

**UNIVERSIDADE FEDERAL DE ITAJUBÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA**

**METODOLOGIA DE OTIMIZAÇÃO MULTIFIDELIDADE BASEADA EM KRIGING  
HIERÁRQUICO APRIMORADO COM APLICAÇÃO AO PROJETO PRELIMINAR DE  
TURBINAS HIDROCINÉTICAS**

**David Cardoso Netto**

**Itajubá, agosto de 2024**

**UNIVERSIDADE FEDERAL DE ITAJUBÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA**

**David Cardoso Netto**

**METODOLOGIA DE OTIMIZAÇÃO MULTIFIDELIDADE BASEADA EM KRIGING  
HIERÁRQUICO APRIMORADO COM APLICAÇÃO AO PROJETO PRELIMINAR DE  
TURBINAS HIDROcinÉTICAS**

**Tese submetida ao Programa de Pós-Graduação em  
Engenharia Mecânica como parte dos requisitos  
para obtenção do Título de Doutor em Engenharia  
Mecânica.**

**Área de Concentração: Térmica, Fluidos e  
Máquinas de Fluxo (TFMF)**

**Orientador: Nelson Manzanares Filho**

**Agosto de 2024  
Itajubá – MG**

**UNIVERSIDADE FEDERAL DE ITAJUBÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA**

**David Cardoso Netto**

**METODOLOGIA DE OTIMIZAÇÃO MULTIFIDELIDADE BASEADA EM KRIGING  
HIERÁRQUICO APRIMORADO COM APLICAÇÃO AO PROJETO PRELIMINAR DE  
TURBINAS HIDROCINÉTICAS**

**Banca Examinadora:**

Prof. Dr. André Luiz Amarante Mesquita – UFPA

Prof. Dr. Antônio Marcos Gonçalves de Lima – UFU

Prof. Dr. Waldir de Oliveira – UNIFEI

Prof. Dr. Ramiro Gustavo Ramirez Camacho – UNIFEI

Prof. Dr. Nelson Manzanares Filho – UNIFEI

**Agosto de 2024**

**Itajubá – MG**

# DEDICATÓRIA

A Deus por me mostrar o caminho a seguir.  
À minha família por sempre acreditar em mim.  
Ao meu amor, Ísis.

# AGRADECIMENTOS

Me dediquei ao desenvolvimento deste trabalho por longos 4 anos e meio, e muita coisa aconteceu nesse tempo. Já não sou quem eu era quando comecei, assim como o próprio trabalho também mudou durante os anos. Uma lição aprendida durante este período é sobre a impermanência. E é interessante pensar sobre isso em um contexto de fluidos. Aquilo que escoar. Escoa de forma ordenada, mas também escoar com a aleatoriedade. Cada um com valor e momento apropriados. E é assim que me senti durante esses anos, fluindo por entre um caminho que não era sempre visível. Às vezes fluí ordenado, em outras, fluí caótico. Mas sempre reverberando a influência do que (quem) estava em contato comigo.

Agradeço ao Professor Nelson, meu orientador e a maior fonte de conhecimento com a qual já tive contato. Admiro como soube conduzir e adaptar a orientação para que eu me sentisse confortável em escoar por mim mesmo, mas também como soube dar direcionamento quando fui turbulento. Sempre me lembrarei com carinho das nossas discussões nas reuniões e nas aulas de Mecânica dos Fluidos. Agradeço imensamente por poder ter sido um pequeno sumidouro perto de tamanha fonte.

Agradeço meus colegas do LHV pelo apoio e pelas conversas, e ao Professor Ramiro que, juntamente com o Professor Waldir, são outras duas fontes constantes de conhecimento pelas quais eu pude absorver parte do que hoje sei. A vocês, o meu muito obrigado!

Agradeço as agências de fomento de pesquisa CAPES e FAPEMIG cujos recursos possibilitaram a realização deste trabalho.

Agradeço à minha família por todo o apoio e compreensão, parte do que sou, sou por vocês.

Por fim, e ainda mais importante, agradeço àquela que sempre foi margem e leito, que soube me acolher mesmo nos momentos de maior turbulência, aquela que sempre foi, ao mesmo tempo, o meu desaguar e o meu brotar, meu amor, Ísis.

# ΕΠΪΓΡΑΦΕ

*E pluribus unum*

# RESUMO

Existe um grande potencial hidrocínético em diferentes cursos de água corrente, como rios, canais e fluxos de usinas. O projeto de uma turbina pode envolver tanto o dimensionamento da turbina a partir de parâmetros previamente estabelecidos por meio de conhecimento empírico, como também a determinação desses parâmetros por meio de técnicas de otimização. Dentre os processos de otimização, o presente trabalho faz uso da otimização bayesiana, mais especificamente, da otimização global eficiente. Assim, este trabalho tem como objetivo principal o desenvolvimento de uma metodologia de otimização global mais eficiente, baseada em metamodelagem multifidelidade, e com aplicação no projeto de turbinas. Um novo método de metamodelagem multifidelidade baseado em Kriging Hierárquico aprimorado por Funções de Base Radial é desenvolvido. É proposto o aprimoramento do metamodelo construído com base nos pontos de alta fidelidade e a correção do metamodelo construído com base nos pontos de baixa fidelidade. O método proposto é aplicado em funções de teste analíticas e no projeto preliminar de uma turbina hidrocínética carenada com o objetivo de maximizar a potência hidráulica pela otimização dos parâmetros geométricos da turbina. Os parâmetros geométricos da ogiva e da nacele, e o ângulo de ataque do difusor são definidos como variáveis de projeto. A turbina é modelada em CFD com dois níveis de refinamento malhas, fina e grosseira, adotadas como as fontes de alta e baixa fidelidades, respectivamente. O rotor é aproximado por um disco atuador com uma queda de pressão fixa. O processo de otimização é feito comparando o método proposto com dois métodos da literatura. Os resultados da aplicação em funções de teste analíticas mostram que o método proposto teve melhor desempenho para o caso de teste de maior número de variáveis e com a mais alta multimodalidade. Para o caso real de engenharia, o método desenvolvido resulta em uma evolução média da potência hidráulica maior que a dos outros dois métodos para o primeiro quarto do orçamento computacional total, além de apresentar menor desvio padrão durante a maior parte do tempo de otimização. A geometria de turbina otimizada resulta no aumento do canal a jusante do rotor e no aumento do diâmetro da ogiva, culminando em um aumento de 21% no coeficiente de potência.

**Palavras-chave:** Otimização Bayesiana, metamodelagem multifidelidade, Kriging Hierárquico aprimorado por Funções de Base Radial, turbinas hidrocínéticas, Dinâmica dos Fluidos Computacional.

# ABSTRACT

There is significant hydrokinetic potential in various flowing water courses, such as rivers, canals, and plant outflows. Turbine design can involve either designing based on empirically established parameters or determining these parameters through optimization techniques. This study focuses on Bayesian optimization, specifically efficient global optimization. The main objective is to develop a more efficient global optimization methodology based on multifidelity surrogate modeling, applied to turbine design. A new multifidelity surrogate modeling method based on Hierarchical Kriging improved by Radial Basis Functions is developed. The proposed method involves improving the surrogate model based on high-fidelity data and correcting the surrogate model based on low-fidelity data. The method is applied to analytical test functions and to the preliminary design of a shrouded hydrokinetic turbine, aiming to maximize hydraulic power by optimizing the turbine's geometric parameters. The geometric parameters of the nose cone, nacelle, and diffuser angle of attack are defined as design variables. The turbine is modeled in CFD with two mesh refinement levels, fine and coarse, used as sources of high and low fidelity, respectively. The rotor is approximated by an actuator disk with a fixed pressure drop. The optimization process compares the proposed method with two methods from the literature. The results from the analytical test functions show that the proposed method performed better for the test case with the highest number of variables and the most multimodality. For the real engineering case, the developed method resulted in an average hydraulic power improvement greater than the other two methods for the first quarter of the total computational budget and showed lower standard deviation during most of the optimization time. The optimized turbine geometry resulted in an increase in the downstream channel of the rotor and an increase in the nose cone diameter, culminating in a 21% increase in the power coefficient.

**Keywords:** Bayesian Optimization, Multifidelity surrogate modeling, Hierarchical Kriging improved by Radial Basis Functions, Hydrokinetic Turbines, Computational Fluid Dynamics.

# LISTA DE FIGURAS

Figura 1.1	- Turbinas hidrocínéticas de eixo vertical (a) e de eixo horizontal (b)	2
Figura 3.1	- Representação qualitativa de um processo de otimização sequencial com o objetivo de minimização	16
Figura 3.2	- Distribuições de probabilidade atreladas ao processo de inferência bayesiana	17
Figura 3.3	- Distribuição fatorial $4^2$	19
Figura 3.4	- Hipercubo latino em três dimensões composto por dez pontos	20
Figura 3.5	- Representação do modelo Kriging e incerteza associada	22
Figura 3.6	- Interpretação gráfica da probabilidade de melhoria	38
Figura 3.7	- Representação de um processo de busca por função de aquisição	40
Figura 4.1	- Geometria base da turbina hidrocínética carenada a ser otimizada	43
Figura 4.2	- Geometria a ser otimizada e variáveis de projeto	44
Figura 4.3	- Domínio de cálculo e condições de contorno	46
Figura 4.4	- Estudo de convergência de malha	47
Figura 4.5	- Malha de maior refinamento e refinamento nas paredes	48
Figura 4.6	- Fluxograma do processo de otimização por FBR-HK	50
Figura 5.1	- Configuração real da turbina eólica de eixo vertical otimizada	52
Figura 5.2	- Modelo computacional da turbina otimizada	52
Figura 5.3	- Conjuntos de treino e de teste	53
Figura 5.4	- Fluxograma do processo de otimização	54
Figura 5.5	- Geometrias de aerofólio ótimas e geometria base	55
Figura 5.6	- $C_p$ versus TSR para geometrias ótimas e base	56
Figura 5.7	- Comparação dos ótimos encontrados pelos diferentes métodos	59
Figura 5.8	- Evolução dos processos de otimização para FBR-HK, IHK e SFK	64
Figura 5.9	- Evolução dos processos de otimização em termos de média e desvio padrão e custo computacional médio associado	66
Figura 5.10	- Comparação da evolução média dos processos de otimização	68
Figura 5.11	- Geometrias base e ótima de turbina hidrocínética	70

Figura 5.12	- Contornos de magnitude do vetor velocidade	71
Figura 5.13	- Distribuição de velocidade axial para geometria base e ótima	72
Figura A.1	- Modelo de disco atuador de uma turbina axial; $U$ representa a velocidade média do escoamento; 1, 2, 3 e 4 indicam a posição	76
Figura C.1	- Esquema analógico de uma célula neural biológica	85
Figura C.2	- Esquema analógico de uma rede neural biológica	86
Figura C.3	- Rede <i>feedforward</i> com uma camada oculta e uma camada de saída	87

## LISTA DE ABREVIATURAS E SIGLAS

<i>CFD</i>	- <i>Computational fluid dynamics</i>	2
<i>EGO</i>	- <i>Efficient Global Optimization</i>	2
<i>FBR-HK</i>	- Kriging hierárquico aprimorado por Funções de Base Radial	4
<i>PAE</i>	- Produção anual de energia	6
<i>MFSKO</i>	- Otimização multifidelidade sequencial por Kriging	8
<i>EI</i>	- Expectativa de melhoria	8
<i>AEI</i>	- Expectativa de melhoria aumentada	8
<i>LHS</i>	- Hipercubo latino	8
<i>HK</i>	- Kriging hierárquico	11
<i>MSE</i>	- Erro médio quadrático	12
<i>IHK</i>	- Kriging hierárquico aprimorado	12
<i>BO</i>	- Otimização Bayesiana	12
<i>MF-MES</i>	- Multifidelidade por busca de máxima entropia	13
<i>VF-EI</i>	- Expectativa de melhoria para multifidelidade	13
<i>MSP</i>	- Minimização da previsão do metamodelo	13
<i>PI</i>	- Probabilidade de melhoria	13
<i>LCB</i>	- Intervalo de confiança inferior	13
<i>TSR</i>	- Razão de ponta de pá	52
<i>RS</i>	- Superfície de resposta	53
<i>ELM</i>	- <i>Extreme Learning Machine</i>	53

## LISTA DE SÍMBOLOS

$p(\phi   x)$	- Distribuição inicial da variável de interesse em inferência Bayesiana	17
$p(y   x, \phi)$	- Distribuição de verossimilhança	17
$p(\phi   x, y)$	- Distribuição posterior	17
$\hat{y}(x)$	- Previsão do metamodelo	21
$Y(x)$	- Incerteza associada à previsão	21
$\mu$	- Média de um processo estocástico	21
$\sigma^2$	- Variância de um processo estocástico	21
$p_l$	- Parâmetro de suavidade da função na direção $l$	22
$\theta_l$	- Medida de taxa de variação correlação na direção $l$	22
$\Psi$	- Matriz de correlações no método de Kriging clássico	23
$\hat{\mu}$	- Previsão da média no processo de Kriging	24
$\hat{\sigma}^2$	- Previsão da variância no processo de Kriging	24
$\psi$	- Matriz coluna de correlações entre os pontos observados e o ponto de estimativa	25
$\hat{\Sigma}^2$	- Erro quadrático médio associado à previsão de um processo Gaussiano	26
$\beta_{0,lf}$	- Média do modelo de Kriging construídos utilizando os pontos de baixa fidelidade	30
$R_{lf}$	- Matriz de correlações entre os pontos de baixa fidelidade	30
$r_{lf}$	- Matriz coluna de correlações entre os pontos de baixa fidelidade e o ponto de estimativa	30
$\hat{y}_{lf}$	- Previsão do metamodelo de Kriging construído utilizando os pontos de baixa fidelidade	31
$R$	- Matriz de correlações entre os pontos de alta fidelidade	31
$r$	- Matriz coluna de correlações entre os pontos de alta fidelidade e o ponto de estimativa	31
$F$	- Vetor coluna composto pelas previsões do modelo de baixa fidelidade para os pontos de alta fidelidade	31

$\hat{S}_{HK}^2$	- Erro quadrático médio associado à previsão do Kriging Hierárquico	32
$h(\mathbf{x})$	- Superfície de resposta polinomial clássica	32
$\beta_{0,h}, \beta_i, \beta_{jk}$	- Coeficientes associados à superfície de resposta clássica	32
$\boldsymbol{\beta}$	- Vetor linha composto pelos coeficientes da superfície de resposta clássica	33
$\mathbf{f}$	- Vetor linha composto pela multiplicação da previsão do modelo de baixa fidelidade pelas variáveis de projeto	33
$\hat{S}_{IHK}^2$	- Erro quadrático médio associado à previsão do Kriging Hierárquico Aprimorado	34
$\boldsymbol{\beta}_{0,lf}$	- Coeficientes associados ao polinômio de segunda ordem da construção do Kriging de baixa fidelidade no Kriging Hierárquico Aprimorado por Funções de Base Radial	35
$g(\mathbf{x})$	- Funções de base radial	35
$\varphi$	- Núcleo da função de base radial	35
$\hat{S}_{FBR-HK}^2$	- Erro quadrático médio associado à previsão do Kriging Hierárquico Aprimorado por Funções de Base Radial	37
$P[I(\mathbf{x}^*)]$	- Probabilidade de melhoria	37
$E[I(\mathbf{x}^*)]$	- Expectativa de melhoria	39
$\alpha$	- Ângulo de ataque do difusor	44
$x_1, x_2$ e $x_3$	- Pontos de controle da geometria da ogiva	44
$d_{x2}, d_{y2},$ $d_{x3}$ e $d_{y3}$	- Variação dos pontos de controle $x_2$ e $x_3$	44
$te\_height$	- Raio do bordo de fuga da nacele	44
$te\_len$	- Comprimento do bordo de fuga da nacele	44
$C_p$	- Coeficiente de potência	52
$C_{P\_saida}$	- Coeficiente de potência calculado com referência à área de saída de turbina	71
$C_{P\_garganta}$	- Coeficiente de potência calculado com referência à área da garganta da turbina	71

## LISTA DE TABELAS

Tabela 4.1	- Intervalo das variáveis de projeto	45
Tabela 4.2	- Parametrização das variáveis de projeto	45
Tabela 5.1	- RMSE dos metamodelos para os pontos de teste	54
Tabela 5.2	- Valores ótimos e coeficiente de potência	55
Tabela 5.3	- Formulação das funções de teste	57
Tabela 5.4	- Definição das funções multifidelidade de teste	58
Tabela 5.5	- Sumário dos processos de otimização dos casos de teste 1, 4, 5 e 6	61
Tabela 5.6	- Configurações ótimas de cada método e respectiva potência hidráulica	62
Tabela 5.7	- Média de avaliações utilizadas pelos métodos	69
Tabela 5.8	- Comparação geometria base e ótima	71

# SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO</b>	1
1.1.	MOTIVAÇÃO DO TRABALHO	3
1.2.	OBJETIVOS DO TRABALHO	4
1.3.	ORGANIZAÇÃO DO TRABALHO	4
<b>2.</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	6
<b>3.</b>	<b>TÓPICOS REFERENTES A OTIMIZAÇÃO</b>	16
3.1.	PLANEJAMENTO DE EXPERIMENTOS	18
3.2.	MÉTODOS DE APROXIMAÇÃO	20
3.2.1.	KRIGING	21
3.2.2.	CO-KRIGING	26
3.2.3.	KRIGING HIERÁRQUICO (HK)	30
3.2.4.	KRIGING HIERÁRQUICO APRIMORADO (IHK)	32
3.2.5.	KRIGING HIERÁRQUICO APRIMORADO POR FUNÇÕES DE BASE RADIAL (FBR-HK)	34
3.2.6.	FUNÇÕES DE AQUISIÇÃO DE NOVOS PONTOS DE AMOSTRAGEM	37
<b>4.</b>	<b>METODOLOGIA: UM CASO DE TESTE REAL DE ENGENHARIA</b>	42
4.1.	DEFINIÇÃO DO PROBLEMA DE OTIMIZAÇÃO	42
4.2.	DESCRIÇÃO DO PROCESSO DE OTIMIZAÇÃO	45
<b>5.</b>	<b>RESULTADOS</b>	51
5.1.	RESULTADOS COMPARATIVOS DE METAMODELOS	51

5.2.	RESULTADOS DA APLICAÇÃO DE FBR-HK EM FUNÇÕES DE TESTE	57
5.3.	RESULTADOS DA APLICAÇÃO DE FBR-HK PARA O CASO REAL DE ENGENHARIA	61
5.4.	RESULTADOS DAS SIMULAÇÕES EM CFD E COMPARAÇÃO COM GEOMETRIA BASE	69
<b>6.</b>	<b>CONCLUSÕES E SUGESTÕES</b>	<b>73</b>
6.1.	CONCLUSÕES	73
6.2.	SUGESTÕES	74
	<b>ANEXO A - TEORIA UNIDIMENSIONAL DO DISCO ATUADOR E CFD</b>	<b>77</b>
A.1.	TEORIA UNIDIMENSIONAL DA QUANTIDADE DE MOVIMENTO E O LIMITE DE BETZ	77
A.2.	EQUAÇÕES DE CONSERVAÇÃO	80
A.2.1.	EQUAÇÃO DA CONTINUIDADE	80
A.2.2.	EQUAÇÃO DA QUANTIDADE DE MOVIMENTO LINEAR	80
A.3.	CFD. MÉTODO DE VOLUMES FINITOS	81
A.4.	CFD. MODELO DE TURBULÊNCIA $k-\omega$ SST	83
	<b>ANEXO B - SUPERFÍCIE DE RESPOSTA CLÁSSICA</b>	<b>85</b>
	<b>ANEXO C - REDES NEURAS ARTIFICIAIS</b>	<b>86</b>
C.1.	ARQUITETURAS DE REDE: <i>FEEDFORWARD</i> MULTICAMADA	87
C.2.	APRENDIZAGEM ASSISTIDA E AS TAREFAS DE APRENDIZAGEM	88
C.3.	REDES NEURAS DE FUNÇÕES DE BASE RADIAL	89
C.4.	TEOREMA UNIVERSAL DA APROXIMAÇÃO	92

C.5.	ESTRATÉGIAS DE APRENDIZADO E <i>EXTREME LEARNING MACHINE</i>	92
	<b>REFERÊNCIAS</b>	94

# 1. INTRODUÇÃO

A queima de combustíveis fósseis libera gases na atmosfera como CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O e o próprio vapor d'água, que agem retendo o calor na atmosfera e intensificando o mecanismo de efeito estufa, primariamente natural. Além disso, o processo de formação dos combustíveis fósseis leva milhares de anos, sendo, assim, fontes de energia de caráter não-renovável e limitado. Ainda, segundo a Agência Internacional de Energia, a previsão é de um crescimento de 80% na demanda por energia até o ano de 2026, quando comparado aos níveis de 2022 (IEA, 2024).

Nesse sentido, a busca por fontes de energias renováveis e com baixa emissão de gases de efeito estufa têm crescido nos últimos anos, atrelada a uma crescente tendência de investimento em geração distribuída, que produz energia perto dos consumidores, reduzindo custos de transmissão e aumentando a confiabilidade do sistema elétrico. Estima-se que as fontes de energia renováveis comporão mais de um terço da matriz energética mundial até 2025 (IEA, 2024).

Dentre as formas de geração distribuída, pode-se citar a energia solar fotovoltaica, a energia proveniente da queima de biomassa, a energia eólica e as pequenas centrais hidrelétricas. Existe um grande potencial hidrocínético em diferentes cursos de água corrente, como rios, canais, fluxo na base de barragens, canais de energia de usinas hidrelétricas de fio d'água, canais de descarga de usinas, canais em plantas de purificação de água e usinas termoelétricas.

As turbinas hidrocínéticas, no que tange aos princípios de funcionamento, podem ser classificadas, de acordo com a posição do seu eixo de rotação em relação à direção do escoamento incidente, em turbinas de eixo horizontal e turbinas de eixo vertical, exemplificadas na Figura 1.1.



Figura 1.1 – Turbinas hidrocínéticas de eixo vertical (a) e de eixo horizontal (b).

Na grande maioria das vezes, o projeto de uma turbina é feito utilizando modelos computacionais, pois o investimento necessário para uma instalação de alta qualidade para experimentos é geralmente mais alto do que o necessário para se obter uma descrição suficientemente completa de um escoamento por meio da Dinâmica dos Fluidos Computacional, da sigla inglesa CFD (VERSTEEG & MALALASEKERA, P. 2, 1995).

O projeto, em si, pode envolver tanto o dimensionamento da turbina a partir de parâmetros previamente estabelecidos por meio de conhecimento empírico, como também a determinação desses parâmetros por meio de técnicas de otimização. Pode-se, ainda, adotar uma abordagem de projeto preliminar, em que, por exemplo, o rotor da turbina é aproximado e simplificado por uma queda de pressão imprimida ao escoamento. O objetivo desse tipo de abordagem é a otimização dos parâmetros geométricos da turbina, preliminarmente ao projeto do rotor propriamente dito. As variáveis escolhidas para um processo de otimização recebem o nome de variáveis de projeto e os limites inferior e superior dessas variáveis definem o espaço de projeto/busca.

Pode-se dizer que o processo de otimização é fundamentalmente uma sequência de decisões. A cada iteração, um ponto do espaço de projeto é escolhido por meio de um método de busca como ponto para a próxima avaliação da função objetivo. Este processo se repete até que um critério de parada seja atingido, terminando o processo. Este tipo de processo de otimização é conhecido como otimização sequencial.

Dentre os processos de otimização sequencial, o presente trabalho tem enfoque no processo de Otimização Global Eficiente, da sigla inglesa EGO, desenvolvido por Jones *et al.* (1998) com

base em processos Gaussianos e na abordagem Bayesiana. A abordagem de otimização bayesiana faz uso de probabilidades e inferência bayesiana no processo de decisão, como, por exemplo, o tratamento da função objetivo como uma variável aleatória (GARNETT, 2023, P. 6).

Na otimização bayesiana baseada em metamodelos, o processo de busca é assistido pela construção e atualização iterativa de um metamodelo baseado em processos estocásticos; o processo de amostragem é guiado por uma função de aquisição de novos pontos (GHOREISHI E ALLAIRE, 2019). A construção do metamodelo auxilia na eficiência do processo de busca, enquanto a abordagem de amostragem sequencial guiada por uma função de aquisição reduz o número total de avaliações da função objetivo no decorrer do processo. Além disso, o uso de metamodelos traz consigo a vantagem da redução do custo computacional, uma vez que os metamodelos tem o objetivo de aproximar as funções objetivo. Assim, a estratégia de amostragem sequencial se mostra particularmente útil para aplicações envolvendo alto custo computacional ou experimental.

## **1.1. MOTIVAÇÃO DO TRABALHO**

O projeto e otimização de sistemas complexos requer geralmente o cálculo das variáveis de interesse em uma gama de configurações. Essas variáveis podem ser calculadas por diferentes meios, como modelos numéricos, medidas por experimentos ou por dados históricos. Essas fontes são chamadas de fontes de informação e cada fonte possui uma fidelidade associada, que diz respeito à sua capacidade de estimar o real valor da variável de interesse. Com o intuito de acelerar o processo de otimização, é interessante aproveitar-se das múltiplas fontes de informação disponíveis. E, uma das formas de se fazer isso se dá pela construção de metamodelos de baixo custo computacional que incorporem os diferentes níveis de fidelidade das diferentes fontes de informação correspondentes (LAM et al. 2015). Os métodos de otimização multifidelidade são os que fazem uso dessa estratégia e possibilitam a escolha de ajustes entre custo computacional e precisão.

Muito tem sido pesquisado em termos de desenvolvimento de novos métodos de metamodelagem, como será visto nas revisões da bibliografia, mas poucos esforços têm se dado na direção de metamodelagem utilizando multifidelidade e há uma carência na literatura quando se

fala da aplicação tanto de metamodelagem e, ainda mais, de métodos de metamodelagem multifidelidade no projeto de turbinas.

## **1.2. OBJETIVOS DO TRABALHO**

Dessa forma, este trabalho tem como objetivo principal o desenvolvimento de uma metodologia de otimização mais eficiente, baseada em metamodelagem multifidelidade, e com aplicação no projeto de turbinas.

Especificamente, este trabalho se propõe a:

- Desenvolver um método de otimização global sequencial que seja mais eficiente em termos do compromisso entre custo computacional e resultado ótimo.
- Desenvolver um método de metamodelagem multifidelidade baseado em Kriging Hierárquico aprimorado por Funções de Base Radial (FBR-HK).
- Aplicar o método desenvolvido no projeto preliminar de uma turbina hidrocínética.
- Otimizar os parâmetros geométricos de uma turbina hidrocínética com o objetivo de maximizar a potência hidráulica.

## **1.3. ORGANIZAÇÃO DO TRABALHO**

No capítulo 2 são mostrados artigos relevantes para melhor entendimento do estado da arte sobre otimização de turbinas, sobre métodos de metamodelagem sequencial usando diferentes métodos de metamodelagem e aplicações desses métodos em projetos de engenharia, além de estudos sobre o desenvolvimento de funções de aquisição.

No capítulo 3 são apresentados alguns tópicos referentes a otimização, como o processo de otimização bayesiana e como se dá o processo de busca por novos pontos, conceitos de planejamento de experimento, os métodos de aproximação e as equações envolvidas no método de Kriging, incluindo suas variações multifidelidade tais como co-Kriging, Kriging hierárquico, Kriging hierárquico aprimorado e o método proposto, Kriging hierárquico aprimorado por funções de base radial, e por fim, são apresentadas algumas funções de aquisição disponíveis na literatura.

O capítulo 4 descreve a metodologia utilizada para a aplicação do método FBR-HK em um caso real de engenharia envolvendo a otimização do projeto preliminar de uma turbina hidrocínética.

O capítulo 5 traz os resultados de um estudo comparativo entre técnicas de metamodelagem com aplicação na otimização de turbinas eólicas, tópico estudado na dissertação de mestrado. Este estudo teve como objetivo avaliar o desempenho do método de Kriging para a otimização no contexto de turbinas. O capítulo também apresenta os resultados da aplicação em funções de teste de dois métodos de metamodelagem baseados em Kriging existentes na literatura e do método FBR-HK. Por fim, o capítulo 5 traz os resultados da aplicação do método FBR-HK no caso real de engenharia.

O capítulo 6 apresenta as conclusões extraídas através dos resultados obtidos no trabalho, tanto pela aplicação nas funções de teste quanto pela aplicação no caso real de engenharia, além de sugestões de trabalhos futuros tanto na direção de melhorar o método FBR-HK quanto no projeto de turbinas hidrocínéticas.

Os Anexos trazem a fundamentação teórica necessária para o entendimento da abordagem de disco atuador, utilizada na modelagem da turbina hidrocínética carenada, e dos métodos de volumes finitos e CFD. A seção também traz alguns tópicos sobre superfície de resposta clássica e redes neurais artificiais. Essas são as técnicas de metadodelagem usadas no Capítulo 5 em comparação com a técnica de Kriging.

## 2. REVISÃO BIBLIOGRÁFICA

Como o comportamento do campo de escoamento de turbinas é altamente turbulento, as simulações numéricas se dão, muitas vezes, em regime transiente, fazendo-se necessário o uso de malhas refinadas e que a solução do escoamento seja feita em pequenos incrementos de tempo. Resulta disso que o custo computacional associado é elevado e dificulta o emprego de estratégias de otimização que necessitem de constantes cálculos de CFD. Nesse sentido, a metamodelagem se apresenta como uma técnica efetiva para a redução do custo computacional.

Além disso, o projeto de uma turbina envolve a consideração de muitas variáveis relacionadas à geometria da pá e às condições de operação. Dessa forma, a avaliação de todas as variáveis simultaneamente pode aumentar significativamente o tempo de análise e o custo computacional.

Neto *et al.* (2018) desenvolveram uma metodologia de otimização de projeto de pás de turbinas eólicas de eixo horizontal por meio do uso de métodos estocásticos de otimização multiobjetivo. Os objetivos da otimização foram maximização da produção anual de energia (PAE) e a minimização da massa da pá. As variáveis de projeto utilizadas foram a distribuição de corda, ângulo de torção e espessura. Os valores encontrados de PAE e de massa da pá para a solução ótima foram próximos às tendências de mercado. A configuração ótima de pá apresentou ângulos de ataque por toda a pá dentro de uma distância segura do ângulo de estol. A distribuição de coeficiente de potência mostrou maiores valores entre a raiz da pá e, aproximadamente, metade do comprimento da pá, com valores mais baixos na ponta, devido aos efeitos de ponta de pá. Os autores concluíram que a configuração ótima encontrada era passível de produção na prática devido à geometria mais suave.

Por meio da modificação dos contornos de seção das pás, Lee e Kwon (2020) otimizaram o desempenho aerodinâmico de uma turbina eólica de eixo horizontal com objetivo de reduzir o custo de energia. Os efeitos de deformação aeroelástica foram avaliados juntamente com os efeitos aerodinâmicos. As otimizações foram feitas por algoritmo genético e redes neurais artificiais foram utilizadas para derivar as funções de aproximação entre as variáveis de projeto e o desempenho aerodinâmico. As variáveis de projeto escolhidas foram as coordenadas das seções da pá perpendiculares à corda em locais determinados. Em baixas velocidades de vento, a potência produzida pela turbina otimizada foi similar à da turbina original. Para maiores velocidades de

vento, houve uma melhora de até 6% aliada a uma redução de cerca de 1% no custo de energia. Os autores concluíram que a abordagem considerando ambos os efeitos aeroelásticos e aerodinâmicos pode ser útil no aperfeiçoamento de rotores de turbinas eólicas de eixo horizontal.

O princípio de operação de uma turbina hidrocínética se assemelha ao de uma turbina eólica no sentido de que ambas utilizam a energia cinética do escoamento como principal fonte de potência. Neste contexto, o trabalho de Rezek *et al.* (2021) teve como objetivo a otimização da geometria do difusor de uma turbina hidrocínética carenada para a maximização do coeficiente de potência. As variáveis de projeto determinadas foram os pontos de controle de curvas de Bezier que representavam a geometria do difusor, o ângulo de ataque do perfil do difusor e a queda de pressão do rotor. O processo de otimização se deu utilizando o método de recozimento simulado. O rotor foi aproximado por um disco atuador e o cálculo do escoamento foi feito por meio de CFD. Os autores estudaram ainda o efeito da adição de uma folga de topo no resultado. O processo de otimização resultou em um coeficiente de potência de 98% sem folga de topo com um aumento para 118% na configuração com folga de topo. Rezek *et al.* (2021) observaram que o tamanho da folga deveria ser apenas o suficiente para amenizar os efeitos de separação no intradorso do difusor, e concluíram que a queda de pressão do disco atuador deveria ser uma fração de 8/9 da pressão dinâmica do escoamento livre, corroborando os estudos de Bagheri-Sadeghi *et al.* (2018) e Jamieson (2009).

Rezek *et al.* (2023) propuseram uma nova metodologia para o projeto de rotores carenados de turbinas hidrocínéticas. O método considerava a influência de toda a grade linear em cada seção anular do rotor. O processo de projeto partia de um projeto preliminar de configuração de geometria do difusor e de nacele e queda de pressão do rotor, que utiliza o método de disco atuador e CFD 2D. A metodologia utilizou-se dos dados de vazão mássica e distribuição de velocidade axial no rotor obtidos do projeto preliminar e dos coeficientes adimensionais de turbomáquinas para projetar as pás. Os autores consideraram dois tipos de configurações de pá, com e sem folga de raiz, para investigar os efeitos no escoamento e no desempenho da turbina. Os resultados obtidos por meio de CFD 3D demonstraram que a metodologia proposta conseguiu estimar com precisão a condição de projeto de máxima potência, validando o método, e que o uso de folga de raiz aumentou o coeficiente de potência. Rezek *et al.* (2023) destacaram, por meio dos resultados, que o método de disco atuador, ainda que seja uma simplificação, conseguiu estimar com adequada precisão o

comportamento da turbina em diversas condições reais de operação. Os autores apontaram como possibilidade de melhoria a otimização do projeto preliminar.

Uma alternativa que também tem se mostrado atrativa é uso de otimização multifidelidade, em que diferentes graus de precisão para o cálculo das funções objetivo são utilizados, reduzindo, assim, ainda mais o custo computacional total do processo de otimização.

Com base nessa ideia, Huang *et al.* (2006) partiram da abordagem de otimização sequencial por Kriging e desenvolveram um algoritmo de otimização que pudesse determinar tanto a localização quanto o nível de fidelidade do próximo ponto de busca a ser calculado. O método foi chamado de otimização multifidelidade sequencial por Kriging (MFSKO). O processo de aquisição de novos pontos era feito por meio da expectativa de melhoria (EI) aumentada (AEI). Os autores observaram um desempenho robusto do método quando aplicado a funções de teste da literatura. Por fim, Huang *et al.* (2006) ressaltaram que o método é melhor indicado quanto maior for a diferença de custo computacional entre as fidelidades.

Lam *et al.* (2015) desenvolveram um metamodelo com base em processos gaussianos e utilizaram-no em um processo de otimização multifidelidade com o objetivo de minimizar o coeficiente de arrasto de um perfil NACA0012 com restrições de coeficiente de sustentação. O método era composto por duas fidelidades com razão de custo computacional de 1:3600. O metamodelo gerado inicialmente pelos autores foi utilizado como função de baixa fidelidade, e um código CFD foi utilizado como função de alta fidelidade. Os autores observaram que, durante o processo de otimização, a função de baixa fidelidade foi utilizada 194 vezes, contra 15 vezes para a de alta fidelidade. Lam *et al.* (2015) relataram que a função barata foi utilizada pelo método para explorar o espaço de busca, ao passo que a função cara foi usada apenas para refinar os resultados, quando necessário. O custo computacional do processo foi reduzido pelo algoritmo proposto.

Amrit *et al.* (2016) investigaram estratégias de aceleração para otimização multiobjetivo de aerofólios em escoamento transônico. Metamodelos baseados em multifidelidade (co-Kriging) e redução do espaço de projeto foram técnicas utilizadas. O cálculo do escoamento transônico foi feito por meio de CFD utilizando dois níveis de refinamento de malha, sendo cada um dos níveis adotado como um nível de fidelidade. O conjunto de treino para construção do metamodelo foi gerado por hipercubo latino (LHS) contendo pontos de ambas as fidelidades. Os autores observaram que o método multifidelidade proposto foi pelo menos uma ordem de grandeza mais rápido que o método utilizando apenas o nível de fidelidade mais alto.

Comparando o método de Kriging e co-Kriging na modelagem da esteira de turbinas eólicas, Wang *et al.* (2016) conseguiram demonstrar como a utilização de metamodelagem multifidelidade poderia acelerar os cálculos de produção de energia e potência do vento. Dois níveis de fidelidade foram utilizados. Um modelo semianalítico e semiempírico foi adotado como baixa fidelidade e um modelo CFD refinado, como alta fidelidade. Amostragem por hipercubo latino (LHS) foi utilizada para construção do conjunto de treino. O método de Kriging necessitou de seis vezes mais pontos de alta fidelidade do que o método de co-Kriging. Por fim, Wang *et al.* (2016) concluem que o método de co-Kriging apresentou erros menores do que o método de Kriging, além de reduzir significativamente o número de avaliações de alta fidelidade necessárias.

Pholdee *et al.* (2021) utilizaram otimização assistida por metamodelagem para maximizar o coeficiente de potência de uma turbina eólica de eixo horizontal. As variáveis de projeto escolhidas foram os ângulos de torção em três diferentes posições ao longo da pá. Kriging foi escolhido como método de metamodelagem. Os pontos para construção do metamodelo foram determinados por meio de LHS e o cálculo do escoamento para esses pontos foi feito utilizando CFD. O metamodelo construído foi otimizado por meio de algoritmo genético para 10 rodadas de otimização. Uma vez encontrado o ponto de ótimo do metamodelo, os autores verificaram o valor de coeficiente de potência por meio de CFD e observaram uma melhora no coeficiente de potência de 4,76% quando comparado com a geometria base, com um erro atrelado de apenas 0,67% entre metamodelo e CFD.

O trabalho de Jasa *et al.* (2022) teve como objetivo estudar as melhores práticas para o uso de métodos de otimização multifidelidade para aplicações de projeto de energia eólica. Os autores utilizaram um método de otimização multifidelidade baseado em região de confiança e propuseram uma nova função de correção entre os níveis de fidelidade baseada em Kriging. O método foi aplicado em três casos de projeto de turbinas eólicas, envolvendo projeto de pá, projeto de controlador e disposição de parque eólico. Além disso, os autores otimizaram os mesmos casos utilizando apenas uma fidelidade, tanto baixa quanto alta, à título de comparação. Os resultados de ótimo encontrados pelo método proposto distanciaram-se do ótimo do processo com alta fidelidade por aproximadamente 1% para os três casos, porém com metade do custo computacional. Jasa *et al.* (2022) apontaram trabalhos futuros na direção da melhora do desempenho do processo utilizando outros métodos de multifidelidade e algoritmos mais inteligentes baseados em expectativa de melhoria. Por fim, concluíram que o desempenho dos algoritmos é bastante

dependente do problema envolvido e que, para outros problemas, outros métodos poderiam se mostrar mais vantajosos.

Os parâmetros escolhidos como variáveis de projeto na grande maioria dos estudos de otimização são definidos como parâmetros com valor constante, isto é, a otimização muitas vezes se dá para uma configuração específica dos parâmetros, quando, em aplicações reais de engenharia, estes parâmetros estão sujeitos a incertezas, tais como propriedade dos materiais, dimensões e condições de operação. E quanto maior a sensibilidade do desempenho otimizado às variações dos parâmetros de projeto, diz-se que menor é a robustez desse projeto/desempenho.

Nesse sentido, a otimização robusta se mostra eficaz na tarefa de otimizar um determinado desempenho ao mesmo tempo em que aumenta a sua robustez. Um exemplo é o trabalho de Zheng *et al.* (2020) que teve por objetivo o desenvolvimento de uma metodologia de otimização robusta com auxílio de metamodelagem por Kriging. A metodologia foi aplicada para a maximização da primeira frequência natural da pá de uma turbina eólica de eixo horizontal. As variáveis de projeto envolviam as espessuras dos diferentes materiais que compõem o laminado e a largura da capa da longarina. Restrições de tensão e peso da pá foram aplicadas. O modelo de Kriging foi construído com base em uma amostra gerada por LHS, e então, utilizado no processo de otimização não-robusta. Os resultados da otimização típica mostraram um aumento de 19% na frequência natural, porém a análise de robustez mostrou que o nível  $6\sigma$  da frequência natural teve uma redução de 61%. Assim, os autores prosseguiram para uma otimização robusta utilizando Monte Carlo e obtiveram um aumento da frequência natural de 15,4%, porém com um aumento de 90% no nível  $6\sigma$  da frequência natural quando comparada com a geometria base. Zheng *et al.* (2020) concluíram que a metodologia de otimização robusta desenvolvida não só aumentou a frequência natural da pá, como também foi capaz de minimizar a sensibilidade do resultado em relação às variáveis de projeto.

Outro método de se tratar o problema das incertezas é chamado de otimização baseada em confiabilidade, que visa a maximização ou minimização de uma função objetivo, satisfazendo restrições de confiabilidade, avaliadas por análises de confiabilidade (CORSO, 2010).

Song *et al.* (2021) propuseram um novo método de otimização baseada em confiabilidade assistida por metamodelagem sequencial com o objetivo de reduzir o custo computacional. Para isso, as funções de restrição e objetivo foram modeladas por Kriging sequencial, alimentado por uma função de aquisição. Este, então, foi utilizado no processo de otimização e nas análises de

confiabilidade. Os autores compararam o método proposto com outros métodos da literatura para três casos de teste e observaram que aquele encontrou ótimos próximos aos encontrados por estes com redução no custo computacional em termos de uma ordem de grandeza. Song *et al.* (2019) apontaram como trabalhos futuros a busca por métodos de metamodelagem de maior precisão e eficiência.

No trabalho de Meng *et al.* (2023) foi desenvolvida uma estratégia de otimização multidisciplinar baseada em confiabilidade assistida por um modelo de Kriging sequencial com o intuito de melhorar a eficiência computacional do método. O modelo de Kriging foi utilizado para aproximar as funções de desempenho e função objetivo. O método foi testado para a otimização dos parâmetros geométricos do suporte do rotor de uma turbina hidráulica com o objetivo de minimizar a massa do rotor. Os autores atribuíram um intervalo de incerteza às variáveis de projeto. Os resultados mostraram que o método proposto encontrou um valor ótimo de massa do rotor 1,1% maior do que o encontrado pelo modelo base, porém com um custo computacional 13 vezes menor.

Wang *et al.* (2024) desenvolveram uma metodologia de otimização sequencial baseada em metamodelagem para a otimização da disposição de um parque eólico. A posição de cada turbina foi otimizada dentro de um espaço delimitado com o objetivo de maximizar a produção anual de energia por meio de um processo de custo computacional reduzido. O metamodelo foi construído pelo método de Kriging e treinado por um conjunto de dados cuja resposta foi calculada usando CFD, tratando as turbinas como discos atuadores. O método proposto foi comparado com o ótimo obtido por um processo de otimização genética convencional. Os autores observaram que a configuração ótima de parque eólico encontrada pelo método assistido por Kriging foi similar ao valor encontrado pela otimização genética, porém com menos da metade do custo computacional. Wang *et al.* (2024) concluíram que, ainda que o método proposto tenha reduzido o custo pela metade, este ainda foi de três a oito dias, a depender das configurações de variáveis de projeto, e que o método seria mais adequado para estágios mais tardios de projeto, para os quais os parâmetros de projeto já teriam sido definidos por métodos mais eficientes de baixa fidelidade.

Outro método de metamodelagem multifidelidade envolvendo Kriging é o chamado Kriging hierárquico (HK). Foi proposto por Han e Görtz (2012) com a motivação de desenvolver um método multifidelidade que fosse tão simples quanto os métodos baseados em correção e tão preciso quanto os métodos baseados em co-Kriging. O método HK utilizava um segundo modelo de Kriging como um modelo de tendência do modelo Kriging primário da função de interesse.

Dessa forma, o modelo é construído hierarquicamente. Han e Görtz (2012) ressaltaram que o método HK possuía melhor capacidade de estimativa do erro médio quadrático (MSE) devido à inclusão direta da previsão de baixa fidelidade no cálculo do MSE. Pela aplicação em funções de teste, os autores observaram que o método HK superou o método de co-Kriging, convergindo mais rapidamente. Han e Görtz (2012) concluíram que a estimativa de MSE melhorada proporcionada pelo método desenvolvido representava uma vantagem para aplicação de métodos de aquisição baseados no erro.

Com o intuito de melhorar o método proposto por Han e Görtz (2012), Hu *et al.* (2016) propuseram um aprimoramento do método HK chamado de Kriging hierárquico aprimorado (IHK). Os autores estudaram a substituição do fator de escala, que reflete a influência do modelo de baixa fidelidade na previsão do modelo HK, de um valor originalmente constante por uma superfície de resposta polinomial clássica. A estimativa de MSE pelo método IHK era composta tanto pela previsão do modelo de baixa fidelidade como pela localidade da previsão. Os autores observaram uma melhor estimativa do MSE e uma maior precisão na aproximação do modelo quando comparado ao HK e Kriging.

Xu *et al.* (2020) utilizaram Kriging hierárquico na proposição de um método de projeto de propulsores de hélice mais eficiente. Dois níveis de fidelidade foram usados, o nível mais baixo composto por um método baseado em elemento de pá e teoria de vórtices e o nível mais alto, por CFD. Diferentes critérios de amostragem sequencial foram utilizados concomitantemente. O plano de experimentos, que corresponde ao conjunto de treino dos metamodelos, foi construído por LHS. Xu *et al.* (2020) compararam dois processos de otimização, usando Kriging e HK, e observaram que ambos os processos obtiveram geometrias semelhantes, porém o processo de otimização com HK foi 17,4% mais rápido.

Li *et al.* (2020) propuseram um processo de otimização bayesiana (BO) multifidelidade baseado em redes neurais profundas com o intuito de capturar todos os tipos de relações complexas entre as fidelidades. Foi atribuída uma rede neural a cada fidelidade. Cada rede neural era alimentada tanto pela entrada original como pela saída da rede neural da fidelidade anterior, como forma de propagar a informação através das fidelidades. O método proposto foi avaliado em 3 funções de teste e em aplicações reais de engenharia e comparado com outros métodos existentes, inclusive a otimização multifidelidade sequencial por Kriging (MFSKO) de Huang *et al.* (2006). Os autores observaram que o método por eles desenvolvido alcançou os objetivos com menores

tempos de investigação em funções de teste. Para aplicações reais de engenharia, Li *et al.* (2020) observaram reduções de tempo na ordem de 25 a 60 vezes em comparação com outros métodos, tais como multifidelidade por busca de máxima entropia (MF-MES) de Takeno *et al.* (2020), e tempo equiparável quando comparado com o MFSKO.

Jones *et al.* (1998) introduziram o uso de processos estocásticos para guiar o procedimento de otimização, sob a denominação de “otimização global eficiente” (EGO), um ramo da otimização bayesiana. O trabalho de Zhang *et al.* (2018) teve como motivação o desenvolvimento de um método adaptativo baseado em EGO para problemas multifidelidade. O método HK de Han e Görtz (2012) foi utilizado como metamodelo. Os autores formularam um processo de aquisição de novos pontos com base na expectativa de melhoria (EI) para multifidelidade (VF-EI). Os conjuntos iniciais para cada fidelidade foram gerado usando LHS. Zhang *et al.* (2018) observaram que o método VF-EI tendeu a acrescentar mais pontos de baixa fidelidade que de alta fidelidade no geral e, quando comparado aos métodos de EI clássico de Jones *et al.* (1998) e EI aumentada de Huan *et al.* (2006), VF-EI fez uso de menos pontos de alta fidelidade que ambos os outros métodos.

Wang *et al.* (2018) utilizaram paralelamente múltiplos critérios de amostragem de novos pontos com o intuito de tornar o processo de otimização sequencial mais eficiente. Os autores propuseram que, com a combinação de vários critérios, os respectivos pontos negativos de cada critério poderiam ser complementados pelos outros. Conjuntos iniciais das fidelidades foram construídos usando LHS. Os critérios de amostragem empregados foram minimização da previsão do metamodelo (MSP), EI, probabilidade de melhoria (PI), erro quadrático médio (MSE) e intervalo de confiança inferior (LCB). A quantidade de amostras de cada critério era decidida pelo algoritmo. Resultados indicaram que a eficiência do processo de otimização foi aumentada significativamente quando comparada tanto ao EGO clássico quanto aos outros critérios usados isoladamente.

Segundo Noè e Husmeier (2018), a expectativa de melhoria não leva em conta a incerteza associada à melhoria esperada, e pontos indicados pela expectativa de melhoria poderiam apresentar uma alta melhoria, porém com uma variabilidade também alta, isto é, a melhoria real poderia ser baixa. Assim, os autores propuseram uma nova função de aquisição com base na variância do quantificador da melhoria que, associado à expectativa de melhoria, constituía a função nomeada de expectativa de melhoria escalada (ScaledEI). Resultados de aplicações em funções de teste indicaram que a função proposta teve desempenho tão bom quanto ou melhor que

outros métodos como EI clássico, PI, LCB e MSE. Noè e Husmeier (2018) compararam, ainda, o método proposto com métodos clássicos de otimização global, como algoritmo genético, busca global e enxame de partículas. ScaledEI precisou de 3 ordens de grandeza a menos de avaliações para chegar ao ótimo quando comparado ao algoritmo genético, o que, para os autores, comprovaria a eficácia da otimização bayesiana.

Ruan *et al.* (2020) desenvolveram um método de amostragem multifidelidade baseado em probabilidade de melhoria estendida utilizando co-Kriging. A função de aquisição de novos pontos era composta por um produto da probabilidade de melhoria pela correlação entre as fidelidades, pela razão de custo computacional das fidelidades e por um termo função de densidade de amostragem, a fim de evitar amostragens desnecessárias. Os conjuntos iniciais foram gerados por LHS. O método, chamado de VF-PI, foi testado e comparado com outros métodos como VF-EI, AEI e o próprio método proposto utilizando apenas pontos de alta fidelidade. Diferentes valores de correlação entre as fidelidades e de razão de custo computacional foram avaliados. O método proposto apresentou melhor desempenho para maiores razões de custo computacional. Para fortes correlações entre as fidelidades, o método VF-PI obteve os melhores resultados dentre os métodos. Já para correlações fracas, VF-EI teve um melhor desempenho que VF-PI.

Xu *et al.* (2021) propuseram um processo de otimização bayesiana híbrida acoplado ao algoritmo de otimização local de Broyden-Fletcher-Goldfarb-Shanno (BFGS) com a utilização de uma nova função de aquisição adaptativa baseada em EI. Para ajustar a capacidade de exploração do método e evitar que o processo ficasse preso em mínimos locais, a função de aquisição era constituída por uma função de penalidade e pela razão de momentum de exploração, além da EI clássica. O método, chamado de HB2O, foi comparado com seis outros otimizadores, como otimização bayesiana clássica, recozimento simulado (SA), enxame de partículas (PSO), colônia de abelhas artificial (ABC), busca de padrões (PS) e estimativa de momento adaptativo (ADAM). A função de aquisição desenvolvida foi comparada com outras quatro funções de aquisição baseadas, também, em EI. Pela aplicação em funções de teste, os autores observaram que HB2O teve desempenho melhor que os demais otimizadores, destacando sua capacidade de não ficar preso em mínimos locais. Por fim, Xu *et al.* (2020) notaram que a função de aquisição proposta foi mais eficiente que as outras avaliadas, para aplicação no método HB2O.

Ainda com o objetivo de pesquisar funções de aquisição mais eficientes, Shu *et al.* (2021) propuseram um método de aquisição de novos pontos baseado em expectativa de melhoria

adicional usando HK. O método foi comparado com o EGO clássico e VF-EI. A função de aquisição era composta pela expectativa de melhoria, a razão do custo computacional das fidelidades e a esperança matemática da expectativa de melhoria de um novo ponto de baixa fidelidade. Os autores ressaltaram que uma diferença chave entre o método proposto e o VF-EI é que aquele consideraria a razão de custo computacional e, assim, tenderia a adicionar mais pontos de baixa fidelidade do que de alta à medida em que a razão aumentasse. Testes mostraram que, de fato, o método proposto pôde usar de mais pontos de baixa fidelidade que os outros métodos quando o modelo de baixa fidelidade possuía menor custo computacional. Shu *et al.* (2021) ressaltaram que o método ainda dependia da razão de custo computacional, que foi definida empiricamente.

### 3. TÓPICOS REFERENTES À OTIMIZAÇÃO

Otimização é fundamentalmente uma sequência de decisões. A cada iteração, um ponto do espaço de projeto é escolhido por meio de um método de busca como ponto para a próxima avaliação da função objetivo. Este processo se repete até que um critério de parada seja atingido, terminando o processo. Este tipo de processo de otimização é conhecido como otimização sequencial. A Figura 3.1 exemplifica esse processo.

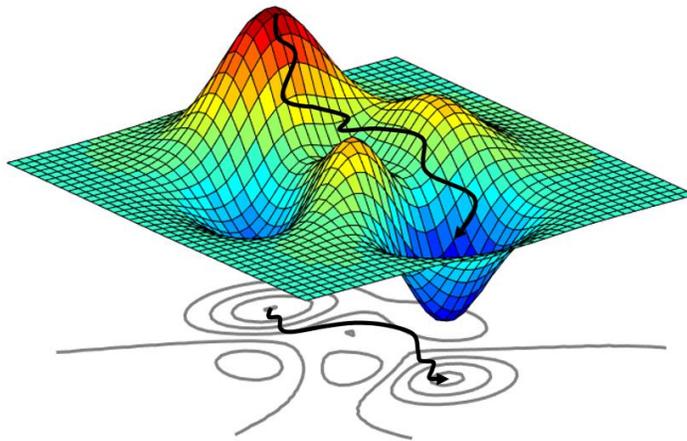


Figura 3.1 – Representação qualitativa de um processo de otimização sequencial com o objetivo de minimização.

Dentre os processos de otimização sequencial, o presente trabalho tem enfoque no processo de Otimização Global Eficiente, da sigla inglesa EGO, desenvolvido por Jones *et al.* (1998) com base em processos Gaussianos e na abordagem Bayesiana. A abordagem de otimização bayesiana faz uso de probabilidades e inferência bayesiana no processo de decisão (GARNETT, 2023, P. 6).

Segundo Jones *et al.* (1998), essa abordagem de aproximação de funções por processos estocásticos possui passado em três campos: geologia matemática, otimização global e em estatística. No campo da geologia matemática, essa abordagem recebeu o nome de Kriging e tem origens da década de 1960. Já no campo da otimização global, tal abordagem recebe o nome de otimização global Bayesiana.

Na abordagem de inferência bayesiana, todas as quantidades desconhecidas são tratadas como variáveis randômicas, isto é, cada variável é representada por uma distribuição de probabilidade, com média e variância atreladas, que refletem os valores possíveis para aquela variável. Assim, o processo começa com uma distribuição inicial, chamada de anterior (*prior* do

inglês),  $p(\phi | x)$  que engloba os valores plausíveis considerados para a variável  $\phi$ , representada pela distribuição da Figura 3.2(a).

A distribuição anterior é, agora, refinada em face de um novo resultado observado. Suponha que a observação mostre que a função objetivo  $\phi$  no ponto  $x$  tem o valor de  $y$ , como toda variável na abordagem bayesiana é representada por uma distribuição de probabilidade, essa observação  $y$  recebe uma distribuição, chamada de verossimilhança (*likelihood* do inglês), representada pela Figura 3.2(b).

Finalmente, dada a nova observação, encontra-se a nova distribuição, chamada de posterior, pela aplicação do teorema de Bayes:

$$p(\phi | x, y) = \frac{p(\phi | x)p(y | x, \phi)}{p(y | x)} \quad (3.1)$$

Assim, a distribuição posterior é proporcional a anterior ponderada pela verossimilhança. Em outras palavras, a posterior representa um ajuste entre a experiência que se tinha, representada pela anterior, e a informação contida nos dados obtidos, representada pela verossimilhança (GARNETT, 2023, P. 7). A Figura 3.2 retrata as três distribuições.

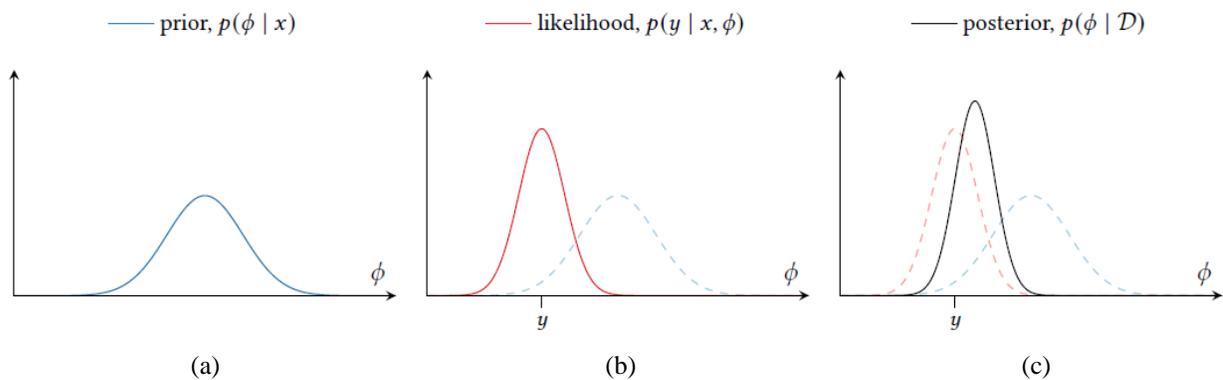


Figura 3.2 – Distribuições de probabilidade atreladas ao processo de inferência bayesiana.

Fonte: Garnett, 2023, p. 7.

Na otimização bayesiana baseada em metamodelos, o processo de busca pelo ótimo é assistido pela construção e atualização, iterativamente, de um metamodelo baseado em processos estocásticos; o processo de amostragem é guiado por uma função de aquisição de novos pontos

(GHOREISHI E ALLAIRE, 2019). A construção do metamodelo auxilia na eficiência do processo de busca, enquanto a abordagem de amostragem sequencial guiada por uma função de aquisição reduz o número total de amostras do processo. Além disso, o uso de metamodelos traz consigo a vantagem da redução do custo computacional, uma vez que os metamodelos tem o objetivo de aproximar as funções objetivo. Assim, a estratégia de amostragem sequencial se mostra particularmente útil para aplicações envolvendo alto custo computacional ou experimental.

O projeto e otimização de sistemas complexos frequentemente requer o cálculo das variáveis de interesse em uma gama de configurações, essas variáveis podem ser calculadas por diferentes meios, como modelos numéricos, medidas por experimentos ou por dados históricos. Essas fontes são chamadas de fontes de informação e cada fonte possui uma fidelidade associada, que diz respeito à sua capacidade de estimar o real valor da variável de interesse.

Ainda no contexto da otimização sequencial assistida por metamodelagem, a literatura destaca a utilidade dos metamodelos de baixo custo computacional que integram diferentes níveis de fidelidade das várias fontes de informação para reduzir ainda mais o custo computacional. Essa abordagem, conhecida como otimização multifidelidade, tem como principal vantagem a capacidade de equilibrar custo computacional e precisão.

### **3.1. PLANEJAMENTO DE EXPERIMENTOS**

Em um processo de otimização, existem dois aspectos que se mostram importantes: o planejamento de experimentos e a análise estatística dos dados. Segundo Montgomery (2009), o processo de planejamento estatístico de experimentos é necessário para garantir que os dados apropriados sejam analisados e para que estes resultem em conclusões válidas e objetivas.

Por meio da análise estatística dos dados é possível observar quais fatores têm maior influência na resposta de interesse. Uma vez escolhidos os fatores de influência do experimento, deve-se selecionar o intervalo no qual esses fatores serão variados. Inicialmente a região de interesse deve ser ampla, e, à medida em que se aprende quais variáveis tem maior influência e quais níveis produzem os melhores resultados, a tendência é que a região de interesse se torne mais restrita (MONTGOMERY, 2009, P. 15).

A forma pela qual se planeja e se conduz o experimento é chamada de estratégia de experimentação. No contexto de metamodelagem, pode-se afirmar que a precisão dos metamodelos

é maior quanto menor for a distância entre o ponto de previsão e um ponto já avaliado da função objetivo. Assim, para que se obtenha, inicialmente, um grau uniforme de precisão através espaço de projeto, é necessário que os pontos avaliados estejam uniformemente distribuídos. Um plano de experimentos que possua essa característica é chamado de preenchedor de espaço (*space-filling*), como o plano fatorial e o hipercubo latino, quando otimizado para tal fim (FORRESTER, SÓBESTER E KEANE, 2008, P. 13).

O plano fatorial é aquele no qual, para cada experimento, todas as combinações possíveis dos níveis dos fatores são investigadas. Este tipo de distribuição é utilizado em experimentos que envolvem diversos fatores e nos quais é necessário o estudo do efeito conjunto desses fatores na resposta. A Figura 3.3 mostra a distribuição de um arranjo fatorial de dois fatores, A e B, a quatro níveis cada, também chamada de distribuição fatorial  $4^2$ .

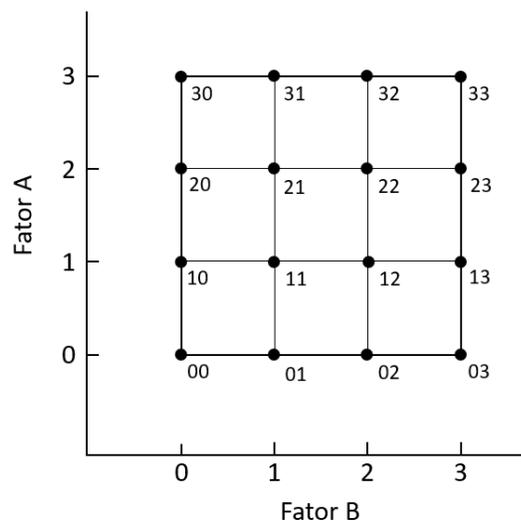


Figura 3.3 – Distribuição fatorial  $4^2$

Pode-se perceber que, quando projetado nos eixos, o conjunto de pontos se sobreporá. Forrester, Sóbester e Keane (2008) sugerem que a amostragem de variáveis pode ser melhorada uma vez que essas projeções se tornem o mais uniformes possível. Nesse sentido, o plano de experimentos por hipercubo latino (LHS) busca a estratificação do conjunto de amostragem em todas as suas dimensões.

A construção de um hipercubo latino se assemelha à divisão do espaço de projeto em hipercubos de mesmo tamanho e se acrescentando um ponto de avaliação em cada hipercubo, de

forma que, partindo de cada hipercubo preenchido, se consiga sair do espaço de projeto por qualquer direção paralela aos eixos sem encontrar outro hipercubo preenchido (FORRESTER, SÓBESTER E KEANE, 2008, P. 16). A figura 3.4 ilustra essa ideia para três dimensões. É válido observar que cada linha e cada coluna da divisão contém exatamente um único ponto.

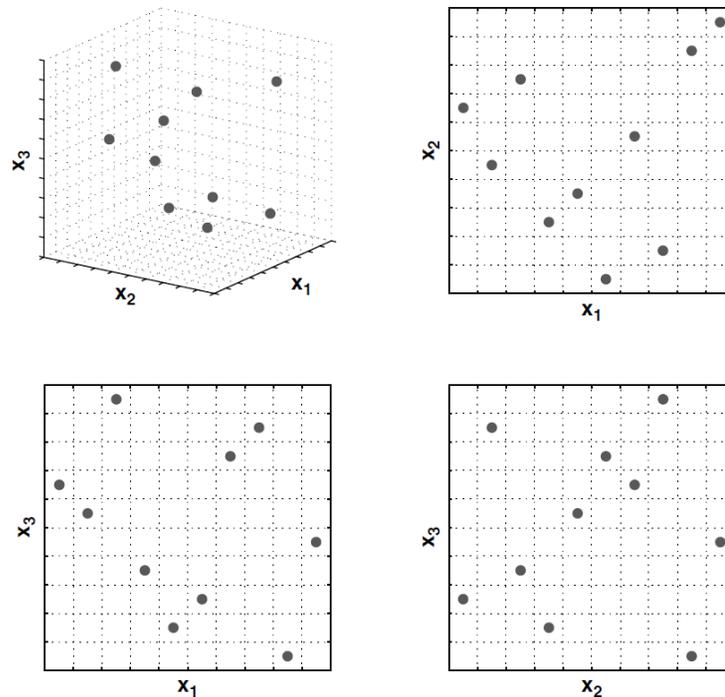


Figura 3.4 – Hipercubo latino em três dimensões composto por dez pontos.

Fonte: Forrester, Sóbester e Keane, 2008, p. 16.

A satisfação do critério de estratificação unicamente não garante que os pontos seguirão a tendência de preencher todo o espaço de projeto. Uma distribuição com pontos posicionados na diagonal principal do espaço de projeto satisfaria o critério de estratificação, por exemplo, porém, não é difícil perceber que essa distribuição não preencheria o espaço de projeto uniformemente.

Forrester, Sóbester e Keane (2008) sugerem o uso de medidas de uniformidade de distribuição para garantir o preenchimento, como o critério de qualidade de amostra desenvolvido por Morris e Mitchell (1995). Nesse caso, a estratificação da distribuição por LHS é determinada por meio da otimização do critério de Morris e Mitchell (1995) para a aplicação determinada.

## 3.2. MÉTODOS DE APROXIMAÇÃO

Segundo Jones *et al.* (1998), quando um processo estocástico é ajustado a um conjunto de dados, essencialmente é feita a calibragem de um modelo que resume como a função se comporta tipicamente, e também o quanto a função tende a mudar à medida em que se move diferentes quantidades em cada coordenada. Portanto, a previsão feita pelo modelo é fundamentalmente um cálculo do valor da função que seja mais consistente com o comportamento típico estimado anteriormente.

Existem diferentes métodos e variações para a modelagem estocástica de funções. A seção a seguir traz os principais métodos derivados do Kriging, especificamente os métodos que utilizam do Kriging para metamodelagem multifidelidade, começando pela descrição da abordagem de Kriging clássico, que servirá de base para o entendimento e desenvolvimento dos métodos subsequentes.

O método de Co-Kriging foi uma das primeiras técnicas desenvolvidas para combinar dados de baixa e alta fidelidade, surgindo no contexto da geologia matemática na década de 1990. Posteriormente, o método de Kriging hierárquico foi proposto para simplificar a combinação dos níveis de fidelidade do Co-Kriging e eliminar a necessidade da matriz de covariância cruzada, que era uma de suas fraquezas. Desde então, os esforços se concentraram em melhorar a construção dos modelos de Kriging, como o aprimoramento por superfície de resposta clássica, proposto pelo Kriging hierárquico aprimorado. O presente trabalho contribui para esses esforços ao propor um método que aprimora o modelo de alta fidelidade por funções de base radial e corrige o modelo de baixa fidelidade com informações dos pontos de alta fidelidade. A modelagem matemática do método proposto é detalhada na seção 3.2.5.

### 3.2.1. Kriging

O processo de Kriging é muitas vezes referido à modelagem de uma função como realização de um processo estocástico. Já Jones (2001) utiliza uma abordagem diferente na derivação das fórmulas utilizadas no Kriging. Suponha que se deseja fazer uma previsão,  $\hat{y}(\mathbf{x})$ , do valor de uma função  $y(\mathbf{x})$  em um ponto  $\mathbf{x}$  do domínio. Antes que a função seja avaliada em alguns dos pontos do domínio, haverá uma incerteza relacionada à esta estimativa, representada pela variável aleatória

$Y(\mathbf{x})$ , cujo valor pode variar dentro de uma distribuição normal de média  $\mu$  e variância  $\sigma^2$ . Essa ideia é representada na Figura 3.5 para uma função unidimensional. A região hachurada em cinza representa o intervalo de confiança de previsão dado pela distribuição normal para a variável  $Y(\mathbf{x})$ .

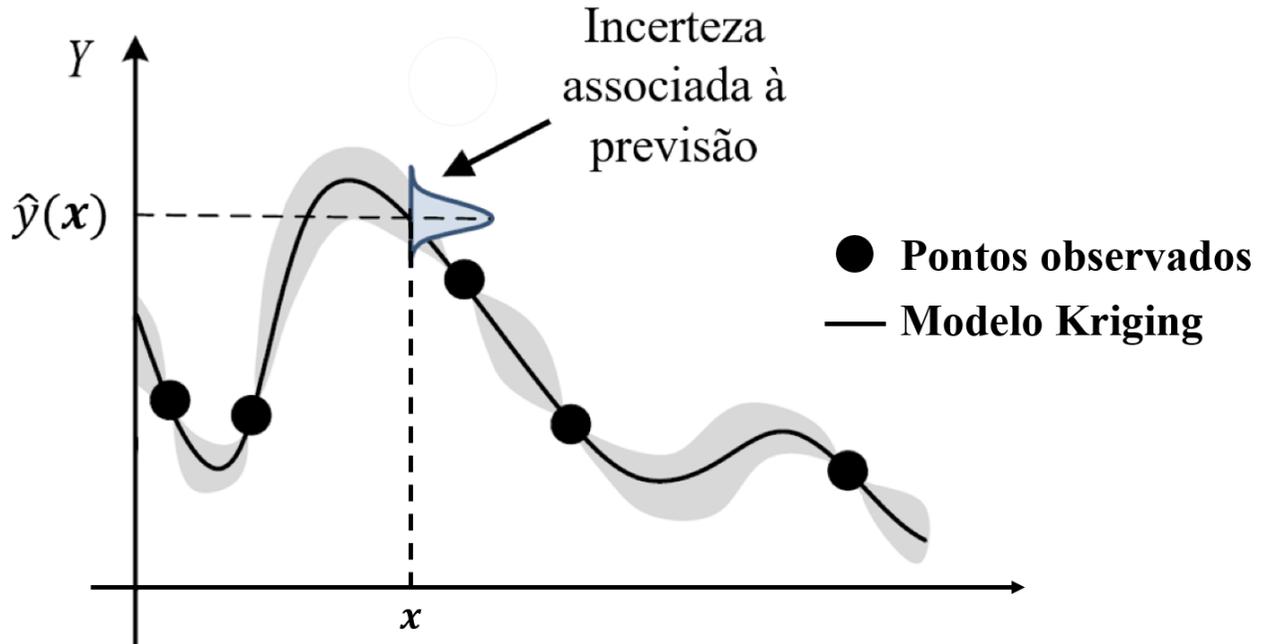


Figura 3.5 – Representação do modelo Kriging e incerteza associada.

Fonte: Adaptado de Zhou *et al.* (2020).

Estendendo o raciocínio agora para dois pontos  $\mathbf{x}_i$  e  $\mathbf{x}_j$ , haverá uma incerteza associada aos valores da função nesses pontos. Porém, se a função modelada for contínua, é esperado que os valores  $y(\mathbf{x}_i)$  e  $y(\mathbf{x}_j)$  se tornem próximos à medida que a distância entre os pontos  $\|\mathbf{x}_i - \mathbf{x}_j\|$  diminua. Em termos estatísticos, pode-se dizer que as variáveis aleatórias  $Y(\mathbf{x}_i)$  e  $Y(\mathbf{x}_j)$  serão altamente correlacionadas se  $\|\mathbf{x}_i - \mathbf{x}_j\|$  for pequena. Tal correlação é dada por

$$\text{corr}[Y(\mathbf{x}_i), Y(\mathbf{x}_j)] = \exp\left(-\sum_{l=1}^d \theta_l |x_{il} - x_{jl}|^{p_l}\right) \quad (3.1)$$

onde  $d$  é o número de variáveis de projeto e  $p_l$  define a suavidade da função na direção da coordenada  $l$ , podendo variar, tipicamente, entre 1 e 2. Para o desenvolvimento e implementação dos métodos abordados na seção 3.2, foi adotado um valor de 2 para  $p_l$ .

Observa-se que se  $\mathbf{x}_i = \mathbf{x}_j$ , a correlação assume o valor de 1 e que se  $\|\mathbf{x}_i - \mathbf{x}_j\| \rightarrow \infty$ , a correlação tende a zero. O hiper-parâmetro  $\theta_l$  determina quão rápido a correlação decai à medida que se caminha na direção  $l$ . Por exemplo, grandes valores de  $\theta_l$  são úteis na modelagem de funções que mudam de valor rapidamente para pequenas variações em  $l$ .

Seguindo o rigor matemático, a incerteza associada aos valores da função para  $n$  pontos pode ser representada pelo vetor aleatório

$$\mathbf{Y} = \begin{pmatrix} Y(\mathbf{x}_1) \\ \vdots \\ Y(\mathbf{x}_n) \end{pmatrix} \quad (3.2)$$

cuja média é igual a  $\mathbf{1}\mu$ , onde  $\mathbf{1}$  é um vetor unitário  $n \times 1$ , e cuja matriz de covariância é dada pela Eq. 3.3.  $\Psi$  é uma matriz de correlação  $n \times n$  cujos elementos  $(i, j)$  são dados pela Eq. 3.1.

$$\text{Cov}(\mathbf{Y}, \mathbf{Y}) = \sigma^2 \Psi \quad (3.3)$$

$$\Psi = \begin{pmatrix} \text{corr}[Y(\mathbf{x}_1), Y(\mathbf{x}_1)] & \cdots & \text{corr}[Y(\mathbf{x}_1), Y(\mathbf{x}_n)] \\ \vdots & \ddots & \vdots \\ \text{corr}[Y(\mathbf{x}_n), Y(\mathbf{x}_1)] & \cdots & \text{corr}[Y(\mathbf{x}_n), Y(\mathbf{x}_n)] \end{pmatrix} \quad (3.4)$$

A distribuição de  $\mathbf{Y}$  depende dos parâmetros  $\mu$ ,  $\sigma^2$ ,  $\theta_l$  e  $p_l$ , que são estimados com o objetivo de maximizar a função de verossimilhança  $L$  dos valores reais  $\mathbf{y}$  da função. O leitor é referido a Forrester *et al.* (2008) para o desenvolvimento das Eq. 3.6, 3.7, 3.8, 3.9 e 3.10.

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad (3.5)$$

$$L = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}|\Psi|^{\frac{1}{2}}} \exp \left[ -\frac{(\mathbf{y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \right] \quad (3.6)$$

A título de simplificação, utiliza-se o logaritmo natural da função de probabilidade, dado por

$$\ln(L) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln|\Psi| - \frac{(\mathbf{y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \quad (3.7)$$

Derivando a Eq. 3.7 em função de  $\sigma^2$  e de  $\mu$ , igualando-se a zero e resolvendo, é possível expressar os valores ótimos de  $\sigma^2$  e  $\mu$  como funções de  $\Psi$

$$\hat{\mu} = \frac{\mathbf{1}^T \Psi^{-1} \mathbf{y}}{\mathbf{1}^T \Psi^{-1} \mathbf{1}} \quad (3.8)$$

$$\hat{\sigma}^2 = \frac{(\mathbf{y} - \mathbf{1}\hat{\mu})^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{n} \quad (3.9)$$

Substituindo as Eq. 3.8 e 3.9 na Eq. 3.7, obtém-se a chamada função de verossimilhança logarítmica concentrada

$$\ln(L) \approx -\frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln|\Psi| \quad (3.10)$$

Percebe-se que a Eq. 3.10 depende apenas da matriz de correlações  $\Psi$  e, conseqüentemente, dos parâmetros de correlação  $\theta_l$  e  $p_l$ . A estimativa desses parâmetros pode ser feita por meio de algum método de busca global, tal como algoritmo genético, uma vez que a função de probabilidade logarítmica concentrada é rápida de se calcular (FORRESTER, SÓBESTER E KEANE, 2008).

Por fim, é possível fazer estimativas para novos pontos  $\mathbf{x}^*$ , desde que seja assumido que as novas estimativas sejam consistentes com o padrão de variação observado dos dados da função.

Assim, a previsão  $\hat{y}$  é encontrada a fim de maximizar a probabilidade dos dados da amostra e da estimativa em si, dados os parâmetros de correlação já determinados.

De acordo com Jones (2001), seja  $\tilde{\mathbf{y}} = (\mathbf{y}^T \hat{y})^T$  o vetor de valores da função quando acrescido da pseudo-observação  $(\mathbf{x}^*, \hat{y})$  e  $\boldsymbol{\psi}$  a matriz coluna de correlações entre os dados observados  $Y(\mathbf{x}_n)$  e a nova estimativa  $Y(\mathbf{x}^*)$

$$\boldsymbol{\psi} = \begin{pmatrix} \text{corr}[Y(\mathbf{x}^*), Y(\mathbf{x}_1)] \\ \vdots \\ \text{corr}[Y(\mathbf{x}^*), Y(\mathbf{x}_n)] \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \quad (3.11)$$

A matriz de correlação para o conjunto de dados acrescido, denotada por  $\tilde{\boldsymbol{\Psi}}$  é dada por

$$\tilde{\boldsymbol{\Psi}} = \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{pmatrix} \quad (3.12)$$

O último elemento da matriz  $\tilde{\boldsymbol{\Psi}}$  é igual a 1 por ser a continuação da diagonal principal da matriz de correlações  $\boldsymbol{\Psi}$ , cujos elementos são unitários, uma vez que representam a correlação dos pontos em relação a eles mesmos.

Voltando à Eq. 3.7, nota-se que apenas o último termo depende do vetor  $\mathbf{y}$  e, conseqüentemente, de  $\hat{y}$ . Substituindo as expressões para  $\tilde{\mathbf{y}}$  e  $\tilde{\boldsymbol{\Psi}}$  nesse último termo

$$\ln(L) \approx \frac{-\begin{pmatrix} \mathbf{y} - \mathbf{1}\hat{\mu} \\ \hat{y} - \hat{\mu} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{y} - \mathbf{1}\hat{\mu} \\ \hat{y} - \hat{\mu} \end{pmatrix}}{2\hat{\sigma}^2} \quad (3.13)$$

Para maximizar a Eq. 3.13 é necessário encontrar a inversa da matriz  $\tilde{\boldsymbol{\Psi}}$  utilizando o método de inversa particionada de Theil (1971)

$$\tilde{\boldsymbol{\Psi}}^{-1} = \begin{pmatrix} \boldsymbol{\Psi}^{-1} + \boldsymbol{\Psi}^{-1}\boldsymbol{\psi}(1 - \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1} & -\boldsymbol{\Psi}^{-1}\boldsymbol{\psi}(1 - \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi})^{-1} \\ -(1 - \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1} & (1 - \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi})^{-1} \end{pmatrix} \quad (3.14)$$

Substituindo a Eq. 3.14 na Eq. 3.13 e removendo os termos que não contém  $\hat{y}$

$$\ln(L) \approx \left( \frac{-1}{2\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu})^2 + \left( \frac{\boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu}) \quad (3.15)$$

Percebe-se que a Eq. 3.15 é uma função quadrática de  $\hat{y}$ . Logo, o valor de  $\hat{y}$  que maximiza a probabilidade é encontrada derivando a Eq. 3.15 e igualando a zero

$$\left( \frac{-1}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu}) + \left( \frac{\boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi})} \right) = 0 \quad (3.16)$$

Resolvendo para  $\hat{y}$ , obtém-se a fórmula para estimar o valor da função em um novo ponto  $\mathbf{x}^*$

$$\hat{y}(\mathbf{x}^*) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}) \quad (3.17)$$

Jones (2001) conclui ainda que, se  $\varphi(\mathbf{z})$  for uma função de base do Kriging dada pela Eq. 3.1, então o elemento  $i$  da matriz coluna  $\boldsymbol{\psi}$  é apenas  $\varphi(\mathbf{x}^* - \mathbf{x}_i)$  e ainda, se  $\hat{\mu} = a$  e  $b_i$  for o elemento  $i$  de  $\boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})$ , então a estimativa do processo de Kriging pode ser escrita como

$$\hat{y}(\mathbf{x}^*) = a + \sum_{i=1}^n b_i \varphi(\mathbf{x}^* - \mathbf{x}_i) \quad (3.18)$$

que nada mais é do que uma combinação linear de funções de base e de termos polinomiais, representados na Eq. 3.18 apenas pela constante  $a$  (JONES, 2001).

Os processos Gaussianos, como o Kriging, permitem ainda o cálculo de uma estimativa de erro associada à previsão do modelo. Essa estimativa pode ser utilizada como guia para processos de aquisição de novos pontos de amostragem. O erro quadrático médio de uma previsão baseada em processo Gaussiano é dado por

$$\hat{s}^2(\mathbf{x}^*) = \sigma^2 \left[ 1 - \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi} + \frac{1 - \mathbf{1}^T \boldsymbol{\Psi}^{-1} \boldsymbol{\psi}}{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{1}} \right] \quad (3.19)$$

### 3.2.2. Co-Kriging

Tomando uma função cujo valores podem ser obtidos por dois métodos com custos computacionais diferentes, e sejam os valores obtidos pelo método de maior custo computacional, ou caro (*expensive*),  $\mathbf{y}_e$  para os pontos  $\mathbf{x}_e$ , e, de menor custo computacional, ou barato (*cheap*),  $\mathbf{y}_c$  para os pontos  $\mathbf{x}_c$ . O conjunto de pontos formado pela concatenação de ambos os métodos é

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_e \end{pmatrix} = (\mathbf{x}_{c,1}, \dots, \mathbf{x}_{c,n_c}, \mathbf{x}_{e,1}, \dots, \mathbf{x}_{e,n_e})^T \quad (3.20)$$

Assim como no Kriging, a incerteza associada ao valor estimado para os pontos  $\mathbf{X}$  é representada pelo vetor de variáveis aleatórias  $\mathbf{Y}$

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_c(\mathbf{x}_c) \\ \mathbf{Y}_e(\mathbf{x}_e) \end{pmatrix} = (Y_c(x_{c,1}), \dots, Y_c(x_{c,n_c}), Y_e(x_{e,1}), \dots, Y_e(x_{e,n_e}))^T \quad (3.21)$$

O processo de construção de uma aproximação da função que correlacione os dados obtidos pelos dois métodos, apresentados nas Eq. 3.20 e 3.21, é chamado de co-Kriging.

Os processos Gaussianos  $Z_c(\cdot)$  e  $Z_e(\cdot)$  representam as características locais dos métodos caro e barato, respectivamente.

$$Z_e(\mathbf{x}) = \rho Z_c(\mathbf{x}) + Z_d(\mathbf{x}) \quad (3.22)$$

Segundo Forrester *et al.* (2008), a Eq. 3.22 basicamente aproxima o processo caro como o processo barato, multiplicado por um fator de escala  $\rho$ , somado ao processo  $Z_d(\cdot)$  que representa a diferença entre  $\rho Z_c(\mathbf{x})$  e  $Z_e(\mathbf{x})$ .

A partir disso, o co-Kriging se assemelha ao Kriging no sentido da construção da matriz de correlação, composta pelas covariâncias entre os dados. Porém, como existem dados obtidos por dois métodos diferentes, além da covariância entre os pontos de um mesmo conjunto de dados, segundo a Eq. 3.3, tem-se a covariância entre os dados dos dois conjuntos, caro e barato.

$$\begin{aligned}
cov\{\mathbf{Y}_c(\mathbf{x}_c), \mathbf{Y}_c(\mathbf{x}_c)\} &= cov\{Z_c(\mathbf{x}_c), Z_c(\mathbf{x}_c)\} = \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c) \\
cov\{\mathbf{Y}_e(\mathbf{x}_e), \mathbf{Y}_c(\mathbf{x}_c)\} &= cov\{\rho Z_c(\mathbf{x}_c) + Z_d(\mathbf{x}_c), Z_c(\mathbf{x}_e)\} = \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_e) \\
cov\{\mathbf{Y}_e(\mathbf{x}_e), \mathbf{Y}_e(\mathbf{x}_e)\} &= cov\{\rho Z_c(\mathbf{x}_e) + Z_d(\mathbf{x}_e), \rho Z_c(\mathbf{x}_e) + Z_d(\mathbf{x}_e)\} \\
&= \rho^2 cov\{Z_c(\mathbf{x}_e), Z_c(\mathbf{x}_e)\} + cov\{Z_d(\mathbf{x}_e), Z_d(\mathbf{x}_e)\} \\
&= \rho^2 \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_e, \mathbf{x}_e) + \sigma_d^2 \boldsymbol{\Psi}_d(\mathbf{x}_e, \mathbf{x}_e)
\end{aligned} \tag{3.23}$$

A notação  $\boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_e)$  por exemplo denota a matriz de correlações da forma  $\boldsymbol{\psi}_c$  entre os dados  $\mathbf{x}_e$  e  $\mathbf{x}_c$ . Assim, a matriz de covariância completa assume a forma

$$\mathbf{C} = \begin{pmatrix} \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c) & \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_e) \\ \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_e) & \rho^2 \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{x}_e, \mathbf{x}_e) + \sigma_d^2 \boldsymbol{\Psi}_d(\mathbf{x}_e, \mathbf{x}_e) \end{pmatrix} \tag{3.24}$$

As correlações possuem a mesma forma da Eq. 3.1, porém são duas matrizes de correlação,  $\boldsymbol{\psi}_c$  e  $\boldsymbol{\psi}_e$ . Logo, tem-se dois vetores de parâmetros de correlação para cada matriz de correlação, são eles  $\boldsymbol{\theta}_c$ ,  $\boldsymbol{\theta}_d$ ,  $\boldsymbol{p}_c$  e  $\boldsymbol{p}_d$  além do fator de escala,  $\rho$ . Assim como no processo de Kriging, os parâmetros de correlação da função barata podem ser encontrados pela maximização da função de verossimilhança logarítmica concentrada, Eq. 3.27. Note a semelhança entre as Eq. 3.25 e 3.26 e as Eq. 3.8 e 3.9. Bem como a semelhança entre as Eq. 3.27 e 3.10.

$$\hat{\mu}_c = \frac{\mathbf{1}^T \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c)^{-1} \mathbf{y}_c}{\mathbf{1}^T \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c)^{-1} \mathbf{1}} \tag{3.25}$$

$$\hat{\sigma}_c^2 = \frac{(\mathbf{y}_c - \mathbf{1} \hat{\mu}_c)^T \boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c)^{-1} (\mathbf{y}_c - \mathbf{1} \hat{\mu}_c)}{n_c} \tag{3.26}$$

$$\ln(L) \approx -\frac{n_c}{2} \ln(\hat{\sigma}_c^2) - \frac{1}{2} \ln |\boldsymbol{\Psi}_c(\mathbf{x}_c, \mathbf{x}_c)| \tag{3.27}$$

Para a estimativa dos parâmetros relacionados ao modelo da diferença,  $\hat{\mu}_d$ ,  $\hat{\sigma}_d^2$ ,  $\boldsymbol{\theta}_d$  e  $\boldsymbol{p}_d$  e do fator de escala  $\rho$ , primeiro define-se a diferença entre os valores da função cara e os valores obtidos pela função barata nos mesmos pontos avaliados pela função cara

$$\mathbf{y}_e = \rho \mathbf{y}_c(\mathbf{x}_e) + \mathbf{d} \quad (3.28)$$

Feito isso, os parâmetros podem ser determinados pelo mesmo processo de maximização da probabilidade logarítmica utilizado anteriormente, onde

$$\hat{\mu}_d = \frac{\mathbf{1}^T \Psi_d(\mathbf{x}_e, \mathbf{x}_e)^{-1} \mathbf{d}}{\mathbf{1}^T \Psi_d(\mathbf{x}_e, \mathbf{x}_e)^{-1} \mathbf{1}} \quad (3.29)$$

$$\hat{\sigma}_d^2 = \frac{(\mathbf{d} - \mathbf{1} \hat{\mu}_d)^T \Psi_d(\mathbf{x}_e, \mathbf{x}_e)^{-1} (\mathbf{d} - \mathbf{1} \hat{\mu}_d)}{n_e} \quad (3.30)$$

e

$$\ln(L) \approx -\frac{n_e}{2} \ln(\hat{\sigma}_d^2) - \frac{1}{2} \ln |\Psi_d(\mathbf{x}_e, \mathbf{x}_e)| \quad (3.31)$$

Por fim, com os parâmetros de correlação estimados, a estimativa do valor da função feita pelo processo de co-Kriging é dada por

$$\hat{y}_e(\mathbf{x}^*) = \hat{\mu} + \mathbf{c}^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{1} \hat{\mu}) \quad (3.32)$$

onde

$$\mathbf{c} = \begin{pmatrix} \hat{\rho} \hat{\sigma}_c^2 \psi_c(\mathbf{x}_c, \mathbf{x}^*) \\ \hat{\rho}^2 \hat{\sigma}_c^2 \psi_c(\mathbf{x}_e, \mathbf{x}^*) + \hat{\sigma}_d^2 \psi_d(\mathbf{x}_e, \mathbf{x}^*) \end{pmatrix} \quad (3.33)$$

denota o vetor de covariância entre os pontos de  $\mathbf{X}$ , definido na Eq. 3.20, e o novo ponto  $\mathbf{x}^*$ . Uma das vantagens dos processos de Kriging e co-Kriging está associada à incerteza do processo: além de estimar o valor da função em um novo ponto, é possível estimar o erro quadrático médio relacionado à essa estimativa

$$\hat{s}^2(\mathbf{x}^*) = \hat{\rho}^2 \hat{\sigma}_c^2 + \hat{\sigma}_d^2 - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c} + \frac{1 - \mathbf{1}^T \mathbf{C}^{-1} \mathbf{c}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \quad (3.34)$$

Usando a Eq. 3.34 é possível a formulação de critérios de busca por novos pontos em um processo de melhoramento da aproximação feita por Kriging ou co-Kriging, ou até a formulação de processos de otimização utilizando esses critérios de busca.

### 3.2.3. Kriging hierárquico (HK)

A construção do co-Kriging se baseia na concatenação dos dados de baixa e de alta fidelidade e se faz necessário o cálculo das correlações cruzadas dos níveis de fidelidade. Han e Görtz (2012) propuseram um método que fosse tão robusto quanto o co-Kriging, porém mais simples. O método foi chamado de Kriging hierárquico. No HK, o modelo de baixa fidelidade é usado para capturar a tendência geral do modelo de alta fidelidade, enquanto os pontos de alta fidelidade são usados para calibração (JIANG, ZHOU E SHAO, P. 80, 2020).

O nome hierárquico se dá pela forma como o modelo é construído no método. Primeiro é feita a construção do metamodelo para a função de baixa fidelidade. Assumindo um processo aleatório correspondente à função de baixa fidelidade  $y_{lf}$

$$Y_{lf}(\mathbf{x}) = \beta_{0,lf} + Z_{lf}(\mathbf{x}) \quad (3.35)$$

onde  $\beta_{0,lf}$  é uma constante até então desconhecida e  $Z_{lf}(\mathbf{x})$ , um processo randômico estacionário. Seguindo o mesmo raciocínio desenvolvido na seção 3.2.1, após a construção do modelo Kriging com base no conjunto de dados de baixa fidelidade, a previsão do modelo para novo ponto  $\mathbf{x}^*$  se dá por

$$\hat{y}_{lf}(\mathbf{x}^*) = \beta_{0,lf} + \mathbf{r}_{lf}^T \mathbf{R}_{lf}^{-1} (\mathbf{y}_{s,lf} - \mathbf{1} \beta_{0,lf}) \quad (3.36)$$

onde  $\beta_{0,lf} = (\mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{y}_{s,lf}$ ;  $\mathbf{1}$  é um vetor coluna unitário (todos componentes iguais a 1).  $\mathbf{R}_{lf} \in \mathbb{R}^{n_{lf} \times n_{lf}}$  é a matriz de correlações entre os pontos de baixa fidelidade avaliados, onde  $n_{lf}$  é o número de pontos do conjunto de baixa fidelidade

$$\mathbf{R}_{lf} = \begin{pmatrix} \text{corr}[Y(\mathbf{x}_1), Y(\mathbf{x}_1)] & \cdots & \text{corr}[Y(\mathbf{x}_1), Y(\mathbf{x}_{n_{lf}})] \\ \vdots & \ddots & \vdots \\ \text{corr}[Y(\mathbf{x}_{n_{lf}}), Y(\mathbf{x}_1)] & \cdots & \text{corr}[Y(\mathbf{x}_{n_{lf}}), Y(\mathbf{x}_{n_{lf}})] \end{pmatrix} \quad (3.37)$$

$\mathbf{r}_{lf}$  é o vetor de correlações entre os pontos de baixa fidelidade avaliados e o novo ponto  $\mathbf{x}^*$ .

$$\mathbf{r}_{lf} = \begin{pmatrix} \text{corr}[Y(\mathbf{x}^*), Y(\mathbf{x}_1)] \\ \vdots \\ \text{corr}[Y(\mathbf{x}^*), Y(\mathbf{x}_{n_{lf}})] \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{n_{lf}} \end{pmatrix} \quad (3.38)$$

Uma vez construído o modelo de baixa fidelidade, assume-se, também, o processo aleatório correspondente à função de alta fidelidade

$$Y(\mathbf{x}) = \beta_0 \hat{y}_{lf} + Z(\mathbf{x}) \quad (3.39)$$

onde  $\hat{y}_{lf}$  é a previsão do metamodelo de baixa fidelidade, dada pela Equação 3.36, multiplicada por um fator de escala constante  $\beta_0$ , que indica a influência do modelo de baixa fidelidade nas previsões do modelo de alta fidelidade. Juntos, representam a função tendência global, que diz respeito ao valor médio da função.  $Z(\mathbf{x})$  é um processo randômico de média zero e covariância

$$\text{Cov}[Z(\mathbf{x}), Z(\mathbf{x}')] = \sigma^2 R(\mathbf{x}, \mathbf{x}') \quad (3.40)$$

$\sigma^2$  é a variância do processo  $Z(\mathbf{x})$  e  $R(\mathbf{x}, \mathbf{x}')$  é a função de correlação dada pela Eq. 3.1, porém para os pontos de alta fidelidade, que depende apenas da distância Euclidiana entre os dois pontos  $\mathbf{x}$  e  $\mathbf{x}'$ . Analogamente às Equações 3.17 e 3.36, a previsão do modelo HK é dada por

$$\hat{y}_{HK}(\mathbf{x}^*) = \beta_0 \hat{y}_{lf} + \mathbf{r}^T \mathbf{R}^{-1} (\mathbf{y}_{s,hf} - \beta_0 \mathbf{F}) \quad (3.41)$$

onde  $\beta_0 = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_{s,hf}$  é um fator de escala que indica o quanto as funções de baixa e alta fidelidade estão correlacionadas entre si.  $\mathbf{R}$  e  $\mathbf{r}$  possuem a mesma definição explorada nas Equações 3.4 e 3.37; e 3.11 e 3.38, respectivamente.  $\mathbf{R} \in \mathbb{R}^{n_{hf} \times n_{hf}}$  onde  $n_{hf}$  é o número de pontos do conjunto de alta fidelidade. O vetor coluna  $\mathbf{F}$  é composto pelas previsões do modelo de baixa fidelidade para os pontos de alta fidelidade avaliados  $\mathbf{x}^{(n_{hf})}$ .

Assim como na seção 3.2.1, o método HK possibilita a estimativa do erro associado à previsão do metamodelo. Segundo Han e Görtz (2012), a inclusão explícita da previsão do metamodelo de baixa fidelidade,  $\hat{y}_{lf}(\mathbf{x}^*)$ , no cálculo do erro quadrático médio possibilita uma melhor estimativa quando comparada com o Kriging clássico.

$$\begin{aligned} \hat{s}_{HK}^2(\mathbf{x}^*) = \sigma^2 & \left[ 1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \right. \\ & \left. + [\mathbf{r}^T \mathbf{R}^{-1} \mathbf{F} - \hat{y}_{lf}(\mathbf{x}^*)] (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} [\mathbf{r}^T \mathbf{R}^{-1} \mathbf{F} - \hat{y}_{lf}(\mathbf{x}^*)]^T \right] \end{aligned} \quad (3.42)$$

### 3.2.4. Kriging hierárquico aprimorado (IHK)

No HK, o fator de escala  $\beta_0$  é considerado constante, assim, a relação de proporção entre baixa e alta fidelidade é assumida como igual em todo o espaço de projeto. Na realidade, porém, a relação entre as fidelidades pode variar em diferentes pontos. Além disso, Jiang, Zhou e Shao (2020) argumentam que  $\beta_0$  é calculado com base apenas nos conjuntos iniciais de pontos. Dessa forma, Hu *et al.* (2016) propuseram um aprimoramento no método HK, chamado de Kriging hierárquico aprimorado (IHK).

Analogamente ao desenvolvimento do HK, assumindo um processo aleatório para a função de alta fidelidade

$$Y(\mathbf{x}) = h(\mathbf{x}) \hat{y}_{lf} + Z(\mathbf{x}) \quad (3.43)$$

Comparando-se as Equações 3.39 e 3.43, percebe-se que o fator  $\beta_0$  é substituído pelo fator variável  $h(\mathbf{x})$ , que corresponde a um modelo de superfície de resposta polinomial clássica (quadrática), vide Anexo B.

$$h(\mathbf{x}) = \beta_{0,h} + \sum_{i=1}^d \beta_i x_i + \sum_{1 \leq j \leq k \leq d} \beta_{jk} x_j x_k \quad (3.44)$$

onde  $d$  denota a dimensão do espaço de projeto;  $x_i$ ,  $x_j$ , e  $x_k$  correspondem às variáveis de projeto (componentes de  $\mathbf{x}$ ) e  $\beta_i$  e  $\beta_{jk}$  os coeficientes associados a elas. Segundo Hu *et al.* (2016) a capacidade de variação de  $h(\mathbf{x})$  com o ponto de previsão possibilita uma melhor expressão da relação entre baixa e alta fidelidades.

Para o caso de duas variáveis, por exemplo, o polinômio  $h(\mathbf{x})$  toma a seguinte forma

$$h(\mathbf{x}) = \beta_{0,h} + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 \quad (3.45)$$

Substituindo 3.45 em 3.43 e reescrevendo na forma compacta

$$Y_{IHK}(\mathbf{x}^*) = \mathbf{f}\boldsymbol{\beta} + Z(\mathbf{x}^*) \quad (3.46)$$

onde

$$\boldsymbol{\beta} = [\beta_{0,h} \ \beta_1 \ \cdots \ \beta_6] \quad (3.47)$$

$$\begin{aligned} \mathbf{f} &= [f_1 \ f_2 \ \cdots \ f_6] \\ &= [\hat{y}_{lf}(\mathbf{x}^*) \ x_1 \hat{y}_{lf}(\mathbf{x}^*) \ x_2 \hat{y}_{lf}(\mathbf{x}^*) \ x_1^2 \hat{y}_{lf}(\mathbf{x}^*) \ x_2^2 \hat{y}_{lf}(\mathbf{x}^*) \ x_1 x_2 \hat{y}_{lf}(\mathbf{x}^*)] \end{aligned} \quad (3.48)$$

A previsão segundo o método IHK segue o mesmo princípio descrito anteriormente e é dada por

$$\hat{y}_{IHK}(\mathbf{x}^*) = \mathbf{f}\boldsymbol{\beta} + \mathbf{r}^T \mathbf{R}^{-1}(\mathbf{y}_{s,hf} - \mathbf{F}\boldsymbol{\beta}) \quad (3.49)$$

onde

$$\mathbf{F} = [f(\mathbf{x}_{hf}^{(1)}), f(\mathbf{x}_{hf}^{(2)}), \dots, f(\mathbf{x}_{hf}^{(n)})]^T \quad (3.50)$$

$$\boldsymbol{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_{S,hf} \quad (3.51)$$

e a estimativa de erro

$$\hat{\sigma}_{IHK}^2(\mathbf{x}^*) = \sigma^2 [1 + (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{r} - \mathbf{f}^T)^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{r} - \mathbf{f}^T) - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}] \quad (3.52)$$

### 3.2.5. Kriging hierárquico aprimorado por funções de base radial (FBR-HK)

No mesmo sentido do desenvolvimento do método IHK, e motivado pela ideia de possibilitar um melhor mapeamento da relação entre as respostas da função de baixa fidelidade e de alta fidelidade para funções de maior dimensionalidade e complexidade, o presente trabalho se propõe a desenvolver um método de Kriging hierárquico aprimorado por funções de base radial (FBR), nomeado FBR-HK.

Diferentemente do processo de construção do Kriging de baixa fidelidade utilizado pelos métodos de HK e IHK, para o FBR-HK, buscou-se também a possibilidade de aprimorar os pontos de baixa fidelidade por meio das observações existentes de alta fidelidade.

Sabe-se dos métodos anteriores que o modelo de Kriging de baixa fidelidade é construído segundo a Equação 3.53 abaixo, assumindo-se um processo aleatório para a função de baixa fidelidade

$$Y_{lf}(\mathbf{x}) = \beta_{0,lf} + Z_{lf}(\mathbf{x}) \quad (3.53)$$

É importante destacar que o termo  $\beta_{0,lf}$ , aqui, é uma constante que independe do ponto no espaço de projeto, calculada com base na matriz de correlações utilizando apenas os pontos de baixa fidelidade.

Uma das melhorias ao qual o presente método se propõe é o aprimoramento da construção do modelo de baixa fidelidade por meio da correção dos pontos de baixa pelos pontos de alta fidelidade. Esta correção é implementada no termo da média  $\beta_{0,lf}$ , que como mencionado anteriormente, é uma constante no Kriging clássico. O aprimoramento foi feito semelhantemente ao que é feito no IHK para a construção do metamodelo de alta fidelidade.

A correção foi desenvolvida concatenando os pontos de alta fidelidade aos de baixa fidelidade. Esse conjunto de pontos é utilizado para o cálculo de uma regressão polinomial de segunda ordem, cuja matriz de coeficientes substitui  $\beta_{0,lf}$ . Assim, este deixa de ser calculado por  $\beta_{0,lf} = (\mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{y}_{s,lf}$  e dá lugar a um polinômio de segunda ordem cujos coeficientes são calculados por mínimos quadrados ordinários, como mostra a Equação 3.54:

$$\boldsymbol{\beta}_{0,lf} = (\mathbf{f}^T \mathbf{f})^{-1} \mathbf{f}^T \mathbf{y}_{s,lf+h_f} \quad (3.54)$$

onde

$$\mathbf{f} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1m} & x_{11}^2 & \cdots & x_{1m}^2 & x_{11}x_{12} & \cdots & x_{1(m-1)}x_{1m} \\ \vdots & & & & & & \ddots & & & \\ 1 & x_{n1} & \cdots & x_{nm} & x_{n1}^2 & \cdots & x_{nm}^2 & x_{n1}x_{n2} & \cdots & x_{n(m-1)}x_{nm} \end{pmatrix} \quad (3.55)$$

em que  $m$  é o número de variáveis de projeto e  $n$  é o número de observações, que para o método é igual a  $n = n_{lf} + n_{hf}$ , isto é, representa a concatenação das observações de baixa e de alta fidelidade. A mesma abordagem é utilizada para a resposta  $\mathbf{y}_{s,lf+h_f}$ .

Por fim, a previsão do modelo para novo ponto  $\mathbf{x}^*$  se dá por

$$\hat{y}_{lf}(\mathbf{x}^*) = \mathbf{f} \boldsymbol{\beta}_{0,lf} + \mathbf{r}_{lf}^T \mathbf{R}_{lf}^{-1} (\mathbf{y}_{s,lf} - \mathbf{F} \boldsymbol{\beta}_{0,lf}) \quad (3.56)$$

onde  $\mathbf{f}$  é um vetor linha composto pela aplicação da Equação 3.55 para o ponto  $\mathbf{x}^*$  para o qual se quer calcular a previsão.  $\mathbf{R}_{lf}$  é a mesma matriz de correlações apresentada na seção 3.2.3 e é dada pela Equação 3.37. O mesmo vale para  $\mathbf{r}_{lf}$ , dada pela Equação 3.38. O vetor  $\mathbf{F}$  é composto pela aplicação da Equação 3.55 para os pontos  $\mathbf{x}_{lf}^{(n_{lf})}$ .

Dessa forma, ao invés de um metamodelo por Kriging construído puramente com pontos de baixa fidelidade, cuja média é uma constante para todo o espaço de projeto, tem-se um metamodelo cuja média varia de acordo com a posição no espaço de busca, além de introduzir informações dos pontos de alta fidelidade.

O modelo de baixa fidelidade construído é usado, hierarquicamente, na construção do modelo de alta fidelidade. Assumindo o mesmo processo aleatório para a função de alta fidelidade

$$Y(\mathbf{x}) = g(\mathbf{x})\hat{y}_{lf} + Z(\mathbf{x}) \quad (3.57)$$

A diferença entre as Equações 3.43 e 3.57 está presente na função que multiplica a previsão do modelo de baixa fidelidade, isto é, na função que faz o mapeamento da relação entre os níveis de fidelidade. A função  $g(\mathbf{x})$  é uma função de base radial, Eq. 3.71, cujo núcleo pode ser, por exemplo, uma função Gaussiana, multiquádrica ou inversa multiquádrica, Eq. 3.59(a), (b), (c), respectivamente. Para as aplicações presentes neste trabalho, adotou-se a função Gaussiana para o FBR-HK, porém o método foi desenvolvido para qualquer uma das funções de base radial abaixo.

$$g(\mathbf{x}) = \sum_{i=1}^{n_{hf}} w_i \varphi(\|\mathbf{x} - \mathbf{c}_i\|) \quad (3.58)$$

$$\varphi = \exp\left(-\frac{1}{\mu}\|\mathbf{x} - \mathbf{c}_i\|^2\right) \quad (3.59a)$$

$$\varphi = \sqrt{1 + \frac{1}{\mu}\|\mathbf{x} - \mathbf{c}_i\|^2} \quad (3.59b)$$

$$\varphi = \frac{1}{\sqrt{1 + \frac{1}{\mu}\|\mathbf{x} - \mathbf{c}_i\|^2}} \quad (3.59c)$$

onde  $n_{hf}$  é o número de centros de FBR, que é igual ao número de pontos de alta fidelidade avaliados,  $w_i$  são os pesos atribuídos a cada FBR,  $\mathbf{x}$  é o ponto no qual se pretende calcular o valor

da FBR e  $\mathbf{c}_i$  são os centros selecionados de cada FBR. Vale lembrar que  $\mu$ , no contexto das FBR, é uma medida da largura das FBR.

A previsão do modelo FBR-HK, então, é dada por

$$\hat{y}_{FBR-HK}(\mathbf{x}^*) = \mathbf{f}\boldsymbol{\beta} + \mathbf{r}^T \mathbf{R}^{-1}(\mathbf{y}_{s,hf} - \mathbf{F}\boldsymbol{\beta}) \quad (3.60)$$

onde, assim como no método IHK

$$\mathbf{f} = \varphi(\|\mathbf{x} - \mathbf{c}_i\|)\hat{y}_{lf}(\mathbf{x}) \quad (3.61)$$

$$\mathbf{F} = \mathbf{f}_i(\mathbf{x}_j); \quad j = n_{hf} \quad (3.62)$$

$$\boldsymbol{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_{s,hf} \quad (3.63)$$

e a estimativa de erro

$$\hat{s}_{FBR-HK}^2(\mathbf{x}^*) = \sigma^2 [1 + (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{r} - \mathbf{f}^T)^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{r} - \mathbf{f}^T) - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}] \quad (3.64)$$

### 3.2.6. Funções de aquisição de novos pontos de amostragem

A construção de modelos baseados em processos estocásticos foi feita, até então, a partir de uma amostragem inicial determinada por um planejamento de experimentos. A ideia por trás da otimização global eficiente é utilizar esses modelos como guias para onde buscar por novos pontos de amostragem. O processo de busca é guiado pela estimativa do erro quadrático médio associado à previsão do modelo.

Durante o processo de busca de novos pontos de amostragem, espera-se que o próximo ponto escolhido seja aquele que apresente alguma melhora em relação ao melhor valor avaliado até então. No caso de um problema de minimização de um objetivo, por exemplo, esse valor seria o valor mínimo atual da função.

Com base nessa abordagem, pode-se desenvolver uma função de probabilidade de melhoria associada ao processo. Segundo Forrester, Sóbester e Keane (2008), considerando  $\hat{y}(x)$  como a

realização de uma variável aleatória, é possível o cálculo de uma probabilidade de melhoria  $I = y_{min} - Y(\mathbf{x}^*)$  sobre  $y_{min}$

$$P[I(\mathbf{x}^*)] = \begin{cases} \frac{1}{\hat{\sigma}\sqrt{2\pi}} \int_{-\infty}^0 e^{\frac{-[I-\hat{y}(\mathbf{x}^*)]^2}{(2\sigma^2)}} dI & \text{se } s > 0 \\ 0 & \text{se } s = 0 \end{cases} \quad (3.65)$$

Tomando como exemplo uma função de uma dimensão, a Equação 3.65 pode ser representada e interpretada graficamente segundo a Figura 3.6, para um problema de minimização. A distribuição Gaussiana vertical de  $Y(\mathbf{x})$  com variância  $\sigma^2(\mathbf{x})$  centrada na previsão  $\hat{y}(\mathbf{x})$  é uma característica de processos estocásticos, e representa a incerteza na previsão  $\hat{y}(\mathbf{x})$ . A porção da distribuição localizada abaixo da linha pontilhada de  $y_{min}$  indica a possibilidade de melhoria em relação ao melhor valor avaliado até então, é a quantia integrada na Equação 3.65. Isto é, a probabilidade de melhoria é a área da distribuição Gaussiana abaixo do melhor valor avaliado até então (FORRESTER, SÓBESTER E KEANE, 2008, P. 89).

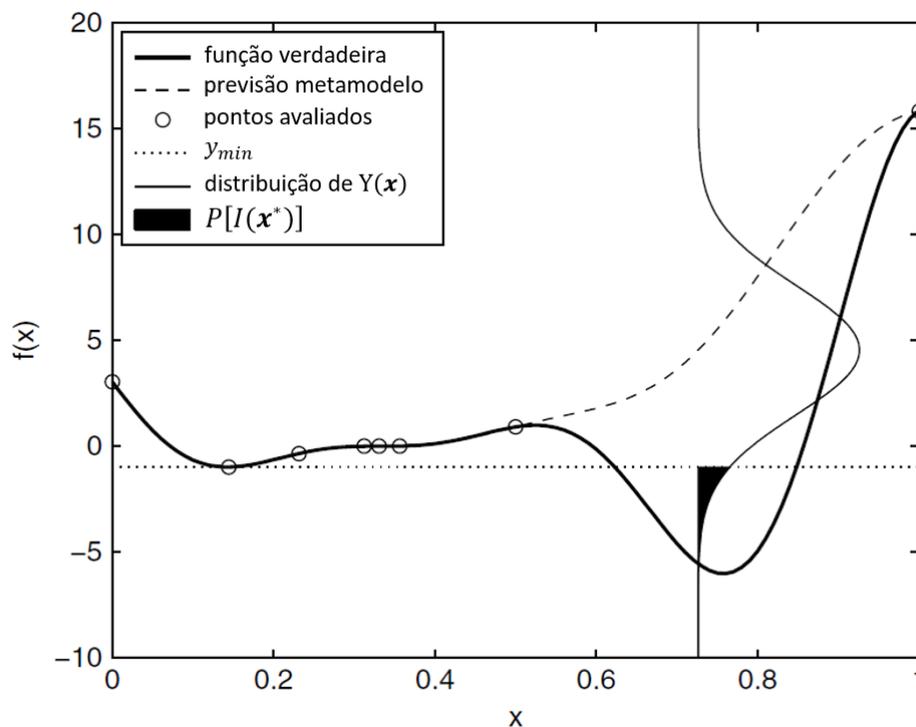


Figura 3.6 – Interpretação gráfica da probabilidade de melhoria.

Fonte: Adaptado de Forrester, Sóbester e Keane, 2008, p. 88.

Ao invés de apenas determinar a probabilidade de que haverá uma melhoria em um determinado local, é possível calcular a quantidade de melhoria esperada, dada a previsão  $\hat{y}(\mathbf{x}^*)$  e a variância  $\sigma^2(\mathbf{x}^*)$ . Essa grandeza é chamada de expectativa de melhoria (EI), e é dada por

$$E[I(\mathbf{x}^*)] = \begin{cases} (y_{min} - \hat{y}(\mathbf{x}^*))\Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x}^*)}{\sigma(\mathbf{x}^*)}\right) + \sigma\phi\left(\frac{y_{min} - \hat{y}(\mathbf{x}^*)}{\sigma(\mathbf{x}^*)}\right) & \text{se } s > 0 \\ 0 & \text{se } s = 0 \end{cases} \quad (3.66)$$

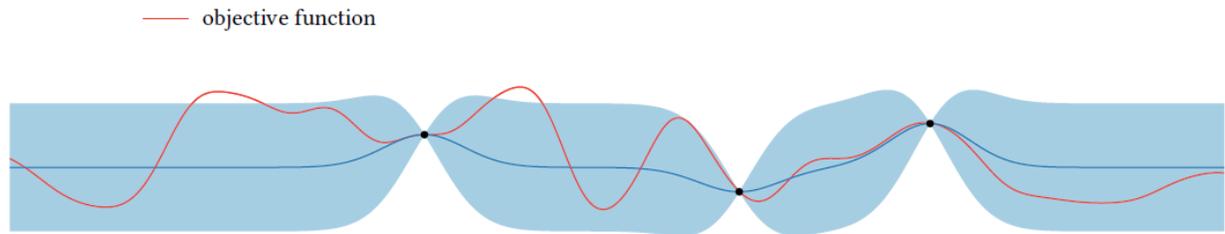
onde  $\Phi(\cdot)$  e  $\phi(\cdot)$  são a função de distribuição normal cumulativa e a função de densidade de probabilidade, respectivamente. Analogamente, pode-se dizer que o termo que contém  $\Phi(\cdot)$  responde pelo processo de busca mais global (*exploration*), e o termo que contém  $\phi(\cdot)$ , pela busca mais local (*exploitation*). Assim, a expectativa de melhoria busca um balanço entre as duas abordagens de adição de pontos baseando-se no julgamento da proximidade de uma região de interesse. Quando em uma região que possa conter um ponto de ótimo, maior peso é dado para *exploitation* a fim de melhor descrever a região em questão. Quando este não é o caso, maior peso é dado para *exploration* para buscar, globalmente, por outra região que possa conter o ótimo.

Essa dinâmica está bastante presente, por exemplo, em funções objetivo que contenham mínimos locais. A função de aquisição atribui maior peso para *exploitation* na região do mínimo local até que o valor de expectativa de melhoria passa a ser maior em outra região candidata a conter ótimo. Nesse momento, o peso maior passa a ser para a parcela de *exploration*, até que se chegue à nova região, onde o método retorna ao maior peso para *exploitation*.

É válido notar que, uma vez que tanto  $P[I(\mathbf{x}^*)]$  quanto  $E[I(\mathbf{x}^*)] = 0$  quando  $\sigma = 0$ , o processo de aquisição de novos pontos evita que um ponto já avaliado seja reavaliado, proporcionando a capacidade de que o processo não fique preso em mínimos locais.

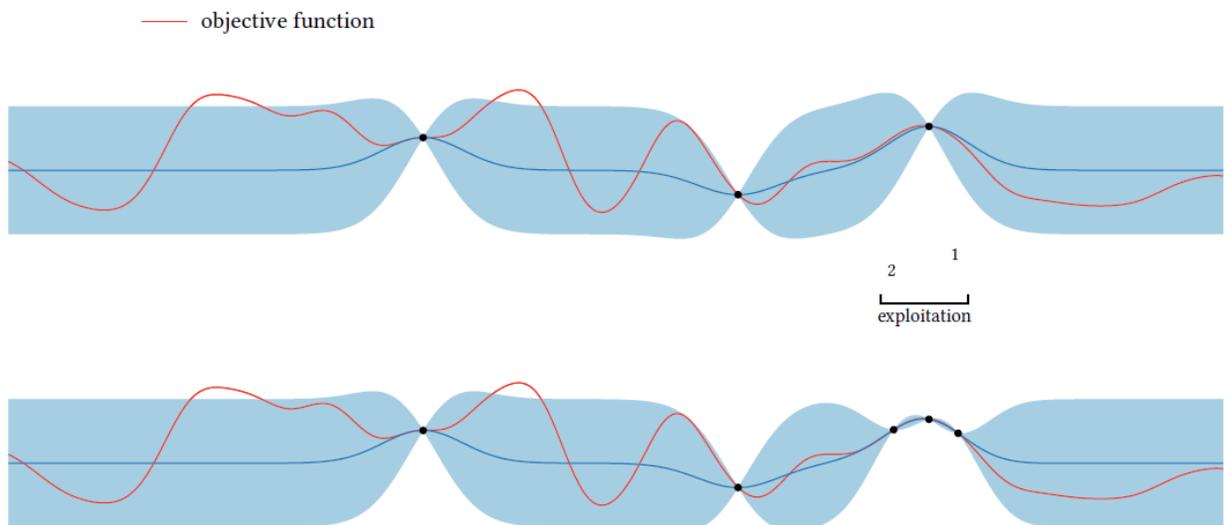
As Figuras 3.7(a), (b), (c) e (d) exemplificam um processo de busca pelo máximo que utiliza expectativa de melhoria como função de aquisição. A linha vermelha representa o comportamento verdadeiro da função objetivo, mas que é desconhecido para o processo.

O modelo é construído primeiramente com os pontos de amostragem iniciais e, assim, apresenta regiões de maior incerteza, representados pelos campos em azul mostrado na Figura 3.7(a). Esses campos serão as porções do espaço de projeto candidatas para a busca de novos pontos através da função de aquisição.



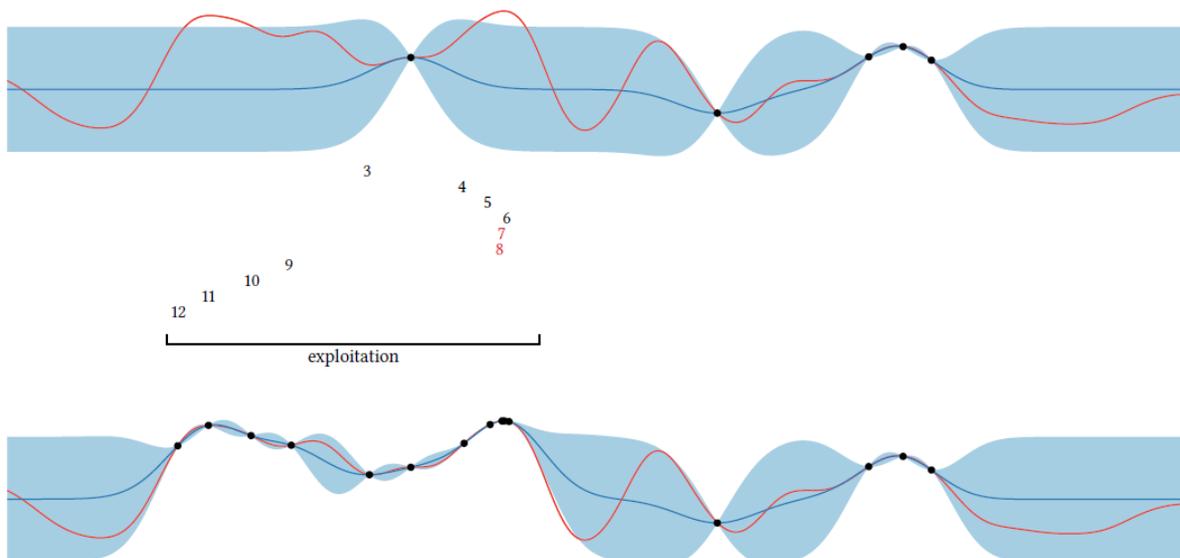
(a) Função objetivo e aproximação por metamodelo no estágio inicial.

O processo de busca inicia com um possível máximo, mas até então o método não tem informação se se trata de um máximo local ou global. Dessa forma, ocorre o processo de busca local em torno do candidato a máximo, como mostra a Figura 3.7(b).



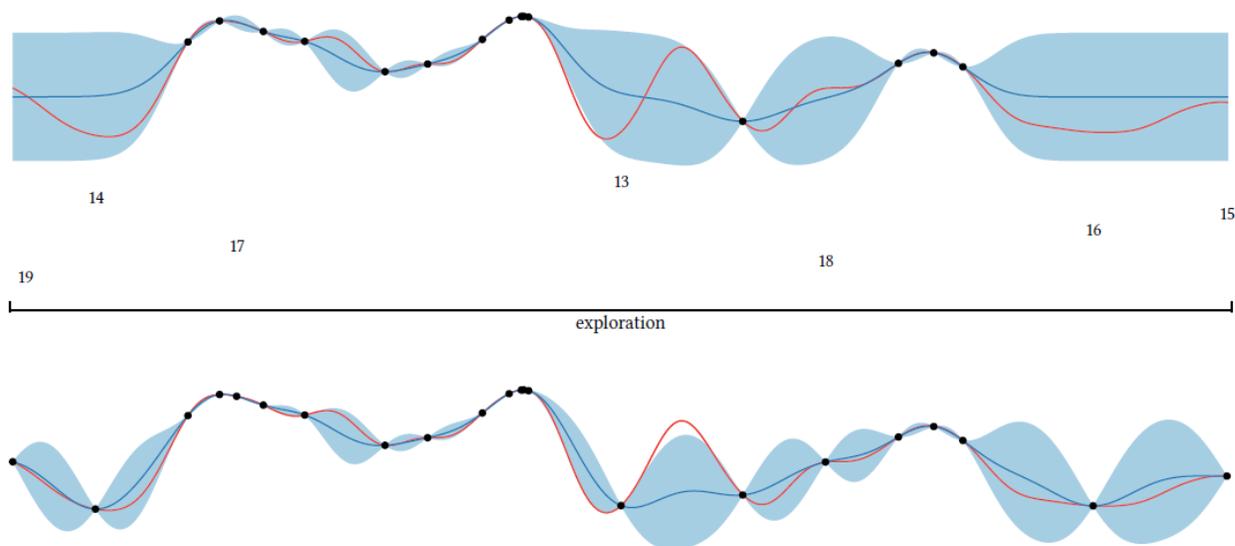
(b) Primeira adição de pontos pela função de aquisição por *exploitation*.

À medida em que novos pontos foram adicionados no processo de busca local, a função de aquisição passou a ter valores mais altos em outra região, fora da região antes explorada localmente, o que indica que o primeiro candidato se tratava de um mínimo local. A busca por uma região fora da inicialmente explorada é possível pelo balanço entre *exploitation* e *exploration*. Assim, começa novamente a exploração local da nova região de interesse, mostrada na Figura 3.7(c). O processo de exploração local acontece até que o máximo global é encontrado nos passos 7 e 8.



(c) Segunda adição de pontos pela função de aquisição por *exploitation*.

Por fim, o método transiciona para exploração global, sistematicamente cobrindo todo o domínio para garantir que o ponto encontrado é de fato o máximo global.



(d) Terceira adição de pontos pela função de aquisição por *exploration*.

Figura 3.7 – Representação de um processo de busca pelo máximo por função de aquisição.

Fonte: Garnett, 2023, p. 12.

## 4. METODOLOGIA: UM CASO DE TESTE REAL DE ENGENHARIA

Além da aplicação em funções de teste analíticas, o presente trabalho se propôs a aplicar o método FBR-HK em um caso real de engenharia. A motivação partiu do trabalho de Rezek *et al.* (2023), e teve como objetivo otimizar a configuração preliminar utilizada pelos autores para o projeto do rotor de uma turbina hidrocínética carenada.

### 4.1. DEFINIÇÃO DO PROBLEMA DE OTIMIZAÇÃO

Segundo Rezek *et al.* (2023), uma das formas de melhorar a metodologia de projeto por eles proposta seria por meio da otimização do projeto preliminar, isto é, otimizar a configuração da turbina hidrocínética carenada com disco atuador. Dessa forma, para a aplicação do método FBR-HK, a geometria inicial do processo de otimização foi definida como a geometria preliminar do trabalho supracitado, mostrada na Figura 4.1, cujas medidas são dadas em milímetros.

Vale ressaltar que a geometria preliminar de Rezek *et al.* (2023) é resultado de um processo de otimização conduzido por Rezek *et al.* (2021). Neste trabalho os autores otimizaram apenas a geometria do difusor e o ângulo de inclinação deste, utilizando apenas um disco atuador no lugar do conjunto nacele e rotor, ou seja, a presença e a variação da geometria da nacele não foram consideradas. Rezek *et al.* (2021) variaram a queda de pressão do disco atuador e o resultado desse estudo serviu de guia para o valor de queda de pressão utilizado tanto em Rezek *et al.* (2023) quanto na aplicação do presente trabalho.

Dessa forma, as variáveis de projeto do processo de otimização aqui conduzido se diferenciam dos trabalhos citados acima por:

1. Considerar a geometria completa da turbina, ainda sem o rotor, com a presença da nacele para a otimização.
2. A geometria da nacele e da ogiva são parâmetros otimizados.
3. O ângulo de inclinação do difusor também é uma variável de projeto, porém os intervalos para esta partem do ângulo ótimo encontrado em Rezek *et al.* (2021).

O problema foi definido pela otimização dos parâmetros geométricos de uma turbina hidrocínética carenada com o objetivo de maximizar a potência hidráulica por meio da alteração das geometrias de ogiva e de nacele e da variação do ângulo de ataque do difusor.

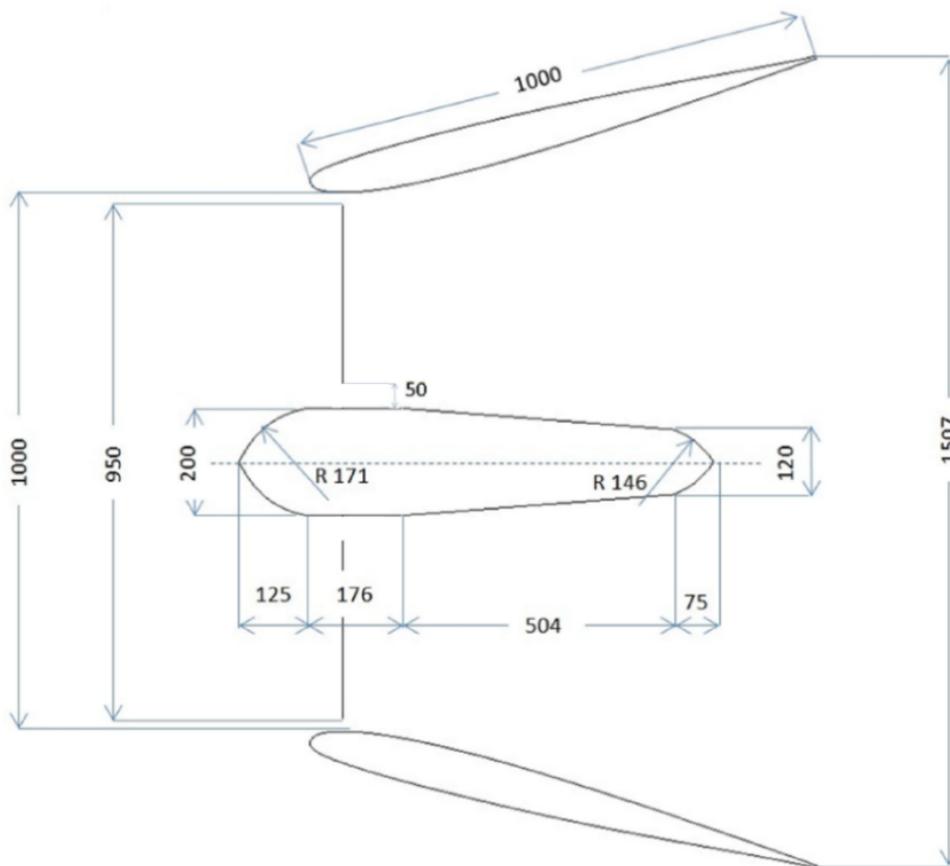


Figura 4.1 – Geometria base da turbina hidrocínética carenada a ser otimizada.

Fonte: Adaptada de Rezek *et al.* (2023).

A partir da geometria inicial foram definidas as variáveis de projeto para o processo de otimização mostradas na Figura 4.2, onde  $O$  representa a origem.

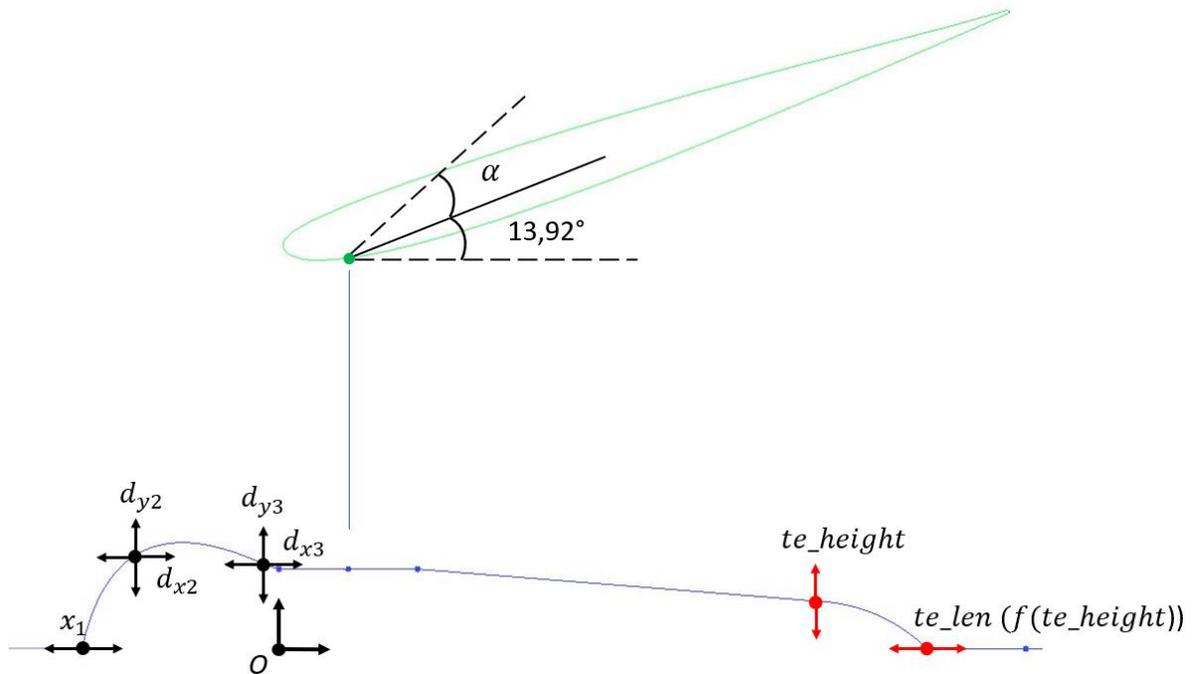


Figura 4.2 – Geometria a ser otimizada e variáveis de projeto.

Para a ogiva, foram definidos três pontos de controle de uma *B-spline*,  $x_1$ ,  $x_2$  e  $x_3$  para investigar a possibilidade de uma geometria de ogiva que atuasse no sentido de acelerar o escoamento antes do rotor. O ângulo de ataque do difusor,  $\alpha$ , também foi escolhido como variável de projeto para investigar a possibilidade de aumento na vazão mássica que passa pela turbina, juntamente com o raio do bordo de fuga da nacele, representado por  $te\_height$  e  $te\_len$ , que também têm influência no descolamento do escoamento à jusante do rotor.

Além disso, vale ressaltar que os valores de folga de topo e de raiz do rotor foram mantidos da geometria base, bem como as dimensões e geometria do difusor e do disco atuador.

Os pontos  $x_2$  e  $x_3$  foram parametrizados em função do ponto  $x_1$  a fim de garantir o formato suave da ogiva e evitar a geração e a avaliação de geometrias inviáveis na realidade. Adicionalmente ao deslocamento em função de  $x_1$ , os pontos  $x_2$  e  $x_3$  se movem na horizontal, representado por  $d_{x2}$  e  $d_{x3}$ , e na vertical,  $d_{y2}$  e  $d_{y3}$ , dentro de um intervalo definido para proporcionar certa movimentação dos pontos de controle e investigar um leque maior de geometrias. Para o bordo de fuga da nacele foi adotada uma abordagem semelhante, em que  $te\_len$  foi parametrizado para variar proporcionalmente a  $te\_height$ , também com o objetivo de evitar a avaliação de geometrias inviáveis na realidade. Totalizando, assim, 7 variáveis de projeto. Os respectivos intervalos, que juntos definem o espaço de projeto, foram definidos de forma arbitrária

levando em consideração os limites de um projeto real e são mostrados na Tabela 4.1, onde as dimensões são dadas em relação à origem. A Tabela 4.2 traz as leis de parametrização de  $x_2$ ,  $x_3$  e de  $te\_len$ .

Variável	Intervalo
$x_1$	[-0,75; -0,25] m
$\alpha$	[-4, 4] °
$d_{x2}$	[0,8; 1,2]
$d_{y2}$	[0,5; 1,5]
$d_{x3}$	[0,5; 1,5]
$d_{y3}$	[0,5; 1,5]
$te\_height$	[0,012; 0,12] m

Tabela 4.1 – Intervalo das variáveis de projeto.

Variável	Lei de Parametrização
$x_2$	$x_2 = ((x_1 * 0,72) * d_{x2}$
$y_2$	$y_2 = ((x_1 * -0,1) * d_{y2}$
$x_3$	$x_3 = ((x_1 * 0,08) * d_{x3}$
$y_3$	$y_3 = ((x_1 * -0,05) * d_{y3}$
$te\_len$	$te\_len = te\_height * 0,2 + 0,89$

Tabela 4.2 – Parametrização das variáveis de projeto.

## 4.2. DESCRIÇÃO DO PROCESSO DE OTIMIZAÇÃO

Uma vez definido o problema de otimização, passou-se para a estruturação do processo. Para a definição do domínio de cálculo, seguindo a mesma abordagem, utilizou-se o domínio definido por Rezek *et al.* (2023), cujas dimensões são mostradas na Figura 4.3. Para o cálculo do escoamento foi utilizado CFD por meio do software Ansys Fluent® (licença disponibilizada pelo LHV/IEM da UNIFEI), em regime transiente, bidimensional e adotando o modelo de turbulência  $k-\omega$  SST. Devido a simetria da turbina, foi possível representá-la por metade da geometria real, aplicando a condição de simetria axial, reduzindo o custo computacional da simulação. O *step time size* foi definido em 0,01 s a fim de possibilitar melhor resolução das pequenas estruturas

turbilhonares. O tempo de escoamento foi definido em 3,15 s para permitir o completo desenvolvimento da esteira, totalizando 315 *time steps*.

As condições de contorno foram definidas com base em Rezek *et al.* (2023). Para a entrada foi adotada a condição de *velocity inlet* com velocidade igual a 2,4 m/s, para o disco atuador foi adotado uma queda de pressão de 2555 Pa, para a saída foi aplicada a condição de *pressure outlet*, para as paredes foi utilizado *no-slip wall* e, por fim, foi aplicada a condição de axissimetria no eixo de simetria da turbina. A teoria por trás da modelagem por disco atuador e por CFD pode ser encontrada no Anexo A.

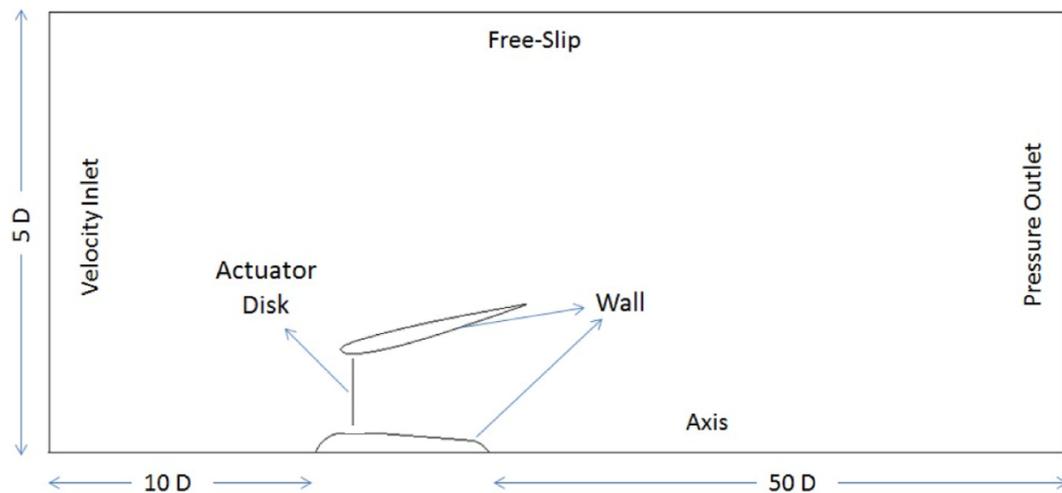


Figura 4.3 – Domínio de cálculo e condições de contorno.

Fonte: Adaptação de Rezek *et al.* (2023).

Para essa aplicação em caso real, foram escolhidos dois níveis de fidelidade associados diretamente ao nível de refinamento da malha, isto é, para alta fidelidade foi utilizada uma malha de maior refinamento e, conseqüentemente, maior custo computacional, e para baixa fidelidade, uma de menor refinamento e menor custo computacional atrelado. A escolha da malha de maior refinamento foi feita através de um estudo de convergência de malha tendo como resposta a potência hidráulica da turbina. A Figura 4.4 traz o comportamento da potência hidráulica em função da evolução do número de elementos da malha.

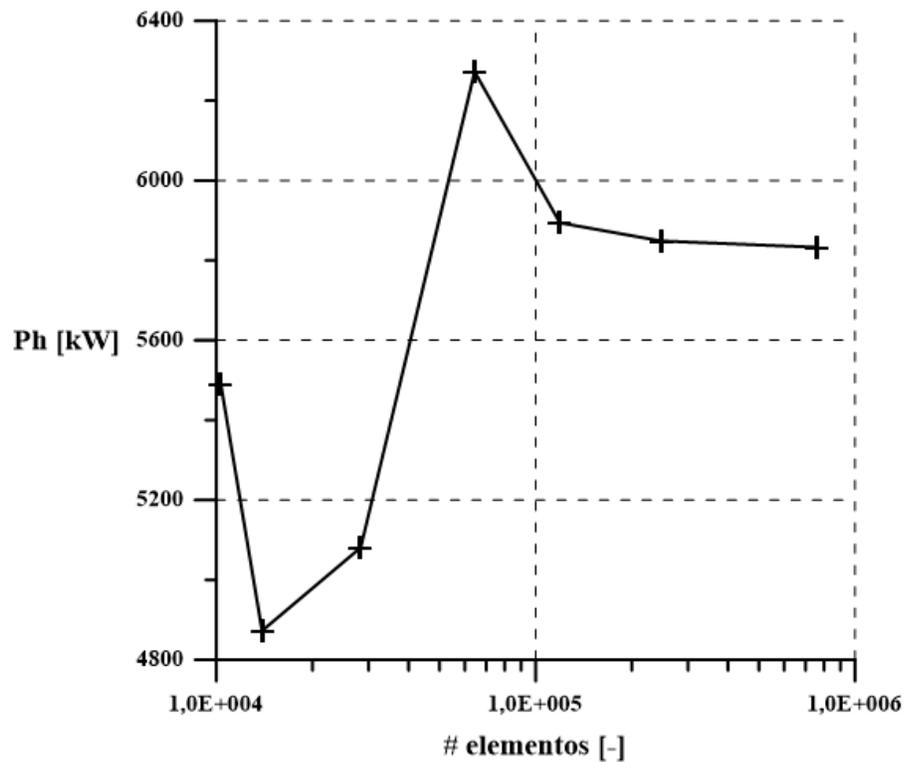
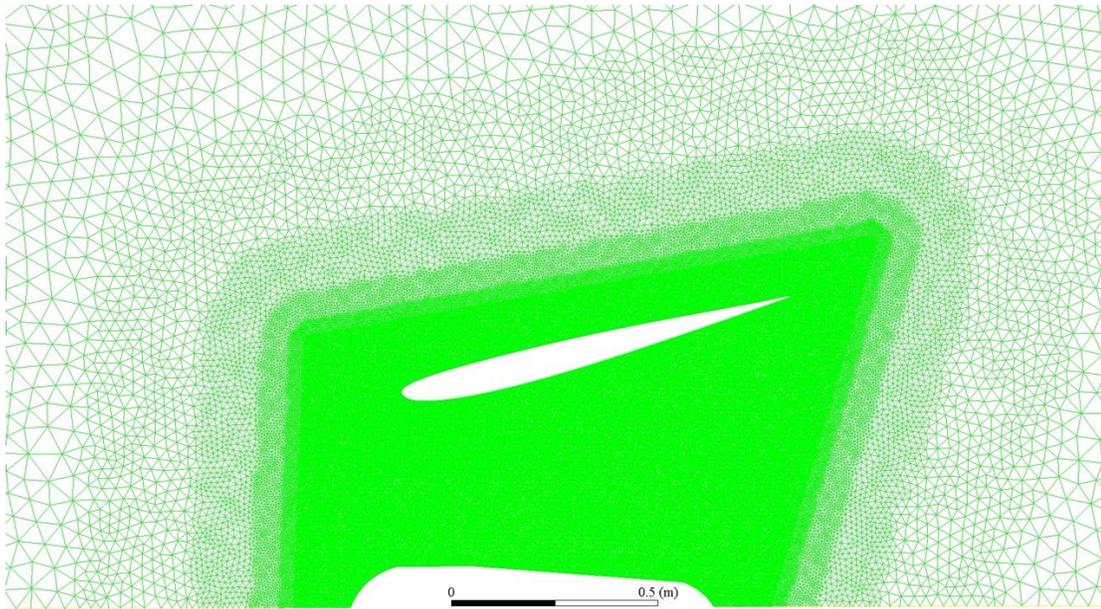
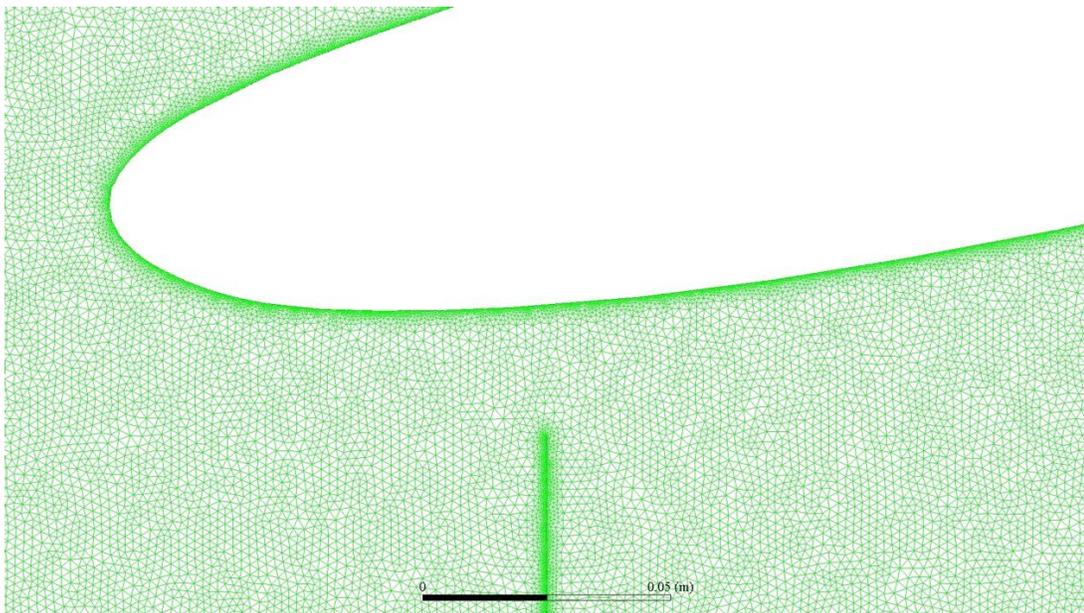


Figura 4.4 – Estudo de convergência de malha.

A discretização do domínio foi feita por uma malha não-estruturada triangular, com a adição de uma região de densidade na região que contém a turbina. A Figura 4.5 apresenta a malha de maior refinamento e o detalhe da região de densidade na proximidade da turbina. A malha de menor refinamento foi constituída por 10333 elementos e a malha de maior refinamento, por 247199 elementos. Foi obtido um valor médio de  $y^+ \approx 3$  para a malha de maior refinamento. O que, dada a configuração da solução mencionada acima, resultou em um tempo total de cálculo do escoamento de cerca de 30 minutos para a baixa fidelidade e de 4 horas e 30 minutos para a alta fidelidade, isto é, uma razão de custo computacional de 1:9. Os cálculos foram feitos utilizando uma máquina com processador de 4 núcleos e 16 GB de memória RAM.



(a) Malha de maior refinamento e região de densidade.



(b) Detalhe do refinamento nas paredes e no disco atuador.

Figura 4.5 – Malha de maior refinamento e refinamento nas paredes.

A resposta de interesse para o processo de otimização é a potência hidráulica, extraída diretamente do Ansys Fluent® por meio da multiplicação da queda de pressão no disco atuador pela vazão volumétrica que passa pelo mesmo.

Uma vez definido o processo que fará o cálculo da resposta de interesse dado um conjunto de variáveis de projeto, passa-se para a estrutura do processo de otimização em si, que consiste na integração do método FBR-HK com o processo de cálculo da resposta.

O primeiro passo do processo de otimização é a definição do conjunto de pontos inicial que é usado de base para a construção da primeira iteração do metamodelo. Esse conjunto foi determinado por meio de hipercubo latino (LHS) com 20 pontos de baixa fidelidade e 10 de alta fidelidade. Foi adotado um pequeno conjunto inicial com o objetivo de possibilitar a melhor utilização do reduzido orçamento computacional disponível por meio da função de aquisição ao invés de gastá-lo preenchendo o espaço de projeto. Assim, o conjunto inicial consumiu cerca de 1/3 do orçamento total estipulado de uma semana, que equivale a 168 horas de cálculo. Este orçamento foi definido tendo em mente o objetivo de aumentar a eficiência do processo de otimização multifidelidade, especialmente em aplicações em que orçamento disponível é pequeno.

O próximo passo do processo é a construção dos metamodelos, IHK ou FBR-HK, sobre os pontos do conjunto inicial, isto foi feito por meio da implementação em MATLAB® (licença disponibilizada pelo LHV/IEM da UNIFEI) do equacionamento apresentado na Seção 3.2.5. Uma vez construído o metamodelo, tem-se a estimativa de erro associado mostrada na Equação 3.64, que é utilizada pela função de aquisição por Expectativa de Melhoria da Equação 3.66, também implementada em MATLAB®, para a aquisição do novo ponto de amostragem. O novo ponto é encontrado por meio de um processo de otimização global com o objetivo de maximizar a Expectativa de Melhoria. O processo é feito para ambos os níveis de fidelidade e o ponto escolhido é aquele com maior valor entre os dois. Vale ressaltar que o algoritmo genético NSGA-II, de Srinivas e Deb (1994), foi utilizado para esta otimização em sua configuração padrão.

Esse novo ponto é transferido para um script em *Tcl/Tk* lido pelo software Ansys ICEM CFD®, que gera a geometria e a malha com base nos valores das variáveis de projeto do novo ponto de amostragem e com base no nível de fidelidade ao qual este novo ponto pertence. A nova malha gerada é encaminhada para o Ansys Fluent®, que faz o carregamento da malha em um arquivo de caso já configurado conforme os parâmetros mostrados no início desta seção e calcula a solução do escoamento juntamente com o valor de potência hidráulica. Este valor é capturado pela rotina do metamodelo e entra como o valor de resposta para o novo ponto de amostragem escolhido previamente. Este processo acontece iterativamente até que o valor de Expectativa de Melhoria atinja o limite estipulado ou até que o processo gaste todo o orçamento computacional

definido no início do processo, dado em horas. O fluxograma da Figura 4.6 apresenta a estrutura do processo de otimização conduzido por FBR-HK.

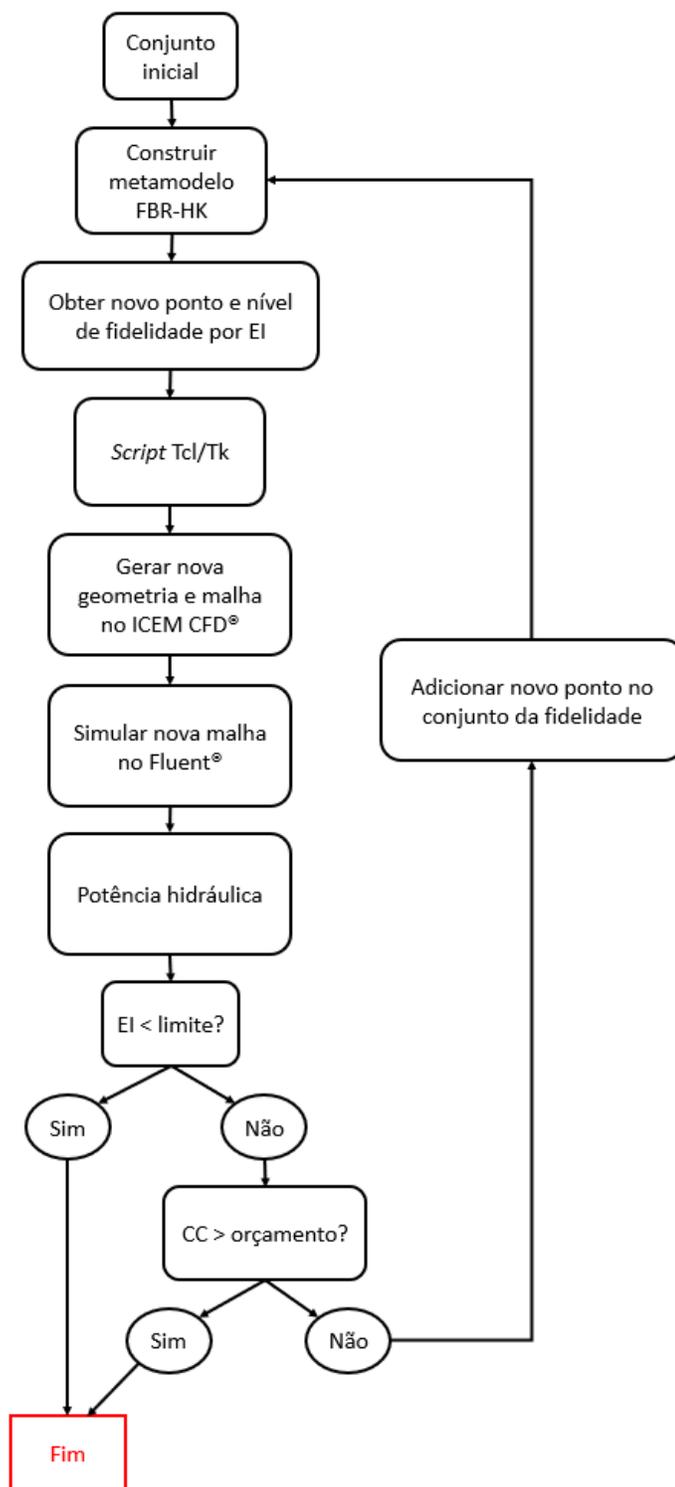


Figura 4.6 – Fluxograma do processo de otimização por FBR-HK.

## 5. RESULTADOS

Preliminarmente aos estudos da presente tese, e como continuação dos desenvolvimentos da dissertação de mestrado do presente autor, foi conduzida uma análise comparativa de metamodelos, ainda no contexto da aplicação da dissertação de mestrado, turbinas eólicas. A análise apresentada a seguir teve como objetivo a avaliação do desempenho do método de Kriging como processo de otimização no contexto de turbinas, área, até então, pouco explorada na literatura da época. Os metamodelos foram empregados na abordagem não-sequencial, onde são construídos com base em um conjunto fixo de pontos e posteriormente otimizados. Nesse cenário, o Kriging destacou-se como uma alternativa competitiva em relação à superfície de resposta clássica e à rede neural artificial. Os resultados apresentados na seção 5.1 foram publicados em um periódico internacional revisado por pares. É importante destacar que o método de Kriging é amplamente utilizado no contexto de otimização sequencial Bayesiana, em que o metamodelo é atualizado a cada adição de novos pontos, sendo essa a sua aplicação mais indicada.

### 5.1. RESULTADOS COMPARATIVOS DE METAMODELOS

O estudo cujos resultados são, aqui, apresentados serviram como forma de avaliação da precisão dos métodos de metamodelagem para uma abordagem não-sequencial. Entenda-se por abordagem não-sequencial aquela na qual o metamodelo é construído com base em um conjunto de pontos fixo, sem a adição subsequente de novos pontos. Com base nesse estudo foi possível comparar o método de Kriging com métodos mais conhecidos na literatura de engenharia: superfície de resposta clássica e funções de base radial, aplicadas no contexto das redes neurais artificiais. A comparação serviu para corroborar a utilização do método de Kriging para metamodelagem de problemas de turbinas, uma vez que a precisão deste foi equiparável a do metamodelo utilizando funções de base radial.

Baseado em otimização assistida por metamodelos, o estudo teve por objetivo de encontrar os parâmetros geométricos ótimos para o aerofólio das pás de uma turbina eólica de eixo vertical. A Figura 5.1 traz uma foto da configuração real desta turbina. A Figura 5.2 mostra o modelo computacional bidimensional utilizado para representar a turbina estudada. Os parâmetros

escolhidos foram corda e espessura do aerofólio. E o objetivo da otimização era maximizar o coeficiente de potência ( $C_p$ ) da turbina para uma razão de ponta de pá (TSR) de 1,5.



Figura 5.1 – Configuração real da turbina eólica de eixo vertical otimizada.

Fonte: Bravo, Tullis e Ziada (2007).

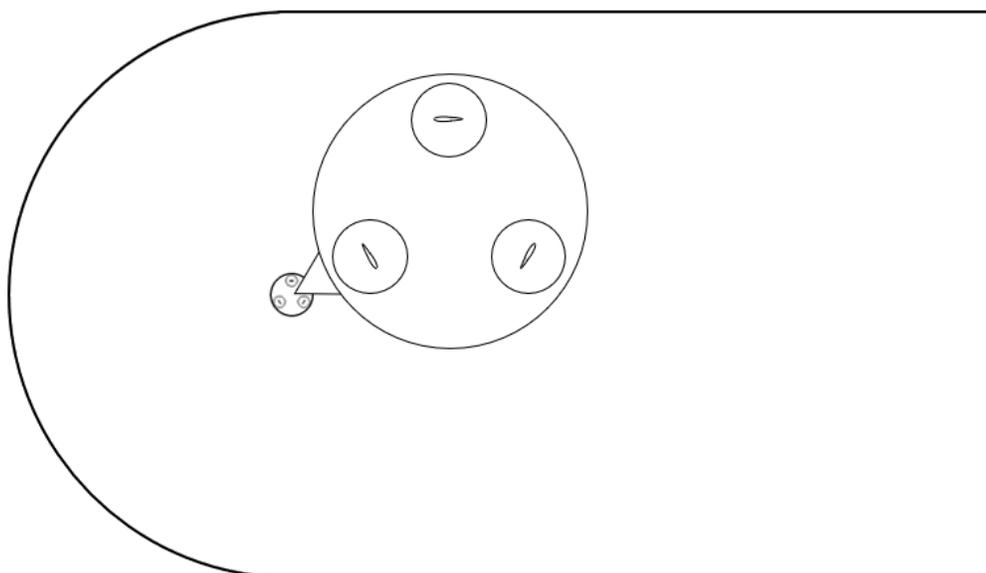


Figura 5.2 – Modelo computacional da turbina otimizada.

Três diferentes tipos de metamodelos foram utilizados e comparados: superfície de resposta polinomial quadrática clássica (RS), redes neurais artificiais do tipo *Extreme Learning Machine* (ELM), que utilizam funções de base radial, e interpolador por Kriging. A modelagem matemática por trás dos metamodelos RS e ELM pode ser encontrada nos Anexos B e C.

O conjunto de dados usado como base para construção dos metamodelos foi feito usando fatorial completo com 4 níveis para cada um dos dois fatores, corda e espessura. A distribuição dos pontos dos conjuntos de treino e de teste é apresentada na Figura 5.3. Os valores de coeficiente de potência correspondentes para cada ponto do conjunto de treino e de teste foram obtidos usando CFD. A Figura 5.4 mostra o fluxograma do processo de otimização.

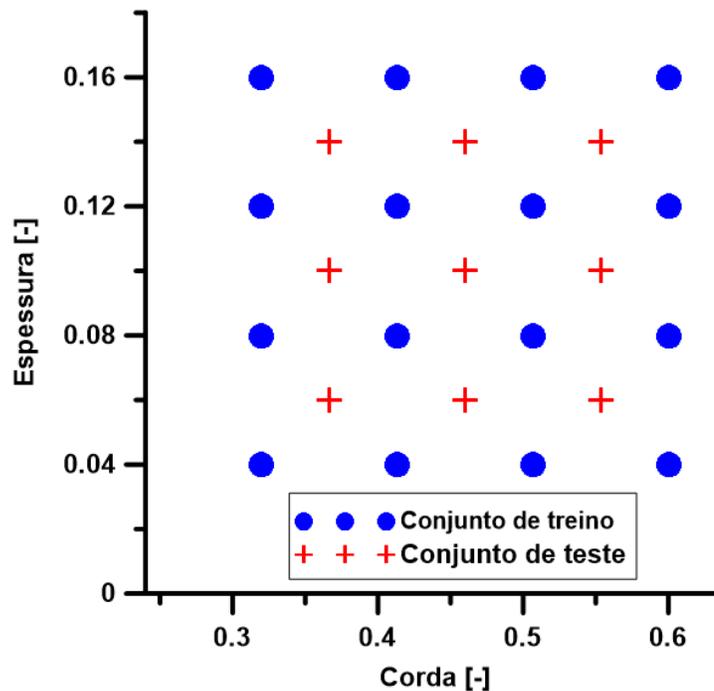


Figura 5.3 – Conjuntos de treino e de teste.

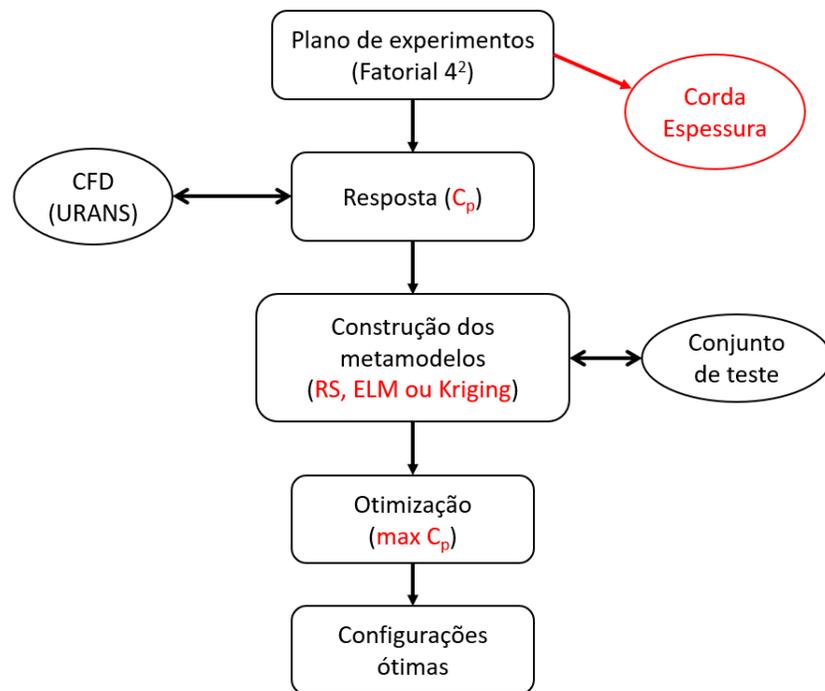


Figura 5.4 – Fluxograma do processo de otimização.

A Tabela 5.1 apresenta os valores de coeficiente de potência previstos por cada um dos metamodelos para os pontos de teste e o erro associado a eles.

Conjunto de teste			Previsão RS	Erro [%]	Previsão ELM	Erro [%]	Previsão Kriging	Erro [%]
Corda	Espessura	$C_p$						
0,367	0,060	0,4207	0,3661	<b>14,913%</b>	0,3984	<b>5,594%</b>	0,3852	<b>9,215%</b>
0,367	0,100	0,4389	0,4113	<b>6,710%</b>	0,4497	<b>2,394%</b>	0,4527	<b>3,049%</b>
0,367	0,140	0,3601	0,3306	<b>8,924%</b>	0,3405	<b>5,771%</b>	0,3249	<b>10,835%</b>
0,460	0,060	0,4156	0,4112	<b>1,081%</b>	0,4092	<b>1,566%</b>	0,4007	<b>3,729%</b>
0,460	0,100	0,4685	0,4875	<b>3,899%</b>	0,4836	<b>3,117%</b>	0,4868	<b>3,761%</b>
0,460	0,140	0,4506	0,4378	<b>2,918%</b>	0,4419	<b>1,953%</b>	0,4420	<b>1,940%</b>
0,553	0,060	0,3865	0,3729	<b>3,655%</b>	0,3692	<b>4,683%</b>	0,3622	<b>6,718%</b>
0,553	0,100	0,4384	0,4802	<b>8,709%</b>	0,4407	<b>0,530%</b>	0,4437	<b>1,200%</b>
0,553	0,140	0,4361	0,4616	<b>5,526%</b>	0,4184	<b>4,236%</b>	0,4168	<b>4,629%</b>
			MSE	<b>4,917%</b>	MSE	<b>1,265%</b>	MSE	<b>3,114%</b>
			RMSE	<b>22,174%</b>	RMSE	<b>11,246%</b>	RMSE	<b>17,647%</b>

Tabela 5.1 – RMSE dos metamodelos para os pontos de teste.

O metamodelo RS resultou no maior RMSE, o que era esperado, uma vez que o comportamento da função objetivo real dificilmente será tão suave quanto o de uma superfície polinomial de segunda ordem. Os metamodelos ELM e Kriging tiveram RMSE de 11,25% e 17,65% respectivamente. A partir disso, pode ser inferido que o metamodelo ELM apresentou a maior precisão para o processo estudado.

Uma vez avaliada a precisão dos metamodelos, prosseguiu-se para o processo de otimização. O algoritmo de busca utilizado foi NSGA-II. A Tabela 5.2 mostra os valores ótimos de corda, de espessura, a proporção espessura sobre corda e os respectivos  $C_p$ , obtidos por CFD, com a geometria base para comparação. A Figura 5.5 compara as geometrias ótimas obtidas através de cada processo de otimização e a geometria base. Comprimentos estão escalados com referência à corda da geometria de base.

#### Geometrias ótimas

Metamodelo	Corda	Espessura	Proporção Espessura/Corda	$C_p$
<b>Base</b>	0,4000	0,0600	0,1500	0,4395
<b>RS Clássica</b>	0,5072	0,1086	0,2142	0,4583
<b>ELM</b>	0,4367	0,0953	0,2183	0,4739
<b>Kriging</b>	0,4353	0,0962	0,2209	0,4743

Tabela 5.2 – Valores ótimos e coeficiente de potência.

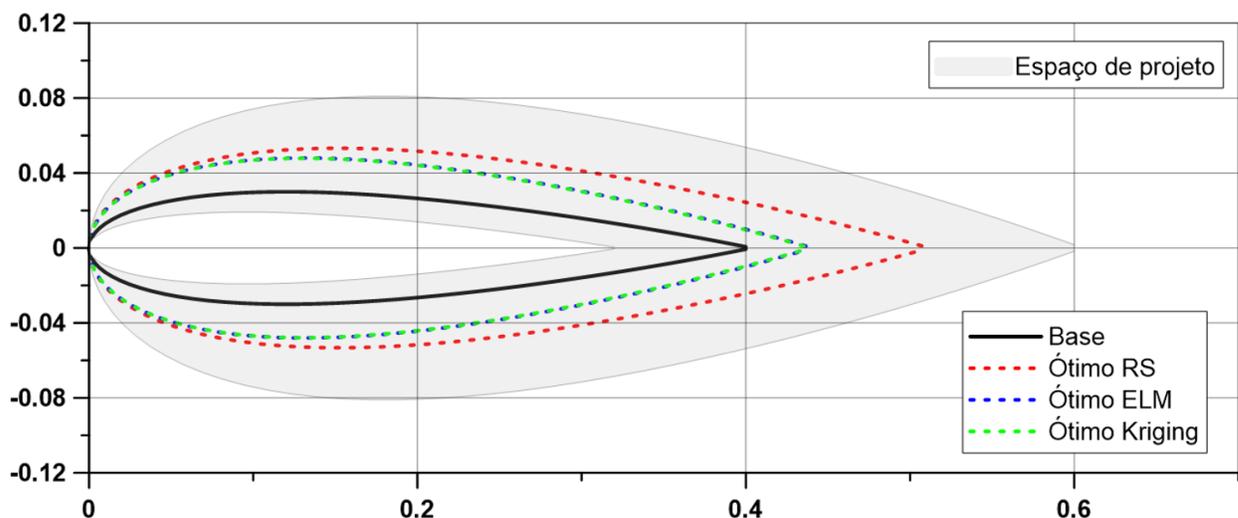


Figura 5.5 – Geometrias de aerofólio ótimas e geometria base.

Da Figura 5.5 pode-se notar que a geometria obtida pelo processo de otimização usando RS levou a um aerofólio maior, no geral, com maiores corda e espessura. Ambos metamodelos ELM e Kriging resultaram em geometrias similares. A tendência dos processos de otimização foi aumentar a espessura e corda do aerofólio em comparação com a geometria base.

Por fim, a Figura 5.6 apresenta a curva de coeficiente de potência em função do TSR das geometrias ótimas e da geometria base de Bravo, Tullis e Ziada (2007).

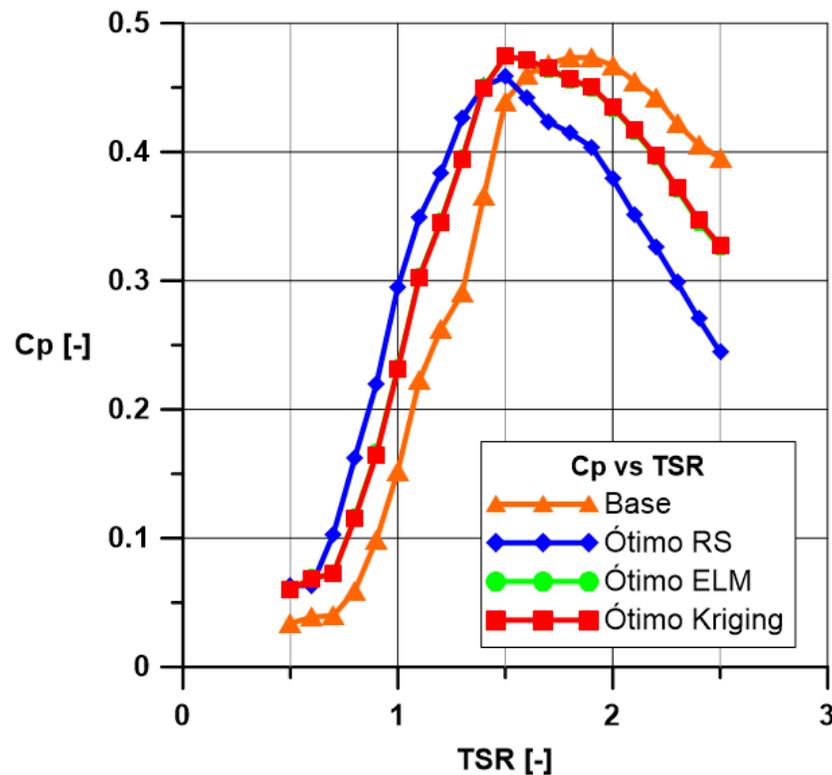


Figura 5.6 –  $C_p$  versus TSR para geometrias ótimas e base.

Pode-se observar que todas as geometrias ótimas resultaram em maiores valores de  $C_p$  quando comparadas com a geometria base para uma faixa de TSR de 0,5 a 1,5 para RS, e até 1,7 para ELM e Kriging. No ponto de otimização, TSR igual a 1,5, os valores de  $C_p$  obtidos pelas geometrias ótimas foram 0,458 para RS, 0,473 para ELM e 0,474 para Kriging, um aumento de 4,29%, 7,86% e 7,92% respectivamente, quando comparado ao  $C_p$  base de 0,439. Como as geometrias encontradas por ELM e Kriging foram próximas, era de se esperar que as curvas de  $C_p$  também fossem. É interessante notar que o ótimo por RS resultou em maiores valores de  $C_p$  entre

todas as geometrias para TSR de 0,5-1,4. Após TSR de 1,4, o comportamento é invertido, com ótimo por RS apresentando os menores valores de  $C_p$  dentre as geometrias ótimas.

## 5.2. RESULTADOS DA APLICAÇÃO DE FBR-HK EM FUNÇÕES DE TESTE ANALÍTICAS

Uma vez implementados em MATLAB® os algoritmos dos métodos de HK, IHK e FBR-HK, foi possível a aplicação dos métodos em funções analíticas de teste utilizadas comumente na literatura, com a dimensão variando de 1 a 5. Cada função de teste possuía dois níveis de fidelidade, tomados como alta e baixa fidelidade, apesar de apresentarem custos computacionais muitas vezes semelhantes. Os casos de teste foram selecionados de Forrester (2007), Currin *et al.* (1988), Hartman (1973) e Ackley (2012), respectivamente por ordem de número de dimensões. A Tabela 5.3 reúne as expressões de cada função. As definições dos casos de teste são mostradas na Tabela 5.4.

Nome e fonte	Descrição das expressões
Forrester <i>et al.</i> (2007)	$f_{hf}(x) = (6x - 2)^2 \text{sen}(12x - 4)$ $x \in [0,1]$ $x^* = 0,75725 \quad f^* = -6,020740$
Forrester 1a	$f_{lf}(x) = 0,5f(x) + 10(x - 0,5) - 5$
Forrester 1b	$f_{lf}(x) = f(x) - 5$
Forrester 1c	$f_{lf}(x) = [6(x + 0,2) - 2]^2 \text{sen}[12(x + 0,2) - 4]$
Currin <i>et al.</i> (1988)	$f_{hf}(\mathbf{x}) = \left[1 - \exp\left(-\frac{1}{2x_2}\right)\right] \frac{2300x_1^3 + 1900x_1^2 + 2092x_1 + 60}{100x_1^3 + 500x_1^2 + 4x_1 + 20}$ $x_1, x_2 \in [0,1]$ $\mathbf{x}^* = [0,21667 \ 0] \quad f^* = -13,7987$
Xiong <i>et al.</i> (2013)	$f_{lf}(\mathbf{x}) = \frac{1}{4}[f(x_1 + 0,05, x_2 + 0,05) + f(x_1 + 0,05, \max(0, x_2 - 0,05))] + \frac{1}{4}[f(x_1 - 0,05, x_2 + 0,05) + f(x_1 - 0,05, \max(0, x_2 - 0,05))]$

Hartman 3 (Hartman (1973))	$f_{hf}(\mathbf{x}) = - \sum_{i=1}^4 c_i \exp \left[ - \sum_{j=1}^3 \alpha_{ij} (x_j - p_{ij})^2 \right]$
	$\text{onde } \alpha_{ij} = \begin{bmatrix} 3 & 10 & 30 \\ 0,1 & 10 & 35 \\ 3 & 10 & 30 \\ 0,1 & 10 & 35 \end{bmatrix}, c_i = \begin{bmatrix} 1 \\ 1,2 \\ 3 \\ 3,2 \end{bmatrix} \text{ e } p_{ij} = \begin{bmatrix} 0,3689 & 0,1170 & 0,2673 \\ 0,4699 & 0,4387 & 0,7470 \\ 0,1091 & 0,8732 & 0,5547 \\ 0,03815 & 0,5743 & 0,8828 \end{bmatrix}$
	$x_i \in [0,1], \text{ para } i = 1, 2, 3 \text{ e } j = 1, \dots, 4$
	$\mathbf{x}^* = [0,114 \ 0,556 \ 0,852] \ f^* = -3,8627$
MA 3	$y(\mathbf{x}) = 0,585 - 0,324x_1 - 0,379x_2 - 0,431x_3 - 0,208x_1x_2 + 0,326x_1x_3 + ,0193x_2x_3 + 0,225x_1^2 + 0,263x_2^2 + 0,274x_3^2$
Ackley 5 (Ackley (2012))	$f_{hf}(\mathbf{x}) = -20 \exp \left[ -0,2 \sqrt{\frac{1}{5} \sum_{i=1}^5 x_i^2} \right] - \exp \left[ \frac{1}{5} \sum_{i=1}^5 \cos(cx_i) \right] + 20 + \exp(1)$
	$x_i \in [-2,2], \text{ para } i = 1, \dots, 5$
	$\mathbf{x}^* = [0, \dots, 0] \ f^* = 0$
MA 5	$y(\mathbf{x}) = 0,585 - 0,324x_1 - 0,379x_2 - 0,431x_3 - 0,208x_1x_2 + 0,326x_1x_3 + ,0193x_2x_3 + 0,225x_1^2 + 0,263x_2^2 + 0,274x_3^2$

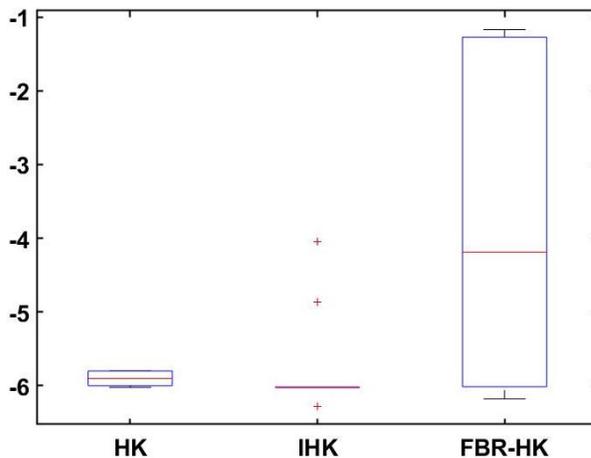
Tabela 5.3 – Formulação das funções de teste.

Caso de teste	Dimensões	Função de alta fidelidade	Função de baixa fidelidade	Ótimo
1	1	Forrester <i>et al.</i> (2007)	Forrester 1a	$f^* = -6,020740$
2	1	Forrester <i>et al.</i> (2007)	Forrester 1b	$f^* = -6,020740$
3	1	Forrester <i>et al.</i> (2007)	Forrester 1c	$f^* = -6,020740$
4	2	Currin <i>et al.</i> (1988)	Xiong <i>et al.</i> (2013)	$f^* = -13,7987$
5	3	Hartman 3	Hartman 3 + 7,6 × MA 3	$f^* = -3,8627$
6	5	Ackley 5	Ackley 5 + 0,74 × MA 5	$f^* = 0$

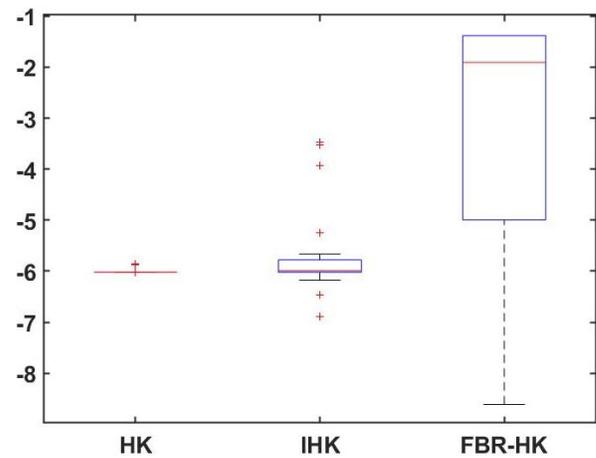
Tabela 5.4 – Definição das funções multifidelidade de teste.

Para todos os casos de teste, o número inicial de pontos para baixa e alta fidelidades foi de 3 e 6 vezes o número de dimensões da função de teste usada, respectivamente, e foram determinados por LHS. O processo de otimização global traz consigo a aleatoriedade dos processos

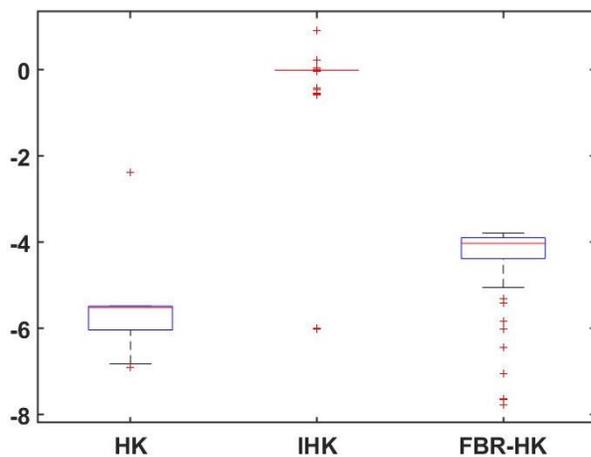
de otimização internos dos hiper-parâmetros e do próprio processo de maximização da expectativa de melhoria, conceito base da otimização global eficiente. Assim, os casos de teste de 1 dimensão foram avaliados 100 vezes e os de maior dimensão, 20 vezes, por cada método. Além disso, como forma de eliminar mais uma fonte de aleatoriedade que adviria do LHS, as avaliações foram feitas partindo dos mesmos conjuntos de pontos de alta e de baixa fidelidade para um mesmo caso de teste. A Figura 5.7 apresenta os diagramas de caixas de cada um dos casos de teste, onde os quartis inferiores (25%) e superiores (75%) são mostrados. A linha vermelha na caixa denota a mediana dos valores. As linhas tracejadas que se estendem dos limites da caixa representam 1,5 vezes a distância entre quartis. Pontos fora dessa distância são considerados outliers e são representados pelo símbolo “+” vermelho.



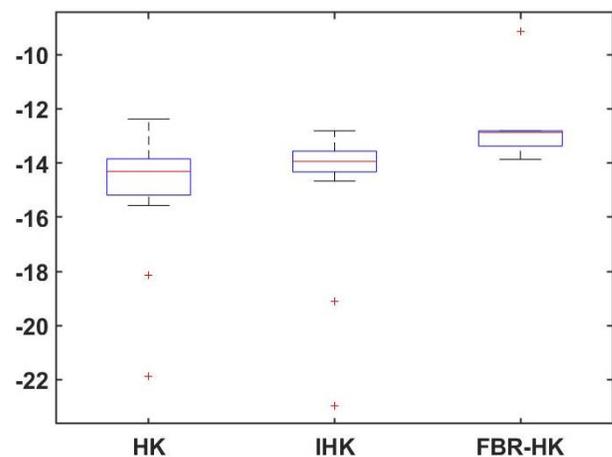
(a) Caso 1



(b) Caso 2



(c) Caso 3



(d) Caso 4

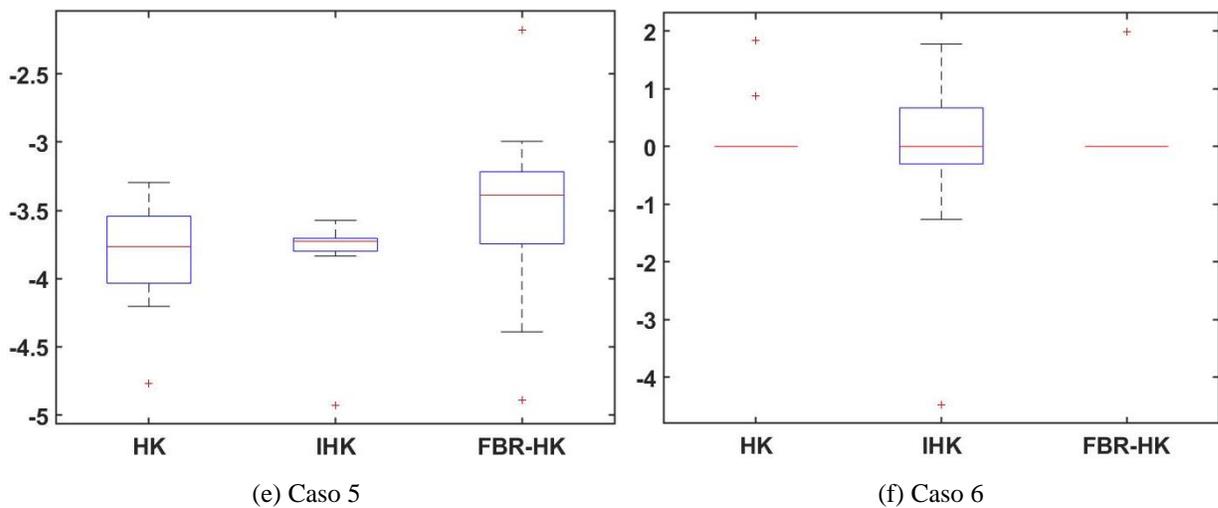


Figura 5.7 – Comparação dos ótimos encontrados pelos diferentes métodos.

Da Figura 5.7, pode-se observar que os métodos HK e IHK obtiveram melhores resultados para o caso 1 e 2, quando comparado com o método proposto. Para o caso 3, FBR-HK apresentou resultados próximos ao mínimo e ao HK, porém com maior presença de outliers, enquanto o método IHK não se aproximou do ótimo. Para o caso 4, de 2 dimensões, o método proposto foi o que mais se aproximou do mínimo, com a menor variância dentre os métodos comparados. Já para o caso de teste 5, IHK se aproximou do mínimo com a menor variância. Por fim, no caso 6, FBR-HK teve o melhor resultado dentre os métodos, com apenas 1 outlier e a menor variância nos resultados. A Tabela 5.5 apresenta um sumário das avaliações dos casos de teste. A taxa de sucesso,  $tx$ , representa a razão entre o número de vezes em que o processo encontrou o ótimo pelo número total de processos de otimização. O número médio de avaliações da função de alta fidelidade,  $\mu_{hf}$ , e de baixa fidelidade,  $\mu_{lf}$ , para cada método, assim como o desvio padrão associado,  $\sigma_{hf}$  e  $\sigma_{lf}$ , respectivamente, são medidas estatísticas calculadas com base no número de avaliações de alta e de baixa fidelidade realizadas ao final do processo. O objetivo é auxiliar na comparação dos diferentes processos de otimização ao longo de uma amostragem de 100 avaliações para os casos unidimensionais e 20 avaliações para casos multidimensionais.

A partir da tabela, observa-se que o método FBR-HK obteve a maior taxa de sucesso no caso com o maior número de variáveis e alta multimodalidade. Isso pode ser atribuído à própria natureza das funções de base radial, especialmente quando se utiliza a função de base radial Gaussiana, que tem uma capacidade maior de aproximar funções com comportamento multimodal. Nos demais casos, o desempenho do FBR-HK foi similar ou inferior ao dos métodos HK e IHK.

Caso	1			4			5			6		
	Método	HK	IHK	FBR-HK	HK	IHK	FBR-HK	HK	IHK	FBR-HK	HK	IHK
tx	49%	83%	37%	35%	55%	40%	50%	60%	24%	90%	32%	95%
$\mu_{lf}$	2,60	2,95	1,08	7,42	10,67	10,5	7,3	12,9	23	0,5	4,75	1,57
$\sigma_{lf}$	1,15	1,35	0,95	3,88	3,86	2,29	5	4,57	5,33	0,5	3,23	0,67
$\mu_{hf}$	1	2,33	1,11	7,71	5,55	1,25	8,6	3,41	4,4	1	4,25	1,1
$\sigma_{hf}$	0	0,47	0,51	4,77	1,89	0,43	8,58	1,32	2,6	0	2,16	0,30

Tabela 5.5 – Sumário dos processos de otimização dos casos de teste 1, 4, 5 e 6.

Ao analisar o número de avaliações de alta e de baixa fidelidade, observa-se que o método FBR-HK fez menos uso, em média, das avaliações de alta fidelidade do que os outros métodos, em troca de utilizar mais avaliações de baixa fidelidade. No caso das funções de teste, a função de baixa fidelidade, muitas vezes, é composta por uma perturbação da função considerada de alta fidelidade, isto somado ao fato do método ter utilizado mais avaliações de baixa fidelidade, ainda que corrigidas pela alta fidelidade, pode ter influenciado negativamente o processo de busca pelo ótimo.

### 5.3. RESULTADOS DA APLICAÇÃO DE FBR-HK PARA O CASO REAL DE ENGENHARIA

Dado o processo de otimização definido na Seção 5, o método proposto (FBR-HK) foi comparado com outros dois métodos da literatura, Kriging hierárquico aprimorado (IHK), apresentado na Seção 4.2.4, e Kriging sequencial monofidelidade (SFK).

Para o método IHK utilizou-se a mesma definição de fonte de alta e de baixa fidelidade do método FBR-HK e para o método monofidelidade SFK foi utilizada apenas a fonte de alta fidelidade.

A fim de estabelecer a mesma base de comparação, todos os métodos partiram do mesmo conjunto inicial e os métodos FBR-HK e IHK tiveram o mesmo orçamento computacional de 168 horas. Para o método SFK, por motivos de disponibilidade de equipamento, foi utilizado um

orçamento menor, de 120 horas. Além disso, todos os métodos utilizaram a mesma função de aquisição baseada em Expectativa de Melhoria.

Como o processo de aquisição de novos pontos se dá pela otimização global da função de Expectativa de Melhoria, a existência de aleatoriedade neste processo é inerente e não se pode desconsiderar a influência disso no próprio processo de otimização sequencial. Assim, a fim de reduzir tais efeitos e levando em consideração a disponibilidade de tempo e de equipamento, o processo por FBR-HK foi conduzido 10 vezes e os processos IHK e SFK, 5 vezes cada.

A Tabela 5.6 apresenta a configuração ótima encontrada por cada método e seus respectivos valores máximos de potência hidráulica dentre toda a amostragem de otimização sequencial. A coluna CC apresenta o custo computacional consumido, em horas, para chegar ao ótimo. O valor de  $\alpha$  mostra o ângulo de inclinação do difusor em termos de adição ao ângulo de  $13,92^\circ$  da geometria inicial. Os valores de  $d_{x2}$ ,  $d_{x3}$ ,  $d_{y2}$  e  $d_{y3}$ , como apresentado na Tabela 5.2, são multiplicadores das coordenadas em  $x$  e em  $y$  dos pontos  $x_2$  e  $x_3$ .

	$\alpha$ [°]	te_height [m]	x1 [m]	dx2 [-]	dy2 [-]	dx3 [-]	dy3 [-]	Ph [W]	CC
<b>FBR-HK</b>	4,000	0,012	-0,25	0,874	0,500	1,500	1,500	<b>7076,642</b>	<b>144</b>
<b>IHK</b>	4,000	0,012	-0,25	0,953	0,501	0,500	1,500	<b>7071,743</b>	<b>165</b>
<b>SFK</b>	4,000	0,012	-0,25	0,800	0,788	0,500	1,500	<b>7087,579</b>	<b>108</b>

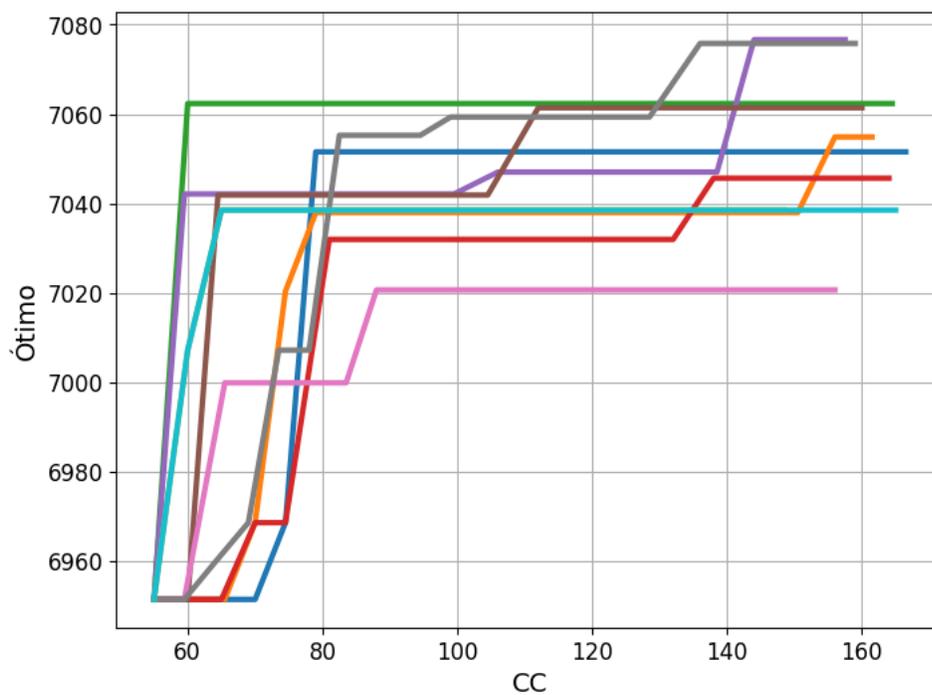
Tabela 5.6 – Configurações ótimas de cada método e respectiva potência hidráulica.

Para todos os métodos, a tendência de otimização foi aumentar o canal da turbina a jusante do rotor, com um ângulo de inclinação do difusor total de  $17,92^\circ$  em relação a horizontal. Outro parâmetro que seguiu a mesma tendência foi o raio do bordo de fuga da nacele, com um total de 2,4 cm, valor este que pode sofrer impacto de restrições estruturais na realidade. A posição ótima encontrada para o ponto de controle  $x_1$  foi de 0,25 m, o dobro da posição original. Os multiplicadores do ponto  $x_2$ ,  $d_{x2}$  e  $d_{y2}$ , mostraram um ajuste no sentido de reduzir o deslocamento proporcional a  $x_1$ . Já para o ponto  $x_3$ , FBR-HK resultou num aumento da posição em  $x$ , ao passo que os outros métodos reduziram essa posição pela metade. Para a posição em  $y$ , todos os métodos resultaram em um aumento de 50% além da proporcionalidade a  $x_1$ . Para trabalhos futuros, sugere-se uma análise de sensibilidade para as variáveis de projeto a fim de avaliar a real influência de cada uma na potência hidráulica.

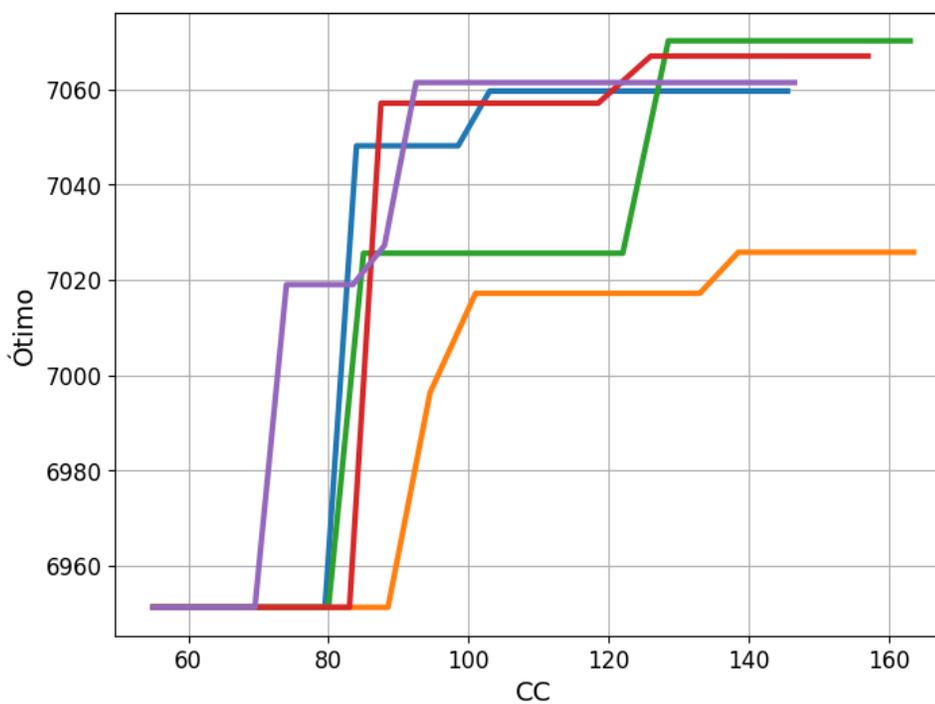
Considerando o melhor desempenho de cada método, observa-se que o FBR-HK alcançou um valor maior de potência hidráulica em comparação com o IHK, associado a um menor custo computacional. Em contrapartida, o SFK obteve o maior valor de potência hidráulica entre os três métodos, além de apresentar o menor custo computacional. Como mencionado anteriormente, existem aleatoriedades inerentes ao processo de aquisição de novos pontos, portanto, as conclusões aqui estabelecidas devem ser interpretadas com um grau de relatividade.

A Figuras 5.8(a), (b) e (c) trazem as curvas de evolução dos processos de otimização com 10 avaliações para o FBR-HK e 5 para IHK e SFK. É possível notar por essas figuras a influência da aleatoriedade do processo de aquisição de novos pontos e a importância de se ter uma amostragem representativa para melhor analisar os resultados e traçar conclusões. Embora cada processo de otimização sequencial comece com o mesmo conjunto inicial, eles seguem caminhos ligeiramente diferentes. É importante ressaltar que todos os métodos começam do mesmo ponto inicial, correspondente ao maior valor de potência hidráulica do conjunto inicial de alta fidelidade gerado por LHS, avaliado por CFD com a malha de maior refinamento.

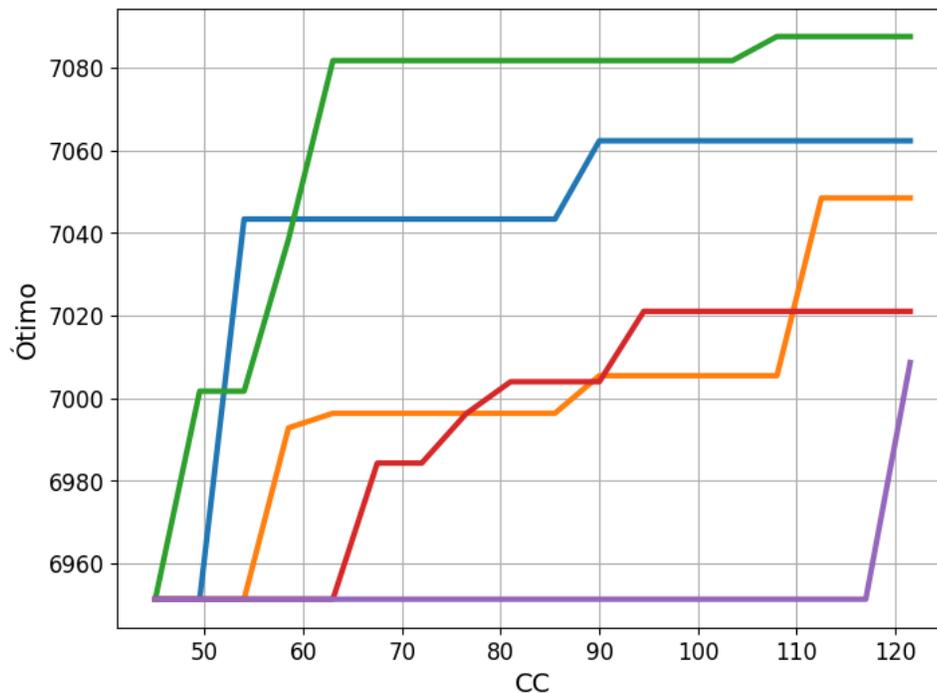
Comparando as figuras em termos da distribuição das curvas de evolução, é possível observar que os métodos multifidelidade, FBR-HK e IHK, apresentam evoluções mais consistentes em direção ao ótimo, frente ao SFK, cujas curvas de evolução são mais esparsas. Isto é um indicador da variabilidade do resultado e da capacidade de se repetir um ótimo dada uma avaliação aleatória. O método de otimização ideal seria aquele que resulta no mesmo ponto de ótimo em qualquer avaliação, isto é, que apresente baixa variabilidade. A análise feita aqui se limita às amostragens de 10 e 5 avaliações, assim sendo, conclusões mais representativas poderiam ser traçadas com uma maior amostragem.



(a) Curvas de evolução do processo de otimização para FBR-HK para 10 avaliações.



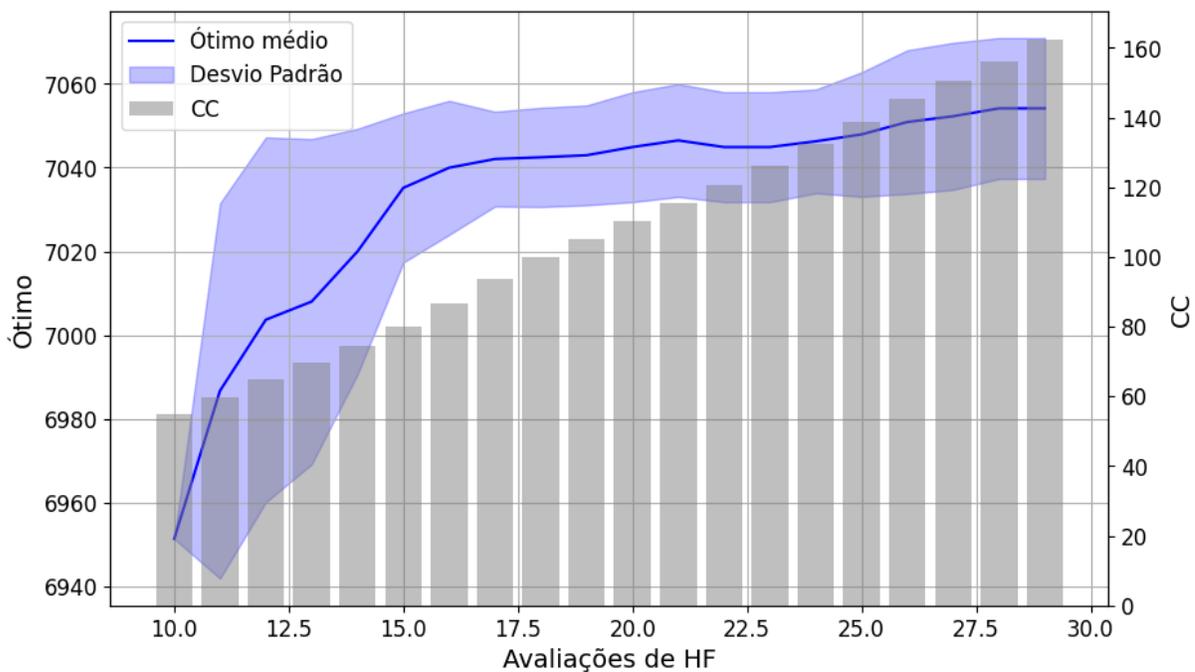
(b) Curvas de evolução do processo de otimização para IHK para 5 avaliações.



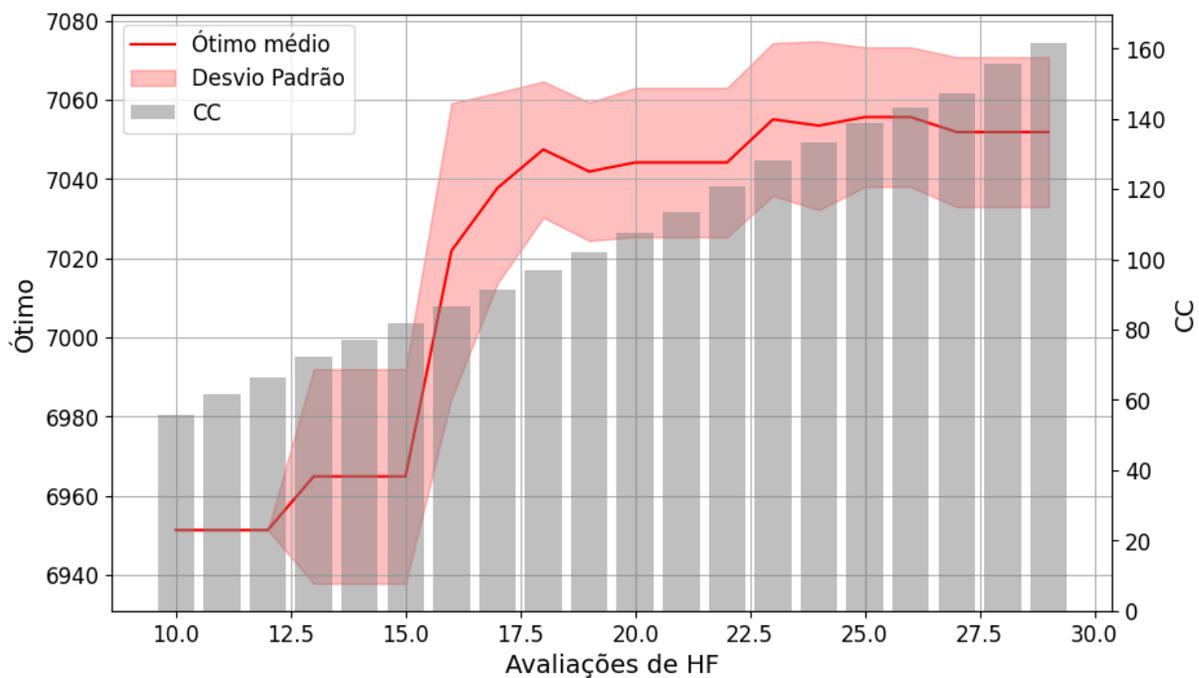
(c) Curvas de evolução do processo de otimização para SFK para 5 avaliações.

Figura 5.8 – Evolução dos processos de otimização para FBR-HK, IHK e SFK.

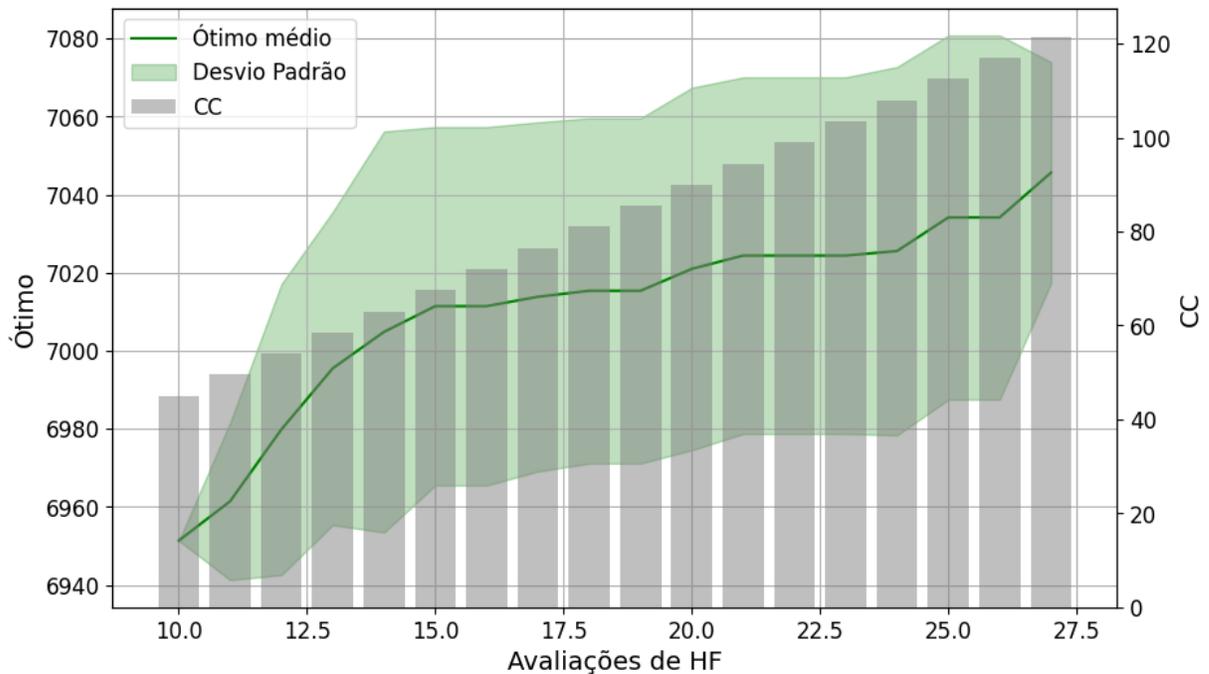
As Figuras 5.9(a), (b) e (c) mostram uma representação estatística dos processos mostrados nas figuras acima. Tomando todas as curvas de evolução para cada método, foi possível calcular a média de potência hidráulica ótima para cada iteração dos processos de otimização, bem como o desvio padrão associado a ela, pois, como foi visto, as avaliações nem sempre seguem o mesmo caminho de evolução. A evolução média da potência hidráulica agora é representada em função do número de avaliações de alta fidelidade e o gráfico traz um segundo eixo das ordenadas, que mostra a evolução do custo computacional total, resultado da soma das avaliações de baixa e de alta fidelidade.



(a) Evolução do processo de otimização para FBR-HK em função de avaliações de HF e evolução do custo computacional médio.



(b) Evolução do processo de otimização para IHK em função de avaliações de HF e evolução do custo computacional médio.



(c) Evolução do processo de otimização para SFK em função de avaliações de HF e evolução do custo computacional médio.

Figura 5.9 – Evolução dos processos de otimização em termos de média e desvio padrão e custo computacional médio associado.

Agora em termos médios, observa-se que o ótimo alcançado pelo SFK no seu melhor desempenho, apesar de ser superior aos outros dois métodos, está associado a um desvio padrão maior, isto é, a uma maior variabilidade do resultado. Isso é um indício de que a probabilidade de reproduzir consistentemente os resultados ótimos aleatoriamente é baixa, especialmente quando comparado às bandas de desvios padrão do FBR-HK e do IHK. Além disso, quando a comparação é feita com base na evolução média, o método SFK apresenta o pior desempenho entre os três métodos.

A Figura 5.10 combina as Figuras 5.7(a), (b) e (c) para melhor comparação dos métodos.

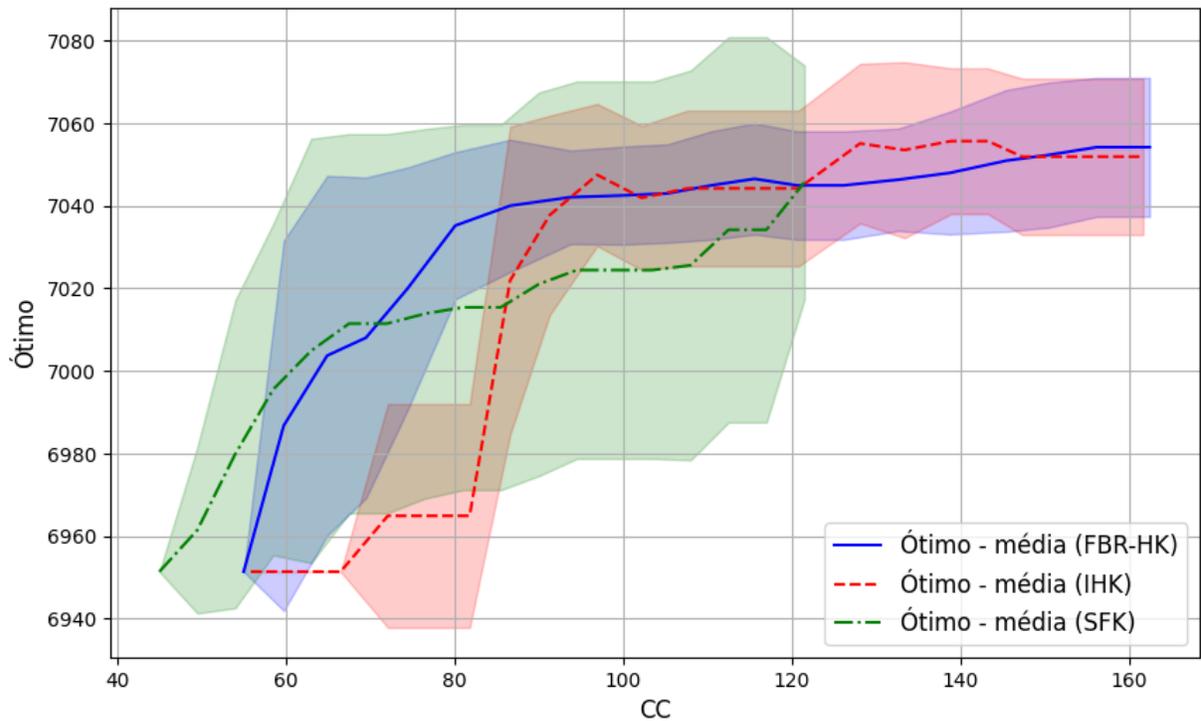


Figura 5.10 – Comparação da evolução média dos processos de otimização.

O método IHK apresenta um desempenho médio superior em comparação com o SFK para orçamentos maiores do que cerca de 85 horas. Para orçamentos menores, o desempenho médio do SFK se mostra mais viável. Em relação ao desvio padrão, o IHK resultou em uma banda mais estreita que o SFK, indicando uma maior capacidade de reprodução consistente dos resultados ótimos.

O FBR-HK demonstrou um desempenho médio mais constante e ascendente quando comparado com IHK, e produziu, em média, valores ótimos maiores que os outros métodos para orçamentos maiores que 70 horas, mas não se limitando a estes, e equiparáveis ao SFK para orçamentos menores. Em termos de desvio padrão, o FBR-HK apresenta a banda mais estreita entre os métodos avaliados, sugerindo que seus resultados ótimos são os mais consistentes.

Adicionalmente, ao observar a evolução do desempenho ao longo do custo computacional, evidencia-se que o FBR-HK não só mantém uma trajetória de melhoria mais estável que IHK como também atinge valores ótimos mais altos de forma mais consistente, isto é, com menor variabilidade. Esse comportamento reflete uma robustez maior no processo de otimização para o cenário avaliado. Além disso, a melhoria média da potência hidráulica maior que a do método IHK para o primeiro quarto do orçamento computacional pode indicar que o método seja ainda mais

vantajoso para processos de otimização em que se disponha de baixo custo computacional, e em que se deseje avaliar relações de troca entre parâmetros em menor tempo.

A Tabela 5.7 traz a média de avaliações de alta e de baixa fidelidades utilizadas pelos métodos IHK e FBR-HK no processo apresentado na Figura 5.10. Observa-se que ambos os métodos distribuíram o orçamento computacional de maneira semelhante, entretanto, quando se olha a Tabela 5.7 em conjunto com as observações da Figura 5.10, pode-se perceber que apesar dos métodos terem utilizado um número semelhante de avaliações em média, o método FBR-HK apresentou desempenho ligeiramente melhor em termos médios e em termos de variabilidade dos resultados, o que pode indicar que o método é capaz de construir o metamodelo de forma mais eficiente dada as mesmas informações disponíveis.

<b>Método</b>	<b>Média de HF</b>	<b>Média de LF</b>
<b>FBR-HK</b>	29,4	75,4
<b>IHK</b>	28	75,2

Tabela 5.7 – Média de avaliações utilizadas pelos métodos.

## **5.4. RESULTADOS DAS SIMULAÇÕES EM CFD E COMPARAÇÃO COM A GEOMETRIA BASE**

A Figura 5.11 compara a configuração base de Rezek et al. (2023) com a configuração ótima obtida pelo método FBR-HK, apresentada na Tabela 5.6. Observa-se que a otimização geométrica foi direcionada para aumentar o canal a jusante do rotor, diminuir o ângulo de ataque do difusor e reduzir o raio do bordo de fuga da nacele. Essas alterações resultam em um aumento da vazão mássica que passa pelo disco atuador e, como a queda de pressão foi mantida fixa em 2555 Pa, o resultado é um aumento da potência hidráulica e, conseqüentemente, do coeficiente de potência hidráulica da turbina. A Tabela 5.8 compara os valores de vazão mássica, potência hidráulica e coeficiente de potência hidráulica, medido na saída da turbina e na garganta, respectivamente, para a configuração base e a ótima. Os coeficientes de potência hidráulica foram obtidos pela divisão da potência hidráulica pela potência disponível do fluxo livre.

Além disso, a elevação da ogiva acelera o escoamento, ainda que ligeiramente, antes de passar pelo rotor, como mostrado na Figura 5.12. Assim, a otimização do projeto preliminar da

turbina hidrocínética utilizando o método FBR-HK resultou em um aumento da potência hidráulica e, conseqüentemente, do coeficiente de potência hidráulica de cerca de 21%.

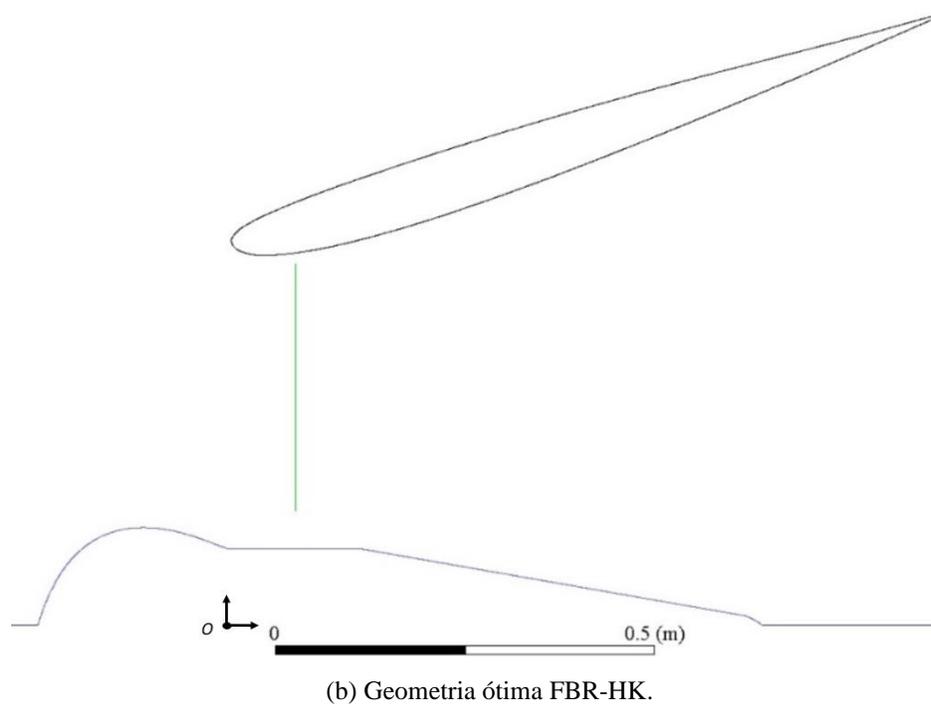
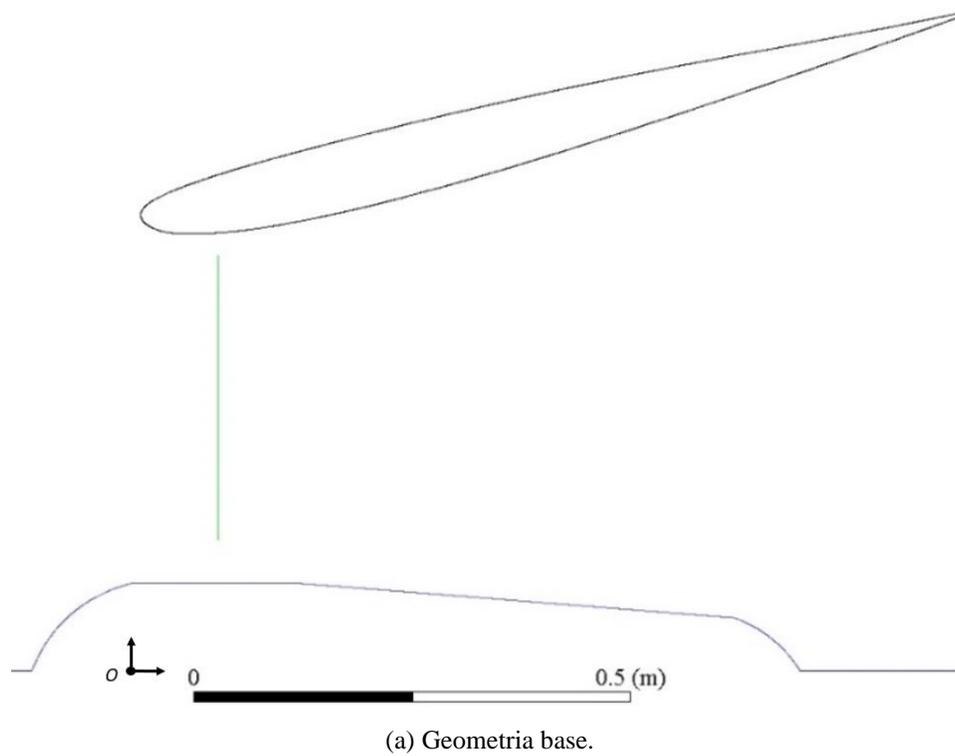
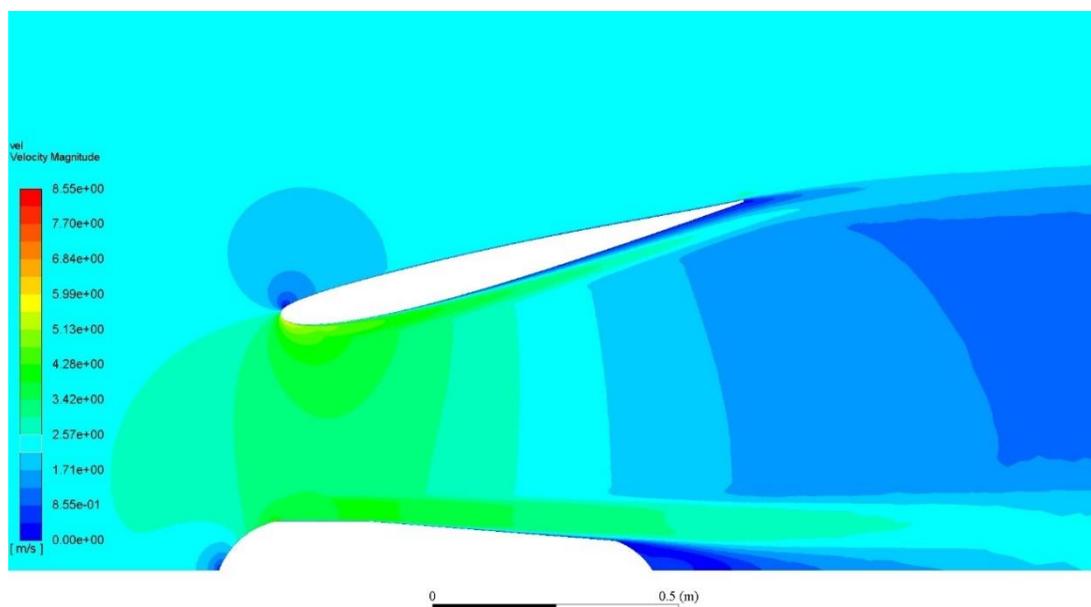


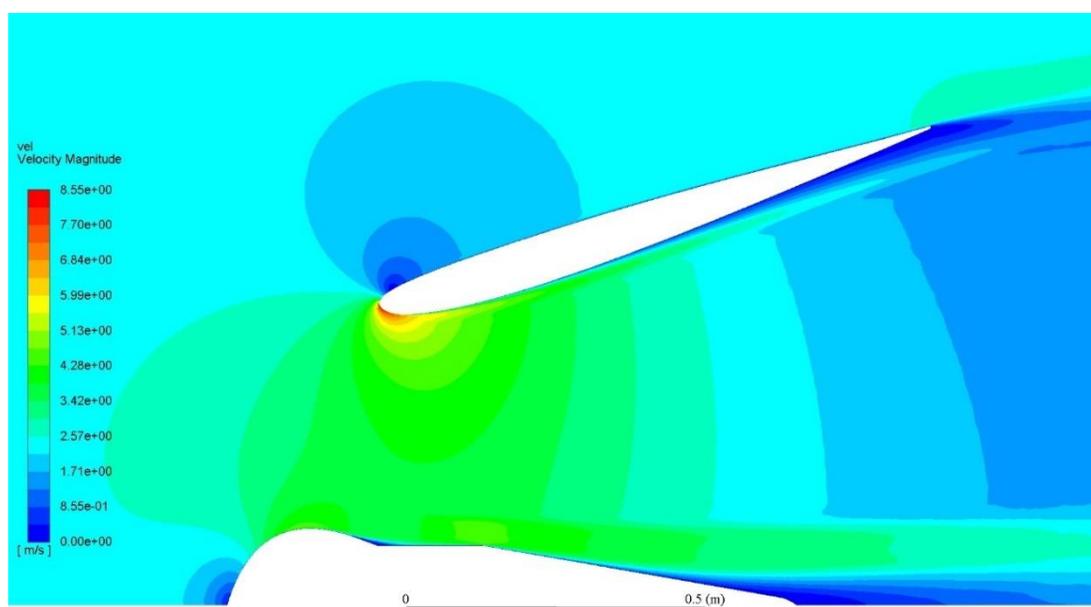
Figura 5.11 – Geometrias base e ótimas de turbina hidrocínética.

	Vazão mássica [kg/s]	Ph [W]	CP_saída	CP_garganta	Aumento
<b>Base</b>	2.284,943	5.848,558	0,4752	1,1243	-
<b>Ótimo</b>	2.764,737	7.076,642	0,5750	1,3603	21%

Tabela 5.8 – Comparação geometria base e ótima.



(a) Geometria base.



(b) Geometria ótima.

Figura 5.12 – Contornos de magnitude do vetor velocidade.

Por fim, pode-se obter a distribuição de velocidade axial pelo rotor para a geometria otimizada, um dos parâmetros de interesse obtidos no projeto preliminar da turbina hidrocínética. Este parâmetro, junto com outros parâmetros como velocidade da corrente livre, vazão volumétrica, queda de pressão do rotor e outros parâmetros geométricos são o ponto de partida para o projeto do rotor seguindo a metodologia proposta por Rezek *et al.* (2023). A Figura 5.13 compara a distribuição de velocidade axial ao longo do disco atuador para a geometria base e a geometria ótima.

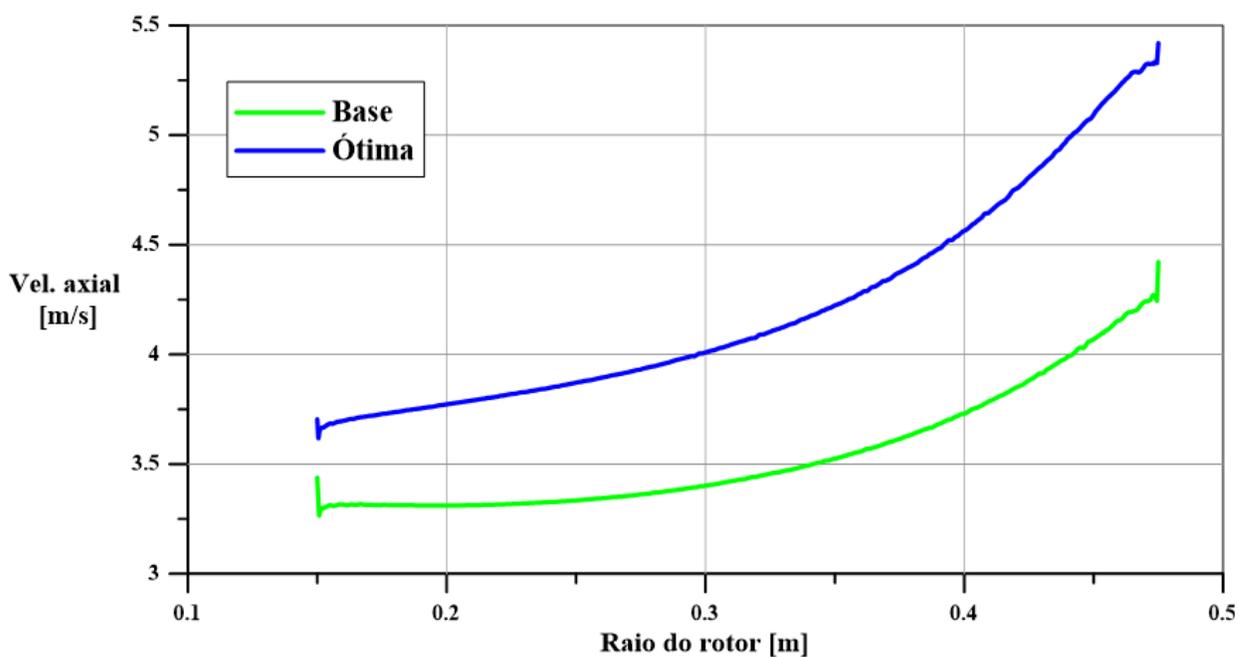


Figura 5.13 – Distribuição de velocidade axial para geometria base e ótima.

## 6. CONCLUSÕES E SUGESTÕES

### 6.1. CONCLUSÕES

Este trabalho se propôs a desenvolver um método de otimização global sequencial que fosse mais eficiente do que os disponíveis na literatura de multifidelidade que fazem uso do Kriging. O método foi desenvolvido por meio do aprimoramento do Kriging hierárquico utilizando funções de base radial. Além disso, o FBR-HK também propôs a correção dos pontos de baixa fidelidade pelos pontos de alta fidelidade na construção do Kriging de baixa fidelidade.

O método FBR-HK foi testado em aplicações em funções de teste analíticas e aplicado em um caso real de engenharia na otimização do projeto preliminar de uma turbina hidrocíntrica carenada.

Da aplicação em funções de teste, observou-se que o método apresentou a maior taxa de sucesso para o caso com maior número de variáveis e com a mais alta multimodalidade. Isto pode sugerir que o método FBR-HK seja mais indicado, quando comparado com o IHK, para a otimização de sistemas mais complexos, nos quais o número de variáveis de projeto é maior e a relação entre elas é altamente não-linear.

Ao analisar o número de avaliações de alta e de baixa fidelidade, observou-se que o método FBR-HK utilizou, em média, menos avaliações de alta fidelidade quando comparado aos outros métodos. Por outro lado, o FBR-HK fez uso de mais avaliações de baixa fidelidade. Como mencionado, para as funções de teste, a função de baixa fidelidade, muitas vezes, é composta por uma perturbação da função de alta fidelidade, o que pode ter influenciado negativamente o processo de busca pelo ótimo.

Para a aplicação no caso real de engenharia, a baixa fidelidade, ainda que forneça uma solução com menor precisão, é representativa do problema. Assim, analisando o gráfico de evolução dos processos de otimização, percebeu-se que ambos os métodos FBR-HK e IHK apresentaram desvios padrão menores do que do SFK. O que diferencia este daqueles é a utilização de avaliações de baixa fidelidade. Assim, o resultado observado poderia indicar que, para a aplicação em questão, as informações obtidas pelos métodos multifidelidade através das avaliações

de baixa fidelidade auxiliaram no processo de otimização, reduzindo a variabilidade, e que o uso da multifidelidade se mostrou vantajoso.

Em termos do desempenho do método FBR-HK, a melhoria média da potência hidráulica maior que a do método IHK para o primeiro quarto do orçamento computacional apontou que o método possa ser ainda mais vantajoso para processos de otimização em que se disponha de baixo orçamento computacional, e em que se deseje avaliar relações de troca entre parâmetros em menor tempo.

A geometria resultante do processo de otimização mostrou que o melhoramento se deu no sentido de aumentar o canal da turbina a jusante do rotor, o que possibilitou maior vazão mássica por este e, conseqüentemente, maior potência hidráulica e coeficiente de potência hidráulica. Outro resultado importante do processo foi o aumento do diâmetro da ogiva, que além de acelerar o escoamento antes de passar pelo rotor, poderia possibilitar o alojamento de um gerador maior, podendo aumentar a potência elétrica gerada pela turbina. No entanto, o valor ótimo encontrado de 2,4 cm para o raio do bordo de fuga da nacele pode não ser viável na realidade devido a restrições estruturais.

## 6.2. SUGESTÕES

Como sugestões de trabalhos futuros, pode-se destacar duas frentes: melhorias no método FBR-HK e melhorias na otimização do projeto preliminar da turbina hidrocínética. Para o primeiro, tem-se:

- Investigação de funções de aquisição mais eficientes e atualizadas: existe na literatura um esforço na busca de funções de aquisição mais eficientes que utilizem não só a expectativa de melhoria, mas uma combinação entre esta e outras funções de aquisição como probabilidade de melhoria e limite inferior de confiança, a fim de que cada método seja utilizado quando a condição for mais adequada. Existem também esforços para combinar a expectativa de melhoria com funções que avaliam a relação de custo computacional entre diferentes fidelidades, a densidade de distribuição de pontos e a correlação entre os níveis de fidelidade. Se mais eficientes, essas funções poderiam tornar o processo de busca sequencial pelo ótimo ainda mais eficiente, reduzindo o custo computacional envolvido.

- Avaliar o desempenho do FBR-HK apenas com o aprimoramento por funções de base radial: o método propôs desde a concepção, simultaneamente, o aprimoramento do Kriging de alta fidelidade por funções de base radial e a correção do Kriging de baixa fidelidade pelos pontos de alta fidelidade, sem avaliar o efeito de cada uma separadamente. Ao avaliar o desempenho do método apenas com o aprimoramento por FBR, seria possível mensurar a real melhoria, se de fato existir, da correção dos pontos de baixa pelos pontos de alta fidelidade.
- Aumentar a amostragem para cada método: para tornar a base de comparação ainda mais justa, sugere-se aumentar a amostragem dos métodos FBR-HK, IHK e SFK para que a comparação da evolução dos processos de otimização seja mais fidedigna ao comportamento médio de cada método.
- Utilizar o método FBR-HK em aplicações de otimização robusta: os resultados indicaram que o FBR-HK seria mais adequado para otimização de sistemas mais complexos e com alta não-linearidade. Dessa forma, utilizar o método em um processo de otimização robusta investigaria, por meio de um caso real de engenharia, o seu desempenho em um processo de otimização de maior complexidade.

Para o projeto preliminar, sugere-se:

- Adicionar a queda de pressão do rotor como variável de projeto: a aplicação do presente trabalho teve como variáveis de projeto apenas os parâmetros geométricos da turbina. Por meio da inclusão de um parâmetro operacional, como a queda de pressão do rotor, seria possível investigar um leque maior de configurações e avaliar a influência da queda de pressão na geometria da turbina.
- Utilizar rotação da esteira para a alta fidelidade: incluir efeitos de rotação para o disco atuador possivelmente aumentaria a precisão da avaliação de alta fidelidade, tornando o processo de otimização mais próximo da realidade.
- Alterar os limites dos pontos de controle da nacele e da ogiva: nos resultados foram levantadas observações sobre a possibilidade de, de fato, alojar-se um maior gerador na ogiva e a restrição estrutural quanto à espessura do bordo de fuga da nacele. Aumentar os limites superiores de variação dos pontos de controle da ogiva e o limite inferior de variação

do raio do bordo de fuga da nacele poderia resultar em uma geometria mais factível na realidade.

- Dimensionar o rotor: utilizando os parâmetros encontrados pela otimização do projeto preliminar, pode-se prosseguir para o dimensionamento do rotor de acordo com a metodologia da literatura. Assim, seria possível, a partir da simulação tridimensional da turbina com o rotor projetado ao invés de um disco atuador, comparar o projeto otimizado pelo FBR-HK com o projeto base.

## ANEXO A. TEORIA UNIDIMENSIONAL DO DISCO ATUADOR E CFD

### A.1. TEORIA UNIDIMENSIONAL DA QUANTIDADE DE MOVIMENTO E O LIMITE DE BETZ

A teoria da quantidade de movimento unidimensional fornece um método simplificado, porém poderoso, para analisar o desempenho de um rotor de turbina axial ideal. O conceito central envolve modelar o rotor como um disco atuador, que induz uma descontinuidade de pressão estática no fluxo que passa por ele. Para isso, considera-se um volume de controle delimitado pela superfície do tubo de corrente e duas seções transversais na entrada e saída do disco atuador, como mostra a Figura A.1.

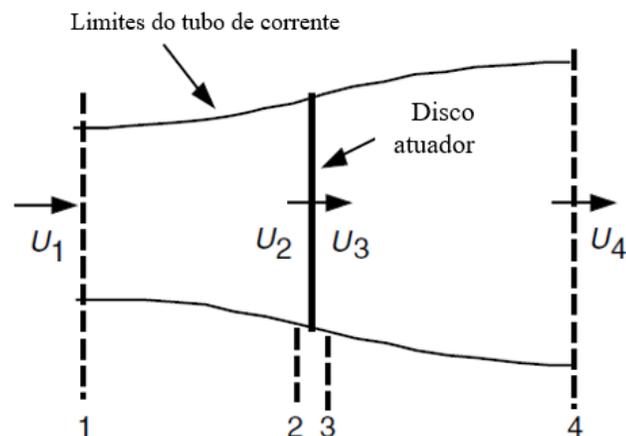


Figura A.1 – Modelo de disco atuador de uma turbina axial;  $U$  representa a velocidade média do escoamento; 1, 2, 3 e 4 indicam a posição.

Fonte: Adaptação de Manwell, McGowan e Rogers, 2009, p. 93.

Para este tipo de análise, adota-se escoamento incompressível e em regime permanente, número infinito de pás do rotor, os efeitos viscosos são desprezados, distribuição uniforme de empuxo pelo disco, esteira sem rotação e que as pressões estáticas a montante e a jusante distante do disco atuador são iguais à pressão estática do escoamento livre.

Aplicando a conservação da quantidade de movimento linear para o volume de controle e assumindo escoamento unidimensional, incompressível e em regime permanente, a força de empuxo,  $T$ , pode ser expressa como a diferença da quantidade de movimento linear do fluxo:

$$T = U_1(\rho AU)_1 - U_4(\rho AU)_4 \quad (\text{A.1})$$

onde  $\rho$  é a densidade do fluido,  $A$  corresponde a área da seção transversal,  $U$  é a velocidade do fluido e os índices indicam a seção transversal correspondente.

Para um escoamento em regime permanente, a vazão mássica,  $\dot{m}$ , se mantém constante, logo a Equação A.1 pode ser reescrita como:

$$T = \dot{m}(U_1 - U_4) \quad (\text{A.2})$$

$T$  pode ainda ser expressa pelo balanço de forças atuantes no disco atuador:

$$T = A_2(p_2 - p_3) \quad (\text{A.3})$$

Aplicando-se a Equação de Bernoulli para os volumes de controle antes e depois do disco atuador, tem-se:

$$p_1 + \frac{1}{2}\rho U_1^2 = p_2 + \frac{1}{2}\rho U_2^2 \quad (\text{A.4})$$

$$p_3 + \frac{1}{2}\rho U_3^2 = p_4 + \frac{1}{2}\rho U_4^2 \quad (\text{A.5})$$

Como a pressão distante a montante,  $p_1$ , e a jusante,  $p_4$ , do disco atuador são iguais à pressão estática do escoamento livre e assumindo que a velocidade através do disco se mantém a mesma, pode-se isolar as pressões  $p_2$  e  $p_3$  nas Equações A.4 e A.5, respectivamente, e reescrever a Equação A.3 como:

$$T = \frac{1}{2}\rho A_2(U_1^2 - U_4^2) \quad (\text{A.6})$$

Igualando as Equações A.2 e A.6 e sabendo que  $\dot{m} = \rho A_2 U_2$ :

$$U_2 = \frac{U_1 + U_4}{2} \quad (\text{A.7})$$

Assim, conclui-se que a velocidade no plano do disco, usando este modelo simplificado, é igual a média das velocidades da corrente livre e da esteira.

A potência extraída pelo disco,  $P$ , pode ser calculada ao se multiplicar a força de empuxo pela velocidade no disco:

$$P = \frac{1}{2} \rho A_2 (U_1^2 - U_4^2) U_2 = \frac{1}{4} \rho A_2 (U_1^2 - U_4^2) (U_1 + U_4) \quad (\text{A.8})$$

Define-se o coeficiente de potência,  $C_p$ , como a razão entre a potência extraída pela turbina e a potência disponível do fluxo que atravessa a área do disco:

$$C_p = \frac{P}{\frac{1}{2} \rho A_2 U_1^3} = \frac{1}{2} \left( \frac{U_1 + U_4}{U_1} \right) \left( \frac{U_1^2 - U_4^2}{U_1^2} \right) \quad (\text{A.9})$$

$$C_p = \frac{1}{2} \left( 1 + \frac{U_4}{U_1} \right) \left( 1 - \frac{U_4^2}{U_1^2} \right) \quad (\text{A.10})$$

Dada que a razão entre a velocidade na esteira e a velocidade da corrente livre seja expressa por uma variável  $v$ , a Equação A.10 pode ser reescrita em função dessa mesma variável:

$$C_p = \frac{1}{2} (1 + v)(1 - v^2) \quad (\text{A.11})$$

Derivando-se a Equação A.11 em função de  $v$  e igualando a zero, é possível encontrar o valor máximo de  $C_p$  e o ponto onde esse valor ocorre. Dadas as devidas operações matemáticas, observa-se que  $C_p$  é máximo quando  $v = 1/3$  e que o valor de  $C_p$  nesse ponto é de  $16/27 \approx 0,5926$ . Esse valor é chamado de limite de Betz e corresponde ao valor teórico máximo de coeficiente de potência para um rotor não-carenado. Fatores como rotação da esteira do rotor, número finito de pás e consideração de efeitos viscosos levam à diminuição desse valor.

## A.2. EQUAÇÕES DE CONSERVAÇÃO

### A.2.1. EQUAÇÃO DA CONTINUIDADE

De forma compacta, a equação da conservação da massa para três dimensões é dada pela Equação A.12. Também chamada de equação da continuidade por não requerer nenhuma hipótese a não ser que a massa específica e a velocidade sejam funções contínuas. Logo, o escoamento pode ser permanente ou não, viscoso ou sem atrito, ou ainda compressível ou incompressível (WHITE, 2011, P. 241).

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{U}) = 0 \quad (\text{A.12})$$

Já para um referencial não-inercial, a Equação A.12 se resume a Equação A.13, onde  $\vec{W}$  representa a velocidade da partícula em relação a um referencial não-inercial.

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{W}) = 0 \quad (\text{A.13})$$

O leitor é indicado às referências para a dedução das equações supracitadas.

### A.2.2. EQUAÇÃO DA QUANTIDADE DE MOVIMENTO LINEAR

A segunda lei de Newton quando aplicada a um volume de controle infinitesimal aponta que a taxa de aumento da quantidade de movimento linear de uma partícula de fluido é igual à soma das forças sobre essa partícula (VERSTEEG & MALALASEKERA, 1995, P. 14). Seguindo esse princípio e após manipulações matemáticas, para as quais o leitor é referenciado a White (2011) e Versteeg e Malalasekera (1995), tem-se a equação da quantidade de movimento linear para um elemento infinitesimal.

$$\rho \frac{D\vec{U}}{Dt} = \rho \vec{g} - \vec{\nabla} p + \vec{\nabla} \cdot \vec{\tau} \quad (\text{A.14})$$

Para um referencial não-inercial, considerando escoamento incompressível e representando o tensor das tensões  $\vec{\nabla} \cdot \vec{\tau}$  pelas relações de Stokes, com viscosidade  $\mu$  constante, resulta:

$$\rho \left( \frac{\partial \vec{W}}{\partial t} + \vec{W} \cdot \vec{\nabla} \vec{W} + \vec{a}^* \right) = \rho \vec{g} - \vec{\nabla} p + \mu \nabla^2 \vec{W} \quad (\text{A.15})$$

onde o termo  $\vec{a}^*$  consiste nas acelerações aparentes, dado por

$$\vec{a}^* = \ddot{\vec{R}}_0 + \dot{\vec{\omega}} \times \vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) + 2\vec{\omega} \times \vec{W} \quad (\text{A.16})$$

### A.3. CFD. MÉTODOS DE VOLUMES FINITOS

A equação de Navier-Stokes, Equação A.15, na forma diferencial, representa o princípio de conservação da quantidade de movimento para um referencial não-inercial, em que as forças estão em equilíbrio com as variações das acelerações convectiva e local. É possível obter a solução analítica para as equações de Navier-Stokes somente para uma pequena variedade de problemas, considerando várias simplificações que eliminam termos não-lineares. Pode-se citar como exemplos de problemas cuja obtenção da solução analítica é possível, o escoamento de Couette e o escoamento de Poiseuille.

No entanto, para a vasta maioria dos problemas de dinâmica dos fluidos, onde devem ser considerados os efeitos viscosos em regime transiente em geometrias complexas (3D), são necessárias outras abordagens para se obter os campos de pressão e de velocidade que satisfazem os princípios fundamentais de conservação. Para se obter solução para esses problemas, têm sido desenvolvidas, desde o século passado, técnicas de abordagem numérica que se baseiam, em geral, na discretização do domínio e de propriedades do escoamento e na solução de sistemas algébricos oriundos da aplicação das leis de conservação nesses contextos. Pode-se citar como métodos numéricos aplicados aos estudos de dinâmica dos fluidos: o Método de Painéis (com ou sem código

de interação com a camada-limite), o Método dos Vórtices Discretos, o Método de Elementos Finitos, o Método de Diferenças Finitas e o Método de Volumes Finitos.

O Método de Volumes Finitos tem como base a representação dos fenômenos físicos por suas equações diferenciais de transporte, que são, em seguida, integradas nos volumes de controle. Posteriormente, aplica-se o teorema de Gauss nas integrais de volume resultantes da divergência dos campos, de forma a transformá-las em integrais de superfície, onde as variáveis do campo de escoamento devem obedecer a certo princípio de conservação.

Aqui, generalizando a variável intensiva representativa de um campo do escoamento escalar ou vetorial pela letra  $\varphi$ , a letra grega  $\Gamma$  representa uma propriedade vinculada ao termo difusivo da equação de transporte, por exemplo, a viscosidade dinâmica no caso da equação da quantidade de movimento linear. Sendo assim, a equação de transporte na forma diferencial pode ser escrita de maneira geral:

$$\frac{\partial(\rho\varphi)}{\partial t} + \vec{\nabla} \cdot (\rho\varphi\vec{U}) = \vec{\nabla} \cdot (\Gamma\vec{\nabla}\varphi) + S_\varphi \quad (\text{A.17})$$

Integrando a Equação A.17, e considerando-se os vários volumes de controle:

$$\int_{VC} \frac{\partial(\rho\varphi)}{\partial t} + \int_{VC} \vec{\nabla} \cdot (\rho\varphi\vec{U}) = \int_{VC} \vec{\nabla} \cdot (\Gamma\vec{\nabla}\varphi) + \int_{VC} S_\varphi \quad (\text{A.18})$$

Aplica-se, então, o teorema de *Gauss* nas integrais contendo as divergências de campos, transformando-as em integrais de superfície referentes aos fluxos das propriedades que atravessam a superfície de controle, resultando:

$$\frac{\partial}{\partial t} \left( \int_{VC} \rho\varphi dV \right) + \int_{SC} (\rho\varphi\vec{U}) \cdot \vec{n} dA = \int_{SC} (\Gamma\vec{\nabla}\varphi) \cdot \vec{n} dA + \int_{VC} S_\varphi dV \quad (\text{A.19})$$

A Equação A.19 representa o princípio de conservação da variável na forma integral aplicada a um volume de controle. Assim, o sistema de equações diferenciais parciais oriundo dos princípios de conservação é transformado em um sistema algébrico linear, que deve ser resolvido

por métodos numéricos diretos, como a eliminação de *Gauss* ou a decomposição LU. Na Equação A.19 podem ser identificados os seguintes termos: variação temporal no volume de controle da variável  $\varphi$  somado à variação convectiva no sistema de controle, igual à variação difusiva no sistema de controle e, finalmente, a variação do termo fonte no volume de controle.

#### A.4. CFD. MODELO DE TURBULÊNCIA $k-\omega$ SST

O modelo de turbulência  $k-\omega$  SST foi proposto por Menter (1994) com o objetivo de superar as deficiências características dos modelos  $k-\varepsilon$  e  $k-\omega$ . Para o primeiro modelo, pode-se ressaltar como principais deficiências a sobre predição das tensões cisalhantes na região externa da camada-limite, a necessidade de modelagem da viscosidade turbulenta na proximidade de paredes e a dificuldade de modelo em simular escoamentos que apresentem gradientes adversos de pressão. Para o último, pode-se destacar a alta sensibilidade à presença de fronteiras livres (NETO, 2020, P. 431).

A fim de superar as deficiências de ambos os modelos, Menter (1994) propôs a fusão dos dois modelos de forma que a modelagem do escoamento nas proximidades de paredes fosse feita utilizando o modelo  $k-\omega$  e em regiões afastadas das paredes fosse utilizado o modelo  $k-\varepsilon$ . A hibridação é feita por meio de uma função  $F_1$  que varia de 0 em regiões distantes das paredes a 1 em regiões próximas das paredes, possibilitando, assim, um balanço entre os dois modelos em regiões intermediárias e a utilização, exclusiva, de cada modelo na condição mais indicada, respectivamente.

Abaixo é apresentado o equacionamento do modelo:

$$\frac{\partial \rho k}{\partial t} + U_j \frac{\partial \rho k}{\partial x_j} = \tau_{ij} \frac{\partial U_i}{\partial x_j} - \beta^* \rho \omega k + \frac{\partial}{\partial x_j} \left[ (\mu + \sigma_k \mu_t) \frac{\partial k}{\partial x_j} \right] \quad (\text{A.20})$$

$$\begin{aligned} \frac{\partial \rho \omega}{\partial t} + U_j \frac{\partial \rho \omega}{\partial x_j} = & \frac{\gamma}{\nu_t} \tau_{ij} \frac{\partial U_i}{\partial x_j} - \beta \rho \omega^2 + \frac{\partial}{\partial x_j} \left[ (\mu + \sigma_\omega \mu_t) \frac{\partial \omega}{\partial x_j} \right] \\ & + 2(1 - F_1) \rho \sigma_{\omega 2} \frac{1}{\omega} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} \end{aligned} \quad (\text{A.21})$$

onde

$$\nu_t = \frac{k}{\omega} \quad (\text{A.22})$$

$$\tau_{ij} = \mu_t \left( \frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} - \frac{2}{3} \frac{\partial U_k}{\partial x_k} \delta_{ij} \right) - \frac{2}{3} \rho k \delta_{ij} \quad (\text{A.23})$$

$\delta_{ij}$  é o chamado Delta de Kronecker.

Os demais termos não representados pelo equacionamento são constantes do modelo.

Por fim, é importante destacar que as condições de contorno do modelo recomendadas por Menter (1994) são condicionadas à utilização de um refinamento de malha de  $y^+ < 3$ .

## ANEXO B - SUPERFÍCIE DE RESPOSTA CLÁSSICA

A metodologia de superfície de resposta é empregada em problemas nos quais se deseja otimizar uma resposta de interesse que é influenciada por diversas variáveis. O método consiste em um grupo de técnicas matemáticas e estatísticas para a modelagem e para a análise do problema (MONTGOMERY, 2009, P. 490).

Na maioria das vezes não se sabe a natureza da relação entre as variáveis de interesse e a resposta analisada. Assim, a finalidade do método de superfície de resposta é encontrar uma aproximação para essa relação, seja ela descrita por uma função linear, ou no caso de suspeita de curvatura no sistema, por polinômios de maior ordem, como por exemplo o modelo de segunda ordem mostrado na Equação B.1, onde  $y$  é a resposta de interesse,  $x_i$  e  $x_j$  representam as variáveis independentes,  $\beta_i$ ,  $\beta_{ii}$  e  $\beta_{ij}$  são os coeficientes a serem estimados,  $k$  é o número de variáveis independentes e  $\epsilon$  representa o erro experimental associado.

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \epsilon \quad (\text{B.1})$$

É pouco provável que o modelo polinomial seja uma boa aproximação da relação entre as variáveis de interesse e a resposta para todo o intervalo, porém para uma pequena região, em especial nas proximidades do ponto ótimo, o modelo se comporta muito bem (MONTGOMERY, 2009, P. 490).

## ANEXO C - REDES NEURAI ARTIFICIAIS

Redes neurais artificiais são redes computacionais criadas com o objetivo de simular, de maneira grosseira, o funcionamento de redes biológicas de neurônios. O uso de redes neurais artificiais possibilita a solução de problemas complexos que envolvem não-linearidade, problemas matematicamente mal definidos e problemas de natureza estocástica por meio do uso de operações matemáticas simples como adição, multiplicação e estruturas lógicas (GRAUPE, 2007, P. 1).

Dessa forma, segundo Graupe (2007), uma das vantagens de uma rede neural artificial está relacionada à sua simplicidade computacional aliada à capacidade de adaptação. A Figura C.1 mostra um esquema analógico de uma célula neural, onde  $w_i$  são os pesos de cada ligação. Dentro da célula neural são feitas operações matemáticas simples que relacionam as entradas com as saídas.

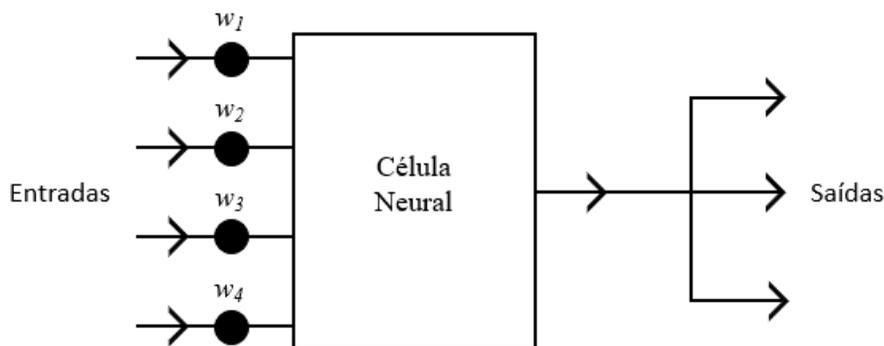


Figura C.1 – Esquema analógico de uma célula neural biológica.

Fonte: Adaptado de Graupe, 2007, p. 7

Como pode-se notar pela Figura C.2, cada interconexão recebe um peso diferente, que está relacionado com prioridade que aquela ligação possui. Maiores pesos significam maiores prioridades. Os pesos são atribuídos por meio do treinamento da rede neural, assim, diferentes problemas geram diferentes pesos para as mesmas ligações, evidenciando-se a capacidade de adaptação do sistema.

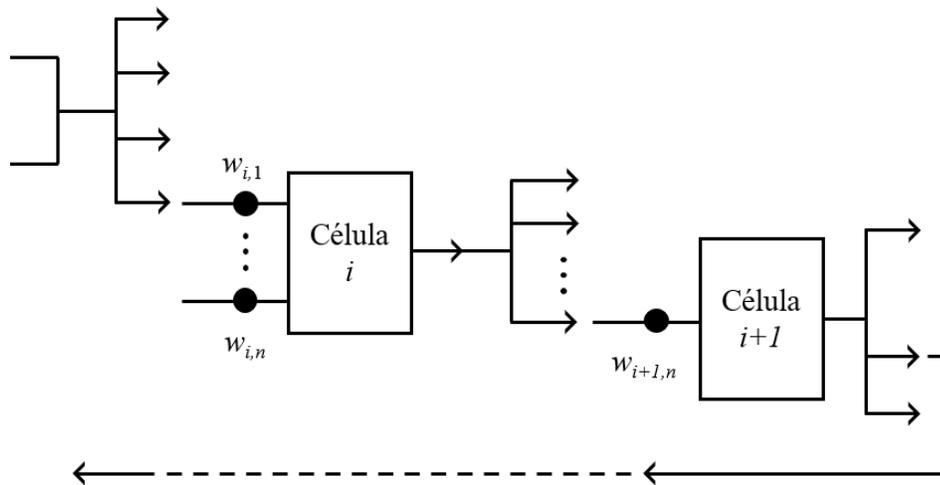


Figura C.2 – Esquema analógico de uma rede neural biológica.

Fonte: Adaptado de Graupe, 2007, p. 7

Quando operando em um ambiente não-estacionário, isto é, onde os parâmetros mudam com o tempo, uma rede neural pode também ser projetada para mudar seus pesos em tempo real (HAYKIN, 1999, P. 25).

### C.1 ARQUITETURAS DE REDE: *FEEDFORWARD* MULTICAMADA

A estrutura de uma rede neural artificial pode ser organizada em forma de camadas. No caso de uma arquitetura multicamada, tem-se a presença de uma camada de entrada, seguida por uma ou mais camadas ocultas, que são ligadas à camada de saída. Segundo Haykin (1999), a camada oculta faz a intermediação entre a camada de entradas externas e a camada de saída da rede através de funções de transferência. E pela adição de uma ou mais camadas ocultas, a rede é capaz de extrair estatísticas de maior ordem. A figura C.3 representa uma rede *feedforward* com uma única camada oculta.

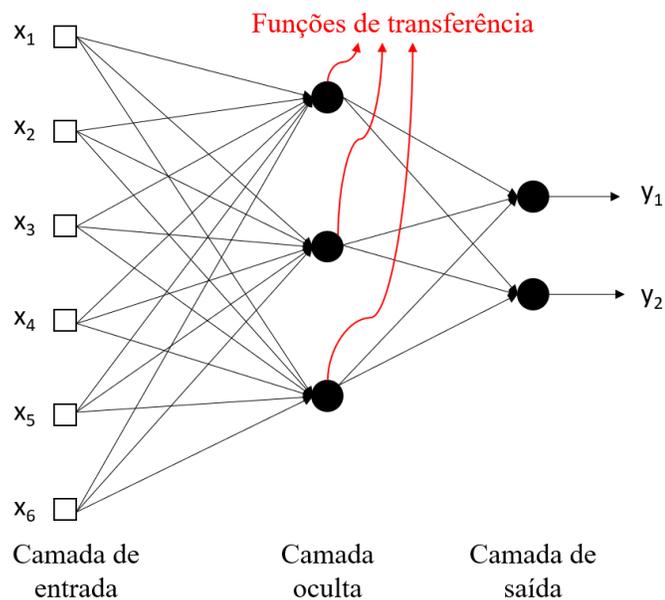


Figura C.3 – Rede *feedforward* com uma camada oculta e uma camada de saída.

Fonte: Adaptado de Haykin, 1999, p. 44

## C.2 APRENDIZAGEM ASSISTIDA E AS TAREFAS DE APRENDIZAGEM

Uma vez definida a arquitetura da rede neural artificial, a atenção se volta, então, para os processos de aprendizagem, mais especificamente para o processo de aprendizagem assistida. Neste tipo de processo, um conjunto de exemplos de entrada-saída é apresentado para a rede, cujos parâmetros são, assim, ajustados com o objetivo de melhor aproximar a saída já conhecida, dada pelo exemplo em questão. O treinamento ocorre iterativamente até que a diferença, também chamada de erro, entre a resposta obtida pela rede e a resposta já conhecida diminua até um limite estipulado (HAYKIN, 1999, P. 85). Pode-se dizer que a rede procura emular o professor, representado pela resposta conhecida. Assim, o conhecimento é transmitido do professor para a rede, e, quando esta condição é totalmente atingida, o professor é dispensado e a rede passa a lidar com o ambiente por ela mesma.

A escolha do tipo de processo de aprendizagem está relacionada ao tipo de tarefa à qual se espera que a rede seja capaz de realizar. Dessa forma, uma rede pode ser treinada para o reconhecimento de padrões, para a classificação e filtragem de dados, para o controle de sistemas ou ainda para a aproximação de funções. No presente trabalho, as redes neurais são usadas para a aproximação de funções, sejam elas conhecidas ou de difícil identificação.

Nesse caso, a rede é projetada com o intuito de aproximar uma função  $f(\cdot)$  desconhecida, de forma que a função  $F(\cdot)$ , que mapeia a relação entre entradas e saídas e calculada pela rede neural, seja próxima o suficiente de  $f(\cdot)$  no sentido Euclidiano para todas as entradas (HAYKIN, 1999, P. 91), isto é:

$$\|F(x) - f(x)\| < \epsilon \quad \text{para todo } x \quad (\text{C.1})$$

Onde  $\epsilon$  é um número positivo pequeno. Dado que o conjunto de treinamento seja grande o suficiente para garantir à rede um número adequado de parâmetros livres, então o erro de aproximação  $\epsilon$  pode se tornar pequeno o suficiente para a tarefa.

### C.3 REDES NEURAIIS DE FUNÇÕES DE BASE RADIAL

A arquitetura de uma rede neural de funções de base radial envolve, basicamente, três camadas com papéis inteiramente diferentes. A camada de entrada é responsável por conectar à rede neural ao seu ambiente, isto é, a fonte de estímulos. A segunda camada, a única oculta da rede, aplica uma transformação não-linear do espaço de entrada para o espaço da camada oculta, de alta dimensionalidade. A camada de saída, por sua vez, é linear e resulta na resposta da rede à entrada em questão (HAYKIN, 1999, P. 278).

Dessa forma, o processo de aproximação de uma função se dá por uma abordagem de alta dimensionalidade. Segundo Haykin (1999), pode-se dizer que o processo de aprendizagem é equivalente a encontrar uma superfície no espaço multidimensional que produz o melhor ajuste ao conjunto de treinamento. Logo, a generalização, ou previsão, feita pela rede é o equivalente ao uso dessa superfície multidimensional para interpolar um conjunto de teste, ou de dados cuja resposta é desconhecida.

Considerando uma rede com uma camada de entrada, uma camada oculta e uma camada de saída com um único neurônio. Seja  $m_0$  a dimensão do espaço de entrada. Assim, a rede neural representa o mapeamento do espaço de entrada  $m_0$ -dimensional para o espaço de saída, unidimensional.

$$s: \mathbb{R}^{m_0} \rightarrow \mathbb{R}^1 \quad (\text{C.2})$$

A função  $s$  pode ser compreendida como uma hipersuperfície  $\Gamma \subset \mathbb{R}^{m_0+1}$ .

Dessa forma, segundo Broomhead e Lowe (1988), o processo de treinamento da rede pode ser visto como a otimização do procedimento de ajuste da superfície  $\Gamma$  com base no conjunto de dados apresentado para a rede na forma de exemplos de entrada-saída. E o processo de generalização, como uma aproximação dos pontos desejados, com base na superfície gerada pelo processo de treinamento.

Segundo a teoria de interpolação multivariável em espaços de alta dimensionalidade (Davis, 1963), dado um conjunto de  $N$  diferentes pontos  $\{\mathbf{x}_i \in \mathbb{R}^{m_0} \mid i = 1, 2, \dots, N\}$  e um conjunto correspondente de  $N$  números reais  $\{d_i \in \mathbb{R}^1 \mid i = 1, 2, \dots, N\}$ , existe uma função  $F: \mathbb{R}^N \rightarrow \mathbb{R}^1$  que satisfaça a condição de interpolação:

$$F(\mathbf{x}_i) = d_i, \quad i = 1, 2, \dots, N \quad (\text{C.3})$$

Assim, de acordo com Powell (1987), o uso de funções de base radial, FBR, se baseia na escolha de uma função  $F$  que tenha a forma:

$$F(\mathbf{x}) = \sum_{i=1}^N w_i \varphi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (\text{C.4})$$

onde  $\{\varphi(\|\mathbf{x} - \mathbf{x}_i\|) \mid i = 1, 2, \dots, N\}$  representa um conjunto arbitrário de  $N$  funções de base radial, tal que  $\|\cdot\|$  denota uma norma, geralmente Euclidiana.

Introduzindo as condições de interpolação dadas pela Equação C.3 na Equação C.4 resulta em um conjunto de equações lineares simultâneas para os coeficientes, ou pesos,  $w_i$ :

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (\text{C.5})$$

onde

$$\begin{aligned}
\varphi_{ji} &= \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|), \quad (j, i) = 1, 2, \dots, N \\
\mathbf{d} &= [d_1, d_2, \dots, d_N]^T \\
\mathbf{w} &= [w_1, w_2, \dots, w_N]^T
\end{aligned} \tag{C.6}$$

Os vetores  $N \times 1$   $\mathbf{d}$  e  $\mathbf{w}$  representam o vetor de resposta desejada e o vetor linear de pesos, respectivamente, onde  $N$  é o tamanho do conjunto de treinamento. Se  $\Phi$  denota uma matriz  $N \times N$  com os elementos  $\varphi_{ji}$ , tal que:

$$\Phi = \{\varphi_{ji} | (j, i) = 1, 2, \dots, N\} \tag{C.7}$$

Então, essa matriz pode ser chamada de matriz de interpolação e a Equação C.5 pode ser reescrita de forma compacta

$$\Phi \mathbf{w} = \mathbf{d} \tag{C.8}$$

Assim, assumindo que  $\Phi$  é não-singular e, conseqüentemente, a matriz inversa  $\Phi^{-1}$  existe, pode-se resolver a Equação C.8 para o vetor de pesos  $\mathbf{w}$

$$\mathbf{w} = \Phi^{-1} \mathbf{d} \tag{C.9}$$

A premissa de que  $\Phi$  é não-singular é garantida pelo Teorema de Micchelli (1986), que afirma que dado  $\{\mathbf{x}_i\}_{i=1}^N$ , um conjunto de pontos distintos em  $\mathbb{R}^{m_0}$ , então a matriz de interpolação  $N \times N$ , cujo  $ji$ -elemento é dado por  $\varphi_{ji} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|)$ , é não-singular. Esse Teorema se aplica a uma grande classe de funções de base radial, como a Multiquadrática, Multiquadrática Inversa e Gaussiana, desde que os pontos  $\{\mathbf{x}_i\}_{i=1}^N$  sejam todos distintos. Haykin (1999) afirma ainda que esse é o único critério que deve ser respeitado para que a matriz de interpolação  $\Phi$  seja não-singular, independentemente do tamanho  $N$  do conjunto de pontos ou da dimensionalidade  $m_0$  do vetor de pontos  $\mathbf{x}_i$ .

## C.4 TEOREMA UNIVERSAL DA APROXIMAÇÃO

Seja  $G: \mathbb{R}^{m_0} \rightarrow \mathbb{R}$  uma função restrita integrável tal que  $G$  seja contínua e

$$\int_{\mathbb{R}^{m_0}} G(\mathbf{x}) d\mathbf{x} \neq 0 \quad (\text{C.10})$$

E  $\mathcal{S}_G$  a família de funções de base radial  $F: \mathbb{R}^{m_0} \rightarrow \mathbb{R}$  representada por

$$F(\mathbf{x}) = \sum_{i=1}^{m_1} w_i G\left(\frac{\mathbf{x} - \mathbf{t}_i}{\sigma}\right) \quad (\text{C.11})$$

onde  $\sigma > 0$ ,  $w_i \in \mathbb{R}$  e  $\mathbf{t}_i \in \mathbb{R}^{m_0}$  para  $i = 1, 2, \dots, m_1$ . Assim, segundo Park e Sandberg (1991), para qualquer função  $f(\mathbf{x})$  que mapeie entradas e saídas, existe uma rede de funções de base radial, de centros  $\{\mathbf{t}_i\}_{i=1}^{m_1}$  e largura  $\sigma > 0$ , cuja função de mapeamento de entrada-saída  $F(\mathbf{x})$  seja próxima de  $f(\mathbf{x})$  na norma  $L_p$ ,  $p \in [1, \infty]$ .

## C.5 ESTRATÉGIAS DE APRENDIZADO E *EXTREME LEARNING MACHINE*

Redes neurais artificiais com base em FBR, normalmente, apresentam três parâmetros a serem calculados: os pesos das ligações entre a camada de entrada e a camada oculta,  $\mathbf{u}_{ij}$ , que fisicamente representam os centros das FBR; a largura  $\mu$  das FBR e os pesos das ligações entre a camada oculta e a camada de saída,  $w_{jk}$ .

Existem diferentes estratégias de aprendizado que podem ser utilizadas a fim de determinar os parâmetros supracitados, a depender da forma como os centros das FBR são especificados. Dentre as estratégias, a utilizada no presente trabalho consiste em fixar os centros aleatoriamente pelo conjunto de dados de treinamento. Segundo Lowe (1989), esta estratégia pode ser considerada uma abordagem sensível o suficiente, uma vez que os dados de treinamento estejam distribuídos de uma forma representativa para o problema em questão.

Assim, Huang e Siew (2004) propuseram um algoritmo de aprendizado chamado de *Extreme Learning Machine* (ELM). Os centros  $\mathbf{t}_i$  das FBR são escolhidos aleatoriamente, e a largura das FBR é escolhida por uma análise de sensibilidade, baseada na distância entre os pontos no conjunto de treinamento e na distribuição dos pontos no espaço.

As funções de transferência da camada oculta são Gaussianas, dadas por

$$G(\mathbf{t}_i, \mu, \mathbf{x}) = \exp\left(-\frac{1}{\mu} \|\mathbf{x} - \mathbf{t}_i\|^2\right), \quad i = 1, 2, \dots, N_0 \quad (\text{C.12})$$

Sendo  $N_0$  o número de pontos do conjunto de treinamento escolhidos como centros.

Dessa forma, o único parâmetro restante a ser aprendido através do algoritmo é composto pelos pesos entre a camada oculta e a camada de saída, calculado pela Equação C.9. Para o algoritmo de ELM, como uma forma mais geral de se evitar o problema de inversão da matriz  $\Phi$ , é utilizado um método de pseudoinversão chamado de decomposição em valores singulares

$$\mathbf{w} = \Phi^\dagger \mathbf{x} \quad (\text{C.13})$$

onde a matriz  $\Phi^\dagger$  é a pseudoinversa da matriz  $\Phi$ .

Logo, para a aproximação de algum conjunto de pontos basta que se calcule a matriz  $\mathbf{G}$  nos pontos desejados, e se multiplique pelos pesos  $\mathbf{w}$  já calculados.

$$\mathbf{Y}_{approx} = G(\mathbf{t}_i, \mu, \mathbf{x})\mathbf{w} \quad (\text{C.14})$$

## REFERÊNCIAS

ACKLEY, D. A connectionist machine for genetic hillclimbing. Vol. 28, Springer science & business media, 2012.

AMRIT, A.; LEIFSSON, L.; KOZIEL, S.; TESFAHUNEGN, Y. A. Efficient multi-objective aerodynamic optimization by design space dimension reduction and co-Kriging. 17<sup>th</sup> AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2016.

BAGHERI-SADEGHI, N.; HELENBROOK, B. T.; VISSER, K. D. Ducted wind turbine optimization and sensitivity to rotor position. Wind Energy Science, Volume 3, p. 221-229, 2018.

BRAVO, R.; TULLIS, S.; ZIADA, S. Performance testing of a small vertical-axis wind turbine. Proceedings of the 21<sup>st</sup> Canadian Congress of Applied Mechanics, CANCAM07, Mechanical Engineering Department, McMaster University, Canada, 2007.

BROOMHEAD, D. S.; LOWE, D. Multivariable functional interpolation and adaptive networks. Complex Systems Publications, Inc. 1988.

CORSO, L. L. Otimização baseada em confiabilidade utilizando métodos com gradientes e de busca global para problemas de análise estrutural. Tese (Doutorado em Engenharia), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.

CURRIN, C.; MITCHELL, T.; MORRIS, M.; YLVISAKER, D. A Bayesian approach to the design and analysis of computer experiments. Technical Report 6498. Oak Ridge National Laboratory, 1988.

DAVIS, P. J. Interpolation and Approximation. Blaisdell Publ. Co. New York. 1963.

FORRESTER, A. I.; SÓBESTER, A.; KEANE, A. J. Multifidelity Optimization Via Surrogate Modelling. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 463, 3251-3269, 2007.

FORRESTER, A. I. J.; SÓBESTER, A.; KEANE, A. J. Engineering Design via Surrogate Modelling, A Practical Guide. John Wiley & Sons Ltd. 2008.

GARNETT, R. Bayesian Optimization. Cambridge University Press, 2023.

- GHOREISHI, S. F.; ALLAIRE, D. Multi-information source constrained Bayesian optimization. *Structural Multidisciplinary Optimization*, 59, 977-991, 2019.
- GRAUPE, D. Principles of artificial neural networks, Advanced series on circuits and systems, vol. 6. World scientific publishing, 2007.
- HAN, Z-H.; GÖRTZ, S. Hierarchical Kriging model for variable-fidelity surrogate modeling. *AIAA Journal*, Vol. 50, No. 9, 2012.
- HARTMAN, J. K. Some experiments in global optimization. *Naval Research Logistics Quarterly*, Vol. 20, issue 3, p. 569-576, 1973.
- HAYKIN, S. Neural Networks a comprehensive foundation. McMaster University. PAErson Education. Second Edition. 1999.
- HUANG, G.; SIEW, C. K. Extreme learning machine: a new learning scheme of feedforward neural networks. *Proceedings of International Joint Conference on Neural Networks*. Budapest, 2004.
- HUANG, D.; ALLEN, T. T.; NOTZ, W. I.; MILLER, R. A. Sequential Kriging optimization using multiple-fidelity evaluations. *Structural Multidisciplinary Optimization*, 32, 369-382, 2006.
- HU, J.; ZHOU, Q.; JIANG, P.; XIE, T. An improved hierarquical Kriging for variable-fidelity surrogate modeling. *International Conference on Cybernetics, Robotics and Control*, 2016.
- IEA, Agência Internacional de Energia. Disponível em: <<https://www.iea.org/reports/electricity-2024/executive-summary>>. Acesso em: 16 maio 2024.
- JAMIESON, P. M. Beating Betz: Energy Extraction Limits in a Constrained Flow Field. *Journal of Solar Energy Engineering*, Volume 131, ISSN 0199-6231, 2009.
- JASA, J.; BORTOLOTTI, P.; ZALKIND, D.; BARTER, G. Effectively using multifidelity optimization for wind turbine design. *Wind Energy Science*, Volume 7, p. 991-1006, 2022.
- JIANG, P.; ZHOU, Q.; SHAO, X. Surrogate model-based engineering design and optimization. Springer nature, ISSN 2195-9862, 2020.
- JONES, D. R. A Taxonomy of Global Optimization Methods Based on Response Surfaces. *Journal of Global Optimization*, 21, 345-383, 2001.

- JONES, D. R.; SCHONLAU, M.; WELCH, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13, 455-492, 1998.
- LAM, R.; ALLAIRE, D.; WILLCOX, K. Multifidelity optimization using statistical surrogate modeling for non-hierarchical information sources. 56<sup>th</sup> AIIA/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 2015.
- LEE, H. M.; KWON, O. J. Performance improvement of horizontal axis wind turbines by aerodynamic shape optimization including aeroelastic deformation. *Renewable Energy*, Volume 147, Part 1, 2020.
- LI, S.; XING, W.; KIRBY, R. M.; ZHE, S. Multi-fidelity Bayesian optimization via deep neural networks. 34<sup>th</sup> Conference on Neural Information Processing Systems. Canada, 2020.
- LOWE, D. Adaptive radial basis function nonlinearities and the problem of generalization. *Proceedings of the 1<sup>st</sup> IEE International Conference on Artificial Neural Networks*, pp: 171-175, 1989.
- MANWELL, J. F.; MCGOWAN, J. G.; ROGERS, A. L. *Wind energy explained: theory design and application*. 2 ed, Wiley and Sons, 2009.
- MENG, D.; YANG, S.; DE JESUS, A. M. P.; ZHU, S-P. A novel Kriging-model-assisted reliability-based multidisciplinary design optimization strategy and its application in the offshore wind turbine tower. *Renewable Energy*, Volume 203, Pages 407-420, ISSN 0960-1481, 2023.
- MENTER, F. R. Two-equation Eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, Vol. 32, p. 1598-1605, 1994.
- MICHELLI, C. A. Interpolation of scattered data: distance matrices and conditionally positive definite functions. *Constructive Approximation*, 2:11-22, 1986.
- MONTGOMERY, D. G. *Design and Analysis of Experiments*. Arizona State University. John Wiley & Sons, Inc. Ninth edition. 2009.
- MORRIS, M. D.; MITCHELL, T. J. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, 43, 381-402, 1995.
- NETO, A. DA S. *Escoamentos Turbulentos: análise física e modelagem teórica*. 1<sup>a</sup> edição, ISBN 978-65-990365-1-4, editora Composer, 2020.

NETO, J. X. V.; JUNIOR, E. J. G.; MORENO, S. R.; AYALA, H. V. H.; MARIANI, V. C.; COELHO, L. S. Wind turbine blade geometry design based on multi-objective optimization using metaheuristics. *Energy*, Volume 168, 2018.

NOË, U.; HUSMEIER, D. On a new improvement-based acquisition function for bayesian optimization. arXiv:1808.06918v, 2018.

PARK, J.; SANDBERG, I. W. Universal approximation using radial-basis-function networks. *Neural Computation*, 3(2), 246-257. 1991.

PHOLDEE, N.; BUREERAT, S.; NUANTONG, W. Kriging surrogate-based genetic algorithm optimization for blade design of a horizontal axis wind turbine. *Computer Modeling in Engineering & Sciences*, 126(1), 261-273, 2021.

POWELL, M. J. D. Radial basis function approximation to polynomials. *Numerical Analysis Proceedings*. Dundee, UK, 1987.

REZEK, T. J.; CAMACHO, R. G. R.; MANZANARES FILHO, N. A novel methodology for the design of diffuser-augmented hydrokinetic rotors. *Renewable Energy*, Volume 210, p. 524-539, ISSN 0960-1481, 2023.

REZEK, T. J.; CAMACHO, R. G. R.; MANZANARES FILHO, N.; LIMACHER, E. J. Design of a Hydrokinetic Turbine Diffuser Based on Optimization and Computational Fluid Dynamics. *Applied Ocean Research*, Volume 107, ISSN 0141-1187, 2021.

RUAN, X.; JIANG, P.; ZHOU, Q.; HU, J.; SHU, L. Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems. *Structural and Multidisciplinary Optimization*, Springer, 2020.

SHU, L.; JIANG, P.; WANG, Y. A multi-fidelity Bayesian optimization approach based on the expected further improvement. *Structural and Multidisciplinary Optimization*, Springer, 2021.

SONG, K.; ZHANG, Y.; ZHUANG, X.; YU, X.; SONG, B. Reliability-based design optimization using adaptive surrogate model and importance sampling-based modified SORA method. *Engineering with Computers* 37, 1295–1314, 2021.

SRINIVAS, N.; DEB, K. Multi-objective function optimization using non-dominated sorting genetic algorithms. *Evolutionary Computation Journal* 2(3), 221-248, 1994.

- TAKENO, S.; FUKUOKA, H.; TSUKADA, Y.; KOYAMA, T.; SHIGA, M.; TAKEUCHI, I.; KARASUYAMA, M. Multi-fidelity Bayesian optimization with max-value entropy search. 2020.
- THEIL, H. Principles of Econometrics. John Wiley, New York. 1971.
- VERSTEEG, H. K.; MALALASEKERA, W. An Introduction to Computational Fluid Dynamics: The finite volume method. Nova York: John Wiley & Sons Inc., 1995.
- WANG, Y. M.; RÉTHORÉ, P-E.; VAN DER LAAN, M. P.; MURCIA LEON, J. P.; LIU, Y. Q.; LI, L. Multi-fidelity wake modelling based on co-Kriging method. Journal of Physics: Conference Series, 753, 2016.
- WANG, Y.; HAN, Z-H.; ZHANG, Y.; SONG, W-P. Efficient global optimization using multiple infill sampling criteria and surrogate models. AIAA Aerospace Sciences Meeting, AIAA SciTech Forum, 2018.
- WANG, Z.; TU, Y.; ZHANG, K.; HAN, Z.; CAO, Y.; ZHOU, D. An optimization framework for wind farm layout design using CFD-based Kriging model. Ocean Engineering, Volume 293, ISSN 0029-8018, 2024.
- WHITE, F. M. Mecânica dos Fluidos. 6 ed. Porto Alegre, AMGH Editora, 2011.
- XIONG, S.; QIAN, P. Z.; WU, C. J. Sequential design and analysis of high-accuracy and low-accuracy computer codes. Technometrics, 55(1), 37-46, 2013.
- XU, J.; HAN, Z.; SONG, W.; LI, K. Efficient aerodynamic optimization of propeller using hierarchical Kriging models. Journal of Physics: Conference Series, 1519, 2020.
- XU, Z.; GUO, Y.; SALEH, J. H. Efficient hybrid bayesian optimization algorithm with adaptive expected improvement acquisition function. Engineering Optimization, Vol. 53, 1786-1804, 2021.
- ZHANG, Y.; HAN, ZH.; ZHANG, KS. Variable-fidelity expected improvement method for efficient global optimization of expensive functions. Struct Multidisc Optim, Vol. 58, 1431–1451, 2018.
- ZHENG, Y.; MA, H.; WEI, J.; ZHU, K. Robust optimization for composite blade of wind turbine based on kriging model. Advanced Composites Letters, Volume 29, 1-8, 2020.
- ZHOU, Q.; WU, Y.; GUO, Z.; HU, J.; JIN, P. A generalized hierarchical co-Kriging model for multi-fidelity data fusion. Struct Multidisc Optimization, Vol. 62, 1885–1904, 2020.

