

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO

Guilherme Freitas Coelho

Otimização para Simulação com Krigagem: Uma Aplicação em
Alocação de Ambulâncias

Belo Horizonte

2015

Guilherme Freitas Coelho

**Otimização para Simulação com Krigagem: Uma Aplicação em
Alocação de Ambulâncias**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Minas Gerais, para a obtenção de Título de Mestre em Engenharia de Produção.

Orientador: Luiz Ricardo Pinto

Belo Horizonte

2015

Freitas Coelho, Guilherme.

Otimização para Simulação com Krigagem: Uma Aplicação em Alocação de Ambulâncias

84 páginas

Dissertação (Mestrado) - Escola de Engenharia da Universidade Federal de Minas Gerais. Belo Horizonte. Departamento de Engenharia de Produção.

1. Krigagem
2. Metamodelagem
3. Otimização para Simulação
4. Saúde

I. Universidade Federal de Minas Gerais. Escola de Engenharia. Departamento de Engenharia de Produção.

Comissão Julgadora:

Prof. Dr.

Marcelo Azevedo Costa

Prof. Dr.

Ricardo Poley Martins Ferreira

Prof. Dr.

Luiz Ricardo Pinto

Você não pode impedir as mudanças, assim como não pode impedir o pôr do sol.

Shmi Skywalker, Guerra nas Estrelas Episódio I: A Ameaça Fantasma

Agradecimentos

Agradeço ao Prof. Luiz Ricardo pela sua valiosa orientação nesses dois anos de mestrado, sempre acreditando no meu potencial e na qualidade do meu trabalho.

Também agradeço aos meus amigos e colegas de mestrado e doutorado, Luiza Real, Alexandre Fortes, Elisangela Martins, Karolly Santos, Aline Veronese, Bárbara Oliveira, João Flávio, Hendrigo Batista, Virginia Casagrande, Álvaro Ferreira e Rejane Corrêa, que me ajudaram bastante no dia-a-dia de pós-graduação. Acredito que novas grandes amizades foram consolidadas.

Agradeço ao apoio dos meus pais, Marco e Elza, e meu irmão, Daniel, pois sem eles eu não seria o que sou e nem estaria nesta posição.

Agradeço em especial à minha namorada, Grazianni, que sempre me incentivou e suportou anos de distância para que este momento pudesse ser alcançado.

Por fim, agradeço à CAPES pela bolsa de mestrado e à UFMG (professores e servidores) pela recepção maravilhosa e auxílio.

Resumo

Metamodelagem é um tema bastante comum na literatura de Otimização para Simulação. Sua aplicabilidade é voltada para a otimização de funções definidas sobre simuladores ou modelos de simulação, de forma que a avaliação de um ponto desconhecido demanda considerável esforço computacional. O uso de metamodelos tem o objetivo de fazer estimações do valor real (simulado) antes mesmo que este ponto seja avaliado pelo modelo de simulação, entretanto a maioria das publicações não aplica o método para modelos de tamanho e complexidade reais. Nesse sentido, esta dissertação de mestrado buscou aplicar metamodelos de Krigagem para a minimização do tempo de resposta do Serviço de Atendimento Móvel de Urgência (SAMU) de Belo Horizonte, a partir da alocação de ambulâncias nas bases da cidade. Krigagem é considerada o estado-da-arte em metamodelagem, pois fornece, além da estimativa em um novo ponto, a informação de incerteza de estimativa (variância de estimativa), proporcional à covariância entre as amostras do seu conjunto de treinamento. O processo de otimização seguiu o algoritmo *Efficient Global Optimization* (EGO), que explora o metamodelo de Krigagem com o critério de desempenho *Expected Improvement* (EI), sendo que, para se trabalhar no âmbito estocástico, foi empregado o Procedimento de Reiterpolação (RI) e um novo modo de estimativa, chamado de KOIC, foi proposto, pela motivação de se levar em conta todo o intervalo de confiança da variável de resposta. Para fazer a alocação das ambulâncias, uma heurística *Simulated Annealing* foi especificada de forma a lidar com suas variáveis discretas. Por fim, RI e KOIC foram comparados entre si e a melhor técnica foi utilizada na obtenção de uma curva que refletiu a relação entre o mínimo tempo de resposta e o número total de ambulâncias alocadas à cidade, informação esta bastante relevante para o projeto e gerenciamento de sistemas públicos de atendimento em Saúde.

Palavras-chave: Krigagem; Metamodelagem; Otimização pra Simulação; Saúde.

Abstract

Metamodeling is a common subject in Optimization for Simulation literature. Its applicability is focused on the optimization of functions defined over simulators or simulation models, so that the evaluation of an unknown point requires considerable computational effort. The use of metamodels aims to estimate the actual value (simulated) even before the point is evaluated by the simulation model. However, most publications do not apply the method to models with real world complexity and size. This dissertation sought to apply Kriging to minimize the response time of the Serviço de Atendimento Móvel de Urgência (SAMU) of Belo Horizonte, while allocating ambulances throughout all city bases. Kriging is considered the state-of-art technique in metamodeling as it provides, in addition to the new point estimation, the level of prediction uncertainty (estimation variance), which is proportional to the covariance between samples of its training set. The optimization process followed the Efficient Global Optimization algorithm (EGO), which explores Kriging by using the performance criterion Expected Improvement (EI), and, in the stochastic case, it was used the Reinterpolation Procedure (RI). Also, a new estimation criterion, called KOIC, was proposed with the motivation of taking into account the whole response variable confidence interval. To allocate the ambulances, a Simulated Annealing heuristic has been specified in order to deal with the discrete variables of the model. Finally, RI and KOIC were compared and the best technique was used to obtain a curve that reflected the relationship between the minimum response time and the total number of ambulances allocated to the city, a very relevant information to health-care public systems managers and designers.

Keywords: Kriging; Metamodeling; Optimization for Simulation; Health-Care.

Lista de Figuras

2.1	Funções de correlação utilizadas em Krigagem para metamodelagem	27
2.2	Comportamento da variância de estimação de Krigagem.	30
2.3	Comparação de ajustes entre Krigagem Ordinária e Krigagem Ordinária estocástica.	31
2.4	Exemplo: Amostras coletadas e suas estimações por validação cruzada para análise de resíduo	35
2.5	Análise de resíduos: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	35
3.1	Fluxograma do processo geral de otimização para simulação	37
3.2	Passos do algoritmo EGO. Baseado em Sasena (2002).	38
3.3	Iterações necessárias para minimizar a função $y = -x \sin(2x) + \sqrt{x^3}$ com o algoritmo EGO e o indicador EI. Melhor solução em $x = 3$ com $y = 2,316$	41
3.4	Iterações necessárias para otimizar a função de média $y = -x \sin(2x) + \sqrt{x^3}$, com variância $Var(y) = \cos(x)^2 + 0.01$, usando o Procedimento RI, com o algoritmo EGO e o indicador EI. Melhor solução em $x = 2.92$, com $\bar{y} = 2.352$ e $\sigma_y^2 = 0.962$	43
3.5	Iterações necessárias para otimizar a função de média $y = -x \sin(2x) + \sqrt{x^3}$, com variância $Var(y) = \cos(x)^2 + 0.01$, usando o modelo KOIC, com o algoritmo EGO e o indicador EI. Melhor solução em $x = 2.99$, $\bar{y} = 2.317$ e $\sigma_y^2 = 0.987$	44
4.1	Mapa de Belo Horizonte dividido nas 9 regiões administrativas. Os número em preto denotam bases de ambulâncias (1-20) e os em vermelho denotam UPAs (1-8), hospitais (9-14) e CERSAMs (15-21). Fonte: Silva e Pinto (2010).	49
4.2	Distribuição média de chamadas de emergência realiadas em Belo Horizonte entre outubro de 2006 e agosto de 2009. Baseado em: Silva e Pinto (2010)	49
4.3	Distribuição média de chamadas por faixa horária. Fonte: Silva e Pinto (2010).	49

5.1	Experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, TotAmb = 17, $N = 100$ e 20 replicações.	55
5.2	Experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, TotAmb = 28, $N = 100$ e 20 replicações.	56
5.3	Tempos de execução dos experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, TotAmb = 17, $N = 100$ e 20 replicações.	57
5.4	Tempos de execução dos experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, TotAmb = 28, $N = 100$ e 20 replicações.	58
6.1	Cinco replicações efetuadas com o otimizador OptQuest para minimizar TR Médio para 17 ambulâncias.	61
6.2	Análise de resíduos para o experimento KOIC-17-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	62
6.3	Análise de resíduos para o experimento KOIC-17-100: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	62
6.4	Análise de resíduos para o experimento RI-17-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	62
6.5	Análise de resíduos para o experimento RI-17-100: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	63
6.6	Cinco replicações realizadas para minimização do tempo de resposta resolver KOIC-17-50.	64
6.7	Cinco replicações realizadas para minimização do tempo de resposta para resolver KOIC-17-100.	65
6.8	Cinco replicações realizadas para minimização do tempo de resposta para resolver RI-17-50.	67
6.9	Cinco replicações realizadas para minimização do tempo de resposta para resolver RI-17-100.	68
6.10	Distribuição geográfica das soluções encontradas com KOIC, RI, OptQuest, além da configuração de 2010 em Belo Horizonte.	70

6.11 Tempo de resposta mínimo <i>versus</i> quantidade de ambulâncias alocadas a Belo Horizonte.	72
A.1 Análise de resíduos para o experimento RI-20-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	81
A.2 Análise de resíduos para o experimento RI-22-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	82
A.3 Análise de resíduos para o experimento RI-24-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	82
A.4 Análise de resíduos para o experimento RI-26-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	83
A.5 Análise de resíduos para o experimento RI-28-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	83
A.6 Análise de resíduos para o experimento RI-30-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	84
A.7 Análise de resíduos para o experimento RI-40-50: (a) Teste de normalidade; (b) resíduos <i>versus</i> estimação; (c) valores reais <i>versus</i> estimação.	84

Lista de Símbolos

$Z(\mathbf{x})$	Processo Gaussiano.
$m(\mathbf{x})$	Média do Processo Gaussiano de Krigagem.
\mathbf{R}	Matriz de covariância do Processo Gaussiano de Krigagem.
R_{ij}	Covariância entre amostras i e j .
θ_k	Peso da variável de entrada k .
ω_k	Diferença absoluta entre duas amostras para a mesma variável k .
$\rho_k(\omega_k, \theta_k)$	Função de correlação para a variável de entrada k .
K	Número de variáveis de entrada.
N	Tamanho do conjunto de treinamento.
β	Média dos experimentos na Krigagem.
\mathbf{c}	Vetor de covariâncias entre as amostras e um novo ponto.
σ_Z^2	Variância do Processo Gaussiano de Krigagem.
\mathbf{y}	Vetor da variável de resposta dos experimentos.
$\bar{\mathbf{y}}$	Vetor da média dos experimentos.
$\mathbf{1}$	Vetor de uns.
σ_K^2	Variância de estimação com Krigagem.
$e(\mathbf{x})$	Ruído aleatório cuja distribuição de probabilidade depende de \mathbf{x} .
\mathbf{I}	Matriz identidade.
ϵ	Efeito pepita.
\mathbf{C}	Matriz de covariância considerando ruído aleatório.
Δ	Matriz diagonal com as variâncias dos experimentos.
$\boldsymbol{\theta}$	Vetor com os pesos das K variáveis de entrada.
$l(\beta, \sigma_Z^2, \boldsymbol{\theta})$	Função log-verossimilhança.
$\sigma_{\mathbf{y}}^2$	Variância do vetor \mathbf{y} .
\mathbf{x}_k	Vetor com valores da variável k para todos os experimentos.

$\sigma_{\mathbf{x}_k}^2$	Variância do vetor \mathbf{x}_k .
δ	Precisão definida para parada do processo de otimização.
$\Phi(\cdot)$	Distribuição acumulada da distribuição normal padronizada.
$\phi(\cdot)$	Distribuição densidade de probabilidade da distribuição normal padronizada.
$y_{\text{interp}}(\mathbf{x}_i)$	(Método RI) Valor de resposta da amostra i no modelo interpolador.
$\hat{y}_{\text{reg}}(\mathbf{x}_i)$	(Método RI) Estimação da amostra i feita com o modelo regressivo.
$\boldsymbol{\tau}$	Vetor com os desvios-padrões das replicações dos experimentos.
$\Phi_{\alpha/2}^{-1}$	Número de desvios-padrões a partir da média para um nível de confiança de $100 \times (1 - \alpha)\%$.
r_{cruz}	Resíduo por validação cruzada.

Lista de Siglas

IC	Intervalo de confiança.
KO	Krigagem Ordinária.
KS	Teste Kolmogorov-Smirnov de normalidade.
EGO	<i>Efficient Global Optimization.</i>
EI	<i>Expected Improvement.</i>
RI	Procedimento de Reinterpolação.
KOIC	Estimação do lado superior do intervalo de confiança com Krigagem.
SA	Meta-heurística <i>Simulated Annealing.</i>
TR Médio	Tempo de resposta médio.
TR IC+	Valor simulado do lado superior do intervalo de confiança da média do tempo de resposta.
TR IC+ Mínimo	Valor mínimo de TR IC+ nas iterações do processo de otimização.
KOIC IC	Intervalo de confiança da estimação com Krigagem.
TR IC	Intervalo de confiança da média do tempo de resposta médio.
TR Mínimo	Valor de mínimo encontrado a cada iteração do processo de otimização com o procedimento RI.
KO IC	Intervalo de confiança da estimação com Krigagem.
TR IC	Intervalo de confiança da média do tempo de resposta.

Sumário

1	Introdução	15
1.1	Definição do problema	17
1.2	Metodologia	17
1.2.1	Objetivos geral e específicos	18
1.3	Estrutura da dissertação	19
2	Metamodelagem	20
2.1	Krigagem	24
2.2	Estimação dos parâmetros de Krigagem	31
2.2.1	Considerações sobre a implementação	32
2.3	Planejamento de Experimentos e Análise de Resíduos	33
3	Otimização para simulação	37
3.1	Critério <i>Expected Improvement</i>	40
3.2	Procedimento de Reinterpolação	42
3.3	Krigagem do Intervalo de Confiança da Média	44
4	Modelo de Simulação: o SAMU de Belo Horizonte	46
4.1	Descrição do sistema	46
4.2	O Modelo de Simulação	48
4.3	Otimização do tempo de resposta	50
5	Heurística de otimização: Simulated Annealing	52
5.1	Definição dos parâmetros de SA	54
6	Otimização para Simulação do modelo do SAMU de Belo Horizonte	59
6.1	Otimização da configuração de ambulâncias de 2010	60

6.2	Tempo de resposta mínimo <i>versus</i> número de ambulâncias alocadas	70
7	Conclusões e Trabalhos Futuros	73
	Referências Bibliográficas	76
Apêndice A	Análises de resíduos e melhores soluções: de 20 a 40 ambulâncias	81

Capítulo 1

Introdução

Otimização para Simulação é um tema de modelagem de simulação voltada para a determinação da melhor configuração de recursos escassos, maximizando ou minimizando uma função objetivo definida. Quando tratando de modelos complexos, a obtenção dos valores dessa função objetivo geralmente demanda considerável esforço computacional e, conseqüentemente, tempo de execução.

Nesse sentido, otimizar a configuração de recursos em um modelo de simulação de alta complexidade não é uma tarefa simples, já que, como a função objetivo não é diferenciável e cada amostragem requer elevado custo computacional, o processo de otimização necessita do uso de estratégias adequadas para escapar de mínimos e máximos locais, mas também para evitar a avaliação excessiva de pontos em regiões pouco significativas do espaço de análise.

As estratégias para facilitar, ou mesmo para tornar viável, estudos de Otimização para Simulação geralmente envolvem o uso de metamodelos, cujo objetivo é fazer aproximações do sistema original (simulado) durante o processo de busca do ponto ótimo. Sem o uso de metamodelos, a heurística de busca ganharia bastante complexidade, pois precisaria ser um método especialista, o que exige o conhecimento de mais informações do que apenas uma variável de saída. Além disso, o método se tornaria menos eficiente, pois todos os pontos candidatos a ótimo precisariam ser avaliados com o modelo de simulação.

Dentre diversos tipos de metamodelos existentes na literatura, o mais popular é a Krigagem, que surgiu na geoestatística como um método de aproximação às concentrações e quantidades de minério em jazidas, baseado na distância entre amostras. Este conceito foi adaptado à metamodelagem e, hoje em dia, Krigagem é considerada como o metamodelo estado-da-arte para otimização de funções *black-box* como os modelos de simulação, em que suas características são

definidas no código computacional.

Modelos de Krigagem fornecem, além da estimação do ponto em análise, a informação de incerteza de estimação que, ao contrário dos outros tipos de metamodelos, é influenciada diretamente pela "distância" entre amostras. Dessa forma, quando a incerteza de estimação é grande, o método de otimização reconhece existência de uma região no espaço de análise com poucas amostras e, conseqüentemente, pouca informação, assim há balanceamento entre buscas locais e globais. O termo distância foi colocado entre aspas, pois em metamodelagem geralmente os problemas não são bidimensionais ou tridimensionais, logo a palavra não pode ser considerada no seu sentido estrito.

O método de otimização mais popular para resolver o problema citado anteriormente é o *Efficient Global Optimization* (EGO). O método explora as características do modelo de Krigagem e, com o indicador *Expected Improvement* (EI), promove a otimização de funções *black-box*, a partir da escolha de um ponto candidato a ótimo em cada iteração e por avaliações sequenciais desses pontos na função real (ou modelo de simulação).

Nesse sentido, o objetivo principal deste trabalho foi aplicar o método de otimização sequencial baseado em Krigagem sobre um modelo de simulação estocástico de grande complexidade. Para tal, o modelo de simulação do SAMU de Belo Horizonte, proposto em 2010 por Silva e Pinto (2010), foi escolhido como um subtrato adequado, por apresentar alta complexidade, e cuja função é retornar o tempo médio de resposta de ambulâncias a partir da localização e alocação de unidades pela cidade.

Para tratar a aleatoriedade do modelo de simulação foram usados dois métodos diferentes: o Procedimento de Reinterpolação (FORRESTER; KEANE; BRESSLOFF, 2006) e a estimação do lado superior do intervalo de confiança, proposto neste trabalho e chamado do KOIC, cuja motivação foi derivada da necessidade de avaliar todo o intervalo de confiança da média. As duas técnicas foram comparadas entre si e a melhor foi escolhida para obter uma curva que faz o relacionamento entre o tempo de resposta mínimo e o número de ambulâncias alocadas à cidade.

Para facilitar a linguagem desta dissertação, o termo "otimização" e suas variações foram usados no sentido heurístico da palavra, ou seja, como tratamos de funções não necessariamente convexas e diferenciáveis, muitas vezes não é possível comprovar a otimalidade das soluções encontradas, entretanto estas soluções são consideradas boas o suficientes pelo critério de parada do processo de otimização.

Nas seções subsequentes serão melhor detalhados o problema estudado nesta dissertação e a

metodologia empregada para sua resolução.

1.1 Definição do problema

O problema estudado aqui é a especificação e implementação de técnicas voltadas para Otimização para Simulação de sistemas estocásticos de alta complexidade. Existem diversas publicações na literatura, como mostrado no Capítulo 2 e Capítulo 3, que estudam a influência de indicadores específicos no EGO e a influência da aleatoriedade inerente à função *black-box*, entretanto é raro encontrar trabalhos que lidem com problemas estocásticos reais com muitas variáveis de decisão.

Nesse sentido, esta pesquisa procura complementar o estado-da-arte de otimização baseada em Krigagem, a partir da análise que envolve as respostas das seguintes perguntas:

- Qual o nível de complexidade agregado ao processo de otimização ao se considerar o caráter estocástico da variável de decisão?
- Em lugar de se avaliar apenas a média da variável de decisão, é viável trabalhar na minimização ou maximização de um dos lados do seu intervalo de confiança?
- Qual é o nível de esforço computacional ao se trabalhar com muitas variáveis de entrada?
- Quais são as dificuldades de se trabalhar com um problema real que possui variáveis discretas?

1.2 Metodologia

A metodologia proposta para resolução do problema apresentado anteriormente foi composta do levantamento de técnicas voltadas a Otimização para Simulação com caráter estocástico, escolha ou determinação de um modelo de simulação adequado para a aplicação, implementação computacional das técnicas, uso do modelo de simulação para verificar a influência de indicadores e tamanho de conjuntos de treinamento diferentes no EGO, determinação do melhor indicador e, por fim, realização do processo de otimização propriamente dito.

O levantamento das técnicas de otimização compreenderam o estudo de metamodelagem, a partir de modelos polinomiais, redes neurais e Krigagem, sendo que Krigagem foi escolhida como o metamodelo mais adequado. Além disso, foram levantados os principais indicadores e algoritmos de otimização sequencial de funções *black-box*, como os critérios EI, *Expected Quantile Improvement*, *Augmented Expected Improvement* e o procedimento RI.

Em relação ao modelo de simulação, por questão praticidade e como citado anteriormente, o mesmo modelo proposto por Silva e Pinto (2010) foi utilizado nessa pesquisa, sendo que o objetivo foi minimizar o tempo de resposta das ambulâncias básicas ao realocar unidades pela cidade de Belo Horizonte.

A implementação das técnicas foi feita em C/C++ e envolveu também a especificação de uma heurística de otimização global *Simulated Annealing*, descrita no Capítulo 5, cujo objetivo foi, usando EI como função objetivo, determinar um novo ponto candidato a ótimo.

Por fim, após avaliar a eficiência de dois indicadores e tamanhos de conjunto de treinamento diferentes em EGO, a melhor técnica foi utilizada para minimização do tempo de resposta das ambulâncias e para determinação de uma curva que fez o relacionamento o tempo de resposta e o número de ambulâncias alocadas a Belo Horizonte.

De forma mais estratificada, os objetivos desta pesquisa são apresentados a seguir.

1.2.1 Objetivos geral e específicos

O objetivo geral deste trabalho foi aplicar e avaliar técnicas diferentes para efetuar Otimização para Simulação de um modelo estocástico de alta complexidade. Além disso, buscou-se trabalhar e otimizar um modelo de um sistema real voltado à área da Saúde.

Para alcançar tal meta, são listados os objetivos específicos responsáveis por garantir a qualidade e coerência de cada etapa da pesquisa. São eles:

- Levantar técnicas de Otimização para Simulação baseadas em metamodelos;
- Escolher um modelo de simulação alinhado com os objetivos da pesquisa;
- Implementar as técnicas de otimização estudadas, identificando as principais possibilidades de melhoria ao estado-da-arte;
- Especificar e validar os conjuntos de treinamento;
- Aplicar EGO para otimizar o modelo de otimização e avaliar a qualidade e eficiência das técnicas levantadas; e
- Executar Otimização para Simulação com a melhor técnica apurada e avaliar os resultados do modelo de simulação.

1.3 Estrutura da dissertação

Esta dissertação foi estruturada em sete capítulos e um apêndice. O Capítulo 1 introduz o contexto deste trabalho, apresenta as motivações para o seu desenvolvimento e a metodologia empregada, com a lista dos objetivos geral e específicos; o Capítulo 2 constitui a primeira parte do levantamento bibliográfico, compreendendo o tema de metamodelagem, mas focado à Krigagem; no Capítulo 3 é apresentada a segunda parte do referencial bibliográfico, abordando o algoritmo EGO e os principais critérios de desempenho e parada; o Capítulo 4 detalha o modelo de simulação utilizado nesta pesquisa; o Capítulo 5 especifica as características da heurística de otimização *Simulated Annealing* empregada para obter pontos candidatos a ótimo; o Capítulo 6 descreve com detalhes a aplicação das técnicas de Otimização para Simulação sobre o modelo de simulação do SAMU de Belo Horizonte; no Capítulo 7 são apresentadas as considerações finais do trabalho, avaliação dos objetivos alcançados e proposição de trabalhos futuros; por fim, o Apêndice A mostra os gráficos das análises de resíduos e as tabelas do processo de otimização relatado ao final do Capítulo 6.

Capítulo 2

Metamodelagem

Modelos computacionais complexos são comumente utilizados para simular sistemas reais, com o intuito de se tomar decisões a partir de certas variáveis de resposta em função de outras de entrada, entretanto estes modelos podem ganhar tamanho e complexidade de tal forma que exijam demasiado tempo computacional para realizar experimentos desejados. Nestes casos, metamodelos (modelos do modelo) podem ser empregados, como uma aproximação ao modelo original. O tema metamodelagem remonta aos trabalhos de Box (1954) e Box et al. (1978), por exemplo, mas somente em Kleijnen (1987) o termo foi propriamente concebido.

A representação matemática de um modelo de simulação computacional (daqui em diante tratado somente como modelo de simulação) pode ser descrita como uma função entrada-saída, na forma:

$$y = f(\mathbf{x}) \tag{2.1}$$

onde $f(\mathbf{x})$ faz a transformação das variáveis de entrada em valores de resposta. A função $f(\mathbf{x})$ também é chamada de caixa-preta, ou seja, a sua estrutura, definida pelo código presente no modelo de simulação, não é explícita. O vetor \mathbf{x} e y contêm variáveis aleatórias que representam, respectivamente, as informações de entrada e saída do metamodelo. Um metamodelo multivariado pode ser desenvolvido caso hajam mais de uma variável de saída, entretanto um metamodelo específico é geralmente desenvolvido para cada componente de γ (BARTON, 1992).

Metamodelos são usados para resolver problemas de otimização de funções no formato de caixa-preta, ou seja, que não definida analiticamente. No caso deste estudo, um função caixa-preta é um modelo de simulação estocástico.

Autores como Simpson et al. (2001a) e Biles et al. (2007) substituem totalmente o modelo de

simulação original por um metamodelo e fazem a otimização do segundo, enquanto outros, como Sasena, Papalambros e Goovaerts (2002), Jones, Schonlau e Welch (1998) e Viana, Haftka e Watson (2010), utilizam o metamodelo como direcionador para um processo sequencial de obtenção de novas observações e realizar otimização para simulação (este tópico é revisitado no Capítulo 3). Citam-se, a seguir, alguns trabalhos importantes que advocam o uso de metamodelagem para resolver problemas de otimização para simulação.

Jin, Du e Chen (2003) aplica regressão polinomial, Krigagem e função de base radial para medir a eficiência de cada metodologia em um problema de otimização sobre incerteza. Biles et al. (2007) aplicam regressão e Krigagem para otimizar o custo mínimo de estocagem, o custo de falta e o custo total, em função da política de estoque, sobre um modelo de estoque simulado. Barton (2009) apresenta um guia para a implementação de métodos locais e globais de otimização utilizando metamodelos. Dellino, Kleijnen e Meloni (2009) propõem uma metodologia que combina o método de Taguchi com a metodologia de superfície de resposta e Krigagem para construção de um método de otimização robusto para simulação.

Aplicações práticas de metamodelagem estão principalmente presentes na resolução de problemas de engenharia. Simpson et al. (2001a) utiliza superfície de resposta e Krigagem no projeto do bocal de um motor *aerospike* (Engenharia Aeroespacial). Garitselov, Mohanty e Koungianos (2012) utiliza redes neurais no projeto de um circuito nano-CMOS (Engenharia Eletrônica). Kumar e Rao (2010) emprega Krigagem para estimar o fator de atrito de canal aluvial (Engenharia Civil). Mais recentemente, Zhang et al. (2015) utiliza metamodelos de Krigagem para otimizar a eficiência e sucção de bombas centrífugas (Engenharia Mecânica), Badhurshah e Samad (2015) usa múltiplos metamodelos para otimizar o fluxo simulado de turbinas de impulso bidirecional (Engenharia Mecânica) e Hussain et al. (2015) aplica regressão polinomial evolucionária e um algoritmo genético para minimizar custos econômicos e ambientais de sistemas aquíferos complexos sujeitos a entrada de água do mar (Engenharia Ambiental).

Fora do mundo da engenharia alguns trabalhos também merecem destaque. Liang, Buongiorno e Monserud (2006) utiliza superfície de resposta para otimizar controles gerenciais sobre efeitos econômicos e ecológicos no crescimento de uma floresta simulada (Gestão Florestal). Ci-mellaro, Reinhorn e Bruneau (2011) aplica metamodelagem para estimar o tempo de permanência em um hospital e avaliar o índice de resiliência de facilidades de saúde em relação a desastres (Gestão em Saúde). Ahuja, Ferreira e Moreira (2004) usa superfície de resposta para otimizar a concentração de três componentes químicos e alcançar crescimento exponencial na cultura de

uma bactéria (Otimização em Saúde).

Os modelos matemáticos mais simples utilizados em metamodelagem geralmente tomam a forma de regressões polinomiais de primeira ou segunda ordem (BILES et al., 2007; LAW, 2007). Tais modelos polinomiais são utilizados como aproximações à superfície de resposta real (Equação 2.1). Alguns textos como Box e Draper (1987), Khuri e Cornell (1996), Myers e Montgomery (2003), além dos trabalhos de Barton (1992), Barton (1998) fazem uma revisão deste e outros tipos de metamodelo.

Pela metodologia de superfície de resposta, a relação entre uma variável de resposta y e um vetor de fatores independentes \mathbf{x} é dada por:

$$y(\mathbf{x}) = \psi(\mathbf{x}) + \xi \quad (2.2)$$

onde $\psi(\mathbf{x})$ é geralmente uma função desconhecida de \mathbf{x} e ξ é o erro aleatório com média zero e variância σ^2 , sendo os erros individuais ξ_i independentes e identicamente distribuídos. Como as verdadeiras formas de $\psi(\mathbf{x})$ e ξ são desconhecidas, uma superfície de resposta $\hat{y}(\mathbf{x})$ é criada como uma aproximação à $\psi(\mathbf{x}) + \xi$ (SIMPSON et al., 2001b).

Uma superfície de resposta obtida a partir de uma regressão de primeira ordem com termo de interação pode ser descrita como:

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^K \beta_i x_i + \sum_{i=1}^K \sum_{\substack{j=2 \\ j>i}}^K \beta_{ij} x_i x_j, \quad (2.3)$$

já uma superfície de resposta quadrática ou de segunda ordem com termo de interação tem a forma

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^K \beta_i x_i + \sum_{ii=1}^K \beta_{ii} x_i^2 + \sum_{i=1}^K \sum_{\substack{j=2 \\ j>i}}^K \beta_{ij} x_i x_j \quad (2.4)$$

onde K denota o número de variáveis de entrada.

Os coeficientes β_0 , β_{ii} e β_{ij} são estimados minimizando a soma de mínimos quadrados dos desvios entre os dados reais e as estimações, cujo resultado tem a forma (MONTGOMERY, 2013):

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.5)$$

onde \mathbf{X} é uma matriz de tamanho $N \times (K + 1)$ com funções de \mathbf{x} sobre todos os dados coletados e cuja primeira coluna é composta apenas de uns, N é o número de amostras coletadas, \mathbf{y} é um

vetor coluna contendo os valores de resposta correspondentes e $\hat{\beta}$ denota o vetor de estimadores dos coeficientes. A estimação de um ponto qualquer é dada por:

$$\hat{y}(\mathbf{x}) = \mathbf{x}^T \hat{\beta} \quad (2.6)$$

Devido a sua simplicidade, os metamodelos polinomiais são comumente utilizados em processos de busca local (BARTON, 2009), de forma que, de acordo com que a região de busca vai sendo explorada, novos metamodelos são construídos e os antigos descartados.

Metamodelos mais sofisticados são recomendados para estimar superfícies não-lineares com muitas variáveis de entrada. Simpson et al. (2001b) fazem algumas recomendações sobre o tipo de metamodelo mais adequado a ser utilizado de acordo com a aplicação. Segundo os autores, modelos polinomiais são mais eficazes para aplicações com erro aleatório e com menos de 10 fatores. Redes neurais são indicadas para problemas, principalmente determinísticos, altamente não lineares e de grande porte (aproximadamente 10.000 fatores).

Neste estudo, apenas modelos de Krigagem foram implementados. Estes modelos são considerados o estado-da-arte em otimização para simulação (PICHENY; GINSBOURGER, 2014). Eles são indicados para estimar superfícies não lineares e são flexíveis para interporlação de dados amostrais ou na filtragem de ruído, mesmo que sejam difíceis de se obter e implementar (WANG; SHAN, 2007). Além disso, modelos de Krigagem fornecem uma informação de variância de estimação, usada para balancear buscas globais e locais no processo de otimização.

Segundo Kleijnen e Sargent (2000), Wang e Shan (2007), Barton (2009) existem seis passos principais para se conduzir um estudo de otimização baseada em metamodelagem. Os passos são:

- Identificar a estrutura do metamodelo;
- Realizar um Planejamento de Experimentos inicial;
- Conduzir os experimentos de simulação;
- Fazer o ajuste das observações ao metamodelo escolhido e validar a qualidade de ajuste;
- Minimizar (ou maximizar) a variável de resposta do metamodelo ou utilizá-lo como direcionador à um processo de otimização sequencial; e
- Verificar a solução ótima obtida a partir o metamodelo com o modelo de simulação.

2.1 Krigagem

Neste trabalho é dado destaque apenas aos metamodelos de Krigagem. Este tipo de metamodelo é usado devido à sua alta flexibilidade, sua habilidade em tratar não linearidade com certa facilidade e pelo fato de fornecer um erro de estimação (variância de estimação) que aumenta em regiões pouco exploradas do espaço de análise. Esta característica específica da Krigagem nos permite determinar regiões do espaço de análise que precisam ser melhor exploradas, ou seja, em um processo de otimização, o metamodelo nos dá informação sobre quando é necessário fazer buscas globais em detrimento de buscas locais.

Os modelos de Krigagem surgiram na geoestatística como métodos de estimação espacial. Nesses modelos, um novo ponto é estimado em função da sua distância em relação à todas as amostras do conjunto de treinamento obtido a partir de um Planejamento de Experimentos.

A Krigagem foi inicialmente desenvolvida pelo engenheiro de minas sul-africano Danie Krige e, posteriormente, por Matheron (1963). Estes modelos ganharam popularidade, fora da área de mineração, principalmente, como aproximação a modelos determinísticos de engenharia, que possuem K dimensões (fatores) de análise, enquanto que na geoestatística geralmente somente duas ou três dimensões são utilizadas (KLEIJNEN, 2009).

Atualmente, os principais disseminadores das técnicas de Krigagem aplicadas à modelos de simulação são Kleijnen e seus colaboradores (BEERS; KLEIJNEN, 2004; BILES et al., 2007; KLEIJNEN, 2009; KLEIJNEN; MEHDDAD, 2014). Estes autores vêm aplicando o método em contextos diferentes, mas sempre relacionados à substituição de modelos de simulação estocásticos.

O modelo determinístico de Krigagem é dividido em dois elementos: uma função de tendência da média dos experimentos e um Processo Gaussiano com média zero, cujos desvios em relação à média são explicados por uma matriz de covariância. As duas variações clássicas de Krigagem são a Krigagem Ordinária (KO), quando a média dos experimentos é constante, e a Krigagem Universal, quando a média é definida por um função de tendência geralmente polinomial.

Algumas novas abordagens para a metodologia de Krigagem foram propostas na tentativa de melhorar a estimação. Citam-se Joseph, Hung e Sudjianto (2008) e Couckuyt et al. (2012) que propõem o método Krigagem Cega (*Blind Kriging* em inglês), utilizando *Bayesian forward selection* na estimação, e Deng et al. (2012) que propõe a incorporação de um esquema de inferência Bayesiana no processo de Krigagem.

No âmbito estocástico, uma componente de ruído é adicionada ao modelo determinístico, entretanto apenas uma minoria dos autores trata Krigagem com uma componente de ruído heterogêneo (PICHENY et al., 2013).

Um ruído é definido como uma perturbação aleatória nos dados coletados, podendo ter origem tanto no método de coleta das amostras quanto na aleatoriedade natural do fenômeno estudado. Quando o ruído é dito homogêneo nos referimos às perturbações que não dependem do ponto de análise, ou seja, ele possui variância constante em todas as amostras coletadas. Já o ruído heterogêneo possui variância dependente do ponto de análise.

A ocorrência de ruído heterogêneo é um fato recorrente em modelos de simulação de filas (modelos de simulação de eventos discretos). Nestes modelos, para cada conjunto de fatores, são realizadas replicações em que a semente de geração números aleatórios é modificada, com isso cada replicação retorna um valor diferente para a variável de resposta. Por fim, após a realização da quantidade necessária de replicações, a média e variância são calculada. Neste trabalho, otimização baseada em Krigagem é aplicada a um modelo de simulação de eventos discretos que simula o atendimento de chamadas de emergência por ambulâncias.

Alguns trabalhos que tratam de Krigagem com ruído homogêneo são: Kleijnen e Beers (2005) aplicam Krigagem Ordinária sobre um esquema de Studentização dos valores da variável de resposta para contabilizar a variância do ruído aleatório; Huang et al. (2006) utilizam a variância do ruído homogêneo ao modificar o critério *Expected Improvement* no processo de otimização sequencial *Efficient Global Optimization*; e Martin (2010) empregam o conceito de efeito pepita para modificar o modelo tradicional de Krigagem. Na geoestatística, efeito pepita é o fenômeno que ocorre quando a variabilidade entre pares de amostras não parece ser nula ao reduzir a distância entre elas a zero (CHILES; DELFINER, 2012).

Já no caso heterogêneo, destacam-se os estudos de Staum (2009), Ankenman, Nelson e Staum (2010) e Chen, Ankenman e Nelson (2012). Os autores propõem o método Krigagem Estocástica (KE), fornecem instruções sobre implementação e mostram o efeito da técnica de redução de variância Commom Random Numbers sobre o método. Em paralelo, Yin, Ng e Ng (2011) propuseram um modelo de Krigagem com efeito pepita modificado para contabilizar ruídos aleatórios heterogêneos e mostram como essa variação permite a estabilização do processo de estimação dos parâmetros do modelo. Além disso, Yin, Ng e Ng (2011) mostram como a não contabilização de ruído heterogêneo sobre um exemplo teste proporcionou um ajuste ruim de um modelo determinístico de KO.

Os principais programas de computador comerciais utilizados para realizar Krigagem pertencem à área da geoestatística e portanto trabalham apenas com 2 ou 3 dimensões, sendo assim bastante limitados para as aplicações desejadas nesta dissertação. Alguns programas neste contexto são o Isatis (Geovariances, <www.geovariances.com>), GS+ Geostatistics Software (Rockware, <www.rockware.com>) e arcGIS (Esri, <www.esri.com>).

Programas voltados para Krigagem como abordado em metamodelagem são geralmente desenvolvidos de forma livre (*opensource*) em C/C++ ou como pacotes de *softwares* de cálculo matemático e estatístico como o Matlab (MathWorks, <www.mathworks.com>) ou o R (comunidade GNU, <www.r-project.org>). Pacotes desenvolvidos neste sentido são o DACE para Matlab (LOPHAVEN; NIELSEN; SØNDERGAARD, 2002) e o DiceKriging para R (ROUSTANT; GINSBOURGER; DEVILLE, 2012).

Dado um conjunto de treinamento com N amostras de K fatores, o modelo determinístico de Krigagem é definido como:

$$y(\mathbf{x}) = m(\mathbf{x}) + Z(\mathbf{x}) \quad (2.7)$$

onde $\mathbf{x} = [x_1, x_2, \dots, x_K]$ é um vetor com variáveis de entrada para uma observação, $y(\mathbf{x})$ é a variável de resposta, $m(\mathbf{x})$ é a média dos experimentos e $Z(\mathbf{x})$ denota realizações de um Processo Gaussiano com média zero e variância desconhecida σ_Z^2 , onde desvios em relação à média são descritos por sua matriz de covariância entre amostras. As covariâncias de $Z(\mathbf{x})$ entre amostras $i \in \{1, 2, 3, \dots, N\}$ e $j \in \{1, 2, 3, \dots, N\}$ são dadas por:

$$R_{i,j} = \sigma_Z^2 \prod_{k=1}^K \rho_k(|x_{ik} - x_{jk}|, \theta_k) \quad (2.8)$$

onde $\rho_k(\cdot)$ é uma função de correlação individual definida para cada variável de entrada $k = 1, 2, \dots, K$, $|x_{ik} - x_{jk}|$ denota a diferença absoluta entre as amostras i e j sobre a mesma variável de entrada k e θ_k é um peso escolhido para cada variável k .

A função mais amplamente utilizada em Krigagem para metamodelagem é função Gaussiana (GANO et al., 2006). Outras funções de correlação populares são a função Linear, Exponencial (KLEIJNEN, 2009), além das funções *Matérn* e Potência (ROUSTANT; GINSBOURGER; DEVILLE, 2012). A formulação das funções de correlação citadas acima são apresentadas a seguir e suas formas plotadas na Figura 2.1. Para facilitar a notação, a distância entre duas amostras em relação à k -ésima variável de entrada foi definida como ω_k , ou seja, $\omega_k = |x_{ik} - x_{jk}|$.

- Função Gaussiana

$$\rho_k(\omega_k, \theta_k) = \exp\left(-\frac{\omega_k^2}{2\theta_k^2}\right) \quad (2.9)$$

- Função Exponencial

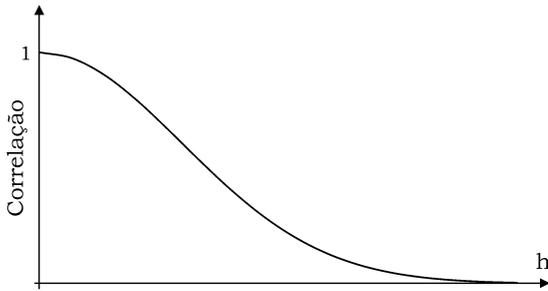
$$\rho_k(\omega_k, \theta_k) = \exp\left(-\frac{\omega_k}{\theta_k}\right) \quad (2.10)$$

- Função Linear

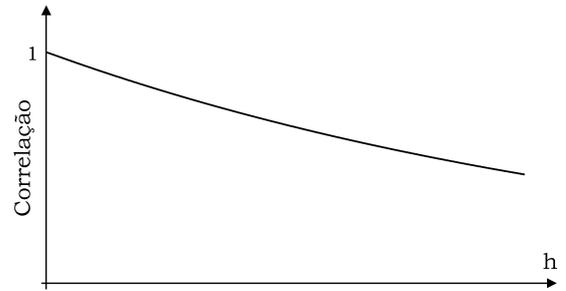
$$\rho_k(\omega_k, \theta_k) = \max\left(1 - \frac{\omega_k}{\theta_k}, 0\right) \quad (2.11)$$

- Função *Matérn* $\nu = 5/2$

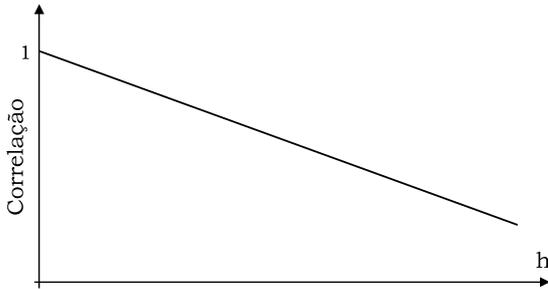
$$\rho_k(\omega_k, \theta_k) = \left[1 + \frac{\sqrt{5}\omega_k}{\theta_k} + \frac{5}{3}\left(\frac{\omega_k}{\theta_k}\right)^2\right] \exp\left[-\frac{\sqrt{5}\omega_k}{\theta_k}\right] \quad (2.12)$$



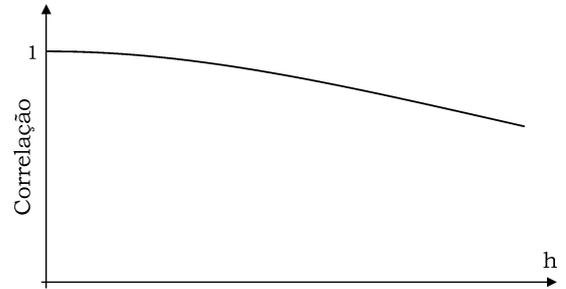
(a) Função Gaussiana



(b) Função Exponencial



(c) Função Linear



(d) Função *Matérn* $\nu = 5/2$

Figura 2.1: Funções de correlação utilizadas em Krigagem para metamodelagem

Por conveniência, a formulação da função Gaussiana é dada segundo Roustant, Ginsbourger e Deville (2012). Em todas as funções, θ_k denota um peso dado à k -ésima variável de entrada, podendo tomar os valores $0 \leq \theta_k < +\infty$. Caso $\theta_k = 0$, mudanças na variável de entrada não afetam a covariância entre amostras. Se $\theta_k \rightarrow \infty$, então a covariância reduz a zero, ou seja, as amostras são independentes.

Na literatura de metamodelagem há pouca informação sobre a escolha de uma ou outra função de correlação. Na geoestatística, entretanto, a adequação de uma função em relação à variabilidade dos dados é definida principalmente pelo seu comportamento na origem (CHILES; DELFINER, 2012). A função Gaussiana é utilizada, por exemplo, para processos onde a variabilidade aumenta de forma lenta com a distância, já a função exponencial ou linear são utilizadas geralmente quando há certa descontinuidade nessa variabilidade (mudanças bruscas).

Existem algumas diferenças entre a metodologia de Krigagem utilizada na geoestatística e em metamodelagem. Além de tratarem problemas com dimensões diferentes, uma das diferenças mais significativas consiste na forma de obtenção da função de correlação a ser utilizada na Equação 2.8, sendo este fato evidente ao comparar o desenvolvimento dos trabalhos voltados à geoestatística de Hengl, Heuvelink e Stein (2004) ou Delhomme (1978) e dos voltados à metamodelagem de Martin e Simpson (2004) ou Zhao, Choi e Lee (2011). Esta mudança de metodologia já foi notada por Beers e Kleijnen (2004) e Sasena (2002).

Na geoestatística a variabilidade espacial é modelada geralmente a partir do semivariograma amostral, ao qual funções teóricas são ajustadas. Um semivariograma é modelado para cada direção em que as amostras estão dispostas, como por exemplo as direções Oeste-Leste ou Norte-Sul. Além disso, a modelagem da correlação espacial é dependente de uma análise visual da variabilidade.

Em metamodelagem são usadas as funções de correlação citadas anteriormente, cujos parâmetros são geralmente obtidos através da maximização da função log-verossimilhança (ver Seção 2.2), sendo desnecessário o uso de análise visual. Esta abordagem requer a tomada de certas suposições sobre o comportamento da variabilidade dos dados, ou seja, a escolha de uma função de correlação depende de como a variabilidade da variável de resposta entre duas amostras se comporta com o aumento da diferença entre as suas variáveis de entrada.

Dadas as considerações anteriores, o estimador de KO é dado por:

$$\hat{y}(\mathbf{x}) = \hat{\beta} + \hat{\mathbf{c}}^T \hat{\mathbf{R}}^{-1} (\mathbf{y} - \mathbf{1}\hat{\beta}) \quad (2.13)$$

onde $\hat{\beta}$ é o estimador da média dos experimentos, $\hat{\mathbf{c}}^T = \hat{\sigma}_Z^2 [\hat{R}_{1,0}, \hat{R}_{2,0}, \dots, \hat{R}_{N,0}]$ é um vetor com os estimadores das covariâncias entre todos os N pontos do conjunto de treinamento e o ponto \mathbf{x}

que se quer estimar, $\hat{\mathbf{R}} = \hat{\sigma}_Z^2 \begin{bmatrix} \hat{R}_{1,1} & \hat{R}_{1,2} & \cdots & \hat{R}_{1,N} \\ \hat{R}_{2,1} & \hat{R}_{2,2} & \cdots & \hat{R}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{N,1} & \hat{R}_{N,2} & \cdots & \hat{R}_{N,N} \end{bmatrix}$ é o estimador da matriz de covariâncias de tamanho $N \times N$, $\hat{\sigma}_Z^2$ é o estimador de σ_Z^2 , $\hat{R}_{i,j}$ é o estimador de $R_{i,j}$, \mathbf{y} é o vetor com as variáveis de resposta dos N pontos do conjunto de treinamento e $\mathbf{1}$ é um vetor de tamanho N com uns. Nota-se que $(\mathbf{y} - \mathbf{1}\hat{\beta})$ representa os desvios em relação à média dos experimentos.

Como citado anteriormente, a variância de estimação por KO, dada por:

$$\hat{\sigma}_K^2(\mathbf{x}) = \hat{\sigma}_Z^2 - \hat{\mathbf{c}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{c}} + \left(1 - \hat{\mathbf{c}}^T \hat{\mathbf{R}}^{-1} \mathbf{1}\right)^T \left(\mathbf{1}^T \hat{\mathbf{R}}^{-1} \mathbf{1}\right)^{-1} \left(1 - \hat{\mathbf{c}}^T \hat{\mathbf{R}}^{-1} \mathbf{1}\right) \quad (2.14)$$

é a grande vantagem da Krigagem em relação a outros tipos de metamodelos. Esta medida de incerteza é muito útil, principalmente em otimização global, pois este indicador apresenta valores elevados em regiões do espaço de solução com pouca ou nenhuma informação, enquanto que valores baixos indicam regiões bem exploradas (PICHENY; GINSBOURGER, 2014).

Após encontrar o valores da variância de estimação, o intervalo de confiança da média é obtido fazendo:

$$\hat{y}(\mathbf{x}) - \Phi_{\alpha/2}^{-1} \hat{\sigma}_K \leq \hat{y}(\mathbf{x}) \leq \hat{y}(\mathbf{x}) + \Phi_{\alpha/2}^{-1} \hat{\sigma}_K \quad (2.15)$$

onde $\hat{\sigma}_K = \sqrt{\hat{\sigma}_K^2}$ denota o desvio-padrão de estimação por Krigagem e $\Phi_{\alpha/2}^{-1}$ denota o número de desvios-padrões a partir da média para um nível de confiança de $100 \times (1 - \alpha)\%$.

A Figura 2.2 mostra como o distanciamento entre as amostras de treinamento faz com que a variância de estimação de Krigagem aumente consideravelmente. Na figura, a função real representada é $y(x) = -x \cos(2x) + \sqrt{x^3}$, com $2 \leq x \leq 6.5$, e o intervalo de confiança da média (IC) foi construído para 95% de confiança. Nota-se que a distância entre o quarto e quinto ponto (da esquerda para direita) resultou em uma incerteza grande na estimação, podendo ser reduzida com a amostragem de um sexto ponto entre os dois.

A abordagem anterior resulta em uma superfície interpoladora, ou seja, o estimador de Krigagem passa exatamente sobre os pontos já amostrados. Quando os dados apresentam ruído, entretanto, fazer o ajuste a um modelo de Krigagem pode ser problemático, pois $\hat{\mathbf{R}}$ pode não ser inversível dependendo dos valores de $\hat{\boldsymbol{\theta}}$. Nestes casos, um efeito pepita pequeno pode ser adicionado à $\hat{\mathbf{R}}$ para que o estimador suavize ligeiramente os dados (SASENA, 2002).

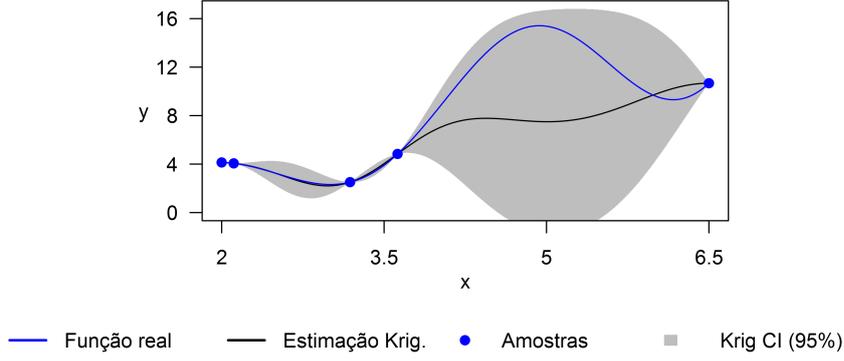


Figura 2.2: Comportamento da variância de estimação de Krigagem.

Assim, o modelo de Krigagem agora toma a forma:

$$y(\mathbf{x}) = m(\mathbf{x}) + Z(\mathbf{x}) + e(\mathbf{x}) \quad (2.16)$$

onde $e(\mathbf{x})$ denota um ruído aleatório cuja distribuição de probabilidade depende de \mathbf{x} .

Para adição do efeito pepita e , conseqüentemente, de um ruído homogêneo à matriz de covariâncias, $\hat{\mathbf{R}}$ é substituída por $\hat{\mathbf{C}} = \hat{\mathbf{R}} + \epsilon \mathbf{I}$ na Equação 2.13 e na Equação 2.14, onde ϵ é o efeito pepita e \mathbf{I} é a matriz identidade de tamanho $N \times N$.

O mesmo modelo de Krigagem definido na Equação 2.16 é usado no caso da presença de ruído heterogêneo. Para isso, $\epsilon \mathbf{I}$ é substituído pela matriz diagonal $\mathbf{\Delta}$ de tamanho $N \times N$, onde cada elemento $\tau_{ii}^2, \forall i \in \{1, 2, \dots, N\}$, de $\mathbf{\Delta}$ denota a variância das replicações em cada amostra i .

As equações do estimador e variância de estimação por KO na abordagem estocástica são então obtidos com:

$$\hat{y}(\mathbf{x}) = \hat{\beta} + \hat{\mathbf{c}}^T \hat{\mathbf{C}}^{-1} (\bar{\mathbf{y}} - \mathbf{1} \hat{\beta}) \quad (2.17)$$

onde $\bar{\mathbf{y}}$ é um vetor com as médias das replicações nas N amostras coletadas e:

$$\hat{\sigma}_K^2 = \hat{\sigma}_Z^2 - \hat{\mathbf{c}}^T \hat{\mathbf{C}}^{-1} \hat{\mathbf{c}} + \left(1 - \hat{\mathbf{c}}^T \hat{\mathbf{C}}^{-1} \mathbf{1}\right)^T \left(\mathbf{1}^T \hat{\mathbf{C}}^{-1} \mathbf{1}\right)^{-1} \left(1 - \hat{\mathbf{c}}^T \hat{\mathbf{C}}^{-1} \mathbf{1}\right) \quad (2.18)$$

onde $\hat{\mathbf{C}} = \hat{\mathbf{R}} + \epsilon \mathbf{I}$ no caso de ruído aleatório homogêneo e $\hat{\mathbf{C}} = \hat{\mathbf{R}} + \mathbf{\Delta}$ no caso de ruído aleatório heterogêneo.

A Figura 2.3 compara o ajuste feito por KO e KO estocástica, onde o intervalo de confiança da média (IC) foi construído para 95% de confiança. A média da função real é dada por $y = -x \sin(2x) + \sqrt{x^3}$, com $2 \leq x \leq 6.5$, e no caso estocástico foi usado um ruído com desvio padrão

de 10% do valor da média, ou seja, $y|x \sim \mathcal{N}(\mu = -x \sin(2x) + \sqrt{x^3}, \sigma = 0.1\mu)$. Nota-se que no caso estocástico o estimador de KO não passa necessariamente pelas amostras e que o intervalo de confiança deixa de ser nulo neste pontos.

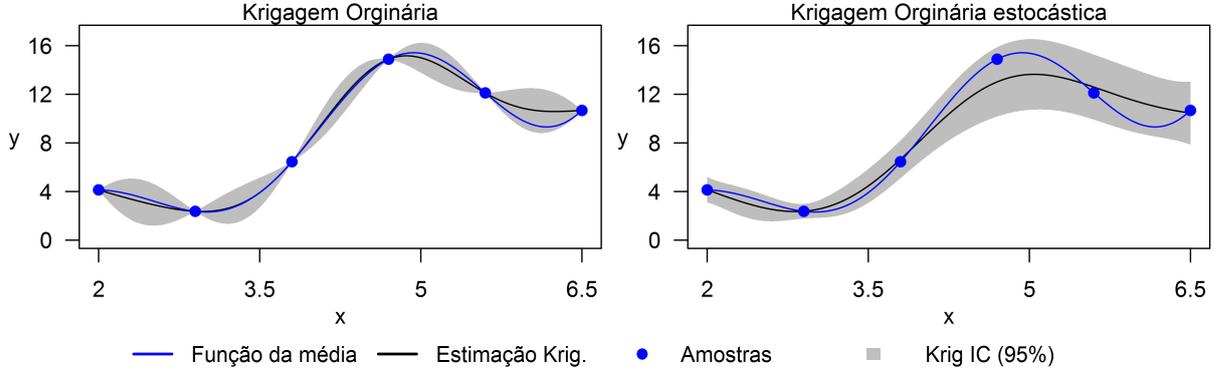


Figura 2.3: Comparação de ajustes entre Krigagem Ordinária e Krigagem Ordinária estocástica.

2.2 Estimação dos parâmetros de Krigagem

Os parâmetros σ_Z^2 e $\theta_k, \forall k \in \{1, 2, \dots, K\}$, são desconhecidos e geralmente escolhidos com a maximização da função log-verossimilhança $l(\beta, \sigma_Z^2, \boldsymbol{\theta})$ da distribuição Gaussiana multivariada, onde $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]^T$ é um vetor com pesos dados a cada variável de entrada K .

A função log-verossimilhança formulada para KO é dada por:

$$l(\beta, \sigma_Z^2, \boldsymbol{\theta}) = -\frac{1}{2} \left[N \ln(2\pi) + \ln |\mathbf{R}| + (\mathbf{y} - \mathbf{1}\beta)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{1}\beta) \right] \quad (2.19)$$

onde $|\mathbf{R}|$ é o determinante da matriz de covariâncias \mathbf{R} .

A maximização da função de log-verossimilhança é de difícil resolução, pois não é convexa (KLEIJNEN; MEHDAD, 2014), a matriz $\hat{\mathbf{R}}$ pode se tornar singular dependendo do vetor $\hat{\boldsymbol{\theta}}$ (GANO et al., 2006) e possui uma forma achatada próxima ao ponto ótimo quando o número de amostras é pequeno (LI; SUDJIANTO, 2005).

Para resolver o problema, Gano et al. (2006) citam: o método de otimização não linear Quase-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS), o método *Pattern Search* e o método *Adaptive Simulated Annealing*. Em seus experimentos, Gano et al. (2006) concluem que não há ganhos significativos sobre a escolha de um método em relação aos outros, entretanto, por ser uma procedimento aleatório, o algoritmo de *Adaptive Simulated Annealing* necessita de muito mais chamadas de função, deixando-o mais lento em relação aos demais.

Simpson et al. (2001a) e Goffe, Ferrier e Rogers (1994) utilizam um algoritmo *Simulated Annealing* para otimização da função log-verossimilhança e Zhao, Choi e Lee (2011) empregam um algoritmo *Generalized Pattern Search*. Li e Sudjianto (2005) propõem um método de máxima verossimilhança penalizado sobre modelos de Krigagem para lidar com a forma achatada da função de verossimilhança. Ankenman, Nelson e Staum (2010) maximiza $l(\beta, \sigma_Z^2, \boldsymbol{\theta})$ a partir de rotinas de otimização não lineares com a restrição $\sigma_Z^2 \geq 0$.

Roustant, Ginsbourger e Deville (2012) definem primeiramente limites inferiores e superiores para σ_Z^2 e $\boldsymbol{\theta}$. Algumas tentativas iniciais são realizadas e o melhor conjunto de parâmetros é usado como ponto de partida para o método de otimização BFGS. A cada iteração os autores calculam $\hat{\beta}$ com:

$$\hat{\beta} = \left(\mathbf{1}^T \hat{\mathbf{R}}^{-1} \mathbf{1} \right)^{-1} \mathbf{1}^T \hat{\mathbf{R}}^{-1} \mathbf{y} \quad (2.20)$$

Como tratado anteriormente, mesmo quando os experimentos são determinísticos, muitas vezes é preferível trabalhar com um efeito pepita pequeno para obter os parâmetros dos modelos de Krigagem. Em função disso, as Equações 2.19 e 2.20 são modificadas para:

$$l(\beta, \sigma_Z^2, \boldsymbol{\theta}) = -\frac{1}{2} \left[N \ln(2\pi) + \ln |\mathbf{C}| + (\bar{\mathbf{y}} - \mathbf{1}\beta)^T \mathbf{C}^{-1} (\bar{\mathbf{y}} - \mathbf{1}\beta) \right], \quad (2.21)$$

onde $|\mathbf{C}|$ é o determinante da matriz \mathbf{C} , e:

$$\hat{\beta} = \left(\mathbf{1}^T \hat{\mathbf{C}}^{-1} \mathbf{1} \right)^{-1} \mathbf{1}^T \hat{\mathbf{C}}^{-1} \bar{\mathbf{y}} \quad (2.22)$$

2.2.1 Considerações sobre a implementação

Os passos seguidos neste trabalho, para fazer o ajuste dos conjuntos de treinamento aos modelos de KO, são semelhantes aos de Roustant, Ginsbourger e Deville (2012), porém com algumas modificações. Um programa em C/C++ foi desenvolvido, sendo seus passos principais descritos a seguir.

Primeiro, defina limites inferiores e superiores para σ_Z^2 e $\boldsymbol{\theta}$. Para este trabalho, foi utilizado, respectivamente, um limite inferior e superior de 10^{-10} e $10 \times \sigma_{\mathbf{y}}^2$ para σ_Z^2 , onde $\sigma_{\mathbf{y}}^2$ é a variância do vetor \mathbf{y} . O limite inferior de $\boldsymbol{\theta}$ foi definido como 10^{-10} e o limite superior para cada θ_k foi escolhido como $10 \times \sigma_{\mathbf{x}_k}^2$, onde \mathbf{x}_k é um vetor coluna com os valores da variável k para todas as N amostras do conjunto de treinamento e $\sigma_{\mathbf{x}_k}^2$ é a variância de \mathbf{x}_k .

A seguir, realize t tentativas aleatórias com os parâmetros $\hat{\sigma}_Z^2$ e $\hat{\boldsymbol{\theta}}$ na tentativa de melhorar

o valor da função log-verossimilhança. Nesta pesquisa, usou-se $t = 100$ para ajustar um modelo com $K = 20$. O melhor conjunto de parâmetros é usado como solução inicial de uma rotina de otimização baseada em gradiente.

Neste sentido, um método de máxima descida com função de penalidade foi implementado, onde os tamanhos dos passos são cuidadosamente obtidos com um procedimento de buscas unidirecionais por *backtracking* (*Backtracking Line Search*), de modo a evitar problemas de inversão de matriz. Além disso, a busca unidirecional deve cumprir a regra de Armijo; ver Luenberger (2003, p. 212-214) ou Wright e Nocedal (1999, p. 41-42). Os gradientes analíticos para a função log-verossimilhança são fornecidos no Apêndice de Roustant, Ginsbourger e Deville (2012).

Com o conjunto inicial de parâmetros, o algoritmo minimiza $-l(\hat{\beta}, \hat{\sigma}_Z^2, \hat{\theta})$, onde os seguintes passos são seguidos a cada iteração:

1. Construir a matriz \hat{R} com os parâmetros $\hat{\sigma}_Z^2$ e $\hat{\theta}$ e computar \hat{C} . Usar \hat{C} no lugar de \hat{R} em todas as equações;
2. Obter $\hat{\beta}$ com a Equação 2.22;
3. Definir a direção de busca com o gradiente da função log-verossimilhança;
4. Realizar uma busca unidirecional para definir um passo suficientemente pequeno (evitar problemas de inversão de matriz e respeitar os limites inferiores e superiores dos parâmetros);
5. Obter um novo conjunto de $\hat{\sigma}_Z^2$ e $\hat{\theta}$ que deve proporcionar um decréscimo em $-l(\hat{\beta}, \hat{\sigma}_Z^2, \hat{\theta})$;
6. Parar se $|l_{n-1}(\cdot) - l_n(\cdot)| \leq \delta$ ou se $n = n_{\max}$, onde $l_{n-1}(\cdot)$ e $l_n(\cdot)$ denotam valores da função log-verossimilhança nas iterações sucessivas $n - 1$ e n do algoritmo e $\delta \geq 0$ é um escalar suficientemente pequeno;
7. Retornar ao Passo 1 se não parar.

Nesta implementação, o programa para quando a diferença absoluta no valor da função log-verossimilhança entre iterações consecutivas é menor que 10^{-16} ou quando alcança um máximo de 200 iterações. O código implementado foi validado com o pacote *DiceKriging* para R e resultados semelhantes foram obtidos.

2.3 Planejamento de Experimentos e Análise de Resíduos

Um metamodelo é somente tão eficiente e eficaz quanto a qualidade dos experimentos utilizados para sua construção. A qualidade e precisão de um metamodelo não estão somente ligadas à

quantidade de amostras coletadas, mas também ao conjunto de dados escolhido e à escolha do procedimento de análise (CRARY, 2002).

Um Planejamento de Experimentos representa um conjunto de experimentos, expressado na forma de fatores (parâmetros) e seus níveis (valores) e organizados como uma tabela onde cada linha representa uma amostra coletada e cada coluna o valor de um parâmetro específico (SIMPSON et al., 2001b; KLEIJNEN, 2005).

Os principais objetivos de um Planejamento de Experimentos são determinar quais fatores têm o maior efeito sobre as variáveis de resposta e realizar observações usando o mínimo possível de amostras ou corridas de simulação (LAW, 2007).

No caso desta pesquisa, assumimos já conhecer quais são os fatores mais importantes para explicação da variável de resposta. Assim, nos resta apenas realizar um Planejamento de Experimentos para amostrar o mínimo necessário de pontos. Uma das técnicas mais populares para este fim é a amostragem por Hipercubo Latino (*Latin Hypercube Design* em inglês, LHD) proposta por McKay, Beckman e Conover (1979). Mesmo que não sendo aplicável a toda situação, a amostragem por Hipercubo Latino é sugerida por Kleijnen (2009), Xiong et al. (2009) e Chen, Ankenman e Nelson (2012).

Após a determinação dos experimentos e a sua execução, é necessário verificar a qualidade destas amostras em relação às suposições tomadas na utilização de um metamodelo específico. No caso de Krigagem precisamos verificar principalmente se os resíduos possuem média zero, se são normalmente distribuídos, se não existe correlação entre os resíduos e as estimações e se existe relação entre valores reais e estimados.

Para este estudo será utilizado o mesmo critério que Jones, Schonlau e Welch (1998). Após amostrar o conjunto de treinamento, os autores calculam o resíduo padronizado de validação cruzada:

$$r_{\text{cruz}} = \frac{y(\mathbf{x}_i) - \hat{y}_{-i}(\mathbf{x}_i)}{\hat{\sigma}_{K-i}(\mathbf{x}_i)} \quad (2.23)$$

onde $\hat{y}_{-i}(\mathbf{x}_i)$ é o estimador de Krigagem retirando a amostra i e $\hat{\sigma}_{K-i}(\mathbf{x}_i)$ é o desvio-padrão de estimacão com Krigagem retirando a amostra i .

De acordo com Jones, Schonlau e Welch (1998), se o modelo é válido, r_{cruz} deve ficar dentro do intervalo $[-3,3]$, o que corresponde a um nível de confiança de, aproximadamente, 99.7%. A ideia por trás da validação cruzada é remover uma observação do conjunto de dados e, então, estimá-la de volta com os $N - 1$ pontos restantes. O procedimento é repetido para todos os N

pontos amostrais.

Para ilustrar o procedimento de análise de resíduos, 10 amostras igualmente espaçadas foram coletadas sobre a função $y = -x \times \cos(2x) + \sqrt{x^3}$. A Figura 2.4 mostra as amostras coletadas e suas estimações por validação cruzada, além do intervalo de confiança da estimação (95% de confiança).

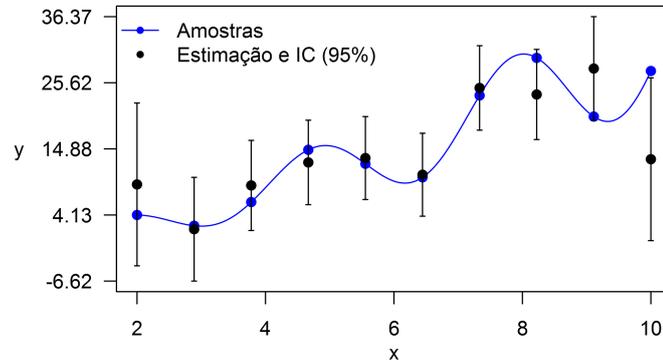


Figura 2.4: Exemplo: Amostras coletadas e suas estimações por validação cruzada para análise de resíduo

O programa Minitab foi usado para verificar a normalidade dos resíduos. O Minitab aplica o teste de normalidade Kolmogorov-Smirnov (KS) para comparar a distribuição empírica acumulada dos dados em relação à distribuição se os dados forem normalmente distribuídos. Se o p-valor do teste for abaixo do que o nível de significância α estabelecido, a hipótese de normalidade é rejeitada. A Figura 2.5 plota a análise de resíduos para o exemplo anterior.

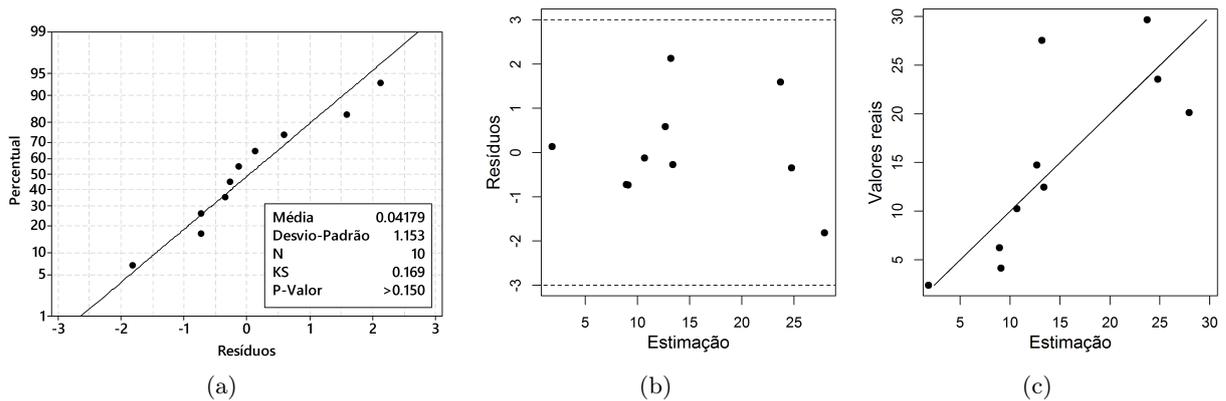


Figura 2.5: Análise de resíduos: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

A Figura 2.5(a) plota o teste de normalidade. Como o p-valor do teste KS é maior do que o nível de significância de 5%, a suposição de normalidade não pode ser rejeitada. A Figura 2.5(b) mostra os resíduos em relação aos valores estimados. Todos os resíduos se encontram dentro do

intervalo $[-3,3]$ e não apresentam nenhuma correlação visível entre os resíduos e as estimações. Finalmente, a Figura 2.5(c) plota os valores reais contra as estimações. Todos os pontos aparentam estar distribuídos aleatoriamente ao redor da linha em 45° , o que sugere precisão suficiente na estimação.

O procedimento de análise de resíduos apresentado anteriormente será usado para todos os Planejamentos de Experimentos realizados na aplicação de otimização para simulação no Capítulo 6.

Capítulo 3

Otimização para simulação

Como enfatizado no capítulo anterior, a metamodelagem surgiu como uma metodologia para realizar otimização de uma função caixa-preta, ou seja, de um modelo de simulação cuja obtenção de uma nova observação demanda esforço e tempo computacional considerável. O processo de otimização para simulação segue, então, os passos abordados na Figura 3.1.

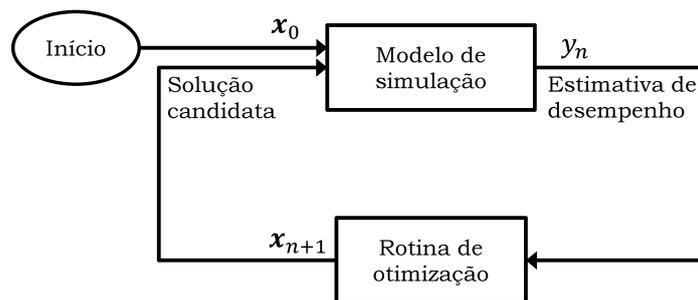


Figura 3.1: Fluxograma do processo geral de otimização para simulação

Um procedimento de otimização para simulação será, portanto, tanto mais eficiente e veloz quanto melhor for a sua rotina de otimização. O uso de metamodelos, principalmente Krigagem, se tornam bastante adequados, pois estimam pontos não simulados, evitando que o processo de otimização passe por regiões não desejadas do seu domínio, além de fornecer um erro de estimação (variância de estimação) que aumenta em regiões menos exploradas do espaço amostral. Essa característica da Krigagem permite que uma rotina de otimização faça buscas locais em regiões bem exploradas e buscas globais em regiões pouco exploradas.

O algoritmo *Efficient Global Optimization* (EGO) proposto por Jones, Schonlau e Welch (1998) é um método popular, se não um dos mais populares, baseado em Krigagem para realizar otimização sequencial de uma função desconhecida (KLEIJNEN; BEERS; NIEUWENHUYSE, 2012), além de ter surgido como uma das abordagens mais promissoras para otimização de

simuladores caros computacionalmente (BASUDHAR et al., 2012).

No algoritmo, dado um modelo de Krigagem construído e validado, novos pontos são iterativamente inseridos às amostras de treinamento de forma a melhorar o ponto de mínimo (dado um problema de minimização). Viana, Haftka e Watson (2010), por outro lado, sugere a implementação do algoritmo EGO com múltiplos metamodelos, em que, a cada iteração, novos pontos são obtidos com cada um deles e todos inseridos às amostras de treinamento no mesmo momento. A Figura 3.2 mostra os passos seguidos no algoritmo EGO.

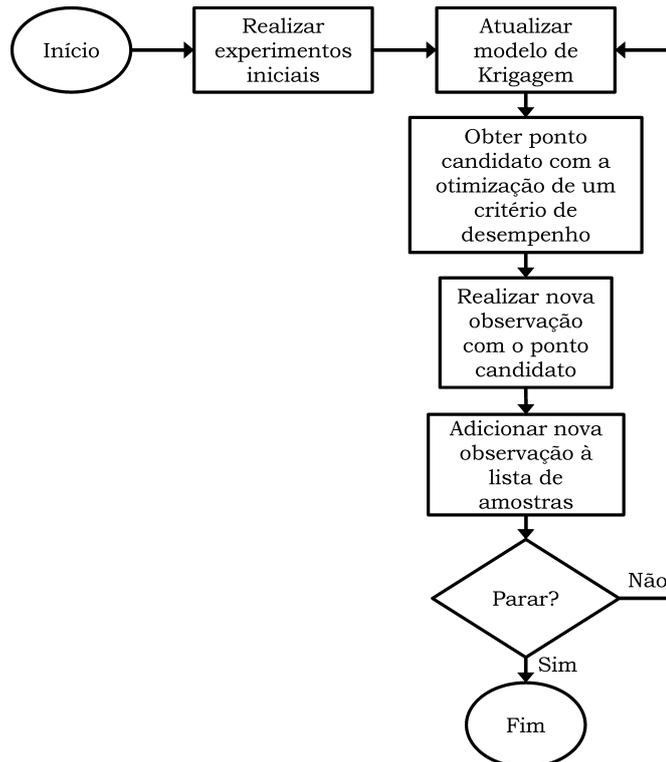


Figura 3.2: Passos do algoritmo EGO. Baseado em Sasena (2002).

O primeiro passo do algoritmo é conduzido com a realização de um Planejamento de Experimentos da escolha do analista. O conjunto de treinamento inicial deve ser obtido de forma a respeitar as suposições do modelo, a partir da análise de resíduos, e proporcionar uma qualidade de estimação suficiente. Em seus experimentos, Picheny et al. (2013) mostram que o tamanho do conjunto de treinamento inicial proporciona pouca influência sobre o resultado da otimização.

A escolha do ponto candidato a ótimo é baseada na maximização ou minimização de um critério de desempenho. Alguns exemplos são: o critério *Expected Improvement* (EI) também proposto por Jones, Schonlau e Welch (1998), utilizado para modelos determinísticos; a função *Lower confidence bounding* (LCB) proposto por Cox e John (1997), que utiliza o intervalo de confiança da estimação de Krigagem tanto para problemas determinísticos quanto estocásticos; o

critério *Augmented Expected Improvement* (AEI) proposto por Huang et al. (2006) para modelos com ruído homogêneo, no qual uma componente multiplicativa é adicionada ao critério EI para contabilizar a variância do ruído homogêneo; o critério *Expected Quantile Improvement* (EQI) proposto por Picheny et al. (2013), empregado em modelos estocásticos com limite na quantidade de pontos a ser obtida, gerando, entretanto, uma complexidade maior de implementação, pois exige o conhecimento ou a suposição de uma função que represente a variância do ruído em uma amostra não coletada; o Procedimento de Reinterpolação (RI) proposto por Forrester, Keane e Bressloff (2006), utilizado em aprendizado de máquina em Koch et al. (2015) e como comparação a outros critérios em Picheny, Wagner e Ginsbourger (2013).

Sasena, Papalambros e Goovaerts (2002) e Basudhar et al. (2012) propuseram, ainda, formas de lidar com o emprego do algoritmo EGO em problemas de otimização com restrições. Sasena, Papalambros e Goovaerts (2002) empregam penalidades grandes caso o novo ponto a ser amostrado pertença a uma região inviável. Por outro lado, Basudhar et al. (2012) utilizam máquinas de vetores de suporte (*support vector machine* em inglês, SVM) que, ao invés de estimar os valores das restrições como nos métodos tradicionais, apenas classificam o novo ponto como viável ou inviável.

Nas seções subsequentes são apresentados o critério EI e variações do algoritmo EGO oriundas da presença de ruído heterogêneo nos experimentos. Relembrando: nosso intuito é otimizar um modelo de simulação de filas, por isso a presença de ruído heterogêneo nos experimentos não pode ser negligenciada. Dessa forma, primeiramente, o critério EI é apresentado, pois é o mais empregado para os modelos determinísticos de Krigagem e é, também, usado como base dos procedimentos para o caso estocástico. O critério EI é detalhado na Seção 3.1.

No âmbito estocástico, destaca-se o procedimento RI proposto por Forrester, Keane e Bressloff (2006). Nele, primeiramente, um modelo estocástico de Krigagem é obtido com os dados originais, depois um modelo de Krigagem determinístico é gerado com os mesmos parâmetros do anterior, mas com o intuito de filtrar o ruído aleatório. O critério EQI também poderia ser utilizado neste caso, entretanto, como este modelo requer o conhecimento de uma função para estimar a variância do ruído em um ponto não amostrado, ele foi desconsiderado. O procedimento RI é detalhado na Seção 3.2.

O procedimento RI trabalha com uma estimação da média dos experimentos após a remoção do ruído aleatório, entretanto, quando tratamos de modelos de simulação de filas, estamos muitas vezes interessados no intervalo de confiança da média dos experimentos e não na média propria-

mente dita. Este fato motiva uma proposta de alteração do modelo estocástico de Krigagem no algoritmo EGO, onde o modelo estocástico é substituído por outro determinístico (interpolador exato) que estima o lado superior (ou inferior) do intervalo de confiança da média. Este novo modelo será chamado daqui em diante de KOIC. A Seção 3.3 mostra com mais detalhes o método proposto e apresenta um exemplo ilustrativo.

3.1 Critério *Expected Improvement*

Tratando de um problema de minimização e utilizando o critério EI, temos o interesse de calcular uma melhoria sobre o ponto de mínimo encontrado em cada iteração. Suponha que o valor de resposta seja uma realização de uma variável aleatória $y(\mathbf{x})$, seguindo um Processo Gaussiano com variância σ_Z^2 . A melhora $I(\mathbf{x})$ em \mathbf{x} em relação ao atual ponto de mínimo $\min(\mathbf{y})$ das amostras de treinamento é dada por:

$$I(\mathbf{x}) = \max(0, \min(\mathbf{y}) - y(\mathbf{x})) \quad (3.1)$$

Usando Krigagem para estimar um ponto candidato, temos $\hat{y}(\mathbf{x})$ como o valor estimado da variável de resposta e $\hat{\sigma}_K^2$ como sua variância de estimação, o valor esperado de melhora em um ponto candidato é dado com:

$$EI = \begin{cases} [\min(\mathbf{y}) - \hat{y}(\mathbf{x})] \Phi\left(\frac{\min(\mathbf{y}) - \hat{y}(\mathbf{x})}{\hat{\sigma}_K(\mathbf{x})}\right) + \hat{\sigma}_K(\mathbf{x}) \phi\left(\frac{\min(\mathbf{y}) - \hat{y}(\mathbf{x})}{\hat{\sigma}_K(\mathbf{x})}\right) & \text{se } \hat{\sigma}_K^2(\mathbf{x}) > 0 \\ 0 & \text{se } \hat{\sigma}_K^2(\mathbf{x}) = 0 \end{cases} \quad (3.2)$$

onde $\Phi(\cdot)$ e $\phi(\cdot)$ são, respectivamente, a função de distribuição acumulada e a função densidade de probabilidade da distribuição normal padronizada $\mathcal{N}(0,1)$. Com $\hat{\sigma}_K^2(\mathbf{x}) > 0$, o primeiro termo do indicador mostra a diferença entre um novo ponto estimado por Krigagem e o mínimo corrente multiplicada pela probabilidade do valor da variável de resposta ser menor ou igual a $\min(\mathbf{y})$, tendo portanto um valor alto quando é provável que $\hat{y}(\mathbf{x})$ é menor ou igual a $\min(\mathbf{y})$. Já o segundo termo denota a multiplicação entre o desvio padrão de estimação multiplicado pela probabilidade do valor da variável de resposta ser igual a $\min(\mathbf{y})$.

Um novo ponto é escolhido para ser adicionado às amostras de treinamento quando maximiza

EI na respectiva iteração, ou seja:

$$\mathbf{x} = \arg \max_{\mathbf{x}} EI(\mathbf{x}) \quad (3.3)$$

além disso o algoritmo é encerrado quando:

$$\max_{\mathbf{x}} EI(\mathbf{x}) < \delta_1, \quad (3.4)$$

ou seja, quando a melhora máxima é pequena o suficiente. Jones, Schonlau e Welch (1998) sugere que o algoritmo pare com $\delta_1 = 0,01 \times \min(\mathbf{y})$. Huang et al. (2006) afirmam que na prática é conveniente utilizar como critério de parada o EI Máximo Relativo, fazendo:

$$\frac{\max_{\mathbf{x}} EI(\mathbf{x})}{\max(\mathbf{y}) - \min(\mathbf{y})} < \delta_2 \quad (3.5)$$

Os autores advogam o uso deste critério por possibilitar que o usuário escolha uma precisão δ_2 sem ter conhecimento prévio da magnitude dos valores de resposta.

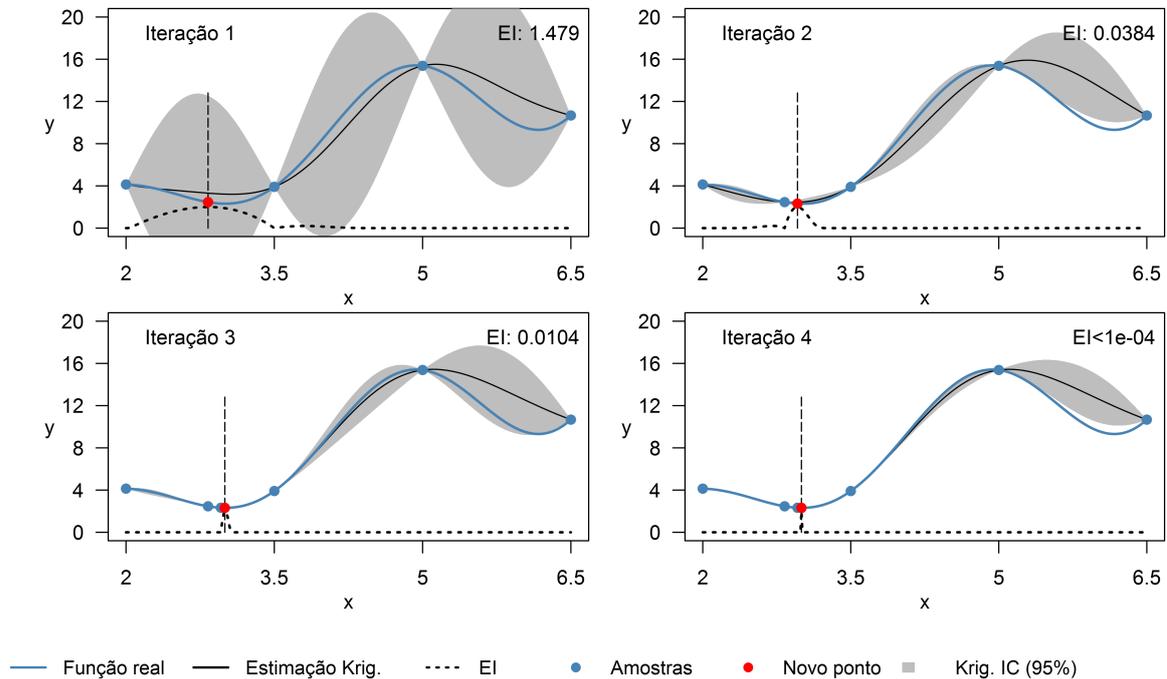


Figura 3.3: Iterações necessárias para minimizar a função $y = -x \sin(2x) + \sqrt{x^3}$ com o algoritmo EGO e o indicador EI. Melhor solução em $x = 3$ com $y = 2,316$.

A Figura 3.3 mostra um exemplo da aplicação do algoritmo EGO com o critério EI no contexto determinístico. Quatro observações iniciais foram realizadas aleatoriamente em relação à função $y(x) = -x \sin(2x) + \sqrt{x^3}$. Apenas 4 iterações foram necessárias para encontrar o ponto de mínimo

da função escolhida, com uma precisão $\delta_1 = 0,001 \min(\mathbf{y})$. O intervalo de confiança da média (IC) foi construído para um nível de confiança de 95% e os valores de EI foram multiplicados por $\frac{2}{\max(ET)}$.

É perceptível, principalmente nas iterações 2, 3 e 4, como o algoritmo EGO se utiliza da informação da variância de estimação de Krigagem para fazer buscas globais em regiões menos exploradas do espaço de análise. A solução ótima real de maximização da função escolhida é $x = 4,934174$ com $y = 15,41695$, com isso o erro percentual do algoritmo é na ordem de $4,85 \times 10^{-4}\%$.

3.2 Procedimento de Reinterpolação

O Procedimento de Reinterpolação foi proposto por Forrester, Keane e Bressloff (2006). Ao contrário de outras técnicas que tratam de Krigagem estocástica, os autores não modificam o critério EI. Neste método, um modelo de Krigagem *regressivo* (não é interpolador exato) e um modelo de Krigagem *interpolador exato* são usados simultaneamente.

O modelo *regressivo* é obtido como descrito na Seção 2.2. Para o modelo *interpolador*, todos os parâmetros e conjunto de treinamento são herdados do anterior, entretanto no seu vetor de variáveis de resposta são inseridas suas estimações, obtidas com o modelo *regressivo*, ou seja, $y_{\text{interp}}(\mathbf{x}_i) = \hat{y}_{\text{reg}}(\mathbf{x}_i)$, $\forall i \in \{1, 2, \dots, N\}$, onde $y_{\text{interp}}(\mathbf{x}_i)$ denota o valor de resposta da amostra i no modelo *interpolador* e $\hat{y}_{\text{reg}}(\mathbf{x}_i)$ se refere à estimação da amostra i feita com o modelo *regressivo*.

Por fim, o algoritmo EGO com o critério EI é utilizado para realizar otimização para simulação e o modelo de Krigagem *interpolador* é aplicado para estimar pontos não amostrados e calcular a variância dessa estimação.

A Figura 3.4 mostra uma aplicação simples do Procedimento RI, com o algoritmo EGO e critério EI. Quatro observações iniciais igualmente espaçadas foram realizadas em relação à função $y|x \sim \mathcal{N}(\mu = -x \sin(2x) + \sqrt{x^3}, \sigma^2 = \cos(x)^2 + 0.01)$, com $2 \leq x \leq 6.5$. O intervalo de confiança da média (IC) foi construído para um nível de confiança de 95% e os valores de EI foram multiplicados por $\frac{2}{\max(ET)}$.

Note-se que, nos quadros da esquerda, a estimação com Krigagem deixa de ser um interpolador exato, sendo verificado pelo fato de alguns círculos pretos (estimação das amostras) não estarem localizados exatamente no centro dos azuis (amostras). Apenas 4 iterações foram necessárias para encontrar o ponto de mínimo da função escolhida, dado uma precisão $\delta_1 = 0,001 \min(\mathbf{y})$.

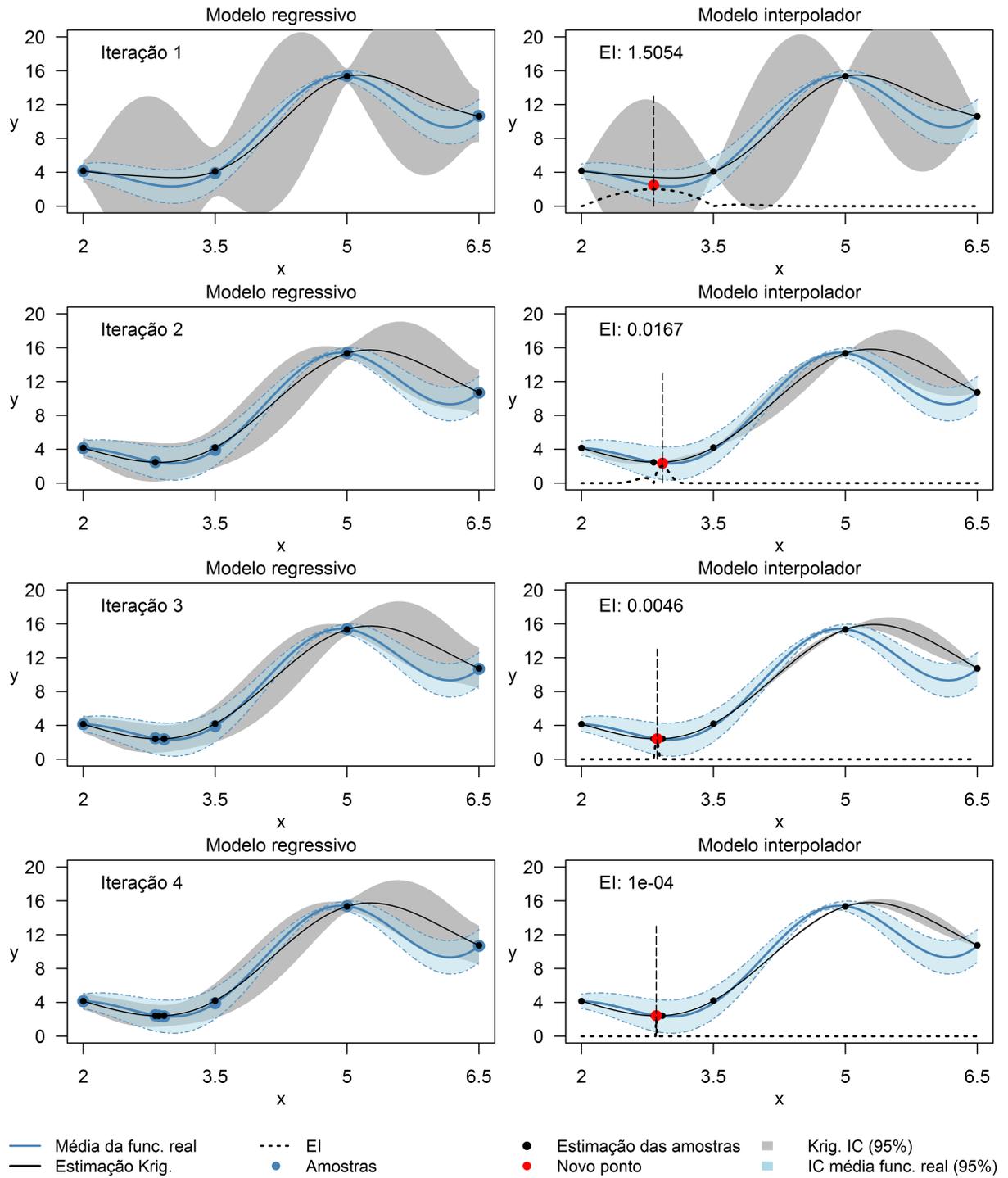


Figura 3.4: Iterações necessárias para otimizar a função de média $y = -x \sin(2x) + \sqrt{x^3}$, com variância $Var(y) = \cos(x)^2 + 0.01$, usando o Procedimento RI, com o algoritmo EGO e o indicador EI. Melhor solução em $x = 2.92$, com $\bar{y} = 2.352$ e $\sigma_y^2 = 0.962$.

O ponto de mínimo foi encontrado em $x = 2.92$, com $\bar{y} = 2.352$ e $\sigma_y^2 = 0.962$.

3.3 Krigagem do Intervalo de Confiança da Média

Ao trabalhar com modelos de simulação de eventos discretos estamos interessados, em grande parte das vezes, no intervalo de confiança da média e não na média propriamente dita, portanto, nesta seção, é proposta a otimização o pior caso possível para a média.

O método proposto nessa seção foi chamado de KOIC e estima um dos lados do intervalo de confiança da média da variável de resposta. Neste caso, o comportamento da variável de resposta é influenciado também pela presença do desvio-padrão, o que gera um vício, pois o comportamento do percentil superior ou inferior é diferente do comportamento da média. Entretanto, essa abordagem permite a utilização de um modelo determinístico de Krigagem, cujo estimador é um interpolador exato, o que diminui substancialmente a complexidade do processo de otimização.

Para o uso do modelo KOIC no lugar do modelo de Krigagem determinística em um problema de minimização, o vetor de variáveis de resposta \mathbf{y} é substituído por $\bar{\mathbf{y}} + \Phi_{\alpha/2}^{-1}\boldsymbol{\tau}$, onde $\bar{\mathbf{y}}$ é um vetor com as médias dos experimentos, $\Phi_{\alpha/2}^{-1}$ denota o número de desvios-padrões em relação à média para um nível de confiança de $100 \times (1 - \alpha)\%$ e $\boldsymbol{\tau}$ é um vetor com os desvios-padrões das replicações dos experimentos. O modelo KOIC é usado com o critério EI no algoritmo EGO e o processo de estimação dos seus parâmetros segue os passos definidos na Seção 2.2.

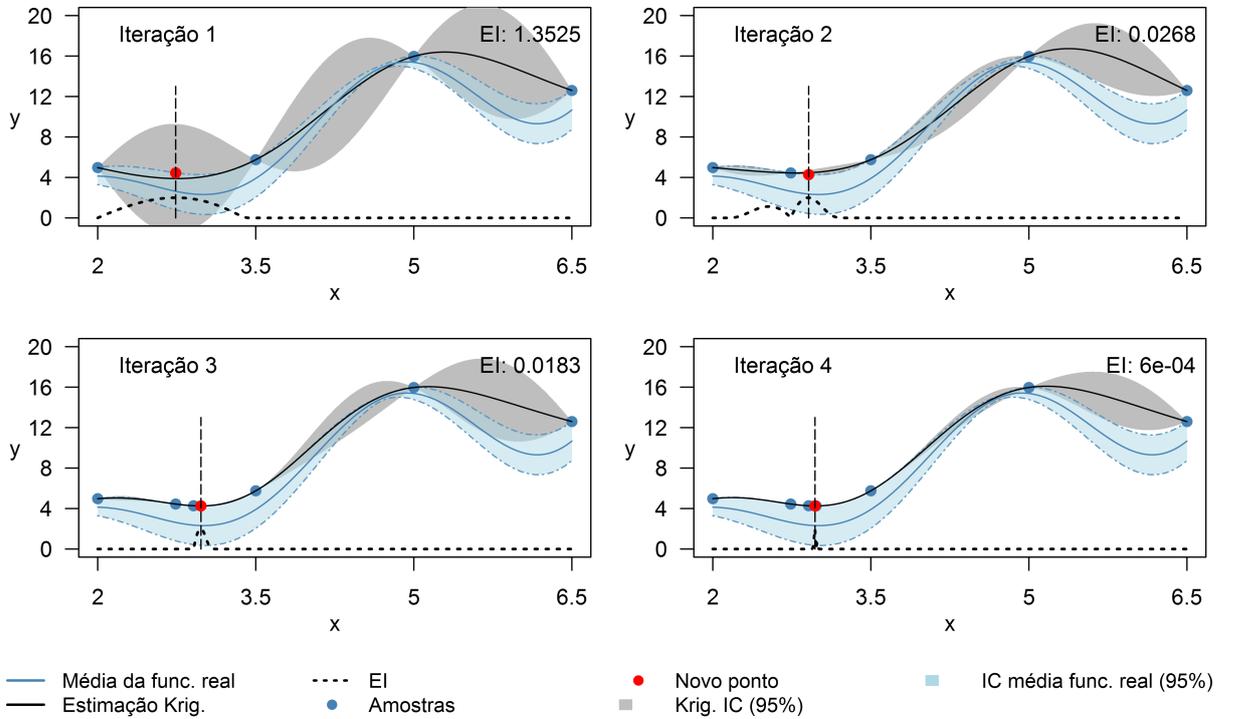


Figura 3.5: Iterações necessárias para otimizar a função de média $y = -x \sin(2x) + \sqrt{x^3}$, com variância $Var(y) = \cos(x)^2 + 0.01$, usando o modelo KOIC, com o algoritmo EGO e o indicador EI. Melhor solução em $x = 2.99$, $\bar{y} = 2.317$ e $\sigma_y^2 = 0.987$.

A Figura 3.5 mostra a utilização do modelo KOIC para a minimização do mesmo esquema de função, ruído e critério de parada apresentados na Seção 3.2 para o exemplo ilustrativo do procedimento RI. Utilizando o modelo KOIC, o algoritmo agora varre o lado superior do intervalo de confiança e não mais a média. Nota-se que o mínimo encontrado para a função utilizada foi ligeiramente menor (1,5%) do que o obtido com o procedimento RI, ao se comparar apenas as médias, entretanto as soluções não são estatisticamente diferentes levando em consideração o intervalo de confiança como um todo.

Capítulo 4

Modelo de Simulação: o SAMU de Belo Horizonte

Todo o embasamento teórico levantado nos capítulos anteriores são usados neste trabalho para realizar otimização sobre um modelo de simulação do sistema de atendimento médico de urgência (SAMU). Neste capítulo são apresentadas as características desse sistema e o modelo de simulação implementado.

4.1 Descrição do sistema

O Sistema de Atendimento Móvel de Urgência (SAMU) foi estabelecido em 2003 pelo Governo Federal é disponível 24 horas por dia. O principal objetivo do sistema é assistir, de maneira rápida e organizada com o mínimo de recursos, todos os eventos de emergência que necessitam serviços médicos. Para atender às emergências, o sistema possui unidades de ambulâncias básicas, avançadas e de tratamento psiquiátrico atuantes em todo país.

O modelo de simulação que representa esse sistema de emergência foi introduzido em Silva e Pinto (2010) e em Pinto, Silva e Young (2015). Os autores forneceram uma metodologia em como construir um modelo de simulação genérico do sistema de atendimento de emergência. Além disso, eles implementaram a metodologia na cidade de Belo Horizonte e validaram o modelo com dados reais. Zhang et al. (2015) desenvolveu um estudo similar, mas o aplicou em Xangai, China.

Belo Horizonte tem 331.401 km² com aproximadamente 2,5 milhões de pessoas. A cidade é dividida em nove regiões administrativas: Barreiro, Oeste, Centro-Sul, Leste, Norte, Venda Nova,

Pampulha, Nordeste e Noroeste. Em 2010, Belo Horizonte possuía 17 ambulâncias básicas, 5 avançadas e 1 de tratamento psiquiátrico com cronograma fixo (PINTO; SILVA; YOUNG, 2015), ou seja, as ambulâncias são mantidas sempre nas mesmas bases independentemente do dia da semana e da hora no dia.

O sistema trabalha também com 22 instalações de atendimento médico distribuídas em 8 Unidades de Pronto Atendimento (UPA) para casos de média complexidade, 6 hospitais para casos de alta complexidade e trauma e 8 Centros de Tratamento em Saúde Mental (CERSAM) para casos de atendimento psiquiátrico e abuso de drogas. A Tabela 4.1 mostra os dados referentes às bases de ambulâncias e suas coordenadas em UTM (Universal Transverse Mercator) e a Tabela 4.2 mostra as coordenadas das instalações médicas em UTM.

Tabela 4.1: Bases de ambulâncias

Base	Tipos de ambulância			Coordenadas em UTM	
	Básicas	Avançadas	Psiquiátricas	X	Y
1	-	1	-	608888	7803873
2	1	1	-	610267	7808237
3	1	1	-	605601	7792252
4	-	1	-	611688	7795582
5	-	1	-	612472	7796754
6	1	-	-	609592	7808252
7	1	-	-	607995	7809288
8	1	-	-	610299	7798655
9	1	-	-	607817	7797622
10	1	-	-	613308	7802579
11	2	-	-	614481	7798991
12	1	-	-	610927	7797568
13	1	-	-	611736	7797381
14	1	-	-	603253	7791096
15	1	-	-	612299	7803342
16	1	-	-	604483	7802483
17	1	-	-	607832	7793931
18	1	-	-	602740	7788779
19	1	-	-	609016	7808665
20	-	-	1	612315	7796598

A Figura 4.1 apresenta o mapa de Belo Horizonte com suas nove regiões administrativas. As localizações de todas as 20 bases são mostradas em preto e das instalações de atendimento médico em vermelho. A Figura 4.2 plota a distribuição média das chamadas de emergência realizadas em Belo Horizonte. E a Figura 4.3 plota a distribuição média de chamada por faixa horária. O modelo de simulação foi construído a partir da análise de dados históricos corresponde a um banco de dados contendo 1.734.416 chamadas realizadas entre outubro de 2006 e agosto de 2009

Tabela 4.2: Coordenadas das instalações médicas utilizadas no modelo de simulação

Instalação médica	Tipo	Coordenadas em UTM	
		X	Y
1	UPA	611767	7796253
2	UPA	611895	7796328
3	UPA	610298	7798654
4	UPA	610267	7808237
5	UPA	605555	7789493
6	UPA	605728	7797968
7	UPA	602739	7788778
8	UPA	607831	7793930
9	Hospital	612471	7796754
10	Hospital	614441	7797973
11	Hospital	612298	7803342
12	Hospital	609592	7808252
13	Hospital	604483	7802482
14	Hospital	611977	7799775
15	CERSAM	608520	7794541
16	CERSAM	602432	7790747
17	CERSAM	607888	7797643
18	CERSAM	613933	7797892
19	CERSAM	611195	7800284
20	CERSAM	609382	7808220
21	CERSAM	607766	7805644

da cidade de Belo Horizonte.

Hoje, Belo Horizonte possui ligeiramente mais bases de ambulâncias e unidades. Além disso, os dados de chegada de chamadas mudaram devido ao crescimento populacional natural. Entretanto dificuldades na atualização dos dados (localização das bases, número total de ambulâncias, distribuição das chamadas pela cidade, etc.) impediram a análise do estado atual do sistema. Dessa forma, optou-se pelo uso dos dados já coletados em 2010 por Silva e Pinto (2010) e Pinto, Silva e Young (2015), o que não prejudicou a realização do trabalho.

4.2 O Modelo de Simulação

A lógica do modelo de simulação é dividida em serviço no Call Center, despacho de ambulância, serviço na cena e movimentação do paciente à instalação médica.

O processo de salvamento começa com a chegada de uma chamada de emergência recebida no Call Center, sendo atendida por um atendente. Antes de iniciar os procedimentos necessários, o modelo primeiramente amostra aleatoriamente as características da chamada: origem geográfica, tipo (clínica, trauma ou psiquiátrica), ação médica (despacho de ambulância ou fornecer

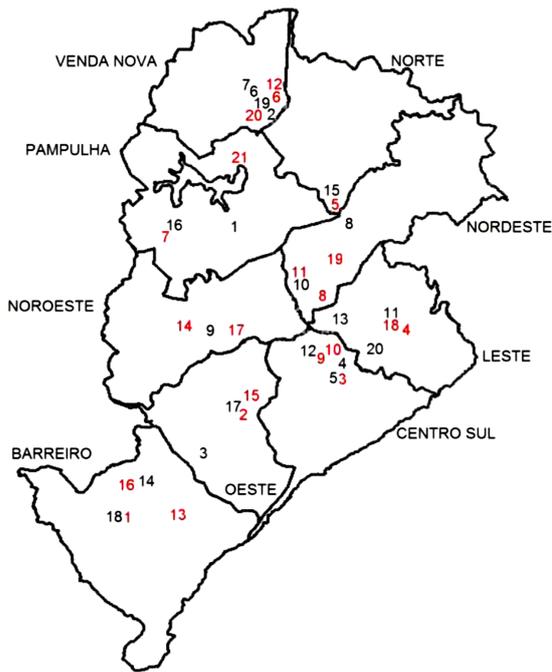


Figura 4.1: Mapa de Belo Horizonte dividido nas 9 regiões administrativas. Os número em preto denotam bases de ambulâncias (1-20) e os em vermelho denotam UPAs (1-8), hospitais (9-14) e CERSAMs (15-21). Fonte: Silva e Pinto (2010).

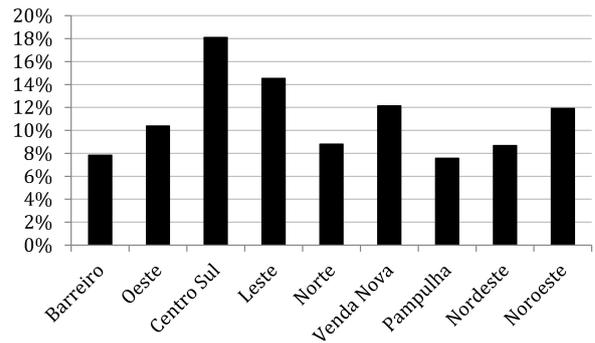


Figura 4.2: Distribuição média de chamadas de emergência realizadas em Belo Horizonte entre outubro de 2006 e agosto de 2009. Baseado em: Silva e Pinto (2010)

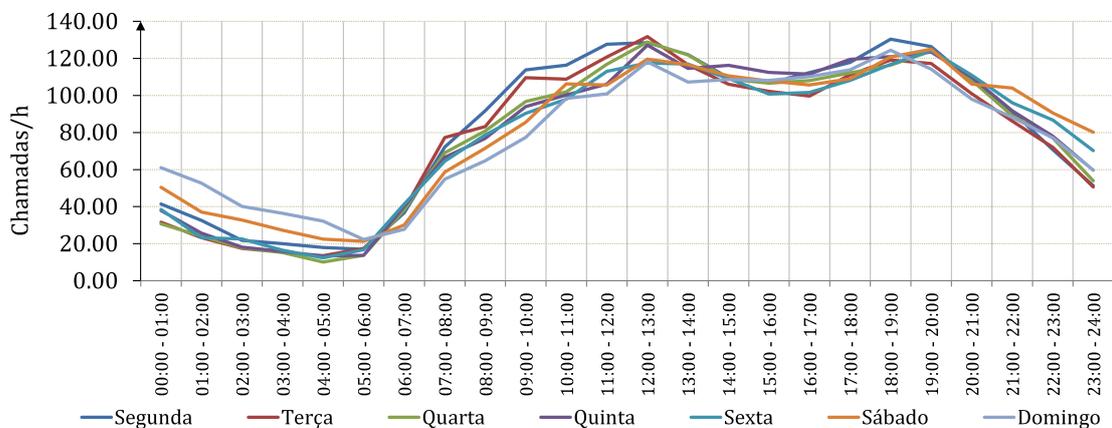


Figura 4.3: Distribuição média de chamadas por faixa horária. Fonte: Silva e Pinto (2010).

orientação por telefone) e as coordenadas.

Um atendente realiza a análise inicial da chamada para identificar sua natureza (emergência real ou engano). Se a emergência for verificada, as informações iniciais, como localização do incidente e dados da vítima, são coletados. Um médico analisa a gravidade da emergência e decide por despachar ou não uma ambulância.

Se uma ambulância for despachada, uma equipe médica realiza o tratamento inicial na cena

e então move o paciente para a instalação médica mais próxima. Após liberá-lo, a equipe determina a necessidade para substituição de material e/ou manutenção. Após, a ambulância se torna disponível. Se outro serviço de atendimento tiver sido agendado, a ambulância procede diretamente para o novo incidente, senão é volta para a respectiva base.

Ambulâncias avançadas e de atendimento psiquiátrico não foram consideradas aqui, pois, comparadas com as básicas, têm relativamente pouca demanda. Além disso, o mesmo modelo de simulação de Silva e Pinto (2010) e Pinto, Silva e Young (2015) foi implementado, mas, enquanto os autores utilizam como simulador o Arena 11, foi empregado aqui o Anylogic 7. Cada rodada de replicações do simulador levou na média aproximadamente 1,34 min.

Com o modelo de simulação é obtido, para as ambulâncias básicas, o tempo de resposta médio após um ano de simulação. O tempo de resposta é calculado pela diferença entre o momento em que a chamada é criada na central de atendimento até o instante em que uma ambulância chega na cena. Para avaliar a aleatoriedade do modelo, 10 replicações foram executadas, de forma a obter a média e variância do tempo de resposta médio, assumindo que a variável segue uma distribuição normal. Para simplificar a nomenclatura, a média e variância do tempo de resposta médio serão tratados, respectivamente, apenas como tempo de resposta médio e variância do tempo de resposta.

4.3 Otimização do tempo de resposta

Com o modelo de simulação e o algoritmo EGO, o tempo de resposta das ambulâncias básicas é minimizado a partir da maximização do critério EI. O problema de otimização a ser resolvido é o seguinte:

$$\begin{aligned}
 & \underset{\mathbf{x}}{\text{maximizar}} && EI(\mathbf{x}) \\
 & \text{sujeito a} && \sum_{k=1}^K x_k = \text{TotAmb} \\
 & && 0 \leq x_k \leq 3, \forall k \in \{1, 2, 3, \dots, K\}
 \end{aligned} \tag{4.1}$$

onde TotAmb é o total de ambulâncias empregadas no sistema, x_k denota o número de ambulâncias alocadas à base k e \mathbf{x} é o um vetor que agrega todas as variáveis x_k . A função objetivo maximiza o critério EI para a obtenção de um ponto candidato a mínimo, a primeira restrição limita o número de ambulâncias a um valor fixo e a segunda representa uma limitação de infraestrutura válida para as bases de ambulâncias.

Para resolver o Problema 4.1 foi implementado um algoritmo *Simulated Annealing* (SA). SA é uma meta-heurística baseada no processo de resfriamento de metal que realiza movimentos aleatórios para escapar de ótimos locais. O algoritmo é melhor detalhado no Capítulo 5.

Capítulo 5

Heurística de otimização: Simulated Annealing

Como dito anteriormente, a obtenção de um ponto candidato a mínimo é feita com a maximização do critério EI, seja com o procedimento RI ou com o uso do modelo KOIC. O pacote `DiceKriging` do R foi usado nas Seções 3.1, 3.2 e 3.3 para executar essa maximização. Entretanto, no caso de otimização para simulação do modelo de atendimento de chamadas de emergência por ambulâncias, um algoritmo SA precisou ser implementado.

Segundo Goffe, Ferrier e Rogers (1994), SA é uma meta-heurística desenvolvida tomando como referência o processo termodinâmico de recozimento de peças de metal. Se resfriado lentamente, flutuações aleatórias inerentes ao processo permitem que o estado energético do metal escape mínimos locais e alcance o mínimo global. Caso contrário, o metal pode não alcançar a mesma configuração mínima de energia.

O algoritmo maximiza o critério EI e retorna um ponto candidato para ser avaliado com o modelo de simulação. O modelo de simulação possui variáveis de entrada discretas que denotam o número de ambulâncias a serem alocadas por base. Por isso, a otimização do EI é, relativamente, de difícil resolução (ARORA; HUANG; HSIEH, 1994) e, especialmente, por EI não ser linear, nem convexo.

SA é um dos métodos revisados por Arora, Huang e Hsieh (1994), usados para otimizar funções não lineares com variáveis discretas. Os autores também fornecem uma revisão de esquemas de *branch-and-bound*, o método de linearização sequencial, abordagens com funções de penalidade, métodos de relaxação Lagrangiana, dentre outros.

A apresentação do algoritmo é feita em três seções: a estrutura de solução, o movimento de

vizinhança e o algoritmo completo.

A estrutura de solução é composta de 20 variáveis, uma para cada base de ambulâncias, onde, em cada uma, um mínimo de zero e um máximo de três unidades podem ser alocados. Além disso, o número total de ambulâncias no sistema é sempre mantido fixo. Para obter a solução inicial, 100 soluções viáveis aleatórias são geradas. A melhor solução é, então, usada como solução inicial.

A cada iteração o algoritmo utiliza um movimento de vizinhança chamado de TRANSFERÊNCIA. No movimento, duas bases diferentes são escolhidas aleatoriamente. Após, uma unidade é retirada de uma base e transferida para outra. Essa estratégia permite que novas soluções sejam obtidas e que o número de ambulâncias seja mantido o mesmo.

Algoritmo 1: Algoritmo Simulated Annealing

Entrada: $T_0, T_t, L_{\max}, \gamma$
Saída: Ponto candidato

```

1  $T \leftarrow T_0$ ;
2  $\mathbf{x} \leftarrow \mathbf{x}_{\text{inicial}}$ ;
3  $\mathbf{x}^{\text{melhor}} \leftarrow \mathbf{x}$ ;
4  $\psi \leftarrow \psi_{\text{inicial}}$ ;
5  $\psi^{\text{melhor}} \leftarrow \psi$ ;
6 repita
7    $L \leftarrow 0$ ;
8   repita
9      $L \leftarrow L + 1$ ;
10     $\mathbf{x}^{\text{tentativa}} \leftarrow \mathbf{x}$ ;
11    Realizar TRANSFERÊNCIA em  $\mathbf{x}^{\text{tentativa}}$ ;
12    Avaliar  $\mathbf{x}^{\text{tentativa}}$  para obter  $\psi^{\text{tentativa}}$ ;
13    se  $\psi^{\text{tentativa}} < \psi$  então
14       $\mathbf{x} \leftarrow \mathbf{x}^{\text{tentativa}}$ ;
15       $\psi \leftarrow \psi^{\text{tentativa}}$ ;
16      se  $\psi^{\text{tentativa}} < \psi^{\text{melhor}}$  então
17         $\mathbf{x}^{\text{melhor}} \leftarrow \mathbf{x}^{\text{tentativa}}$ ;
18         $\psi^{\text{melhor}} \leftarrow \psi^{\text{tentativa}}$ ;
19      fim
20    senão
21      se  $\text{Uniforme}(0,1) \leq \exp^{(\psi^{\text{tentativa}} - \psi)/T}$  então
22         $\mathbf{x} \leftarrow \mathbf{x}^{\text{tentativa}}$ ;
23         $\psi \leftarrow \psi^{\text{tentativa}}$ ;
24      fim
25    fim
26    até  $L = L_{\max}$ ;
27     $T \leftarrow T \times \gamma$ ;
28 até  $T \geq T_t$ ;
29 retorna  $\mathbf{x}^{\text{melhor}}$ ;
```

O procedimento completo é mostrado no Algoritmo 1, onde T_0 é a temperatura inicial, T_t é

a temperatura de término, T é a temperatura corrente, L_{\max} é o número máximo de iterações no *loop* interior, L é o número corrente de iterações no *loop* interior, γ é a fator de redução de temperatura e ψ , $\psi^{\text{tentativa}}$, ψ^{melhor} são, respectivamente, o valor atual, valor tentativa e melhor valor da função objetivo.

O Algoritmo 1 aplica movimentos TRANSFERÊNCIA na esperança de melhorar a melhor solução \mathbf{x} . Os movimento geram uma solução tentativa $\mathbf{x}^{\text{tentativa}}$ que é avaliada com a função objetivo. Se $\mathbf{x}^{\text{tentativa}}$ for melhor que \mathbf{x} , a solução corrente se torna a solução tentativa, e, se o valor de EI de $\mathbf{x}^{\text{tentativa}}$ for menor do que o da melhor solução, $\mathbf{x}^{\text{melhor}}$ se torna $\mathbf{x}^{\text{tentativa}}$. Por outro lado, se $\mathbf{x}^{\text{tentativa}}$ for pior do que \mathbf{x} , a probabilidade $e^{(\psi^{\text{tentativa}}-\psi)/T}$ é verificada. Se o número aleatório gerado entre zero e um for menor do que $e^{(\psi^{\text{tentativa}}-\psi)/T}$, \mathbf{x} recebe $\mathbf{x}^{\text{tentativa}}$.

Uma das forças do SA é fato de que ele não descarta imediatamente uma solução ruim. Se a probabilidade $\exp^{(\psi^{\text{tentativa}}-\psi)/T}$ for verificada, esta solução ruim é mantida para que mais um movimento seja executado na tentativa de melhorá-la. Este esquema também permite que o algoritmo escape de pontos de máximo locais.

5.1 Definição dos parâmetros de SA

Para definir os valores dos parâmetros do algoritmo SA (T_0 , T_t , L_{\max} , γ), alguns testes foram realizados variando seus valores, em relação a dois conjuntos de treinamento específicos. Quanto maior for o valor alocado aos parâmetros, maior será a precisão do algoritmo, entretanto o tempo computacional também aumentará exponencialmente.

Deve-se levar em consideração que o algoritmo SA é executado a cada iteração, dessa forma quanto menor o tempo computacional necessário para resolver o Problema 4.1, menor será o tempo total da otimização. Tanto o modelo KOIC quanto o procedimento RI foram usados no critério EI e algoritmo EGO para realizar os testes. Devido a aleatoriedade inerente do algoritmo SA, 20 replicações foram realizadas em cada experimento. Além disso, para facilitar a análise, T_0 foi mantido igual a 2 em todos os experimentos. Finalmente, os outros parâmetros foram variados conforme se segue:

- $T_t = \{0,01; 0,001; 0,0001; 0,00001\}$;
- $L_{\max} = \{40; 60; 80; 100\}$;
- $\gamma = \{0,9; 0,925; 0,95; 0,975\}$

A Figura 5.1 mostra todos os experimentos realizados para um conjunto de treinamento de

tamanho 100, onde todas as amostras somam 17 ambulâncias. Pela figura o parâmetro γ é o que mais influencia nos resultados tanto com o modelo KOIC quanto com o procedimento RI. É perceptível o considerado ganho de precisão com o uso de $\gamma = 0,975$. Neste mesmo contexto, a precisão dos resultados aumenta substancialmente com incrementos no parâmetro L_{\max} , entretanto o tempo computacional também é acrescido aumenta.

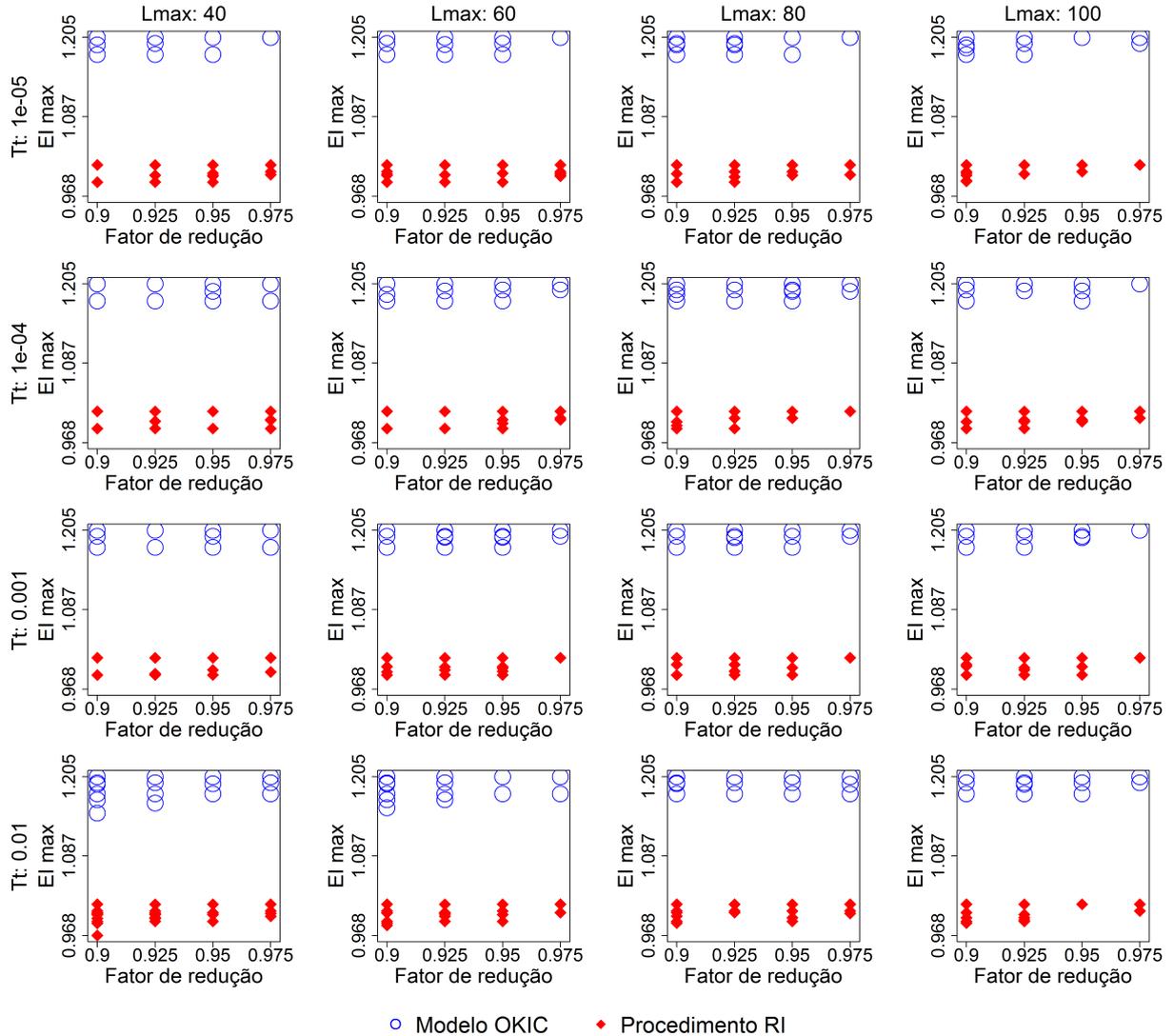


Figura 5.1: Experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, $TotAmb = 17$, $N = 100$ e 20 replicações.

A Figuras 5.2 mostra os mesmos testes das figuras anteriores, mas considerando um conjunto de treinamento de tamanho 100 com 28 ambulâncias no sistema. Neste caso a maioria das combinações de parâmetros gerou o mesmo valor de EI máximo. Desta vez o uso de $T_t = 0,01$ não proporcionou resultados minimamente razoáveis em nenhum combinação de valores de parâmetros.

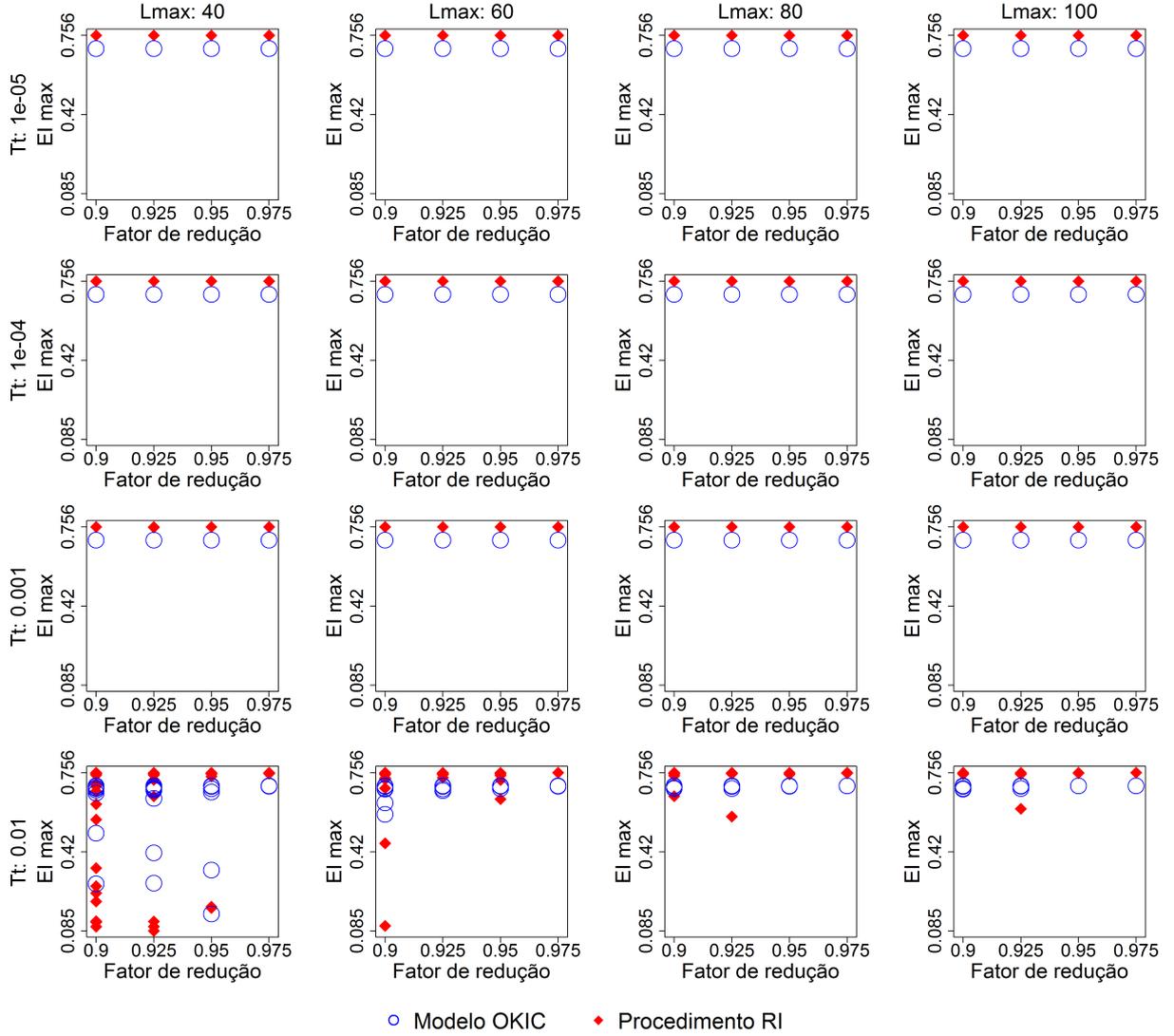


Figura 5.2: Experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, $\text{TotAmb} = 28$, $N = 100$ e 20 replicações.

Para dar ideia de tempo computacional, a Figura 5.3 e a Figura 5.4 plotam os tempos de execução necessários para realizar os experimentos mostrados, respectivamente, na Figura 5.1 e Figura 5.2. É possível verificar que o tempo computacional cresce exponencialmente principalmente com o aumento dos valores do parâmetros.

Segundo os resultados dos experimentos, parece ser suficiente utilizar $L_{\max} = 40$, com $\gamma = 0.95$ e $T_t = 10^{-5}$, além de $T_0 = 2$ (escolhido inicialmente e deixado fixo para todos os testes) no processo de otimização para simulação do modelo de ambulâncias descrito no Capítulo 4. Com tal configuração e utilizando o modelo KOIC, o intervalo de confiança da média (95% de confiança) do tempo computacional é de $0,31105 \pm 0,0166$ segundos com $\text{TotAmb} = 17$ e $0,2799 \pm 0,00476$ com $\text{TotAmb} = 28$. Utilizando o procedimento RI, o mesmo intervalo de confiança fica em

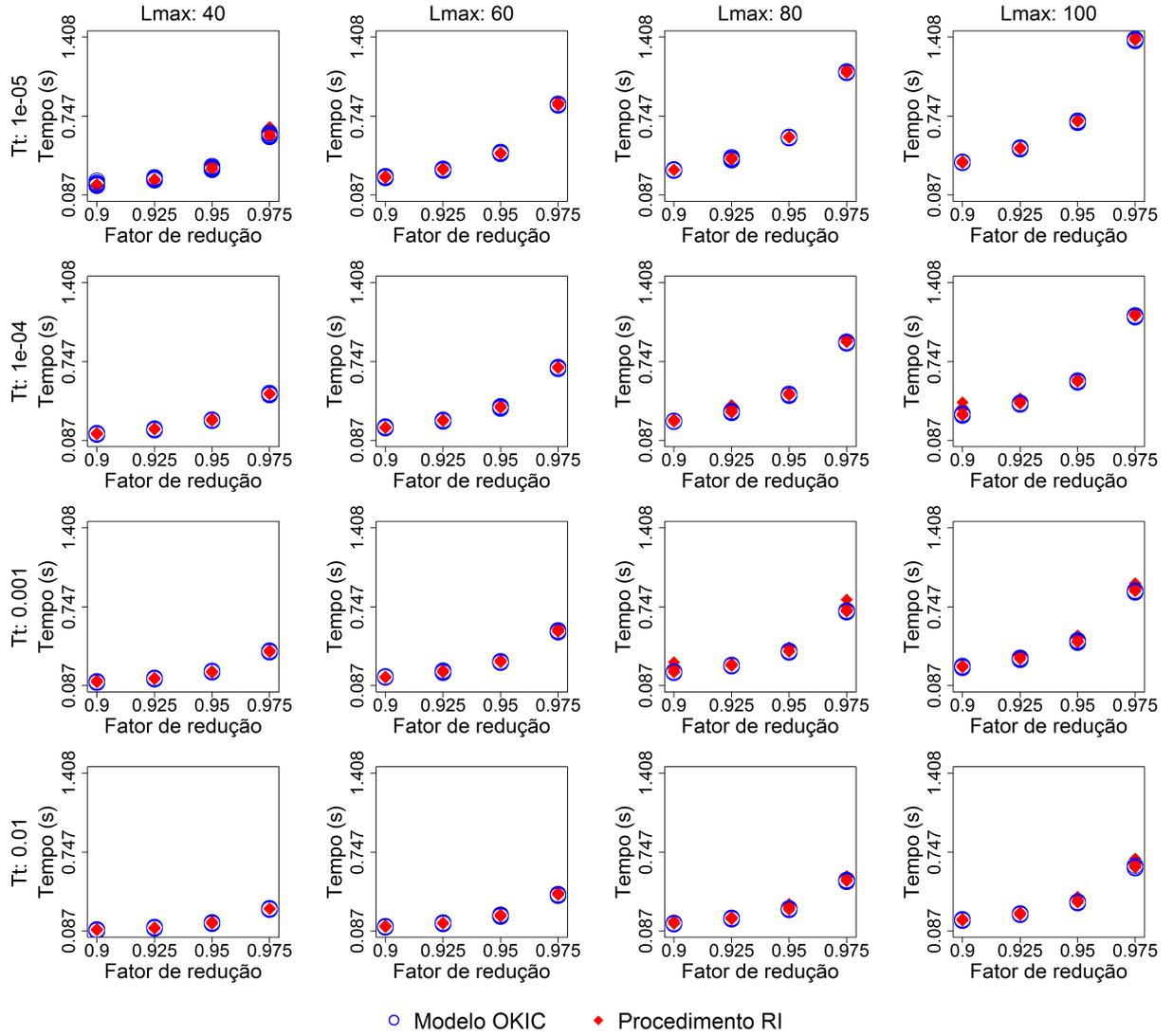


Figura 5.3: Tempos de execução dos experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, $\text{TotAmb} = 17$, $N = 100$ e 20 replicações.

$0,3103 \pm 0,0173$ segundos com $\text{TotAmb} = 17$ e $0,2825 \pm 0,0075$ segundos com $\text{TotAmb} = 28$.

Ao analisar, entretanto, com mais detalhe o tempo computacional necessário para executar os experimentos, é possível ser mais conservador em relação à qualidade da resposta. Ao utilizar $L_{\max} = 100$, $\gamma = 0,975$, $T_t = 10^{-5}$ e $T_0 = 2$ o intervalo de confiança da média do tempo computacional com o modelo KOIC será de apenas $1,38085 \pm 0,0089$ segundo para $\text{TotAmb} = 17$ e $1,36885 \pm 0,0057$ segundo para $\text{TotAmb} = 28$. Já com o procedimento RI este intervalo fica em $1,3919 \pm 0,00995$ segundo com $\text{TotAmb} = 17$ e $1,406 \pm 0,0382$ segundo com $\text{TotAmb} = 28$.

Como mostrado, o tempo de execução do algoritmo não é um limitador que mereça tanto destaque. Dessa forma, no processo de otimização para simulação todos os parâmetros serão utilizados em suas configurações máximas, o que possibilita uma certeza maior da obtenção de

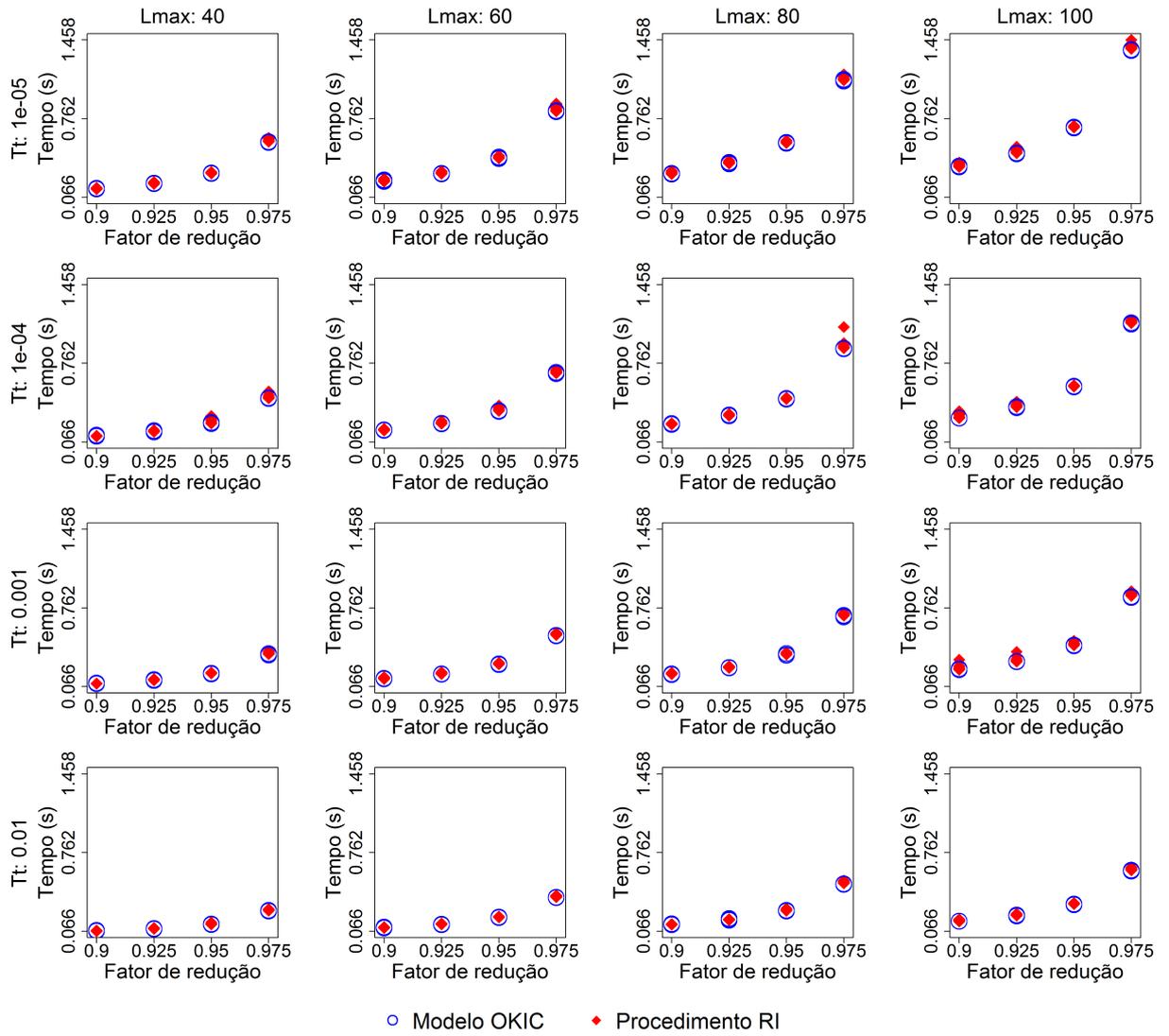


Figura 5.4: Tempos de execução dos experimentos realizados com o algoritmo SA, variando T_t , L_{\max} e γ , para resolver o Problema 4.1 com $T_0 = 2$, $\text{TotAmb} = 28$, $N = 100$ e 20 replicações.

pontos candidatos de melhor qualidade.

Capítulo 6

Otimização para Simulação do modelo do SAMU de Belo Horizonte

Neste capítulo é apresentada a aplicação dos métodos de Otimização para Simulação utilizados na minimização do tempo de resposta das ambulâncias básicas do SAMU de Belo Horizonte. Para isso, o estudo foi dividido em duas fases: a determinação do melhor método de Otimização para Simulação, levando em consideração os critérios definidos na Seção 6.1, e a determinação do tempo de resposta mínimo *versus* a quantidades de ambulâncias alocadas a Belo Horizonte.

Primeiramente, considerando a configuração de ambulâncias de 2010, o otimizador OptQuest do Anylogic foi usado para obter uma avaliação inicial do tempo de resposta mínimo das ambulâncias básicas. Em seguida, o modelo KOIC e o procedimento RI foram comparados entre si, de forma a determinar o método baseado em Krigagem mais adequado para a resolução do problema. Além disso, a influência do tamanho do conjunto de treinamento inicial também foi avaliada, com realização de testes com um conjunto de 50 pontos e outro com 100 pontos.

O método (KOIC, RI ou Optquest) escolhido na etapa anterior foi, então, utilizado na determinação do tempo mínimo de resposta para uma série de configurações de ambulâncias. Com os resultados, uma curva foi obtida representando o relacionamento entre estas duas informações: o intervalo de confiança do tempo mínimo de resposta e o número de ambulâncias alocadas à cidade.

É importante ressaltar, ainda, que todos os intervalos de confiança de média foram construídos para 95% de confiança e que cinco replicações foram realizadas na aplicação de cada método de otimização. Sendo assim, para facilitar a compreensão, a nomenclatura de cada solução foi

constituída de três termos:

$$\text{MODELO} - \text{TotAmb} - N \#i$$

onde o termo MODELO representa o uso do modelo KOIC ou o procedimento RI, TotAmb é número de ambulâncias consideradas no conjunto de treinamento, N denota o tamanho do conjunto de treinamento, podendo ter 50 ou 100 pontos, e $\#i$ é utilizado, quando necessário, para identificar o índice da replicação do processo de Otimização para Simulação.

6.1 Otimização da configuração de ambulâncias de 2010

Para efetuar a otimização da configuração de ambulâncias de 2010 de Belo Horizonte, primeiramente, empregou-se o otimizador OptQuest do Anylogic, sendo que o programa minimiza a média do tempo de resposta das ambulâncias básicas, com a restrição de utilizar um total de 17 unidades (número de unidades alocadas à Belo Horizonte em 2010, conforme a Tabela 4.1). Cinco replicações, cada uma com 200 iterações, foram realizadas e os resultados são apresentados na Figura 6.1 e Tabela 6.1, onde IC- e IC+ denotam, respectivamente, o limite inferior e superior do intervalo de confiança, e TR Médio é o tempo de resposta médio.

Tabela 6.1: Caparação das 5 replicações utilizando o otimizador OptQuest.

Replicação	IC- (min)	TR Médio (min)	IC+ (min)	Iterações	Duração (min)
1	19,896	20,412	20,928	200	245,8
2	20,157	20,466	20,775	200	265,43
3	20,221	20,511	20,801	200	290,87
4	20,071	20,458	20,845	200	265,25
5	19,960	20,428	20,896	200	276,9
Média:				200	268,85

Considerando apenas TR Médio, a melhor solução encontrada foi 20,412 min na primeira replicação do OptQuest, entretanto todas as soluções são estatisticamente equivalentes, pois os intervalos de confiança se sobrepõem. Observa-se que o tempo de execução varia razoavelmente, mesmo que o número de replicações se mantenha constante. Isto acontece, pois algumas soluções buscadas provocam a permanência de muitas entidades em espera no modelo de simulação, exigindo maior memória do computador e, conseqüentemente, aumentando o tempo de execução.

Para utilizar os modelos baseados em Krigagem, um Planejamento de Experimentos foi conduzido, onde dois conjuntos de treinamento foram amostrados: um com 50 pontos e outro com 100 pontos (mantendo os 50 pontos do conjunto anterior), ainda considerando a restrição de 17 ambulâncias. A definição do número de ambulâncias por base seguiu uma distribuição de

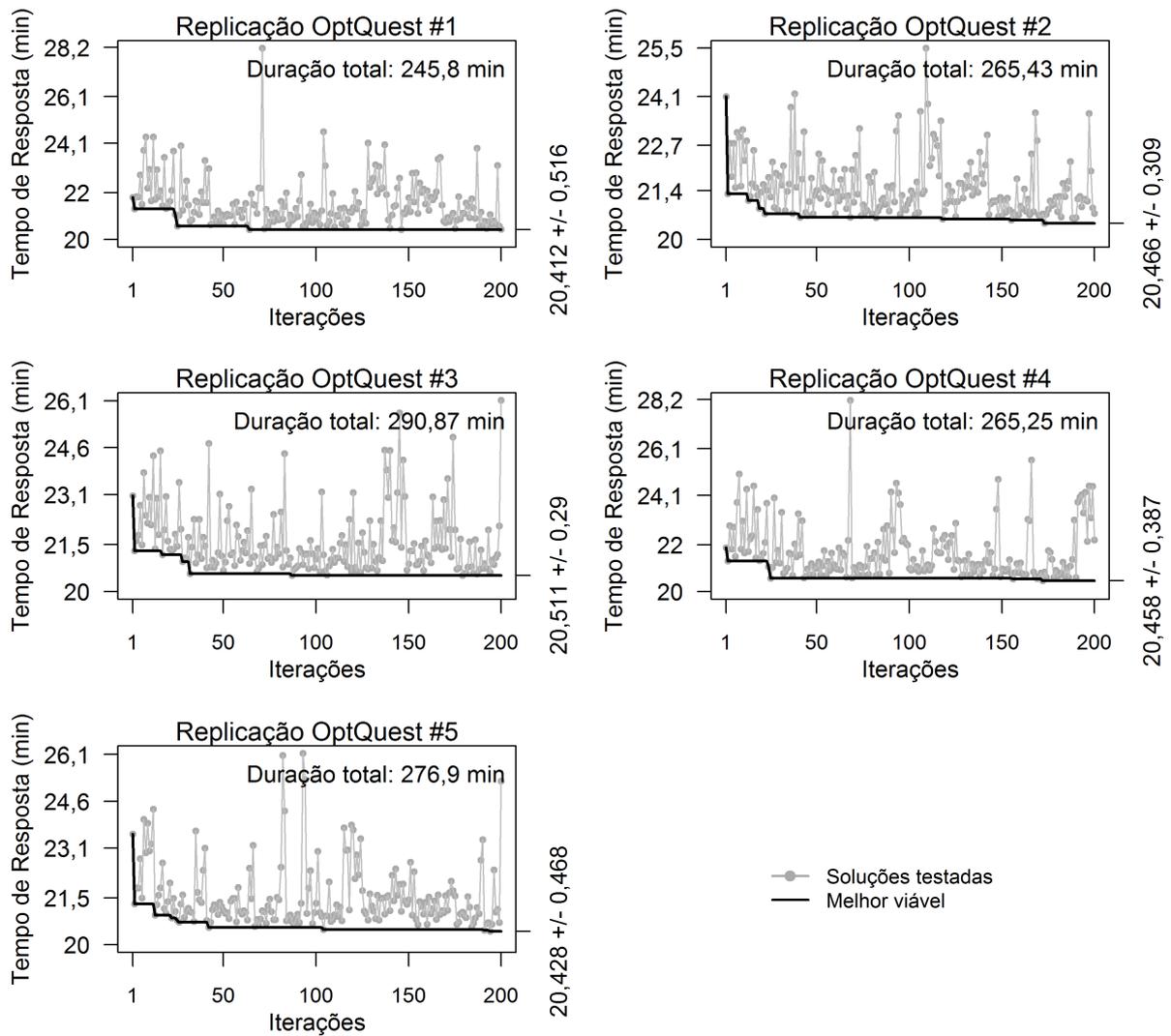


Figura 6.1: Cinco replicações efetuadas com o otimizador OptQuest para minimizar TR Médio para 17 ambulâncias.

probabilidade uniforme, ou seja, em cada base, existe a mesma probabilidade de serem sorteadas 0, 1, 2 ou 3 ambulâncias.

Para validar os conjuntos, uma análise de resíduos foi efetuada levando em consideração cada método. Dessa forma, a Figura 6.2, Figura 6.3, Figura 6.4 e Figura 6.5 plotam, respectivamente, as análises dos resíduos dos experimentos KOIC-17-50, KOIC-17-100, RI-17-50 e RI-17-100.

Todos os experimentos foram validados quanto à normalidade da distribuição dos resíduos (quadro a), não apresentaram correlação entre os resíduos e a estimação, além de se situarem dentro do intervalo $[-3,3]$ (quadro b) e possuem qualidade suficiente de estimação ao se organizarem uniformemente ao redor da linha em 45° (quadro c). No caso do experimento RI-17-100, houve uma exceção ao critério de análise, pois existe um ponto ultrapassando os limites $[-3,3]$. Entretanto, este problema não trouxe quaisquer dificuldades ao processo de otimização sequencial, já

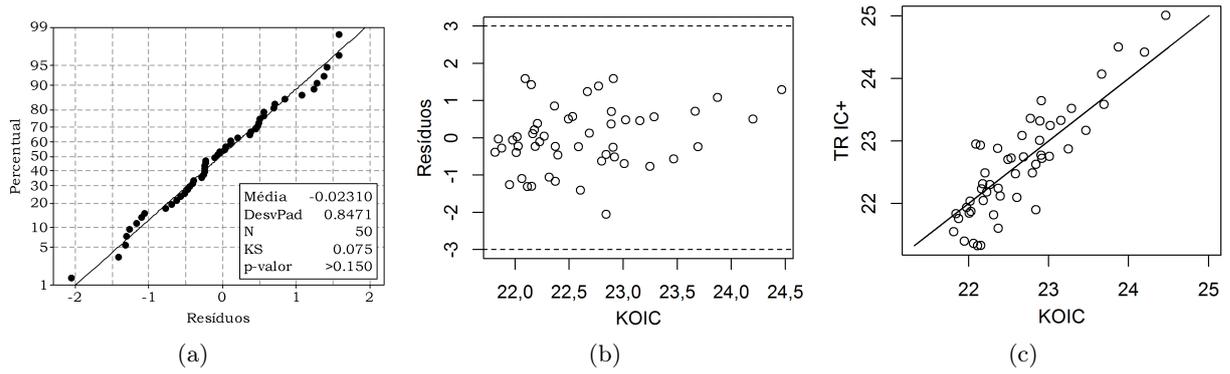


Figura 6.2: Análise de resíduos para o experimento KOIC-17-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

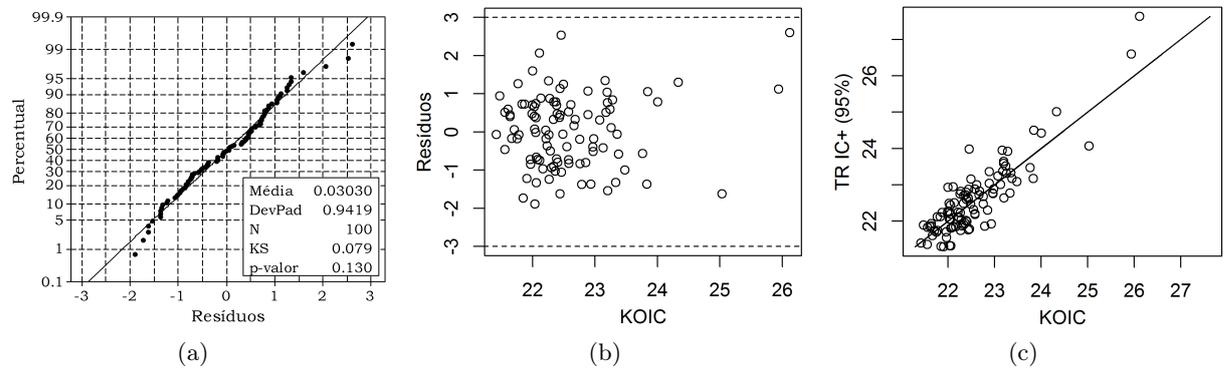


Figura 6.3: Análise de resíduos para o experimento KOIC-17-100: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

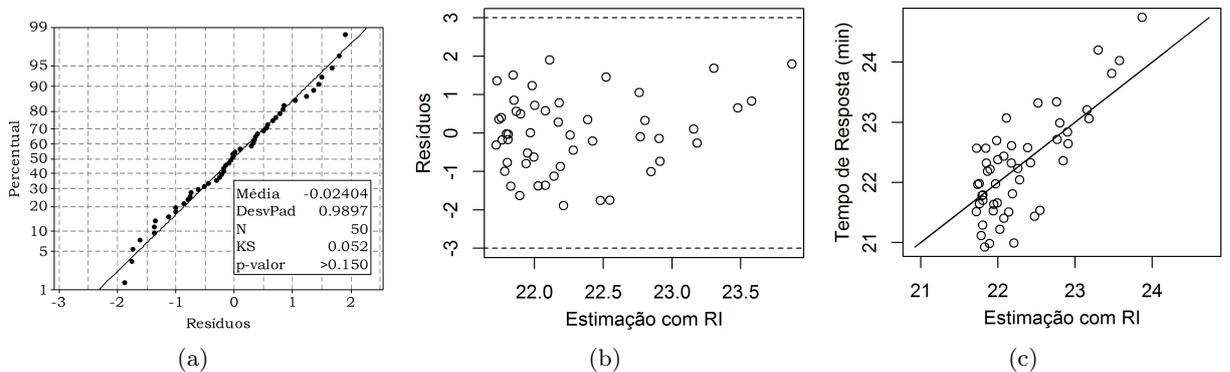


Figura 6.4: Análise de resíduos para o experimento RI-17-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

que ele se encontrava em uma região do espaço amostral pouco significativa.

Com os conjuntos de treinamento propriamente validados, os processos de otimização baseados em Krigagem foram finalmente realizados, sendo que o primeiro método testado foi o uso do modelo KOIC. A Figura 6.6 e a Figura 6.7 plotam 5 replicações do algoritmo EGO com o modelo em questão, utilizando, respectivamente, conjuntos de treinamento iniciais de 50 e 100

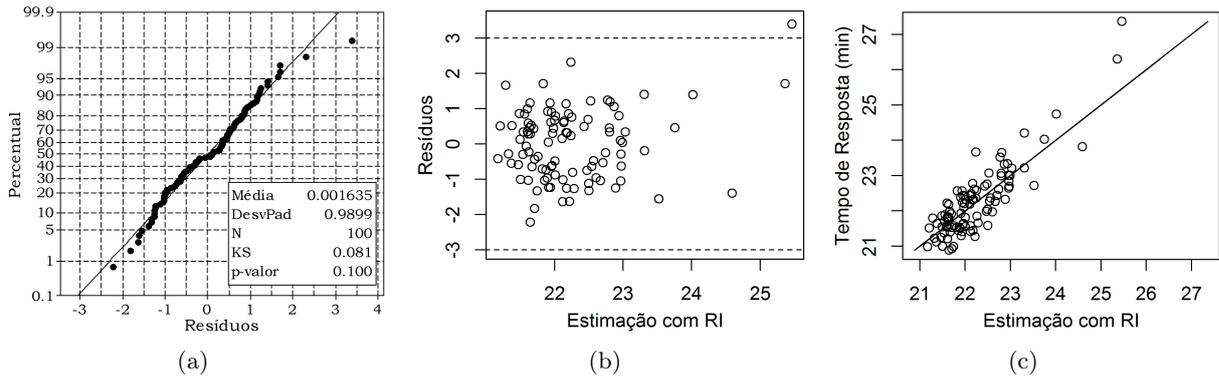


Figura 6.5: Análise de resíduos para o experimento RI-17-100: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

pontos. Por fim, a Tabela 6.2 e Tabela 6.3 resumem os resultados.

Tabela 6.2: Caparação das 5 replicações de solução de KOIC-17-50.

Replicação	IC- (min)	TR Médio (min)	IC+ (min)	Iterações	Duração (min)
1	20,013	20,254	20,496	134	132,61
2	20,251	20,401	20,551	171	191,24
3	20,080	20,334	20,587	160	175,16
4	19,909	20,270	20,631	190	220,34
5	19,996	20,262	20,528	125	123,19
		Média:		156	168,508

Tabela 6.3: Caparação das 5 replicações de solução de KOIC-17-100.

Replicação	IC- (min)	TR Médio (min)	IC+ (min)	Iterações	Duração (min)
1	20,192	20,372	20,552	162	139,53
2	20,126	20,390	20,653	187	172,08
3	20,148	20,398	20,648	201	192,15
4	20,065	20,336	20,606	165	138,03
5	20,135	20,392	20,648	199	198,17
		Média:		182,8	167,992

Cada quadro da Figura 6.6 e Figura 6.7 plota uma replicação do algoritmo EGO usando KOIC, onde TR IC+ é o valor simulado do lado superior do intervalo de confiança da média do tempo de resposta, TR IC+ Mínimo denota o mínimo corrente das iterações, KOIC IC é o intervalo de confiança da estimação com Krigagem e TR IC é o intervalo de confiança da média do tempo de resposta. Por fim, o valor mínimo inicial de TR IC+ é mostrado na iteração 50 ou iteração 100, dependendo do tamanho do conjunto de treinamento, e compreende o valor de mínimo obtido com a amostragem inicial.

O melhor valor encontrado para TR IC+ Mínimo foi 20,496 min na solução KOIC-17-50 #1, entretanto, do mesmo modo que com os resultados do OptQuest, a Tabela 6.2 e a Tabela 6.3

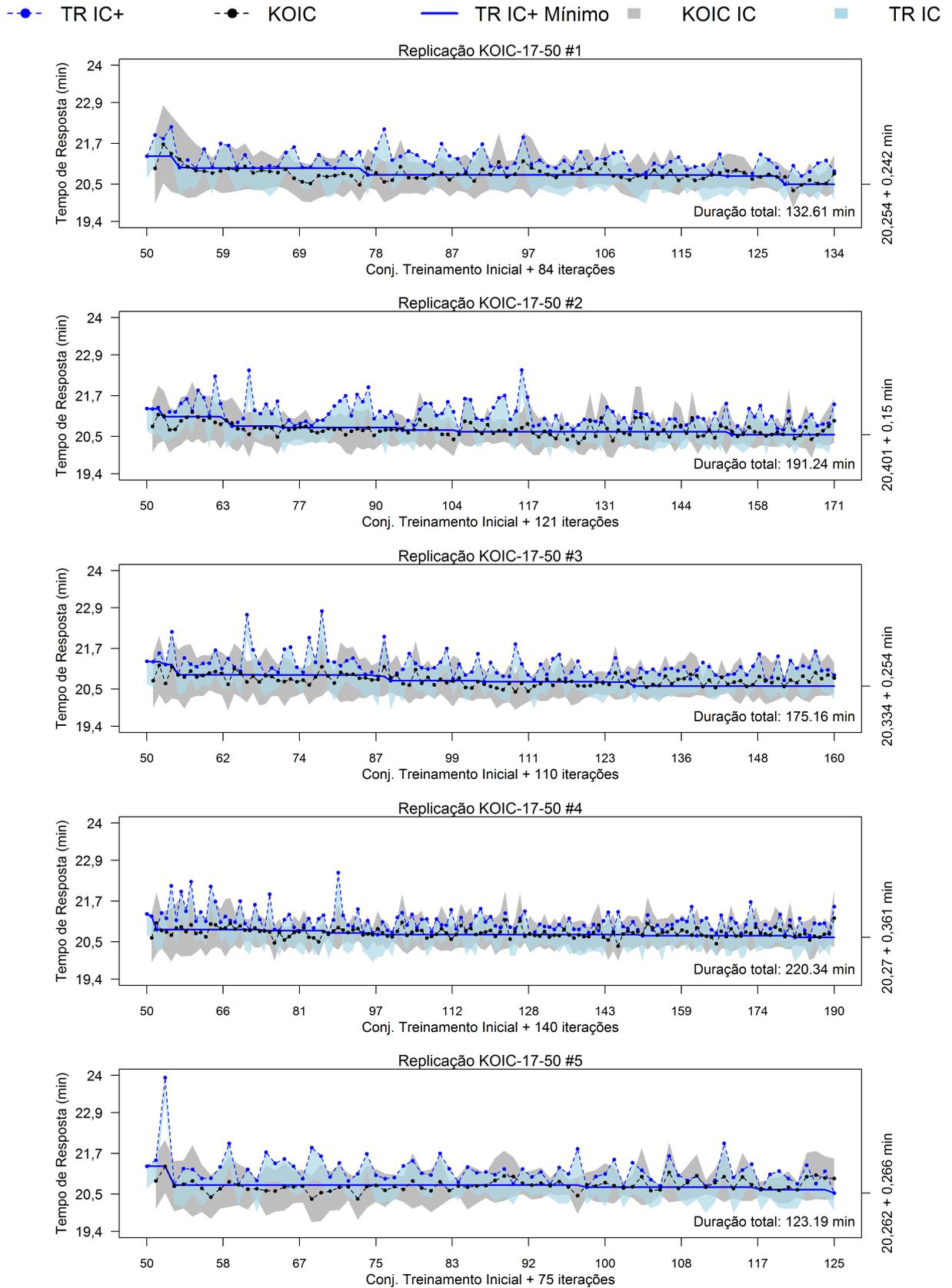


Figura 6.6: Cinco replicações realizadas para minimização do tempo de resposta resolver KOIC-17-50.

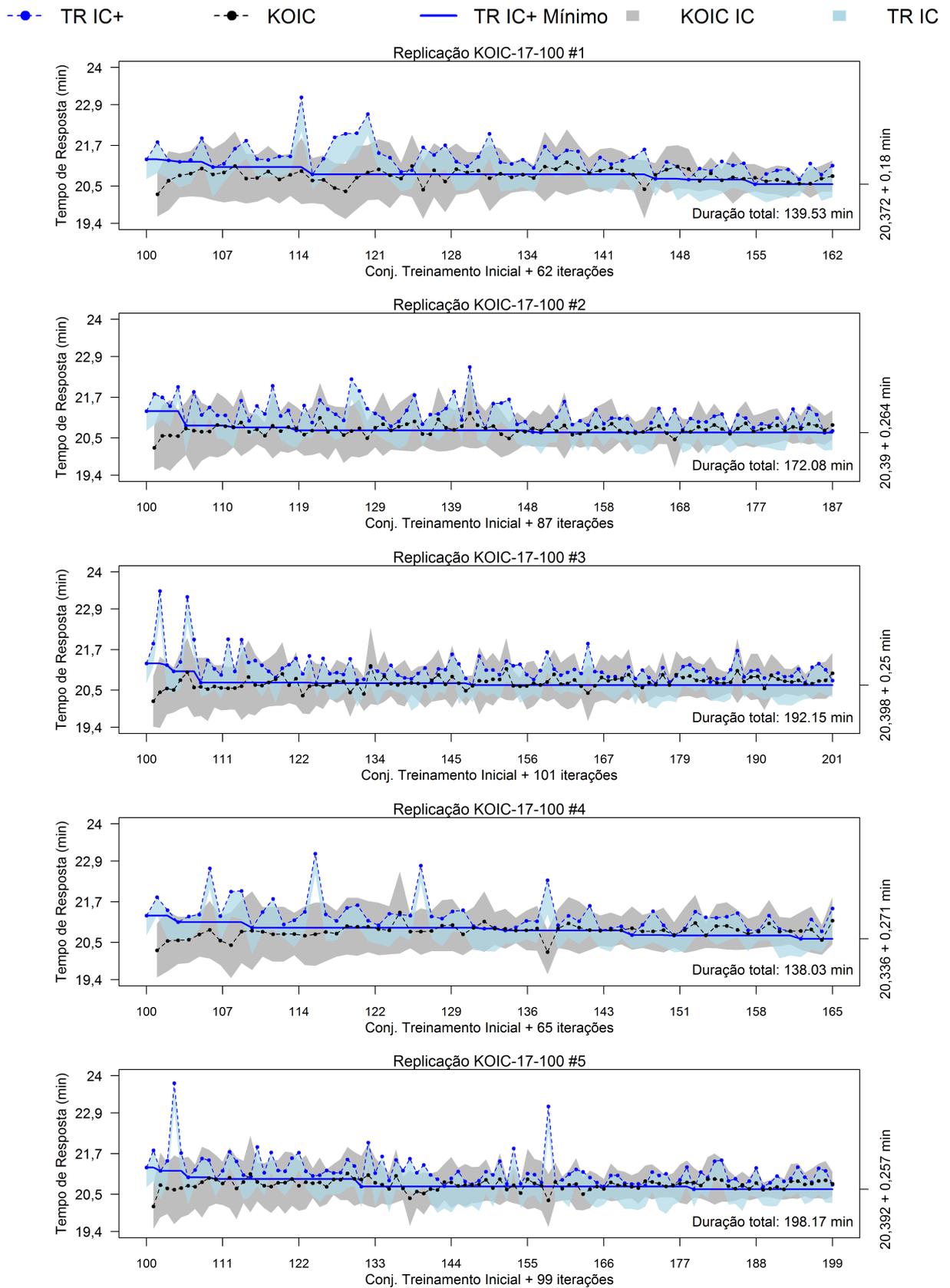


Figura 6.7: Cinco replicações realizadas para minimização do tempo de resposta para resolver KOIC-17-100.

mostram que os intervalos de confiança de todas as replicações se sobrepõem, seja considerando replicações com o mesmo conjunto de treinamento ou com conjuntos diferentes. Dessa forma, todas as soluções encontradas são equivalentes no uso do modelo KOIC, independentemente do tamanho do conjunto de treinamento inicial.

Repetindo a análise feita para o modelo KOIC, as mesmas etapas foram aplicadas para o procedimento RI, sendo que agora a Krigagem é usada para estimar a média do tempo de resposta e não o lado superior do seu intervalo de confiança. A Figura 6.8 e a Figura 6.9 plotam os resultados do processo de otimização com, respectivamente, 50 e 100 pontos no conjunto de treinamento inicial, e a Tabela 6.4 e Tabela 6.5 resumem os resultados obtidos.

Tabela 6.4: Caparação das 5 replicações de solução de RI-17-50.

Replicação	IC- (min)	TR Médio (min)	IC+ (min)	Iterações	Duração (min)
1	20,053	20,301	20,549	100	86,16
2	19,866	20,319	20,772	122	125,04
3	19,957	20,319	20,681	115	116,62
4	20,008	20,27	20,531	159	181,31
5	19,989	20,293	20,597	135	123,39
Média:				126,2	126,504

Tabela 6.5: Caparação das 5 replicações de solução de RI-17-100.

Replicação	IC- (min)	TR Médio (min)	IC+ (min)	Iterações	Duração (min)
1	20,178	20,386	20,593	154	121,62
2	19,940	20,242	20,543	152	119,53
3	19,858	20,288	20,717	164	137,33
4	19,903	20,312	20,720	164	136,25
5	19,825	20,259	20,694	154	125,45
Média				157,6	128,036

Na Figura 6.8 e Figura 6.9, cada quadro denota uma replicação do processo de otimização, onde TR Médio é o tempo de resposta médio obtido com o modelo de simulação, KO é a sua estimativa com o procedimento RI, TR Mínimo é o valor de mínimo encontrado a cada iteração, KO IC é o intervalo de confiança da estimativa com Krigagem e TR IC denota o intervalo de confiança da média do tempo de resposta.

Com o procedimento RI, o valor mínimo encontrado para a média foi de 20,242 minutos na solução RI-17-100 #2, entretanto novamente é necessário avaliar o intervalo de confiança da média de todas as soluções. Segundo a Tabela 6.4 e a Tabela 6.5, todas as soluções são equivalentes já que os intervalos de confiança se sobrepõem, seja considerando o mesmo conjunto de treinamento ou conjuntos diferentes.

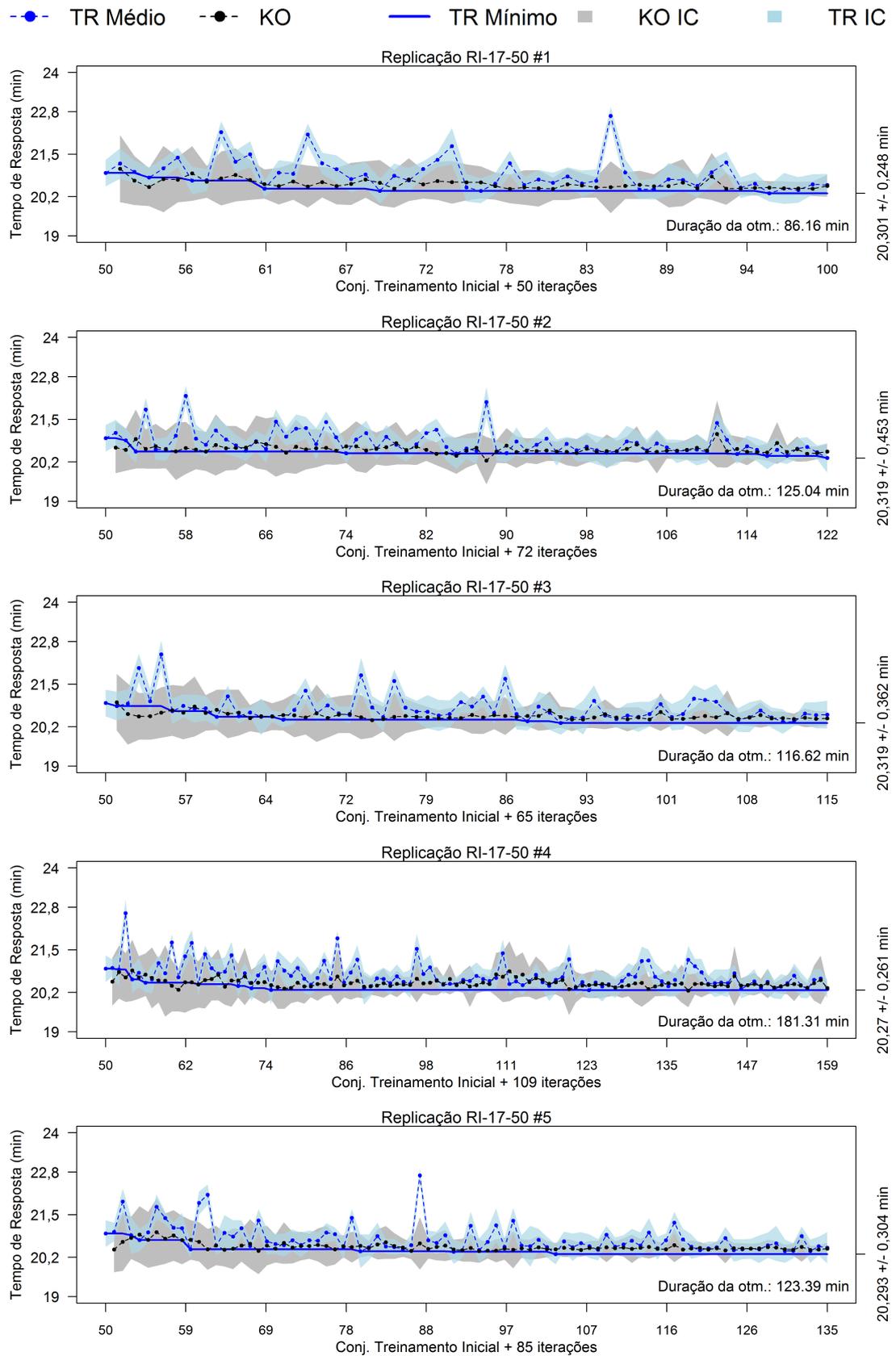


Figura 6.8: Cinco replicações realizadas para minimização do tempo de resposta para resolver RI-17-50.

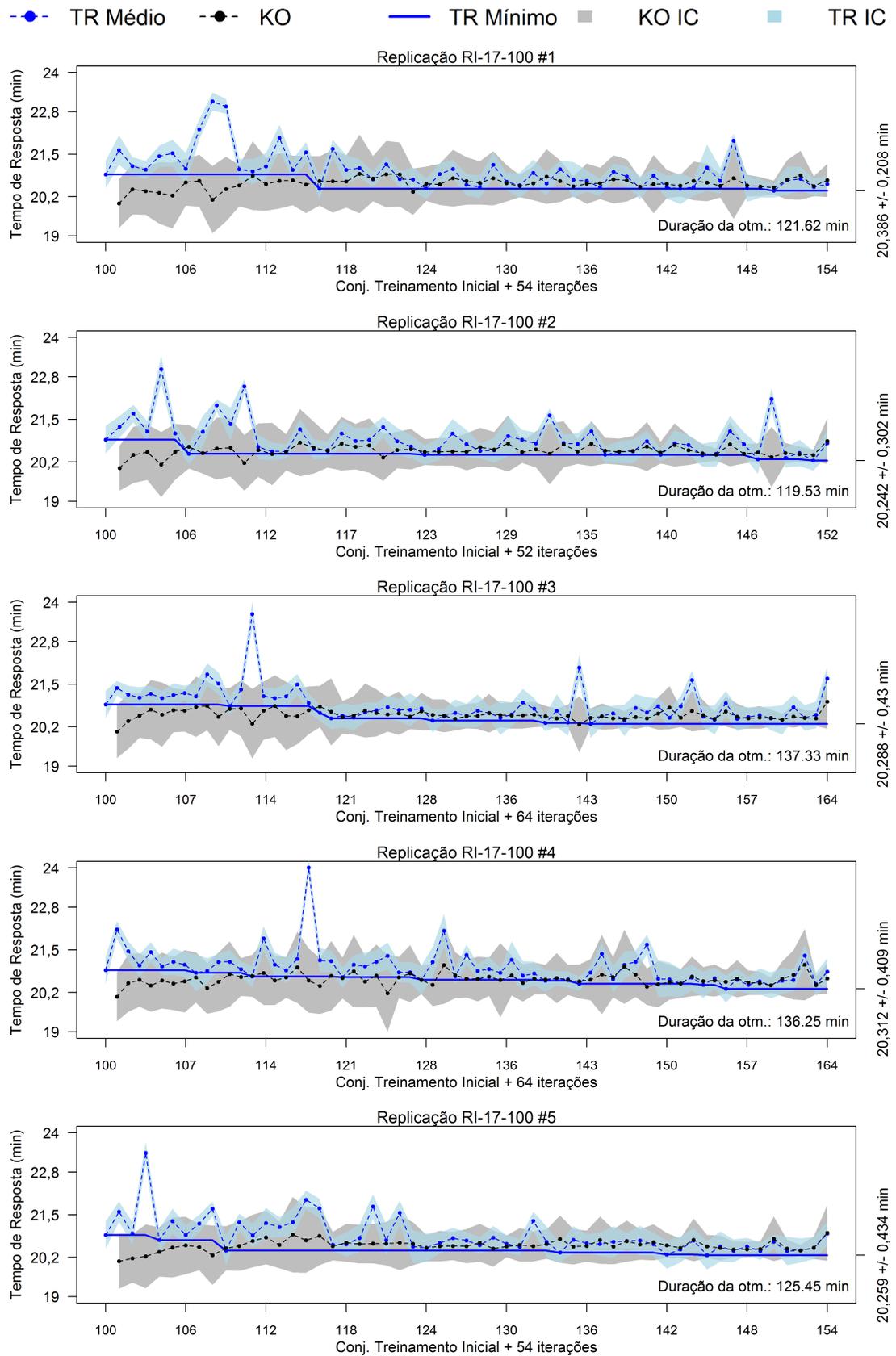


Figura 6.9: Cinco replicações realizadas para minimização do tempo de resposta para resolver RI-17-100.

Comparando os resultados dos dois métodos baseados em Krigagem e do OptQuest, vemos que os valores de mínimo encontrados também são estatisticamente equivalentes, assim não é possível determinar o melhor método, baseando-se apenas na qualidade das soluções encontradas. Entretanto, com o procedimento RI o tempo do processo de otimização foi consideravelmente menor do que com o modelo KOIC ou com o OptQuest.

A partir dos resultados, nota-se que o modelo KOIC tem mais dificuldade em estimar TR IC+ do que o procedimento RI tem em estimar TR Médio. Assim, é razoável supor que TR Médio possui um comportamento consideravelmente mais suave do que TR IC+, o que facilita sua estimação com o procedimento RI.

Além disso, pode-se concluir que TR Médio e TR IC+ possuem uma forma achatada próxima ao ótimo, o que justifica a redução de menos de 1 minuto no tempo de resposta com o processo de otimização. Nesse sentido, o maior ganho acontece com o OptQuest, devido o otimizador não realizar uma amostragem inicial.

Como em todos os métodos os resultados obtidos foram equivalentes, a melhor solução escolhida para o modelo KOIC foi aquela com menor TR IC+. Já para o procedimento RI e OptQuest, a solução escolhida foi aquela com menor TR Médio. A Tabela 6.6 apresenta os melhores pontos encontradas com cada método, onde as bases de ambulâncias foram ordenadas conforme a Tabela 4.1, e a Figura 6.10 plota a distribuição geográfica das soluções encontradas.

Tabela 6.6: Melhores soluções encontradas com o modelo KOIC, o procedimento RI e OptQuest, além da configuração utilizada em 2010 na cidade de Belo Horizonte (IC-, TR Médio e IC+ em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+
KOIC-17-50 #1	1 0 1 2 1 2 1 1 2 1 1 0 0 0 1 0 2 0 0 1	20,013	20,254	20,496
RI-17-100 #2	0 1 2 3 2 1 1 1 1 1 0 0 0 0 1 0 2 0 1 0	19,940	20,242	20,543
OptQuest #1	0 1 1 1 0 0 1 1 1 1 0 1 1 1 1 0 2 0 2 2	19,896	20,412	20,928
Configuração de 2010	0 1 1 0 0 1 1 1 1 1 2 1 1 1 1 1 1 1 1 0	20,744	21,200	21,656

Da Tabela 6.6, nota-se que as melhores soluções obtidas com KOIC e RI são estatisticamente melhores que a configuração de 2010, o que não é verdade para a melhor solução obtida com o OptQuest. Além disso, ao observar a Figura 6.10, nota-se que o processo de otimização eliminou ambulâncias das regiões Barreiro e Pampulha e as concentrou em regiões com mais ocorrências de chamadas, como Centro Sul e Venda Nova.

No caso da configuração de 2010, é provável que os gestores do sistema de saúde da cidade acreditassem que distribuir uniformemente as unidades pela cidade traria melhores benefícios, entretanto, com os resultados da Tabela 6.6 e Figura 6.10, mostrou-se que essa hipótese não é

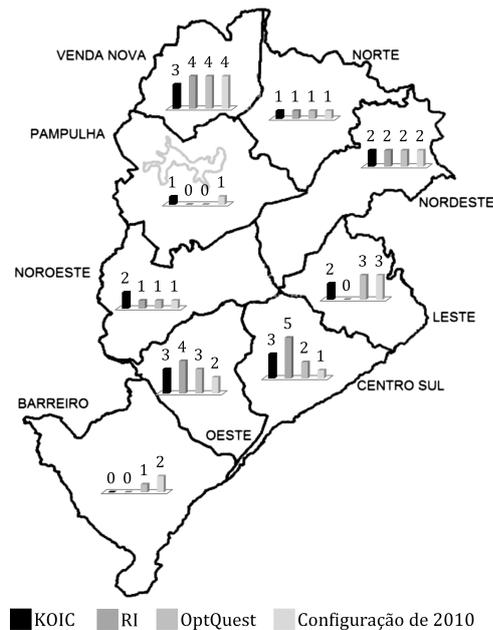


Figura 6.10: Distribuição geográfica das soluções encontradas com KOIC, RI, OptQuest, além da configuração de 2010 em Belo Horizonte.

verdadeira, pois, como as ambulâncias podem ser compartilhadas, nenhuma região fica de fato desabastecida por completo.

Por fim, considerando que todas as melhores soluções entre KOIC e RI são equivalentes e que o procedimento RI foi o método mais eficiente em função do tempo de processamento, ele foi utilizado nos processos de Otimização para Simulação para determinação da curva que faz a relação entre o tempo mínimo de resposta e o número de ambulâncias alocadas a Belo Horizonte.

6.2 Tempo de resposta mínimo *versus* número de ambulâncias alocadas

Para determinação da curva que fará a relação entre o tempo de resposta mínimo e o número de ambulâncias, novamente um Planejamento de Experimento foi conduzido. Neste caso, como o tamanho do conjunto de treinamento não influenciou os resultados dos experimentos realizados anteriormente, um conjunto específico de 50 pontos foi amostrado para 20, 22, 24, 26, 28 e 30 ambulâncias.

Do mesmo modo que na Seção 6.1, os resíduos foram validados quanto à normalidade, estarem dentro do intervalo $[-3,3]$ e por se organizarem uniformemente ao redor da linha em 45° . Os gráficos resultantes das análises, para facilitar a estruturação desta seção, são mostrados no Apêndice A, pela Figura A.1, Figura A.2, Figura A.3, Figura A.4, Figura A.5 e Figura A.6.

Novamente 5 replicações do processo de otimização, com o procedimento RI, foram realizadas para cada quantidade de ambulâncias alocadas a Belo Horizonte e os resultados são mostrados também no Apêndice A, na Tabela A.1, Tabela A.2, Tabela A.3, Tabela A.4, Tabela A.5, e Tabela A.6.

Agrupando as melhores soluções obtidas, com a Tabela 6.7 é possível concluir que o tempo de resposta se comporta como uma exponencial negativa em relação ao número de ambulâncias, ou seja, de acordo com que a cidade fica saturada de ambulâncias o tempo de resposta médio tem decréscimos cada vez menores.

Tabela 6.7: Melhores soluções encontradas para minimizar TR Médio com 20, 22, 24, 26, 28, 30 e 40 ambulâncias (IC-, TR Médio, IC+ e Duração em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-20-50 #2	1 1 1 2 0 0 1 1 2 1 2 0 0 0 1 0 3 1 2 1	15,140	15,262	15,383	193,792
RI-22-50 #5	0 1 2 2 1 1 2 1 2 1 2 1 0 0 2 1 2 1 0 0	13,594	13,706	13,817	202,639
RI-24-50 #4	0 1 2 2 0 1 2 1 2 1 2 1 0 0 2 1 2 2 1 1	12,817	12,857	12,897	142,672
RI-26-50 #1	1 1 2 3 2 0 3 1 2 2 2 0 0 0 1 1 2 2 1 0	12,290	12,340	12,390	125,943
RI-28-50 #5	0 1 3 3 1 1 2 1 2 3 2 1 0 0 1 2 2 2 1 0	11,955	12,022	12,089	101,942
RI-30-50 #2	1 2 2 3 1 0 3 1 2 2 2 0 1 1 1 2 3 2 1 0	11,696	11,721	11,747	123,075
RI-40-50 #5	1 3 3 3 1 0 3 1 3 3 3 1 1 2 2 2 3 3 1 1	11,196	11,226	11,257	87,710

Como esta curva é assintótica, é também interessante tentar descobrir o limite mínimo do tempo médio de resposta, dada as configurações adotadas neste trabalho para o modelo de simulação. Dessa forma, o mesmo procedimento anterior foi empregado para a minimização do tempo de resposta, mas agora considerando 40 ambulâncias.

Um Planejamento de Experimentos foi conduzido e os resíduos foram validados, sendo os gráficos da análise de resíduos apresentados no Apêndice A, pela Figura A.7, e o resultado da Otimização para Simulação é apresentado na Tabela A.7, sendo a melhor solução apresentada ao final da Tabela 6.7.

Finalmente, unindo todos os resultados obtidos com os processos de Otimização para Simulação, a Figura 6.11 plota o tempo de resposta mínimo em relação ao número de ambulâncias alocadas. Na figura, as melhores soluções foram identificadas ao lado ou acima dos pontos, juntamente com a configuração de 2010.

Como dito anteriormente, esta curva se comporta como uma exponencial negativa, dessa forma o esforço necessário, para diminuir 1 minuto no tempo médio de resposta, cresce exponencialmente, acarretando um alto nível de investimento para que pouca melhora seja propriamente verificada.

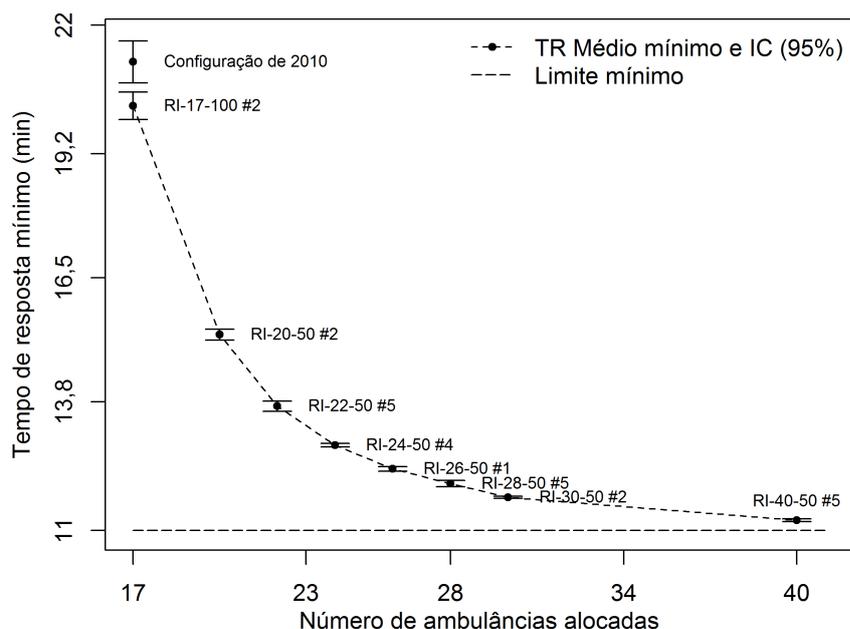


Figura 6.11: Tempo de resposta mínimo *versus* quantidade de ambulâncias alocadas a Belo Horizonte.

Por fim, pela Figura 6.11 é possível perceber que, com o número de ambulâncias tendendo ao infinito, o tempo de resposta mínimo se aproxima a 11 minutos. Esta é uma conclusão importante, pois, a partir de certo ponto, a inserção de ambulâncias no sistema do SAMU de Belo Horizonte deixa de ser uma solução viável, ou seja, o gargalo não é mais o número de ambulâncias.

Para reverter esse quadro, novas bases de ambulâncias poderiam ser instaladas na cidade, distribuindo melhor espacialmente as ambulâncias em espera. Além disso, as ambulâncias têm escalas de horário fixas nas bases, por isso seria interessante alterá-las para um configuração dinâmica, de modo que um número específico de ambulâncias seja alocado não só por base, mas também por turno de trabalho (matutino, vespertino e noturno).

Capítulo 7

Conclusões e Trabalhos Futuros

O objetivo dessa dissertação de mestrado foi explorar metamodelagem direcionada à Otimização para Simulação e ampliar o estado-da-arte sobre a aplicação do método em modelos estocásticos de grande complexidade, onde há necessidade de avaliação de um número considerado de variáveis de entrada. Nesse sentido, todos os objetivos propostos no Capítulo 1 foram alcançados com êxito.

A contribuição principal da pesquisa foi a análise completa de um estudo de Otimização para Simulação sobre um modelo estocástico da área da Saúde, em que as etapas de determinação das variáveis de entrada e saída, especificação e validação de conjuntos de treinamento iniciais, construção de uma heurística de otimização e análise do processo de otimização foram realizadas com o devido rigor teórico.

Para facilitar a compreensão do trabalho, praticamente todos os métodos apresentados foram melhor esclarecidos a partir de exemplos numerosos e com o auxílio de figuras e gráficos, quando adequados, mesmo que apenas um número restrito de variáveis tenha sido empregado.

Durante o desenvolvimento, foram levantados os principais métodos de Otimização para Simulação baseados em metamodelagem, sendo que Krigagem foi escolhida por apresentar as características mais apropriadas para a realização do trabalho, com destaque ao seu parâmetro variância de estimação, que aumenta ou diminui conforme a incerteza do espaço em análise.

Após, os critérios de desempenho mais adequados foram escolhidos, levando em consideração o caráter estocástico do problema. Uma heurística *Simulated Annealing* foi especificada como um otimizador global, sendo utilizada para maximizar EI nas duas variações, KOIC e RI, apresentadas no Capítulo 3. Por fim, o processo de otimização foi executado, tendo como resultado final uma curva que fez o relacionamento entre o tempo de resposta mínimo e o número de

ambulâncias alocadas a Belo Horizonte.

A aleatoriedade do modelo de simulação foi tratada com o procedimento RI e com a proposição do modelo KOIC, que estima o lado superior do intervalo de confiança da variável de resposta. A partir dos resultados obtidos, foi verificado que o modelo KOIC teve certa dificuldade na estimação, o que parece ter ocorrido devido a influência da variância do tempo de resposta. O procedimento RI foi escolhido como o melhor método dentre KOIC e OptQuest, por ter sido o mais eficiente em otimizar o tempo de resposta na configuração de 17 ambulâncias básicas, com menores tempos de execução.

Mesmo não tendo sido escolhido como o melhor método, o modelo KOIC ainda não pode ser descartado como um bom tratamento a modelos de simulação estocásticos, pois os experimentos realizados aqui foram voltados apenas para uma situação específica. Para chegar a uma conclusão mais adequada, será preciso testar o modelo em funções diferentes, como Branin-Hoo, Goldstein-Price, Rosenbrock4, dentre outras, apresentadas em Picheny, Wagner e Ginsbourger (2013), e para diversas condições de ruído aleatório.

Na literatura, um dos principais problemas se tratando de Krigagem se dá pelo aumento do tamanho do conjunto de treinamento no processo de Otimização para Simulação e, conseqüentemente, pelo aumento no tempo computacional necessário para realizar as operações matemáticas inerentes ao metamodelo. Assim, como uma forma de evitar este problema, os conjuntos de treinamento iniciais não foram dimensionados com muitos pontos, entretanto, como mostrado no Capítulo 6 e corroborando Picheny et al. (2013), o tamanho do conjunto de treinamento inicial não teve impacto significativo nos resultados.

Em relação a trabalhos futuros, duas vertentes principais podem ser percorridas: a utilização de múltiplos metamodelos no processo de otimização, como sugerido por Viana, Haftka e Watson (2010), e a estimação, com Krigagem, da região Pareto-ótima na otimização de funções multiobjetivo, como trabalhado por Binois, Ginsbourger e Roustant (2015).

No primeiro caso, o uso de múltiplos metamodelos permite a utilização de processamento paralelo do computador, o que proporciona um ganho considerado de eficiência no processo, já que agora dois ou mais pontos são inseridos ao conjunto de treinamento em cada iteração. Dentre os diversos tipos de metamodelos, destaca-se para este tipo de aplicação a Krigagem e redes neurais, de forma que, para empregar EGO adequadamente, a variância de estimação obtida com o modelo de Krigagem é importada pelo modelo de rede neural e, assim, um ponto candidato é obtido normalmente, de acordo com o Capítulo 3.

A estimação da região Pareto-ótima de funções multiobjetivo a partir de Krigagem é uma das linhas de estudo mais recentes em metamodelagem. No caso do modelo do SAMU de Belo Horizonte, por exemplo, uma análise multiobjetivo seria interessante para obter a região Pareto-ótima do tempo médio e máximo de resposta. Neste caso, a minimização de apenas o tempo médio com 17 ambulâncias, como mostrado nos resultados do Capítulo 6, concentrou unidades em regiões com alta densidade de chamadas, entretanto, como duas regiões ficaram desguarnecidas, é provável que o tempo máximo de resposta tenha aumentado. Assim, a determinação da região Pareto-ótima da combinação destas duas variáveis de resposta proporcionaria um melhor entendimento do sistema.

Além das possibilidades de trabalho futuro citadas anteriormente, esta pesquisa abriu espaço para alternativas interessantes de execução. Uma possibilidade seria trocar o método de obtenção dos conjuntos de treinamento para Krigagem, alterando a distribuição uniforme usada no trabalho por uma distribuição multinomial, que consideraria a distribuição de probabilidade de geração de chamadas nas 9 regiões administrativas de Belo Horizonte. Outra possibilidade seria alterar o procedimento de construção e inversão da matriz de covariância \mathbf{C} por um método mais sofisticado, mantendo apenas um número fixo de pontos no conjunto de treinamento e avaliando se o ponto candidato a ótimo tornaria a matriz \mathbf{C} singular.

A aplicação de Otimização para Simulação baseada em metamodelagem sobre um modelo voltado à Saúde deu ao trabalho um apelo importante, já que as publicações relacionadas são voltadas a resolução de problemas de Engenharia e, principalmente, por ser um problema real enfrentado em qualquer metrópole do mundo. Assim, a análise deste problema e dos resultados desta pesquisa também podem ajudar gestores públicos de Saúde a planejarem de forma mais adequada a capacidade dos sistemas de atendimento.

Referências Bibliográficas

- AHUJA, S. K.; FERREIRA, G. M.; MOREIRA, A. R. Application of plackett-burman design and response surface methodology to achieve exponential growth for aggregated shipworm bacterium. *Biotechnology and Bioengineering*, 2004. Wiley Subscription Services, Inc., A Wiley Company, v. 85, n. 6, p. 666–675, 2004. ISSN 1097-0290. Disponível em: <<http://dx.doi.org/10.1002/bit.10880>>.
- ANKENMAN, B.; NELSON, B. L.; STAUM, J. Stochastic kriging for simulation metamodeling. *Operations research*, 2010. INFORMS, v. 58, n. 2, p. 371–382, 2010.
- ARORA, J.; HUANG, M.; HSIEH, C. Methods for optimization of nonlinear problems with discrete variables: a review. *Structural Optimization*, 1994. Springer, v. 8, n. 2-3, p. 69–85, 1994.
- BADHURSHAH, R.; SAMAD, A. Multiple surrogate based optimization of a bidirectional impulse turbine for wave energy conversion. *Renewable Energy*, 2015. Elsevier, v. 74, p. 749–760, 2015.
- BARTON, R. R. Metamodels for simulation input-output relations. In: ACM. *Proceedings of the 24th conference on Winter simulation*. [S.l.], 1992. p. 289–299.
- BARTON, R. R. Simulation metamodels. In: IEEE COMPUTER SOCIETY PRESS. *Proceedings of the 30th conference on Winter simulation*. [S.l.], 1998. p. 167–176.
- BARTON, R. R. Simulation optimization using metamodels. In: WINTER SIMULATION CONFERENCE. *Winter Simulation Conference*. [S.l.], 2009. p. 230–238.
- BASUDHAR, A. et al. Constrained efficient global optimization with support vector machines. *Structural and Multidisciplinary Optimization*, 2012. Springer, v. 46, n. 2, p. 201–221, 2012.
- BEERS, W. C. V.; KLEIJNEN, J. P. Kriging interpolation in simulation: a survey. In: IEEE. *Simulation Conference, 2004. Proceedings of the 2004 Winter*. [S.l.], 2004. v. 1.
- BILES, W. E. et al. Kriging metamodeling in constrained simulation optimization: an explorative study. In: IEEE PRESS. *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*. [S.l.], 2007. p. 355–362.
- BINOIS, M.; GINSBOURGER, D.; ROUSTANT, O. Quantifying uncertainty on pareto fronts with gaussian process conditional simulations. *European Journal of Operational Research*, 2015. Elsevier, v. 243, n. 2, p. 386–394, 2015.
- BOX, G. E. The exploration and exploitation of response surfaces: some general considerations and examples. *Biometrics*, 1954. JSTOR, v. 10, n. 1, p. 16–60, 1954.
- BOX, G. E.; DRAPER, N. R. *Empirical model-building and response surfaces*. [S.l.]: John Wiley & Sons, 1987.
- BOX, G. E. et al. *Statistics for experimenters*. New York, USA: John Wiley & Sons, 1978.

- CHEN, X.; ANKENMAN, B. E.; NELSON, B. L. The effects of common random numbers on stochastic kriging metamodels. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 2012. ACM, v. 22, n. 2, p. 7, 2012.
- CHILES, J.-P.; DELFINER, P. *Geostatistics: modeling spatial uncertainty*. New York, USA: John Wiley & Sons, 2012.
- CIMELLARO, G. P.; REINHORN, A. M.; BRUNEAU, M. Performance-based metamodel for healthcare facilities. *Earthquake Engineering & Structural Dynamics*, 2011. Wiley Online Library, v. 40, n. 11, p. 1197–1217, 2011.
- COUCKUYT, I. et al. Blind kriging: Implementation and performance analysis. *Advances in Engineering Software*, 2012. Elsevier, v. 49, p. 1–13, 2012.
- COX, D. D.; JOHN, S. Sdo: A statistical method for global optimization. *Multidisciplinary design optimization: state of the art*, 1997. SIAM, Philadelphia, p. 315–329, 1997.
- CRARY, S. B. Design of computer experiments for metamodel generation. *Analog Integrated Circuits and Signal Processing*, 2002. Springer, v. 32, n. 1, p. 7–16, 2002.
- DELHOMME, J. Kriging in the hydrosociences. *Advances in water resources*, 1978. Elsevier, v. 1, n. 5, p. 251–266, 1978.
- DELLINO, G.; KLEIJNEN, J. P.; MELONI, C. Robust simulation-optimization using metamodels. In: WINTER SIMULATION CONFERENCE. *Winter Simulation Conference*. [S.l.], 2009. p. 540–550.
- DENG, H. et al. Bayesian metamodeling for computer experiments using the gaussian kriging models. *Quality and Reliability Engineering International*, 2012. Wiley Online Library, v. 28, n. 4, p. 455–466, 2012.
- FORRESTER, A. I. J.; KEANE, A. J.; BRESSLOFF, N. W. Design and analysis of "noisy" computer experiments. *AIAA journal*, 2006. v. 44, n. 10, p. 2331–2339, 2006.
- GANO, S. E. et al. Update strategies for kriging models used in variable fidelity optimization. *Structural and Multidisciplinary Optimization*, 2006. Springer, v. 32, n. 4, p. 287–298, 2006.
- GARITSELOV, O.; MOHANTY, S. P.; KOUGIANOS, E. Fast-accurate non-polynomial meta-modeling for nano-cmos pll design optimization. In: IEEE. *VLSI Design (VLSID), 2012 25th International Conference on*. [S.l.], 2012. p. 316–321.
- GOFFE, W. L.; FERRIER, G. D.; ROGERS, J. Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, 1994. v. 60, n. 1–2, p. 65 – 99, 1994. ISSN 0304-4076.
- HENGL, T.; HEUVELINK, G.; STEIN, A. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 2004. Elsevier, v. 120, n. 1, p. 75–93, 2004.
- HUANG, D. et al. Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of global optimization*, 2006. Springer, v. 34, n. 3, p. 441–466, 2006.
- HUSSAIN, M. S. et al. A surrogate model for simulation–optimization of aquifer systems subjected to seawater intrusion. *Journal of Hydrology*, 2015. Elsevier, v. 523, p. 542–554, 2015.
- JIN, R.; DU, X.; CHEN, W. The use of metamodeling techniques for optimization under uncertainty. *Structural and Multidisciplinary Optimization*, 2003. Springer, v. 25, n. 2, p. 99–116, 2003.

- JONES, D. R.; SCHONLAU, M.; WELCH, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 1998. Springer, v. 13, n. 4, p. 455–492, 1998.
- JOSEPH, V. R.; HUNG, Y.; SUDJIANTO, A. Blind kriging: A new method for developing metamodels. *Journal of mechanical design*, 2008. American Society of Mechanical Engineers, v. 130, n. 3, p. 031102, 2008.
- KHURI, A. I.; CORNELL, J. A. *Response surfaces: designs and analyses*. New York, USA: CRC press, 1996.
- KLEIJNEN, J. P. An overview of the design and analysis of simulation experiments for sensitivity analysis. *European Journal of Operational Research*, 2005. Elsevier, v. 164, n. 2, p. 287–300, 2005.
- KLEIJNEN, J. P. Kriging metamodeling in simulation: A review. *European Journal of Operational Research*, 2009. Elsevier, v. 192, n. 3, p. 707–716, 2009.
- KLEIJNEN, J. P.; BEERS, W. van. Robustness of kriging when interpolating in random simulation with heterogeneous variances: some experiments. *European Journal of Operational Research*, 2005. Elsevier, v. 165, n. 3, p. 826–834, 2005.
- KLEIJNEN, J. P.; BEERS, W. van; NIEUWENHUYSE, I. V. Expected improvement in efficient global optimization through bootstrapped kriging. *Journal of global optimization*, 2012. Springer, v. 54, n. 1, p. 59–73, 2012.
- KLEIJNEN, J. P.; MEHDAD, E. Multivariate versus univariate kriging metamodels for multi-response simulation models. *European Journal of Operational Research*, 2014. Elsevier, v. 236, n. 2, p. 573–582, 2014.
- KLEIJNEN, J. P.; SARGENT, R. G. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research*, 2000. Elsevier, v. 120, n. 1, p. 14–29, 2000.
- KLEIJNEN, J. P. C. *Statistical Tools for Simulation Practitioners*. New York, USA: Marcel Dekker, Inc., 1987.
- KOCH, P. et al. Efficient multi-criteria optimization on noisy machine learning problems. *Applied Soft Computing*, 2015. Elsevier, 2015.
- KUMAR, B.; RAO, A. R. Metamodeling approach to predict friction factor of alluvial channel. *Computers and electronics in agriculture*, 2010. Elsevier, v. 70, n. 1, p. 144–150, 2010.
- LAW, A. M. *Simulation modeling and analysis*. New York, USA: McGraw-Hill, 2007.
- LI, R.; SUDJIANTO, A. Analysis of computer experiments using penalized likelihood in gaussian kriging models. *Technometrics*, 2005. v. 47, n. 2, 2005.
- LIANG, J.; BUONGIORNO, J.; MONSERUD, R. A. Bootstrap simulation and response surface optimization of management regimes for douglas-fir/western hemlock stands. *Forest science*, 2006. Society of American Foresters, v. 52, n. 5, p. 579–594, 2006.
- LOPHAVEN, S. N.; NIELSEN, H. B.; SØNDERGAARD, J. A matlab kriging toolbox. *Technical University of Denmark, Kongens Lyngby, Technical Report No. IMM-TR-2002-12*, 2002. 2002.
- LUENBERGER, D. G. *Linear and nonlinear programming*. 2. ed. [S.l.]: Springer Science & Business Media, 2003.

- MARTIN, J. D. Robust kriging models. In: *14th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*. [S.l.: s.n.], 2010.
- MARTIN, J. D.; SIMPSON, T. W. On the use of kriging models to approximate deterministic computer models. In: AMERICAN SOCIETY OF MECHANICAL ENGINEERS. *ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. [S.l.], 2004. p. 481–492.
- MATHERON, G. Principles of geostatistics. *Economic geology*, 1963. Society of Economic Geologists, v. 58, n. 8, p. 1246–1266, 1963.
- MCKAY, M. D.; BECKMAN, R. J.; CONOVER, W. J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 1979. Taylor & Francis, v. 21, n. 2, p. 239–245, 1979.
- MONTGOMERY, D. C. *Applied Statistics and Probability for Engineers 6th edition*. New York, USA: John Wiley & Sons, 2013.
- MYERS, W. R.; MONTGOMERY, D. C. Response surface methodology. *Encycl Biopharm Stat*, 2003. Taylor & Francis, v. 1, p. 858–869, 2003.
- PICHENY, V.; GINSBOURGER, D. Noisy kriging-based optimization methods: A unified implementation within the diceoptim package. *Computational Statistics & Data Analysis*, 2014. Elsevier, v. 71, p. 1035–1053, 2014.
- PICHENY, V. et al. Quantile-based optimization of noisy computer experiments with tunable precision. *Technometrics*, 2013. Taylor & Francis Group, v. 55, n. 1, p. 2–13, 2013.
- PICHENY, V.; WAGNER, T.; GINSBOURGER, D. A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 2013. Springer, v. 48, n. 3, p. 607–626, 2013.
- PINTO, L.; SILVA, P.; YOUNG, T. A generic method to develop simulation models for ambulance systems. *Simulation Modelling Practice and Theory*, 2015. v. 51, n. 0, p. 170 – 183, 2015. ISSN 1569-190X.
- ROUSTANT, O.; GINSBOURGER, D.; DEVILLE, Y. *DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization*. 2012. Acessado em: 02/09/2014. Disponível em: <<http://hal.archives-ouvertes.fr/hal-00495766>>.
- SASENA, M. J. *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*. Tese (Doutorado) — University of Michigan, 2002.
- SASENA, M. J.; PAPALAMBROS, P.; GOOVAERTS, P. Exploration of metamodeling sampling criteria for constrained global optimization. *Engineering optimization*, 2002. Taylor & Francis, v. 34, n. 3, p. 263–278, 2002.
- SILVA, P. M. S.; PINTO, L. R. Emergency medical systems analysis by simulation and optimization. In: WINTER SIMULATION CONFERENCE. *Proceedings of the Winter Simulation Conference*. [S.l.], 2010. p. 2422–2432.
- SIMPSON, T. W. et al. Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA journal*, 2001. v. 39, n. 12, p. 2233–2241, 2001.
- SIMPSON, T. W. et al. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with computers*, 2001. Springer, v. 17, n. 2, p. 129–150, 2001.

- STAUM, J. Better simulation metamodeling: The why, what, and how of stochastic kriging. In: IEEE. *Simulation Conference (WSC), Proceedings of the 2009 Winter*. [S.l.], 2009. p. 119–133.
- VIANA, F. A.; HAFTKA, R. T.; WATSON, L. T. Why not run the efficient global optimization algorithm with multiple surrogates. In: *Proceedings of the 51th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, AIAA, Orlando, FL, USA. AIAA–2010–3090*. [S.l.: s.n.], 2010.
- WANG, G. G.; SHAN, S. Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design*, 2007. American Society of Mechanical Engineers, v. 129, n. 4, p. 370–380, 2007.
- WRIGHT, S. J.; NOCEDAL, J. *Numerical optimization*. [S.l.]: Springer New York, 1999.
- XIONG, F. et al. Optimizing latin hypercube design for sequential sampling of computer experiments. *Engineering Optimization*, 2009. Taylor & Francis, v. 41, n. 8, p. 793–810, 2009.
- YIN, J.; NG, S.; NG, K. Kriging metamodel with modified nugget-effect: The heteroscedastic variance case. *Computers & Industrial Engineering*, 2011. Elsevier, v. 61, n. 3, p. 760–777, 2011.
- ZHANG, Y. et al. Modeling and multi-objective optimization of double suction centrifugal pump based on kriging meta-models. In: *Advances in Global Optimization*. [S.l.]: Springer, 2015. p. 251–261.
- ZHAO, L.; CHOI, K.; LEE, I. Metamodeling method using dynamic kriging for design optimization. *AIAA journal*, 2011. v. 49, n. 9, p. 2034–2046, 2011.

Apêndice A

Análises de resíduos e melhores soluções: de 20 a 40 ambulâncias

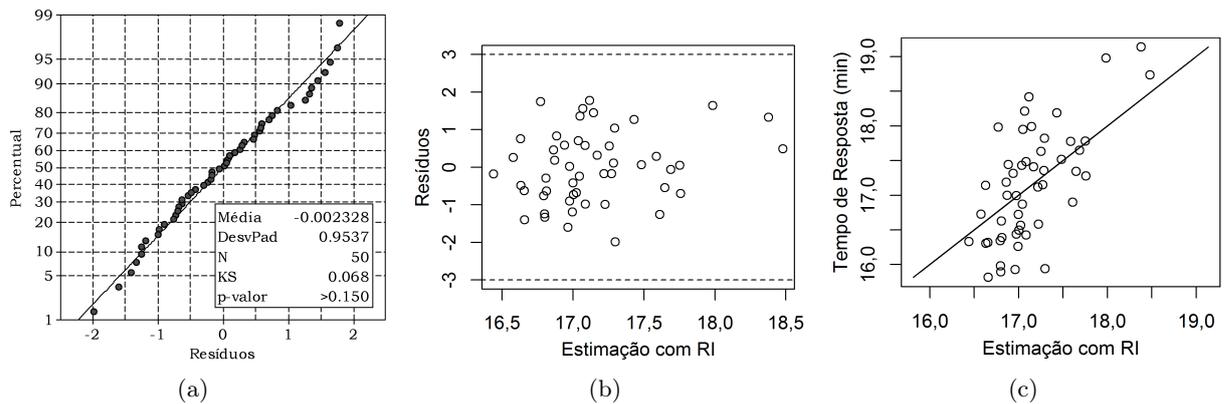


Figura A.1: Análise de resíduos para o experimento RI-20-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.1: Soluções encontradas para minimizar TR Médio com 20 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-20-50 #1	1 2 1 2 2 0 1 1 2 1 1 0 0 0 1 0 3 1 1 0	15,141	15,275	15,410	131,483
RI-20-50 #2	1 1 1 2 0 0 1 1 2 1 2 0 0 0 1 0 3 1 2 1	15,140	15,262	15,383	193,792
RI-20-50 #3	0 1 1 3 1 0 2 1 2 1 1 1 0 0 1 1 2 1 1 0	15,103	15,264	15,424	143,250
RI-20-50 #4	0 1 2 2 0 0 1 1 1 2 2 1 0 0 0 1 2 1 2 1	15,074	15,291	15,507	149,475
RI-20-50 #5	1 2 1 2 1 0 1 1 2 1 2 0 0 1 1 0 2 1 1 0	15,135	15,291	15,447	140,583

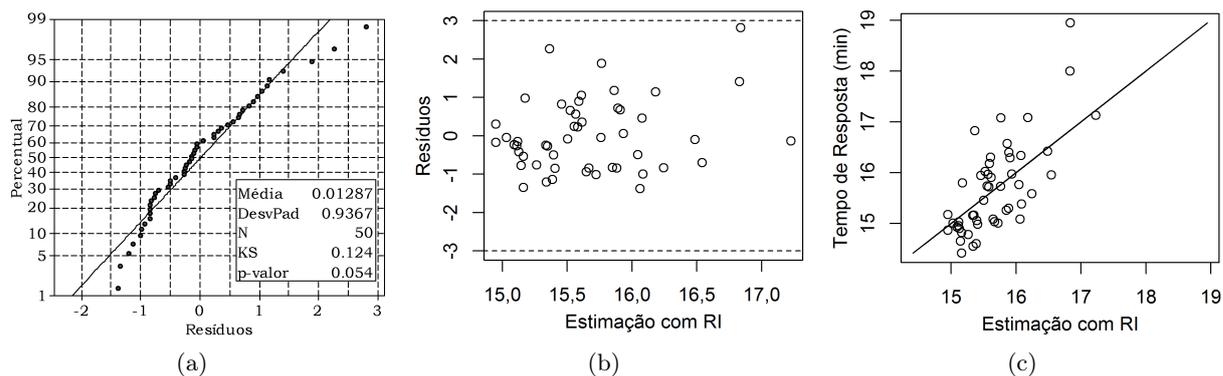


Figura A.2: Análise de resíduos para o experimento RI-22-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.2: Soluções encontradas para minimizar TR Médio com 22 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-22-50 #1	1 1 2 3 0 1 2 2 1 1 2 0 0 0 1 1 2 1 0 1	13,618	13,790	13,962	141,788
RI-22-50 #2	0 2 2 2 0 1 2 2 1 1 2 0 0 0 1 1 2 1 0 2	13,659	13,779	13,899	130,894
RI-22-50 #3	1 1 2 2 1 1 2 1 2 1 1 0 0 0 1 1 2 1 0 2	13,688	13,767	13,847	153,948
RI-22-50 #4	0 2 1 2 0 2 1 0 2 1 2 0 1 1 1 1 3 1 0 1	13,769	13,889	14,009	115,933
RI-22-50 #5	0 1 2 2 1 1 2 1 2 1 2 1 0 0 2 1 2 1 0 0	13,594	13,706	13,817	202,639

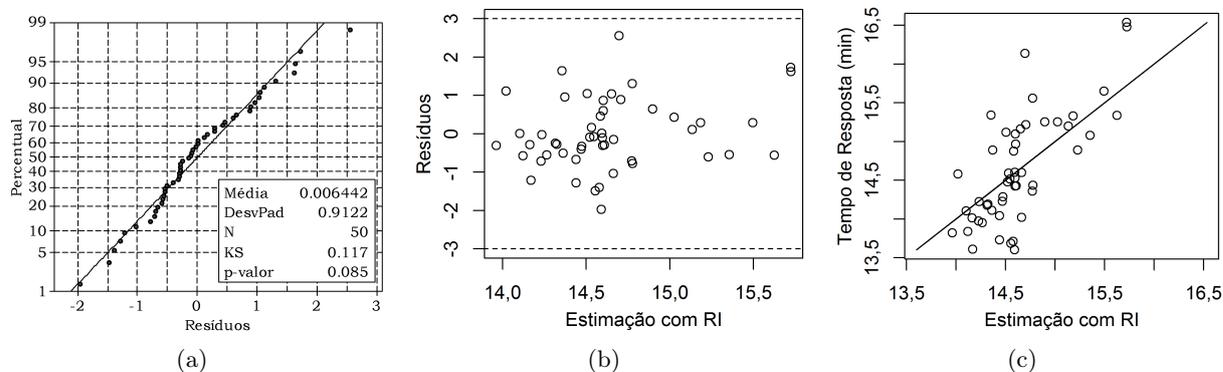


Figura A.3: Análise de resíduos para o experimento RI-24-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.3: Soluções encontradas para minimizar TR Médio com 24 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-24-50 #1	1 0 1 2 1 3 2 1 2 1 2 1 0 0 1 1 2 2 0 1	12,938	12,980	13,021	138,813
RI-24-50 #2	1 1 2 2 1 2 2 1 1 0 3 1 0 1 2 1 2 1 0 0	12,908	12,971	13,035	100,484
RI-24-50 #3	0 1 1 2 1 2 2 1 2 1 2 0 0 2 2 1 2 1 0 1	12,897	12,963	13,029	159,208
RI-24-50 #4	0 1 2 2 0 1 2 1 2 1 2 1 0 0 2 1 2 2 1 1	12,817	12,857	12,897	142,672
RI-24-50 #5	0 1 1 2 0 2 2 1 2 1 2 0 0 0 2 1 3 2 0 2	12,852	12,906	12,960	119,289

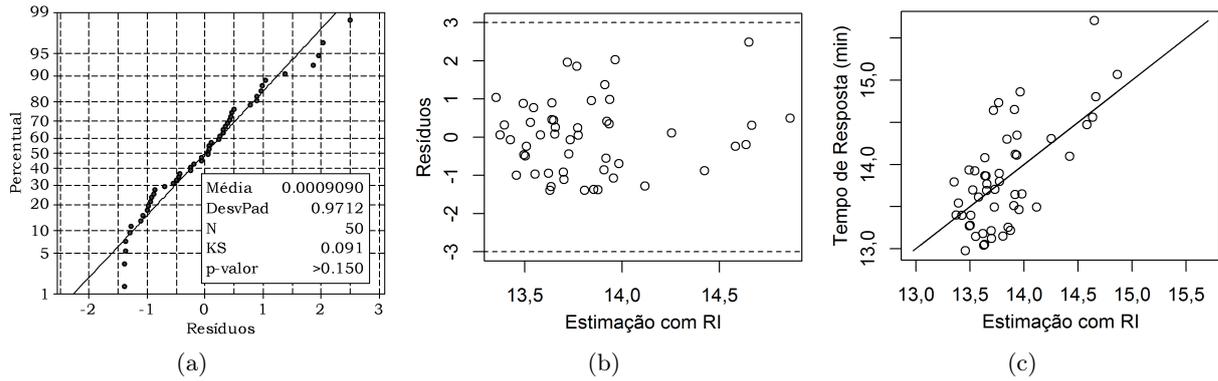


Figura A.4: Análise de resíduos para o experimento RI-26-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.4: Soluções encontradas para minimizar TR Médio com 26 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-26-50 #1	1 1 2 3 2 0 3 1 2 2 2 0 0 0 1 1 2 2 1 0	12,290	12,340	12,390	125.943
RI-26-50 #2	0 2 2 3 1 0 3 1 2 2 2 0 0 0 2 2 2 2 0 0	12,299	12,357	12,415	143.758
RI-26-50 #3	1 2 2 3 1 0 1 1 3 2 2 0 0 0 1 1 2 2 2 0	12,347	12,392	12,437	116.655
RI-26-50 #4	0 2 2 3 1 0 2 1 2 2 2 0 1 0 1 1 2 2 2 0	12,315	12,377	12,438	118.700
RI-26-50 #5	0 2 2 3 1 0 2 1 1 2 2 0 1 0 1 2 3 2 1 0	12,306	12,358	12,410	118.455

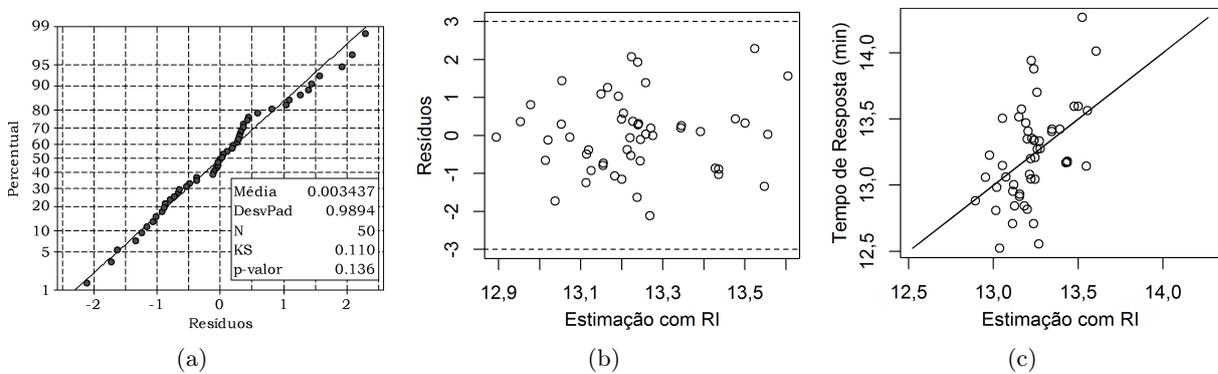


Figura A.5: Análise de resíduos para o experimento RI-28-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.5: Soluções encontradas para minimizar TR Médio com 28 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-28-50 #1	0 1 2 3 1 0 2 0 2 2 2 1 1 0 2 2 3 2 2 0	12,027	12,076	12,126	94,489
RI-28-50 #2	1 1 3 3 1 1 2 0 2 3 2 1 0 0 1 1 2 2 1 1	12,042	12,075	12,108	122,498
RI-28-50 #3	1 1 2 3 0 1 3 0 3 2 2 1 1 0 1 1 3 2 0 1	12,050	12,086	12,123	86,853
RI-28-50 #4	0 1 3 3 1 0 3 0 2 3 2 0 1 0 1 2 2 2 2 0	12,102	12,130	12,157	76,542
RI-28-50 #5	0 1 3 3 1 1 2 1 2 3 2 1 0 0 1 2 2 2 1 0	11,955	12,022	12,089	101,942

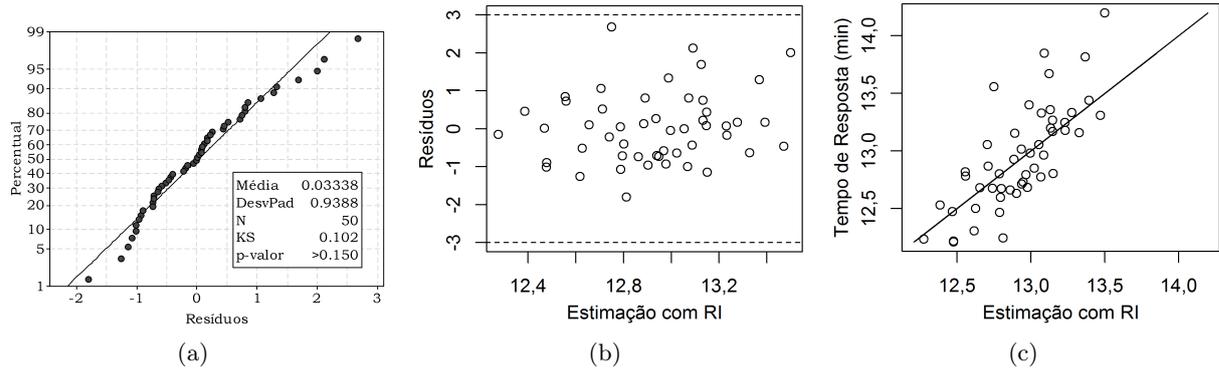


Figura A.6: Análise de resíduos para o experimento RI-30-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.6: Soluções encontradas para minimizar TR Médio com 30 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-30-50 #1	1 1 2 3 1 0 2 1 2 2 2 1 0 1 1 2 3 3 2 0	11,803	11,836	11,869	80,023
RI-30-50 #2	1 2 2 3 1 0 3 1 2 2 2 0 1 1 1 2 3 2 1 0	11,696	11,721	11,747	123,075
RI-30-50 #3	1 2 2 3 1 0 3 0 2 2 2 1 0 1 2 2 3 2 1 0	11,748	11,796	11,845	93,689
RI-30-50 #4	1 2 2 2 1 0 2 0 2 1 2 1 1 0 2 2 3 3 2 1	11,791	11,818	11,845	95,149
RI-30-50 #5	0 2 2 3 1 0 3 0 2 2 2 1 0 1 2 2 3 2 1 1	11,706	11,816	11,872	88,258

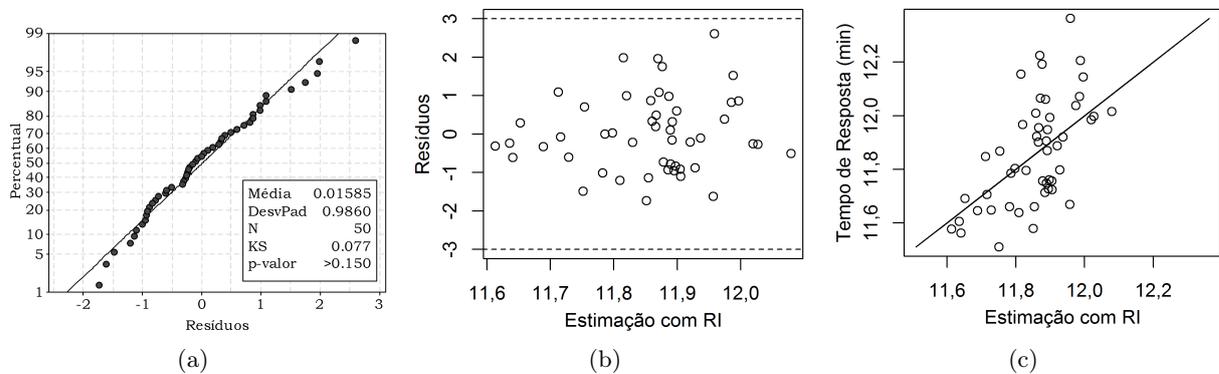


Figura A.7: Análise de resíduos para o experimento RI-40-50: (a) Teste de normalidade; (b) resíduos *versus* estimação; (c) valores reais *versus* estimação.

Tabela A.7: Soluções encontradas para minimizar TR Médio com 40 ambulâncias (tempos em minutos).

Solução	Ambulâncias por base	IC-	TR Médio	IC+	Duração
RI-40-50 #1	1 2 3 3 1 1 3 2 3 2 3 1 1 2 2 3 3 3 0 1	11,236	11,257	11,277	81,598
RI-40-50 #2	1 2 3 3 1 1 3 1 3 2 3 0 1 2 3 3 3 3 1 1	11,219	11,245	11,271	67,237
RI-40-50 #3	1 2 3 3 2 1 3 2 3 3 3 0 1 2 2 2 3 3 1 0	11,195	11,241	11,287	77,616
RI-40-50 #4	2 2 2 3 2 1 3 1 3 2 3 0 1 2 3 3 3 3 1 0	11,231	11,253	11,275	77,079
RI-40-50 #5	1 3 3 3 1 0 3 1 3 3 3 1 1 2 2 2 3 3 1 1	11,196	11,226	11,257	87,710