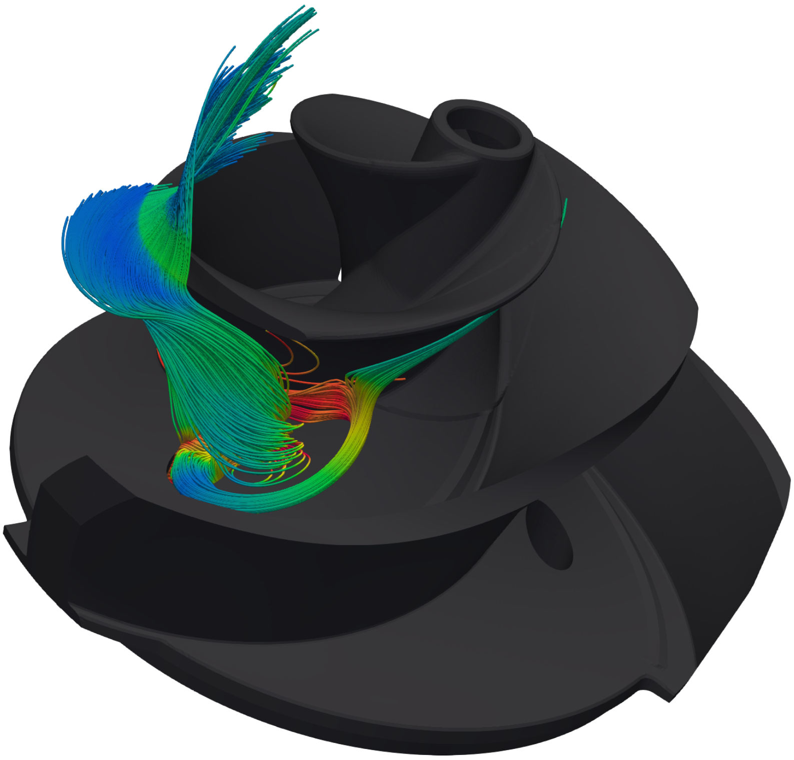


Design Optimization of Balancing Holes in a Centrifugal Pump

Master's thesis in Applied Mechanics

BALA KUMARESH THILEEP KUMAR

DEPARTMENT OF MECHANICS AND MARITIME SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Design Optimization of Balancing Holes in a Centrifugal Pump

BALA KUMARESH THILEEP KUMAR

Department of Mechanics and Maritime Sciences
Division of Fluid Dynamics
Chalmers University of Technology
Gothenburg, Sweden 2025

Design Optimization of Balancing Holes in a Centrifugal Pump
BALA KUMARESH THILEEP KUMAR

Examiner:
Håkan Nilsson
Professor at Division of Fluid Dynamics
Department of Mechanics and Maritime Sciences
Chalmers University of Technology

Industrial Supervisor:
Arash Soltani
Christian Wollblad
Xylem Water Solutions Sweden

Master's Thesis 2025
Department of Mechanics and Maritime Sciences
Division of Fluid Dynamics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Streamline visualization generated in ParaView, illustrating internal jet-like flow behavior through balancing holes in a semi-open centrifugal pump impeller. Streamlines are colored by relative flow angle magnitude (β).

Typeset in LaTeX
Gothenburg, Sweden 2025

Design Optimization of Balancing Holes in a Centrifugal Pump
BALA KUMARESH THILEEP KUMAR
Department of Mechanics and Maritime
Chalmers University of Technology

Abstract

Centrifugal pumps are widely used in fluid transport systems due to their ability to deliver high flow rates in a compact, low-maintenance design. Despite their well-defined performance characteristics, a major operational challenge is the generation of axial force, a force imbalance caused by the pressure difference across the front (suction) and back (pressure) sides of the impeller. If the axial force is not properly investigated during the design process, it can cause bearing failure, increase maintenance requirements, and significantly reduce both hydraulic efficiency and pump lifespan. Balancing holes are commonly introduced to mitigate axial force by reducing the pressure imbalance. While effective in reducing axial force, balancing holes introduce fluid recirculation from the pressure side to the suction side of the impeller, leading to a reduction in hydraulic performance. Therefore, identifying an optimal balancing hole configuration is essential to minimize axial force without compromising overall efficiency.

This study investigates the development of a multi-objective target function using both scalarization and Pareto-based methods to identify an optimal balancing hole design configuration that reduces axial force while maintaining the hydraulic performance. CFD simulations were carried out using a steady-state approach with implicit local time stepping, in conjunction with a non-linear $k-\varepsilon$ turbulence model and the frozen rotor approach, to capture the internal flow behaviour and to extract performance metrics—head, power, and axial force—for all generated configurations during the optimization process. To explore the design space effectively, Sobol sequence sampling was used to generate a diverse set of initial configurations. The target function was formulated by normalizing head and power at the Best Efficiency Point (BEP) and axial force across part-load, BEP, and over-load conditions, then iteratively optimized using Bayesian optimization and target function scoring to identify an optimal balancing hole design. A Pareto front analysis was finally conducted to analyze the trade-offs between hydraulic performance and axial force reduction for the generated configurations throughout this study.

The findings revealed that the formulated target function using scalarized method effectively guided the optimization process toward configurations that reduced axial force by up to 40.59 % while maintaining head and power within 99.98 % and 99.87 % of the reference, respectively. Iterative Bayesian optimization and Pareto front analysis consistently led to a key design feature, which was also supported by the flow field visualization.

Keywords: Balancing hole, Multi-objective target function, Sobol sequence sampling, Bayesian optimization, Pareto front analysis.

Acknowledgements

First and foremost, I would like to express my heartfelt gratitude to my supervisors, Arash Soltani and Christian Wollblad, at Xylem for their continuous support, encouragement, and guidance throughout my master's thesis. From helping me with fundamental details to complex decision-making, their support has been truly instrumental in every stage of this journey. Their mentorship went beyond technical supervision—they were always available and patient, and I am deeply thankful for the knowledge and experience I gained working alongside them.

I would also like to sincerely thank Fredrik Söderlund, my manager at Xylem, for providing me this incredible opportunity and for fostering a supportive environment throughout my time at the company. I am grateful to Tobias Pilstrand for his valuable assistance with Bayesian optimization techniques and for always being approachable when I needed help. A special thanks to Viktor Bredvad for providing the impeller geometry and helping me understand the design aspects involved—his support enabled me to create multiple geometries in a short time, which was crucial for generating a wide range of configurations used in this thesis.

My thanks extend to the entire Xylem team for welcoming me and for their technical insights, discussions, and encouragement, which made my thesis work both meaningful and enjoyable.

I would also like to thank Chalmers University of Technology for providing an excellent academic environment and resources that supported my learning and growth throughout the program. I am grateful to Professor Håkan Nilsson, my examiner, for his valuable feedback and for supporting the academic evaluation of this thesis. A special mention goes to Professor Hudong Yao, whose referral played a key role in helping me secure this thesis opportunity. I would also like to thank the entire Department of Mechanics and Maritime Sciences and all the professors whose courses helped build the foundation for this work.

Finally, I wish to thank my family and friends for their constant encouragement, support, and belief in me during this wonderful journey. And last but not least, I would like to acknowledge myself—for staying committed, believing in my abilities, and pushing through every challenge along the way.

Bala Kumaresh Thileep Kumar, Gothenburg, June 2025

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

BEP	Best Efficiency Point
BH	Balancing Hole
CDF	Cumulative Distribution Function
CFD	Computational Fluid Dynamics
CRPS	Continuous Ranked Probability Score
EI	Expected Improvement
PDF	Probability Density Function
RSE	Relative Squared Error
TF	Target Function

Nomenclature

Below is the nomenclature of symbols that have been used throughout this thesis.

Symbols

Centrifugal Pump

$\dot{W}_c$	Impeller power (work rate on the fluid)
τ_A	Torque acting on the impeller
Ω	Angular velocity of the impeller
$\dot{m}$	Mass flow rate
U_1	Blade speed at inlet
U_2	Blade speed at outlet
$c_{\theta 1}$	Tangential component of the absolute velocity at inlet
$c_{\theta 2}$	Tangential component of the absolute velocity at outlet
ΔW	Specific work (work done on fluid per unit mass)
H_{theo}	Theoretical head developed by the pump
H	Total head developed by the pump
g	Acceleration due to gravity
P_{Shaft}	Shaft power
$P_{\text{Hydraulic}}$	Hydraulic power imparted to the fluid
ρ	Density of the fluid
η	Hydraulic efficiency of the pump
Q	Volumetric flow rate
Q^*	Normalized flow rate
Q_{BEP}	Flow rate at BEP condition
P_2	Pressure at outer radius of the impeller
H_{p2}	Static pressure
C_2	Absolute velocity of the liquid at impeller outlet

F	Axial force acting on the impeller
T	Net axial force acting along the pump shaft
F_1	Pressure force on the rear shroud
F_2	Pressure force on the impeller hub
F_3	Force acting across the blade tip clearance
F_4	Dynamic reaction force at the impeller inlet
$A_{\text{Clearance}}$	Clearance area of the impeller
D_{Mean}	Mean diameter of the impeller
s	Radial clearance between rotating and stationary components
D_{min}	Minimum diameter of the balancing hole
D_{max}	Maximum diameter of the balancing hole
A_{Total}	Total area of the balancing holes
$A_{\text{per, hole}}$	Area of balancing hole

Computational Fluid Dynamics

k	Turbulent kinetic energy
ε	Turbulent dissipation rate
μ	Dynamic viscosity
v_i	Instantaneous velocity component
p	Instantaneous pressure
ν_t	Eddy viscosity
S_{ij}	Mean strain rate tensor
a_{ij}	Anisotropic part of the Reynolds stress tensor
δ_{ij}	Kronecker delta

Multi-Objective Optimization

N	Number of samples
$f_i(x_j)$	Value of i^{th} objective for j^{th} sample
$f_{i,ref}$	Reference value for the i^{th} objective
Δ_i	Mean absolute deviation
I_i	Inverse of the mean absolute deviation
w_i	Normalized weight assigned to i^{th} objective
m	Number of objectives

I_k	Inverse deviation of k^{th} objective
$f_i(x)$	i^{th} objective function
$F(x)$	Scalarized objective function
$y_{\text{ref, i}}$	Reference value for i^{th} data point
$\hat{y}_i$	Observed or simulated value for the i^{th} data point
$\overline{y}_{\text{ref}}$	Mean of all reference values
n	Total number of data points
y_i	Observed or simulated value for the i^{th} configuration
$f(x)$	Objective function
$\mu(x)$	Mean function
$\mathbb{E}[f(x)]$	Expected value
$K(x, y)$	Covariance function (or kernel)
k	Variance parameter
σ	Length-scale parameter of the function
Z	Standardized improvement
f^*	Best value observed
$\Phi(Z)$	Cumulative distribution function
$\phi(Z)$	Probability density function
D_N	Discrepancy of the first N Sobol points
θ	Model parameters
$\mathcal{D}$	Observed dataset
x_*	New input
y_*	Predicted output at the new input x_*
X_i	Simulated result of the quantity (Head or Power) for i^{th} impeller configuration with balancing hole at BEP ($Q^* = 1.0$) condition
X_{ref}	Simulated result of the quantity (Head or Power) for the reference impeller without balancing holes at BEP ($Q^* = 1.0$) condition
σ_X	Standard deviation of the simulated results (Head or Power) across the 12 configurations generated using Sobol sequence at BEP ($Q^* = 1.0$) condition
F_i	Simulated axial force for the i^{th} impeller configuration with balancing hole at BEP ($Q^* = 1.0$) condition
$F_{\text{ref, i}}$	Simulated axial force for the reference impeller without balancing hole at BEP ($Q^* = 1.0$) condition
$F_{\text{i, 0.5}}$	Simulated axial force for the i^{th} impeller configuration with balancing hole at part-load ($Q^* = 0.5$) condition

$F_{\text{i, 1.0}}$	Simulated axial force for the i^{th} impeller configuration with balancing hole at part-load ($Q^* = 1.0$) condition
$F_{\text{ref, 0.5}}$	Simulated axial force for the reference impeller without balancing hole at BEP ($Q^* = 0.5$) condition
$\text{Normalized}_{\text{P}}$	Normalized value of power
$\text{Normalized}_{\text{H}}$	Normalized value of head
$\text{Normalized}_{\text{AF}}$	Normalized value of axial force
Inv_P	Inverse deviation for power
Inv_H	Inverse deviation for head
Inv_{AF}	Inverse deviation for axial force
w_P	Weightage value for power
w_H	Weightage value for head
w_{AF}	Weightage value for axial force
$TF_{\text{Predicted}}$	Predicted target function score obtained using Bayesian optimization
$TF_{\text{Scalarized}}$	Target function score obtained using the scalarized form
$(\Delta\eta)_{\text{max}}$	Maximum efficiency loss
η_{Ref}	Efficiency for the reference impeller
η_i	Efficiency for the i^{th} impeller configuration with balancing hole

Contents

1

Introduction

In the modern world of fluid transport, few machines are as indispensable as pumps. They serve as the driving force behind countless industrial and domestic systems, ensuring the efficient and reliable liquid transport through complex networks of pipelines and processing units. From delivering clean water to households and wastewater management to supporting the rigorous demands of industrial operations, pumps play a significant role in contemporary infrastructure. Pumps function by transforming mechanical energy into fluid energy and operate primarily through two mechanisms: fixed-volume displacement (positive displacement pumps) or continuous kinetic energy transfer (dynamic pumps). Among the various types of dynamic pumps, centrifugal pumps stand out due to their superior performance, compact design, and ability to handle high flow rates with relatively low maintenance requirements. Their broad range of applications—from municipal water systems to high-demand industrial processes—underscores their critical role in modern engineering and fluid transport systems.

As previously noted, centrifugal pumps operate by converting mechanical energy into hydraulic energy. In this process, the fluid gains kinetic energy through the rotating impeller, which is then gradually transformed into pressure energy as the fluid moves through the pump. Despite their well-defined performance characteristics (like efficiency and handling high flow rates), centrifugal pumps face certain operational challenges, with this study focusing specifically on axial loads. These loads result from a pressure imbalance between the front and back sides of the impeller. If not addressed properly, axial loads can cause bearing failures, increased maintenance requirements, and a decline in both pump performance and service life [1].

Among the various options available to reduce the axial load, Shepherd [2] considered the use of balancing holes to be the simplest and cheapest option. These holes connect the high-pressure region behind the impeller to the low-pressure region at the front, allowing a portion of the fluid to recirculate [3]. While this pressure equalization helps reduce axial force, it comes at the cost of hydraulic performance. The recirculated fluid introduces energy losses, which can negatively affect head generation, power consumption, and overall efficiency [4].

The central engineering challenge lies in optimizing the design of balancing holes to effectively reduce axial load without significantly compromising pump performance. Achieving this balance is complex and depends heavily on the configuration of the balancing holes—including their number, diameter [5], and placement—as well as the intricate internal flow behavior within the pump. As industries continue to demand higher-performing pumps, the refinement of balancing hole design remains a key focus in the development of modern centrifugal pump technology.

1.1 Problem Statement

Despite their widespread application, there are no robust methods for optimally designing balancing holes that both reduce the axial load and preserve the overall pump performance. In industries, balancing holes are drilled based on guesswork [6], which leads to lower pump performance without significantly reducing the axial force [4, 5]. These limitations highlight the need for a more advanced and reliable optimization framework —one that leverages CFD techniques and performance-driven algorithms to optimize the hole configuration for improved pump performance.

Previous studies by Xylem have adapted similar techniques to address these challenges, providing valuable insights into the flow behavior and design trade-offs. However, they encountered difficulties in defining a comprehensive multi-objective target function capable of simultaneously minimizing axial load and maintaining critical pump performance metrics. This thesis addresses that gap by formulating a customized multi-objective target function specifically for the Flygt N-centrifugal pump, designed and developed by Xylem, aiming to overcome those limitations.

1.2 Objective

Continuing the work initiated by Xylem, this thesis aims to formulate a multi-objective target function that supports the identification of optimal balancing hole design configurations. Instead of delivering a single optimal design for the pump considered in this study, the primary objective is to establish a consistent framework using both scalarization and Pareto-based methods to evaluate and compare multiple design configurations. This formulated function is intended to facilitate future optimization efforts by targeting maximum possible axial force reduction while maintaining the overall hydraulic performance of the pump. To achieve this, detailed CFD simulations will be conducted to analyze the hydraulic performance of the generated configurations throughout the optimization process. Based on these, the objectives of this study are classified into primary and secondary objectives:

- **Primary Objective:** Develop a multi-objective target function that integrates the axial force and the key performance metrics (Head, Power Consumption & Efficiency) at the Best Efficiency Point (BEP) of the pump. Evaluate the effectiveness of the formulated target function by analyzing the generated configurations at BEP but also across a range of operating condition. Iteratively refine the target function until the selected configuration consistently maintains the pump's characteristic performance at all investigated operating conditions while minimizing axial force.

- **Secondary Objective:** Analyze CFD simulation results of the generated configurations, to gain insights into internal flow phenomena with and without balancing holes. The focus is to provide a physical insight by evaluating pressure distribution and relevant hydrodynamic effects to identify the relationship between design configurations of balancing holes and their influence on axial force and performance metrics of the pump.

1.3 Limitations of the Study

This study is subject to the following limitations based on the defined scope, computational resources, and time constraints:

- The development of the multi-objective target function and the associated analysis results are specific to the geometry and operating range of the centrifugal pump with a vaneless diffuser configuration. Generalization to other pump series or types is beyond the scope of this work.

- All simulations are carried out under steady-state condition using the Multiple Reference Frame model with the Frozen Rotor approach. Transient simulations are not conducted in this study for CFD simulations.

- The study exclusively focuses on the reduction of axial forces using optimization of balancing hole configuration. Radial forces are not considered.

- Cavitation is not considered in this study. The simulations assume single phase, incompressible flow.

- Temperature effects and thermal-related fluid properties are not included in the simulations. All analyses are performed under isothermal conditions.

- The study does not include structural evaluations such as stress analysis, fatigue assessment, or deformation of the impeller or surrounding regions near the balancing holes.

- A mesh independence study has been carried out in the region around the balancing holes. A full domain mesh independence study for the computational domain was not carried out, as the baseline mesh without balancing holes provided by Xylem was used without modification.

- Validation of the optimized target function results was not possible due to the absence of experimental or real-world data. However, the baseline configuration without balancing hole is validated against the transient CFD simulation results previously conducted by Xylem which in turn have been validated against measured performance.

These limitations are acknowledged to clarify the scope of the study, which focuses on analyzing the hydraulic performance and internal flow behavior associated with balancing hole configurations under defined and controlled conditions.

2

Theory

This chapter outlines the theoretical concepts relevant to this study and is organized into four main sections: the fundamentals of centrifugal pump, including the energy transfer process, the development of axial forces in a centrifugal pump, an general overview of CFD and turbulence modeling, and the frame-work for multi-objective optimization.

2.1 Centrifugal Pump

Centrifugal pumps are hydraulically driven turbomachines designed to transport fluids, primarily liquids, by converting mechanical energy into hydraulic energy through the action of centrifugal force [7]. Their main function is to transfer the fluid by increasing its velocity and pressure.

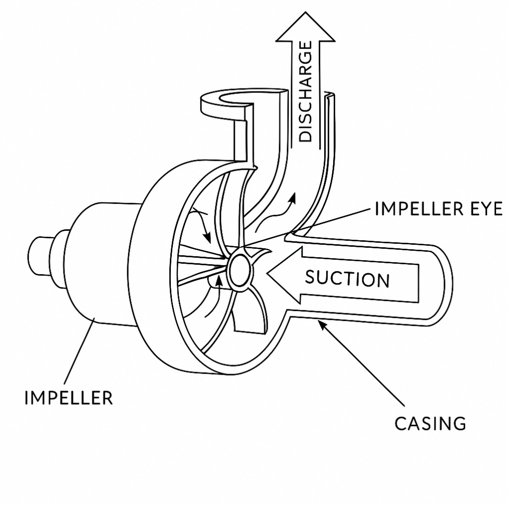


Figure 2.1: Schematic representation of a centrifugal pump. Generated using ChatGPT.

A typical centrifugal pump consists of an impeller, suction nozzle, discharge nozzle and pump casing [8] as shown in figure 2.1. The impeller comprises a series of blades, increases both the static pressure and velocity of the incoming fluid as it passes through during rotation [9]. The pump casing (either a diffuser or volute type) encloses the impeller and converts the fluid's kinetic energy into pressure energy. Additionally, the casing is also equipped with suction and discharge sections to guide the flow into and out of the pump efficiently. Despite various types of centrifugal pumps, the fundamental operating principle remains consistent, which is discussed in detail in the following section.

2.1.1 Working Principle

The fundamental working principle of a centrifugal pump is based on the centrifugal force generated by the rotating motion of the impeller. Fluid enters the pumps through the suction nozzle at the center of the impeller (eye of the impeller). As the fluid enters, the rotational motion of the impeller imparts kinetic energy onto the fluid, causing it to accelerate. Due to centrifugal force, which is constantly acting on the fluid particle, the fluid is pushed towards the outer edge of the impeller, thereby increasing its kinetic energy [7]. The curvature and the orientation—especially the arrangement of impeller blades—guide the fluid that is being pushed by the centrifugal force smoothly, thereby minimizing energy loss. Simultaneously, due to the geometry of the impeller channel, part of this kinetic energy is converted into pressure energy. By the time the fluid reaches the outlet of the impeller, its velocity is high, and its pressure has also increased substantially [3]. Upon leaving the impeller, the fluid enters the pump casing. The pump casing is designed in such a way to collect the fluid discharged from the impeller and convert part of the remaining fluid's kinetic energy into pressure energy, typically using a volute type or diffuser type design [3]. In a volute type, the gradually expanding spiral shape of the pump casing reduces the velocity of the fluid leaving the impeller, thereby increasing the fluid pressure towards the discharge. In contrast, the diffuser casing uses a stationary set of guide vanes surrounding the impeller to achieve the same effect by by gradually decelerating the fluid. According to Bernoulli's principle, this reduction in velocity results in an increase in static pressure, allowing for efficient pressure recovery before the fluid exits the pump [9].

2.1.2 Classification of Centrifugal Pump considered in this study

Centrifugal pumps are classified based on several factors, including their application, flow direction, impeller design, and the number of stages [8]. This section describes the classification of the centrifugal pump considered in this study.

- **Flow Type:** Radial: The fluid enters the pump axially, along the shaft axis, and is discharged perpendicularly to the direction of the shaft axis. This is a common configuration for pumps handling moderate to high heads.

- **Impeller Design:** Semi-Open Impeller: The impeller blades are attached to a hub on one side, while the other side is exposed to the flow.

- **Stage Count:** Single Stage: A single impeller is mounted on the pump shaft, meaning all energy transfer occurs within one stage. Single-stage pumps are typically used in applications where moderate pressure rise is sufficient.

- **Shaft Orientation:** Vertical: The pump shaft is vertically oriented with respect to the ground plane. This configuration is commonly used in applications with limited horizontal space or where submersible installation is required.

- **Casing Type:** Vaneless Diffuser: The pump casing features a vaneless diffuser design, characterized by an open annular passage without stationary guide vanes.

2.1.3 Energy Transfer within a Centrifugal Pump

The energy transfer within a centrifugal pump can be mathematically described using fundamental principles of fluid mechanics and rotational dynamics. The conservation law, particularly the conservation of angular momentum, relates the amount of torque required to impart a change in angular momentum to the fluid, ultimately leading to the fundamental Euler's equation for turbomachines [9]. This equation governs the energy transfer between the rotating component and the working fluid within a centrifugal pump.

Based on this, Euler's Equation 2.1 quantifies the amount of work done on the fluid by the rotating impeller. The impeller's work rate ($\dot{W}_c$) on the fluid is defined as the product of torque and angular velocity:

$$\dot{W}_c = \tau_A \Omega = \dot{m}(u_2 c_{\theta 2} - u_1 c_{\theta 1}) \tag{2.1}$$

where, the blade speed $u = \omega \cdot r$, τ_A is the torque exerted by the impeller, Ω is the angular velocity of the impeller, $\dot{m}$ is the mass flow rate, u_1 and u_2 are the blade speeds at the inlet and outlet respectively, and $c_{\theta 1}$ and $c_{\theta 2}$ are the tangential components of the absolute velocity at the inlet and outlet respectively.

Dividing Equation 2.1 by the mass flow rate gives the specific work (or) work done on the fluid per unit mass. Equation 2.2 is referred to as the Euler's pump equation [9].

$$\Delta W = \frac{\dot{W}}{\dot{m}} = \frac{\tau_A \Omega}{\dot{m}} = U_2 c_{\theta 2} - U_1 c_{\theta 1} \tag{2.2}$$

2.1.3.1 Head Developed

The specific work expressed in Equation 2.2 can be expressed in terms of head, which represents the energy transferred to the fluid as an equivalent head. By dividing the specific work from Euler's equation by the gravitational acceleration, the theoretical head developed by the pump can be expressed as:

$$H_{\text{theo}} = \frac{\Delta W}{g} = \frac{U_2 c_{\theta 2} - U_1 c_{\theta 1}}{g} \tag{2.3}$$

Equation 2.3 represents the theoretical estimate of the head developed by the pump based solely on the impeller geometry and the tangential components of the absolute velocity at the inlet and outlet, meaning no energy losses due to friction, turbulence, leakage, or other secondary effects [9].

2.1.3.2 Power Consumption & Efficiency

The power consumption of a centrifugal pump refers to the mechanical energy required to drive the pump shaft. It is referred to as shaft power, and it is greater than the hydraulic power due to the presence of various losses within the pump, such as mechanical, volumetric, and hydraulic losses [9]. Equation 2.4 expresses P_{shaft} the as:

$$P_{\text{shaft}} = \frac{P_{\text{hydraulic}}}{\eta} \tag{2.4}$$

where η is the pump's hydraulic efficiency. The hydraulic power imparted to the fluid, which represents the useful power transmitted through the fluid to overcome the head [9], is defined by Equation 2.5:

$$P_{\text{hydraulic}} = \rho g Q H \tag{2.5}$$

where ρ is the density of the fluid, g is the gravitational acceleration, Q is the volumetric flow rate, and H is the total head developed by the pump.

2.1.4 Performance Characteristics of a Centrifugal Pump

In a centrifugal pump, energy transfer is accomplished by changing the state of motion of the liquid as it passes through the rotating impeller. The magnitude of this energy transfer depends upon the geometry of the impeller and the pump's operating conditions [8]. While the impeller geometry and the rotational speed of the pump remain constant, the fluid velocity varies with changes in flow rate. Pumps often are subjected to operate under varying conditions to meet different system demands, making it essential to evaluate their performance across a range of flow rates. This evaluation is carried out using performance characteristics, which describe how the pump responds to different flow rates in terms of head development, power consumption, efficiency, and Net Positive Suction Head (NPSH) required to avoid cavitation. In this study, the NPSH is excluded from the analysis, as the focus is on optimizing balancing hole design based on internal flow behaviour and its influence on axial force and hydraulic performance. To enable consistent comparison, the flow rate in the performance curves is normalized relative to the flow rate corresponding to the BEP, as defined in Equation 2.6. Each performance characteristics curve is briefly explained in the upcoming subsections to provide a theoretical background of the parameters evaluated in this study.

$$Q^{*} = \frac{Q}{Q_{\text{BEP}}} \tag{2.6}$$

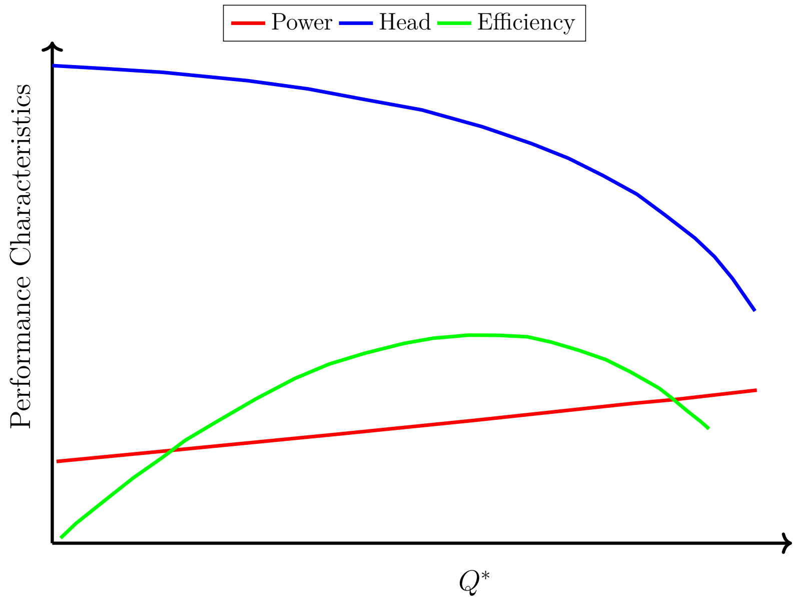


Figure 2.2: Typical performance chart of a centrifugal pump.

2.1.4.1 $Q^* - H$ Curve

The $Q^* - H$ curve represents the amount of head developed by the pump for different flow rates. As illustrated in Figure 2.2, this curve typically exhibits a negative slope, indicating that the head decreases as the flow rate increases. This decline results from increasing internal energy losses within the pump at higher flow rates. Evaluating this curve helps in understanding the pump's capability to develop the desired head under different operating conditions.

2.1.4.2 $Q^* - P$ Curve

The $Q^* - P$ curve describes the power consumed by the pump (i.e., shaft power) at different flow rates. As seen in Figure 2.2, this curve follows a positive trend, where the power consumption increases with an increase in flow rate. This increase is due to the fact that a pump requires higher power to transfer greater volume of fluid . Evaluating the Q–P curve is essential for selecting an appropriately sized motor and estimating power requirements under various operating scenarios.

2.1.4.3 Q^*-η Curve

The Q^*-η curve represents the proportion of mechanical energy converted into useful hydraulic energy. As shown in Figure 2.2, this curve typically exhibits a peak: efficiency increases with flow rate up to a maximum point—the BEP—and then declines beyond it. Identifying this peak is crucial for determining the pump's optimal operating point and maximizing overall performance.

2.2 Forces Acting on a Centrifugal Pump

Centrifugal pumps, like all rotating turbomachines, are subjected to various fluid dynamic forces. These forces arise inherently from the energy transfer process (via angular momentum exchange between the impeller and the working fluid) as well as from internal flow dynamics (including pressure imbalance, momentum changes and unsteady flow effects induced by blade-passage interactions). In a rotating impeller, these forces act in both radial and axial directions and may also include transient components from such interactions.

While all of these forces can negatively impact pump reliability and operational lifespan, this study focuses on axial force due to its critical significance in causing bearing failure and reduction in pump life [1]. The following section presents a detailed analysis of axial force generation, emphasizing the fluid dynamic mechanisms involved and the associated mathematical formulations.

2.2.1 Axial Force on Semi–Open Impeller

Axial force acting on an impeller, regardless of configuration, arises from pressure imbalance between its front (suction) side and the back (pressure) side of the impeller. A mathematical understanding of this pressure imbalance begins with the head developed by the pump as described in Equation 2.3. This head represents the total energy per unit weight imparted to the fluid, comprising both the static pressure head (H_{p2}) and dynamic (velocity) head ($C_2^2/2g$), where C_2 is the absolute velocity of the liquid at the impeller outlet.

Specifically, the pressure P_2 at the impeller outlet is equal to H_{p2}. This pressure acts in all directions, including within the cavity between the impeller hub and pump casing. However, due to the rotational motion of the liquid in this space, the pressure acting on the back side of the hub is not uniform and varies with the radius. For practical purposes, it is often assumed that the liquid within that space rotates with half the angular velocity of the impeller. Using this assumption, the pressure distribution on either side of the semi–open impeller can be derived by applying Bernoulli's equation in a rotating reference frame. This pressure distribution varies radially inward towards the pump shaft as illustrated in the upper part of Figure 2.3, due to the reduction in the circumferential velocity resulting from centrifugal effects, and can be expressed as:

$$P(R) = \frac{P_2 - \gamma(U_2^2 - U^2)}{8g} \tag{2.7}$$

where, P_2 is the pressure at the outer radius of the impeller R_2, $U_2 = \omega R_2$ is the circumferential speed at the impeller outlet, $U = \omega R$ is the circumferential speed at any arbitrary radius R, and $\gamma = \rho g$ is the specific weight of the liquid. This equation explains how the centrifugal effect of the rotating fluid generates a radially varying pressure distribution on both side of the impeller. However, the pressure distribution on the suction side of the impeller follows a more complex profile, typically taking the form of a curve, as illustrated in the lower part part of Figure 2.3. The shape of this curve is heavily influenced by the geometry of the impeller blades. While in theory, the pressure distribution could be computed using the usual approach as discussed above, in practice this is highly challenging. This difficulty arises because pressure within the impeller passages is not uniform at a given radius; instead, it varies along the arc between consecutive blades due to complex internal flow phenomena. For practical estimation, it has therefore been suggested that the pressure variation within the impeller passage may linearly vary with radius or proportionally to the square of the radius. Both assumptions have been evaluated, and while neither produces a perfect balance, they have been proven adequate to prevent excessive bearing overload [8].

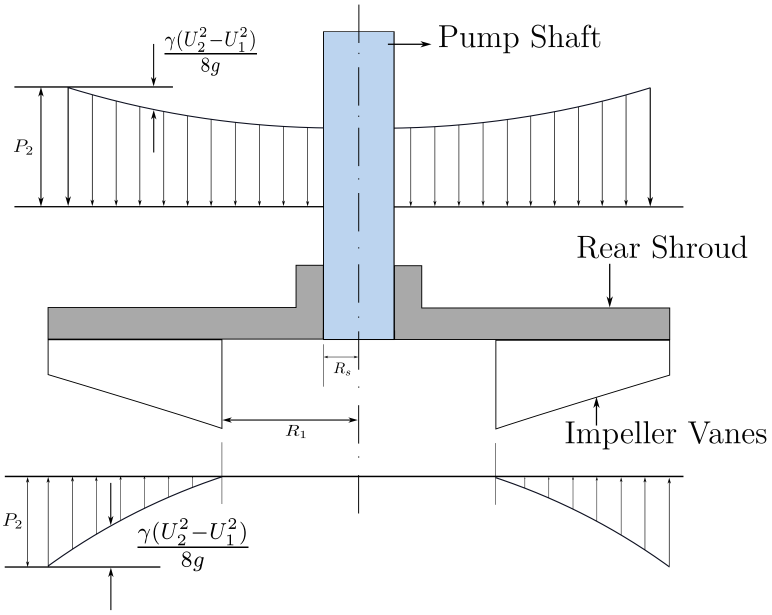


Figure 2.3: Schematic diagram of pressure distribution on a semi-open impeller. Reproduced from [8].

To determine the net axial force, the pressure distribution is integrated over the respective surface areas on the suction and back side of the impeller. For a given pressure profile $P(R)$, the axial force F acting on the impeller can be estimated using the following expression:

$$F = 2\pi \int_{R_s}^{R_1} P(R) \cdot R\,dR \tag{2.8}$$

where R_s is the radius of the pump shaft and R_1 is the radius of the inner edge of the vane (as marked in Figure 2.3). From Figure 2.3, it is clear that the axial force that will act on the suction and the back side of the impeller are not equal, due to the difference in pressure distribution. To further illustrate, a schematic diagram of the axial forces acting on the semi-open impeller is shown in Figure 2.4. The net axial force T can be approximated by combining those individual contributions as:

$$T = F_1 - (F_2 + F_3 + F_4) \tag{2.9}$$

where, F_1 is the pressure force on the rear shroud (back side of the impeller), F_2 is the opposing pressure force on the impeller hub, F_3 is the force caused by the pressure differential across the blade tip clearance (the tiny gap between the blade tips and the stationary casing), and F_4 is the dynamic reaction force resulting from the momentum change of the incoming fluid at the impeller inlet.

This force imbalance leads to a net axial thrust directed along the shaft. In vertically oriented pumps, such unbalanced axial forces can exert a considerable amount of force on the bearing, potentially affecting the mechanical reliability of the pump. To mitigate this axial load, several balancing techniques, such as wear rings, back vanes, and balancing holes, are employed to redistribute pressure and reduce the load on the bearing. Notably, the axial force tends to decrease with increasing flow rate [10]. The following sections cover the underlying theory and function of balancing holes in a semi-open impeller.

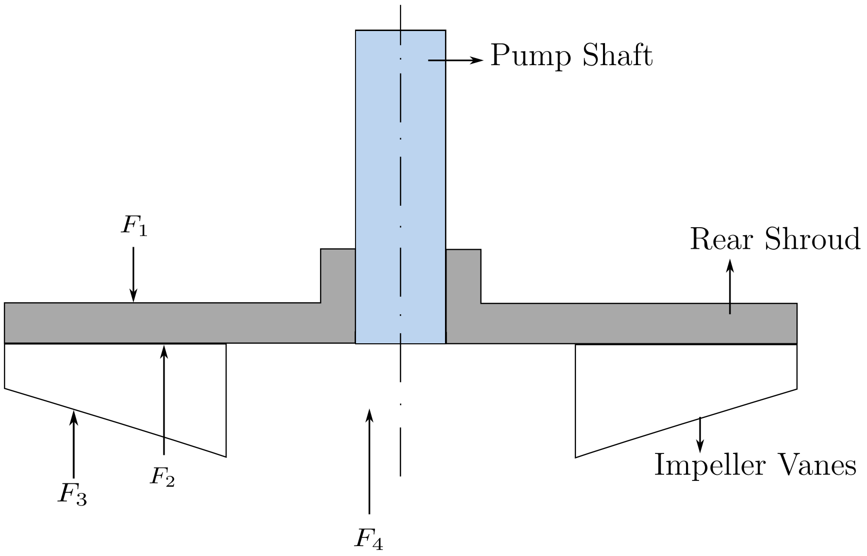


Figure 2.4: Schematic diagram of axial force on a semi-open impeller.

2.2.2 Balancing Holes

Balancing holes (as illustrated in Figure 2.5) are typically small openings drilled through the back side (rear shroud) of the impeller in a centrifugal pump. These holes hydraulically connect the high-pressure region on the back side of the impeller to the low-pressure region near the suction side. By allowing a controlled reverse flow from the back side to the suction side, the balancing holes help equalize the pressure across the impeller, thereby significantly reducing the axial load [11]. This makes them a simple and effective method for reducing axial force in centrifugal pumps [2].

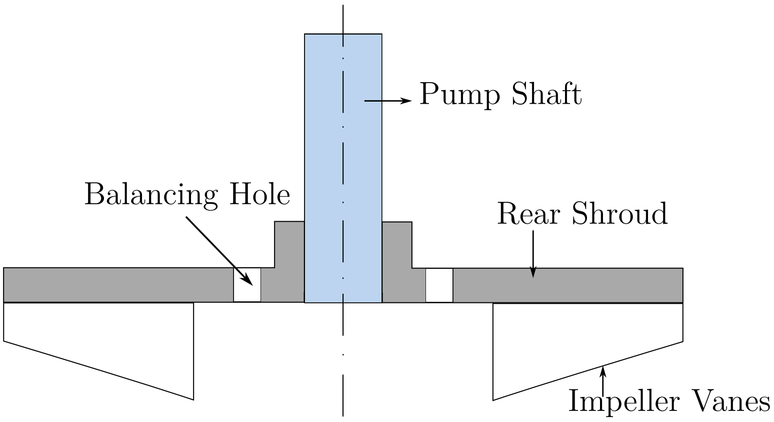


Figure 2.5: Schematic showing the balancing hole in a semi–open centrifugal impeller.

2.2.2.1 Effect of Balancing Holes on Pressure Distribution and Axial Force

As discussed in the previous section, the introduction of balancing holes in the impeller results in a reduced pressure distribution compared to the case without balancing holes (refer to Figures 2.3 and 2.4). This modification produces a less steep radial pressure gradient across the back shroud. This reduction in backside pressure lowers the axial force acting in that region, thereby reducing the net axial force acting on the impeller. This demonstrates the effectiveness of balancing holes in mitigating axial thrust and improving the mechanical balance of semi-open impellers.

2.2.2.2 Design Configurations

While adding balancing holes is beneficial in mitigating the axial force, it might lead to performance loss (head and efficiency) due to the secondary leakage flow within the pump. The severity of this trade-off primarily depends on the geometric configuration of the balancing hole, and to a lesser extent, on the operating condition of the centrifugal pump. Therefore, the design of balancing holes must be optimally configured to mitigate as much axial force as possible while minimizing the adverse effects on the pump performance.

The parameters that mainly define the design configurations of balancing hole include hole diameter, number of holes, chamfer (the beveled edge at hole entrance), radial and circumferential position. Each of these design parameters has a significant effect on both axial force and pump performance.

2.2.3 Clearance Area

In centrifugal pumps with semi-open impellers, the clearance area refers to the narrow radial gap between the rear shroud of the impeller and the stationary casing (bottom of the motor). This region, shown in Figure 2.6, forms a leakage path through which high-pressure fluid from the discharge side can move toward the backside of the impeller. When fluid leaks through this gap, it enters the back side of the impeller, affecting the pressure distribution in that region. The pressure on the back side of the impeller is determined by the balance of leakage flow and the dynamics of the rotating impeller and stationary casing, and it is typically between the suction and discharge pressures.

The clearance area ($A_{\text{Clearance}}$), which governs the leakage flow rate and contributes to axial thrust, is given by the expression in Equation 2.10

$$A_{\text{Clearance}} = D_{\text{Mean}} \cdot s \tag{2.10}$$

where D_{Mean} is the mean diameter of the impeller and s is the radial clearance between the rear shroud of the impeller and the stationary casing (bottom of the motor).

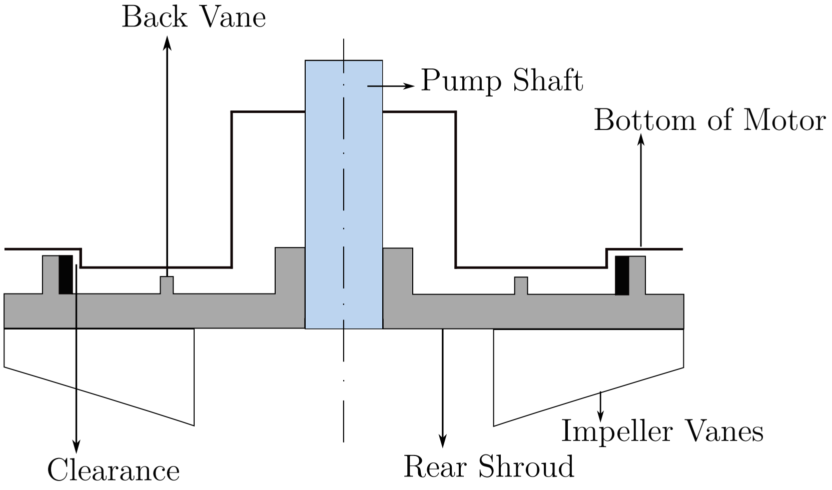


Figure 2.6: Clearance area on a semi-open impeller.

This area plays a critical role in estimating internal leakage and is a key consideration in the design of other pump components, such as balancing holes, to ensure proper pressure distribution across the impeller to minimize axial force.

2.3 Computational Fluid Dynamics

CFD is a branch of fluid mechanics that uses powerful numerical techniques to analyze and predict fluid flow behavior in complex engineering problems. The foundation of CFD lies in the fundamental governing equations of a fluid motion, which for isothermal incompressible flow with constant viscosity are the following:

- Continuity equation (Conservation of Mass):

$$\frac{\partial v_i}{\partial x_i} = 0 \tag{2.11}$$

- Navier–Stokes equation (Conservation of Momentum):

$$\underbrace{\rho\frac{\partial v_i}{\partial t}}_{\text{Transient term}} + \underbrace{\rho\frac{\partial v_i v_j}{\partial x_j}}_{\text{Convective term}} = \underbrace{-\frac{\partial p}{\partial x_i}}_{\text{Pressure gradient}} + \underbrace{\mu\frac{\partial}{\partial x_j}(\frac{\partial v_i}{\partial x_j})}_{\text{Viscous (Diffusion) term}} \tag{2.12}$$

where, v_i and v_j denote the components of the velocity vector, p is the pressure, ρ is the fluid density, and μ is the dynamic viscosity.

These equations are discretized and solved iteratively to predict and analyze the distribution of velocity and pressure within the fluid domain. In case of laminar flow, these equations are straightforward to solve. However, for turbulent flows—which are common in centrifugal pumps—directly solving these equations become much more complex due to their chaotic, nonlinear and multiscale nature.

To make the problem solvable, the Reynolds-Averaged Navier–Stokes (RANS) equations are used. In this method, the flow variables such as velocity and pressure are split into mean and fluctuating component. Here, the mean component represents time-averaged or ensemble-average value of the flow variable and the fluctuating component represents the deviation of the instantaneous value from the mean.

$$v_i = \bar{v}_i + v_i' \tag{2.13}$$

$$p = \bar{p} + p' \tag{2.14}$$

$$\bar{v} = \frac{1}{2T}\int_{-T}^{T} v\,dt \tag{2.15}$$

where v_i is the instantaneous velocity component, $\bar{v}_i$ is the time-averaged (mean) velocity, v_i' is the fluctuating velocity component, p is the instantaneous pressure, $\bar{p}$ is the mean pressure, p' is the pressure fluctuation, and $\bar{v}$ is the time-averaged scalar velocity.

Substituting Equation 2.13 & 2.14 into Equation 2.12 and performing time averaging (along with some mathematical manipulation) gives the Reynolds Averaged Navier-Stokes equation:

$$\frac{\partial \bar{v}_i}{\partial x_i} = 0 \tag{2.16}$$

$$\rho\frac{\partial \overline{v_i}}{\partial t} + \rho\frac{\partial \bar{v}_i \bar{v}_j}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu\frac{\partial \bar{v}_i}{\partial x_j} - \rho\overline{v_i' v_j'}\right) \tag{2.17}$$

In the resulting RANS momentum equation, a new additional term ($\tau_{ij} = -\rho\overline{v_i' v_j'}$) has appeared in the diffusion term. This term is known as the Reynolds Stress Tensor, which accounts for the additional momentum transfer caused by turbulent fluctuations. This term represents the effect of turbulence on the mean flow and introduces 6 additional unknown terms, on top of the four original unknowns (3 velocity components and pressure) in the system of equations. To solve this problem, the additional term that appeared in Equation 2.17 must be modeled. The process of modeling those terms is known as turbulence modeling, which will be explained in the upcoming section.

2.3.1 Turbulence Modeling

Turbulence modeling is a mathematical approach used in CFD to predict turbulent flow behavior within a fluid system. Due to its inherently chaotic, unstable, and fluctuating nature, accurately modeling this is a complex task. Instead of resolving all the turbulent scales directly (Direct Numerical Simulation method), which is computationally expensive, a turbulence model has been employed to approximately represent their influence on the mean flow. Hence, different turbulence models have been developed to model the unknown additional term (Reynolds stress) and close the system of equations.

In this study, the cubic $k - \varepsilon$ turbulence model is used to analyze the internal flow effects within the centrifugal pump, due to its improved accuracy in capturing the anisotropic effects which is common in turbomachinery. The upcoming section provides a brief overview of standard $k - \varepsilon$ turbulence model then discussing about cubic $k - \varepsilon$ model.

2.3.2 Standard $k - \varepsilon$ Turbulence Model

The standard $k - \epsilon$ model is the most commonly used turbulence model to provide a mathematical representation of the turbulent flow field. It is a two-equation model, meaning it uses two additional transport equations, namely, turbulent kinetic energy (K) and turbulent dissipation rate (ϵ) along with the RANS equation, to help solve and predict the turbulent flow properties. The Boussinesq hypothesis models/relates the additional term ($-\overline{v_i' v_j'}$ - Reynolds Stress Tensor) in the RANS equation to the mean velocity gradient using eddy (ν_t) viscosity, through a linear-relationship to close the RANS equation. The mathematical expression of the hypothesis is expressed as:

$$-\overline{u_i' u_j'} = 2\nu_t S_{ij} - \frac{2}{3} k \delta_{ij} \tag{2.18}$$

$$S_{ij} = \frac{1}{2}\left(\frac{\partial \overline{v_i}}{\partial x_j} + \frac{\partial \overline{v_j}}{\partial x_i}\right) \tag{2.19}$$

$$a_{ij} = -\overline{u_i' u_j'} + \frac{2}{3} k \delta_{ij} = 2\nu_t S_{ij} \tag{2.20}$$

$$\nu_t = C_\mu \frac{k^2}{\varepsilon} \tag{2.21}$$

where C_μ is a constant with a value of 0.09, a_{ij} is the anisotropic part of the Reynolds stress tensor, δ_{ij} is the Kronecker delta and S_{ij} is the mean strain rate tensor. By applying the Boussinesq hypothesis, the 10 unknown terms in the RANS equations are reduced to a solvable system with 2 additional unknowns: the turbulent kinetic energy and the turbulent dissipation rate. These can be determined by solving their respective transport equations. However, this linear model as said earlier assumes an isotropic (which means the turbulence has the same intensity and statistical properties in all 3 directions) relationship between the turbulent stresses and the mean strain rate, which limits its ability to accurately predict complex flows with strong anisotropy, streamline curvature, or separation.

The transport equation for the turbulent kinetic energy:

$$\underbrace{\frac{\partial(\rho k)}{\partial t}}_{\text{Unsteady term}} + \underbrace{\frac{\partial(\rho k \overline{v}_j)}{\partial x_j}}_{\text{Convection}} = \underbrace{\frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right]}_{\text{Diffusion}} + \underbrace{P_k}_{\text{Production}} - \underbrace{\rho\varepsilon}_{\text{Dissipation}} \tag{2.22}$$

The transport equation for the turbulent dissipation rate:

$$\underbrace{\frac{\partial(\rho \varepsilon)}{\partial t}}_{\text{Unsteady term}} + \underbrace{\frac{\partial(\rho \varepsilon \overline{v}_j)}{\partial x_j}}_{\text{Convection}} = \underbrace{\frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_j}\right]}_{\text{Diffusion}} + \underbrace{C_{1\varepsilon}\frac{\varepsilon}{k}P_k}_{\text{Production}} - \underbrace{C_{2\varepsilon}\rho\frac{\varepsilon^2}{k}}_{\text{Destruction}} \tag{2.23}$$

Where $C_{1\varepsilon}$, $C_{2\varepsilon}$, and C_3 are model coefficients whose values depend on the specific form of the k-ε model chosen.

2.3.3 Cubic $k-\varepsilon$ Turbulence Model

In this model, the Reynolds stress is modeled using a more general non-linear relation that includes both the mean strain rate tensor and the mean rotation rate tensor, also known as the vorticity tensor [12]. This general form is expressed as:

$$a_{ij} = \sum_n G_n T_{ij}^{(n)} \tag{2.24}$$

$$\begin{aligned} a_{ij} = & -2C_\mu \tau \hat{S}_{ij} \\ & + c_1\tau^2\left(\hat{S}_{ik}\hat{S}_{kj} - \frac{1}{3}\hat{S}_{\ell k}\hat{S}_{k\ell}\delta_{ij}\right) + c_2\tau^2\left(\hat{\Omega}_{ik}\hat{S}_{kj} - \hat{S}_{ik}\hat{\Omega}_{kj}\right) \\ & + c_3\tau^2\left(\hat{\Omega}_{ik}\hat{\Omega}_{jk} - \frac{1}{3}\hat{\Omega}_{\ell k}\hat{\Omega}_{k\ell}\delta_{ij}\right) + c_4\tau^3\left(\hat{S}_{ik}\hat{S}_{k\ell}\hat{\Omega}_{\ell j} - \hat{\Omega}_{i\ell}\hat{S}_{\ell k}\hat{S}_{kj}\right) \\ & + c_5\tau^3\left(\hat{\Omega}_{i\ell}\hat{\Omega}_{\ell m}\hat{S}_{mj} + \hat{S}_{i\ell}\hat{\Omega}_{\ell m}\hat{\Omega}_{mj} - \frac{2}{3}\hat{\Omega}_{mn}\hat{\Omega}_{n\ell}\hat{S}_{\ell m}\delta_{ij}\right) \\ & + c_6\tau^3\hat{S}_{k\ell}\hat{S}_{k\ell}\hat{S}_{ij} + c_7\tau^3\hat{\Omega}_{k\ell}\hat{\Omega}_{k\ell}\hat{S}_{ij} \end{aligned} \tag{2.25}$$

$$\hat{\Omega}_{ij} = \frac{1}{2}\left(\frac{\partial \overline{v}_i}{\partial x_j} - \frac{\partial \overline{v}_j}{\partial x_i}\right) \tag{2.26}$$

where $T_{ij}^{(n)}$ are the basis tensors and G_n are scalar coefficients. This nonlinear Boussinesq approximation allows the model to more accurately capture anisotropic effects, secondary flows, and curvature-related phenomena, which are especially important for simulating the complex turbulent flows in centrifugal pumps.

2.3.4 Multiple Reference Frame model - Frozen Rotor Approach

The Multiple Reference Frame model is the most widely used CFD approach for simulating rotating machinery like turbines, fans and pumps. It is a steady-state simulation method where the entire computational domain is divided into multiple zones – a rotating zone (typically an impeller or rotor) and a stationary zone (volute or casing). The rotating and stationary zones are mathematically coupled through a non-sliding interface (fixed interface) where the flow variables are interpolated between the frames of reference.

Frozen Rotor Approach is a specific implementation of the MRF model, where the rotating region is held fixed at a certain angular position with respect to the stationary region. The governing equations are solved in a rotating reference frame for the rotating region and a stationary reference frame for the stationary region. Even though the rotating region does not physically rotate during the simulation, the rotational effects are fully accounted for in the governing equations by additional source terms, especially centrifugal and Coriolis forces in the momentum equations.

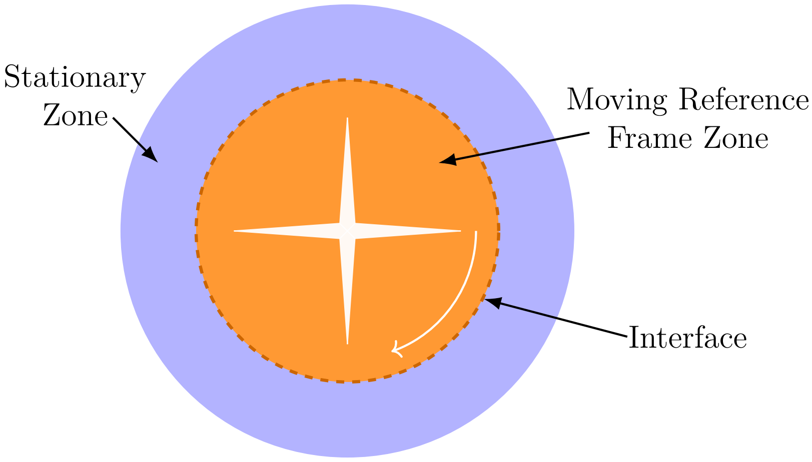


Figure 2.7: Geometry with one rotating impeller.

Taking an example of a mixing tank with a single impeller as illustrated in Figure 2.7, the impeller along with the surrounding flow (orange region) is defined to be the moving frame, and the flow outside the impeller is defined as the stationary frame. The interface between the two frames is assumed to have a steady-state condition, where it serves as a surface of revolution for the two frames [13]. The interface (non-sliding) between the two zones allows momentum and pressure to be transferred without the need for any transient time-stepping scheme.

The Frozen Rotor Approach provides a computationally effective way to evaluate the performance of rotating machines. It is suitable for steady-state performance predictions, initial design and performance assessment and a precursor to more detailed transient CFD analysis. However, this approach is not applicable for simulations involving transient effects like unsteady interactions between the rotor and the stator, blade passing phenomenon, and flow-induced fluctuations that require time-accurate modeling.

2.4 Multi-Objective Optimization

Multi-objective optimization refers to an optimization problem that involves more than a single objective function, which needs to be optimized simultaneously, often with conflicting goals. In such problems, the aim is to find solutions that represent the best possible trade-offs between those objectives, since one objective may lead to the degradation of the other objective. These solutions are simply called Pareto-optimal solutions or non-dominated solutions.

2.4.1 Pareto Optimality

In multi-objective optimization problems, it is possible to obtain multiple feasible solutions rather than a single optimal solution. Identifying a single feasible solution in such optimization processes is quite challenging due to the inherently conflicting nature of each objective. Instead, a more meaningful strategy is to determine a set of feasible solutions that represent the best possible trade-offs, a concept known as Pareto optimality.

A solution is said to be Pareto optimal if no other solution in the objective space exists that improves one objective without worsening the other objective. These types of solutions available in the design / objective space are also called non-dominated solutions. These non-dominated solutions collectively form the Pareto front, which defines the boundary of optimal trade-offs in the objective space. This is defined more precisely as:

> A point x^* in the feasible design space S is **Pareto optimal** if and only if there does not exist another point x in the set S such that $f(x) \leq f(x^*)$ with at least one $f_i(x) < f_i(x^*)$ [14].

The collection of such non-dominated solutions can be visualized as forming a Pareto front in the objective space. In problems with two objectives, as shown in Figure 2.8, the Pareto front typically appears as a smooth curve connecting the non-dominated solutions. Thus, the Pareto optimal set is always on the boundary of the feasible objective space. Each point on the Pareto front represents a valid trade-off solution, and no point is strictly better than another across all objectives. Therefore, the final selection among Pareto-optimal solutions depends on the decision maker's preferences and contextual priorities.

While the Pareto-optimal set is always on the boundary of the feasible objective space in a constrained scenario, it is significant to note that this boundary is not always defined by the explicit constraints. The Pareto front also emerges in unconstrained problems, where the optimal trade-offs purely arise from the inherent conflict between objectives. In such problems, the Pareto optimal set is determined by comparing the non-dominated solutions across the objective space or analytically by identifying the points where the gradients of the objective functions oppose each other, indicating that no common descent direction exists for simultaneous improvement.

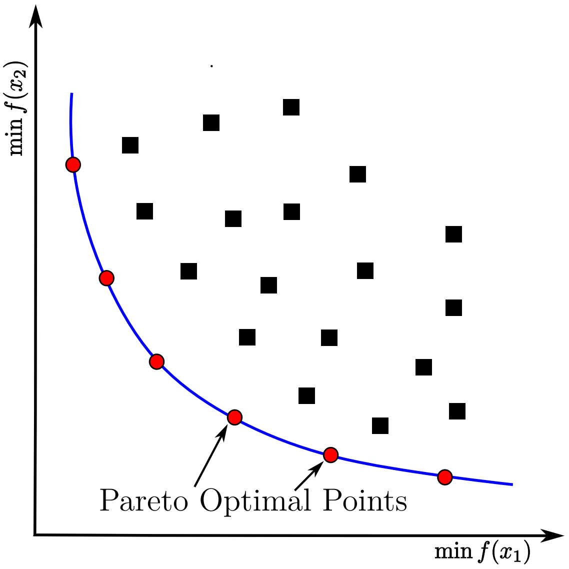


Figure 2.8: Pareto front with non-dominated optimal points.

2.4.1.1 Preferences and Utility Function

Mathematically, there are infinitely many Pareto-optimal solutions available within an objective space. This makes the multi-objective optimization problem more challenging to solve and find an optimal solution. Apparently, it becomes necessary to make decisions regarding which solution is preferred. This can be done by the process of providing preference to the objective function. Preferences reflect the decision maker's priorities or opinions concerning different solutions within the objective space. Ideally, a multi-objective method should incorporate these preferences so that the outcome better aligns with the decision maker's needs. However, representing preferences mathematically is a tedious and non-trivial task, especially when aiming for a multi-objective target function that covers all the nuances of the decision maker's need.

However, various methods have been developed to incorporate the preferences mathematically into an objective function. Among those, the most traditional and mathematically straightforward approach is to assign weights to each objective based on their relative importance, allowing them to reflect the decision maker's preference [14]. Although straightforward, it is not an optimal strategy, as it may cause the optimization process to overly focus on a single objective, completely deviating from the core principle of a multi-objective optimization. To tackle this, Steuer [15] provided a mathematical formulation using a hyperplane for interpreting and assigning weights in a multi-objective target function. Based on this, the mathematical formulation for calculating weights in a multi-objective optimization process can be expressed as:

$$\Delta_i = \frac{1}{N} \sum_{j=1}^{N} |f_i(x_j) - f_{i,\mathrm{ref}}| \tag{2.27}$$

where $f_i(x_j)$ is the value of the i^{th} objective for the j^{th} sample, $f_{i,\mathrm{ref}}$ is the reference value for the i^{th} objective, and N is the total number of samples. Equation 2.27 represents the mean absolute deviation from its reference value. Then, the inverse of the mean absolute deviation is computed to emphasize objectives that are closer to their reference values:

$$I_i = \frac{1}{\Delta_i} \tag{2.28}$$

The normalized weights w_i are obtained by:

$$w_i = \frac{I_i}{\sum_{k=1}^{m} I_k} \tag{2.29}$$

where w_i is the normalized weight assigned to the i^{th} objective, I_i represents the inverse of the mean absolute deviation for the i^{th} objective, I_k denotes the inverse deviation of the k^{th} objective (with k ranging over all objectives), and m is the total number of objectives considered.

While the weightage method represents the preferences among different objectives, the utility function plays the role of combining these multi-objective functions into a single-objective framework (i.e., by scalarizing multiple objectives into a single objective function), thereby simplifying the optimization process. Scalarization is the mathematical process of combining or transforming a set of different objective optimization problems into a single objective function [14]. By scalarizing, the optimization process becomes easier, and it provides a structured and accurate way to analyze the trade-offs between different sets of objectives in multi-objective optimization. The general form of the scalarized objective function is expressed as shown in Equation 2.30.

$$F(x) = \sum_{i=1}^{m} f_i(x) \tag{2.30}$$

where $F(x)$ is the scalarized objective function, $f_i(x)$ is the i^{th} objective function, and m is the total number of objectives. This transformation allows the selection and evaluation of the optimal solution to be carried out more mathematically without neglecting the decision maker's demands while properly taking into account the trade-offs between different objectives within a unified approach.

2.4.2 Normalization

Normalization is the process of scaling or transforming a set of absolute data sets relative to a reference value, typically to make the data sets into dimensionless or comparable form, allowing consistent comparison across magnitudes, units, or ranges. Normalization is widely used to omit the effects of scale, making analyses based on relative differences rather than absolute values.

In this study, different normalization technique were used to quantify the deviation between simulated results and a reference data. Below is a brief explanation of the techniques used in this study.

2.4.2.1 Relative Squared Error (RSE)

The Relative Squared Error (RSE) is a widely used metric system to analyze or quantify how close the observed (in this case, simulated result) value deviates or matches with the reference value. It is simply the ratio of the total deviation of all observed values from the reference values to the total deviation of the same observed values from their mean values. By this, the RSE brings the deviation to a relative scale. The general form of RSE, often used for assessing the performance across an entire dataset, is expressed as:

$$\text{RSE} = \frac{\sum_{i=1}^{n}(y_{\text{ref},i} - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_{\text{ref},i} - \overline{y}_{ref})^2} \tag{2.31}$$

where $y_{\text{ref},i}$ is the reference value for the i^{th} data point, $\hat{y}_i$ is the observed or simulated value for the i^{th} data point, $\overline{y}_{ref}$ is the mean of all reference values, and n is the total number of data points.

In this study, a modified version of RSE, known as the Pointwise Relative Squared Error, has been used to normalize and evaluate the deviation of individual data points with respect to the reference values.

2.4.2.2 Pointwise Relative Squared Error

The modified version (for this study) known as the Pointwise RSE focusses on analyzing the deviation of each individual dataset from the reference value. This offers a more detailed understanding of the performance at each specific observation, rather than presenting a single overall measure of error of the observed data. The Pointwise RSE for an individual dataset is expressed as:

$$\text{Pointwise RSE} = \frac{(y_{ref,i} - y_i)^2}{(y_{\text{ref, i}})^2} \tag{2.32}$$

where y_i is the observed or simulated value for the i^{th} configuration, and $y_{\text{ref, i}}$ is the corresponding reference value for the i^{th} configuration. The necessity and the practical implementation of this technique are discussed in Section 3.4.1.

2.4.3 Bayesian Optimization

Bayesian optimization is a technique used for finding optimal solutions for objective functions that are expensive to evaluate since they provide no information about their gradients [16]. This function can be noisy or non-convex in nature further complicating the traditional optimization like the first order and second order based methods. Instead of directly evaluating and optimizing the true objective function, which can be computationally expensive and time-consuming, Bayesian optimization relies on constructing a surrogate model to approximate the function and guide the search for the optimal solution of the objective function.

In Bayesian Optimization, the Gaussian process as the surrogate model provides a probabilistic prediction of the true black-box objective function value and also an estimate of the uncertainty at any point in the design space. This dual information associated with an acquisition function enables mathematical decision-making on where to explore or sample next based on the current information by balancing exploration (where the model is uncertain) and exploitation (where the model predicts high performance) [16].

2.4.3.1 A Bayesian Approach to Optimization

The Bayesian optimization is fundamentally concerned with finding the minimizer of the objective function without needing the access to the Gradient or Hessian of the function [16]. Rather, it queries on the function f at a given point x and returning the value of $f(x)$. In the Bayesian approach, instead of attempting to find the optimal point x for a function f, it aims to learn a best model of the function based on previously observed points x. By conditioning on these observed points, the Bayesian optimization extracts an expected function f and a confidence interval, which are used to guide the selection of future sampling points.

This makes Bayesian optimization particularly well suited for complex design optimization and computationally expensive simulation tasks. In this study, the method is employed to optimize different balancing hole configurations in a centrifugal pump, where each design evaluation involves computationally intensive CFD simulations. The following sections describe the two core components of Bayesian optimization: the Gaussian Process and the acquisition function.

2.4.3.2 Gaussian Process

A Gaussian process is a probabilistic non-parametric model used in Bayesian optimization to construct a surrogate model over the objective function. It defines a probability distribution over the functions and can be updated with new observations to refine its predictions [16]. The Gaussian process provides not only an estimate of the objective function's value at any input points but also an uncertainty measure at those same points in the input space.

$$f(x) \sim \mathcal{GP}(\mu(x), K(x, y)) \tag{2.33}$$

where $\mu(x)$ is the *mean function*, assigning to each input point x the expected value $\mathbb{E}[f(x)]$, and $K(x, y)$ is the *covariance function* (or kernel), assigning to each pair of input points x and y the covariance between $f(x)$ and $f(y)$.

The covariance function is also called the kernel of the Gaussian process. It simply measures the similarity or correlation between the values of $f(x)$ and $f(x')$. Among various kernel functions available, the most widely used kernel in Gaussian process is the Radial Basis Function (RBF) [16] kernel also known as the squared exponential kernel or Gaussian kernel, defined as:

$$K(x, y) := k \cdot \exp\left\{-\frac{\|x - y\|_2^2}{2\sigma^2}\right\}, \tag{2.34}$$

where k is the variance parameter that sets the overall scale of variation, and σ is the length-scale parameter that controls the smoothness of the function.

This kernel function ensures that the input closer in the design space have highly correlated function values, which distant inputs have low level of correlation. By using this kernel function, Gaussian process surrogate model captures the structure of the underlying true objective function effectively, allowing Bayesian Optimization to propose new promising candidates [17].

2.4.3.3 Acquisition Function

The acquisition function is a guideline that suggests where to query the next sample points based on the posterior predictive distribution (refer to Section 2.4.5) of the function provided by the Gaussian process. One of the most commonly used acquisition functions is the expected improvement (EI), which quantifies the expected amount by which a new sample will improve the best observed objective value [16].

Let f^* be the best value observed so far, and let $\tilde{f}(x) \sim \mathcal{N}(\mu(x), \sigma^2(x))$ be the posterior predictive distribution of the Gaussian Process at input x. The expected improvement (EI) is defined as:

$$\text{EI}(x) = \mathbb{E}[\max(0, f^* - \tilde{f}(x))] = (f^* - \mu(x)) \cdot \Phi(Z) + \sigma(x) \cdot \phi(Z) \tag{2.35}$$

$$Z = \frac{f^* - \mu(x)}{\sigma(x)} \tag{2.36}$$

where Z is the standardized improvement, $\Phi(Z)$ is the cumulative distribution function (CDF) of the standard normal distribution, and $\phi(Z)$ is the probability density function (PDF) of the standard normal distribution.

The next sampling point is chosen by maximizing the acquisition function defined in Equation 2.35, where the goal is to identify the location in the design space that offers the greatest expected improvement over the current best observation. This formulation balances exploitation and exploration, making expected improvement (EI) a powerful and practical choice for design optimization tasks.

2.4.3.4 Bayesian Optimization Work Flow

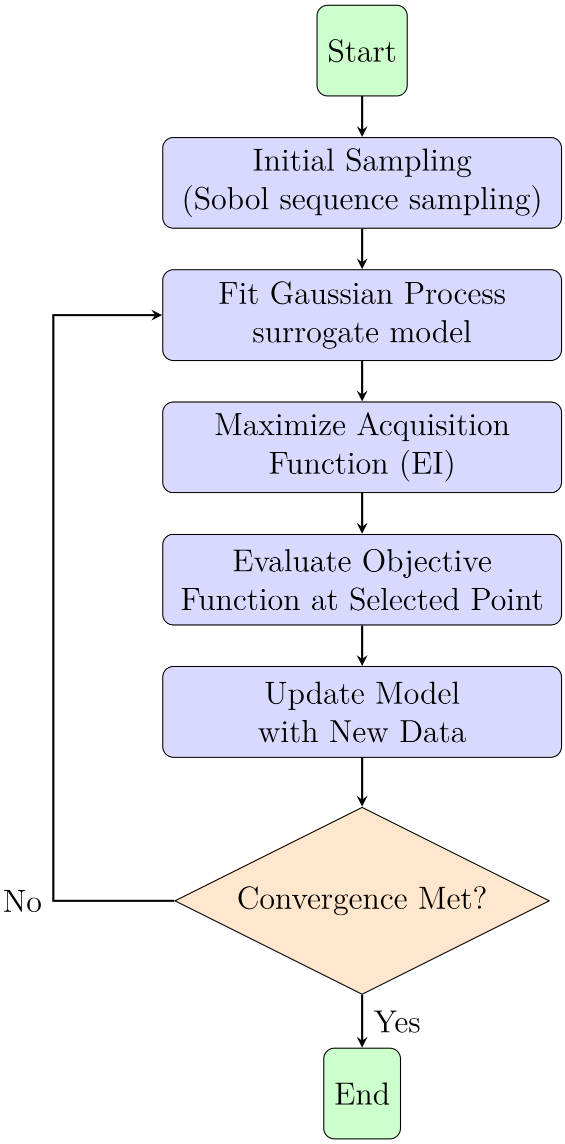


Figure 2.9: Bayesian optimization workflow.

Figure 2.9 illustrates the overall process involved in the Bayesian optimization workflow. It can be seen from the flow chart that the workflow starts with an initial sampling step using Sobol sequence. This initial step ensures that the optimization starts with a set of input points that are uniformly (evenly) distributed in the design space, providing a solid foundation for Bayesian Optimization to build the surrogate model using Gaussian process. The theoretical background for the Sobol sequence sampling is presented in the following section.

2.4.4 Sobol Sequence Sampling

In high-dimensional design and optimization problems, especially those involving surrogate modeling or Bayesian optimization, the initial sampling of design points plays a critical role in determining the efficiency and accuracy of the optimization process. Traditional random sampling often leads to clustering and poor coverage of the design space. To overcome this, low-discrepancy sequences, such as the Sobol sequence, are employed to provide better uniformity and space-filling properties.

A Sobol sequence is a type of quasi-random sequence designed to fill the unit hypercube $[0,1]^d$ as uniformly as possible. Unlike purely random (Monte Carlo) samples, which suffer from uneven spacing and gaps in the domain, Sobol sequences exhibit low discrepancy, meaning the generated points are distributed uniformly within the design space [18]. This makes them highly effective in areas such as numerical integration, global sensitivity analysis, and initial design of experiments (DoE) for machine learning and optimization. Sobol' sequences are a type of digital (t,s)-sequence in base 2, according to the definition or framework given by Niederreiter [19].

2.4.4.1 Discrepancy and Uniformity Properties

Sobol sequences are constructed to minimize discrepancy—a measure of the deviation of a sequence from a perfectly uniform distribution. The discrepancy D_N of the first N Sobol points satisfies:

$$D_N = \mathcal{O}\left(\frac{(\log N)^d}{N}\right)$$

This is asymptotically superior to the convergence rate of Monte Carlo sampling:

$$D_N = \mathcal{O}(N^{-1/2})$$

2.4.5 Posterior Predictive Distribution

The posterior predictive distribution, a fundamental concept in Bayesian statistics, is the distribution of possible unobserved outcomes based on the Bayesian model [20]. It is obtained by integrating the entire posterior distribution of the model parameters by averaging over the likelihood of all possible parameter values weighted by their posterior probability. The posterior predictive distribution for a new input x_* is given by:

$$p(y_* \mid x_*, \mathcal{D}) = \int p(y_* \mid x_*, \theta)\, p(\theta \mid \mathcal{D})\, d\theta \tag{2.37}$$

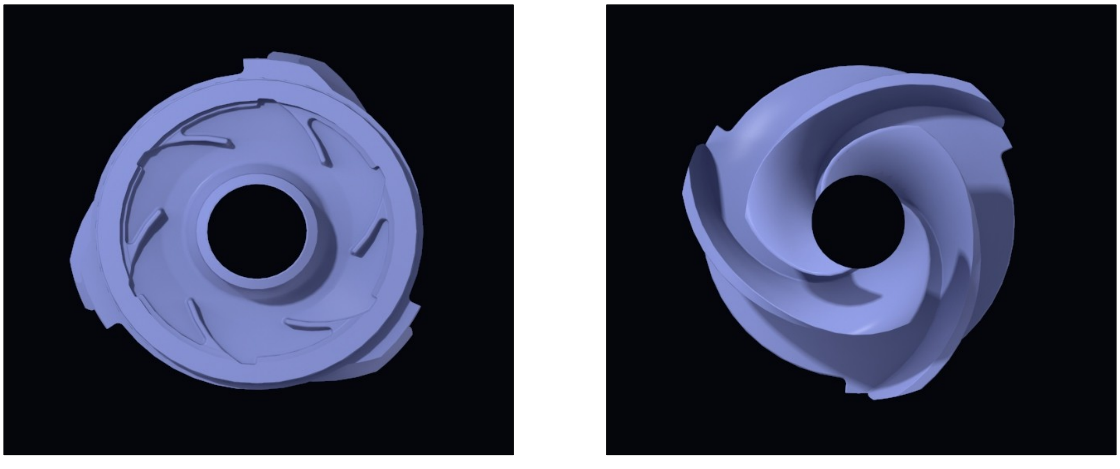
where θ denotes the model parameters, $\mathcal{D}$ is the observed dataset, and y_* is the predicted output at the new input x_*.

3

Methodology

3.1 Design configurations for balancing hole optimization

To develop a robust multi-objective target function for the design optimization of balancing hole, a set of well-defined suitable design parameters were considered in this study. This set of design parameters includes the diameter of the balancing hole, the number of holes, their radial position relative to the center of the impeller, and their circumferential position relative to the impeller's front and rear vanes. The selection of these design parameters was highly influenced by their impact on axial force reduction, leakage flow management and overall hydraulic performance of the pump.



(a) Pressure / Back side. **(b)** Suction / Front side.

Figure 3.1: Semi-open impeller without balancing hole.

The 3D CAD model of the impeller along with the balancing hole layout, was provided by Xylem. The geometrical representation of the working impeller without balancing hole is shown in Figure 3.1. The design configurations explored in this study were defined using the Relation feature (an inbuilt option) within PTC Creo, allowing for a systematic control over the diameter, radial and circumferential positioning of the balancing holes. These parametric relations were predefined by Xylem and reused in this study to design multiple balancing hole configurations.

The following subsections describe how each of these parameters was defined.

3.1.1 Number of balancing holes

The number of balancing holes used in this optimization study was fixed at 3. This is because the impeller has three blade passages, and each hole is placed to match one passage. So, there is one hole for each blade passage, as shown in Figure 3.3a.

3.1.2 Balancing hole diameter

The diameter of the balancing hole range was determined using the clearance area ($A_{\text{Clearance}}$), which was provided by Xylem. According to the guideline proposed in Gülich's centrifugal pump design framework [3], the total area of the balancing hole should be 4–5 times the clearance area. Figure 3.2, illustrates red arrow highlights the clearance between the impeller and the station casing (bottom of motor), while the blue arrow indicates the washout channel designed to remove debris entering through the clearance. Together, these two regions constitute the total clearance area used for balancing hole sizing.

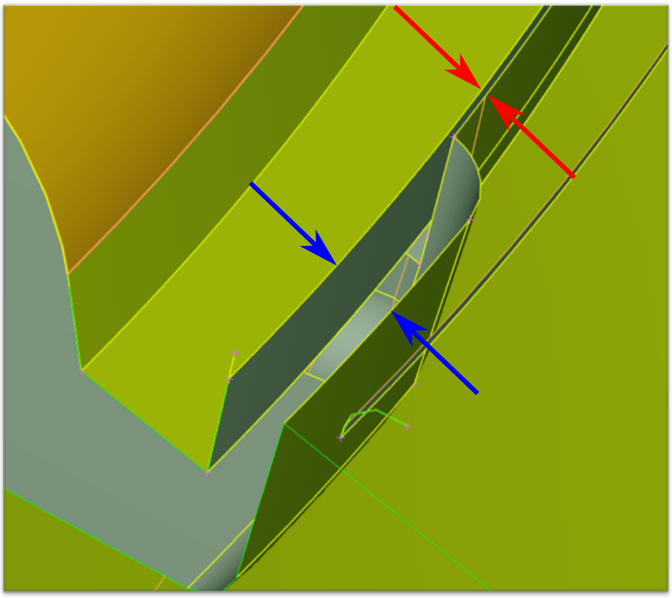


Figure 3.2: Clearance area of the semi-open impeller.

Given three balancing holes fixed, the corresponding minimum and maximum range of total area were calculated as:

$$A_{\text{Total}} \in [1 \times A_{\text{Clearance}},\ 5 \times A_{\text{Clearance}}]$$

$$A_{\text{per Hole}} = \left[\frac{1 \times A_{\text{Clearance}}}{3},\ \frac{5 \times A_{\text{Clearance}}}{3}\right] = [A_{\text{per Clearance,min}},\ A_{\text{per Clearance,max}}]$$

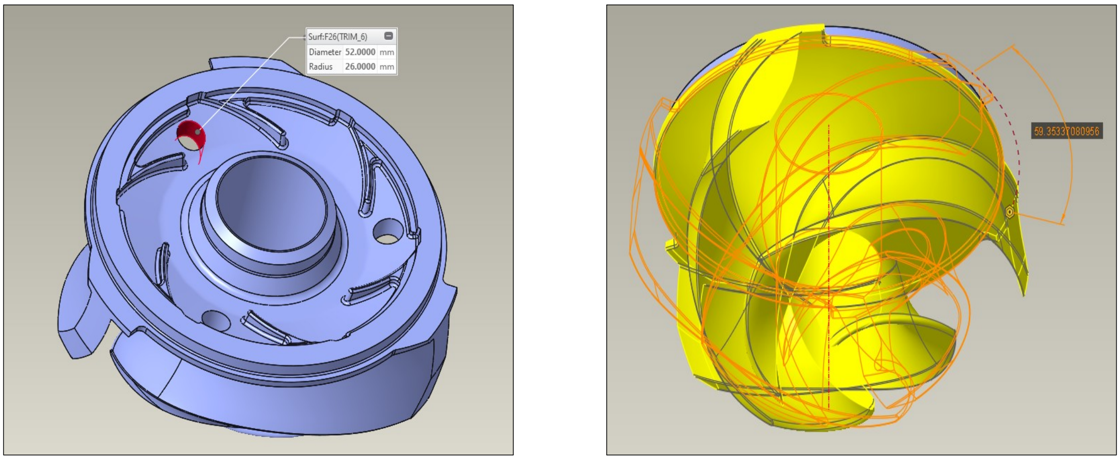
Given that the holes are circular, the diameter range is computed as:

$$D = \sqrt{\frac{4 \times A_{\text{per Hole}}}{\pi}} \quad \Rightarrow \left[\sqrt{\frac{4 \times A_{\text{per Hole,min}}}{\pi}}, \sqrt{\frac{4 \times A_{\text{per Hole,max}}}{\pi}} \right] = [D_{\text{min}}, D_{\text{max}}]$$

Thus, the balancing hole diameter considered in this study ranges from D_{min} and D_{max}.

3.1.3 Circumferential position of balancing hole

Defining the circumferential positioning of the holes required a two-part approach due to the geometry of the impeller. The impeller was conceptually divided into the back shroud, and the front blade section. Drilling through the impeller based solely on the back-side layout would risk interference with the front blades. To address this, the front section of the impeller was rotated about its central axis to align the front and rear hole positions as illustrated in Figure 3.3b.



(a) Number of holes.

(b) Rotation in front side.

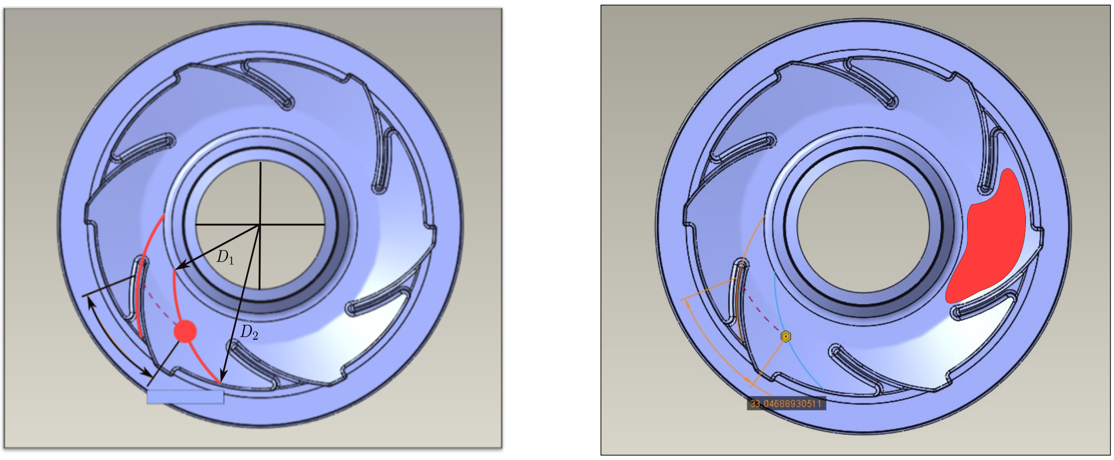
Figure 3.3: Semi-open impeller with balancing holes.

3.1.3.1 Back Side of the Impeller

To define hole positions on the rear shroud, a geometric method based on back vane spacing was adopted. A centerline was drawn along one of the back vanes and rotated around the impeller's central axis. Since the back vanes are uniformly spaced, the angular separation between two vanes is 60°. Therefore, the circumferential positioning of holes will be explored between the two adjacent vanes as shown in Figure 3.4a

3.1.4 Radial Position of Balancing Hole

The radial position of the balancing hole was defined using two diameters, D_1 and D_2, as shown in Figure 3.4a. These diameters represent the radial limit within which the hole positioning was explored. Within this bounded radial region, a curved guide path—shown as the red arc in Figure 3.4a—was drawn to follow the same curvature used for the circumferential placement of the holes. This arc is similar to the shape of the back vane and maximizes the available surface area for hole placement. This curvature-based method ensures that the hole positioning effectively explores the available space between the adjacent vanes (the red zone in Figure 3.4b), enabling a broader coverage of the impeller rear shroud region. The radial distance from the impeller center to each hole was then systematically varied along this curved path during the design exploration process, within a range of 0.203 to 0.342. Here these two values represent the normalized value of D_1 and D_2 with respect to the impeller diameter.



(a) Radial position. **(b)** Feasible region for hole placement.

Figure 3.4: Balancing hole placement and feasible region in the semi-open impeller.

3.1.5 Finalization of Design Parameter Ranges

Although the diameter, radial, and circumferential position of the balancing holes were initially defined independently, they were ultimately coupled during the optimization process. In practice, this coupling introduced new geometrical constraints that necessitated refinement of the original parameter ranges. Specifically, the circumferential angle range between 0° and 60° was excluded, as the placements within this range would intersect with the back vanes, potentially compromising the vane structure. Similarly, the radial positioning range was adjusted to prevent the hole position from being too close to either the impeller center or its peripheral edge. In addition to this, the varying range of hole diameter contributed to these constraints, as larger diameters increased the risk of interference with key structural features. Furthermore, the front-side rotation required for through-hole drilling was adjusted to ensure geometric feasibility across all configurations. To effectively explore the geometric space while avoiding interference with critical impeller features, the circumferential and radial position ranges were carefully refined to avoid interference with any structural features within the impeller geometry. The final constrained ranges used in this study for the design parameters are summarized in Table 3.1.

Table 3.1: Design parameter ranges before and after geometric refinement.

Design Parameter	Initial Range	Constrained Range
Circumferential Position	60°	24°
Radial Position normalized with Impeller Diameter	0.203 to 0.342	0.245 to 0.3
Rotation	120°	66°

3.2 Computational Modeling

This section outlines the computational modeling strategy used for simulating the internal flow within the centrifugal pump considered in this study. Both the simulation setup and meshing strategies were provided by Xylem. While a base- line mesh for the entire computational domain was provided , local mesh refinement was performed near the balancing hole regions to re- solve the complex flow phenomena more accurately. All simulations were conducted using CFD++, with mesh generation and pre-processing performed in ANSA.

3.2.1 Computational Domain

The computational domain, as shown in Figure 3.5, represents the internal geometry of the centrifugal pump used in this study. The Frozen Rotor method was employed to model the rotor and stator interactions within the computational domain. This approach minimizes computational cost and time compared to transient simulations. To facilitate this modeling strategy, the computational domain is divided into two regions (refer to Figure 3.5):

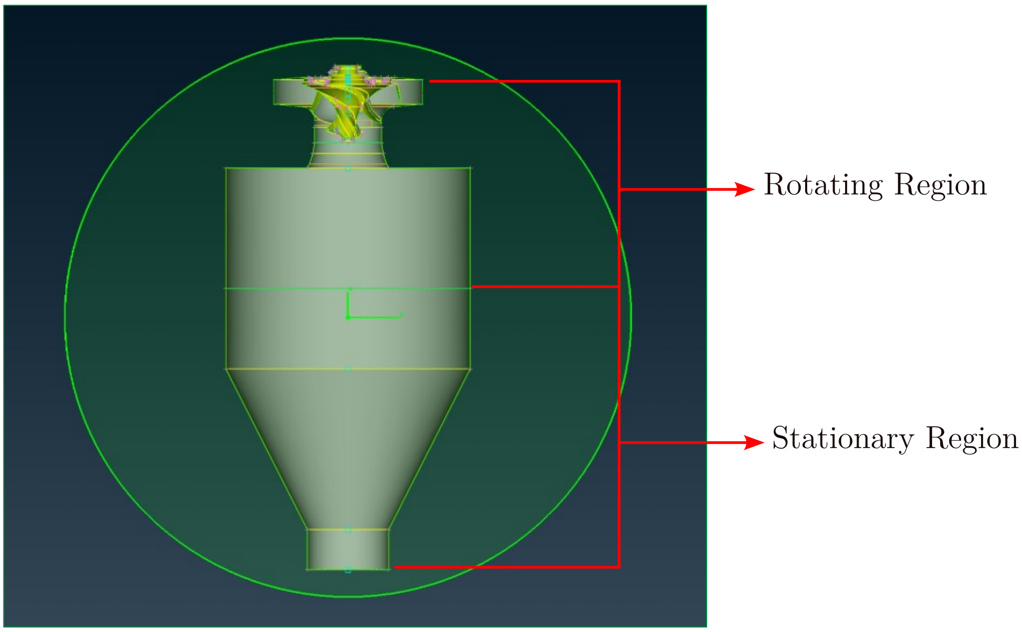


Figure 3.5: Computational domain.

- **Rotating Region:** The rotating region represents the part of the computational domain that is subjected to rotational motion. It was assigned a fixed rotation speed of 990 RPM to replicate actual pump operating conditions. Flow within this region is solved in a rotating frame of reference. However, it can be seen from Figure 3.5 that certain boundaries included within this region do not physically rotate in reality. This will be summarized in the upcoming Section 3.2.2. The computational domain is extended upstream, as shown in Figure 3.5, to include a stationary inflow section that minimizes the influence of numerical boundary conditions on the physical flow development near the impeller. Additionally, the rotating region is extended slightly upstream of the impeller to accurately capture the incoming flow characteristics and to provide a sufficient buffer for the smooth interpolation of flow variables across the rotating–stationary interface. This transition zone, necessitated by the use of different reference frames, helps maintain continuity in flow variable gradients, reduces numerical artifacts, and ensures accurate flow prediction at the impeller inlet.

- **Stationary Region:** This region represents the stationary part of the computational domain. The components assigned to this region remain fixed during the simulation and interact with the rotating region through interface boundary.

Although no volute is included in the current model and the domain appears axisymmetric, the stationary region is retained to allow for potential inclusion of a volute in future studies, which would break axisymmetry and require a fixed region.

3.2.2 Boundary Conditions

To ensure accurate flow modeling, appropriate boundary conditions were assigned to each surface of the computational domain. The conditions were assigned based on the physical behaviour of each component and the simulation strategy. Since the Frozen Rotor simulation strategy were used to handle the interaction between the rotating and stationary regions, boundaries in those regions were defined either in the global (stationary) frame of reference or local (rotating) frame of reference, depending on the actual motion of the components.

As shown in Figure 3.6, the boundaries were labeled according to the Property IDs (PIDs) in the ANSA pre-processing environment. The names are consistent with the physical surfaces represented in the geometry and distinguish between rotating and stationary regions. The assigned condition for each boundary is summarized as:

1. **Outlet**: Pressure Outlet.

2. **Rotating Smooth**: Wall (Viscous, Wall Function).

3. **Slip**: Wall (Inviscid).

4. **Inflow**: Mass Flow Rate (Value depends on Operating condition of the pump).

5. **Stat Rough In Rot**: Wall (Viscous, Wall Function, Rotation about Z–axis, Rotation Rate: 0, Roughness: 0.00011 m).

6. **Rotating Rough Back**: Wall (Viscous, Wall Function, Stationary wrt Mesh Motion, Roughness: 0.00011 m).

7. **Rotating Rough Channel**: Wall (Viscous, Wall Function, Stationary wrt Mesh Motion, Roughness: 0.00005 m).

8. **Stationary Smooth In Rot**: Wall (Viscous, Wall Function, Rotation about Z–axis, Rotation Rate: 0).

9. **Diffuser Slip**: Wall (Inviscid).

10. **Inlet Interface**: Zonal (Patched Only, Simple Flow Through).

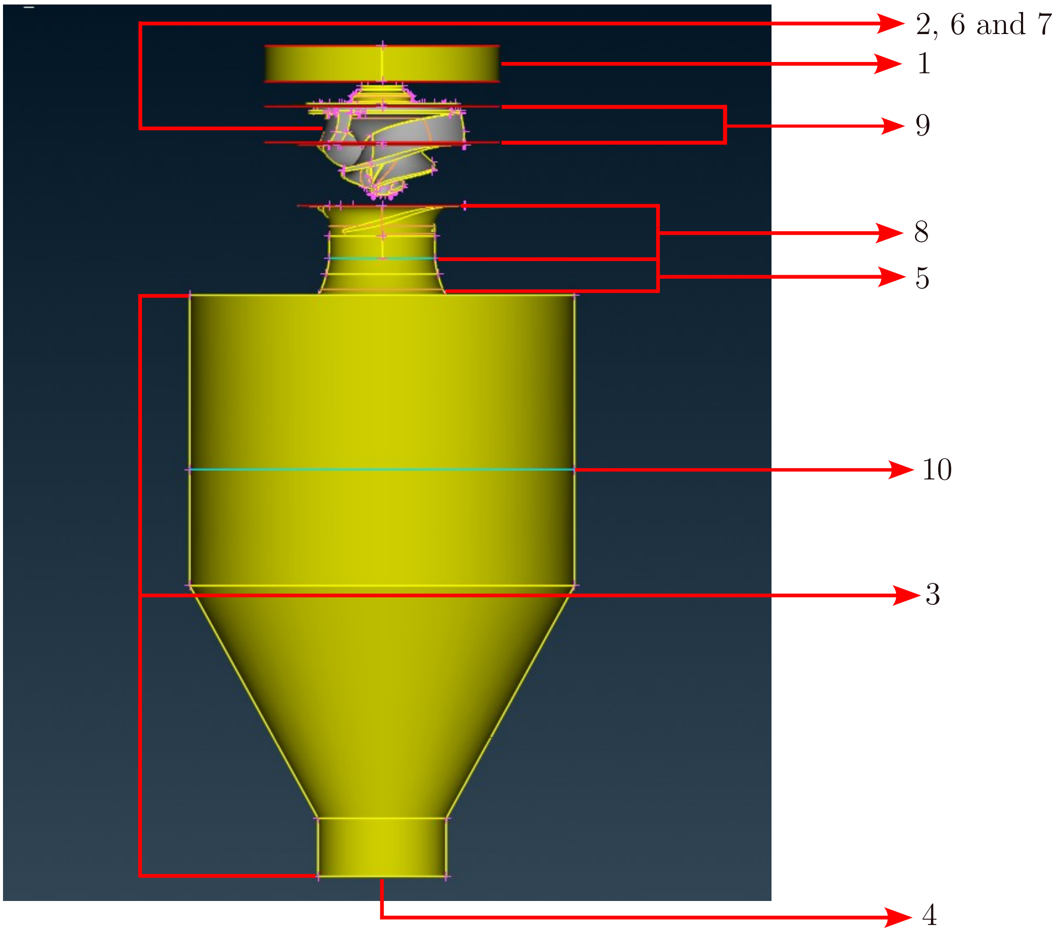


Figure 3.6: Boundary condition for the computational domain.

Some boundaries located within the rotating region are assigned zero rotation to reflect the real case where only the impeller rotates. The surface roughness values were provided by Xylem and are based on experimental characterization of the pump component's surfaces. The boundary condition types and recommended settings are based on the CFD++ user manual [21].

3.2.3 Computational Mesh

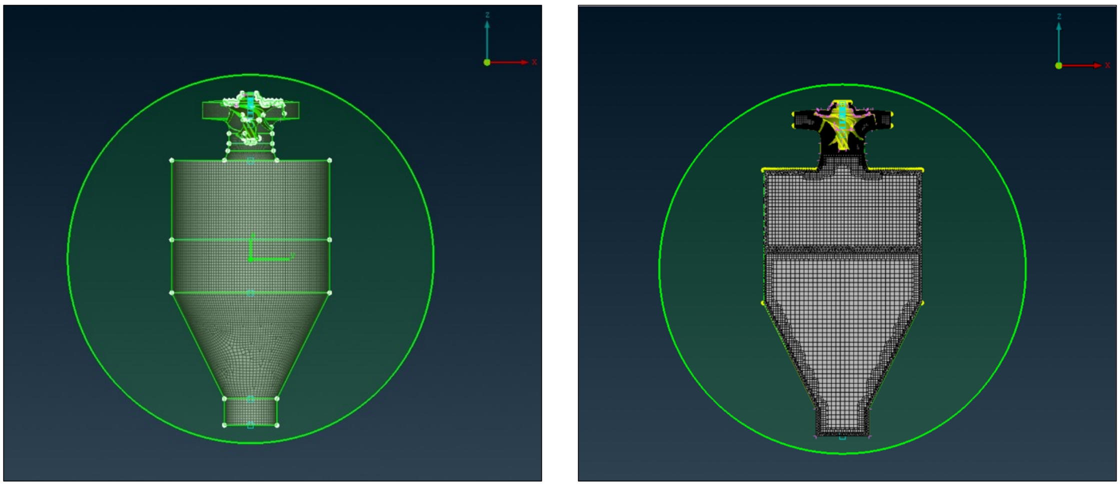
As previously mentioned in Section 3.2, the generated computational mesh in ANSA was provided by Xylem and validated through transient CFD simulations and measured performance. Since the baseline mesh did not include balancing holes, local refinement was applied in those regions. This section first details the baseline meshing strategy for the original geometry, followed by a mesh independence study focused on the balancing hole.

3.2.3.1 Baseline Mesh Configuration (Without Balancing Hole)

Figure 3.7 shows the sectional view of the computational domain with surface and volume meshes. The surface mesh uses a combination of quadrilateral and triangular elements to effectively mesh complex areas like the impeller blades, back vanes, narrow passages, and sharp trailing edges. This hybrid approach ensures a good mesh quality with reduced skewness in complex regions and enables the creation of high-quality surface discretization. For the volume mesh, a combination of hexahedral, tetrahedral, and pyramidal elements was used throughout the computational domain. These elements are chosen for their favorable numerical properties, where hexahedral elements offer improved accuracy, reduced numerical diffusion, and enhanced solver stability due to their structured nature, while tetrahedral and pyramidal elements provide flexibility for meshing complex geometries. This meshing strategy helps in better resolution of the pressure and velocity gradients in intricate regions of a centrifugal pump geometry. Additional mesh details near the impeller and boundary layer regions are shown in Figures 3.8 and 3.9. The mesh settings are summarized in Table 3.2:

Table 3.2: Mesh settings used in the original configuration

Parameter	**Description / Value**
Surface Mesh Elements	Hybrid mesh (Quadrilateral + Triangular)
Surface Element Size	Ranges from 0.8 mm to 30 mm
Surface Mesh Growth Rate	1.2
Volume Mesh Type	Hybrid mesh (Hexa + Tetra + Pyramid)
Volume Element Size	Maximum element size of the surface mesh
Volume Mesh Growth Rate	1.2
Number of Boundary Layers	7
First Cell Height	0.2 mm
Boundary Layer Growth Rate	1.2
Total Element Count	13.5 million cells



(a) Surface mesh.

(b) Volume mesh.

Figure 3.7: Computational mesh of the domain

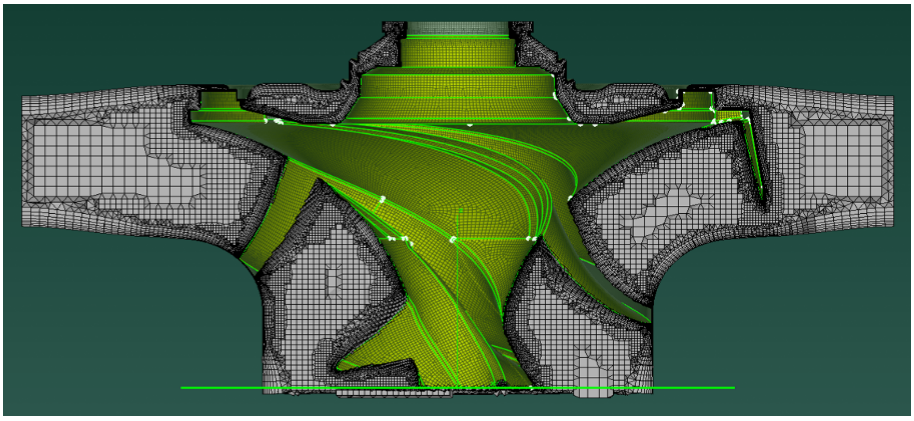
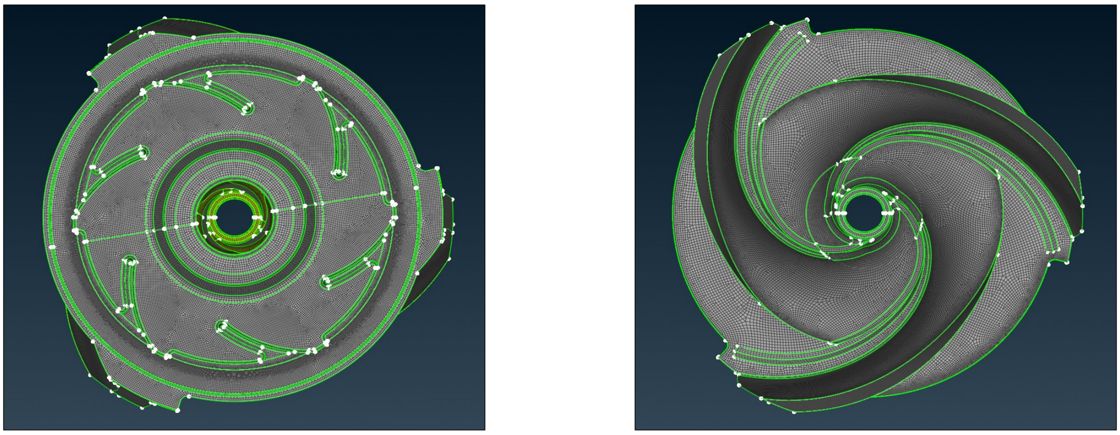


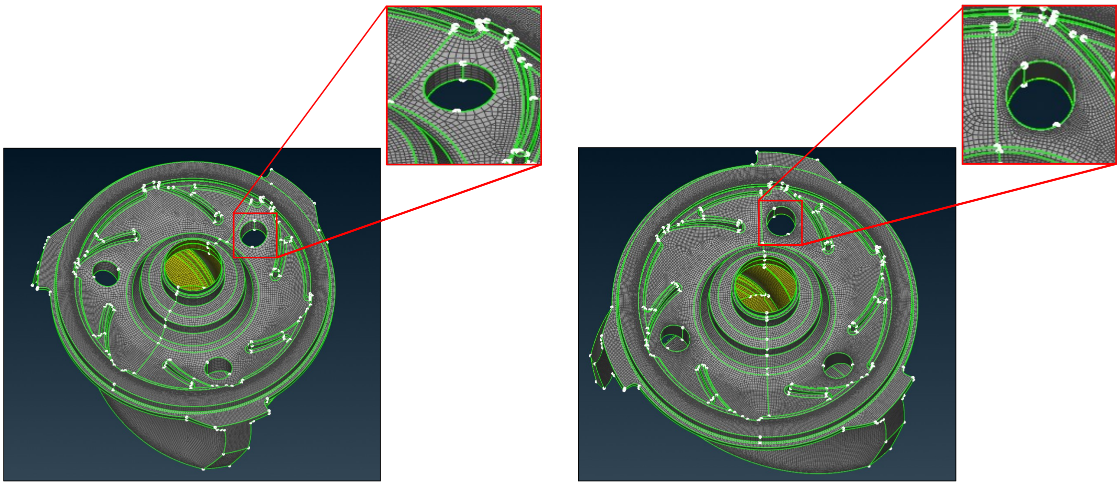
Figure 3.8: Boundary layer mesh.



(a) Pressure / Back side.

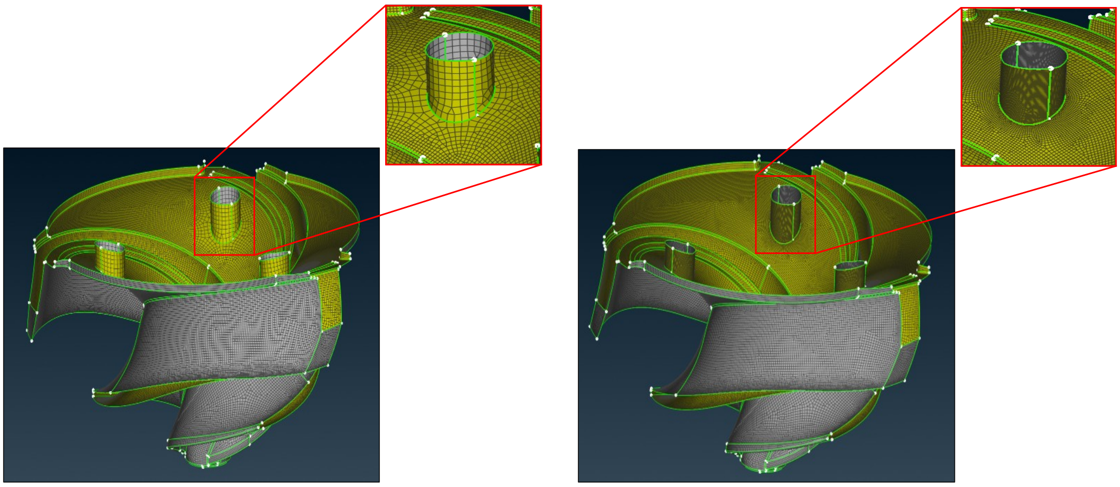
(b) Suction / Front side.

Figure 3.9: Surface mesh on the semi-open impeller without balancing hole.


(a) Coarse mesh.

(b) Fine mesh.

Figure 3.10: Surface mesh refinement – back side (Balancing hole diameter = $D_{\max}$).

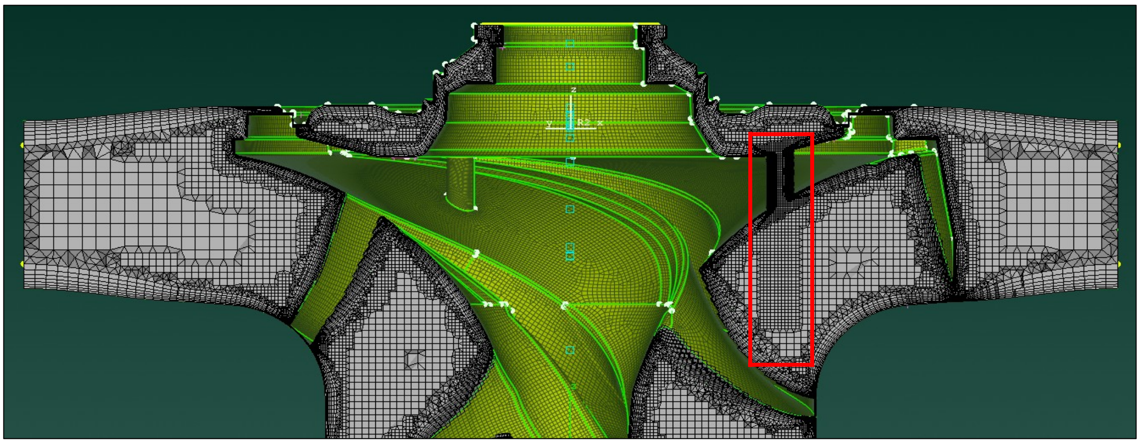

(a) Coarse mesh.

(b) Fine mesh.

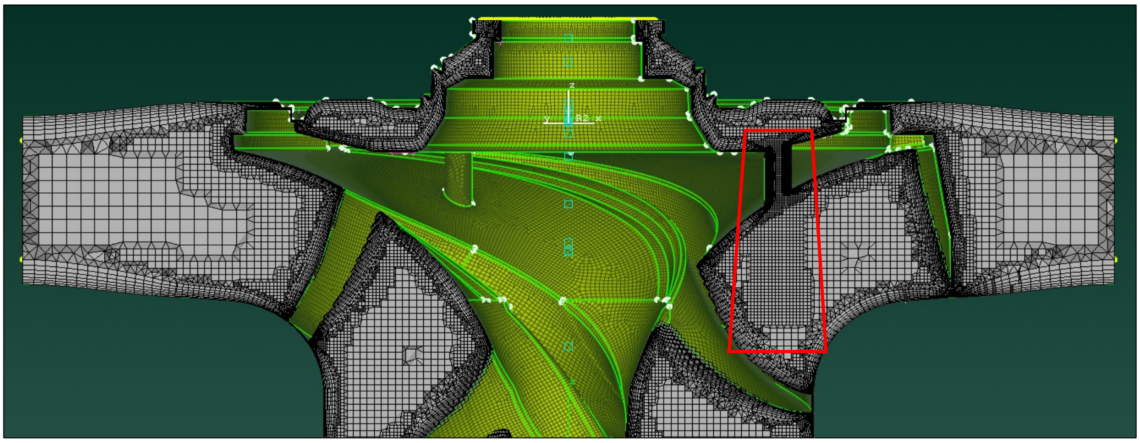
Figure 3.11: Surface mesh refinement – inner side (Balancing hole diameter = $D_{\max}$).

3.2.3.2 Mesh Refinement Study around Balancing Hole

To ensure accurate resolution of flow characteristics near the balancing hole, a mesh independence study was carried out with a focus on localized refinement. Initially, surface mesh refinement was applied around the balancing hole on the impeller while maintaining the baseline volume mesh settings. Element sizes ranged from 0.8 mm to 7 mm. Figures 3.10 and 3.11 presents surface mesh examples for coarse and fine refinements, applied to the maximum diameter ($D_{\max}$) range of the balancing hole.



(a) Cylindrical shape.



(b) Conical shape.

Figure 3.12: Volume mesh refinement (Balancing hole diameter = D_{min}).

Due to the expected strong pressure gradients and jet-like flow behavior caused by throttling at the hole, volumetric mesh refinement was also tested. Two refinement zone geometries were evaluated.a cylindrical shape and a conical shape (refer to Figure 3.12). These refinements increased local mesh density and improved resolution of the flow structures within and around the balancing hole.

The final mesh settings selected through this study provided sufficient resolution for all tested balancing hole configurations and were used for simulation results discussed in Section 4.2. This mesh refinement study was performed at the BEP and included surface refinement for both the smallest (D_{min}) and largest (D_{max}) balancing hole diameters, while volume refinement was evaluated only for the smallest diameter configuration using a surface mesh element size of 1 mm.

3.2.4 Turbulence Modeling and Solver Settings

The CFD simulations in this study were performed using a steady-state approach with solver settings and turbulence model configurations based on Xylem's industrial CFD++ setup. A non-linear cubic k-epsilon model (refer to Section 2.3.3) was used to model the turbulence. This model was chosen to accurately resolve the complex flow features that arise in centrifugal pumps with balancing holes, particularly the secondary flow caused by rotational effects and the related turbulent anisotropy, flow separation, and recirculation near both the balancing hole and the impeller blades. In addition to this, a non-equilibrium Launder-Spalding wall function was used to ensure accurate near-wall modeling in regions affected by strong pressure gradients and boundary layer separation, which are dominant near impeller blades and balancing holes. To maintain physical consistency in the turbulence quantities throughout the domain, the Schwartz realizability criterion was also selected.

The numerical solver used an implicit time-stepping scheme with local time- step control to accelerate convergence while maintaining numerical stability. A second-order upwind scheme was used for spatial discretization to improve numerical accuracy and effectively resolve the flow gradients. Further, under-relaxation factors were used for both the coupled pressure-velocity solver and turbulence equations to ensure stability during iterations. The solver and the turbulence model settings used in this study are summarized in Table 3.3. These settings ensured stable convergence, accurate turbulence prediction, and robust performance across all operating conditions.

Table 3.3: Turbulence modeling and solver settings

Setting	**Value / Description**
Solver Type	Pseudo-Steady-State
Time Integration	Local time stepping (Implicit)
Turbulence Model	Cubic k–ε
Wall Function Methodology	Launder–Spalding
Wall Function Type	Non-Equilibrium
Wall Function Pressure Gradient Term	Enabled
Discretization	Second Order
Under-Relaxation Factor	0.5 (Coupled Solver), 0.95 (Turbulence)
Working Fluid	Incompressible water at 293.15 K

3.3 Initial Sampling for the Optimization Process using Sobol Sequence

To initiate the optimization process and to ensure comprehensive coverage of the design space, the Sobol sequence was employed for sampling. As previoiusly mentioned, four design parameters were considered in this study: balancing hole diameter (**BH_Diameter**), the circumferential position (**BH_Circumferential_Back & BH_Circumferential_Front**) in the front side of the impeller and back side of the impeller as well as its radial position (**BH_Radial_Position**).

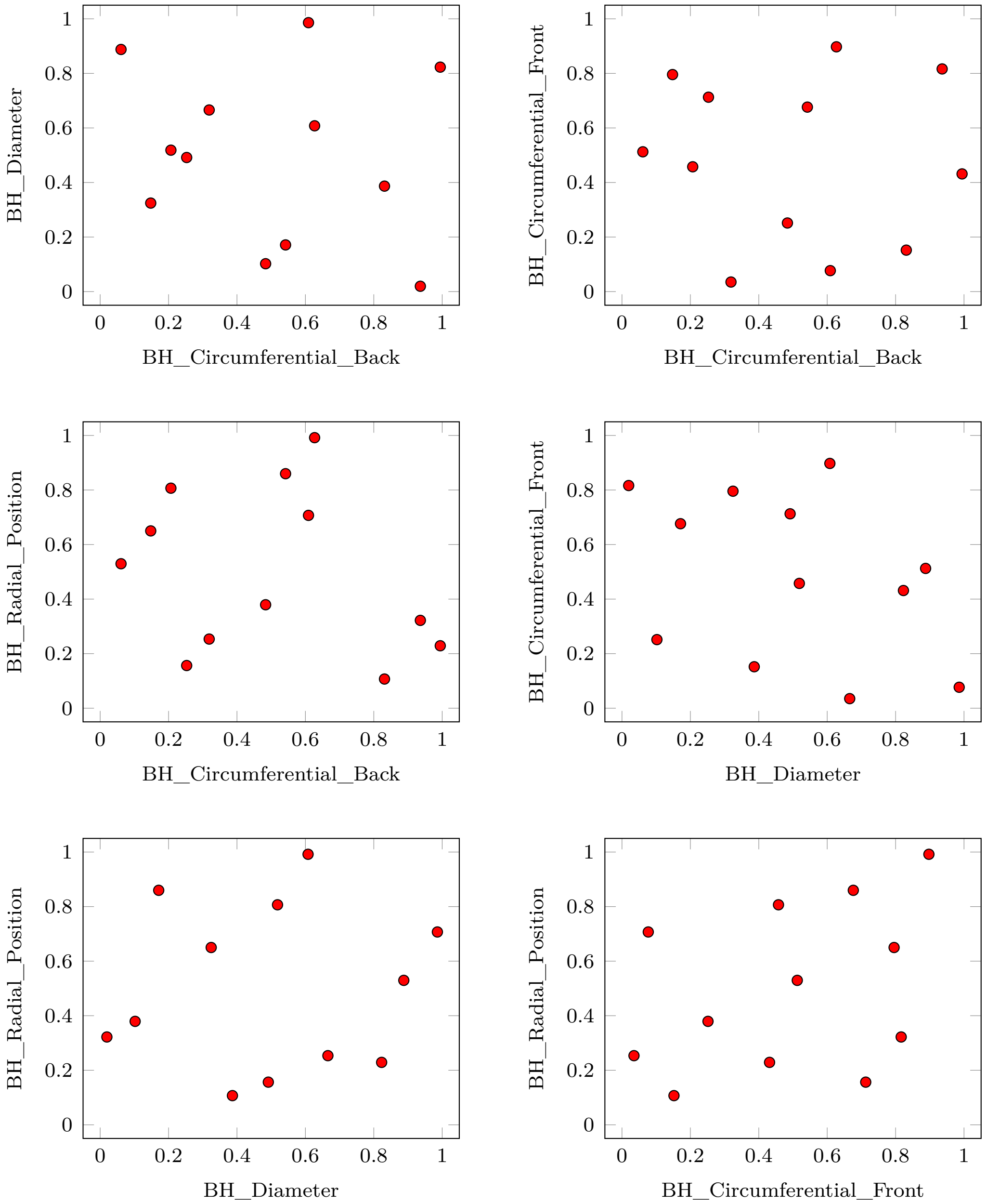


Figure 3.13: Pairwise Scatter Plots of Initial Sobol Samples.

Due to the high computational cost associated with CFD simulations, 12 initial design points were considered and it corresponds to 3 times the total number of design parameters. This sampling strategy aimed to strike a balance between adequate exploration of the design space and computational efficiency, particularly in the early phases of the optimization. Furthermore, it ensured that the surrogate model was initialized with sufficient data to support the subsequent Bayesian optimization process.

The 12 initial seed samples, each representing a unique design configuration, were generated using the Python code script provided in appendix A.1. The actual code for the Sobol sequence samples, developed by Xylem, is not included in this report due to confidentiality restrictions. Figure 3.13 illustrates how the initial seed samples are distributed across the design space. Table 3.4 summarizes the normalized design parameter values—such as back, diameter, front, and radial locations—for a subset of these Sobol-generated configurations. The performance results of these 12 design configurations, evaluated at four different flow rates ($Q^* = 0.50,\ 0.75,\ 1.00,$ and 1.25), are discussed in detail in the results chapter under section 4.8.

Table 3.4: Design parameters and their spatial positions for the 12 initial Sobol-seeded configurations.

Configuration	Back position	Hole diameter	Front position	Radial position
1	0.678	0.858	0.903	0.719
2	0.038	0.067	0.166	0.007
3	0.333	0.534	0.705	0.885
4	0.944	0.258	0.480	0.330
5	0.781	0.704	0.118	0.443
6	0.426	0.495	0.826	0.778
7	0.188	0.904	0.304	0.142
8	0.581	0.179	0.504	0.571
9	0.512	0.618	0.329	0.872
10	0.156	0.331	0.598	0.411
11	0.481	0.802	0.031	0.541
12	0.874	0.030	0.792	0.236

3.4 Target Function Development Strategy

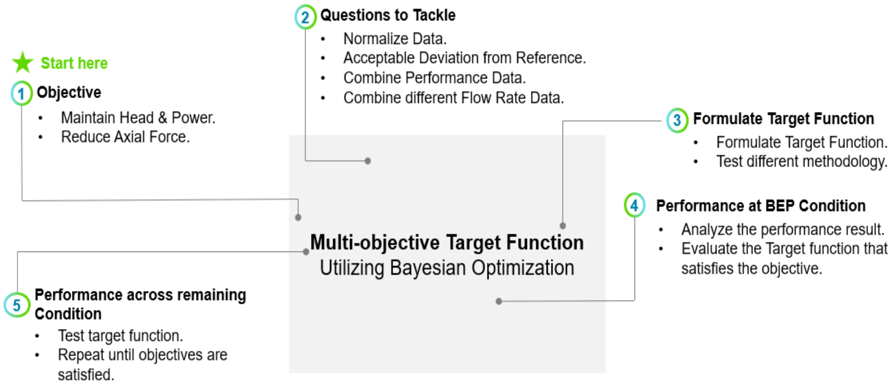


Figure 3.14: Workflow for multi-objective target function development.

To construct a robust multi-objective target function for the optimization of balancing hole design configurations, a systematic approach was employed. This approach involved the integration of key performance metrics, such as head developed, power consumption, and axial force, into a unified evaluation framework. The resulting target function enables the rankings of design configurations based on their ability to maintain hydraulic performance while minimizing axial force. The development and refinement process of the target function is outlined in Figure 3.14, which highlights the key steps from the evaluation of the initial samples generated by the Sobol sequence to function refinement. This section describes the approach under two main categories: normalization strategy and target function formulation.

3.4.1 Normalization Strategy

Normalization procedure is a crucial component of this study, since the considered performance parameters have different units and magnitudes. Without appropriate normalization, the target function could be biased toward parameters with larger numerical values, potentially undermining the optimization process. The normalization procedures adopted are discussed in the following subsections.

3.4.1.1 Normalization of Head and Power

Initial normalization was carried out using the standard deviation derived from the initial design configurations data obtained using the Sobol sequence sample. Although statistically valid, this approach introduced challenges due to the uneven spread of normalized values, which made it challenging to construct a stable and interpretable target function. The normalization formula using the standard deviation is expressed as:

$$\text{Normalized Parameter} = 1 - \left(\frac{(X_i - X_{ref})^2}{\sigma_X^2} \right) \tag{3.1}$$

where, X_i is the simulated result of the quantity (Head or Power) for the i^{th} impeller configuration with balancing holes at BEP ($Q^* = 1.0$) condition, X_{ref} is the simulated result of the same quantity for the reference impeller without balancing holes at BEP ($Q^* = 1.0$) condition, and σ_X is the standard deviation of the simulated results (Head or Power) across the 12 initial configurations generated using the Sobol sequence at BEP ($Q^* = 1.0$) condition.

To represent this issue more clearly, two extreme results from the BEP ($Q^* = 1.0$) condition were considered. These results are presented in Table 3.5. The head (H_i) was normalized with the head (H_{ref}) for the BEP condition.

Table 3.5: Illustration of normalization error using standard deviation.

Configuration	$\frac{H_i}{H_{\text{ref}}}$ (%)	σ_X	**Normalized Data**
1	100.03	0.129	0.9759
7	99.37	0.129	-7.2222

This example highlights how minor differences in head can yield disproportionately skewed normalized values. A similar behavior was observed for power. To address this, re-normalization of the already normalized data using min-max scaling was explored. This method maps the data into a $[0, 1]$ range using the transformation:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3.2}$$

Applying this to the normalized head values in Table 3.5, the re-normalized values become:

$$x'_{\text{Config 1}} = \frac{0.9759 - (-7.2222)}{0.9759 - (-7.2222)} = 1.000, \quad x'_{\text{Config 7}} = \frac{-7.2222 - (-7.2222)}{0.9759 - (-7.2222)} = 0.000$$

Although this re-scaling constrains the data within the desired range, it does not correct the core distortion. The original normalization already exaggerated differences due to the small standard deviation. As a result, min-max normalization merely rescales an already skewed distribution, compressing real variations and flattening intermediate configurations to values near 0 or 1.

Therefore, this re-normalization method was ultimately not adopted, as it further reduced the interpretability of the normalized data and was inconsistent with the goal of constructing a smooth and meaningful target function.

A more practical method was adopted: normalizing the data using a fixed percentage of the reference value (the impeller without balancing hole). The new normalization formula used is:

$$\text{Normalized Parameter} = 1 - \left(\frac{(X_i - X_{\text{ref}})^2}{(X_{\text{ref}} \cdot \delta)^2}\right) \tag{3.3}$$

where, δ is the fixed percentage. $\delta = 0.2\ \%$ and $\delta = 1\ \%$ are used for power and head respectively.

This approach ensured that the normalized data consistently fall within the range $[0, 1]$, preserved the natural variation between samples, and enabled the construction of a more robust and interpretable multi-objective target function. Hence, all subsequent results and optimization strategies in this study are based solely on this fixed-percentage normalization method.

3.4.1.2 Normalization of Axial Force

In centrifugal pumps, axial force varies with flow rate, increasing at low flow rates due to the larger pressure imbalance across the impeller. Thus, normalization must consider multiple operating conditions.

Initially, the axial force was normalized with respect to the reference value of the impeller at the BEP condition:

$$\text{Normalized Axial Force} = \left(\frac{(F_i - F_{\text{ref, 1.0}})^2}{F^2_{\text{ref, 1.0}}}\right) \tag{3.4}$$

where, F_i is the simulated axial force for the i^{th} impeller configuration with balancing holes at BEP ($Q^* = 1.0$) condition, $F_{ref,1.0}$ is the simulated axial force for the reference impeller without balancing holes at BEP ($Q^* = 1.0$) condition.

While simple and useful for preliminary analysis, this approach fails to account for behavior under off-design conditions, particularly at partial flow rates where axial force is more critical. This could result in suboptimal configurations being favored

To address this limitation, a second normalization method was introduced. This method is particularly focused on identifying the worst-case axial force across two operating conditions: $Q^* = 0.50$ and 1.0. Equation 3.5 was used to normalize the axial force based on the maximum force:

$$\text{Normalized Axial Force} = \frac{F_{\text{ref, 0.5}} - \max(F_{\text{i, 0.5}}, F_{\text{i, 1.0}})}{|F_{\text{ref, 0.5}}|} \tag{3.5}$$

where, $F_{\text{i, 0.5}}$ is the simulated axial force for the i^{th} impeller configuration with balancing holes at part-load ($Q^* = 0.50$) condition, $F_{\text{i, 1.0}}$ is the simulated axial force for the i^{th} impeller configuration with balancing holes at BEP condition, and $F_{ref,0.5}$ is the axial force for the reference impeller without balancing holes at part-load condition.

By focusing on worst-case across operating ranges, the normalized value obtained using Equation 3.5 prioritized the configurations that achieved more reduction in axial force. Thus, this normalization strategy based on the trade-off across operating conditions ensured a meaningful comparison and reflected the true nature of balancing hole configuration, that is crucial in developing a multi-objective target function.

3.4.2 Target Function Formulation

Once the normalization strategies were finalized for the considered quantities (head, power, and axial force), the next step was to combine these into a single target function for ranking design configurations. Two different methods were explored:

3.4.2.1 Geometric Mean Method

The geometric mean method combines the normalized quantities of head, power, and axial force by taking the cube root of the product, and it is expressed as:

$$TF = \left((\text{Normalized}_{\text{P}}) \cdot (\text{Normalized}_{\text{H}}) \cdot (\text{Normalized}_{\text{AF}})\right)^{1/3} \tag{3.6}$$

The advantage of this method is that it is easy to construct, and treats each objective equally, making it suitable when no specific priority is offered to any one quantity. However, this has a significant limitation since it is overly sensitive to objectives that are performing low. For example, a configuration that performs well in two objectives (head and power) might perform poorly in another objective (axial force). This leads to a lower score in the target function, which might shift the optimization process towards non-optimal solutions. Recalling the objective of this study, it is to minimize axial force and maintain the overall pump performance. So, this method is not a good option.

3.4.2.2 Scalarized Form with Inverse Deviation Weightage Method

To overcome the limitations of geometric mean-based target functions, an alternative method using weightage was adopted. As discussed briefly in Section 2.4.1.1, the weight for each considered quantity is based on the inverse of its deviation from the reference configuration across all configurations.

The inverse deviation for each quantity is calculated as:

$$\Delta = \frac{1}{N}\sum_{i=1}^{N}|X_i - X_{\text{ref}}| \tag{3.7}$$

$$Inv = \frac{1}{\Delta} \tag{3.8}$$

Then, the weight for each considered quantity is computed as:

$$w_P = \frac{1/\Delta_P}{(1/\Delta_P) + (1/\Delta_H) + (1/\Delta_{AF})} \tag{3.9}$$

$$w_H = \frac{1/\Delta_H}{(1/\Delta_P) + (1/\Delta_H) + (1/\Delta_{AF})} \tag{3.10}$$

$$w_{AF} = \frac{1/\Delta_{AF}}{(1/\Delta_P) + (1/\Delta_H) + (1/\Delta_{AF})} \tag{3.11}$$

Thus, the final scalarized form of the target function is:

$$TF = w_P \cdot (\text{Normalized}_{\text{P}}) + w_H \cdot (\text{Normalized}_{\text{H}}) + w_{AF} \cdot (\text{Normalized}_{\text{AF}}) \tag{3.12}$$

This method gives greater importance to quantities that deviate less from the reference configuration value and assigns weightage based on the data available and not on arbitrary weights. Although it performs well for head and power, this weightage based on inverse deviation is not reliable for axial force as it largely deviates from the reference value (illustrated in Figure 4.11). The weightage assigned to the axial force is very low comparatively, meaning its effect was ignored in the final ranking. To rectify this, a modified inverse was introduced:

$$Inv_{AF} = \frac{Inv_P + Inv_H}{2} \tag{3.13}$$

This approach ensures that axial force retains meaningful influence in the optimization process, better capturing the trade-offs among objectives.

3.4.3 Bayesian Optimization & Pareto Front Analysis

The earlier methods for creating target functions were tested repeatedly by making new samples that used the same target function to rank them in the Bayesian optimization. This optimization implicitly helped the development of the target function and also allowed exploration of the design space with minimal computational cost. Once the sampling process had reached a certain limit, a Pareto front analysis was performed to evaluate the trade-offs between those considered quantities and helped identify the non-dominated configurations within the design space, offering more informed choices when balancing multiple objectives.

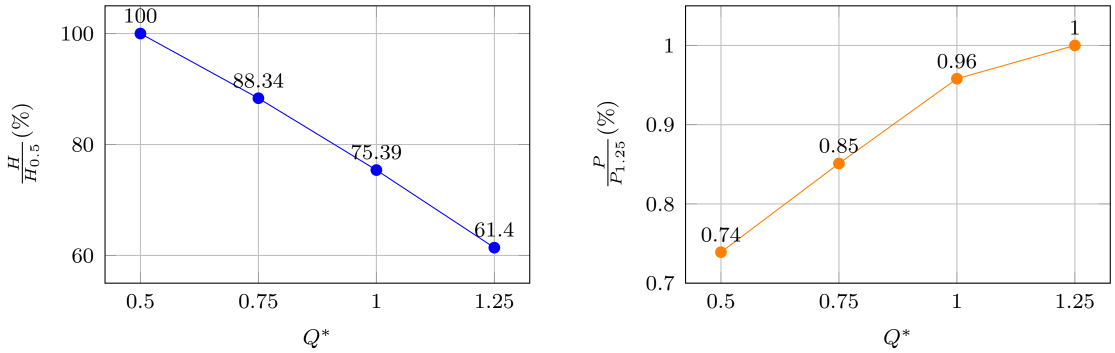
In this study, the Pareto front analysis was carried out using the Python code attached in appendix A.2, and the Bayesian optimization process was implemented using the inbuilt framework provided by Xylem.

4

Results

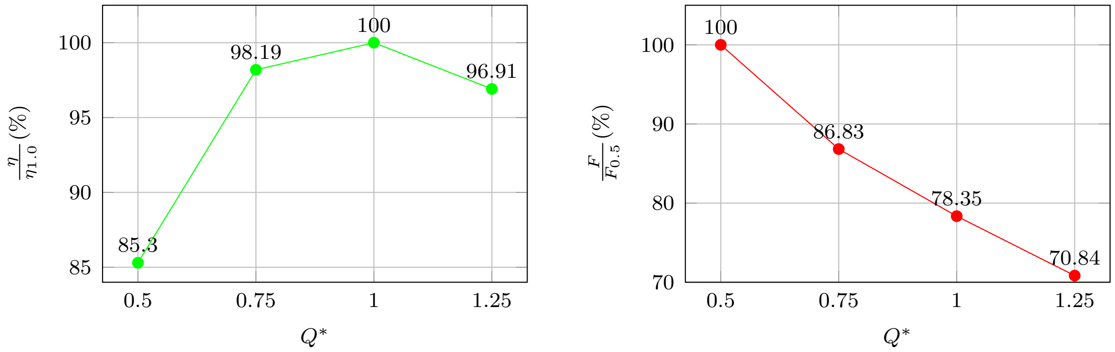
4.1 Performance of Reference Impeller (without Balancing Hole)

The hydraulic performance of the reference impeller—including head developed, power consumption, efficiency, and axial force—was analyzed at four different flow rates: $Q^* = 0.50,\ 0.75,\ 1.00,$ and 1.25, under a constant rotational speed of 990 RPM. The corresponding results are presented in Figure 4.1.



(a) Head Developed

(b) Power



(c) Efficiency

(d) Axial Force

Figure 4.1: Performance results of the reference impeller.

As expected, the head decreased with increasing flow rate, while power consumption and axial force decreased. The pump efficiency peaked around the design flow rate

($Q^* = 1.0$), confirming it as the BEP. These trends are consistent with the typical performance characteristics of a centrifugal pump, as discussed in Section 2.2 and serve as the baseline for evaluating all balancing hole configurations generated in this study.

4.2 Results of Mesh Refinement Study Around the Balancing Hole

This section discusses the mesh refinement study results for the two balancing hole diameters (D_{min} and D_{max}). Table 4.1 below summarizes the element sizes and corresponding number of elements for both diameters.

Table 4.1: Number of elements for corresponding element size.

Element Size	**Number of Elements**	
	D_{min}	D_{max}
Surface Refinement		
7 mm	13.53	13.37
5 mm	13.56	13.42
3 mm	13.60	13.52
1 mm	14.15	14.77
0.8 mm	14.32	15.14
Volume Refinement		
Conical	14.16	—
Cylindrical	14.18	—

As mentioned in Section 3.2.3.2, volume mesh refinement was only applied to the smallest balancing hole diameter. Figures 4.2 and 4.3 illustrate that the variation of head, power, efficiency, and axial force for different mesh sizes stays within acceptable limits for both small and large balancing hole diameters. Therefore, additional volume refinement was deemed unnecessary.

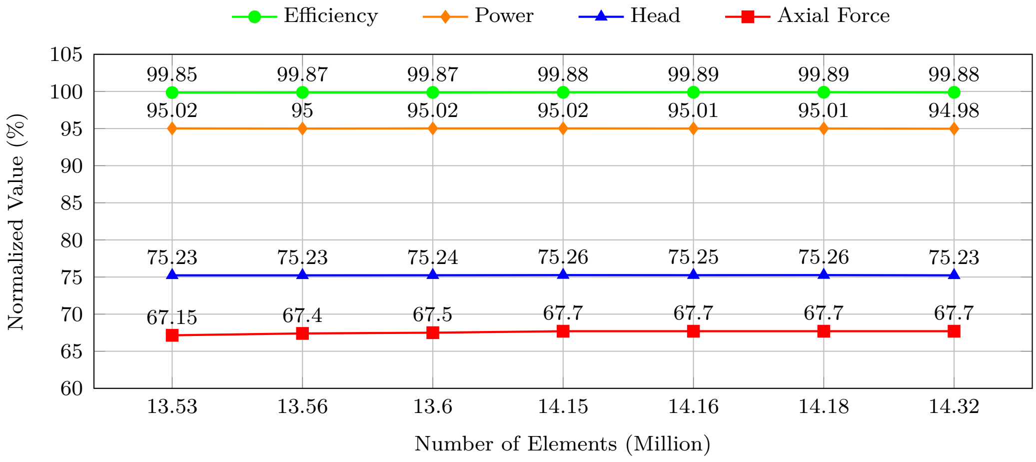


Figure 4.2: Results of mesh sensitivity analysis (Balancing hole diameter = D_{min}).

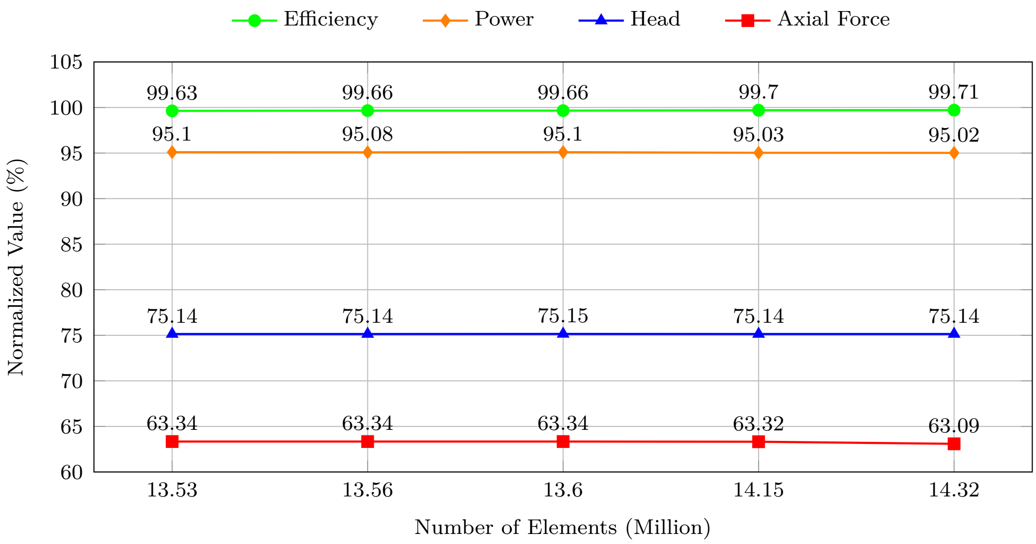
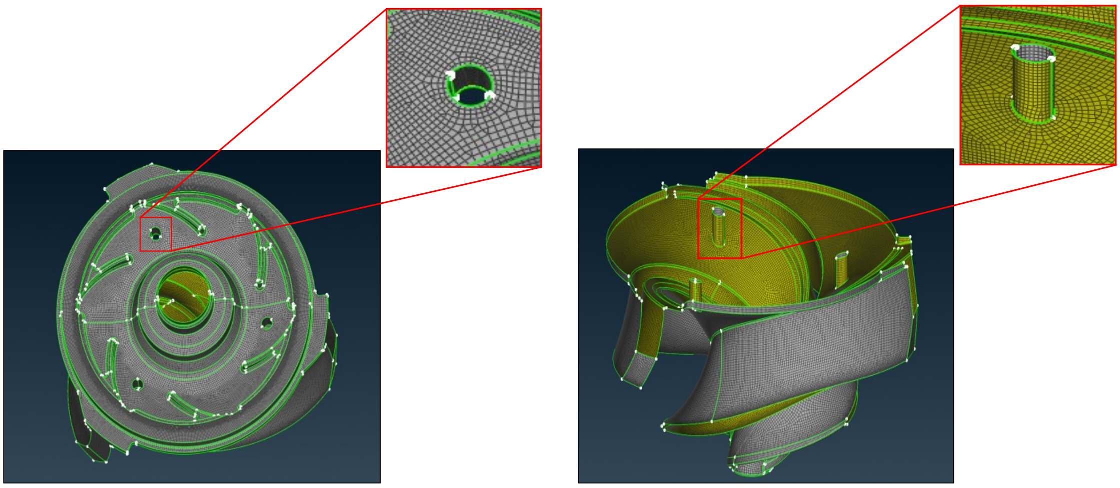


Figure 4.3: Results of mesh sensitivity analysis (Balancing hole diameter = $D_{\max}$).

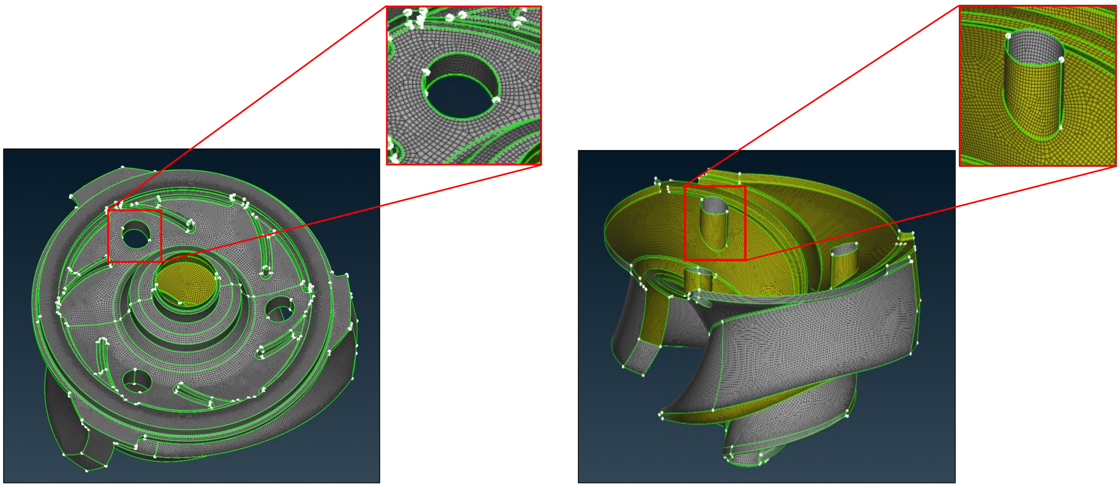
Based on these observations, a surface mesh element size of 3 mm was chosen for both small and larger balancing hole diameters. This element size provided superior structured topology with better mesh quality near complex regions and smoother transitions. Figures 4.4 to 4.7 show the finalized surface and volume mesh for both the smallest and largest balancing hole diameters. In addition to the mesh settings summarized in Table 3.2 for the impeller without balancing holes, a 3 mm surface mesh element size with no additional volume mesh refinement was used for the impeller with balancing holes. This ensured consistency across all impeller configurations with balancing holes while maintaining sufficient mesh resolution.



(a) Pressure / Back side.

(b) Inner side.

Figure 4.4: Surface mesh around balancing hole ($D_{\min}$).



(a) Pressure / Back side. **(b)** Inner side.

Figure 4.5: Surface mesh around balancing hole ($D_{\max}$).

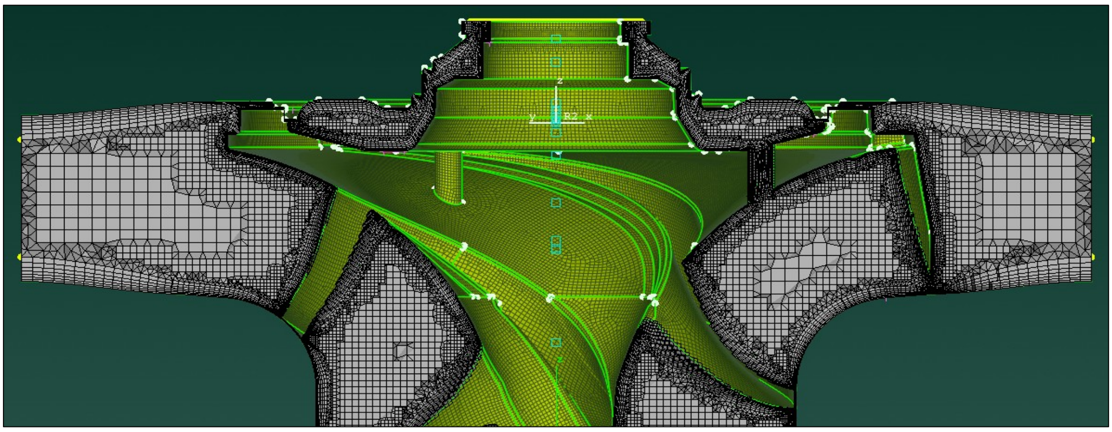


Figure 4.6: Volume mesh around balancing hole ($D_{\min}$).

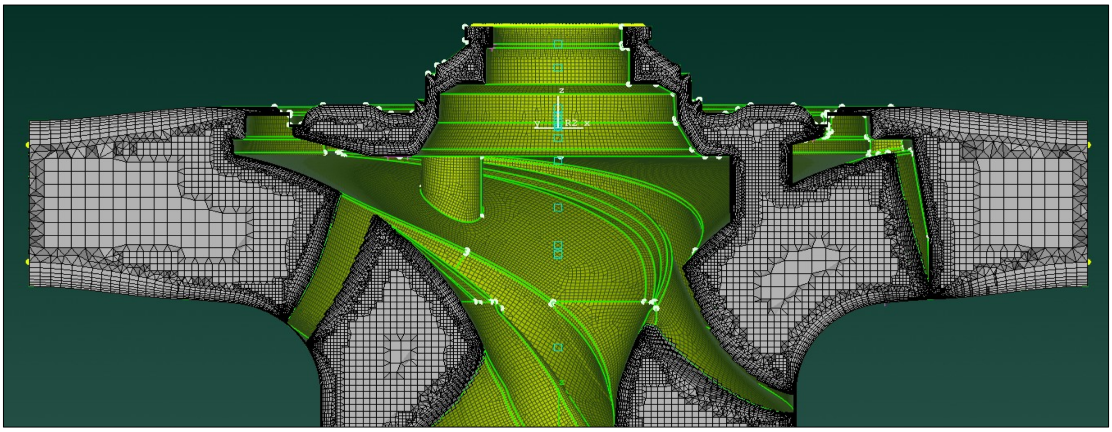


Figure 4.7: Volume mesh around balancing hole ($D_{\max}$).

4.3 Performance Evaluation of Sobol-Seeded Configurations

Twelve initial design configurations of the balancing hole were generated using Sobol sequence sampling and simulated. Their performance was evaluated relative to the reference impeller without a balancing hole at four different flow rates. The results were normalized with respect to the reference impeller to highlight the deviations within the design space and analyze the influence of the initial parameter variations. These configurations provide a basis for efficiently exploring the design space and serve as a foundation for subsequent surrogate modeling and optimization. The normalized results in percentage at BEP condition are shown in Figures 4.8 to 4.11. The additional flow rates, including $Q^* = 0.5$, 0.75, and 1.25, will be presented in the following sections.

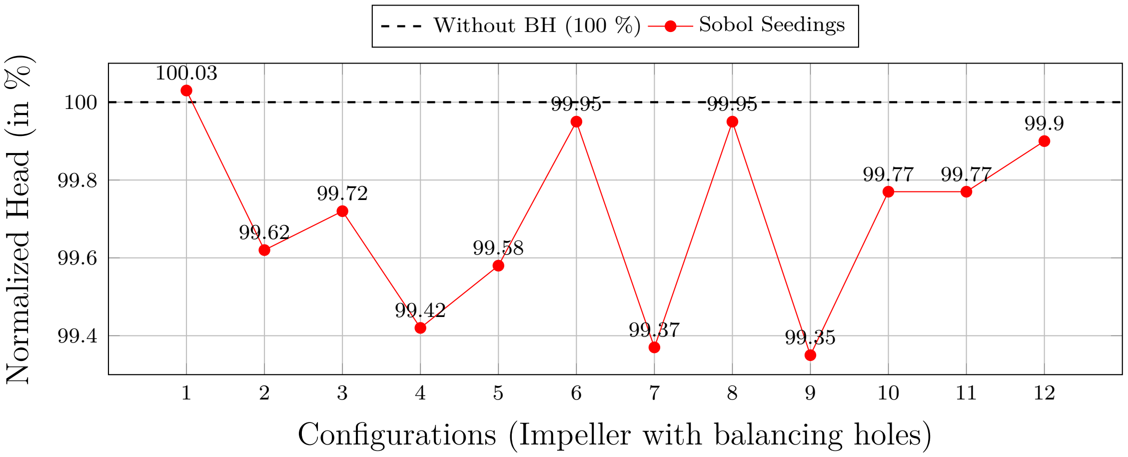


Figure 4.8: Normalized head for $Q^* = 1.0$.

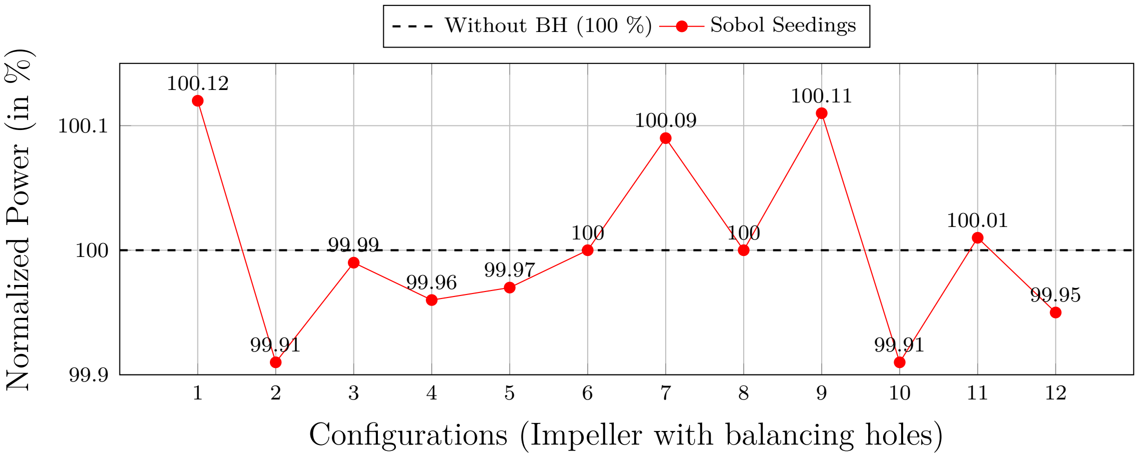


Figure 4.9: Normalized power for $Q^* = 1.0$.

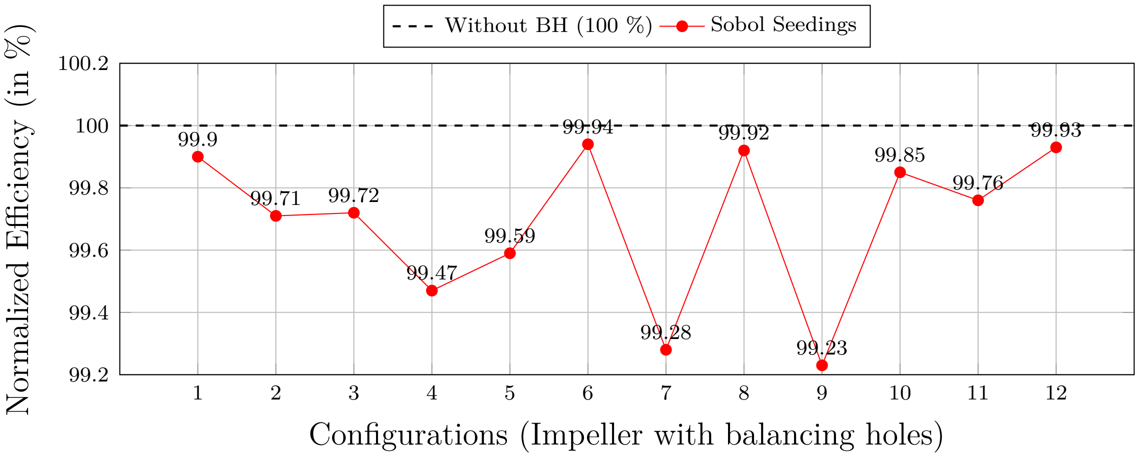


Figure 4.10: Normalized efficiency for $Q^* = 1.0$.

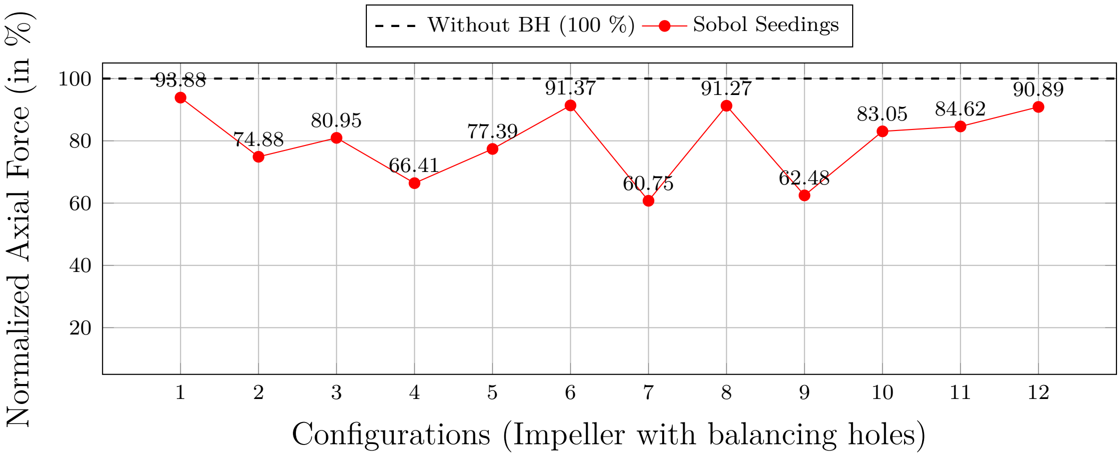


Figure 4.11: Normalized axial force for $Q^* = 1.0$.

Some observations can be drawn from the results of the initial Sobol-seeded configurations, as shown in Figures 4.8 to 4.11. For instance, configurations 6 and 8 maintain the power consumption at 100 % and slightly reduce the generated head, indicating that these configurations do not significantly alter the efficiency relative to the reference impeller.

As shown in Figure 4.11, the normalized axial force is expressed as a percentage relative to the axial force of the reference impeller without a balancing hole (100%). Hence, values close to 100% indicate lower reduction, meaning the axial force remains nearly unchanged, whereas lower values represent greater reductions. For instance, a value of 93.88%—as seen in configuration 1—implies that only 6.12% of the axial force has been reduced compared to the reference impeller. The variation across configurations—from 93.88% down to 60.75%—highlights that the generated Sobol seed designs span a wide range of performance, achieving both modest and substantial axial force reductions.

4.4 Normalization Results of Head & Power for Target Function

Based on the preceding results, the head and power values were normalized using the fixed percentage method described in Section 3.4.1.1 of the methodology chapter, and their corresponding results are illustrated in Figures 4.12 and 4.13. The resulting normalized values lie within the range [0, 1], where a higher score indicates performance closer to the desired target; that is, a minimal deviation from the reference impeller's performance.

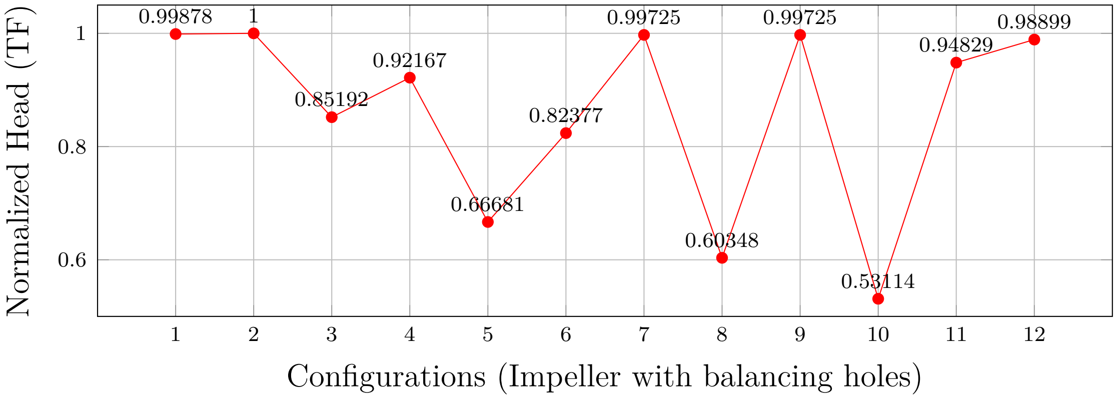


Figure 4.12: Head ranking for $Q^* = 1.0$.

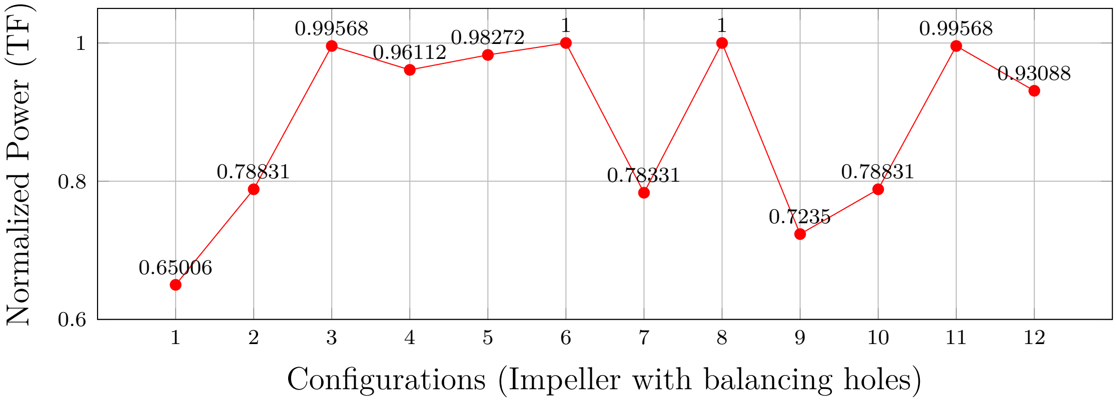


Figure 4.13: Power ranking for $Q^* = 1.0$.

4.5 Normalization Results of Axial Force for Target Function

To evaluate the effectiveness of different normalization approaches for axial force, two strategies introduced in Section 3.4.1 were compared. This evaluation was performed using the simulated results of the 12 Sobol-seeded configurations.

4.5.1 Baseline Normalization using BEP Reference ($Q^* = 1.0$):

As shown in Figure 4.14, the normalized axial force values obtained using Equation 3.4 fall within a very narrow range, resulting in reduced sensitivity between configurations, thereby diminishing the ability of the multi-objective target function to distinguish among design alternatives effectively.

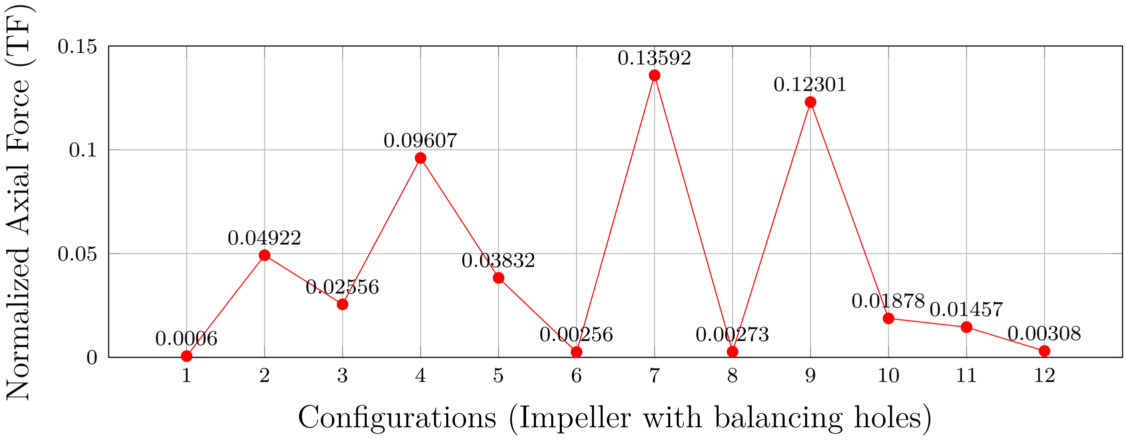


Figure 4.14: Axial force ranking $Q^* = 1.0$.

4.5.2 Worst-Case Normalization across Operating conditions

The normalized values obtained using Equation 3.5 are illustrated in Figure 4.15. For example, configuration 5 ranks higher than configuration 2 under the worst-case normalization, whereas the opposite was true using the baseline method. This is due to the fact that the reduction in axial force at part-load is much higher for configuration 5 than for the configuration 2 (refer to Figure 4.30). By capturing such behaviour, this approach better reflects performance under critical operating conditions and provides a more meaningful input to the multi-objective target function.

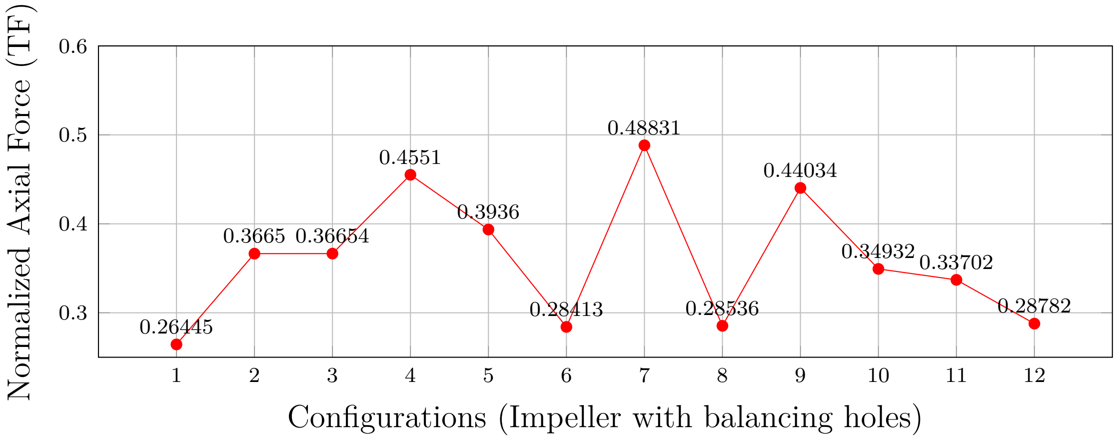


Figure 4.15: Axial Force Ranking $Q^* = 1.0$.

Figures 4.12, 4.13 and 4.15 provide a perspective on how each configuration ranks relative to the reference impeller, offering insights into which designs are more favorable. The outcomes presented here form the basis for the formulation of the multi-objective target function used in the subsequent Bayesian optimization process.

4.6 Optimization Using the Scalarized Form of the Target Function

This section describes and analyzes the results obtained during the optimization process based on the scalarized inverse deviation weightage method. This method combines normalized values of head, power, and axial force into a single multi-objective target function. Head and power were consistently normalized using the fixed-percentage method throughout the optimization process. However, axial force normalization was applied using two distinct approaches across different optimization cycles: Equation 3.4 for the initial cycle and Equation 3.5 for the subsequent cycle. The optimization process was carried out using Bayesian optimization to generate improved design configurations based on the ranking system provided by the target function.

As discussed Section 3.4.2.2, the scalarized target function used a weightage scheme based on the inverse deviation of each objective. However, the initial weightage obtained using Equation 3.7 resulted in a disproportionately low weight for axial force. To address this imbalance, a modified inverse formulation (Equation 3.13) was adopted to ensure that axial force contributes comparably to head and power. The weightage values obtained from both equations are summarized in Table 4.2. Figure 4.16 illustrates the scalarized form of target function values for all 12 configurations computed using the modified weightage formulation.

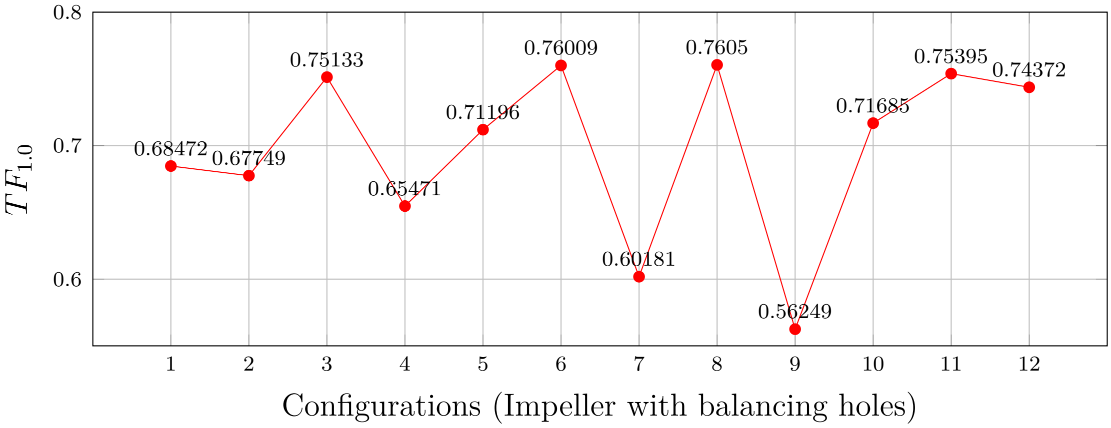


Figure 4.16: Scalarized form of target function (Inverse deviation weightage method).

Table 4.2: Comparison of weightage values for scalarized target function.

Type	$w_{\textbf{Power}}$	$w_{\textbf{Head}}$	$w_{\textbf{Axial Force}}$
Initial Inverse Method	0.2967	0.6990	0.0041
Modified Inverse Method	0.1986	0.4680	0.3333

4.6.1 First Optimization Cycle using Baseline Normalization of Axial Force

Following the evaluation of the 12 initial Sobol-seeded configurations, Bayesian optimization was used to generate new configurations of the balancing holes. The four design parameters were used as input, and the scalarized target function value (Figure 4.16) was the output to guide the optimization process.

Figure 4.40 provides a scatter plot in the design space, with the initial Sobol samples in red and the newly generated configurations in blue. Out of those 12 newly generated configurations, 6 configurations were selected for further simulation. The predicted target function scores obtained using Bayesian optimization are presented in Table 4.3.

Table 4.3: Selected configurations from bayesian optimization with target function scores.

Configuration	**Back**	**Diameter**	**Front**	**Radial**	$TF_{\textbf{Predicted}}$
13	0.23692	0.345456	0.650277	0.743552	0.998035
14	0.46840	0.019591	0.786387	0.608900	0.978371
15	0.06097	0.288414	0.681344	0.991838	0.964715
16	0.699475	0.354108	0.679049	0.590004	0.968395
17	0.225048	0.019591	0.631849	0.621766	0.968033
18	0.18765	0.664809	0.675100	0.847455	0.965946

The performance results of the 6 chosen configurations are displayed in Figures 4.23 to 4.26 for BEP and part-load conditions, respectively. At the BEP condition, the configurations show minimal deviation from the reference impeller. In contrast, the power consumption exhibits slightly more variation but remains within a narrow band observed from the initial 12 Sobol-seeded configurations. The most significant observation lies in the reduction of axial force. The maximum reduction in axial force achieved among the 6 configurations is only 10.68 %, compared to a 39.25 % reduction achieved by the initial 12 Sobol-seeded configurations. This indicates that while hydraulic performance was preserved, axial force reduction was limited.

Figure 4.27 to 4.38 show the performance results for the remaining three operating conditions. Similar to the BEP results, head and power closely followed the reference impeller's performance, but axial force reduction remains minimal across all conditions, confirming the observed trade-off. The trend remains consistent with the earlier observation: at all operating conditions, the objective of maintaining hydraulic performance across the selected configurations comes at the cost of a small axial force reduction.

4.6.2 Second and Third Optimization Cycle (Worst Case Normalization of Axial Force)

The first optimization cycle prioritized the hydraulic performance; however, this effort came at the cost of axial force reduction. To address this limitation, the second and third optimization cycles focused more dominantly on minimizing the axial force while maintaining acceptable hydraulic performance.

The updated design space for these two cycles is shown in Figure 4.40, and Table 4.4 lists the design parameters and predicted target function score values obtained using Bayesian optimization. In this phase, the target function and Bayesian optimization were based on the worst-case scenario for axial force under various operating conditions.

Table 4.4: Configurations selected in second and third optimization cycles with target function scores.

Configuration	**Back**	**Diameter**	**Front**	**Radial**	$TF_{\textbf{Predicted}}$
Second Optimization Cycle					
19	0.99434	0.01959	0.74123	0.99183	0.74377
20	0.06097	0.56079	0.61066	0.67648	0.76000
21	0.99434	0.47697	0.62423	0.69565	0.75298
Third Optimization Cycle					
22	0.99434	0.339989	0.61663	0.991838	0.762797
23	0.06097	0.01959	0.89739	0.99183	0.714914
24	0.99434	0.985948	0.89739	0.10701	0.696063

For the second optimization, Equation 3.5 was used for the normalization of axial force. The corresponding performance results for the selected configurations are presented in Figures 4.23 to 4.38, shown in green. Compared to the first cycle, the hydraulic performance remains within a similar range, with only minor deviations from the reference impeller. However, axial force reduction improved, with at least one configuration in this cycle outperforming those from the first cycle in this metric.

Despite these improvements, a key limitation became evident at the overload condition ($Q^* = 1.25$). As seen from Figure 4.38, the reduction in axial force across increasing flow rates was relatively minimal. More critically, the trend of axial force decreasing with increasing flow rate—typically expected in centrifugal pumps [10]. Instead, the reduction becomes less pronounced or even reverses, contradicting the fundamental operating behavior of centrifugal pumps.

This issue is further illustrated in Figure 4.17, which compares normalized axial force for configuration 16 (from the first optimization cycle) and configuration 20 (from the second optimization cycle), alongside the reference impeller. It is clear that for these optimized configurations, the axial force at $Q^* = 1.25$ slightly increases compared to lower flow rates, and in some cases approaches or even exceeds the reference value. This unexpected shift indicates that the current target function, which excluded this flow condition, failed to capture worst-case axial loading effectively.

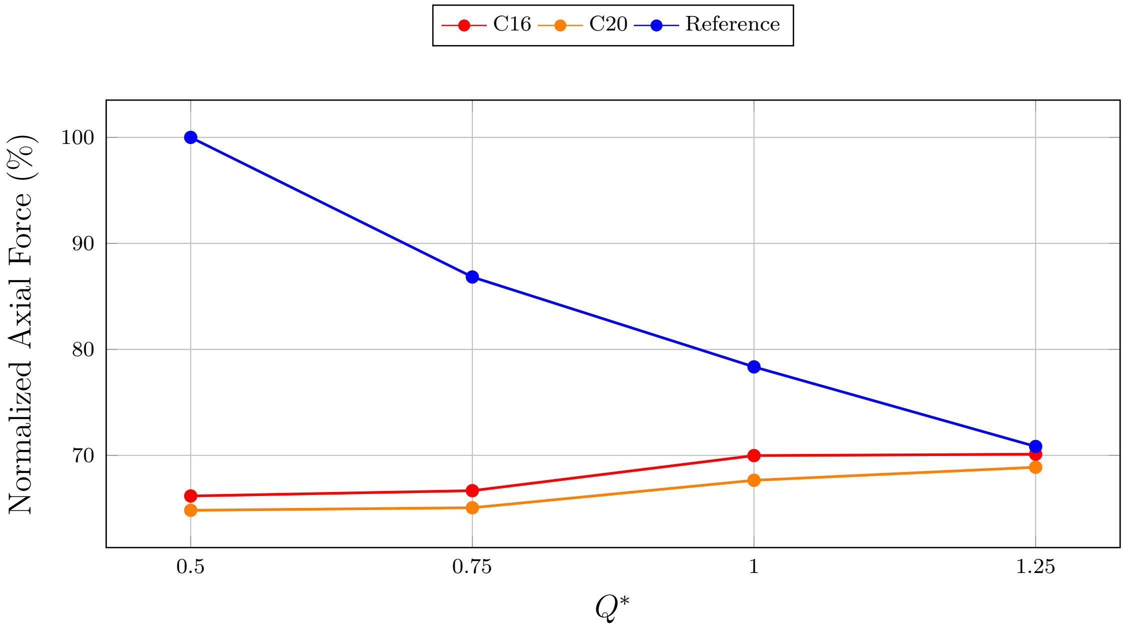


Figure 4.17: Axial force comparison across different Q^* for first and second optimization cycle.

As a result, the third optimization cycle explicitly incorporated the $Q^* = 1.25$ operating condition's axial force values into the Equation 3.5, leading to a modified formulation shown in equation 4.1, to ensure broad applicability. The performance results shown in Figures 4.23 to 4.38, highlighted in orange, indicate that configuration 24 achieved a noticeable reduction in axial force while consistently maintaining hydraulic performance across all operating conditions, outperforming all configurations generated from the optimization cycles.

$$\text{Normalized Axial Force} = \frac{F_{\text{ref, }0.5} - \max(F_{0.5}, F_{0.5}, F_{1.25})}{|F_{\text{ref, }0.5}|} \tag{4.1}$$

Compared to the reference impeller, all optimized configurations from the third optimization cycle (refer to Figure 4.18) exhibit a clear reduction in axial force across the entire operating range, without compromising hydraulic performance. It is known that at $Q^* = 0.5$ and 0.75, where axial force is severe due to the uneven pressure distribution, configuration 24 achieves the most pronounced reductions. This behavior highlights the effectiveness of the balancing hole design in reducing pressure buildup behind the impeller.

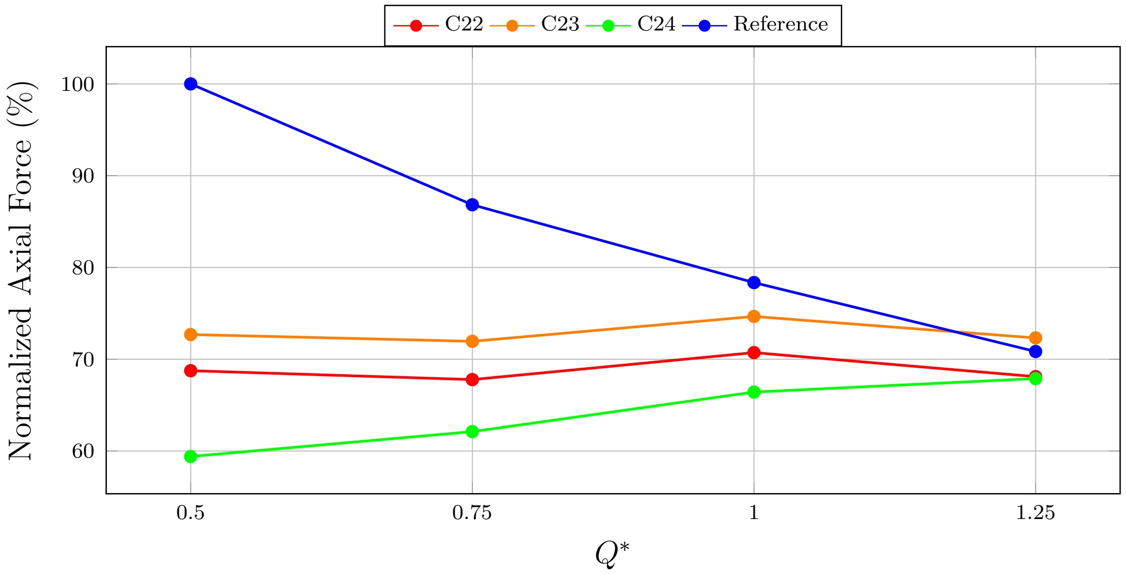


Figure 4.18: Axial force comparison across different Q^* for third optimization cycle.

Notably, the previously observed issue of insufficient axial force reduction at $Q^* = 1.25$ has been partially addressed. While configuration 24 shows a steady decrease across all flow rates, configuration 23 maintains nearly equal axial force between $Q^* = 1.0$ and 1.25, and configuration 22 exhibits a slight increase. This indicates that although the updated formulation improves performance in the overload regime, some sensitivity still remains in specific designs.

Among those, configuration 24 achieves the lowest axial force and maintains relatively stable hydraulic performance across all operating conditions. Interestingly, the trend indicates the potential for a design configuration that maintains uniform axial force across all operating conditions—a straight-line trend—which would represent balanced performance.

These results highlight the importance of refining the target function to align with specific optimization priorities across the full operating range. They also confirm the critical role of balancing holes in reducing axial force without compromising hydraulic stability—an essential outcome for improving pump longevity and operational robustness.

To further validate this, Figures 4.19 and 4.20 compare the normalized performance curves of configurations 20 and 24. As seen, there is no significant difference in hydraulic performance between the two. Both configurations exhibit closely matching trends in head, power, and efficiency, with only minimal deviation from the reference case. This reinforces that the observed axial force reduction in configuration 24 is achieved without sacrificing hydraulic performance, confirming the effectiveness of target function based design optimization process.

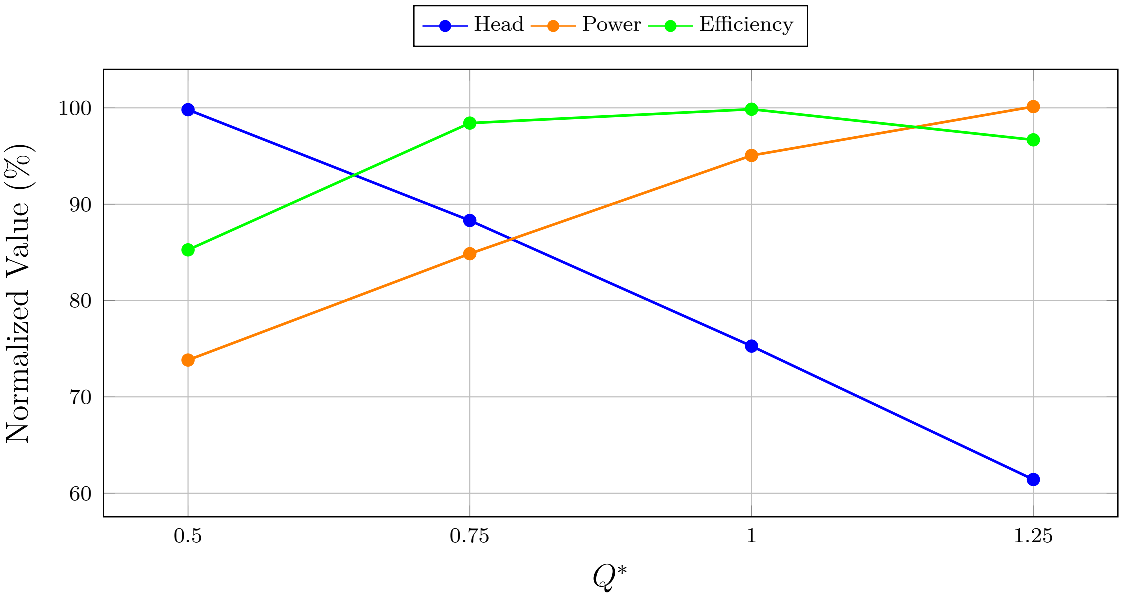


Figure 4.19: Performance curve of configuration 20.

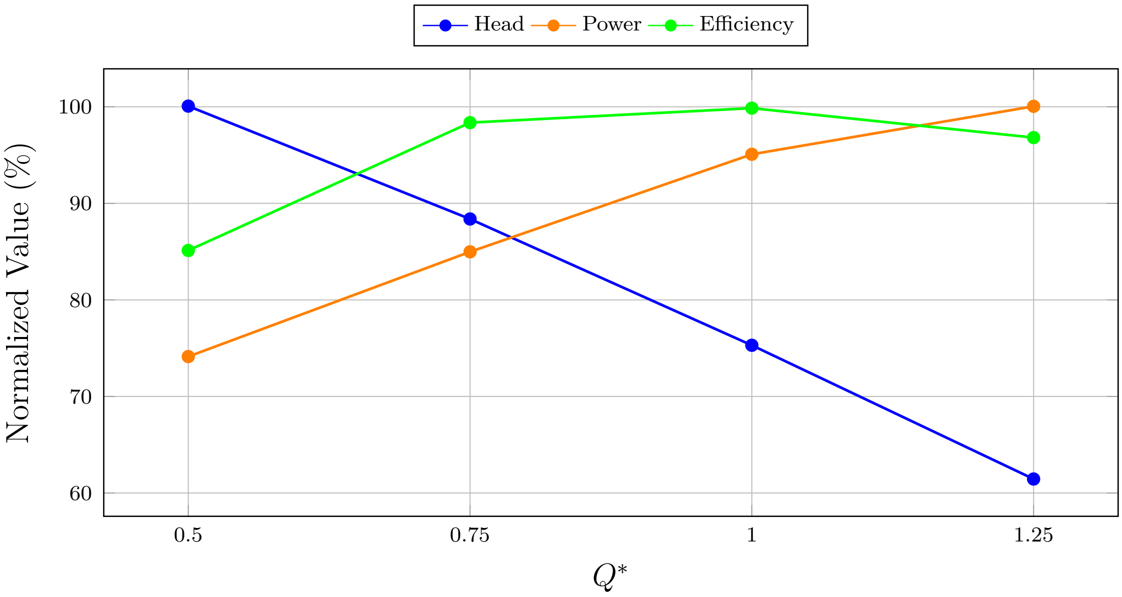


Figure 4.20: Performance curve of configuration 24.

4.7 Pareto Front Analysis

The final stage of the optimization involved a Pareto front analysis to evaluate trade-offs between axial force and hydraulic performance (combined head and power). Rather than relying on a scalarized target function, this approach aimed to identify non-dominated solutions where improvements in one objective could not be achieved without compromising the other.

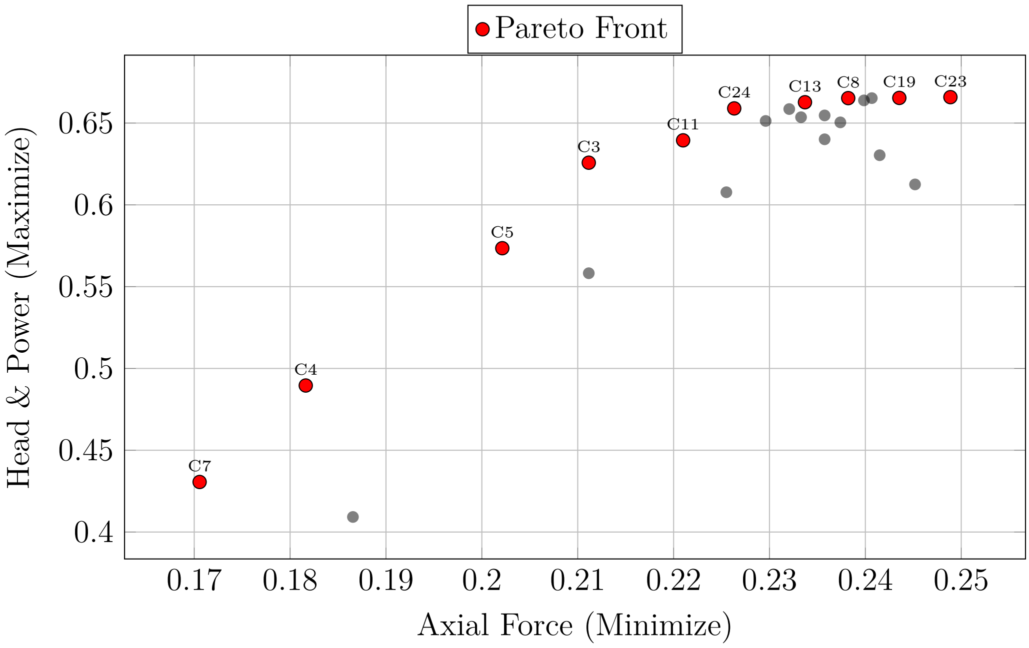


Figure 4.21: Pareto front analysis.

Figure 4.21 illustrates the resulting Pareto front for all configurations that have been studied till now. The red-highlighted points represent the configurations lying on the Pareto front (non-dominated configurations in the objective space), indicating the best trade-offs between axial force and hydraulic performance between each objective considered in this study. As expected, it reflects the underlying design trade-offs between each objective.

Notably, configurations 23 and 19 exhibit strong hydraulic performance but relatively higher axial force, indicating that moving right towards the objective space might not significantly reduce axial force, whereas configurations 7 and 4 represent a completely opposite scenario, indicating that moving left towards the objective space might hinder the improvement of hydraulic performance further. The most notable feature in figure 4.21 is the horizontal segment of the Pareto front involving configurations 24, 13, 8, 19, and 23. This section demonstrates a well-balanced trade-off, combining substantial axial force reduction with near-peak performance in both head and power.

This highlights the significance of Pareto front analysis as both an evaluation and decision-making tool. By identifying a set of non-dominated configurations, the method provides a clear view of the best trade-offs achieved between conflicting objectives. These results reinforce the robustness of the optimization framework and offer practical guidance for selecting balanced designs based on specific operational goals.

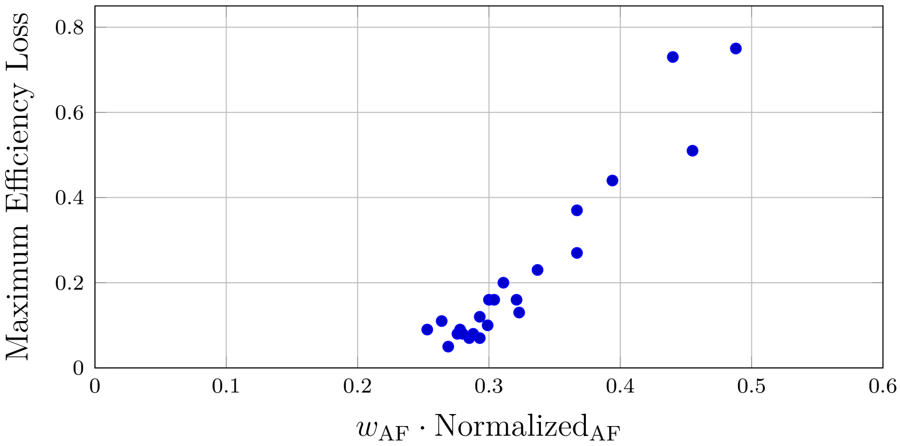


Figure 4.22: Trade-off between axial force and maximum efficiency loss.

To further illustrate this trade-off quantitatively, Figure 4.22 presents a scatter plot comparing the axial force contribution in the scalarized target function and the maximum efficiency loss across all configurations. The x-axis represents the weighted axial force term $w_{\text{AF}} \cdot \text{Normalized}_{\text{AF}}$, directly taken from the scalarized target function formulation shown in Equation 3.12, and it represents that a higher value means a higher reduction in axial force. The y-axis shows the maximum efficiency loss, computed using the following expression:

$$(\Delta\eta)_{\max} = \max_{Q^*=0.5,\,0.75,\,1.0,\,1.25} (\eta_{\text{Ref}} - \eta_i)$$

where η_i is the efficiency for the i^{ith} impeller configuration with balancing hole and η_{Ref} is the efficiency for the reference impeller at the same operating condition. The trend clearly from the above figure clearly indicate that configurations with a stronger axial force reduction tend to suffer greater efficiency loss. It highlights the challenge associated with balancing conflicting objectives. This analysis complements the earlier Pareto front analysis by quantitatively highlighting the performance boundaries and visually confirming the trade-offs involved in achieving balanced solutions.

4.8 Performance Results of Impeller Configurations

The following figures present the detailed performance trends—head, power, efficiency, and axial force—across all generated impeller configurations during this study for all operating conditions. While each configuration shows minimal deviation from the reference case in terms of hydraulic performance, this visualization ensures traceability of how specific configurations evolved through the optimization cycles.

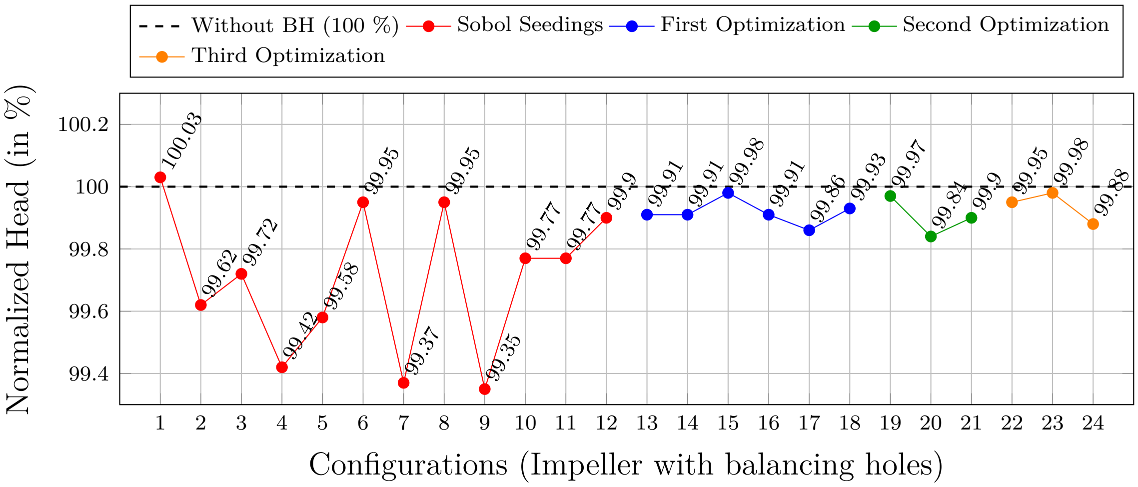


Figure 4.23: Normalized head for $Q^* = 1.0$.

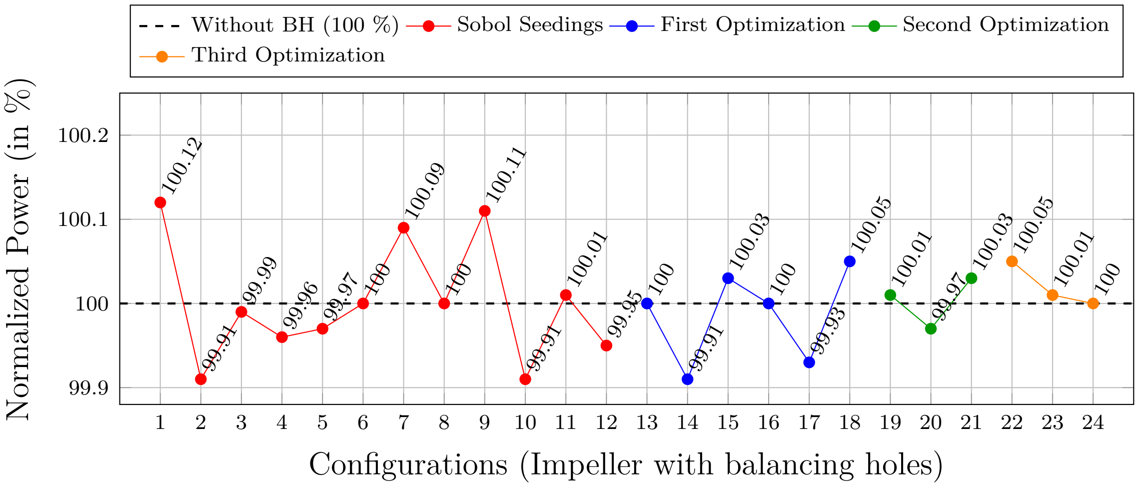


Figure 4.24: Normalized power for $Q^* = 1.0$.

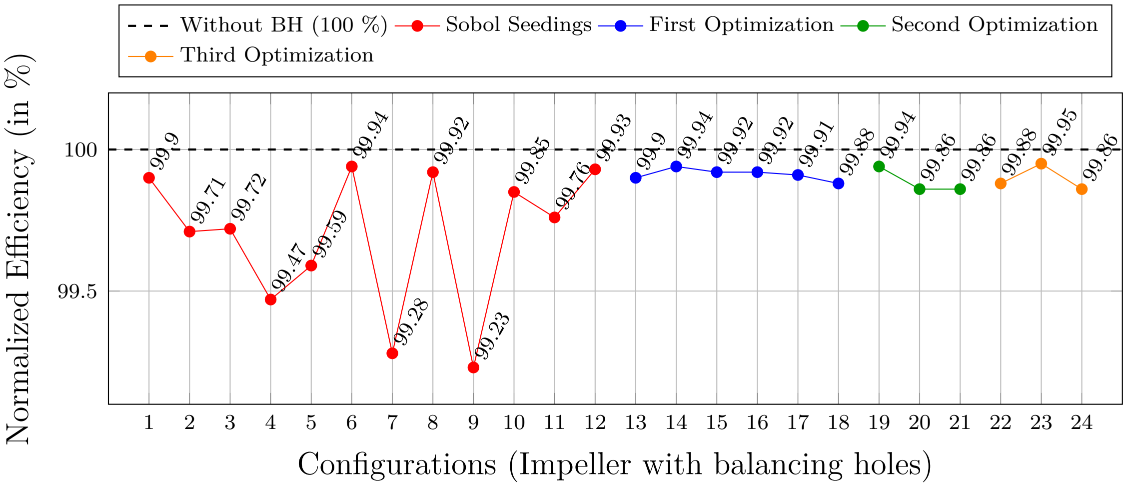


Figure 4.25: Normalized efficiency for $Q^* = 1.0$.

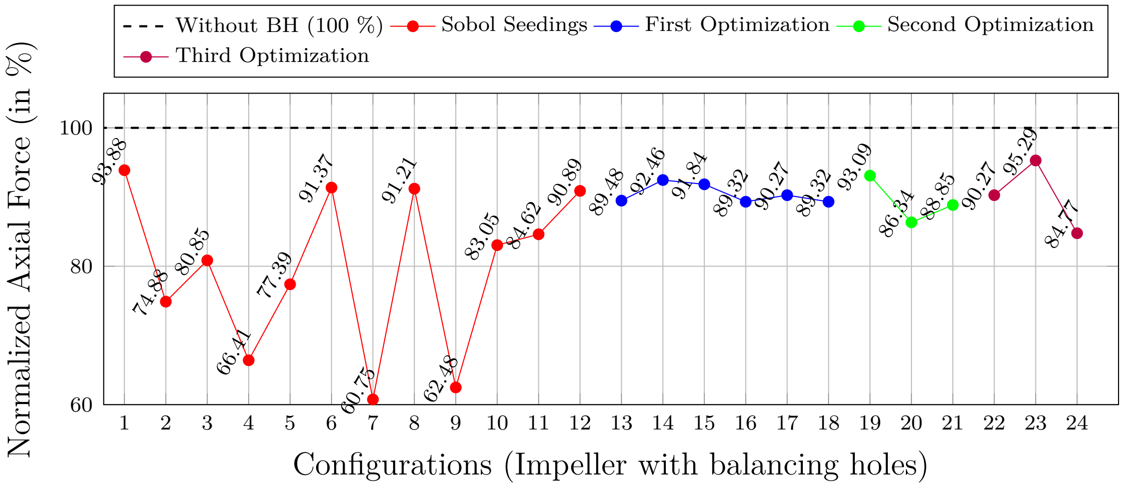


Figure 4.26: Normalized axial force for $Q^* = 1.0$.

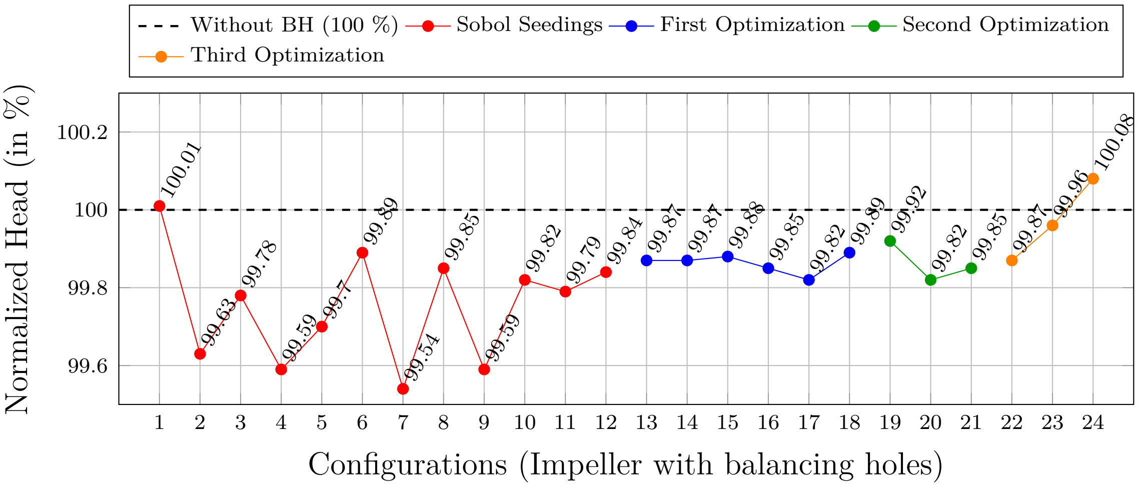


Figure 4.27: Normalized head for $Q^* = 0.5$.

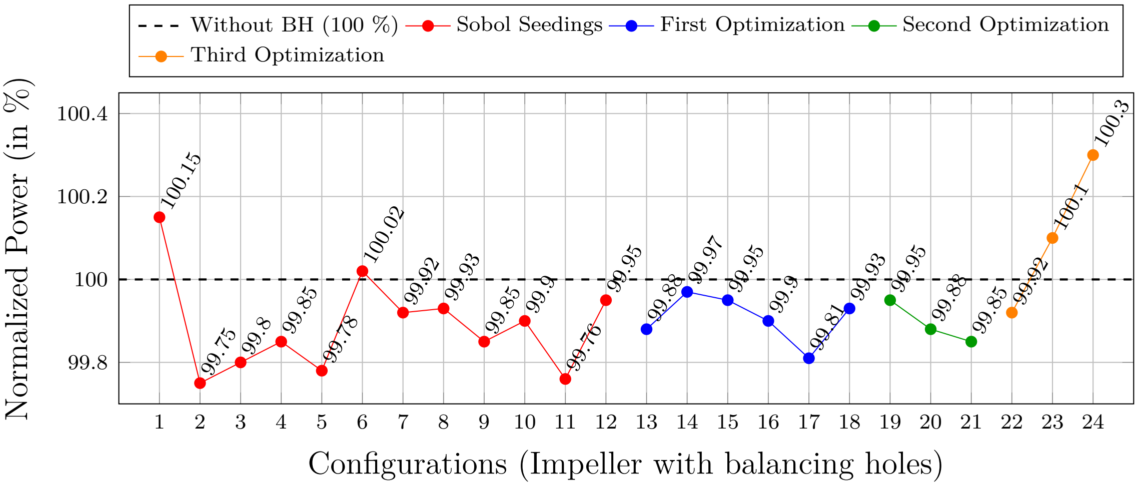


Figure 4.28: Normalized power for $Q^* = 0.5$.

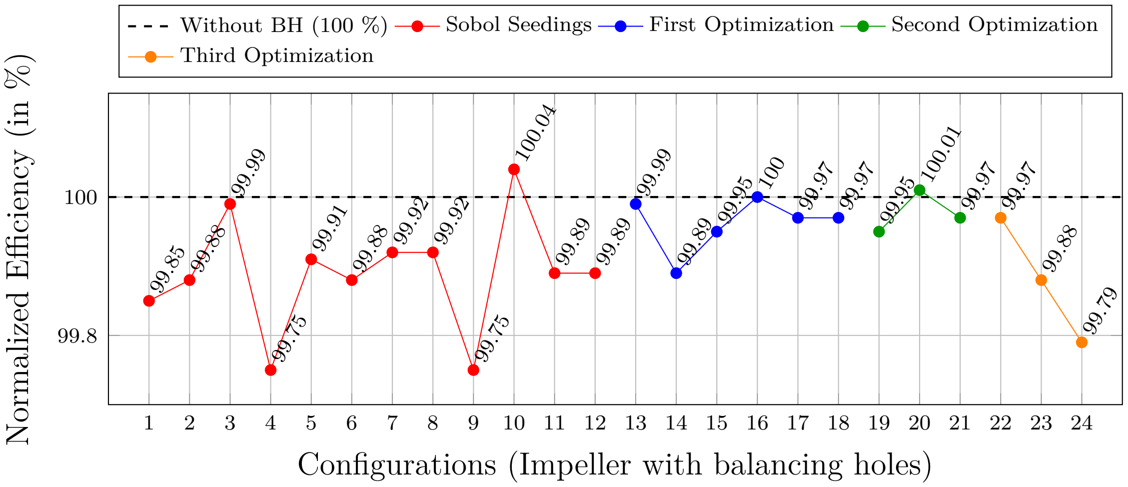


Figure 4.29: Normalized efficiency for $Q^* = 0.5$.

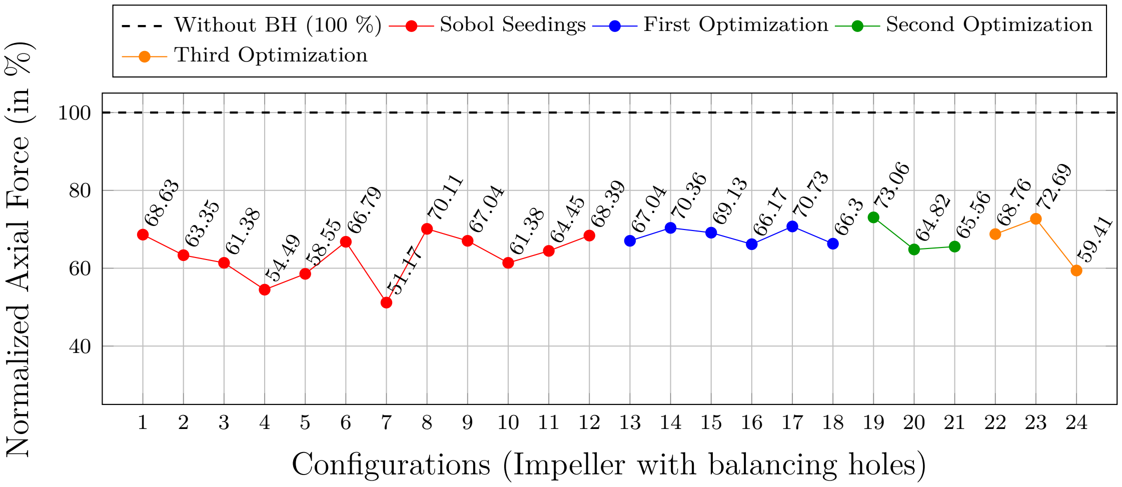


Figure 4.30: Normalized axial force for $Q^* = 0.5$.

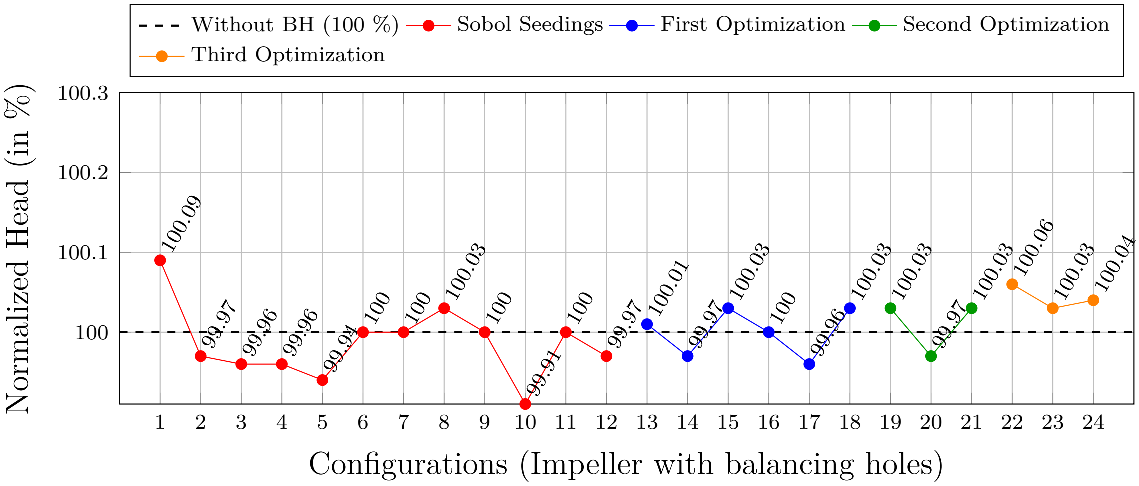


Figure 4.31: Normalized head for $Q^* = 0.75$.

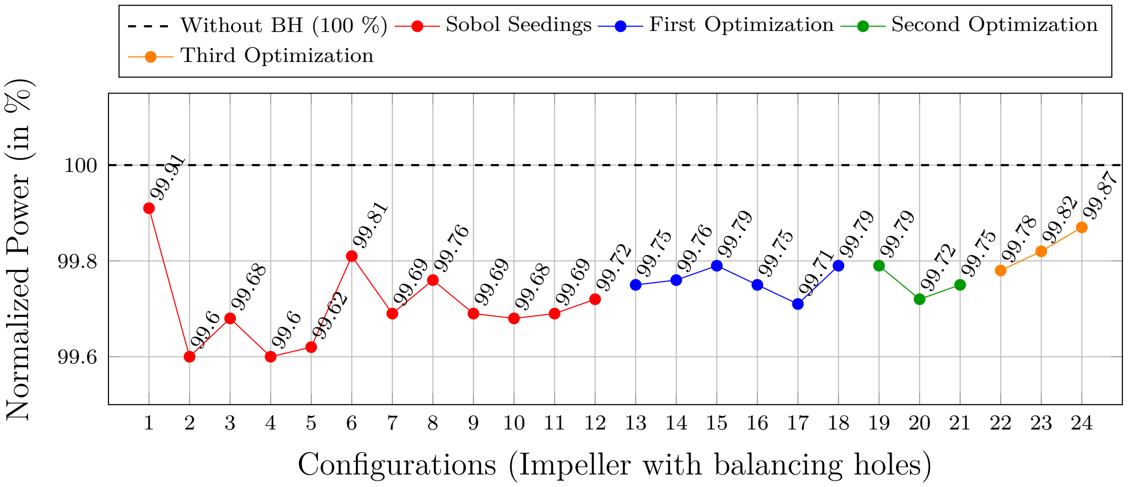


Figure 4.32: Normalized power for $Q^* = 0.75$.

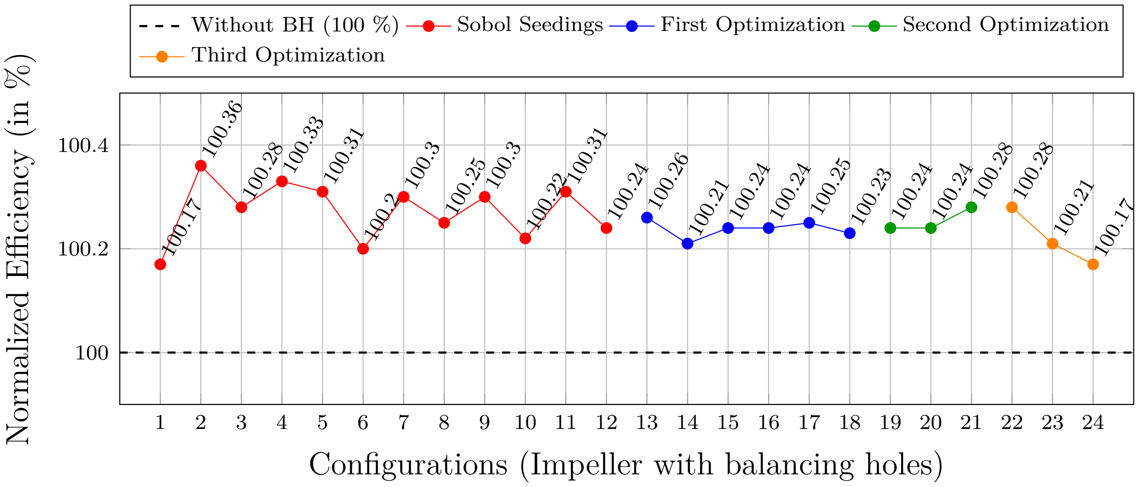


Figure 4.33: Normalized efficiency for $Q^* = 0.75$.

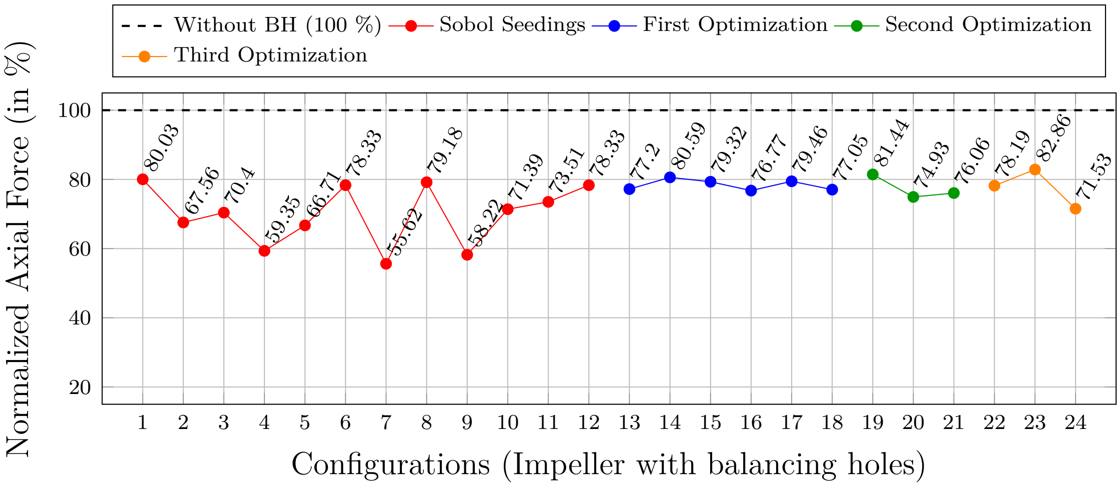


Figure 4.34: Normalized axial force for $Q^* = 0.75$.

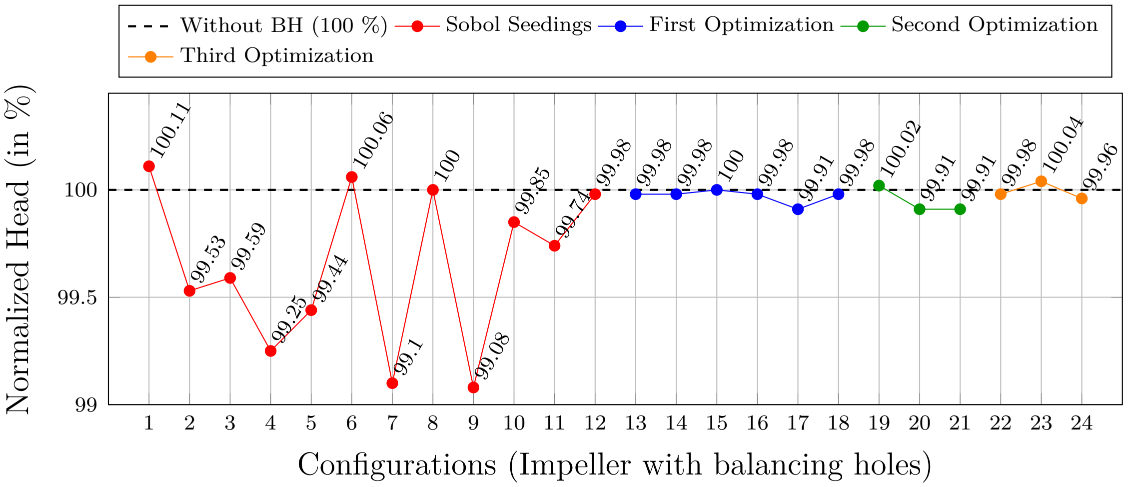


Figure 4.35: Normalized head for $Q^* = 1.25$.

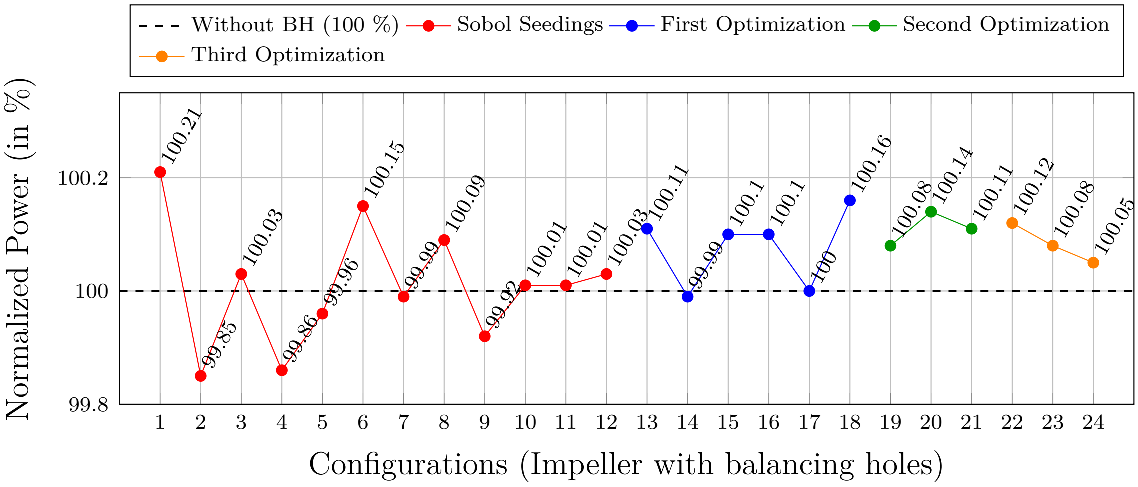


Figure 4.36: Normalized power for $Q^* = 1.25$.

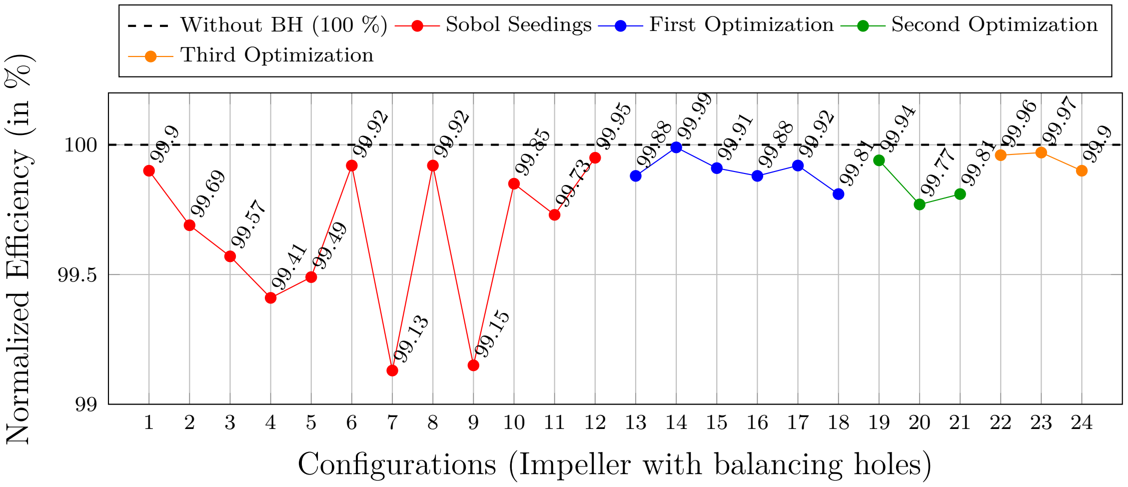


Figure 4.37: Normalized efficiency for $Q^* = 1.25$.

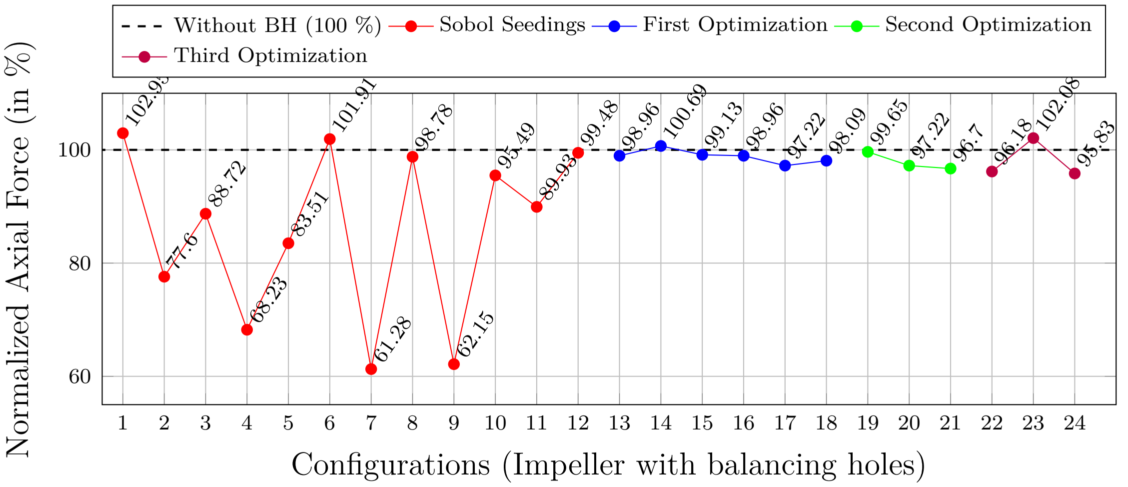


Figure 4.38: Normalized axial force for $Q^* = 1.25$.

4.9 Influence of Design Parameters during Optimization Process

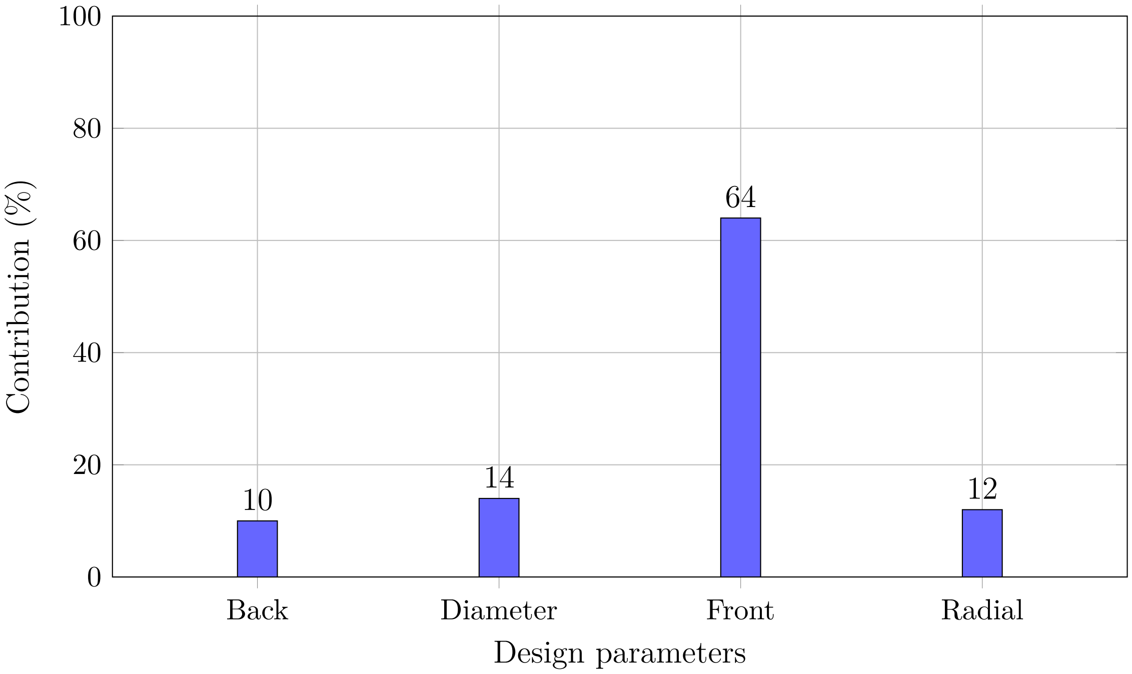


Figure 4.39: Design parameter contribution.

Figure 4.39 shows the relative influence of each design parameter on the scalarized target function. **BH__Circumferential__Front** was the most dominant, contributing nearly 60 % to model prediction, whereas the other design parameters had a noticeably lower impact, suggesting that the axial force and hydraulic performance trade-off is significantly controlled by the front side position of the balancing hole.

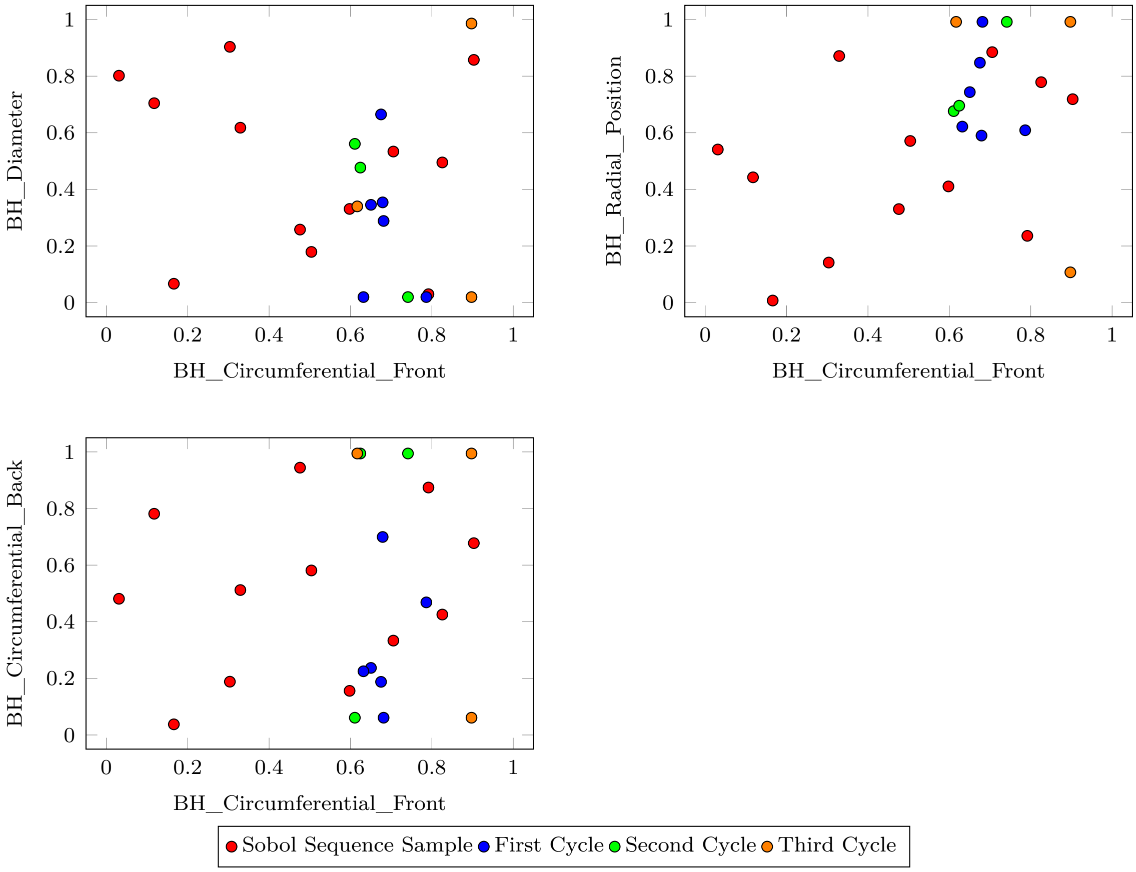


Figure 4.40: Pairwise scatter plot of generated configurations.

To explore this dominance further, Figure 4.40 presents pairwise scatter plots showing how the selected configurations from each optimization cycle are distributed across the design space. Configurations generated throughout the optimization cycle begin clustering towards specific regions in the design space. In particular, when compared against **BH_Circumferential_Front**, the distributions of the other parameters—especially in the third optimization cycle—show a clear tendency to concentrate near the lower or upper bounds of the design space. This indicates that the scalarized form of the target function, along with Bayesian optimization, concentrated on regions offering better performance. Additionally, this narrowing of points towards a certain region in the design space hints at a potential correlation between **BH_Circumferential_Front** and the other design parameters, which may be indirectly influencing the optimization path.

4.10 Comparison of predicted and scalarized values of Target Function

Figure 4.41 compares the target function value obtained using the scalarized form of the target function ($TF_{\text{Scalarized}}$) and from the predicted Bayesian optimization ($TF_{\text{Predicted}}$). It can be seen from the same figure that the predicted and scalarized values align closely, confirming that the surrogate model was reasonably accurate during the second optimization cycle. However, starting from configuration 22 onward (third optimization cycle), a significant deviation is observed—particularly for configurations 23 and 24. While the scalarized target function assigns a higher score to configuration 24, consistent with its superior multi-objective performance (as seen in Figures 4.23 to 4.38), the predicted values from Bayesian optimization exhibit the opposite trend, incorrectly ranking configuration 24 lower than 23.

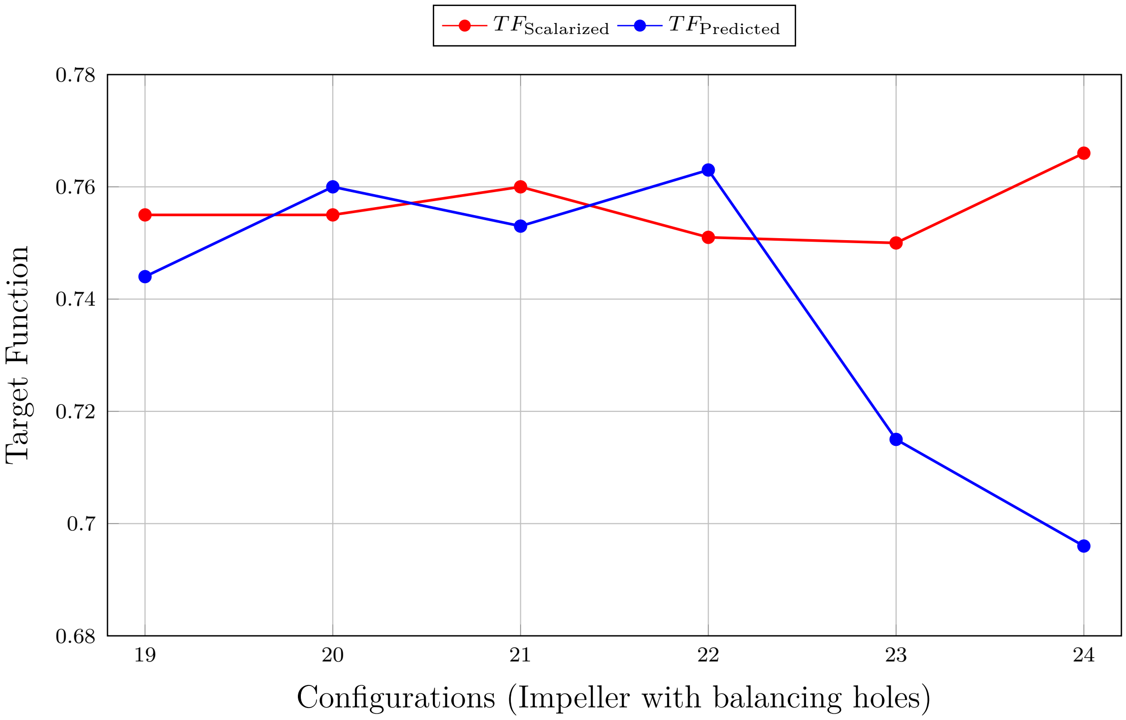


Figure 4.41: Comparison of target function values.

This difference may be attributed to the design parameters explored during the third optimization cycle becoming increasingly concentrated within a localized region within the design space, as discussed in Section 4.9. Since the surrogate model prediction relies on the Gaussian process trained on earlier samples, its accuracy might have degraded when extrapolating to sparsely sampled regions of the design space, such as those encountered during the third optimization cycle. In contrast, during the second optimization cycle, the surrogate model predictions matched more closely with the scalarized values because the samples were more densely clustered in regions where prior data was already abundant, resulting in higher model reliability. Therefore, the deviation between the scalarized and predicted values is likely due to the reduced model reliability in the highly localized region of the exploration.

4.11 Flow Visualization and Interpretation of Balancing Hole Effects

To investigate internal flow behaviour and the pressure equalization role of different balancing hole configurations, a comprehensive post-processing was carried out using

ParaView. Selected configurations, configurations 7 and 24, were analyzed based on their contrasting performance in axial force and hydraulic characteristics. These analyses, at flow rate $Q^* = 1.0$, included surface pressure distributions, streamline visualizations, and flow angle variation to understand jet-mainstream interaction. The reference impeller served as the baseline for comparison and mass flow rate discussed in this section will be normalized with the flow condition.

4.11.1 Pressure Distribution and Mass Flow Rate Analysis through Balancing Holes

4.11.1.1 BEP Condition — $Q^* = 1.0$

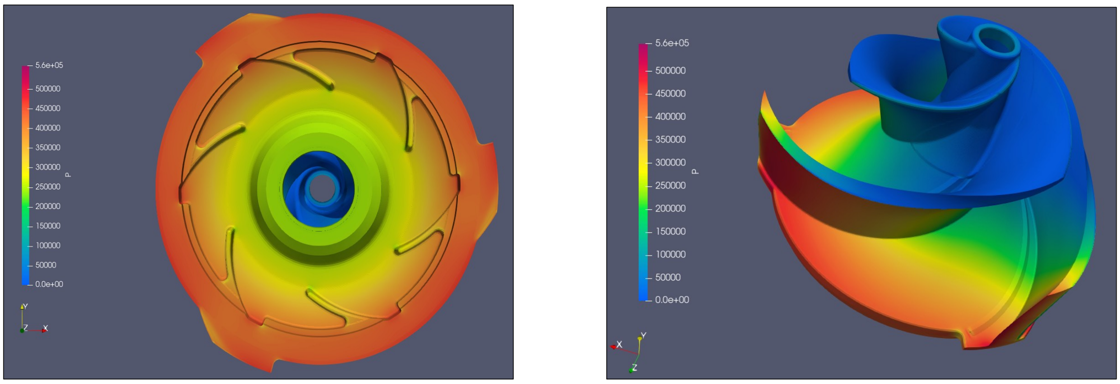
Surface pressure distribution on the front and back sides of the impeller are shown in Figures 4.42 to 4.44 for the chosen configurations along with the reference case. In the reference case (refer to Figure 4.42), a clear pressure buildup is observed on the back side of the impeller compared to the front side of the impeller.

In contrast, the addition of balancing holes on the impeller introduces a visible reduction in pressure on the back side of the impeller (Figures 4.43 to 4.44). Although both configurations (Configurations 7 and 24) reduced pressure on the back side of the impeller, Configuration 7 shows a slightly greater degree of relief compared to configuration 24, and this is not distinguishable from the surface contour alone. However, Table 4.5 reveals a substantial difference in axial force reduction between the two.

Table 4.5: Performance results of the impeller with and without balancing holes at BEP ($Q^* = 1.0$) condition.

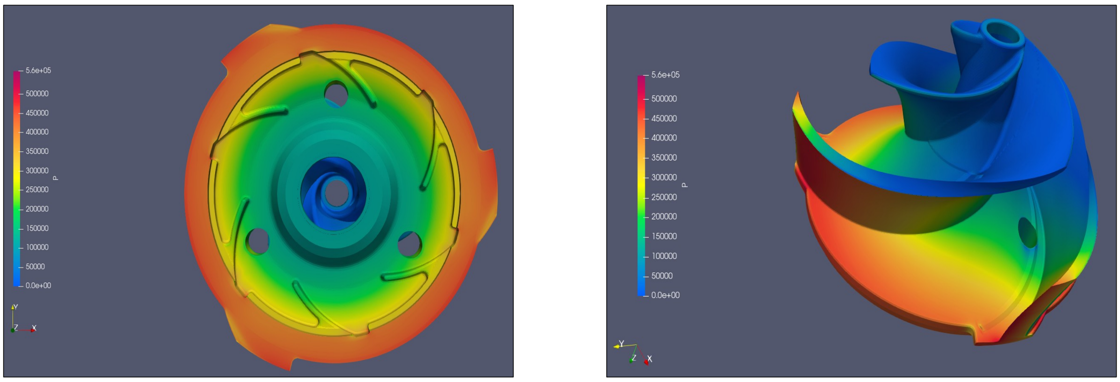
Normalized Parameter	Reference	Configuration 7	Configuration 24
Head (%)	100	99.37	99.88
Power (%)	100	100.09	100.0
Efficiency (%)	100	99.28	99.86
Axial Force (%)	100	60.75	84.77

Mid-plane pressure contours (Figure 4.45) provide deeper insight into the flow physics behind axial force mitigation. Configuration 7 (Figure 4.45b) shows pronounced pressure relief across the impeller back side, indicating effective flow through the balancing hole. This is supported by the high normalized mass flow rate of 8.275×10^{-3}, which reflects a strong pressure gradient between the impeller disc region and the hole outlet into the rear channel. The enhanced flow through the hole promotes pressure equalization across the disc surfaces, thereby significantly reducing axial force. In contrast, configuration 24 (Figure 4.45c) exhibits a smaller pressure drop and a lower normalized mass flow rate of 3.34×10^{-3}, which leads to moderate axial force relief. However, the reduced leakage through the balancing hole contributes to better hydraulic efficiency due to more stable pressure conditions in the impeller and lower energy dissipation.



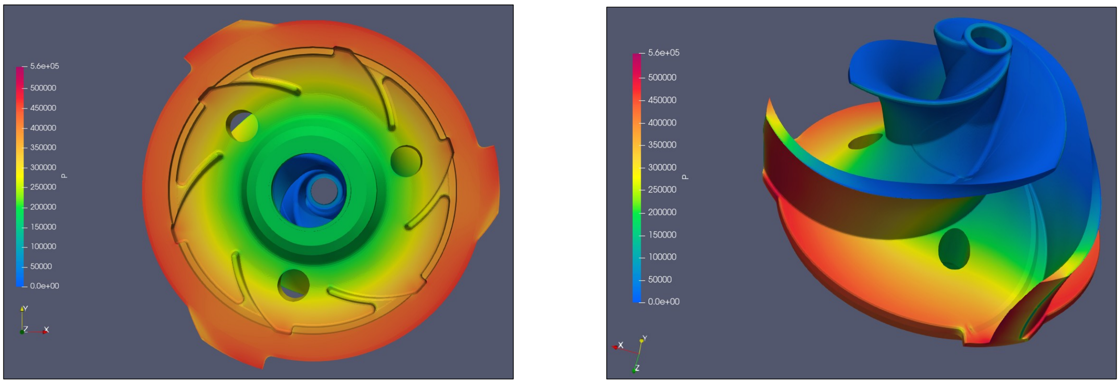
(a) Pressure Side. **(b)** Suction Side.

Figure 4.42: Pressure contour on the surface of the reference impeller at BEP ($Q^* = 1.0$) condition.



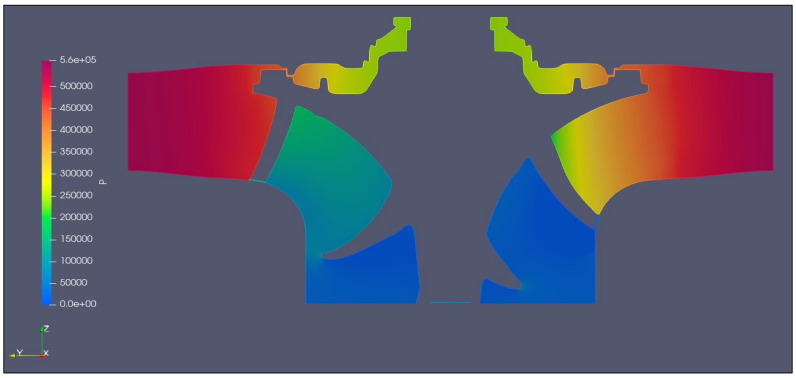
(a) Pressure side. **(b)** Suction side.

Figure 4.43: Pressure contour on the surface of the impeller with balancing hole (Configuration - 7) at BEP ($Q^* = 1.0$) condition.

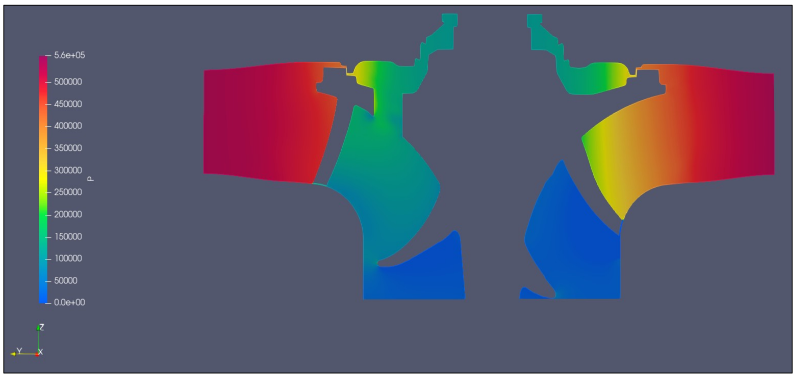


(a) Pressure side. **(b)** Suction side.

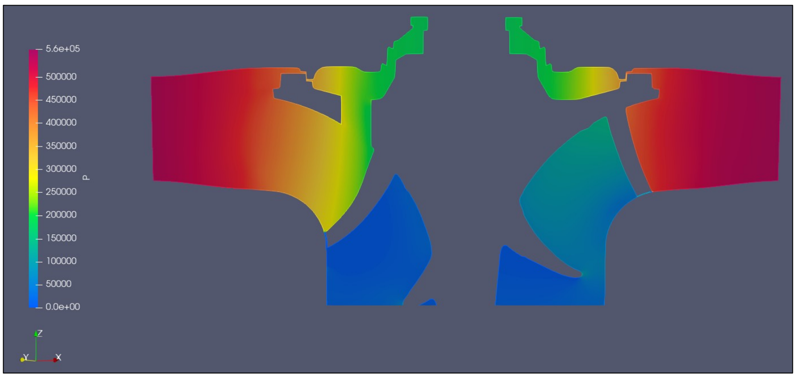
Figure 4.44: Pressure contour on the surface of the impeller with balancing hole (Configuration - 24) at BEP ($Q^* = 1.0$) condition.



(a) Impeller without balancing hole.



(b) Impeller with balancing hole (Configuration - 7).



(c) Impeller with balancing hole (Configuration - 24).

Figure 4.45: Pressure contour along the mid-plane of the flow domain, with emphasis on the impeller region at BEP ($Q^* = 1.0$) condition.

4.11.2 Jet Interaction and Flow Redistribution through Balancing Hole

To understand the influence of the balancing hole on internal flow behaviour, especially the interaction between the jet emerging from the balancing hole and the main stream, streamline visualizations coloured by the relative flow angle were analyzed for both configurations and the reference impeller under BEP condition.

4.11.2.1 BEP Condition — $Q^* = 1.0$

At BEP condition, the streamline visualizations scaled by flow angle reveal that the jet emerging from the balancing hole does not significantly disrupt the main flow in either configuration (refer to Figures 4.47 to 4.48). This is due to the high incoming velocity at BEP condition that facilitates the smooth integration of the jet from the balancing hole into the main flow passage.

In Configuration 7, the holes are located near the suction side of the impeller blades, where the local velocity is higher, and are also positioned upstream facing the incoming flow. As a result, the jet merges seamlessly with the main flow without causing significant disturbances.

Due to the positioning of the holes in configuration 7, as shown in Figure 4.47, there is a slight increase in the relative flow angle near the suction side of the impeller blade compared to the reference case (refer to Figure 4.46). This localized change in flow angle may contribute to a marginal reduction in head, as summarized in Table 4.5. Furthermore, the upstream placement of the hole, along with its position on the suction side of the blade, allows for more effective pressure relief due to the strong pressure gradients between the front and back sides of the impeller. As a result, configuration 7 achieves better axial force reduction (39.25 %).

In configuration 24 (refer to Figure 4.48), although the jet also merges smoothly with the main flow, the flow angle remains higher than that of the reference impeller. This may explain the potential drop in head as indicated in Table 4.5. However, the downstream positioning of the holes closer to the pressure side of the impeller blades results in a smaller pressure difference between the rear and front sides, limiting the pressure equalization capacity, yielding in a less effective reduction in axial force (15.23 %). Overall, the minor reduction in head across both configurations remains negligible, indicating that the jet emerging from the balancing hole has minimal influence on hydraulic performance while still contributing to significant axial force reduction.

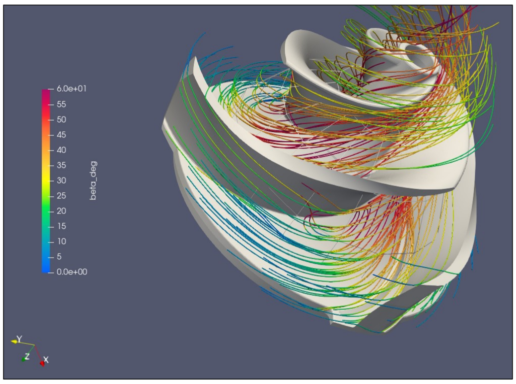
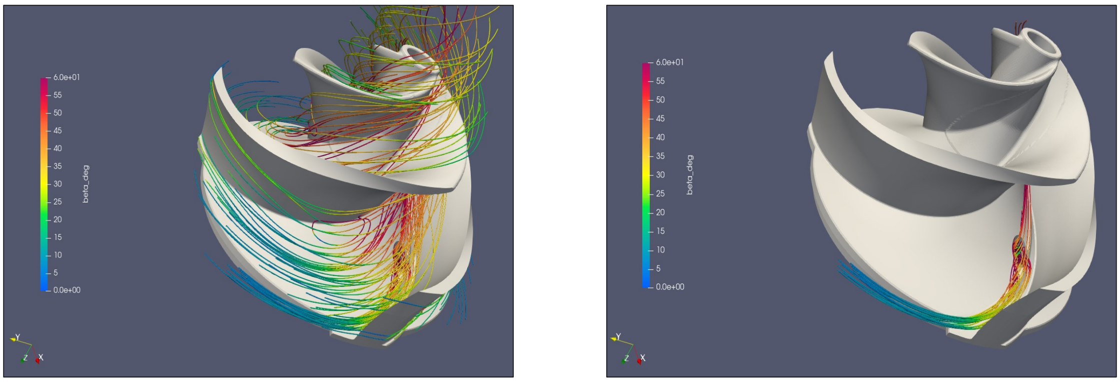


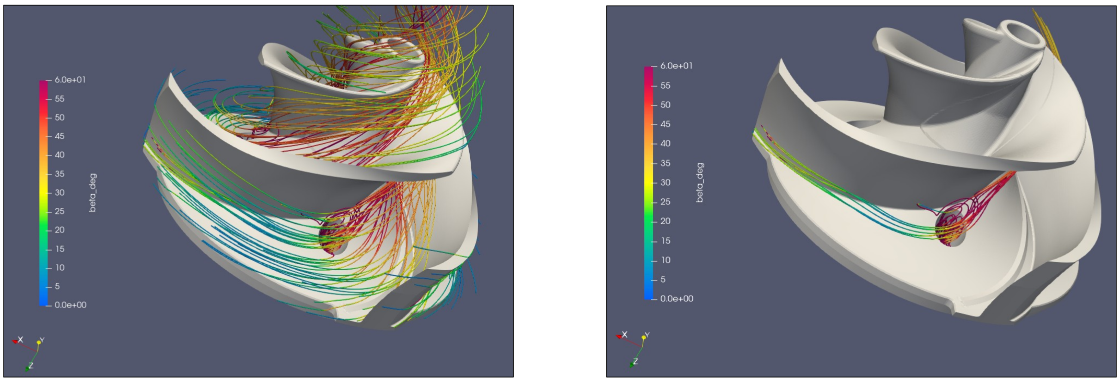
Figure 4.46: Streamline visualization coloured by relative flow angle (β) for the reference Impeller at BEP ($Q^* = 1.0$) condition.



(a) Without highlighting balancing hole jet interaction.

(b) With focus on jet interaction near the balancing hole.

Figure 4.47: Streamline visualization coloured by relative flow angle (β) for configuration 7 at BEP ($Q^* = 1.0$) condition.



(a) Without highlighting balancing hole jet interaction.

(b) With focus on jet interaction near the balancing hole.

Figure 4.48: Streamline visualization coloured by relative flow angle (β) for configuration 24 at BEP ($Q^* = 1.0$) condition.

From this analysis, it is evident that the circumferential position of the balancing holes on the front side of the impeller plays a dominant role in axial force reduction. These findings underscore that, even when both configurations have approximately the same diameter and similar radial placements, differences in circumferential positioning resulted in significantly varied performance. This reinforces the strategic placement for achieving an optimal balance between hydraulic efficiency and axial force reduction in centrifugal pumps.

5

Conclusion

This thesis focused on developing a robust and reliable multi-objective target function to guide the optimization process for designing balancing holes in the Flygt N-centrifugal pump, using CFD to numerically evaluate the hydraulic performance of the generated configurations throughout this study. The target function was formulated to balance the trade-offs between maximum possible axial force reduction while maintaining the overall hydraulic performance of the pump. For this reason the performance metrics, such as head and power, including axial forces, were normalized based on various techniques. The multi-objective target function was formulated using two different methods: the geometric mean method and the scalarized form with the inverse deviation weightage method.

The findings revealed that the target function formulated with the scalarized approach using the inverse deviation weightage method worked well as a scoring tool in the Bayesian optimization process, consistently guiding the search for optimal configurations that reduced axial force by up to 40.59% while maintaining head and power within 99.98% and 99.87% of the reference case, respectively. A clear trade-off was observed between axial force and efficiency loss, where configurations with significant reduction in axial force typically showed higher efficiency drop, highlighting the complexity of the multi-objective optimization process.

Additionally, the step-by-step Bayesian optimization, along with Pareto front analysis, showed that the position around the impeller's suction side plays a major role, accounting for almost 60% of the optimization process. This was further supported by pressure distribution and streamline visualizations, which showed that the jet emerging from the balancing hole on the suction side played a key role in altering the local pressure field and reducing axial force.

Overall, this study provides a systematic approach to balancing hole optimization using CFD and multi-objective scoring strategies, offering valuable insights for future design improvements in centrifugal pumps. The combined use of scalarized optimization, Pareto front analysis, and flow physics analysis proved essential in optimizing the design configuration of a balancing hole in a centrifugal pump that offers a balance between hydraulic performance and axial force reduction.

5.1 Future Work

Building upon the findings of this study, several potential research directions are proposed to further enhance the design and evaluation of balancing holes:

- **Inclined and Filleted Configurations:** Investigate the influence of inclined orientations and filleted geometries for balancing holes to improve flow guidance and reduce local pressure peaks that may lead to losses or instabilities.
- **Transient Flow Simulations:** Extend the current steady-state analysis to transient simulations in order to capture unsteady phenomena such as flow oscillations, vortex shedding, or start-up behavior, particularly in the optimized configurations.
- **Cavitation Risk Assessment:** Analyze the susceptibility of the optimized balancing hole configurations to cavitation, considering pressure drop zones and vapor formation, especially near the impeller hub and rear side where balancing holes are positioned.
- **Application to Other Pump Designs:** Apply the proposed design and optimization methodology to pumps with different specific speeds to evaluate its generalizability and effectiveness across varying operating conditions and geometrical scales.

Bibliography

[1] Nourbakhsh, A., Jaumotte, A., Hirsch, C., and Parizi, H. B., *Turbopumps and Pumping Systems*, Springer Science & Business Media, 2007.

[2] Shepherd, D. G. (1987). *Principles of Turbomachinery.* John Wiley & Sons, Inc.

[3] Gülich, J. F. (2019). Centrifugal Pumps (4th ed.). Springer International Publishing.

[4] O. Babayigit, M. Ozgoren, M.H. Aksoy, and O. Kocaaslan, "The effect of balance holes to centrifugal pump performance," AIP Conference Proceedings, vol. 1863, no. 1, pp. 030004–030004-4, 2017. doi:10.1063/1.4992157.

[5] M. Fathi, M. Raisee, S. A. Nourbakhsh, and H. A. Arani, "The effect of balancing holes on performance of a centrifugal pump: numerical and experimental investigations," IOP Conf. Ser.: Earth Environ. Sci., vol. 240, no. 3, p. 032017, 2019, doi:10.1088/1755-1315/240/3/032017.

[6] S. Sefaci, O. Babayigit, and S. Koçak, "Experimental design methods in a centrifugal pump impeller balancing hole design optimization," *Bitlis Eren University Journal of Science*, vol. 11, no. 2, pp. 436–448, 2022, doi:10.17798/bitlisfen.1017804.

[7] Loewenstein, L. C., and Crissey, C. P. *Centrifugal Pumps: Their Design and Construction.* McGraw-Hill Book Company, 1911.

[8] Yedidiah, S. (1996). Centrifugal Pump User's Guidebook: Problems and Solutions. Chapman & Hall.

[9] Dixon, S. L., and Hall, C. A. *Fluid Mechanics and Thermodynamics of Turbomachinery.* 7th ed., Butterworth-Heinemann, 2014.

[10] Pehlivan, H. and Parlak, Z. (2019). Investigation of parameters affecting axial load in an end suction centrifugal pump by numerical analysis. *Journal of Applied Fluid Mechanics*, 12(5), pp.1615–1627. doi:10.29252/jafm.12.05.29623

[11] Hong, S. and Kang, S. (2014). Effects of balancing holes on the performance and axial thrust of a centrifugal pump. *Transactions of the Korean Society of Mechanical Engineers*, **23**, 443–451.

[12] Davidson, L. (2024). *Fluid mechanics, turbulent flow and turbulence modeling.* Division of Fluid Dynamics, Department of Mechanics and Maritime Sciences, Chalmers University of Technology. Available at: `http://www.tfd.chalmers.se/~lada/postscript_files/solids-and-fluids.turbulent-flow.turbulence-modelling.pdf`

[13] Patel, M. (2014, November 24). Using multiple reference frame model for turbomachinery analysis. HiTech CFD. Retrieved from

https://www.hitechcfd.com/cfd-knowledgebase/using-multiple-reference-frame-model-for-turbomachinery-analysis.html

[14] Arora, J. S. (2005). Introduction to Optimum Design (3rd ed.). Elsevier Academic Press.

[15] Steuer, R.E. (1989). *Multiple Criteria Optimization: Theory, Computation, and Application.* Krieger, Malabar.

[16] Farina, G. (2024, April 25). Lecture 16: Bayesian optimization. MIT 6.7220/15.084 — Nonlinear Optimization. Massachusetts Institute of Technology.

[17] C. A. Micchelli, Y. Xu, and H. Zhang, "Universal Kernels," *Journal of Machine Learning Research*, vol. 7, no. 12, pp. 2651–2667, 2006.

[18] S. Burhenne, D. Jacob, and G. P. Henze, "Sampling based on Sobol' sequences for Monte Carlo techniques applied to building simulations," in *Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*, Sydney, Australia, Jan. 2011.

[19] H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, Society for Industrial and Applied Mathematics, 1992.

[20] DeepAI, *Posterior Predictive*, Machine Learning Glossary. Available at: `https://deepai.org/machine-learning-glossary-and-terms/posterior-predictive` [Accessed: 15-Jun-2025].

[21] Metacomp Technologies, *CFD++ User Manual*, Version 21.5, Metacomp Technologies, Agoura Hills, CA, USA, 2024.

A

Appendix

A.1 Python Code for Sobol Sequence Sampling

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats.qmc import Sobol

# Number of design parameters (dimensions)
dimensions = 4

# Recommended number of samples = 3 to 4    dimensions
num_samples = 12

# Initialize Sobol engine
sobol_engine = Sobol(d=dimensions, scramble=True)

# Generate normalized samples in [0, 1]^4
sobol_samples = sobol_engine.random(n=num_samples)

# Plot pairwise projections (2D scatter plots)
fig, axes = plt.subplots(2, 3, figsize=(12, 6))
axes = axes.flatten()
titles = [’Dim 1 vs Dim 2’, ’Dim 1 vs Dim 3’, ’Dim 1 vs Dim 4’,
          ’Dim 2 vs Dim 3’, ’Dim 2 vs Dim 4’, ’Dim 3 vs Dim 4’]
pairs = [(0,1), (0,2), (0,3), (1,2), (1,3), (2,3)]

for ax, (i, j), title in zip(axes, pairs, titles):
    ax.scatter(sobol_samples[:, i], sobol_samples[:, j], color=’
       green’, edgecolors=’black’)
    ax.set_xlabel(f’Dimension {i+1}’)
    ax.set_ylabel(f’Dimension {j+1}’)
    ax.set_title(title)

plt.tight_layout()
plt.show()
```

A.2 Python Code for Pareto Front Analysis

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

def is_pareto_efficient(costs):
    """Find the Pareto-efficient points (faster version)."""
    is_efficient = np.arange(costs.shape[0])
    n_points = costs.shape[0]
    next_point_index = 0
    while next_point_index < len(costs):
        nondominated_point_mask = np.any(costs < costs[next_point_index], axis=1)
        nondominated_point_mask[next_point_index] = True
        is_efficient = is_efficient[nondominated_point_mask]
        costs = costs[nondominated_point_mask]
        next_point_index = np.sum(nondominated_point_mask[:next_point_index + 1])
    mask = np.zeros(n_points, dtype=bool)
    mask[is_efficient] = True
    return mask

# Load your Excel (.csv) data
file_path = 'Target_Function_Pareto.csv'
df = pd.read_csv(file_path)

force = df['TF_AF']
head = df['TF_PH']

# Prepare data: Force (minimize), Head  & Power (maximize) is inverted
data = np.vstack((force, -head)).T

# Find Pareto front
pareto_mask = is_pareto_efficient(data)
pareto_indices = np.where(pareto_mask)[0]
print("Indices of Pareto-efficient points (Python index):", pareto_indices)
print("Corresponding SP numbers (1-based):", pareto_indices + 1)

pareto_points = data[pareto_mask]
force_pareto = pareto_points[:, 0]
head_pareto = -pareto_points[:, 1]

# Plot
plt.figure(figsize=(8, 6))
plt.scatter(force, head, color='lightgray', label='All Points including Red')
plt.scatter(force_pareto, head_pareto, color='red', label='Pareto Front')

# Annotate Pareto points
for i, (x, y) in enumerate(zip(force_pareto, head_pareto)):
    sp_number = pareto_indices[i] + 1
```

```
    plt.text(x, y, f'SP{sp_number}', fontsize=9, ha='right', va='bottom')

# Highlight SP24, SP25, SP26
for specific_index in [24, 25, 26]:
    if specific_index < len(df):
        force_val = df.loc[specific_index, 'TF_AF']
        head_val = df.loc[specific_index, 'TF_PH']
        sp_label = f'SP{specific_index + 1}'
        plt.scatter(force_val, head_val, s=40, color='blue', edgecolor='black', label=sp_label)
        plt.text(force_val, head_val, sp_label, fontsize=9, ha='left', va='top', color='blue')
        print(f"{sp_label} (Index {specific_index}): Force = {force_val:.4f}, Head = {head_val:.4f}")
    else:
        print(f"Index {specific_index} is out of bounds.")

plt.xlabel('Axial Force (Minimize)')
plt.ylabel('Head & Power (Maximize)')
plt.title('Pareto Front with 0.5, 1 & 1.25 Q* Combined')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

DEPARTMENT OF MECHANICS AND MARITIME SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY