

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

Avaliação de Desempenho de Malhas de Controle

DISSERTAÇÃO DE MESTRADO

Ariel de Oliveira Kempf

Porto Alegre

2003

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

Avaliação de Desempenho de Malhas de Controle

Ariel de Oliveira Kempf

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

Área de concentração: Controle de Processos

Orientador:
Prof. Dr. Jorge Otávio Trierweiler

Co-orientador:
Prof. Dr. Argimiro Resende Secchi

Porto Alegre

2003

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 *Avaliação de Desempenho de Malhas de Controle*, elaborada por Ariel de Oliveira Kempf, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:

Prof. Dr. Jorge Rodolfo Silva Zabadal

Prof. Dr. Luís Gustavo Soares Longhi

Eng. Dr. Mario Cesar Mello Massa de Campos

Eng. Herbert Campos G. Teixeira

Resumo

Avaliação de desempenho de malhas de controle, na abordagem baseada em controle com variância mínima, é uma tecnologia bastante promissora e de grandes potencialidades. No entanto ainda não é devidamente explorada no meio científico nacional e muito menos em nível industrial. Neste trabalho é apresentado um estudo abrangente na área de avaliação de desempenho de malhas de controle. Um vasto levantamento bibliográfico é feito e os fundamentos teóricos necessários para o entendimento das técnicas existentes são detalhadamente desenvolvidos. As características positivas e as principais falhas são apontadas. Diferentes abordagens são propostas, técnicas complementares à avaliação de desempenho baseada em variância mínima são sugeridas e questões de cunho prático são levantadas. Simulações e estudo de casos reais são levados a cabo, nos casos SISO e MIMO, com a finalidade de colocar em prática os procedimentos estudados, fornecer subsídios a uma avaliação crítica dos métodos apresentados e também possibilitar conclusões a respeito dos resultados obtidos. Como resultado, este trabalho introduz o assunto no meio acadêmico nacional, fornece diretrizes para a implementação industrial das técnicas apresentadas e demonstra, através da interpretação de resultados obtidos, que esta tecnologia é valiosa no sentido que pode resultar em grande retorno econômico ao resolverem-se os problemas detectados nos diagnósticos produzidos pela análise de desempenho de malhas de controle.

Palavras Chave: controle com variância mínima, avaliação de desempenho de malhas, auditoria de malhas *feedback*, matriz de interação, estimação de tempo morto

Abstract

Control loops performance assessment, in terms of control with minimum variance, is a very promising technology and shows big potential. However, it is still not properly exploited by the national scientific community and far less at industrial level. In this work is presented a huge study on the field of performance assessment of control loops. A comprehensive bibliographic review is done and the theoretical basis necessary to an appropriated understanding of the existing techniques is developed in details. The positive characteristics and the main drawbacks are appointed. Different approaches are proposed, complementary techniques to the performance assessment based on minimum variance are suggested and practical implementation remarks are discussed. Simulations and real case studies are carried out, in SISO and MIMO cases, to test the studied procedures, to provide means of critically evaluate the presented methods and to allow conclusions about the acquired results. As main results, this work introduces the subject to the national academic community, provides guidelines for an industrial implementation of the techniques presented and shows, through interpretation of the produced results, that this technology is valuable in the sense the it can result in large economic benefits by solving the problems detected by the performance assessment of control loops.

Keywords: minimum variance control, performance assessment of control loops, assessment of feedback control loops, interactor matrix, time delay estimation.

Agradecimentos

Meus agradecimentos vão para diversas pessoas que de alguma forma influenciaram no rumo desta dissertação. A todos os meus colegas do mestrado pela convivência agradável, que tornou mais leve a realização deste trabalho. Em especial aos amigos Liliane Pollo, André Muniz, Rafael Soares, Tiago Finkler, Vinícius Machado e Marcelo Farenzena, que me acompanham desde a graduação. Outras pessoas também contribuíram com seu tempo ou convivência em alguma parte desta dissertação, e recebem aqui meus agradecimentos.

Ao pessoal da TriSolutions, que motivou e deu força à realização deste estudo. Sucesso para todos nós: Simone Marcon, Leandro P. Lusa, Luciano Farina, Maurício Posser, Pedro Fernandes e Ricardo Duraiski. Ao pessoal da Petrobrás, na pessoa de Herbert Teixeira e de Euclides Almeida Neto, pelas oportunidades de contado com o ambiente industrial e pelas contribuições valiosas.

Também sou grato a todo o Departamento de Engenharia Química da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, por toda sua infra-estrutura de ponta, seu ambiente acolhedor e seu ensino de qualidade.

Agradeço ainda ao meu orientador, professor e colega, Jorge O. Trierweiler e ao meu co-orientador, professor e colega Argimiro R. Secchi pela instrução e suporte, pelas discussões proveitosas e pelo exemplo de profissionalismo e dedicação à pesquisa.

Agradeço finalmente a todos os meus familiares que souberam compreender minha ausência durante tanto tempo na realização de meus estudos. Espero compensar minha falta em breve. Um agradecimento especial à minha irmã Helena que suportou os piores momentos e esteve junto dos melhores e aos meus pais, Dieter e Arlete, a quem dedico este trabalho. Devo a vocês o que eu sou hoje.

Dedicação especial a Maria Dalira Hoffman de Oliveira

Sumário

| | |
|--|-----------|
| Capítulo 1 Introdução | 1 |
| 1.1. Relevância do tema desta dissertação..... | 2 |
| 1.2. Revisão histórica..... | 3 |
| 1.3. Produtos de software comerciais existentes | 6 |
| 1.4. Motivação | 7 |
| 1.5. Estrutura da dissertação | 9 |
| Capítulo 2 Revisão Bibliográfica | 11 |
| 2.1. Introdução..... | 12 |
| 2.2. Teoria de predição ótima | 13 |
| 2.3. Teoria de controle com variância mínima (CVM)..... | 18 |
| 2.4. Índices de desempenho baseados em CVM..... | 24 |
| 2.4.1. Índice de desempenho AGENER e ASUBS | 25 |
| 2.4.2. Índice de desempenho normalizado (ANORM) | 27 |
| 2.4.3. Índice de desempenho FCOR..... | 29 |
| 2.4.4. Outros índices | 31 |
| 2.5. Procedimentos complementares | 34 |
| 2.5.1. Procedimento de detecção de oscilações..... | 34 |
| 2.5.2. Índices de eficiência operacional | 36 |
| 2.5.3. Técnicas estatísticas..... | 38 |
| 2.6. Conclusões..... | 40 |
| Capítulo 3 Aplicações em malhas SISO..... | 43 |
| 3.1. Formas de implementação | 44 |
| 3.2. Efeito dos parâmetros e questões práticas | 46 |
| 3.2.1. Tempo morto ou horizonte de predição (d ou h) | 46 |
| 3.2.2. Tamanho do conjunto de dados (n)..... | 48 |
| 3.2.3. Ordem dos modelos autoregressivos (m)..... | 49 |
| 3.2.4. Intervalo de amostragem..... | 50 |
| 3.2.5. Fator de esquecimento..... | 50 |
| 3.2.6. Referência desejada..... | 51 |
| 3.2.7. Demais questões práticas | 52 |
| 3.3. Aplicação em exemplos | 54 |
| 3.3.1. Sistemas escolhidos..... | 54 |
| 3.3.2. Análise dos resultados | 60 |
| 3.3.3. Efeito do modelo de distúrbio..... | 65 |
| 3.3.4. Máxima sensibilidade | 68 |
| 3.3.5. Aplicação com dados de processo industrial..... | 70 |
| 3.4. Conclusões..... | 75 |
| Capítulo 4 Sistemas Multivariáveis | 79 |
| 4.1. Introdução..... | 80 |
| 4.2. A Matriz de Interação | 80 |
| 4.2.1. Definições..... | 80 |
| 4.2.2. Estimativa da Matriz de Interação..... | 83 |

| | |
|---|------------|
| 4.3. Índices de desempenho multivariáveis..... | 84 |
| 4.3.1. <i>Abordagem baseada em fatoração espectral</i> | 84 |
| 4.3.2. <i>Abordagem baseada em filtragem e correlação</i> | 88 |
| 4.4. Aplicação em sistemas multivariáveis..... | 91 |
| 4.4.1. <i>Simulações</i> | 92 |
| 4.5. Conclusões..... | 95 |
| Capítulo 5 Testes em unidade experimental..... | 97 |
| 5.1. Introdução..... | 98 |
| 5.2. Descrição da unidade..... | 98 |
| 5.2.1. <i>Configuração e instrumentação</i> | 99 |
| 5.2.2. <i>Zonas de operação e comportamento dinâmico</i> | 100 |
| 5.3. Estrutura de controle..... | 102 |
| 5.3.1. <i>Malhas de controle</i> | 102 |
| 5.3.2. <i>Ajuste empregado</i> | 104 |
| 5.4. Procedimento de testes..... | 105 |
| 5.5. Desempenho com controladores Multi-SISO..... | 106 |
| 5.5.1. <i>Caso 1</i> | 106 |
| 5.5.2. <i>Caso 2</i> | 111 |
| 5.6. Desempenho com controlador NMPC..... | 115 |
| 5.6.1. <i>Caso 1</i> | 115 |
| 5.6.2. <i>Caso 2</i> | 116 |
| 5.7. Influência da matriz de pesos..... | 120 |
| 5.8. Conclusões..... | 121 |
| Capítulo 6 Conclusões gerais e sugestões..... | 123 |
| 6.1. Principais Conclusões..... | 123 |
| 6.2. Sugestões para Futuros Trabalhos..... | 125 |
| Apêndice A - Cálculo da matriz de interação..... | 129 |
| A.1 Definições..... | 129 |
| A.2 O algoritmo..... | 131 |
| Apêndice B - Estimação da matriz de interação..... | 135 |
| B.1 Introdução..... | 135 |
| B.2 Determinação da ordem da matriz de interação..... | 136 |
| B.3 Estimação da matriz de interação em malha fechada..... | 138 |
| B.4 Problemas de posto dos parâmetros de Markov..... | 139 |
| Apêndice C - Algoritmo ARMarkov-LS..... | 143 |
| C.1 Introdução..... | 143 |
| C.2 Determinação da ordem da matriz de interação..... | 144 |
| C.3 Extensão para o caso MIMO..... | 146 |
| C.4 Propriedades estatísticas..... | 148 |
| C.4.1 <i>Consistência do mínimos-quadrados</i> | 148 |
| C.4.2 <i>Covariância das estimativas para sistemas com ruído branco</i> | 149 |
| C.4.3 <i>Covariância das estimativas com ruído não-branco</i> | 150 |
| Apêndice D - Figuras Complementares..... | 151 |
| D.1 Sinais de excitação..... | 151 |
| D.2 Análise de Correlação..... | 152 |

| | |
|---|------------|
| Apêndice E - Funções Polinomiais | 157 |
| Referências Bibliográficas | 161 |

Lista de Figuras

| | |
|--|----|
| Figura 1.1. Distribuições amostrais da temperatura do forno antes (linha cheia) e após (linha pontilhada) a redução da variabilidade de malha de controle e da mudança de ponto de operação. | 8 |
| Figura 2.1. Exemplo de análise de componentes principais. | 26 |
| Figura 2.2. Diagrama esquemático do sistema de estimação da seqüência de inovações..... | 31 |
| Figura 2.3. Seqüência do procedimento manual de diagnóstico das causas de oscilação..... | 36 |
| Figura 3.1. Janela móvel dos dados de um forno industrial, com estimação a cada 1750 pontos. | 44 |
| Figura 3.2. Avaliação recursiva do índice de desempenho do mesmo conjunto de dados da Figura 3.1. A linha de cima apresenta a variável de processo subtraída do <i>setpoint</i> | 45 |
| Figura 3.3. Gráfico de horizonte de predição para uma malha de temperatura de uma refinaria. | 47 |
| Figura 3.4. Gráfico de horizonte de predição para uma malha de pressão de uma refinaria. Aqui fica clara a dificuldade de se escolher um patamar no qual o índice de desempenho não varia rapidamente. | 48 |
| Figura 3.5. Índice de desempenho em função do horizonte de predição para malha de nível onde não existe patamar algum na variação de η | 48 |
| Figura 3.6. Índice de desempenho em função do número de parâmetros do modelo para diferentes tipos de malha, mostrando que a influência deste fator deixa de ser importante para ordens maiores que 5. | 50 |
| Figura 3.7. Esquema para avaliação de desempenho com referência desejada. A malha superior apresenta a compensação com a referência desejada para um degrau no <i>setpoint</i> . A malha inferior é o caso normal. Os resultados encontram-se na Figura 3.8. | 52 |
| Figura 3.8. Resultados utilizando a referência desejada (dois primeiros gráficos) comparativamente com a mudança normal de <i>setpoint</i> . O erro da saída é bem menor no caso da referência desejada. | 52 |
| Figura 3.9. Malha <i>feedback</i> montada no Simulink® para simulação dos sistemas propostos. | 58 |
| Figura 3.10. Resultados para o sistema G_{11} | 59 |
| Figura 3.11. Resultados para o sistema G_{12} | 59 |
| Figura 3.12. Resultados para o sistema G_{13} | 59 |
| Figura 3.13. Resultados para o sistema G_{21} | 59 |
| Figura 3.14. Resultados para o sistema G_{31} | 59 |
| Figura 3.15. Resultados para o sistema G_{32} | 59 |
| Figura 3.16. Resultados para o sistema G_{41} | 60 |
| Figura 3.17. Resultados para o sistema G_{42} | 60 |
| Figura 3.18. Resultados para o sistema G_{43} | 60 |
| Figura 3.19. Comparação da saída do modelo de distúrbio e a saída do sistema em malha fechada para o sistema G_{43} | 61 |
| Figura 3.20. Peça da resposta ao impulso do sistema G_{31} . O ganho estacionário do sistema foi normalizado em uma unidade. | 64 |
| Figura 3.21. Sistema G_{11} com distúrbio N_1 | 66 |
| Figura 3.22. Sistema G_{11} com distúrbio N_2 | 66 |
| Figura 3.23. Sistema G_{11} com distúrbio N_3 | 66 |
| Figura 3.24. Sistema G_{11} com distúrbio N_4 | 66 |
| Figura 3.25. Sistema G_{11} com distúrbio N_1 e distúrbio na carga em $t = 200$ | 68 |

| | |
|---|-----|
| Figura 3.26. Sistema G_{11} com distúrbio N_2 e distúrbio na carga em $t = 200$. | 68 |
| Figura 3.27. Sistema G_{11} com distúrbio N_3 e distúrbio na carga em $t = 200$. | 68 |
| Figura 3.28. Sistema G_{11} com distúrbio N_4 e distúrbio na carga em $t = 200$. | 68 |
| Figura 3.29. Diagrama de Nyquist ilustrando a função complementar de sensibilidade, P . | 69 |
| Figura 3.30. Variação do índice de desempenho com a ordem do polinômio de ajuste e com o tempo morto para uma malha de vazão. | 72 |
| Figura 3.31. Variação do índice de desempenho com a ordem do polinômio de ajuste e com o tempo morto para uma malha de pressão. | 73 |
| Figura 3.32. Gráficos gerados pelo procedimento de detecção de oscilação para uma malha de vazão não oscilatória. | 75 |
| Figura 3.33. Gráficos gerados pelo procedimento de detecção de oscilação para uma malha de vazão oscilatória. | 75 |
| Figura 4.1. Diagrama em malha fechada do sistema utilizado para identificação de parâmetros de Markov. | 83 |
| Figura 4.2. Avaliação de desempenho de um sistema MIMO com variação do grau de interação entre os canais. | 94 |
| Figura 4.3. Avaliação de desempenho de um sistema MIMO com saídas ponderadas e variação do grau de interação entre os canais. | 95 |
| Figura 5.1. Unidade de 2 tanques de níveis acoplados. | 98 |
| Figura 5.2. As três regiões de operação da unidade de dois tanques. | 100 |
| Figura 5.3. Esquema da unidade de dois tanques explicitando a estrutura de controle utilizada. | 102 |
| Figura 5.4. Diagrama esquemático da malha de controle com o uso do NMPC. | 104 |
| Figura 5.5. Região 0 em operação normal. | 107 |
| Figura 5.6. Região 0 com perturbações. | 107 |
| Figura 5.7. Região 1 em operação normal. | 107 |
| Figura 5.8. Região 1 com perturbações. | 107 |
| Figura 5.9. Região 2 em operação normal. | 107 |
| Figura 5.10. Região 2 com perturbações. | 107 |
| Figura 5.11. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 0. | 108 |
| Figura 5.12. Índices de desempenhos fragmentado e para todo o conjunto de dados, estimados para a Região 1. | 109 |
| Figura 5.13. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 2. | 109 |
| Figura 5.14. Mudança no comportamento da atuação de VP (esquerda) e conseqüente alteração no padrão da vazão de alimentação (direita). | 111 |
| Figura 5.15. Região 0 em operação normal. | 112 |
| Figura 5.16. Região 0 com perturbações. | 112 |
| Figura 5.17. Região 1 em operação normal. | 112 |
| Figura 5.18. Região 1 com perturbações. | 112 |
| Figura 5.19. Região 2 em operação normal. | 112 |
| Figura 5.20. Região 2 com perturbações. | 112 |
| Figura 5.21. Índices de desempenho fragmentados e para todo o conjunto de dados, Região 0. | 113 |
| Figura 5.22. Índices de desempenho fragmentados e para todo o conjunto de dados, Região 1. | 114 |
| Figura 5.23. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 2. | 114 |
| Figura 5.24. Dados coletados para o experimento com o NMPC na Região 0. | 115 |

| | |
|--|-----|
| Figura 5.25. Índices de desempenho fragmentados e para todo o conjunto de dados, como resultado da utilização do NMPC na Região 0. | 116 |
| Figura 5.26. Dados coletados para o experimento com o NMPC na Região 0 com adição de tempo morto extra. | 117 |
| Figura 5.27. Índices de desempenho fragmentados e para todo o conjunto de dados para a Região 0, Caso 2, utilizando NMPC. | 117 |
| Figura 5.28. Dados coletados com o uso do NMPC na Região 1. | 118 |
| Figura 5.29. Índices de desempenho fragmentados e para todo o conjunto de dados coletados na Região 1. | 118 |
| Figura 5.30. Dados relativos ao experimento com NMPC na Região 2. | 119 |
| Figura 5.31. Índices de desempenho fragmentados e para todo o conjunto de dados coletados na Região 2. | 119 |
| Figura 6.1. Perturbações inseridas no sistema operando na Região 0 nos Casos 1 (esquerda) e 2 (direita). | 151 |
| Figura 6.2. Perturbações inseridas no sistema operando na Região 1 nos Casos 1 (esquerda) e 2 (direita). | 152 |
| Figura 6.3. Perturbações inseridas no sistema operando na Região 2 nos Casos 1 e 2. | 152 |
| Figura 6.4. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 1, com controladores PI. | 152 |
| Figura 6.5. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 2, com controladores PI. | 153 |
| Figura 6.6. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 1, com controladores PI. | 153 |
| Figura 6.7. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 2, com controladores PI. | 153 |
| Figura 6.8. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 1, com controladores PI. | 154 |
| Figura 6.9. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 2, com controladores PI. | 154 |
| Figura 6.10. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 1, com NMPC. | 154 |
| Figura 6.11. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 2, com NMPC. | 155 |
| Figura 6.12. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 2, com NMPC. | 155 |
| Figura 6.13. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 2, com NMPC. | 155 |

Lista de Tabelas

| | |
|---|-----|
| Tabela 3.1. Análise de pólos e zeros dos sistemas discretos..... | 57 |
| Tabela 3.2. Variância da saída dos sistemas sob ação dos controladores e em malha aberta.. | 61 |
| Tabela 3.3. Variâncias produzidas pelo CVM comparadas com as variâncias estimadas resultantes das análises de índices de desempenho com as saídas de um CVM..... | 62 |
| Tabela 3.4. Variâncias mínimas estimadas pelos diferentes métodos de avaliação quando alimentados com dados provenientes de um PI, em comparação com a variância do controlador de variância mínima..... | 63 |
| Tabela 3.5. Erros médios de estimação da variância mínima desconsiderando o sistema G_{31} | 65 |
| Tabela 3.6. Resultados para as variâncias do sistema G_{11} com diferentes modelos de distúrbio sob ação de um CVM..... | 66 |
| Tabela 3.7. Resultados para as variâncias do sistema G_{11} com diferentes modelos de distúrbio sob ação de um controlador PI ajustado por um método baseado no ITAE. | 67 |
| Tabela 3.8. Máxima sensibilidade dos sistemas em malha fechada. | 70 |
| Tabela 3.9. Índices de desempenho de diferentes malhas antes e depois de sua re-sintonia... | 71 |
| Tabela 3.10. Resultados do procedimento de detecção de oscilações em conjunto com os índices de desempenho..... | 74 |
| Tabela 5.1. Configuração das válvulas da unidade..... | 99 |
| Tabela 5.2. Parâmetros de ajuste dos controladores PI empregados nos testes. | 105 |
| Tabela 5.3. Valores para os <i>setpoints</i> utilizados nos testes para as 3 regiões de operação.... | 105 |

Nomenclatura

| | |
|----------------------------|--|
| a | ordem dos polinômios A e C |
| A | polinômio do denominador de um sistema ARMA que relaciona entrada/saída |
| B | polinômio correspondente ao numerador de um sistema ARMA que relaciona entrada/saída |
| C | polinômio do numerador da função de transferência do de um sistema ARMA do distúrbio para saída |
| $\text{cov}(\cdot)$ | operador covariância |
| D | matriz de interação |
| D_w | matriz de interação ponderada |
| d | tempo morto em número de intervalos de amostragem ou ordem da matriz de interação (caso multivariável) |
| $\text{deg}(\cdot)$ | operador grau de uma matriz polinomial ou de um polinômio |
| $\text{det}(\cdot)$ | operador determinante |
| $\text{diag}\{\cdot\}$ | operador de extração da diagonal de uma matriz |
| $\text{diag}(\cdot)$ | operador que gera uma matriz diagonal a partir de um vetor |
| $E\{\cdot\}$ ou $E(\cdot)$ | média (esperança) |
| e_t ou $e(k)$ | sinal de ruído branco |
| \hat{e}_t | estimativa de e_t |
| F | polinômio da resposta ao impulso de um sistema em malha fechada sob ação de um CVM |
| f_i | coeficientes do polinômio F |
| G | polinômio resultante da equação de Diophantine |
| G_i | função de transferência genérica de um sistema i ou parâmetros de Markov MIMO (Apêndice B) |
| \hat{H} | estimativa do índice de desempenho multivariável |
| h | horizonte de predição |
| h_i | parâmetros de Markov SISO |
| I | matriz identidade |
| J | função objetivo de otimização |
| k | variável representativa do tempo atual |
| L | representação RMF do numerador de T |
| \lim | operador limite dos argumentos |
| N | matriz de transferência de e_t para y_t |
| $\text{null}(\cdot)$ | operador de extração do espaço nulo de uma matriz |
| m | ordem do polinômio do erro de predição ou número de intervalos de amostragem a frente no tempo em que é feita uma predição |
| $\min(\cdot)$ | valor mínimo dos argumentos |
| n | tamanho de um conjunto de dados |
| p_i | coeficientes de um polinômio qualquer |
| Q | matriz de transferência do controlador |
| q | operador deslocamento no tempo |
| R | representação RMF do denominador de T |
| r_{ij} | covariância entre os sinais i e j |
| $\text{rank}(\cdot)$ | posto de uma matriz |
| s_e^2 | erro médio quadrático residual |

| | |
|---------------------|---|
| $\text{svd}(\cdot)$ | fatoração SVD dos argumentos |
| T | matriz de transferência de u_t para y_t |
| \tilde{T} | matriz de transferência fatorada (sem a estrutura de tempo morto) |
| t | tempo |
| $\text{tr}(\cdot)$ | operador traço de uma matriz |
| u_t | sinal de entrada |
| U_t | vetor de sinais de entrada |
| $\text{var}(\cdot)$ | operador variância |
| W | matriz de pesos |
| x | estados de um sistema |
| x_t | sinal resultante de $[B/A]u_t$ |
| X | matriz regressora de mínimos quadrados |
| y_{SET} | sinal de referência |
| y_t | sinal de saída |
| \bar{y}_t | média amostral da saída |
| \tilde{y}_t | sinal do erro de predição |
| \hat{y}_t | sinal do preditor ótimo |
| Y_t | vetor de sinais de saída |
| Y_t^{MV} | vetor de sinais de saída sob ação de um CVM |
| Z | vetor de coeficientes da correlação cruzada entre y_t e e_t . |

Símbolos gregos

| | |
|--------------------|---|
| r_{ij} | correlação cruzada entre sinais i e j |
| \mathbf{x} | índice de desempenho dado pela razão entre \mathbf{s}_y^2 e \mathbf{s}_{MV}^2 |
| l | fator de esquecimento em procedimentos recursivos |
| \mathbf{S}_e | variância do sinal de ruído multivariável |
| \mathbf{S}_{MIN} | variância mínima multivariável |
| \mathbf{G} | matriz real unitária |
| \mathbf{W}_t | sinal de saída deslocado no tempo |
| \mathbf{n}_t | sinal resultante de $[C/A]e_t$ |
| \mathbf{s}^2 | variância do sinal e_t |
| \mathbf{k} | índice de desempenho dado pela razão entre \mathbf{s}_{MV}^2 e \mathbf{s}_y^2 |
| \mathbf{h} | índice de desempenho dado por $1 - \mathbf{k}$ |
| m_j | desvio médio do <i>setpoint</i> |
| \mathbf{a} | parâmetros estimados por mínimos quadrados |
| t_I | fator integral de um controlador PI/PID |
| t_i | constante de tempo característica de um sistema |
| t_{res} | tempo de residência |

Abreviaturas

| | |
|---------------|---|
| <i>ACF</i> | função de autocorrelação |
| <i>AGENER</i> | abordagem generalizada para cálculo do índice de desempenho |
| <i>ANORM</i> | abordagem normalizada para cálculo do índice de desempenho |

| | |
|--------------|--|
| <i>AR</i> | modelo autorregressivo |
| <i>ARIMA</i> | modelo autorregressivo, integrativo, com média móvel |
| <i>ARMA</i> | modelo autorregressivo com média móvel |
| <i>ARMAX</i> | modelo autorregressivo com média móvel e sinal exógeno |
| <i>ARX</i> | modelo autorregressivo com sinal exógeno |
| <i>ASUBS</i> | abordagem baseada em identificação no subespaço de estado para cálculo do índice de desempenho |
| <i>CLP</i> | controlador lógico programável |
| <i>CVM</i> | controlador de variância mínima ou controle com variância mínima |
| <i>FCOR</i> | abordagem baseada em filtragem e correlação para cálculo do índice de desempenho |
| <i>FIR</i> | resposta ao impulso finito |
| <i>IAE</i> | integral do erro absoluto |
| <i>IMC</i> | controlador de modelo interno |
| <i>LLT</i> | algoritmo de linearizações ao longo da trajetória |
| <i>MA</i> | modelo de média móvel |
| <i>MIMO</i> | sistema com múltiplas entradas e múltiplas saídas |
| <i>MPC</i> | controlador preditivo multivariável |
| <i>NMPC</i> | controlador preditivo não linear multivariável |
| <i>PCA</i> | análise de componentes principais |
| <i>PI</i> | controlador proporcional e integral |
| <i>PID</i> | controlador proporcional, integral e derivativo |
| <i>RMF</i> | matriz fração do lado direito |
| <i>RVI</i> | índice de desempenho baseado em variância relativa |
| <i>SDCD</i> | sistema digital de controle distribuído |
| <i>SISO</i> | sistema com uma entrada e uma saída |
| <i>SVD</i> | decomposição em valores singulares |

Capítulo 1 Introdução

Avaliação do desempenho regulatório de controladores é uma área relativamente nova no meio acadêmico e pouco conhecida no ambiente industrial. No entanto é um tópico bastante promissor e proporciona muitos desafios em termos de pesquisa e aplicação industrial. De forma alguma é uma tecnologia consolidada, como demonstram as dezenas de publicações anuais com novas propostas e diferentes abordagens para o problema do monitoramento de qualidade de malhas de controle. Ainda assim, sua aplicação, mesmo na forma mais básica, em indústrias de vanguarda tecnológica, demonstra ganhos de tal porte que estimulam o investimento nestas técnicas por mais simples que possam ser.

1.1. Relevância do tema desta dissertação

A maioria dos processos industriais utiliza dezenas, centenas ou até milhares de malhas de controle do tipo *feedback*. Com o passar dos anos, mesmo as malhas corretamente sintonizadas apresentam uma deterioração de seu desempenho devido a alterações naturais e contínuas na dinâmica dos processos. Surge assim a necessidade de manutenção de sistemas de controle. Como esta tarefa é árdua e cara (tão mais cara quanto maior o número de malhas), é necessário saber identificar, de um modo simples e rápido, quais malhas não estão desempenhando seu papel adequadamente.

Várias técnicas foram desenvolvidas com a finalidade de monitorar o desempenho de controladores através da análise não-invasiva de dados de processo. Apesar destas técnicas estarem bem desenvolvidas, na grande maioria das plantas não há um sistema de monitoramento do desempenho de malhas de controle (Harris et al. 1999).

Junto com o desenvolvimento teórico do assunto também surgem ferramentas de software que se propõe a implementar estas técnicas e permitir sua aplicação no ambiente industrial. Ferramentas deste tipo já estão bastante desenvolvidas no mercado, contando com uma gama significativa de recursos e tecnologia agregados e com o suporte de pesquisadores de renome na área de avaliação de desempenho de malhas de controle. Uma ferramenta deste tipo tem potencial de aplicação em qualquer processo industrial automatizado, desde processos de pequeno porte até grandes indústrias de *commodities* tais como a indústria química e petroquímica.

O uso de tal ferramenta não requer complicados sistemas de implementação; é totalmente independente do processo em si por ser não invasiva (ou pelo menos permite que se escolha diferentes níveis de invasão, dependendo da necessidade do usuário); produz resultados imediatos e alega resultar em benefícios como o aumento de produtividade, aumento de qualidade do produto final, aumento da vida útil de equipamentos, redução dos custos operacionais e de manutenção. Enfim, todas as vantagens que são objetivo final do processo produtivo, podem ser influenciados pela implementação desta tecnologia.

Cabe ressaltar que em uma aplicação real de métodos de redução de variabilidade em fornos de pirólise foram relatados ganhos de US\$ 4 milhões/ano (Emerson, 2002). Resultados de diversas auditorias mostram que apenas 20% das malhas de controle reduzem as variabilidades, 30% oscilam e aumentam as variabilidades em função da sintonia, 30% oscilam e aumentam as variabilidades em função do uso de transmissores e válvulas de controle de desempenho questionável, 15% requerem novas estratégias de controle e 5% requerem novo projeto (Ender, 1993).

Uma nova área, ainda pouco explorada pela comunidade científica é a avaliação de desempenho de controladores preditivos multivariáveis (MPC), que é uma estratégia de controle em crescente popularidade no âmbito industrial, principalmente devido à evolução da

capacidade de processamento computacional instalado nas grandes indústrias. A avaliação de desempenho de MPC's apresenta grandes lacunas no campo teórico a serem preenchidas.

Todos estes fatores e situações analisados neste capítulo já são provas de que é mais que necessário o investimento em pesquisa, desenvolvimento, implementação e suporte industrial desta tecnologia, justificando assim a confecção desta dissertação.

1.2. Revisão histórica

Avaliação de desempenho de controladores é um assunto que sempre esteve presente desde que o controle automático de processos químicos começou a se tornar uma exigência na indústria.

Técnicas bastante rudimentares de avaliação de desempenho quanto as suas características determinísticas sempre foram bem dominadas. Análises visuais ou comparativas das respostas dos controladores relativas a degraus ou pulsos na entrada permitem conclusões subjetivas ou pouco quantitativas a respeito do desempenho do controlador na sua atuação servo. A observação do tempo de subida, tempo de assentamento, *overshoot*, integral do erro quadrático entre outras características sempre permitem que o engenheiro de processo formule conclusões a respeito da ação do controlador. No entanto, estas técnicas não são de prática implementação para monitoramento *on-line* do processo em operação rotineira, pois requerem a introdução de distúrbios no processo e/ou operação em malha aberta. Adicionalmente, a grande quantidade de malhas de controle em uma unidade industrial de porte tornariam a tarefa de monitoramento do desempenho um trabalho árduo e caro, pois a impossibilidade de automatização destas técnicas requereria a utilização de recursos humanos em tempo integral nesta função. Além disto, estas técnicas não são capazes de fornecer explicitamente alguma informação sobre o desempenho em malha fechada sob mudanças aleatórias, e na maioria das vezes, não fornecem um índice de avaliação que seja apropriado para implementação *on-line*.

Na direção de avaliação automática do desempenho de malhas de controle com função regulatória o primeiro estudo publicado data de 1967 (Åström, K. "Computer control of a paper machine – an application of linear stochastic control theory", IBM Journal, July, 1967). No entanto, a base para a formulação de metodologias para avaliação de desempenho de malhas de controle foi definida em 1970 através da contribuição de Karl J. Åström. Em seu livro *Introduction to Stochastic Control Theory* (Åström, 1970), são apresentadas formulações para a predição ótima de processos estacionários lineares discretos, da qual a teoria de controle com variância mínima é então derivada.

Poucos trabalhos foram desenvolvidos na área nos anos subseqüentes. No fim dos anos 80 surgem várias ferramentas que formariam um novo campo, chamado Controle Estatístico de Processos, com o objetivo de monitorar a variabilidade de processos e detectar a

presença e fontes de distúrbios. Estas técnicas, no entanto, não avaliam a qualidade das saídas em função das ações de controle.

É natural que, quando se tenta quantificar desempenho, compare-se a avaliação feita contra algum tipo de referência. É assim que se mede, por exemplo, desempenho escolar (esta forma embora consagrada, é questionável), qualificação de candidatos a emprego, beleza e demais características humanas, produtos no mercado (onde a referência é a concorrência), etc. Seguindo esta linha, surgiu a proposta feita por Harris (1989), que usou o controle com variância mínima como referência para comparar o desempenho de controladores. A idéia foi criar um índice que quantificasse o desempenho do controlador em relação ao mínimo teórico possível de variância. Basicamente o índice consiste na razão entre a variância na saída que um controlador de variância mínima produziria e a variância real da saída da malha em questão. Este trabalho conseguiu grande destaque por mostrar como as técnicas estudadas por Åström somadas com simples análises de séries temporais poderiam ser utilizadas para encontrar uma expressão apropriada para a variância do controlador de variância mínima (CVM) a partir de dados rotineiros de processos SISO. Sua contribuição foi significativa por definir uma nova direção para a área de monitoramento de desempenho de malhas de controle.

Dando continuidade a este trabalho, Desborough e Harris (1992) propuseram a utilização de regressão linear simples para calcular uma estimativa da variância mínima de forma automatizada, que não havia sido desenvolvida no trabalho anterior. Também um procedimento recursivo, apropriado para cálculo *on-line* do índice foi apresentado. Uma metodologia automática para monitoramento da qualidade de malhas deve ser simples o suficiente para não requerer intervenções humanas e deve ser não invasiva. Isto significa que não deve promover alterações no sistema onde o monitoramento é introduzido e deve se utilizar da mínima informação possível, de preferência apenas a informação disponível no sistema digital de controle. Este é o grande mérito da metodologia proposta por Desborough e Harris, que requer apenas dados de processo rotineiro em malha fechada ou aberta e o conhecimento prévio do tempo morto.

A partir destes trabalhos, diversos pesquisadores começaram a explorar os mais diversos aspectos da avaliação de desempenho baseada em controladores com variância mínima. Tyler e Morari (1995) estenderam a idéia a processos SISO de fase não mínima. Tyler e Morari (1996) propõem o uso de diversas técnicas baseadas em verosimilhança para avaliação de desempenho. Tais técnicas têm seu devido valor, mas não são apropriadas para monitoramento automático ou em larga escala. Huang e Shah (1996) desenvolveram um método baseado em filtragem e correlação (FCOR) para estimar o valor da variância mínima e assim calcular o índice de desempenho. Em Harris et. al. (1996) e Huang et. al. (1996) a metodologia SISO original é estendida à sistemas multivariáveis. Huang et al. (1997b) estendem o algoritmo FCOR para o caso multivariável.

Paralelamente, novas ferramentas ou modificações nos procedimentos de avaliação de malhas foram introduzidas, tais como procedimentos de detecção de oscilações (Hägglund, 1995; Thornhill e Hägglund, 1997; Hägglund, 2000); seleção de pólos do CVM para desempenho desejado (Horch e Isaksson, 1999) e detecção de desempenho lento de

controladores (Hägglund, 1999, 2000). Em Bezergianni e Georgakis (2000) e Bezergianni (2001) é apresentada uma modificação do índice de desempenho, chamado índice de variância relativo. O método utiliza identificação em malha fechada no subespaço de estado para determinar modelos para controlador, processo e perturbações para calcular um índice que leva em conta também a variância em malha aberta.

Sistemas que utilizam técnicas *feedforward* / *feedback* são inicialmente abordados em Desborough e Harris (1993). Huang e Shah (1999, 2000) demonstram a utilização do algoritmo FCOR para esta estrutura de controle. A idéia é avaliar o desempenho de malhas que empregam simultaneamente as duas estratégias de controle, ou então, quantificar os benefícios de uma hipotética inclusão de um componente *feedforward* à um sistema *feedback*.

Embora o tempo morto exerça um papel importante na estimativa da variância mínima, somente em 1996 começou-se a realizar estudos para estimativa do tempo morto a partir de dados em malha fechada. Lynch e Dumont (1996) discutem o uso de regressores variáveis de modelos fixos para estimar o tempo morto do processo. Huang et al. (1997a) propõem a estimação do tempo morto a partir da obtenção dos primeiros parâmetros de Markov do processo com dados de malha fechada e a aplicação da metodologia proposta por Rogozinski et al. (1987) e Peng e Kinnaert (1992).

Ko e Edgar (2000) são os primeiros a abordar com maior profundidade a avaliação de desempenho de malhas em cascata. Wan e Huang (2002) propõem a avaliação de desempenho de malhas levando em conta a robustez do controlador. Gimble (2002) propõe o uso da variância mínima generalizada (GMV). A variância mínima gera ações de controle muito violentas, não podendo ser utilizada na prática, o que faz com que a comparação com a saída do processo seja desleal. O GMV utiliza um critério de variância mínima no qual a ação de controle é ponderada, fazendo com que se torne um *benchmark* mais realístico, segundo o autor. Tentando fugir da necessidade de se utilizar o controle com variância mínima como referência de desempenho, Thyagarajan et al. (2003) propõem a utilização de perturbações do tipo relé para formular um procedimento de avaliação de desempenho e ao mesmo tempo propor os parâmetros de reajuste da malha se necessário. Esta técnica além de requerer perturbações no processo em malha fechada, não é de fácil automação, pois requer a resolução de expressões analíticas e comparações gráficas para obtenção dos resultados.

No âmbito do emprego prático de metodologias de avaliação de desempenho, diversos trabalhos relatam experiências de implementação destas técnicas em larga escala em sistemas industriais. Harris et al. (1996b) relatam o desenvolvimento de um sistema especialista que automatiza o processo de coleta de dados, avaliação de desempenho e demonstração de resultados, com a implementação deste sistema em uma unidade industrial.

No entanto, metodologias de avaliação direta de sistemas MPC já são encontradas na literatura em estágio inicial. Huang e Shah (1999) propõem uma metodologia baseada na minimização da função-objetivo do controlador, chamada LQG Benchmark. Patwardhan e Shah (2002) abordam diversos aspectos práticos da auditoria de desempenho de MPC's.

No Brasil estas auditorias de desempenho começaram a ser realizadas somente a partir de 1998, nos setores de papel e celulose, químico e petroquímico, siderúrgico e de açúcar e álcool (Tecnologia e Negócios nº18). No entanto, no meio acadêmico não se tem conhecimento até a presente data de algum estudo ou desenvolvimento produzido nesta área.

1.3. Produtos de software comerciais existentes

Foram encontrados na bibliografia e em pesquisas na Internet apenas duas empresas que já se propuseram a transformar a tecnologia e o conhecimento desenvolvidos em produtos de software e serviços de suporte e que estão no mercado.

A Control Arts Inc. (CONTROLARTS, www.controlartsinc.com) alega possuir um sistema de avaliação de desempenho de controladores chamado Control Monitor que permite monitorar todas as malhas do sistema distribuído de controle, diagnosticando problemas e reajustando controladores. Aparentemente permite o monitoramento agendado da malhas e o monitoramento detalhado de alguma malha de interesse maior. Avalia tanto controladores PID SISO quanto controladores multivariáveis preditivos, mas não é informado se existe análise recursiva ou se ela é só janela móvel.

Já a Matrikon Solutions for Industrial Agility (MATRIKON, www.matrikon.com) fornece bastante detalhes a respeito do seu software, chamado ProcessDoctor. Este software permite cálculos personalizados; integração com Matlab®, Visual Basic/C++/Java; conta com suporte de conexão a diversos sistemas historiadores comerciais tais como os da Aspen, Honeywell e OSI. Suas interfaces gráficas são bastante amigáveis, permitindo a geração de relatórios detalhados de avaliação de malhas. Dentre a tecnologia embutida no produto, estão: índice de desempenho tradicional e relativo, recursivos e não-recursivos, SISO e MIMO, índices de eficiência operacional (índices de serviço), medidas da saturação do controlador e índices baseados em técnicas de malha aberta, como tempo de subida e resposta ao pulso. Também conta com ferramentas estatísticas clássicas para auxiliar na análise, tais como seis sigma, análise de oscilações, respostas no domínio da frequência, análise espectral e de correlação cruzada. A empresa ainda conta com o suporte de dois pesquisadores de renome na área de controle de processos, com ampla gama de trabalhos publicados sobre o assunto.

A Honeywell (HONEYWELL) também entrou no mercado com um produto de monitoramento global de desempenho de unidades industriais. A grande inovação deste produto é a possibilidade de acesso aos dados do processo, relatórios e resultados através da Internet. No entanto, a coleta de dados automática era limitada a sistemas fabricados pela Honeywell apenas. Isto pesa decisivamente para a não adoção de tal ferramenta em larga escala, como ocorreu na Eastman Chemical Company em 1999 (Paulonis e Cox, 2003).

1.4. Motivação

Desta dissertação espera-se que resulte um estudo bastante abrangente e didático da área de avaliação de desempenho de malhas de controle. Nisto estão implícitos o estudo e a aplicação das metodologias para avaliação da qualidade do controle. Desta maneira, surge a análise crítica das diversas proposições existentes na bibliografia, desde suas formas mais básicas e simples até as mais elaboradas.

Ao longo da pesquisa desenvolvida não foi encontrada nenhuma ocorrência de trabalhos publicados sobre este assunto no âmbito do meio acadêmico nacional. Espera-se assim, com esta dissertação, abrir caminho para futuros desenvolvimentos ao apontar tópicos a serem explorados (pois este trabalho, por ser abrangente, não tem este foco) e desta forma, motivar a comunidade científica nacional se interessar pela pesquisa nesta área. Diversas lacunas ainda devem ser preenchidas tendo como objetivo a obtenção de informações de maior valor e mais precisas, e serão apontadas no decorrer dessa dissertação.

É dada atenção especial para os métodos de avaliação baseados no controle com variância mínima (CVM), visto que são os mais interessantes tanto em termos de implementação prática tanto quanto no sólido embasamento teórico, permitindo por isto mesmo, um maior desenvolvimento de idéias novas na área. Também são procedimentos relativamente fáceis de serem implementados nos sistemas digitais de controle. Serão discutidas e comparadas, tanto teoricamente quanto em termos práticos, as técnicas de auditoria de malhas mais difundidas. Na literatura não encontra-se uma comparação desta natureza.

Nesta dissertação será abordada com mais ênfase apenas a questão da avaliação de controle de malhas na sua função regulatória. Para processos que operam em vários pontos estacionários em campanhas rápidas, esta abordagem pode ser relativamente limitada. No entanto, para a maioria dos processos industriais, o controle regulatório é a condição normal de operação. Para processos multipropósito, as técnicas de avaliação de desempenho servo encontram-se bem dominadas e difundidas no meio industrial, não sendo abordadas neste trabalho. Outra observação importante, é que as técnicas estudadas são baseadas em teoria de controle linear. Isto traz consigo todas as implicações e problemas clássicos de se considerar processos não-lineares como aproximadamente lineares.

A variabilidade total do produto é o resultado cumulativo da variância inerente às diversas etapas da produção. Atacar os pontos onde ela pode ser reduzida é subtrair parcelas da variabilidade do produto final. Os controladores de processo têm como um de seus papéis a atenuação das oscilações de baixa frequência existentes, que é geralmente um dos principais fatores causadores de variabilidade. Seu mau funcionamento faz com que ao contrário de se obter ganhos com seu uso, tenha-se perdas. As vantagens de se conseguir diminuir a variabilidade de um processo podem ser ilustradas através da análise de um sistema hipotético em que a temperatura média de um forno é fundamental na qualidade do produto final. Quanto mais elevada a temperatura de operação, maior é a qualidade do produto. Por questões de resistência do material do forno, a temperatura máxima é limitada em T_{lim} , ou seja, em

nenhuma hipótese a temperatura deve ultrapassar o limite. Ao conseguir-se reduzir a variância da temperatura do forno, pode-se elevar a temperatura de operação sem que se ultrapasse T_{lim} , obtendo-se um produto final hipotético de maior qualidade com esta elevação da temperatura de operação. Este resultado pode ser ilustrado na Figura 1.1, com $T_{lim} = 560^{\circ}\text{C}$.

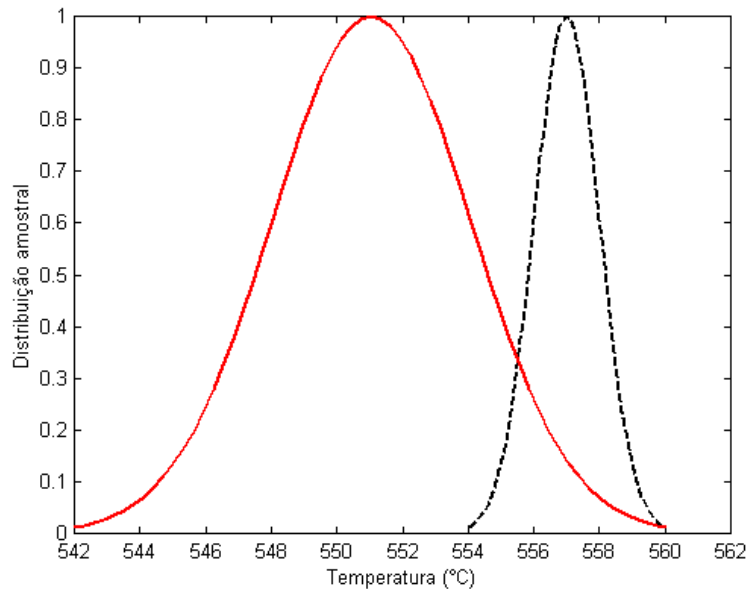


Figura 1.1. Distribuições amostrais da temperatura do forno antes (linha cheia) e após (linha pontilhada) a redução da variabilidade de malha de controle e da mudança de ponto de operação.

Neste exemplo hipotético, a temperatura média de operação pôde ser elevada de 551°C para 557°C sem acarretar prejuízos à estrutura do forno, com as conseqüentes melhorias na qualidade do produto.

Este é um exemplo típico de uma situação onde uma a variabilidade de uma variável tem impacto significativo na qualidade do produto final. Esta é uma situação que pode nem sempre ocorrer na prática. Saber identificar quais variáveis podem produzir melhores resultados se tiverem sua variância reduzida é uma tarefa para o engenheiro de processo. Na prática, a redução de variabilidade no processo é sempre recomendável, uma vez que, se não resulta em maior qualidade do produto final, pode resultar em redução de custos de matéria-prima ou de manutenção do equipamento por exemplo. Uma exceção à regra é a análise do desempenho de malhas de controle em que o objetivo é manter um outra variável, dependente da variável controlada em questão, dentro da especificação. Assim, a variável controlada é permitida variar livremente dentro de uma faixa ampla para que a variável dependente mantenha a especificação. Neste caso, a redução da variabilidade da malha de controle não é o objetivo direto.

Adicionalmente esta dissertação propõe alternativas para as limitações ainda existentes nas metodologias mais comumente abordadas em estudos e empregadas em indústrias, bem como oferece novas contribuições e caminhos para esta área.

1.5. Estrutura da dissertação

Esta dissertação encontra-se dividida em 6 capítulos. O Capítulo 2 apresenta uma extensiva revisão bibliográfica onde são desenvolvidas as bases teóricas da avaliação de desempenho de controladores SISO baseada em variância mínima. Outros tópicos e diferentes abordagens na área de avaliação de desempenho também são tratadas nesse capítulo.

O Capítulo 3 consiste de uma aplicação teórico/prática dos índices de avaliação de desempenho para malhas SISO. Nele são inicialmente tratadas questões pertinentes a aplicação prática dos procedimentos apresentados na revisão bibliográfica. Logo em seguida é feito o uso de simulações de diversos sistemas SISO com o principal intuito de verificar o comportamento de controladores de variância mínima, compará-los com controladores do tipo PID e por fim verificar o grau de confiabilidade dos índices de desempenho estimados pelos diversos algoritmos implementados. Com estas simulações é, enfim, possível verificar todos aspectos teóricos e práticos com detalhes. Na última parte do capítulo é feita a aplicação das diversas técnicas de avaliação da qualidade de malhas em dados provenientes de unidades industriais reais, de onde mais conclusões úteis quanto a utilização prática dos métodos podem ser obtidas.

O Capítulo 4 trata da extensão das principais abordagens de avaliação de desempenho para o caso multivariável. Nisto encontra-se incluída também a extensão do conceito de tempo morto para sistemas MIMO e uma metodologia de estimação do tempo morto multivariável em malha fechada é apresentada. Os algoritmos são testados com simulações onde verifica-se a validade de suas estimações através da comparação com resultados teóricos. Por último, as técnicas apresentadas são empregadas na avaliação de controladores em uma unidade experimental. Os resultados são interpretados e conclusões acerca da extensão do caso SISO são formuladas.

O Capítulo 5 propõe um conjunto de testes de aplicação sistemática das metodologias apresentadas no Capítulo 4 em uma unidade experimental. Todas as características da unidade que envolvem sua dinâmica, instrumentação e estrutura de controle são estudadas, para que os resultados obtidos possam ser interpretados. Os índices de desempenho são estimados frente a diferentes condições de operação e comparações de desempenho são feitas entre estas diferentes condições e também entre diferentes sistemas de controle: PID descentralizados e controlador preditivo multivariável.

Finalmente, no Capítulo 6 são sumariadas as principais conclusões deste trabalho, bem como são apontados alguns caminhos a serem seguidos em trabalhos futuros.

Os apêndices encontram-se divididos da seguinte maneira: o Apêndice A apresenta o algoritmo desenvolvido por Rogozinski et al. (1987) para a fatoração da matriz de interação a partir dos coeficientes do numerador de uma função de transferência; o Apêndice B apresenta um procedimento que pode ser empregado para estimação da matriz de interação utilizando dados de processo em malha fechada; o Apêndice C apresenta um algoritmo apropriado para identificação de parâmetros de Markov de um processo em malha fechada, com os quais

utilizando-se as técnicas apresentadas no Apêndice A e no Apêndice B, pode-se completar a identificação da matriz de interação da maneira mais apropriada; no Apêndice D são apresentadas figuras que complementam a análise das malhas de controle da unidade experimental testada no Capítulo 5; por fim, no Apêndice E encontra-se descrito um conjunto de funções para lidar com polinômios, desenvolvidas na forma de uma classe no Matlab®, que ajudaram na execução de grande parte dos testes realizados neste trabalho.

Capítulo 2 Revisão Bibliográfica

Neste capítulo são detalhados os principais métodos de avaliação e monitoramento de desempenho de controladores e de qualidade de malhas para o caso SISO sob ação de perturbações aleatórias. O caso MIMO será abordado em detalhes em um capítulo a parte. Será apresentada a teoria em que se baseiam as abordagens propostas na literatura, ou seja, teoria de predição ótima e controladores de variância mínima, bem como o desenvolvimento teórico dos índices de avaliação de desempenho. Também serão apresentados os procedimentos de detecção de oscilações e os índices de eficiência operacional, que são alguns dos métodos complementares que fornecem informações úteis a respeito do processo em diferentes etapas do procedimento global de avaliação.

Nesta seção, por questões de simplicidade, não será feita diferenciação entre os operadores q e q^{-1} e as variáveis representativas da transformada Z, z e z^{-1} .

2.1. Introdução

A introdução dos sistemas digitais de controle na indústria a partir do meio dos anos 70 e o aumento da capacidade de processamento computacional fez com que a facilidade de obtenção de informação do processo em tempo real se tornasse uma tarefa trivial e a adoção de estratégias avançadas de controle melhorasse muito a qualidade do produto final.

No entanto, pode-se verificar que muito do potencial que a tecnologia disponibiliza atualmente é sub-utilizada em grande parte da indústria. Ender (1993) aponta uma série de problemas que vão desde a não observância de regras básicas da teoria de controle até problemas de manutenção de válvulas e sensores, passando até mesmo pela sobrecarga de tarefas da equipe de controle.

Assim sendo, é de se esperar que muitas unidades industriais ainda apresentem grande potencial de otimização através da simples melhoria das condições das malhas de controle. Desta forma é importante primeiramente apontar as possíveis causas de baixo desempenho e em seguida formular as demandas para um sistema de auditoria de desempenho que aponte malhas problemáticas e possibilite o seu diagnóstico.

Harris et al. (1999) aponta as principais causas de desempenho insatisfatório:

- Limitações devido à presença de distúrbios, limites nas variáveis manipuladas, tempo morto e fatores de fase não-mínima;
- Ajuste impróprio do controlador. Isto pode ocorrer por motivos variados, tais como inadequação do projeto ou alterações na dinâmica do processo ou na natureza das perturbações.
- Problemas de mal funcionamento dos sistemas mecânicos e eletrônicos, devidos a desgastes de material ou falhas nos sistemas de aquisição de dados entre outros.

Segundo Hugo (2002), a técnica de monitoramento de desempenho deve possuir os seguintes atributos:

- I. Deve ser independente das perturbações e do *setpoint*;
- II. Não deve requerer testes na planta;
- III. Deve ser passível de implementação automática;
- IV. Deve requerer mínima especificação da dinâmica do processo;

- V. A métrica deve fornecer uma medida não arbitrária ou absoluta, ou seja deve comparar a qualidade do processo com algum padrão absoluto;
- VI. Deve ser sensível apenas ao desajuste do controlador, incertezas de modelo ou problemas de equipamento;
- VII. Deve fornecer um indicativo do porquê de o controlador estar desempenhando insatisfatoriamente;

Já em Harris et al. (1999) pode-se encontrar requisitos que complementam a lista anterior, apontando que um sistema de monitoramento do controle e avaliação de desempenho deve ter as seguintes propriedades:

- i. Operação automática, incluindo a coleta remota de dados da malha e verificação da integridade dos dados;
- ii. Embasamento teórico;
- iii. Suporte à decisões;
- iv. Suporte técnico;
- v. Interface gráfica com o usuário apropriada;

Tendo consciência destes requerimentos pode-se começar a explorar o estágio atual de desenvolvimento das técnicas existentes e verificar se elas satisfazem os requerimentos I. a VII. e i. a iv. Os pontos i e v. dependem exclusivamente de engenharia de *software* e não entra no escopo desta dissertação.

Com esta finalidade e seguindo estas diretrizes, será apresentado o embasamento teórico das técnicas mais empregadas na avaliação de desempenho, que é a Teoria de Predição Ótima na qual se fundamenta a Teoria de Controle com Variância Mínima. Após serão apresentados detalhes da implementação destas técnicas.

2.2. Teoria de predição ótima

A teoria de controle com variância mínima deriva da teoria de predição ótima. Desta forma, todas as considerações acerca do desenvolvimento matemático do controle com variância mínima serão feitas neste capítulo.

Primeiramente, é assumido que o processo a ser controlado é linear e invariante com o tempo, e que o processo possui uma entrada (u) e uma saída (y), ou seja, o processo é SISO. A

dinâmica do processo é caracterizada por uma combinação de um tempo morto e uma função de transferência racional. As perturbações são assumidas como sendo ruído branco filtrado.

A dinâmica do processo é dada então pela seguinte expressão:

$$x(k) = \frac{B_1(q)}{A_1(q)} u(k) \quad (2.1)$$

Onde A_1 e B_1 são polinômios no operador avanço q .

A influência externa do ambiente pode ser caracterizada por distúrbios que são processos aleatórios. Pelo fato de o sistema ser linear, pode-se utilizar o princípio da superposição para reduzir todos os distúrbios a um distúrbio equivalente, v , na saída do sistema.

$$y(k) = x(k) + v(k) \quad (2.2)$$

O distúrbio v pode ser representado como a saída de um sistema linear excitado por ruído branco:

$$v(k) = \frac{C_1(q)}{A_2(q)} e(k) \quad (2.3)$$

Onde $C_1(q)$ e $A_2(q)$ são polinômios no operador avanço q , e $e(k)$ é uma seqüência de valores aleatórios independentes (não correlacionados) com média zero e desvio padrão σ . O distúrbio v pode ser um processo aleatório estacionário ou um processo flutuante, uma vez que o polinômio A_2 pode ser instável¹.

O modelo pode ser reduzido a uma forma padrão, substituindo (2.1) e (2.3) em (2.2) para se obter:

$$A(q) y(t) = B(q) u(t) + C(q) e(t) \quad (2.4)$$

Onde: $A = A_1 A_2$

$$B = B_1 A_2$$

$$C = C_1 A_1$$

O modelo em (2.4) é a forma canônica na qual se basearão os desenvolvimentos neste trabalho.

¹ Por polinômio instável entende-se o polinômio com raízes fora do círculo unitário.

A equação (2.4) pode ser normalizada, de forma que os primeiros coeficientes de A e C sejam unitários. Tais polinômios são chamados monômicos. O polinômio C também pode ser multiplicado por uma potência arbitrária de q que isto não altera a estrutura de correlação de $C(q)e(t)$ (Åström & Wittenmark, 1997). Com isto, pode-se também normalizar o grau de C de forma a igualar-se ao grau de A . Os polinômios $A(q)$ e $B(q)$ podem ter zeros dentro ou fora do círculo unitário. É assumido que todos os zeros do polinômio $C(q)$ estão dentro do círculo unitário. Se alguma identificação de sistemas resultar num polinômio $C(q)$ instável, este pode ser alterado por fatoração espectral de forma a que todos seus zeros resultem dentro do círculo unitário. Isto é conseguido fatorando-se C da seguinte forma:

$$C = C^+ C^- \quad (2.5)$$

Onde C^+ contém as raízes estáveis (fase mínima) e C^- contém as raízes instáveis (fase não mínima). Assim, pode-se substituir C por $C^+ C^{-*}$, onde C^{-*} é o recíproco de C^- , isto é: $C^{-*}(q^{-1}) = q^{-a} C^-(q)$ e a é a ordem de C . $C^+ C^{-*}$ é o polinômio equivalente a C no sentido que não altera suas propriedades espectrais.

No estado estacionário, é conveniente expressar o critério de otimalidade em termos das variâncias em estado estacionário da variável controlada e da saída do processo. Para controle regulatório de sistemas SISO, o critério pode ser escolhido como sendo minimizar a variância da saída, o que leva à:

$$J_{MV} = \mathbf{E}y^2(t) \quad (2.6)$$

Onde é assumido que o ponto de operação desejado é $y = 0$.

As propriedades do sinal de controle sob controle com variância mínima dependem criticamente do tempo de amostragem. Tempos de amostragem curtos resultam em grande variância enquanto que longos períodos de amostragem resultam em pequena variância (Åström & Wittenmark, 1997).

A teoria de predição implica então, em que o preditor ótimo é aquele que minimiza a média do erro de predição ao quadrado.

Considerando um sinal y gerado por um modelo do tipo ARMA:

$$y(k) = \frac{C(q)}{A(q)} e(k) = \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k) \quad (2.7)$$

onde A^* e C^* são os polinômios recíprocos de A e C , e q^{-1} é o operador atraso no tempo. É conveniente que a análise seja feita em termos deste operador porque esta é baseada em causalidade. C e A são de ordem a .

No tempo k a situação é a seguinte: as variáveis $y(k)$, $y(k-1)$, ... foram observadas e deseja-se prever $y(k+m)$. Uma expansão em séries de C^*/A^* em q^{-1} resulta em:

$$\begin{aligned} y(k+m) &= \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k+m) \\ &= \underbrace{e(k+m) + p_1 e(k+m-1) + \dots + p_{m-1} e(k+1)}_{\text{Desconhecido no tempo } k} + \\ &\quad + \underbrace{p_m e(k) + p_{m+1} e(k-1) + \dots}_{\text{Conhecido no tempo } k} \end{aligned} \quad (2.8)$$

onde os termos p_i são equivalentes aos coeficientes da resposta ao impulso do sistema representado por C^*/A^* .

Os dois termos do lado direito de (2.8) são independentes entre si, pois $e(k)$ é uma seqüência de variáveis aleatórias independentes (Åström & Wittenmark, 1997). Os primeiros $m-1$ termos de (2.8) são independentes dos dados no tempo k . A segunda parte é função dos dados disponíveis no tempo k . Segue-se que o preditor ótimo é dado por:

$$\hat{y}(k+m | k) = p_m e(k) + p_{m+1} e(k-1) + p_{m+2} e(k-2) \dots \quad (2.9)$$

E o erro de predição é dado por:

$$\tilde{y}(k+m | k) = e(k+m) + p_1 e(k+m-1) + \dots + p_{m-1} e(k+1) \quad (2.10)$$

Os coeficientes p_i formam um polinômio P que pode ser calculado da seguinte maneira: sendo $e(k)$ uma seqüência de variáveis aleatórias independentes e $C(q)$ possuindo todos os zeros dentro do círculo unitário, a divisão de $q^{m-1}C$ por A tem como quociente um polinômio $F(q)$ cujos coeficientes são os primeiros $m-1$ coeficientes de P em (2.10) e como resto um polinômio $G(q)$ cujos coeficientes são os demais coeficientes de P em (2.9), ou seja:

$$q^{m-1}C(q) = A(q)F(q) + G(q) \quad (2.11)$$

A equação (2.11) também pode ser resolvida como uma equação de Diophantine (Åström, 1970). O termo "Diophantine" deve-se ao aritmético Diophantus de Alexandria, conhecido como o "pai da álgebra", nascido em aproximadamente 200 A.C.

O preditor de mínima variância em m passos a frente é dado por:

$$\hat{y} = \hat{y}(k+m | k) = \frac{qG(q)}{C(q)} y(k) = \frac{G^*(q^{-1})}{C^*(q^{-1})} y(k) \quad (2.12)$$

E o erro de predição, \tilde{y} , é uma média móvel do tipo:

$$\tilde{y}(k + m | k) = y(k + m) - \hat{y}(k + m | k) = F(q)e(k + 1) \quad (2.13)$$

O erro de predição tem média zero e variância dada por:

$$\mathbf{E}\tilde{y}(k + m | k)^2 = \left(1 + f_1^2 + \dots + f_{m-1}^2\right) \mathbf{S}^2 \quad (2.14)$$

A prova das equações (2.12) e (2.13) pode ser obtida executando-se os seguintes equacionamentos:

O polinômio F é monômico de grau $m - 1$ e G é de grau menor que a . Seus coeficientes são equivalentes aos $m - 1$ coeficientes p_i de (2.8).

$$\begin{aligned} F(q) &= q^{m-1} + f_1 q^{m-2} + \dots + f_{m-1} \\ G(q) &= g_0 q^{a-1} + g_1 q^{a-2} + \dots + g_{a-1} \end{aligned} \quad (2.15)$$

Equivalentemente:

$$\begin{aligned} F^*(q^{-1}) &= 1 + f_1 q^{-1} + \dots + f_{m-1} q^{-m+1} \\ G^*(q^{-1}) &= g_0 + g_1 q^{-1} + \dots + g_{a-1} q^{-a+1} \end{aligned} \quad (2.16)$$

Assim, a equação de Diophantine (2.11) fica:

$$C^*(q^{-1}) = A^*(q^{-1})F^*(q^{-1}) + q^{-m}G^*(q^{-1}) \quad (2.17)$$

A equação (2.8) pode então ser escrita como:

$$y(k + m) = \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k + m) = F^*(q^{-1})e(k + m) + \frac{G^*(q^{-1})}{A^*(q^{-1})} e(k) \quad (2.18)$$

Usando a equação (2.7) o sinal $e(k)$ no último termo de (2.18) pode ser expresso em termos dos dados disponíveis no tempo k .

$$y(k + m) = F^*(q^{-1})e(k + m) + \frac{G^*(q^{-1})}{C^*(q^{-1})} y(k) \quad (2.19)$$

O primeiro termo de (2.19), por analogia com (2.8), é o erro de predição e o segundo termo é o preditor ótimo [ver equações (2.12) e (2.13)].

Sendo y uma função arbitrária de $y(k)$, $y(k - 1)$, $y(k - 2)$, \dots , que se esteja querendo prever, então, a variância do erro de predição é dada por:

$$\begin{aligned}
E(y(k+m) - \hat{y})^2 &= E(F^*(q^{-1})e(k+m))^2 + E\left(\frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) - \hat{y}\right)^2 + \\
&+ 2E\left\{F^*(q^{-1})e(k+m)\left(\frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) - \hat{y}\right)\right\}
\end{aligned} \tag{2.20}$$

O último termo em (2.20) é zero pois todos os $e(t)$ com $t > k$ têm média zero e são independentes de y (Åström, 1970). Este termo é equivalente à covariância entre os dois termos de (2.19). Desta forma, resulta:

$$E(y(k+m) - \hat{y})^2 = E(F^*(q^{-1})e(k+m))^2 + E\left(\frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) - \hat{y}\right)^2 \tag{2.21}$$

Assim sendo:

$$E(y(k+m) - \hat{y})^2 \geq E(F^*(q^{-1})e(k+m))^2 \tag{2.22}$$

A igualdade é obtida quando o segundo termo do lado direito de (2.21) é zero, ou seja:

$$\hat{y} = \hat{y}(k+m | k) = \frac{G^*(q^{-1})}{C^*(q^{-1})}y(k)$$

Que é a mesma equação que a equação (2.12) e o erro de predição é dado por (2.13).

É importante ressaltar que o preditor é linear, sendo assim, o ponto ótimo é único e que o preditor é um sistema dinâmico de ordem a cujo polinômio característico é $C(q)$.

Também pode-se interpretar $e(k)$ como sendo inovações (resíduo) do processo $y(k)$, pois de (2.13) segue-se que:

$$\tilde{y}(k+1 | k) = y(k+1) - \hat{y}(k+1 | k) = e(k+1) \tag{2.23}$$

2.3. Teoria de controle com variância mínima (CVM)

Para determinar a equação para o controlador de variância mínima será primeiro considerado o caso em que o polinômio B da equação (2.4) tem todas suas raízes dentro do círculo unitário. Isto significa que a dinâmica do processo tem uma função de transferência estável, sem resposta inversa.

Com o uso do operador q^{-1} o modelo em (2.4) pode ser escrito como:

$$y(k) = \frac{B(q)}{A(q)} u(k) + \frac{C(q)}{A(q)} e(k) = \frac{B^*(q^{-1})}{A^*(q^{-1})} q^{-d} u(k) + \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k) \quad (2.24)$$

Onde $d = \deg(A) - \deg(B) > 0$ é o excesso de pólos do sistema, ou seja, o seu tempo morto.

Assume-se, como detalhado na Teoria de Predição Ótima, que $\deg(A) = \deg(C) = a$. Segue-se, de (2.24), que:

$$\begin{aligned} y(k+d) &= \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k+d) + \frac{B^*(q^{-1})}{A^*(q^{-1})} u(k) \\ &= \underbrace{F^*(q^{-1}) e(k+d)}_{\text{Da teoria de predição ótima}} + \frac{G^*(q^{-1})}{A^*(q^{-1})} e(k) + \frac{B^*(q^{-1})}{A^*(q^{-1})} u(k) \end{aligned} \quad (2.25)$$

Onde a equação (2.17) com $m = d$ foi usada para obter (2.25) a partir de (2.24).

O primeiro termo do lado direito de (2.25) é independente dos dados disponíveis no tempo k e assim sendo, também é independente dos outros dois termos. O segundo termo pode ser computado exatamente a partir dos dados disponíveis no tempo k . Para tanto, resolve-se (2.24) para $e(k)$:

$$e(k) = \frac{A^*}{C^*} y(k) - q^{-d} \frac{B^*}{C^*} u(k) \quad (2.26)$$

Usando (2.26) e (2.17) pode-se transformar (2.25) em:

$$y(k+d) = F^* e(k+d) + \frac{G^*}{C^*} y(k) + \frac{B^* F^*}{C^*} u(k) \quad (2.27)$$

Considerando que $u(k)$ é uma função arbitrária dos seus próprios valores passados e da saída atual e seus valores passados, pode-se escrever:

$$E y^2(k+d) = E (F^* e(k+d))^2 + E \left(\frac{G^*}{C^*} y(k) + \frac{B^* F^*}{C^*} u(k) \right)^2 \quad (2.28)$$

Que pode ser escrita como:

$$s_y^2 = \text{var}(\tilde{y}) + \text{var}(\hat{y}) \quad (2.29)$$

ou seja, a soma das variâncias dos valores preditos e do erro de predição representam a variância da saída.

O termo misto resultante da transformação da equação (2.27) para a (2.28) desaparece, já que representam a covariância entre os termos do lado direito de (2.27), e $e(k+d), \dots, e(k+1)$ são independentes de $y(k), y(k-1), \dots$ e de $u(k-1), u(k), \dots$ resultando em covariância nula. Como o último termo em (2.28) é positivo, tem-se que:

$$E y^2(k+d) \geq (1 + f_1^2 + f_2^2 + \dots + f_{d-1}^2) \mathbf{s}^2 \quad (2.30)$$

A igualdade é obtida quando:

$$u(k) = - \frac{G^*(q^{-1})}{B^*(q^{-1})F^*(q^{-1})} y(k) = - \frac{G(q)}{B(q)F(q)} y(k) \quad (2.31)$$

Que é a equação para o controlador que minimiza a variância da saída.

Nota-se assim, que o equacionamento para se obter o CVM está intimamente relacionado com o problema de predição ótima. As identidades (2.11) ou (2.17) foram utilizadas em ambos os casos. Os últimos dois termos de (2.27) podem ser interpretados como a predição da saída d passos adiante. Desta forma, a estratégia de mínima variância pode ser obtida predizendo a saída d tempos de amostragem adiante e escolhendo um controlador que faça com que esta predição seja igual à saída desejada. Usando a equação (2.31) a saída resultante sob controle com variância mínima é uma média móvel de ordem $d-1$ dada por:

$$y(k) = F^*(q^{-1})d(k) \quad (2.32)$$

A equação de Diophantine para este caso pode ser resolvida por simples comparação de potências iterativamente, de posse de A, C , e d , ou através da formação de um sistema de equações lineares algébricas envolvendo uma matriz de Sylvester com os coeficientes de A e C . Em Åström (1970) e Åström e Wittenmark (1997) e na Internet (MATHWORLD, PLANETMATH) podem ser encontradas referências em como resolver equações de Diophantine na forma de (2.17). O controlador dado por (2.31) cancela todos os zeros do processo.

Da expressão para $u(k)$ pode-se mostrar que a equação característica para o sistema em malha fechada é dada por:

$$1 + \frac{G^*}{B^* F^*} \frac{B^*}{A^*} q^{-d} = 0 \quad (2.33)$$

Rearranjando os termos:

$$B^* F^* A^* + G^* B^* q^{-d} = 0 \quad (2.34)$$

Colocando B^* em evidência:

$$B^*(A^*F^* + q^{-d}G^*) = 0 \quad (2.35)$$

Nota-se que o termo entre parênteses corresponde exatamente à equação de Diophantine (2.17) para $C^*(q^{-1})$. Substituindo (2.17) em (2.35) obtém-se:

$$B^*(q^{-1})C^*(q^{-1}) = 0 \quad (2.36)$$

Assim, pode-se ver que, se C^* possui modos instáveis, a resposta será instável.

A resposta do sistema em malha fechada é dada por (2.32). Desta forma, a variância da saída será:

$$\mathbf{s}_{MV}^2 = \left(\sum_{i=1}^{d-1} f_i^2 \right) \mathbf{s}^2 \quad (2.37)$$

Em que f_i são os coeficientes de $F^*(q^{-1})$.

Enquanto que a saída do controlador é dada por:

$$u(k) = - \frac{G(q^{-1})}{B(q^{-1})} e(k) \quad (2.38)$$

Para exemplificar os resultados obtidos, segue-se um exemplo. Considerando que a identificação ARX de um sistema gerou o seguinte modelo:

$$(1 - 0.9q^{-1})y(k) = 0.5u(k - 2) + (1 + 0.7q^{-1})e(k)$$

A equação de Diophantine (2.17) fica então, para $m = 2$:

$$A = (1 - 0.9q^{-1})$$

$$B = 0.5$$

$$C = (1 + 0.7q^{-1})$$

$$(1 + 0.7q^{-1}) = (1 - 0.9q^{-1})F(q^{-1}) + q^{-2}G(q^{-1})$$

Resolvendo por comparação de potências:

$$F = 1 + 1.6q^{-1}$$

$$G = 1.44$$

A saída do sistema (2.27) é dada por:

$$y(k+2) = \frac{1.44y(k) + (0.5 + 0.8q^{-1})u(k)}{(1 + 0.7q^{-1})} + (1 + 1.6q^{-1})e(k+2)$$

O controlador de variância mínima é determinado escolhendo-se para $u(k)$ uma expressão tal que o primeiro termo do lado direito da equação da saída do sistema seja zero. Deste modo, só o que resta é o erro de previsão, que não pode ser reduzido por ser dependente de valores futuros de $e(k)$. A expressão do controlador de variância mínima é então obtida igualando-se o numerador do primeiro termo do lado direito a zero:

$$\begin{aligned} 1.44y(k) + (0.5 + 0.8q^{-1})u(k) &= 0 \\ u(k) &= -1.6u(k-1) - 2.88y(k) \end{aligned}$$

que é um caso da equação genérica dada por (2.31).

O polinômio B^* influi diretamente na estabilidade do controlador, como pode-se ver pela equação (2.38). A metodologia apresentada anteriormente aplica-se somente a processos cujos únicos zeros fora do círculo unitário estão no infinito, ou seja, tempo morto. Se a modelagem de um sistema resulta em um polinômio B^* com raízes fora do círculo unitário, o controlador conterà modos instáveis, que serão excitados pelas perturbações. Desta forma, apesar da resposta do sistema em malha fechada ser limitada, pois os modos instáveis não estão acoplados à saída, a ação de controle se torna ilimitada. Em uma simulação esta instabilidade logo causaria problemas de *overflow* e se transferiria para a saída também (Åström & Wittenmark, 1997). Em suma, a malha de controle não possuiria estabilidade interna.

Desta forma, para este tipo de sistema, é necessário que a abordagem seja diferente. $B(q)$ deve ser fatorado da seguinte forma:

$$B(q) = B^+(q)B^-(q) \quad (2.39)$$

Onde $B^+(q)$ é a parte de fase mínima e $B^-(q)$ é a parte de fase não-mínima de B . O recíproco de B^- , B^{-*} deve ser sempre monômico. Isto é possível já que na fatoração de B , B^+ pode ser multiplicado por um número e B^- pode ser dividido pelo mesmo número, de forma a que esta condição seja satisfeita. A expressão para o controlador de variância mínima fica então:

$$u(k) = \frac{-G(q)}{B^+(q)F(q)} y(k) \quad (2.40)$$

G e F são provenientes da resolução da seguinte equação de Diophantine:

$$q^{d-1}C(q)B^{-*}(q) = A(q)F(q) + B^-(q)G(q) \quad (2.41)$$

A ordem de F é igual a $\deg(B^-) + d - 1$, enquanto que a ordem de G deve ser menor que a ordem de A . O controlador dado por (2.40) cancela apenas os zeros estáveis do processo. A equação (2.41) pode ser resolvida através do cálculo do maior divisor comum para os polinômios A e B , com um algoritmo conhecido como *Algoritmo de Euclides Estendido* (Åström & Wittenmark, 1997; MATHWORLD; PLANETMATH). Este algoritmo foi desenvolvido para inteiros, mas como as operações entre inteiros e entre polinômios são similares, o algoritmo pode ser aplicado sem restrições.

A prova para o equacionamento deste caso do CVM é baseada no trabalho de Wiener sobre predição. Um método alternativo para prova pode ser encontrado em Åström & Wittenmark (1997).

A expressão para a equação característica do sistema em malha fechada, empregando-se agora as equações (2.40) e (2.41), resulta em:

$$B^+(q^{-1})B^{-*}(q^{-1})C(q^{-1}) = 0 \quad (2.42)$$

A resposta na saída neste caso é:

$$y(k) = \frac{F(q)}{B^{-*}(q)} e(k - d + 1) = \frac{F(q^{-1})}{B^{-*}(q^{-1})} q^{-2d+2} e(k) \quad (2.43)$$

ou

$$y(k) = \frac{F(q)}{B^{-*}(q)} q^{-d+1} e(k) = \frac{F(q)}{q^{d-1}B^{-*}(q)} e(k) \quad (2.44)$$

De onde se pode tirar por expansão em séries que:

$$y(k) = q^{d-1} + f_1 q^{d-2} + \dots + f_{d-1} + \frac{F_2(q)}{B^{-*}(q)} \quad (2.45)$$

onde $\deg(F_2) < \deg(B^{-*})$ e:

$$F_1 = q^{d-1} + f_1 q^{d-2} + \dots + f_{d-1} \quad (2.46)$$

é o quociente da divisão de $q^{d-1}C(q)$ por $A(q)$.

Pode-se ver que F_1 é exatamente o resultado para sistemas inversíveis, de onde a variância é calculada pela equação (2.37). Assim, para sistemas não inversíveis, a variância do sinal de saída será sempre maior quando for aplicado um controlador estável dado por (2.40) do que um controlador dado por (2.31) que neste caso seria instável. Este acréscimo na variância da saída é então devido a existência de zeros instáveis no sistema.

Pode-se sintetizar este resultado da seguinte forma: o CVM modificado resultante de processos instáveis em malha aberta possui a menor variância entre todos os controladores estáveis. A variância da saída é dada por uma média móvel de ordem $(d - 1 + n)$, onde n é o número de zeros de B fora do círculo unitário (Harris et al., 1999).

2.4. Índices de desempenho baseados em CVM

Devido à importância do monitoramento do desempenho de malhas de controle, diversos pesquisadores lançaram-se a pesquisar e desenvolver ferramentas para esta finalidade. Enquanto alguns, como Åström, concentraram seus esforços no problema servo, Huang e Shah, Harris et al., Ko e Edgar entre outros focaram seus esforços no problema regulatório.

Dentre as diversas metodologias de avaliação de malhas de controle existentes na literatura, as que encontram maior destaque e apelo prático são aquelas baseadas no controle com variância mínima. Isto se deve principalmente por sua sólida fundamentação teórica, sua fácil implementação em sistemas de controle, ser um método não invasivo e o fato de requerer pouco conhecimento de informações do processo. De fato, estes métodos utilizam dados rotineiros do processo em malha fechada, não requerendo intervenções especiais na operação normal da unidade. O único conhecimento extra sobre o processo é a duração do tempo morto, mas diversas proposições existem no sentido de se fazer estimações deste número, interpretar os resultados com valores especificados dos mesmos ou usar valores comuns a diferentes tipos de malhas (Thornhill et al., 2000).

Sabendo-se que a variância do controlador de variância mínima é a mínima variância teórica alcançável por um sistema de controle retroalimentado (*feedback*), linear, pode-se elaborar um índice adimensional que quantifique o desempenho de um controlador *feedback* linear qualquer.

Para tanto basta comparar a variância produzida pelo sistema com o uso deste controlador *feedback* com a variância mínima teórica possível para este sistema. Esta foi a idéia de Harris em seu artigo "Assessment of Control Loop Performance" (Harris, 1989). Naquele trabalho, Harris demonstrou que pode-se obter uma estimativa para a variância mínima sem a necessidade de se calcular um CVM, o que implicaria em obter o modelo exato do processo. A idéia do índice proposto é fazer a razão entre uma estimativa da variância que um controlador de variância mínima produziria e a variância real da saída do processo:

$$k(d) = \frac{\mathbf{s}_{MV}^2}{\mathbf{s}_y^2} \quad (2.47)$$

ou

$$\mathbf{h}(d) = 1 - \frac{\mathbf{s}_{MV}^2}{\mathbf{s}_y^2} \quad (2.48)$$

Onde \mathbf{s}_{MV}^2 é a estimativa da variância mínima teórica e \mathbf{s}_y^2 é a variância da saída do sistema. O índice de (2.47) é zero para processo sem controle ou 1 para controle com variância mínima, enquanto que o índice de (2.48) é o oposto.

Para se estimar a variância mínima é necessário primeiro, de posse da saída do processo e do conhecimento do tempo morto, estimar os polinômios $A(q)$ e $C(q)$ através de modelos ARMA ou AR, resolver a equação de Diophantine (2.17) por divisão longa ou comparação de potências para obter o polinômio $F(q)$ e utilizar a equação (2.37) para obter o valor da variância mínima. Desborough e Harris (1992) propõem uma metodologia (aqui referida como Abordagem Normalizada, ANORM), que utiliza uma simples regressão linear para estimar a variância mínima.

Neste trabalho foram implementados quatro algoritmos para cálculo do índice de desempenho utilizando as idéias acima mencionadas. Dois deles utilizam a metodologia original do primeiro trabalho de Harris (1989) que é a estimação dos polinômios A e C com um método de análise de séries temporais (uma referência introdutória em séries temporais pode ser encontrada em HARSHAM). O primeiro usa estimação com modelos ARMA e é aqui denominado AGENER (Abordagem Generalizada). O segundo é uma abordagem proposta pela primeira vez neste trabalho. Utiliza estimação no subespaço de estado e análise de componentes principais (Principal Component Analysis, PCA), com um algoritmo elaborado por Peter Overshee (Overshee, 1996) e é denominado aqui como ASUBS. Esta variante é proposta pela primeira vez neste trabalho. Ambos métodos resolvem uma equação de Diophantine (2.17) ou a divisão longa de C por A .

Os outros dois algoritmos implementados são o método ANORM e o método de Filtragem e Correlação (FCOR) de Huang et al. (1997), que serão detalhados na seqüência.

2.4.1. Índice de desempenho AGENER e ASUBS

Da Teoria de Predição Ótima apresentada na Seção 2.1 pode-se concluir que para a obtenção dos coeficientes que formam o polinômio F^* da equação da resposta em malha fechada sob variância mínima (2.32) basta resolver a divisão longa entre o polinômio C^* e o polinômio A^* , segundo a equação (2.18). Esta divisão tem como quociente o polinômio F^* e como resto o polinômio G^* dividido por A^* .

$$\frac{C^*(q^{-1})}{A^*(q^{-1})} = F^*(q^{-1}) + q^{-d} \frac{G^*(q^{-1})}{A^*(q^{-1})} \quad \text{ou} \quad q^{d-1} \frac{C(q)}{A(q)} = F(q) + \frac{G(q)}{A(q)} \quad (2.49)$$

Devido ao efeito do tempo morto, os primeiros $d - 1$ termos da resposta ao impulso em malha fechada são os mesmos da resposta ao impulso em malha aberta, ou seja, são equivalentes aos $d - 1$ primeiros termos da divisão longa de $C^*(q^{-1})$ por $A^*(q^{-1})$. Portanto, identificando um modelo do sistema em malha fechada do tipo:

$$y(k) = \frac{\mathbf{j}(q^{-1})}{\mathbf{f}(q^{-1})} e(k)$$

e efetuando-se a divisão longa entre os polinômios identificados pode-se obter os $d - 1$ primeiros termos equivalentes ao polinômio F^* , responsáveis pela variância mínima. Os demais coeficientes representados pelo lado direito de (2.49) são funções das funções de transferência do controlador, do processo e do distúrbio. A divisão longa em (2.49) é equivalente à resolução da equação de Diophantine (2.17).

O método AGENER consiste da simples estimativa de um modelo ARMA para os dados da saída do processo, seguida da divisão longa entre os polinômios \mathbf{j} e \mathbf{f} , ou da resolução da equação de Diophantine (2.17) envolvendo estes dois polinômios. Os modelos ARMA são obtidos resolvendo-se sistemas de equações através de mínimos quadrados. Ljung (1999) demonstra em detalhes o equacionamento e o funcionamento deste e de outros métodos de identificação.

O método ASUBS utiliza um método de identificação no subespaço de estado para gerar uma representação do sistema, que é então convertida para a forma de função de transferência. A seguir equaciona-se a divisão longa entre o numerador e o denominador para a obtenção do polinômio F . O algoritmo de identificação desenvolvido por Overschee (ver referência Overschee, 1996) tem a opção de fornecer uma estimativa para a ordem adequada do modelo estimado através da análise de componentes principais (PCA).

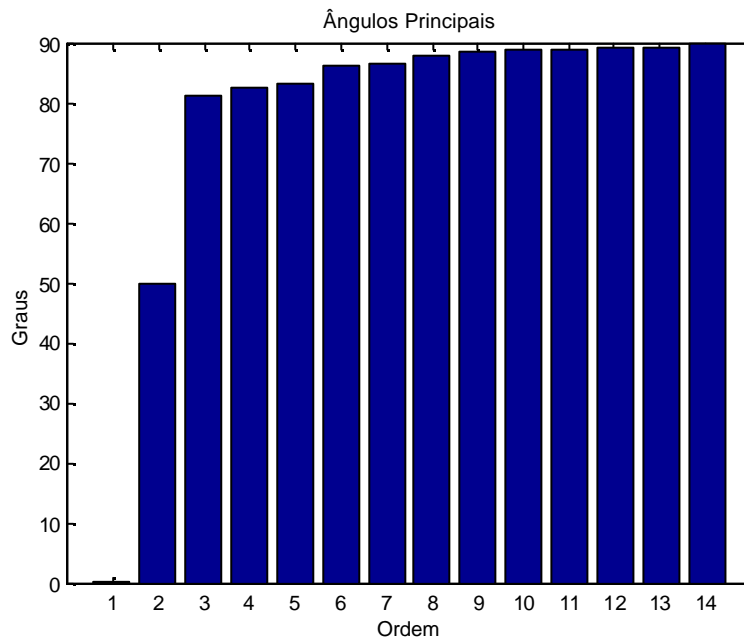


Figura 2.1. Exemplo de análise de componentes principais.

A idéia da análise de componentes principais é escolher a ordem para o sistema como a menor ordem que contém quantidade considerável de informações sobre o sistema. Assim,

no exemplo da Figura 2.1, a ordem 3 é a menor ordem e apresenta quantidade de informações suficientes, sendo assim, a ordem sugerida para o sistema analisado.

2.4.2. Índice de desempenho normalizado (ANORM)

Desborough e Harris (1992) propõem uma metodologia que permite a fácil estimativa da variância mínima a partir de dados rotineiros de processo e conseqüentemente a elaboração de um índice normalizado de desempenho de malhas de controladores. O índice de desempenho é calculado através de uma simples regressão linear. Um índice recursivo apropriado para medidas em tempo real também é proposto, baseado em mínimos quadrados recursivo.

Esta metodologia elimina a necessidade de resolver a equação de Diophantine e requer apenas o conhecimento do tempo morto da malha. O índice proposto também pode ser estimado *on-line* através do algoritmo recursivo. O seu valor mostra-se equivalente ao coeficiente de correlação R^2 , utilizado em análises de regressão. Isto possibilita que se examine a dependência explícita que o tamanho da amostra, o tempo morto do processo e a estrutura de autocorrelação do processo tem na incerteza associada a esta estatística.

A equação (2.29) também pode ser escrita como:

$$\mathbf{s}_y^2 = \mathbf{s}_{MV}^2 + \mathbf{s}_{\hat{y}}^2 \quad (2.50)$$

onde o primeiro termo do lado direito é a variância mínima e o segundo termo é a predição d termos a frente. Quando um controlador está atuando como um CVM, o segundo termo do lado direito de (2.50) é zero. Desborough e Harris indicam ser mais útil avaliar o controle não em termos da variância de y_t , mas em termos do erro médio quadrado de y_t :

$$mse(y_t) = \mathbf{s}_y^2 + \mathbf{m}_{y-y_{SET}}^2 \quad (2.51)$$

onde $\mathbf{m}_{y-y_{SET}}^2$ é o desvio médio do *setpoint* e por conveniência e simplicidade, a notação $y(k)$ (e demais variáveis dependentes de k) será representada por y_t . Esta formulação tem a vantagem de levar em conta *offsets* que porventura existam.

Com estas definições, é possível equacionar a estimação do índice de desempenho proposta.

De (2.29) e usando-se a convenção em (2.50) tem-se que:

$$y_t - \mathbf{m}_y = \tilde{y}_t + \hat{y}_t \quad (2.52)$$

que é igual à:

$$y_t - \mathbf{m}_y = \tilde{y}_t + \frac{\mathbf{y}_1(q^{-1})}{\mathbf{y}_2(q^{-1})} (y_{t-d} - \mathbf{m}_y) \quad (2.53)$$

Como o sistema em malha fechada é presumivelmente estável, a razão $\mathbf{y}_1/\mathbf{y}_2$ deve formar uma série convergente no operador q^{-1} . Conseqüentemente, a malha fechada pode ser representada por uma expressão com a forma:

$$y_t - \mathbf{m}_y = \tilde{y}_t + \sum_{k=1}^{\infty} \mathbf{a}_k (y_{t-d-k+1} - \mathbf{m}_y) \quad (2.54)$$

Na prática, a série infinita é truncada após m termos. A variável m seria o equivalente à ordem dos modelos AR, MA ou ARMA utilizados para estimar os polinômios $A(z)$ e $C(z)$ da equação (2.7), que levam à resolução da equação de Diophantine (2.17) para o cálculo da variância mínima.

A idéia principal desta metodologia é então estimar os parâmetros autoregressivos $\{\mathbf{a}\}$ usando uma amostra de dados de operação em malha fechada $\{y_1, \dots, y_n\}$ através de uma regressão linear atrasada, utilizando a equação (2.54). Esta equação, em notação matricial pode ser escrita como:

$$\underline{y} = \mathbf{X}\underline{\mathbf{a}} + \underline{\tilde{y}} \quad (2.55)$$

Onde:

$$\underline{\dot{y}} = \begin{bmatrix} \dot{y}_n \\ \dot{y}_{n-1} \\ \vdots \\ \dot{y}_{d+m} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \dot{y}_{n-d} & \dot{y}_{n-d-1} & \cdots & \dot{y}_{n-d-m+1} \\ \dot{y}_{n-d-1} & \dot{y}_{n-d-2} & \cdots & \dot{y}_{n-d-m} \\ \vdots & \vdots & \ddots & \vdots \\ \dot{y}_m & \dot{y}_{m-1} & \cdots & \dot{y}_1 \end{bmatrix}, \underline{\mathbf{a}} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \quad (2.56)$$

Em que n é o tamanho da amostra e $\dot{y}_n = y_n - \mathbf{m}_n$ é o desvio do *setpoint* corrigido pela média.

A estimativa dos parâmetros é encontrada resolvendo um conjunto de equações lineares:

$$(\mathbf{X}^T \mathbf{X})\underline{\mathbf{a}} = \mathbf{X}^T \underline{\dot{y}} \quad (2.57)$$

A média populacional \mathbf{m} é usualmente substituída pela média amostral, \bar{y} . O erro médio quadrado residual é dado então por:

$$s_e^2 = \frac{(\underline{\dot{y}} - \mathbf{X}\underline{\mathbf{a}})^T (\underline{\dot{y}} - \mathbf{X}\underline{\mathbf{a}})}{n - d - 2m + 1} \quad (2.58)$$

Que é uma estimativa para \mathbf{s}_{MV}^2 .

A estimativa por mínimos quadrados para o índice de desempenho normalizado, $\mathbf{h}(d)$, é:

$$\mathbf{h}(d) = 1 - \frac{n - d - m + 1}{n - d - 2m + 1} \frac{(\underline{\dot{y}} - \mathbf{X}\underline{\mathbf{a}})^T (\underline{\dot{y}} - \mathbf{X}\underline{\mathbf{a}})}{\underline{\dot{y}}^T \underline{\dot{y}} + \overline{y}^2} \quad (2.59)$$

A versão recursiva do índice de desempenho é mais apropriada para estimação em tempo real do desempenho da malha de controle, requerendo a implementação de um algoritmo de mínimos quadrados recursivo. O índice recursivo pode ser utilizado para monitorar mudanças no desempenho de controladores e como uma ferramenta de ajuste de controladores, ao mostrar imediatamente o efeito de mudanças nos parâmetros do controlador. A implementação do índice de desempenho na forma recursiva será abordada na Seção 3.1.

Esta metodologia é aplicável para sistemas estáveis. Quando a função de transferência não é inversível a metodologia deve ser modificada (Desborough e Harris, 1992), para se chegar na equação (2.44), de onde pode-se tirar a variância mínima.

2.4.3. Índice de desempenho FCOR

O algoritmo FCOR é a metodologia proposta por Huang et. al. (1997) (ver também Huang e Shah, 1999) para estimação da variância mínima. FCOR significa *filtering and correlation*, que é exatamente a idéia do método.

O cálculo do índice de desempenho pode ser feito a partir de (2.8), ou seja, o sistema em malha fechada escrito como uma média móvel (MA) de ordem infinita.

$$y_t = \left(\underbrace{1 + f_1 + \dots + f_{m-1}}_{\text{Desconhecido no tempo } k} + \underbrace{p_m + p_{m+1} + \dots}_{\text{Conhecido no tempo } k} \right) e_t \quad (2.60)$$

Multiplicando (2.60) por $e_t, e_{t-1}, \dots, e_{t-d+1}$ respectivamente e tomando a média dos dois lados da equação, resulta:

$$\begin{aligned}
r_{ye}(0) &= E[y_t e_t] = f_0 \mathbf{s}_e^2 \\
r_{ye}(1) &= E[y_t e_{t-1}] = f_1 \mathbf{s}_e^2 \\
r_{ye}(2) &= E[y_t e_{t-2}] = f_2 \mathbf{s}_e^2 \\
&\vdots \\
r_{ye}(d-1) &= E[y_t e_{t-d+1}] = f_{d-1} \mathbf{s}_e^2
\end{aligned} \tag{2.61}$$

Onde $r_{ye}(k)$ é a covariância entre os dados da amostra, y_t , e os dados da perturbação deslocados k unidades na amostra.

Assim sendo, a parte de mínima variância, ou parte invariante com o controlador da variância da saída é dada por:

$$\begin{aligned}
\mathbf{s}_{MV}^2 &= (f_0^2 + f_1^2 + f_2^2 + \dots + f_{d-1}^2) \mathbf{s}^2 \\
&= \left[\left(\frac{r_{ye}(0)}{\mathbf{s}^2} \right)^2 + \left(\frac{r_{ye}(1)}{\mathbf{s}^2} \right)^2 + \dots + \left(\frac{r_{ye}(d-1)}{\mathbf{s}^2} \right)^2 \right] \mathbf{s}^2 \\
&= \frac{[r_{ye}^2(0) + r_{ye}^2(1) + r_{ye}^2(2) + \dots + r_{ye}^2(d-1)]}{\mathbf{s}^2}
\end{aligned} \tag{2.62}$$

Sendo o índice de desempenho dado por (2.48) e substituindo-se (2.61) em (2.48) obtém-se:

$$\begin{aligned}
\hat{\mathbf{H}}(d) &= \frac{[r_{ye}^2(0) + r_{ye}^2(1) + r_{ye}^2(2) + \dots + r_{ye}^2(d-1)]}{\mathbf{s}_y^2 \mathbf{s}^2} \\
&= \mathbf{r}_{ye}^2(0) + \mathbf{r}_{ye}^2(1) + \mathbf{r}_{ye}^2(2) + \dots + \mathbf{r}_{ye}^2(d-1) \equiv \mathbf{Z}\mathbf{Z}^T
\end{aligned} \tag{2.63}$$

Onde \mathbf{Z} é o vetor de coeficientes de correlação cruzada entre y_t e e_t para atrasos de 0 até $d-1$ e os $\mathbf{r}_{ye}(k)$ são os coeficientes de correlação entre y_t e e_t para atrasos de 0 até $d-1$, dados por:

$$\mathbf{r}_{ye}(k) = \frac{\frac{1}{n} \sum_{t=1}^n y_t e_{t-k}}{\sqrt{\frac{1}{n} \sum_{t=1}^n y_t^2 \frac{1}{n} \sum_{t=1}^n e_t^2}} \tag{2.64}$$

Embora e_t seja desconhecido, pode ser estimado como uma seqüência de inovações, \hat{e}_t . Esta estimativa é feita passando a saída do processo, y_t , através de um filtro branco (procedimento chamado *pre-whitening*) via análise de séries temporais.

A fonte original de variação em um processo regulatório em malha fechada pode sempre ser reduzida à um ruído branco e_t , conforme mostrado na Figura 2.2.

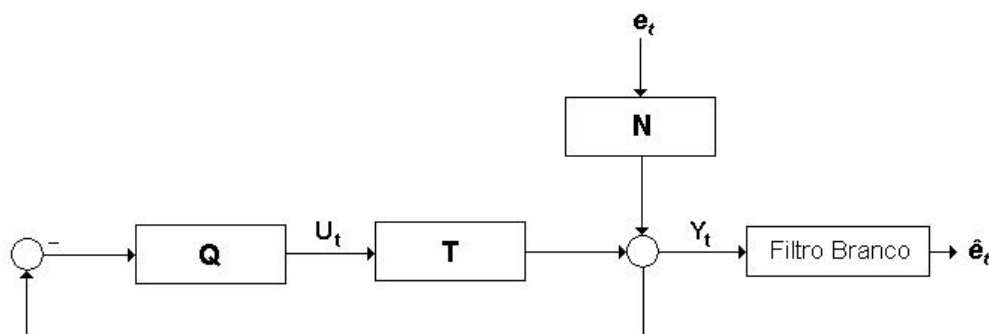


Figura 2.2. Diagrama esquemático do sistema de estimação da seqüência de inovações.

Onde \hat{e}_t é uma estimativa de e_t . A relação entre y_t e e_t é estabelecida pela função de transferência em malha fechada:

$$G_{ye} = \frac{N}{1 + TQ} \quad (2.65)$$

Assim, a variação de y_t é devida a excitação de e_t através de G_{ye} . Revertendo o processo, a seqüência de ruído branco pode ser vista como a saída de um filtro cuja entrada é a saída do processo, y_t . O processo de obtenção do filtro branco é análogo à modelagem com séries temporais, onde o resíduo obtido é o ruído branco desejado. Dependendo dos dados, modelos AR, ARMA ou filtros de Kalman podem ser empregados como filtro branco (Huang e Shah, 1999).

O método FCOR tem a vantagem sobre o método ANORM baseado em mínimos quadrados no fato de não requerer inversão de matrizes, o que acarreta problemas de estabilidade numérica e por vezes inviabiliza a implementação em sistemas de escassos recursos computacionais, tais como SDCD's antigos ou CLP's.

2.4.4. Outros índices

Diversos pesquisadores (Harris et al., 1999) introduziram diferentes formas de se efetuar o *benchmark* da variância do processo com a variância mínima. Enquanto a forma de estimar a variância mínima não difere muito dos métodos aqui apresentados, o cálculo dos índices pode ser feito da forma das equações (2.47) e (2.48) como:

$$\mathbf{x}(d) = \frac{\mathbf{s}_y^2}{\mathbf{s}_{MV}^2} \quad (2.66)$$

Enquanto os índices \mathbf{k} e \mathbf{h} dados pelas equações (2.47) e (2.48) são limitados entre 0 e 1, o índice em (2.66) é sempre maior que a unidade.

$\mathbf{h}(d)$ corresponde diretamente à fração de redução da variância da saída que pode ser alcançada com a implementação de um controlador de variância mínima. O índice nesta forma tem a vantagem de ser diretamente análogo ao coeficiente múltiplo de determinação, R^2 , encontrado em regressão linear (Desborough e Harris, 1992). O índice $\mathbf{k}(d)$ é referido como “potencial de malha fechada” (Huang e Shah, 1999) e indica a fração da qualidade máxima que o sistema está atingindo. O índice $\mathbf{x}(d)$ é igual a unidade quando o processo está em variância mínima aumenta conforme a qualidade da saída piora. Pode ser interpretado como o quanto de variância a mais em relação à variância mínima o sistema está produzindo. A utilização de qualquer um dos três índices é uma opção que o engenheiro deve fazer de acordo com a interpretação que achar mais apropriada ao seu entendimento da teoria envolvida na obtenção dos índices.

Se o interesse é fazer o *benchmark* da saída do controlador com o controlador de variância mínima, os índices devem ser calculados com d igual ao número de intervalos de amostragem equivalentes ao tempo morto, já que o índice de desempenho fornece uma medida do componente predizível no sinal de saída d intervalos de amostragem no futuro. Esta estratégia nem sempre é a mais apropriada. Desborough e Harris (1992) descrevem vários exemplos onde o horizonte de predição adequado é diferente do tempo morto. Por exemplo, um controlador que busca minimizar o erro de predição 10 tempos de amostragem à frente deve usar horizonte igual a 10. Um controlador que faz uso de uma estratégia *feedforward* que reduz o tempo morto em três unidades, deve calcular $\mathbf{h}(d - 3)$.

Thornhill et al. (1999) refere-se a técnica avaliar o desempenho com um horizonte de predição maior que o do tempo morto ($d + h$) como “índice de desempenho de horizonte estendido”. Neste caso, o *benchmark* não é mais a variância mínima e a qualidade do controle entre o tempo morto e o horizonte de predição não é avaliada. A grande vantagem de se utilizar tal abordagem é não precisar de metodologias para estimar o tempo morto.

A comparação com a variância mínima pode muitas vezes ser uma avaliação injusta do desempenho do controlador. Controladores do tipo PID, por exemplo, possuem limitações que vão além dos fatores de fase não mínima: estão limitados até por sua baixa ordem. O tipo de abordagem com horizonte estendido pode compensar estas limitações.

A interpretação de $\mathbf{h}(d+h)$ é que este índice corresponde ao quadrado da correlação entre o erro atual e a estimativa por mínimos quadrados da predição feita $(d + h)$ períodos no passado. Thornhill et al. (1999) sugere que ao invés de refletir o tempo morto real, o horizonte de predição seja utilizado como um critério de engenharia, representando uma demanda que o engenheiro de controle faz sobre a malha de controle. Esta demanda é que “a malha deve eliminar componentes predizíveis do erro do controlador dentro do horizonte especificado”.

Também é sugerido que alguns problemas em diagnosticar o desempenho podem ser resolvidos fazendo-se o uso de $\mathbf{h}(d+h)$ e $\mathbf{h}(d)$ simultaneamente.

Harris et al. (1999) também demonstra uma relação dos coeficientes da resposta em malha fechada (resposta ao impulso, equação (2.8)) com o tempo de assentamento. O tempo de assentamento é freqüentemente utilizado como uma medida de desempenho para sistemas determinísticos. O tempo de assentamento pode ser definido como o tempo t , além do qual os coeficientes da resposta são menores que um valor predefinido. Em termos da resposta ao impulso, o tempo de assentamento é interpretado como o tempo além do qual mudanças no índice de desempenho estendido são pequenas. O uso desta métrica baseada no horizonte estendido permite o diagnóstico de tempos de assentamento demasiadamente longos, que e que não são detectados pelo *benchmark* com a variância mínima.

Bezergianni e Georgakis (2000) e Bezergianni (2001) introduzem uma abordagem de avaliação de desempenho diferente ao propor um novo índice, chamado Índice de Variância Relativo (RVI, Relative Variance Index). A idéia do RVI é comparar não apenas a variância de saída em malha fechada com a variância mínima, mas também com a variância na ausência de controle, ou seja, a variância da saída em malha aberta.

$$RVI = \frac{\mathbf{s}_{OL}^2 - \mathbf{s}_y^2}{\mathbf{s}_{OL}^2 - \mathbf{s}_{MV}^2} \quad (2.67)$$

Com isso, espera-se também levar em conta a dinâmica inerente do processo em malha aberta e conseqüentemente valorizar o esforço do controlador em eliminar distúrbios.

A idéia do índice teoricamente vem a sanar diversas limitações existentes nas metodologias clássicas, mas introduz algumas dificuldades práticas de implementação. O algoritmo é baseado em procedimentos de identificação de modelos para o processo em malha fechada, para o controlador e para o distúrbio. Desta forma, é proposta a identificação no subespaço de estado baseada em PCA (Principal Component Analysis) para identificação dos modelos do processo e do distúrbio e um procedimento baseado em decomposição SVD de estruturas de controladores pré-selecionadas (STFA, Structured Target Factor Analysis) para identificação do controlador. A identificação neste formato requer perturbações na entrada que sejam suficientemente persistentes e um procedimento manual de seleção dos dados talvez tenha que ser empregado ou então perturbações adicionais tenham que ser inseridas no processo. Isto descaracteriza em um certo grau a idéia dos procedimentos originais, que é a de avaliar o desempenho de uma maneira passiva, não intrusiva e automática.

Uma vez de posse dos modelos em malha fechada e aberta, os coeficientes das respostas ao impulso podem ser facilmente obtidos através de divisão longa e com estes coeficientes pode-se calcular as variâncias da saída em malha fechada e aberta e a variância mínima.

Procedimentos para avaliação de desempenho na presença de perturbações determinísticas e *setpoint* variável não serão detalhados neste trabalho, mas diversas propostas

são feitas na literatura (Harris et. al. 1999; Huang e Shah, 1999). As abordagens mais comuns são alterações no modelo do distúrbio para incluir um termo determinístico aditivo ao termo aleatório; modelagem utilizando análise de intervenção; utilização de trajetória desejada no lugar do modelo de intervenção ou simplesmente utilizar a abordagem convencional aqui apresentada.

Avaliação de desempenho de malhas que utilizam estratégias *feedforward* também estão apenas referenciadas nesta dissertação. Desborough e Harris (1992), Qin (1998), Huang e Shah (1999) e Harris et al. (1999) fornecem indicações dos caminhos a serem tomados neste caso.

Horch & Isaksson (1999) e Huang & Shah (1998, 1999) propõem diferentes índices de avaliação de desempenho, onde dinâmicas de malha fechada desejadas são tomadas como o *benchmark* com o qual a variância do processo seja avaliada.

A avaliação de desempenho no controle de tanques-pulmão não pode aplicar diretamente os índices de desempenho na forma em que foram apresentados. O objetivo de tanques pulmões não é o de manter o nível no *setpoint* e sim atenuar a variabilidade da corrente da saída frente a perturbações na entrada. Um controlador ajustado muito agressivamente produziria índices de desempenho satisfatórios, mas acabaria fazendo com que o tanque atuasse como uma continuação da tubulação, descaracterizando o seu papel original. Hugo (2001) propõe que se use para o monitoramento destes sistemas um índice baseado em um controlador-ótimo de nível. Um controlador-ótimo de nível é um otimizador não-linear com restrições, onde o objetivo é minimizar a taxa de mudança máxima da vazão de saída, com o nível sujeito a permanecer dentro das restrições. Assim, o índice pode ser obtido da maneira convencional:

$$h^{Nivel} = 1 - \frac{s_{Otimo}^2}{s_y^2} \quad (2.68)$$

onde s_y^2 é a variância na vazão de saída controlada.

2.5. Procedimentos complementares

Além dos índices que se propõe a avaliar a qualidade do controlador através da análise direta de suas ações de controle ou do resultado delas na saída do processo, existem outras abordagens de análises que visam realizar outras formas de medida de desempenho.

2.5.1. Procedimento de detecção de oscilações

Os procedimentos de auditoria de desempenho de malhas de controle baseados em variância mínima fornecem informações muito úteis para a melhoria das condições do

produto final. No entanto, uma malha de controle pode indicar não estar desempenhando adequadamente por outros fatores alheios ao ajuste do controlador. Uma das causas de desempenho baixo pode ser, por exemplo, problemas de fricção (comumente chamado de “agarramento”) ou até mesmo um mau dimensionamento da válvula de controle. Para tentar identificar estes tipos de problema, alguns estudos foram publicados visando a detecção de oscilações em malhas de controle, que ao ser empregado juntamente com o índices de desempenho podem fornecer um suporte mais consistente à tomada de decisões.

Hägglund (1995) propõe uma metodologia bastante interessante, baseada no monitoramento da integral do erro absoluto (IAE). A idéia é monitorar a magnitude do IAE entre duas intersecções da abscissa zero no erro do controlador:

$$IAE = \int_{t_{i-1}}^{t_i} |y(k) - y_{SET}(k)| dt \quad (2.69)$$

onde t_{i-1} e t_i são duas instâncias sucessivas onde o erro do controlador cruza o zero. Este procedimento assume que o controlador tem ação integral, de modo que a média do erro é zero. Hägglund sugere que, se o controlador for proporcional, um valor médio do sinal de medida possa ser obtido através de um filtro. O valor de IAE é zerado a cada vez que o erro intercepta a abscissa.

Durante controle adequado, o valor de IAE se torna pequeno, pois os tempo entre as intersecções se torna curto. Se um distúrbio na carga ocorre, o valor de IAE se torna grande e a diferença $t_{i-1} - t_i$ também é grande. Quando a integral do erro ultrapassa um certo limite pré-estabelecido, IAE_{LIM} , o procedimento considera que um distúrbio na carga foi detectado. Observando a taxa em que distúrbios na carga são detectados, pode-se concluir que não são distúrbios normais de operação, e sim problemas de oscilação. Então a malha é considerada oscilatória naquele período de supervisão T_{sup} . Para uma versão não recursiva do algoritmo, se o número de distúrbios na carga detectados durante T_{sup} excede um certo limite, n_{LIM} , a malha é caracterizada como oscilatória. Em uma implementação recursiva, o algoritmo de detecção de oscilações utiliza um “fator de esquecimento”, que é calculado a partir de T_{sup} .

O procedimento de detecção de oscilações requer que alguns parâmetros sejam estabelecidos para seu funcionamento. Em Hägglund (1995, 2002) estão sugeridos alguns métodos para determinar mais adequadamente estes valores, bem como pode ser encontrada uma listagem do algoritmo em pseudocódigo. No Capítulo 3 são feitas algumas considerações práticas em relação à determinação destes parâmetros.

A utilização de um procedimento de detecção de oscilações é um importante complemento para a avaliação de desempenho. Pode ser usado da seguinte maneira: uma vez detectado um baixo desempenho na ação reguladora do controlador, deve-se verificar se a malha está oscilatória. Se a malha for oscilatória uma investigação deve ser feita a fim de verificar as possíveis causas. Controladores regulados com ganhos excessivos também podem resultar em controle oscilatório, bem como a presença de perturbações externas oscilatórias. Assim, um procedimento metódico não automático deve ser realizado e está sumarizado na Figura 2.3, de acordo com o proposto em Hägglund (1995).

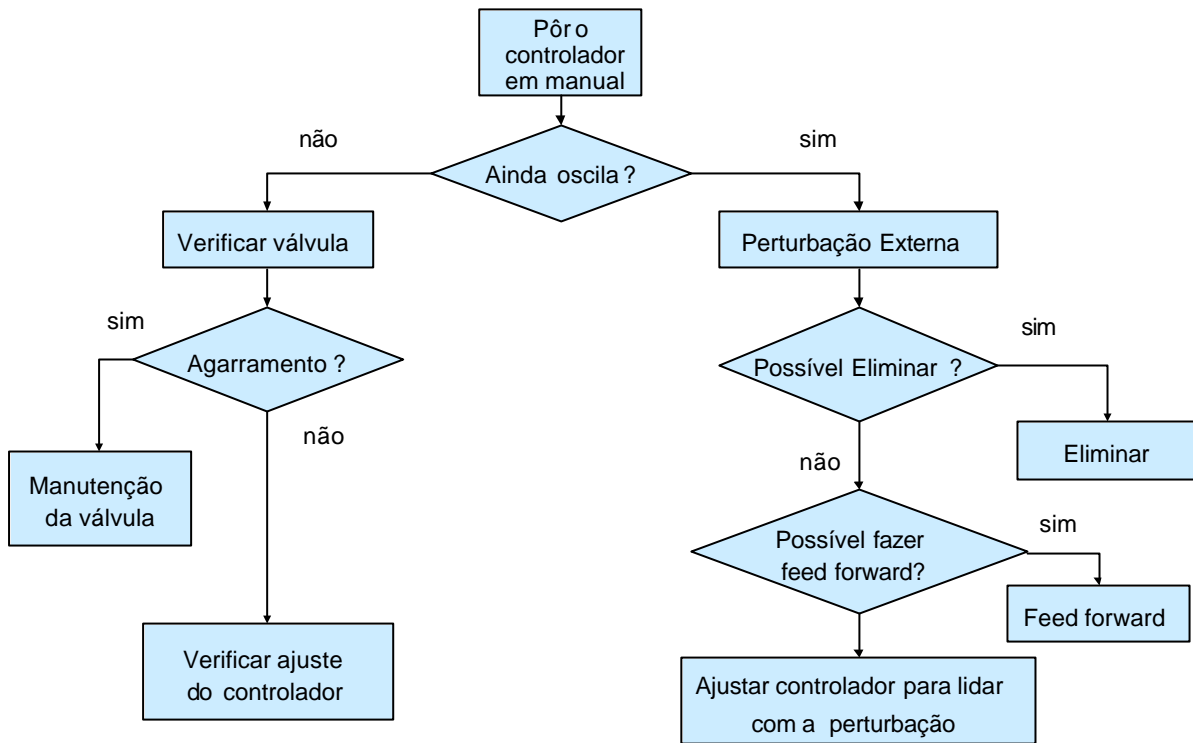


Figura 2.3. Seqüência do procedimento manual de diagnóstico das causas de oscilação.

2.5.2. Índices de eficiência operacional

A proposta dos índices de eficiência operacional é quantificar o desempenho de controladores através de uma avaliação indireta. Esta avaliação não é feita tomando como base as ações do controlador ou o seu resultado na saída propriamente, mas sim a forma em que este é utilizado.

Estes índices vêm sendo largamente empregados no controle avançado das unidades de refino da Petrobras Petróleo do Brasil S.A. e os resultados levam a conclusões significativas sobre o gerenciamento destas unidades em nível de controle. Seus controladores preditivos operam num sistema de faixas, onde a variável controlada não possui um *setpoint* propriamente dito, e sim opera entre dois limites. Quando uma variável atinge um destes extremos, a variável passa a ser classificada como “ativa”.

Através de procedimentos de monitoramento simples, capazes de serem implementados em qualquer unidade industrial com algum grau de automação, são avaliados quesitos como, por exemplo, tempo de utilização de variáveis manipuladas pelo controlador, quantidade de variáveis manipuladas utilizadas e de controladas ativas.

O Fator Operacional dos controladores (FOP) é dado pela razão entre o tempo em que o controlador permanece com pelo menos uma variável manipulada ligada (T_{ma}) e o tempo total do mês (T).

$$FOP = 100 \frac{T_{ma}}{T} \quad (2.70)$$

O Grau de Utilização dos controladores (GUT) é uma medida do percentual de variáveis manipuladas ligadas. Com este índice pode-se ter uma noção do aproveitamento dos graus de liberdade disponíveis no processo.

$$GUT = 100 \frac{M_{ut}}{M} \quad (2.71)$$

Onde M_{ut} é o número de manipuladas utilizadas pelo controlador e M é o número total de manipuladas disponíveis.

Um índice que monitora variáveis controladas é o Percentual de Controladas Ativas (PCAT). É dado pela razão entre o número de variáveis controladas ativas (variáveis que estão na restrição) C_{at} e o número total de manipuladas.

$$PCAT = 100 \frac{C_{at}}{M} \quad (2.72)$$

O Percentual de Manipuladas Ativas é dado pela razão entre o número de variáveis manipuladas sendo efetivamente utilizadas pelo controlador e o número total de variáveis manipuladas disponíveis ao controlador.

$$PMAT = 100 \frac{M_{ef}}{M_{ut}} \quad (2.73)$$

Relatórios contendo estes índices e mais detalhes da análise de operação do controlador são gerados automaticamente pelo sistema de controle. A comparação de desempenho para cada variável controlada e manipulada é feita, por exemplo, comparando-se os valores de uma outra variável que expresse a qualidade do produto final, como a octanagem de uma gasolina, durante o período em que a variável controlada/manipulada esteve ligada e desligada.

Os relatórios são atualizados diariamente e contém a análise da operação do controlador desde o início do mês corrente até a data atual. Com frequência diária também são feitas as seguintes auditorias:

- Verificação da existência de variáveis manipuladas desligadas;
- Verificação da violação dos limites das variáveis manipuladas (geralmente saturação de válvulas);
- Verificação do estado ligado/desligado das variáveis controladas;

- Verificação de violações nas restrições das variáveis controladas;

A popularização dos novos sistemas de instrumentação do tipo Fieldbus e similares que empregam sensores inteligentes também permite que se adquira informações operacionais com maior relevância e diretamente do instrumento. Tais sensores são capazes de informar a respeito do modo de operação da malha, da qualidade da operação (bom, suspeito e ruim, em geral), da razão do mau funcionamento do sensor e sobre as violações de restrições ou limites da variável. Estas informações podem ser coletadas e apresentadas na forma de índices tais como os demais índices operacionais aqui apresentados.

2.5.3. Técnicas estatísticas

É bastante prudente e recomendável embasar as conclusões obtidas através dos índices de avaliação de desempenho e do procedimento de detecção de oscilações com algumas técnicas estatísticas tradicionais. Tais ferramentas são um importante complemento para diagnóstico e avaliação de malhas de controle.

Função de autocorrelação (ACF): a ACF da saída do processo é uma medida aproximada do quão perto da variância mínima encontra-se o sistema, ou seja, do quão predizíveis são os dados ao longo do horizonte de predição de interesse.

A ACF surge calculando-se os valores de correlação da série temporal em relação a ela mesma para deslocamentos de zero até um horizonte de predição desejado. Na avaliação de malhas de controle é interessante que este horizonte seja maior que o tempo morto (d). Se um sistema produz elevados valores de autocorrelação significa que os valores no tempo t da série temporal são altamente influenciados por seus valores passados ($t - 1, t - 2, \dots$). Para um controlador bem ajustado, a função de autocorrelação deve cair a zero rapidamente tão logo o tempo morto tenha se esgotado, pois a adequada ação do controlador elimina a dependência da série temporal de seus valores passados. Assim, não deve haver informação predizível além do atraso $d - 1$. A função de autocorrelação também pode ser utilizada como uma ferramenta para avaliar a qualidade do conjunto de dados sendo analisado com respeito a presença de componentes periódicos. A ACF varia entre -1 e 1 , com valores perto de 1 ou -1 indicando forte correlação.

$$ACF(d) = \frac{\sum_{i=1}^{n-d} (x_i - \bar{x})(x_{i+d} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.74)$$

Onde d é o horizonte de predição, n é o tamanho total do conjunto de dados, x_i é o valor da variável no atraso i e \bar{x} é a média da amostra. Referências: Huang e Shah (1999), BOURKE e GEOS.

Correlação cruzada: A correlação cruzada é especialmente útil para verificar a influência que uma determinada malha tem sobre outra. Isto ocorre em malhas cascadeadas, onde o

setpoint da malha interna é dado pela saída de uma outra malha (externa), ou em malhas com forte interação. Muitas vezes o desempenho de uma malha é prejudicado pela interação com outra malha, embora o controlador esteja teoricamente bem ajustado. Este fator deve ser levado em conta em uma posterior avaliação de desempenho em relação a variância mínima. A função de correlação cruzada entre duas amostras x e y para um atraso d é dada por:

$$r(d) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_{i-d} - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_{i-d} - \bar{y})^2}} \quad (2.75)$$

Referências: Huang e Shah (1999), BOURKE e GEOS.

Resposta ao impulso: a função de resposta ao impulso representa um relacionamento dinâmico entre a série temporal do distúrbio na forma de ruído branco e a saída do processo. Pode ser calculada diretamente usando análise de séries temporais. De fato, a metodologia original de Harris (1989) calcula a variância do processo e a variância mínima para a determinação do índice de desempenho através da estimação de respostas ao impulso da malha fechada. Pode servir como um indicativo da velocidade da dinâmica e agressividade do controlador: a resposta pode ter um decaimento lento ou rápido e suave ou oscilatória para diferentes ajustes do controlador. Para calcular a resposta ao impulso basta modelar a série temporal como uma função do tipo ARMA, MA ou AR e efetuar a divisão longa do numerador pelo denominador. Referência: Huang e Shah (1999).

Análise espectral: a resposta do sistema em malha fechada no domínio da frequência é uma maneira alternativa de verificar a qualidade da malha. Com procedimentos espectrais pode-se detectar oscilações, *offset's* e ruídos de medida de modo simples. Para se efetuar uma análise espectral é necessário transformar os dados de saída do processo no domínio do tempo, y_t , para o domínio da frequência, Y_f , através da transformada de Fourier.

A comparação da resposta no domínio da frequência (FRD) real do sistema em malha fechada com a FRD do sistema em CVM permite avaliar a possibilidade de melhora do desempenho através de ajuste nos parâmetros do controlador. Esta comparação também permite verificar em quais frequências o controlador se desvia significativamente da variância mínima. Desvios em baixa frequência normalmente indicam ação integral insuficiente ou ganho proporcional baixo. Picos largos em frequências médias tipicamente indicam um controlador ajustado de forma muito sensível ou presença de distúrbios oscilatórios.

Com a função densidade de espectro de potência (PSD) também pode-se tirar conclusões a respeito de questões como: verificar se é possível obter melhora com re-sintonia de parâmetros do controlador e se há uma oscilação escondida no ruído. Referências: Huang e Shah (1999), GEOS, EXPERTUNE e STATSOFT.

Teste residual: Os algoritmos para cálculo do índice de desempenho baseados em variância mínima supõem que o sistema é linear ou possa ser considerado aproximadamente linear. Os

resíduos são obtidos ajustando-se um modelo linear aos dados. Se os resíduos passam no teste do “ruído branco”, ou seja, são uma seqüência aleatória, independente, igualmente distribuída e possuem média zero, então os modelos estimados que foram usados para avaliação de desempenho são válidos. Para serem considerados como ruído branco, a função de autocorrelação dos resíduos deve ficar dentro do intervalo de confiança de 95%. Esta característica pode ser utilizada para verificar a qualidade da estimação do índice de desempenho. Referências: Huang e Shah (1999) e GEOS.

2.6. Conclusões

A abrangente revisão feita neste capítulo mostra que nos últimos dez anos tem-se desenvolvido bastante pesquisa na área de avaliação de desempenho de malhas de controle e grandes avanços foram obtidos com a introdução de ferramentas que permitem auditar de uma maneira simples e sistemática a qualidade dos sistemas de controle.

As estimativas da variância mínima calculadas com as abordagens propostas, em teoria não dependem da estrutura de controle existente, visto que a resposta do sistema em variância mínima pode ser obtida através da resolução de uma equação de Diophantine envolvendo apenas o tempo morto e o modelo do distúrbio. No Capítulo 3 esta afirmação será verificada na prática. No entanto, ao utilizar-se o horizonte de predição estendido, o critério de referência de desempenho passa a depender do controlador, já que a comparação não está mais considerando apenas variância mínima. Com isso a variância do erro de predição é afetada pela estrutura e o ajuste do controlador *feedback*.

O índice de desempenho normalizado é independente da magnitude da força motriz da perturbação, ou seja, a variância do sinal de excitação do distúrbio, já que a razão entre a variância mínima dada por (2.37) e a variância real da saída do processo dada por:

$$\mathbf{s}_y^2 = \left(\sum_{i=0}^{\infty} p_i \right) \mathbf{s}^2 \quad (2.76)$$

faz com que \mathbf{s}^2 se cancele. Em (2.76) Os primeiros $d - 1$ coeficientes p_i são iguais aos coeficientes de F, f_i , em (2.37).

Após a revisão bibliográfica, resta analisar se os requisitos de uma ferramenta de monitoramento e avaliação de desempenho de malhas de controle colocados na Introdução são preenchidos pelas técnicas apresentadas.

Os índices, ao menos na sua forma mais básica não requerem testes na planta para sua estimação e são independentes do *setpoint*. A dependência da métrica em relação as perturbações será analisada no Capítulo 3. Os índices podem perfeitamente ser implementados automaticamente, inclusive com versões recursivas, já que isto requer apenas

a utilização de algoritmos de mínimos quadrados recursivos. A única especificação da dinâmica dos processos que as metodologias na sua forma mais básica requerem é a do tempo morto, mas diversas propostas podem ser encontradas na literatura que visam estimar o tempo morto a partir de dados em malha fechada. Algumas destas técnicas serão apresentadas no Capítulo 4. A medida da qualidade dos processos é baseada na comparação com a variância mínima, sendo portanto uma medida não arbitrária e absoluta da qualidade para uma determinada malha. Ainda que se possa considerar a avaliação em relação a variância mínima um tanto injusta em alguns casos para controladores simples do tipo PID, não pode-se rotular o critério como subjetivo. O procedimento de avaliação é sensível ao desajuste do controlador, incertezas de modelo e problemas de equipamento, bem como a características da perturbação.

Uma das limitações destas técnicas, se não empregadas em conjunto com outras técnicas, estatísticas e procedimentos, é que não têm como fornecer um indicativo do porquê de o controlador estar desempenhando insatisfatoriamente. Por esta razão não é indicada a tomada de decisões sem uma análise mais aprofundada do problema e sem o suporte de outras ferramentas, baseado somente na estimativa do índice de desempenho. De fato, a avaliação de desempenho de controladores pode ser encarada como uma extensão do Controle Estatístico de Processos (CEP), que utiliza monitoramento de capacidade de processo e análise da raiz das causas, Análise de Variância (ANOVA) entre outras ferramentas.

Quando um controlador não atinge seus objetivos, pode ser necessária a re-sintonia de seus parâmetros, manutenção do *hardware* de controle, implementação de algoritmos *feedback* / *feedforward* alternativos, uso de variáveis manipuladas alternativas, modificações no processo para reduzir ou eliminar o tempo morto ou distúrbios de processo.

Quando um controlador está produzindo variância muito próxima da mínima, então a re-sintonia do controlador não poderá produzir melhorias. Se neste caso a variância da saída não satisfaz as especificações do processo ou do produto final, então as outras soluções acima apresentadas devem ser tentadas.

Capítulo 3 Aplicações em malhas SISO

Este capítulo é o resultado de um conjunto de testes simulados e com dados de processos industriais. Foram postos a prova os métodos de avaliação de desempenho apresentados no Capítulo 2 bem como alguns procedimentos complementares. A avaliação dos resultados produzidos por esta batelada de testes permite formular algumas conclusões interessantes a respeito das questões que envolvem a utilização destes procedimentos em ambiente industrial.

As simulações foram realizadas utilizando-se Matlab/Simulink® e os testes com dados reais foram executados a partir de dados provenientes de fornos de craqueamento de uma unidade de primeira geração e de uma unidade de refino da Petrobrás Petróleo do Brasil S.A.

3.1. Formas de implementação

Os algoritmos de avaliação de desempenho de malhas podem ser implementados em um sistema automático basicamente de duas formas: janela móvel (ou batelada) e recursiva.

A implementação em janela móvel utiliza os algoritmos apresentados no Capítulo 2 recebendo um conjunto de dados de tamanho apropriado através de um sistema de coleta de dados ou de um historiador. O sistema de coleta de dados é o responsável por acumular dados coletados com um intervalo de amostragem adequado (regras heurísticas para seleção de intervalo de amostragem, tamanho de amostra entre outros parâmetros serão apresentados na Seção 3.2) até que o conjunto de dados atinja o tamanho apropriado. Neste momento o índice de desempenho é calculado para todo o conjunto acumulado. Se um historiador estiver disponível, como por exemplo, o Aspen InfoPlus ou o OSI PI, os dados correspondentes a janela podem ser adquiridos no momento em que chegar o intervalo de tempo correspondente ao armazenamento do conjunto de dados.

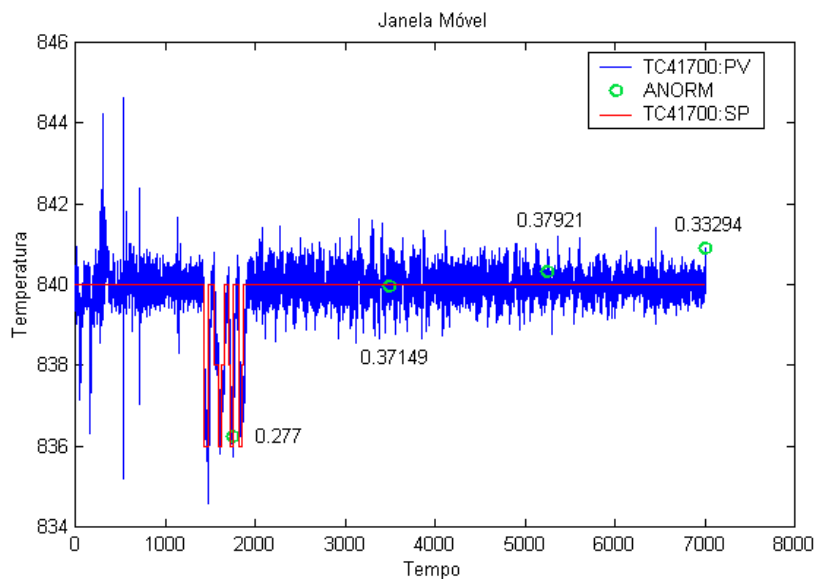


Figura 3.1. Janela móvel dos dados de um forno industrial, com estimação a cada 1750 pontos.

A implementação recursiva está baseada na utilização de um algoritmo de mínimos quadrados recursivo. Este algoritmo é mais apropriado para fornecer uma estimativa *on-line* dos parâmetros \mathbf{a}_k da Equação (2.54). Para a estimação de \mathbf{a}_k foi utilizada a média amostral de y . No caso do algoritmo recursivo, isto não é possível e o modelo de malha fechada deve ser mudado para incluir um termo constante \mathbf{a}_0 . Os algoritmos de mínimos quadrados recursivos empregam um fator de esquecimento, λ , que faz com que os dados mais recentes tenham mais peso na estimação. Assim, o algoritmo minimiza uma função-objetivo do tipo:

$$J = (\mathbf{y} - \mathbf{X}\mathbf{a})^T \Lambda (\mathbf{y} - \mathbf{X}\mathbf{a}) \quad (3.1)$$

em que Λ é uma matriz diagonal com elementos $(\lambda, \lambda^2, \dots, \lambda^n)$.

Uma estimativa da variância mínima no tempo t é dada por:

$$\hat{\mathbf{S}}_{MV,t}^2 = \mathbf{I} \hat{\mathbf{S}}_{MV,t-1}^2 + e_t^2 \quad (3.2)$$

onde e_t^2 é o erro de predição no tempo t .

A média do erro quadrado ponderada exponencialmente é dada por:

$$\hat{\mathbf{S}}_{y,t}^2 = \mathbf{I} \hat{\mathbf{S}}_{y,t-1}^2 + y_t^2 \quad (3.3)$$

O índice de desempenho normalizado é calculado como:

$$\mathbf{R}(d) = 1 - \frac{\hat{\mathbf{S}}_{MV,t}^2}{\hat{\mathbf{S}}_{y,t}^2} \quad (3.4)$$

O índice de desempenho normalizado estimado recursivamente é útil para a detecção de pontos de mudança de desempenho, que ficariam indefinidos no caso da implementação em janela móvel. Também é vantajoso na análise de desempenho de sistemas muito não-lineares. A metodologia não recursiva é válida somente para processos aproximadamente lineares, conforme demonstrado no Capítulo 2. No entanto, se a dinâmica do processo for lenta o suficiente de tal forma que o processo possa ser considerado localmente linear, a estimação recursiva fornece uma estimativa local do desempenho do controlador. O índice estimado também pode ser utilizado como uma ferramenta de ajuste, mostrando imediatamente se as alterações nos parâmetros do controlador melhoram ou pioram o desempenho da malha (Desborough e Harris, 1992).

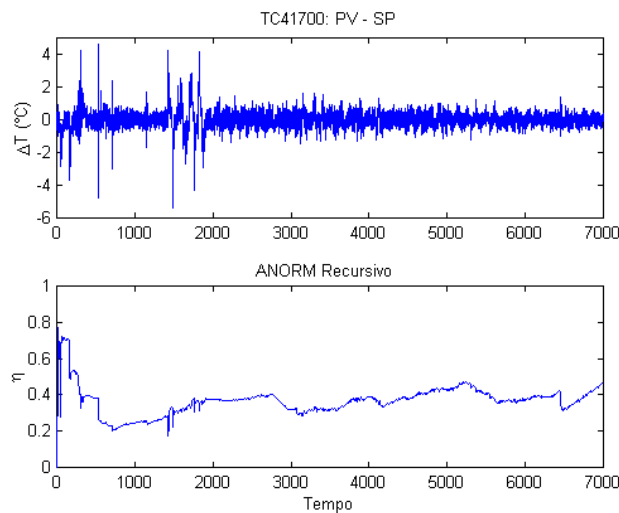


Figura 3.2. Avaliação recursiva do índice de desempenho do mesmo conjunto de dados da Figura 3.1. A linha de cima apresenta a variável de processo subtraída do *setpoint*.

Com o surgimento dos sensores “inteligentes”, funcionando através de sistemas com processamento distribuído (Fieldbus, Profibus etc), surge uma possibilidade interessante, que é a de deixar o cálculo dos índices de desempenho (especialmente no modo recursivo) à cargo da unidade de processamento do instrumento. O baixo requerimento computacional em termos de processamento e memória dos algoritmos recursivos baseados em variância mínima permite com que esta possibilidade seja viável, muito embora não tenha-se encontrado alguma referência na literatura a respeito de tal implementação. A principal vantagem de implementar-se estes algoritmos diretamente na unidade de processamento do sensor está em casos de dinâmicas bastante rápidas que requerem frequência de amostragem elevada, de modo que SDCD's ou bases de dados históricos não preencham este requisito.

3.2. Efeito dos parâmetros e questões práticas

Enquanto a teoria por trás dos métodos de avaliação de desempenho se mostra bastante atraente e permite vislumbrar resultados significativos, o emprego prático das diversas metodologias ainda apresenta uma série de pequenos empecilhos. Diversos parâmetros para que a metodologia funcione na sua forma mais automatizada devem ser obtidos com base em testes não automáticos ou segundo regras heurísticas. Neste sentido, vários pesquisadores propõem maneiras sistemáticas ou baseadas em conhecimento prévio para a escolha destes parâmetros e várias destas propostas serão apresentadas a seguir. Notadamente, Thornhill et al. (1999) realiza um amplo estudo em diversas refinarias e sugere valores empíricos baseados em experiência de operadores ou em regras heurísticas para os principais parâmetros de interesse para avaliação de desempenho. Na Seção 3.3 diversas proposições serão utilizadas em casos práticos e algumas novas considerações serão feitas.

3.2.1. *Tempo morto ou horizonte de predição (d ou h)*

A teoria na qual se baseiam os índices de desempenho apresentados neste trabalho (Capítulo 2) requer que o horizonte de predição seja escolhido como sendo o número de intervalos de amostragem equivalentes ao tempo morto. A existência de tempo morto é a mais fundamental limitação ao desempenho de controladores *feedback*. O tempo morto e o erro da saída da malha são as variáveis-chave para o procedimento de avaliação de desempenho de malhas. Enquanto o erro da saída é facilmente obtido do processo, o valor do tempo morto de uma malha é uma variável que raramente encontra-se disponível para uso, isso para não falar de processos que se caracterizam por tempo morto variável com o ponto de operação. Dentro da filosofia de que o monitoramento de desempenho deva ser efetuado de modo que a produção normal de um processo seja afetada o mínimo possível, as metodologias de estimação de tempo morto que requerem colocar a malha em modo manual não são muito recomendadas. Mas se, por exemplo, dados históricos de experimentos em malha aberta existirem ou algum teste em malha aberta for imprescindível para outros fins, pode-se utilizá-los para obter uma aproximação boa do tempo morto.

Diversas metodologias que se propõem a estimar o tempo morto em malha fechada, tanto no domínio contínuo como no discreto, existem na literatura (Tuch et al. 1994; Huang et al., 1997; Kamrunnahar et al., 2000). Ferreira e Fernandes (1997) é uma boa fonte de referências sobre o assunto. No Capítulo 4 uma destas técnicas é estudada e aplicada em testes.

Não existe técnica para estimação do tempo morto que não requeira um mínimo de intervenção no processo. Thornhill et al. (1999) apresenta valores recomendados para diferentes tipos de malhas genéricas em refinarias. Neste artigo o autor também sugere que o desempenho não seja avaliado em relação à variância mínima, o que requer conhecimento do tempo morto, mas sim que o horizonte de predição seja utilizado como um critério de engenharia que representa o desempenho desejado para uma determinada malha. O autor ainda sugere que sejam feitos gráficos do índice de desempenho *versus* o horizonte de predição, tal como na Figura 3.3.

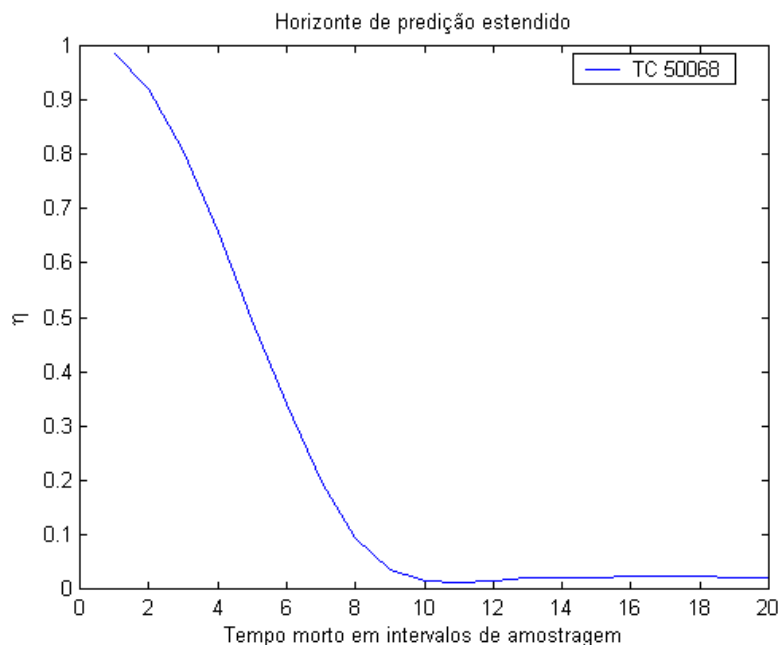


Figura 3.3. Gráfico de horizonte de predição para uma malha de temperatura de uma refinaria.

De posse de gráficos deste tipo, é sugerido que o horizonte de predição seja escolhido como o valor em que o índice de desempenho cai em um patamar no qual não varia consideravelmente. A intenção é escolher valores para o tempo morto de forma a eliminar a dependência do índice com o horizonte de predição. O valor do horizonte de predição em que o índice de desempenho atinge valores próximos a zero é interpretado como o número de intervalos de amostragem que o controlador precisa tomar para rejeitar completamente um distúrbio. No entanto, esta escolha ainda é bastante subjetiva, uma vez que em muitos casos não chega a haver um patamar definido ou em outros há o surgimento de vários patamares, conforme ilustra a Figura 3.4.

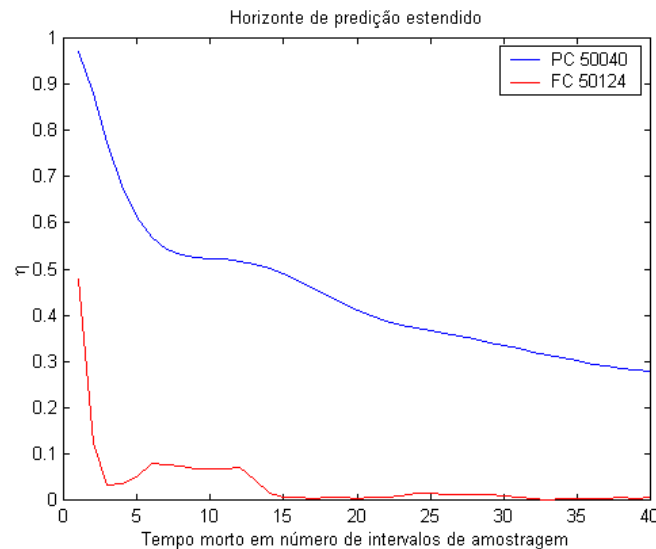


Figura 3.4. Gráfico de horizonte de predição para uma malha de pressão de uma refinaria. Aqui fica clara a dificuldade de se escolher um patamar no qual o índice de desempenho não varia rapidamente.

Em casos em que o controlador apresenta um desempenho ruim esta análise fica absolutamente impraticável, conforme mostrado na Figura 3.5.

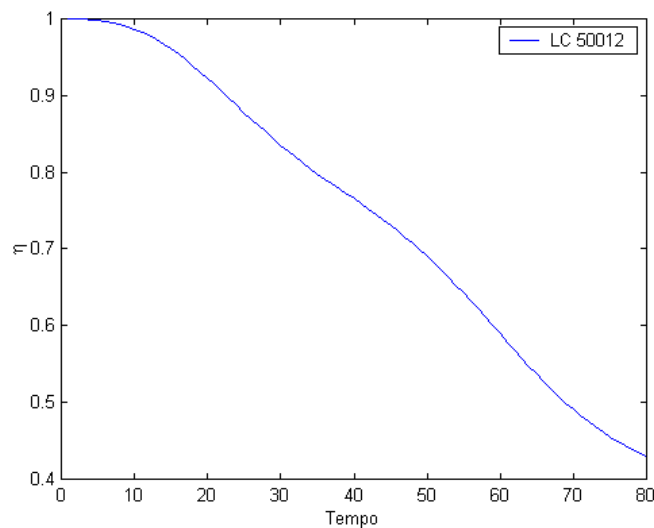


Figura 3.5. Índice de desempenho em função do horizonte de predição para malha de nível onde não existe patamar algum na variação de η .

3.2.2. Tamanho do conjunto de dados (n)

As implicações do tamanho do conjunto de dados nas propriedades estatísticas do índice de desempenho são discutidas em Desborough e Harris (1992). Deste estudo pode-se concluir que amostras com tamanho pequeno devem ser evitadas pois as propriedades estatísticas tornam-se mais sensíveis a pontos ruins e a variabilidade de $\mathbf{h}(d)$ aumenta. Por outro lado, amostras muito grandes podem mascarar mudanças no desempenho do controlador ou no processo, fazendo com que o índice represente uma média do desempenho.

Desta forma, a escolha do tamanho da amostra pode ser encarada como um compromisso entre intervalos de confiança estreitos para as estimações e a perda de informação sobre a real situação do desempenho da malha. Os intervalos de confiança em função do tamanho da amostra n podem ser obtidos através do cálculo da variância de \mathbf{h} :

$$\text{var}[\hat{\mathbf{h}}(d)] = \frac{4}{n} [1 - \mathbf{h}]^2 \left(\sum_{k=1}^{d-1} (\mathbf{r}_k - \mathbf{r}_{e,k})^2 + \sum_{k=d}^{\infty} \mathbf{r}_k^2 \right) \quad (3.5)$$

onde \mathbf{r} e \mathbf{r}_e são a autocorrelação da saída e do erro do controlador (saída menos *setpoint*) respectivamente.

Thornhill et al. (1999) aponta que, para dados de refinarias, um bom balanço entre a confiança estatística e a adequada caracterização da malha é obtida com amostras com 1500 pontos.

3.2.3. Ordem dos modelos autoregressivos (m)

A ordem dos modelos da saída em malha fechada é outro importante parâmetro a ser definido para a estimação do índice de desempenho. No entanto a influência da sua definição sobre o índice não é tão significativa como o tempo morto, de forma que as técnicas comuns de seleção de estruturas de modelos tais como o critério de Akaike (AIC, Akaike Information Criteria) não sejam efetivamente empregados na literatura.

Fazendo-se variar o índice de desempenho estimado em relação à ordem dos modelos, nota-se que ordens abaixo de 5 não produzem resultados satisfatórios na maioria dos casos. De fato, Desborough e Harris (1992) sugerem que para determinar a ordem, deve-se ir aumentando a ordem do modelo a partir de 5 até um valor em que o índice não apresente mudanças consideráveis. Com isso espera-se escolher um valor em um ponto tal que $\mathbf{h}(d)$ torna-se independente da ordem.

Thornhill et al. (1999) aponta que o número de parâmetros do modelo e o intervalo de amostragem não são independentes entre si pois ambos influenciam no tamanho da região do tempo que o modelo é capaz de capturar. Por isso no referido trabalho é sugerido fixar a ordem dos modelos em 30 para todos os casos em seu estudo e ajustar as propriedades amostrais para cada tipo genérico de malha. Um modelo com este tamanho tem consideráveis conseqüências em termos de condicionamento numérico devido a necessidade de inversão de matrizes para cálculo dos parâmetros do modelo segundo a Equação (2.57). A matriz tende a tornar-se singular com o aumento do número de parâmetros e com o aumento da preditabilidade dos dados.

Este problema pode ser contornado ao utilizar-se o algoritmo FCOR (Capítulo 2), que não requer inversão de matrizes.

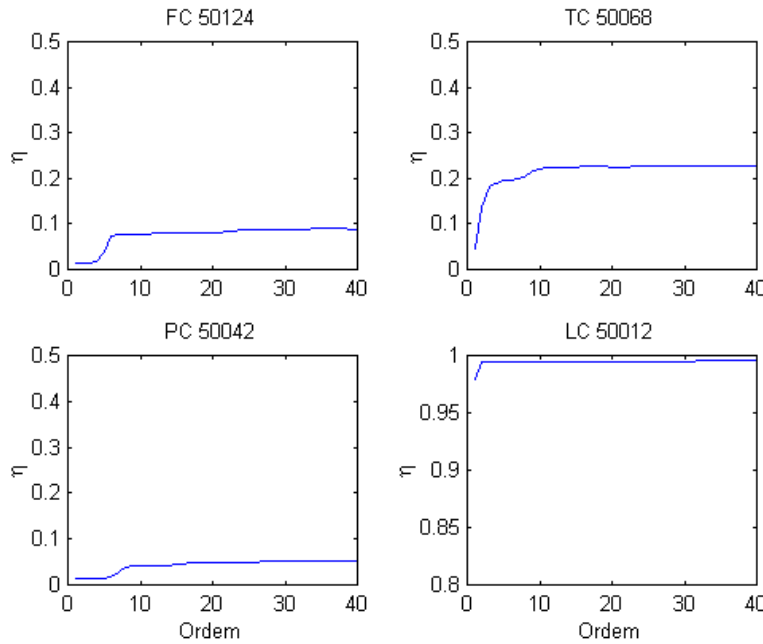


Figura 3.6. Índice de desempenho em função do número de parâmetros do modelo para diferentes tipos de malha, mostrando que a influência deste fator deixa de ser importante para ordens maiores que 5.

De qualquer forma, a Figura 3.6 mostra que não é necessário utilizar-se modelos de ordem tão elevada já que modelos com 10 parâmetros fornecem uma ótima estimativa do índice de desempenho, de forma que é preferível alterar-se o intervalo de amostragem para cada tipo de malha genérica e desta forma também obter melhor desempenho computacional.

3.2.4. Intervalo de amostragem

Uma vez que o tempo morto ou horizonte de predição são empregados na estimação do índice de desempenho em termos de número de intervalos de amostragem, o índice não deve ser sensível ao intervalo de amostragem. No entanto Thornhill et al. (1999) indica ser importante otimizar o valor do intervalo de amostragem quando se deseja realizar uma análise espectral dos dados para detecção das causas de baixo desempenho. Para seleccionar o intervalo de amostragem, é sugerido que se empregue um valor tal que a resposta ao impulso do sistema em malha fechada seja capturada dentro de m amostragens.

Assim, se os dados são amostrados com frequência excessiva, a parte transiente da resposta ao impulso não se assenta dentro dos m intervalos de amostragem. Se os dados são subamostrados, a resposta ao impulso em malha fechada assenta-se com poucas amostragens e não é adequadamente capturada, já que o comportamento dinâmico do sistema perde-se entre um intervalo de amostragem e outro. Em ambos os casos, as propriedades espectrais do conjunto de dados não são adequadamente resolvidas.

3.2.5. Fator de esquecimento

Na implementação recursiva do índice de desempenho também é necessário ser especificado um valor para o fator de esquecimento, I . Desborough e Harris (1992) sugerem que se escolha I como uma função do número de pontos da janela de dados:

$$I = \frac{n - 1}{n + 1} \quad (3.6)$$

de forma que as propriedades estatísticas inerentes ao índice façam com que a variância calculada nas formas recursiva e não-recursiva sejam comparáveis.

A escolha do valor de I está intimamente relacionada às questões referentes a escolha do tamanho do conjunto de dados apresentadas na Seção 3.2.2. Um valor de I muito próximo a 1 equivale a uma amostra de dados muito grande enquanto que um valor muito menor que 1 equivale a uma amostra de dados muito pequena.

3.2.6. Referência desejada

Uma técnica simples que possibilita uma avaliação de desempenho mais realística em sistemas com *setpoint* variável é a utilização da referência desejada no lugar do *setpoint* para o cálculo do erro do controlador e posterior estimação do índice de desempenho. Em mudanças bruscas de *setpoint*, tais como degraus consideráveis, o índice estimado pode acusar perda de desempenho do controlador, pois seu erro torna-se significativo até o assentamento da saída. A técnica consiste da simples passagem do sinal do *setpoint* através de uma função de transferência que produz a referência desejada para a avaliação de desempenho.

A Figura 3.7 e a Figura 3.8 ilustram um exemplo da aplicação. A Figura 3.7 apresenta um diagrama de blocos com um sistema utilizando como trajetória de referência de primeira ordem com o mesmo tempo morto da planta e um sistema sem esta estratégia. Os resultados produzidos estão ilustrados na Figura 3.8, onde por comparação visual entre os *setpoints* e as saídas das duas estratégias evidencia que o erro da saída no caso com trajetória de referência é consideravelmente menor que o erro na estratégia tradicional. A trajetória de referência faz com que o índice de desempenho não acuse perda de desempenho significativa do controlador, tornando o *benchmark* muito mais justo para sistemas com *setpoint* variável.

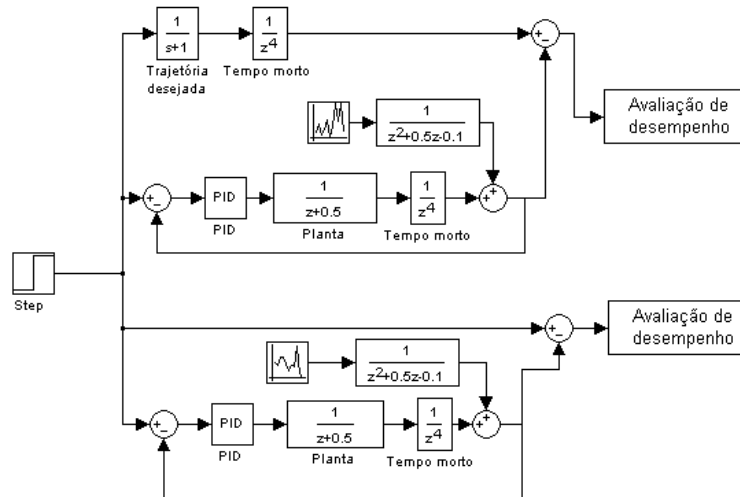


Figura 3.7. Esquema para avaliação de desempenho com referência desejada. A malha superior apresenta a compensação com a referência desejada para um degrau no *setpoint*. A malha inferior é o caso normal. Os resultados encontram-se na Figura 3.8.

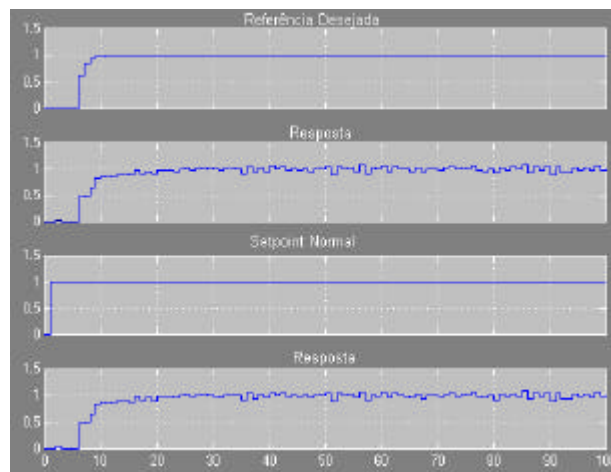


Figura 3.8. Resultados utilizando a referência desejada (dois primeiros gráficos) comparativamente com a mudança normal de *setpoint*. O erro da saída é bem menor no caso da referência desejada.

3.2.7. Demais questões práticas

O procedimento de avaliação de desempenho requer, para sua correta interpretação, que os dados sendo analisados sejam representativos do processo. Dados coletados durante operações de manutenção, testes com a malha, ou equipamento com problemas podem levar a conclusões errôneas sobre o desempenho do controlador. Problemas em equipamentos podem ser detectados quando a análise de desempenho conta com técnicas de detecção de oscilações e demais técnicas auxiliares apresentadas no Capítulo 2.

Thornhill et al. (1999) também aborda questões relativas ao uso de dados quantizados ou arquivados por sistemas historiadores de dados. As bases de dados de tais sistemas de

arquivamento de dados trabalham com algoritmos de filtragem para remoção de ruído e algoritmos de compressão que inevitavelmente modificam os dados, com o objetivo de otimizar o espaço de armazenamento no banco de dados. Esta alteração nos dados acaba por alterar os valores dos índices de desempenho em relação aos dados originais do processo.

Dados “suavizados” por filtros tornam-se uma série temporal com maior predibilidade do que os não filtrados. Os índices de desempenho podem ser interpretados como uma medida da predibilidade dos dados, já que seqüências com maior autocorrelação indicam que o controlador não está realizando bem o seu papel, enquanto que seqüências com baixa autocorrelação (ou predibilidade) indicam que o controlador desempenha bem. Assim, a filtragem dos dados tem um impacto no índice na medida em que uma análise de desempenho dos dados filtrados indicaria que o controlador correspondente tem um desempenho pior do que o real.

Os algoritmos de compressão utilizam diversas regras para decidir se arquivam ou não um determinado ponto baseado nos valores dos pontos coletados anteriormente. Este procedimento pode vir a remover componentes predizíveis no conjunto de dados, prejudicando desta forma a avaliação de desempenho através dos índices.

Thornhill et al. (1999) aponta que o uso de dados quantizados afeta o cálculo do índice de desempenho pelo fato destes conterem erros de arredondamento e serem menos predizíveis que o sinal contínuo. Desta forma, a quantização dos dados faz com que as malhas apresentem uma falsa melhora no desempenho.

Outra questão importante surge quando se analisa os objetivos de controle para uma determinada malha. Uma malha pode apresentar desempenho considerado satisfatório para os objetivos de controle e no entanto acarretar grande diferença em relação à variância mínima. Malhas com índices de desempenho próximas a 1 podem ter desempenho bom aos olhos do pessoal de operação e controle, enquanto que malhas com índices mais baixos podem não satisfazer as necessidades de controle. É o caso, por exemplo, de malhas de controle de nível em tanques pulmão (índices altos não significam mau desempenho) ou malhas de vazão (índices relativamente baixos não significam bom desempenho). Desta forma, se a análise da qualidade de malhas puder envolver a experiência e o conhecimento do pessoal da operação na qualificação de determinadas malhas como sendo “satisfatórias” ou “ruins”, alguns erros de avaliação podem ser evitados.

Por fim um procedimento de tratamento de dados que remova dados ruins, tais como valores não numéricos, valores ausentes e pontos extremamente desviados da média (*outliers*), é indispensável em um sistema de avaliação de desempenho de grande escala e alto grau de automatização.

3.3. Aplicação em exemplos

Os objetivos desta seção são:

- Estudar o comportamento de controladores de variância mínima (CVM) quando aplicados para diversos sistemas;
- Comparar as respostas dos CVM's com as respostas dadas por PI's e PID's elaborados com algum método de ajuste de controladores, bem como comparar as respostas das saídas dos sistemas quando submetidos a estes controladores;
- Comparar a variância produzida pelos CVM e PI's/PID's com as estimativas de variância dadas pelos diversos algoritmos de cálculo do índice de desempenho apresentados no Capítulo 2;
- Verificar a máxima sensibilidade de todos os sistemas gerados, visando quantificar a robustez das malhas de controle;
- Avaliar para um determinado sistema o efeito de diversos modelos de distúrbios na qualidade do controlador e analisar a ação de controle gerada pelo CVM para estes distúrbios;
- Testar os diversos algoritmos de cálculo de índice de desempenho com dados provenientes de unidades industriais.

Com esta série de testes com dados simulados e industriais deseja-se formular mais questões a respeito do uso destas técnicas, a respeito de suas limitações e tirar conclusões de cunho prático a respeito das questões levantadas ao longo de toda a Seção 3.2.

3.3.1. Sistemas escolhidos

Para a realização da batelada de testes propostos foram escolhidos quatro sistemas que podem ser caracterizados como os que representam a maioria dos processos industriais mais comuns (Åström & Hägglund, 1998).

Tais sistemas estão representados pelas seguintes classes de funções de transferência contínuas:

$$G_1 = \frac{1}{(as + 1)^2} \quad \text{para } a = 0.1, 3 \text{ e } 10 \quad (3.7)$$

$$G_2 = \frac{1}{(s + 1)^4} \quad (3.8)$$

$$G_3 = \frac{1}{(s+1)(s+m)(s+m^2)(s+m^3)} \quad \text{para } m=0.2 \text{ e } 0.7 \quad (3.9)$$

$$G_4 = \frac{1-bs}{(s+1)^3} \quad \text{para } b=0.1, 0.5 \text{ e } 1.2 \quad (3.10)$$

Com os diferentes parâmetros para cada classe de função de transferência tem-se ao todo nove sistemas. Cada sistema foi discretizado com tempo de amostragem igual a uma unidade de tempo. As equações de (3.11) a (3.19) representam os sistemas discretizados. A convenção é que G_{11} é G_1 com $a = 0.1$, G_{12} é G_1 com $a = 3$, e assim por diante.

$$G_{11} = \frac{q + 0,0004086}{q^2 - 9,08 \cdot 10^{-5}q + 2,061 \cdot 10^{-9}} = \frac{q^{-2}(1 + 0,00041q^{-1})}{(1 - 4,54 \cdot 10^{-1}q^{-1})^2} \quad (3.11)$$

$$G_{12} = \frac{0,04462q + 0,03573}{q^2 - 1,433q + 0,5134} = \frac{0,0446q^{-2}(1 + 0,8q^{-1})}{(1 - 0,72q^{-1})^2} \quad (3.12)$$

$$G_{13} = \frac{0,004679q + 0,004377}{q^2 - 1,81q + 0,8187} = \frac{0,0047q^{-2}(1 + 0,94q^{-1})}{(1 - 0,905q^{-1})^2} \quad (3.13)$$

$$\begin{aligned} G_2 &= \frac{0,01899q^3 + 0,09595q^2 + 0,04301q + 0,00172}{q^4 - 1,472q^3 + 0,812q^2 - 0,1991q + 0,01832} = \\ &= \frac{0,019(q + 4,56)(q + 0,448)(q + 0,0443)}{q(q - 0,37)^4} \end{aligned} \quad (3.14)$$

$$\begin{aligned} G_{31} &= \frac{0,03284q^3 + 0,287q^2 + 0,224q + 0,01555}{q^4 - 3,139q^3 + 3,572q^2 - 1,719q + 0,2871} = \\ &= \frac{0,033(q + 7,88)(q + 0,78)(q + 0,077)}{q(q - 0,992)(q - 0,961)(q - 0,819)(q - 0,368)} \end{aligned} \quad (3.15)$$

$$\begin{aligned} G_{32} &= \frac{0,02534q^3 + 0,1706q^2 + 0,1028q + 0,005544}{q^4 - 2,187q^3 + 1,76q^2 - 0,6174q + 0,07942} = \\ &= \frac{0,025(q + 6,07)(q + 0,60)(q + 0,06)}{q(q - 0,71)(q - 0,61)(q - 0,50)(q - 0,37)} \end{aligned} \quad (3.16)$$

$$G_{41} = \frac{0,06191q^2 + 0,166q + 0,02465}{q^3 - 1,104q^2 + 0,406q - 0,04979} = \frac{0,062(q + 2,52)(q + 0,16)}{q(q - 0,37)^3} \quad (3.17)$$

$$\begin{aligned}
G_{42} &= \frac{-0,01167q^2 + 0,2125q + 0,05171}{q^3 - 1,104q^2 + 0,406q - 0,04979} = \\
&= \frac{-0,012(q - 18,45)(q + 0,24)}{q(q - 0,37)^3}
\end{aligned} \tag{3.18}$$

$$\begin{aligned}
G_{43} &= \frac{-0,1404q^2 + 0,2939q + 0,09908}{q^3 - 1,104q^2 + 0,406q - 0,04979} = \\
&= \frac{-0,140(q - 2,39)(q + 0,30)}{q(q - 0,37)^3}
\end{aligned} \tag{3.19}$$

Cada um dos sistemas pode ser representado na notação ARMAX:

$$A(q) y(k) = B(q) u(k - d) + C(q) e(k) \tag{3.20}$$

Onde d é o tempo morto e os polinômios A , B e C derivam diretamente dos modelos das equações (3.11) a (3.19) e de modelos para perturbações. O tempo morto de todos os sistemas foi escolhido como sendo dois intervalos de amostragem. Cada sistema G_{ij} pode ser representado como uma função impulso onde o numerador é o polinômio $B(q)$ e o denominador é $A(q)$. Da mesma forma, a função de transferência dos modelos para perturbações pode ser representada como uma função impulso com numerador $C(q)$ e denominador $A(q)$.

Para cada classe de sistema escolheu-se um modelo de distúrbio de forma que o numerador resultasse em um polinômio estável para o cálculo do controlador de variância mínima. Isso implica que o polinômio $C(q)$ deve ter todas suas raízes dentro ou sobre o círculo unitário.

Os modelos de distúrbio possuem o mesmo denominador dos modelos das equações (3.11) a (3.19), ou seja, o polinômio $A(q)$. Os polinômios $C(q)$ que compõe os modelos de distúrbios estão descritos pelas equações de (3.7) a (3.10).

$$C_1 = q^2 + 1,5q + 0,9 \tag{3.21}$$

$$C_2 = q^2 - 0,9q + 0,2 \tag{3.22}$$

$$C_3 = q^3 - 2,4q^2 + 1,91q - 0,504 \tag{3.23}$$

$$C_4 = q^2 - 1,1q + 0,3 \tag{3.24}$$

As ordem dos polinômios $C(q)$ podem ser aumentadas pela multiplicação sucessiva por q de forma a igualar-se a ordem dos polinômios $A(q)$. Este procedimento permite o

cálculo dos controladores de variância mínima através da resolução da equação de Diophantine apropriada e não altera a estrutura de correlação de $C(q)e(k)$ (Åström & Wittenmark, 1997).

Os polinômios C de cada sistema foram alterados por uma normalização da função de transferência do modelo de distúrbio de modo a terem ganho estacionário igual a 50 dada pela seguinte equação:

$$v(k) = 50 \frac{v(k)}{\text{dcgain}(v(k))}$$

onde $\text{dcgain}(v(k))$ é o ganho estacionário da função de transferência do modelo, $v(k) = [C(q)/A(q)]e(k)$.

Com esta normalização, todos os modelos de distúrbios gerados com os polinômios A e C têm o mesmo ganho estacionário, de 50 unidades. Desta forma, a comparação entre os diferentes modelos pode ser feita mais diretamente.

Com os modelos definidos, pode-se proceder à análise metódica dos mesmos. Em malha aberta todos os sistemas se demonstram estáveis, pois não possuem pólos positivos na versão contínua (ou fora do círculo unitário na versão discreta). Quanto aos zeros, exceto os sistemas da classe 1 (G_{1i} , $i = 1, 2, 3$), os sistemas discretos apresentam zeros fora do círculo unitário. As implicações desta característica serão exploradas posteriormente. Uma análise dos pólos e zeros dos sistemas discretos encontra-se na Tabela 3.1.

Tabela 3.1. Análise de pólos e zeros dos sistemas discretos.

| Sistema | Pólos | | | | Zeros | |
|----------|---------------------|---------------------|--------|--------|----------------------|---------|
| G_{11} | $4,5 \cdot 10^{-5}$ | $4,5 \cdot 10^{-5}$ | | | $-4,1 \cdot 10^{-4}$ | |
| G_{12} | 0,7165 | 0,7165 | | | -0,8 | |
| G_{13} | 0,9048 | 0,9048 | | | -0,9355 | |
| G_2 | $0,3680 + 0,0001i$ | $0,3680 - 0,0001i$ | | | -4,5607 | -0,4479 |
| | $0,3678 + 0,0001i$ | $0,3678 - 0,0001i$ | | | -0,0443 | |
| G_{31} | 0,9920 | 0,9608 | 0,8187 | 0,3679 | -7,8800 | -0,7811 |
| | | | | | -0,0769 | |
| G_{32} | 0,7096 | 0,6126 | 0,4966 | 0,3679 | -6,0681 | -0,6025 |
| | | | | | -0,0598 | |
| G_{41} | 0,3679 | 0,3679 | 0,3679 | | -2,5241 | -0,1577 |
| G_{42} | 0,3679 | 0,3679 | 0,3679 | | 18,4546 | -0,2402 |
| G_{43} | 0,3679 | 0,3679 | 0,3679 | | 2,3885 | -0,2954 |

Os sistemas das classes 2, 3 e 4 possuem zeros fora do círculo unitário. Isto implica que o cálculo do controlador de variância mínima para estes sistemas deve utilizar a equação (2.39), já que os zeros do sistema são as raízes do polinômio B .

Para cada sistema também foram projetados controladores PI e PID utilizando o método ITAE (Integral of Time Absolute Error, Ogunnaike e Ray, 1994) apropriado para eliminar distúrbios na carga.

Os sistemas foram simulados em controle *feedback* empregando-se o Matlab/Simulink (Mathworks). Para implementar os sistemas com uma abordagem polinomial foi implementada em Matlab uma classe de polinômios (Apêndice E), com diversos métodos que permitem manipular polinômios, com a finalidade de facilitar a realização do conjunto de testes. Esta classe contém funções que permitem entre outras coisas, todas as operações básicas entre polinômios, a resolução de equações de Diophantine, cálculo de mínimo divisor comum, fatoração de polinômios, avaliação do polinômio em pontos, conversão em objetos da classe LTI do Matlab (objetos TF) e cálculo de raízes. O diagrama de blocos da malha simulada encontra-se apresentado na Figura 3.9. Os resultados das simulações encontram-se na Figura 3.10 até a Figura 3.18.

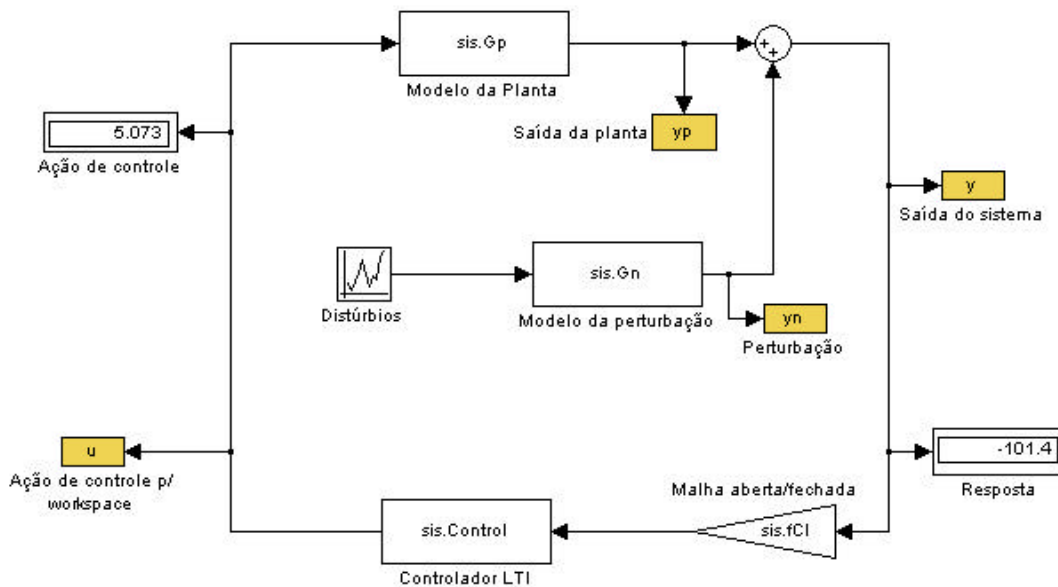


Figura 3.9. Malha *feedback* montada no Simulink® para simulação dos sistemas propostos.

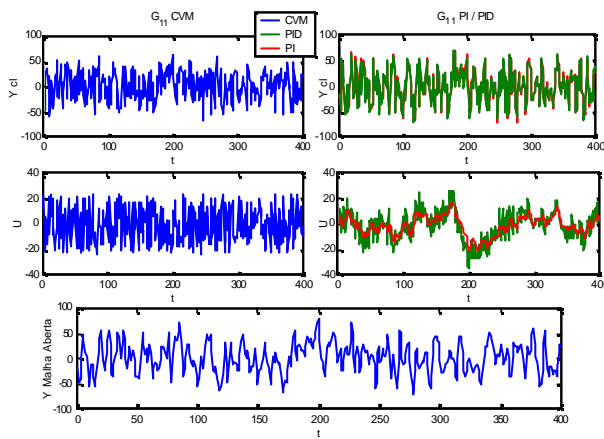


Figura 3.10. Resultados para o sistema G_{11} .

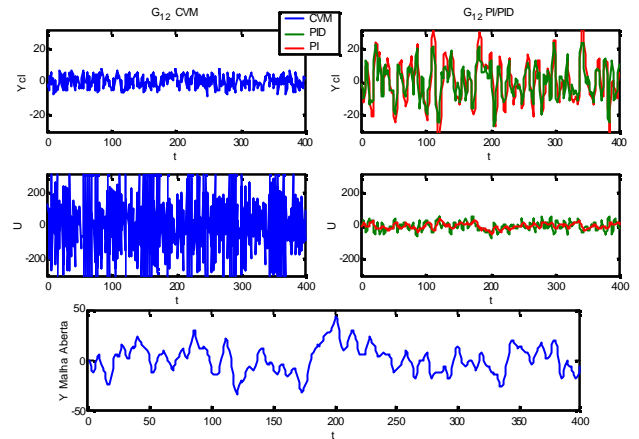


Figura 3.11. Resultados para o sistema G_{12} .

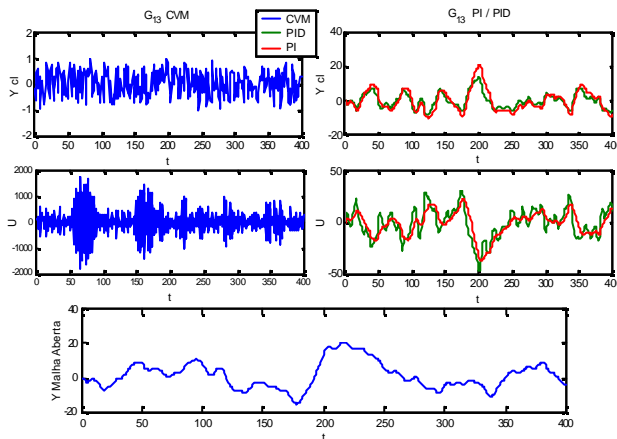


Figura 3.12. Resultados para o sistema G_{13} .

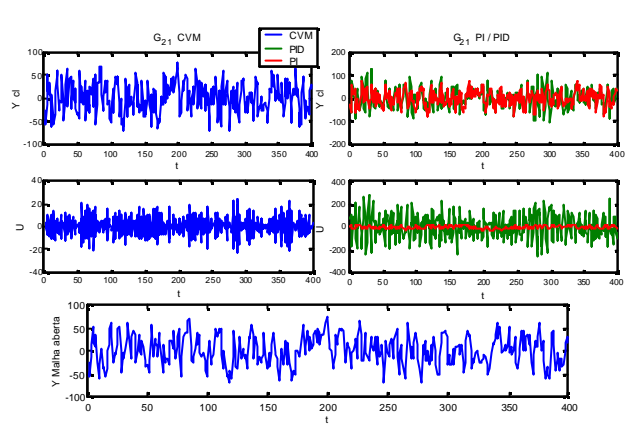


Figura 3.13. Resultados para o sistema G_{21} .

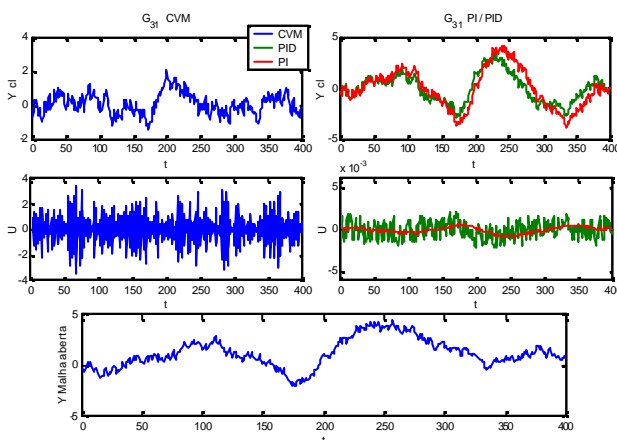


Figura 3.14. Resultados para o sistema G_{31} .

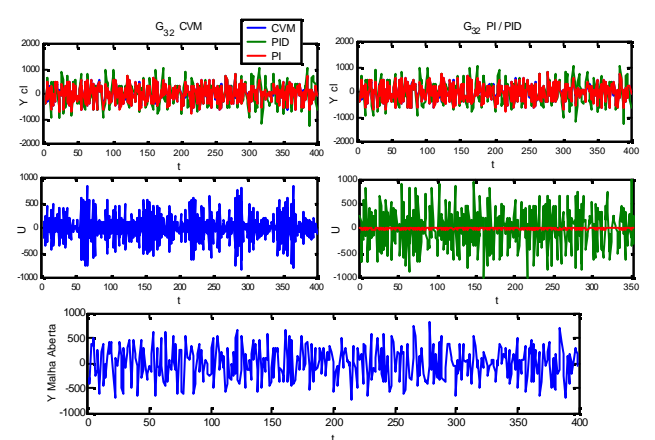


Figura 3.15. Resultados para o sistema G_{32} .

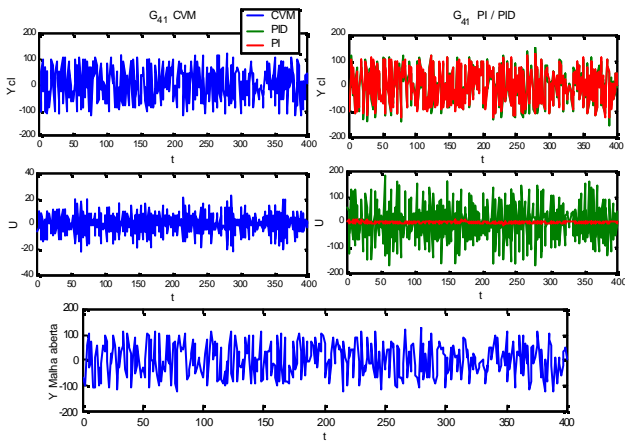


Figura 3.16. Resultados para o sistema G_{41} .

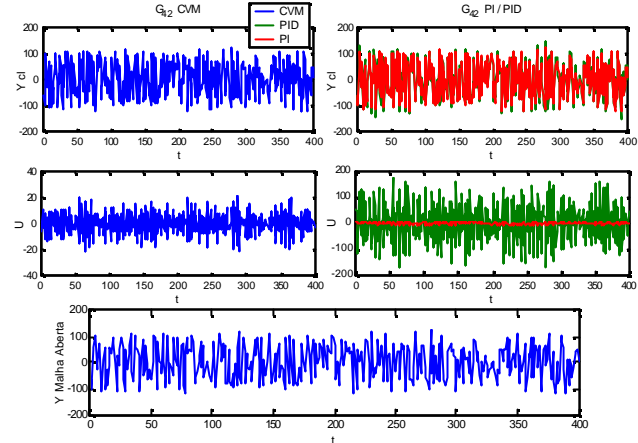


Figura 3.17. Resultados para o sistema G_{42} .

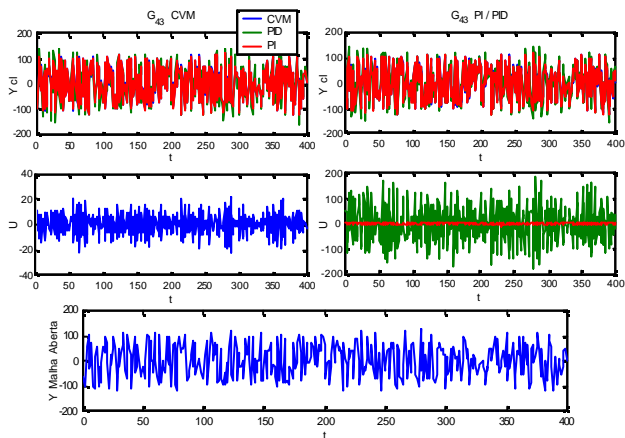


Figura 3.18. Resultados para o sistema G_{43} .

3.3.2. Análise dos resultados

Uma análise visual dos gráficos da Figura 3.10 até a Figura 3.18 mostra que as ações de controle do CVM são bastante bruscas, tornando praticamente inviável seu emprego na maioria dos casos práticos.

Para os sistemas com inversa instável, curiosamente o ajuste de PID produziu controladores com ações de controle piores que as do CVM e saída do sistema com variância muito mais alta que os PI. Isto pode ser uma característica inerente ao algoritmo de ajuste de PID's baseado no método ITAE empregado.

Também é possível constatar visualmente que o CVM resultou em menor variância que PI/PID's, especialmente nos sistemas da classe 1 e o G_{31} . Os sistemas da classe 4 e o G_{32} já encontram-se próximos da variância mínima em malha aberta, fazendo com que a redução de variância fosse insignificante para os três tipos de controladores.

A Tabela 3.2 mostra os resultados do cálculo de variância da saída de todos os sistemas.

Tabela 3.2. Variância da saída dos sistemas sob ação dos controladores e em malha aberta.

| Sistemas | CVM | PI | PID | Malha Aberta |
|----------|--------|--------|--------|--------------|
| G_{11} | 714,18 | 926,76 | 894,71 | 885,63 |
| G_{12} | 13,56 | 166,25 | 96,48 | 191,03 |
| G_{13} | 0,2143 | 38,83 | 19,62 | 61,46 |
| G_2 | 988,35 | 1085,4 | 1717,1 | 990,32 |
| G_{31} | 0,3650 | 3,642 | 2,052 | 1,988 |
| G_{32} | 105820 | 109510 | 184390 | 107460 |
| G_{41} | 4065,3 | 4102,1 | 4327,0 | 4073,6 |
| G_{42} | 4073,6 | 4099,4 | 4357,4 | 4073,6 |
| G_{43} | 4082,1 | 4095,1 | 4557,2 | 4073,6 |

Nota-se que no caso do sistema G_{43} o sistema em malha aberta obteve a menor variância, sendo menor inclusive que a variância mínima. Conclui-se então que este sistema encontra-se com mínima variância em malha aberta e a presença de qualquer ação de controle, por mínima que seja, só pode aumentar a variância.

Uma comparação da saída da função de transferência $v(k) = [C / A]e(k)$ (em outras palavras, do sistema em malha aberta) com a saída em malha fechada mostra que ambas respostas são iguais, o que significa que v não altera o sinal do ruído branco. Qin (1998) aponta que se o distúrbio é ruído branco, a variância mínima é atingida em malha aberta, o que corrobora com o resultado prático obtido. A Figura 3.19 ilustra este ponto. Nela são mostradas a saída do sistema em malha fechada e a saída de $v(k)$, onde se pode observar que os dois sinais são iguais.

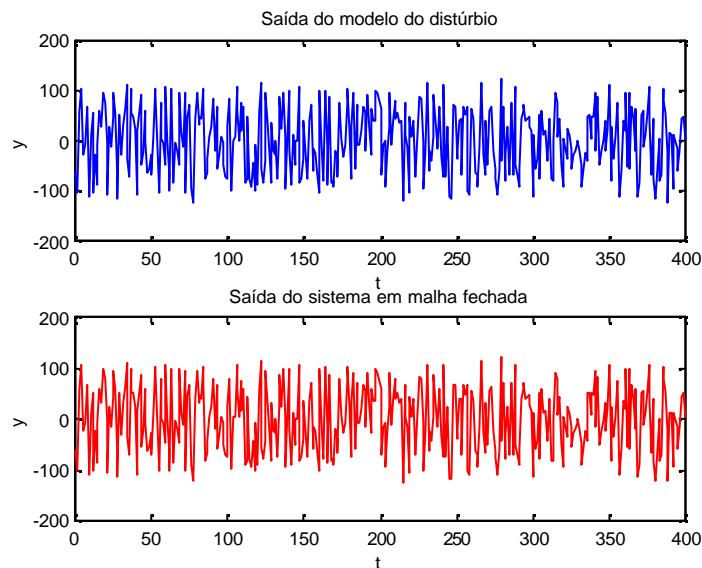


Figura 3.19. Comparação da saída do modelo de distúrbio e a saída do sistema em malha fechada para o sistema G_{43}

Com estes dados de variâncias torna-se interessante comparar as variâncias produzidas pelos sistemas sob ação de CVM com as variâncias que são estimadas pelos índices de desempenho tradicionais, como os descritos no Capítulo 2.

As ordens das estimativas utilizadas nos cálculos dos índices de desempenho foram as mesmas ordens dos polinômios $A(q)$ e $C(q)$ simulados. Na Tabela 3.3, para a estimação da variância mínima pelos métodos de avaliação (AGENER, ASUBS, ANORM e FCOR) foram empregados dados produzidos pelos CVM calculados para cada sistema

Tabela 3.3. Variâncias produzidas pelo CVM comparadas com as variâncias estimadas resultantes das análises de índices de desempenho com as saídas de um CVM.

| Sistemas | S^2_{CVM} | S^2_{AGENER} | % _{cvm} | S^2_{ASUBS} | % _{cvm} | S^2_{ANORM} | % _{cvm} | S^2_{FCOR} | % _{cvm} |
|-------------------|-------------|----------------|------------------|---------------|------------------|---------------|------------------|--------------|------------------|
| G_{11} | 714,18 | 717,03 | 0,40 | 714,64 | 0,06 | 714,77 | 0,08 | 730,32 | 2,26 |
| G_{12} | 13,56 | 13,63 | 0,52 | 13,58 | 0,12 | 13,58 | 0,08 | 13,56 | 0,04 |
| G_{13} | 0,2143 | 0,2150 | 0,34 | 0,2147 | 0,18 | 0,2144 | 0,07 | 0,2145 | 0,15 |
| G_2 | 988,35 | 968,03 | 2,06 | 962,01 | 2,66 | 980,92 | 0,75 | 970,32 | 1,82 |
| G_{31} | 0,3650 | 0,1746 | 52,2 | 0,1810 | 50,4 | 0,1768 | 51,6 | 0,1733 | 52,5 |
| G_{32} | 105820 | 102700 | 2,95 | 103080 | 2,59 | 104450 | 1,29 | 105700 | 0,12 |
| G_{41} | 4065,3 | 4031,4 | 0,83 | 4053,6 | 0,29 | 4063,3 | 0,05 | 4025,3 | 0,98 |
| G_{42} | 4073,6 | 4039,6 | 0,77 | 4057,0 | 0,34 | 4062,0 | 0,22 | 4048,5 | 0,55 |
| G_{43} | 4082,1 | 4055,8 | 0,64 | 4063,9 | 0,44 | 4075,2 | 0,17 | 4061,1 | 0,52 |
| <i>Erro médio</i> | - | - | 6,75 | - | 6,34 | - | 6,03 | - | 6,55 |
| <i>Erro máx.</i> | - | - | 52,2 | - | 50,4 | - | 51,6 | - | 52,5 |
| <i>Erro mín.</i> | - | - | 0,34 | - | 0,06 | - | 0,05 | - | 0,04 |

Na Tabela 3.3 as respostas inversas dos sistemas com zeros fora do círculo unitário foram incorporadas como tempos-mortos no cálculo dos índices de desempenho. Denominando-se w um zero fora do círculo unitário qualquer, pode-se demonstrar (Harris et. al., 1999), que processos com apenas um zero fora do círculo unitário, $w \leq -1$, contribuem com uma unidade de acréscimo no tempo morto, enquanto que processos com $w \geq 1$ incrementam o tempo morto em $w / (w - 1)$ unidades.

Nota-se que há uma discrepância entre a variância real produzida pelo controlador de variância mínima em relação à variância estimada pelas análises de índice de desempenho no sistema G_{31} . A explicação para isto está no fato de que o modelo do distúrbio deste sistema produz perturbações muito não estacionárias (modelo ARIMA) que não conseguem ser capturadas pelos modelos ARMA ou AR utilizados pelas metodologias implementadas.

Os resultados para os sistemas da classe 1 mostram-se compatíveis com os resultados produzidos pela aplicação do CVM, considerando-se que os sistemas da classe 1 não possuem zeros instáveis. Para este caso, a estimação de parâmetros dos modelos ARMA mostra-se bastante acurados em relação ao controlador de variância mínima.

Também se torna interessante avaliar se as metodologias de avaliação de desempenho estimam a variância mínima corretamente quando os dados fornecidos não são de um CVM, mas de um PI, por exemplo. A Tabela 3.4 estabelece este comparativo. Na estimação da

variância mínima pelos métodos de avaliação foram empregados dados produzidos pelos PI ajustados para cada sistema.

Tabela 3.4. Variâncias mínimas estimadas pelos diferentes métodos de avaliação quando alimentados com dados provenientes de um PI, em comparação com a variância do controlador de variância mínima.

| Sistemas | S^2_{CVM} | S^2_{AGENER} | % _{cvm} | S^2_{ASUBS} | % _{cvm} | S^2_{ANORM} | % _{cvm} | S^2_{FCOR} | % _{cvm} |
|-------------------|-------------|----------------|------------------|---------------|------------------|---------------|------------------|--------------|------------------|
| G_{11} | 714,18 | 800,27 | 12,1 | 789,28 | 10,5 | 839,01 | 17,5 | 879,13 | 23,2 |
| G_{12} | 13,56 | 15,50 | 14,3 | 29,13 | 114,8 | 24,59 | 81,3 | 25,69 | 89,4 |
| G_{13} | 0,2143 | 0,2364 | 10,3 | 0,5809 | 171 | 0,484 | 126 | 0,5406 | 152 |
| G_2 | 988,35 | 983,74 | 0,47 | 984,7 | 0,37 | 1025,4 | 3,75 | 1027,5 | 3,96 |
| G_{31} | 0,3650 | 0,1812 | 50,4 | 0,1964 | 46,2 | 0,2131 | 41,6 | 0,2127 | 41,7 |
| G_{32} | 105820 | 104280 | 1,46 | 103201 | 2,47 | 107070 | 1,18 | 108640 | 2,67 |
| G_{41} | 4065,3 | 4050,1 | 0,37 | 4084,5 | 0,47 | 4090,9 | 0,63 | 4059,2 | 0,15 |
| G_{42} | 4073,6 | 4047,4 | 0,58 | 4079,9 | 0,22 | 4075,1 | 0,10 | 4071,0 | 0,00 |
| G_{43} | 4082,1 | 4046,4 | 0,87 | 4077,1 | 0,12 | 4069,3 | 0,31 | 4067,0 | 0,37 |
| <i>Erro médio</i> | - | - | 10,09 | - | 38,46 | - | 30,26 | - | 34,83 |
| <i>Erro máx.</i> | - | - | 50,4 | - | 171 | - | 126 | - | 152 |
| <i>Erro mín.</i> | - | - | 0,37 | - | 0,12 | - | 0,1 | - | 0,0 |

Na Tabela 3.4, as mesmas observações quanto a inversibilidade dos modelos são válidas. Pode-se concluir, desta tabela, que apesar de haver discrepâncias entre os valores reais de variância mínima e os valores estimados, estas diferenças não chegam a desvalorizar as conclusões que podem ser feitas acima dos índices de desempenho obtidos.

Uma observação importante a ser feita aqui é que as ordens dos modelos ARMA/AR utilizadas pelos índices de desempenho na Tabela 3.4 foram as mesmas dos denominadores dos sistemas (3.7) a (3.10). Melhores resultados são conseguidos se utiliza-se ordens entre 5 e 10 para modelar os dados, com uma conseqüente redução na discrepância entre as variâncias mínimas estimadas e a variância mínima real. Sem o uso de ordens mais elevadas, os sistemas com perturbações menos estacionárias (caso dos sistemas G_{12} , G_{13} e G_{31}) causaram os maiores erros de estimação da variância mínima no caso da Tabela 3.4 (saída produzida por um PI), a ponto de reduzir a confiança nos resultados obtidos para este tipo de malha. No entanto, na prática, costuma-se empregar ordens maiores para a estimação, de forma que o erro para o modelo G_{13} cai de 126% para 14,8% ao se utilizar ordem 7 com o algoritmo ANORM.

Na Tabela 3.3 e na Tabela 3.4 o método baseado em identificação no subespaço de estado empregou para a estimação 14 matrizes-bloco de Hankel. A determinação adequada deste parâmetro (número de matrizes-bloco) é um outro problema que o método de identificação no subespaço de estado introduz na automatização do monitoramento do desempenho de malhas, fazendo com que seu emprego prático seja preterido pelos métodos baseados em identificação ARMA/AR. A idéia da utilização do método possui grande potencial, mas técnicas para determinação do número de matrizes-bloco ainda precisam ser desenvolvidas para a apropriada utilização deste potencial.

Cabe salientar ainda, que estes resultados abrem a possibilidade de em análises de desempenho de controladores surgirem índices negativos, pois a variância produzida por um determinado controlador pode ser menor que a variância mínima estimada, que está acima da variância mínima real.

Uma análise dos percentuais também permite que se recomende cautela quanto a confiança excessiva no resultado dos índices de desempenho, uma vez que os erros percentuais das estimativas de variância mínima em relação à variância mínima real foram consideravelmente grandes. Estes erros médios elevados devem-se especialmente ao sistema G_{31} . Este sistema apresenta um comportamento anômalo na estimação da variância mínima. A explicação deste fato pode ser obtida com a análise da resposta ao degrau deste sistema (representada na Figura 3.20).

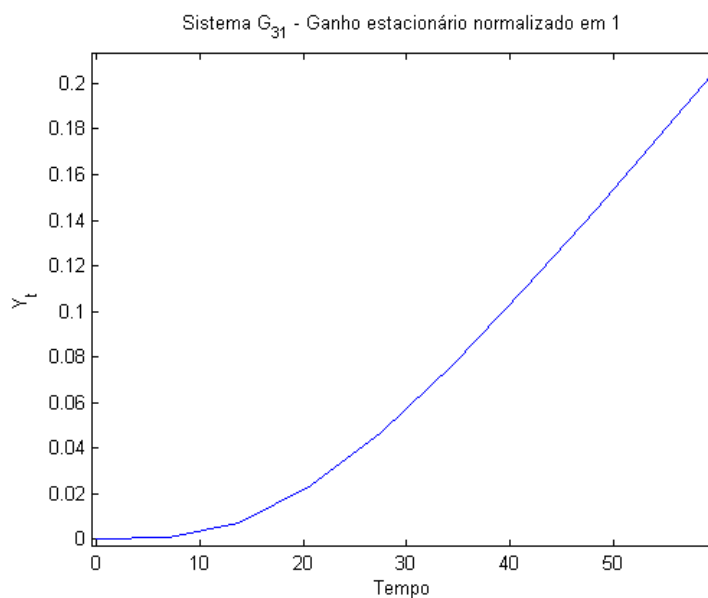


Figura 3.20. Pedaco da resposta ao impulso do sistema G_{31} . O ganho estacionário do sistema foi normalizado em uma unidade.

A resposta ao degrau demonstra que este sistema possui um tempo morto “aparente” muito maior do que o tempo morto real (2 intervalos de amostragem). Este comportamento é devido especialmente à sua alta ordem combinada com os zeros mais próximos da origem que o sistema G_{32} por exemplo. Para resultados corretos na avaliação do desempenho de uma malha de controle com este sistema, deve-se considerar um tempo morto da ordem de 15 a 25 intervalos de amostragem. De fato, ao estimar-se o índice de desempenho deste sistema com o método ANORM com ordem do modelo autorregressivo igual a 4 e considerando-se o tempo morto total como 20 intervalos de amostragem, os erros gerados são de 2,98% quando o método é alimentado com a saída do sistema controlado por um CVM. Com a exceção dos sistemas com este tipo de comportamento, em que maior atenção é necessária, as demais classes de sistemas apresentam erros bastante aceitáveis, da ordem de até 8%, conforme ilustra a Tabela 3.5, onde os resultados do sistema G_{31} são omitidos. Assim, pode-se concluir que, para uma ampla classe de sistemas os resultados de avaliação de desempenho com as abordagens apresentadas são bastante confiáveis. A exceção é uma restrita classe de sistemas de altas ordens e zeros perto da origem.

Tabela 3.5. Erros médios de estimação da variância mínima desconsiderando o sistema G_{31} .

| | AGENER | ASUBS | ANORM | FCOR |
|-----|---------------|--------------|--------------|-------------|
| CVM | 7,50 | 1,48 | 0,78 | 2,44 |
| PI | 6,52 | 5,06 | 4,02 | 3,58 |

3.3.3. Efeito do modelo de distúrbio

O objetivo desta seção é verificar o efeito que diferentes modelos de distúrbios (polinômios C diversos) têm sobre a qualidade do controlador. Uma análise da ação de controle gerada pelo CVM e uma comparação com a variância estimada pelos índices de desempenho também é feita.

Foi escolhido um sistema da classe 1 pelo fato de os sistemas desta classe não apresentarem problemas de estabilidade e a relação entre o controlador de variância mínima e os índices de desempenho ser direta. Mais especificamente selecionou-se o sistema G_{11} . O tempo morto empregado foi de 2 tempos de amostragem.

Os diversos modelos de distúrbios, retirados de Qin, 1998, são dados por:

$$N_1 = \frac{1 - 0,2q^{-1}}{(1 - 0,3q^{-1})(1 + 0,4q^{-1})(1 - 0,5q^{-1})}$$

$$N_2 = \frac{1 + 0,6q^{-1}}{(1 - 0,6q^{-1})(1 - 0,5q^{-1})(1 + 0,7q^{-1})}$$

$$N_3 = \frac{1 - 0,2q^{-1}}{(1 - 0,3q^{-1})(1 + 0,4q^{-1})(1 - 0,5q^{-1})(1 - q^{-1})}$$

$$N_4 = \frac{1 + 0,6q^{-1}}{(1 - 0,6q^{-1})(1 - 0,5q^{-1})(1 + 0,7q^{-1})(1 - q^{-1})}$$

Os resultados das simulações encontram-se na Figura 3.21 à Figura 3.24, e as variâncias obtidas via simulação e aplicação de índices de desempenho para cada caso encontram-se na Tabela 3.6.

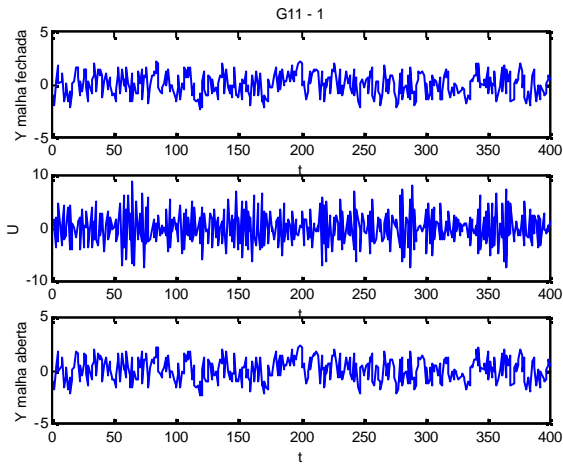


Figura 3.21. Sistema G_{11} com distúrbio N_1 .

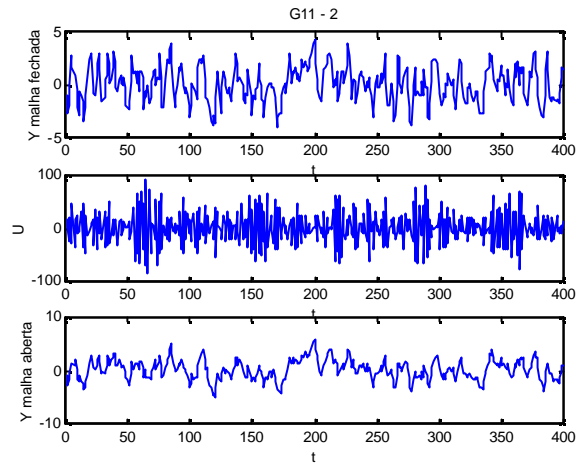


Figura 3.22. Sistema G_{11} com distúrbio N_2 .

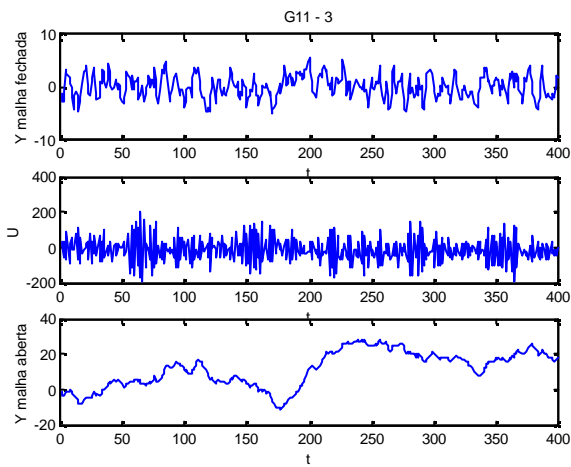


Figura 3.23. Sistema G_{11} com distúrbio N_3 .

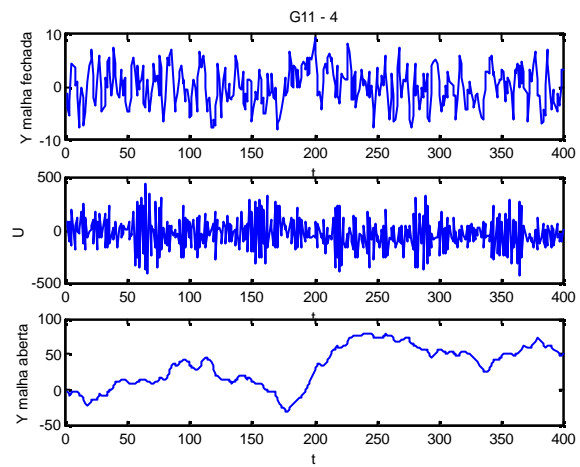


Figura 3.24. Sistema G_{11} com distúrbio N_4 .

Tabela 3.6. Resultados para as variâncias do sistema G_{11} com diferentes modelos de distúrbio sob ação de um CVM.

| Sistemas | S^2_{CVM} | S^2_{AGENER} | $\%_{cvm}$ | S^2_{ASUBS} | $\%_{cvm}$ | S^2_{ANORM} | $\%_{cvm}$ | S^2_{FCOR} | $\%_{cvm}$ | $S^2_{Malha\ aberta}$ | $\%_{cvm}$ |
|-------------------|-------------|----------------|-------------|---------------|-------------|---------------|-------------|--------------|-------------|-----------------------|--------------|
| G_{11-1} | 1,047 | 1,041 | 0,72 | 1,053 | 0,44 | 1,048 | 0,02 | 1,045 | 0,37 | 1,106 | 5,51 |
| G_{11-2} | 2,035 | 2,074 | 1,92 | 2,045 | 0,52 | 2,036 | 0,07 | 2,118 | 4,07 | 3,411 | 67,6 |
| G_{11-3} | 2,182 | 2,492 | 0,41 | 2,487 | 0,23 | 2,484 | 0,08 | 2,571 | 3,59 | 100,0 | 3931 |
| G_{11-4} | 5,072 | 5,122 | 0,99 | 5,074 | 0,04 | 5,076 | 0,09 | 5,113 | 0,80 | 821,6 | 16099 |
| <i>Erro médio</i> | - | - | 1,01 | - | 0,31 | - | 0,07 | - | 2,21 | | 5026 |
| <i>Erro máx.</i> | - | - | 1,92 | - | 0,52 | - | 0,09 | - | 4,07 | | 16099 |
| <i>Erro mín.</i> | - | - | 0,41 | - | 0,04 | - | 0,02 | - | 0,37 | | 5,51 |

Tabela 3.7. Resultados para as variâncias do sistema G_{11} com diferentes modelos de distúrbio sob ação de um controlador PI ajustado por um método baseado no ITAE.

| Sistemas | $S^2_{REAL PI}$ | S^2_{AGENER} | $\%_{cvm}$ | S^2_{ASUBS} | $\%_{cvm}$ | S^2_{ANORM} | $\%_{cvm}$ | S^2_{FCOR} | $\%_{cvm}$ |
|-------------------|-----------------|----------------|-------------|---------------|--------------|---------------|--------------|--------------|-------------|
| G_{11-1} | 1,337 | 1,169 | 11,5 | 1,134 | 8,15 | 1,246 | 18,8 | 1,250 | 19,2 |
| G_{11-2} | 4,245 | 1,170 | 6,66 | 2,285 | 12,3 | 2,390 | 17,5 | 2,430 | 19,3 |
| G_{11-3} | 10,13 | 2,455 | 1,07 | 2,567 | 3,43 | 2,610 | 5,15 | 2,653 | 6,91 |
| G_{11-4} | 68,15 | 5,125 | 1,04 | 7,135 | 40,7 | 5,748 | 13,3 | 5,792 | 14,2 |
| <i>Erro médio</i> | - | - | 5,07 | - | 16,15 | - | 13,69 | - | 14,9 |
| <i>Erro máx.</i> | - | - | 11,5 | - | 40,7 | - | 18,8 | - | 19,3 |
| <i>Erro mín.</i> | - | - | 1,04 | - | 3,43 | - | 5,15 | - | 6,91 |

Os resultados mostram claramente que o modelo que se estima para o distúrbio tem grande influência na qualidade da saída do processo e que o acréscimo de pólos em $q^{-1} = 1$ deste modelo aumenta drasticamente a variabilidade do processo. Estes pólos fazem com que o distúrbio seja não-estacionário (média móvel integrada), aumentando a dificuldade em reduzir a variância. A dinâmica não aleatória ocasionada pela presença destes integradores é evidenciada através da análise dos gráficos da resposta em malha aberta dos sistemas com distúrbio N_3 e N_4 (Figura 3.23 e Figura 3.24). As ações de controle do CVM também se tornam mais agressivas na medida em que as perturbações se tornaram mais não-estacionárias.

No entanto, a adição destes integradores no modelo do ruído fazem com que o controlador calculado tenha ação integral, uma vez que o sistema G_{11} por si só não possui esta característica. Isto traz como vantagem o fato de o CVM rejeitar distúrbios não aleatórios e mudanças de ponto de operação. A Figura 3.25 e a Figura 3.28 evidenciam esta característica dos sistemas G_{11-3} e G_{11-4} que rejeitam um distúrbio na carga, enquanto que os sistemas G_{11-1} e G_{11-2} não conseguem trazer a saída de volta ao ponto de operação.

A Tabela 3.7 mostra ainda que os distúrbios não exercem grande influência na qualidade do índice de desempenho estimado para a malha pelos métodos apresentados. Ao examinar-se, por exemplo, o caso do sistema G_{11-3} , seu desempenho teórico seria dado pela comparação da variância real produzida na saída pela ação de um PI (10,13) com a variância que seria produzida por um controlador de variância mínima calculado com base no modelo exato do sistema (2,182). Subtraindo-se da unidade a razão entre estas variâncias obtêm-se o índice real de desempenho: 0,78. No entanto, a abordagem normalizada (ANORM) estimou para a variância mínima um valor de 2,610, o que levaria a um índice de 0,74. Esta diferença não altera a interpretação do índice ou as conclusões que se chegaria quanto a qualidade do controle para esta malha.

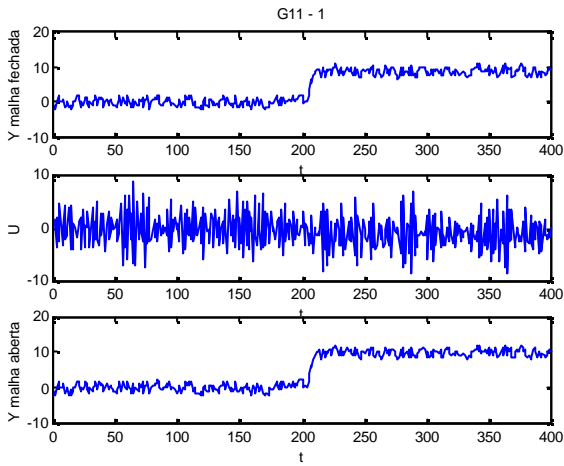


Figura 3.25. Sistema G_{11} com distúrbio N_1 e distúrbio na carga em $t = 200$.

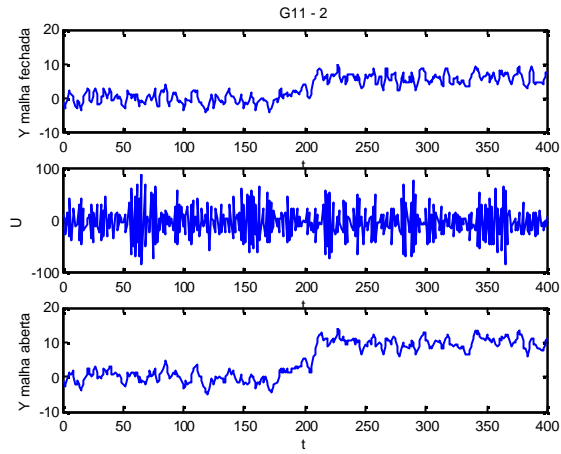


Figura 3.26. Sistema G_{11} com distúrbio N_2 e distúrbio na carga em $t = 200$.

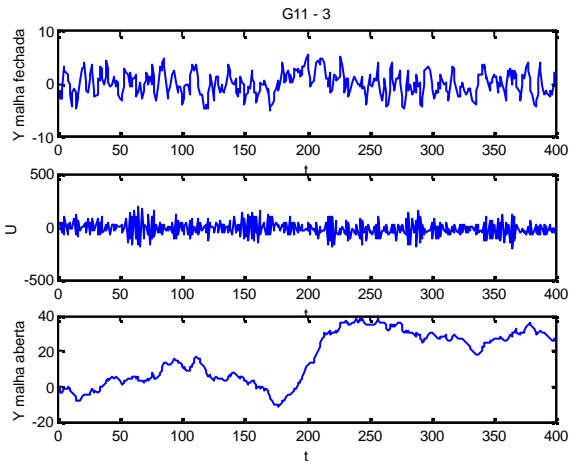


Figura 3.27. Sistema G_{11} com distúrbio N_3 e distúrbio na carga em $t = 200$.

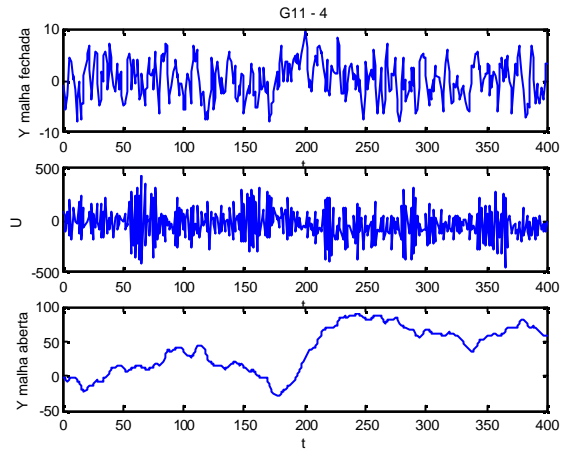


Figura 3.28. Sistema G_{11} com distúrbio N_4 e distúrbio na carga em $t = 200$.

3.3.4. Máxima sensibilidade

A análise da máxima sensibilidade permite que se tenha uma idéia bastante correta a cerca da robustez de um determinado sistema. Seu valor representa o inverso da menor distância do ponto -1 que a resposta do sistema em malha fechada, no domínio da frequência, produz, ao observar-se um gráfico de Nyquist. Quanto menor a distância mínima do ponto -1 que um sistema apresentar, maior será sua suscetibilidade à instabilidade, pois menor vai se tornando sua margem de ganho. Se a resposta intercepta o ponto -1 , o sistema é dito marginalmente estável e se a curva da resposta engloba o ponto -1 , o sistema é instável. A função de sensibilidade pode ser esquematizada através do diagrama da Figura 3.29. A função complementar de sensibilidade, P , de um ponto qualquer da curva $L(j\omega)$ para a frequência ω_i é dada pela diferença entre os dois segmentos Q e R:

$$\begin{aligned} P &= Q - R \\ &= L(j\omega_i) + 1 \end{aligned} \quad (3.25)$$

O valor mínimo da distância do ponto -1 (ou seja, do segmento P) pode ser representado como:

$$P_{MIN} = \min_w |S(jw)| = \min_w |1 + L(jw)| \quad (3.26)$$

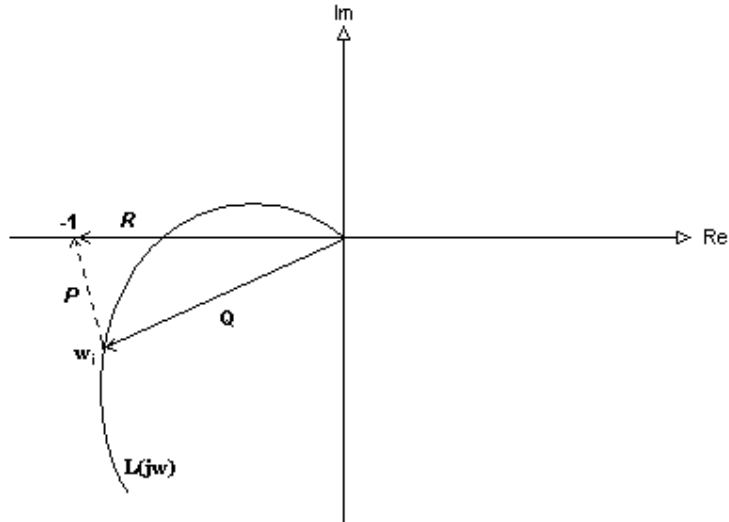


Figura 3.29. Diagrama de Nyquist ilustrando a função complementar de sensibilidade, P .

Equivalentemente, pode-se tomar o valor máximo do inverso da função P :

$$P_{MIN} = \max_w \left| \frac{1}{P(jw)} \right| = \max_w \left| \frac{1}{1 + L(jw)} \right| \quad (3.27)$$

Para um sistema em malha fechada composto de uma planta T , com um controlador Q e um distúrbio aditivo representado por um sinal $e(k)$, a função de sensibilidade, S , é a função de transferência do distúrbio para a saída:

$$S = \frac{1}{QT + 1} = \frac{1}{1 + L} \quad (3.28)$$

Então, a máxima sensibilidade, M_S , é dada por:

$$M_S = \max_w |S(jw)| = \frac{1}{P_{MIN}} \quad (3.29)$$

Onde P_{MIN} é a menor distância do ponto -1 no diagrama de Nyquist. Assim, um sistema cuja resposta no domínio da frequência passa próximo ao ponto -1 tem alto valor da máxima sensibilidade, enquanto que um sistema cuja menor distância de -1 é grande, tem baixo valor de máxima sensibilidade.

Os valores da Tabela 3.8 foram obtidos para os nove sistemas utilizados na análise de variância e mostram que, especialmente para os sistemas da classe 1, o controlador de variância mínima resulta muito sensível a variações na dinâmica do processo. Assim mostra-se que a baixa robustez deste tipo de controlador é um grande empecilho para seu emprego real. Novamente, este fato deve-se a inversão completa do sistema quando este é inversível, causada pelo CVM. Os sistemas das demais classes por serem não completamente inversíveis, resultariam em CVM's mais robustos, mas não a ponto de serem mais robustos que os controladores do tipo PI/PID, como mostra a Tabela 3.8. A explicação para tal comportamento deve estar no fato de que os CVM obtidos a partir de distúrbios modelados como ARMA não resultam como controladores com ação integral. A ação integral é a responsável pelo acréscimo em sensibilidade nos controladores PI/PID.

Tabela 3.8. Máxima sensibilidade dos sistemas em malha fechada.

| Sistemas | CVM | PI | PID |
|----------|-------|------|------|
| G_{11} | 8,64 | 1,06 | 1,52 |
| G_{12} | 50,17 | 1,56 | 1,98 |
| G_{13} | 71,26 | 1,59 | 1,33 |
| G_2 | 1,01 | 1,21 | 2,94 |
| G_{31} | 1,12 | 1,68 | 1,21 |
| G_{32} | 1,12 | 1,28 | 5,02 |
| G_{41} | 1,03 | 1,03 | 1,36 |
| G_{42} | 1,02 | 1,03 | 1,45 |
| G_{43} | 1,01 | 1,04 | 1,66 |

3.3.5. Aplicação com dados de processo industrial

Dados de operação de uma unidade FCC da Refinaria de Capuava (RECAP - Petrobrás) foram utilizados para testar a aplicação industrial dos índices de avaliação de malhas bem como levantar questões práticas de sua implementação.

Os dados foram coletados em duas etapas: um conjunto antes e um conjunto após um reajuste dos parâmetros do controlador da malha. Todos os dados coletados eram representativos de um período de 12 horas de operação sob condições normais do processo, com tempo de amostragem de 20 segundos (totalizando 2160 pontos) adquiridos através da base de dados do PI (OSI Software).

Foram analisadas malhas de diferentes tipos, tais como temperatura, pressão, nível e vazão, com o objetivo de determinar-se parâmetros genéricos para as metodologias, eliminando-se a necessidade de levantar valores para cada malha individualmente. A determinação de valores-padrão para cada tipo de malha permite que um procedimento de monitoramento das malhas de controle seja implementado de modo automático, sem a necessidade de haver alguém do corpo de engenharia inspecionando malha por malha. O desejável é aplicar-se os valores genéricos para todas as malhas e deixar que o sistema de supervisionamento descubra os pontos problemáticos. Assim, para cada tipo de malha observou-se:

- a sensibilidade à escolha da ordem dos modelos AR/ARMA/MA empregados nos índices de desempenho;
- um tempo morto aproximado representativo para cada tipo de malha;
- valores característicos do índice de desempenho para cada tipo de malha nos casos em que o controlador está satisfatoriamente ajustado e mal ajustado;
- Sensibilidade do procedimento de detecção de oscilação aos parâmetros de cada tipo de malha, tais como a faixa (range) apropriada da variável de processo;

A Tabela 3.9 ilustra os resultados obtidos para algumas malhas de tipos diferentes utilizando controladores PI em condições representativas de processo normal, antes e depois de uma re-sintonia dos controladores.

Tabela 3.9. Índices de desempenho de diferentes malhas antes e depois de sua re-sintonia.

| Método: | Temperatura | | Pressão | | Nível | | Vazão | |
|---------|-------------|--------|---------|--------|-------|--------|-------|--------|
| | Antes | Depois | Antes | Depois | Antes | Depois | Antes | Depois |
| ANORM | 0,83 | 0,69 | 0,98 | 0,20 | 0,98 | 0,70 | 0,72 | 0,09 |
| FCOR | 0,88 | 0,77 | 0,96 | 0,17 | 0,99 | 0,72 | 0,71 | 0,08 |

Os valores observados na Tabela 3.9, para cada tipo de malha, encontram reprodutibilidade quando aplicados a outras malhas do mesmo tipo, indicando por exemplo, que valores típicos de uma malha de vazão bem sintonizada devem ser próximos a zero e que para malhas de temperatura o ganho no reajuste dos parâmetros dos controladores é pequeno em termos de variação de $h(d)$. Para malhas de temperatura ou de nível (principalmente) os operadores relatam um bom desempenho mesmo com índices em torno de 0,7. Isto porque não é o objetivo de uma malha de nível de tanque-pulmão reduzir a variância da saída e sim mantê-la em torno do *setpoint* sem tomar ações de controle muito bruscas. Assim, uma comparação da variância de malhas de nível deste tipo com a variância mínima é injusta e a simples observação do índice de desempenho pode levar a conclusões enganosas quanto a necessidade de reajustar seus parâmetros.

Uma limitação fundamental na aplicação prática dos índices de avaliação de desempenho está na determinação de alguns parâmetros que devem ser especificados ou conhecidos para o cálculo. Os principais empecilhos recaem em questões como tempo de amostragem, quantidade de pontos experimentais, uso de dados compactados ou interpolados pela base de dados, determinação da ordem apropriada dos polinômios empregados nos métodos e conhecimento ou especificação do tempo morto.

Os testes realizados na RECAP corroboram para alguns dos resultados apresentados em Thornhill et al. (1999), como por exemplo os valores típicos sugeridos de tempo morto para diferentes tipos de malhas. No entanto, comprovou-se que não há a necessidade de utilizar-se modelo com 30 parâmetros. A Figura 3.30 e a Figura 3.31 ilustram que ordens

entre 5 e 10 representam muito bem os dados. Ordens maiores que esta faixa aumentam o esforço computacional e o erro numérico dos métodos desnecessariamente.

A Figura 3.30 e a Figura 3.31 estão organizadas da seguinte maneira: a primeira linha contém dois gráficos onde é apresentado o índice de desempenho variando com a ordem do polinômio de ajuste, mantendo-se o tempo morto constante e igual a 5. As duas curvas em cada gráfico representam um período antes e um depois da re-sintonia das malhas de controle. A segunda linha de gráficos mostra o comportamento do índice de desempenho em relação ao tempo morto.

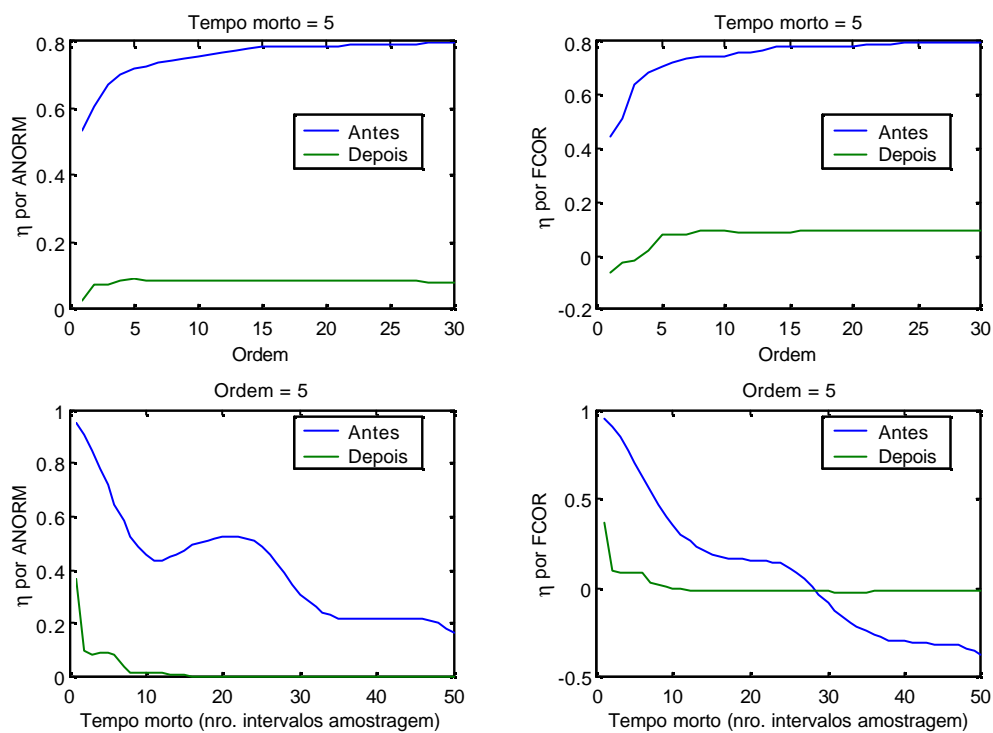


Figura 3.30. Variação do índice de desempenho com a ordem do polinômio de ajuste e com o tempo morto para uma malha de vazão.

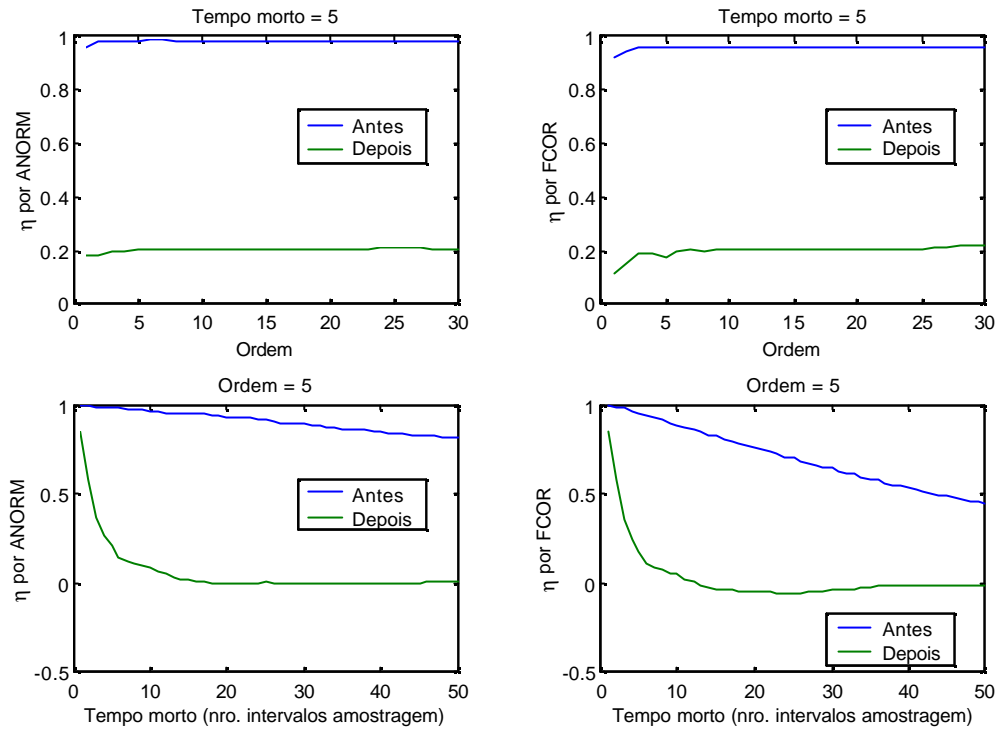


Figura 3.31. Variação do índice de desempenho com a ordem do polinômio de ajuste e com o tempo morto para uma malha de pressão.

Estas duas figuras também servem para expor a dificuldade de se obter uma estimativa para o horizonte de predição de acordo com a abordagem proposta por Thornhill et al. (1999) quando as malhas estão mal ajustadas. Ao calcular-se os índices de desempenho variando-se o horizonte de predição nota-se que não há a formação de um patamar definido nas malhas de baixo desempenho, fazendo com que esta análise torne-se casual e subjetiva.

A análise dos gráficos de índice de desempenho contra o tempo morto também permite observar que um controlador bem ajustado rejeita em bem menos tempo os distúrbios que um controlador mal ajustado. Isto é equivalente a dizer que em um intervalo de tempo menor os dados da saída do processo deixam de ter correlação com os dados do distúrbio de entrada devido a ação do controlador, levando o índice rapidamente a valores próximos de zero. Esta informação também pode ser utilizada como um critério de engenharia quando não se conhece o tempo morto de uma malha (ou discorda-se dos valores sugeridos na literatura). Pode-se escolhê-lo como sendo o tempo em termos de número de intervalos de amostragem em que se deseja que o controlador rejeite uma perturbação. Assim, a comparação da variância do controlador implementado com a deste controlador desejado vai resultar num valor mais “justo” do índice de desempenho para esta malha.

Para testar-se o procedimento de detecção de oscilações, selecionou-se algumas malhas sabidamente oscilatórias e outras com as mesmas características mas sabidamente não oscilatórias. Hagglund (2001) observa ser necessária a implementação deste procedimento diretamente em um SDCD, tendo em vista que certas informações necessárias para a detecção de oscilações não serem disponibilizadas em computadores de engenharia. No entanto, todas

as informações necessárias para o procedimento estavam disponíveis e foram obtidas através do PI, permitindo que se efetuassem os testes em qualquer computador de processo.

A Figura 3.32 ilustra o emprego do método em uma malha de vazão com parâmetros do controlador $t_i = 1\text{min}$; período de amostragem $h = 0,33\text{min}$; limite aceitável de oscilações $n_{Lim} = 10$; máxima amplitude de oscilação aceitável $a = 13,09$ unidades de vazão e sem comportamento oscilatório. A Figura 3.33 ilustra o emprego do procedimento em outra malha de vazão, com parâmetros do controlador $t_i = 4\text{min}$; período de amostragem $h = 0,33\text{min}$; limite aceitável de oscilações $n_{Lim} = 10$; máxima amplitude de oscilação aceitável $a = 40,26$ unidades de vazão e comportamento oscilatório detectável por simples inspeção visual.

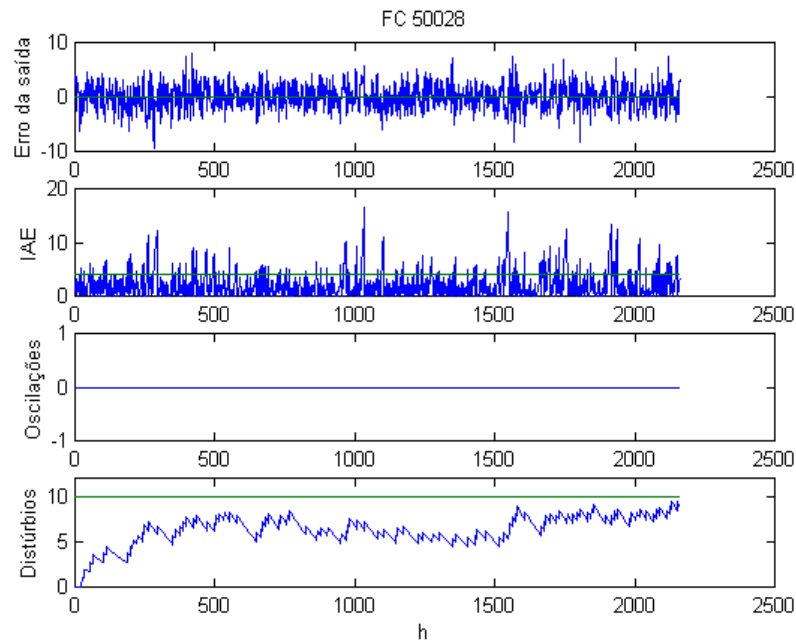
O primeiro gráfico da Figura 3.32 e da Figura 3.33 apresenta o erro da variável de processo em relação ao *setpoint*. O segundo gráfico apresenta a integral do erro (zerada a cada troca de sinal do erro da saída). O terceiro gráfico indica os pontos onde o procedimento caracterizou a malha como oscilatória. O último gráfico apresenta a contagem de oscilações durante a supervisão da malha, corrigido pelo fator de esquecimento que faz o papel de tempo de supervisão, já que este é um método recursivo.

Observa-se que no caso da malha da Figura 3.33, o procedimento detecta comportamento oscilatório por três vezes na janela de dados, enquanto que no caso da Figura 3.32, a contagem de distúrbios não atinge em nenhum momento o limite máximo para que não se caracterize comportamento oscilatório.

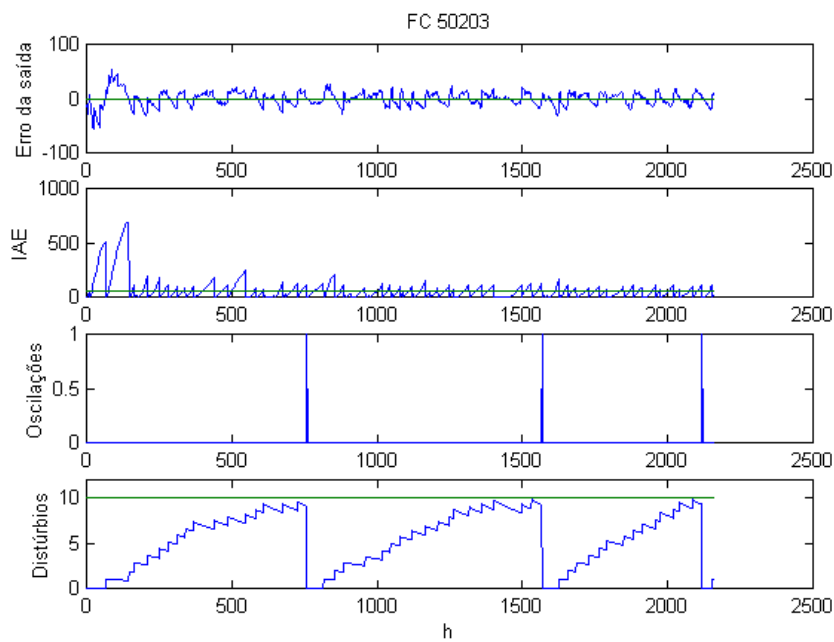
Os índices de desempenho também foram avaliados para as duas malhas. Os resultados encontram-se na Tabela 3.10.

Tabela 3.10. Resultados do procedimento de detecção de oscilações em conjunto com os índices de desempenho.

| Procedimento | Malha 1 | Malha 2 |
|--------------|---------|---------|
| ANORM. | 0,009 | 0,61 |
| FCOR | 0,000 | 0,61 |
| Oscilatória? | Não | Sim |

Figura 3.32. Gráficos gerados pelo procedimento de detecção de oscilação para uma malha

de vazão não oscilatória.

**Figura 3.33.** Gráficos gerados pelo procedimento de detecção de oscilação para uma malha de vazão oscilatória.

3.4. Conclusões

O resultado mais direto que se pode obter a partir dos testes com controladores de variância mínima é que seu emprego prático na indústria química/petroquímica é na grande maioria dos casos inviável. Suas ações de controle são excessivamente bruscas e possuem

uma amplitude tamanha que fariam as válvulas de controle atuarem sempre nos limites de saturação. Outro empecilho seria a baixa robustez, quando implementados a partir de sistemas sem resposta inversa e com componentes integrativos. No caso prático nunca se consegue obter um modelo do processo com a exatidão necessária para o funcionamento de um CVM, de tal forma que a mínima incerteza no modelo instabilizaria o sistema. A sensibilidade poderia ser reduzida com uma taxa de amostragem menor (maior intervalo de amostragem) em relação a sua constante de tempo dominante, mas mesmo estimando-se um modelo exato para o processo, o sistema pode estar sujeito a alterações nos seus parâmetros decorrente do tempo de operação, fazendo com que os modelos tivessem que ser sempre atualizados.

As metodologias empregadas são de relativa facilidade de implementação e uso, através de abordagens simples e atrativas. Por outro lado, a redução da variabilidade é crítica para processos industriais. No entanto, não é aconselhável investir muito esforço em atingir-se variância mínima às custas de deterioração de outras características importantes como redução na robustez do sistema à mudanças no processo, desempenho determinístico (servo) e diagnóstico baseado em conhecimento prévio. Supondo-se sistemas lineares, o ponto de mínimo da variância é único, de forma que para se atingir a variância mínima o controlador PI/PID deve aproximar-se do CVM. Com isso, o controlador começará a manifestar também os efeitos indesejáveis da implementação de um CVM.

O erro das estimativas da variância mínima produzido pelos métodos de avaliação de desempenho encontra-se dentro do limite aceitável, como demonstram os dados da Tabela 3.3 até a Tabela 3.7.

Os testes mostram que fatores limitantes ao desempenho do controlador, adicionais ao tempo morto, tais como zeros de fase não mínima e pólos instáveis podem levar a um resultado enganoso da interpretação dos índices de desempenho. Estes fatores devem ser levados em conta na avaliação da qualidade do controle, mas na prática requerem um conhecimento a mais do sistema, que é justamente a localização destes pólos e zeros. Esta é outra limitação encontrada nas ferramentas de avaliação de desempenho atuais.

Os testes demonstram que existem certas limitações no emprego das metodologias apresentadas. Um índice de desempenho de malhas de controle por si só não pode ser tomado como justificativa para intervenções em uma malha de controle. Em conjunto com outras técnicas e usado em um sistema de monitoramento automático do processo permite levantar com clareza pontos críticos no conjunto de malhas de controle. O procedimento de detecção de oscilação, por exemplo, mostra-se complementar no diagnóstico de desempenho de uma malha. Outras técnicas estatísticas, tais como a análise de correlação dos dados, são recomendados como auxiliares na avaliação de malhas de controle. O grande benefício destas metodologias (e também a idéia principal por trás delas) está no monitoramento automático das condições das malhas, indicando à engenharia os pontos problemáticos a serem explorados.

As dificuldades de implementação dos algoritmos que calculam os índices estão dentro do aceitável para um corpo de engenharia com recursos humanos limitados.

Como pode ser visto não há um índice absoluto que permita taxar com certeza o desempenho de uma malha como bom ou ruim com a simples verificação de seu valor. Mas quando usados comparativamente a períodos de ajustes diferentes, estes índices demonstram-se claramente como um indicativo de melhora ou deterioração do controle.

Algumas melhorias ainda podem ser efetuadas nos procedimentos de cálculo dos índices. Alternativas de índice de desempenho para malhas de nível ainda devem ser propostas e testadas. Como demonstrado neste trabalho, as ordens dos polinômios de ajuste dos métodos de avaliação de desempenho podem ficar entre 5 e 10, sem necessidade de ordens maiores, como sugerido na bibliografia, desde que o tempo de amostragem seja compatível com a dinâmica do sistema.

Quanto ao procedimento de detecção de oscilações, algumas modificações também podem ser efetuadas. O procedimento proposto por Hägglund (1995) requer a determinação de diversos parâmetros através de testes experimentais na malha ou de informações que muitas vezes não se encontram disponíveis. É interessante então que se verifique a possibilidade de determinar valores comuns para tipos de malhas específicos. Para malhas de temperatura, por exemplo, o uso da faixa de operação (*range*, informação geralmente configurada e disponível nos SDCD's) da variável não é adequado para o método. Sugere-se aqui usar ao invés de 1% da faixa de operação, um percentual de 1 a 5% da faixa que compreende ± 4 desvios padrão da variável no período. Um estudo mais abrangente ainda deve ser levado a cabo com a finalidade de determinar parâmetros deste tipo comuns a cada tipo de malha de diferentes unidades industriais.

Capítulo 4 Sistemas Multivariáveis

A extensão dos métodos de avaliação de desempenho ao caso multivariável se faz necessária devido a inevitável existência de interação entre variáveis na grande maioria dos processos químicos ou petroquímicos. Esta característica dos sistemas se reflete na avaliação da qualidade do controle, onde um diagnóstico completo e mais esclarecedor só é possível se a interação entre as malhas for levada em conta.

Este capítulo apresenta esta extensão ao caso multivariável (MIMO). Duas das principais técnicas apresentadas no Capítulo 2 são aqui adaptadas. O ponto chave para estas extensões está na matriz de interação, que é para o caso MIMO o equivalente ao tempo morto no caso SISO.

4.1. Introdução

Em todo o Capítulo 4 será adotada a convenção de que as potências negativas de q não serão explicitadas nas variáveis, enquanto que as potências positivas serão mostradas explicitamente.

Ao longo deste capítulo um sistema MIMO será representado por um modelo do tipo:

$$Y_t = TU_t + Ne_t \quad (4.1)$$

onde T e N são matrizes de transferência racionais e próprias, no operador q^{-1} , Y_t , U_t e e_t são os vetores de saída, entrada e ruídos, com dimensões apropriadas. e_t será assumido como sendo ruído branco com média zero e variância representada por S_e .

Para resolver problemas de controle com variância mínima envolvendo sistemas multivariáveis é necessário introduzir a matriz de interação, D , que é a generalização do tempo morto SISO para o caso MIMO.

4.2. A Matriz de Interação

4.2.1. Definições

Assim como o tempo morto é a limitação fundamental ao desempenho nos casos monovariáveis, a matriz de interação caracteriza a mais fundamental limitação ao desempenho alcançável de um controlador multivariável. Deste modo, a noção da matriz de interação é melhor entendida fazendo-se a analogia com o tempo morto do caso SISO. Para um sistema SISO, com função de transferência dada por:

$$T = \frac{B}{A} q^{-d}$$

o tempo morto q^{-d} insere d zeros infinitos no sistema. A matriz T não é inversível já que sua inversa não é própria. No entanto, multiplicando a função de transferência pela matriz de interação do caso SISO, $D = q^d$, então a função de transferência livre do tempo morto, $\tilde{T} = DT$, é inversível. A multiplicação da função de transferência pela matriz de interação remove os zeros infinitos da função de transferência original, gerando uma função de transferência livre do tempo morto.

A matriz de interação pode ser definida como o operador que extrai a estrutura de atraso no tempo de uma matriz de transferência, levando em conta todas as interações existentes. Desta forma, a equação (4.1) pode ser representada como:

$$Y_t = D^{-1}\tilde{T}U_t + Ne_t \quad (4.2)$$

A matriz de interação possui a seguinte propriedade:

Para uma função de transferência polinomial T , própria, racional, de dimensões $n \times m$, existe uma única matriz polinomial D não singular, triangular inferior, de dimensões $n \times n$ tais que:

$$|D(q)| = q^r \quad (4.3)$$

$$\lim_{q^{-1} \rightarrow 0} D(q)T = \lim_{q^{-1} \rightarrow 0} \tilde{T} = K \quad (4.4)$$

Onde K é uma matriz constante de posto cheio, r é um inteiro definido como o número de zeros infinitos de T e \tilde{T} possui apenas zeros finitos.

A matriz D , definida como a matriz de interação, pode ser escrita como:

$$D(q) = D_0q^d + D_1q^{d-1} + \dots + D_{d-1}q \quad (4.5)$$

onde d é denotado a ordem da matriz de interação e D_i são matrizes constantes representando os coeficientes da matriz polinomial D .

Uma matriz de interação pode existir de três formas distintas. Se D é da forma:

$$D(q) = q^d I \quad (4.6)$$

então a matriz de interação é chamada de “matriz de interação simples”.

Se D é uma matriz diagonal:

$$D(q) = \text{diag}(q^{d_1}, q^{d_2}, \dots, q^{d_n}) \quad (4.7)$$

então é referida como sendo uma “matriz de interação diagonal”

De outra forma, D é chamada de “matriz de interação genérica”.

Rogozinski et al. (1987) introduziu um algoritmo para o cálculo de uma “matriz de interação nilpotente”. Uma matriz D é nilpotente se:

$$\exists_p \text{ tal que } D^p = 0 \quad \text{ou seja,} \quad \det(A(q)) = q^p \text{ e } A(q) = qI_p - D$$

Nestes termos, a matriz D nilpotente é uma matriz quadrada em que todos os valores característicos são nulos.

Para a obtenção da matriz de interação utilizando este algoritmo é necessário o conhecimento da matriz de transferência T do sistema. O algoritmo encontra-se descrito no Apêndice A.

Baseado no trabalho de Rogozinski et al. (1987), Peng e Kinnaert (1992) introduzem a “matriz de interação unitária”. Este tipo de matriz de interação satisfaz a seguinte identidade:

$$D^T(q^{-1})D(q) = I \quad (4.8)$$

A matriz de interação definida na sua forma original era diagonal e isto implicava que a equação do controlador de mínima variância obtido resolvendo-se (4.1) tornava-se dependente da ordem em que as variáveis de saída são arrançadas. No entanto a fatoração de T com uma matriz de interação unitária não altera as propriedades espectrais da matriz de interação. Assim, a obtenção de uma matriz de interação na forma unitária elimina esta preocupação com o ordenamento das variáveis.

A matriz de interação unitária possui outra propriedade, definida em Peng e Kinnaert (1992): para uma matriz de transferência polinomial T , de posto cheio e própria, existe uma matriz de interação unitária não-única, pois quaisquer duas matrizes de interação D e \bar{D} satisfazem:

$$\bar{D}(q) = \Gamma D(q) \quad (4.9)$$

onde Γ é uma matriz real unitária $n \times n$, ou seja, $\Gamma^T \Gamma = I$.

Também é possível definir uma “matriz de interação espectral”, com as mesmas propriedades da matriz de interação unitária. Esta abordagem é utilizada por Harris et al. (1996a) para desenvolvimento de seu algoritmo de cálculo do índice de desempenho multivariável, a ser apresentado na Seção 4.3.1.

Se o esquema de controle com mínima variância visa minimizar uma função-objetivo ponderada, seja por questões de escalonamento devido a unidades, seja por diferentes importâncias dos canais, pode ser definida uma matriz de interação ponderada. Tal matriz é definida como a matriz de interação unitária que satisfaz a seguinte identidade:

$$D_w^T(q^{-1})D_w(q) = W \quad (4.10)$$

onde W é uma matriz de ponderação positiva e simétrica.

As mesmas propriedades de D se aplicam à D_w . Para uma matriz de transferência ponderada, $W^{1/2}T$, tendo D como matriz de interação unitária, pode-se escrever a matriz de interação ponderada como:

$$D_w(q) = D(q)W^{1/2} \quad (4.11)$$

4.2.2. Estimação da Matriz de Interação

Devido à interação entre os canais, em geral não é possível determinar a estrutura de atraso no tempo do sistema através apenas do conhecimento do tempo morto de cada canal individualmente. Existem técnicas para construir-se a matriz de interação através de métodos que envolvem operações elementares ou decomposição em valores singulares do numerador da matriz de transferência T (Rogozinski et al., 1987). Um destes métodos é desenvolvido no Capítulo 5

Esta metodologia para a determinação de D requer a determinação explícita da matriz de transferência T . Métodos de identificação de sistemas em malha fechada podem ser empregados para tal finalidade, ou então uma identificação em malha aberta também pode ser efetuada.

Em Huang et al. (1997a) e Huang e Shah (1999) é descrita uma metodologia que permite substituir o conhecimento da matriz de transferência pela determinação de alguns poucos parâmetros de Markov² do processo. A noção de atraso no tempo corresponde ao número de coeficientes de resposta ao impulso ou parâmetros de Markov cuja combinação linear é não-singular. Uma técnica para a determinação do número de parâmetros necessários para se obter a informação completa sobre a estrutura de atraso (ou seja, d), utilizando decomposição SVD é apresentada e encontra-se no Apêndice B.

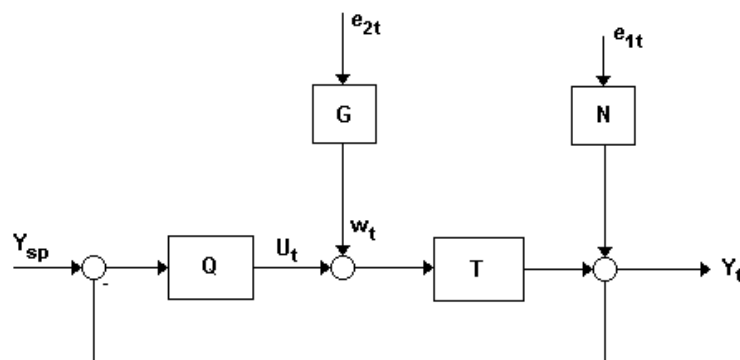


Figura 4.1. Diagrama em malha fechada do sistema utilizado para identificação de parâmetros de Markov.

² Parâmetros de Markov correspondem aos coeficientes da resposta ao impulso de um sistema.

Esta técnica permite que se utilize o mesmo algoritmo proposto por Rogozinski et al. (1987) para calcular a matriz de interação unitária, substituindo o numerador de T na forma polinomial pelos d parâmetros de Markov (resposta ao impulso) na forma de matriz polinomial. Assim, pode-se calcular os parâmetros de Markov com dados rotineiros de processo utilizando técnicas em malha fechada e alimentar o algoritmo de cálculo de D com estes parâmetros. Utilizando um sistema em malha fechada e inserindo-se uma perturbação conhecida conforme a Figura 4.1 é possível provar (Huang e Shah, 1999; Kamrunnahar et al., 2000) que a matriz de interação de um processo em malha fechada de w_t para Y_t é a mesma que para a matriz de transferência do processo em malha aberta (T).

Uma técnica de identificação que forneça boas estimativas dos parâmetros de Markov em altas frequências é desejada, já que é a faixa do espectro onde se encontra a informação sobre a estrutura de tempo morto. Huang e Shah (1999) sugerem análise de correlação cruzada ou alternativamente o ajuste de modelos paramétricos. Kamrunnahar et al. (2000) propõem uma técnica mista que envolve identificação de modelos auto-regressivos juntamente com parâmetros de resposta ao impulso (Markov) através de mínimos quadrados multivariável. Esta técnica provê vantagens sobre as técnicas convencionais tais como simplicidade computacional e menor erro na determinação dos parâmetros. O algoritmo, chamado ARMarkov-LS, encontra-se desenvolvido no Apêndice B.

4.3. Índices de desempenho multivariáveis

Nesta seção serão apresentadas basicamente duas abordagens de determinação de índices de desempenho para processos multivariáveis. Variantes destas abordagens podem ser implementadas, se desejado for, seguindo-se as considerações feitas no Capítulo 2. A primeira, desenvolvida por Harris et al. (1996a), é baseada em fatoração espectral da inversa da matriz de interação (D^{-1}) na sua forma triangular inferior para resolver o problema linear-quadrático. Um controlador linear-quadrático (LQ) ótimo para o processo descrito em (4.1) pode ser projetado como o controlador que minimiza a seguinte função objetivo:

$$J_{LQ} = \lim_{n \rightarrow \infty} \frac{1}{n} E \left\{ \sum_{t=1}^n [Y_t^T W_1 Y_t + \Delta U_t^T W_2 \Delta U_t] \right\} \quad (4.12)$$

onde W_1 é a matriz de pesos na saída e W_2 é a matriz de ponderação das entradas.

Já o método desenvolvido por Huang e Shah utiliza-se diretamente da matriz de interação na sua forma unitária para estender seu método FCOR ao caso multivariável.

4.3.1. Abordagem baseada em fatoração espectral

Harris et al. (1996a) estendem o método de avaliação de desempenho original (Desborough e Harris, 1992) para o caso multivariável utilizando-se de uma fatoração espectral da inversa da matriz de interação triangular inferior. Esta fatoração tem por objetivo

normalizar a matriz de interação, colocando-a em uma representação causal. A matriz de transferência T pode ser representada como uma fração (RMF, Right Matrix Fraction) da seguinte forma:

$$T = LR^{-1} \quad (4.13)$$

Em malha aberta, \tilde{T} pode ser escrito então como:

$$\tilde{T} = \mathbf{g}\tilde{L}R^{-1} \quad (4.14)$$

em que \mathbf{g} é uma matriz polinomial única de ordem finita, causal em termos do operador q^{-1} , obtida resolvendo-se a seguinte equação de fatoração espectral:

$$[\mathbf{g}(q)]^T W_1 \mathbf{g}(q^{-1}) = [D^{-1}(q)]^T W_1 D^{-1}(q^{-1}) \quad (4.15)$$

Esta fatoração espectral pode ser resolvida através de iterações com fatorações Cholesky, extração de fatores simétricos ou um procedimento de projeção causal utilizando transformada rápida de Fourier (Harris et al., 1996a).

A matriz de transferência do distúrbio para a saída pode ser também representada em termos de uma RMF:

$$N = \Theta\Phi^{-1} \quad (4.16)$$

de forma que o processo em (4.1) pode ser reescrito como:

$$Y_t = \mathbf{g}\tilde{L}R^{-1}U_t + \Theta\Phi^{-1}e_t \quad (4.17)$$

A equação para o controlador que minimiza a variância da saída para o caso multivariável (com $W_2 = 0$) é dada assim por:

$$U_t = -H_1\Theta\Phi^{-1}e_t \quad (4.18)$$

que é obtida resolvendo-se uma equação de Diophantine definida a seguir.

H_1 é um filtro dado por:

$$H_1 = \tilde{T}F_1 \quad (4.19)$$

e F_1 é resolvido de:

$$F_1 = \Xi\Theta^{-1} \quad (4.20)$$

onde Ξ é uma matriz polinomial de ordem finita surge da equação de Diophantine elaborada a partir da separação do termo $\mathbf{g}D(q)\Theta\Phi^{-1}$ em dois termos: um envolvendo apenas potências positivas de q e outro envolvendo um termo constante e potências negativas de q :

$$\mathbf{g}D(q)\Theta\Phi^{-1} = \mathbf{t}\Phi^{-1} + qP(q) \quad (4.21)$$

Rearranjando (4.21):

$$\mathbf{g}D(q)\Theta = \mathbf{t} + qP(q)\Phi \quad (4.22)$$

Sob controle com variância mínima, o comportamento do desvio da saída em relação ao *setpoint* é dado por:

$$Y_t^{MV} = D^{-1}\mathbf{g}^{-1}qP(q)e_t \quad (4.23)$$

Esta equação pode ser escrita como uma média móvel com ordem f , dada pela maior potência de q na matriz de interação menos uma unidade:

$$Y_t^{MV} = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \dots + \Psi_f e_{t-f} \quad (4.24)$$

A matriz de variância-covariância de Y_t^{MV} é dada por:

$$E\left\{(Y_t^{MV})(Y_t^{MV})^T\right\} = \Sigma_e + \Psi_1 \Sigma_e \Psi_1^T + \Psi_2 \Sigma_e \Psi_2^T + \Psi_f \Sigma_e \Psi_f^T \quad (4.25)$$

Também pode ser mostrado (Harris et al., 1996a) que o valor mínimo alcançável da função objetivo quadrática (4.12) com $W_2 = 0$ é dado por:

$$E\left\{(Y_t^{MV})W_1(Y_t^{MV})^T\right\} = \text{tr} W_1 \sum_{k=0}^f P_k \Sigma_e P_k^T \quad (4.26)$$

onde P_k é o k -ésimo coeficiente da matriz polinomial $P(q)$.

O desenvolvimento até a equação (4.26) considera sistemas sem zeros não-inversíveis. Para as alterações necessárias para sistemas de fase não-mínima, veja Harris et al. (1996a).

Para o cálculo do índice de desempenho é definida uma nova variável de saída deslocada no tempo:

$$\Omega_t \equiv \mathbf{g}D(q)Y_t \quad (4.27)$$

Sob controle com variância mínima, substituindo-se (4.23) em (4.27) obtém-se:

$$\Omega_t^{MV} \equiv qP(q)e_t \quad (4.28)$$

onde Ω_t^{mv} é função apenas de valores futuros. Desta forma, segue-se que, sob controle com variância mínima, as matrizes de covariância entre Ω_t e Y_{t-k} serão nulas para todos os deslocamentos positivos ($k \geq 0$):

$$E\{\Omega_t Y_{t-k}^T\} = 0, \quad k \geq 0 \quad (4.29)$$

Para um controlador qualquer dado por:

$$U_t = -KY_t \quad (4.30)$$

a expressão para Ω_t pode ser desenvolvida de (4.17) e (4.27) para chegar em:

$$\Omega_t = -\tilde{g}\tilde{T}K(I + D^{-1}\tilde{T}K)^{-1}\Theta\Phi^{-1}e_t + \tilde{g}D(q)\Theta\Phi^{-1}e_t \quad (4.31)$$

O primeiro termo no lado direito de (4.31) é uma função apenas de valores presentes e passados de e_t . O segundo termo pode ser fatorado em uma parcela envolvendo apenas valores presentes e passados de e_t e outra parcela envolvendo apenas valores futuros de e_t . O segundo termo do lado direito de (4.31) é exatamente o encontrado em (4.21). Desta forma as duas parcelas podem ser obtidas resolvendo-se a equação de Diophantine (4.22).

Resolvendo a equação de Diophantine, a equação (4.31) pode ser modificada para agrupar todos os termos dependentes do presente e do passado de e_t , Γ , e outro termo que não pode ser afetado por nenhum controlador, já que depende apenas de valores futuros de e_t .

$$\Omega_t = \Gamma e_t + qP(q)e_t \quad (4.32)$$

Com isso, a matriz de variância-covariância de Y_t e a função objetivo (4.26) sob controle com variância mínima dependem apenas dos elementos da matriz $P(q)$ e de Σ_e (assumindo que as matrizes de peso e de interação são conhecidas). Para avaliar a função objetivo é necessário então estimar $P(q)$ a partir de dados do processo em malha fechada. Isto pode ser feito de uma maneira similar ao caso monovariável descrito no Capítulo 2.

Dado um conjunto de dados de saída representativos do processo, pode-se ajustar um modelo de série temporal do tipo ARMA:

$$Y_t = BA^{-1}e_t \quad (4.33)$$

Assim, Ω_t fica:

$$\Omega_t = \tilde{g}D(q)BA^{-1}e_t \quad (4.34)$$

que pode ser fatorado em dois termos seguindo o mesmo procedimento desenvolvido para a equação (4.27), resolvendo-se uma equação de Diophantine.

$$\Omega_t = MA^{-1}e_t + qP(q)e_t \quad (4.35)$$

Finalmente, uma estimativa para o valor mínimo da função objetivo em (4.26) pode ser obtido, bem como uma estimativa para a matriz de variância-covariância em (4.25). As matrizes Ψ_k em (4.25) são os coeficientes do polinômio $D^{-1}\gamma^{-1}qP(q)$.

Utilizando todo este ferramental é possível definir um índice de desempenho global para o controlador multivariável como sendo a unidade subtraída da razão entre o valor ótimo da função objetivo e a matriz de variância-covariância da saída:

$$\hat{H} = 1 - \frac{E\left\{\left(Y_t^{MV}\right)^T W_1 \left(Y_t^{MV}\right)\right\}}{\frac{1}{n} \sum_{t=1}^n Y_t^T W_1 Y_t} = 1 - \frac{\text{tr} W_1 \sum_{k=0}^f P_k \Sigma_e P_k^T}{\frac{1}{n} \sum_{t=1}^n Y_t^T W_1 Y_t} \quad (4.36)$$

Índices de desempenho para cada malha individualmente podem ser obtidos estimando-se a matriz de variância-covariância da saída sob controle com variância mínima e fazendo-se a razão com a matriz de variância-covariância da saída real:

$$\begin{aligned} \left[h_{y_1}, \dots, h_{y_m} \right]^T &= 1 - \text{diag} \left\{ E \left\{ \left(Y_t^{MV} \right) \left(Y_t^{MV} \right)^T \right\} \text{diag} \left(\frac{1}{n} \sum_{t=1}^n Y_t^T W_1 Y_t \right)^{-1} \right\} \\ &= 1 - \text{diag} \left\{ \sum_{k=0}^f \Psi_k \Sigma_e \Psi_k^T \text{diag} \left(\frac{1}{n} \sum_{t=1}^n Y_t^T W_1 Y_t \right)^{-1} \right\} \end{aligned} \quad (4.37)$$

4.3.2. Abordagem baseada em filtragem e correlação

A extensão do método FCOR para o caso MIMO foi originalmente apresentada em Huang et al. (1997) e pode ser encontrada também em Huang e Shah (1999).

Considerando um sistema como descrito em (4.1), um controlador que minimiza a função objetivo LQ da saída original do sistema, Y_t :

$$J_{LQ,1} = E\left\{Y_t^T Y_t\right\} \quad (4.38)$$

também minimiza a função objetivo da saída filtrada pela matriz de interação, \tilde{Y}_t :

$$J_{LQ,2} = E\left\{\tilde{Y}_t^T \tilde{Y}_t\right\} \quad (4.39)$$

e $J_{LQ,1} = J_{LQ,2}$. A prova pode ser encontrada em Huang e Shah (1999) ou Huang et al. (1997). \tilde{Y}_t é dado por:

$$\tilde{Y}_t = q^{-d} D(q) Y_t \quad (4.40)$$

de forma que o sistema em (4.2) ao ser multiplicado em ambos os lados por $q^{-d} D(q) Y_t$ pode ser reescrito como:

$$\tilde{Y}_t = q^{-d} \tilde{T} U_t + \tilde{N} e_t \quad (4.41)$$

em que $\tilde{N} \equiv q^{-d} D(q) N$. Com este algebrismo conseguiu-se transformar um sistema com uma matriz de interação genérica em um sistema com uma matriz de interação do tipo simples.

A matriz de transferência do distúrbio filtrada pela matriz de interação pode ser separada em dois termos através de divisão longa, gerando a seguinte equação de Diophantine (note que na metodologia FCOR não é necessário resolver a equação de Diophantine para estimar a variância mínima).

$$\tilde{N} = \underbrace{F_0 + F_1 q^{-1} \cdots + F_{d-1} q^{-d+1}}_{F(q^{-1})} + q^{-d} R \quad (4.42)$$

R é uma matriz de transferência racional e própria.

Substituindo-se (4.42) em (4.41) obtém-se:

$$\tilde{Y}_t = \tilde{T} U_{t-d} + R e_{t-d} + F e_t \quad (4.43)$$

O termo $F e_t$ depende apenas de valores futuros de e_t e não pode assim ser afetado por algum controlador *feedback*. Assim, este termo corresponde ao termo de variância mínima do processo, pois é o único que resta em (4.43) na implementação de um controlador ótimo.

A variância da saída do processo sob controle com variância mínima é então dado por:

$$E\{\tilde{Y}_t^T \tilde{Y}_t\}_{\min} = E\{E_t^T E_t\} = \text{tr}(\text{var}(F e_t)) \equiv \Sigma_{\min} \quad (4.44)$$

onde $E_t \equiv F e_t$.

Seguindo o desenvolvimento feito para o caso monovariável do Capítulo 2, a covariância entre a saída e a seqüência de ruído branco para o atraso i (para $i < d$):

$$E\{\tilde{Y}_t e_{t-i}^T\} = F_i \Sigma_e \equiv \Sigma_{\tilde{Y}_e}(i) \quad (4.45)$$

A partir de:

$$\tilde{Y}_t^{MV} = E_t = F_0 e_t + \dots + F_{d-1} e_{t-d+1} \quad (4.46)$$

pode-se resolver para Y_t^{MV} :

$$Y_t^{MV} = q^d D(q)^{-1} (F_0 e_t + \dots + F_{d-1} e_{t-d+1}) \quad (4.47)$$

De (4.8) tem-se que:

$$\begin{aligned} D(q)^{-1} &= D^T(q^{-1}) = (D_0 q^d + \dots + D_{d-1} q)^{-1} \\ &= D_0^T q^{-d} + \dots + D_{d-1}^T q^{-1} \end{aligned} \quad (4.48)$$

de modo que:

$$\begin{aligned} Y_t^{MV} &= (D_0^T + \dots + D_{d-1}^T q^{d-1}) (F_0 + \dots + F_{d-1} q^{-d+1}) e_t \\ &\equiv (E_0 + E_1 q^{-1} \dots + E_{d-1} q^{-d+1}) e_t \end{aligned} \quad (4.49)$$

Para o caso de matriz de ponderação diferente da identidade, $D_w^{-1}(q) \neq D_w'(q^{-1})$, mas $D_w = DW^{1/2}$, de forma que:

$$D_w^{-1}(q) = W^{-1/2} D^T(q) \quad (4.50)$$

assim, a equação (4.48) deve ser modificada para levar em conta a matriz de pesos.

Representando E da equação (4.49) na forma matricial:

$$[E_0, E_1, \dots, E_{d-1}] = [D_0^T, D_1^T, \dots, D_{d-1}^T] \begin{bmatrix} F_0 & F_1 & F_2 & \dots & F_{d-1} \\ F_1 & F_2 & \dots & \dots & 0 \\ F_2 & \vdots & \ddots & \ddots & \vdots \\ \vdots & F_{d-1} & \ddots & \ddots & \vdots \\ F_{d-1} & 0 & \dots & \dots & 0 \end{bmatrix} \quad (4.51)$$

Da equação (4.49) a variância da saída sob controle com variância mínima pode ser escrita como:

$$\Sigma_{\min} = E_0 \Sigma_e E_0^T + E_1 \Sigma_e E_1^T + \dots + E_{d-1} \Sigma_e E_{d-1}^T \equiv XX^T \quad (4.52)$$

com X dado por:

$$X \equiv \left[E_0 \Sigma_e^{1/2} \quad E_1 \Sigma_e^{1/2} \quad \cdots \quad E_{d-1} \Sigma_e^{1/2} \right] \quad (4.53)$$

Da equação (4.45), tem-se diretamente que:

$$F_i = \Sigma_{\tilde{Y}_e}(i) \Sigma_e^{-1} \quad (4.54)$$

Substituindo (4.54) em (4.51) e utilizando o resultado em (4.53):

$$X = \left[D_0^T, D_1^T, \dots, D_{d-1}^T \right] \begin{bmatrix} \Sigma_{\tilde{Y}_e}(0) \Sigma_e^{-1/2} & \Sigma_{\tilde{Y}_e}(1) \Sigma_e^{-1/2} & \cdots & \Sigma_{\tilde{Y}_e}(d-1) \Sigma_e^{-1/2} \\ \Sigma_{\tilde{Y}_e}(1) \Sigma_e^{-1/2} & \Sigma_{\tilde{Y}_e}(2) \Sigma_e^{-1/2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \Sigma_{\tilde{Y}_e}(d-1) \Sigma_e^{-1/2} & \vdots & \vdots \\ \Sigma_{\tilde{Y}_e}(d-1) \Sigma_e^{-1/2} & 0 & \cdots & 0 \end{bmatrix} \quad (4.55)$$

Com a variância mínima (ponto ótimo da função objetivo) dada por (4.52), pode-se finalmente definir o índice de desempenho global do sistema multivariável:

$$\hat{H} = \frac{\text{tr} \Sigma_{\min}}{\text{tr} \{E\{Y_t T_t^T\}\}} = \frac{\text{tr}(XX^T)}{\text{tr} \Sigma_Y} \quad (4.56)$$

Os índices de desempenho das saídas individuais são obtidos dos elementos diagonais:

$$\left[\mathbf{h}_{y_1}, \dots, \mathbf{h}_{y_m} \right]^T = \text{diag} \{ XX^T \text{diag}(\Sigma_Y)^{-1} \} \quad (4.57)$$

Da mesma forma que no caso monovariável, os resíduos e_t são determinados através de um procedimento de filtragem utilizando o mesmo raciocínio descrito no Capítulo 2. Algoritmos do tipo ARX ou identificação no espaço de estado multivariável podem ser facilmente implementados ou obtidos para efetuar este procedimento de filtragem.

Para processos multivariáveis com zeros de fase não mínima, uma ligeira modificação no procedimento de cálculo é introduzida, de forma a que a matriz de interação fatorada incorpore não só os zeros infinitos referentes a estrutura de tempo morto, mas também os zeros de fase não mínima. Mais detalhes a respeito do procedimento pode ser encontrado em Huang e Shah (1999, 2000).

4.4. Aplicação em sistemas multivariáveis

Este capítulo trata da aplicação da metodologia de avaliação de desempenho a sistemas multivariáveis. As técnicas de avaliação de desempenho multivariável e univariável são conceitualmente similares. Por isso é de se esperar que a maioria das considerações feitas

no Capítulo 2 e no Capítulo 3 a respeito do caso SISO são válidas para o caso MIMO, como por exemplo, sistemas de fase não mínima, emprego de horizonte estendido, limitações das técnicas, suposição de sistemas lineares ou aproximadamente lineares e as considerações estatísticas como tamanho do conjunto de dados, tempo de amostragem e ordem de modelos. Modelos AR/ARMA/MA também podem ser substituídos por identificação no subespaço de estado ou filtro de Kalman, assim como no caso SISO.

A extensão da metodologia FCOR ao caso MIMO, baseada na utilização direta da matriz de interação do tipo unitária, é notadamente mais simples matematicamente e computacionalmente, e, portanto mais fácil de implementar, que a abordagem de Harris et al. (1996a) baseada em fatoração espectral de matrizes de interação diagonais. Uma implementação da abordagem normalizada (Harris et al., 1996a) em Matlab foi obtida diretamente com um dos autores (François Boudreau). No entanto, o algoritmo está implementado em uma versão antiga do Matlab e mesmo adaptado para versões posteriores, é de difícil utilização. Estas são, portanto, as razões da utilização do método FCOR para testes com sistemas multivariáveis.

4.4.1. Simulações

Para fins de teste da implementação do algoritmo FCOR multivariável, reproduziu-se alguns exemplos descritos em Huang e Shah (1999). O primeiro sistema é um processo MIMO 2 x 2 com matriz de transferência em malha aberta dada por:

$$T = \begin{bmatrix} \frac{q^{-1}}{1 - 0.4q^{-1}} & \frac{K_{12}q^{-2}}{1 - 0.1q^{-1}} \\ \frac{0.3q^{-1}}{1 - 0.1q^{-1}} & \frac{q^{-2}}{1 - 0.8q^{-1}} \end{bmatrix}$$

e matriz de transferência do distúrbio para a saída dado por:

$$N = \begin{bmatrix} \frac{1}{1 - 0.5q^{-1}} & \frac{-0.6}{1 - 0.5q^{-1}} \\ \frac{0.5}{1 - 0.5q^{-1}} & \frac{1}{1 - 0.5q^{-1}} \end{bmatrix}$$

Neste sistema, o ruído é uma seqüência normalmente distribuída com variância $S_e = I$. A qualidade da saída é medida por $J = E\{Y_t^T Y_t\}$. Pode-se representar o numerador de T como uma RMF ($K_{12} = 0$):

$$L = \begin{bmatrix} -0.065q^{-2} + 0.65q^{-1} & 0 \\ -0.078q^{-2} + 0.2q^{-1} & 0.62q^{-2} \end{bmatrix}$$

de onde, utilizando-se o algoritmo de Rogozinski et al. (1987) (Apêndice A), pode-se extrair uma matriz de interação unitária como:

$$D = \begin{bmatrix} -0.9578q & -0.2873q \\ -0.2873q^2 & 0.9578q^2 \end{bmatrix}$$

Multiplicando D por N e resolvendo-se a equação de Diophantine (4.42) obtém-se F como:

$$F = \begin{bmatrix} -1.1014q^{-1} & 0.2874q^{-1} \\ 0.1916 + 0.0958q^{-1} & 1.1302 + 0.5651q^{-1} \end{bmatrix}$$

A equação da saída sob variância mínima teórica pode então ser calculada de (4.47) como sendo:

$$Y_t^{MV} = \begin{bmatrix} 1 - 0.02752q^{-1} & -0.6 - 0.1623q^{-1} \\ 0.5 + 0.09176q^{-1} & 1 + 0.5412q^{-1} \end{bmatrix} e_t$$

Este será o *benchmark* teórico contra o qual vai-se comparar a variância mínima estimada pelo algoritmo FCOR a partir de dados simulados do sistema.

Para a simulação, um controlador de variância mínima sem compensação de interações é projetado:

$$Q = \begin{bmatrix} \frac{0.5 - 0.2q^{-1}}{1 - 0.5q^{-1}} & 0 \\ 0 & \frac{0.25 - 0.2q^{-1}}{(1 - 0.5q^{-1})(1 + 0.5q^{-1})} \end{bmatrix}$$

A simulação é rodada para valores de K_{12} variando de 0 até 10. O parâmetro K_{12} atua como um regulador do grau de interação do sistema. $K_{12} = 0$ resulta num sistema com fraca interação, de forma que o CVM projetado sem compensação de distúrbio deve ter alto desempenho. K_{12} tendendo a 10 resulta num sistema com alto grau de interação, produzindo um desempenho baixo. O índice de desempenho é calculado tanto para a variância mínima teórica quanto para a variância mínima estimada pelo algoritmo, e o gráfico encontra-se na Figura 4.2.

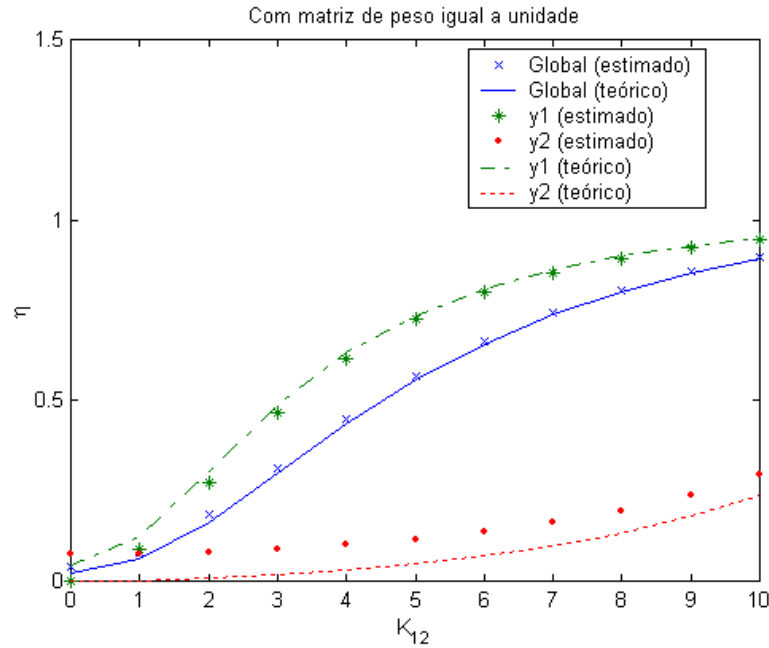


Figura 4.2. Avaliação de desempenho de um sistema MIMO com variação do grau de interação entre os canais.

Estes resultados são os mesmos encontrados por Huang e Shah (1999).

O segundo exemplo ilustra a utilização de uma matriz de peso nas saídas. Para tanto, utilizou-se o mesmo sistema do exemplo anterior, com uma matriz de ponderação dada por:

$$W = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Uma matriz de interação unitária ponderada é dada por (4.11), onde D é uma matriz unitária da matriz de transferência ponderada $W^{1/2}T$. Utilizando-se o algoritmo apresentado no Apêndice A, a matriz D pode ser calculada como:

$$D = \begin{bmatrix} -0.8575q & -0.5145q \\ 0.5145q^2 & -0.8575q^2 \end{bmatrix}$$

De forma que a matriz de interação unitária ponderada seja:

$$D_w = \begin{bmatrix} -0.8575q & -0.1029q \\ 0.5145q^2 & -1.7155q^2 \end{bmatrix}$$

Utilizando o mesmo procedimento do exemplo anterior, pode-se obter a saída do sistema de mínima variância ponderado como:

$$Y_t^{MV} = \begin{bmatrix} 1 - 0.08824q^{-1} & -0.6 - 0.5207q^{-1} \\ 0.5 + 0.07354q^{-1} & 1 + 0.4339q^{-1} \end{bmatrix}$$

A Figura 4.3 ilustra o resultado obtido. Novamente os índices estimados são bastante aproximados dos valores teóricos. Neste caso nota-se que o peso maior dado à saída Y_2 faz com que o desempenho global se desloque em direção a esta saída, diferentemente do exemplo anterior em que o desempenho global era mais influenciado por Y_1 . Nota-se que enquanto o desempenho global é limitado por $[0, 1]$, os índices de desempenho individuais podem ser menores que zero, efeito causado pela ponderação com a matriz de pesos.

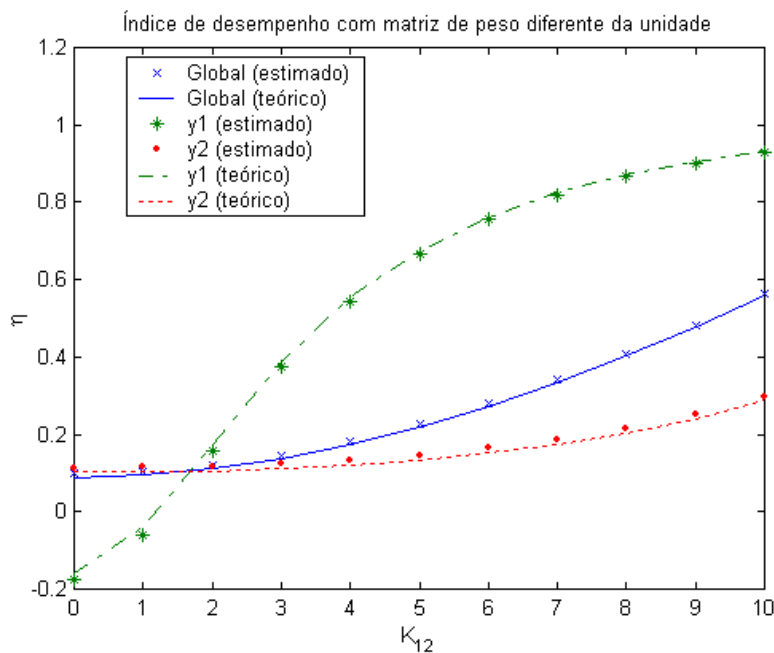


Figura 4.3. Avaliação de desempenho de um sistema MIMO com saídas ponderadas e variação do grau de interação entre os canais.

4.5. Conclusões

O embasamento matemático que estende os conceitos de tempo morto e desempenho baseado em variância mínima foi apresentado. O índice de desempenho estimado na forma multivariável é condizente com o esperado baseando-se na teoria. Uma técnica apropriada para estimação da estrutura de tempo morto multivariável também foi apresentada. Avaliação de desempenho multivariável e monovariável são conceitualmente similares. No entanto, no segundo caso, a análise dos resultados e o diagnóstico da malha são significativamente mais complicados, devido principalmente a presença de interações no controlador e no processo.

A abordagem baseada em filtragem e correlação é computacionalmente mais apropriada, uma vez que se utiliza apenas de cálculo matricial. A abordagem baseada em fatoração espectral requer cálculos com matrizes polinomiais, o que não é tão trivial de se

implementar. A mesma observação pode ser feita em relação às matrizes de interação. A forma unitária é a mais apropriada computacionalmente, tendo em vista que utiliza apenas matrizes constantes, reais, correspondentes aos parâmetros de Markov ou aos coeficientes do numerador na representação RMF da matriz de transferência.

Uma propriedade interessante do índice de desempenho no caso multivariável é que, além de calcular um desempenho global e os desempenhos individuais levando em conta o acoplamento do sistema, ainda permite que se analise a influência do desempenho de determinadas saídas no índice global, conforme demonstrado em simulações. Também é interessante notar que os algoritmos apresentados tornam possível trabalhar com matrizes de peso na saída e seus efeitos são capturados pelo índice de desempenho.

Ainda não há registro na literatura da integração de técnicas de avaliação de desempenho multivariáveis em larga escala em unidades industriais utilizando uma estratégia consideravelmente automatizada. A complexidade de sistemas multivariáveis torna difícil tanto a correta implementação de uma estratégia como a interpretação dos resultados obtidos.

Capítulo 5 Testes em unidade experimental

Neste capítulo realizou-se um estudo detalhado da aplicação das técnicas multivariáveis estudadas anteriormente em uma unidade experimental. Para que os resultados da aplicação dos métodos de avaliação de desempenho pudessem ser interpretados corretamente, foi feito um estudo detalhado das características da unidade. Os testes na unidade consistiram basicamente da identificação de matrizes de interação e avaliação do desempenho com índices baseados em variância mínima e análise de correlação.

5.1. Introdução

Neste capítulo serão apresentados resultados de testes em uma unidade experimental em escala piloto. Uma descrição detalhada da unidade será feita e testes sistemáticos serão realizados. O principal objetivo deste estudo é verificar se as conclusões obtidas através da interpretação dos resultados dos índices de desempenho, bem como da identificação da matriz de interação, estão de acordo com os resultados esperados pelo conhecimento prévio e detalhado das características da unidade. Outro objetivo deste capítulo é comparar o desempenho das malhas de controle do tipo PI utilizadas em relação a um controlador não-linear preditivo multivariável (NMPC) em controle regulatório e analisar os resultados.

Uma unidade experimental existente no Departamento de Engenharia Química da Escola de Engenharia da Universidade Federal do Rio Grande do Sul (UFRGS) foi utilizada para avaliação do algoritmo multivariável em condições mais próximas da realidade industrial.

5.2. Descrição da unidade

O objetivo de controle da unidade é controlar o nível de dois tanques de geometrias diferentes. A alimentação é proveniente de uma única linha, comum aos dois tanques, que é ramificada para abastecer cada tanque. Os dois tanques encontram-se em desnível e também estão acoplados por duas linhas que os interligam (V4 e V5). A unidade é controlada através de duas válvulas de controle, uma pneumática (VP), que regula a vazão de alimentação dos dois tanques (F_{SET}) e outra elétrica (VE) que regula a vazão de saída do tanque cilíndrico.

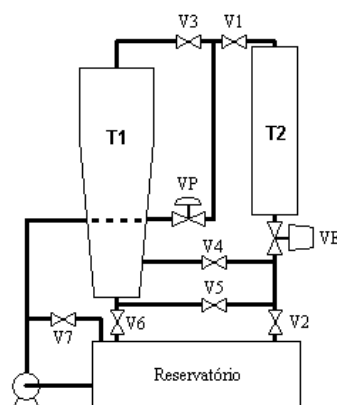


Figura 5.1. Unidade de 2 tanques de níveis acoplados.

A unidade pode ser configurada de diversas maneiras através da abertura ou fechamento das válvulas abre-fecha V1, V2, V3, V4, V5, V6 e V7, de forma a possibilitar a

elaboração de diversos sistemas com as mais variadas características dinâmicas e complexidades de controle. Também o formato do tanque T1, com uma seção cônica e outra cilíndrica, influencia consideravelmente no comportamento dinâmico da unidade. Os níveis serão referenciados como Y1 e Y2 para T1 e T2 respectivamente. O fluido empregado é água à temperatura ambiente. As válvulas de controle são denominadas VP (válvula pneumática) e VE (válvula elétrica)

5.2.1. Configuração e instrumentação

Dentre as diversas configurações possíveis para unidade, a configuração escolhida para efetuar os testes empregava as válvulas abre-fecha operando abertas ou fechadas de acordo com o descrito na Tabela 5.1. Esta configuração permite uma grande interação entre os níveis dos dois tanques.

Tabela 5.1. Configuração das válvulas da unidade.

| Válvula | Descrição | Aberta/fechada |
|---------|---|----------------|
| V1 | Válvula que habilita a alimentação ao tanque T2. | ABERTA |
| V2 | Válvula que habilita a descarga de T2 (OU T1 E T2) por caminho alternativo. | FECHADA |
| V3 | Válvula que habilita a alimentação ao tanque T1. | ABERTA |
| V4 | Válvula que habilita a interação intermediária entre os tanques. | ABERTA |
| V5 | Válvula que habilita a interação inferior entre os tanques. | FECHADA |
| V6 | Válvula de descarga de T1 (ou T1 e T2). | ABERTA |
| V7 | Válvula que habilita o reciclo da bomba. | ABERTA |
| VP | Válvula pneumática de controle da linha de alimentação dos dois tanques. | – |
| VE | Válvula elétrica de controle da descarga de T2 | – |

O reciclo máximo, proporcionado pela válvula V7, é necessário devido a um superdimensionamento da bomba. O reciclo exerce um papel importante na estrutura de controle. Como a bomba encontra-se superdimensionada, a inexistência do reciclo faria com que a válvula pneumática ficasse sem resolução de *range* para controlar a unidade, já que a pressão muito alta proporcionada pela bomba faz com que uma leve alteração na posição da válvula acarrete uma grande variação na vazão de alimentação. A faixa de vazão necessária para controlar a unidade fica entre 4 e 9 l/min.

O medidor de vazão encontra-se na linha de alimentação a montante da válvula pneumática. Utiliza como princípio de medição a rotação de uma turbina pela qual passa o líquido. A quantidade de ciclos por unidade de tempo é detectada utilizando uma propriedade eletromagnética.

O tanque de geometria mista (T1) emprega um medidor de nível do tipo *strain gauge*, que utiliza a diferença de pressão entre dois orifícios à diferentes colunas de líquido. O tanque cilíndrico (T2) utiliza um medidor capacitivo. Como a bomba centrífuga utilizada nesta unidade está trabalhando numa região de baixa eficiência, a maior parte da energia que a bomba transfere ao fluido não é aproveitada na forma de energia cinética, e sim na forma de calor, o que acarreta um problema para o medidor de nível de T2. Este medidor capacitivo demonstrou-se bastante sensível à mudança de temperatura e ao longo de um experimento a temperatura do líquido aumenta consideravelmente. Isto poderia ser resolvido através de calibração do medidor com a variação de temperatura e uma posterior correção do valor medido, mas a ausência de um medidor de temperatura *on-line* fez com que o único remédio aplicável fosse a troca, de modo manual, da água do reservatório.

O tanque cilíndrico (T2) é construído em acrílico, com diâmetro de 14,5cm e altura útil de 74 cm. O tanque T1, também de acrílico, possui altura útil de 96cm, com uma seção cônica de 70cm de altura e seção cilíndrica de 26cm. O diâmetro inferior é de 15cm e o diâmetro superior de 28cm. A linha de interação intermediária entre os dois tanques se dá a 30cm de altura em T1 e aproximadamente 15cm abaixo do fundo de T2.

Os dados de entrada/saída da instrumentação passam por uma placa de conversão de sinais e por uma placa de aquisição de dados instalada em um PC. Do PC, através de uma interface no Matlab/Simulink, é possível executar algoritmos de controle e operação da unidade. Uma biblioteca de comunicações possibilita a execução e visualização remota dos testes através de rede. A taxa de amostragem é de um segundo para todos os instrumentos.

5.2.2. Zonas de operação e comportamento dinâmico

A Unidade de 2 Tanques de Níveis Acoplados apresenta forte integração mássica, causada pela divisão da vazão de alimentação para cada tanque e pela ligação entre os dois vasos na posição intermediária. Devido principalmente as características geométricas de T1 e à linha de interação que conecta os dois tanques, a unidade apresenta então, 3 regiões de operação com comportamentos distintos quanto as suas dinâmicas, conforme demonstrado na Figura 5.2.

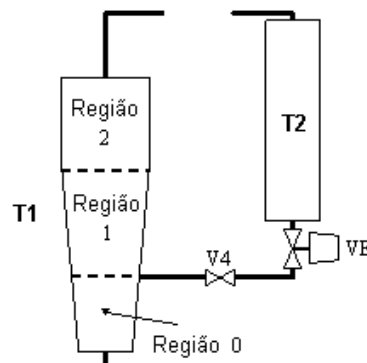


Figura 5.2. As três regiões de operação da unidade de dois tanques.

As três zonas de operação apresentam dinâmicas diferentes e a existência de interação entre os tanques torna o controle da unidade mais acoplado. A Região 0 é a região de dinâmica mais rápida, tendo em vista que a variação de sua altura de coluna d'água com a vazão é a mais intensa dentre as três regiões, já que possui a menor área da seção transversal. A Região 2 é a zona de dinâmica mais lenta dentre as três e é, portanto, a zona escolhida para ajuste dos controladores de Y1 empregados na unidade. Controladores projetados nesta região são robustos o suficiente para atuar nas demais regiões.

As Regiões 0 e 1 possuem um componente de não-linearidade a mais que a Região 2, em função de sua área de seção transversal variável com a altura de coluna d'água. No entanto, aproximações lineares em estado estacionário são razoáveis para emprego das metodologias de avaliação de desempenho. A identificação de parâmetros de Markov do sistema para posterior fatoração da matriz de interação considera modelos lineares e pode, portanto, ser afetada pela não-linearidade do sistema, uma vez que o sinal de perturbação utilizado na identificação faz com que o ponto de operação seja variável, obedecendo a dinâmica não-linear presente.

A Região 0, por possuir dinâmica mais rápida e pelas características do *range* da válvula VP, apresenta um comportamento ligeiramente oscilatório em malha fechada. A Região 2, de dinâmica mais lenta, acaba por filtrar naturalmente esta oscilação. A Região 1 apresenta um comportamento intermediário às demais regiões. A presença de oscilações na alimentação também pode ser sentida na medida de Y2, sendo este nível ainda mais sensível a distúrbios que Y1. Na presença de interação com T2, os comportamentos podem ser bastante diferentes dos descritos. Através de uma modelagem simples, baseada no balanço material, pode-se obter que a constante de tempo característica de T2 é uma função da altura de T2:

$$t_i = 2 \frac{Ah}{C_v \sqrt{h_0}} = 2 \frac{V}{F_s} = 2t_{res} \quad (5.1)$$

onde A é a área da seção transversal, C_v é a constante da válvula, h_0 é a altura do nível do tanque no ponto da linearização, V é o volume do tanque, F_s é a vazão de saída e t_{res} é o tempo de residência do tanque.

Uma função de transferência para este sistema pôde ser desenvolvida como:

$$\frac{\Delta h(s)}{\Delta F_e(s)} = \frac{2 \sqrt{h_0} / C_v}{t_i s + 1} \quad (5.2)$$

Em que ΔF_e é a variação da vazão de entrada de T2. Esta função de transferência funciona como um filtro das perturbações na vazão de alimentação para a descarga de T2.

A informação do valor da constante de tempo é importante para o ajuste adequado do controlador da malha de Y2. O controlador deve ter uma constante integral, t_I , da ordem de grandeza do maior valor de t_i no tanque, para garantir robustez em toda a faixa de nível de T2.

Valores de vazão entre 4 e 6,5 l/min fazem com que o sistema possa operar na Região 0. Valores acima de 6,5 l/min forçam a unidade a operar a partir da Região 1 ou 2.

5.3. Estrutura de controle

Através de dados experimentais prévios, e até mesmo intuitivamente, constata-se a maior influência da posição da VP no nível do T1 e da posição da VE no nível do tanque T2. A proposta de controle da unidade, por consequência, baseia-se no emparelhamento VP–Y1 e VE–Y2. Este emparelhamento foi utilizado tanto nos testes com malhas PI como nos testes com o NMPC.

5.3.1. Malhas de controle

A implementação de controladores SISO descentralizados simples em testes prévios não deu bons resultados na rejeição de distúrbios, especialmente na malha VP–Y1. Desta forma, optou-se pela implementação de uma malha em cascata para controlar o nível de T1. Em Machado et al. (2001) é demonstrado que, para esta unidade, a aplicação da estratégia em cascata reduz significativamente o tempo para rejeição completa de um distúrbio, com redução do erro absoluto integrado na mesma proporção. Mas a principal vantagem que a malha em cascata apresenta é a redução da não linearidade da válvula de controle.

Um esquema da estrutura de controle empregada para os testes é apresentado na Figura 5.3. A cascata tem como malha interna um controlador de vazão que utiliza como variável manipulada a posição de VP e recebe como *setpoint* o valor de vazão, F_{SET} , calculado pela malha externa. A malha externa tem como variável controlada o nível de T1 e como variável manipulada o valor de vazão passado como *setpoint* para a malha interna.

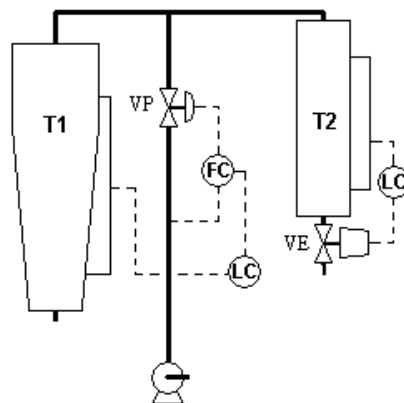


Figura 5.3. Esquema da unidade de dois tanques explicitando a estrutura de controle utilizada.

Os controladores da malha em cascata são do tipo PI simples, enquanto que o controlador da malha de Y2 é um PI baseado no algoritmo da velocidade (Åström e

Hägglund, 1998) com *anti-windup*, o que melhora significativamente a qualidade do controle desta malha. Esta melhoria resulta da ação *anti-windup* de minimizar o problema de saturação da válvula V2, que sem esta ação, passa a trabalhar constantemente nas restrições.

A malha de Y2 é bem mais rápida que a malha de Y1, permitindo um ajuste mais agressivo do controlador, o que leva a esperar um desempenho melhor para esta malha no cálculo dos índices de desempenho.

O controlador preditivo (NMPC) testado na unidade para comparação de desempenho com as malhas PI caracteriza-se por ser um controlador não linear e utilizar um algoritmo de linearizações ao longo da trajetória (LLT) (Duraiki, 2001). O modelo utilizado pelo controlador é um modelo fenomenológico não linear, resultante de balanços de massa ajustado a dados experimentais do sistema. A modelagem utilizada para este sistema é descrita pelo seguinte conjunto de equações:

$$\frac{dh_1}{dt} = K_{11} \frac{F_1 + F_3 - F_4}{A_1}$$

$$\frac{dh_2}{dt} = K_{11} \frac{F_2 - F_3}{A_2}$$

$$\frac{dx_1}{dt} = 0$$

$$\frac{dx_2}{dt} = 0$$

Onde:

$$F_1 = X \cdot F_T + \frac{x_1}{100}$$

$$F_2 = (1 - X) \cdot F_T + \frac{x_2}{100}$$

$$F_3 = V_E \cdot 0,0064817 \cdot \text{sign}(Fm) \cdot (\text{abs}(Fm))^{0,55378}$$

$$Fm = (h_2 + 15) - (h_1 - 30) \cdot \frac{\text{sign}(h_1 - 30) + 1}{2}$$

$$F_4 = -0,00034123 \cdot h_1^2 + 0,084646 \cdot h_1 + 4,1861$$

$$A_1 = \left(K_1 + \frac{h_1}{K_2} \right)^2; \quad \text{para } h_1 < 70$$

$$A_1 = 616; \quad \text{para } h_1 > 70$$

$$A_2 = 165;$$

$$K_1 = 7,5 \cdot \sqrt{p}$$

$$K_2 = \frac{10,77}{\sqrt{p}};$$

$$K_{11} = 16,67;$$

Os valores de x_1 e x_2 dão grandezas estimadas por um filtro de Kalman e dão uma estimativa do desvio existente na vazão de alimentação dos tanques T1 e T2. Ambos são limitados por $[0, 1]$ e são complementares. h_1 e h_2 e representam os níveis dos tanques cônico e cilíndrico respectivamente bem como A_1 e A_2 as suas áreas transversais e F_1 e F_2 as suas vazões de alimentação. F_3 é a vazão intermediária aos dois tanques (descarga de T2 em T1) e F_4 a vazão de descarga do tanque cônico. F_T e VE são as variáveis manipuladas vazão total de alimentação e abertura de válvula elétrica respectivamente.

A malha de controle, na presença do NMPC, encontra-se esquematizada na Figura 5.4. A malha interna do sistema em cascata de Y1, empregada nos testes com controladores PI, continua existindo, mas agora recebe o *setpoint* do controlador preditivo. Este toma então, o lugar da malha externa de Y1 e da malha de Y2.

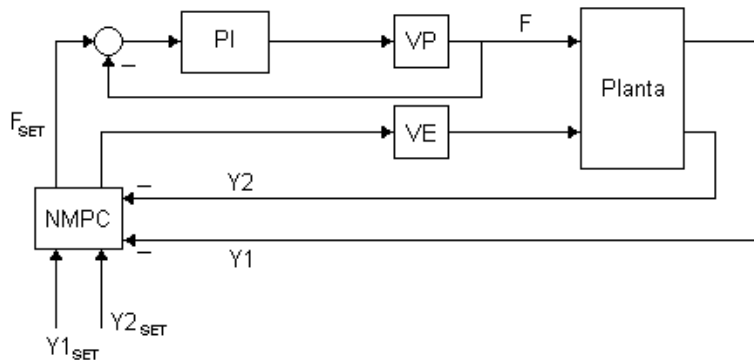


Figura 5.4. Diagrama esquemático da malha de controle com o uso do NMPC.

5.3.2. Ajuste empregado

Os ajustes dos controladores PI discretos foram feitos com um procedimento simples baseado em heurísticas do método IMC, cujos parâmetros foram obtidos a partir de testes feitos em malha aberta para cada malha separadamente. Portanto, a interação entre os tanques não foi considerada, resultando em controle multivariável do tipo PID descentralizado.

Os parâmetros de ajuste dos controladores encontram-se na Tabela 5.2. Os parâmetros dos controladores PI seguem o seguinte formato:

$$u(k) = K \left(1 + \frac{h}{t_I(q-1)} \right) e(k) \quad (5.3)$$

onde h é o intervalo de amostragem e K é o ganho do controlador.

Tabela 5.2. Parâmetros de ajuste dos controladores PI empregados nos testes.

| Parâmetro | Malha Y2 | Malha Y1 - Interna | Malha Y1 - Externa |
|-----------|----------|--------------------|--------------------|
| K | -10 cm | 0,4 l/min | 0,3 cm |
| t_I (s) | 50 | 2,5 | 200 |

Os *setpoints* de cada malha encontram-se listados na Tabela 5.3.

Tabela 5.3. Valores para os *setpoints* utilizados nos testes para as 3 regiões de operação.

| Setpoint | Região 0 | Região 1 | Região 2 |
|------------|----------|----------|----------|
| $Y1_{SET}$ | 20 cm | 50 cm | 80 cm |
| $Y2_{SET}$ | 20 cm | 40 cm | 65 cm |

Os parâmetros de ajuste do NMPC foram obtidos empregando-se a metodologia RPN (Farina et al., 2001).

5.4. Procedimento de testes

O procedimento de testes consistiu na coleta de 2 horas de dados com a unidade operando em condições normais, em torno do *setpoint*, com atuação regulatória dos controladores. Como o tempo de amostragem foi de um segundo, coletou-se 7200 pontos sob estas condições de operação dentro de cada região de operação característica da unidade. Para cada região, ao finalizar-se a coleta em estado estacionário, iniciou-se o procedimento de inserção de perturbações medidas na carga, com o objetivo de coletar dados adequados para a identificação dos parâmetros de Markov e posterior fatoração da matriz de interação do sistema em cada região de operação.

As perturbações inseridas consistiam de sinais aleatórios e independentes, somados ao *setpoint*, com média zero e variância de 10 unidades. O tempo de amostragem das perturbações foi de 60s para Y2 e 40s para Y1. O tamanho dos dados coletados com o sistema nestas condições foi de 3600 pontos, equivalente à uma hora de aquisição.

As seções de dados do processo em operação normal foram divididas em 7 janelas de aproximadamente 1000 pontos para fins do cálculo do índice de desempenho na forma de janela móvel (índices de desempenho fragmentados, Capítulo 3). O índice de desempenho multivariável também foi estimado para o conjunto completo de dados, incluindo os dados perturbados utilizados para a identificação da matriz de interação. Com isto pode-se demonstrar a deterioração da qualidade do controle devido à inserção de perturbações no sistema. Neste caso, 10 janelas móveis de aproximadamente 1000 pontos cada foram utilizadas.

Como era sabido de antemão que a unidade apresentava tempos mortos baixos, também procedeu-se uma nova rodada de testes nas três regiões, desta vez com a adição artificial de tempo morto nas saídas (medições de nível). Este tempo morto adicional consiste de uma matriz diagonal da seguinte forma:

$$D_{ADIC} = \begin{bmatrix} q^{-15} & 0 \\ 0 & q^{-10} \end{bmatrix} \quad (5.4)$$

de modo que 15s de tempo morto foram adicionados na malha do tanque cônico (T1) e 10s de tempo morto foram adicionados ao tanque cilíndrico (T2). Assim, denominar-se-à Caso 1 os testes realizados com o sistema original e Caso 2 os testes realizados com o sistema alterado pela adição do tempo morto D_{ADIC} .

O procedimento de testes com o NMPC foi um pouco diferente do procedimento para os controladores PI. Nos testes com o controlador preditivo, não foram coletados dados com perturbações medidas para a estimação da matriz de interação, sendo o cálculo dos índices de desempenho feito com as matrizes de interação obtidas nos testes com PI's para cada região de operação da unidade. O tempo de amostragem nos testes com o NMPC foi de 2s, sendo que as ações de controle eram tomadas de 5 em 5 segundos.

5.5. Desempenho com controladores Multi-SISO

Os procedimentos de avaliação de desempenho foram aplicados juntamente com procedimentos estatísticos de análise de correlação tanto para os dados do processo em operação normal quanto para os dados do processo perturbado. Os resultados para os dois casos (sem adição de tempo morto e com adição de tempo morto) encontram-se nas seções a seguir.

5.5.1. Caso 1

Para o caso em que a estrutura de tempo morto natural do processo não foi alterada, os dados gerados para as três regiões de operação normal encontram-se ilustrados a seguir.

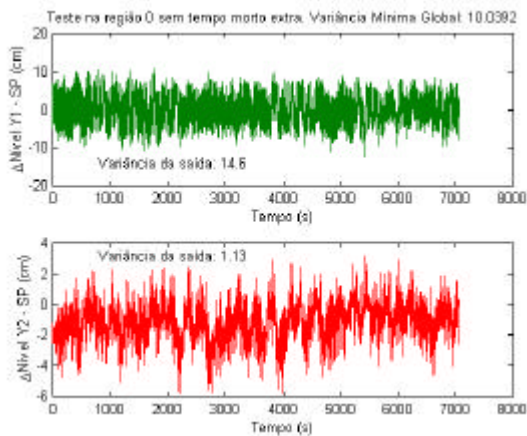


Figura 5.5. Região 0 em operação normal.

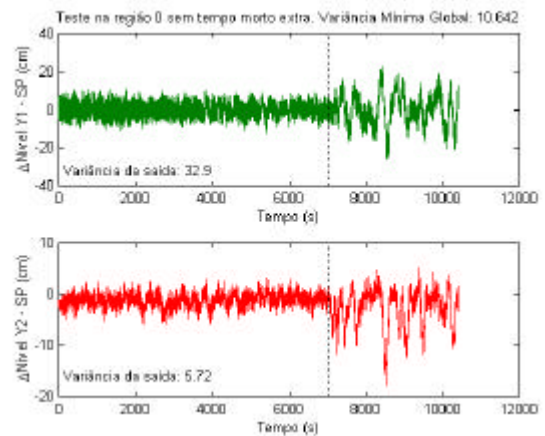


Figura 5.6. Região 0 com perturbações.

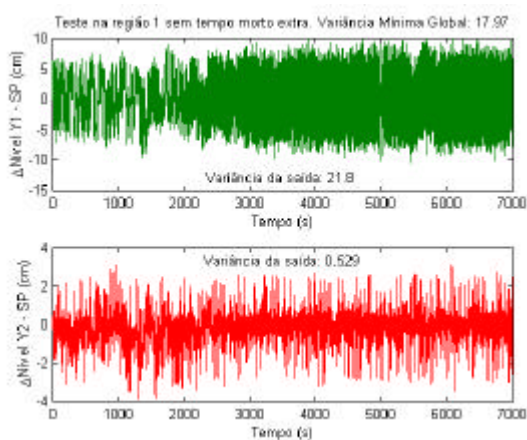


Figura 5.7. Região 1 em operação normal.

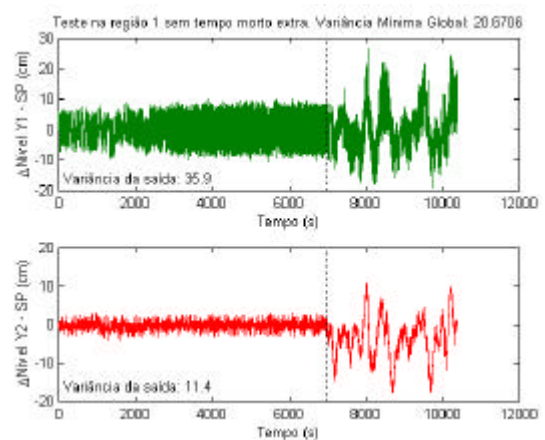


Figura 5.8. Região 1 com perturbações.

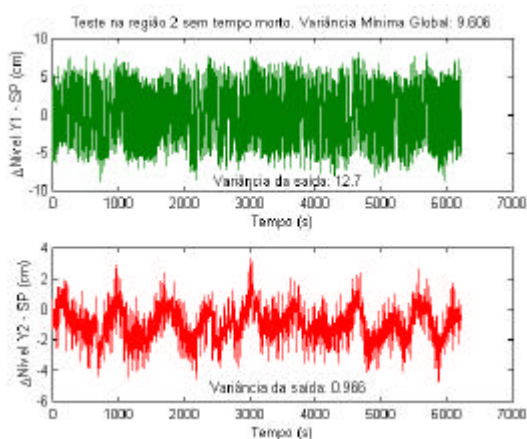


Figura 5.9. Região 2 em operação normal.

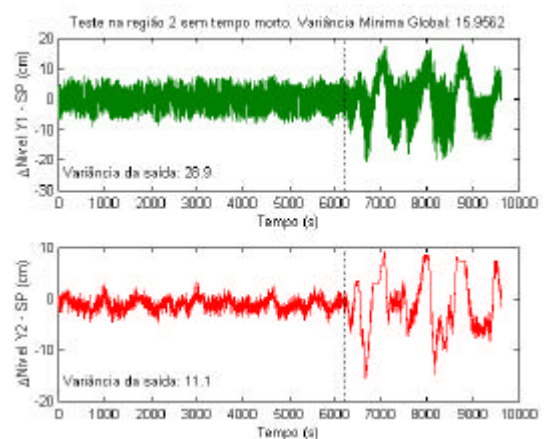


Figura 5.10. Região 2 com perturbações.

A linha pontilhada nos gráficos dos dados das regiões com as perturbações (coluna a direita) é a linha que separa os dados de operação normal dos dados utilizados na identificação das matrizes de interação (direita da linha). As matrizes de interação para as regiões 0, 1 e 2 foram identificadas a partir dos dados perturbados e são dadas respectivamente pelas equações a seguir:

$$D_0 = \begin{bmatrix} -q & 0.022q \\ 0.022q^3 & q^3 \end{bmatrix} \quad (5.5)$$

$$D_1 = \begin{bmatrix} -0.082q & q \\ -q^2 & 0.082q^2 \end{bmatrix} \quad (5.6)$$

$$D_2 = \begin{bmatrix} -0.062q & q \\ -q^3 & 0.082q^3 \end{bmatrix} \quad (5.7)$$

Com os dados do processo em condições normais de operação e as respectivas matrizes de interação, foram calculados os índices de desempenho para as malhas nas três regiões. Para cada região foi estimado o índice de desempenho fragmentado em janelas de aproximadamente 1000 pontos, bem como o índice de desempenho para o conjunto todo de dados.

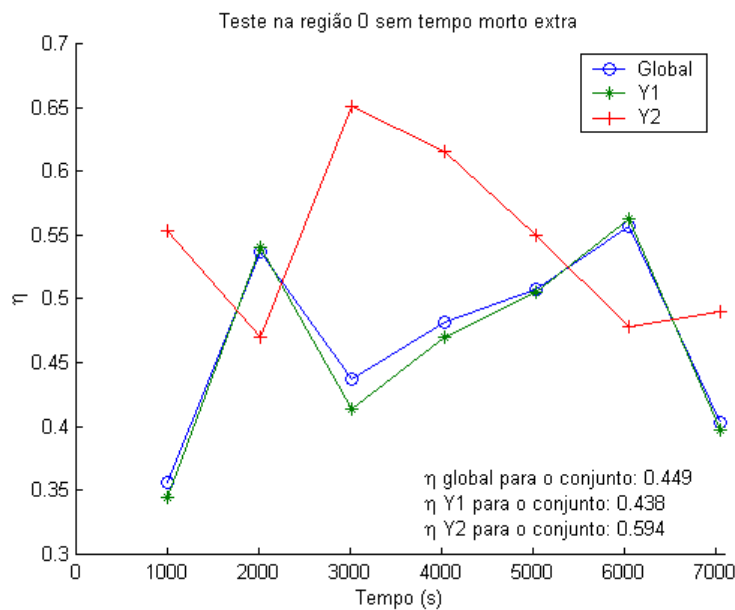


Figura 5.11. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 0.

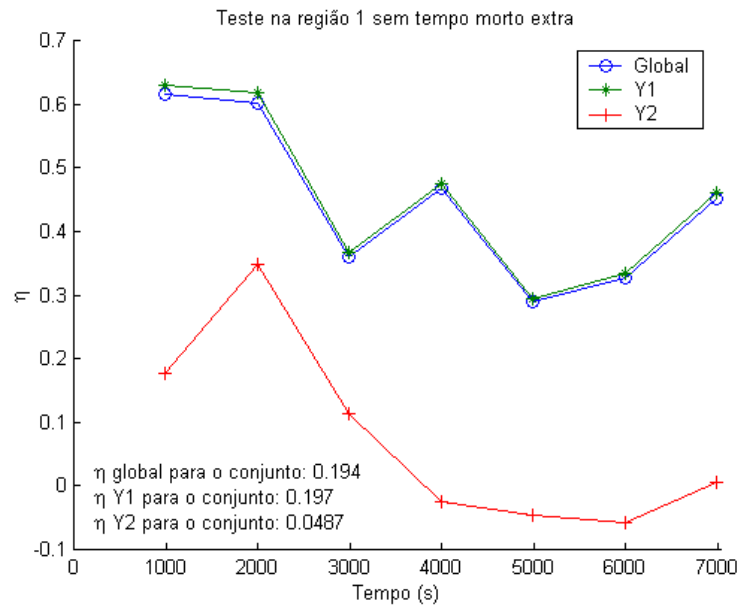


Figura 5.12. Índices de desempenhos fragmentado e para todo o conjunto de dados, estimados para a Região 1.

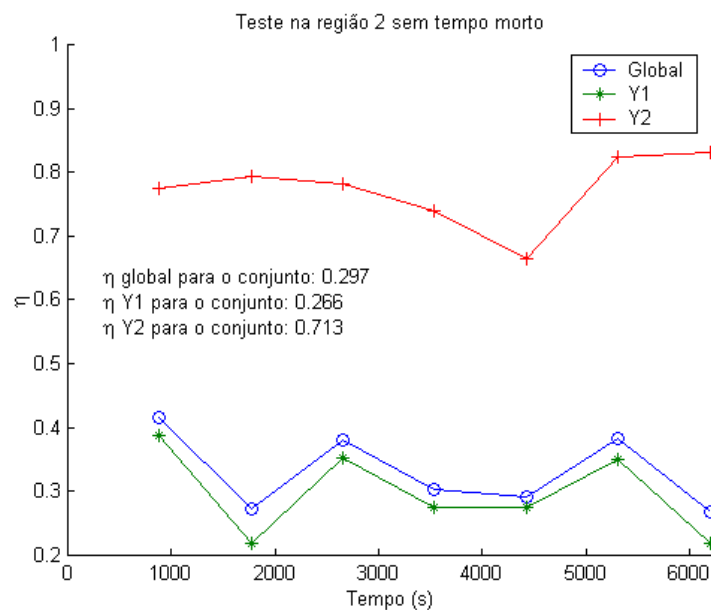


Figura 5.13. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 2.

Os dados referentes às perturbações inseridas no *setpoint* para estimação da matriz de interação, bem como a análise de correlação dos dados encontram-se no Apêndice D.

Os resultados dos índices de desempenho estimados mostram que o desempenho global do controle da unidade é especialmente influenciado pelo desempenho da malha de Y1, indicando que o investimento em melhoras nesta malha trará resultados mais significativos para o desempenho global que na malha de Y2. Devido ao acoplamento, pode não ser possível

melhorar o desempenho global com um simples reajuste de Y1 apenas. Desta maneira, se necessário for, pode-se relaxar os parâmetros da malha Y2 e ajustar a malha de Y1 mais agressivamente. Estas considerações também dependeriam da relativa importância das variáveis para o objetivo de controle. Se o objetivo principal da unidade é controlar Y1, pode-se sacrificar Y2. Se o objetivo é controlar Y2, uma matriz de pesos poderia ser utilizada para dar mais importância à Y2 na avaliação de desempenho e um ajuste mais cuidadoso desta malha deveria ser explorado, em detrimento do ajuste de Y1.

A identificação das matrizes de interação resultou em uma estrutura de tempo morto de ordem 3 para as regiões 0 e 2 e de ordem 2 para a Região 1. O comportamento dos desempenhos também é da mesma forma distinto. Enquanto que na Região 1 o desempenho da malha de Y2 é superior ao desempenho da malha de Y1, nas regiões 0 e 2 o inverso ocorre. Ao analisar-se as saídas de Y2 nas três regiões, nota-se que para a Região 0 e especialmente para a Região 2, o comportamento do nível é oscilatório. Este fato repercute nos índices de desempenho estimados para esta malha. Isto pode ser explicado através da análise da dinâmica do sistema. O sistema segue a equação (5.1), na qual em malha fechada, F_s é também dependente da abertura de VE. Para um *setpoint* de nível e vazão de alimentação baixos (Região 0), VE permanece aberta e o tempo de residência é baixo. Consequentemente, t_i é pequeno. Para um *setpoint* de nível e vazão altos (Região 2), a válvula elétrica permanece aberta e o tempo de residência também é baixo. Assim ocorre uma competição de fatores para definir o valor da constante de tempo característica do sistema. As relações das variáveis F_e , *setpoint* do nível e abertura de VE na Região 1 são tais que conferem um tempo de residência maior ao tanque. Seguindo a política de ajustar os parâmetros do controlador para a Região mais lenta do sistema, o controlador resulta mais apropriado para a Região 1, enquanto que para as regiões 0 e 2 o controlador se torna muito lento, a ponto de não ser rápido o suficiente para compensar os distúrbios.

A Região em que os controladores PI apresentaram o pior desempenho global foi a Região 0. A explicação deste fato deve-se a dois fatores: a forte dependência do desempenho global em relação ao da malha de Y1 e o ajuste do PI desta malha, que por ser realizado tomando como base a região de maior constante de tempo (Região 2), acaba por produzir fraco desempenho na Região 0.

Na Figura 5.5 pode-se notar um aumento na amplitude de oscilações de Y2 entre os tempos 2000s e 5000s. Esta piora na variabilidade da saída é devidamente capturada pela avaliação do desempenho na Figura 5.11. Na Região 1 uma mudança no padrão do distúrbio de Y2 e Y1 a partir de 2500s (Figura 5.7) também é refletida na melhora do índice de desempenho das duas malhas, conforme demonstrado na Figura 5.12. Esta mudança no padrão do distúrbio para as duas malhas é explicado pela mudança de comportamento na atuação de VP a partir de 2500s. A Figura 5.14 ilustra claramente esta mudança.

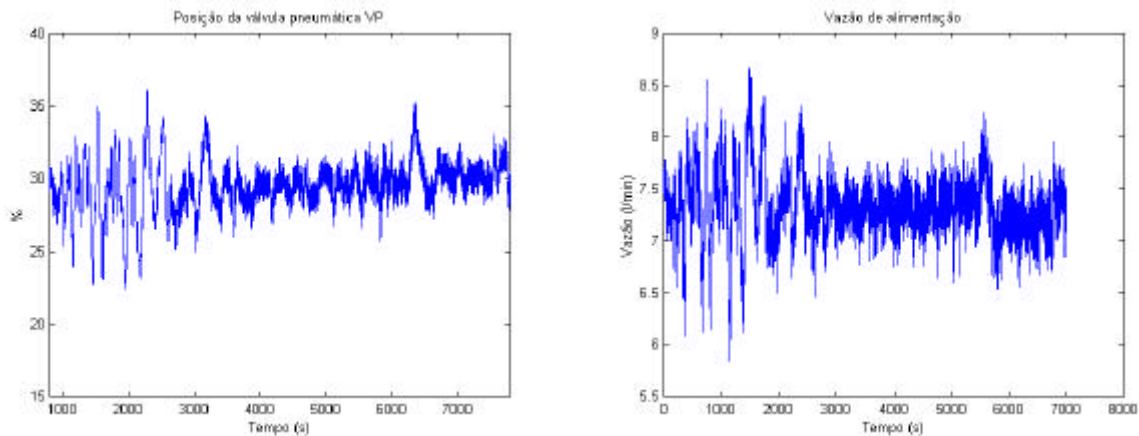


Figura 5.14. Mudança no comportamento da atuação de VP (esquerda) e conseqüente alteração no padrão da vazão de alimentação (direita).

Já o motivo da alteração na atuação da válvula pode ser justificado pela mudança na pressão do ar comprimido de alimentação de VP ocasionado por uma ativação ou desativação do respectivo compressor. Isto demonstra em que nível o desempenho da malha interna da cascata influi no desempenho da malha externa. Um procedimento de avaliação como o proposto por Ko e Edgar (2000) que avalia a saída da malha cascata levando em conta a qualidade da malha interna poderia ser aplicado nesta unidade para mensurar este efeito. A simples aplicação dos métodos aqui apresentados na malha interna também apontaria a deterioração no desempenho da malha cascata devido a malha interna, mas de um modo indireto.

5.5.2. Caso 2

O procedimento de testes foi repetido para o caso em que a estrutura de tempo morto do sistema foi alterado pela adição de um tempo morto extra na saída, dado por (5.4). Os resultados encontram-se ilustrados nas figuras a seguir, no mesmo esquema da Seção 5.5.1.

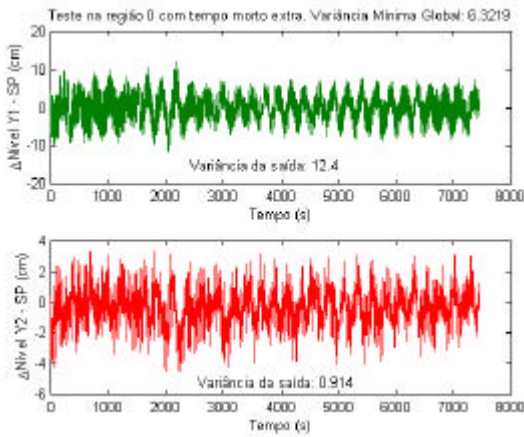


Figura 5.15. Região 0 em operação normal.

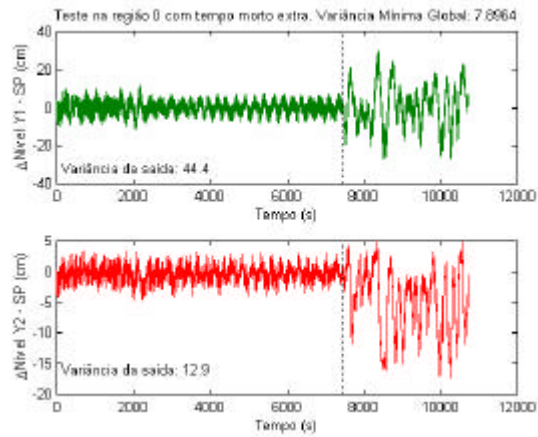


Figura 5.16. Região 0 com perturbações.

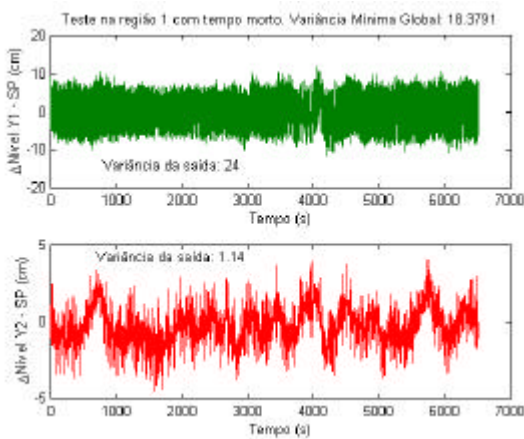


Figura 5.17. Região 1 em operação normal.

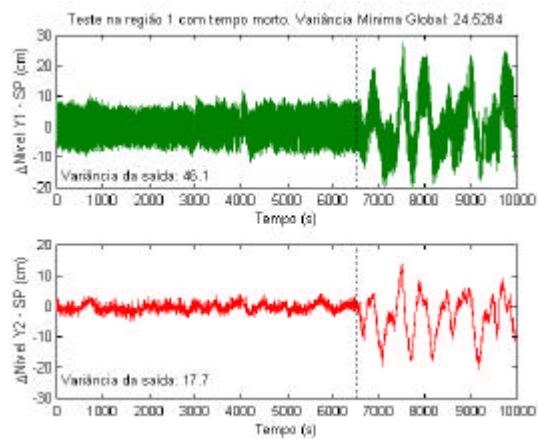


Figura 5.18. Região 1 com perturbações.

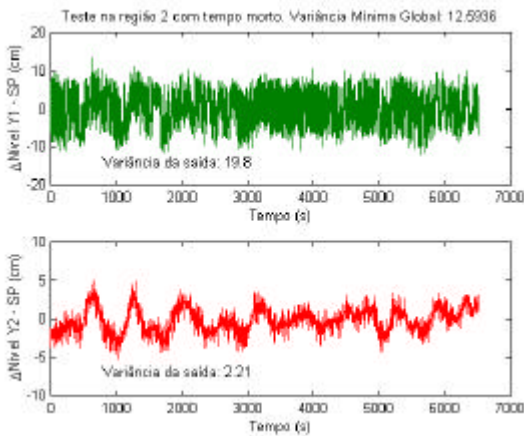


Figura 5.19. Região 2 em operação normal.

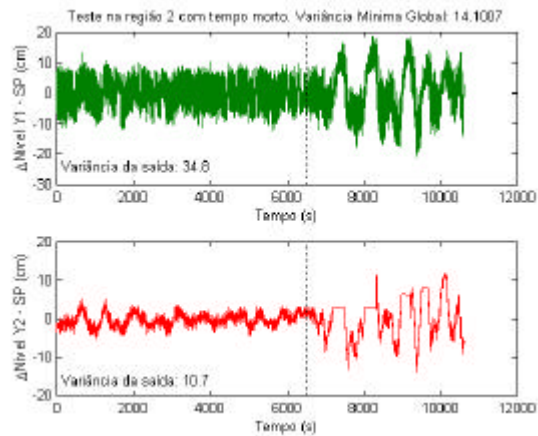


Figura 5.20. Região 2 com perturbações.

As matrizes de interação para as regiões 0, 1 e 2 foram identificadas a partir dos dados perturbados e são dadas respectivamente pelas equações a seguir:

$$D_0 = \begin{bmatrix} -0.88q^6 & 0.47q^6 \\ 0.47q^6 & 0.88q^6 \end{bmatrix} \quad (5.8)$$

$$D_1 = \begin{bmatrix} -0.15q & 0.99q \\ -0.99q^3 & -0.15q^3 \end{bmatrix} \quad (5.9)$$

$$D_2 = \begin{bmatrix} -q^6 & 0.074q^6 \\ 0.074q^6 & q^6 \end{bmatrix} \quad (5.10)$$

Com estas matrizes de interação e os dados de operação em condições normais, foi aplicado o algoritmo FCOR multivariável. Os resultados estão ilustrados nas figuras a seguir.

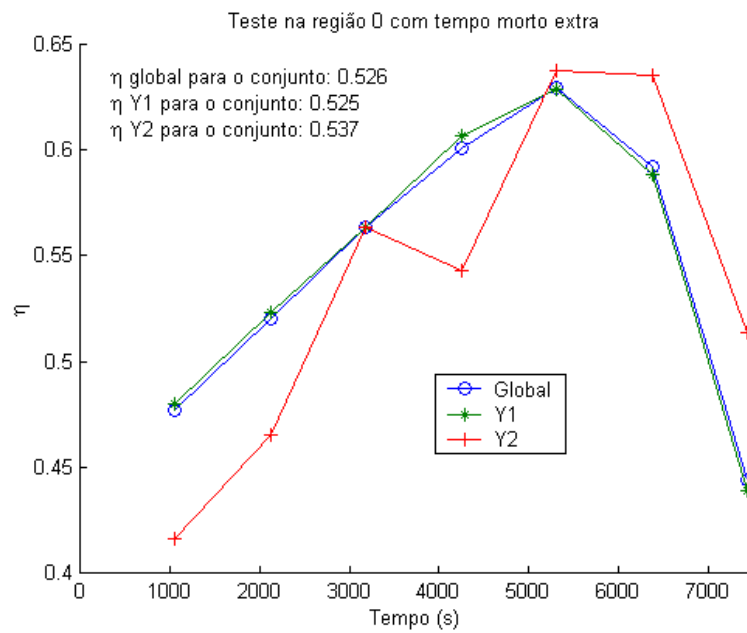


Figura 5.21. Índices de desempenho fragmentados e para todo o conjunto de dados, Região 0.

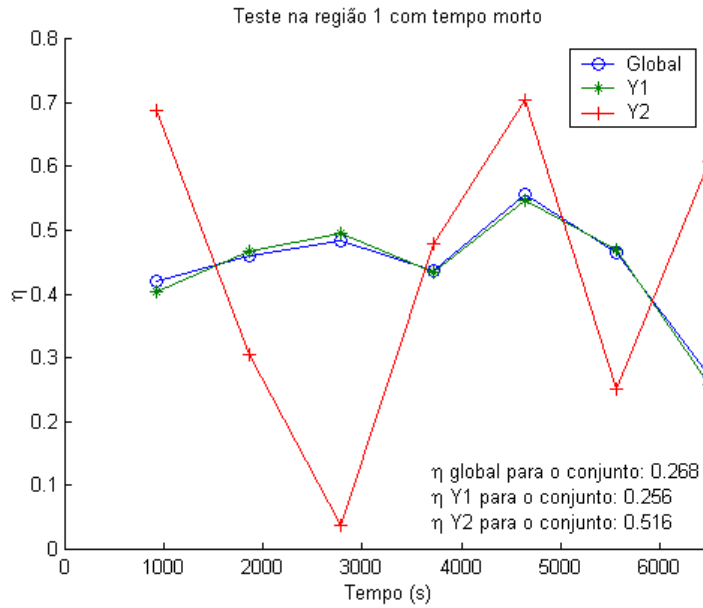


Figura 5.22. Índices de desempenho fragmentados e para todo o conjunto de dados, Região 1.

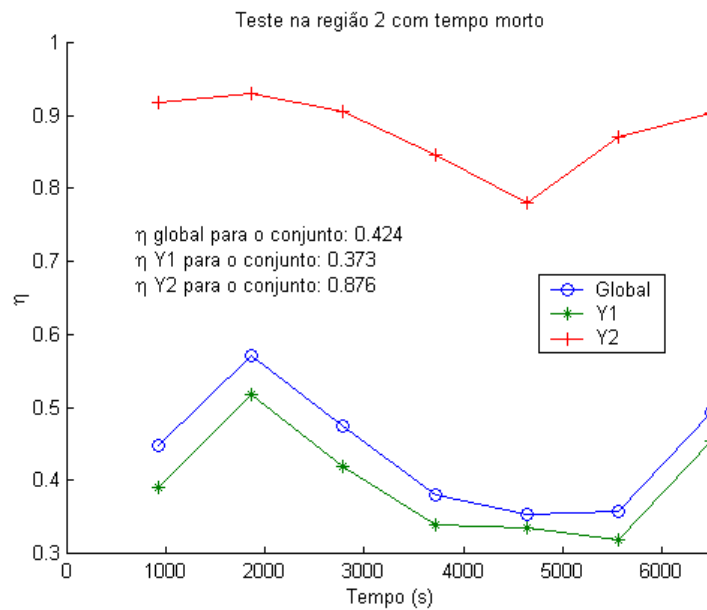


Figura 5.23. Índices de desempenho fragmentado e para todo o conjunto de dados, estimados para a Região 2.

Como era de se esperar, a adição de tempo morto extra às saídas resulta em matrizes de interação de maior ordem. Novamente a Região 1 apresenta resultados distintos quanto à estimação das matrizes de interação. Já o desempenho global da unidade foi relativamente parecido nas três regiões. Novamente a característica marcante é a forte influência do desempenho de Y1 no desempenho global do sistema. Nos dois casos apresentados pode-se verificar que esta influência é mais acentuada nas regiões 0 e 1. A adição do tempo morto extra causou uma leve deterioração do desempenho das malhas se comparado com o Caso 1. Esta deterioração era esperada, pois o tempo morto é o principal fator limitante do desempenho de uma malha de controle.

A Figura 5.20 demonstra que, durante o procedimento de perturbações inseridas no *setpoint* para estimação da matriz de interação, T2 sofreu inundação em alguns períodos. Embora este fato possa prejudicar a qualidade da obtenção da matriz de interação, ao menos a sua ordem é condizente com a esperada por analogia com a Região 0 e com o Caso 1.

5.6. Desempenho com controlador NMPC

5.6.1. Caso 1

Para o Caso 1 foi realizado um experimento apenas na Região 0. Esta foi a situação em que os controladores PI do Caso 1 apresentaram o pior desempenho global. É de se esperar que o controlador preditivo não-linear não sofra os efeitos de ajuste de parâmetros para as diferentes regiões de operação, uma vez que o modelo empregado no NMPC leva em conta as não-linearidades e dinâmicas distintas de cada região. No Caso 2 esta particularidade será explorada, com o teste sendo realizado nas três regiões de operação.

O tempo de duração dos experimentos com o NMPC foi o mesmo do tempo de duração dos experimentos com PI's. Como o tempo de amostragem do NMPC era de 2s, a metade do número de pontos em relação aos testes com PI's foi coletada. Isto implica que, para resultar em uma estimativa do índice de desempenho com o mesmo número de pontos das estimativas para os PI's, o conjunto de dados foi fragmentado em janelas de aproximadamente 2000s (1000 pontos). Assim, o número estimativas também é reduzido pela metade. Os dados coletados no experimento estão ilustrados na Figura 5.24.

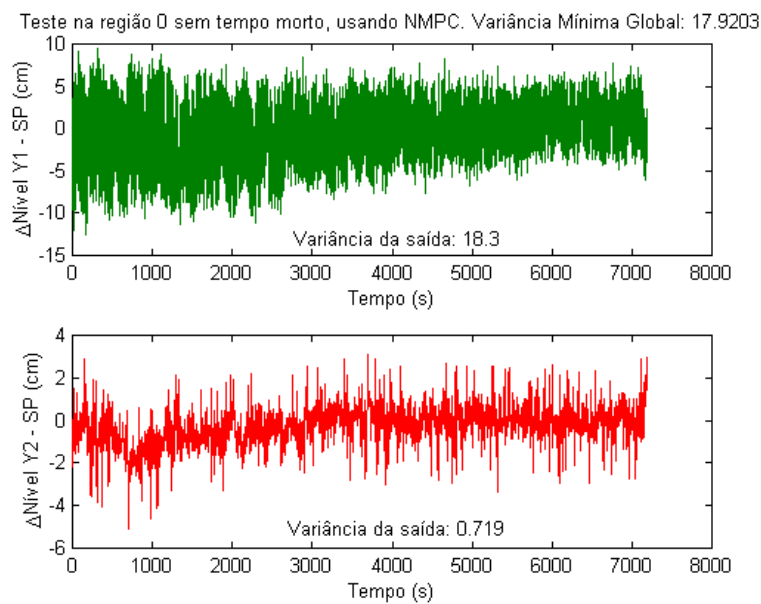


Figura 5.24. Dados coletados para o experimento com o NMPC na Região 0.

O resultado dos índices de desempenho do único teste realizado para este Caso encontra-se na Figura 5.25. Ao analisar-se os dados coletados nota-se que, por alguma razão ocorreu um desvio do *setpoint* na região compreendida entre 500s e 1500s na malha de Y2. A malha de Y1 também apresenta uma alteração na faixa entre 500s e 2500s. Dados de vazão de alimentação e da abertura de VE não apresentam anomalias que possam justificar este comportamento. Estes desvios refletem-se nos índices de desempenho calculados nesta região, especialmente na malha de Y2.

Apesar desta alteração, comparativamente ao resultado obtido com controlador PI na Região 0, Caso 1, os índices de desempenho produzidos pelo NMPC são consideravelmente melhores.

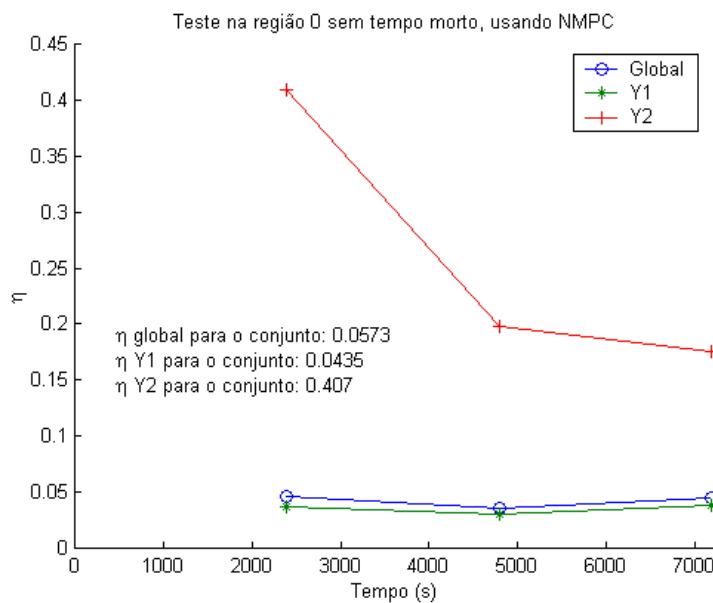


Figura 5.25. Índices de desempenho fragmentados e para todo o conjunto de dados, como resultado da utilização do NMPC na Região 0.

5.6.2. Caso 2

No Caso 2, com o sistema alterado pela adição de tempo morto dado por (5.4), os testes foram realizados para as três regiões de operação. As matrizes de interação utilizadas para o cálculo dos índices de desempenho são as mesmas estimadas com dados dos testes com controladores PI no Caso 2. Os dados obtidos com o experimento encontram-se ilustrados nas figuras a seguir, iniciando pela Região 0.

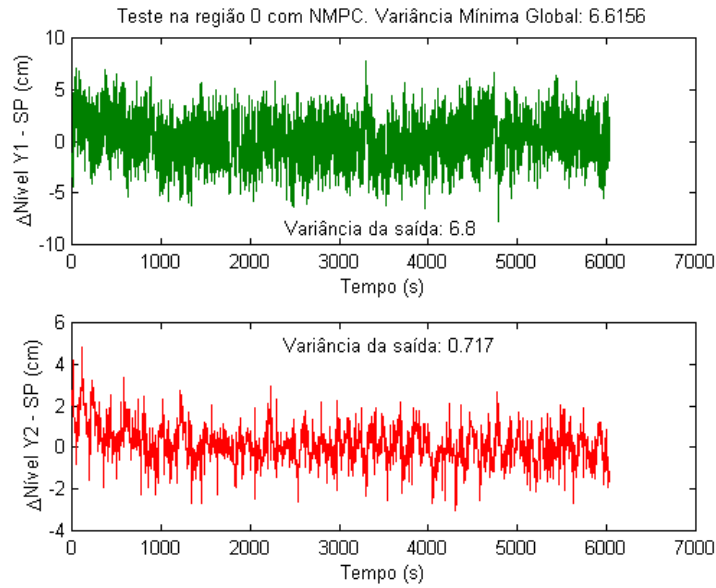


Figura 5.26. Dados coletados para o experimento com o NMPC na Região 0 com adição de tempo morto extra.

Pode-se notar claramente que os dados apresentam um ligeiro desvio do *setpoint* na região inicial do teste, até 1000s. Problemas na execução do teste levaram o sistema a sair do *setpoint* no meio do experimento e, para que o tamanho da amostra de dados fosse suficiente, dados do sistema em estado transiente tiveram que ser incluídos na análise de desempenho. Este desvio do ponto de operação também é capturado pela estimativa dos índices de desempenho, conforme ilustrado na Figura 5.27.

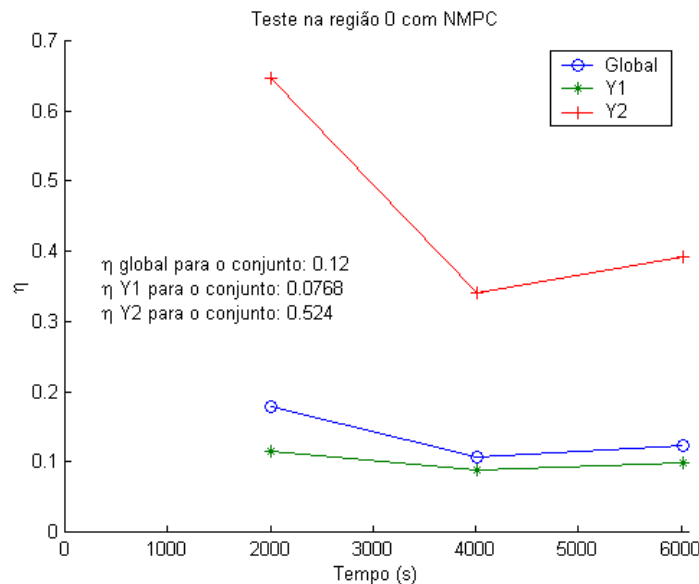


Figura 5.27. Índices de desempenho fragmentados e para todo o conjunto de dados para a Região 0, Caso 2, utilizando NMPC.

Apesar do problema ocorrido durante o experimento, os índices de desempenho para a Região 0, Caso 2 mostram-se melhores quando comparados aos equivalentes utilizando controlador PI.

Os resultados referentes à Região 1 são apresentados a seguir.

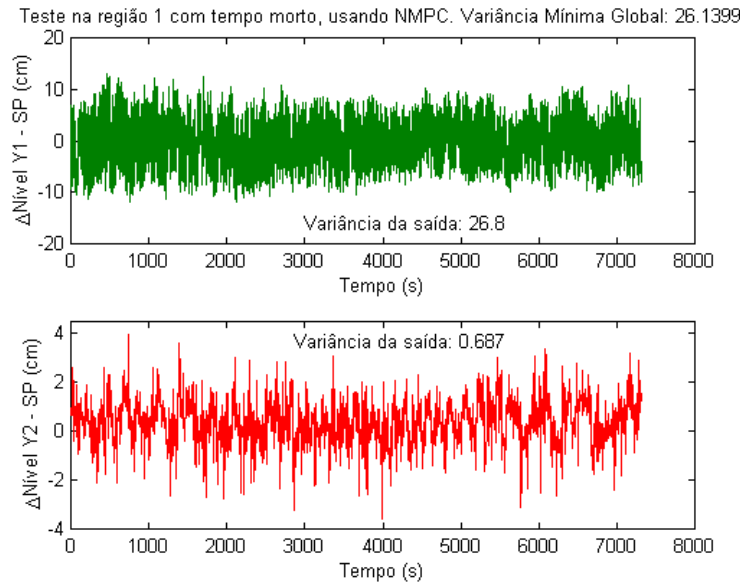


Figura 5.28. Dados coletados com o uso do NMPC na Região 1.

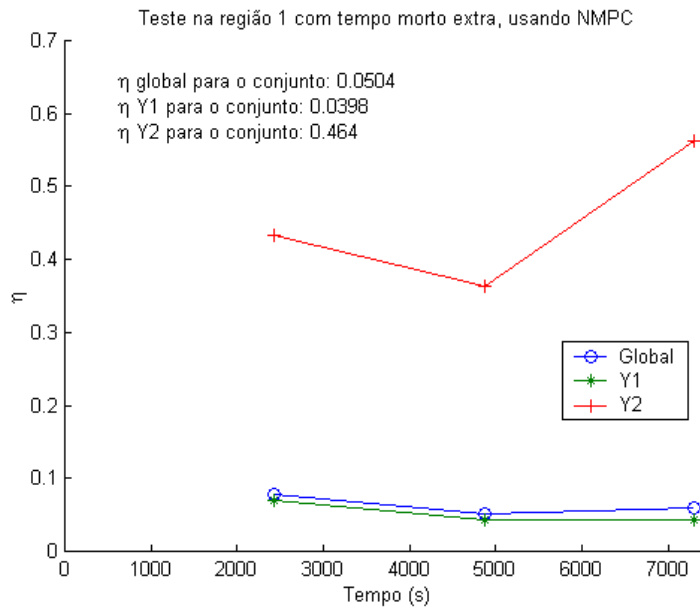


Figura 5.29. Índices de desempenho fragmentados e para todo o conjunto de dados coletados na Região 1.

Novamente os índices calculados são melhores que os correspondentes na utilização de controladores PI. Uma diminuição na amplitude do padrão oscilatório entre 3000s e 5000s é refletido na ligeira melhora dos índices de desempenho, especialmente da malha de Y2, que a propósito, demonstrou-se mais sensível à este tipo de alteração ao longo de todos os experimentos realizados.

Por fim, os resultados referentes à Região 2 são ilustrados nas duas figuras a seguir.

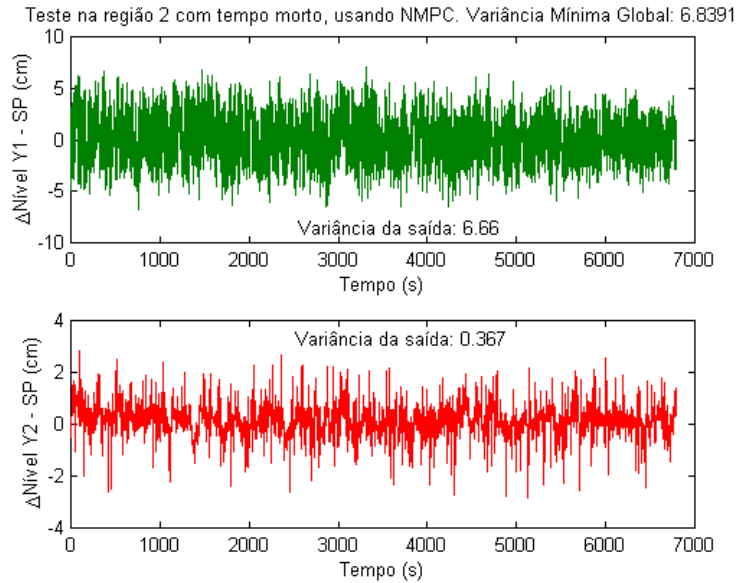


Figura 5.30. Dados relativos ao experimento com NMPC na Região 2.

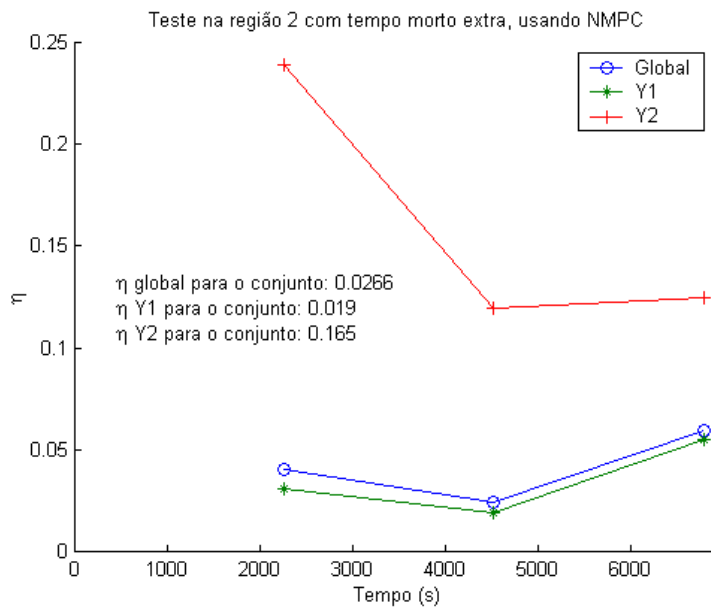


Figura 5.31. Índices de desempenho fragmentados e para todo o conjunto de dados coletados na Região 2.

Os índices de desempenho são significativamente melhores que os produzidos por controladores PI na mesma região, com a estrutura de tempo morto adicional.

Analisando-se o comportamento dos índices nas três regiões nota-se que os resultados apresentam-se bastante similares. O índice de desempenho global e da malha Y1 permanecem no mesmo patamar nas três regiões, enquanto que o índice correspondente à malha Y2 é similar nas regiões 0 e 1 e apresenta melhora na região 2.

5.7. Influência da matriz de pesos

Para verificar-se a influência que a matriz de pesos tem sobre o resultado do índice de desempenho, realizou-se mais um teste na unidade.

Como a vazão de alimentação da unidade, F_{SET} é a variável manipulada utilizada pelo controlador da malha de Y1 para controlar o nível de T1, mas que também influencia no nível de T2, é de se esperar que o desempenho global da unidade seja mais fortemente influenciado pelo desempenho da malha de Y1. Os gráficos da **Erro! A origem da referência não foi encontrada.** e da **Erro! A origem da referência não foi encontrada.** confirmam esta hipótese, na medida em que o índice de desempenho global é mais próximo da malha de Y1 e é mais afetado por mudanças no seu desempenho. Assim, é interessante verificar a influência de uma matriz de pesos neste conjunto de dados. Supondo-se então que, hipoteticamente, a saída Y2 é dez vezes mais importante que Y1 ($W = \text{diag}(1, 10)$), os resultados do índice de desempenho para o conjunto de dados representativo da operação normal da unidade (**Erro! A origem da referência não foi encontrada.**) mostram que o desempenho global mostra uma melhora, deslocando-se agora em direção à curva de Y2. No entanto, o desempenho global ainda segue os movimentos da curva de Y1, mostrando que embora Y2 tenha agora mais importância, Y1 ainda é a variável que afeta mais o desempenho da unidade. A nova matriz de interação, calculada com os dados ponderados é:

$$D = \begin{bmatrix} -0.99q & -0.1q \\ -0.1q^2 & 0.99q^2 \end{bmatrix}$$

Com estes resultados pode-se concluir que, para este sistema, não há nenhuma razão para investir-se na melhoria da malha de Y2, uma vez que em condições normais de operação, seu desempenho está próximo da variância mínima. Já para a malha de Y1, uma melhora no seu desempenho é recomendada e ocasionaria uma melhora substancial no desempenho global, pois como foi visto, esta malha influencia o índice de desempenho global fortemente e também o desempenho da malha Y2. Como este é um sistema com bastante interação, pode ser necessário reajustar a malha Y2 para aliviar as ações de controle de forma com que a malha de Y1 possa ser ajustada mais fortemente. Também é importante levar em conta que a malha Y1 é um sistema em cascata, no qual o desempenho da malha externa também depende do desempenho da malha interna, que não foi avaliado. Uma metodologia de avaliação de desempenho de malhas cascata que leva em conta este fator pode ser encontrado em Ko e Edgar, 2000.

5.8. Conclusões

Nesta série de testes um diagnóstico que se pode fazer da avaliação da qualidade dos objetivos de controle é que o desempenho geral das malhas de controle da unidade não é função exclusiva do ajuste dos controladores. Na ocasião hipotética em que a unidade de testes fosse uma unidade industrial de importância, utilizando as mesmas malhas PI, a melhora global do controle poderia ser atingida corrigindo-se problemas tais como a bomba mal dimensionada e a válvula atuando numa faixa estreita e sem posicionador. Em casos de impossibilidade da parada da unidade para manutenção, um remédio seria a utilização de filtros na retroalimentação das malhas de modo a cortar as frequências de oscilação características das regiões de operação.

Um procedimento passível de automação foi implementado computacionalmente e testado em escala experimental. O procedimento de avaliação apresentado possui o inconveniente de precisar de uma interferência no processo para a estimação das matrizes de interação. Apesar disto, unidades industriais com múltiplo propósito operam geralmente em mais de um ponto estacionário e as mudanças de um ponto para outro podem ser utilizadas como dados apropriados para estimação da estrutura de tempo morto. De outra maneira, também pode-se empregar dados históricos provenientes de testes de identificação com outras finalidades para este fim. Até mesmo o conhecimento prévio e a experiência prática com o processo podem ser empregados como uma estimativa da matriz de interação. Este argumento é reforçado pelo fato de que um número significativo de unidades industriais apresentar matrizes de interação diagonais ou aproximadamente diagonais (Huang e Shah, 1999), onde basta conhecer o tempo morto, em termos de intervalos de amostragem, de cada malha individualmente.

A própria adoção de outras estratégias de controle também pode se demonstrar eficaz na melhoria da variabilidade de um processo, conforme evidenciado neste capítulo com a adoção de controle preditivo não linear. Foi demonstrado que a adoção da estratégia com NMPC produz reduções de variabilidade da ordem de até 500 a 1000% no caso do sistema testado. Dependendo dos objetivos de controle e do produto final em questão, reduções de variabilidade de centésimos destes valores podem resultar em retornos econômicos vultuosos. Tais melhorias poderiam ser conseguidos com simples recalibração de controladores PI, projeto de um controlador PI multivariável e/ou manutenção de instrumentação, não sendo nem necessária a adoção de outras estratégias de controle. Um teste deste tipo não foi feito neste trabalho e deve ser levado a cabo para quantificar tal benefício.

No Apêndice D encontram-se os resultados das análises de correlação feitas em cima das saídas produzidas pelos controladores e dos resíduos das estimativas do método FCOR multivariável. O teste residual mostra que todos os resíduos são brancos, pois as suas funções de autocorrelação (ACF) permanecem dentro do intervalo de confiança de 95%. Isto indica que uma aproximação linear na região em torno dos pontos de operação é válida e que os modelos lineares ajustados pelo algoritmo FCOR e os índices estimados representam bem o sistema e o desempenho real frente à variância mínima. As funções de autocorrelação das saídas evidenciam mais uma vez a superioridade do NMPC frente aos PI's utilizados nos

testes. O NMPC traz mais rapidamente a ACF das saídas para próximo de zero. Enquanto a ACF da saída Y2 geralmente atinge um patamar entre 0,6 e 0,8 para os controladores PI, indicando componente oscilatório mais rápido do que o controlador pode eliminar, a ACF utilizando o NMPC tem uma queda lenta e permanente até aproximar-se de zero. Já a ACF de Y1 é levada à zero muito mais rapidamente com o NMPC que com os controladores PI utilizados no teste. Já os dados de correlação cruzada entre as duas saídas apenas comprovam que o sistema tem uma considerável interação entre as duas malhas, recomendando o uso de estratégias com controladores acoplados.

Capítulo 6 Conclusões gerais e sugestões

Ao concluir este trabalho, espero ter conseguido transmitir de uma maneira abrangente e didática todo o conhecimento adquirido ao longo deste estudo. Uma visão compreensiva do assunto e praticamente toda a base teórica necessária para sua compreensão foi fornecida nesta dissertação, de modo que futuros trabalhos a serem desenvolvidos nesta área possam tomá-la como referência para eliminar entraves iniciais e atingir objetivos mais específicos mais rapidamente.

6.1. Principais Conclusões

Como resultado dos diversos experimentos realizados e a avaliação crítica das técnicas existentes, pode-se formular diversas conclusões e sugestões sobre a análise de desempenho de malhas de controle.

Existem diversas lacunas, pontos que devem ser melhor desenvolvidos e técnicas que podem ser melhoradas. Fornecer sugestões para sanar estas deficiências também foi um dos objetivos desta dissertação. Do abrangente estudo feito neste trabalho, pode ser elaborado um procedimento metódico de aplicação das técnicas apresentadas. Deste modo é possível produzir resultados bastante condizentes com a realidade das unidades industriais, facilitar sua interpretação e permitir conclusões bem suportadas sobre os problemas de variabilidade, de modo a garantir total segurança na tomada de decisões. É preciso manter sempre a postura crítica de um bom engenheiro, pois “não trata-se de mágica, e sim de engenharia”. Isto quer dizer que o simples acompanhamento de um índice de desempenho sozinho não permite chegar com certeza a conclusões definitivas sobre a qualidade de uma malha. Por esta razão, foram apresentadas neste trabalho diversas técnicas auxiliares e complementares aos índices de desempenho, tais como procedimentos de detecção de oscilações, análises estatísticas, análise espectrais, horizonte estendido etc que permitem a obtenção de diagnósticos mais precisos sobre o problema de uma malha. Mesmo a experiência de operadores ou engenheiros não é dispensável e deve ter peso na avaliação final do problema de variabilidade quando

possível. Apesar da quantidade ínfima de informações que esta auditoria de malhas requer, é de se admirar a quantidade de informações que fornece em retorno, se os resultados forem devidamente interpretados. Este fato pode gerar outro problema, que é justamente a qualificação técnica necessária para tirar as devidas conclusões sobre os resultados apresentados. É sempre importante salientar que as conclusões não devem ser formuladas unicamente baseando-se na metodologia de variância mínima. As condições em que os valores dos índices foram obtidos devem principalmente ser levadas em conta. Também é de vital importância certificar-se dos resultados de outras metodologias, tais como as técnicas complementares apresentadas neste trabalho, para suportar as conclusões finais sobre desempenho de malhas.

Diversas questões práticas, pertinentes a problemas que surgem durante implementação industrial automatizada em larga escala, foram levantadas e recomendações foram sugeridas ou compiladas, de forma a orientar a configuração de um sistema que realize a tarefa de monitoramento das malhas sem a necessidade de grandes intervenções do usuário ou de conhecimento detalhado sobre cada malha.

Controle multivariável levando em conta interações é uma estratégia que oferece grandes benefícios em relação às malhas SISO simples. No entanto sua utilização não é tão comum devido ao seu nível de complexidade e à conseqüente necessidade de conhecimento mais especializado e caro na sua implementação e manutenção. Apesar disto, a compreensão dos sistemas multivariáveis tem se popularizado e seu uso tende a se tornar mais comum. Técnicas de avaliação de desempenho de malhas multivariáveis foram apresentadas. As ferramentas mais adequadas para um procedimento o mais automatizado possível foram estudadas e este procedimento, supondo-se desconhecimento da estrutura de tempo morto, deve ser composto de uma identificação em malha fechada dos parâmetros de Markov do sistema com o algoritmo ARMarkov-LS, seguido da determinação de sua ordem utilizando técnicas baseadas em SVD e posterior fatoração da matriz de interação a partir dos primeiros coeficientes de Markov cuja combinação linear é não singular.

Baseado nos estudos realizados neste trabalho, uma série de conclusões podem ser tiradas referentes ao atual desenvolvimento das técnicas de auditoria de malhas. A seguir serão compiladas as principais conclusões:

- **CVM como referência de desempenho.** Há vantagens e desvantagens em se utilizar um controlador de variância mínima (CVM) como referência de desempenho. Como é sabido o CVM é tanto uma função do modelo da planta que se está controlando quanto do tipo de distúrbio. Para uma dada planta e um padrão de distúrbios, o CVM será o controlador capaz de minimizar o erro quadrático médio da malha de controle sem penalizar as ações de controle. Desta forma, se houver uma modificação do padrão de distúrbio ou uma modificação da dinâmica da planta haverá uma modificação do CVM. Esta dupla dependência dificulta, até certo ponto, a caracterização da origem da queda de desempenho de uma determinada malha de controle, uma vez que um indicativo de mau desempenho pode ser decorrente de uma modificação do padrão de distúrbio ou da dinâmica da planta. Variações no padrão de distúrbio podem ser produzidos tanto por mudanças de *setpoint* quanto por distúrbios externos (por exemplo, composição, temperatura, vazão das

correntes de entrada) e ocorre frequentemente em plantas industriais. Isto dificulta a correta definição da necessidade ou não de se reajustar o controlador, a qual deveria ocorrer principalmente como consequência de uma mudança do comportamento dinâmico da planta do que uma modificação momentânea e transitória do padrão de distúrbio vigente. A principal vantagem de se utilizar o CVM como referência está em este ser um limite máximo de diminuição de variância, o qual consiste em um critério de desempenho bastante utilizado. Sem sombra de dúvida, a principal vantagem reside no fato de se poder estimar variância mínima utilizando-se apenas dados corriqueiros de planta. Este é o principal motivo para a sua utilização.

- **Ações de Controle.** As técnicas de avaliação de desempenho de malhas baseadas em CVM não levam em conta as ações de controle. Esta importante informação da malha de controle é praticamente desconsiderada nas atuais técnicas de auditorias de malhas. A partir das ações de controle pode-se ter um ótimo indicador da robustez e do desempenho da malha de controle. Poucas técnicas e propostas exploram esta importante informação.
- **Características do Controlador.** Questões como o impacto da ordem do controlador e características de ajuste não são levados em conta nas atuais técnicas de auditoria de malhas. As técnicas não permitem que se diga o quanto o atual PI está distante em termos de desempenho do melhor desempenho alcançável por um PI. Até que ponto um aumento da ordem do controlador irá melhorar o desempenho da malha. Muitas vezes no processo (por exemplo, tanques pulmões) respostas lentas são requeridas e benéficas para o funcionamento do processo. Estas respostas são utilizadas como amortecedores e filtros dos distúrbios dos processos. Desta forma, um CVM está longe de ser considerado como o padrão de desempenho adequado.
- **Falta de um índice Absoluto.** Sem um prévio conhecimento da malha de controle que está sendo avaliada, os índices baseados em CVM não permitem determinar se o desempenho é precário ou não. Por exemplo, uma malha pode ser considerada ruim tendo $h = 0,7$, enquanto que outra com este mesmo índice poderá ser considerada ideal. Também existe uma baixa sensibilidade para uma determinada malha para se diferenciar uma malha boa ou ruim, isto é uma malha considerada boa com $h = 0,5$ pode ser considerada ruim com $h = 0,55$, enquanto que há casos em que a diferença entre o bom e ruim é capturada numa faixa muito mais ampla de variação do índice. Em suma, hoje os índices necessitam ser comparados de forma relativa, enquanto que o ideal seria ter-se um indicador absoluto de desempenho, de tal forma que um determinado valor levasse sempre a mesma conclusão independentemente da malha de controle que está sendo utilizada.

6.2. Sugestões para Futuros Trabalhos

Diversos pontos ficam em aberto para posterior desenvolvimento. Uma sumarização é apresentada a seguir:

- Integração de todas as técnicas apresentadas neste estudo em uma ferramenta de *software* que também disponha de um sistema automatizado de coleta remota de dados e verificação de sua integridade, bem como de uma interface gráfica com o usuário apropriada. Também é desejável que tal sistema possua um sistema de alertas que detectam mudanças estatisticamente significativas no processo e elabore relatórios detalhados ao seu usuário. Sistemas similares já existem e foram citados no Capítulo 1, mas desenvolvimento de tecnologia nacional própria é sempre atraente.
- Avaliação de controle de estratégias de controle diversas foi apenas referenciada neste trabalho. Estudos detalhados ainda precisam ser apropriadamente efetuados para sistemas *feedforward*, sistemas em cascata, sistemas em que o usuário especifica o desempenho desejado como *benchmark*, sistemas operando com restrições nas variáveis e sistemas tais como tanques-pulmão, onde o objetivo não é a variabilidade mínima. Propostas podem ser encontradas na literatura, mas seu emprego prático ainda não foi reportado.
- Outras técnicas complementares à abordagem baseada em variância mínima podem ser agregadas à “caixa de ferramentas” da avaliação de desempenho. Como exemplo, pode-se citar o procedimento de detecção de malhas de controle excessivamente lentas proposto por Hägglund (1999) e o procedimento qualitativo de análise de forma para avaliar especificamente problemas de agarramento de válvulas de controle, proposto por Rengaswamy et al. (2001). As questões práticas de implementação destas ferramentas ainda devem ser verificadas.
- Avaliação de desempenho de controladores preditivos baseados em modelo. O uso de controladores preditivos baseados em modelos (MPC) vem se tornando cada vez mais popular com o crescimento da capacidade de processamento dos computadores de processo e especialização do corpo técnico no sítio industrial. Contudo, controladores do tipo PID ainda são necessários no nível mais básico de controle, de forma que o sucesso de estratégias MPC está frequentemente limitado ao desempenho destes controladores SISO. Assim, os procedimentos de avaliação destes controladores PID podem servir de ferramenta para auditoria pré-instalação de MPC's. A utilização de um método baseado em variância mínima, tal como o FCOR multivariável empregado neste trabalho, não provê no entanto, informações a respeito do MPC em si, da qualidade do modelo e do esforço de controle utilizado. Se um controlador preditivo apresenta desempenho insatisfatório, é interessante contar com uma metodologia que possibilite diagnosticar se o problema está na validade do modelo utilizado ou no ajuste dos parâmetros do MPC, bem como uma metodologia que leve em conta restrições impostas na otimização das ações de controle.

A propósito do último ponto, para levar em conta restrições no esforço de controle, o método proposto por Huang e Shah (1999), *LQG Benchmark*, é uma alternativa. Uma solução completa para o problema de auditoria de MPC's pode passar pela utilização do *LQG Benchmark* como um complemento à um procedimento que diagnostique onde encontra-se o

problema do MPC, no ajuste ou no modelo, para formar uma ferramenta abrangente de avaliação de desempenho de MPC's.

Apêndice A - Cálculo da matriz de interação

Este algoritmo, originalmente concebido por Rogozinski et al. (1987), calcula a estrutura de tempo morto de um sistema linear multivariável a partir dos coeficientes do numerador de uma representação RMF da matriz de transferência do sistema. A matriz de interação nilpotente resultante é não-única para o sistema e é um produto de matrizes polinomiais nilpotentes de primeira ordem. É assumido que o denominador polinomial na representação RMF é monômico, o que pode ser sempre satisfeito com o uso de alguma álgebra. O algoritmo opera nos coeficientes do numerador da representação RMF do sistema e não em matrizes polinomiais ou funções racionais. Assim, cálculos envolvendo polinômios são substituídos por cálculos com matrizes reais, o que facilita a implementação do algoritmo em métodos computacionais. Peng e Kinnaert (1992) estendem o conceito de matriz de interação nilpotente e mostram a existência da matriz de interação unitária.

A.1 Definições

Seja $T(q)$ uma matriz de transferência $p \times m$ que descreve um sistema linear invariante com o tempo com m entradas e p saídas: Assume-se que:

- I. $T(q)$ tem posto cheio: $\text{rank } T(q) = \min(m, p)$
- II. $T(q)$ é próprio: $\lim_{q \rightarrow \infty} T(q) < \infty$

Definindo uma matriz polinomial de primeiro grau, $U(q)$, chamada *matriz polinomial de deslocamento de linhas (row shift polynomial matrix, r.s.p.m)* de ordem k_i , onde:

$$U(q) = U_0 q + U_1 = \begin{bmatrix} 0 & I_r \\ qI_{k_i} & 0 \end{bmatrix} \quad (\text{A.1})$$

As matrizes U_0 e U_1 são definidas através das matrizes de coeficientes:

$$U = \begin{bmatrix} U_0 \\ U_1 \end{bmatrix} = \begin{bmatrix} 0_r \\ I_p \\ 0_{k_i} \end{bmatrix}, \quad p = r + k_i \quad (\text{A.2})$$

em que U_0 e U_1 são de dimensão $p \times p$, I_p é a matriz identidade $p \times p$ e 0_r é uma matriz de zeros com r linhas.

O determinante da r.s.p.m. de ordem k_i é dado por: $\det(U(q)) = q^{k_i}$

Se $T(q)$ satisfaz (I) e (II), então uma matriz polinomial $D(q)$, $p \times p$ que possui as seguintes propriedades:

$$\lim_{q \rightarrow \infty} D(q)T(q) = K \quad (\text{A.3})$$

onde K é uma matriz constante de posto cheio, e:

$$\det(D(q)) = cq^k \quad (\text{A.4})$$

é chamada de “matriz de interação nilpotente” para o sistema definido por $T(q)$. c é uma constante.

A matriz de transferência $T(q)$ pode ser fatorada (não unicamente) como RMF da forma:

$$T(q) = L(q)R^{-1}(q) \quad (\text{A.5})$$

Onde: $R(q) = Iq^n + R_1q^{n-1} + \dots + R_n$ é uma matriz polinomial monômica, correspondendo a realização não-mínima do sistema. O numerador da descrição RMF é assumido ser uma matriz polinomial:

$$L(q) = L_0q^n + L_1q^{n-1} + \dots + L_n \quad (\text{A.6})$$

e pode ser escrito como uma matriz bloco de coeficientes:

$$L = \begin{bmatrix} L_0 \\ \vdots \\ L_n \end{bmatrix} \quad (\text{A.7})$$

onde n é o grau da matriz polinomial do denominador R na representação RMF. Em geral, L_i não é de posto cheio e, em particular, os primeiros coeficientes podem ser zero.

Como o denominador da matriz em (A.5) é assumidamente monômico, a matriz de interação nilpotente pode ser avaliada do numerador $L(q)$:

$$\begin{aligned} \lim_{q \rightarrow \infty} D(q)L(q)R^{-1}(q) &= \lim_{q \rightarrow \infty} D(q)q^{-n}L(q)R^{-1}(q) \lim_{q \rightarrow \infty} q^n R^{-1}(q) \\ &= \lim_{q \rightarrow \infty} D(q)q^{-n}L(q) = K \end{aligned}$$

Para a definição de matriz nilpotente, referir-se à Seção 4.2.1.

A.2 O algoritmo

Para uma matriz de interação $T(q)$ satisfazendo (I) e (II) existe uma matriz de interação nilpotente consistindo de t fatores:

$$D(q) = S^{(t)}(q)S^{(t-1)}(q) \cdots S^{(1)}(q) \quad (\text{A.8})$$

onde:

$$S^{(i)}(q) = U^{(i)}(q)Q^{(i)} \quad (\text{A.9})$$

e $U^{(i)}(q)$ é uma r.s.p.m. de ordem k_i e $Q^{(i)}$ é uma matriz real constante $p \times p$. A prova pode ser encontrada em Rogozinski et al. (1987).

O algoritmo pode ser então iniciado fornecendo-se a matriz bloco de coeficientes do numerador.

Fazendo $i = 0$, $L^{(0)}(q) = L(q)$, e $D^{(0)} = I_p$ para iniciar o algoritmo. A i -ésima iteração no cálculo de $D(q)$ baseada em (A.8) é considerada:

Passo 1:

- Se $r_i = \text{rank } L_0^{(i-1)} = \min(m, p)$, o algoritmo termina e a matriz de interação nilpotente é $D(q) = D^{(i-1)}(q)$, com $t = i - 1$.
- Se $r_i < \min(m, p)$, fatorar $L_0^{(i-1)}$ utilizando um método de fatoração tal como ortogonalização Householder QR, eliminação Gaussiana LDU ou decomposição em valores singulares SVD, em:

$$L_0^{(i-1)} = (Q^{(i)})^{-1} \begin{bmatrix} 0_i \\ L^{(i)} \\ L_{0D}^{(i)} \end{bmatrix}, \quad \text{ou seja, } Q^{(i)}L_0^{(i-1)} = \begin{bmatrix} 0_i \\ L^{(i)} \\ L_{0D}^{(i)} \end{bmatrix} \quad (\text{A.10})$$

onde $Q^{(i)}$ é uma matriz $p \times p$ não-singular (geralmente unitária), $k_i = p - r_i$ e 0_i é uma matriz de zeros com k_i linhas.

Passo 2:

- Pré-multiplicar $L^{(i-1)}(q)$ pela matriz $Q^{(i)}$:

$$\tilde{L}(q) = Q^{(i)}L^{(i-1)}(q) \quad (\text{A.11})$$

O primeiro coeficiente de $\tilde{L}(q)$ é agora igual ao lado direito de (A.10).

Passo 3:

- Pré-multiplicar $\tilde{L}(q)$ pela r.s.p.m. de ordem k_i :

$$L^{(i)}(q) = U^{(i)}(q)\tilde{L}(q) \quad (\text{A.12})$$

Esta multiplicação desloca os coeficientes de $\tilde{L}(q)$, L , para cima por k_i linhas de zeros. Atualizar a matriz de interação:

$$D^{(i)}(q) = S^{(i)}(q)D^{(i-1)}(q) \quad (\text{A.13})$$

Este passo termina a i -ésima iteração. Combinando (A.10) até (A.13), a i -ésima iteração resulta em:

$$L^{(i)}(q) = U^{(i)}(q)Q^{(i)}L^{(i-1)}(q) = S^{(i)}(q)L^{(i-1)}(q) = D^{(i)}(q)L(q) \quad (\text{A.14})$$

onde $S^{(i)}(q)$ e $D^{(i)}(q)$ são definidos por (A.9) e (A.14).

A iteração final ($t = i - 1$) resulta:

$$L^{(t)}(q) = D(q)L(q) \quad (\text{A.15})$$

onde $D(q)$ definido por (A.8) é $D(q) = D^{(t)}(q)$. Assim, (A.3) é satisfeito e $K = L^{(t)}$. Notar também que $\det(U^{(i)}(q)) = z^{k_i}$.

Denotando $c_i = \det(Q^{(i)})$, tem-se que:

$$\det(D^{(t)}(q)) = \prod_{i=1}^t c_i \prod_{i=1}^t z^{k_i} \quad (\text{A.16})$$

De onde tira-se que as constantes c e k definidas em (A.4) são dadas por:

$$c = \prod_{i=1}^t c_i \quad (\text{A.17})$$

e:

$$k = \text{deg}(\det(D(q))) = \sum_{i=1}^t k_i \quad (\text{A.18})$$

Peng e Kinnaert (1992) mostram a existência da matriz de interação unitária. Se uma matriz nilpotente conforme definida em Rogozinski et al. (1987) satisfaz:

$$D^T(q^{-1})D(q) = I_p \quad (\text{A.19})$$

então $D(q)$ é uma matriz de interação unitária. Em Rogozinski et al. (1987) foi demonstrada a existência da matriz de interação nilpotente. $Q^{(i)}$ é unitária se for calculada com uma fatoração QR normalizada. Além disto, $U^{(i)}(q)$ é também uma matriz polinomial unitária, uma vez que satisfaz $U^T(q^{-1})U(q) = I_p$. Já que o produto de matrizes unitárias deve ser unitário, $D(q)$ é então uma matriz de interação unitária.

Em Peng e Kinnaert (1992) é observado que embora o algoritmo proposto por Rogozinski et al. (1987) opere no numerador da representação RMF da matriz de transferência, ele também pode operar nas matrizes de espaço de estado do sistema com uma pequena modificação.

Como o algoritmo depende da fatoração RMF não-única da matriz de interação, a matriz de interação unitária é também não-única. No entanto, pode-se demonstrar que duas matrizes de interação unitárias estão relacionadas por uma matriz unitária real:

$$\bar{D}(q) = \Gamma D(q) \quad (\text{A.20})$$

onde Γ é uma matriz unitária $p \times p$ real.

Apêndice B - Estimação da matriz de interação

No Apêndice A foi apresentada uma maneira de se calcular a matriz de interação unitária. O conhecimento da matriz de interação é importante não só na avaliação de desempenho de malhas, mas em diversas estratégias de controle de alta complexidade. No entanto, o algoritmo demonstrado requer a representação RMF do numerador da matriz de transferência, o que é equivalente a identificação completa do sistema. Para o monitoramento de desempenho de controladores na sua proposta mais automatizada e não invasiva, é desejável que se empregue uma metodologia capaz de determinar uma boa aproximação da matriz de interação utilizando-se de variáveis medidas e com o mínimo de intervenção no processo. Esta idéia, proposta em Huang et al. (1997a) e Huang e Shah (1999), encontra-se aqui reproduzida.

B.1 Introdução

No caso multivariável, a noção de tempo morto corresponde aos primeiros coeficientes da resposta ao impulso ou parâmetros de Markov matriciais cuja combinação linear é não singular. Esta combinação linear de matrizes da resposta ao impulso pode ser expressada na forma de uma matriz polinomial. O determinante desta matriz polinomial tem como suas raízes os zeros infinitos do sistema multivariável discreto. O algoritmo descrito a seguir visa então determinar o número de parâmetros de Markov necessários para que o tempo morto seja estimado, ou seja, o número de coeficientes da resposta ao impulso cuja combinação linear tem posto cheio e calcular a matriz de interação utilizando dados de malha aberta ou fechada com uma excitação. Uma vez determinado o número de parâmetros de Markov necessários, este método de estimação da matriz de interação unitária utiliza o mesmo algoritmo do Apêndice A entrando com os parâmetros de Markov ao invés do numerador da matriz de transferência.

B.2 Determinação da ordem da matriz de interação

Uma matriz de transferência $p \times m$ pode ser representada em termos de parâmetros de Markov como:

$$T = \sum_{i=0}^{\infty} G_i q^{-i-1} \quad (\text{B.1})$$

Onde o coeficiente de ordem zero, G_{-1} é omitido por ser zero. De (A.3) tem-se que:

$$\lim_{q \rightarrow \infty} D(q)T(q) = \lim_{q \rightarrow \infty} [D_0 q^d + D_1 q^{d-1} + \dots + D_{d-1} q] [G_0 q^{-1} + G_1 q^{-2} + \dots] = K \quad (\text{B.2})$$

onde K tem posto cheio: $\text{rank}(K) = \min(p, m)$. Deste modo:

$$\begin{aligned} D_0 G_0 &= 0 \\ D_1 G_0 + D_0 G_1 &= 0 \\ &\vdots \\ D_{d-1} G_0 + \dots + D_1 G_{d-2} + D_0 G_{d-1} &= K \end{aligned} \quad (\text{B.3})$$

Resolvendo-se as equações algébricas acima chega-se a solução geral da matriz de interação. Escrevendo (B.3) em forma matricial:

$$\begin{bmatrix} D_{d-1} & \dots & D_0 \end{bmatrix} \begin{bmatrix} G_0 & 0 & 0 & \dots & 0 \\ G_1 & G_0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \\ G_{d-2} & G_{d-3} & \dots & \ddots & 0 \\ G_{d-1} & G_{d-2} & \dots & \dots & G_0 \end{bmatrix} = \begin{bmatrix} K & 0 & \dots & 0 \end{bmatrix} \quad (\text{B.4})$$

ou simplesmente:

$$\underline{D}\underline{G} = \underline{K} \quad (\text{B.5})$$

onde \underline{G} é uma matriz Toeplitz-bloco. Se G_0 não tem posto cheio, então em adição aos zeros infinitos decorrentes do *hold* de ordem zero, pelo menos mais um zero infinito existe na matriz de transferência.

A inversão direta para resolver (B.5) não é possível já que \underline{G} tem posto deficiente. A existência de solução para (B.5) também depende da ordem da matriz de interação, ou seja, o tamanho de \underline{G} tal que exista ao menos uma solução exata de \underline{D} . Para determinar a ordem da matriz de interação, pode ser usada uma técnica baseada em decomposição SVD.

A decomposição SVD da matriz Toeplitz-bloco resulta em:

$$\underline{G} = USV^T = [U_1 \ U_2] \begin{bmatrix} S_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \quad (\text{B.6})$$

onde $[U_1 \ U_2]$ e $[V_1 \ V_2]^T$ são matrizes ortogonais. As colunas de U_2 sobrepõe o espaço nulo de \underline{G} (de forma que $U_2^T \underline{G} = 0$), S_r é uma matriz diagonal de posto cheio com os r valores singulares de \underline{G} , e as linhas de V_1^T sobrepõe o espaço das linhas de \underline{G} . Assim, pode-se calcular V_2 e U_2 da seguinte forma:

$$\begin{aligned} V_2 &= \text{null}(\underline{G}) \\ U_2 &= \text{null}(\underline{G}^T) \end{aligned} \quad (\text{B.7})$$

A existência de solução para (B.5) também requer que:

1. $\text{rank}(\underline{G}) \geq \text{rank}(\underline{K}) = \text{rank}(K) = \min(p, m)$
2. cada linha de \underline{K} precisa estar dentro do espaço nulo produzido por V_1^T ou ortogonal ao espaço das linhas sobreposto por V_2^T :

$$\underline{K}V_2 = [K \ 0 \ \dots \ 0] \begin{bmatrix} V_{21} \\ V_{22} \\ \vdots \\ V_{2d} \end{bmatrix} = KV_{21} \quad (\text{B.8})$$

onde V_{21} é a partição superior de V_2 com número de linhas igual ao número de colunas de T . Então a condição expressada por (B.8) é equivalente à:

$$KV_{21} = 0 \quad (\text{B.9})$$

Se K (ou T) é uma matriz quadrada ou é uma matriz $p \times m$ com $p > m$, a equação (B.9) pode ser simplificada para:

$$V_{21} = 0 \quad (\text{B.10})$$

Se, no entanto, estas condições não são satisfeitas, a matriz \underline{G} pode ser expandida adicionando-se mais parâmetros de Markov até que sejam satisfeitas. Desta forma, a ordem da matriz de interação, d , pode ser determinada das equações (B.9) ou (B.10).

B.3 Estimação da matriz de interação em malha fechada

O algoritmo para cálculo da matriz de interação proposto por Rogozinski et al. (1987) e apresentado no Apêndice A requer a representação RMF do numerador da matriz de transferência. Isto requer o conhecimento exato da matriz de transferência. A seguir é mostrado (Huang e Shah, 1999) que se a representação em termos de parâmetros de Markov é utilizada, o algoritmo pode ser empregado substituindo-se a matriz bloco dos coeficientes do numerador por uma matriz bloco com os parâmetros de Markov.

Para tanto, assume-se que:

- T é uma matriz de transferência polinomial, $p \times m$, racional e tem posto cheio: $\text{rank}(T(q^{-1})) = \min(p, m)$.
- T é própria: $\lim_{q^{-1} \rightarrow \infty} T(q^{-1}) < \infty$.

Uma matriz bloco com os d primeiros coeficientes de Markov é expressa como:

$$\Lambda = \begin{bmatrix} G_0 \\ G_1 \\ \vdots \\ G_d \end{bmatrix} \quad (\text{B.11})$$

Uma vez montada, pode-se alimentar o algoritmo do Apêndice A para fatorar a matriz de interação. Mesmo sem o conhecimento da ordem da matriz de interação, d , é possível obter a estimativa por tentativa e erro, aumentando o número de parâmetros de Markov até obter uma combinação linear não singular.

O algoritmo de fatoração requer agora apenas os d primeiros parâmetros de Markov. Como os primeiros parâmetros de Markov contribuem para a resposta transiente inicial do sistema, estes parâmetros caracterizam a dinâmica de alta frequência do processo. Assim, a matriz de interação representa o ganho de alta frequência de um sistema. Desta forma é desejável que uma técnica de identificação que produza boas estimativas a altas frequências seja empregada na estimação dos parâmetros de Markov a partir de dados de operação em malha fechada.

Para tanto, é necessário excitar o sistema com um sinal de relativa alta frequência. Computacionalmente, uma análise de correlação ou um modelo paramétrico que utiliza as entradas e saídas e pode ser posteriormente convertido para a forma de parâmetros de Markov, podem ser empregados para obter boas estimativas em altas frequências.

Em Huang et al. (1997a) e Huang e Shah (1999) é demonstrado que, para um processo multivariável como o da Figura 4.1, a matriz de interação em malha fechada, D_{mf} de W_t para Y_t é a mesma que a matriz de interação da matriz de transferência em malha aberta, T .

Se o sinal de excitação é inserido a partir do *setpoint*, a mesma conclusão pode ser feita para a matriz de transferência em malha fechada de Y_t^{SP} para Y_t . Isto ilustra o fato de que a estrutura de tempo morto é invariante com o controle *feedback*, ou seja, os parâmetros de Markov da malha aberta e da malha fechada são diferentes, mas suas combinações lineares geram a mesma matriz de interação. Com este resultado, a matriz de interação pode ser estimada diretamente em malha fechada ou em malha aberta. Levando-se em conta que alguns processos industriais podem ter altas ordens ou serem bastante não-lineares, estimação utilizando dados em malha aberta não são tão apropriados quanto dados provenientes da malha fechada operando em torno de um ponto de operação nominal no qual o sistema pode ser considerado aproximadamente linear. A matriz de interação assim obtida é uma boa estimativa então para este ponto de operação.

Quanto ao sinal de excitação, pode-se dizer que se o sinal é “branco” ou pode ser filtrado por análise de séries temporais, uma análise de correlação pode ser empregada para obter os parâmetros de Markov. Para plantas reais, um sinal aleatório tal qual um ruído branco, pode não ser adequado. Para tais casos é mais interessante uma série de mudanças do tipo degrau. Também é importante frisar que um sinal de excitação com baixa frequência pode resultar em uma estimação pior que quando utilizado um sinal de alta frequência.

Uma estratégia de identificação que forneça uma boa estimativa dos primeiros parâmetros de Markov é mais importante que uma boa estimativa de toda a matriz de transferência do sistema, a qual geralmente se compromete com um ajuste balanceado ao longo de uma faixa de frequência mais ampla. Computacionalmente, uma identificação direta dos primeiros d parâmetros de Markov é também mais desejável que uma identificação de toda a matriz de transferência para posterior obtenção das respostas ao impulso. Similarmente, a fatoração da matriz de interação dos primeiros parâmetros de Markov é preferível à fatoração da matriz de interação a partir da matriz de transferência.

B.4 Problemas de posto dos parâmetros de Markov

Os parâmetros de Markov estimados a partir de dados do processo em operação não são exatos devido à existência de distúrbios e isto faz com que a determinação numérica do posto da matriz Toeplitz-bloco \underline{G} seja arbitrária. Huang e Shah (1999) propõem que alguns resultados da literatura sejam utilizados para definir-se uma tolerância na determinação numérica do posto de \underline{G} .

Sendo H uma matriz hipotética com seus r valores singulares exatos provenientes de uma decomposição SVD, USV^T . Supondo que uma aproximação numericamente construída de H é obtida como $\hat{H} = H + \Delta H$, onde é sabido que $\|H - \hat{H}\| \leq a\|\hat{H}\|$. A constante a

representa uma medida da exatidão dos dados. Em termos de valores singulares de H e \hat{H} , esta inequação pode ser escrita como:

$$|V_i - \hat{V}_i| \leq aV_1$$

onde V_i e \hat{V}_i denotam os i -ésimos valores singulares teóricos e calculados respectivamente. Se \hat{V}_r é maior que aV_1 , mas \hat{V}_{r+1} é menor que este número, então V_i é positivo. Então o posto da matriz é pelo menos r . O próximo valor singular, V_{r+1} , pode ser zero. Tal valor de r pode ser escolhido como o posto numérico da verdadeira mas desconhecida matriz H .

Outro resultado útil (ver Huang e Shah, 1999, para referência) que pode ser utilizado afirma que a diferença dos valores singulares de $H + P$ e H é limitada pelo maior valor singular de P , onde P é uma matriz de perturbação. Este valor singular é referido como *valor limite*.

Os dois resultados podem ser aplicados para encontrar o posto de G_i e \underline{G} . O primeiro requer um conhecimento *a priori* de a . Para encontrar o valor de a pode-se recorrer a um valor empírico ou estatístico, tal como $2 / \sqrt{n}$, onde n é o tamanho da amostra. A abordagem em termos de valor limite é útil se algum conhecimento prévio da perturbação é conhecido. Análises de correlação ou ajustes de modelos FIR geralmente fornecem este resultado.

Em adição aos parâmetros de Markov expressados por (B.1), a análise de correlação também fornece o parâmetro de Markov correspondente a ordem zero de q , escrito como G_{-1} . Devido ao *hold* de ordem zero, este coeficiente é teoricamente zero e por isso não aparece em (B.1). Sua estimação no entanto não é zero devido ao efeito dos distúrbios. Desta forma, o valor estimado de G_{-1} fornece uma aproximação da matriz de perturbação P , e pode ser utilizado para determinação do posto dos parâmetros de Markov estimados.

A determinação da ordem da matriz de interação pode então seguir os seguintes passos da iteração i :

Passo 1:

Calcular o valor limite de tolerância a partir do coeficiente de ordem zero de G para decidir se um valor característico é significativamente diferente de zero: $tol = \max(\text{svd}(G_{-1}))$.

Alternativamente pode-se utilizar como valor limite $a = 2 / \sqrt{n}$.

Passo 2:

Montar \underline{G} com G_0, G_1, \dots, G_i e calcular o espaço nulo: $V_2 = \text{null}(\underline{G})$.

Passo 4:

Verificar se algum elemento de V_{21} é maior que o valor limite: $\text{abs}(V_{21}) > tol$?

Se V_{21} estiver dentro dos limites de tolerância, \underline{G} está correto e a ordem da matriz de interação é o número de coeficientes da resposta ao impulso utilizados para formar \underline{G} . O algoritmo está encerrado.

Do contrário, um novo parâmetro de Markov, G_{i+1} , deve ser acrescentado para formar uma nova matriz \underline{G} . Mas antes é conveniente modificar G_i fazendo seu menor valor característico igual a zero:

$$G'_i = US'V \quad (\text{B.12})$$

ou seja, fatora-se G_i com SVD em U , S e V , altera-se S para S' e monta-se G'_i novamente com (B.12). Esta aproximação é razoável quando um valor singular é significativamente menor que outros valores singulares e a matriz foi testada e mostrou-se ter posto deficiente.

O procedimento prossegue a partir do *Passo* 1, com a iteração $i + 1$ e a dimensão de \underline{G} acrescida de mais um parâmetro de Markov.

Com mais esta metodologia em mãos, a única ferramenta que falta para completar o processo de estimar a matriz de interação com dados de operação em malha fechada é um método adequado de estimação dos parâmetros de Markov nestas condições. Huang et al. (1997a) e Huang e Shah (1999) propõem o uso de análise de correlação ou ajuste de modelos FIR com tal finalidade. Porém, como é demonstrado no Apêndice C, uma técnica mais apropriada que estes dois métodos, denominada ARMarkov – LS e desenvolvida por Kamrunnahr et al. (2000), é preferível.

Apêndice C - Algoritmo ARMarkov-LS

O algoritmo ARMarkov-LS como proposto em Kamrunnahar et al. (2000) propõem-se a ser uma técnica mais apropriada que outros métodos tais como análise de correlação, ARX e FIR. Suas estimativas são computacionalmente mais simples que as obtidas com modelos ARX e os resultados são mais exatos que FIR ou análise de correlação. Excepcionalmente neste apêndice, a variável n representará ordens de polinômios de modelos autorregressivos, enquanto N representará número de pontos num conjunto de dados.

C.1 Introdução

O método ARMarkov-LS de estimação de parâmetros de Markov na sua forma SISO é conhecido de longa data (ver referências em Kamrunnahar et al., 2000). O procedimento utiliza um algoritmo de mínimos-quadrados (LS, *Least Square*) para estimar os parâmetros de Markov na representação ARMarkov, que tem a mesma forma que a de um modelo ARX, exceto que o modelo ARMarkov contém explicitamente o número de parâmetros de Markov que se deseja obter. Os modelos ARMarkov e sua determinação podem ser encarados como uma mescla dos métodos ARX e FIR.

Em Kamrunnahar et al. (2000) o método é estendido para o caso MIMO e os resultados da estimação são avaliados estatisticamente frente aos métodos ARX, FIR e análise de correlação para mostrar as vantagens de se empregá-lo na estimação dos parâmetros de Markov. Neste apêndice, por simplicidade, primeiro será apresentado o algoritmo na forma SISO, que será após estendido por analogia para a forma MIMO, conforme Kamrunnahar et al. (2000).

C.2 Determinação da ordem da matriz de interação

Considerando um sistema SISO $G(q)$ representado por uma função de transferência de um sistema linear, discreto e invariante com o tempo:

$$G(q) = \begin{bmatrix} A_x & B_x \\ C_x & D_x \end{bmatrix} \quad (\text{C.1})$$

onde $A_x \in \mathfrak{R}^{s \times s}$, $B_x \in \mathfrak{R}^{s \times 1}$, $C_x \in \mathfrak{R}^{1 \times s}$ e $D_x \in \mathfrak{R}^{1 \times 1}$ são os coeficientes da representação em espaço de estado do sistema e s é o número de estados do modelo SISO. Os parâmetros de Markov, h_j , são definidos como (ver referências em Kamrunnahar et al., 2000):

$$\begin{aligned} h_j &\equiv D_x && \text{para } j = -1 \\ h_j &\equiv C_x A_x^j B_x && \text{para } j \geq 0 \end{aligned} \quad (\text{C.2})$$

A representação de $G(q)$ correspondente ao espaço de estado é:

$$G(q) = C_x (qI - A_x)^{-1} B_x + D_x \quad (\text{C.3})$$

A representação de $G(q)$ em termos de parâmetros de Markov é:

$$G(q) = \sum_{j=-1}^{\infty} h_j q^{-(j+1)} \quad (\text{C.4})$$

A representação de $G(q)$ na forma de função de transferência ARX é:

$$G_{ARX}(q) = \frac{b_0 q^n + b_1 q^{n-1} + \dots + b_n}{q^n + a_1 q^{n-1} + \dots + a_n} \quad (\text{C.5})$$

com representação no domínio do tempo dada por:

$$y(k) = -a_1 y(k-1) - \dots - a_n y(k-n) + b_0 u(k) + \dots + b_n u(k-n) \quad (\text{C.6})$$

para $k \geq 0$

Aqui n é a ordem dos modelos. Os coeficientes do modelo ARX em (C.6) e os parâmetros de Markov em (C.4) satisfazem a seguinte relação:

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} h_{-1} & 0 & \cdots & 0 \\ h_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{n-1} & \cdots & h_0 & h_{-1} \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \quad (\text{C.7})$$

A representação de $G(q)$ na forma ARMarkov com \mathbf{m} parâmetros é dada por:

$$G_{ARMark}(q) = \frac{h_{-1}q^{m+n-1} + \cdots + h_{m-2}q^n + \mathbf{b}_{m,1}q^{n-1} + \cdots + \mathbf{b}_{m,n}}{q^{m+n-1} + \mathbf{a}_{m,1}q^{n-1} + \cdots + \mathbf{a}_n} \quad (\text{C.8})$$

A representação ARMarkov no domínio do tempo é dada por:

$$\begin{aligned} y(k) = & -\sum_{j=1}^n \mathbf{a}_{m,j} y(k - \mathbf{m} - j + 1) + \sum_{j=1}^m h_{j-1} u(k - j + 1) \\ & + \sum_{j=1}^n \mathbf{b}_{m,j} u(k - \mathbf{m} - j + 1) \end{aligned} \quad (\text{C.9})$$

e envolve apenas os primeiros \mathbf{m} parâmetros de Markov, h_{-1}, \dots, h_{m-2} . Os parâmetros $\mathbf{a}_{m,i}$ são funções dos coeficientes ARX e dos parâmetros de Markov e $\in \mathfrak{R}^{1 \times n}$. n é a ordem do modelo ARMarkov. A representação ARX no domínio do tempo é um caso especial da representação ARMarkov no domínio do tempo com $\mathbf{m} = 1$.

Utilizando a notação típica de algoritmos de identificação baseados em mínimos-quadrados, o vetor regressor do modelo ARMarkov, \mathbf{f}_m , para (C.9) é escrito como:

$$\mathbf{f}_m(k) = \begin{bmatrix} y(k - \mathbf{m}) \\ \vdots \\ y(k - \mathbf{m} - n + 1) \\ u(k) \\ \vdots \\ u(k - \mathbf{m} - n + 1) \end{bmatrix} \quad (\text{C.10})$$

A saída do processo pode ser expressa como:

$$y(k) = \mathbf{W}_m \mathbf{f}_m(k) \quad (\text{C.11})$$

onde a matriz de peso \mathbf{W}_m é dada por:

$$\mathbf{W}_m = \begin{bmatrix} -\mathbf{A}_m & h_{-1} & \cdots & h_{m-2} & \mathbf{B}_m \end{bmatrix} \quad (\text{C.12})$$

$$\mathbf{A}_m = [\mathbf{a}_{m,1} \quad \cdots \quad \mathbf{a}_{m,n}] \in \mathfrak{R}^{1 \times n} \quad (\text{C.13})$$

$$\mathbf{B}_m = [\mathbf{b}_{m,1} \quad \cdots \quad \mathbf{b}_{m,n}] \in \mathfrak{R}^{1 \times n} \quad (\text{C.14})$$

de forma que a dimensão resultante de \mathbf{W}_m é $1 \times 2n + m$

Sendo $\hat{\mathbf{W}}_m$ uma estimativa da matriz de peso e $\hat{y}(k)$ a saída estimada, então:

$$\hat{y}(k) = \hat{\mathbf{W}}_m \mathbf{f}(k) \quad (\text{C.15})$$

Definindo o erro de saída, $\mathbf{e}(k)$ como:

$$\mathbf{e}(k) = y(k) - \hat{y}(k) \quad (\text{C.16})$$

e a função objetivo, J , em termos do erro de saída como:

$$J = \frac{1}{N} \sum_{k=1}^N \frac{1}{2} \mathbf{e}^2(k) = \frac{1}{2N} \sum_{k=1}^N (y(k) - \hat{\mathbf{W}}_m \mathbf{f}(k))^2 \quad (\text{C.17})$$

onde N é o número total de pontos do conjunto de dados. Assim, \mathbf{W}_m é uma minimização estrita de J se e somente se:

$$\hat{\mathbf{W}}_m = \left[\frac{1}{N} \sum_{k=1}^N y(k) \mathbf{f}_m^T(k) \right] \left[\frac{1}{N} \sum_{k=1}^N \mathbf{f}_m(k) \mathbf{f}_m^T(k) \right]^{-1} \quad (\text{C.18})$$

O termo inverso de (C.18):

$$\left[\frac{1}{N} \sum_{k=1}^N \mathbf{f}_m(k) \mathbf{f}_m^T(k) \right]$$

contém a estimativa das covariâncias de $u(k)$ e $y(k)$ e precisa ser não-singular para que a inversa exista. Quando $m = 1$, (C.18) é idêntico ao algoritmo de identificação ARX baseado em mínimos quadrados.

C.3 Extensão para o caso MIMO

A metodologia desenvolvida em Kammrunnahar et al. (2000) permite que o algoritmo ARMarkov-LS seja reformulado para o caso MIMO diretamente. A relação de entrada/saída do sistema é expressa como:

$$\begin{aligned} \mathbf{Y}(k) = & \sum_{j=1}^n -\mathbf{a}_{m,j} \mathbf{Y}(k - \mathbf{m} - j + 1) + \sum_{j=1}^m \mathbf{H}_{j-1} \mathbf{U}(k - j + 1) \\ & + \sum_{j=1}^n \mathbf{b}_{m,j} \mathbf{U}(k - \mathbf{m} - j + 1) \end{aligned} \quad (\text{C.19})$$

onde:

$$\mathbf{Y}(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_p(k) \end{bmatrix}, \quad \mathbf{U}(k) = \begin{bmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_m(k) \end{bmatrix} \quad (\text{C.20})$$

e $\mathbf{a} \in \mathfrak{R}^{p \times p}$, $\mathbf{b} \in \mathfrak{R}^{p \times m}$ e $\mathbf{H}_j \in \mathfrak{R}^{p \times m}$. O número de saídas e entradas é p e m respectivamente.

O vetor regressor é expresso como:

$$\mathbf{F}(k) \equiv \begin{bmatrix} \mathbf{Y}(k - \mathbf{m})^T \\ \vdots \\ \mathbf{Y}(k - \mathbf{m} - n + 1)^T \\ \mathbf{U}(k)^T \\ \vdots \\ \mathbf{U}(k - \mathbf{m} - n + 1)^T \end{bmatrix} \quad (\text{C.21})$$

Seguindo os mesmo procedimento de mínimos quadrados do caso SISO, a função objetivo é:

$$J = \frac{1}{N} \sum_{k=1}^N \mathbf{e}^T \mathbf{e} \quad (\text{C.22})$$

onde \mathbf{e} é o vetor do erro da saída. A estimativa da matriz peso é:

$$\hat{\mathbf{W}}_m = \left[\frac{1}{N} \sum_{k=1}^N \mathbf{Y}(k) \mathbf{F}_m^T(k) \right] \left[\frac{1}{N} \sum_{k=1}^N \mathbf{F}_m(k) \mathbf{F}_m^T(k) \right]^{-1} \quad (\text{C.23})$$

onde:

$$\hat{\mathbf{W}}_m = \left[-\hat{\mathbf{A}}_m \quad \hat{\mathbf{H}}_{-1} \quad \cdots \quad \hat{\mathbf{H}}_{m-2} \quad \hat{\mathbf{B}}_m \right] \quad (\text{C.24})$$

$$\mathbf{A}_m = \left[\hat{\mathbf{a}}_{m,1} \quad \cdots \quad \hat{\mathbf{a}}_{m,n} \right] \in \mathfrak{R}^{p \times nm} \quad (\text{C.25})$$

$$\mathbf{B}_m = \begin{bmatrix} \hat{\mathbf{b}}_{m,1} & \cdots & \hat{\mathbf{b}}_{m,n} \end{bmatrix} \in \mathfrak{R}^{p \times nm} \quad (\text{C.26})$$

$$\hat{\mathbf{H}}_j \in \mathfrak{R}^{p \times m}, \quad \hat{\mathbf{H}} \in \mathfrak{R}^{p \times mm}$$

$\hat{\mathbf{W}}_m$ tem então dimensão final dada por: $p \times m(2n + m)$.

C.4 Propriedades estatísticas

Em Ljung (1999) é apresentada uma base teórica completa das técnicas de regressão linear bem como a formulação de suas propriedades estatísticas. Tendo em vista que o método ARMarkov é uma junção dos métodos FIR e ARX, é de se esperar que suas propriedades estatísticas estejam entre as propriedades destes dois últimos. Os modelos destes métodos podem ser escritos em termos dos vetores regressores e de parâmetros como:

$$\text{ARX:} \quad \hat{y}(k) = \mathbf{f}_{ARX}^T(k) \mathbf{W}_{m,ARX} + v_{ARX}(k) \quad (\text{C.27})$$

$$\text{FIR:} \quad \hat{y}(k) = \mathbf{f}_{FIR}^T(k) \mathbf{W}_{m,FIR} + v_{FIR}(k) \quad (\text{C.28})$$

$$\text{ARMarkov:} \quad \hat{y}(k) = \mathbf{f}_{ARM}^T(k) \mathbf{W}_{m,ARM} + v_{ARM}(k) \quad (\text{C.29})$$

Onde v é o ruído de processo.

C.4.1 Consistência do mínimos-quadrados

Assumindo que o sinal de entrada é persistentemente excitante e o modelo não contém *bias*, a estimativa em mínimos-quadrados é consistente (Ljung, 1999), ou seja, $\hat{\mathbf{W}}_m$ converge para \mathbf{W}_m conforme o número de pontos do conjunto de dados tende ao infinito ($N \rightarrow \infty$), se:

- v é ruído branco (o que não é o caso na maioria dos sistemas reais), ou
- a seqüência de entrada é independente da seqüência de ruído e não haja termos de saída no vetor regressor.

No modelo FIR não há termos de saída no vetor regressor, então os parâmetros estimados convergem para os reais conforme N tende ao infinito. No método ARX no entanto, a estimativa dos parâmetros não é consistente na presença de ruído não branco devido aos termos de saída no regressor. Os parâmetros de Markov estimados com o método ARMarkov são consistentes, embora os parâmetros equivalentes aos parâmetros ARX no modelo ARMarkov não o sejam (para prova, referir-se a Kamrunnahar et al., 2000).

C.4.2 Covariância das estimativas para sistemas com ruído branco

Os termos correspondentes ao erro da saída, ε , na predição do erro de saída (C.16), são estatisticamente independentes, tendo média zero e variância dada por I^2 . As estimativas $\hat{\mathbf{W}}_m$ são as estimativas sem desvio (*bias*) dos parâmetros reais, \mathbf{W}_m , ou seja, $E\{\hat{\mathbf{W}}_m\} = \mathbf{W}_m$. A matriz de covariância das estimativas é (Ljung, 1999):

$$\text{cov}(\hat{\mathbf{W}}_m) = I^2 \left[\frac{1}{N} \sum_{k=1}^N \mathbf{f}(k)\mathbf{f}(k)^T \right]^{-1} \quad (\text{C.30})$$

Para que se calcule a covariância das estimativas, é necessário o conhecimento de I^2 , o que nem sempre é disponível diretamente. Assim, é preciso uma estimativa da variância do erro de predição:

$$\hat{I}^2 = \frac{SSE}{N - 2n - m} \quad (\text{C.31})$$

onde SSE é o erro padrão quadrático (SSE , *squared standard error*), dado por:

$$SSE = \sum_{k=1}^N \mathbf{e}^T(k)\mathbf{e}(k) \quad (\text{C.32})$$

De acordo com Ljung (1999) a variância total dos parâmetros estimados é proporcional à razão entre o número de parâmetros do modelo e o número de pontos do conjunto de dados empregados na identificação. Desta forma:

$$\text{var}(\hat{G}_{ARX}) \propto \frac{2n}{N} \quad (\text{C.33})$$

$$\text{var}(\hat{G}_{ARMark}) \propto \frac{2n + m}{N} \quad (\text{C.34})$$

$$\text{var}(\hat{G}_{FIR}) \propto \frac{f}{N} \quad (\text{C.35})$$

onde f é o número de parâmetros do modelo FIR obtido por análise de correlação. Isto leva a concluir que, para $f \geq m + 2n$ e $m > 1$:

$$\text{var}(\hat{G}_{FIR}) > \text{var}(\hat{G}_{ARMark}) > \text{var}(\hat{G}_{ARX}) \quad (\text{C.36})$$

A variabilidade da estimativa dos parâmetros de Markov através do método ARMarkov é maior que através de ARX, mas o algoritmo ARMarkov é computacionalmente

mais eficiente que o ARX para o cálculo dos parâmetros de Markov, uma vez que já os provê diretamente como resultado do mínimos-quadrados, diferentemente do ARX que requer a divisão longa (ou deconvolução) dos parâmetros a e b .

C.4.3 Covariância das estimativas com ruído não-branco

Assumindo agora que o processo real é sujeito a ruído não-branco, v , com covariância dada por:

$$E\{vv^T\} = \mathbf{V} \quad (\text{C.37})$$

a matriz de covariância dos parâmetros estimados é:

$$\text{cov}(\hat{\mathbf{W}}_m) = \left[\frac{1}{N} \sum_{k=1}^N \mathbf{f}(k)\mathbf{f}^T(k) \right]^{-1} \mathbf{F} \mathbf{V} \mathbf{F}^T \left[\frac{1}{N} \sum_{k=1}^N \mathbf{f}(k)\mathbf{f}^T(k) \right]^{-1} \quad (\text{C.38})$$

Na prática, \mathbf{V} nem sempre é conhecida e é substituída por uma estimativa da matriz de covariância do erro de saída:

$$\hat{\mathbf{V}} = \frac{1}{N} \sum_{k=1}^N \mathbf{e}(k)\mathbf{e}(k)^T \quad (\text{C.39})$$

Devido as diferentes estruturas do vetor regressor nos métodos ARMarkov, ARX e FIR, estes não podem ser comparados diretamente. Pode-se tentar comparar a soma total das variâncias correspondentes aos m parâmetros de Markov determinados por cada método. Como no método ARX os parâmetros de Markov não são determinados diretamente, deve-se tentar uma abordagem indireta de medida da variância. Para tanto converte-se os parâmetros ARX para o formato ARMarkov e então as variâncias dos parâmetros do modelo podem ser mapeadas às variâncias dos parâmetros de Markov. Kamrunnahar et al. (2000) demonstra, através de simulações, que neste caso:

$$\sum_{i=1}^m \text{var}(f_i^{(FIR)}) > \sum_{i=1}^m \text{var}(h_i^{(ARMark)}) > \sum_{i=1}^m \left(\text{var}(a_i^{(ARKX)}) + \text{var}(b_i^{(ARKX)}) \right) \quad (\text{C.40})$$

As mesmas considerações do caso de sistema com ruído branco são válidas para este resultado.

Como se pode concluir destes resultados, o método ARX falha em termos de uma medida analítica da consistência dos parâmetros estimados e o método FIR tem a maior variância dos três algoritmos. Os parâmetros de Markov estimados utilizando o método ARMarkov têm ambas as propriedades em níveis aceitáveis, entre os produzidos por ARX e FIR.

Apêndice D - Figuras Complementares

Este apêndice apresenta um conjunto de figuras que por conveniência foram omitidas do Capítulo 5. Estas figuras ilustram o sinal de excitação das perturbações inseridas na carga durante o procedimento de estimação das matrizes de interação e os resultados obtidos nas análises de correlação.

D.1 Sinais de excitação

Os sinais de excitação w_1 e w_2 correspondem a perturbações inseridas nas malhas dos níveis Y1 e Y2 respectivamente.

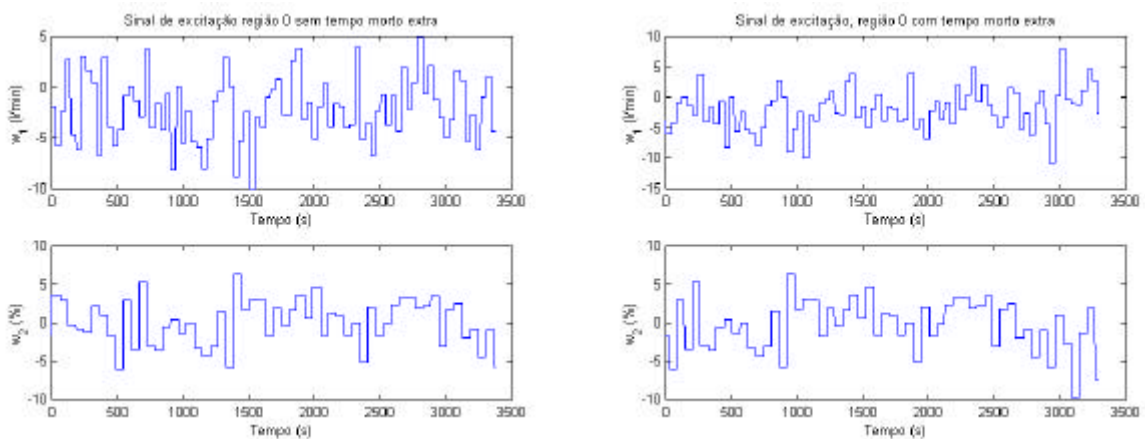


Figura 6.1. Perturbações inseridas no sistema operando na Região 0 nos Casos 1 (esquerda) e 2 (direita).

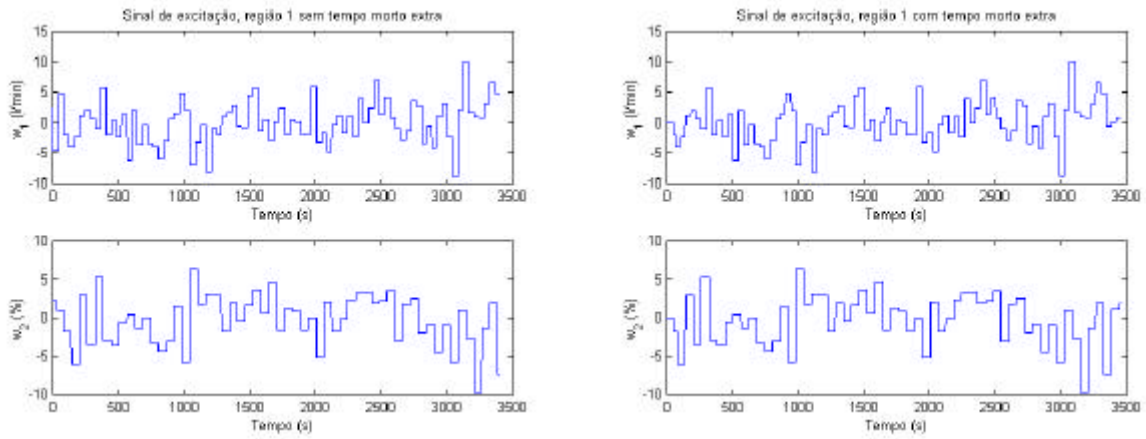


Figura 6.2. Perturbações inseridas no sistema operando na Região 1 nos Casos 1 (esquerda) e 2 (direita).

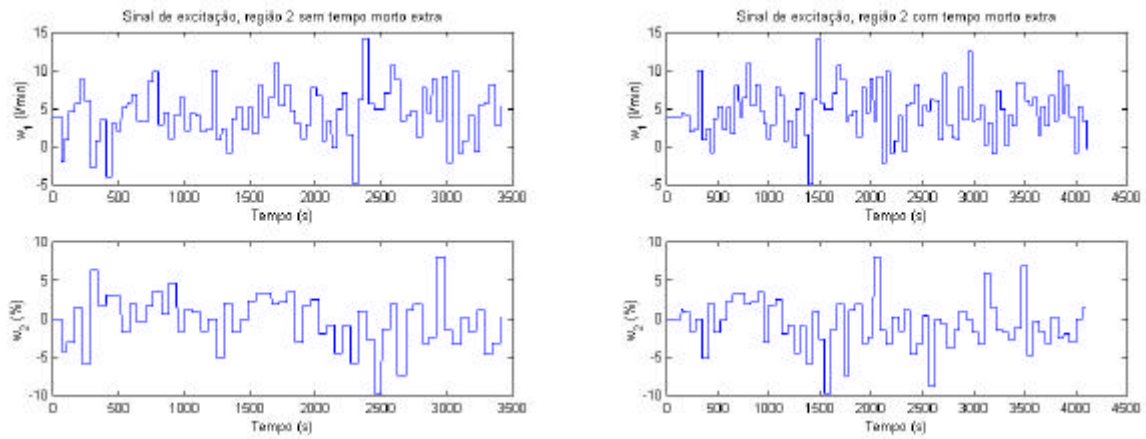


Figura 6.3. Perturbações inseridas no sistema operando na Região 2 nos Casos 1 e 2.

D.2 Análise de Correlação

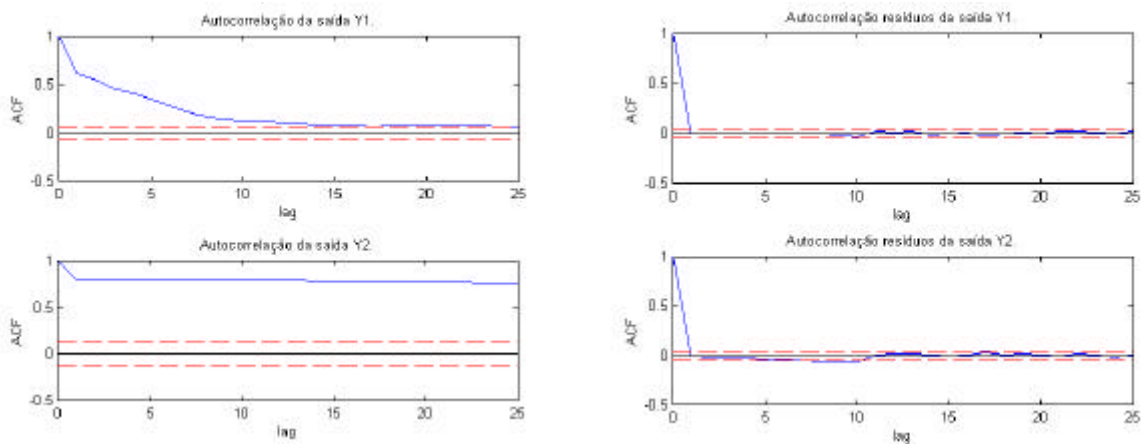


Figura 6.4. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 1, com controladores PI.

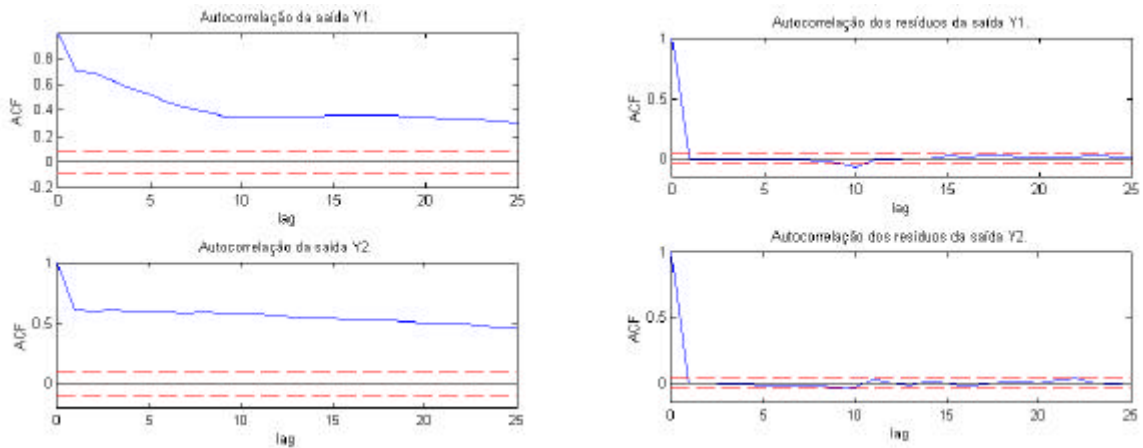


Figura 6.5. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 2, com controladores PI.

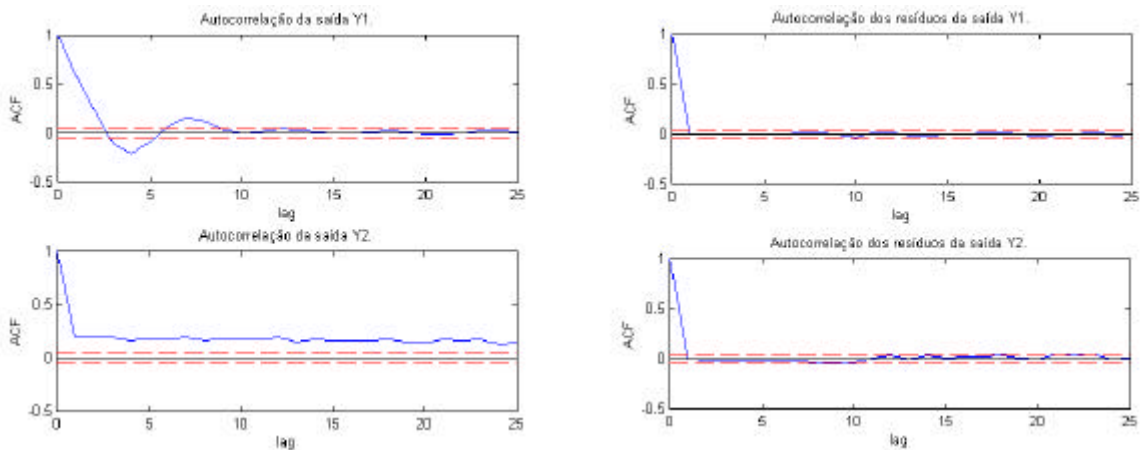


Figura 6.6. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 1, com controladores PI.

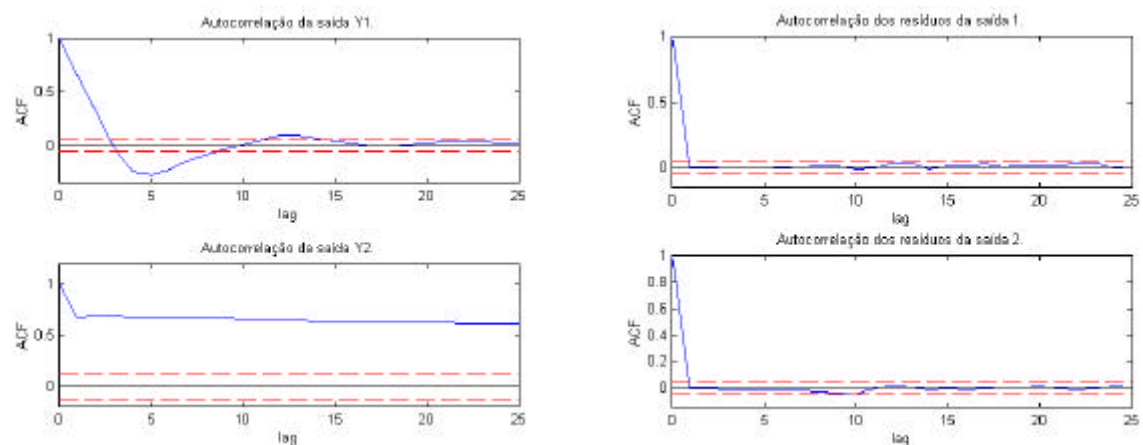


Figura 6.7. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 2, com controladores PI.

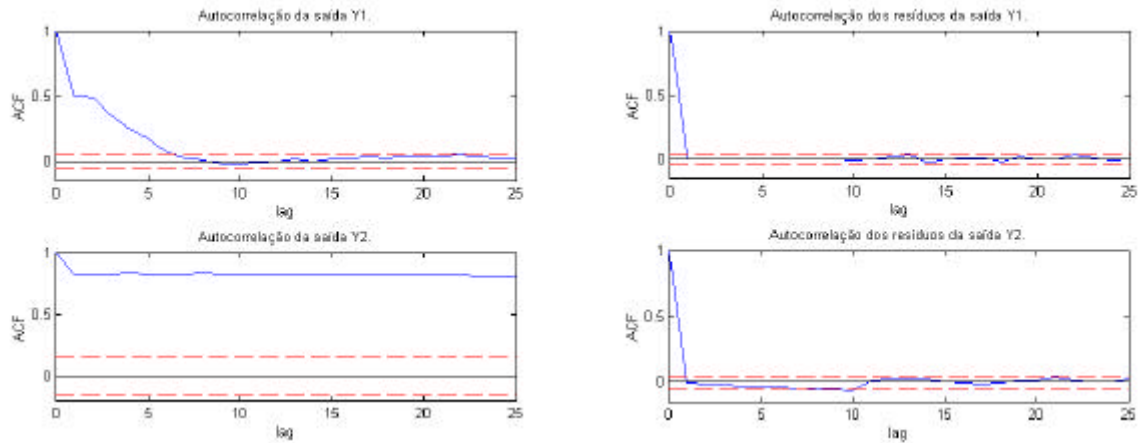


Figura 6.8. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 1, com controladores PI.

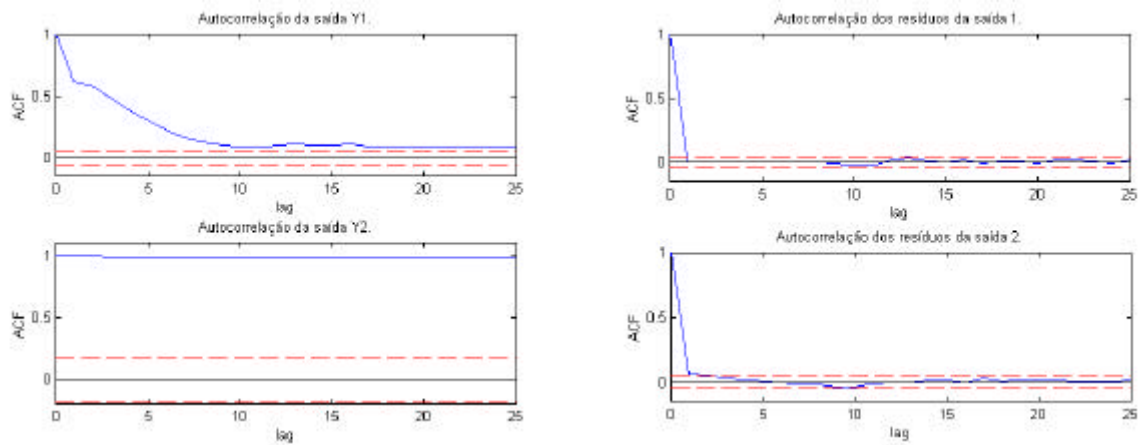


Figura 6.9. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 2, com controladores PI.

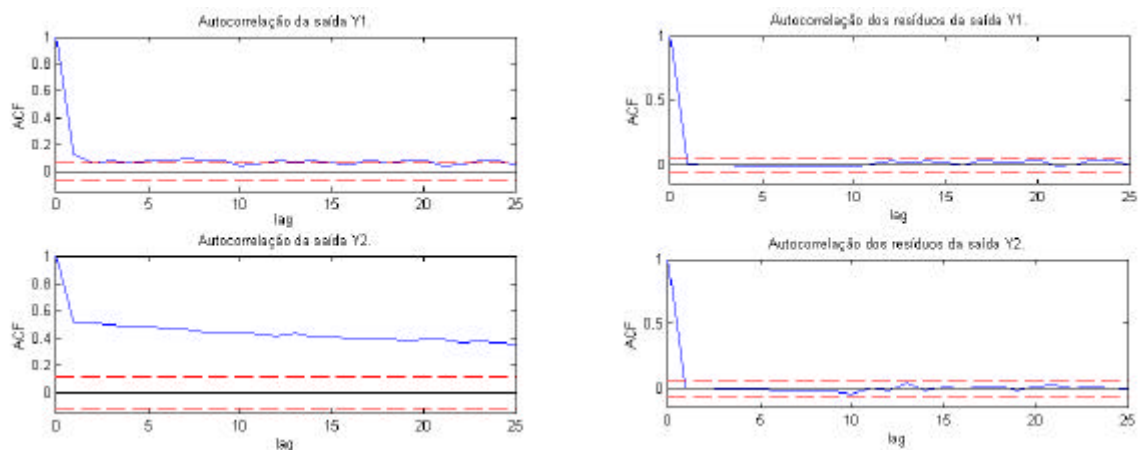


Figura 6.10. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 1, com NMPC.

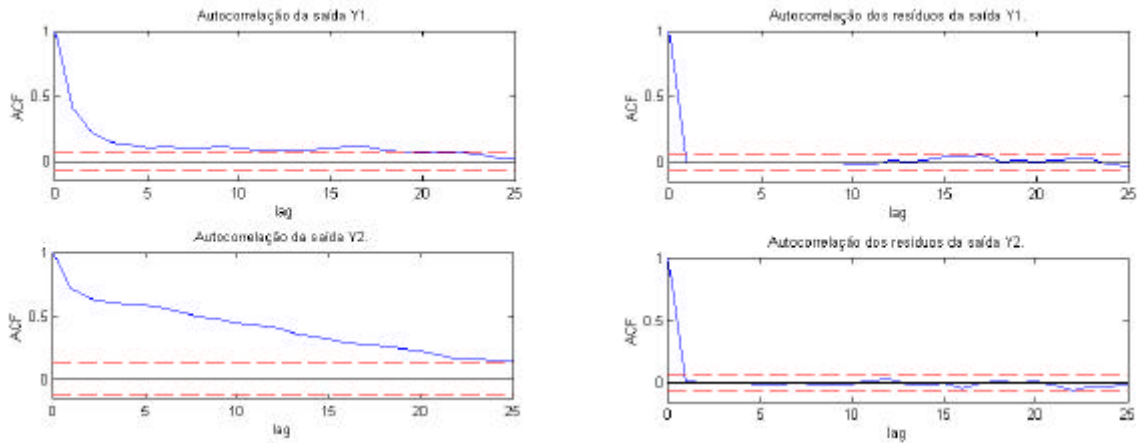


Figura 6.11. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 0, Caso 2, com NMPC.

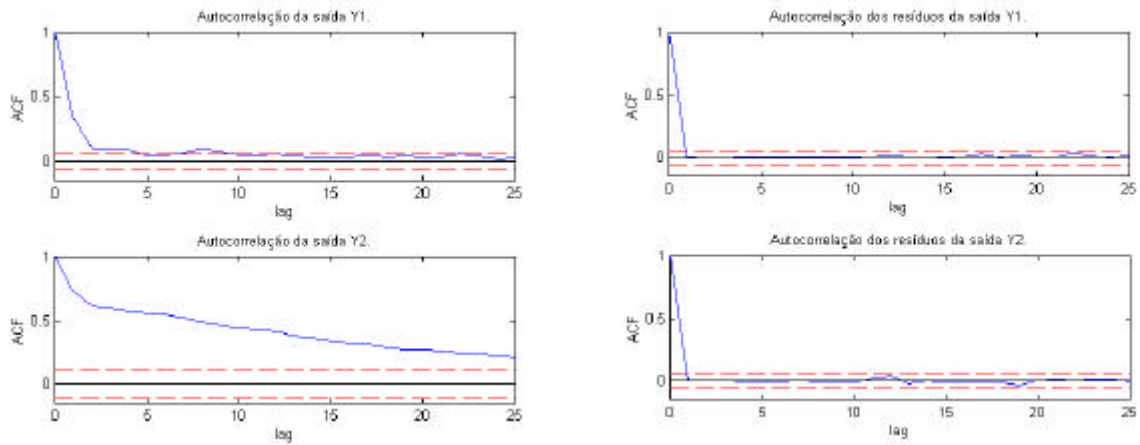


Figura 6.12. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 1, Caso 2, com NMPC.

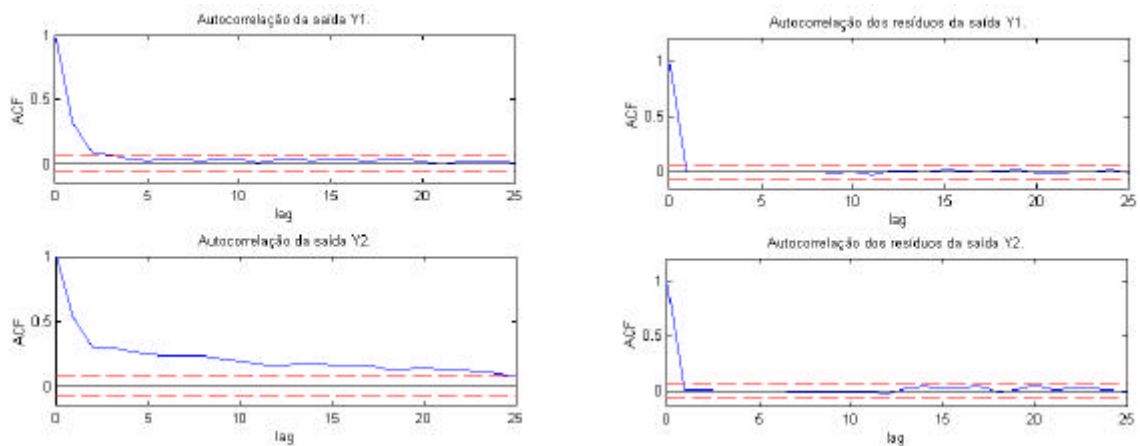


Figura 6.13. Função de autocorrelação da saída (esquerda) e dos resíduos (direita) para a Região 2, Caso 2, com NMPC.

Apêndice E - Funções Polinomiais

Para auxiliar na realização dos testes com controladores de variância mínima, realizados no Capítulo 3, foi desenvolvida uma classe polinomial simples em Matlab®. Esta classe consiste de um conjunto de funções que operam sobre um objeto polinomial denominado *polynom*. Além das operações básicas com polinômios SISO, diversas funções especializadas estão implementadas.

A listagem destas funções é dada a seguir, juntamente com uma descrição de sua utilização.

| | |
|---|---|
| <code>s = char(p)</code> | converte o polinômio <i>p</i> em uma <i>string</i> formatada <i>s</i> . |
| <code>q = diff(p)</code> | retorna o derivativo <i>q</i> do polinômio <i>p</i> . |
| <code>[F, G] = diophantine(A, C, d)</code> | resolve uma equação de Diophantine na forma da equação (2.17). O algoritmo é baseado em comparação de potências e pode ser deduzido a partir de exemplos em Åström e Wittenmark (1997) |
| <code>display(p)</code> | envia o polinômio <i>p</i> para a saída padrão do Matlab na forma de <i>string</i> . |
| <code>q = double(p)</code> | converte o polinômio <i>p</i> para um vetor <i>q</i> com seus coeficientes. |
| <code>q = dzero(p)</code> <code>q = dzero(p, tol)</code> | zera os coeficientes de <i>p</i> próximos a zero. Se <i>tol</i> é especificada, os coeficientes menores que <i>tol</i> são zerados. Se <i>tol</i> não é especificada, os coeficientes menores que $1e-8$ são zerados. |
| <code>q = fliplr(p)</code> | alterna os coeficientes do polinômios <i>p</i> . Para um polinômio <i>p</i> com coeficientes dados por $[1\ 2\ 3]$ é retornado um polinômio <i>q</i> com coeficientes $[3\ 2\ 1]$. |

```
G = gcd(A, B)
[G, X, Y] = gcd(A, B)
[G, X, Y, U, V] = gcd(A, B)
```

```
get(p)
propvalue = get(p, propnam)
n = length(p)
```

```
m = order(p)
q = ordreduce(p)
```

```
plot(p)
p = polynom(coefs)
p = polynom(coefs, var)
```

```
O = polysf(p)
[O, Q, R] = polysf(p)
```

```
y = polyval(p, x)
q = reciproc(p)
```

```
r = roots(p)
set(p, pname, pvalue)
set(p, pname1, pvalue1,
     pname2, pvalue2, ...)
T = tf(pnum, qden)
T = tf(pnum, qden, ts)
```

G é o maior divisor comum entre A e B . X e Y são tais que $G = AX + BY$. U e V são tais que $0 = AU + BV$. A função utiliza o Algoritmo de Euclides Estendido.

acessa os atributos do polinômio p . `propvalue` é o valor da propriedade `propnam`.

retorna o tamanho de um polinômio, ou seja, o número de coeficientes de p .

retorna a ordem do polinômio p .

retorna um polinômio q com a ordem reduzida, isto é, remove os coeficientes iguais a zero no final (polinômios em x, z, s ou q) ou no início (polinômios em z^{-1} ou q^{-1}).

faz o gráfico dos coeficientes do polinômio p .

construtor de um objeto `polynomial`. Dado um vetor de coeficientes, retorna um polinômio p . A variável `var` pode ser escolhida como x, s, z, q, q^{-1} ou z^{-1} . Se não for informada, o padrão é x .

Fatoração espectral de um polinômio. Fatora um polinômio p separando a fase mínima da fase não mínima: $p = S \times R$, onde S é o polinômio recíproco de Q . Q contém as raízes de fase mínima, R contém as raízes de fase não mínima, O é o polinômio fatorado resultante, recíproco de p , sem as raízes de fase não mínima, mas com as mesmas propriedades espectrais de p .

retorna o valor do polinômio p no ponto x .

retorna um polinômio q recíproco de p . O recíproco de p é definido como: $q(z) = z^n p(z^{-1})$, onde n é a ordem de p .

retorna as raízes de p .

altera os atributos de p de nome `pname` para os valores `pvalue`

converte dois polinômios representando o numerador e o denominador em uma função de transferência do Control System Toolbox. `ts` é o tempo de amostragem para sistemas discretos. A sintaxe é a mesma da função equivalente do toolbox de controle.

Alguns exemplos de utilização:

```
» p = polynom([1, 2, 3], 'v', 'z')
p = POLYNOM object
```

```
z^2 + 2*z + 3
```



```
» r = polynom([1 2 0], 'v', 'z')
r = POLYNOM object

      z^2 + 2*z

» p + r
ans = POLYNOM object

      2*z^2 + 4*z + 3

» [Q, R] = mrdivide( p, r ) % Divisão longa de p por r,
                          % produzindo o quociente Q e o resto R.
Q = POLYNOM object

      1

R = POLYNOM object

      3

» gcd(p, r)
ans = POLYNOM object

      3

[F, G] = diophantine(p, r, 2) % p = rF + Gz-2
F = POLYNOM object

      z

G = POLYNOM object

      -3*z

» polyval(r, -4)
ans =

      8

» get(p) % Acessa os atributos de p
      c: [1 2 3]
      var: 'z'
»
```


Referências Bibliográficas

- Åström, K. J.;** “Introduction to Stochastic Control Theory”, Academic Press; **1970.**
- Åström, K. J.; Wittenmark, B.;** “Computer Controlled Systems-Theory and Design”, Prentice Hall International; **1997.**
- Åström, K. J.; Hägglund, T.;** “PID Controllers: Theory, Design and Tuning”, 2nd Edition; Instrument Society of America, **1998.**
- Bezergianni, S.;** “Assessment of Multivariable Controller Performance”, Doctor of Philosophy’s Thesis, Lehigh University, Department of Chemical Engineering; **2001.**
- Bezergianni, S.; Georgakis, C.;** “Controller performance assessment based on minimum and open-loop output variance”, Control Engineering Practice, Vol. 8, pp. 791-797; **2000.**
- Desborough, L.; Harris, T.** “Performance Assessment Measures for Univariate Feedback Control”, The Canadian Journal Of Chemical Engineering, Vol. 70, pp. 1186-1197; December **1992.**
- Desborough, L.; Harris, T.;** “Performance Assessment measures for univariate feedforward/feedback control”. The Canadian Journal Of Chemical Engineering, Vol. 71, pp. 605-616; **1993.**
- Duraiski, R. G.;** “Controle preditivo não linear utilizando linearizações ao longo da trajetória”. MSc. Thesis, Universidade Federal do Rio Grande do Sul, Departamento de Engenharia Química; **2001.**
- Emerson News;** Informativo da Emerson Process Management – Brasil; Nº 12; **2002.**
- Ender, D.;** “Process control performance: not as good as you think”, Control Engineering, 180; **1993.**

- Farina, L. A.; Trierweiler, J. O.; Duraiski, R. G.;** “Rpn tuning strategy for model predictive control”, Preprints of DYCOPS-2001, pp. 283-288; **2001.**
- Ferreira, A. R.; Fernandes, J. M.;** “A Survey on Time Delay System Estimation”, Proceedings to the European Control Conference; **1997.**
- Grimble, M. J.;** “Controller performance benchmarking and tuning using generalised minimum variance control”, *Automatica*, Vol. 38, pp. 2111-2119; **2001.**
- Hägglund, T.;** “A Control Loop Performance Monitor”, *Control Engineering Practice*, Vol. 3, pp. 1543-1551; **1995.**
- Hägglund, T.;** “Automatic detection of sluggish control loops”, *Control Engineering Practice*, Vol. 7, pp. 1505-1511; **1999.**
- Hägglund, T.;** “Industrial application of automatic monitoring tools”, 15th IFAC, **2002.**
- Harris, T. J.;** “Assessment of Control Loop Performance”, *The Canadian Journal Of Chemical Engineering*, Vol. 67, 856-861; October, **1989.**
- Harris, T. J.; Boudreau, F.; MacGregor, J.;** “Performance assessment of multivariable feedback controllers”, *Automatica*, Vol 32, No. 11; pp. 1505-1518; **1996a.**
- Harris, T. J.; Seppala, C. T.; Jofriet, P. J.; Surgenor, B. W.;** “Plant-Wide Feedback Control Performance Assessment Using An Expert-System Framework”, *Control Engineering Practice*, Vol. 4, No. 9, pp. 1297-1303; **1996b.**
- Harris, T. J.; Seppale, C. T.; Desborough, L. D.;** “A review of performance monitoring and assessment techniques for univariate and multivariate control systems”, *Journal of Process Control*, Vol 9, pp. 1-17; **1999.**
- Horch, A.; Isaksson, A. J.;** “A modified index for control performance assessment”, *Journal of Process Control*, Vol. 9, pp. 475-483; **1999.**
- Huang, B.; Shah, S. L.;** “Performance Assessment of Control Loops – Theory and Applications”, Springer-Verlag; **1999.**
- Huang, B.; Shah, S. L.;** “Performance limits: practical control loop performance assessment”, Proceedings of AIChE Annual Meeting, **1996.**
- Huang, B.; Shah, S. L.;** “Practical issues in multivariable feedback control performance assessment”, *Journal of Process Control*, Vol. 8, Nos. 5-6, pp. 421-430; **1998.**
- Huang, B.; Shah, S. L.; Fujii, H.;** “The unitary matrix and its estimation using closed loop data”, *Journal of Process Control*, Vol. 7, No. 3; pp. 195-207; **1997a.**

- Huang, B.; Shah, S. L.; Kwok, E. K.;** “Good, Bad or Optimal? Performance Assessment of Multivariable Processes”, *Automatica*, Vol. 33, No. 6, pp. 1175-1183; **1997b**.
- Huang, B.; Shah, S. L.; Kwok, K.;** “How good is your controller? Application of control loop performance assessment techniques to MIMO processes”, *Proc. 13th IFAC World Congress*, Vol. M, pp. 229-239, **1996**.
- Huang, B.; Shah, S. L.; Miller, R.** “Feedforward Plus Feedback Controller Performance Assessment of MIMO Systems”, *IEEE Transactions on Control Systems Technology*, Vol. 8, No. 3, pp. 580-587; **2000**.
- Hugo, A. J.;** “Process Controller Performance Monitoring and Assessment”, Control Arts Inc. (www.controlartsinc.com);
- Kamrunnahr, M.; Huang, B.; Fisher, D. G.;** “Estimation of Markov parameters and time-delay/interactor matrix”, *Chemical Engineering Science*, Vol. 55, pp. 3353-3363; **2000**.
- Ko, B. S.; Edgar, T. F.;** “Performance Assessment of Cascade Control Loops”, *AIChE Journal*, Vol. 46, No. 2, pp. 281-291; **2000**.
- Ljung, L.;** “System identification – Theory for the User”, 2nd Ed. Prentice Hall, **1999**.
- Lynch, C. B.; Dumont, G. A.;** “Control loop performance monitoring”, *IEEE Transactions on Control Systems Technology*, 4(2), pp. 185-192; **1996**.
- Machado, V. C.; Trierweiler, J. O.; Secchi, A. R.;** “Utilização de plantas-piloto acessadas via Internet no ensino de engenharia química”, *Anais do IV COBEQ-IC*, **2001**.
- Ogunnaike, B. A.; Ray, W. H.;** “Process Dynamics, Modeling and Control”, Oxford University Press, New York, Oxford; **1994**.
- Overschee, P. Van; de Moor, B.** “Subspace Identification for Linear Systems, Theory - Implementation – Applications”, Kluwer Academic Publishers; **1996**.
- Paulonis, M. A.; Cox, J. W.;** “A practical approach for large-scale controller performance assessment, diagnosis, and improvement”, *Journal of Process Control*, Vol. 13, pp. 155-168; **2003**.
- Peng, Y.; Kinnaert, M.;** “Explicit solution to the Singular LQ Regulation Problem”, *IEEE Transactions on Automatic Control*, Vol. 37, No. 5; **1992**.
- Qin, S. J.;** “Control Performance Monitoring – a review and assessment”, *Computers and Chemical Engineering*, 23, 173-186; **1998**.

- Rengaswamy, R.; Hägglund, T.; Venkatasubramanian, V.;** “A qualitative shape analysis formalism for monitoring control loop performance”, *Engineering Applications of Artificial Intelligence*, Vol. 14, pp. 23-33; **2001**.
- Rogozinski, M. W.; Paplinski, A. P.; Gibbard, M. J.;** “An Algorithm for the Calculation of a Nilpotent Interactor Matrix for Linear Multivariable Systems”, *IEEE Transactions on Automatic Control*, Vol. AC-32, No. 3, pp. 234-237; **1987**.
- Tecnologia e Negócios – Automação e Controle. Suplemento Especial: Automação.** Revista Brasileira de Tecnologia e Negócios de Automação, Instrumentação, Sistemas de Controle, Monitoramento, Informática e Robótica. N. 18.
- Thornhill N. F.; Oettinger, M.; Fedenkzuk, P.;** “Refinery-wide control loop performance assessment”, *Journal of Process Control*, Vol. 9, pp. 109-124; **1999**.
- Thornhill, N. F.; Hägglund, T.;** “Detection and diagnosis of oscillation in control loops”, *Control Engineering Practice*, Vol. 5, No. 10, pp. 1343-1347; **1997**.
- Thyagarajan, T.; Yu, C. C.; Huang, H. P.;** “Assessment of controller performance: a relay feedback approach”, *Chemical Engineering Science*, Vol. 58, pp. 497-512; **2003**.
- Tuch, J.; Feuer, A.; Palmor, Z. J.;** “Time Delay Estimation in Continuous Linear Time-Invariant Systems” *IEEE Transactions on Automatic Control*, Vol. 39, No. 4, pp. 823-827; **1994**.
- Tyler, M. L.; Morari, M.** “Performance monitoring of control systems using likelihood ratio methods”, *Proceedings of American Control Conference*, pp 1245-1249, **1995**.
- Tyler, M. L.; Morari, M.;** “Performance Monitoring of Control Systems Using Likelihood Methods”, *Automatica*, Vol. 8, pp. 1145-1162; **1996**.
- Wan, S.; Huang, B.;** “Robust performance assessment of feedback control systems”, *Automatica*, Vol. 38, pp. 33-46; **2002**.

Referências na Internet:

- BOURKE;** <http://astronomy.swin.edu.au/~pbourke/analysis/correlate>: Referência em métodos de análise de correlação.
- CONTROLARTS;** <http://www.controlartsinc.com>: Desenvolvedor de software para avaliação de malhas de controle.
- EXPERTUNE;** www.expertune.com/usingpower.doc: Referência em análise de espectro de potências.

-
- GEOS;** <http://www.ltrr.arizona.edu/~dmeko/geos595e.html>: Curso de Séries temporais com aulas em arquivos PDF. Excelente referência em análise de séries temporais e análises estatísticas.
- HARSHAM;** <http://ubmail.ubalt.edu/~harsham/stat-data/opre330Forecast.htm>: Referência em análise de séries temporais.
- HONEYWELL;** http://www.tata-honeywell.com/html/am_ls_ls.htm: Página de uma ferramenta para avaliação de malhas de controle desenvolvida pela Honeywell.
- MATHWORLD;** <http://mathworld.wolfram.com>. Definições matemáticas e algoritmos para cálculo das equações de Diophantine e maior divisor comum (GCD).
- MATRIKON;** <http://www.matrikon.com>: Desenvolvedor do ProcessDoctor, uma ferramenta para avaliação, auditoria e monitoramento de malhas de controle.
- PLANETMATH;** <http://planetmath.org>: Endereço alternativo para definições matemáticas e algoritmos para cálculo das equações de Diophantine e maior divisor comum (GCD).
- STATSOFT;** <http://www.statsoft.com/textbook/sttimser.html>: Referência em análise de séries temporais e análise espectral.