Application of computational fluid dynamics and surrogate-coupled evolutionary computing to enhance centrifugal-pump performance

Sayed Ahmed Imran Bellarya, Rohit Adhava, M. Hamid Siddiquea, Bo-Hyun Chon, Frank Kenyery and Abdus Samad

aDepartment of Ocean Engineering, IIT Madras, Chennai, India; bDepartment of Energy Resources Engineering, Inha University, Incheon, South Korea; cDepartment of Transport of Energy, Universidad Simon Bolivar (USB), Caracas, Venezuela

ABSTRACT

To reduce the total design and optimization time, numerical analysis with surrogate-based approaches is being used in turbomachinery optimization. In this work, multiple surrogates are coupled with an evolutionary genetic algorithm to find the Pareto optimal fronts (PoFs) of two centrifugal pumps with different specifications in order to enhance their performance. The two pumps were used a centrifugal pump commonly used in industry (Case I) and an electrical submersible pump used in the petroleum industry (Case II). The objectives are to enhance head and efficiency of the pumps at specific flow rates. Surrogates such as response surface approximation (RSA), Kriging (KRG), neural networks and weighted-average surrogates (WASs) were used to determine the PoFs. To obtain the objective functions' values and to understand the flow physics, Reynolds-averaged Navier–Stokes equations were solved. It is found that the WAS performs better for both the objectives than any other individual surrogate. The best individual surrogates or the best predicted error sum of squares (PRESS) surrogate (BPS) obtained from cross-validation (CV) error estimations produced better PoFs but was still unable to compete with the WAS. The high CV error-producing surrogate produced the worst PoFs. The performance improvement in this study is due to the change in flow pattern in the passage of the impeller of the pumps.

Nomenclature

Abbreviations

BPS Best PRESS surrogate
CFD Computational fluid dynamics
CV Cross validation
ESP Electrical submersible pump
KRG Kriging
LE Leading edge
NSGA-II Non-dominated sorting of genetic algorithm-II
POD Pareto optimal design
PRESS Predicted error sum of squares
RANS Reynolds-averaged Navier–Stokes
RBF Radial basis function
RMSE Root mean square error
RSA Response surface approximation
TE Trailing edge

WAS Weighted-average surrogate

Greek Symbols

\( b \) Blade width, mm
\( c \) Absolute fluid flow velocity, m/s
\( d \) Diameter, mm
\( t \) Blade thickness, mm
\( v \) Absolute fluid flow velocity, m/s
\( w \) Weighting function
\( x \) Decision variable

ARTICLE HISTORY

Received 20 June 2015
Accepted 2 December 2015

KEYWORDS

surrogate model; genetic algorithm; multi-objective optimization; electrical submersible pump; centrifugal pump impeller

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
β Blade angle, °
η Hydraulic efficiency, %

### Subscripts
1 Inlet
2 Outlet
avg Average
CFD Computational fluid dynamics
C Combined
c Absolute velocity, m/s
dv Design variables
Err Error
H Head generated
krg Kriging
o Outer diameter of impeller
POD Pareto optimal design
RANS Reynolds-averaged Navier–Stokes
REF Reference
rbf Radial basis function
rsa Response surface approximation
SM Surrogate model
u Peripheral component
η Hydraulic efficiency

### 1. Introduction

With the advancement of computational capabilities, the three-dimensional simulation of any complex geometries such as turbomachines is possible. Even an analysis of multistage pumps is possible, although a detailed analysis takes a long time. The optimization of a turbomachines requires a large amount of performance data at different design points in order to generate an objective function. This process may take a long time to simulate, and thus the cost of optimization can be high. Derakhshan, Pourmahdavi, Abdolahnejad, Reihani, and Ojaghi (2013) reported an approach for enhancing the performance of turbomachines by optimizing its parameters via experimental and numerical methods. Hundreds of parameters of an impeller can be altered and the performance can be upgraded. The parameter such as number of blades affects the low-pressure area at the blade inlet and the jet wake formation in the blade passages (Houlin, Yong, Shouqi, Minggao, & Kai, 2010). A change in blade angles changes the hydraulic efficiency and the cavitation formation (Kamimoto & Matsuoka, 1956; Luo, Zhang, Peng, Xu, & Yu, 2008; Sanda & Daniela, 2012). Rutter, Sheth, and O’Bryan (2013) optimized the hydraulic performance of an electrical submersible pump (ESP) handling a single-phase fluid, with selected head, efficiency and input power as the objective functions. Cao, Peng, and Yu (2004), Shi, Long, Li, Leng, and Zou (2010) and Marsis, Pirouzpanahand, and Morrison (2013) also reported an improvement in pump performance by changing the design parameters.

However, the evaluation of all the parameters for optimization requires a high-dimensional analysis that is even more time consuming. To avoid this issue, a recently developed regression-based technique which generates a low-fidelity model also known as a surrogate is being used to assist engineers to optimize geometry in less time. The issue of the 'dimensionality curse' of surrogate models is usually helped by a sensitivity analysis, which can reduce the number of design parameters (Samad, 2008). Here, the low-fidelity model specifies an approximating function which approximately mimics the high-fidelity computational fluid dynamics (CFD) or computer aided engineering (CAE) model.

The surrogate models used for system optimization (also known as simply 'surrogates' or low-fidelity models) are highly dependent on the responses produced by the high-fidelity model (Forrester, Sobester, & Keane, 2008). If the wrong surrogate is selected for a low-fidelity construction then an erroneous result may be produced. There are many articles which investigate surrogate predictions and their comparative results (Acar & Rais-Rohani, 2009; Forrester & Keane, 2009; Goel, Haftka, Shyy, & Queipo, 2007; Lee & Choi, 2014; Nguyen, Reiterand, & Rigo, 2014; Peter & Onera, 2008; Robinson, 2007; Samad, Kim, Goel, Haftka, & Shyy, 2008; Sanchez, Pintos, & Queipo, 2008; Viana, Simpson, Balabanov, & Toropov, 2014; Zerpa, Queipo, Pintos, & Salager, 2005). The commonly-used surrogates are response surface approximation (RSA), Kriging (KRG), and radial basis function (RBF) or neural network. On the other hand, real-life engineering problems are being solved using single- or multi-objective optimization methodologies (Acar & Rais-Rohani, 2009; Forrester & Keane, 2009; Lee & Choi, 2014; Nguyen et al., 2014; Samad et al., 2008; Samad & Kim, 2008; Sanchez et al., 2008; Viana et al., 2014). However, multi-objective optimization methodology is better in terms of the performance enhancement of engineering systems and the objectives under consideration come under a 'win-win' strategy. The conflicting objectives or independent objectives can be correlated through the Pareto optimal front (PoF; Collette & Siarry, 2003; Deb, 2001). Multi-objective problems can be solved via the weighted averaging of the objectives and the formulation is given as:

\[ F_C = \sum w_i F_i, \]  

where \( i = 1, 2, \ldots n \), and \( n \) is the number of objectives. The non-dominated sorting of genetic algorithm-II (NSGA-II) and other optimizers such as particle swarm
intelligence for multi-objective optimizations were compared in the literature and it was observed that NSGA-II is extensively used (Couckuyt, Turck, Dhaene, & Gorissen, 2011; Cravero & Macelloni, 2010; Gorissen, Couckuyt, Laermans, & Dhaene, 2010; Husain & Kim, 2010; Husain, Lee, & Kim, 2011; Marjavaara, Lundstrom, Goel, Mack, & Shyy, 2007; Taormina & Chau, 2015). There are several commercial as well as non-commercial software programs like ModeFrontier, SUMO, etc. which use surrogate-based methods along with the multi-objective Pareto optimal designs (PODs; Couckuyt et al., 2011; Cravero & Macelloni, 2010). Jin (2011) compared surrogates and shows that a neural network can be applied where all the ensemble members are of the same type of feed-forward neural networks.

Ensemble surrogate methods have been developed by several authors (Acar & Rais-Rohani, 2009; Goel et al., 2007; Lee & Choi, 2014; Samad et al., 2008; Sanchez et al., 2008; Zerpa et al., 2005; Zhang & Chau, 2009) and assessed and executed for different application areas (Marjavaara et al., 2007; Samad et al., 2008; Samad & Kim, 2008; Zerpa et al., 2005) for single objective optimizations. Global data-based error measurement gives the fidelity information for the surrogates and can be used to ensemble the surrogates or to create a weighted-average surrogate (WAS). Such a comprehensive WAS-based strategy is required to check the performance of evolutionary algorithms to generate PoFs.

Surrogates are problem-dependent; thus, a single surrogate cannot be applied to solving every problem. To handle this challenge, an ensemble of surrogates or the concept of the WAS was introduced. The WAS is based on the average of the predicted error sum of squares (PRESS), which is implemented from cross validation (CV) error estimations. As the WAS includes the contribution from all the surrogates, it gives quite satisfactory results.

In this paper, a strategy using multiple-surrogate-assisted multi-objective optimization is developed to enhance the performances of two pumps: an ESP and a centrifugal impeller. CFD simulations were performed to find the objective function values at different points and a PoF was generated via a genetic algorithm. The PoFs obtained from all surrogates were compared to find a robust surrogate that performs well for both cases. The results and discussion section contain an analysis of the surrogate fidelity and flow physics of the pumps.

2. Optimization strategy

Figure 1 shows the optimization strategy used in this study. It starts with problem formulations which include understanding the problem and deciding on the objective functions which need to be optimized and the design variables which influence these functions. In this study, two cases are used.

2.1. Problem formulation and numerical schemes

Case I involves a centrifugal pump impeller with a diameter ($d_o$) of 365 mm. Three design variables and two objectives were considered based on the literature (Bellary & Samad, 2014; Gulich, 2010; Luo et al., 2008); inlet angle ($\beta_1$), exit angle ($\beta_2$) and number of blades ($G$). The objective functions were to maximize efficiency and head. The geometric features of the impeller are shown in Table 1 (Lazarkiewicz & Troskolanski, 1965).

Computer-aided design (CAD) modeling and meshing of the flow domain were carried out using Ansys BladeGen and TurboGrid modules, respectively (Ansys CFX 13.0, 2010). Ansys CFX 13.0 was used for the flow simulations. Table 2 shows the meshing and boundary conditions. Further details regarding geometry,
Table 1. Features of impeller.

| Parameters          | Case I          | Case II         |
|---------------------|-----------------|-----------------|
| Inlet diameter, \(d_1\) | 160 mm          | 40 mm           |
| Outlet diameter, \(d_2\) | 365 mm          | 100 mm          |
| Inlet blade width, \(b_1\) | 54 mm           | 9.84 mm         |
| Outlet blade width, \(b_2\) | 30 mm           | 8.25 mm         |
| Blade thickness, \(t\) | 5 mm            | 3 mm            |
| Inlet blade angle, \(\beta_1\) | 17°             | 30°             |
| Exit blade angle, \(\beta_2\) | 70°             | 22.5°           |
| Blade number, \(z\) | 7               | 7               |

Table 2. Meshing and boundary conditions.

| Parameter               | Case I                      | Case II                     |
|-------------------------|-----------------------------|-----------------------------|
| Flow domain             | Single impeller             | Impeller and diffuser       |
| Interface               | Periodic                    | Periodic and frozen rotor   |
| Mesh/nature             | Structured/Hexahedral       | Structured for rotor/diffuser|
|                         | Unstructured for intermediate zone |                            |
| Nodes/elements          | 550,000/521,185             | 266,084/227,416             |
| Fluids                  | Water                       | Water                       |
| Turbulence model        | \(k-\varepsilon\)           | \(k-\varepsilon\)           |
| Inlet/outlet            | Pressure/mass flow rate     | Pressure/mass flow rate     |
| Iteration steps         | 2000                        | 1600                        |
| Y+ value                | 38                          | 193                         |

Boundary conditions, validation and grid independence test can be found in Bellary and Samad (2014).

Case II involved an ESP with a diameter of 100 mm. The impeller and diffuser were considered for the analysis, and two variables (inlet and outlet blade angles) along with two objectives (head and efficiency) were used for the optimization in this case. Figures 2 & 3 show the geometry and the flow domain of the ESP, and the impeller features are shown in Table 1. The simulations were carried out for an impeller at a speed of 3000 rpm with a discharge rate of \(Q = 0.0037\, m^3/s\). The grid independence test was performed and an optimum number of nodes equal to 227,416 was selected. Mass flow inlet and pressure outlet boundary conditions are used at the inlet and outlet, respectively. Further details regarding the boundary conditions and flow simulation can be found in Table 2. The turbulence model used was \(k-\varepsilon\) which uses the scalable wall-function approach to improve robustness and accuracy as it efficiently captures near-wall boundary layer effects (Wilcox, 1994). The numerical results were validated with the experimental results given in Caridad, Asuaje, Kenery, Tremante, and Aguillon (2008).

The working fluid for both the cases was water of standard properties at a normal temperature and pressure. The performance of the pumps was calculated by solving the Reynolds-averaged Navier–Stokes (RANS) equations and used as design variable responses.

2.2. Design of experiment

After the completion of the problem formulation stage, a design space bounded by the two extreme (top and bottom) limits of the variables was created. The lower and upper bounds were determined via a literature survey (Bellary & Samad, 2014; Gulich, 2010; Li, 2002; Luo et al., 2008; Marsis et al., 2013; Ohta & Aoki, 1996), followed by simulations at the corners of the design space. A three-level full factorial design (Myers & Montgomery, 1995) was employed to choose the designs or design points from the design space. Once again, RANS equations were solved to obtain the design variable response that those design points. These responses were then used to construct the surrogate model.

2.3. Formulating the surrogate modeling

The surrogates used in this study are RSA, RBF, KRG and WAS. RSA is a polynomial function for responses generated from numerical calculations. The basic concept of a neural network or RBF is to learn from events. The RBF (Samad et al., 2008; Samad & Kim, 2008) comprises of a two-layer network. The first layer is a hidden layer of a radial basis function and the second is a linear output layer. The variables required to generate this surrogate model are a spread constant and a user-specified error goal, which comes from the mean input responses.

![Figure 2. Geometry showing impeller, rotor and diffuser.](image-url)
The KRG (Husain & Kim, 2010) interpolation employs a known function for a large-scale variation and a function for a small-scale departure. The KRG model evaluated at some unsampled location has a global model and a systematic departure. The algorithm to construct a PoF through WAS and best PRESS surrogate (BPS) is as follows:

**Algorithm 1.** Algorithm for BPS- and WAS-assisted PoF.

**Step 1:** Generate data from CFD analysis
**Step 2:** Construct surrogates
**Step 3:** Calculate PRESS

For BPS-BPSPoF

**Step 4a:** Find least error and select surrogates for objective 1 and objective 2.

**Step 5a:** Generate population for NSGA-II using the selected surrogates

**Step 6a:** Generate PoF and cluster points
**Step 7a:** Check errors with CFD analysis

For WAS-WASPoF

**Step 4b:** Find weights for the surrogates and construct WAS for objective 1 and WAS for objective 2.

**Step 5b:** Generate population for NSGA-II using the WAS models

**Step 6b:** Generate PoF and cluster points
**Step 7b:** Check errors with CFD analysis

### 2.3.1. Weighted-average surrogate (WAS)

Generally, a large amount of data distributed in the design space fits a surrogate well, and the surrogate reflects the trend of the objectives. Any poorly-fitted surrogate may lead to unreliable results. The ensemble or averaged surrogates adequately refines the response trend of the data and insulates against bad surrogates (Goel et al., 2007; Queipo et al., 2005). A WAS model generated from the PRESS or CV errors (Sanchez et al., 2008) of the basic surrogates (RSA, KRG and RBF) was adopted in this problem. The predicted response of the WAS model is defined in Equation (1). In simplified form, the function can be written as

\[
F_{\text{avg}} = w_{\text{rsa}}F_{\text{rsa}} + w_{\text{krg}}F_{\text{krg}} + w_{\text{rbf}}F_{\text{rbf}}
\]

where \(F_{\text{rsa}}, F_{\text{krg}}, \) and \(F_{\text{rbf}}\) are constructed using the CFD evaluated responses. A lower weight is assigned to the surrogate with a higher CV error, which means that the surrogate makes a reduced contribution to the construction of the WAS model.

In CV, the data is divided into \(k\) subsets of equal size. A surrogate is constructed \(k\) times, with each time excluding one of the subsets from training, and the excluded subset is used to evaluate the error. The \(k\) errors finally produce the error sum of squares. This process is known as leave-one-out CV. The calculation of the weights used in WAS is given in Goel et al. (2007).

### 2.3.2. Best PRESS surrogate (BPS)

To implement a surrogate ensemble, the best surrogate must be selected from the available options. This can be achieved using a weighting scheme, and a weight equal to 1 is assigned to the surrogate with the least CV error. A 0 weight is allocated to the other models. This scheme is termed the BPS model (Goel et al., 2007) and is based on the concept of selecting a surrogate that gives the least CV error. In the present problem, the CV errors for the objectives were calculated separately and the surrogate which produced the least error was used for the PoF generation.

### 2.4. Non-dominated sorting of genetic algorithm-II (NSGA-II)

NSGA-II (Deb, 2001) is widely used in the scientific community. In the present problem, the algorithm is coupled with the Surrogates Toolbox (Viana, 2011; Viana et al., 2014) in order to make handling multiple functions easier. The relationships among the competitive objectives were established via the PoF produced by NSGA-II. The problem relates to two competing objectives (for each pump case), where the enhancement of one objective results in the degradation of the other. All feasible solutions related to the multi-objective problem are grouped into two types: dominated and non-dominated. The non-dominated solutions are designated as Pareto-optimal solutions. The objective functions are defined mathematically and estimated by numerical simulation.

In this approach, an initial or parent population is arbitrarily established, and the objective function values are estimated for each design. Here, the term ‘population’ refers to the designs in the design space and the non-dominance condition and crowding distance determine the rank of this population. Intermediate populations are
created via selection, crossover and mutation by implementing a genetic operator. The populations (parent and intermediate) are combined, and the rank and crowding distance of the combined population is obtained. New populations are selected depending on the rank and crowding distance and are again ranked for the next generation based on the non-domination criteria. This process is reiterated until the limiting number of generations is attained.

### 2.5. k-means clustering

Five clusters in the global Pareto-optimal solutions were generated by the $k$-means clustering algorithm. This is a repeated-fitting process that generates the number of specified clusters. The clusters are distributed across the whole range of the PoF.

### 3. Results and discussion

#### 3.1. Case I

Initially, three different variables ($\beta_1$, $\beta_2$ and $G$) affecting the performance of the pump (Lazarkiewicz & Troskolaniski, 1965; Srinivasan, 2008) were selected and a parametric study was performed. The allocation of the angles was assured based on the literature (Li, 2002; Ohta & Aoki, 1996). Design points within the design space were selected from a three-level full factorial design method. The design space defined by the lower and upper limits of the variables is given in Table 3. The impeller passage was generated by polynomial curves at the inlet and the outlet, so that a small change in angle produces a continuous surface.

The surrogate which gives the least CV error was selected as the BPS. In Table 4, RSA shows the least CV error for both objectives, so it was selected as the BPS in both cases. The BPS was then used to generate populations using NSGA-II to build a BPS-BPS PoF (Figure 4). Figure 4(a) shows the comparison of PoFs for different surrogates. The KRG-produced PoF designs are greatly distant from reality and can also be compared in the PRESS (Table 5) which shows that the KRG-produced errors are the highest for both objectives. In Figure 4(b), the PoF at some cluster points of the KRG-KRG PoF show an efficiency of more than 100%, which is unrealistic for this type of simulation. The WAS-WAS PoF has higher fidelity.

The design variable values (Figure 5) are calculated for the cluster points. To get a higher efficiency, the number of blades should be increased and the inlet angle should be reduced. The exit angle initially increases the efficiency, but after attaining a certain efficiency it is reduced. Similarly for the head, the opposite phenomenon is observed (Figure 5(b)). Among different cluster points for the BPS, the one with a 59.63 exit blade angle ($\beta_2$) is chosen and compared with other cluster points of different $\beta_2$ values. In Table 6 for different numbers of blades other than the designed one, the error is greater even though the exit blade angle has increased. For the other cases, the same surrogates, i.e., KRG-KRG and RBF-RBF were selected for both objectives. The combined error ($C_{Err}$), which is the root mean square error (RMSE) of

| Table 3. Design variables with design space. |
|--------------------------------------------|
| Case | Variables | Lower limit | Upper limit |
|------|------------|--------------|-------------|
| Case I | $\beta_1$ | 10 | 20 |
| | $\beta_2$ | 30 | 50 |
| | $G$ | 5 | 9 |
| Case II | $\beta_1$ | 17 | 42 |
| | $\beta_2$ | 14 | 41 |

| Table 4. PRESS for weights calculation. |
|----------------------------------------|
| Case | Objective | KRG | RSA | RBF |
|------|-----------|-----|-----|-----|
| Case I | $\eta$ | 8.683 | 1.343 | 2.850 |
| | $H$ | 18.028 | 2.691 | 3.158 |
| Case II | $\eta$ | 0.402 | 0.961 | 0.34 |
| | $H$ | 0.107 | 0.104 | 0.092 |

![Figure 4. Case I PoF and cluster points for (a) different surrogates and (b) selected cluster points.](image-url)
Table 5. Surrogate-produced errors.

| Case  | Surrogate | η<sub>REF</sub> | H<sub>REF</sub> | η<sub>POD</sub> | H<sub>POD</sub> | η<sub>CFD</sub> | H<sub>CFD</sub> | η<sub>Err</sub> | H<sub>Err</sub> | C<sub>Err</sub> |
|-------|-----------|----------------|---------------|----------------|----------------|----------------|---------------|----------------|----------------|-------------|
| Case I| BPS-BPS   | 90             | 40            | 92.12          | 48.48          | 92.52          | 48.60         | 0.4            | 0.12           | 0.29        |
|       | WAS-WAS  | 90             | 40            | 91.41          | 49.93          | 91.79          | 50.12         | 0.38           | 0.19           | 0.30        |
|       | KRG-KRG  | 90             | 40            | 90.46          | 51.12          | 91.89          | 50.3          | 1.54           | −0.82          | 1.23        |
|       | RBF-RBF  | 90             | 40            | 91.35          | 50.63          | 91.86          | 50.65         | 0.51           | 0.02           | 0.36        |
| Case II| BPS-BPS | 66             | 10.3          | 66.12          | 10.27          | 66.30          | 10.29         | 0.18           | 0.02           | 0.13        |
|       | WAS-WAS  | 66             | 10.3          | 66.31          | 10.35          | 66.35          | 10.39         | 0.04           | 0.04           | 0.04        |
|       | KRG-KRG  | 66             | 10.3          | 67.3           | 10.15          | 66.93          | 10.09         | −0.37          | −0.06          | 0.26        |
|       | RSA-RSA  | 66             | 10.3          | 66.65          | 10.52          | 66.71          | 10.55         | 0.06           | 0.03           | 0.05        |

In Figure 6, the CFD results are compared for the two extreme cases (cluster points A and E in Figure 5(b)). The shock losses, secondary losses and hydraulic losses which reduce the hydraulic efficiency are greater at the off-design points (Lazarkiewicz & Troskolanski, 1965; Srinivasan, 2008). The head generation is augmented by an increase in β<sub>2</sub> at the reference point. Here, an absolute flow velocity at the exit (c<sub>2</sub>) increases with an increase in the peripheral velocity (c<sub>u2</sub>), which enables a further increment in the dynamic head (Bellary & Samad, 2014; Gulich, 2010). The contours of total pressure at two extreme cluster points show that the highest pressure is generated at the impeller exit where the kinetic energy of flow attains a maximum level. This is because the pressure augments constantly with the mechanical energy

![Figure 5](image-url)  
Figure 5. Case I design variables trend for WAS-WAS PoF (a) efficiency and (b) head.

η<sub>Err</sub> and H<sub>Err</sub>, is high for the KRG-KRG results and thus it made the least contribution to constructing the WAS model. The surrogates produce different values of PRESS; hence, the use of multiple surrogates or weighted surrogates is suggested. Table 5 shows the PRESS errors and RMSEs of the BPS and WAS at different cluster points. All five clustered designs were simulated for comparison.

Table 6. Surrogate-produced errors for Case I.

| Model | Cluster point | η<sub>POD</sub> | H<sub>POD</sub> | η<sub>CFD</sub> | H<sub>CFD</sub> | η<sub>CFD</sub> − η<sub>POD</sub> | H<sub>CFD</sub> − H<sub>POD</sub> | RMSE |
|-------|---------------|----------------|---------------|----------------|---------------|-----------------------------|-----------------------------|------|
| BPS-BPS | A             | 92.54          | 47.5          | 92.24          | 49.68         | −0.28                       | 2.18                        | 1.55 |
|        | B             | 92.12          | 48.48         | 92.52          | 48.6          | 0.4                         | 0.12                        | 0.30 |
|        | C             | 90.32          | 51.1          | 91.12          | 53.1          | 0.8                         | 2.0                         | 1.52 |
|        | D             | 89.33          | 51.8          | 91.13          | 52.79         | 1.8                         | 0.99                        | 1.45 |
|        | E             | 87.15          | 53.81         | 89.91          | 45.74         | 2.76                        | −8.07                       | 6.03 |
| WAS-WAS | A             | 86.65          | 53.18         | 87.43          | 51.8          | 0.78                        | −1.38                       | 1.12 |
|        | B             | 88.10          | 52.58         | 91.15          | 48.58         | 3.05                        | −4.0                        | 3.55 |
|        | C             | 91.41          | 49.93         | 91.79          | 50.12         | 0.38                        | 0.19                        | 0.30 |
|        | D             | 92.56          | 47.57         | 91.96          | 48.51         | −0.6                        | 0.94                        | 0.78 |
|        | E             | 92.74          | 46.57         | 92.4           | 50.00         | −0.34                       | 3.43                        | 2.43 |
supplied by the rotation of impeller and is then transformed into pressure energy. In the vicinity of the leading edge (LE) of the blade at the suction side, minimum pressure is observed (Figure 6). An increase in blade angle increases the pressure difference between the outlet and inlet of the blade. The absolute velocity component \(c_{u2}\) at the exit of the impeller increases with an increase in the discharge angle (Srinivasan, 2008).

This matches with the existing results (Lazarkiewicz & Troskolanski, 1965; Srinivasan, 2008), which state that an increase in \(\beta_2\) decreases the relative velocity and increases the total pressure.

### 3.2. Case II

In this case, a three-level full factorial design was also employed to find the designs from the design space (Table 3) obtained from the literature (Li, 2002; Ohta & Aoki, 1996; Srinivasan, 2008). A polynomial curve makes a smooth impeller surface. The variables and responses were used to construct the KRG and RBF surrogates and the WAS was formulated using weighted averaging on these surrogates. The PRESS shows that the RBF is the best one for both objectives (Table 4) and thus it selected for the BPS-BPS PoF construction. RSA and KRG produced the higher errors and so their contributions to the construction of the WAS model were least. Table 5 shows the CFD results of the cluster points and their respective errors. The WAS-WAS generated points show the least error, while the KRG-KRG show the most error.

Figure 7(a) shows that the RSA-RSA PoF has a different trend in PoF unlike the WAS-WAS PoF and the BPS-BPS PoF trends. The errors in Table 4 show that RSA produced the highest error for the efficiency. The cluster points and respective design variable values are shown in Figures 7 and 8. The inlet angle \((\beta_1)\) increases if a higher efficiency is expected but the exit angle behaves in the opposite way. For a higher head, the opposite behavior of the variables can be observed (Figure 8(b)). Table 5 shows that the WAS-WAS PoF has the least combined error \((C_{Err})\). Table 7 shows that the error is low in the WAS-WAS PoF and the BPS-BPS PoF.

Figure 9 shows the CFD calculations for two extreme points (designs A and E) from the clustered points of Figure 8(b). On the suction side and at 50% span, a low-pressure region appears in both the designs. The impingement of the flow on the LE causes a sudden loss in pressure. Also, a wider low-pressure region, which causes more flow losses and reduced efficiency, is present in design A as compared to design E. In Figure 9, the pressure increases from the LE to the trailing edge (TE) at different rates in both cases (designs A and E). A rapid.
pressure increase from LE to TE is observed in design A compared to design E, which generates more head in design A (Gulich, 2010). The values of efficiency and pressure head for these two cases are given in Table 7.

From the above analysis, it can be seen that the surrogates predict differently with different sets of data. The best surrogate for one set of data (Case I) produced the worst results for the other set (Case II). In Case I, the best surrogate was RSA, while it was RBF in Case II. The WAS model performed reliably or had higher fidelity with respect to CFD.

4. Conclusion

In this paper, a multiple-surrogate-assisted multi-objective optimization procedure is presented. The capability of surrogates to produce PoFs was evaluated with two examples of pump or turbomachinery design and optimization.

It was observed that the CV error is different for different surrogates and gives an indication of the degree of fit for each surrogate. A surrogate producing a high CV error is highly unreliable for generating a PoF. The verification with the CFD model shows that the error in Pareto designs is high for the surrogates which produce a high CV error. The BPS-produced PoF is better and can be used instead of any individual surrogate. The WAS-WAS PoF, which shows a robust result, is comparable to the BPS-BPS PoF. Hence, instead of using a single surrogate, the authors suggest the use of multiple surrogates to generate PoFs. A change in exit angle influences the augmentation of head and efficiency. The improvement in efficiency was due to the mitigation in flow loss in the disc friction losses outside of the impeller shroud and hub.

The BPS performs well for both cases and a wider design space or sparsely-distributed design points may require sequential sampling or multiple surrogate analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by an Indian Institute of Technology Madras NFSC grant [grant number OEC/10-11/529/NFSC/ABDU]; the Ministry of Trade, Industry and Energy (MOTIE) for the Korea Energy and Mineral Resources Engineering Program; and the Special Education Program for Offshore Plant.
References

Acar, E., & Rais-Rohani, M. (2009). Ensemble of metamodels with optimized weight factors. Structural and Multidisciplinary Optimization, 37(3), 279–294.

Ansys-CFX 13.0. Ansys Inc. 2010.

Bellary, S. A. I., & Samad, A. (2014). An alternate approach. Ansys-CFX13.0. Ansys Inc. 2010.

Couckuyt, I., Turck, F. D., Dhaene, T., & Gorissen, D. (2011). Deb, K. (2001).

Cravero, C., & Macelloni, P. (2010). Caridad, J., Asuaje, M., Kenyere, F., Tremante, A., & Aguillon, O. (2008). Characterization of a centrifugal pump impeller under two phase flow conditions. Journal of Fluids Engineering, 127(2), 330–338.

Collette, Y., & Siarry, P. (2003). Multiobjective optimization: Principles and case studies. New York: Springer.

Collette, Y., Turk, F. D., Dhaene, T., & Gorissen, D. (2011). Automatic surrogate model type selection during the optimization of expensive black-box problems. Proceedings of the 2011 Winter Simulation Conference, December 11–14, Phoenix, AZ, USA.

Cravero, C., & Macelloni, P. (2010). Design optimization of a multistage axial turbine using a response surface based strategy. Proceedings of the 2nd Int. Conference on Eng. optimization. 1–11.

Deb, K. (2001). Multi-objective optimization using evolutionary algorithms (1st ed.). New York: John Wiley & Sons.

Derakhshan, S., Pourmahdavi, M., Abdolahnejad, E., Reihani, A., & Ojaghi, A. (2013). Numerical shape optimization of a centrifugal pump impeller using artificial bee colony algorithm. International Journal of Computers & Fluids, 81, 145–151.

Forrester, A. I. J., & Keane, A. J. (2009). Recent advances in surrogate-based optimization. Progress in Aerospace Sciences, 45(1–3), 59–79.

Forrester, A. I. J., Sobester, A., & Keane, A. J. (2008). Engineering design via surrogate modeling: A practical guide. Chichester: John Wiley & Sons Ltd.

Goel, T., Haftka, R. T., Shyy, W., & Queipo, N. V. (2007). Ensemble of surrogates. Structural and Multidisciplinary Optimization, 33(3), 199–216.

Gorissen, D., Couckuyt, I., Laermans, E., & Dhaene, T. (2010). Multiobjective global surrogate modeling dealing with the 5-percenter problem. Engineering with Computers, 26(1), 81–98.

Gulich, J. F. (2010). Centrifugal pumps (2nd ed.). Berlin: Springer Publications.

Houlin, L., Yong, W., Shouqi, Y., Minggao, T., & Kai, W. (2010). Effects of blade number on characteristics of centrifugal pumps. Chinese Journal of Mechanical Engineering, 23, 1–6.

Husain, A., & Kim, K. Y. (2010). Enhanced multi-objective optimization of a microchannel heat sink through evolutionary algorithm coupled with multiple surrogate models. Applied Thermal Engineering, 30(13), 1683–1691.

Husain, A., Lee, K. D., & Kim, K. Y. (2011). Enhanced multi-objective optimization of a dimpled channel through evolutionary algorithms and multiple surrogate methods. International Journal for Numerical Methods in Fluids, 66(6), 742–759.

Jin, M. (2011). Surrogate-assisted evolutionary computation: Recent advances and future challenges. Swarm and Evolutionary Computation, 1(2), 61–70.

Kamimoto, G., & Matsuoka, Y. (1956). On the flow in the impeller of centrifugal type hydraulic machinery (The 2nd report). Transactions of the Japan Society of Mechanical Engineers, Series 3. 22(113), 55–59.

Lazarkiewicz, S., & Toskolanksi, A. T. (1965). Impeller pumps (1st ed.). Oxford, London: Pergamon press Ltd.

Lee, Y., & Choi, D. H. (2014). Point wise ensemble of metamodels using v nearest points cross-validation. Structural Multidisciplinary Optimization. doi:10.1007/s00158-014-1067-1

Li, W. G. (2002). The influence of number of blades on the performance of centrifugal oil pumps. WorldPumps, 427, 32–35.

Luo, X., Zhang, Y., Peng, J., Xu, H., & Yu, W. (2008). Impeller inlet geometry effect on performance improvement for centrifugal pumps. Journal of Mechanical Science and Technology, 22, 1971–1976.

Marjavaara, B. D., Lundstrom, T. S., Goel, T., Mack, Y., & Shyy, W. (2007). Hydraulic turbine diffuser shape optimization by multisupergre model approximations of Pareto fronts. Transactions of the ASME, 129, 1228–1240. September 2007.

Marsis, E., Pirouzpanahand, S., & Morrison, G. (2013). CFD-based design improvement for single-phase and two-phase fluids inside an electrical submersible pump. ASME 2013 Fluids Engineering Division Summer Meeting. Incline Village, Nevada, USA. FEDSM2013-16060, V01BT1A006; 8 pages. July 7–11.

Myers, R. H., & Montgomery, D. C. (1995). Response surface methodology-process and product optimization using designed experiments. New York: John Wiley & Sons, Inc.

Nguyen, A. T., Reiterand, S., & Rigo, P. A. (2014). Review on simulation-based optimization methods applied to building performance analysis. Applied Energy, 113, 1043–1058.

Ohta, H., & Aoki, K. (1996). Effect of impeller angle on performance and internal flow of centrifugal pump for high viscosity liquids. Proceedings of the School of Engineering, Tokai University, 36(1), 159–168. 28.

Peter, J., & Onera, M. M. (2008). Comparison of surrogate models for turbomachinery design. WSEAS Transactions on Fluid Mechanics, 3(1), 10–17.

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Tucker, P. K. (2005). Surrogate based analysis and optimization. Progress in Aerospace Sciences, 41(1), 1–28.

Robinson, T. D. (2007). Surrogate based optimization using multifidelity models with variable parameterization (PhD thesis). Massachusetts Institute of Technology.

Rutter, R., Sheth, K., & O’Bryan, R. (2013). Numerical flow simulation and validation of an electrical submersible pump. ASME Proceedings, Symposium Application in CFD. Incline Village, Nevada, USA. 1A(FEDSM201316078), V01 AT03A005; 7 pages. July 7–11, 2013. doi:10.1115/FEDSM 2013-16078

Samad, A. (2008). Numerical optimization of turbomachinery blade using surrogate models (PhD thesis). Inha University, Republic of Korea.

Samad, A., & Kim, K. Y. (2008). Shape optimization of an axial compressor blade by multi-objective genetic algorithm.
Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 222, 599–611.
Samad, A., Kim, K. Y., Goel, T., Haftka, R. T., & Shyy, W. (2008). Multiple surrogate modeling for axial compressor blade shape optimization. Journal of Propulsion and Power, 24(2), 301–310.
Sanchez, E., Pintos, S., & Queipo, N. V. (2008). Toward an optimal ensemble of kernel-based approximations with engineering applications. Structural and Multidisciplinary Optimization, 36, 247–261.
Sanda, B., & Daniela, C. V. (2012). The influence of the inlet angle over the radial impeller geometry design approach with Ansys. Journal of Engineering Studies and Research, 18(4), 32–39.
Shi, W., Long, F., Li, Y., Leng, H., & Zou, P. (2010). Non-overload design of low specific speed submersible pump. ASME 2010 3rd Joint US-European Fluids Engineering Summer Meeting. Montreal, Quebec, Canada. I(FEDSM-ICNMM2010–30368), 563–567. August 1–5.
Srinivasan, K. M. (2008). Rotodynamic pumps (1st ed.). New Delhi: New Age International (P) Ltd.
Taormina, R., & Chau, K. (2015). Neural network river forecasting with multi-objective fully informed particle swarm optimization. Journal of Hydroinformatics, 17(1), 99–113.
Viana, F. A. C. (2011). SURROGATES toolbox user’s guide version 3. Retrieved March 1, 2014, from http://sites.google.com/site/felipeacviana/surrogatestoolbox.
Viana, F. A. C., Simpson, T. W., Balabanov, V., & Toropov, V. (2014). Metamodeling in multidisciplinary design optimization: How far have we really come? AIAA Journal, 52(4), 670–690.
Wilcox, D. C. (1994). Turbulence modeling for CFD (2nd ed.). La Canada, California: DCW Industries, Inc.
Zerpa, L. E., Queipo, N. V., Pintos, T. S., & Salager, J. L. (2005). An optimization methodology of alkaline surfactant polymer flooding processes using field scale numerical simulation and multiple surrogates. Journal of Petroleum Science and Engineering, 47, 197–208.
Zhang, J., & Chau, K. (2009). Multilayer ensemble pruning via novel multi-sub-swarm particle swarm optimization. Journal of Universal Computer Science, 15(4), 840–858.