Inverse Problems in Engineering: Theory and Practice
3rd Int. Conference on Inverse Problems in Engineering
June 13-18, 1999, Port Ludlow, WA, USA

3ICIPE99/ID11

NONLINEARITY, SCALE, AND SENSITIVITY FOR PARAMETER ESTIMATION PROBLEMS: SOME IMPLICATIONS FOR ESTIMATION ALGORITHMS

Trond Mannseth*
Reservoir Department
RF-Rogaland Research
Thormøhlensgt. 55
N-5008 Bergen, Norway
Email: trond.mannseth@rf.no

Kari Brusdal†
Department of Mathematics
University of Bergen
Johs. Brunsgt. 12
N-5008 Bergen, Norway

Alv-Arne Grimstad‡
Department of Physics
University of Bergen
Allégt. 55
N-5007 Bergen, Norway

Jan-Erik Nordtvedt
Reservoir Department
RF-Rogaland Research
Thormøhlensgt. 55
N-5008 Bergen, Norway
Email: jan-erik.nordtvedt@rf.no

Geir Nævdal
Reservoir Department
RF-Rogaland Research
Thormøhlensgt. 55
N-5008 Bergen, Norway
Email: geir.naevdal@rf.no

ABSTRACT

Both sensitivity and nonlinearity are important for the efficiency of an estimation algorithm. Knowledge of a general nature on sensitivity and/or nonlinearity for some class of models can perhaps be utilized to improve the estimation efficiency for this class.

For an ODE model, a correlation between high nonlinearity, low sensitivity, and small-scale perturbations, has been reported. Also, it was found that representing the unknown function by a multi-scale basis lead to faster estimation convergence than use of a single-scale local basis. This was explained referring to the above-mentioned correlation. Recently, the existence of such a correlation for a large class of nonlinear models, including the above-mentioned ODE model, was found.

Here, we further investigate into utilization of the correlation between nonlinearity, scale, and sensitivity within parameter estimation. Results from numerical experiments with alternative optimizers on ODE and PDE models describing flow in porous media, are presented.

1 INTRODUCTION

The mathematical modeling and simulation of the flow of fluids through porous media are important for designing and controlling a number of industrial processes, including the production of oil and gas from underground reservoirs and remediation of underground water resources. The equations describing porous-media flow contain several coefficient functions which are inaccessible to measurement, but their specification is crucial for the predictive power of the models. This gives rise to several inverse problems of different structure and complexity.

In this paper, we study two such problems. A closer description of the problems is given later (see Section 4), but a common feature is that the amount of detail with which it is possible to determine the coefficient functions, is not known a priori. An

* Address all correspondence to this author.
† Now with Nansen Environmental and Remote Sensing Center, Edvard Griegsvei 3a, N-5037 Solheimsviken, Norway, E-mail: karib@nrsc.no
‡ Now with RF-Rogaland Research, Thormøhlensgt. 55, N-5008 Bergen, Norway, Email: alv-arne.grimstad@rf.no

important issue is therefore to be able to exert some control over the scale for which one can be able to obtain accurate determination of the desired property for a given data set and expected data accuracy.

In a multi-scale representation – such as use of the Haar basis (Chui, 1992; Daubechies, 1992) to represent piecewise constant functions – variation on many length scales can be accounted for. The Haar basis is the first-order spline-wavelet basis. Multi-scale representation of continuous functions, and functions with higher degree of smoothness can be provided by higher-order spline-wavelets (Chui and Quak, 1992).

In (Chavent and Liu, 1989; Liu, 1993), a correlation between high nonlinearity, small-scale perturbations, and low sensitivity was found for a 1D model problem, in the context of estimation of a spatially dependent conductivity function. A multi-scale approach to estimation of the conductivity in that model was investigated in (Chavent and Liu, 1989; Liu, 1993), with positive results.

In this paper we attempt to present some general features of the influence of a correlation between nonlinearity, scale, and sensitivity, on estimation efficiency. It is based on results in (Grimstad and Mannseth, 1999; Brusdal and Mannseth, 1999; Brusdal et al., 1999; Nævdal et al., 1999). We consider several optimizers: The BFGS quasi-Newton algorithm and three variants of the Levenberg-Marquardt algorithm. The algorithms are applied to two different model problems related to fluid flow in petroleum reservoirs: The first is a further study with the ODE-model studied in, e.g., (Chavent and Liu, 1989; Liu, 1993), attempting to recover a spatially dependent coefficient function. In the second, the task is to recover a state dependent coefficient function, i.e., a function varying with the dependent variable in the forward model.

In (Grimstad and Mannseth, 1999), the above-mentioned correlation was found to be valid for a large class of parameter estimation problems, including the one studied in (Chavent and Liu, 1989; Liu, 1993). The structure of the class of problems considered in (Grimstad and Mannseth, 1999) indicates that also the problem involving estimation of the state dependent function might be included. However, due to the complexity of the forward model in that problem, this can only be investigated into through numerical experiments.

In Section 2 some background on the correlation between nonlinearity, scale, and sensitivity is given. Scale discriminate estimation is discussed in Section 3. This includes a brief presentation of spline-wavelets. In Section 4, the model equations for both of the porous-media flow models considered are presented, and some basic reservoir physics is briefly discussed. Finally, we present and discuss some results of our estimations in Section 5.

2 NONLINEARITY, SCALE, AND SENSITIVITY FOR GENERIC MODELS

In this section, we give a brief background on, and present some results illustrating, a correlation between nonlinearity, scale, and sensitivity, for a certain class of models. It is included primarily as a motivation for the remaining sections in the paper. Readers who are particularly interested in the correlation as such, will find a more thorough description in (Grimstad and Mannseth, 1999).

2.1 Theory

Let x_i denote external variables in a model (such as time, temperature, etc). Let $F_i(c) = F(x_i, c)$ be the outcome of the model when evaluated at the conditions x_i with the parameter values c.

The directional derivatives of F in the direction h in parameter space are denoted F_h for the first order derivatives and F_{hh} for the second order derivatives. If we decompose the second order derivative F_{hh} into F^t_{hh} parallel to the tangent plane defined by F_h for all directions h, and F^n_{hh} normal to this plane, the nonlinearity measures γ^n and γ^t of Bates and Watts (Bates and Watts, 1980) are defined as the maximum values of the curvatures $\kappa^n_h = \|F^n_{hh}\|/\|F_h\|^2$ and $\kappa^t_h = \|F^t_{hh}\|/\|F_h\|^2$ for any direction h. In this work we will use the parameter effects curvature, $\kappa^t_h = \|F^t_{hh}\|/\|F_h\|^2$, as our measure of nonlinearity.

Suppose that the task is to estimate the parameters, c, associated with the function $f(x,c)$ from measured data on some nonlinear function of f. In (Grimstad and Mannseth, 1999), the following model classes are studied:

$$m(x,c) = g(f(x,c)), \tag{1}$$

for models in the non-integrated class, and

$$M(x,c) = \int_0^x m(t,c)\,dt, \tag{2}$$

for the class of integrated models. The function g can be any smooth nonlinear function. Note that recovery of a coefficient function, f, in an ODE- or a PDE-model have strong similarities to the latter class, since solutions to such model equations are in some sense integrals involving a nonlinear function of f.

2.1.1 Model derivatives

For the calculation of sensitivity and curvature of the models, the first- and second-order directional derivatives in an arbitrary direction h are needed. If the function f is parameterized by a linear expansion in a set of basis functions, $f(x,c) = \sum_i c_i \psi_i(x)$, these are

$$m_h(x,c) = \sum_i h_i \frac{\partial m(x,c)}{\partial c_i} = g'(f(x,c))\,f(x,h), \tag{3}$$

$$m_{hh}(x,c) = \sum_{i,j} h_i h_j \frac{\partial^2 m(x,c)}{\partial c_i \partial c_j} = g''(f(x,c))\, f^2(x,h). \qquad (4)$$

for the non-integrated models. The derivatives of any integrated model, M, are obtained by integrating Equations 3, 4 with respect to x.

2.1.2 Sensitivity and curvature We want to investigate if there is a general correlation between low sensitivity and large curvature seen only within the class of integrated models. This means that we look for reasons causing $||M_h||$ to be small and $||M_{hh}||/||M_h||^2$ to be large, that would not produce similar effects in $||m_h||$ and $||m_{hh}||/||m_h||^2$.

Suppose that h is the direction vector of maximum curvature for a model. There is one possibility that h is also a direction of low sensitivity which is a unique feature of integral models: If the integrand is oscillating rapidly about zero as a function of the integration variable. From Equations 3, 4, it is seen that if $g'(f(x,c))$ is slowly varying, such that rapid oscillations in m_h are caused only by rapid oscillations in $f(x,h)$, similar oscillations in m_{hh} would not occur. This would have the desired influence on the sensitivities and curvatures in both model classes, indicating the existence of a correlation between high nonlinearity, low sensitivity, and small scale (in the sense of rapid oscillations) for the class of integrated models. In (Grimstad and Mannseth, 1999), several other scenarios leading to rapid oscillations of the integrand are discussed.

2.2 Results for generic models

To check the predictions of the theory, numerical experiments with the selection of nonlinear model functions listed in Table 1, have been performed. Also, several normalized multi-scale expansion bases for $f(x,h)$ were tested, offering a variety of possibilities to generate both slowly varying and rapidly oscillating functions. See (Grimstad and Mannseth, 1999) for details.

Fig. 1 shows the curvature and the norm of the sensitivity vector in the unit parameter directions when combining model functions G and H in Table 1 with a basis of five sine-functions with half-periods ranging from 1 to $2.5 \cdot 10^{-2}$. The parameters are ordered according to half-period lengths of the associated basis functions, such that parameter no. 1 corresponds to half-period 1, and parameter no. 5 corresponds to half-period $2.5 \cdot 10^{-2}$. The function $f(x,c)$ is proportional to the sine-function with half-period 1, i.e., a slowly varying function on [0,1].

All choices for $g(f)$, also those not shown here (Grimstad and Mannseth, 1999), follow the same trend in behavior: For the integrated models, perturbations $f(x,h)$ with shorter characteristic length have associated a lower sensitivity and a higher curvature. For the non-integrated models, the variation in sensitivity and curvature is minor and difficult to relate to the characteristic length of $f(x,h)$.

In (Grimstad and Mannseth, 1999), several other scenarios were considered, and among other things, it was found that the correlation is less evident if $f(x,c)$ is not slowly varying.

Table 1. Nonlinear model functions.

A:	$g(f) = f + \exp(f)$
B:	$g(f) = f/(1+f^2)$
C:	$g(f) = f^2/2$
D:	$g(f) = 1/2 \ln(1+f^2)$
E:	$g(f) = 2\exp(f+1) - \exp(2f-3)$
F:	$g(f) = 1 + \exp(-f^2/5)$
G:	$g(f) = [1+\exp(-f^2/5)]^{1/2}$
H:	$g(f) = (f+5)^{f/5}$

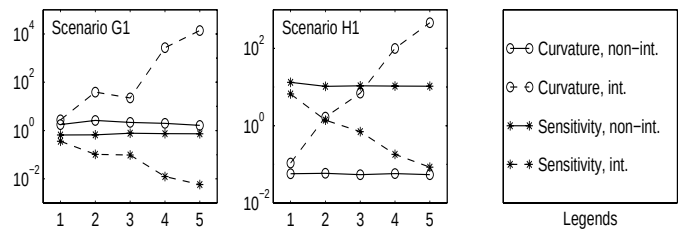

Figure 1. Correlation between nonlinearity (curvature), scale, and sensitivity. Scale decreases from left to right.

3 SCALE-DISCRIMINATE ESTIMATION

Many estimation algorithms that utilize sensitivity information to construct the next step, like quasi-Newton and Levenberg-Marquardt, rely on a linearization of the model function about the current point in parameter space. In this section, we briefly describe some strategies which seek to utilize the correlation between nonlinearity, scale, and sensitivity, when applying such an estimation algorithm. But first, a short introduction to spline-wavelets is given.

3.1 Spline-wavelets

Spline-wavelets are multi-scale functions. Essentially all length scales are represented in the support, and in the characteristic length of variation. This facilitates an investigation into the role of various length scales on the estimation problem, see

Sections 3.2 and 3.3.

The construction of spline-wavelets is closely related to B-splines, see (Chui and Quak, 1992) for details. A few basic properties of spline-wavelets are given below.

Spline-wavelets on $[0,1]$ of order m are defined on a nested sequence of subspaces

$$V_0 \subset V_1 \subset V_2 \dots \quad . \tag{5}$$

The space, V_0, is spanned by m'th-order B-splines without interior knots, on $[0,1]$. A basis for the space V_j is the set of m'th-order B-splines with interior knots $\left\{k2^{-j}\right\}_{k=1,\dots,2^j-1}$. The number of basis functions spanning V_j is 2^j+m-1. The spline-wavelets span the subspaces $W_j = V_{j+1} \setminus V_j$. For $m=1$, the spline-wavelets constitute the Haar basis.

Here, we represent the coefficient functions using first- and third-order spline-wavelets on the interval $[0,1]$. In Figure 2 we show such spline-wavelets spanning V_2.

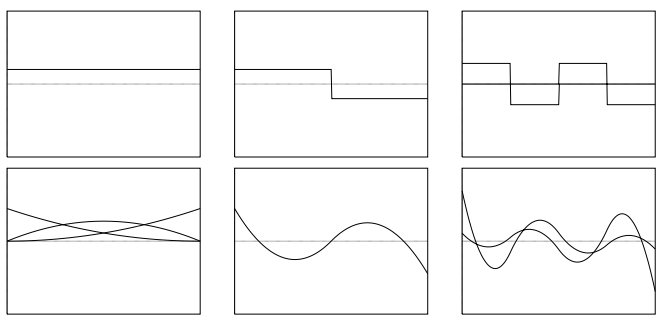

Figure 2. Spline-wavelets of order 1 (upper row) and 3 (lower row) spanning $V_2 = V_0 \oplus W_0 \oplus W_1$: V_0 (left), W_0 (middle) and W_1 (right).

3.2 Hierarchical estimation

Non-uniqueness and instability problems associated with over-parametrization can be avoided by seeking the simplest (in terms of the least number of parameters) estimated functions reconciling the data (Watson et al., 1988). To this end, an approach where a sequence of estimation problems with an increasing number of unknowns are solved, was first suggested in (Watson et al., 1988), and later verified in a series of cases. (See, e.g., (Grimstad et al. 1997; Kulkarni et al., 1998; Richmond and Watson, 1990).) With a multi-scale representation of the coefficient functions, this approach can be pursued in a systematic manner. This involves introducing parameters corresponding to smaller and smaller scales-of-variation in the basis functions into the estimation problem, sequentially. Such hierarchical estimation was performed in (Liu, 1993). Note that hierarchical estimation is an extreme form of scale-discriminate estimation.

3.3 Scale discriminate standard estimation

In standard estimation, estimation of all relevant parameters simultaneously, is attempted. A simple strategy to systematically enhance sensitivities to large-scale parameters on behalf of sensitivities to small-scale parameters, within standard estimation, was introduced in (Brusdal and Mannseth, 1999) for the BFGS quasi-Newton algorithm (see, e.g., (Gill et al., 1981)). It was termed *norm rescaling*. With this approach, a sliding scale-discriminate estimation, controlled by a single quantity, γ, results. This quantity is the ratio of sensitivity enhancement between neighboring scale levels. That is, basis elements in W_j are divided by γ^j. Hence, with $\gamma \neq 1$, scale-discriminate estimation results, and $\gamma > 1$ corresponds to increased emphasize on large-scale components.

In (Brusdal et al., 1999; Nævdal et al., 1999), norm rescaling was applied to define and test three variants of the Levenberg-Marquardt algorithm (see, e.g., (Gill et al., 1981)). It was shown that pure norm rescaling corresponds to modifications both in the trust-region test, and in the step formula. This variant will be referred to as (TS). The two other variants correspond to modifications in either the trust-region test (T), or the step formula (S). For details, see (Brusdal et al., 1999; Nævdal et al., 1999).

Norm rescaling can also be applied along with hierarchical estimation, since large-scale parameters are not fixed when smaller-scale parameters are introduced. The large-scale parameters merely provide a good initial value for the estimation involving also the next scale level. Although scale-discriminate standard estimation does not remove the risk of over-parameterization, it will be of interest to compare convergence speeds of this algorithm to that of hierarchical estimation, for various values of γ.

4 POROUS-MEDIA FLOW MODELS

In porous media there exists, as the name indicates, an interconnected matrix of pores and channels through which fluids may flow. We study macroscopic models for single- and two-phase fluid flow in porous media. The fraction of porous medium available for fluid flow is called the porosity, ϕ, and is an example of a volume averaged quantity, i.e., a quantity characteristic of a macroscopic point containing many individual pores. All quantities entering the model equations below should be interpreted in this sense. A standard reference on porous-media flow in the context of petroleum reservoirs is (Aziz and Settari, 1979).

4.1 Single-phase flow

Steady-state, 1D, horizontal flow of a single incompressible fluid phase in a porous medium is described by an equation formally identical to that describing steady-state, 1D heat conduction

$$-\frac{d}{dx}\left(a(x)\frac{dp}{dx}\right) = r(x). \tag{6}$$

In this equation, $a(x)$ represents the *permeability* (i.e., the fluid conductivity) of the porous medium, $r(x)$ is a source term representing injection or production of fluid, and $p(x)$ represents the fluid pressure. The permeability is assumed to be piecewise constant. The inverse problem considered for single-phase flow in Section 5.1, is to recover $a(x)$ from spatially distributed data on $p(x)$.

4.1.1 Relation to the class of integrated models

The solution to Equation 6 can be expressed in closed form, see e.g., (Liu, 1993). The result is an integral in x, but more complex than Equation 2. Still, the correlation between nonlinearity, scale, and sensitivity can be shown to hold for this model also (Grimstad and Mannseth, 1999). (See also (Chavent and Liu, 1989; Liu, 1993).) However, at stage n in an estimation, the current point in parameter space, c_n, may not always correspond to a slowly varying function, $a_n(x)$. As mentioned in Section 2, the correlation can be expected to be less evident when $a_n(x)$ is not slowly varying (Grimstad and Mannseth, 1999).

4.2 Two-phase flow

When two fluid phases flow simultaneously in a porous medium, each phase obstructs the flow of the other phase. To account for this, a quantity, k_i, called the *relative permeability* of phase i, is introduced. The relative permeability is a monotonically increasing function of the fluid saturation, S_i, i.e., the fraction of pore space occupied by phase i. It may assume values in the range [0,1].

Due to interfacial tension, two phases may coexist in the porous medium at different phase pressures, p_i. The pressure difference is given by the *capillary pressure*, $P_c = p_2 - p_1$, which is a function of fluid saturation. With a nonzero P_c, the process of fluid flow in porous media is not symmetrical with respect to the two phases.

Denoting the viscosity of phase i by μ_i, 1D, horizontal flow of two immiscible, incompressible fluid phases in a porous medium is then described by

$$\phi\frac{\partial S_i}{\partial t} = \frac{\partial}{\partial x}\left(\frac{a\,k_i(S_i)}{\mu_i}\frac{\partial p_i}{\partial x}\right), \quad i = 1,2, \tag{7}$$

$$p_2 - p_1 = P_c(S_1), \tag{8}$$

$$S_1 + S_2 = 1. \tag{9}$$

The inverse problem considered for two-phase flow in Section 5.2, is to recover k_2 from data measured on a small sample (typical dimensions are in the range of a few cm) of the porous medium when fluid 2 is displaced by fluid 1, which is injected at one end of the sample. The data will consist of time series of produced volume of fluid 2, and pressure drop across the sample. For simplicity, it will be assumed that k_1 and P_c are known functions of fluid saturation.

4.2.1 Relation to class of integrated models

The estimated function of the current model is a not a direct function of the integration variable, as is the case for the models described by Equation 2, and in Section 4.1. Rather, it is a function of x and t through the dependence on fluid saturation. To study this kind of dependence, Equation 2 should perhaps be replaced by generic models of the form

$$M(t,c) = \int_0^t \int_0^L g(f(S(u,v),c))\,du\,dv. \tag{10}$$

However, an explicit correspondence between such a model and Equations 7–9 is not known. Hence, Equation 10 is only intended as an indication as to how the PDE-model, Equations 7–9, might be related to the class of models where a correlation between nonlinearity, scale, and sensitivity exists. However, an important prerequisite for a strong correlation is fulfilled; relative permeability is known to be a slowly varying function of saturation. With a slowly varying initial function, an iteration sequence of slowly varying functions, $\{k_{2,n}\}$, where $k_{2,n} \to k_{2,true}$, is a possibility.

5 RESULTS

Through numerical experiments, we investigate into the possible influence of the correlation considered in Section 2 on the estimation efficiency of the quasi-Newton and Levenberg-Marquardt optimizers. The selection of results presented in this section are concerned with both single-phase flow (Section 5.1), and two-phase flow (Section 5.2). The selection is made to illustrate some general trends seen in (Brusdal and Mannseth, 1999; Brusdal et al., 1999; Nævdal et al., 1999).

In all cases, synthetic observation data with an added error vector drawn from a Gaussian distribution, have been applied. For the cases shown here, the standard deviations of the Gaussian distribution have been kept smaller than those that can be anticipated in practice. Nevertheless, use of an added random error allows for the application of statistically based solution criteria (see, e.g., (Grimstad et al. 1997)) for the inverse problem. This concerns the closeness of the final value of the objective function to its expected value.

5.1 Single-phase flow model: Estimation of permeability

Results with the single-phase flow model (see, Section 4.1) are presented. The permeability is assumed to be a piecewise constant function of the spatial coordinate. Hence, first-order spline-wavelets (i.e., the Haar basis) have been applied to represent $a(x)$. The number of parameters needed to represent the true permeability is 32, i.e., $a_{true}(x) \in V_5$. Equidistantly distributed data for $p(x)$ have been applied.

5.1.1 Quasi-Newton optimizer In this section, some results from applying norm rescaling to the BFGS quasi-Newton optimizer are shown. Relevant details concerning the setup of the numerical experiments are found in (Brusdal and Mannseth, 1999).

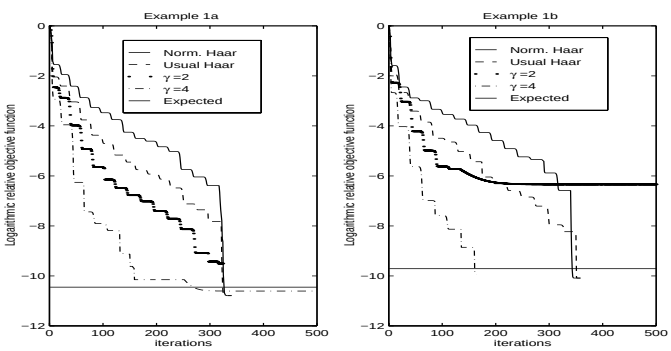

Figure 3. Logarithmic relative objective functions for problem with small fine-scale variation in $a_{true}(x)$ (left), and for problem with large fine-scale variation in $a_{true}(x)$ (right).

Figure 3 shows results from one experiment with small fine-scale variation in $a_{true}(x)$, and one experiment with large fine-scale variation. The horizontal line on the plots in the right column corresponds to the expected value of the logarithmic relative objective function. The label 'Norm. Haar' corresponds to use of the normalized Haar basis (i.e., $\gamma = 1$), while the label 'Usual Haar' corresponds to $\gamma = \sqrt{2}$. Some benefit from enhancing large-scale sensitivities (i.e., selecting $\gamma > 1$) is seen in both cases. However, as indicated by the results for $\gamma = 2$ to the right on Figure 3, stable improvement with $\gamma > 1$ was not always the case for the quasi-Newton algorithm. For details and more results, see (Brusdal and Mannseth, 1999).

5.1.2 Levenberg-Marquardt variants In this section, some results from estimation with three variants of the Levenberg-Marquardt optimizer are shown. In essence, the variants, labeled (S), (T), and (TS), correspond to different forms of scale-discriminate estimation, see also Section 3.3. More details, including a motivation for each of the variants, are found in (Brusdal et al., 1999; Nævdal et al., 1999). Details on the setup of the numerical experiments are found in (Brusdal et al., 1999).

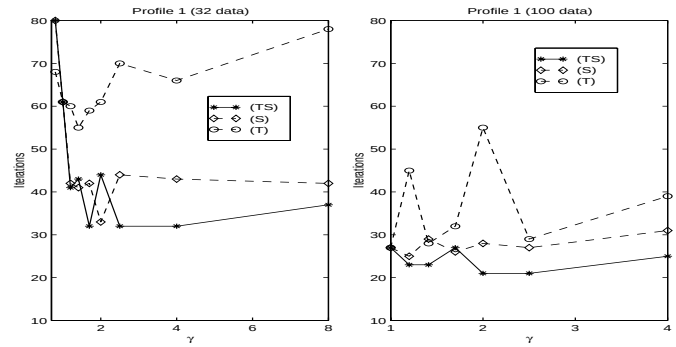

Figure 4. The number of iterations versus γ for 32 data (left), and for 100 data (right).

The Levenberg-Marquardt variants were always able to bring the objective function down to the desired level, and for small data errors there were no visible differences between the true and estimated $a(x)$. Figure 4 shows the number of iterations needed to reach the desired level for the objective function versus the value of γ, for 32 data points (i.e., equal to the number of parameters) and for 100 data points. A benefit from enhancing large-scale sensitivities is seen for (S) and (TS) with 32 data points, but this benefit almost vanish as the number of data points is increased to 100. For details and more results, see (Brusdal et al., 1999).

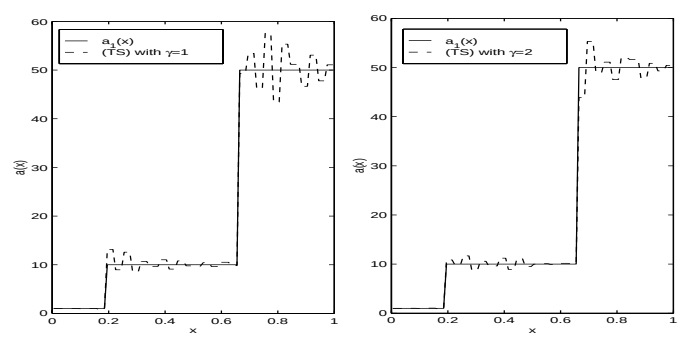

Figure 5. The estimated $a(x)$ closest to $a_{true}(x)$ obtained with $\gamma = 1$ (after 67 iterations) and $\gamma = 2$ (after 26 iterations).

For larger data errors visible differences between the true and estimated $a(x)$ are found. For such data errors, the minimum error in the estimated function is not found at the final value of the objective function. Figure 5 shows the best estimates (in this sense) with (TS) for $\gamma = 1$ (after 67 iterations) and $\gamma = 2$ (after 26 iterations). However, the estimated $a(x)$ with $\gamma = 1$ and $\gamma = 2$ at the final value of the objective function were not visually different, but the difference in the number of iterations persisted.

5.2 Two-phase flow model: Estimation of relative permeability

Results with the variants of the Levenberg-Marquardt optimizer applied to the two-phase flow model (see, Section 4.2) are presented. The relative permeability is a smooth function of saturation. Third-order spline-wavelets have been applied to represent $k_2(S_2)$. The data are time-series of produced volume of phase 2, and of applied pressure drop across the core sample. Details on the setup of the numerical experiments are found in (Nævdal et al., 1999).

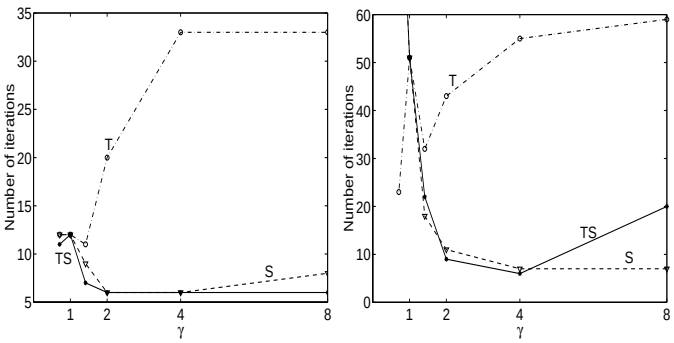

Figure 6. Number of iterations versus γ for a good initial value (left) and for a bad initial value (right)

Figure 6 shows the number of iterations needed to reach the desired level for the objective function versus the value of γ. The left plot corresponds to a better initial value than the right. It is seen that (S) and (TS) benefit from selecting $\gamma > 1$, especially if the initial value is not very good. As for the single-phase flow model, there was no gain in using the (T)-variant.

The influence that the choice of γ has on the scale structure of the initial steps taken by the optimizer is illustrated in Figure 7. The next few steps have similar structures. Hence, selecting $\gamma > 1$, ensures that the first few steps are dominated by large-scale components. A similar scale structure were found for the steps when estimating the permeability in Section 5.1, see, (Brusdal et al., 1999).

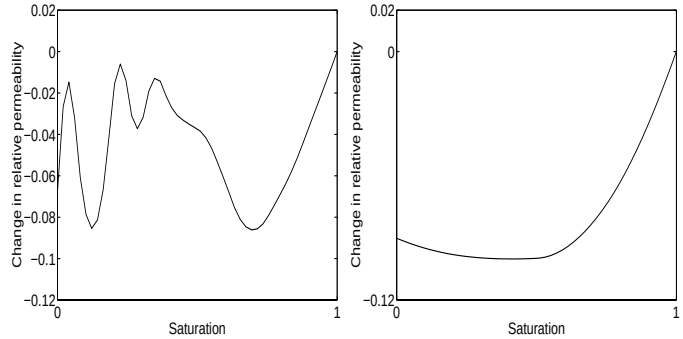

Figure 7. Changes in the oil relative permeability curve corresponding to the first step with (TS), for $\gamma = 1$ (left) and $\gamma = 4$ (right).

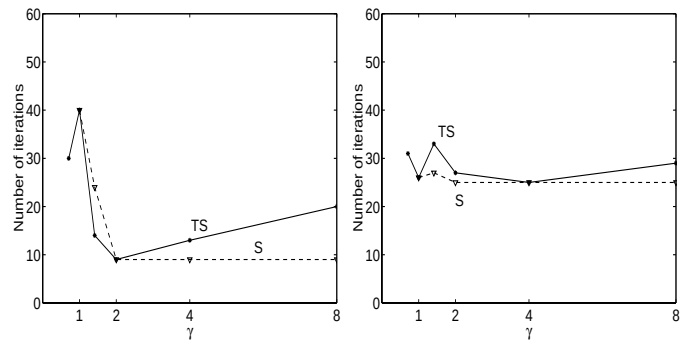

Figure 8. Number of iterations versus γ with standard estimation (left), and with hierarchical estimation (right).

Figure 8 shows results from applying standard and hierarchical estimation to the same case, with the (S) and (TS) variants. It is seen that the better performances are obtained with a good choice of γ and standard estimation. Hierarchical estimation seems more indifferent to the choice of γ. However, hierarchical estimation perform better than standard Levenberg-Marquardt (i.e., corresponding to $\gamma = 1$ on the left plot in Figure 8) for all values of γ. This is largely the impression also when estimating the permeability (Brusdal et al., 1999). Thus, if convergence speed with an a priori number of identifiable parameters was the only issue, the better choice seems to be scale-discriminate standard estimation with a proper choice of γ. However, in practice, the number of identifiable parameters is often not known a priori, making hierarchical estimation an attractive alternative.

6 SUMMARY AND CONCLUSIONS

The focus of this paper has been on the possible influence of a correlation between nonlinearity, scale, and sensitivity,

on estimation efficiency for the quasi-Newton and Levenberg-Marquardt algorithms.

The correlation was demonstrated for a class of integrated nonlinear models of a specific form. The correlation was found to be strong at points in parameter space corresponding to slowly varying functions, and less evident for other points.

Multi-scale spline-wavelet bases were used to represent the function to be estimated, to facilitate various forms of scale-discriminate estimation strategies. These strategies aim at enhancement of sensitivities to parameters associated with low nonlinearity. The correlation guides us to enhance sensitivities to parameters associated with large-scale basis elements.

Two estimation problems, stemming from porous-media flow, was considered: (1) Estimation of a spatially dependent, piecewise constant function (the permeability) in a single-phase flow ODE-model. (2) Estimation of a state dependent, smooth function (a relative permeability) in a two-phase flow, coupled PDE-model.

Both model problems were found to be related to the class of models for which the correlation has been demonstrated. Problem (1) has been shown to belong to this model class, but the permeability is not always slowly varying, making the strength of the correlation uncertain. The relative permeability is slowly varying, but due to the complexity of the two-phase model, problem (2) can not be shown to belong to the model class, by analytical means.

Numerical results indicated that there is a benefit of applying scale-discriminate estimation, although the effect was somewhat problem dependent. Both scale-discriminate standard estimation, and hierarchical estimation, usually performed better than ordinary estimation (i.e., straightforward quasi-Newton and Levenberg-Marquardt). The best scale-discriminate standard estimation performed better than hierarchical estimation. However, in practice, the number of identifiable parameters is often not known a priori, making hierarchical estimation an attractive alternative.

REFERENCES

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

Bates, D. M. and Watts, D. G., Relative curvature measures of nonlinearity, Journal of the Royal Statistical Society B, 42 No. 1 (1980) 1–25.

Brusdal, K., and Mannseth, T., Effects of basis-element norm rescaling on the convergence speed for a nonlinear parameter estimation problem, Submitted (1999).

Brusdal, K., Mannseth, T., and Nævdal, G., Modified Levenberg Marquardt algorithms for a nonlinear parameter estimation problem, Submitted (1999).

Chavent, G., and Liu, J., Multiscale parametrization for the estimation of a diffusion coefficient in elliptic and parabolic problems, in *Fifth IFAC Symposium on Control of Distributed Parameter Systems* Perpignan, France, June (1989).

Chui, C. K., *An introduction to wavelets*, Academic Press, San Diego, California, (1992).

Chui, C. K., and Quak, E., Wavelets on a bounded interval, in *Numerical methods of Approximation Theory, Vol. 9* eds. Braess, D., and Schumaker, L. L., vol. 105 of International Series of Mathematics, Birkhäuser Verlag, Basel, (1992) pp. 53–75.

Daubechies I., *Ten lectures on wavelets*, SIAM, Philadelphia, Pennsylvania, (1992).

Gill, P. E., Murray, W., and Wright, M. H., *Practical optimization*, Academic Press, New York, (1981).

Grimstad, A. A., Kolltveit, K., Nordtvedt J. E., Watson, A. T., Mannseth, T., and Sylte, A., The uniqueness and accuracy of porous media multiphase properties estimated from displacement experiments, in *Reviewed Proceedings of the Society of Core Analysts 1997 International Symposium*, Calgary (1997).

Grimstad, A. A., and Mannseth, T., Nonlinearity, scale, and sensitivity for parameter estimation problems, Submitted, (1999).

Kulkarni, R., Watson, A.T., Nordtvedt, and J.E., Sylte, A., Two-Phase Flow in Porous Media: Property Identification and Model Validation, AIChE Journal, 44, No. 11 (1998) 2337-2350.

Liu, J., A multiresolution method for distributed parameter estimation, SIAM J. Sci. Comput. 14 (1993) 175–182.

Nævdal, G., Mannseth, T., Brusdal, K., Nordtvedt, J. E. , Multiscale estimation with spline-wavelets: The case of relative permeability, Submitted (1999).

Richmond, P. C., and Watson, A. T., Estimation of multiphase flow functions from displacement experiments, SPE Reservoir Engineering (1990) 121–127.

Watson, A.T., Richmond, P.C., Kerig, P.D., and Tao, T.M., A Regression-Based Method for Estimating Relative Permeabilites from Displacement Experiments, SPE Reservoir Engineering 3, No. 3, (1988) 953-958.