Optimization of a single-stage double-suction centrifugal pump

Aljaž Škerlavaj¹,², Mitja Morgut¹,², Dragica Jošt¹ and Enrico Nobile²
¹ Kolektor Turboinštitut d.d., Rovšnikova 7, 1210 Ljubljana (Slovenia)
² Dipartimento di Ingegneria e Architettura, Università degli Studi di Trieste, via A. Valerio 10, 34127 Trieste (Italy)
E-mail: aljaz.skerlavaj@kolektor.com

Abstract. In this study, the objective of the optimization of a double-suction pump is the maximization of its hydraulic efficiency. The optimization is performed, by means of the modeFRONTIER optimization platform, in steps. At first, by means of a DOE (Design of Experiments) strategy, the design space is explored, using a parameterized CAD representation of the pump. Suitable metamodels (surrogates or Response Surfaces), which represent an economical alternative to the more expensive 3D CFD model, are built and tested. Among different metamodels, the evolutionary design, radial basis function and the stepwise regression models seem to be the most promising ones. Finally, the stepwise regression model, trained on a set of 200 designs and constructed with only five the most influential input design parameters, was chosen as a potentially applicable metamodel.

1. Introduction
Centrifugal pumps are widely used in industrial applications. Double-suction (also called: double-entry) centrifugal pumps allow transportation of greater flow rates than single-entry pumps [1] because they are less prone to cavitation problems (smaller NPSHR). Another advantage is counter-balancing of axial hydraulic forces due to double-entry design [1]. The pump geometry considered in this study, together with a sketch of the impeller, is depicted in figure 1.

Usually, hydraulic turbomachines are designed by performing modifications of a nearly-suitable and experimentally validated baseline geometry. The modifications are either performed directly in a CAD software by the designer, or by modifying input variables in in-house codes using direct or inverse singularity methods. Nowadays, modified geometries are compared by means of Computational Fluid Dynamics (CFD) systems. Only the best design is then usually manufactured and experimentally tested. There are numerous benefits of using optimization techniques to perform modifications in automatized way instead of using traditional human-based ‘trial-and-error’ technique. First of all, no occasional hard-to-identify human-based errors are present, which guarantees the same design and CFD procedure for each design. Secondly, after the (large) initial effort is performed to include the design creation in an optimization loop, the human interaction is reduced and a lot of variants can be tested, which in turn allows a designer to understand the effect of the design variables on the objective (e.g., pump efficiency) much better. In turbomachines, usually multiple objectives have to be optimized. One of the first multi-objective optimization study was performed by Lipej and Poloni [2]. Due to importance of centrifugal pumps in industrial applications, as well as of pump-turbines, many recent studies
aim to optimize their geometry to improve their performance. Shingai et al. [3] performed a multi-objective optimization of pump-turbine, where the objectives were efficiency of turbine, pump efficiency, cavitation characteristics for both types of operation, torque to hold runner blades and total pump head. Xuhe et al. [4] optimized a pump turbine for efficiency in both modes of operation whereas Zhang et al. [5] performed optimization of a centrifugal pump for vibration, using fluid-structure interaction. The purpose of this work is to compare different RSM (Response Surface Methods) for the case of optimization of a centrifugal pump. We recall that RSM (also called metamodels or surrogates) represent an economical alternative to the more expensive 3D CAE models, CFD in this case. In particular, the following RSM algorithms were considered: polynomial singular value decomposition with the second order of polynomial degree, stepwise regression [6], K-nearest RSM [7], Kriging RSM [8], anisotropic Kriging [9], Gaussian processes [10], Hardy’s multiquadric radial basis function [11], neural networks [12] and evolutionary design [13]. In this study, the specific speed $n_q = n \sqrt{Q/2/H^{3/4}}$ of the pump was equal to 62 (specified per impeller side). Only the impeller geometry was modified while the rest of the geometry was considered as fixed, according to the limitations given in a commercial project, where only the impeller could be modified.

2. Numerical procedure

In figure 2, the logic of the optimization strategy, where the CFD simulations were performed to evaluate the different designs, is sketched.

For the impeller, the shape of the meridional channel was based on past experience and was not modified during the optimization procedure. On the other hand, the blade geometry was generated from a certain set of input parameters (figure 3) provided to an in-house code that uses an inverse-singularity method. The data was exported to ANSYS DesignModeler, where a 3-D geometry of an impeller was automatically created. For each geometry, a tetrahedral-based, unstructured mesh of the impeller was automatically created in ANSYS ICEM, which also provided the generation of 15 prism layers near the walls. The CFD model consisted of a single-channel impeller (using a rotational periodicity boundary condition) and a symmetric half of a volute with an outlet pipe, as depicted in figure 4. During the optimization, only the single-channel impeller mesh was being modified. At the impeller inlet, the velocity components were prescribed as a function of radius around rotation axis. The components were obtained
Figure 2. Logic of the numerical procedure.

Figure 3. Optimized geometric parameters (in red): (a) front view; (b) side view.

from a previously performed numerical simulation in a ‘traditional’ way (without optimization, using a full geometry model). At the outlet, average pressure boundary condition was used. A stage condition [14] was used as a frame change / mixing model at the general grid interface between the impeller and the volute. A symmetry condition was used to take into account both sides of the impeller and another half of the volute.

The ANSYS CFX software was used for CFD simulations. These were performed only for the design point, as steady-state single-phase simulations with shear-stress-transport $k-\omega$ turbulence model [15] with a curvature correction [16] (SST-CC). The SST model uses blending between the $k-\omega$ turbulence model close to walls and the $k-\varepsilon$ turbulence model elsewhere, while limiting the eddy viscosity at the boundary layer and the turbulence production term. The high-resolution advection scheme was used, which is an upwind adaptive scheme, based on the Barth and Jespersen’s limiter [17], and limits the numerical advection correction in order to suppress possible oscillations due to large gradients. The scheme tends to be locally as close as possible to the second order accuracy, and provides a compromise between robustness and accuracy. A local time-scale factor equal to 10 was used for 550 iterations. The result data (head, torque and efficiency) were averaged over the last 50 iterations to produce meaningful result. The
optimization process was driven by modeFRONTIER 2014 optimization platform. It is worth to remark that the objective of the optimization was the pump efficiency. Ten input variables were used to define the pump shape: length of the blade ($\varphi$ in figure 3) in polar coordinate system for hub, middle and shroud chord; three parameters to define the position and the shape of the impeller blade leading edge in the meridional view ($P_{LE,s}, P_{LE,h}$ and $\delta_{LE,h}$ in figure 3); two parameters to define the position of the blade trailing edge in the meridional view ($P_{TE,s}$ and $P_{TE,h}$ in figure 3); and two parameters that influence the impeller trailing edge angle inclination (in the polar view), at hub and at shroud ($\beta_2$ in figure 3). The length of the blade in polar coordinate system for middle chord was defined as an average of the length of the hub and the shroud chord, so only nine effective variables were used. Design of experiments (DOE), which fills the design space with the initial set of design variables, was based on the Latin Hypercube [19] samplings. The latter guarantees uniform random distribution of points over the variable range. The initial generation represented 20 sets of variables. Afterward, the Multi-Objective Genetic Algorithm (MOGA-II) [20] based on generational evolution was used for 10 generations. Probability of directional cross-over, probability of selection and probability of mutation were set to 0.5, 0.05 and 0.1, respectively. After performing CFD simulations on 200 cases, each one taking half an hour to complete, the design space was approximated by training a response surface method (RSM). Several RSM algorithms were evaluated for this particular case: polynomial singular value decomposition with the second order of polynomial degree (SVD2), stepwise regression (STEP), K-nearest RSM (KN), Kriging RSM (KR), anisotropic Kriging (AKR), Gaussian processes (GP), Hardy’s multiquadric radial basis function (RBF), neural networks (NN) and evolutionary design (ED). For the most promising RSM algorithms, a virtual optimization was performed. After the final verification, performed with CFD simulations, a suitable RSM was proposed.

3. Results and discussion

3.1. Optimization with MOGA-II algorithm

After the CFD simulations were performed on a set of initial design parameters (DOE), subsequent CFD simulations were performed during optimization process with MOGA-II algorithm for 10 generations. The objective of the optimization was overall (hydraulic) pump efficiency. The hydraulic efficiency of the pump is the ratio between the absorbed hydraulic power in fluid (velocity and pressure) and the provided power to the rotor, and is defined as $\eta_{\text{pump}} = \frac{Q \rho g H}{(M \omega)}$, where $Q$ is flow rate (in m$^3$/s), $\rho$ is fluid density (in kg/m$^3$), $g$ is standard acceleration due to gravity (in m/s$^2$), $H$ is pump head (in m), $M$ it torque on the shaft (in Nm) and $\omega$ is angular velocity (in rad/s). As the purpose of this work was to identify eventual RSM procedure for our future research, no constraints and additional objective variables were used. It
Table 1. Error estimate for different RSM: evolutionary design (ED), radial basis function (RBF), stepwise regression (STEP), polynomial singular value decomposition (SVD2), K-nearest (KN), neural networks (NN), Kriging (KR), anisotropic Kriging (AKR) and Gaussian processes (GP).

| RSM algorithm | ED   | RBF  | STEP  | SVD2 | KN   | NN   | KR   | AKR  | GP   |
|---------------|------|------|-------|------|------|------|------|------|------|
| Mean absolute error | 0.097 | 0.135 | 0.139 | 0.157 | 0.199 | 0.450 | 0.571 | 0.573 | 0.819 |

should be stressed that for a real-case optimization the pump head should be constrained, as the efficiency was increasing in parallel with increasing head, which was noticed also by Kim et al. [21]. The history chart of optimization procedure for the efficiency is presented in figure 5. The relative efficiency is defined as the ratio of actual efficiency to the efficiency of the first design. The efficiency of the first (baseline) design was close to the values of efficiency usually obtained by the first traditional design. During the first 20 designs, the highest relative efficiency of 100.89 % was obtained. During the MOGA-II optimization, the highest relative efficiency was equal to 101.12 %. The trend in figure 5 shows that efficiency could still be increased by further MOGA-II optimization process.

3.2. Training and validation of RSM algorithms

As described previously, numerous RSM algorithms, available within the modeFRONTIER optimization software, were trained on nine initial variables. For the training, 80 % of all previously computed designs were used (160 designs, out of which 23 were faulty). The trained singular value decomposition polynomial was of the second order (SVD2). For the validation of RSM algorithms, 20 % of all previously computed designs were used (40, out of which 5 were faulty). Results of validation are presented in Table 1.

In figure 6, relative efficiency predicted by four most promising algorithms is compared to the relative efficiency obtained by CFD simulations.

Although the most promising algorithm by means of absolute error was ED, it can be noticed in figure 6 that at large values of efficiency the curve seemed to be flat, meaning that it might not distinguish high values of efficiency from the lower ones, which might in turn make it useless for the virtual optimization process. On the contrary, in case of the RBF, STEP and SVD algorithms, the high efficiency values predicted by CFD resulted in high efficiency values predicted by these algorithms as well (higher than the rest of the curve). The ED algorithm was used for virtual
optimization for 40 generations (virtual designs between 200 and 1000), as presented in figure 7. The highest relative efficiency obtained was equal to 101.064 %, which is smaller than that obtained in the first 200 designs with CFD.

Nine designs with values of efficiency greater than 100.95 % with distinctly different input variable combinations were used for validation purposes, and this is presented in figure 6. It can be noticed that some of the points were quite accurately predicted, especially at the right-hand side of the curve. However, the highest values of efficiency was predicted for a different setup of input variables than the CFD simulations. It can be concluded that although the ED algorithm appeared as the best choice, it has to be trained on more than 160 CFD simulation cases. The largest noticed difference in efficiency between CFD and ED model was approximately 1 %. It is interesting to note that for one of the cases in [4], where RSM algorithm was trained on 40 cases, such difference was approximately 2.6 % for a case on a Pareto front. The ED model was improved by training it on 200 designs (marked as ED200) in figure 8. However, the results were similar as with training on 160 designs, which means that the same conclusions can be drawn.

The behavior of the other three RSM algorithms (RBF, STEP and SVD2) looks quite similar to figure 6. Therefore, the second-best model in table 1, the RBF algorithm, was tested. As evident from figure 9, the relative efficiency after 40 generations reached 101.38 %, which is higher than that obtained by ED algorithm in figure 7. The reason might be due to the fact that in figure 6 the RBF algorithm was able to generate higher values of efficiency than the ED. It
should be noted that the efficiency predicted by RBF is also higher than that obtained during the first 200 designs.

Nine designs with values of efficiency greater than 101.28 % with different input variable combinations were used for validation purposes. Results are presented in figure 10. In general, the results represent almost a flat curve over the observed range of efficiency obtained by CFD results. Secondly, the values (of curve RBF) are much higher than those obtained by CFD, from 1.2 % to 1.84 %. The results improved when RBF was trained on 200 designs instead of on 160 CFD simulations (curve RBF200). In addition, when only the five most influential input variables were used for training on 200 designs (curve RBF200_5var), the results seemed to improve.

The last algorithm tested was STEP algorithm. After 40 generations the highest relative efficiency obtained was equal to 101.35 (figure 11), which is similar to that of the RBF.

From the validation with CFD results (figure 12), similar conclusions as for RBF function can be drawn: the curve STEP is flat and too high. Trained on 200 designs, the curve STEP200 is closer to CFD simulations. However, when only 5 design variables were used for training on 200 designs, the result was improved considerably (curve STEP200_5var).

Finally, in figure 13 the most promising STEP and RBF models were compared to all CFD simulations that were performed in figures 8, 10 and 12. It can be noticed that the prediction for both algorithms improved when only 5 variables were used for training. In addition, it seems that the STEP200_5var curve follows the CFD curve very well, which means that it has a potential
Figure 10. Validation of the RBF algorithm.

Figure 11. Virtual optimization of efficiency with the STEP algorithm.

Figure 12. Validation of the STEP algorithm.

to be used for virtual optimization of the pump case with 200 initial design points simulated with CFD.

The efficiency prediction of the RSM STEP algorithm, trained on a set of 200 designs by using five variables, could not be improved by using just the high-efficiency values in training phase. Even when only three low-efficiency designs were discarded (when efficiency was smaller than 95%), as presented by the curve STEP200_5var_95 in figure 12, large discrepancy occurred at
efficiency value 99.94 %. The predicted value was approximately 100.8 %, which is very close to the highest efficiency values at the right side of the curve. Such discrepancies might lead to improper virtually-optimized designs.

Finally, the initial (baseline) and optimized geometries of the impeller, characterized by 100 % and 100.9 % relative efficiencies respectively, are illustrated in figure 14.

![Figure 13. Comparison of STEP and RBF models with CFD results.](image)

![Figure 14. Impeller geometries. Left: initial (baseline); Right: optimized](image)

The results obtained by the simplified steady CFD model were confirmed by running a full 3D unsteady simulation, using the SAS-SST (Shear Stress Transport Scale Adaptive Simulation) DES-like turbulence model of Egorov and Menter [22].

The experimental validation, by measuring the performances of the optimized impeller, is in progress at Kolektor Turboinštitut.

4. Conclusions
Several metamodels, which represent an economical alternative to the more expensive 3D CFD model, were built and tested. Based on comparison with CFD simulations, the following conclusions were drawn:

- The most promising metamodels were the ED, the RBF and the STEP;
- A set of 200 designs was not sufficient for the three RSM, when built on 9 design variables;
- For such set, the STEP metamodel, built on five most influential input design variables, seemed a potentially applicable metamodel for virtual optimization of a centrifugal pump;
- Prediction of metamodels was improved by increasing the number of initial (CFD) designs and by building/training the metamodels on only the most influential design variables;
• The five most influential design variables for this case, by order of importance, were: blade trailing edge inclination angle at hub, location of leading edge at hub in the meridional view, length of the blade at hub in the polar view, location of leading edge at shroud in the meridional view, length from the axis of rotation to the impeller blade trailing edge, at hub.

• Discarding low-efficiency designs in the training phase of the metamodels can lead to false high-efficiency design predictions.

Acknowledgments
This research has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 612279 and from the Slovenian Research Agency ARRS - Contract No. 1000-15-0263.

References
[1] Gülich JF 2008 Centrifugal Pumps (Berlin: Springer-Verlag) p 926
[2] Lipej A and Poloni C 2000 Design of Kaplan runner using multiobjective genetic algorithm optimization J. Hydraul. Res. 38(4) pp 73-9
[3] Shingai K, Omura Y, Tani K, Tominaga H and Ito S 2008 Multi-objective optimization of diagonal flow type reversible pump-turbines Proc. 24th Symp. on Hydraulic Machinery and Systems 27-31 Foz do Iguaçu (Madrid: IAH) [4] Xuhe W, Baoshan Z, Lei T, Jie Z and Shuliang C 2014 Development of a pump-turbine runner based on multiobjective optimization J. Phys. Conf. Series: Earth Environ. Sci. 22 012028
[5] Zhang Y, Hu S, Zhang Y, Chen L 2014 Optimization and analysis of centrifugal pump considering fluid-structure interaction Scientific World J. 2014 131802
[6] Venables WN and Ripley BD 2002 Modern Applied Statistics with S (New York: Springer-Verlag) p 498
[7] Allef P 1989 Scattered data interpolation in three or more variables Mathematical Methods in Computer Aided Geometric Design ed T Lyche and LL Schumaker (New York: Academic Press) pp 1-34
[8] Deutsch CV and Journel AG 1992 GSLIB: Geostatistical Software Library and User’s Guide ed AG Journel and A Press (New York: Oxford University Press) p 340
[9] Rasmussen CE and Williams CKI 2006 Gaussian Processes for Machine Learning ed T Dietterich, C Bishop et al (Cambridge: MIT Press) p 262
[10] Gibbs M and MacKay DJC 1997 Efficient Implementation of Gaussian Processes (Cambridge: Cavendish Laboratory) p 17
[11] Buhmann MD 2003 Radial Basis Functions: Theory and Implementations (Cambridge: Cambridge University Press)
[12] Hagan MT and Menhaj MB 1994 Training feedforward networks with the Marquardt algorithm IEEE Trans. on Neural Networks 5 989-93
[13] Fillon C 2008 New Strategies for Efficient and Practical Genetic Programming (Trieste: Università degli Studi di Trieste) p 145
[14] ANSYS 2016 ANSYS CFX-Solver Theory Guide (ANSYS: Canonsburg) p 362
[15] Menter F R Kuntz M and Langtry R 2003 Ten years of industrial experience with the SST turbulence model Turbulence, Heat and Mass Transfer 4. ed Hanjalić K Nagano Y and Tummers M (Begell House, New York) pp 625-632.
[16] Smirnov P E and Menter F 2009 Sensitization of the SST turbulence model to rotation and curvature by applying the Spalart-Allmaras correction term. J. Turbomachinery 131, no. 4, p 041010.
[17] Barth T J and Jespersen D C 1989 The design and application of upwind schemes on unstructured meshes 27th Aerospace Sciences Meeting AIAA Paper 89-0366.
[18] ESTECO 1999-2016 modeFRONTIER User Manual (ESTECO: Trieste)p 504
[19] McKay MD, Conover WJ and Beckman RJ 1979 Latin hypercube sampling: a comparison of three methods for selecting values of input variables in the analysis of output from a computer code Technometrics 21 239-45
[20] Poles S 2003 MOGA-II An improved Multi-Objective Genetic Algorithm (Trieste: ESTECO)
[21] Kim JH, Oh KT, Pyun KB, Kim CK, Choi YS and Yoon JY 2012 Design optimization of a centrifugal pump impeller and volute using computational fluid dynamics J. Phys. Conf. Series: Earth Environ. Sci. 15 032025
[22] Egorov Y and Menter F 2008 Development and application of SST-SAS turbulence model in the DESIDER project. Advances in Hybrid RANS-LES Modelling ed Peng S-H and Haase W (Springer, Heidelberg), pp 261-270.