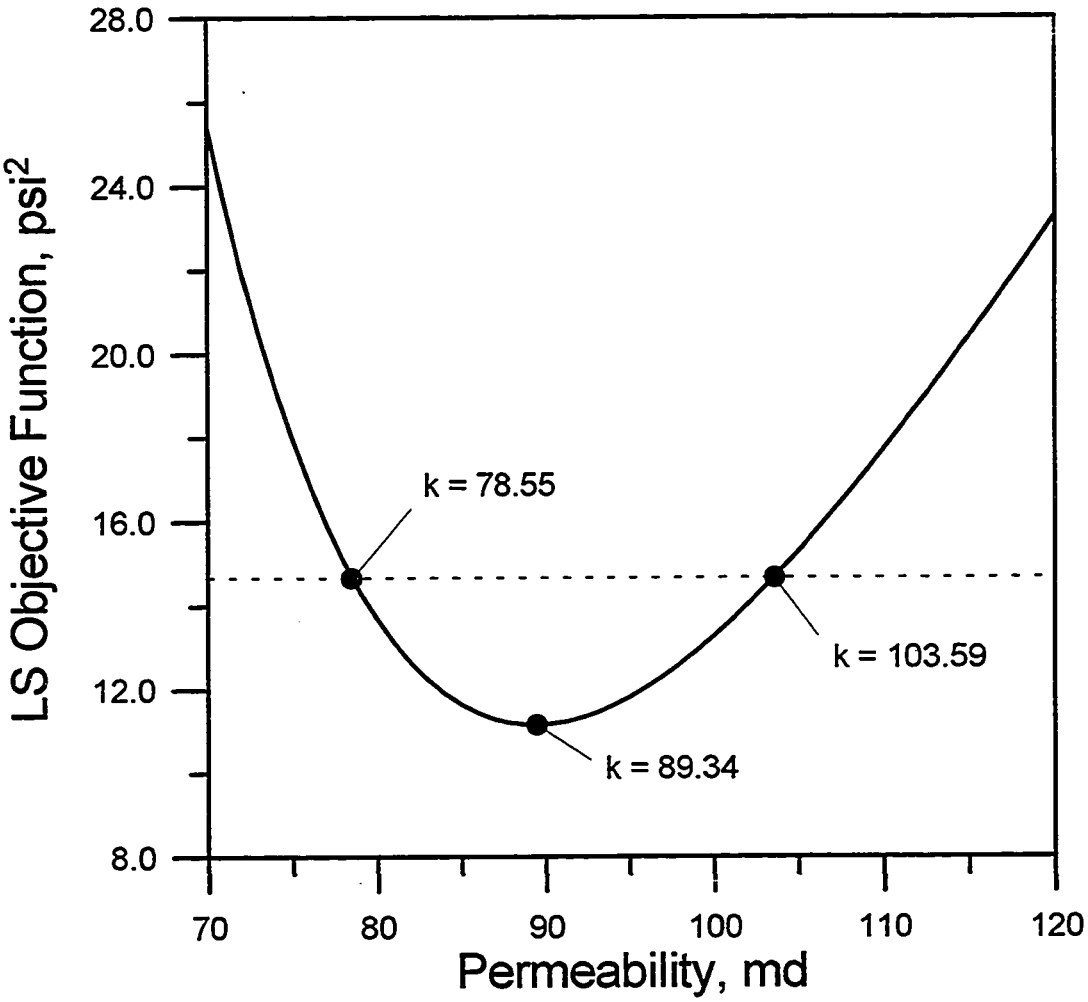

Fig. 3.3.13 - Likelihood-ratio confidence interval for permeability - Field Example #1

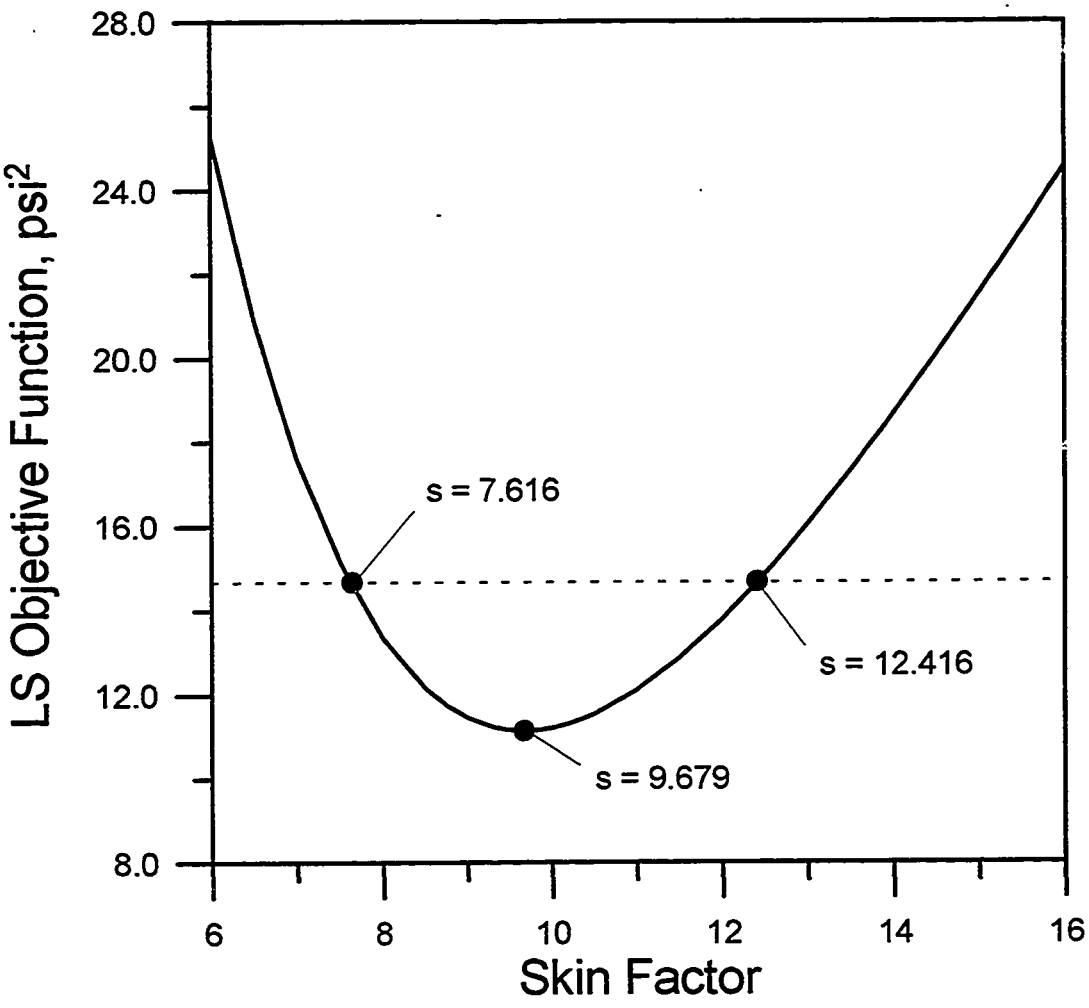

Fig. 3.3.14 - Likelihood-ratio confidence interval for the skin factor - Field Example #1

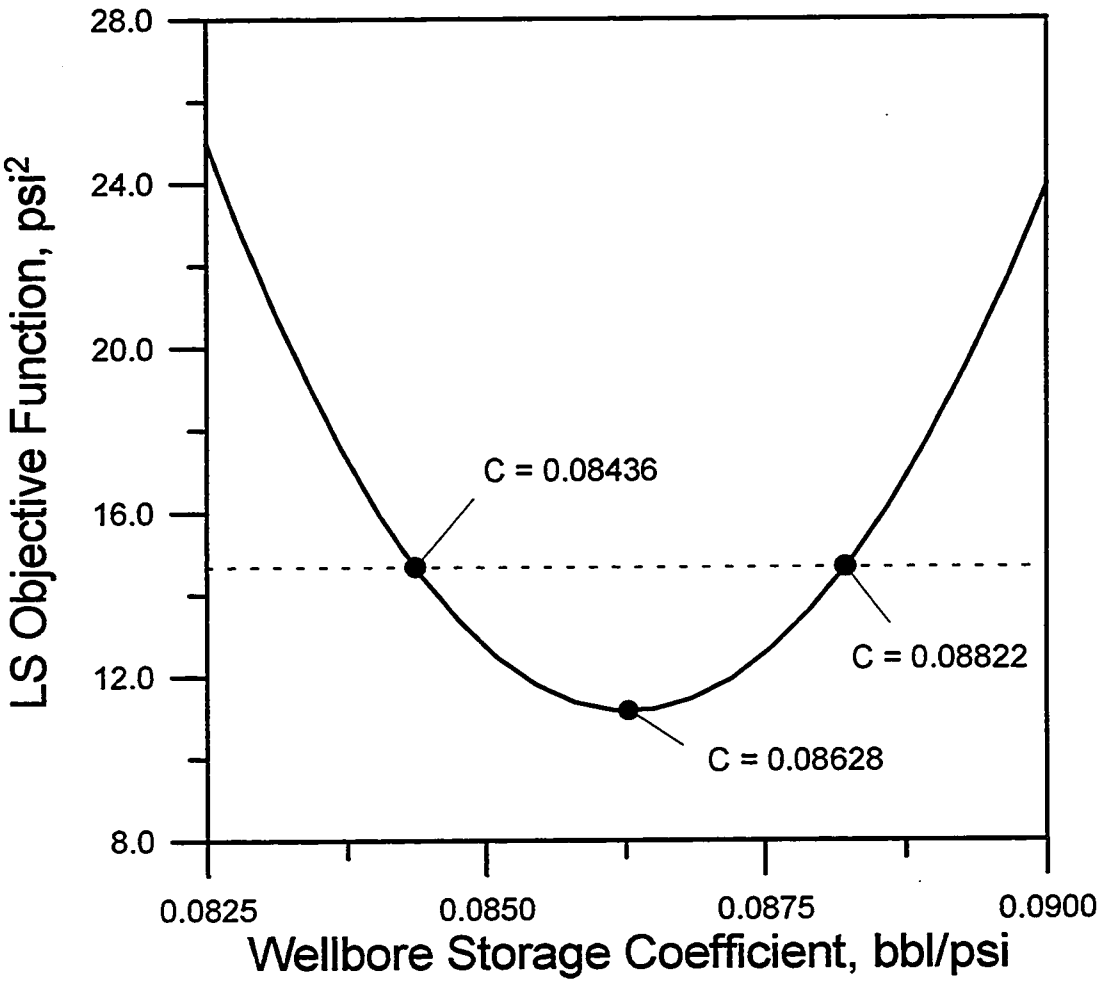

Fig. 3.3.15 - Likelihood-ratio confidence interval for the wellbore storage coefficient - Field Example #1

confidence intervals for all three parameters were conservative, i.e., the conventional confidence intervals accommodate the likelihood ratio confidence intervals.

3.4 Field Example #2

For our second field example, we analyze the drawdown data that was used in Ref. 8 to illustrate the applicability of the LAV method. In order to demonstrate the robustness of LAV analysis, Ref. 8 deliberately introduced three outliers into the data; the entire resulting pressure drop data set is shown in Figs. 3.4.1 through 3.4.3. The solid black circular points which fall below the solid curve in Fig. 3.4.1 are the artificially introduced outlier data; these are points #12, 19 and 25 in the data set. Table 3.4.1 summarizes the results of various analysis runs on the pressure data. We used the same initial guess as in Ref. 8, i.e., $k = 100$ md, $C = 0.001$ bbl/psi and $s = 100$. Figs. 3.4.1 through 3.4.3 compare computed model responses with the "measured" pressure data. The first two rows of Table 3.4.1 show the results obtained when the entire data set is analyzed using the LAV and LS methods. Note that the results of both runs are significantly different, since LS regression is sensitive to the presence of the outliers; this point is more clearly emphasized by the large degree of uncertainty associated with each LS parameter estimate. Our outlier detection analysis (based on the LAV fit) indicated that points #12, 15, 19 and 25 are outliers. Note that point #15 belongs to the original measured data. Figure 3.4.4 shows a plot of the scaled residuals for the LAV fit (r_i/s^*) versus producing time. The dashed lines correspond to $r_i/s^* = \pm 3.0$; the circular data points for $r_i/s^* < 3.0$ are the artificially introduced outliers, while the circular point above $r_i/s^* = 3.0$ is point #15, the detected outlier in the original measured data set. Figure 3.4.5 shows a normalized residual plot for the original data *without* the artificially introduced outliers.

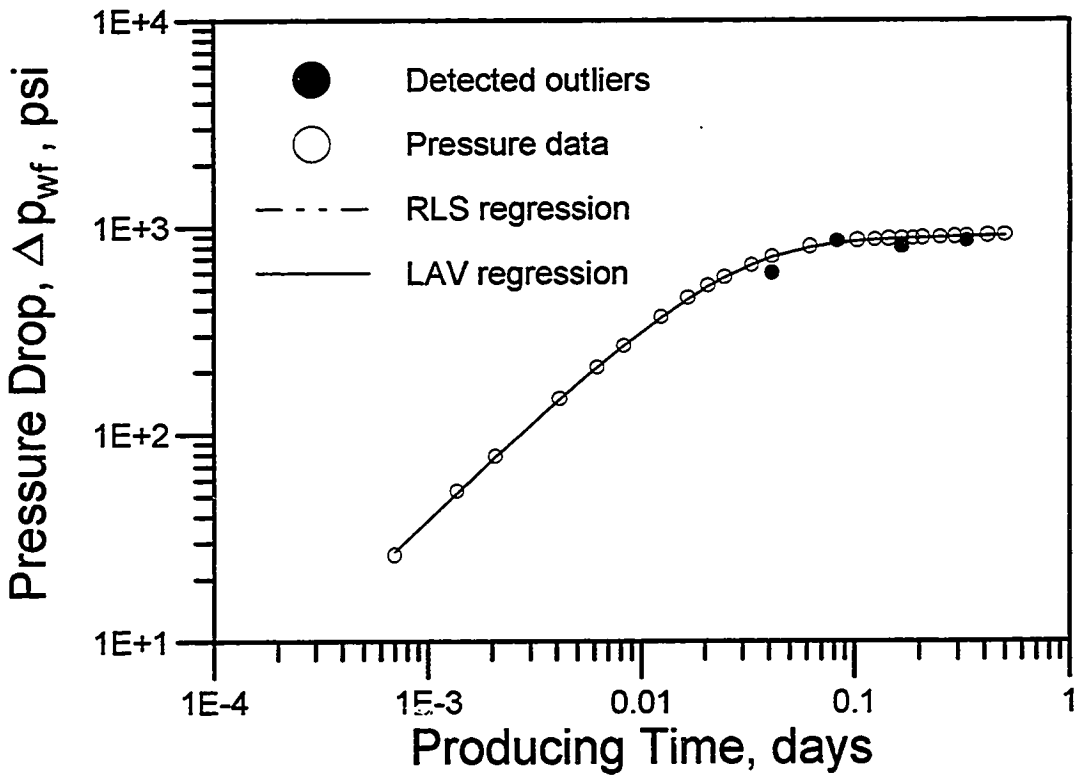

Fig. 3.4.1 - Comparison between LAV regression and RLS regression - Field Example #2

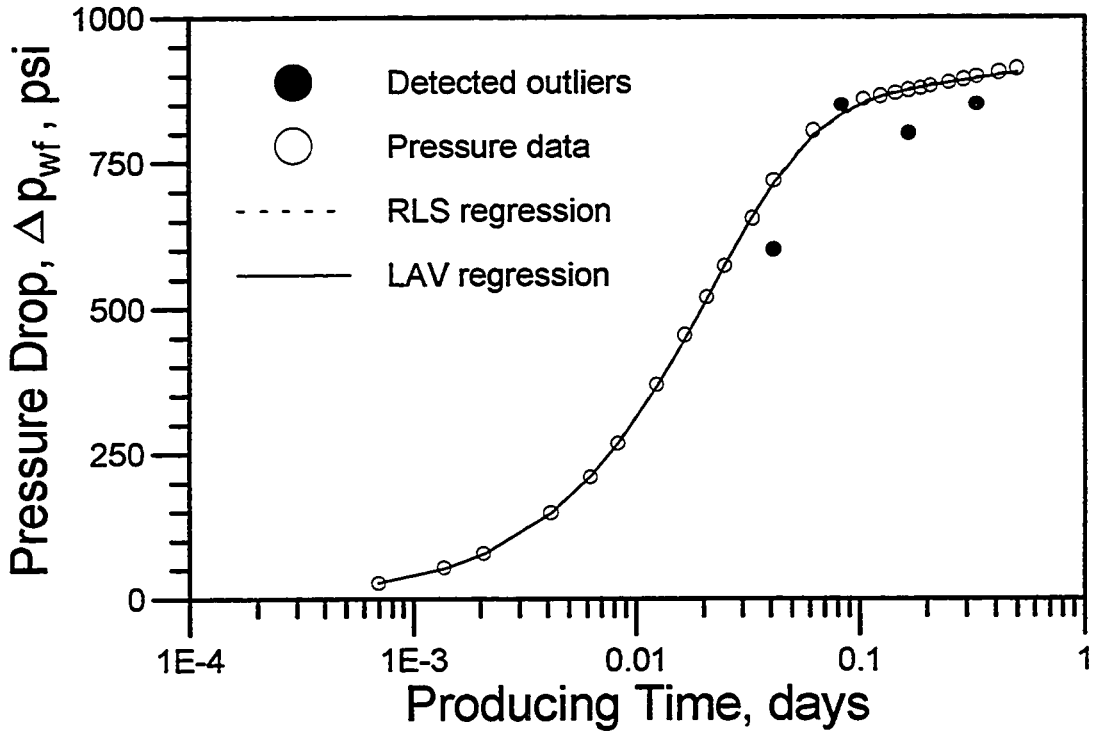

Fig. 3.4.2 - Comparison between LAV regression and RLS regression - Field Example #2

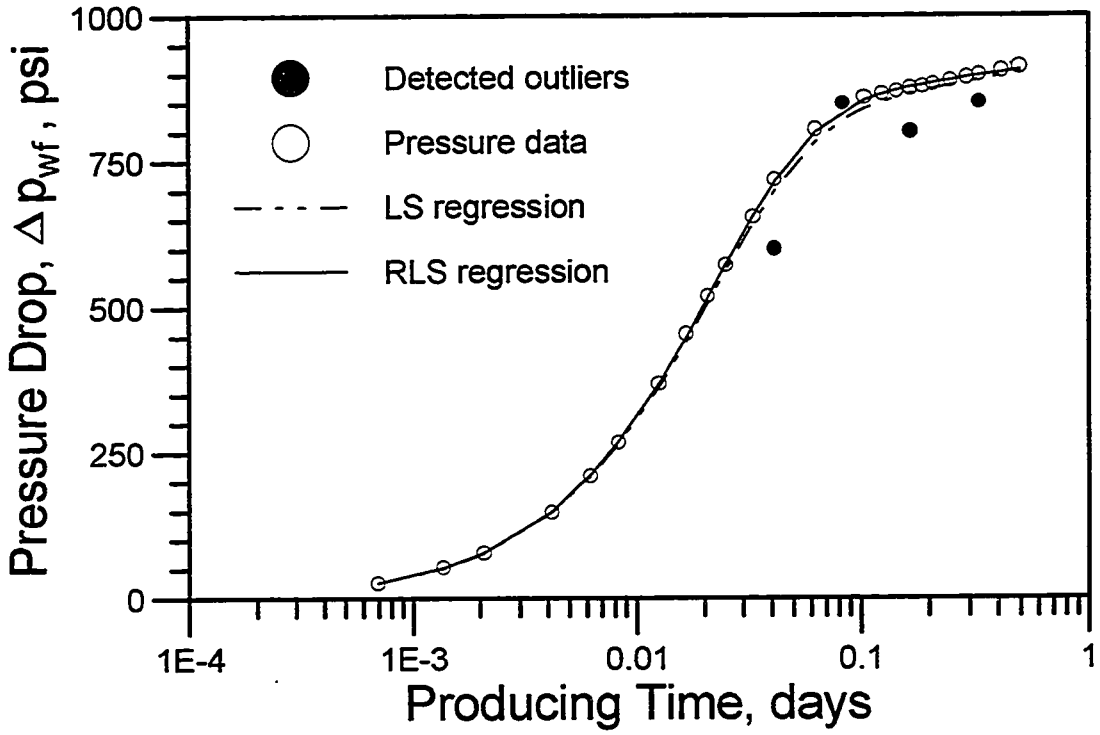

Fig. 3.4.3 - Comparison between LS regression and RLS regression - Field Example #2

Table 3.4.1 - Summary of the results for Field Example #2

Type of Regression (Pressure Only)	Parameter Estimates with Respective Confidence Intervals			Number of Iterations	STD Deviastion (psi)	Outliers Detected
	k (md)	C (bbl/psi)	s			
LAV with outliers	13.9	0.00845	12.6	18	3.12	points # 12 15, 19 and 25
LS with outliers	11.1 ± 77.%	0.00847 ± 11.%	8.6 ± 138.%	12	27.8	–
LAV without points # 12, 19 and 25	13.6	0.00842	12.2	49	2.84	point # 15
LS without points # 12, 19 and 25	15.1 ± 15.%	0.00847 ± 1.5%	14.3 ± 22.%	12	3.75	–
LAV without points # 12, 15, 19 and 25	13.3	0.00841	11.8	14	2.91	none
LS without points # 12, 15, 19 and 25	13.9 ± 11.%	0.00843 ± 1.1%	12.6 ± 17.%	5	2.81	–

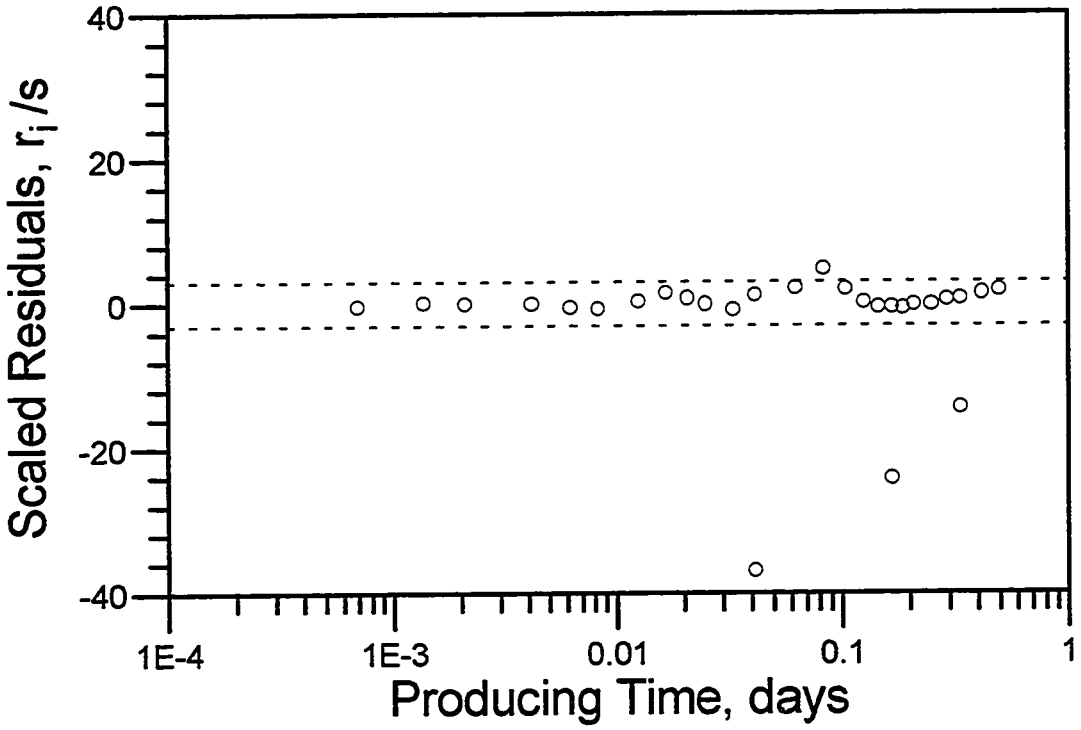

Fig. 3.4.4 - Scaled residual plot of the original pressure data with outliers - Field Example #2

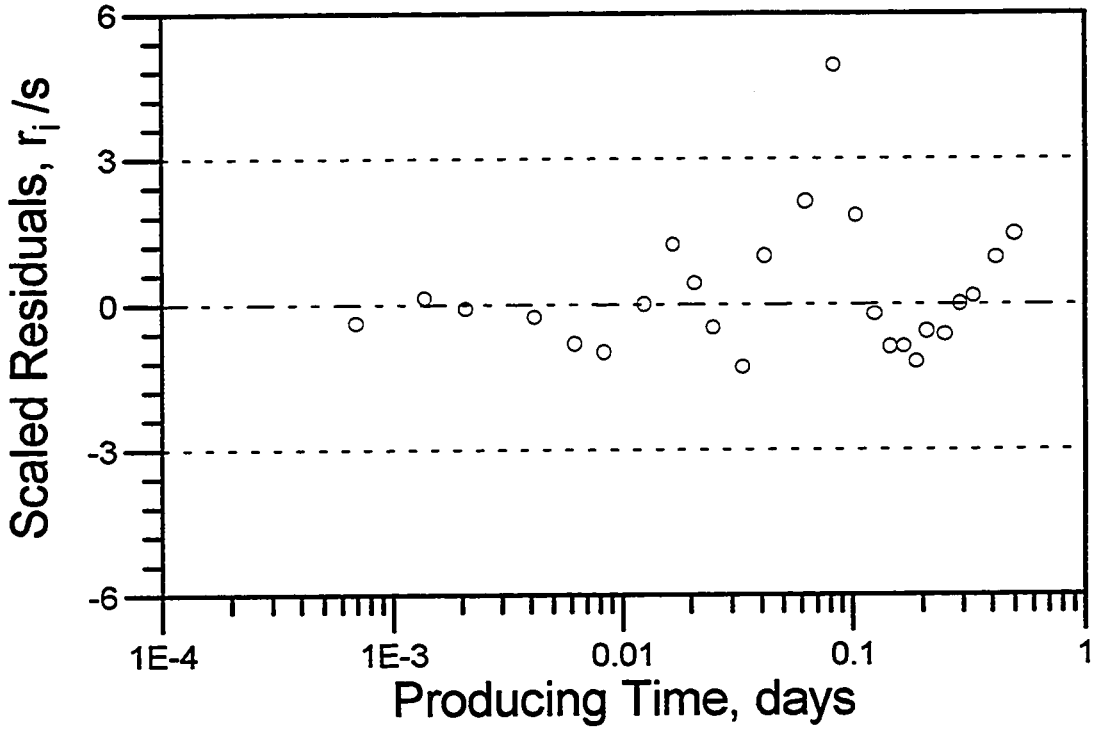

Fig. 3.4.5 - Scaled residual plot of the original pressure data - Field Example #2

Rows 3 and 4 of Table 3.4.1 compare the parameter estimates using LAV and LS regression on the original pressure data set; i.e., with points #12, 19 and 25 removed. Two points are worth noting: (i) the LAV parameter estimates are virtually unchanged, i.e., the LAV results are insensitive to the presence of outliers, and (ii) the LAV and LS results are still different, even though the confidence intervals on the LS estimates are greatly improved. Outlier analysis on this LAV run continued to indicate that point #15 is an outlier.

The remaining rows of Table 3.4.1 show the final parameter estimates with all outliers removed. In this case, LAV analysis and LS (now denoted "RLS") analysis yield similar results. Note that the final parameter confidence intervals for the RLS analysis are significantly smaller than the confidence intervals obtained with all outliers present. Figures 3.4.1 and 3.4.2 compare measured and model pressure responses using LAV and RLS parameter estimates; note that the LAV and RLS matches overlay each other. In contrast, Fig. 3.4.3 shows that the model response computed with the original LS parameters is affected by the introduction of the artificial outliers.

3.5 Field Example #3

Here, we illustrate that the analysis methods devised for wellbore storage and skin problems can be applied equally to slug test data which has been appropriately transformed. If we denote the measured pressure drop from a slug test by Δp_{slug}, and the analogous pressure drop from a well with skin and wellbore storage by Δp_{SS}, Ref. 36 shows that the slug test data can be transformed to an equivalent unit rate drawdown skin and wellbore storage pressure and pressure derivative data via the following relationships:

$$\Delta p_{SS} = \frac{I(\Delta p_{\text{slug}})}{24C(p_i - p_o)}, \tag{3.5.1}$$

where

$$I(\Delta p_{\text{slug}}) = \int_0^t \Delta p_{\text{slug}}(\tau)d\tau, \tag{3.5.2}$$

and

$$\Delta p'_{SS} = \frac{t\Delta p_{\text{slug}}}{24C(p_i - p_o)}. \tag{3.5.3}$$

Following Ref. 36, we transformed the slug test data of Ref. 37 to an equivalent unit rate storage and skin data, and optimized parameter estimates by regressing on $I(\Delta p_{\text{slug}})$. The resulting match is shown in Fig. 3.5.1, and results from the analysis runs are summarized in Table 3.5.1. Note that in this case, the wellbore storage coefficient C is assumed to be known, and we only regressed on k and s. (For this example, the initial guesses of the parameters were $k = 100$ md and $s = 50$). The agreement between LAV and LS methods is excellent (see Fig. 3.5.2), since no outliers were detected in the data. The results of Table 3.5.1 also illustrate a common phenomenon, that the LAV method frequently converges much more slowly than the least squares method.

Table 3.5.1 - Summary of the results for Field Example #3

Type of Regression (pressure data)	Parameter Estimates with Respective Confidence Intervals			Number of Iterations	STD Deviation (psi)
	k (md)	C (bbl/psi)	s		
LAV all data points	889.	0.0365	–2.4	48	0.393
LS all data points	900.±9.8%	0.0365	–2.4±12%	18	0.377

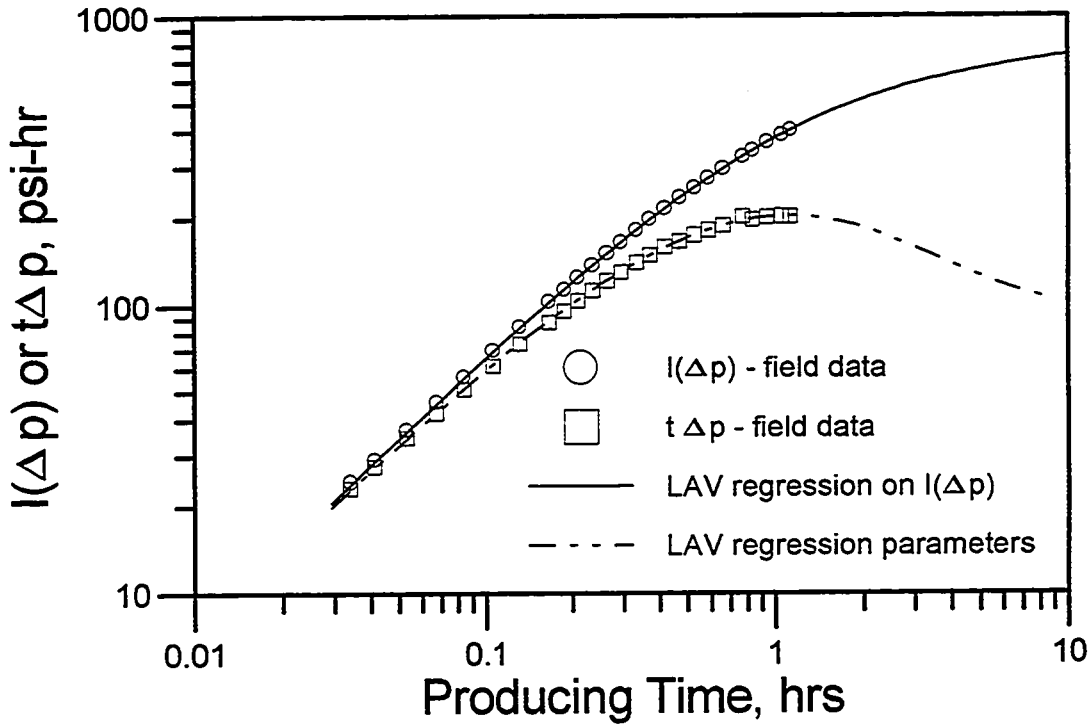

Fig. 3.5.1 - LAV regression applied to slug test data - Field Example #3

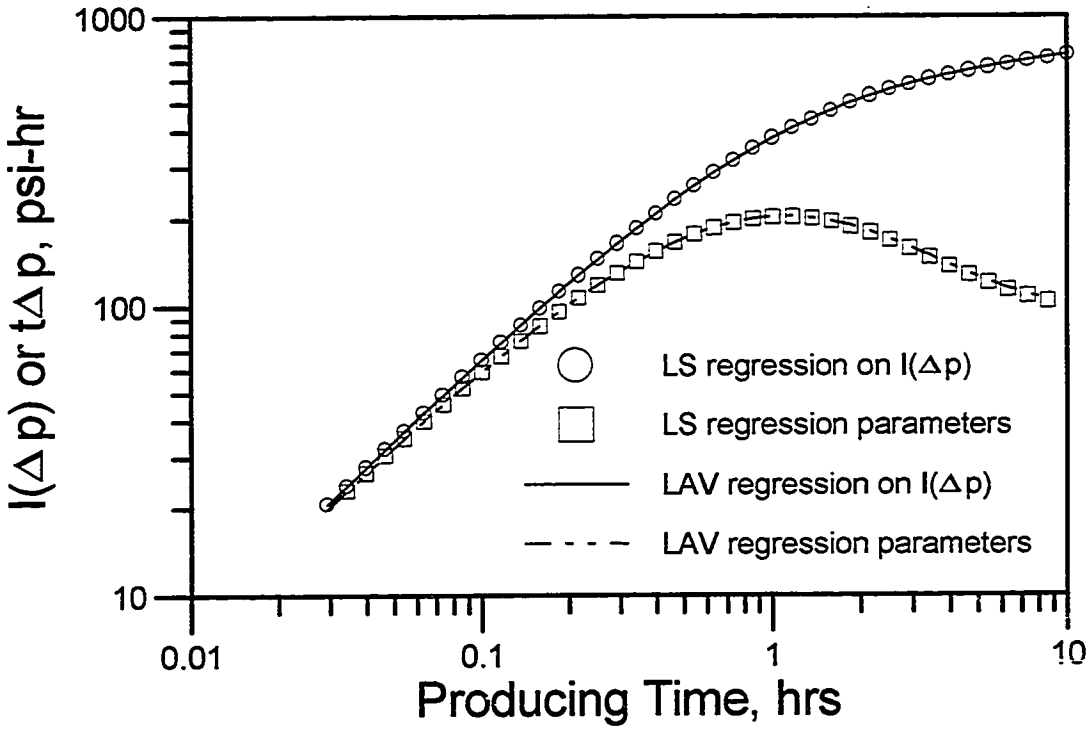

Fig. 3.5.2 - LS and LAV regressions applied to slug test data - Field Example #3

3.6 Field Example #4

In our final field example, we demonstrate that the method is applicable to the analysis of pressure buildup data if the model pressure response is generated using the principle of superposition[1]. For a multirate test consisting of m distinct constant surface rates ($q = q_l$ for $t_{l-1} < t \le t_l$), the model pressure response at any time $t = t_m > t_{m-1}$ is given by,

$$p_{wf}(t) = p_i - \sum_{l=0}^{m-1} q_{l+1}\left\{\Delta p_{wc}(C, s, t - t_l) - \Delta p_{wc}(C, s, t - t_{l+1})\right\}, \qquad (3.6.1)$$

where Δp_{wc} denotes the unit surface rate pressure drop for a fixed storage constant C. It is clear from Eq. 3.6.1 that p_{wf} is a linear function of p_i; thus, nonlinear regression analysis can be used to estimate initial pressure[1], provided wellbore storage effects are present in the data. If there are no wellbore storage effects, p_i and s are simply additive constants in the solution, and cannot be determined independently.

We analyze the buildup data of Ref. 27, shown as data points in Figs. 3.6.1 and 3.6.2. We applied both LAV and LS analysis to the pressure data only using the following initial guesses: $k = 50$ md, $C = 0.001$ bbl/psi, $s = 20$ and $p_i = 4000$ psi. The results of these analyses are shown in Table 3.6.1. No data outliers were detected, and LAV and LS analyses yielded essentially the same results.

Figure 3.6.1 shows semilog plots of the LAV and LS model pressure (which overlay on a single solid curve) and the measured field data. Similarly, Fig. 3.6.2 shows a log-log plot of measured versus model pressure and pressure derivative data. Note that the model derivative data matches closely with the measured derivative data even though derivative data was not considered in the nonlinear regression analysis.

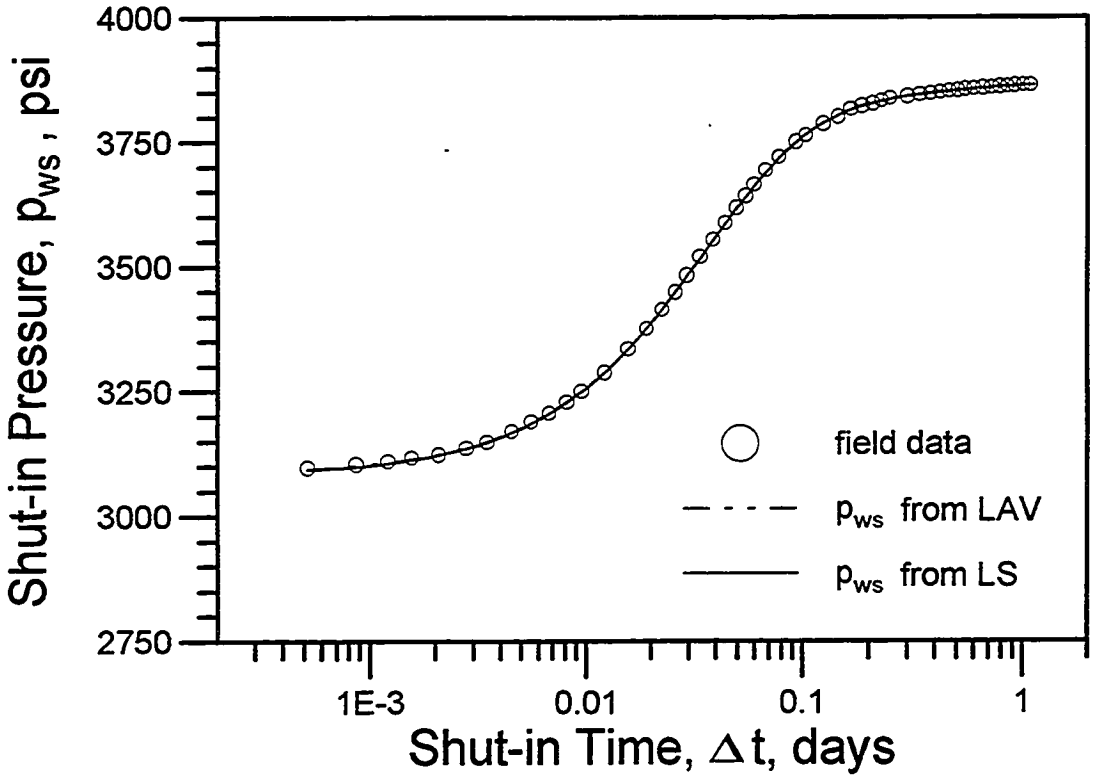

Fig. 3.6.1 - LS and LAV regressions applied to buildup data - Field Example #4

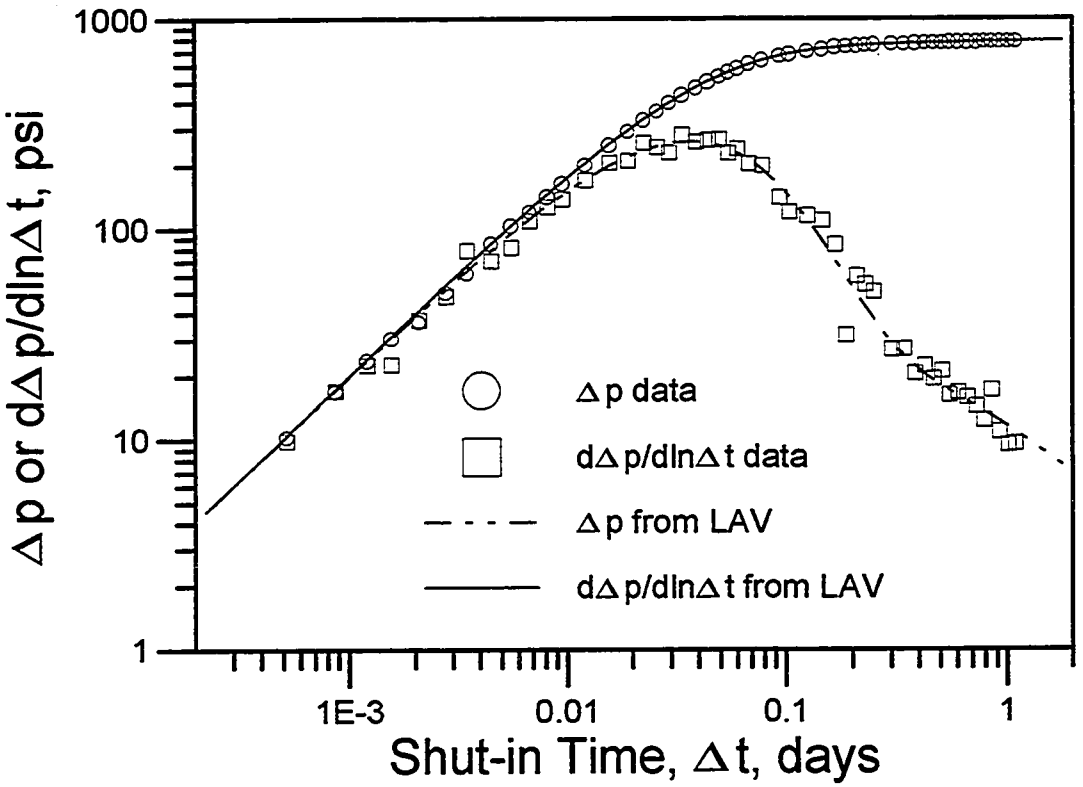

Fig. 3.6.2 - Type-curve match from LAV regression applied on pressure data - Field Example #4

Table 3.6.1 - Summary of the results for Field Example #4
(Tolerance = 10^{-6} and Stehfest # = 8)

Type of Regression (pressure data)	Parameter Estimates with Respective Confidence intervals				Number of Iterations	STD Deviation (psi)
	k (md)	C (bbl/psi)	s	p_i (psi)		
LAV all points	11.57	0.00903	8.66	3877.3	28	1.936
LS all points	11.86 $\pm$ 0.424	0.00910 $\pm 5.96\times10^{-5}$	8.97 $\pm$ 0.508	3876.6 $\pm$ 1.159	10	1.774

The parameter estimates obtained from regression analysis agree with those we obtained from conventional analysis (i.e., type curve matching, and semilog analysis using Agarwal equivalent time or Horner analysis). However, because of the producing time effects on the buildup data, conventional analysis requires an iterative process to obtain accurate estimates and thus is extremely tedious and time consuming, whereas, regression analysis yields equivalent results in a few seconds.

The values of the objective function at the end points of the four axes of the ellipsoidal confidence region of the parameter estimates lie between 338.7 psi^2 and 341.9 psi^2. Thus, in a practical sense, the linearity assumption is valid in this example.

Figures 3.6.3 through 3.6.8 show plots of the confidence region and the confidence intervals of the parameters (as the sides of the dashed rectangles) using the same procedure discussed in Field Example #1. Again, k and s are highly correlated to each other. The initial pressure does not seem to be strongly correlated to any other parameter. Table 3.6.2 presents the correlation coefficients for this example.

For the wellbore storage and skin problems we have studied in this chapter, a correlation coefficient less than approximately 0.8 usually means a small correlation between

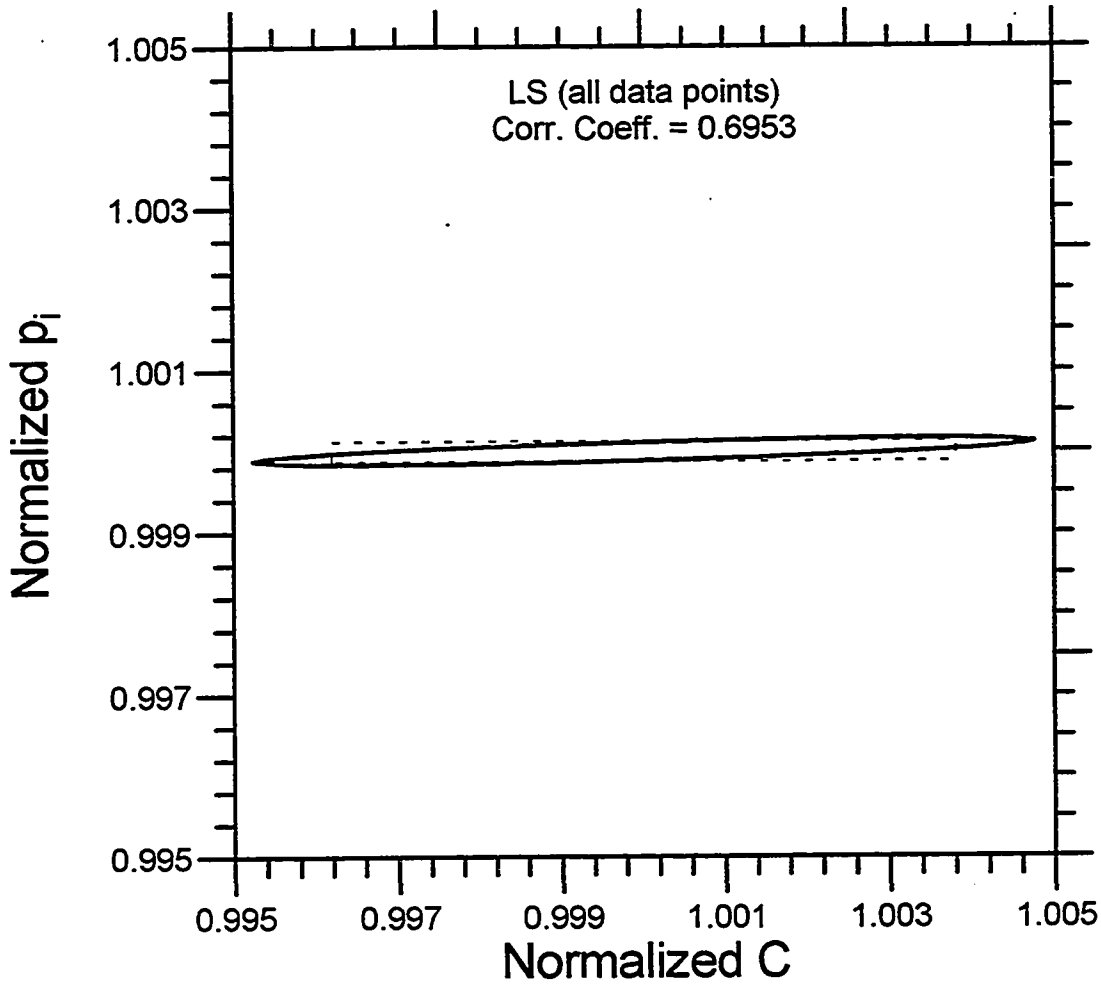

Fig. 3.6.3 - 95% confidence region versus
95% confidence intervals - Field Example #4

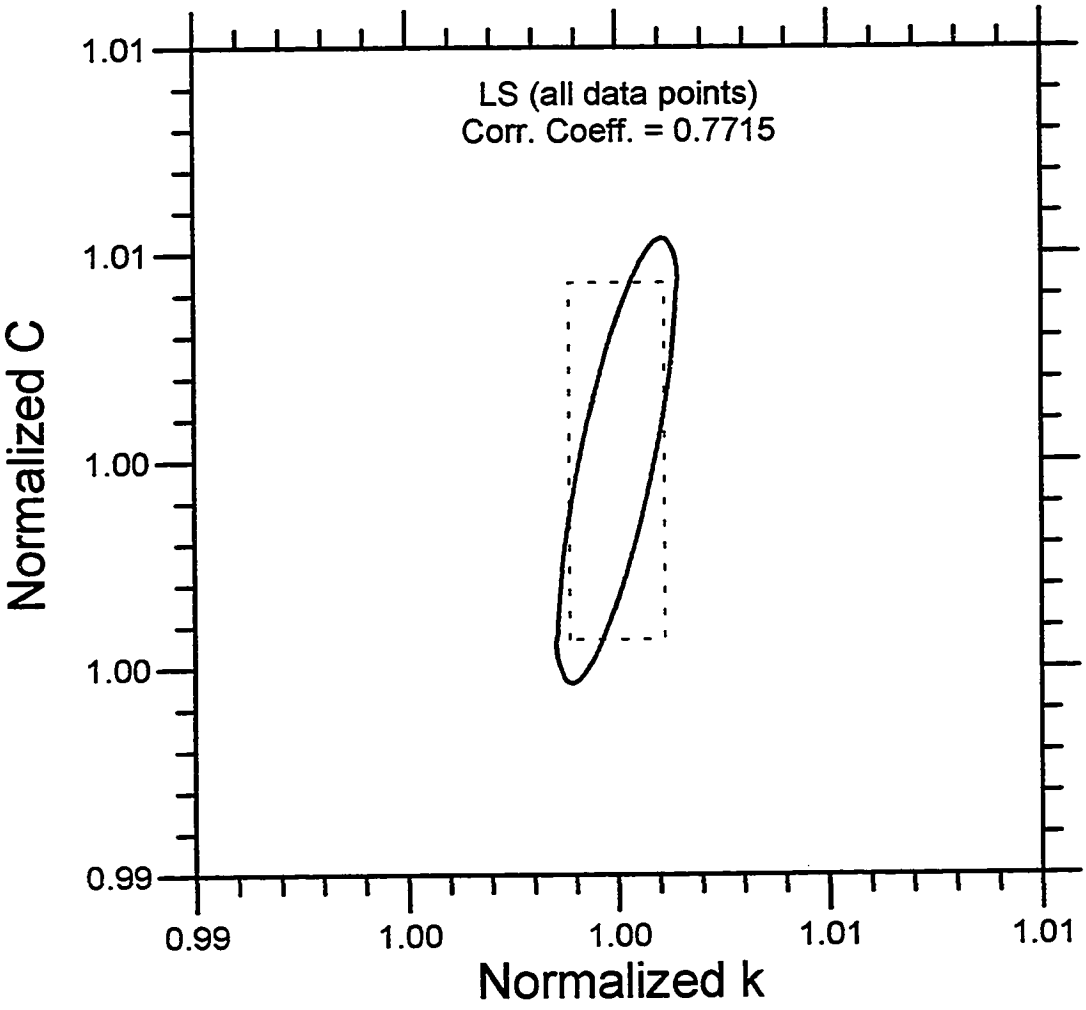

Fig. 3.6.4 - 95% confidence region versus
95% confidence intervals - Field Example #4

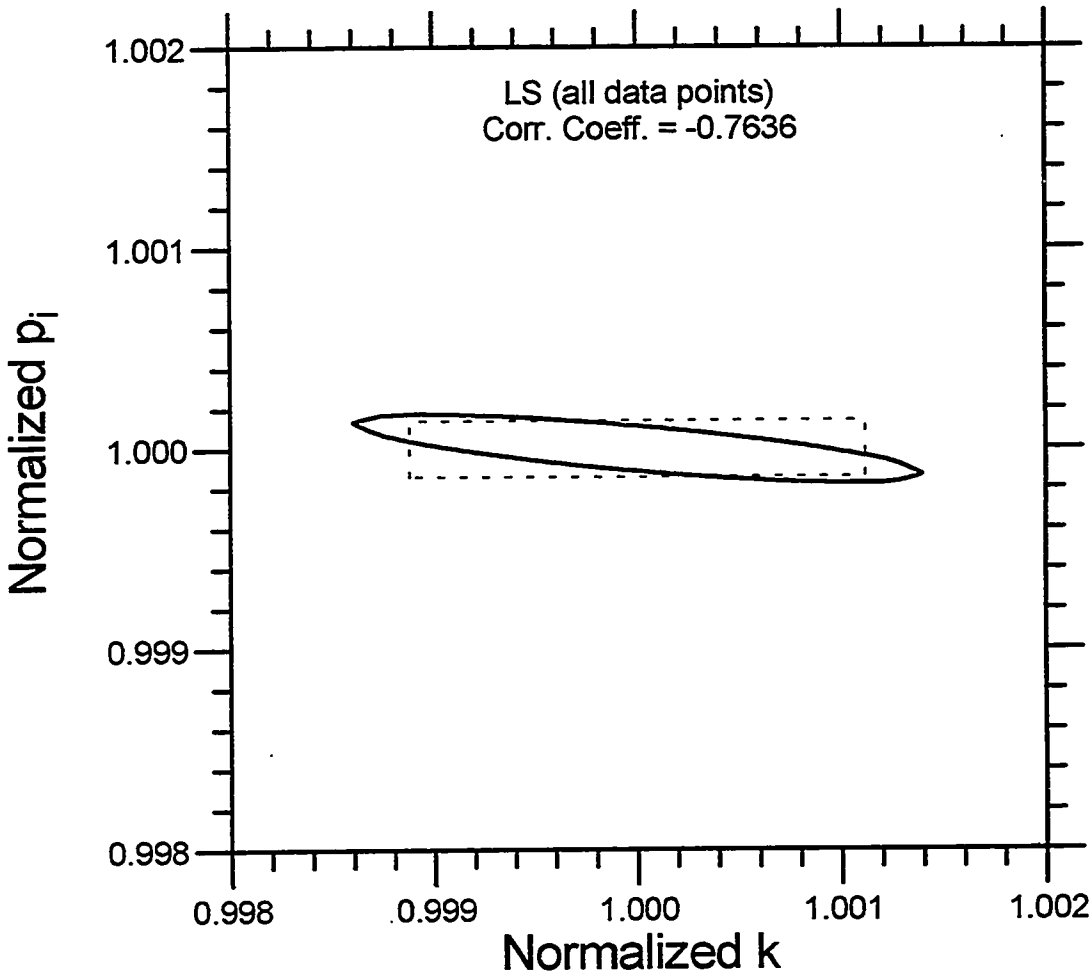

Fig. 3.6.5 - 95% confidence region versus
95% confidence intervals - Field Example #4

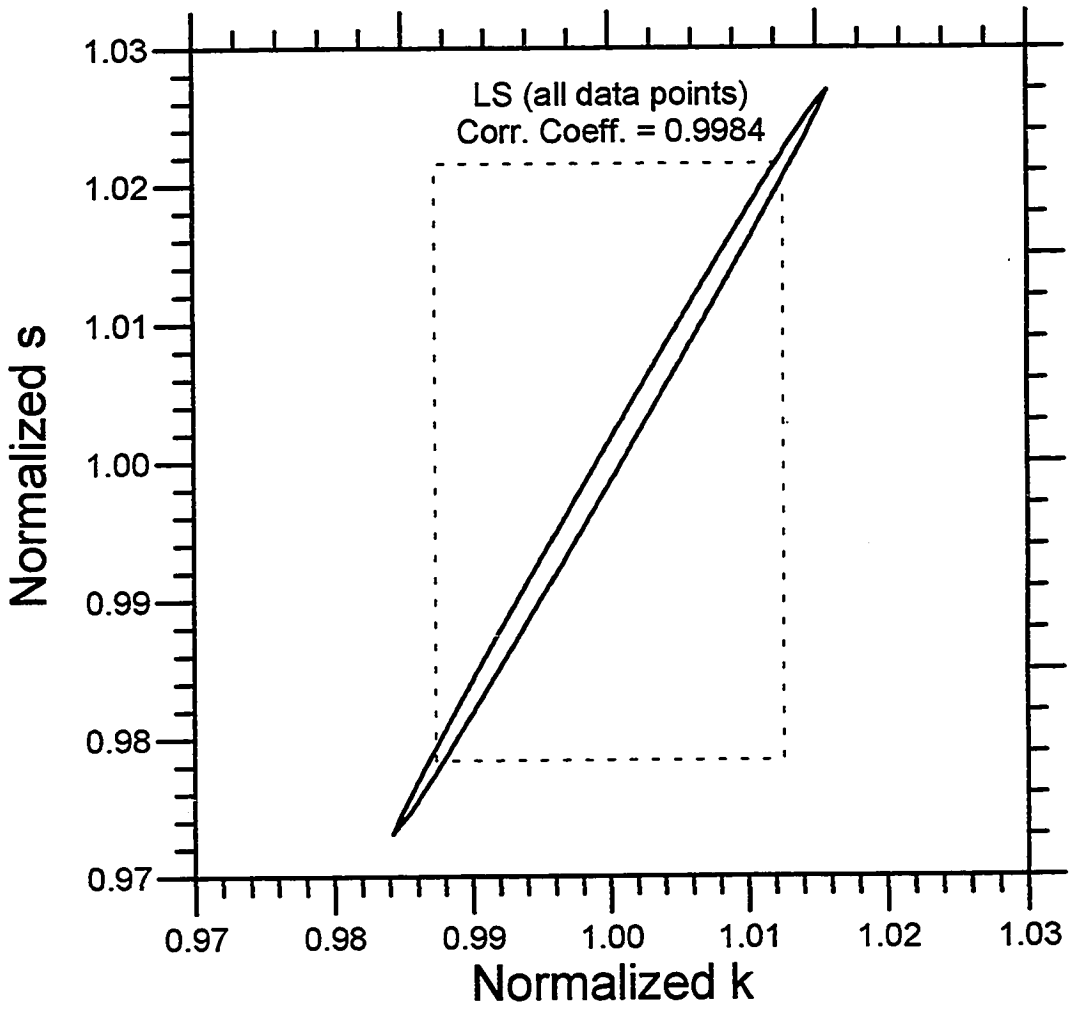

Fig. 3.6.6 - 95% confidence region versus
95% confidence intervals - Field Example #4

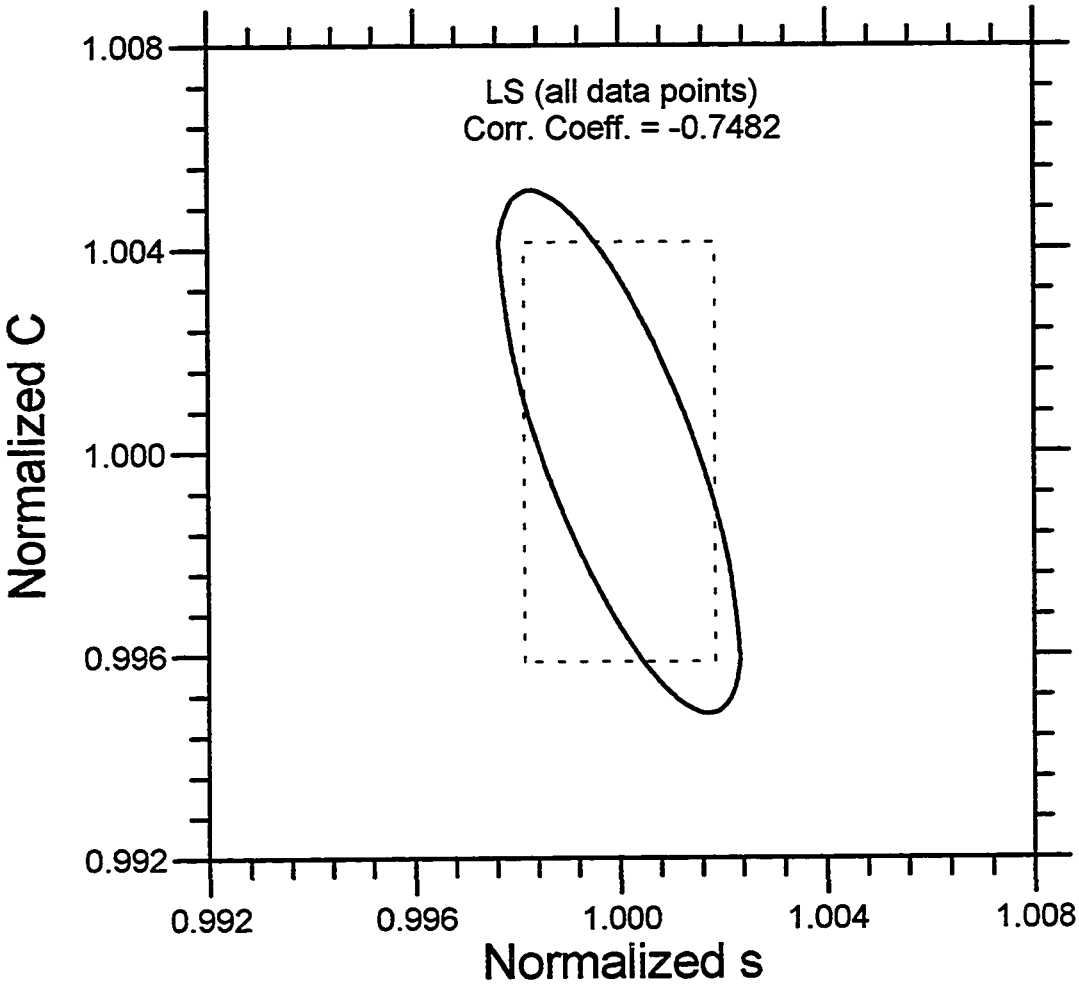

Fig. 3.6.7 - 95% confidence region versus
95% confidence intervals - Field Example #4

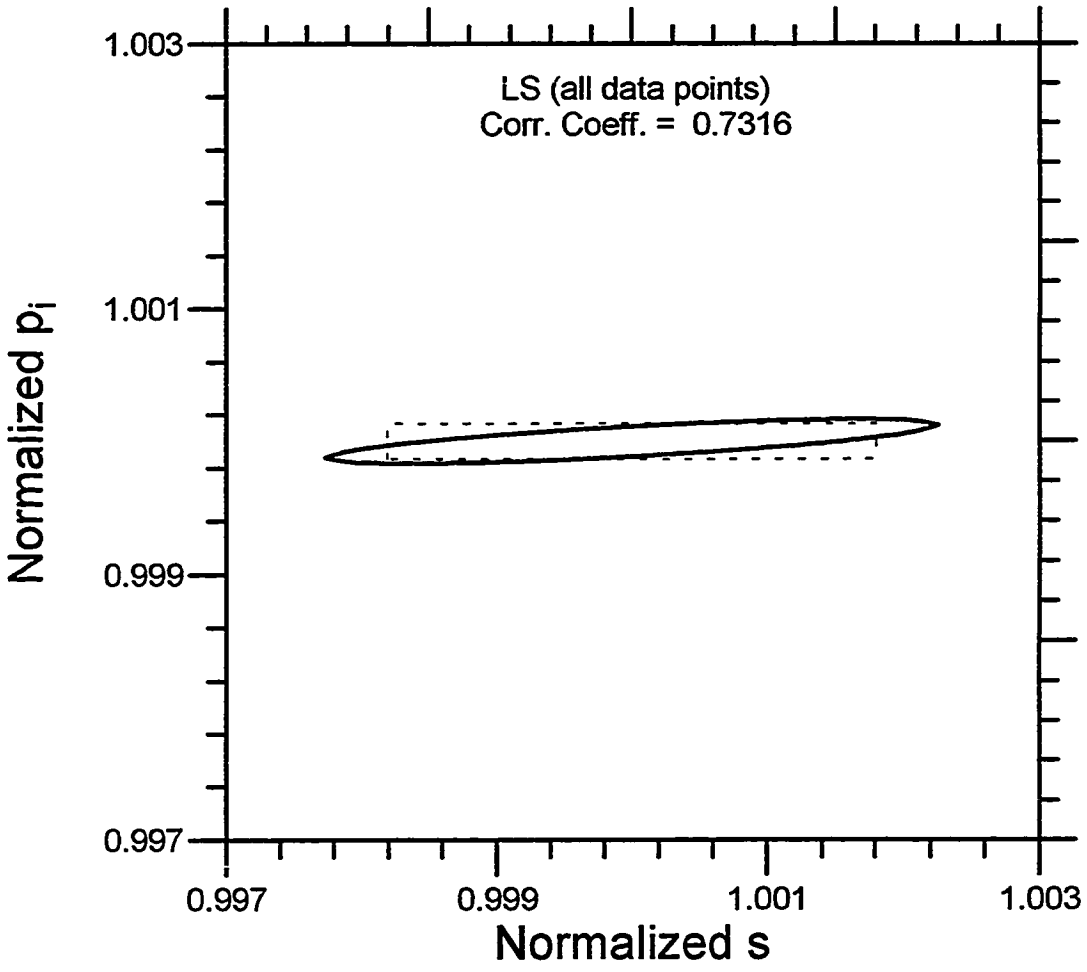

Fig. 3.6.8 - 95% confidence region versus
95% confidence intervals - Field Example #4

the parameters, in the sense that a good approximation to the correct two-dimensional confidence region can be constructed from the corresponding one-dimensional confidence intervals.

Table 3.6.2 - Correlation coefficients for Field Example #4

Pair of Parameters	Correlation Coefficient
(k, s)	0.9933
(k, C)	0.7429
(C, s)	0.7303
(k, p_i)	-0.8852
(s, p_i)	-0.8771
(C, p_i)	-0.5569

There is a final important remark. Once we have removed outliers and performed least squares regression, we can estimate the standard deviation of the measurement errors by Eq. 2.4.5, which we repeat here as

$$s^* = \sqrt{\frac{\sum_{u=1}^{n} r_u^2}{\hat{n} - p}}, \tag{3.6.2}$$

where $\hat{n}$ is the number of data points left after we have removed outliers. This estimate of σ should correspond to the accuracy of the pressure measurement. Note that in the synthetic example application of Table 3.2.4, we introduced errors into the data where the random error variable had a Gaussian distribution of mean zero and standard deviation $\sigma = 0.5$ psi. In all cases, the estimates of the standard deviation, s^*, computed from the

regression estimates using Eq. 3.6.2, was close to the value of $\sigma = 0.5$ psi. However, in our field applications, we obtained, $s^* = 0.555$ psi for Field Example 1, $s^* = 2.81$ psi for Field Example 2 (see Table 3.4.1), $s^* = 0.377$ psi for Field Example 3 (see Table 3.5.1), and $s^* = 1.774$ psi for Field Example 4 (see Table 3.6.1). We conjecture that these results indicate that the actual pressure measurements are much less accurate than the reported pressure gauges resolutions indicate; however, there could be other causes, for example that the homogeneous reservoir wellbore storage and skin model used is not quite correct, due for example, to heterogeneities, restricted-entry or multiphase flow effects in the reservoir or in the wellbore. It is also possible that the variances of the measurements are not all equal.

CHAPTER IV

APPLICATIONS TO MORE COMPLEX MODELS

Nonlinear regression techniques are most helpful for complex reservoir/wellbore models due to the large number of reservoir/well parameters, which makes classical type-curve analysis difficult. In this chapter, we apply nonlinear regression techniques to the horizontal well and dual porosity models.

4.1 Horizontal Well Model

Measured data are modeled using the analytical solution for a uniform-flux horizontal well in an anisotropic, bounded, rectangular-shaped reservoir.[38] The well axis is parallel to the y-direction (see Fig. 4.1.1A). The derivatives of the model function with respect to the model parameters are evaluated analytically as presented in Ref. 30.

We analyze two horizontal-well tests; the first one is a synthetic drawdown data and the second is the field data set presented in Ref. 39. In analyzing these tests, we perform preliminary regressions on specific parts of the data, where one or more of the parameters have little or no influence on the pressure response. We show that this procedure tends to enhance the chance of obtaining convergence to the correct estimates for the remaining parameters, and can be used to devise better initial estimates for regressing on all parameters using the entire span of the test data. Both pressure and pressure derivative data are used in the analysis. Parts of the data that are suitable for preliminary analysis are as follows:

1) *Early radial flow period.*[38,40] This flow period is not affected by any of the boundaries, so the size of the reservoir and the position of the well are not parameters.

Moreover, during this flow period there is no flow in the y-direction, consequently, k_y is also not a parameter. Also, it is well known that we cannot obtain correct estimates of k_x and k_z individually from early radial flow data, but the product $k_x k_z$ can be determined.

2) *Boundary-dominated data.* Here, wellbore storage effects are negligible. Estimates of the pore volume of the reservoir and the drainage area can be obtained, especially if the test exhibits pseudo-steady state flow period. Although Δp_s, the pressure drop due to a mechanical skin, and k_z affect the pressure data, they cannot be determined from this flow period because they are lumped together in the pseudoskin factor.

3) *Data after wellbore storage effects become negligible.* Here, if the regression is done on the pressure derivative data, p_i and Δp_s (and C) are not parameters. Also, k_z may not have much effect on the pressure derivative data after we reach the pseudoradial flow period (in this case, k_z affects only the pseudoskin factor). In this flow period, we usually obtain estimates of the permeabilities k_x and/or k_y, depending on the reservoir flow regimes. In particular, if the data during this period includes pseudoradial data, then a good estimate of the effective horizontal permeability $k_h = \sqrt{k_x k_y}$ is obtained.

4) *Pseudoradial flow.*[38,40] In this flow period the product $k_x k_y$ can be determined with high confidence.

If we regress on a subset of the data corresponding to any one of the four preceding periods, we have fewer parameters to determine accurately. However, to apply the procedure, we must be able to determine which part of the data corresponds to which flow period. This will normally require the application of deconvolution to account for wellbore storage effects. After we regress on two or more of the data subsets listed above, we combine results to obtain initial guesses for a final regression on all data. The applicability of this procedure is demonstrated in the analysis of a synthetic data.

With the field data, we illustrate the advantages of nonlinear regression for the analysis of a multi-rate horizontal well test with a flow period approximately one tenth the duration of the shut-in time period. We also compare our results (based on nonlinear regression) with the results presented by Ref. 39 (based on classical analysis).

4.1.1 Horizontal-Well Synthetic Example

Our first example is the synthetic drawdown test, which was generated from the Odeh and Babu[38] horizontal-well solution using the data of Table 4.1.1. Note that the horizontal well is in the center of a rectangular box-shaped drainage area. Figures 4.1.1a through 4.1.1c show a schematic of the reservoir/well system, and introduce the notation for the horizontal well model. In an attempt to be more realistic, we introduced gaussian-distributed noise in the pressure data with standard deviation $\sigma = 0.25$ psi, and used the resulting pressures to obtain the pressure derivative data numerically. Figures 4.1.2 through 4.1.4 show log-log plots for this example.

The data set is analyzed for two different spans of time corresponding to two tests with different duration. In the first test, we consider data (15 points per log-cycle) in the time interval from 1×10^{-4} to 30 days; and in the second one, we consider data only in the time interval from 1×10^{-4} to 1.5 days, which eliminates the boundary-dominated flow data. The initial estimates of the parameters for the various runs are shown in Table 4.1.2. The length of the well, L_w, and its vertical position, z', are assumed to be known for this example. For all runs, *wellbore pressures* are computed at a distance r_w from the well axis, with the y-coordinate (along the well axis) given by[41] $y_w = 0.84L_w + y_1$, see Fig. 4.1.1. At this point the infinite-conductivity and uniform-flux horizontal well solutions are equivalent. Thus we use the simpler solution (uniform flux) to model the more complex infinite-conductivity case.

Table 4.1.1 - Input variables to generate the Horizontal-Well Synthetic Example (drawdown test with $q = 2500$ bbl/d)

Variable	Value
k_x (md)	10.
k_y (md)	20.
k_z (md)	0.5
L_w (ft)	500.
L_x (ft)	1250.
L_y (ft)	1750.
x' (ft)	625.
y_1 (ft)	625.
p_i (psi)	7000.
L_z or h (ft)	50.
z' (ft)	25.
C (bbl/psi)	0.020
ϕ	0.30
Δp_s (psi/bbl/d)	0.1
r_w (ft)	0.35
μ_o (cp)	1.0
c_t (psi^{-1})	1.0×10^{-5}
N(0,σ) noise, σ (psi)	0.25

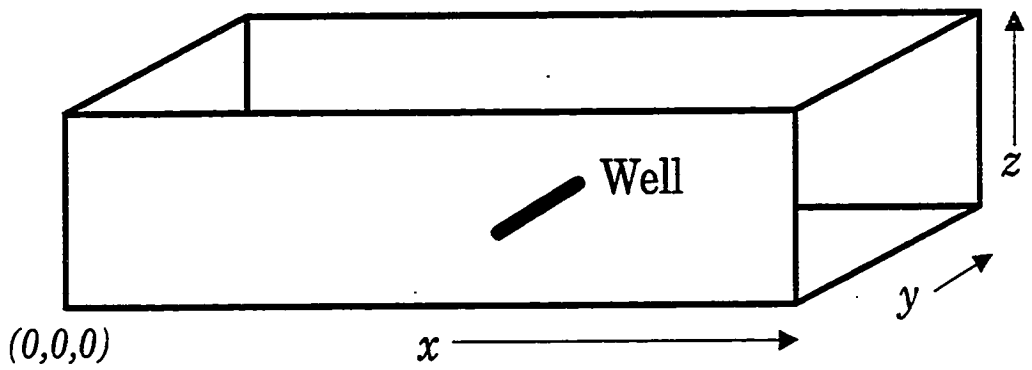

Figure 4.1.1a: 3-D Scheme of the Well-Reservoir Geometry.

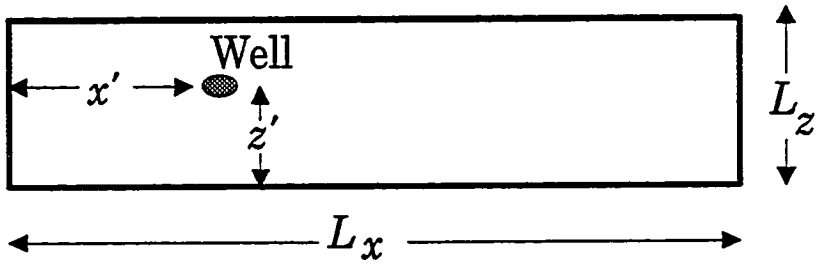

Figure 4.1.1b - Position of well axis in the *x*-*z* plane.

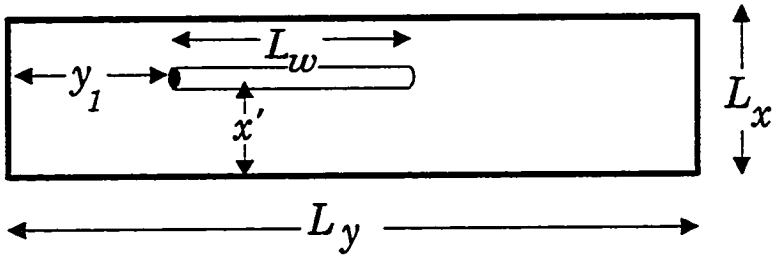

Figure 4.1.1c - Position of well axis in the *x*-*y* plane.

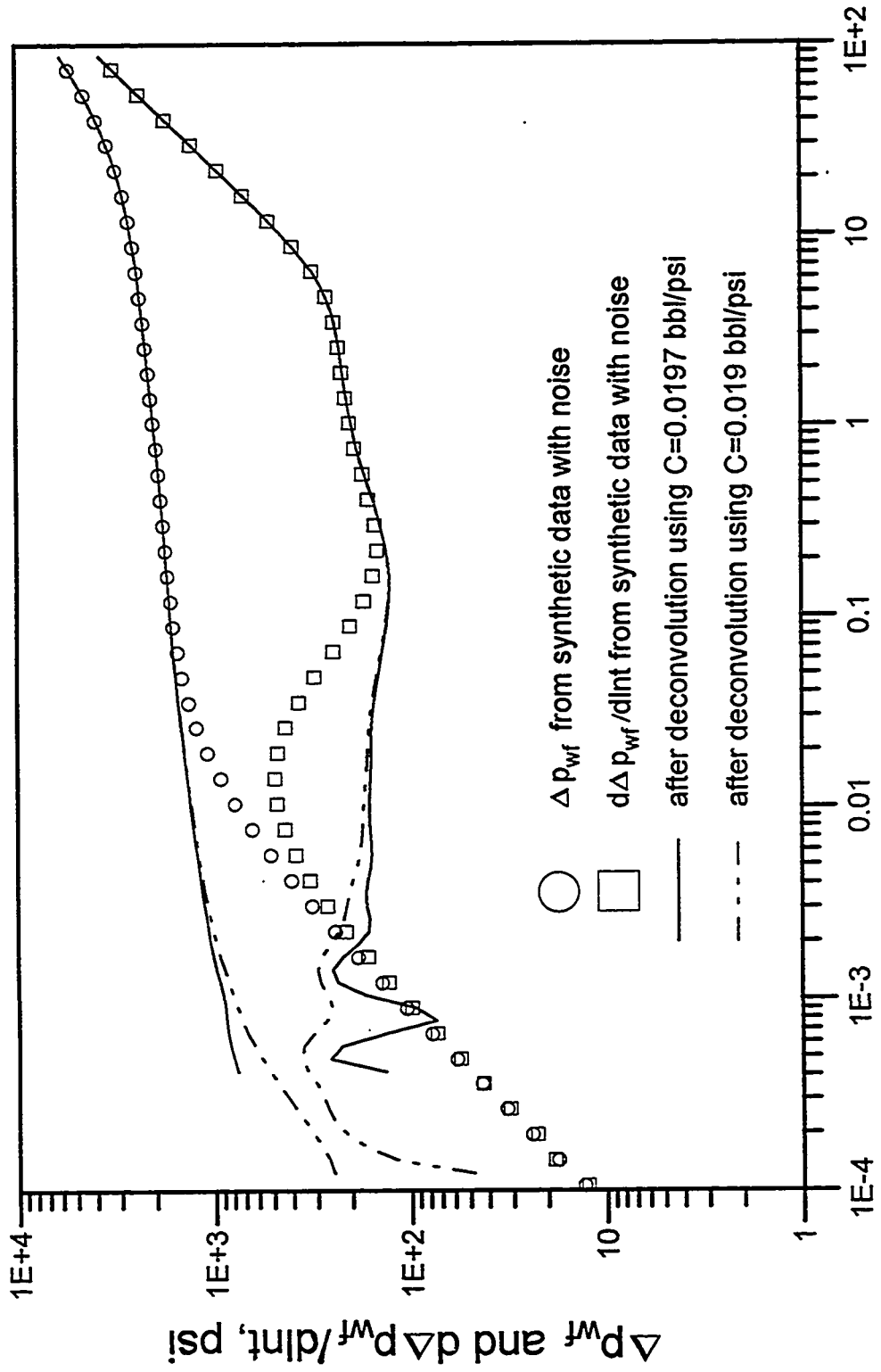

Fig. 4.1.2 - Horizontal well synthetic drawdown example

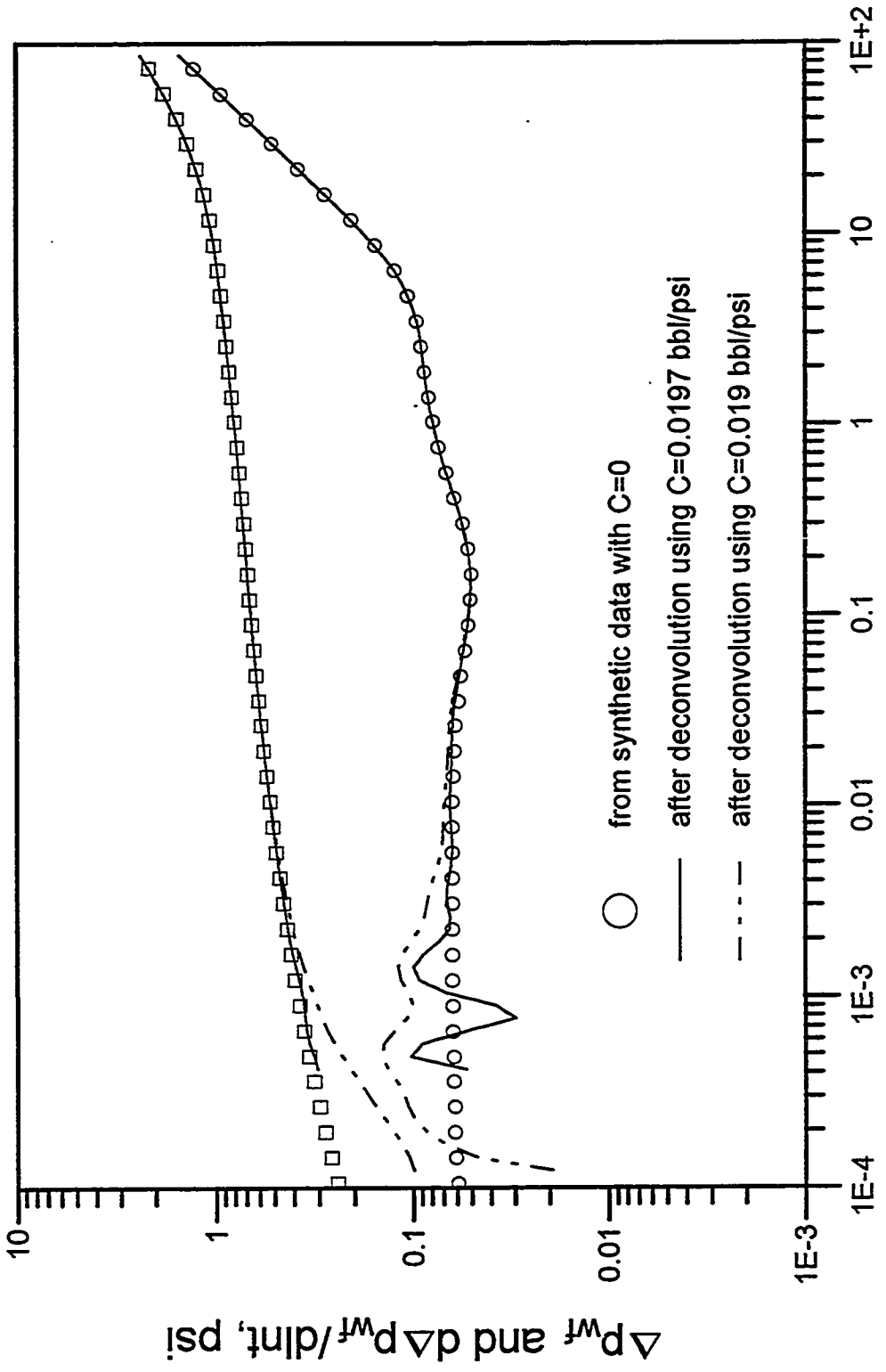

Fig. 4.1.3 - Horizontal well synthetic drawdown example

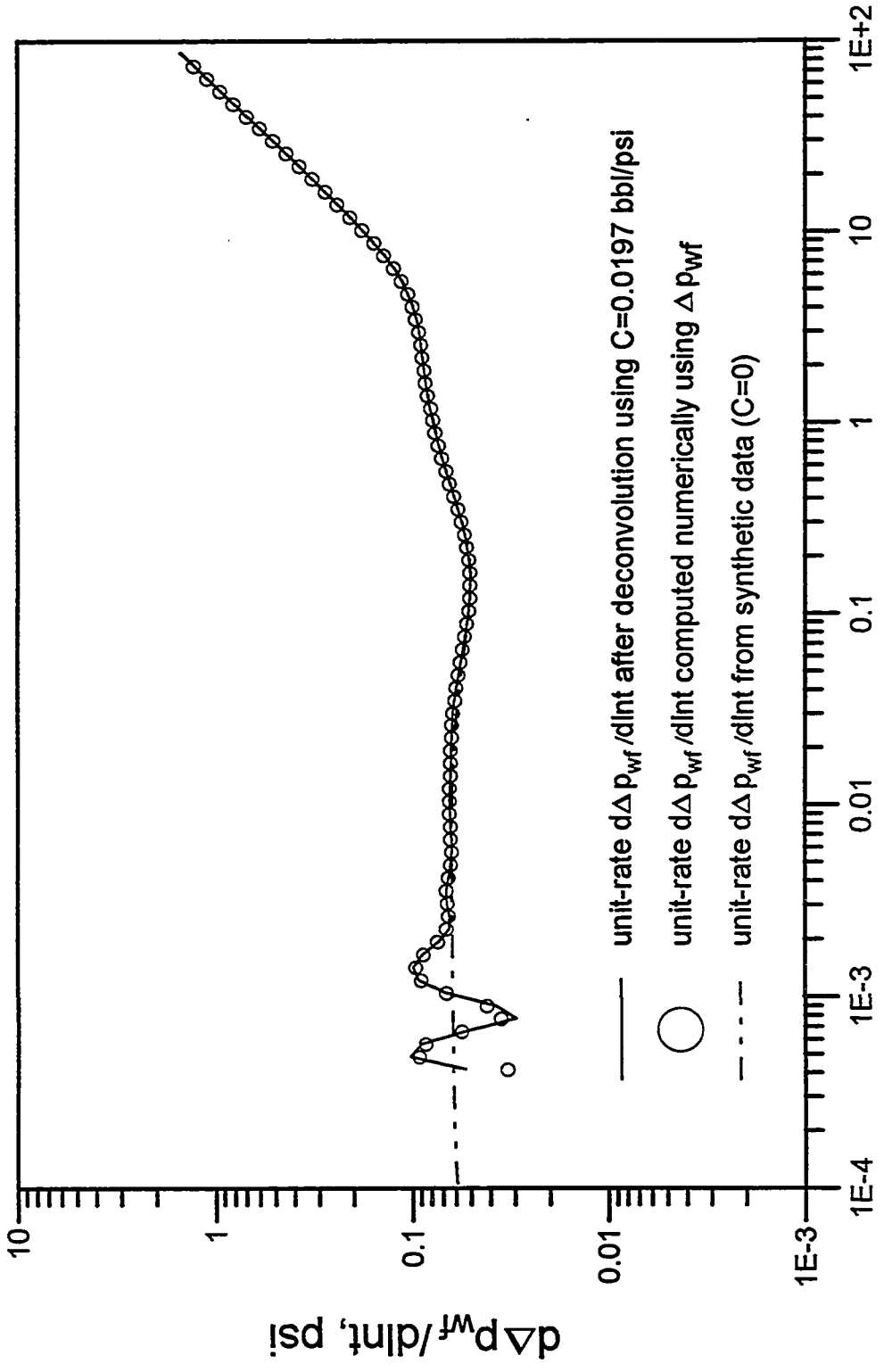

Fig. 4.1.4 - Horizontal well synthetic drawdown example comparison between pressure derivatives

Table 4.1.2 - Initial estimates of the parameters for the various runs of the Horizontal-Well Synthetic Example

Run #	k_x (md)	k_y (md)	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	L_x (ft)	L_y (ft)
actual values	10.	20.	–	0.5	0.020	0.10	7000.	1250.	1750.
A1	2.	2.	–	2.	0.015	0.	7300.	2500.	2500.
A2	2.	2.	–	2.	0.015	0.	–	2500.	2500.
B1	2.	2.	–	2.	0.015	0.	7300.	2500.	2500.
B2	2.	2.	–	2.	0.015	0.	7300.	∞	∞
B3	–	–	2.	2.	0.015	0.	7300.	∞	∞
B4	–	–	2.	2.	0.015	0.01999	–	∞	∞
C1	–	–	3.865	1.295	0.01999	0.01999	–	2500.	2500.
C2	3.865	3.865	–	1.295	0.01999	0.0999	–	2500.	2500.
D1	13.5	13.5	–	0.37	0.01999	0.0999	7000.	1346.	1620.
D2	13.5	13.5	–	0.37	0.01999	0.0999	–	1346.	1620.
E1	3.865	3.865	–	1.295	0.01999	0.0999	–	∞	∞
E2	3.865	3.865	–	1.295	0.01999	0.0999	7000.	∞	∞
E3	–	–	3.865	1.295	0.01999	0.0999	–	∞	∞
A3	2.	2.	–	2.	0.015	0.	7300.	∞	∞
A4	2.	2.	–	2.	0.015	0.	–	∞	∞
F1	12.8	12.8	–	0.39	0.01999	0.0999	7000.	∞	∞
F2	12.8	12.8	–	0.39	0.01999	0.0999	–	∞	∞

The dashed and solid lines of Figures 4.1.2 through 4.1.4 show the pressure and pressure derivative curves after performing the wellbore-storage deconvolution[12] by using $C = 0.0197$ and $C = 0.019$ bbl/psi, respectively. From Fig. 4.1.2, we estimate the duration of the early radial flow to be 0.02 days, the end of the wellbore storage effects at 0.9 days, the beginning of the boundary effects at 2 days, and the beginning of the pseudo-steady state flow by 10 days. It seems that the pseudoradial flow was reached just about when the boundaries came in. These approximate times were used to define the time intervals on which we applied the regression techniques using the initial guesses shown in Table 4.1.2. The results for the various regressions are summarized in Table 4.1.3. The range below the estimates in Table 4.1.3 corresponds to the 95% confidence limits of the parameters. Column 2 of Table 4.1.3 shows the span of data used, and the type of regression: least-squares estimation (LS) or least-absolute-value estimation (LAV) applied on pressure data (Press.) or pressure derivative data (Deriv.). Note that in some cases, for comparison purposes, instead of regressing on k_x and k_y individually, we regress instead on $k_h = k_x = k_y$ assuming an isotropic horizontal permeability k_h.

In the first two runs (A1 and A2), we applied the LS regression on pressure data and on pressure derivative data, respectively, in an attempt to determine all 8 parameters by using all the data from 1×10^{-4} to 30 days. Our regression routine fails to converge to the right answers for those two cases, because the original initial estimates are not close enough to the solution. The problem of not having good initial estimates is aggravated by the large number of parameters. In the next four runs (B1 to B4), we consider data in the time interval from 1×10^{-4} to 0.02 days only, which approximately corresponds to the early radial flow period (see Table 4.1.3). As expected, all four cases give good estimates of the correct value of the product $k_x k_z$, 5 md^2, the wellbore storage coefficient $C = 0.020$ bbl/psi, and initial pressure $p_i = 7000$psi. We also obtain a good estimate of Δp_s. We estimated $\Delta p_s = 0.0199$ psi/bbl/day, when the correct value is 0.020 psi/bbl/day. The high correlation between k_x and k_z (or k_h and k_z) confirms that the data are in the early radial

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example

Run #	Type	k_x (md)	k_y (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	L_x (ft)	L_y (ft)	Obs.
A1	LS Press. 10^{-4} to 30 days	55.9	0.0030	0.121	0.0207	0.161	6998.1	2.1×10^{6}	1.15×10^{6}	did not converge
A2	LS Deriv. 10^{-4} to 30 days	40.2	0.0036	0.163	0.0207	0.152	--	4.6×10^{4}	2.1×10^{4}	did not converge

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficient
B1	LS Press. 10^{-4} to 0.02 days	5.262 (*) 0. to 659.	2.553 0. to 5059.	0.9486 0. to 119.	0.01998 0.0191 to 0.0208	0.0992 -3.095 to 3.294	7000.0 6999.6 to 7000.5	0.0643	k_x k_z=-1.000 k_x Δp_s=1.000 k_x C=0.994 Δp_s C=0.994
B2	LS Press. 10^{-4} to 0.02 days	5.262 0. to 1322.	2.997 0. to 7989.	0.9447 0. to 237.	0.01998 0.0181 to 0.0219	0.0985 -6.401 to 6.598	7000.1 6999.3 to 7000.8	0.0597	k_x k_z=-1.000 k_x s=1.000 k_x C=0.999 Δp_s C=0.999

(*) the range below the estimates are the 95% confidence intervals

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients
B3	LS Press. 10^{-4} to 0.02 days	3.865 0. to 75.1	1.295 0. to 25.3	0.01999 0.01992 to 0.02006	0.0999 −0.196 to 0.395	7000.0 6999.8 to 7000.3	0.0552	k_h k_z=−1.000 k_z Δp_s=0.9995 k_h Δp_s= 0.9996
B4	LS Deriv. 10^{-4} to 0.02 days	2.090 0. to 13.2	2.314 0. to 14.4	0.01987 0.01960 to 0.02014	0.0920 0.0705 to 0.1134	–	1.410	k_h k_z=−0.9999 C Δp_s=0.9479

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_h (md)	L_x (ft)	L_y (ft)	Obj. Function (psi)	Correlation Coefficients
C1	LS Deriv. 2 to 30 days	–	–	13.53 13.40 to 13.67	1346.4 1286. to 1407.	1619.8 1554. to 1686.	5.041	L_x L_y=−0.9990
C2	LS Deriv. 0.9 to 30 days	9.920 9.28 to 10.56	20.07 18.2 to 21.9	–	1257.4 1248. to 1267.	1732.6 1721. to 1745.	0.676	k_x k_y=−0.9963 L_x L_y=−0.9954

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_z (md)	Δp_s (psi/bbl/d)	p_i (psi)	L_x (ft)	L_y (ft)	Obj. Function (psi)	Correlation Coefficients
D1	LS Press. 10^{-4} to 30 days	10.11 10.0 to 10.2	19.69 19.5 to 19.9	0.495 0.492 to 0.498	0.09976 0.0992 to 0.1003	7000.0 6999.9 to 7000.1	1165.6 1131 to 1200.	1879.4 1829. to 1930.	0.0553	L_x L_y −0.9998 k_x k_y −0.9852
D2	LS Deriv. 10^{-4} to 30 days	9.965 9.75 to 10.2	20.11 19.4 to 20.8	0.497 0.488 to 0.506	0.09714 0.0945 to 0.0998	–	1264.7 1254 to 1276.	1722.6 1709. to 1736.	1.303	L_x L_y −0.9945 k_x k_y −0.9670

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_h (md)	p_i (psi)	Δp_s (psi/bbl/d)	Obj. Function (psi)	Correlation Coefficients
E1	LS Deriv. 0.2 to 1.5 days	10.45 9.76 to 11.14	19.89 17.35 to 22.43	–	–	0.8779 0.789 to 0.966	4.198	k_x k_y=−0.9642
E2	LS Press. 0.2 to 1.5 days	10.89 10.47 to 11.32	18.62 17.09 to 20.15	–	8705.7 8542. to 8870.	0.997 0.927 to 1.068	0.0821	k_x k_y=−0.9842 Δp_s p_i=0.9982
E3	LS Deriv. 0.2 to 1.5 days	–	–	12.78 12.45 to 13.11	–	1.008 0.896 to 1.120	20.739	Δp_s k_h=0.505

Table 4.1.3 - Results of the regression for the Horizontal-Well Synthetic Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients / Obs.
A3	LS Press. 10^{-4} to 1.5 days	6.93	40.93	0.616	0.01978	0.06970	7000.3	–	did not converge within 100 function evaluations
A4	LS Deriv. 10^{-4} to 1.5 days	9.89 9.54 to 10.23	20.42 19.02 to 21.82	0.497 .485 to .508	0.01993 .0198 to .0200	0.09534 0.0903 to 0.1003	–	1.522	k_x k_y=–0.9715
F1	LS Press. 10^{-4} to 1.5 days	9.953 9.81 to 10.10	20.14 19.52 to 20.77	0.502 .497 to .507	–	0.09970 0.0983 to 0.1011	7000.0 6999.9 to 7000.1	0.0583	k_x k_y=–0.9875 k_x k_z=–0.9741
F2	LS Deriv. 10^{-4} to 1.5 days	9.970 9.65 to 10.29	20.06 18.79 to 21.33	0.497 .486 to .509	–	0.09725 0.0933 to 0.1012	–	1.542	k_x k_y=–0.9666 k_x k_z=–0.9193

period, where the product $k_x k_z$ is an independent parameter. The results of run B1 (which assumes a closed system) and run B2 (which assumes an infinite-acting system solution) are equivalent, indicating that L_x *and* L_y are not parameters in this time interval of the data.

In run C1, we use the pressure derivative data to regress on k_h, L_x *and* L_y during the boundary-dominated flow, $t > 2$ days. The objective of this run is to obtain a good estimate of the product $L_x L_y$. The horizontal well is assumed to be in the center of the drainage area for all cases. Run C1 provides a very good estimate of the product $L_x L_y$ around 2.18×10^6 ft^2. The high correlation between L_x and L_y indicates that the individual values of these parameters may not be accurate. In run C2 (see Table 4.1.3), we use the pressure derivative data after the effect of wellbore storage vanishes. Thus, p_i and Δp_s are not parameters. Here, k_z and C also do not affect the computed pressure derivative (k_z and C are fixed at the latest estimates). Note that the remaining parameters are very well determined with small confidence intervals.

In the final two runs (D1 and D2), we regress on all parameters using results of runs B3 and C1 as initial estimates. Note that the initial guesses for L_x and L_y are given by the results of run C1. Also, the initial guesses for k_x and k_y are set equal to the resulting value of k_h in run C1. The initial estimates for Δp_s, C and p_i are given by the results of run B3. The initial guess for k_z is calculated by using the initial guess for k_x and the value of the product $k_x k_z$ obtained in run B3. In runs D1 and D2, we obtained very good estimates (with small confidence intervals) for all parameters by regressing on the entire span of the test.

From this point on, we assume that we have measurements only up to 1.5 days (this eliminates the boundary effects from the data). In runs E1, E2 and E3 we use data in the time interval 0.2 to 1.5 days only, where k_z is not an important parameter. Here, the objective is to obtain a good estimate of the product $k_x k_y$ (or, $k_h = \sqrt{k_x k_y}$). Runs E1 and

E3 use pressure derivative data and run E2 employs pressure data to estimate a k_h of about 13 to 14 md. The primary results of runs B3 and E3 are combined to obtain a better set of initial estimates of the parameters that is used in runs F1 and F2 (see Tables 4.1.2 and 4.1.3). Note that we have obtained very good estimates for all the parameters by regressing on the entire span of the pressure data (run F1) or pressure derivative data (run F2). Moreover, the values of the products $k_x k_y$, $k_x k_z$, and $L_x L_y$ are compatible with our previous runs. For the last two runs (A3 and A4) we use all the data (up to 1.5 days) and our original initial estimates (see Table 4.1.2) to regress on all 6 parameters in just one regression. Note that whereas run A4 converged, run A3 (which uses pressure data) did not converge within 100 function evaluations. This example demonstrates that, for the horizontal well problem, pressure derivative data are also important for regression purposes. In fact, it seems that the hyper-surface of the pressure derivative objective function has more features than the pressure objective function. This example demonstrates the following:

1) The application of nonlinear regression analysis on specific parts of the test, where one or more of the parameters are not important, enhances the chance of obtaining convergence to the correct estimates for the remaining parameters, or at least a combination of some parameters (e.g., $k_x k_y$, $k_x k_z$, or $L_x L_y$). This information can be used to devise better initial estimates for a final regression on all parameters using the entire span of the test data.

2) For horizontal-well tests, which usually exhibit a long wellbore storage dominated flow, deconvolution procedures are especially important to reveal the theoretical behavior of the early time flow regimes, as well as the ending time of the wellbore storage effects.

3) The correlation coefficient between some of the parameters can be used to help define a specific flow regime, for example, during the early radial flow we expect that k_x and k_z are highly correlated.

4) Pressure derivative data are important in horizontal well applications. We theorize that the hyper-surface of the pressure derivative objective function has more features than the pressure objective function, which tends to yield faster convergence in the regression procedure.

4.1.2 Horizontal-Well Field Example

This example is the DST data presented in Ref. 39. It is a horizontal-well buildup test following a very short producing period. The pressure measurements were taken by a strain gage placed in the vertical part of the well above a multiple shut-in cable operated test valve. Actually, the strain gage collected information just below the valve, so it was possible to obtain drawdown and buildup measurements. Table 4.1.4 shows the estimated flow rate history prior to the buildup. The flow rate history was computed by Ref. 39 using the drawdown pressure data, and the known capacity of the string and the density of the incoming fluids. In Table 4.1.4, the flow rate, q_i, for a given producing time, t_i, corresponds to the average flow rate in the interval $[t_{i-1}, t_i]$ (also see Fig. 4.1.5).

Table 4.1.5 presents the pressure and pressure derivative data during the buildup. All other necessary information is shown in Table 4.1.6. The results of the analysis of Reference 39 are given in Table 4.1.7.

The relationship between the unit-rate flowing pressure response, $\Delta p_{su}(t)$, and the general wellbore pressure change, $\Delta p(t)$, that occurs in response to a variable sandface flow rate, $q_w(t)$, is given by Duhamel's principle (multirate superposition),[42]

$$\Delta p(t) = \int_0^t q_w(\tau)\Delta p'_{su}(t-\tau)d\tau. \qquad (4.1.1)$$

Table 4.1.4 - Flowrate history for the Horizontal-Well Field Example
(avg. flowrate 264.6 bbl/d)

t (hrs)	q (bbl/d)
0.0	0.0
0.0113	17.3100
0.0130	37.2742
0.0148	105.9372
0.0201	140.3264
0.0236	170.0996
0.0306	193.0642
0.0341	225.7224
0.0376	245.2250
0.0428	260.6886
0.0463	283.0762
0.0498	294.2700
0.0551	305.4638
0.0586	292.6544
0.0655	301.4248
0.0967	344.5844
0.1001	354.9704
0.1122	359.8172
0.1276	365.3564
0.1701	350.4698
0.2306	336.6218
0.2555	317.3500
0.2867	306.2716
0.3587	294.8470
0.4024	282.3838
0.4901	266.2278
0.5484	252.0336
0.6109	234.7236
0.6820	222.0296
0.7646	207.0276
0.8330	200.7960

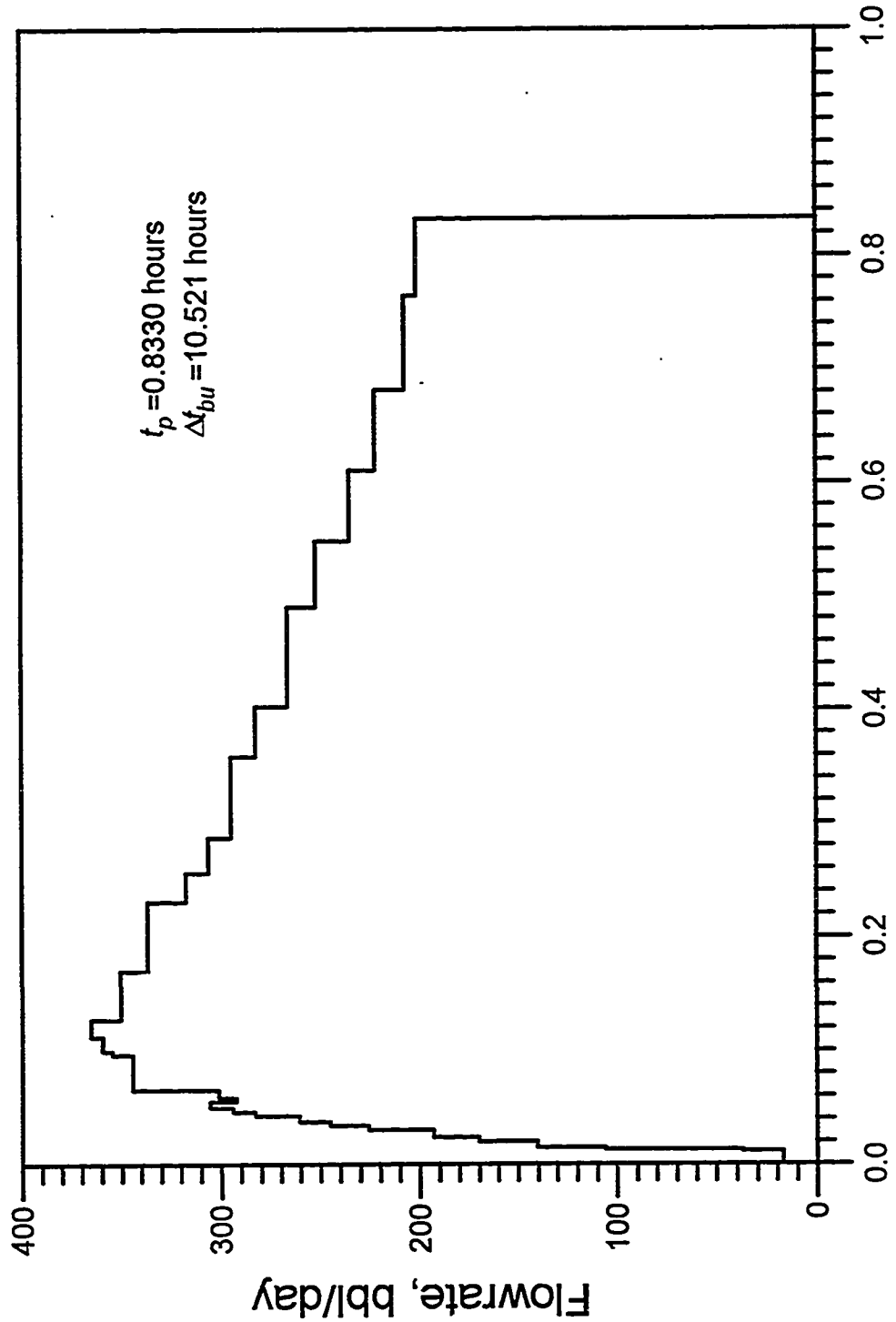

Fig. 4.1.5 - Flowrate history (Horizontal-Well Field Example)

Table 4.1.5 - Buildup pressure data for the Horizontal Well Field Example

Δt (hrs)	p_{ws} (psi)	$\frac{dp_{ws}}{d \ln \Delta t}$
0.E0	0.82747E+03	–
0.17000E-02	0.83896E+03	0.11490E+02
0.32000E-02	0.84851E+03	0.19477E+02
0.47000E-02	0.85702E+03	0.25554E+02
0.62000E-02	0.86478E+03	0.30890E+02
0.77000E-02	0.87196E+03	0.37985E+02
0.91000E-02	0.87893E+03	0.42867E+02
0.10600E-01	0.88563E+03	0.45968E+02
0.12100E-01	0.89195E+03	0.50476E+02
0.13600E-01	0.89813E+03	0.54936E+02
0.15100E-01	0.90407E+03	0.59677E+02
0.18000E-01	0.91541E+03	0.66466E+02
0.19500E-01	0.92080E+03	0.69088E+02
0.21000E-01	0.92604E+03	0.73674E+02
0.23900E-01	0.93624E+03	0.81629E+02
0.26800E-01	0.94587E+03	0.88399E+02
0.29700E-01	0.95535E+03	0.94090E+02
0.32600E-01	0.96427E+03	0.99057E+02
0.35500E-01	0.97297E+03	0.10125E+03
0.38400E-01	0.98086E+03	0.10404E+03
0.42800E-01	0.99268E+03	0.11300E+03
0.47100E-01	0.10038E+04	0.11858E+03
0.51400E-01	0.10144E+04	0.12394E+03
0.57100E-01	0.10278E+04	0.12754E+03
0.62800E-01	0.10399E+04	0.13277E+03
0.70100E-01	0.10553E+04	0.13899E+03
0.75900E-01	0.10663E+04	0.14277E+03
0.84300E-01	0.10818E+04	0.14836E+03
0.94100E-01	0.10982E+04	0.14777E+03
0.98300E-01	0.11046E+04	0.15023E+03
0.10660	0.11172E+04	0.15573E+03
0.11910	0.11345E+04	0.15774E+03
0.13290	0.11520E+04	0.15904E+03
0.14670	0.11677E+04	0.16009E+03
0.16440	0.11861E+04	0.16008E+03
0.18350	0.12035E+04	0.15851E+03

Table 4.1.5 - Buildup pressure data for the Horizontal Well Field Example (continued)

Δt (hrs)	p_{ws} (psi)	$\frac{dp_{ws}}{d\ln\Delta t}$
0.20510	0.12211E+04	0.15659E+03
0.22670	0.12367E+04	0.15307E+03
0.25360	0.12536E+04	0.15040E+03
0.28030	0.12686E+04	0.14638E+03
0.31360	0.12846E+04	0.14089E+03
0.35070	0.13002E+04	0.13579E+03
0.39300	0.13152E+04	0.12935E+03
0.44040	0.13296E+04	0.12359E+03
0.49290	0.13432E+04	0.11687E+03
0.55290	0.13562E+04	0.10997E+03
0.61940	0.13683E+04	0.10221E+03
0.69340	0.13793E+04	0.94294E+02
0.77110	0.13890E+04	0.89537E+02
0.86420	0.13990E+04	0.83369E+02
0.96600	0.14078E+04	0.74219E+02
0.10818E+01	0.14157E+04	0.69670E+02
0.12088E+01	0.14235E+04	0.65885E+02
0.13562E+01	0.14305E+04	0.59363E+02
0.15199E+01	0.14370E+04	0.55605E+02
0.17039E+01	0.14432E+04	0.48511E+02
0.19093E+01	0.14481E+04	0.42960E+02
0.21410E+01	0.14530E+04	0.41042E+02
0.24020E+01	0.14575E+04	0.36765E+02
0.26933E+01	0.14615E+04	0.33076E+02
0.30207E+01	0.14651E+04	0.29835E+02
0.33885E+01	0.14683E+04	0.27639E+02
0.37995E+01	0.14715E+04	0.27057E+02
0.42483E+01	0.14745E+04	0.23750E+02
0.47546E+01	0.14768E+04	0.21028E+02
0.53224E+01	0.14792E+04	0.20913E+02
0.59644E+01	0.14815E+04	0.21127E+02
0.66773E+01	0.14840E+04	0.21729E+02
0.74766E+01	0.14864E+04	0.18338E+02
0.83883E+01	0.14882E+04	0.20674E+02
0.93796E+01	0.14911E+04	0.24590E+02
0.10521E+02	0.14937E+04	–

In Eq 4.1.1, the subscript prime denotes derivative with respect to time. Note that $\Delta p_{su}(t)$ includes the skin factor s, i.e., denoting the fundamental unit-rate solution with zero skin by $\Delta p_u(t)$, then, for the horizontal well problem,

$$\Delta p_{su}(t) = \Delta p_u(t) + \Delta p_{skin}. \tag{4.1.2}$$

Table 4.1.6 - Reservoir/fluid parameters for the Horizontal-Well Field Example (data from reference 39)

Parameter	Value
r_w (ft)	0.27
ϕ	0.18
L_z or h (ft)	13.1
B_o (bbl/RB)	1.154
μ_o (cp)	5.6
c_t (psi^{-1})	4.79×10^{-5}
Avg. core permeability (md)	1.1

Table 4.1.7 - Results from the analysis of Horizontal-Well Field Example

Parameter	Reference 1	This work
L_w (ft)	344.	346.2
z' (ft)	4.0	4.02
k_h (md)	3.5	5.79
k_z (md)	1.75	0.753
s	0.5	–0.026
Δp_s (psi/bbl/d)	0.4644	–0.02823
p_i (psi)	1519.6 (*)	1512.6
C (bbl/psi)	0.0015	0.00126

(*) Reference 39 did not give p_i. This value was computed by LS regression with all other parameters fixed.

Duhamel's principle can be applied at the surface, or at any point in the wellbore (see, for example, Refs. 30 and 43). The downhole pressure change, $\Delta p_b(t)$, at any point in the wellbore can be written as

$$\Delta p_b(C_b,s,t)=\int_0^t q_b(\tau)\Delta p'_{bsu}(C_b,s,t-\tau)d\tau, \tag{4.1.3}$$

where Δp_{bsu} is the unit-rate response of the system including the effect of skin and the wellbore storage, C_b, which is due to the wellbore volume below the tool (measuring point). The downhole flow rate, $q_b(t)$, and $\Delta p_b(t)$ are measured at that same reference point. Note that Eq. 4.1.3 is strictly valid only for the case of constant C_b. We assume that the wellbore storage coefficient (during the drawdown and/or buildup periods) is constant, and perform the superposition by using Eq. 4.1.3. In order to simplify the notation, we delete the subscript b throughout.

Table 4.1.8 shows the initial estimates of the parameters, and Table 4.1.9 summarizes the results for the various regressions. Actually, we do not need all these runs to analyze the test. Most of them are presented for the purpose of comparing the effects of different initial estimates and analysis techniques. The value 1.1 md for the initial estimates of the permeabilities comes from the reported average core permeability. The initial wellbore storage estimate, $C=0.00127$ bbl/psi, was obtained from the slope of the early time Cartesian plot. L_w and z' are given in Ref. 39, and Δp_s and p_i are somewhat arbitrarily estimated.

In runs G1 and G2, we use data only at very early times (in an attempt to obtain the early radial flow period), where k_y and z' (see Fig. 4.1.1) are not parameters. The high correlation between k_x and k_z supports the assumption that, in the span of the test used by runs G1 and G2, the data is approximately early radial. The initial estimates for the next runs are based on the results of run G1. For most runs, we selected $k_h=k_x=k_y$ (instead of k_x and k_y individually) as a regressing parameter because this is a very short test, which tends to make k_y an ill-determined parameter. In run H3, we employed the LAV

Table 4.1.8 - Initial estimates of the parameters for the various runs of Horizontal-Well Field Example

Run #	k_x (md)	k_y (md)	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	L_w (ft)	z' (ft)
G1	1.1	–	–	1.1	0.00127	0	1510	344	4.019
G2	1.1	–	–	1.1	0.00127	0	1510	344	4.019
H1	–	–	5.701	0.7314	0.00118	–0.09295	–	344	4.019
H2	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
H3	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
H4	–	–	5.733	0.7697	0.00126	–0.03085	1512.9	344	4.019
H5	5.733	5.733	–	0.7697	0.00126	–0.03085	1512.9	344	4.019
H6	5.734	4.330	–	0.7688	0.00126	–0.03291	1512.9	344	4.019
H7	–	–	5.734	0.7688	0.00126	–0.03291	1512.9	344	4.019
H7A	–	–	5.734	0.7688	0.00126	–0.03291	1512.9	344	4.019
I1	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
I1A	–	–	1.1	1.1	0.00126	0	1510	344	4.019
I2	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
I2A	–	–	1.1	1.1	0.00126	0	1510	344	4.019
I3	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
I3A	–	–	1.1	1.1	0.00126	0	1510	344	4.019
I4	–	–	5.701	0.7314	0.00118	–0.09295	1514.5	344	4.019
J1	–	–	3.5	1.75	0.0015	0.4644	1519.6	344	4.019
J2	–	–	3.5	1.75	0.0015	0.4644	1519.6	344	4.019
J3	–	–	3.5	1.75	0.0015	0.4644	1519.6	344	4.019

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example

Run #	Type	k_x (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients
G1	LS Press. 0. to 0.8 hrs	5.701 4.54 to 6.86	0.7314 0.570 to 0.893	0.00118 0.00114 to 0.00122	−0.09295 −0.148 to −0.0381	1514.5 1512.2 to 1516.8	1.7840	k_x k_z=−0.9852
G2	LS Deriv. 0.0015 to 0.2 hrs	5.612 0.862 to 10.36	0.8350 0. to 1.736	0.001303 0.00113 to 0.00147	0.02296 −0.0703 to 0.1162	–	4.1573	k_x k_z=−0.9951

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	L_w (ft)	z' (ft)	Obj. Function (psi)	Correlation Coefficient
H1	LS Deriv. 0.0015 to 10 hrs	5.934 3.78 to 8.09	0.7505 0.627 to 0.874	0.001206 0.00115 to 0.00126	−0.1299 −0.167 to −0.0925	332.88 298.2 to 367.6	4.0308 4.017 to 4.045	4.812	k_h L_w −0.8969

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients or Outliers
H2	LS Press. 0. to 11 hrs	5.792 5.20 to 6.38	0.7349 0.664 to 0.806	0.001197 0.00118 to 0.00122	−0.06954 −0.107 to −0.0325	1513.2 1511.9 to 1514.6	1.2833	k_h k_z=−0.9851 k_h p_i=−0.9477
H3	LAV Press. 0. to 11 hrs	5.733	0.7697	0.001263	-0.03085	1512.9	1.8910	outliers #1, 2, 3, 4, 72
H4	LS Press. 0.005 to 10 hrs	5.872 5.53 to 6.21	0.7521 0.712 to 0.793	0.001264 0.00125 to 0.00128	−0.03196 −.0538 to −.0101	1512.6 1511.8 to 1513.4	0.1788	k_h k_z=−0.9837 k_h p_i=−0.9586

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients
H5	LS Press. 0.005 to 10 hrs	5.734 5.31 to 6.16	4.330 0. to 13.49	0.7688 0.719 to 0.819	0.001264 0.00125 to 0.00128	−0.03291 −.0622 to −.0036	1512.9 1511.8 to 1514.1	0.1888	k_x k_z=−0.9859 k_x p_i=−0.9687 Δp_s p_i=−0.9557 k_x k_y=0.6698

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_x (md)	k_y (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	z' (ft)	L_w (ft)	Obj. Function (psi)	Correlation Coefficients
H6	LS Press. 0.005 to 10 hrs	5.820 0 to 1160.	5.433 0. to 6717.	0.7490 0 to 1.882	0.001264 .00117 to .00136	–0.02573 –1.25 to 1.20	1512.5 1505.2 to 1519.9	4.018 0 to 8.91	346.33 0 to 34400.	0.2334	$k_x\, k_y$=–1.000 $k_x\, L_w$=–1.000 $k_x\, k_z$=–0.9980 $k_x\, z'$ =–0.9993 $k_z\, z'$ = 0.9979

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	z' (ft)	L_w (ft)	Obj. Function (psi)	Correlation Coefficients
H7 tol=10^{-4}	LS Press. 0.005 to 10 hrs	5.778 4.08 to 7.48	0.7574 0.684 to 0.831	0.001264 .00121 to .00132	–0.02803 –1.11 to 1.06	1512.7 1510.8 to 1514.5	4.018 3.84 to 4.19	345.5 311.1 to 379.9	0.2192	$k_h\, p_i$=–0.9449 $k_z\, p_i$= 0.9498 $\Delta p_s\, z'$ =–1.000 $\Delta p_s\, C$ = 0.9515
H7A tol=10^{-5}	LS Press. 0.005 to 10 hrs	5.793 4.33 to 7.26	0.7528 0.690 to 0.815	0.001264 .00121 to .00131	–0.02823 –.946 to .889	1512.6 1511.0 to 1514.1	4.018 3.87 to 4.17	346.2 316.7 to 375.7	0.1848	$k_h\, p_i$=–0.9449 $k_z\, p_i$=–0.9507 $\Delta p_s\, z'$ =–1.000 $\Delta p_s\, C$ = 0.9505

Table 4.1.9 - Results of the regression for the Horizontal-Well Field Example (continued)

Run #	Type	k_h (md)	k_z (md)	C (bbl/psi)	Δp_s (psi/bbl/d)	p_i (psi)	Obj. Function (psi)	Correlation Coefficients / Outliers
I1	LS Press. 0 to 11 hrs avg. flowrate	10.722 8.76 to 12.68	0.4597 0.385 to 0.534	0.001368 0.00132 to 0.00141	–0.3302 –0.421 to –0.239	1505.6 1503.9 to 1507.3	2.2991	k_h k_z=–0.9880 k_h p_i=–0.9268 k_h s= 0.9464 Δp_s p_i=–0.9564
I1A	LS Press. 0 to 11 hrs avg. flowrate	3.373 2.88 to 3.87	1.395 1.139 to 1.651	0.001419 0.00133 to 0.00151	–0.3320 –0.402 to –0.262	1520.3 1517.6 to 1522.9	3.5373	k_h k_z=–0.9667 k_h p_i=–0.9460 C Δp_s= 0.9606
I2	LAV Press. 0 to 11 hrs avg. flowrate	6.818	0.6829	0.001342	–0.3898	1510.4	2.2724	outliers #1, 2, 3 and 72
I2A	LAV Press. 0 to 11 hrs avg. flowrate	5.044	0.9043	0.001349	–0.4027	1514.1	2.6748	outliers #1, 2, 3,4 and 5
I3	LS Press. 0.004 to 10 avg. flowrate	6.002 5.50 to 6.51	0.7532 0.694 to 0.812	0.001305 0.00128 to 0.00133	–0.4230 –0.458 to –0.388	1512.1 1511.0 to 1513.3	0.3616	k_h k_z=–0.9809 k_h p_i=–0.9574
I3A	LS Press. 0.007 to 11 avg. flowrate	5.710 5.36 to 6.06	0.7905 0.743 to 0.838	0.001316 0.00129 to 0.00134	–0.4229 –0.448 to –0.398	1512.9 1512.0 to 1513.7	0.2987	k_h k_z=–0.9768 k_h p_i=–0.9518
I4	LS Press. 0.007 to 11 last flowrate	6.561 5.91 to 7.21	0.6039 0.551 to 0.657	0.001262 0.00125 to 0.00128	0.0635 0.0204 to 0.107	1511.8 1510.6 to 1513.0	0.2981	k_h k_z=–0.9918 k_h p_i=–0.9725
J1	LS Press. 0. to 11 hrs	4.882 4.35 to 5.41	0.8667 0.774 to 0.960	0.001206 0.00118 to 0.00124	–0.06804 –0.104 to –0.0325	1515.4 1513.8 to 1517.0	1.4800	k_h k_z=–0.9797 k_h p_i=–0.9472
J2	LAV Press. 0. to 11 hrs	5.206	0.8356	0.001258	–0.04462	1514.5	2.073	outliers #1, 2, 3 and 4
J3	LS Press. 0.005 to 11	5.617 5.28 to 5.96	0.7802 0.735 to 0.825	0.001259 0.00124 to 0.00128	–0.03911 –.0609 to –.01734	1513.3 1512.5 to 1514.1	0.2399	k_h k_z=–0.9815 k_h p_i=–0.9540

estimation using all pressure data. The first four data points and the last one were considered outiliers by our outlier detection routine (see Chapter II, Section 2.4). In run H4, we use LS estimation to regress on all data points excluding the five outliers previously detected. Note that the LS results are very close to the LAV results of run H3. Moreover, run H4 yielded very good estimates with small confidence intervals for all parameters. In run H5, we attempt to regress on k_x and k_y individually. Note that the confidence interval for k_y is relatively large and that k_x and k_z are highly correlated. These facts agree with the point that k_y is an ill-defined parameter due to the brief duration of the test. So in run H4, k_h is basically reflecting the permeability k_x instead of $(k_x k_y)^{1/2}$. In run H7, we include L_w and z' as regressing parameters. The comparison between the results of run H7 and run H4 (which does not regress on L_w and z') shows that the parameter estimates did not change, however, their confidence intervals significantly increased with the inclusion of L_w and z' in the parameter list.

In the next runs (I1 through I3A), we use the average flow rate $q = 264.6$ bbl/d, instead of the variable flow rate history shown in Table 4.1.4. If we compare the results of run I3 (which uses an average flow rate) with the results of run H4 (which uses the flow rate history), we can see that all parameter estimates are reasonably close, except for the Δp_s estimates which vary by more than a factor of 10. As expected, since the final rate at the end of the drawdown is lower than the average rate, the estimate of the wellbore storage coefficient for run H4 is lower than that for run I3. This causes an inaccuracy in the estimates of Δp_s since C and Δp_s are strongly correlated. As shown in run I4 (see Table 4.1.9), we can obtain estimates for Δp_s and C that are closer to previous results (after superposition of all flowrate information) by using the last flow rate, $q = q(\Delta t = 0)$, and adjusting the producing time as follows,

$$t = \frac{Q_{cum}}{q(\Delta t = 0)}, \tag{4.1.4}$$

where Q_{cum} denotes the cumulative production.

In the last three runs (J1 through J3), we use the results of Ref. 39 as initial estimates. This results are similar with those obtained with our initial estimates (see run H7A in Table 4.1.9). We should point out that Ref. 39 did not apply any regression procedure in their analysis, and that their analysis is based on the Goode and Thambynayagam[44] horizontal-well solution, whereas ours is based on the solution of Odeh and Babu.[38] Good and Thambynayagam modeled the horizontal well as a very thin partially penetrating vertical fracture, and used Fourier transforms to obtain the solution. In the Babu and Odeh solution, the horizontal well is represented by a line source.

Figures 4.1.6 through 4.1.11 show semilog plots and log-log plots of our match as well as Ref. 39 match. One can see that our results yield a slightly better match (run H7A) for pressure and a much better match for pressure derivative data than those of Ref. 39, even though our match was obtained by regressing on pressure data only.

4.2. Dual-Porosity Reservoir Model

We consider here the dual-porosity model as defined by Barenblatt *et al.*[45,46] and Warren and Root.[47] The dual porosity reservoir is characterized by having two mediums with different porosities and permeabilities. The first one, the fracture (or fissure) system, has high conductivity (i.e., high permeability) and transports the fluids to the producing wells. The second is the matrix system which has low conductivity and only feeds into the fractures. It behaves as a uniformly distributed source. In Appendix E, we extend the so-called Warren and Root model to anisotropic fracture systems, and show the relation between dual porosity and homogeneous solutions. We also show how to generate dual-porosity anisotropic solutions based on the knowledge of the corresponding homogeneous solutions.

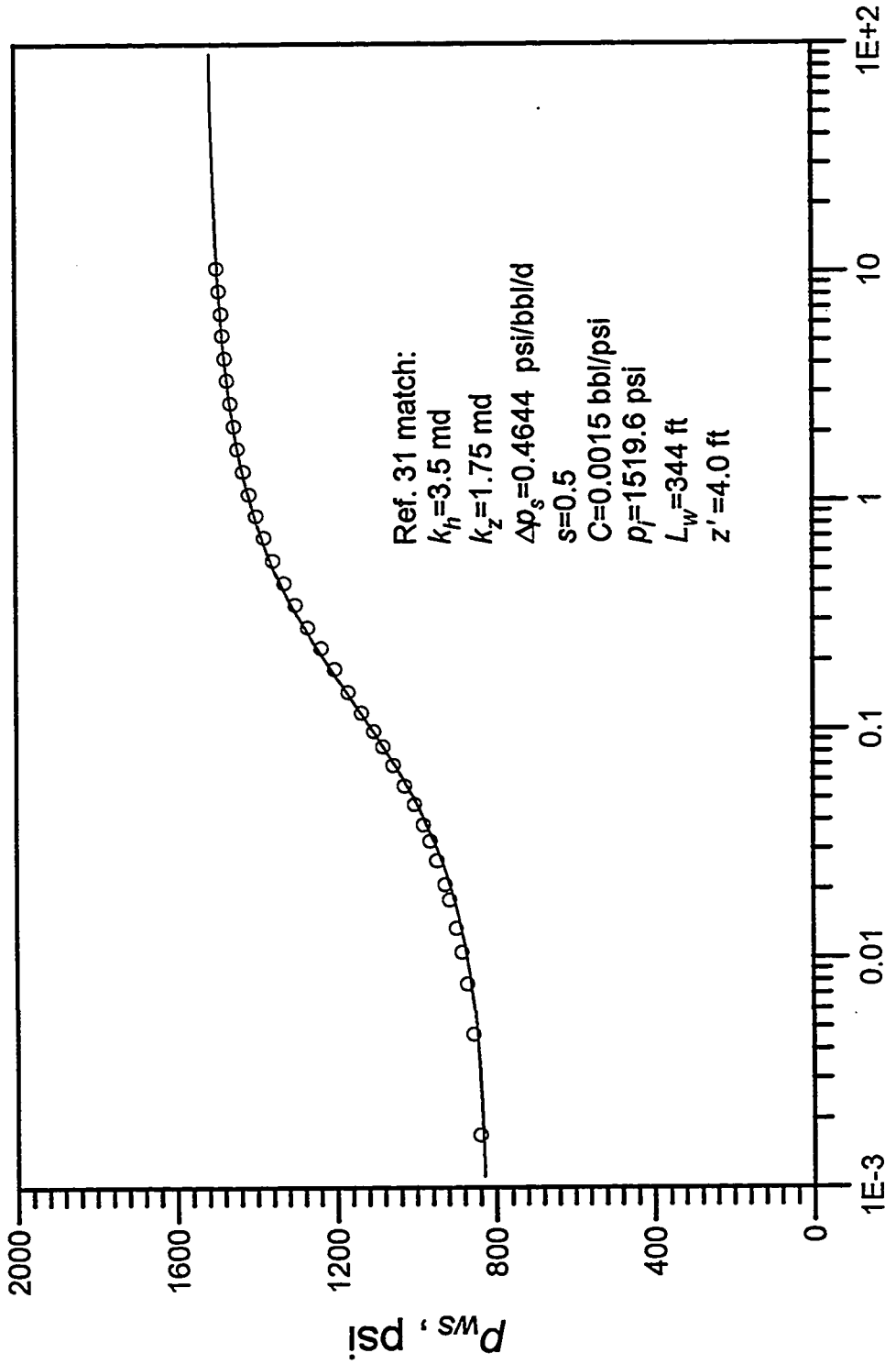

Fig. 4.1.6 - Horizontal-Well Field Example (match from Reference 31)

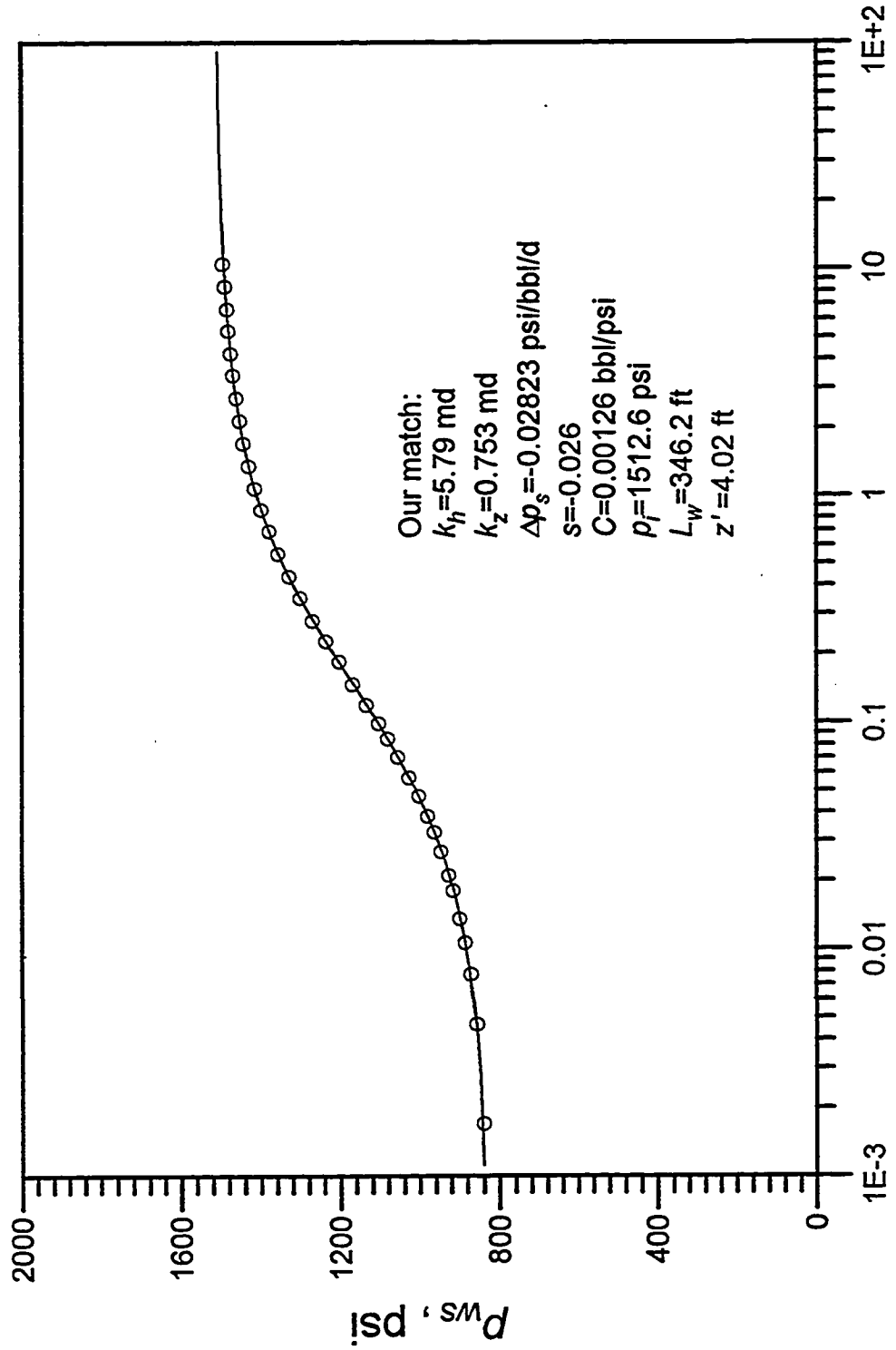

Fig. 4.1.7 - Horizontal-Well Field Example (LS regression on pressure data)

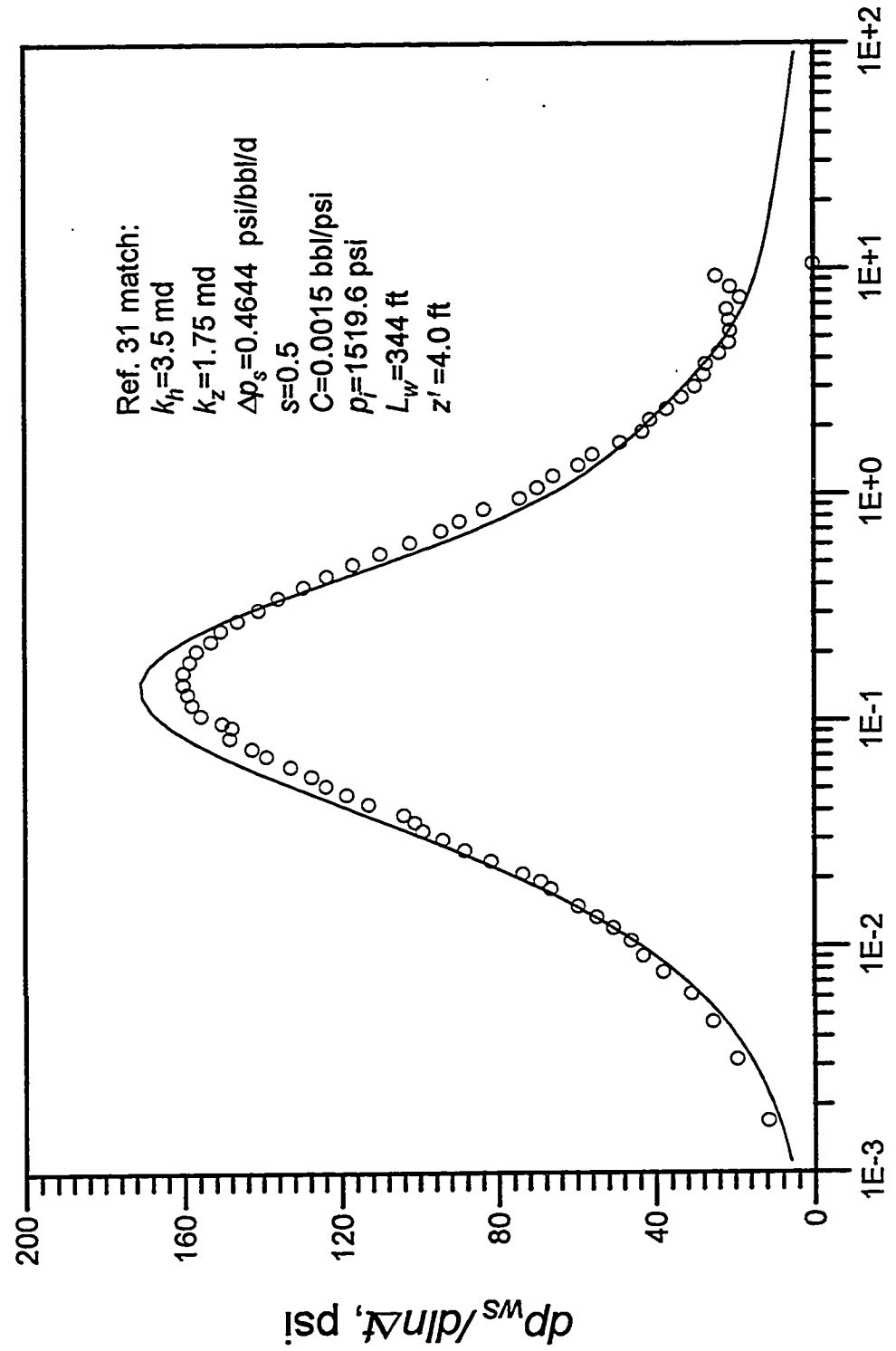

Fig. 4.1.8 - Horizontal-Well Field Example (match from Reference 31)

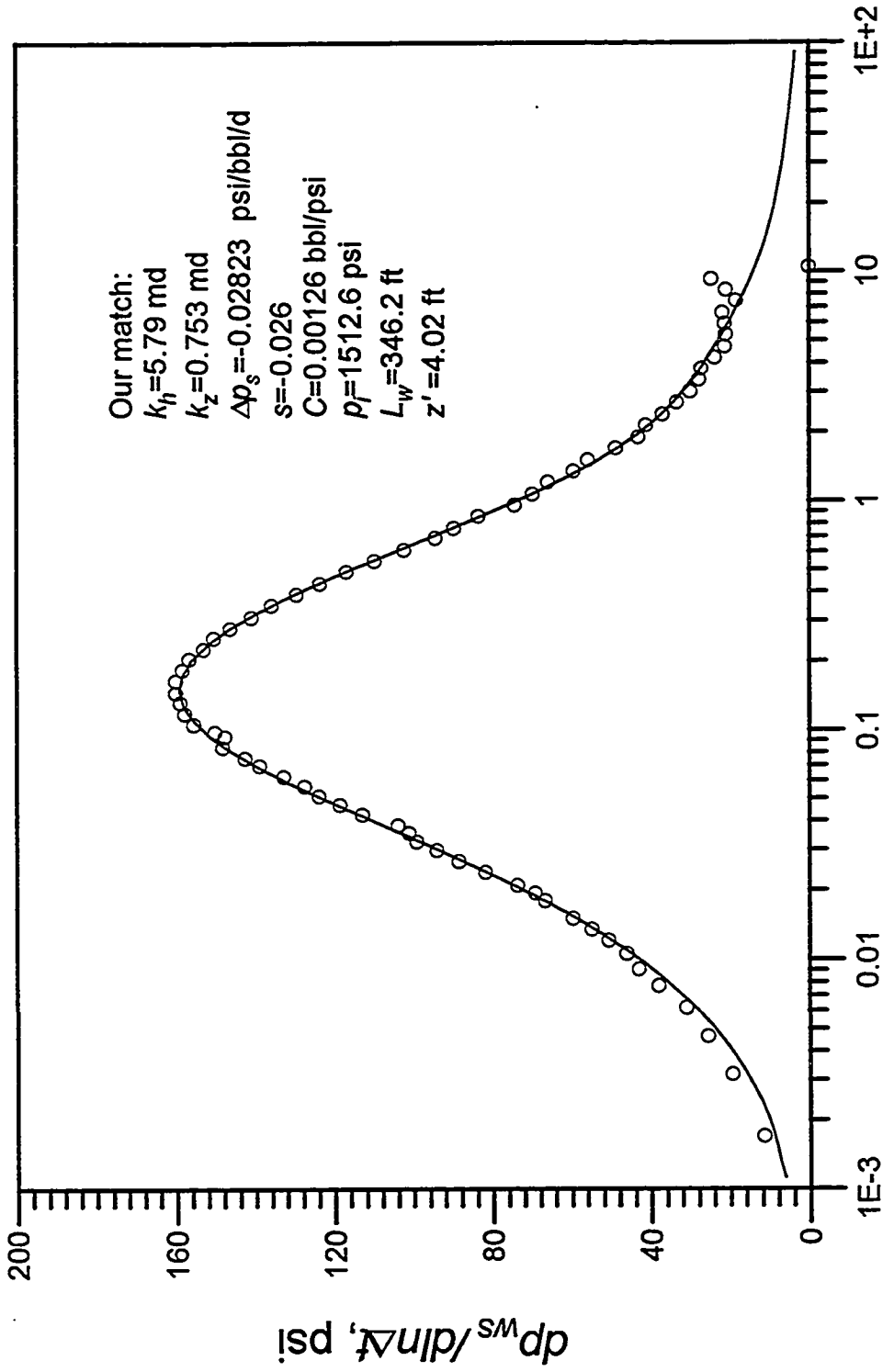

Fig. 4.1.9 - Horizontal-Well Field Example (LS regression on pressure data)

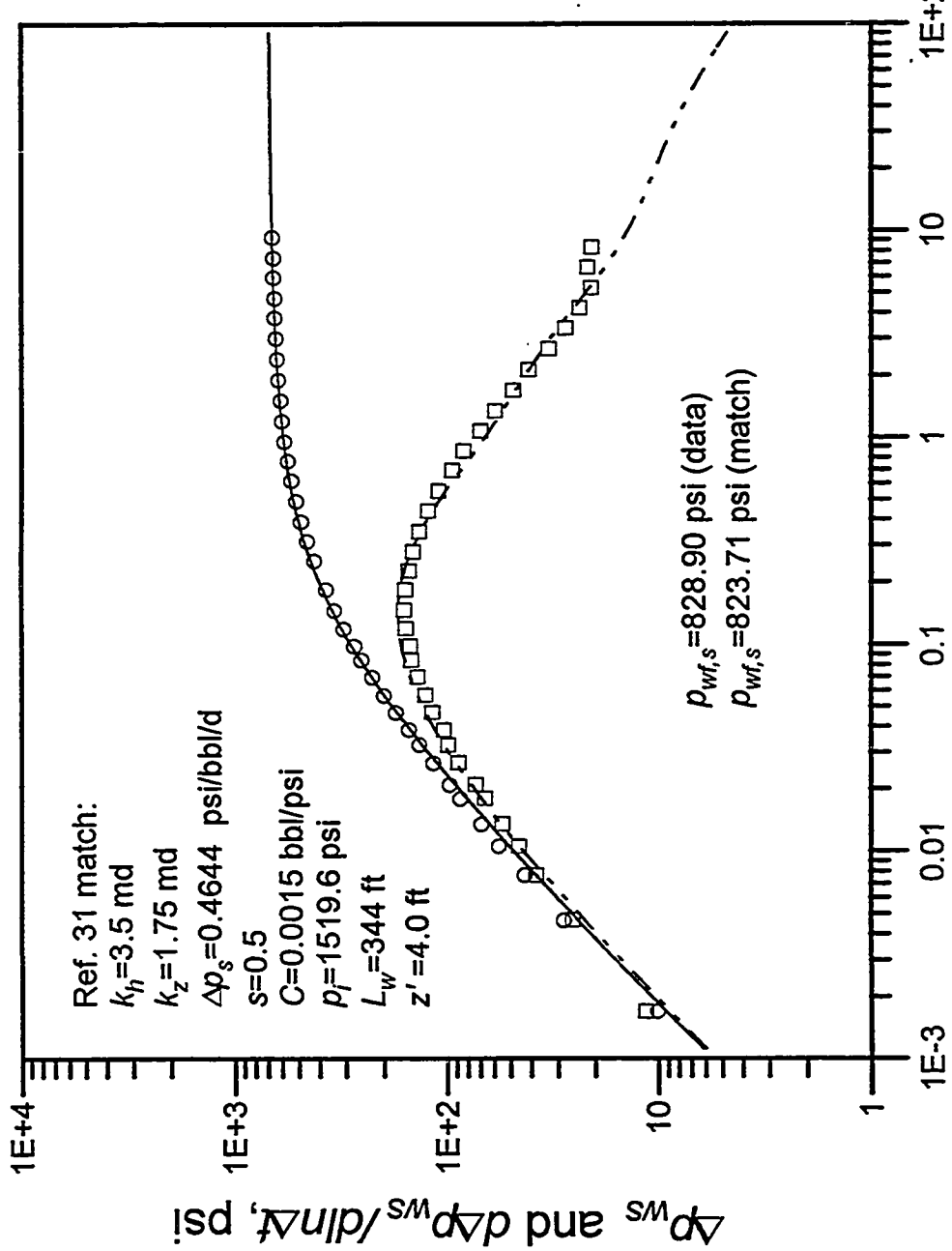

Fig. 4.1.10 - Horizontal-Well Field Example (match from Reference 31)

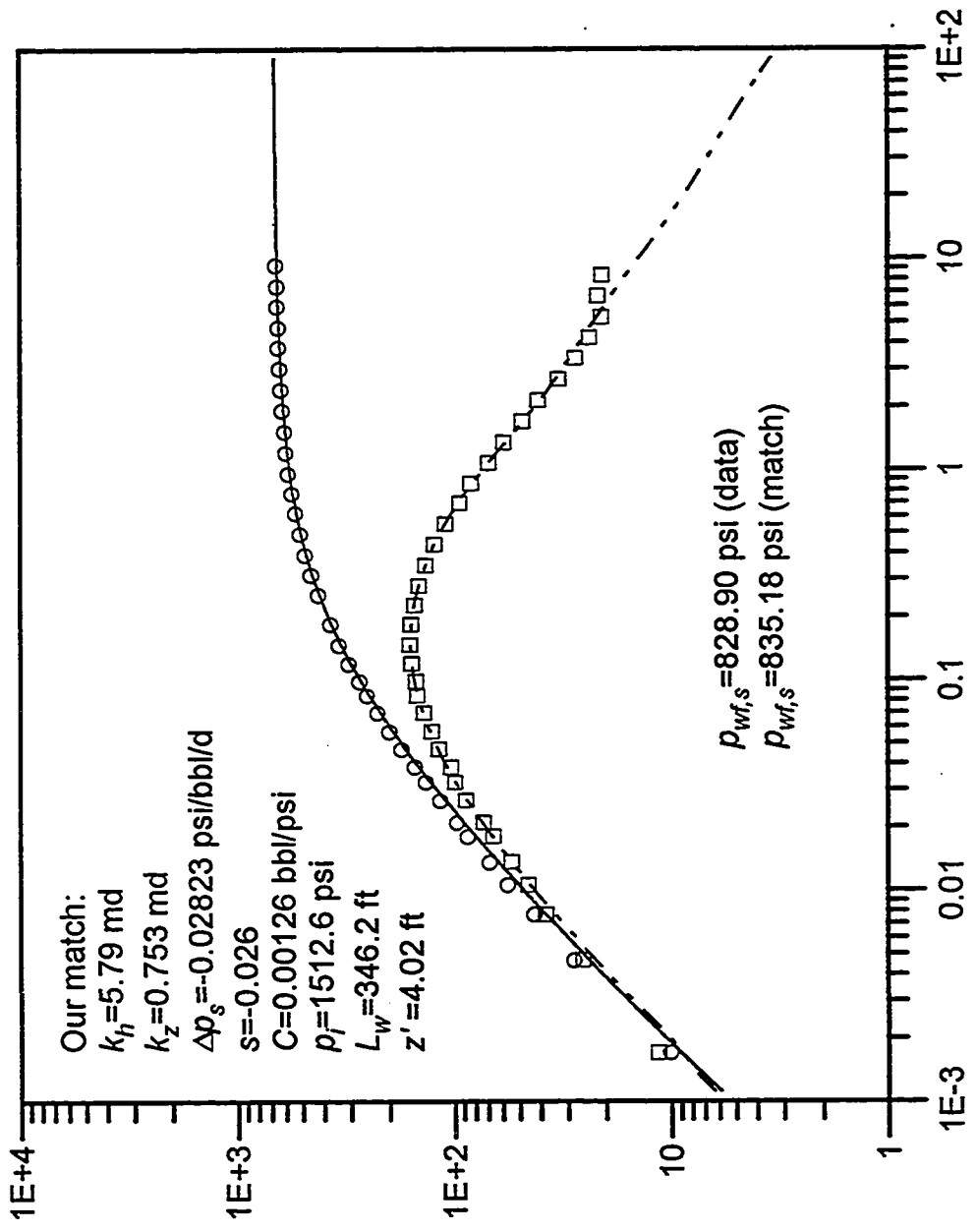

Fig. 4.1.1.11 - Horizontal-Well Field Example (LS regression on pressure)

A set of model parameters for the dual porosity model can be written as

$$\left\{\hat{k}_f, \alpha\hat{k}_m, \left(\hat{\phi}c_t\right)_f, \left(\hat{\phi}c_t\right)_{f+m}, C, \Delta p_s\right\}, \tag{4.2.1}$$

or, alternatively, as

$$\left\{\hat{k}_f, \lambda, \omega, \left(\hat{\phi}c_t\right)_{f+m}, C, \Delta p_s\right\}, \tag{4.2.2}$$

where $\hat{k}_f$ and $\hat{k}_m$ are the fracture and matrix bulk permeabilities defined in Eqs. E.6 and E.7, respectively, and α is a characteristic parameter of the fractured rock related to the shape of the matrix blocks. In Eq. 4.2.2, λ and ω are the Warren and Root dimensionless parameters defined in Eqs. E.60 and E.61.

$$\left(\hat{\phi}c_t\right)_f = \hat{\phi}_f c_{tf} \tag{4.2.3}$$

is the fracture storativity, and

$$\left(\hat{\phi}c_t\right)_{f+m} = \hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm} \tag{4.2.4}$$

is the total (fracture plus matrix) storativity. Refer to Appendix E, Eqs. E.4 and E.5.

In Eqs. 4.2.1 and 4.2.2, it is assumed that the parameters are independent; i.e., it is assumed that the partial derivatives of a given parameter with respect to the others, in the same set, are zero. We had concerns about this hypothesis in Eq. 4.2.2, since λ (see Eq. E.60) is a function of $\hat{k}_f$. However, we successfully implemented regressions for both sets of parameters with similar results. The relevant expressions for the derivatives with respect to the parameters are presented in Appendix E.

4.2.1 Dual Porosity Synthetic Example #1

This example is a synthetic drawdown test, which was generated from the Warren and Root[47] model using the data of Table 4.2.1. With this example, we demonstrate that

the use of pressure derivative data can improve convergence of the regression procedure to the correct solution. In an attempt to be more realistic, we introduce Gaussian-distributed noise in the pressure data with standard deviation of the errors $\sigma = 4$ psi (using -1 as the value of the seed for the random sequence). Pressure derivative data are computed numerically using the noisy pressure data.

Table 4.2.1 - Input variables to generate the data for the Dual-Porosity Synthetic Examples

Fracture Permeability:	$\hat{k}_f = 5$ md
Inter-porosity Flow Coefficient:	$\lambda = 3\times10^{-7}$
Fracture-total Storativity Ratio:	$\omega = 0.035$
Wellbore Storage Coefficient:	$C = 0.003$ bbl/psi
Skin Factor:	$s = 0$
Thickness:	$h = 20$ ft
Viscosity:	$\mu = 1.0$ cp
Wellbore Radius:	$r_w = 0.25$ ft
Total Compressibility:	$c_t = 1\times10^{-5}$ psi^{-1}
Flow Rate:	$qB = 100$ bbl/day
Initial Pressure:	$p_i = 10000$ psi
STD Deviation of $N(0,\sigma)$ Noise:	$\sigma = 4$ psi
Stehfest Number	8
# of Data Points per Cycle:	10

Fig. 4.2.1 shows the log-log plot for this example. Note that the pressure derivative data are deliberately very noisy.

As initial estimates of the parameters, we used all 54 possible combinations of the following: $\hat{k}_f$ =2, 10 and 20 md; $\lambda = 1.\times10^{-6}$, $1.\times10^{-7}$ and $1.\times10^{-8}$; ω =0.01, 0.05 and

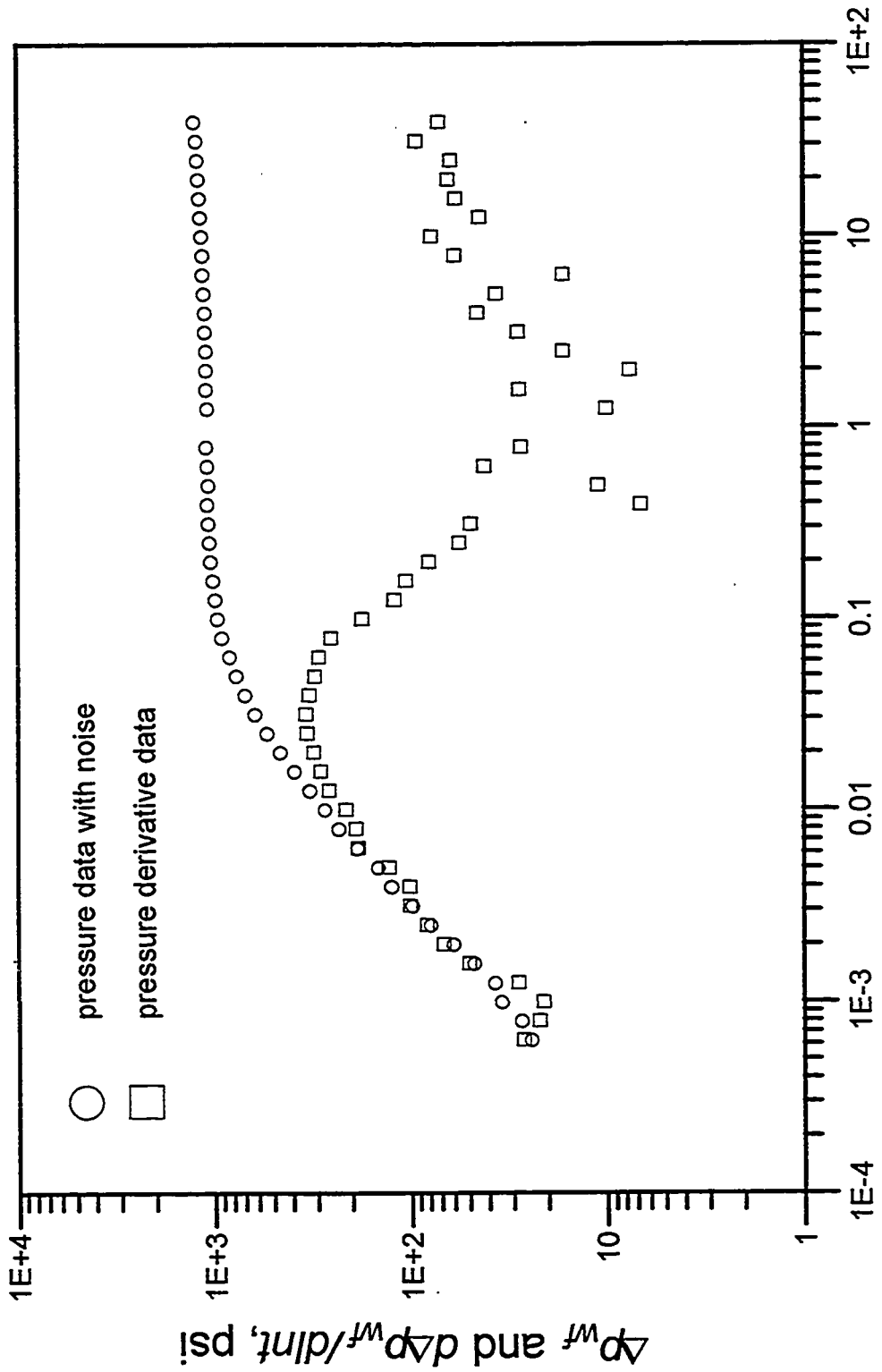

Fig. 4.2.1 - Dual-porosity synthetic drawdown example (seed=–1)

0.15; $\Delta p_s = -0.1$ and 0.3 psi/bbl/d; and $C = 0.03$ bbl/psi. Note from Table 4.2.1 that the actual values (without noise) are $\hat{k}_f = 5$ md, $\lambda = 3.\times 10^{-7}$, $\omega = 0.035$, $\Delta p_s = 0$ and $C = 0.03$ bbl/psi. Table 4.2.2 summarizes the results of the various runs.

Table 4.2.2 - Results from regressions on pressure data and pressure derivative data
Dual Porosity Synthetic Example #1

	Regression on Pressure Data	Regression on Pressure Derivative Data
$\hat{k}_f$ (md)	4.95 4.51 to 5.38	4.93 4.06 to 5.81
C (bbl/psi)	0.00297 0.00293 to 0.00301	0.00296 0.00279 to 0.00314
Δp_s (psi/bbl/d)	−0.131 −1.13 to 0.872	−0.147 −2.20 to 1.91
λ	2.91×10^{-7} 2.62×10^{-7} to 3.21×10^{-7}	2.92×10^{-7} 1.85×10^{-7} to 3.99×10^{-7}
ω	0.0345 0.0254 to 0.0435	0.0371 0.0105 to 0.0636
avg. # of functions evaluations	19.8	14.6
avg. # of gradients evaluations	19.0	10.8
convergence to right solution (%)	72.2	92.6
combinations that did not converge to the right solution	1211, 1212, 2111, 2112, 2132, 3111, 3112, 3121, 3122, 3131, 3132, 3221, 3222, 3231, 3232	2112, 3121, 3122, 3131, 3132

In the last row of Table 4.2.2, we present the combinations of parameters for which the regression procedure did not converge to the right solution; e.g., combination of parameters 1211 corresponds to $\hat{k}_f = 2$ md, $\lambda = 1.\times 10^{-7}$, $\omega = 0.01$, $\Delta p_s = -0.1$ psi/bbl/d,

i.e., first initial guess for $\hat{k}_f$, second initial guess for λ, and first initial guesses for ω and Δp_s. Rows 1-5 of Table 4.2.2 show the final parameter estimates when the regression converged to the correct solution.

The following points should be noted:

1. The results from the regression on pressure data were slightly closer to the original input values (with smaller confidence intervals), which suggests that the pressure data is less influenced by noise than is pressure derivative data.
2. Regression on pressure data failed to converge to the right solution in 15 cases out of 54 (72.2% convergence) whereas regression on pressure derivative data failed only in 5 cases out of 54 (92.6% convergence).
3. For the 5 failures obtained when regressing on pressure derivative data, the initial estimates for λ and $\hat{k}_f$ were bigger than the input values.

This example suggests that for dual porosity models, we should first regress on pressure derivative data, and then use the results of this run as initial estimates for the final regression on pressure data.

4.2.2 Dual Porosity Synthetic Example #2

It is a common practice in well testing to smooth noisy pressure derivative before performing a type-curve matching. In this example, we show that smoothing procedures are not recommended in regression analysis, since they may contaminate the pressure derivative data.

The input data for this example are the same as shown in Table 4.2.1. We generate four different data sets corresponding to four different seeds to start the random sequence of the errors. The seed values are –1, –15, –225 and –3375. Note that each set of data is a

different realization of the same transient test. Each set of data is analyzed in three ways. Table 4.2.3 presents the parameter estimates for the various runs, and Figs. 4.2.2 through 4.2.13 show the log-log plots of the results. First, we regress on pressure derivative data without smoothing the data. Second, we regress on pressure derivative data after smoothing the data. Third, we regress on the noisy pressure data. The smoothing is performed by using Bourdet's smoothing algorithm with $L = 0.6$.

In Figs. 4.2.2 through 4.2.13, the solid lines are the synthetic pressure derivatives without noise and the dashed lines are the results from the regression fit. Note that regression on pressure data yielded better results (even for the pressure derivative match), which confirms the results from the previous example. For regression on pressure derivative data, the data without smoothing produced better results; i.e., the dashed derivative lines for this case better fit the actual derivative (also see Table 4.2.3). For example, note in Fig. 4.2.3 that the square symbols (which represent the smooth derivatives) do not properly follow the actual pressure derivative data (solid line), especially during the transition. The noisy pressure derivative (see the square symbols in Fig. 4.2.2) shows much more variability, but better fit the actual response. The results show that it is not necessary to smooth the data; indeed, the regression will provide the best possible smooth pressure derivative if the model is right.

4.2.3 Dual Porosity Field Example

Our field example is the build-up test presented by Ref. 51 in which the well was produced at three different rates prior to shut in (see Table 4.2.4). All other information is presented in Table 4.2.5.

Table 4.2.3 - Dual-Porosity Synthetic Example - Results from regression on pressure derivative data

parameters	simulated	seed = −1		seed = −15		seed = −225		seed = −3375	
		after smoothing	no smoothing	after smoothing	no smoothing	after smoothing	no smoothing	after smoothing	no smoothing
$\hat{k}_f$ (md)	5.	4.33	4.90	4.19	4.69	4.43	5.28	4.32	5.22
C (bbl/psi)	0.003	0.00286	0.00296	0.00282	0.00291	0.00292	0.00304	0.00293	0.00305
Δp_s (psi/bbl/d)	0.	−1.87	−0.276	−2.21	−0.645	−1.53	0.624	−1.88	0.521
λ	$3.\times10^{-7}$	2.44×10^{-7}	2.90×10^{-7}	2.61×10^{-7}	3.05×10^{-7}	2.73×10^{-7}	3.32×10^{-7}	2.55×10^{-7}	3.30×10^{-7}
ω	0.035	0.032	0.036	0.033	0.037	0.0338	0.040	0.030	0.040

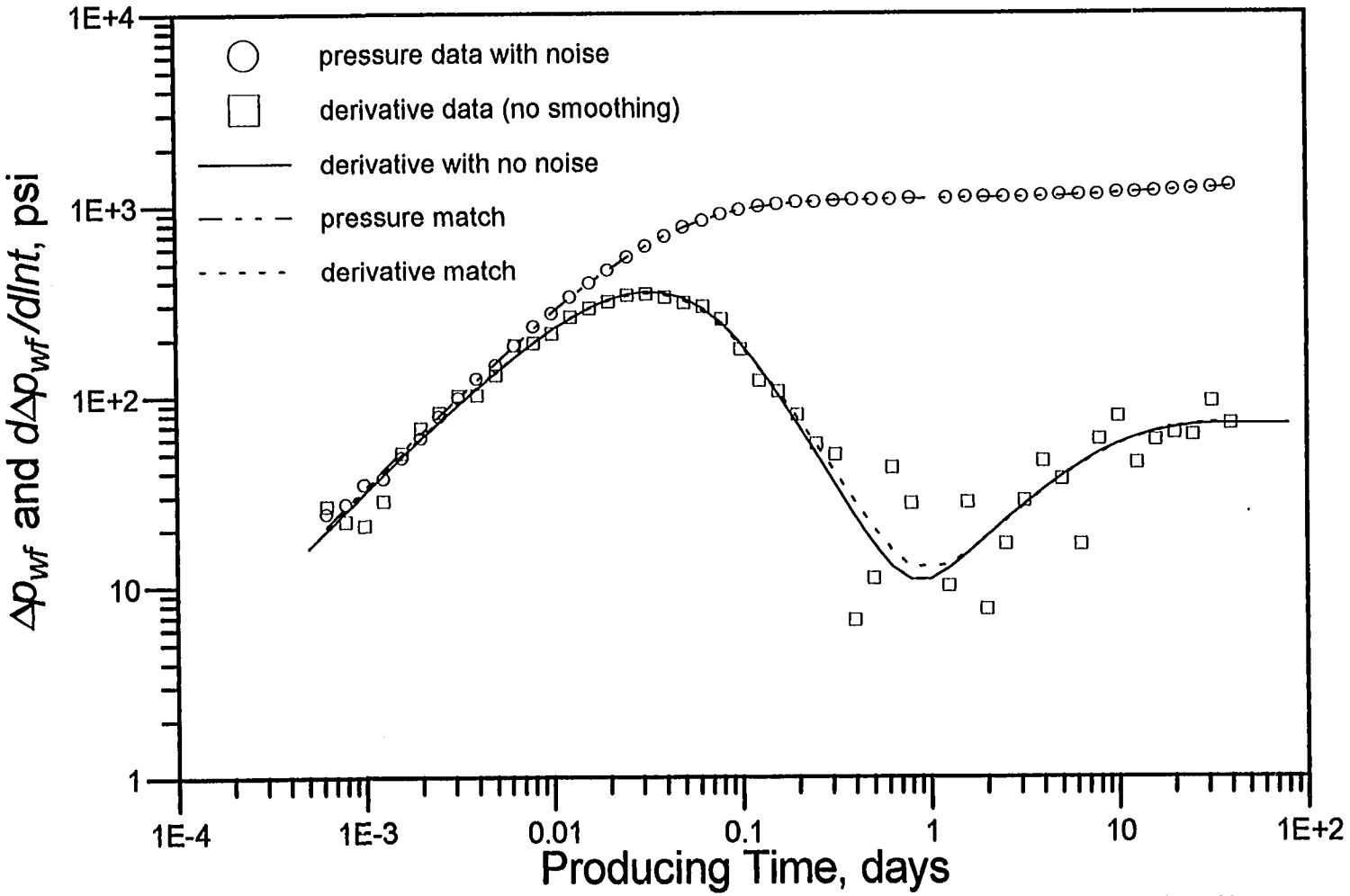

Fig. 4.2.2 - Dual-porosity synthetic drawdown example (seed=–1) results from regression on pressure derivative data

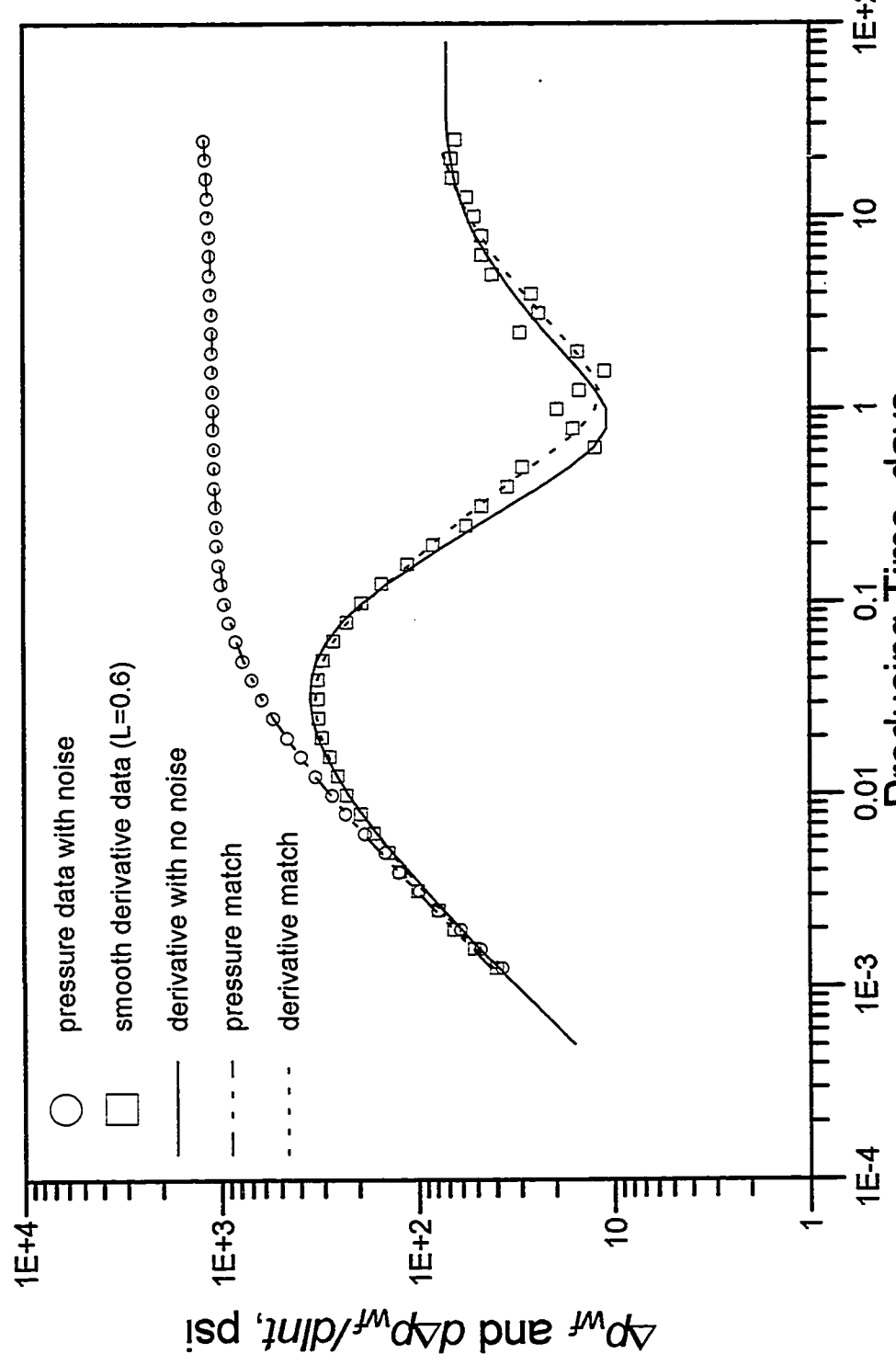

Fig. 4.2.3 - Dual-porosity synthetic drawdown example (seed=–1) results from regression on pressure derivative data

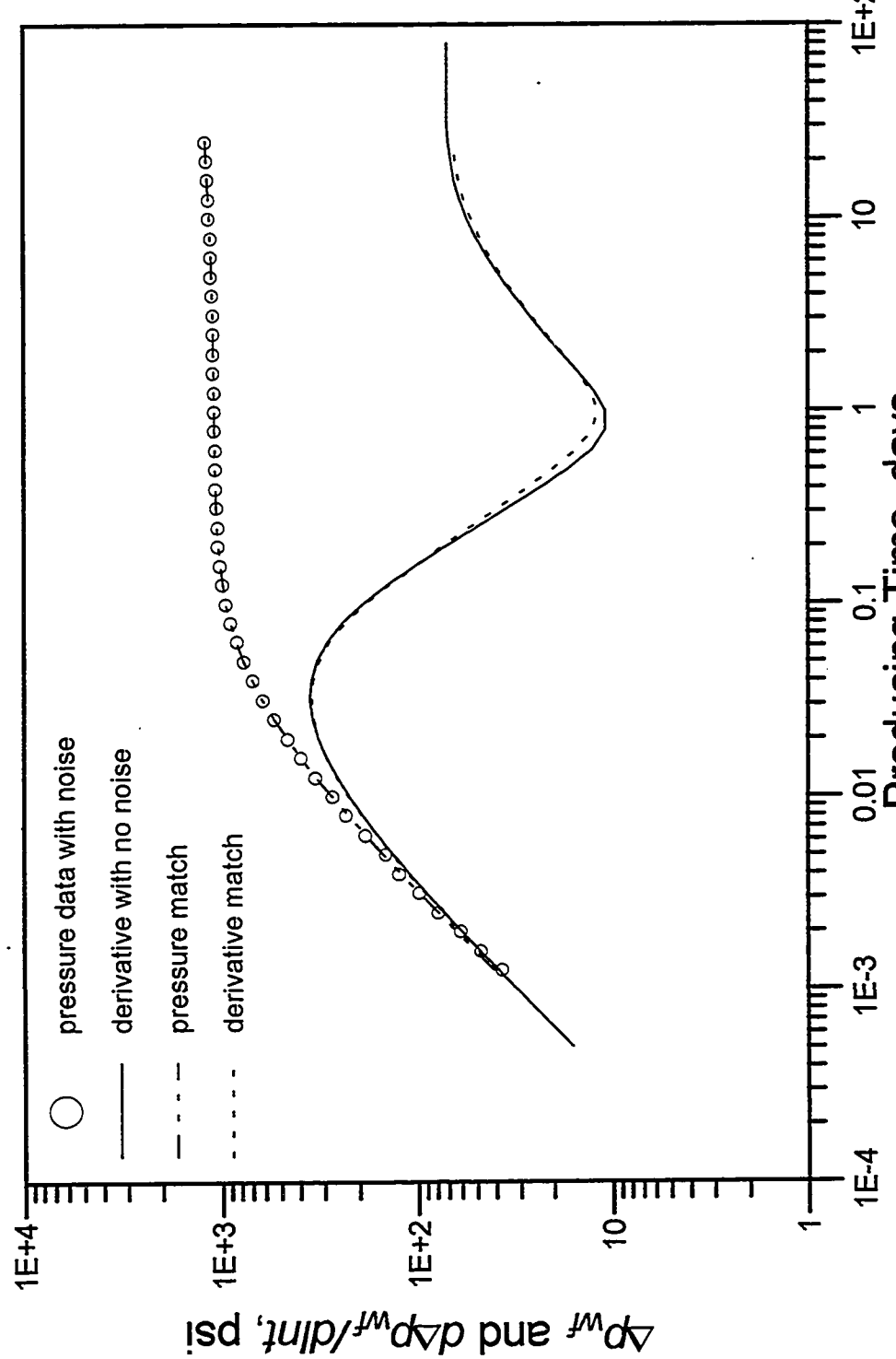

Fig. 4.2.4 - Dual-porosity synthetic drawdown example (seed=-1) results from regression on pressure data

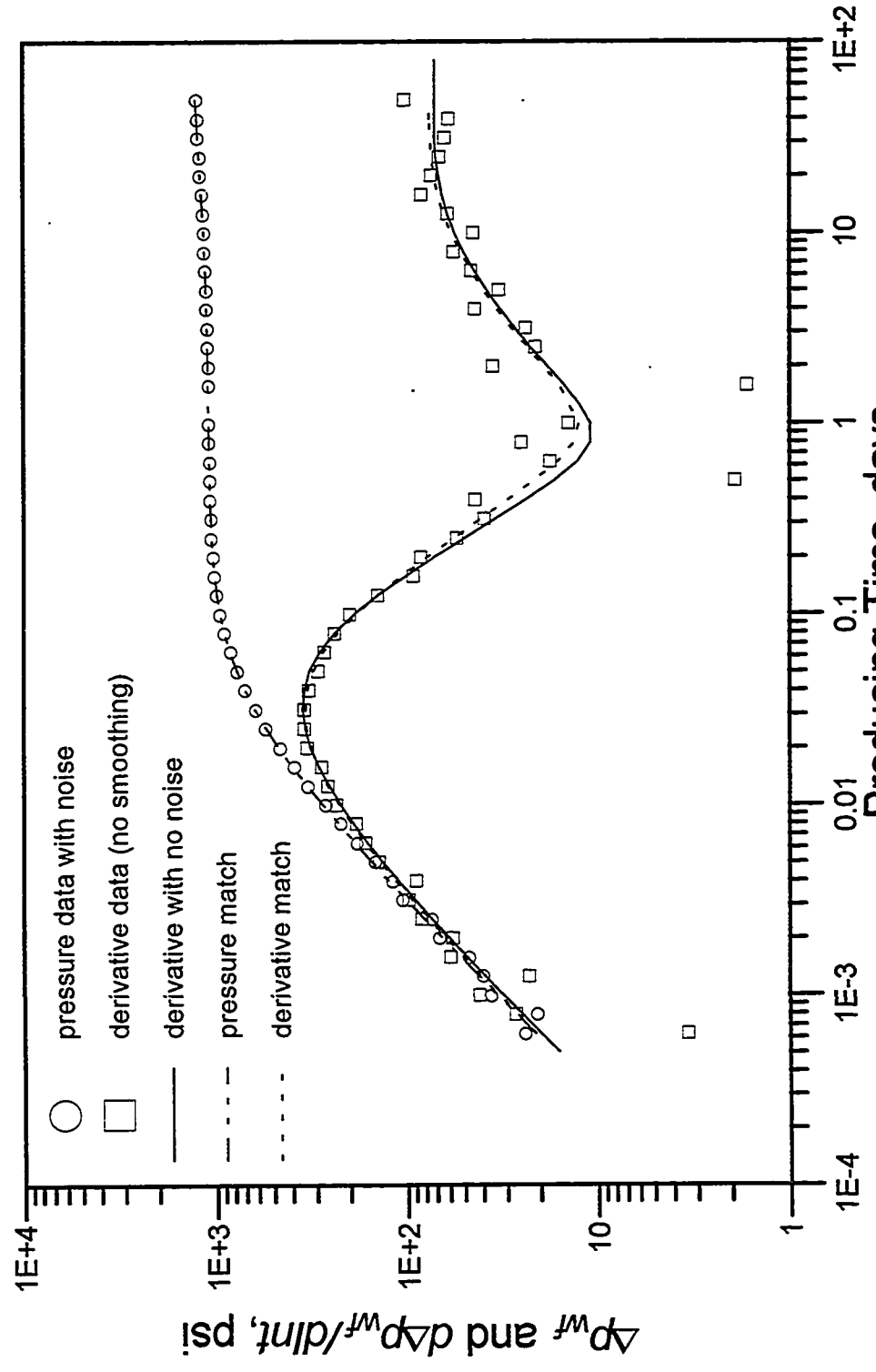

Fig. 4.2.5 - Dual-porosity synthetic drawdown example (seed=–15) results from regression on pressure derivative data

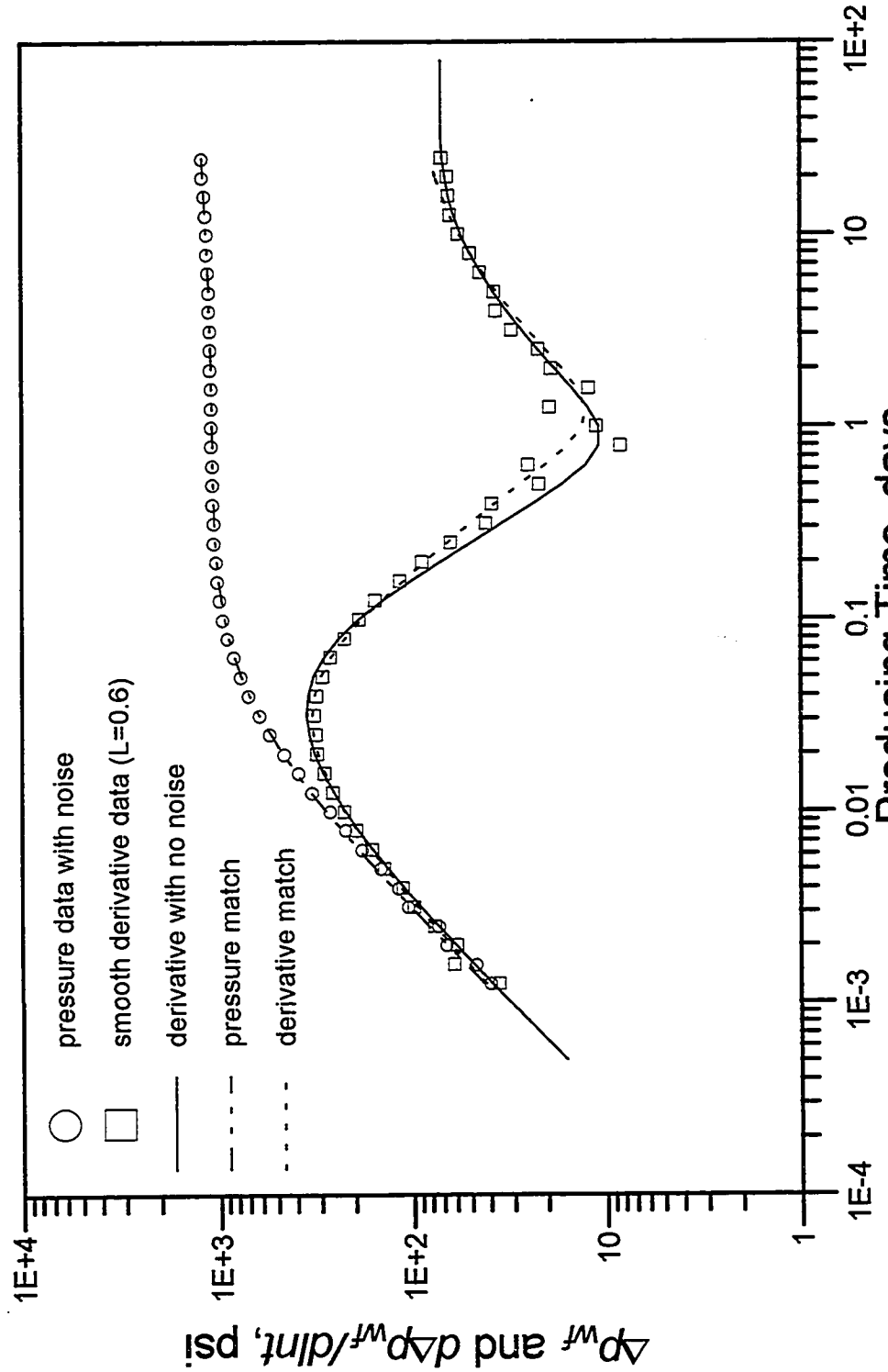

Fig. 4.2.6 - Dual-porosity synthetic drawdown example (seed=–15) results from regression on pressure derivative data

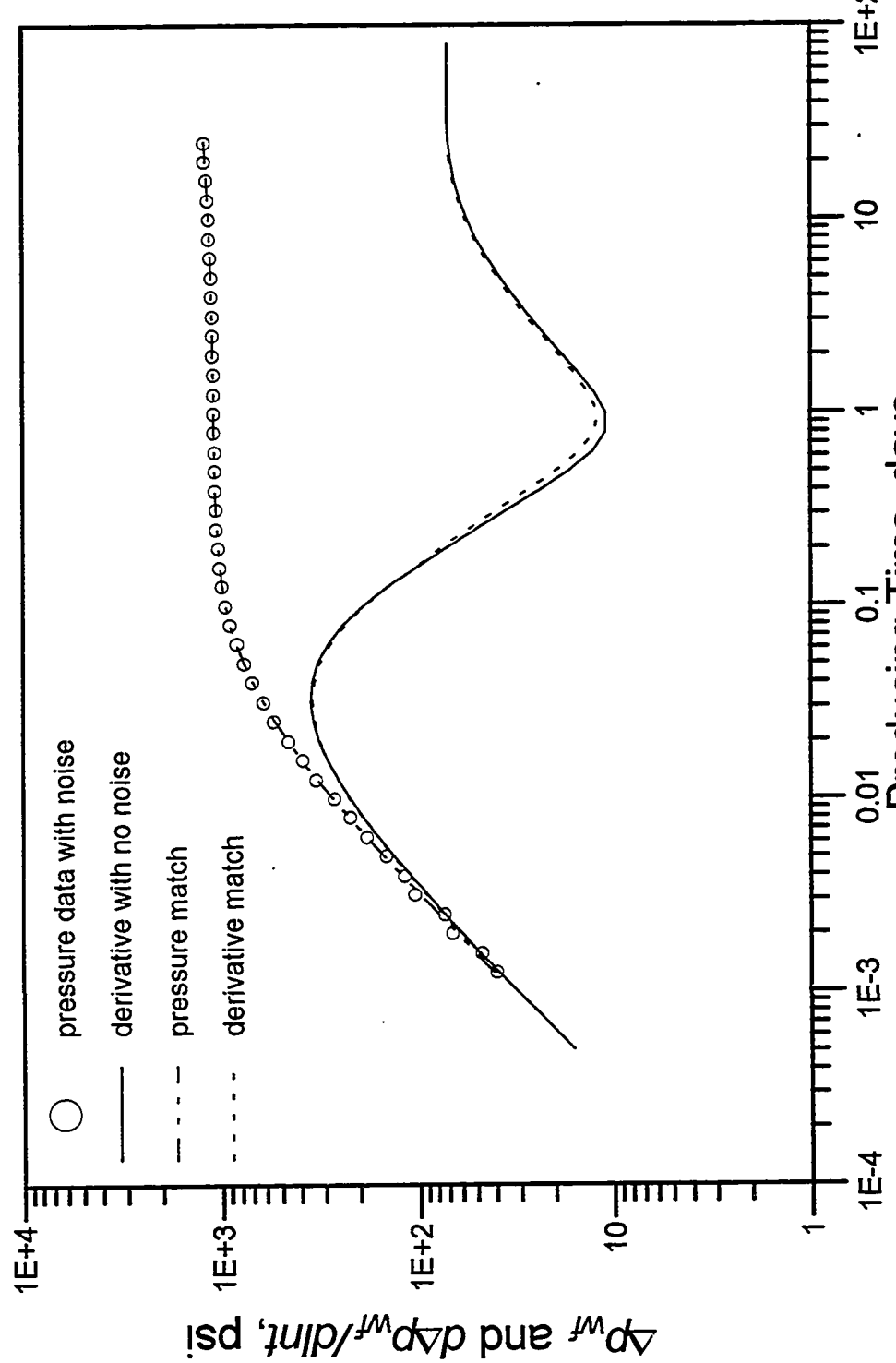

Fig. 4.2.7 - Dual-porosity synthetic drawdown example (seed=–15) results from regression on pressure data

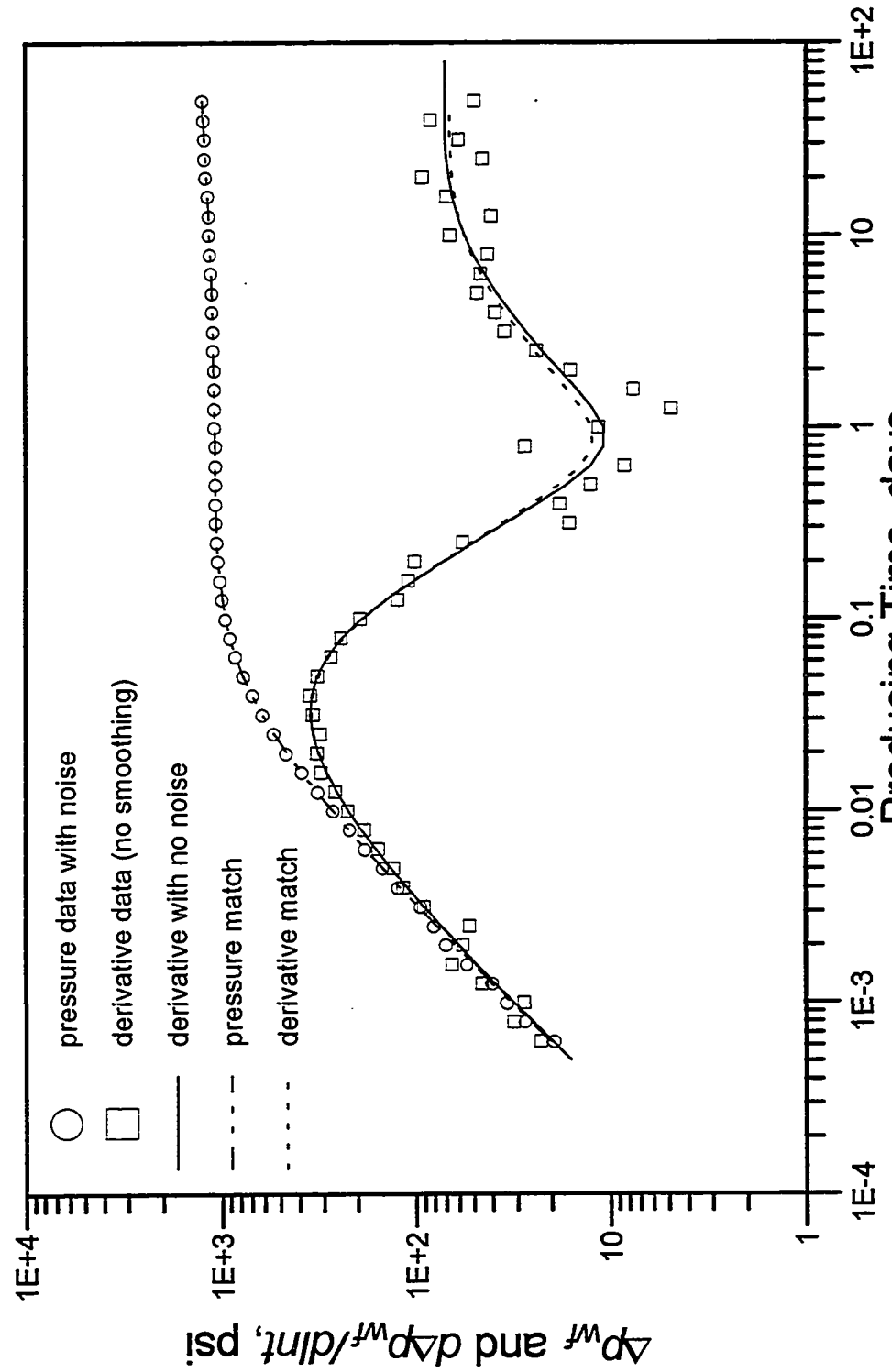

Fig. 4.2.8 - Dual-porosity synthetic drawdown example (seed=–225) results from regression on pressure derivative data

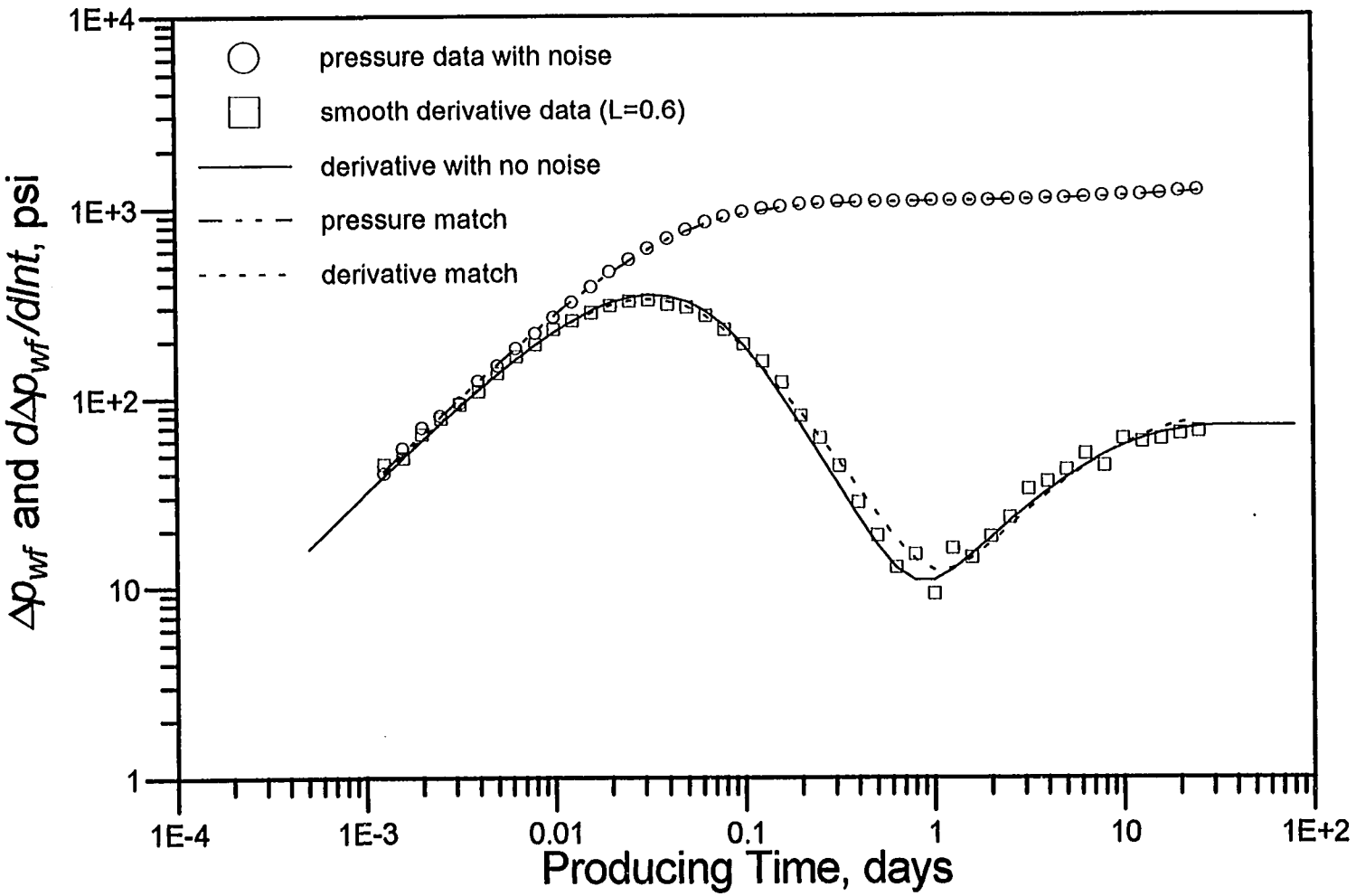

Fig. 4.2.9 - Dual-porosity synthetic drawdown example (seed=-225) results from regression on pressure derivative data

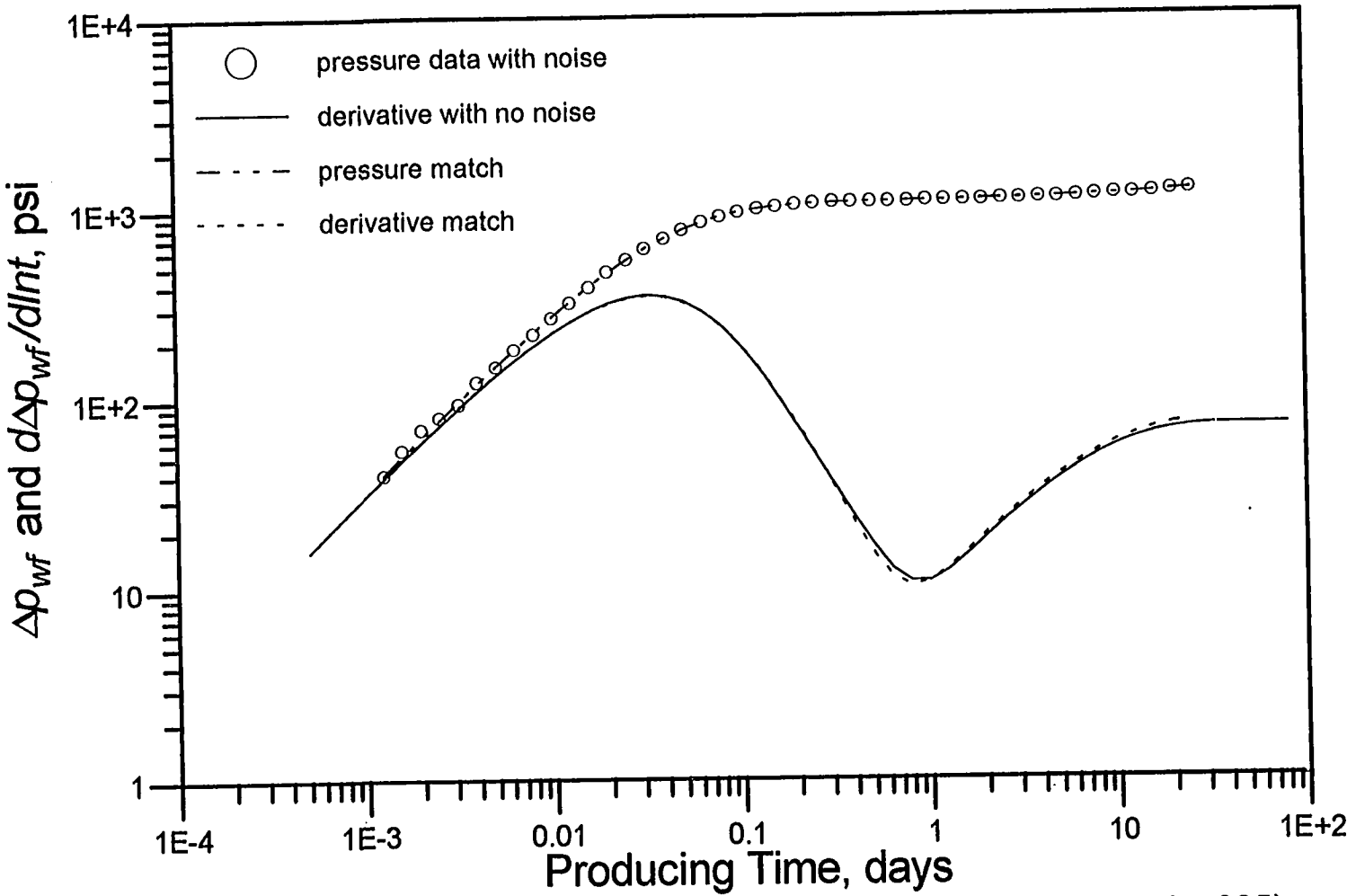

Fig. 4.2.10 - Dual-porosity synthetic drawdown example (seed=–225) results from regression on pressure data

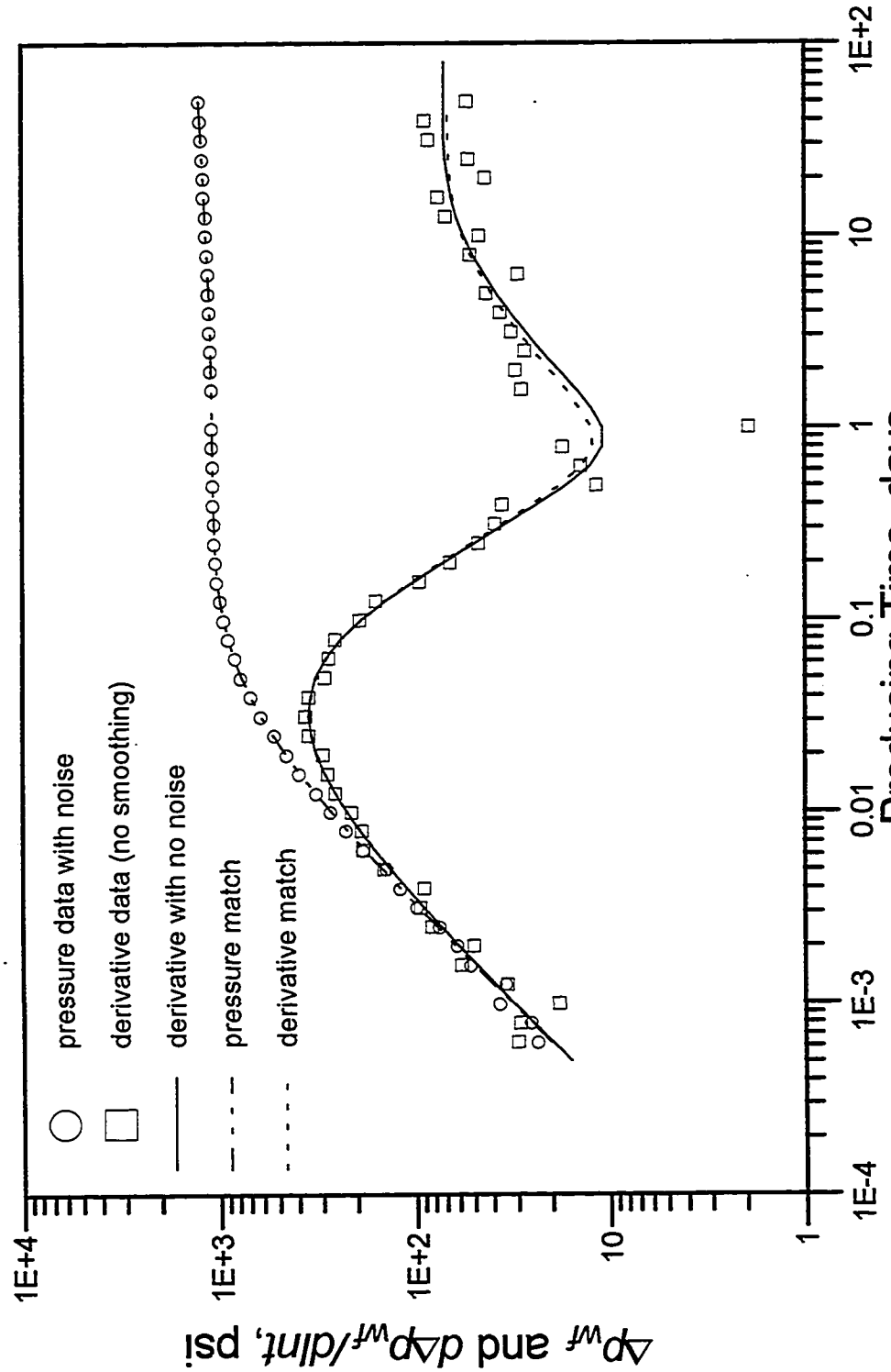

Fig. 4.2.11 - Dual-porosity synthetic drawdown example (seed=–3375) results from regression on pressure derivative data

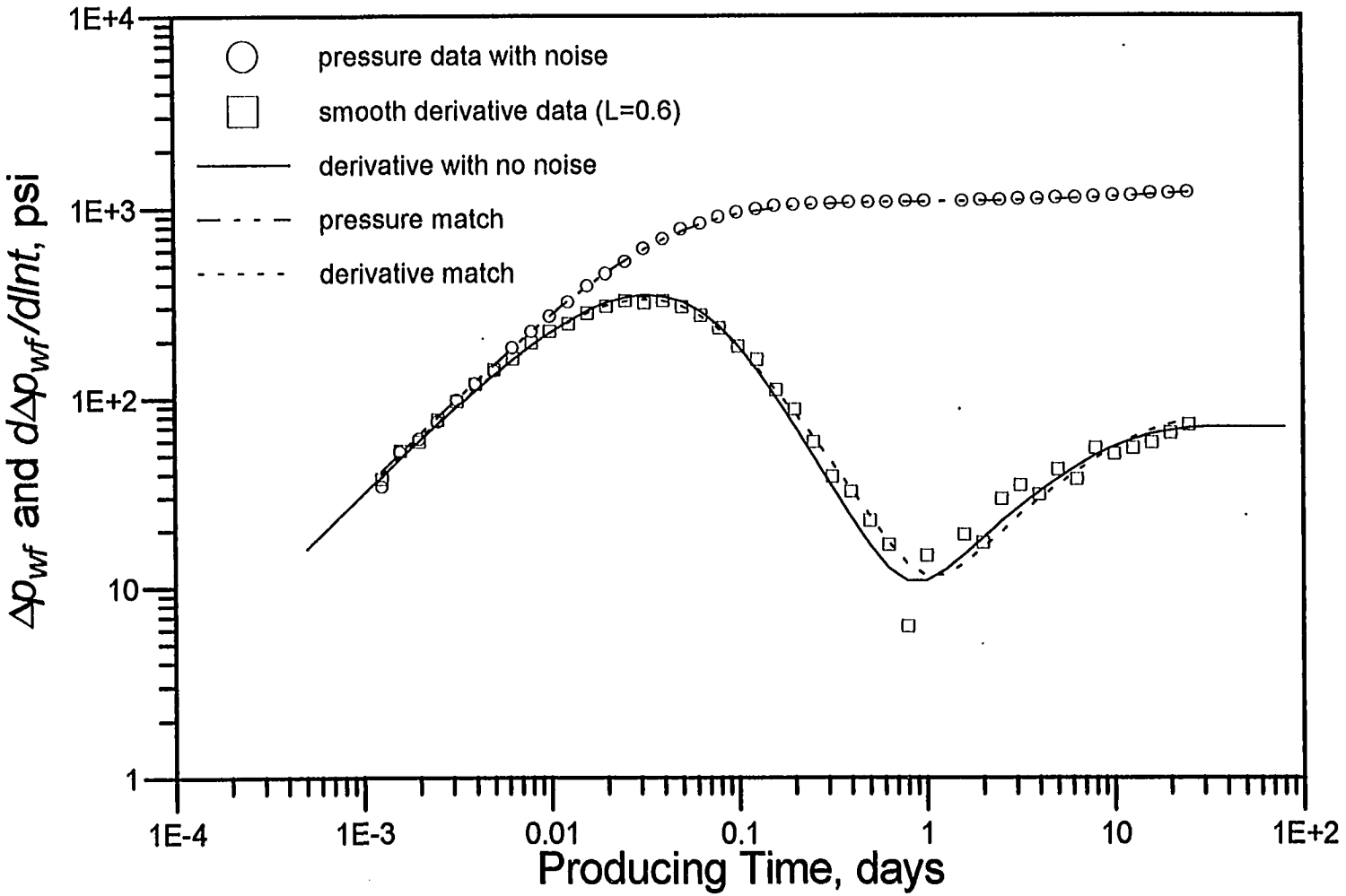

Fig. 4.2.12 - Dual-porosity synthetic drawdown example (seed=–3375) results from regression on pressure derivative data

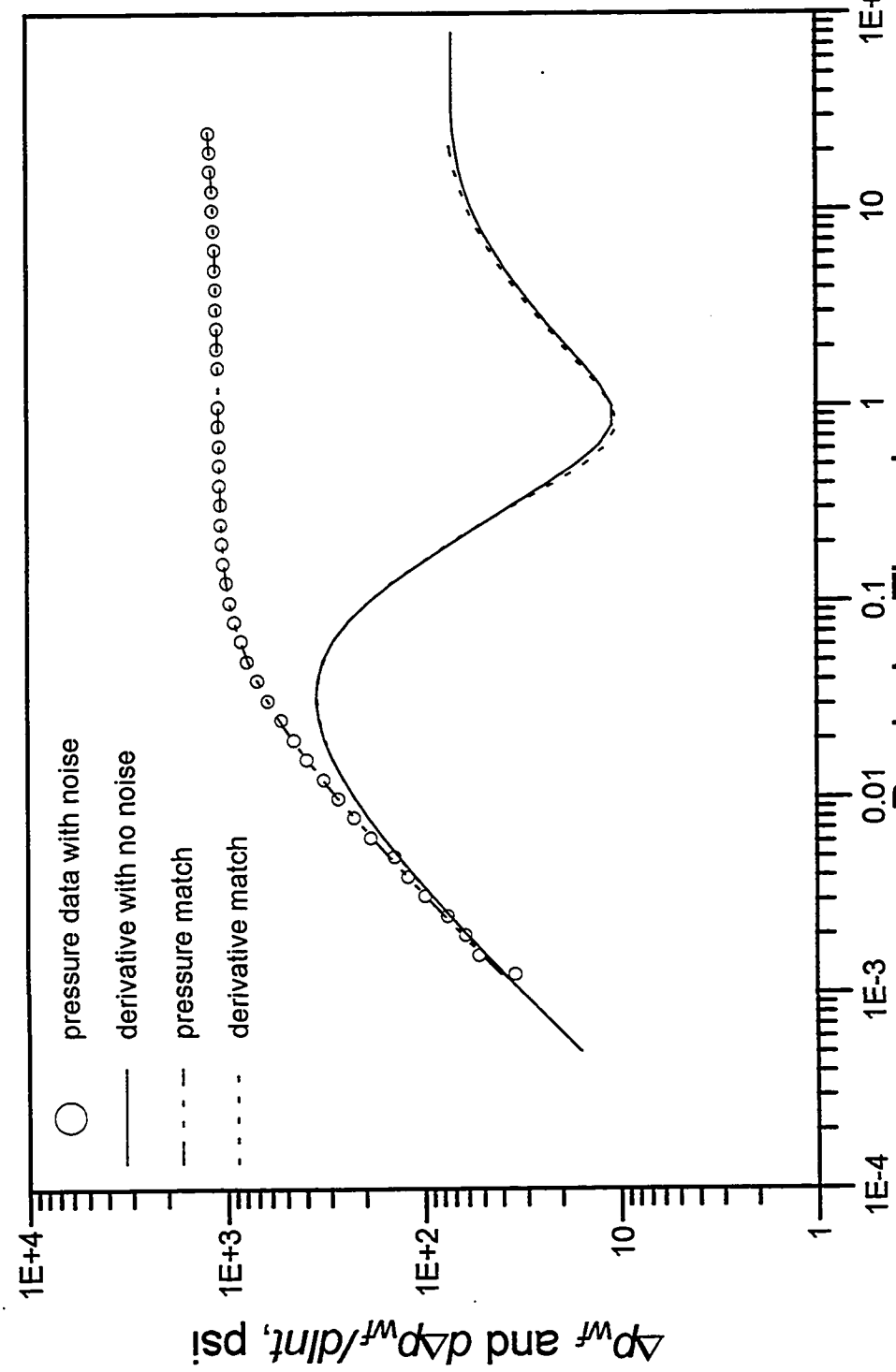

Fig. 4.2.13 - Dual-porosity synthetic drawdown example (seed=–3375) results from regression on pressure data

Fig. 4.2.14 shows the pressure change and the pressure change derivative versus shut-in time on a log-log scale. Note that the pressure derivative exhibits the characteristic dual porosity behavior.

Table 4.2.4 - Flowrate history for the Dual Porosity Field Example (data from reference 51)

Period	Duration (hr.)	Flowrate (stb/d)
1	3.8	800
2	3.3	2500
3	16.45	830
4	18	0

Table 4.2.5 - Reservoir/fluid parameters for the Dual Porosity Field Example (data from reference 51)

Parameter	Value
r_w (ft)	0.29
h (ft)	7
B_o (bbl/RB)	1.5
μ_o (cp)	0.3
c_t (psi^{-1})	2×10^{-5}
$\hat{\phi}_{f+m}$	0.05

The results of the nonlinear LS regressions applied to both pressure data and pressure derivative data are presented in Table 4.2.6. The initial estimates for both runs were somewhat arbitrarily chosen as $\hat{k}_f = 600$ md, $C = 0.015$ bbl/psi, $\Delta p_s = 0$ psi/bbl/d, $\lambda = 1\times10^{-9}$ and $\omega = 0.1$. Here, the total storativity $(\hat{\phi}c_t)_{f+m}$ is not a regression parameter; i.e., it is fixed to the given value during the regression procedure. The

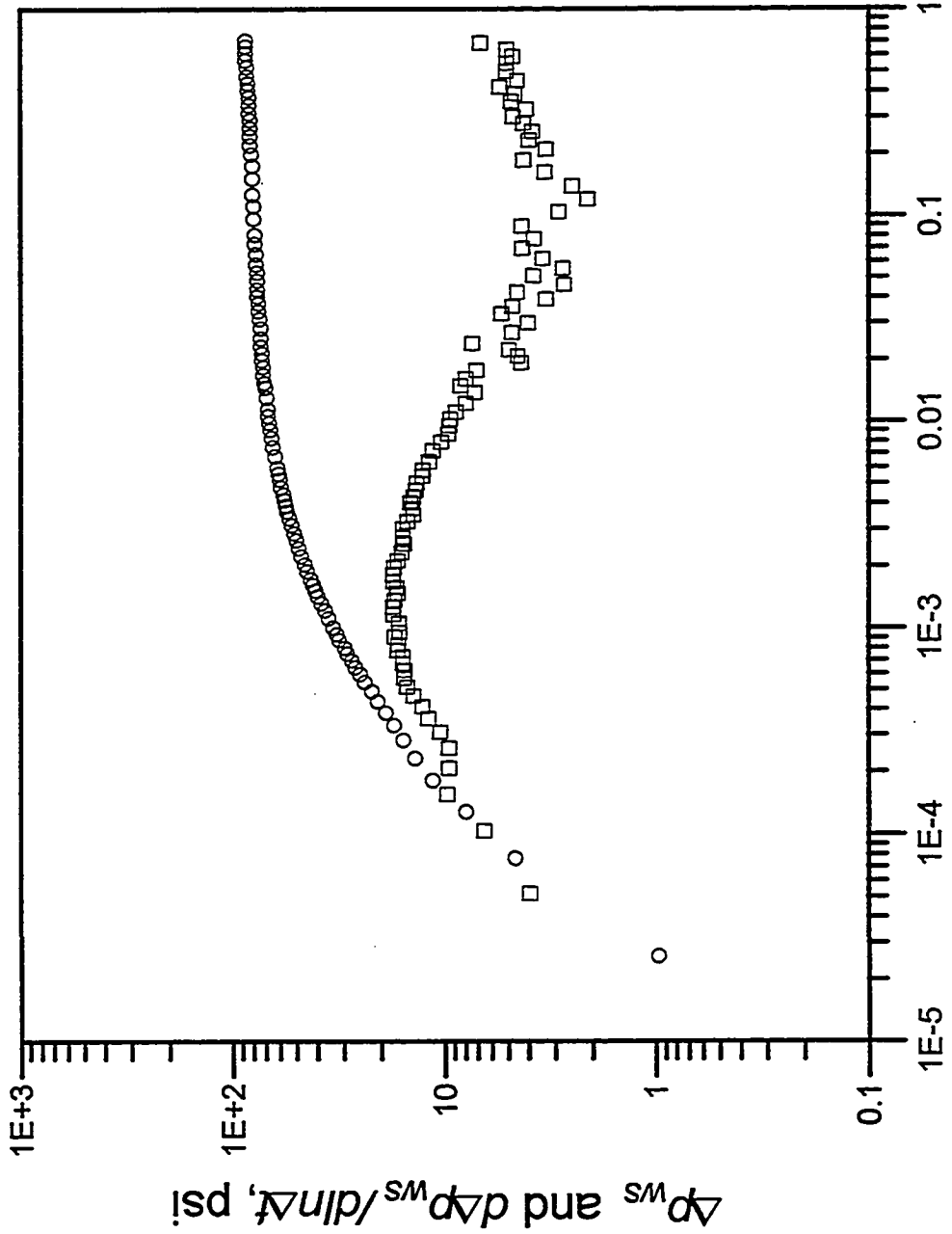

Fig. 4.2.14 - log-log plot - Dual Porosity Field Example

Table 4.2.6 - Results from LS regressions on pressure data and pressure derivative data
Dual-Porosity Field Example

	Regression on Pressure Data	Regression on Pressure Derivative Data
$\hat{k}_f$ (md)	588.33 510.88 to 665.79	442.94 435.90 to 449.99
C (bbl/psi)	0.02448 0.02162 to 0.02735	0.02073 0.01979 to 0.02166
Δp_s (psi/bbl/d)	−0.02620 −0.03773 to −0.01467	−0.05966 −0.06124 to −0.05808
λ	1.222×10^{-8} 0. to 2.615×10^{-8}	2.965×10^{-8} 2.704×10^{-8} to 3.225×10^{-8}
ω	0.6425 0.3263 to 0.9586	0.1420 0.1261 to 0.1579
$p_i - p_{wf,s}$ (psi)	97.331 95.723 to 98.939	—
# of functions evaluations	12	19
# of gradients evaluations	4	15
std. deviation (psi)	3.788	0.7213

regression on the pressure data is for comparison only; according to our discussion in the previous section, for dual porosity systems, we always perform a preliminary regression on the derivative data. Figs. 4.2.15 and 4.2.16 show that the match for the regression on the pressure data is poor (i.e., it seems that the regression converged to a local minimum). Conversely, we obtained an adequate match for the regression on the pressure derivative data with small confidence intervals (see Fig. 4.2.17 and Table 4.2.6). The results of this run were used as initial estimates for the next regressions.

Finally, we use the results of the previous run as initial estimates to regress on the pressure data using both LS and RLS regressions. Table 4.2.7 summarizes the results. Recall that RLS refers to the procedure of applying the LS regression after neglecting the outliers in the data. The number of model (function or gradient) evaluations for the RLS regression shown in Table 4.2.7, includes the model evaluations for the LAV estimation and the final LS regression (after neglecting the outliers based on the LAV residuals). Note that the first four data points were considered outliers by our LAV-based outlier detection (see last row of Table 4.2.7).

Although the final parameter estimates are not that different, the matches shown in Figs. 4.2.18-19 (for the LS regression) and Figs. 4.2.20-21 (for the RLS regression), indicate that the RLS results clearly give a better match. Note that the LS regression suffers from lack of fit at the end of the recorded data (where it is most important). However, except at the beginning of the wellbore storage dominated flow, the RLS results give an excellent match.

It is worth mentioning that the LAV regression (or the RLS regression) does not converge to the right solution using the previous initial estimates; i.e., we still need the preliminary regression on the pressure derivative data.

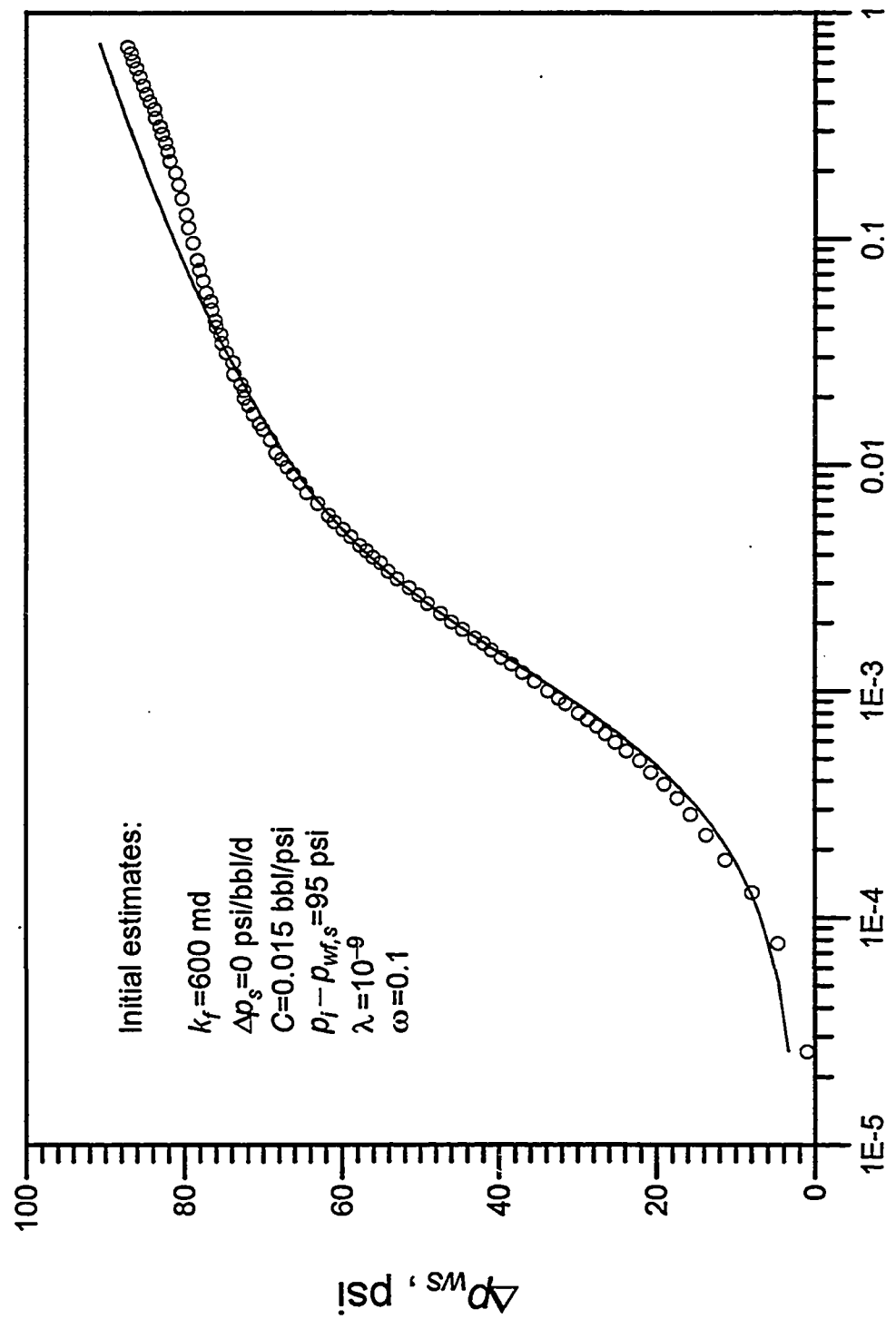

Fig. 4.2.15 - Dual Porosity Field Example - LS regression on pressure data

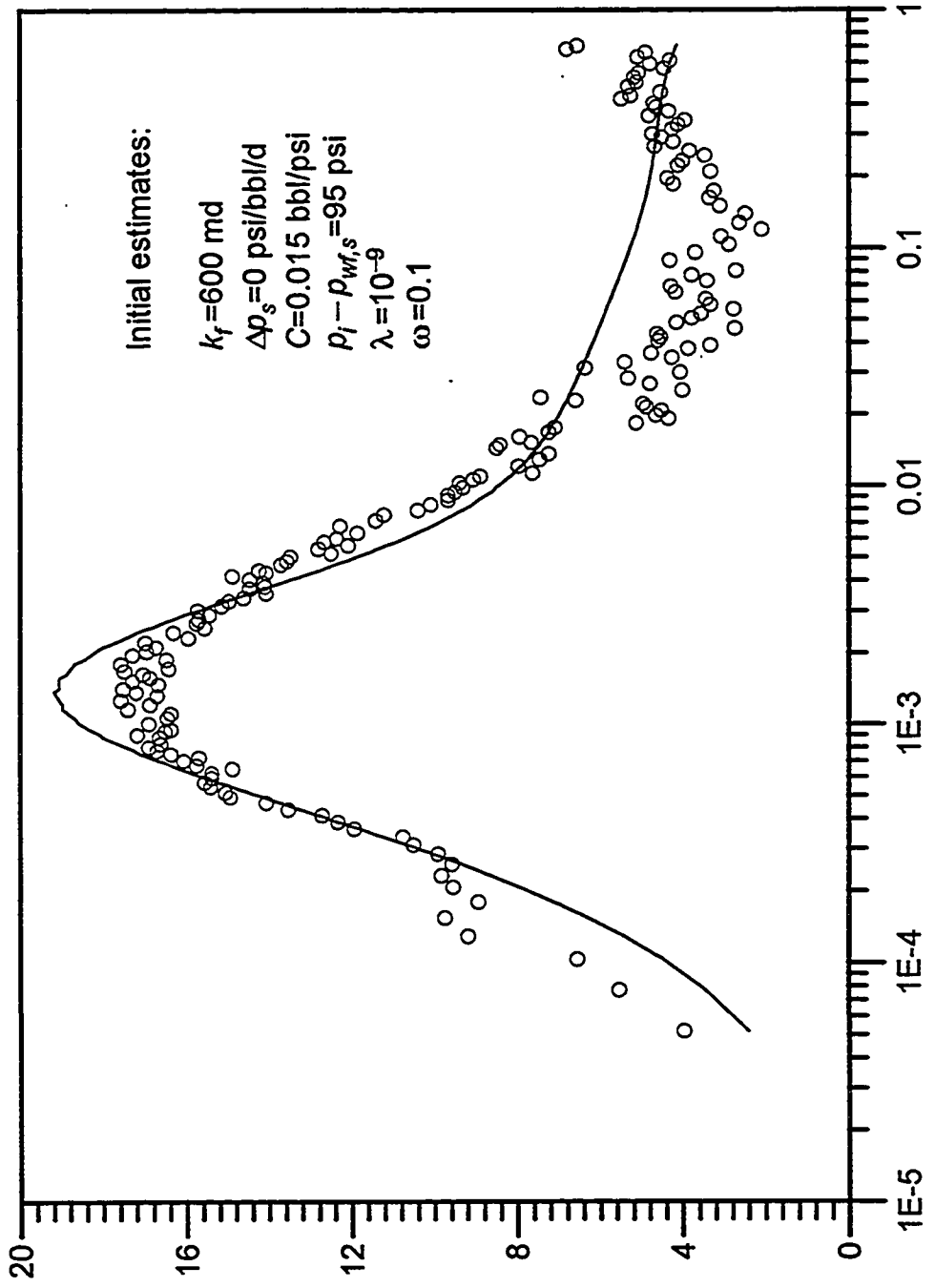

Fig. 4.2.16 - Dual Porosity Field Example - LS regression on pressure data

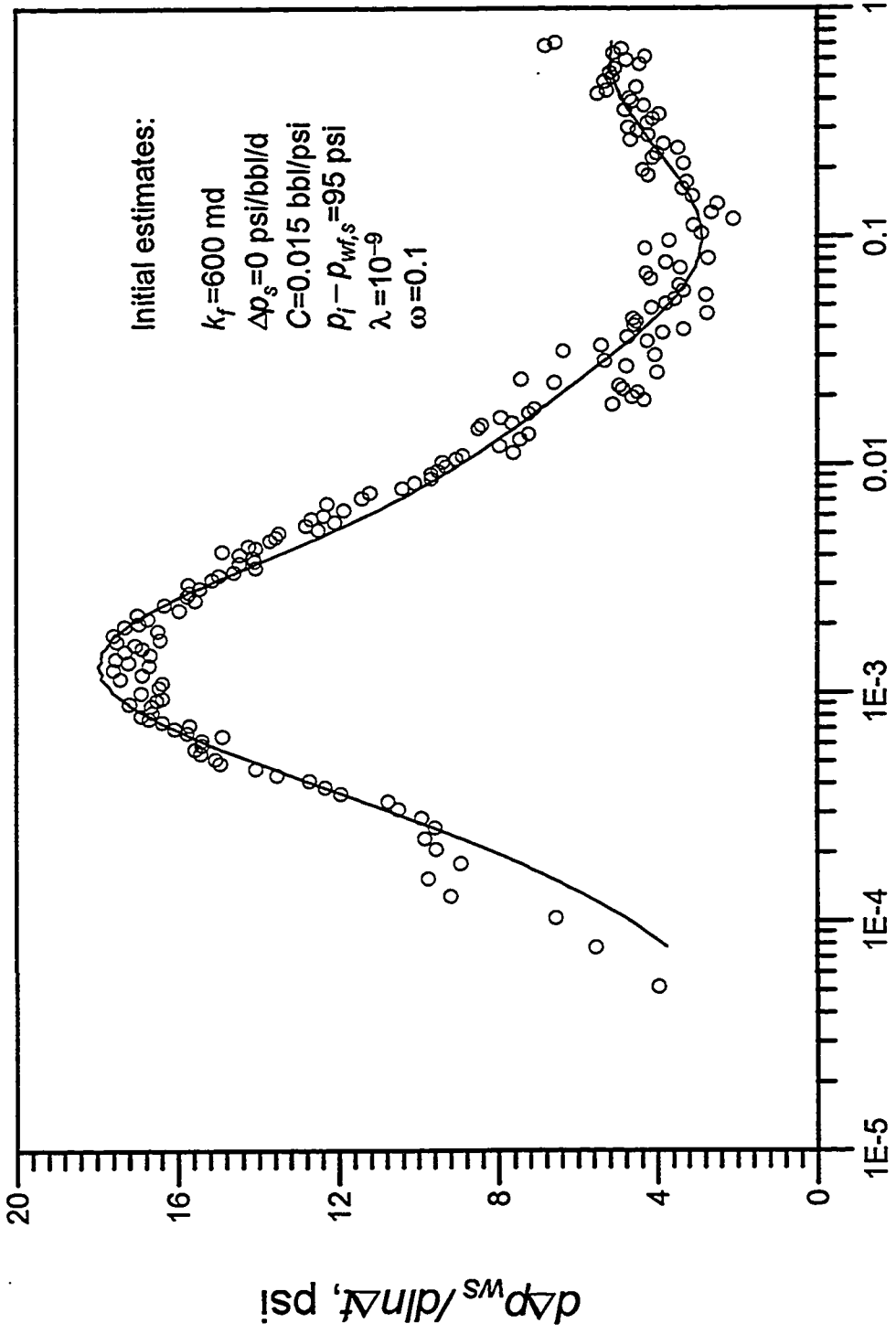

Fig. 4.2.17 - Dual Porosity Field Example - LS regression on pressure derivative

Table 4.2.7 - Results from LS and RLS regressions on pressure data
Dual-Porosity Field Example

	LS Regression on Pressure Data	RLS Regression on Pressure Data
$\hat{k}_f$ (md)	424.65 417.79 to 431.50	442.91 437.85 to 447.96
C (bbl/psi)	0.01931 0.01876 to 0.01986	0.02099 0.02054 to 0.02143
Δp_s (psi/bbl/d)	–0.06323 –0.06527 to –0.06120	–0.05958 –0.06091 to –0.05825
λ	3.389×10^{-8} 3.234×10^{-8} to 3.544×10^{-8}	2.966×10^{-8} 2.880×10^{-8} to 3.052×10^{-8}
ω	0.1452 0.1358 to 0.1545	0.1462 0.1402 to 0.1522
$p_i - p_{wf,s}$ (psi)	97.210 96.853 to 97.568	96.022 95.812 to 96.231
# of functions evaluations	13	67
# of gradients evaluations	17	40
std. deviation (psi)	0.1499	0.0478
Outliers Detected from LAV residuals	—	data points #1, #2, #3 and #4

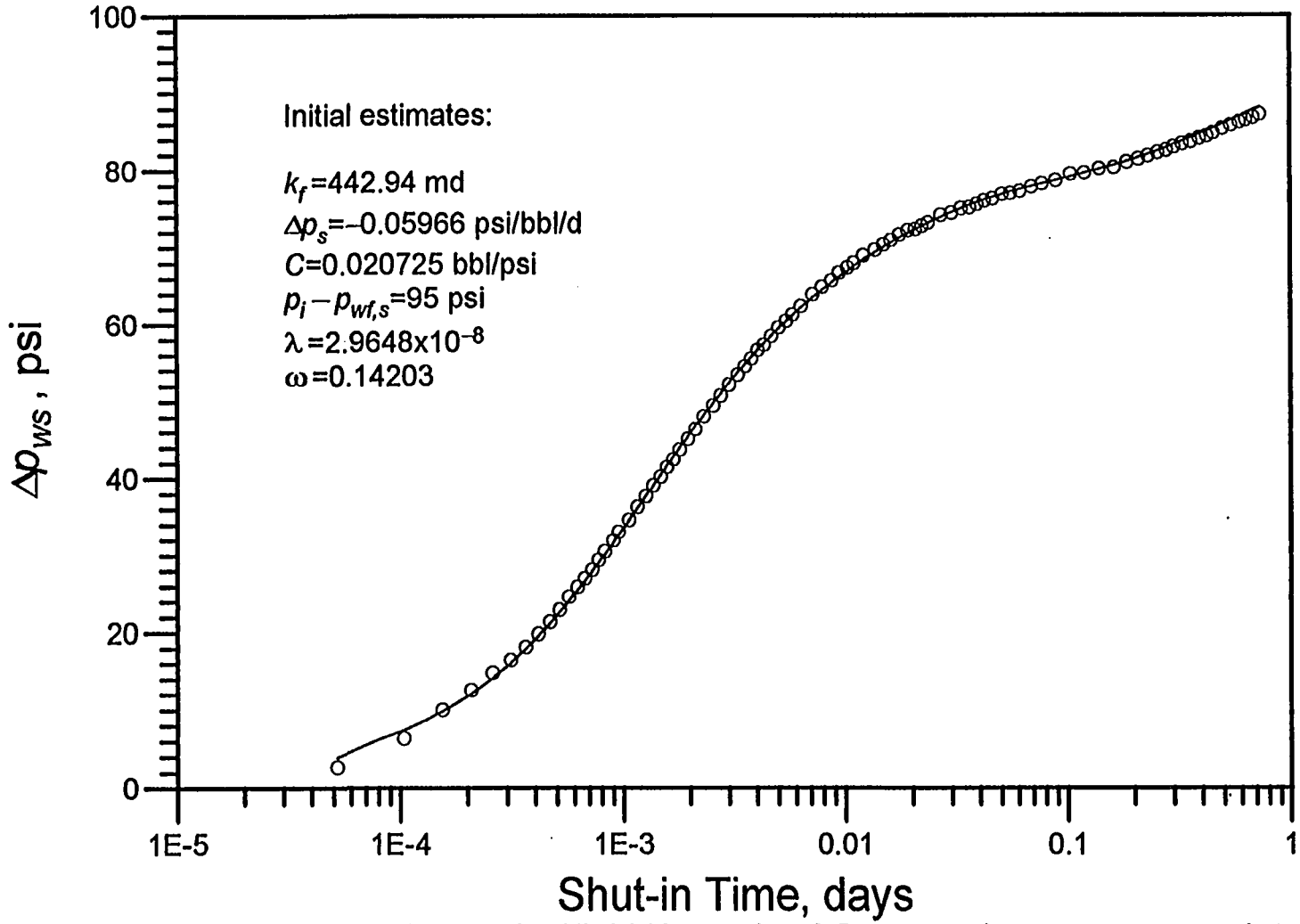

Fig. 4.2.18 - Dual Porosity Field Example - LS regression on pressure data

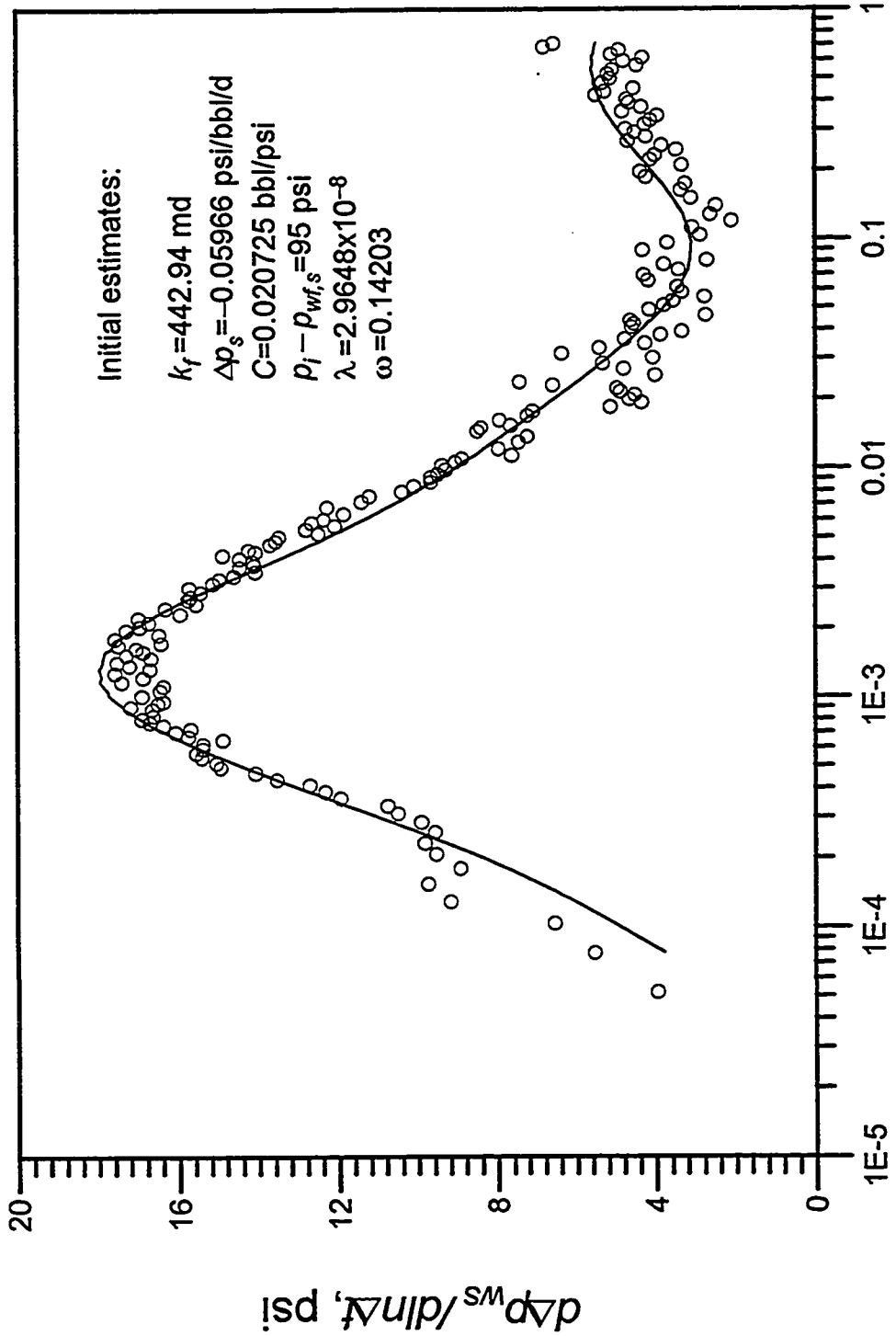

Fig. 4.2.19 - Dual Porosity Field Example - LS regression on pressure data

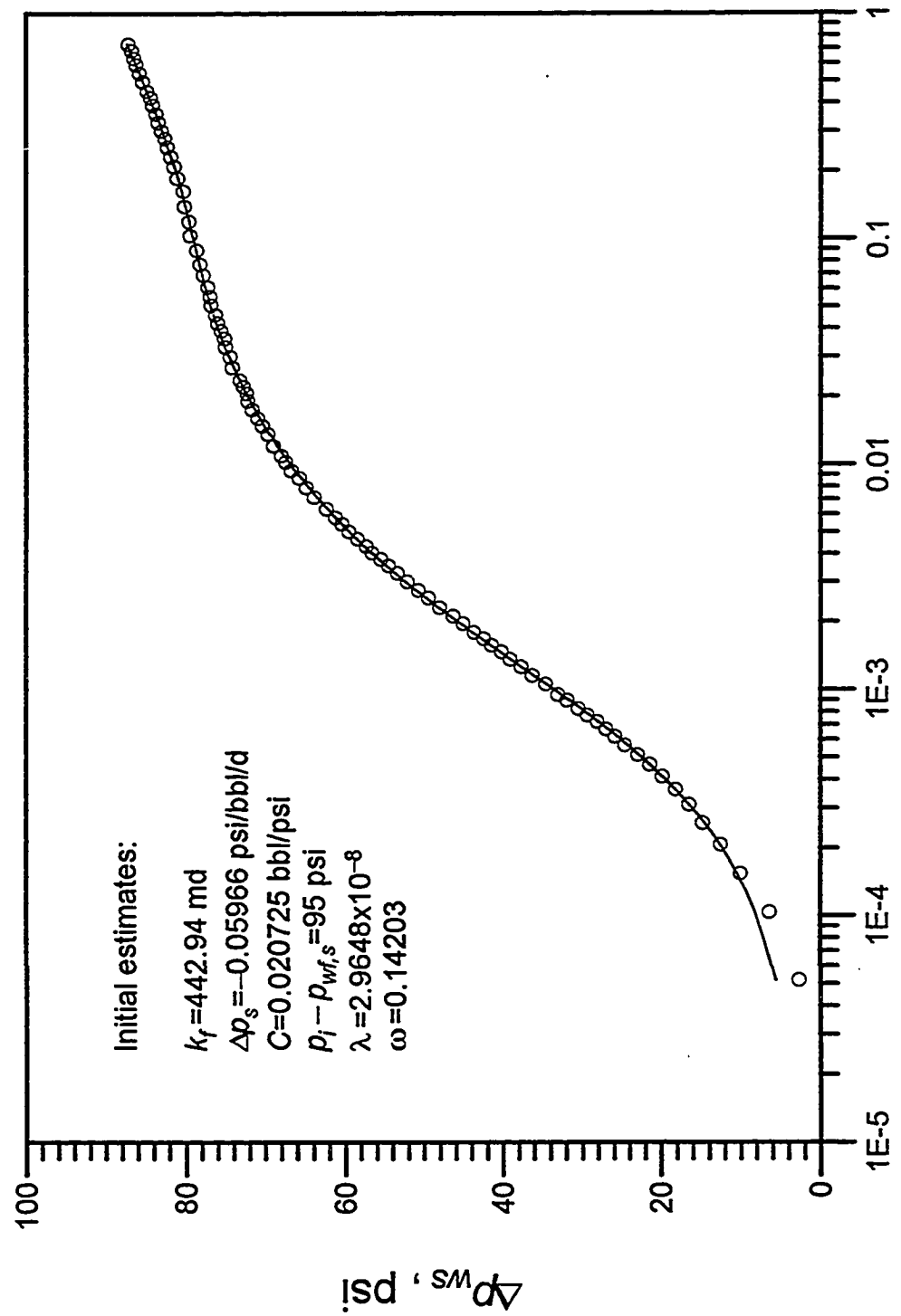

Fig. 4.2.20 - Dual Porosity Field Example - RLS regression on pressure data

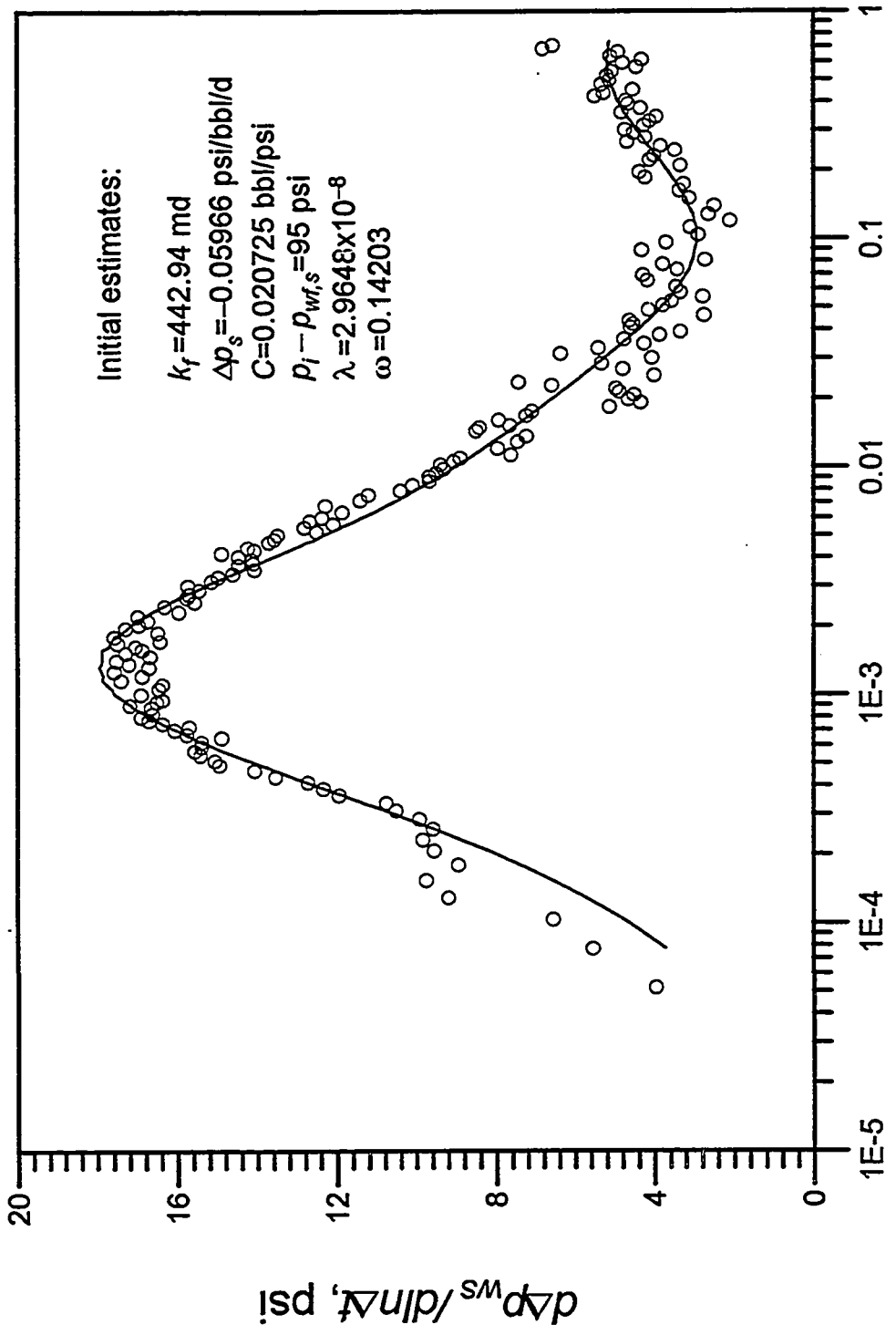

Fig. 4.2.21 - Dual Porosity Field Example - RLS regression on pressure data

CHAPTER V

TOPICS ON DIAGNOSTICS METHODS BASED ON THE RESIDUALS

5.1 Introduction

In this chapter, we review existing methods for verifying the assumptions we have made about the errors. Recall that the usual assumptions are that the errors follow a Gaussian distribution with zero mean and constant variance, σ^2. We expect that, if the model is correct, the residuals should exhibit tendencies that confirm or contradict the assumptions we have made. Some diagnostics are based on statistical quantities computed from the residuals; however, most of them are based on visual analysis of different plots of the residuals (diagnostic plots). Although the last approach is subjective, it can reveal when the assumptions are violated, especially if the analyst systematically applies the technique.

The vector of the residuals, **r**, (sometimes called *ordinary residuals*) is defined by

$$\mathbf{r} = (r_u) = \mathbf{Y} - \hat{\mathbf{Y}}, \tag{5.1.1}$$

where $\mathbf{Y} = (y_u)$, $u = 1, 2, \ldots, n$ is the vector of the n observations and $\hat{\mathbf{Y}} = (\hat{y}_u)$ is the corresponding vector of the fitted values obtained from the regression. If the model is perfect, these residuals must be related to the errors in the data. However, since by construction the regression makes the residuals as small as possible, it is expected that on the average, the residuals are smaller than the errors. For unweighted linear least squares, Eq. 5.1.1 is given by (see Eq. B.7)

$$\mathbf{r} = \mathbf{Y} - \mathbf{X}\mathbf{b}$$

$$= \mathbf{Y} - \mathbf{X}\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}\right)$$

$$= (\mathbf{I} - \mathbf{R})\mathbf{Y}, \tag{5.1.2}$$

where $\mathbf{R} = (R_{ij}) = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T$ is symmetric and idempotent ($\mathbf{R}^2 = \mathbf{R}$). Note that the matrix $\mathbf{R}$ depends only on the form of $\mathbf{X}$. Similarly, for the linearized nonlinear model, Eq. 5.1.1 can be written as

$$\mathbf{r} = (\mathbf{I} - \mathbf{R})\mathbf{Y}^0. \tag{5.1.3}$$

Here $\mathbf{R} = (R_{ij}) = \mathbf{Z}_0\left(\mathbf{Z}_0^T\mathbf{Z}_0\right)^{-1}\mathbf{Z}_0^T$. The definitions of $\mathbf{Y}^0$ and $\mathbf{Z}_0$ are presented in Eq. 2.3.5. The variance-covariance matrix of $\mathbf{r}$ can be obtained from Eqs. 5.1.2 or 5.1.3. Since $V(\mathbf{Y}) = V(\mathbf{Y}^0) = \mathbf{I}s^2$, it follows that

$$\begin{aligned} V(\mathbf{r}) &= (\mathbf{I} - \mathbf{R})\mathbf{I}\sigma^2(\mathbf{I} - \mathbf{R})^T \\ &= (\mathbf{I} - \mathbf{R})\sigma^2 \end{aligned} \tag{5.1.4}$$

This identity demonstrates that the residuals can be correlated to each other even when the observations are not. The $cov(r_i,r_j)$ is given by the (i,j)th element of the matrix $(\mathbf{I} - \mathbf{R})\sigma^2$. Thus, even though the theoretical random errors ε_u are assumed to be independent with mean zero and constant variance σ^2, the residuals are in general not independent and they also do not have the same variance.

The $n \times n$ matrix $\mathbf{R}$ is called the *hat matrix*. It is important to understand the behavior of $\mathbf{R}$ for a sound understanding of the relationship between the errors, ε, and the residuals, $\mathbf{r}$. The fact that $\mathbf{R}$ is idempotent and symmetric implies that[5]

$$\text{rank}(\mathbf{R}) = \text{trace}(\mathbf{R}) = p \quad \left(\text{i.e., } \sum_{i=1}^{n} R_{ii} = p\right), \tag{5.1.5}$$

$$R_{ii} = \sum_{j=1}^{n} R_{ij}^2 \quad \text{for all } i. \tag{5.1.6}$$

It follows from Eqs. 5.1.5 and 5.1.6, respectively, that the average value of R_{ii} is p/n and that $0 \le R_{ii} \le 1$ for all i. A diagonal element of zero indicates a point with no influence on

the fit; this is the case of a point with all independent variables equal zero in a regression model without intercept. On the other hand when $R_{ii} = 1$, the ith observation is fitted exactly by the regression model. Some authors suggest that high R_{ii} values are associated with possible influential or leverage points; specifically, particular attention should be paid to points for which $R_{ii} > 2p/n$ (see Refs. 5 and 25 for further discussion). If we neglect the correlation between the residuals, the variance of the residuals given by Eq. 5.1.4 can be approximated in an average sense by

$$V(\mathbf{r}) = \left(1 - \frac{p}{n}\right)\mathbf{I}\sigma^2 = \frac{(n-p)}{n}\mathbf{I}\sigma^2. \tag{5.1.7}$$

Bard[21] proposes a similar equation based on different arguments.

For diagnostic analysis (i.e., identifying influencial points and testing model hypotheses), many scaling transformations of the ordinary residuals have been suggested in the literature. One of them is the *standardized residuals* defined as

$$\frac{r_i}{s}, \tag{5.1.8}$$

where s, the estimate of the standard deviation of the errors, σ, is given by Eq. B.9. The term r_i/s is also called the *unit normal deviate form* of the residual r_i, since we often assume that $\varepsilon_i \approx N(0,\sigma^2)$, so that $\varepsilon_i/\sigma \approx N(0,1)$. The idea is to examine the standardized residuals r_i/s to see if the assumption $\varepsilon_i/\sigma \approx N(0,1)$ is violated. Some authors take Eq. 5.1.4 into account for the scaling of the residuals. They recommend the *studentized residuals*

$$\frac{r_i}{\sqrt{(1-R_{ii})s^2}}. \tag{5.1.9}$$

Draper and Smith[16] recognize that Eq. 5.1.9 is more general; however, they point out that very little is lost by using the simpler form of Eq. 5.18 for most practical problems, except

when the ratio $(n-p)/n$ is quite small (see Eq. 5.1.7). Our experience is that, if the model is correct, the effect of correlation between residuals can be ignored as long as the data is free of outliers.

5.2 Statistical Tests

5.2.1 Testing the Zero-Mean Hypothesis of the Residuals

Suppose that we have a model that does not guarantee zero average residuals. In this case, the statistic

$$t = \bar{r}\sqrt{n-1} \Big/ \sqrt{\sum_{i=1}^{n}(r_i - \bar{r})^2 \Big/ n} \tag{5.2.1}$$

is distributed as a t-distribution with $n-1$ degrees of freedom. This is the well-known t-test for the mean (zero, in this case).[18] The statistic given by Eq. 5.2.1 can be compared with the tabulated t-value for a specific confidence level. If the absolute value of the calculated t-value is smaller than the tabulated value, then we accept the zero-mean hypothesis.

5.2.2 Testing the Normality Assumption of the Errors

We are interested in testing the normality assumption of a population based on a sample data. The χ^2 test can handle this kind of problem. Let us first look at the chi-squared frequency function. It is given by

$$f(\chi^2)=\frac{(\chi^2)^{\upsilon/2-1}e^{-\chi^2/2}}{2^{\upsilon/2}\Gamma(\upsilon/2)}, \tag{5.2.2}$$

where $\Gamma(x)$ denotes the gamma function of x. Note that the distribution depends only on υ, which is called the number of degrees of freedom. The mean and the variance of the distribution are υ and 2υ, respectively. In general, the statistic used to measure and to test the significance of the compatibility of observed and theoretical frequencies is defined as

$$\chi^2=\sum_{i=1}^{m}\frac{(o_i-e_i)^2}{e_i}, \tag{5.2.3}$$

where m is the number of pairs of frequencies, and o_i and e_i comprise the ith pair of observed and theoretical frequencies. Clearly, for a continuous distribution, first we need to sort the events into bins or classes. In that case, m represents the number of bins.

It is apparent from Eq. 5.2.3 that the χ^2 statistic is a measure of deviation from expected values. Larger values of χ^2 correspond to poorer agreement between observed and expected frequencies. More importantly, experience and theoretical investigations show that if the expected frequencies are not too small in number, this quantity is very closely distributed like the χ^2 distribution given by Eq. 5.2.2, with $m-1$ degrees of freedom. According to Hoel,[18] the approximation is satisfactory provided that each $e_i \geq 5$ and $m \geq 5$. The values of e_i should be somewhat larger than 5 if $m<5$. For best results it may be necessary to use bin widths of varying sizes, such that the bin frequencies are roughly the same.

Since large values of χ^2 correspond to poor agreement between the observed and expected frequencies, one may compare the calculated χ^2 value with a critical value χ_0^2 corresponding to a specified type I error probability (usually 1% or 5%). Thus for a 1% significance level, χ_0^2 is the value that cuts off 1% of the right tail of the χ^2 distribution given by Eq. 5.2.2. If the actual χ^2 value is less than the critical value, we can not reject

the hypotheses of compatibility between the frequencies. A remarkable property of the χ^2 test is that it is still applicable when the bin frequencies depend upon unknown parameters, provided that the unknown parameters are replaced by their maximum likelihood estimates and provided that one degree of freedom is deducted for each parameter estimated. Thus $\upsilon = m-1-p$ when the bin probabilities depend upon p parameters.

If the error term satisfies the usual assumption $\varepsilon_i \approx N(0,\sigma^2)$, then the scaled residuals (see Eqs 5.1.8 and 5.1.9) will be approximately normally distributed. Thus we can examine the validity of the normality assumption by an examination of the residuals. Note that χ^2 is not a robust statistic in the sense that it is highly affected by outliers—the test should be performed after detecting and neglecting all outliers. First we classify the residuals (say n) in classes. It is recommended that each class has the same expected frequency. For example, if we want to use five classes, the boundaries are such that the expected relative frequencies are 20% for each class (which corresponds to an expected frequency of $0.2n$).

The cumulative frequency of the normalized gaussian distribution can be derived as,

$$F(z) = \frac{1}{2}\left[1 + erf\left(z/\sqrt{2}\right)\right]. \tag{5.2.4}$$

$F(z)$ is the integral of the normalized gaussian distribution from $-\infty$ to z. It represents the probability of drawing a random sample value less than z from a normal population with mean zero and unit standard deviation. Eq. 5.2.4 can be used to obtain the proper class boundaries (z-values). Note that since we want the z value that gives a specified $F(z)$ value, the procedure will require the interpolation of the error function.

Once we get the m intervals, Eq. 5.2.1 can be used to compute the χ^2 statistic. The computed χ^2 value is then compared with the tabulated value for a specified significance level and $(m-1-p)$ degrees of freedom. Recall that in order to have a meaningful

test the expected frequencies in each class should be greater than 5. This can be a limitation in problems with few data points, say less than 25.

5.2.3 Testing the Serial Correlation Patterns in the Residuals

Here we want to test the randomness of the residuals by counting the number of *runs*. A run is defined as a sequence of residuals of the same sign. If the number of runs is much smaller than expected, the randomness of the residuals is suspected.

Let n_1 and n_2 be the numbers of negative and positive residuals, respectively. Under the randomness assumption, the expected number of runs is given by[16]

$$\mu = \frac{2n_1 n_2}{n_1 + n_2} + 1, \tag{5.2.5}$$

with variance

$$\sigma^2 = \frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}. \tag{5.2.6}$$

When both n_1 and n_2 are greater than about 10, the normal approximation to the actual distribution of the number of runs provides good accuracy. In this case, the quantity

$$z = \frac{(u - \mu + 0.5)}{\sigma} \tag{5.2.7}$$

is approximately a unit normal deviate, $N(0,1)$. In Eq. 5.2.7, u is the actual number of runs and the half is used to compensate for the fact that the discrete distribution is being approximated by a continuous distribution in the lower tail. When n_1 and n_2 are smaller than 10, one should use a table of the actual distribution for the runs (see for example Draper[16], p. 160).

Note that the test for the runs of the residuals is exactly applicable only if the residuals are independent. Eq. 5.1.4 shows that the residuals are correlated; thus will somehow affect the significance level of the test. In practice, the effect of correlation between the residuals can be ignored, except when the ratio $(n-p)/n$ is quite small.[16]

5.3 Graphical Analysis of the Residuals

Graphical analysis of the residuals relates to the visual examination of different plots of the residuals. Although this is a subjective approach, it is very revealing when the assumptions about the errors are violated. For graphical purposes, it would be better to use the studentized residuals, $r_i/\sqrt{(1-R_{ii})s^2}$, rather than the standardized residuals, r_i/s, or the ordinary residuals r_i. However, on the basis of our previous discussion, while we recommend use of the studentized residuals whenever they are available, we will use r_i/s and r_i for most practical problems. The following residual plots are usually considered:

1. *Overall.* The overall plot is made in terms of a histogram plot. The classes or bins of the histogram have the same length and are bounded by $\pm 3s$. For example, if we want six classes and $s = 0.5$, the boundaries of the classes would be -1.5, -1.0, -0.5, 0., 0.5, 1.0, and 1.5. If our model is correct (i.e., the well test model and the assumptions of the error term are valid), the histogram should approximately resemble a sample of a gaussian distribution with zero mean.

2. *Against the independent variables* (time in our problem). The residuals in this plot should look like a horizontal band with the residuals randomly placed inside the band. Linear, quadratic, or periodic trends may happen and will call for modifications in the model. For example, a systematic increase in the absolute value of the residuals with time indicates that the variance of the errors is not constant with time, and a weighted least

squares analysis should have been used. A quadratic trend implies that the model is inadequate, i.e., extra terms are missing in the linearized model.

3. *Against the fitted values* $\hat{y}_i$. Again, this plot should look like a horizontal band with the residuals randomly placed inside the band. If not, possible problems with the model are identified in exactly the same way as in the previous plot of residuals versus time..

4. *Against the preceding residual* (except for the first one). This plot reveals if the residuals are serially correlated; more specifically, it shows if a lag-1 serial correlation is present in the data. If a lag-1 serial correlation exists, the successive residuals tend to be more alike than otherwise, and the plot will exhibit a diagonal trend. Similarly, we can view a lag-l correlation by plotting the residuals l steps apart, i.e., plotting r_i versus r_{i-1} for $i = l+1, \ldots, n$. If serial correlations exist in the residuals, Ref. 16 recommends redoing the analysis using the weighted least squares estimation, as described in Appendix B. In this case, the elements $(\mathbf{W}^{-1})_{ij}$ of the variance-covariance matrix of the errors $\mathbf{W}^{-1}$, would be estimated by ρ_l where $l = |i-j|$. The correlation coefficient ρ_l is a measure of the correlation between the residuals l steps apart in time sequence, and is obtained via

$$\rho_l = \frac{\sum_{i=1}^{n-l}\left(r_i - \frac{1}{n-l}\sum_{j=1}^{n-l} r_j\right)\left(r_{i+l} - \frac{1}{n-l}\sum_{j=1}^{n-l} r_{j+l}\right)}{\left\{\sum_{i=1}^{n-l}\left(r_i - \frac{1}{n-l}\sum_{j=1}^{n-l} r_j\right)^2 \sum_{i=1}^{n-l}\left(r_{i+l} - \frac{1}{n-l}\sum_{j=1}^{n-l} r_{j+l}\right)^2\right\}^{1/2}} \qquad (5.3.1)$$

The diagonal elements of $\mathbf{W}^{-1}$ are all equal to one.

5.3.1 Field Example #1

This example is the Earlougher and Kersch drawdown that we analyzed in Chapter III, Section 3.3. Here we focus on the residuals of the regression fit in order to test the assumptions that the errors are independently normally distributed with zero mean and constant variance, σ^2.

Fig. 5.3.1 shows the histogram of the residuals when the residuals are classified in seven bins. It appears that the normality assumption was not violated, confirmed by the fact that the residuals passed the zero-mean test (Section 5.2.1) and the normality test (Section 5.2.2). From the plot of r_i versus r_{i-1} for $i = 2, \ldots, n$, shown in Fig. 5.3.2, we conclude that the residuals are not serially correlated. The residuals also pass the serial correlation test (Section 5.2.3). The plot of the residuals versus the fitted response (Fig. 5.3.3) and the plot of the residuals versus time (Fig. 5.3.4) give the impression that the residuals are randomly placed inside a horizontal band, which does not contradict the wellbore storage and skin model used to fit the data. Thus, in this case, the analysis of the residuals gives us more confidence in the results of the regression.

5.3.2 Field Example #2

This is the Bourdet buildup test which was considered previously in Chapter III, Section 3.6. Fig. 5.3.5 is the histogram plot of the LS residuals that were classified in nine bins. The histogram plot shows that the normality assumption of the errors cannot be rejected; moreover, the residuals passed the zero-mean test (Section 5.2.1) and the normality test (Section 5.2.2).

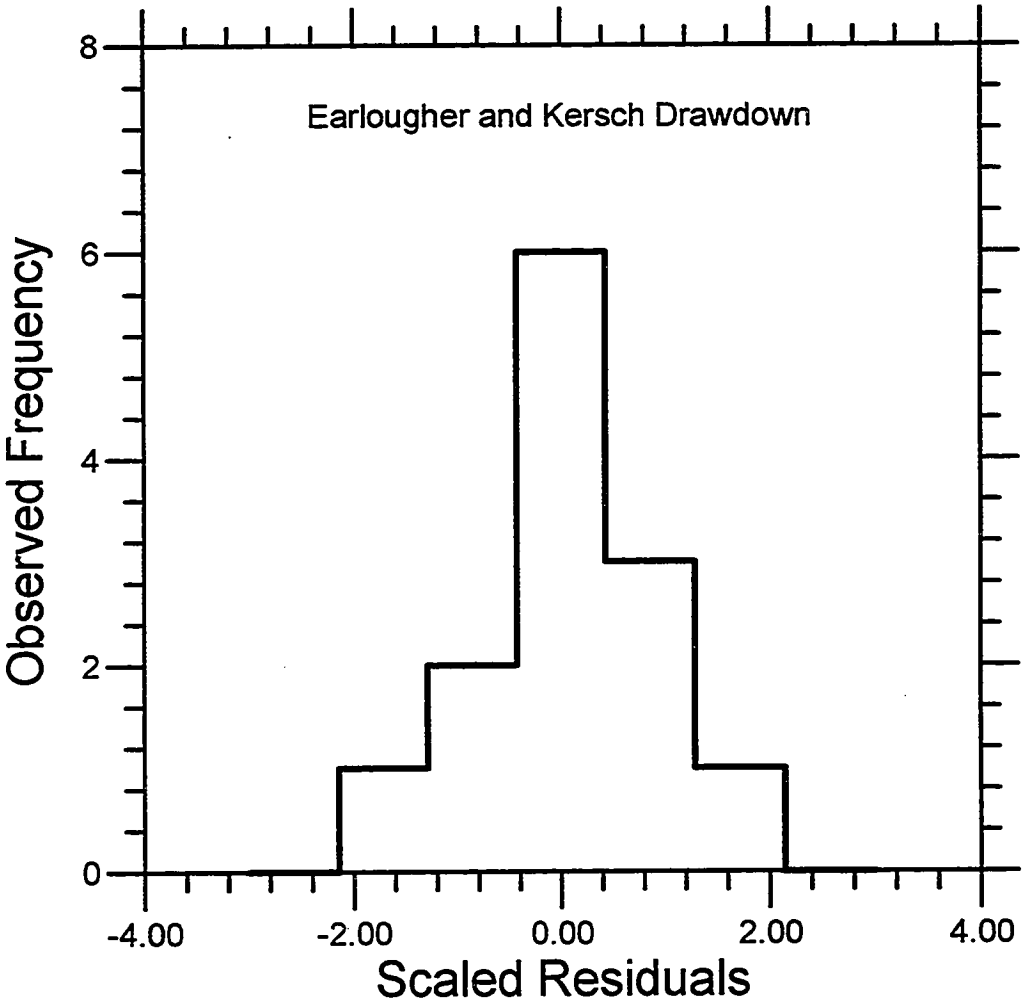

Fig. 5.3.1 - Histogram Plot - Field Example #1

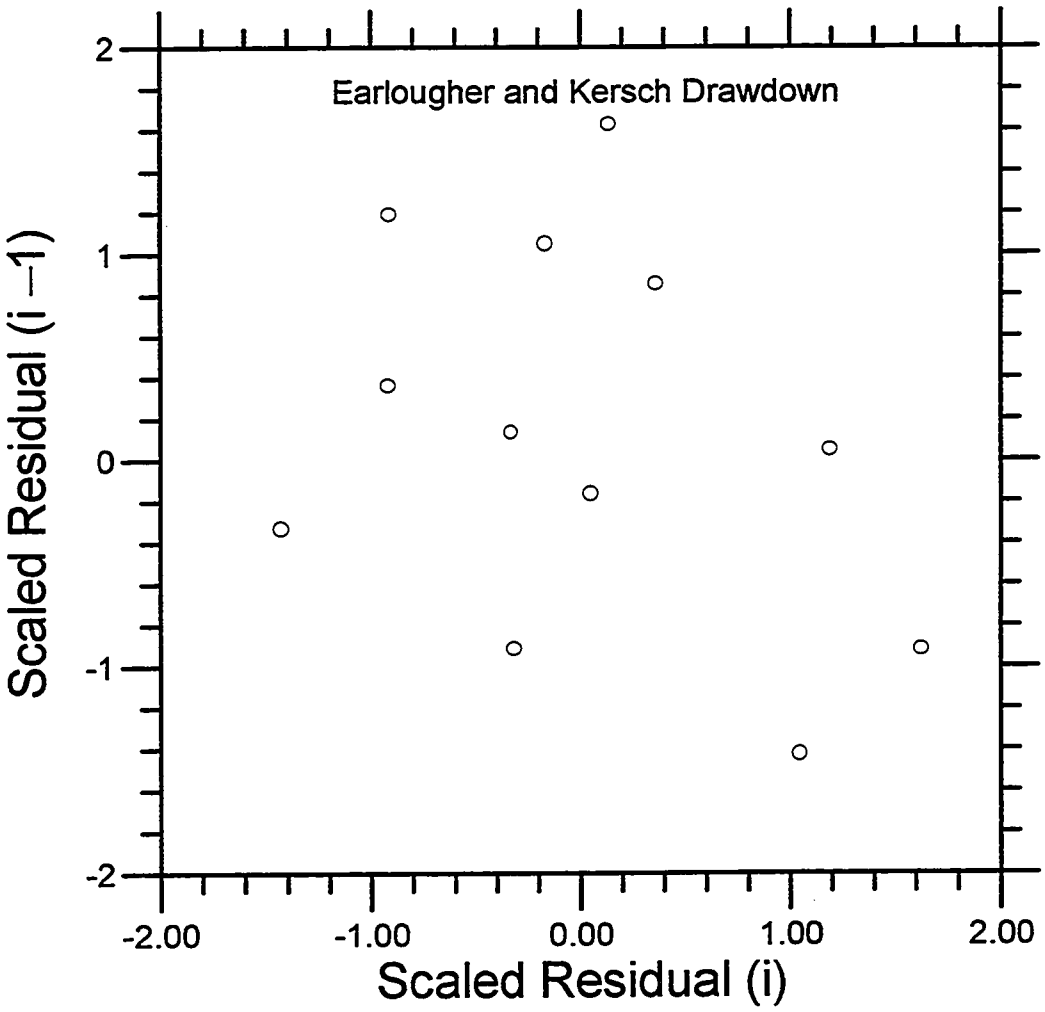

Fig. 5.3.2 - Serial Correlation Plot - Field Example #1

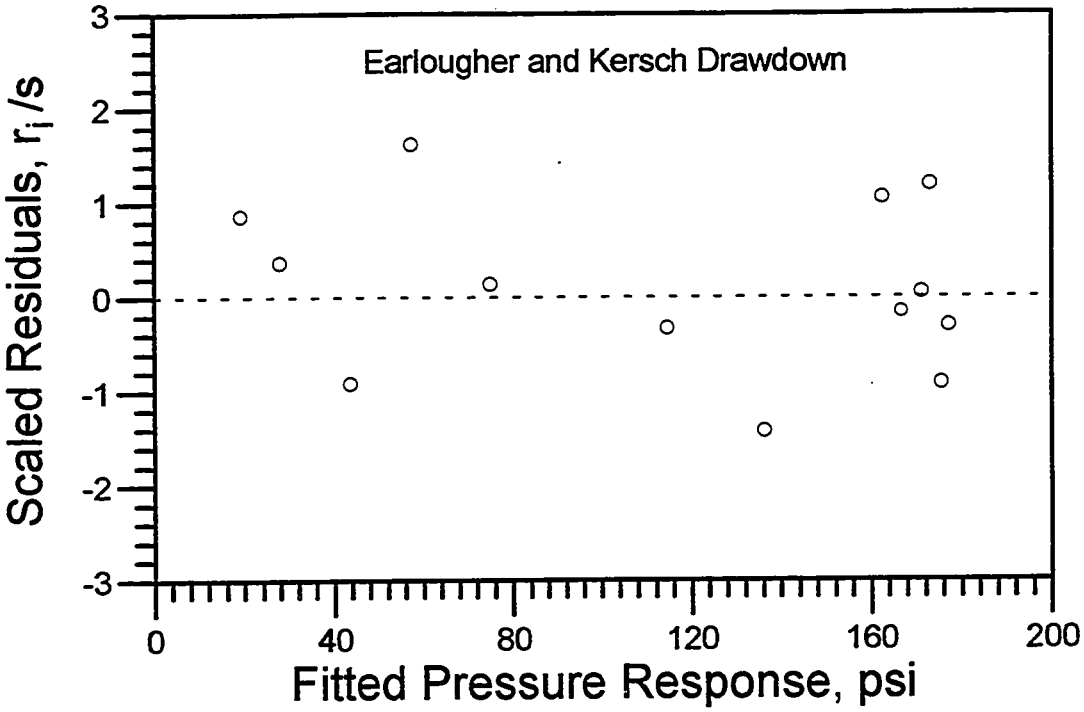

Fig. 5.3.3 - Diagnostic plot - Field Example #1

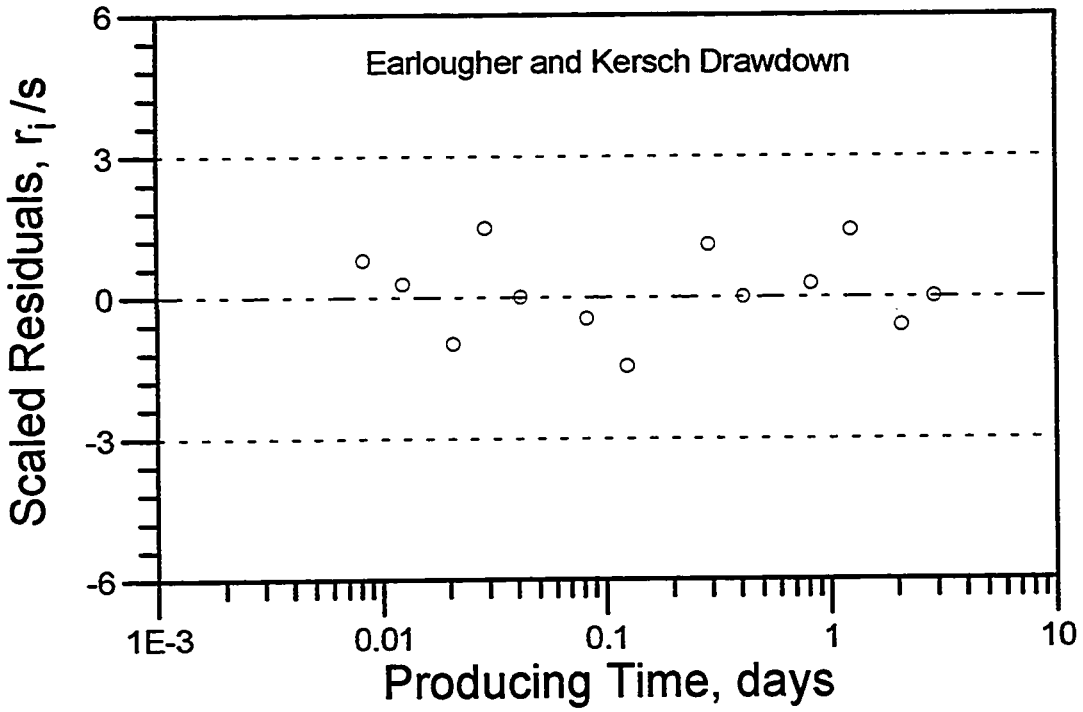

Fig. 5.3.4 - Scaled residual plot - Field Example #1

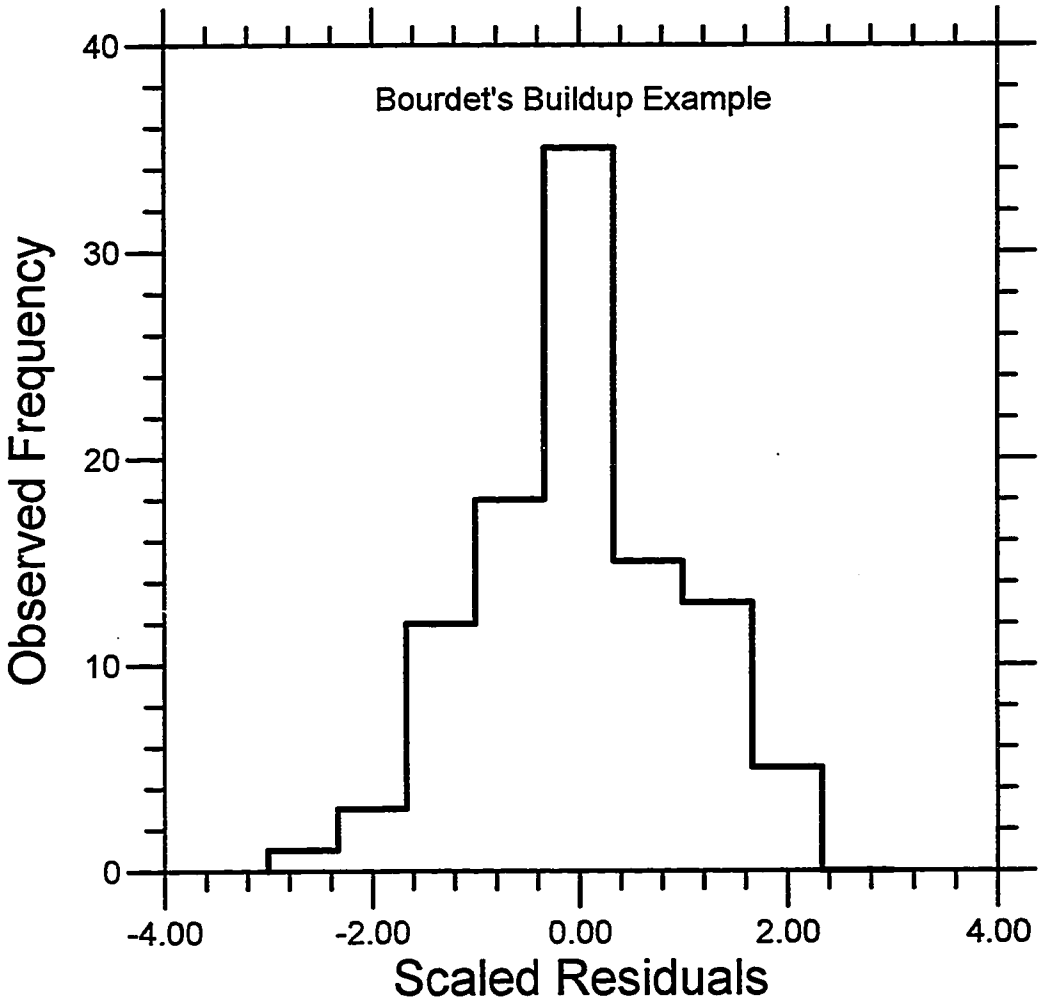

Fig. 5.3.5 - Histogram Plot - Field Example #2

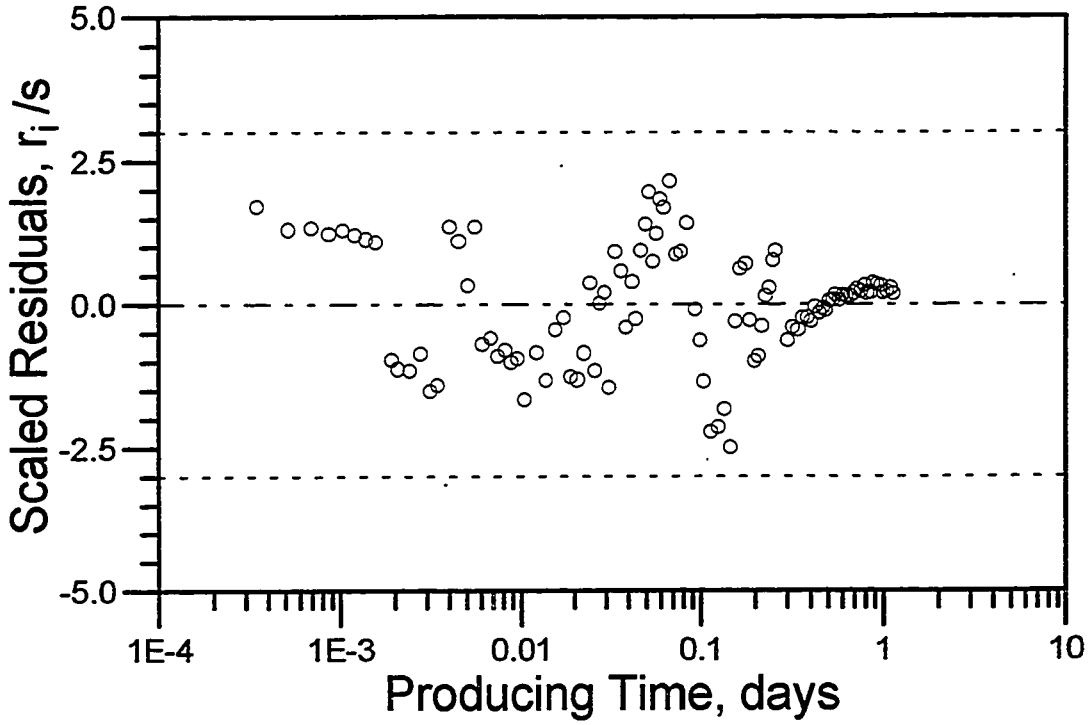

Fig. 5.3.6 - Scaled residual plot - Field Example #2

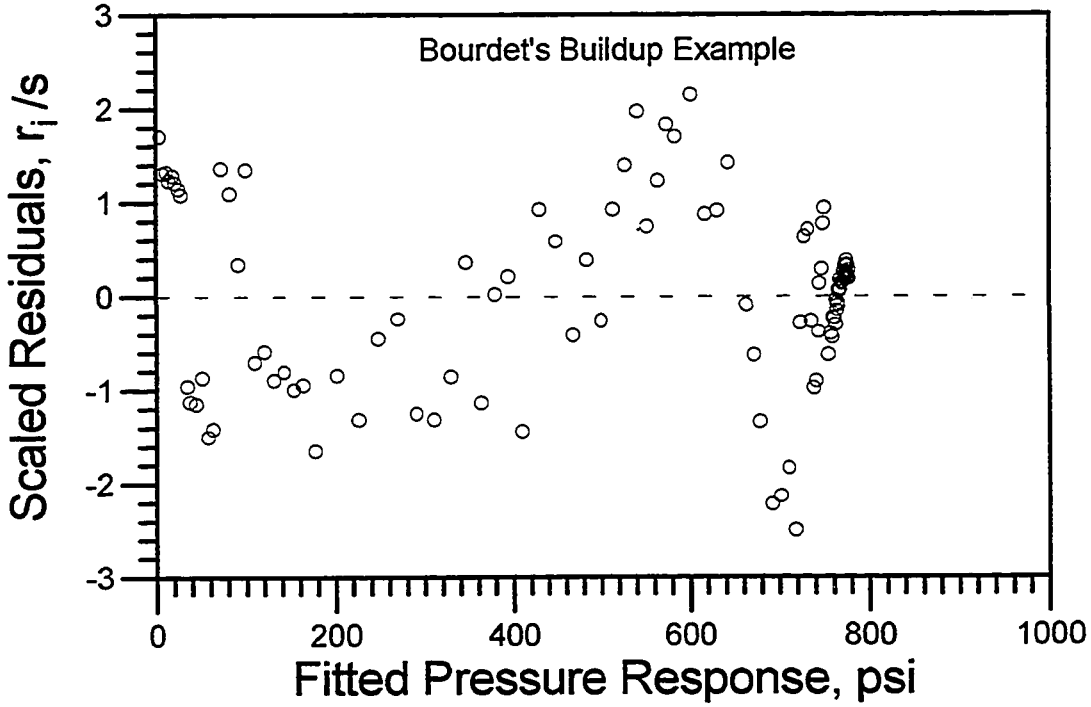

Fig. 5.3.7 - Diagnostic plot - Field Example #2

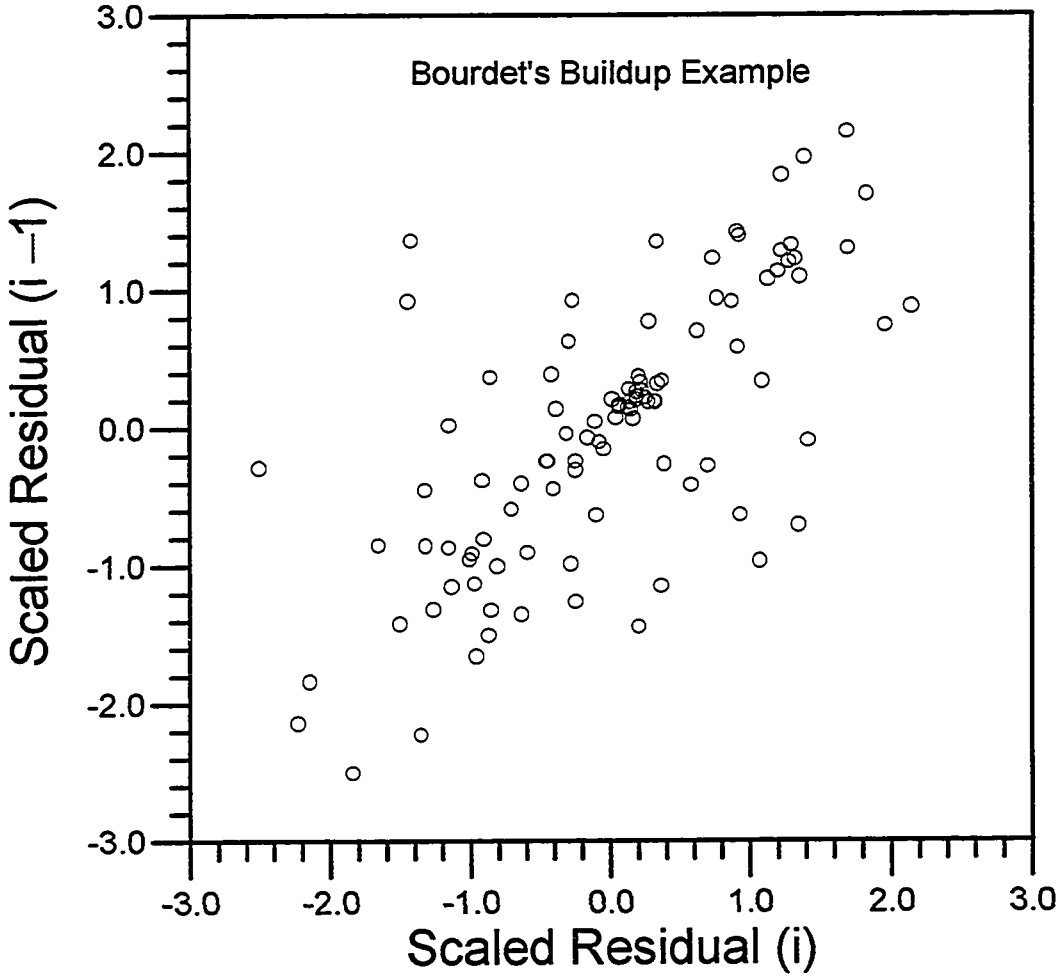

Fig. 5.3.8 - Serial Correlation Plot - Field Example #2

Figures 5.3.6 through 5.3.8 present different diagnostic plots of the scaled residuals, r_i/s. All three plots exhibit some structure, especially the serial correlation plot (see Fig. 5.3.8), which evidence that the residuals are somewhat correlated. Also, Fig. 5.3.6 shows that the assumption of constant variance of the errors may be violated since the variance of residuals are smaller at the end of the test. For such a problem, it would be preferable to apply a weighted least squares estimation, but this would require specific knowledge of the variance of each residual.

Even though the correlation of the residuals evidenced in this example does not invalidate the estimates of the parameters, more work is needed to properly account for correlation between errors and a non-constant variance of the error term.

5.3.3 Synthetic Example

Here, we show the diagnostic plots of the residuals for the synthetic example (with the seed for the random sequence equal to −21) which was analyzed in Chapter III, Section 3.2.3. Fig. 5.3.9 present the histogram plot of the residuals. Figures 5.3.10 through 5.3.12 show various diagnostic plots. As expected, no structure or correlation is evident.

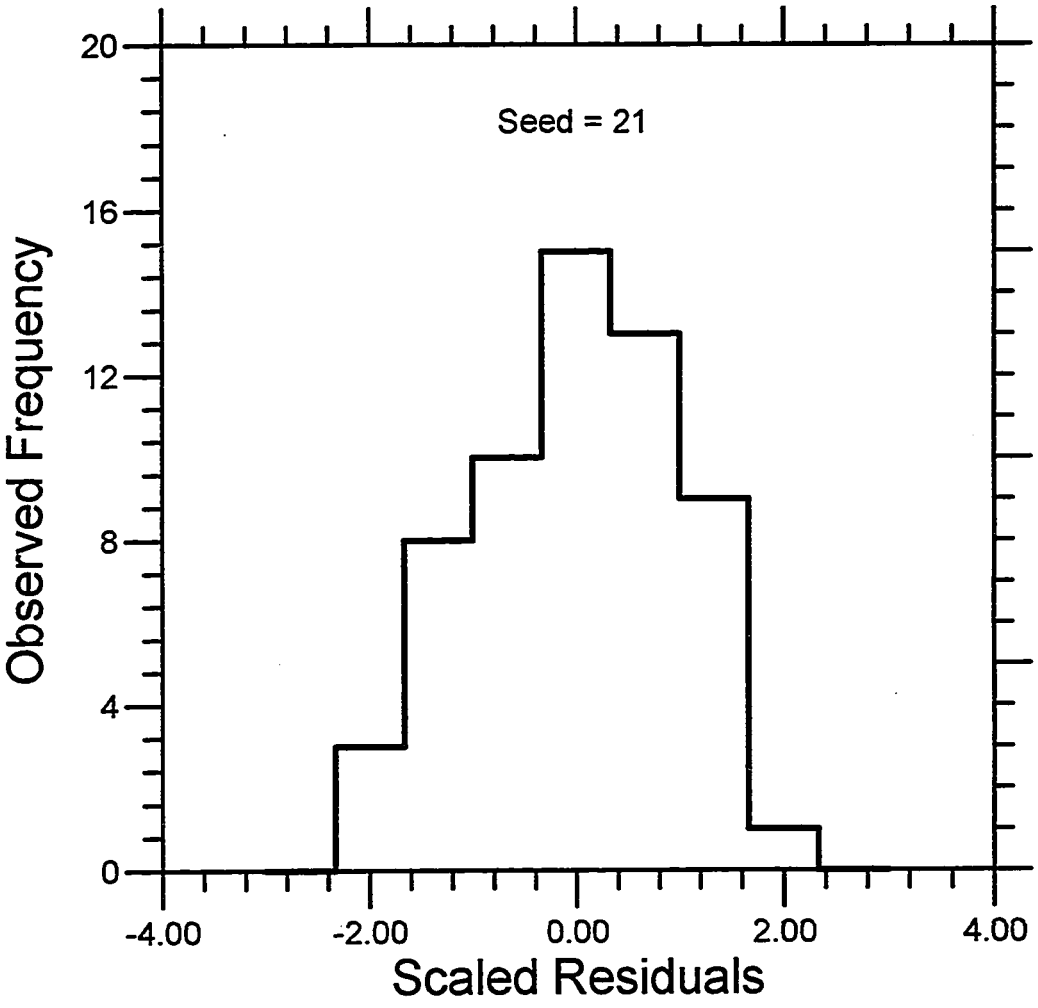

Fig. 5.3.9 - Histogram Plot - Synthetic Example

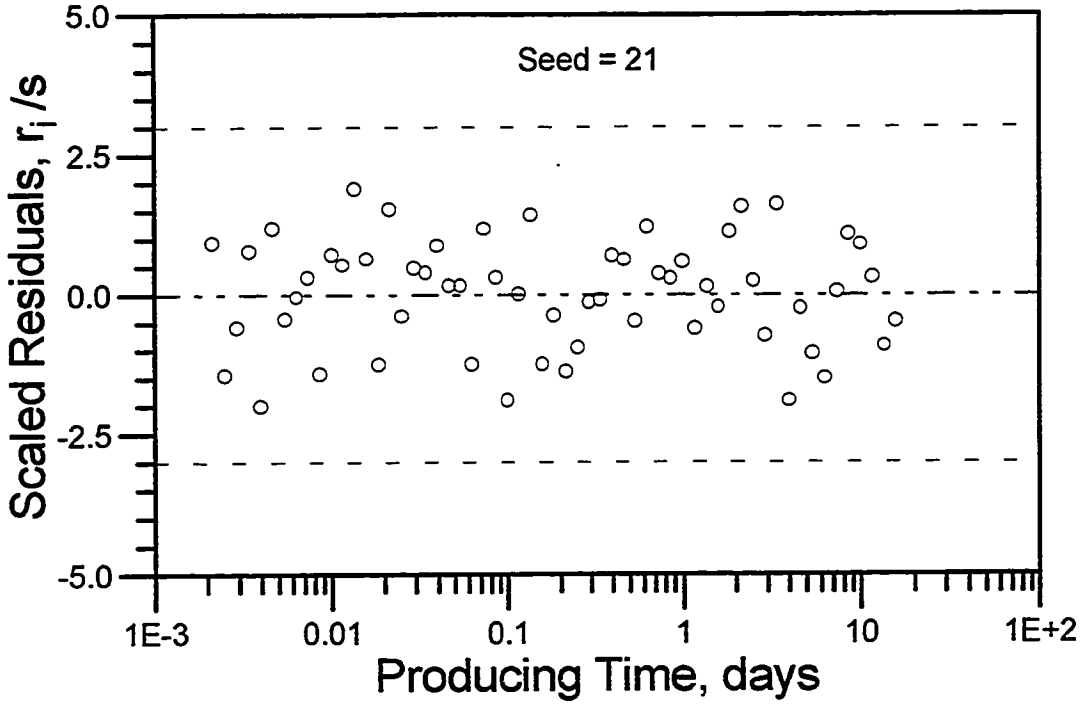

Fig. 5.3.10 - Scaled residual plot - Synthetic Example

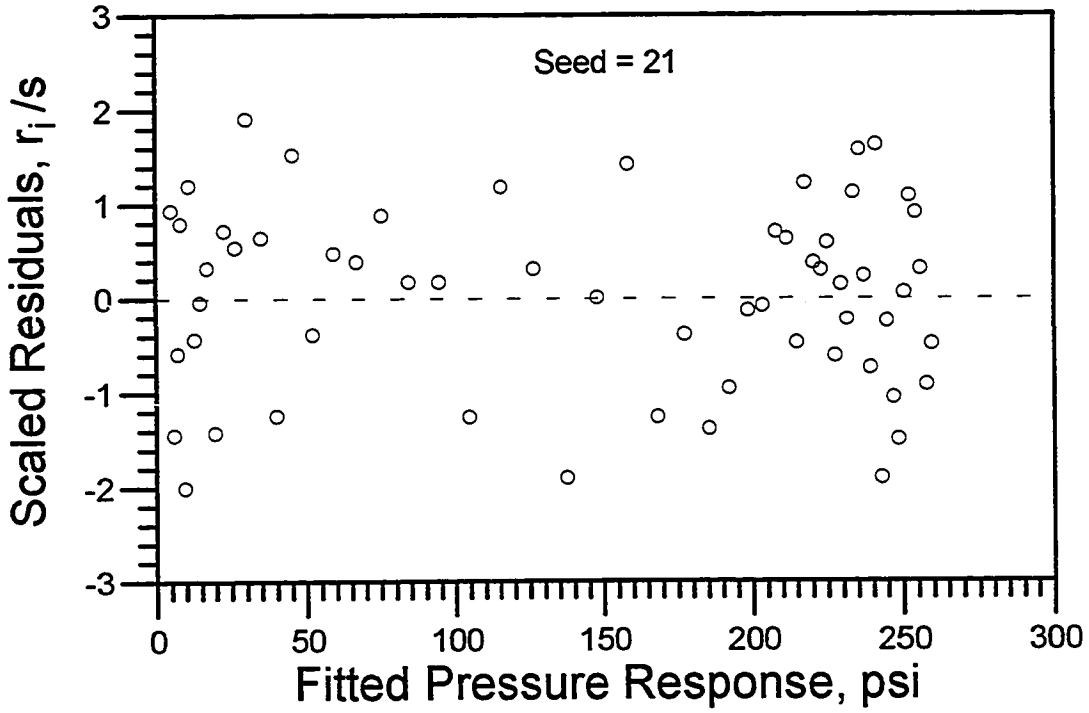

Fig. 5.3.11 - Diagnostic plot - Synthetic Example

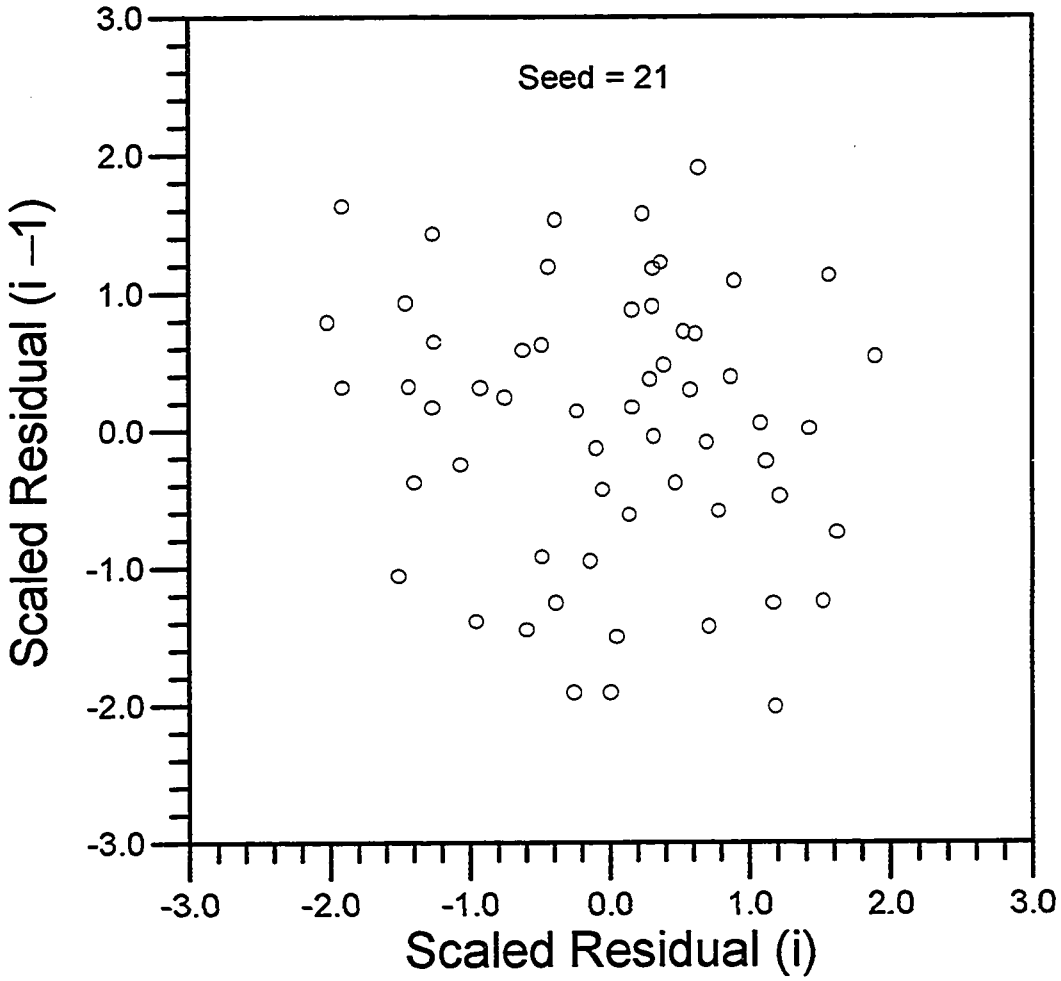

Fig. 5.3.12 - Serial Correlation Plot
Synthetic Example

CHAPTER VI

ADDITIONAL TOPICS IN WELL-TEST ANALYSIS

6.1 Convolution and Deconvolution

The relationship between the unit-rate flowing pressure response, $\Delta p_{su}(t)$, and the general wellbore pressure change, $\Delta p(t)$, that occurs in response to a variable sandface flow rate, $q(t)$, is given by Duhamel's principle (multirate superposition).[42] If we assume that $q(t) = 0$ for $t < 0$, then

$$\Delta p(t) = \int_0^t q'(\tau)\Delta p_{su}(t-\tau)d\tau + q(0^+)\Delta p_{su}(t). \tag{6.1.1}$$

The subscript prime denotes derivative with respect to time. Eq. 6.1.1 can be written in the following equivalent form after integration by parts,

$$\Delta p(t) = \int_0^t q(\tau)\Delta p'_{su}(t-\tau)d\tau. \tag{6.1.2}$$

In Eqs. 6.1.1 and 6.1.2, $\Delta p_{su}(t)$ includes the skin factor s, i.e., denoting the fundamental unit-rate solution with zero skin by $\Delta p_u(t)$, then

$$\Delta p_{su}(t) = \Delta p_u(t) + \Delta p_{skin}. \tag{6.1.3}$$

The Duhamel principle can also be applied at the surface or at any point in the wellbore (see for example Refs. 30 and 52). The downhole pressure change at any point in the wellbore, $\Delta p_b(t)$, can be written as

$$\Delta p_b(C_b, s, t) = \int_0^t q_b(\tau)\Delta p'_{bsu}(C_b, s, t-\tau)d\tau, \tag{6.1.4}$$

where Δp_{bsu} is the unit-rate response of the system including the effects of skin and wellbore storage, C_b, which is due to the wellbore volume below the point at which the rate is measured. Note that Eq. 6.1.4 is strictly valid only for the case of constant C_b. The relationship between the sandface flowrate and the downhole flowrate (at the measuring point) is[52]

$$q_b(t) = q(t) + C_b \frac{d\Delta p_w}{dt}, \tag{6.1.5}$$

where time is in days and q_b and q are in RB/D.

Convolution refers to the direct problem of calculating the variable rate response by superposition of the unit-rate response with the time dependent inner boundary condition, $q(t)$. For example, convolution can be used to compute the variable-rate model function. *Deconvolution* refers to the inverse problem of determining the unit-rate response based on the knowledge (or measurements) of $\Delta p(t)$ and $q(t)$. In practice, the fundamental difference between deconvolution and convolution is that the former does not assume any specific model for the unit-rate response, which makes deconvolution (more specifically the pressure derivative of the unit-rate response) an important tool for model recognition. Many authors (see for example Refs. 12, 52 and 53) have reported that deconvolution methods can improve the analysis of multi-rate tests by removing the undesirable rate variations from the pressure measurements. This is especially true for horizontal-well tests where the wellbore storage is usually large and masks important early-time pressure response.

It is accepted today that the deconvolution is performed better in Laplace space where it is much easier and more stable in most cases. This is in contrast to real time deconvolution algorithms, such as described by Kuchuk[52,54] and Thompson and Reynolds.[55] Taking the Laplace transform of Eqs. 6.1.1 or 6.1.2 with respect to time, the convolution integral reduces to a simple multiplication of the transformed pressure and rate variables,

$$\overline{\Delta p}(u) = u\overline{q}(u).\overline{\Delta p}_{su}(u), \tag{6.1.6}$$

where u is the Laplace variable with respect to t. Thus the real time unit-rate response can be computed via,

$$\Delta p_{su}(t) = \mathcal{L}^{-1}\left\{\frac{\overline{\Delta p}(u)}{u\overline{q}(u)}\right\}. \tag{6.1.7}$$

The Stehfest algorithm[7] is normally employed to numerically invert Eq. 6.1.7. In 1988, Roumboutsos and Stewart[12] presented a numerical algorithm to take the Laplace transform of pressure and rate measurements, and proposed using Eq. 6.1.7 to deconvolve pressure transient data. We will discuss the forward Laplace transform algorithms in the next section.

For the case where the surface flow rate (converted to bottom hole conditions by multiplying by the formation volume factor), q_{surf}, is constant and the sandface rate variation is only due to afterflow, the classical constant-wellbore-storage model[34] predicts the sandface flow rate as

$$q(t) = q_{surf} - C\frac{d\Delta p}{dt}. \tag{6.1.8}$$

Note that Eq. 6.1.8 is also valid if q_{surf} is a function of time. Substituting Eq. 6.1.8 into Eq. 6.1.7, the deconvolution equation for wellbore storage can be expressed as

$$\Delta p_{su}(t) = \mathcal{L}^{-1}\left\{\frac{\overline{\Delta p}(C,s,u)}{q_{surf} - Cu^2\overline{\Delta p}(C,s,u)}\right\}. \tag{6.1.9}$$

Eqs. 6.1.7 and 6.1.8 assume Δp is measured at the sandface (inside the wellbore). If Δp is measured in any other location in the wellbore between the sandface and the surface, the proper deconvolution equations come from applying the Laplace transform to Eq. 6.1.4. Thus, the deconvolution equations become

$$\Delta p_{bsu}(C_b,s,t)=\mathcal{L}^{-1}\left\{\frac{\overline{\Delta p_b}(u)}{u\overline{q_b}(u)}\right\} \tag{6.1.10}$$

and

$$\Delta p_{bsu}(C_b,s,t)=\mathcal{L}^{-1}\left\{\frac{\overline{\Delta p}_b(u)}{q_{surf}-(C-C_b)u^2\overline{\Delta p}_b(u)}\right\}. \tag{6.1.11}$$

For most tests, the pressure recorder and the flowmeter are placed just above the interval to be tested. In spite of that, C_b may be significantly large in horizontal-well tests and wells with rat holes. $(C-C_b)$ corresponds to the wellbore-storage volume between the production tool and the surface. It should be emphasized that Δp_{bsu} is the response of the system for a unit-rate boundary condition at the measurement point, therefore it includes the wellbore storage effects below that point.

Some authors (see for example Refs. 52 and 56) pointed out that the Laplace space deconvolution exhibits stability problems at early times for wellbore-storage deconvolution. The instability arises because it is difficult to evaluate the denominator of Eq. 6.1.9 accurately at early times. First, the denominator of Eq. 6.1.9 is virtually zero at very early times, when the sandface flow rate (i.e., the reservoir contribution to the total flow rate) is negligible. In addition, instability occurs because, at early times, the sandface flow rate is too small to be accurately computed by subtracting two numbers very close to each other. Thus, at early times, even a small inaccuracy in $\overline{\Delta p}$, or in our estimate of C, is enough to cause a very large error in Δp_{su}. As pointed out in Refs. 52 and 56, the instability problem may be minimized by removing only part of the wellbore storage. That is, instead of performing the deconvolution at the sandface (Eq. 6.1.9), one may use Eq. 6.1.11 that keeps a residual storage in Δp_{bsu}, which tends to stabilize the deconvolution. However, clearly, at the beginning, $\Delta p_{bsu}(t)$ will not follow the theoretical unit-rate response with zero wellbore storage, $\Delta p_{su}(t)$. This fact is shown in Figs. 4.1.2 through 4.1.4, where we applied deconvolution to a synthetic horizontal-well buildup test that was

generated with $C = 0.02$ bbl/psi. Note that the deconvolution is not accurate for the first log-cycle. Fig. 4.1.4 shows that the derivative data can be computed numerically in real space (using the deconvolved pressures), or in Laplace space. In this case, we use the Stehfest algorithm to invert the data to real space.

Fair and Simmons[53] have recently proposed a trial and error deconvolution, which applies Eq. 6.1.9 to progressively remove the afterflow effects by using increasingly larger values of wellbore storage C.

6.1.2 Numerical Laplace Transform of Tabulated Data

We discussed in the previous section the necessity of an algorithm to convert the tabulated pressure data to Laplace space. By definition, the Laplace transform of a general function $\Delta p(t)$ requires the knowledge of the function Δp over the whole time domain, $0 \le t < \infty$. Thus, some kind of extrapolation must be performed when the function is known only in a given time range as in the case with measured well test pressure and/or rate data. The key point of a good forward Laplace transform algorithm is to properly account for the extrapolation of the data so that it does not affect the interpretation and analysis of the data.

In 1988, Romboutsos and Stewart[12] presented an algorithm based on a linear interpolation of the measured data and a linear extrapolation using the cord slope of the last interval. Clearly, their algorithm is more suitable for boundary dominated flow where the pressure behaves linearly with respect to time. During transient flow, instead of a linear extrapolation, Bourgeois and Horne[56] assume a semilog extrapolation of the data. They also considered a semilog extrapolation with respect to the Horner time ratio in the case of buildup. Of course this algorithm is more appropriate for data that exhibit a semilog straight line at the end of the test.

In Appendix F, we derive an algorithm that considers a linear interpolation of the function (in the same way as proposed by Romboutsos and Stewart[12]), and a linear extrapolation of the data in terms of its *pressure derivative*. Here, pressure derivative means derivative of pressure data with respect to the logarithm of time. Our algorithm for the Laplace transform of a function and its logarithmic derivative are given by Eqs. F.13 and F.15, respectively. In Appendix F, we also show that the proposed algorithm incorporates both the Romboutsos-Stewart algorithm (linear extrapolation of the function) and the Bourgeois-Horne algorithm[56] (semilog extrapolation of the data) as particular cases.

After we developed our method, Bourgeois and Horne[13] considered an algorithm that assumes the logarithmic pressure derivative curve is extrapolated linearly on a log-log plot. (The algorithm they presented did not appear in the original version[56] of their paper.) Note that they extrapolate (linearly) the logarithmic pressure derivative on a log-log plot, whereas we extrapolate the logarithmic pressure derivative on a Cartesian plot. Their algorithm can be written as

$$\overline{\Delta p}(u)=\frac{\Delta p_0}{u}+\frac{\Delta p_0'}{u^2}\left(1-e^{-ut_1}\right)+\sum_{i=1}^{n-1}\frac{\Delta p_i'}{u^2}\left(e^{-ut_i}-e^{-ut_{i+1}}\right)$$

$$+\frac{m}{\alpha(ut_n)^{\alpha}}\frac{1}{u}\left[\Gamma(\alpha+1,ut_n)-(ut_n)^{\alpha}e^{-ut_n}\right], \qquad (6.1.12)$$

where $\Gamma(a,x)$ is the incomplete Gamma function defined as

$$\Gamma(a,x)=\int_x^{\infty}e^{-t}t^{a-1}dt. \qquad (6.1.13)$$

Letting $v=\ln(t)$, $m=\Delta p'(v_n)$, i.e., m is the slope of the semilog straight-line extension, and

$$\alpha=\Delta p''(v_n)/\Delta p'(v_n). \qquad (6.1.14)$$

The superscripts prime and double prime denote first and second derivatives with respect to time, respectively. Throughout, the Romboutsos-Stewart algorithm is called Method A, our proposed algorithm (see Appendix F) is called Method B, and the Bourgeois-Horne algorithm (Eq. 6.1.12) is called Method C. In the next sections, we compare the performance of the three methods via three examples. The comparison is done in Laplace space by plotting $u\bar{p}$ and $ud(-u\bar{p})/du$ versus $1/u$, where u is the Laplace variable and $\bar{p}$ is the Laplace transform of the pressure data. We demonstrate that either Method B or Method C can be applied successfully, and that these methods are superior to Method A.

6.1.2.1 Synthetic Wellbore Storage and Skin Example

The pressure data for this example were generated from the Agarwal *et al.*[34] solution using the data of Table 6.1.1.

Table 6.1.1 - Input variables to generate the Synthetic Wellbore Storage and Skin Example

Permeability:	$k = 15$ md
Skin Factor:	$s = 10$
Wellbore Storage Coefficient:	$C = 0.02$ bbl/psi
Porosity:	$\phi = 0.25$
Thickness:	$h = 20$ ft
Viscosity:	$\mu = 1.0$ cp
Wellbore Radius:	$r_w = 0.25$ ft
Total Compressibility:	$c_t = 1\times10^{-5}$ psi^{-1}
Flow Rate:	$qB = 100$ bbl/day
Time period of the test:	0.0001 to 100days

The pressure data are numerically transformed to Laplace space for four different spans of the data, all of them starting at 10^{-4} days. The ending times are 0.341, 1.00, 2.15 and 10.0 days, respectively. Figs 6.1.1 through 6.1.8 show the results; $L = 0.35$ means that only the last 0.35 log cycles of data is used for extrapolation purposes.

Note that Method A (Romboutsos-Stewart algorithm) does not yield accurate results at the end of the data, especially for the pressure derivative. However, Method B (our algorithm) and Method C (Bourgeois-Horne algorithm), worked satisfactorily for all spans of the data. Note that for the shortest test (10^{-4} to 0.341 days), shown in Fig. 6.1.1 and 6.1.2, Method B deviates downwards from the correct solution at the very end of the derivative curve, while Method C deviates upwards.

6.1.2.2 Synthetic Horizontal-Well Example

This is a buildup test that was generated from the Odeh and Babu[38] horizontal-well solution using the data of Table 6.1.2. Figures 4.1.1A through 4.1.1C presents the notation for the horizontal well model. Again, the computed pressure data is numerically transformed to Laplace space for four different spans of the data, all of them starting at $.10^{-4}$ days. The ending times for each test are 0.0133, 0.113, 1.13 and 9.61 days, respectively. The plots for this example are presented in Figs. 6.1.9 through 6.1.16. As expected, Method A worked very well only when the well reached pseudosteady state (Fig. 6.1.15). Methods B and C give satisfactory results for all spans of the data, with Method B slightly better than Method C for estimating the derivative (see for example Fig. 6.1.14).

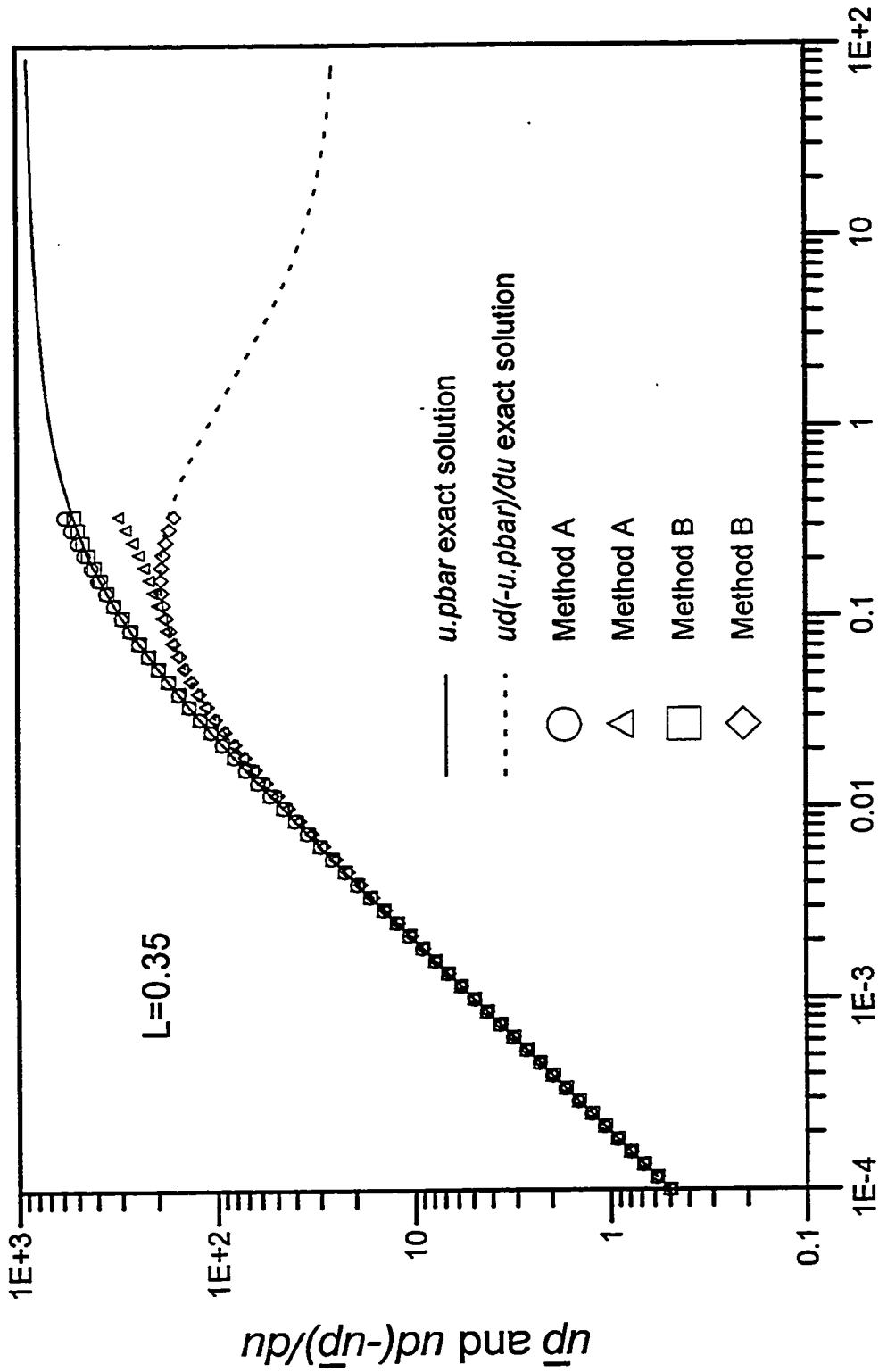

Fig. 6.1.1 - Numerical Laplace transform - Wellbore Storage & Skin Example

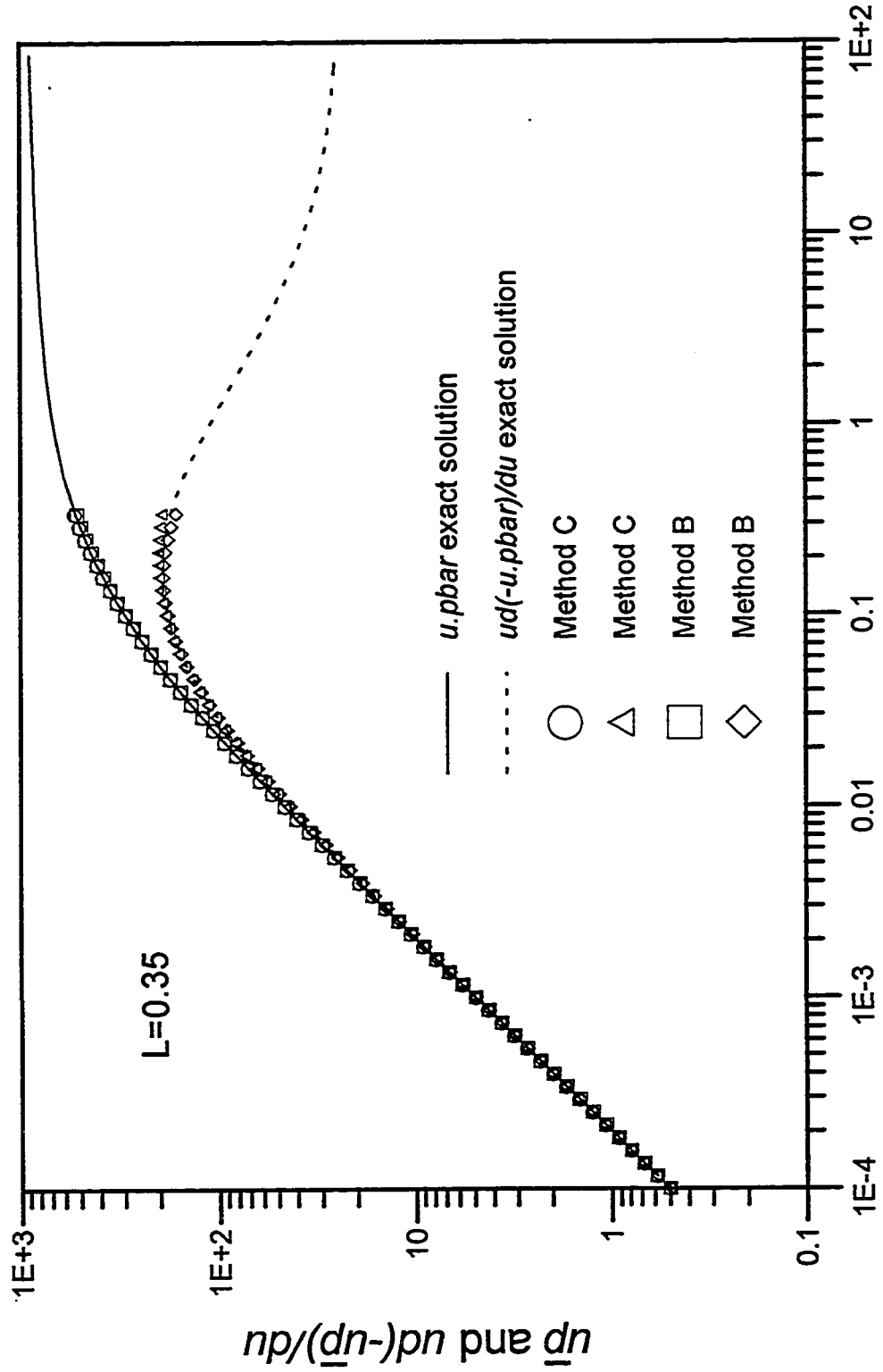

Fig. 6.1.2 - Numerical Laplace transform - Wellbore Storage & Skin Example

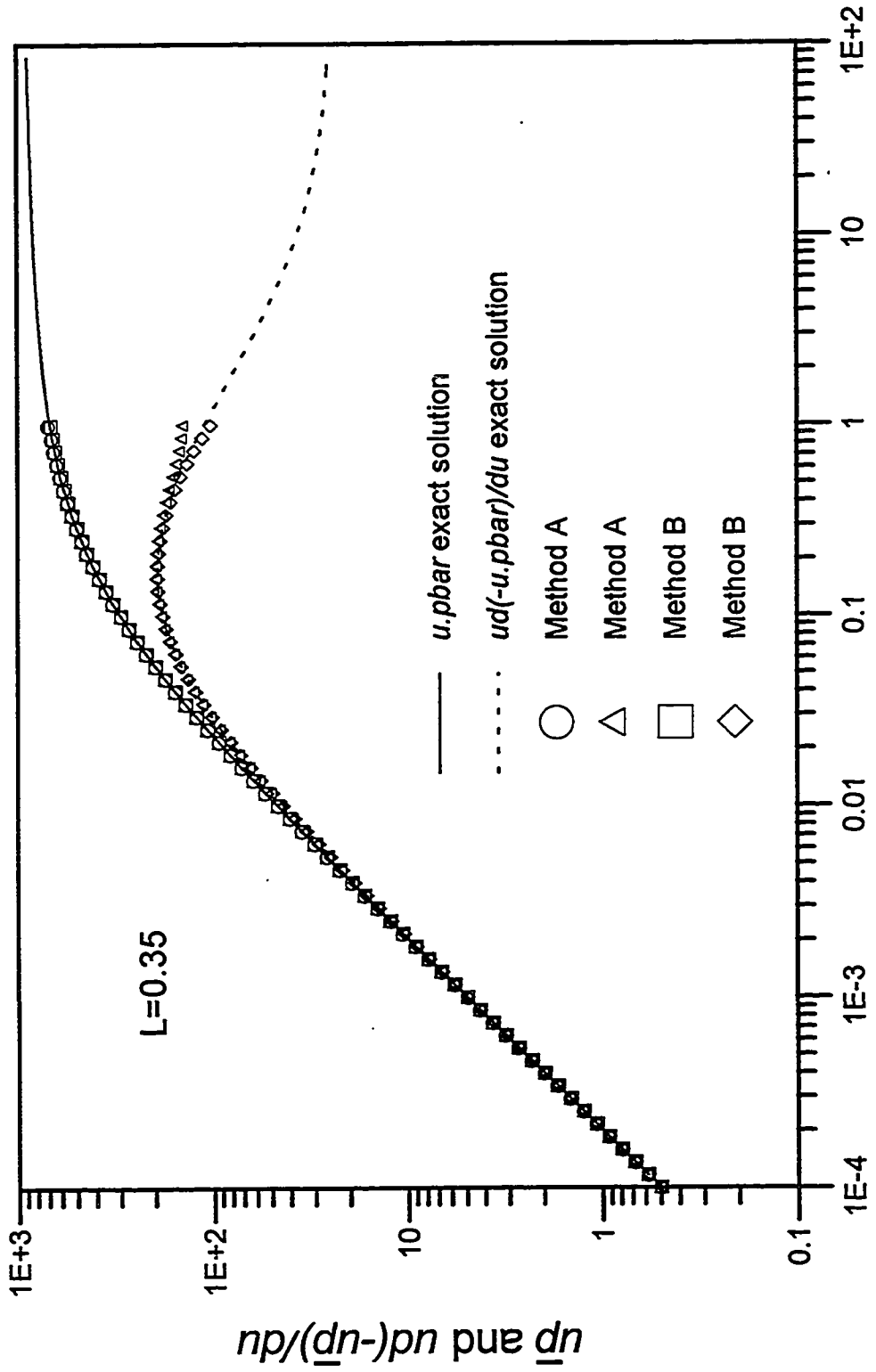

Fig. 6.1.3 - Numerical Laplace transform - Wellbore Storage & Skin Example

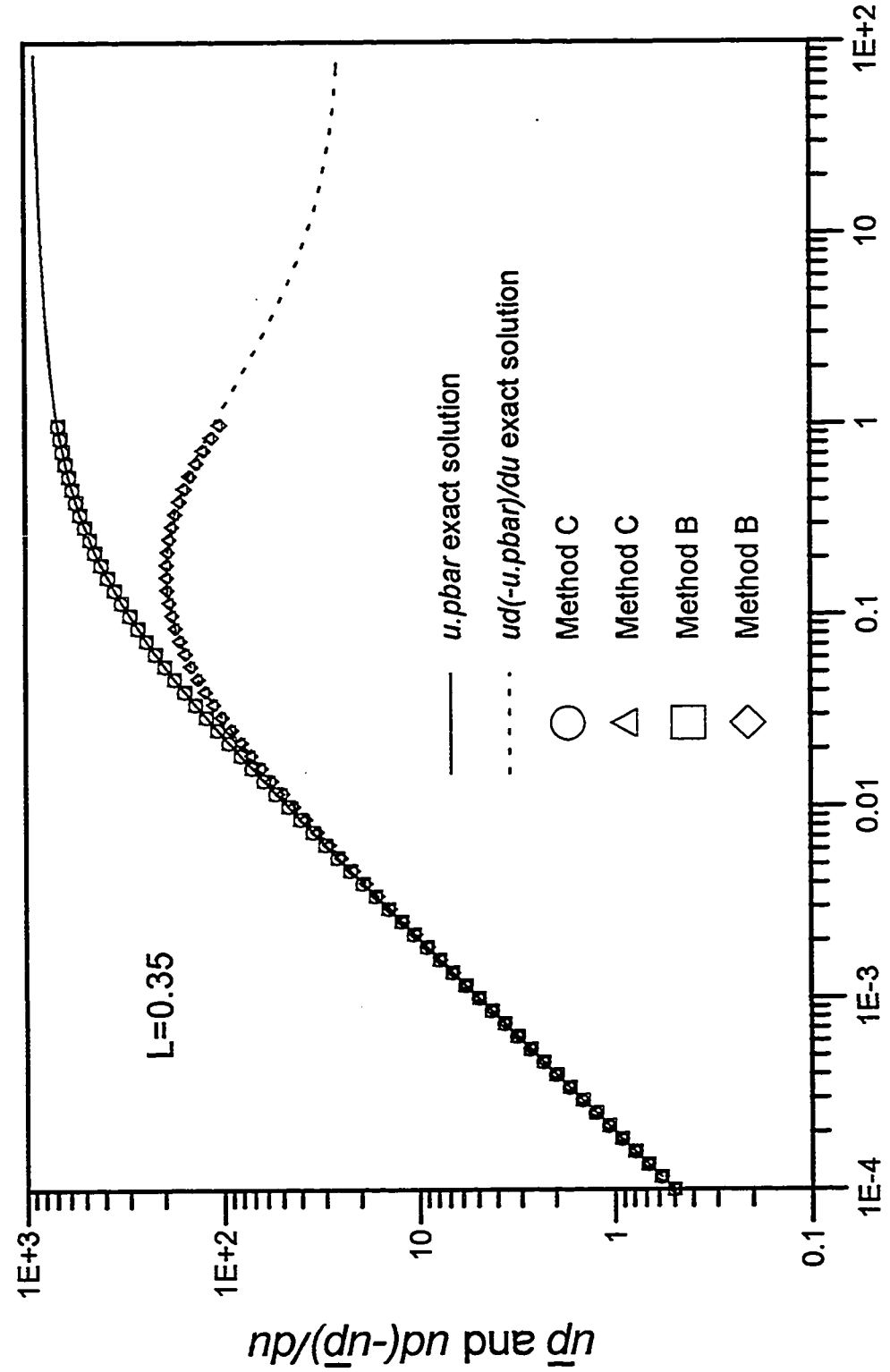

Fig. 6.1.4 - Numerical Laplace transform - Wellbore Storage & Skin Example

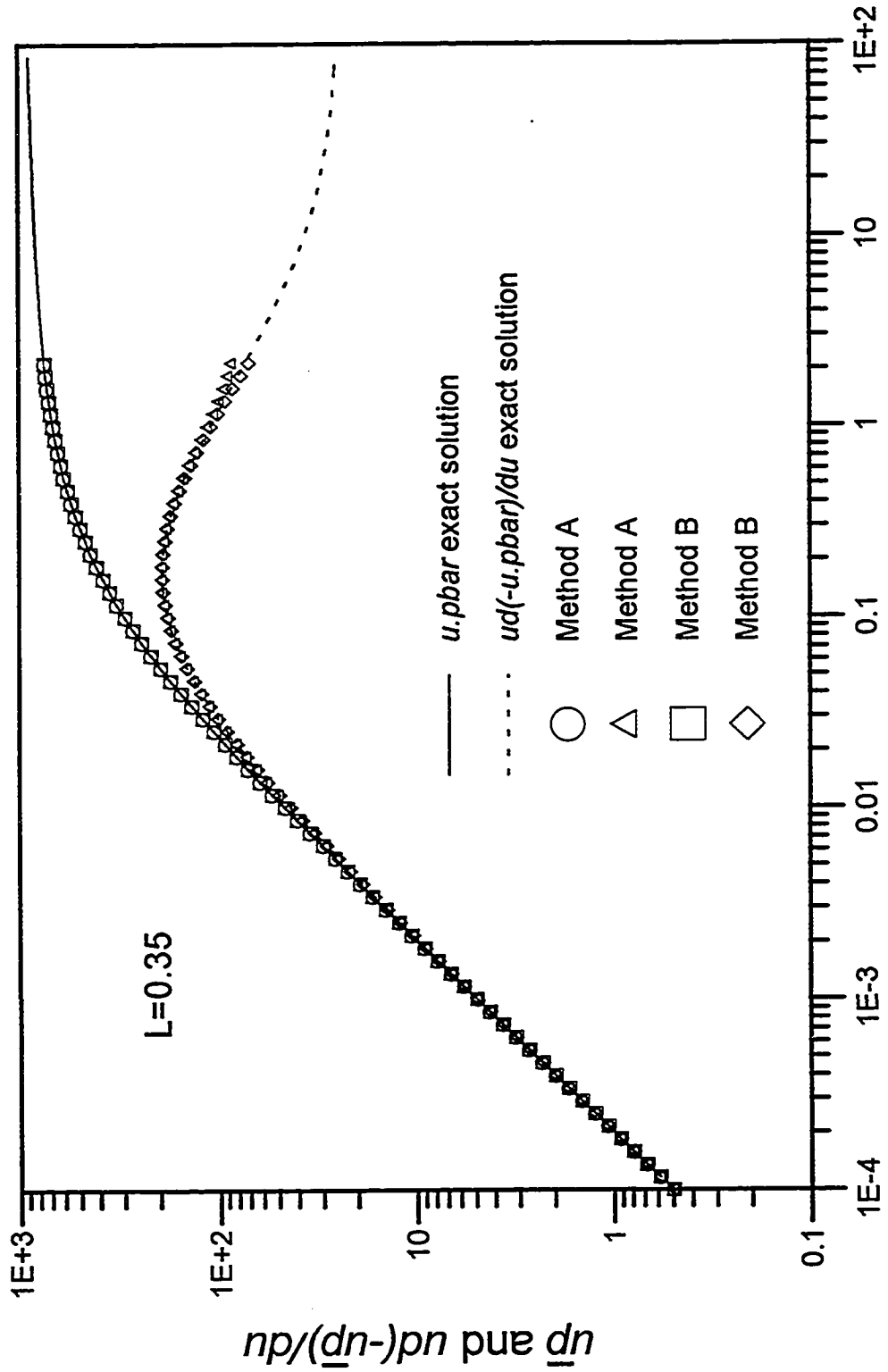

Fig. 6.1.5 - Numerical Laplace transform - Wellbore Storage & Skin Example

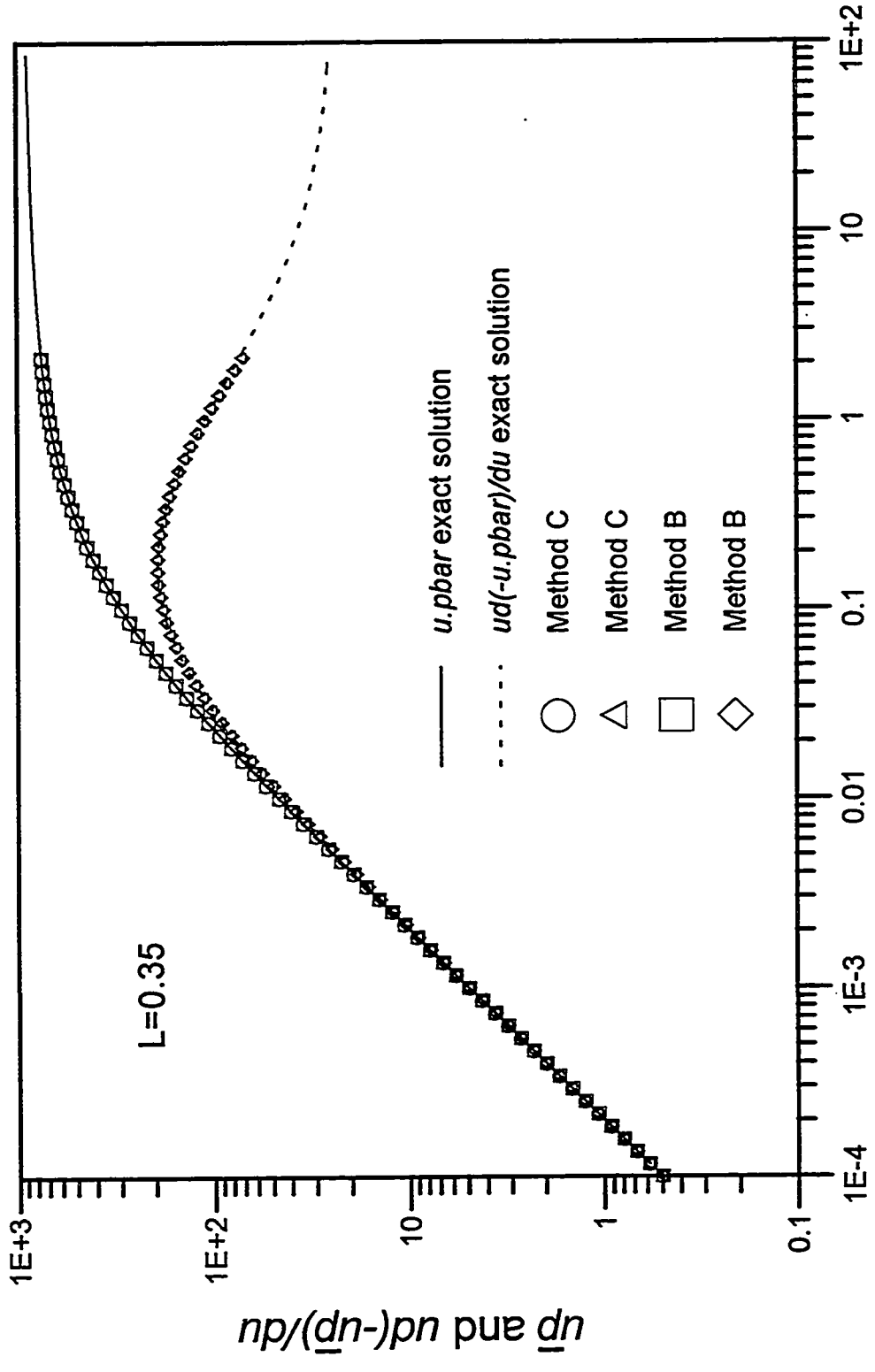

Fig. 6.1.6 - Numerical Laplace transform - Wellbore Storage & Skin Example

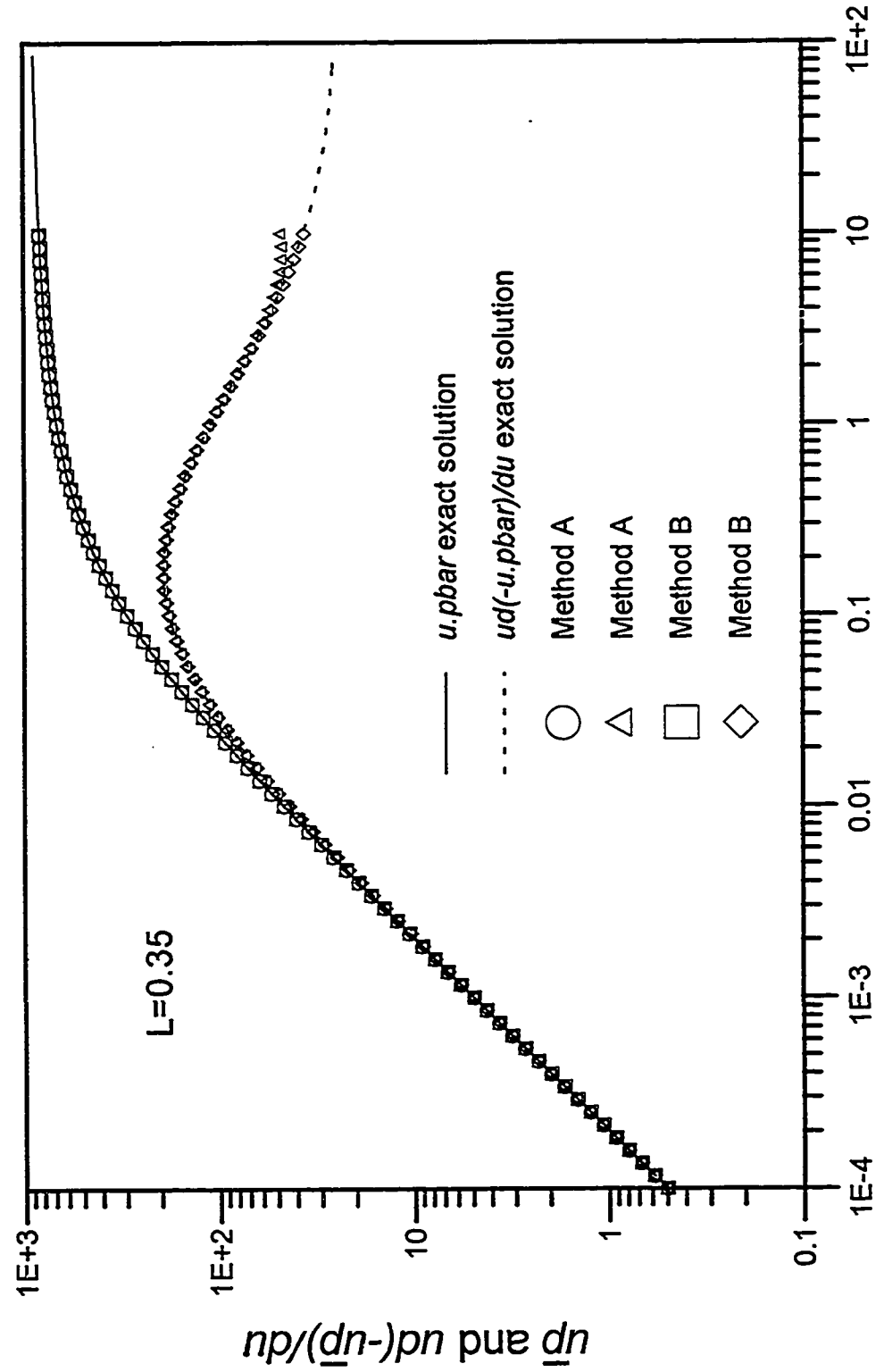

Fig. 6.1.7 - Numerical Laplace transform - Wellbore Storage & Skin Example

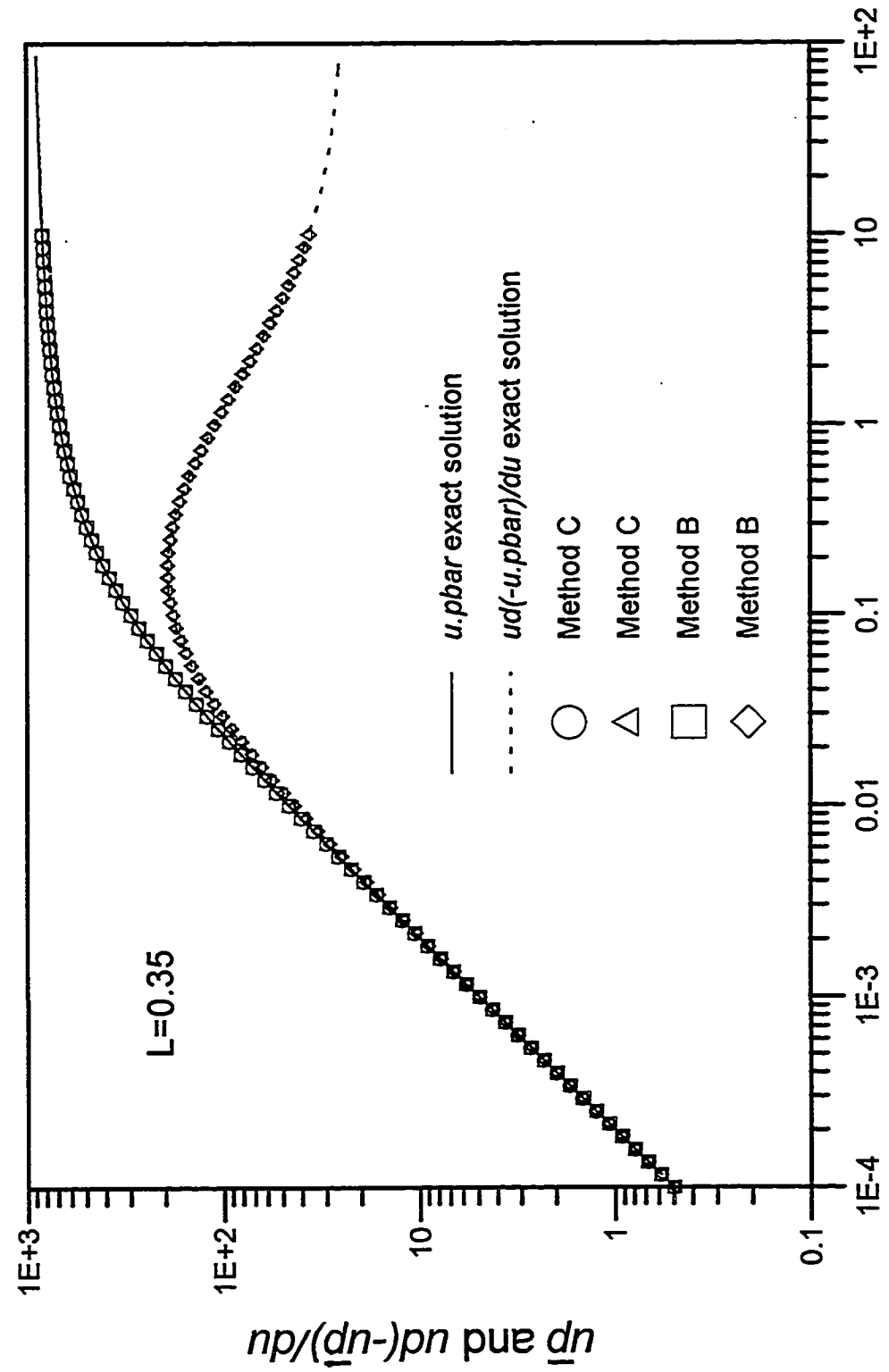

Fig. 6.1.8 - Numerical Laplace transform - Wellbore Storage & Skin Example

Table 6.1.2 - Input variables to generate the Synthetic Horizontal-Well Example (buildup test $q = 1000$ bbl/d)

Variable	Value
k_x (md)	5.
k_y (md)	10.
k_z (md)	1.
L_w (ft)	400.
L_x (ft)	2000.
L_y (ft)	2000.
x' (ft)	500.
y_1 (ft)	500.
L_z or h (ft)	50.
z' (ft)	10.
C (bbl/psi)	0.010
ϕ	0.30
Δp_s (psi/bbl/d)	0.1
r_w (ft)	0.35
μ_o (cp)	1.0
c_t (psi^{-1})	1.0×10^{-5}
# points per log-cycle	15

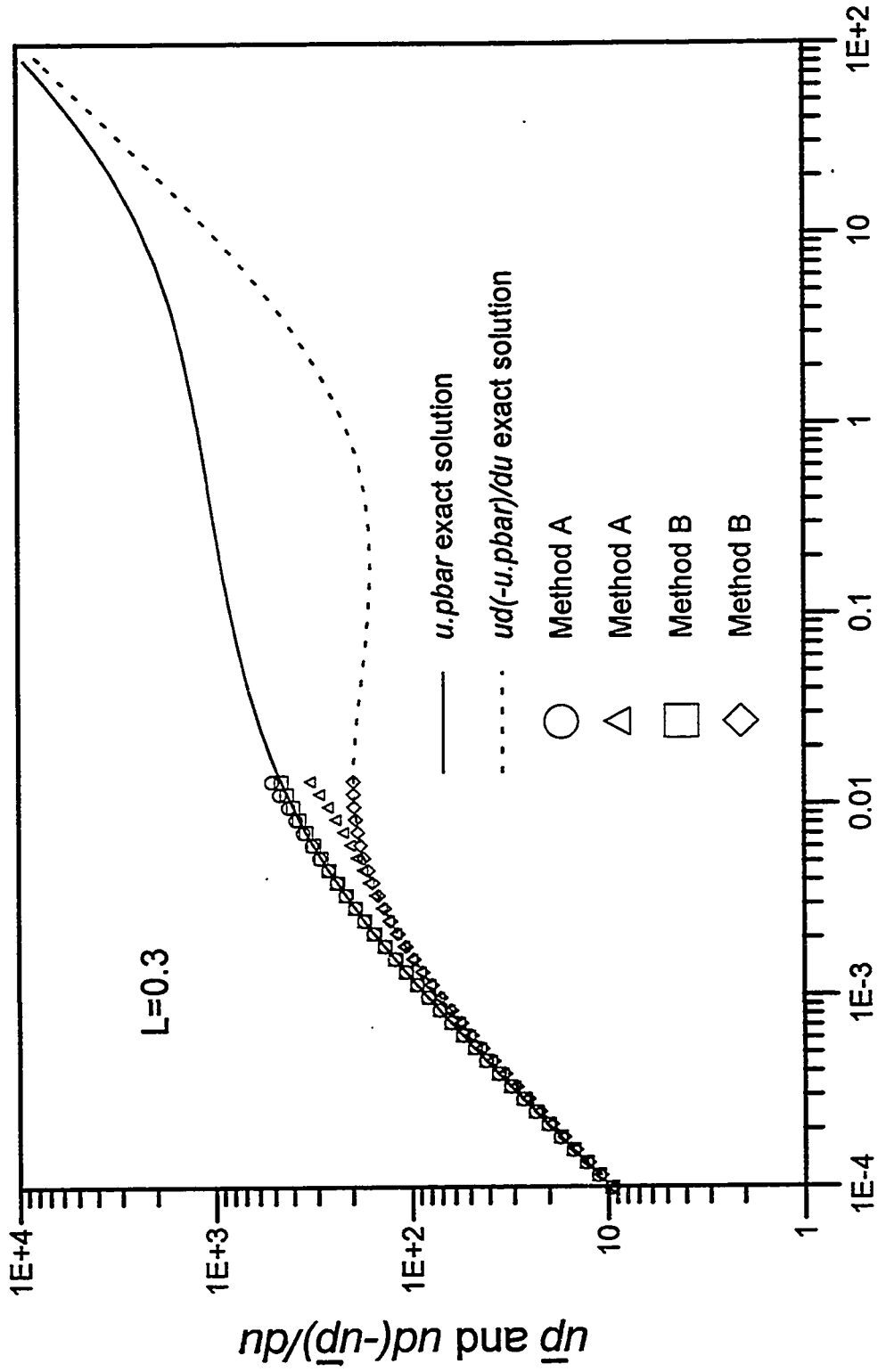

Fig. 6.1.9 - Numerical Laplace transform - Horizontal Well Example

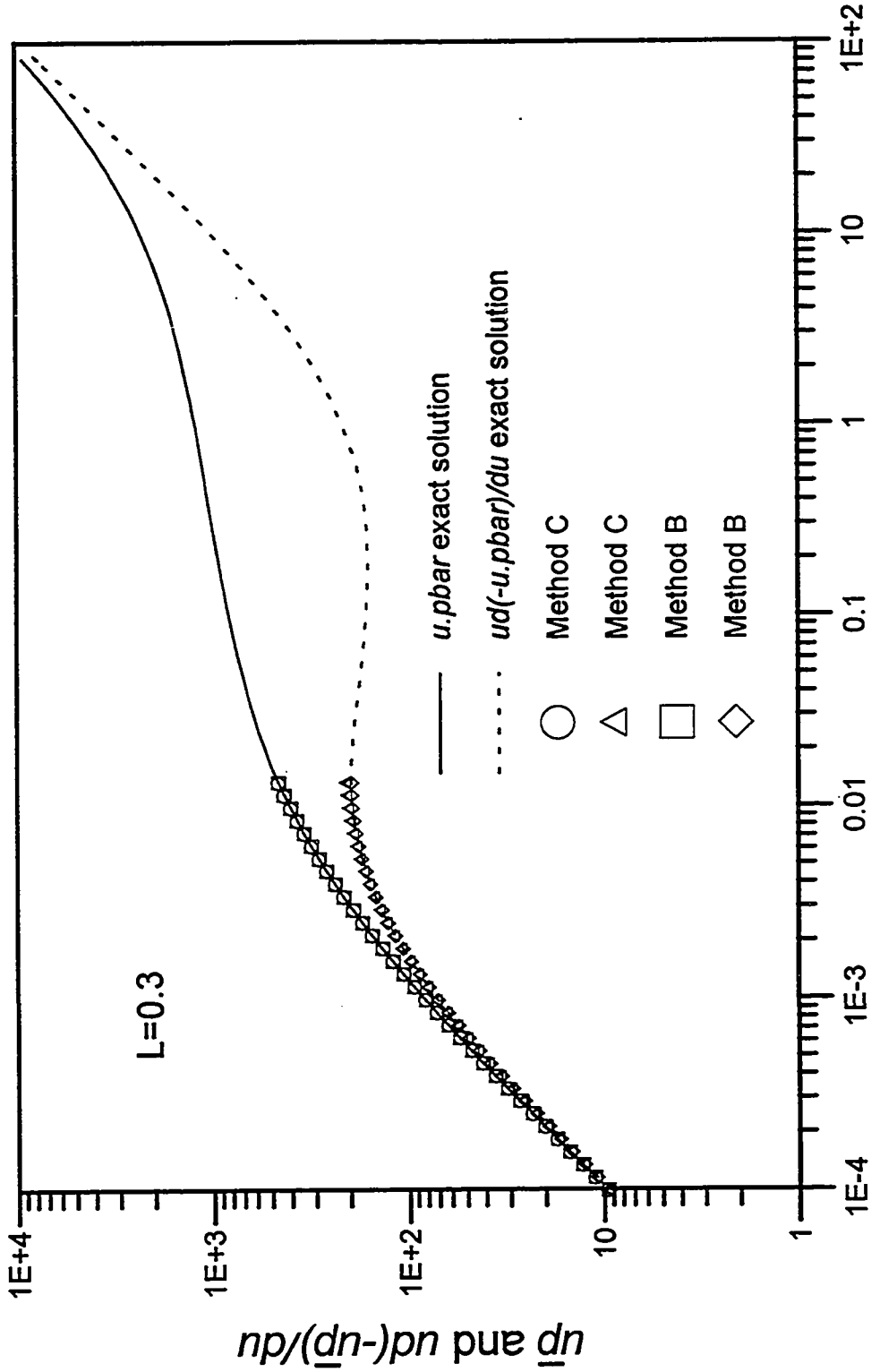

Fig. 6.1.10 - Numerical Laplace transform - Horizontal Well Example

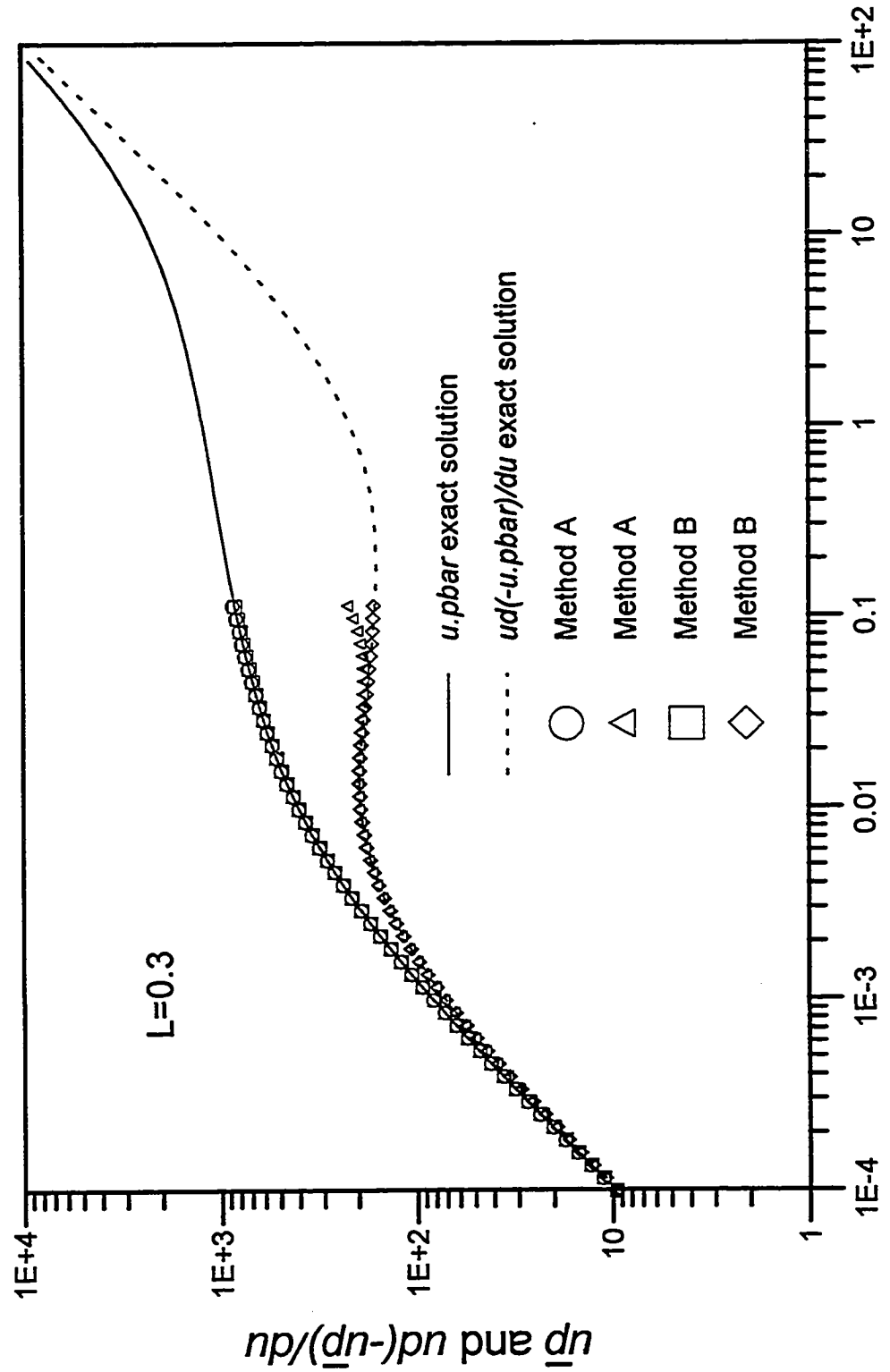

Fig. 6.1.11 - Numerical Laplace transform - Horizontal Well Example

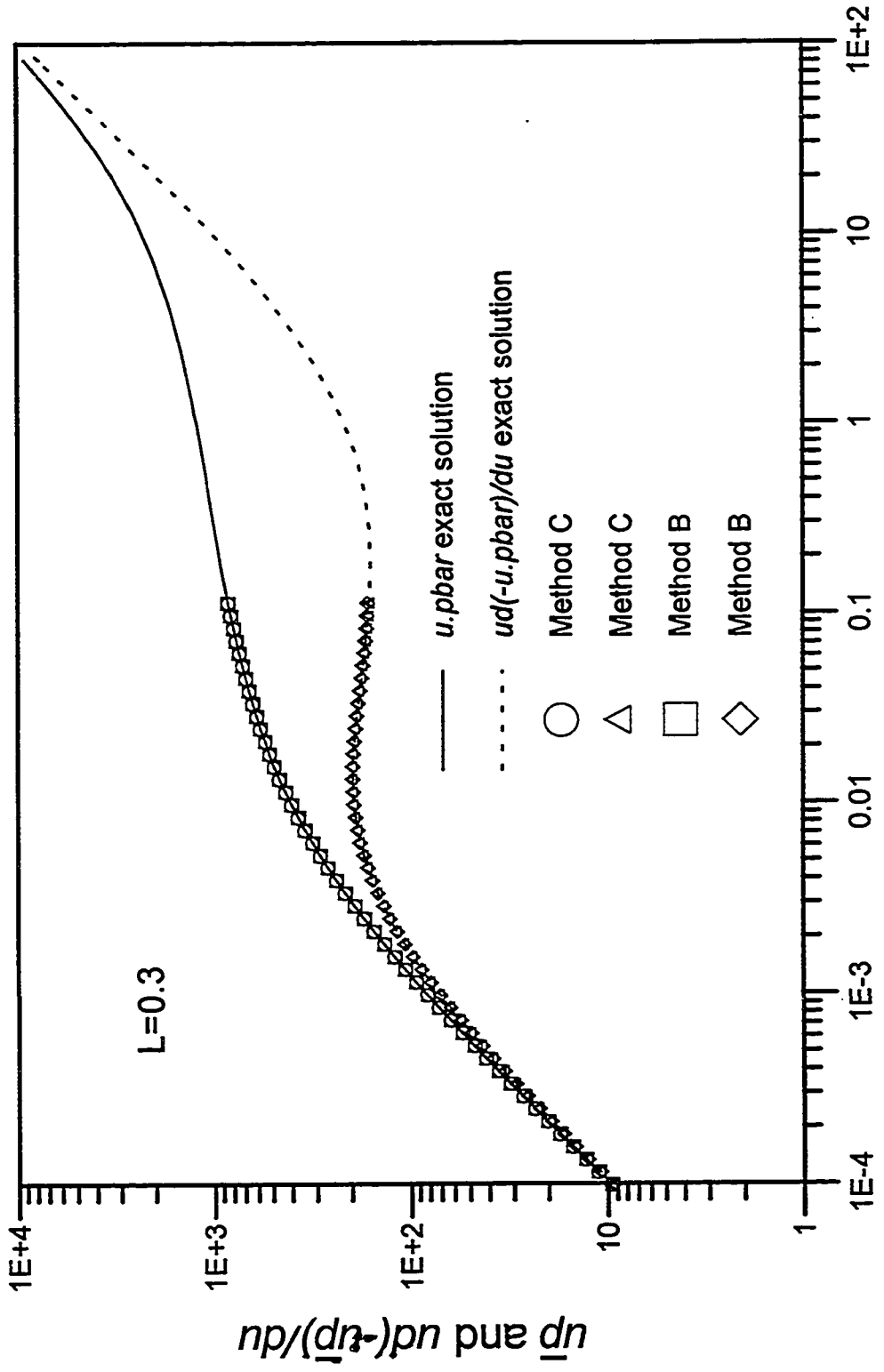

Fig. 6.1.12 - Numerical Laplace transform - Horizontal Well Example

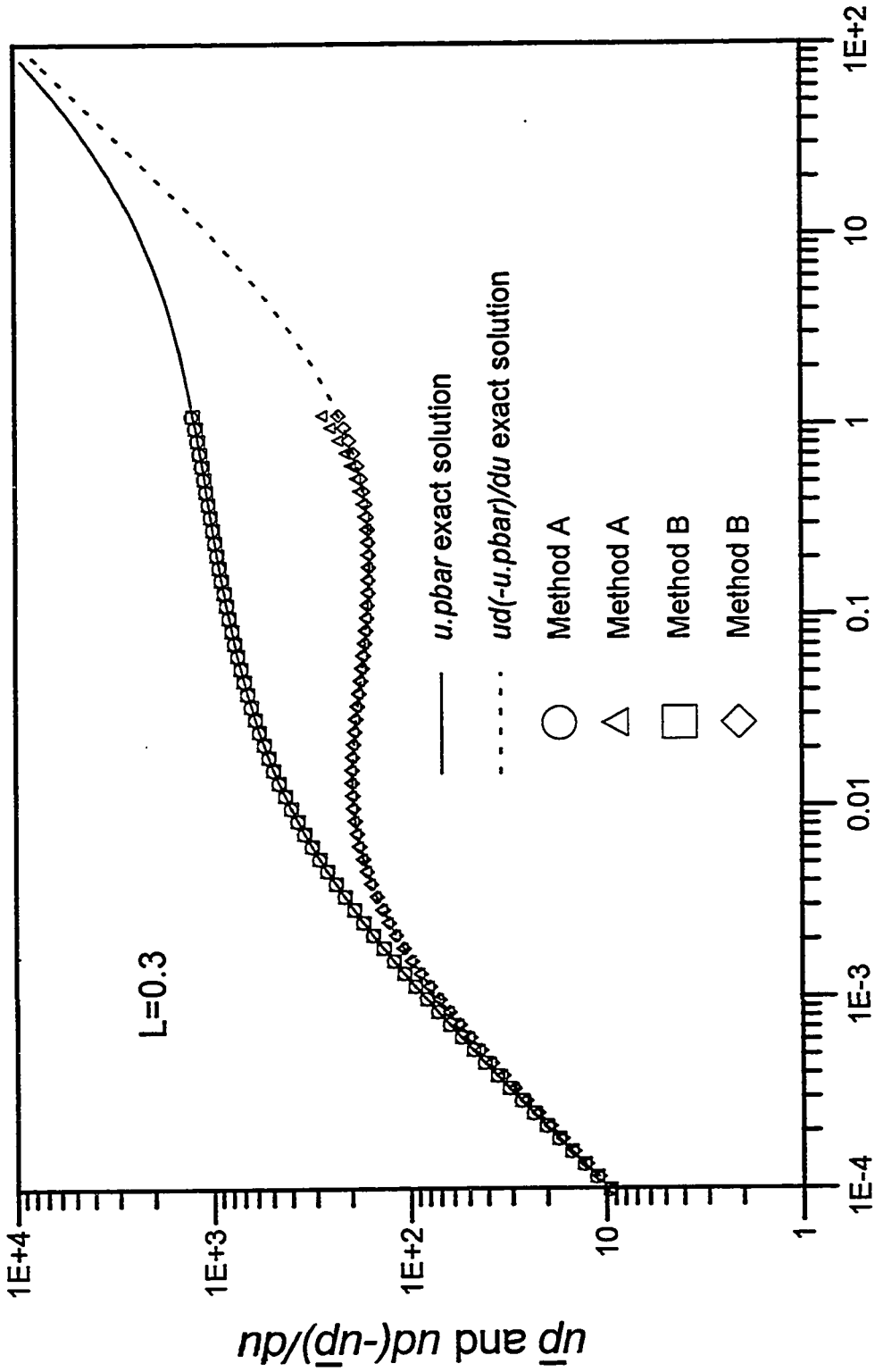

Fig. 6.1.13 - Numerical Laplace transform - Horizontal Well Example

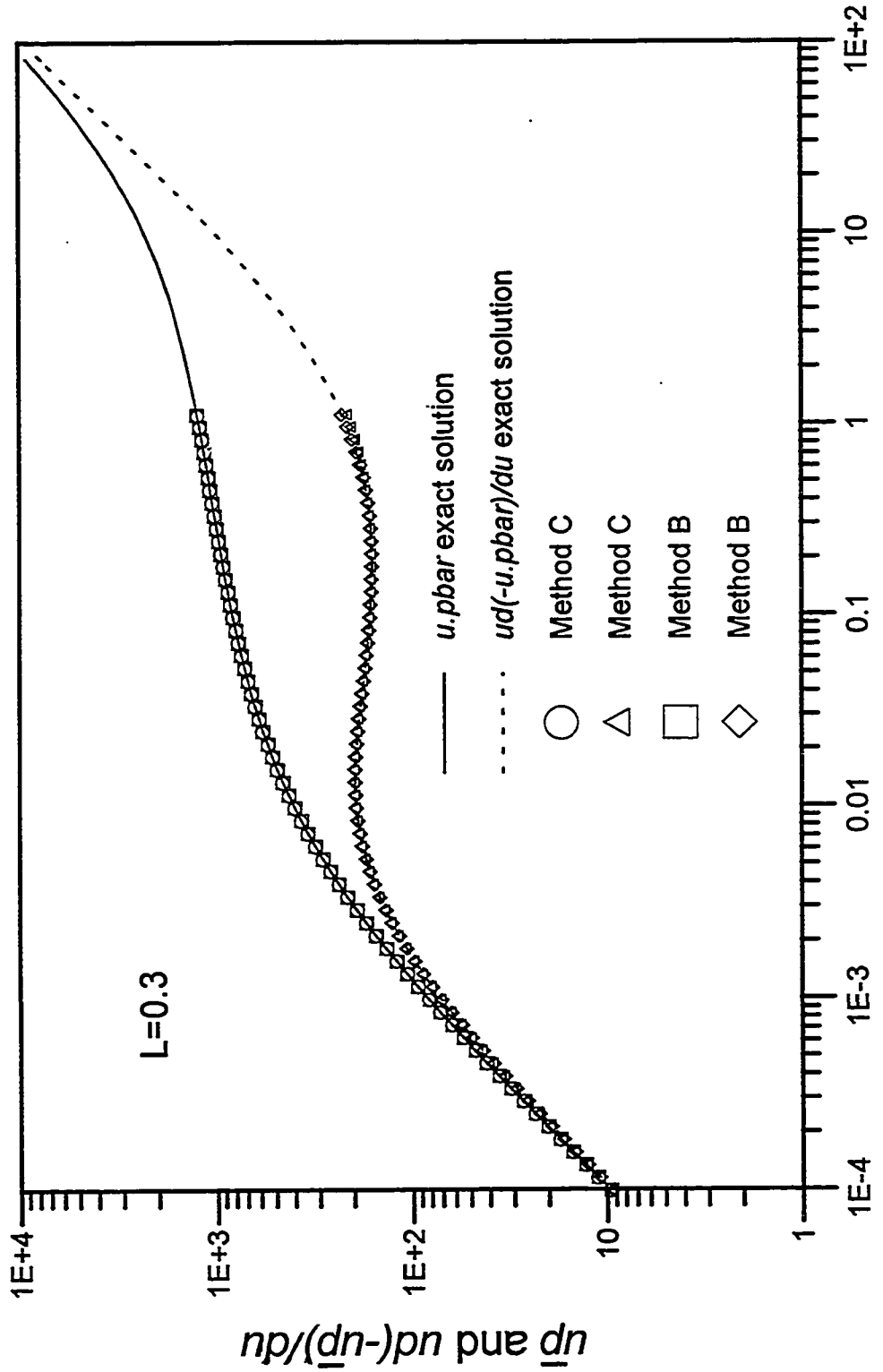

Fig. 6.1.14 - Numerical Laplace transform - Horizontal Well Example

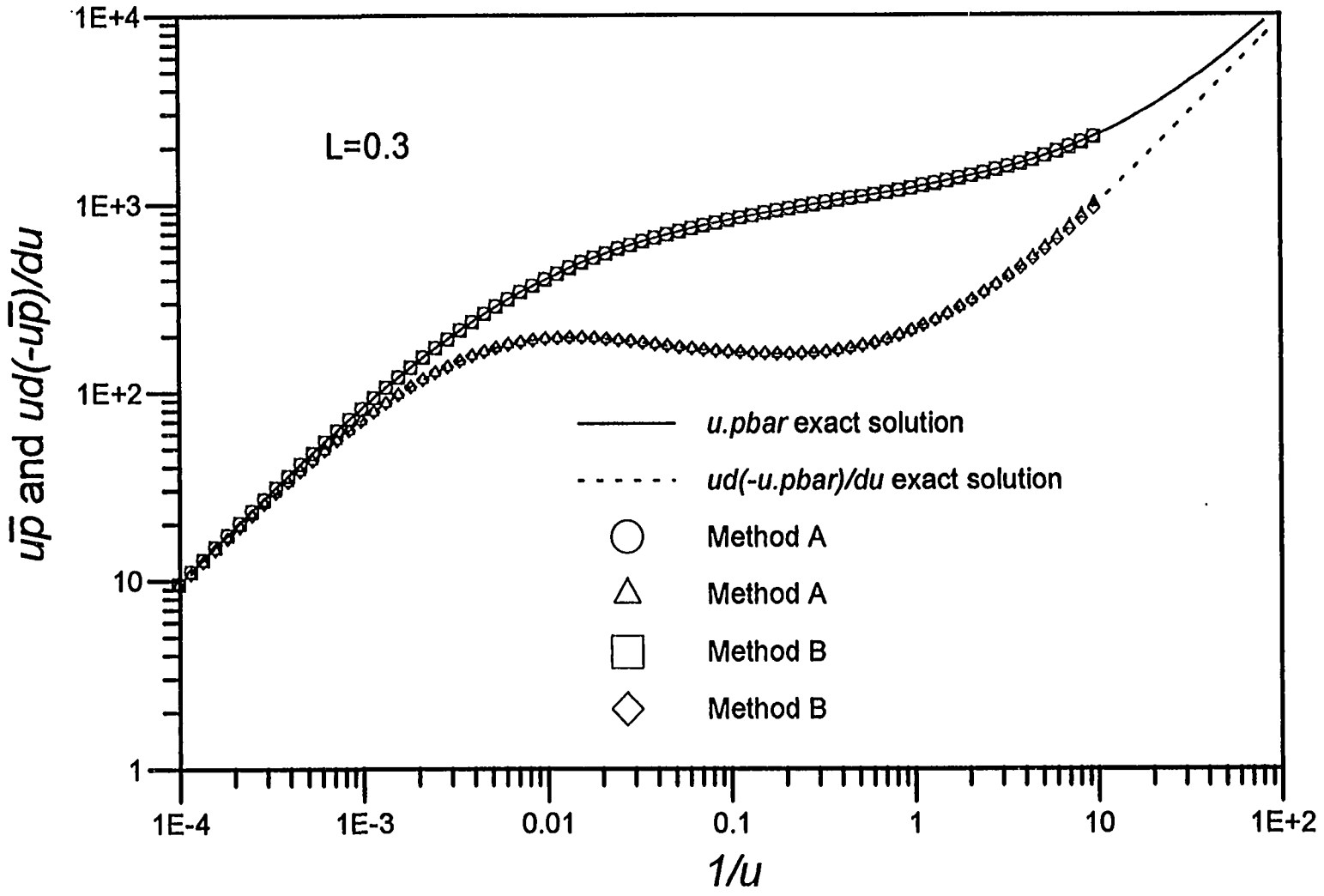

Fig. 6.1.15 - Numerical Laplace transform - Horizontal Well Example

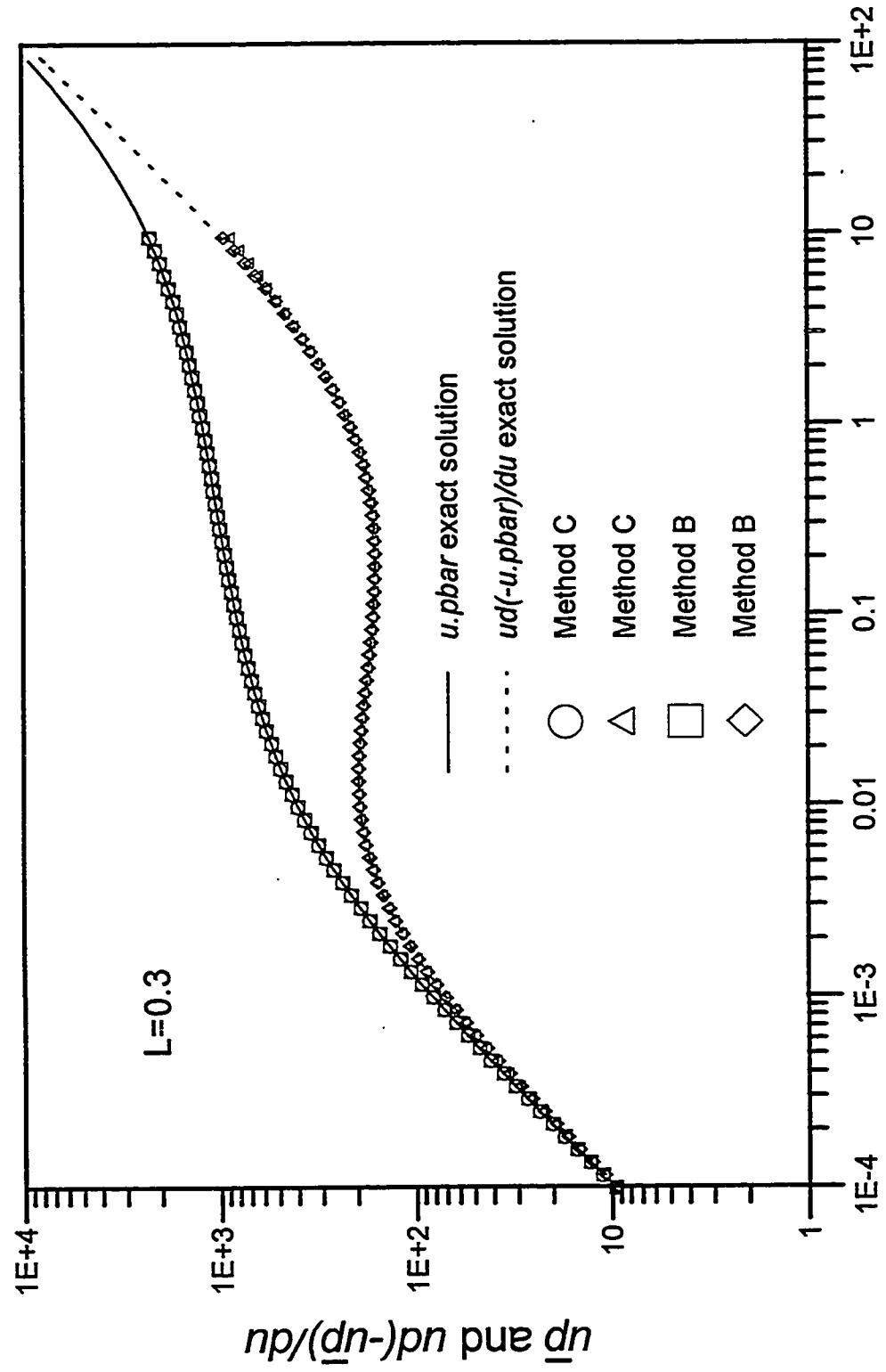

Fig. 6.1.16 - Numerical Laplace transform - Horizontal Well Example

6.1.2.3 Buildup Field Example

As our final example, we use the buildup data presented by Ref. 27. Recall that this is the same test that was analyzed in Chapter III, Section 3.6. The solid lines (as well as the dashed lines) in Figs. 6.1.17 and 6.1.18 were obtained analytically using the parameter estimates from the least squares regression (see Table 3.6.1).

The symbols in Figs. 6.1.17 and 6.1.18 were obtained by applying Methods A, B or C to the original pressure data. Note that the performances of all three methods are equivalent for the pressure curve. For the derivative curve, Method A deviates upwards a little bit at the end of the curve, whereas Methods B and C deviate a little bit downwards. In this case, Methods B and C will match the derivative perfectly if we use $L = 0.75$.

6.2 Parameter Estimation in Laplace Space

Many analytical solutions in well testing are available in closed analytical form only in Laplace space. Thus, it is common to compute real space pressure transient solutions by numerically inverting solutions from Laplace space. Another reason to compute the solutions in Laplace space is that dual porosity effects and wellbore storage can be conveniently incorporated once we know the unit-rate homogeneous solution in Laplace space. Ref. 13 showed that the plot of $u\overline{p_w}(u)$ versus $1/u$ in Laplace space displays the same characteristic signature as the plot of $p_w(t)$ versus t in real space. In fact, since $u\overline{p_w}(u)$ is defined with an integral, the plot of $u\overline{p_w}(u)$ versus $1/u$ looks similar to the integral type curves described by Onur and Reynolds.[57] Note that the variables $u\overline{p_w}(u)$ and $1/u$ have the same dimensions as $p_w(t)$ and t.

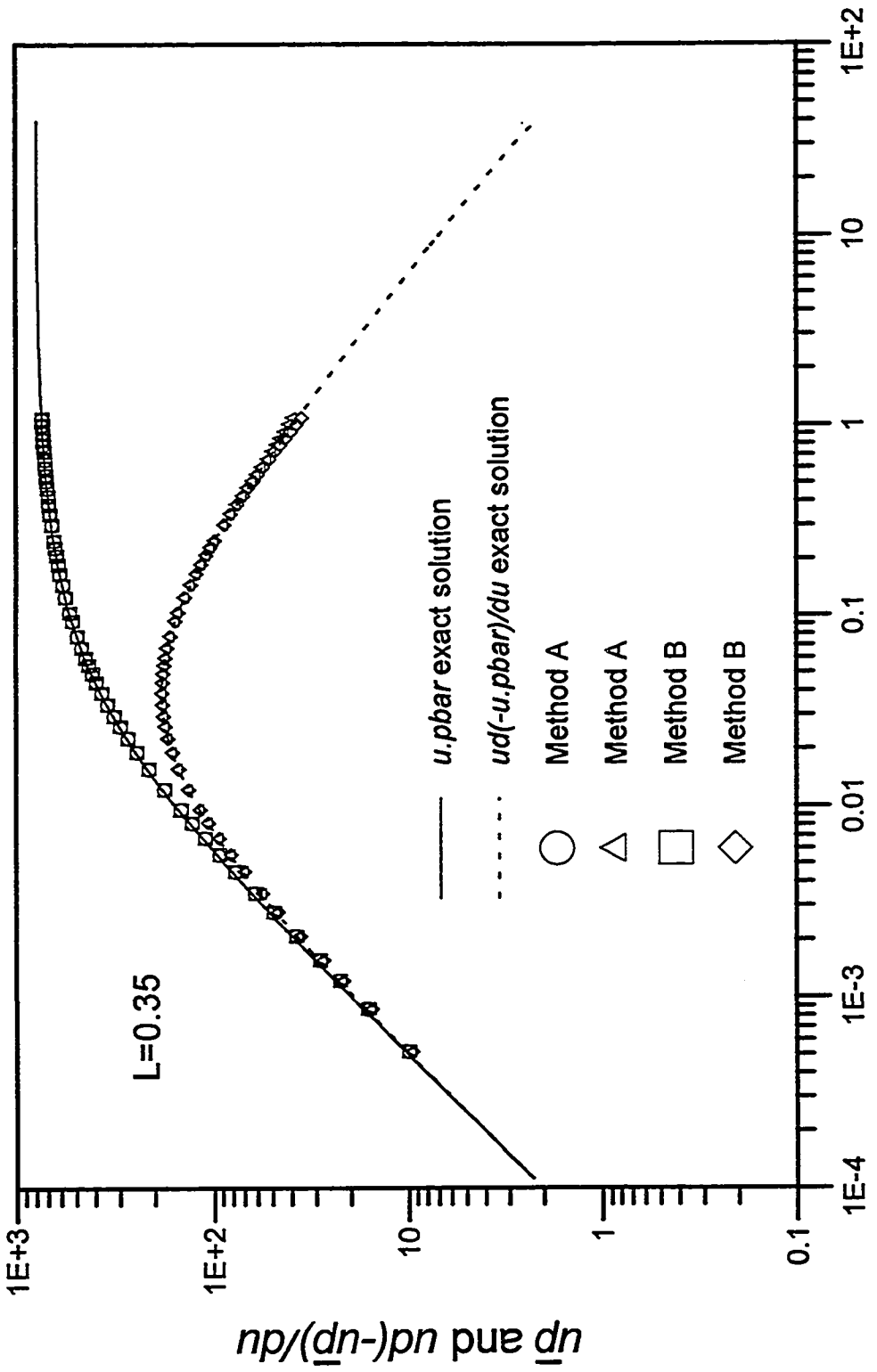

Fig. 6.1.17 - Numerical Laplace transform - Buildup Field Data

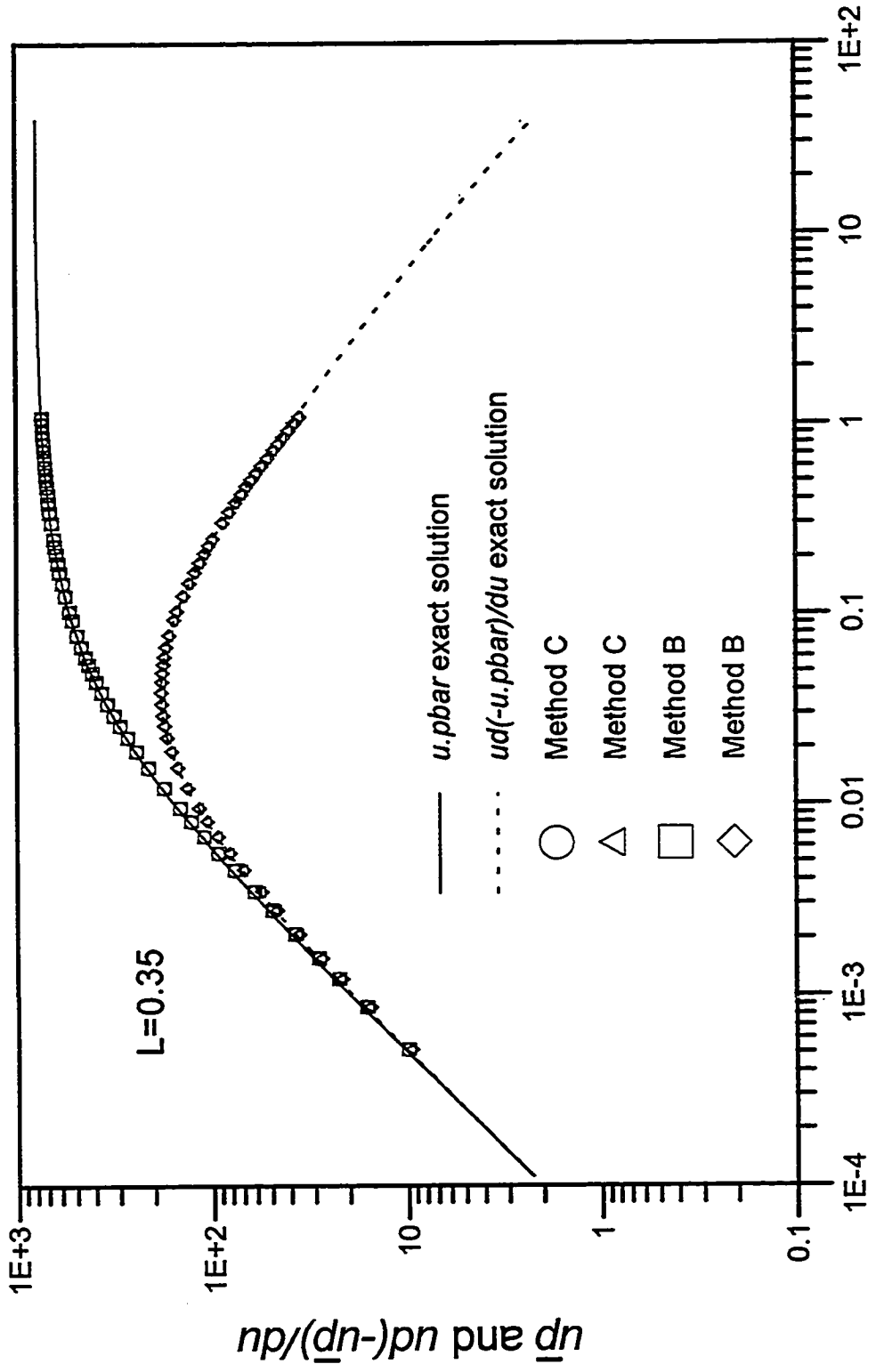

Fig. 6.1.18 - Numerical Laplace transform - Buildup Field Data

In this section, we will briefly discuss regression analysis in Laplace space. The parameter estimation can be performed in Laplace space by minimizing the appropriate objective function. For LS regression on pressure data, the objective function in Laplace space can be written as[56]

$$S(\alpha)=\sum_{i=1}^{n} u_i^2\left[\bar{y}(u_i)-\bar{F}(\alpha,u_i)\right]^2, \tag{6.2.1}$$

where $u_i = 1/t_i$ is the Laplace variable, $\bar{F}$ is the Laplace transform of the functional model F, and $\bar{y}$ is the corresponding Laplace transform of the measurements. In the case of variable flowrate, $\bar{F}$ can be computed via the convolution equation in Laplace space given by Eq. 6.1.6. In this case, $\bar{q}$ can be computed assuming that the rate function is piecewise constant with the last flowrate extrapolated to infinity.

In general, we have pressure measurements only in a given time interval, say (t_b,t_e). Thus, if we want to compute $\bar{y}$, we need to extrapolate the measurements to the time intervals $(0,t_b)$ and (t_e,∞). Since the idea is to match the data during the test interval only, we use the model function (evaluated at the current values of the parameters) for extrapolation purposes; i.e., outside the test interval, the unobserved pressure measurements are given by the current model function. Thus, $\bar{y}$ in Eq. 6.2.1 is given by

$$\bar{y}(\alpha,u_i)=\int_0^{t_b} e^{-u_i t}F(\alpha,t)dt+\int_{t_b}^{t_e} e^{-u_i t}y(t)dt+\int_{t_e}^{\infty} e^{-u_i t}F(\alpha,t)dt. \tag{6.2.2}$$

The second integral in Eq. 6.2.2 is computed just once using the piecewise linear approximation of the measurements. The last integral in Eq. 6.2.2, can be approximated by

$$\int_{t_e}^{\infty} e^{-u_i t}F(\alpha,t)dt \approx \int_{t_e}^{10t_e} e^{-u_i t}F(\alpha,t)dt, \tag{6.2.3}$$

i.e., for $t>10t_e$, the contribution to the integral is neglegible because $e^{-u_i t}$ is essentially zero. Following Ref. 9, we call this last integral as the "tail" contribution of $\bar{y}$. If the

model is known only in Laplace space, F can be computed by using a Laplace inversion algorithm and convolution in real space (in the case of variable rate). When t_b is large (say $t_b > 0.01\sqrt{t_e}$, t in days), there is no apparent advantage in regressing in Laplace space, since at each iteration, the computation of the first integral in Eq. 6.2.2 will require more function evaluations than the regression in real space. Clearly, there is no advantage in regressing in Laplace space if we know the analytical solution of the functional model in the real space.

Wilkinson[9] recently presented procedures for parameter estimation in the Laplace space when $t_b = 0$. He uses the same idea of extrapolating the data with the function model evaluated at the current value of the parameters. He also discusses a procedure to compute the tail contribution of $\bar{y}$ (given by the last integral in Eq. 6.2.2) approximately, without performing the Laplace inversion and the convolution at every iteration. Conversely, we propose to compute $\bar{y}$ using Method B or C of the previous section to estimate the Laplace transform of tabulated data. As we showed in the previous section, these two methods yield accurate Laplace transforms for both pressure and pressure derivative. In this approach, $\bar{y}$ is computed just once, which expedites the regression. A final regression in real space is recommended for computing the confidence intervals of the parameter estimates.

CHAPTER VII

CONCLUSIONS

This chapter summarizes the work we have presented in the previous chapters, and enumerates the contributions and conclusions of this study.

The objective of this work has been to study the application of nonlinear regression and statistical techniques in the analysis of pressure transient tests.

In Chapter II, we summarized the nonlinear regression and statistical techniques we found meaningful for well test applications, which includes: least squares (LS) estimation, robust least absolute value (LAV) estimation, confidence intervals of the parameters, confidence regions, incorporation of parameter constraints, verification of the linearity assumption, and residual-based methods for outlier detection. The conclusions of this part of the work are as follows.

(i) The algorithm we use for L_1-norm minimization has the ability to account for possible outliers in the data, and proved to be efficient for well test problems. In this approach, we minimize a least squares objective function which is equivalent to the least absolute value objective function. As a result, commonly available LS algorithms (such as the LMDER routine we use) are readily applicable for LAV estimation.

(ii) Since the LAV residuals are less affected by outliers, they embody powerful information for detecting all outliers in the data. We present a method that uses the LAV residuals for outlier detection. Once the outliers are identified and removed from the original data set, we can perform the LS parameter estimation on the remaining data. The estimates of this LS regression agree closely with the LAV estimates. This approach has the advantage that statistical analysis and confidence intervals of the parameter

estimates are well defined for the LS regression. The size of the confidence region (or the confidence intervals) can be significantly affected by outliers.

(iii) The classical confidence regions and confidence intervals of the parameter estimates are important tools for determining the uncertainty in the results, however, they rely upon the linear approximation of the model within the confidence region. An existing method for constructing likelihood ratio confidence intervals (which does not assume that the model is linear) was considered. In contrast to the classical confidence interval that requires only one minimization, the likelihood ratio confidence intervals require multiple regressions. We also presented a method for testing the linearity assumption.

(iv) We presented a new method for constraining the parameters within the feasible region. In contrast to the penalty function method, which may not avoid a non-physical parameter value in some circumstances, the proposed method guarantees that the estimates satisfy the physical constraints imposed. A nice feature of our imaging extension method is that it is readily applied in conjunction with existing unconstrained regression routines.

In Chapter III, we focused on the application of regression analysis to the classical wellbore storage and skin model. Motivated by the classical type curve matching techniques, we investigated the benefits of using grouped parameters (following the Gringarten *et al.* type curves) and the advantages of including pressure derivative information in the model. For this problem, the results presented in this chapter lead to the following conclusions.

(i) We have introduced a two-step regression procedure that uses pressure derivative information. This procedure works better for short tests than the common regression on pressure data. In the first step, we defined a new objective function in terms of the pressure-pressure derivative ratio (i.e., having $\Delta p/(2\Delta p')$ as the dependent variable) to obtain estimates of k/C and Ce^{2s} by regression on these two grouped

parameters. In the second step we fix k/C (from the first step) and regress on Ce^{2s} and k. We found that we can also regress on the primitive parameters $\{k, C, s\}$ in the first step. In this approach, different initial estimates gave very different solutions for the individual parameters k, C and s, however all solutions gave good estimates of the correct values of k/C and Ce^{2s}.

(ii) For the wellbore storage and skin problem studied, the use of grouped parameters offered no advantages over analysis based on "primitive" parameters. In fact, when regressing on group parameters, a significantly larger number of iterations are required to obtain convergence.

In Chapter IV, we applied nonlinear regression techniques to the horizontal well and dual porosity models. Regression techniques are especially helpful for these more complex models due to the large number of reservoir parameters, which makes classical type curve analysis difficult. The observations and conclusions for the dual porosity part are as follows.

(i) A general procedure for introducing anisotropic dual porosity effects in the isotropic homogeneous solution was presented.

(ii) We demonstrated through synthetic examples that for dual porosity models, we should first regress on pressure derivative data, and then use the results of this run as initial estimates for the final regression on pressure data.

(iii) When regression is applied to pressure derivative data, the data should not be smoothed (even for very noisy data), because smoothing can contaminate the pressure derivative data, and consequently the parameter estimates; indeed, the regression fit will provide the best possible smooth pressure derivative solution if the model is right. (This does not mean we should not smooth derivative data for model identification purposes.)

(iv) The LAV estimation (or the LS regression after deleting the outliers) gave better results for the dual porosity field example. Our experience is that the dual porosity problems may exhibit multiple local minima.

The following conclusions apply for horizontal well regression analysis.

(i) The application of nonlinear regression analysis on specific parts of the test, where one or more of the parameters are not important, enhances the chance of obtaining convergence to the correct estimates for the remaining parameters, or at least a combination of some parameters (e.g., $k_x k_y$, $k_x k_z$, or $L_x L_y$). This information can be used to devise better initial estimates for a final regression on all parameters using the entire span of the test data.

(ii) For horizontal-well tests, which usually exhibit a long wellbore storage dominated flow period, deconvolution procedures are especially important to reveal the theoretical behavior of the early time flow regimes, as well as the ending time of the wellbore storage effects.

(iii) The correlation coefficient between some of the parameters can be used to define a specific flow regime, for example, during the early radial flow we expect that k_x and k_y are highly correlated.

(iv) The hyper-surface of the pressure derivative objective function has more features than the pressure objective function, which tends to yield faster convergence in the regression procedure.

In Chapter V, we reviewed existing methods for verifying whether the basic assumptions of the error term are violated. This includes application of the chi-squared tests, and various plots of the scaled residuals.

A brief study of some convolution and deconvolution techniques is presented in Chapter VI. These techniques are necessary for generating multi-rate solutions, or for

removing the effects of the rate variation (especially afterflow) from the pressure response. The contributions and conclusions of this part of the work are as follows.

(i) We present a new algorithm to take the Laplace transform of tabulated real data, which assumes a linear extrapolation of the pressure derivative data. We demonstrate that the proposed method yields better results than the Romboustsos-Stewart algorithm, and slightly better results than a recent method presented by Bourgeois and Horne. Our method yields accurate Laplace transforms for both pressure and pressure derivative data.

(ii) The parameter estimation can be performed directly in Laplace space. One advantage of regressing in Laplace space is that instead of inverting the functional model from Laplace space at every iteration, we take the Laplace transform of the pressure data just once. The proposed method can be used to obtain the Laplace transform of pressure data if we have measurements in the interval $(0,t)$.

(iii) In general, we may have pressure measurements only in a given time interval (t_b,t_e). For regression purposes, we can extrapolate the measurements to the time intervals $(0,t_b)$ and (t_e,∞) at each iteration by using the model function evaluated at the current value of the parameters. It appears that there is no apparent advantage in regressing in Laplace space when t_b is large.

(iv) A final regression in real space (using the results from the regression in Laplace space as initial estimates) is recommended for computing the confidence intervals of the parameters.

NOMENCLATURE

a, b, c	constants
B	formation volume factor, *RB/STB*
$\mathbf{b}, \{b_i\}$	estimate of β, vector of the parameters in the linear model
$\mathbf{b}^0, \{b_i^0\}$	estimate of β^0, vector of the parameters in the linearization method
C	wellbore storage coefficient, *bbl/psi*
C_b	wellbore storage below the measurement point, *bbl/psi*
Cov	covariance
c_t	total compressibility, psi^{-1}
c_{tf}	total compressibility of the fracture system, psi^{-1}
c_{tm}	total compressibility of the matrix system, psi^{-1}
$\mathbf{d}$	direction vector
E^+	objective function of the modified unconstrained problem
E_1	LAV (or L_1-norm) objective function (Eq. 2.2.1)
$\hat{E}_1$	modified objective function for LAV parameter estimation
E_2	LS (or L_2-norm) objective function (Eq. 2.1.2)
$erf(x)$	error function
F	functional model; F-distribution
$f(u)$	dual-porosity characteristic function in Laplace space
$g(u)$	dual-porosity characteristic function in Laplace space
$\mathbf{H}, H$	Hessian matrix (Eq. A.8)
h	formation thickness, *ft*
$\mathbf{I}, I$	identity matrix

$I(\Delta p)$	pressure integral, $\int_0^t \Delta p(\tau)d\tau$, *psi-day*
k	permeability, *md*; or iteration counter
$\hat{k}_f$	bulk permeability of the fracture system, *md*
$\hat{k}_m$	bulk permeability of the matrix system, *md*
k_h	effective horizontal permeability, *md*
k_x	permeability in the *x*-direction, *md*
k_y	permeability in the *y*-direction, *md*
k_z	permeability in the *z*-direction, *md*
L	log-likelihood function; reference length
$\mathcal{L}$	Laplace transform operator
$\mathcal{L}^{-1}$	inverse Laplace transform operator
LAV	least absolute value
LS	least squares, L_2-norm
L_w	horizontal well length, *ft*
L_x	reservoir length in the *x*-direction, *ft*
L_y	reservoir length in the *y*-direction, *ft*
L_z	reservoir thickness, *ft*
n	number of data points
p	number of regression parameters; pressure, *psi*
p_o	initial wellbore pressure, *psi*
p_i	initial reservoir pressure, *psi*
p_{wf}	wellbore flowing pressure, *psi*
p_{ws}	wellbore shut-in pressure, *psi*
q	sandface flow rate, *STB/day*
$\mathbf{R}$, $\{R_{i,j}\}$	hat matrix defined in Eq. 2.4.2
r	radial distance, ft

RLS	re-weighted least squares
r_u	uth residual (Eqs. 2.1.1 and 5.1.1)
r_w	wellbore radius, *ft*
s	skin factor, *dimensionless*; or estimate of the standard deviation of the error term
$S(\alpha)$	error sum of squares at α
s^*	robust estimate of the standard deviation of the error term
s^2	estimate of the variance of the error term
SAR	sum of the absolute value of the residuals
SSR	sum of the residual squares
t	student distribution; time, *days*
t_e	ending time of the pressure test, *days*; or time when the logarithimic pressure extrapolation becomes zero, *days*
u	Laplace variable
V, Var	variance
V_B	total bulk volume, ft^3
V_f	ratio of the bulk volume of fractures to the total volume
V_m	ratio of the bulk volume of matrix to the total volume
$\mathbf{W}$	variance matrix of the error term
w	weight
$\mathbf{X}$, $\{X_{iu}\}$	$n \times p$ matrix of the independent variables
x'	x-coordinate of the well axis, ft
$\mathbf{Y}$	dependent variable vector
y, Y	dependent variable, measurement
$\mathbf{Y}^0$, $\{Y_u^0\}$	vector of the dependent variables in the linearization method (Eq. 2.3.4)
y_1	y-coordinate of the closest well tip, ft

$\mathbf{Z}$, $\{Z_{iu}\}$	$n \times p$ matrix of the partial derivatives of the model function (Eq. 2.3.5c)
z'	z-coordinate of the well axis, ft

Greek Symbols

∇_α	gradient with respect to α
α	characteristic parameter of the fractured rock, $1/ft^2$
α, $\{\alpha_i\}$	vector of unknown parameters in the nonlinear model
$\hat{\alpha}$	estimate of the unknown vector of the parameters
α_0, $\{\alpha_{i0}\}$	vector of the initial estimates of the parameters
α_{eq}	vector of the parameters of the unconstrained problem
α^k, $\{\alpha_i^k\}$	vector of the parameter estimates at the kth iteration
$\alpha_{\max,i}$	upper limit of the feasible interval for parameter α_i
$\alpha_{\min,i}$	lower limit of the feasible interval for parameter α_i
β, $\{\beta_i\}$	vector of the unknown parameters in the linear model
β^0, $\{\beta_i^0\}$	parameter vector of the linearization method (Eq. 2.3.5b)
χ^2	chi-squares distribution; chi-squares statistic
$\Delta\alpha_i$	step change in the parameter α_i
Δp	pressure change, $p_i - p$, psi
Δp_{fu}	unit-rate pressure response of the dual-porosity model, psi
Δp_s	steady state skin pressure drop, psi
Δp_{su}	unit-rate pressure response including skin, psi
Δp_{wf}	drawdown pressure change, $p_i - p_{wf}$, psi
Δp_{ws}	buildup pressure change, $p_i - p_{ws}$, psi
$\Delta p'$	logarithmic pressure derivative, $\partial \Delta p / \partial \ln(t)$, psi

$\Delta p_i'$	cord slope of the *i*th interval (Eq. F.2)
Δt_{bu}	shut-in time period, *hrs*
ε	small number in the L_1-norm algorithm
ε, $\{\varepsilon_u\}$	vector of the error term
ϕ, phi	porosity, *fraction*
$\hat{\phi}_f$	bulk porosity of the fracture system, $\phi_f V_f$, *fraction*
$\hat{\phi}_m$	bulk porosity of the matrix system, $\phi_m V_m$, *fraction*
$\hat{\phi}_{f+m}$	total porosity of the dual porosity system, $\hat{\phi}_f + \hat{\phi}_m$, *fraction*
γ	confidence level
η	hydraulic diffusivity, $k/\phi\mu c_t$, ft^2/day
λ	inter-porosity flow coefficient, *dimensionless*; or relaxation parameter, *dimensionless*
μ	fluid viscosity, *cp*; or expected number of runs
ω	ratio of the fracture storativity to the total system storativity
π	constant ≈ 3.14159
ρ	correlation coefficient (Eq. B.13); or fluid density, lbm/ft^3
σ	standard deviation of the error term

Subscripts

0	initial estimate
u	data point number; or unit-rate
meas	measured
i	parameter number; or initial
o	oil
w	wellbore
b	at the measuring point (bottom-hole); or beginning

m	matrix
f	fracture
$f+m$	total system (fracture system plus matrix system)
fx	x-direction in the fracture system
fy	y-direction in the fracture system
x	x-direction
y	y-direction
z	z-direction

Superscripts

*	minimizer; or modified variable
0	initial estimate
T	transpose
^	at the solution of the minimization procedure; bulk property
c	current estimate
−	Laplace space; average; effective

REFERENCES

1. Abbaszadeh, M. and Kamal, M.M.: "Automatic Type-Curve Matching for Well Test Analysis," *SPEFE* (Sept. 1988) 567–577.

2. Rosa, A.J. and Horne, R.N.: "Automated Type-Curve Matching in Well Test Analysis Using Laplace Space Determination of Parameter Gradients", paper SPE 12131 presented at the 58th SPE Annual Technical Conference and Exhibition, San Francisco, CA (October 5–8, 1983).

3. Barua, J., Horne, R.N., Greenstadt, J.L. and Lopez, L.: "Improved Estimation Algorithms for Automated Type Curve Analysis of Well Tests," *SPEFE* (March, 1988) 186–196.

4. Nanba, T. And Horne, R.N.: "An Improved Regression Algorithm for Automated Well Test Analysis," paper SPE 18161 presented at the 63rd SPE Annual Technical Conference and Exhibition, Houston, TX (October 2-5, 1988).

5. Rousseeuw, P.J. and Leroy, A.M.: *Robust Regression and Outlier Detection*, John Wiley & Sons, New York (1987).

6. Huber, P.J.: *Robust Statistics*, John Wiley & Sons, New York (1981).

7. Stehfest, H.: "Numerical Inversion of Laplace Transforms," Algorithm 368, *Communications of ACM* (Jan., 1970) **13**, No. 1, 47-49.

8. Rosa, A.J. and Horne, R.N.: "Automated Well Test Analysis Using Robust (LAV) Nonlinear Parameter Estimation," paper SPE 22679 presented at the 66th SPE Annual Technical Conference and Exhibition, Dallas, TX (October 6-9, 1991).

9. Wilkinson, D.J.: "Pressure Transient Parameter Estimation in the Laplace Domain," paper SPE 24720 presented at the 1992 SPE Annual Technical Conference and Exhibition, Washington, DC, Oct. 4-7.

10. Horne, R.N.: "Advances in Computer-Aided Well Test Interpretation," paper SPE 24731 presented at the 1992 SPE Annual Technical Conference and Exhibition, Washington D.C., Oct. 4-7.

11. Garbow, B.S., Hillstrom, K.E. and More, J.J.: Subroutine LMDER, Argonne National Laboratory, MINPACK project, March, 1980.

12. Romboutsos, A. and Stewart, G.: "A Direct Deconvolution or Convolution Algorithm for Well Test Analysis," paper SPE 18157 presented at the 1988 SPE Annual Technical Conference and Exhibition, Houston, TX, Oct. 2-5.

13. Bourgeois, M.J. and Horne, R.N.: "Well Test Model Recognition Using Laplace Space," *SPEFE* (March 1993).

14. Marquardt, D.W.: "Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Indust. Appl. Math. (June 1963) 11, No. 2.

15. Wittink, Dick. R.: *The Application Of Regression Analysis*, Allyn & Bacon Corporation, Boston.

16. Draper, N.R. and Smith, H.: *Applied Regression Analysis*, Second Edition, John Wiley & Sons, New York (1981).

17. Davies, O.L. and Goldsmith, P.L. (ed.): *Statistical Methods in Research and Production*, Fourth Edition, Longman Inc., New York (1984).

18. Hoel, Paul G.: *Introduction To Mathematical Statistics*, John Wiley & Sons, New York (1961).

19. Faddeev, D.K. and Faddeeva, V.N.: *Computational Methods of Linear Algebra*, W. H. Freeman and Company, San Francisco (1963).

20. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P.: *Numerical Recipes in FORTRAN - The Art of Scientific Computing*, Second Edition, Cambridge University Press, New York (1992).

21. Bard, Yonathan: *Nonlinear Parameter Estimation*, Academic Press, San Diego (1974).

22. Carroll, R.J. and Ruppert, D.: *Transformation and Weighting in Regression*, Chapman and Hall, New York (1988).

23. El-Attar, R.A., Vidyasagar, M. and Dutta, S.R.K.: "An Algorithm for L_1-Norm Minimization with Application to Nonlinear L_1 Approximation," *SIAM J. Num Anal.* (Feb. 1979), **16**, No. 1, 70-86.

24. Dogru, A.H., Dixon, T.N. and Edgar, T.F.: "Confidence Limits on Parameters and Prediction of Slightly Compressible, Single Phase Reservoirs," *SPEJ* (Feb. 1977) 42-56; *Trans*, AIME, **263**.

25. Cook, R.D. and Weisberg, S.: *Residuals and Influence in Regression*, Chapman and Hall, New York (1982).

26. Barnett, V. and Lewis, T.: *Ouliers in Statistical Data*, Second Edition, John Wiley & Sons, New York (1984).

27. Bourdet, D., Whittle, T.M., Douglas, A.A., and Pirard, Y.M.: "A New Set of Type Curves Simplifies Well Test Analysis," *World Oil* (May 1983) 95-106.

28. Onur, M. and Reynolds, A.C.: "A New Approach for Constructing Type Curves for Well Test Analysis," *SPEFE* (March 1988) 197-206.

29. Gringarten, A.C., Bourdet, D., Landel, P.A. and Kniazeff, V.J.: "A Comparison Between Different Skin and Wellbore Storage Type-Curves for Early-Time Transient Analysis," paper SPE 8205 presented at the 1979 SPE Annual Technical Conference and Exhibition. Las Vegas, Sept. 23-26.

30. Thompson, L.G. and Temeng, K.O.: "Automatic Type-Curve Matching for Horizontal Wells," SPE 25507, presented at the 1993 SPE Production Operations Symposium, Oklahoma City, OK, March 21-23.

31. Earlougher, R.C.: *Advances in Well Test Analysis*, Monograph Series, SPE of AIME, Dallas (1977).

32. Fair, B.W., Jr.: "Pressure Buildup Analysis With Wellbore Phase Redistribution," *SPEJ* (Apr. 1981) 259-270.

33. Hegeman, P.S., Hallford, D.L. and Joseph, J.A.: "Well Test Analysis With Changing Wellbore Storage," paper SPE 21892 presented at the 1991 Rocky Mountain Regional Meeting and Low-Permeability Reservoirs Symposium, Denver, CO (April 15-17, 1991).

34. Agarwal, R.G., Al-Hussainy R. and Ramey, H.J., Jr.: "An Investigation of Wellbore Storage and Skin Effect in Unsteady Liquid Flow: I. Analytical Treatment," *Trans.* AIME, **249** (1970) 279-290.

35. Earlougher, R.C., Jr. and Kersch, K.M.: "Analysis of Short-Time Transient Test Data by Type-Curve Matching." *JPT* (July 1974) 793-97; *Trans*, AIME, **257**.

36. Peres, A., Onur, M. and Reynolds, A.C.: "A New General Pressure Analysis Procedure for Slug Tests," paper SPE 18801 presented at the SPE 1989 California Regional Meeting. Bakersfield, CA, April 5-7.

37. Correa, A.C. and Ramey, H.J., Jr.: "A Method for Pressure Buildup Analysis of Drillstem Tests," paper SPE 16802 presented at the 1987 SPE Annual Technical Conference and Exhibition, Dallas, Sept. 27-30.

38. Odeh, A.S. and Babu, D.K.: "Transient Flow Behavior of Horizontal Wells, Pressure Drawdown, and Buildup Analysis," SPE 18802, presented at the 1989 SPE California Regional Meeting, Bakersfield, CA, April 5-7, 1989.

39. Donoso, J.M. and Daniel, G.A.: "Application on Horizontal-Test Methods," SPE 21099, presented at the SPE Latin American Petroleum Engineering Conference, Rio de Janeiro, Brazil, Oct. 14-19, 1990.

40. Babu, D.K. and Odeh, A.S.: "Productivity of a Horizontal Well," SPE 18298, presented at the 1988 SPE Annual Technical Conference and Exhibition, Houston, TX, Oct. 2-5.

41. Rosa, A.J. and Carvalho, R.S.: "A Mathematical Model for Pressure Evaluation in an Infinite Conductivity Horizontal Well," *SPEFE* (Dec. 1989) 559-566.

42. van Everdingen, A.F. and Hurst, W.: "The application of the Laplace Transformation to Flow Problems in Reservoirs," *Trans.* AIME (1949) **186**, 305-324.

43. Kuchuk, F.J. and Goode, P.A.: "Pressure Transient Analysis and Inflow Performance for Horizontal Wells," paper SPE 18300 presented at the 1988 SPE Annual Technical Conference and Exhibition, Houston, TX, Oct. 2-5.

44. Goode, P.A. and Thambynayagam, R.K.M.: "Pressure Drawdown and Build-up Analysis for Horizontal Wells in Anisotropic Media," *SPEFE* (Dec. 1987) 683-97.

45. Barenblatt, G.I. and Zheltov, Yu.P.: "Fundamental Equation of Homogeneous Liquids in Fissured Rocks," *Dokl. Akad. Nauk SSR* (June 1960) **132**, No. 3, 545-48.

46. Barenblatt, G.I., Zheltov, Yu.P. and Kochina, I.N.: "Basic Concepts in the Theory of Seepage of Homogenous Liquids in Fissured Rocks (Strata)," *J. Appl. Math. Mech.* (USSR) (1960) **24**, No. 5, 1286-1303.

47. Warren, J.E. and Root, P.J.: "The Behavior of Naturally Fractured Reservoirs," *SPEJ* (Sept. 1963) 245-55; *Trans.*, AIME, **228**.

48. Houzé, O.P., Horne, R.N. and Ramey, H.J., Jr.: "Pressure-Transient Response of Infinite-Conductivity Vertical Fracture in a Reservoir with Double-Porosity Behavior," *SPEFE* (Sept. 1988) 510–18.

49. Carvalho, R.S. and Rosa, A.J.: "Transient Pressure Behavior for Horizontal Wells in Naturally Fracture Reservoir," paper SPE 18302 presented at the 1988 SPE Annual Technical Conference and Exhibition, Houston, TX, Oct. 2-5.

50. Gringarten, A.C. and Ramey, H.J., Jr.: "The Use of Source and Green's Functions in Solving Unsteady-Flow Problems in Reservoirs," *SPEJ* (Oct. 1973) 285-296.

51. Bourdet, D., Ayoub, J.A., Whittle, T.M., Pirard, Y.M. and Kniazeff, V.: "Interpreting Well Tests in Fractured Reservoirs," *World Oil* (Oct. 1983).

52. Kuchuk, F.J.: "Applications of Convolution and Deconvolution to Transient Well Tests," *SPEFE* (Sept. 1992).

53. Fair, P.S. and Simmons, J.F.: "Novel Well Testing Applications of Laplace Transform Deconvolution," paper SPE 24716 presented at the 1992 SPE Annual Technical Conference and Exhibition, Washington, DC, Oct. 4-7.

54. Kuchuk, F.J., Carter, R.G. and Ayestaran, L.: "Numerical Deconvolution of Wellbore Pressure and Flow Rate," SPE 13960, presented at the SPE Annual Technical Conference and Exhibition, Las Vegas, NV, 1985.

55. Thompson, L.G. and Reynolds, A.C.: "Analysis of Variable-Rate Well-Test Pressure Data Using Duhamel's Principle," *SPEFE* (Oct. 1986) 453-69.

56. Bourgeois, M.J. and Horne, R.N.: "Well Test Model Recognition Using Laplace Space ," paper SPE 22682 presented at the 1991 SPE Annual Technical Conference and Exhibition, Dallas, TX, Oct. 6-9.

57. Onur, M., Peres, A.M.M., and Reynolds, A.C.: "New Pressure Functions for Well Test Analysis," paper SPE 19819 presented at the 1989 SPE Annual Technical Conference and Exhibition, San Antonio, TX, Oct. 8-11.

58. Carvalho, R.S., Redner, R.A., Thompson, L.G. and Reynolds, A.C.: "Robust Procedures for Parameter Estimation by Automated Type-Curve Matching," paper SPE 24732 presented at the 1992 SPE Annual Technical Conference and Exhibition, Washington, DC, Oct. 4-7.

APPENDIX A

DESCRIPTION OF THE ALGORITHM USED IN THE LMDER PROGRAM

The LMDER[11] computer program utilizes a modified Newton's method for the solution of the nonlinear problem with a restricted step instead of a line search. This appendix will describe the basic algorithm on which this program is based.

The algorithm in the LMDER program is closely related to Newton's method for the minimization of a function of several variables. Newton's method for the minimization of a nonlinear function $E(\alpha)$ uses a quadratic approximation to the function. If we let $H(\alpha)$ denote the matrix of second order partial derivatives of E with respect to α, where α is a column vector containing the model parameters, then the quadratic approximation of E about a fixed point α_0 is

$$\begin{aligned} E(\alpha) \approx E(\alpha_0) + \nabla_\alpha E(\alpha_0)^t (\alpha - \alpha_0) \\ + \frac{1}{2}(\alpha - \alpha_0)^t H(\alpha_0)(\alpha - \alpha_0), \end{aligned} \tag{A.1}$$

where $\nabla_\alpha E(\alpha_0)$ denotes the gradient of E at α_0.

The function on the right side of Eq. A.1 is quadratic in α and if $H(\alpha_0)$ is positive definite then this function has a unique minimum satisfying

$$H(\alpha_0)(\alpha - \alpha_0) = -\nabla_\alpha E(\alpha_0), \tag{A.2}$$

i.e., the unique minimum of the expression on the right hand side of Eq. A.1 is given by

$$\alpha = \alpha_0 - H^{-1}(\alpha_0)\nabla_\alpha E(\alpha_0). \tag{A.3}$$

For the nonlinear optimization of a nonquadratic function, we define an iterative procedure as follows. Given a current estimate of the minimizer of E, say α^c, we set

$$\mathbf{d}^c = -H^{-1}(\alpha^c)\nabla_\alpha E(\alpha^c), \tag{A.4}$$

and try to choose a value of $\lambda > 0$ so that $E(\alpha^c + \lambda \mathbf{d}^c)$ is minimized, or at least is significantly smaller than $E(\alpha^c)$. When $H(\alpha^c)$ is positive definite, then $\mathbf{d}^c$ is a descent direction meaning that the directional derivative of E in the direction $\mathbf{d}^c$ is negative so that $E(\alpha^c + \lambda \mathbf{d}^c)$ is a decreasing function when λ is sufficiently small. Newton's method is an iterative procedure which generates a sequence of estimates of the minimizer. The algorithm stops when the program determines that no further progress is being made or when the program determines that a good estimate of the solution has been found.

We now specialize our discussion to the least squares nonlinear minimization problem. Let

$$E(\alpha) = \sum_{u=1}^{n} [y_u - F(\alpha, t_u)]^2, \tag{A.5}$$

where y_u is the data measured at time t_u. Then

$$\nabla_\alpha E(\alpha) = -2 \sum_{u=1}^{n} [y_u - F(\alpha, t_u)] \nabla_\alpha F(\alpha, t_u), \text{ and} \tag{A.6}$$

$$\begin{aligned} (H(\alpha))_{jk} = -2 \sum_{u=1}^{n} [y_u - F(\alpha, t_u)] \frac{\partial^2 F(\alpha, t_u)}{\partial \alpha_j \partial \alpha_k} + \\ 2 \sum_{u=1}^{n} \frac{\partial F(\alpha, t_u)}{\partial \alpha_j} \frac{\partial F(\alpha, t_u)}{\partial \alpha_k}. \end{aligned} \tag{A.7}$$

In the Gauss modification to Newton's Method we observe that $y_u - F(\alpha, t_u)$ is often small near the minimum of E in which case the Hessian matrix can be approximated by

$$(H(\alpha))_{jk} = 2 \sum_{u=1}^{n} \frac{\partial F(\alpha, t_u)}{\partial \alpha_j} \frac{\partial F(\alpha, t_u)}{\partial \alpha_k}. \tag{A.8}$$

This matrix is positive semi-definite and is positive definite if and only if

$$Span\{\nabla_\alpha F(\alpha,t_1),\ldots,\nabla_\alpha F(\alpha,t_n)\} = R^p. \tag{A.9}$$

Note that in most cases Eq. A.9 holds since, for nonlinear regression problems, n is always much bigger than p which makes the Hessian matrix positive definite. However, it is also possible that these gradient vectors almost lie in a subspace of dimension k, $k < p$, in which case H would be nearly singular (i.e., has a bad condition number or an eigenvalue close to zero). If the gradient vectors involved all lie in a subspace of dimension less than n, $H(\alpha)$ will be singular.

When the Hessian or the estimate of the Hessian H is not positive definite, then the vector $\mathbf{d} = -H^{-1}\nabla E$ may not be a descent direction. If the matrix H has a poor condition number then $\mathbf{d}$ may contain a lot of error and so again may not be a descent direction. In order to correct this, the Levenberg-Marquardt[14] idea is to replace H with the matrix $H + D$ where D is a diagonal matrix and is chosen so that $H + D$ is not only positive definite but also has smallest eigenvalue not too close to zero. D is usually written as εI so that we have $H + \varepsilon I$ and ε is chosen to be greater that or equal to zero and chosen large enough so that $H + \varepsilon I$ is a well-conditioned positive definite matrix. In this formulation, the descent direction is set to $\mathbf{d}^c = -\left(H(\alpha^c) + \varepsilon I\right)^{-1}\nabla_\alpha E(\alpha^c)$.

Now that a descent direction $\mathbf{d}^c$ is defined, a value of $\lambda > 0$ must be chosen so that $E(\alpha^c + \lambda \mathbf{d}^c)$ is significantly reduced. One alternative is to perform a one dimensional search in order to find an acceptable value of λ. A second alternative is the restricted step method.

In the restricted step method, we begin with an initial step size λ. At each step, we have a current value of λ, say λ^c and we define $\alpha^+ = \alpha^c + \lambda^c \mathbf{d}^c$ and compute the value of $E(\alpha^+)$ and the predicted value of the function according to our quadratic model

$$\begin{aligned} Q(\alpha^+) = E(\alpha^c) + \nabla_\alpha E(\alpha^c)^t(\alpha^+ - \alpha^c) + \\ \frac{1}{2}(\alpha^+ - \alpha^c)^t H(\alpha^c)(\alpha^+ - \alpha^c). \end{aligned} \tag{A.10}$$

The ratio

$$r^{+}=\frac{E(\alpha^{c})-E(\alpha^{+})}{E(\alpha^{c})-Q(\alpha^{+})} \tag{A.11}$$

is then computed and one of the following actions is adopted:

(1) If $r^{+}<0.25$ (which includes the case that the $E(\alpha^{+})>E(\alpha^{c})$) then the reduction is considered small and the step length is reduced.

(2) If $0.25\leq r^{+}\leq 0.75$ then the reduction is considered to be acceptable and the step size remains the same, the iteration is accepted and α is updated.

(3) If $r^{+}>0.75$ then the prediction is excellent and the step size is increased, the iteration is accepted and α is updated.

Subroutine **LMDER** is used in conjunction with the constraining algorithms discussed in Section 2.3.6.

APPENDIX B

LINEAR MODEL THEORY

Since confidence limits are based on the assumption that the model behaves linearly in the vicinity of the solution, we consider here the linear model case.

The linear functional model between the independent variables, X_i, and the dependent variable Y is given by

$$Y_u = \beta_1 X_{1u} + \beta_2 X_{2u} + \ldots + \beta_p X_{pu} + \varepsilon_u, \tag{B.1}$$

where β_i, $i = 1, \ldots, p$, are the parameters of the model. X_{iu} and Y_u ($u = 1, \ldots, n$), are a set of n observations, where the X_{iu} values are assumed to be exact and each Y_u is measured with a corresponding error ε_u. The model is said to be linear because Y_u is a linear function of each β_i, for example,

$$Y_u = \beta_0 + \beta_1 X_u + \beta_2 X_u^2 + \varepsilon_u \tag{B.2}$$

is a linear model even though it involves X_u^2. We can write Eq. B.1 in matrix form as follows:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{B.3}$$

where $\mathbf{X}$ is an $n \times p$ matrix that is defined as the *matrix of the independent variables* or *sensitivity matrix*.

Here, β is a p-dimensional column vector defined by $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^T$, where throughout the "superscript" T refers to the operation of taking the transpose of a vector or matrix. The error vector, ε, and the dependent variable vector, $\mathbf{Y}$, are n-dimensional column vectors defined respectively by $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$ and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T$. The variable X_{iu} represents the element of $\mathbf{X}$ in the uth row and ith column, *not* the element in the ith row and uth column.

Although we can not determine the parameters, β_i's, exactly because of the randomness of the errors, ε_i's, we can find their estimates by using the observed data and the following model:

$$\hat{\mathbf{Y}} = \mathbf{Xb}, \tag{B.4}$$

where **b** is an estimate of the parameter vector, β, and $\hat{\mathbf{Y}}$ is the so-called *predicted* or *criterion variable*. For least-squares (LS) regression, **b** is the value of β which minimizes the *error sum of squares*, $S(\beta)$, which is defined as

$$\begin{aligned} S(\beta) &= \sum_{u=1}^{n}\left\{Y_u - \sum_{i=1}^{p}\beta_i X_{iu}\right\}^2 \\ &= (\mathbf{Y}-\mathbf{X}\beta)^T(\mathbf{Y}-\mathbf{X}\beta) \\ &= \mathbf{Y}^T\mathbf{Y} - \mathbf{Y}^T\mathbf{X}\beta - \beta^T\mathbf{X}^T\mathbf{Y} + \beta^T\mathbf{X}^T\mathbf{X}\beta \\ &= \mathbf{Y}^T\mathbf{Y} - 2\beta^T\mathbf{X}^T\mathbf{Y} + \beta^T\mathbf{X}^T\mathbf{X}\beta. \end{aligned} \tag{B.5}$$

Note that $S(\beta)$ represents the error sum of squares at the true (but unknown) values of the parameters.

In order to perform statistical analysis, we need to make specific assumptions about the error term. The basic assumptions are as follows.

1) The error term, ε_u, is a random variable with mean zero, i.e., the *expected value* of the error term, $E(\varepsilon_u)$, is zero. This is by far the most important assumption since we can only claim that $E(b_i) = \beta_i$ if $E(\varepsilon_u) = 0$.[15]

2) The unobservable variance of the error term for each observation is an unknown constant σ^2, i.e., $Var(\varepsilon_u) = \sigma^2$ for all u. This assumption implies that $Var(Y_u) = \sigma^2$.

3) The possible values for the errors for any two different observations are uncorrelated, so that the covariance of ε_i and ε_j is zero, i.e., $Cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$,.

4) Finally, ε_u is distributed according to the Gaussian distribution, with mean zero and variance σ^2, i.e., $\varepsilon_u \approx N(0,\sigma^2)$, or equivalently we can say $\varepsilon_u / \sigma \approx N(0,1)$.

According to Wittink,[15] a violation of the third assumption, if it is not the result of model misspecification, is not critical. As long as the first assumption holds true, the parameter estimates are unbiased. The second and the third assumptions are necessary only to obtain a simple formula for the standard deviation of the parameter estimates. The fourth assumption is needed to justify the use of particular distributions (e.g. χ^2, F and t-distributions) to construct confidence regions (or intervals).

In order to minimize $S(\beta)$, given by Eq. B.5, we differentiate Eq. B.5 with respect to β, and set the resulting equation equal to zero to obtain the p normal (least squares) equations given by

$$\mathbf{X}^T\mathbf{X}\beta = \mathbf{X}^T\mathbf{Y}, \tag{B.6}$$

The $p \times p$ symmetric matrix $\mathbf{X}^T\mathbf{X}$ is usually referred to as the *information matrix*. If $\mathbf{X}^T\mathbf{X}$ is nonsingular, then Eq. B.6 has a unique solution. Letting $\mathbf{b}$ denote the value of β satisfying Eq. B.6, we see that

$$\mathbf{b} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}, \tag{B.7}$$

where $(\mathbf{X}^T\mathbf{X})^{-1}$ is the inverse matrix of $\mathbf{X}^T\mathbf{X}$.

The variance-covariance matrix of the parameter estimates is denoted by $V(\mathbf{b})$, and is a $p \times p$ symmetric matrix with the variances as the elements in the main diagonal, and the covariances between each pair of parameters as the off-diagonals. From Eq. B.7, and the assumption that $Var(Y_u) = \sigma^2$, it can be seen that

$$V(\mathbf{b}) = V\left[\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}\right]$$

$$= \left[\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]V(\mathbf{Y})\left[\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]^T$$

$$=\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\sigma^2. \tag{B.8}$$

Eq. B.8 is important for computing individual confidence intervals for elements of **b**. Since σ^2 is usually unknown, we estimate σ^2 from the parameter estimates. Letting s^2 denote this approximation of σ^2, then

$$s^2 = \frac{\sum_{u=1}^{n}\left(Y_u - \hat{Y}_u\right)^2}{n-p} = \frac{S(\mathbf{b})}{n-p} \tag{B.9}$$

is an unbiased estimate of the variance of the error term, ε_u.

From Eq. B.7, we see that each estimated parameter is a linear function of the observed values of the dependent variable Y. Recall that we have assumed that at each fixed set of X values, the corresponding Y_u is an independent normally distributed variable about the regression line with $Var(Y_u)=\sigma^2$; hence, **b** is a combination of independent normal variables which is also a normal variable. It follows that

$$\mathbf{b} \approx N\left[\beta, (\mathbf{X}^T\mathbf{X})^{-1}\sigma^2\right]. \tag{B.10}$$

Confidence Limits

Under the preceding conditions, it can be shown that each of the statistical quantities

$$\frac{b_i - \beta_i}{\sqrt{s^2\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)_{ii}}}, \qquad i=1,\ldots,p, \tag{B.11}$$

has a t-distribution with $n-p$ degrees of freedom. Throughout, $((\mathbf{X}^T\mathbf{X})^{-1})_{ii}$ is the ith diagonal element of the inverse matrix $(\mathbf{X}^T\mathbf{X})^{-1}$. If we denote by $t(\gamma, n-p)$ the value that

cuts off $\gamma/2\times100\%$ in the upper tail of this distribution, then a $(1-\gamma)\times100\%$ confidence interval of each parameter, β_i, is given by

$$b_i - t(\gamma, n-p)\sqrt{s^2\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)_{ii}} \le \beta_i \le b_i + t(\gamma, n-p)\sqrt{s^2\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)_{ii}}, \quad \text{(B.12)}$$

for $i = 1,\ldots,p$. Sometimes the confidence limits are presented in terms of a percentage uncertainty in the estimate b_i.

Note in Eq. B.12 the size of the confidence interval of parameter β_i depends on the magnitude of the estimated measurement error s^2, i.e., on the residual scatter about the fit, on the number of data points (to determine the t-value), and on the ith diagonal element of the matrix $(\mathbf{X}^T\mathbf{X})^{-1}$, which depends on the range of the independent variable X_i.[16, 17]

The length of the confidence interval for a given parameter β_i, (i.e., the uncertainty), comes in part from the uncertainty about the true values of the other parameters. Thus, the rectangular region defined by the corresponding confidence intervals for parameters β_i and β_j is not really appropriate to define the joint, two-dimensional confidence region for (β_i, β_j). When the correlation coefficient between b_i and b_j, defined by

$$\rho_{ij} = \frac{Cov(b_i, b_j)}{\left[Var(b_i)Var(b_j)\right]^{1/2}} \quad \text{(B.13)}$$

is close to 1 or -1, the difference between the correct two-dimensional confidence region for (β_i, β_j) and the rectangular region obtained from the individual confidence intervals is large. In Eq. B.13, the quantities $Var(b_i)$, $Var(b_j)$ and $Cov(b_i,b_j)$ are elements of the variance-covariance matrix defined in Eq. B.8. The correlation coefficient has a limit of $-1 \le \rho_{ij} \le 1$. A zero value of ρ_{ij} indicates absolute noncorrelation between parameters b_i and b_j, while ρ_{ij} of unity (in absolute value) reflects a perfect correlation. In this last case, b_i and b_j cannot be uniquely determined.

Joint Confidence Regions

In the following, we summarize all relevant theorems used later in our analysis.

Theorem 1: If z is normally distributed with zero mean and unit variance, then the sum of the squares of n random sample values of z has a χ^2 distribution with n degrees of freedom.[18]

Theorem 2: If χ_1^2 and χ_2^2 possess independent χ^2 distributions with υ_1 and υ_2 degrees of freedom, respectively, then $\chi_1^2 + \chi_2^2$ will possess a χ^2 distribution with $\upsilon_1 + \upsilon_2$ degrees of freedom.[18]

Theorem 3: If V_1 and V_2 are independent random variables, and V_1 and $V = V_1 + V_2$ have χ^2 distributions with υ_1 and υ degrees of freedom, respectively, and $\upsilon_1 < \upsilon$, then V_2 has a χ^2 distribution with $\upsilon - \upsilon_1$ degrees of freedom.[18]

Theorem 4: If X is a normally distributed random variable with variance σ^2, and s^2 is the sample variance based on a random sample of size n, then the statistic ns^2/σ^2 has a χ^2 distribution with $n-1$ degrees of freedom.[18]

Theorem 5: If u and v possess independent χ^2 distributions with υ_1 and υ_2 degrees of freedom, respectively, then

$$F = \frac{u/\upsilon_1}{v/\upsilon_2}$$

has a F distribution with υ_1 and υ_2 degrees of freedom.[18]

It follows directly from Theorem 1 and the normality assumption of the errors that if the model is correct, then the statistic,

$$\chi_n^2 = \sum_{u=1}^{n} \frac{\left[Y_u - \sum_{i=1}^{p} \beta_i X_{iu} \right]^2}{\sigma^2} = \frac{S(\beta)}{\sigma^2}, \tag{B.14}$$

possesses a χ^2 distribution with n degrees of freedom. A remarkable property of the χ^2 statistic in Eq. B.14 is that it is still applicable when the unknown parameters, β_i, are replaced by their maximum likelihood estimates, provided that one degree of freedom is deducted for each parameter estimated (see Hoel[18] p. 170). Thus, we can also say that

$$\chi_{n-p}^2 = \sum_{u=1}^{n} \frac{\left[Y_u - \sum_{i=1}^{p} b_i X_{iu} \right]^2}{\sigma^2} = \frac{S(\mathbf{b})}{\sigma^2} \tag{B.15}$$

has a χ^2 distribution with $n-p$ degrees of freedom.

Now we can apply Theorem 3 to the χ^2 distributions given by Eqs. B.14 and B.15 to conclude that

$$\chi_p^2 = \frac{S(\beta) - S(\mathbf{b})}{\sigma^2} \tag{B.16}$$

follows a χ^2 distribution with p degrees of freedom. Finally, since $S(\beta) - S(\mathbf{b})$ and $S(\mathbf{b})$ are distributed independently, it follows from Theorem 5 and Eqs. B.15 and B.16 that

$$\frac{[S(\beta) - S(\mathbf{b})]/p}{S(\mathbf{b})/(n-p)} \approx F(p, n-p). \tag{B.17}$$

Before we examine Eq. B.17 in terms of confidence region levels, we derive an expression for $S(\beta) - S(\mathbf{b})$ for the linear model we are considering. Substituting $\mathbf{b}$ for β in the last expression of Eq. B.5, and using the fact that $\mathbf{Y} = \mathbf{Xb}$, we get

$$S(\mathbf{b}) = \mathbf{Y}^T\mathbf{Y} - \mathbf{b}^T\mathbf{X}^T\mathbf{Xb}. \tag{B.18}$$

Now subtracting Eq. B.18 from Eq. B.5, we obtain

$$S(\beta)-S(\mathbf{b})=-2\beta^T\mathbf{X}^T\mathbf{Y}+\beta^T\mathbf{X}^T\mathbf{X}\beta+\mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{b}$$
$$=-2\beta^T\mathbf{X}^T\mathbf{X}\mathbf{b}+\beta^T\mathbf{X}^T\mathbf{X}\beta+\mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{b}$$
$$=(\beta-\mathbf{b})^T\mathbf{X}^T\mathbf{X}(\beta-\mathbf{b}). \tag{B.19}$$

Finally, the substitution of Eqs. B.9 and B.19 in Eq. B.17 yields the following expression:

$$(\beta-\mathbf{b})^T\mathbf{X}^T\mathbf{X}(\beta-\mathbf{b})\approx ps^2F(p,n-p). \tag{B.20}$$

At this point, it is easy to interpret Eqs. B.19 and B.20 in terms of a joint confidence region for the parameters. All values of β^0 which satisfy $S(\beta^0)$ = constant = a are given by the contour

$$(\beta^0-\mathbf{b})^T\mathbf{X}^T\mathbf{X}(\beta^0-\mathbf{b})=a-S(\mathbf{b}). \tag{B.21}$$

Note that $S(\mathbf{b})$ is the minimum value that $S(\beta^0)$ can assume, which implies that $a\geq S(\mathbf{b})$. It can be shown that Eq. B.19 is the equation of a closed ellipsoidal (or a hyper-ellipsoidal) contour surrounding the point $\mathbf{b}$. When $a_1>a_2$, the contour $S(\beta)=a_1$ completely encloses the contour $S(\beta)=a_2$, and $\mathbf{b}$ lies in the center of all ellipsoids.

From Eq. B.20, it follows that the contour of a $(1-\gamma)\times 100\%$ confidence region for the true (but unknown) value of β is enclosed by the contour

$$(\beta-\mathbf{b})^T\mathbf{X}^T\mathbf{X}(\beta-\mathbf{b})=ps^2F(p,n-p,1-\gamma). \tag{B.22}$$

Eq. B.22 is equivalent to

$$\frac{[S(\beta)-S(\mathbf{b})]/p}{S(\mathbf{b})/(n-p)}=F(p,n-p,1-\gamma), \tag{B.23}$$

which can be rearranged to obtain

$$S(\beta)=S(\mathbf{b})\left\{1+\frac{p}{n-p}F(p,n-p,1-\gamma)\right\}. \tag{B.24}$$

The right-hand-side of Eq. B.24 is the constant value that defines the $(1-\gamma)\times 100\%$ confidence contour for β. Eqs. B.22 to B.24 are identical for the linear model we are studying.

Ill-Conditioned Problem

In the numerical solution of the least-squares equations, Eq. B.6, basically two sources of inaccuracy may arise. The first one is the round-off errors in the computational process. The second source relates to the inaccuracy of the measurements and the inaccuracy of the elements of the information matrix, which may significantly affect the solution.

A system is called *ill-conditioned* if its inverse matrix is unstable, i.e., if its solution is significantly affected by small variations of the independent variable as well as variations in the coefficients. Reference 19 shows that if the elements of the inverse matrix are large, then a small change in the coefficients of the system or the independent variables produces a significant change in the solution. This is the case, for example, when some of the parameters are strongly correlated.

The condition of a matrix can be characterized by its *condition number*. Formally, the condition number (*P*-number) is defined as the ratio of the largest eigenvalue (in absolute value) to the smallest one (in absolute value). An infinite condition number means that the matrix is singular. Large condition number indicates that the matrix is ill-conditioned. Reference 20 states that a matrix is ill-conditioned if the reciprocal of its condition number approaches the machine's floating-point precision.

Weighted Least Squares

The traditional least squares estimation is often unsatisfactory when some observations are less reliable than others. In this case, we want to make sure that the parameter estimates will be influenced more by the more reliable observations than by the less accurate ones. This is, for example, the case where the variances of the observations are not all equal, i.e., the case where the variance matrix of the error term in Eq. B.3, $V(\varepsilon)$, is of the form $\mathbf{W}^{-1}\sigma^2$ instead of the classical form $\mathbf{I}\sigma^2$. Here, $\mathbf{W}$ is a known real symmetric positive definite or semidefinite $n \times n$ matrix. Since it is often assumed that the observations are uncorrelated, $\mathbf{W}$ is usually a diagonal matrix with $w_{ii} > 0$. The constants w_{ii} are generically called *weights*, a terminology we explain below.

The weighted least squares estimation can be written in terms of the usual (unweighted) LS problem if we properly redefine the model. Pre-multiplying both sides of Eq. B.3 by $\mathbf{W}^{1/2}$ we obtain a new model

$$\mathbf{W}^{1/2}\mathbf{Y} = \mathbf{W}^{1/2}\mathbf{X}\beta + \mathbf{W}^{1/2}\varepsilon \tag{B.25}$$

or

$$\mathbf{Y}^* = \mathbf{X}^*\beta + \varepsilon^*, \tag{B.26}$$

with an obvious notation. Note that since $E(\varepsilon) = 0$ and $V(\varepsilon) = \mathbf{W}^{-1}\sigma^2$, it follows, respectively, that $E(\varepsilon^*) = 0$ and $V(\varepsilon^*) = \mathbf{I}\sigma^2$. The matrix $\mathbf{W}^{1/2}$ in Eq. B.25 is given by[21]

$$\mathbf{W}^{1/2} = \left(\mathbf{P}\mathbf{D}^{1/2}\mathbf{P}^T\right)^{-1}, \tag{B.27}$$

where $\mathbf{D}$ is diagonal; $D_{ii} = d_i$, the ith eigenvalue of $\mathbf{W}^{-1}$; and the columns of matrix $\mathbf{P}$ are the normalized eigenvectors of $\mathbf{W}^{-1}$.

It is also true that ε^* is normally distributed, i.e., $\varepsilon^* \approx N(0, \mathbf{I}\sigma^2)$, since the elements of ε^* are linear combinations of elements of ε which is itself normally distributed.

Thus we can apply the basic least squares theory to Eq. B.26, particularly, in analogy to Eqs. B.7 and B.8, we can write the parameter estimates as

$$\mathbf{b} = \left(\mathbf{X}^{*T}\mathbf{X}^{*}\right)^{-1}\mathbf{X}^{*T}\mathbf{Y}^{*} = \left(\mathbf{X}^{T}\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^{T}\mathbf{W}\mathbf{Y}, \tag{B.28}$$

and the variance-covariance matrix of **b** as

$$V(\mathbf{b}) = \left(\mathbf{X}^{*T}\mathbf{X}^{*}\right)^{-1}\sigma^2 = \left(\mathbf{X}^{T}\mathbf{W}\mathbf{X}\right)^{-1}\sigma^2. \tag{B.29}$$

The general weighted LS objective function can be written as

$$\begin{aligned} S(\beta) &= \left(\mathbf{Y}^{*} - \mathbf{X}^{*}\beta\right)^{T}\left(\mathbf{Y}^{*} - \mathbf{X}^{*}\beta\right) \\ &= (\mathbf{Y} - \mathbf{X}\beta)^{T}\mathbf{W}(\mathbf{Y} - \mathbf{X}\beta). \end{aligned} \tag{B.30}$$

The residual sum of squares, $S(\mathbf{b})$ (see Eq. B.18), becomes[16]

$$\begin{aligned} S(\mathbf{b}) &= \mathbf{Y}^{*T}\mathbf{Y}^{*} - \mathbf{b}^{T}\mathbf{X}^{*T}\mathbf{Y}^{*} \\ &= \mathbf{Y}^{T}\mathbf{W}\mathbf{Y} - \mathbf{Y}^{T}\mathbf{W}\mathbf{X}\left(\mathbf{X}^{T}\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^{T}\mathbf{W}\mathbf{Y}. \end{aligned} \tag{B.31}$$

Under the assumption that the observations are uncorrelated, i.e, assuming that **W** is a diagonal matrix, Eq. B.30 reduces to

$$S(\beta) = \sum_{u=1}^{n} w_u \left[Y_u - \sum_{i=1}^{p} \beta_i X_{iu} \right]^2, \tag{B.32}$$

where w_u is the uth element of the diagonal matrix **W**. The interpretation of Eq. B.32 as weighted least squares is straightforward. Note that the weights are proportional to the inverse of the variances, thus weighted least squares gives more weight to data points with low variability, which is desirable. In this case, an estimate of σ^2 can be derived as[22]

$$s^2 = \frac{S(\mathbf{b})}{n-p} = \frac{\sum_{u=1}^{n} w_u \left(Y_u - \hat{Y}_u\right)^2}{n-p}. \tag{B.33}$$

At this point, one can easily write the expressions for computing confidence intervals and confidence regions in analogy to Eqs. B.12 and B.22. For example, the $(1-\gamma)\times 100\%$ confidence interval of each parameter, β_i, is calculated via

$$b_i - t(\gamma, n-p)\sqrt{s^2\left(\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\right)_{ii}} \le \beta_i \le b_i + t(\gamma, n-p)\sqrt{s^2\left(\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\right)_{ii}} \quad \text{(B.34)}$$

where s is obtained from Eq. B.31 or Eq. B.33.

If the model is linear, or if the number of observations is large and the errors are normally distributed, then the weighted least squares, with weights given by the elements of the inverse of the covariance matrix of the errors, is the estimator which leads to least-variance estimates.[21] Even though these optimal properties can not be proved for the general case (non-normal distributions with nonlinear models), the use of the result above is still reasonable. When the weights are unknown, one may use alternative methods such as the *generalized least squares estimate*.[21,22]

All the equations presented for the linear model case also apply for nonlinear models provided that the model can be approximated by a linear function of the parameters in a vicinity of the solution (see Chapter II, Section 2.3). In this case the matrix $\mathbf{X}$ should be replaced by the matrix $\mathbf{Z}$ defined in Eq. 2.3.5.

APPENDIX C

FUNCTIONAL MODEL AND PARAMETER GRADIENTS FOR THE CLASSICAL WELLBORE STORAGE AND SKIN PROBLEM

In this appendix, we develop working expressions for required partial derivatives of the model pressure-pressure derivative response with respect to the system parameters. The model pressure (and pressure derivative) responses are available in closed functional forms only in Laplace space. Since the model parameters are independent of time, we use the following relationships to generate real time derivatives from corresponding Laplace space expressions.

$$\Delta p(\alpha,t) = \mathcal{L}^{-1}\left[\overline{\Delta p}(\alpha,u)\right]. \tag{C.1}$$

Here $\mathcal{L}^{-1}$ denotes the inverse Laplace transform operator, and u represents the Laplace variable. Similarly,

$$\frac{\partial \Delta p(\alpha,t)}{\partial \alpha_i} = \mathcal{L}^{-1}\left[\frac{\partial \overline{\Delta p}(\alpha,u)}{\partial \alpha_i}\right] \quad \text{for } \alpha_i \in \alpha, \tag{C.2}$$

$$\begin{aligned}\frac{\partial \Delta p(\alpha,t)}{\partial \ln t} &= t\mathcal{L}^{-1}\left[u\overline{\Delta p}(\alpha,u)\right] \\ &= -\mathcal{L}^{-1}\left[\overline{\Delta p}(\alpha,u) + u\frac{\partial \overline{\Delta p}(\alpha,u)}{\partial u}\right]\end{aligned} \tag{C.3}$$

and finally,

$$\frac{\partial^2 \Delta p(\alpha,t)}{\partial \alpha_i \partial \ln t} = t\mathcal{L}^{-1}\left[u\frac{\partial \overline{\Delta p}(\alpha,u)}{\partial \alpha_i}\right]. \tag{C.4}$$

Laplace space expressions are numerically inverted to real space using the Stehfest algorithm.[7]

The solution in Laplace space for the classical wellbore storage and skin problem in SI units (or any other consistent system of units) is given by,

$$\overline{\Delta p}=\frac{\frac{1}{u}K_0\left(\sqrt{\frac{\phi u}{bk}}\right)+s\sqrt{\frac{\phi u}{bk}}K_1\left(\sqrt{\frac{\phi u}{bk}}\right)}{\frac{k}{a}\sqrt{\frac{\phi u}{bk}}K_1\left(\sqrt{\frac{\phi u}{bk}}\right)+\frac{C}{qB}u\left[K_0\left(\sqrt{\frac{\phi u}{bk}}\right)+s\sqrt{\frac{\phi u}{bk}}K_1\left(\sqrt{\frac{\phi u}{bk}}\right)\right]}. \tag{C.5}$$

Where K_0 and K_1 are the modified Bessel functions of the second kind of zero and unit orders; a and b are constants that depend on the fluid/ reservoir properties:

$$a=\frac{qB\mu}{2\pi h} \tag{C.6}$$

and

$$b=\frac{1}{\mu c_t r_w^2}. \tag{C.7}$$

With simple algebraic manipulation, it can be shown that the gradients of the functional model with respect to the primitive set of parameters, i.e., $\alpha=\{k,C,s,\phi\}$, can be written as

$$\frac{\partial\overline{\Delta p}}{\partial s}=\frac{a}{ku}\left(1-C\frac{u^2\overline{\Delta p}}{qB}\right), \tag{C.8}$$

$$\frac{\partial\overline{\Delta p}}{\partial C}=-\frac{u^2\overline{\Delta p}}{qB}, \tag{C.9}$$

$$\frac{\partial\overline{\Delta p}}{\partial\phi}=\frac{a\left(1-Cu^2\overline{\Delta p}/qB\right)^2}{2bk^2}\left\{\left[\frac{ku\overline{\Delta p}}{a\left(1-Cu^2\overline{\Delta p}/qB\right)}-s\right]^2-\frac{bk}{\phi u}\right\}, \tag{C.10}$$

$$\frac{\partial\overline{\Delta p}}{\partial k}=-\frac{\overline{\Delta p}}{k}\left(1-\frac{Cu^2\overline{\Delta p}}{qB}\right)-\frac{\phi}{k}\frac{\partial\overline{\Delta p}}{\partial\phi}, \tag{C.11}$$

and finally

$$\frac{\partial\overline{\Delta p}}{\partial u}=-2\frac{\overline{\Delta p}}{u}-\frac{k}{u}\frac{\partial\overline{\Delta p}}{\partial k}. \tag{C.12}$$

(In this work, we assume that ϕ is a known parameter.)

For the case where we perform the minimization procedure in terms of the combined parameters,

$$\alpha=\left\{k,\frac{k}{C},\frac{C\exp(2s)}{\phi}\right\},$$

the gradients of the functional model with respect to the combined parameters can be written in terms of the gradients with respect to the primitive parameters (see Eqs. C.8 to C.12) by using the concept of the total differential. The relevant formulas are as follows:

$$\begin{aligned}\frac{\partial\overline{\Delta p}}{\partial\alpha_1}&=\frac{\partial\overline{\Delta p}}{\partial k}+\frac{1}{k}\left[C\frac{\partial\overline{\Delta p}}{\partial C}-\frac{1}{2}\frac{\partial\overline{\Delta p}}{\partial s}\right]\\&=\frac{\partial\overline{\Delta p}}{\partial k}+\frac{1}{k}\left[C\frac{\partial\overline{\Delta p}}{\partial C}+\phi\frac{\partial\overline{\Delta p}}{\partial\phi}\right],\end{aligned} \tag{C.13}$$

$$\begin{aligned}\frac{\partial\overline{\Delta p}}{\partial\alpha_2}&=-\frac{C}{k}\left[C\frac{\partial\overline{\Delta p}}{\partial C}-\frac{1}{2}\frac{\partial\overline{\Delta p}}{\partial s}\right]\\&=-\frac{C}{k}\left[C\frac{\partial\overline{\Delta p}}{\partial C}+\phi\frac{\partial\overline{\Delta p}}{\partial\phi}\right]\end{aligned} \tag{C.14}$$

and

$$\frac{\partial \overline{\Delta p}}{\partial \alpha_3} = -\frac{\phi^2}{C\exp(2s)}\frac{\partial \overline{\Delta p}}{\partial \phi} = \frac{\phi}{2C\exp(2s)}\frac{\partial \overline{\Delta p}}{\partial s}. \qquad \text{(C.15)}$$

APPENDIX D

WELL TEST REGRESSION PACKAGE

In this appendix, we provide details on the FORTRAN code which implements nonlinear parameter estimation. As mentioned previously, we use the public-domain routine LMDER (a modified Levenberg-Marquardt minimization routine) from the Argonne National Laboratories.

The program handles multirate production during and prior to the well test (assuming that the rate history is a sequence of constant rates) by applying Duhamel's Principle. Note that buildup is a special case of the multirate production where the last production rate (measured at the surface) is zero. The program can be run in either a predictive mode, where model pressure and pressure derivative responses for given values of the model parameters are generated, or in an optimization mode, where initial estimates of the parameters are refined until convergence is achieved or the maximum number of iterations is exceeded. The program can handle tests where the measurements are associated with one or more flow rates. For example, we can apply the minimization procedure to the drawdown and buildup periods simultaneously. The program consists of approximately 14,000 lines (instruction and commentary lines) including the LMDER routine and other small public-domain routines.

Figure D.1 presents a schematic overview of the program design that is based on the work presented in Ref. 30.

Figure D.2 presents a schematic overview of the logic involved in generating the model responses, i.e., model pressure, pressure derivatives and partial derivatives of the functional model with respect to the regressing parameters.

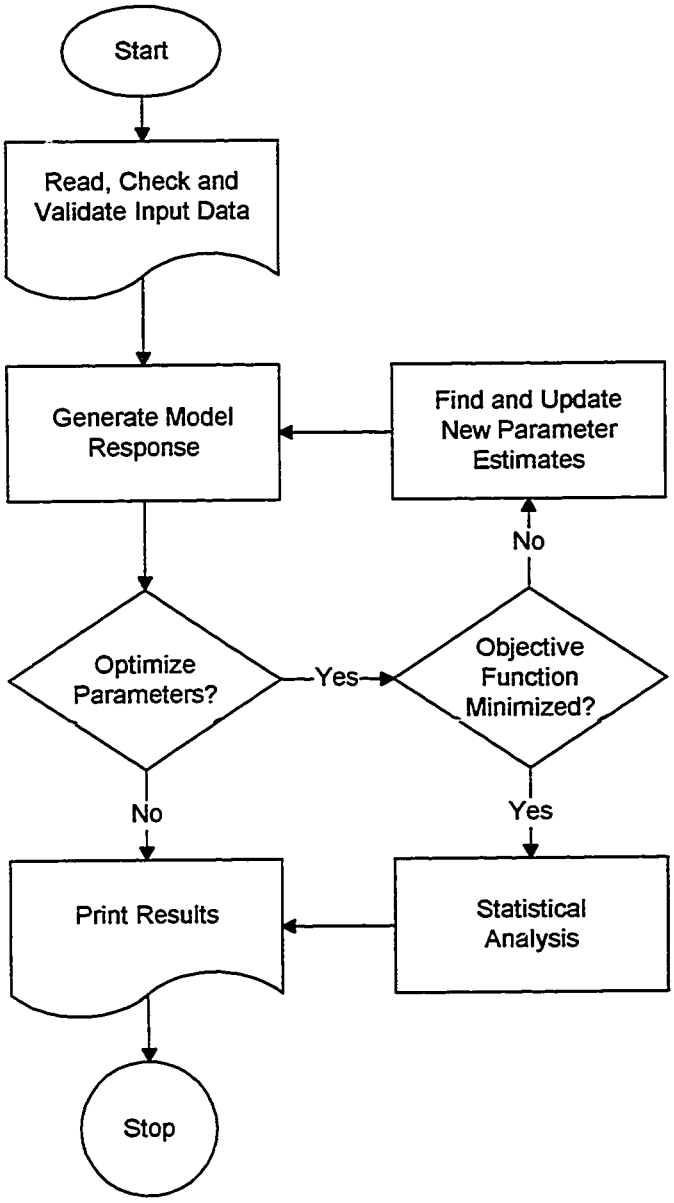

Fig. D.1 - Schematic Overview of the logic of the computer program

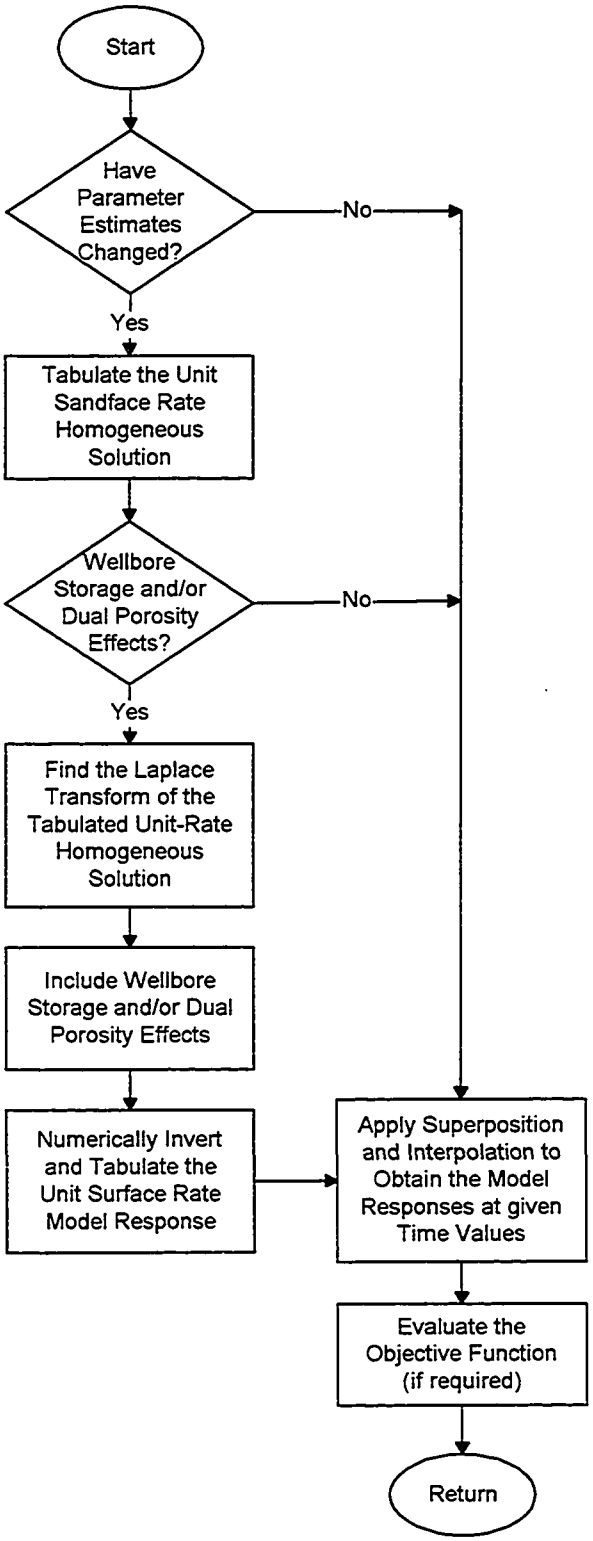

Fig. D.2 - Flow chart for generation of model response

Wellbore storage and dual porosity effects are incorporated into the model in Laplace space, based on the Laplace transform of the unit-sandface-rate homogeneous solution. In general, the Laplace transform of the unit-sandface-rate homogeneous solution is obtained by numerical integration of the tabulated real space solution (assuming that the function is piecewise linear). The set of time points at which the homogeneous model response is calculated, is different from the set of times at which we have measured data. To ensure accurate numerical Laplace transforms of the model, it is required that the span of the tabulated data exceeds the span of the measured data and that the initial time of the tabulated is not greater than 10^{-4} days; we use 1.5 log-cycles prior to the beginning of the measured data and 1.5 log cycles after the maximum time for the test, assuming that the origin of the time variable is the instant when the well is first opened for production. We use a fixed number of geometrically spaced time values per log cycle (15 points per log cycle is usually adequate); this information is used in the algorithms for numerical Laplace transform and for interpolation, which speeds up the computations.

Having generated the set of time points at which we will compute the tabulated data, we proceed to form the unit-sandface-rate pressure drop and its derivatives with respect to the regressing parameters. The unit-rate skin pressure drop is simply added to the unit-rate pressure drop. The model is chosen by the user from the list of models presented in Table D.1. Since the skin pressure drop is treated as steady state, the pressure drop may be negative at early times when the skin is negative. To bypass this problem, we either use the equivalent wellbore radius concept (see for example Ref. 31) or, alternatively, we set the unit-rate pressure drop to zero in the event it is negative. Both procedures produced equivalent results when applied to the horizontal well model. We use the latter procedure in the final version of the program.

If dual porosity and/or wellbore storage effects are required, then the tabulated data are numerically transformed to Laplace space where these effects are easily introduced into the model. The real time responses with wellbore storage and/or dual porosity

effects (at the tabulated data points) are obtained by numerical inversion using the Stehfest algorithm.

Table D.1 - List of the models implemented in the program code

Reservoir/Well Geometry	Bounded radial system (finite wellbore radius) Infinite radial system (finite wellbore radius) Fully penetrating line source well in a bounded reservoir Uniform-flux horizontal well in a bounded reservoir Fully penetrating vertically fractured well in a bounded system Partially penetrating vertically fractured well in a bounded system
Reservoir System	Anisotropic homogeneous system Anisotropic dual porosity system (Warren and Root model)
Wellbore Condition	Constant wellbore storage Fair's model for changing wellbore storage (Ref. 32) Hegeman's model for changing wellbore storage (Ref. 33)

The next step in obtaining the tabulated model response is the superposition of the tabulated unit-rate functions to account for the multirate production. Finally, the tabulated pressure drop and its derivatives with respect to the parameters are then interpolated to obtain the functions at the values of time where we have measured pressure or pressure derivatives. We use the Lagrange or polynomial interpolation with 3, 4 or 5-point interpolation (user defined). At the time of the interpolation, all required time derivatives are obtained using the interpolating polynomials.

Since changing wellbore storage solutions (Fair's or Hegeman's models) cannot be superposed for variable rate test, these two models are available only with single rate tests. In this case, buildup data must be analyzed (approximately) as an equivalent single rate drawdown test.

The parameter estimates are constrained as discussed in Chapter II, Section 2.3.6. The program may exhibit oscillations in the generated model response (especially pressure derivative data) at very early times, when the skin pressure drop attains a large negative value and the wellbore storage coefficient is small. This problem arises because of the Stehfest algorithm. We allow the user to specify the minimum allowable value for the skin pressure drop, and constrain the estimates so that they always exceed this value.

APPENDIX E

ANISOTROPIC DUAL-POROSITY SYSTEMS

In this appendix, we show how we generate dual-porosity solutions for anisotropic systems, and how these solutions are related to the corresponding homogeneous anisotropic problems.

Consider the two-dimensional control volume of dimensions Δx, Δy, h (depth is normal to the plane of the paper) shown in Fig. E.1. The reservoir consists of two systems—the fracture system that transports the fluids and the matrix system that feed the fractures.

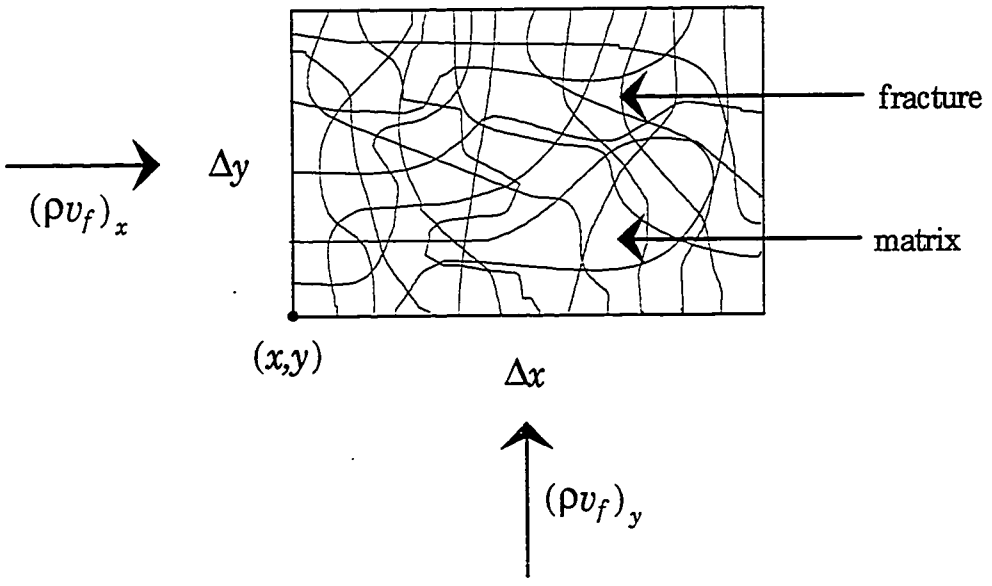

Fig. E.1 - Illustration of the dual-porosity control volume

Let us first introduce the following definitions:

$$V_f = \frac{\text{bulk volume of fractures}}{V_B} = \frac{V_{Bf}}{V_B}, \tag{E.1}$$

$$V_m = \frac{\text{bulk volume of matrix}}{V_B} = \frac{V_{Bm}}{V_B} = 1 - V_f, \tag{E.2}$$

where V_B is the total bulk volume.

The conservation of mass principle applied to the control volume of Fig. E.1 is given by

$$5.615 V_f \left[\frac{\partial}{\partial x}(\rho v_{fx}) + \frac{\partial}{\partial y}(\rho v_{fy}) \right] - \frac{5.615 \rho q_m}{V_B} = -\frac{\partial}{\partial t}(\rho \phi_f V_f), \tag{E.3}$$

where v_f $(bbl/ft^2 day)$ is the fluid velocity in the fractures, and $5.615\rho q_m$ (lbm/day) is the mass flow from the matrix to the fractures. Flow from the matrix to the fractures, q_m, is considered positive. The bulk properties of the rock are defined as follows:

$$\hat{\phi}_f = \phi_f V_f, \tag{E.4}$$

$$\hat{\phi}_m = \phi_m V_m, \tag{E.5}$$

$$\hat{k}_f = k_f V_f, \tag{E.6}$$

$$\text{and } \hat{k}_m = k_m V_m. \tag{E.7}$$

Using Darcy's equation to relate velocities and pressure gradients in the fracture system, i.e.,

$$v_{fx} = -1.127 \times 10^{-3} \frac{k_{fx}}{\mu} \frac{\partial p_f}{\partial x} \tag{E.8}$$

and

$$v_{fy} = -1.127 \times 10^{-3} \frac{k_{fy}}{\mu} \frac{\partial p_f}{\partial y}, \tag{E.9}$$

Eq. E.3 can be written as

$$0.00633\left[\frac{\partial}{\partial x}\left(\rho\frac{\hat{k}_{fx}}{\mu}\frac{\partial p_f}{\partial x}\right)+\frac{\partial}{\partial y}\left(\rho\frac{\hat{k}_{fy}}{\mu}\frac{\partial p_f}{\partial y}\right)\right]+\frac{5.615\rho q_m}{V_B}=\frac{\partial}{\partial t}\left(\rho\hat{\phi}_f\right), \qquad \text{(E.10)}$$

where p_f denotes pressure in the fracture system.

We define a specific flow rate from the matrix to the fractures, q^* $(bbl/ft^3\,day)$, as

$$q^* = \frac{\text{flow from matrix to fractures}}{\text{bulk volume of matrix}} = \frac{q_m}{V_{Bm}}. \qquad \text{(E.11)}$$

Using Eq. E.11 and noting that $V_{Bm} = V_m V_B$ (see Eq. E.2), Eq. E.10 becomes

$$0.00633\left[\frac{\partial}{\partial x}\left(\rho\frac{\hat{k}_{fx}}{\mu}\frac{\partial p_f}{\partial x}\right)+\frac{\partial}{\partial y}\left(\rho\frac{\hat{k}_{fy}}{\mu}\frac{\partial p_f}{\partial y}\right)\right]+5.615\rho V_m q^* = \frac{\partial}{\partial t}\left(\rho\hat{\phi}_f\right). \qquad \text{(E.12)}$$

Similarly, the flow equation within each matrix element can be written as

$$0.00633\bar{\nabla}.\left[\rho\frac{\hat{k}_m}{\mu}\nabla p_m\right]-5.615\rho V_m q^* = \frac{\partial}{\partial t}\left(\rho\hat{\phi}_m\right). \qquad \text{(E.13)}$$

Note we have assumed that the matrix is isotropic.

In Barenblatt's model,[45,46] the right-hand-side term in Eq. E.13 is neglected, i.e., they assume that change in mass due to fluid compressibility and change in fracture porosity is negligible compared to change in mass due to the mass flux. We will follow Warren and Root[47] and not neglect this term.

Finally, using the standard assumptions of slightly compressible fluid of constant compressibility and constant viscosity, and also assuming that $\hat{k}_m$, $\hat{k}_{fx}$, $\hat{k}_{fy}$, $\hat{\phi}_m$, and $\hat{\phi}_f$ are constant, we modify Eqs. E.12 and E.13 to obtain the final partial differential equations which govern the flow in the reservoir.

Fracture system pde:

$$\eta_{fx}\frac{\partial^2 \Delta p_f}{\partial x^2}+\eta_{fy}\frac{\partial^2 \Delta p_f}{\partial y^2}-\frac{5.615}{\hat{\phi}_f c_{tf}}q^* V_m=\frac{\partial \Delta p_f}{\partial t}. \tag{E.14}$$

Matrix system pde:

$$\eta_m \nabla^2 \Delta p_m+\frac{5.615}{\hat{\phi}_m c_{tm}}q^* V_m=\frac{\partial \Delta p_m}{\partial t}. \tag{E.15}$$

In Eqs. E.14 and E.15, the matrix diffusivity and the fracture diffusivity are defined, respectively, as

$$\eta_m=\frac{0.00633\hat{k}_m}{\hat{\phi}_m \mu c_{tm}} \tag{E.16}$$

and

$$\eta_f=\frac{0.00633\hat{k}_f}{\hat{\phi}_f \mu c_{tf}}. \tag{E.17}$$

Eq. E.15 is a general equation for the flow in the matrix system. The Laplacian, ∇^2, is related to the shape of the matrix elements (e.g., spheres, slabs, cubes, etc.). Note that both matrix and fracture systems are assumed to be homogeneous, however, the fracture system can be anisotropic. Since in many dual-porosity applications the matrix system consists of small blocks, the assumption of an isotropic matrix is reasonable. In this case, k_m represents the effective permeability of the matrix.

At this point, we consider the basic hypotheses of Barenblatt's model[45,46] that any infinitesimal volume of the reservoir has a large number of matrix blocks. They assumed that each point of the reservoir has two pressures associated with it—p_f, fracture pressure, and p_m, matrix pressure. Another hypothesis is that the flow from the matrix to the fractures is given by

$$q_m = 1.127\times10^{-3}\alpha\frac{\hat{k}_m}{\mu}\left(p_m - p_f\right)V_B. \tag{E.18}$$

which is often referred to in the literature as *the pseudo-steady state inter-porosity flow* model. The shape factor, α $(1/ft^2)$, is a characteristic parameter of the fractured reservoir, which reflects the geometry of the matrix. The units of q_m is bbl/day.

The pseudo-steady state flow model (Eq. E.18) can also be written in terms of the specific flow rate (see Eq. E.11) as

$$q^* = 1.127\times10^{-3}\alpha\frac{\hat{k}_m}{\mu}\frac{1}{V_m}\left(\Delta p_f - \Delta p_m\right). \tag{E.19}$$

Substituting Eq. E.19 into Eq. E.14, we obtain

$$\eta_{fx}\frac{\partial^2\Delta p_f}{\partial x^2} + \eta_{fy}\frac{\partial^2\Delta p_f}{\partial y^2} - 0.00633\alpha\frac{\hat{k}_m}{\mu}\frac{1}{\hat{\phi}_f c_{tf}}\left(\Delta p_f - \Delta p_m\right) = \frac{\partial\Delta p_f}{\partial t}. \tag{E.20}$$

Another assumption, (see Refs. 46 and 47) is that the flux term in Eq. E.15 (i.e., the term $\eta_m\nabla^2\Delta p_m$) is negligible. With this assumption, the partial differential equation for the matrix system becomes

$$0.00633\alpha\frac{\hat{k}_m}{\mu}\frac{1}{\hat{\phi}_m c_{tm}}\left(\Delta p_f - \Delta p_m\right) = \frac{\partial\Delta p_m}{\partial t}. \tag{E.21}$$

Eqs. E.20 and E.21 are the anisotropic versions of the so-called Warren and Root model. The extension of Eq. E.20 to three-dimensional flow is straightforward; we simply add another flux term corresponding to the z-direction. Throughout, the initial condition of the system is assumed to be

$$\Delta p_m(t=0) = p_i - p_m(t=0) = 0 \tag{E.22}$$

and

$$\Delta p_f(t=0) = p_i - p_f(t=0) = 0. \tag{E.23}$$

Taking the Laplace transform with respect to t of Eq. E.21 and manipulating the resulting transformed equation, we obtain

$$\overline{\Delta p_m} = \left[\frac{1}{1+(u/\alpha\eta_m)}\right]\overline{\Delta p_f}, \tag{E.24}$$

where u is the Laplace variable. Note that in order to obtain Eq. E.22 we have used the initial condition (Eq. E.23).

Taking the Laplace transform of Eq. E.20, and substituting Eq. E.24 in the resulting equation, we obtain the following expression after some algebraic manipulations

$$\eta_{fx}\frac{\partial^2\overline{\Delta p_f}}{\partial x^2}+\eta_{fy}\frac{\partial^2\overline{\Delta p_f}}{\partial y^2}=uf(u)\overline{\Delta p_f}, \tag{E.25}$$

where $f(u)$ is given by

$$f(u)=1+0.00633\frac{\hat{k}_m}{\hat{\phi}_f\mu c_{tf}}\left(\frac{\alpha}{u+\alpha\eta_m}\right)$$

$$=1+\frac{1}{\hat{\phi}_f c_{tf}}\left(\frac{1}{\hat{\phi}_m c_{tm}}+\frac{u}{0.00633\alpha\hat{k}_m/\mu}\right)^{-1}. \tag{E.26}$$

It is important to note that $f(u)$ is not function of $\hat{k}_{fx}$ or $\hat{k}_{fy}$. For homogeneous systems, it is well known that the flow equation in Laplace space (which corresponds to the homogeneous version of Eq. E.25) is given by

$$\eta_x\frac{\partial^2\overline{\Delta p}}{\partial x^2}+\eta_y\frac{\partial^2\overline{\Delta p}}{\partial y^2}=u\overline{\Delta p}, \tag{E.27}$$

where η_x and η_y are the hydraulic diffusivities in the x and y directions, respectively (see Nomenclature).

Following References 48 and 49, we seek a relationship between the homogenous and dual-porosity solutions, i.e., we wish to generate the anisotropic dual-porosity solutions from the anisotropic homogeneous solutions. Without loss of generality, we

consider the problem of an instantaneous point source in a three-dimensional anisotropic reservoir. "Sources and sinks" theory[50] shows that this is a fundamental problem whose solution can be integrated (superposed) with respect to space and time to generate the solutions to many different problems, e.g., vertical wells, horizontal wells, partial penetration, etc.

The diffusivity equation (pde) for an homogeneous three-dimensional anisotropic reservoir can be written as

$$k_x \frac{\partial^2 \Delta p}{\partial x^2} + k_y \frac{\partial^2 \Delta p}{\partial y^2} + k_z \frac{\partial^2 \Delta p}{\partial z^2} = \frac{\phi \mu c_t}{0.00633} \frac{\partial \Delta p}{\partial t}. \tag{E.28}$$

Eq. E.28 can also be written as an *equivalent* isotropic system via

$$\frac{\partial^2 \Delta p}{\partial \tilde{x}^2} + \frac{\partial^2 \Delta p}{\partial \tilde{y}^2} + \frac{\partial^2 \Delta p}{\partial \tilde{y}^2} = \frac{\phi \mu c_t}{\bar{k}} \frac{\partial \Delta p}{\partial t}, \tag{E.29}$$

where

$$\bar{k} = \left(k_x k_y k_z\right)^{1/3}, \tag{E.30}$$

$$\tilde{x} = \left(\bar{k}/k_x\right)^{1/2} x, \tag{E.31}$$

$$\tilde{y} = \left(\bar{k}/k_y\right)^{1/2} y, \tag{E.32}$$

and

$$\tilde{z} = \left(\bar{k}/k_z\right)^{1/2} z. \tag{E.33}$$

In order to express the inner-boundary condition in a convenient easy way, we use Eq. E.29 in spherical coordinates. The coordinates of the point source are (x', y', z') for the original problem (Eq. E.28), or $(\tilde{x}', \tilde{y}', \tilde{z}')$ for the equivalent isotropic problem (Eq. E.29). If we place the point source at the origin, Eq. E.29 can be written in the spherical coordinate system as

$$\frac{\partial^2 \Delta p}{\partial r^2} + \frac{2}{r} \frac{\partial \Delta p}{\partial r} = \frac{1}{\bar{\eta}} \frac{\partial \Delta p}{\partial t}, \tag{E.34}$$

where

$$r=\sqrt{(\tilde{x}-\tilde{x}')^2+(\tilde{y}-\tilde{y}')^2+(\tilde{z}-\tilde{z}')^2}\ , \text{ and} \tag{E.35}$$

$$\overline{\eta}=\frac{0.00633\bar{k}}{\phi\mu c_t}=\frac{0.00633\left(k_x k_y k_z\right)^{1/3}}{\phi\mu c_t}. \tag{E.36}$$

$\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ are defined in Eqs. E.31-35.

The instantaneous point source problem (i.e., the Green's function for a point source) can be found by solving the diffusivity equation subject to the condition of an instantaneous removal of Q barrels of fluid from the source. For a three-dimensional anisotropic reservoir with the source at the origin of the coordinate system, the instantaneous point source problem is given by the following IBVP.

$$\text{pde:}\quad \frac{\partial^2\Delta p}{\partial r^2}+\frac{2}{r}\frac{\partial\Delta p}{\partial r}=\frac{1}{\overline{\eta}}\frac{\partial\Delta p}{\partial t}, \tag{E.37}$$

$$\text{Initial condition:}\quad \Delta p(r,t=0)=0, \tag{E.38}$$

$$\text{Outer-boundary condition:}\quad \lim_{r\to\infty}\Delta p(r,t)=0, \tag{E.39}$$

$$\text{Inner-boundary condition:}\quad \lim_{r\to 0}\left[r^2\frac{\partial\Delta p}{\partial r}\right]=\frac{-Q\mu}{4\pi\bar{k}}\frac{\delta(t)}{1.127\times10^{-3}}, \tag{E.40}$$

where $\delta(t)$ is the Dirac delta function, r is defined in Eq. E.35 and $\overline{\eta}$ in Eq. E.36. The inner-boundary condition (Eq. E.40) comes directly from Darcy's law. The same IBVP (Eqs. E.37 through E.40) in Laplace space takes the form

$$\text{ode:}\quad \frac{d^2\overline{\Delta p}}{dr^2}+\frac{2}{r}\frac{d\overline{\Delta p}}{dr}=\frac{u}{\overline{\eta}}\overline{\Delta p}, \tag{E.41}$$

$$\text{Outer-boundary condition:}\quad \lim_{r\to\infty}\overline{\Delta p}(r,u)=0, \tag{E.42}$$

$$\text{Inner-boundary condition:}\quad \lim_{r\to 0}\left[r^2\frac{d\overline{\Delta p}}{dr}\right]=\frac{-Q\mu}{4\pi\bar{k}}\frac{1}{1.127\times10^{-3}}, \tag{E.43}$$

where u is the Laplace variable. Note that $\mathcal{L}\{\delta(t)\}=1$. The solution to this problem in Laplace space is

$$\overline{\Delta p}(r,u)=\frac{\underline{Q}\mu}{4\pi\bar{k}}\frac{1}{1.127\times10^{-3}r}\exp\left(-\sqrt{\frac{u}{\bar{\eta}}}r\right). \tag{E.44}$$

In real space, Eq. E.44 takes the form

$$\Delta p(r,t)=\frac{5.615\underline{Q}}{8\phi c_t(\pi\bar{\eta}t)^{3/2}}\exp\left(\frac{-r^2}{4\bar{\eta}t}\right), \tag{E.44a}$$

or

$$\Delta p(x,y,z,t)=\frac{5.615\underline{Q}}{8\phi c_t\sqrt{\eta_x\eta_y\eta_z}(\pi t)^{3/2}}\exp\left(-\frac{(x-x')^2}{4\eta_x t}-\frac{(y-y')^2}{4\eta_y t}-\frac{(z-z')^2}{4\eta_z t}\right) \tag{E.45}$$

in Cartesian coordinates, after substituting the definitions of r and $\bar{\eta}$ given by Eq. E.35 and Eq. E.36, respectively. Recall that (x',y',z') are the coordinates of the source. Eq. E.45 is exactly the instantaneous point source function (multiplied by the strength of the source) presented by Gringarten.[50]

The problem of an instantaneous point source in an infinite anisotropic reservoir with dual-porosity behavior (Warren and Root model) is given, in Laplace space, by

$$\text{ode:}\quad \frac{d^2\overline{\Delta p_f}}{dr^2}+\frac{2}{r}\frac{d\overline{\Delta p_f}}{dr}=\frac{uf(u)}{\bar{\eta}_f}\overline{\Delta p_f}, \tag{E.46}$$

$$\text{Outer-boundary condition:}\quad \lim_{r\to\infty}\overline{\Delta p_f}(r,u)=0, \tag{E.47}$$

$$\text{Inner-boundary condition:}\quad \lim_{r\to0}\left[r^2\frac{d\overline{\Delta p_f}}{dr}\right]=\frac{-\underline{Q}\mu}{4\pi\bar{k}_f}\frac{1}{1.127\times10^{-3}}, \tag{E.48}$$

where

$$\bar{k}_f=\left(\hat{k}_{fx}\hat{k}_{fy}\hat{k}_{fz}\right)^{1/3}, \tag{E.49}$$

$$\overline{\eta}_f = \frac{0.00633\bar{k}_f}{\hat{\phi}_f \mu c_{tf}} = \frac{0.00633\left(\hat{k}_{fx}\hat{k}_{fy}\hat{k}_{fz}\right)^{1/3}}{\hat{\phi}_f \mu c_{tf}}. \tag{E.50}$$

$f(u)$ is given by Eq. E.26 and the space variable, r, is still given by Eq. E.35, provided that $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ (given by Eqs. E.31-33, respectively) are defined in terms of fracture properties, i.e., in Eqs. E.31-33, $\bar{k}$ is replaced by $\bar{k}_f$ and k_v is replaced by $\hat{k}_{fv}$, for $v = x, y, z$.

If we compare the homogeneous problem in Laplace, Eqs. E.41-43, with the dual-porosity problem, Eqs. E.46-48, there is a clear analogy between the two problems—the dual porosity problem is equivalent to the homogeneous problem provided that $u/\overline{\eta}$ is replaced by $uf(u)/\overline{\eta}_f$ and $\bar{k}$ is replaced by $\bar{k}_f$. The key point is that the boundary conditions are independent of u. The way the problem is set up, the parameters of the dual-porosity model are $\left\{\bar{k}_f, \alpha\hat{k}_m, \hat{\phi}_f c_{tf}, \hat{\phi}_m c_{tm}\right\}$. Note that the fluid viscosity, μ, is not in the parameter list because it is assumed to be known.

We can also define the dual-porosity flow equation (Eq. E.46) in terms of $\overline{\eta}_{fm}$, i.e., Eq. E.46 can be written as

$$\frac{d^2\overline{\Delta p_f}}{dr^2} + \frac{2}{r}\frac{d\overline{\Delta p_f}}{dr} = \frac{ug(u)}{\overline{\eta}_{f+m}}\overline{\Delta p_f}, \tag{E.51}$$

where

$$g(u) = f(u)\hat{\phi}_f c_{tf} / (\hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm}), \tag{E.52}$$

$$\overline{\eta}_{f+m} = \frac{0.00633\bar{k}_f}{(\hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm})} = \frac{0.00633\bar{k}_f}{\mu(\hat{\phi}c_t)_{f+m}}. \tag{E.53}$$

In this case, the dual porosity problem is equivalent to the homogeneous problem if $u/\overline{\eta}$ is replaced by $ug(u)/\overline{\eta}_{f+m}$ and $\bar{k}$ is replaced by $\bar{k}_f$. Here, the model parameters can be written as $\left\{\bar{k}_f, \alpha\hat{k}_m, (\hat{\phi}c_t)_{f+m}, \hat{\phi}_f c_{tf}\right\}$, where

$$(\hat{\phi}c_t)_{f+m} = \hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm}. \tag{E.54}$$

Thus, we can easily obtain the Green's function (in Laplace space) for an instantaneous point source in an anisotropic reservoir with dual-porosity behavior if we know the Laplace space solution for the homogeneous system. The dual-porosity solution for the instantaneous point source can be integrated (superposed) to generate many other sources. For example, we can generate the Green's function for the well-known line source problem based on the fundamental instantaneous point source solution. First, we find the solution for the instantaneous point source in a slab reservoir by using the classical imaging method,[31] i.e., the slab solution can be written as an infinite summation of fundamental instantaneous point source solutions evaluated at each image. Then this solution can be integrated over the slab thickness to generate the instantaneous line source solution.

The general procedure to obtain the instantaneous source function for an anisotropic dual-porosity reservoir based on the corresponding solution for the homogeneous system is summarized as follows:

1. Replace u in the homogenous solution in Laplace space by $uf(u)$ (see Eq. E.26) and replace the product ϕc_t by $\hat{\phi}_f c_{tf}$. Alternatively, we can also replace u in the homogenous solution in Laplace space by $ug(u)$ (see Eq. E.52). In this case, the product ϕc_t is replaced by $(\hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm})$.
2. Replace the homogenous permeability by the fracture system properties, i.e., in the homogeneous solution replace k_v by $\hat{k}_{fv}$, for $v = x, y, z$.
3. Take the inverse Laplace transform.

The Stehfest algorithm is normally used to numerically invert the solution from Laplace space in step 3.

It can be shown (see for example Ref. 48) that the instantaneous source function (i.e., the Green's function of the source) can be obtained via

$$\Delta p_G(x,y,z,t) = \frac{\partial \Delta p_u(x,y,z,t)}{\partial t} = \Delta p'_u(x,y,z,t), \tag{E.55}$$

where Δp_G is the Green's function and Δp_u is the constant unit rate solution for the same source. Suppose we know, in real space, the instantaneous source function for a homogeneous reservoir (Δp_G). By definition, its Laplace transform is

$$\overline{\Delta p_G}(u) = \int_0^\infty e^{-ut} \Delta p_G(r, \bar{k}, \phi c_t, t) dt. \tag{E.56}$$

Applying the procedure discussed previously, the dual-porosity instantaneous function is given by

$$\overline{\Delta p_{fG}}(u) = \int_0^\infty e^{-uf(u)t} \Delta p_G(r, \bar{k}_f, \hat{\phi}_f c_{tf}, t) dt, \tag{E.57}$$

where $\overline{\Delta p_{fG}}$ is the Green's function for the dual-porosity behavior. Also, we can obtain the dual-porosity unit-rate solution, $\overline{\Delta p_{fu}}$, via

$$\overline{\Delta p_{fu}}(u) = \frac{1}{u} \int_0^\infty e^{-uf(u)t} \Delta p'_u(r, \bar{k}_f, \hat{\phi}_f c_{tf}, t) dt, \tag{E.58}$$

or equivalently, via

$$\overline{\Delta p_{fu}}(u) = f(u) \int_0^\infty e^{-uf(u)t} \Delta p_u(r, \bar{k}_f, \hat{\phi}_f c_{tf}, t) dt. \tag{E.59}$$

In order to obtain Eqs. E.58 and E.59, we assume that the initial conditions of both problems are $\Delta p_G(t=0) = \Delta p_u(t=0) = 0$ and $\Delta p_{fG}(t=0) = \Delta p_{fu}(t=0) = 0$. Wellbore storage and skin effects are also included in the model in Laplace space.[30, 42]

Warren and Root[47] defined two dimensionless parameters for the dual-porosity model,

$$\lambda = L^2\left(\alpha \hat{k}_m / \bar{k}_f\right) \tag{E.60}$$

and

$$\omega = \hat{\phi}_f c_{tf} \Big/ (\hat{\phi}_f c_{tf} + \hat{\phi}_m c_{tm}). \tag{E.61}$$

In Eq. E.60, L is a reference length; $L = r_w$ for vertical-well problems and L is the well length for horizontal well problems. λ is the inter-porosity flow coefficient and determines how rapidly the matrix starts contributing to the flow, and ω is the ratio of the fracture storativity to the total reservoir storativity.

If we use λ and ω as model parameters, $f(u)$ in Eq. E.46 becomes

$$f(u) = \frac{\lambda + (1-\omega) u L^2 / \overline{\eta}_f}{\lambda\omega + (1-\omega) u L^2 / \overline{\eta}_f}, \tag{E.62}$$

which is associated with the parameter list $\left\{\bar{k}_f, \lambda, \omega, \hat{\phi}_f \hat{c}_{tf}\right\}$. Also, in this case $g(u)$ (see Eq. E.52) becomes

$$g(u) = \omega f(u) = \frac{\lambda + \omega(1-\omega) u L^2 / \overline{\eta}_{f+m}}{\lambda + (1-\omega) u L^2 / \overline{\eta}_{f+m}}, \tag{E.63}$$

which corresponds to the parameter list $\left\{\bar{k}_f, \lambda, \omega, (\hat{\phi}_f \hat{c}_{tf} + \hat{\phi}_m \hat{c}_{tm})\right\}$. We have implemented both options.

The regression procedure requires partial derivatives of the model with respect to the parameters. The parameter derivatives in Laplace space are obtained by differentiating Eq. E.59, i.e., for a generic parameter α, the partial derivative of the model in Laplace space with respect to α can be written as

$$\frac{\partial}{\partial\alpha}\left(\overline{\Delta p_{fu}}(u)\right) = \frac{\partial f(u)}{\partial\alpha}\left[\int_0^\infty e^{-uf(u)t} \Delta p_u dt - uf(u) \int_0^\infty e^{-uf(u)t} t \Delta p_u dt\right]$$

$$+ f(u) \int_0^\infty e^{-uf(u)t} \frac{\partial \Delta p_u}{\partial\alpha} dt. \tag{E.64}$$

In Eq. E.64, the integrals are evaluated numerically assuming that Δp_u is piecewise linear[12] (we normally use 15 points per log-cycle). All partial derivatives of the unit-rate solution, $\partial \Delta p_u / \partial \alpha$, are the same as used for the homogeneous model provided we replace k by $\hat{k}_f$ and ϕc_t by $\hat{\phi}_f c_{tf}$. Note that if $\alpha \in \{\hat{k}_m, \lambda, \omega\}$ then $\partial \Delta p_u / \partial \alpha = 0$. Here, we need an expression for $\partial f(u)/\partial \alpha$, or $\partial g(u)/\partial \alpha$, where $f(u)$ and $g(u)$ are given by Eqs. E.26, E.52, E.62 and E.63. For $f(u)$ given by Eq. E.26, the partial derivatives are

$$\frac{\partial f(u)}{\partial \bar{k}_f} = 0, \tag{E.65}$$

$$\frac{\partial f(u)}{\partial (\hat{\phi}_f c_{tf})} = \frac{1 - f(u)}{\hat{\phi}_f c_{tf}}, \tag{E.66}$$

$$\frac{\partial f(u)}{\partial \hat{k}_m} = \frac{u[1 - f(u)]^2 \bar{k}_f}{\alpha \hat{k}_m^2 \bar{\eta}_f}, \tag{E.67}$$

$$\frac{\partial f(u)}{\partial (\hat{\phi}_m c_{tm})} = \frac{[1 - f(u)]^2 (\hat{\phi}_f c_{tf})}{(\hat{\phi}_m c_{tm})^2}, \tag{E.68}$$

For $g(u)$ given by Eq. E.52, the partial derivatives are

$$\frac{\partial g(u)}{\partial \bar{k}_f} = 0, \tag{E.69}$$

$$\frac{\partial g(u)}{\partial (\hat{\phi}_f c_{tf})} = \frac{\alpha \eta_m [1 - f(u)] / [(\hat{\phi} c_t)_{f+m} - (\hat{\phi}_f c_{tf})] + u / (\hat{\phi} c_t)_{f+m}}{(\alpha \eta_m + u)}, \tag{E.70}$$

$$\frac{\partial g(u)}{\partial \hat{k}_m} = \frac{\alpha \eta_m [1 - f(u)]}{(\alpha \eta_m + u) \hat{k}_m}, \tag{E.71}$$

$$\frac{\partial g(u)}{\partial (\hat{\phi} c_t)_{f+m}} = -\frac{\alpha \eta_m [1 - f(u)] (\hat{\phi}_f c_{tf}) / [(\hat{\phi} c_t)_{f+m} - (\hat{\phi}_f c_{tf})] + u f(u)}{(\alpha \eta_m + u)(\hat{\phi} c_t)_{f+m}}, \tag{E.72}$$

where $(\hat{\phi}c_t)_{f+m}$ is given by Eqs. E.54 and

$$\eta_m = \frac{0.00633\hat{k}_m/\mu}{(\hat{\phi}c_t)_{f+m} - (\hat{\phi}_f c_{tf})}. \tag{E.73}$$

For $f(u)$ given by Eq. E.62, the partial derivatives are

$$\frac{\partial f(u)}{\partial \bar{k}_f} = \frac{L^2(1-\omega)^2 u\lambda\bar{\eta}_f}{\left[\omega\lambda\bar{\eta}_f + L^2(1-\omega)u\right]^2 \bar{k}_f}, \tag{E.74}$$

$$\frac{\partial f(u)}{\partial \lambda} = \frac{L^2(1-\omega)^2 u\bar{\eta}_f}{\left[\omega\lambda\bar{\eta}_f + L^2(1-\omega)u\right]^2}, \tag{E.75}$$

$$\frac{\partial f(u)}{\partial \omega} = \frac{\lambda^2\bar{\eta}_f^2}{\left[\omega\lambda\bar{\eta}_f + L^2(1-\omega)u\right]^2}, \tag{E.76}$$

and

$$\frac{\partial f(u)}{\partial(\hat{\phi}_f c_{tf})} = \frac{L^2(1-\omega)^2 u\lambda\bar{\eta}_f}{\left[\omega\lambda\bar{\eta}_{fm} + L^2(1-\omega)u\right]^2 \hat{\phi}_f c_{tf}}. \tag{E.77}$$

Finally, for $g(u)$ given by Eq. E.63, the partial derivatives are

$$\frac{\partial g(u)}{\partial \bar{k}_f} = \frac{L^2(1-\omega)^2 u\lambda\bar{\eta}_{f+m}}{\left[\lambda\bar{\eta}_{f+m} + L^2(1-\omega)u\right]^2 \bar{k}_f}, \tag{E.78}$$

$$\frac{\partial g(u)}{\partial \lambda} = \frac{L^2(1-\omega)^2 u\bar{\eta}_{f+m}}{\left[\lambda\bar{\eta}_{f+m} + L^2(1-\omega)u\right]^2}, \tag{E.79}$$

$$\frac{\partial g(u)}{\partial \omega} = \frac{\left[L^2(1-\omega)u\right]^2 + 2L^2(1-\omega)u\lambda\overline{\eta}_{f+m}}{\left[\lambda\overline{\eta}_{f+m} + L^2(1-\omega)u\right]^2}, \tag{E.80}$$

and

$$\frac{\partial g(u)}{\partial(\hat{\phi}c_t)_{f+m}} = \frac{-L^2(1-\omega)^2 u\lambda\overline{\eta}_{f+m}}{\left[\lambda\overline{\eta}_{f+m} + L^2(1-\omega)u\right]^2(\hat{\phi}c_t)_{f+m}}, \tag{E.81}$$

where $\overline{\eta}_{f+m}$ and $(\hat{\phi}c_t)_{f+m}$ are given by Eqs. E.53 and Eqs. E.54, respectively.

APPENDIX F

AN ALGORITHM FOR NUMERICAL TRANSFORM OF TABULATED DATA

In this appendix, we derive an algorithm to take the Laplace transform of tabulated real data. The proposed algorithm is especially applicable to pressure data, dimensionless or not.

In 1988, Romboutsos and Stewart[12] presented an algorithm based on a linear interpolation of the measured data and a linear extrapolation using the cord slope of the last interval. In 1991, Bourgeois and Horne[56] presented an algorithm which assumes a semilog extrapolation of the data. They also considered a semilog extrapolation with respect to the Horner time ratio for buildup data. Of course this algorithm is more appropriate for data that exhibit a semilog straight line at the end of the test.

We will consider a piecewise linear approximation to Δp as a function of time. Thus, between the measured data, the pressure is interpolated linearly in the same way as proposed by Romboutsos and Stewart.[12] Beyond the measured data, we consider a linear extrapolation of the data in terms of its derivative with respect to the logarithm of time. As we shall demonstrate, in general, the proposed algorithm fits better the late time behavior of the data, which yields a better extrapolation. The proposed algorithm also incorporates both procedures mentioned above as particular cases.

The pairs $(t_j,\ \Delta p_j)$, for $j = 0,\ 1,\ldots,\ n$ represent the tabulated data, where n is the total number of data points. In each interval $[t_{i-1},\ t_i]$, the function is approximated by the following straight line:

$$\Delta p(t) = \Delta p_{i-1} + \Delta p'_{i-1}(t - t_{i-1}) \quad \text{for} \quad t_{i-1} \le t \le t_i, \tag{F.1}$$

where $\Delta p'$ is given by the chord slope, i.e.,

$$\Delta p'_{i-1} = \frac{\Delta p_i - \Delta p_{i-1}}{t_i - t_{i-1}}. \tag{F.2}$$

The contribution of the $(i+1)th$ interval to the Laplace transform of $\Delta p(t)$ is given by

$$I_{i+1} = \int_{t_i}^{t_{i+1}} e^{-ut} \Delta p(t) dt \approx \int_{t_i}^{t_{i+1}} e^{-ut} \left[\Delta p_i + \Delta p'_i (t - t_i) \right] dt. \tag{F.3}$$

Performing the integration in Eq. F.3 and using Eq. F.2, we obtain

$$I_{i+1} = \frac{\Delta p_i}{u} e^{-ut_i} + \frac{\Delta p'_i}{u^2} \left(e^{-ut_i} - e^{-ut_{i+1}} \right) - \frac{\Delta p_{i+1}}{u} e^{-ut_{i+1}}. \tag{F.4}$$

The contributions of all intervals from $[0, t_1]$ to $[t_{n-1}, t_n]$ to the Laplace transform of $\Delta p(t)$ is given by

$$I(t_n) = \int_0^{t_n} e^{-ut} \Delta p(t) dt \approx \sum_{i=0}^{n-1} I_{i+1}, \tag{F.5}$$

or, using Eq. F.4 in F.5, we obtain

$$I_n = \frac{\Delta p_0}{u} + \frac{\Delta p'_0}{u^2} \left(1 - e^{-ut_1} \right) + \sum_{i=1}^{n-1} \frac{\Delta p'_i}{u^2} \left(e^{-ut_i} - e^{-ut_{i+1}} \right) - \frac{\Delta p_n}{u} e^{-ut_n} \tag{F.6}$$

At this point, we consider the extrapolation issue.

The assumption of the linear extrapolation of the pressure derivative with respect to the logarithm of time can be expressed as

$$\frac{d\Delta p}{d \ln t} = t \frac{d\Delta p}{dt} \approx at + b \quad \text{for } t \geq t_n \tag{F.7}$$

where a and b are obtained from the last few pressure derivative data points. We usually fit a best straight line through the pressure derivative data points on the last 0.3 logarithmic cycle. The actual number of pressure derivative data to be used for extrapolation purposes depends on the quality of the data. Separating variables in Eq. F.7 and integrating from t_n to t, we obtain the corresponding pressure extrapolation as

$$\Delta p_{ext}(t) = \Delta p(t_n) + a(t - t_n) + b\ln(t/t_n) \quad \text{for} \quad t \geq t_n. \tag{F.8}$$

If we also assume that the log-pressure derivative given by Eq. F.7 will never cross the zero value (which means that the pressure extrapolation is monotonic, either increasing or decreasing), then the contribution of the extrapolation to the Laplace transform of $\Delta p(t)$ is given by

$$I_{ext} = \int_{t_n}^{t_e} e^{-ut} \Delta p_{ext}(t)dt + \int_{t_e}^{\infty} e^{-ut} \Delta p_{ext}(t_e)dt, \tag{F.9}$$

where $t_e > t_n$ is the time when the log-pressure derivative extrapolation becomes zero, and $\Delta p(t_e)$ is the pressure obtained by extrapolation to t_e. In this case, from Eq. F.7, we can see that $t_e = -b/a$. Eq. F.9 says that once we reach the zero derivative plateau, the pressure is extrapolated as a constant equal to $\Delta p(t_e)$. In the case where the log-pressure derivative extrapolation never reaches the zero plateau, Eq. F.9 reduces to

$$I_{ext} = \int_{t_n}^{\infty} e^{-ut} \Delta p_{ext}(t)dt, \tag{F.10}$$

which is equivalent to setting t_e to infinity in Eq. F.9. Now, substituting Eq. F.8 into Eq. F.9, and performing the integration, we obtain after many algebraic manipulations:

$$I_{ext} = \frac{a}{u^2}\left(e^{-ut_n} - e^{-ut_e}\right) + \frac{\Delta p_n}{u} e^{-ut_n} + \frac{b}{u}\left[E_i(-ut_e) - E_i(-ut_n)\right], \tag{F.11}$$

where $-E_i(-x)$ is the exponential integral function defined as

$$-E_i(-x) = \int_x^{\infty} \frac{e^{-\xi}}{\xi} d\xi. \tag{F.12}$$

Finally, the Laplace transform of $\Delta p(t)$ is obtained by summing up Eqs. F.6 and F.11 as

$$\begin{aligned} \overline{\Delta p}(u) = {} & \frac{\Delta p_0}{u} + \frac{\Delta p_0'}{u^2}\left(1 - e^{-ut_1}\right) + \sum_{i=1}^{n-1} \frac{\Delta p_i'}{u^2}\left(e^{-ut_i} - e^{-ut_{i+1}}\right) + \frac{a}{u^2}\left(e^{-ut_n} - e^{-ut_e}\right) \\ & + \frac{b}{u}\left[E_i(-ut_e) - E_i(-ut_n)\right]. \end{aligned} \tag{F.13}$$

It is interesting to note that if $t_e \to \infty$, $b = 0$ and $a = \Delta p'_{n-1}$, Eq. F.13 reduces to the well-known Romboutsos and Stewart algorithm. For $t_e \to \infty$ and $a = 0$, Eq. F.13 reduces to the Bourgeois and Horne semilog extrapolation method.[56] Thus, the proposed algorithm is more general than the previous ones.

We are also interested in computing the Laplace transform of the log-pressure derivative data. This can be achieved by using the following expression,

$$\mathcal{L}\left\{\frac{d\Delta p}{d\ln t}\right\} = \frac{d}{du}\left(-u\overline{\Delta p}\right). \tag{F.14}$$

Substituting Eq. F.13 in Eq. F.14 and differentiating with respect to u, we obtain

$$\begin{aligned}\mathcal{L}\left\{\frac{d\Delta p}{d\ln t}\right\} = {} & \frac{\Delta p'_0}{u^2}\left(1-e^{-ut_1}\right) - \frac{\Delta p'_0}{u}t_1 e^{-ut_1} + \sum_{i=1}^{n-1}\frac{\Delta p'_i}{u^2}\left(e^{-ut_i} - e^{-ut_{i+1}}\right) \\ & + \sum_{i=1}^{n-1}\frac{\Delta p'_i}{u}\left(t_i e^{-ut_i} - t_{i+1}e^{-ut_{i+1}}\right) + \frac{a}{u^2}\left(e^{-ut_n} - e^{-ut_e}\right) + \frac{(at_n+b)}{u}e^{-ut_n}\end{aligned} \tag{F.15}$$

Again, in the case where the log-pressure derivative extrapolation does not reach the zero derivative plateau, the expression for the Laplace transform of the log-pressure derivative function is obtained by taking the limit of Eq. F.15 as $t_e \to \infty$. In this case, the exponential term that contains t_e in Eq. F.15 disappears.