

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**METODOLOGIA PARA QUANTIFICAÇÃO E
ACOMPANHAMENTO DE INDICADORES-CHAVE DE
DESEMPENHO OPERACIONAL**

DISSERTAÇÃO DE MESTRADO

Cláudia Daniela Melo Giaquinto

Porto Alegre

2017

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**METODOLOGIA PARA QUANTIFICAÇÃO E
ACOMPANHAMENTO DE INDICADORES-CHAVE DE
DESEMPENHO OPERACIONAL**

Cláudia Daniela Melo Giaquinto

Dissertação de Mestrado apresentada como requisito parcial para obtenção do título de Mestre em Engenharia

Área de concentração: Pesquisa e Desenvolvimento de Processos

Linha de Pesquisa: Projeto, Simulação, Modelagem, Controle e Otimização de Processo Químicos e Bioprocessos.

Orientadores:

Prof. Dr. Marcelo Farenzena e

Prof^a Dr^a Luciane Ferreira Trierweiler

Porto Alegre

2017

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

A Comissão Examinadora, abaixo assinada, aprova a Dissertação *Metodologia para Quantificação e Acompanhamento de Indicadores-Chave de Desempenho Operacional* elaborada por Cláudia Daniela Melo Giaquinto, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:



Eng. Dr. Antônio Carlos Zanin – Petrobras



Prof. Dr. Michel Anzanello – PPGEQ – UFRGS



Prof. Dr. Nilo Sergio Medeiros Cardozo – PPGEQ – UFRGS

Resumo

Indicadores-chave de desempenho (KPIs) exercem um papel de extrema importância na indústria de processos, auxiliando na tomada de decisão. No entanto, para serem representativos precisam ser calculados de forma confiável. O presente trabalho propõe uma metodologia para o cálculo destes KPIs com base em técnicas de detecção do estado estacionário, remoção de ruído, propagação de erros e análise de sensibilidade. Estes KPIs foram apresentados, de acordo com o que consta na literatura, em uma nova ferramenta gráfica de acompanhamento proposta pelos autores, denominada *StatSSCandlePlot*. O *StatSSCandlePlot* apresenta os KPIs no padrão *candlestick*, que é bastante utilizado no mercado de ações, incluindo informações adicionais. O grande diferencial do *StatSSCandlePlot* é que os indicadores e suas respectivas propriedades exibidas são calculadas a partir de técnicas que englobam o tratamento de dados e análises estatísticas. A metodologia proposta foi aplicada em um estudo de caso de um chuveiro contendo dois princípios de aquecimento, gás e energia elétrica. Para este estudo, foi criado o Índice de Qualidade do Banho (IQB), que é um indicador dependente da temperatura e da vazão de saída, cujos dados foram avaliados em três cenários distintos, o primeiro quando o sistema é submetido a distúrbios na vazão, no segundo ocorre uma queda na temperatura da água fria e no último, o IQB foi avaliado quando o sistema foi submetido a distúrbios na vazão sob uma nova estratégia de controle da planta. A partir do *StatSSCandlePlot*, foi possível identificar as tendências do indicador nos diferentes cenários, a parcela de cada janela no estado estacionário, os valores a serem considerados do indicador e, de forma complementar, identificar a variável que mais influenciou na variação do indicador, através da análise de sensibilidade.

Palavras-chave: Indicadores-Chave de Desempenho, detecção do estado estacionário, remoção de ruído, propagação de erros, gráfico candlestick, avaliação de processos.

Abstract

Key performance indicators (KPIs) play an extremely important role in the process industry, aiding in decision-making. However, to be representative they need to be calculated reliably. The present work proposed a methodology for the calculation of these KPIs based on steady state detection, noise removal, error propagation and sensitivity analysis techniques. These KPIs were presented, as far as it is known, in a new graphical KPIs monitoring tool proposed by the authors, called StatSSCandlePlot. StatSSCandlePlot introduces KPIs in the candlestick standard, which is widely used in the stock market, including additional information. The major difference of StatSSCandlePlot is that the indicators and their respective displayed properties are calculated from techniques that encompass data processing and statistical analysis. The proposed methodology was applied in a case study of a shower containing two principles of heating, gas and electric energy. For this study the Bath Quality Index (BQI) was created, which is a temperature and output flow dependent indicator, whose data were evaluated in three different scenarios, the first one when the system was submitted to flow disturbances, in the second one, a decrease in the temperature of the cold water and in the last one, the IQB was evaluated when the system was submitted to disturbances in the flow under a new strategy of control of the plant. From the StatSSCandlePlot, it was possible to identify the trends of the indicator in the different scenarios, the portion of each window in the steady state, the values to be considered in the indicator and, in a complementary way, to identify the variable that most influenced the variation of the indicator, through the sensitivity analysis.

Keywords: Key Performance Indicators, steady state detection, noise removal, error propagation, candlestick chart, processes assessment.

“No meio da dificuldade encontra-se a oportunidade”.
(Albert Einstein)

Agradecimentos

Agradeço primeiramente a Deus, por tudo.

Aos meus pais, Alberto e Helena, e aos meus irmãos, André, Lucas e Gabriel, pelo amor, apoio e compreensão, sobretudo nesta nova etapa da minha vida.

Ao meu namorado Filipe, pela paciência, amor e companhia em todos os momentos.

Aos Professores Jorge Trierweiler, Luciane Trierweiler e Marcelo Farenzena, pelos ensinamentos e orientação.

Aos meus amigos Diego e Karinne, por todo apoio diário e leveza nos momentos mais difíceis.

A todos os meus amigos da UFRGS, especialmente, Denilson de Oliveira, José Eduardo, Juliano Sebben e Maria Andressa, pela ajuda, convivência e momentos de descontração.

Aos professores do Departamento de Engenharia Química da UFRGS, pela contribuição para a minha formação profissional.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pelo apoio financeiro.

SUMÁRIO

| | |
|--|-----------|
| Capítulo 1 – Introdução | 1 |
| 1.1 Objetivos do trabalho..... | 2 |
| 1.2 Estrutura da dissertação | 2 |
| Capítulo 2 – Revisão Bibliográfica | 3 |
| 2.1 Indicadores-chave de desempenho | 3 |
| 2.2 Tratamento de dados | 5 |
| 2.2.1 Remoção de ruído..... | 5 |
| 2.2.2 Detecção do estado estacionário | 8 |
| 2.3 Propagação de erros | 10 |
| 2.4 Análise de sensibilidade | 14 |
| 2.5 Visualização de dados | 18 |
| Capítulo 3 – Metodologia | 25 |
| 3.1 Função <i>ssdet</i> | 26 |
| 3.1.1 Remoção de ruído e detecção do estado estacionário | 26 |
| 3.1.2 Propagação de erros..... | 28 |
| 3.1.3 Análise de Sensibilidade | 29 |
| 3.2 StatSSCandlePlot | 30 |
| 3.2.1 Descrição do plot | 30 |
| 3.2.2 Exemplo de uso..... | 31 |
| Capítulo 4 – Estudo de Caso | 35 |
| 4.1 Escolha dos parâmetros e métodos..... | 39 |
| 4.1.1 Detecção do estado estacionário e remoção de ruído..... | 39 |
| 4.1.2 Método de expansão de séries de Taylor de 1ª ordem..... | 46 |
| 4.1.3 Determinação de N_A na análise de sensibilidade..... | 49 |
| 4.2 Cenários..... | 49 |
| 4.2.1 Cenário 1: Distúrbios na vazão | 49 |
| 4.2.2 Cenário 2: Queda da temperatura fria | 52 |
| 4.2.3 Cenário 3: Mudança na estratégia de controle | 55 |
| Capítulo 5 – Considerações Finais | 59 |
| 5.1 Conclusões..... | 59 |
| 5.2 Sugestões para trabalhos futuros | 61 |
| Referências | 63 |
| Apêndice A - Estados dos cenários | 69 |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 2.1. <i>Wavelet</i> spline quadrática (adaptada de Mallat (1999)). | 8 |
| Figura 2.2. Exemplo de aplicação do método DER ($l_{ee} = 0,8$). | 10 |
| Figura 2.3. Algoritmo simplificado do método de Sobol (adaptado de Zhang <i>et al.</i> (2015)). | 17 |
| Figura 2.4. Gráfico de dispersão da concentração de enxofre na gasolina. | 18 |
| Figura 2.5. Gráfico de dispersão com linha da concentração de enxofre na gasolina. | 19 |
| Figura 2.6. Gráfico de dispersão com barras de erro da concentração de enxofre na gasolina. | 19 |
| Figura 2.7. Diagrama de caixa da concentração de enxofre na gasolina. | 20 |
| Figura 2.8. Gráfico <i>candlestick</i> do dólar (valores fictícios). | 21 |
| Figura 2.9. Interpretação do <i>candlestick</i> . | 21 |
| Figura 3.1. Fluxograma representativo da metodologia utilizada no trabalho. | 25 |
| Figura 3.2. Esquema utilizado para a determinação do desvio-padrão da janela i . | 28 |
| Figura 3.3. Exemplo do <i>StatSSCandlePlot</i> . | 31 |
| Figura 3.4. Exemplo do <i>StatSSCandlePlot</i> com ênfase em um período. | 32 |
| Figura 3.5. Exemplo do <i>StatSSCandlePlot</i> ocultando os <i>candlesticks</i> . | 33 |
| Figura 4.1. Esquema representativo do chuveiro utilizado no estudo de caso. | 35 |
| Figura 4.2. Representação gráfica do IQB. | 37 |
| Figura 4.3. Estratégia de controle utilizada nos cenários 1 e 2. | 38 |
| Figura 4.4. Estratégia de controle utilizada no cenário 3. | 38 |
| Figura 4.5. (a) Gráfico do distúrbio na vazão em $t = 300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho). | 40 |
| Figura 4.6. Gráficos da identificação do estado estacionário para o IQB com $N = 4$ e diferentes valores de l_{ee} para $t_1 = 200$ s e $t_2 = 500$ s. | 41 |
| Figura 4.7. Gráficos da identificação do estado estacionário para o IQB com $l_{ee} = 0,7$ e diferentes valores de N para $t_1 = 200$ s e $t_2 = 500$ s. | 42 |
| Figura 4.8. Gráficos da identificação do estado estacionário para o IQB com $N = 4$ e diferentes valores de l_{ee} para $t_1 = 0$ s e $t_2 = 300$ s. | 44 |
| Figura 4.9. Gráficos da identificação do estado estacionário para o IQB com $l_{ee} = 0,7$ e diferentes valores de N para $t_1 = 0$ s e $t_2 = 300$ s. | 45 |
| Figura 4.10. Histograma e gráfico de probabilidade de F_s (vazão da água do banho). | 46 |
| Figura 4.11. Histograma e gráfico de probabilidade de T_4 (temperatura do banho). | 46 |
| Figura 4.12. Comparação entre as incertezas do IQB. | 48 |
| Figura 4.13. (a) Gráfico do distúrbio na vazão em $t = 500$ s e $t = 1300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho). | 50 |
| Figura 4.15. Análise de sensibilidade do IQB no cenário 1. | 52 |
| Figura 4.16. (a) Gráfico do distúrbio na temperatura da água fria (T_f), (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho). | 53 |
| Figura 4.18. Análise de sensibilidade do IQB no cenário 2. | 55 |
| Figura 4.19. (a) Gráfico do distúrbio na vazão em $t = 500$ s e $t = 1300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho). | 56 |
| Figura 4.21. Análise de sensibilidade do IQB no cenário 3. | 58 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 2.1: Descrições de <i>candlesticks</i> simples (adaptado de Marshall <i>et al.</i> (2006))..... | 22 |
| Tabela 4.1: Percentual da janela no EE de acordo com N e I_{ee} para $t_1 = 200$ s e $t_2 = 500$ s. | 43 |
| Tabela 4.2: Percentual da janela no EE de acordo com N e I_{ee} para $t_1 = 0$ s e $t_2 = 300$ s..... | 45 |
| Tabela 4.3: Erros relativos das incertezas de acordo com os coeficientes de similaridade | 48 |
| Tabela 4.4: Índices de sensibilidade total de acordo com o N_A | 49 |

NOTAÇÃO E SIMBOLOGIA

Nomenclatura

| | |
|----------------|---|
| B | produção no fundo da coluna de destilação |
| c | coeficiente de similaridade |
| CWT | transformada contínua de <i>Wavelet</i> |
| D | produção no topo da coluna de destilação |
| $dados$ | DataFrame de entrada da função <i>ssdet</i> |
| DWT | transformada discreta de <i>Wavelet</i> |
| EDC | capacidade equivalente de destilação |
| EE | estado estacionário |
| F_d | distúrbio na vazão |
| F_f | vazão da corrente fria |
| FI | medidor de vazão |
| F_q | vazão da corrente quente |
| F_s | vazão da corrente de saída |
| $I_{dif,i}$ | índice de estacionariedade |
| I_{ee} | estacionariedade limiar de projeto |
| IQB | índice de qualidade do banho |
| K | transformada de Fourier de uma função de suavização |
| KPI | indicador-chave de desempenho |
| $KPI_{est,j}$ | KPI do estacionário j |
| $KPI_{jani,i}$ | KPI da janela i |
| LI | medidor de nível do tanque |
| N | nível de decomposição de <i>Wavelet</i> |
| N_A | nº de amostras geradas |
| N_C | termo de correção da normalidade |
| n_{var} | nº de variáveis de entrada do indicador |
| P | potência do chuveiro |
| $p_{est,j}$ | peso do estacionário j |
| PR | indicador de taxa de produção |
| S_i | índice de sensibilidade de 1ª ordem |
| S_{ij} | índice de sensibilidade de 2ª ordem |
| S_{Ti} | índice de sensibilidade de ordem total |
| t | tempo |

| | |
|----------------------|---|
| t_{est_j} | período de tempo do estacionário j |
| T_d | distúrbio na temperatura |
| T_f | temperatura da corrente fria (caixa d'água) |
| TI | medidor de temperatura |
| T_{inf} | temperatura do ambiente |
| T_q | temperatura da corrente quente |
| T_4 | temperatura da corrente de saída |
| UEC | indicador de consumo de energia |
| $UEDC$ | capacidade de destilação equivalente utilizada |
| u_i | vetor aleatório para a variável i |
| v | vetor aleatório comum às variáveis |
| $Var(y)$ | variância de y |
| w | tamanho da janela na função $ssdet$ |
| $W_{\acute{a}gua}$ | quantidade de água de resfriamento no condensador |
| $wavfunc$ | função de <i>Wavelet</i> da função $ssdet$ |
| W_{vapor} | quantidade de vapor de aquecimento no refeedor |
| x_{est} | critério para a detecção do estado estacionário |
| \bar{x}_i | média de x_i |
| \hat{x}_i | primeira derivada suavizada |
| \ddot{x}_i | segunda derivada suavizada |
| x_f | posição da válvula na corrente fria |
| x_q | posição da válvula na corrente quente |
| x_s | posição da válvula na corrente de saída |
| $y(\bar{x})^{dwt}$ | transformada discreta de <i>Wavelet</i> de $y(x)$ |
| <i>Letras gregas</i> | |
| $\mu_{1,i}$ | primeiro momento central de x_i |
| σ_{est_j} | desvio-padrão do estacionário j |
| σ_{jan_i} | desvio-padrão da janela i |
| $\sigma_{x_i}^2$ | variância de x_i |
| $\sigma_{x_i x_j}^2$ | covariância entre x_i e x_j |
| σ_y^2 | variância de y |
| $\psi(t)$ | função de <i>Wavelet</i> |

Capítulo 1 – Introdução

Os processos industriais são avaliados por meio de indicadores de caráter estratégico calculados a partir de variáveis de processo. A estes indicadores dá-se o nome de indicadores-chave de desempenho (KPIs, do inglês *Key Performance Indicators*). No entanto, estas variáveis estão sujeitas a diversas fontes de variação, sendo elas resultantes de mudanças nas propriedades de um produto, falhas em equipamentos, erros de medição e ruído. O aumento da incidência destas variações interfere no valor final do indicador, tornando-o menos representativo, podendo levar a interpretações incorretas acerca do processo e decisões errôneas.

Em se tratando da determinação destes indicadores, algumas questões são levantadas: como representar um indicador de um período em que os dados assumiram valores muito diferentes? Diante de dados dinâmicos, como calcular um indicador estático? Qual a variável que mais está influenciando na variação do indicador?

Tomando-se como exemplo a eficiência de uma caldeira, que pode ser determinada pela razão entre a energia que sai e a energia que entra. Este indicador depende de parâmetros como a vazão de água que alimenta a caldeira, do dessuperaquecedor e de correntes secundárias; pressão e temperatura de todas as correntes que entram e saem da caldeira, vazão de combustível e poder calorífico do combustível, de acordo com a ASME (2008). Cada parâmetro está associado a uma incerteza e assume valores diferentes no decorrer do processo, que são refletidos na incerteza e no valor final da eficiência da caldeira. Adicionalmente, cada parâmetro possui uma magnitude diferente de influência na eficiência.

Diante desta problemática, o presente trabalho buscou responder estas questões partindo de conceitos bem estabelecidos na literatura, para o propósito específico de determinação e acompanhamento de KPIs.

1.1 Objetivos do trabalho

O objetivo principal deste trabalho foi desenvolver uma metodologia para o cálculo e o acompanhamento de KPIs operacionais envolvendo quatro etapas, a saber:

- 1) Tratamento de dados: utilizando técnicas de remoção de ruído e detecção do estado estacionário, visto que diante de um KPI estático os dados de entrada devem estar no estado estacionário;
- 2) Análises estatísticas: empregando métodos de propagação de erros, dado que os KPIs operacionais estão sujeitos à variabilidade dos dados;
- 3) Visualização de dados: através de uma ferramenta gráfica proposta pelos autores denominada *StatSSCandlePlot*, visando uma apresentação apropriada dos dados para facilitar o desenvolvimento de análises. Diante do que foi pesquisado na literatura, esta é a primeira ferramenta de acompanhamento de KPIs aplicada à engenharia química;
- 4) Diagnóstico: utilizando a análise de sensibilidade para identificar as variáveis que mais impactam na variação do KPI.

1.2 Estrutura da dissertação

Este trabalho está dividido em cinco capítulos. O capítulo 2 inclui uma revisão bibliográfica a respeito de técnicas para remoção de ruído e detecção do estado estacionário, métodos utilizados para a propagação de erros, análise de sensibilidade e apresenta algumas ferramentas gráficas de visualização de dados. O terceiro capítulo descreve a metodologia empregada para cálculo e acompanhamento de KPIs. O quarto capítulo apresenta um estudo de caso para a aplicação da metodologia e no capítulo 5 estão as conclusões e sugestões para trabalhos futuros.

Capítulo 2 –Revisão Bibliográfica

Este capítulo contém uma abordagem teórica dos principais assuntos necessários como base para o desenvolvimento deste trabalho. O capítulo está subdividido em 5 seções que abordaram os indicadores-chave de desempenho, técnicas para o tratamento de dados, análises estatísticas e ferramentas gráficas de visualização de dados.

2.1 Indicadores-chave de desempenho

O gerenciamento de ativos em uma indústria é realizado através da medição do desempenho de seus processos. Esta medição é quantificada pela diferença entre os KPIs e os valores de referência pré-estabelecidos. Os KPIs servem de base para a tomada de decisões em relação à análise de qualidade de produto e serviço, investimentos, controle de processos e implantações de melhorias. No entanto, algumas dificuldades são encontradas na efetivação de um sistema de medição, a saber: disponibilidade, consistência e fidedignidade em obter as informações necessárias. A ausência dessas características pode gerar um sistema propenso a falhas e até inviável (Fischmann e Zilber, 1999; Müller, 2003; Weber e Thomas, 2005; Nader *et al.*, 2012).

Na literatura, existem inúmeros KPIs para diferentes finalidades em diversas áreas. Na área de qualidade, por exemplo, utiliza-se o número de reclamações, número de produtos defeituosos, custo da não-qualidade e perdas durante o processo, como apresentado por Torkko *et al.* (2013) para as indústrias farmacêutica e alimentícia. Com relação à manutenção, tem-se os KPIs de confiabilidade, como o tempo médio entre as

falhas, disponibilidade de equipamento e razão de manutenção preventiva, e os KPIs baseados em custos, como o custo de manutenção de emergência e o cumprimento do orçamento de manutenção, como discutidos por Assaf *et al.* (2015) para a indústria petroquímica. Montequín *et al.* (2013) apresentaram alguns KPIs da área financeira para fabricação de softwares, a saber: retorno de investimento, valor agregado, eficiência, percentual de custo de desenvolvimento e receita de vendas de clientes externos e internos. Adicionalmente, os autores também usam KPIs de recursos humanos, como rotatividade de pessoal, rotatividade de empresas terceirizadas e capital humano.

Zhu *et al.* (2014) avaliaram dois KPIs de uma unidade de destilação, a saber: indicador de taxa de produção, PR , e o indicador de consumo de energia, UEC , apresentados, respectivamente nas Equações 2.1 e 2.2.

$$PR = \frac{D}{D+B} \quad (2.1)$$

$$UEC = \frac{W_{vapor} + W_{\acute{a}gua}}{D+B} \quad (2.2)$$

onde: D e B correspondem, respectivamente, à produção no topo e na base da coluna, W_{vapor} é a quantidade de vapor de aquecimento no refeedor e $W_{\acute{a}gua}$ é a quantidade de água de resfriamento utilizada no condensador.

O indicador PR estabelece a relação entre a produção alvo e o que de fato foi produzido, portanto, quanto maior o seu valor, melhor o desempenho da coluna. É um KPI que avalia o nível de trabalho e a operação do equipamento. Já o UEC , é a razão entre a energia consumida e o que foi produzido, logo, quanto menor o seu valor, menor o impacto nos custos de produção (Zhu *et al.*, 2014).

Narvaes-Garcia *et al.* (2015) e Aurangzeb e Jana (2016) também avaliaram diferentes KPIs referentes a colunas de destilação. No primeiro trabalho são apresentados diversos indicadores relacionados com o rendimento, capacidade, eficiência, custo e lucro de uma coluna de destilação em batelada. Já no segundo, são propostos KPIs de eficiência térmica, economia de energia e desempenho econômico de uma coluna de parede dividida.

Ceric (2012) descreve alguns KPIs voltados para refinarias de petróleo baseados na empresa Solomon Associates, a qual faz o cálculo dos indicadores e estabelece um *ranking* mundial entre as refinarias que fazem parte do estudo, auxiliando na tomada de decisões internas como, por exemplo, investimento em novos equipamentos, manutenção e outras mudanças nos processos. Nesse trabalho, destaca-se um importante KPI para a refinaria, que é o Refinery EDC, que consiste na capacidade de destilação equivalente da refinaria. Este indicador considera todas as unidades instaladas e existentes na refinaria. Outro indicador é o Refinery UEDC, que considera apenas as unidades que são utilizadas.

Visto que os KPIs têm um papel de grande importância para a avaliação de um determinado processo, os dados fornecidos para a sua determinação precisam ser consistentes para que o KPI seja representativo e as decisões tomadas sobre este valor calculado sejam as mais próximas do ótimo. Logo, antes de calcular os KPIs, os dados precisam ser pré-processados, ou seja, é necessário um tratamento prévio dos dados.

2.2 Tratamento de dados

Em indústrias de processos, as variáveis estão sujeitas a ruídos e erros inerentes aos processos e aos instrumentos de medição. Logo, é de grande importância realizar o tratamento dos dados obtidos para minimizar a interferência desses erros nos valores finais dos indicadores. Dentre os tratamentos de dados existentes, destacam-se a remoção de ruído e a detecção do estado estacionário.

2.2.1 Remoção de ruído

Industrialmente, os sinais existentes possuem ruídos e somente em condições ideais os mesmos poderiam ser negligenciados e sua remoção se tornar desnecessária. Infelizmente, os ruídos corrompem o sinal e necessitam ser removidos para que se possa prosseguir com a análise de dados e a extração de informações úteis (Justen e Ramlau, 2009). Na literatura, consta uma vasta aplicação de diferentes técnicas de remoção de ruídos em diversas áreas como, por exemplo, na área médica com a finalidade de melhorar a qualidade dos dados de ultrassonografia e eletrocardiografia (Pratik *et al.*; Luo *et al.*, 2006), área geofísica (To *et al.*, 2009), automotiva (Lee e Kim, 2015), imagens (Nasri e Nezamabadi-Pour, 2009) e materiais (Staszewski, 2002) .

Com relação aos métodos de remoção de ruído existentes, tem-se o método da fração máxima de ruído (MNF, do inglês *Maximum Noise Fraction*) (Sun *et al.*, 2009), o método de Fourier (Justen e Ramlau, 2009; To *et al.*, 2009) e o método de supressão de ruídos por *Wavelets*, também conhecido como “*Wavelet Shrinkage Denoising*”. Este último tem apresentado bons resultados, além da possibilidade de aliar a remoção de ruído à estimação de derivadas, por meio de transformadas de *Wavelet* (Taswell, 2001; Luo *et al.*, 2006; Nasri e Nezamabadi-Pour, 2009).

O método MNF, desenvolvido originalmente para a remoção de ruído em imagens e estendido para séries temporais, assume que o sinal está combinado com uma fonte de ruído não correlacionada com o sinal verdadeiro. O método descreve o sinal como sendo composto pelo sinal puro e pelo ruído, separando-os em subspaços diferentes em termos de frações máximas de ruído, obtidas através da transformada de MNF. À medida que a fração de ruído diminui, a qualidade da imagem ou dos dados vai melhorando (Green *et al.*, 1988; Anderle e Kirby, 2001; Sun *et al.*, 2009).

Entende-se por *Wavelet* uma função ondulatória de forma irregular e assimétrica capaz de decompor uma série de dados no domínio do tempo e da frequência, diferentemente da análise de Fourier, na qual se analisa o sinal decomposto apenas no domínio da frequência. Esta transformação de um sinal apenas para o domínio da frequência pode desprezar informações temporais da série, restringindo a análise de Fourier para séries com comportamentos estacionários, problemas lineares ou funções periódicas (Rocha, 2008).

A transformada de *Wavelet* tem duas variantes: contínua (CWT, do inglês *Continuous Wavelet Transform*) e discreta (DWT, do inglês *Discrete Wavelet Transform*), ambas possuem propriedades de remoção de ruído e diferenciação, sendo aplicadas para a estimativa do cálculo de derivadas em diversas áreas (Luo *et al.*, 2006; Rocha, 2008). De acordo com Luo *et al.* (2006), o método DWT tem a vantagem de apresentar maior eficiência computacional na remoção do ruído do sinal e estimativa das derivadas, quando utilizado o algoritmo *à trous* (Shensa, 1992; Wink e Roerdink, 2010).

O princípio de resolução do método DWT consiste na decomposição do sinal em componentes de baixa e alta frequência da série, de acordo com uma função base (família de *Wavelet*) e da quantidade de níveis de decomposição. O componente de alta escala ou baixa frequência da série é analisado em cada nível e, de acordo com um limiar, os coeficientes desta série que apresentam valores menores que um valor crítico são eliminados, sendo mantidos os mais representativos, oriundos em sua maioria do sinal e não do ruído. Após a decomposição, tem-se a reconstrução do sinal através da transformada inversa, gerando um sinal com uma menor quantidade de ruídos quando comparado ao original (Luo *et al.*, 2006; Rocha, 2008).

Na Equação 2.3, encontra-se a relação utilizada no método DWT para a estimação da derivada de uma função $y(x)$ (Luo *et al.*, 2006).

$$\frac{dy(x)}{dx} = \frac{y(\bar{x})^{dwt}}{K \cdot 2^{3N/2}} \quad (2.3)$$

onde: $y(\bar{x})^{dwt}$ é a transformada discreta de *Wavelet* de $y(x)$, que é calculada em relação a uma função de *Wavelet*; N é o nível de decomposição de *Wavelets* e K é um parâmetro não-nulo correspondente à transformada de Fourier de uma função de suavização. O fator 2^N corresponde ao parâmetro de dilatação, para o qual um baixo valor resulta em uma elevada sensibilidade à ruídos e um alto valor resulta em um forte cancelamento de ruídos (Luo *et al.*, 2006).

Na Equação 2.3, diversas famílias de *Wavelet* podem ser empregadas, a saber: Haar, spline quadrática, Daubechies, dentre outras. A família Daubechies possui dez formas diferentes, sendo a primeira, conhecida como db1, a função de Haar. A função de Haar é a mais simples função de *Wavelet* e corresponde a uma função descontínua do tipo degrau, que analisa um sinal constante por partes (Lee e Yamamoto, 1994; Mallat, 1999; Campo *et al.*, 2016). Na Equação 2.4, tem-se a transformada de *Wavelet* da função Haar.

$$\psi(t) = \begin{cases} -1 & \text{se } 0 \leq t < 1/2 \\ 1 & \text{se } 1/2 \leq t < 1 \\ 0 & \text{caso contrário} \end{cases} \quad (2.4)$$

A transformada de *Wavelet* da spline quadrática tem forma semelhante à função seno, com diferente amplitude, como pode ser visto na Figura 2.1 (Lee e Yamamoto, 1994; Zhang *et al.*, 2001).

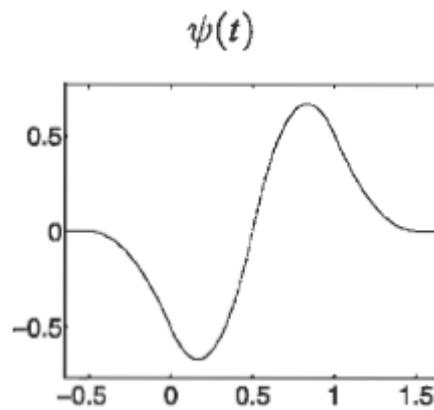


Figura 2.1. Wavelet spline quadrática (adaptada de Mallat (1999)).

Em Misiti et al. (2016) podem ser encontradas mais informações acerca das famílias de Wavelet e do método DWT, bem como da sua programação no *software* MatLab®.

2.2.2 Detecção do estado estacionário

A etapa subsequente à remoção de ruído é a de detecção do estado estacionário, que é crítica para a avaliação do desempenho, controle e otimização de um processo, detecção de falhas e reconciliação de dados (Mejia et al., 2010a; 2010b; Korbel et al., 2014). Jiang et al. (2003) ressaltam a importância da utilização de dados apenas quando o estado estacionário é atingido, principalmente na construção e aplicações de modelos, de forma a garantir resultados significantes e modelos menos propensos a falhas.

O método proposto por Cao e Rhinehart (1995) (CAO) é um método consolidado na literatura em identificação do estado estacionário e consiste na comparação da razão das variâncias de ruído estimadas, R , com uma razão crítica de projeto, R_C , que está associada a um nível de significância. Se a média da variável do processo está variando, $R > R_C$, do contrário, se a média é constante, $R \leq R_C$, o processo está no estado estacionário. O ajuste dos parâmetros, por não ser trivial, pode ser realizado com base nas tabelas e orientações reportadas em Cao e Rhinehart (1997) em função das propriedades do sinal (Mejia et al., 2010b; Korbel et al., 2014).

Mejia et al.(2010) propuseram um novo método baseado na correlação local (LOC), comparando-o com o método CAO. Este novo método possui um índice baseado na relação sinal ruído obtido a partir da correlação local. O método proposto teve um desempenho maior do que o CAO, possuindo um número menor de parâmetros com um

ajuste mais fácil dos mesmos. Apesar do método CAO apresentar vantagem na velocidade de processamento e baixo esforço computacional, exige um ajuste não trivial dos parâmetros.

Em outro trabalho, Mejia *et al.* (2010b), fizeram uma comparação entre os métodos CAO, LOC e um novo método proposto para a identificação do estado estacionário, baseado na estimativa de derivadas (DER). Após realizar simulações em diferentes cenários de ruídos e tipos de sinal, os autores concluíram que o método DER possui o melhor desempenho dentre os métodos em todos os cenários utilizando duas formas diferentes de ajuste dos parâmetros, além de se destacar pela simplicidade.

O método DER baseia-se no conceito de variação nula ou aproximadamente nula das variáveis do processo. Esta proposta faz uma análise comparativa do índice de estacionariedade, denotado por $I_{dif,i}$ (Equação 2.5) com um valor de estacionariedade limiar de projeto, I_{ee} (Mejia *et al.*, 2010b).

$$I_{dif,i} = 1 - \left[\frac{|\hat{x}_i| + 10|\ddot{x}_i|}{\max(|\hat{x}_i| + 10|\ddot{x}_i|)} \right] \quad (2.5)$$

A Equação 2.5 está baseada na estimativa das primeiras e segundas derivadas do sinal após a remoção de ruídos, denotadas por \hat{x}_i e \ddot{x}_i , respectivamente. A Equação 2.5 normaliza a soma $|\hat{x}_i| + 10|\ddot{x}_i|$ e retorna o seu valor complementar. Após a determinação do $I_{dif,i}$, o seu valor é comparado com o I_{ee} e, para o caso em que $I_{dif,i} > I_{ee}$, o sistema está no estado estacionário, caso contrário, se $I_{dif,i} \leq I_{ee}$, o sistema não se encontra no estado estacionário. Logo, quanto maior a rigurosidade, ou seja, quanto maior o I_{ee} , uma menor porção será considerada como estacionária. A seguir, na Figura 2.2, tem-se um exemplo de aplicação do método DER utilizando um I_{ee} igual a 0,8.

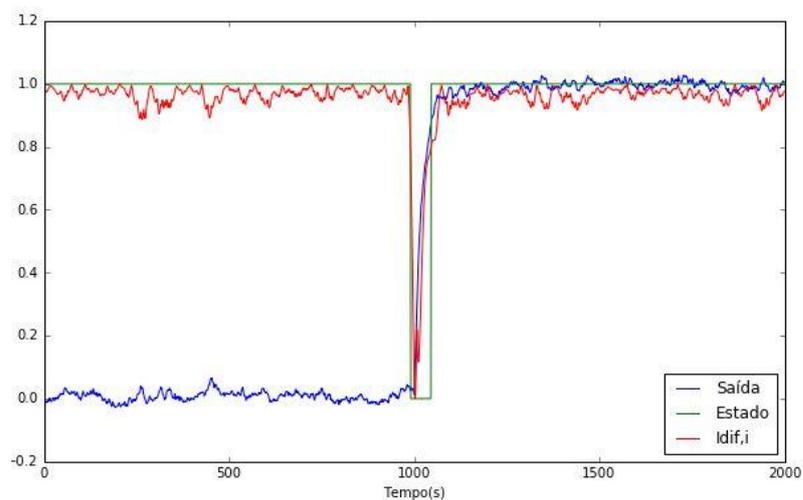


Figura 2.2. Exemplo de aplicação do método DER ($I_{ee} = 0,8$).

Na Figura 2.2, observa-se que quando o $I_{dif,i}$, representado pela curva vermelha, está próximo da unidade e, portanto, maior do que o I_{ee} , o sistema está no estado estacionário, assumindo um valor igual a 1. Do contrário, observa-se que na região próxima ao tempo de 1000 segundos, o $I_{dif,i}$ fica abaixo de 0,8 e o estado igual a zero, considerado não estacionário.

O método DER aliado ao método DWT torna-se uma ferramenta simples e eficiente para a detecção do estado estacionário, visto que este último integra a estimativa das derivadas com a remoção de ruídos do sinal, fornecendo os dados de entrada necessários para o método DER.

2.3 Propagação de erros

As variáveis de processo estão sujeitas a erros de medição e erros aleatórios, os quais nem sempre podem ser removidos em sua totalidade, portanto, torna-se essencial a rastreabilidade de suas ocorrências e a realização do estudo de propagação de erros. Este estudo, também conhecido como análise de tolerância, consiste em uma técnica para examinar o efeito das tolerâncias parciais, referentes às variáveis de entrada, na saída do sistema (Cox, 1979). Mais precisamente, de acordo com Xue *et al.* (2015), trata-se de um problema de estimação estatística da distribuição de saída a partir de variações devido aos erros aleatórios nos componentes de entrada.

Em virtude de sua elevada importância, o estudo de propagação de erros tem sido extensivamente aplicado em diferentes áreas como, por exemplo, reatores nucleares (Cox, 1979), montagens mecânicas (Chase e Parkinson, 1991; Glancy e Chase, 1999; Yan *et al.*, 2015), análise de ciclo de vida (Heijungs e Lenzen, 2014), sistemas de informações geográficas (Xue *et al.*, 2015) e explorações espaciais (Sun e Zhang, 2016).

Existem diversos métodos para a propagação de erros em um sistema, a saber: análise por intervalos, metodologia da superfície de resposta (RSM, do inglês *Response Surface Methodology*), teste da sensibilidade de amplitude de Fourier (FAST, do inglês *Fourier Amplitude Sensitivity Test*), sendo mais comumente utilizados os métodos de Monte Carlo (MC) e expansões de séries de Taylor (Helton e Davis, 2003; Xue *et al.*, 2015; Sun e Zhang, 2016).

Através da análise por intervalos é possível utilizar intervalos de entrada ao invés de valores nominais e submetê-los às operações matemáticas para encontrar o intervalo referente à saída. É um método utilizado principalmente para erros de arredondamento e de medição. No entanto, é limitado às quatro operações básicas e às vezes, por ser um método conservativo, resulta em intervalos de resposta superestimados (Anderson e Mattson, 2012; Jiang *et al.*, 2016).

No método RSM, as incertezas das variáveis de entrada são propagadas de acordo com um planejamento experimental para selecionar a entrada do modelo, permitindo uma exploração empírica e eficiente da superfície de resposta. O planejamento experimental a ser escolhido depende das propriedades do modelo e do tipo de resultados desejados (Kuo e Uppuluri, 1983; Helton e Davis, 2003). Este método pode apresentar dificuldades relacionadas ao desenvolvimento de um planejamento experimental apropriado, à detecção de limiares, descontinuidades e não linearidades, dificuldades em incluir correlações e restrições entre variáveis de entrada e, adicionalmente, pode ser difícil construir uma aproximação da superfície de resposta apropriada (Helton, 1993).

A abordagem do método FAST está baseada em uma decomposição das séries de Fourier que representam o modelo para obter a contribuição individual de cada variável de entrada para a variância do modelo predito (Helton, 1993). Xue *et al.* (2015) destacam

que os métodos RSM e FAST apresentam dificuldades quando as variáveis de entrada são dependentes.

O método de Monte Carlo consiste em um dos métodos mais simples e bem estabelecidos para análise estatística de tolerância através de números aleatórios gerados para simular variações naturais do processo. Este método é flexível, modela com precisão os efeitos não lineares e não exige que as distribuições de entrada e saída sejam Gaussianas. No entanto, tem um alto custo computacional e a cada modificação no modelo e/ou nos valores de entrada requer uma nova execução do método, sendo proibitivo em processos iterativos (Anderson e Mattson, 2012). Outra desvantagem, segundo Xue *et al.* (2015), é o fato deste método não dispor de uma relação analítica entre as incertezas das entradas e das saídas. O tamanho das amostras para este método pode variar de 5.000 a 100.000, a depender da acurácia requerida na simulação, que aumenta de acordo com o tamanho da amostra, aumentando também o esforço computacional (Glancy e Chase, 1999; Anderson e Mattson, 2012). Informações adicionais sobre o método de Monte Carlo podem ser encontradas em Kuo e Uppuluri (1983), Helton (1993), Helton e Davis (2003) e Yan *et al.* (2015).

O uso de expansões de séries de Taylor para a propagação de erros está bem consolidado na literatura e tem como ideia principal a determinação de uma expressão analítica para as incertezas das saídas em função das entradas, cujos efeitos podem ser explicados e analisados através desta relação (Xue *et al.*, 2015).

Seja y uma função de n variáveis como definida pela Equação 2.6. A aproximação por séries de Taylor de primeira ordem em torno das médias das variáveis de entrada é dada pela Equação 2.7, na qual as derivadas parciais são avaliadas na média.

$$y = f(x_1, x_2, \dots, x_n) \quad (2.6)$$

$$y \approx y_1 = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} (x_i - \bar{x}_i) \quad (2.7)$$

A partir da Equação 2.7, tem-se a fórmula da variância de y (σ_y^2) em função das variâncias das variáveis de entrada ($\sigma_{x_i}^2$), como mostrado na Equação 2.8.

$$\sigma_y^2 \approx \sum_{i=1}^n \left[\left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 + 2 \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \sigma_{x_i x_j}^2 \right] \quad (2.8)$$

O somatório interno na Equação 2.8 está relacionado com a covariância entre as variáveis de entrada x_i e x_j ($\sigma_{x_i x_j}^2$). Este termo é empregado quando as variáveis de entrada possuem interdependência, assumindo um valor igual a zero quando as variáveis são independentes e reduzindo a expressão de σ_y^2 à Equação 2.9.

$$\sigma_y^2 \approx \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 \quad (2.9)$$

Anderson e Mattson (2012) afirmam que a simplificação da covariância é tipicamente feita na literatura e na prática. No entanto, é necessário avaliar se ao ignorar a relação entre as variáveis de entrada, a precisão da variância diminui.

A Equação 2.8 assume que as variáveis de entrada possuem distribuição Normal. De acordo com Anderson *et al.* (2012), uma correção deve ser inserida na Equação 2.8 quando estas variáveis não possuem distribuição Gaussiana. Na Equação 2.10, encontra-se o termo de correção e na Equação 2.11, a relação resultante para calcular a variância de y após a adição do referido termo.

$$N_C = \left(\frac{\partial f}{\partial x_i} \right)^2 (\mu_{1,i})^2 + 2 \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} (\mu_{1,i})(\mu_{1,j}) \quad (2.10)$$

$$\sigma_y^2 \approx \sum_{i=1}^n \left[\left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 - N_C + 2 \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \sigma_{x_i x_j}^2 \right] \quad (2.11)$$

onde: $\mu_{1,i}$ corresponde ao primeiro momento central da variável x_i .

A partir dos valores da média e do desvio-padrão da distribuição de saída, admite-se que a sua distribuição é Gaussiana. Xue *et al.* (2015) afirmam que estes dois parâmetros podem fornecer informações suficientes sobre o erro envolvido. Cox (1979), Glancy e Chase (1999) e Anderson e Mattson (2012) apresentam as expansões de séries de Taylor de segunda ordem para descrever completamente a distribuição de saída através da determinação dos momentos de terceira e de quarta ordem, que correspondem à assimetria e à curtose, respectivamente. Os autores afirmam que a determinação destes parâmetros é fundamental quando o objetivo é obter uma representação significativa da distribuição de saída.

A aproximação por séries de primeira ordem oferece, no geral, resultados significativos e facilidade em sua aplicação, sendo recomendada para funções lineares e polinomiais. No entanto, para funções não lineares os resultados podem ficar menos precisos, sendo sugerido o truncamento da série com termos de ordem mais elevada (Anderson e Mattson, 2012; Xue *et al.*, 2015).

Visando o aumento da acuracidade dos resultados, Xue *et al.* (2015) compararam as variâncias obtidas pelas expansões de primeira, terceira e quinta ordem de séries de Taylor tendo como referência o método de Monte Carlo. Eles observaram que a expansão de quinta ordem é o método mais preciso. Porém, a diferença é pequena quando comparada com a expansão de terceira ordem, menor do que 0,5%, e apresenta um aumento significativo do esforço computacional devido ao número de derivadas parciais que precisa ser calculado. Logo, é preciso estabelecer um balanço de interesses a depender da necessidade de aplicação, pois o aumento da precisão vem acompanhado do aumento da complexidade, sobretudo quando o número de variáveis de entrada é elevado. Anderson *et al.* (2012) mostram que o custo computacional cresce exponencialmente com o número de entradas do sistema, além dos cálculos requeridos de momentos de ordem superior e covariância entre as variáveis, podendo tornar o método proibitivo para sistemas complexos.

2.4 Análise de sensibilidade

Os indicadores nem sempre possuem fórmulas simples e poucos parâmetros, portanto, ao se deparar com um elevado número de variáveis e equações complexas, é importante saber a quais parâmetros o indicador é mais sensível, ou seja, quais as variáveis de entrada que mais influenciam na variável de saída. Desta forma, os esforços para a melhoria do desempenho de um determinado processo passam a ser concentrados e mais rigorosos nas variáveis mais críticas, contribuindo para a identificação da causa raiz de um eventual problema, redução de gastos e focando na melhoria contínua.

A análise de sensibilidade consiste em um estudo de como a incerteza da saída do modelo é influenciada pelas incertezas das entradas do mesmo. Portanto, este estudo é complementar ao de propagação de erros, o qual quantifica a incerteza da saída a partir

das incertezas da entrada (Helton e Davis, 2003; Saltelli *et al.*, 2008). Este estudo possui diversas aplicações como, por exemplo, em montagens mecânicas (Chase e Parkinson, 1991), análise de risco em alimentos (Christopher Frey e Patil, 2002) e em sistemas bioquímicos e farmacológicos (Kent *et al.*, 2013; Zhang *et al.*, 2015).

No caso das montagens mecânicas, Chase e Parkinson (1991) afirmam que a análise de sensibilidade auxilia o *designer* a ter conhecimento de quais variações dos parâmetros possuem o efeito mais significativo nas características críticas da peça, auxiliando, desta forma, no diagnóstico de qual componente deve-se focar para a melhoria do *design*. A partir do conhecimento da sensibilidade do componente, pode-se reduzir a variância total devido à redução da variância do componente mais sensível e também reduzir custos através da folga de tolerância do componente menos sensível.

Existem duas vertentes da análise de sensibilidade, a saber: local e global. A análise de sensibilidade local avalia as mudanças na saída do modelo baseando-se na variação individual de cada variável de entrada mantendo as demais variáveis do domínio constantes, considerando, desta forma, que as variáveis são independentes. Esta abordagem deve ser usada quando a saída é linearmente relacionada com os parâmetros de entrada e tem como principal limitação o fato de não avaliar as mudanças simultâneas em todos os parâmetros do modelo, assim como as interações entre os parâmetros. Já na análise de sensibilidade global, todos os parâmetros são variados simultaneamente no domínio, permitindo avaliar tanto a contribuição relativa de cada parâmetro quanto as interações entre os mesmos para o modelo de saída (Saltelli *et al.*, 2008; Zhang *et al.*, 2015).

Diversos métodos para a análise de sensibilidade estão disponíveis na literatura, dentre eles, destacam-se o método de Morris (Morris, 1991; Campolongo *et al.*, 2007), o método FAST (Cukier *et al.*, 1973; Saltelli *et al.*, 1999; Zhang *et al.*, 2015) e o de Sobol (Sobol, 2001; Saltelli, 2002; Saltelli *et al.*, 2010; Zhang *et al.*, 2015), os quais serão subsequentemente detalhados.

O método de Morris está baseado no cálculo de uma série de taxas incrementais, também chamadas de Efeitos Elementares (EE, do inglês *Elementary Effects*), para cada variável de entrada. A partir do valor médio dessas taxas, avalia-se a importância geral de cada entrada específica. Este método é adequado quando o número de variáveis é alto e

o modelo exige um alto custo computacional (Morris, 1991; Campolongo *et al.*, 2007). Campolongo *et al.* (2007) ressaltaram a importância deste método e propuseram alterações no mesmo para estender o seu uso para sistemas com múltiplas saídas e para aumentar a sua eficiência através de uma nova estratégia de amostragem.

Os métodos FAST e Sobol são baseados em técnicas de decomposição da variância para quantificar as contribuições das variáveis e das suas interações para o modelo. A principal diferença entre os dois métodos está no algoritmo utilizado para integração multidimensional dos índices de sensibilidade, que utiliza uma função senoidal no método FAST e uma integração de Monte Carlo no método de Sobol. Ambos os métodos não exigem relação entrada-saída linear e monótona, são robustos e exploram todo a extensão do conjunto de entrada, mas tem alta demanda computacional (Helton e Davis, 2003; Zhang *et al.*, 2015).

Zhang *et al.* (2015) compararam cinco métodos comumente usados para a análise de sensibilidade global, utilizando diferentes critérios, e concluíram que o método de Sobol é um dos mais poderosos, por apresentar independência de modelo, capacidade de decompor a variância de saída e por apresentar elevada ordem de interação entre os parâmetros.

Na Figura 2.3, tem-se o algoritmo simplificado do método de Sobol para a análise de sensibilidade. O primeiro passo do algoritmo consiste na definição dos limites dos parâmetros, referindo-se ao intervalo que compreende todos os valores que os parâmetros podem assumir na análise. A execução do método inicia com a geração dos conjuntos de variáveis de entrada utilizando a sequência de Sobol de baixa discrepância, caracterizada por gerar números mais uniformes (Zhang *et al.*, 2015). Os conjuntos gerados são então executados no modelo em questão para posterior determinação dos índices de sensibilidade de Sobol, os quais fornecem informação quantitativa com relação à sensibilidade individual dos parâmetros e também com relação à sensibilidade devido às interações entre os parâmetros.

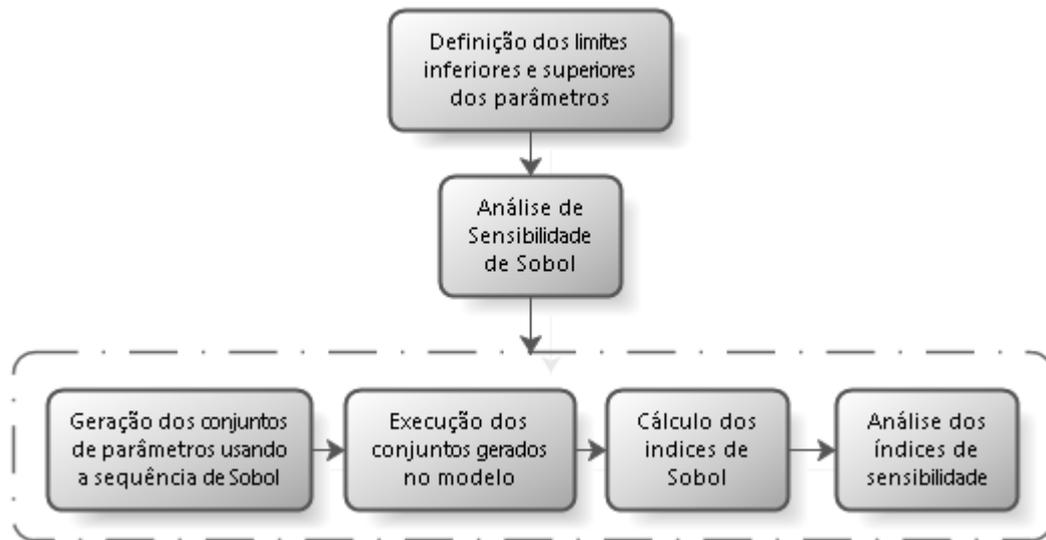


Figura 2.3. Algoritmo simplificado do método de Sobol (adaptado de Zhang *et al.*(2015)).

Seja y uma função descrita como na Equação 2.6, a variância de y , $\text{Var}(y)$, pode ser decomposta como segue na Equação 2.12 (Helton e Davis, 2003; Zhang *et al.*, 2015).

$$\text{Var}(y) = \sum_{i=1}^n \text{Var}_i + \sum_{i<j} \text{Var}_{ij} + \sum_{i<j<k} \text{Var}_{ijk} + \dots + \text{Var}_{1,2,\dots,k} \quad (2.12)$$

onde: Var_i corresponde à parte de $\text{Var}(y)$ devido somente à x_i , Var_{ij} é a parte de $\text{Var}(y)$ devido à interação de x_i e x_j e assim por diante.

A partir da Equação 2.12, podem ser determinados os índices de sensibilidade de Sobol, S , como apresentados nas Equações 2.13, 2.14 e 2.15 (Helton e Davis, 2003; Zhang *et al.*, 2015).

$$S_i = \frac{\text{Var}_i}{\text{Var}(y)} \quad (2.13)$$

$$S_{ij} = \frac{\text{Var}_{ij}}{\text{Var}(y)} \quad (2.14)$$

$$S_{Ti} = S_i + S_{ij_{i \neq j}} + \dots + S_{1\dots i\dots n} \quad (2.15)$$

onde: S_i fornece a contribuição de primeira ordem de x_i para a variável de saída, S_{ij} a contribuição de segunda ordem da interação entre os parâmetros x_i e x_j para a variável de saída e, por fim, S_{Ti} é o índice de sensibilidade de ordem total, o qual quantifica a contribuição total de x_i e das interações de x_i com as outras variáveis para a saída do modelo. Quanto maior o valor destes índices, maior será a influência na saída do modelo.

2.5 Visualização de dados

Após a determinação dos KPIs, a forma como eles serão expostos é um diferencial para a extração de informações dos mesmos, para isto é preciso apresentá-los de uma forma clara e fácil de ser acompanhada. A seguir serão apresentados alguns tipos de gráfico comumente utilizados para tanto. Na Figura 2.4, tem-se a maneira mais simples de apresentar os KPIs, que é utilizando o gráfico de dispersão ou *scatter plot*.

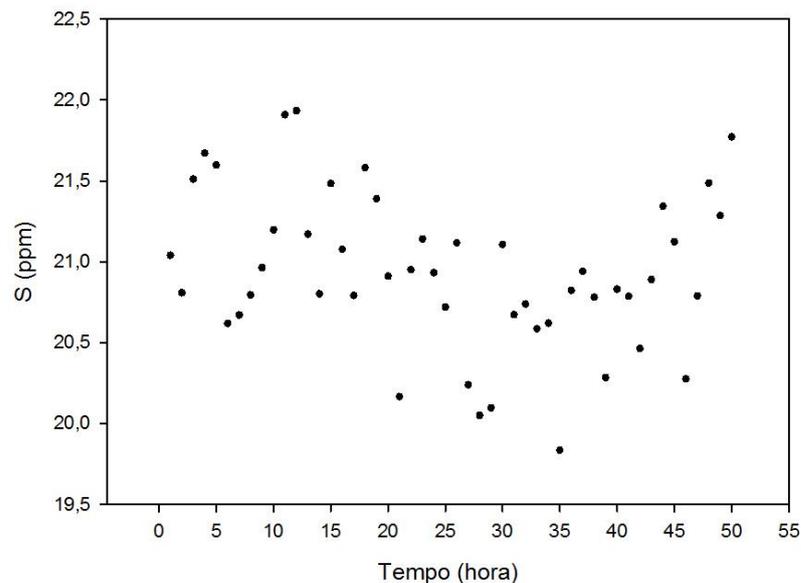


Figura 2.4. Gráfico de dispersão da concentração de enxofre na gasolina.

Como pode ser visto no gráfico de dispersão da Figura 2.4, os dados correspondentes à concentração de enxofre na gasolina são dispostos em função do tempo, neste caso a cada hora, podendo ser expressos como valores instantâneos ou médios e sendo representados por marcadores. Este gráfico possui a versão com linhas para dados contínuos para facilitar a visualização da tendência do indicador, como mostrado na Figura 2.5.

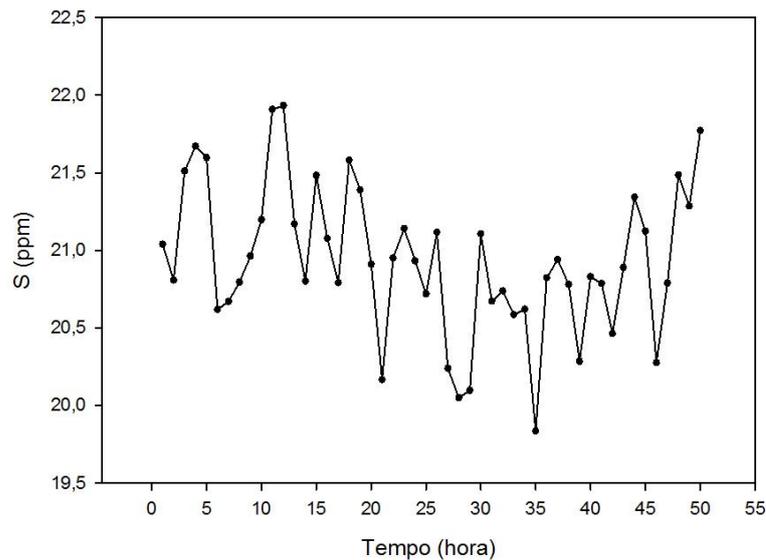


Figura 2.5. Gráfico de dispersão com linha da concentração de enxofre na gasolina.

Como o próprio nome do gráfico sugere, a visualização do mesmo permite saber apenas qualitativamente o quão dispersos os dados estão ao longo do tempo. Uma forma de quantificar esta dispersão é utilizando gráficos de dispersão com barras de erros, agrupando os dados da Figura 2.4 em janelas de 5 horas, como pode ser visto na Figura 2.6.

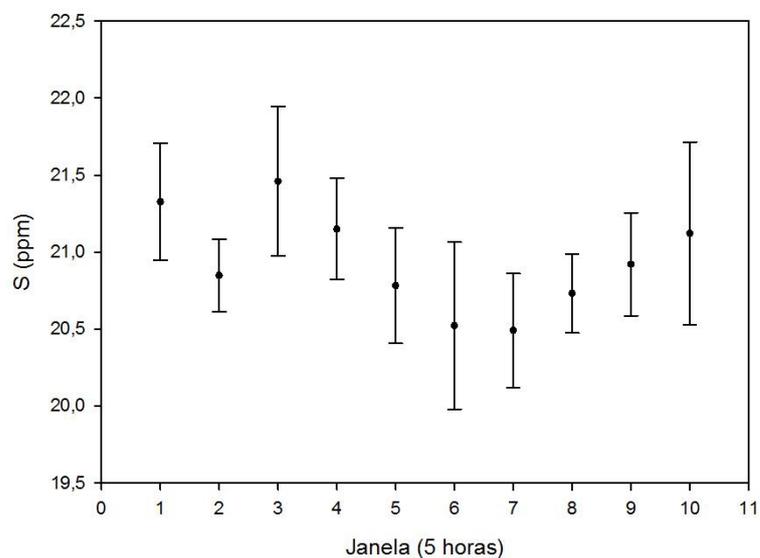


Figura 2.6. Gráfico de dispersão com barras de erro da concentração de enxofre na gasolina.

No gráfico da Figura 2.6, os dados são expressos como valores médios de cada janela em conjunto com os seus respectivos desvios-padrões. Neste caso, os dados apresentados anteriormente na Figura 2.4 foram agrupados em janelas de cinco horas. É uma prática comum se aglutinar dados relativos a um determinado período para se fazer

análises. O desvio-padrão de cada janela é indicado pelas barras de erro; quanto maior o tamanho destas barras, maior a variabilidade do indicador na janela em questão.

Outra ferramenta que pode ser utilizada para o acompanhamento de KPIs é o diagrama de caixa ou *boxplot*, apresentado na Figura 2.7 com os dados da Figura 2.4 agrupados em janelas de 10 horas.

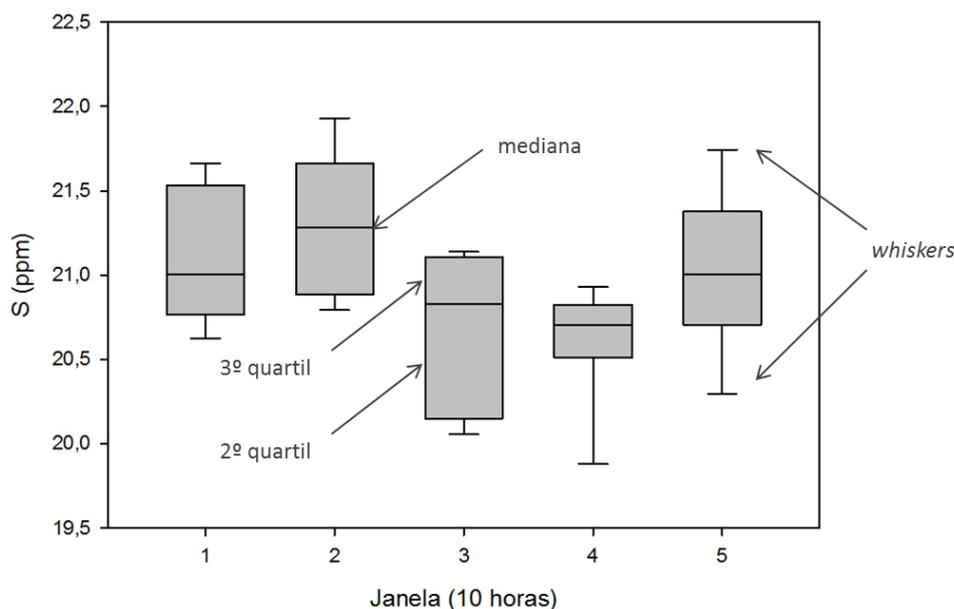


Figura 2.7. Diagrama de caixa da concentração de enxofre na gasolina.

No diagrama de caixa da Figura 2.7, cada janela de 10 horas está representada por uma caixa, que é dividida em termos de quartis da distribuição. As barras, também conhecidas como *whiskers*, representam os extremos da distribuição e a caixa está dividida em duas partes: segundo e terceiro quartil, que representam 50% dos dados e são separados pela mediana, que é o valor central dos dados. A principal aplicação deste tipo de gráfico é na identificação da forma da distribuição e possíveis *outliers*, que são dados considerados atípicos e se situam acima ou abaixo, respectivamente, do limite superior ou inferior (Abuzaid *et al.*, 2012; Bruffaerts *et al.*, 2014). Ferreira *et al.* (2016) fornecem maiores detalhes de como proceder para a construção do diagrama de caixa, bem como a determinação de *outliers*.

Um gráfico bastante popular no mercado financeiro de ações entre negociadores, sobretudo no acompanhamento da cotação de moedas, é o gráfico *candlestick*. Devido à alta variação dos preços, é importante saber qual o melhor momento para comprar ou

vender ações visando um elevado retorno de investimento (Lee e Jo, 1999; Kurita, 2014). Na Figura 2.8, tem-se um exemplo de um *candlestick* para o dólar com valores fictícios.



Figura 2.8. Gráfico *candlestick* do dólar (valores fictícios).

Na Figura 2.8, cada *candlestick* representa um dia e fornece quatro informações dos valores assumidos pelo dólar, a saber: abertura (O), máximo (H), mínimo (L) e fechamento (C). Os valores extremos do *candlestick* correspondem aos valores mínimo e máximo. O corpo do *candlestick* é dado pelo intervalo entre os valores de abertura e fechamento e a sua cor está relacionada com a tendência da moeda naquele dia, como mostrado na Figura 2.9.

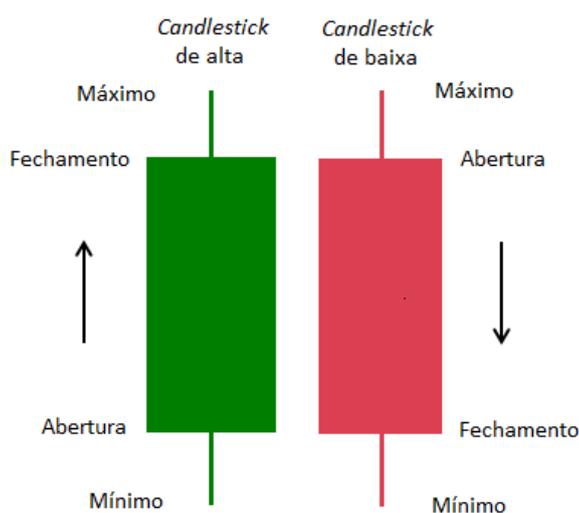


Figura 2.9. Interpretação do *candlestick*.

De acordo com a Figura 2.9, se o indicador tem tendência de crescimento, ou seja, fechou em alta (fechamento > abertura), o mesmo é representado com a cor verde (ou branca), do contrário, quando o dólar fecha o dia em baixa (abertura > fechamento), o

corpo é da cor vermelha (ou preta). Adicionalmente, neste gráfico têm-se duas sequências possíveis para cada *candlestick*: abertura-máximo-mínimo-fechamento ou abertura-mínimo-máximo-fechamento (Marshall *et al.*, 2006).

A partir do gráfico *candlestick*, os investidores podem detectar padrões que podem ser utilizados para prever movimentos em curto prazo dos preços do mercado. Baseado na diferença entre o máximo e o mínimo, sugere-se o grau de tendência e a volatilidade intradiária nos movimentos dos preços. Já entre o mínimo e o fechamento, tem-se a direção em que os preços se movem até o horário de fechamento do dia. Estas relações podem conter informações correlacionadas com o movimento do preço central do dia presente para o dia seguinte, podendo conter evidências da tendência interdía (Kurita, 2014).

Marshall *et al.* (2006) realizaram análises independentes e rigorosas a respeito de estratégias de negócios, a partir de gráficos de *candlestick*, para auxiliar na interpretação dos resultados. Os autores apresentam uma descrição detalhada de *candlesticks* simples (diários) e de padrões observados numa análise, contendo dados de um período de onze anos do Índice Industrial Dow Jones (DJIA, do inglês *Dow Jones Industrial Index*), identificando *candlesticks* de continuação ou reversão, indicando, respectivamente, a predominância ou a mudança de tendência, podendo ter variações de alta (*bullish*) e de baixa (*bearish*). Na Tabela 2.1, encontram-se alguns exemplos de *candlesticks* simples com suas respectivas descrições.

Tabela 2.1: Descrições de *candlesticks* simples (adaptado de Marshall *et al.* (2006)).

| <i>Candlestick</i> | Descrição |
|---|--|
|  | Fechamento bem acima da abertura e próximo do máximo, indicando sentimento positivo de que o preço pode subir no futuro. |
|  | <i>Candlestick</i> muito forte devido à subida de preço no dia e fechamento no máximo. |



Fechamento no máximo do dia, similar ao caso anterior.



Abertura no mínimo. Os preços sobem ao longo do dia, mas não fecham no máximo.



Abertura e fechamento no máximo. Os preços caem durante o dia, mas fecham próximo ao valor de abertura.



Ambos possuem tendência de alta. No *candlestick* verde, os preços caem, mas o fechamento é maior do que a abertura. No vermelho, os preços caem ao longo do dia, mas fecham acima do mínimo (único *candlestick* vermelho de alta).



Fechamento bem abaixo da abertura e próximo do mínimo do dia, indicando sentimento negativo de que o preço pode cair no futuro.



Candlestick muito fraco devido à queda de preço no dia e fechamento no mínimo.



Fechamento no mínimo, similar ao caso anterior.



Abertura no máximo e fechamento próximo ao valor mínimo.



Abertura e fechamento no mínimo. Os preços sobem durante o dia, mas fecham próximo ao valor de abertura.



Ambos possuem tendência de queda. No *candlestick* verde, os preços sobem, voltam a cair, mas o fechamento é maior do que a abertura (único *candlestick* verde de baixa). No vermelho, os preços sobem ao longo do dia, mas fecham abaixo do valor de abertura.

A utilização da ferramenta gráfica *candlestick* é de grande importância no auxílio da tomada de decisões a respeito do mercado de ações, através da identificação de tendências a partir de *candlesticks* simples e padrões de *candlesticks*. A extensão de aplicação deste gráfico para ser utilizado com outros indicadores poderá fornecer informações adicionais sobre o comportamento dos mesmos, auxiliando no diagnóstico de problemas. Não foi encontrado na literatura nenhum trabalho que usa *candlesticks* na engenharia química e no acompanhamento de KPIs.

Capítulo 3 – Metodologia

Neste capítulo, serão detalhadas as etapas seguidas para o desenvolvimento da metodologia empregada no presente trabalho para quantificação e acompanhamento de KPIs. Na Figura 3.1, encontra-se um fluxograma representativo desta metodologia.

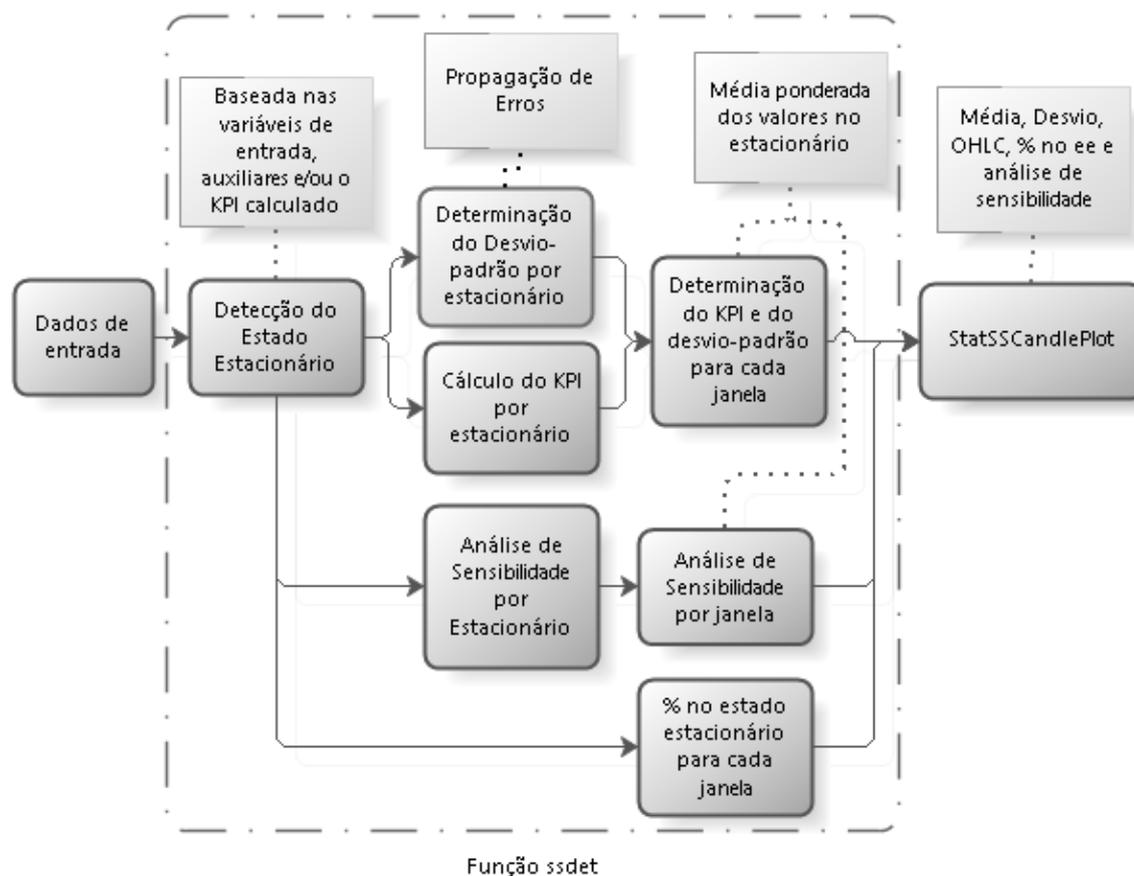


Figura 3.1. Fluxograma representativo da metodologia utilizada no trabalho.

Os dados brutos de entrada, como pode ser visto na Figura 3.1, foram processados na função *ssdet*, na qual foram inicialmente divididos em janelas e submetidos à detecção do estado estacionário. Após a determinação dos intervalos que estão no estado estacionário em cada janela, foi determinado o percentual de cada janela no estacionário, assim como o valor do KPI, seu desvio-padrão através da propagação de erros e dos índices de sensibilidade para cada estacionário. Para determinar os valores por janela, foi calculada a média ponderada dos valores no estacionário do KPI, desvio-padrão e índices de sensibilidade. Como mostrado na Figura 3.1, estes dados serviram de base para a construção do *StatSSCandlePlot*, a ferramenta gráfica de acompanhamento de KPIs proposta neste estudo.

A seguir, estão detalhadas as etapas sumarizadas acima.

3.1 Função *ssdet*

A função *ssdet* foi construída em *Python 3* e seus parâmetros de entrada estão apresentados na Equação 3.1.

$$ssdet = f(n_{var}, dados, w, x_{est}, I_{ee}, N, wavfunc) \quad (3.1)$$

onde: n_{var} é o número de variáveis de entrada, *dados* é um DataFrame composto pelos valores das variáveis de entrada, variáveis auxiliares e o KPI calculado em cada ponto, w é o tamanho da janela, x_{est} é o critério utilizado para a detecção do estado estacionário, I_{ee} é o limiar de estacionariedade de projeto (discutido na Seção 2.2.2), N é o nível de decomposição das *Wavelets* e *wavfunc* é a função de *Wavelet* escolhida.

O primeiro passo da função *ssdet* é a divisão dos dados em janelas, conforme o número de dados por janela, w , informado pelo operador. Após esta divisão, os dados são submetidos à detecção do estado estacionário, como descrito na Seção 3.1.1.

3.1.1 Remoção de ruído e detecção do estado estacionário

O método escolhido para a detecção no estado estacionário foi DER (Seção 2.2.2) com as derivadas estimadas pelo método DWT (Seção 2.2.1), que também é responsável pela remoção do ruído branco dos dados. Inicialmente, as derivadas suavizadas foram estimadas pela Equação 2.3 e, em seguida, utilizou-se a Equação 2.5 para determinar o

índice de estacionariedade, $I_{dif,i}$, que foi posteriormente comparado com o I_{ee} para verificar o estado da janela em questão. Quatro parâmetros de entrada da função $ssdet$ estão diretamente relacionados com estes métodos, a saber: x_{est} , I_{ee} , N e $wavfunc$.

O parâmetro x_{est} da função $ssdet$ fornece o critério de detecção do estacionário, ou seja, o operador pode escolher uma ou mais variáveis para ser o critério, que pode ser composto por variáveis de entrada, auxiliares ou o próprio KPI. As variáveis auxiliares são aquelas que não fazem parte da expressão do KPI, mas podem ser de extrema importância para o processo, para caracterizá-lo como estacionário. No caso de escolher apenas uma variável, a detecção é baseada nesta e somente os trechos em que esta estiver no estacionário são considerados para as etapas subsequentes. Em caso de mais de uma variável, somente os trechos em que todas as variáveis possuem em comum no estacionário são considerados para as análises posteriores.

O I_{ee} é um valor entre 0 e 1 fornecido pelo usuário e está relacionado com a rigorosidade da detecção do estado estacionário. No caso do parâmetro N , uma decomposição em um nível muito elevado pode resultar em um efeito contrário, cancelando fortemente os ruídos e descaracterizando o sinal, logo, é preciso ajustá-lo com cautela. Para uma dimensão de x_{est} maior do que 1 ($\dim(x_{est}) > 1$), podem ser escolhidos valores de I_{ee} e N correspondentes a cada variável.

Por fim, o parâmetro $wavfunc$ determina qual a função de *Wavelet* a ser utilizada na estimação das derivadas e remoção de ruídos. Neste caso, o operador pode optar entre duas funções, a saber: Haar ou spline quadrática. Em caso de $\dim(x_{est}) > 1$, também podem ser escolhidas diferentes funções para cada variável.

A partir destes parâmetros, o método é executado e para cada janela são determinados os trechos no estado estacionário. Em seguida, são calculados os percentuais de dados no estado estacionário ($\%EE$), fornecendo uma noção de estabilidade da planta em cada janela. O $\%EE$ é dado pela Equação 3.2.

$$\%EE = \frac{\text{n}^\circ \text{ de dados no EE}}{\text{n}^\circ \text{ de dados na janela}} \cdot 100 \quad (3.2)$$

De posse dos intervalos no EE a serem considerados para cada janela, prossegue-se com o estudo da propagação de erros para a determinação do KPI e do desvio-padrão por estacionário e por janela.

3.1.2 Propagação de erros

Para o estudo da propagação de erros utilizou-se o método de expansão em séries de Taylor de 1ª ordem com a correção da não-normalidade e o cálculo da covariância, como apresentado na Seção 2.3, para determinar a variância por estacionário. Considerando um KPI com n variáveis de entrada, cuja janela i possui k regiões em estado estacionário, a sequência de cálculos utilizada até a determinação do desvio-padrão da janela i é realizada como segue na Figura 3.2.

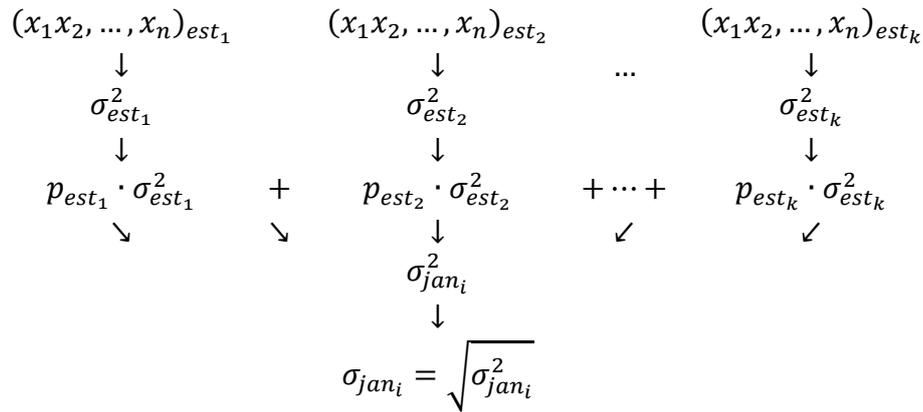


Figura 3.2. Esquema utilizado para a determinação do desvio-padrão da janela i .

Na Figura 3.2, a partir das variáveis de entrada correspondentes ao estacionário e da equação do KPI para a determinação das derivadas, utiliza-se a Equação 2.11 para calcular a variância de cada estacionário, $\sigma_{est_j}^2$ ($j \in [1, k]$). A variância da janela i , $\sigma_{jan_i}^2$, é determinada a partir da média ponderada das variâncias por estacionário, cujo peso, p_{est_j} , é determinado segundo a Equação 3.3.

$$p_{est_j} = \frac{t_{est_j}}{\sum_{j=1}^k t_{est_j}} \quad (3.3)$$

onde: t_{est_j} corresponde ao período de tempo do estacionário j .

Para calcular o KPI médio da janela, inicialmente determina-se o KPI para cada estacionário j , de acordo com a Equação 3.4.

$$KPI_{est_j} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)_{est_j} \quad (3.4)$$

A partir dos valores obtidos das Equações 3.3 e 3.4 para todos os estacionários, utiliza-se a Equação 3.5 para determinar o indicador médio da janela i , KPI_{jan_i} .

$$KPI_{jan_i} = p_{est_1} \cdot KPI_{est_1} + p_{est_2} \cdot KPI_{est_2} + \dots + p_{est_k} \cdot KPI_{est_k} \quad (3.5)$$

3.1.3 Análise de Sensibilidade

O método de Sobol, descrito na Seção 2.4, é utilizado para fazer a análise de sensibilidade. Este método está disponível na biblioteca *SALib* (*Sensitivity Analysis Library*) para o *Python* e foi executado conforme as etapas em seguida:

- (1) Definição do problema: nesta primeira etapa do método, define-se um conjunto de parâmetros (denominado dicionário no *Python*) contendo o número de variáveis do KPI, suas respectivas notações e seus limites inferior e superior, denominados por *'num_vars'*, *'names'* e *'bounds'* na biblioteca;
- (2) Geração das amostras das entradas do modelo: a partir do problema definido, são geradas as variáveis de entrada do KPI através da extensão de Saltelli para a sequência de Sobol, que reduz as taxas de erro no cálculo dos índices de sensibilidade. A função utilizada no *Python* é a *saltelli.sample*, que tem como entradas o problema, o número de amostras a serem geradas, N_A , e a variável booleana *'calc_second_order'* para o cálculo dos índices de segunda ordem (Equação 2.13). Para um KPI com n_{var} variáveis, esta função gera uma matriz com $N_A \cdot (n_{var} + 2)$ linhas, se os índices de segunda ordem não forem calculados (*'calc_second_order' = False*), e com $N_A \cdot (2n_{var} + 2)$ linhas se os índices de segunda ordem forem calculados (*'calc_second_order' = True*). De acordo com Zhang *et al.* (2015), para um modelo de baixa complexidade, $N_A = 1.000$ amostras é suficiente, para modelos mais complexos, com cerca de 20 parâmetros, no mínimo deve-se ter $N_A = 100.000$ amostras;
- (3) Cálculo do KPI: o KPI é calculado em cada ponto da matriz resultante da função *saltelli.sample*, na qual cada coluna corresponde aos valores de uma variável, respeitando a ordem de declaração das variáveis na etapa (1);
- (4) Cálculo dos índices de sensibilidade por estacionário: nesta etapa, utiliza-se a função *sobol.analyze* para realizar a análise de sensibilidade nas saídas do modelo.

A partir do problema e dos valores calculados do KPI, esta função retorna um dicionário com os índices de sensibilidade de primeira, segunda (para *'calc_second_order' = True*) e total ordem, como descrito nas Equações 2.13, 2.14 e 2.15, assim como suas respectivas incertezas;

- (5) Cálculo dos índices de sensibilidade por janela: a partir dos índices de sensibilidade de ordem total calculados na etapa (4) para cada estacionário, $S_{T_{est_j}}$, prossegue-se da mesma forma para determinar os índices de cada variável por janela, $S_{Tx_{jan_i}}$, a partir do esquema apresentado na Figura 3.2 e da Equação 3.3.
- (6) Avaliação da sensibilidade por janela: quanto maior o $S_{Tx_{jan_i}}$, maior é a sensibilidade do indicador à variável x , desta forma, determina-se qual a variável com maior responsabilidade pela variação do KPI na janela i .

3.2 StatSSCandlePlot

Visando uma nova abordagem para o acompanhamento de KPIs, integrando dados provenientes de análises estatísticas, da detecção do estado estacionário e o gráfico *candlestick*, se propõe neste trabalho o *StatSSCandlePlot*. Esta ferramenta gráfica fornece diretrizes sobre a tendência do indicador através do padrão *candlestick*, apresenta valores consistentes do KPI empregando a estatística da propagação de erros aliada ao estudo da estabilidade do indicador pela detecção do estado estacionário. Adicionalmente, fornece um diagnóstico inicial da causa raiz da variação do indicador através da análise de sensibilidade. Tudo isto apresentado de uma forma interativa com a ferramenta *Plotly* para o *Python*.

3.2.1 Descrição do plot

O *StatSSCandlePlot* contém as seguintes informações:

- (1) Candlestick: informação sobre todos os valores assumidos pelo KPI na janela, destacando-se o mínimo e o máximo, assim como os valores de abertura e fechamento. A partir destas informações é possível saber a tendência do indicador na janela, crescente ou decrescente;

- (2) Percentual no estado estacionário (%EE) (gráfico de dispersão): indicativo da dinamicidade do indicador em cada janela, informação importante para saber o quanto da janela é considerado para o cálculo do KPI, do desvio-padrão e dos índices de sensibilidade;
- (3) Média e desvio-padrão (gráfico de linha): baseados na propagação de erros, indicam os valores confiáveis do KPI e seu respectivo desvio levando em consideração apenas os valores assumidos no estado estacionário. A partir da média e do desvio-padrão, determinaram-se os limites inferior e superior, correspondentes a três desvios abaixo e acima da média, respectivamente;
- (4) Referência (linha tracejada): meta do KPI pré-estabelecida informada pelo usuário para identificação do valor desejável para o KPI. Valores acima desta reta são desejáveis se o aumento do KPI favorecer o desempenho;
- (5) Índices de sensibilidade global (gráfico de barras): análise de causa raiz da variação do KPI indicando a influência das variáveis no seu comportamento de forma quantitativa. Estes resultados são apresentados de forma complementar em um gráfico de barras agrupadas para cada janela com as respectivas confianças calculadas pelo método de Sobol.

3.2.2 Exemplo de uso

Na Figura 3.3, tem-se um exemplo do *StatSSCandlePlot* para um determinado KPI.



Figura 3.3. Exemplo do *StatSSCandlePlot*.

Na Figura 3.3, observa-se que o *StatSSCandlePlot* possui três regiões: a principal, que expõe os dados, a legenda, à direita, e a barra de rolagem horizontal na parte inferior. Neste gráfico, estão expostos os *candlesticks*, vermelho para o KPI em baixa e verde para o KPI em alta, o %EE de cada janela, em azul, referindo-se ao eixo da direita, o KPI médio e os limites inferior e superior e, adicionalmente, a linha de referência, representada pela cor laranja. Ao navegar pelo gráfico, os dados por janela são destacados em caixas separadas de acordo com a cor referente a cada dado, como mostrado na Figura 3.3, facilitando, desta forma, a visualização dos valores das variáveis. A partir da barra de rolagem horizontal, é possível focar em um período determinado dos dados e ao mesmo tempo saber onde o mesmo está situado no contexto geral, como mostrado na Figura 3.4.



Figura 3.4. Exemplo do *StatSSCandlePlot* com ênfase em um período.

Na Figura 3.4, ao mover os dois lados da barra de rolagem horizontal, aparecem na janela principal apenas os dados do período escolhido. Esta ferramenta é importante para permitir a visualização em um período específico e facilitar a análise, visto que uma menor quantidade de dados fica exposta. Também é possível mover a barra mantendo um período fixo, por exemplo, de 10 dias, para isto basta ajustar o período desejado movendo os lados da barra e depois clicar no espaço compreendido entre os limites, movendo-o para o sentido desejado.

O *StatSSCandlePlot* é um gráfico que reúne muitas informações, portanto, em alguns momentos isto pode dificultar a visualização de um parâmetro específico. No entanto,

utilizando a ferramenta *Plotly*, é possível desativar alguns dados ao clicar na legenda, como pode ser visto nas Figuras 3.5 e 3.6.



Figura 3.5. Exemplo do *StatSSCandlePlot* ocultando os *candlesticks*.



Figura 3.6. Exemplo do *StatSSCandlePlot* ocultando os *candlesticks* e o *%EE*.

Nas Figuras 3.5 e 3.6, pode-se verificar que algumas variáveis foram ocultadas, reduzindo a quantidade de informações que aparecem na área principal do gráfico, conforme a necessidade do usuário. Para habilitar estas variáveis, é só selecioná-las novamente na legenda.

O *StatSSCandlePlot* é uma ferramenta de grande potencial para o acompanhamento de KPIs que, aliada às interpretações corretas, auxilia positivamente nas decisões acerca de mudanças no processo.

Capítulo 4 – Estudo de Caso

Após o desenvolvimento da metodologia referente ao cálculo e ao acompanhamento de KPIs, a mesma foi aplicada em um estudo de caso. A escolha deste estudo foi baseada na facilidade de entendimento do sistema e representatividade para aplicar a metodologia. O sistema estudado foi um chuveiro composto por duas fontes de aquecimento, gás e energia elétrica, válvulas, um tanque e uma bomba, como ilustrado na Figura 4.1.

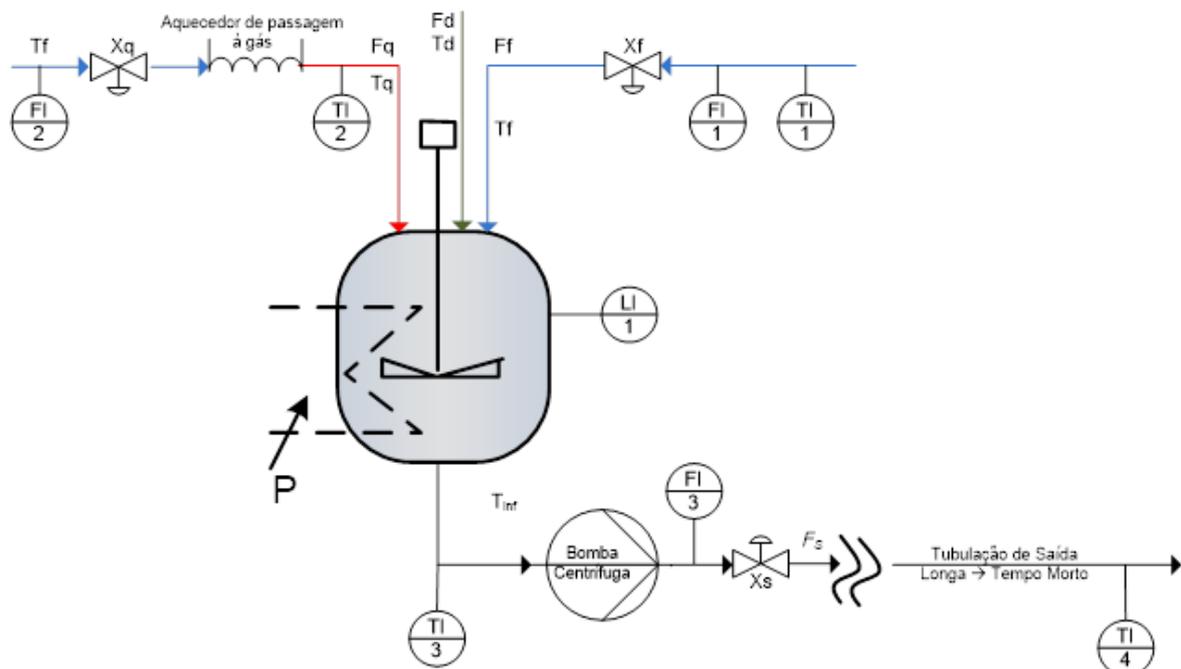


Figura 4.1. Esquema representativo do chuveiro utilizado no estudo de caso.

Na Figura 4.1, os índices subscritos q, f, d, s e inf correspondem, respectivamente, a quente, fria, distúrbio, saída e infinito (ambiente). F corresponde à vazão, T à

temperatura, x é a posição da válvula e P , a potência referente ao aquecimento elétrico. Fls são os medidores de vazão, LI é o medidor de nível do tanque e TIs correspondem aos medidores de temperatura.

A partir de duas correntes oriundas da mesma caixa d'água, uma é denominada de corrente fria e outra de corrente quente, esta última passa por um aquecedor de passagem a gás, como mostrado na Figura 4.1. As duas correntes são misturadas em um tanque de agitação que contém dois aquecedores elétricos, acionados de acordo com o nível do tanque. Visto que a tubulação da saída do tanque possui um longo comprimento, o tempo morto do sistema é elevado, ocasionando a perda de calor para o meio ambiente que afeta a temperatura de saída da planta.

Este chuveiro foi programado no MatLab® R2014a de forma que o tanque não transborde ou esvazie, através do controle do nível do tanque, e também que as resistências não queimem. Como a potência de aquecimento é constante, a mesma funciona no modo liga-desliga, logo, quando o tanque estiver com o nível inferior a 30% da capacidade do tanque, a resistência é desligada. Quando o nível ultrapassa 60%, uma segunda resistência elétrica é ligada para duplicar a capacidade de aquecimento.

Um indicador foi criado para medir a qualidade do banho, ou seja, baseando-se em valores de vazão e temperatura de saída da água, F_S e T_4 , respectivamente, criou-se uma relação empírica denominada de Índice de Qualidade do Banho (IQB). Na Equação 4.1, encontra-se a expressão do IQB para o chuveiro da Figura 4.1.

$$IQB = \frac{1}{e} \cdot e^{\left[1 - \left(\frac{T_4 - 38 + 0,02F_S^2}{2}\right)\right]} \cdot \left[0,506 + \log\left(\log\left(\frac{10000\sqrt{F_S}}{10 + F_S + 0,004 \cdot F_S^4}\right)\right)\right]^{20} \quad (4.1)$$

Na Equação 4.1, pode-se observar que o IQB depende apenas de F_S e T_4 , pois estas são as variáveis que influenciam na satisfação do consumidor. Adicionalmente, observa-se que esta equação envolve operações não lineares e que o seu valor foi normalizado, ou seja, assume apenas valores de 0 a 1. Na Figura 4.2, tem-se a representação gráfica do IQB para F_S variando de 0,1 a 10 L/min e T_4 de 32°C a 44°C.

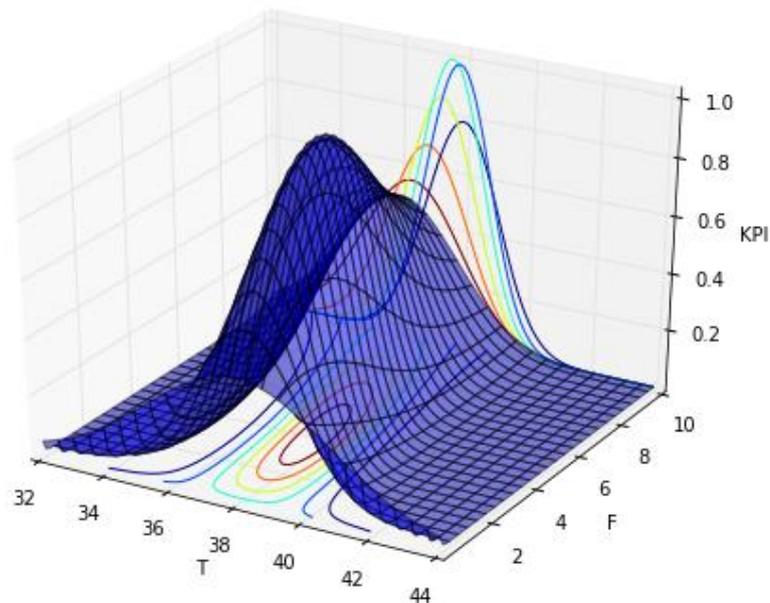


Figura 4.2. Representação gráfica do *IQB*.

Na Figura 4.2, verifica-se que o *IQB* possui um ótimo global ao atingir o valor máximo de 1, que ocorre quando as suas variáveis de entrada são aproximadamente iguais a $3,8 L/min$ para a vazão do banho e $37,5 ^\circ C$ para a temperatura do banho. Observa-se que o KPI em questão é penalizado tanto pelo aumento quanto pela diminuição da vazão, isto acontece porque na prática, ao aumentar demais a vazão, a temperatura do banho não acompanha este aumento, tornando o banho frio. Já ao diminuir demais a vazão, a qualidade do banho também cai pela elevação excessiva da temperatura. O indicador também é afetado negativamente pelo aumento ou diminuição da temperatura, dado que um banho muito frio ou demasiadamente quente não resulta em uma situação confortável.

O *IQB* foi avaliado em três cenários distintos, apresentando: 1) distúrbios na vazão, F_d , 2) queda da temperatura fria, T_f , e 3) avaliando o *IQB* diante de distúrbios na vazão, mas em uma malha com estratégia de controle diferente da utilizada nos dois primeiros cenários. Em todos os três cenários o controlador foi ajustado manualmente. Nos cenários 1 e 2, a potência é constante e funciona no modo liga-desliga de acordo com o nível do tanque, como mencionado anteriormente. Neste caso, o nível do tanque é controlado pela válvula de entrada de água fria e a temperatura de saída é controlada pela válvula de saída, como pode ser visto na Figura 4.3.

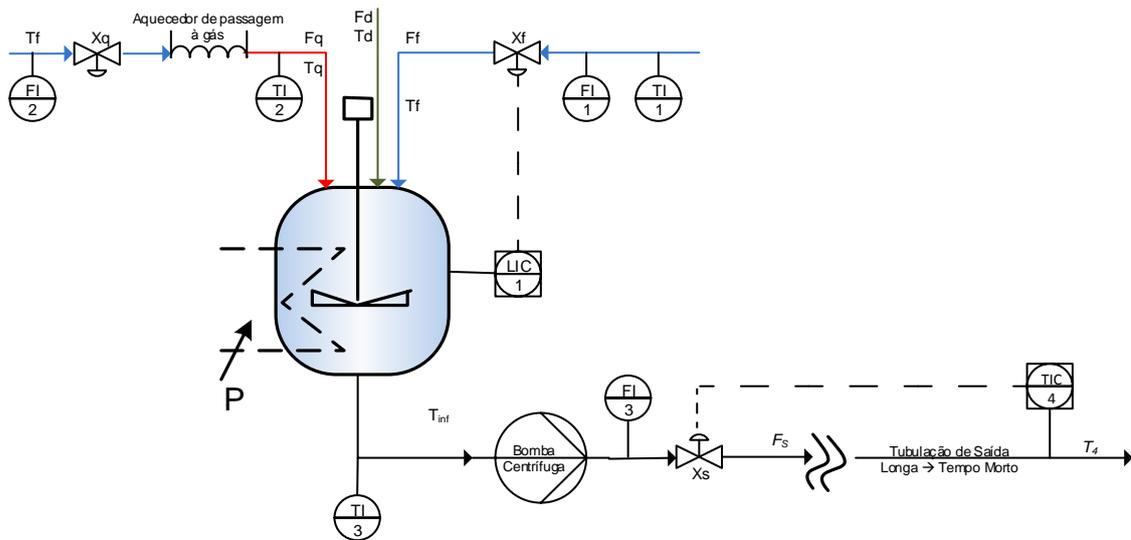


Figura 4.3. Estratégia de controle utilizada nos cenários 1 e 2.

No cenário 3, a estratégia de controle foi alterada, neste caso o nível é controlado pela válvula de saída e a temperatura de saída é controlada pela potência, que é variável. Foi aplicada uma estratégia tipo *split range*, na qual quando a variável manipulada corresponde de 0 a 50 % utiliza-se apenas o aquecimento a gás e de 50 a 100 %, o aquecimento elétrico é ligado. Esta estratégia de controle está apresentada na Figura 4.4.

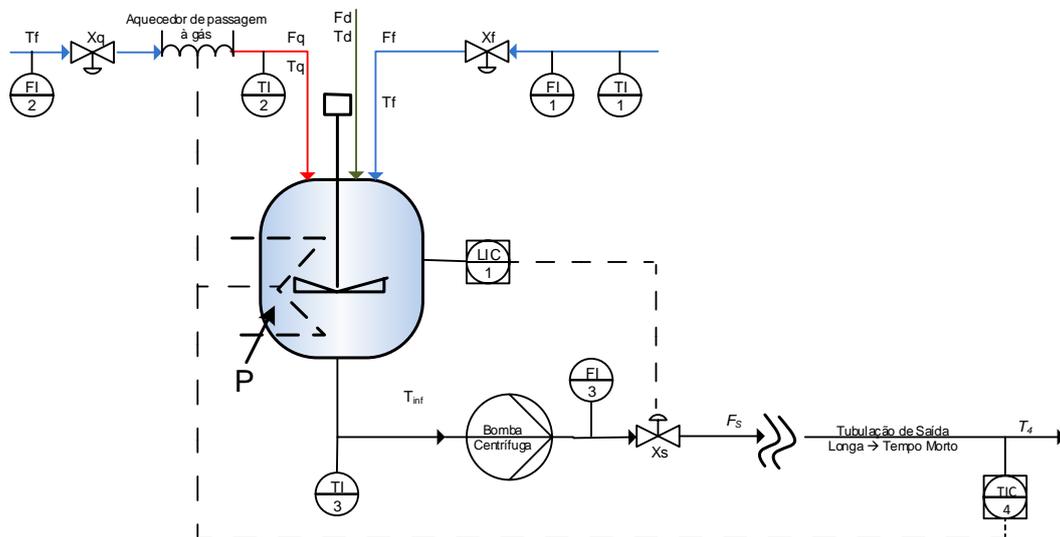


Figura 4.4. Estratégia de controle utilizada no cenário 3.

Foi adicionado um ruído nas variáveis F_s e T_4 para representar um processo real, no qual a presença de flutuações nos valores obtidos é comum. No *Simulink*, o bloco

utilizado para adicionar o ruído a estas variáveis foi o *Band-Limited White Noise*, com uma magnitude de 0,0005 e normalmente distribuído.

Para aplicar a metodologia, foram comparados os resultados obtidos com parâmetros e métodos diferentes para justificar a escolha dos mesmos utilizados nos cenários do estudo de caso. A seguir, será apresentado como foi escolhido o limiar do índice de estacionariedade, I_{ee} , o nível de decomposição de *Wavelets*, N , o método de séries de Taylor de 1ª ordem para a propagação de erros e a escolha do número de amostras a serem geradas, N_A , na análise de sensibilidade.

4.1 Escolha dos parâmetros e métodos

4.1.1 Detecção do estado estacionário e remoção de ruído

Na Figura 4.5, tem-se os gráficos referentes ao IQB , F_s e T_4 para uma situação em que um distúrbio do tipo degrau unitário negativo na vazão, F_d , incide no tempo de 300 segundos para a estratégia de controle da Figura 4.3.

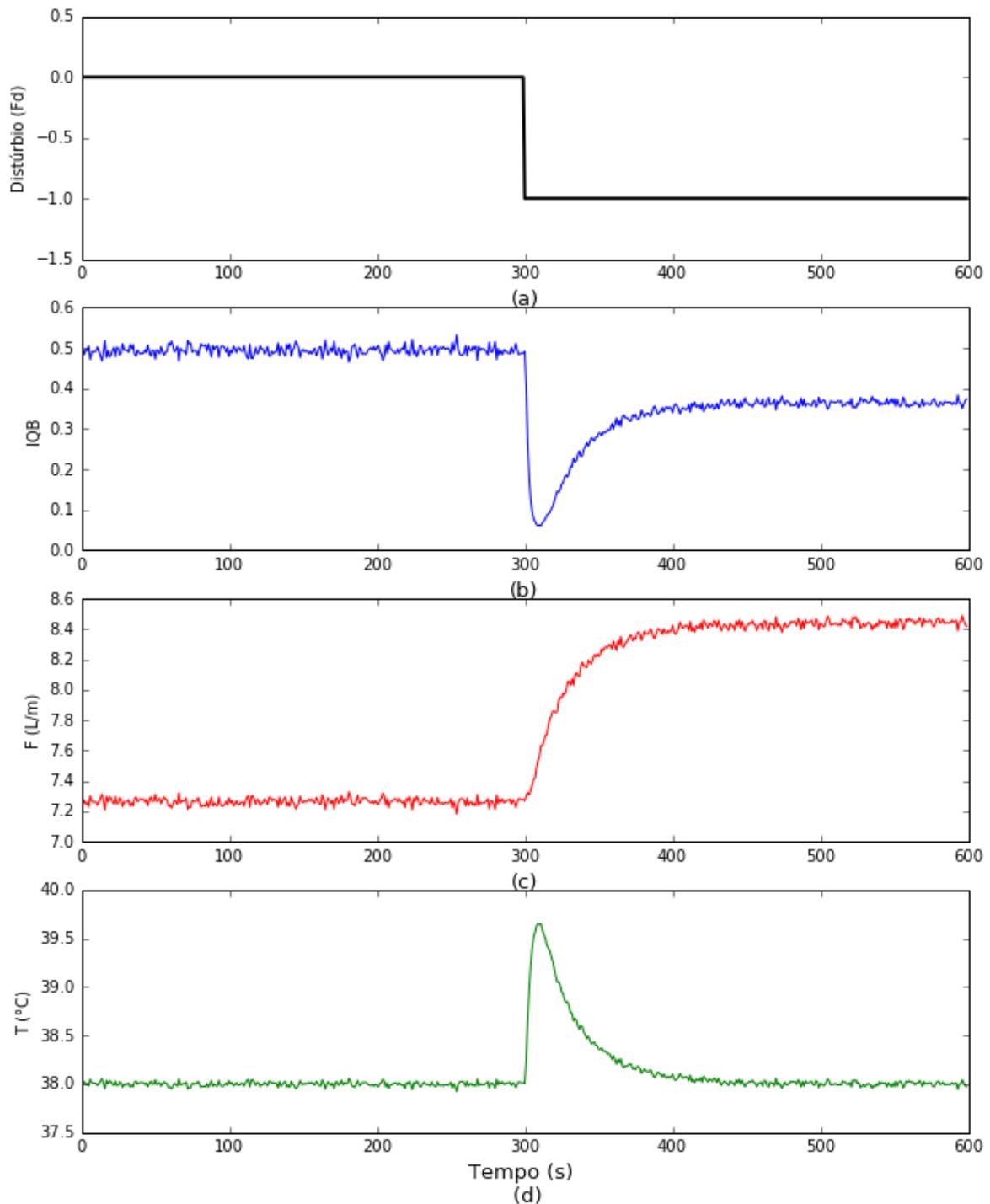


Figura 4.5. (a) Gráfico do distúrbio na vazão em $t = 300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho).

Na (a), observa-se o distúrbio na vazão, que inicialmente é igual a zero, mas a partir do tempo igual a 300 segundos, assume o valor de -1. Este distúrbio provoca uma redução no IQB, como pode ser visto na Figura 4.5(b). O degrau negativo na vazão

também provoca um aumento na vazão de saída como pode ser visto na Figura 4.5(c), causado pelo aumento repentino da temperatura de saída, como mostrado na Figura 4.5(d). Isto aconteceu porque o tanque tem o seu nível reduzido e, com uma menor quantidade de água para ser aquecida, uma mesma potência eleva a temperatura da água. Ao enfrentar uma queda na temperatura de saída, o controlador reage aumentando a vazão de saída, restabelecendo, posteriormente, o equilíbrio em torno da temperatura desejada.

Avaliando-se o período entre $t_1 = 200$ s e $t_2 = 500$ s na , foi utilizado o método DER para a detecção do estado estacionário no IQB para diferentes valores de I_{ee} , fixando $N = 4$, utilizando a função de *Wavelet* spline quadrática e o próprio indicador como critério para a detecção do estado estacionário ($x_{est} = 'IQB'$). Estes cenários estão apresentados a seguir na , na qual a linha verde é a classificação feita pelo método DER e a linha azul é o índice calculado.

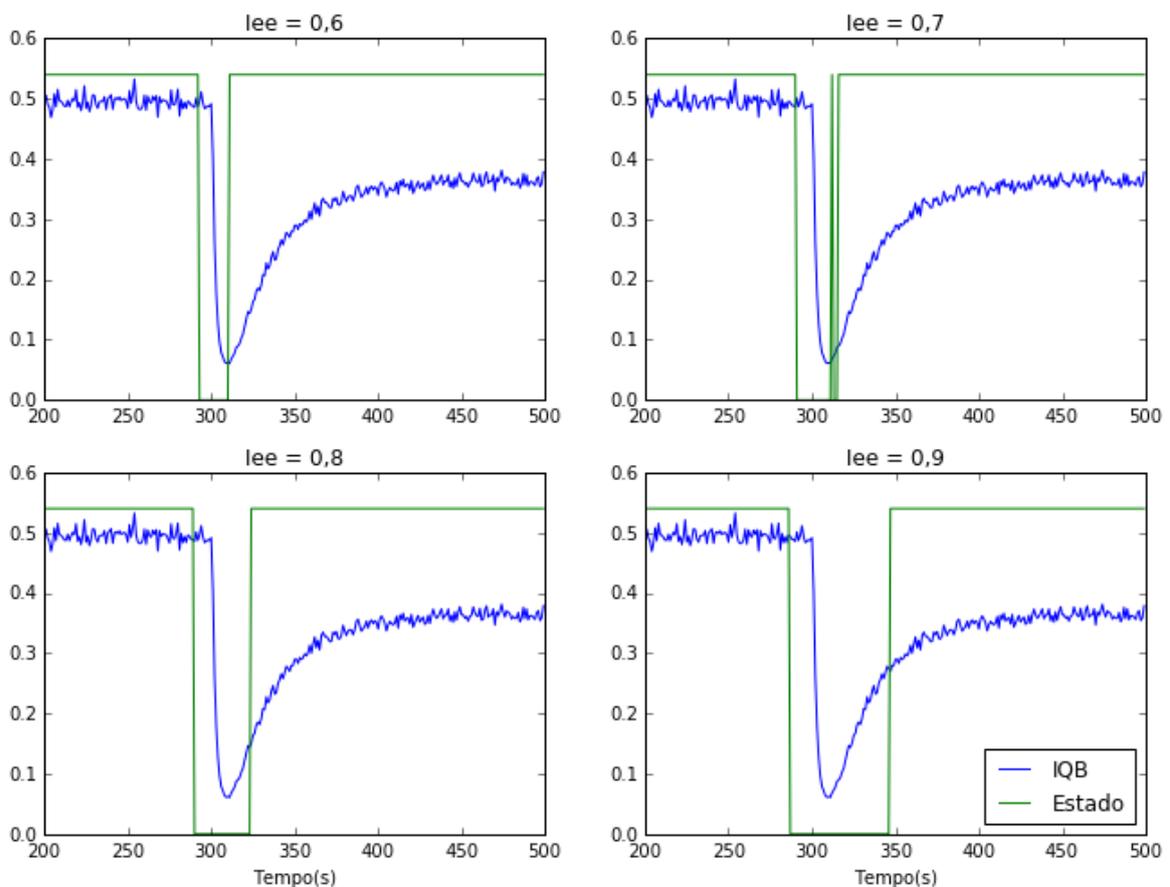


Figura 4.6. Gráficos da identificação do estado estacionário para o IQB com $N = 4$ e diferentes valores de I_{ee} para $t_1 = 200$ s e $t_2 = 500$ s.

Na , o estado foi multiplicado por 0,54 apenas para melhor visualização, assumindo um valor zero quando o EE não é identificado e 0,54 quando está no EE. A partir da , observa-se que quanto maior o I_{ee} , menor a porção da janela no estado estacionário. Em todos os casos, o método identificou a mudança de patamar do indicador e caracterizou esta transição como não estacionária, no entanto, apresenta diferentes amplitudes para a transição de acordo com o I_{ee} , sendo a maior e, conseqüentemente, a mais rigorosa, a faixa obtida quando $I_{ee} = 0,9$.

Utilizando o método DER com I_{ee} fixado em 0,7 e diferentes valores de N para $t_1 = 200$ s e $t_2 = 500$ s na Figura 4.5, foram obtidos os resultados apresentados na Figura 4.7.

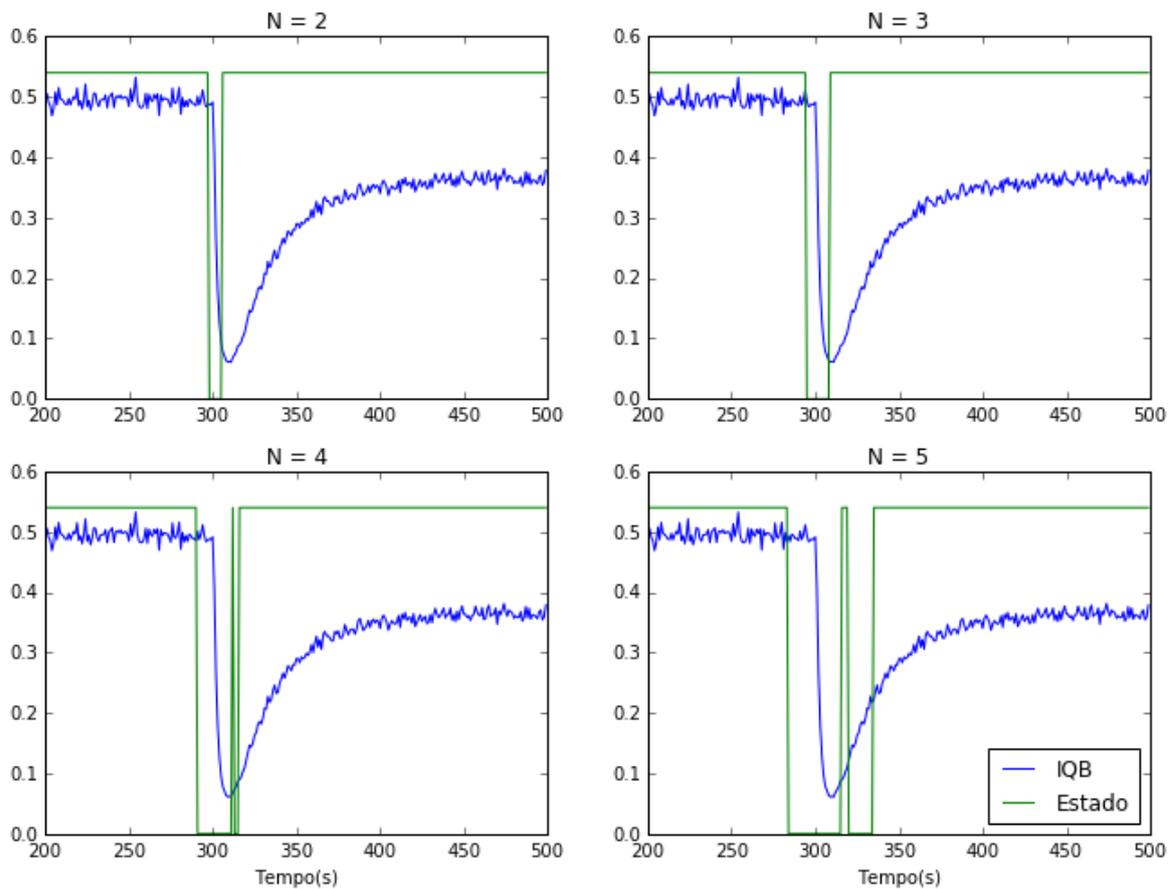


Figura 4.7. Gráficos da identificação do estado estacionário para o IQB com $I_{ee} = 0,7$ e diferentes valores de N para $t_1 = 200$ s e $t_2 = 500$ s.

Na Figura 4.7 observa-se que o aumento do número de níveis de decomposição por *Wavelet* reduz a porção da janela no estado estacionário. À medida que N aumenta, há

uma maior eliminação de ruídos, tornando mais claros os intervalos que podem ser considerados como estacionários.

Na Tabela 4.1, encontra-se o percentual da janela no estado estacionário (%EE) de acordo com diferentes valores de I_{ee} e N .

Tabela 4.1: Percentual da janela no EE de acordo com N e I_{ee} para $t_1 = 200$ s e $t_2 = 500$ s.

| $N \backslash I_{ee}$ | 0,6 | 0,7 | 0,8 | 0,9 |
|-----------------------|------|------|------|------|
| 2 | 98 % | 97 % | 97 % | 90 % |
| 3 | 96 % | 95 % | 95 % | 89 % |
| 4 | 94 % | 92 % | 89 % | 80 % |
| 5 | 88 % | 84 % | 77 % | 69 % |

Comparando-se a com a linha referente ao $N = 4$ na Tabela 4.1, observa-se que o %EE é igual a 80 % para $I_{ee} = 0,9$, aumentando para 94 % quando $I_{ee} = 0,6$. Ao observar a coluna referente ao $I_{ee} = 0,9$, observa-se que o %EE é de 90% quando $N = 2$ e cai para 69 % quando $N = 5$. O melhor cenário para a detecção do estado estacionário, após analisar as Figuras 4.6 e 4.7 em conjunto com a Tabela 4.1, seria com $N = 5$ e $I_{ee} = 0,9$, pois desconsidera os valores muito próximos da transição de patamar, resultando em uma escolha rigorosa e altamente confiável com respeito ao EE. No entanto, como os dados podem apresentar diferentes comportamentos em janelas distintas, nem sempre a melhor escolha é a mais rigorosa. Na Figura 4.8, tem-se a identificação do EE na janela do período entre $t_1 = 0$ s e $t_2 = 300$ s da considerando $N = 4$.

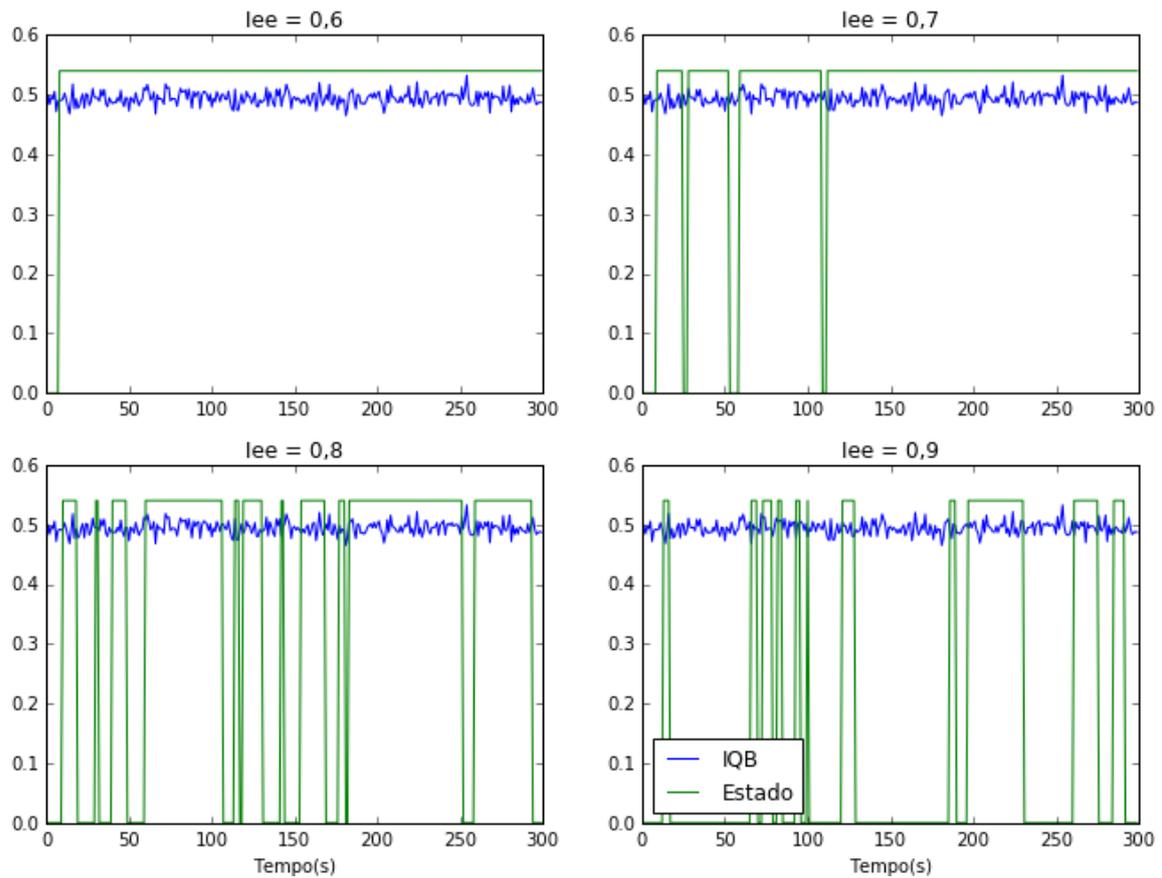


Figura 4.8. Gráficos da identificação do estado estacionário para o *IQB* com $N = 4$ e diferentes valores de I_{ee} para $t_1 = 0$ s e $t_2 = 300$ s.

Na Figura 4.8, observa-se que os valores do *IQB* flutuam em torno de um valor médio, não apresentando mudanças bruscas de valores. Neste caso, a detecção do EE é mais sensível, reduzindo a porção da janela no estado estacionário. O mesmo pode ser observado na Figura 4.9, na qual o I_{ee} está fixado em 0,7 e o N é variável.

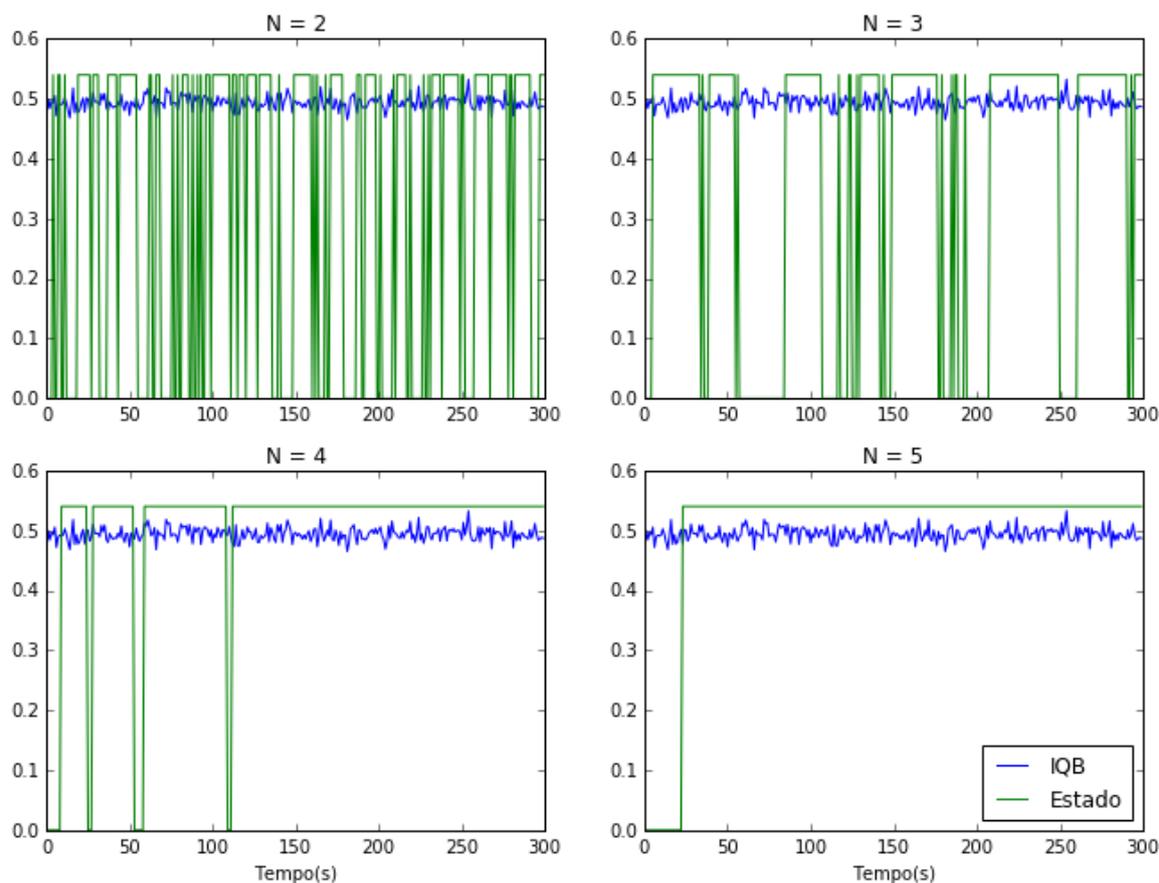


Figura 4.9. Gráficos da identificação do estado estacionário para o IQB com $I_{ee} = 0,7$ e diferentes valores de N para $t_1 = 0$ s e $t_2 = 300$ s.

Na Tabela 4.2, tem-se o %EE de acordo com os diferentes valores de I_{ee} e N para a faixa de tempo entre zero e 300 segundos.

Tabela 4.2: Percentual da janela no EE de acordo com N e I_{ee} para $t_1 = 0$ s e $t_2 = 300$ s.

| $N \backslash I_{ee}$ | 0,6 | 0,7 | 0,8 | 0,9 |
|-----------------------|------|------|------|------|
| 2 | 77 % | 59 % | 36 % | 13 % |
| 3 | 83 % | 66 % | 43 % | 15 % |
| 4 | 98 % | 93 % | 69 % | 30 % |
| 5 | 94 % | 92 % | 77 % | 28 % |

Na Tabela 4.2, observa-se que ao escolher o melhor cenário ($N = 5$ e $I_{ee} = 0,9$) resultaria em uma janela com apenas 28 % no EE, podendo chegar a valores menores ou até mesmo iguais a zero. Baseando-se nesta análise preliminar, foram adotados os valores intermediários de 0,6 para o I_{ee} e de 4 para o N , com a finalidade de representar

a janela com um maior número de valores e de evitar que algumas janelas fiquem sem um valor representativo por falta de intervalos no EE, sendo preferível valores próximos ao reduzir o I_{ee} .

4.1.2 Método de expansão de séries de Taylor de 1ª ordem

O método baseado na expansão de séries de Taylor de 1ª ordem foi comparado com o Método de Monte Carlo para verificar se o mesmo é adequado para o estudo de caso. Para F_s e T_4 no período de tempo entre 0 e 200 segundos nas Figuras 4.5(c) e 4.5(d), respectivamente, inicialmente foi verificado se a distribuição de ambas pode ser considerada Normal, para posterior geração de números aleatórios. Nas Figuras 4.10 e 4.11, encontram-se o histograma e o gráfico de probabilidade para F_s e T_4 , respectivamente.

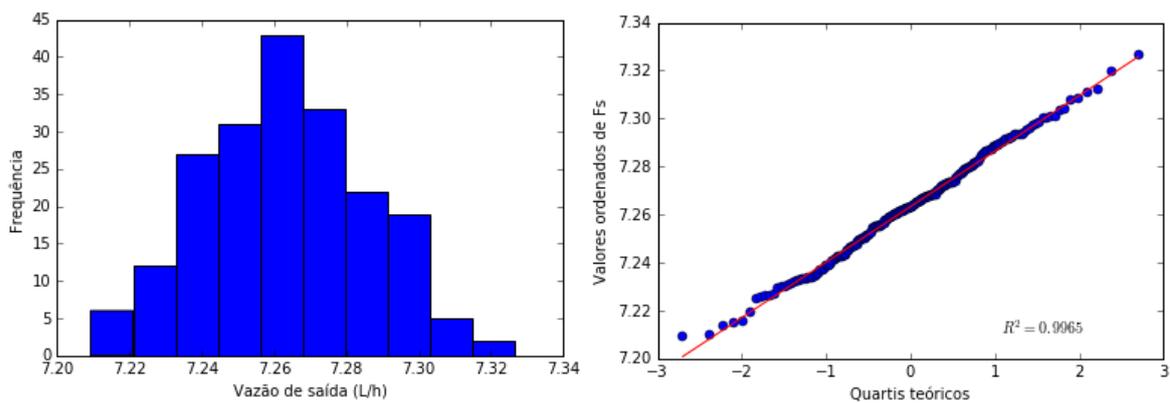


Figura 4.10. Histograma e gráfico de probabilidade de F_s (vazão da água do banho) entre 0 e 200 s.

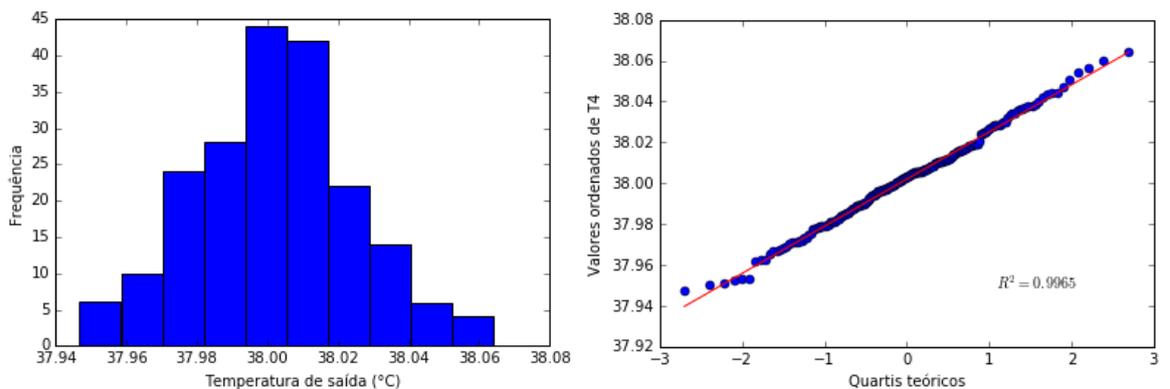


Figura 4.11. Histograma e gráfico de probabilidade de T_4 (temperatura do banho) entre 0 e 200 s.

A partir das Figuras 4.10 e 4.11, observa-se que ambas as variáveis podem ser consideradas com distribuição Gaussiana no intervalo em questão, visto que nos gráficos de probabilidade os dados apresentaram um bom ajuste ao da distribuição Normal, dispersando-se minimamente apenas nos valores extremos. Para corroborar o que as Figuras 4.10 e 4.11 indicam, foi utilizado o teste de normalidade de Shapiro-Wilk no *Python*, no qual a hipótese nula de que os dados são provenientes de uma distribuição Normal não pode ser rejeitada nos dois casos, apresentando valores P iguais a 0,75 para a vazão e 0,76 para a temperatura com um nível de significância de 5 %.

Foram gerados três vetores aleatórios com 100.000 elementos e valores de 0 a 1 segundo a distribuição Normal, a saber: \mathbf{u}_F para a vazão, \mathbf{u}_T para a temperatura e \mathbf{v} um vetor comum aos dois, que foi incluído para representar a correlação entre as duas variáveis. Na Equação 4.2, encontra-se a relação que descreve a maneira como o vetor da variável de entrada i foi determinado.

$$\mathbf{x}_i = \bar{x}_i + \sigma_i \cdot [(1 - c) \cdot \mathbf{u}_i + c \cdot \mathbf{v}] \quad (4.2)$$

onde: \bar{x}_i é o valor médio da variável i no intervalo de 0 a 200 s, σ_i é o desvio padrão e c é o coeficiente de similaridade ($0 < c < 1$).

O IQB foi calculado tendo como entrada os valores aleatórios dos vetores determinados pela Equação 4.2. Os desvios-padrões destes valores do IQB calculado são referentes ao método MC, que é o de referência. Os desvios-padrões referentes ao método de séries de Taylor de primeira ordem foram calculados segundo as Equações 2.9 e 2.11, onde a primeira considera as variáveis independentes e a segunda considera a dependência entre as variáveis.

Na Figura 4.12, encontra-se o gráfico comparativo entre os métodos de Monte Carlo e Séries de Taylor de 1ª ordem, considerando as variáveis independentes (var. indep.) e dependentes (var. dep.), para três valores de c : 0,25 (amostra 1), 0,50 (amostra 2) e 0,75 (amostra 3).

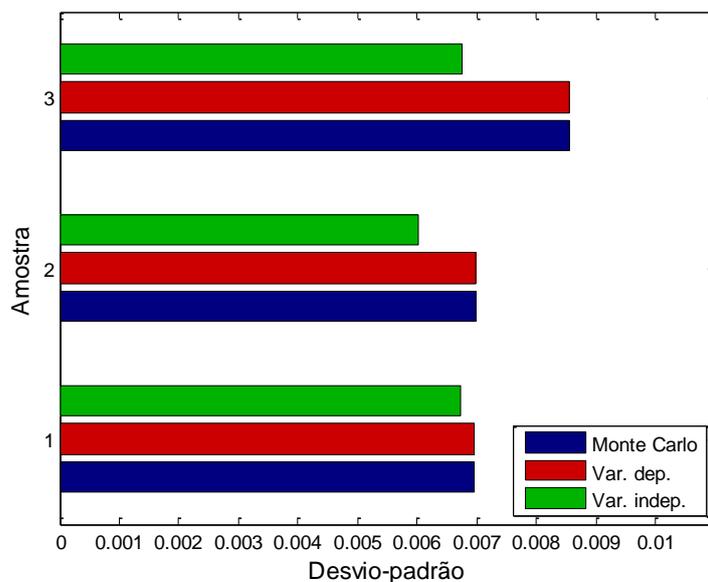


Figura 4.12. Comparação entre as incertezas do IQB.

Observa-se, na Figura 4.12, que quando foi considerada a dependência entre as variáveis, o método de séries de Taylor de 1ª ordem apresentou valores de desvios-padrões muito próximos aos do método MC. Quando a interdependência não é considerada, o desvio-padrão diverge do valor calculado pelo método MC, sendo esta diferença maior quando os dados são determinados com maiores valores de c . Na Tabela 4.3, estão apresentados os erros relativos percentuais com relação ao método MC quando as variáveis são consideradas dependentes e independentes.

Tabela 4.3: Erros relativos das incertezas de acordo com os coeficientes de similaridade

| c | Erro (ind) | Erro (dep) |
|------|------------|------------|
| 0,25 | 3,2243 % | 0,0170 % |
| 0,50 | 13,5665 % | 0,0169 % |
| 0,75 | 21,1461 % | 0,0348 % |

Observa-se, na Tabela 4.3, que quando é considerada a dependência entre as variáveis, o método de séries de Taylor de 1ª ordem apresenta erros menores do que 0,04 % em relação ao método MC para todos os valores de c . Adicionalmente, observa-se que quanto maior o valor de c , maior a penalização por não considerar a dependência entre as variáveis, logo, mostra-se que é de extrema importância incluir o termo da covariância no cálculo da incerteza.

O método de séries de Taylor de 1ª ordem é mais indicado para modelos lineares e polinomiais, no entanto, por apresentar resultados muito próximos do método MC, quando utilizada a Equação 2.11, e incorrer em baixo custo computacional, o mesmo foi utilizado para determinação do desvio-padrão do ***IQB*** nos cenários do estudo de caso.

4.1.3 Determinação de N_A na análise de sensibilidade

Na análise de sensibilidade, o método de Sobol foi comparado considerando-se $N_A = 1.000$, $N_A = 10.000$ e $N_A = 100.000$ para F_S variando de 0,1 a 10 L/m e T_4 de 32°C a 44°C. Na Tabela 4.4, encontram-se os valores obtidos do índice de sensibilidade total para cada variável com seus respectivos desvios-padrões.

Tabela 4.4: Índices de sensibilidade total de acordo com o N_A .

| N_A | $S_{T_{F_S}}$ | $\sigma_{T_{F_S}}$ | $S_{T_{T_4}}$ | $\sigma_{T_{T_4}}$ |
|---------|---------------|--------------------|---------------|--------------------|
| 1.000 | 0,1744 | 0,0248 | 0,9977 | 0,0908 |
| 10.000 | 0,1740 | 0,0062 | 0,9976 | 0,0246 |
| 100.000 | 0,1745 | 0,0024 | 0,9980 | 0,0078 |

Na Tabela 4.4, observa-se que os índices de sensibilidade total apresentam diferenças entre si apenas a partir da terceira casa decimal quando o valor de N_A é alterado, não apresentando um padrão de proporção. No entanto, os valores dos desvios-padrões diminuem à medida que N_A aumenta. O ***IQB*** tem apenas duas variáveis de entrada e, como mostrado na Tabela 4.4, este indicador é mais sensível à variável T_4 , apresentando um índice muito maior do que o de F_S . Logo, para o estudo de caso em questão, escolheu-se $N_A = 1.000$, visto que um desvio padrão da ordem de 10^{-2} não influenciaria na determinação de qual variável é a mais sensível e reduziria o esforço computacional.

4.2 Cenários

4.2.1 Cenário 1: Distúrbios na vazão

No cenário 1, foram adicionados dois distúrbios do tipo degrau na vazão (F_d). No tempo $t = 500$ s, o F_d foi inserido em forma de degrau unitário negativo e em $t = 1300$ s, um degrau unitário positivo. Na Figura 4.13, estão apresentados os gráficos do distúrbio, do ***IQB***, F_S e T_4 , respectivamente, para este cenário.

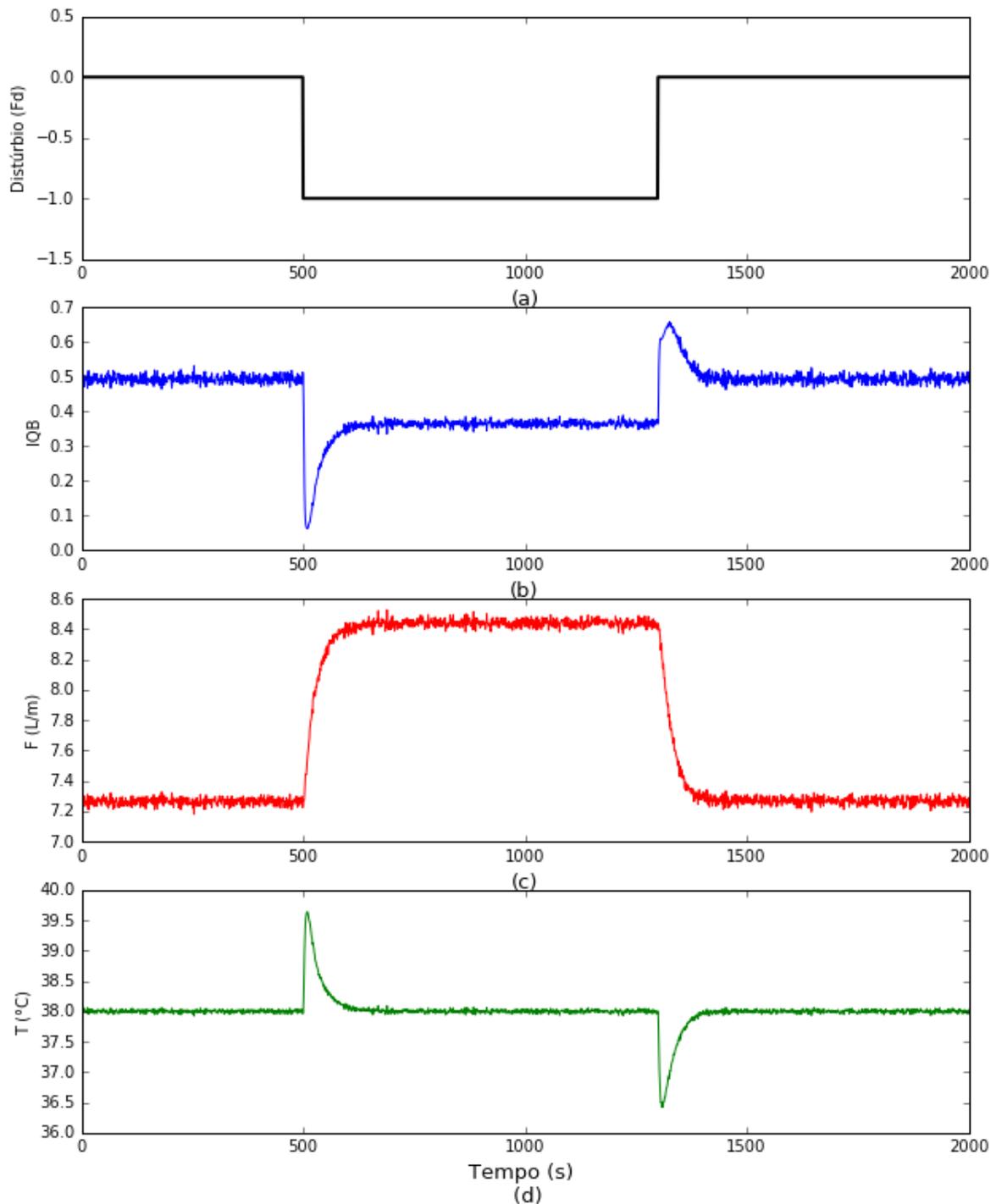


Figura 4.13. (a) Gráfico do distúrbio na vazão em $t = 500$ s e $t = 1300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho).

Na , observa-se que o distúrbio negativo na vazão no tempo de 500 segundos provoca uma redução do IQB e o distúrbio positivo em 1300 s ocasiona um aumento do indicador. O distúrbio negativo na vazão provoca a redução do nível do tanque,

provocando um aumento na temperatura de saída (), visto que uma menor quantidade de água está sendo aquecida. A redução do nível aciona a válvula de entrada de água fria, que é aberta para aumentar o nível. Ao aumentar a temperatura de saída, a válvula de saída abre mais para corrigir a temperatura, logo, neste ponto, a vazão de saída começa a aumentar, como visto na . O oposto acontece quando o distúrbio é positivo, o nível do tanque aumenta e a válvula de entrada de água fria é fechada. Devido à redução da temperatura de saída, a válvula de saída fecha mais, reduzindo a vazão de saída.

Na Figura 4.14, encontra-se o *StatSSCandlePlot* do *IQB* para o cenário 1 em um período de tempo de 2000 segundos dividido em 10 janelas.

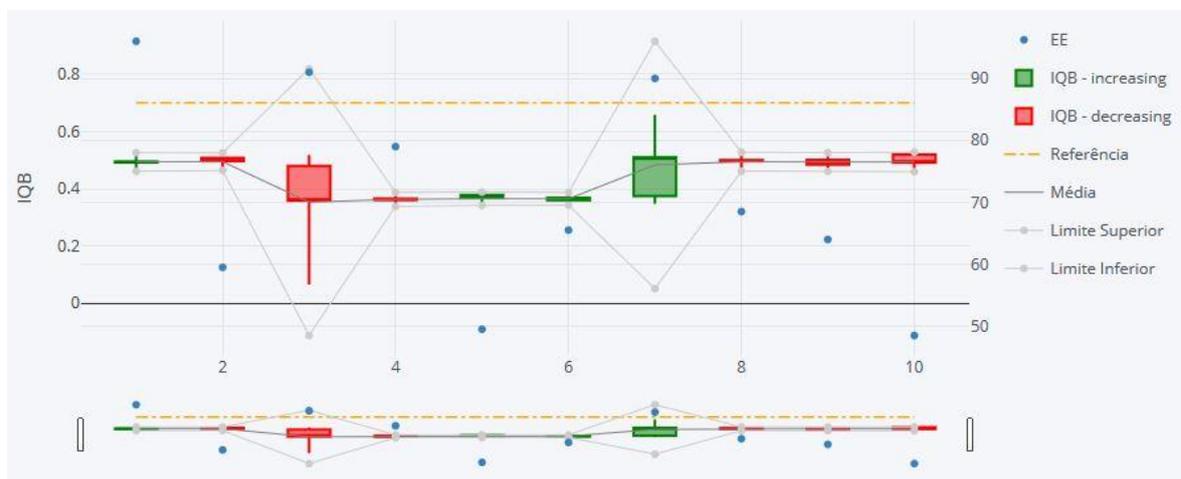


Figura 4.14. *StatSSCandlePlot* do *IQB* no cenário 1.

Na Figura 4.14, observa-se que na janela 3 o valor de fechamento do *IQB* foi bastante inferior ao valor de abertura, no entanto, o fechamento foi consideravelmente superior ao mínimo, indicando uma tendência de recuperação. No caso da janela 7, ocorreu o inverso, o *IQB* finalizou o período com um valor maior do que o seu valor de abertura, indicando uma tendência de alta, no entanto, como o seu valor de fechamento ficou bem abaixo do máximo da janela, nota-se que o indicador poderia ter permanecido em um valor maior, mas sofreu uma queda antes do fechamento do período. As demais janelas apresentaram tendências muito baixas de alta e queda, visto que os valores assumidos foram muito próximos. Adicionalmente, observa-se que os valores médios do *IQB* ficaram abaixo do valor de referência em todas as janelas.

Como observado na Figura 4.14, as janelas possuem diferentes percentuais no estado estacionário que estão indicados pelos pontos azuis. Estes estacionários estão detalhados na Figura A.1 do Apêndice A.

Ao se deparar com o gráfico da Figura 4.14, as janelas 3 e 7 são as que mais chamam atenção e é de extrema importância saber qual das variáveis foi a mais impactante em cada uma destas variações significantes no processo. Para isto, foi feito o gráfico com os resultados da análise de sensibilidade, que é complementar ao *StatSCandlePlot* e está apresentado na Figura 4.15.

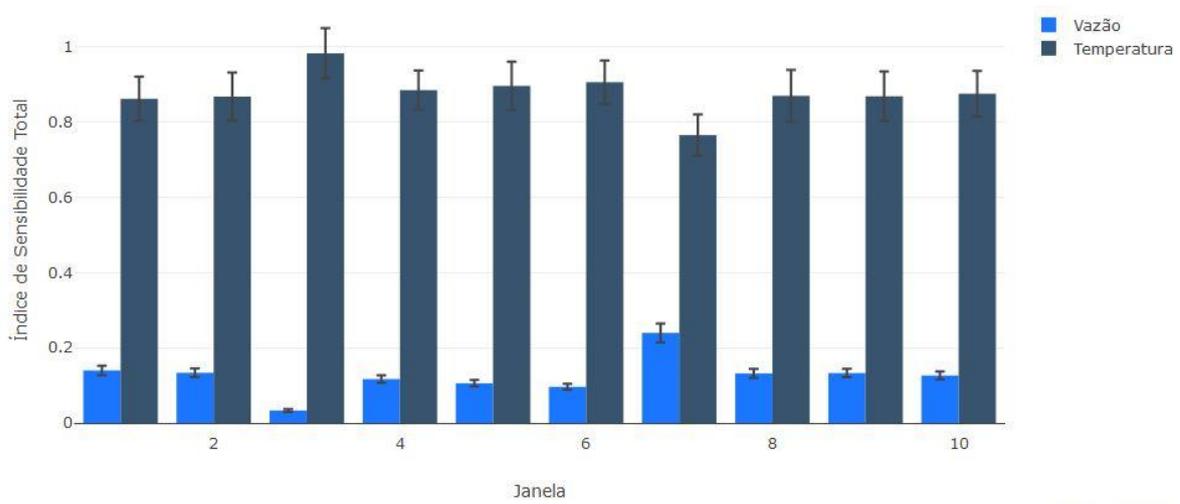


Figura 4.15. Análise de sensibilidade do *IQB* no cenário 1.

Na Figura 4.15, observa-se que em todos os casos a temperatura é a variável com maior índice de sensibilidade total. Logo, a variação do *IQB* é, em sua maior parte, devido à variação da temperatura. De posse desta informação, o foco para a melhoria do indicador é a temperatura de saída do processo.

4.2.2 Cenário 2: Queda da temperatura fria

Neste cenário, a temperatura de água fria na caixa d'água caiu de forma linear seguindo um comportamento do tipo rampa. A Figura 4.16 apresenta os gráficos do distúrbio, do *IQB*, F_5 e T_4 , respectivamente, para este cenário.

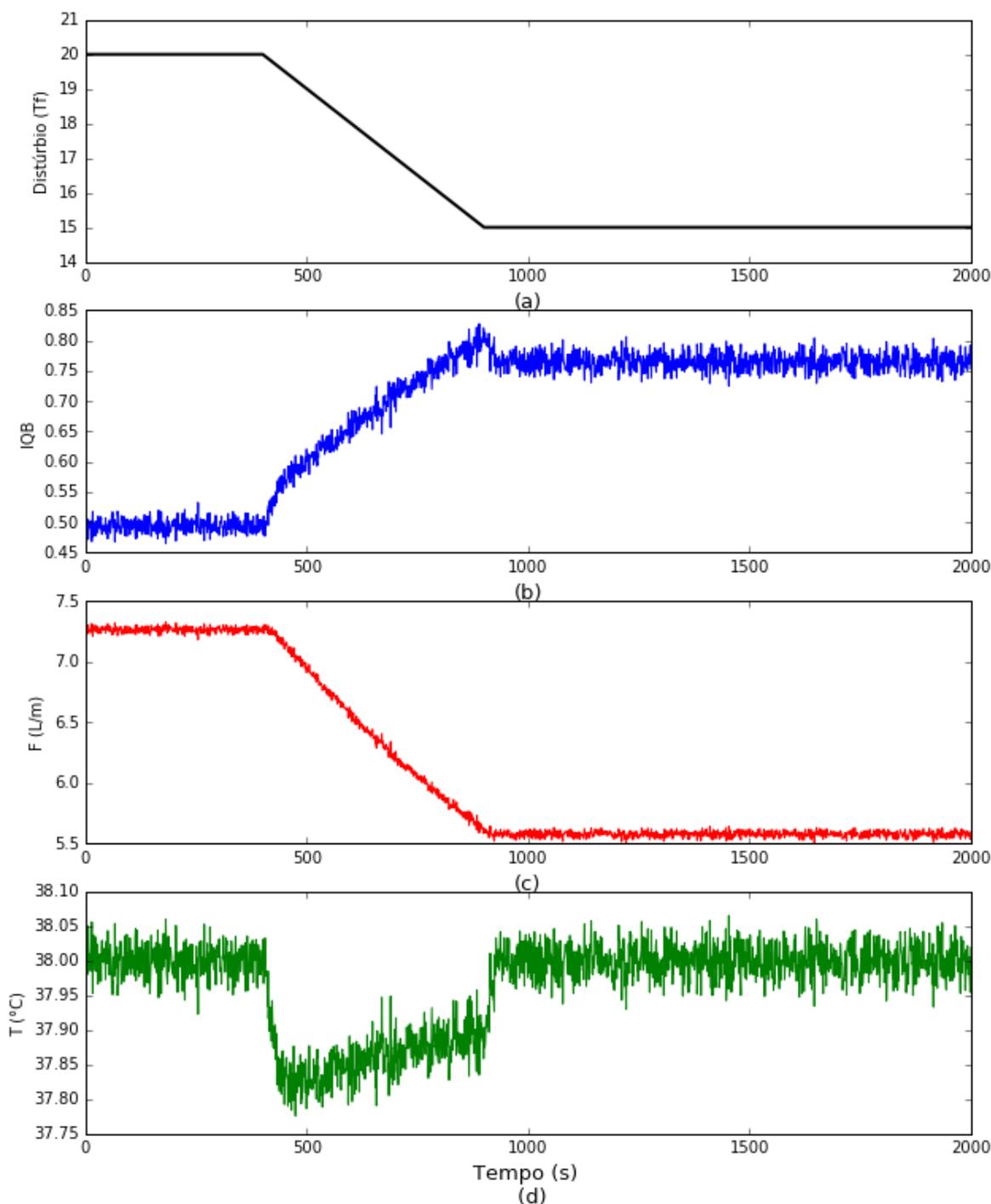


Figura 4.16. (a) Gráfico do distúrbio na temperatura da água fria (T_f), (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho).

A queda na temperatura fria provoca um aumento gradual do IQB, como visto na Figura 4.16(b). Isto aconteceu porque a queda da temperatura de saída reflete a queda da temperatura da caixa d'água. Como T_4 é controlada pela válvula de saída, esta é fechada, provocando a queda da vazão, como visto na Figura 4.16(c), para que o tanque tenha

tempo suficiente de aquecer mais a água e retornar à temperatura desejada (Figura 4.16(d)).

Na Figura 4.17, tem-se o *StatSSCandlePlot* referente ao cenário 2.

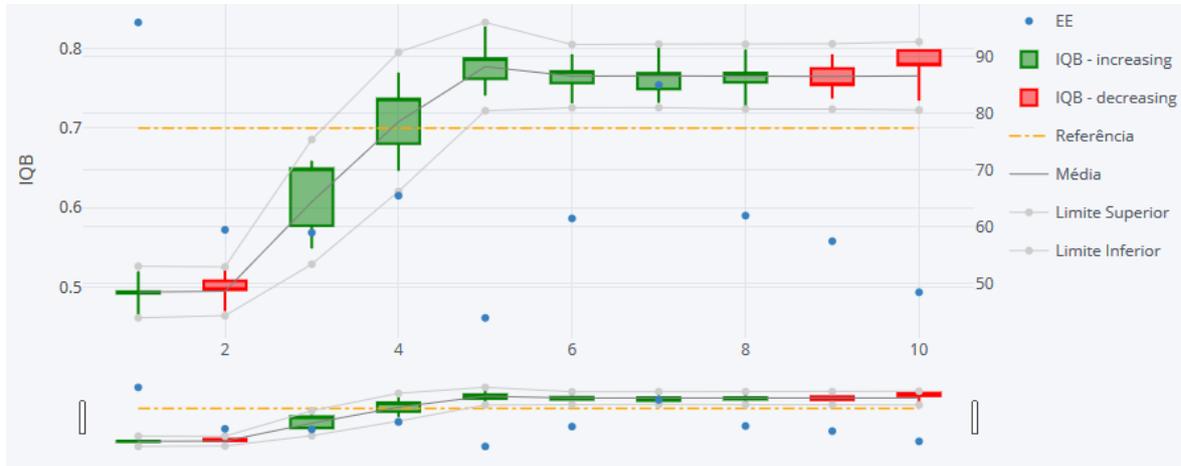


Figura 4.17. *StatSSCandlePlot* do IQB no cenário 2.

Inicialmente, na Figura 4.17, o IQB está em torno de 0,5, abaixo do valor de referência, assumindo valores próximos nas janelas 1 e 2. A partir da janela 3, o IQB cresce para cerca de 0,63 e o corpo do *candle* aumenta. Adicionalmente, o seu valor de fechamento fica superior ao valor de abertura, indicando um sentimento positivo em torno do IQB sugerindo que o mesmo pode aumentar no futuro. Na janela 4, o indicador possui valor de fechamento maior do que o de abertura e uma média superior ao valor de referência. Ao avançar para a janela 5, percebe-se que os valores de abertura e fechamento foram muito próximos e, adicionalmente, que o indicador cresceu consideravelmente na janela, mas voltou a cair. Nas janelas seguintes, a média do IQB ficou aproximadamente constante e os seus respectivos *candles* indicaram leves tendências de queda e alta alternadas, mantendo o valor médio do indicador acima do valor de referência. Os estacionários estão detalhados na Figura A.2 do Apêndice A.

A análise de sensibilidade do cenário 2 encontra-se a seguir na Figura 4.18.

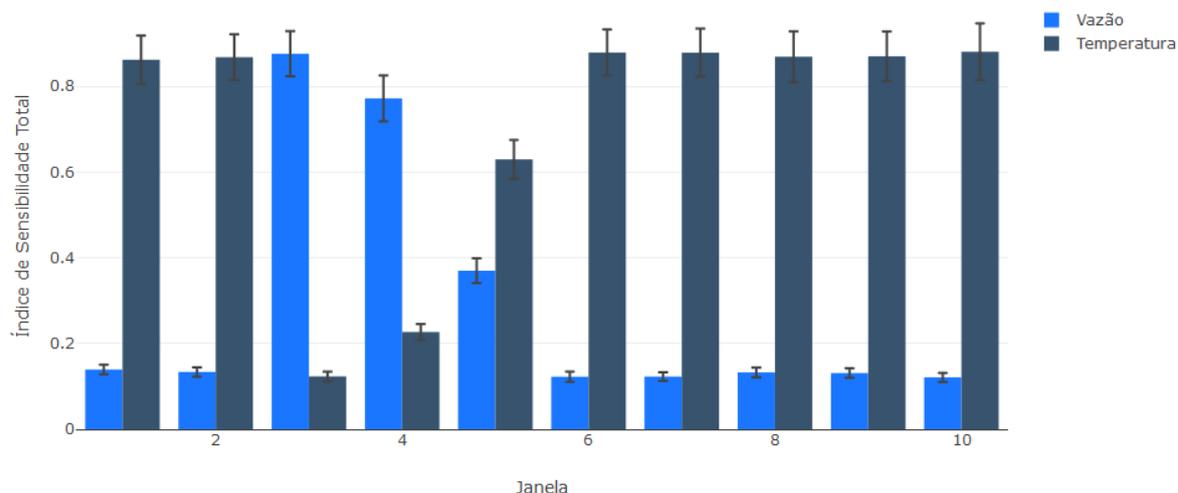


Figura 4.18. Análise de sensibilidade do IQB no cenário 2.

Na Figura 4.18, verifica-se que a vazão passa a ter uma contribuição mais significativa na variação do IQB quando a temperatura da caixa d'água começa a cair, tornando-se a variável com maior índice de sensibilidade total nas janelas 3, 4 e 5. Nas demais janelas, o IQB continua sendo mais afetado pela variação de T_4 .

4.2.3 Cenário 3: Mudança na estratégia de controle

No cenário 3, a estratégia de controle foi alterada, como visto na Figura 4.4. A Figura 4.19 apresenta os gráficos do distúrbio, do IQB , F_s e T_4 quando o sistema é submetido a distúrbios na vazão (F_d), um degrau unitário negativo em $t = 500 s$ e um degrau unitário positivo no tempo $t = 1300 s$.

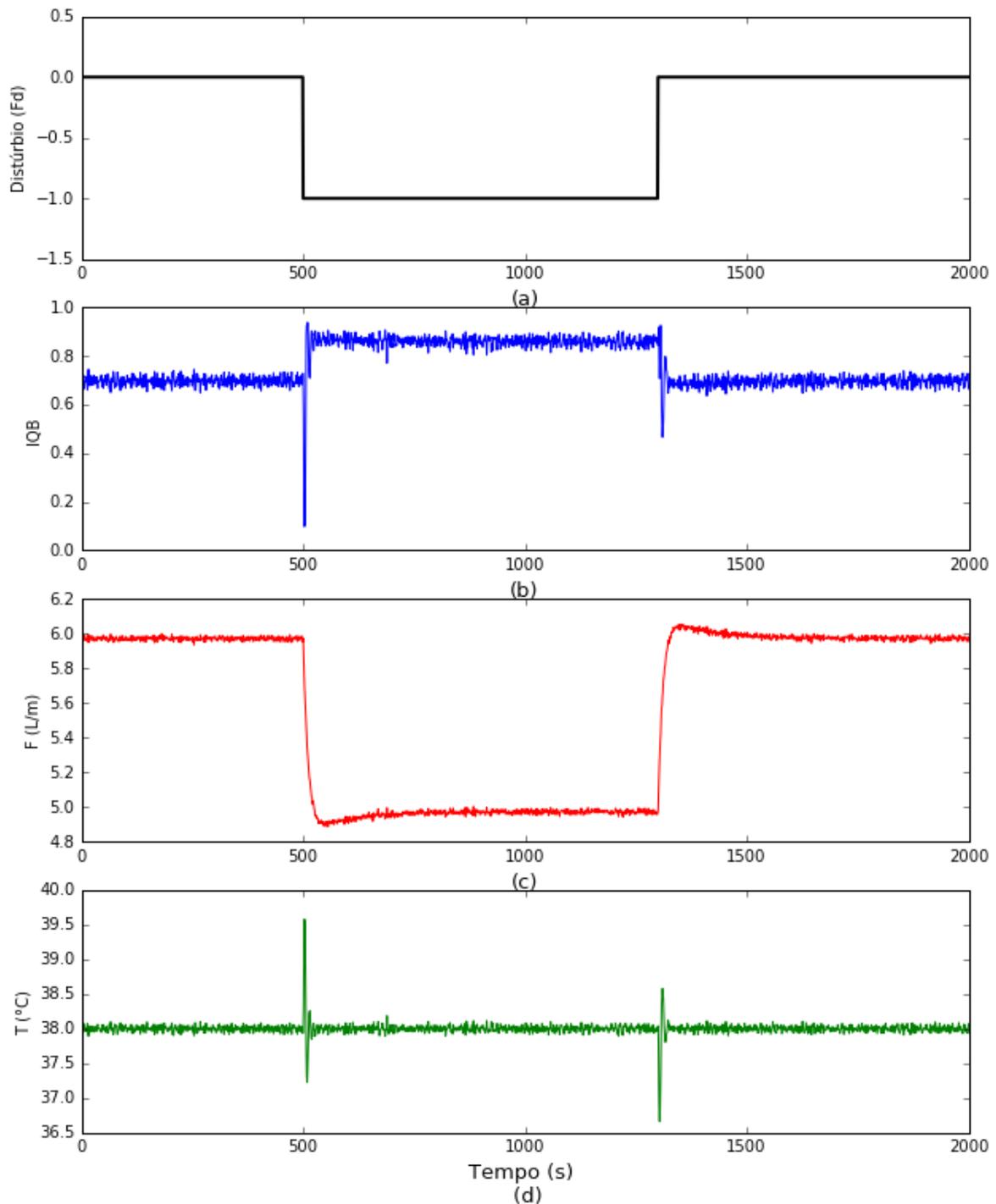


Figura 4.19. (a) Gráfico do distúrbio na vazão em $t = 500$ s e $t = 1300$ s, (b) IQB, (c) F_s (vazão da água do banho) e (d) T_4 (temperatura do banho).

A partir do gráfico da Figura 4.19(b), observa-se que o IQB foi beneficiado com o distúrbio negativo na vazão e após o distúrbio positivo, retornou ao patamar anterior. Ao submeter o sistema da Figura 4.4 a um distúrbio negativo na vazão, o nível do tanque

reduz e a válvula de saída é fechada, reduzindo a vazão de saída, como mostrado na Figura 4.19(c). Como esta estratégia de controle tem uma potência variável, as alterações na temperatura de saída se restabelecem mais rapidamente, como visto na Figura 4.19(d), quando comparadas ao caso apresentado no cenário 1 (Figura 4.13(d)).

Na Figura 4.20, encontra-se o *StatSSCandlePlot* do *IQB* referente ao cenário 3.

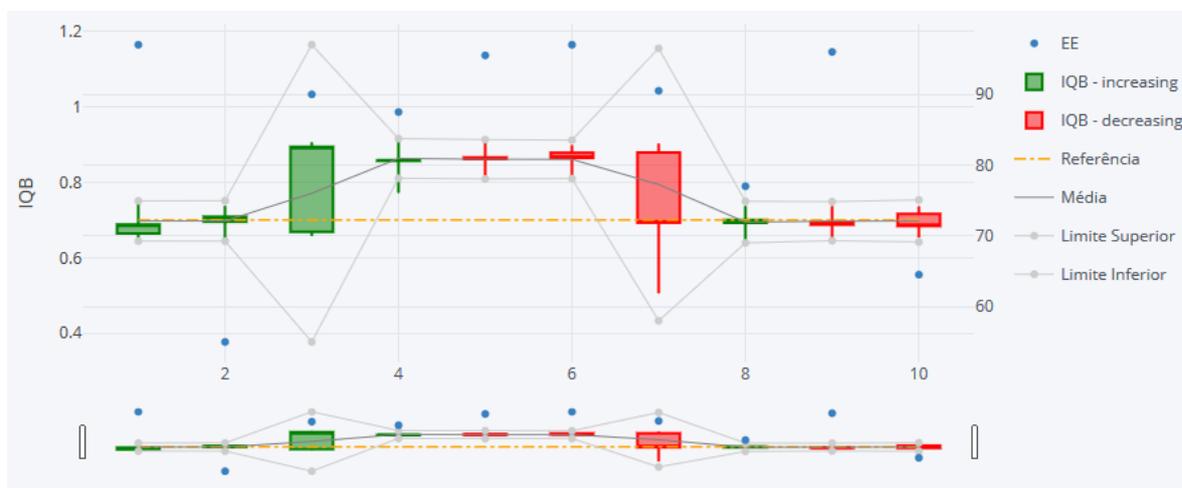


Figura 4.20. *StatSSCandlePlot* do *IQB* no cenário 3.

A partir da Figura 4.20, observa-se que na janela 3, o indicador apresentou tendência de alta no período e que o mesmo teve seu valor de abertura (fechamento) muito próximo do mínimo (máximo), sugerindo aumento futuro do indicador. Na janela 7, o valor de fechamento está consideravelmente acima do mínimo, indicando uma tendência de recuperação do *IQB*. Os demais *candles* são neutros e não fornecem informações relevantes a respeito dos movimentos futuros. Ainda no gráfico, observa-se o elevado desvio-padrão nas janelas 3 e 7, resultante da mudança repentina de patamares do indicador, e também que todas as janelas apresentaram valores médios iguais ou superiores ao valor de referência. O detalhamento dos estados em cada janela encontra-se na Figura A.3 do Apêndice A.

Na Figura 4.21, encontram-se os resultados da análise de sensibilidade para as janelas do cenário 3.

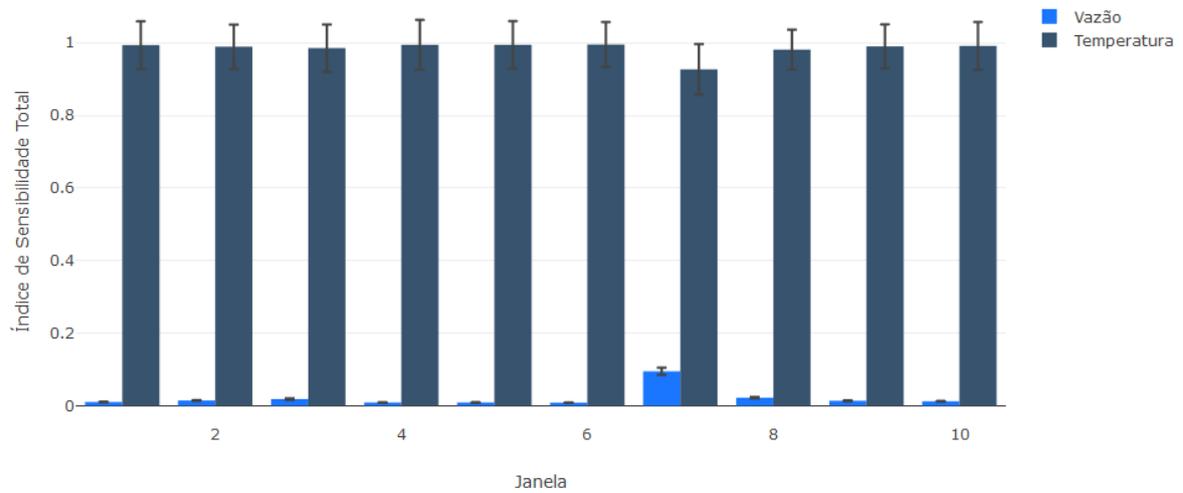


Figura 4.21. Análise de sensibilidade do *IQB* no cenário 3.

Na Figura 4.21, observa-se que, assim como no cenário 1, a temperatura é a variável que mais influencia no *IQB* em todas as janelas. Uma variação, mesmo que pequena nesta variável, é capaz de mudar substancialmente o *IQB*.

Capítulo 5 – Considerações Finais

5.1 Conclusões

Este trabalho teve como objetivo a elaboração de uma metodologia para a quantificação de KPIs, envolvendo tratamento de dados, através da remoção de ruído e da detecção do estado estacionário, e análises estatísticas, como a propagação de erros e a análise de sensibilidade. Adicionalmente, foi apresentada uma nova ferramenta de acompanhamento de KPIs, ainda não encontrada na literatura, denominada *StatSSCandlePlot*.

Antes de aplicar a metodologia, foram realizadas análises para ajuste manual dos parâmetros da remoção de ruído e detecção do estado estacionário pelos métodos DWT e DER, respectivamente, e da análise de sensibilidade pelo método de Sobol. Adicionalmente, o método de propagação de erros via séries de Taylor de primeira ordem foi comparado ao método Monte Carlo, que é considerado um método de referência. Após analisar diferentes cenários variando o índice de estacionariedade limiar e o número de níveis de decomposição de *Wavelets*, escolheu-se um valor de I_{ee} igual a 0,6 e N igual a 4, sendo estes os parâmetros mais críticos de ajuste. Para o método de Sobol, foram testados três valores diferentes para o número de amostras a serem geradas, a saber: 1.000, 10.000 e 100.000, e decidiu-se utilizar $N_A = 1.000$, visto que valores maiores implicavam em maior custo computacional e diferenças irrisórias no índice de sensibilidade, com um desvio-padrão da ordem de 10^{-2} que, para este estudo de caso, não implicaria em grandes mudanças nos resultados. Comparando o método de séries de

Taylor de primeira ordem com o método de Monte Carlo, observou-se que quando foi considerada a dependência entre as variáveis, o erro relativo foi menor do que 0,04%. No entanto, ao desconsiderar o termo da covariância, os erros aumentaram para aproximadamente 3%, 14% e 21% considerando um coeficiente de similaridade de 0,25, 0,50 e 0,75, respectivamente.

A metodologia proposta foi aplicada em três cenários distintos de um chuveiro contendo dois princípios de aquecimento, gás e energia elétrica. No primeiro cenário, o sistema foi submetido a distúrbios na vazão, no segundo, houve uma queda de temperatura na caixa d'água e, no terceiro, distúrbios na vazão incidiram no sistema com uma nova estratégia de controle. Para cada cenário, os dados foram divididos em janelas e foram calculados os valores de média, desvio-padrão, percentual no estado estacionário e índice de sensibilidade total do índice de qualidade de banho, o indicador estudado que é dependente da temperatura e da vazão de saída do chuveiro. A partir destes dados, os mesmos foram expostos no *StatSSCandlePlot*, que também apresenta o padrão *candlestick* dos dados.

A partir do padrão *candlestick* foi possível identificar algumas tendências do *IQB* em cada cenário, baseando-se nos valores de abertura, fechamento, mínimo e máximo assumidos em cada janela. De forma complementar a este gráfico principal, foi apresentado o gráfico da análise de sensibilidade para cada cenário que, para os cenários 1 e 3, indicou que a temperatura é a variável que mais influencia no *IQB*, apresentando índices de sensibilidade, no geral, maiores do que 0,8 frente a valores menores do que 0,4 para a vazão, fornecendo um diagnóstico inicial a respeito da principal causa de variação do indicador. No cenário 2, o *IQB* se tornou mais sensível à vazão de saída em 30% das janelas.

A metodologia proposta, em conjunto com o *StatSSCandlePlot*, é uma ferramenta promissora na área de acompanhamento de processos por apresentar informações importantes dos indicadores apenas quando o mesmo está no estado estacionário, além de informar suas incertezas e um diagnóstico rápido e visual da variação do processo.

5.2 Sugestões para trabalhos futuros

Para aprimoramento da metodologia e da ferramenta gráfica propostas neste trabalho, sugere-se a incorporação de outros métodos de propagação de erros e análise de sensibilidade, o ajuste automático dos parâmetros da detecção do estado estacionário e a adição de mais funções de *Wavelet* na função *ssdet*, visto que a depender do comportamento da série de dados e do indicador estudado, existem diferentes parâmetros e métodos mais adequados para serem escolhidos.

Outro ponto importante é a adição de limites inferior e superior no *StatSSCandlePlot* baseando-se em distribuições de saída não-Gaussianas. Para incrementar a parte de interpretação do padrão *candlestick*, seria interessante programar os padrões assumidos no gráfico para, ao mover o cursor do mouse, identificar a tendência do indicador baseando-se nas interpretações do mercado financeiro.

Referências

ABUZAID, A. H.; MOHAMED, I. B.; HUSSIN, A. G. Boxplot for circular variables. **Computational Statistics**, v. 27, n. 3, p. 381-392, 2012/09/01 2012.

ANDERLE, M.; KIRBY, M. **An Application of the Maximum Noise Fraction Method to Filtering Noisy Time Series**. 5th International Conference on Mathematics in Signal Processing. University of Warwick, Coventry, UK 2001.

ANDERSON, T. V.; MATTSON, C. A. Propagating Skewness and Kurtosis Through Engineering Models for Low-Cost, Meaningful, Nondeterministic Design. **Journal of Mechanical Design**, v. 134, n. 10, p. 100911-100911, 2012.

ANDERSON, T. V. et al. Efficient Propagation of Error Through System Models for Functions Common in Engineering. **Journal of Mechanical Design**, v. 134, n. 1, p. 014501-014501-6, 2012.

ASME. **Fired Steam Generators: Performance Test Codes**. New York: The American Society of Mechanical Engineers. ASME PTC 4: 294 p. 2008.

ASSAF, S. et al. Performance evaluation and benchmarking for maintenance decision making units at petrochemical corporation using a DEA model. **International Journal of Advanced Manufacturing Technology**, v. 76, n. 9-12, p. 1957-1967, 2015.

AURANGZEB, M.; JANA, A. K. Dividing wall column: Improving thermal efficiency, energy savings and economic performance. **Applied Thermal Engineering**, v. 106, p. 1033-1041, 2016/08/05/ 2016.

BRUFFAERTS, C.; VERARDI, V.; VERMANDELE, C. A generalized boxplot for skewed and heavy-tailed distributions. **Statistics & Probability Letters**, v. 95, p. 110-117, 2014/12/01/ 2014.

CAMPO, D.; QUINTERO, O. L.; BASTIDAS, M. Multiresolution analysis (discrete wavelet transform) through Daubechies family for emotion recognition in speech. **Journal of Physics: Conference Series**, v. 705, n. 1, p. 012034, 2016.

CAMPOLONGO, F.; CARIBONI, J.; SALTELLI, A. An effective screening design for sensitivity analysis of large models. **Environmental Modelling & Software**, v. 22, n. 10, p. 1509-1518, 2007/10/01/ 2007.

CAO, S.; RHINEHART, R. R. An efficient method for on-line identification of steady state. **Journal of Process Control**, v. 5, n. 6, p. 363-374, 1995/12/01/ 1995.

_____. Critical values for a steady-state identifier. **Journal of Process Control**, v. 7, n. 2, p. 149-152, 1997/01/01/ 1997.

CERIC, E. **Crude Oil, Processes and Products**. Saravejo: IBC, 2012. 440.

CHASE, K. W.; PARKINSON, A. R. A survey of research in the application of tolerance analysis to the design of mechanical assemblies. **Research in Engineering Design**, v. 3, n. 1, p. 23-37, March 01 1991.

CHRISTOPHER FREY, H.; PATIL, S. R. Identification and Review of Sensitivity Analysis Methods. **Risk Analysis**, v. 22, n. 3, p. 553-578, 2002.

COX, N. D. TOLERANCE ANALYSIS BY COMPUTER. **Journal of Quality Technology**, v. 11, n. 2, p. 80-87, 1979.

CUKIER, R. I. et al. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. **The Journal of Chemical Physics**, v. 59, n. 8, p. 3873-3878, 1973/10/15 1973.

FERREIRA, J. E. V. et al. Graphical representation of chemical periodicity of main elements through boxplot. **Educación Química**, v. 27, n. 3, p. 209-216, 2016/07/01/ 2016.

FISCHMANN, A. A.; ZILBER, M. A. **Utilização de Indicadores de Desempenho como Instrumento de Suporte à Gestão Estratégica**. Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração. Rio de Janeiro 1999.

GLANCY, C. G.; CHASE, K. W. **A Second-Order Method for Assembly Tolerance Analysis**. ASME Design Engineering Technical Conferences. Las Vegas 1999.

GREEN, A. A. et al. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. **IEEE Transactions on Geoscience and Remote Sensing**, v. 26, n. 1, p. 65-74, 1988. ISSN 0196-2892.

HEIJUNGS, R.; LENZEN, M. Error propagation methods for LCA—a comparison. **The International Journal of Life Cycle Assessment**, v. 19, n. 7, p. 1445-1461, July 01 2014.

HELTON, J. C. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. **Reliability Engineering & System Safety**, v. 42, n. 2, p. 327-367, 1993/01/01/ 1993.

HELTON, J. C.; DAVIS, F. J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. **Reliability Engineering & System Safety**, v. 81, n. 1, p. 23-69, 2003/07/01/ 2003.

JIANG, C. et al. Interval arithmetic operations for uncertainty analysis with correlated interval variables. **Acta Mechanica Sinica**, v. 32, n. 4, p. 743-752, August 01 2016.

JIANG, T. et al. Application of steady-state detection method based on wavelet transform. **Computers & Chemical Engineering**, v. 27, n. 4, p. 569-578, 2003/04/15/ 2003.

JUSTEN, L.; RAMLAU, R. A general framework for soft-shrinkage with applications to blind deconvolution and wavelet denoising. **Applied and Computational Harmonic Analysis**, v. 26, n. 1, p. 43-63, 2009/01/01/ 2009.

KENT, E. et al. What Can We Learn from Global Sensitivity Analysis of Biochemical Systems? **PLOS ONE**, v. 8, n. 11, p. e79244, 2013.

KORBEL, M. et al. Steady state identification for on-line data reconciliation based on wavelet transform and filtering. **Computers & Chemical Engineering**, v. 63, p. 206-218, 2014/04/17/ 2014.

KUO, W.; UPPULURI, V. R. R. A review of error propagation analysis in systems. **Microelectronics Reliability**, v. 23, n. 2, p. 235-248, 1983/01/01/ 1983.

KURITA, T. Dynamic characteristics of the daily yen-dollar exchange rate. **Research in International Business and Finance**, v. 30, p. 72-82, 2014/01/01/ 2014.

LEE, D. T. L.; YAMAMOTO, A. Wavelet Analysis: Theory and Applications. **Hewlett-Packard Journal**, 1994.

LEE, K. H.; JO, G. S. Expert system for predicting stock market timing using a candlestick chart. **Expert Systems with Applications**, v. 16, n. 4, p. 357-364, 1999/05/01/ 1999.

LEE, S.; KIM, J. Discrete wavelet transform-based denoising technique for advanced state-of-charge estimator of a lithium-ion battery in electric vehicles. **Energy**, v. 83, p. 462-473, 2015/04/01/ 2015.

LUO, J.; BAI, J.; SHAO, J. **Application of the wavelet transform on axial strain calculation in ultrasound elastography**. Progress in Natural Science. China: China Academic Journal Electronic Publishing House. 16 2006.

MALLAT, S. **A Wavelet Tour of Signal Processing**. USA: 1999.

MARSHALL, B. R.; YOUNG, M. R.; ROSE, L. C. Candlestick technical trading strategies: Can they create value for investors? **Journal of Banking & Finance**, v. 30, n. 8, p. 2303-2323, 2006/08/01/ 2006.

MEJIA, R. I. G. et al. **Novo Método para a Identificação de Estado Estacionário Baseada na Estimativa da Autocorrelação Local**. XVIII Congresso Brasileiro de Automática. Bonito - MS 2010.

MEJIA, R. I. G.; DUARTE, M. B.; TRIERWEILER, J. O. **Avaliação do Desempenho de Métodos de Identificação do Estado Estacionário**. XVIII Congresso Brasileiro de Automática. Bonito, MS: 9 p. 2010a.

_____. **Avaliação do Desempenho e Ajuste Automático de Métodos de Identificação de Estado Estacionário**. XVIII Congresso Brasileiro de Engenharia Química. Foz do Iguaçu, PR: 10 p. 2010b.

MISITI, M. et al. **Wavelet Toolbox™ User's Guide** *Wavelet Toolbox™ User's Guide: The MathWorks*: 736 p. 2016.

MONTEQUÍN, V. R. et al. Scorecard and KPIs for monitoring software factories effectiveness in the financial sector. **International Journal of Information Systems and Project Management**, v. 1, n. 3, p. 29-43, 2013.

MORRIS, M. D. Factorial Sampling Plans for Preliminary Computational Experiments. **Technometrics**, v. 33, n. 2, p. 161-174, 1991/05/01 1991.

MÜLLER, C. J. **Modelo de Gestão Integrando Planejamento Estratégico, Sistemas de Avaliação de Desempenho e Gerenciamento de Processos (MEIO - Modelo de Estratégia, Indicadores e Operações)**. 2003. 292 (Doutorado). Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Porto Alegre.

NADER, B.; TOMI, G. D.; PASSOS, A. O. Indicadores-chave de desempenho e a gestão integrada da mineração. **Rem: Revista Escola de Minas**, v. 65, p. 537-542, 2012.

NARVAES-GARCIA, A. et al. PERFORMANCE INDICES TO DESIGN A MULTICOMPONENT BATCH DISTILLATION COLUMN USING A SHORTCUT METHOD. **Brazilian Journal of Chemical Engineering**, v. 32, p. 595-608, 2015.

NASRI, M.; NEZAMABADI-POUR, H. Image denoising in the wavelet domain using a new adaptive thresholding function. **Neurocomputing**, v. 72, n. 4, p. 1012-1025, 2009/01/01/2009.

PRATIK; PRADHAN, G.; SHAHNAWAZUDDIN, S. Denoising of ECG signal by non-local estimation of approximation coefficients in DWT. **Biocybernetics and Biomedical Engineering**. 2017.

ROCHA, V. B. D. **Uma abordagem de Wavelets aplicada à combinação de previsões: uma análise teórica e experimental**. 2008. 155 Universidade Federal do Paraná, Curitiba.

SALTELLI, A. Making best use of model evaluations to compute sensitivity indices. **Computer Physics Communications**, v. 145, n. 2, p. 280-297, 2002/05/15/ 2002.

SALTELLI, A. et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. **Computer Physics Communications**, v. 181, n. 2, p. 259-270, 2010/02/01/ 2010.

_____. Front Matter. In: (Ed.). **Global Sensitivity Analysis. The Primer**: John Wiley & Sons, Ltd, 2008. p.i-xi. ISBN 9780470725184.

SALTELLI, A.; TARANTOLA, S.; CHAN, K. P. S. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. **Technometrics**, v. 41, n. 1, p. 39-56, 1999.

SHENSA, M. J. The discrete wavelet transform: wedding the a trous and Mallat algorithms. **IEEE Transactions on Signal Processing**, v. 40, n. 10, p. 2464-2482, 1992. ISSN 1053-587X.

SOBOL, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. **Mathematics and Computers in Simulation**, v. 55, n. 1, p. 271-280, 2001/02/15/ 2001.

STASZEWSKI, W. J. Intelligent signal processing for damage detection in composite materials. **Composites Science and Technology**, v. 62, n. 7, p. 941-950, 2002/06/01/ 2002..

SUN, L.; RIEGER, J.; HINRICHS, H. Maximum noise fraction (MNF) transformation to remove ballistocardiographic artifacts in EEG signals recorded during fMRI scanning. **NeuroImage**, v. 46, n. 1, p. 144-153, 2009/05/15/ 2009.

SUN, Y.; ZHANG, J. Error propagation characteristic analysis of halo orbit based on condition number of state transition matrix. **Astrophysics and Space Science**, v. 361, n. 9, p. 297, 2016.

TASWELL, C. **Experiments in Wavelet Shrinkage Denoising**. Journal of Computational Methods in Science and Engineering. 1: 12 p. 2001.

TO, A. C.; MOORE, J. R.; GLASER, S. D. Wavelet denoising techniques with applications to experimental geophysical data. **Signal Processing**, v. 89, n. 2, p. 144-160, 2009/02/01/ 2009.

TORKKO, M. et al. Quality KPIs in Pharmaceutical and Food Industry. **Journal of Pharmaceutical Innovation**, v. 8, n. 4, p. 205-211, 2013/12/01 2013.

WEBER, A.; THOMAS, R. **Key Performance Indicators: Measuring and Managing the Maintenance Function**. Ontario. 2005

WINK, A. M.; ROERDINK, J. B. T. M. Polyphase decompositions and shift-invariant discrete wavelet transforms in the frequency domain. **Signal Processing**, v. 90, n. 6, p. 1779-1787, 2010/06/01/ 2010.

XUE, J.; LEUNG, Y.; MA, J.-H. High-order Taylor series expansion methods for error propagation in geographic information systems. **Journal of Geographical Systems**, v. 17, n. 2, p. 187-206, 2015.

YAN, H.; WU, X.; YANG, J. Application of Monte Carlo Method in Tolerance Analysis. **Procedia CIRP**, v. 27, p. 281-285, 2015/01/01 2015.

ZHANG, S. et al. Application of spline wavelet transform in differential of electroanalytical signal. **Chinese Science Bulletin**, v. 46, n. 7, p. 550, 2001.

ZHANG, X. Y. et al. Sobol Sensitivity Analysis: A Tool to Guide the Development and Evaluation of Systems Pharmacology Models. **CPT: Pharmacometrics & Systems Pharmacology**, v. 4, n. 2, p. 69-79, 2015.

ZHU, L. et al. Coordinating and Evaluating of Multiple Key Performance Indicators for Manufacturing Equipment: Case Study of Distillation Column. **Chinese Journal of Chemical Engineering**, v. 22, n. 7, p. 805-811, 2014/07/01/ 2014.

Apêndice A - Estados dos cenários

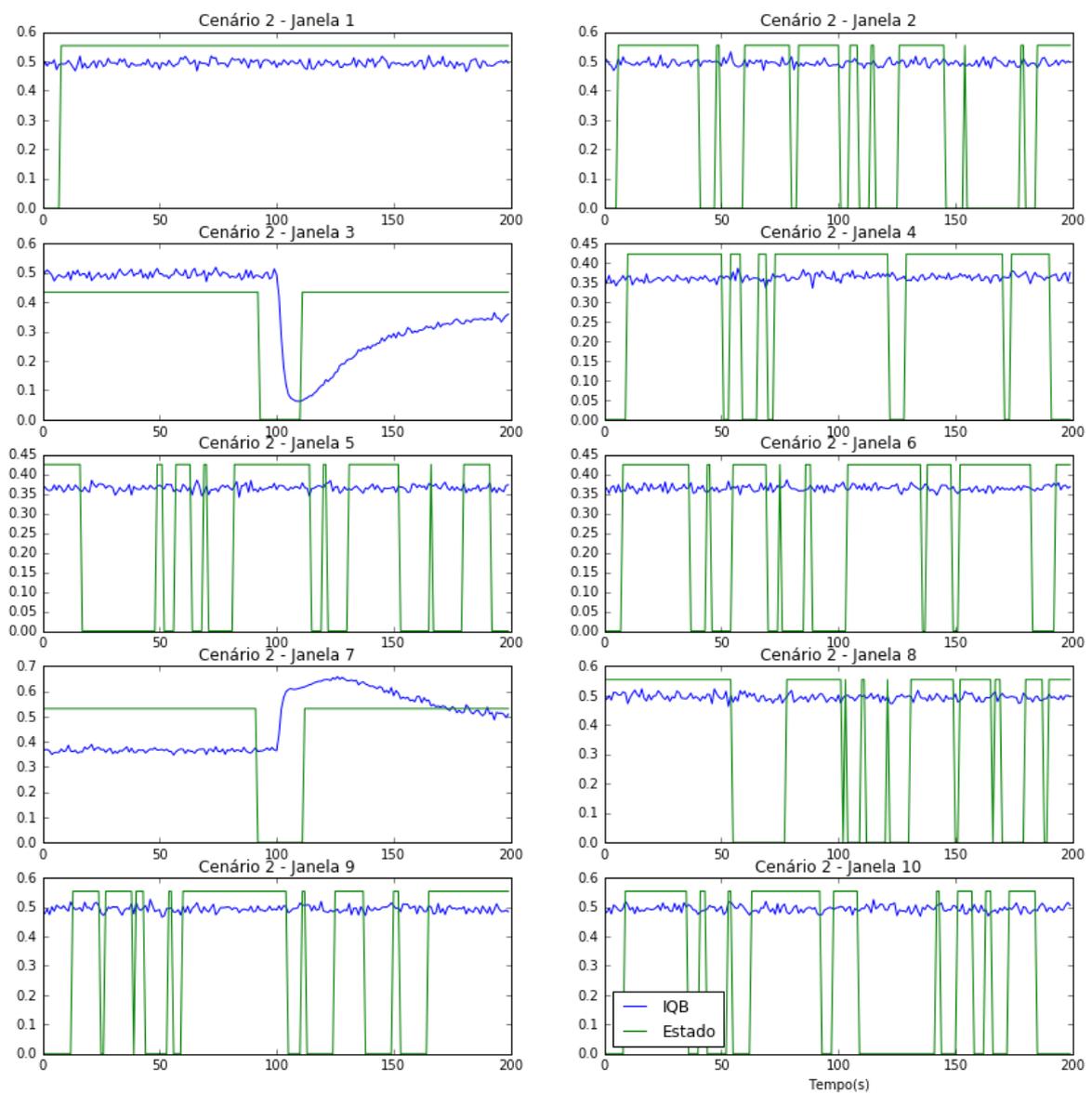


Figura A.1. Estacionários por janela no cenário 1.

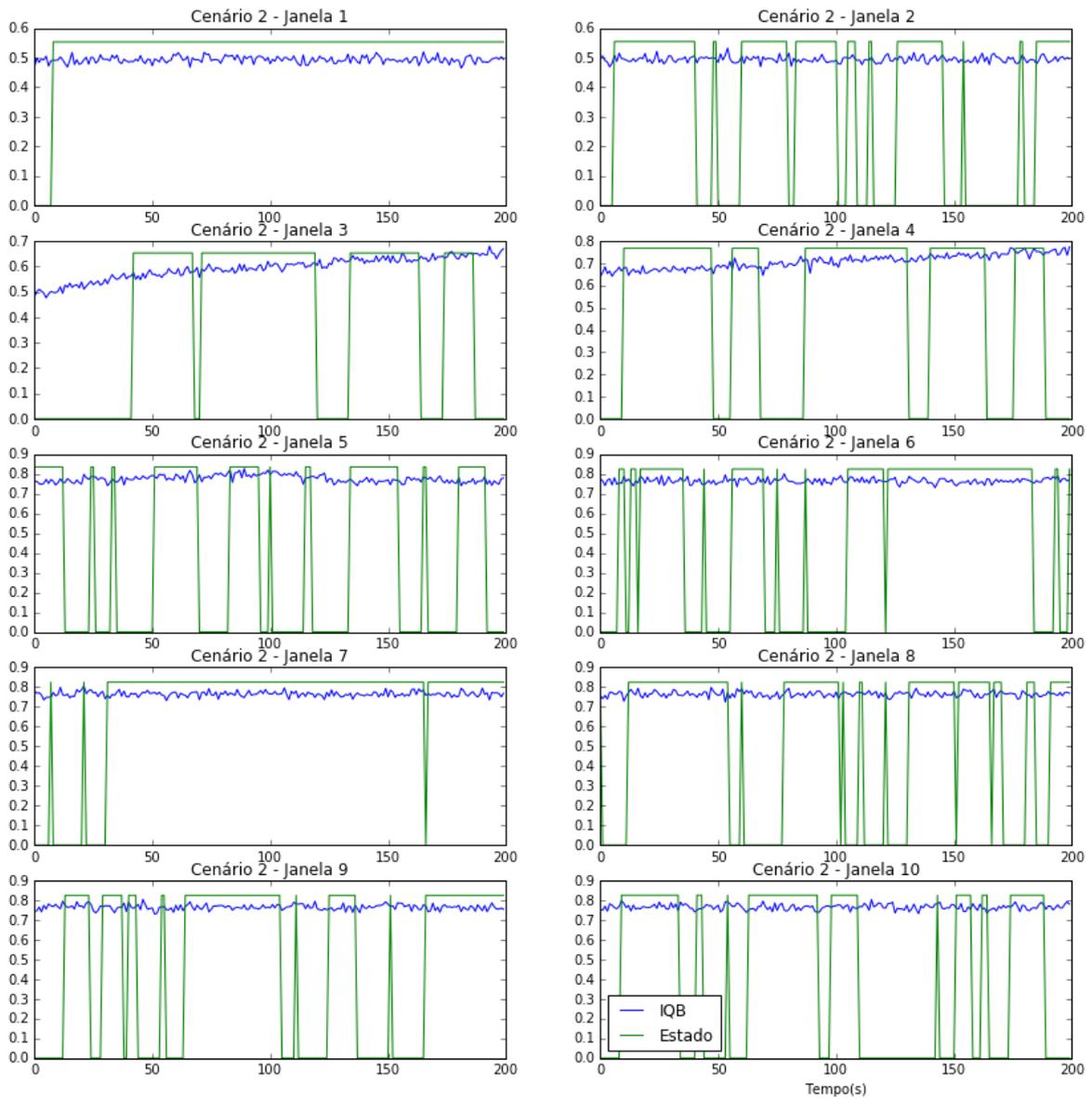


Figura A.2. Estacionários por janela no cenário 2.

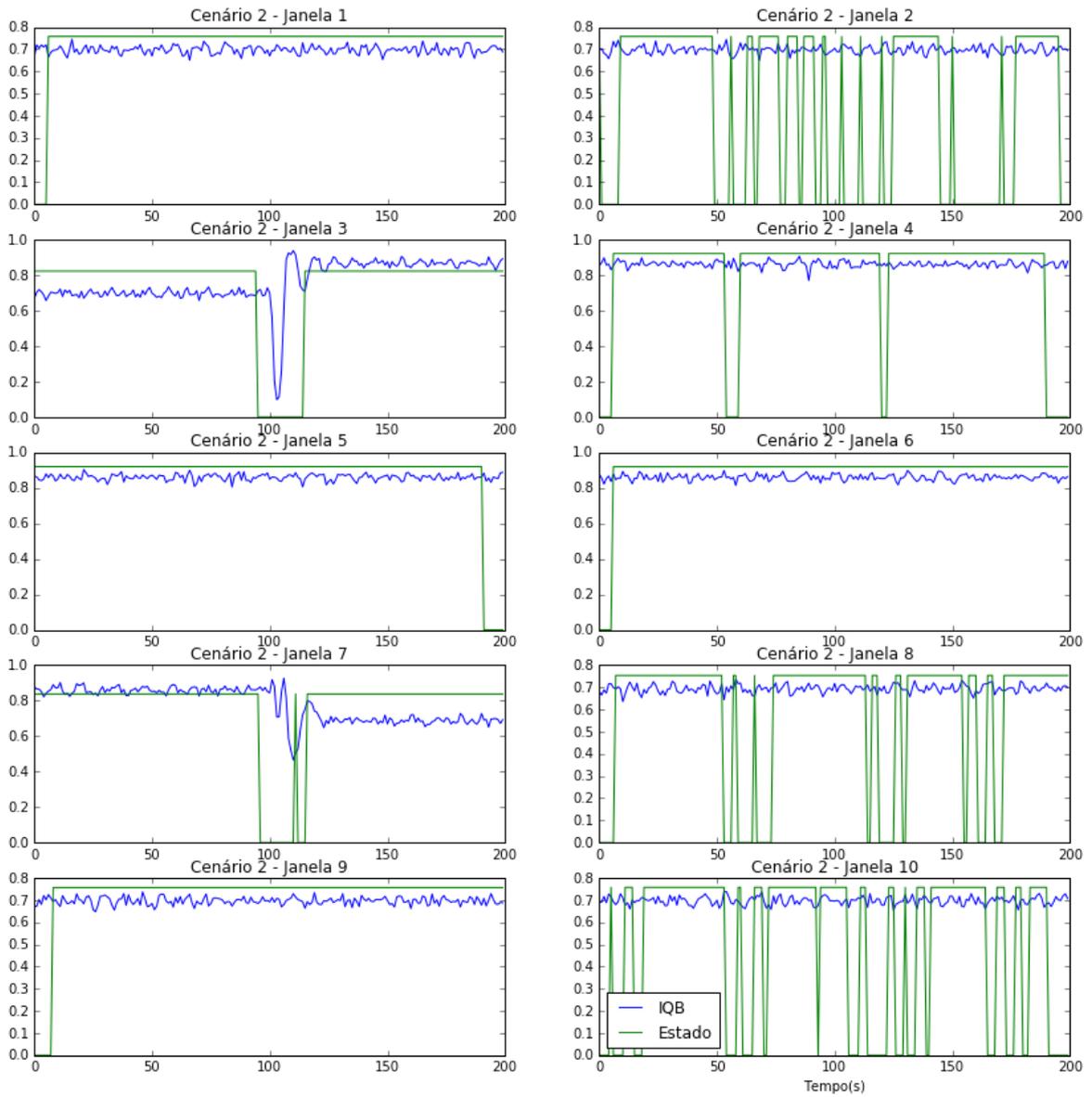


Figura A.3. Estacionários por janela no cenário 3.