

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

**Aplicação e Cálculo da Derivada de Sinais
de Processos Industriais**

DISSERTAÇÃO DE MESTRADO

Letícia Caumo

Porto Alegre

2006

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

Aplicação e Cálculo da Derivada de Sinais de Processos Industriais

Letícia Caumo

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

Área de concentração: Pesquisa e
Desenvolvimento de Processos

Orientador:
Prof. Dr. Jorge Otávio Trierweiler

Porto Alegre

2006

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 *Aplicação e Cálculo da Derivada de Sinais de Processos Industriais*, elaborada por Letícia Caumo, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:

Prof. Dr. Argimiro Resende Secchi

Dr. Pedro Rafael Bolognese Fernandes

Prof. Dr. Luís Gustavo Soares Longhi

Agradecimentos

A Deus.

À UFRGS e ao DEQUI, pelo apoio financeiro, pela estrutura e pelo corpo docente. Ao GIMSCOP, nas pessoas dos professores Jorge e Argimiro, pelo trabalho dedicado ao grupo. Ao prof. Jorge, que me ensinou a ser uma pessoa mais independente. À Petrobras, pelo financiamento deste projeto, e ao engenheiro Herbert Teixeira, por acreditar no potencial do grupo e pela condução do projeto.

Aos colegas do GIMSCOP, amigos que ganhei nesta convivência diária, pelo ânimo, pelas festas, pelas viagens e pelo apoio. À Vanessa, amiga sempre presente. Não poderia deixar de agradecer às “ladies biológicas”, pela amizade, pelas jantãs, pelos desabafos, pelas conversas, pelos almoços e cafés-da-tarde, pelo incentivo e pelos risos. E ao pessoal da sala 16, pela ajuda, pela força e pela amizade.

Aos amigos e colegas Ariel, Ricardo, Escobar, Farenzena, Pedro e Márcio, pelo apoio técnico e pela orientação.

À minha família, especialmente aos meus pais Gladir e Aidê, que entendem meu tempo, e que tudo tem o seu tempo. Seu apoio e carinho me deram forças para seguir neste caminho. Que seus valores e seu amor sempre me acompanhem.

Ao meu querido Ariel, pela sua simples presença, pela sua paciência, pela sua compreensão e pelo seu amor. Sua tranquilidade me acalmou nos momentos mais difíceis, e seu apoio e carinho me ajudaram a seguir em frente. Ari, obrigada por ser sempre tão companheiro e amigo. Agradeço também à sua família, por me acolher de forma tão especial.

Às minhas queridas maravilhosas amigas da “sexta-sim”, não poderia deixar de agradecer por tanto amor e união, mesmo que às vezes as distâncias sejam grandes. Cada elétron de vocês traz o brilho desta união, louca, insensata, irracional, ponderada, mágica, feliz.

Ao pai-postiço Gerson, que me salvou muitas vezes, mesmo sem saber.

Por fim, a todas as pessoas que colaboraram de alguma forma para este trabalho ou que me incentivaram neste projeto, meu obrigada e minha eterna gratidão.

Resumo

Metodologias para processamento de sinais têm sido amplamente pesquisadas com o objetivo de se extrair informação útil a partir de dados de processo. Nesta categoria incluem-se os filtros digitais, os quais atenuam os ruídos ou erros aleatórios, componentes de alta frequência do sinal. Técnicas de atenuação de ruído como aproximações polinomiais, filtros de Fourier e Ondaletas recebem destaque dentre as técnicas de filtragem. Os filtros das Ondaletas, por exemplo, baseiam-se nos componentes de frequência de um sinal, realizando a transformada deste sinal de modo a eliminar as frequências acima de um limite especificado. Este sinal suavizado (tendência), de ruído atenuado, também pode ser processado de modo a obter a sua derivada. A diferenciação numérica de sinais digitais tem muitas aplicações em processamento de sinais analíticos, devido às suas propriedades, e em várias técnicas de engenharia. Um sinal suavizado desempenha papel importante na identificação de estados estacionários por exemplo, que encontra grande aplicação em análise de processos, otimização, identificação de modelos e reconciliação de dados. Como estas aplicações requerem dados sob condição estacionária, ou muito próxima dela, um método eficiente de detecção de estacionários se faz necessário. A derivada de um sinal ruidoso fica mascarada pelo próprio ruído e não apresenta de forma clara os pontos de inflexão desejados para a análise do sinal. Faz-se necessário, portanto, tratar o sinal com um filtro de modo a reduzir ou eliminar o ruído antes de processá-lo ou, então, obter a tendência que este sinal apresenta. Neste trabalho busca-se comparar algumas metodologias comumente utilizadas para filtragem de sinais de processos e obtenção de tendências, e analisa-se a aplicação de uma destas metodologias para análise e identificação de estados estacionários. Uma metodologia multi-variável simples e eficiente baseada em PCA é proposta, tomando como referência o método proposto por JIANG *et al.* (2003), que determina um índice de estado estacionário pontual. Para validação do método, estudaram-se alguns sistemas industriais e os resultados obtidos são apresentados, mostrando-se satisfatórios.

Abstract

Methodologies for signal processing are being widely used to extract useful information from process data. In this category are included the digital filters, which attenuate noise or random errors (i.e., high frequency components of the signal). Techniques to attenuate noise as polynomial approximation, Fourier filters and Wavelets are commonly used. The Wavelet filters, for example, are based on the high frequency components of the signal, performing the wavelet transform on it to eliminate the frequencies above a specified level. This smooth signal, with less noise (trend), can be also processed to obtain its derivative. Numerical differentiation of digital signals has many applications in analytic signal processing, due to its properties, and in several engineering techniques. A smooth signal plays an important role in steady-state identification, for example, in applications for process analysis, process optimization, model identification and data reconciliation. As these applications require data under steady-state condition, or very close to it, an efficient method for steady-state detection is necessary. The derivative of a noisy signal is hidden by the noise itself and does not show clearly the inflexion points desired for the signal analysis. Thus it is necessary to filter the signal to reduce or even eliminate the noise before processing it. In this work, some commonly used filter techniques are used and compared, and the application of one of them to steady-state detection is analysed. An easy and efficient multivariable methodology based on PCA is proposed, according to the method presented by JIANG *et al.* (2003), which determines a pointwise steady-state index. To validate the method, some industrial systems were studied and the obtained results are presented, showing that they are satisfactory.

Sumário

1. Introdução	1
1.1 Histórico	3
1.2 Motivação.....	4
1.3 Estrutura da Dissertação.....	4
2. Revisão Bibliográfica	7
2.1 Erros de Medidas.....	7
2.2 Filtros do tipo Passa-Baixa.....	8
2.2.1 Filtros de Resposta Finita ao Impulso (FIR).....	10
2.2.2 Filtros Butterworth	12
2.3 Splines	14
2.3.1 B-splines.....	15
2.4 Ondaletas.....	16
2.4.1 Histórico	17
2.4.2 Transformada das Ondaletas	19
2.4.3 Comparação com a Transformada de Fourier.....	32
2.5 Análise por Componente Principal	33
2.5.1 Variáveis Escalonadas.....	36
3. Derivada de Sinais Ruidosos.....	37
3.1 Geração dos Sinais	37
3.2 Algoritmo para as Ondaletas.....	38
3.2.1 Determinação dos Parâmetros.....	40
3.3 Cálculo das Derivadas.....	40
3.4 Sinal 1	42
3.5 Sinal 2	44
3.6 Sinal 3	45
3.7 Sinal 4	46
3.8 Comparação dos Métodos	47
3.9 Conclusões	49
4. Detecção de Estados Estacionários	51
4.1 Metodologia de Referência	52
4.1.1 Sistemas Multi-variáveis.....	53
4.2 PCA e Sistemas Multi-variáveis	54
4.3 Estudos de Caso	55
4.3.1 Metodologias.....	55
4.3.2 Influência da Constante de Tempo.....	55
4.3.3 Influência do Período Histórico	58
4.3.4 Coluna de Destilação 1.....	62
4.3.5 Coluna de Destilação 2.....	72
4.3.6 Unidade de Craqueamento Catalítico.....	83
4.4 Conclusões	88

5. Conclusões.....	91
5.1 Considerações Finais.....	91
5.2 Sugestões para Trabalhos Futuros.....	92
Apêndice A - Considerações Matemáticas e Conceitos.....	95
A.1 Propriedades Matemáticas	95
A.2 Convolução	96
A.3 Invariância no Tempo.....	98
A.4 Efeitos de Borda.....	98
Apêndice B - Algoritmo das Ondaletas.....	101
B.1 Decomposição do Sinal (Transformada das Ondaletas).....	101
B.2 Remoção de Ruído Branco.....	103
B.3 Remoção de Picos Anormais.....	103
B.4 Reconstrução (Transformada Inversa das Ondaletas)	105
B.5 Valores de Corte para Detecção de Estacionários.....	107
Apêndice C - Propriedades das Ondaletas.....	109
C.1 Admissibilidade.....	109
C.2 Momentos Nulos	110
C.3 Suporte Compacto	112
C.4 Regularidade.....	112
C.5 Simetria	113
C.6 Energia.....	114
Referências Bibliográficas.....	115

Lista de Figuras

Figura 2.1: Desvios típicos de um filtro passa-baixa ideal quando aproximado por um filtro FIR, com $\omega_c = 0,4 \pi$: oscilações das bandas de passagem e rejeição (<i>passband ripple</i> e <i>stopband ripple</i>) e largura de transição (<i>transition width</i>).	11
Figura 2.2: Resposta em frequência para os filtros eletrônicos clássicos.	12
Figura 2.3: Resposta em frequência de um filtro Butterworth de primeira ordem.	14
Figura 2.4: Alguns exemplos de famílias de ondaletas. (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat.	17
Figura 2.5: Processo de decomposição: envolve etapas de deslocamento e posterior escalonamento (dilatação) da ondaleta (ALTMANN, 1996).	20
Figura 2.6: Representação das ondaletas nos domínios da frequência-escala e do tempo.	20
Figura 2.7: Representação de amostragens diádicas no plano tempo-escala.	22
Figura 2.8: Exemplo de um banco de filtros para um sinal cuja variação de frequência é de 0 a w_n (VALENS, 2004).	23
Figura 2.9: Banco de filtros do processo de decomposição (POLIKAR, 2001).	24
Figura 2.10: Filtros espelhados em quadratura (QMF).	26
Figura 2.11: L é invariante no tempo se as duas saídas forem iguais.	27
Figura 2.12: Esquema de computação dos filtros passa-baixa e passa-alta na transformada discreta estacionária das ondaletas.	28
Figura 2.13: Banco de filtros segundo o <i>Algorithme à Trous</i> proposto por Mallat. a) etapa de decomposição. b) etapa de reconstrução.	28
Figura 2.14: Função de escala <i>spline</i> cúbica e função ondaleta <i>spline</i> quadrática.	30
Figura 2.15: Representação da STFT.	33
Figura 2.16: Ilustração da técnica de PCA.	34
Figura 3.1: Geração do sinal para análise.	42
Figura 3.2: a) Sinal gerado e suas tendências. b) Derivadas.	43
Figura 3.3: Ampliação da Figura 3.2 entre o período de 80 e 240 para facilitar a visualização.	43
Figura 3.4: a) Sinal gerado e suas tendências. b) Derivadas.	44
Figura 3.5: a) Sinal gerado e suas tendências. b) Derivadas.	45
Figura 3.6: Ampliação da Figura 3.5 para facilitar a visualização.	46
Figura 3.7: Sinal com descontinuidades próximas.	46
Figura 3.8: Ampliação do sinal com descontinuidades próximas.	47
Figura 4.1: Sinal original e suas tendências a vários níveis de decomposição.	57
Figura 4.2: Para os 6 níveis de decomposição (ou $\tau = 2, 5, 10, 15, 30$ e 50): sinal original, tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B	58
Figura 4.3: Sinal original e suas tendências para diferentes períodos históricos selecionados.	59
Figura 4.4: Detecção de estacionários baseada em três períodos históricos distintos.	60
Figura 4.5: Sinal original e suas tendências para diferentes períodos históricos selecionados.	61
Figura 4.6: Índices de estado estacionário de acordo com as três tendências obtidas.	62
Figura 4.7: Coluna de destilação de tolueno e seus medidores.	63

Figura 4.8: Variáveis-chave selecionadas.....	64
Figura 4.9: Variável-chave temperatura do estágio 2 (TI02), representada por f , tendência do sinal fs , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B	64
Figura 4.10: Variável-chave temperatura do estágio 21 (TI21), representada por f , tendência do sinal fs , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B	65
Figura 4.11: Variável-chave pressão no topo (PI18), representada por f , tendência do sinal fs , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B	65
Figura 4.12: Variável-chave nível de fundo (LIC09), representada por f , tendência do sinal fs , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B	66
Figura 4.13: Índice de estado estacionário B_m para a coluna de destilação, obtido por seleção de variáveis-chave.	66
Figura 4.14: Perfil de temperatura da coluna de tolueno.	67
Figura 4.15: Temperaturas escalonadas da coluna de tolueno, deslocadas para facilitar a visualização.	67
Figura 4.16: Novas variáveis ortogonais t_1 e t_2	69
Figura 4.17: Resultados para a variável ortogonal t_1 , representada por f , fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado para t_1	70
Figura 4.18: Resultados para a variável ortogonal t_2 , representada por f , fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado para t_2	70
Figura 4.19: Índices de estado estacionário para as variáveis t_1 , t_2 e para a combinação B_m (para toda a coluna de destilação), obtidos pela utilização de PCA para combinação das variáveis.	71
Figura 4.20: Índices de estado estacionário obtidos pela utilização de variáveis- chave (B_{VC}) e PCA (B_{PCA}).	71
Figura 4.21: Coluna de destilação atmosférica, com as indicações de medições de temperatura e vazão.....	73
Figura 4.22: Variáveis-chave selecionadas.....	74
Figura 4.23: Resultados para a vazão de reciclo RCI, representada por f , fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	74
Figura 4.24: Resultados para a temperatura de topo, representada por f , fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	75
Figura 4.25: Resultados para a vazão de refluxo de diesel, representada por f , fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	75

Figura 4.26: Resultados para a temperatura de entrada, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	76
Figura 4.27: Índices de estado para as quatro variáveis-chave selecionadas (vazão de reciclo, temperatura do topo, refluxo de diesel e temperatura de entrada) e o índice global da coluna B_m	76
Figura 4.28: Temperaturas da coluna de destilação.	77
Figura 4.29: Vazões relacionadas à coluna de destilação.	78
Figura 4.30: Variáveis de temperatura e de vazão escalonadas e plotadas deslocadas para facilitar a visualização.	78
Figura 4.31: Variável ortogonal gerada a partir da combinação linear das temperaturas e vazões.	79
Figura 4.32: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	79
Figura 4.33: Temperaturas escalonadas e índice de estado estacionário B	80
Figura 4.34: Variáveis de temperatura e de vazão escalonadas.	81
Figura 4.35: Variável ortogonal que é uma combinação linear das temperaturas e vazões.	81
Figura 4.36: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	82
Figura 4.37: Temperaturas escalonadas e índice de estado estacionário B	82
Figura 4.38: Unidade de craqueamento catalítico de leito fluidizado (FCC).	84
Figura 4.39: Representação das variáveis de entrada e saída utilizando um modelo não-linear para a unidade de FCC.	85
Figura 4.40: Distúrbios inseridos em cada variável de entrada do modelo.	86
Figura 4.41: Variáveis de saída do modelo não-linear. a) temperaturas b) pressão e variáveis de qualidade dos produtos.	86
Figura 4.42: Variáveis de saída do modelo não-linear normalizadas, as quais são as variáveis de interesse para análise de estacionários.	87
Figura 4.43: Variável ortogonal gerada.	87
Figura 4.44: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.	88
Figura B.1: Decomposição do sinal em aproximações e detalhes.	102
Figura B.2: Sinal original e suas aproximações e detalhes, após aplicação de valores de corte (<i>soft-thresholding</i>).	103
Figura B.3: Tratamento de pico anormal: $f(t)$ são os dados originais, $f'(t)$ são os dados tratados por extensão simétrica, $WT(f)$ são os coeficientes de detalhe, T_1 é o valor de corte para identificação de pico anormal, T_2 é o valor de corte para cálculo da duração de pico anormal, t_a é o instante de tempo de início do pico anormal, t_b é o tempo final da duração do pico anormal, p_1 e p_2 são os	

pontos de sinais contrários nos coeficiente de detalhe que identificam um pico anormal.....	104
Figura B.4: Sinal com picos anormais e o sinal após tratamento com técnica de extensão simétrica e remoção de ruído.	105
Figura B.5: Sinal original e sinais reconstruídos aos níveis de reconstrução completa, nível 1 e nível 2.	106
Figura B.6: Sinal original, tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado estacionário B	107

Lista de Tabelas

Tabela 2.1: Respostas ao impulso dos filtros H , G e K	31
Tabela 2.2: Coeficientes de normalização λ_j para a ondaleta <i>spline</i> quadrática. Para $j > 5$, $\lambda_j = 1$	32
Tabela 3.1: Comparação do desempenho dos métodos: SPL (<i>Splines</i>), OND (Ondaletas) e BUT (filtro Butterworth). 1- Melhor, 2- Intermediário e 3- Pior ...	48
Tabela 4.1: Relação entre a constante de tempo e o número de decomposições.....	56

Capítulo 1

Introdução

A busca por metodologias que auxiliem na análise de sinais de processo tem sido cada vez maior. O conhecimento do comportamento de um sinal é de fundamental importância para aplicação de técnicas estatísticas, de auditoria, de otimização ou de identificação de sistemas. Por isso, técnicas de processamento de sinais têm sido amplamente desenvolvidas e estudadas.

Extrair informação útil de uma série de informações conflitantes é a forma mais simples de processamento de sinais. Genericamente, processamento de sinais é uma ferramenta criada para extrair, melhorar, armazenar ou transmitir informação útil. Como distinguir entre informação útil e indesejada é subjetivo, normalmente processamento de sinais depende da aplicação.

Uma das aplicações importantes do processamento de sinais digitais é o filtro digital, visto que medidas de processo são comumente afetadas por erros, que podem ser causados durante as medições, processamento ou transmissão do sinal medido. Dois tipos de erros são conhecidos: erros sistemáticos, no qual o valor medido é maior ou menor do que o valor “correto”, e erros aleatórios, no qual há variações não-preditas no sinal de medida em medida. Este último tipo de erro é conhecido como ruído, o qual provém de várias fontes, como flutuações de energia elétrica, correntes de ar, movimento térmico aleatório das moléculas, entre outras. Distinguir o ruído do sinal nas medidas é um grande problema. Enquanto o sinal possui alguma reprodutibilidade, o ruído não é o mesmo entre uma medida do sinal e a próxima, por exemplo.

Por isso, diminuir ou até mesmo suprimir os erros aleatórios, normalmente caracterizados por altas frequências, é uma tarefa desejada. Em muitos tipos de sinais, a amplitude real varia mais suavemente como uma função dos valores do eixo das ordenadas, enquanto muitos tipos de ruídos são rápidos, apresentando mudanças randômicas na amplitude de ponto a ponto de um dado sinal. Neste caso, uma prática comum para redução

do ruído é o uso de um processo conhecido por “suavização”. Nesta metodologia, o sinal é modificado de modo que os pontos maiores que pontos imediatamente adjacentes são reduzidos, e pontos que são menores que os imediatamente adjacentes são aumentados. Isto leva, naturalmente, a um sinal mais suave. Se o sinal é realmente suave, então o sinal real não será distorcido pela suavização, mas o ruído será removido. À medida que aumenta a suavização, maior é a redução de ruído, mas isto causa uma redução na amplitude dos picos e um aumento das suas larguras (O’HAVER, 2005). Existem outras maneiras de reduzir o ruído de um sinal, obtendo-se uma tendência, como a aplicação de interpolações ou aproximações polinomiais e ainda os filtros de Fourier ou das Ondaletas. Os filtros de Fourier e das Ondaletas baseiam-se nos componentes de frequência de um sinal e funcionam tomando a transformada deste sinal, eliminando todas as frequências acima de um limite especificado pelo usuário, e então realizando a transformada inversa. Assume-se que os componentes de frequência do sinal se localizem na região predominantemente de baixas frequências, e os componentes de ruído nas frequências altas. O usuário tenta achar uma frequência de corte que permita que a maior parte do ruído seja eliminada sem distorcer o sinal significativamente (O’HAVER, 2005).

Este sinal suavizado (tendência) também pode ser processado de modo a obter a sua derivada. A diferenciação numérica de sinais digitais tem muitas aplicações em processamento de sinais analíticos, devido às suas propriedades, e em várias técnicas de engenharia. Por exemplo, um sinal suave que apresente um ponto de inflexão oblíquo apresentará um pico máximo na sua derivada de primeira ordem e um “zero” (ponto onde o sinal cruza o eixo das abscissas) na derivada de segunda ordem. Este comportamento pode ser importante na localização precisa de um ponto de inflexão de uma curva sigmoideal, por exemplo, pela localização do zero da derivada de segunda ordem. Da mesma forma, a localização de um pico no sinal pode ser obtida pela verificação dos zeros na derivada de primeira ordem. Outras aplicações da derivação são para análise de alguns sinais típicos de análises térmicas e experimentos cinéticos que apresentam comportamento bem característicos (como de uma série de segmentos de linha reta com diferentes inclinações, onde o objetivo é determinar quantos segmentos de reta existem, qual é a inclinação destas retas, etc), diferenciação de espectros, análise de traços, entre outros.

Quando um sinal caracterizado por vários picos é analisado, o efeito da diferenciação deve ser levado em conta. Quanto maior a largura de um pico, menor será a amplitude da sua derivada, e este efeito fica ainda mais destacado quanto maior a ordem da derivação. Assim, a diferenciação distingue picos largos e estreitos, o que pode ser útil para aplicações analítico-quantitativas na detecção de picos superpostos, por exemplo (O’HAVER, 2005).

Portanto, distinguir os componentes de frequência de um sinal para extrair apenas a informação útil deste sinal é de extrema importância, como também é obter um sinal de processo mais suavizado, que traga informações mais claras sobre o processo e cuja derivada também fique mais suave, de modo que se possa facilitar a distinção entre os diferentes comportamentos do sinal.

Um sinal suavizado, com menos ruído, é de fundamental importância para ajudar na identificação de estados estacionários, que encontra grande aplicação em análise de processos, otimização, identificação de modelos e reconciliação de dados. Como estas aplicações requerem dados sob condição estacionária, ou muito próxima dela, um método eficiente de detecção se faz necessário.

1.1 Histórico

Existem vários métodos para se filtrar um sinal, obtendo sua tendência com menos ruído randômico. Como exemplo, na literatura, existe um método de obtenção de tendências e análise para monitoramento de sistemas e suporte à decisão (CHARBONNIER *et al.*, 2005), no qual as tendências são caracterizadas de três formas: crescente, estacionário e decrescente.

Com relação às ondaletas, que é um assunto relativamente novo no meio científico, elas têm sido largamente aplicadas para análise de imagens, devido a sua capacidade de resolução e compactação de informação sem perdas para a reconstrução (MALLAT e ZHONG, 1992). Em análise de sinais, são bastante utilizadas para identificação de anormalidades ou descontinuidades em sinais de processo (DUAN e OWEIS, 2005), sinais sísmicos e sinais eletrocardiográficos, principalmente.

Várias técnicas de filtração de sinais de processo baseadas em ondaletas surgiram ao longo dos últimos anos, como as propostas por DOYMAZ *et al.* (2001), AKBARYAN e BISHNOI (2000) e outra aplicada inclusive para estimação do tempo morto apresentada por CHING *et al.* (1999). Sinais sonoros também foram filtrados (WATSON e ADDISON, 2002) utilizando transformada das ondaletas baseada no módulo máximo.

No que se refere à detecção de estacionários, normalmente as técnicas são baseadas em métodos estatísticos, como em BROWN e RHINEHART (2000). NARASIMHAN *et al.* (1986) apresentaram um Teste Estatístico Composto (*Composite Statistical Test - CST*) e um Teste Matemático de Evidências (*Mathematical Test of Evidence - MTE*). No método CST, períodos consecutivos de tempo são definidos e avaliados de acordo com a matriz de covariância e média amostral. No método MTE, diferenças nas médias são comparadas com a variabilidade de cada período. Mais recentemente, CAO e RHINEHART (1995) propuseram um método baseado em média móvel ou filtro convencional de primeira ordem para reposicionar a média amostral.

Porém, estes métodos consideram apenas a presença de erros randômicos e não se mostraram eficientes em vários estudos práticos. Além disto, estes métodos indicam regiões estacionárias ou não, ou seja, não contemplam um índice que indique a situação do processo, a distância da condição estacionária. De modo a se calcular um índice pontual de estado estacionário, e não uma condição para um período, as ondaletas surgiram como uma alternativa de metodologia. FLEHMIG *et al.* (1998) propuseram um método baseado em ondaletas para identificar e localizar tendências polinomiais em medidas ruidosas. O trabalho propõe uma busca hierárquica de intervalos nos quais as medidas podem ser bem aproximadas

por um polinômio de grau limitado. Os intervalos são identificados como polinômios aproximados por mínimos quadrados e uma medida quantitativa da qualidade de ajuste é informada. Entretanto, esta técnica identifica regiões estacionárias e não-estacionárias, ou tendências lineares ou não. Outra técnica baseada em ondaletas para identificação de estacionários foi proposta por JIANG *et al.* (2003), na qual um índice é calculado para cada instante de tempo, sendo que este índice assume valores entre 0 e 1. Esta técnica, diferentemente da proposta por FLEHMIG *et al.* (1998), quantifica a proximidade do estacionário através de um índice, permitindo que se conheça o real comportamento do processo.

1.2 Motivação

Técnicas em tempo real baseadas em modelos, como reconciliação de dados, detecção de erros grosseiros ou otimização em tempo real, ganharam significativa importância ao longo dos últimos anos, de modo que a identificação de tendências nas medidas de processo para aplicações como em detecção de estados estacionários tem sido uma tarefa importante.

Estes modelos são geralmente estacionários, o que justifica a importância de conhecer quais são, de fato, os dados de processo estacionários. Para outras aplicações, como visualização ou detecção de flutuações no processo, identificar e localizar tendências polinomiais lineares ou de alta ordem pode ser bastante importante, da mesma forma como distinguir entre estas flutuações e estados estacionários.

Assim, técnicas baseadas em ondaletas ganharam destaque em processamento de sinais. Apesar do desempenho superior para certas aplicações, a grande vantagem dos métodos baseados em ondaletas é sua capacidade de concentrar em uma única técnica uma variedade de tarefas de processamento de sinais, de modo que esta sinergia aumenta o desempenho da solução geral em comparação com a solução de cada tarefa individualmente (FLEHMIG *et al.*, 1998).

Em virtude disto, uma metodologia baseada em ondaletas (JIANG *et al.*, 2003) foi avaliada. A representação dos dados em várias escalas tem se apresentado como uma grande habilidade quando utilizada para filtrar dados, pois apresenta a vantagem de ajudar na separação do ruído de características importantes do sinal. Além do mais, a técnica é utilizada para detecção de estacionários mediante o cálculo de um índice que indica a distância da condição estacionária. Este índice é mais um fator importante que motivou o emprego da metodologia proposta por JIANG *et al.* (2003) e, além disto, as técnicas baseadas em métodos estatísticos apresentam-se pouco robustas e pouco flexíveis. Estes métodos estatísticos avaliam o estado do processo em um período de tempo, ao invés de avaliarem pontualmente no tempo. Para aplicações em tempo real isto é, obviamente, uma condição fundamental.

1.3 Estrutura da Dissertação

Além deste capítulo introdutório, que buscou apresentar uma revisão histórica da utilização das tendências e derivadas de sinais e a motivação para este trabalho, esta dissertação apresenta mais quatro capítulos. O Capítulo 2 contempla uma breve revisão baseada nos principais trabalhos publicados a respeito do tema, procurando descrever de forma didática os principais conteúdos abordados em cada etapa deste trabalho.

O Capítulo 3 faz um estudo comparativo das técnicas que são comumente utilizadas para eliminação ou diminuição do ruído presente nas medidas de processo. São comparados sinais frequentemente encontrados em aplicações práticas com relação à obtenção de tendências destes sinais e suas derivadas. Esta comparação inclui os métodos de filtros Butterworth, aproximação por splines e filtração por ondaletas.

O Capítulo 4 consiste na aplicação de um algoritmo baseado em ondaletas para a obtenção de um sinal de tendência, suas derivadas e a detecção de pontos estacionários nestes sinais. Os parâmetros de entrada deste algoritmo são avaliados quanto a sua influência nos resultados, ou seja, uma análise de sensibilidade destes parâmetros é analisada. Para sistemas multi-variáveis, um método baseado em PCA (Análise por Componente Principal) é proposto, e sua aplicabilidade é analisada em exemplos reais. São analisadas duas colunas de destilação e uma unidade de craqueamento catalítico.

Por fim, o Capítulo 5 sumariza as principais conclusões sobre o trabalho desenvolvido. As contribuições desta tese são apresentadas, bem como sugestões para trabalhos que poderiam dar continuidade a ele.

Apêndices foram acrescentados ao final para completar informações sobre alguns conteúdos apresentados durante o desenvolvimento deste trabalho. Algumas propriedades matemáticas e conceitos relevantes para o trabalho são apresentados no Apêndice A. No Apêndice B, uma idéia do algoritmo baseado nas ondaletas utilizado para obtenção das tendências é apresentado. O Apêndice C mostra algumas propriedades das ondaletas, as quais são importantes para definição de qual tipo de ondaleta deve ser escolhida de acordo com a finalidade/aplicação desejada.

Capítulo 2

Revisão Bibliográfica

Neste capítulo são apresentadas algumas metodologias que foram utilizadas no desenvolvimento do trabalho. Alguns conceitos que serão abordados nos demais capítulos serão aqui brevemente apresentados, tendo como foco apenas os aspectos que serão necessários para compreensão posterior. Assim, será dado um enfoque amplo, destacando os aspectos importantes utilizados neste trabalho.

2.1 Erros de Medidas

Erros nas medidas de processo são comuns na instrumentação de processos, e podem ser causados durante as medições, processamento e transmissão do sinal medido. O erro e é a diferença entre o valor medido y e o valor real da variável x , conforme Eq. (2.1). Alguns erros são aleatórios (erros randômicos), outros sistemáticos ou erros grosseiros. A soma das contribuições destes erros caracteriza o erro total (NARASIMHAN e JORDACHE, 2000).

$$y = x + e \quad (2.1)$$

Erros randômicos não podem ser completamente eliminados e estão sempre presentes nas medidas. Como o próprio nome indica, nem a magnitude nem o sinal podem ser preditos, de forma que, se a medição é repetida com o mesmo instrumento sob condições idênticas de processo, um valor diferente pode ser obtido dependendo do efeito do erro randômico. Erros randômicos normalmente correspondem a componentes da alta frequência do sinal medido, e normalmente são de pequena magnitude exceto por alguns pulsos ocasionais (*spikes*). Estes erros podem ser causados por um número diferente de fontes como flutuações no suprimento de energia, rede de transmissão, filtração de entrada analógica, mudanças nas condições ambientais, etc (NARASIMHAN e JORDACHE, 2000).

Erros grosseiros, por outro lado, estão associados a eventos não-aleatórios, significando que, se a medição é repetida com o mesmo instrumento sob condições idênticas de processo, a contribuição do erro grosseiro para o valor medido será a mesma. Erros grosseiros estão relacionados, portanto, ao mau funcionamento de instrumentos (devido à instalação imprópria dos instrumentos de medida), falta de calibração, corrosão nos sensores e depósitos de sólidos. Mantendo bons procedimentos de instalação e manutenção, é possível assegurar que erros grosseiros não estejam presentes nas medidas pelo menos por algum tempo. Erros grosseiros causados por falta de calibração nos sensores podem ocorrer rapidamente num tempo particular e depois permanecer num nível ou magnitude constantes. Outras causas de erros grosseiros, como desgaste ou obstrução dos sensores, podem ocorrer gradualmente em um período de tempo e então a magnitude dos erros grosseiros aumenta lentamente em um período relativamente longo de tempo. São representados, então, por algum viés (*bias*) sistemático nas medidas e vazamentos (*leaks*). Portanto, erros grosseiros ocorrem com menos frequência, mas sua magnitude é tipicamente maior do que erros aleatórios. A identificação de erros grosseiros é muito mais difícil do que de erros aleatórios. Alguns dos métodos existentes são o Teste Global, Teste de Componente Principal e Teste de Medida. Porém, os métodos existentes normalmente não são satisfatórios para todos os cenários que se apresentam (NARASIMHAN e JORDACHE, 2000).

Na tentativa de diminuir ou atenuar os erros aleatórios, normalmente caracterizados por frequências mais altas, apresentou-se vários métodos de redução de ruído ou extração de tendências. Alguns tipos de filtros usualmente empregados são filtros exponenciais, de média móvel e polinomiais. Nas seções seguintes, apresentam-se mais detalhadamente os filtros do tipo passa-baixa, *splines* e ondaletas, os quais serão utilizados posteriormente em estudos de caso para verificação de suas aplicações, vantagens e desvantagens.

2.2 Filtros do tipo Passa-Baixa

O processo de filtragem de sinais consiste em reduzir erros nos dados de processo. Um filtro pode selecionar, ou suprimir, certos componentes de frequência de um sinal. Um filtro passa-baixa (*lowpass filter*) atenua, ou reduz, a amplitude das frequências maiores que um valor estabelecido, chamado de frequência de corte ω_c , permitindo a passagem apenas das frequências baixas. Em geral, a largura da banda de passagem dá uma idéia das propriedades transientes da resposta, visto que uma grande largura da banda de passagem corresponde a uma resposta mais rápida (LEVINE, 1996).

Um filtro passa-baixa ideal possui como característica um ganho constante unitário até a frequência de corte ω_c , e então o ganho subitamente cai a zero para $\omega > \omega_c$ (LATHI, 1998). Ou seja, o filtro passa-baixa permite que os componentes do sinal de frequência abaixo de ω_c passem, e retém os componentes do sinal com frequência acima de ω_c . Um filtro passa-baixa ideal é um filtro de ordem infinita. Assim, quanto maior a ordem de um filtro, mais ele se aproxima do ideal (LATHI, 1998).

Os filtros digitais¹, além de selecionar componentes de frequências, podem atenuar (ou suprimir) cada componente de frequência por uma quantia desejada. São caracterizados pela sua resposta a um impulso (um sinal consistindo de um único valor e seguido de zeros): a resposta ao impulso é uma indicação de quanto tempo o filtro leva para chegar a um estado estacionário, sendo uma indicação da estabilidade do filtro, visto que uma resposta ao impulso que oscila continuamente por um longo tempo indica que o filtro pode ter tendência à instabilidade (BORES, 2004).

A saída de um filtro digital é uma composição das últimas entradas e das últimas saídas, segundo a Eq. (2.2).

$$y[n] = \sum c[k] * x[n-k] + \sum d[j] * y[n-j] \quad (2.2)$$

onde $y[n]$ é a saída do filtro no instante de tempo n , $c[k]$ e $d[j]$ são coeficientes, $x[n-k]$ é a entrada em k instantes passados, $y[n-j]$ são as últimas j saídas do filtro e $*$ denota a operação de convolução² (Apêndice A).

Duas convoluções estão envolvidas, portanto: uma das últimas entradas e outra das últimas saídas. Em cada caso, a função de convolução é chamada de "coeficientes" dos filtros (BORES, 2004).

Como a filtragem é um processo de seleção de frequências, é importante pensar em um filtro digital com relação a sua resposta no domínio da frequência. Neste caso, a resposta na frequência do filtro pode ser calculada a partir da equação do filtro, dada pela transformada- z de sua resposta ao impulso. Em geral, a transformada- z $Y(z)$ da saída $y[n]$ de um filtro digital está relacionada com a transformada- z da entrada $X(z)$ por

$$Y(z) = H(z)X(z) = \frac{b(1) + b(2)z^{-1} + \dots + b(n_b + 1)z^{-n_b}}{a(1) + a(2)z^{-1} + \dots + a(n_a + 1)z^{-n_a}} X(z) \quad (2.3)$$

onde $H(z)$ é a função de transferência do filtro. Portanto, as constantes $b(i)$ e $a(i)$ são os coeficientes do filtro e a ordem do filtro é o máximo entre n_a e n_b .

¹ Um filtro digital é um filtro que processa sinais digitais, sendo, portanto implementado através de software. Utilizando um conversor analógico-digital, digitaliza-se o sinal a filtrar, em seguida este é processado por um microprocessador, onde está programado o filtro digital. Para se obter de novo um sinal analógico já filtrado, coloca-se um conversor digital-analógico (LEVINE, 1996).

² Qualquer filtro LTI (linear time invariant filter) pode ser implementado pela convolução do sinal de entrada pela resposta ao impulso do filtro. Convolução é um operador matemático que, a partir de duas funções, produz uma terceira. O conceito de convolução está ligado ao de média móvel e é fundamental no estudo de sistemas LTI.

Quando $n_a = 0$ (ou seja, a é um escalar), o filtro é do tipo Resposta Finita ao Impulso (*Finite Impulse Response - FIR*), não-recursivo, ou filtro de média móvel (NARASIMHAN e JORDACHE, 2000).

Para aplicação em tempo real, estes filtros não existem, pois o filtro teria que prever o futuro e ter conhecimento infinito do passado para realizar a convolução. Isto é possível para sinais digitais pré-gravados, ou perfeitamente cíclicos, que se repetem infinitamente. Para aplicações em tempo real, os filtros reais aproximam-se do filtro ideal por atrasarem o sinal por um período de tempo, permitindo uma pequena "visão" do futuro. Isto é manifestado como a mudança de fase. Uma maior precisão na aproximação requer um atraso maior.

Os filtros de média móvel são bons para codificar sinais no domínio do tempo. Entretanto, são os piores filtros para codificar sinais no domínio da frequência, pois possuem pequena capacidade de separar uma banda de frequências da outra (SMITH, 1997).

2.2.1 Filtros de Resposta Finita ao Impulso (FIR)

É muito mais fácil resolver o problema de cálculo dos coeficientes do filtro se há uma simplificação da equação do filtro de modo que se trabalhem somente com as últimas entradas, i.e., excluir a possibilidade de retroalimentação (*feedback*). Isto equivale a desprezar o segundo termo da soma da Eq. (2.2) anteriormente apresentada e tornar o filtro não-recursivo. Isto implica em considerar apenas as entradas passadas, porém muito mais valores passados são necessários para "compensar" o efeito de não necessitar de dados passados do sinal de saída (WHITE, 2000).

Se tal filtro é então sujeito a um impulso, então sua saída deve necessariamente ir a zero depois que o impulso tenha passado. Então a resposta ao impulso de tal filtro deve ter duração finita necessariamente (KRAUSS *et. al*, 1999). Tal filtro é chamado de filtro de Resposta Finita ao Impulso, ou FIR, e é representado por

$$\begin{aligned}
 y(n) &= b_0 x(n) + b_1 x(n-1) + b_2 x(n-2) + \dots + b_M x(n-M) \\
 &= \sum_{k=0}^M b_k x(n-k) \\
 &= \sum_{k=-\infty}^{\infty} h(k) x(n-k) \\
 &= (h * x)(k)
 \end{aligned}
 \tag{2.4}$$

onde o atraso $k = 0, 1, 2, \dots, M$ e $x(n)$ é a amplitude do sinal de entrada no tempo n , e $*$ denota a convolução entre os sinais h e x .

A resposta do filtro na frequência é simplesmente a transformada de Fourier dos coeficientes do filtro. Assim, os coeficientes de um filtro FIR podem ser calculados simplesmente pela transformada inversa de Fourier da resposta na frequência desejada (SMITH, 2004).

Estes tipos de filtros apresentam algumas vantagens: não levam a *overshoot*, alcançando corretamente estado-estacionário após uma mudança do tipo degrau; podem ter sua fase especificada como linear; são fáceis de implementar e são rápidos computacionalmente; sendo úteis para estimar um valor fixo ou tendência linear (NARASIMHAN e JORDACHE, 2000).

No entanto, na prática, os filtros passa-baixa ideais não são possíveis de serem projetados. A aproximação finita a estes filtros leva as bandas do filtro de passagem ($\omega < \omega_c$) e de rejeição ($\omega > \omega_c$) a apresentarem oscilações, bem como a terem uma largura de transição não-zero entre si (veja **Figura 2.1**), os quais são efeitos indesejáveis (LOSADA, 2004).

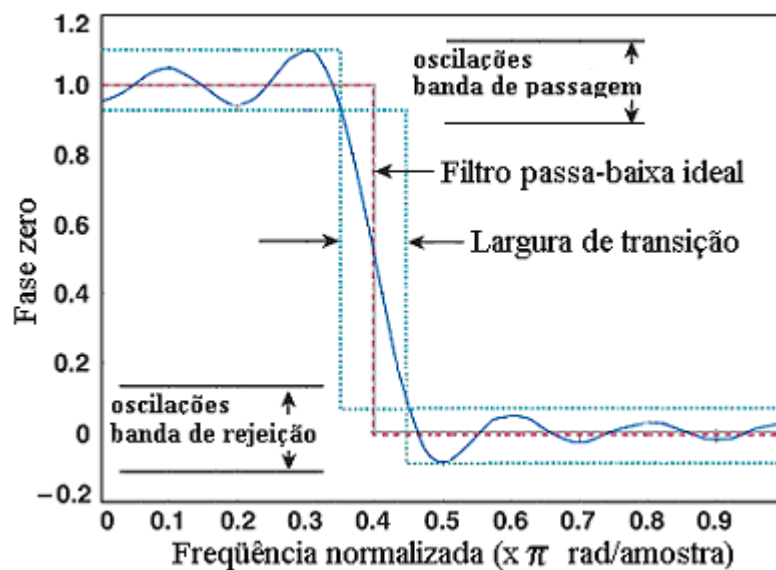


Figura 2.1: Desvios típicos de um filtro passa-baixa ideal quando aproximado por um filtro FIR, com $\omega_c = 0,4 \pi$: oscilações das bandas de passagem e rejeição (*passband ripple* e *stopband ripple*) e largura de transição (*transition width*).

Para um projeto de filtro FIR o que se seleciona, então, são as faixas de oscilação de banda e faixa de transição que não excedam determinado valor. Pode-se também selecionar a ordem do filtro, ou equivalentemente, o comprimento da resposta ao impulso truncada.

Um filtro desejável é o que mais se aproxima do ideal e que apresenta resposta de fase linear (LOSADA, 2004). Isto significa que a resposta do filtro é uma função linear da frequência e o filtro, assim, gera um atraso, mas não gera distorção de fase, que é uma vantagem deste tipo de filtro. Um filtro FIR é de fase linear se seus coeficientes são simétricos em relação ao coeficiente central, ou seja, o primeiro coeficiente é igual ao último, o segundo é igual ao penúltimo, e assim por diante.

Um tipo de filtro FIR são os filtros de interpolação, que são filtros de banda limitada. O conceito é baseado na noção de um sinal contínuo no tempo sendo amostrado. A interpolação ideal de banda limitada tomará o sinal digital (amostrado) e produzirá uma

interpolação do sinal que será idêntica ao sinal que seria obtido a uma taxa de amostragem maior. Na realidade, é colocado em prática por meio de aumento de amostragem e uso de um filtro passa-baixa ideal (LOSADA, 2004).

Dentre os filtros de interpolação existe o filtro de meia-banda (*halfband filters*). Este é um filtro de Nyquist (*Lth-band filters*) particular, pois a banda de passagem da sua magnitude de resposta ocupa aproximadamente $1/L$ do intervalo de Nyquist. No caso especial de uma interpolação por um fator de 2, os filtros são chamados de meia-banda. Filtros de meia-banda são normalmente usados quando a interpolação (ou decimação) é de um fator de 2. A frequência de corte destes filtros é sempre $0,5 \pi$. Além disso, as ondulações da banda de passagem e da banda de rejeição são idênticas.

2.2.2 Filtros Butterworth

O filtro Butterworth é um tipo de projeto de filtro eletrônico e é desenvolvido de modo a ter uma resposta em frequência o mais plana o quanto for matematicamente possível na banda de passagem. Uma resposta plana é aquela que não apresenta ondulações na banda de passagem, e se aproxima do zero na banda de rejeição (KRAUSS *et. al*, 1999).

Quando visto em um gráfico logarítmico, esta resposta desce linearmente até o infinito negativo. Isto significa que os filtros Butterworth possuem uma queda na sua magnitude como uma função linear com ω , sendo que o filtro de Butterworth é o único filtro que mantém o mesmo formato para ordens mais elevadas (porém com uma inclinação mais íngreme na banda atenuada) enquanto outras variedades de filtros (Bessel, Chebyshev, Elíptico) possuem formatos diferentes para ordens mais elevadas (KRAUSS *et. al*, 1999). A **Figura 2.2** apresenta os filtros eletrônicos mais clássicos.

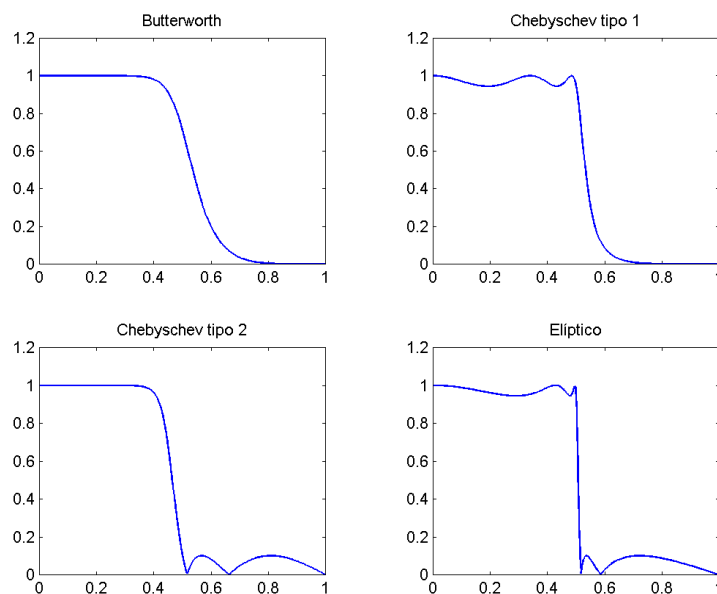


Figura 2.2: Resposta em frequência para os filtros eletrônicos clássicos.

Na **Figura 2.2**, o primeiro filtro é o de Butterworth, que é o mais suave, não apresentando oscilações. O último é o filtro Elíptico, que é o mais brusco (menor largura de transição), mas apresenta oscilações tanto na banda de passagem quanto na de rejeição. Os dois filtros Chebyshev do meio têm um comportamento intermediário, sendo mais ou menos bruscos e com oscilações em parte do espectro. Os gráficos foram obtidos usando os mesmos parâmetros e o mesmo número de coeficientes, mostrando, portanto, somente as propriedades dos diferentes filtros (KRAUSS *et. al*, 1999).

Observa-se que o filtro de Butterworth é mais plano que os outros não apresentando oscilações. Entretanto, sua queda é um pouco mais lenta, requerendo, assim, uma ordem maior para implementar uma banda de rejeição específica. Por outro lado, sua resposta em fase será mais linear na banda de passagem do que os outros filtros.

Nos filtros Butterworth, de ordem n , a amplitude de resposta $|H(j\omega)|$ é dada por

$$|H(i\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}} \quad (2.5)$$

onde i é a constante imaginária, ω é a frequência angular do sinal (rad/s) e ω_c é a frequência de corte (OPPENHEIM e WILLSKY, 1997).

Pode-se observar que em $\omega = 0$ o ganho $|H(0)|$ é unitário e em $\omega = \omega_c$ o ganho $|H(i\omega_c)|$ é igual a $1/\sqrt{2}$ ou -3 dB. Ou seja, a frequência de corte é aquela na qual a magnitude da resposta cai para 0,707 do valor na frequência zero. Como o ganho tem uma queda de um fator $\sqrt{2}$ em $\omega = \omega_c$ e a energia é proporcional ao quadrado da amplitude, a razão entre as energias (de saída pela de entrada) cai por um fator de 2 em $\omega = \omega_c$. Por esta razão, ω_c é chamada de frequência de meia-potência (*half-power frequency*) ou frequência de corte de 3 dB (LATHI, 1998; LEVINE, 1996; OPPENHEIM e WILLSKY, 1997).

A **Figura 2.3** apresenta a resposta em frequência de um filtro de Butterworth de primeira ordem normalizado (frequência de corte normalizada para 1 rad/s e ganho normalizado para 0 dB na banda de passagem), a qual mostra o ganho e a fase. Como se observa, o ganho tem uma queda de -20 dB a cada decado³ (-6 dB por oitava).

No Matlab®, os filtros Butterworth foram testados utilizando-se a função *idfilt*, que realiza uma filtragem no sentido direto e depois no sentido reverso. É necessário informar a ordem do filtro, em termos da frequência de Nyquist, e a frequência de corte.

³ Decado representa uma razão entre frequências, correspondente a um valor de 10. Por exemplo, a amplitude pode cair de -20 dB entre as frequências de 10 e 100 Hz ou entre as frequências de 2.5 e 25 Hz (LEVINE, 1996).

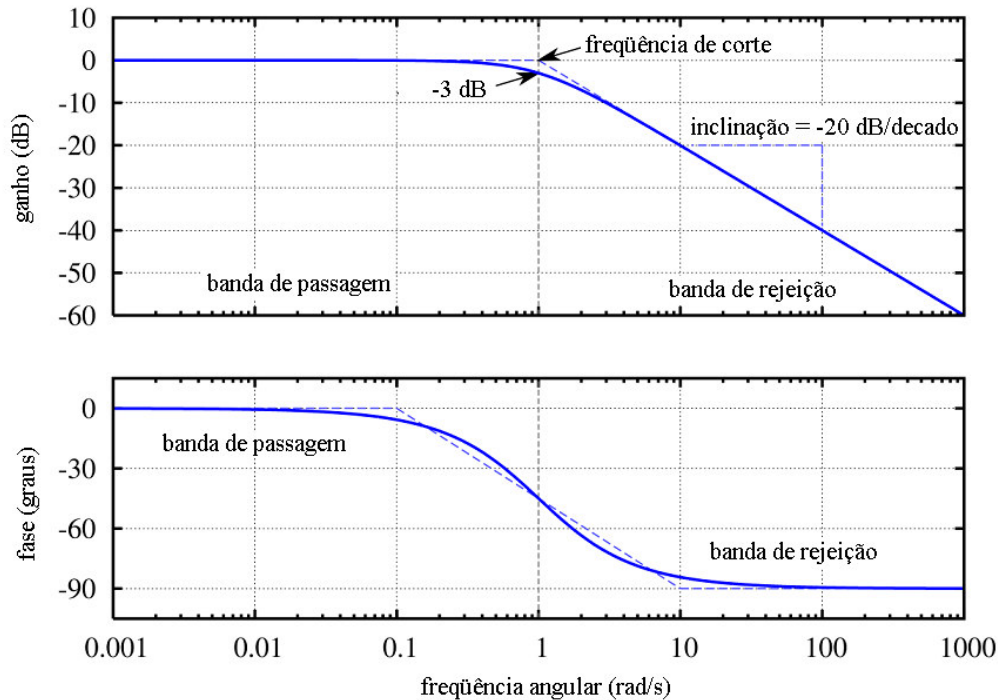


Figura 2.3: Resposta em frequência de um filtro Butterworth de primeira ordem.

2.3 Splines

Splines são funções polinomiais que podem ser usadas para ajuste de uma curva ou função aos dados. Um ajuste exato aos pontos de dados envolve interpolação, enquanto um ajuste aproximado envolve aproximação por mínimos quadrados, resultando em um sinal suavizado (BOOR, 1999). Assim, para *splines* de interpolação, a curva passa por cada ponto, dito "ponto de controle", enquanto para *splines* de aproximação, não necessariamente.

As *splines* são, por natureza, de dimensão finita, o que é a razão principal da sua utilidade em computação e representação. Também são de simples construção, avaliação fácil e acurada, e possuem capacidade de aproximar formas complexas através de um ajuste de curva e projeto iterativo.

Se uma função deve ser aproximada em um grande intervalo, o grau do polinômio de aproximação deverá ser inaceitavelmente grande. A alternativa é subdividir o intervalo de aproximação $[a, b]$ em intervalos suficientemente pequenos $[x_j, x_{j+1}]$ (com $a = x_1 < \dots < x_{l+1} = b$) de modo que, em cada intervalo, um polinômio de grau relativamente baixo possa trazer uma boa aproximação para a função. Qualquer seção de função polinomial suave é chamada *spline* (BOOR, 1999).

Os pontos de união entre estes subintervalos são denominados de "nós". Se a união entre dois subintervalos acontece da maneira mais suave possível, caracteriza-se uma *spline* conhecida por B-*spline*, e normalmente de ordem cúbica. Para isto, as derivadas de primeira e segunda ordem ao final da curva devem ser iguais às derivadas de primeira e segunda ordem

da próxima. Para se assegurar que isto aconteça, as *splines* cúbicas naturais computacionalmente envolvem a resolução simultânea de um sistema de equações.

2.3.1 B-splines

A forma básica das *splines* (B-form ou *basic splines*) tem sido a forma padrão de representar uma *spline* visto que busca suavizações nos nós. A forma básica (B-form) descreve uma *spline* como combinação linear de beta *splines* (B-splines). Assim, podem ser representadas conforme a Eq. (2.6).

$$f = \sum_{i=1}^n B_{ik} a_i \quad (2.6)$$

onde $B_{ik} = B(t_i, \dots, t_{i+k})$ é a i -ésima B-*spline* de ordem k para a seqüência de *knots* $t_1 \leq t_2 \leq \dots \leq t_{n+k}$, onde *knots* são os valores da abscissa nos nós. Assim, uma *spline* f é especificada por sua seqüência t de *knots* e por sua seqüência de coeficientes B-*splines* a , sendo que os coeficientes são chamados de "pontos de controle" da curva. Uma B-*spline* é nula fora do seu intervalo básico (BOOR, 1999).

Uma das vantagens da B-*spline* é que o grau do polinômio é independente do número de pontos. Assim, a forma é localmente controlada. Porém, possuem grande complexidade computacional.

Uma *spline* cúbica natural, muito utilizada, é um polinômio cúbico para cada seção de modo que a função, sua derivada e sua segunda derivada são contínuas nos nós de interpolação, e onde "natural" significa que a segunda derivada é definida nula nos extremos do intervalo de interpolação.

$$S''(a) = S''(b) = 0 \quad (2.7)$$

Isto obriga a *spline* a ser uma linha contínua reta fora do intervalo, sem comprometer sua suavidade. É, então, a mais suave possível curva de interpolação, visto que minimiza a integral do quadrado da derivada segunda (BOOR, 1999).

Uma *spline* típica de ordem n , usada em análise numérica, tem continuidade

$$S(t) \in C^{n-1}[a, b] \quad (2.8)$$

Isto significa que quaisquer dois subintervalos polinomiais adjacentes encontram-se no valor e nas suas $n-1$ derivadas em cada *knot*. *Splines* com poucos *knots* são geralmente mais suaves que *splines* com muitos *knots*. Entretanto, o aumento do número de *knots* normalmente aumenta o ajuste da função *spline* aos dados. *Knots* dão à curva liberdade para mudar de direção de maneira a se aproximar mais dos dados (SAS, 1999). A mais comum das *splines* é cúbica ($n = 3$), duas vezes diferenciável (C^2).

A função do Matlab® utilizada neste trabalho foi *spaps*, que trabalha com *splines* suaves (de aproximação). A função retorna a B-*spline* cuja distância da função f aos dados é menor do que um valor de tolerância tol (BOOR, 1999).

$$E(f) = \sum_{j=1}^n w_j (y(j) - f(x(j)))^2 \leq tol \quad (2.9)$$

onde w_j são pesos calculados, ou informados como argumento de entrada, de modo que $E(f)$ seja uma aproximação para $\int_{x(1)}^{x(n)} (y - f)^2 dt$ segundo a Regra Trapezoidal Composta.

2.4 Ondaletas

As ondaletas são ondas de energia finita, de integral nula, conforme Eq. (2.10), que podem ser simétricas ou assimétricas, suaves ou não, de expressão matemática simples ou não (DAUBECHIES, 1992). Alguns conceitos básicos sobre as principais propriedades das ondaletas pode ser visto no Apêndice C. Veja alguns exemplos de tipos de ondaletas na **Figura 2.4**.

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.10)$$

A idéia na qual se baseia a transformada das ondaletas é originada de Fourier: aproximar uma função complexa por uma soma ponderada de funções mais simples, que são obtidas de uma função protótipo simples. A função protótipo, também conhecida por função base, pode ser imaginada como um tijolo de construção, e a função original pode ser aproximada, ou sob certas condições plenamente representada, usando tijolos de construção similares, com larguras e alturas diferentes (POLIKAR, 1999).

Destaca-se aqui que "ondaletas" é a tradução adotada neste trabalho para *wavelets*. Ondaletas seriam ondas pequenas (*ondelettes*, do francês). Não está ainda bem definida, no meio científico nacional, uma correta tradução para a Língua Portuguesa. Alguns autores utilizam ainda o termo em inglês e outros a traduzem como "ondeletas", devido à palavra francesa.

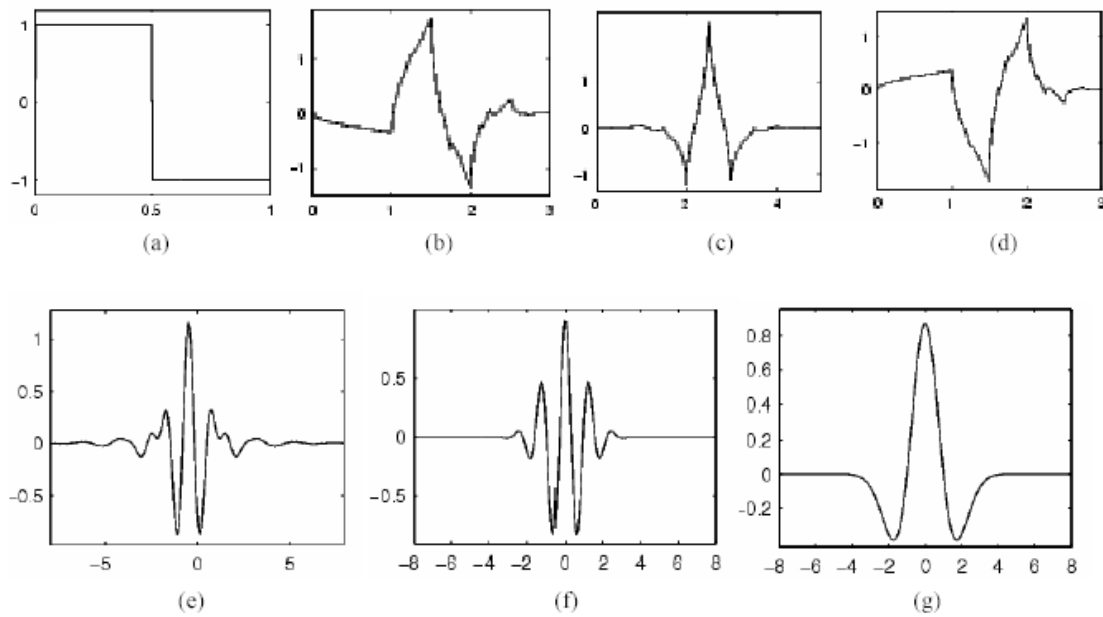


Figura 2.4: Alguns exemplos de famílias de ondaletas. (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat.

2.4.1 Histórico

A história científica das ondaletas é incomum e marcada por várias descobertas independentes e redescobertas. O progresso mais rápido têm ocorrido desde o começo dos anos 80, quando uma teoria matemática coerente de ondaletas finalmente emergiu (MACKENZIE, 2001).

Em 1807, Jean Baptiste Joseph Fourier anuncia que qualquer função periódica pode ser representada como uma soma infinita de senos e cossenos em várias frequências. Sérias dúvidas sobre a certeza de seus argumentos fizeram com que seus resultados fossem publicados somente 15 anos mais tarde.

Os matemáticos, baseados na teoria de Fourier, expandiram sua idéia para funções não-periódicas (ou ondas) que mudam com o tempo, ao invés de repetirem sempre a mesma forma. Alfred Haar, um matemático húngaro, em 1909, descobriu uma base de funções que são hoje reconhecidas como as primeira ondaletas. Estas funções são simplesmente um pequeno degrau positivo seguido de um negativo.

Em 1946, Dennis Gabor, um físico britânico-húngaro, introduziu a transformada de Gabor, análoga à transformada de Fourier, a qual separa uma onda em "pacotes tempo-frequência" ou "estados coerentes" que tem a grande possibilidade de localização simultânea no tempo e na frequência.

Nas décadas de 70 e 80, as comunidades de processamento de sinais e imagens introduziram suas próprias versões de análise das ondaletas, dando os nomes de "*subband coding*", "*quadrature mirror filters*" e "*pyramidal algorithm*".

Em 1981, o engenheiro de petróleo Jean Morlet encontrou uma maneira de decompor sinais sísmicos no que ele chamou de "ondaletas de forma constante". No seu método, ele podia separar uma onda em componentes de ondaletas e então reagrupá-las obtendo novamente o sinal original. Ele pediu ajuda ao físico quântico Alex Grossmann, do Centre de Physique Théorique in Marseilles, para ajudá-lo a provar que o método funcionava. Trabalharam juntos por um ano, até confirmarem que as ondas podem ser reconstruídas a partir de suas decomposições em ondaletas. Em 1984, um artigo conjunto de Morlet e Grossmann trouxe a palavra "*wavelet*" ao vocabulário matemático pela primeira vez.

Um ano mais tarde, Yves Meyer, da Universidade de Paris, descobriu as primeiras ondaletas suaves ortogonais. A propriedade matemática da ortogonalidade torna a transformada das ondaletas tão fácil de trabalhar e manipular quanto a transformada de Fourier. Ortogonalidade significa que a informação capturada por uma ondaleta é completamente independente da informação capturada por outra. Ele é conhecido como um dos fundadores da teoria das ondaletas. Realizou conexão entre as ondaletas de Morlet e ondaletas anteriores (encontrou 16 referências anteriores a Morlet para ondaletas).

Em 1986, Stéphane Mallat, então doutorando em visão computacional na Universidade da Pensilvânia e aluno de Meyer, mostrou que a base de Haar, as oitavas de Littlewood-Paley, os gorjeios de Gabor e os filtros de sub-banda de Galand e Esteban estavam todos relacionados com os algoritmos baseados em ondaletas. Mostrou que existe relação entre a teoria das ondaletas e a literatura existente de códigos de sub-banda e filtros espelhados em quadratura, que são a versão de ondaletas da comunidade de processamento de imagens. A idéia de análise multi-resolução, que é olhar os sinais em diferentes escalas de resolução, já era familiar aos especialistas em processamento de imagem. Mallat, colaborando com Meyer, mostrou que ondaletas estão implícitas no processo de análise multi-resolução.

No ano seguinte, Ingrid Daubechies descobriu toda uma nova classe de ondaletas, as quais não são só ortogonais, mas podem ser implementadas usando simples idéias de filtração digital ou usando filtros digitais pequenos. Suas ondaletas são quase tão simples de programar e usar quanto as ondaletas de Haar, porém são suaves e sem os saltos da ondaleta de Haar, ou seja, Daubechies "construiu" as primeiras ondaletas suaves ortogonais com suporte compacto. Os processadores de sinais tinham então uma ferramenta dos sonhos: um modo de dividir dados digitais em contribuições de várias escalas. Combinando as idéias de Daubechies e Mallat, há uma transformada ortogonal simples que pode ser rapidamente usada com computadores digitais modernos. Suas ondaletas tornam a teoria em uma ferramenta prática que pode ser facilmente programada e usada por qualquer cientista com um mínimo de conhecimento matemático. Suas ondaletas são as mais utilizadas na prática (MORETTIN, 1997).

Em 1990, David Donoho e Iaim Johnstone da Universidade de Stanford usaram as ondaletas para "eliminar ruído" de imagens, fazendo-as mais reconhecidas que as originais.

Na década de 90 surgiram as primeiras aplicações de grande porte das ondaletas, como seu uso pelo FBI para compressão de dados de impressões digitais, formas feitas por

subdivisão de superfícies utilizadas no filme Toy Store 2 e compressão de imagens digitais chamada JPEG-2000.

Ainda existem dúvidas quanto a esta nova teoria das ondaletas. Alguns a vêem como uma grande promessa, outros, com dúvidas. Uma boa parte da incredulidade é devido aos vários nomes que a metodologia ganhou ao longo dos anos, como *constant-Q analysis*, *wide-band cross-ambiguity functions*, *Frazier-Jawerth transforms*, *perfect reconstruction octave-band filter banks* ou *variation of Laplacian pyramid decomposition*. Uma das vantagens das ondaletas seria que ela busca unificar todas estas teorias em uma rede comum, permitindo novas idéias e desenvolvimentos (RIOUL e VETTERLI, 1991).

2.4.2 Transformada das Ondaletas

A transformada das ondaletas (*Wavelet Transform* - WT) tem sido largamente aplicada, em áreas como processamento de sinais e imagens, detecção de singularidades, análise de fractais, extração de tendências, análises numéricas, eliminação de ruído e compressão de dados, devido à sua aplicação matemática simples e porque é uma ferramenta que determina simultaneamente a localização no tempo e frequência.

A transformada das ondaletas é uma ferramenta que divide um sinal em componentes de frequências diferentes, e então estuda cada componente com uma resolução que combina com sua escala (DAUBECHIES, 1992). Este processo de divisão do sinal consiste em multiplicar o mesmo por ondaletas escalonadas e deslocadas, sobre todo o tempo, com relação a uma ondaleta mãe (ondaleta original), que é a função base. A **Figura 2.5** representa o processo de decomposição do sinal, também chamado de análise.

Escalonar uma ondaleta equivale a dilatá-la (ou comprimi-la), e deslocar uma ondaleta no tempo é o mesmo que transladá-la. Assim, uma ondaleta escalonada e deslocada pode ser representada de acordo com a Eq. (2.11).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.11)$$

onde a é o fator de escala, b é o fator de translação ou deslocamento e ψ representa a ondaleta mãe (MISITI *et al.*, 1997). O fator $1/\sqrt{a}$ é para normalização da energia do sinal (para que tenha a mesma energia em todas as escalas).

Quanto maior o fator de escala, mais dilatada é a ondaleta. Isto implica trabalhar com baixas frequências. Por outro lado, se o fator de escala decresce, há uma compressão da ondaleta no tempo, aumentando as oscilações e exibindo componentes de alta frequência. Portanto, o fator de escala está inversamente relacionado com a frequência. Pode-se observar melhor esta relação escala-tempo pela análise da **Figura 2.6**. Assim, a resolução no tempo se torna arbitrariamente boa a altas frequências, enquanto a resolução na frequência se torna boa em baixas frequências (MISITI *et al.*, 1997).

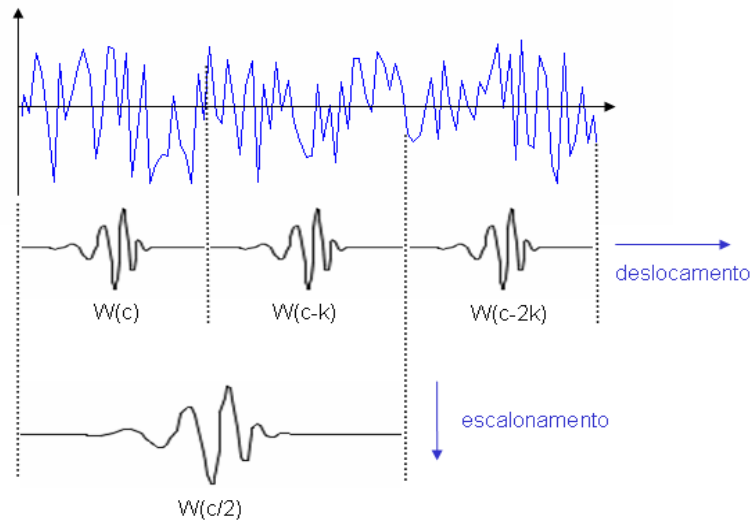


Figura 2.5: Processo de decomposição: envolve etapas de deslocamento e posterior escalonamento (dilatação) da ondaleta (ALTMANN, 1996).

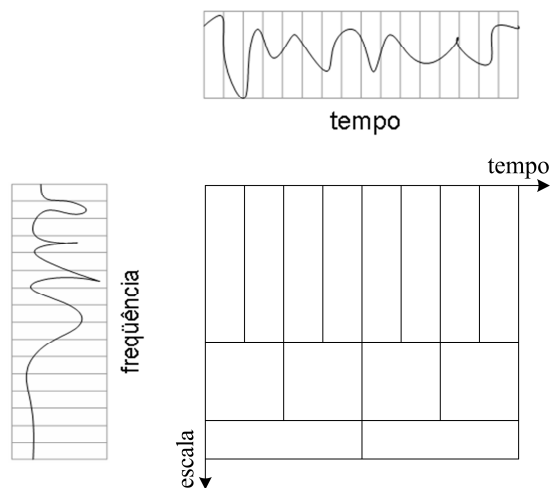


Figura 2.6: Representação das ondaletas nos domínios da frequência-escala e do tempo.

Devido a esta característica das ondaletas de análise de componentes de alta frequência em intervalos de tempo curtos e análise de componentes de baixa frequência em intervalos de tempo longos, elas são indicadas para análise de sinais transientes e que podem apresentar singularidades, significando que apresentam caráter de análise global e local. A performance das ondaletas é, contudo, destacada para análise de sinais não-estacionários. A escala de resolução mais baixa (*coarsest scale*) é a de maior escala, e ir de uma escala maior para uma menor equivale a dar um *zoom* no sinal (MORETTIN, 1997).

O processo de divisão ou decomposição do sinal pode ser, então, expressado conforme a Eq. (2.12).

$$C(a,b) = W_{a,b} f(t) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.12)$$

onde $a, b \in \mathfrak{R}$ e $a \neq 0$.

Esta multiplicação produz coeficientes $C(a,b)$ que são como um índice de semelhança entre a seção analisada do sinal e a ondaleta. Este processo de decomposição, conhecido como Transformada Contínua das Ondaletas (*Continuous Wavelet Transform - CWT*), nada mais é, portanto, do que a soma no tempo do sinal multiplicado por versões escalonadas e deslocadas da função ondaleta ψ . Entretanto, a CWT produz uma análise redundante, gerando muitos dados desnecessários.

Surge, então, a Transformada Discreta das Ondaletas (*Discrete Wavelet Transform - DWT*), que realiza a análise apenas com alguns escalonamentos e deslocamentos. Normalmente, os escalonamentos são de dilatação binária $a = 2^j$ e as translações são diádicas $b = ka$ (DAUBECHIES, 1992; RIOUL e VETTERLI, 1991), daí o nome de Transformada Diádica das Ondaletas (*Dyadic Wavelet Transform*). Uma das vantagens do algoritmo discreto é sua relativa baixa complexidade, a qual aumenta linearmente com o número de amostras de entrada, por um fator constante que depende do comprimento do filtro (RIOUL e VETTERLI, 1991). Neste caso discreto, estamos considerando um caso particular da Eq. (2.11), onde as ondaletas podem então ser representadas por

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k) \quad (2.13)$$

onde j é o nível associado com a escala e k é o parâmetro de deslocamento associado com a posição no tempo (DAUBECHIES, 1992).

Estas ondaletas formam uma base ortogonal $\{\psi_{j,k}(x), j, k \in \mathbb{Z}\}$, que apresenta algumas vantagens, como a reconstrução perfeita do sinal original a partir dos coeficientes da transformada e que cada coeficiente pode ser calculado como o produto interno do sinal com a função base (MORETTIN, 1997). O processo de reconstrução, também chamado de síntese, consiste na obtenção de um sinal contínuo no tempo a partir de uma seqüência de números obtidos pelo processo de amostragem. Salienta-se que amostragem sempre implica em perda de informação. Entretanto, a quantidade de perda depende do método de amostragem e dos parâmetros associados. É importante observar que a taxa de amostragem deve ser maior se o sinal apresenta componentes de alta freqüência (GOODWIN *et al.*, 2001).

Na transformada diádica, o plano tempo-escala pode ser representado como o da **Figura 2.7**, onde cada ponto corresponde a uma função base da ondaleta com escala 2^{-j} e deslocamento $2^{-j}k$ (RIOUL e VETTERLI, 1991).

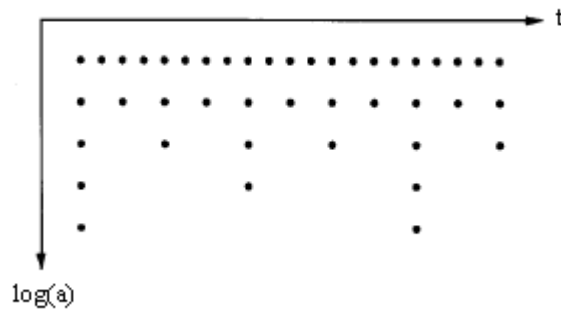


Figura 2.7: Representação de amostragens diádicas no plano tempo-escala.

Neste caso, uma aproximação de $f(x)$ em uma resolução 2^j significa que alguma informação sobre $f(x)$ foi perdida. Entretanto, à medida que a resolução aumenta para o infinito, o sinal aproximado deve convergir com o sinal original. Ao contrário, se a resolução tende para zero, o sinal aproximado contém cada vez menos informação e também tende a zero (MALLAT, 1989).

A DWT pode ser representada pela Eq. (2.14), discretizando a Eq. (2.12).

$$C(j, k) = \langle f, \psi_{j,k} \rangle = 2^{-j/2} \sum_k f(x) \psi(2^{-j} x - k) \quad (2.14)$$

onde $\psi_{j,k}$ é dado pela Eq. (2.13) ou então por

$$W_{2^j} f(x) = f * \psi_{2^j}(x) \quad (2.15)$$

onde $W_{2^j} f(x)$ representam coeficientes das ondaletas.

No processo de síntese discreta, ocorre o inverso da decomposição, ou seja, o sinal original é obtido a partir dos coeficientes das ondaletas.

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} c(j, k) \psi_{j,k}(x) \quad (2.16)$$

Às vezes é conveniente considerar uma função de escala ϕ , ou também chamada de ondaleta pai, que é uma solução da Eq. (2.17).

$$\phi(x) = \sum_k c_k \phi(2x - k) \quad (2.17)$$

Esta função gera uma família ortonormal definida por⁴

$$\phi_{j,k}(x) = 2^{-j/2} \phi(2^{-j}x - k) \quad (2.18)$$

Esta propriedade adicional de escalonamento está relacionada com as frequências baixas, conforme apresentado na **Figura 2.8**, permitindo que sejam implementadas por um banco de filtros (*filter bank*). Um banco de filtros é composto por filtros passa-banda com largura de banda relativa constante, também chamada de análise "constante-Q", na qual a frequência central e a largura de banda da ondaleta caem pela metade em cada escala. Ainda pode-se dizer que a resposta do filtro no domínio da frequência é regularmente espaçada em uma escala logarítmica (RIOUL e VETTERLI, 1991).

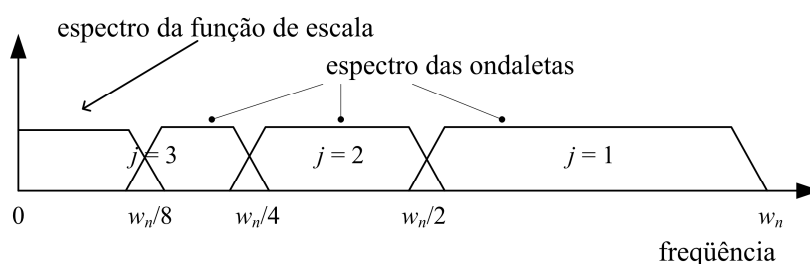


Figura 2.8: Exemplo de um banco de filtros para um sinal cuja variação de frequência é de 0 a w_n (VALENS, 2004).

Um filtro passa-baixa meia-banda, como o da **Figura 2.8**, remove todas as frequências acima da metade da maior frequência do sinal. Por exemplo, se um sinal tem um componente máximo de 1000 Hz, então o filtro remove as frequências acima de 500 Hz, e permite a passagem das frequências de 0 a 500 Hz (POLIKAR, 2001).

Contudo, na prática, a filtragem não é feita pela convolução indicada na Eq. (2.15), mas se calculam os coeficientes através de algoritmos que consistem em uma seqüência de filtros passa-alta $g[n]$ e passa-baixa $h[n]$, no processo denominado de código sub-bandas (*subband coding*). Existem vários algoritmos, como o ilustrado na **Figura 2.9**, que podem ser utilizados, de acordo com a finalidade desejada.

⁴As Eq. (2.13) e (2.18) são representadas, em algumas literaturas, de modo um pouco diferente (MORETTIN, 1997). As equações podem aparecer sem os sinais de "negativo" junto ao nível j . Isto implica numa inversão de conceitos da relação entre o fator de escala e a resolução, e conseqüente alteração de outras equações apresentadas neste trabalho.

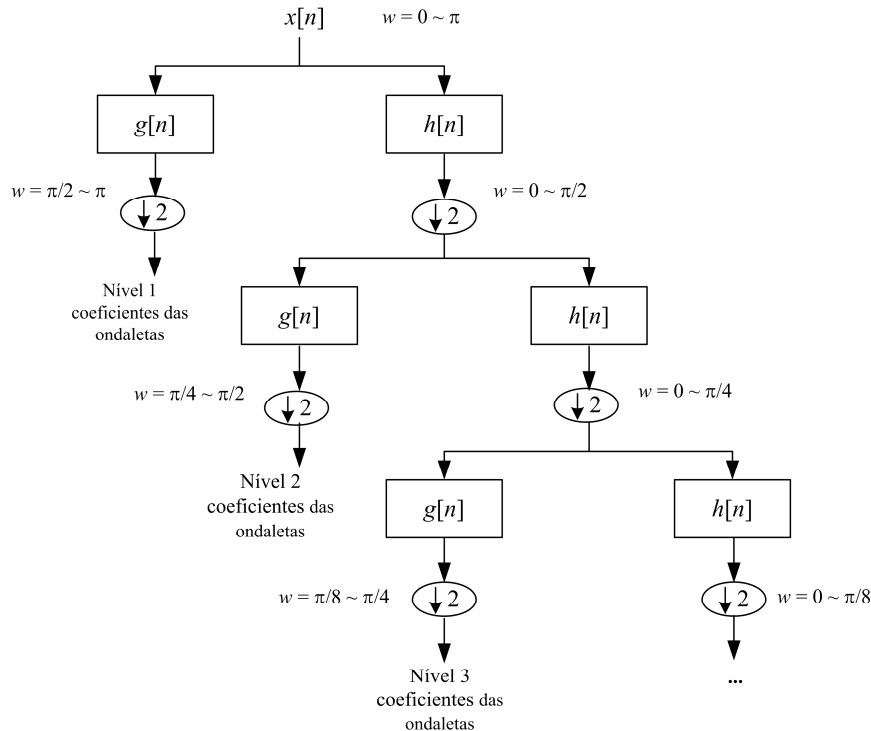


Figura 2.9: Banco de filtros do processo de decomposição (POLIKAR, 2001).

Na **Figura 2.9**, $x[n]$ é o sinal original, com frequências ω que variam entre 0 e π rad/s⁵. O símbolo $\downarrow 2$ indica o processo de decimação ou redução de amostragem (*downsampling*), o qual faz uso do que se chama de amostragem crítica, em um processo no qual não há redundância de amostras. A cada duas saídas do filtro, uma é desprezada. Neste exemplo, 3 níveis de decomposição foram representados. Portanto, se o sinal original tivesse 16 pontos amostrados, resultariam 4 coeficientes de cada filtro no terceiro nível. A resolução na frequência está representada na **Figura 2.8**. Este processo de redução de amostragem é realizado de modo a reduzir a quantidade de dados a serem analisados e para garantir a incerteza tempo-frequência implícita da análise (à medida que a análise se torna mais precisa para componentes da frequência do sinal, ela se torna menos precisa para o instante de tempo em que ocorrem) (BRADLEY e WILSON, 2004).

Após passar o sinal por um filtro passa-baixa meia-banda, metade das amostras podem ser eliminadas, de acordo com a regra de Nyquist, visto que o sinal tem agora uma frequência máxima de $\pi/2$ rad/s ao invés de π radianos. Descartando uma em cada duas amostragens fará o sinal ter metade do número de pontos e a escala dobrada. Note que o filtro passa-baixa remove a informação de alta frequência, mas deixa a escala igual. Apenas o processo de redução de amostragem muda a escala. A resolução, por outro lado, está relacionada com a quantidade de informação do sinal, e portanto, é afetada pelas operações de filtração. Filtros

⁵Em sinais discretos, a frequência é expressa em termos de radianos. Assim, a frequência de amostragem de um sinal é igual a 2π radianos em termos da frequência radial. Por isso, o componente de frequência máxima que existe em um sinal será π radianos, se o sinal é amostrado na taxa de Nyquist (maior que o dobro da maior frequência contida no sinal a ser amostrado, para que possa ser reproduzido integralmente sem erro de aliasing). Assim, a taxa de Nyquist corresponde a π rad/s no domínio discreto da frequência (POLIKAR, 2001).

passa-baixa meia-banda removem metade das frequências, o que pode ser interpretado como perda da metade da informação. Portanto, a resolução é reduzida à metade após a operação de filtração. Note também que a operação de redução de amostragem após filtração não afeta a resolução, visto que a remoção de metade dos componentes espectrais do sinal faz o número de amostras redundantes cair pela metade de qualquer forma. Metade das amostras podem ser descartadas sem perda de informação. Em resumo, o filtro passa-baixa reduz a resolução à metade, mas mantém a escala constante. O sinal é então reduzido à metade visto que metade do número de amostras é redundante. Isto dobra a escala.

Em cada nível de decomposição do sinal, como apresentado na **Figura 2.9**, resultam coeficientes das ondaletas (ou de detalhe), d_j ou $W_{2^j} f$, e coeficientes de aproximação, a_j .

$$a_j[n] = \langle f(x), \phi_{2^j}(x-n) \rangle \quad (2.19)$$

$$d_j[n] = Wf(n, 2^j) = \langle f(x), \psi_{2^j}(x-n) \rangle \quad (2.20)$$

Os coeficientes de detalhe resultam dos filtros passa-alta $g[n]$, e estão, portanto, relacionados às características de alta frequência do sinal, que dão nuances ao sinal. Os coeficientes de aproximação estão relacionados aos componentes de baixa frequência presentes no sinal, responsáveis pela identidade do sinal, que são os componentes mais suaves. Assim, se uma mudança brusca anormal ocorre no sinal, os coeficientes de detalhe serão afetados. Estes coeficientes também indicam o incremento de resolução necessário para ir de uma aproximação mais suave para uma aproximação de maior resolução. Esta idéia de aproximações consecutivas a resoluções melhores forma o conceito de "análise multi-resolução" (*multiresolution approximation* ou *multiresolution analysis* - MRA) (MALLAT, 1999).

Então, dois eventos bruscos muito próximos podem ser sempre separados na análise indo para uma frequência de análise mais alta de modo a aumentar a resolução no tempo, conforme a **Figura 2.6**. Este tipo de análise funciona melhor, logicamente, para sinais que possuem componentes de alta frequência de curta duração, o que é normalmente o caso dos sinais encontrados na prática (RIOUL e VETTERLI, 1991).

O filtro h é determinado pela função de escala e suas propriedades. Como o número de escalas aumenta, o número de bancos de filtros vai para o infinito, como pode ser visto na **Figura 2.8**. Em particular, a função de escala é compactamente suportada se e somente se h tem um número finito de coeficientes não-zero. Os filtros h e g são filtros de Resposta Finita ao Impulso (FIR), que caracterizam ondaletas de suporte compacto (Apêndice C).

Pode-se verificar que os níveis são representados por dilatações da base das funções, como a da Eq. (2.17), e, portanto suas resoluções podem ser relacionadas entre si. Assim, existe uma "equação de dilatamento" que relaciona dois níveis de resolução consecutivos (DAUBECHIES, 1992). Há necessariamente uma seqüência de números reais $h[n]$ tal que

$$\phi(x) = \sqrt{2} \sum_n h[n] \phi(2x - n) \quad (2.21)$$

Da mesma forma, a relação entre os filtros passa-baixa e passa-alta pode ser expressa por

$$\psi(x) = \sqrt{2} \sum_n g[n] \phi(2x - n) \quad (2.22)$$

sendo que os filtros h e g são relacionados, então, por

$$g(L-1-n) = (-1)^n h(n) \quad (2.23)$$

onde L é o comprimento do filtro (que deve ser par). O fator $(-1)^n$ converte o filtro passa-baixa em um passa-alta, de modo que g é o espelho de h . Assim, os filtros h e g são filtros espelhados em quadratura (*quadrature mirror filters* - QMF) (MALLAT, 1989).

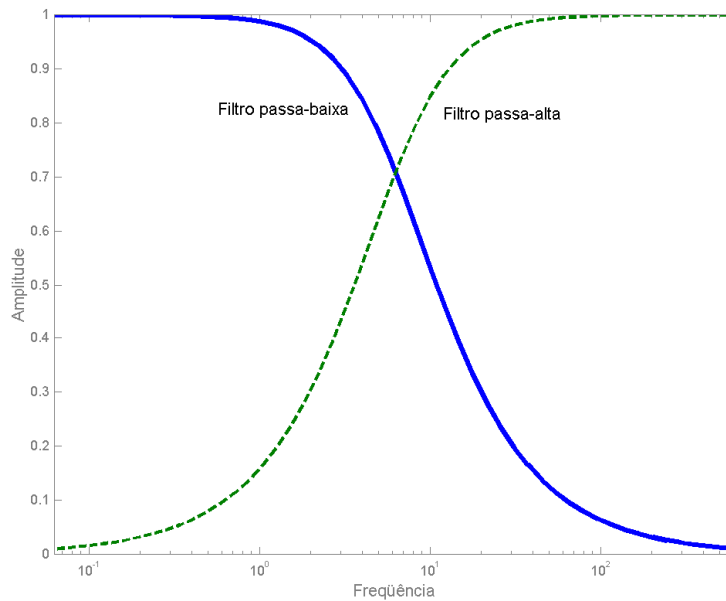


Figura 2.10: Filtros espelhados em quadratura (QMF).

A desvantagem da DWT é a variância no tempo, causada pela redução de amostragem (geração de *aliasing*). Este efeito é cancelado quando o processo inverso é executado usando todos os coeficientes das ondaletas (ou seja, quando o sinal original é reconstruído). Isto significa que a DWT seguida de reconstrução é invariante no tempo somente quando todos os coeficientes das ondaletas são utilizados para realizar a reconstrução. Como coeficientes não são incluídos na reconstrução, quer dizer, como uma sub-banda única é extraída, o efeito de *aliasing* não é mais cancelado e a saída não é mais invariante no tempo. Isto quer dizer que, para tornar a DWT ortogonal, amostragem crítica deve ser realizada. Entretanto, ao se realizar

qualquer sub-amostragem um certo grau de variância no deslocamento irá ocorrer (BRADLEY e WILSON, 2004).

Este conceito de invariância no tempo está representado pela **Figura 2.11**, onde $f(t)$ é uma função e a é um número real, de modo que $f_a(t) = f(t-a)$. Ou seja, f_a é um deslocamento no tempo, por a unidades, do sinal f (BOGGESS e NARCOWICH; INGLE e PROAKIS, 1997).

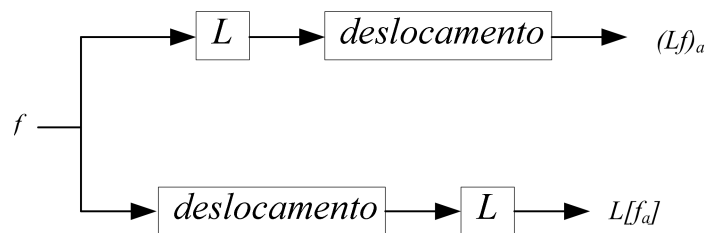


Figura 2.11: L é invariante no tempo se as duas saídas forem iguais.

A vantagem de se utilizar uma metodologia redundante é a obtenção da invariância no tempo (mais detalhes veja Apêndice A). Se uma velha gravação for tocada às 15h do dia de hoje e o sinal for passado em um filtro de redução de ruído, se ouvirá um sinal de saída limpo, sem interferências, aproximadamente no mesmo instante de tempo que a gravação for tocada. Se a mesma gravação for tocada às 10h da manhã do dia seguinte, e o mesmo filtro for utilizado, a mesma saída deveria ser escutada, aproximadamente no mesmo instante de tempo.

Uma das maneiras de tornar a análise das ondaletas invariante no tempo é realizando uma transformada redundante das ondaletas. Esta idéia foi desenvolvida independentemente por vários pesquisadores ao longo dos últimos anos e acabou recebendo diferentes nomes, incluindo “*algorithme à trous*” (HOLSCHNEIDER *et al.*, 1989; DUTTILEX, 1989), “*redundant wavelet transform*” (BURRUS *et al.*, 1997), “*stationary wavelet transform*” (MISITI *et al.*, 2006), “*undecimated DWT*” (LANG *et al.*, 1996), “*overcomplete DWT*” (ZACIU *et al.*, 1996), “*shift-invariant DWT*” (LANG *et al.*, 1995) e “*discrete wavelet frames*” (UNSER, 1995). Observa-se que a transformada contínua (CWT) é também invariante no tempo, pois não realiza nenhuma redução de amostragem.

Assim, existem várias maneiras de implementar a DWT redundante e de representar o conjunto supercompleto de coeficientes resultantes. A implementação original e mais simples foi na forma do “*algorithme à trous*” (HOLSCHNEIDER *et al.*, 1989; DUTTILEX, 1989) que, basicamente, remove o operador de decimação (redução de amostragem) da implementação usual da DWT. Neste caso, ao invés de decimar o sinal, zeros são inseridos (*holes* em inglês ou *trous* em francês) entre os coeficientes não-nulos dos filtros, segundo mostra a **Figura 2.12**. Ou seja, o filtro (ondaleta-mãe) é “dilatado” para levar em conta o reescalonamento, e a convolução é realizada sem redução e aumento de amostragem (MALLAT, 1999). O processo é dito redundante pois cada conjunto de coeficientes contém o mesmo número de amostras da entrada, de modo que, para uma decomposição de n níveis, a redundância é de $2n$. É nesta metodologia que se baseia este trabalho, segundo o mecanismo apresentado na **Figura 2.13**, onde a_j são os coeficientes de aproximação e d_j os coeficientes

de detalhe no nível j . Uma descrição mais detalhada também pode ser encontrada em MERTINS (1999).

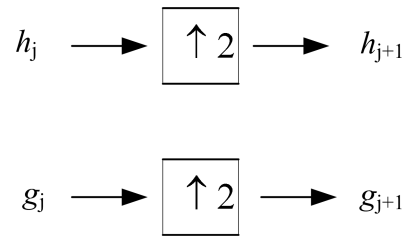


Figura 2.12: Esquema de computação dos filtros passa-baixa e passa-alta na transformada discreta estacionária das ondaletas.

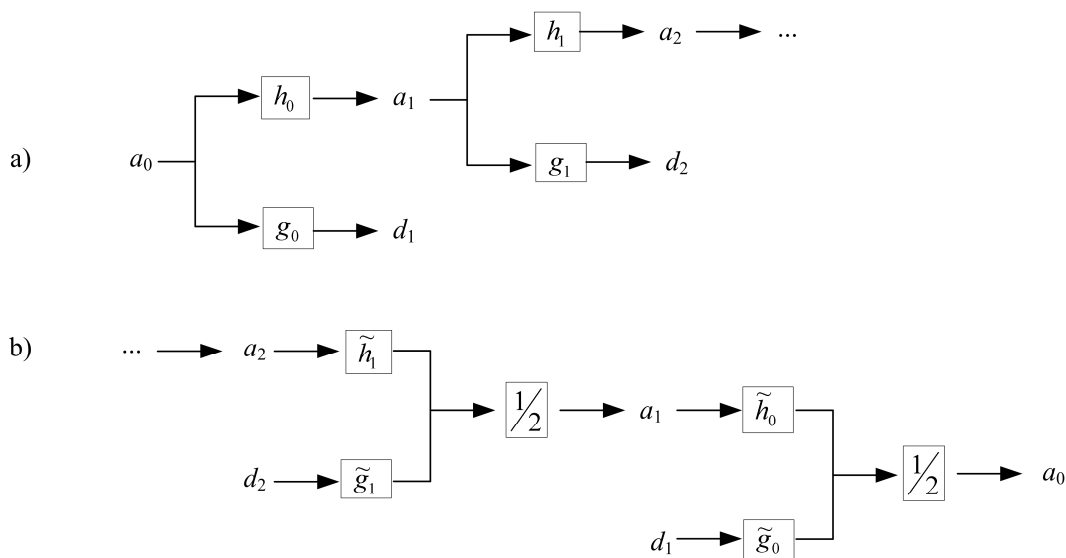


Figura 2.13: Banco de filtros segundo o *Algorithme à Trous* proposto por Mallat. a) etapa de decomposição. b) etapa de reconstrução.

Uma das desvantagens da DWT redundante é que requer maior esforço computacional, memória e correções de efeito de borda significativas, visto que a ondaletamãe dobra seu comprimento a cada nível de decomposição. Além disto, a DWT redundante é estritamente invariante no deslocamento sob convolução circular, ou seja, quando extensão de borda periódica é utilizada (BRADLEY e WILSON, 2004).

A decomposição do sinal a_j , conforme **Figura 2.13**, gera os sinais a_{j+1} e d_{j+1} . O sinal a_{j+1} pode ser decomposto gerando um sinal a_{j+2} com menor resolução e um sinal de detalhes d_{j+2} , e assim por diante. Da mesma forma, na etapa de reconstrução, um sinal a_j pode ser obtido a partir dos sinais a_{j+1} e d_{j+1} .

Os coeficientes de aproximação e detalhe são obtidos (MISITI *et al.*, 1997), para $j \geq 0$ e $h_j[n] = h[-n]$, por

$$a_{j+1}[n] = a_j * h_j[n] \quad (2.24)$$

$$d_{j+1}[n] = a_j * g_j[n] \quad (2.25)$$

A saída a_{j+1} de um filtro FIR para qualquer sinal de entrada pode ser calculada pela convolução do sinal de entrada a_j com a resposta ao impulso do filtro, expressa pelos coeficientes do filtro $h_j[n]$. Genericamente, para um dado filtro X com coeficientes $X[n]$, $X_j[n]$ é denominado o filtro obtido pela inserção de 2^j-1 zeros entre cada coeficiente X (MALLAT, 1999).

De modo equivalente, o processo de reconstrução discreta é calculado por:

$$a_j[n] = \frac{1}{2} (a_{j+1} * \tilde{h}_j[n] + d_{j+1} * \tilde{g}_j[n]) \quad (2.26)$$

No processo de reconstrução do sinal, a função $f(x)$ associada ao vetor de dados original pode ser calculada segundo a Eq. (2.27).

$$f(x) = \sum_{i \in I_j} a_{j,i} \phi_{j,i}(x) + \sum_{j=J}^1 \sum_{k \in k_j} d_{j,k} \psi_{j,k}(x) \quad (2.27)$$

No caso específico do trabalho que será apresentado nos capítulos seguintes, o algoritmo das ondaletas é baseado no trabalho de JIANG *et al.* (2003), onde ψ é uma ondaleta de primeira ordem, i.e., é a derivada de primeira ordem da função de escala ϕ , segundo a Eq. (2.28).

$$\psi(x) = \frac{d\phi(x)}{dx} \quad (2.28)$$

Substituindo a Eq. (2.28) na Eq.(2.15) resulta:

$$W_{2^j} f(x) = f\left(2^j \frac{d\phi_{2^j}}{dx}\right) = 2^j \frac{d}{dx} (f\phi_{2^j})(x) \quad (2.29)$$

onde

$$\phi_a(x) = \frac{1}{\sqrt{a}} \phi\left(\frac{x}{a}\right) \quad (2.30)$$

Como a ondaleta é a derivada de uma função suave, (2.29) prova que ψ é proporcional à derivada do sinal original suavizado na escala 2^j . O fato de a ondaleta ser a derivada da função de escala traz uma vantagem importante para o processo de decomposição, que é a obtenção direta da derivada do sinal aproximado nos coeficientes de detalhe. Algumas outras ondaletas poderiam ter sido escolhidas para analisar o sinal, como a ondaleta de Daubechies apresentada na **Figura 2.4**. No caso da seleção da ondaleta-mãe de Daubechies, ela é ótima

em termos de possuir comprimento de suporte mínimo para um dado número de momentos nulos, mas não é ótima com relação à simetria e suavidade (Apêndice C). Suporte compacto implica em ter complexidade computacional reduzida (pois ondaletas menores possuem menos multiplicações na integral da convolução) e comprimento de qualquer extensão de borda reduzido (BRADLEY e WILSON, 2004). Entretanto, a seleção da ondaleta do tipo *spline* se deve ao fato da ondaleta-pai ser suave e simétrica, e da ondaleta-mãe possuir suporte compacto, ser o mais regular possível e anti-simétrica (MALLAT e ZHONG, 1992). Isto confere boas propriedades à ondaleta, como localização de singularidades e descontinuidades (Apêndice C), o que é importante principalmente em sinais ruidosos. Na verdade, há poucos comentários na literatura sobre a escolha de determinado tipo de ondaleta para uma determinada aplicação, mas não deve haver grande diferença entre sinais reconstruídos a partir de ondaletas-mães diferentes.

Como apresentado anteriormente, aqui também a aplicação prática se faz por meio de uma seqüência de filtros passa-alta e passa-baixa, propostos por MALLAT e ZHONG (1992), que caracterizam funções do tipo *spline*. A função ondaleta é uma *spline* quadrática, derivada da função de escala, que é uma *spline* cúbica. As funções são apresentadas na Figura 2.14, e são representadas, no domínio da frequência, pelas Eq. (2.31) e (2.32).

$$\hat{\psi}(\omega) = i\omega \left(\frac{\sin(\omega/4)}{\omega/4} \right)^4 \quad (2.31)$$

$$\hat{\phi}(\omega) = \left(\frac{\sin(\omega/4)}{\omega/4} \right)^4 \quad (2.32)$$

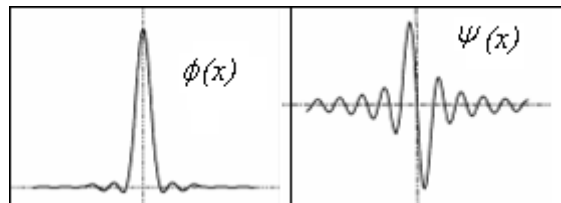


Figura 2.14: Função de escala *spline* cúbica e função ondaleta *spline* quadrática

Estas funções são representadas pelos coeficientes dos filtros passa-alta e passa-baixa apresentados na Tabela 2.1, os quais foram calculados por MALLAT e ZHONG (1992). Os filtros H e G são os filtros de decomposição passa-baixa e passa-alta, enquanto K é o filtro de reconstrução passa-alta. O filtro de reconstrução passa-baixa é o complexo conjugado de H .

Tabela 2.1: Respostas ao impulso dos filtros H , G e K .

n	H	G	K
-3			0,0078125
-2			0,054685
-1	0,125		0,171875
0	0,375	-2,0	-0,171875
1	0,375	2,0	-0,054685
2	0,125		-0,0078125

Estes filtros são utilizados diretamente no algoritmo desenvolvido por MALLAT e ZHONG (1992), para decomposição e reconstrução do sinal, o qual é descrito esquematicamente através do seguinte grupo de equações:

decomposição :

$$j = 0$$

while $j < J$

$$W_{2^{j+1}} f = \frac{1}{\lambda_j} \cdot S_{2^j} f * G_j$$

$$S_{2^{j+1}} f = S_{2^j} f * H_j$$

$$j = j + 1$$

end

reconstrução :

$$j = J$$

while $j > 0$

$$S_{2^{j-1}} f = \lambda_j \cdot W_{2^j} f * K_{j-1} + S_{2^j} f * \tilde{H}_{j-1}$$

$$j = j - 1$$

end

(2.33)

onde os fatores λ_j são coeficientes de normalização de energia em cada escala 2^j . A norma de qualquer versão transladada ou deslocada da ondaleta-mãe é a mesma da própria ondaleta-mãe, ou seja, assegura que todas as ondaletas em todas as escalas tenham a mesma área e contenham a mesma energia (ALSBERG *et al.*, 1997), pois, devido à discretização, o módulo máximo da ondaleta em um ponto de máximo não tem a mesma amplitude para todas as escalas como deveria no modelo contínuo. A constante λ_j compensa este efeito discreto (MALLAT e ZHONG (1992)). Os valores de λ_j que correspondem aos filtros da **Tabela 2.1** estão mostrados na **Tabela 2.2**.

Tabela 2.2: Coeficientes de normalização λ_j para a ondaleta *spline* quadrática.
Para $j > 5$, $\lambda_j = 1$.

j	λ_j
1	1,50
2	1,12
3	1,03
4	1,01
5	1,00

2.4.3 Comparação com a Transformada de Fourier

Cabe aqui alguns comentários para a comparação entre as transformadas das ondaletas e de Fourier, visto suas similaridades e metodologia equivalente.

A transformada de Fourier (*Fourier Transform*) discreta no tempo pode ser representada como

$$X(\omega) = \Delta \sum_{k=-\infty}^{\infty} x[k] e^{-j\omega\Delta k} \quad (2.34)$$

onde $x[k]$ é um sinal discreto obtido por amostragem de $x(t)$ nos instantes de tempo $t = k\Delta$ ($k = 0, 1, 2, \dots$), de modo que $x[k] = x(k\Delta)$, e Δ é o intervalo de tempo de amostragem associado (LEVINE, 1996).

Como mostra a Eq. (2.34), os coeficientes de Fourier são computados como produto interno do sinal por uma função base senoidal de duração infinita. Assim, a análise de Fourier trabalha bem para sinais compostos de uns poucos componentes estacionários (como ondas senoidais). Entretanto, qualquer mudança brusca no tempo em um sinal não-estacionário estará espalhada sobre todo o eixo de frequência. Portanto, uma análise adaptada para sinais não-estacionários requer mais do que uma transformada de Fourier (RIOUL e VETTERLI, 1991).

A transformada das ondaletas é semelhante à STFT (*Short-Time Fourier Transform*), uma técnica de transformada de Fourier que trabalha com janelas móveis no tempo. A STFT, porém, trabalha com regiões de tamanho constante, como mostra a **Figura 2.15**, enquanto a WT possui regiões de tamanho variável, conforme a **Figura 2.6**.

Assim, a principal diferença é que as distribuições das funções no tempo e na frequência são constantes na STFT, enquanto na WT a distribuição do tempo é função da escala e a distribuição da sua frequência é proporcional ao inverso da escala. Assim, se uma janela foi escolhida para a STFT, então a resolução tempo-frequência é definida e fixa em todo o plano tempo-frequência, já que a mesma janela é usada para todas as frequências.

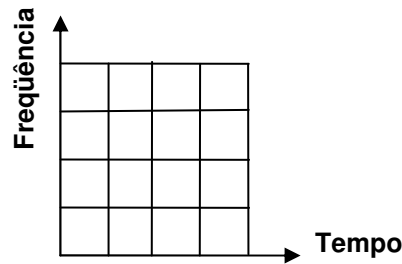


Figura 2.15: Representação da STFT.

Isto gera uma propriedade importante das ondaletas. Enquanto a transformada de Fourier é uma representação global do sinal, isto é, ela não pode analisar comportamentos de frequências locais, WT pode analisar tanto comportamentos globais quanto locais (regularidade local do sinal). A STFT tem uma resolução de frequência e tempo constantes, e esta resolução pode ser alterada através da mudança da janela. Portanto é uma representação completa, estável e redundante do sinal. Ao contrário, WT tem uma resolução de frequência e tempo que dependem da escala, de modo que em escalas finas há uma melhor resolução no tempo (MALLAT, 1999).

Bons artigos de referência que comparam Fourier com ondaletas de maneira didática e com ilustrações são ABSBERG et al. (1997) e RIOUL e VETTERLI (1991).

2.5 Análise por Componente Principal

Uma das técnicas multi-variáveis mais utilizadas para simplificar e melhorar os procedimentos de monitoramento de processos é a Análise por Componente Principal (*Principal Component Analysis - PCA*), visto que é uma técnica ótima em capturar a variabilidade dos dados. PCA é uma técnica linear de redução de dimensionalidade, na qual os dados são projetados em um espaço dimensional menor que o original, mas que representa o estado do processo. Isto significa que a estrutura de correlação entre as variáveis de processo é preservada neste espaço de dimensão reduzida (CHIANG *et al.*, 2001).

Na técnica de PCA, os dados são decompostos em um novo conjunto de vetores ortogonais, denominados vetores de carga, que são combinações lineares dos vetores de dados originais. Estes novos vetores são ordenados de acordo com a informação contida ou variância compreendida, conforme a **Figura 2.16**.

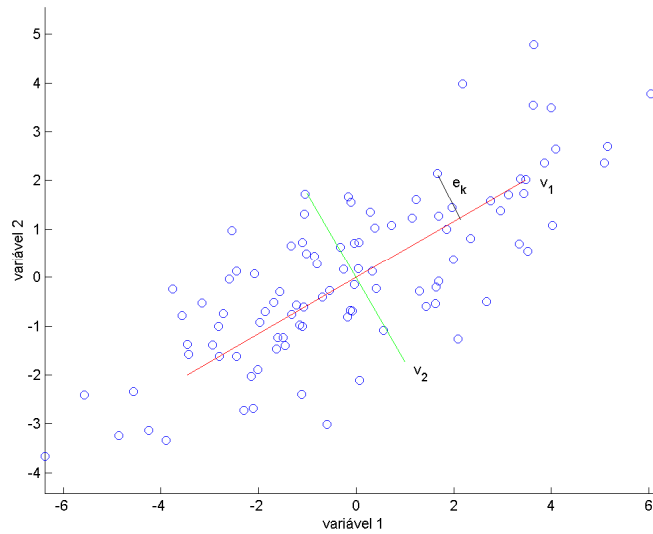


Figura 2.16: Ilustração da técnica de PCA.

No exemplo da **Figura 2.16**, duas medidas de processo foram realizadas. Observa-se que os dados apresentam uma elevada correlação pois, à medida que a variável 1 aumenta, a variável 2 também aumenta. O primeiro autovetor v_1 é a direção na qual os dados apresentam maior variabilidade, enquanto o segundo vetor v_2 é ortogonal ao primeiro, na direção da maior variabilidade residual. Observa-se que v_2 é necessário apenas para explicar o ruído nos dados. Assim, se apenas o vetor v_1 é usado para representar os dados, então uma reconstrução suave dos dados é possível (LEVINE, 1996).

Assim, dado um conjunto de dados X representado na matriz abaixo, onde as variáveis representam as colunas e as amostras representam as linhas,

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (2.35)$$

deseja-se determinar um conjunto de vetores de carga v (CHIANG et al., 2001) pela resolução dos pontos estacionários do problema de otimização

$$\max_{v \neq 0} \frac{v^T X^T X v}{v^T v} \quad (2.36)$$

A obtenção destes vetores de carga pode ser realizada por decomposição em valores singulares (singular value decomposition - SVD). Assim, se v são as colunas da matriz V , U e V são matrizes unitárias, e S é uma matriz diagonal com elementos não-negativos em ordem decrescente, tem-se que

$$X = U S V^T \quad (2.37)$$

onde

$$V = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1m} \\ v_{21} & v_{22} & \dots & v_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mm} \end{bmatrix} \quad (2.38)$$

e

$$S = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_m \end{bmatrix} \quad (2.39)$$

Para reduzir a dimensão da matriz de dados X , determina-se o número de fatores significativos desejados a ($a \leq m$) e constrói-se a matriz P a partir da matriz V .

$$P = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1a} \\ v_{21} & v_{22} & \dots & v_{2a} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{ma} \end{bmatrix} \quad (2.40)$$

A seguir são obtidas as projeções das variáveis originais no novo espaço formado por $T(t_k)$, onde T é a matriz que contém os componentes principais t_k , que são as novas variáveis ortogonais. O primeiro vetor da matriz $T(t_1)$ contém a informação da maior variabilidade, enquanto o segundo vetor contém a seguinte maior quantidade de variância compreendida após a remoção do primeiro efeito, e assim por diante. Estes componentes principais são determinados pela multiplicação das medidas reais pelos vetores de carga, e também são chamados de escores (LEVINE, 1996).

$$T = X \cdot P \quad (2.41)$$

Se apenas alguns vetores de carga foram utilizados, então X será aproximada por

$$\hat{X} = T \cdot P^T \quad (2.42)$$

e então

$$E = X - \hat{X} \quad (2.43)$$

Ou seja, se todos os autovetores foram utilizados ($a = m$), as matrizes X e \hat{X} são idênticas, de modo que o erro E é nulo, e a reconstrução perfeita de X é obtida (LEVINE, 1996).

2.5.1 Variáveis Escalonadas

Como uma etapa de pré-processamento, realiza-se a centralização e escalonamento dos dados de processo (LEVINE, 1996).

A centralização tem o objetivo de eliminar o viés (*bias*) dos dados, e é necessário porque PCA captura a informação (variância) que há em comum nos dados. Portanto, se há um viés (*bias*) essa será a principal informação que os dados têm em comum. A centralização é realizada subtraindo da cada variável X_i a sua média \bar{X}_i .

O escalonamento é realizado para que todas as variáveis adquiram a mesma importância, ou o mesmo grau de magnitude, no caso de variáveis que tenham medidas com valores grandes e outras que possuam valores pequenos. É realizado pela divisão dos dados pelo desvio padrão σ_i de cada variável.

Assim, o resultado da etapa de pré-processamento é uma variável \tilde{X}_i centralizada e escalonada, segundo a Eq. (2.44).

$$\tilde{X}_i = \frac{X_i - \bar{X}_i}{\sigma_i} \quad (2.44)$$

Capítulo 3

Derivada de Sinais Ruidosos

A análise e processamento de sinais com ruído é um problema comum para dados reais de processos, principalmente quando este sinal é aplicado em metodologias que requerem sua derivada. Ferramentas que podem fazer uso da derivada de um sinal são identificação de modelos, determinação de estados estacionários, compressão de dados, cálculo de variabilidades, entre outras.

A derivada de um sinal ruidoso fica mascarada pelo próprio ruído e não apresenta de forma clara os pontos de inflexão desejados para a análise do sinal. Faz-se necessário, portanto, passar o sinal por um filtro de modo a reduzir ou eliminar o ruído antes de processá-lo ou, então, obter a tendência que este sinal apresenta. Neste trabalho, portanto, considera-se que a tendência é uma aproximação do sinal original, o sinal que é obtido após a eliminação do ruído.

Neste trabalho são apresentadas comparativamente algumas maneiras de obter a tendência de um sinal e sua derivada. Três alternativas comumente empregadas para remoção de ruídos são comparadas: (i) uso de filtros discretos do tipo passa-baixa (filtros *Butterworth*); (ii) aproximação por *Splines* e (iii) tendências obtidas por Ondaletas (*Wavelets*). Alguns tipos de sinais foram criados no Simulink / MATLAB® e tiveram sua tendência estimada, sendo, então, derivados e comparados. Aqui serão apresentados quatro tipos de sinais gerados: degrau com ruído, seqüência de degraus, sinal senoidal e rampa com descontinuidades. Os métodos foram comparados segundo diversos critérios, tais como: qualidade da tendência, qualidade da derivada, simplicidade de configuração/ajuste e esforço computacional.

3.1 Geração dos Sinais

Alguns tipos de sinais, com características diferentes, foram gerados no Simulink / MATLAB®. Os sinais estão detalhados nas seções seguintes.

As técnicas de obtenção de tendência de um sinal foram implementadas no MATLAB. Para o caso de filtros Butterworth, utilizou-se a função *idfilt*; para *Splines*, as funções utilizadas foram *spaps* e *spap2* e, para Ondaletas, um algoritmo foi desenvolvido, baseado nos filtros criados por MALLAT e ZHONG (1992) e descrito em JIANG *et al.*(2003), conforme será apresentado na seção 3.2.

Utilizando o filtro de Butterworth, através da função *idfilt*, os dados filtrados são obtidos após filtração nos sentidos direto e reverso, e é necessário informar dois parâmetros: a ordem do filtro n e a frequência de corte ω_c . Aqui deve-se lembrar que, quanto maior a ordem do filtro, mais ele se aproxima de um filtro ideal ($n \rightarrow \infty$), porém mais amostras são necessárias para se efetuar o cálculo, ou seja, o tamanho do filtro aumenta.

Splines, implementada com a função *spaps*, utiliza B-*spline* para a obtenção de sinal suave aproximado, e necessita de um parâmetro, que é uma tolerância *tol*. Esta tolerância, conforme descrito no Capítulo 2, é o limite aceitável entre a curva ajustada aos dados e os dados amostrados em si.

A implementação de *splines* na função *spap2* se dá por mínimos quadrados, onde são informados dois parâmetros de entrada: a ordem da *spline* k e o número de pontos de controle (*knots*) $nknt$, que compõe uma seqüência de pontos de controle de modo que a função seja aproximada tentando minimizar a diferença entre a função aproximada e os dados reais.

Por fim, para a utilização do algoritmo baseado em Ondaletas, apresentado na seção seguinte, são necessários dois parâmetros de entrada, que são a constante de tempo τ e um período de dados históricos l .

Estes parâmetros necessários para a obtenção das tendências foram escolhidos de modo que se obtivesse a melhor tendência com relação à referência. Para cada método, os valores dos parâmetros foram variados aleatoriamente e, de acordo com a resposta, escolheram-se os valores mais adequados.

3.2 Algoritmo para as Ondaletas

A técnica de extração da tendência de um sinal foi baseada no trabalho de JIANG *et al.* (2000 e 2003), tendo como base a análise multi-resolução das ondaletas. O algoritmo decompõe o sinal em várias escalas, elimina o ruído branco e os picos anormais, e então reconstrói o sinal.

O processo inicia com a decomposição do sinal original (transformada das ondaletas no sinal) gerando a_j e d_j em cada escala j . O algoritmo é baseado nos filtros espelhados em quadratura h e g propostos por MALLAT e ZHONG (1992), onde h_j e g_j são filtros com 2^{j-1} zeros interpolados entre dois coeficientes sucessivos. Os filtros caracterizam uma ondaleta do tipo *spline* quadrática.

Na etapa seguinte, aplica-se um valor de corte em d_j para as escalas $1 < j < J$, conforme Eq. (3.1), obtendo-se d'_j . O valor de corte para a primeira escala d_1 é atribuído como a média do módulo dos coeficientes de detalhe dos dados históricos (δ_1), pois, na escala $j = 1$, a WT é completamente dominada pelo ruído.

$$d'_{j,k} = \begin{cases} 0 & \text{se } |d_{j,k}| \leq \delta_j \\ \text{sgn}(d_{j,k})(|d_{j,k}| - \delta_j) & \text{se } |d_{j,k}| > \delta_j \end{cases} \quad (3.1)$$

onde, para $2 \leq j \leq J$,

$$\delta_j = \delta_1 \cdot 2^{-(j-1)/2} \quad (3.2)$$

Em seguida, picos anormais, como pulsos (*spikes*), são detectados e removidos com a técnica de extensão simétrica para as escalas $2 < j < J$, resultando em novos d'_j e a'_j (Apêndice B). Os pulsos são identificados (Eq. (3.3)) quando ocorre um par de coeficientes de detalhe com sinal contrário e de magnitude superior a um valor de corte T_1 e cuja duração é menor que um intervalo de tempo t_p , estimado como 1/5 da constante de tempo. Isto corresponde a uma mudança brusca nos dados de processo.

$$\begin{aligned} WT_1(p_1), WT_1(p_2) &\geq T_1 \\ p_2 - p_1 &\leq t_p \\ \text{sign}[WT_1(p_1)] \cdot \text{sign}[WT_1(p_2)] &= -1 \end{aligned} \quad (3.3)$$

O valor de corte T_1 para identificação de um pulso p é calculado pelo desvio padrão $\sigma_{|WT_1|}$ do módulo dos coeficientes de detalhe de dados históricos numa escala J definida.

$$T_1 = 3 \cdot \lambda_1 \cdot \sigma_{|WT_1|} \quad (3.4)$$

O parâmetro λ_1 é ajustado em torno de 1,0, e determinado de acordo com a variabilidade das medidas e sua sensibilidade à operação. Se uma indústria trabalha com a regra de 3σ , se $\lambda_1 = 1,0$, ou $T_1 = 3 \cdot \sigma_{|WT_1|}$, isto indica que o percentual de pontos fora de especificação (anormalidades) é de 0,3% (OAKLAND, 2003).

A duração do pulso $[t_a, t_b]$ é determinada de modo que t_a seja o ponto esquerdo mais próximo de p_1 e t_b seja o ponto a direita mais próximo de p_2 e que satisfazem

$$|W_f(t_a)|, |W_f(t_b)| \geq T_2 \quad e \quad |W_f(t_a - 1)|, |W_f(t_b + 1)| < T_2 \quad (3.5)$$

onde T_2 é um valor de corte determinado como a média do módulo dos coeficientes de detalhe de dados históricos.

Na etapa seguinte, o sinal é reconstruído usando os coeficientes a_j' e d_j' , a partir da escala $j = J$ até 1.

A tendência do sinal f_s é extraída na escala característica $j = s$, onde s é determinada pela constante de tempo τ e pelo tempo de amostragem t_s , conforme Eq. (3.6).

$$s = \text{int} \left(\log_2 \left(\frac{\tau}{t_s} + 0,5 \right) + 0,5 \right) \quad (3.6)$$

3.2.1 Determinação dos Parâmetros

A constante de tempo pode ser estimada através da relação entre variáveis manipuladas e controladas de configuração do algoritmo utilizando, por exemplo, um modelo de identificação ARX de ordem 10 obtido com o MATLAB[®] *System Identification Toolbox*. A seguir a resposta degrau deste modelo de elevada ordem é aproximada pelo método SK (SUNDARESAN e HRISHNASWAMY, 1977), o qual permite estimar a constante de tempo dominante do sistema.

Como parâmetro de entrada no algoritmo, a constante de tempo deve ser informada em função do número de amostragens. Por exemplo, se o intervalo entre duas amostragens é de 10 segundos, e a constante de tempo calculada foi de 2 minutos, o valor a ser informado ao algoritmo é 12 (6 amostragens/min * 2 minutos).

Por período histórico considera-se uma seqüência de dados de processo com boa variabilidade, que represente bem as características da variável. Se o conjunto selecionado possui pouca variabilidade e é muito bem comportado, o algoritmo calculará uma estreita faixa para que os dados sejam considerados estacionários. Ao contrário, se a variabilidade dos dados é a maior possível dentro de um limite aceitável, então a faixa calculada pelo algoritmo será maior, e mais dados poderão ser considerados estacionários. Uma verificação detalhada deste comportamento está apresentada na seção 4.3.2.

3.3 Cálculo das Derivadas

As derivadas dos sinais gerados e das tendências obtidas foram calculadas numericamente baseadas em diferenças finitas para todos os sinais e para todas as metodologias. Este cálculo foi realizado através da função *diff* do MATLAB[®].

A diferenciação por diferenças finitas, que é uma maneira de se obter numericamente as derivadas, pode reduzir bastante a precisão dos resultados, visto que a derivada possui alta sensibilidade a valores imprecisos da função que está sendo diferenciada. O mesmo não ocorre para integração numérica, que é muito mais insensível às funções porque possui um efeito atenuante que diminui o efeito das imprecisões nos valores da função que está sendo integrada (BOURG, 2006).

Quando se possui uma expressão analítica para a função considerada, prefere-se utilizá-la ao invés de realizar-se a diferenciação numérica. Por outro lado, se dados foram experimentalmente obtidos (sem expressão analítica para a função), então só resta a utilização de diferenciação numérica. Para estes casos, apresentam-se a seguir algumas considerações.

Por definição, a derivada de uma função é representada como:

$$\frac{df(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} \quad (3.7)$$

Quando os dados tabelados representam uma função, não se pode considerar que Δt tende a zero. Neste caso, a função é avaliada em algum ponto t e depois em outro ponto $t + \Delta t$, que é uma distância finita Δt a partir de t . Neste caso, a derivada da função pode ser aproximada por diferenças finitas, de acordo com a Eq. (3.8).

$$\frac{df(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} \approx \frac{f(t + \Delta t) - f(t)}{\Delta t} \quad (3.8)$$

À medida que Δt é menor, esta aproximação deve ser mais acurada. O problema de dados de processo é que normalmente o intervalo de amostragem Δt não é tão pequeno quanto se gostaria que fosse para efeitos de obtenção de derivadas. Mas ainda assim pode-se pelo menos aproximar a derivada para uma função.

Existem algumas maneiras diferentes para trabalhar com um conjunto de dados de processo de modo a obter-se a derivada. Podem-se utilizar diferenças ascendentes, diferenças decrescentes ou diferenças centrais. Pode-se facilmente mostrar que o método por diferenças centrais leva a resultados mais adequados do que os métodos de diferenças ascendentes ou decrescentes (BOURG, 2006). Por esta razão, normalmente esta metodologia é preferida com relação às demais. A aproximação por diferenças centrais tem o efeito de suavizar a derivada no ponto visto que considera os valores anterior e próximo ao ponto considerado. Ou seja, a metodologia considera os valores “vizinhos” dos pontos. As curvas das derivadas não são tão suaves quanto as curvas obtidas pelas derivadas teóricas, mas são melhores que as curvas obtidas por diferenciação ascendente ou decrescente.

Outras técnicas se basearam nestes efeitos das derivadas centrais, de modo a buscarem suavizar os dados tabelados antes de calcular a derivada numérica com o objetivo de aumentar a precisão. Por exemplo, pode-se interpolar uma função ao conjunto de dados e então usar a expressão analítica resultante da curva de ajuste para derivar uma expressão para obter a derivada. Na realidade, estas técnicas implicam em resultados mais suaves para a derivada e servem de base para diversos métodos numéricos que requerem cálculos de derivadas.

Outro modo é realizar a derivada numérica como apresentado e então ajustar uma curva aos dados para obter uma curva suave para a derivada. Pode-se mostrar que os resultados assim obtidos são melhores do que o próprio método de diferenças centrais (BOURG, 2006).

3.4 Sinal 1

O primeiro sinal gerado para análise e comparação do melhor método de obtenção de tendências foi um sinal com um degrau positivo num dado instante de tempo e outro degrau negativo, de mesma magnitude, 100 intervalos de tempo depois, e tomado por ruído. O sinal também foi gerado para apresentar dois picos anormais nos instantes de tempo de 240 e 720. O sinal foi gerado de acordo com o sistema apresentado na **Figura 3.1**.

No caso deste primeiro sinal gerado, a ordem do filtro Butterworth n e a frequência de corte ω_n foram considerados 2 e 0,15 respectivamente. Para as *Splines* a tolerância tol estabelecida foi de 500, a ordem k igual a 4 e o número de pontos de controle $nknt$ igual a 100, e, para as Ondaletas, a constante de tempo τ estimada foi de 5 minutos (considerando tempo de amostragem de 1 minuto) e o período histórico l considerado foi do tempo 800 a 1000 minutos. Para as Ondaletas, isto corresponde a 2 níveis de decomposição. Para comparação da *spline* quadrática com outro tipo de ondaleta, escolheu-se também comparar com uma ondaleta clássica Daubechies de ordem 2, também com 2 níveis de decomposição.

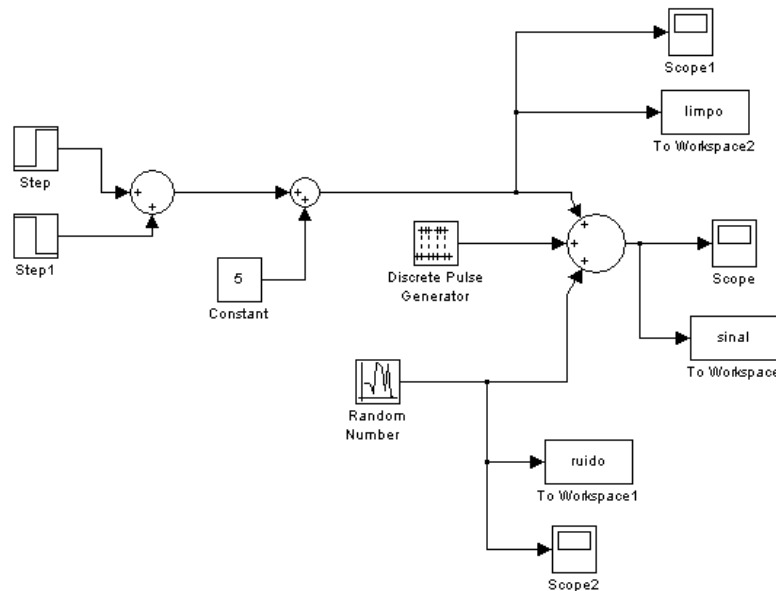


Figura 3.1: Geração do sinal para análise.

O sinal gerado está representado na **Figura 3.2**, assim como as tendências resultantes por cada técnica. O sinal foi gerado com uma composição de um sinal *limpo* e componentes de ruído. Assim, este sinal livre de ruído serve como referência para a obtenção das tendências e derivadas.

Nas **Figura 3.2** e **Figura 3.3**, a tendência e derivada do sinal original é chamada de *refer*, e as tendências e derivadas obtidas por *splines*, utilizando as funções *spaps* e *spap2*, são *ySPL* e *ySPL2*, respectivamente. A tendência e derivada obtida por Ondaletas é chamada de *yOND* e, por filtro de Butterworth, *yBUT*.

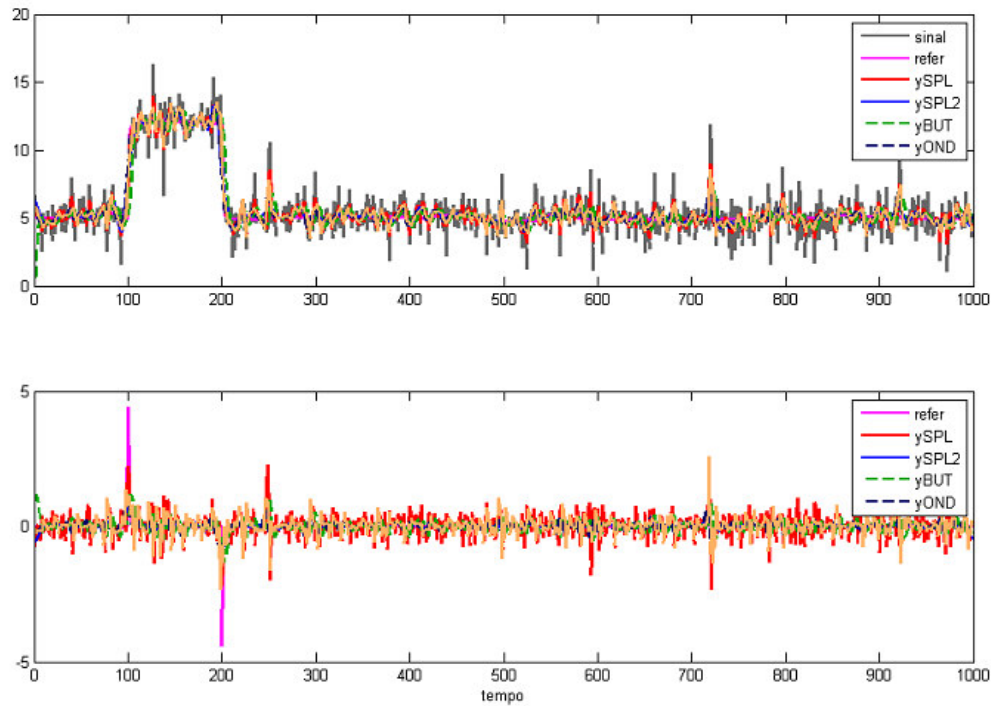


Figura 3.2: a) Sinal gerado e suas tendências. b) Derivadas.

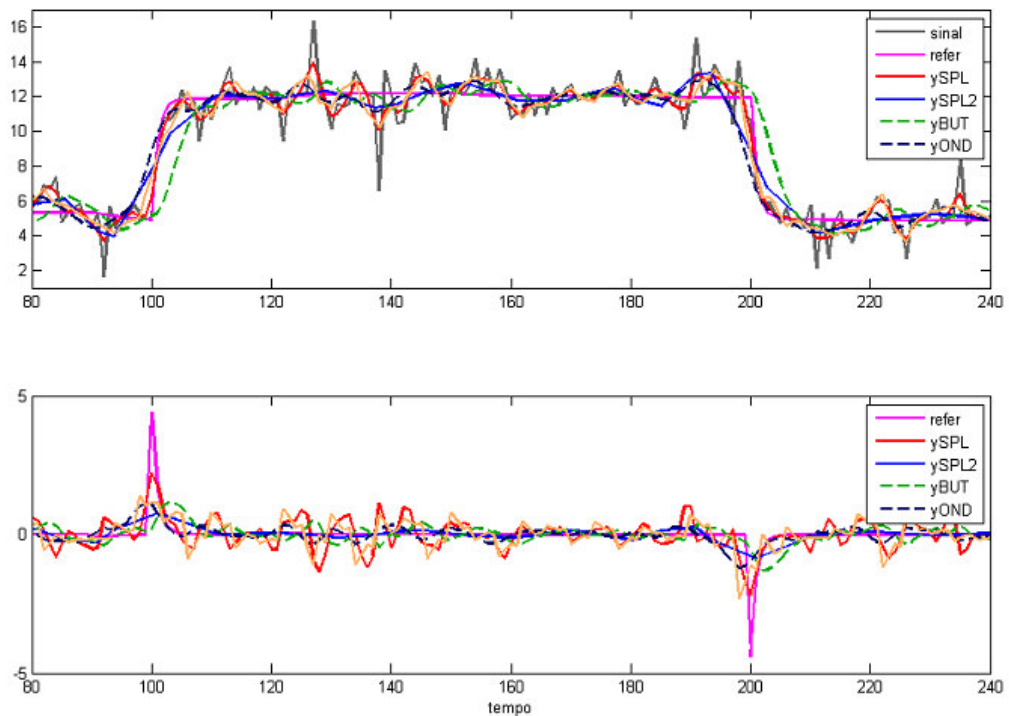


Figura 3.3: Ampliação da **Figura 3.2** entre o período de 80 e 240 para facilitar a visualização.

Os gráficos apresentados indicam que o filtro Butterworth introduz um certo atraso na tendência e, conseqüentemente, na sua derivada. A tendência resultante da *spline* de função *spap2* mostra-se não-suave, apresentando certas discontinuidades em seu sinal. As tendências

resultantes da função *spaps* (*splines*) e da aplicação das ondaletas clássica e *spline* quadrática apresentaram resultados similares, seguindo um comportamento aproximado. Entretanto, as ondaletas estacionárias (*splines*) parecem que adiantam um pouco o sinal, como também pode ser percebido nas suas derivadas.

3.5 Sinal 2

Um outro tipo de sinal que pode ser encontrado é uma extensão do sinal anterior, e consiste em uma série de degraus consecutivos, com processo de dinâmica rápida. Este sinal, pode ser visto na **Figura 3.4**, bem como suas tendências e derivadas, obtidas considerando $n = 3$, $\omega_n = 0,5$, $tol = 300$, $k = 3$, $nknt = 80$, $\tau = 2$ (1 nível de decomposição) e $l = 60$ a 80 .

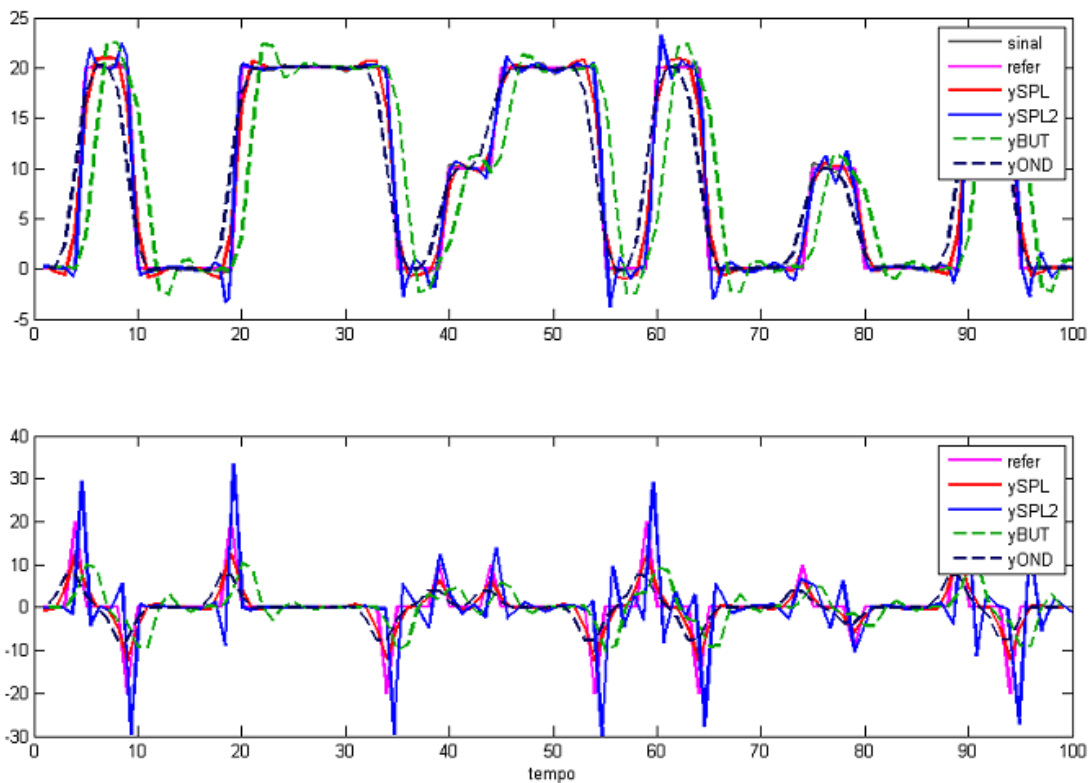


Figura 3.4: a) Sinal gerado e suas tendências. b) Derivadas.

Neste sinal, novamente percebe-se a introdução de um atraso pelos filtros Butterworth e o comportamento descontínuo gerado pela função *spap2* das *splines*. A função *spaps* gerou uma tendência muito boa, e uma derivada coerente com o esperado. A tendência obtida por ondaletas pode ser considerada boa, entretanto levemente adiantada. Sua derivada mostra-se bastante suavizada.

3.6 Sinal 3

Sinais do tipo senoidais são comumente encontrados em aplicações práticas e, por isto, um sinal com estas características foi estudado. O sinal está representado na **Figura 3.5**, com as tendências extraídas e suas respectivas derivadas. Selecionou-se a parte inicial deste sinal para facilitar a comparação, e o resultado está na **Figura 3.6**. Os melhores parâmetros obtidos foram para $n = 2$, $\omega_n = 0,1$, $tol = 1000$, $k = 4$, $nknt = 100$, $\tau = 5$ (2 níveis de decomposição) e $l = 150$ a 250.

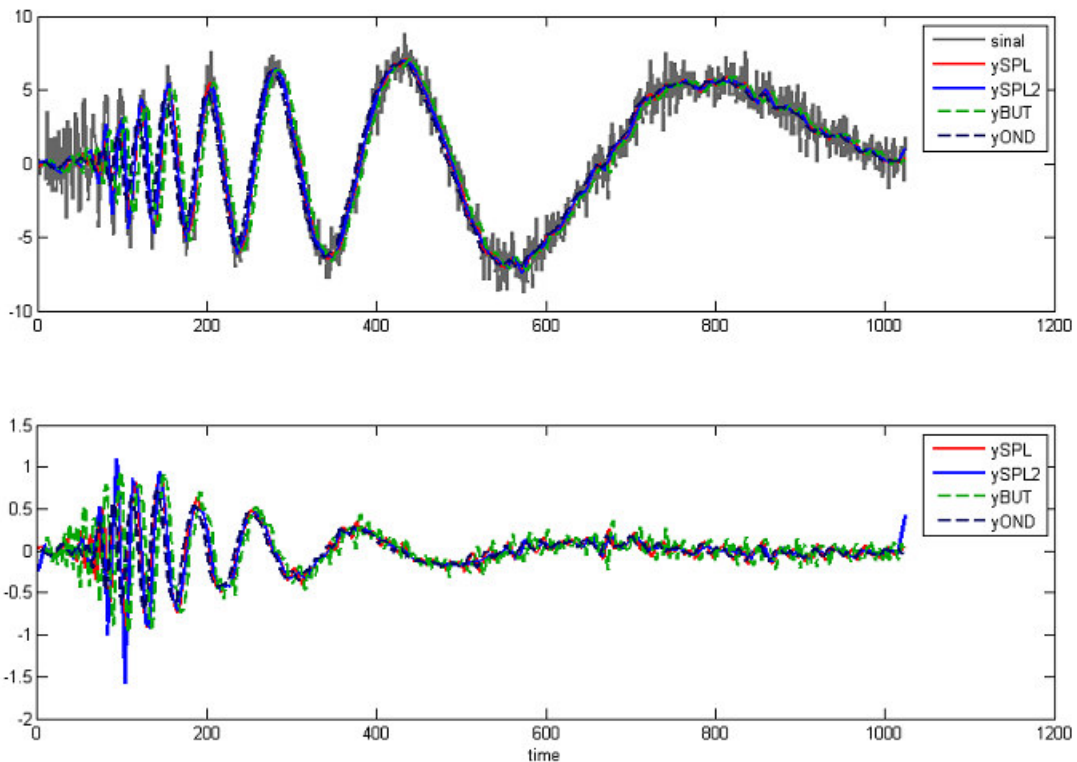


Figura 3.5: a) Sinal gerado e suas tendências. b) Derivadas.

Analisando a **Figura 3.6** pode-se observar novamente alguns efeitos já constatados nos exemplos anteriores, com o atraso gerado pelo filtro Butterworth e o leve adiantamento gerado pelas Ondaletas, bem como um sinal não-suave obtido por *spap2*. Todos os métodos suprimiram a parte inicial do sinal, de mais alta frequência, e acompanharam bem a parte de mais baixa frequência.

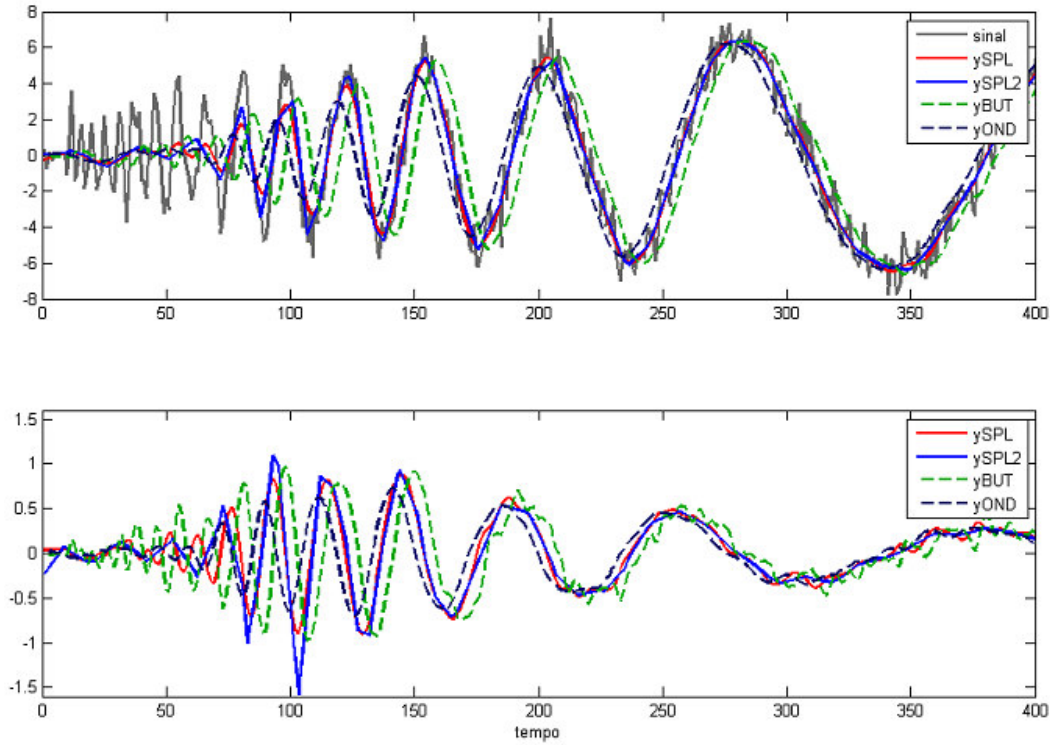


Figura 3.6: Ampliação da **Figura 3.5** para facilitar a visualização.

3.7 Sinal 4

Um sinal talvez não tão comum, mas importante para ser analisado também, consiste de uma rampa com sucessivas descontinuidades, próximas umas das outras, conforme mostra a **Figura 3.7**.

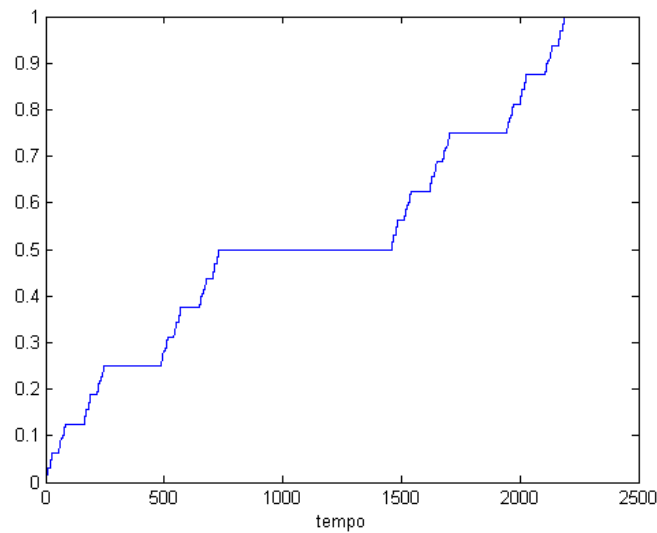


Figura 3.7: Sinal com descontinuidades próximas.

O sinal foi analisado tendo como parâmetros $n = 2$, $\omega_n = 0,3$, $tol = 3000$, $k = 3$, $nknt = 300$, $\tau = 2$ (1 nível de decomposição) e $l = 1$ a 700, e uma ampliação do resultado pode ser vista na **Figura 3.8**.

Os resultados não são satisfatórios, pois os métodos não apresentam bom desempenho. As mesmas conclusões dos exemplos anteriores podem ser observadas neste exemplo, porém, neste caso, a função *spaps* para *splines* apresentou o pior resultado, pois não conseguiu resolver as diferentes descontinuidades. Por mais que o parâmetro *tol* fosse alterado, o seu comportamento permaneceu o mesmo.

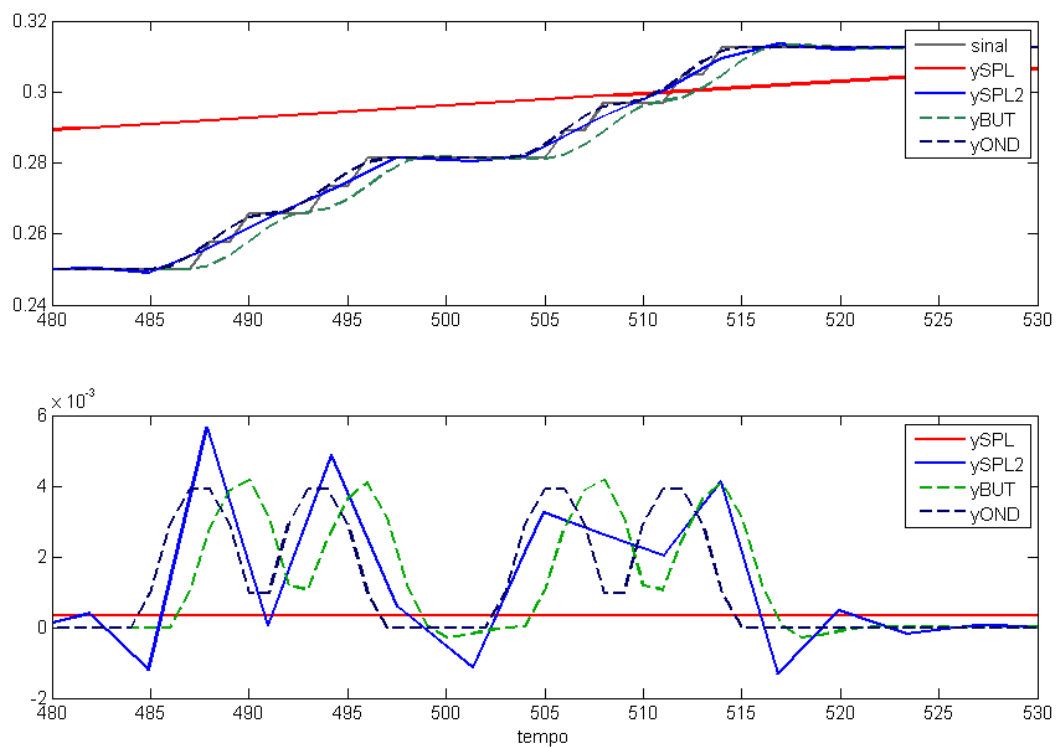


Figura 3.8: Ampliação do sinal com descontinuidades próximas.

3.8 Comparação dos Métodos

A aplicação das técnicas de extração de tendências para os sinais propostos e a obtenção de suas derivadas podem ser qualitativamente comparadas de acordo com a qualidade da tendência e da derivada, a dificuldade de implementação do algoritmo e efeitos de defasagem e de borda, conforme mostra a **Tabela 3.1**. Para efeitos desta comparação, estabeleceu-se que o índice 1 representa o melhor dos métodos para uma determinada característica, 2 é o método de qualidade intermediária e 3 é o pior método.

Tabela 3.1: Comparação do desempenho dos métodos: SPL (*Splines*), OND (Ondaletas) e BUT (filtro Butterworth).
1- Melhor, 2- Intermediário e 3- Pior

	SPL	OND	BUT
Esforço computacional	2	3	1
Facilidade de ajuste	1	1	2
Qualidade da filtração	1	2	3
Qualidade da derivada	1	2	3
Defasagem	1	2	2
Efeitos de borda	1	1	2

No que se refere ao esforço computacional, considera-se a dificuldade de implementação dos algoritmos e o tempo de cálculo. O filtro de Butterworth é o mais simples, enquanto as Ondaletas exigem um certo conhecimento e uma implementação mais apurada, visto que os cálculos são um pouco mais complexos. *Splines* ficam num plano intermediário, exigindo também um pouco mais de esforço computacional.

Com exceção da função *spaps*, que possui um parâmetro de ajuste, as demais técnicas exigem a estimativa de dois parâmetros de entrada. As Ondaletas têm um parâmetro mais fácil de informação, que é o período de dados históricos, mas por outro lado o parâmetro da constante de tempo de resposta nem sempre é fácil de ser estimado ou obtido na prática. Entretanto, as *Splines*, embora possua um parâmetro de entrada, também perde pois este parâmetro de tolerância não é intuitivo e não possui uma referência absoluta.

A análise das tendências e das derivadas obtidas indica que as *Splines* são as que apresentam a melhor qualidade, considerando a função *spaps*. Embora não tenham se adequadado ao último exemplo (sinal 4), o qual é um tipo de sinal mais raro de ser encontrado em sinais reais, elas apresentaram resultados superiores nos demais exemplos estudados.

No quesito defasagem, as Ondaletas e o filtro Butterworth apresentaram adiantamento e atraso, respectivamente, do sinal de tendência e das derivadas, o que é, obviamente, indesejável. Por isto, *Splines* tem a melhor pontuação nesta característica.

Por fim, quanto aos efeitos de borda, que são efeitos gerados no começo e no final de um sinal devido ao truncamento da janela de dados, todos os métodos parecem apresentar resultados semelhantes. Em um dos exemplos, o filtro Butterworth apresentou efeitos de borda e, por isso, foi penalizado na sua pontuação.

3.9 Conclusões

A decisão sobre qual é o melhor método para ser aplicado para a obtenção da tendência de um sinal passa por vários fatores, e depende do tipo de sinal que se deseja analisar.

Como foi mostrado, *Splines (spaps)* parece ser a ferramenta mais indicada para a maioria dos sinais. Entretanto, não é indicada se o sinal apresentar descontinuidades próximas. A outra função baseada em *Splines (spap2)* apresentou resultados ruins para os sinais estudados e não se apresenta como uma boa alternativa para a obtenção de tendências.

Da mesma forma, os filtros Butterworth parecem não serem os mais apropriados, devido ao atraso que geram ao sinal, a não ser que se deseje um menor esforço computacional.

Se o processo em análise é conhecido, as Ondaletas ainda podem ser uma boa alternativa, devido aos seus parâmetros de entrada serem mais facilmente estimados. Porém, trazem um pequeno adiantamento no sinal de saída. Esta antecipação pode ser facilmente corrigida, porém não foi realizada neste trabalho para facilitar a visualização dos dados.

Capítulo 4

Detecção de Estados Estacionários

Uma importante aplicação da obtenção de tendências e derivadas é a determinação dos estados estacionários de uma variável ou unidade. A detecção de estados estacionários assume um papel importante em otimização em tempo real, identificação de modelos estacionários, análise de processos, reconciliação de dados, entre outras aplicações. Estas aplicações requerem dados sob condições estacionárias ou próximas delas e, para isto, uma técnica eficiente se faz necessária.

Com este objetivo, várias metodologias de detecção de estados estacionários têm sido propostas. Muitos métodos são baseados em testes estatísticos. NARASIMHAN *et al.* (1986) apresentou o *Composite Statistical Test* - CST (1986) e o *Mathematical Test of Evidence* - MTE (1987). No método CST, períodos sucessivos de tempo são definidos e avaliados de acordo com as matrizes de covariância e média amostral. No método MTE, diferenças nas médias são comparadas com a variabilidade entre os períodos. Mais recentemente, CAO e RHINEHART (1995) propuseram um método baseado em média móvel ou filtro de primeira ordem convencional usado para reposicionar a média da amostra. Estas técnicas, entretanto, não são pontuais, ou seja, elas determinam a situação do processo sobre períodos de tempo. Este é um fato importante para aplicações *on-line*. Além disto, estas técnicas consideram apenas a presença de erros randômicos, e sabe-se que erros não randômicos também estão presentes, na forma de pulsos, por exemplo.

Neste trabalho, propõe-se uma metodologia baseada em ondaletas, a qual é uma adaptação da técnica proposta por JIANG *et al.* (2003), que determina um índice de estado estacionário pontual. O artigo propõe uma técnica de transformada de ondaletas e foca-se, principalmente, no caso mono-variável, citando uma proposta para sistemas multi-variáveis. Entretanto, uma abordagem muito mais simples e eficiente para sistemas multi-variáveis é apresentada neste trabalho, através do uso da técnica de Análise por Componente Principal (*Principal Component Analysis* – PCA).

Nesta metodologia, as ondaletas são utilizadas para detecção de pontos estacionários em um sinal de processo, em um algoritmo baseado em código sub-banda de dois canais (*two channel subband coder*) usando filtros espelhados em quadratura, segundo a metodologia de transformada das ondaletas estacionária. A tendência do sinal de processo é extraída das medidas por um processo multi-escalas baseado nas ondaletas, com eliminação de ruído randômico e erros não aleatórios. A tendência ("sinal limpo") ainda preserva as nuances do sinal original, porém sem picos anormais e ruído branco. O estado do processo é então medido usando um índice com variação entre 0 e 1, de acordo com o módulo da transformada da ondaleta da tendência do sinal de processo e dos dados históricos. Este índice traz uma série de vantagens visto que pode ser utilizado, por exemplo, para compressão de dados e determinação de pontos ótimos de operação.

Entretanto, como a maioria dos processos químicos industriais são multi-variáveis, era necessário obter um procedimento que tornasse possível quantificar quão perto a unidade ou o equipamento está da condição estacionária. Normalmente, um índice único para todo o processo é recomendado e desejável, desde que seja de fácil análise. JIANG *et al.* (2003) sugere a seleção de variáveis-chave e o uso da regra de combinação de Dempster para a obtenção deste índice global. Deste modo, é necessário calcular um índice de estado para cada variável-chave e, pela regra de combinação, é necessário atribuir um peso para cada variável. Ao invés disto, propõe-se, neste trabalho, o uso de PCA, que evita atribuição de pesos e estima um índice de estado apenas, representativo de todo o processo.

Nas seções seguintes, apresenta-se a metodologia original de JIANG *et al.* (2003) e a técnica proposta neste trabalho, com exemplos de aplicações práticas em colunas de destilação industriais e em uma unidade de craqueamento catalítico.

4.1 Metodologia de Referência

Na metodologia proposta por JIANG *et al.* (2000 e 2003), a detecção de estados estacionários se baseia na transformada discreta das ondaletas. A técnica por eles apresentada consiste em um processo de extração de tendência de dados brutos baseado na análise multi-resolução das ondaletas e posterior determinação do estado do processo baseado nesta tendência extraída a várias escalas. O estado do processo é medido usando um índice com valor entre 0 e 1, de acordo com o módulo da derivada do sinal extraída pela transformada das ondaletas da tendência do sinal de processo, sendo que, quanto mais próximo de 1, mais perto do estado estacionário.

O índice de estado estacionário B é determinado basicamente pelas derivadas de primeira e segunda ordem da tendência f_s , representadas por WT_1 e WT_2 , conforme (4.1). A extração da tendência de um sinal, segundo a metodologia de JIANG *et al.* (2000 e 2003), foi apresentada no Capítulo 3.

$$\begin{aligned} |WT_1| < T_s \text{ e } |WT_2| < T_w \text{ então } B(t) &= 1 \\ |WT_1| > T_u \text{ então } B(t) &= 0 \end{aligned} \quad (4.1)$$

Os valores de corte T_s , T_w e T_u são estimados a partir de dados históricos utilizando-se as seguintes equações:

$$T_s = \sigma_{WT_1} \quad (4.2)$$

$$T_u = 3 \cdot \lambda_2 \cdot T_s \quad (4.3)$$

$$T_w = \overline{WT_2} \quad (4.4)$$

onde σ_{WT_1} é o desvio padrão do módulo dos coeficientes de detalhe de dados históricos ao último nível de decomposição e λ_2 é um parâmetro ajustável de qualidade, semelhante a λ_1 do Capítulo 3 (seção 3.2). O valor de corte T_u também pode ser informado diretamente (e não precisando ser calculado), se houver conhecimento do processo, ou pode ser calculado por métodos estatísticos e, então, informado ao algoritmo. O parâmetro T_w é um valor de corte para os coeficientes de detalhe de segunda ordem, e consiste no valor mediano destes ao último nível de decomposição.

A Eq. (4.1) calcula o índice de estado estacionário para os valores limites. De modo geral, o índice de estado estacionário pode ser calculado de acordo com a Eq. (4.5), segundo formulação proposta por JIANG *et al.* (2003).

$$B(t) = \begin{cases} 0 & \text{se } \theta(t) \geq T_u \\ \xi[\theta(t)] & \text{se } T_s < \theta(t) < T_u \\ 1 & \text{se } \theta(t) \leq T_s \end{cases} \quad (4.5)$$

onde as funções $\xi(\theta(t))$ e $\theta(t)$ são dadas por

$$\xi(\theta(t)) = \frac{1}{2} \left[\cos \left(\frac{\theta(t) - T_s}{T_u - T_s} \cdot \pi \right) + 1 \right] \quad (4.6)$$

$$\theta(t) = |WT_1(t)| + \gamma \cdot |WT_2(t)| \quad (4.7)$$

sendo

$$\gamma = \begin{cases} 0 & \text{se } |WT_2| \leq T_w \\ (|WT_2| - T_w) / (2T_w) & \text{se } T_w < |WT_2| < 3T_w \\ 1 & \text{se } |WT_2| \geq 3T_w \end{cases} \quad (4.8)$$

4.1.1 Sistemas Multi-variáveis

Para os sistemas multi-variáveis, JIANG *et al.* (2003) sugere a seleção de variáveis-chave, calculando o índice de estado estacionário para cada uma. As variáveis-chave são aquelas representativas do sistema analisado, que englobam todo o processo, mas que não sejam correlacionadas. Após esta seleção de variáveis e cálculo do índice, de acordo com o procedimento descrito na seção anterior, precisa-se combinar estes índices de modo a obter um índice único para todo o sistema em estudo. Isto é feito segundo a regra de combinação de Dempster (SHAFER, 1976).

$$B_m(t) = \prod_{i=1}^N [B_i(t)]^{w_i / \sum w_i} \quad (4.9)$$

onde B_i é o índice de estado estacionário de cada variável i em cada instante de tempo t , w_i é a parcela de influência de cada variável no processo (peso) e B_m é o índice global.

Esta metodologia, entretanto, é *off-line* e apresenta algumas questões que devem ser consideradas. O primeiro ponto é a seleção das variáveis-chave, que depende do conhecimento de processo, pois as variáveis devem ser não-correlacionadas e devem englobar todo o sistema. Outro ponto é que, na regra de combinação de Dempster, alguns pesos w_i devem ser atribuídos, como se observa na Eq. (4.9), considerando a contribuição que cada variável exerce no processo.

Em virtude dos pontos acima mencionados, desenvolveu-se uma nova metodologia para sistemas multi-variáveis, baseada em Análise por Componente Principal (PCA), conforme será apresentado na próxima seção.

4.2 PCA e Sistemas Multi-variáveis

A técnica proposta baseada em PCA consiste na combinação das variáveis do processo multi-variável em novas variáveis, ortogonais, que são combinações lineares das variáveis originais. A escolha das variáveis não requer que sejam não-correlacionadas, e não há necessidade de atribuição de pesos para a combinação delas. Além disso, menos esforço computacional é requerido. Estas vantagens tornam a metodologia facilmente aplicável para sistemas multi-variáveis.

A detecção de pontos estacionários baseada em componentes principais para uma janela de dados inicia com uma redução de dimensão através do uso da técnica de PCA. As variáveis selecionadas, por exemplo temperatura e vazão (que fecham os balanços de energia e massa de um processo), são escalonadas (centralizadas na média e normalizadas pelo desvio padrão) e, então, linearmente combinadas pela técnica de PCA. Resultam novas variáveis, cada qual contendo informação de todas as variáveis originais. Não são consideradas todas as novas variáveis, mas apenas uma, que é a que apresenta a maior quantidade de informação. Se, por exemplo, um sistema apresenta 10 variáveis de controle de interesse, a combinação

por PCA reduzirá em novas variáveis, das quais apenas uma é suficiente para análise de estacionários, conforme poderá ser visto na seção 4.3.4.

Esta nova variável pode servir, então, de parâmetro de entrada para o cálculo de transformada de ondaletas para identificação de estados estacionários, conforme procedimento descrito nas seções 3.2 e 4.1. O índice de estado estacionário calculado para esta variável já representa o estado de todo o processo em análise, não sendo necessário calcular para outras variáveis ou estimar pesos para as variáveis.

4.3 Estudos de Caso

Para comparar as duas técnicas apresentadas e validar a metodologia proposta, alguns estudos de caso foram realizados, com exemplos de aplicações industriais. Foram estudadas duas colunas de destilação e uma unidade de craqueamento catalítico.

4.3.1 Metodologias

A primeira metodologia, denominada neste trabalho de técnica de variáveis-chave, é a proposta multi-variável de JIANG *et al.* (2003), que faz uso da regra de Dempster para combinação das variáveis. A seleção das variáveis é baseada em conhecimento do processo e, para verificar se são independentes, uma análise simples de correlação é feita, utilizando a função *corrcoef* do MATLAB®. Nos exemplos que serão apresentados, as variáveis foram consideradas com o mesmo grau de importância sobre o processo, ou seja, receberam o mesmo peso na regra de combinação de Dempster, resultando na Eq. (4.10).

$$B_m(t) = \prod_{i=1}^N [B_i(t)]^{1/N} \quad (4.10)$$

A Eq. (4.10) indica que uma coluna de destilação será considerada em estado estacionário ($B_m = 1$) se todas as variáveis estiverem estacionárias no mesmo instante de tempo. Ao contrário, se uma das variáveis estiver transiente, a condição de estado não-estacionário para o sistema será apresentada por $B_m = 0$.

A outra metodologia é a proposta de utilização de PCA para a obtenção de uma única variável global e um único índice de estado estacionário, na técnica que aqui chama-se simplesmente de PCA.

4.3.2 Influência da Constante de Tempo

Alguns parâmetros de entrada são necessários para o algoritmo de detecção de estacionários, o qual foi implementado no MATLAB®. Além do vetor de dados da variável, deve-se informar a constante de tempo, estimada conforme descrito na seção 3.2.1, e um período de dados considerados históricos.

A constante de tempo é calculada para cada variável. Como estamos trabalhando com sistemas multi-variáveis, o valor considerado da constante de tempo é uma média aritmética das constantes de tempo das variáveis.

Entretanto, se a constante de tempo não puder ser estimada com tanta precisão, deve-se ter idéia de qual o erro que isto acarreta. Para tanto, avaliou-se qual é o comportamento da constante de tempo com relação à qualidade da tendência obtida, ou seja, como a constante de tempo influencia na aproximação final do sinal original. Sabe-se (Eq. (3.6)) que quanto maior o valor da constante de tempo de resposta τ , ou seja, quanto mais lento é o sistema, maior é o número de decomposições que são realizadas, conforme mostra a **Tabela 4.1**, e, portanto, mais suave fica a tendência final.

Tabela 4.1: Relação entre a constante de tempo e o número de decomposições.

τ	s
2	1
3 a 5	2
6 a 11	3
12 a 22	4
23 a 30	5
46 a 90	6

A **Figura 4.1** exemplifica melhor o que representa a Eq. (3.6). A figura apresenta um sinal ruidoso com duas variações do tipo degrau (tempos 100 e 200) e um pico anormal (tempo 242), o qual foi filtrado, segundo o algoritmo da transformada das ondaletas, a vários níveis de decomposição, de acordo com a constante de tempo de resposta.

Os resultados da **Figura 4.1** comprovam que quanto menor o valor da constante de tempo, ou seja, menor o número de decomposições, menos suave é a tendência e mais resquícios de ruído ela pode apresentar. Isto pode ser entendido pelo fato de que um processo que tem uma dinâmica rápida não pode ser suavizado muito, pois senão perderá suas características dinâmicas. Por outro lado, um sistema mais lento, se pouco decomposto, não trará as informações reais de variação do processo, pois será tomado por ruído, que pode mascarar a verdadeira característica do sinal.

Entretanto, observa-se que, com o aumento do número de decomposições, a aproximação pode se tornar tão suave a ponto de alterar o tempo de algum evento ocorrido no sinal. Por exemplo, se a decomposição for feita com 6 níveis, a tendência será tão suave que o degrau do instante de tempo 100 do sinal original começa em torno do instante de tempo 80 no sinal aproximado. Da mesma forma, o degrau negativo do tempo 200 do sinal original termina, na tendência extraída, no tempo 220.

Além disso, observa-se que o pico anormal do tempo 240 só é removido da tendência a partir do terceiro nível de decomposição. Isto ocorre pois a detecção de picos anormais é baseada na constante de tempo de resposta, que determina o número de níveis de decomposição.

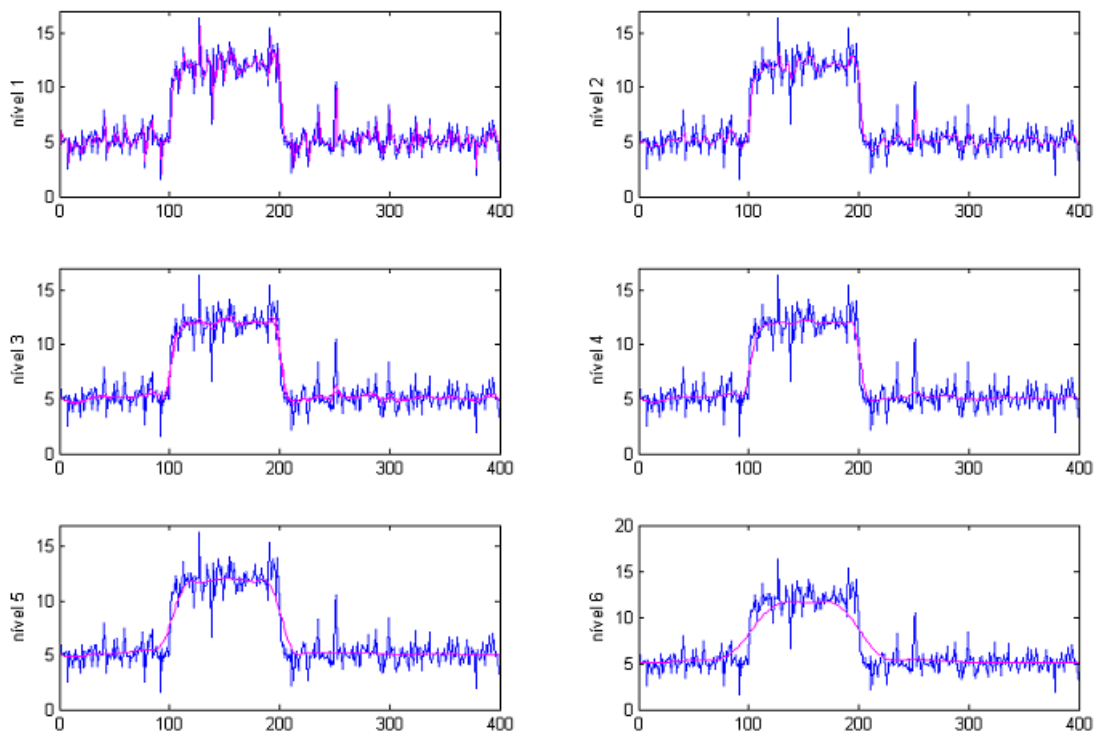


Figura 4.1: Sinal original e suas tendências em vários níveis de decomposição.

Portanto, se observa que a constante de tempo apresenta grande influência na tendência do sinal a ser obtida e na remoção de picos anormais. Como consequência, a detecção de estacionários é completamente alterada, conforme mostra a **Figura 4.2**, a qual foi obtida através da aplicação da metodologia a cada uma das seis decomposições ilustradas na **Figura 4.1**.

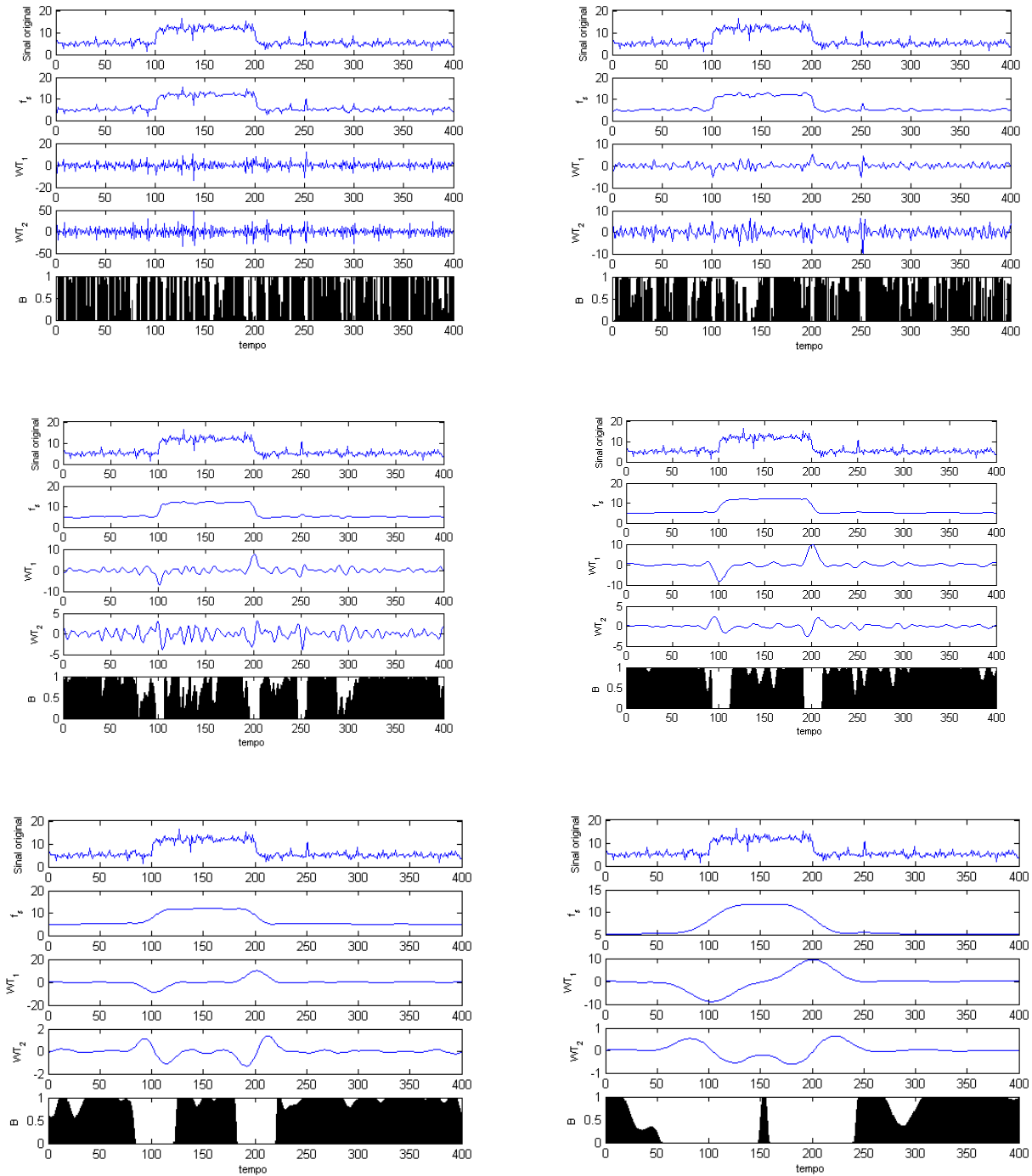


Figura 4.2: Para os 6 níveis de decomposição (ou $\tau = 2, 5, 10, 15, 30$ e 50): sinal original, tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B .

4.3.3 Influência do Período Histórico

Outro parâmetro de entrada do algoritmo das ondaletas é um período de dados históricos de processo. Portanto, é importante também avaliar a sua influência na obtenção da tendência do sinal.

A **Figura 4.3** apresenta um sinal semelhante ao apresentado na seção anterior, e três tendências obtidas para três períodos históricos diferentes. Neste caso, fixou-se a constante de tempo de resposta como 10 (3 níveis de decomposição) e modificou-se o período de dados históricos informado ao algoritmo. Foram avaliadas três situações. Na primeira, considerou-se o período de dados históricos como sendo os pontos de tempo de 1 a 200 deste mesmo sinal, o qual compreende uma mudança do tipo degrau. Nas duas outras situações, os períodos históricos foram considerados como sendo de 211 a 410 e 501 a 700 deste mesmo sinal. Este último é um período composto apenas de ruído basicamente, e, no anterior, um pico anormal (em 242) é englobado no período histórico.

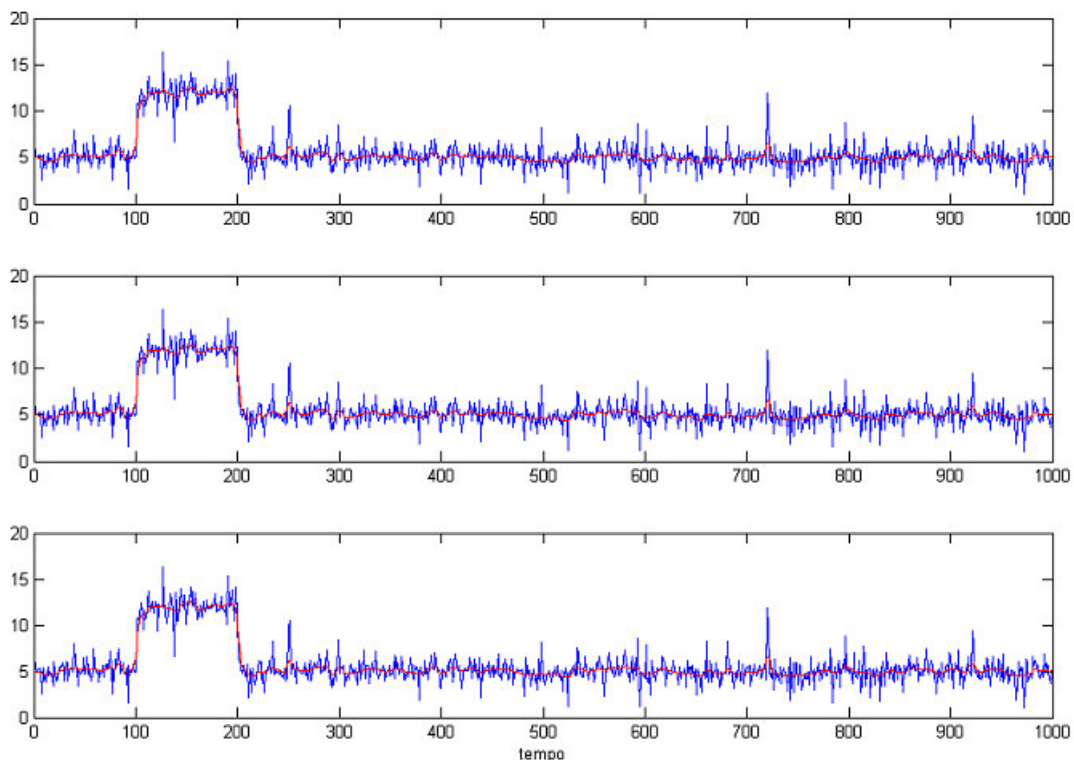


Figura 4.3: Sinal original e suas tendências para diferentes períodos históricos selecionados.

O que se observa é que as tendências obtidas para os três períodos históricos foram praticamente as mesmas. Ou seja, neste exemplo, não se verificou influência de períodos históricos diferentes na obtenção de tendências. A presença de um pico anormal ou mesmo de um degrau no período histórico não afeta a tendência do sinal.

Entretanto, com o índice de estado estacionário se verifica que a seleção é importante, conforme pode ser visto na **Figura 4.4**. Isto ocorre porque o cálculo de valores de corte para detecção de estacionários é determinado em função do desvio padrão dos coeficientes de detalhe de dados históricos e, portanto, as características deste período alteram a magnitude destes valores de corte.

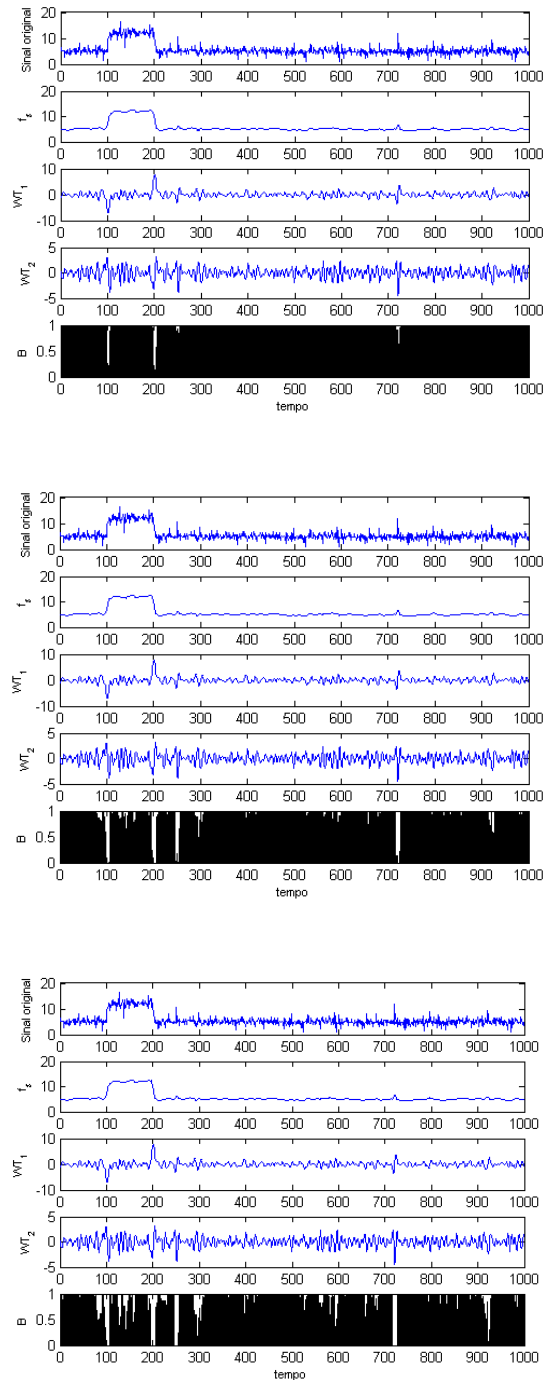


Figura 4.4: Detecção de estacionários baseada em três períodos históricos distintos.

O que se pode observar a partir da **Figura 4.4** é que o período histórico que continha a informação de um degrau (caso 1) gerou índices estacionários para todo o período, com exceção dos tempos no qual o degrau ocorreu. Nestes instantes, o índice se aproximou de zero, caracterizando uma condição não-estacionária. Os dois picos anormais (tempos 240 e 720) foram corretamente rejeitados e o índice nestes instantes também foi considerado nulo.

No segundo caso, onde o período histórico continha informação de pico anormal, o índice ficou unitário para a maioria dos pontos de tempo. Foram detectadas quatro situações

de condição não estacionária correspondentes aos instantes de ocorrências dos degraus e dos dois picos anormais. Os picos anormais, neste caso, não foram descartados por estarem contidos também no período informado como histórico.

Na terceira situação, o índice de estado foi encontrado mais vezes como não-estacionário, além dos pontos de ocorrência de degraus e picos anormais. O período histórico para esta situação foi considerado como só possuindo ruído e, por isto, a detecção se torna mais “rígida” e mais instantes de tempo são calculados como fora do estacionário.

Um sinal com diferentes magnitudes de ruído, conforme o da **Figura 4.5**, também foi analisado para três condições de períodos históricos. No primeiro caso, o período histórico foi considerado como sendo do intervalo de tempo de 1 a 500, no segundo caso foi considerado de 1001 a 1500 e no último, de 2001 a 2500.

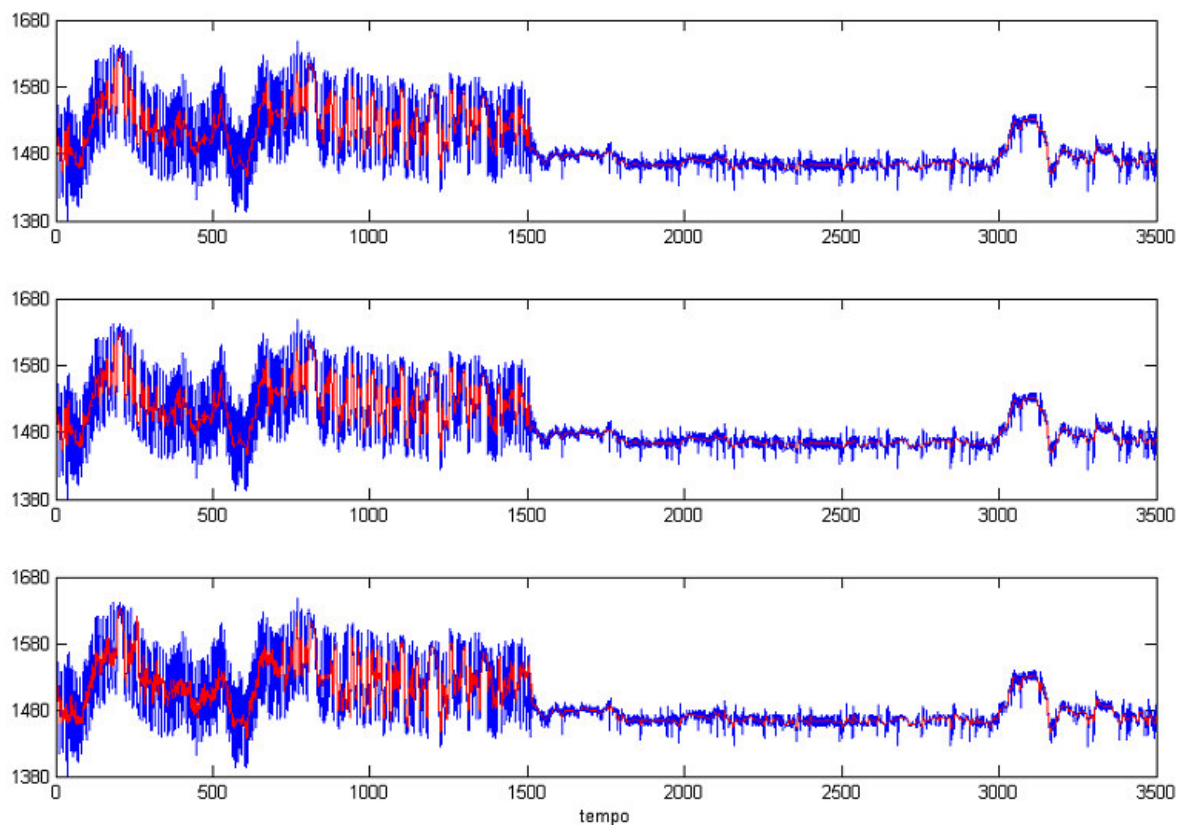


Figura 4.5: Sinal original e suas tendências para diferentes períodos históricos selecionados.

Para este exemplo, o que se observa é que existe pouca diferença entre as tendências obtidas. Nas duas primeiras situações, elas são praticamente iguais, enquanto a terceira tendência apresenta uma sutil diferença para as duas anteriores, mais visível na região em torno dos tempos 200 a 500.

Entretanto, devido aos valores de corte calculados pelo algoritmo, o cálculo de índices de estados estacionários gera resultados bastante diferentes, conforme mostra a **Figura 4.6**.

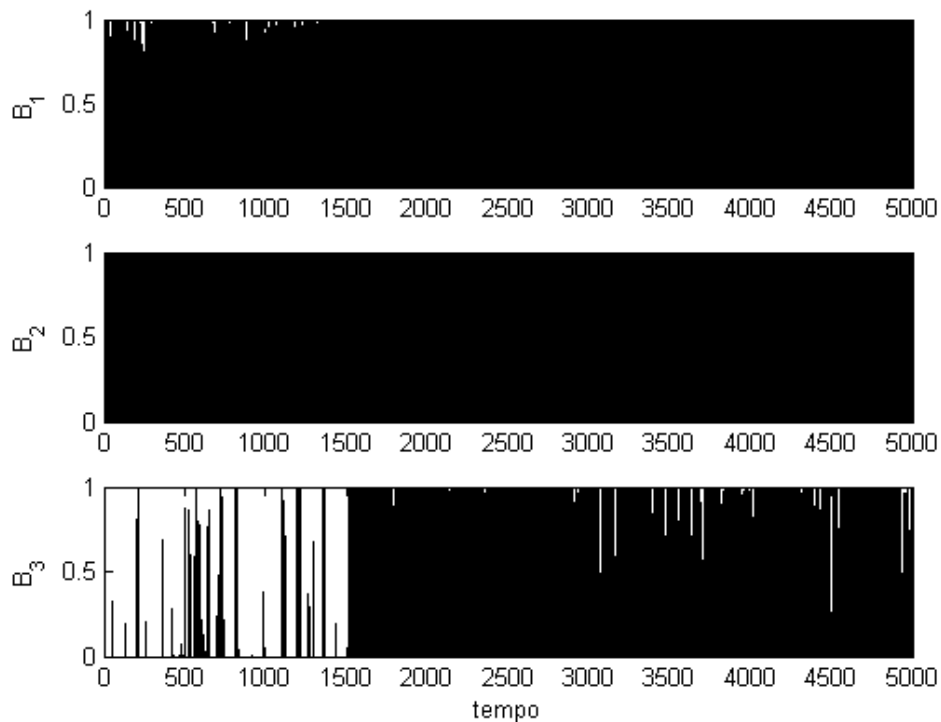


Figura 4.6: Índices de estado estacionário de acordo com as três tendências obtidas.

Os resultados da **Figura 4.6** são coerentes com o que se espera. Quando o período histórico foi considerado como os primeiros 500 pontos do sinal, que possui componentes de ruído branco e colorido, o algoritmo calculou todo o período como estacionário. O mesmo aconteceu para a segunda situação, onde o período de dados históricos era composto por ruído branco de grande amplitude. Por fim, para o período histórico que continha ruído branco, mas de baixa amplitude, o algoritmo calculou regiões não-estacionárias, compreendidas entre os tempos 1 e 1500, que é um intervalo de tempo em que o sinal de processo apresenta ruído de grande magnitude. Como o período histórico continha ruído de baixa amplitude, os valores de corte calculados foram baixos. Ao verificar uma região de maior variabilidade, então, os valores de processo excederam estes valores de corte e a condição de pontos não-estacionários foi verificada.

Pode-se concluir, portanto, que as características do período histórico selecionado não influenciam na obtenção da tendência do sinal, mas influenciam no cálculo dos índices de estado estacionário.

4.3.4 Coluna de Destilação 1

Um bom exemplo para estudo comparativo das metodologias de variáveis-chave e PCA é uma coluna de destilação. Neste caso, selecionou-se uma coluna atmosférica de tolueno de uma unidade industrial petroquímica. Esta coluna é alimentada por uma corrente de fundo de uma coluna de benzeno. A coluna de tolueno possui 60 pratos valvulados e a alimentação ocorre nos pratos 30 ou 36. A temperatura do estágio 20 é controlada através da

vazão do refeedor. Esta coluna possui 5 medições de vazão, 9 medições de temperatura ao longo da coluna e uma medida de pressão no topo, conforme mostra a **Figura 4.7**. O número no *tag* dos medidores de temperatura é de acordo com o estágio de equilíbrio em que se encontram.

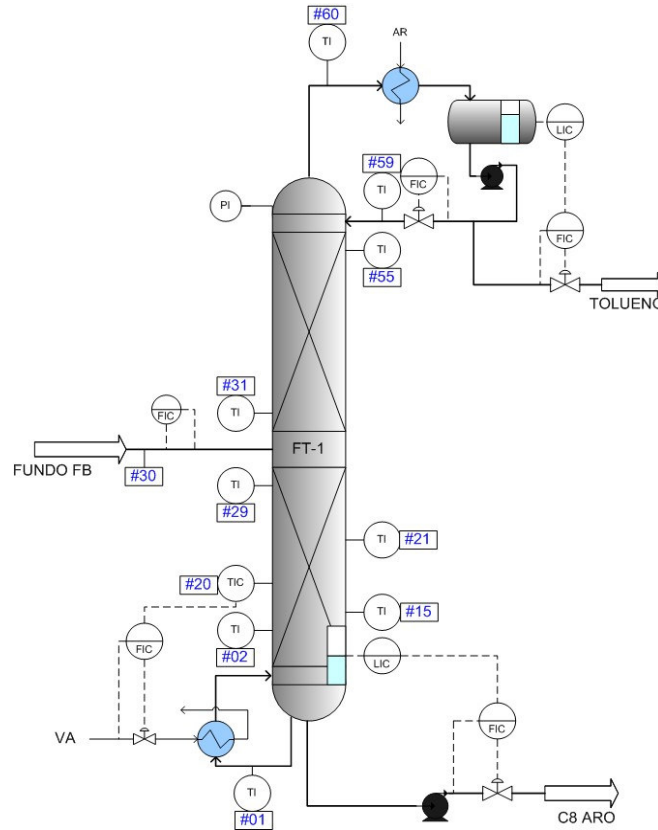


Figura 4.7: Coluna de destilação de tolueno e seus medidores.

Parâmetros do algoritmo: O parâmetro de tempo de resposta considerado foi $\tau = 30$ min.

Determinação de estacionários por variáveis-chave: As variáveis-chave foram selecionadas através de uma análise tradicional de correlação, de modo que as variáveis escolhidas foram as temperaturas TI02 e TI21, a pressão no topo da coluna PI18 e o nível no fundo LIC09. Estas variáveis foram selecionadas dentre todas as variáveis da coluna de destilação, após passarem pela análise de correlação cruzada. As variáveis-chave estão representadas na **Figura 4.8**. Foram escolhidos 2000 pontos de dados, equivalentes a quase um dia e meio de processo (tempo de amostragem de 1 minuto). O período de dados considerados históricos foi entre os pontos de tempo de 1 a 600.

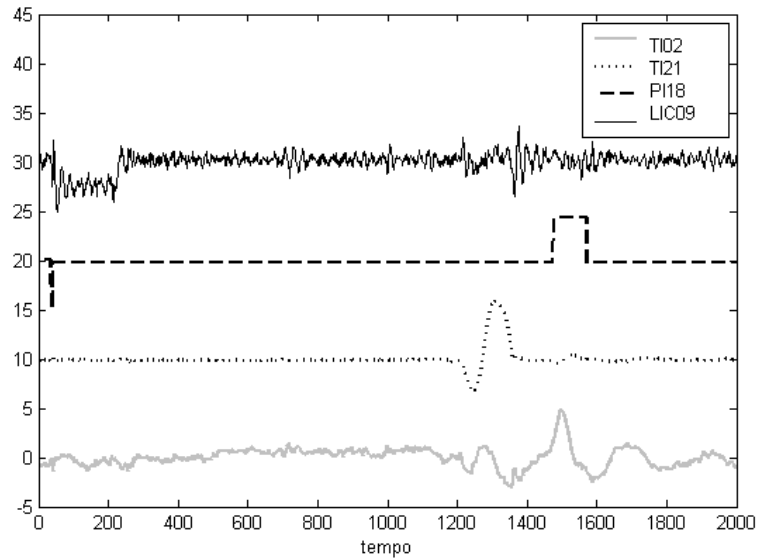


Figura 4.8: Variáveis-chave selecionadas.

Para cada uma das variáveis, então, realizou-se o processo de detecção de estacionários. Os resultados são apresentados nas **Figura 4.9**, **Figura 4.10**, **Figura 4.11** e **Figura 4.12**.

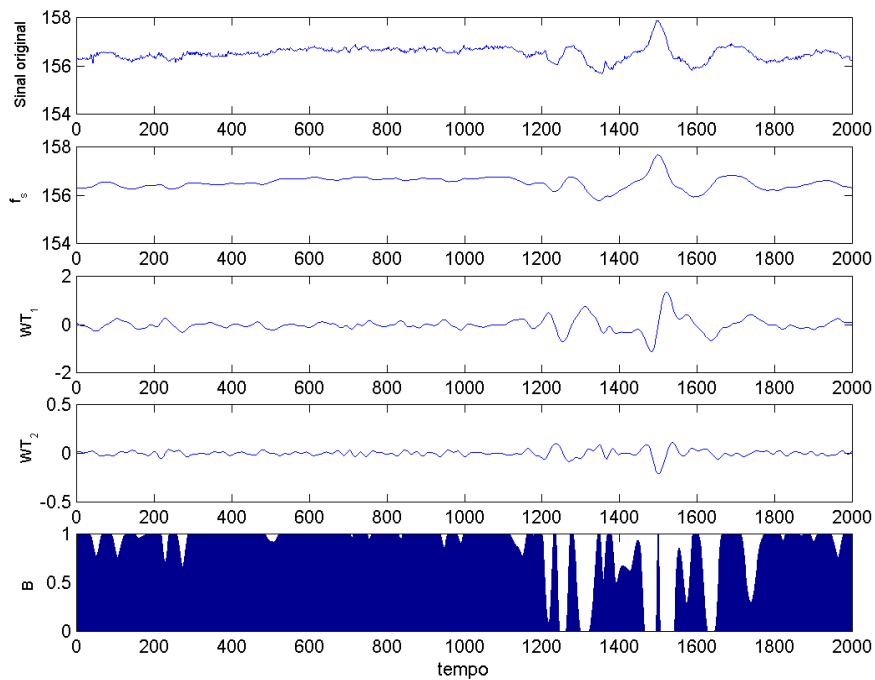


Figura 4.9: Variável-chave temperatura do estágio 2 (TI02), representada por f , tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B .

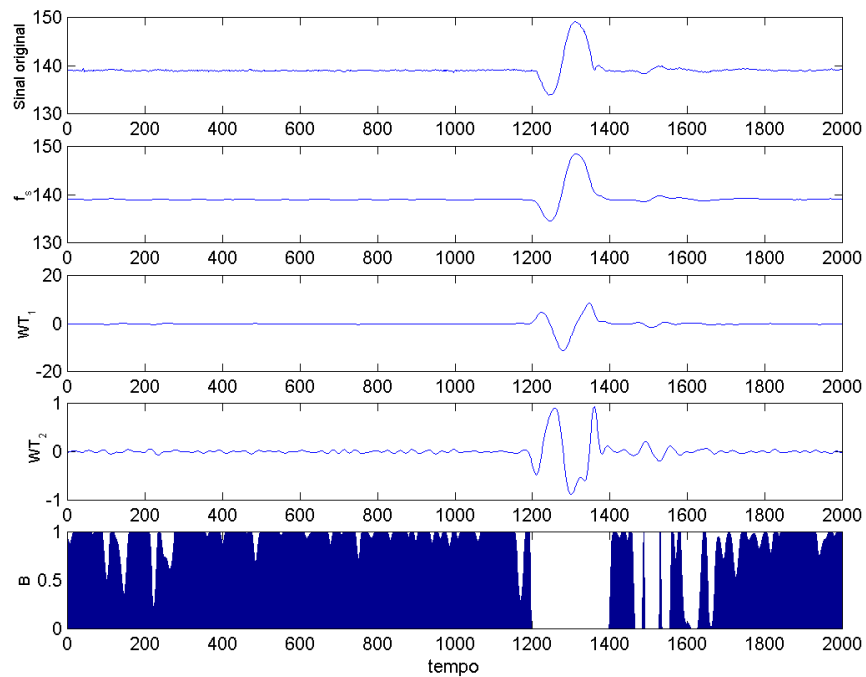


Figura 4.10: Variável-chave temperatura do estágio 21 (TI21), representada por f , tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B .

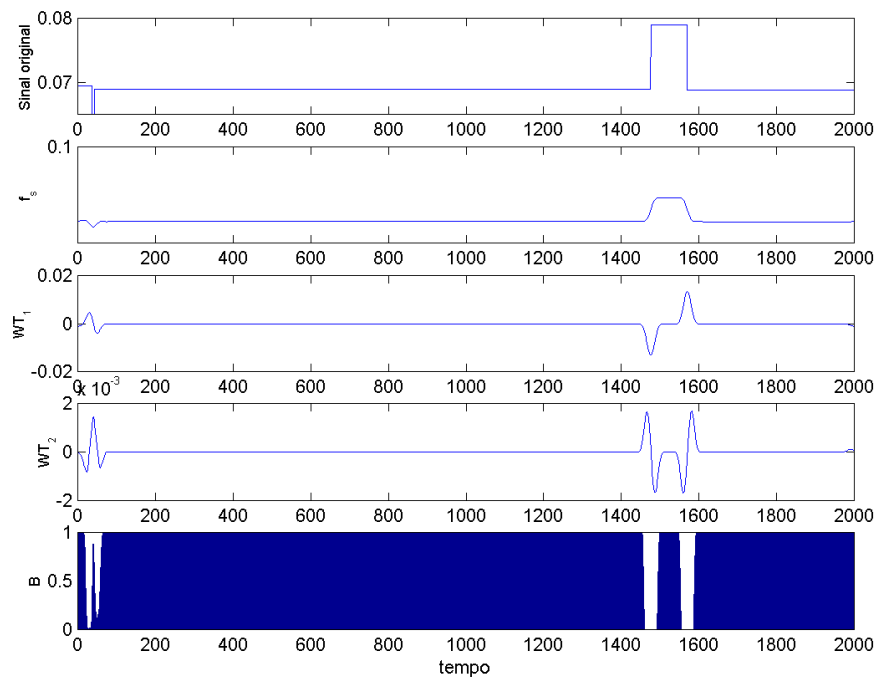


Figura 4.11: Variável-chave pressão no topo (PI18), representada por f , tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B .

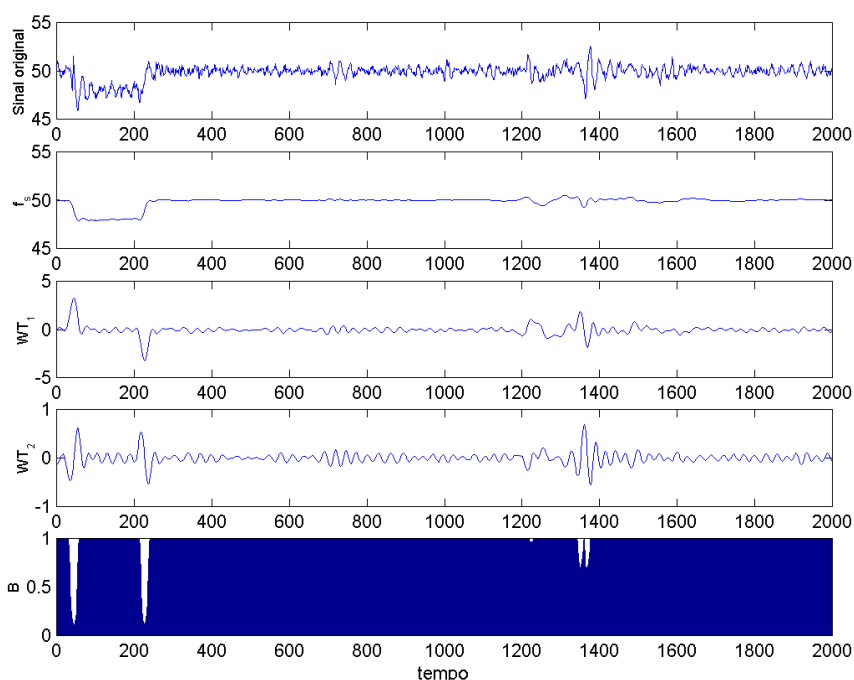


Figura 4.12: Variável-chave nível de fundo (LIC09), representada por f , tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado B .

Como resultado da combinação destes 4 índices, tem-se um índice único B_m representativo de toda a coluna, apresentado na **Figura 4.13**.

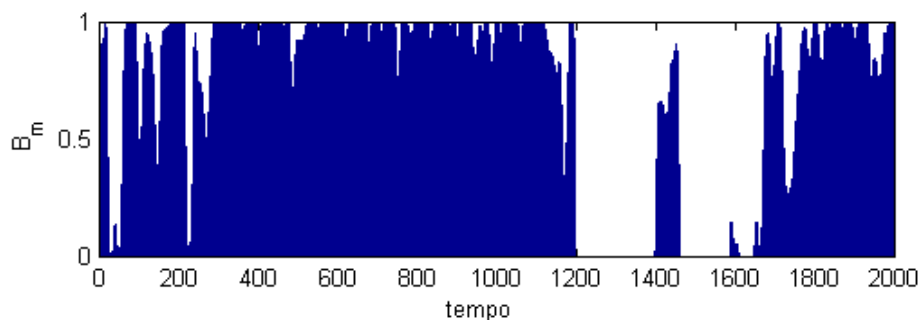


Figura 4.13: Índice de estado estacionário B_m para a coluna de destilação, obtido por seleção de variáveis-chave.

Determinação de estacionários por PCA: As variáveis selecionadas para a análise por PCA foram as temperaturas que formam o perfil da coluna de destilação (menos as temperaturas indicadas por TI30 e TI59, por não estarem atreladas ao equilíbrio termodinâmico). As medições de temperatura desta coluna, para um dado período de tempo selecionado, estão representadas na **Figura 4.14** e na **Figura 4.15**, onde estão centralizadas na média e normalizadas pelo desvio padrão (conforme mostrado no Capítulo 2), e foram plotadas deslocadas uma com relação à outra para uma melhor visualização.

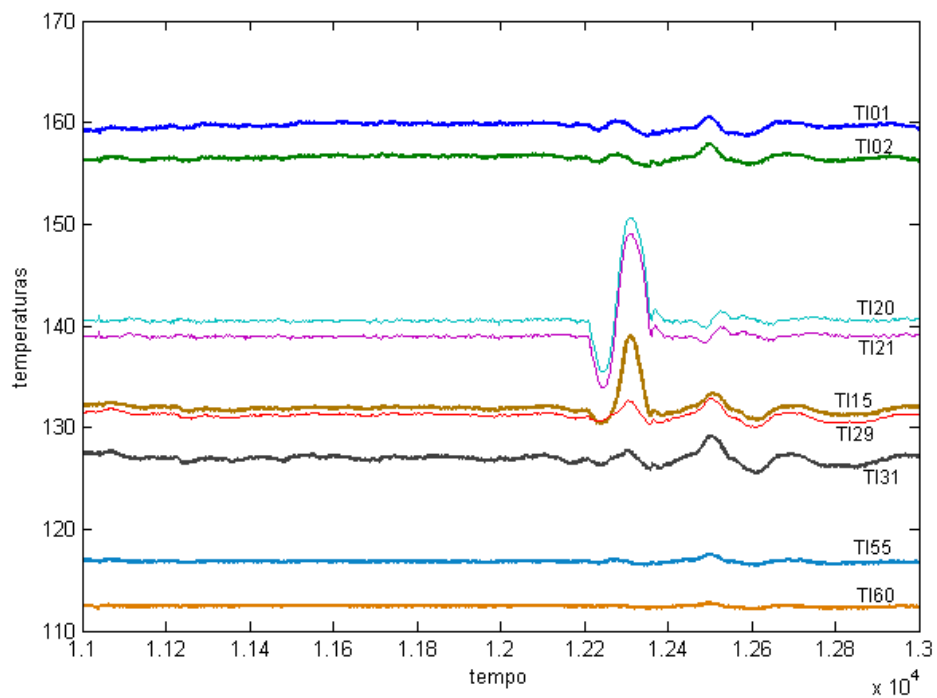


Figura 4.14: Perfil de temperatura da coluna de tolueno.

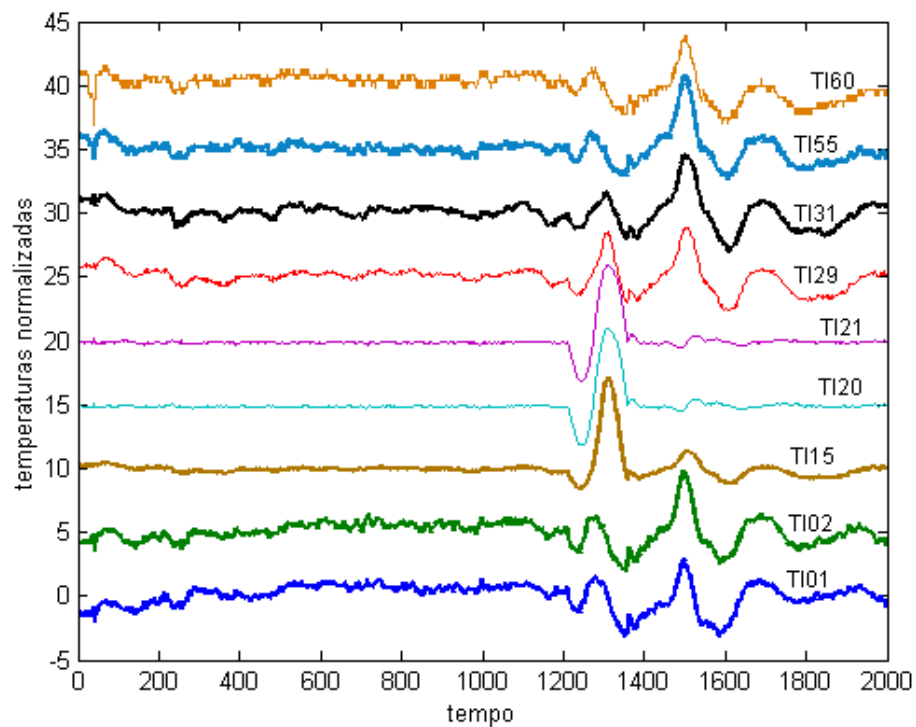


Figura 4.15: Temperaturas escalonadas da coluna de tolueno, deslocadas para facilitar a visualização.

Salienta-se que as variáveis de análise por PCA são diferentes das analisadas por variáveis-chave pois, neste caso, as variáveis devem ser pouco correlacionadas, o que não

ocorre no perfil de temperaturas. Isto quer dizer que o perfil de temperaturas não foi empregado no método das variáveis-chave por ser altamente correlacionado. Por outro lado, variáveis de pressão, vazão ou nível poderiam também ser selecionadas na análise por PCA (como foi feito no exemplo que será apresentado na seção seguinte, em que variáveis de temperatura e vazão foram analisadas concomitantemente).

Estas variáveis foram linearmente combinadas por PCA, gerando outras nove variáveis ortogonais, cujos valores singulares indicaram que duas variáveis apenas são significativas em termos de variância do processo, conforme apresentado no vetor *base* em (4.11), que consiste nos nove valores singulares da matriz utilizada para calcular o PCA.

$$base = [97,2 \quad 78,9 \quad 37,7 \quad 21,2 \quad 16,0 \quad 10,2 \quad 7,9 \quad 5,0 \quad 0,4] \quad (4.11)$$

Das nove variáveis considera-se que apenas duas sejam significativas, pois dois valores singulares (97,2 e 78,9) são bem maiores que os demais. Estas duas novas variáveis obtidas por PCA são uma combinação linear das variáveis originais, segundo a proporção apresentada em (4.12), onde a primeira coluna representa a combinação que gera a primeira variável ortogonal t_1 e a segunda coluna a segunda variável ortogonal t_2 . Ou seja, t_1 é uma combinação de 12% da primeira variável original, 15% da segunda variável original e assim por diante. Da mesma forma, t_2 é uma combinação de 6% da primeira variável original, 7% da segunda variável original e assim sucessivamente. Estas duas novas variáveis ortogonais são mostradas na **Figura 4.16**. A segunda variável foi plotada deslocada para cima visando facilitar a visualização. Para este primeiro exemplo, estas duas variáveis foram selecionadas para análise de estacionário para mostrar que os resultados são semelhantes, indicando que apenas uma delas seja suficiente para a estimação do índice de estado estacionário.

$$comb = \begin{bmatrix} 0,1206 & 0,0618 \\ 0,1520 & 0,0712 \\ 0,0745 & 0,2042 \\ 0,0165 & 0,2266 \\ 0,0166 & 0,2266 \\ 0,1572 & 0,0718 \\ 0,1613 & 0,0161 \\ 0,1521 & 0,0580 \\ 0,1492 & 0,0636 \end{bmatrix} \quad (4.12)$$

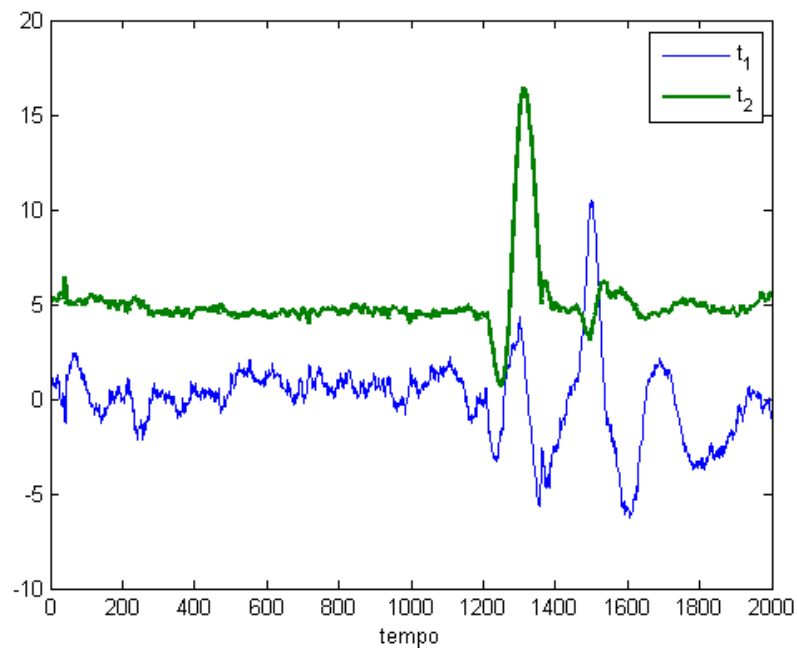


Figura 4.16: Novas variáveis ortogonais t_1 e t_2 .

Então, para cada variável t_1 e t_2 foi feita uma análise de estado estacionário e um índice de estado foi calculado para cada uma em cada instante de tempo. O período histórico foi considerado como os primeiros 600 pontos de t_1 e t_2 .

As **Figura 4.17** e **Figura 4.18** apresentam a análise de estacionários, onde f é o sinal original (t_1 ou t_2), fs é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado estacionário.

A combinação dos dois índices de estado estacionário, se preciso, poderia ser de acordo com a Eq. (4.10), e resultaria em um índice único B_m , mostrado na **Figura 4.19**, o qual representa o estado de toda a coluna. Como pode ser visto, as variáveis t_1 e t_2 apresentam uma faixa comum onde os resultados são os mesmos, ou seja, os mesmos períodos de tempo estacionários e não-estacionários. Entretanto, a variável t_1 apresenta mais períodos de condição não estacionária, em virtude de capturar a maior variabilidade. Pela **Figura 4.19**, podemos observar que B_m é muito semelhante ao B obtido para t_1 . Em virtude disto, pode-se considerar que não é necessário analisar o estado de todas as variáveis ortogonais. A análise de somente a primeira variável ortogonal (t_1) já contém a informação suficiente para a determinação dos estados estacionários da coluna, visto que a informação que as demais variáveis trazem está em grande parte contida nesta primeira variável ortogonal.

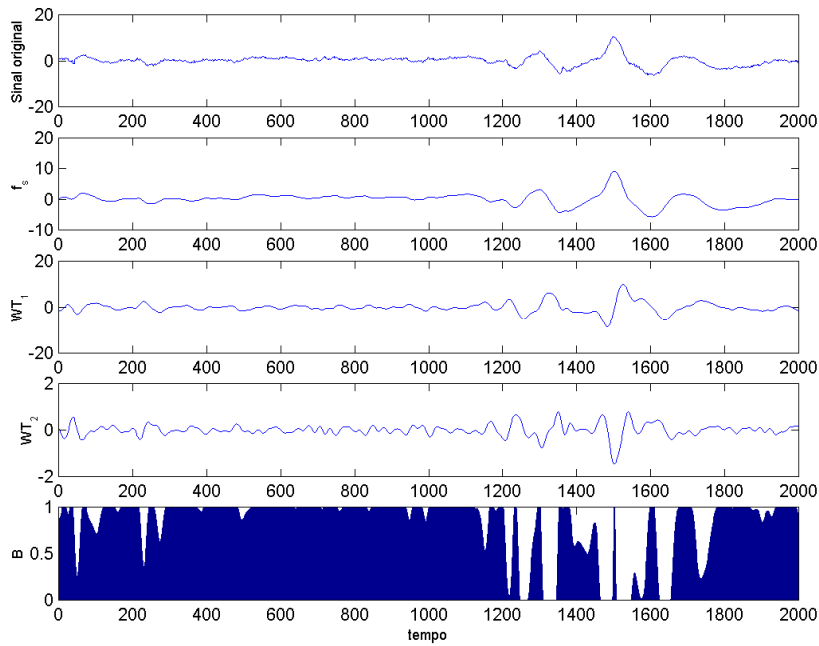


Figura 4.17: Resultados para a variável ortogonal t_1 , representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado para t_1 .

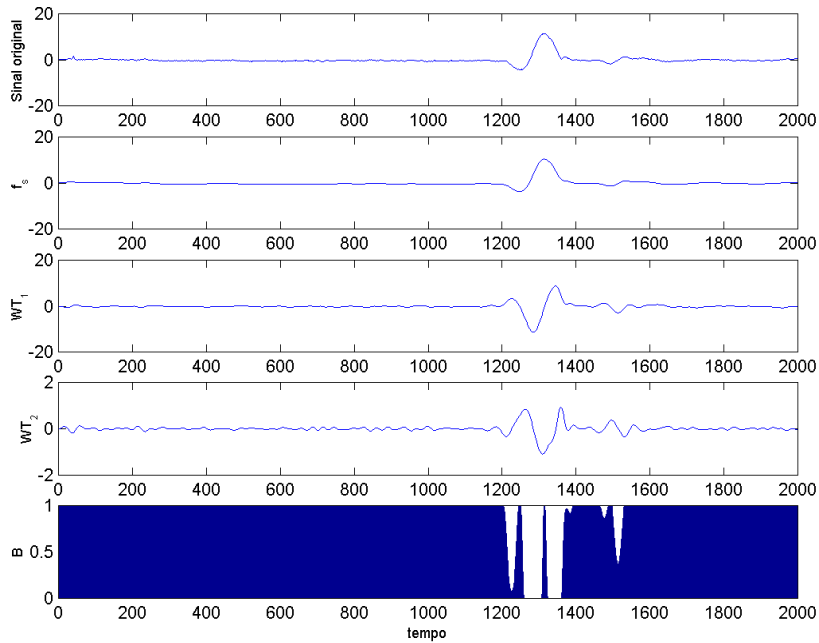


Figura 4.18: Resultados para a variável ortogonal t_2 , representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado para t_2 .

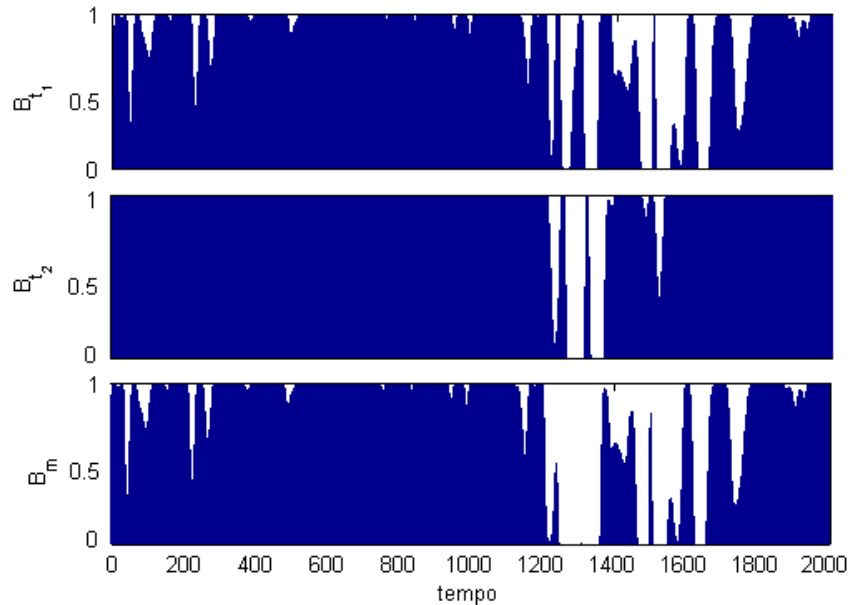


Figura 4.19: Índices de estado estacionário para as variáveis t_1 , t_2 e para a combinação B_m (para toda a coluna de destilação), obtidos pela utilização de PCA para combinação das variáveis.

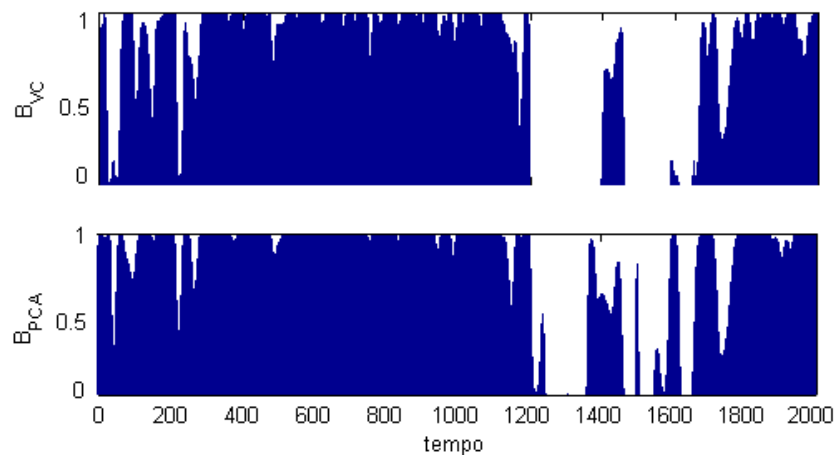


Figura 4.20: Índices de estado estacionário obtidos pela utilização de variáveis-chave (B_{VC}) e PCA (B_{PCA}).

Pela **Figura 4.20**, pode-se observar que os resultados do índice por PCA e variáveis-chave são similares para este exemplo, isto é, os instantes de tempo detectados estacionários por PCA são praticamente os mesmos encontrados por variáveis-chave, e a magnitude deles também é semelhante. Porém, PCA apresenta uma maneira muito mais fácil de lidar com as variáveis com relação à utilização de variáveis-chave, e não requer a atribuição de pesos na regra de combinação para sistemas multi-variáveis. As variáveis informadas são linearmente combinadas por PCA, sem a necessidade de se conhecer quais são as variáveis mais importantes e representativas do sistema, e qual é sua influência no processo. Ou seja, uma das vantagens de PCA é poder trabalhar com variáveis correlacionadas. Outra vantagem é a

estimação do estado para uma variável apenas, e não para várias, como é o caso de variáveis-chave, reduzindo o esforço computacional.

4.3.5 Coluna de Destilação 2

Um outro exemplo prático é o de uma coluna atmosférica de uma unidade de destilação da UN-RECAP. A unidade de destilação compreende os sistemas de dessalgação, torre de pré-fracionamento atmosférica, forno atmosférico e torre atmosférica. O petróleo cru reduzido, após ser aquecido no forno, é enviado para a fracionadora atmosférica, a qual tem por objetivo principal a produção de diesel linha, que é uma combinação de diesel pesado (diesel DD) com querosene, conforme pode ser visto na **Figura 4.21**. Outros produtos são a nafta pesada, produto de topo, e o resíduo atmosférico (RAT), produto de fundo.

A coluna atmosférica apresenta uma seção de refluxo circulante intermediário (RCI) para retirar calor da parte superior da coluna, indicado por Bateria de RCI (conjunto de condensadores). Assim, um aumento no refluxo de topo é desejado para diminuir a necessidade de RCI e, assim, aumentar a fração de nafta leve para ser incorporada ao querosene.

A nafta produzida é utilizada na produção de solventes e pode também ser incorporada à gasolina, e sua vazão mínima é restringida pela programação de produção. Mas o objetivo da produção é minimizá-la, pois o interesse é o diesel linha. O querosene, normalmente, incorporado ao diesel pesado, é maximizado quanto à parte leve, com restrição de ponto de fulgor. O diesel DD é sempre maximizado, sujeito às restrições de densidade, enxofre, T85%, T50% e ponto de entupimento.

A coluna possui medidores de temperatura e vazão, além de alguns medidores de pressão. Os medidores de temperatura e vazão considerados para detecção de estacionários por PCA (azul) estão destacados, na **Figura 4.21**, com os números de 1 a 8 e 9 a 13, respectivamente, e os medidores utilizados para análise com variáveis-chave (em laranja), são os de número 2 e 4 e 1 e 3, respectivamente.

Deseja-se que a temperatura de topo seja baixa para que a fração de nafta pesada seja reduzida. Ou seja, é esperado que a vazão FIC12 seja baixa. Do mesmo modo, o FIC de refluxo de diesel para a coluna deve ser minimizado, a ponto de apenas garantir o molhamento da panela de diesel.

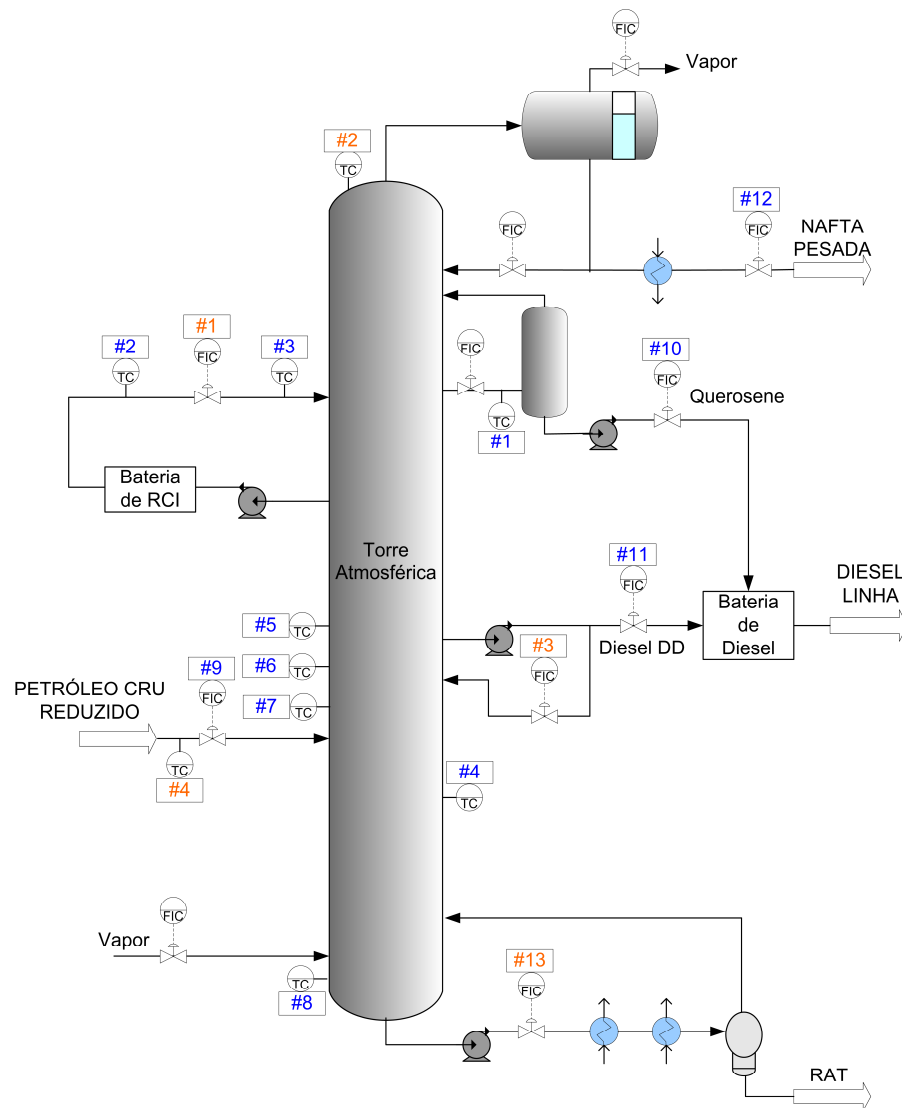


Figura 4.21: Coluna de destilação atmosférica, com as indicações de medições de temperatura e vazão.

Parâmetros do algoritmo: O parâmetro de tempo de resposta considerado foi $\tau = 80$ min.

Determinação de estacionários por variáveis-chave: As variáveis-chave selecionadas, que apresentam pouca correlação cruzada, foram as vazões de RCI (FIC1) e de refluxo de diesel (FIC3) e as temperaturas do topo (TC2) e de entrada na torre (TC4), que são variáveis manipuladas. As variáveis estão representadas na **Figura 4.22**. Foram escolhidos 5000 pontos de dados, equivalentes a quase três dias e meio de operação (tempo de amostragem de 1 minuto). O período de dados considerados históricos foram os pontos de tempo de 800 a 1500, que é uma região que está dentro da faixa de variação de temperatura de $\pm 1,5^\circ\text{C}$, considerada aceitável pela operação.

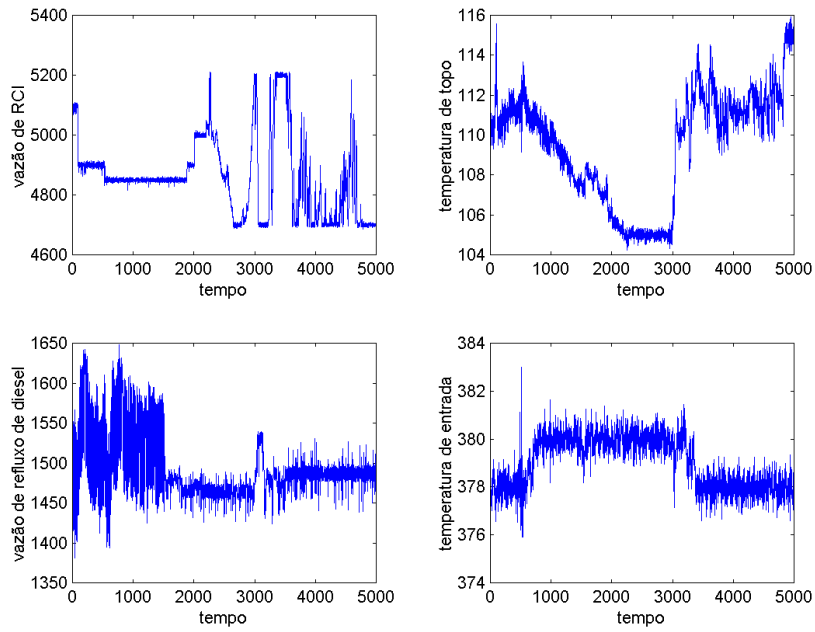


Figura 4.22: Variáveis-chave selecionadas.

Para cada variável se faz uma análise de estacionários, resultando nas **Figura 4.23**, **Figura 4.24**, **Figura 4.25** e **Figura 4.26**. A combinação dos quatro índices resulta no índice da coluna B_m , apresentado na **Figura 4.27**.

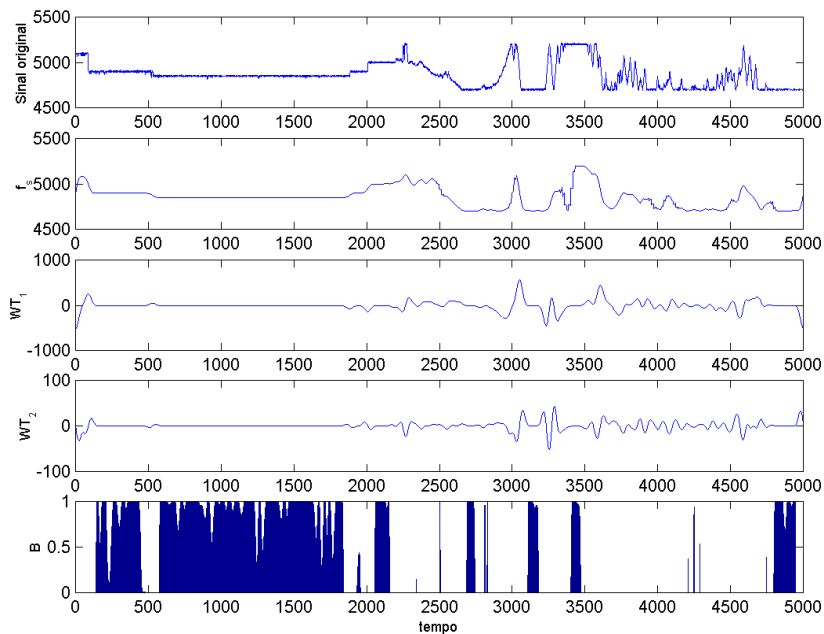


Figura 4.23: Resultados para a vazão de reciclo RCI, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

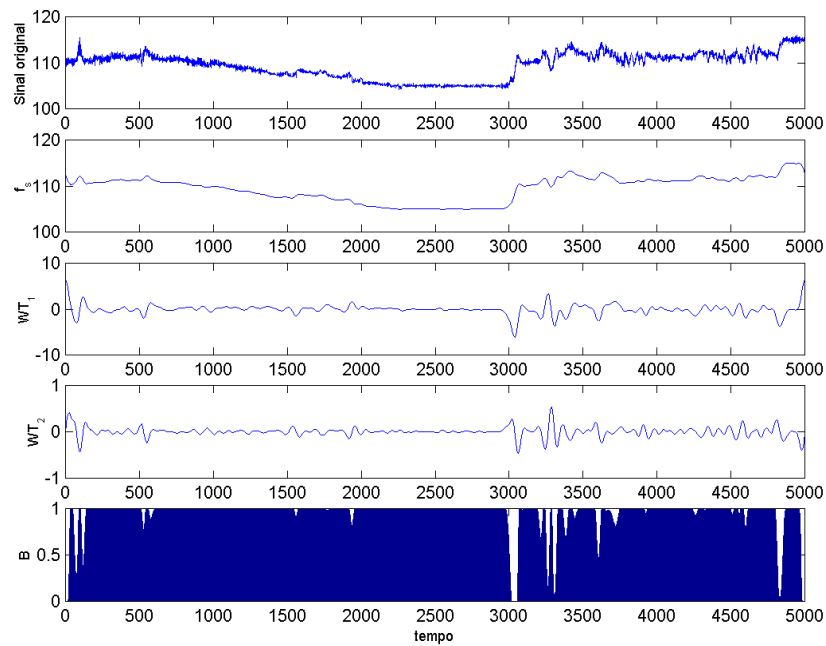


Figura 4.24: Resultados para a temperatura de topo, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

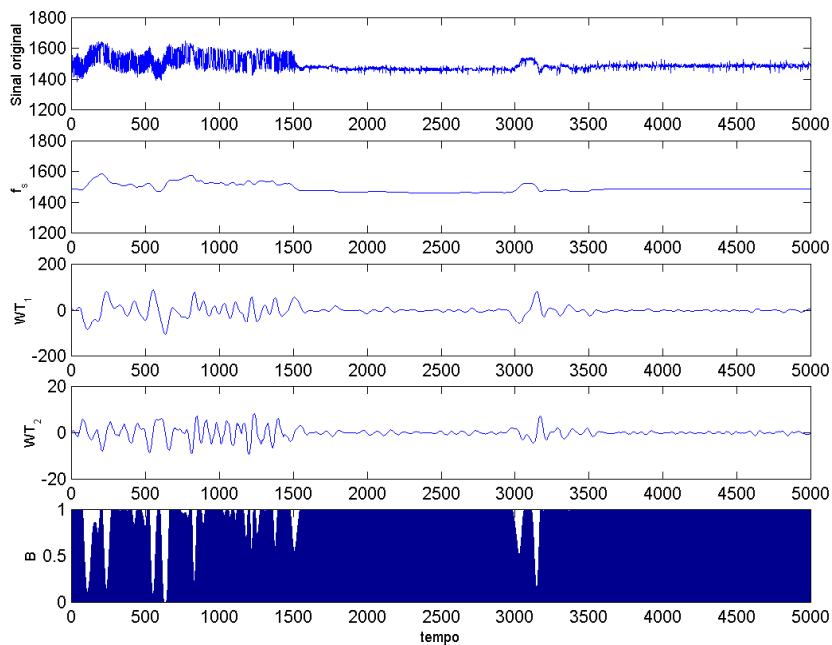


Figura 4.25: Resultados para a vazão de refluxo de diesel, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

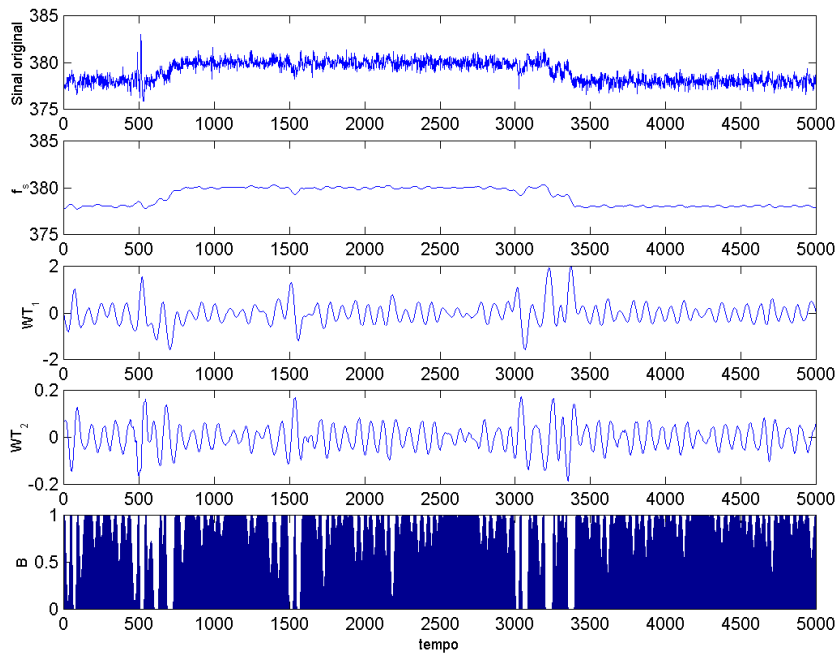


Figura 4.26: Resultados para a temperatura de entrada, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

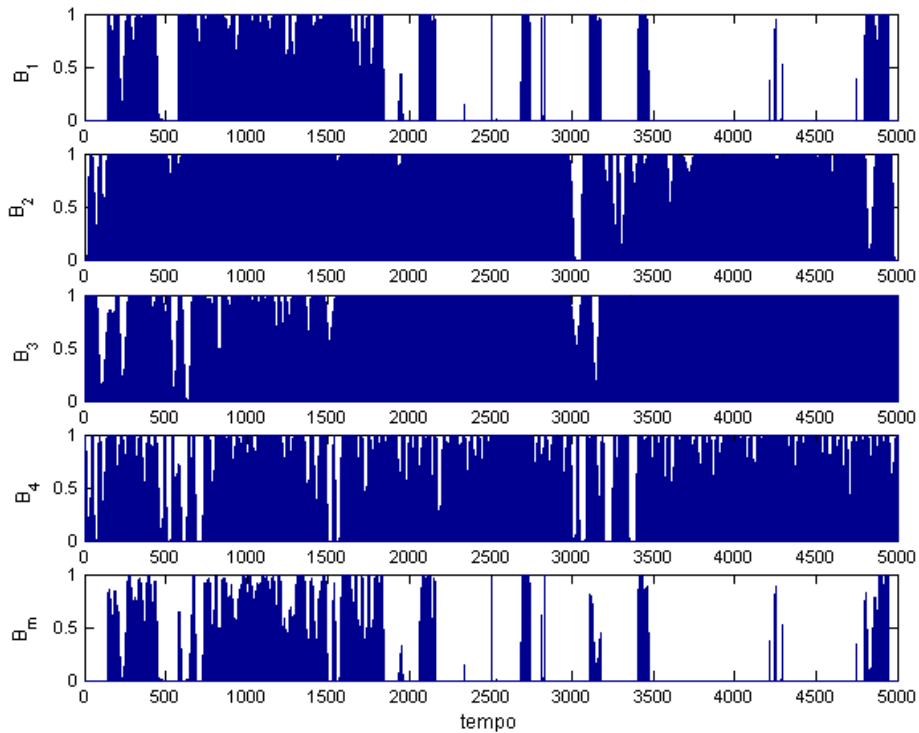


Figura 4.27: Índices de estado para as quatro variáveis-chave seleccionadas (vazão de reciclo, temperatura do topo, refluxo de diesel e temperatura de entrada) e o índice global da coluna B_m .

O índice global apresentou vários períodos de condição não estacionária, o que não está coerente com o modo como a coluna foi operada neste período. Cada variável independentemente analisada apresenta índices coerentes, com exceção da temperatura de entrada, que apresenta comportamento “oscilatório”, o qual foi capturado pelo índice de estado. Porém, a combinação destas variáveis em um único índice não foi adequada. Tendo as variáveis recebido o mesmo peso na combinação, observa-se que a coluna só é considerada estacionária se todas as variáveis estão estacionárias no mesmo instante de tempo.

Determinação de estacionários por PCA: Foram selecionadas as 8 medições de temperatura da coluna e mais 5 de vazões, para complementar com o balanço de massa. Estas variáveis foram avaliadas no mesmo período de tempo que foi realizada análise por variáveis-chave e o período histórico também foi considerado o mesmo.

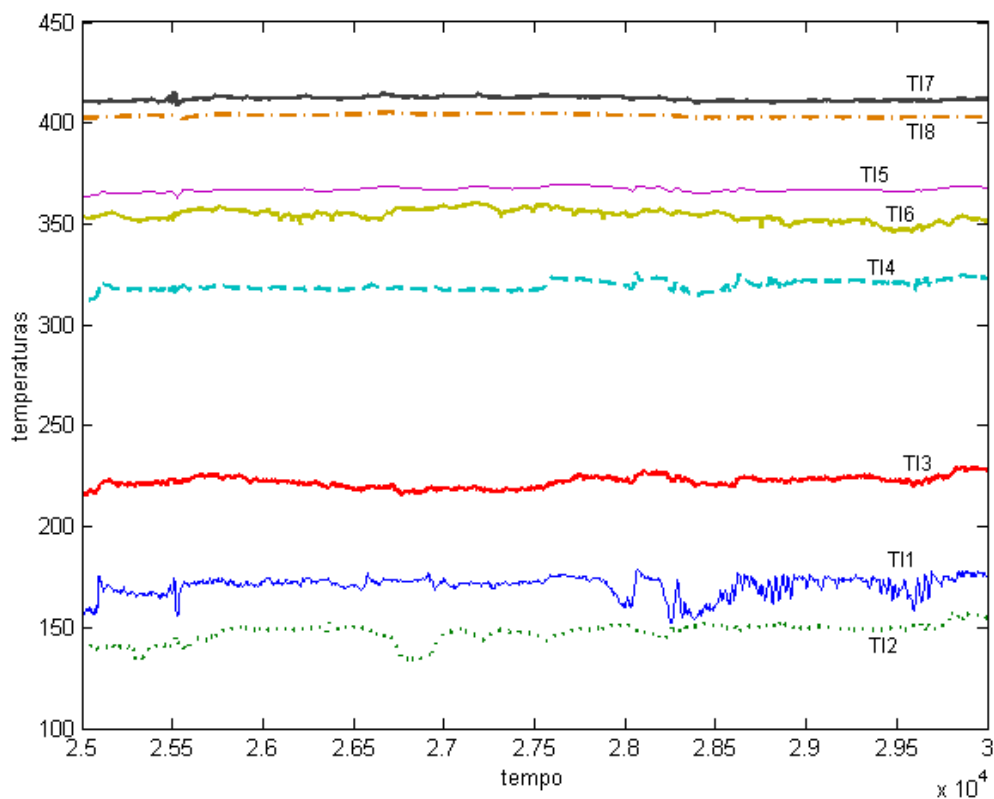


Figura 4.28: Temperaturas da coluna de destilação.

As variáveis de temperatura estão representadas na **Figura 4.28**, e as de vazão na **Figura 4.29**. As variáveis escalonadas estão apresentadas na **Figura 4.30**, deslocadas na ordenada a cada 5 unidades para facilitar a visualização. A **Figura 4.31** apresenta a variável ortogonal gerada para análise de estacionários, a qual é combinação linear das demais. A **Figura 4.32** apresenta a variável ortogonal, suas derivadas WT_1 e WT_2 e o índice B . A **Figura 4.33** mostra as variáveis originais escalonadas, para comparação com o índice de estado estacionário B da coluna.

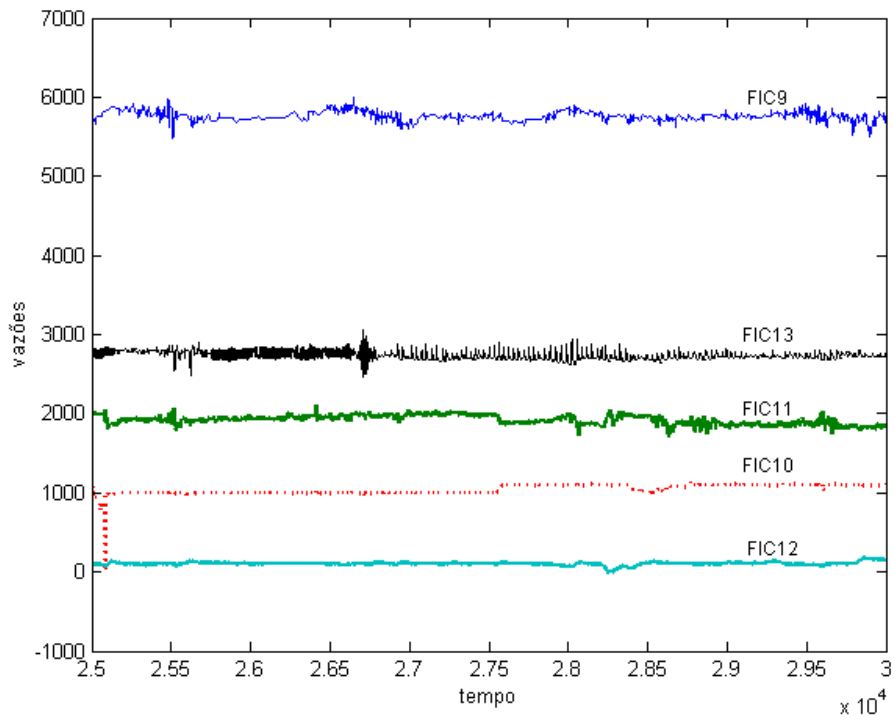


Figura 4.29: Vazões relacionadas à coluna de destilação.

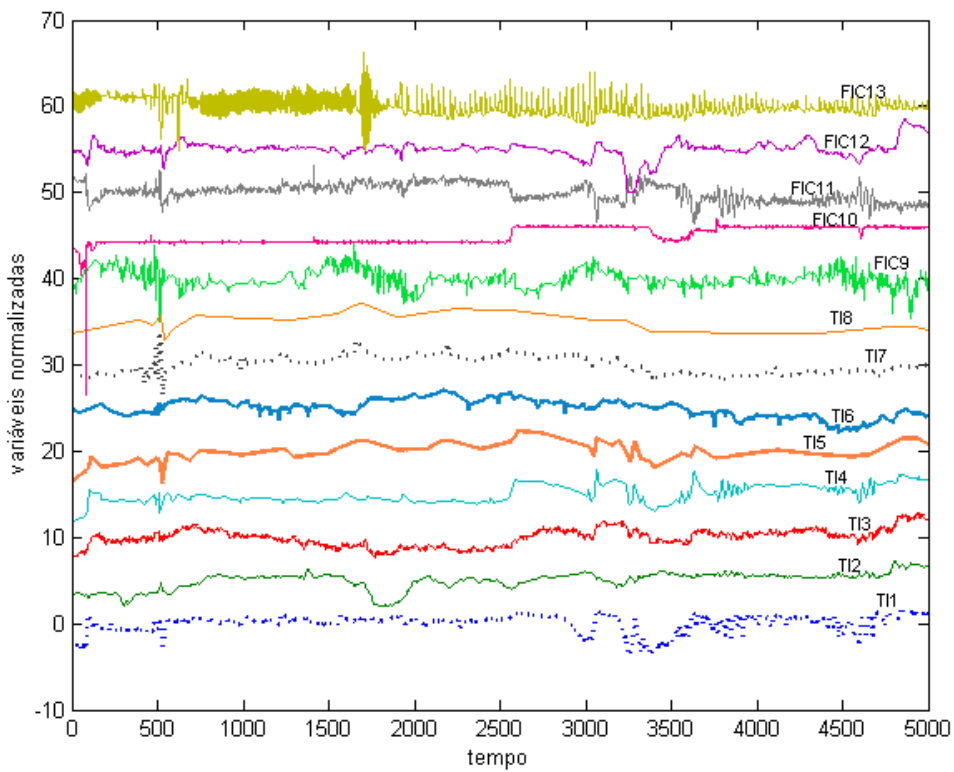


Figura 4.30: Variáveis de temperatura e de vazão escalonadas e plotadas deslocadas para facilitar a visualização.

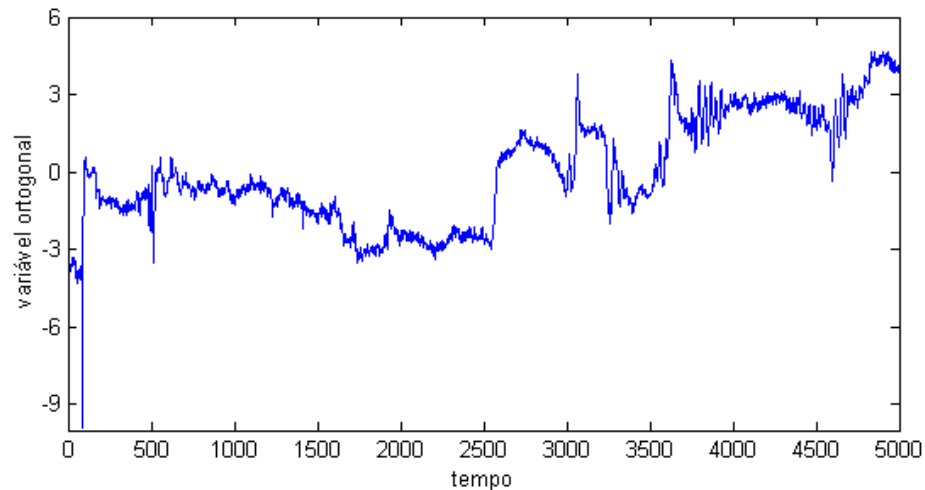


Figura 4.31: Variável ortogonal gerada a partir da combinação linear das temperaturas e vazões.

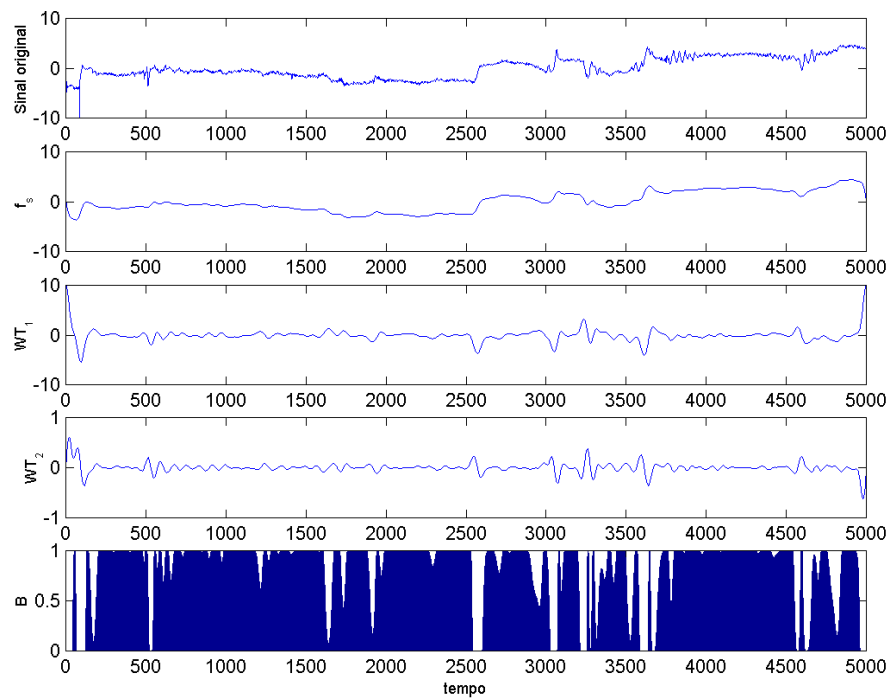


Figura 4.32: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

A **Figura 4.33** permite verificar que os índices obtidos estão de acordo com o esperado, que é uma faixa de variação de temperatura para a coluna de até $1,5^\circ\text{C}$. E o índice de estado é coerente com as alterações ocorridas na coluna.

Portanto, uma comparação entre os resultados alcançados por variáveis-chave e PCA mostra que os obtidos por PCA são mais coerentes com a real situação da coluna neste intervalo de tempo.

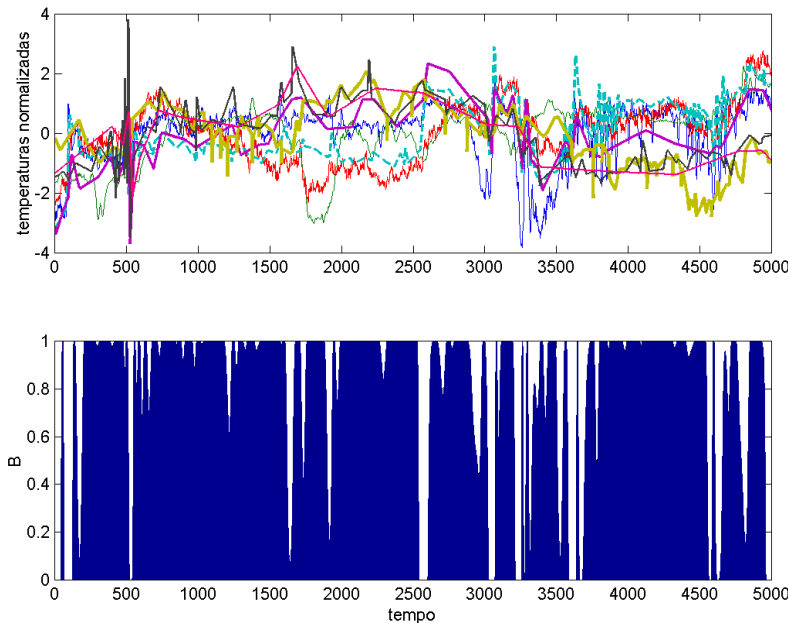


Figura 4.33: Temperaturas escalonadas e índice de estado estacionário B .

Extrapolação para outro período de tempo utilizando técnica de PCA: A mesma análise baseada em PCA foi realizada para outro período de tempo, considerando as mesmas variáveis, ou seja, as 8 temperaturas da coluna e mais as 5 vazões.

O objetivo desta análise era verificar se a variável gerada a partir da mesma combinação linear entre as variáveis poderia ser aplicada satisfatoriamente a este período de dados, e verificar se a detecção de estacionários se mantinha coerente. Também o período histórico de dados foi mantido igual, para que os valores de corte para detecção de estacionários fossem os mesmos.

As próximas figuras ilustram os resultados obtidos para o novo conjunto de dados. O que pode se observar é que a variável ortogonal gerada representa bem as variáveis de temperatura e vazão, e que os resultados dos índices de estado estão coerentes com as variações significativas que ocorreram na coluna neste intervalo de tempo considerado.

Deve-se observar que, para cada variável, apenas uma vez se calculam os valores de corte no algoritmo (com a informação de um período histórico e constante de tempo) e apenas uma vez o vetor de pesos é calculada por PCA. Estes resultados, então, podem ser utilizados para qualquer outro período de tempo dentro da mesma região de operação, diminuindo o número de cálculos e sem prejudicar a identificação dos estacionários destes outros períodos.

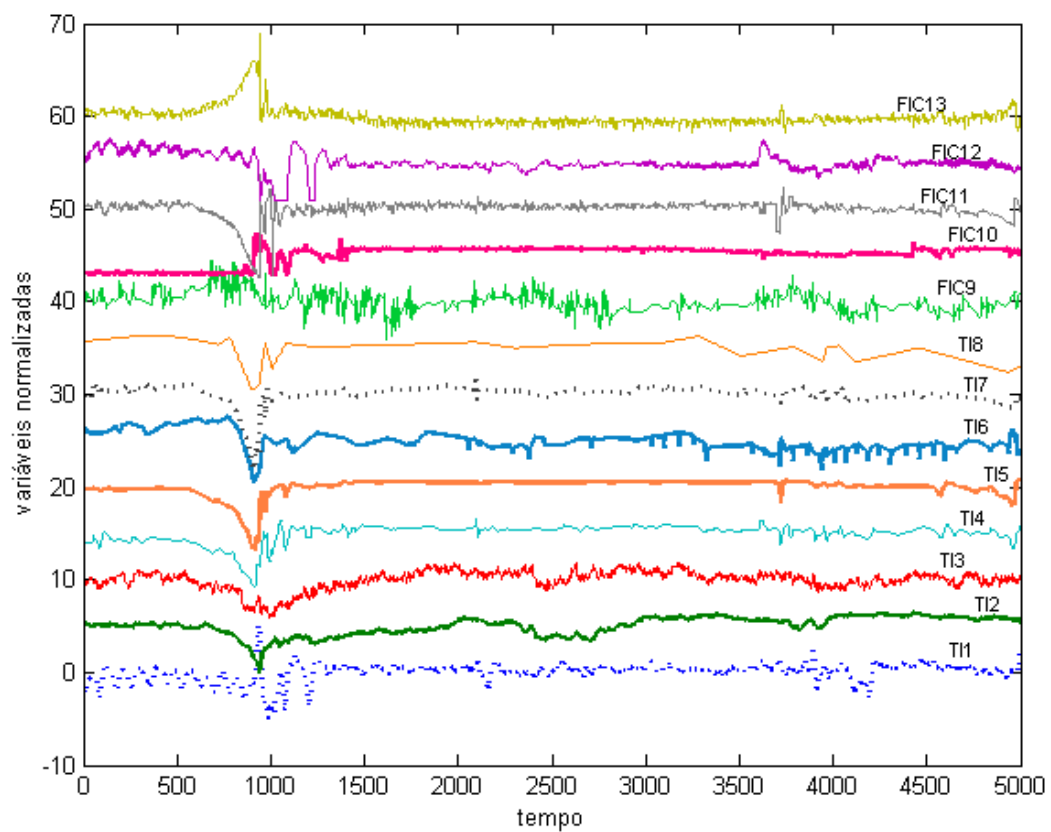


Figura 4.34: Variáveis de temperatura e de vazão escalonadas.

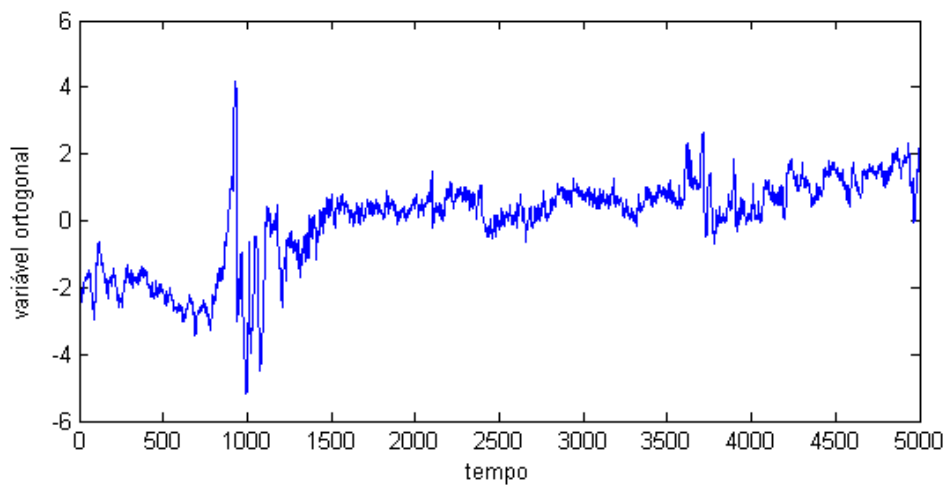


Figura 4.35: Variável ortogonal que é uma combinação linear das temperaturas e vazões.

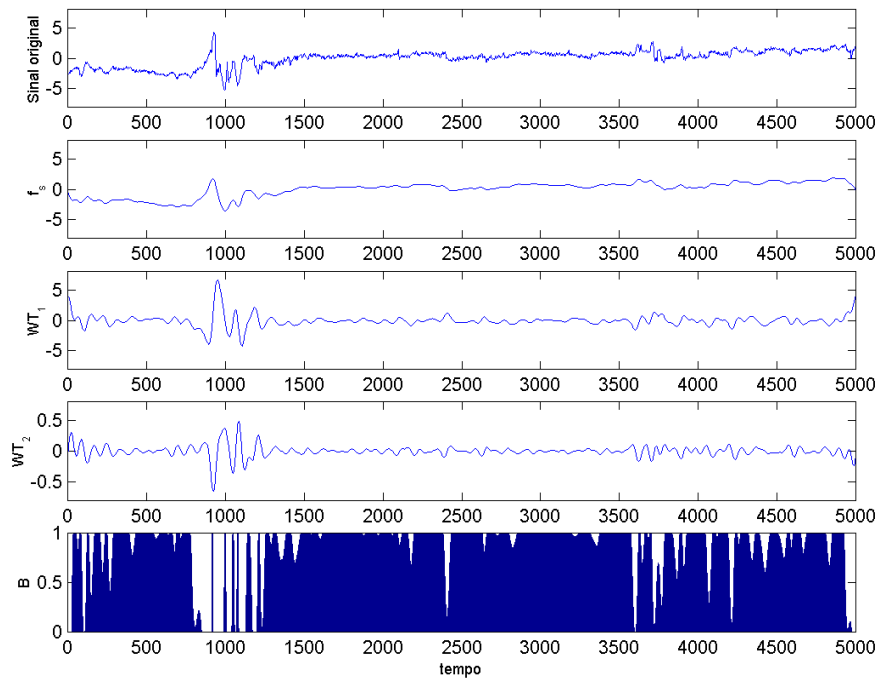


Figura 4.36: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

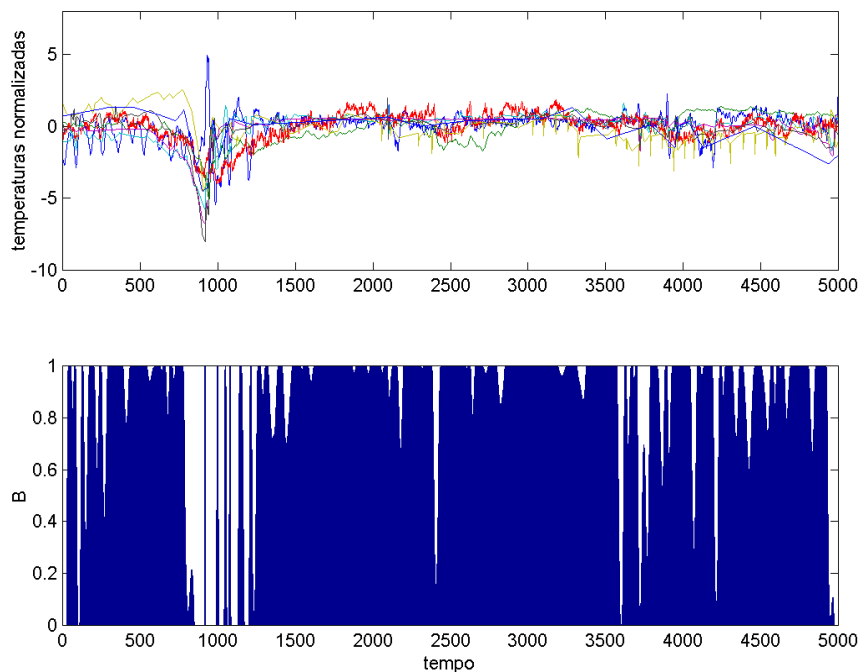


Figura 4.37: Temperaturas escalonadas e índice de estado estacionário B .

4.3.6 Unidade de Craqueamento Catalítico

Nas seções anteriores, dois exemplos de colunas de destilação foram apresentados. Para verificar os resultados de outro tipo de equipamento, estudou-se também uma unidade de craqueamento catalítico de leito fluidizado (FCC) (SECCHI *et al.*, 2001).

A unidade de FCC consiste de um processo de refino para aumentar a produção de gasolina e gás liquefeito de petróleo (GLP) a partir de um gás-óleo pesado e resíduo de conversão de frações leves. O processo de craqueamento da nafta ocorre no reator tubular (*riser*) com fluxo ascendente de vapor, em uma alta temperatura. Os produtos deste craqueamento são o gás liquefeito de petróleo (GLP), gasolina, gás efluente, óleo de ciclo leve (OCL) e coque. Este último é responsável pela desativação do catalisador, pois se deposita sobre este. Por isso, existe uma seção na unidade chamada regenerador, que é onde ocorre a queima do coque em um reator de leito fluidizado. Denomina-se de conversor o sistema *riser*/regenerador. Uma representação deste sistema é apresentada na **Figura 4.38**.

A unidade apresenta vários medidores de vazão e temperatura, principalmente, e um analisador da qualidade dos produtos. As variáveis medidas são a temperatura do *riser*, as temperaturas do regenerador, a pressão de sucção do compressor de gás úmido, o nível de catalisador do separador, a pressão diferencial do separador-regenerador e a temperatura do gás efluente do regenerador. As variáveis manipuladas são a vazão de alimentação (F_{alim}), a temperatura de alimentação pré-aquecida, as taxas de circulação de catalisador (na TCV e LCV), a vazão de ar de combustão (F_{ar}) e a taxa de reciclo de gás úmido no compressor. Os distúrbios medidos são características da alimentação e as temperaturas de alimentação (T_{alim}) e do ar (T_{ar}). Conforme SECCHI *et al.* (2001), pode-se obter um bom modelo (não-linear) para a unidade, de modo que as variáveis de saída de interesse podem ser estimadas a partir das entradas (vazões de alimentação e ar e distúrbios), conforme representação da **Figura 4.39**.

Neste caso, então, um modelo não-linear foi obtido, considerando como pontos de operação estacionários $F_{\text{ar}} = 90$ ton/h e $F_{\text{alim}} = 2750$ m³/dia. A partir deste modelo, simularam-se as condições desejadas, ou seja, as entradas foram variadas e as saídas obtidas foram combinadas linearmente pela técnica de PCA para identificação dos pontos estacionários.

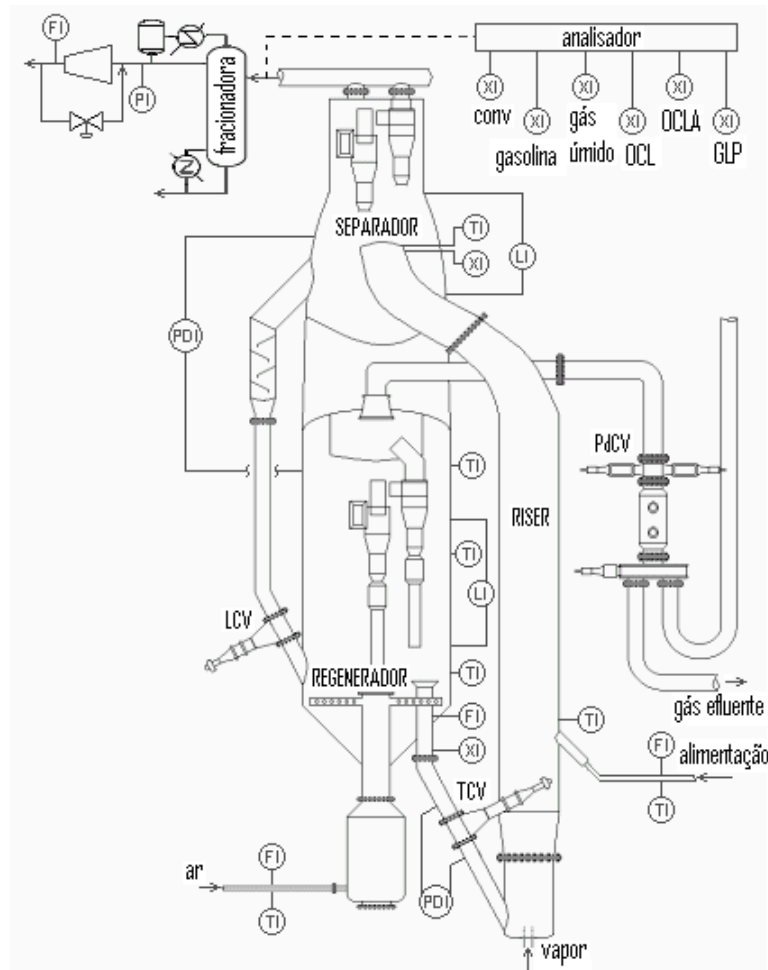


Figura 4.38: Unidade de craqueamento catalítico de leito fluidizado (FCC).

No caso das variáveis de entrada (F_{alim} , F_{ar} , T_{alim} e T_{ar}), buscou-se reproduzir o que seriam comportamentos e condições típicas de operação. Assim, as vazões de alimentação e do ar consideraram degraus positivos e negativos, componente ascendente linear e componente senoidal. Os distúrbios (T_{alim} e T_{ar}) foram considerados puramente senoidais, representando as variações de temperatura que podem ocorrer ao longo de um dia ou do ano, por exemplo. A frequência da componente senoidal nos sinais é igual, com exceção da temperatura do ar que, além desta componente, possui também uma componente senoidal de frequência menor. Nos primeiros instantes de tempo, ocorrem variações na vazão do ar (F_{ar}), depois começam as variações na vazão de alimentação (F_{alim}) e, por fim, ocorrem as variações nas temperaturas simultaneamente. Uma representação de cada variável está na **Figura 4.40**.

Como saída do modelo da unidade de FCC resultam as variáveis de interesse para análise, as quais são mostradas na **Figura 4.41** (de acordo com a **Figura 4.39**).

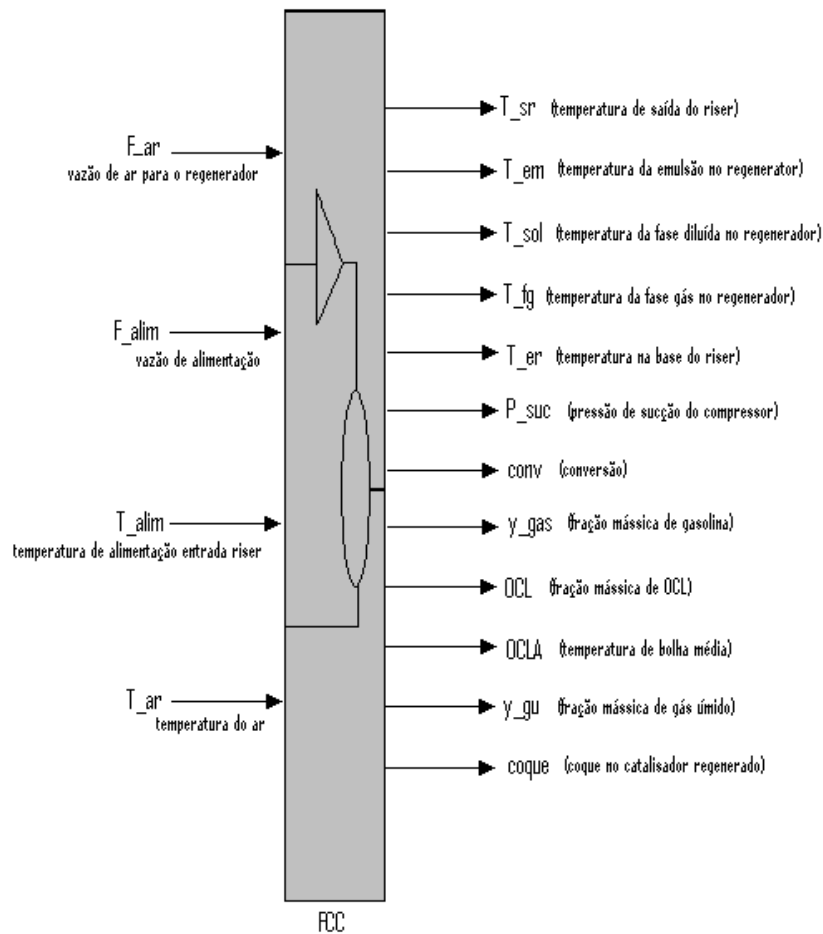


Figura 4.39: Representação das variáveis de entrada e saída utilizando um modelo não-linear para a unidade de FCC.

As variáveis mostradas na **Figura 4.41** foram normalizadas e o resultado é apresentado na **Figura 4.42**. Na etapa seguinte, estas variáveis foram linearmente combinadas pelo uso da técnica de PCA, obtendo-se uma nova variável, apresentada na **Figura 4.43**.

Como já descrito nas seções anteriores, esta nova variável foi analisada para detecção dos pontos estacionários. Para isto, considerou-se que a constante de tempo de resposta deste sistema é de aproximadamente 15 minutos, e que o período histórico é igual aos instantes de tempo de 2000 a 2600 da variável normalizada. Lembra-se que os sinais de entrada possuíam componentes senoidais, de frequência menor. Isto significa que a constante de tempo de resposta a ser estimada deve ser um valor elevado (e, conseqüentemente, maior o número de níveis de decomposição na transformada das ondaletas). O período de dados históricos selecionado contempla uma região com dois períodos senoidais, um de menor amplitude e outro de maior amplitude, ambos com frequências semelhantes. O resultado para a aplicação do algoritmo de detecção de estacionários está apresentado na **Figura 4.44**.

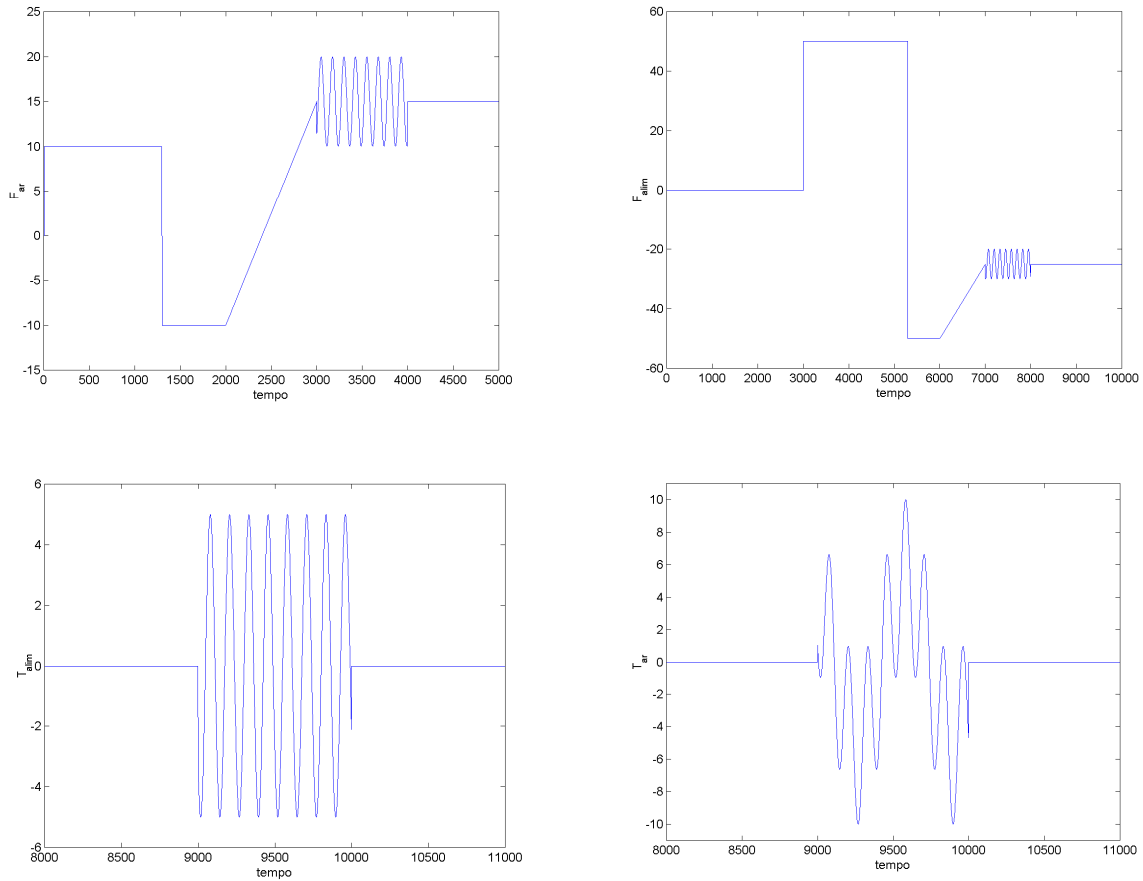


Figura 4.40: Distúrbios inseridos em cada variável de entrada do modelo.

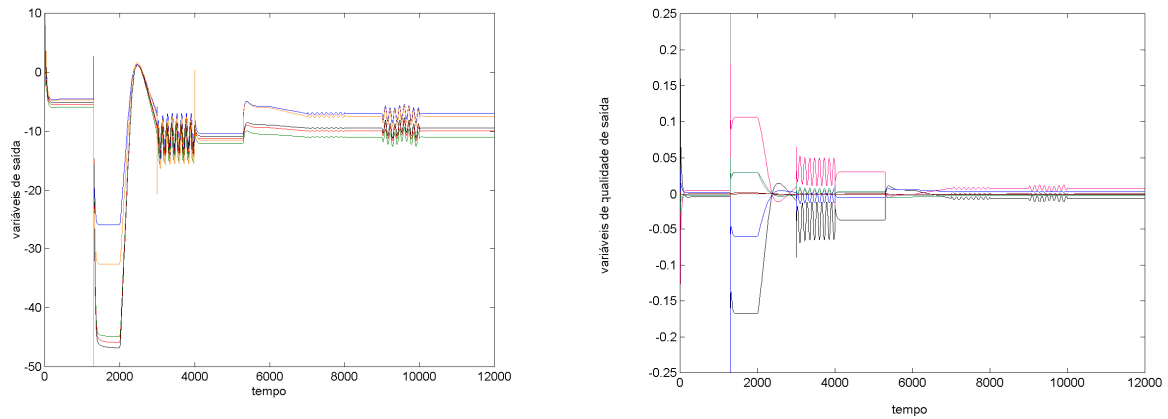


Figura 4.41: Variáveis de saída do modelo não-linear. a) temperaturas b) pressão e variáveis de qualidade dos produtos.

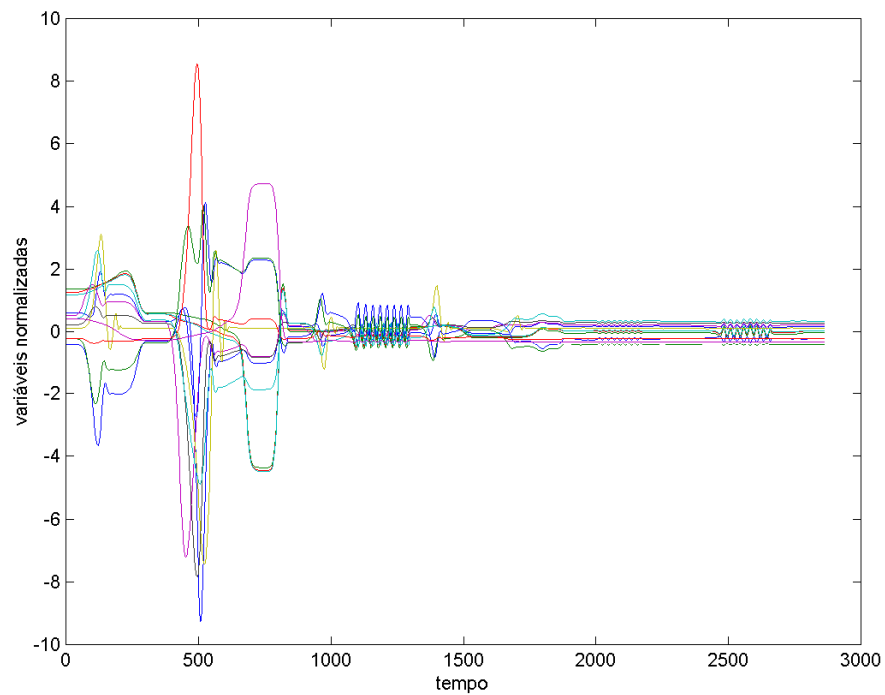


Figura 4.42: Variáveis de saída do modelo não-linear normalizadas.

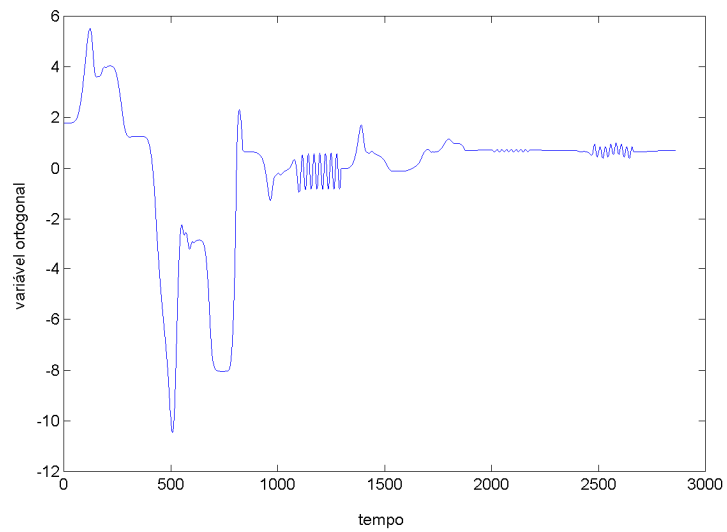


Figura 4.43: Variável ortogonal gerada pela técnica de PCA.

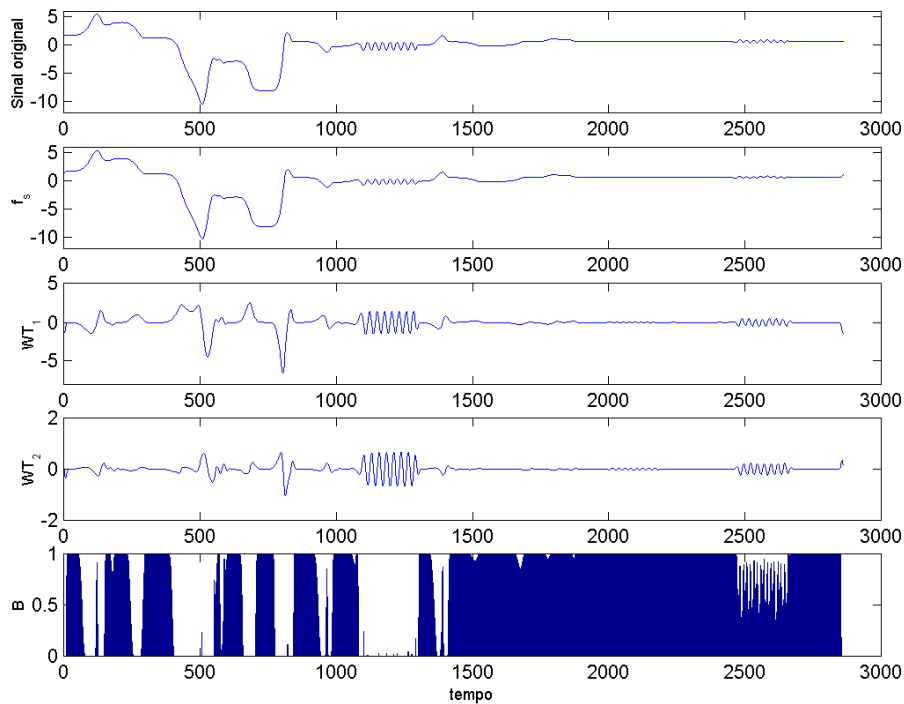


Figura 4.44: Resultados para a variável ortogonal, representada por f , f_s é a tendência extraída do sinal original, WT_1 é a transformada de primeira ordem das ondaletas, WT_2 é a transformada de segunda ordem das ondaletas, e B é o índice de estado.

A seleção do período relativo à frequência dos senos menor do que a constante de tempo indica que a ocorrência das oscilações deve ser reconhecida pelo algoritmo de detecção de estacionários. Até uma certa magnitude, componentes rápidos são considerados como eventos normais de operação, e componentes mais lentos passam a ser detectados como mudanças no processo. Como pode ser visto na **Figura 4.44**, dos três períodos com oscilações, dois são detectados fora da condição estacionária, os quais são os dois períodos de maior amplitude. O seno do período intermediário (tempo entre 2000 e 2200, aproximadamente) não é detectado por possuir pequena magnitude, menor que o valor de corte calculado pelo algoritmo, de modo que este intervalo de tempo é considerado estacionário. As demais variações que ocorrem no sinal, do tempo inicial até o instante 1000, são coerentemente detectadas pelo algoritmo como condições não-estacionárias. Estes períodos não-estacionários estão coerentes com os distúrbios inseridos nas quatro variáveis de entrada (**Figura 4.40**).

4.4 Conclusões

A metodologia de JIANG *et al.* (2003) para casos monovariáveis parece ser realmente eficiente e apresenta, além de tudo, um índice de estado estacionário, que não é 0 ou 1, mas

sim uma faixa entre 0 e 1. Isto é importante para aplicações práticas, pois permite saber o quão perto ou longe se está da condição estacionária.

A influência da constante de tempo e do período histórico foi estudada, e verificou-se que a detecção de estacionários é dependente destes parâmetros. Quanto maior a constante de tempo, mais decomposições são realizadas e mais suave se torna a tendência extraída do sinal. Conseqüentemente, a detecção de estacionários é também influenciada por esta tendência e pela magnitude dos coeficientes de detalhe de primeira ordem.

Embora a tendência calculada seja a mesma para diferentes períodos de dados históricos, o índice de estado estacionário é fortemente alterado por este parâmetro. Se o período histórico contém características de degrau e picos anormais, o índice de estado poucas vezes detectará estas mudanças no sinal de processo e, assim, a condição será considerada estacionária na maior parte do tempo. Por outro lado, se a região histórica for composta somente de ruído branco de baixa magnitude, freqüentemente a condição será considerada transiente. O que se espera, portanto, é considerar o período histórico como uma região coerente com a condição normal de operação do equipamento, apresentado variabilidade e ocorrências características. Como esta tarefa de escolher o período histórico é feita apenas uma vez para cada variável de processo, é importante que ela traduza de forma adequada os efeitos desejados na determinação de estacionários.

A metodologia baseada em PCA para detectar estacionários se mostrou bastante eficiente para os casos estudados. Ela é mais simples de ser aplicada do que a proposta por JIANG *et al.* (2003), além de apresentar resultados mais coerentes. Como não necessita de atribuição de pesos às variáveis, ela se torna mais facilmente aplicável e exige menos esforço computacional. Além disso, uma vez calculada a combinação linear entre as variáveis (vetor de pesos) para cada sistema, este cálculo não precisa mais ser refeito para uma mesma região de operação, e apenas a computação da nova variável ortogonal (que é uma combinação linear das variáveis medidas) e a detecção de estacionários são realizadas. Da mesma forma, apenas uma vez se calculam os limites de corte do algoritmo (para identificação de picos anormais e detecção de estacionários), os quais podem ser utilizados para quaisquer outros intervalos de tempo.

Assim, esta metodologia de identificação de estacionários baseada em PCA pode ser considerada como uma boa alternativa para identificação de pontos estacionários e pode ser aplicada na prática sem a exigência de grandes informações ou conhecimento sobre o processo.

Capítulo 5

Conclusões

Neste último capítulo serão ressaltados os principais resultados encontrados neste trabalho.

5.1 Considerações Finais

Os resultados dos testes realizados e sua avaliação crítica permitem algumas conclusões e sugestões sobre a obtenção de tendências de um sinal de processo. A comparação das metodologias para obtenção de tendências de sinais (filtros Butterworth, Splines e Ondaletas) indicou que *splines* ajustadas se apresentam como uma boa alternativa para a maioria dos casos, por obter boa qualidade na aproximação e por ser de fácil ajuste, não sendo indicada para casos com descontinuidades próximas. O filtro Butterworth e a versão interpolada de *splines* não apresentaram bons resultados para os exemplos analisados, não sendo então recomendados. As ondaletas apresentaram comportamento intermediário, e podem ser utilizadas dependendo da aplicação ou finalidade.

Embora as *splines* apresentaram-se como uma boa alternativa, as Ondaletas possuem uma característica importante de sua transformada decompor o sinal em diferentes escalas, o que permite analisar a tendência do sinal e seu componente de ruído em diferentes frequências. No caso do objetivo ser detectar estacionários, Ondaletas são então a ferramenta indicada, pois possuem a capacidade de separar a informação útil da informação que não traz características do processo. Para esta aplicação, a metodologia apresentada mostrou resultados muito bons. Além de um índice de estado estacionário, no caso multi-variável a técnica baseada em PCA foi eficiente em combinar as variáveis de modo a detectar pontos estacionários corretos. Os parâmetros de entrada do método são estimados uma única vez para cada variável, o que torna o processo subsequente de fácil aplicação. Portanto, a metodologia proposta neste trabalho se apresenta como uma boa técnica para detecção de pontos estacionários.

5.2 Sugestões para Trabalhos Futuros

Alguns aspectos relacionados a este trabalho podem ser desenvolvidos ou aprimorados no futuro.

Como sugestões de alterações e melhorias, propõe-se testar o algoritmo das ondaletas utilizando outro tipo de função base para a ondaleta, como a chapéu mexicano, por exemplo. Sabe-se que determinadas ondaletas não são apropriadas para determinadas aplicações, como a ondaleta de Haar para descrever os casos de funções suaves. Entretanto, embora tendam a apresentar resultados semelhantes, uma ondaleta mais adequada para a aplicação de interesse pode resultar em uma convergência mais rápida. A utilização de uma ondaleta mais similar possível com a função que se tenta aproximar implica em menos escalas necessárias para obter uma descrição satisfatória (ALSBERG *et al.*, 1997).

No caso da aplicação do algoritmo baseado em ondaletas para a identificação de estados estacionários para sistemas multi-variáveis, sugere-se testar a combinação das variáveis por outras técnicas, como NPCA ou PLS. A utilização de combinação por PCA mostrou-se satisfatória. Entretanto, a combinação de variáveis de variabilidades muito diferentes pode acarretar alguns erros na determinação dos pesos de cada uma. Além disso, para sistemas com muitas variáveis é recomendável que sejam eliminadas da combinação do PCA as variáveis que são pouco significativas, ou seja, aquelas que recebem baixo peso. Ainda nos casos multi-variáveis, uma outra alternativa é que cada variável possa ter sua constante de tempo específica (e não um valor médio, como é o caso), o que faz com que cada variável tenha um nível de decomposição diferente.

A utilização de ondaletas e derivadas encontra um espaço importante na compressão de dados. O monitoramento computadorizado moderno e os sistemas de controle em plantas resultam em uma grande quantidade de dados sendo coletados. Estes possuem informação de valor sobre o processo e podem levar a tomadas de decisão efetivas. Entretanto, devem ser armazenados sem distorção das tendências, através de uma técnica de compressão de dados que seja eficiente. Armazenamento e aquisição de dados são processos dispendiosos. Quanto maior é a quantidade de informação com que se lida, mais cara ela é. Ao invés disto, grande parte de sinais digitais não são armazenados na sua forma mais compacta, mas sim na maneira que os torna mais fáceis de serem utilizados ou manuseados, como os formatos de texto ASCII e códigos binários (SMITH, 2004). Todavia, a compressão afeta os resultados de análise de dados de processo, como mostraram THORNHILL *et al.* (2004) e, portanto, deve-se cuidar na escolha de um bom método de compressão. Várias técnicas de compressão têm sido utilizadas, como algumas mais tradicionais baseadas em interpolação, como BoxCar e Backward-Slope (HALE e SELLARS, 1981) e uma combinação das duas (AspenTech, 2001). Estes métodos, entretanto, resultam em baixa compressão. Alguns outros métodos têm sido propostos, como o swimming door (BRISTOL, 1990), utilizado pelo historiador PI da OSI, uma técnica *on-line* de tendências lineares nos segmentos PLOT (OSI Software, 2002), método de interpolação em linhas retas SLIM, uma técnica baseada em B-splines utilizando representação multi-escalas (VEDAM e VENKATASUBRAMANIAN, 1998).

A maioria dos métodos aplicados industrialmente utiliza apenas a informação temporal como critério para compactação, o que obviamente não é ótimo. Neste sentido, métodos de compressão baseados em ondaletas têm sido apontados como boas alternativas, e é uma sugestão de trabalho futuro testar a eficiência destes métodos, como a técnica *on-line* de compressão baseada na ondaleta de Haar proposta por MISRA *et al.* (2000). As metodologias para compressão utilizando ondaletas basicamente consistem inicialmente da aplicação da transformada das ondaletas ao sinal. Após, os valores dos coeficientes das ondaletas que são insignificantes, ou seja, que são menores do que um valor de corte, são considerados iguais a zero. Então apenas os valores significativos são considerados, representando um conjunto de dados bem menor do que o original. Ao final, realiza-se a transformada inversa das ondaletas com os dados da etapa anterior, no processo de descompressão, o qual deve produzir uma aproximação do sinal original sem perda de qualidade.

Apêndice A

Considerações Matemáticas e Conceitos

A.1 Propriedades Matemáticas

Escalonamento: $f_a(x)$ denota a dilatação da função $f(x)$ por um fator de escala a .

$$f_a(x) = \frac{1}{a} f\left(\frac{x}{a}\right) \quad (\text{A.1})$$

Convolução: a convolução de duas funções $f(t) \in L^2(\mathbf{R})$ ¹ e $g(t) \in L^2(\mathbf{R})$ é definida como

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau \quad (\text{A.2})$$

Função Delta e Resposta ao Impulso: Todo sinal discreto pode ser interpretado como um conjunto de impulsos unitários δ , chamados de Deltas de Dirac. Impulso é um sinal composto de zeros, exceto por um único ponto não-nulo (SMITH, 2004). Assim, qualquer impulso pode ser representado como uma função delta deslocada e escalonada. Quando um sinal deste tipo é considerado, um procedimento descrito como convolução é utilizado (veja seção seguinte), de modo que fazer a convolução de um sinal com uma função delta resulta no mesmo sinal.

¹ $L^p(\mathbf{R})$ é o espaço de medida de Hilbert das funções, de modo que

$$\int_{-\infty}^{\infty} |f(x)|^p dx < \infty$$

Portanto, como o nome indica, a resposta ao impulso é o sinal que o sistema exhibe quando uma função delta (impulso unitário) é a entrada. Se dois sistemas são diferentes, eles terão diferentes respostas ao impulso. Como sinais de entrada e saída são usualmente representados por $x[n]$ e $y[n]$, a resposta ao impulso é normalmente definida por $h[n]$.

A resposta ao impulso recebe nomes diferentes, conforme a aplicação. Se o sistema em consideração é um filtro, a resposta ao impulso pode ser chamada de *filter kernel*, *convolution kernel* ou simplesmente *kernel*.

A.2 Convolução

Do ponto de vista matemático, a convolução é um operador matemático que toma duas funções f e g e produz uma terceira função que representa qual é a sobreposição existente entre f e uma versão revertida e transladada de g . Portanto, a convolução relaciona três sinais de interesse: sinal de entrada, sinal de saída e resposta ao impulso.

Se F é um operador linear invariante no tempo, atuando em seqüências temporais, ele tem a forma de uma convolução (BOGGESS e NARCOWICH, 2001). Assim, existe uma seqüência f tal que

$$F(x) = f * x \quad (\text{A.3})$$

onde $*$ denota a operação de convolução entre f e x . Inversamente, se a Eq. (A.3) acima é válida, então F é um operador linear discreto, invariante no tempo.

Tais operadores de convolução são chamados de “filtros discretos” (BOGGESS e NARCOWICH, 2001). A seqüência f , que satisfaz $F(e^0) = f$, portanto, é a resposta a um impulso no tempo discreto 0, e é chamada de “resposta ao impulso”. Se f possui um número infinito de termos não-nulos, é chamado de resposta infinita ao impulso (*infinite impulse response* - IIR). Por outro lado, se possui apenas um número finito de termos não-nulos é uma resposta finita ao impulso (*finite impulse response* - FIR).

A convolução é definida, conforme Eq. (A.2), como a integral do produto de uma das funções com uma cópia revertida e deslocada da outra. A função resultante depende do valor deste deslocamento τ . Para se adquirir uma boa visão intuitiva da convolução, é preciso entender que diversas cópias transladadas e tomadas de trás-para-frente de uma das funções são ponderadas pelo valor da outra função, e somadas, produzem o resultado. Na Eq. (A.2), f seria a função de ponderação, enquanto que cópias revertidas de g estariam sendo deslocadas e somadas ao resultado. Entretanto, a convolução não depende da ordem das funções, ou seja, a função de ponderação pode ser tanto f quanto g , produzindo o mesmo resultado.

Considere um sistema linear discreto no tempo e uma entrada arbitrária $x[n]$ para este sistema. Pode-se expressar $x[n]$ como uma combinação linear de impulsos unitários deslocados, segundo a Eq. (A.4).

$$x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n-k] \quad (\text{A.4})$$

Utilizando a propriedade de superposição de sistemas lineares, a saída $y[n]$ pode ser expressa como uma combinação linear das respostas do sistema a impulsos unitários deslocados (OPPENHEIM e WILLSKY, 1997). Especificamente, se $h_k[n]$ é a resposta de um sistema linear a um impulso unitário deslocado $\delta[n-k]$, então a resposta do sistema a uma entrada arbitrária é (KRAUSS *et. all*, 1999):

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] h_k[n] \quad (\text{A.5})$$

A acima Eq. (A.5) expressa, portanto, a resposta de um sistema linear invariante no tempo (LTI) a uma entrada arbitrária em termos da sua resposta a um impulso unitário.

Na área de processamento de sinais digitais, a convolução pode ser usada para gerar a saída de um filtro FIR pela convolução do sinal de entrada com a resposta ao impulso. A convolução é, então, uma medida de similaridade entre dois sinais, normalmente utilizada para encontrar características em um sinal desconhecido pela comparação com um sinal conhecido, e pode ser utilizada como um algoritmo genérico e bem eficiente de suavização e diferenciação.

Ainda pode-se dizer que a convolução é uma operação realizada com dois sinais que envolve a multiplicar um sinal por uma versão atrasada de outro sinal, integrar ou fazer a média do produto, e repetir o processo para diferentes atrasos. Desta forma, a convolução é um processo útil porque descreve acuradamente alguns efeitos que ocorrem largamente em medidas científicas, como a influência de um filtro passa-baixa em um sinal elétrico ou de um espectrofotômetro na forma de um espectro (O'HAVER, 2005).

Na prática, o cálculo é normalmente realizado pela multiplicação de dois sinais no domínio de Fourier. Primeiro, a transformada de Fourier de cada sinal é obtida. Então, as duas transformadas são multiplicadas de acordo com as regras de multiplicação complexa e o resultado é então inversamente transformado por Fourier. Esta é uma maneira de se obter uma convolução mais rápida que o algoritmo de deslocamento-e-multiplicação quando o número de pontos do sinal é grande, entretanto ela introduz um certo atraso.

Implementando a convolução utilizando FFT se obtém uma convolução cíclica (MADISSETTI e WILLIAMS, 1998). De maneira a utilizar a FFT, zeros são adicionados no final do sinal ou da sequência do filtro até que ambos possuam o mesmo comprimento. Se a FFT do sinal $x[n]$ é multiplicada termo a termo pela FFT do filtro $h[n]$, o resultado é a FFT da saída $y[n]$. Entretanto, o comprimento de $y[n]$ obtido pela inversa da FFT é igual ao comprimento da entrada. O *aliasing* (quando o sinal contínuo é discretamente amostrado e não pode ser reconstruído devido à perda de informação) causado pela convolução cíclica pode ser superado pela adição de zeros tanto em $x[n]$ quanto em $y[n]$ até que seus

comprimentos sejam $N+L-1$ e então pelo uso da FFT. A parte do sinal de saída que é conhecida/reconhecida (*aliased*) é zero, e o resultado da convolução cíclica é exatamente o mesmo da convolução não-cíclica. O preço é realizar FFT de seqüências com comprimentos maiores, mas seqüências as quais aproximadamente metade dos números são zeros.

No caso das ondaletas, a CWT é uma convolução contínua da função ondaleta com o sinal ao longo das escalas contínuas (ou seja, convolução circular de $f(t)$ com $W_{a,b}(t)$). Lembra-se que a CWT é uma análise muito redundante, pois mapeia uma função 1D em uma função 2D. A discretização da WT permite que esta seja calculada como uma série de convoluções discretas no tempo em intervalos de escala discretos.

A.1 Invariância no Tempo

Um sistema é invariante no tempo se um deslocamento no tempo no sinal de entrada causa um deslocamento no tempo no sinal de saída. Precisamente, se $y[n]$ é a saída em um tempo discreto para uma entrada $x[n]$, então um sistema é invariante no tempo se $y[n-k]$ é a saída quando $x[n-k]$ é aplicado (OPPENHEIM e WILLSKY, 1997).

A Eq. (2.15), que representa a transformada diádica, não é invariante no tempo (MERTINS, 1999) porque um atraso no sinal de entrada $f(x-l)$ leva a

$$\begin{aligned} C(2^j(k-2^{-j}l), 2^j) &= 2^{-j/2} \sum_x f(x-l) \psi(2^{-j}x-k) \\ &= 2^{-j/2} \sum_i f(i) \psi(2^{-j}i - [k - 2^{-j}l]) \end{aligned} \quad (\text{A.6})$$

Somente se l é múltiplo de 2^j obtém-se versões deslocadas dos mesmos coeficientes das ondaletas. Entretanto, para muitas aplicações como reconhecimento de padrões e estimação do movimento no domínio das ondaletas é desejável que se atinja a invariância de translação. Este problema pode ser resolvido computando todos os valores

$$C(k, 2^j) = 2^{-j/2} \sum_x f(x) \psi(2^{-j}(x-k)) \quad (\text{A.7})$$

Geralmente este cálculo é computacionalmente dispendioso, mas quando se usa o *Algorithm à Trous* o cálculo torna-se tão eficiente quanto o da DWT.

A.2 Efeitos de Borda

Quando uma convolução é realizada em sinais de comprimento finito, normalmente acontecem distorções nas bordas. Para se adequar a este problema, as bordas devem ser tratadas de uma maneira diferenciada com relação as outras partes do sinal. Maneiras simples

de estender o sinal nas bordas são técnicas preferíveis, pois envolvem computação de uns poucos coeficientes extras em cada estágio do processo de decomposição de modo a atingir a reconstrução perfeita. Salienta-se que a extensão é necessária em cada estágio do processo de decomposição (MISITI *et al.*, 2006).

Estes tratamentos para as bordas do sinal assumem algum tipo de comportamento para além de suas extremidades. Se estes comportamentos estiverem incorretos, então mudanças bruscas podem ocorrer nos coeficientes das ondaletas e pontos espúrios podem aparecer no sinal reconstruído (BRADLEY e WILSON, 2004).

Algumas técnicas de extensão de borda (MISITI *et al.*, 2006) comumente utilizadas são *zero padding*, *periodic padding*, *smooth padding* e *even periodic*. Estas técnicas estão brevemente descritas abaixo.

- *Zero padding* assume que o sinal é zero fora do suporte original (preenchimento com zeros) e isso claramente implica em baixo esforço computacional. A desvantagem é que descontinuidades são artificialmente criadas nas bordas.
- *Periodic padding* (ou *odd periodic*) é uma extensão periódica do sinal, pois insere cópias do sinal no começo e no final do sinal original para produzir uma série de sinais um depois do outro. A desvantagem é cria descontinuidades artificiais nas bordas. Por outro lado, produz o menor comprimento para a decomposição das ondaletas.
- *Symmetrization* (ou *boundary value replication* ou *even periodic*) é reflexão simétrica dos valores das bordas, pela adição de cópias reversas do sinal no começo e no final do sinal original para produzir uma série de sinais que são espelhos uns dos outros. Possui a desvantagem de criar descontinuidades artificiais nas bordas da derivada de primeira ordem.
- *Smooth padding* de ordem 1 é a simples extrapolação da derivada de primeira ordem, usando ajuste de extensão linear aos dois primeiros e dois últimos valores. É um método bom para sinais suaves. Também há o método de *smooth padding* de ordem 0, que é extrapolação de um valor constante, o que significa repetir o primeiro valor na esquerda e o último valor na direita.

Para uma ondaleta ortogonal e um sinal definido em um intervalo de tempo infinito, a DWT teórica possui as seguintes propriedades (MISITI *et al.*, 2006): preservação da norma ($\|X\|^2 = \|cA\|^2 + \|cD\|^2$), ortogonalidade ($\|X\|^2 = \|A\|^2 + \|D\|^2$) e reconstrução perfeita ($X = A + D$). Como a DWT é aplicada a sinais definidos em um intervalo de tempo finito, um método de extensão é necessário para a decomposição, e truncamento é necessário para a reconstrução.

Para assegurar a reconstrução perfeita para um sinal de comprimento arbitrário, se a ondaleta e o modo de extensão são arbitrários, a preservação da norma e a ortogonalidade

podem ser perdidas. Se estas propriedades são preservadas para um sinal estendido de comprimento normalmente maior do que o do sinal original, então somente a propriedade de reconstrução perfeita é sempre preservada. Por outro lado, se a DWT é realizada utilizando o modo de extensão periódica e se o sinal possui comprimento divisível por 2^J (onde J é o nível máximo de decomposição), então todas as propriedades são verdadeiras.

Apêndice B

Algoritmo das Ondaletas

Nesta seção, descreve-se o algoritmo que faz uso das ondaletas para os cálculos de detecção de estados estacionários e extração da tendência de um sinal. Os algoritmos foram implementados no MATLAB®.

Os coeficientes dos filtros passa-baixa e passa-alta foram extraídos do artigo de MALLAT e ZHONG (1992) e estão apresentados abaixo.

```
H = [ 0.125 0.375 0.375 0.125 ];  
G = [ 0 -2.0 2.0 0 ];  
K = [ 0.0078125 0.054685 0.171875 -0.171875 -0.054685 -0.0078125 ];  
Htil = conj( H );
```

Cabe lembrar que estes coeficientes são para a primeira decomposição, que resulta no nível $j = 1$. Para a decomposição do nível 1 tem-se

```
H = [ 0.1250 0 0.3750 0 0.3750 0 0.1250 ]
```

E assim, de modo genérico, são inseridos $(2^j - 1)$ zeros entre cada coeficiente, de acordo com o nível j .

B.1 Decomposição do Sinal (Transformada das Ondaletas)

A função que calcula o sinal de saída de detalhe w_{t1} tem como parâmetros de entrada o vetor de dados, os coeficientes do filtro passa-alta G , o comprimento do vetor de dados l_x e do filtro l_g e um fator $lambda$ de normalização de energia, conforme apresentado na seção 2.4.5. O algoritmo primeiramente estende o vetor de dados $data$ usando a técnica de extensão periódica, para evitar efeitos de borda, por $l_g - 1$ pontos. A convolução é então realizada entre

o vetor de dados estendido y e o filtro G , resultando nos coeficientes de detalhe $wt1d$. Como o vetor resultante desta convolução é maior do que o vetor de dados do sinal, mantém-se apenas a parte central do vetor $wt1d$, com tamanho do vetor dos dados originais lx .

```
function wt1 = wtd( data,G,lx,lg,lambda )
    y = wextend( 'ld','per',data,lg-1 );
    wt1d = wconv( 'ld', y, G ) / lambda;
    wt1 = wkeep( wt1d,lx );
```

De forma semelhante, os coeficientes de aproximação são calculados, apenas alterando para filtro passa-baixa H .

```
function sf = wta( data,H,lx,lh )
    ya = wextend( 'ld','per',data,lh-1 );
    sfa = wconv( 'ld', ya, H );
    sf = wkeep( sfa,lx );
```

É interessante lembrar que, deste modo, resultam vetores $wt1$ e sf de mesmo tamanho.

O resultado desta etapa é mostrado na **Figura B.1**, para um sinal analisado a 3 níveis de decomposição.

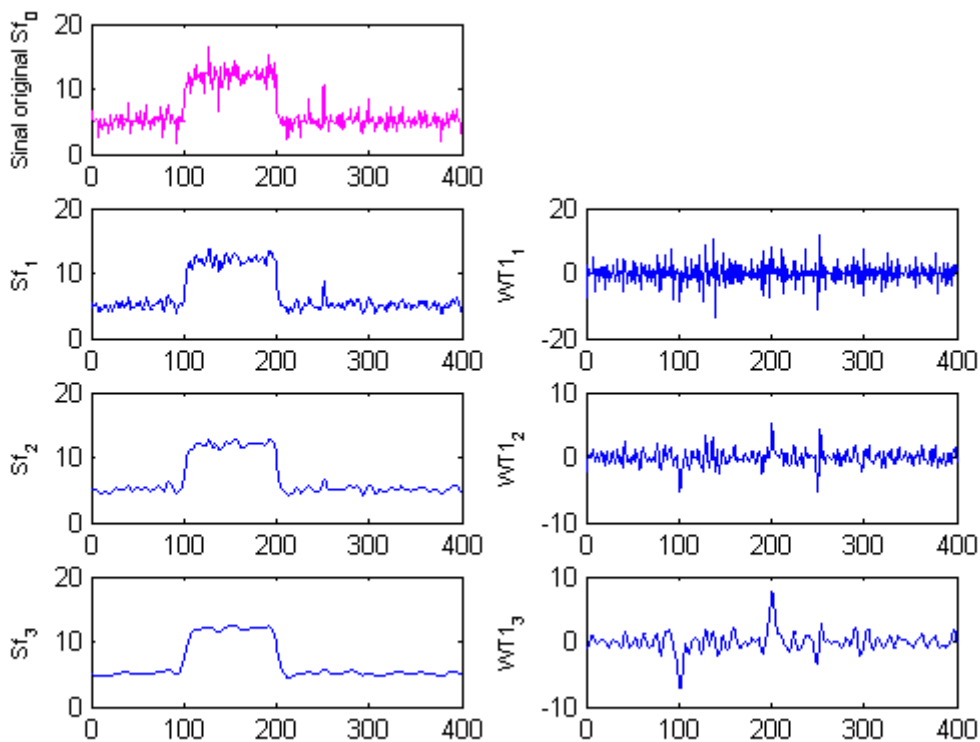


Figura B.1: Decomposição do sinal em aproximações e detalhes.

B.2 Remoção de Ruído Branco

O ruído é eliminado do sinal pela aplicação de valores de corte (*soft thresholding*), o que significa que os coeficientes de aproximação e detalhe que são abaixo de um valor estabelecido são zerados, permanecendo apenas os que excedem tal valor, conforme Eq. (3.1) apresentada na seção 3.2. O valor de corte varia conforme o nível de decomposição. O resultado pode ser visto na **Figura B.2**.

No caso do primeiro nível, dominado pelo ruído, o valor de corte foi estabelecido como a média do valor absoluto dos coeficientes de detalhe de nível 1 resultantes da decomposição de um período histórico considerado. Para os outros níveis, o valor de corte é determinado segundo a Eq. (3.2) também apresentada na seção 3.2.

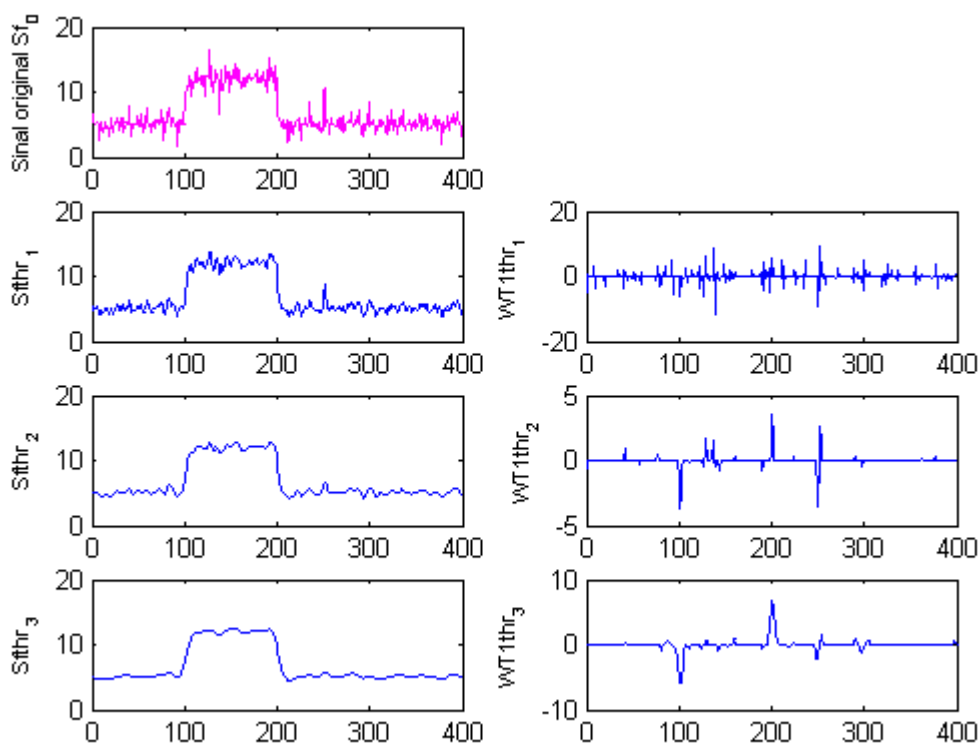


Figura B.2: Sinal original e suas aproximações e detalhes, após aplicação de valores de corte (*soft-thresholding*).

B.3 Remoção de Picos Anormais

Em um sinal de processo, pode haver alguns picos que sejam extremamente desviados da média (*outliers*), como pulsos, por exemplo. Estes picos, aqui denominados de picos anormais, devem ser eliminados do sinal para análise, pois podem trazer falsas informações, não coerentes com a realidade. Por exemplo, se ocorreu uma queda instantânea de voltagem, o medidor registrará esta queda que, se não for eliminada dos dados de análise, pode ser

interpretada com uma região de instabilidade. Isto não é real, pois o processo pode continuar estacionário, visto que este pico é um pulso de curta duração de tempo, quando comparado a uma perturbação normal, característica do processo.

Por isso, estes picos anormais são detectados, em todos os níveis, com exceção do primeiro, pela ocorrência de um par de coeficientes de detalhe com sinal contrário p_1 e p_2 , de magnitude superior a um valor de corte T_1 e com uma duração que é menor que um intervalo de tempo t_p (estimado como $1/5$ da constante de tempo), conforme mostra a **Figura B.3**. A duração do pico $[t_a, t_b]$ é correspondente ao primeiro coeficiente de detalhe imediatamente antes de p_1 e pelo último ponto imediatamente após p_2 que ficam dentro da faixa estabelecida pelos valores de corte T_1 e T_2 , conforme indica a Eq. (**Erro! Fonte de referência não encontrada.**) e também representado na **Figura B.3**.

$$\begin{aligned} |WT(t_a)| \text{ e } |WT(t_b)| &\geq T_2 \\ |WT(t_a - 1)| \text{ e } |WT(t_b + 1)| &< T_2 \end{aligned} \quad (\text{B.1})$$

Os valores de T_1 e T_2 são obtidos a partir de dados históricos. T_2 é a média do valor absoluto dos coeficientes de detalhe na escala característica obtidos pela transformada das ondaletas dos dados históricos. T_1 é calculado como 3 vezes o desvio padrão destes coeficientes de detalhe, e corrigido por um fator de variabilidade λ_1 , aqui considerado igual a 1.

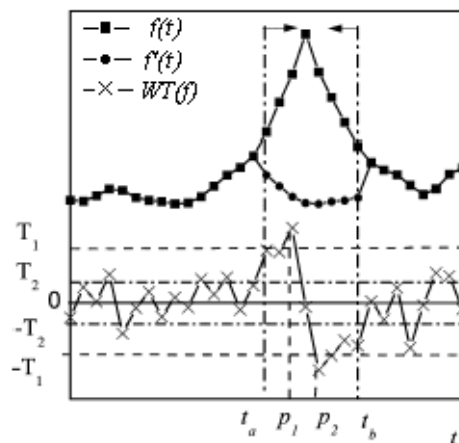


Figura B.3: Tratamento de pico anormal: $f(t)$ são os dados originais, $f'(t)$ são os dados tratados por extensão simétrica, $WT(f)$ são os coeficientes de detalhe, T_1 é o valor de corte para identificação de pico anormal, T_2 é o valor de corte para cálculo da duração de pico anormal, t_a é o instante de tempo de início do pico anormal, t_b é o tempo final da duração do pico anormal, p_1 e p_2 são os pontos de sinais contrários nos coeficiente de detalhe que identificam um pico anormal.

Para remover ou tratar picos anormais, uma das técnicas mais utilizadas é a extensão simétrica periódica (JIANG *et al.*, 2003). A técnica consiste em reproduzir simetricamente, em torno do ponto t_a-1 , os últimos pontos anteriores ao início do pico anormal. Se, então, um pico anormal tem duração n (intervalo $[t_a, t_b]$), o pico é removido tratando seus coeficientes de detalhe e aproximação segundo

$$d_{t_a+j} = d_{t_a-1-j} \text{ e } a_{t_a+j} = a_{t_a-1-j} \quad (\text{B.2})$$

para $1 \leq j \leq n$.

A **Figura B.4** apresenta um exemplo de sinal com picos anormais, em 250 e em 720, os quais foram removidos pelo tratamento de extensão simétrica.

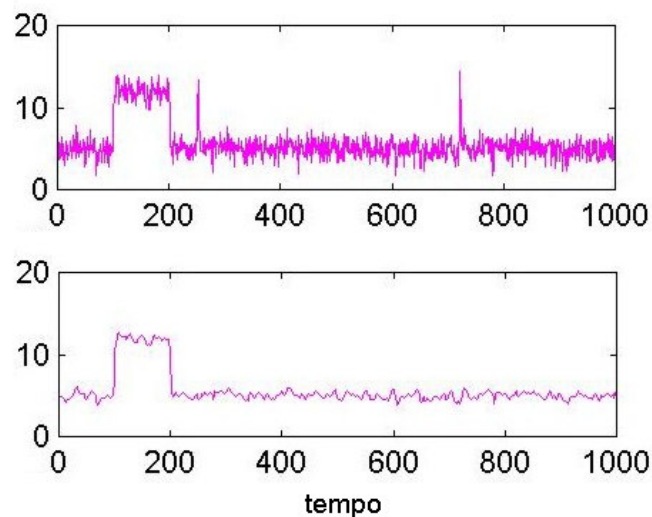


Figura B.4: Sinal com picos anormais e o sinal após tratamento com técnica de extensão simétrica e remoção de ruído.

B.4 Reconstrução (Transformada Inversa das Ondaletas)

O processo de síntese ou reconstrução é realizado da escala característica J até a primeira escala ($j = 1$), considerando os coeficientes modificados, ou seja, após a aplicação de valores de corte e eliminação de picos anormais. A soma do sinal de aproximação com o de detalhe do nível j resulta na aproximação do nível $j-1$. Isto significa fazer a convolução das aproximações e detalhes do nível j pelos filtros passa-baixa e passa-alta, respectivamente, do nível $j-1$.

```
fr1 = lambda * wta( cd,ck,lcd,lck );
fr2 = wta( ca,chttil,lca,lchttil );
f = fr1 + fr2;
```

O vetor $fr1$ representa os detalhes do nível j , onde a função w_{ta} é a mesma exemplificada anteriormente, porém agora o vetor de entrada é o vetor dos coeficientes de detalhe modificados cd , cujo comprimento é l_{cd} , e o filtro passa-alta ck , de comprimento l_{ck} .

O vetor $fr2$ representa, por sua vez, as aproximações do nível j , sendo a convolução realizada com os coeficientes de aproximação modificados ca , de comprimento l_{ca} , e o filtro passa-baixa cht_{il} , de comprimento $l_{cht_{il}}$.

Assim, o sinal f é o sinal reconstruído no nível $j-1$, a partir da soma das aproximações e dos detalhes.

A **Figura B.5** mostra um exemplo de sinais reconstruídos. Para o exemplo anterior de 3 níveis de decomposição, pode-se reconstruir o sinal aos níveis 2, 1 e 0, que representa o sinal reconstruído integralmente (o que seria o sinal original).

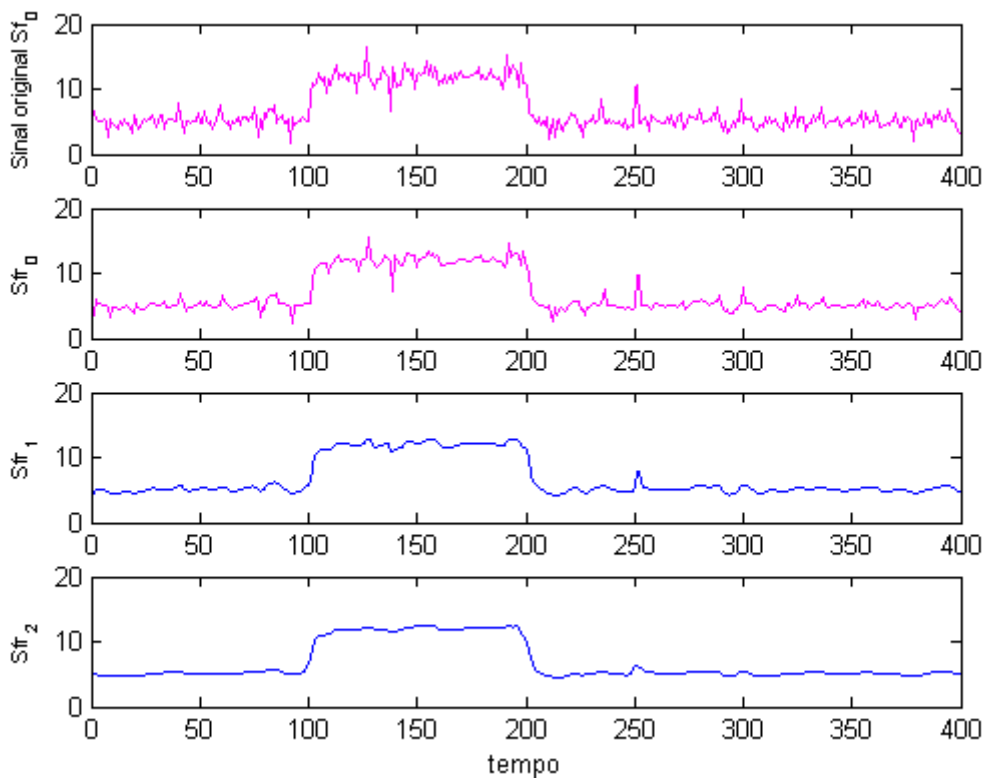


Figura B.5: Sinal original e sinais reconstruídos aos níveis de reconstrução completa, nível 1 e nível 2.

B.5 Valores de Corte para Detecção de Estacionários

Se o objetivo é a detecção de estados estacionários em um sinal de processo, alguns valores de corte devem ser calculados pelo algoritmo, como mencionado na seção 4.1. No algoritmo, são calculados T_s , T_u e T_w a partir de dados históricos.

O valor de corte abaixo do qual se considera estacionário, T_s , é calculado como o desvio padrão do módulo dos coeficientes de detalhe, na escala característica, obtidos pela transformada das ondaletas dos dados históricos. Acima de T_u , os dados são completamente não-estacionários, e este valor de corte é calculado como 3 vezes o valor de T_s , corrigido também por um fator de variabilidade λ_2 , considerado igual a 2 neste algoritmo. T_u também pode ser informado diretamente, se o processo e sua variabilidade são conhecidos. Por fim, o valor de T_w , que é o valor de corte para a derivada de segunda ordem (transformada das ondaletas da transformada das ondaletas de dados históricos na escala característica), é determinado como a média do módulo dos coeficientes de detalhe de segunda ordem.

Estes valores de corte são determinados apenas uma vez para cada variável, e, então, para qualquer determinação de estacionários desta variável usam-se estes valores de corte inicialmente estimados. Como resultado de sua aplicação, tem-se o exemplo da **Figura B.6**.

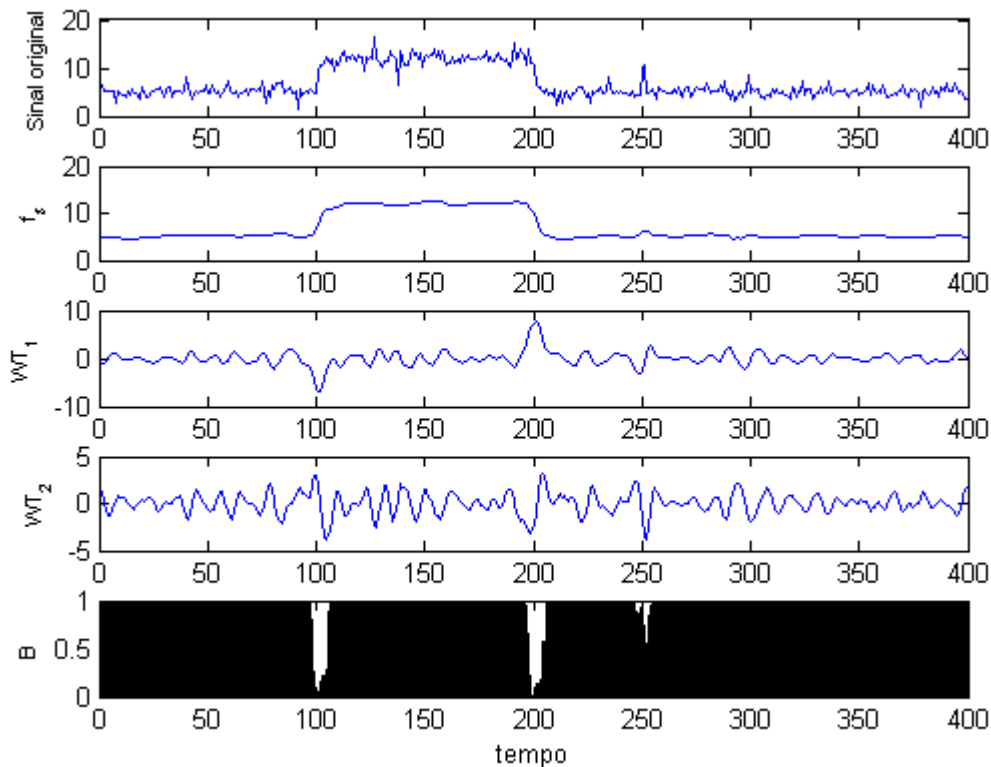


Figura B.6: Sinal original, tendência do sinal f_s , derivada de primeira ordem WT_1 , derivada de segunda ordem WT_2 , índice de estado estacionário B .

Apêndice C

Propriedades das Ondaletas

As propriedades mais importantes das ondaletas são as condições de admissibilidade e a regularidade, e são elas que dão às ondaletas o seu nome. Estas propriedades, entre outras, serão apresentadas a seguir. As propriedades de cada ondaleta é que definem o tipo de sinal aproximado e sua melhor aplicação. Isto quer dizer que, de acordo com a aplicação desejada, a ondaleta deve seguir algumas características específicas.

C.1 Admissibilidade

Quando a transformada de um sinal é utilizada para se obterem melhores informações sobre o mesmo, deve-se assegurar que este sinal possa ser perfeitamente reconstruído a partir de sua representação, pois, senão, a representação torna-se sem significado relevante.

Para a WT, a condição que deve ser seguida para garantir a reconstrução perfeita é (C.1). Esta condição é conhecida como “condição de admissibilidade” para a ondaleta $\psi(t)$. A prova de (C.1) pode ser encontrada em MERTINS (1999), seção 8.3.

Funções quadráticas integradas $\psi(t)$ que satisfazem a condição de admissibilidade

$$\int \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty \quad (\text{C.1})$$

podem ser usadas para analisar primeiramente o sinal e então reconstruí-lo sem perda de informação (VALENS, 2004). Em C.1, $\Psi(\omega)$ representa a transformada de Fourier da ondaleta $\psi(t)$.

A condição de admissibilidade implica que a transformada de Fourier de $\psi(t)$ desapareça na frequência zero, ou seja,

$$|\Psi(\omega)|^2 \Big|_{\omega=0} = 0 \quad (\text{C.2})$$

Isto significa que as ondaletas devem ter um espectro do tipo passa-banda. Esta é uma observação importante, que pode ser usada para a construção de uma transformada de ondaletas eficiente.

Um zero na frequência zero também significa que o valor médio da ondaleta no domínio do tempo deve ser zero

$$\Psi(0) = \int \psi(t) dt = 0 \quad (\text{C.3})$$

Além disso, $|\Psi(\omega)|$ deve decair rapidamente para $|\omega| \rightarrow 0$ e para $|\omega| \rightarrow \infty$. Ou seja, $\psi(t)$ deve ser uma resposta ao impulso de banda. Como uma resposta ao impulso de banda parece como uma onda pequena, a transformada é chamada de “transformada das ondaletas” (MERTINS, 1999).

C.2 Momentos Nulos

A decomposição de sinais é muitas vezes realizada com a intenção de compactação de sinais. Muitos destes sinais a serem comprimidos são de natureza de baixa frequência e podem ser bem aproximados localmente por polinômios de baixa ordem (MERTINS, 1999). Portanto, é interessante buscar por ondaletas com boas propriedades de aproximação para polinômios de baixa ordem. Assim, pode-se ver que estas propriedades de aproximação de uma decomposição multi-resolução estão intimamente relacionadas ao número de momentos nulos de uma ondaleta.

Da maneira como a transformada contínua das ondaletas é formulada, uma ondaleta é simplesmente uma função de energia finita com média zero (MALLAT, 1999). Uma das características mais importantes das ondaletas é o seu número de momentos nulos (*vanishing moments*) ou ordem de aproximação,

$$m_k = \int_{-\infty}^{\infty} t^k \psi(t) dt = 0 \quad (\text{C.4})$$

para $0 \leq k < n$, onde m_k representa o k -ésimo momento, cujo número ideal depende da aplicação desejada. Se uma ondaleta possui n momentos nulos, então a ordem de aproximação da transformada das ondaletas é também n . Os momentos não precisam ser exatamente zero, pois um valor pequeno já é suficiente o bastante.

De uma maneira análoga à Transformada de Fourier, os momentos também podem ser representados pelo decaimento rápido da ondaleta ψ com n momentos nulos se existe uma função θ com decaimento rápido tal que

$$\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n} \quad (\text{C.5})$$

Assim, se a função θ tem n zeros em $t = 0$, a ondaleta tem n momentos nulos, o que significa que a Eq. (C.4) tem $k = 0, 1, \dots, n-1$ (MERTINS, 1999).

Como consequência,

$$Wf(u, s) = s^n \frac{d^n}{du^n} (f * \bar{\theta}_s)(u) \quad (\text{C.6})$$

onde $\bar{\theta}_s(t) = 1/\sqrt{s} \theta(-t/s)$. Além disso, ψ não possui mais do que n momentos nulos se, e somente se, $\int_{-\infty}^{+\infty} \theta(t) dt \neq 0$. Então, uma ondaleta com n momentos nulos pode estimar derivadas de n -ésima ordem do sinal. Assim, ondaletas com um grande número de momentos nulos podem representar sinais com poucos coeficientes não-nulos, proporcionando uma representação esparsa apropriada para compressão de dados e cálculos rápidos (MALLAT, 1999).

A propriedade de momentos nulos torna possível a análise da regularidade local do sinal. Uma ondaleta possui m momentos nulos se e somente se sua função de escala pode gerar polinômios de grau menor ou igual a m . Se as estruturas polinomiais do sinal até ordem m são completamente reproduzidas pelas funções de escala, os coeficientes das ondaletas são anulados na representação multi-resolução.

Enquanto esta propriedade é usada para descrever o "poder de aproximação" das funções de escala, no caso das ondaletas ela tem um duplo sentido, isto é, a possibilidade de caracterizar a ordem de singularidades isoladas. O número de momentos nulos é determinado somente pelos coeficientes $h[n]$ do filtro h que é relacionado à função de escala. Se a transformada de Fourier da ondaleta é p vezes continuamente diferenciável, então 3 condições são equivalentes: a ondaleta ψ possui p momentos nulos; a função de escala j pode gerar polinômios de grau menor ou igual a p ; e a função de transferência do filtro h e suas primeiras $p-1$ derivadas desaparecem em $\omega = p$.

A escolha do número de momentos nulos na construção de uma ondaleta, bem como mais informações sobre esta propriedade, está detalhada na seção 8.6.5 de MERTINS (1999).

C.3 Suporte Compacto

Se os filtros passa-baixa e passa-alta $h[n]$ e $g[n]$ são filtros FIR, então as funções de escala e ondaleta resultantes têm suporte finito (DAUBECHIES, 1998). A prova pode ser vista na seção 8.6.7 de MERTINS (1999).

A função de escala possui suporte compacto se e somente se o filtro h possui suporte finito, e seus suportes sejam de mesmo tamanho. Se o suporte da função de escala é $[N_1, N_2]$, então o suporte da ondaleta é $[(N_1-N_2+1)/2, (N_2-N_1+1)/2]$ (MALLAT, 1999).

C.4 Regularidade

Regularidade é uma medida da suavidade de uma função, associada, portanto, com a derivação do sinal (WORNELL, 1996). Em particular, uma função é dita regular de ordem R se sua transformada de Fourier $F(\omega)$ decai de acordo com

$$F(\omega) \sim O(|\omega|^{-R}), \quad |\omega| \rightarrow \infty \quad (\text{C.7})$$

Usa-se o termo “regular” para definir uma função que é, pelo menos, regular de primeira ordem, e observe que uma função regular de ordem R tem $R-1$ derivadas regulares.

Da mesma forma, para ondaletas

$$\Psi(\omega) \sim O(|\omega|^{-1}), \quad |\omega| \rightarrow \infty \quad (\text{C.8})$$

A notação $O(\cdot)$ é interpretada como

$$A(\omega) \sim O(B(\omega)), \quad |\omega| \rightarrow \infty$$

$$\lim_{\omega \rightarrow \infty} \frac{A(\omega)}{B(\omega)} < \infty \quad (\text{C.9})$$

Normalmente, funções suaves para $\phi(t)$ são preferidas, pois estas podem possuir várias derivadas contínuas. Regularidade é uma propriedade associada aos filtros passa-baixa. Dada uma função $\phi(t)$, a função $\psi(t)$ de acordo com Eq. (2.23) será contínua para qualquer seqüência $g[n]$ (MERTINS, 1999).

Em geral, a suavidade da ondaleta-mãe é proporcional ao comprimento do suporte, ou seja, um suporte mais longo implica em uma ondaleta mais suave. Entretanto, deve-se observar a relação implícita na análise tempo-frequência entre resolução no domínio do tempo (aumenta com diminuição do comprimento do suporte) e resolução no domínio da frequência

(que aumenta com aumento do comprimento do suporte) (BRADLEY e WILSON, 2004). Assim, DAUBECHIES (1992) mostrou que a suavidade é incompatível com as propriedades de ortonormalidade e suporte compacto.

Mais conceitos matemáticos sobre esta propriedade podem ser encontrados em MALLAT (1999).

C.5 Simetria

A resposta de fase de um filtro, e conseqüentemente a ondaleta, é definida pelas suas propriedades de simetria. Se uma ondaleta-mãe tem simetria par ou ímpar (tanto simétrica quanto anti-simétrica) então ela terá fase linear (BRADLEY e WILSON, 2004).

Funções de escala e ondaletas simétricas são importantes pois são usadas para construir bases de ondaletas regulares em um intervalo, ao invés de todo o eixo real (MALLAT, 1999). Daubechies provou que, para uma ondaleta ser simétrica ou anti-simétrica, seu filtro deve ter uma fase linear complexa, e o único filtro espelhado conjugado simétrico compactamente suportado é o filtro de Haar (uma B-spline de ordem zero), que corresponde a uma ondaleta descontínua com somente um momento nulo. Além da ondaleta de Haar, não há outra que seja simétrica, ortogonal e de suporte compacto. Assim, uma ondaleta de suporte compacto (ortogonal) não pode ser simétrica e contínua (MALLAT, 1999).

A simetria é importante para a análise de sinais finitos e indica que o filtro possui fase linear, que é uma propriedade desejada pois significa que o filtro tem atraso de fase constante, ou seja, todos os componentes de frequência do sinal são atrasados pelo mesmo número de amostragens quando passam pelo filtro (OPPEHEIM e WILLSKY, 1997). Isto significa que se um número de componentes de frequência em fase passa através de um filtro de fase linear, eles estarão em fase na saída também. Isto não acontece para um filtro de fase não-linear, gerando distorção de fase.

Há uma vantagem adicional em se usar uma ondaleta com fase linear: como a ondaleta é simétrica, o número de multiplicações na integral da convolução pode ser reduzido à metade (adicionando amostras ao sinal antes da multiplicação pelos coeficientes do filtro) (BRADLEY e WILSON, 2004).

A melhor maneira de obter simetria é usando filtros espelhados em quadratura (QMF). QMF são pares combinados de filtros passa-baixa e passa-alta com características de corte similares, de modo que quando são aplicados simultaneamente (combinação dos filtros equivale a uma multiplicação) o filtro passa-banda resultante tem forma simétrica em ambos lados da sua frequência central. O par de QMF divide o sinal original em componentes de baixa e alta frequências, sendo que a divisão é o ponto que divide o intervalo entre 0 Hz e metade da frequência de amostragem dos dados pelo meio (ALSBERG *et al.*, 1997).

C.6 Energia

Uma questão importante no processo de decomposição é pensar na reconstrução do sinal. A reconstrução perfeita de um sinal nem sempre é possível. A reconstrução estável do sinal a partir das suas séries de decomposição por ondaletas é possível (DAUBECHIES, 1992), e tem como condição necessária e suficiente que a energia dos coeficientes da ondaleta deve ficar entre dois limites positivos, $0 < A \leq B < \infty$, conforme Eq. (C.10).

$$A\|f\|^2 \leq \sum_{j,k} \left| \langle f, \psi_{j,k} \rangle \right|^2 \leq B\|f\|^2 \quad (\text{C.10})$$

onde $\|f\|^2$ é a energia de $f(t)$, e A e B são independentes de $f(t)$.

Ao avaliar a qualidade da reconstrução, se falamos que duas funções são parecidas, deve-se esperar observar esta semelhança também no espaço dos coeficientes das ondaletas. Matemáticos dizem que estes dois espaços têm a mesma topologia. O efeito disto é que podemos estabelecer limites inferiores e superiores para a soma de todos os produtos internos entre a função e a função da ondaleta. No caso, estes limites são os valores A e B da Eq. (C.10).

Quando a Eq. (C.10) é satisfeita, a família de funções base $\psi_{j,k}(t)$ é chamada de *frame* (VALENS, 2004). Se os dois limites A e B são iguais, diz-se que o *frame* é estreito e, neste caso, as ondaletas discretas comportam-se exatamente como uma base ortonormal. Assim, *frames* constituem um tipo de ondaleta que não é ortonormal, e podem ser vistos como algo entre a WT ortonormal não-redundante e a muito redundante CWT (ALSBERG *et al.*, 1997).

Referências Bibliográficas

- AKBARYAN, F. e BISHNOI, P. R.. Smooth representation of trends by a wavelet-based technique. *Computers and Chemical Engineering*, **24**, p. 1913-1943 (2000).
- ALSBERG, B. K.; WOODWARD, A. M. e KELL, D. B.. An introduction to wavelet transforms for chemometricians: A time-frequency approach. *Chemometrics and Intelligent Laboratory Systems*, **37**, p. 215-239 (1997).
- ALTMANN, J.. *Surfing the wavelets* (1996). Disponível em: <http://www.wavelet.org/tutorial/>.
- AspenTech. Analysis of data storage technologies for the management of real-time process manufacturing data. Disponível em: http://www.advanced-energy.com/Upload/symphony_wp_infoplus.pdf (2001).
- BOGGESESS, A. e NARCOWICH, F. J.. *A first course in wavelets with Fourier analysis*. Prentice Hall, Texas (2001).
- BOOR, C. de. *Spline Toolbox for use with Matlab*. Version 2. MathWorks, Inc. (1999).
- BORES Signal Processing: Fordwater Pond Road Woking*, United Kingdom (2004). Disponível em: <http://www.bores.com/courses/intro/filters/>.
- BOURG, D. M.. *Excel Scientific and Engineering Cookbook*. O'Reilly (2006).
- BRADLEY, A. P. e WILSON, W. J.. On wavelet analysis of auditory evoked potentials. *Clinical Neurophysiology*, **115**, p. 1114–1128 (2004).
- BRISTOL, E.H.. Swimming door trending: adaptative trend recording, in: ISA National Conference Proceedings, p. 749-753 (1990).
- BROWN, P. R. e RHINEHART, R. R.. Automated steady-state identification in multivariate systems. *Process Control and Information Systems*, Hydrocarbon Processing, september (2000).
- BURRUS, C. S.; GOPINATH, R. A. e GUO, H.. *Introduction to Wavelets and Wavelet Transforms: a primer*. Prentice Hall, Englewood Cliffs, NJ (1997).

- CAO, S. e RHINEHART, R. R.. An efficient method for on-line identification of steady state. *J. Process Control*, **5** (6), p. 363-374 (1995).
- CHARBONNIER, S.; BELTAN, C. G., CADET, C. e GENTIL, S.. Trends extraction and analysis for complex system monitoring and decision support. *Engineering Applications of Artificial Intelligence*, **18**, p. 21–36 (2005).
- CHIANG, L. H.; RUSSELL, E. L. e BRAATZ, R. D.. *Fault detection and diagnosis in industrial systems*. Springer-Verlag, Great Britain (2001).
- CHING, P. C.; SO, H. C. e WU, S. Q.. On Wavelet Denoising and its Applications to Time Delay Estimation. *IEEE Transactions on Signal Processing*, **47** (10), p. 2879-2882 (1999).
- DAUBECHIES, I. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PE (1992).
- DAUBECHIES, I.. Orthonormal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.*, **41**, p. 909-996 (1998).
- DOYMAZ, F.; BAKHTAZAD, A.; ROMAGNOLI, J. A. e PALAZOGLU, A.. Wavelet-based robust filtering of process data. *Computers and Chemical Engineering*, **25**, p. 1549-1559 (2001).
- DUAN, J. e OWEIS, I. S.. Dyadic wavelet analysis of PDA (pile dynamic analyzer) signals. *Soil Dynamics and Earthquake Engineering*, **25**, p. 661-677 (2005).
- DUTTILEX, P.. *An implementation of the algorithme à trous to compute the wavelet transform*, em *Wavelets: Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman e P. Tchamitchian, Eds. Berlin, Germany: Springer-Verlag, p. 298–304 (1989).
- FLEHMIG, F., WATZDORF, R. v. e MARQUARDT, W.. Identification of trends in process measurements using the Wavelet Transform. *Computers and Chemical Engineering*, **22**, p. S491-S496 (1998).
- GOODWIN, G.; GRAEBE, S. F. e SALGADO, M. E.. *Control System Design*. Prentice Hall, Upper Saddle River, New Jersey, p.319-320 (2001).
- HALE, J.C., SELLARS, H.L.. Historic data recording for process computers. *Chemical Engineering Progress*, **77**, p. 38-43, November (1981).
- HOLSCHNEIDER, M.; KRONLAND-MARTINET, R.; MORLET, J. e TCHAMITCHIAN, P.. *A real-time algorithm for signal analysis with the help of the wavelet transform*, em *Wavelets: Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman e P. Tchamitchian, Eds. Berlin, Germany: Springer-Verlag, p. 286–297 (1989).

- INGLE, V. K. e PROAKIS, J. G.. *Digital Signal Processing Using MATLAB*. Version 4. PWS Publishing Company, Boston, MA (1997).
- JIANG, T.; CHEN, B. e HE, X.. Industrial application of Wavelet Transform to the on-line prediction of side draw qualities of crude unit. *Computers and Chemical Engineering*, **24**, p. 507-512 (2000).
- JIANG, T.; CHEN, B.; HE, X. e STUART, P.. Application of steady-state detection method based on wavelet transform. *Computers and Chemical Engineering*, **27**, p. 569-578 (2003).
- KRAUSS, T. P.; SHURE, L. e LITTLE, J. N.. *Signal Processing Toolbox User's Guide*, for use with Matlab®. Versão 4.2. The MathWorks, Inc., Natick, MA (1999).
- LANG, M.; GUO, H.; ODEGARD, J. E.; BURRUS, C. S. e R. O. WELLS Jr. *Nonlinear processing of a shift invariant DWT for noise reduction* em Proc. SPIE Wavelet Applications II, vol. 2491, H. H. Szu, Ed., Orlando, FL, p. 640–651 (1995).
- LANG, M.; GUO, H.; ODEGARD, J. E.; BURRUS, C. S. e R. O. WELLS Jr. Noise reduction using an undecimated discrete wavelet transform. *IEEE Signal Process. Lett.*, **3** (1), p. 10–12 (1996).
- LATHI, B. P.. *Signal processing and linear systems*. Berkeley-Cambridge Press, Carmichael, CA, p. 500-506 (1998).
- LEVINE, W. S.. *The Control Handbook*. CRC and IEEE Press, Boca Raton, Florida, p.239,240,253-4,261-3,313-4,1033 (1996).
- LJUNG, L.. *MATLAB System Identification Toolbox, User's Guide*. Versão 4.0.5. The MathWorks, Inc., Natick, MA (1997).
- LOSADA, R. A.. *Practical FIR Filter Design in Matlab*. Revision 1.1. The MathWorks, Inc. (2004).
- MACKENZIE, D.. Wavelets: seeing the forest and the trees. *Beyond Discovery™: The Path from Research to Human Benefit*, National Academy of Sciences (2001).
- MADISETTI, V. e WILLIAMS, D. B.. *The digital signal processing handbook*. CRC and IEEE Press, Boca Raton, Florida (1998).
- MALLAT, S.. *A wavelet tour of signal processing*. Academic Press (1999). Uma versão adaptada por F. Chaplais é disponibilizada em: http://cas.ensmp.fr/~chaplais/Wavetour_presentation/Wavetour_presentation_US.html.
- MALLAT, S.. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2** (7), p. 674-693 (1989).

- MALLAT, S. e ZHONG, S. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14** (7), p. 710-732 (1992).
- MATLAB® é uma marca registrada de The Mathworks, Inc.. Natick, MA, USA.
<http://www.mathworks.com>.
- MERTINS, A.. *Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transforms and Applications*. John Wiley and Sons, England (1999).
- MISITI, M.; MISITI, Y.; OPPENHEIM, G. e POGGI, J.. *Wavelet Toolbox, for use with Matlab*. Version 3. The MathWorks, Inc. (2006).
- MISRA, M., KUMAR, S., QIN, S. J. E SEEMANN, D.. Online data compression and error analysis using wavelet technology. *AIChE Journal*, **46**, p. 119-132 (2000).
- MORETTIN, P. A.. *Ondaletas e seus usos em estatística*. Universidade de São Paulo, SP (1997).
- NARASIMHAN, S.; MAH, S. H. e TAMHANE, A. C.. A Composite Statistical Test for Detecting Changes of Steady States. *AIChE Journal*, **32** (9), p. 1409-1418 (1986).
- NARASIMHAN, S. e JORDACHE, C.. *Data Reconciliation & Gross Error Detection: An Intelligent Use of Process Data*. Gulf Publishing Company, Houston, Texas (2000).
- OAKLAND, J. S.. *Statistical Process Control*. 5ª ed. Butterworth-Heinemann, Burlington, MA (2003).
- O'HAVER, T. *An Introduction to Signal Processing in Chemical Analysis* (2005).
Department of Chemistry and Biochemistry, University of Maryland. Disponível em:
<http://www.wam.umd.edu/~toh/spectrum/TOC.html>.
- OPPEHEIM, A. V. e WILLSKY, A. S.. *Signals & Systems*. Prentice-Hall, Upper Saddle River, N.J. (1997).
- OSI Software Inc.. PI data storage component overview. Disponível em:
<http://www.osisoft.com/270.htm> (2002).
- POLIKAR, R.. The Story of Wavelets. in *Physics and Modern Topics in Mechanical and Electrical Engineering*, (ed. Mastorakis, N), World Scientific and Eng. Society Press, p. 192-197 (1999).
- POLIKAR, R.. *The wavelet tutorial: the engineer's ultimate guide to wavelet analysis*. Rowan University, College of Engineering (2001). Disponível em:
<http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>.
- INGLE, V. K. e PROAKIS, J. G.. *Digital Signal Processing using Matlab*. PWS Publishing Company, ITP (1997).

- RIOUL, O. e VETTERLI, M.. Wavelets and signal processing. *IEEE Signal Processing Magazine*, p. 14-38 (1991).
- SAS OnlineDoc®, Version 8, Cary, NC: SAS Institute Inc., USA (1999). Disponível em: <http://www.sasdocs.utoledo.edu/stat/chap65/sect43.htm>.
- SECCHI, A. R.; SANTOS, M.G.; NEUMANN, G.A. e TRIERWEILER, J.O.. A dynamic model for a FCC UOP stacked converter unit. *Computers and Chemical Engineering*, 25, p. 851–858 (2001).
- SMITH, J. O.. *Introduction to Digital Filters*. "FIR Filters" section (2004). Disponível em: http://ccrma.stanford.edu/~jos/filters/Finite_Impulse_Response_FIR.html.
- SUNDARESAN, K. R. e HRISHNASWAMY, P. R.. Estimation of time delay time constant parameters in time, frequency, and Laplace domains. *Can. J. Chem. Eng.*, 56, p. 257 (1977).
- The discrete wavelet transform*. Cap. 2, p. 8. Disponível em: http://etd.lib.fsu.edu/theses/available/etd-11242003-185039/unrestricted/09_ds_chapter2.pdf.
- THORNHILL, N. F., CHOUDHURY, M.A.A.S. e SHAH, S.L.. The impact of compression on data-driven process analyses. *Journal of Process Control*, 14, p. 389-398 (2004).
- TRIERWEILER, J. O. e MACHADO, V. C.. Which is the best criterion for identification of dynamic models?, in Proceedings of the 7th DYCOPS (IFAC Symposium on Dynamics and Control of Process Systems), Boston (2004).
- UNSER, M.. Texture classification and segmentation using wavelet frames. *IEEE Trans. Image Process.*, 4 (11), p. 1549–1560 (1995).
- VALENS, C. *A really friendly guide to wavelets* (2004). Disponível em: <http://perso.wanadoo.fr/polyvalens/clemens/wavelets/wavelets.html>.
- VEDAM, H. e VENKATASUBRAMANIAN, V.. A B-spline based method for data compression, process monitoring and diagnosis. *Computers and Chemical Engineering*, 22, p. S827-S830 (1998).
- ZACIU, R.; LAMBA, C.; BURLACU, C. e NICULA, G.. Image compression using an overcomplete discrete wavelet transform. *IEEE Trans. Consum. Electron.*, 42 (3), p. 800–807 (1996).
- WHITE, S. *Digital Signal Processing, A Filtering Approach*. 1 ed. Thomson Delmar Learning (2000).
- WORNELL, G. W.. *Signal processing with fractals: a wavelet-based approach*. Upper Saddle River, NJ, Prentice Hall (1996).

