OPTIMAL MEASUREMENT LOCATIONS FOR PARAMETER ESTIMATION OF DISTRIBUTED PARAMETER SYSTEMS

A thesis submitted to The University of Manchester
for the degree of
Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2011

JORGE ENRIQUE ALAÑA

School of Chemical Engineering and Analytical Science

CONTENTS

ABSTRACT

"Errors using inadequate data are much less than those using no data at all"

Charles Babbage

Identifying the parameters with the largest influence on the predicted outputs of a model reveals which parameters need to be known more precisely to reduce the overall uncertainty on the model output. A large improvement of such models would result when uncertainties in the key model parameters are reduced. To achieve this, new experiments could be very helpful, especially if the measurements are taken at the spatio-temporal locations that allow estimate the parameters in an optimal way. After evaluating the methodologies available for optimal sensor location, a few observations were drawn. The method based on the Gram determinant evolution can report results not according to what should be expected. This method is strongly dependent of the sensitivity coefficients behaviour. The approach based on the maximum angle between subspaces, in some cases, produced more that one optimal solution. It was observed that this method depends on the magnitude of outputs values and report the measurement positions where the outputs reached their extrema values. The D-optimal design method produces number and locations of the optimal measurements and it depends strongly of the sensitivity coefficients, but mostly of their behaviours. In general it was observed that the measurements should be taken at the locations where the extrema values (sensitivity coefficients, POD modes and/or outputs values) are reached. Further improvements can be obtained when a reduced model of the system is employed. This is computationally less expensive and the best estimation of the parameter is obtained, even with experimental data contaminated with noise. A new approach to calculate the time coefficients belonging to an empirical approximator based on the POD-modes derived from experimental data is introduced. Additionally, an artificial neural network can be used to calculate the derivatives but only for systems without complex non-linear behaviour. The latter two approximations are very valuable and useful especially if the model of the system is unknown.

Key words: Optimal measurement locations, parameter estimation, optimal experiment design, distributed parameter system, proper orthogonal decomposition, sensitivity coefficients.

University of Manchester

PhD by Published work Candidate Declarations

Candidate Name:	Jorge Enrique Alaña
Faculty:	Engineering and Physical Sciences
Thesis title:	Optimal measurement locations for parameter estimation of distributed parameter systems.

DECLARATIONS

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

I personally undertook all the digital experimental analyses contained in the listed publications, and was responsible for authoring them. The only co-author, C. Theodoropoulos, was responsible of supervising the work. My contributions to each of the papers are as follow.

Publication 1. I proposed the methodology whereby the spatial locations of the largest absolute values of the sensitivity coefficients should be used as optimal experimental measurement placements to estimate parameters of a distributed parameter system. This consideration not only reduces the uncertainty in the parameter estimated, but cost is decreased also, which is significant, especially in situations where experimental observations are expensive.

Publication 2. I proposed an optimal sensor location based on the spatial positions of the extrema values of the empirical eigenfunctions derived by applying the Proper Orthogonal Decomposition method. Another of my contributions is to demonstrate numerically and analytical that the extrema values of the predominant sensitivity coefficients coincide at the same locations of the extrema values of the POD-modes. Additionally, a comparison of the different methods reported in the literature for optimal sensor positions and our proposed approach is introduced.

Publication 3. A new approach to determine the temporal coefficients of an empirical approximator is introduced. This method is based on the use of the empirical eigenfunctions obtained from applying the Proper Orthogonal Decomposition method. It is demonstrated that this approach is valuable when the model of the system is unknown or very complex. This application can be used straightforward by means of experimental observations.

All the work presented in this thesis has been completed whilst a member of staff at The University of Manchester.

Jorge Enrique Alaña

COPYRIGHT STATEMENT

THE AUTHOR

EDUCATION

- Master Degree in Chemical Engineering. La Universidad del Zulia, Maracaibo, Venezuela, 2003.
- Master Degree in Applied Mathematics. La Universidad del Zulia, Maracaibo, Venezuela, 2002.
- Bsc. Chemical Engineering. La Universidad del Zulia, Maracaibo, Venezuela, 1993.

RESEARCH EXPERIENCE

- Senior Lecturer and Researcher. School of Chemical Engineering, La Universidad del Zulia, Maracaibo, Venezuela, 1997–2005.

RECENT PUBLICATIONS

- Alaña, J., Theodoropoulos C. (2011), Optimal location of measurements for parameter estimation of distributed parameter systems. Computer and Chemical Engineering, 35(1), 106.
- Alaña, J. (2010), Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes. Computer and Chemical Engineering, 34(8), 1220.
- Alaña, J. (2010), Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems. Brazilian Journal of Chemical Engineering, 27(4), 627.
- Alaña, J. et al. (2010), Influence of a thermoelectric facility on V and Ni concentration in sediments in the city of Maracaibo, Venezuela. Interciencia, 35(3), 217.
- Alaña, J. et al. (2009), Diffusional analysis of acid leaching from oil fly ash. Revista Venezolana de Tecnología y Sociedad, 2, 11.
- Alaña, J. et al. (2008), Moistening factor effect on the determination of viscosities for crude oil-solvent mixtures. Rev. Téc. Ing. Univ. Zulia, 31(2), 142.

- Alaña, J. et al. (2007), Determinación y estudio de dispersión de las concentraciones de monóxido de carbono en la zona oeste de la ciudad de Maracaibo-Venezuela. Multiciencias, 7(2), 115.
- Alaña, J. et al. (2007), Nernst-Planck Model applied to the kinetics of ionic exchange of the system PO_4^{-3}-Cl^- in milk. Rev. Téc. Ing. Univ. Zulia, 30(3), 236.
- Alaña, J. et al. (2006), Kinetics of ion-exchange in raw and commercial pasteurized milk. Rev. Téc. Ing. Univ. Zulia, 29(2), 182.
- Alaña, J. et al. (2006), Regression models for the prediction of high impact polystyrene properties using operational variables. Rev. Téc. Ing. Univ. Zulia, 29(2), 144.
- Alaña, J., Vinante, C. (2005), Multistep predictive controllers based on neural network. Rev. Téc. Ing. Univ. Zulia, 28(3), 210.

INDUSTRIAL EXPERIENCE

- Process Control Asesor. Petrochemical Complex El Tablazo, Venezuela, 2002–2004.
- Process Control Engineer. Petrochemical Complex El Tablazo, Venezuela, 1993–1997.

LIST OF SUBMITTED PUBLICATIONS

1. Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems.
   Alaña, J., Brazilian Journal of Chemical Engineering, 27(4), 627(2010).
2. Optimal location of measurements for parameter estimation of distributed parameter systems.
   Alaña, J., Theodoropoulos, C., Computer and Chemical Engineering, 35(1), 106(2011).
3. Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes.
   Alaña, J., Computer and Chemical Engineering, 34(8), 1220(2010).

LIST OF PUBLICATIONS UNDER REVIEW

1. Optimal measurement locations for Parameter Estimation of Distributed Parameter Systems based on the use of Artificial Neural Networks. Alaña, J., Applied Soft Computing, (2010)
2. Optimal Spatial Sampling Scheme for Parameter Estimation of Non Linear Distributed Parameter Systems. Alaña, J., Theodoropoulos, C., Computer and Chemical Engineering, (2010).

SUMMARY OF THE SUBMITTED PUBLICATIONS

1. INTRODUCTION

"So far as the theories of mathematics are about reality, they are not certain; so far as they are certain they are not about reality"

Albert Einstein

1.1 INTRODUCTORY BACKGROUND

Modelling, as a technique of analysing diverse phenomena, emerges in several different areas of scientific research. A model, which comprises a mathematical depiction of a process under study, plays the main role in such an analysis. An ample spectrum of techniques can be used in building models. These mappings are employing to corroborate physical observations and to predict future behaviours of real systems and provide guides to a deeper understanding of practical problems.

In formulating mathematical models of dynamical systems, obtaining a high degree of predictive capability may not be the only objective pursuit in several areas of science and engineering. The model must be useful for its intended applications, and models of reduced complexity are attractive in most cases.

One of the most general and significant group of systems is that of dynamic distributed parameter system also termed spatial-temporal dynamic systems. They can be found in diverse practical engineering areas such material science, nanotechnology, medical bioengineering and micro-reactor technology. These kind of systems also emerge in several common and conventional biological and industrial processes that vary from the structure and dynamics of living cells (Slepchenko et al., 2002) to the design and analysis of chemical reacting flows (Belfiore, 2003).

The real difficulty in the design and operation of most processes rests on the appropriate dynamic modelling of the system. Even if steady state behaviour is assumed the governing equations for the phenomena are usually described by nonlinear distributed parameter systems defined on complicated domains, and it is not easy to get practical dynamic models of these kinds of processes which can be implemented without undue complications, even for linear systems.

The spatial domain of these distributed parameter systems can be either continuous or discrete. If the system is on a continuous domain, it is well modelled by partial differential equations and some representative examples include: the Navier-Stokes and transport-reaction equations for a non-reactive and reactive flow processes (Byrd et al., 2002; Kee et al., 2003), and the Fitzhugh-Nagumo model (Fitzhugh, 1961; Nagumo et al., 1962) for the study of pattern formation and chemical oscillations in reactive-diffusive systems. By contrast, discrete distributed parameter systems are generally described by stochastic differential equations where the dynamic transition is defined by

evolution rules (for example transition probabilities) and governing system equations without a closed form. Examples of such systems vary from the macroscopic scale including models for vehicle road traffic (Maeriovoet and De Moor, 2005), the epidemic of a disease in a human population (Filipe and Maule, 2004; Lara-Sagahón et al., 2006), etc., to the microscopic level including the hydrodynamics of bubble flow (Sankaranarayan et al., 1999), the kinetics of non-well mixed biochemical reacting mechanisms (Battaile and Srolovitz, 2002; Gilmer et al., 2000, Zhdanov, 2002), and the molecular modelling of protein structures (Neumaier, 1997).

From a global perspective, a logical effect of the increases of the complexity on modern systems is the fact that lumped parameter descriptions frequently become inadequate as they may not provide an adequate approximation of the process. Consequently, there exists a strongly motivated requirement for applying supplementary, sophisticated and efficient mathematical methods to development new modelling techniques. Nevertheless, depending on the circumstances, corresponding integral or mixed integro-differential systems of equations can be considered. Despite the refined formulation in the spirit of the parameter distributed systems, such models present a high quality and efficiency for simulations and control procedures.

Optimisation and control of distributed parameter systems are active and expanding research areas (Kowalewski, 2001a; Kowalewski, 2001b; Lasiecka and Triggiani, 2000; Li et al., 1995; Robinson, 2001; Uciński, 1999a). Simultaneously, progress in computational and applied mathematics, combined with the accessibility of rapidly escalating computer power, progressively expands the range of applications that can be

simulated numerically. These sophisticated techniques direct to new challenges in the field of modelling. Furthermore, the set of naturally distributed engineering systems for which estimation and control is needed has recently been expanded. As an after-effect, the distributed parameter systems have an extremely important position in system analysis and control theory, establishing a separate field of research with a plenty of publications addressed to this subject (Curtain and Zwart, 1995; Gil, 1998; Grabowski, 1999; Korbicz and Zgurowski, 1991; Kowalewski, 1991; Luo et al., 1999; Malanowski et al., 1996; Mitkowski, 1991; Omatu and Seinfeld, 1989; Sasane, 2002; Sokolowski and Zolesio, 1992; Zwart and Bontsema, 1997).

In typical applications for mathematical modelling of physical systems, two main problems can be mentioned;

- Forward problem, which consist in determining the system state when the system parameters, spatial region, time observation interval, subsidiary conditions and control variables are already identified.
- Inverse problem, also called parameter estimation, consists in determining some of the model parameters using a collected experimental data.

The inverse problem in the framework of distributed parameter systems comprises the estimation of the parameters, forcing inputs or initial and boundary conditions. The major difficulty is that frequently the estimation problem is ill-conditioned even if the forward problem is well-posed (Isakov, 1998; Sikora, 2000; Sun, 1994). The exceptional significance from solving inverse problems is evident because a simulation

model obtained from solving the forward problem could be unreliable. Consequently, more attention should be paid to this topic, which requires more effective and robust analysis techniques.

The literature associated to the subject of parameter estimation in distributed parameter systems is abundant (Banks and Kunisch, 1989; Kubrusly, 1977; Polis, 1982). The interested reader can be also referred to works (Banks, 1992; Chavent, 1991; Gibson et al., 2000; Kunisch, 1988; Uciński and Korbicz, 1990, Alaña et al., 2006, Alaña et al., 2007). However, most of contributions are focused on off-line approaches, and merely a few results concern on-line methods (Aihara, 1997; Demetriou, 2000).

An important aspect in parameter estimation and control of chemical processes is the monitoring system of concentrations and temperature profiles which warrant defining completely the mass and energy balances in a reliable and cost effective manner. However, in many practical applications the measurements of the state variables, critical for quality control and identification, are spatially constrained. This may be possibly caused directly by physical limitations of the system under consideration, but more frequently this result due the sensing instrumentation available. Even though distributed measurements over the whole spatial domain are in general not available, there usually subsists a possibility of making measurement continuously in time.

A fundamental problem towards parameter estimation of distributed parameter systems is to appropriately design the process of data acquisition from the system. This task comprises of obtaining an optimal arrangement of a limited number of measurements

along the spatial domain in such a way as to maximise the estimation efficiency of the process parameters. Generally, the dependence between the sensor placement and the system performance is not intuitive and has confusing characters. Due to the cost of obtaining experimental information, constraints related with measurements and needs of increased efficiencies of identifiers, there is a crucial requirement for developing some systematic approaches of selecting suitable sensor configurations and measurement schemes.

The motivations to study sensor placement problem stem from real-world engineering applications. One of the most fascinating ones is computer-assisted tomography which comprises of reconstructing material parameters which characterise the unreachable interior of an object under examination based on measurements made at the boundary avoiding any harm of the subject. Given that the observations must be non-invasive, the problem of proper data acquisition happens to be very difficult and the positions of the measurements electrodes are of great significance as they should offer the most informative measurements (Williams and Beck, 1995).

A different inspirational application regards optimisation of air quality monitoring networks, due the fast expansion of industry around the world, air pollution turn into a large societal problem. Safeguard and restitution of the natural environment needs a high precision of prediction and detection. To grant them, a correct calibration of models which describe the pollutant emission processes is required (Berliner et al., 2000; Sydow et al., 1997; 1998; van Loon, 1994). Generally, the variations in pollutant concentrations over a given area are described by partial differential equations of the

advection-diffusion category. Given that a number of coefficients of the equations are not assessable, accurate modelling becomes enormously complicated. Additionally to this, the monitoring stations are relatively costly and the problem of selecting a suitable observation scheme is of great practical significance mentioned in many publications (Andó et al., 1999; Müller, 1998; Nychka and Saltzman, 1998; Sturm et al., 1994; van Loon, 1995, Alaña et al., 2010, Alaña et al., 2007).

Analogous problems can be found in several other engineering areas, for example in ground-water sources management (Kovarik, 2000; Sun, 1996), in assembly measurement records for calibration on models employed in meteorology and oceanography (Bennett, 1992; Daley, 1991; Hogg, 1996; Malanotte-Rizzoli, 1996), in automated examination of hazardous environments (Korbicz et al., 1993), in prediction of radioactive pollution and emerging smart material systems (Banks et al., 1996; Lasiecka, 1998).

For numerous chemical processes, identifying the model parameters with the largest influence on the predicted outputs reveals which ones need to be known more accurately to reduce the overall uncertainty on the outputs. If the predictions are highly uncertain, a large improvement is expected when uncertainties in the key model parameters are reduced. In this case making new experiments could be very helpful; especially if measurements are made at locations where these parameters can be estimated efficiently.

Many industrial important diffusion–convection–reaction processes are naturally

described by non–linear parabolic partial differential equation systems. Examples include diffusion of heat and/or mass, rapid thermal processing, plasma reactors, crystal growth processes to name a few. The main feature of these equations is that their dominant dynamic behaviour is usually characterised by a finite (typically small) number of degrees of freedom. In the case of systems with linear spatial differential operators their eigenspectrum can be partitioned into a finite–dimensional slow one and an infinite–dimensional stable fast complement. This implies that the dynamic behaviour of such systems can be approximately described by a set of ordinary differential equation. Therefore, the standard approach to depict such systems involves the application of Galerkin's method. Here the basis used to expand the solution of the system is typically the eigenfunctions of the spatial differential operator. A set of ordinary differential equations that accurately describe the dynamics of the dominant (slow) modes of these systems is obtained (Chen and Chang, 1992; Ray, 1981; Teman, 1988).

The degree of freedom of a distributed parameter system is essentially infinite and the relevant mathematical theory is too complicated to be implemented in an industrial application. Moreover, a satisfactory mathematical theory of nonlinear distributed parameter system is still lacking. Most of the techniques reduce a distributed parameter system to a lumped parameter system with a finite number of degrees of freedom by using eigenfunctions of the system. The resulting description may be used in control or parameter estimation of the original systems with the aids of advanced mathematical techniques.

The model reduction technique discussed in this work is based on these lumping techniques, and it is known as the Proper Orthogonal Decomposition combined with the Galerkin procedure. Proper orthogonal decomposition, also known as the Karhunen–Loeve decomposition, is a model reduction method used to obtain low dimensional dynamic model of systems described by partial differential equations. Roughly speaking, this method is an optimal technique of finding a basis which spans an ensemble of data, collected from an experiment or a numerical simulation of a dynamical system. When these basis functions are used in a collocation formulation of Galerkin's projection will yield a finite dimensional system with the smallest possible degrees of freedom. Thus this technique is well suited to treat optimal control and parameter estimation of systems described by partial differential equations (Benerjee et al., 1998; Bangia et al., 1997; Park and Cho, 1996b).

In the present work, this method is applied to analyse several different examples, starting from the less complex and increasing the degree of complexity of the system. This was done in order to determine the versatility of this low dimensional technique. The reduced model obtained was used for parameter estimation and field reconstruction purposes.

1.2 CONTRIBUTIONS OF THIS DISSERTATION TO THE STATE-OF-THE-ART

The main objective of this work is to expand significantly the existing results and to develop new approaches to determine optimal observations strategies for distributed parameter systems, particularly in the case of sensor location. Predominantly, the

problem is to develop new or implement existing algorithms for different schemes and measurement error correlations. A secondary objective is to present some efficient techniques when a parametric uncertainty has to be taken into account.

The main thesis of this work can be proposed as;

> For a wide group of distributed parameter systems a considerable enhancement in the parameter estimation quality is possible through the development of effective and robust approaches in the sense of statistical uncertainty, using an optimal sensor location strategy.

In the process of testing this thesis, it was essential to develop a theoretical basis for new approaches and to build some new algorithms for different types of calculations. The subsequence is a concise summary of the contributions provided by this work to the state-of-the-art in sensor location methods.

THEORETICAL FEATURES

- Generalisation of the established results of optimal experiment design to multiple input multiple output systems with potential output correlation.

- Variation and generalisation of a number of algorithms of nonlinear programming and optimal experiment design to resolve sensor location problems.

- Development of an approach to solve the sensor placement problem in the case of correlated observations.

- Adaptation of established techniques to optimal sensor allocation in the existence of model parametric uncertainties.

- Improvement of the parameter estimation procedure using a reduced model.

APPLICATION FEATURES

- Application of optimal measurement strategies to heat and mass transfer systems considering linear and nonlinear forms of transfer mechanisms.

- Application of optimal measurement strategies to chemical reactions in a non-isothermal tubular reactor, considering stable and unstable operation conditions.

This submission consists of a group of papers published and/or under publication. They are interrelated according to the findings derived of my research work.

PAPERS PUBLISHED.

The first paper (see Section 2.1) introduces the analysis of the sensitivity coefficients in the spatial domain in order to define the optimal sensor locations that further will be used to estimates the parameters of the system in the best possible way. In this paper the optimal positions are defined by the allocations of the largest absolute value of the sensitivity coefficients (Alaña, 2010b). In the second paper (see Section 2.2) a comparative analysis of the different techniques available is introduced. These techniques are compared with our approach consisting of the extrema analysis of the sensitivity coefficients and/or POD-modes. An analytical and numerical demonstration is presented to show how the extrema values of the sensitivity coefficients and the POD modes coincide at the same spatial locations (Alaña, and Theodoropoulos, 2010). The third paper (see Section 2.3) shows the results obtained when a fast and reliable calculation of the temporal coefficients belonging to an empirical approximator was needed. This was carried out because these coefficients were required to define the optimal sampling time complementing the whole spatio-temporal optimal experiment design proposed in this work. The optimal time measurement allocations can be defined by means of the evaluation of the extrema values of these temporal coefficients in the time domain (Alaña, 2010a).

PAPERS UNDER REVIEW.

Additionally to the papers mentioned above, two more manuscripts are considered in this work. These works are under review and some comments have been received from the reviewers.

The first paper (4th in this submission, see Section 3.1) entitled; "**Optimal measurement locations for Parameter Estimation of Distributed Parameter Systems based on the use of Artificial Neural Networks**" has been submitted to the Applied Soft Computing. This work presents the use of the Artificial Neural Networks (ANNs) to estimate the sensitivity coefficients of a system without knowledge of the governing equations. The results were obtained during the execution of my research, and our intentions were based on calculating these sensitivity coefficients of the process to further determination the optimal sensor positions for parameter estimation purposes. This approach is limited to the complexity of the system under study. The second manuscript (5th in this submission, see Section 3.2) and entitled; "**Optimal Spatial Sampling Scheme for Parameter Estimation of Non Linear Distributed Parameter Systems**" is a complement of the papers presented in Section 2.1 and 2.2. In this work the concepts and new approaches introduced in previous papers are extended to systems containing high non-linear components. New analytical and numerical findings are introduced and explained to support our hypothesis which delineates the optimal measurement locations for parameter estimation by the direct analysis of the extrema values of the sensitivity coefficients and/or POD-modes.

2. DISCUSSION OF THE SUBMITTED PUBLICATIONS

2.1 PUBLICATION 1. **Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems**. Alaña, J., Brazilian Journal of Chemical Engineering, 27(4), 627(2010).

This paper was accepted for publication in May 2010. It is believed to be the first where the spatial locations of the largest absolute value of the sensitivity coefficients of a system described by a highly non-linear distributed parameter system are considered for parameter estimation purposes.

A sensor placement strategy to estimate parameters is presented. The spatial locations selection is essential in the parameter estimation procedure. To optimally estimate the parameters of the system, the measurements can be taken where the extrema values of the sensitivity coefficients are reached in the spatial domain. The parameters obtained with the highest deviation from the nominal values are usually the ones showing the lowest absolute sensitivity coefficients. To improve this estimation it is advisable to measure at the locations where the sensitivity functions reach their extrema values. The results are strongly affected by the presence of noise, but this can be partly solved using common filtering techniques. Not only the uncertainty in parameter estimation can be

reduced by using these sensor locations, but cost is decreased also, which is significant, especially in situations where experimental observations are expensive.

2.2 PUBLICATION 2. **Optimal location of measurements for parameter estimation of distributed parameter systems**. Alaña, J., Theodoropoulos, C., Computer and Chemical Engineering, 35(1), 106(2011).

This paper was accepted for publication in April 2010. Here, it is demonstrated, for first time, that the spatial locations of the extrema values corresponding to the predominant sensitivity coefficients coincide with the positions where the POD-modes reach their extrema values. This was done numerically and analytically.

In this work, a comparison between different methods reported in the literature on estimating optimal sensor locations to compute system parameters was presented. Three relevant methods from the most recent references were presented, and a number of observations based on case studies were made. Furthermore, a method where sensors placed at the positions where POD functions capturing most of the system's energy (and sensitivity coefficients) reach their extrema values was discussed. The method based on the Gram determinant evolution, in some cases, might not produce accurate results. This method is strongly dependent on the sensitivity coefficients behaviour, and requires extensive calculations. The max-min optimisation method can produce more than one optimal solution, and requires more calculations as the number of spatial intervals is increased for a fixed number of sensors. It was observed that this method depends on the magnitude of output values, assigning the measurements points to the positions

where the outputs reached their extrema values. The D-optimal design method produces as results the number and positions where the measurements should be made in the spatial domain. This method depends strongly on the sensitivity coefficients, but mostly on their spatial evolution rather than their exact magnitudes. The results are strongly affected by the presence of noise in the system, but this can be partly solved using common filtering techniques. The methods that are based on the calculation of sensitivity coefficients are the ones most affected by the presence of noise. While the D-optimal design requires post-processing of the sensitivity coefficients to determine positions and the optimal number of measurements, a mere observation of the extrema values of the sensitivity coefficients can produce the same results. Furthermore, it was observed in all cases studied that the measurements taken at the spatial positions where sensitivity coefficients, POD modes, and/or output values reach their extrema values produce best estimates of the system parameters. Not only the uncertainty in parameter estimation was reduced by using these sensor locations, but cost is reduced also, which is significant, especially in situations where experimental observations are expensive.

2.3 PUBLICATION 3. **Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes**. Alaña, J., Computer and Chemical Engineering, 34(8), 1220(2010).

This paper was accepted for publication in February 2010. In this work a new approach to identify processes without any knowledge of the mathematical model of the system has introduced and presented. An empirical approximator is used by means of the empirical eigenfunctions obtained from the application of the POD method. Only a few

experimental data are required to determine these empirical eigenfunctions and the subsequent temporal coefficients. For the latter very good results are guaranteed for any case study and/or system under study. This can be explained because it involves the minimisation of a linear objective function subject to a linear constraint defined by the linear combination of the empirical eigenfunctions and the temporal coefficients. A nonlinear chemical process described by a distributed parameter system is considered. This method, which is easy and fast to implement, produces the lowest deviation from the original experimental data and is very useful when the theoretical model of the system is unknown or difficult to determine. The optimal solution is obtained independently of the values used as initial guess. It is worth mentioning that the same advantages of this approach are observed for any other empirical approximator or ansatz used.

This approach was developed to determine in a fast and efficiently way the temporal coefficients of an empirical approximator using experimental data and without deriving a reduced model based on the Galerkin projection method. These temporal coefficients will be used, later on, to define an optimal measurement location in the time domain for parameter estimation purposes. A future manuscript exploiting this approach is under revision.

3. DISCUSSION OF THE PUBLICATIONS UNDER REVIEW

3.1 PUBLICATION 1. **Optimal measurement locations for Parameter Estimation of Distributed Parameter Systems based on the use of Artificial Neural Networks**. Alaña, J., Applied Soft Computing, (2010).

The Applied Soft Computing accepted this manuscript for review in May 2010. Here, a new approach using artificial neural network is presented. These artificial neural networks were trained using data generated from simulations. These data were obtained introducing changes in the parameters of the system (defined as inputs to the neural network). The chain rule was applied and an analytical expression relating the predicted output and the inputs of the neural networks was found. These relationships can be used as an approximation of the sensitivity coefficients. Therefore, it is possible to calculate the Fisher Information Matrix and/or the covariance matrix of parameter estimates, which is very helpful for optimal experiment design purposes. Additionally, it can be used to predict the outputs of the system during the parameter estimation process. It is established that the optimal experimental samplings should be made at the positions where the sensitivity coefficients reached their largest absolute values in the spatial domain. A very well trained neural network can be used as a model predictor and

sensitivity coefficients estimator at the same time.

3.2 PUBLICATION 2. **Optimal Spatial Sampling Scheme for Parameter Estimation of Non Linear Distributed Parameter Systems**. Alaña, J., Theodoropoulos, C., Computer and Chemical Engineering, (2010).

The Computer and Chemical Engineering accepted this manuscript for review in November 2010. In this paper we introduced a novel technique to define an optimal sampling scheme for a process described by a nonlinear distributed parameter system. This is carried out for parameter estimation purposes. This scheme is based on the Proper Orthogonal Decomposition (POD) and it is verified by a numerical example regarding the chemical reaction occurring in a tubular reactor for two possible scenarios; stable and unstable operation conditions. An extrema evaluation of POD modes can be used directly to define optimal sensor locations for parameter estimation of nonlinear distributed parameter system. This is a computational efficient approach that allows defining the optimal measurement placements without the laborious calculation of the sensitivity coefficient Jacobian matrices generally required by the D-optimal experiment design. Moreover, only merely system responses and/or experimental observations are required and used straightforwardly. The underlying combination of model reduction techniques and sensor location problem in this approach becomes even more relevant as the complexity of the system under consideration increases. By this method, not only the uncertainty in parameter estimates can be reduced, but cost is decreased also, which is significant, particularly in circumstances where experimental observations are expensive.

5. CONCLUSIONS

"God does not care about our mathematical difficulties. He integrates empirically"

Albert Einstein

A comparison between different methods, reported in the literature, for optimal sensor location to estimate parameters of distributed parameter systems was presented, observations and conclusions were derived. New approaches were developed, implemented and proved using an ample range of examples extracted from the literature. In the next paragraphs a series of conclusions derived from the results are presented.

The approach based on the Gram determinant evolution, in some cases, might not produce accurate results since it is strongly dependent on the behaviour of the sensitivity coefficients in the spatial domain.

The max-min optimisation method can produce more than one optimal solution depending on the number of sensors prescribed. Its computational cost increases significantly as the size of the DPS increases.

The D-optimal experiment design produces as results the optimal number and

measurements positions where theses samplings should be made in the spatial domain. This method depends strongly on the sensitivity coefficients, but mostly on their spatial evolution rather than their exact magnitudes.

The results, reported by the methods mentioned above, are strongly affected by the presence of noise, an aspect that can be partly solved using common filtering techniques. The approaches based on the sensitivity coefficients information are the most affected by this presence of noise.

It was verified, analytically and numerically, that the extrema values, in the spatio-temporal domain, of the sensitivity coefficients and the POD-modes coincide. Additionally, it was demonstrated, that in the time domain, the extrema values of these sensitivity coefficients concur at the same time positions of the temporal coefficients belonging to an empirical approximator.

As a general rule the optimal measurements positions can be allocated in the spatial locations where the sensitivity coefficients and/or POD-modes reached their extrema values.

The uses of the POD-Galerkin procedure can limit the function space to describe the observed solutions. Consequently this reduces the original nonlinear partial differential equations to a macro-model with small number of degrees of freedom. The resulting low-dimensional dynamic model simulates the system accurately using either data noise-free and/or contaminated with certain degree of noise.

Comparative numerical experiments show that the proposed reduced models, using the empirical eigenfunctions obtained from the POD-snapshot method, have numerical advantages over the uses of the original governing equations as a model. This includes; stability, robustness and a fast convergence of iterations during the parameter estimation process. The simulations show that this model reduction technique provides another feasible way for system designers to formulate, optimise and determine parameters of the process, efficiently and effectively.

One of the most important stages in the POD-Galerkin procedure is how to obtain appropriate snapshots. These snapshots must include information about the dynamic characteristics of the system to produce empirical eigenfunctions that will represent the dynamics of the system. The accuracy of the dynamic reduced model improves as the number of empirical eigenfunctions employed in the POD-Galerkin procedure increases. Even though that most authors suggest to extract the average from the snapshots, previous to the calculation of the empirical eigenfunctions, it was observed that a reduced model reconstructs the system satisfactorily with or without extracting this average.

The procedure of solving inverse problems, for parameter estimation purposes, employing low-dimensional model is faster than employing the original nonlinear partial differential equations. The parameters estimated using these reduced models are more accurate when the experimental data is taken at the locations where the sensitivity coefficients and/or POD-modes reached their extrema values, even for scenarios of measurements contaminated with a certain degree of noise. This procedure can be

improved even more when the samples are taken at the time intervals defined either by the positions of the extrema values of the sensitivity coefficients and/or the temporal coefficients belonging to the empirical approximator.

Using the reduced model for parameter estimation purposes the computation time is lowered in almost 60% of the time required when the original partial differential equations are used. A very good estimation of the parameters is obtained using the reduced model, and for simplex and stable case study, do not depends that much of the positions of the measurements, but mostly of the number of eigenfunctions used and the number of measurements taken in the time domain.

The temporal coefficients calculated from the new approach based in the minimisation of the objective function agreed the coefficients determined from the reduced model. The residuals obtained employing these temporal coefficients in the empirical approximator, are smaller than the ones produced by the uses of the conventional reduced model derived from the Galerkin-POD combination.

In the specific case of snapshots contaminated with a certain degree of noise it was observed that the reduced model acts as a filter and could reconstructs a field with a high accuracy. Then, using the reduced model for parameter estimation purposes produce better results than when the original model is used.

The performance of the Nelder-Mead algorithm varies from problem to problem that a careful collection of the experimental data will probably save time and produce better

results. Choosing and calibrating optimisation algorithms for effective and efficient solutions to parameter identification problems is important. However, there are many other important issues in parameter estimation problems that merit investigation. Model parameters can be correlated with one another. The issue of correlations among parameters causes several interesting problems. If parameters are correlated, then small or even large changes in two or more parameter values could result in the same objective function value. This phenomena creates the possibility of multiple, or even infinite global minima.

Noise in the experimental data affects the parameter estimation problem by increasing the minimum objective function value. In most cases, noisy versions of smooth optimisation problems require fewer simulations than the smooth problem itself. Poor choices of algorithmic parameters can result in fitting values many orders of magnitude higher than if a good set of parameter values are used. There are combinations of parameter samples where one parameter sample requires fewer simulations than another parameter sample and yet results in a lower fit value. Thus, a parameter sample that requires a large number of simulations does not necessarily result in a low fit value.

4. FUTURE WORK

In future, improvements to the experimental data collection scheme can be achieved by incorporating the optimal sampling time to the spatial measurement array defined by the extrema analysis of the POD-modes/temporal coefficients. In this matter a paper is under preparation, which will show analytically and numerically how the determinant of the covariance matrix can be minimised considering the experimental observations at the spatial locations where the empirical eigenfunctions show extrema values, and the sampling times where the temporal coefficients derived from the combined Galerkin-POD method reach their extrema values. Additionally, another manuscript introducing the uses of the spatio-temporal locations of the largest absolute value of the sensitivity coefficients to define the spatio-temporal experimental measurements scheme, and the equivalence of these sensor locations with the POD modes/temporal coefficients derived from the POD method is under preparation.

New studies regarding the analytical and numerical evaluation of the objective function dynamic behaviour respect to the parameters of interest need to be conducted. This aspect is especially interesting when a reduced model is used instead of the original governing equations.

New approach to calculate the spatio-temporal locations of the experimental measurements need to be developed, considering artificial neural networks and/or calculating the POD modes based on the sensitivity coefficients. Here, the POD modes can be determined using the spatio-temporal dynamic behaviour of the sensitivity coefficients instead of the experimental observations.

6. REFERENCES

Aihara, S. I., Consistency of extended least square parameter estimate for stochastic distributed parameter systems. In G. Bastin and M. Gevers (eds), Proceeding 4th European Control Conference, Brussels, Belgium, 1-4 July, EUCA(1997).

Alaña, J., Kinetics of ion-exchange in raw and commercial pasteurized milk. Rev. Téc. Ing. Univ. Zulia, 29(2), 182(2006).

Alaña, J., Determinación y estudio de dispersión de las concentraciones de monóxido de carbono en la zona oeste de la ciudad de Maracaibo-Venezuela. Multiciencias, 7(2), 115(2007).

Alaña, J., Nernst-Planck Model applied to the kinetics of ionic exchange of the system PO_4^{-3}-Cl^- in milk. Rev. Téc. Ing. Univ. Zulia, 30(3), 236(2007).

Alaña, J., Influence of a thermoelectric facility on V and Ni concentration in sediments in the city of Maracaibo, Venezuela. Interciencia, 35(3), 217(2010).

Alaña, J., Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes. Computer and Chemical Engineering, 34(8), 1220(2010a).

Alaña, J., Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems. Brazilian Journal of Chemical Engineering, 27(4), 627(2010b).

Alaña, J., Theodoropoulos, C., Optimal location of measurements for parameter estimation of distributed parameter systems. Computer and Chemical Engineering, 35(1), 106(2011).

Andó, B., Cammarata, G., Fichera, A., Graziani, S., Pitrone, N., A procedure for the optimization of air quality monitoring networks. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and reviews, 29(1), 157(1999).

Banks, H. T., Computational issues in parameter estimation and feedback control problems for partial differential equation systems. Physica D, 60, 226(1992).

Banks, H. T., Kunish, K., Estimation techniques for distributed parameter systems. Systems and Control: Foundations and Applications, Boston (1989).

Banks, H. T., Smith, R. C., Wang, Y., Smart material structures: Modeling, estimation and control. Research in applied Mathematics, Paris (1996).

Battaile, C. C., Srolovitz, D. J., Kinetic Monte Carlo simulation of chemical vapor deposition. Annual Review of Materials Research, 2,297(2002).

Belfiore, L. A., Transport Phenomena for Chemical Reactor Design. John Wiley and Sons, USA (2003).

Bennet, A. F., Inverse methods in physical oceanography. Cambridge monographs on mechanics and applied mathematics. Cambridge University Press (1992).

Berliner, M. L., Nychka, D., Hoar, T., Studies in the atmospheric sciences. Vol. 144 of lecture Notes in Statistics, Berlin: Springer-Verlag (2000).

Byrd, R. H., Stewart, W. E., Lightfoot, E. N., Transport Phenomena. John Wiley and Sons, USA (2002).

Chavent, G., On the theory and practice of non-linear least squares. Advances in water resources, 14(2), 55(1991).

Chen, C. C., Chang, H. C., Accelerated disturbance damping of an unknown distributed system by nonlinear feedback. AIChE Journal, 38(9), 1461(1992).

Curtain, R. F., Zwart, H., An introduction to infinite-dimensional linear systems theory. Texts in applied mathematics, New York: Springer-Verlag (1995).

Demetriou, M. A., Activation policy of smart controllers for flexible structures with multiple actuator/sensor pairs. In A. El Jai and M. Fliers (Eds), Proceeding 14th International Symposium of Mathematical Theory of Networks and Systems, France, 19-23 June (2000).

Filipe, J. A. N., Maule, M. M., Effects of dispersal mechanisms on spatio-temporal development of epidemics. Journal of Theoretical Biology, 226, 125(2004).

Fitzhugh, R., Impulses and physiological states in theoretical models of nerve membrane. Biophysical Journal, 1, 445(1961).

Gibson, J. S., George, H., Wu, C., Least-square estimation of input/output models for distributed linear systems in the presence of noise. Automatica, 36, 1427(2000).

Gil, M. I., Stability of finite and infinite dimensional systems. Boston: Kluwer Academic Publishers (1998).

Gilmer, G. H., Huang, H., Diaz de la Rubia, T., Dallas Torres, J., Bauman, F., Lattice Monte Carlo models of thin film deposition. Thin Solid Films, 365, 189(2000).

Grabowski, P., Lecture Notes on Optimal Control Systems. Cracow: University of Mining and Metallurgy Publishers (1999).

Hogg, N. G., Oceanographic data for parameter estimation. In P. Malanotte-Rizzoli (ed.), Modern Approaches to Data Assimilation in Ocean Modeling, Elsevier Oceanography, Amsterdam: Elsevier, 57(1996).

Isakov, V., Inverse problems for partial differential equations. Applied Mathematical Science, New York: Springer-Verlag, (1998).

Kee, R. J., Coltrin, M. E., Blarborg, P., Chemically reacting flow: Theory and practice. John Wiley and Sons, USA (2003).

Korbicz, J., Uciński, D., Pieczyński, A., Marczewska, G., Knowledge-based fault detection and isolation system for power plant. Applied Mathematics and Computer Science, 3(3), 613(1993).

Korbicz, J., Zgurowsky, M. Z., Estimation and control of stochastic distributed parameter systems. Warsaw: Polish Scientific Publishers (1991).

Kovarik, K., Numerical models in groundwater pollution. Berlin, New York: Springer-Verlag (2000).

Kowalewski, A., Optimal control of infinite dimensional distributed parameter systems with delays. Cracow: University of Mining and Metallurgy Press, (2001).

Kowalewski, A., Optimal control of time delay parabolic systems. Optimization, 50(1-2), 205(2001).

Kowalewski, A., Optimal control problems of distributed parameter systems with boundary conditions involving time delays. Scientific Bulletins of the University of Mining and Metallurgy, series Automatics, Cracow, University of Mining and Metallurgy Publishers (1991).

Kubrusly, C. S., Distributed parameter system identification: A survey. International Journal of Control, 26, 509(1977).

Kunisch, K., A review of some recent results on the output least squares formulation of parameter estimation problems. Automatica, 24(4), 531(1988).

Lara-Sagahón, A., Govezensky, T., Méndez-Sánchez, R. A., José, M. V., A lattice-based model of Rotavirus Epidemics. Physica A, 359, 525(2006).

Lasiecka, I., Active noise control in an acoustic chamber: Mathematical Theory. In S. Domek, R. Kaszyński and L. Tarasiejski (Eds.), proceeding 5th Int. Symposium of Methods and Models in Automation and Robotics, Poland, 25-29 August (1998).

Lasiecka, I., Triggiani, R., Control theory for partial differential equations: Continuous and approximation theories. Vol. I and II of Encyclopedia of Mathematics and Its Applications, Cambridge: Cambridge University Press (2000).

Li, X., Yong, J., Li, H., Optimal control theory for infinite dimensional systems (Systems and Control). Berlin: Springer-Verlag, (1995).

Luo, Z., Guo, B., Morgul, O., Stability and Stabilization of infinite dimensional systems with applications. Heidelberg: Springer-Verlag, (1999).

Maerivoet, S., De Moor, B., Cellular automata models of road traffic. Physics Reports, 419, 1(2005).

Malanotte-Rizzoli, P., Modern approaches to data assimilation in Ocean Modeling. Elsevier Oceanography, Amsterdam, Elsevier, (1996).

Malanowski, K., Nahorski, Z., Pesyńska, M., Modelling and optimization of distributed parameter systems. International Federation for Information Processing, Boston: Kluwer Academic Publishers (1996).

Mitkowski, W., Stabilization of dynamic systems. Warsaw: Wydawnictwa Naukowo-Techniczne, (1991).

Müller, W. G., Collecting spatial data. Optimum design of experiments for random fields. Contributions to Statistics, Heidelberg: Physica-Verlag, (1998).

Nagumo, J. S., Arimoto, S., Yoshizawa, S., An active pulse transmission like simulating nerve axon. Proceedings IRE, 50, 2061(1962).

Neumaier, A., Molecular modeling of proteins and mathematical prediction of protein structure. SIAM Review, 39(3), 407(1997).

Nychka, D., Saltzman, N., Design of air-quality monitoring networks. In D. Nychka, W. W. Peigorsch and L. H. Cox (Eds). Cases Studies in Environmental Statistics, Lecture Notes in Statistics, 132, New York: Springer-Verlag, 51(1998).

Omatu, S., Seinfeld, J. H., Distributed parameter systems: Theory and Applications. Oxford Mathematical Monographs, New York: Oxford University Press (1989).

Park, H. M., Cho, D. H., The use of the Karhunen-Loeve decomposition for the modelling of distributed parameter systems. Chemical Engineering Science, 51(1), 81(1996).

Polis, M. P., The distributed system parameter identification problem: A survey of recent results. Proc. 3rd IFAC Symp. Control of Distributed Parameter Systems, Toulouse, France, 45(1982).

Ray, W. H., Advance process control. McGraw-Hill, New York, (1981)

Robinson, J. C., Infinite-dimensional dynamical systems: An introduction to dissipative parabolic PDEs and the theory of global attractors. Cambridge: Cambridge University Press (2001).

Sankaranarayan, K., Shan, X., Keverekidis, I. G., Sudaresan, S., Bubble flow simulations with the Lattice Boltzmann method. Chemical Engineering Science, 54, 4817(1999).

Sasane, A., Hankel norm approximation for infinite-dimensional systems. Vol. 227 of Lecture Notes in Control and Information Sciences, Heidelberg: Springer-Verlag, (2002).

Sikora, J., Numerical approaches to the impedance and eddy-currental tomography. Warsaw: Technical University Press (2000).

Slepchenko, B. M., Schaff, J. C., Carson, J. H., Loew, L. M., Computational Cell Biology: Spatiotemporal simulation of cellular events. Annual Review of Biophysics and Bimolecular Structure, 31, 423(2002).

Sokolowski, J., Zoleiso, J., Introduction to shape optimization: Shape sensitivity analysis. Computational Mathematics, Berlin: Springer-Verlag, (1992).

Sturm, P. J., Almbauer, R. A., Kunz, R., Air quality study for the city of Graz, Austria. In H. Power, N. Moussiopoulos and C. A. Brebbia (Eds), Urban Air Pollution, Vol. 1, Southampton: Computational Mechanics Publications, 43(1994).

Sun, N., Inverse problems in groundwater modeling. Theory and applications of transport in porous media. Dordrecht: Kluwer Academic Publishers (1994).

Sun, N., Mathematical modeling of groundwater pollution. New York: Springer-Verlag (1996).

Sydow, A., Lux, T., Mietch, P., Schmidt, M., Unger, S., The DYMOS model system for the analysis and simulation of regional air pollution. In R. Grützner (ed.) Modellierung und Simulation im Umweltbereich, Wiesbaden: Vieweg-Verlag, 209(1997).

Sydow, A., Lux, T., Rosé, H., Rufeger, W., Walter, B., Conceptual design of the Branch-oriented system DYMOS (Dynamic Models for Smog Analysis). Transactions of the Society for Computer Simulation International, 15(3), 95(1998).

Uciński, D., Korbicz, J., Parameter identification of two-dimensional distributed systems. International Journal of Systems Science, 21(2), 2441(1990)

Uciński, D., Measurement optimization for parameter estimation in distributed systems.

Zielona Góra: Technical University Press, (1999)

Van Loon, M., Numerical Smog Prediction, I: The physical and chemical model. Technical Report NM-R9411, Centrum voor Wiskunde en Informatica, Amsterdam, (1994).

Van Loon, M., Numerical Smog Prediction, II: Grid refinement and its application to the Dutch Smog prediction model. Technical Report NM-R9523, Centrum voor Wiskunde en Informatica, Amsterdam, (1995).

Williams, R. A., Beck, M. S., Process tomography: principles, techniques, and applications. Oxford: Butterworth-Heinemann, (1995).

Zhdanov, V. P., Monte Carlo simulations of oscillations, chaos and pattern formation in heterogeneous catalytic reactions. Surface Science Reports, 45,231(2002).

Zwart, H., Bontsema, J., An application-driven guide through infinite-dimensional systems theory. In G. Bastin and M. Gerves (Eds), European Control Conference 1997: Plenaries and Mini-courses, Ottignies/Louvain-la-Neuve: CIACO, (1997).

THE SUBMITTED PUBLICATIONS

"Discovery consists of seeing what everybody has seen, and thinking what nobody has thought"

Albert Szent - Gyorgyi

Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems. Alaña, J., Brazilian Journal of Chemical Engineering, 27(4), 627(2010).

Brazilian Journal
of Chemical
Engineering

ISSN 0104-6632
Printed in Brazil
www.abeq.org.br/bjche

Vol. 27, No. 04, pp. 627 - 642, October - December, 2010

OPTIMAL MEASUREMENT LOCATIONS FOR PARAMETER ESTIMATION OF NON LINEAR DISTRIBUTED PARAMETER SYSTEMS

J. E. Alaña

Universidad Del Zulia, Facultad de Ingeniería, Escuela de Ingeniería Química,
Maracaibo, Venezuela.
School of Chemical Engineering and Analytical Science, Fax: +58 261 7598100,
The University of Manchester, Manchester, UK.
E-mail: joalana@luz.edu.ve

(Submitted: April 8, 2009 ; Revised: April 26, 2010 ; Accepted: May 10, 2010)

Abstract - A sensor placement approach for the purpose of accurately estimating unknown parameters of a distributed parameter system is discussed. The idea is to convert the sensor location problem to a classical experimental design. The technique consists of analysing the extrema values of the sensitivity coefficients derived from the system and their corresponding spatial positions. This information is used to formulate an efficient computational optimum experiment design on discrete domains. The scheme studied is verified by a numerical example regarding the chemical reaction in a tubular reactor for two possible scenarios; stable and unstable operation conditions. The resulting approach is easy to implement and good estimates for the parameters of the system are obtained. This study shows that the measurement location plays an essential role in the parameter estimation procedure.
Keywords: Sensor placement; Distributed parameter system; Parameter estimation; Sensitivity coefficients.

INTRODUCTION

An essential problem in parameter estimation is the selection of the experimental measurement locations. This problem consists of an arrangement of a limited number of measurements over the spatial domain that guarantees the best estimates of the system parameters.

The allocation of sensors is a task not necessarily dictated by physical considerations or by intuition and, therefore, some systematic approaches should still be developed to reduce the cost of instrumentation and to increase the efficiency of the parameter estimation procedures. Although the parameter estimation accuracy in a distributed parameter system depends significantly on the selection of sensor positions, only a few contributions to the experimental designs for such systems have been reported. The effects of either number or spatial allocation of measurements for the parameter estimation problem were considered by Kubrusly (1980), Carotenuto and Raiconi (1980), Kitamura and Taniguchi (1981), Courdesses et al. (1981), Courdesses and Amouroux (1982), and Nakagiri (1983).

Le Pourhiet and Le Letty (1976) offered two algorithms, as an optimal sensor location for parameter estimation applied to a deterministic distributed parameter system. The main idea is to maximise the identification error sensitivity with respect to the location of a new sensor. The first algorithm concerns the enhancement in the sensitivity criterion by adding a new sensor to the set of all sensors allocated in preceding iterations. The second takes into account the position of the new sensor at the previous iteration. Both algorithms stop when the assignment of a new sensor adds no

*To whom correspondence should be addressed

substantial improvement as far as the identification error sensitivity is concerned.

Sokollik (1976) considered the number and position of sensors, as well as the measurement times, to solve the parameter estimation problem of a distributed parameter system. The model was approximated by a lumped system using finite-differences. In this method both time and space domains were discretized with invariable sampling intervals. The optimal set of space-time was given by minimising the parameter estimate covariance.

Qureshi et al. (1980) presented a technique for scheming optimal experiments for distributed parameter system identification with observations contaminated with noise. The boundary perturbations were considered. The optimisation criterion to be maximised was the determinant of the Fisher Index Matrix (FIM) associated to the parameters to be identified, which depends on both boundary perturbations and spatial positions of the measurements. This method was developed for hyperbolic and parabolic partial differential equations.

Carotenuto and Raiconi (1981) considered the parameter identification of the one-dimensional static diffusion equation. The effects of an additional measurement point in terms of a potential improvement in the parameter estimate were analysed, and a criterion for selecting the location was presented. Rafajlowicz (1981) presented a method for optimal experiment of a distributed parameter system identification problem, which comprises sensor location and determination of classes of random inputs. A searching of an optimal probability measure corresponding to the position of the sensors was studied. This approach is equivalent to that considered in Qureshi et al. (1980), where the determinant of the information matrix is maximised. However, the information matrix was correlated to the system eigenvalues rather than to the system parameters. The conditions for optimality of the experimental design were derived, including an upper bound for the number of sensors (Rafajlowicz, 1983, Singh and Hahn, 2005, 2006, Zamprohna et al., 2005, Papadimitriou, 2004).

More recently, a number of computational approaches were introduced. (Kubrously and Malebranche 1985, Uciński 1992, 2000a, 2000b, 2003, 2005, Uciński and Bogacka, 2002, Uciński and Demetriou, 2004). The central idea is to define a design criterion to be minimised as a scalar measure of the FIM associated to the estimated parameters. This is followed by methods of optimum experimental design for nonlinear models to solve the sensor location problem at hand or, alternatively, by employing standard nonlinear programming procedures. Olanrewaju and Al-Arfaj (2006) presented a state estimation scheme based on Kalman filters for reactive distillation systems without extensively addressing the sensor selection problem. Nahor et al. (2003) minimised the ratio of the largest to the smaller eigenvalue of the FIM to compute optimal temperature sensor positions for food processes. To the best of our knowledge, these techniques have not been applied yet to systems described by partial differential equations, in spite of their resolute advantages (Uciński, 1999, Löhner and Camelli, 2005, Waterhouse et al., 2009, Vande, 2000, Venkateswarlu and Kumar, 2006, Punithakumar et al., 2006, Peng, 2005, Cotae et al., 2008, Balsa-Canto et al., 2008).

Christofides and Antoniades (2000, 2001, 2002) presented new approaches to calculate the optimal actuator/sensor locations of uncertain transport-reaction systems under control. They established that the solution to this problem is near-optimal for the closed-loop infinite-dimensional processes analysed. Similar and other kinds of techniques have been reported by Van den Berg et al. (2000) and Harries et al. (2004). These aspects will be analysed in future works regarding the optimal measurement positions for parameter estimation of distributed parameter systems.

This paper proposes an experimental design for measurement placement of distributed parameter systems. This is done by taking into consideration the spatial locations where the sensitivity coefficients of the system reached theirs extrema values.

THE SYSTEM GOVERNING AND SENSITIVITY EQUATIONS

The system considered is a homogeneous tubular reactor without catalyst packing. An irreversible first–order chemical reaction $A \rightarrow B$ takes place and is described by the concentration rate $C_A(t,z)$ and the temperature $T(t,z)$. This reaction is exothermic and a cooling jacket is used to remove heat from the reactor. A fraction of the products can be recycled, r, and mixed with the reactants at the inlet of the reactor ($z=0$) (Bendersky and Christofides 2000, Li and Chistofides 2007, Alaña, 2010).

Derived from mass and energy balances, and under the standard assumptions of constant density (ρ), heat capacity of the reacting fluid (cp), and constant axial fluid velocity (v), the dynamic behaviour of the reactor is described by a set of partial differential equations defined on a spatial domain $z \in (0,1)$, in dimensionless form,

$$\frac{\partial y_1(t,z)}{\partial t} = -\frac{\partial y_1(t,z)}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_1(t,z)}{\partial z^2} + R_{i3} \quad (1)$$

$$\frac{\partial y_2(t,z)}{\partial t} = -\frac{\partial y_2(t,z)}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_2(t,z)}{\partial z^2} + R_{i1} + R_{i2} \quad (2)$$

In these equations; $y_1(t,z)$, $y_2(t,z)$ are the state variables that represent concentration and temperature, respectively. The control variable u_C represents the cooling water temperature. The terms Pe_C, Pe_T are the Peclet numbers of mass and energy flows, respectively, B_C is the Damkohler number, γ is the activation energy, B_T and β_T are the parameters related to heat of reaction and heat transfer, respectively, and $t \in (t_0, t_f)$, where t_0 and t_f are initial and final times, respectively. The terms R_{ij} correspond to the examples; i=1 stable case with a recycle r=0.00, and i=2 unstable case with a recycle r=0.50, j stands for the following expression,

$$R_{11} = B_T B_C (1 - y_1(t,z)) \exp\left(\frac{y_2(t,z)}{1 + \frac{y_2(t,z)}{\gamma}}\right), \quad (3a)$$

$$R_{21} = B_T B_C (1 + y_1(t,z)) \exp\left(\frac{\gamma y_2(t,z)}{1 + y_2(t,z)}\right)$$

$$R_{12} = R_{22} = -\beta_T (u_C - y_2(t,z)) \quad (3b)$$

$$R_{13} = -B_C (1 - y_1(t,z)) \exp\left(\frac{y_2(t,z)}{1 + \frac{y_2(t,z)}{\gamma}}\right), \quad (3c)$$

$$R_{23} = B_C (1 + y_1(t,z)) \exp\left(\frac{\gamma y_2(t,z)}{1 + y_2(t,z)}\right)$$

The boundary and initial conditions are,

a. Stable case; $t \in (0,5)$

$$z = 0 \quad \begin{aligned} &\frac{\partial y_1(t,0)}{\partial z} = Pe_C y_1(t,0) \rightarrow \\ &\frac{\partial y_2(t,0)}{\partial z} = Pe_T y_2(t,0) \end{aligned} \quad (4a)$$

$$z = 1 \quad \frac{\partial y_1(t,1)}{\partial z} = 0 \quad \frac{\partial y_2(t,1)}{\partial z} = 0 \quad (4b)$$

$$t = 0 \qquad y_{10} = y_{20} = u_C = 0 \quad (4c)$$

b. Unstable case; assuming negligible reaction in the recycle loop and instantaneous mixing of fresh feed and recycle feed at the reactor inlet, the boundary and initial conditions in the interval $t \in (0,25)$ take the form,

$$z = 0 \quad \begin{aligned} &\frac{\partial y_1(t,0)}{\partial z} = Pe_C \left[y_1(t,0) - (1-r) y_{10} - r y_1(t,1) \right] \\ &\frac{\partial y_2(t,0)}{\partial z} = Pe_T \left[y_2(t,0) - (1-r) y_{20} - r y_2(t,1) \right] \end{aligned} \quad (5a)$$

$$z = 1 \quad \frac{\partial y_1(t,1)}{\partial z} = 0 \quad \frac{\partial y_2(t,1)}{\partial z} = 0 \quad (5b)$$

$$t = 0 \quad y_{10} = y_{20} = u_C = 0 \quad (5c)$$

The sensitivity coefficients can be calculated based on the local sensitivity analysis (Morbidelli and Varma, 1988, Vajda and Rabitz, 1992, Tildem et al., 1981, Juncu and Floarea, 1995). Indicating any of the reactor parameters [Pe_C, Pe_T, B_C, γ, B_T and β_T] as ϕ, the first-order sensitivity coefficient of Y, with Y being y_1 and y_2, is defined as,

$$s(Y,\varphi) = s_{Y,\varphi} = \frac{\partial Y}{\partial \varphi} \quad (6)$$

In this work, to take into account the pseudoadiabatic behaviour of the system, the sensitivity coefficients are computed using the direct method (Bauman et al., 1990, Coste and Aris, 2004). According to this method, the sensitivity coefficients can be expressed by the following equations,

$$\frac{\partial s_{y_1,\varphi}}{\partial t} = -\frac{\partial s_{y_1,\varphi}}{\partial z}\frac{\partial}{\partial \varphi}\left(\frac{1}{Pe_C}\frac{\partial^2 y_1(t,z)}{\partial z^2}\right) + \frac{\partial R_{i3}}{\partial y_2} s_{y_2,\varphi} + \frac{\partial R_{i3}}{\partial y_1} s_{y_1,\varphi} \quad (7)$$

$$\frac{\partial s_{y_2,\varphi}}{\partial t} = -\frac{\partial s_{y_2,\varphi}}{\partial z}\frac{\partial}{\partial \varphi}\left(\frac{1}{Pe_T}\frac{\partial^2 y_2(t,z)}{\partial z^2}\right) + \frac{\partial R_{i1}}{\partial y_2} s_{y_2,\varphi} + \frac{\partial R_{i1}}{\partial y_1} s_{y_1,\varphi} + \frac{\partial R_{i2}}{\partial y_2} s_{y_2,\varphi} \quad (8)$$

The boundary and initial conditions for each case study are,

a. Stable case,

$$z = 0 \quad \frac{\partial s_{y_1,\varphi}}{\partial z} = \frac{\partial}{\partial \varphi}\left(Pe_C y_1(t,0)\right)$$

$$z = 0 \quad \frac{\partial s_{y_2,\varphi}}{\partial z} = \frac{\partial}{\partial \varphi}\left(Pe_T y_2(t,0)\right) \tag{9a}$$

$$z = 1 \quad \frac{\partial s_{y_1,\varphi}}{\partial z} = 0 \quad \frac{\partial s_{y_2,\varphi}}{\partial z} = 0 \tag{9b}$$

$$t = 0 \qquad s_{y_1,\varphi} = s_{y_2,\varphi} = 0 \tag{9c}$$

b. Unstable case,

$$z = 0 \quad \frac{\partial s_{y_1,\varphi}}{\partial z} = \frac{\partial}{\partial \varphi}\left(Pe_C\left[y_1(t,0) - (1-r)y_{10} - r y_1(t,1)\right]\right)$$

$$\frac{\partial s_{y_2,\varphi}}{\partial z} = \frac{\partial}{\partial \varphi}\left(Pe_T\left[y_2(t,0) - (1-r)y_{20} - r y_2(t,1)\right]\right) \tag{10a}$$

$$z = 1 \quad \frac{\partial s_{y_1,\varphi}}{\partial z} = 0 \quad \frac{\partial s_{y_2,\varphi}}{\partial z} = 0 \tag{10b}$$

$$t = 0 \qquad s_{y_1,\varphi} = s_{y_2,\varphi} = 0 \tag{10c}$$

Equations (7) – (10) are obtained by differentiating the model equations (1) – (2), and (4) or (5) (depending of the case study), with respect to ϕ. The direct approach to obtain the sensitivity coefficients consists of solving (1) – (2) and (4), together with (7) – (9) for the stable case, and solving (1) – (2) and (5), together with (7) – (8) and (10) for the unstable example. For each case study, a system of fourteen coupled nonlinear partial differential equations is formed, with their respective initial and boundary conditions. This system can be solved using the Matlab® partial differential equation solver "pdepe", which solves initial-boundary value problems for systems of parabolic and elliptical partial differential equations in one space variable and time. The solver converts the partial differential equations to ordinary differential equations using a second-order accurate spatial discretization based on a set of nodes specified by the user. Time integration is then performed with a multistep variable-order method based on the numerical differentiation formulae. The spatial domain was divided into 21 equidistant intervals, and n time reporting intervals were considered. The nominal values of the parameter for the systems studied in this work are shown in Table 1.

Table 1: Nominal values of the parameters

Stable case			Unstable case		
$Pe_C = 1.0$	$B_C = 1.0$	$\gamma = 20.0$	$Pe_C = 7.0$	$B_C = 0.1$	$\gamma = 10.0$
$Pe_T = 1.0$	$B_T = 2.0$	$\beta_T = 1.0$	$Pe_T = 7.0$	$B_T = 2.5$	$\beta_T = 2.0$

In many cases, it can be assumed that some, albeit rough, a priori value of the parameters vector ϕ is on hand, e.g., determined from preliminary experiments. The aim in this work is to select a state sampling schedule to maximise the expected accuracy of the parameter estimates to be obtained from data generated in new experiments. This parameter should be determined with the lower uncertainties.

SENSITIVITY OF THE OPTIMUM PARAMETER $\hat{\varphi}$ WITH RESPECT TO THE EXPERIMENTAL DATA

A knowledge of how the solution of a system can vary with respect to small changes in the data or the parameters can yield insights into the model behaviour and can assists the modelling process.

Christopher and Fathalla (1999) proposed a new method to estimate the sensitivity of the state variables to the parameter estimates and the sensitivity of the parameter estimates to the observation. Using a set of analytical and numerical approaches, they concluded that the sensitivity of the parameter estimates to the observations is low if the sensitivity of the state variable to the parameter estimates is high. Similar results have been reported by Binder (2007). Next, the demonstration of the observation mentioned before can be found.

In order to compute $\frac{\partial \hat{\varphi}}{\partial u_{ij}}$, the sensitivity of the parameter $\hat{\varphi}$ to the experimental data u_{ij}, the objective function in the parameter estimation process has the form,

$$\Im(\varphi) \equiv \Im(\varphi, u) = \sum_i \sum_j \left(y\left(t_j, z_i, \varphi\right) - u_{ij}\right)^2 \tag{11}$$

This function can be considered to be a smooth function of φ in the vicinity of the optimal parameter $\hat{\varphi}$. Then the first and second derivatives are,

$$\frac{\partial \Im(\varphi,u)}{\partial \varphi_k} = 2\sum_i \sum_j \left(y\left(t_j,z_i,\varphi\right) - u_{ij}\right)\frac{\partial y\left(t_j,z_i,\varphi\right)}{\partial \varphi_k} \tag{12}$$

$$\frac{\partial^2 \Im(\varphi,u)}{\partial \varphi_l \partial \varphi_k} = 2\sum_i \sum_j \frac{\partial y\left(t_j,z_i,\varphi\right)}{\partial \varphi_l}\frac{\partial y\left(t_j,z_i,\varphi\right)}{\partial \varphi_k} + 2\sum_i \sum_j \left(y\left(t_j,z_i,\varphi\right) - u_{ij}\right)\frac{\partial^2 y\left(t_j,z_i,\varphi\right)}{\partial \varphi_l \partial \varphi_k} \tag{13}$$

In order to minimise the objective function (11), the right hand side of Equation (12) vanishes at $\varphi = \hat{\varphi}$ (where $\hat{\varphi} = \hat{\varphi}(u)$),

$$\sum_i \sum_j \left(y\left(t_j,z_i,\hat{\varphi}(u)\right) - u_{ij}\right) s_k\left(t_j,z_i,\hat{\varphi}(u)\right) = 0 \tag{14}$$

with, $s_k\left(t_j,z_i,\hat{\varphi}(u)\right) = \dfrac{\partial y\left(t_j,z_i,\varphi\right)}{\partial \varphi_k}$

At this moment, the left hand side of Equation (14) is a function of $\hat{\varphi}(u)$ and u_{ij}; differentiating both sides with respect to u_{ij} yields, for $k = 1,...,L_p$,

$$\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{l=1}^{L_p}\left(s_k\left(t_j,z_i,\hat{\varphi}\right)s_l\left(t_j,z_i,\hat{\varphi}\right) + \left(y\left(t_j,z_i,\hat{\varphi}\right) - u_{ij}\right)\frac{\partial^2 y\left(t_j,z_i,\hat{\varphi}\right)}{\partial \varphi_l \partial \varphi_k}\right)\frac{\partial \hat{\varphi}}{\partial u_{ij}} = s_k\left(t_j,z_i,\hat{\varphi}\right) \tag{15}$$

If we assume that $y\left(t_j,z_i,\hat{\varphi}\right)$ is close to the observed measured values u_{ij}, then the term $\left(y\left(t_j,z_i,\hat{\varphi}(u)\right) - u_{ij}\right)$ on the left hand side of Equation (15) can be neglected. Finally, the sensitivity of the best fit parameter estimate $\hat{\varphi}$ to the experimental measurements u_{ij} can be approximated using the following compact form,

$$\frac{\partial \hat{\varphi}}{\partial u_{ij}} \approx \left[\Theta(\hat{\varphi})\right]^{-1} s\left(t_j,z_i,\hat{\varphi}\right) \tag{16}$$

Here $s\left(t,z,\hat{\varphi}\right)$ is a L_px1 vector, given by the derivatives of the governing equations of the system with respect to the parameters, and $\Theta(\hat{\varphi}) = \sum_{i=1}^{N}\sum_{j=1}^{M} s\left(t_j,z_i,\hat{\varphi}\right)s^T\left(t_j,z_i,\hat{\varphi}\right)$ is a L_pxL_p non-singular matrix.

A desirable property of the model is that the sensitivity of the parameter estimate to the experimental measurement, $\dfrac{\partial \hat{\varphi}}{\partial u_{ij}}$, should be small in order to minimise the effect of the noise present in the experimental observation on the parameter estimation process. Equation (16) suggests that increasing $s\left(t_j,z_i,\hat{\varphi}\right)$ (the sensitivity of the state variable with respect to the unknown parameter) decreases the sensitivity of the parameter estimate to the experimental measurement. Then it is reasonable to expect that the best estimation of the parameters should be obtained when the experimental data are taken at the locations where the sensitivity coefficients reach their extrema values (the largest absolute values). It is our intention to verify here how useful these observations are to define a set of measurement locations for parameter estimation purposes using numerical experimentation.

PARAMETER ESTIMATION

Model calibration consists of finding a set of parameter values that produces the best model output which fit the observed data. This calibration is usually done by the minimisation of the objective function (17).

Once the sensor locations have been determined, the parameter estimation problem can be viewed as matching the model to the real system through the

minimisation of an error criterion over a set of admissible parameters. This can be defined as,

$$J(\varphi)=\frac{1}{m}\left[\frac{1}{t_f}\int_0^{t_f}\sum_{i=1}^{m}\left(y_e(z_i,t)-y_p(z_i,t)\right)^2 dt\right] \quad (17)$$

where, $y_p(z_i,t)$ is defined as the predicted response of the model at location z_i, and $y_e(z_i,t)$ as the experimental response at the same location, m is the number of optimal measurement locations determined by the different methods studied in this work. Unfortunately, elements in the vector $y_p(z_i,t)$ are not linear functions of the parameters ϕ, and multiple solutions of (17) are possible.

It is well known that problem (17) is challenging for numerous reasons. The existence of a solution is not certain, particularly if the observed data contain errors or if the model is grossly incorrect. It is unusual that any parameter set can accurately match the experimental data used, especially when these data are contaminated with noise. The presence of noise can promote difficulties during the optimisation process, causing spurious local minima and discontinuities. The gradient-based methods are the most affected of all. In this work, several different optimisation techniques (already existing in the Matlab® library) were used. The method of Nelder–Mead produced the best results for the scenarios studied. This can be explained by the fact that, using this technique, poor gradient approximations are not a problem, and continuity and differentiability of the objective function are not required.

Covariance Matrix of Parameter Estimates

When assessing the quality of an identified dynamic model, the covariance matrix of the estimated parameters gives an important measure (Ljung, 1999, Söderström and Stoica, 1989). In a maximum likelihood context, the inverse of the expected value of the negative of the Hessian provides the Cramer-Rao lower bound on the variance-covariance matrix of the parameter estimates. Large variances imply imprecise parameter estimates. The arrival of new data, or data measurement, can lead to substantial change in parameter estimates if the variance-covariance matrix is ill-conditioned. So, the goal of the estimation procedure is to determine unknown parameters in such a way that the difference between the sample covariance matrix and the implied covariance matrix is minimised in a certain sense (Davidson and MacKinnon, 1980, Sorenson, 1980, Zhun and Stein, 2005).

Based on the system studied here, the approximation of the parameter covariance matrix (Yen and Yoon, 1981) can be approximated by,

$$\mathrm{Cov}\left(\hat{\varphi}_{nv}\right)\approx\left(\frac{1}{\sigma_{y_1}^2}J_{y_1}^T J_{y_1}+\frac{1}{\sigma_{y_2}^2}J_{y_2}^T J_{y_2}\right)^{-1} \quad (18)$$

where J_{y_1} and J_{y_2} represent the Jacobian matrix of estimated variables; y_1 and y_2, respectively, with respect to changes in the parameters. nv is the parameter dimension, and σ_k^2 is the variance of the variable k. In this equation it is easy to view the influence of field data quality and quantity in parameter uncertainty. Since these Jacobian matrices are evaluated at those locations for which observations are available, any experimental design should aim at sampling at those locations where the variables are most sensitive to the estimated parameters. Such a design is said to provide the maximum amount of information about the unknown parameters (Knopman and Voss, 1987).

It is intuitively obvious that the experimental design objective should be intended to minimise the norm of the covariance matrix (18), i.e., to make matrices $\left(J_{y_1}^T J_{y_1}\right)^{-1}$ and $\left(J_{y_2}^T J_{y_2}\right)^{-1}$ as small as possible. The variance terms $\sigma_{y_1}^2$ and $\sigma_{y_2}^2$ are constants, and can be dropped from the formulation. To measure the accuracy of the estimates, we prefer to summarise the information about the variability in the covariance matrix into a single number. Here we used the determinant of (18) as the function that transforms a matrix into a scalar. This is quite informative since in fact it is related to the volume of the multidimensional simplex defined by the column/row vectors of the matrix (Kumar and Seinfeld, 1978, Rutzler, 1987).

Experimental design for parameter estimation deals with the problem of defining experimental conditions that increase the reliability of a simulation model. This can be formulated by using a measure of the covariance matrix of the parameter estimates. Among these, the most widely used design criteria are,

- A-optimality: A design is said to be A-optimal if it minimises the trace of matrix (18).
- D-optimality: A design is said to be D-optimal if it minimises the determinant of (18).
- E-optimality: A design is said to be E-optimal if it minimises the maximal eigenvalue of (18).

Using different norms leads to slightly different conclusions regarding the optimal design. The D-optimality criterion minimises the volume of the hyper-ellipsoid in the parameter space, which makes no consideration of the relationship between the ellipsoid's axes lengths, which are in turn proportional to the square root of the covariance matrix eigenvalues.

In general, it can be shown that, under some assumptions of regularity and for a sufficiently large sample size N, the vector $\left(\hat{\varphi}_{(N)} - \overset{*}{\varphi}\right)$ (with $\overset{*}{\varphi}$ denoting the 'true' but unknown value of the parameters and $\hat{\varphi}_{(N)}$ the least square parameter estimates) has approximately a normal distribution with zero mean and covariance matrix,

$$\frac{\sigma^2}{N} M^{-1}\left(\xi, \overset{*}{\varphi}\right) \tag{19}$$

where σ^2 denotes the standard deviation of the errors in the model, ξ represent a design experiment (including the measurement locations) and $M\left(\xi, \overset{*}{\varphi}\right)$ is the FIM, defined by,

$$M\left(\xi, \overset{*}{\varphi}\right) = \left(\sum_{k=1}^{n} \frac{\partial y\left(z_k, t, \varphi\right)}{\partial \varphi_i} \frac{\partial y\left(z_k, t, \varphi\right)}{\partial \varphi_j}\right)_{i,j=0}^{N} \tag{20}$$

In principle, the covariance matrix is a measure for the precision of the least square estimator for the unknown parameter $\overset{*}{\varphi}$ and a 'smaller' matrix yields more precise estimates. For example, the i^{th} diagonal element of (19) will be denoted by $\frac{\sigma^2}{N} M^{-1}\left(\xi, \overset{*}{\varphi}\right)_{ii}$ and is an approximation of the variance or mean squared error for the i^{th} component $\hat{\varphi}_{i,(N)}$ of the least squared estimator $\hat{\varphi}_{(N)}$. An approximate confidence interval for the i^{th} component φ_i of the vector φ is given by,

$$\left[\hat{\varphi}_{i,(N)} - \frac{\hat{\sigma} u_{1-\alpha/2}}{\sqrt{N}} \sqrt{M^{-1}\left(\xi, \overset{*}{\varphi}\right)_{ii}}, \hat{\varphi}_{i,(N)} + \frac{\hat{\sigma} u_{1-\alpha/2}}{\sqrt{N}} \sqrt{M^{-1}\left(\xi, \overset{*}{\varphi}\right)_{ii}}\right] \tag{21}$$

where $u_{1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of the standard normal distribution and $\hat{\sigma}^2$ is an estimate of the unknown variance of the error. For most cases, it was shown that, for moderate sample sizes N, the sampling variances of the parameter estimates are well approximated.

The precision of the estimates can always be decreased by increasing the sample size N, which yields a 'smaller' covariance matrix and smaller variances of the least square estimates. However, in practice the sample size is usually fixed, due to cost considerations of each additional experiment. To improve the quality of the estimates or, from a different point of view, to reduce the number of experimental measurements needed to obtain the estimates with a given accuracy, we note that the variances of the estimates $\hat{\varphi}_{i,(N)}$ and the covariance matrix of the vector $\hat{\varphi}_{(N)}$ also depend on the given design, ξ, which determines the relative proportion of total observations to be taken at the experimental locations.

It is advisable to check that the confidence interval magnitude agrees with the sensitivity analysis, where the parameter that has the largest sensitivity coefficient should have the smallest confidence interval. Because of the consideration of the measurement error only, if the FIM is well-defined, the confidence intervals may result in very small values. Grimstadt and Mannseth (1998) indicate that the use of such an approximation of the confidence intervals was almost always justified, even for the highly nonlinear model they analysed.

NUMERICAL RESULTS

The spatial variability of sensitivities has a significant impact on parameter estimation and sampling design for studies of distributed parameter systems. Information about a physical parameter will be most accurately gained at points in space with a high sensitivity to the parameter.

The set of partial differential equations described above was solved using nominal values for the parameters and random noise with zero mean, $N(0,\sigma^2)$, was added to the outputs, the variance, σ^2, was manipulated in order to produce up to ±15% of stochastic deviation from the outputs. For each case 10 replications of the digital experiment were collected.

Data set averaging is used to reduce the effect of random noise in the measurements, which typically

arises from the data acquisition system. In practice, this implies that several data sets of the same signal, in response to identical perturbations, have been taken. Data set averaging may be obtained as,

$$\hat{y}(t_i) = \frac{1}{nt}\sum_{j=1}^{nt} y_j(t_i) \qquad i = 1,...,N_S \qquad (22)$$

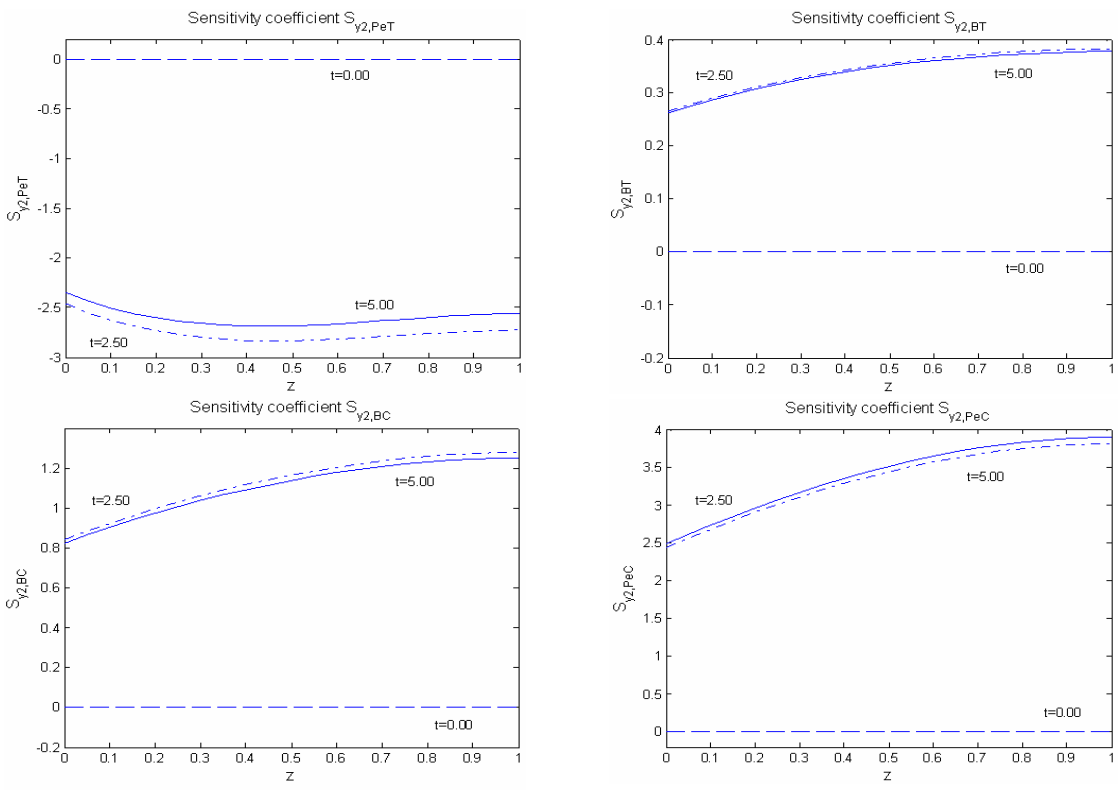
where nt is the number of data sets and N_S is the number of observations per set. The effects of the random noise tend to cancel when several sets of data are averaged. However, the noise can only be reduced to a certain degree and not totally eliminated. In this work, only minimal improvements were achieved by averaging together more than ten sets of data. In addition to averaging, the output can be low pass filtered to remove the residual random noise and the systematic noise.

The average of the resulting ten responses was then treated as experimental data and was used in the minimisation of the objective function (17). The calculated parameters were compared with their corresponding nominal values.

The analysis of the sensitivity coefficients and of the determinant of the FIM reveals the most appropriate sensor locations for the estimation of the unknown parameters (Storch et al, 2007). It is well known that sensitivity analysis quantifies the dependence of the system behaviour on the parameters that affect the process dynamics. Prasad and Vlachos (2010) presented results that show that high values of the FIM are correlated with large normalised sensitivity coefficients. These results can be very helpful to reinforce the analysis carried out here.

Figures 1 and 2 show the results of the sensitivity coefficients obtained for the stable tubular reactor case, i=1. From these figures the positions where the extrema values are reached can be extracted easily.

The extrema values of the sensitivity coefficients for temperature are located at 0.00, 0.45, and 1.00 and at 0.00, 0.30, and 1.00 for concentration. The parameter showing the lowest absolute sensitivity coefficient is γ, which implies that more sampling effort would be required to estimate this parameter. These measurement locations were considered and their effects on the parameter estimation procedure were verified. The minimum of the determinant of the covariance matrix, Equation (18) was used to measure the quality of the parameter estimates considering the spatial positions established by the extrema analysis of the sensitivity coefficients. Table 2 shows the scenarios considered for the parameter estimation of the stable example.



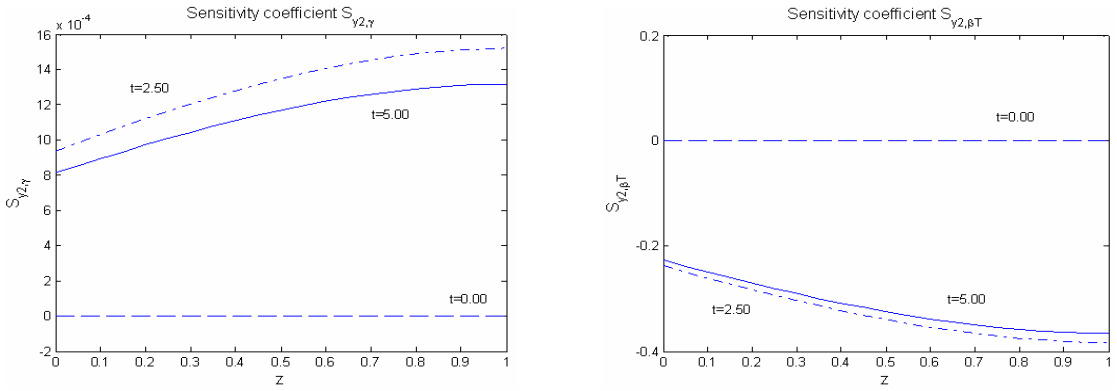


Figure 1: Sensitivity coefficients behaviour in the spatial domain, stable case.

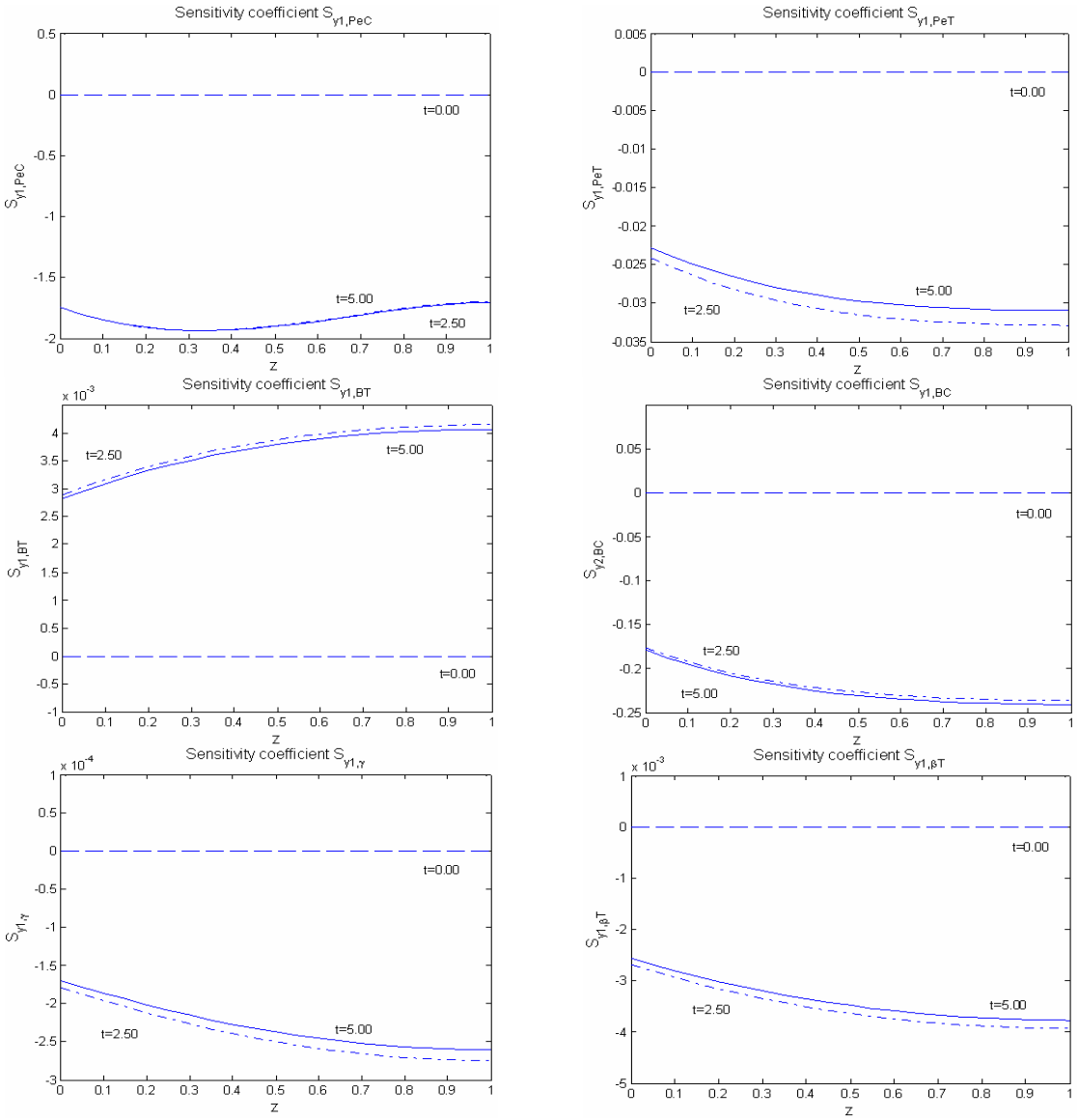


Figure 2: Sensitivity coefficients behaviour in the spatial domain, stable case.

It is our intention to consider the spatial locations where the sensitivity coefficients show extrema values to verify the accuracy of the parameter estimates using experimental data contaminated with noise. At these positions it would be logical to expect that the confidence intervals for each parameter should be small, according to Equation (21).

Matrix $\Theta(\hat{\varphi})$ in Equation (16) resembles the FIM, and Equation (20), the inverse of (16) can be considered as an approximation of the inverse of the FIM, which is a lower bound on the parameter covariance matrix. Even though Christopher and Fathalla (1999) derived their conclusions for purposes other than sensor location for parameter estimation, it is our intent to use their results to reinforce our hypothesis that a tentative set of optimal measurement positions could be defined considering the spatial locations of the extrema values of the sensitivity coefficients, which should result in small confidence intervals for the parameter estimates.

The procedure followed in this work resembles the conventional D-optimal experimental design. In the D-optimal design, the sensors are allocated by the minimisation of the determinant of the parameter covariance matrix. In our approach, the measurement positions are defined by a mere extrema analysis of the sensitivity coefficients, which can be defined as a locally optimal design, with a further evaluation of the determinant of the parameter covariance matrix to verify the quality of the parameter estimates. Both approaches need a previous estimate of the parameters.

Table 3 shows the values of the parameters estimated taking the experimental data at the locations mentioned in Table 2. The data estimated from the model at these same positions are used to minimise the objective function (17). The initial guess used corresponds to 50% of the nominal values. During each iteration the sensitivity coefficients were calculated using (7) and (8) together with (9) and/or (10). The values obtained for each case studied are shown in Tables 3 and 5. Finally, the parameters calculated were compared with their nominal values.

It can be seen in Table 3 that the parameter with the highest deviation was γ in most cases. The best estimation is obtained when the experimental measurements are taken at the locations where the sensitivity coefficients reached the extrema values for each variable of the system. The sampling scheme experienced a significant reduction from the original process, consisting of 21xn measurement locations for each state variable, to 3xn (based on the positions where the sensitivity coefficients reached extrema values). This reduction represents 86% of the original spatial domain.

For the tubular reactor with recycle operating under unstable conditions, Figures 3 and 4 show the behaviour of the sensitivity coefficients in the spatial domain for temperature and concentration, respectively.

Table 2: Locations considered for each variable, stable case

Case	Concentration locations	Temperature locations
A	Whole domain	Whole domain
B	0.00, 0.15, 1.00	0.00, 0.15, 1.00
C	0.00, 0.15, 0.95, 1.00	0.00, 0.15, 0.95, 1.00
D[a]	0.00, 0.30, 1.00	0.00, 0.45, 1.00

[a]Based on the extrema values of the sensitivity coefficients and considering the largest absolute value

Table 3: Parameters estimated, stable case

Case	Pe_C	Pe_T	B_C	B_T	γ	β_T	SSE	det[Cov]
A	1.0030 0.30%	1.0035 0.35%	0.9954 0.46%	2.0039 0.20%	20.0085 0.04%	0.9992 0.08%	1.33e-8	6.32×10^{-8}
B	0.9841 1.59%	0.9964 0.36%	1.0301 3.01%	1.9905 0.48%	6.8791 65.61%	0.9780 2.20%	0.0034	81.27×10^{-2}
C	0.9910 0.90%	1.0073 0.73%	1.0317 3.17%	1.9849 0.76%	7.2747 63.63%	0.9740 2.60%	0.0042	73.33×10^{-2}
D	1.0099 0.99%	0.9984 0.16%	1.0107 1.07%	1.9688 1.56%	19.0229 4.89%	0.9982 0.18%	2.15e-4	4.06×10^{-5}

Numbers in bold represent the deviation of the parameter from the nominal value.

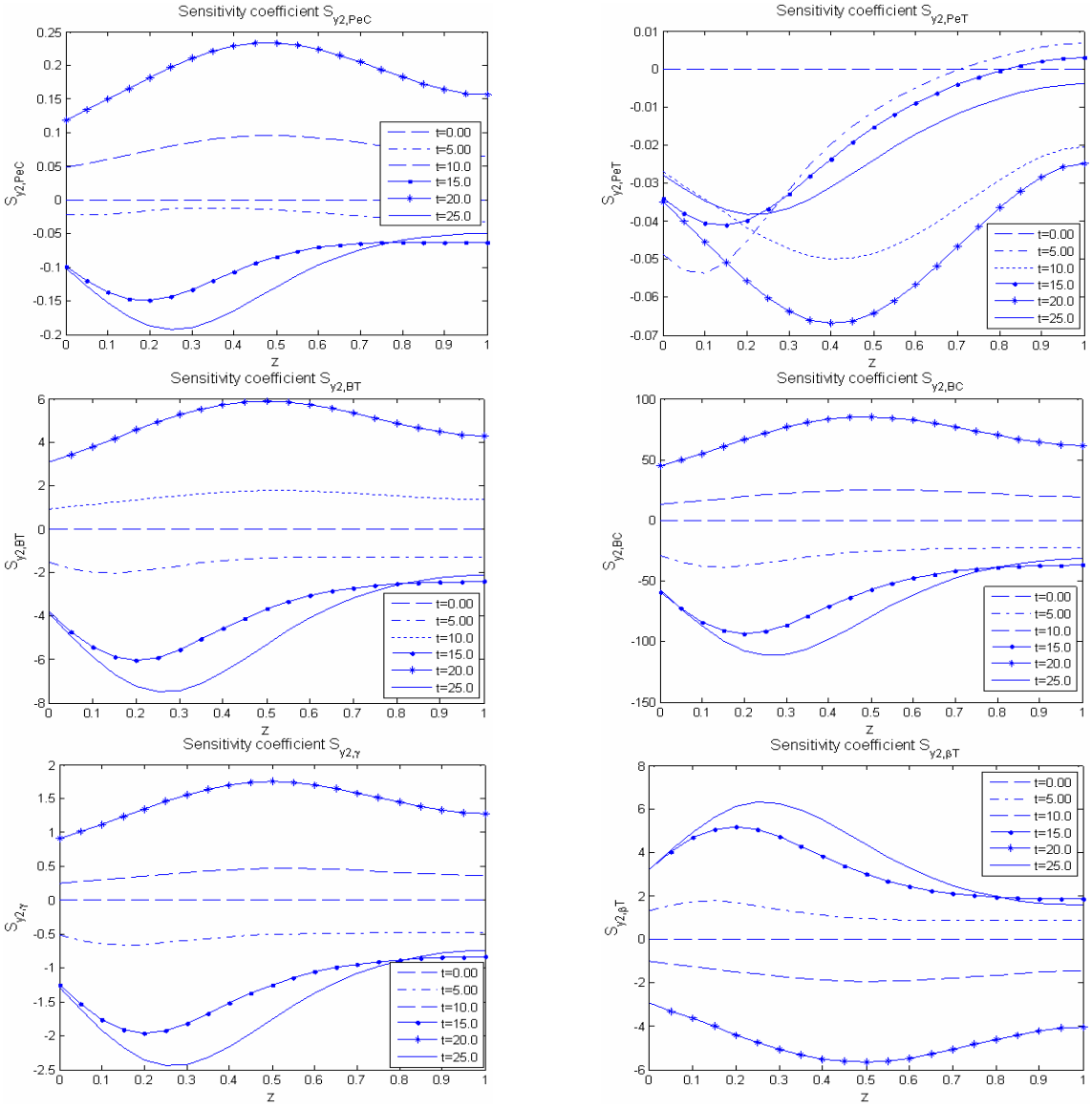
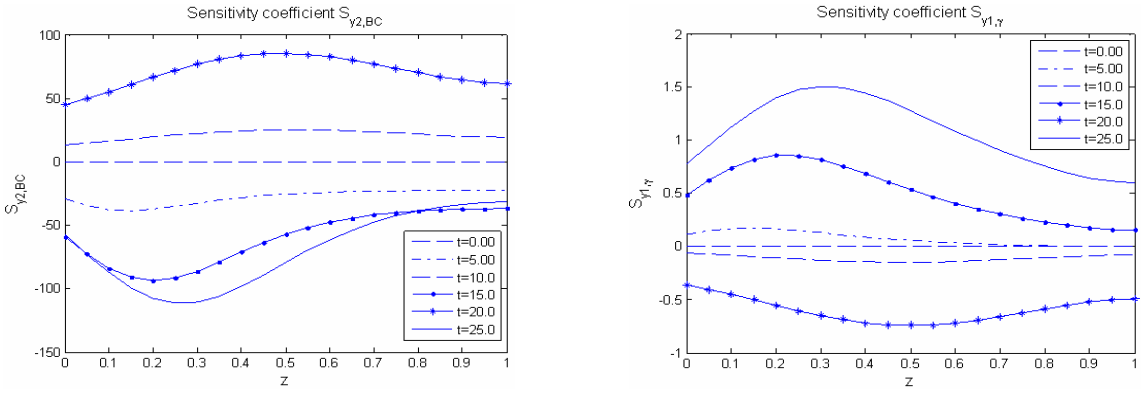


Figure 3: Sensitivity coefficients behaviour in the spatial domain, unstable case.



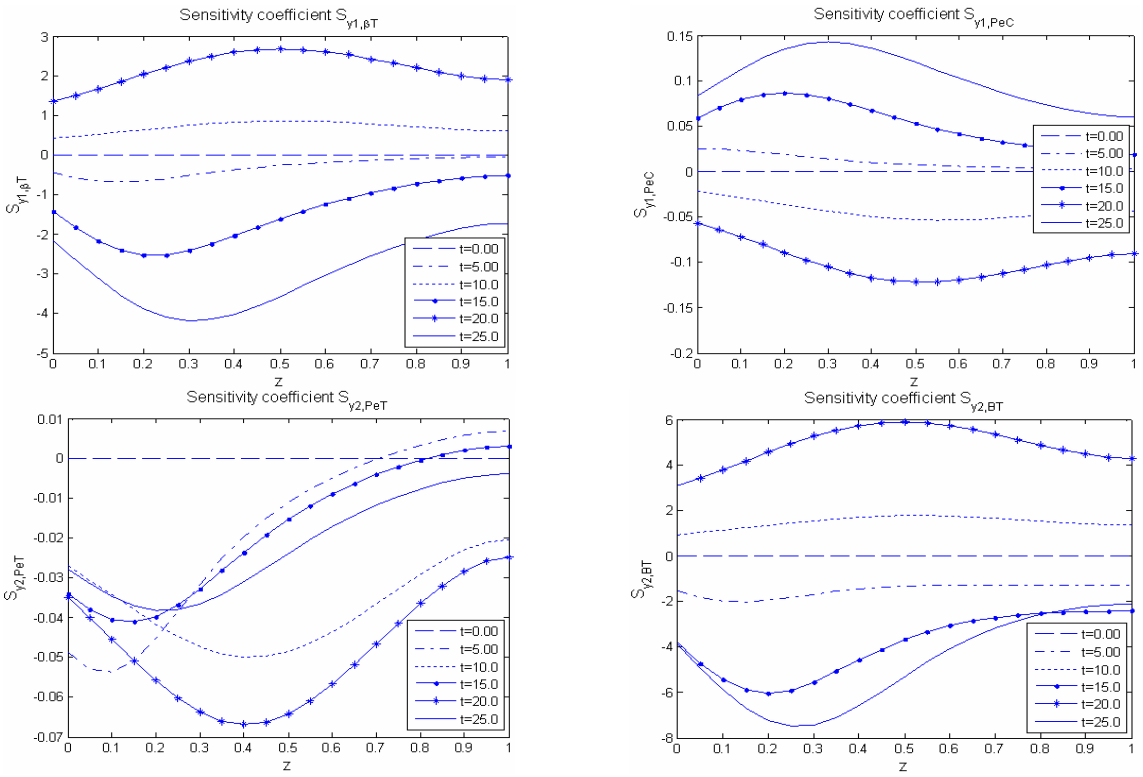


Figure 4: Sensitivity coefficients behaviour in the spatial domain, unstable case.

Examining these figures, the variety of extrema for different values of time can be seen. The largest absolute values for temperature are located at 0.15, 0.20, and 0.30 and at 0.15, 030, and 0.35 for concentration. The parameter showing the lowest absolute sensitivity coefficient is Pe_T. This system is very unstable, showing numerous extrema values in the whole spectrum of constant time lines, as can be seen in Figures 3 and 4. Even though that is not shown here, a quantification of the most recurrent extrema was carried out, which are resumed as Case D in Table 4. This table shows the cases considered to analyse the parameter estimation procedure for the unstable case.

The parameters estimated, taking the experimental measurements at the locations mentioned in Table 4 and following the procedure explained above, are shown in Table 5.

Table 4: Locations considered for each variable, unstable case

Case	Concentration locations	Temperature locations
A	Whole domain	Whole domain
B	0.05 0.10 … 0.65 0.70	0.05 0.10 … 0.65 0.70
C	0.15 0.30 0.35 0.45	0.15 0.20 030 0.40
D[a]	0.15 0.30 0.35	0.15 0.20 0.30

[a] Based on the extrema values of the sensitivity coefficients and considering the largest absolute value

Table 5: Parameter estimated, unstable case

Case	Pe_C	Pe_T	B_C	B_T	γ	β_T	10^{-3}*SSE	det[Cov]
A	7.2560 3.66%	6.3010 9.997%	0.1005 0.50%	2.4799 0.80%	9.9243 0.76%	2.0068 0.34%	3.049	$2.10\text{x}10^{-7}$
B	7.5560 7.93%	6.0010 14.27%	0.0974 2.60%	2.4480 2.08%	10.2121 2.12%	2.0799 3.40%	15.008	$7.95\text{x}10^{-4}$
C	7.7381 10.54%	4.6780 33.17%	0.1026 2.60%	2.3040 7.84%	10.2191 2.19%	1.9354 3.23%	19.482	$45.38\text{x}10^{-1}$
D	7.2279 3.26%	6.1370 12.33%	0.0998 0.20%	2.5233 0.93%	10.0330 0.33%	2.0231 1.16%	3.098	$3.89\text{x}10^{-7}$

Bold numbers represent the deviation of the parameter from the nominal value.

Due to the complex nature of this example, not even considering the whole experimental data in the spatial domain, the nominal values of the parameters are reproduced. Pe_T is the parameter showing the highest deviation from its nominal value. This gives another motivation for using sensor placement strategies based on the extrema evaluation of the predominant sensitivity functions showing the largest absolute value. Thus, not only the uncertainty in parameter estimation was reduced (see the deviation from the nominal values, number in parentheses, shown in Tables 3 and 5), but also the cost associated with such measures was reduced as well, which can be significant, especially in situations where experimental measurements are very expensive. Information about a physical parameter may be most accurately gained at points in space with a high sensitivity to the parameter. Taking observations at locations showing high sensitivity to the parameter tends to yield relatively low variance, as can be seen in Tables 3 and 5. Level of noise, above the ones considered here, affect considerably the parameter estimates.

The parameters with the lowest absolute sensitivity coefficients are the most difficult to estimate, especially if the experimental measurements are taken far from the optimal locations. When at least one parameter shows a very low absolute sensitivity, an optimisation algorithm based on the gradient and/or Hessian matrix generally gets stuck in a local minimum, distant from the optimal solution. Considering the measurement locations where the sensitivity coefficients show extrema values provided very good parameter estimates.

Taking into account the largest absolute sensitivity coefficients of the parameter of interest through a careful choice of observation points in a sampling design will lower the variance of the parameter estimate. This is the motivation for examining the behaviour of sensitivities in the system when refinement of parameter estimates is an objective of field sampling.

CONCLUSIONS

A sensor placement strategy to estimate parameters is presented. The spatial location selection is essential in the parameter estimation procedure. To optimally estimate the parameters of the system, the measurements can be taken where the extrema values of the sensitivity coefficients are reached in the spatial domain. The parameters determined with the highest deviation from the nominal values are usually the ones showing the lowest absolute sensitivity coefficients. To improve this estimation, it is advisable to measure at the locations where the sensitivity functions reach their extrema. The results are strongly affected by the presence of noise in the system, but this can be partly solved by using common filtering techniques. Not only the uncertainty in parameter estimation can be reduced by using these sensor locations, but cost is decreased also, which is significant, especially in situations where experimental observations are expensive.

NOMENCLATURE

$C_A(t,z)$	Concentration
$T(t,z)$	Temperature
$y_1(t,z)$	Dimensionless concentration
$y_2(t,z)$	Dimensionless temperature
r	Recycle to reactor
z	Spatial variable
Pe_T	Temperature Peclet number
Pe_C	Concentration Peclet number
B_C	Damkohler number
γ	Activation energy
B_T	Heat of reaction
β_T	Heat transfer
t	Time variable
t_0	Initial time
t_f	Final time
ϕ	A vector parameter of the system
Y	Output of the system
$s(Y,\varphi)$ or $s_{Y,\varphi}$	Local sensitivity of Y with respect parameter ϕ
$\frac{\partial\hat{\varphi}}{\partial u_{ij}}$	Sensitivity of the parameter $\hat{\varphi}$ to the experimental data u_{ij}
u_{ij}	Experimental data
$\Im(\varphi)$ or $\Im(\varphi,u)$	Objective function
$s(t,z,\hat{\varphi})$	Vector of local sensitivities
$\Theta(\hat{\varphi})$	Non-singular matrix
$\overset{*}{\varphi}$	the 'true' but unknown value of the parameters
$\hat{\varphi}_{(N)}$	the least square parameter estimates
ξ	the design experiment (including the measurement locations)
$M\left(\overset{*}{\xi,\varphi}\right)$	Fisher Index Matrix
$Cov(\hat{\varphi}_{nv})$	Parameter covariance matrix

$N(0,\sigma^2)$	random noise with zero mean and variance σ^2
$u_{1-\alpha/2}$	the 1-α/2 quantile of the standard normal distribution
J_y	Jacobian of the variable y
J_{y_k}	Variance of the variable k
SSE	Sum of square errors
$y_p(z_i,t)$	Predicted response at location z_i
$y_e(z_i,t)$	Experimental response at location z_i
$\hat{y}(t_i)$	Averaged data
m	Optimal number of spatial locations
nt	Number of data sets
N_S	Number of observations per data set

Acronyms

FIM	Fisher Index Matrix

REFERENCES

Alaña, J. E., Estimation of the Temporal Coefficients for an Empirical Approximator. New Approach based on the Proper Orthogonal Decomposition Modes. Computer and Chemical Engineering, doi:10.1016/j.compchemeng.2010.02.001 (2010).

Balsa-Canto, E., Alonso, A. A., Banga, J. R., Computing optimal dynamic experiments for model process calibration in predictive microbiology. Journal of Food Process Engineering, 31 (2), 186 (2008).

Bauman, E., Varma, A., Lorusso, J., Dente, M., Morbidelli, M., Parametric sensitivity in tubular reactors with co-current external cooling. Chemical Engineering Science, 45, 1301 (1990).

Bendersky, E., Christofides, P. D., Optimisation of transport-reaction processes using nonlinear model reduction. Chemical Engineering Science, 55, 4349 (2000).

Binder, T., Parameter estimation for dynamic process described with DDEs. Theses, Fakultät für Mathematik und Informatik, Ruprecht-Karls-Universität Heidelberg, Germany (2007).

Carotenuto, L., Raiconi, G., Identifiability and identification of a Galerkin approximation for a class of distributed parameter systems. Int. Journal Syst. Sci., 11, 1035 (1980).

Carotenuto, L., Raiconi, G., On the identification of a random parameter function in the one-dimensional diffusion equation and related location. Proc. 8th IFAC Triennial World Congress, Kyoto, Vol. II, 76 (1981).

Chistopher, T. H., Fathalla, A. R., Sensitivity analysis of parameters in modelling with delay-differential equations. Numerical Analysis Report No. 349, Manchester Centre for Computational Mathematics, Departments of Mathematics, The University of Manchester, England (1999).

Christofides, P. D., Antoniades, C., Computation of optimal actuator locations for nonlinear controllers in transport-reaction processes. Computer and Chemical Engineering, 24, 577 (2000).

Christofides, P. D., Antoniades, C., Integrated optimal actuator/sensor placement and robust control of uncertain transport-reaction processes. Computer and Chemical Engineering, 26, 187 (2002).

Christofides, P. D., Antoniades, C., Integrating nonlinear output feedback control and optimal actuator/sensor placement for transport-reaction processes. Chemical Engineering Science, 56, 4517 (2001).

Coste, J., Aris, R., Amundson, N. R., Tubular reactor sensitivity. AIChE Journal, 7, (1), 124 (2004).

Cotae, P., Yalamnchili, S., Chen, C. L., Ayon, A., Optimization of sensor locations and sensitivity analysis for engine health monitoring using minimum interference algorithms. EURASIP Journal on Advances in Signal Processing, 9 (2008).

Courdesses, M., Amouroux, M., Identifiabilité et identification des systémes á paramétres repartis. Proc. 3rd IFAC Symp. On Control of D. P. S., Toulouse, XXI. 1 (1982).

Courdesses, M., Polis, M. P., Amouroux, M., On identifiability of parameters in a class of parabolic distributed systems. IEEE Trans. Aut. Control, AC-26, 474 (1981).

Davidson, R., MacKinnon, J. G., Estimating the covariance matrix for regression models with ar(1) errors and lagged dependent variables. Economics Letters, 6, (2), 119 (1980).

Grimstaadt, A. A., Mannseth, T., Nonlinearity, Scale, and Sensitivity for a class of parameter estimation problem. Proc. 9th Conf. Inverse Problems in Eng., Muncie, IN (1998).

Harris, T. J., Macgregor, J. F., Wright, J. D., Optimal sensor location with an application to a packed bed tubular reactor. AIChE Journal, 26, (6), 910 (2004).

Hernandez, M. L., Kirubarajan, T., Bar-Shalom, Y., Multisensor resource deployment using posterior Cramer-Rao bounds. IEEE Transactions on Aerospace and Electronic Systems, 40, 399 (2004).

Juncu, G., Floarea, O., Sensitivity analysis of tubular packed-bed reactor by pseudohomogeneous 2D model. AIChE Journal, 41, (12), 2625 (1995).

Kitamura, S., Taniguchi, Y., Parameter estimation of one- and two–dimensional heat conduction processes using a minicomputer. Int. Journal Control, 34, 955 (1981).

Knopman, D. S., Voss, C. I., Behavior of sensitivities in the one-dimensional advection-dispersion equation: implications for parameter estimation and sampling design. Water Resour. Res., 23, (2), 253 (1987).

Kubrously, C. S., Malebranche, H., Sensors and controllers location in distributed systems–A survey. Automatica, 21, (2), 117 (1985).

Kubrusly, C. S., Linear DPS identification in a stochastic environment. R. A. I. R. O. System Anal. and Control, 14, 343 (1980).

Kumar, S., Seinfeld, S. H., Optimal location of measurements in tubular reactors. Chemical Engineering Science, 33, 1507 (1978).

Le Pourhiet, A. M., Le Letty, L., Optimization of sensors location in distributed parameter system identification. Proc. 4th IFAC Symp. On Identification & System Parameter Estimation, Tibilisi, 1581 (1976).

Li, M., Chistofides, P. D., Optimal control of diffusion–convection–reaction processes using reduced–order models. Computers and Chemical Engineering, doi: 10.1016/j.compchemeng. 2007.10.018 (2007).

Ljung, L., System Identification–Theory for the user. 2nd Edition, Prentice Hall, Upper Saddle River, NJ, USA (1999).

Löhner, R., Camelli, F., Optimal Placement of Sensors for Contaminant Detection Based on Detailed 3D CFD. Simulations Engineering Computations. International Journal for Computer-Aided Engineering, 22, (3), 260 (2005).

Morbidelli, M., Varma, A., A generalized criterion for parametric sensitivity application to thermal explosion theory. Chem. Eng. Sci., 43, 91 (1988).

Nahor, H. B., Scheerlinck, N., Van Impe, J. F., Nicolaï, B. M., Optimization of temperature sensor position in a hot wire probe set up for estimation of the thermal properties of foods using optimal experimental design. J. Food. Eng., 57, 103 (2003).

Nakagiri, S., Identifiability of linear systems in Hilbert spaces. SIAM J. Control & Optimization, 21, 501 (1983).

Olanrewaju, M. J., Al-Arfaj, M. A., Estimator-based control of reactive distillation system: Application of an extended Kalman filtering. Chemical Engineering Science 61, 3386 (2006).

Papadimitriou, C., Optimal sensor placement methodology for parametric identification of structural systems. J. Sound Vibration 278, 923 (2004).

Peng, T., Fault Detection Methods Based on Optimal Sensor Location. Hunan: Central South University (2005).

Prasad, V., Vlachos, D. G., Combined multiscale model and informatics based design of experiments: Application to the catalytic decomposition of Ammonia on Ruthenium. Ind. Eng. Chem. Res., (accepted).

Punithakumar, K., Kirubarajan, T., Hernandez, M., Multisensor deployment using PCRLBs, incorporating sensor deployment and motion uncertainties. IEEE Transactions on Aerospace and Electronic Systems 42, 1474 (2006).

Qureshi, Z. H., Ng, T. S., Goodwin, G. C., Optimum experimental design for identification of distributed parameter systems. International Journal of Control, 31, (1), 21 (1980).

Rafajlowics, E., Optimal experiment design for identification of linear distributed-parameter systems: Frequency domain approach. IEEE Transactions on Automatic Control, 28, (7), 806 (1983).

Rafajlowicz, E., Design of experiments for eigenvalues identification in DPS, International. Journal of Control, 34, (6), 1079 (1981).

Rutzler, W., Nonlinear and adaptive parameter estimation methods for tubular reactors. Ind. Eng. Chem. Res., 26, (2), 325 (1987).

Rutzler, W., Nonlinear and adaptive parameter estimation methods for tubular reactors. Ind. Eng. Chem. Res., 26, (2), 325 (1987).

Singh, A. K., Hahn, J., Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems. Industrial and Eng. Chem. Res., 44, (15), 5645 (2005).

Singh, A. K., Hahn, J., Sensor Location for Stable Nonlinear Systems: Placing Multiple Sensors. Proceedings Chemical Process Control 7, Lake Louise, Canada (2006).

Söderström, T., Stoica, P., System Identification. Prentice Hall International, Hemel Hempstead, UK (1989).

Sokollik, F., Zur Parameterschätzung von Systemen mit verteilten Parametern mittels on–line. Verfahren. Messen. Steuern. Regeln., 19, 84 (1976).

Sorenson, H. W., Parameter estimation: principles and problems. New York, M. Dekker, 183 (1980).

Storch, R. B., Pimentel, L. C. G., Orlande, H. R. B., Identification of atmospheric boundary layer parameters by inverse problem. Atmospheric Environment, 41, (7), 1417 (2007).

Tilden, J. W., Constanza, V., McRae, G. J., Seinfeld, J. H., Sensitivity analysis of chemically reacting systems. In: K. H. Ebert, P. Deuflhard and W. Jager, Editors, Modeling of Chemical Reaction Systems, Springer, New York, 69 (1981).

Uciński, D., Bogacka, B., Construction of T-optimum designs for multiresponse dynamic models. In COMPSTAT 2002 (Berlin), Heidelberg: Physica, pp. 267 (2002).

Uciński, D., Demetriou, M., An approach to the optimal scanning measurement problem using optimum experimental design. Proceeding of the 2004 American Control Conference, June 30-July 2, Boston, Massachusetts (2004).

Uciński, D., Measurement Optimization for Parameter Estimation of Distributed Systems. Technical University Press, Zielona Gora (1999).

Uciński, D., On the determination of moving sensor policies for parameter estimation of distributed systems. 11th Mediterranean IEEE Conference on Control and Automation MED'03, June 18-20, Rhodes, Greece (2003).

Uciński, D., Optimal measurement method for distributed parameter system identification. CRC Press. USA (2005).

Uciński, D., Optimal sensor location for parameter estimation of distributed processes. International Journal of Control, Vol. 73, No. 13, 1235 (2000a).

Uciński, D., Optimal sensor location for parameter identification of distributed systems. Applied Mathematics and Computer Science, 2, (1), 119 (1992).

Uciński, D., Optimization of sensors' allocation strategies for parameter estimation in distributed systems. System Analysis Modell. Simul., Vol. 37, 243 (2000b).

Vajda, S., Rabitz, H., Parametric sensitivity and self-similarity in thermal explosion theory. Chem. Engng. Sci., 47, 1063 (1992).

Van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M., Smilde, A. K., Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. Chemical Engineering Science, 55, (4), 827 (2000).

Vande Wouwer, A., Point, N., Porteman, S., Remy, M., An approach to the selection of optimal sensor locations in distributed parameter systems. Journal of Process Control, 10, (4), 291 (2000).

Venkateswarlu, C. B. J., Kumar, B. J., Composition estimation of multi component reactive batch distillation with optimal sensor configuration. Chemical Engineering Science, 61, 5560 (2006).

Waterhouse, T. H., Eccleston, J. A., Duffull, S. B., Optimal design criteria for discrimination and estimation in nonlinear models. J. Biopharmaceutical Statistics, 19, (2) 386 (2009).

Yen, W. W-G., Yoon, Y. S., Aquifer parameter identification with optimum dimension in parameterization. Water Resour. Res., 17, (3), 664 (1981).

Zamprohna, E., Barolo, M., Seborg, D. E., Optimal Selection of Sensors Inputs for Batch Distillation Columns Using Principal Component Analysis. Journal of Process Control, 15, 39 (2005).

Zhu, Z., Stein, M. L., Spatial sampling design for parameter estimation of the covariance function. Journal of Statistical Planning and Inference, 134, (2), 583 (2005).

Optimal location of measurements for parameter estimation of distributed parameter systems. Alaña, J., Theodoropoulos, C., Computer and Chemical Engineering, 35(1), 106(2011).

Computers and Chemical Engineering 35 (2011) 106–120

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Optimal location of measurements for parameter estimation of distributed parameter systems

Jorge E. Alaña[a,b], Constantinos Theodoropoulos[a,*]

[a] School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, UK
[b] Facultad de Ingeniería, Universidad Del Zulia, Maracaibo, Venezuela

ARTICLE INFO

Article history:
Received 3 May 2009
Received in revised form 7 April 2010
Accepted 16 April 2010
Available online 24 April 2010

Keywords:
Sensitivity coefficients
Gram determinant
D-optimal design
Measurement points
Max–min optimisation
Proper Orthogonal Decomposition

ABSTRACT

Systematic methodologies for the optimal location of spatial measurements, for efficient estimation of parameters of distributed systems, are investigated. A review of relevant methods in the literature is presented, and a comparison between the results obtained with three distinctive existing techniques is given. In addition, a new approach based on the Proper Orthogonal Decomposition (POD), to address this important problem is introduced and discussed with the aid of illustrative benchmark case studies from the literature. Based on the results obtained here, it was observed that the method based on the Gram determinant evolution (Vande Wouwer et al., 2000), does not always produce accurate results. It is strongly dependent on the behaviour of sensitivity coefficients and requires extensive calculations. The method based on max–min optimisation (Alonso, Kevrekidis, Banga, & Frouzakis, 2004) assigns optimal sensor locations to the positions where system outputs reach their extrema values; however, in some cases it produces more than one optimal solution. The D-optimal design method, Uciński (2003, June 18–20), produces as results the optimal number and spatial positions of measurements based on the behaviour (rather than the magnitude) of the sensitivity functions. Here we show that the extrema values of POD modes can be used directly to compute optimal sensor locations (as opposed e.g. to Alonso, Kevrekidis, et al., 2004, where PODs are merely used to reduce the system and further calculations are needed to compute sensor locations). Furthermore, we demonstrate the equivalence between the extrema of POD modes and of sensitivity functions. The added value of directly using PODs for the computation of optimal sensor locations is the computational efficiency of the method, side-stepping the tedious computation of sensitivity coefficient Jacobian matrices and using only system responses and/or experimental results directly. Furthermore, the inherent combination of model reduction and sensor location estimation in this method becomes more important as the complexity of the original distributed parameter system increases.

1. Introduction

Many contributions on optimal sensor location for state estimation have been published in the last three decades. In this section, we first present an overview of these methods and then we focus on works dealing with the closely related problem of sensor placement for parameter estimation. In the works addressing the former problem, numerous metrics have been developed for the spatial placement of sensors in a distributed parameter system (DPS). Most of the techniques rely on exhaustive searches over a pre-defined set of candidates (e.g. Keller & Bovin, 1992). These approaches, while valid for a small number of locations, become impractical when the number of possible position candidates increases. Even with the current increase in computational power there is an essential need for fast computations in terms of sensor placement to address operational changes, variations in production, etc.

Omatu, Seinfeld, and Soeda (1978) presented an optimality criterion based on the error covariance matrix of Kalman filters. More recently Olanrewaju and Al-Arfaj (2006) presented a state estimation scheme based on Kalman filters for reactive distillation systems without extensively addressing the sensor selection problem. Harris, Macgregor, and Wright (1980), Kumar and Seinfeld (1978), and Colantuoni and Padmanabhan (1977), have used analogous metrics based on the trace and the determinant of error covariance. Some techniques are based on the qualitative knowledge of the system variables in addition to the variance of the state prediction errors (Jørgensen, Goldschmindt, & Clement, 1984). Morari and Stephanopoulos (1980), and Morari and O'Dowd (1980) employed non-stationary noise models and used criteria based on minimising the estimation error by the unobservable sub-space.

* Corresponding author at: School of Chemical Engineering and Analytical Science, The University of Manchester, Sackville St, Manchester M60 1QD, UK.
E-mail address: k.theodoropoulos@manchester.ac.uk (C. Theodoropoulos).

0098-1354/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2010.04.014

Another type of metric used for sensor location is based on the observability matrix and/or the observability gramian. Muller and Weber (1972) used the smallest eigenvalue, the determinant, and the trace of the inverse of the gramian. Dochain, Tali-Mammar, and Babary (1997) and Damak, Babary, and Nihtila (1992) presented the condition number of the observability matrix, and Van den Berg, Hoefsloot, Boelens, and Smilde (2000) made use of the trace and the spectral norm of the observability gramian as a criterion for sensor location in their work. Waldraff, Dochain, Bourrel, and Magnus (1998) have presented results considering the smallest singular value and the condition number of the observability matrix. Alvares, Romangnoli, and Stepanopoulos (1981) and Romagnoli, Alvares, and Stepanopoulos (1981) presented a technique that considers variable measurement structures to solve the problem of sensor location. More recently Hernandez, Kirubarajan, and Bar-Shalom (2004) and Punithakumar, Kirubarajan, and Hernandez (2006) have developed metrics based on the posterior Cramer–Rao lower bound to quantify and ultimately to control the accuracy of state estimation in conjunction with optimization techniques to control the measurement process.

In addition to the metrics mentioned above, there are contributions that take into account measurement cost, sensor failure and uncertain sensor movement in addition to system information (Fahroo & Demetriou, 2000; Peng, 2005; Punithakumar et al., 2006). Among the works dealing with nonlinear observability analysis for sensor location, which is computationally expensive, those by Hermann and Kerner (1977), Lopez and Alvares (2004), Damak et al. (1992), Singh and Hahn (2006) and Van den Berg et al. (2000) should be noted. The relevant measures introduced reduce to a criterion based on the linear observability matrix, if the system is linear and time invariant. For nonlinear systems, the nonlinear observability presented by Scherpen (1993) is another alternative. Georges (1995) and Sumana and Venkateswarlu (2009) have used nonlinear observability gramians for sensor selection and positioning. However, the applicability of this method is limited to relatively low-order systems due to numerical difficulties.

Vande Wouwer, Point, Poterman, and Remy, (2000), Alonso, Frouzakis, and Kevrekidis (2004) and Alonso, Kevrekidis, Banga, and Frouzakis (2004) presented other significant contributions on sensor placement for nonlinear systems, based on the Gram determinant evolution in the spatial domain and on a max–min optimisation process, respectively. The first contribution discusses the maximisation of the Gram determinant (which is a measure of the linear independence of the sensitivity coefficients evaluated at sensor locations) as a procedure that guarantees that the parameters are identifiable and that the correlation between the sensor outputs is minimised. The form of the criterion itself resembles the D-optimality criterion proposed by Qureshi, Ng, and Goodwin (1980) and Rafajlowicz (1978), but the counterpart of the Fisher Information Matrix (FIM) (see e.g. Walter & Prozanto, 1997) is much larger, which suggests that the approach involves calculations that are more cumbersome. Alonso, Kevrekidis, et al. (2004) state that the most appropriate type (and number) of measurements is found by the solution of a max–min optimisation problem, where the number of measurement locations has been defined by the user. All three methods are discussed in detail in the following sections.

While there are ample contributions on optimal sensor location, most works deal with state estimation. There are, however, very few contributions that address optimal sensor location for parameter estimation, some reported by Vande Wouwer et al. (2000) and Basseville, Benveniste, Moustakides, and Rougee (1987). This is mainly because a direct extension of the methodologies for sensor location for parameter estimation from the results obtained from state estimation is not straightforward and has not been pursued. The problem of sensor location for parameter identification is essentially different from the optimal measurements problem, because in the first case the current state usually depends strongly nonlinearly on unknown parameters (Chmielewski, Tasha, & Manousiouthakis, 2002).

Most of the existing optimality criteria for sensor location for parameter estimation are based on scalar measures of the FIM (Mehra, 1974). For example, the modified E-criterion (Mehra, 1974; Walter & Prozanto, 1990), which minimises the ratio of the largest to the smallest eigenvalue of the FIM, was used by Nahor, Scheerlinck, Van Impe, and Nicolaï (2003) to compute optimal temperature sensor positions for food processes. Heredia-Zavoni, Montes-Iturizzaga, and Esteva (1998) have used the trace of the inverse of the FIM to calculate the minimum of the Bayesian loss function. Also, the determinant of the FIM has been used by Qureshi et al. (1980), Vande Wouwer et al. (2000) and Papadimitriou (2004) who have given an asymptotic approximation for the information entropy and have used a sequential approach to optimally place sensors one at a time.

Using a related approach relying on principal component analysis of the output sensitivity matrix, Li, Henson, and Kurtz (2004) computed the best set of parameters to be estimated if a set of measurements positions is already given. All the techniques require computation of the parametric output sensitivity coefficients based on local sensitivity analysis. Therefore, the results obtained by these methods might not capture the nonlinearity and may only be suitable for small changes in the parameters. Singh and Hahn (2005) have computed sensor location with information derived from observability covariance matrices combined with already existing measures they proposed, either for state or parameter estimation, and computed the degree of observability of a nonlinear system over a certain operating region (Zamprohna, Barolo, & Seborg, 2005). Furthermore, Waterhouse, Eccleston, and Duffull (2009) have developed a combination of T-optimal designs (Uciński & Bogacka, 2002), which are computationally tedious even for relatively simple systems with D-optimality criteria to combine parameter estimation and model discrimination for dynamic pharmacokinetic models.

In summary, the state of the art in the sensor placement indicates that more attention should be paid to the case of sensor position for parameter estimation, especially for DPS. From an engineering point of view, the use of the existing scarce methods is restricted due to computational and/or realising difficulties (Korbicz, Zgurovsky, & Novikov, 1988; Löhner & Camelli, 2005; Malebranche, 1988; Waterhouse et al., 2009).

Methods to determine the optimal sensor positions of distributed parameter systems are especially helpful in research and engineering practice. This is an essential problem especially for parameter estimation purposes. Unfortunately, there still exists some distance between theoretical/computational researchers, who explore models of systems, and experimental researchers. This incentivises the improvement of methods which allow the identification of optimal measurement location techniques accounting for the spatially changing nature of the system. Additionally, practice dictates that these methods should use information collected from the process by simple experimental procedures. This is the cornerstone of the approach presented here, which uses direct system responses to determine sensor locations. In this paper we introduce the use of the extrema values of the system's sensitivity coefficients and/or POD modes to compute the optimal spatial positions (and number) of experimental measurements for parameter estimation. The rest of the paper is organised as follows: In Section 2, three existing distinct techniques for optimal sensor placement are reviewed and discussed. In Section 3 our POD-based method for sensor placement is discussed, and a necessary and sufficient condition for the equivalence between the extrema of POD modes and of sensitivity coefficients is given. Section 4 presents three illustrative case studies from the literature used for benchmarking purposes,

and Section 5 presents numerical results and comparisons between the three methodologies reviewed and our proposed technique.

2. Sensor location methods

In this section, three relevant techniques from the literature used to determine optimal sensor positions are reviewed and their theoretical backgrounds are given.

2.1. Optimal experiment design

Optimal experiment design can be defined as the set of conditions under which an experiment should be conducted in order to maximise the accuracy of the obtained results, i.e. the system's response can be measured in such a way that its parameters can be best estimated (Emery & Nenarokomov, 1998; Rafajlowicz, 1983).

In order to understand this concept, let us consider a general DPS. Let $\Omega \subset R^2$ be a bounded simply connected open domain with a sufficiently smooth boundary, $\partial\Omega$. A given DPS can be represented by a system of n (possibly nonlinear) partial differential equations (PDEs) as follows (Uciński, 1992, 2005):

$$G\left(x, u, t, \theta, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \frac{\partial^2 u}{\partial {x_1}^2}, \frac{\partial^2 u}{\partial {x_2}^2}\right) = 0, \quad x \in \Omega \tag{1}$$

where $x = (x_1, x_2) \in \tilde{\Omega}$ is the vector of spatial coordinates, $u = u(x, t)$ the vector of state variables $\in R^n$ and G some function mapping its arguments to R^n, possibly including terms accounting for *a priori* known forcing inputs. The boundary conditions are:

$$\omega\left(x, u, t, \theta, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) = 0, \quad x \in \Omega \tag{2}$$

Here ω is some known function that maps its arguments to R^n. The model above contains a vector of unknown parameters $\theta \in R^m$ to be estimated based on the observations:

$$\eta_{ij} = u(x^i, \theta) + \varepsilon_{ij}, \quad i = 1, \ldots, p, \ j = 1, \ldots, r_i \tag{3}$$

made at point x^i, where ε_{ij} denotes the noise in measurements satisfying a Gaussian distribution:

$$E\{\varepsilon_{ij}\} = 0, \qquad E\{\varepsilon_{ij}\varepsilon_{qs}^T\} = \delta_{iq}\delta_{js}C(x^i) \tag{4}$$

δ_{ij} being the Kronecker delta and $C(x^i) \in R^{n\times n}$ a known positive-definite matrix. Replicated measurements can be admitted, i.e., some values x^i may appear several times in the optimal solution (an unavoidable consequence of independent measurements). Then it is sensible to distinguish only the components of the sequence $x^1, \ldots, x^N$, which are different and, if there are p such components, to re-label them as $x^1, \ldots, x^p$ while introducing $r^1, \ldots, r^p$ as the corresponding numbers of replications.

It is assumed that some, albeit rough, a priori estimate of the parameter vector θ^0 is available e.g. from preliminary experiments. The goal here is to select the state sampling policy, which will maximise the accuracy of the parameter estimates from data generated in the corresponding new experiments. It has been shown that this task corresponds to the maximisation of a scalar concave function of the FIM whose inverse gives a lower bound on the covariance matrix of the estimates (Atkinson & Donev, 1992; Chmielewski et al., 2002; Emery & Nenarokomov, 1998; Rafajlowicz, 1981, 1983). The average per observation (or normalised) FIM takes the following form:

$$M(\xi_N) = \sum_{i=1}^{p} w_i F^T(x^i) C^{-1}(x^i) F(x^i) \tag{5}$$

where $w_i = r_i/N$, $N = \sum_{i=1}^{p} r_i$ and $F(x^i) = \partial y(x^i, \theta_0)/\partial\theta$ is the Jacobian matrix consisting of the sensitivity coefficients. The collection of variables, ξ_N that define the exact design of the experiment is given below:

$$\xi_N = \begin{Bmatrix} x^1, & x^2, & \ldots & , x^p \\ w_1, & w_2, & \ldots & , w_p \end{Bmatrix} \tag{6}$$

The proportion w_i of observations made at x^i can be considered as the percentage of experimental effort spent at that location. It can be assumed that w_i's can be any real numbers in the interval [0, 1] such that $\sum_{i=1}^{p} w_i = 1$ and in terms of the design they can be considered as probability distributions on $\{x^1, x^2, \ldots, x^p\}$. The purpose of the optimal experimental design can be then defined as the search for a design $\xi^* \in \Xi(X)$ which satisfies (Uciński, 2005; Walter & Prozanto, 1990):

$$\Psi[M(\xi^*)] = \max_{\xi \in \Xi(X)} \Psi[M(\xi)] \tag{7}$$

Here $\Xi(X)$ is the set of all admissible designs (i.e. all probability distributions on the spatial domain X, where measurements are allowed). In the literature various choices for Ψ are presented (Atkinson & Donev, 1992; Uciński & Demetriou, 2004; Uciński, 2005; Walter & Prozanto, 1990). The most popular ones are:

- The D-optimality criterion:

$$\Psi[M(\xi)] = \det[M(\xi)] \tag{8}$$

- The A-optimality criterion:

$$\Psi[M(\xi)] = -\text{trace}\left[(M(\xi))^{-1}\right] \tag{9}$$

A very important step in the design of optimal sensor locations is to use an effective procedure for the computation of the sensitivity coefficients, which are necessary when determining the FIM elements, defined by Eq. (5). Obviously the construction of such Jacobian matrices is a "computationally expensive" step, which makes the computational cost of the corresponding methods relatively "high" (see also Table 7 in the results section). Furthermore, the model equations are explicitly needed for the computation of sensitivities (note that numerical differentiation in the absence of explicit models when possible will significantly increase computational costs).

2.2. Gram determinant-based method

Point, Vande Wouwer, and Remy (1996) proposed a method, which is a direct application of the idea of a measure of the linear independence between the sensitivity functions. In many cases, the solution of a least-squares problem is not unique and the parameters cannot be estimated from the experimental data considered. It is therefore necessary to determine experimental data so that all the sensitivity functions are linearly independent. A measure of the linear independence of the P sensitivity functions $S_{u(x,t),\theta j}$ for $j = 1, \ldots, P$ (evaluated at point x) can be defined as the minimum value of the quadratic form (10) where α_j are real parameters subject to the condition $\sum_{i=1}^{P} a_i^2 = 1$:

$$m(x) = \min_{a_i}\left(\int_0^{\Gamma} (a_1 S_{u(x,t),\theta_1} + a_2 S_{u(x,t),\theta_2} + \cdots + a_p S_{u(x,t),\theta_p})^2\, dt\right) \tag{10}$$

Here $[0, \Gamma]$ is the time span of interest. This optimality criterion is based on a measure of independence between the quantities of interest at the several measurement points, i.e. the sensor responses in the context of state estimation or the sensitivity

functions in the context of parameter estimation. The sensitivity functions are linearly dependent if and only if $m(x)$ is equal to zero or, equivalently, if and only if the Gram determinant $g(x)$ vanishes, i.e.:

$$g(x) = \int_0^{\Gamma} \begin{bmatrix} (S_{u(x,t),\theta_1})^2 & \cdots & S_{u(x,t),\theta_1} S_{u(x,t),\theta_P} \\ \vdots & \vdots & \vdots \\ S_{u(x,t),\theta_P} S_{u(x,t),\theta_1} & \cdots & (S_{u(x,t),\theta_P})^2 \end{bmatrix} dt = 0 \tag{11}$$

Calculating the evolution of the Gram determinant (11) in the spatial domain allows for the determination of the regions where the sensitivity functions are linearly independent and where the measurement points should be located. Alternatively, input signals can be computed by maximising the Gram determinant in the interval where the sensor is located. The determinant of a positive semi-definite matrix increases with its diagonal elements and decreases with its off-diagonal elements. The maximisation of $\det[g(x)]$, with respect to the sensor locations $x_1, \ldots, x_M$, provides the spatial positions where the elements of the matrix in (11) take their largest values and are independent of one another (Vande Wouwer et al., 2000). This criterion is closely related to the one proposed by Qureshi et al. (1980), based on information theory, that makes use of the determinant of the FIM which, has a form similar to that of the Gram determinant. Obviously here, too, the computation of sensitivity matrices makes the cost of this method also "high".

2.3. Max–min optimisation procedure

Alonso, Kevrekidis, et al. (2004) reported a very interesting approach for sensor location. In this work the performance of the observers can be improved by placing measurements at locations that maximise the eigenvalues of the QQ^T matrix, where $Q^T = P_m \Phi$, and P_m is an operator that projects any vector $\phi_i \in R^n$ on m of its n coordinates (m being the size of the sub-space of the measurements). The minimum eigenvalue of QQ^T is then used to select the appropriate spatial sensor arrangement, for a given number of sensors m. Formally, this can be stated as follows:

$$\max_{P_m} \min_{i=1,\ldots,k} \lambda_i(QQ^{\mathrm{T}}) \tag{12}$$

Here the operator P_m indicates the spatial sensor location. This problem can be solved by exhaustive search among all possible m combinations of the n coordinates. However, this approach, although feasible for small problems, becomes unsuitable for most cases of practical interest due to the high dimensionality of the search space. Alternatively, by taking advantage of the underlying structure of Q, a systematic approach was used to approximate the solution. The approach is based on the following facts:

1. The effect of P_m^T on Φ^T is that of deleting elements of the basis vectors ϕ_i at the positions where the columns of P_m are zeros.
2. The scalar products of the resulting sub-vectors $P_m\phi_i$ are the diagonal elements of QQ^T.
3. The eigenvalues of matrix QQ^T are located inside circles centred at the positions given by the diagonal elements with radii satisfying:

$$r_i = \sum_{j \neq i}^{n} \left| (QQ^T)_{ij} \right| \tag{13}$$

Thus, when the radii are much smaller than the diagonal elements (denoted as s_i), the solution of (12) approximates that of maximising the minimum diagonal element. Formally, this can be written as:

$$\max_{P_m} \min(s_1, \ldots, s_k) \tag{14}$$

s_k being of the form:

$$s_k = \sum_{j=1}^{m} (\phi_{kj})^2 \tag{15}$$

The optimal sensor location computed with this method provides a reliable field reconstruction from a limited number of measurements. For large systems a low-dimensional representation of the original system (based on the POD method) is used, combined with a guided search algorithm that minimises orthonormality distortions. This method uses directly system outputs and avoids the computation of large sensitivity matrices. However, for large-scale problems first a model reduction (POD-based) step is performed and additional calculations on the reduced system are required, hence the computational cost for this method is characterised as "moderate" (see Table 7). Furthermore, the number of sensors needs to be pre-specified.

3. Methodology

In this section the Proper Orthogonal Decomposition method is first described. Since our objective is to demonstrate that the optimal locations to place sensors are at the points where POD basis functions reach their extrema values, we then show the equivalence between the extrema of the POD modes and those of the sensitivity coefficients of the system.

3.1. Proper Orthogonal Decomposition

The POD method has received much attention in the last two decades as a tool to simulate, optimise and control complex large-scale systems. The underlying idea is to use a reliable number of responses from the system (snapshots in time), to compute a small set of global basis functions (the POD modes) that can accurately capture the spatio-temporal behaviour of the system.

Several researchers including Loeve (1945), Karhunen (1946), Pougachev (1953), and Obukhov (1954) independently proposed the POD method. Some mathematical theories of POD can be found in articles by Aubry, Lian, and Titi (1993) and Graham, Lane, and Luss (1993). The POD technique has been applied to numerous applications, such as the attraction of spatial organised motions in fluid flows (Berkooz, 1992; Lumley, 1970). Lumley (1967), Lumley (1970), Aubry, Holmes, Lumley, and Stone (1988), Sirovich (1991), Berkooz, Holmes, Lumley, and Mattingly (1997) and Chambers, Adrian, Moin, Stewart, and Sung (1988) have adapted the POD technique to study turbulent flows. Other applications of POD include complex geometry flows (e.g. Deane, Kevrekidis, Karniadakis, & Orszag, 1991) channel flows (Ball, Sirovich, & Keefe, 1991; Moin & Moser, 1989), square-duct flows (Reichert, Hatay, Biringers, & Husser, 1994) and shear flows (Kirby and Sirovich, 1990; Rajaee, Karlson, & Sirovich, 1994). Other researchers have applied POD to the Burgers' equation (Chambers et al., 1988) and to the Ginsburg–Landau equation and the Bérnard convection (Sirovich, 1989). The nonlinear control of DPS with the use of POD has been extensively studied (e.g. Ly & Tran, 2005; Shvartsman et al., 2000) while POD has also been used for the optimisation of nonlinear DPS (Bendesrky & Christofides, 2000). Other interesting applications of the POD technique include the characterisation of human faces by Kirby, Boris, and Sirovich (1990), and image recognition by Hilai and Rubinstein (1994).

In this work the snapshots approach was used for the determination of the POD basis functions following the procedure explained

below (Newman, 1996; Sirovich, 1987). Let $u(x,t)$ be a given output field and $\{u(x,t_i)\}_{i=1}^{N}$ be the corresponding output fields at N different times steps t_k, i.e. the snapshots. Next $u(x,t)$ is decomposed as follows:

$$u(x,t) = u_m(x) + v(x,t) \tag{16}$$

with u_m being the average snapshot, $u_m(x) = (1/N)\sum_{k=1}^{N} u(x,t_k)$. The spatial correlation matrix C can be defined as:

$$C_{ij} = \frac{1}{N}\int_{\Omega} v^{(i)}v^{(j)}\, d\Omega \tag{17}$$

where $v^{(i)} = v(x,t_i)$. The eigenvalues and eigenvectors of C are then computed by the following eigenproblem:

$$CA = \lambda A \tag{18}$$

The M eigenvectors corresponding to the M largest eigenvalues capturing most (in practice more than 99%) of the system's energy are then used to compute the POD basis vectors $\varphi_k(x)$:

$$\varphi_k(x) = \sum_{m=1}^{M} A_m^{(k)} v^{(m)}(x) \tag{19}$$

Here $A_m^{(k)}$ is the mth component of the kth eigenvector A.

3.2. Sensitivity coefficient and POD modes extrema equivalence

In this section it is proven that the extrema values of the sensitivity coefficients and the POD modes coincide at the same locations. It should be noted that the same reasoning can be followed using any other Ansatz.

Starting from the following approximation of the system response, where $\alpha_k(t)$ are M the time coefficients calculated by projecting the original system onto the computed M POD modes:

$$u(x,t) = \sum_{k=1}^{M} \alpha_k(t)\varphi_k(x) \tag{20}$$

Taking derivatives with respect to parameter θ_i:

$$\frac{\partial u(x,t)}{\partial \theta_i} = \sum_{k=1}^{M} \frac{\partial \alpha_k(t)}{\partial \theta_i}\varphi_k(x) + \sum_{k=1}^{M} \alpha_k(t)\frac{\partial \varphi_k(x)}{\partial \theta_i} \tag{21}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial u(x,t)}{\partial \theta_i}\right) = \sum_{k=1}^{M} \frac{\partial \alpha_k(t)}{\partial \theta_i}\frac{\partial \varphi_k(x)}{\partial x} + \sum_{k=1}^{M} \alpha_k(t)\frac{\partial}{\partial x}\left(\frac{\partial \varphi_k(x)}{\partial \theta_i}\right) \tag{22}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial u(x,t)}{\partial \theta_i}\right) = \frac{\partial S_{u,\theta_i}}{\partial x} = \sum_{k=1}^{M} \frac{\partial}{\partial \theta_i}\left(\alpha_k(t)\frac{\partial \varphi_k(x)}{\partial x}\right) \tag{23}$$

Let

$$\frac{\partial}{\partial x}\left(\frac{\partial u(x,t)}{\partial \theta_i}\right) = \frac{\partial S_{u,\theta_i}}{\partial x},$$

be the sensitivity coefficient of variable $u(x,t)$ with respect to parameter θ_i, then:

$$\frac{\partial S_{u,\theta_i}}{\partial x} = \sum_{k=1}^{M} \frac{\partial}{\partial \theta_i}\left(\alpha_k(t)\frac{\partial \varphi_k(x)}{\partial x}\right) \tag{24}$$

Then if

$$\frac{\partial \varphi_k(x)}{\partial x} = 0 \tag{25a}$$

$$\frac{\partial S_{u,\theta_i}}{\partial x} = 0 \tag{25b}$$

To prove that the condition is also sufficient let us assume that Eq. (25b) holds while

$$\frac{\partial \varphi_k(x)}{\partial x} \neq 0 \tag{26}$$

From Eq. (24) we have

$$\frac{\partial S_{u,\theta_i}}{\partial x} = \sum_{k=1}^{M} \frac{\partial}{\partial \theta_i}\left(\alpha_k(t)\frac{\partial \varphi_k(x)}{\partial x}\right) = \frac{\partial}{\partial \theta_i}\sum_{k=1}^{M}\left(\alpha_k(t)\frac{\partial \varphi_k(x)}{\partial x}\right) = 0 \tag{27}$$

Since both $\alpha_k(t)$ and $\partial\varphi_k(x)/\partial x$ are functions of θ_i, for Eq. (27) to be satisfied:

$$\sum_{k=1}^{M}\left(\alpha_k(t)\frac{\partial \varphi_k(x)}{\partial x}\right) = 0 \tag{28}$$

The empirical basis functions $\varphi_k(x)$ are given by Eq. (19). Hence,

$$\frac{\partial \varphi_k(x)}{\partial x} = \sum_{m=1}^{M}\left(A_m^{(k)}\frac{\partial v^{(m)}(x)}{\partial x}\right) \tag{29}$$

From Eqs. (28) and (29) we get:

$$\sum_{k=1}^{M}\left(\alpha_k(t)\sum_{m=1}^{M}\left(A_m^{(k)}\frac{\partial v^{(m)}(x)}{\partial x}\right)\right) = 0 \tag{30}$$

Re-arranging Eq. (30) we have

$$\sum_{m=1}^{M}\left(\frac{\partial v^{(m)}(x)}{\partial x}\sum_{k=1}^{M}(\alpha_k(t)A_m^{(k)})\right) = 0 \tag{31}$$

For Eq. (31) to hold if Eq. (26) is true, Eq. (32) below should also hold:

$$\sum_{k=1}^{M}\alpha_k(t)A_m^{(k)} = 0 \quad \forall m = 1,\ldots,M \tag{32}$$

It is, however, obvious that if this was the case eigenvectors, A_m would be linearly dependent. Therefore Eq. (26) cannot be true. This concludes our proof. Hence, the extrema values of the sensitivity coefficients and those of the POD modes, capturing most of the energy of the system, coincide at the same spatial locations.

3.3. Parameter estimation

Once the sensor locations have been determined, the parameter estimation problem can be viewed as matching the model to the real system through the minimisation of an error criterion over a set of admissible parameters. This can be defined as

$$J(\theta) = \frac{1}{\Gamma}\frac{1}{m}\int_0^{\Gamma}\sum_{i=1}^{m}\left(u_e(x_i,t) - u_p(x_i,t)\right)^2 dt \tag{33}$$

where $u_p(x_i,t)$ is defined as the predicted response of the model at location x_i, $u_e(x_i,t)$ as the experimental response at the same location and m as the number of optimal measurement locations determined by the different methods studied in this work.

It is well known that minimising (33) is challenging for numerous reasons. The existence of a solution is not certain, particularly if the observed data contains errors or if the model is grossly incorrect. It is unusual that any parameter set can accurately match the experimental data used, especially when these data are contaminated with noise. The presence of noise can promote difficulties

during the optimisation process causing spurious local minima and discontinuities. The gradient-based methods are the most affected of all. In this work, several different optimisation techniques (already existing in the Matlab® library) were used, and it was observed that the method of Nelder–Mead produced the best results. This can be explained by the fact that using this technique, poor gradient approximations are not a problem, and continuity and differentiability of the objective function are not required.

4. Case studies

To compare the existing methods for sensor location described above and our method based on the extrema values of POD modes, illustrative DPS examples from the literature are employed.

Example 4.1. This problem is described by the following one-dimensional PDE:

$$\frac{\partial u}{\partial t} = \theta_1 \frac{\partial^2 u}{\partial x^2} \quad t > 0, \ \ x \in (0, 1) \tag{34}$$

With the boundary and initial conditions:

$$\frac{\partial u}{\partial x} = -\theta_2 (u_a - u) \quad t > 0, \ \ x = 0 \tag{35a}$$

$$\frac{\partial u}{\partial x} = \theta_3 (u_a - u) \quad t > 0, \ \ x = 1, \ \text{and} \ \ u_a = 600 \tag{35b}$$

$$u = u_0(x) = 500, \quad t = 0, \ \ x \in [0, 1] \tag{35c}$$

Here $0 \le t \le 1$. This simple example by Vande Wouwer et al. (2000), is used to demonstrate in an intuitive way the effectiveness and some of the characteristics of the experiment design procedure and optimal sensor location analysed. Here, the problem consists of the determination of the best sensor location for the estimation of

1. *Example 4.1a*: The two parameters, θ_2 and θ_3, for the particular case of θ_1 known and kept constant.
2. *Example 4.1b*: The three parameters, θ_1, θ_2 and θ_3.

Example 4.2. Consider a nonlinear system described by the following PDE, solved and analysed in detail by Uciński (2005):

$$\frac{\partial u}{\partial t} = \theta_1 \frac{\partial^2 u}{\partial x^2} - \theta_2 u^3, \quad t \in (0, t_f], \ \ x \in (0, 1) \tag{36}$$

Subject to the conditions:

$$u(x, 0) = \theta_3 \Psi(x), \quad x \in (0, 1) \tag{37a}$$

$$u(0, t) = u(1, t) = 0, \quad t \in (0, t_f] \tag{37b}$$

where $t_f = 0.8$ and

$$\Psi(x) = \sin(\pi x) + \sin(3\pi x) \tag{38}$$

The problem consists of the determination of the best sensor location for the estimation of θ_1, θ_2 and θ_3.

Example 4.3. This system involves coupled mass and heat transfer mechanism, and it is described by the following couple system of partial differential equations with nonlinear source terms. It is represents realistic chemical engineering reaction systems (such as tubular reactors) and it can produce interesting nonlinear parametric behaviour including different types of bifurcations, such as folds and oscillations:

$$\frac{\partial u_1(x, t)}{\partial t} = \frac{\partial^2 u_1(x, t)}{\partial x^2} - \theta_1 \frac{\partial u_2(x, t)}{\partial x} - \theta_2 u_1(x, t) u_2(x, t) \tag{39a}$$

$$\frac{\partial u_2(x, t)}{\partial t} = \frac{\partial^2 u_2(x, t)}{\partial x^2} - \theta_3 \frac{\partial u_2(x, t)}{\partial x} + \theta_4 u_1(x, t) \tag{39b}$$

Subject to the conditions:

$$u_1(x, 0) = u_2(x, 0) = 1, \quad x \in (0, 1) \tag{40a}$$

$$u_1(0, t) = 1 \qquad \frac{\partial u_2(0, t)}{\partial x} = 0, \quad t \in (0, t_f] \tag{40b}$$

$$u_1(1, t) = u_2(1, t) = 1, \quad t \in (0, t_f] \tag{40c}$$

This example was solved for time $t \in [0, 120]$. This time span was used because the dynamic behaviour of sensitivity functions of u_1 with respect to θ_1 and θ_2 is quite complex. This can be seen in Fig. 3, where S_{u_1,θ_1} and S_{u_1,θ_2} change shape after $t = 10$.

The model equations were solved using the Matlab® partial differential equation solver "pdepe" which solves initial-boundary value problems for systems of parabolic and elliptic partial differential equations in one spatial dimension. The solver converts the partial differential equations to ordinary differential equations using a second-order accurate spatial discretisation based on a set of nodes specified by the user. Time integration is then performed with a multistep variable-order method based on the numerical differentiation formulae. This method was used to solve all the case studies considered in this work.

The spatial domains were divided into 21 equidistant intervals, and the partial differential equations were solved simultaneously with the equations describing the transient behaviour of the sensitivity coefficients. n time "reporting" intervals were considered. Hence, the spatio–temporal distributions of the systems' responses and of the sensitivity coefficients are obtained. Three algorithms were studied to estimate the optimal positions of the measurements; using the Gram determinant (Point et al., 1996), the D-optimal design theory (Uciński, 2005) and the method based on max–min optimisation (Alonso, Kevrekidis, et al., 2004). For all the case studies, these three methods were used, and compared with our POD-based technique.

In addition using the method of snapshots (Sirovich, 1987), the POD modes $\phi_i(x)$ for each system were calculated, implementing the procedure discussed above, their extrema values satisfying $d\phi_i(x)/dx = 0$ were computed and the corresponding locations x were evaluated for optimal sensor placement.

5. Results and discussion

5.1. Sensor locations

The spatial variability of sensitivities has a significant impact on parameter estimation and sampling design for studies of distributed parameter systems. Information about a physical parameter will be most accurately gained at points in space with a high sensitivity to the parameter.

The sensitivity coefficients, S_{u,θ_i}, calculated for each parameter corresponding to the examples considered in this study are depicted in Figs. 1–3. These sensitivities were normalised in order to observe clearly the extrema values in the spatial domain.

For Example 4.1a the extrema values of the sensitivity coefficients S_{u,θ_2} and S_{u,θ_3} were found to reach extrema values at $x = 0.00$ and 1.00 (as shown in Fig. 1b and c and in Table 2). Note that here as well as in all examples shown it is the *highest* maxima value and the *lowest* minima of the spatio-temporal distributions of the sensitivity functions that are taken into account (here for $t = 0.1$ and 1.0, for S_{u,θ_2} and for S_{u,θ_3}, respectively. In the case of considering parameter θ_1 in the analysis (Example 4.1b) S_{u,θ_1} has the highest extrema at $x = [0.00 \quad 0.50 \quad 1.00]$ (see Fig. 1a and Table 2). Even though it is not shown here (since all the sensitivity functions in Figs. 1–3 were normalised) it was observed that this parameter sensitivity coefficient showed the largest changes. For Example 4.2 it was observed that the predominant sensitivity coefficient is S_{u,θ_3},

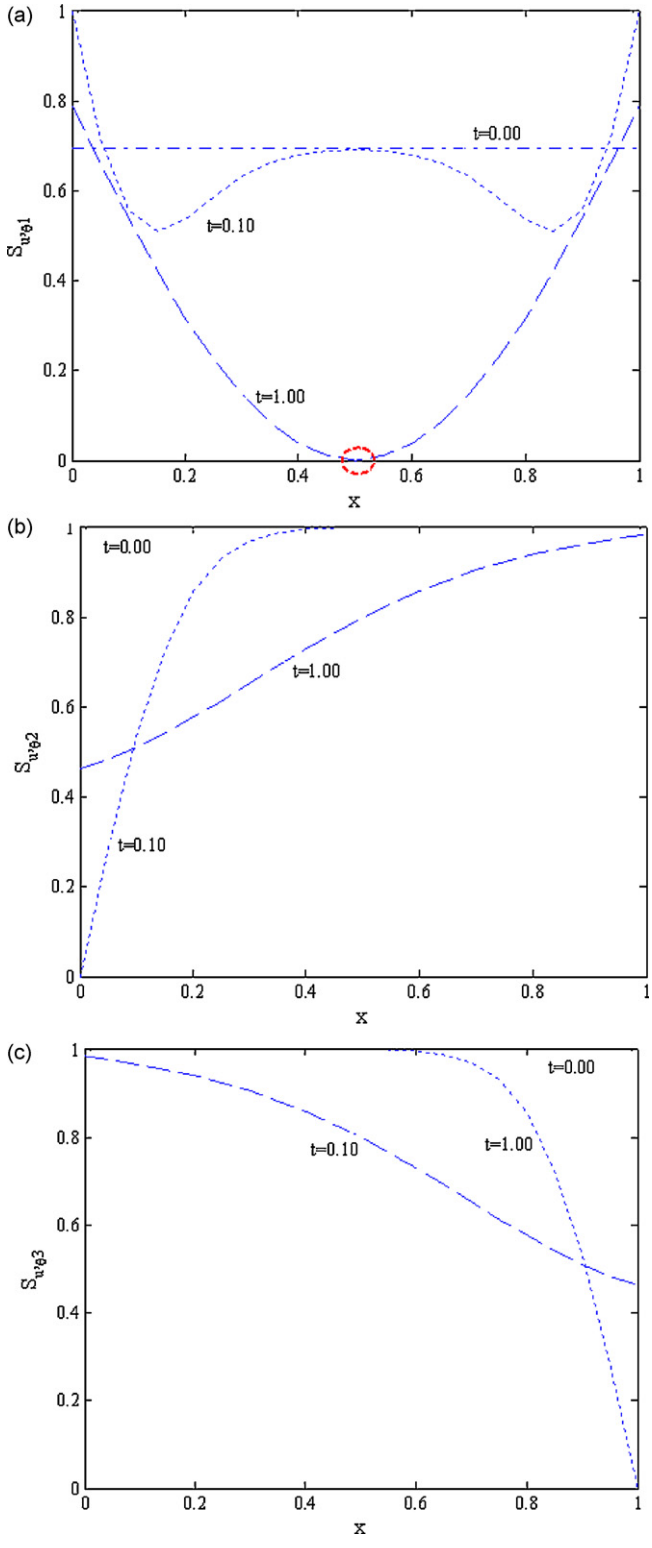


Fig. 1. Spatio-temporal distributions of sensitivity functions: (a) S_{u,θ_1}, (b) S_{u,θ_2} and (c) S_{u,θ_3} for Example 4.1a and b. Circles indicate locations of extrema values.

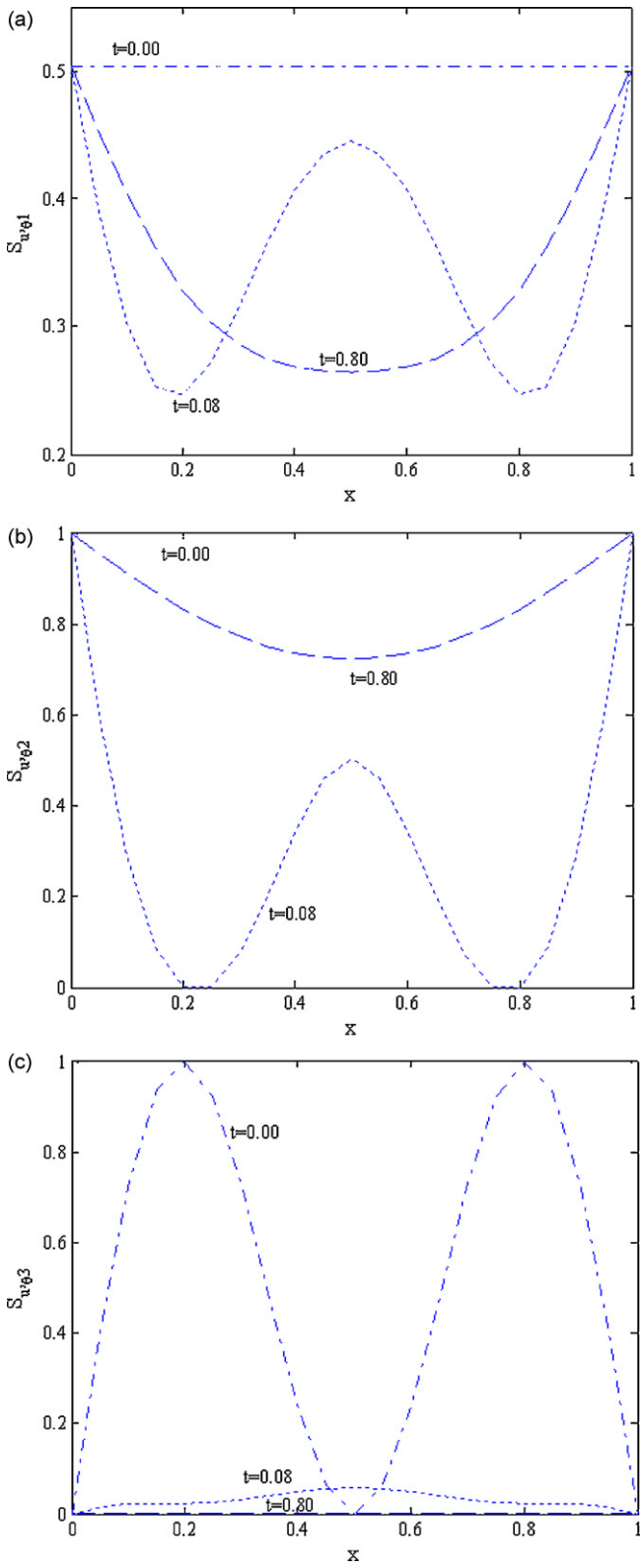


Fig. 2. Spatio-temporal distributions of sensitivity functions: (a) S_{u,θ_1}, (b) S_{u,θ_2}, (c) S_{u,θ_3} for Example 4.2.

showing extrema values at $x = [0.20 \quad 0.50 \quad 0.80]$ as it can be seen in Fig. 2c. In Example 4.3 the spatial positions of the extrema values of S_{u_1,θ_1} and S_{u_1,θ_2} change in time as shown in Fig. 3, where the dynamic behaviour of sensitivity functions for different equidistant time intervals for $t \in [0, 60]$ is depicted. The corresponding extrema values at $t = 120$, where the system reaches complete steady state are shown in Table 2.

Example 4.3 shows extrema values at $x = [0.00, 0.20, 0.45, 0.55, 0.70, 0.80, 1.00]$ for field $u_1(x, t)$ as shown in Fig. 3a–d and $x = [0.00 \; 1.00]$ for field $u_2(x, t)$ (Fig. 4a–d), for this example the predominant sensitivity coefficients correspond to parameter θ_4 for u_1 and u_2. These locations are collectively listed in Table 2.

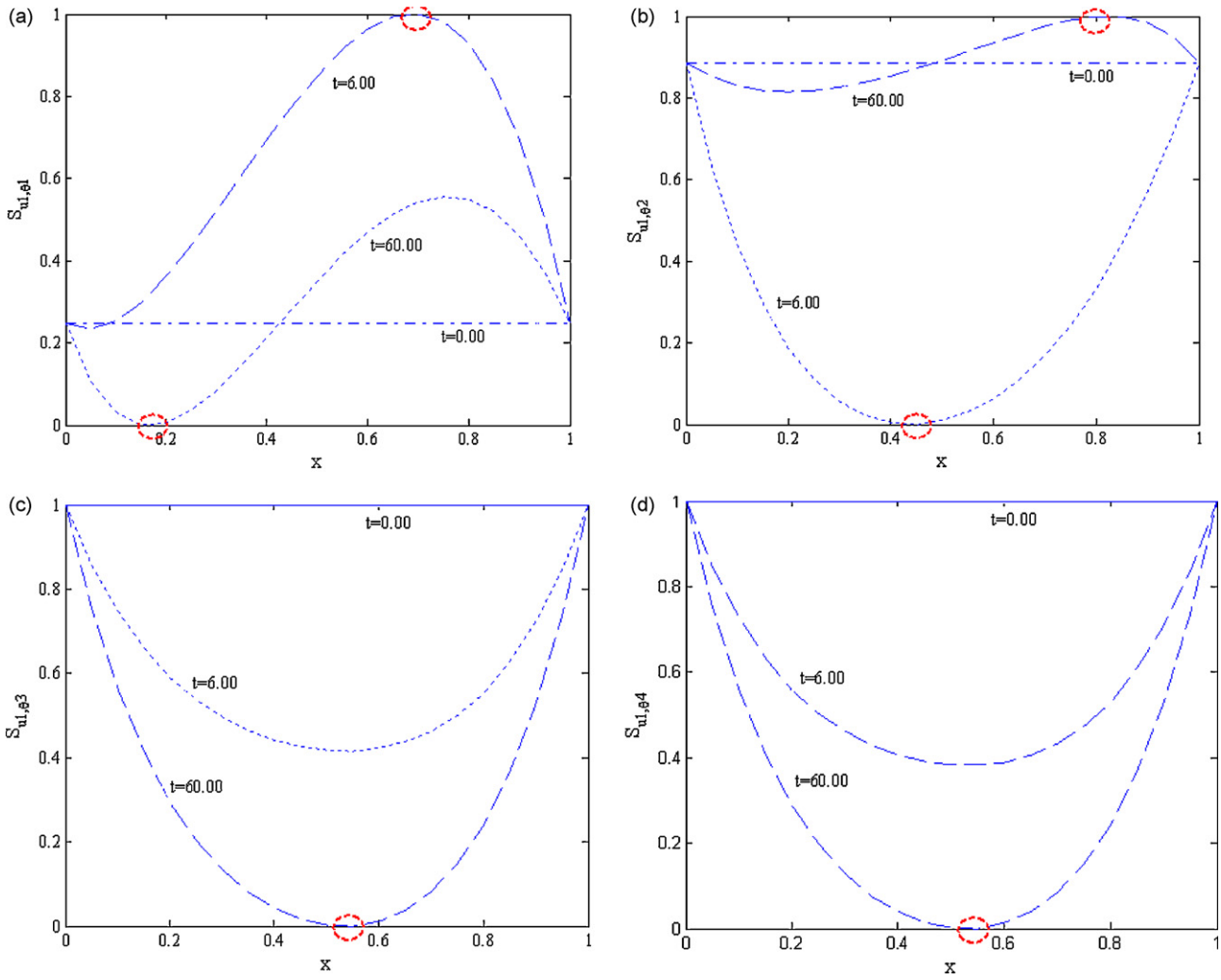


Fig. 3. Spatio–temporal distribution of the sensitivity coefficients of u_1 with respect to parameters: (a) θ_1, (b) θ_2, (c) θ_3 and (d) θ_4 for Example 4.3. Circles indicate locations of extrema values.

The sensitivity coefficients computed in the previous step were used to calculate the Gram determinant distribution of each system, using Eq. (11). Fig. 5 shows the results obtained.

Based on the Gram determinant criterion the sensor locations are; [0.20 0.80] for Example 4.1a, [0.15 0.85] for Example 4.1b, [0.25 0.75] for Example 4.2, and at $x = 0.55$ for field $u_1(x, t)$ but no evident solution is observed for field $u_2(x, t)$ for Example 4.3. The corresponding results on sensor placement using the D-optimal design and the max–min optimisation method, prescribing 1, 2 and 3 sensors, are listed in Table 2.

For each system studied here, the POD using the method of snapshots (Sirovich, 1987) was applied, and the relative and total energy of the system captured by the empirical eigenfunctions were calculated, see Table 1.

The field reconstruction for Examples 4.1 and 4.2 can be achieved by considering the first two empirical eigenfunctions, which capture 99.99% of the system's energy. Example 4.3 requires the first two POD modes to reproduce $u_1(x, t)$ (again capturing 99.99% of the system's energy) and the first one to reproduce $u_2(x, t)$. Fig. 6 shows the spatial distribution of these dominant POD empirical eigenfunctions for each example studied here. In this figure, it can be seen that the extrema values of the *dominant* POD modes coincide at the same positions where sensitivity functions reach their *highest* extrema values. This is particularly more evident for the dominant sensitivity functions of each system. Note that for field $u_1(x, t)$ in Example 4.3 two more locations, namely $x = 0.55$ and 0.70, are obtained by the extrema of the sensitivity functions. These two locations can be captured if more POD modes as well as the dominant ones are considered.

The advantage of using POD modes is that their calculation is more computationally efficient than the calculations involved and required in all the other methods analysed in this work sidestepping the need to construct large-scale sensitivity matrices. Furthermore, the POD-based method can then be used to directly reduce the system at hand, leading to efficient linking of sensor placement with on-line optimisation and control strategies for DPS. See Table 7 where a qualitative comparison of the computational requirements of each technique is given. Another interesting feature is that we can, thus, directly link the POD extrema with the corresponding system's energy each mode captures.

Table 2 lists the optimal sensor locations obtained from the different methods. As it can be seen, several feasible solutions are reported for the max–min optimisation approach. This method would produce the same results reported by the extrema value analysis of the POD modes (and/or sensitivity coefficients) if the *proper* number of sensors is considered, however, this cannot be explicitly provided by the technique. Furthermore, it can be observed that D-optimal experiment design predicts the same locations as the ones computed from the extrema of POD modes/sensitivity coefficients for Examples 4.1 and 4.2 while for Example 4.3 we can see some differences in the locations predicted from the different methods.

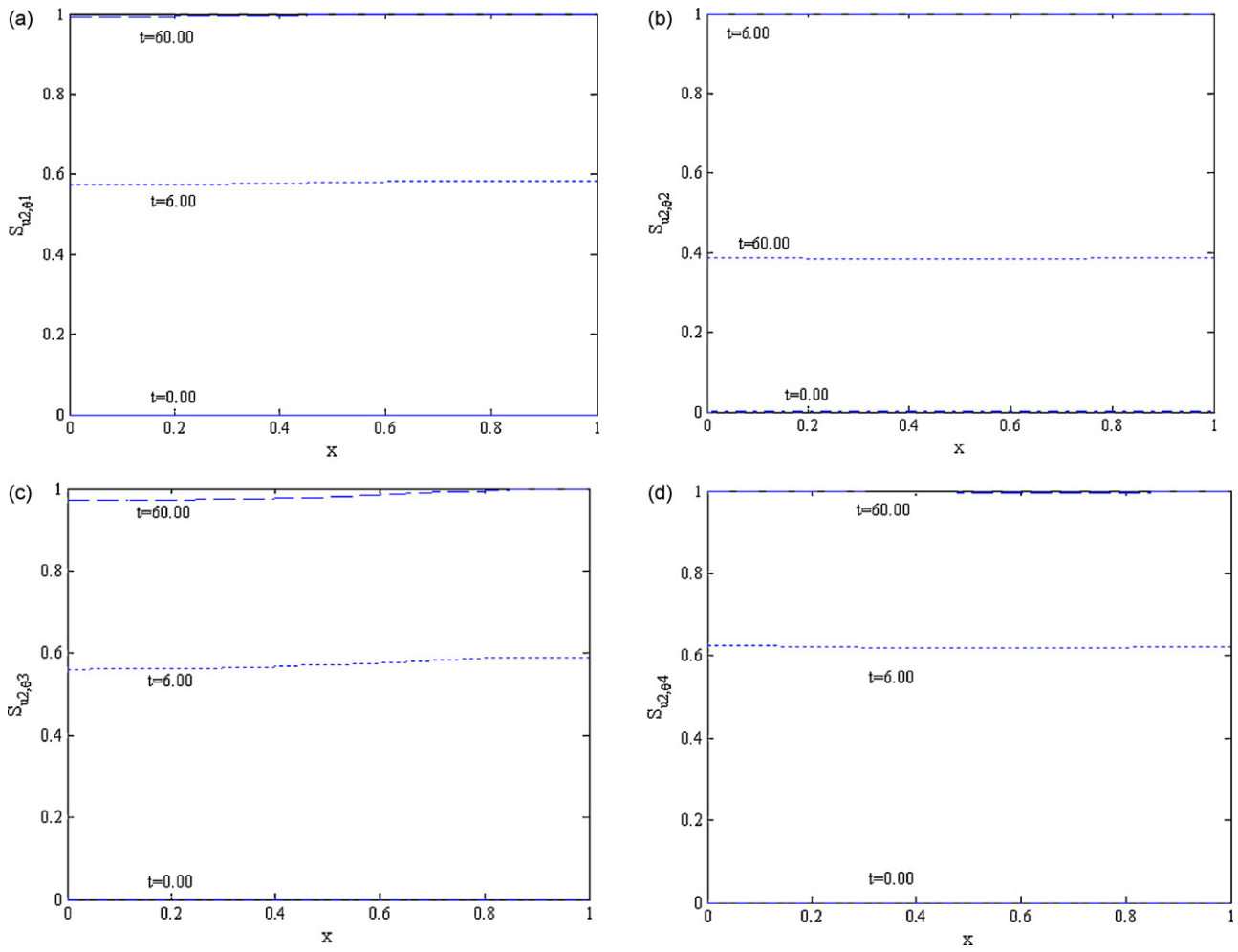


Fig. 4. Spatio-temporal distribution of the sensitivity coefficients of u_2 with respect to parameters: (a) θ_1, (b) θ_2, (c) θ_3 and (d) θ_4 for Example 4.3.

In order to verify the results obtained from the different methods followed here to determine optimal measurement positions, a parameter estimation procedure was implemented. Each system was solved using nominal values for the parameters and random noise with zero mean, $N(0, \sigma^2)$, was added to the outputs. The variance, σ^2, was manipulated in order to produce up to ±20% of stochastic deviation from the outputs. For each example 10 replications of the digital experiment were collected. The resulting ten responses were then averaged and treated as experimental data, which were used in the minimisation of the objective function (33). The square deviations between the outputs estimated through the model and the experimental data measured at the positions reported by the different the optimal sensor location methodologies are computed and reported as the sum of square errors–SSE normalised with the number of measurements used. The parameters calculated were compared with the nominal values. The results obtained from the parameter estimation procedure for each of the examples are shown in Tables 3–6, respectively. The nominal values of the parameters are shown in bold.

It can be seen from the reported SSE values in Table 3 that no further improvement is introduced in the parameter estimation process when experimental data at $x = 0.50$ are considered. The D-optimal design and the extrema evaluation of the POD modes provide the measurement locations that ensure the best parameter estimates.

In case 4.1b as can be seen in Table 4 considering experimental measurements at $x = 0.50$ improved considerably the parameter estimates. This can be observed when the *additional scenario* (i.e. measuring only at $x = [0.00 \quad 1.00]$) is compared with measuring at $x = [0.00 \quad 0.50 \quad 1.00]$ as the D-optimal design and the extrema

Table 1
Relative and total system's energy captured by the POD modes.

		λ_1	λ_2	Total kinetic energy
Example 4.1	Field $u(x, t)$	1.7777 (0.9594)	0.0751 (0.0405)	99.99%
Example 4.2	Field $u(x, t)$	35.4749 (0.9910)	0.3204 (0.0090)	99.99%
Example 4.3	Field $u_1(x, t)$	0.23677 (0.9997)	0.00006 (0.0002)	99.99%
	Field $u_2(x, t)$	0.22160 (0.99994)	0.00001 (0.00005)	99.99%[a]

Number in parentheses represent the relative energy captured by the ith POD-mode, defined by; $E_i = \lambda_i / \sum_{k=1}^{m} \lambda_k$.

[a] This is achieved considering just the first POD-mode.

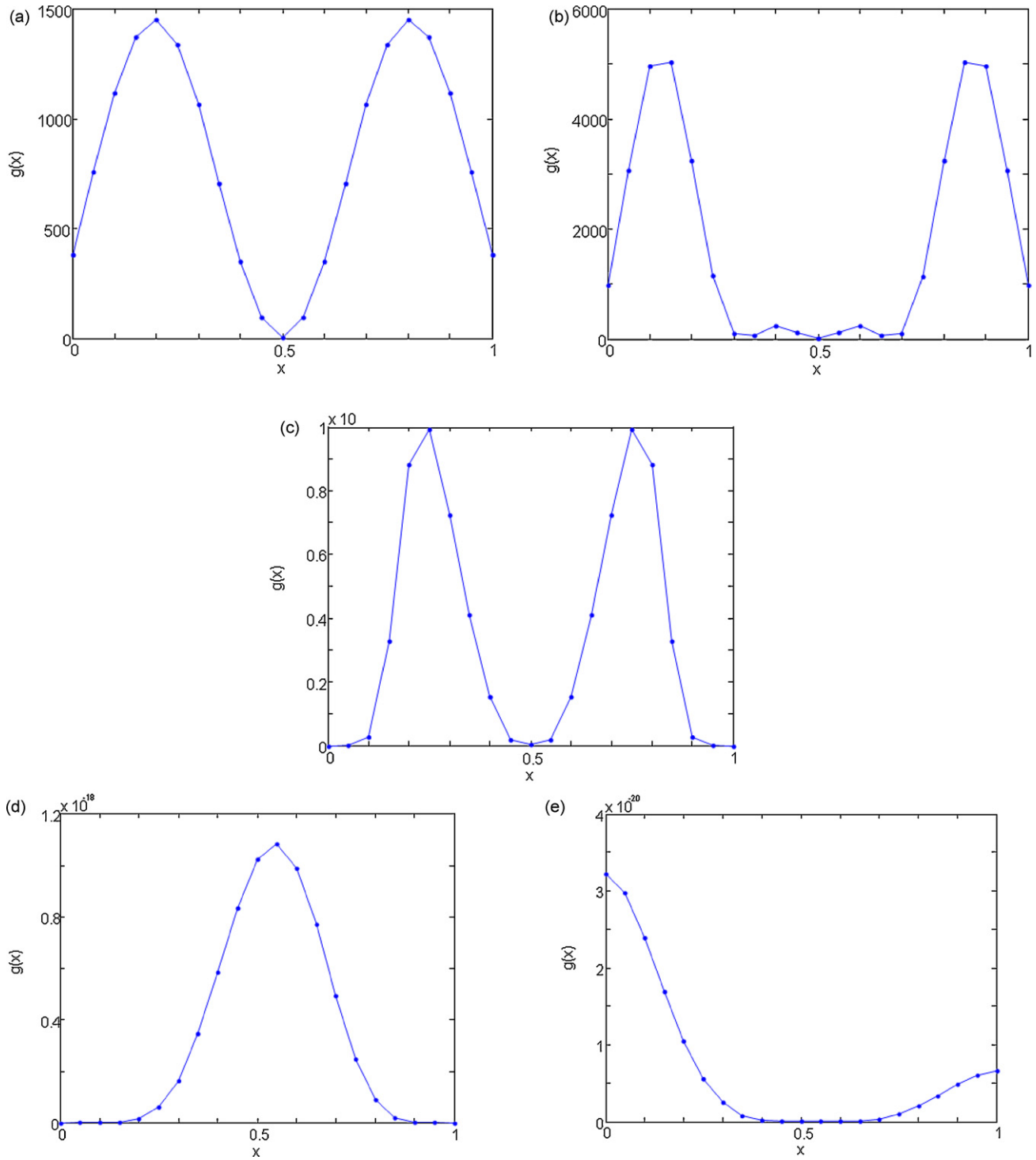


Fig. 5. Spatial distributions of the Gram determinant for example: (a) 4.1a, (b) 4.1b, (c) 4.2 (d) 4.3, $u_1(x, t)$ and (e) 4.3, $u_2(x, t)$.

values of the POD modes reported as optimal sensor positions. It is also worth mentioning that measuring only at $x = 0.50$ (where the predominant POD mode reaches a minimum as sees in Fig. 6a, provided better results than measuring at $x = [0.00 \quad 1.00]$, which reinforces the statement that experimental information at this location enhances parameter estimation. Hence, using the dominant POD mode here we were able to obtain very good results (SSE = 20.331 compared to SSE = 18.995 obtained with 3 measurements) with a single measurement. Obtaining the minimum number of measurements required is obviously important in situations where experimental measurements are very expensive. It should be mentioned here that the max–min approach reports the same result when the number of sensors sought is set to one.

In Table 5 it can be noticed that the best estimates (lowest SSE) are obtained when the experimental data is measured at the positions reported by the extrema value evaluation of the POD modes and/or sensitivity coefficients, which, here, are the same as the

Table 2
Optimal sensor locations.

	D-optimal design	Gram-determinant	Max–min optimisation	Extrema sensitivity coefficients	Extrema POD empirical eigenfunctions
Example 4.1a	0.00 1.00	0.15 0.85	1 sensor 0.50 2 sensors (1) 0.45 0.50 (2) 0.50 0.55 3 sensors 0.45 0.50 0.55	0.00 1.00	0.50[c] 0.00 0.50 1.00
Example 4.1b	0.00 0.50 1.00	0.15 0.85	1 sensor 0.50 2 sensors (1) 0.45 0.50 (2) 0.50 0.55 3 sensors 0.45 0.50 0.55	0.00 0.50 1.00 0.50[c]	0.50[c] 0.00 0.50 1.00
Example 4.2	0.20 0.50 0.80	0.25 0.75	1 sensor 0.50 2 sensors 0.50 0.55 3 sensors 0.45 0.50 0.55	0.50[c] 0.20 0.50 0.80	0.50[d] 0.20 0.50 0.80
Example 4.3	Field $u_1(x,t)$ 0.20 0.55 0.80 0.50 Field $u_2(x,t)$ 0.00 0.50 1.00	Field $u_1(x,t)$ 0.55 Field $u_2(x,t)$ –[b]	Field $u_1(x,t)$[a] 0.00 0.05 0.90 0.95 1.00 Field $u_2(x,t)$[a] 0.00 0.05 0.10 0.15 0.20	Field $u_1(x,t)$ 0.00 0.20 0.45 0.55 0.70 0.80 1.00 Field $u_2(x,t)$ 0.00 1.00	Field $u_1(x,t)$ 0.00 0.20 0.45 0.80 1.00 Field $u_2(x,t)$ 0.00 1.00

[a] For this particular example five sensors were defined.
[b] No particular solution was observed from the Gram-determinant evaluation for this case.
[c] This location corresponds to the point where the predominant sensitivity coefficient and POD mode reaches an extrema.
[d] This location corresponds to the point where the predominant sensitivity coefficient and all the POD modes coincide at the same extrema.

Table 3
Parameter estimation Example 4.1a.

Method	Measurement locations	θ_1 (**0.100**)	θ_2 (**1.000**)	θ_3 (**1.000**)	SSE
D-optimal design	0.00 1.00[b]	Fixed at 0.10	1.0315 (3.15%)	1.1039 (10.24%)	20.873
Gram-determinant	0.15 0.85	Fixed at 0.10	1.1403 (14.03%)	0.9449 (5.51%)	29.645
Extrema POD modes	0.00 0.50 1.00	Fixed at 0.10	1.0334 (3.34%)	1.1058 (10.58%)	20.903
	0.50[a]	Fixed at 0.10	1.0504 (5.04%)	1.0961 (9.61%)	21.100
Max–min optimisation	0.45 0.50	Fixed at 0.10	0.8340 (16.60%)	1.3078 (30.78%)	30.801
	0.50 0.55	Fixed at 0.10	1.9183 (91.83%)	0.6182 (38.18%)	37.566
	0.45 0.50 0.55	Fixed at 0.10	1.5245 (52.45%)	0.6738 (32.62%)	33.262

Numbers in parentheses represent the percentage deviation of the estimated parameter from the nominal value (shown in bold).
[a] The same result is obtained when the number of sensors is fixed to 1 during the max–min optimisation method.
[b] This same result is obtained when the extrema values in the spatial domain of the sensitivity coefficients are evaluated.

ones produced by the D-optimal design. An additional calculation, not shown, was made considering experimental sampling only at $x = 0.50$. This extra case produced worse results, corroborating the fact that for this example more that one measurement location is needed.

Table 6 shows the parameter estimates obtained for Example 4.3. The best results (lowest SSE value) are produced when the sensor locations are derived from the extrema values of the POD modes. Also the corresponding parameters obtained have low deviation from the nominal ones, including a very good estimate of θ_1, while

Table 4
Parameter estimation Example 4.1b.

Method	Measurement locations	θ_1 (**0.100**)	θ_2 (**1.000**)	θ_3 (**1.000**)	SSE
D-optimal design	0.00 0.50 1.00[b]	0.1005 (0.05%)	1.0295 (2.95%)	1.0295 (2.95%)	18.955
Gram-determinant	0.15 0.85	0.1176 (17.60%)	0.9042 (9.58%)	0.8142 (18.58%)	52.852
Extrema POD modes	0.50[a]	0.1071 (7.10%)	0.9149 (8.51%)	1.0032 (0.32%)	20.331
Max–min optimisation	0.45 0.50	0.0946 (0.54%)	1.0028 (0.28%)	1.4860 (48.60%)	49.528
	0.50 0.55	0.0830 (1.70%)	1.3797 (37.97%)	1.0665 (6.65%)	31.778
	0.45 0.50 0.55	0.0836 (1.64%)	1.5954 (59.54%)	0.9670 (3.30%)	49.009
Additional scenario [c]	0.00 1.00	0.1174 (1.74%)	1.0556 (5.56%)	1.1358 (13.58%)	27.995

Number in parentheses represents the percentage deviation of the parameter from the nominal value (shown in bold).
[a] The same result is obtained when the number of sensors is fixed to one during the max–min optimisation method.
[b] The same result is obtained when the extrema values of the predominant sensitivity coefficients, S_{u,θ_1}, and of the POD modes are evaluated.
[c] This additional scenario was considered in order to compare the results obtained for Example 4.1a measuring at the same positions but estimating all three parameters.

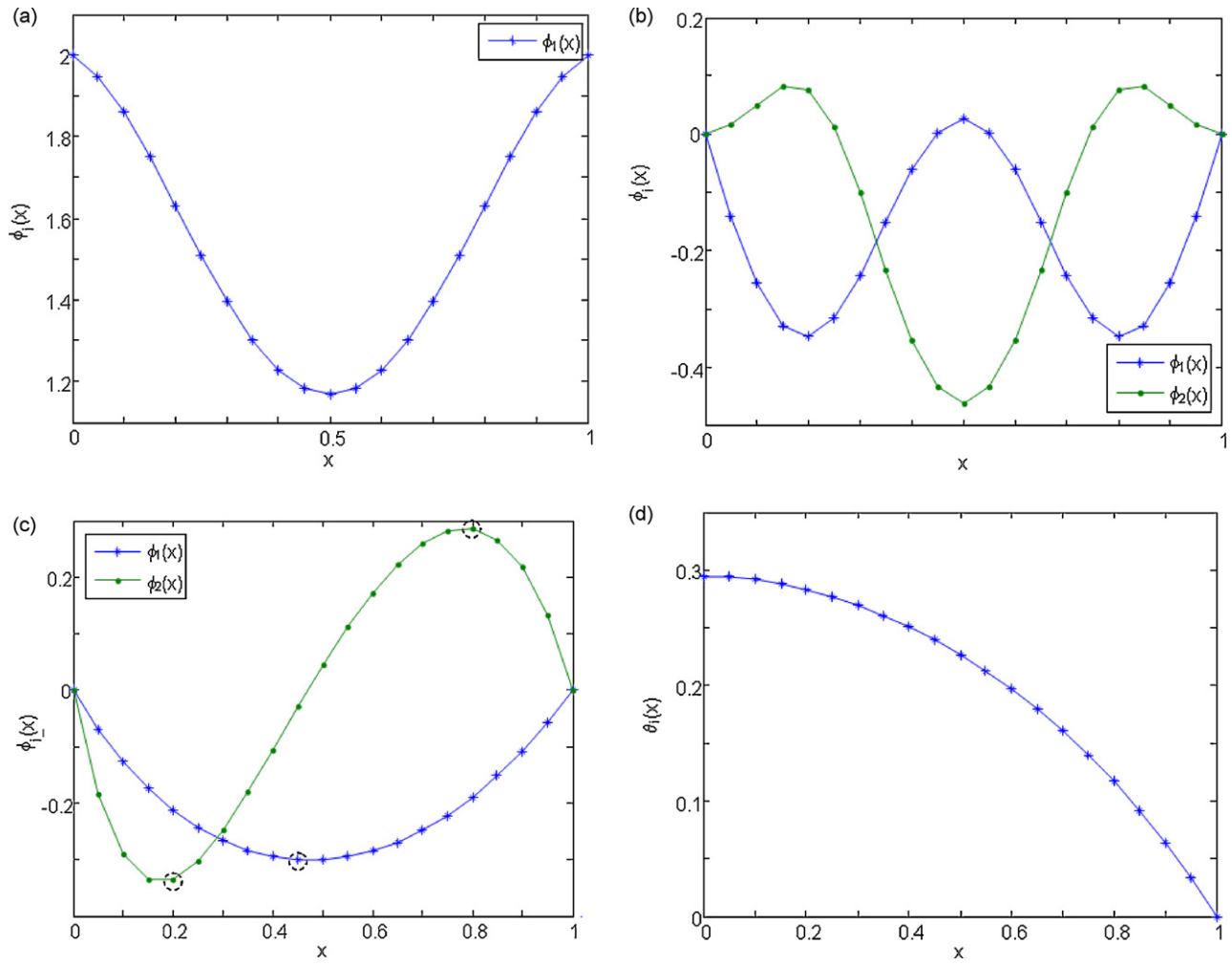


Fig. 6. Empirical eigenfunctions capturing most energy of the system for example: (a) 4.1, (b) 4.2, (c) $u_1(x, t)$ for Example 4.3 and (d) $u_2(x, t)$ for Example 4.3.

Table 5
Parameter estimation Example 4.2.

Method	Measurement locations	θ_1 (**0.100**)	θ_2 (**2.000**)	θ_3 (**5.000**)	SSE
D-optimal design and extrema POD modes	0.20 0.50 0.80[a]	0.0875 (1.25%)	2.1098 (5.49%)	4.9109 (1.78%)	0.030
Gram-determinant	0.25 0.75	0.0860 (1.40%)	2.4231 (21.16%)	5.0521 (1.04%)	0.203
Max–min optimisation	0.50 0.55	0.0939 (0.61%)	1.9896 (0.52%)	5.9534 (19.07%)	0.109
	0.45 0.50 0.55	0.1037 (0.37%)	1.9453 (2.74%)	5.7122 (14.24%)	0.096

Numbers in parentheses represent the percentage deviation of the parameter from the nominal value (shown in bold).
[a] The same result is obtained when the extrema values of the predominant sensitivity coefficients, S_{u,θ_3}, and of the dominant POD modes are evaluated.

Table 6
Parameter estimation Example 4.3.

Method	Measurement locations	θ_1 (**1.000**)	θ_2 (**6.000**)	θ_3 (**1.000**)	θ_4 (**1.000**)	SSE
D-optimal design	$u_1(x, t)$: 0.20 0.50 0.55 0.80 $u_2(x, t)$: 0.00 0.50 1.00	1.6333 (63.33%)	6.3206 (5.34%)	0.9679 (3.21%)	1.0288 (2.88%)	0.0039
Gram-determinant	$u_1(x, t)$: 0.55 $u_2(x, t)$: No information was obtained	–	–	–	–	–
Extrema POD modes	$u_1(x, t)$: 0.00 0.20 0.45 0.80 1.00 $u_2(x, t)$: 0.00 1.00	1.0125 (1.25%)	6.2197 (3.66%)	0.8653 (13.47%)	1.0759 (7.59%)	0.0026
Max–min optimisation	$u_1(x, t)$: 0.00 0.05 0.90 0.95 1.00 $u_2(x, t)$: 0.00 0.05 0.10 0.15 0.20	1.4316 (43.16%)	6.3435 (5.73%)	0.3064 (69.36%)	1.3148 (31.48%)	0.0094
Extrema sensitivity coefficients	$u_1(x, t)$: 0.00 0.20 0.45 0.55 0.70 0.80 1.00 $u_2(x, t)$: 0.00 1.00	1.0287 (2.87%)	6.2558 (4.26%)	1.1980 (19.80%)	1.1859 (18.59%)	0.0036
All the data	$u_1(x, t)$: whole domain $u_2(x, t)$: whole domain	1.1387 (13.87%)	6.0930 (1.55%)	1.0980 (9.80%)	0.9617 (3.83%)	0.0025

Numbers in parentheses represent the percentage deviation of the parameter from the nominal value (shown in bold).

Table 7
Optimal sensor position methods. Requirements.

Method	Numerical optimal sensor location—requirements				
	Experimental data	The original governing equations	Initial value of parameters	Sensitivity coefficients	Calculation cost
D-optimal design	√	√	√	√	High
Gram-determinant[a]	√	√	√	√	High
Min–max optimisation	√	×	×	×	Moderate
Sensitivity coefficients extrema analysis	√	√	√	√	High
POD modes extrema analysis	√	×	×	×	Low

[a] There is no information in the literature for the uses of this method in the case of systems consisting of more than one output or multi-state variables.

Table 8
Optimal sensor position methods. Features.

Method	Produce optimal locations	Produce number of optimal locations	Effect of noise	Observations
D-optimal design	√	√	Highly affected	Needs a previous estimation of parameters
Min–max optimisation	√	×	Moderately affected	Can produce more than one solution, generally in the neighbourhood where the outputs reach extrema values. The number of sensors needs to be specified
Gram-determinant	√	√	Highly affected	Needs sensitivity coefficients
Sensitivity coefficient extrema analysis	√	√	Highly affected	The computation of these coefficients for PDE models is a highly demanding operation
POD modes extrema analysis	√	√	Moderately affected[a]	The empirical eigenfunctions capturing the most energy from the system are needed

[a] The POD modes capturing the lowest level of energy of the system are the most affected by the presence of noise in the experimental data. This problem can be solved using common filtering techniques applied either to the experimental data or to the POD modes calculated.

the parameter with the highest deviation form the nominal one is θ_3. The worse parameter values are reported when the experimental measurements are taken in positions different from these locations. Here also, results obtained when experimental measurements at all 21 (computational) points along the spatial domain were considered. This case produces only a very small improvement in the error estimate (SSE = 0.0025 compared with 0.0026 in the POD case) at the extra cost of 35 more measurements.

Table 7 shows the computational requirements for each method to solve the optimal sensor location problem. All the methods require experimental data. The POD extrema and the min–max optimisation method are straightforward determinations of the optimal sensor location considering a set of experimental data. However, in the min–max optimisation approach the number of sensors needs to be established by the user. Furthermore, for large-scale systems model reduction needs to be performed first and additional calculations are needed to determine sensor locations. Hence the computational cost for this method is characterised as "moderate". "High" computational cost is attributed to the methods requiring (large) sensitivity matrices to be constructed, especially for DPS. Table 8 shows the main features of each method considered for optimal sensor position purposes. All the methods are affected by the presence of high levels of noise in the experimental data.

6. Conclusions

In this work, a comparison between different methods reported in the literature on estimating optimal sensor locations to compute system parameters was presented. Three relevant methods from the literature were presented, and a number of observations based on three case studies were made. Moreover, a new method where sensors placed at the positions where POD functions capturing most of the system's energy reach their extrema values was discussed. Furthermore, the equivalence between the extrema values of sensitivity functions and of POD modes was proven and illustrated. The method based on the Gram determinant evolution, might not produce accurate results in some cases, since is strongly dependent on the behaviour of sensitivity coefficients and it also, requires extensive calculations. The max–min optimisation method can produce more than one optimal solution, depending on the number of sensors prescribed. Its computational cost increases significantly as the size of the distributed parameter system increases. Hence model reduction is required to handle large systems. This method depends on the magnitude of output values, assigning the measurements points to the positions where the outputs reach their extrema values. The D-optimal design method produces as results the number and positions where the measurements should be made in the spatial domain. This method depends strongly on the sensitivity coefficients, but mostly on their spatial evolution rather than their exact magnitudes. The results are strongly affected by the presence of noise in the system, but this can be partly solved using common filtering techniques. The methods that are based on the calculation of sensitivity coefficients are the ones most affected by the presence of noise. While the D-optimal design requires post-processing of the sensitivity coefficients to determine positions and the optimal number of measurements, a mere observation of the extrema values of the sensitivity coefficients can produce equally good results. It was observed in all cases studied that using the extrema values of the dominant POD modes to estimate measurement positions not only produced very good results in all cases with small computational cost, but also associating the computed measurement locations with the energy captured by each mode, an estimate of the minimum measurements required could be implicitly obtained. This can be important especially in situations where experimental observations are expensive.

Acknowledgements

CT would like to acknowledge the financial support of the EU project CONNECT (COOP-2006-31638).

References

Alonso, A., Frouzakis, C., & Kevrekidis, I. (2004). Optimal sensor placement for state reconstruction of distributed process systems. *AIChE Journal*, *50*(7), 1438–1452.

Alonso, A., Kevrekidis, I., Banga, J., & Frouzakis, C. (2004). Optimal sensor location and reduced order observer design for distributed process systems. *Computer and Chemical Engineering*, *28*, 27–35.

Alvares, J., Romangnoli, J., & Stepanopoulos, G. (1981). Variable measurements structures for the control of a tubular reactor. *Chemical Engineering Science*, *36*, 1695–1712.

Atkinson, A. C., & Donev, A. N. (1992). *Optimum experimental design*. Oxford: Oxford University Press.

Aubry, N., Holmes, P., Lumley, J. L., & Stone, E. (1988). The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Mechanics*, *192*, 115–173.

Aubry, N., Lian, W. Y., & Titi, E. S. (1993). Preserving symmetries in the proper orthogonal decomposition. *SIAM Journal on Scientific Computing*, *14*(2), 483–505.

Ball, K. S., Sirovich, L., & Keefe, L. R. (1991). Dynamical eigenfunctions decomposition of turbulent channel flow. *International Journal for Numerical Methods in Fluids*, *12*, 585–604.

Basseville, M., Benveniste, A., Moustakides, G. V., & Rougee, A. (1987). Optimal sensor location for detecting changes in dynamical behaviour. *IEEE Transactions on Automatic Control*, *23*, 1067–1075.

Bendesrky, E., & Christofides, P. D. (2000). Optimisation of transport-reaction using nonlinear model reduction. *Chemical Engineering Science*, *55*, 4349–4366.

Berkooz, G. (1992). Observations on the proper orthogonal decomposition. In *Studies in turbulence*. New York: Springer-Verlag., pp. 229–247.

Berkooz, G., Holmes, P., Lumley, J. L., & Mattingly, J. C. (1997). Low dimensional models of coherent structures in turbulence. *Physics Reports—Review Section of Physics Letters*, *287*, 338–384.

Chambers, D. H., Adrian, R. J., Moin, P., Stewart, D. S., & Sung, H. J. (1988). Karhunen–Loeve expansion of Burgers' model of turbulence. *Physics of Fluids*, *31*, 2573–2582.

Chmielewski, D. J., Tasha, P., & Manousiouthakis, V. (2002). On the theory of optimal sensor placement. *AIChE Journal*, *48*(5), 1001–1012.

Colantuoni, G., & Padmanabhan, L. (1977). Optimal sensor location for tubular flow reactor systems. *Chemical Engineering Science*, *32*, 1035–1049.

Damak, T., Babary, J. P., & Nihtila, M. T. (1992). Observer design and sensor location in distributed parameter bioreactors. *Proceedings of DYCORD, Maryland*, *87*, 315–320.

Deane, A. E., Kevrekidis, I. G., Karniadakis, G. E., & Orszag, S. A. (1991). Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders. *Physics of Fluids*, *3*, 2337.

Dochain, D., Tali-Mammar, N., & Babary, J. (1997). On modeling, monitoring and control of fixed bed bioreactors. *Computer and Chemical Engineering*, *21*, 1255–1266.

Emery, A., & Nenarokomov, A. V. (1998). Optimal experiment design. *Measurement Science & Technology*, *9*, 864–876.

Fahroo, F., & Demetriou, M. A. (2000). Optimal actuator/sensor location for active noise regulator and tracking control problems. *Journal of Computational and Applied Mathematics*, *114*(1), 15, 137–158 (22).

Georges, D. (1995). The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems. In *Proceeding of the 34th conference on decision and control* New Orleans, LA, (pp. 3319–3324).

Graham, M. D., Lane, S. L., & Luss, D. (1993). Proper orthogonal decomposition analysis of spatiotemporal temperature patterns. *Journal of Physical Chemistry*, *97*(4), 889–894.

Harris, T. J., Macgregor, J. F., & Wright, J. D. (1980). Optimal sensor location with an application to a packed bed tubular reactor. *AIChE Journal*, *26*(6), 910–916.

Heredia-Zavoni, E., Montes-Iturizzaga, R., & Esteva, L. (1998). Optimal instrumentation of uncertain structural systems subject to earthquake motions. *Earthquake Engineering & Structural Dynamics*, *27*, 343–362.

Hermann, R., & Kerner, A. J. (1977). Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, *22*, 728–740.

Hernandez, M. L., Kirubarajan, T., & Bar-Shalom, Y. (2004). Multisensor resource deployment using posterior Cramer–Rao bounds. *IEEE Transactions on Aerospace and Electronic Systems*, *40*, 399–416.

Hilai, R., & Rubinstein, J. (1994). Recognition of rotated images by invariant Karhunen–Loeve expansion. *Journal of the Optical Society of America A: Optics Image Science and Vision*, *11*(5), 1610–1618.

Jørgensen, S., Goldschmindt, L., & Clement, K. (1984). A sensor location procedure for chemical process. *Computer and Chemical Engineering*, *8*(3/4), 195–204.

Karhunen, K. (1946). Zur spectral theorie stochasticher prozess. *Annales Academiae Scientiarum Fennicae Series A1-Mathematica*, 37.

Keller, J. P., & Bovin, D. (1992). Selection of input and output variables as a model reduction problem. *Automatica*, *28*(1), 171–177.

Kirby, M., Boris, J. P., & Sirovich, L. (1990). A proper orthogonal decomposition of a simulated supersonic shear layer. *International Journal for Numerical Methods in Fluids*, *10*(4), 411–428.

Kirby, M., & Sirovich, L. (1990). Application of the Karhunen–Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(1), 103–108.

Korbicz, J., Zgurovsky, M. Z., & Novikov, A. N. (1988). Suboptimal sensors location in the state estimation problem for stochastic nonlinear distributed parameter system. *International Journal of Systems Science*, *19*(9), 1871–1882.

Kumar, S., & Seinfeld, S. H. (1978). Optimal location of measurements in tubular reactors. *Chemical Engineering Science*, *33*, 1507–1516.

Li, R., Henson, M. A., & Kurtz, M. J. (2004). Selection of model parameters for off-line parameter estimation. *IEEE Transactions on Control Systems Technology*, *12*(3), 402–412.

Loeve, M. (1945). Sur les Fonctions aléatories stationaires du second ordre. *Revue de Scientifique*, *83*, 297.

Löhner, R., & Camelli, F. (2005). Optimal placement of sensors for contaminant detection based on detailed 3D CFD. Simulations engineering computations. *International Journal for Computer-Aided Engineering*, *22*(3), 260–273.

Lopez, T., & Alvares, J. (2004). On the effect of the estimation structure in the functioning of a nonlinear copolymer reactor estimator. *Journal of Process Control*, *14*(1), 99–109.

Lumley, J. L. (1967). *The structure of inhomogeneous turbulent flows in atmospheric turbulence and ratio wave propagation*. Moscow: Nauka., p. 166.

Lumley, J. L. (1970). *Stochastic tools in turbulence*. New York: Academic Press.

Ly, H. V., & Tran, H. T. (2005). Reduced-order-based feedback control of the Kuramoto–Sivashinsky equation. *Journal of Computational and Applied Mathematics*, *173*(1), 1–9.

Malebranche, J. (1988). Simultaneous state and parameter estimation and location of sensor for distributed systems. *International Journal of Systems Science*, *19*(8), 1387–1405.

Mehra, R. K. (1974). Optimal inputs for linear system identification. *IEEE Transactions on Automatic Control*, *19*(3), 192–200.

Moin, P., & Moser, R. D. (1989). Characteristic-eddy decomposition of turbulence in a channel. *Journal of Fluids Mechanics*, *200*, 417–509.

Morari, M., & O'Dowd, M. J. (1980). Optimal sensor location in the presence of nonstationary noise. *Automatica*, *16*, 463–480.

Morari, M., & Stephanopoulos, G. (1980). Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances. *AIChE Journal*, *26*, 247–259.

Muller, P. C., & Weber, H. I. (1972). Analysis and optimization of certain quantities of controllability and observability for linear dynamic system. *Automatica*, *8*, 237–246.

Nahor, H. B., Scheerlinck, N., Van Impe, J. F., & Nicolaï, B. M. (2003). Optimization of temperature sensor position in a hot wire probe set up for estimation of the thermal properties of foods using optimal experimental design. *Journal of Food Engineering*, *57*, 103–110.

Newman, A. J. (1996). *Model reduction via the Karhunen–Loeve expansion. Part II. Some elementary examples* (Tech. Rep. T.R. 96-32). Inst. Systems Research.

Obukhov, A. M. (1954). Statistical description of continuous fields. *Transactions Geophysical International Academii Nauk USSR*, *24*, 3–42.

Olanrewaju, M. J., & Al-Arfaj, M. A. (2006). Estimator-based control of reactive distillation system: Application of an extended Kalman filtering. *Chemical Engineering Science*, *61*, 3386–3399.

Omatu, S., Seinfeld, J. H., & Soeda, T. (1978). Optimal sensor location for a linear distributed parameter system. *IEEE Transactions on Automatic Control*, *23*, 665.

Papadimitriou, C. (2004). Optimal sensor placement methodology for parametric identification of structural systems. *Journal of Sound Vibration*, *278*, 923–947.

Peng, T. (2005). *Fault detection methods based on optimal sensor location*. Hunan: Central South University.

Point, N., Vande Wouwer, A., & Remy, M. (1996). Practical issues in distributed parameter estimation: Gradient computation and optimal experiment design. *Control Engineering Practice*, *4*(11), 1553–1562.

Pougachev, V. S. (1953). General theory of the correlations of random functions. *Izvestia Akademii Nauk USSR, Seriya Matematicheskaya*, *17*, 1401.

Punithakumar, K., Kirubarajan, T., & Hernandez, M. (2006). Multisensor deployment using PCRLBs, incorporating sensor deployment and motion uncertainties. *IEEE Transactions on Aerospace and Electronic Systems*, *42*, 1474–1485.

Qureshi, Z. H., Ng, T. S., & Goodwin, G. C. (1980). Optimum experimental design for identification of distributed parameter systems. *International Journal of Control*, *31*(1), 21–29.

Rafajlowicz, E. (1983). Optimal experiment design for identification of linear distributed-parameter systems: Frequency domain approach. *IEEE Transactions on Automatic Control*, *28*(7), 806–808.

Rafajlowicz, E. (1978). Design of experiments for parameter identification of the static distributed systems. *Systems Science*, *4*(4), 349–361.

Rafajlowicz, E. (1981). Design of experiments for eigenvalues identification in DPS. *International Journal of Control*, *34*(6), 1079–1094.

Rajaee, M., Karlson, S. K. F., & Sirovich, L. (1994). Low-dimensional description of free-shear-flow structures and their dynamical behaviour. *Journal of Fluid Mechanics*, *258*, 1–29.

Reichert, R. S., Hatay, F. F., Biringers, S., & Husser, A. (1994). Proper orthogonal decomposition applied to turbulent flows in a square duct. *Physics of Fluids Mechanics*, *6*, 3086–3092.

Romagnoli, J., Alvares, J., & Stepanopoulos, G. (1981). Variable measurements structures for process control. *Chemical Engineering Science*, *36*, 1695–1712.

Scherpen, J. M. A. (1993). Balancing of nonlinear systems. *Systems & Control Letters*, *21*, 143–153.

Shvartsman, S. Y., Theodoropoulos, C., Rico-Martinez, R., Titi, E. S., Mountziaris, T. J., & Kevrekidis, I. G. (2000). Order reduction for nonlinear dynamic models of distributed reacting systems. *Journal of Process Control*, *10*, 177–184.

Singh, A. K., & Hahn, J. (2005). Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems. *Industrial & Engineering Chemistry Research*, *44*(15), 5645–5659.

Singh, A. K., & Hahn, J. (2006). Sensor location for stable nonlinear dynamic systems: Multiple sensor case. *Industrial & Engineering Chemistry Research*, *45*(10), 3615–3623.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. *Quarterly of Applied Mathematics*, *5*, 561.

Sirovich, L. (1989). Chaotic dynamics of coherent structures. *Physica D*, *37*(1), 126–145.
Sirovich, L. (1991). Analysis of turbulent flows by means of the empirical eigenfunctions. *Fluid Dynamics Research*, *8*, 85–100.
Sumana, C., & Venkateswarlu, Ch. (2009). Optimal selection of sensors for state estimation in a reactive distillation process. *Journal of Process Control*, *19*, 1024–1035.
Uciński, D. (1992). Optimal sensor location for parameter identification of distributed systems. *Applied Mathematics and Computer Science*, *2*(1), 119–134.
Uciński, D. (2003, June 18–20). On the determination of moving sensor policies for parameter estimation of distributed systems. In *11th Mediterranean IEEE conference on control and automation MED'03* Rhodes, Greece.
Uciński, D. (2005). *Optimal measurement method for distributed parameter system identification*. USA: CRC Press.
Uciński, D., & Bogacka, B. (2002). Construction of T-optimum designs for multiresponse dynamic models. In *COMPSTAT 2002* (pp. 267–272). Berlin, Heidelberg: Physica.
Uciński, D., & Demetriou, M. (2004, June 30–July 2). An approach to the optimal scanning measurement problem using optimum experimental design. In *Proceeding of the 2004 American control conference* Boston, Massachusetts.
Van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M., & Smilde, A. K. (2000). Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. *Chemical Engineering Science*, *55*(4), 827–837.
Vande Wouwer, A., Point, N., Poterman, S., & Remy, M. (2000). An approach to the selection of optimal sensor locations in distributed parameter systems. *Journal of Process Control*, *10*(4), 291–300.
Waldraff, W., Dochain, D., Bourrel, S., & Magnus, A. (1998). On the use of observability measures for sensor location in tubular reactor. *Journal of Process Control*, *8*(5), 497–505.
Walter, E., & Prozanto, L. (1990). Qualitative and quantitative experiment design for phenomenological models—A survey. *Automatica*, *26*(5–6), 195–213.
Walter, E., & Prozanto, L. (1997). *Identification of parameter models from experimental data*. Springer.
Waterhouse, T. H., Eccleston, J. A., & Duffull, S. B. (2009). Optimal design criteria for discrimination and estimation in nonlinear models. *Journal of Biopharmaceutical Statistics*, *19*(2), 386–402.
Zamprohna, E., Barolo, M., & Seborg, D. E. (2005). Optimal selection of sensors inputs for batch distillation columns using principal component analysis. *Journal of Process Control*, *15*, 39–52.

Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes. Alaña, J., Computer and Chemical Engineering, 34(8), 1220(2010).

Computers and Chemical Engineering 34 (2010) 1220–1225

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Estimation of the temporal coefficients for an empirical approximator. New approach based on the Proper Orthogonal Decomposition modes

Alaña Jorge Enrique[a,b,*]

[a] Facultad de Ingeniería, Universidad Del Zulia, Maracaibo, Venezuela
[b] School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, UK

ARTICLE INFO

Article history:
Received 22 April 2009
Received in revised form 15 June 2009
Accepted 1 February 2010
Available online 6 February 2010

Keywords:
Proper Orthogonal Decomposition
Empirical eigenfunction
Temporal coefficient

ABSTRACT

A new approach to identify the temporal coefficients of an empirical approximator of a process without any knowledge of the mathematical model of the system is introduced. This approach is based on the Proper Orthogonal Decomposition method, and only a few experimental values are required to determine the empirical eigenfunction, and the temporal coefficients. The scheme studied is verified by a numerical example regarding the chemical reaction in a tubular reactor. The approach consists of the minimisation of an objective function, which is based on the sum of the square errors between the original snapshots and the values predicted from the linear combination of the empirical eigenfunctions and the time coefficients. This method is easy and fast to implement, produces the lowest deviation from the original experimental data, and is very useful when the theoretical model of the system is unknown or difficult to determine.

1. Introduction

This paper introduces a useful approach that uses experimental data taken from a process to identify an empirical approximator of the system. The sorts of process studied here are described by a distributed parameter system. In general, a model can be obtained using knowledge of the physics and chemistry of the system, and a diversity of theoretical methods to arrive at a solution. However, the explicit functional form and parameters required are often unknown or demand sophisticated estimation methods to arrive at realistic results. In addition, the conventional solution techniques often lead to high-order realisations, which are sometimes incongruent to the objective of finding simple and easy to solve models.

In a distributed parameter system, the process variables, and the parameters might change as a function of time and space. Usually, the exact description of these sorts of systems consists of partial differential equations with either mixed or homogeneous boundary constraints. Such a system by its very nature has distributed properties, and it is not surprising that there are an infinite number of spatially dependent state variables. In practical applications, a distributed parameter system is usually modelled by disregarding the spatial nature of the process and employing lumped system techniques. However, by doing so, interactions and noticeable time delays due to the underlying convection and diffusion phenomena are neglected.

Methods to solve distributed parameter system are especially helpful in research and engineering practice. There exists, unfortunately, some distance between mathematicians who explore models of systems and practical researchers who execute experiments. The beautiful, proficient, and potentially very useful mathematical results are frequently not accessible for the wide range of experimental researchers and engineers which each time need to face similar problems.

This incentivises the improvement of methods, which allow the identification of models accounting for the spatially changing nature of the system. Additionally, practice dictates that these methods should use information collected from the process by simple experimental procedures. This is the cornerstone of the approach presented here.

Many authors have focused on data-driven techniques. Gay and Ray (1995) proposed a data-driven system identification approach to obtain a low-order model using singular value decomposition theory for linear distributed parameter system. They found a general pseudomodal model that reproduces the dominant behaviour of the system by means of data collected in a stable nominal operation condition. These data are approximated by spline functions and subsequently employed to determine the kernel operator that relates the process variables.

A different data-driven method, presented by Sirovich (1987), uses the Karhunen–Loève expansion. This technique yields empirical eigenfunctions that represent the dominant characteristics of the system in an ordered fashion. That is, the first empirical eigenfunction captures the dominant behaviour; the second one captures the next dominant, and so forth. Next, these eigenfunctions are

* Tel.: +58 261 7598100.
E-mail addresses: joalana@luz.edu.ve, jorge_alana@hotmail.co.uk.

0098-1354/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2010.02.001

used as basis functions, in a method such a Galerkin, to arrive at a low-order model.

The approach presented here uses the approximation for the specific case of unknowing the model of the process. This approximation is defined as the linear combination of the empirical eigenfunctions and the temporal coefficients. The model obtained provides an accurate spatial and temporal representation of the process. Local measurements provide the necessary information to generate the kernel operator. The Proper Orthogonal Decomposition is then applied to determine a basis set that is a function only of the spatial dimension and a corresponding set of coefficients that is a function of time. All this can be accomplished without a previous knowledge of the equations, parameters, and boundary conditions that constitute the conventional distributed parameter system models.

2. Proper Orthogonal Decomposition—Galerkin procedure

The proper orthogonal decomposition has received much attention in recent years as a tool to analyse complex physical systems. The idea is to use a reliable solver to produce a number of snapshots from the physical model. This technique is used to produce an optimal representation of these snapshots in an average sense. The power of this method lies in its mathematical properties suggesting that it is the preferred method to use in many applications.

Proper orthogonal decomposition was independently proposed by several researchers including Karhunen (1946), Loève (1945), Pougachev (1953), and Obukhov (1954). For surveys in this area, the works of Lumley (1970) and Berkooz (1992) can be mentioned. This technique has been applied for numerous purposes. One such important application was the attraction of spatial organised motions in fluid flows. Sirovich (1991), Berkooz, Holmes, Lumley, and Mattingly (1997) adapted this method to study turbulent flows. Other applications include channel flows by Moin and Moser (1989), Ball, Sirovich, and Keefe (1991), square-duct flows by Reichert, Hatay, Biringers, and Husser (1994) and shear flows by Rajaee, Karlson, and Sirovich (1994). Proper orthogonal decomposition has also been applied to fluid related problems. For instance, to the Burgers' equation by Chambers, Adrian, Moin, Stewart, and Sung (1988), the Ginsburg–Landau equation and the Bérnard convection by Sirovich (1989). Interesting non-fluid applications of this approach are the characterisation of human faces by Kirby and Sirovich (1990) and image recognition by Hilai and Rubinstein (1994).

The Proper Orthogonal Decomposition had been formulated as a reasonable method of representing a stochastic field with a minimum degree of freedom. This technique is going to be used here to approximate an experimental field without knowing the model of the system.

Consider a set of N arbitrary shaped functions. This ensemble it is called snapshots $\{u_i\}$, $i = 1, 2, \ldots, N$. The objective is to obtain the most typical structure $\phi(z)$ among these snapshots $\{u_i\}$. Here some notations used in this work,

$$u_i(z): \quad \text{a snapshot in } L^2(\Omega) \tag{1}$$

$$\{u_i\}: \quad \text{ensemble of snapshots} \tag{2}$$

$$(f,g) \equiv \int_\Omega f(z)g(z)d\Omega: \quad \text{inner product in } L^2 \tag{3}$$

$$\langle u_i \rangle \equiv \frac{1}{N}\sum_{i=1}^{N} u_i(z): \quad \text{Ensemble average of snapshots} \tag{4}$$

Then the objective is equivalent expressed as to find a set of functions $\phi(z)$ such that,

$$\max \lambda \equiv \frac{\langle (\phi, u_i)^2 \rangle}{(\phi,\phi)} \tag{5}$$

The numerator of this equation can be written as,

$$\langle (\phi, u_i)^2 \rangle = \int_\Omega \left\{ \int_\Omega \langle u_i(z), u_i(z') \rangle \phi(z) dz \right\} \phi(z') dz' \tag{6}$$

Eq. (6) can be represented compactly by introducing a two-point correlation function $K(z,z')$ and an integral operator R having $K(z,z')$ as the kernel,

$$K(z,z') \equiv \langle u_i(z), u_i(z') \rangle = \frac{1}{N}\sum_{i=1}^{N} u_i(z) u_i^T(z') \tag{7}$$

$$R \equiv \int_\Omega K(z,z')dz' \tag{8}$$

Now (6) takes the form,

$$\langle (\phi, u_i)^2 \rangle = \int_\Omega \{R\phi\}\{\phi\} dz = (R\phi, \phi) \tag{9}$$

Substituting (9) into (5), it is found that the maximisation problem is equivalent to a eigenvalue problem,

$$R\phi = \lambda\phi \tag{10}$$

or,

$$\int_\Omega K(z,z')\phi(z')dz' = \lambda\phi(z) \tag{11}$$

The function that maximises λ on (5) is the eigenfunction of (11) with the largest eigenvalue. The eigenfunction $\phi(z)$ is expressed linearly in terms of the N snapshots,

$$\phi(z) = \sum_{k=1}^{N} \omega_k u_k(z) \tag{12}$$

Substituting (12) into (11) yields a matrix eigenvalue problem,

$$C_{ik}\omega_k = \lambda\omega_k \tag{13}$$

where

$$C_{ik} = \frac{1}{N}\int_\Omega u_i^T(z')u_i(z')dz \tag{14}$$

The eigenvector ω of (13), is then substituted into (12) to produce a set of empirical eigenfunctions $\phi(z)$.

Considering the eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_N$ and the corresponding eigenfunctions $\phi_1, \phi_1, \ldots, \phi_N$ in the order of magnitude of the eigenvalues. The eigenfunction ϕ_1 corresponds to the largest eigenvalue λ_1 and is the most representative structure of the snapshots $\{u_i\}$, the eigenfunction ϕ_2 with the next largest eigenvalue λ_2 is the next typical structure, and so forth. The span of these eigenfunctions is precisely the span of all the realisation of snapshots. In consequence, any feasible solution can be represented as a linear combination of these eigenfunctions.

3. The system and governing equations

A homogeneous tubular reactor without catalyst packing is considered. An irreversible first-order chemical reaction A → B takes place and it is described by the concentration rate $C_A(t,z)$ and the temperature $T(t,z)$. This reaction is exothermic and a cooling jacket is used to remove heat from the reactor. A fraction of the products can be recycled, r, and mixed with the reactants at the inlet of

Table 1
Parameter values.

Pe_C	Pe_T	B_C	γ	B_T	β_T
7.0	7.0	0.1	10.0	2.5	2.0

the reactor ($z=0$) (Bendersky & Christofides, 2000; Li & Chistofides, 2007).

Derived from mass and energy balances and under the standard assumptions of constant density (ρ); heat capacity of the reacting fluid (cp) and constant axial fluid velocity (v), the dynamic behaviour of the process is described by a set of partial differential equations defined on a spatial domain $z\in(0,1)$, in dimensionless form;

$$\frac{\partial y_1(t,z)}{\partial t} = -\frac{\partial y_1(t,z)}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1(t,z)}{\partial z^2} + B_T B_C(1+y_2(t,z))\exp\left(\frac{\gamma y_1(t,z)}{1+y_1(t,z)}\right) - \beta_T(u_C - y_1(t,z)) \tag{15}$$

$$\frac{\partial y_2(t,z)}{\partial t} = -\frac{\partial y_2(t,z)}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2(t,z)}{\partial z^2} + B_C(1+y_2(t,z))\exp\left(\frac{\gamma y_1(t,z)}{1+y_1(t,z)}\right) \tag{16}$$

Here $y_1(t,z)$, $y_2(t,z)$ are the temperature and concentration, respectively. The control variable u_C is the cooling water temperature. The terms Pe_C, Pe_T are the Peclet numbers of mass and energy flows, correspondingly, B_C is the Damkohler number, γ is the activation energy, B_T and β_T are related to heat of reaction and heat transfer, in that order, and $t\in[t_0,t_f]$, where t_0 and t_f are initial and final times.

Assuming negligible reaction in the recycle loop and instantaneous mixing of fresh feed and recycle at the reactor inlet, the boundary and initials conditions take the form,

$$z=0;\quad \frac{\partial y_1(t,0)}{\partial z} = Pe_T\,[y_1(t,0)-(1-r)y_1 - r y_1(t,1)]$$
$$\frac{\partial y_2(t,0)}{\partial z} = Pe_C\,[y_2(t,0)-(1-r)y_{20} - r y_2(t,1)] \tag{17a}$$

$$z=1;\quad \frac{\partial y_1(t,1)}{\partial z}=0,\quad \frac{\partial y_2(t,1)}{\partial z}=0 \tag{17b}$$

$$t=0;\quad y_1 = y_{20} = u_C = 0 \tag{17c}$$

The parameter values for the system studied in this work are shown in Table 1.

4. The low-order dynamic model

In this section the low dimensional dynamic model is developed. The temperature and concentration fields can be represented as a linear combination of the empirical eigenfunctions as follows:

$$y_1(z,t) = \sum_{i=1}^{M_1}\alpha_i(t)\phi_i(z) \tag{18a}$$

$$y_2(z,t) = \sum_{i=1}^{M_2}\beta_i(t)\theta_i(z) \tag{18b}$$

Here $\phi_i(z)$ and $\theta_i(z)$ are the ith empirical eigenfunction for temperature and concentration respectively, $\alpha_i(z)$ and $\beta_i(z)$ are their corresponding temporal coefficients. M_1 and M_2 are the number of empirical eigenfunctions employed in the Proper Orthogonal Decomposition—Galerkin procedure. The residuals for Eqs. (15) and (16) are,

$$R_1 \equiv -\frac{\partial y_1(t,z)}{\partial t} - \frac{\partial y_1(t,z)}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1(t,z)}{\partial z^2} + B_T B_C(1+y_2(t,z))\exp\left(\frac{\gamma y_1(t,z)}{1+y_1(t,z)}\right) - \beta_T(u_C - y_1(t,z)) \tag{19}$$

$$R_2 \equiv -\frac{\partial y_2(t,z)}{\partial t} - \frac{\partial y_2(t,z)}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2(t,z)}{\partial z^2} + B_C(1+y_2(t,z))\exp\left(\frac{\gamma y_1(t,z)}{1+y_1(t,z)}\right) \tag{20}$$

Applying the Galerkin principle, which enforces these residuals to be orthogonal to each of the trial functions, and exploiting the boundary conditions, the reduced model has the form,

$$M_T\dot{\alpha}(t) = -N_T\alpha(t) + F_T\beta(t) + V_T \tag{21a}$$

$$M_C\dot{\beta}(t) = -N_C\beta(t) + V_C \tag{21b}$$

where,

$$m_{Cij} = \int_0^1 \theta_j(z)\theta_i(z)dz \tag{22a}$$

$$m_{Tij} = \int_0^1 \phi_j(z)\phi_i(z)dz \tag{22b}$$

$$n_{Cij} = \int_0^1 \theta_j(z)\frac{\partial\theta_i(z)}{\partial z}dz + \frac{1}{Pe_C}\int_0^1 \frac{\partial\theta_j(z)}{\partial z}\frac{\partial\theta_i(z)}{\partial z}dz + \theta_j(0)\theta_i(0) + B_C\int_0^1 \theta_j(z)\theta_i(z)\exp(w_i)dz \tag{22c}$$

$$n_{Tij} = \int_0^1 \phi_j(z)\frac{\partial\phi_i(z)}{\partial z}dz + \frac{1}{Pe_T}\int_0^1 \frac{\partial\phi_j(z)}{\partial z}\frac{\partial\phi_i(z)}{\partial z}dz + \beta_T\int_0^1 \phi_j(z)\phi_i(z)z - \phi_j(0)\phi_i(0) \tag{22d}$$

$$f_{Tj} = -\int_0^1\left(\phi_j(z) + \frac{1}{Pe_T}\frac{\partial\phi_j(z)}{\partial z}\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz - \beta_T\int_0^1 \phi_j(z)\bar{y}_1(z)dz - \phi_j(0)\bar{y}_1(0) + B_T B_C\int_0^1 \phi_j(z)(1-\bar{y}_2(z))\exp(w_i)dz \tag{22e}$$

$$v_{Cj} = -\int_0^1\left(\theta_j(z) + \frac{1}{Pe_C}\frac{\partial\theta_j(z)}{\partial z}\right)\frac{\partial\bar{y}_2(z)}{\partial z}dz + \theta_j(0)\,[\bar{y}_2(0) - r\bar{y}_2(1)] - B_C\int_0^1 \theta_j(z)(1+\bar{y}_2(z))\exp(w_i)dz \tag{22f}$$

$$\nu_{Tj} = -\int_0^1 \left(\phi_j(z) + \frac{1}{Pe_T}\frac{\partial\phi_j(z)}{\partial z}\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz - \beta_T\int_0^1 \phi_j(z)\bar{y}_1(z)dz + \phi_j(0)\left[\bar{y}_1(0) - r\bar{y}_1(1)\right] + B_T B_C\int_0^1 \phi_j(z)(1+\bar{y}_2(z))\exp(w_i)dz \quad (22g)$$

with

$$w_i = \left(\frac{\gamma\left[\sum_{i=1}^{M_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{M_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)$$

The relevant initial conditions are,

$$\alpha_i(0) = M_T^{-1}\int_0^1 \phi_i(z)(y_{10} - \bar{y}_1(z))dz \quad (23a)$$

$$\beta_i(0) = M_C^{-1}\int_0^1 \theta_j(z)(y_{20} - \bar{y}_2(z))dz \quad (23b)$$

5. Empirical eigenfunctions

The Proper Orthogonal Decomposition—Galerkin procedure to reduce the degree of freedom of the system requires a set of empirical eigenfunctions capturing the system behaviour adequately. These eigenfunctions are obtained from an ensemble of snapshots, which are representative of the process characteristics.

Eqs. (15)–(17) were solved using the Matlab partial differential equation solver "pdepe" which solves initial-boundary value problems for systems of parabolic and elliptic type equations in one space variable and time. The spatial domain was divided into 21 equidistant intervals, and the system was solved in the time interval $t \in (0,25)$ with a sampling time of 0.125. The transient concentration and temperature fields were recorded and contaminated with noise normally distributed $N(0,1)$. An ensemble of $N = 200$ snapshots, selected randomly from the whole data, was constructed to calculate the empirical eigenfunctions; $\phi_i(z)$ and $\theta_i(z)$.

To attain model reduction, $M \ll N$ has to be selected and a nonlinear Galerkin procedure is carried out with the set of elements $\{\phi_1, \phi_2, \ldots, \phi_M\}$. The fundamental question is how to select M. The term $\sum_{i=1}^{M}\lambda_i$ represents the average kinetic energy contained in the first M modes and consequently to capture most of the energy of the system enclosed in the N Proper Orthogonal Decomposition elements, it suffices to choose M so that $\sum_{i=1}^{M}\lambda_i \approx \sum_{i=1}^{N}\lambda_i$. In fact, the ratio $\sum_{i=1}^{M}\lambda_i / \sum_{i=1}^{N}\lambda_i$ yields the percentage of the total kinetic energy in the N elements that is contained in the first M elements. Given that the associated eigenvalues are ordered $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$, it is reasonably expected to achieve a high percentage of the total kinetic energy in a reduced model of order M with M sufficiently smaller than N.

Generally, if the samplings of the fields are obtained by numerical simulation, the evaluation of $K(z,z')$ is computationally demanding. To reduce such calculations, the snapshots method proposed by Sirovich (1987) can be used instead. This technique is based on the fact that the eigenfunctions can be expressed in terms of the original set of data,

$$\phi_n(z) = \sum_{k=1}^{M_1} y(z,t_k)A_{nk} \quad (24)$$

The following matrix eigenvalue problem must be solved,

$$\sum_{k=1}^{M_1} C_{kj}A_{nk} = \lambda A_{nj} \quad \text{for} \quad j = 1,\ldots,M \quad (25)$$

where

A_{nk} is a constant associated to the nth eigenvector, and C is a symmetric and positive semi-definite temporal correlation tensor defined by,

$$C_{kj} = \frac{1}{M}\int_\Omega y_i(z,t_k)y_i(z,t_j)dz \quad (26)$$

It was observed for the case study considered here that $M_1 = M_2 = 4$ yielded a kinetic energy ratio of 0.999. These first four empirical eigenfunctions were taken into account to obtain a reduced model of the system expressed by Eqs. (21)–(22).

6. Process predictive model. New approach to calculate the temporal coefficients

In some cases, to obtain a reduced model defined by (21) is extremely difficult or even impossible because the original model is not even known. The estimation of the temporal or spectral

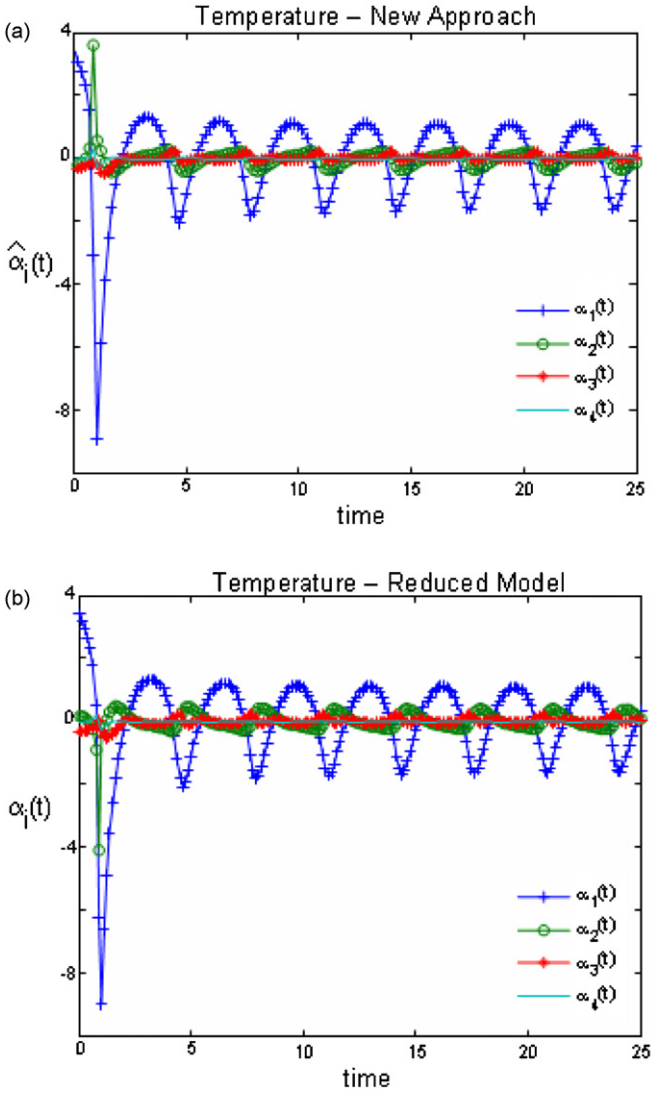


Fig. 1. Temporal coefficients. (a) Approximation (18) and (b) reduced model (21).

coefficients to define completely the reduced model (21) is a very difficult task and sometimes even impossible to achieve. A new approach to determine such coefficients was implemented in this work. This approach consists of the minimisation of an objective function based on the sum of the square errors between the original snapshots and the values predicted from (18),

$$\min_{\substack{\hat{\alpha}_i(t),\,\hat{\beta}_i(t)\\ \text{s.t.}}} \sum_{w=1}^{nv}\sum_{k=1}^{n}\sum_{j=1}^{m}(y_w(z_k,t_j)-y_{wp}(z_k,t_j))^2 \tag{27a}$$

$$y_{1p}(z,t)=\sum_{i=1}^{M_1}\hat{\alpha}_i(t)\phi_i(z) \tag{27b}$$

$$y_{2p}(z,t)=\sum_{i=1}^{M_2}\hat{\beta}_i(t)\theta_i(z) \tag{27c}$$

Here, nv stands for the number of state variables considered in the problem, k and j corresponds to the snapshot in the spatial location k and the time j for the variable $y_w(z,t)$.

Using the empirical eigenfunctions calculated in Section 5, the temporal or spectral coefficients of (18); $\hat{\alpha}_i(t)$ and $\hat{\beta}_i(t)$ are determined so the objective function (27) is minimised. These coefficients will be compared and validated with the ones obtained from the solution of the reduced model (21); $\alpha_i(t)$ and $\beta_i(t)$. The initials guess vectors used were; $\hat{\alpha}_0=\hat{\beta}_0=0$. After establishing these coefficients, the model of the system is completely defined by (18), which is the linear combination of them and the empirical eigenfunctions obtained previously.

It is well know that problem (27) is challenging for numerous reasons. The existence of a solution is not certain, particularly if the observed data contains errors or if the model is grossly incorrect. It is unusual that any parameter set can accurately match the experimental data used, especially when these data are contaminated with noise. The presence of noise can promote difficulties during the optimisation process that can cause spurious local minima and discontinuities. The gradient-based methods are the most affected of all. In this work, several different optimisation techniques (already existing in the Matlab library) were used, and it was observed that the method of Nelder–Mead produced the best results. This can be explained by the fact that using this technique, poor gradient approximations are not a problem, and continuity and differentiability of the objective function are not required.

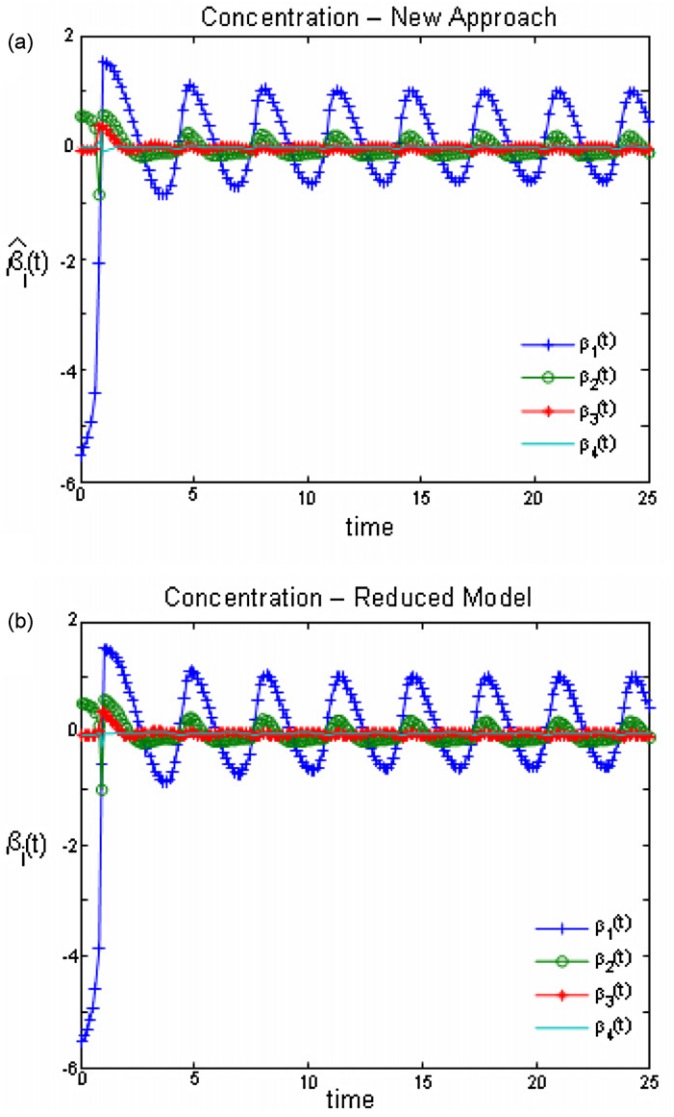


Fig. 2. Temporal coefficients. (a) Approximation (18) and (b) reduced model (21).

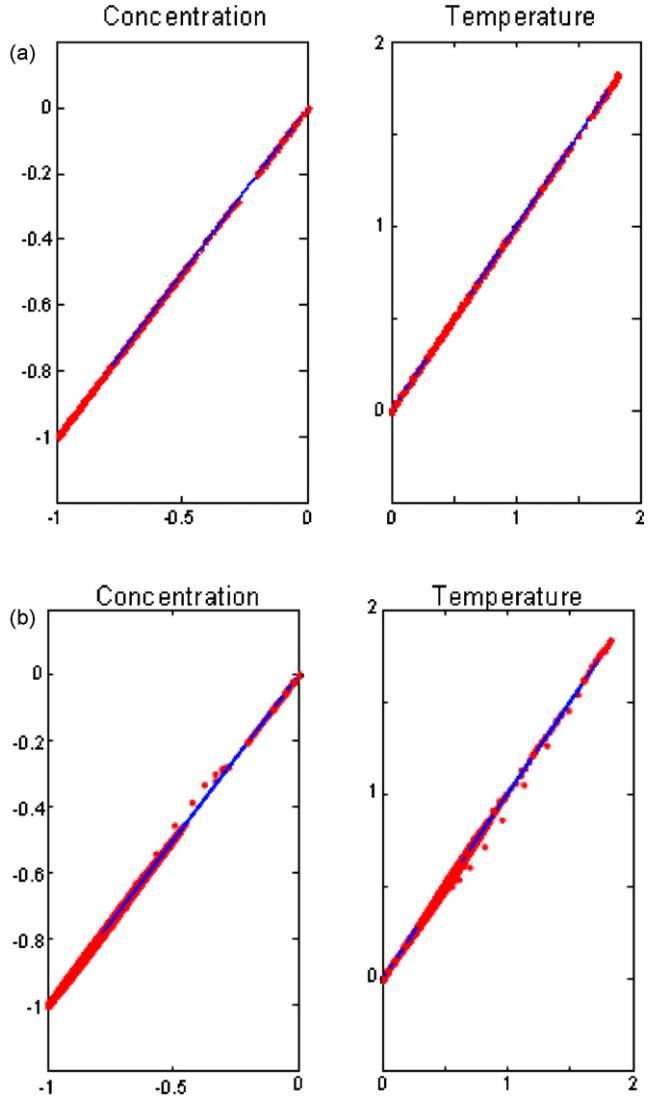


Fig. 3. Parities plot. (a) Approximation (18) and (b) reduced model (21).

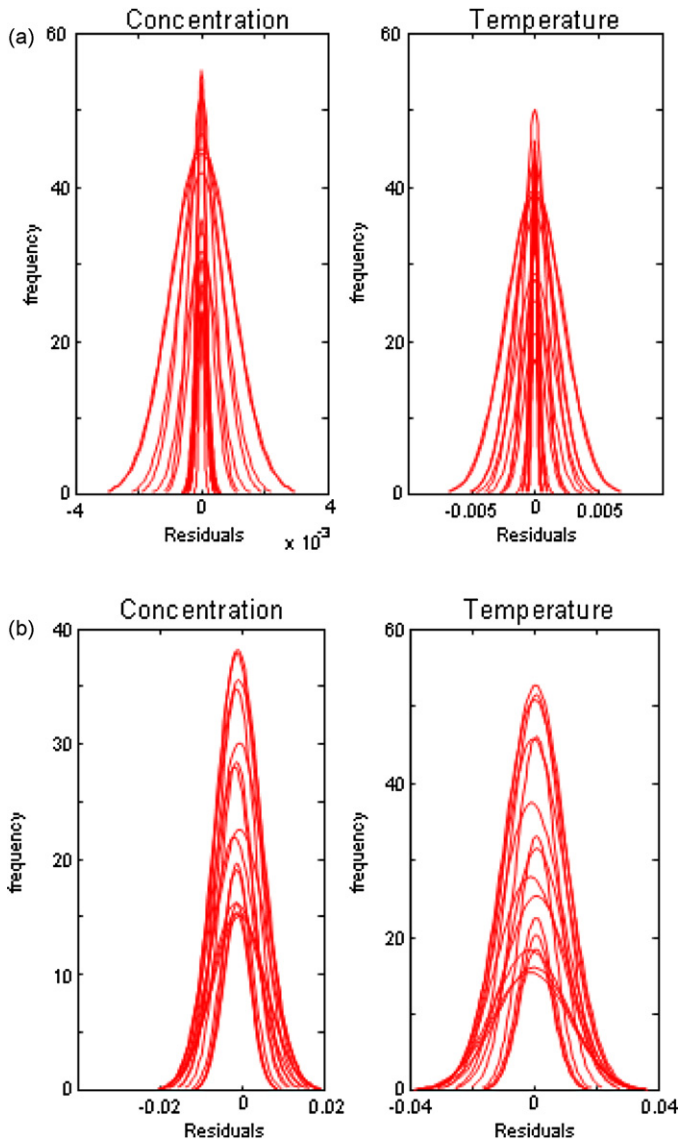


Fig. 4. Residuals distribution. (a) Approximation (18) and (b) reduced model (21).

7. Numerical results

Figs. 1 and 2 depict the temporal coefficients for the temperature and concentration approximations obtained using Eq. (18) and the reduced model (21).

From Figs. 1 and 2, it can be seen that the results agreed satisfactorily using the two different calculation methods in the same time interval.

The predicted variables are compared with the original fields in order to verify the accuracy of the reduced model and the approximation considered here. The parity plot and the residuals for each field are depicted in Figs. 3 and 4.

Both approaches produced very good results, but it is approximation (18), with temporal coefficients calculated from the optimisation procedure that reports the best solution.

The results obtained using the temporal coefficients from the minimisation procedure produced the lowest residuals on the reconstruction of the field. The solution of the reduced model (21) involves several numerical methods to calculate the temporal coefficients (numerical integration, matrix inversion, numerical derivation, etc.). The errors accumulate during all these calculations to solve the model and reproduce the field, which increase the deviations between the predicted outputs and the experimental observations.

The optimisation problem (27) consists of the minimisation of a linear objective function, the temporal coefficients estimation can be done straightforwardly and easily with the lowest deviations from the experimental snapshots. This aspect was verified using several different systems, linear and nonlinear, and similar results were obtained. Additionally, it was observed that the optimisation process takes, on average, 40% less time than to solve the reduced model (21). The optimal solution of (27) is reached independently of the values used as initial guess.

8. Conclusions

This work has introduced and presented a new approach to identify processes without any knowledge of the mathematical model of the system. Only a few experimental data are required to determine the empirical eigenfunction and the temporal coefficients. Very good results are guaranteed for any case study and/or system. This is because it involves the minimisation of a linear objective function subject to a linear constraint defined by the linear combination of the empirical eigenfunctions and the temporal coefficients. This method, which is easy and fast to implement, produces the lowest deviation from the original experimental data and is very useful when the theoretical model of the system is unknown or difficult to determine. The optimal solution is obtained independently of the values used as initial guess. It is worth mentioning that the same advantages of this approach are observed for any other empirical approximator or ansatz used.

References

Ball, K. S., Sirovich, L., & Keefe, L. R. (1991). Dynamical eigenfunctions decomposition of turbulent channel flow. *International Journal for Numerical Methods in Fluids*, *12*, 585.

Bendersky, E., & Christofides, P. D. (2000). Optimisation of transport-reaction processes using nonlinear model reduction. *Chemical Engineering Science*, *55*, 4349.

Berkooz, G., Holmes, P., Lumley, J. L., & Mattingly, J. C. (1997). Low-dimensional models of coherent structures in turbulence. *Physics Reports—Review Section of Physics Letters*, *287*, 338.

Berkooz, G. (1992). Observations on the proper orthogonal decomposition. In *Studies in turbulence*. New York: Springer Verlag., p. 229.

Chambers, D. H., Adrian, R. J., Moin, P., Stewart, D. S., & Sung, H. J. (1988). Karhunen–Loève expansion of Burgers' model of turbulence. *Physics of Fluids*, *31*, 2573.

Gay, D. H., & Ray, W. H. (1995). Identification and control of distributed parameter systems by means of the singular value decomposition. *CES*, *50*(10), 1519.

Hilai, R., & Rubinstein, J. (1994). Recognition of rotated images by invariant Karhunen–Loève expansion. *Journal of the Optical Society of America A—Optics Image Science and Vision*, *11*, 1610.

Karhunen, K. (1946). Zur spectral theorie stochasticher prozess. *Annales Academiae Scientiarum Fennicae, Series A-I (Mathematica-Physica)*, 37.

Kirby, M., & Sirovich, L. (1990). Application of the Karhunen–Loève procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 103.

Li, M., & Chistofides, P. D. (2007). Optimal control of diffusion-convection-reaction processes using reduced-order models. *Computers and Chemical Engineering*, doi:10.1016/j.compchemeng. 2007.10.018

Loève, M. (1945). Sur les Fonctions aléatories stationaires du second ordre. *Revue Scientifique*, *83*, 297.

Lumley, J. L. (1970). *Stochastic tools in turbulence*. New York: Academic Press.

Moin, P., & Moser, R. D. (1989). Characteristic-eddy decomposition of turbulence in a channel. *Journal of Fluids Mechanics*, *200*, 417.

Obukhov, A. M. (1954). Statistical description of continuous fields. *Geophys. Inst. Acad. Sci. SSSR*, *24*(151), 3.

Pougachev, V. S. (1953). General theory of the correlations of random functions. *Izvestiya Akademi Nauk USSR*, *17*, 1401.

Rajaee, M., Karlson, S. K. F., & Sirovich, L. (1994). Low-dimensional description of free-shear-flow structures and their dynamical behaviour. *Journal of Fluid Mechanics*, *258*, 1.

Reichert, R. S., Hatay, F. F., Biringers, S., & Husser, A. (1994). Proper orthogonal decomposition applied to turbulent flows in a square duct. *Physics of Fluids Mechanics*, *6*, 3086.

Sirovich, L. (1991). Analysis of turbulent flows by means of the empirical eigenfunctions. *Fluid Dynamics Research*, *8*, 85.

Sirovich, L. (1989). Chaotic dynamics of coherent structures. *Physica D*, *37*, 126.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. *Quarterly of Applied Mathematics*, *5*, 561.

THE PUBLICATIONS UNDER REVIEW

“Perfect numbers like perfect men are very rare”

Descartes

Optimal measurement locations for Parameter Estimation of Distributed Parameter Systems based on the use of Artificial Neural Networks

Alaña Jorge Enrique
Universidad Del Zulia, Facultad de Ingeniería, Escuela de Ingeniería Química, Maracaibo, Venezuela
School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, UK
Fax: +58 261 7598100, Email: jalana@fing.luz.edu.ve

Abstract

The optimal measurement location problem on the spatial domain of a distributed parameter system is considered. Artificial neural networks were trained using data generated from simulations. These data were obtained introducing changes in the parameters of the system (defined as inputs to the neural network). The chain rule was applied and an analytical expression relating the predicted output and the inputs of the neural networks was found. These relationships can be used as an approximation of the sensitivity coefficients. Therefore, it is possible to calculate the Fisher Information Matrix and/or the covariance matrix of parameter estimates, which is very helpful for optimal experiment design purposes. Additionally, it can be used to predict the outputs of the system during the parameter estimation process. It is established that the optimal experimental samplings should be made at the positions where the sensitivity coefficients reached their largest absolute values in the spatial domain. A very well trained neural network can be used as a model predictor and sensitivity coefficients estimator at the same time.

Key words: sensitivity coefficient, artificial neural network, measurement location, distributed parameter system

1. Introduction

The literature on optimal sensor position is plentiful, but mostly for state estimation. A variety of metrics has been proposed to place transducers in a *distributed parameter system* (DPS). Most of those techniques are based on an exhaustive exploration over a pre-defined set of candidates [28]. These approaches, suitable for a reduced number of locations, become

unpractical when the number of possible positions increase. In the last two decades, this problem has been studied considering different kind of criterion to establish the optimal sensor locations. Omatu et al. [41] and Harris et al. [21] used some metrics of the error covariance matrix of parameters estimates, its trace, and determinant. A qualitative information of the system as well as the minimisation of the estimated error by unobservable subspaces were taken into account in the works of Jørgensen et al. [26], Morari and Stephanopoulos [38], and Morari and O'Dowd [39]. Another category of techniques considered observability matrix and/or gramian, including their smallest eigenvalue, determinant, and trace of their inverses [16,52,54]. These approaches are based on very complex mathematical concepts and require a large number of calculations.

Lopez and Alvarez [34] and Damak et al. [15] introduced the nonlinear observability analysis for sensor location. Their contributions considered the measurement cost and transducer failure as well as the system information. These measures are reduced to a criterion based on the linear observability matrix, if the system is linear and time invariant. To solve this problem, Georges [19] used nonlinear observability gramians, but due to numerical difficulties, this method is limited to low-order systems.

Wouwer et al. [57] and Alonso et al. [6,7] offered significant contributions on sensor placement for nonlinear systems, based on the Gram determinant evolution in the spatial domain and using a max-min optimisation process, respectively. The maximisation of the Gram determinant is a measure of the linear independence of the sensitivity functions evaluated at sensor locations. This procedure guarantees that the parameters are identifiable and the correlation between the sensor outputs is minimised. The form of the criterion itself resembles the D-optimality approach proposed by Qureshi et al. [44] and Rafajlowicz [45]. The counterpart of the *Fisher Information Matrix* (FIM) used takes on much larger

dimensions, which suggest that the approach involves more calculations that are cumbersome. On the other hand, Alonso et al. [7] showed that the most appropriated location for the measurements for state estimation is found by the solution of a max-min optimisation problem, with the number of sensors predefined by the user.

A closely related problem is sensor placement for parameter estimation, but only few contributions for this subject can be found in the literature [9,29,59]. This is mainly because a direct extension of sensor location for parameter estimation from the results obtained for state estimation is not straightforward and has not been pursued. These two problems are essentially different, in the first for instance the current state usually depends strongly on nonlinearly of the unknown parameters.

Most of the existing optimality criteria for sensor location for parameter identification are based on scalar measures of the FIM, e.g. its determinant. Using a related approach relying on principal component analysis of the output sensitivity matrix, Li et al. [31] computed the best set of parameters to be estimated if a collection of measurement positions is already given. These techniques require computation of the parametric sensitivity coefficients based on local sensitivity analysis. Therefore, the results obtained might not capture the nonlinearity and may only be suitable for small changes in the parameters. Singh and Hahn [47] have computed sensor position with information derived from observability covariance matrices. They proposed this approach based on the degree of observability of a nonlinear system over a certain operating region [55]. Nahor et al. [40] minimised the ratio of the largest to the smaller eigenvalue of the FIM to compute optimal temperature sensor positions for food process. To the best of our knowledge, most of these techniques have not been applied yet to systems described by DPS, in spite of their resolute advantages [13,53,56].

In summary, the state of the art indicates that more attention should be paid to the sensor position for parameter estimation problem, especially for DPSs. From an engineering point of view, the use of the existing scarce methods is restricted due to computational difficulties [30,36].

More recently, some new approaches introduce the use of *artificial neural networks* (ANN) to estimate parameters [1,5,32,37]. These ANNs are considered universal approximator capable of reproduce an unknown mapping and its derivatives arbitrarily well [24].

Approximating the derivatives of the output with respect to the inputs is of significant importance in many applications. For instance, in process optimisation, the first and second order derivatives may be used in estimating the gradient vector and the Hessian matrix. This allows the use of techniques such as Newton's method to optimise the process under consideration [61].

The capability of a multilayer ANNs to replicate an unknown mapping $f : \Re^n \to \Re$ arbitrarily well has been investigated by Cybenko [14], Funahashi [18], Hecht-Nielson [22], Hornik, Stinchcombe and White [24]. Moreover, Hornik et al. [25] showed that a multilayer *feedfoward neural network* (FFNN), with as few as a single hidden layer with appropriately smooth activation function, is capable of accurately approximate to a function and its derivatives. The realisation to be approximated need not be differentiable in the classical sense as long as it possesses a generalised derivative, which is the case for certain piecewise differentiable functions [10,27]. This fundamental result provides the necessary theoretical foundation for sensitivity analysis and definition of the optimal sensor locations for parameter estimation purposes presented.

In this paper, the largest absolute value analysis of the sensitivity coefficients to define the optimal number and the spatial positions of the experimental measurements is introduced. The sensitivity coefficients are determined using an ANN. The data used to train these ANNs correspond to simulations of the system varying the value of the parameters. The approximations obtained are then used to calculate the sensitivity coefficients of the system, which allow defining a set of optimal sensor locations. This information can be employed to compute the FIM and the covariance matrix of the parameter estimates. Additionally, the ANN can be utilised as predictive model during the parameter identification procedure.

2. Optimal experiment design

Optimal experiment design for precise parameters estimation can be defined as the determination of the experimental conditions which leads to parameter estimates with the best possible precision [17,45].

To understand this concept, let us consider a general DPS with $\Omega \subset R^2$ as a bounded simply connected open domain with a sufficiently smooth boundary, $\partial\Omega$. A given DPS can be represented by a system of n (possibly non-linear) partial differential equations (PDEs) as follows [50,51],

$$G\left(x, y, t, \theta, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial^2 y}{\partial x_1^{\,2}}, \frac{\partial^2 y}{\partial x_2^{\,2}}\right) = 0 \qquad x \in \Omega \tag{1}$$

Where $x=(x_1,x_2) \in \overline{\Omega}$ is the vector of spatial coordinates, $y=y(x,t)$ the vector of state variables $\in R^n$ and G some function mapping its arguments to R^p, possibly including terms accounting for *a priori* known forcing inputs. The boundary conditions are,

$$\vartheta\left(x, y, t, \theta, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}\right) = 0 \qquad x \in \Omega \tag{2}$$

Here ϑ is some known function that maps its arguments to R^r. The model above contains a vector of unknown parameters $\theta \in R^{nv}$ to be estimated based on the observations,

$$\eta_{ij} = y\left(x^i, \theta\right) + \varepsilon_{ij}, \qquad i = 1,..., p \qquad j = 1,..., r_i \tag{3}$$

Made at point x^i, where ε_{ij} denotes the noise in measurements satisfying a Gaussian distribution,

$$E\left\{\varepsilon_{ij}\right\} = 0, \qquad E\left\{\varepsilon_{ij}\varepsilon_{qs}^T\right\} = \delta_{iq}\delta_{js}C\left(x^i\right) \tag{4}$$

With δ_{ij} being the Kronecker delta and $C(x^i) \in R^{nxn}$ a known positive-definite matrix. Replicated measurements can be admitted, i.e., some values x^i may appear several times in the optimal solution (an unavoidable consequence of independent measurements). Then it is sensible to distinguish only the components of the sequence $x^1,\ldots,x^N$, which are different and, if there are p of such components, to re-label them as $x^1,\ldots,x^p$ while introducing $r^1,\ldots,r^p$ as the corresponding numbers of replications.

It is assumed that some, albeit rough, a priori estimate of the parameter vector θ^0 is available e.g. from preliminary experiments. The goal here is to select the state sampling policy, which will maximise the accuracy of the parameter estimates from data generated in the corresponding new experiments. It has been shown that this task corresponds to the maximisation of a scalar concave function of the FIM whose inverse gives a lower bound on the covariance matrix of the estimates [8,12]. The average per observation (or normalised) FIM takes the following form,

$$M\left(\xi^*\right) = \sum_{i=1}^{p} \omega_i J_y{}^T\left(x^i\right) C^{-1}\left(x^i\right) J_y\left(x^i\right) \tag{5}$$

Where $\omega_i = r_i / N$, $N = \sum_{i=1}^{p} r_i$ and $J_y\left(x^i\right) = \dfrac{\partial y\left(x^i, \theta\right)}{\partial \theta}$ is the Jacobian matrix consisting of the sensitivity coefficients. The collection of variables, ξ^* that define the exact design of the experiment is given below,

$$\xi^* = \begin{Bmatrix} x^1, & x^2, & \dots & ,x^p \\ \omega_1, & \omega_2, & \dots & ,\omega_p \end{Bmatrix} \tag{6}$$

The proportion ω_i of observations made at x^i can be considered as the percentage of experimental effort spent at that location. It can be assumed that ω_i's can be any real numbers in the interval [0,1] such that $\sum_{i=1}^{p} \omega_i = 1$ and in terms of the design they can be considered as probability distributions on $\left(x^1, x^2, ..., x^p\right)$. The purpose of the optimal experimental design can be then defined as the search for a design $\xi^* \in \Xi(X)$, which satisfies [55],

$$\Psi\left[M\left(\xi^*\right)\right] = \max_{\xi \in \Xi(X)} \Psi\left[M(\xi)\right] \tag{7}$$

Here $\Xi(X)$ is the set of all admissible designs (i.e. all probability distributions on the spatial domain X, where measurements are allowed). This can be formulated using a metric of the covariance matrix of the parameter estimates. Among these, the most widely used design criteria are [50],

- A-optimality: A design is said to be A-optimal if it minimises the trace of the covariance matrix.
- D-optimality: A design is said to be D-optimal if it minimises the determinant of the covariance matrix.
- E-optimality: A design is said to be E-optimal if it minimises the maximal eigenvalue of the covariance matrix.

3. Parameter estimation

Model calibration consists of finding a set of parameter values that produces the best model output, which fit the observed data. This calibration is usually done by the minimisation of the objective function presented in equation (8).

Once the sensor locations have been determined, the parameter estimation problem can be viewed as matching the model to the real system through the minimisation of an error criterion over a set of admissible parameters. This can be defined as,

$$J(\theta) = \frac{1}{m}\left[\frac{1}{t_f}\int_0^{t_f}\sum_{i=1}^{m}\left(y_e(x_i,t) - y_p(x_i,t,\theta)\right)^2 dt\right] \tag{8}$$

Where, $y_p(x_i,t,\theta)$ is defined as the predicted response of the model at location x_i, and $y_e(x_i,t)$ as the experimental response at the same location, m is the number of optimal measurement locations. Unfortunately, elements in the vector $y_p(x_i,t,\theta)$ are not linear functions of the parameters θ, and multiple solutions of the equation (8) are possible. This vector can be either obtained by the use of the original governing equations of the system or from a very well trained non-parametric model (as will be shown later in the equation (14)).

3.1. Covariance matrix of parameter estimates

When assessing the quality of an identified dynamic model, the covariance matrix of the estimated parameters gives an important measure [33,48]. In a maximum likelihood context, the inverse of the expected value of the negative of the Hessian provides the Cramer-Rao lower bound on the variance-covariance matrix of the parameter estimates. Large variances imply imprecise parameter estimates. The arrival of new data, or data measurement, can lead to substantial change in parameter estimates if the variance-covariance matrix is ill-posed. Therefore, the goal of the estimation procedure is to determine unknown parameters in such a way that the difference between the sample covariance matrix and the implied covariance matrix is minimised in a certain sense [60].

For the systems studied here, it is usual to obtain the covariance matrix of parameters estimates [58] by the following approximation,

$$Cov\left(\hat{\theta}_{nv}\right) \approx \left(\frac{1}{\sigma_y^2} J_y^T J_y \right)^{-1} \tag{9}$$

Where, J_y represents the Jacobian matrix of estimated variable *y* with respect to changes in the parameters, *nv* is the parameter dimension, and σ_y^2 is the variance of the variable *y*. In this equation, it is easy to view the influence of field data quality and quantity in parameter uncertainty. Since these Jacobian matrices are evaluated at those locations for which observations are available, any experimental design should aim at sampling at those locations where the variables are most sensitive to the estimated parameters. Such a design is said to provide the maximum amount of information about the unknown parameters [29].

It is intuitively obvious that the experimental design objective should be intended at minimising the norm of (9); that is, to make matrix $\left(J_y^T J_y\right)^{-1}$ as small as possible. The variance term σ_y^2, it is often admitted constant for simply the numerical procedure, and can be dropped from the formulation. To measure the accuracy of the estimates we like to summarise the information about the variability in the covariance matrix into a single number. Here we used the determinant of (9) as the function that transforms a matrix into a scalar. This is quite informative as in fact it relates to the volume of the multidimensional simplex defined by the column/row vectors of the matrix [23,46].

Using different norms leads to different conclusions regarding the optimal design. The D-optimality criterion minimises the volume of the hyper-ellipsoid in the parameter space, with no consideration of the relationship between the ellipsoid's axes lengths, which are in turn proportional to the square root of the covariance matrix eigenvalues.

It can be shown that under some assumptions of regularity or for a sufficiently large sample size N the vector $\left(\hat{\theta}_{(N)} - \overset{*}{\theta}\right)$ (with $\overset{*}{\theta}$ denoting the 'true' but unknown value of the parameters and $\hat{\theta}_{(N)}$ obtained using the least square parameter estimates) has approximately a normal distribution with zero mean and covariance matrix,

$$\frac{\sigma^2}{N} M^{-1}\left(\xi, \overset{*}{\theta}\right) \tag{10}$$

Where σ^2 denotes the standard deviation of the errors in the model. In principle, the covariance matrix is a measure for the precision of the least square estimator for the unknown parameter $\overset{*}{\theta}$ and a 'smaller' matrix yields estimates that are more precise. For example, the *ith* diagonal element of (10) will be denoted by $\frac{\sigma^2}{N} M^{-1}\left(\xi, \overset{*}{\theta}\right)_{ii}$ and approximates the variance or mean squared error for the *ith* component $\hat{\theta}_{i(N)}$ of the least squared estimator $\hat{\theta}_{(N)}$. A confidence interval for the *ith* component θ_i of the vector θ is given by,

$$\left[\hat{\theta}_{i(N)} - \frac{\hat{\sigma} u_{1-\alpha/2}}{\sqrt{N}} \sqrt{M^{-1}\left(\xi, \overset{*}{\theta}\right)_{ii}}, \hat{\theta}_{i(N)} + \frac{\hat{\sigma} u_{1-\alpha/2}}{\sqrt{N}} \sqrt{M^{-1}\left(\xi, \overset{*}{\theta}\right)_{ii}}\right] \tag{11}$$

Where $u_{1-\alpha/2}$ denotes the 1-α/2 quantile of the standard normal distribution. For most cases, it was shown that for moderate sample sizes N the sampling variances of the parameter estimates are well approximate.

If the parameters values do not change significantly the imprecision of the estimates can always be decreased by increasing the sample size N, which yields a 'smaller' covariance matrix and smaller variances of the least square estimates. However, in practice the sample size is usually fixed, due to cost considerations of each additional experiment. To improve the quality of the estimates or, from a different point of view, to reduce the number of experiment

measurements needed to obtain the estimates with a given accuracy, we note that the variances of the estimates $\hat{\theta}_{i,(N)}$ and the covariance matrix of the vector $\hat{\theta}_{(N)}$ also depend on the given design, ξ, which determines the relative proportion of total observations to be taken at the experimental locations.

It is desirable to check that the confidence interval magnitude agrees with the sensitivity analysis, the parameter with the largest sensitivity coefficient should have the smallest confidence interval. Because of the consideration of the measurement error only, if the FIM is well-defined, the confidence intervals may result in very small values. Grimstat and Mannseth [20] indicate that the use of such approximation of the confidence intervals was usually justified even for the highly nonlinear models analysed.

4. Sensitivity coefficients using Artificial Neural Networks

Among the many ways to develop mathematical models, artificial neural networks, have found successful applications in the approximation of multivariable nonlinear functions, their modelling, and control. A model based on an ANN is a typical black-box system where the system structure is implied by its weights. These weights carry no obvious meaning and the ANN only establishes a mapping of inputs to outputs. Therefore, the model cannot be formulated by explicit expressions. The sensitivity coefficients $s_{y,\theta_i} = \frac{\partial f(\theta)}{\partial \theta_i}$ in the model cannot be calculated, which makes it impossible to estimate measurement uncertainty on the results. This problem can be solved by setting up a parameterised model based on an ANN using *Radial Basis Functions* (RBF) and/or FFNN. An analytic equation of the model and the computing formula of its sensitivity coefficients can be obtained. The effectiveness and precision of this approach can be proved by simulation [25,61].

An indirect measurement model can be denoted by,

$$y_p(x,t,\theta) = f(\theta_1,\theta_2,...,\theta_n) \tag{12}$$

Where $y(x,t,\theta)$ is the predicted output and $\theta_1,\theta_2,...,\theta_n$ are the input variables.

The model (12) is usually unknown and sometimes is difficult to establish. Therefore, a RBF neural network with n inputs and one output (vector) can be used as indirect measurement model. The mapping relation of the RBF neural network is,

$$y_p(x,t,\theta) = f(\theta) = w_0 + \sum_{i=1}^{r} w_i \phi(\|\theta - c_i\|) \tag{13}$$

Substituting the transfer function in equation (13),

$$y_p(x,t,\theta) = w_0 + \sum_{i=1}^{r} w_i \, exp\left(- \frac{\sum_{j=1}^{n} |\theta_j - c_{ij}|^2}{2\sigma_i^2} \right) \tag{14}$$

This equation is the analytic form of the model based on RBF neural network. θ is the input vector $[\theta_1,\theta_2,...,\theta_n]$, $y_p(x,t,\theta)$ is the predicted output, c_i is the centre vector $[c_{i1},c_{i2},...,c_{in}]$ of RBF neural network, r is the number of joints at the hidden layer of the ANN, σ_i is the spread of centre *ith*, and $W = [w_0,w_1,...,w_r]$ is the weight vector of output layer. This vector W can be determined by the learning and training process of the ANN, and given them, an analytic denotation of the model can be obtained. In this work, a K-mean clustering [35] and the orthogonal least square learning algorithm [11] were used to train the RBF ANN and calculate its centres and weights..

4.1 Sensitivity coefficient calculation

First order output sensitivity is computed by applying a simple backward chaining differentiation rule. For the equation (14), the sensitivity coefficients which reflect the

propagation law of measurement uncertainty of each θ_j *(j=1,2,…,n)* can be calculated as following,

$$s_{y,\theta_j} = \frac{\partial y_p(x,t)}{\partial \theta_j} = \sum_{i=1}^{r} w_i \, exp\left(- \frac{\sum_{j=1}^{n} \left| \theta_j - c_{ij} \right|^2}{b} \right) \frac{-2\left(\theta_j - c_{ij}\right)}{b} \tag{15}$$

That is,

$$s_{y,\theta_j} = \sum_{i=1}^{r} \frac{-2w_i\left(\theta_j - c_{ij}\right)}{b} exp\left(- \frac{\sum_{j=1}^{n} \left| \theta_j - c_{ij} \right|^2}{b} \right) \tag{16}$$

This can be expressed as,

$$s_{y,\theta_j} = \sum_{i=1}^{r} \frac{-2w_i\left(\theta_j - c_{ij}\right)}{b} exp\left(- \frac{\left\| \theta - c_i \right\|^2}{b} \right) \tag{17}$$

Equation (17) is the calculation formula of the sensitivity coefficient of the model based on a RBF neural network. Here *b* represents the spread of the corresponding centre. This approximation allows calculating the FIM, which can be used to estimate the confidence intervals given by (11) and to determine the optimal locations following a D-optimal design and any other approach requiring the sensitivity coefficients of the system.

A forward predictor is obtained when the ANN is trained (14), which it can be used to calculate the outputs of the system in (8) during the parameter estimation process. Here some optimisation techniques based on gradient information were implemented using (17) to approximate some of its elements. It is our intention to introduce the use of well-trained ANNs to carry out the calculations required to determine the optimal sensor locations for parameter estimation purposes.

5. Case studies

To study the capabilities of the ANN to calculate the sensitivity coefficients of a system, without a strict previous knowledge of the parameters, the following examples from the literature were employed.

Example 5.1. The heat equation problem described by the one-dimensional partial differential equation [42],

$$\frac{\partial y(x,t)}{\partial t} = \theta_1 \frac{\partial^2 y(x,t)}{\partial x^2} \qquad t > 0, x \in (0,1) \tag{18}$$

With the boundary and initial conditions,

$$\frac{\partial y(x,t)}{\partial x} = -\theta_2 (y_a - y(x,t)) \qquad t > 0, x = 0 \tag{19.a}$$

$$\frac{\partial y(x,t)}{\partial x} = \theta_3 (y_a - y(x,t)) \qquad t > 0, x = 1, \quad and \quad y_a = 600 \tag{19.b}$$

$$y(x,0) = 500 \qquad t = 0, x \in (0,1) \tag{19.c}$$

This example is used to demonstrate the effectiveness of the optimal sensor location approach introduced in this work. Here the problem consists of determining the best measurement positions to estimate the parameters θ_1, θ_2 and θ_3. The nominal values for these parameters are; $\theta_1 = 0.10$, $\theta_2 = 1.00$ and $\theta_3 = 1.00$.

Example 5.2. A system described by the following nonlinear partial differential equation [50],

$$\frac{\partial y(x,t)}{\partial t} = \theta_1 \frac{\partial^2 y(x,t)}{\partial x^2} - \theta_2 y(x,t)^3 \qquad t \in (0,t_f), x \in (0,1) \tag{20}$$

Subject to the conditions,

$$y(x,0) = \theta_3 \Psi(x) \qquad x \in (0,1) \tag{21.a}$$

$$y(0,t) = y(1,t) = 0 \qquad t \in (0,t_f) \tag{21.b}$$

Where; t_f=0.8 and;

$$\Psi(x) = sin(\pi x) + sin(3\pi x) \tag{22}$$

This problem consists on the determination of the best measurement positions to estimates θ_1, θ_2 and θ_3 optimally. The 'true' values for such parameters, reported in the literature, are; $\theta_1 = 0.10$, $\theta_2 = 2.00$ and $\theta_3 = 5.00$.

6. Numerical experimentation

The original governing equations were solved using the Matlab® partial differential equation solver "pdepe". This command solves initial-boundary value problems for systems of parabolic and elliptic partial differential equations in one space variable and time. The system is solved several times using different set of values for the parameters, far from their nominal values, introducing uncertainties, which factors are shown in Table 1. These values are far from the nominal values of the system. This will allow evaluating how good the estimation of the sensitivity coefficients can be. The spatial domain was divided into *m*=11 equidistant intervals. The system was solved simultaneously with the equations describing the transient behaviour of the sensitivity functions, as the direct method for sensitivity analysis is usually implemented [4]. A sampling time of 0.25 was used and a mesh solution of dimensions 11x4 is obtained. The spatial–temporal distributions of the responses and the sensitivity coefficients were obtained and compared.

Table 1
Set of parameters and uncertainties considered to generate the training data

	θ_1	θ_2	θ_3
Example 5.1	0.10 [±0.050]	0.10 [±0.050]	0.01 [±0.005]
	3.00 [±1.50]	3.00 [±1.50]	1.00 [±0.50]
	3.00 [±1.50]	3.00 [±1.50]	0.01 [±0.005]
	0.10 [±0.050]	0.10 [±0.050]	1.00 [±0.500]
Example 5.2	0.05 [±0.005]	3.00 [±1.50]	3.00 [±1.50]
	1.00 [±0.50]	1.00 [±0.50]	10.0 [±2.50]
	1.00 [±0.50]	1.00 [±0.50]	3.00 [±1.50]
	0.05 [±0.005]	3.00 [±1.50]	10.0 [±2.50]

Numbers in brackets are the uncertainty factors considered

An ANN was trained using the 70% of the data produced in the previous step. Two different kind of ANN were trained; FFNN and RBF, and their performance and prediction capabilities were compared. It was assumed that the activation functions of the ANN are differentiable. The remaining 30% of the data collected from the simulations were used to validate the ANNs in order to test their prediction capabilities using observations unknown by the ANNs. Under this assumption, the formulas to compute the first order derivatives from the trained network can be obtained easily as it was shown in section 4.1.

The ANNs were initialised with independent random weights. Each network had as inputs the parameters of the systems [$\theta_1,\ldots, \theta_n$] and as outputs the response of the system at x_i and the interval of time t; [$y(x_1,t),\ldots,y(x_m,t)$]. The activation function for the hidden units was the logistic sigmoid function. The unconstrained optimal weights were estimated based on the data in the training set. The number of hidden neurons was changed until good estimates of the outputs were obtained. The centres for each processing unit are calculated using the K-mean clustering algorithm. These ANNs are then used to approximate the first derivatives of the system by equation (17).

7. Numerical Results

7.1 Sensitivity coefficients and the optimal measurement locations

The sensitivity coefficients for each parameter, $s_{y,\theta}$, calculated by the well trained and validated ANN, and the ones obtained from the application of the Direct method considering the nominal values of the parameters are compared. The normalised values, $s_{y,\theta}$, for each example are shown Fig. 1 and 2 [2,4].

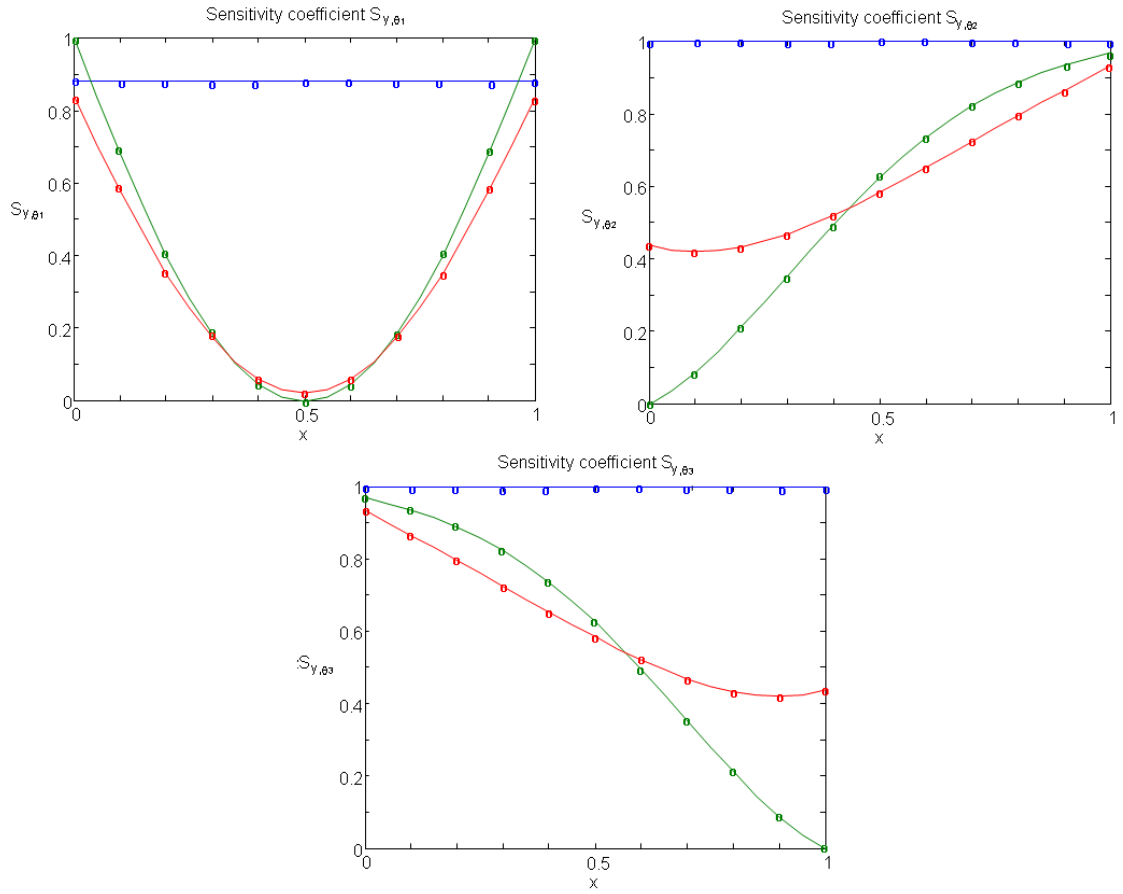


Fig. 1. Spatial – temporal behaviour of the sensitivity coefficients, Example 5.1. Circles represent the values estimated by the ANN.

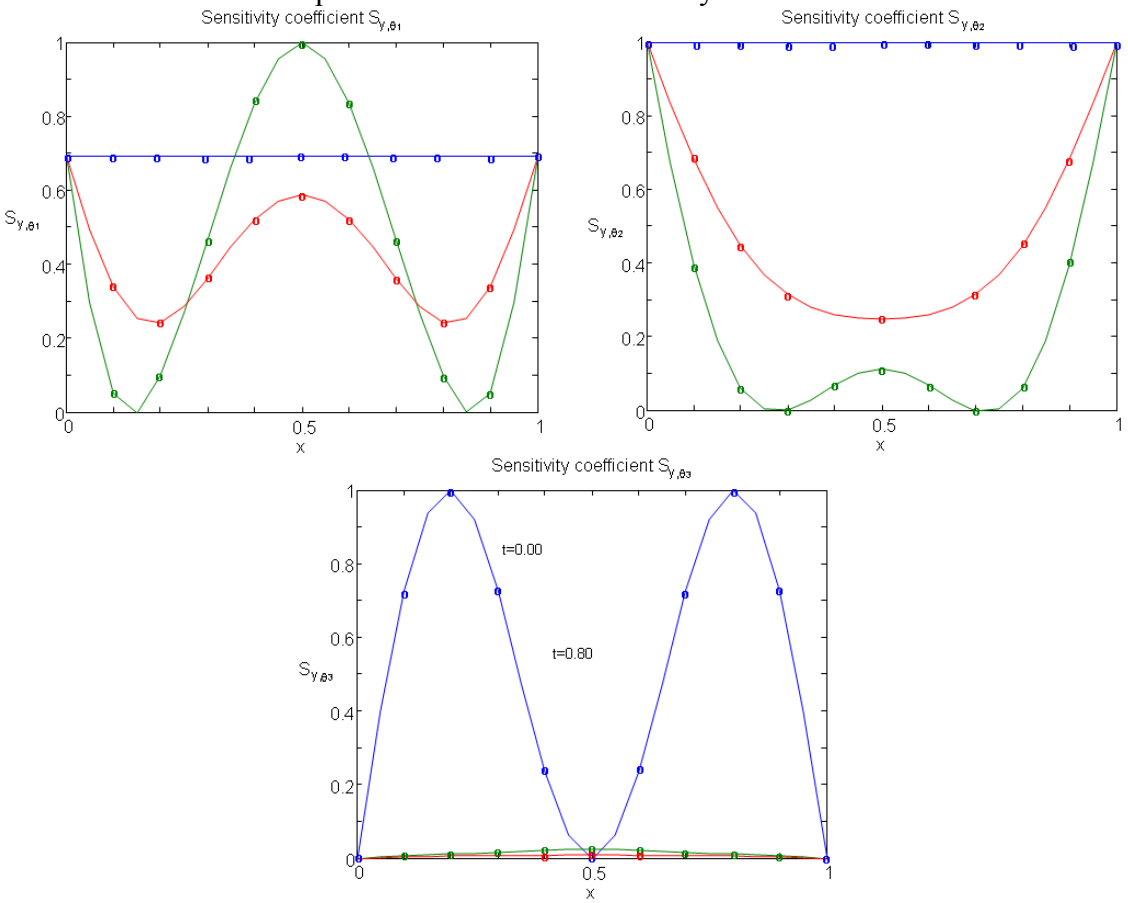


Fig. 2. Spatial – temporal behaviour of the sensitivity coefficients, Example 5.2. Circles represent the values estimated by the ANN.

From these figures, it can be observed how efficiently the ANN can reproduce the sensitivity coefficients for any set of inputs used. It seems that a well-trained ANN, which adequately approximates a given function, will provides good estimate of the derivatives of such function. In addition, can be seen clearly, where the largest absolute values of the sensitivity coefficients are located. For all the combinations of parameter's values, the extrema of the sensitivity coefficients were recorded. Example 5.1 showed that these extrema are located at positions x=(0.00, 0.50, 1.00) for any value considered on the parameters. For example 5.2, the locations obtained are x=(0.20, 0.50, 0.80). For each case study, an ANN with 40 processing units was required to reproduce the targets of the systems satisfactorily.

A D-optimal experiment design was performed for each case study following the procedure depicted in Fig. 3 and explained in section 2. The results obtained coincided with the ones derived from the extrema sensitivity analysis carried out here.

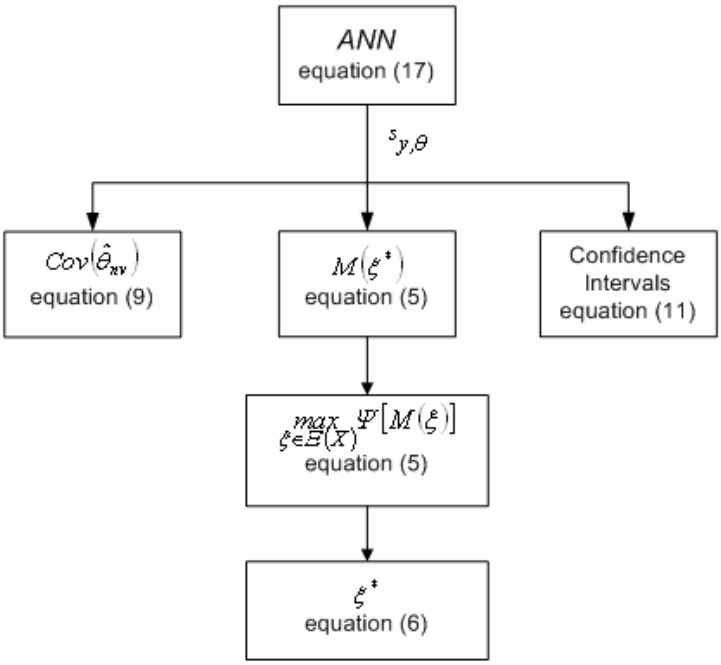


Fig. 3. Post-processing of the ANN predictions

Two different kind of ANN were used in the approaches explained above and it was observed

that the RBF ANN showed a better performance predicting the sensitivity coefficients than the FFNN. It is worth to mention that the methods explained and implemented in this work do not depend of the accuracy of the sensitivity coefficients calculated but of their spatio-temporal behaviour.

7.2. Parameter estimation

To estimate the parameters either the original governing equations or the trained ANN can be used as predictive model (Fig. 4) or any other kind of reduced model [3]. In this work, both types of predictive models were employed reporting practically the same results. The only difference is established in the time of execution, which is lower when an ANN is used as predictive model, especially if such ANN is built with a low number of hidden neurons.

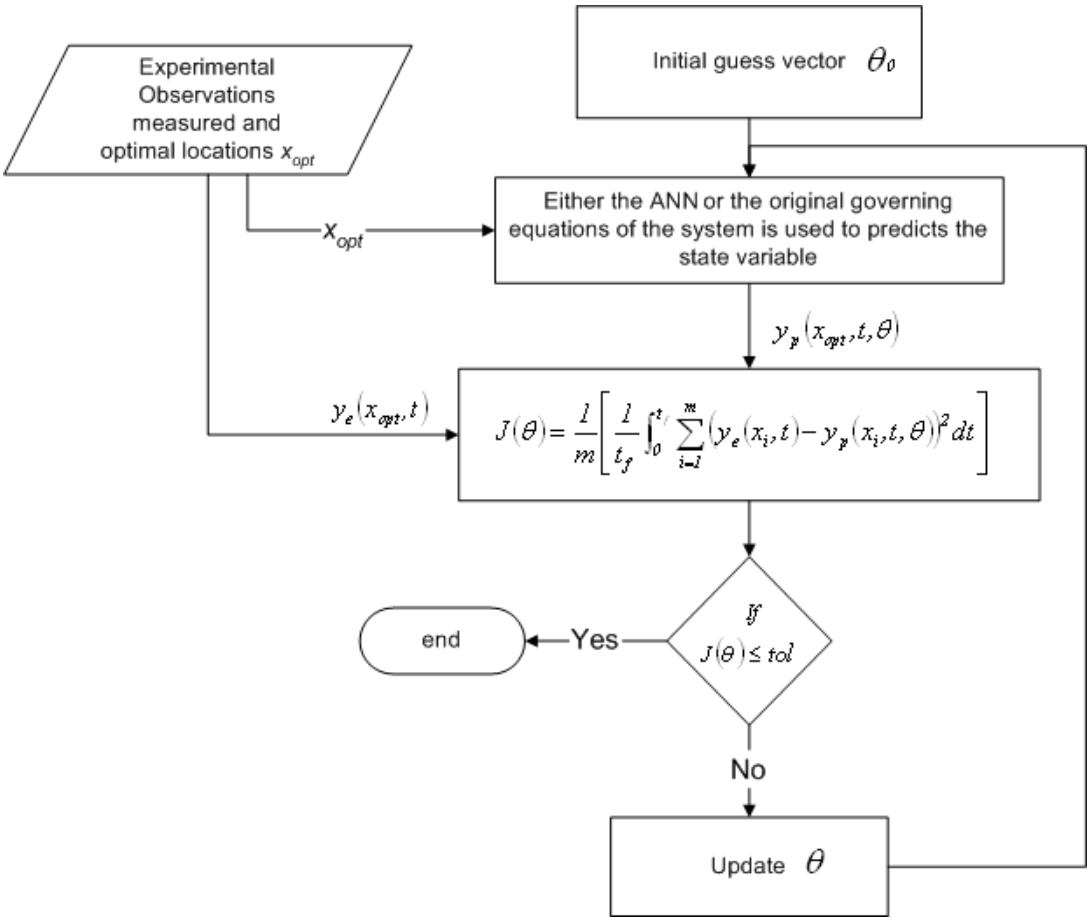


Fig. 4. Parameter estimation procedure

The system was solved using the nominal values of the parameters. Noise, $N(0,\sigma^2)$, was added to the outputs, and the giving results are treated as experimental data. The output from the model together with the experimental data both taken at the optimal sensor location, x_i, are used to minimise (8). The parameters calculated were compared with their corresponding nominal values (Table 2).

Data set averaging is used to reduce the effect of random noise in the measurements, which typically arise from the data acquisition system. In practice, this implies that several data sets of the same signal, in response to identical perturbations, have been taken. Data set averaging may be obtained as,

$$\hat{y}(t_i) = \frac{1}{nt}\sum_{j=1}^{nt} y_j(t_i) \qquad i = 1,...,N_S \tag{23}$$

Where nt is the number of data sets and N_S is the number of observations per set. The effects of the random noise tend to cancel as several sets of data are averaged. However, the noise can only be reduced to a certain degree and not totally eliminated. In this work, only minimal improvements were achieved by averaging together, more than ten sets of data. In addition to averaging, the output can be low pass filtered to remove the residual random noise and the systematic noise. The parameters calculated were compared with their corresponding nominal values.

The analysis of the sensitivity coefficients and of the determinant of the FIM reveals the most appropriate sensor locations for the estimation of the unknown parameters [49]. It is well known that sensitivity analysis quantifies the dependence of the system behaviour on the parameters that affect the process dynamics. Prasad and Vlachos [43] presented results that show that high values of the FIM are correlated with large normalised sensitivity coefficients. These results can be very helpful to reinforce the analysis carried out here.

These measurement locations were considered and their effects on the parameter estimation procedure were verified. The minimum of the determinant of the covariance matrix, equation (9) was used to measure the quality of the parameter estimates considering the spatial positions established by the extrema analysis of the sensitivity coefficients. It is our intention to consider the spatial locations where the sensitivity coefficients show largest absolute values to verify the accuracy of the parameter estimates using experimental data contaminated with noise [2]. At these positions, it could be logical to expect that the confidence intervals for each parameter should be small, according to (11) [4].

Table 2
Parameter estimation results

Example	Measurement Locations	θ_1	θ_2	θ_3	$J(\theta)$	$det[Cov(\theta)]$
5.1	0.00 0.50 1.00[a]	0.1016**±0.050** (1.60%)	1.0375**±0.078** (3.755%)	1.0459**±0.096** (4.59%)	20.999	$1.75e^{-3}$
	Whole spatial domain	0.1008**±0.032** (1.60%)	1.0166**±0.049** (1.66%)	1.0209**±0.065** (2.09%)	18.573	$0.58e^{-3}$
5.2	0.20 0.50 0.80[a]	0.0096**±0.017** (0.80%)	2.1966**±0.401** (9.83%)	5.0889**±0.989** (1.78%)	0.0951	$8.39e^{-3}$
	Whole spatial domain	0.0109**±0.014** (1.09%)	2.0733**±0.279** (3.67%)	5.0497**±0.775** (0.99%)	0.0897	$2.44e^{-3}$

Number in parentheses represents the percentage deviation of the parameter from the nominal value
Numbers in bold represent the uncertainties calculated using (11)
[a] The same locations are obtained when the largest absolute values in the spatial domain of the predominant sensitivity coefficients are evaluated and/or when a D-optimal experiment design is carried out

From Table 2 it is clear that a low deviation on the parameter estimates is obtained measuring at the positions where the sensitivity coefficients reached their largest absolute values. The poorest results are obtained when the measurements are taken at positions far from those largest absolute values (see [2] for more details). Additionally, the sensitivity coefficients calculated using the ANN were employed to compute the covariance matrix and the confidence intervals reporting excellent results as they were compared with the results obtained using the sensitivity coefficients calculated by the conventional ways.

It is well know that problem (8) is challenging for numerous reasons. The existence of a solution is not certain, particularly if the experimental data contains errors or if the model is grossly incorrect. The presence of noise can promotes difficulties during the optimisation process that can cause spurious local minima and discontinuities. In this work, several different optimisation techniques (already existing in the Matlab® library) were used, and it was observed that the method of Nelder – Mead provides the best performance. This can be explained by the fact that using this technique, poor gradient approximations are not a problem, and continuity and differentiability of the objective function are not required. This gives another motivation for using sensor placement strategies based on the extrema evaluation of the sensitivity functions, because not only the uncertainties in parameter estimates were reduced, but the cost associated to such measures as well, which is significant, especially in situations where experimental measurements are very expensive. Information about a physical parameter may be most accurately gained at points in space with a high sensitivity to the parameter.

The parameters with the lowest absolute sensitivity coefficients are the most difficult to estimate, especially if the experimental measurements are taken far from the optimal locations. When at least one parameter shows a very low absolute sensitivity, an optimisation algorithm based on the gradient and/or Hessian matrix generally get stuck in a local minimum, distant from the optimal solution.

Taking into account the largest absolute sensitivities coefficients of the parameter of interest through a careful choice of observation points in a sampling design will lower the variance of the parameter estimate. This is the motivation for examining the behaviour of sensitivities in the system when refinement of parameter estimates is an objective of field sampling. The use

of the ANNs during the parameter identification process improve remarkably the computation cost, which it was reduced in approximately 60% of the time required when the original governing equations are used to predict the outputs. The only drawbacks derived for the use of such non-parametric models consist in the large number of processing units required to fit the training data. This approach needs to be extended to systems showing dynamic behaviours highly nonlinear.

8. Conclusions

The optimal sensor location approach presented here depends strongly on the sensitivity coefficients, but mostly on their spatial and temporal evolution rather than their exact magnitudes. The measurements should be made at the positions where the largest absolute values of the sensitivity coefficients are reached in the spatial domain. In many cases, these locations can be obtained without information of the exact value of the parameters. Well-trained ANNs may yield accurate approximations to an unknown function and its derivatives, which can be used for optimal sensor position purposes. The computational effort required for calculating the optimal weights of the ANN is modest and no information about the values of derivatives of the function at the training points is needed. Even though the FFNNs reported good performances, the RBF neural networks produced the best estimation of the first-order sensitivity coefficients. The number of hidden neurons required to fit the training data need to be incremented as the complexity or nonlinearity of the system increases. The ANN used to approximate the sensitivity coefficients could be used as a predictive model during the parameter estimation process. Not only the uncertainties in the parameter estimates were reduced, but the cost associated to such measures as well, which is significant, especially in situations where experimental measurements are very expensive.

References

[1] T. Ahmed, G. Kenne, F. Lamnabhi-Lagarrige, Nonlinear systems parameter estimation using neural networks: Application to synchronous machines,

Mathematical and Computer Modelling of Dynamical Systems 13 (4) (2007) 365-382.
[2] J. Alaña, C. Theodoropoulos, Optimal location of measurements for parameter estimation of distributed parameter systems, Computer and Chemical Engineering 35(1) (2011) 106.
[3] J. Alaña, Estimation of the temporal coefficients for and empirical approximator. New approach based on the Proper Orthogonal Decomposition modes, Computer and Chemical Engineering 34(8) (2010) 1220-1225.
[4] J. Alaña, Optimal measurement locations for parameter estimation of nonlinear distributed parameter systems, Brazilian Journal of Chemical Engineering Accepted Manuscript (2010).
[5] J. Almeida, E.O. Voit, Neural-networks-based parameter estimation in S-system models of biological networks, Genome Information 14 (2003) 114-123.
[6] A. Alonso, C. Frouzakis, I. Kevrekidis, Optimal sensor placement for state reconstruction of distributed process systems, AIChE Journal 50 (7) (2004a) 1438-1452.
[7] A. Alonso, I. Kevrekidis, J. Banga, C. Frouzakis, Optimal sensor location and reduced order observer design for distributed process systems, Computer and Chemical Engineering 28 (2004b) 27-35.
[8] A. Atkinson, A. Donev, Optimum Experimental Design, Oxford University Press (1992) Oxford.
[9] M. Basseville, A. Benveniste, G. V. Moustakides, A. Rougee, Optimal sensor location for detecting changes in dynamical behaviour, IEEE Transactions on Automatic Control 23 (1987) 1067-1075.
[10] P. Cardaliaguet, G. Eurard, Approximation of a function and its derivative with a neural network, Neural Networks 5 (2) (1992) 207-220.
[11] S. Chen, C.F.N. Cowan, P.M. Grant, Orthogonal least square learning algorithm for radial basis functions networks, IEEE Transactions on Neural Network 2 (2) (1991) 302-309.
[12] D. Chmielewski, P. Tasha, V. Manousiouthakis, On the theory of optimal sensor placement. AIChE Journal 48(5) (2004)1001-1012.
[13] P. Cotae, S. Yalamnchili, C. Chen, A. Ayon, Optimization of sensor locations and sensitivity analysis for engine health monitoring using minimum interference algorithms, EURASIP Journal on Advances in Signal Processing, 9 (2008).
[14] G. Cybenko, Approximation by superposition of a sigmoidal function, Mathematics of Control, Signals, and Systems, Springer-Verlag New York Inc. 2 (1989) 303-314.
[15] T. Damak, J.P. Babary, M.T. Nihtila, Observer design and sensor location in distributed parameter bioreactors, Proc. of DYCORD Maryland 87 (1992) 315-320.
[16] D. Dochain, N. Tali-Mammar, J. Babary, On modeling, monitoring and control of fixed bed bioreactors, Computer and Chemical Engineering 21 (1997) 1255-1266.
[17] A. Emery, A. Nenarokomov, Optimal Experiment Design, Measurement Science and Technology 9 (1998) 864-876.
[18] K. Funahashi, On the approximate realization of continuous mappings by Neural Networks, Neural Networks, 2 (1989) 183-192.
[19] D. Georges, The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems, Proceeding of the 34th Conference on Decision and Control New Orleans (1995) 3319-3324.
[20] A. Grimstat, T. Mannseth, Nonlinearity, Scale, and Sensitivity for a class of parameter estimation problem, SIAM Journal on Scientific Computer 21(6) (1999) 2096.
[21] T. Harris, J.F. Macgregor, J.D. Wright, Optimal sensor location with an application to a packed bed tubular reactor, AIChE J. 26 (6) (1980) 910-916.
[22] R. Hecht-Nielson, Theory of the backpropagation Neural Network, Proceedings of the 1989 International Joint Conference on Neural Networks (1989) 593-605.

[23] M. Hernandez, T. Kirubarajan, Y. Bar-Shalom, Multisensor resource deployment using posterior Cramer-Rao bounds, IEEE Transactions on Aerospace and Electronic Systems 40 (2004) 399-416.
[24] K. Hornik, M. Stinchcomne, H. White, Multilayer feedfoward networks are universal approximators, Neural Networks 2 (5) (1990) 359-366.
[25] K. Hornik, M. Stinchcomne, H. White, Universal approximation of an unknown mapping and its derivatives using multilayer feedfoward networks, Neural Networks 3 (1990) 551-560.
[26] S. Jørgensen, L. Goldschmidt, K. Clement, A sensor location procedure for chemical process, Computer and Chemical Engineering 8 (3-4) (1984) 195-204.
[27] P. Jung-Wook, Ch. Byoung-Kon, S. Kyung-Bin, First Derivatives Estimation of Nonlinear Parameters in Hybrid System, IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences 89 (12) (2006) 3736-3738.
[28] J. Keller, D. Bovin, Selection of input and output variables as a model reduction problem, Automatica 28 (1) (1992) 171-177.
[29] D. Knopman, C.I. Voss, Behavior of sensitivities in the one-dimensional advection-dispersion equation: implications for parameter estimation and sampling design, Water Resources Research 23 (2) (1987) 253-272.
[30] J. Korbicz, M.Z. Zgurovsky, A.N. Novikov, Suboptimal sensors location in the state estimation problem for stochastic nonlinear distributed parameter system, International Journal of Systems Science 19 (9) (1988) 1871-1882.
[31] R. Li, M.A. Henson, M.J. Kurtz, Selection of model parameters for off-line parameter estimation, IEEE Transactions on Control Systems Technology 12 (3) (2004) 402-412.
[32] A. Liaqat, M. Fukuhara, T. Takeda, Optimal estimation of parameters of dynamical systems by neural networks collocation method, Computer Physics Communications 150 (3) (2003) 215-234.
[33] L. Ljung, System Identification–Theory for the user, 2nd Edition, Prentice Hall Upper Saddle River (1999) USA.
[34] T. Lopez, J. Alvares, On the effect of the estimation structure in the functioning of a nonlinear copolymer reactor estimator, Journal of Process Control 14 (1) (2004) 99-109.
[35] J. MacQueen, Some methods for classification and analysis of multivariable observations, Proceeding of 5th Berkeley Symposium on Mathematical Statistics and Probability University of California Press (1967) 281-297.
[36] J. Malebranche, Simultaneous state and parameter estimation and location of sensor for distributed systems, Int. Journal of Syst. Sci. 19 (8) (1988) 1387-1405.
[37] Y. Matsubara, S. Kikuchi, M. Sugimoto, M. Tomita, Parameter estimation for stiff equations of biosystems using radial basis function networks, BMC Bioinformatics 7 (230) (2006) 1-11.
[38] M. Morari, G. Stephanopoulos, Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances, AIChE J. 26 (2) (2004) 247-260.
[39] M. Morari, M.J. O'Dowd, Optimal sensor location in the presence of nonstationary noise, Automatica 16 (5) (1980) 463-480.
[40] H. Nahor, N. Scheerlinck, J. Van Impe, B. Nicolaï, Optimization of the temperature sensor position in a hot wire probe set up for estimation of the thermal properties of foods using optimal experimental design, Journal of Food Engineering 57 (2003) 103-110.
[41] S. Omatu, J.H. Seinfeld, T. Soeda, Optimal sensor location for a linear distributed parameter system, IEEE Transactions on Automatic Control 23 (1978) 665.
[42] N. Point, W.A. Vande, M. Remy, Practical issues in distributed parameter estimation: Gradient computation and optimal experiment design, Control Engineering Practice 4

(11) (1996) 1553-1562.
[43] V. Prasad, D. Vlachos, Combined multiscale model and informatics based design of experiments: Application to the catalytic decomposition of Ammonia on Ruthenium, Industrial and Engineering Chemistry Research (accepted).
[44] Z. Qureshi, T.S. Ng, G.C. Goodwin, Optimum experimental design for identification of distributed parameter systems, International Journal of Control 31 (1) (1980) 21-29.
[45] E. Rafajlowics, Optimal experiment design for identification of linear distributed-parameter systems: Frequency domain approach, IEEE Transactions on Automatic Control 28 (7) (1983) 806-808.
[46] W. Rutzler, Nonlinear and adaptive parameter estimation methods for tubular reactors, Industrial and Engineering Chemistry Research. 26(2) (1987) 325-333.
[47] A. Singh, J. Hahn, Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems, Industrial and Engineering Chemistry Research 44 (15) (2005) 5645-5659.
[48] T. Söderström, P. Stoica, System Identification, Prentice Hall International Hemel Hempstead (1989) UK.
[49] R. Storch, L. Pimentel, H. Orlande, Identification of atmospheric boundary layer parameters by inverse problem, Atmospheric Environment 41(7) (2007) 1417-1425.
[50] D. Uciński, Optimal measurement method for distributed parameter system identification, CRC Press. USA (2005).
[51] D. Uciński, Optimal sensor location for parameter identification of distributed systems, Applied Mathematics and Computer Science 2(1) (1992) 119-134.
[52] F. van den Berg, H.C.J. Hoefsloot, H.F.M. Boelens, A.K. Smilde, Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria, Chemical Engineering Science 55 (4) (2000) 827-837.
[53] C. Venkateswarlu, B. Kumar, Composition estimation of multicomponent reactive batch distillation with optimal sensor configuration, Chemical Engineering Science 61 (2006) 5560-5574.
[54] W. Waldraff, D. Dochain, S. Bourrel, A. Magnus, On the use of observability measures for sensor location in tubular reactor, Journal of Process Control 8 (5) (1998) 497-505.
[55] E. Walter, L. Prozanto, Qualitative and quantitative experiment design for phenomenological models – A survey, Automatica 26 (2) (1990) 195-213.
[56] T. Waterhouse, J. Eccleston, S. Duffull, Optimal design criteria for discrimination and estimation in nonlinear models, Journal of Biopharmaceutical Statistics 19 (2) (2009) 386-402.
[57] A.Wouwer, N. Point, S. Poterman, M. Remy, An approach to the selection of optimal sensor locations in distributed parameter systems, Journal of Process Control 10 (4) (2000) 291-300.
[58] W. Yeh, Y. Yoon, Aquifer parameter identification with optimum dimension in parameterization, Water Resources Research 17 (3) (1981) 664-672.
[59] E. Zamprogna, M. Barolo, D. Seborg, Optimal selection of soft sensors inputs for batch distillation columns using principal component analysis, Journal of Process Control 15 (2005) 39-52.
[60] Z. Zhu, M. Stein, Spatial sampling design for parameter estimation of the covariance function, Journal of Statistical Planning and Inference 134(2) (2005) 583-603.
[61] J. Zhu, Z. Wang, X. Xia, J. Lei, One parameterized model of indirect measurement based on neural network and its sensitivity coefficient computing, Journal of Physics: Conference Series 13 (1) (2005) 202-205.

Optimal Spatial Sampling Scheme for Parameter Estimation of Non Linear Distributed Parameter Systems

Jorge E. Alaña [1,2] and Constantinos Theodoropoulos [1]
[1]School of Chemical Engineering and Analytical Science,
The University of Manchester, Manchester, M60 1QD, U.K.
Email: k.theodoropoulos@manchester.ac.uk
[2] Escuela de Ingeniería Química, Facultad de Ingeniería, Universidad Del Zulia, Maracaibo, Zulia, Venezuela
Email: jalana@fing.luz.edu.ve

Abstract

A new approach for the optimal location of spatial measurements, for efficient estimation of parameters of nonlinear distributed systems, is introduced. This scheme is based on the Proper Orthogonal Decomposition (POD) and it is verified by a numerical example regarding the chemical reaction occurring in a tubular reactor for two possible scenarios; stable and unstable operation conditions. An extrema evaluation of POD modes can be used directly to define optimal sensor locations for parameter estimation of nonlinear distributed parameter system. This is a computational efficient approach that allows defining the optimal measurement placements without the laborious calculation of the sensitivity coefficient Jacobian matrices generally required by the D-optimal experiment design. Moreover, only merely system responses and/or experimental observations are required and used straightforwardly. The underlying combination of model reduction techniques and sensor location problem in this approach becomes even more relevant as the complexity of the system under consideration increases. By this method, not only the uncertainty in parameter estimates can be reduced, but cost is decreased also, which is significant, particularly in circumstances where experimental observations are expensive.

Key words: sensitivity coefficients, measurement points, proper orthogonal decomposition, empirical eigenfunctions, nonlinear distributed parameter system

1. Sensor position problem. Review

The sensor position problem has been analysed from various points of view, but the results reported by most authors are restricted to the selection of stationary sensor locations and the framework of state estimation (Amouroux and Babary, 1988; El Jai, 1991; El Jai and Pritchard, 1988; Kubrusly and Malebranche, 1985). A broadening of state estimation techniques to the parameter estimation problem is not straightforward and somewhat difficult, given that in the latter the actual state usually depends nonlinearly on unidentified parameters (Korbicz and Uciński, 1994) even if the system is linear on unknown parameters. This important distinction makes the parameter estimation more complicated and for that cause the number of corresponding results up to date is fewer (Uciński, 2000).

From the angle of the optimal sensor placement problem for parameter estimation, the existing approaches can be categorised into three major groups;

- Methods leading to state estimation.
- Methods making use of random field theory.
- Methods originating in optimal experiment design.

The first group convert the original problem into an estate-estimation one in such a way as to raise the likelihood of applying powerful techniques of optimal sensor location for estate estimation (Damak et al., 1992; Korbicz, 1991). The main disadvantage of such an approach is that simultaneous state and parameter estimation promotes a strong nonlinearity of the problem. Several endeavours concerning sequential linearisation at the successive state trajectories were performed with suboptimal filtering. However, this approach can be implemented only in simple

applications (Basseville et al, 1987; Korbicz et al., 1988; Malebranche, 1988).

The effectiveness of the second group in the case of distributed parameter systems (DPSs) is relatively limited, since transformations between system descriptions, in the form of partial differential equations (PDEs), and appropriate random field characteristics is not straightforward. In addition, such conversions rely on statistics up to a given order (for example, mean, covariance, skewness, kurtosis, etc.) (Coate et al., 2008; Georges, 1995; Balsa-Canto et al., 2008).

The classical theory of optimal experimental design (Atkinson and Donev, 1992; Balsa-Canto et al., 2008; Fedorov and Hackl, 1997; Pázman, 1986; Pukelsheim and Rieder, 1992; Rafajlowicz, 1996; Walter and Prozanto, 1997) and its extensions to dynamic systems (Titterington, 1980) comprises a foundation of the third category of methods. Here the problem is cast as an optimisation one. The performance index is defined in the form of some scalar metric operating on the Fisher Index Matrix (FIM), whose inverse, based on the Cramer-Rao inequality, plays the role of an estimate of the parameter dispersion matrix. This guide to considerable simplification, since even if the exact dispersion matrix is hard to obtain, the inverse of the FIM can be calculated without difficulty (Uciński and Demetriou, 2004; Bernaerts et al., 2005 Chmielewski et al., 2002; Prasad and Vlachos, 2008; Peng, 2005; Waterhouse et al., 2009).

In the case of DPSs the first formulation in this character was intended by Qureshi et al. (1980) whose scheme based on the maximisation of the determinant of the FIM was employed to allocate sensors and boundary perturbations in the dynamic of the system. This approach was implemented for hyperbolic linear systems with identified eigenvalues and eigenfunctions, and a

further general outline in terms of Green's functions was considered later for DPSs (Rafajlowicz, 1981, 1983; Uciński, 2003; Uciński and Patan, 2007; Vande et al., 1996; Wouwer et al., 200).

A fundamental problem towards identification, state estimation, and control of process described by DPSs is the optimal location of the measurements. This problem consists of the arrangement of a limited number of sensor and control devices along the spatial domain in an optimal fashion (Van den Berg et al., 2000; Mossberg, 2004; Papadimitriou, 2004; Rensfelt et al., 2008; Sigh and Hahn, 2005, 2006; Walter and Prozanto, 1990). Nahor et al. (2003) minimised the ratio of the largest to the smallest eigenvalue of the FIM to compute optimal temperature sensor positions for food processes.

Christofides and Antoniades (2000, 2001, 2002) presented new approaches to calculate the optimal actuator/sensor locations of uncertain transport-reaction systems under control. They established that the solution to this problem is near-optimal for the closed-loop infinite-dimensional processes analysed. Similar and other kind of techniques have been reported by Van den Berg et al. (2000), Fahroo and Demetriou (2000) and Harris et al. (1980).

More recently, a number of computational approaches to address this problem have been introduced (Alonso et al., 2004b; Cohen et al., 2006; Löhner and Camelli, 2005; Punithakumar et al., 2006; Uciński and Bogacka, 2002, 2005; Zumoffen and Basualdo, 2009). The central idea is to define a design criterion to be minimised as a scalar of the FIM associated to the estimated parameter. Followed by techniques of optimum experimental plan for nonlinear systems setting the sensor location problem at hand, or alternatively, using standard nonlinear programming procedures. Olanrewaju and Al-Arfaj (2006) introduced, for a reactive distillation process, a state

estimation scheme based on Kalman filters without widely addressing the sensor selection problem. Similar applications were developed by Sumana and Venkateswarlu (2009), Venkateswarlu and Kumar (2006), Wahl et al. (2002), among others. For the best of our knowledge, these sort of techniques have not been broadly applied to system described by PDEs, regardless of their resolute advantages.

In this paper we make use of the extrema values of the system's POD-modes to define the optimal spatial locations (and number) of experimental measurements for parameter estimation purposes. This technique has been proven to be very efficient and practical in the case of systems described by DPSs but only a couple of references can be found in the literature (Yildirim et al., 2009; Alaña and Theodoropoulos, 2010). The rest of the paper is organised as follows: section 2 presents two illustrative case studies from the literature used for benchmarking purposes. In section 3, the methodology followed to define the optimal sensor locations. Section 4 is dedicated to parameter estimation aspects including the covariance matrix calculations and section 5 presents the results derived from the approach considered in this work.

2. The system governing and Sensitivity equations

The system considered is a homogeneous tubular reactor without catalyst packing. An irreversible first–order chemical reaction A→B takes place, and it is described by the concentration rate $C_A(t,z)$ and the temperature $T(t,z)$. This reaction is exothermic and a cooling jacket is used to remove heat from the reactor. A fraction of the products can be recycled, r, and mixed with the reactants at the inlet of the reactor ($z=0$) (Bendersky and Christofides 2000; Li and Chistofides, 2008). This example has been used to address the sensor location using the

techniques available and reported in the literature (Alonso et al., 2004a; Alvares et al., 1981; Dochain et al., 1997; Harris et al., 1980; Waldraff et al., 1998; Zamprohna et al., 2005).

Derived from mass and energy balances, and under the standard assumptions of constant density (ρ), heat capacity of the reacting fluid (cp), and constant axial fluid velocity (v), the dynamic behaviour of this system is described by a set of PDEs defined on a spatial domain $z \in (0,1)$, in dimensionless form,

$$\frac{\partial y_1(t,z)}{\partial t} = -\frac{\partial y_1(t,z)}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_1(t,z)}{\partial z^2} + R_{i3} \tag{1a}$$

$$\frac{\partial y_2(t,z)}{\partial t} = -\frac{\partial y_2(t,z)}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_2(t,z)}{\partial z^2} + R_{i1} + R_{i2} \tag{1b}$$

Here $y_1(t,z)$ and $y_2(t,z)$ represent concentration and temperature, respectively, and u_C is the cooling water temperature. Pe_C, Pe_T are the Peclet numbers of mass and energy flows, in that order, B_C is the Damkohler number, γ is the activation energy, B_T and β_T are the parameters related to heat of reaction and heat transfer, respectively, and $t \in (t_0, t_f)$, where t_0 and t_f are initial and final times, correspondingly. The terms R_{ij} correspond to the examples; $i=1$ stable case with a recycle $r=0.00$, and $i=2$ unstable case with $r=0.50$, j stands for the following expression,

$$R_{11} = B_T B_C (1 - y_1(t,z)) exp\left(\frac{y_2(t,z)}{1 + \frac{y_2(t,z)}{\gamma}}\right), \qquad R_{21} = B_T B_C (1 + y_1(t,z)) exp\left(\frac{\gamma y_2(t,z)}{1 + y_2(t,z)}\right) \tag{2a}$$

$$R_{12} = R_{22} = -\beta_T (u_C - y_2(t,z)) \tag{2b}$$

$$R_{13} = -B_C(1 - y_1(t,z))exp\left(\frac{y_2(t,z)}{1+\frac{y_2(t,z)}{\gamma}}\right), \qquad R_{23} = B_C(1 + y_1(t,z))exp\left(\frac{\gamma y_2(t,z)}{1+y_2(t,z)}\right) \tag{2c}$$

The boundary and initial conditions are,

a.- For the stable case; $t \in (0,5)$

$$z = 0 \qquad \frac{\partial y_1(t,0)}{\partial z} = Pe_C y_1(t,0) \qquad \frac{\partial y_2(t,0)}{\partial z} = Pe_T y_2(t,0) \tag{3a}$$

$$z = 1 \qquad \frac{\partial y_1(t,1)}{\partial z} = 0 \qquad \frac{\partial y_2(t,1)}{\partial z} = 0 \tag{3b}$$

$$t = 0 \qquad y_{10} = y_{20} = u_C = 0 \tag{3c}$$

b.- For the unstable case; assuming negligible reaction in the recycle loop and instantaneous mixing of fresh feed and recycle feed at the reactor inlet, the boundary and initial conditions in the interval $t \in (0,25)$ take the form,

$$z = 0 \qquad \frac{\partial y_1(t,0)}{\partial z} = Pe_C\left[y_1(t,0) - (1-r)y_{10} - ry_1(t,1)\right] \tag{4a}$$

$$\frac{\partial y_2(t,0)}{\partial z} = Pe_T\left[y_2(t,0) - (1-r)y_{20} - ry_2(t,1)\right]$$

$$z = 1 \qquad \frac{\partial y_1(t,1)}{\partial z} = 0 \qquad \frac{\partial y_2(t,1)}{\partial z} = 0 \tag{4b}$$

$$t = 0 \qquad y_{10} = y_{20} = u_C = 0 \tag{4c}$$

The nominal values of the parameter for the systems studied in this work are shown in Table 1.

Table 1
Nominal values of the parameters

Stable case, i=1	Unstable case, i=2
$Pe_C = 1.0 \quad B_C = 1.0 \quad \gamma = 20.0$ $Pe_T = 1.0 \quad B_T = 2.0 \quad \beta_T = 1.0$	$Pe_C = 7.0 \quad B_C = 0.1 \quad \gamma = 10.0$ $Pe_T = 7.0 \quad B_T = 2.5 \quad \beta_T = 2.0$

In most cases study it can be assumed that some, albeit rough, a priori value of the parameters vector θ is available, e.g. determined from preliminary experiments. The aim in this work is to select a state sampling scheme to maximise the expected accuracy of the parameter estimates to be obtained from data generated in new experiments. These parameters should be determined with the lower uncertainties.

3. Methodology

The system governing equations (1) – (4) were solved using the Matlab® PDE solver "pdepe" which solves initial-boundary value problems for systems of parabolic and elliptic PDEs in one spatial dimension. The solver converts the PDEs to ordinary differential equations using a second-order accurate spatial discretisation based on a set of nodes specified by the user. Time integration is then performed with a multistep variable-order method based on the numerical differentiation formulae.

The spatial domain was divided into 21 equidistant intervals, and the PDEs were solved simultaneously with the equations describing the transient behaviour of the sensitivity coefficients considering *n* time reporting intervals. Hence, the spatio–temporal distributions of the response and the sensitivity coefficients are obtained.

The method of snapshots (Sirovich, 1987) was performed and the empirical eigenfunctions corresponding to concentration and temperature; $\phi_i(z)$ and $\theta_i(z)$, respectively, for each system were calculated, their extrema values satisfying $\partial\phi_i(z)/\partial z=0$ and $\partial\theta_i(z)/\partial z=0$ were evaluated and the resultant locations were considered for optimal sensor placement purposes (Alaña and Theodoropoulos, 2010).

4. Parameter estimation

Model calibration consists of finding a set of parameter values that produces the best model output which fit the observed data. This calibration is usually done by the minimisation of the objective function (5).

Once the sensor locations have been determined, the parameter estimation problem can be viewed as matching the model to the real system through the minimisation of an error criterion over a set of admissible parameters. This can be defined as,

$$\Im(\theta) = \frac{1}{m}\left[\frac{1}{t_f}\int_0^{t_f}\sum_{i=1}^{m}\left(y_e(t,z_i) - y_p(t,z_i,\theta)\right)^2 dt\right] \tag{5}$$

Where, $y_p(t,z_i,\theta)$ is defined as the predicted response of the model at location z_i, and $y_e(t,z_i)$ as the experimental response at the same location, m is the number of optimal measurement locations determined by the POD extrema analysis studied in this work. Unfortunately, elements in the vector $y_p(t,z_i,\theta)$ are not linear functions of the parameters θ, and multiple solutions of (5) are possible. This vector can be obtained either by the use of the original governing equations of the system or by a reduced model derived from the application of the Galerkin projection in conjunction with the empirical eigenfunctions calculated by the snapshots procedure.

It is well known that problem (5) is challenging for numerous reasons. The existence of a solution is not certain, particularly if the observed data contains errors or if the model is grossly incorrect. It is unusual that any parameter set can accurately match the experimental data used, especially when these data are contaminated with noise. The presence of noise can promote difficulties during the optimisation process causing spurious local minima and discontinuities. The gradient-based methods are the most affected of all. In this work, several different optimisation techniques (already existing in the Matlab® library) were used. The method of Nelder–Mead produced the best results for the scenarios studied. This can be explained by the fact that using this technique, poor gradient approximations are not a problem, and continuity and differentiability of the objective function are not required.

4.1. Covariance matrix of parameter estimates

When assessing the quality of an identified dynamic model, the covariance matrix of the estimated parameters gives an important measure (Ljung, 1999; Söderström and Stoica, 1989). In a maximum likelihood context, the inverse of the expected value of the negative of the Hessian provides the Cramer-Rao lower bound on the variance-covariance matrix of the parameter estimates. Large variances imply imprecise parameter estimates. The arrival of new data, or data measurement, can lead to substantial change in parameter estimates if the variance-covariance matrix is ill-conditioned. So, the goal of the estimation procedure is to determine unknown parameters in such a way that the difference between the sample covariance matrix and the implied covariance matrix is minimised in a certain sense (Sorenson, 1980; Zhun and Stein, 2005).

Based on the system studied here, the approximation of the parameters covariance matrix (Yen and Yoon, 1981) can be approximated by,

$$Cov\left(\hat{\theta}_{nv}\right) \approx \left(\frac{1}{\sigma_{y_1}^2} J_{y_1}^T J_{y_1} + \frac{1}{\sigma_{y_2}^2} J_{y_2}^T J_{y_2} \right)^{-1} \tag{6}$$

Where, J_{y_1} and J_{y_2} represent the Jacobian matrix of estimated variables; y_1 and y_2, respectively, with respect to changes in the parameters. *nv* is the parameter dimension, and σ_k^2 is the variance of the variable *k*. In this equation it is easy to view the influence of field data quality and quantity in parameter uncertainty. Since these Jacobian matrices are evaluated at those locations for which observations are available, any experimental design should aim at sampling at those locations where the variables are most sensitive to the estimated parameters. Such a design is said to provide the maximum amount of information about the unknown parameters (Knopman and Voss, 1987).

It is intuitively obvious that the experimental design objective should be intended at minimising the norm of (6); that is, to make matrices $\left(J_{y_1}^T J_{y_1}\right)^{-1}$ and $\left(J_{y_2}^T J_{y_2}\right)^{-1}$ as small as possible. The variance terms $\sigma_{y_1}^2$ and $\sigma_{y_2}^2$ are constants, and can be dropped from the formulation. To measure the accuracy of the estimates we like to summarise the information about the variability in the covariance matrix into a single number. Here we used the determinant of (6) as the function that transforms a matrix into a scalar. This is quite informative as in fact it relates to the volume of the multidimensional simplex defined by the column/row vectors of the matrix (Hernandez et al., 2004; Rutzler, 1987).

Experimental design for parameter estimation deals with the problem of defining experimental conditions that increase the reliability of a simulation model. This can be formulated using a measure of the covariance matrix of the parameter estimates. Among these, the most widely used design criteria are,

- A-optimality: when the trace of (6) is minimised.
- D-optimality: when the determinant of (6) is minimised.
- E-optimality: when the maximal eigenvalue of (6) is minimised.

Using different norms leads to slightly different conclusions regarding the optimal design. The D-optimality criterion minimises the volume of the hyper-ellipsoid in the parameter space, which no consideration of the relationship between the ellipsoid's axes lengths, which are in turn proportional to the square root of the covariance matrix eigenvalues.

In general, it can be shown that under some assumptions of regularity and for a sufficiently large sample size N the vector $\left(\hat{\theta}_{(N)} - \overset{*}{\theta}\right)$ (with $\overset{*}{\theta}$ denoting the 'true' but unknown value of the parameters and $\hat{\theta}_{(N)}$ the least square parameter estimates) has approximately a normal distribution with zero mean and covariance matrix,

$$\frac{\sigma^2}{N} M^{-1}\left(\xi, \overset{*}{\theta}\right) \tag{7}$$

Where σ^2 denotes the standard deviation of the errors in the model, ξ represent a design experiment (including the measurement locations) and the $M\left(\xi, \overset{*}{\theta}\right)$ is the FIM, defined by,

$$M\left(\xi,\overset{*}{\theta}\right)=\left(\sum_{k=1}^{n}\frac{\partial y(t,z_k,\theta)}{\partial\theta_i}\frac{\partial y(t,z_k,\theta)}{\partial\theta_j}\right)_{i,j=0}^{N} \tag{8}$$

In principle the covariance matrix is a measure for the precision of the least square estimator for the unknown parameter $\overset{*}{\theta}$ and a 'smaller' matrix yields more precise estimates. For example, the *ith* diagonal element of (7) will be denoted by $\frac{\sigma^2}{N}M^{-1}\left(\xi,\overset{*}{\theta}\right)_{ii}$ and is an approximation of the variance or mean squared error for the *ith* component $\hat{\theta}_{i(N)}$ of the least squared estimator $\hat{\theta}_{(N)}$. An approximate confidence interval for the *ith* component θ_i of the vector θ is given by,

$$\left[\hat{\theta}_{i(N)}-\frac{\hat{\sigma}u_{1-\alpha/2}}{\sqrt{N}}\sqrt{M^{-1}\left(\xi,\overset{*}{\theta}\right)_{ii}},\hat{\theta}_{i(N)}+\frac{\hat{\sigma}u_{1-\alpha/2}}{\sqrt{N}}\sqrt{M^{-1}\left(\xi,\overset{*}{\theta}\right)_{ii}}\right] \tag{9}$$

Where $u_{1-\alpha/2}$ denotes the 1-α/2 quantile of the standard normal distribution and $\hat{\sigma}^2$ is an estimate of the unknown variance of the error. For most cases it was shown that for moderate sample sizes *N* the sampling variances of the parameter estimates are well approximate.

The precision of the estimates can always be decreased by increasing the sample size *N*, which yields a 'smaller' covariance matrix and smaller variances of the least square estimates. However, in practice the sample size is usually fixed, due to cost considerations of each additional experiment. To improve the quality of the estimates or, from a different point of view, to reduce the number of experiment measurements needed to obtain the estimates with a given accuracy, we note that the variances of the estimates $\hat{\theta}_{i(N)}$ and the covariance matrix of the vector $\hat{\theta}_{(N)}$ also depend on the given design, ξ, which determines the relative proportion of total observations to be taken at the experimental locations.

It is advisable to check that the confidence interval magnitude agrees with the sensitivity analysis, which is the parameter which has the largest sensitivity coefficient should have the smallest confidence interval. Because of the consideration of the measurement error only, if the FIM is well-defined, the confidence intervals may result in very small values. Grimstadt and Mannseth (1998) indicate that the use of such approximation of the confidence intervals was almost always justified even for the highly nonlinear model they analysed.

4.2 Experimental sampling design

Data set to be used in the enhancement of parameter estimates should be preceded by initial assumptions concerning the physical model, taking account of geometry, boundary conditions, and parameters. This is the fundamental idea of sampling design. Design of an optimal sampling scheme consequently requires the ability to mathematically model the physical system either analytically or numerically.

A sampling design is usually described as,

$$\overset{*}{\xi} = \left\{ \left(z_i, t_i \right); z_i \in Z, t_i \in T, i = 1,...,m \right\} \tag{10}$$

The set $\overset{*}{\xi}$ identifies the locations in space and time from which *m* observations of the system are taken and Z and T define the spatial and temporal domains, respectively, of design space. Therefore a collection of n measurement points is delineated for the system transient response as pairs $\left(z_i, t_i \right)$ for $i = 1,...,m$. In addition to the z_i and t_i coordinates, the number of observations in an optimal design is also unknown.

Sampling design, like parameter estimation, is an optimisation problem. Here the problem is how to select the time, location, and number of observations to maximise the value of the data collected while minimising the sampling effort. The meaning of "value" depends on the use to which the data is put. For applications to parameter estimation, a practical aim would be to identify a sampling design that is most likely to yield parameter estimates with low variances and covariances. In theory, such estimates would imply a model that more accurately represents system behaviour.

An optimal design is based on preliminary assumptions about the physical model including the true parameters, which paradoxically, are the target of the estimation process. This difficulty may be circumvented, first, by viewing sampling design as a sequential process in which model assumptions are updated following each round of sampling and, second, by robust designs that perform well over a range of possible true parameter values and boundary conditions.

One possible objective of sampling design is a scalar performance criterion related to the covariance matrix, theoretically defined as,

$$Cov\left(\overset{*}{\theta}\right) \approx Cov(\theta) = \sigma^2 \left(J^T J\right)^{-1} \tag{11}$$

Where σ^2 is the common variance of the random error terms and J is the Jacobian matrix. The covariance matrix $Cov\left(\overset{*}{\theta}\right)$ is exactly equals to $Cov(\theta)$ when the regression model is linear respect to the parameters. For a nonlinear model, $Cov\left(\overset{*}{\theta}\right) = Cov(\theta)$ only in a neighbourhood near θ in parameter space when the predictive approximation of the system, $y_p(z,t,\theta)$, is a smooth function.

An example of scalar performance criterion is the maximisation of the determinant of $\left(J^T J\right)$ known as D-optimality. Assuming the theoretical expression for the covariance matrix (11) is a valid approximation in the neighbourhood of θ, then it can be seen that the elements of the J matrix determine the variance and covariance of parameters in addition to an estimate of variance σ^2 in the random errors in the experimental observations. Recall that the inverse of a matrix is defined as,

$$\left(J^T J\right)^{-1} = \frac{1}{\left|J^T J\right|} adj\left(J^T J\right) \tag{12}$$

Where adj is the adjoint of $\left(J^T J\right)$.

The solution to the D-optimality problem is mathematically equivalent to the solution of the problem of minimising the determinant of the covariance matrix $\sigma^2\left(J^T J\right)^{-1}$, assuming the validity of (11) near the estimated parameters set θ. In parameter space, the determinant of the covariance matrix is proportional to the volume of the confidence ellipsoid surrounding the parameter estimates. Thus the determinant of $\left(J^T J\right)$ may be used as a surrogate measure of overall uncertainty in parameter estimates. Although algorithms for identifying D-optimal designs have been developed, none guarantees a global optimum. An enumerative exploration through design space, although inefficient, is nonetheless a satisfactory means of finding local optima for the purpose of the present work.

The D-optimality criterion depends on an adequate approximation of the covariance matrix for the parameter set θ, either by the theoretical expression (11) or by estimation from a Monte

Carlo experiment. Each trial in a Monte Carlo experiment is a different realisation of observations from which parameters are then estimated. As many as 1000 trials may be required to arrive at a stable estimate of variance for a single design. As a consequence it would be attractive to establish the validity of the theoretical covariance matrix over the range of variance in random error and parameter values for which the regression model would be applied. To test the validity of the least squares assumption of linearity in the neighbourhood of the solution θ^* for a particular regression model, the theoretical approximation of the covariance matrix should be compared with a sample covariance matrix generated from a large number of trials. If the theoretical approximation of $Cov(\theta)$ is sufficiently close to the sampling covariance matrix, then the theoretical approximation may be considered adequate for the particular regression model. Nevertheless, this is not known a priori. If the theoretical and sampling covariance matrices are calculated for a range of variance in random error and a range of parameter values, then the validity of the theoretical approximation is proven for the complete range.

To illustrate the basic relationship between the covariance matrix and sampling design, it is useful to examine the case of a simple, two-parameter regression model. For a given design the appropriate entries to the *J* matrix in (11) are as follow;

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial \theta_1} & \frac{\partial y_1}{\partial \theta_2} \\ \frac{\partial y_2}{\partial \theta_1} & \frac{\partial y_2}{\partial \theta_2} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \frac{\partial y_m}{\partial \theta_1} & \frac{\partial y_m}{\partial \theta_2} \end{bmatrix} \tag{13}$$

The square of the *J* matrix $\left(J^T J\right)$ is therefore,

$$\left(J^T J\right)=\begin{bmatrix}\kappa_{11} & \kappa_{12}\\ \kappa_{21} & \kappa_{22}\end{bmatrix} \tag{14}$$

Where,

$$\kappa_{11}=\sum_{j=1}^{m}\left(\frac{\partial y_j}{\partial \theta_1}\right)^2 \tag{15a}$$

$$\kappa_{12}=\kappa_{21}=\sum_{j=1}^{m}\left(\frac{\partial y_j}{\partial \theta_1}\right)\left(\frac{\partial y_j}{\partial \theta_2}\right) \tag{15b}$$

$$\kappa_{22}=\sum_{j=1}^{m}\left(\frac{\partial y_j}{\partial \theta_2}\right)^2 \tag{15c}$$

The determinant of $\left(J^T J\right)$ is thus given by,

$$\left|J^T J\right|=\kappa_{11}\kappa_{22}-\left(\kappa_{12}\right)^2 \tag{16}$$

Assuming independence in the rows and columns of J, it can be shown that $\left(J^T J\right)$ is positive definite and hence $det\left(J^T J\right)>0$. The determinant of $\left(J^T J\right)$ will increase as the sum of squared sensitivity coefficients for each parameter increases, and decrease as the correlation between the sensitivity of the two parameters increases.

The adjoint of the $\left(J^T J\right)$ matrix for the two-parameter problem is,

$$adj\left(J^T J\right)=\begin{bmatrix}\kappa_{22} & -\kappa_{12}\\ -\kappa_{21} & \kappa_{11}\end{bmatrix} \tag{17}$$

Both $\left(J^T J\right)$ and its adjoint are symmetric.

Using parameter θ_1 as an example, its relative variance can be expressed as,

$$\frac{Cov(\theta_1)}{\sigma^2} = \frac{\kappa_{22}}{\kappa_{11}\kappa_{22} - (\kappa_{12})^2} \tag{18}$$

Dividing through by κ_{11} and rearranging,

$$\frac{Cov(\theta_1)}{\sigma^2} = \frac{1}{\kappa_{11}\left(1 - \zeta_{12}{}^2\right)} \tag{19}$$

Where ζ_{12} is a simple correlation coefficient for the sensitivities defined by,

$$\zeta_{12} = \frac{\kappa_{12}}{(\kappa_{11}\kappa_{22})^{1/2}} \tag{20}$$

This coefficient is indicative of the linear relationship between the sensitivities for θ_1 and θ_2. The extent to which maximising κ_{11} will minimise $\frac{Cov(\theta_1)}{\sigma^2}$ depends on the collinearity among the sensitivities. As collinearity increases, the influence of a large κ_{11} on $\frac{Cov(\theta_1)}{\sigma^2}$ diminishes, although increasing κ_{11} will continue to reduce variance.

The connotations of the previous argument also apply to θ_2. For problems with more than two parameters, the analysis may be easily generalised. Maximising the sum of absolute sensitivities to the parameter of interest through a careful choice of observation points in a sampling design will lower the variance of the parameter estimate. This is the incentive for examining the behaviour of sensitivities in the system studied when improvement of parameter estimates is a purpose of field sampling.

To maximise κ_{11}, consider the following optimisation problem,

$$max \quad \kappa_{11} = max \sum_{j=1}^{m} \left(\frac{\partial y_j}{\partial \theta_1} \right)^2 \tag{21}$$

Or alternatively,

$$max \quad \kappa_{11} = max \left\{ \left(\frac{\partial y_1}{\partial \theta_1} \right)^2 + \left(\frac{\partial y_2}{\partial \theta_1} \right)^2 + ... + \left(\frac{\partial y_m}{\partial \theta_1} \right)^2 \right\} \tag{22}$$

Or,

$$max \quad \kappa_{11} = max \left\{ S_{11}{}^2 + S_{12}{}^2 + ... + S_{1m}{}^2 \right\} \qquad \text{with } S_{1j} = \frac{\partial y_j}{\partial \theta_1} \tag{23}$$

Taking derivative with respect to variable z,

$$\frac{\partial \kappa_{11}}{\partial z} = 2S_{11} \frac{\partial S_{11}}{\partial z} + 2S_{12} \frac{\partial S_{12}}{\partial z} + ... + 2S_{1m} \frac{\partial S_{1m}}{\partial z} \tag{24a}$$

$$\frac{\partial \kappa_{11}}{\partial z} = 2 \sum_{j=1}^{m} S_{1j} \frac{\partial S_{1j}}{\partial z} \tag{24b}$$

Then to maximise κ_{11} it is necessary that,

$$\frac{\partial \kappa_{11}}{\partial z} = 0 \tag{25}$$

Which it can be achieved when,

$$1) \frac{\partial S_{j1}}{\partial z} = 0 \qquad j = 1,...,m \tag{26a}$$

$$2) S_{j1} = 0 \qquad j = 1,...,m \tag{26b}$$

Then if we select carefully the locations j for which $\frac{\partial S_{j1}}{\partial z} = 0$ we could warrantee that the $\frac{Cov(\theta_i)}{\sigma^2}$ is minimised, which is the same to consider the locations of the extrema values corresponding to each POD-mode capturing most of the kinetic energy of the system.

Considering the equivalence between the sensitivity coefficients of a general system and the same given by an approximator based on the POD-modes (Alaña and Theodoropoulos, 2010),

$$\frac{\partial \kappa_{11}}{\partial z} = \sum_{j=1}^{m}\sum_{k=1}^{M}\sum_{w=1}^{M} \frac{\partial}{\partial \theta_1}\left(\alpha_k(t)\phi_k(z_j)\right)\frac{\partial}{\partial \theta_1}\left(\alpha_k(t)\frac{\partial \phi_k(z_j)}{\partial z}\right) \tag{27}$$

Here it can be seen that (25) is accomplished when,

$$\frac{\partial \phi_k(z_j)}{\partial z} = 0 \quad \textit{and / or} \quad \phi_k(z_j) = 0 \quad \textit{for} \quad k = 1,...,M \tag{28}$$

Then it can be observed that the variance of the parameter estimates can be lowered by means of an optimal selection of measurement points, z_j for $j = 1,...,m$ where the sum of the absolute POD-modes is maximised, according to (22) and (27).

5. Numerical Results

The spatial variability of sensitivities and/or POD-modes has a significant impact on parameter estimation and sampling design for studies of DPSs. Information about a physical parameter will be most accurately gained at points in space with a high sensitivity to the parameter.

The set of PDEs described in Section 2 above was solved using nominal values for the parameters (Table 1) and random noise with zero mean, $N(0,\sigma^2)$, was added to the outputs, the variance used here, $\sigma^2=0.0005$. For each case 10 replications of the digital experiment were collected.

Data set averaging is used to reduce the effect of random noise in the measurements which typically arise from the data acquisition system. In practice, this implies that several data sets of

the same signal, in response to identical perturbations, have been taken. Data set averaging may be obtained as,

$$\hat{y}(t_i, z) = \frac{1}{n}\sum_{j=1}^{n} y_j(t_i, z) \qquad i = 1,...,N_s \tag{29}$$

Where n is the number of data sets and N_s is the number of observations per set. The effects of the random noise tend to cancel as several sets of data are averaged. However, the noise can only be reduced to a certain degree and not totally eliminated. In this work only minimal improvements were achieved by averaging together more than ten sets of data. In addition to averaging, the output can be low pass filtered to remove the residual random noise and the systematic noise.

An ensemble of $N=50$ snapshots, selected randomly from the whole data, was constructed to calculate the empirical eigenfunctions; $\phi_i(z)$ and $\varphi_i(z)$ for each field. A reduced model was derived for each example applying Galerkin projection to the original governing equations and using as basis the empirical eigenfunctions obtained from the previous step. For each example the relative and total kinetic energy captured by the POD-modes are shown in Tables 2 and 3.

The average responses obtained applying (29) were treated as experimental data and used during the minimisation of (5). The parameters calculated were compared with their corresponding nominal values and some extra analyses were carried out (Alaña and Theodoropoulos, 2010).

The parameter estimation was carried out considering the governing equations as predictive model and the equivalent reduced model obtained for each example studied and the results obtained were compared. The initial guess used, in most of cases, corresponds to the 50% of the

nominal values given in Table 1, otherwise it is mentioned in the corresponding results.

Table 2
Relative and total system's energy captured by the POD-modes. Tubular reactor stable operation conditions

		λ_1	λ_2	**Total kinetic energy (%)**
Tubular Reactor Stable Operation	Field $y_1(t,z)$	0.41082 (0.99995)	0.00002 (5.12e-5)	99.999
	Field $y_2(t,z)$	0.41786 (0.99980)	0.00009 (2.04e-4)	99.999

Number in parentheses represent the relative energy captured by the *ith* POD-mode, defined by,

$$E_i = \lambda_i \Big/ \sum_{k=1}^{m} \lambda_k$$

The accumulative sum of the relative energies approaches 1.00 as the number of modes in the reconstruction increases to *m*.

Table 3
Relative and total system's energy captured by the POD-modes. Tubular reactor unstable operation conditions

		λ_1	λ_2	λ_3	λ_4	**Total kinetic energy (%)**
Tubular Reactor Unstable Operation	Field $y_1(t,z)$	0.61849 (0.97479)	0.01449 (0.02283)	0.001460 (0.00229)	0.000050 (7.30e-5)	99.999
	Field $y_2(t,z)$	1.17224 (0.97860)	0.01994 (0.01665)	0.005578 (0.00466)	0.000105 (8.78e-5)	99.999

Number in parentheses represent the relative energy captured by the *ith* POD-mode, defined by,

$$E_i = \lambda_i \Big/ \sum_{k=1}^{m} \lambda_k$$

The accumulative sum of the relative energies approaches 1.00 as the number of modes in the reconstruction increases to *m*.

The field reconstruction for the stable scenario can be achieved by considering the first two empirical eigenfunctions. The unstable case requires the first four POD-modes to reproduce satisfactorily the concentration and temperature fields. Figures 1 and 2 show these empirical eigenfunctions in the spatial domain for each example. The locations of the extrema values can be extracted by mere observations of these figures and are depicted in corresponding tables. This information is particularly important if the POD-based method is used to directly reduce the system at hand leading to efficient linking the sensor placement problem with the on-line

optimisation and control strategies for DPSs. The spatial positions of the POD-modes extrema values and the parameter estimation scenarios performed in this work are show in Table 4.

Table 4
Extrema spatial locations of the empirical eigenfunctions

Scenario	Tubular reactor stable operation conditions		Tubular reactor unstable operation conditions	
	Concentration $y_1(t,z)$	**Temperature** $y_2(t,z)$	**Concentration** $y_1(t,z)$	**Temperature** $y_2(t,z)$
A	Whole spatial domain	Whole spatial domain	Whole spatial domain	Whole spatial domain
B[a]	0.00 0.05 1.00	0.00 0.05 1.00	0.00 0.05 0.15 0.20 0.40 0.50 0.55 1.00	0.00 0.05 0.20 0.45 0.50 1.00
C[c]	0.00 1.00	0.00 1.00	N/a	N/a
D[b]	0.00 0.05 1.00	0.00 0.05 1.00	0.00 0.05 0.15 0.20 0.40 0.50 0.55 1.00	0.00 0.05 0.20 0.45 0.50 1.00

[a] The experimental measurements positions are based in the allocation of the extrema values of the POD-modes in the spatial domain
[b] Same case as B but the reduced model was used as predictive model
[c] This case was carried out to quantify the effect of considering or not the sensor location at z=0.05 stable operation conditions

Additionally, a D-optimal experiment design (Uciński, 2005; Alaña and Theodoropoulos, 2010) was carried out for each example considering the normalised and non-normalised sensitivity coefficients during the computation of the FIM. The results reported by this approach are shown in Table 5.

Table 5
Optimal sensor positions. D-optimal experiment design

D-optimal experiment design	Tubular reactor stable operation conditions		Tubular reactor unstable operation conditions	
	Concentration $y_1(t,z)$	**Temperature** $y_2(t,z)$	**Concentration** $y_1(t,z)$	**Temperature** $y_2(t,z)$
Normalising the sensitivity coefficients **Case E**	0.00 1.00	0.00 1.00	0.00 0.10 1.00	0.00 0.55 0.95
Non normalising the sensitivity coefficients **Case F**	0.10 1.00	0.15 1.00	0.00 0.25 0.30	0.20 1.00

In this table it can be observed that the D-optimal method reported slightly different results when the FIM uses normalised and/or non-normalised sensitivity coefficients. These optimal locations were considered to estimate parameters for each example and the results were compared with the ones obtained using the method introduced here.

To verify the quality of the results produced by the different approaches followed to define the optimal positions of the observations, a parameter estimation procedure, explained in Section 4, was implemented. The parameter estimates were compared with the nominal values shown in Table 1. The determinant of the parameter covariance matrix, equation (6), was used as a measure of identifiability and the design aim of the parameter estimates considering the scenarios reported in Tables 4 and 5. The confidence intervals, given by (9), were also calculated and the results are shown in Tables 6 and 7.

Table 6
Parameter estimates. Tubular reactor stable case

Scenario	Parameter estimates – Tubular reactor stable operation conditions						
	Pe_C	Pe_T	B_C	B_T	γ	β_T	SSE and $det[Cov(\theta)]$
A	0.9947 **(0.53%)** [± 0.053]	1.0002 **(0.02%)** [± 0.003]	1.0026 **(0.26%)** [± 0.024]	1.9960 **(0.20%)** [± 0.036]	19.1941 **(4.03%)** [± 0.368]	0.9933 **(0.67%)** [± 0.061]	$1.04e^{-4}$ **$2.10e^{-3}$**
B	0.9887 **(1.13%)** [± 0.102]	1.0009 **(0.09%)** [± 0.008]	0.9963 **(0.37%)** [± 0.034]	2.0043 **(0.22%)** [± 0.044]	21.3920 **(6.96%)** [± 0.709]	1.0021 **(0.21%)** [± 0.019]	$7.57e^{-4}$ **$5.33e^{-3}$**
C	0.9957 **(0.43%)** [± 0.039]	1.0018 **(0.18%)** [± 0.016]	1.0027 **(0.26%)** [± 0.024]	2.0035 **(0.18%)** [± 0.017]	15.4678 **(22.66%)** [± 7.670]	1.0001 **(0.01%)** [± 0.002]	$8.37e^{-3}$ **$25.67e^{-1}$**
D	0.9917 **(0.83%)** [± 0.075]	0.9983 **(0.17%)** [± 0.016]	1.0022 **(0.22%)** [± 0.020]	1.9976 **(0.12%)** [± 0.011]	19.1229 **(4.39%)** [± 0.400]	0.9985 **(0.15%)** [± 0.013]	$1.12e^{-4}$ **$5.59e^{-3}$**

Number in parentheses represents the percentage deviation of the parameter from the nominal value (Table 1)
Number in brackets represents the confidence interval.

The parameter more difficult to fit was γ. This can be explained by the fact of being the parameter with the lowest absolute sensitivity in the spatio-temporal domain. The presence of noise in the experimental data complicates even more its estimation. A large improvement is obtained when the measurements at location z=0.05 is considered, especially to identify γ. The uses of a reduced model to predict the outputs during the optimisation process enhance remarkably the estimates and the computation cost. This computation cost it correspond to the 60% of using the original governing equations. The covariance matrix is implicitly dependent on the sensor location. The variance of the estimates is thus influenced by the choice of sensor positions and the number of sensors *m*, which in turn means that an optimal choice of measurement locations will improve the accuracy of the estimates

Table 7
Parameter estimates. Tubular reactor unstable case

Scenario	Parameter estimates – Tubular reactor unstable operation conditions						
	Pe_C	Pe_T	B_C	B_T	γ	β_T	SSE and $det[Cov(\theta)]$
A[a]	6.9932 **(0.10%)** [± 1.100]	6.8257 **(2.49%)** [± 1.248]	0.0995 **(0.50%)** [± 0.017]	2.4685 **(1.26%)** [± 0.244]	10.0926 **(0.93%)** [± 0.854]	1.9910 **(0.45%)** [± 0.344]	$1.89e^{-4}$ **$4.09e^{-3}$**
B	7.0116 **(0.17%)** [± 1.053]	7.0317 **(0.43%)** [± 0.655]	0.0994 **(0.60%)** [± 0.020]	2.3963 **(4.15%)** [± 0.705]	10.2110 **(2.11%)** [± 1.002]	1.9470 **(2.65%)** [± 0.997]	$5.98e^{-4}$ **$9.65e^{-3}$**
D	7.0328 **(0.47%)** [± 1.201]	7.0199 **(0.29%)** [± 0.322]	0.0992 **(0.80%)** [± 0.038]	2.4155 **(3.38%)** [± 0.414]	10.1098 **(1.10%)** [±0.899]	1.9684 **(1.58%)** [± 0.505]	$3.24e^{-4}$ **$7.01e^{-3}$**
E[b]	5.2873 **(24.47%)** [± 3.001]	5.8595 **(16.29%)** [± 2.444]	0.0995 **(0.50%)** [± 0.053]	2.4077 **(3.69%)** [± 0.400]	10.2595 **(2.60%)** [± 1.323]	1.9553 **(2.24%)** [± 0.873]	$7.22e^{-2}$ **$141.33e^{-1}$**

[a] In this particular case the initial guess vector for the parameter needed to be changed as 70% of the nominal value (Table 1)
[b] Results obtained considering the optimal locations reported the D-optimal experiment design using the normalised sensitivity coefficients
Number in parentheses represents the percentage deviation of the parameter from the nominal value.
Number in brackets represents the confidence interval.

For this example, considering the measurements in the whole spatial domain, Case A, required initial parameters as a guess vector closer to the real ones, using the 50% of the nominal values do not produced good results, which indicates how the identification depends of the initial guess used during the optimisation process. This aspect can be overcome when the observations are taken at the locations established by the extrema values of the POD-modes and/or using a reduced model to predict the outputs.

Comparison of optimisation runs for the global objective function show that the information gained from the POD-modes extrema analysis was effective in both reducing the computational

load of the optimisation and finding an optimum parameter set. This analysis yielded that the measurement locations plays an essential role in the parameter estimation problem.

Due the complex nature of this example, the nominal values of the parameters are not reproduced satisfactorily, not even considering the whole experimental data in the spatial domain. Pe_T and Pe_C are the parameters showing the highest deviation from its nominal value. This gives another motivation for using sensor placement strategies based on the extrema evaluation of the predominant empirical eigenfunctions, because not only the uncertainty in parameter estimation was reduced (See the deviation from the nominal values, number in parentheses, shown in Tables 6 and 7), but the cost associated to such measures as well, which is significant, especially in situations where experimental measurements are very expensive. Information about a physical parameter may be most accurately gained at points in space with a high sensitivity to the parameter. Taking observations at locations defined by the extrema values of the POD-modes tends to yield relatively low variance.

The tubular reactor operating at unstable conditions it is highly nonlinear and difficult to control and indentify. In the particular case analysed here it was observed that parameters Pe_T and Pe_C show the lowest absolute sensitivity coefficients. Preliminary results have proven that the inverse problem it is usually bad conditioned; as some sensitivity coefficients are very small. In practice, this implies that any method based on gradient and/or Hessian information only modifies a few parameters (parameters with the highest values of the sensitivity coefficients). This problem can be overcome using optimisation techniques that do not use such information, as the Nealder-Mead method. A further improvement can be achieved if the experimental observations are taken at the locations where the POD-modes show extrema values combined with the use of a reduced

model to predict the outputs of the system. Here the objective function was evaluated varying Pe_T and Pe_C and keeping the other parameters constant using the original governing equations or the reduced model as predictive model (Fig. 3).

It can be noticed in Fig. 3 how the surface of $\Im(\theta)$ changed when a reduced model is used. On the three-dimensional surface of $\Im(\theta)$, when the original governing equations were used as predictive model, a flat valley can be seen as well as multiple areas with local minima. Each combination of parameter values in this valley corresponds to a model simulation which is able to describe the measured data equally well. This poses problems for the optimisation algorithms because no clear direction can be found in which $\Im(\theta)$ value improves and this result in a premature convergence of the algorithm and wrongly estimated parameters. When the reduced model is used to predict the outputs the areas with local minima disappeared which can improves the optimum search. Also, new curvatures in $\Im(\theta)$ are produced which can enhance to reach the global minimum of the objective function.

It is valuable to investigate, in future works, the influence of the predictive model used over the objective function for parameter estimation of highly nonlinear problems.

6. Conclusions

A sampling placement scheme to estimate parameters is introduced. The definition of the spatial positions is crucial in the parameter estimation process. The latter can be performed optimally when the observations are taken at the locations in the spatial domain where the empirical eigenfunctions reached their extrema values. The parameters estimates with the highest deviation from the nominal values are generally the ones exhibiting the lowest absolute sensitivity

coefficients. This estimation can be improved measuring at the locations where the POD-modes show extrema values. The results are strongly influenced by the presence of noise in the experimental data, but this can be solved using common filtering techniques. Not only the uncertainty in parameter estimation can be reduced by considering these sensor locations, but cost is decreased also, which is significant, especially in situations where experimental observations are expensive. While the D-optimal design requires post-processing of the sensitivity coefficients to determine positions and the optimal number of measurements, a mere observation of the extrema values of the POD-modes can produce the same results.

Literature cited

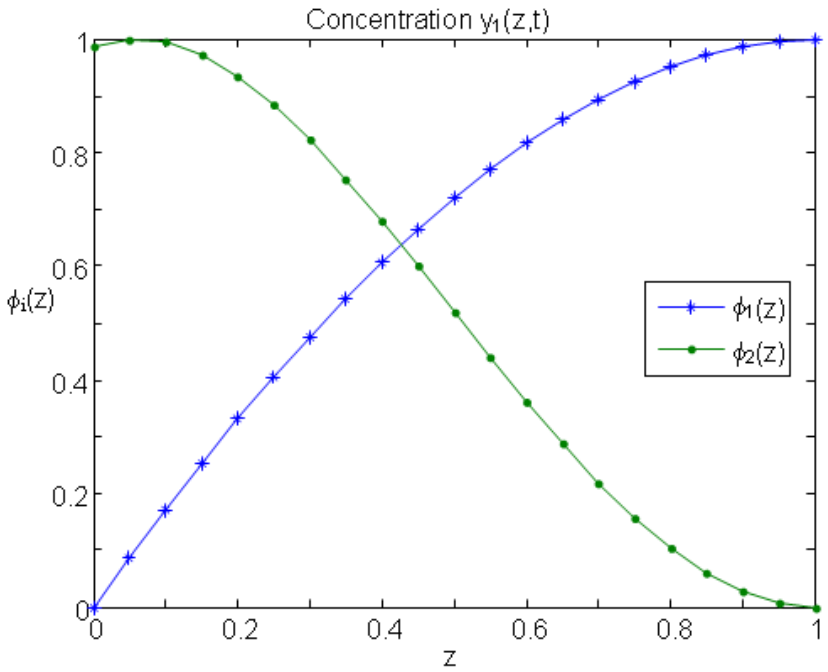
1. Alaña, J. E., Theodoropoulos, C., *Optimal location of measurements for Parameter Estimation of Distributed Parameter Systems*. Computers & Chemical Engineering, 35(1), 106(2011)
2. Alonso, A., Frouzakis, C., Kevrekidis, I., *Optimal sensor placement for state reconstruction of distributed process systems*. AIChE Journal, 50(7), 1438-1452(2004a).
3. Alonso, A., Kevrekidis, I., Banga, J., Frouzakis, C., *Optimal sensor location and reduced order observer design for distributed process systems*. Computers and Chemical Engineering, 28, 27-35(2004b).
4. Alvares, J., Romangnoli, J., Stepanopoulos, G., *Variable measurements structures for the control of a tubular reactor*. Chemical Engineering Science, 36, 1695(1981).
5. Amouroux, M., Babary, J. P., *Sensor and control location problems*. In M. G. Singh (ed.), Systems and Control Encyclopedia. Theory, Technology, Applications, Oxford: Pergamon Press, 4238(1988).
6. Atkinson, A. C., Donev, A. N., *Optimum Experimental Design*. Oxford University Press, Oxford (1992).
7. Balsa-Canto, E., Alonso, A. A., Banga, J. R., *Computing optimal dynamic experiments for model process calibration in predictive microbiology*. Journal of Food Process Engineering, 31(2), 186-206(2008).
8. Basseville, M., Benveniste, A., Moustakides, G. V., Rougee, A., *Optimal sensor location for detecting changes in dynamical behaviour*. IEEE Transactions on Automatic Control, 23, 1067-1075(1987).
9. Bendersky, E., Christofides, P. D., *Optimisation of transport-reaction processes using nonlinear model reduction*. Chemical Engineering Science, 55, 4349(2000).
10. Bernaerts, K., Gysemans, K. P. M., Nhan, T., Van Impe, J. F., *Optimal experiment design for cardinal values estimation: guidelines for data collection*. International Journal of Food Microbiology, 100(1-3), 153-185(2005).
11. Chmielewski, D. J., Tasha, P., Manousiouthakis V., *On the Theory of Optimal Sensor Placement*. AIChE Journal, 48(5):1001-1012(2002).

12. Christofides, P. D., Antoniades, C., *Computation of optimal actuator locations for nonlinear controllers in transport-reaction processes*. Computer and Chemical Engineering, 24, 577-583(2000).
13. Christofides, P. D., Antoniades, C., *Integrated optimal actuator/sensor placement and robust control of uncertain transport-reaction processes*. Computer and Chemical Engineering, 26, 187-203(2002).
14. Christofides, P. D., Antoniades, C., *Integrating nonlinear output feedback control and optimal actuator/sensor placement for transport-reaction processes*. Chemical Engineering Science, 56, 4517-4535(2001).
15. Cohen, K., Siegel, S., McLaughlim, T., *A heuristic approach to effective sensor placement for modelling of a cylinder wake*. Computers and Fluids, 35, 103-120(2006).
16. Cotae, P., Yalamnchili, S., Chen, C. L., Ayon, A., *Optimization of sensor locations and sensitivity analysis for engine health monitoring using minimum interference algorithms*. EURASIP Journal on Advances in Signal Processing, 9(2008).
17. Damak, T., Babary, J. P., Nihtila, M. T., *Observer design and sensor location in distributed parameter bioreactors*. Proceedings of DYCORD, Maryland, 87, 315-320(1992).
18. Dochain, D., Tali-Mammar, N., Babary, J., *On modeling, monitoring and control of fixed bed reactors*. Computers and Chemical Engineering, 21, 1255-1266(1997).
19. El Jai, A., *Distributed systems analysis via sensors and actuators*. Sensor and actuators A, 29, 1-11(1991).
20. El Jai, A., Pritchard, A. J., *Sensors and controls in the analysis of distributed systems*. New York: John Wiley and Sons (1988).
21. Fahroo, F., Demetriou, M.A., *Optimal actuator/sensor location for active noise regulator and tracking control problems*. Journal of Computational and Applied Mathematics, 114(1), 137-158(2000).
22. Feodorov, V. V., Hackl, P., *Model-oriented design of experiments*. Lecture Notes in Statistics, New York: Springer-Verlag (1997).
23. Georges, D., *The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems*. Proceeding of the 34th Conference on Decision and Control, New Orleans, LA, 4, 3319-3324(1995).
24. Grimstaadt, A. A., Mannseth, T., *Nonlinearity, Scale, and Sensitivity for a class of parameter estimation problem*. Proceeding 9th Conf. Inverse Problems in Engineering, Muncie, IN, (1998).
25. Harris, T. J., Macgregor, J. F., Wright, J. D., *Optimal sensor location with an application to a packed bed tubular reactor*. AIChE Journal, 26(6), 910-916(1980).
26. Hernandez, M. L., Kirubarajan, T., Bar-Shalom, Y., *Multisensor resource deployment using posterior Cramer-Rao bounds*. IEEE Transactions on Aerospace and Electronic Systems, 40(2), 399-416(2004).
27. Knopman, D. S., Voss, C. I., *Behavior of sensitivities in the one–dimensional advection-dispersion equation: implications for parameter estimation and sampling design*. Water Resources Research, 23(2), 253-272(1987).
28. Korbicz, J., *Discrete-scanning observation problem for stochastic non-linear discrete-time distributed parameter systems*. International Journal of Systems Science, 22(9), 1647-1662(1991).
29. Korbicz, J., Uciński, D., *Sensors allocation for state and parameter estimation of distributed systems*. In W. Gutkowski and J. Bauer (eds), Proc. IUTAM Symposium, Zakopane, Poland, 31 August–3 September, Berlin: Springer-Verlag, 178(1994).
30. Korbicz, J., Zgurovsky, M. Z., Novikov, A. N., *Suboptimal sensors location in the state*

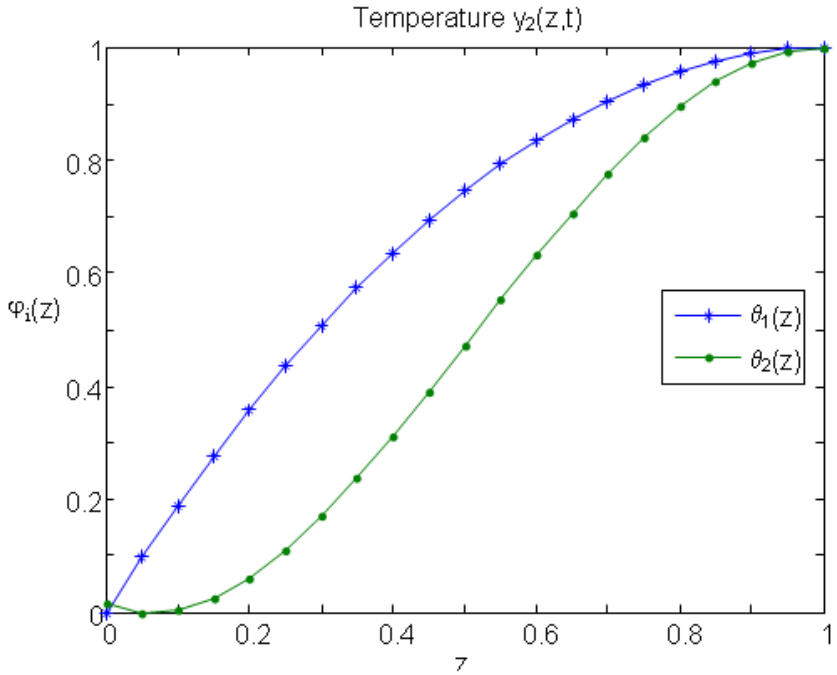
estimation problem for stochastic non-linear distributed parameter systems. International Journal of Systems Science, 19(9), 1871-1882(1988).
31. Kubrusly, C. S., Malebranche, H., *Sensors and controllers location in distributed systems-A survey*. Automatica, 21(2), 117-128(1985).
32. Li, M., Chistofides, P. D., *Optimal control of diffusion–convection–reaction processes using reduced–order models*. Computers and Chemical Engineering, 32(9), 2123-2135(2008).
33. Ljung, L., *System Identification–Theory for the user*. 2nd Edition, Prentice Hall, Upper Saddle River, NJ, USA (1999).
34. Löhner, R., Camelli, F., *Optimal Placement of Sensors for Contaminant Detection Based on Detailed 3D CFD*. Simulations Engineering Computations. International Journal for Computer-Aided Engineering, 22 (3), 260-273(2005).
35. Malebranche, J., *Simultaneous state and parameter estimation and location of sensor for distributed systems*. International Journal of Systems Science, 19(8), 1387-1405(1988).
36. Mossberg, M., *Optimal experimental design for identification of viscoelastic materials*. IEEE Transactions on Control Systems Technology, 12(4), 578-582(2004).
37. Nahor, H. B., Scheerlinck, N., Van Impe, J. F., Nicolaï, B. M., *Optimization of temperature sensor position in a hot wire probe set up for estimation of the thermal properties of foods using optimal experimental design*. Journal of Food Engineering, 57(1), 103-110(2003).
38. Olanrewaju, M. J., Al-Arfaj, M. A., *Estimator-based control of reactive distillation system: Application of an extended Kalman filtering*. Chemical Engineering Science, 61(10), 3386–3399(2006).
39. Papadimitriou, C., *Optimal sensor placement methodology for parametric identification of structural systems*. Journal of Sound Vibration, 278(4-5), 923-947(2004).
40. Pázman, A., *Foundations of optimum experimental design*. Mathematics and Its Applications, Dordrecht: D. Reidel Publishing Company, (1986).
41. Peng, T., *Fault Detection Methods Based on Optimal Sensor Location*. Hunan: Central South University, (2005).
42. Prasad, V., Vlachos, D. G., *Multiscale model and informatics based design of experiments: Application to the catalytic decomposition of Ammonia on Ruthenium*. Industrial and Engineering Chemistry Research, 47(17), 6555-6567(2008).
43. Pukelsheim, F., Rieder, S., *Efficient rounding of approximate designs*. Biometrica, 79(4), 763-770(1992).
44. Punithakumar, K., Kirubarajan, T., Hernandez, M., *Multisensor deployment using PCRLBs, incorporating sensor deployment and motion uncertainties.* IEEE Transactions on Aerospace and Electronic Systems, 42(4), 1474-1485(2006).
45. Qureshi, Z. H., Ng, T. S., Goodwin, G. C., *Optimum experimental design for identification of distributed parameter systems*. International Journal of Control, 31(1), 21-29(1980).
46. Rafajlowicz, E., *Algorithms of experimental design with implementations in MATHEMATICA*. Warsaw: Academic Publishers PLJ, (1996).
47. Rafajlowicz, E., *Design of experiments for eigenvalues identification in DPS*. International. Journal of Control, 34(6), 1079-1094(1981).
48. Rafajlowicz, E., *Optimal experiment design for identification of linear distributed-parameter systems: Frequency domain approach*. IEEE Transactions on Automatic Control, 28(7), 806-808(1983).
49. Rensfelt, A., Mousavi, S., Mossberg, M., Söderström, T., *Optimal sensor locations for nonparametric identification of viscoelastic materials*. Automatica, 44(1), 28-38(2008).

50. Rutzler, W., *Nonlinear and adaptive parameter estimation methods for tubular reactors*. Industrial and Engineering Chemistry Research, 26(2), 325-333(1987).
51. Singh, A. K., Hahn, J., *Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems*. Industrial and Engineering Chemistry Research, 44(5), 5645-5659(2005).
52. Singh, A. K., Hahn, J., *Sensor Location for Stable Nonlinear Systems: Placing Multiple Sensors*. Proceedings Chemical Process Control 7, Lake Louise, Canada, (2006).
53. Sirovich, L., *Turbulence and the dynamics of coherent structures*. Quarterly of Applied Mathematics, 45, 561-590(1987).
54. Söderström, T., Stoica, P., *System Identification*. Prentice Hall International, Hemel Hempstead, UK (1989).
55. Sorenson, H. W., *Parameter estimation: principles and problems.* New York, M. Dekker, (1980).
56. Sumana, C., Venkateswarlu, C., *Optimal selection of sensors for state estimation in a reactive distillation process*. Journal of Process Control, 19(6), 1024–1035(2009).
57. Titterington, D. M., *Aspects of optimal design in dynamic systems*. Technometrics, 22(3), 287-299(1980).
58. Uciński, D., Bogacka, B. *Construction of T-optimum designs for multiresponse dynamic models*. In COMPSTAT 2002, Berlin, Heidelberg: Physica, 267-272(2002).
59. Uciński, D., Bogacka, B. *T-optimum designs for discrimination between two multivariable dynamic models.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67, 3-18(2005).
60. Uciński, D., Demetriou, M., *An approach to the optimal scanning measurement problem using optimum experimental design*. Proceeding of the 2004 American Control Conference, June 30 – July 2, Boston, Massachusetts, (2004).
61. Uciński, D., *On the determination of moving sensor policies for parameter estimation of distributed systems*. 11th Mediterranean IEEE Conference on Control and Automation MED'03, June 18-20, Rhodes, Greece (2003).
62. Uciński, D., *Optimal measurement method for distributed parameter system identification*. CRC Press. USA, (2005).
63. Uciński, D., *Optimal sensor location for parameter estimation of distributed processes*. International Journal of Control, 73(13), 1235-1248(2000).
64. Uciński, D., Patan, M., *D-optimal design of a monitoring network for parameter estimation of distributed systems*. Journal of Global Optimization, 39(2), 291-322(2007).
65. Van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M., Smilde, A. K., *Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria*. Chemical Engineering Science, 55(4), 827-837(2000).
66. Vande, W. A., Point, N., Remy, M., *Practical issues in distributed parameter estimation: Gradient computation and optimal experiment design*. Control Engineering Practice, 4(11), 1553-1562(1996).
67. Venkateswarlu, C., Kumar, B. J., *Composition estimation of multi component reactive batch distillation with optimal sensor configuration*. Chemical Engineering Science, 61(7), 5560–5574(2006).
68. Wahl, T., Jones, W. E., Wilson, J. A., *A scheme for whole temperature profile control in distributed parameter systems*. Computer Aided Chemical Engineering, 10, 577-582(2002).
69. Waldraff, W., Dochain, D., Bourrel, S., Magnus, A., *On the use of observability measures for sensor location in tubular reactor*. Journal of Process Control, 8(5-6), 497-505(1998).
70. Walter, É., Prozanto, L., *Identification of parametric models from experimental data*.

Communications and Control Engineering, Berlin: Springer-Verlag, (1997).
71. Walter, É., Prozanto, L., *Qualitative and quantitative experiment design for phenomenological models-A survey*. Automatica, 26(2), 195-213(1990).
72. Waterhouse, T. H., Eccleston, J. A., Duffull, S. B., *Optimal design criteria for discrimination and estimation in nonlinear models*. Journal of Biopharmaceutical Statistics, 19(2) 386-402(2009).
73. Wouwer, A. V., Point, N., Poterman, S., Remy, M., *An approach to the selection of optimal sensor locations in distributed parameter systems*. Journal of Process Control, 10(4), 291-300(2000).
74. Yen, W. W-G., Yoon, Y. S., *Aquifer parameter identification with optimum dimension in parameterization*. Water Resources Research, 17(3), 664-672(1981).
75. Yildirim, B., Chryssostomidis, C., Karniadakis, G., *Efficient sensor placement for ocean measurements using low-dimensional concepts*. Ocean Modelling, 27(3-4), 160-173(2209).
76. Zamprohna, E., Barolo, M., Seborg, D. E., *Optimal Selection of Sensors Inputs for Batch Distillation Columns Using Principal Component Analysis*. Journal of Process Control, 15, 39-52(2005).
77. Zhu, Z., Stein, M. L., *Spatial sampling design for parameter estimation of the covariance function*. Journal of Statistical Planning and Inference, 134(2), 583-603(2005).
78. Zumoffen, D., Basualdo, M., *Optimal sensor location for chemical process accounting the best control configuration*. 10th International Symposium on Process Systems Engineering – PSE2009(2009).

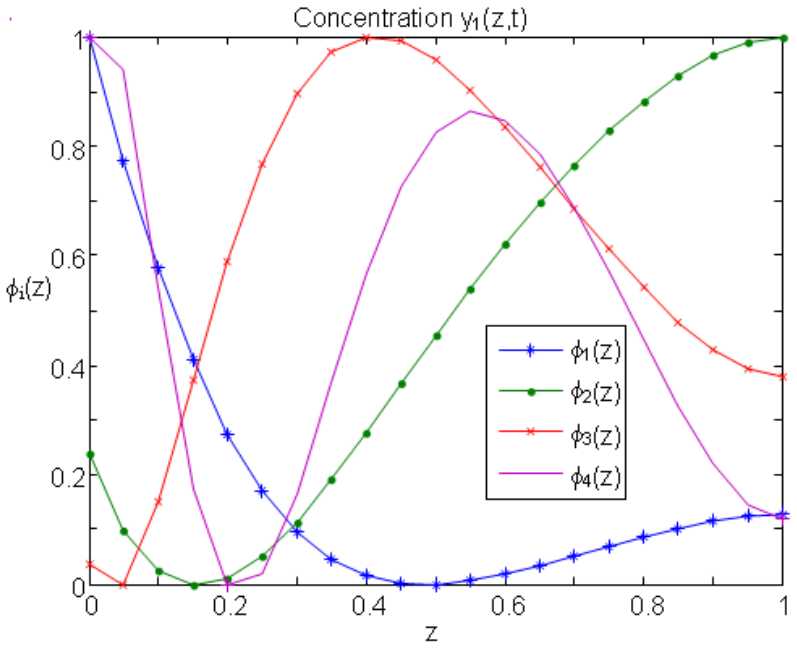


	$\phi_1(z)$	$\phi_2(z)$
Extrema locations	0.05 1.00	0.00 1.00

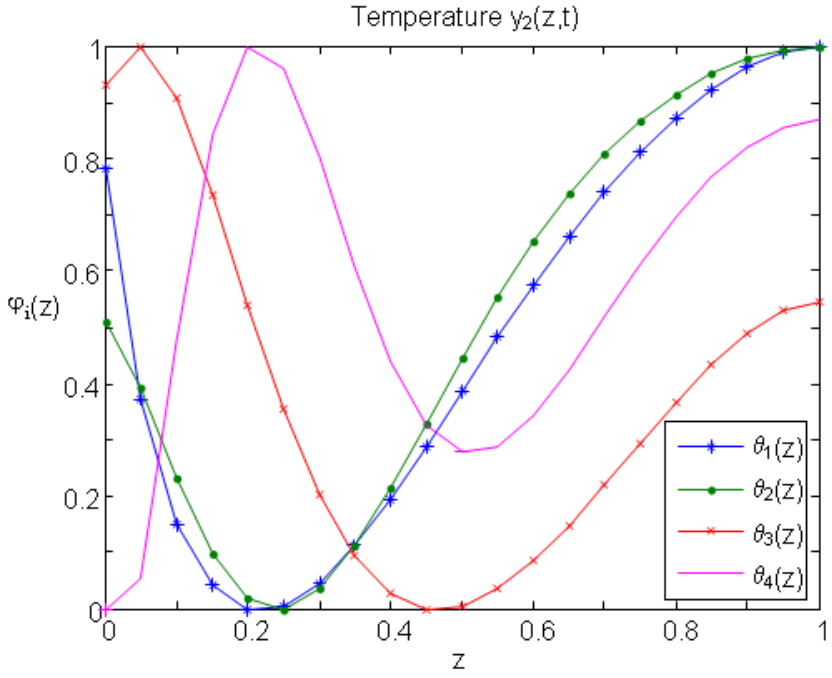


	$\varphi_1(z)$	$\varphi_2(z)$
Extrema locations	0.05 1.00	0.00 1.00

Fig. 1. Empirical eigenfunctions and extrema locations. Tubular reactor stable case

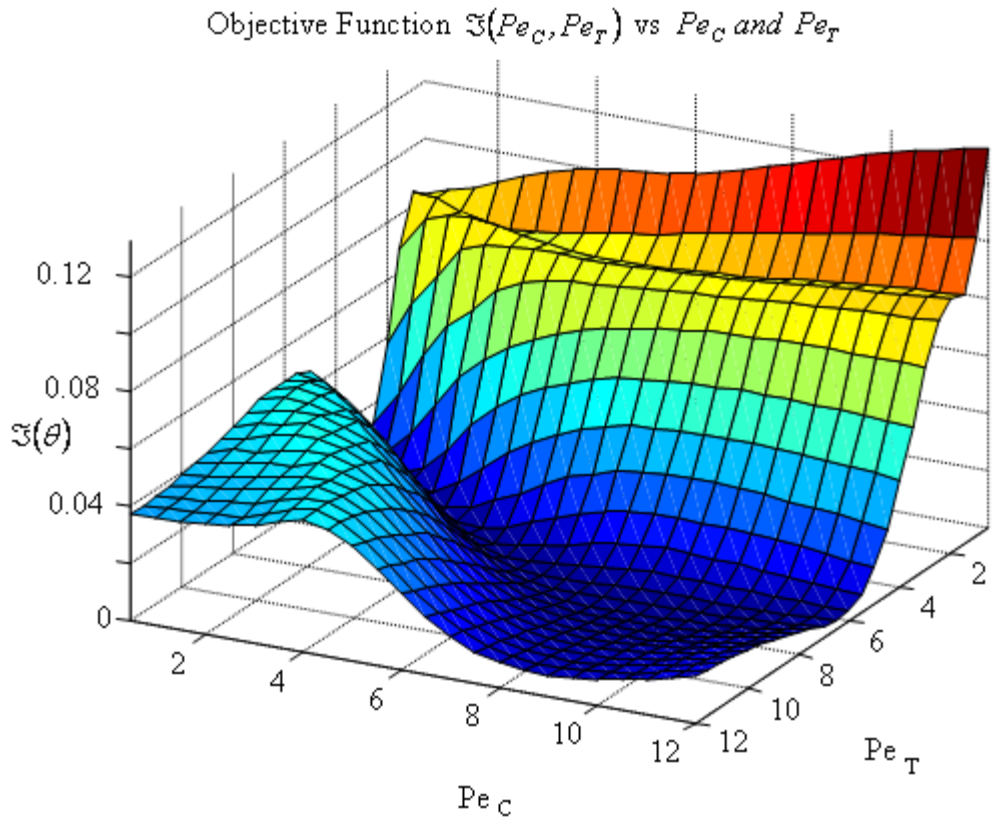


	$\phi_1(z)$	$\phi_2(z)$	$\phi_3(z)$	$\phi_4(z)$
Extrema locations	0.00 0.50	0.15 1.00	0.05 0.40	0.00 0.20 0.55

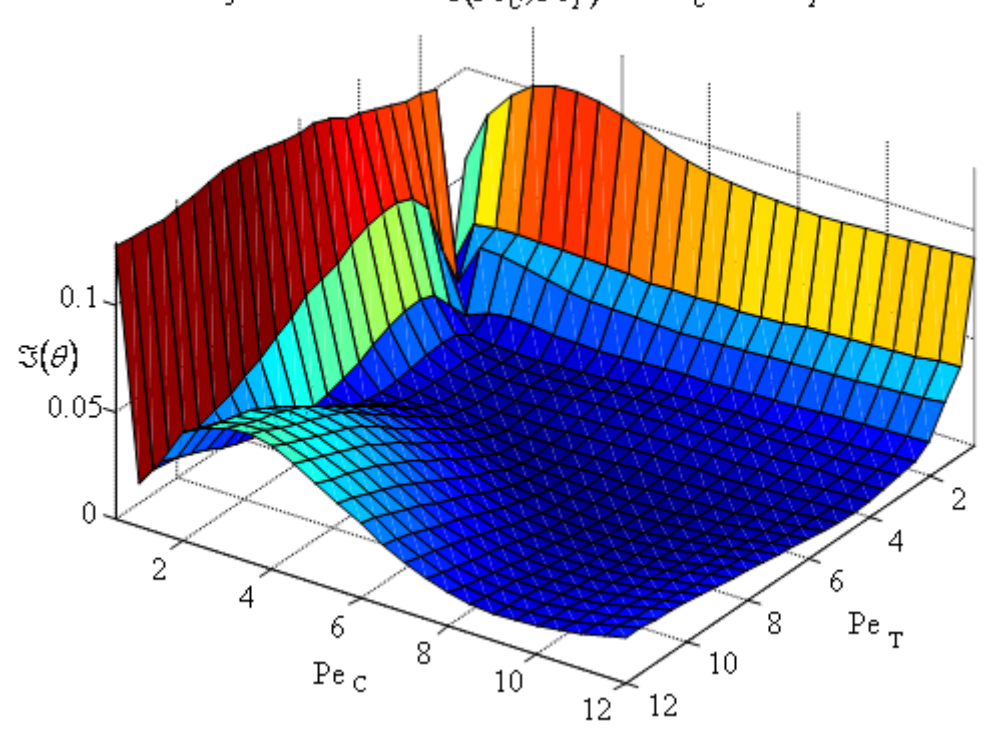


	$\varphi_1(z)$	$\varphi_2(z)$	$\varphi_3(z)$	$\varphi_4(z)$
Extrema locations	0.20 1.00	0.25 1.00	0.05 0.45	0.00 0.20 0.50

Fig. 2. Empirical eigenfunctions and extrema locations. Tubular reactor unstable case



a) **Using the original governing equations as predictive model in (5)**



a) **Using the reduced model as predictive model in (5)**

Fig. 3. Objective function surface. Tubular reactor unstable operation conditions

APPENDIXES

"Models, of course, are never true, but fortunately
it is only necessary that they be useful

Box

APPENDIX A

For the reaction system introduce in Chapter VI, Figure 6.2. Under the standard assumptions of constant density (ρ) and heat capacity (cp) of the reacting fluid, and constant axial fluid velocity (v), the dynamic behaviour of the reactor is described by the following set of partial differential equations defined on a spatial domain $z \in (0,1)$ derived from the mass and energy balances;

$$\frac{\partial T(z,t)}{\partial t} = -v\frac{\partial T(z,t)}{\partial z} + \frac{k}{\rho c_p}\frac{\partial^2 T(z,t)}{\partial z^2} + \frac{-\Delta H}{\rho c_p}\exp\left(-\frac{E}{RT(z,t)}\right)C_A(z,t) - \frac{hA_S}{\rho c_p}\left(T(z,t) - T_C(t)\right) \tag{A.1a}$$

$$\frac{\partial C_A(z,t)}{\partial t} = -v\frac{\partial C_A(z,t)}{\partial z} + D_A\frac{\partial^2 C_A(z,t)}{\partial z^2} - k_0\exp\left(-\frac{E}{RT(z,t)}\right)C_A(z,t) \tag{A.1b}$$

Where $C_A(z,t)$ and $T(z,t)$ stand for temperature and concentration of species A in the reactor, respectively, k and D_A are the thermal conductivity and mass diffusivity of the reacting fluid, respectively, k_0, E and $(-\Delta H)$ represent the pre-exponential constant, activation energy and the heat of the reaction, respectively, h is the heat transfer coefficient between the reactor and the cooling jacket, A_s is the surface area of the reactor walls, and Tc(t) corresponds to the temperature of the cooling medium . Assuming negligible reaction in the recycle loop and instantaneous mixing of fresh feed

and the recycle at the inlet, the boundary conditions of the system take the form;

At z=0;

$$\frac{\partial T(0,t)}{\partial z} = \frac{\rho c_p v}{k}\left[T(0,t) - (1-r)T_0(t) - rT(1,t)\right] \tag{A.2a}$$

$$\frac{\partial C_A(0,t)}{\partial z} = \frac{v}{D_A}\left[C_A(0,t) - (1-r)C_{A0}(t) - rC_A(1,t)\right] \tag{A.2b}$$

At z=1;

$$\frac{\partial T(1,t)}{\partial z} = 0 \tag{A.3a}$$

$$\frac{\partial C_A(1,t)}{\partial z} = 0 \tag{A.3b}$$

Where $T_0(t)$ and $C_{A0}(t)$ denote the inlet temperature and concentration of species A in the reactor, r is the recycle ratio, which varies from zero to one, with one corresponding to total recycle and zero to fresh feed or no recycle. To simplify this representation, the following dimensionless variables can be introduced;

$$Pe_C = \frac{v}{D_A} \tag{A.4a}$$

$$Pe_T = \frac{\rho c_p v}{k} \tag{A.4b}$$

$$y_1 = \frac{T(z,t) - T_R}{T_R} \tag{A.4c}$$

$$y_2 = \frac{C_A(z,t) - C_{AR}}{C_{AR}} \tag{A.4d}$$

$$y_{10} = \frac{T_0 - T_R}{T_R} \tag{A.4e}$$

$$y_{20} = \frac{C_{A0} - C_{AR}}{C_{AR}} \tag{A.4f}$$

$$y_C = \frac{T_C - T_R}{T_R} \tag{A.4g}$$

$$\gamma = \frac{E}{RT_R} \tag{A.4h}$$

$$\beta_T = \frac{hA_s}{C_{AR}} \tag{A.4i}$$

$$B_T = \frac{-(\Delta H)C_{AR}}{\rho c_p T_R} \tag{A.4j}$$

$$B_C = \frac{k_0 \exp\left(-\frac{\mathrm{E}}{\mathrm{RT_R}}\right)}{v} \tag{A.4k}$$

The system of equations (A.1) can be written in the following form;

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) - \beta_T (y_C - y_1) \tag{A.5a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} - B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \tag{A.5b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = Pe_T\left[y_1(0,t) - (1-r)y_{10} - ry_1(1,t)\right] \tag{A.6a}$$

$$\frac{\partial y_2}{\partial z} = Pe_C\left[y_2(0,t) - (1-r)y_{20} - ry_2(1,t)\right] \tag{A.6b}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \tag{A.7a}$$

$$\frac{\partial y_2}{\partial z} = 0 \tag{A.7b}$$

For the particular case studied here, the operating conditions of the reactor are; C_{A0}=T_0=Tc=0 with a recycling relation r=0.5. The initial conditions in dimensionless form are;

$$y_{10} = y_{20} = y_C = 0 \tag{A.8}$$

In the specific case of no recycle in the system (r=0.00). The equations are;

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C(1-y_2)\exp\left(\frac{y_1}{1+\frac{y_1}{\gamma}}\right) - \beta_T(u_C - y_1) \tag{A.9a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} + B_C(1-y_2)\exp\left(\frac{y_1}{1+\frac{y_1}{\gamma}}\right) \tag{A.9b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = Pe_T y_1(0,t) \qquad \text{(A.10a)}$$

$$\frac{\partial y_2}{\partial z} = Pe_C y_2(0,t) \qquad \text{(A.10b)}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \qquad \text{(A.11a)}$$

$$\frac{\partial y_2}{\partial z} = 0 \qquad \text{(A.11b)}$$

With the same initial conditions presented above.

APPENDIX B

pdepe

Solve initial-boundary value problems for systems of parabolic and elliptic partial differential equations (PDEs) in one space variable and time.

Syntax

sol = pdepe(m,pdefun,icfun,bcfun,xmesh,tspan)

sol = pdepe(m,pdefun,icfun,bcfun,xmesh,tspan,options)

sol = pdepe(m,pdefun,icfun,bcfun,xmesh,tspan,options,p1,p2...)

Arguments

m A parameter corresponding to the symmetry of the problem. m can be slab = 0, cylindrical = 1, or spherical = 2.

pdefun A function that defines the components of the PDE.

icfun A function that defines the initial conditions.

bcfun A function that defines the boundary conditions.

xmesh A vector $[x_0,x_1,...,x_n]$ specifying the points at which a numerical solution is requested for every value in tspan. The elements of xmesh must satisfy $x_0<x_1<...<x_n$. The length of xmesh must be >= 3.

tspan A vector $[t_0,t_1,...,t_f]$ specifying the points at which a solution is requested for every value in xmesh. The elements of tspan must satisfy $t_0<t_1<...<t_f$. The length of tspan must be >= 3.

options Some options of the underlying ordinary differential equations solver are available in pdepe: RelTol, AbsTol, NormControl, InitialStep, and MaxStep. In most cases, default values for these options provide satisfactory solutions.

p1, p2,... Optional parameters to be passed to pdefun, icfun, and bcfun.

Description

sol = pdepe(m,pdefun,icfun,bcfun,xmesh,tspan) solves initial-boundary value problems for systems of parabolic and elliptic partial differential equations in the one space variable x and time t. The ordinary differential equations resulting from discretization in space are integrated to obtain approximate solutions at times specified in tspan. The pdepe function returns values of the solution on a mesh provided in xmesh.

pdepe solves PDEs of the form:

$$c\left(x,t,u,\frac{\partial u}{\partial x}\right)\frac{\partial u}{\partial t}=x^{-m}\frac{\partial}{\partial x}\left(x^m f\left(x,t,u,\frac{\partial u}{\partial x}\right)\right)+s\left(x,t,u,\frac{\partial u}{\partial x}\right) \qquad \text{(B.1)}$$

The PDEs hold for $t_0 \le t \le t_f$ and $a \le x \le b$. The interval $[a,b]$ must be finite. m can be 0, 1, or 2, corresponding to slab, cylindrical, or spherical symmetry, respectively. If m>0, then must be >= 0. In Equation B.1, $f\left(x,t,u,\frac{\partial u}{\partial x}\right)$ is a flux term and $s\left(x,t,u,\frac{\partial u}{\partial x}\right)$ is a source term. The coupling of the partial derivatives with respect to time is restricted to multiplication by a diagonal matrix $c\left(x,t,u,\frac{\partial u}{\partial x}\right)$. The diagonal

elements of this matrix are either identically zero or positive. An element that is identically zero corresponds to an elliptic equation and otherwise to a parabolic equation. There must be at least one parabolic equation. An element of "c" that corresponds to a parabolic equation can vanish at isolated values of x if those values of x are mesh points. Discontinuities in "c" and/or "s" due to material interfaces are permitted provided that a mesh point is placed at each interface. For $t=t_0$ and all x, the solution components satisfy initial conditions of the form,

$$u(x,t_0)=u_0(x) \tag{B.2}$$

For all t and either x=a or x=b, the solution components satisfy a boundary condition of the form,

$$p(x,t,u)+q(x,t)f\left(x,t,u,\frac{\partial u}{\partial x}\right)=0 \tag{B.3}$$

Elements of q are either identically zero or never zero. Note that the boundary conditions are expressed in terms of the flux "f" rather than $\frac{\partial u}{\partial x}$. Also, of the two coefficients, only "p" can depend on u.

Example. This example illustrates the straightforward formulation, computation, and plotting of the solution of a single PDE.

$$\pi^2\frac{\partial u}{\partial t}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) \tag{B.4}$$

This equation holds on an interval $0 \le x \le 1$ for times $t \ge 0$.

The PDE satisfies the initial condition,

$$u(x,0)=\sin(\pi x) \tag{B.5}$$

And boundary conditions,

$$u(0,t)=0 \tag{B.6a}$$

$$\pi e^{-t} + \frac{\partial u(1,t)}{\partial x} = 0 \qquad \text{(B.6b)}$$

It is convenient to use subfunctions to place all the functions required by pdepe in a single M-file.

```
function pdex1
m = 0;
x = linspace(0,1,20);
t = linspace(0,2,5);
sol = pdepe(m,@pdex1pde,@pdex1ic,@pdex1bc,x,t);
% Extract the first solution component as u.
u = sol(:,:,1);
% A surface plot is often a good way to study a solution.
surf(x,t,u)
title('Numerical solution computed with 20 mesh points.')
xlabel('Distance x')
ylabel('Time t')
% A solution profile can also be illuminating.
figure
plot(x,u(end,:))
title('Solution at t = 2')
xlabel('Distance x')
ylabel('u(x,2)')
% --------------------------------------------------
function [c,f,s] = pdex1pde(x,t,u,DuDx)
c = pi^2;
f = DuDx;
s = 0;
% --------------------------------------------------
function u0 = pdex1ic(x)
u0 = sin(pi*x);
% --------------------------------------------------
function [pl,ql,pr,qr] = pdex1bc(xl,ul,xr,ur,t)
pl = ul;
ql = 0;
pr = pi * exp(-t);
qr = 1;
```

APPENDIX C

SENSITIVITY ANALYSIS. TUBULAR REACTOR WITH RECYCLE

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + \beta_T (u_C - y_1) \tag{C.1a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} - B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \tag{C.1b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = Pe_T\left[y_1(t,0) - (1-r)y_{10} - r y_1(t,1)\right] \tag{C.2a}$$

$$\frac{\partial y_2}{\partial z} = Pe_C\left[y_2(t,0) - (1-r)y_{20} - r y_2(t,1)\right] \tag{C.2b}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \tag{C.3a}$$

$$\frac{\partial y_2}{\partial z} = 0 \tag{C.3b}$$

The parameters values are,

$$\begin{aligned} Pe_C &= 7.0 \quad B_C = 0.1 \quad \gamma = 10.0 \\ Pe_T &= 7.0 \quad B_T = 2.5 \quad \beta_T = 2.0 \end{aligned} \tag{C.4}$$

The reactor is operating at C_{A0}=To=Tc=0 with a recycling relation r=0.5.

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) - \beta_T y_1 \tag{C.5a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} - B_C(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \tag{C.5b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = Pe_T\left[y_1(t,0) - r y_1(t,1)\right] \tag{C.6a}$$

$$\frac{\partial y_2}{\partial z} = Pe_C\left[y_2(t,0) - r y_2(t,1)\right] \tag{C.6b}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \tag{C.7a}$$

$$\frac{\partial y_2}{\partial z} = 0 \tag{C.7b}$$

The derivatives of equations (C.5) – (C.7) with respect to each parameter;

$$y_3 = \frac{\partial y_1}{\partial Pe_T} \tag{C.8a}$$

$$\begin{aligned}\frac{\partial y_3}{\partial t} = &-\frac{\partial y_3}{\partial z} - \frac{1}{Pe_T^{\,2}}\frac{\partial^2 y_1}{\partial z^2} + \frac{1}{Pe_T}\frac{\partial^2 y_3}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial Pe_T} \\ &+ B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_3 - \beta_T y_3\end{aligned} \tag{C.9a}$$

$$y_4 = \frac{\partial y_1}{\partial B_T} \tag{C.8b}$$

$$\begin{aligned}\frac{\partial y_4}{\partial t} = &-\frac{\partial y_4}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_4}{\partial z^2} + B_C(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial B_T} \\ &+ B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_4 - \beta_T y_4\end{aligned} \tag{C.9b}$$

$$y_5 = \frac{\partial y_1}{\partial B_C} \tag{C.8c}$$

$$\frac{\partial y_5}{\partial t} = -\frac{\partial y_5}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_5}{\partial z^2} + B_T(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial B_C} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_5 - \beta_T y_5 \tag{C.9c}$$

$$y_6 = \frac{\partial y_1}{\partial \gamma} \tag{C.8d}$$

$$\frac{\partial y_6}{\partial t} = -\frac{\partial y_6}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_6}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial \gamma} + B_T B_C \frac{(1+y_2)}{(1+y_1)} y_1 \exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_6 - \beta_T y_6 \tag{C.9d}$$

$$y_7 = \frac{\partial y_1}{\partial \beta_T} \tag{C.8e}$$

$$\frac{\partial y_7}{\partial t} = -\frac{\partial y_7}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_7}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial \beta_T} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_7 - y_1 - \beta_T y_7 \tag{C.9e}$$

$$y_8 = \frac{\partial y_1}{\partial Pe_C} \tag{C.8f}$$

$$\frac{\partial y_8}{\partial t} = -\frac{\partial y_8}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_8}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_2}{\partial Pe_C} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_8 - \beta_T y_8 \tag{C.9f}$$

$$y_9 = \frac{\partial y_2}{\partial Pe_T} \tag{C.8g}$$

$$\frac{\partial y_9}{\partial t} = -\frac{\partial y_9}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_9}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_9 - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial Pe_T} \tag{C.9g}$$

$$y_{10} = \frac{\partial y_2}{\partial B_T} \tag{C.8h}$$

$$\begin{aligned}\frac{\partial y_{10}}{\partial t} = &-\frac{\partial y_{10}}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_{10}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{10} \\ &- B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial B_T}\end{aligned} \tag{C.9h}$$

$$y_{11} = \frac{\partial y_2}{\partial B_C} \tag{C.8i}$$

$$\begin{aligned}\frac{\partial y_{11}}{\partial t} = &-\frac{\partial y_{11}}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_{11}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{11} - (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \\ &- B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial B_C}\end{aligned} \tag{C.9i}$$

$$y_{12} = \frac{\partial y_2}{\partial \gamma} \tag{C.8j}$$

$$\begin{aligned}\frac{\partial y_{12}}{\partial t} = &-\frac{\partial y_{12}}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_{12}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{12} - B_C \frac{(1+y_2)}{(1+y_1)} y_1 \exp\left(\frac{\gamma y_1}{1+y_1}\right) \\ &- B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial \gamma}\end{aligned} \tag{C.9j}$$

$$y_{13} = \frac{\partial y_2}{\partial \beta_T} \tag{C.8k}$$

$$\begin{aligned}\frac{\partial y_{13}}{\partial t} = &-\frac{\partial y_{13}}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_{13}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{13} + \\ &+ B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial \beta_T}\end{aligned} \tag{C.9k}$$

$$y_{14} = \frac{\partial y_2}{\partial Pe_C} \tag{C.8m}$$

$$\frac{\partial y_{14}}{\partial t} = -\frac{\partial y_{14}}{\partial z} - \frac{1}{Pe_C{}^2}\frac{\partial^2 y_2}{\partial z^2} + \frac{1}{Pe_T}\frac{\partial^2 y_{14}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{14} - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right)\frac{\partial y_1}{\partial Pe_T} \quad \text{(C.9m)}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = Pe_T\left[y_1(t,0) - r y_1(t,1)\right] \quad \text{(C.10a)}$$

$$\frac{\partial y_2}{\partial z} = Pe_C\left[y_2(t,0) - r y_2(t,1)\right] \quad \text{(C.10b)}$$

$$\frac{\partial y_3}{\partial z} = \left[y_1(t,0) - r y_1(t,1)\right] + Pe_T\left[y_3(t,0) - r y_3(t,1)\right] \quad \text{(C.10c)}$$

$$\frac{\partial y_4}{\partial z} = Pe_T\left[y_4(t,0) - r y_4(t,1)\right] \quad \text{(C.10d)}$$

$$\frac{\partial y_5}{\partial z} = Pe_T\left[y_5(t,0) - r y_5(t,1)\right] \quad \text{(C.10e)}$$

$$\frac{\partial y_6}{\partial z} = Pe_T\left[y_6(t,0) - r y_6(t,1)\right] \quad \text{(C.10f)}$$

$$\frac{\partial y_7}{\partial z} = Pe_T\left[y_7(t,0) - r y_7(t,1)\right] \quad \text{(C.10g)}$$

$$\frac{\partial y_8}{\partial z} = Pe_T\left[y_8(t,0) - r y_8(t,1)\right] \quad \text{(C.10h)}$$

$$\frac{\partial y_9}{\partial z} = Pe_C\left[y_9(t,0) - r y_9(t,1)\right] \quad \text{(C.10i)}$$

$$\frac{\partial y_{10}}{\partial z} = Pe_C\left[y_{10}(t,0) - r y_{10}(t,1)\right] \quad \text{(C.10j)}$$

$$\frac{\partial y_{11}}{\partial z} = Pe_C\left[y_{11}(t,0) - r y_{11}(t,1)\right] \quad \text{(C.10k)}$$

$$\frac{\partial y_{12}}{\partial z} = Pe_C\left[y_{12}(t,0) - r y_{12}(t,1)\right] \quad \text{(C.10n)}$$

$$\frac{\partial y_{13}}{\partial z} = Pe_C\left[y_{13}(t,0) - r y_{13}(t,1)\right] \quad \text{(C.10p)}$$

$$\frac{\partial y_{14}}{\partial z} = \left[y_2(t,0) - r y_2(t,1)\right] + Pe_C\left[y_{14}(t,0) - r y_{14}(t,1)\right] \quad \text{(C.10q)}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = \frac{\partial y_2}{\partial z} = \ldots = \frac{\partial y_{13}}{\partial z} = \frac{\partial y_{14}}{\partial z} = 0 \quad \text{(C.11)}$$

$$\frac{\partial y_3}{\partial t} = -\frac{\partial y_3}{\partial z} - \frac{1}{Pe_T{}^2}\frac{\partial^2 y_1}{\partial z^2} + \frac{1}{Pe_T}\frac{\partial^2 y_3}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_9 + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_3 - \beta_T y_3 \quad \text{(C.12a)}$$

$$\frac{\partial y_4}{\partial t} = -\frac{\partial y_4}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_4}{\partial z^2} + B_C(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{10} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_4 - \beta_T y_4 \quad \text{(C.12b)}$$

$$\frac{\partial y_5}{\partial t} = -\frac{\partial y_5}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_5}{\partial z^2} + B_T(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{11} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_5 - \beta_T y_5 \quad \text{(C.12c)}$$

$$\frac{\partial y_6}{\partial t} = -\frac{\partial y_6}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_6}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{12} + B_T B_C \frac{(1+y_2)}{(1+y_1)} y_1 \exp\left(\frac{\gamma y_1}{1+y_1}\right) + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_6 - \beta_T y_6 \quad \text{(C.12d)}$$

$$\frac{\partial y_7}{\partial t} = -\frac{\partial y_7}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_7}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{13} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_7 - y_1 - \beta_T y_7 \quad \text{(C.12e)}$$

$$\frac{\partial y_8}{\partial t} = -\frac{\partial y_8}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_8}{\partial z^2} + B_T B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{14} + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_8 - \beta_T y_8 \quad \text{(C.12f)}$$

$$\frac{\partial y_9}{\partial t} = -\frac{\partial y_9}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_9}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_9 - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_3 \quad \text{(C.12g)}$$

$$\frac{\partial y_{10}}{\partial t} = -\frac{\partial y_{10}}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_{10}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{10} - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_4 \quad \text{(C.12h)}$$

$$\frac{\partial y_{11}}{\partial t} = -\frac{\partial y_{11}}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_{11}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{11} - (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_5 \quad \text{(C.12i)}$$

$$\frac{\partial y_{12}}{\partial t} = -\frac{\partial y_{12}}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_{12}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{12} - B_C \frac{(1+y_2)}{(1+y_1)} y_1 \exp\left(\frac{\gamma y_1}{1+y_1}\right) - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_6 \quad \text{(C.12j)}$$

$$\frac{\partial y_{13}}{\partial t} = -\frac{\partial y_{13}}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_{13}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{13} + + B_T B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_7 \quad \text{(C.12k)}$$

$$\frac{\partial y_{14}}{\partial t} = -\frac{\partial y_{14}}{\partial z} - \frac{1}{Pe_C{}^2}\frac{\partial^2 y_2}{\partial z^2} + \frac{1}{Pe_T}\frac{\partial^2 y_{14}}{\partial z^2} - B_C \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_{14} - B_C \frac{(1+y_2)}{(1+y_1)^2}\gamma \exp\left(\frac{\gamma y_1}{1+y_1}\right) y_8 \quad \text{(C.12m)}$$

APPENDIX D

REDUCED MODEL. TUBULAR CHEMICAL REACTOR WITH RECYCLE

This system is described by the following non – linear partial differential equations,

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) + \beta_T (u_C - y_1) \tag{D.1a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} - B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \tag{D.1b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = -Pe_T\left[y_1(t,0) - (1-r)y_{10} - r y_1(t,1)\right] \tag{D.2a}$$

$$\frac{\partial y_2}{\partial z} = -Pe_C\left[y_2(t,0) - (1-r)y_{20} - r y_2(t,1)\right] \tag{D.2b}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \tag{D.3a}$$

$$\frac{\partial y_2}{\partial z} = 0 \tag{D.3b}$$

The parameters values are,

$$\begin{aligned} Pe_C &= 7.0 \quad B_C = 0.1 \quad \gamma = 10.0 \\ Pe_T &= 7.0 \quad B_T = 2.5 \quad \beta_T = 2.0 \end{aligned} \tag{D.4}$$

The reactor is operating at C_{A0}=To=Tc=0 with a recycling relation r=0.5.

$$\frac{\partial y_1}{\partial t} = -\frac{\partial y_1}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 y_1}{\partial z^2} + B_T B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) - \beta_T y_1 \tag{D.5a}$$

$$\frac{\partial y_2}{\partial t} = -\frac{\partial y_2}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 y_2}{\partial z^2} - B_C (1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right) \tag{D.5b}$$

At z=0;

$$\frac{\partial y_1}{\partial z} = -Pe_T\left[y_1(t,0) - r y_1(t,1)\right] \tag{D.6a}$$

$$\frac{\partial y_2}{\partial z} = -Pe_C\left[y_2(t,0) - r y_2(t,1)\right] \tag{D.6b}$$

At z=1;

$$\frac{\partial y_1}{\partial z} = 0 \tag{D.7a}$$

$$\frac{\partial y_2}{\partial z} = 0 \tag{D.7b}$$

Applying Galerkin projection,

$$\begin{aligned}\int_0^1 \phi_j(z)\frac{\partial y_1}{\partial t}dz = &-\int_0^1 \phi_j(z)\frac{\partial y_1}{\partial z}dz + \frac{1}{Pe_T}\int_0^1 \phi_j(z)\frac{\partial^2 y_1}{\partial z^2}dz \\ &+ B_T B_C\int_0^1 \phi_j(z)(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right)dz - \beta_T\int_0^1 \phi_j(z) y_1 dz\end{aligned} \tag{D.8a}$$

$$\begin{aligned}\int_0^1 \theta_j(z)\frac{\partial y_2}{\partial t}dz = &-\int_0^1 \theta_j(z)\frac{\partial y_2}{\partial z}dz + \frac{1}{Pe_C}\int_0^1 \theta_j(z)\frac{\partial^2 y_2}{\partial z^2}dz \\ &- B_C\int_0^1 \theta_j(z)(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right)dz\end{aligned} \tag{D.8b}$$

$$y_1(z,t) = \sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z) \tag{D.9a}$$

$$y_2(z,t) = \sum_{i=1}^{N_2} \beta_i(t)\theta_i(z) + \bar{y}_2(z) \tag{D.9b}$$

$$\frac{\partial y_1(z,t)}{\partial t} = \sum_{i=1}^{N_1} \dot{\alpha}_i(t)\phi_i(z) \tag{D.10a}$$

$$\frac{\partial y_2(z,t)}{\partial t} = \sum_{i=1}^{N_2} \dot{\beta}_i(t)\theta_i(z) \tag{D.10b}$$

$$\frac{\partial y_1(z,t)}{\partial z} = \sum_{i=1}^{N_1} \alpha_i(t)\dot{\phi}_i(z) + \frac{\partial \bar{y}_1(z)}{\partial z} \tag{D.11a}$$

$$\frac{\partial y_2(z,t)}{\partial z} = \sum_{i=1}^{N_2} \beta_i(t)\dot{\theta}_i(z) + \frac{\partial \bar{y}_2(z)}{\partial z} \tag{D.11b}$$

For the mass balance,

$$\begin{aligned}\int_0^1 \theta_j(z)\left[\sum_{i=1}^{N_2} \dot{\beta}_i(t)\theta_i(z)\right]dz = &-\int_0^1 \theta_j(z)\left[\sum_{i=1}^{N_2} \beta_i(t)\dot{\theta}_i(z) + \frac{\partial \bar{y}_2(z)}{\partial z}\right]dz \\ &+ \frac{1}{Pe_C}\left[\theta_j(z)\frac{\partial y_2}{\partial z}\bigg|_0^1 - \int_0^1 \dot{\theta}_j(z)\left[\sum_{i=1}^{N_2} \beta_i(t)\dot{\theta}_i(z) + \frac{\partial \bar{y}_2(z)}{\partial z}\right]dz\right] \\ &- B_C\int_0^1 \theta_j(z)\left[1 + \sum_{i=1}^{N_2} \beta_i(t)\theta_i(z) + \bar{y}_2(z)\right]\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1} \alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1 + \sum_{i=1}^{N_1} \alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz\end{aligned} \tag{D.12}$$

$$\int_0^1 \theta_j(z)\left[\sum_{i=1}^{N_2}\dot{\beta}_i(t)\theta_i(z)\right]dz = -\int_0^1 \theta_j(z)\sum_{i=1}^{N_2}\beta_i(t)\dot{\theta}_i(z)dz$$
$$-\int_0^1 \theta_j(z)\frac{\partial \overline{y}_2(z)}{\partial z}dz$$
$$+\frac{1}{Pe_C}\left[\theta_j(z)\frac{\partial y_2}{\partial z}\right|_0^1$$
$$-\int_0^1 \dot{\theta}_j(z)\sum_{i=1}^{N_2}\beta_i(t)\dot{\theta}_i(z)dz - \int_0^1 \dot{\theta}_j(z)\frac{\partial \overline{y}_2(z)}{\partial z}dz\Bigg]$$
$$-B_C\int_0^1 \theta_j(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)}\right)dz \qquad \text{(D.13)}$$
$$-B_C\int_0^1 \theta_j(z)\sum_{i=1}^{N_2}\beta_i(t)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)}\right)dz$$
$$-B_C\int_0^1 \theta_j(z)\overline{y}_2(z,t)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)}\right)dz$$

$$\int_0^1 \theta_j(z)\left[\sum_{i=1}^{N_2}\dot{\beta}_i(t)\theta_i(z)\right]dz = -\sum_{i=1}^{N_2}\beta_i(t)\left[\int_0^1 \theta_j(z)\dot{\theta}_i(z)dz\right.$$
$$\left.+\frac{1}{Pe_C}\int_0^1 \dot{\theta}_j(z)\dot{\theta}_i(z)dz + B_C\int_0^1 \theta_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)}\right)dz\right]$$
$$-\int_0^1\left(\theta_j(z)+\frac{1}{Pe_C}\dot{\theta}_j(z)\right)\frac{\partial \overline{y}_2(z)}{\partial z}dz \qquad \text{(D.14)}$$
$$-B_C\int_0^1 \theta_j(z)\left(1+\overline{y}_2(z,t)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\overline{y}_1(z)}\right)dz$$
$$+\frac{1}{Pe_C}\left[\theta_j(z)\frac{\partial y_2}{\partial z}\right|_0^1$$

$$\left.\frac{\partial y_2}{\partial z}\right|_0^1 = Pe_C\left[y_2(t,0) - r y_2(t,1)\right] \tag{D.15a}$$

$$\left.\frac{\partial y_2}{\partial z}\right|_0^1 = Pe_C\left[\sum_{i=1}^{N_2}\beta_i(t)\theta_i(0) + \bar{y}_2(0) - r\left(\sum_{i=1}^{N_2}\beta_i(t)\theta_i(1) + \bar{y}_2(1)\right)\right] \tag{D.15b}$$

$$\left.\frac{\partial y_2}{\partial z}\right|_0^1 = Pe_C\left[\sum_{i=1}^{N_2}\beta_i(t)\left(\theta_i(0) - r\theta_i(1)\right) + \bar{y}_2(0) - r\bar{y}_2(1)\right] \tag{D.15c}$$

$$\begin{aligned}\sum_{i=1}^{N_2}\dot{\beta}_i(t)\int_0^1\theta_j(z)\theta_i(z)dz = -\sum_{i=1}^{N_2}\beta_i(t)\Bigg[&\int_0^1\theta_j(z)\dot{\theta}_i(z)dz \\ &+ \frac{1}{Pe_C}\int_0^1\dot{\theta}_j(z)\dot{\theta}_i(z)dz + B_C\int_0^1\theta_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz\Bigg] \\ &- \int_0^1\left(\theta_j(z) + \frac{1}{Pe_C}\dot{\theta}_j(z)\right)\frac{\partial\bar{y}_2(z)}{\partial z}dz \\ &- B_C\int_0^1\theta_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz \\ &+ \theta_j(0)\left[\sum_{i=1}^{N_2}\beta_i(t)\left(\theta_i(0) - r\theta_i(1)\right) + \bar{y}_2(0) - r\bar{y}_2(1)\right]\end{aligned} \tag{D.16}$$

$$\begin{aligned}\sum_{i=1}^{N_2}\dot{\beta}_i(t)\int_0^1\theta_j(z)\theta_i(z)dz = -\sum_{i=1}^{N_2}\beta_i(t)\Bigg[&\int_0^1\theta_j(z)\dot{\theta}_i(z)dz + \theta_j(0)\left(\theta_i(0) - r\theta_i(1)\right) \\ &+ \frac{1}{Pe_C}\int_0^1\dot{\theta}_j(z)\dot{\theta}_i(z)dz + B_C\int_0^1\theta_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz\Bigg] \\ &- \int_0^1\left(\theta_j(z) + \frac{1}{Pe_C}\dot{\theta}_j(z)\right)\frac{\partial\bar{y}_2(z)}{\partial z}dz + \theta_j(0)\left[\bar{y}_2(0) - r\bar{y}_2(1)\right] \\ &- B_C\int_0^1\theta_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz\end{aligned} \tag{D.17}$$

$$m_{Cij} = \int_0^1 \theta_j(z)\theta_i(z)dz \tag{D.18a}$$

$$n_{Cij} = \int_0^1 \theta_j(z)\dot{\theta}_i(z)dz + \theta_j(0)\big(\theta_i(0) - r\theta_i(1)\big) + \frac{1}{Pe_C}\int_0^1 \dot{\theta}_j(z)\dot{\theta}_i(z)dz$$
$$+ B_C \int_0^1 \theta_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz \tag{D.18b}$$

$$v_{Cj} = -\int_0^1 \left(\theta_j(z) + \frac{1}{Pe_C}\dot{\theta}_j(z)\right)\frac{\partial \bar{y}_2(z)}{\partial z}dz + \theta_j(0)\big[\bar{y}_2(0) - r\bar{y}_2(1)\big]$$
$$- B_C \int_0^1 \theta_j(z)\big(1+\bar{y}_2(z)\big)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right) \tag{D.18c}$$

$$\boxed{M_C\dot{\beta}(t) = -N_C\beta(t) + V_C} \tag{D.19}$$

For the energy balance,

$$\int_0^1 \phi_j(z)\frac{\partial y_1}{\partial t}dz = -\int_0^1 \phi_j(z)\frac{\partial y_1}{\partial z}dz + \frac{1}{Pe_T}\int_0^1 \phi_j(z)\frac{\partial^2 y_1}{\partial z^2}dz$$
$$+ B_T B_C \int_0^1 \phi_j(z)(1+y_2)\exp\left(\frac{\gamma y_1}{1+y_1}\right)dz + \beta_T\int_0^1 \phi_j(z)y_1 dz \tag{D.20}$$

$$\int_0^1 \phi_j(z)\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\phi_i(z)dz = -\int_0^1 \phi_j(z)\left[\sum_{i=1}^{N_1}\alpha_i(t)\dot{\phi}_i(z) + \frac{\partial \bar{y}_1(z)}{\partial z}\right]dz$$
$$+ \frac{1}{Pe_T}\left[\phi_j(z)\frac{\partial y_1}{\partial z}\bigg|_0^1 - \int_0^1 \dot{\phi}_j(z)\left[\sum_{i=1}^{N_1}\alpha_i(t)\dot{\phi}_i(z) + \frac{\partial \bar{y}_1(z)}{\partial z}\right]dz\right]$$
$$+ B_T B_C \int_0^1 \phi_j(z)\left[1 + \sum_{i=1}^{N_2}\beta_i(t)\theta_i(z) + \bar{y}_2(z)\right]\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)}\right)dz \tag{D.21}$$
$$- \beta_T\int_0^1 \phi_j(z)\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z) + \bar{y}_1(z)\right]dz$$

$$\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\int_0^1\phi_j(z)\phi_i(z)dz=-\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\phi_j(z)\dot{\phi}_i(z)dz-\int_0^1\phi_j(z)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+\frac{1}{Pe_T}\phi_j(z)\frac{\partial y_1}{\partial z}\bigg|_0^1-\frac{1}{Pe_T}\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz-\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+B_TB_C\int_0^1\phi_j(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$+B_TB_C\int_0^1\phi_j(z)\left[\sum_{i=1}^{N_2}\beta_i(t)\theta_i(z)\right]\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \qquad \text{(D.22)}$$
$$+B_TB_C\int_0^1\phi_j(z)\bar{y}_2(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$-\beta_T\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\phi_j(z)\phi_i(z)dz-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz$$

$$\frac{\partial y_1}{\partial z}\bigg|_0^1=Pe_T\left[y_1(t,0)-ry_1(t,1)\right] \qquad \text{(D.23a)}$$

$$\frac{\partial y_1}{\partial z}\bigg|_0^1=Pe_T\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(0)+\bar{y}_1(0)-r\left(\sum_{i=1}^{N_2}\alpha_i(t)\phi_i(1)+\bar{y}_1(1)\right)\right] \qquad \text{(D.23b)}$$

$$\frac{\partial y_1}{\partial z}\bigg|_0^1=Pe_T\left[\sum_{i=1}^{N_1}\alpha_i(t)\left(\phi_i(0)-r\phi_i(1)\right)+\bar{y}_1(0)-r\bar{y}_1(1)\right] \qquad \text{(D.23c)}$$

$$\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\int_0^1\phi_j(z)\phi_i(z)dz=-\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\phi_j(z)\dot{\phi}_i(z)dz$$
$$-\int_0^1\phi_j(z)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+\frac{1}{Pe_T}\phi_j(z)\frac{\partial y_1}{\partial z}\bigg|_0^1$$
$$-\frac{1}{Pe_T}\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz$$
$$-\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+B_TB_C\int_0^1\phi_j(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \qquad \text{(D.24)}$$
$$+B_TB_C\int_0^1\phi_j(z)\left[\sum_{i=1}^{N_2}\beta_i(t)\theta_i(z)\right]\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$+B_TB_C\int_0^1\phi_j(z)\bar{y}_2(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$-\beta_T\sum_{i=1}^{N_1}\alpha_i(t)\int_0^1\phi_j(z)\phi_i(z)dz$$
$$-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz$$

$$\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\int_0^1\phi_j(z)\phi_i(z)dz=-\sum_{i=1}^{N_1}\alpha_i(t)\left[\int_0^1\phi_j(z)\dot{\phi}_i(z)dz+\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz+\beta_T\int_0^1\phi_j(z)\phi_i(z)dz\right]$$
$$-\int_0^1\left(\phi_j(z)+\frac{1}{Pe_T}\dot{\phi}_j(z)\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+B_TB_C\int_0^1\phi_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \qquad \text{(D.25)}$$
$$+B_TB_C\sum_{i=1}^{N_2}\beta_i(t)\int_0^1\phi_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz+\frac{1}{Pe_T}\phi_j(z)\frac{\partial y_1}{\partial z}\Big|_0^1$$

$$\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\int_0^1\phi_j(z)\phi_i(z)dz=-\sum_{i=1}^{N_1}\alpha_i(t)\left[\int_0^1\phi_j(z)\dot{\phi}_i(z)dz+\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz+\beta_T\int_0^1\phi_j(z)\phi_i(z)dz\right]$$
$$-\int_0^1\left(\phi_j(z)+\frac{1}{Pe_T}\dot{\phi}_j(z)\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz$$
$$+B_TB_C\int_0^1\phi_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \qquad \text{(D.26)}$$
$$+B_TB_C\sum_{i=1}^{N_2}\beta_i(t)\int_0^1\phi_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz+\phi_j(0)\left[\sum_{i=1}^{N_1}\alpha_i(t)\left(\phi_i(0)-r\phi_i(1)\right)+\bar{y}_1(0)-r\bar{y}_1(1)\right]$$

$$\sum_{i=1}^{N_1}\dot{\alpha}_i(t)\int_0^1\phi_j(z)\phi_i(z)dz=-\sum_{i=1}^{N_1}\alpha_i(t)\left[\int_0^1\phi_j(z)\dot{\phi}_i(z)dz+\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz+\beta_T\int_0^1\phi_j(z)\phi_i(z)dz\right.$$
$$\left.-\left(\phi_i(0)-\phi_j(0)r\phi_i(1)\right)\right]+B_TB_C\sum_{i=1}^{N_2}\beta_i(t)\int_0^1\phi_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz$$
$$-\int_0^1\left(\phi_j(z)+\frac{1}{Pe_T}\dot{\phi}_j(z)\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz+B_TB_C\int_0^1\phi_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \quad (D.27)$$
$$-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz+\phi_j(0)\left[\bar{y}_1(0)-r\bar{y}_1(1)\right]$$

$$m_{Tij}=\int_0^1\phi_j(z)\phi_i(z)dz \quad (D.28a)$$

$$n_{Tij}=\int_0^1\phi_j(z)\dot{\phi}_i(z)dz+\frac{1}{Pe_T}\int_0^1\dot{\phi}_j(z)\dot{\phi}_i(z)dz+\beta_T\int_0^1\phi_j(z)\phi_i(z)dz-\left(\phi_i(0)-\phi_j(0)r\phi_i(1)\right) \quad (D.28b)$$

$$f_{Tij}=B_TB_C\int_0^1\phi_j(z)\theta_i(z)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \quad (D.28c)$$

$$v_{Tj}=-\int_0^1\left(\phi_j(z)+\frac{1}{Pe_T}\dot{\phi}_j(z)\right)\frac{\partial\bar{y}_1(z)}{\partial z}dz-\beta_T\int_0^1\phi_j(z)\bar{y}_1(z)dz+\phi_j(0)\left[\bar{y}_1(0)-r\bar{y}_1(1)\right]$$
$$+B_TB_C\int_0^1\phi_j(z)\left(1+\bar{y}_2(z)\right)\exp\left(\frac{\gamma\left[\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)\right]}{1+\sum_{i=1}^{N_1}\alpha_i(t)\phi_i(z)+\bar{y}_1(z)}\right)dz \quad (D.28c)$$

$$M_T\dot{\alpha}(t)=-N_T\dot{\alpha}(t)+F_T\beta_i(t)+V_T \quad (D.29)$$

With initial conditions,

$$y_1(z,0)=\sum_{i=1}^{N_1}\alpha_i(0)\phi_i(z)+\bar{y}_1(z)=y_{10} \quad (D.30a)$$

$$\int_0^1 \phi_j(z) y_1(z,0) dz = \int_0^1 \phi_j(z) \sum_{i=1}^{N_1} \alpha_i(0) \phi_i(z) dz = \int_0^1 \phi_j(z) \left(y_{10} - \bar{y}_1(z) \right) dz \tag{D.30b}$$

$$\sum_{i=1}^{N_1} \alpha_i(0) \int_0^1 \phi_j(z) \phi_i(z) dz = \int_0^1 \phi_j(z) \left(y_{10} - \bar{y}_1(z) \right) dz \tag{D.30c}$$

$$\alpha_i(0) = M_T^{-1} \int_0^1 \phi_i(z) \left(y_{10} - \bar{y}_1(z) \right) dz \tag{D.31}$$

$$y_2(z,0) = \sum_{i=1}^{N_2} \beta_i(0) \theta_i(z) + \bar{y}_2(z) = y_{20} \tag{D.32a}$$

$$\int_0^1 \theta_j(z) y_2(z,0) dz = \int_0^1 \theta_j(z) \sum_{i=1}^{N_1} \beta_i(0) \theta_i(z) dz = \int_0^1 \theta_j(z) \left(y_{20} - \bar{y}_2(z) \right) dz \tag{D.32b}$$

$$\sum_{i=1}^{N_1} \beta_i(0) \int_0^1 \theta_j(z) \theta_i(z) dz = \int_0^1 \theta_j(z) \left(y_{20} - \bar{y}_2(z) \right) dz \tag{D.32c}$$

$$\beta_i(0) = M_C^{-1} \int_0^1 \theta_j(z) \left(y_{20} - \bar{y}_2(z) \right) dz \tag{D.33}$$