Parameter quality conditions in open-channel inverse problems

Conditions sur la Qualité des paramètres dans les problèmes inverses dans le cas des écoulement à surface libre

RAHMAN H. KHATIBI, *BSc, MSc, PhD, CEng, MCIWEM, Consultant mathematical modeller, Academic Visitor, University of Oxford, Department of Engineering Science, Park Road, Oxford, OX13PJ, UK; Email: rahman@khatibi.freeserve.co.uk; formerly Ph.D. Student at Queen Mary and Westfield College, London University*

PETER R. WORMLEATON, *BSc, PhD, CEng, MICE, Senior Lecturer, Queen Mary and Westfield College, London University, Mile End Road, London E1 4NS. Tel 0044 (0)171 975 5304, Fax 0044 (0)171 9831007; Email p.r.wormleaton@qmw.ac.uk*

JOHN J. R. WILLIAMS, *BSc, PhD, CEng, MICE, Reader, Queen Mary and Westfield College, London University, Mile End Road, London E1 4NS. Tel 0044 (0)171 975 5306, Fax 0044 (0)171 9831007; Email j.j.r.williams@qmw.ac.uk*

ABSTRACT
Some open-channel empirical parameters lack exact values due to (a) not having definitive measurement methods and/or (b) not rendering a "properly-posed" system of equations in the formulation of their inverse problems. The values of these parameters are therefore susceptible to data-errors, imperfections in governing equations or insufficiency of the gauged data used in inverse problems. Best values of such parameters in a mathematical sense can be identified by the implementation of inverse problems using optimisation methods but there are potential pitfalls. A comprehensive review of inverse problems is presented in this paper outlining some of the pitfalls in their implementation and illustrating the parameter quality conditions of "identifiability, uniqueness and stability". The uniqueness condition was approached by statistical methods in this paper. Other considerations of equal importance were found to include the compliance with the assumptions underlying optimisation methods and a careful selection of the objective function.
Key words: improperly-posed, properly-posed, empirical parameters, inverse problems, quality, assumptions, statistics

RÉSUMÉ
Dans les problèmes d'écoulements à surface libre, quelques paramètres empiriques font défaut a) parce qu'ils ne peuvent être mesurés, et/ou b) parce que leurs valeurs ne peuvent être extraites d'un système d'équations proprement posées dans la formulation du problème inverse associé. Les valeurs de ces paramètres sont par conséquent sujettes à des erreurs de données, des imperfections dans les équations qui les régissent ou des manques de données de calage dans les problèmes inverses. Les meilleures valeurs de tels paramètres, au sens mathématique du terme, peuvent être identifiées par la mise en œuvre de problèmes inverses utilisant des méthodes d'optimisations mais il existe alors des pièges potentiels. Cet article présente une large revue des problèmes inverses en soulignant quelques uns des pièges possibles dans leur mise en œuvre et en éclaircissant les conditions d' "identifiabilité, unicité et stabilité". La condition d'unicité sont abordées ici par des méthodes statistiques. D'autres considérations d'égale importance ont été trouvées pour assurer la conformité avec les hypothèses sous-tendant les méthodes d'optimisation et une soigneuse sélection de la fonction objective.

1 Introduction

Open-channel friction parameters lack definitive measurement methods. These parameters describe particular systems and their values are an *a priori* requirement for the implementation of simulation problems based on the Saint-Venant equations. Given the values of these parameters, the governing equations are generally "properly-posed" in a mathematical sense. However, "improperly-posed" equations are normally rendered if they are formulated to ascertain the values of the parameters. The formulation of flow-state equations to this end is often referred to under the generic name of inverse problems. Inverse problems are an integral part of simulation problems but the insight into their complexity is comparatively poor and the versatility of modelling tools is often rudimentary. This may largely be attributed to the mathematical complexity of improperly-posed inverse problems and also, to some extent, to scarcity of appropriate data. As a result there is a considerable scope for improvement of inverse problems, even though such a need has not been widely reflected in recent research publications.

Practising hydraulic engineers are largely familiar with calibration methods that are in fact an inverse problem technique. Implementation of such a technique is often based on subjective visual comparison of simulated and gauged values. In fact calibration for roughness coefficient in open-channels has some two centuries of history, for which Leliavsky (1965) presents a comprehensive account of the evolution on the subject of open-channel friction. More reliable objective techniques are available, which suit the improperly-posed nature of inverse problems. These techniques employ optimisation methods and are based on automatic algorithms seeking to minimise a prescribed error criterion. The level of mathematical complexity of these methods is often too advanced for practising hydraulic engineers. The focus of this paper is to explain the underlying issues and help to stimulate research in this field. The strength and weaknesses of objective automatic methods are illustrated through a range of test cases. The subject will be presented in a comprehensive manner, where appropriate.

Revision received July, 1999. Open for discussion till June 30, 2001.

2 Overview of inverse problems

Steady and unsteady flow-state equations describe open-channel hydrodynamics through a number of complex equations in terms of discharge and stage, a host of geometric and hydrometric parameters and a number of empirical parameters specifying the system particulars, see (9a-9b), Appendix I. There are no known direct methods for the evaluation of the values of some of these empirical parameters and therefore flow-state equations are used instead as outlined below.

2.1 *Open-channel Empirical Parameters*

For a complete mathematical description of open-channel flow systems, their hydraulic behaviour can be broken down into boundary, locally-governed and spatially-distributed processes. Each of these processes is governed by appropriate equations derived using conservation of mass and/or momentum/energy principles. Equations derived purely using laws of nature are referred to as theoretical equations and are preferred where possible. A purely theoretical framework for open-channel simulations based on the Navier-Stokes equations is possible but they are not particularly useful for practical problems since they require extensive amount of data and involve complex algorithms and enormous amount of computing times. Empirical or semi-empirical formulae, embedded with empirical parameters, are normally preferred for simplifying or replacing complex mathematical descriptions. In “mathematical modelling”, theoretical and empirical approaches are combined by replacing complex theoretical equations with empirical approximations for simplicity. While mathematical modelling has created a proactive problem-solving capability, this has come about at the expense of using empirical equations and thereby creating the category of inverse problems.

In open-channels, spatially-distributed effects are modelled by the Saint-Venant equations. Their continuity equation establishes a balance among storage effects comprising wedge, prism and rate of rise terms and likewise their momentum equation establishes a balance among dynamic effects comprising inertial, pressure, gravitational and frictional terms. The momentum equation employs the Manning formula or a similar one for an empirical description of friction effects, which is embedded by a roughness coefficient, $\boldsymbol{P_f}$ (denoting friction parameters such as Manning *n*). This is the only parameter of interest in this paper but other empirical parameters can equally be necessary due to: (a) local effects in connection with energy losses caused by hydraulic control structures or by afflux due to constrictions; (b) parameters describing boundary hydrometric effects. The numerical schemes employed in this study for the solution of the saint-Venant equations are outlined in Appendix I.

2.2 *Review of Ascertainment of Open-channel Parameters*

The methods of ascertainment of the values of empirical friction parameters in open-channels have been reviewed by Khatibi *et al*, (1998) outlining the following methods: (a) selection of the values based on published guidelines; (b) direct methods by using steady state equations, which will also be used below to explain the improperly-posed nature of open-channel friction parameters; (c) calibration methods; (d) identification methods. Both calibration and identification methods are “referencing techniques”, whereby the values of the empirical parameters are determined by a comparison of gauged values against their corresponding simulated ones. These two approaches, however, differ in that calibration methods using visual comparisons are often subjective, whilst identification methods using optimisation methods are objective. This distinction is not a matter of semantics but it is necessary in the same way as it is for distinguishing between “average” and “mean”.

2.3 *Improperly-posed Nature of the Problem*

The direct use of the Saint-Venant equations is ideal for the ascertainment of empirical parameters but there are problems associated with this, which are best explained by using the steady state equation:

$$\frac{\beta Q^2}{A^2}\frac{\partial A}{\partial x} + gA\frac{\partial y}{\partial x} + gn^2\frac{Q|Q|}{AR^{4/3}} - gAS_0 = 0 \qquad (1a)$$

Using the numerical procedure outlined in Appendix I and depicted in Figure 1, (1*a*) can be transformed into a system of *I*-number of simultaneous difference equations, as follows:

$$F(y_i, Q_i, y_{i+1}, n_i) = a_i' y_i + b_i' Q_i + c_i' y_{i+1} - e_i' \qquad (1b)$$

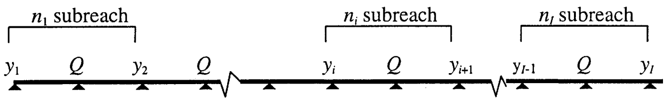

Fig. 1. The 6-point staggered cheme of a single reach channel for the formulation of inverse friction parameter.

For steady state flows, (1*b*) can be solved for the *Y*-values given the values of the boundary condition for y_1, the measured flow value of $Q = Q_1 = Q_2 = \ldots = Q_i = \ldots Q_I$ and the Manning *n* values of n_1, n_2, n_3, ... n_i for the subreaches 1, 2, 3, …*I*. In order for the subsequent system of equations to be a properly-posed problem, as described by Gustafson (1980), it is required that: (a) the solution must exist; (b) the solution must be unique by specifying boundary conditions and (c) it must be stable. These three conditions have been thoroughly investigated over the years concerning the continuous form of underlying hyperbolic partial differential equations, their transformation into difference equations and the solution of the subsequent difference equations. As a result, it can be taken for granted that these equations belong to the family of properly-posed problems. As it will be referred to later in this paper, the above conditions for properly-posed equations were given by Hadamard in 1932.

A formulation of (1*b*) for the calculation of the values of Manning *n* (*i.e.* n_1, n_2,…n_i…n_I) will render a system of *I*-number of equations, as follows:

$$p_i n_i = q_i, \text{ with } i = 1, I \qquad (1c)$$

where, p_i, q_i, are non-linear coefficients of n_i when (1*a*) is applied to the i^{th} subreach shown in Figure 1 with n_i treated as an unknown. The mathematical behaviour of (1*c*) has not been investigated. The condition for existence of a solution for these equations is that the number of equations must be equal to that of unknowns but it is not known if these equations are consistent (the uniqueness of the solution) and stable. It is useful to draw the analogy between this problem and that in groundwater flows where, the Cauchy problem is classic, rendering a properly-posed inverse problem. However, mathematicians have not produced a similar case for open-channel flows. The Cauchy problem in groundwater is often trivial and the lack of its equivalence in open-channels is not a setback.

Even if it can be shown that the formulation of (1*c*) for the calculation of the *I*-values of the Manning *n* is properly-posed, a direct measurement of *Y*-values at all the nodes will be required but this is extremely rare. One way to overcome this problem would be to reduce the node numbers and thereby reduce the requirements for the measurement of stage values but this would be at the expense of accuracy. In practice, the friction parameter is treated as a lumped parameter for the whole reach, collectively accounting for energy losses due to friction, irregular boundaries, bends and variations in channel material. If the above *I*-equations are used to identify the value of the lumped roughness coefficient, the number of equations will far exceed that of the single unknown and this alone violates the assumptions underlying the properly-posed equations. If one or more of the conditions for properly-posed equations are not satisfied, the problem is referred to as an improperly-posed problem (Emsellem, 1972), as the case is here. The above arguments were presented for steady state flows for simplicity and can be extended to unsteady flow conditions, leading to the same conclusion.

For minimising measurement requirements of inverse problems, identification/calibration problems are formulated by lumping together the parameters. However, this is at the expense of the loss of the sense for local variations in the values of the parameters. The mathematical basis of these referencing methods may conveniently be stated as follows:

$\underline{P}_f$ = vector of empirical friction parameters to be identified,

$F(x,t)$ = simulated flow-state (depth and discharge) at any point in space and time.

Simulation (flood routing or prediction) problems are expressed symbolically as:

$$F(x, t) = S(\underline{P}_f) \tag{2a}$$

Inverse problems are expressed symbolically as:

$$\underline{P}_f = I[F(x, t)] \tag{2b}$$

Where, *S*[ ] is a functional operator for simulation problems and *I*[ ] for inverse problems. For properly-posed inverse problems the operator $I = S^{-1}$ renders rigorous solutions but for improperly-posed ones normally there are no rigorous solutions and therefore the operator is transformed into curve-fitting or optimisation algorithm. As this category of mathematical problems lacks exact solutions, there is a risk of producing inconsistent and meaningless results if the mathematical formulation of these problems is not carefully handled. For the implementation of improperly-posed problems (2*b*) is modified as:

$$\underline{P}_f = I[G(x, t)] \tag{2c}$$

where, $G(x,t)$ denotes gauged data.

2.4 *Parameter Identification Procedure*

The fundamental idea underlying identification of friction parameters is: (a) discharge and stage are intertwined such that discharge attenuation is particularly sensitive to storage but can also be affected by dynamic effects; whereas stage is more sensitive to dynamic than to storage effects; (b) some of the empirical parameters embedded in boundary, locally-governed and spatially-distributed processes may interact with one another. The procedure for the formulation of inverse problems is quite different and more complex than that of simulation models. The following guidelines have a theoretical basis: (a) flow distribution networks and significant inflows into the system have to be identified and represented in the model; (b) each significant inflow represents a degree of freedom and has to be referenced by a corresponding set of gauged data that is not affected by attenuation features; (c) empirical parameters associated with energy losses have to be referenced with respect to gauged depth values; (d) each set of empirical parameters may be determined by one set of gauged data, but the use of one set of gauged data for the determination of more than one set of parameters can suffer from potential interaction problems; (e) objective methods should be preferred in inverse problems, as detailed in this paper.

The specific procedure followed in this research will be detailed in due course but the main steps are:

1. The simulation problem was implemented by assuming the geometric details of the channel, boundary conditions, roughness coefficient and space and time steps.
2. A node in the model was selected as the gauging station and the gauged values were generated by introducing Gaussian noise into the simulated depth and/or discharge values at this site.
3. Assuming the roughness coefficient is unknown, the inverse problem was implemented for identifying its values.

2.5 *Mathematical Basis of Inverse Problems*

Models by their natures are approximate representations of physical systems and therefore the results commonly suffer from undesirable effects of errors and noise. The following sources may be recognised: (a) measurement errors including gross errors and random errors; (b) "model errors" caused by imperfections of the governing equations, simplification of some of the physical processes, lack of adequate data, and simplification and schematisation of governing equations; (c) "numerical errors" associated with finite difference methods.

There is no unique way of defining errors but this has an important bearing on the formulation of objective functions. Two methods are employed in this paper to define errors; the first method uses Functional Error Residuals (FER) and the second method uses Direct Error Residuals (DER). FER is defined as:

$$FER = \phi'(Q_i^{j+1}, Y_i^{j+1}, Q_{i+1}^{j+1}, Y_{i+1}^{j+1}, \underline{P})$$
$$\Phi = \sum_{j=1}^{J} [\mathrm{FER}]^2 = \sum_{j=1}^{j} [Q_i^{j+1}, Y_i^{j+1}, Q_{i+1}^{j+1}, Y_{i+1}^{j+1}, \underline{P}]^2 \quad (3)$$

where, = ϕ' is a function of stage/discharge values and the parameter $\underline{P}$ is defined by (10*b*). Given the values of depth and discharge and also of the friction parameter, it is possible to evaluate ϕ'. A substitution of the true values of the above quantities in (3) renders the ϕ'-values to be zero. If the value of the friction parameter is unknown, ϕ' can serve as its functional representation. In practice the true values of depth and discharge are generally not available and therefore a combination of gauged and simulated values, often suffering from errors, may be used instead. Owing to the presence of data-errors, the ϕ' values are unlikely to be zero and therefore ϕ' can be regarded as a functional measure of data-errors and of the values of the parameters.

The objective functions formulated by DER are widely used, which reference gauged and simulated values by their direct comparisons. Sum of squares of error criteria are additive methods and those reported in this paper are:

$$\Phi = \sum_{j=1}^{J} (y_{g,j} - y_{s,j})^{\beta} \quad \text{with} \quad \beta = 2 \quad (4a)$$

$$\Phi = \sum_{j=1}^{J} (y_{g,j} - y_{s,j})^{\beta} \quad \text{with} \quad \beta = 2 \quad (4b)$$

$$\Phi = \sum_{j=1}^{J} \left(\frac{y_{g,j} - y_{s,j}}{y_{s,j}}\right)^{\beta} \quad \text{with} \quad \beta = 2 \quad (4c)$$

2.6 *Implementation of Optimisation Methods*

Each of the above objective functions can be used to uncover the values of system parameters by their minimisation. This could be carried out by a standard non-linear least squares technique and the method used in this study was based on a modified Gauss-Newton method similar to that of Powell (1965) and employed a scheme given by Marquardt (1964). These techniques are essentially gradient methods and their applications are well established with the added advantage that assessment of their statistical properties becomes possible. Application of least squares methods must comply with the following assumptions: (i) the error vector (i.e. Φ in (4a-4c) but with $\beta = 1$) has zero mean and constant variance, (ii) the errors are mutually uncorrelated, and (iii) the error distribution in a statistical sense is normal, Clarke, (1973). Figure 2 outlines iterative optimisation procedures, showing that parameter identification programs have two components: (i) model simulation and (ii) parameter correction through optimisation procedures. The program starts with an initial estimate of the parameters and performs a complete simulation run. The objective function is then evaluated according to the selected error criterion, which normally involves observed and simulated data. If the value of the function is above a prescribed tolerance value, the process is continued iteratively through a full scale simulations and then computing a correction to the parameters by using the optimisation algorithm.

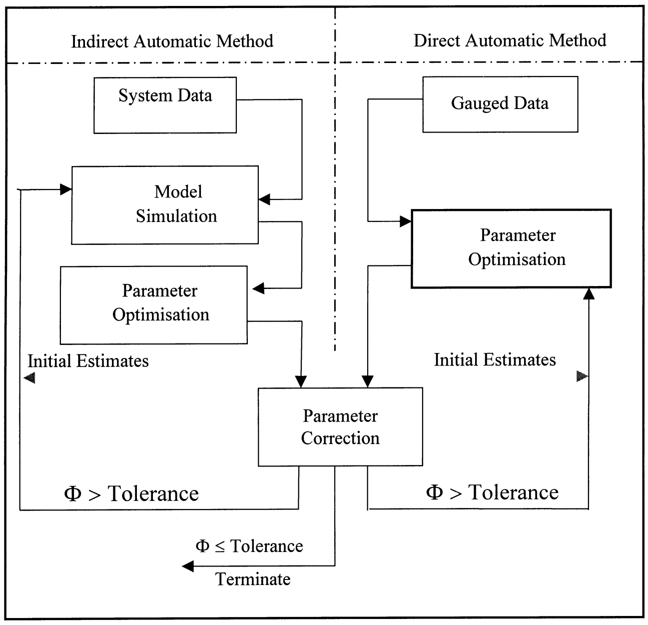

Fig. 2. Implementation of direct and iterative optimisation methods.

3 Problems underlying objective automatic methods

Although calibration of roughness coefficients in open-channel flows is now some two centuries old, objective approaches in this field have not taken off the ground yet. Allison (1979) explains the core of the problem in a different context stating that "*the respect for Hadamard was so great that incorrectly posed problems were considered 'taboo' for generations of mathematicians. Until relatively recently, it became clear that there are a number of quite meaningful problems, the so called 'inverse problems …'.*" Optimisation methods have been recognised as the tools for producing meaningful results from improperly-posed inverse problems. Whilst there are potential improvements in the application of optimisation methods in terms of efficiency and objectivity, clearly there are potential pitfalls concerning their applications, as detailed below.

3.1 *Induced errors*

Gauged data from different historic events often leads to different values of identified parameters and yet there is no indication as to which identified value may be the best. These differences are related to errors induced in the parameter values, which are caused by the various sources of error normally associated with mathematical modelling. It may intuitively be assumed that normally distributed data-errors induce normally distributed errors in the identified parameters. Evidence for this will be presented through a simplified numerical test. The induced errors in the parameters are then studied in terms of their statistics descriptors of expected values, confidence intervals and dispersion indices. Throughout this study, a confidence level of 95% was used to work out confidence intervals as defined by Student's *t*-distribution. Expected values and confidence intervals are convenient for studying the performance of a given objective function. For the expected values behaving in a single consistent manner, the response of a test case to different parameters may be studied by the dispersion index (defined as the confidence interval divided by the true value of the parameters). In general, the smaller the index, the narrower the confidence interval and the better the quality of the identified parameters.

3.2 *Quality Conditions*

The formal definitions of identifiability, uniqueness and stability conditions are complex but, before their presentation below, improperly-posed problems need to be explained in plain language. These three conditions interrelate three "spaces": (**i**) the simulation or flow-state space, $F(x,t)$, (**ii**) the parameter space, $\boldsymbol{P}$, and (**iii**) the space of gauged values, $G(x,t)$. Identifiability associates simulation spaces with parameter spaces in a forward direction. In an identifiable inverse problem, the system particulars described by the parameters must make a significant contribution to the make up of the flow-state. Conversely, the system particulars described by different values of the parameters are descriptions of different systems. Uniqueness associates the parameter space with the space of gauged values in a backward direction. As a result, one or more sets of gauged values must facilitate the identification of only one set of parameter values. Conversely, the system particulars described by a set of parameters must result in an identical behaviour from one gauged event to another. Stability is concerned with the effects of data-errors in the gauged values on the identified parameter values. Small errors in the gauged values must not lead to large differences in the values of the identified parameters. These are depicted in Figure 3 and are detailed below.

Identifiability: This refers to the relationship between simulated flow-states, $F(x,t)$, and the values of the identified parameters $\boldsymbol{P}$ used in their simulation. Carrera and Neuman (1986) reported a definition of identifiability based on the metric norm of two flow-state solutions $F(x, t) = S(\underline{P})$ and $F'(x, t) = S(\underline{P}')$ such that:

$$||F(x, t) - F'(x, t)|| = 0 \Leftrightarrow ||\underline{P} - \underline{P}'|| = 0 \quad (5a)$$

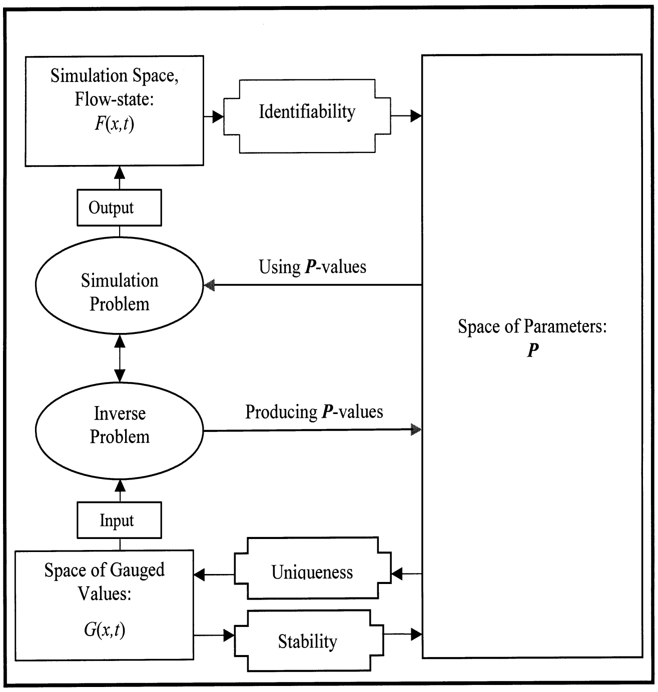

Fig. 3. Illustration of parameter quality conditions in inverse problems.

Equation (5*a*) expresses that when the differences in the metric norm of two flow-states in two systems approach zero for every other flow condition being identical, then they must be originating from the same values of the parameters. In non-identifiable problems, large changes in parametric values would be resulted from minor changes on flow-states. The above definition is often implemented through carrying out sensitivity tests to establish a feel for the identifiability of the problems. The identifiability condition can play a useful role in implementation of open-channel inverse problems. The prominence of the constituent terms in governing equations or the various hydraulic effects can change according to flow-states and also according to interactions of flow-states with system configurations. The terms containing empirical parameters can therefore be important in one case but not in others. The identifiability condition can be used to unravel such potential problems. For example, under high tidal flow conditions, inertia is expected to dominate over friction, although friction and other effects are not completely suppressed.

Uniqueness: This refers to the relationship between the values of identified parameters $\boldsymbol{P}$ and gauged values $G(x,t)$ used for their identification. According to Carrera and Neuman, if the identified parameters of an inverse problem are unique, the differences in the metric norm of two solutions $\underline{P} = I[G(x, t)]$ and $\underline{P}' = I[G'(x, t)]$ tend to zero given that the differences in the metric norm of their gauged values also tend to zero as follows:

$$\text{for } ||\underline{P} - \underline{P}'|| = 0 \Leftrightarrow ||G(x, t) - G'(x, t)|| = 0 \quad (5b)$$

The coverage of uniqueness by Carrera and Neuman is confined to the consideration of initial values of the parameters as inputs to the optimisation methods. This is particularly important if there are local minima. In this paper, it is argued that the uniqueness condition should also be posed in a statistical sense as follows. Since optimum values inevitably suffer from induced errors, it is highly likely that the optimum values obtained from one set of gauged data will differ from that obtained by another set. In this sense the identified parameters may not seem to be unique and therefore statistical analyses are required to ensure their uniqueness.

The uniqueness condition in this paper is treated statistically through the following procedure. (a) The consistency of the behaviour of the identified parameters is shown in terms of their mean. (b) Confidence interval for each value of the mean is calculated and the variation of its width with noise level is investigated. (c) It is expected that the true value of the parameter to be contained within the confidence intervals, else the identification problem suffers from pitfalls. The true value is not generally known when using field data but this is possible when the test data is generated numerically as detailed later.

Stability: This condition states that small errors in the gauged values used for identification must not produce large changes in the values of the identified parameters. This is expressed mathematically as the case where for every small value of δ there exists another small value of ε such that for $\underline{P} = I[G(x, t)]$ and $\underline{P}' = I[G'(x, t)]$, the following holds:

$$||G(x,t) - G'(x,t)|| < \delta \Rightarrow ||\underline{P} - \underline{P}'|| < \varepsilon \qquad (5c)$$

Instability arises in problems where the error criterion being minimised is too flat near the optimum values. The optimum region can then be undermined by data-errors or noise. If a simulation run, performed towards investigating the identifiability condition, reveals this characteristic problem (*i.e.* large changes in the parameters produce minor changes in flow-states), instability should be expected in the identification process. Under unstable conditions the parameters are excessively sensitive to data-errors.

3.3 *Compliance with the Assumptions Underlying Optimisation Methods*

Bias: The error vector, defined by (4a-4c) with β = 1, is required to have a zero mean, else bias will be introduced into the identified parameters. Although the mean of the normally distributed data-errors in the test runs to be presented later approximated to zero, this was not sufficient to ensure that the mean of the error vectors also approximated to zero. Objective functions can give way to the introduction of bias, as reported by, Khatibi et al, (1998). The reported test cases, as to be confirmed here by a different set of results, showed that the objective function formulated by (4a) was free of bias but those using (4b-4c) gave way to the introduction of inadvertent bias in the identified parameters. For open-channel databases, which are subjected to normally distributed data-errors in gauged data, it was shown in the above reference (and numerically will be shown here) that:

$$\lambda_a = \frac{\sum N_j Y_{t,j}^2}{y_{t,j}^2} \qquad (6a)$$

$$\lambda_g = \frac{\sum \frac{1}{N}}{\sum \frac{1}{N^2}} \qquad (6b)$$

$$\lambda_s = \frac{\sum N}{\sum N^2} \qquad (6c)$$

where N is defined later by (8a), λ is a measure of bias with λ_a corresponding to (4a) and approximating to one; λ_g corresponding to (4b) and tending to less than one; λ_s corresponding to (4c) and tending to greater than one.

Autocorrelation:- Autocorrelated error residuals may arise from gauged data used for identification purposes that suffer from periodicity. Their inverse problems may then be implemented by using the maximum likelihood method by minimising the variance and co-variance of error residuals. Using ordinary least squares techniques at the presence of autocorrelation, which are based on minimising variances, the quality of the identified parameters will deteriorate. This will be demonstrated numerically.

Collinearity:- Implicit in the assumptions underlying least squares techniques is that the parameters to be identified are independent of one another. When there is a relationship between them, the parameters are said to be collinear. Treating the friction parameter in a distributed manner, *i.e.* assigning different values to floodplain and main channel roughness values and varying them along the course of the river is often desirable. These parameters can, however, be correlated, in which case this will lead to ill-conditioned optimisation problems but by using a coarser schematisation of friction parameters the problem may be controlled. Similar problems may also arise due to interactions of the parameters with the gauged data used for their identification. A simple manifestation of this effect will be presented in the results.

Statistical distribution of the error vector:- When error vectors are not normally distributed, generalised least squares techniques are used, which are based on the maximum likelihood method. This is based on minimising variance and covariance of the error residuals.

4 Numerical test procedures

This paper aims to investigate the fundamental aspects of inverse problems in open-channel friction parameters. This is only possible if the exact values of the parameters are known which is not generally the case when using field data. Even then the available gauged events are often few with no quantitative information on their quality. Synthetic data were used in this study for the following advantages: (a) the true value of the parameter is known, (b) the quality of the data can be expressed numerically, (c) test conditions can be controlled and (d) there is no restrictions on the number of available gauged events – some 440 events were used in this study.

4.1 *Test Cases*

A series of test cases were devised to investigate some of the issues raised above concerning the lumped friction parameter of a single reach channel. The gauged data to be used in the identification process was generated by simulating different hydrodynamic waves through a 50 km long model river. A Manning *n* value of 0.03 was selected to describe the roughness coefficient. Each test run was then formulated by changing some of the model details to cover a range of physical and flow conditions. A representative flood wave was simulated through the channel of bed gradient of 1:2000 whilst that for the tidal waterway was 1:100000. The specific details of each test run were as follows.

Test Case 1 describes flow events through a model river of 5-m wide simulated using the 6-point staggered scheme (see Appendix I). The downstream boundary condition was specified by the Manning equation and this was made possible by using a scheme proposed by the first author, see (Khatibi 1999). The roughness coefficient for this boundary condition was taken as an integral part of the river as in natural boundary conditions. The upstream boundary inflow Q (m^3s^{-1}) at time t (hours) was generated by:

$$Q = 100 + 800\left[\frac{t}{24}\exp\left(1 - \frac{t}{24}\right)\right]^{16} \qquad (7a)$$

Test Case 2: is similar to Test Case 1 except that the downstream boundary condition was implemented as a rating relationship. This was achieved by setting the Manning *n* of the boundary condition to a constant value of 0.06.

Test Case 3 corresponds to a comparatively high tidal wave travelling upstream where at its upstream boundary the depth of flow is assumed to remain constant. The 4-point box scheme, as referenced in Appendix I, was used for the simulation runs of this test case. The results of this case are reported by Khatibi, *et al* (1999, unpublished paper), in a different context but some are reproduced here to illustrate identifiability and stability conditions. The flow in the system was simulated using the following boundary conditions:

$$y_{u/s} = 6.0 \text{ and } y_{d/s} = 6 + 5\sin\left(2\pi\frac{t}{24}\right) \qquad (7b)$$

4.2 *Generation of gauged data*

The reaches in the study models were divided into subreaches each of 5 km. The simulation was carried out using a time step, Δt, of one hour and an implicit finite difference weighting factor, θ, of 0.55. The gauged data necessary for the identification of the roughness coefficient was then obtained by running the model to predict the discharge and stage throughout the system. The true simulated values, y_t, at a selected node such as the downstream end of the reach, were contaminated with Gaussian noise, N for the emulation of noise normally contained in field data. The gauged data, y_g, was generated as

$$y_{g,j} = y_{t,j} * N_j[\mu, \sigma] \qquad (8a)$$

Each test run corresponds to a sample of data-errors selected by changing the seeding of the random number generator with a prescribed noise level, σ, and mean noise of, μ. The μ-value was assigned to 1 to suit the method of noise introduction into the parameters. A systematic study of the effect of data-errors on the identified Manning *n* was carried out by varying σ from 0.05 to 0.30 in steps of 0.05. By open-channel field data standards, noise level at $\sigma = 0.05$ may be considered low, moderate at $\sigma = 0.1$–0.15 but severe at $\sigma > 0.2$. A few individual errors greater than ±100% at $\sigma = 0.3$ were regarded as outliers and clipped to within the ±100% range, since this level of error is larger than might reasonably be expected from reliable field data. Errors may also be subjected to a first order Markovian chain at autocorrelation level, ρ, as:

$$y_{g,j} = y_{t,j} * (N_j[\mu, \sigma] + \rho N_{j-1}[\mu, \sigma]) \qquad (8b)$$

5 Results

5.1 *Statistical distribution of induced errors*

Numerical runs were formulated to test the intuitive assumption on the normal distribution of the errors induced in the identified parameters. In order to avoid time consuming test procedures, the direct optimisation procedure, as outlined in Figure 3, was implemented as follows. The hydrographic values of discharge and stage of a subreach in Test Case 1 were generated as above and were then regarded as gauged data. These values were substituted in (3), rendering it as a function of the Manning *n* of the subreach. Using optimisation methods the roughness coefficient of 55 samples were then identified at a range of noise levels. A χ^2 goodness-of-fit test of the induced errors as depicted in Figure 4 showed that they were normally distributed. A moving average of the expected values of each five sample out of 55 data-sets are shown in Figure 5 showing that, at each noise level the expected value is fairly insensitive to sample size, a size of 10 was taken as a compromise between simulation times and the reliability of the results. A detailed investigation was then carried out using Student's *t*-distribution to calculate confidence intervals by employing 10 samples for each series of test runs.

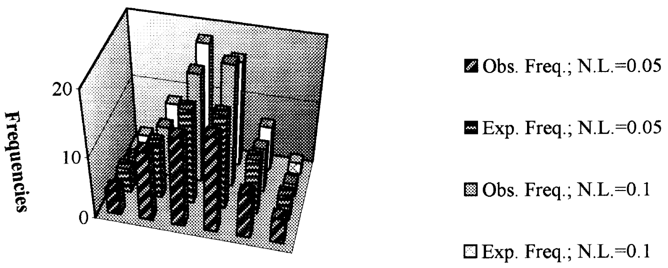

Fig. 4. Goodness-of-bit by the Chi2 test.

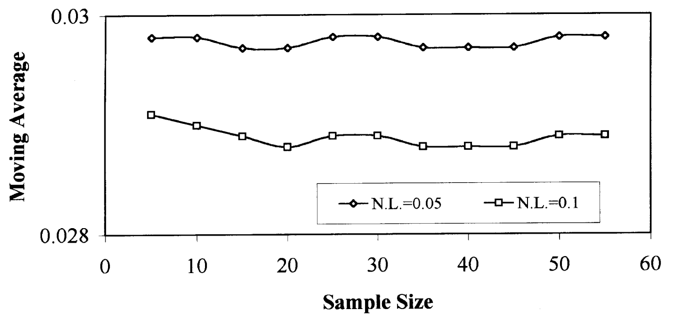

Fig. 5. Variation of moving average of identified parameters with sample size – (N.L. = Noise level).

5.2 *Illustration of Parameter Quality Conditions*

Identifiability: In tidal waterways under high tidal conditions discharge values are generally expected to be more sensitive to the variations in channel friction than water levels. The results of Test Case 3, reproduced in Figures 6a-6b, clearly indicate that large changes in the values of friction parameters only produce minor changes in the values of water levels but significant changes in discharge values. These results set the scene for : (a) expecting a non-identifiable problem when gauged water levels are used to identify the friction parameter (b) suggesting an identifiability condition if discharge values are the basis for formulating the objective function of this test case.

Uniqueness:- Table 1 presents a selection of typical results based on Test Case 1 to highlight the importance of this condition. It clearly shows that the induced errors in the identified parameter of one event differ from those of the others, where the differences not only depend on the level of data-errors, but also on the objective function selected. The results are depicted in Figure 7 and for each noise level it clearly shows that the expected values of the identified Manning *n*, identified using (3) and (4*a*-4*c*), behave consistently. Three trends are observed : (a) the expected values of the parameters identified using (3) and (4*a*) closely follow the true value; (b) those due to (4*b*) follow a clear trend of strong deviations from the true value leading to excessive underestimation of the parameter; (c) the behaviour of (4*c*) is similar to (4*b*) except that the deviation is milder and the trend is toward overestimation.

It is clear that in a statistical sense, the expected values behave uniquely owing to their consistent trends but not all of them are necessarily accurate. Figure 8 shows that the expected values of the parameters identified by (4*a*) closely follow the true value and this value is contained within the confidence intervals of the identified parameters. This is also true for (3). As depicted in Figure 9, however, the true value lies outside the confidence intervals when (4*b*) is used; this is also true for (4*c*). The selection of objective function therefore has an important bearing on the identified parameters. The contrasting behaviours are mathematically explained below.

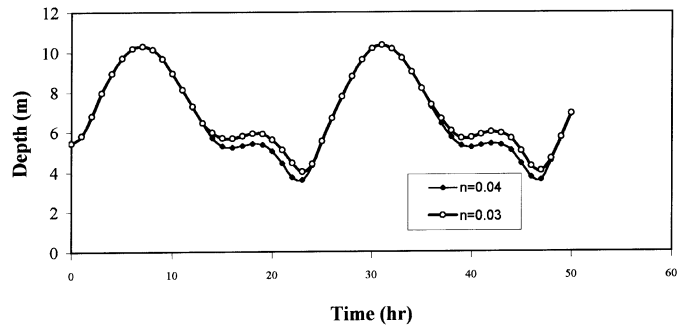

Fig. 6a. Sensitivity of simulated depth to Manning *n* – Test case 3, middle of model river.

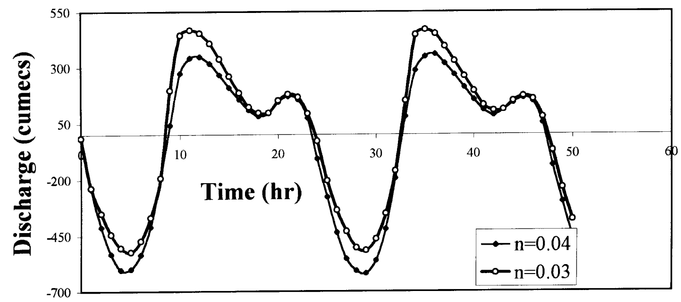

Fig. 6b. Sensitivity of simulated discharge to Manning *n* – Test case 3, middle of model river.

Table 1. Illustration of the Uniqueness Condition Based on Test case 1 Using (3) and (4a-4c) – the true value of Manning *n* = 0.03.

Sample	Error percentages[1] in the Manning *n* at noise level: σ=0.1			Error percentages[1] in the Manning *n* at noise level: σ=0.20		
	(4*a*)	(4*b*)	(4*c*)	(4*a*)	(4*b*)	(4*c*)
1	+1.7	-2.0	+3.3	+3.3	-9.7	**+10.0**
2	-0.7	-3.7	+0.0	-1.0	**-12.7**	+2.3
3	+5.0	+0.7	+5.0	+9.7	-5.3	**+13.0**
4	+3.0	-2.0	+3.3	+6.7	**-13.3**	+9.7
5	-1.7	-5.3	-0.3	-3.3	**-20.0**	+2.3
6	+2.0	+0.7	+4.3	+4.0	-4.3	**+11.0**
7	-0.1	-3.7	+1.7	-0.1	**-23.0**	+7.0
8	+2.0	-2.3	+2.3	+4.3	**-14.3**	+7.7
9	-6.7	-6.7	-2.3	**-12.7**	**-21.3**	-1.7
10	-0.1	-3.3	+1.3	-0.2	**-14.3**	+5.7

1 Error percentages in the identified Manning *n* are defined as: $\frac{n_{true} - n_{opt}}{n_{true}}$ and the values in excess of 10% are in bold.

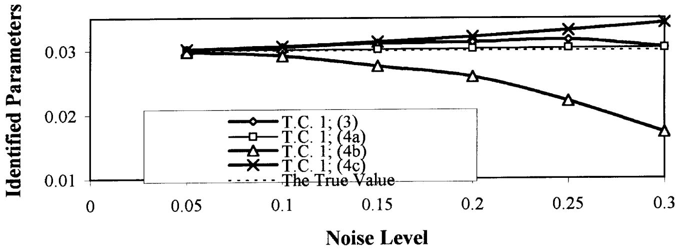

Fig. 7. Variation of mean of identified parameters with noise level – Test case 1 (T.C.1) with (3) and (4a)–(4c).

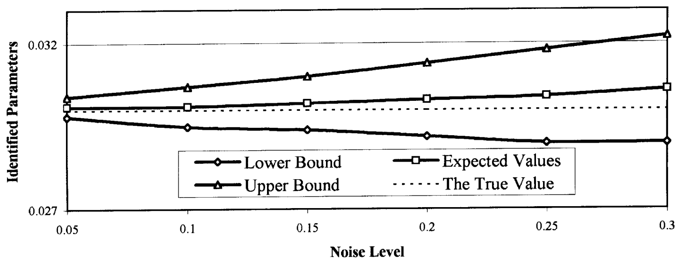

Fig. 8. Variation of mean of identified parameters and confidence intervals with noise level – Test case 1 using (4a).

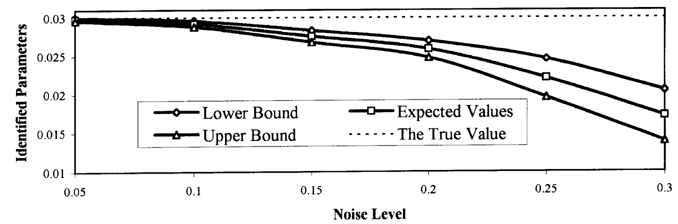

Fig. 9. Variation of mean identified parameters and confidence intervals with noise level – Test case 1 using (4b).

Table 2. Effects of Data-Errors on Identified Parameters Using Levels (4a) to Formulate the Objective Functions- Test Case 3.

SAMPLE	$\Phi = \sum (y_g - y_s)^2$			
	σ=0.001	σ=0.01	σ=0.05	σ=0.10
1	+0.4	+3.7	**+25.3**	**+124.7**
2	+0.5	+5.0	**+34.3**	**+137.3**
3	+0.3	+2.7	**+14.3**	**+33.3**
4	-0.1	-0.7	-3.3	-6.3
5	-0.6	-5.3	**-22.3**	**-35.7**
6	-0.1	-1.0	-4.7	-8.7
7	+0.5	+5.7	**+37.0**	**+84.7**
8	+0.2	+1.7	**+14.0**	**+293.7**
9	+0.2	+1.7	+8.3	**+16.0**
10	-0.3	-1.3	**+13.0**	**-22.3**

Error percentages are presented in the Table, which are calculated as $\frac{n_{opt} - n_{true}}{n_{true}}$ and the values in excess of 10% are in bold.

Stability: Optimisation of empirical parameters of an unidentifiable inverse problem will inevitably suffer from instability problems, particularly when data-errors are present. Some of the results corresponding to Test Case 3, are presented in Table 2. The hydrographic results depicted in Figures 6*a*-6*b*, led to the anticipation of instability problems. These results clearly demonstrate that the identified parameters suffer from instability even at very low noise levels, as already anticipated from the investigations for identifiability. Further details of this test case are presented by Khatibi, *et al* (2000), who report that stable inverse problems may be formulated using discharge values in the objective function.

5.3 *Illustration of Problems due to Lack of Compliance with the Assumptions*

Bias:- The bias scalars associated with (6a-6c) are presented in Table 3 corresponding to the results shown in Figure 7. These provide a numerical confirmation for the introduction of inadvertent bias into the identified parameters. It is clear that λ_a shows a random but minor oscillation whilst remaining close to one; λ_g and λ_s have a clear tendency for a value less or greater than one respectively explaining the bias.

Table 3. Sources of Introducing Bias into the Identified Parameters – Test Case 1; True Manning n = 0.03.

Samples	Noise Level: σ=0.10					
	(4*a*) $\Phi=\Sigma[y_g - y_s]^2$		(4*b*) $\Phi=\Sigma[(y_g-y_s)/y_g]^2$		(4*c*) $\Phi=\Sigma[(y_g-y_s)/y_s]^2$	
	Error[1] (%)	λ_a	Error[1] (%)	λ_g	Error[1] (%)	λ_s
1	+1.7	1.014	-2.0	0.992	+3.3	1.026
2	-4.0	0.888	-3.7	0.980	0.0	1.007
3	+5.0	1.021	+0.7	0.998	+5.0	1.031
4	+2.7	1.007	-2.0	0.981	+3.3	1.018
5	-1.7	0.983	-5.3	0.961	-0.3	0.994
6	+2.0	1.015	+0.7	0.998	+4.3	1.024
7	-0.1	1.010	-3.7	0.976	+1.7	1.013
8	+2.0	1.002	-2.3	0.987	+2.3	1.017
9	-6.7	1.007	-6.7	0.963	-2.3	0.997
10	-0.1	0.999	-3.3	0.978	+1.3	1.010

[1] Error % is defined as $\frac{n_{opt} - n_{true}}{n_{true}}$

$$\lambda_a = \frac{\sum N_j y_{t,j}^2}{\sum y_{t,j}^2} \qquad \lambda_g = \frac{\sum \frac{1}{N}}{\sum \frac{1}{N^2}} \qquad \lambda_s = \frac{\sum N}{\sum N^2}$$

Collinearity: The results of Test Cases 1-2 in relation to the implementation of their boundary conditions are presented to indirectly illustrate the potential problem of collinearity (interactions among the parameters to be optimised). The gauged water levels, $y(t)$, in the vicinity of the downstream boundary are a function of both the friction parameter of the river reach P_f and the downstream boundary parameter, P_b (*i.e.* the roughness coefficient of the Manning equation used to specify the bound-

ary condition as in Case Studies 1–2). Any credible inverse problem procedure is expected to be capable of identifying both P_f and P_b when data-errors are absent. In response to data-errors, P_f and P_b may interact and this is explained as follows.
In Test Case 1 P_f is equated to P_b and therefore the inverse problem has one degree of freedom. In Test case 2 both P_f and P_b are discriminated and therefore the problem has two degrees of freedom. The results for both test cases are reported in Figure 10 in terms of dispersion indices (see Section 3.1 for the definition), in which a downstream and middle point is selected for the gauge site. The results show that, (**i**) the effect of gauging site on the results of Test Case 1 is marginal but significant on that of Test case 2 and (**ii**) the results of Test Case 2 are more sensitive to data-errors than those of Test Case 1. Thus, these results confirm that Test Case 2 is associated with the risk of collinearity.
The results of Test Cases 1-2 reveal that if a number of hydraulic processes, each represented by its own parameters, interact with one another and affect the gauged values, their inverse problems can suffer from collinearity problems. The best method of dealing with collinearity is by representing each of the hydraulic processes in the objective function by gauged values. The use of search methods of optimisation or the more specialised technique of ridge analysis is normally recommended for the problems suffering from collinearity.

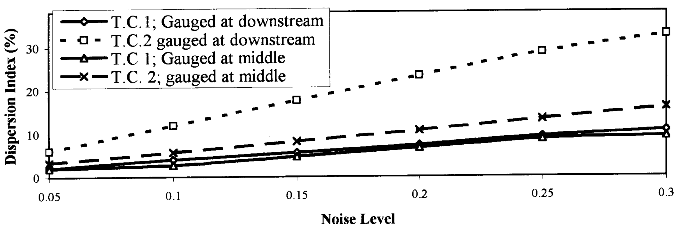

Fig. 10. Interaction of boundary parameter with channel friction – Test case 1 and 2 (T.C).

Autocorrelation: The response of the identified parameters to autocorrelated data-errors at scalar levels of $\rho = 0.0$, 0.1 and 0.5 is depicted in Figure 11. The dispersion indices became larger when higher ρ-values were selected, thus signifying the deterioration of the quality of the identified parameters. This was particularly the case for $\rho = 0.5$.

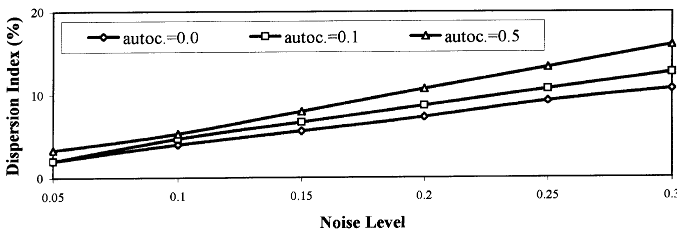

Fig. 11. Effects of autocorrelated (autoc.) gauged data on dispesion index of identified parameters.

6 General discussion

Open-channel modellers seem generally content with traditional trial-and-error calibration procedures. The capability of commercial software tools is often based on simulation problems but their identification facilities are often rudimentary. There seems little concern raised over the issue of the quality of calibrated parameters, although uncertainties are normally treated promptly through sensitivity tests. The reluctance to improve the practice may in part be attributed to the mathematical complexity of inverse problems and to the scarcity of gauged data. However, a poor level of insight into the underlying issues is perhaps a major contributory factor. The authors' work opens up the argument for a better understanding by reviewing the strength and weaknesses of open-channel inverse problems.
The simulation runs of Test Cases 1-3 were used to generate test data for the identification of friction parameters in a controlled manner. The data serving as gauged values were then contaminated to emulate measurement errors that are normally contained in field data. The use of such an approach may lack a certain amount of realism, but it does provide a flexible control over the conditions affecting the test runs. As the true values of the parameters are known, this approach is therefore powerful in facilitating definitive comparisons. The conclusions drawn from the tests must not be confined to a limited range of runs, *i.e.* if they are of value, they must reveal the underlying issues irrespective of the procedure employed. This is traditionally catered for by assessing the effects of all the significant test conditions on the conclusions and to this end, the generality of the findings of this study have been reported elsewhere (see, Khatibi *et al* (1998)). A notable exception is the better performance of the staggered scheme using (3), as opposed to the 4-point box scheme, which led to a considerable inadvertent bias when using (3) to formulate the identification problem.
The non-linear least squares technique as applied in this study is suitable for the cases, where the cross product of error-residuals in the objective function is negligible, otherwise the maximum likelihood method should be used. The statistical properties of data-errors in open-channel flows are an area where a better understanding is needed. Of particular concern is the nature of the statistical distribution of their random component and the presence of bias or autocorrelation. Without a clear understanding of these basic concepts research on the application of, say, the maximum likelihood method would lack a firm basis, which would reinforce any reluctance to improve current practices. Another consideration is the conjunctive identification of the parameters embedded in hydrological, locally-governed and spatially-distributed effects. Investigation is clearly needed to improve the efficiency of otherwise painstaking and tedious calibration practices. Test data provides an invaluable approach to proactively investigate the quality of the identified parameters with respect to collinearity, identifiability, uniqueness and stability of the identified parameters.
The results on Test Cases 1-2 are based on the use of the staggered scheme, which involves an approach by the first author for implementing head-discharge relationships as a downstream boundary condition. This scheme is notorious for its inherent difficulties on treating this type of boundary condition. The details of its implementation are presented elsewhere (see Khatibi, 1999) but it may be reported here that this scheme produced almost identical results to those of the 4-point box scheme.

7 Conclusion

The implementation of open-channel inverse problems has been presented, which shows that the underlying parameters generally lack exact values because: (a) inverse problems are improperly-posed and (b) the parameters often lack definitive measurement methods. The paper argues that the mathematical insight into the complexity of inverse problems is poor and that versatility of software tools in current practices for simulation and inverse problems are not in tandem. The mathematical complexity of the problem has been considered in this paper in terms of a number of important conditions that must be given attention if the current practice is to be improved. The specifications for modularity architecture of inverse problems should be outlined in order to provide guidelines for software developers so that versatile software tools for inverse problems are developed. The fundamental concepts underlying inverse problems are argued to be due to : (a) empirical friction parameters are embedded in the governing equations; (b) discharge and stage are intertwined such that discharge attenuation is particularly sensitive to storage and can also be sensitive to dynamic effects; whereas stage is more sensitive to dynamic than storage effects. A procedure was outlined in Section 2.4 that could enable to envisage the various significant contributions in the make up of inverse problems.

The reliability of empirical parameters in modelling practices is often overlooked and therefore the results suffer from uncertainties. The importance of parameter quality conditions should not be underestimated but a full understanding of the subject in this field awaits to be realised. Implementation of objective inverse problems alone does not guarantee reliability, as there are potential pitfalls. However, these methods facilitate a determination of the compliance with the conditions underlying optimisation methods. Without understanding the true implication of these conditions, the integrity of the model results can be undermined. This paper argued that further research and development areas concerning open-channel inverse problems should include : (a) investigating the statistical nature of data-errors; (b) investigating interactions of the various parameters involved in hydrological, locally-governed and spatially-distributed effects; (c) determining the provisions for modularity architectures of coupled inverse and simulation software tools.

Acknowledgement

The authors thank the editors and reviewers for their constructive comments.

Appendix I discretization of the saint-venant equations

Hydrodynamic simulation of free surface open-channel shallow water waves is based on the full Saint-Venant equations as given below:

Continuity equation:

$$W\frac{\partial Y}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{9a}$$

Dynamic equation

$$\frac{\partial Q}{\partial t} + (1+\beta)\frac{QW}{A}\frac{\partial A}{\partial t} + \frac{BQ^2}{A^2}\frac{\partial A}{\partial x} + gA\frac{\partial y}{\partial x} + gn^2\frac{Q[Q]}{AR^{4/3}} + gAS_0 = 0 \tag{9b}$$

Closed form unconditionally stable solutions of these equations are obtained by the application of the 6-point staggered (as referenced in detail by Khatibi, 1999) or 4-point box (as referenced in detail by Khatibi *et al*, 1998) implicit finite difference schemes. Authors' implementation of the staggered scheme is based on that given by Vreugdenhill (1973) but involves a scheme by the first author to overcome its inherent problem in treating head-discharge boundary relationships. Both the staggered and box schemes transform (9*a*-9*b*) into the following difference equations:

$$\phi(Q_i^{j+1}, Y_i^{j+1}, Q_{i+1}^{j+1}, Y_{i+1}^{j+1}) = a_i Q_i^{j+1} + b_i Y_i^{j+1} + c_i Q_{i+1}^{j+1} + d_i Y_{i+1}^{j+1} - e_i' \tag{10a}$$

$$\phi'(Q_i^{j+1}, Y_i^{j+1}, Y_{i+1}^{j+1}, Q_{i+1}^{j+1}) = a'_i Q_i^{j+1} + b'_i Y_i^{j+1} + c'_i Q_{i+1}^{j+1} + d'_i Y_{i+1}^{j+1} - e_i' \tag{10b}$$

Note that in the 6-Point staggered scheme *d* and *a'* are zero and of course the expressions for each of the coefficients are different from those of the box scheme. Case Studies 1-2 were investigated by the staggered scheme that incorporated first author's method of treating rated downstream boundary conditions.

Appendix II References

ALLISON, H. (1979), "Inverse Unstable Problems and some of their Applications", Math. Sci. 4, 9–30.

CARRERA, J. and NEUMAN S.P., (1986), "Estimation of Aquifer Parameter Under Transient and Steady State Conditions: Uniqueness, Stability and Solution Algorithms," Water Resources Research, Vol. 22, No. 2.

CLARKE, R.T., and BOWER, C.E., (1973). "A Review of Optimisation Methods Used in Hydrology with Observations on their Calibration and Use," Journ. Of Hydrology, 19.

EMSELLEM, Y., and de MARSILY, G., (1974). "An Automatic Solution for the Inverse Problem," Water Resources Research, Vol. 7, No. 4.

HADAMARD, J., (1932), "Le Probleme de Cauchy et les Equitions aux Derivees Partielles Lineaires Hyperboliques", Hermann, Paris.

GUSTAFSON, K.E., (1980). "Introduction to Partial Differential Equations and Hilbert Space Methods," John Wiley and sons, New York.

KHATIBI, R.H., WILLIAMS, J.J.R. and WORMLEATON, P.R. (1998) "Identification Problem of Open-channel Friction Parameters" J. Hydr. Div., ASCE, 123(12), 1078–1088.

KHATIBI, R.H. (1999), "Incorporation of Head-Discharge Relationships in 6-Point Staggered Schemes", J. Hydr. Div., ASCE, Vol. 125, No. 1

KHATIBI, R.H., WILLIAMS, J.J.R. and WORMLEATON P.R. (2000) "Friction Parameters for Flows in Nearly Flat Channels"; J. Hydr. Div., ASCE, Vd.126, No. 10, pp. 741-749.

LELIAVSKY, S. (1965), "River and Canal Hydraulics", Chapman and Hall.

MARQUARDT D.W., (1964). “An Algorithm for Least Squares Estimation of Non-linear Parameters,” Journ. of the Society for Industrial and Applied Mathematics, Vol. 11, No. 2.
POWELL, M.J.D., (1965). “A Method for Minimizing a Sum of Squares of Non-linear Function without Calculating Derivatives,” Computer Journal, 7, pp 303-307.
VREUGDENHILL, CB (1973) “Computational Methods for Channel Flow”, Chapter 3, of Hydraulic Research for Water Management, Proc. of Technical Meeting, 26.

Appendix III Notation

a (and similarly: *b, c, d, e, a', b', c', d'* and *e'*) non-linear coefficients in the difference Saint-Venant equations;
A cross sectional Area, m^2;
F[ ] flow-state in terms of discharge and/or water levels;
g gravitational acceleration, ms^{-2};
G[ ] Gauged flow-state in terms of discharge or water levels;
I[ ] operator for inverse problems;
n Manning *n*;
N[ ] random Number generated by a random number generator;
$\underline{P}$ (or *P*) parameter vector comprising P_f, P_b: friction, local and boundary parameters;
q a non-linear coefficient
Q flow, (Q_g = gauged, Q_s = simulated), m^3s^{-1};
S[ ] operator for simulation problems;
S slope, (S_f = friction slope defined by the Manning equation, S_0 = channel gradient);
y flow depth, (and also depths: y_g = noisy gauged, y_s = simulated, y_t = noise-free (true) simulated, $y_{u/s}$ = upstream, $y_{d/s}$ = downstream) m;
Y stage, m;
t time, sec; (t_p=time to the peak of the upstream flow hydrograph, hours);
x *x*-co-ordinate, m;
W width, m.
β momentum correction factor, or a number;
δ a small number
Δ step (e.g. Δt = time step, Δx = space step);
σ noise level, standard deviation of noises within a sample;
ε small error;
μ the mean of noise values within a sample;
θ weighting factor;
λ a ratio;
ρ autocorrelation level;
Φ objective function;

Subscripts
i spatial subscript, which varies recursively from 1 to I;
j temporal subscript, which varies recursively from 1 to J.