

Kriging meta-model assisted calibration of computational fluid dynamics models

Olumayowa T. Kajero, Rex B. Thorpe and Tao Chen

Department of Chemical and Process Engineering, University of Surrey, Guildford, GU2 7XH, UK

Bo Wang

Department of Mathematics, University of Leicester, Leicester LE1 7RH, UK

Yuan Yao

Department of Chemical Engineering, National Tsing Hua University, Hsinchu, Taiwan, ROC

Computational fluid dynamics (CFD) is a simulation technique widely used in chemical and process engineering applications. However, computation has become a bottleneck when calibration of CFD models with experimental data (also known as model parameter estimation) is needed. In this research, the kriging meta-modelling approach (also termed Gaussian process) was coupled with expected improvement (EI) to address this challenge. A new EI measure was developed for the sum of squared errors (SSE) which conforms to a generalised chi-square distribution and hence existing normal distribution-based EI measures are not applicable. The new EI measure is to suggest the CFD model parameter to simulate with, hence minimising SSE and improving match between simulation and experiments. The usefulness of the developed method was demonstrated through a case study of a single-phase flow in both a straight-type and a convergent-divergent-type annular jet pump, where a single model parameter was calibrated with experimental data.

Keywords: Calibration, computational fluid dynamics, expected improvement, Gaussian process, parameter estimation

Introduction

Computational fluid dynamics (CFD) simulations are usually large complex computer programs representing real life fluid systems. The complexity of such simulations can make them quite a herculean task to solve, with the run time associated with the simulation being of great concern. Such simulation run time could be for hours or days.^{1,2} In addition, large scale or complex systems usually place a great demand on computer memory.³ Computation is especially a problem when the CFD simulation is required to run for a large number of times, for example when sensitivity analysis is carried out^{1,4}, or the models are used within a certain optimisation problem.^{4,5} Of particular interest in this research project is calibration of CFD models (also known as model parameter estimation), which can be cast into an optimisation problem.

CFD model calibration can be defined as the process of adjusting numerical or physical parameters in the CFD model which, on obtaining the optimal values⁶, helps to improve agreement with experimental data.^{7,8,9} The parameters in this context are typically turbulence model coefficients or turbulence model choice. The default values given by CFD software providers for turbulence parameters may apply only in standard circumstances. In this paper, calibration is formulated as an optimisation problem to minimise the mismatch between simulated and experimental outputs (responses). The calibration process may need tens to hundreds of simulation runs to evaluate the objective function, the sum of squared errors (where the errors quantify the differences between the simulated and the experimental outputs) over the region of interest, a requirement that is often infeasible for CFD.

These difficulties faced by CFD simulations have led to the application of meta-modelling, which is a further abstraction of the simulation model. The basic concept is to treat the simulation as “computer experiments”, from which the simulation “data” are used to develop a relatively simple empirical model termed meta-model (also known as surrogate, emulator, model of a model, approximate or auxiliary model or response surface).^{1,4} Then, the fast-to-run meta-model can be used in place of the CFD for analysis and optimisation purposes. Meta-modelling has been widely applied in different fields; Table 1 provides a sample.

However, the application of meta-modelling in chemical process engineering has been less common. Palmer and Realff reported the early use of meta-modelling in chemical process design via ammonia synthesis plant steady state flow sheet simulation.^{31,32} Later, the use of meta-models for real time optimisation was proposed by Gomes et al. with the consideration of a typical alkylation process.³³ Furthermore, Gomes et al. applied the same methodology to a crude distillation unit and solvents units.³⁴ More recently, Hoque developed CFD based multiple linear and neural network meta-models for bio-aerosol transport in indoor environments.³⁵ Coetzee et al. investigated the development of kriging meta-model (also called Gaussian process in the literature with slightly different formulation^{1,5,6}) for the CFD simulation of gas–liquid flow in bubble columns.³⁶ From their findings, the kriging meta-modelling approach is effective in summarising or approximating non-linear data. The usefulness of kriging is discussed later in this article. Li et al. proposed two different types of surrogate models (which are not empirical, but rather mechanistic) to replace the detailed but computationally expensive full-order simulated moving bed chromatography model for optimisation purposes.³⁷ Wang et al. developed a meta-model using Gaussian process regression for fast prediction and uncertainty analysis in a CFD simulated natural gas vapour dispersion process.³⁸

Despite these progresses in meta-model assisted process design and analysis, the topic of calibrating CFD models using meta-model has not been well explored. Hence, the current research seeks to further this line of work to develop a method for a fast calibration of CFD simulation with experimental data. The kriging method is to be used for meta-modelling. The principle of expected improvement (EI), a measure for global optimisation³⁹⁻⁴¹, is utilised to formulate the calibration into an optimisation problem. However, in previous studies, EI was developed assuming a Gaussian process meta-model for the output (response) of the underlying simulation model; this is not applicable to calibration problems in which the objective is to minimise the sum of squared errors. Assuming Gaussian distributed simulation output implies that this sum has a generalised chi-square distribution (as we shall see below equation 15).

In this paper, the usefulness of the developed methodology will be demonstrated by calibrating a CFD model of a straight-type and convergent-divergent-type annular jet pump. Jet pumps have the advantages of simplicity in structure, absence of moving parts, convenience of operation and maintenance, easy to machine and low capital cost. Hence, they are ideal for corrosive environments and can be used to handle poisonous, explosive, flammable or radioactive substance.⁴² They can also be used in de-liquefaction of oil and gas wells, boosting of production in such wells, boosting pressure of low pressure gas in a process system, and can also help to eliminate intermediate compressors.⁴³ The developed meta-modelling methodology can be used in the design and optimisation of jet pumps.

Meta-model based CFD calibration

Meta-model is also referred to as surrogate model or response surface model in the literature.^{44,45} It is a further abstract representation of the CFD simulation model via construction of approximations to the computationally expensive simulation.⁷ It can also be defined as an explicit and simplified model of the underlying complex simulation model, where the simulation model implies an implicit input/output function that maps the simulation parameters onto the output (response).⁴⁵ These surrogate models are for fast prediction, verification and validation, calibration, sensitivity analysis and uncertainty analysis^{31,32,45,46} hence helping with decision making process.³

An overview of the proposed calibration methodology is given in Figure 1.

[Figure 1 about here]

The starting point is usually the CFD simulation results at the default parameter values. The error obtained when the simulation results are compared with the experimental data can be calculated in terms of the sum of squared errors (SSE) which serves as a benchmark for the calibration. Other criteria used in modelling practice include the sum of relative error, mean squared residuals, and average absolute relative error.⁴⁵ The SSE needs to be minimised for a better match of the CFD simulation results with experimental data.

$$SSE = \sum_{i=1}^m (w_i - r_i)^2 \quad (1)$$

where r_i is the i th experimental datum and w_i the simulated value, and m is the number of data points.

It is essential to identify the CFD simulation inputs and outputs which will be used for calibration. The inputs will be the main input variables used in running the simulation and which are most important, and can likewise be included as inputs into the meta-model, while the outputs are the outputs of interest in both real life and computer experiment. The most important input variables can be determined using factor screening which is defined as the search for the important inputs (also called factors) that should be varied in the computer experiment.⁴⁵ In this paper, the inputs were selected based on experience and knowledge of the simulation problem. In both the straight-type and convergent-divergent-type annular jet pump case studies considered, the inputs into the meta-model are turbulent parameter, C_μ , spatial location, X/D_0 , and flow ratio, M ; while the experimental measurement and therefore, the simulation output is pressure coefficient, C_p . These parameters will be defined under the sub-heading “Experimental data of Shimizu et al.⁴⁷”.

The design of experiments (DoE) is then used to determine the values of input variables that should be used for CFD simulation. DoE is a statistical approach used in designing and selection of the inputs that can be utilised for computer simulations. Examples of widely used DoE methods are grid layouts plan, random sampling, Latin hypercube sampling^{48,49,50}, Hammersley sequence sampling^{1,46,51}, and uniform design^{52,53}. The key feature which makes DoE useful is that it is used to sample the design space (region of interest bounded by the upper and lower limits of the design input variables being studied), hence generating sample data to fit the meta-model to each of the response variables of interest.⁴⁶

The meta-model is then developed from the CFD simulation data, and the SSE (or equivalently the root mean square error, RMSE) is calculated to check the accuracy of the meta-model against CFD simulation. This should be carried out by using the leave-one-out cross validation (LOOCV) method. The LOOCV is a statistical technique which involves taking a single run from the entire dataset which is used as the validation data (test data), while the remaining runs (training data) are used to build the meta-model. The procedure is repeated such that each run is used once for validation, and the overall

validation error is evaluated. Subsequently, the meta-model is used to predict the SSE at untried combination of values (or points in the input space) of model parameters and the expected improvement (EI) can be obtained. The EI is needed because the meta-model prediction is probabilistic, i.e. it gives prediction mean and variance.⁵⁵ By using the prediction mean only, the optimisation problem can be stuck in some local optimum.⁵⁵ If the maximised EI is less than a certain threshold, then the iteration stops, otherwise, the process is repeated from simulations at the new design points. An alternative stopping criterion could be reaching maximum computational resources (e.g. maximum number of iterations).

The rest of this section gives an overview of CFD simulation, kriging meta-model, and maximisation of expected improvement.

Computational fluid dynamics (CFD) simulation

The governing equations of fluid flow are a representation of the mathematical statements of the conservation laws of physics. There are different turbulence models, but in this project, we have focused on the k - ε models. The equations governing the fluid flow are: the continuity equation, the momentum equation, and equations for turbulent kinetic energy and turbulent dissipation rate. The default model parameters for three k - ε turbulence models are given in Table 2. RNG in the table refers to Renormalization Group. The standard k - ε model was chosen for the calibration methodology because it is reasonably accurate for a wide range of applications.⁵⁶ ANSYS Fluent CFD software was used in this research. Its solvers are based on the finite volume method.

[Table 2 about here]

According to Launder and Spalding⁵⁸, the turbulent parameters, C_μ and $C_{2\varepsilon}$, can be adjusted for the case of axisymmetric jets as no single best value can be found. However, carrying out calibration via CFD directly involves a lot of simulation runs which are time consuming and expensive. This perhaps is the reason why little has been done to adjust these parameters for practical applications. In this work involving the use of jet pump as case study, meta-modelling assisted calibration of CFD simulation was carried out by adjusting the turbulent parameter, C_μ . In the preliminary study, we attempted

calibration for both C_μ and $C_{2\varepsilon}$; however the results suggested that adjusting C_μ only can provide satisfactory results for this particular application, and adjusting both parameters did not give improved results.

The kriging meta-model

The meta-model of choice here is kriging, which is also termed Gaussian process in the literature with slightly different formulation.^{1,5,6,45,54} It approximates the output of a computer model as a random Gaussian process. It is probabilistic giving both an expected value and a variance in the prediction.^{54,59} Kriging obtains the prediction at unobserved input values by assigning higher weights to simulation output data, the more similar the corresponding input combinations are.^{45,59} It is an excellent way of interpolation compared with other meta-modelling techniques such as polynomial, regression splines, radial basis functions and artificial neural network.^{1,31,32,33,44,60,61,62,63}

Let $\mathbf{x}_i, i = 1, 2, \dots, m$ be a set of m training data points (also known as “design sites” in kriging) for developing the kriging model, each of which is a vector of input variables with dimension n ($\mathbf{x}_i = [x_{i1}, x_{i2} \dots, x_{in}]^T$), and $\mathbf{y} = [y_1, y_2 \dots, y_m]^T$ be the corresponding response variables. Then, the prediction for a new data point, \mathbf{x}_* , in the kriging model is given by

$$y_*(\mathbf{x}_*) = \mathbf{f}^T(\mathbf{x}_*)\boldsymbol{\beta} + \mathbf{r}^T(\mathbf{x}_*)\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}) \quad (2)$$

which is a general type of kriging called universal kriging^{33,34,45,54,64}. Here $\mathbf{f}(\mathbf{x}_*)$ contains a set of regression functions of the input variables as determined by the modeller, and $\boldsymbol{\beta}$ is the corresponding regression coefficients to be estimated. Usually, \mathbf{f} contains polynomials of up to second order.⁶⁴ For example, if only the zero-th order polynomial is used, \mathbf{f} reduces to a scalar function with fixed value of unity: $f(\mathbf{x}_*) = 1$. If additional first order polynomial is included, then $\mathbf{f}(\mathbf{x}_*) = [1, x_{*1}, x_{*2}, \dots, x_{*n}]^T$ and so on. $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1), \mathbf{f}(\mathbf{x}_2), \dots, \mathbf{f}(\mathbf{x}_m)]^T$ is a matrix containing the regression functions calculated for all the m training data points. In this work, only the zero-th order polynomial is used which is termed “ordinary kriging” in the literature.^{31,44,45,50,54,59,64,65}

\mathbf{R} is the correlation matrix which is obtained from correlation functions evaluated at each pair of the

training points:

$$\mathbf{R} = \begin{bmatrix} R(\mathbf{x}_1, \mathbf{x}_1) & R(\mathbf{x}_1, \mathbf{x}_2) & \cdots & R(\mathbf{x}_1, \mathbf{x}_m) \\ R(\mathbf{x}_2, \mathbf{x}_1) & R(\mathbf{x}_2, \mathbf{x}_2) & \cdots & R(\mathbf{x}_2, \mathbf{x}_m) \\ \vdots & \vdots & \ddots & \vdots \\ R(\mathbf{x}_m, \mathbf{x}_1) & R(\mathbf{x}_m, \mathbf{x}_2) & \cdots & R(\mathbf{x}_m, \mathbf{x}_m) \end{bmatrix} \quad (3)$$

where the correlation function can be parameterised in various ways.^{45,64} A widely used specification of R is the Gaussian function^{40,44,65}:

$$R(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\sum_{k=1}^n \theta_k (x_{ik} - x_{jk})^2\right) \quad (4)$$

The above correlation function is appropriate for deterministic simulations. If stochastic simulations are of interest, a random noise term needs to be added to the correlation function.⁴⁵

\mathbf{r} is a vector representing the correlation between the new point, \mathbf{x}_* , and the training set:

$$\mathbf{r}(\mathbf{x}_*) = [R(\mathbf{x}_*, \mathbf{x}_1), R(\mathbf{x}_*, \mathbf{x}_2), \dots, R(\mathbf{x}_*, \mathbf{x}_m)]^T \quad (5)$$

The parameters in kriging include those in the correlation function ($\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_n]^T$), the process variance (σ_p^2), and the regression coefficients ($\boldsymbol{\beta}$). First, $\boldsymbol{\beta}$ is estimated by using the maximum likelihood method⁶⁶, giving rise to

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y} \quad (6)$$

The process variance is then obtained as

$$\hat{\sigma}_p^2 = \frac{1}{m} (\mathbf{y} - \mathbf{F} \hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F} \hat{\boldsymbol{\beta}}) \quad (7)$$

Lastly, the correlation parameter can be estimated by solving the following optimisation problem:

$$\min_{\boldsymbol{\theta}} (|\mathbf{R}|^{1/m} \hat{\sigma}_p^2) \quad (8)$$

The estimated parameters are then plugged into the kriging model in equation (2) to obtain prediction mean (denoted μ_*). The corresponding prediction variance, σ_*^2 is obtained as the estimated mean square error of the predictor^{44,45,64,65} which is given as:

$$\sigma_*^2 = \hat{\sigma}_p^2 (1 + \mathbf{u}^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{u} - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}) \quad (9)$$

where

$$\mathbf{u} = \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r} - \mathbf{f}. \quad (10)$$

Equation (9) is only an approximation because it ignores the uncertainty of the estimated kriging parameters.⁴⁵

Mathematical details of kriging and its construction can be found in Sacks et al.⁵⁴ and other meta-modelling studies.^{19,31,36,44,45,59,64,65}

Maximising expected improvement

Expected improvement (EI) indicates which input combination to simulate next in order to have a better match of the experimental data with the CFD model. It is the computation of the extent of improvement when a given point is sampled. It is a widely used statistical measure which can help to decide the subsequent function evaluations in global optimisation.^{39,40,41} Here “global optimisation” refers to that (i) first, the method guarantees to find the global optimum when the number of iterations tends to infinity⁵⁵; and (ii) the search is not limited to a local region of the current point. Although in practice, the number of iterations is always finite, the capability of searching globally gives rise to higher probability of finding the global optimum, when compared with the traditional gradient-based optimisation methods.^{39,55,67}

The EI for SSE needs to be derived, because the objective functions in existing studies have been assumed to be normally distributed^{39,55}; however the SSE is not normally distributed (clearly SSE must be non-negative whilst the normal distribution is defined on the entire real axis⁶⁸). This will be discussed in the rest of the section.

To minimise an objective function, EI is defined as:

$$\text{EI} = \int_0^\infty I p(I) dI \quad (11)$$

where $I = z_{\min} - z$, a continuous random variable which is the improvement, z is the objective function to be minimised and z_{\min} is the best z that has been achieved so far, and $p(I)$ is a probability

density function.

The meta-model is to be used for calibration, and the variable to be minimised is the sum of squared errors, SSE. So, it might be possible to build a meta-model between calibration parameters (e.g. turbulent model parameter, C_μ) and SSE, but such a meta-model loses the capability of predicting the physical measurements of interest (e.g. pressure coefficients, concentration field, etc.) due to its empirical nature and the inability to establish and describe the physical or chemical process of the considered system.⁵ We choose to build meta-model to predict the measurement, and calculate SSE according to equation (1).

Let the minimum SSE achieved so far be SSE_{\min} (which is a constant), then the improvement becomes $I = SSE_{\min} - SSE$. Since SSE must be non-negative, the greatest improvement can only be SSE_{\min} . From equation (11), EI becomes:

$$EI = \int_0^{SSE_{\min}} [SSE_{\min} - SSE] p(SSE_{\min} - SSE) d(SSE_{\min} - SSE) \quad (12)$$

which is equivalent to:

$$EI = \int_0^{SSE_{\min}} [SSE_{\min} - SSE] p(SSE) dSSE \quad (13)$$

Equation (13) can be further expressed as:

$$EI = SSE_{\min} \int_0^{SSE_{\min}} p(SSE) dSSE - \int_0^{SSE_{\min}} SSE p(SSE) dSSE \quad (14)$$

The integral part of the first expression in equation (14) can be related to the cumulative distribution function (*cdf*) which describes the probability that SSE takes on a value less than SSE_{\min} . Hence, we have:

$$EI = SSE_{\min} \text{Prob}(SSE < SSE_{\min}) - \int_0^{SSE_{\min}} SSE p(SSE) dSSE. \quad (15)$$

The distribution of SSE can be derived from the fact that the kriging prediction is normally distributed, ignoring the fact that the kriging parameters are unknown and must be estimated so the predictor becomes non-linear.⁴⁵ As such, the SSE conforms to a generalised chi-square distribution.⁶⁹ This can be seen as follows.

By definition, the sum of squares of independent, standard normal random variables conforms to a chi-square distribution. In the context of model calibration, suppose there are m experimental data points with inputs $\mathbf{x}_i, i = 1, 2, \dots, m$, against which the kriging meta-model predictions are compared to calculate SSE. Recall that the prediction from kriging model for data i , denoted y_i , is normally distributed with prediction mean being μ_i and variance $\sigma_i^2, i = 1, \dots, m$, i.e. $y_i \sim N(\mu_i, \sigma_i^2)$. When calculating SSE, the simulated value in equation (1) is replaced by the meta-model prediction y_i and therefore $\text{SSE} = \sum_{i=1}^m (y_i - r_i)^2$. This SSE is not chi-square distributed because $(y_i - r_i) \sim N(\mu_i - r_i, \sigma_i^2)$ is not a standard normal random variable (its mean is not necessarily zero and variance is not unity).

Fortunately, the chi-square distribution can be extended to its *generalised* form: if a multivariate random variable \mathbf{z} has a multivariate normal distribution with a fixed mean vector \mathbf{a} and covariance matrix $\mathbf{\Sigma}$, then the quadratic form $\mathbf{z}^T \mathbf{z}$ has a *generalised* chi-square distribution.⁶⁹ Using the vector-matrix form, it can be seen that $\text{SSE} = (\mathbf{y} - \mathbf{x})^T (\mathbf{y} - \mathbf{x})$ is such a quadratic form if we let $\mathbf{z} = \mathbf{y} - \mathbf{x}$, $\mathbf{a} = \boldsymbol{\mu} - \mathbf{x}$, and $\mathbf{\Sigma} = \text{diag}([\sigma_1^2, \dots, \sigma_m^2])$, where $\mathbf{y} = [y_1, \dots, y_m]^T$, and $\boldsymbol{\mu} = [\mu_1, \dots, \mu_m]^T$. Note that when making predictions at m new data points using kriging, the predictions are correlated and thus to be rigorous $\mathbf{\Sigma}$ is not a diagonal matrix. We choose to ignore this correlation in order to simplify the calculations.

There is no analytical form of this distribution function, but a numerical procedure exists to calculate the following cumulative distribution function (*cdf*): $\text{Prob}(\text{SSE} < \text{SSE}_{\min})$.⁶⁹ This numerical procedure involves the numerical computation of multivariate normal distribution based on ellipsoidal sets. It computes the multivariate normal value to a given relative accuracy for an ellipsoid at a given centre and radius, with a positive definite matrix. Its mathematical basis can be found in the book by Genz and Bretz⁷⁰ while its implementation is given by Sheil and O’Muircheartaigh.⁶⁹ With this *cdf*, the first term in the EI expression in equation (15) can be directly calculated; the second term can be calculated through numerical integration.

The maximum value of EI was obtained by using the interior point optimisation method combined with the multi-start method, i.e. the optimisation algorithm was randomly initialised multiple times (10

times in this study) and the best solution of these multiple optimisation runs is used. The input variables corresponding to this maximum indicate the next point where simulation will be carried out. The interior point has been extensively discussed in the literature.^{71,72,73} The optimisation process eventually results in a new set of simulation data. The meta-model is then updated using the new simulation data, with the process repeated until EI is less than a specified threshold in which there is no further improvement. Being a probabilistic measure, EI does not guarantee monotonic decrease of SSE in each iteration; rather it intends to provide a search direction towards which SSE is likely to improve.³⁹ When the entire process is completed, we choose the values of the calibrated parameters corresponding to the minimum SSE, since this gives the best match between the simulation and experimental data.

Case study: CFD simulation of annular jet pump

The case studies used to illustrate the novel methodology are for the turbulent single-phase flow in a straight-type and in a convergent-divergent-type annular jet pump. Generally speaking, the jet pump transfers energy from a liquid or gas primary fluid to a secondary fluid. Momentum is transferred from the high velocity primary jet flow to the secondary flow of lower velocity.^{42,74} In this case study, both the primary and secondary fluid are water. The primary fluid passes through a nozzle where pressure energy is converted into kinetic energy. The high kinetic primary fluid mixes with low kinetic energy secondary fluid, thus drawing the low kinetic energy to a higher pressure. Recirculation can also be observed. The observed recirculation can be explained based on the concept of entrainment. The jet flows through the mixing region and expands in diameter due to the entrainment of the secondary flow. Recirculation comes into play when all the secondary flow has been entrained before the expansion of the jet to reach the walls of the mixing region.⁷⁵

In the convergent-divergent-type annular jet pump, converging and diverging parts (which are the suction chamber and diffuser region respectively) differentiate it from the straight-type. In the converging part, there is a decrease in flow pressure, and mixing and recirculation commences here. In the diverging part, there is more efficient pressure recovery than in the straight-type. Schematics of the jet pumps used, with the dimensions are shown in Figure 2 corresponding to the jet pumps studied by

Shimizu et al.⁴⁷ who also reported experimental measurements as outlined in the next section. The geometries are symmetrical along the x-axis (axisymmetric).

[Figure 2 about here]

Experimental data of Shimizu et al.⁴⁷

Wall pressures in the suction pipe (secondary flow pipe or throat), annular nozzle and delivery pipe (outlet) were experimentally taken from the respective pressure tapping and were measured by the mercury manometers. These pressure tapings were placed at different locations on the jet wall based on a dimensionless spatial location or axial distance, X/D_o , where X is the distance measured from the nozzle exit and D_o is the throat diameter. The obtained pressures at these points were subsequently used to compute the pressure coefficient given as:

$$C_p = \frac{P - P_o}{\left(\frac{\rho U_j^2}{2}\right)} \quad (16)$$

where P is static pressure at the wall, P_o is static pressure at the secondary flow stream, ρ is the density of the flow medium and U_j is the jet flow velocity.

In the straight-type annular jet pump case, four different flow ratios were used, where the flow ratio, M is the ratio of secondary flow rate to primary flow rate. The flow ratios are 0.01, 0.11, 0.19 and 0.34. In the convergent-divergent-type, three flow ratios were used, which are 0.04, 0.3 and 0.58. The flow ratios were used to investigate their effect on the flow field in the two geometries. More details of the experiments can be found in Shimizu et al.⁴⁷, from where the experimental data were extracted from their graphs for use in the present study.

CFD simulation

Prior to CFD simulation, a suitable DoE is essential to determine the value of input parameters. Since in this case, there is only one parameter to be adjusted for calibration (C_μ), we can simply choose a lower bound ($C_\mu = 0.05$) and upper bound ($C_\mu = 0.2$) with an intermediate value ($C_\mu = 0.09$). All these points (lower, intermediate and upper bound) constitute the initial CFD simulation. The turbulent parameter, C_μ , spatial location, X/Do and flow ratio, M constitute the meta-model inputs.

The CFD simulation was done using the ANSYS Workbench version 14.5 which comprises of the ANSYS Design-modeler, ANSYS Meshing and ANSYS Fluent. The geometries used are shown in Figure 2 and 3. The inlet boundary for both the primary flow and secondary flow inlet were defined as mass flow inlet, outlet boundary as pressure outlet, the duct wall as wall, and the centreline as axis corresponding to a 2D axisymmetric case. The no-slip boundary condition was defined for the wall. The meshing was done using ANSYS Meshing in the workbench. For the straight-type, three computational grids with grid numbers of approximately 32,000, 58,000 and 113,000 (taken as coarse, medium and fine grid respectively) were used to perform mesh or grid independence tests. Likewise for the convergent-divergent-type, three computational grids with grid numbers of approximately 173,000, 367,000 and 629,000 were used as the coarse, medium and fine grid respectively. In both cases, the results were grid independent.

The fluid flow in the annular jet pump was assumed to be steady and incompressible under the influence of Reynolds Averaged Navier Stokes Equations and continuity equations. The pressure and mean velocity fields were coupled by the SIMPLE algorithm.⁵⁶ The $k-\varepsilon$ turbulence model was used and a comparison was made between three classes of the model such as the standard $k-\varepsilon$, the realizable $k-\varepsilon$ and Renormalization Group (RNG) $k-\varepsilon$ with the standard $k-\varepsilon$ chosen for the meta-modelling methodology.

Figure 3 and 4 present the simulation results for the straight-type and convergent-divergent pumps respectively using the default C_μ . A reasonably good fit with experimental data was obtained within an

axial distance, X/D_o of 2 and 7. But between X/D_o of 0 and 2, a good agreement was not obtained especially for $M = 0.01$ and $M = 0.34$, hence the need for model calibration to give a better match between the simulation model and experimental data. The need for calibration is also clearly seen from the plot of CFD simulation versus experimental data as shown in Figure 5a in which it is clear that some points do not lie on the 45° line.

[Figure 3 about here]

It can be seen in Figure 3 that from the jet exit to the mixing region, the wall pressure increases, and then flattens out as it moves towards the outlet of the annular jet pump. This trend was observed for all the flow ratios. The low pressure obtained within an axial distance, X/D_o of 0 and 2 could be due to the formation of recirculation zone, the jet expansion and secondary flow mixing zone that partly overlaps with it.

For the convergent-divergent-type, the results from the standard $k-\varepsilon$ simulation are shown in Figure 4 and 5b. The need for calibration can be clearly seen with some data points not lying on the 45° line in Figure 5b. The pattern of the pressure coefficient was seen to vary with flow ratio more than for the straight case (see Figure 4).

[Figure 4 about here]

[Figure 5 about here]

Meta-models developed from initial CFD data

The 12 initial CFD runs (three C_μ values combined with four flow ratios) for the straight-type jet pump were used to develop a kriging meta-model. Likewise, a kriging meta-model for the convergent-divergent jet pump was developed from the 9 CFD runs (three C_μ values combined with three flow ratios). The LOOCV results, by taking one run out for validation, while the remaining runs (training data) are used to build the kriging meta-model are presented in Figure 6. The LOOCV results show excellent accuracy for straight-type (Figure 6a). The results for convergent-divergent-type were reasonably good as shown in Figure 6b, though that of the straight-type is far better. The reason for

this is probably due to the difference in the observed trend between the plot of pressure coefficient versus X/D_o for both the straight-type and the convergent-divergent-type. That of straight-type in Figure 3 gave similar trend, while there is an obvious variability in the plot trend of convergent-divergent-type (Figure 4) as the flow ratio, M changes. This variability seems to be challenging to the kriging meta-model.

[Figure 6 about here]

Results and discussion

Straight-type annular jet pump

The calibration results for C_μ independent on flow ratio and C_μ dependent on flow ratio have been presented. C_μ independent on flow ratio implies the C_μ does not vary with flow ratio. So, all the flow ratios in the straight-type case study were combined to give a single C_μ for the purpose of calibration. In contrast, C_μ dependent on flow ratio implies the C_μ varies with flow ratio. So, calibration was done for each flow ratio, which eventually results into different C_μ for different flow ratio.

The calibration results obtained were based on the procedure described previously which involved the calculation of EI using SSE. Figure 7a shows the maximum EI which suggests a value of turbulent parameter, $C_\mu = 0.095$. This optimum value of C_μ is suggested for all the four flow ratios combined together i.e. $M = 0.01, 0.11, 0.19$ and 0.34 . It corresponds to the second iteration where each iteration is equivalent to four simulations (from four flow ratios). It is important to mention that the maximum EI suggests the C_μ value of 0.095 and the corresponding SSE (or root mean squared error, RMSE) is then calculated, which gives the minimum SSE (or minimum RMSE). SSE leads to the generalised chi-square distribution, whereas the RMSE does not have a known distribution; therefore SSE is used to find the maximum EI. The RMSE, however, is used for ease of interpretation of the results for the case study; it has the same units (scale) as the simulated and experimental responses and has been used in the literature as a standard measure for model performance.

A plot of all the RMSE versus number of CFD simulations obtained for each calibration simulation iterations is shown in Figure 7b, which clearly depicts the number of CFD simulations at respective RMSE values and subsequently at minimum RMSE which gives $C_\mu = 0.095$. This minimum RMSE obtained was 0.0193 which was previously 0.0199 from the initial simulation. This could probably be due to the fact that four flow ratios were combined, which resulted into an optimum C_μ value of 0.095 not far from that of the initial simulation, 0.09. The remaining points in Figure 7b indicating other numbers of CFD simulations are for confirmatory studies. In essence, we could have stopped at the minimum RMSE, but the remaining points indicate no further improvement is obtained. Figure 7b also shows that one iteration is equal to four simulations. A plot of CFD simulation (default and calibrated model) versus experimental data is given in Figure 7c. The plot shows that at RMSE = 0.0196, the data points lie slightly closer to the 45° line as compared with RMSE = 0.0199 which is for the default CFD simulation. However, the improvement is not seen in terms of the coefficient of determination, R^2 (0.948 and 0.950 for default CFD model and calibrated model respectively).

[Figure 7 about here]

The flow ratios were also considered individually in terms of the calibration results which gives different values of C_μ which are 0.1385, 0.0935, 0.083 and 0.056 for $M = 0.01, 0.11, 0.19$, and 0.34 respectively. The SSE and hence, the root mean square error (RMSE) for each of these C_μ values was minimised, hence a better match with experimental data. This is seen in Table 3. The RMSE values were taken to three significant figures.

[Table 3 about here]

The relationship between the use of a single turbulent model parameter C_μ and for C_μ values dependent on flow ratio is shown in Figure 8. The best match of the CFD model with experimental data was given by C_μ values dependent on flow ratio. This is indicated by the root mean square errors. It shows that for better calibration, C_μ has to be dependent on flow ratio.

[Figure 8 about here]

Comparison between the proposed meta-model-assisted method and the traditional calibration method

for CFD simulation has also been attempted. In principle, the calibration problem can be formulated to minimise SSE, but without meta-modelling the optimisation problem needs to be solved directly with the optimiser (e.g. a gradient-based method) being able to run the CFD simulation automatically. However, to the best of the author's knowledge (and from a survey of the literature), there is no ready-to-use tool that can directly calibrate turbulent parameters in CFD. Hence, we used the "trial-and-error" approach as a benchmark for comparison. It is worth noting that this trial-and-error approach is still widely used in the CFD and other simulation communities for model calibration purposes. Specifically, the simulation started from the default value for C_μ (0.09), and based upon the mismatch between simulation and experimental data, the modeller (in this case the first author) decided the next value of C_μ that should be tried. This iteration goes on until the same (within 1% difference) RMSE is achieved as the meta-model-assisted approach for a fair comparison. Relying on the modeller's experience, the comparison is subjective; however it does provide a good demonstration (at least qualitatively) of how the meta-modelling approach compares with the traditional approach.

The results for the straight-type annular jet pump is given in Table 4 (the case for flow ratio-dependent C_μ). The comparison gives an indication of how much computational time could have been saved using the approach proposed in this work. The traditional and the proposed meta-modelling approaches utilised 29 and 12 simulation runs, respectively, to arrive at the same minimum RMSE of 0.0104 which was originally 0.0199.

[Table 4 about here]

Convergent-divergent-type annular jet pump

The same methodology was applied to the convergent-divergent annular jet pump case study. Calibration was done for C_μ independent on flow ratio, where all the flow ratios, $M = 0.04, 0.3$ and 0.58 were combined to give a single C_μ value and also for C_μ dependent on flow ratio, where calibration was done for individual flow ratios. For the combined case, the default C_μ value, 0.09 gave the best match with experimental data, with minimum RMSE. This corresponds to the first iteration, where one iteration is equivalent to three simulations. This can be seen in Figure 9.

[Figure 9 about here]

For C_μ dependent on flow ratio, the sum of squared errors, and hence the root mean square error was minimised to give better agreement with experimental data. This is seen in Table 5.

[Table 5 about here]

C_μ dependent on flow ratio gave better match with experimental data as compared with a single C_μ . This can be seen in Figure 10 where RMSE reduces from 0.0993 to 0.0835.

[Figure 10 about here]

The comparison between the traditional CFD calibration method and the proposed approach for the convergent-divergent annular jet pump is given in Table 6. Similar to the straight-type jet pump, the proposed meta-model-assisted approach could significantly save the computation needed for calibration: it used 14 simulation runs (whilst the trial-and-error approach used 25) to arrive at the same minimum RMSE of 0.0835.

[Table 6 about here]

The relationship between C_μ and flow ratio for both case studies considered was obtained as shown in Figure 11. It can be seen that the straight-type and convergent-divergent-type both gave exponential correlation, with the relationship between C_μ and flow ratio for straight-type being:

$$C_\mu = 0.1354e^{-2.646M} \quad (17)$$

while that of convergent-divergent-type is:

$$C_\mu = 0.1412e^{-1.677M} \quad (18)$$

These two equations (17) and (18) can be used to estimate likely values of C_μ for other flow ratios which could give good simulation predictions of the measured parameter (in this case the pressure coefficient).

[Figure 11 about here]

Conclusion

From the foregoing, it can be seen that in the case studies of the straight-type and convergent-divergent annular jet pump considered, involving one calibration parameter (C_μ), the objective of the research was achieved. The expected improvement which was expressed as a function of the sum of squared errors (the objective function to be minimised) helped to suggest values of the calibration parameters (C_μ) to simulate with, thereby reducing the sum of squared errors when compared with the default CFD simulation, hence, a better match with experimental data. In addition, the simulation time using the proposed approach was reduced as compared to using traditional CFD calibration approach. This confirms that the proposed novel methodology of maximising expected improvement to minimise the sum of squared errors, for the purpose of calibrating CFD simulations worked. It is a potential contribution to CFD simulation studies in different fields of study, as it helps to provide a supportive hand in cases where calibration is required.

It was also observed that in both cases of the annular jet pump, it is better to carry out calibration on the basis that C_μ is dependent on flow ratio rather than trying to use a single C_μ value for all the flow ratios. The proposed relationship between C_μ and flow ratio can also be used to estimate likely values of C_μ for required flow ratios. As only one parameter was considered for calibration in the case studies, future study can try to address two or more parameters. For instance, the adjustment of both C_μ and $C_{2\varepsilon}$. We also plan to investigate the impact of data quality and model-reality mismatch on the calibration results.

Acknowledgment

O. Kajero's PhD is funded by the Tertiary Education Trust Fund (TETFUND), Nigeria in collaboration with University of Lagos, Nigeria. This work was also partially supported by an International Exchange grant co-funded by the UK Royal Society (Grant number: IE140859) and the Taiwan Ministry of Science & Technology (Grant number: MOST 104-2911-I-007-516).

Literature Cited

1. Chen T, Hadinoto K, Yan W, Ma Y. Efficient Meta-modelling of complex process simulation with time-space-dependent outputs, *Computers and Chemical Engineering*. 2011;35:502–509.

2. Keating EH, Doherty J, Vrugt JA, Kang OJ. Optimization and uncertainty assessment of strongly non-linear groundwater models with high parameter dimensionality. *Water Resources Research*. 2010;46.
3. Popplewell K, Yu B. Metamodels in manufacturing: a review. *International Journal of Production Research*. 1994;32:787–796.
4. Tolson A, Burn DH. Numerical Assessment of meta-modelling strategies in computationally intensive optimization. *Environmental Modelling & Software*. 2012;34:67–86.
5. Manfren M, Aste N, Moshksar R. Calibration and uncertainty analysis for computer models – A meta-model based approach for integrated building energy simulation. *Applied Energy*. 2013;103:627–641.
6. Campbell K. Statistical calibration of computer simulations. *Reliability Engineering and System Safety*. 2006;91:1358–1363.
7. AIAA. Guide for the verification and validation of computational fluid dynamics simulations. American Institute of Aeronautics and Astronautics. 1998.
8. Kennedy MC, O'Hagan A. Bayesian calibration of computer models, *Journal of Royal Statistical Society B*. 2001; 63: 425-464.
9. Ierardi JA, Barnett JR. A quantitative method for calibrating CFD Model Calculations. *Proceedings of the CIB-CTBUH International Conference on Tall Buildings*. 2003:20–23.
10. Chang KJ, Haftka RT, Giles GL, Kao PJ. Sensitivity-based Scaling for Approximating Structural Response. *Journal of Aircraft*. 1993;30:283-287.
11. Liu D, Zhou X, Toropov V. Metamodels for composite lattice fuselage design. *International Journal of Materials, Mechanics and Manufacturing*. 2016;4:175-178.
12. Hutchison MG, Unger ER, Mason WM, Grossman B, Haftka RT. Variable Complexity Aerodynamic Optimization of a High-Speed Civil Transport Wing. *Journal of Aircraft*. 1994;31:110-120.
13. Safari A, Younis A, Wang G, Lemu H, Dong Z. Development of metamodel assisted sampling approach to aerodynamic shape optimization problems. *Journal of Mechanical Science and Technology*. 2015;29:2013-2024.

14. Khalfallah S, Ghenaiet A. Radial basis function-based shape optimization of centrifugal impeller using sequential sampling. *Journal of Aerospace Engineering*. 2015;229:648-665.
15. Ghani AO, Agelin-chaab M, Barari M. Application framework for aero-based design optimization of passenger cars using NURBS. *Journal of Applied Mechanical Engineering*. 2015;4:188.
16. Madsen JJ, Langthjem M. Multifidelity Response Surface Approximations for the Optimum Design of Diffuser Flows. *Optimization and Engineering*. 2001;2:453-468.
17. Wang W, Yuan S, Pei J, Zhang J. Optimization of a diffuser in a centrifugal pump by combining response surface method with multi-island genetic algorithm. *Journal of Process Mechanical Engineering*. 2015;0:1-11.
18. Simpson TW, Peplinski JD, Koch PN, Allen JK. Meta-models for computer-based engineering design: Survey and recommendations. *Engineering with Computers*. 2001;17:129-150.
19. Simpson TW, Mauery TM, Korte JJ, Mistree F. Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA Journal*. 2001;39:2233-2241.
20. Glaz B, Friedmann PP, Liu L. Surrogate based optimization of helicopter rotor blades for vibration reduction in forward flight. *Structural and multidisciplinary optimization*. 2007;35:341-363.
21. Broad DR, Dandy GC, Maier HR. Water distribution system optimisation using meta-models, *Journal of Water Resources Planning and Management*. 2005;3:172–180.
22. Zou R, Lung WS, Wu J. An adaptive neural network embedded genetic algorithm approach for inverse water quality modelling. *Water Resources Research*. 2007;43:1-13.
23. Wu W, Dandy GC, Maier HR. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. *Environmental Modelling and Software*. 2014;54:108-127.
24. Fen CS, Chan CC, Cheng HC. Assessing a response surface-based optimization approach for soil vapor extraction system design. *Journal of Water Resources Planning and Management*. 2009;135:198-207.
25. Khu ST, Savic D, Liu Y, Madsen H. A fast evolutionary-based meta-modelling approach for the calibration of a rainfall-runoff model. In. *The first Biennial meeting of the International Environmental Modelling and Software society Osnabruck. Germany, 2004;147–152.*

26. Broad DR, Dandy GC, Maier HR. A systematic approach to determining meta-model scope for risk-based optimization and its application to water distribution system design. *Environmental Modelling and Software*. 2015;69:382-395.
27. Khu ST, Werner MGF. Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modelling. *Hydrology and Earth System Sciences*. 2003;7:680–692.
28. Yadzi J, Salehi Neyshabouri SAA. Adaptive surrogate modelling for optimization of flood control detention dams. *Environmental Modelling and Software*. 2014;61:106-120.
29. Goswami S, Ghosh S, Chakraborty S. Reliability analysis of structures by iterative improved response surface method. *Structural Safety*. 2016;60:56-66.
30. Beck J, Friedrich D, Brandani S, Fraga ES. Multi-objective optimization using surrogate models for the design of VPSA systems. *Computers and Chemical Engineering*. 2015;82:318-329.
31. Palmer K, Realff M. Meta-modeling Approach to Optimization of Steady-state flowsheet Simulation: Model Generation. Institution of Chemical Engineers. *Trans IChemE*. 2002a;80:760–772.
32. Palmer K, Realff M. Optimization and Validation of Steady-state Flowsheet Simulation Meta-models. Institution of Chemical Engineers. *Trans IChemE*. 2002b;80:773–782.
33. Gomes MVC, Bogle IDL, Oldloak D, Biscaia Jr EC. An application of meta-models for Process Optimization. 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering. 2006;1449–1454.
34. Gomes MVC, Bogle DL, Biscaia Jr EC, Oldloak D. Using Kriging models for real-time process optimisation. 18th European Symposium on Computer Aided Process Engineering. ESCAPE 18. 2008.
35. Hoque S. Development of Computational Fluid Dynamics based Multiple Linear and Neural Network Meta-models for Bioaerosol Fate and Transport in Indoor Environments. Ph.D. Thesis. Drexel University U.S.A, 2010.
36. Coetzee W, Coetzer RIJ, Rawatlal R. Response surface strategies in constructing statistical bubble flow models for the development of a novel bubble column, *Computers and Chemical Engineering*, 2012;36:22–34.
37. Li S, Feng L, Benner P, Seidel-Morgenstern A. Using surrogate models for efficient optimization of simulated moving bed chromatography. *Computers and Chemical Engineering*. 2014;67:121–132.

38. Wang K, Chen T, Kwa ST, Ma Y, Lau R. Meta-modelling for fast Analysis of CFD-Simulated Vapour Cloud Dispersion Processes. *Computers and Chemical Engineering*. 2014;69:89–97.
39. Jones DR. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*. 2001;21:345–383.
40. Kleijnen JPC, Van Beers W, Van Nieuwenhuyse I. Expected improvement in efficient global optimization through bootstrapped kriging. *Journal of Global Optimization*. 2012;54:59–73.
41. Yan W, Hu S, Yang Y, Gao F, Chen T. Bayesian migration of Gaussian process regression for rapid process modelling and optimization. *Chemical Engineering Journal*. 2011;166:1095–1103.
42. Long X, Han N, Chen Q. Influence of nozzle exit tip thickness on the performance and flow field of jet pump. *Journal of Mechanical Science and Technology*. 2008;22:1959–1965.
43. Sarshar S. De-liquefaction and revival of oil and gas wells using jet pump technology. 6th European well-deliquefaction conference Groningen, 2011.
44. Caballero JA, Grossmann IE. An Algorithm for the use of Surrogate Models in Modular Flowsheet Optimization, *AIChE Journal*. 2008;54:2633–2650.
45. Kleijnen JPC. *Design and analysis of simulation experiments* (2nd Edition). New York: Springer, 2015.
46. Simpson TW, Lin DKJ, Chen W. Sampling strategies for Computer Experiments: Design and Analysis. *International Journal of Reliability and Applications*. 2000;2:209-240.
47. Shimizu Y, Nakamura S, Kuzuhara S, Kurata S. Studies of the Configuration and Performance of Annular Type Jet Pumps. *Journal of Fluids Engineering*. 1987;109:205-212.
48. McKay. MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of the output from a computer code. *Technometrics*, 1979;21:239-245.
49. McKay MD. Latin Hypercube Sampling as a tool in uncertainty Analysis of computer Models. *Proceedings of the 1992 Winter Simulation Conference*. 1992;554–564.
50. Welch WJ, Yu TK, Kang SM, Sacks J. Computer experiments for quality control by parameter design. *Journal of Quality Technology*. 1990;22:15–22.
51. Kalagnanam JR, Diwekar UM. An efficient sampling technique for off-line technometrics. 1997;39:308–319.

52. Fang KT. Experimental Design by Uniform distribution. *Acta Mathematica Applicata Sinica*. 1980;3:363–372.
53. Fang KT, Lin DKJ, Winker P, Zhang Y. Uniform Design; theory and applications. *Technometrics*. 2000;42:237–248.
54. Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and Analysis of Computer Experiments. *Statistical Science*. 1989;4:409–423.
55. Jones DR, Schonlau M, Welch WJ. Efficient Global Optimization of Expensive Black Box functions. *Journal of Global Optimization*. 1998;13:455–492.
56. ANSYS Fluent. ANSYS Fluent Theory Guide. Release 15.0. 2014a.
57. ANSYS Fluent. Introduction to ANSYS Fluent, Lecture 2: Introduction to the CFD Methodology. Release 15.0. 2014b.
58. Launder BE, Spalding DB. The Numerical Computation of Turbulent Flows. *Computer Methods in Applied Mechanics and Engineering*. 1974;3:269–289.
59. Piazza AD, Piazza MCD, Vitale G. A Kriging-Based Partial Shading Analysis in a Large Photovoltaic Field for Energy Forecast. *International Conference on Renewable Energies and Power Quality (ICREPQ'09) Valencia, Spain, 2009*.
60. Chen VCP, Tsui K, Barton RR, Meckesheimer M. A review on design, modelling and applications of computer experiments, *IIE Transactions*. 2006;38:273–291.
61. Farshidi A, Rakai L, Samimi B, Behjat L, Westwick D. A new a priori net length estimation technique for integrated circuits using radial basis functions. *Computers and Electrical Engineering*. 2013;39:1204–1218.
62. El Tabach E, Lancelot L, Shahrour I, Najjar Y. Use of artificial neural network simulation metamodeling to assess groundwater contamination in a road project. *Mathematical and Computer Modeling*. 2007;45:766–776.
63. Fonseca DJ, Navaresse DO, Moynihan GP. Simulation metamodeling through artificial neural networks. *Engineering Applications of Artificial Intelligence*. 2003;16:177–183.
64. Lophaven SN, Nielsen, HB, Sondergaard J. DACE, A MATLAB Kriging Toolbox. Version 2.0. Technical Report IMM-TR-2002-12, 2002.

65. Martin J, Simpson T. Use of Kriging models to approximate deterministic computer models. *American Institute of Aeronautics and Astronautics Journal*. 2005;43:853–863.
66. Myung IJ. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*. 2003;47:90–100.
67. Quan N, Yin J, Ng SH, Lee LH. Simulation optimization via kriging: a sequential search using expected improvement with computing budget constraints. *IEE Transactions*. 2013;45:763-780.
68. Limpert E, Stahel WA. Problems with using the normal distribution – and ways to improve quality and efficiency of data analysis. *PLoS ONE*. 2011;6.
69. Sheil J, O'Muircheartaigh I. Algorithm AS 106: The Distribution of Non-Negative Quadratic Forms in Normal Variables. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1977;26:92-98.
70. Genz A, Bretz F. Computation of multivariate normal and t probabilities: Lectures notes in statistics. New York: Springer, 2009.
71. Bryd RH, Gilbert JC, Nocedal J. A Trust Region Method based on Interior Point techniques for Non-linear Programming. *Mathematical Programming*. 2000;89:149-185.
72. Waltz RA, Morales JL, Nocedal J, Orban D. An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical Programming*. 2006;107:391-408.
73. Byrd RH, Mary EH, Nocedal J. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*. 1999;9:877-900.
74. Green AJ. Jet Pumps and Ejectors, A-to-Z Guide to Thermodynamics. *Heat and Mass Transfer and Fluids Engineering*. Thermopedia. 2011.
75. Elger DF, Taylor SJ, Liou CP. Recirculation in an Annular-Type Jet Pump. *Journal of Fluids Engineering*. 1994;116:735-740.

Table Legends

Table 1. Application of meta-modelling in different fields

Table 2. Default k - ε Turbulence models constant

Table 3. Calibration results for straight-type annular jet pump

Table 4. Comparison between traditional CFD and the proposed approach (straight-type)

Table 5. Calibration results for convergent-divergent annular jet pump

Table 6. Comparison between traditional CFD and the proposed approach
(convergent-divergent type)

Figure Legends

Figure 1. Overview of calibration procedure.

Figure 2. Schematics of the annular jet pump. (a) Straight-type; (b) Convergent-divergent-type.

Figure 3. Pressure distribution variation with axial distance for straight-type annular jet pump k - ε model.

Figure 4. Pressure distribution variation with axial distance for convergent-divergent-type annular jet pump k - ε model.

Figure 5. CFD simulation versus experimental data for the standard k - ε model (Parameter measured: C_p which is dimensionless). (a) Straight-type (b) Convergent-divergent-type.

Figure 6. Leave-one-out cross validation results. (a) Straight-type
(b) Convergent-divergent-type.

Figure 7. Straight-type calibration results: C_μ independent on flow ratio.

Figure 8. Default C_μ , and calibration using single C_μ and
dependence of C_μ on flow ratio, M for straight-type.

Figure 9. Convergent-divergent-type calibration results: C_μ independent on flow ratio.

Figure 10. Single C_μ and calibration with C_μ dependent on flow ratio for
convergent-divergent type.

Figure 11. Relationship between C_μ and flow ratio.

Table 1. Application of meta-modelling in different fields

Research Area	Aspect studied	Reference
Mechanical and aerospace systems	• Aircraft structural performance	10,11
	• Aerodynamics Optimisation	12,13,14,15
	• Optimising a flow diffuser using CFD	16,17
	• Optimising an aerospace nozzle in a rocket engine	18,19
	• Optimisation of helicopter rotor blades	20
Water resources management	• Optimisation and uncertainty assessment of strongly non-linear groundwater models	2
	• Water distribution system optimisation	21
	• Water quality modelling	22,23
	• Soil vapour extraction system design	24
	• Calibration of rainfall runoff model	25,26
	• Simulation runs reduction in uncertainty estimation in hydrological modelling	27
	• Optimisation of flood control detention dams	28
	• Water distribution system design	26
Civil engineering	• Development of a meta-model based approach for integrated building energy simulation	5
	• Reliability analysis of structures	29
Chemical engineering	• Design of vacuum/pressure swing adsorption systems	30

Table 2. Default k - ε Turbulence models constant⁵⁷

Model Parameters	Standard k - ε	Realizable k - ε	RNG k - ε
$C_{1\varepsilon}$	1.44	1.44	1.42
$C_{2\varepsilon}$	1.92	1.9	1.68
C_μ	0.09	-	0.0845
σ_k	1.0	1.0	-
σ_ε	1.3	1.2	-

Table 3. Calibration results for straight-type annular jet pump

	Standard $k\text{-}\varepsilon$ model	Calibrated model
M = 0.01		
C_μ	0.09	0.1385
RMSE	0.0255	0.00325
M = 0.11		
C_μ	0.09	0.0935
RMSE	0.00450	0.00448
M = 0.19		
C_μ	0.09	0.083
RMSE	0.00700	0.00512
M = 0.34		
C_μ	0.09	0.056
RMSE	0.0270	0.0195

Table 4. Comparison between traditional CFD and the proposed approach (straight-type)

Method	Number of simulation runs	Total simulation time
Traditional CFD	29	145 hours
Proposed approach	12	60 hours

Table 5. Calibration results for convergent-divergent annular jet pump

	Standard $k\text{-}\varepsilon$ model	Calibrated model
M = 0.04		
C_μ	0.09	0.1294
RMSE	0.1128	0.0391
M = 0.3		
C_μ	0.09	0.0888
RMSE	0.1284	0.1279
M = 0.58		
C_μ	0.09	0.0524
RMSE	0.1088	0.0597

Table 6. Comparison between traditional CFD and the proposed approach
(convergent-divergent type)

Method	Number of simulation runs	Total simulation time
Traditional CFD	25	125 hours
Proposed approach	14	70 hours

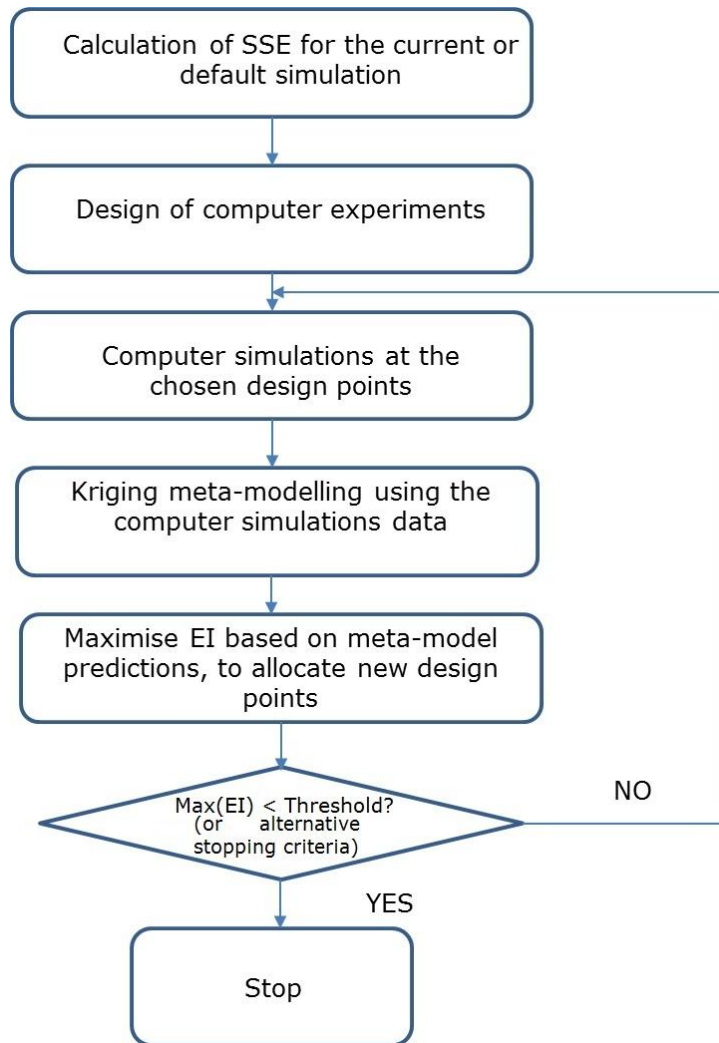


Figure 1. Overview of calibration procedure.

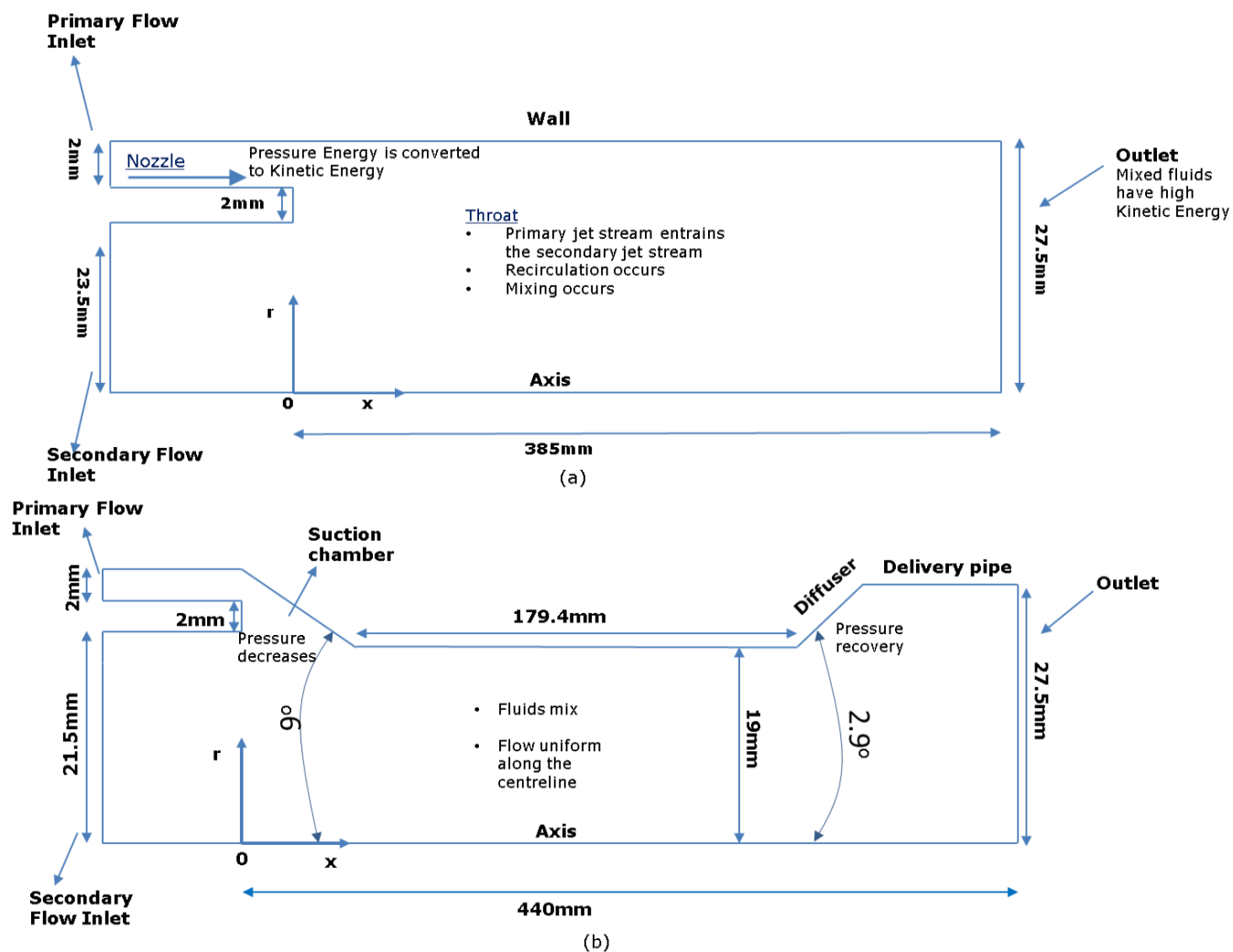


Figure 2. Schematics of the annular jet pump. (a) Straight-type; (b) Convergent-divergent-type.

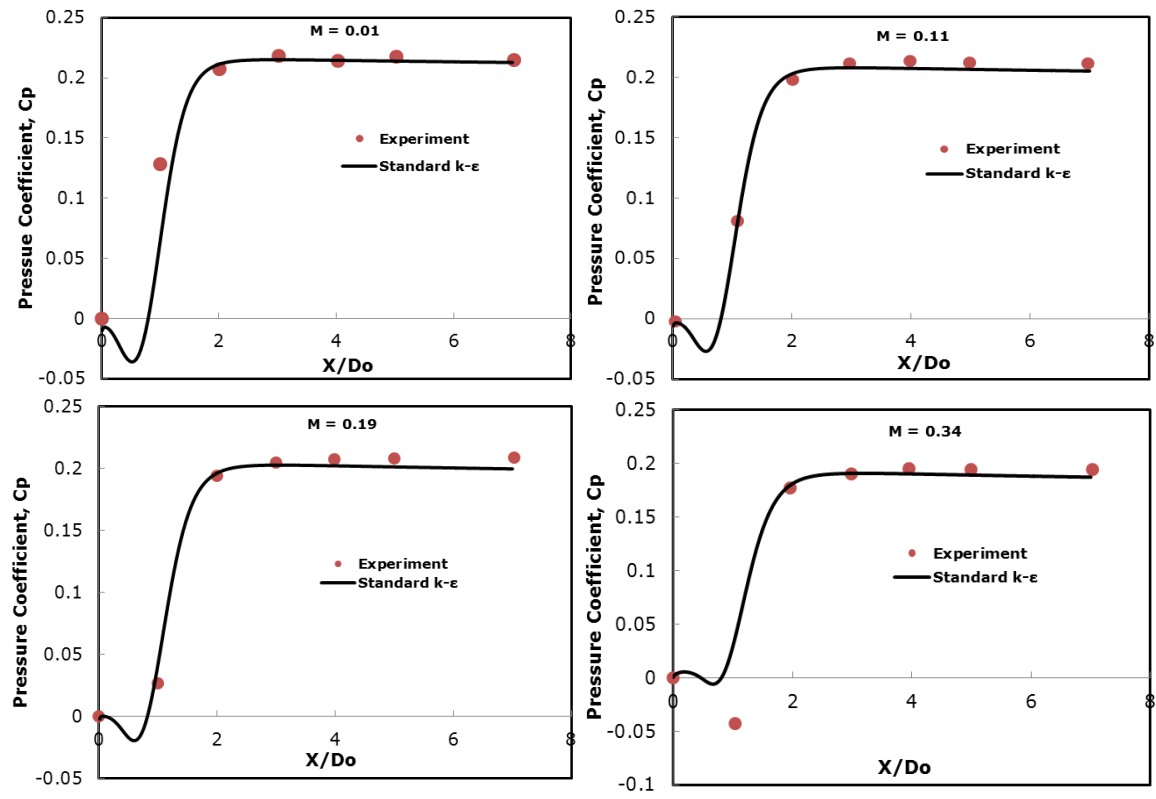


Figure 3. Pressure distribution variation with axial distance for straight-type annular jet pump $k-\epsilon$ model.

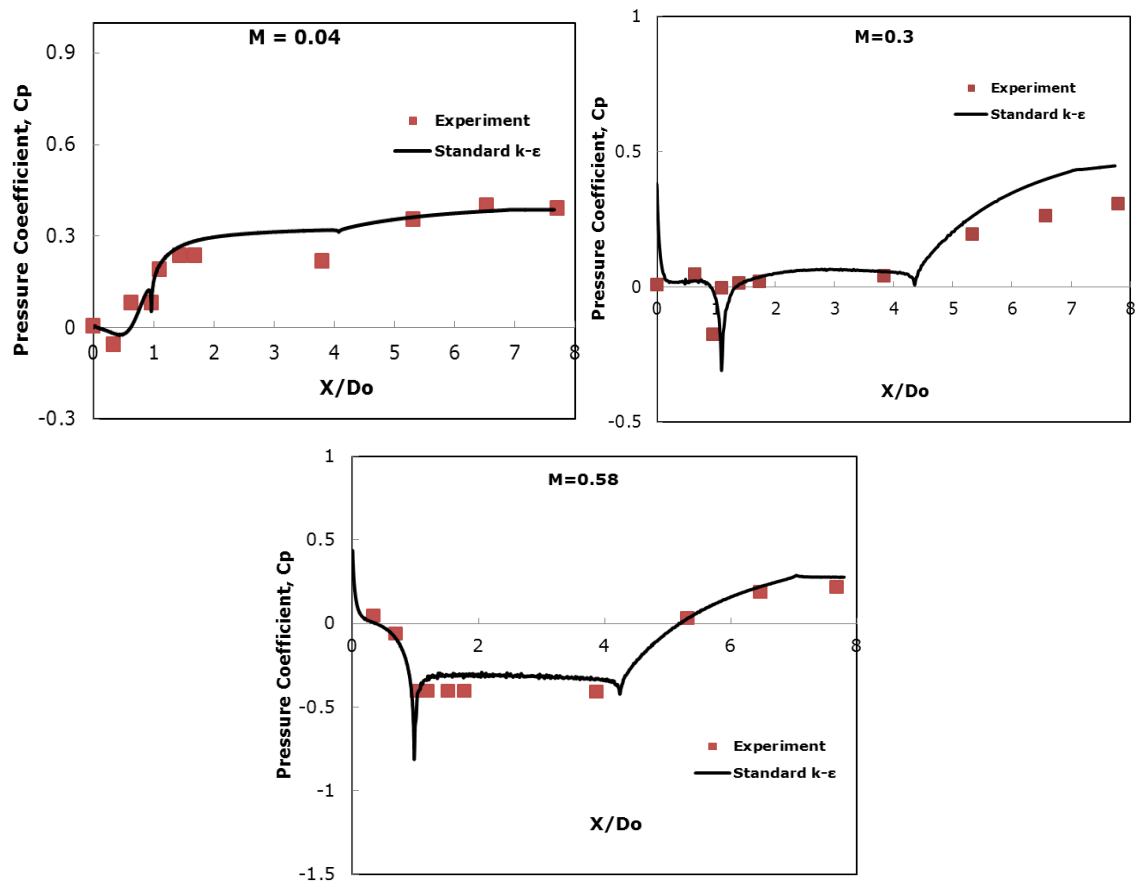


Figure 4. Pressure distribution variation with axial distance for convergent-divergent-type annular jet pump $k-\epsilon$ model.

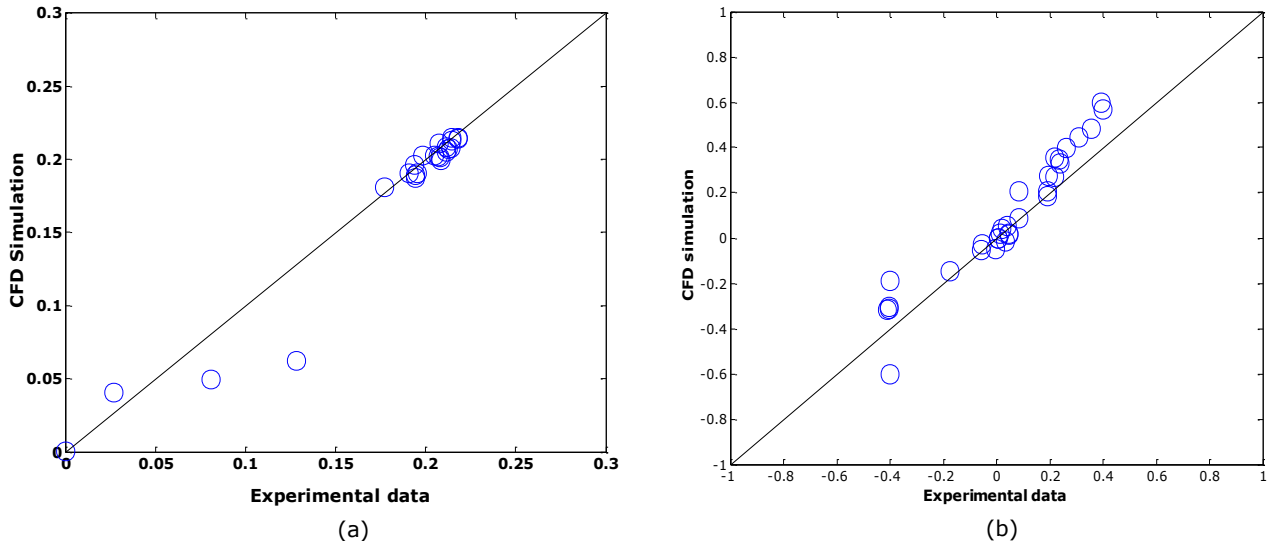


Figure 5. CFD simulation versus experimental data for the standard $k-\epsilon$ model (Parameter measured: C_p which is dimensionless). (a) Straight-type (b) Convergent-divergent-type.

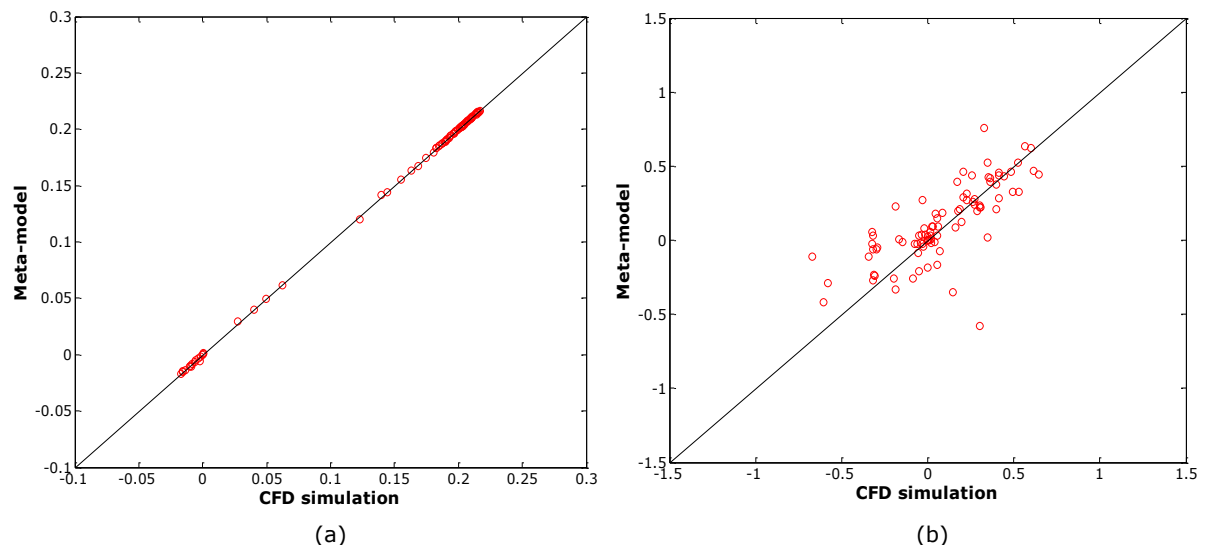


Figure 6. Leave-one-out cross validation results. (a) Straight-type
(b) Convergent-divergent-type.

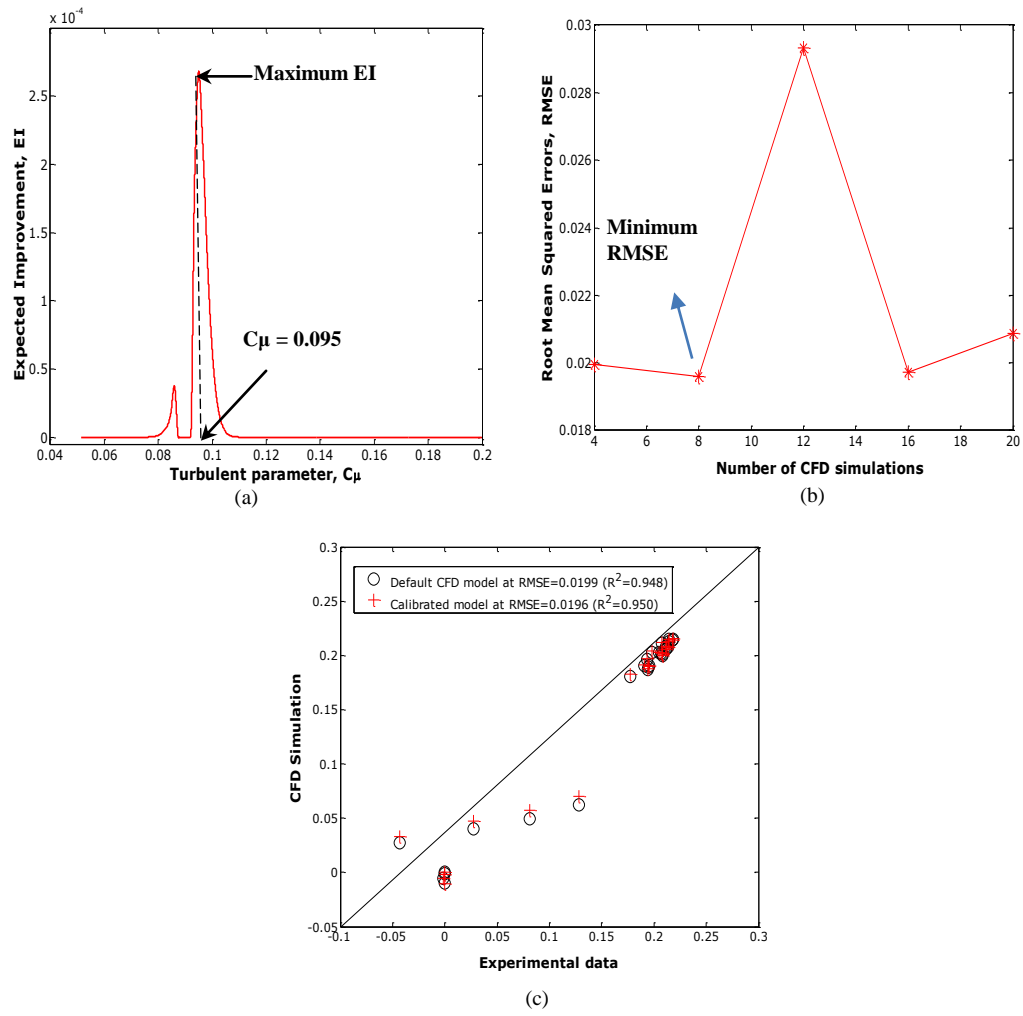


Figure 7. Straight-type calibration results: C_μ independent on flow ratio.

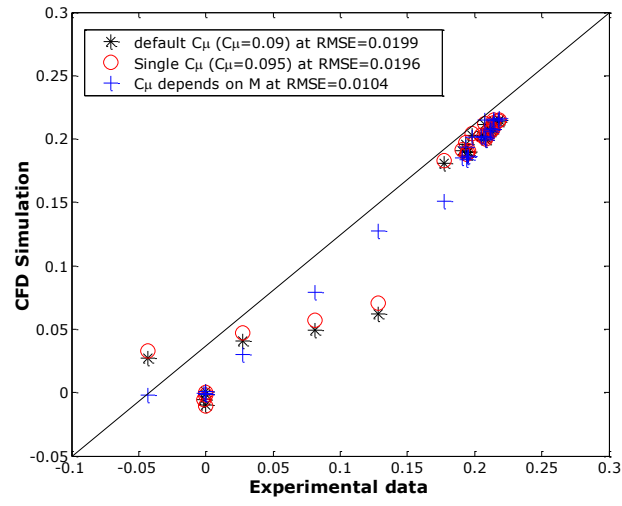


Figure 8. Default C_μ , and calibration using single C_μ and dependence of C_μ on flow ratio, M for straight-type.

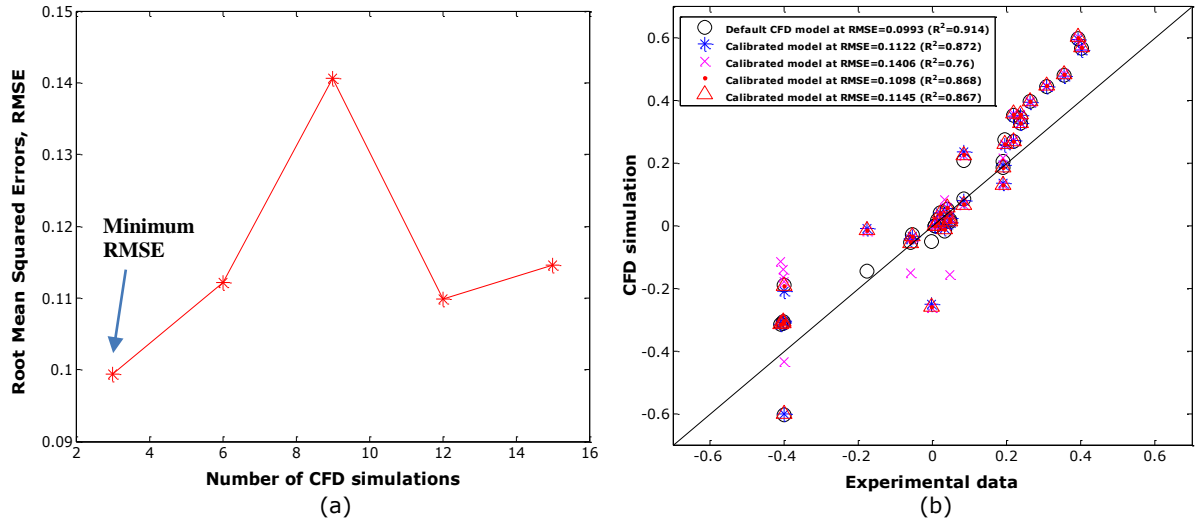


Figure 9. Convergent-divergent-type calibration results: C_μ independent on flow ratio.

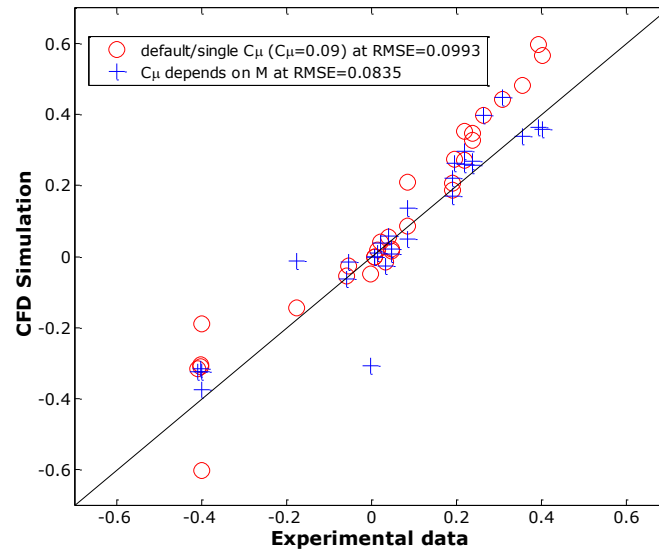


Figure 10. Single C_μ and calibration with C_μ dependent on flow ratio for convergent-divergent type.

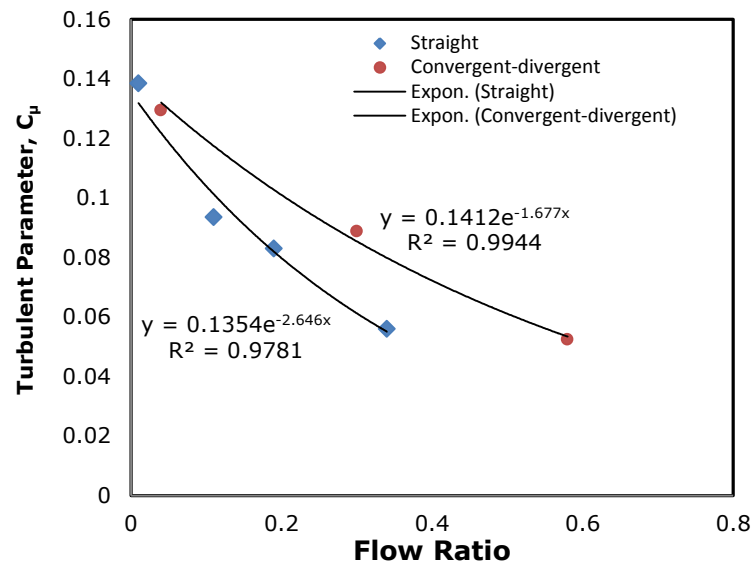


Figure 11. Relationship between C_μ and flow ratio.