

ENXAME DE PARTÍCULAS COMO FERRAMENTA DE OTIMIZAÇÃO EM
PROBLEMAS COMPLEXOS DE ENGENHARIA NUCLEAR

Jose Antonio Carlos Canedo Medeiros

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM
ENGENHARIA NUCLEAR.

Aprovada por:

Prof. Roberto Schirru, D. Sc.

Prof. Aquilino Senra Martinez, D. Sc.

Prof. Cláudio Márcio do Nascimento Abreu Pereira, D. Sc.

Dr. Antônio César Ferreira Guimarães, D. Sc.

Dr. Sérgio Queiróz Bogado Leite, D. Sc.

RIO DE JANEIRO, RJ - BRASIL

JUNHO DE 2005

MEDEIROS, JOSE ANTONIO CARLOS
CANEDO

Enxame de Partículas como Ferramenta de
Otimização em Problemas Complexos de
Engenharia Nuclear [Rio de Janeiro] 2005

XI, 108 p. 29,7 cm (COPPE/UFRJ, D.Sc.,
Engenharia Nuclear, 2005)

Tese - Universidade Federal do Rio de
Janeiro, COPPE

- 1 – Otimização
- 2 – Enxame de Partículas
- 3 – Classificação de Acidentes
- 4 – Projeto de Reator
- 5 – Controle Preditivo

I. COPPE/UFRJ II. Título (série)

A MEUS PAIS

MÁRIO CANEDO LOPEZ E

TERESA MEDEIROS DE CANEDO

A MEUS FILHOS

LUIZA CASTELLAN CANEDO MEDEIROS E

FABIO CASTELLAN CANEDO MEDEIROS

AGRADECIMENTOS

Ao amigo, professor e orientador, Dr. Roberto Schirru, não só pela sua orientação e o senso de direção, sem o que a realização deste trabalho não teria sido possível, mas também pela sua ajuda e incentivo no plano profissional.

Ao professor e amigo Cláudio M N A Pereira, pela inestimável colaboração em diversas fases do trabalho, com discussões e idéias sobre o tema da tese.

Ao amigo Wagner Sacco, com quem por diversas vezes conversamos sobre o assunto otimização, em parte responsável pelo aumento de meu interesse e direcionamento de meus esforços para essa área.

Ao amigo José Airton, principalmente pela sua amizade, e pelo apoio na manutenção dos sistemas computacionais essenciais para a elaboração da tese.

Aos amigos Wagner Esteves e Roberto Domingues pelas discussões e sua colaboração com idéias a respeito de temas desta tese.

Aos amigos e colegas do LMP e da COPPE e a todos, que, de alguma forma contribuíram para a realização deste trabalho.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D Sc.)

ENXAME DE PARTÍCULAS COMO FERRAMENTA DE OTIMIZAÇÃO EM
PROBLEMAS COMPLEXOS DE ENGENHARIA NUCLEAR

Jose Antonio Carlos Canedo Medeiros

Junho/2005

Orientador: Roberto Schirru

Programa: Engenharia Nuclear

Devido ao seu baixo custo computacional, técnicas de busca baseadas no gradiente associadas a técnicas de programação linear são utilizadas como ferramentas de otimização. Essas técnicas, no entanto, quando aplicadas a espaços de busca multimodais complexos não são adequadas podendo em muitos casos conduzir a, e estacionar em, ótimos locais. Na busca de soluções em domínios multimodais, técnicas de busca randômica tem sido desenvolvidas e usadas com grande eficácia. Neste trabalho exploramos a capacidade do algoritmo de otimização por enxame de partículas como ferramenta de otimização para solução de problemas complexos na área nuclear. Devido à facilidade e natural representação de espaço reais de qualquer dimensão, foi aplicado com sucesso a problemas na área nuclear e mostrou-se eficaz na busca de soluções em espaços de busca multimodais de alta dimensão. Em uma das aplicações foi possível elaborar um modelo de solução simples e obter uma solução trivial não obtida por outros métodos confirmando assim a validade de sua aplicação.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

PARTICLE SWARM AS OPTIMIZATION TOOL IN COMPLEX NUCLEAR
ENGINEERING PROBLEMS

Jose Antonio Carlos Canedo Medeiros

June/2005

Advisor: Roberto Schirru

Department: Nuclear Engineering

Due to its low computational cost, gradient-based search techniques associated to linear programming techniques are being used as optimization tools. These techniques, however, when applied to multimodal search spaces, can lead to local optima. When finding solutions for complex multimodal domains, random search techniques are being used with great efficacy. In this work we exploit the swarm optimization algorithm search power capacity as an optimization tool for the solution of complex high dimension and multimodal search spaces of nuclear problems. Due to its easy and natural representation of high dimension domains, the particle swarm optimization was applied with success for the solution of complex nuclear problems showing its efficacy in the search of solutions in high dimension and complex multimodal spaces. In one of these applications it enabled a natural and trivial solution in a way not obtained with other methods confirming the validity of its application.

ÍNDICE

1. INTRODUÇÃO	1
1.1. Objetivo	1
1.2. Motivação	1
1.3. Metodologia	4
1.3.1. Projeto do Núcleo de um Reator	5
1.3.2. Controle Preditivo da Oscilação Axial de Xenônio	6
1.3.3. Classificação de Acidentes em Usinas Nucleares	7
1.4. Organização do Trabalho	10
2. OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS	13
2.1. Introdução	13
2.2. Algoritmo Básico	14
2.3. Modelo Determinístico	16
2.4. Dinâmica da Partícula	19
2.4.1. Convergência	20
2.4.2. Oscilações Harmônicas	22
2.4.3. Movimento Zig-Zag	22
2.5. Seleção de Parâmetros	23
3. OTIMIZAÇÃO DO PROJETO DO NÚCLEO DE UM REATOR	28
3.1. Introdução	28
3.2. Projeto do Núcleo de um Reator Nuclear	29
3.3. Modelagem do Problema com o PSO	32
3.4. Configuração do PSO	33
3.5. Resultados	35
4. CONTROLE PREDITIVO DA OSCILAÇÃO DE XENÔNIO	38
4.1. Introdução	38
4.2. Controle Preditivo	40
4.2.1. Preditor com Redes Neurais	41
4.2.2. Otimizador	43
4.2.3. Algoritmo do Controlador Preditivo	44
4.3. Configuração e Treinamento do Preditor	45
4.4. Otimizador com Enxame de Partículas	49
4.5. Configuração do PSO	49
4.6. Resultados	50

5. CLASSIFICAÇÃO DE ACIDENTES EM USINAS NUCLEARES	53
5.1. Introdução	53
5.2. Diagnóstico de Transientes em Usinas Nucleares	59
5.3. Assinaturas de Acidentes	60
5.4. Classificação de Eventos	65
5.5. Método do Número Mínimo de Protótipos	67
5.6. Configuração do PSO	68
5.7. Resultados	69
6. CONCLUSÕES E RECOMENDAÇÕES	75
6.1. Conclusões	75
6.2. Sugestões para Trabalhos Futuros	78

ANEXOS

ANEXO 1– ALGORITMOS DE OTIMIZAÇÃO BASEADOS EM GRADIENTE	80
ANEXO 2– OSCILAÇÃO DE POTÊNCIA AXIAL	88
ANEXO 3– REDES NEURAIS ARTIFICIAIS	95
REFERÊNCIAS BIBLIOGRÁFICAS	102

LISTA DE FIGURAS

Figura 2.1. Enxame de partículas	14
Figura 2.2. Atualização de uma partícula	15
Figura 2.3. Convergência do PSO	16
Figura 2.4. Regiões de convergência	21
Figura 2.5. PSO convergente não oscilatório	24
Figura 2.6. PSO, oscilação harmônica, convergência lenta	24
Figura 2.7. PSO, oscilação harmônica, convergência rápida	25
Figura 2.8. PSO, oscilação harmônica, com zig-zag	25
Figura 2.9. PSO, zig-zag convergente assimétrico	26
Figura 2.10. PSO, zig-zag convergente simétrico	26
Figura 3.1. Reator e célula característica	29
Figura 3.2. Convergência com 200 iterações de 100 partículas	35
Figura 3.3. Convergência com 100 iterações de 200 partículas	35
Figura 3.4. Convergência com 50 iterações de 400 partículas	36
Figura 4.1. Controle preditivo	41
Figura 4.2. Estrutura de um controlador preditivo	41
Figura 4.3. Aprendizado supervisionado	42
Figura 4.4. Predição com rede feed-forward e linhas de retardo	43
Figura 4.5. Dados simulados para identificação da planta	46
Figura 4.6. Dados de treinamento	47
Figura 4.7. Dados de teste	47
Figura 4.8. Dados de validação	48
Figura 4.9. Performance de treinamento (MSE)	48

Figura 4.10. Perturbação da planta	50
Figura 4.11. Predição neural e otimização com PSO	50
Figura 4.12. Ação do controlador preditivo	51
Figura 4.13 Saída da planta sob ação do controlador	51
Figura 5.1. Assinaturas de acidentes a 50%	63
Figura 5.2. Assinaturas de acidentes a 100%	64
Figura 5.3. Convergência para NPC = 3	70
Figura 5.4. Convergência para NPC = 2	70
Figura 5.5. Convergência para NPC = 1	71
Figura 5.5. Protótipos para NPC = 3	72
Figura 5.6. Protótipos para NPC = 2	72
Figura 5.7. Protótipos para NPC = 1	73
Figura A2.1 – Modelo de reator a 2 pontos	88
Figura .A2.2 – Oscilação de potência do reator	93
Figura .A2.3 – Diferenças axiais	94
Figura A3.1 – Neurônio básico	95
Figura A3.2 – Camada de neurônios	97
Figura A3.3. Treinamento supervisionado	99
Figura A3.4 – Filtro linear com rede neural	101

TABELAS

Tabela 3.1. Limites operacionais do reator	30
Tabela 3.2. Limites dos parâmetros de projeto	30
Tabela 3.3 Parâmetros de configuração do PSO	34
Tabela 3.4 Resultados do algoritmo PSO	36
Tabela 3.5 Comparação de resultados	37
Tabela 4.1. Configuração do preditor	45
Tabela 4.2. Parâmetros do otimizador com PSO	49
Tabela 5.1. Acidentes postulados para um reator PWR	61
Tabela 5.2. Lista de variáveis de acidentes	61
Tabela 5.3. Conjunto de transientes para treinamento	65
Tabela 5.4. Método MCP com 6 partições	66
Tabela 5.5. Método MCP com 1 partição	67
Tabela 5.6. Parâmetros do algoritmo PSO	68
Tabela 5.7. Acertos para diferentes números de protótipos	69
Tabela 5.8. Particionamento para o MCP com 2 partições	72
Tabela A2.1 - Parâmetros do reator	91

CAPÍTULO 1

INTRODUÇÃO

1.1. Objetivo

Este trabalho explora a capacidade do algoritmo de otimização por enxame de partículas, traduzido do inglês Particle Swarm Optimization (PSO), (KENNEDY, EBERHART, SHI, 2001), como ferramenta de otimização, mostrando a viabilidade do algoritmo como ferramenta de otimização na busca de soluções em espaços de busca multimodais complexos e de alta dimensão, como os problemas encontrados na área de engenharia nuclear.

O algoritmo de otimização por enxame de partículas foi aplicado na solução de três problemas na área nuclear e mostrou-se eficaz na busca de soluções sem necessidade de conhecimento prévio sobre a complexidade dos espaços de busca envolvidos.

1.2. Motivação

A solução de inúmeros problemas da engenharia envolve, invariavelmente, um problema de otimização. Devido ao seu baixo custo computacional, as técnicas de busca tradicionais e as técnicas baseadas no gradiente, ver ANEXO 1, tem sido utilizadas como ferramentas de otimização. Em muitos casos, no entanto, não é possível calcular facilmente o gradiente o que dificulta a implementação do otimizador.

Essas técnicas, no entanto, quando aplicadas a domínios de aplicação multimodais complexos, não são adequadas, podendo estacionar em ótimos locais e conduzir a soluções sub ótimas. Quando a dimensão do espaço de busca for grande, e a função objetivo e as restrições forem não lineares ou descontínuas, a obtenção de soluções por métodos analíticos torna-se muito difícil ou até mesmo impossível, especialmente nos casos em que não se dispõe explicitamente a função analítica do problema.

No caso de funções multidimensionais, a determinação do ponto que otimiza uma função objetivo pode ser obtida por um mecanismo de busca gerando-se sistematicamente uma seqüência de pontos do espaço de busca e testando a cada passo se o ponto gerado é uma solução ou se conduz para a solução. Dentre os métodos de busca destacam-se os métodos de busca por enumeração, os métodos de busca baseados em gradiente e no hessiano e os métodos de busca probabilísticos.

Os métodos de otimização baseados na busca sistemática exigem conhecimento prévio sobre o espaço de busca e os métodos de otimização baseados no gradiente dependem da existência de limites e das derivadas.

No entanto, devido à grande capacidade computacional disponível nos sistemas atuais, cada vez mais as técnicas de otimização tem sido substituídas por técnicas desenvolvidas na área de inteligência artificial como otimização por sistemas neuro-fuzzy (PASSINO, K.M., YURKOVICK, S., 1998) e algoritmos de otimização evolucionários como algoritmos genéticos (GOLDBERG, 1989) e programação genética (KOZA, J.R., 1992). Por essa razão os métodos de busca probabilística tem sido desenvolvidos na tentativa de promover soluções que não dependam de conhecimento prévio sobre o espaço de busca.

Nos métodos de busca randômicos a regra de busca é dependente de números randômicos mas não é aleatória . Dentre esses métodos destacamos os algoritmos genéticos e os algoritmos de otimização baseados em enxame de partículas.

Os algoritmos genéticos, (HOLLAND, 1975 e GOLDBERG, 1989), são algoritmos de busca e otimização baseados nos mecanismos teoria da evolução das espécies de Darwin e da genética. Em um algoritmo genético os indivíduos de uma população, representados por um string de dígitos binários denominado cromossomo, representam soluções candidatas a solução do problema. Inicialmente é gerada, de forma totalmente aleatória, uma população inicial de indivíduos que representam soluções candidatas a solução do problema. A cada iteração do algoritmo genético todos os indivíduos da população são avaliados segundo uma função objetivo, que determina sua adequação, e recebem uma pontuação que será utilizada na determinação da geração da população para a próxima iteração. Os indivíduos melhor adaptados e os seus descendentes carregam consigo informação genética para as próximas gerações, e tem mais chance de sobreviver através de gerações sucessivas. Ao final de um grande número de iterações a população do algoritmo genético tende a conter apenas os indivíduos melhores adaptados. Os algoritmos genéticos tem sido utilizados com eficácia como métodos de busca em processos de otimização e não exigem conhecimento do espaço de busca nem sobre a existência de limites derivadas da função a ser otimizada.

Os algoritmos de otimização baseados em enxame de partículas são algoritmos baseados no paradigma da interação entre as partículas de um enxame (ou indivíduos de uma população) que procuram um ótimo global (KENNEDY e EBERHART, 1995).

No algoritmo baseado em enxame partículas uma população de indivíduos, ou partículas, representa soluções candidatas como pontos do espaço de solução do problema que são avaliadas a cada iteração do algoritmo de acordo com uma função objetivo. Cada indivíduo, representado como um vetor de números de dimensão igual à dimensão do espaço de busca, é inicializado com uma posição e velocidade aleatórias, e guarda consigo informação sobre a melhor posição por ele já visitada. Por outro lado o algoritmo guarda a melhor posição já alcançada pelo enxame. A cada iteração os indivíduos da população são avaliados e as melhores posições de cada indivíduo e do enxame são atualizadas. As posições de cada indivíduo são atualizadas levando em conta a influência da melhor posição já atingida pelo enxame (influência social) e a melhor posição já atingida pelo próprio indivíduo (decisão individual). Assim os indivíduos do enxame se movem pelo espaço de busca procurando dessa forma atingir o máximo ou mínimo global com resultado da decisão individual e da influência social.

Da mesma forma que os algoritmos genéticos, os algoritmos baseados em enxame de partículas não exigem conhecimento prévio sobre o espaço de busca nem sobre a existência de limites derivadas da função objetivo e ou as funções de restrição, bastando para sua aplicação a existência de uma função de avaliação dos indivíduos da população.

1.3. Metodologia

A fim de comprovar as características do algoritmo por enxame de partículas, as propriedades de convergência do algoritmo são estudadas.

A viabilidade e eficácia de sua utilização como ferramenta de otimização são comprovadas por meio da aplicação do algoritmo na solução de três problemas na área nuclear como otimizador sem a necessidade de conhecimento prévio sobre a complexidade dos espaços de busca envolvidos. A primeira aplicação consiste em utilizar o algoritmo de otimização por enxame de partículas na otimização do projeto do núcleo um reator nuclear em que se procura obter a melhor combinação de dimensões e constituição dos elementos do núcleo, e ao mesmo tempo obedecer a um conjunto de restrições relativas a limites operacionais de variáveis bem como dos limites permissíveis para as dimensões e composição material dos elementos do núcleo. A fim de corroborar sua robustez na otimização de funções, a segunda aplicação envolve a substituição do otimizador de um controlador preditivo, aplicado no controle da oscilação de xenônio de um reator PWR, por um otimizador baseado no algoritmo por enxame de partículas. A terceira e última aplicação do algoritmo por enxame de partículas, envolve a determinação, de forma inovadora, do número mínimo de vetores de referência capazes de permitir a melhor separação e classificação de acidentes durante a operação de um reator nuclear.

1.3.1. Projeto do Núcleo de um Reator

O projeto do núcleo de um reator nuclear envolve não linearidades e descontinuidades e um grande número de parâmetros a serem maximizados e sujeitos a um grande número de restrições. A obtenção de uma solução exige, como consequência, a busca de soluções em um espaço de alta dimensão envolvendo funções multimodais (PEREIRA, MARTINEZ e SCHIRRU, 1999).

Os resultados obtidos com o PSO foram comparados com a solução obtida por meio de um algoritmo genético em (SCHIRRU, MARTINEZ, PEREIRA, DOMINGOS e MACHADO, 1999).

1.3.2. Controle Preditivo da Oscilação Axial de Xenônio

O xenônio 135, um dos principais subprodutos da fissão do urânio, é produzido em pequena parte como subproduto da fissão do urânio do núcleo e em grande parte como resultado do decaimento do iodo 135 precursor (GLASSTONE, 1960 e LAMARSH, 1966). O xenônio, devido a sua alta seção de choque de absorção, e como efeito da variação do nível de produção de xenônio no núcleo, é propenso a induzir oscilações espaciais de potência do núcleo (DUDERSTAD e HAMILTON, 1976). As oscilações induzidas na potência axial no núcleo, ver ANEXO 2, são não lineares e influenciadas por parâmetros variantes no tempo tais como: nível de potência, posição das barras e nível de boro. A distribuição axial de potência em um reator PWR é função do tempo, da posição axial (z) e depende do nível de potência, do nível de queima e da variação da concentração de xenônio, apresentando oscilações com ciclos de período entre 20 e 30 horas. O controle da oscilação axial de potência se faz por meio do controle da posição das barras de controle de modo.

Os sistemas de controle preditivo (CLARKE, MOHTADI, 1987) são sistemas que utilizam o modelo da planta para predição do comportamento futuro da planta e um otimizador que, considerando a evolução do controle e da saída da planta, procura determinar o melhor valor do sinal de controle de modo a minimizar o erro na saída da planta e ao mesmo tempo minimizar a ação do controle (CAMACHO e BORDONS, 1995).

A predição da saída da planta pode ser realizada por meio de redes neurais (HAYKIN, 1994) e (HAYKIN, 1996) configuradas e treinadas para predição de séries temporais. O otimizador procura otimizar uma função objetivo que induz uma minimização dos erros de saída e ao mesmo tempo minimiza a ação de controle. Usualmente a minimização é implementada por algoritmos de otimização baseados em gradiente que dependem do gradiente e do hessiano do modelo da planta, ver ANEXO 1. Neste trabalho procuramos explorar a capacidade do PSO na determinação de ótimos globais sem a necessidade de conhecimento a priori do espaço de busca nem da existência de limites e derivadas da função possibilitando assim o uso de qualquer modelo de rede para a planta sem a necessidade de reprogramar o otimizador.

1.3.3. Classificação de Acidentes em Usinas Nucleares

A operação de uma usina envolve a operação de vários sistemas cuja falha pode provocar uma evolução adversa das condições operacionais da usina que, se não forem diagnosticadas em tempo ou se não forem corretamente diagnosticadas, podem produzir conseqüências não só para a usina como também para o público e meio ambiente. A fim de assegurar uma operação segura e reduzir os custos de operação de uma usina nuclear, diversos sistemas da usina devem ser monitorados fornecendo ao operador informações sobre o estado desses sistema para a tomada de decisão. Para operar uma usina nuclear, os operadores monitoram um grande número de medidas, obtidas a partir da instrumentação instalada nesses sistemas, e, em caso de um evento anormal de operação da usina, devem identificar e diagnosticar a situação da usina a partir da análise e classificação das medidas correspondentes ao evento em curso.

A evolução das condições operacionais da usina usualmente apresenta um comportamento transiente com padrões temporais bem definidos (assinaturas) dessas medidas para cada situação operacional. O grande número de instrumentos e a dinâmica de variação de cada grandeza medida associada ao evento em curso dificultam o processo de identificação e tomada de decisão por parte do operador que quando pressionado a tomar uma decisão reduz ainda mais a sua capacidade cognitiva aumentando assim a chance de uma tomada de decisão errada por parte do operador.

A fim de auxiliar o operador na tomada de decisões e reduzir ou minimizar a carga cognitiva do operador, sistemas de apoio à tomada de decisão, tais como sistemas de classificação de transientes e diagnóstico de acidentes, são incorporados ao sistema de operação da usina com o objetivo de aumentar o tempo disponível do operador para a tomada de uma resposta rápida e correta minimizando o risco de uma identificação errada. Um sistema de classificação de transientes deve fornecer ao operador um diagnóstico rápido e correto do evento em curso e assim aumentar o tempo disponível do operador para a tomada de ações corretivas de modo a manter a usina em uma condição segura.

Para identificar um evento, um sistema de classificação de transientes utiliza um conjunto de assinaturas de um conjunto de variáveis de operação da usina para cada evento anômalo da usina. Quando um evento de operação anormal deve ser identificado, o sistema de diagnóstico procura comparar a evolução das variáveis correspondentes às condições operacionais da usina com a assinatura de evolução das variáveis para cada um dos eventos anômalos postulados da usina e classificar o evento em curso pelo evento ao qual o evento está mais próximo usando para isso uma medida de distância.

Em (ALVARENGA, 1998) encontramos uma lista de acidentes postulados para uma usina nuclear bem como as variáveis de processo consideradas como as que mais contribuem para a caracterização desses acidentes. Os transientes postulados caracterizam as classes de transiente segundo a qual um determinado evento poderá ser classificado.

A maneira mais simples e rápida de classificar os eventos temporais da usina consiste em representar as assinaturas dos transientes postulados por protótipos (usualmente o centróide) e usando uma métrica, como por exemplo a distância euclidiana, comparar as distancias dos protótipos do evento aos protótipos dos transientes postulados. Cada evento será classificado segundo a classe do transiente a cujo protótipo o protótipo do evento estiver mais próximo. Esse método de classificação, denominado de Método de Classificação Simples (MCS), embora simples apresenta algumas deficiências de classificação, ver (SCHIRRU, PEREIRA e MARTINEZ, 1999) e (ALMEIDA, 2001), como o caso em que as classes são muito próximas ou os centróides são coincidentes.

Para superar as deficiências do MCS, foi desenvolvido o Método do Conjunto Mínimo de Centróides (CMC), (SCHIRRU, PEREIRA e MARTINEZ, 1999), que, particionando o espaço segundo o eixo do tempo, procura determinar duas ou mais subclasses que forneçam o melhor conjunto de centróides de modo que o MCS possa distinguir as classes e classificar os eventos corretamente. Na busca desse particionamento, o método de classificação de mínimos centróides (CMC), utiliza um algoritmo genético, (SCHIRRU, PEREIRA e MARTINEZ, 1999), para maximizar o número de acertos e minimizar o número de partições.

Nos modelos de minimização de particionamento do espaço, o que se procura é obter o número mínimo de partições, sendo cada partição representada por um protótipo, o centróide da partição. Uma vez que nada obriga a que os protótipos para classificação sejam os centróides resultantes do particionamento, e o que se procura é o menor número de partições (menor número de partições é igual a menor número de protótipos) que maximizem o número de acertos, pensou-se em deixar que um algoritmo de busca resolva o problema através de um processo de busca “cega” desses protótipos sendo o desempenho de cada protótipo avaliado por uma função que mede o número de acertos para classificação de todos os acidentes para um número prefixado de protótipos para cada acidente.

Para resolver o problema de número mínimo de protótipos, usamos aqui o algoritmo de otimização por enxame de partículas para o qual fixamos diferentes valores para o número mínimo de protótipos por classe a ser procurado. Esta abordagem estabelece um modelo de solução, que corresponde ao processo de busca dos vetores de Voronoi (RUSSEL e NORVIG,1995) para separação e identificação das classes. Os resultados obtidos mostram soluções triviais não obtida por outros métodos e foram comparados com os resultados obtidos por meio de algoritmos genéticos com base na determinação do número mínimo de partições relatados em (ALMEIDA, 2001).

1.4. Organização do Trabalho

O presente trabalho está organizado em seis capítulos e 3 anexos sendo o conteúdo dos demais capítulos como descrito a seguir.

O presente capítulo, apresenta o objetivo da tese, a metodologia para desenvolvimento dos trabalhos, descreve os principais testes realizados e apresenta a organização do documento da tese.

O capítulo 2 apresenta o formalismo do método de otimização baseado em enxame de partículas. Inicialmente, são introduzidos os conceitos e definições e é apresentado o equacionamento que descreve o modelo canônico do algoritmo. A seguir é apresentada uma análise da convergência do método considerando a versão determinística do algoritmo mostrando condições sob as quais a convergência pode ser obtida. Finalmente são feitas considerações sobre o a influência do ajuste dos parâmetros para obtenção de resultados.

O capítulo 3 apresenta e descreve os resultados da aplicação do método de otimização por enxame de partículas ao problema de otimização do projeto do núcleo de um reator nuclear. Inicialmente é apresentado o problema de otimização do projeto de um reator nuclear. A seguir é apresentada a modelagem do problema para solução pelo PSO e finalmente são mostrados os resultados obtidos com a aplicação do PSO.

O capítulo 4 apresenta e descreve os resultados da aplicação do método ao módulo de otimização de um controlador preditivo aplicado ao controle da oscilação de xenônio de um reator nuclear PWR. Inicialmente é apresentado o problema de controle da oscilação do xenônio do reator PWR. A seguir são apresentados os fundamentos e os mecanismos envolvidos no controle preditivo. Finalmente é apresentada a aplicação do algoritmo de otimização por enxame de partículas como substituto ao módulo de otimização com apresentação dos resultados obtidos.

O capítulo 5 apresenta e descreve os resultados da aplicação do método ao problema de classificação de acidentes em uma usina nuclear a partir da análise da evolução de um conjunto de variáveis. Inicialmente é apresentado o problema de classificação de acidentes e algumas considerações a respeito da validação do método de otimização baseado na minimização de centróides CMC. O trabalho apresenta uma solução de classificação baseada exclusivamente na capacidade do algoritmo encontrar o menor número de protótipos capazes de produzir o maior número de acertos na classificação de acidentes, ou seja, os vetores de Voronoi, que no caso limite são a solução ideal do problema.

No capítulo 6 são apresentadas as conclusões do presente trabalho e são apresentadas sugestões para trabalhos futuros.

O Anexo 1, ALGORITMOS DE OTIMIZAÇÃO BASEADOS EM GRADIENTE, apresenta um resumo dos principais algoritmos de otimização por busca com base no gradiente.

O Anexo 2, OSCILAÇÃO DE POTÊNCIA AXIAL DE UM REATOR PWR, apresenta o modelo de oscilação de potência bem como apresenta as equações que descrevem a oscilação de xenônio de um reator PWR.

O Anexo 3, REDES NEURAIAS ARTIFICIAIS, apresenta os conceitos básicos e descreve os principais mecanismos envolvidos no aprendizado e na implementação de redes neurais artificiais para utilização como modelos de processos.

CAPÍTULO 2

OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS

2.1. Introdução

A técnica de otimização PSO (Particle Swarm Optimization), traduzida para o português como Otimização por Enxame de Partículas, é uma técnica de computação estocástica baseada em populações, desenvolvida por (KENNEDY e EBERHART, 1995), que implementa uma metáfora do comportamento social da interação entre indivíduos (partículas) de um grupo (enxame).

A metáfora foi desenvolvida a partir da observação de bandos de pássaros e cardumes de peixes em busca de alimento em uma determinada região. Ao observar o comportamento desses grupos, verifica-se que o comportamento do grupo é influenciado pela experiência individual acumulada por cada indivíduo bem como pelo resultado da experiência acumulada pelo grupo.

No algoritmo por enxame de partículas, que também designaremos como algoritmo PSO, cada candidato a solução do problema corresponde a um ponto no espaço de busca. Essas soluções, denominadas partículas, tem associados um valor, que é avaliado individualmente para cada partícula e que indica a adequação da partícula como solução para o problema, e uma velocidade que define a direção do movimento da partícula. Cada partícula modifica sua velocidade levando em conta a melhor posição da partícula e a melhor posição do grupo, e ao longo do tempo o grupo acaba alcançando o alimento.

2.2. Algoritmo Básico

No algoritmo PSO o enxame é inicializado randomicamente, ver figura 2.1, com uma população de soluções candidatas sendo, cada partícula, inicializada com uma posição e uma velocidade randômicas.

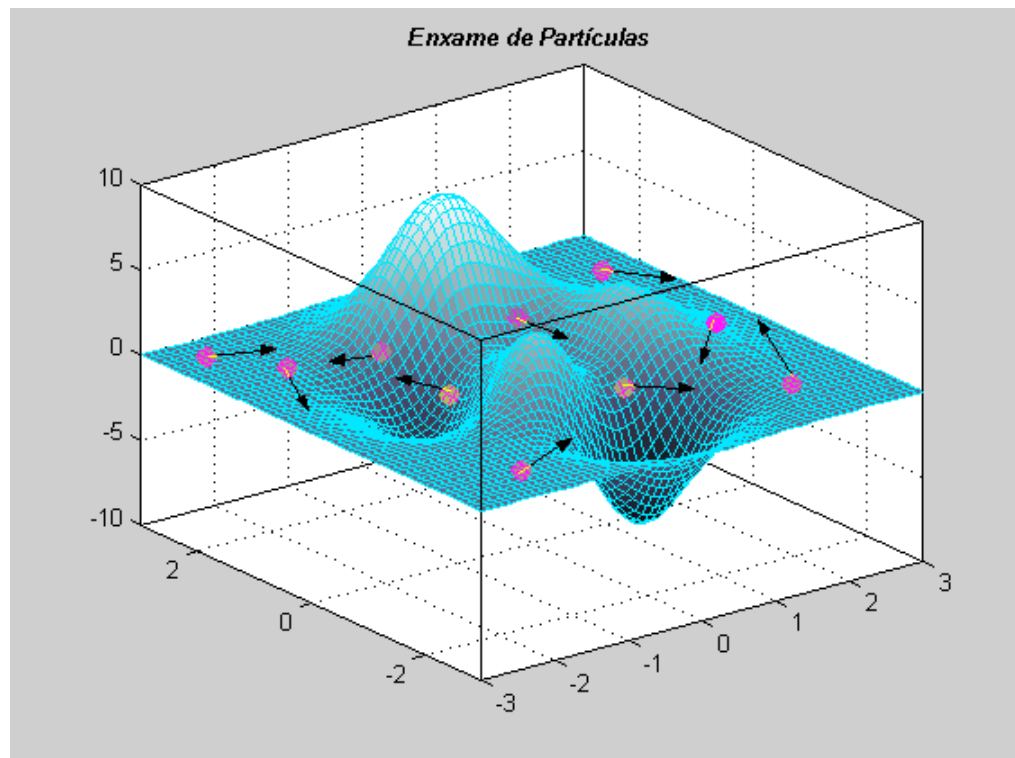


Figura 2.1 – Enxame de partículas

O algoritmo básico do PSO, denominado modelo canônico do PSO, pode ser descrito em notação vetorial como a seguir:

$$\vec{v}_{k+1} = a \otimes \vec{v}_k + b_1 \otimes r_1 \otimes (\vec{p}_1 - \vec{x}_k) + b_2 \otimes r_2 \otimes (\vec{p}_2 - \vec{x}_k) \quad \text{eq. (1)}$$

$$\vec{x}_{k+1} = c \otimes \vec{x}_k + d \otimes \vec{v}_{k+1} \quad \text{eq. (2)}$$

onde:

\otimes	produto vetorial elemento por elemento
\vec{v}	vetor velocidade da partícula
\vec{x}	vetor posição da partícula
\vec{a}	fator de momento
\vec{p}_1	melhor posição anterior da partícula
\vec{p}_2	melhor posição anterior do enxame
\vec{b}_1, \vec{b}_2	coeficientes de força de atração
\vec{r}_1, \vec{r}_2	vetores de números randômicos no intervalo [0,1]
\vec{c}, \vec{d}	vetores coeficientes

A figura 2.2 ilustra a atualização de uma partícula em um espaço de busca bidimensional. Para simplificar a ilustração, o fator de momento não foi considerado. Cada partícula modifica sua velocidade levando em conta a melhor posição da partícula e a melhor posição do grupo.

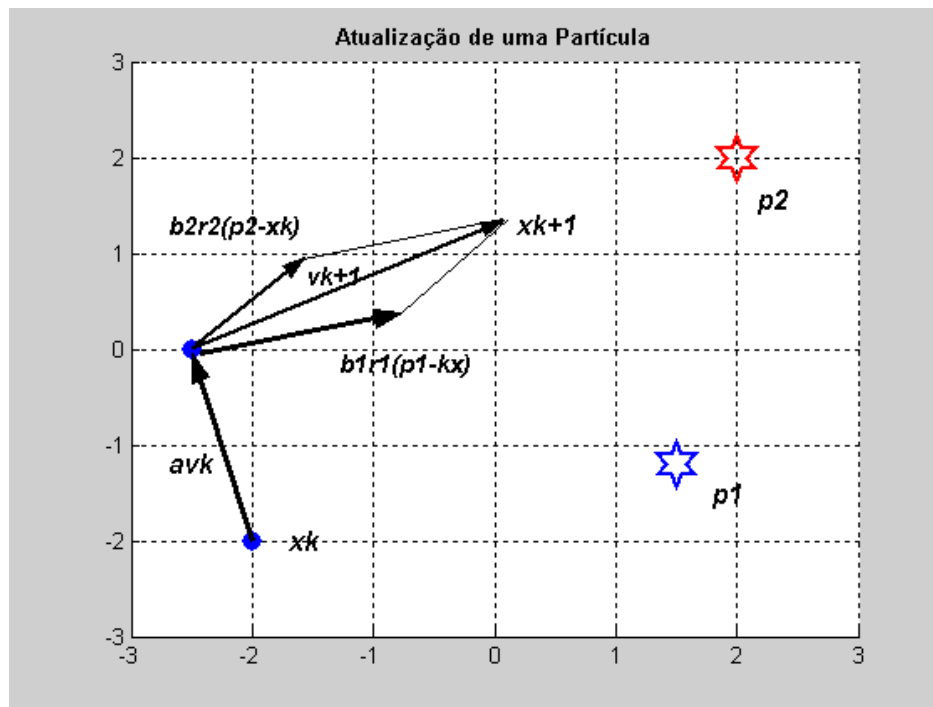


Figura 2.2 – Atualização de uma partícula

Nas equações (1) e (2) cada dimensão é atualizada independentemente das outras. O único vínculo entre as dimensões se dá por meio da função objetivo através da influência exercida pelas melhores posições anteriores \vec{p}_1 e \vec{p}_2 . Essa independência permite reduzir a análise do caso multidimensional ao caso unidimensional:

$$v_{k+1} = av_k + b_1r_1(p_1 - x_k) + b_2r_2(p_2 - x_k) \quad \text{eq. (3)}$$

$$x_{k+1} = cx_k + dv_{k+1} \quad \text{eq. (4)}$$

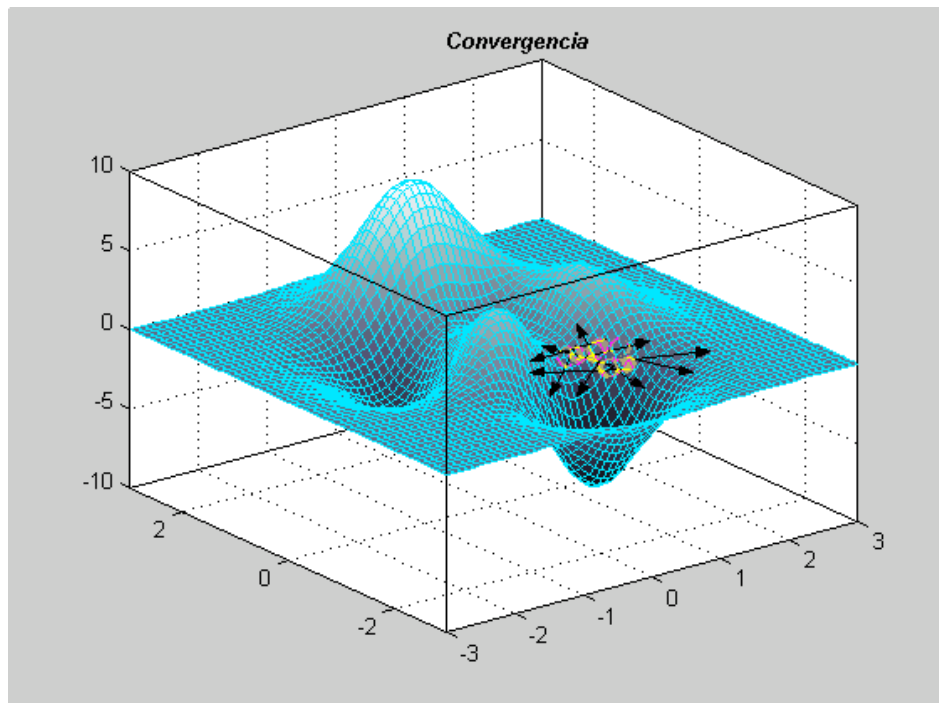


Figura 2.3. Convergência do PSO

A mudança de velocidade das partículas modifica a posição das mesmas fazendo-as mover-se através do espaço do problema. Ao longo de sucessivas iterações, como resultado da decisão individual e da influência social, o enxame de partículas acaba convergindo, ver figura 2.3, para uma solução ótima.

2.3. Modelo Determinístico

A versão determinística do PSO (TRELEA, 2000), é obtida das equações (3) e (4) atribuindo-se aos números randômicos seus valores esperados, isto é, $r_1 = r_2 = \frac{1}{2}$. A equação pode ainda ser simplificada fazendo:

$$b = \frac{b_1 + b_2}{2} \quad \text{eq. (5)}$$

$$p = \frac{b_1}{b_1 + b_2} p_1 + \frac{b_2}{b_1 + b_2} p_2 \quad \text{eq. (6)}$$

de modo que a expressão da versão determinística pode ser expressa como:

$$v_{k+1} = av_k + b(p - x_k) \quad \text{eq. (7)}$$

$$x_{k+1} = cx_k + dv_{k+1} \quad \text{eq. (8)}$$

O coeficiente de atração equivalente b , assim definido, representa a média dos coeficientes de atração individual e de atração social b_1 e b_2 . O ponto de atração equivalente, p é a média ponderada de p_1 e p_2 .

A versão determinística do algoritmo definida pelas equações (7) e (8) depende do ajuste de quatro parâmetros a, b, c e d . Eliminando-se a velocidade dessas equações, obtém-se uma equação recursiva de segunda ordem, ver equação (9), que apenas envolve posições sucessivas da partícula e que só depende do produto bd :

$$x_{k+1} + (bd - a - c)x_k + acx_{k-1} = bdp \quad \text{eq. (9)}$$

Fixando-se o valor de $d=1$, qualquer seqüência de posições da partícula geradas pelo algoritmo PSO, descritas pelas equações (7) e (8), pode ser gerada com $d = 1$ e algum valor para b .

$$x_{k+1} + (b - a - c)x_k + acx_{k-1} = bp \quad \text{eq. (10)}$$

Embora a seqüência de velocidades seja diferente, a convergência do algoritmo de otimização não será afetada pois a função objetivo apenas depende da posição sendo a velocidade apenas uma variável auxiliar. Do ponto de vista de otimização, o que se espera é que no longo prazo a população de partículas termine convergindo para o ótimo local anteriormente encontrado, isto é:

$$\lim_{k \rightarrow \infty} x_k = p \quad \text{eq. (11)}$$

Levando-se isto na equação (10) obtém-se a condição necessária:

$$(a - 1)(c - 1) = 0 \quad \text{eq. (12)}$$

As soluções $a = 1$ ou $c = 1$ são equivalentes no que se refere à seqüência de partículas definida pela equação (9); isso se deve ao fato de que a equação(9) é simétrica em relação a a e c . Escolhendo-se $c = d = 1$, a velocidade da partícula, equação (7), pode ser interpretada como a diferença entre duas posições diferentes da partícula, equação (8).

2.4. Dinâmica da Partícula

O comportamento dinâmico da partícula pode ser avaliado recorrendo-se à teoria de sistemas dinâmicos lineares discretos, (CADZOW, 1970 e LUENBERGER, 1975), que proporciona mecanismos de análise do comportamento dinâmico de um sistema e que serão usadas para avaliar o comportamento dinâmico de uma partícula.

Considerando-se $c = d = 1$, as equações (7) e (8) podem ser escritas na forma matricial como a seguir:

$$y_{k+1} = Ay_k + Bp \quad \text{eq. (13)}$$

onde:

$$y_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix}, \quad A = \begin{bmatrix} 1-b & a \\ -b & a \end{bmatrix}, \quad B = \begin{bmatrix} b \\ b \end{bmatrix} \quad \text{eq. (14)}$$

y_k estado da partícula, constituído da posição e velocidade vigentes.

A matriz dinâmica, determina o comportamento temporal da partícula (assintótico, convergente, oscilatório, etc.)

p entrada externa, usada para mover a partícula em direção de uma posição especificada

B matriz de entrada, determina o efeito da entrada externa sobre a partícula.

O ponto de equilíbrio de um sistema é o ponto atingido pelo sistema em ausência de excitação externa (entrada externa $p = \text{constante}$). O ponto de equilíbrio, se existir, deve satisfazer a condição $y_{k+1}^{eq} = y_k^{eq}$ para qualquer k .

Assim, para uma partícula do enxame, a condição de equilíbrio é dada por:

$$y^{eq} = [p \ 0]^T, \text{ ou seja,}$$

$$x^{eq} = p \text{ e } v^{eq} = 0 \quad \text{eq. (15)}$$

O que significa que, quando a partícula atinge o ponto de equilíbrio, ou seja, não há melhores posições, então a partícula deve estar posicionada no ponto de atração e a velocidade da partícula deve ser nula.

A análise de convergência procura determinar se uma partícula partindo de um ponto fora de equilíbrio irá atingir o equilíbrio (convergência), e como essa partícula se moverá pelo espaço de estados na busca de melhores pontos.

De acordo com a teoria dos sistemas dinâmicos, (CADZOW, 1970 e LUENBERGER, 1975), o comportamento temporal de uma partícula depende dos dois autovalores da matriz dinâmica A . Os dois autovalores (reais ou complexos) são soluções da equação:

$$\lambda^2 - (a - b + 1)\lambda + a = 0 \quad \text{eq. (16)}$$

2.4.1. Convergência

A condição necessária e suficiente para que o ponto de equilíbrio seja estável (isto é, o ponto de equilíbrio seja um atrator) é que os dois autovalores λ_1 e λ_2 da matriz A , calculados segundo a equação (16), tenham valor menor do que 1.

Da análise das raízes da equação (16) resultam as seguintes condições:

$$a < 1, \quad b > 0, \quad 2a - b + 2 > 0 \quad \text{eq. (17)}$$

Assim, dada uma posição e velocidade iniciais, a partícula irá convergir para uma posição de equilíbrio, dada pelas equações (6) e (15), se os parâmetros do algoritmo forem escolhidos dentro da região triangular do plano (a,b), ver gráfico (A) da figura 2.4, determinada pelas condições definidas pela equação (17).

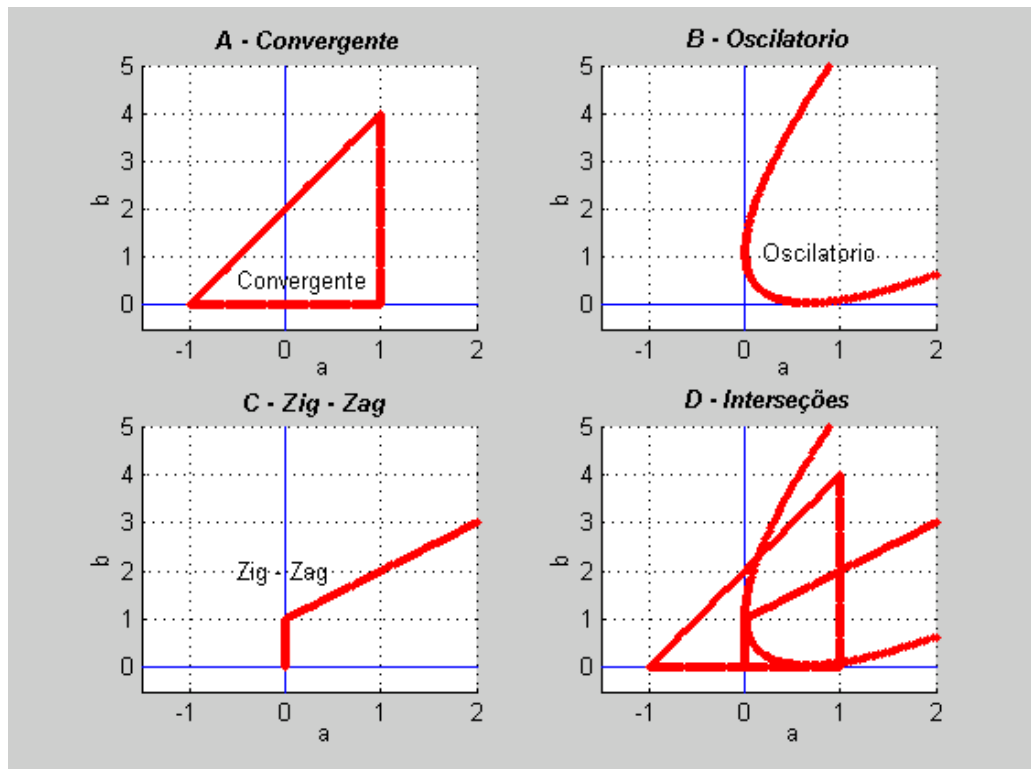


Figura 2.4. Regiões de convergência

2.4.2. Oscilações Harmônicas

De acordo com a teoria de sistemas lineares, se os autovalores λ_1 e λ_2 da matriz A forem complexos, a partícula irá exibir comportamento oscilatório harmônico antes de atingir a convergência. Da análise das raízes da equação (16) resulta a seguinte condição:

$$a^2 + b^2 - 2ab - 2a - 2b + 1 < 0 \quad \text{eq. (18)}$$

O gráfico (B) da figura 2.4 apresenta a região do plano (a,b), determinada pela equação (18), para a qual a partícula apresenta comportamento oscilatório harmônico.

2.4.3. Movimento Zig-Zag

Se pelo menos um dos autovalores λ_1 e λ_2 da matriz A apresentar parte real negativa, a partícula irá exibir movimento em zig-zag em torno do ponto de equilíbrio. Da análise das raízes da equação (16) resultam as seguintes condições:

$$a < 0 \quad \text{ou} \quad a - b + 1 < 0 \quad \text{eq. (19)}$$

O gráfico (C) da figura 2.4 apresenta a região do plano (a,b), definido pela equação (19), para o qual a partícula apresenta movimento em zig-zag.

2.5. Seleção de Parâmetros

O gráfico (D) da figura 2.4 mostra as interseções resultantes das diferentes regiões correspondentes às diferentes condições de convergência.

Escolhendo-se diferentes valores para os parâmetros do algoritmo pode-se obter diferentes comportamentos para a dinâmica da partícula.:

- O fator de momento a tem o papel de estabelecer um compromisso entre diversidade e refinamento da busca. Valores maiores do momento favorecem buscas abrangentes e valores menores favorecem o refinamento da busca.
- O ajuste adequado dos parâmetros b_1 e b_2 pode produzir uma rápida convergência e reduzir a possibilidade de convergência prematura.
- Os fatores randômicos r_1 e r_2 permitem manter a diversidade da população e são distribuídos uniformemente no intervalo $[0, 1]$.

A seguir são apresentados gráficos com diferentes exemplos de comportamento dinâmico de uma partícula para diferentes escolhas dos parâmetros a e b na otimização de uma função unidimensional por meio de um algoritmo com 30 indivíduos e um total de 50 iterações. Em cada gráfico são mostrados o comportamento da função a ser otimizada, a evolução da melhor posição da partícula, a evolução da melhor posição do enxame e a evolução da função objetivo do enxame.

A figura 2.5 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento convergente não oscilatório.

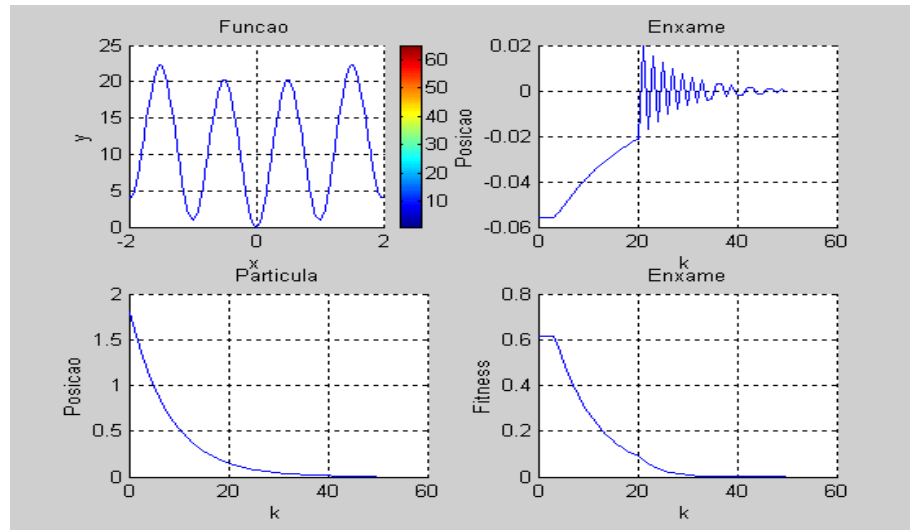


Figura 2.5. PSO convergente não oscilatório
 $a=0.1$, $b=0.1$

A figura 2.6 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento oscilatório convergente, com convergência lenta.

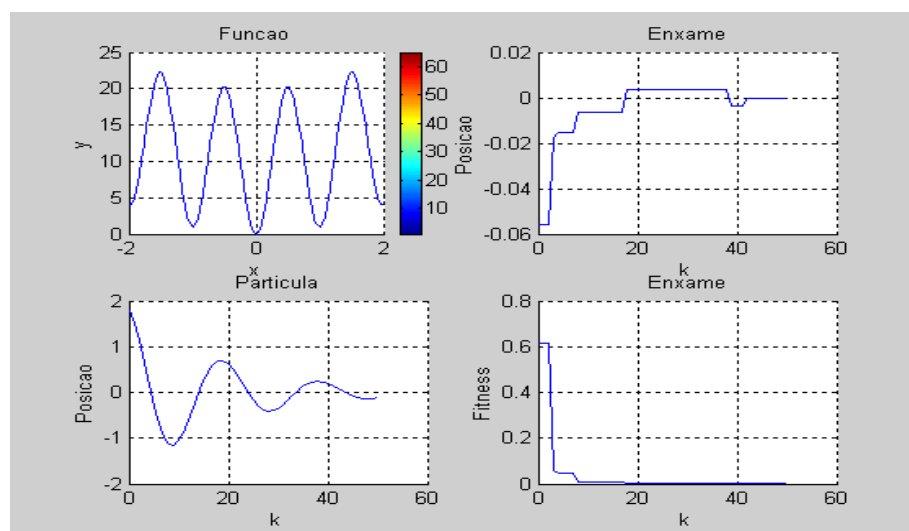


Figura 2.6. PSO, oscilação harmônica , convergência lenta
 $a=0.9$ e $b=0.1$

A figura 2.7 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento oscilatório convergente, com convergência rápida.

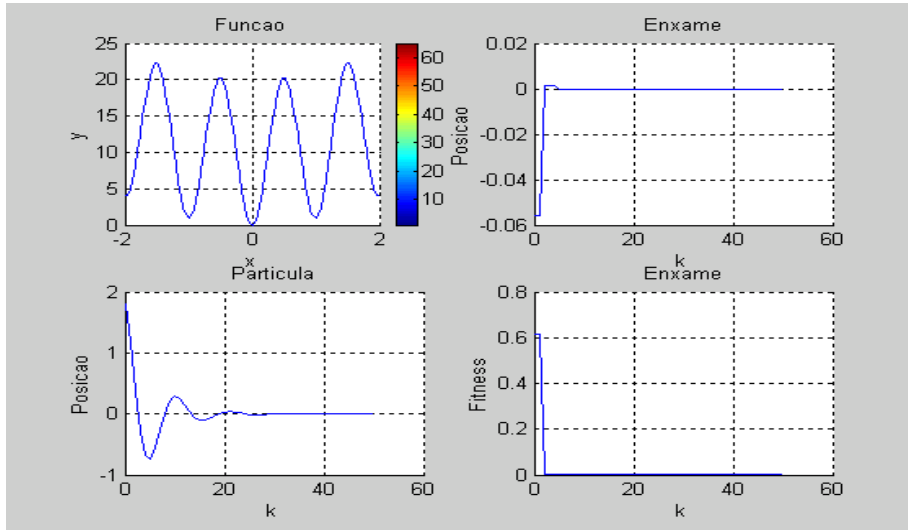


Figura 2.7. PSO, oscilação harmônica, convergência rápida
 $a=0.7$ e $b=0.3$

A figura 2.8 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento oscilatório com zig-zag.

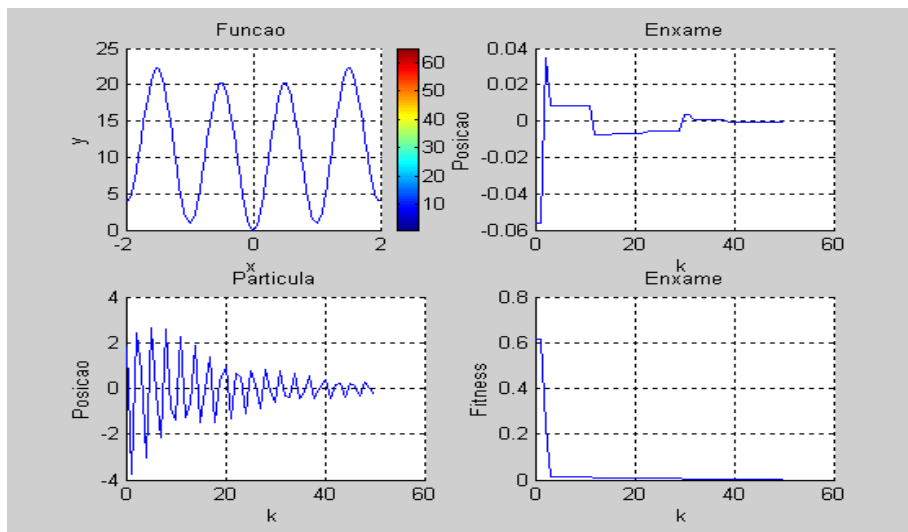


Figura 2.8. PSO, oscilação harmônica com zig-zag
 $a=0.9$ e $b=3.0$

A figura 2.9 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento em zig-zag convergente assimétrico.

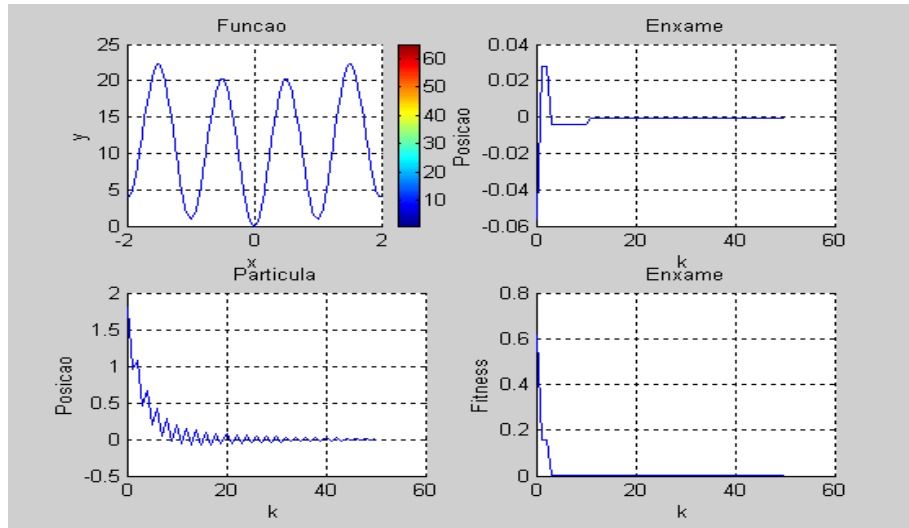


Figura 2.9 – PSO, zig-zag, convergente assimétrico
 $a = -0.7$ e $b = 0.5$

A figura 2.10 apresenta os resultados para valores de a e b escolhidos na região do plano (a,b) correspondente ao comportamento em zig-zag convergente simétrico.

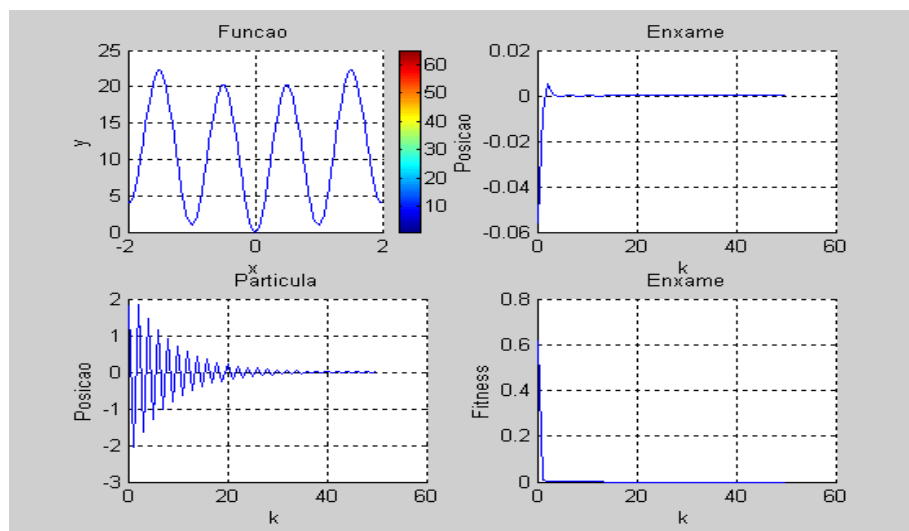


Figura 2.10 – PSO, zig-zag convergente simétrico
 $a = -0.1$ e $b = 2.1$

Como pode ser visto pelos exemplos anteriores, no caso do modelo determinístico do algoritmo, o comportamento oscilatório e o comportamento em zig-zag tendem a favorecer uma exploração maior do espaço, embora isso seja feito às custas de um retardo na convergência. Por outro lado, o comportamento convergente assimétrico deve ser evitado já que assim apenas uma parte do espaço é explorada.

A análise do comportamento dinâmico das partículas para o modelo randômico pode ser realizada, do ponto de vista qualitativo usando as mesmas considerações adotadas na análise do comportamento dinâmico do modelo determinístico. O ponto de atração equivalente, no caso do modelo randômico do algoritmo é dado por:

$$p = \frac{b_1 r_1}{b_1 + b_2} p_1 + \frac{b_2 r_2}{b_1 + b_2} p_2 \quad \text{eq. (20)}$$

Se $p_1 \neq p_2$ o ponto de atração muda de iteração para iteração mesmo que não ocorram melhores soluções, ou seja, mesmo que p_1 e p_2 permaneçam constantes. A longo prazo, espera-se que p_1 e p_2 convirjam para um mesmo ponto na medida em que todas as partículas da população convirjam para um mesmo ponto que passará a ser o único atrator. A equação (20) garante que o algoritmo converge para $p = p_1 = p_2$ independente da seqüência de números randômicos gerados.

A presença de números randômicos acentua a tendência de zig-zag e retarda a convergência, melhorando a exploração do espaço e prevenindo convergência prematura para pontos não ótimos, especialmente quando o ponto de auto atração p_1 está situado distante do ponto de atração da população p_2 .

CAPÍTULO 3

OTIMIZAÇÃO DO PROJETO DO NÚCLEO DE UM REATOR

3.1. Introdução

A fim de verificar e confirmar a eficácia do algoritmo de otimização por enxame de partículas, descrevemos neste capítulo a aplicação do algoritmo PSO na otimização do projeto do núcleo de um reator.

A otimização do projeto de um reator nuclear é um problema complexo que envolve não linearidades e descontinuidades, e sujeito a um grande número de restrições (PEREIRA, SCHIRRU e MARTINEZ, 1998).

A fim de obter uma operação segura e econômica de um reator, o projeto do núcleo deve obter a melhor configuração de parâmetros de projeto relativos à configuração e constituição dos elementos do núcleo e ao mesmo tempo obedecer a um conjunto de restrições relativos aos limites de operação de determinadas variáveis bem como dos limites dos parâmetros de configuração e constituição do núcleo (SILVENNOINEN, 1976).

A obtenção de uma solução exige, como consequência, a busca de soluções em um espaço de alta dimensão envolvendo funções multimodais. Em (PEREIRA, SCHIRRU e MARTINEZ, 1998) e (SCHIRRU, MARTINEZ, PEREIRA, DOMINGOS e MACHADO, 1999) encontramos a aplicação de algoritmos genéticos na otimização do projeto de um reator experimental.

A fim de poder avaliar o resultado obtido na otimização do projeto do núcleo com exame de partículas, utilizaremos os dados de projeto de (PEREIRA, SCHIRRU e MARTINEZ, 1998) comparando os resultados da aplicação do PSO com os resultados obtidos com a aplicação de algoritmos genéticos.

3.2. Projeto do Núcleo de um Reator Nuclear

O núcleo de reator a ser otimizado, ver figura 3.1, consiste de um núcleo com três regiões de enriquecimento definidas por $R1= 86$ cm, $R2=38$ cm e $R3=18$ cm.

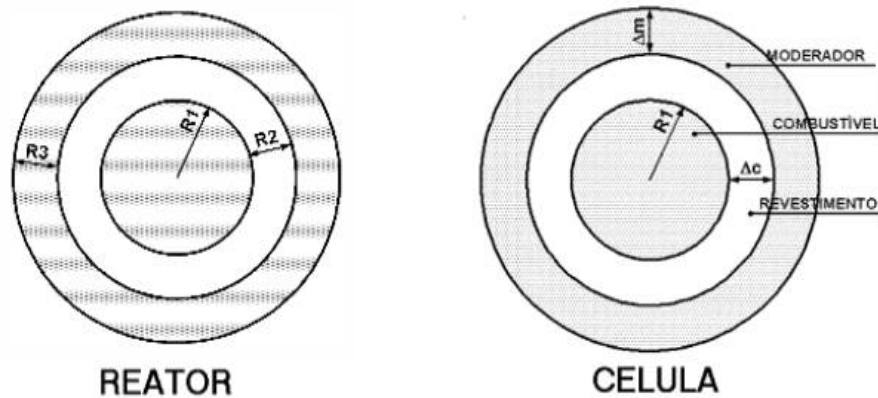


Figura 3.1 – Reator e célula característica

O objetivo da otimização é o de construir um núcleo que minimize o fator de pico médio f_m e mantenha dentro dos limites, ver tabela 3.1, o fluxo médio $\phi_m = 8.0 \times 10^{-5} \pm 1\%$ e fator de multiplicação $k_{ef} = 1.0 \pm 0.01\%$ submoderado. O fluxo de neutrons é normalizado considerando-se a fonte de neutrons de fissão igual a unidade.

Tabela 3.1 – Limites operacionais do reator

Variável	Mínimo	Máximo
ϕ_m [$\text{cm}^{-2} * \text{sec}^{-1}$]	7.92×10^{-5}	8.08×10^{-5}
k_{ef}	0,99	1.01

A fim de atingir esse objetivo, o raio do combustível, as espessuras e materiais do revestimento e moderador, o raio equivalente, bem como o enriquecimento das três zonas podem variar obedecendo os limites especificados na tabela 3.2.

Tabela 3.2 – Limites dos parâmetros do projeto

Parâmetro	Mínimo	Máximo
R_f [cm]	0,508	1,27
Δ_e [cm]	0,0254	0,254
Δ_m [cm]	0,0254	0,762
E_1 [%]	2.0	5.0
E_2 [%]	2.0	5.0
E_3 [%]	2.0	5.0
Mf	1= Urânio Metal, 2=Óxido de Urânio	
Mc	1=Alumínio 2= Aço Inox 3= Zircaloy	

Onde:

R_f	Raio do combustível
Δ_c	Espessura do revestimento
Δ_m	Espessura do moderador
E_1	Enriquecimento da região 1
E_2	Enriquecimento da região 2
E_3	Enriquecimento da região 3
Mf	Material do elemento combustível
Mc	Material do revestimento

O problema de otimização é expresso como o problema de minimizar:

$$\phi_m(\cdot) \equiv \phi_m(R_f, \Delta_c, \Delta_m, E_1, E_2, E_3, Mf, Mc) \quad \text{eq. (1)}$$

sujeito ao seguinte conjunto de restrições:

$$\phi_{\min} < \phi_m(\cdot) < \phi_{\max} \quad \text{eq. (2)}$$

$$k_{ef\min} < k_{ef}(\cdot) < k_{ef\max} \quad \text{eq. (3)}$$

$$\frac{dk_{ef}(\cdot)}{dV_m} > 0 \quad \text{eq. (4)}$$

$$R_{f\min} < R_f < R_{f\max} \quad \text{eq. (5)}$$

$$\Delta_{c\min} < \Delta_c < \Delta_{c\max} \quad \text{eq. (6)}$$

$$\Delta_{m\min} < \Delta_m < \Delta_{m\max} \quad \text{eq. (7)}$$

$$E_{1\min} < E_1 < E_{1\max} \quad \text{eq. (8)}$$

$$E_{2\min} < E_2 < E_{2\max} \quad \text{eq. (9)}$$

$$E_{3\min} < E_3 < E_{3\max} \quad \text{eq. (10)}$$

$$Mf = \{UMetal \quad UO_2\} \quad \text{eq. (11)}$$

$$Mc = \{Al \quad AçoInox \quad Zircaloy\} \quad \text{eq. (12)}$$

onde V_m representa o volume do moderador e os subscritos min e max se referem aos mínimos e máximos das tabelas 3.1 e 3.2.

3.3. Modelagem do problema com o PSO

Cada indivíduo do enxame no algoritmo PSO consiste de um vetor n dimensional, onde n representa a dimensão do espaço do problema, que representa uma solução do problema .

Para otimização do núcleo do reator temos 8 dimensões correspondendo aos parâmetros do núcleo $R_f, \Delta_c, \Delta_m, E_1, E_2, E_3, Mf, Mc$ a serem ajustados. Embora o PSO trabalhe com espaços reais, para os parâmetros correspondentes aos materiais do combustível e do revestimento (Mf, Mc) apenas consideramos a parte inteira dos números reais.

Cada indivíduo do enxame é avaliado de acordo com a função de performance a ser minimizada a qual depende dos parâmetros e das restrições do problema do projeto do núcleo do reator. No caso, a função a ser minimizada é a função:

$$f = \begin{cases} f_m & se \Delta k_{ef} \leq 0.01; \Delta \phi \leq 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} > 0 \\ f_m + r_1 \Delta k_{ef} & se \Delta k_{ef} > 0.01; \Delta \phi \leq 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} > 0 \\ f_m + r_2 \Delta \phi & se \Delta k_{ef} \leq 0.01; \Delta \phi > 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} > 0 \\ f_m + r_3 \frac{\Delta k_{ef}}{\Delta V_m} & se \Delta k_{ef} \leq 0.01; \Delta \phi \leq 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} < 0 \\ f_m + r_1 \Delta k_{ef} + r_2 \Delta \phi & se \Delta k_{ef} > 0.01; \Delta \phi > 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} > 0 \\ f_m + r_1 \Delta k_{ef} + r_3 \frac{\Delta k_{ef}}{\Delta V_m} & se \Delta k_{ef} > 0.01; \Delta \phi \leq 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} < 0 \\ f_m + r_2 \Delta \phi + r_3 \frac{\Delta k_{ef}}{\Delta V_m} & se \Delta k_{ef} \leq 0.01; \Delta \phi > 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} < 0 \\ f_m + r_1 \Delta k_{ef} + r_2 \Delta \phi + r_3 \frac{\Delta k_{ef}}{\Delta V_m} & se \Delta k_{ef} > 0.01; \Delta \phi > 0.01\phi_0; \frac{\Delta k_{ef}}{\Delta V_m} < 0 \end{cases}$$

eq. (13)

onde

$$\Delta k_{ef} = [1.0 - k_{ef}] \quad \text{eq. (14)}$$

$$\Delta \phi = [\phi - \phi_0] \quad \text{eq. (15)}$$

$$\frac{\Delta k_{ef}}{\Delta V_m} = \frac{k_{ef} - k_{ef}^!}{0.03 V_m} \quad \text{eq. (16)}$$

$$r_1 = r_2 = r_3 = 10 \quad \text{eq. (17)}$$

As equações (14) e (15) representam os valores das penalizações para configurações com k_{ef} e ϕ fora dos limites de 1%, e a equação (16) representa a penalização para a configuração de reator submoderado onde $k_{ef}^!$ é o fator de multiplicação obtido com 3% de variação do volume do moderador V_m (enquanto o volume do combustível V_f permanece constante, a razão $\frac{V_m}{V_f}$ se modifica dentro de limites determinados pelos parâmetros da tabela 3.2). As constantes da equação (17) representam os pesos correspondentes às importâncias relativas dessas penalizações.

3.4. Configuração do PSO

Os valores dos limites máximos e mínimos das $n=8$ dimensões, correspondentes aos parâmetros $R_f, \Delta_c, \Delta_m, E_1, E_2, E_3, Mf, Mc$ do núcleo, são definidos na tabela 3.2. O modelo de algoritmo PSO adotado para a solução do projeto do núcleo do reator corresponde ao modelo canônico do PSO.

Os parâmetros de configuração utilizados para o algoritmo PSO, escolhidos iguais para todas as dimensões, estão mostrados na tabela 3.3. Em cada iteração todos os indivíduos da população são avaliados utilizando para isso o programa HAMMER (SUICH e HONEC, 1967), para avaliação de configuração do núcleo.

Tabela 3.3 – Parâmetros de configuração do PSO

Parâmetro	Descrição	Valor
DIM	Dimensão de cada indivíduo	8
MAX ITER	Número máximo de iterações	Tabela 3.4
MAX POP	Numero de indivíduos do enxame	Tabela 3.4
<i>a</i>	Fator de inércia	0.73
<i>b1</i>	Coeficiente de força de atração individual	1.5
<i>b2</i>	Coeficiente de força de atração social	1.5
<i>c e d</i>	Coeficientes	1.0

O modelo adotado para o algoritmo PSO corresponde ao modelo canônico. De modo a permitir melhor exploração do espaço, os valores de *b1* e *b2* foram ajustados com o mesmo valor 1.5 e os valores dos coeficientes *c* e *d* são iguais a 1.

3.5. Resultados

As figuras 3.2, 3.3 e 3.4, a seguir, mostram a evolução e convergência da fitness do algoritmo PSO para diferentes populações (MAX POP) e diferentes números de iterações (MAX ITER), conforme tabela 3.4., mas mantendo o número total de avaliações (MAX ITER * MAX POP = 20000) .

Em todos os gráficos o eixo das abcissas representa o número de avaliações e o

eixo das ordenadas representa a fitness obtida.

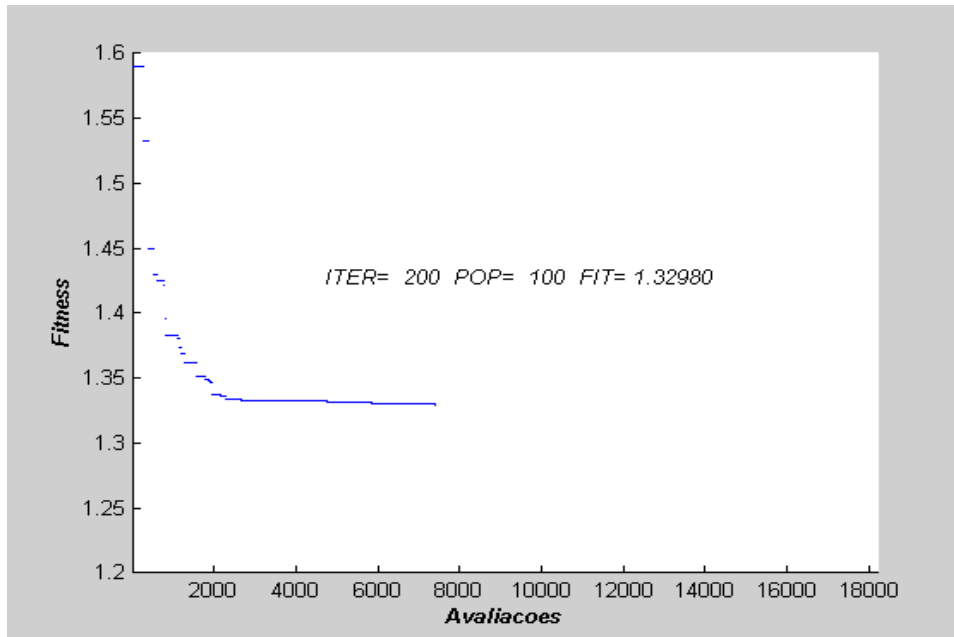


Figura 3.2 – Convergência com 200 Iterações de 100 partículas

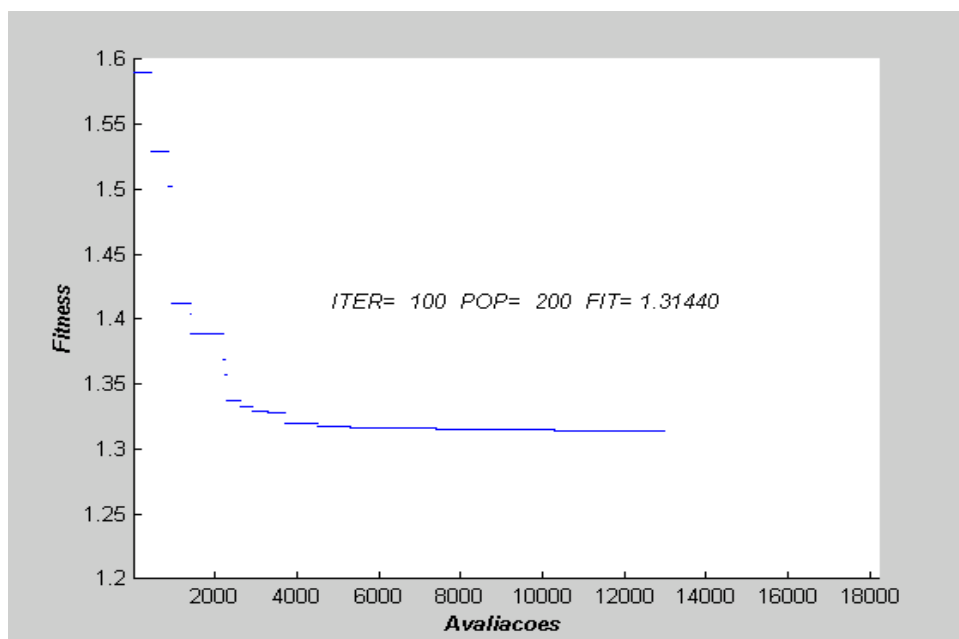
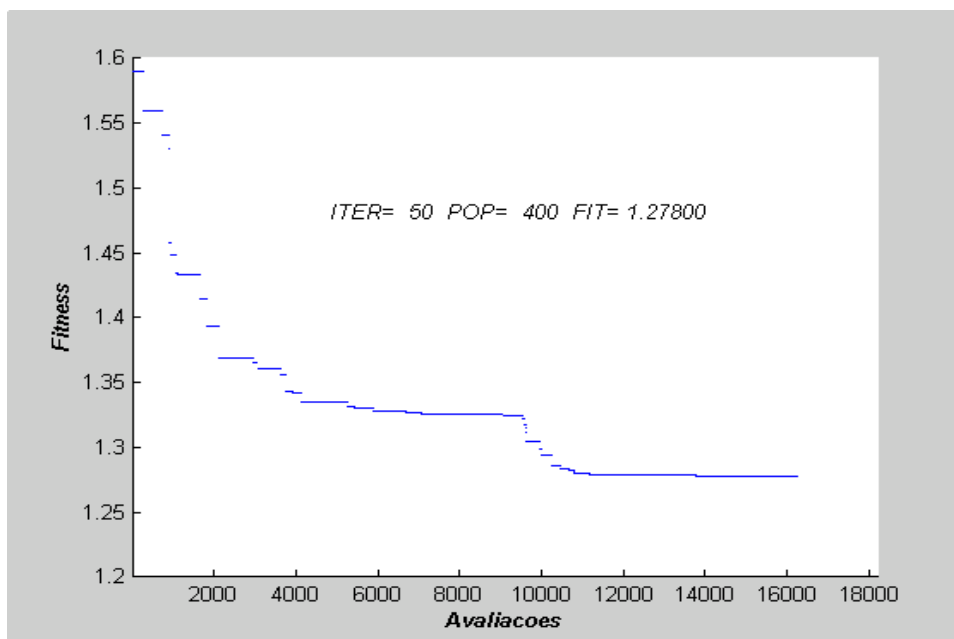


Figura 3.3 – Convergência com 100 Iterações de 200 partículas



Como pode se visto pelos resultados mostrados nas tabela 3.4, os melhores resultados foram obtidos com populações maiores, mostrando a sensibilidade do PSO ao tamanho população inicial.

Tabela 3.4 – Resultados do algoritmo PSO

MAX ITER	MAX POP	FITNESS
200	100	1,3298
100	200	1,3144
50	400	1,2780

Nos três casos o número total de avaliações ($\text{MAX ITER} * \text{MAX POP} = 20000$) é o mesmo. No entanto, a melhor fitness é obtida com um maior número de indivíduos e um menor número de iterações.

O melhor resultado obtido pelo PSO, obtido com 50 gerações de 400 indivíduos cada, pode ser visto na tabela 3.5 que mostra o resultado obtido com o PSO bem como o resultado obtido por meio do uso de algoritmo genético em (SCHIRRU, MARTINEZ, PEREIRA, DOMINGOS e MACHADO, 1999).

Tabela 3.5 – Comparação de resultados

Objetivos	AG	PSO
Mínimo fator de pico médio (f_m)	1.2949	1.2781
Fluxo médio (ϕ_m)	8.023×10^{-05}	8.023×10^{-05}
Fator de multiplicação (k_{ef})	0.9948	0.99026
Parâmetros do Núcleo	AG	PSO
R_f [cm]	0.6280	0.7968
Δ_c [cm]	0.1604	0.0706
Δ_m [cm]	0.6808	0.7592
E_1 [%]	2.7087	2.0000
E_2 [%]	3.0394	2.0993
E_3 [%]	4.7638	3.6207
Mf	1 = U metálico	1 = U metálico
Mc	2 = Aço Inox	2 = Aço Inox
Fitness	1.2949	1.2781

Como pode ser visto o algoritmo PSO obteve uma configuração, ainda dentro dos limites de 1% para o fluxo e fator de multiplicação, cuja fitness de 1.2780, inferior ao valor de fitness de 1.2949 da configuração obtida pelo algoritmo genético. Tanto o PSO como o algoritmo genético convergem em torno de 12000 avaliações, o que mostra uma eficiência maior do PSO sobre o algoritmo genético nessa comparação.

O algoritmo por enxame de partículas foi testado com diferentes sementes e convergiu valores de fitness consistentes com os valores obtidos confirmando sua robustez.

Uma vez que o que se deseja é minimizar o valor da função objetivo, e dada a complexidade do problema, concluímos que o algoritmo PSO foi eficaz na solução desse problema obtendo resultados consistentes e até mesmo superando o resultado obtido com o algoritmo genético.

CAPÍTULO 4

CONTROLE PREDITIVO DA OSCILAÇÃO DE XENÔNIO

4.1. Introdução

Neste capítulo apresentamos a aplicação do algoritmo por enxame de partículas na otimização de um sistema de controle preditivo aplicado no controle da oscilação de xenônio de um reator PWR.

Durante a operação de um reator nuclear, muitas vezes torna-se necessário modificar a distribuição axial de potência de um certo limiar para outro em um espaço de tempo relativamente curto. Usualmente essas operações são acompanhadas de uma mudança do nível de potência. O envenenamento do núcleo de um reator nuclear por produtos de fissão é um mecanismo inerente ao processo de fissão e um dos principais responsáveis pela distorção de reatividade do reator (GLASSTONE, 1960 e LAMARSH, 1966). O xenônio 135, um dos principais subprodutos da fissão do urânio, é produzido em pequena parte como subproduto da fissão do urânio do núcleo e em grande parte como resultado do decaimento do iodo 135 (precursor). O xenônio, devido a sua alta seção de choque de absorção, e como efeito da variação do nível de produção de xenônio no núcleo, é propenso a induzir oscilações espaciais de potência do núcleo (DUDERSTAD e HAMILTON, 1976). Essas oscilações são não lineares e são influenciadas por parâmetros variantes no tempo tais como: nível de potência, posição das barras e nível de boro.

Devido á variação da concentração de xenônio e á grande seção de choque de absorção do xenônio, oscilações de potência, ver ANEXO 2, são induzidas na potência axial do núcleo. A distribuição axial de potência em um reator PWR é função do tempo, da posição axial (z) e depende do nível de potência, do nível de queima e da variação da concentração de xenônio, apresentando oscilações com ciclos de período entre 20 e 30 horas. O controle da oscilação da potência axial do reator é realizado pela variação da seção de choque no núcleo do reator por meio da inserção e remoção das barras de controle do reator. O controle deve procurar acompanhar e manter o valor de potência da saída através da minimização do erro de potência e , ao mesmo tempo, procurando minimizar o tempo de estabilização, minimizar o “overshoot” e o “undershoot” de potência e reduzindo o esforço de movimentação das barras de controle.

O sistema de controle de xenônio deve estabelecer uma estratégia de controle da distribuição de potência obedecendo aos limites operacionais e de modo a não gerar qualquer oscilação na concentração de xenônio.

Em (SHIMAZU, Y., 1992 e 1995) encontramos uma proposta de controle da oscilação axial de xenônio com base em três “offsets” axiais de potência calculados a partir da informação do transiente. O objetivo da estratégia de controle é o de estabelecer uma trajetória que mantenha a distribuição de potência dentro dos limites operacionais sem gerar qualquer oscilação de xenônio. Nenhuma oscilação de xenônio pode ser gerada se os três “offsets” axiais forem iguais. A solução ótima pode ser obtida por meio de uma estratégia de controle que iguale os três “offsets” axiais em um espaço de tempo tão curto quanto possível.

(DOMINGOS, R.P., 2003) utiliza programação genética para o projeto automático de um sistema de controle capaz de estabelecer uma estratégia de controle da oscilação axial de xenônio em reatores PWR baseado na estratégia proposta por (SHIMAZU, 1992). Em (DOMINGOS, R.P., CALDAS, G.H.F., PEREIRA, C.M.N.A, SCHIRRU, R., 2003) programação genética é aplicada para projeto automático de um sistema especialista fuzzy para estabelecer uma estratégia de controle da oscilação axial de xenônio em reatores PWR. A estratégia de controle desenvolvida é aprendida a partir da estratégia proposta por (SHIMAZU, 1992).

Em (NA, G.M., UDADHYAYA., B.R., CHOI, J.I., 2000) encontramos o desenvolvimento de um algoritmo de controle adaptativo baseado em modelo da planta, que utiliza um modelo analítico para modelo do reator, aplicado no controle do fluxo axial em um reator PWR. Os parâmetros de controle são atualizados “on line” usando o método de mínimos quadrados para ajustar as condições operacionais variantes no tempo a fim de permitir tratar as condições operacionais variantes no tempo.

4.2. Controle Preditivo

Os sistemas de controle preditivo são sistemas de controle baseados no modelo do sistema controlado (planta) que utilizam o modelo da planta para prever o seu comportamento futuro, ver figura 4.1, permitindo assim, estabelecer um esquema de controle antecipativo a partir da previsão do controle futuro da planta. Os sistemas de controle preditivo possuem um preditor e um otimizador que, considerando um horizonte de tempo futuro de saída da planta e um horizonte recente da ação de controle, determina a saída do controlador de modo a otimizar um índice de performance do sistema controlador-planta.

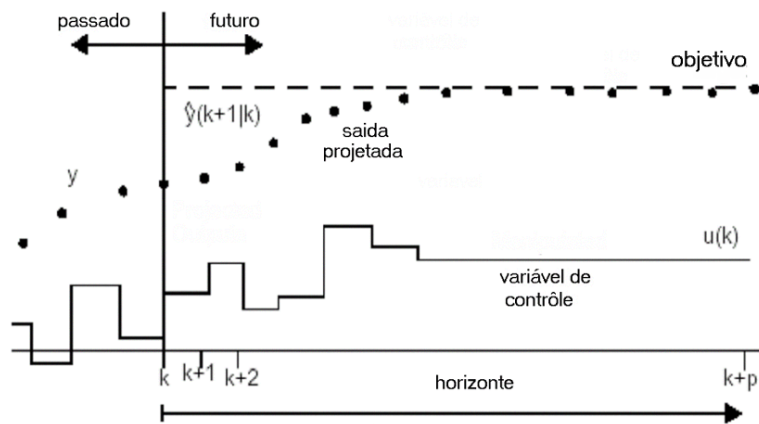


Figura 4.1 – Controle preditivo

A estrutura básica de um sistema de controle preditivo pode ser vista na figura 4.2, onde podem ser vistos os principais elementos envolvidos.

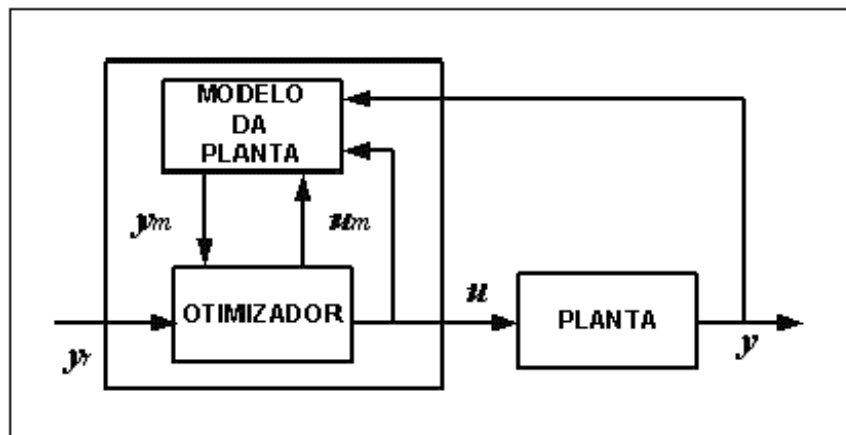


Figura 4.2 – Estrutura de um controlador preditivo

4.2.1. Preditor com Redes Neurais

O preditor do sistema de controle preditivo tem como objetivo, dada uma seqüência de controle, presente e futuro, determinar o comportamento futuro da planta considerando um horizonte p :

$$\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+m-1) \quad \text{eq.(1)}$$

$$y(k+1 | k), y(k+2 | k), \dots, y(k+p | k) \quad \text{eq.(2)}$$

onde :

$\Delta u(k+m-1)$ seqüência de controle

$y(k+p | k)$ comportamento futuro da planta

p horizonte de predição

O comportamento futuro da seqüência de saída planta pode ser previsto dentro de um horizonte p com auxílio de redes neurais (ROJAS,1996). As redes neurais realizam o mapeamento $f()$ de um conjunto de entrada x em um conjunto de saída y :

$$y = f(x) \quad \text{eq. (3)}$$

Por meio de um processo de treinamento supervisionado, as redes neurais, ver figura 4.3, aprendem o mapeamento $f()$ por através do ajuste dos pesos da rede a partir da apresentação para a rede dos conjuntos de dados de entrada e saída (x,y) .

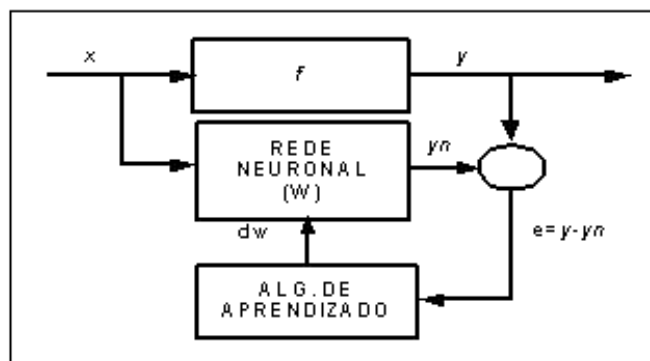


Figura 4.3 – Aprendizado supervisionado

Com a utilização de redes neurais, o sistema de controle não requer modelos analíticos da planta; o modelo da planta pode ser aprendido a partir de dados históricos de operação da planta.

Introduzindo-se linhas de retardo, ver figura 4.4, as redes neurais do tipo “feed-forward” podem ser usadas para prever o estado futuro da planta, ver ANEXO 3.

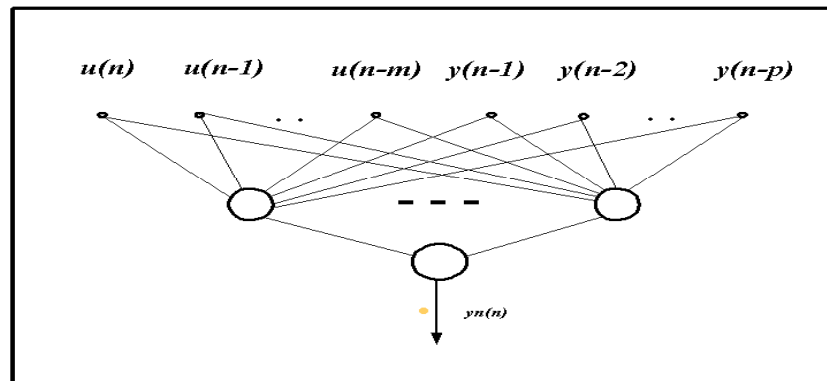


Figura 4.4 – Predição com rede feed-forward e linhas de retardo

4.2.2. Otimizador

O otimizador de um controlador preditivo estabelece a estratégia do controle: através da minimização de um índice de performance J as m ações de controle ($m \leq p$), presentes e futuras, são computadas de modo a minimizar uma função objetivo quadrática da forma:

$$J = \sum_{j=N1}^{N2} (y_r(t+j) - y_m(t+j))^2 + \rho \sum_{j=1}^{Nu} (\Delta u(t+j-1) - \Delta u(t+j-2))^2 \quad \text{eq. (4)}$$

onde:

- J índice de performance
- y_r valor de referência
- y_m predição de saída do modelo
- u controle tentativo
- $N1$ horizonte mínimo de custo ($N1 = 1$)
- $N2$ horizonte máximo de custo ($N2 = p$)
- N_U horizonte de custo do incremento de controle ($N_U = m$)
- ρ ponderação das ações de controle

A estratégia de controle, implementada pelo otimizador, induz a formação de um mecanismo de controle ótimo que procura reduzir o tempo de estabilização, reduzir o “overshoot” e “undershoot” bem como busca minimizar a atuação do controle estabelecendo para isso uma solução de compromisso entre esses objetivos.

4.2.3. Algoritmo do Controlador Preditivo

A cada passo de execução, o algoritmo do controlador preditivo gera seqüências de referência para o sinal de controle e determina o comportamento futuro da planta. O otimizador, usando as trajetórias de referência, calcula a performance da planta, determina a seqüência que minimiza a função objetivo, e então utiliza o primeiro valor da seqüência de controle para controlar a planta.

4.3. Configuração e Treinamento do Preditor

A tabela 4.1 mostra os parâmetros de configuração da rede neural que implementa o preditor. A rede neural consiste de uma rede “feed-forward” de três camadas com linhas de retardo para os sinais de entrada conforme mostrado na figura 4.3, considerando um horizonte de predição da saída $p=5$ e um horizonte de custo do controle $m = 3$. Todas as camadas utilizam como função de transferência a função sigmóide (logística).

Tabela 4.1. Configuração do preditor

Parâmetro	Descrição	Valor
p	Horizonte de predição da saída da planta	5
m	Horizonte de controle	3
N_L	Numero de camadas	3
S_0	Neurônios na camada de entrada	8
S_1	Neurônios na camada intermediária	200
S_2	Neurônios na camada de saída	1
$e_{training}$	Erro máximo de treinamento	2×10^{-4}
e_{test}	Erro máximo de teste	2×10^{-4}

Para predizer o comportamento futuro da planta, a rede neural deve ser treinada por um mecanismo de treinamento supervisionado. Para isso, dados de treinamento foram gerados com auxílio de um simulador da oscilação da potência axial da planta, ver ANEXO 2, simulando sinais randômicos na entrada da planta, e obtendo-se através do modelo as correspondentes saídas da planta.

A figura 4.5 mostra os dados simulados: a entrada da planta consiste da variação da seção de choque e a saída da planta consiste do sinal de fluxo (potência).

Os dados foram simulados, para de variação da seção de choque dentro de limites impostos pelo modelo da planta, tendo sido obtidos assim um total de 10000 pontos para treinamento e teste da rede. Os dados simulados para o treinamento da rede consistem de pares de valores $(u(t-1), y(t))$.

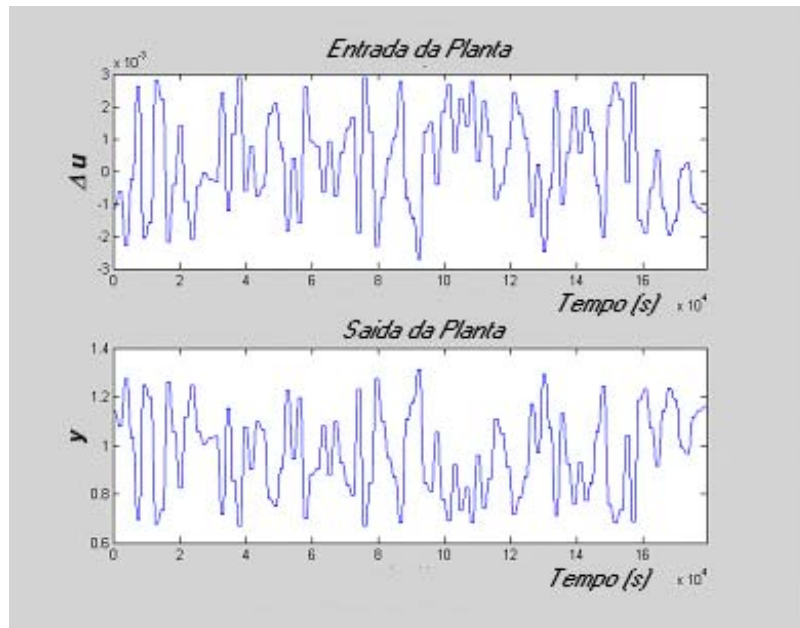


Figura 4.5. Dados simulados para identificação da planta

Os dados simulados são divididos em três conjuntos: dados de treinamento, dados de teste e dados de validação:

- Os dados de treinamento e de teste são utilizados para ajustar os pesos da rede durante o treinamento a cada época.
- Os dados de validação são utilizados para verificar a capacidade de generalização da rede quando submetida a dados que não foram apresentados durante o treinamento.

A figura 4.6 mostra a resposta da rede e o erro e resultantes da aplicação dos dados de treinamento.

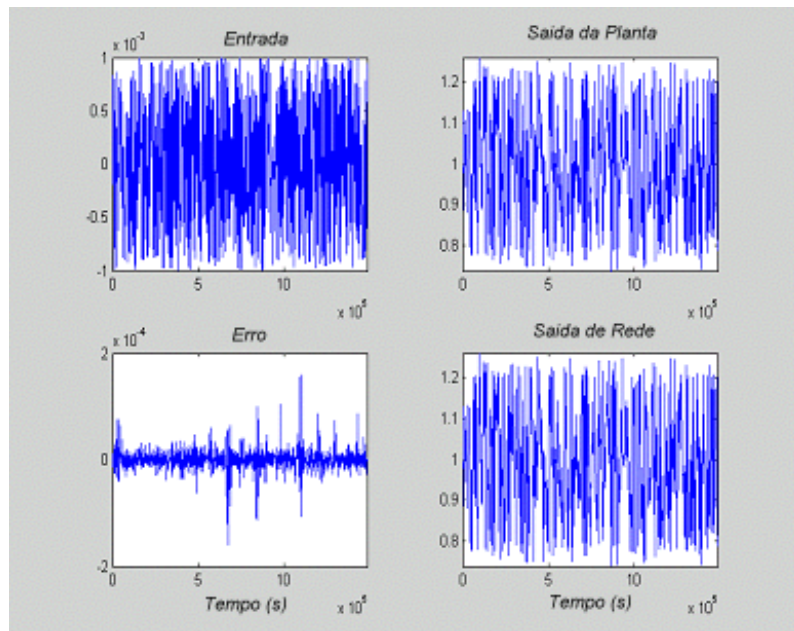


Figura 4.6 - Dados de treinamento

A figura 4.7, a seguir, mostra a resposta da rede e o erro e resultantes da aplicação dos dados de teste.

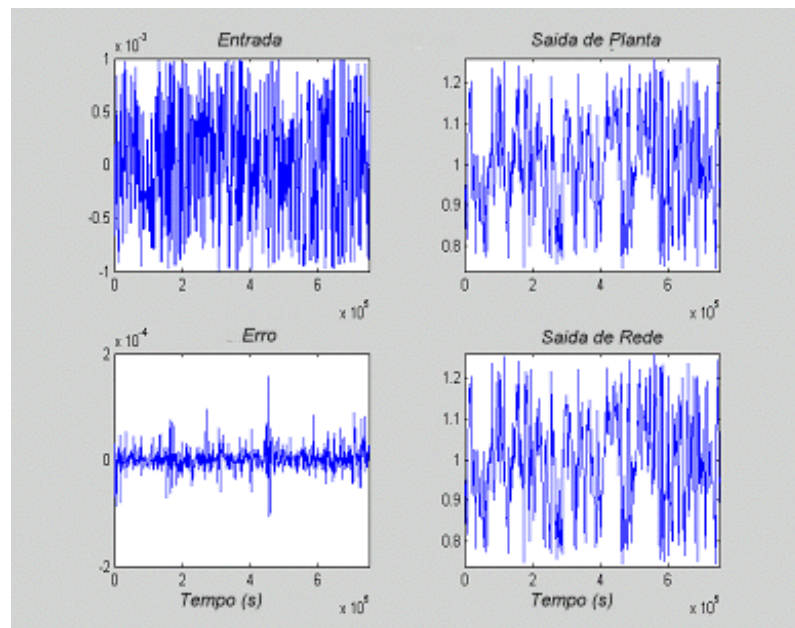


Figura 4.7 - Dados de teste

A figura 4.8, a seguir, mostra a resposta da rede e o erro e resultantes da



aplicação dos dados de validação.

A figura 4.9, a seguir, mostra a evolução do comportamento do índice de performance do treinamento (critério de erros mínimos quadrados - MSE) da rede neural mostrando a convergência do processo de aprendizado.

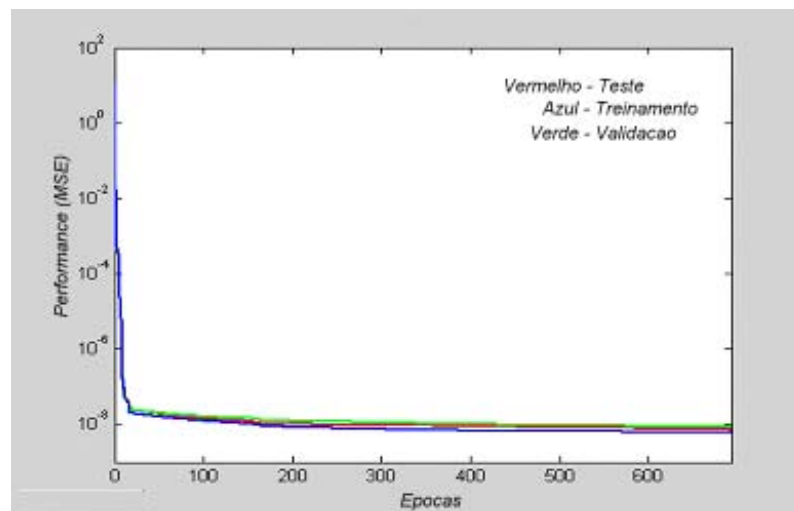


Figura 4.9. Performance de treinamento (MSE)

4.4. Otimizador com Enxame de Partículas

O algoritmo de otimização do controlador preditivo pode ser implementado por qualquer técnica de otimização. As técnicas descritas na introdução de algoritmos baseados no gradiente e suas variantes, ver ANEXO 1, são frequentemente utilizadas para essa finalidade.

4.5. Configuração do PSO

Nesta aplicação vamos usar o algoritmo de otimização baseado em enxame de partículas para testar a eficácia do método. Os parâmetros de configuração do otimizador estão mostrados na tabela 4.2.

Tabela 4.2. Parâmetros do otimizador com PSO

Parâmetro	Descrição	Valor
m	Horizonte de controle	5
DIM	Dimensão do indivíduo	m
X Min	Valores mínimos de cada dimensão do problema	0.0
X Max	Valores máximo de cada dimensão do problema	1.0
MAX ITER	Número máximo de iterações	50
MAX POP	Numero de indivíduos do enxame	100
a	Fator de inércia	0.7
b1	Coefficiente de força de atração individual	1.5
b2	Coefficiente de força de atração social	1.5
c e d	Coefficientes	1.0

4.6. Resultados

A figura 4.10 mostra o comportamento da planta submetida a uma perturbação da seção de choque de absorção $\Delta\Sigma_a = 0.001 \text{ cm}^{-1}$ e sem ação do controle. O valor da perturbação da seção de choque, corresponde a 1% dos limites da seção de choque absorção cujos limites vão de 1.50 a 1.54 cm^{-1} .

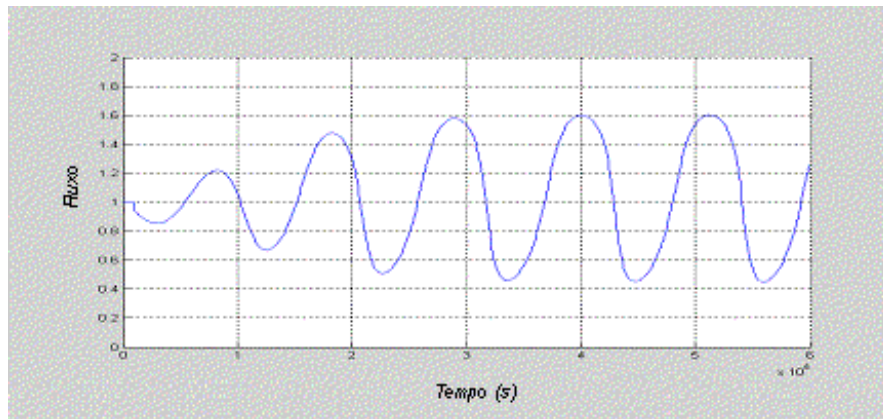


Figura 4.10 - Perturbação da planta

A figura 4.11 mostra a planta sob ação do controlador preditivo considerando a rede neural e o otimizador PSO conforme configurado no item anterior.

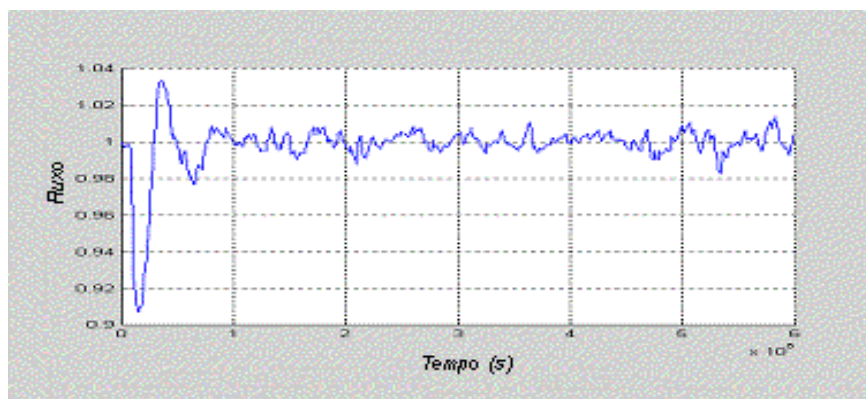


Figura 4.11- Predição neural e otimização com PSO

Como se vê pela figura 4.11, observa-se que embora o controlador preditivo atue rapidamente, a saída da planta não estabiliza e mantém uma oscilação em torno do ponto de equilíbrio. Isto se deve ao fato de que embora o erro obtido no treinamento da rede neural seja pequeno, basta uma pequena variação no sinal de controle para provocar uma alteração na saída da planta. Assim, qualquer erro, por pequeno que seja, na saída do preditor irá se refletir na ação de controle, provocando assim essa instabilidade.

A figuras 4.12 e 4.13 mostram o resultado utilizando como preditor um modelo preciso da planta e o algoritmo PSO como otimizador. A figura 4.12 mostra a ação de controle do controlador.

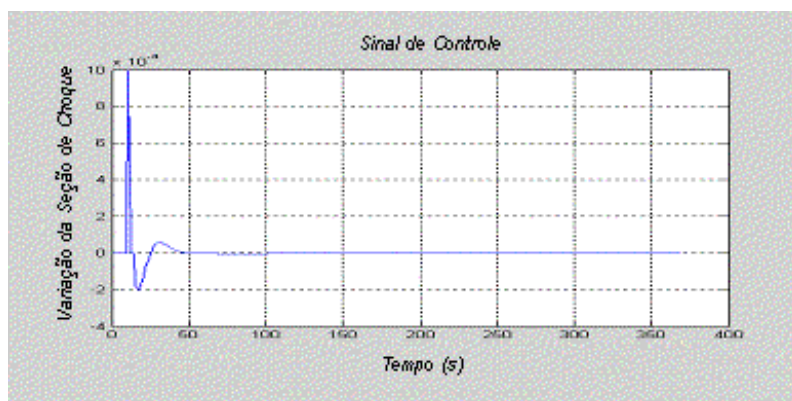


Figura 4.12 - Ação do controlador preditivo

A figura 4.13 mostra a saída da planta sob ação do controlador.



Figura 4.13 - Saída da planta sob ação do controlador

O resultado obtido mostra que o PSO é capaz de otimizar a função de performance sem conhecer a priori sua forma analítica e independentemente da existência de suas derivadas.

Como pode ser visto, o controlador preditivo atua rapidamente eliminando o desvio do fluxo fazendo com que a oscilação de xenônio seja suprimida. O controlador preditivo reduz o “overshoot” e o “undershoot”, elimina o erro de regime permanente e o tempo de estabilização é da ordem de 11 horas, , comparável ao tempo de estabilização (10 horas) de um controlador obtido por (DOMINGOS,2003) por meio de programação genética..

CAPÍTULO 5

CLASSIFICAÇÃO DE ACIDENTES EM USINAS

NUCLEARES

5.1. Introdução

Com o fim de mostrar a capacidade do algoritmo por enxame de partículas na otimização de sistemas complexos da área nuclear, apresentamos neste capítulo a solução com o algoritmo PSO para o problema de classificação de um conjunto de 3 acidentes postulados dentre os acidentes postulados para a usina nuclear de Angra 2. Nesta aplicação, além de corroborar a viabilidade e confirmar a eficiência do algoritmo como visto nos capítulos 3 e 4, foi possível propiciar uma solução inovadora e inesperada para tal problema. Em particular, na busca da solução do problema de identificação de acidentes, obteve uma solução que se aproxima da solução ideal, um protótipo por classe, ou seja, os vetores de Voronoi, das classes para a classificação de acidentes.

O diagnóstico de acidentes em usinas nucleares apresenta dificuldades para sua identificação pelo operador, especialmente em situações de pressão em que a capacidade cognitiva do operador fica prejudicada. Por essa razão sistemas de classificação de transientes para apoio à operação tem sido desenvolvidos para liberar o operador de tarefas cansativas e aumentar o tempo disponível para a tomada de ações corretivas (JEONG, E., FURUTA K., KONDO S., 1996).

Por essa razão, sistemas de diagnóstico de acidentes tem que ser rápidos e prover uma resposta correta (a melhor possível), a despeito da relativa insuficiência de informações disponíveis. Adicionalmente, a identificação deve ser robusta com relação a ruído nos dados bem como fornecer uma resposta “*não sei*” para transientes desconhecidos, não pertencentes ao conjunto de aprendizado.

Na construção de sistemas de classificação de transientes em usinas nucleares, diversas técnicas de inteligência artificial, envolvendo sistemas especialistas, sistemas neuro-fuzzy e algoritmos genéticos tem sido relatadas na literatura.

Um sistema utilizando redes neurais para diagnósticos de falhas em usinas nucleares proposto em (BARTLETT e UHRIG, 1992) com o trabalho “Nuclear Power Plant Diagnostics Using an Artificial Neural Network”, posteriormente aperfeiçoado por (BASU e BARTLETT, 1993) para identificação de transientes em um reator BWR, considera apenas os valores assumidos pelas variáveis escolhidas em um único instante de tempo, isto é, não é levada em consideração a evolução temporal das variáveis observadas, exigiu um grande número de variáveis necessárias para alcançar uma classificação satisfatória (no caso, 97 variáveis).

A fim de proporcionar uma resposta “*não sei*” quando apresentado a um novo transiente que não estava contido na base de conhecimento acumulado no sistema, (BARTAL, LIN e UHRIG, 1995) desenvolveram um classificador baseado em redes neurais probabilísticas que foi utilizado para classificar 72 cenários de 13 diferentes tipos de transientes, a partir dos valores observados de 76 variáveis ao longo do tempo. Eles também introduziram um mecanismo denominado acumulação de evidência mediante o qual resultados de classificações obtidos em instantes anteriores são usados como suporte de evidência para a classificação final.

Usando o mecanismo de acumulação de evidências, o classificador continua trabalhando em instantes de tempo independentes, mas a classificação final é computada usando uma votação majoritária dos valores obtidos em cada instante de tempo.

(FURUKAWA, UEDA E KITAMURA (1995) propuseram um método de classificação de eventos baseado em um classificador independente para cada variável observada. O classificador recebe em sua entrada a série temporal completa da variável selecionada e produz na saída a melhor classificação parcial possível em conjuntos de classes (isto é, um evento é incluído em uma das super classes que são separáveis a partir da informação trazida por uma simples variável). A interseção dos conjuntos de classes (super classes) gerada por todos os classificadores produz a classificação final. Este método tem a vantagem de ser muito robusto, já que a classificação final é baseada em múltiplos classificadores independentes. A dificuldade ocorre nos casos para os quais a função discriminação (critério de classificação) venha a ser muito complexa e seja necessária uma análise da interação entre duas ou mais variáveis.

JEONG, FURUTA E KONDO (1995) propuseram um modo alternativo para identificação de transientes usando o que eles chamaram de “adaptive template matching” (casamento de molde adaptativo) que utiliza informações ocorridas no passado. Este método usa redes neurais do tipo “feed-forward” (sem realimentações internas) para aproximar as trajetórias contínuas que representam os transientes no espaço de estado das variáveis, em vez de usar um número finito de padrões de treinamento das referidas trajetórias. O algoritmo permite não só identificar transientes diferentes, como avaliar vários transientes do mesmo tipo com diferentes níveis de severidade.

ROVERSO (1999) desenvolveu três métodos para classificação de eventos dinâmicos representados por séries temporais. Os dois primeiros utilizam algoritmos de agrupamento de padrões (“clustering”), baseados em conjuntos nebulosos, e redes neurais artificiais que conseguem avaliar as distâncias das amostras às classes do conjunto de treinamento, no caso, Funções de Base Radial (“RBF–Radial Basis Functions”) e Mapas de Kohonen. No terceiro método, foi usado um tipo especial de RNA recorrente (com realimentações internas) –o Classificador de Elman– que tem capacidade de lidar diretamente com séries temporais, dispensando a etapa de agrupamento dos padrões. Os resultados obtidos pelos três métodos foram comparados e o Classificador de Elman apresentou o melhor desempenho.

Em (ALVES, 1993) um sistema de diagnóstico utiliza um conjunto de RNA, do tipo “backpropagation”, para análise de algumas variáveis da usina, previamente escolhidas, de tal forma a poder identificar os transientes rapidamente. No caso de ocorrência de um evento que não seja reconhecido pelo sistema, um sistema especialista, com interface amigável, faz perguntas aos operadores procurando obter informações adicionais sobre as condições da usina e, assim, poder completar o diagnóstico.

Em (ALVARENGA, 1998) encontramos um sistema de diagnóstico para acidentes utilizando uma combinação de redes neurais artificiais, algoritmos genéticos e lógica nebulosa. Redes neurais do tipo AVQ (“Adaptive Vector Quantization”) serviram para gerar os protótipos (centróides) das classes representativas dos acidentes postulados. Estes centróides foram utilizados para particionar os eixos das variáveis em conjuntos nebulosos e estabelecer as zonas de influência de cada acidente.

A maneira mais simples e rápida de classificar os eventos temporais da usina consiste em representar as assinaturas dos transientes por protótipos (usualmente o centróide) e usando uma métrica, como por exemplo a distância euclidiana, comparar as distâncias dos protótipos do evento aos protótipos dos transientes postulados. Cada evento será classificado segundo a classe do transiente a cujo protótipo o protótipo do evento estiver mais próximo. Esse método de classificação, denominado de Método de Classificação Simples MCS, embora simples apresenta algumas deficiências de classificação, ver (SCHIRRU, PEREIRA e MARTINEZ, 1999) e (ALMEIDA, 2001), como o caso em que as classes são muito próximas ou os centróides são coincidentes.

Para superar as deficiências do MCS (SCHIRRU, PEREIRA e MARTINEZ, 1999) propõem o Método do Conjunto Mínimo de Centróides (CMC) que, particionando o espaço segundo o eixo do tempo, procura determinar duas ou mais subclasses que forneçam o melhor conjunto de centróides (subclasses) de modo que o MCS possa distinguir as classes e classificar os eventos corretamente.

Em (ALMEIDA, 2001), o sistema de classificação propõe um método de classificação possibilística como forma de melhorar a separação dos subconjuntos das classes e assim fornecer a resposta “não sei”, em caso de eventos não conhecidos no processo de aprendizado de classificação. Adicionalmente aprimora e utiliza o conceito de acumulação de evidências proposto em (BARTAL, LIN e UHRIG, 1995).

Em sua tese de doutorado (MOL, 2002) propõe um sistema de identificação de transientes baseado em redes neurais artificiais para identificação de transientes com capacidade de resposta não sei na identificação dinâmica de eventos desconhecidos, ou melhor, para eventos não pertencentes ao conjunto de aprendizado utilizados durante seu treinamento.

Nesse método são utilizadas duas redes sendo uma rede responsável pela identificação dinâmica, de um conjunto recente de valores de entrada, através de uma janela de tempo móvel e a segunda rede responsável por validar a identificação realizada pela primeira rede através da validação de cada variável, permitindo assim uma resposta não sei para eventos desconhecidos. Para aumentar a robustez a rede é treinada acrescentando ruído aos dados de treinamento. Nesse trabalho são consideradas 17 variáveis como conjunto mínimo de variáveis capazes de caracterizar 16 condições operacionais de uma usina PWR.

Nos modelos de minimização do particionamento do espaço o que se procura é obter o número mínimo de partições sendo cada partição representada por um protótipo, o centróide da partição. Uma vez que nada obriga a que os protótipos para classificação sejam os centróides resultantes do particionamento, e o que se procura é o menor número de partições (menor número de partições é igual a menor número de protótipos) que maximizem o número de acertos, pensou-se em deixar que um algoritmo de busca resolva o problema através de um processo de busca cega desses protótipos sendo o desempenho de cada protótipo avaliado por uma função que mede o numero de acertos para classificação de todos os acidentes para um número prefixado de protótipos para cada acidente.

Para resolver o problema de número mínimo de protótipos, usamos aqui o algoritmo de otimização por enxame de partículas para o qual fixamos diferentes valores para o número mínimo de protótipos por classe a ser procurado. Esta abordagem estabelece um modelo de solução, que corresponde ao processo de busca dos vetores de Voronoi para separação e identificação das classes, ou seja, o menor número de vetores que permitem classificar essas classes.

Os resultados obtidos mostram soluções triviais não obtida por outros métodos e foram comparados com os resultados obtidos por meio de algoritmos genéticos com base na determinação do número mínimo de partições relatados em (ALMEIDA, 2001).

5.2. Diagnóstico de Transientes em Usinas Nucleares

A operação segura e econômica de uma central nuclear envolve a monitoração e atuação de diversos sistemas constituídos por inúmeros componentes cuja falha pode provocar uma evolução adversa das condições operacionais da usina que, se não for corretamente diagnosticada, pode ter como resultado consequências não só para a central como também para o público.

A evolução das condições operacionais da usina, usualmente, apresenta um comportamento transiente com padrões temporais bem definidos (assinaturas) dessas medidas para cada situação operacional. Para operar a central, os operadores monitoram um grande número de medidas, obtidas a partir da instrumentação instalada nesses sistemas, e, em caso de um evento anormal de operação da usina, devem identificar e diagnosticar a situação da usina a partir da análise e classificação das medidas correspondentes ao evento em curso.

O grande número de instrumentos e a dinâmica de variação de cada grandeza medida associada ao evento em curso dificultam o processo de identificação e tomada de decisão por parte do operador que, quando pressionado a tomar uma decisão, reduz ainda mais a sua capacidade cognitiva aumentando assim a chance de uma tomada de decisão errada por parte do operador.

A fim de reduzir e minimizar a carga cognitiva do operador, e auxiliar o operador, sistemas de apoio à tomada de decisão, tais como sistemas de classificação de transientes, são incorporados ao sistema de operação da usina com o objetivo de aumentar o tempo disponível do operador para a tomada de uma resposta rápida e correta minimizando o risco de uma identificação errada.

5.3. Assinaturas de Acidentes

O objetivo de um sistema de diagnóstico de transientes é o de, em caso de um evento anômalo da usina, identificar dentre um conjunto de transientes postulados para uma usina, qual transiente melhor caracteriza o evento em curso. Para identificar um evento, um sistema de diagnóstico de transientes utiliza um conjunto de assinaturas de um conjunto de variáveis de operação da usina para cada evento anômalo da usina.

Em (ALVARENGA, 1998) e em (MOL, 2002) encontramos uma lista de acidentes postulados para uma usina nuclear bem como uma lista das variáveis de processo consideradas como as que mais contribuem para a caracterização desses acidentes.

Em (MOL, 2002) encontramos também uma descrição da filosofia de projeto e construção dos sistemas em níveis de segurança para assegurar o controle com segurança de uma usina nuclear bem como uma descrição da maneira como esses sistemas se comportam durante a ocorrência de um conjunto de acidentes postulados para um reator nuclear PWR e que devem fazer parte de um relatório final de análise de segurança (FSAR).

A tabela 5.1 a seguir apresenta a lista de acidentes preconizada em (ALVARENGA. 1998):

Tabela 5.1 – Acidentes Postulados para um reator PWR

Evento	Descrição
NORMAL	Condição normal de potência.
BLACKOUT	Perda de alimentação elétrica externa.
LOCA	Perda de refrigerante do sistema primário.
SGTR	Ruptura de tubos do gerador de vapor.
MFWBR	Ruptura da alimentação principal.
TRIPTUR	Desligamento da turbina
MEFWISO	Isolamento da alimentação principal e auxiliar
MFWISO	Isolamento da alimentação principal
MSTMISO	Isolamento da linha de vapor principal
STMLIBR	Ruptura da linha de vapor principal
BLACKSEM	Perda de alimentação elétrica sem desligamento do reator
MFBRWSEM	Perda de alimentação principal sem desligamento do reator
MSTMISEM	Isolamento da linha de vapor principal sem desligamento do reator
MEFWISEM	Isolamento da alimentação principal/auxiliar sem desligamento do reator
MFWISEM	Isolamento da alimentação principal sem desligamento do reator
TRIPTURSEM	Desligamento da turbina sem desligamento do reator

A tabela 5.2., a seguir, mostra a lista de variáveis de processo consideradas como necessárias e suficientes para o reconhecimento desses transientes. Todas variáveis são normalizadas no intervalo de 0 a 1.

Tabela 5.2- Lista de variáveis de acidentes

Variável	Unidade
Vazão do núcleo	%
Temperatura da perna quente	°C
Temperatura da perna fria	°C
Nível no gerador de vapor – Faixa larga	%
Nível no gerador de vapor – Faixa estreita	%
Pressão no gerador de vapor	Mpa
Vazão de água de alimentação	Kg/s

Tabela 5.2- Lista de variáveis de acidentes (continuação)

Variável	Unidade
Vazão de vapor	Kg/s
Vazão na ruptura	Kg/s
Vazão no circuito primário	Kg/s
Pressão no sistema primário	Mpa
Potência térmica	%
Potência Nuclear	%
Margem de subresfriamento	°C
Nível do pressurizador	%
Temperatura média no primário	°C

Os transientes postulados caracterizam as classes de transiente segundo a qual um determinado evento poderá ser classificado. Quando um evento de operação anormal deve ser identificado, o sistema de diagnóstico procura comparar a evolução das variáveis correspondentes às condições operacionais da usina com a assinatura de evolução das variáveis para cada um dos eventos anômalos postulados da usina e classificar o evento em curso pelo evento ao qual o evento está mais próximo usando para isso uma medida de distância.

As figuras 5.1 e 5.2, a seguir, mostram a evolução (assinatura) de cada variável para cada um dos 18 acidentes postulados considerando a usina operando a 50% e 100% da potência nominal respectivamente.

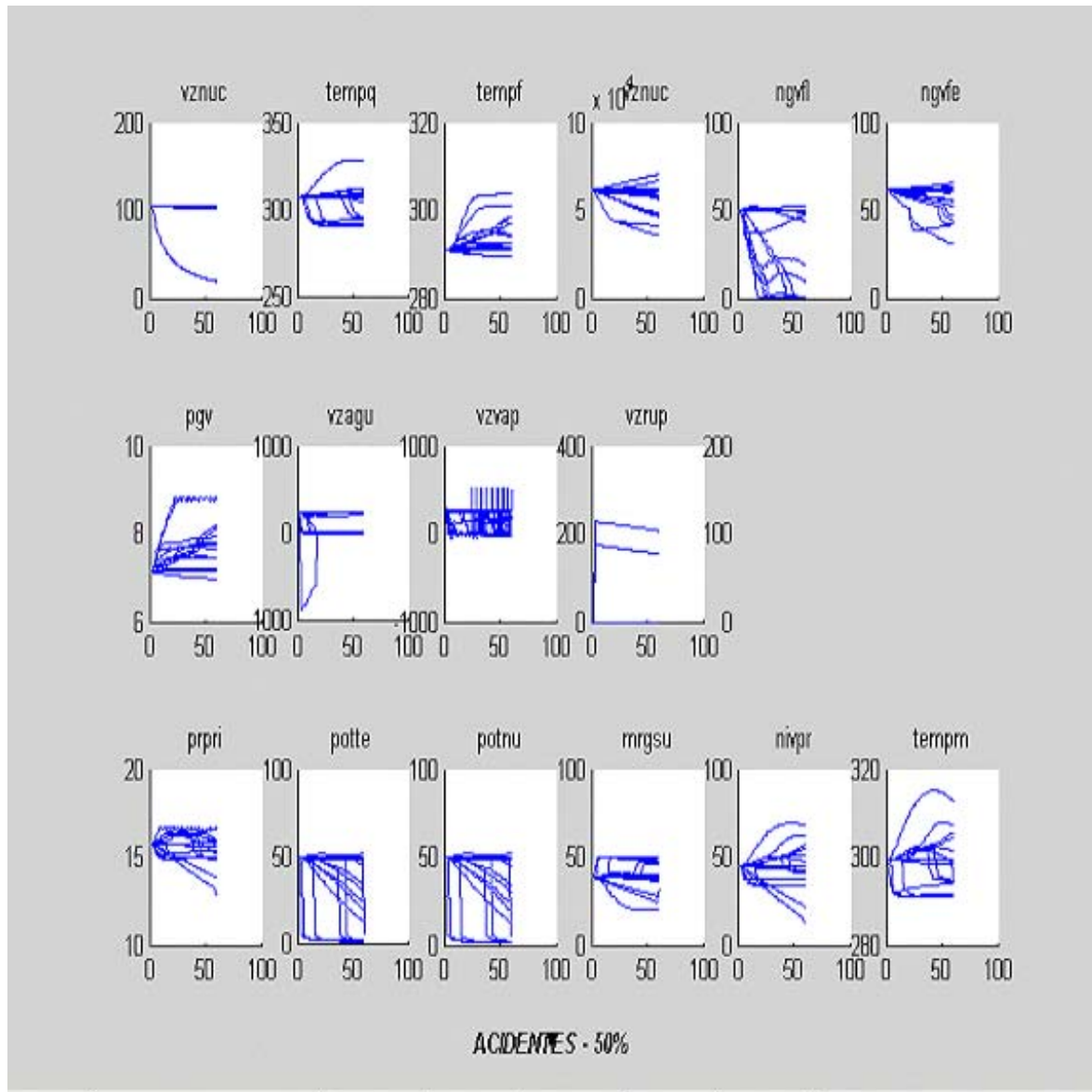


Figura 5.1 Assinaturas de acidentes a 50%

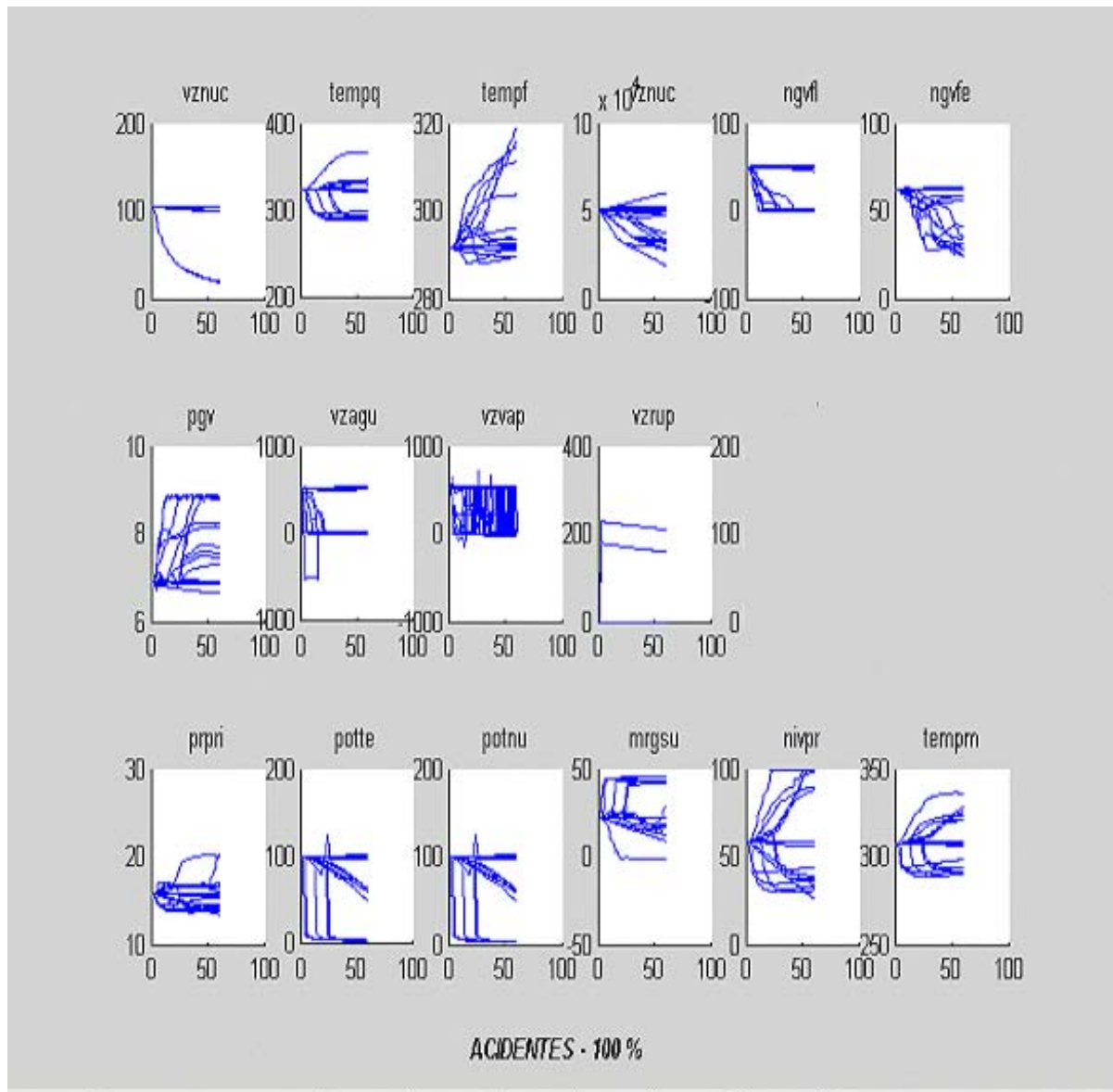


Figura 5.2 Assinaturas de acidentes a 100%

5.4. Classificação de Eventos

Em (ALMEIDA, 2001) encontramos o método de classificação possibilística (MCP) para na classificação de acidentes de uma usina nuclear. Este método mostrou-se superior aos métodos anteriormente encontrados na literatura.

Para o treinamento foram usados três dos 18 transientes postulados para uma usina nuclear operando a uma potência de 100%. Cada transiente é representado pela evolução de 17 variáveis, que por hipótese, são assumidas como necessárias e suficientes para o reconhecimento destes transientes.

Tabela 5.3. Conjunto de transientes para treinamento

1	BLACKOUT	Perda de alimentação elétrica externa
2	LOCA	Perda de refrigerante do sistema primário
3	SGTR	Ruptura de tubos do gerador de vapor

O MCP utiliza uma função objetivo que pondera favoravelmente o número de classificações corretas e penaliza o número de partições conforme equação:

$$f = k_A \cdot A - k_C \cdot C \quad \text{eq. (1)}$$

onde

f valor da função objetivo

k_A fator de ponderação do número de acertos

A número de acertos

k_C fator de penalização do número de partições = 1

C número de partições

Na fase de treinamento são gerados centróides representativos das subclasses dos transientes escolhidos; os transientes, com suas respectivas classificações segundo uma função objetivo que calcula o desempenho (“fitness”) de cada solução candidata levando em conta o número de classificações corretas e o número de partições usadas na classificação. Os protótipos assim obtidos são utilizados posteriormente na fase de operação para classificação instantânea das amostras do evento a ser classificado.

O método de classificação MCP, com resultados superiores ao CMC, obteve NP = 6 partições, com um índice de acertos de 98,4% (A=180 acertos em 183 classificações) no reconhecimento dos padrões. A tabela 5.4 a seguir mostra o particionamento obtido pelo método MCP:

Tabela 5.4. Método MCP com 6 partições

NP	A	Particionamento GA
6	180	00011000001111111111100000000000000000000000000001111111111111111

De acordo com a fitness estabelecida para o método MCP, ver equação 1, a solução trivial deveria ser (a menos da solução de número de partições igual a 0 que não tem sentido) numero de partições igual a 1.

Isto sugere que se utilizássemos o particionamento definido na tabela 5.5 deveríamos obter o a melhor fitness. No entanto, como pode ser visto, o método de classificação possibilística apresenta como resultado um número de acertos 176 é inferior aos 180 acertos obtidos com 6 centróides.

Tabela 5.5. Método MCP com 1 partição

NP	A	Particionamento GA
1	176	00

5.5. Método do Número Mínimo de Protótipos

Não existe na definição de um protótipo nada que obrigue o mesmo a ser o centróide dos pontos da classe que o mesmo representa. A não ser o fato de que o mesmo deve maximizar o acerto das classificações, o protótipo ou protótipos que representam uma classe podem, em tese ser quaisquer vetores que representem bem a classe, isto é permitam a classificação correta da classe.

Uma vez que o que se quer é determinar o menor número de partições (menor número de partições é igual a menor número de protótipos) que maximizem o número de acertos, pensou-se em deixar que um algoritmo de busca resolva o problema através de um processo de busca cega desses protótipos.

Cada protótipo consiste de um vetor de dimensão igual ao número de variáveis de acidente. Assim, por meio de uma função objetivo, em que a adequação é medida pelo numero de acertos para classificação de todos os acidentes, resolvemos utilizar um algoritmo de busca que procure uma solução que maximize o número de acertos para um número prefixado NPC de protótipos para cada acidente.

5.6. Configuração do PSO

Para resolver o problema de número mínimo de protótipos, usamos aqui o algoritmo de otimização por enxame de partículas para o qual fixamos diferentes valores para o número mínimo de protótipos por classe NPC a ser procurado. Cada indivíduo do algoritmo PSO consiste de NPC vetores de dimensão igual a NV . Assim, a dimensão de cada indivíduo do algoritmo PSO é então igual a $DIM = NPC * NV$ onde:

DIM dimensão de cada indivíduo do algoritmo PSO

NPC número de protótipos por classe

NV número de variáveis de acidente

Uma vez que todas as variáveis, ver tabela 5.1, são normalizadas no intervalo $[0,1]$, os limites das variáveis usadas para limites das dimensões dos indivíduos do PSO também são definidos nesse intervalo.

A tabela 5.6, a seguir, mostra os principais parâmetros de configuração do algoritmo PSO.

Tabela 5.6. Parâmetros do algoritmo PSO

Parâmetro	Descrição	Valor
NPC	Número de protótipos por classe	1,2 e 3
NV	Número de variáveis de acidente	17
DIM	Dimensão do indivíduo	$NPC * NV$
$X Min$	Valores mínimos de cada dimensão do problema	0.0
$X Max$	Valores máximo de cada dimensão do problema	1.0

Tabela 5.6. Parâmetros do algoritmo PSO (continuação)

<i>MAX ITER</i>	Número máximo de iterações	200
<i>MAX POP</i>	Numero de indivíduos do enxame	2000
<i>TIPO</i>	1=Canônico 2= Individual 3= Social 4=Determinístico	1
<i>a</i>	Fator de inércia	0.729
<i>b1</i>	Coeficiente de força de atração individual	1.494
<i>b2</i>	Coeficiente de força de atração social	1.494
<i>c e d</i>	Coeficientes	1.0

5.7. Resultados

Os resultados obtidos pelo algoritmo PSO acima configurado, estão mostrados na tabela 5.7 indicando, para cada número de protótipos NPC prefixado, o número de acertos obtido

Tabela 5.7. Acertos para diferentes números protótipos

NPC	Acertos
3	183 = 100%
2	183 = 100%
1	181 = 98.9%

O número de instantes de tempo após o início do acidente é igual a 60. Como pode ser visto, o número de acertos foi o maior possível (100%) para NPC = 2 e NPC = 3. Para NPC = 1 o número de acertos foi igual a 181 (98.9%).

A figura 5.3 a seguir mostra a evolução da fitness para NPC = 3.

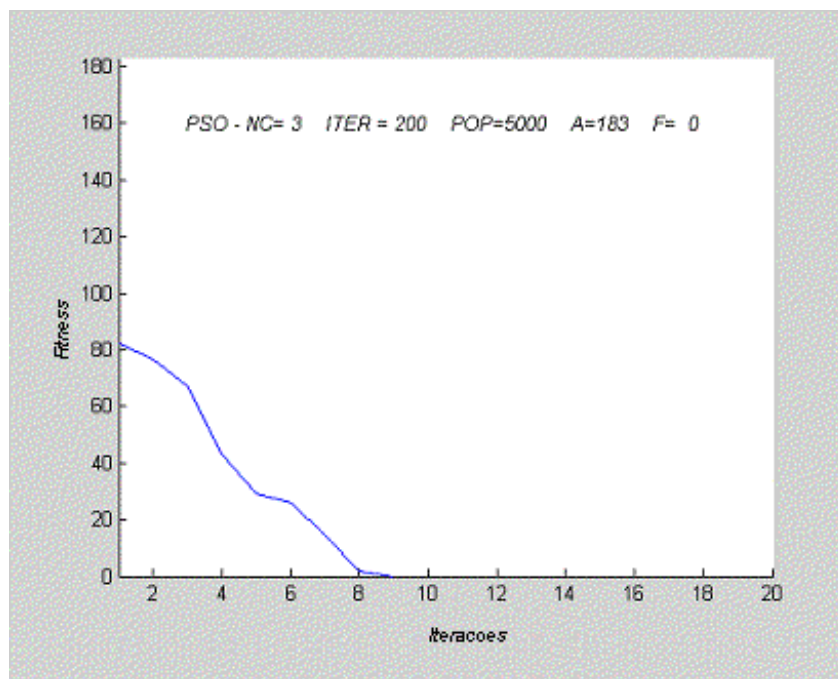


Figura 5.3. Convergência para NPC=3

A figura 5.4 a seguir mostra a evolução da fitness para NPC = 2.

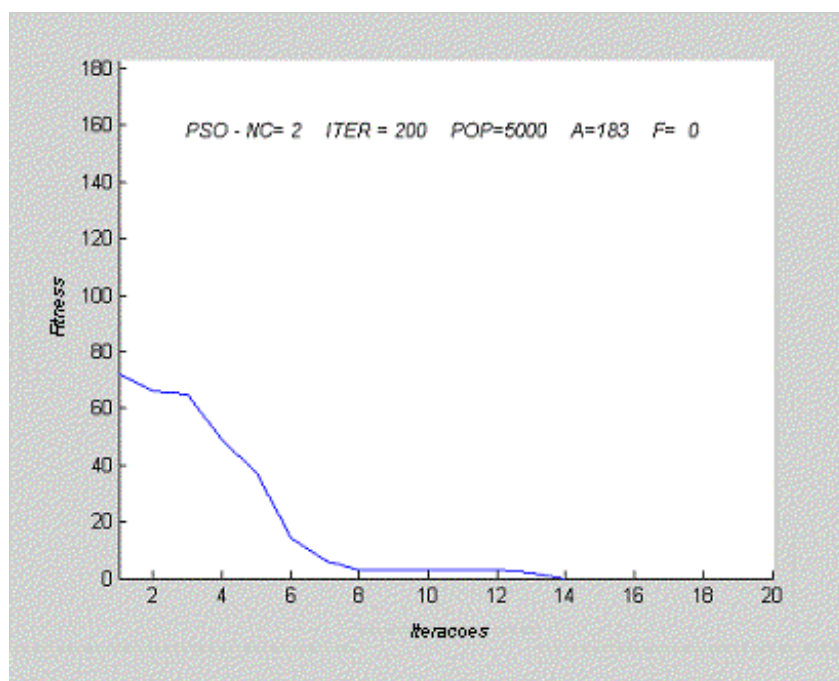


Figura 5.4. Convergência para NPC=2

A figura 5.5 a seguir mostra a evolução da fitness para NPC = 1.

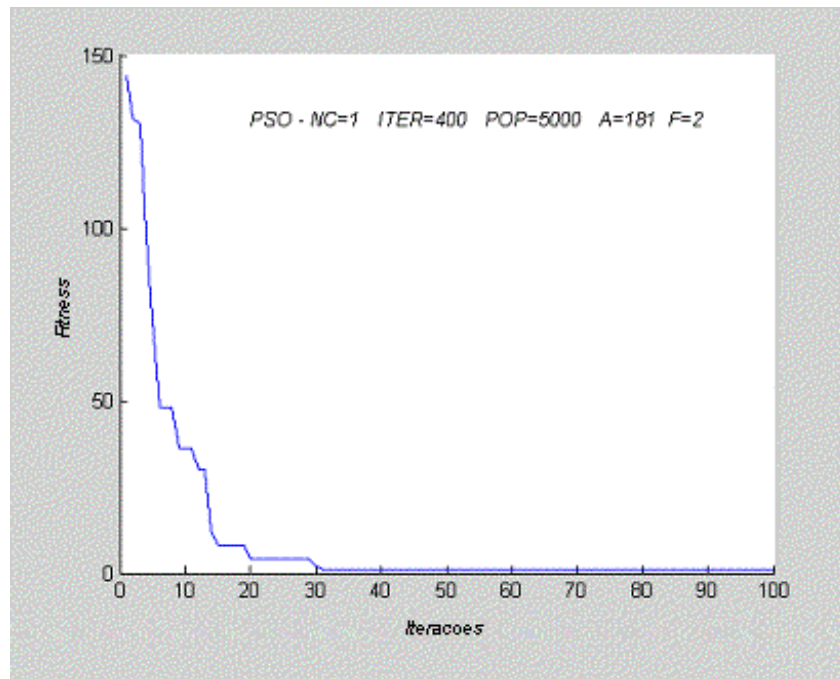


Figura 5.5. Convergência para NPC = 1

As figuras 5.5 e 5.6 mostram uma convergência mais lenta, o que se justifica por serem problemas cuja solução é mais difícil de ser encontrada.

As figuras 5.7, 5.8 e 5.9, a seguir, mostram, para uma das dimensões, a localização dos protótipos obtidos com o PSO para diferentes números de protótipos por classe (NPC).

A figura 5.6 a seguir mostra os protótipos obtidos com o PSO correspondentes a um número de protótipos por classe NPC = 3 e número de acertos igual a A=183.

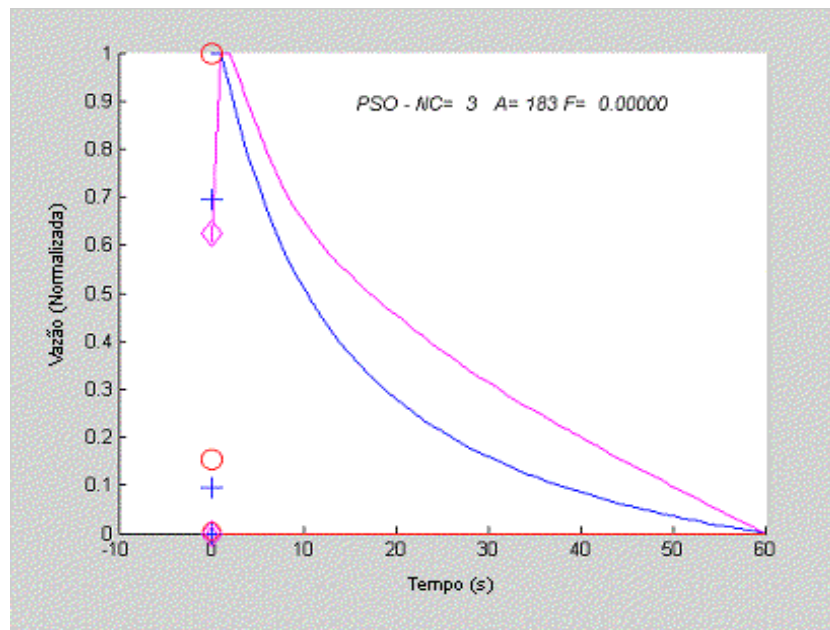


Figura 5.6. Protótipos para NPC = 3

A figura 5.7 a seguir mostra os protótipos obtidos com o PSO correspondentes a um número de protótipos por classe NPC = 2, para o qual o número de acertos foi igual a A=183.

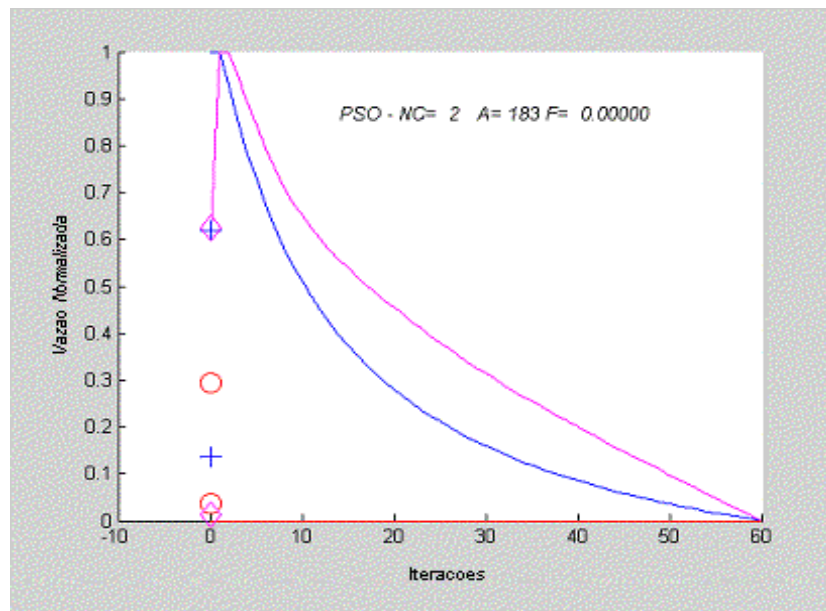


Figura 5.7. Protótipos para NPC = 2

A figura 5.8 a seguir mostra os protótipos obtidos com o PSO correspondentes a

De fato avaliando-se esse particionamento pelo método de classificação possibilística, obtém-se o número máximo de acertos $A = 183$, o que mostra a consistência da solução proposta com o método de mínimos protótipos.

Analisando os resultados verificamos que para $NPC = 3$ e $NPC = 2$ obtemos 183 (100% de) acertos o que sugere que os acidentes podem ser identificados a partir do instante $t = 0$. do início do transiente.

Para $NPC = 1$ obtemos 181 acertos de um total de 183 possíveis acertos, o que equivale a 98.9% de acertos, ou seja, próximo do máximo possível de acertos, que na realidade caracterizaria os vetores de Voronoi das classes de acidentes.

CAPÍTULO 6

CONCLUSÕES E RECOMENDAÇÕES

6.1. Conclusões

Nesta pesquisa exploramos a capacidade do algoritmo de otimização por enxame de partículas, mostrando a viabilidade do algoritmo como ferramenta de otimização na busca de soluções em espaços de busca multimodais complexos e de alta dimensão, como os problemas encontrados na área nuclear.

Inicialmente devemos ressaltar a simplicidade do algoritmo cuja implementação é facilmente traduzida a partir das equações que definem o modelo canônico do algoritmo. Por outro lado, devido a seu equacionamento, torna-se natural a representação de modelos de funções reais em espaços reais de qualquer dimensão.

Devido à naturalidade e facilidade de representação de espaços reais de qualquer dimensão, bem como a sua característica de busca do ótimo global, o algoritmo de otimização por enxame de partículas foi aplicado com sucesso na solução de três problemas na área nuclear e mostrou-se eficaz na busca de soluções em espaços de busca multimodais de alta dimensão sem necessidade de conhecimento prévio sobre a complexidade dos espaços de busca envolvidos.

Quando aplicado no projeto do núcleo de reatores PWR, o algoritmo de enxame de partículas resolve com eficiência a otimização dos parâmetros de projeto do núcleo.

A aplicabilidade do método a sistemas contínuo/discretos fica demonstrada pela aplicação do algoritmo ao problema de otimização do projeto do núcleo do reator em que duas das variáveis, material do combustível e material do revestimento, assumem valores discretos e as demais podem assumir valores dentro de uma faixa de valores reais. Nesse caso, o algoritmo PSO adota o artifício de só utilizar os valores inteiros das variáveis.

Quando aplicado como módulo de otimização de um controlador preditivo, o algoritmo PSO não depende do conhecimento do modelo de representação da planta nem a obtenção do gradiente de variáveis da planta o que permite que diferentes modelos de rede possam ser utilizados para representar a planta sem que seja necessário, por isso, reprogramar o otimizador.

O otimizador de um controlador preditivo por algoritmos baseados em gradiente requer a otimização de uma função objetivo cuja maximização ou minimização implica na obtenção de derivadas de primeira e de segunda ordem de variáveis da planta. Assim, se no controlador preditivo o modelo de rede utilizado para representar a planta tiver que ser modificado, o algoritmo de otimização deverá ser modificado também. A aplicação do algoritmo PSO, no entanto, requer, da mesma forma que outros métodos baseados em população, um tempo de execução maior do que os algoritmos tradicionais. Assim, sua utilização só é possível nos casos em que o passo de atualização do controle em tempo real é maior do que o tempo de execução da otimização pelo algoritmo.

Quando aplicado ao problema de classificação de acidentes, o algoritmo PSO foi utilizado para implementar um modelo de solução que procura um número prefixado de protótipos por classe sem impor qualquer restrição sobre o que esses protótipos representam a não ser o de maximizar o número de classificações.

Uma vez que nada obriga a que os protótipos para classificação sejam os centróides resultantes do particionamento do espaço de busca, o algoritmo foi utilizado na busca dos protótipos que maximizam o número de acertos para classificação de todos os acidentes. Os resultados obtidos com enxame de partículas mostram soluções triviais não obtidas por outros métodos usando algoritmos genéticos com base na determinação do número mínimo de partições em que se procura obter o número mínimo de partições sendo cada partição representada por um protótipo, o centróide da partição. Esta abordagem corresponde ao processo de busca dos vetores de Voronoi que representam as classes e maximizam a separação entre as mesmas.

Com base nos resultados obtidos fica evidenciada a viabilidade da aplicação do método de otimização baseado em enxame de partículas a problemas de otimização complexos de alta dimensionalidade sem a necessidade de conhecimento a priori a respeito do espaço de busca, o que permite a otimização de problemas sem preocupação com aspectos de continuidade e existência de limites e/ou derivadas. Uma vez que o método não depende da existência de limites e/ou derivadas o algoritmo pode ser aplicado a qualquer tipo de problema para o qual se possa representar cada solução candidata por um vetor de dimensão n e essas soluções possam ser avaliadas por uma função objetivo.

Finalmente, verificamos que para os problemas apresentados, o PSO apresentou soluções com resultados compatíveis, e até superiores, aos resultados obtidos nas soluções da literatura usada como referência. Em particular, na solução do problema de identificação de acidentes, permitiu uma solução que se aproxima da solução ideal, um protótipo por classe, ou seja, os vetores de Voronoi, das classes para a classificação de acidentes.

6.2. Sugestões para Trabalhos Futuros

Apesar dos resultados obtidos na identificação de transientes, o foco do trabalho se concentrou na minimização do número de protótipos usando três acidentes.

Uma vez que o número mínimo de protótipos se aproxima dos vetores de Voronoi e uma vez que os mesmos estão localizados na origem, isto significa que os acidentes podem ser identificados a partir do instante de tempo de início do transiente. Por essa razão deve ser possível obter evidências que possam ser transformadas em regras para definir as condições que caracterizam os acidentes.

A fim de estender os resultados recomenda-se aplicar o método a um número maior de transientes e análise da evolução do número mínimo de vetores de referência de classificação a fim de estudar a existência de um limite mínimo para o conjunto de vetores de classificação no próprio espaço dos sinais dos transientes.

Por outro lado pode ser necessário estabelecer um mecanismo de separação que permita introduzir o conceito de classificação “não sei” para classificar eventos não conhecidos durante o treinamento do sistema de classificação bem como a identificação de acidentes com introdução de ruído e determinar a sensibilidade da classificação em presença de ruído.

No caso do controle preditivo, uma vez que o objetivo desta tese é o de verificar a viabilidade do otimizador preditivo, e uma vez que a obtenção de um modelo suficientemente preciso do preditor com a rede neural pode ter outros desdobramentos que poderiam se desviar o objetivo do trabalho, deixamos como uma sugestão de tarefa o problema de obtenção de uma rede suficientemente precisa para o preditor do controlador.

A viabilidade de aplicação do PSO se concentrou na aplicação do mesmo para busca de soluções em espaços reais. No entanto existem classes de problemas cuja solução requer uma modificação do PSO, como no caso de problemas combinatórios. Assim, a fim de identificar as limitações do método e obter soluções para superar eventuais deficiências, o método deve ser aplicado na solução dessas classes de problemas.

ANEXO 1

ALGORITMOS DE OTIMIZAÇÃO BASEADOS EM GRADIENTE

A1.1. Otimização Numérica

Os métodos de otimização numérica, (LUENBERGER, 1986) e (KAYLATH, 1980), encontram inúmeras aplicações em diversas aplicações envolvendo sistemas de inteligência artificial como por exemplo, o treinamento de redes neurais artificiais, a otimização de sistemas de controle adaptativo e otimização em sistemas de classificação. O problema de otimização consiste em determinar o valor de um ponto (vetor) $\mathbf{x} = [x_1, x_2, \dots, x_n]$ de um espaço dimensional \mathfrak{R}^n que optimize (minimize ou maximize) um benefício medido por uma função objetivo $J: \mathfrak{R}^n \rightarrow \mathfrak{R}$ muitas vezes sujeito a restrições de igualdade e/ou desigualdade sobre \mathbf{x} do tipo $\mathbf{x} \in S$ onde $S \subset \mathfrak{R}^n$:

$$\min J(\mathbf{x}) \quad \text{eq. (1)}$$

$$\mathbf{x} = [x_1, x_2, \dots, x_n] \in R^n \quad \text{eq. (2)}$$

sujeito a restrições do tipo

$$\leq = [g_1, g_2, \dots, g_{ng}] \quad \text{eq. (3)}$$

$$= [h_1, h_2, \dots, h_{nh}] \quad \text{eq. (4)}$$

O problema de otimização consiste em determinar o valor de $\mathbf{x} \in R^n$ que

produza um valor máximo (ou mínimo) para a função objetivo $J()$ e que atenda as condições de desigualdade definidas pelas funções f_1 e f_2 .

Os métodos de otimização analíticos necessitam que a função a ser otimizada seja explicitamente conhecida e derivável, ou que possa ser aproximada por alguma função derivável. Para encontrar o ponto de máximo (ou de mínimo) no espaço de busca, devem ser resolvidas as equações que surgem quando se iguala a zero o gradiente da própria função, ou da função aproximante utilizada.

A1.2. Métodos de Busca por Enumeração

Os métodos de otimização de funções baseados em busca por enumeração procuram gerar e testar pontos do espaço em busca dos pontos que otimizem a função desejada. O método de busca uniforme divide cada direção do espaço de maneira uniforme e gera todos os pontos do espaço correspondentes ao produto cartesiano dos conjuntos de pontos em cada direção.

Os métodos de bisseção (método de busca binária e método da seção áurea) são uma alternativa que procura acelerar o processo de busca e dividem cada direção do espaço em duas seções e procuram estimar em que seção se encontra o ponto de mínimo. O processo é repetido a cada passo dividindo-se sucessivamente a seção em que se encontra o ponto de mínimo.

De qualquer forma, a enumeração do espaço requer conhecimento prévio do espaço de busca e em espaços de busca muito grandes, estes métodos não são eficazes.

A13. Métodos de Busca Baseados no Gradiente

Os métodos de máxima declividade, baseados no gradiente podem ser usados nos casos em que se pode determinar o gradiente da função objetivo em torno de cada ponto.

O método gradiente descendente é um método da máxima declividade que parte de um ponto inicial e busca o ponto de condição ótima como o ponto em que o gradiente é nulo. A cada passo do algoritmo, o gradiente é calculado e se for nulo, ou bastante próximo de zero, então o algoritmo termina. Senão, o ponto é movido na direção contrária ao vetor gradiente. O procedimento continua até que a condição de gradiente nulo ou próximo de nulo seja obtida:

O método de Newton utiliza uma aproximação quadrática do hessiano para acelerar a velocidade da convergência. O método utiliza uma expansão do índice de performance em série de Taylor até segunda ordem em torno de cada ponto x e requer o cálculo de derivadas de segunda ordem para cálculo do hessiano e inversão da matriz hessiano.

A1.3.1. Método Gradiente Descendente

- - Estabelecer uma tolerância $\varepsilon > 0$
- - Estabelecer uma tolerância $\alpha \in (0..1)$
- - Iniciar a partir de um ponto $x = x_0$
- Se $\nabla J(x) \leq \varepsilon$ então *exit*
- - $v = -\alpha \nabla J(x)$
- - $x = x + v$

A1.3.2. Método de Newton

$$J(x+h) = J(x) + [\nabla J(x)]^T h + (1/2) h^T H(x) h + o(h^3)$$

O hessiano é a matriz $H(x) = [H_{i,j}]$ onde cada elemento $H_{i,j} = \frac{\partial^2 J(w)}{\partial x_i \partial x_j}$

O mínimo $J(x+h)$ se dá em $h = -H^{-1}(x)\nabla J(x)$

A1.3.3. Método de Newton

- Estabelecer uma tolerância $\varepsilon > 0$
- Iniciar a partir de um ponto $x = x_0$
- Calcular $\nabla J(x)$ e $\nabla^2 J(x)$
- Se $\|\nabla J(x)\| \leq \varepsilon$ então $\rightarrow exit$
- $v = -\nabla^2 J(x_0)^{-1} \nabla J(x)$
- $x = x + v$

A1.3.4. Método de Newton - Funções Convexas

- $\nabla J(x) = 0$. e $\nabla^2 J(x) = pos. def.$
- $\nabla J(x) = \nabla J(x_0) + \nabla^2 J(x)(x - x_0) = 0$
- $v = x_0 - \nabla^2 J(x_0)^{-1} \nabla J(x)$
- $x = x + v$

A1.3.5. Método de Newton com Backtracking

- Estabelecer uma tolerância $\varepsilon > 0$
- Estabelecer $\alpha \in (0,1/2)$
- Iniciar a partir de um ponto $x = x_0$
- Calcular $\nabla J(x)$ e $\nabla^2 J(x)$
- Se $\|\nabla J(x)\| \leq \varepsilon$ então $\rightarrow exit$
- Se $J(x+tv) > J(x) + \alpha t \nabla J(x)^T v$ então $t = t/2$
- $v = -\nabla^2 J(x_0)^{-1} \nabla J(x)$
- $x = x + tv$

A1.3.6. Método de Newton - Funções Não Convexas

- Estabelecer uma tolerância $\varepsilon > 0$
- Estabelecer $\alpha \in (0,1/2)$
- Iniciar a partir de um ponto $x = x_0$
- Calcular $\nabla J(x)$ e $\nabla^2 J(x)$
- Se $\|\nabla J(x)\| \leq \varepsilon$ então $\rightarrow exit$
- Se $\nabla^2 J(x) = pos\ def$
 - então $v = -\nabla^2 J(x)^{-1} \nabla J(x)$
 - senão $v = -\nabla J(x)$
- Se $J(x+tv) > J(x) + \alpha t \nabla J(x)^T v$ então $t = t/2$
- $x = x + tv$

A1.3.7. Mínimos Quadrados Não Lineares

- $J(x) = \sum_{i=1}^m r_i(x)^2 = \sum_{i=1}^m (a_i^T x - b)^2$
- $J(x) = \sum_{i=1}^m r_{LIN,i}(x)^2 = \sum_{i=1}^m (r_i(x_0) + \nabla r_i(x)^T (x - x_0))^2$
- $r_i(x) \approx r_{LIN,i}(x) = r_i(x_0) + \nabla r_i(x)^T (x - x_0)$
- $x = x_0 - \nabla^2 J(x_0)^{-1} \nabla J(x)$
- $\nabla J(x) = \nabla J(x_0) + \nabla^2 J(x)(x - x_0) = 0$
- $x = x + v$

A1.3.7.1. Método de Newton

- $\nabla J(x) = 2 \sum_{i=1}^m r_i(x) \nabla r_i(x)$
- $\nabla^2 J(x) = 2 \sum_{i=1}^m (r_i(x) \nabla^2 r_i(x) + \nabla r_i(x)^T \nabla r_i(x))$

A1.3.7.2 .Método de Gauss-Newton

- $J(x) = \sum_{i=1}^m r_{LIN,i}(x)^2 = \|Ax - b\|^2$

- $A = \begin{bmatrix} \nabla r_1(x)^T \\ \nabla r_2(x)^T \\ \dots \\ \nabla r_m(x)^T \end{bmatrix}$, $b = \begin{bmatrix} \nabla r_1(x)^T x - r_1(x) \\ \nabla r_2(x)^T x - r_2(x) \\ \dots \\ \nabla r_m(x)^T x - r_m(x) \end{bmatrix} = Ax - r$
- $\nabla J(x) = 2 \sum_{i=1}^m r_i(x) \nabla r_i(x) = 2A^T r$

A1.3.7.3. Método de Gauss-Newton

- Estabelecer uma tolerância $\varepsilon > 0$
- Iniciar a partir de um ponto $x = x_0$
- Avaliar
 - $r_i(x)$ e $\nabla r_i(x)$ $i = 1, 2, \dots, m$

- $r = \begin{bmatrix} r_1(x) \\ r_2(x) \\ \dots \\ r_m(x) \end{bmatrix}$, $A = \begin{bmatrix} \nabla r_1(x)^T \\ \nabla r_2(x)^T \\ \dots \\ \nabla r_m(x)^T \end{bmatrix}$, $b = Ax - r$

- Se $\|\nabla J(x)\| = \|2A^T r\| \leq \varepsilon$ então $\rightarrow exit$
- $x = \arg(\min(\|Ax - b\|^2)) = (A^T A)^{-1} A^T b$

A1.3.7.4. Método de Gauss-Newton com Backtracking

- Estabelecer uma tolerância $\varepsilon > 0$
- Estabelecer $\alpha \in (0, 1/2)$
- Iniciar a partir de um ponto $x = x_0$
- Avaliar

- $r_i(x)$ e $\nabla r_i(x)$ $i = 1, 2, \dots, m$
- $r = [r_i(x)]$, $A = [\nabla r_i(x)^T]$
- Se $\|\nabla J(x)\| = \|2A^T r\| \leq \varepsilon$ então $\rightarrow exit$
- $v = \arg(\min(\|Ax - b\|^2)) = (A^T A)^{-1} A^T b$
- Se $\sum_{i=1}^m r_i(x + tv)^2 > \|r\|^2 + \alpha t (2A^T A v)$ então $t = t/2$
- $x = x + tv$

ANEXO 2

OSCILAÇÃO DA POTÊNCIA AXIAL DE UM REATOR

A2.1 Modelo de oscilação axial da potência do reator

O modelo de reator considerado consiste em um modelo a “dois pontos” sendo o reator dividido em duas seções idênticas no sentido axial (eixo z), ver figura A2.1. O modelo é descrito por equações não lineares de balanço da produção de xenônio e de iodo e por meio da equação de difusão de neutros a uma dimensão e um grupo de energia.

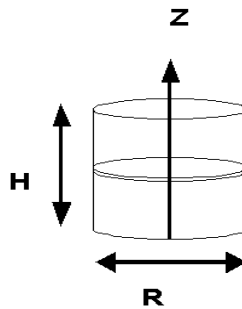


Figura A2.1 – Modelo de reator a 2 pontos

As oscilações axiais de xenônio, de iodo e de potência são funções do tempo e da posição axial.

$$\phi(z,t) = \phi_0 \psi(z,t) \quad \text{eq. (1)}$$

$$X(z,t) = \phi_0 x(z,t) \quad \text{eq. (2)}$$

$$I(z,t) = I_0 y(z,t) \quad \text{eq. (3)}$$

onde

ϕ	Fluxo
X	Concentração de xenônio
I	Concentração de iodo

A equação de difusão unidimensional a um grupo é dada por :

$$D \frac{\partial^2 \psi(z,t)}{\partial z^2} + \left[\frac{v \Sigma_f}{k} - \Sigma_a(z) \right] \psi(z,t) - \sigma_X X_0 x(z,t) \psi(z,t) - \alpha_F \bar{\Sigma}_a \phi_0 \psi^2(z,t) = 0$$

eq. (4)

As equações de balanço de xenônio e de iodo são definidas por:

$$\frac{\partial y(z,t)}{\partial t} = \left(\gamma_I \Sigma_f \frac{\phi_0}{I_0} \right) \psi(z,t) - \lambda_I y(z,t)$$

eq. (5)

$$\frac{\partial y(z,t)}{\partial t} = \left(\gamma_I \Sigma_f \frac{\phi_0}{I_0} \right) \psi(z,t) - \left(\lambda_I \frac{I_0}{X_0} \right) y(z,t) - \lambda_X x(z,t) - \sigma_X \phi_0 x(z,t) \psi(z,t)$$

eq. (6)

As equações para as distribuições de fluxo, xenônio e iodo são obtidas a partir de séries harmônicas espaciais a 2 termos:

$$\psi(z,t) = \cos(\pi z / H) + A(t) \sin(2\pi z / H)$$

eq. (7)

$$x(z,t) = \cos(\pi z / H) + B(t) \sin(2\pi z / H)$$

eq. (8)

$$y(z,t) = \cos(\pi z / H) + C(t) \sin(2\pi z / H)$$

eq. (9)

A seção de choque nas metades inferior e superior é dada por:

$$\Sigma_a(z) = \begin{cases} \Sigma_{a1}, & z \in (-\frac{H}{2}, 0) \\ \Sigma_{a2}, & z \in (0, \frac{H}{2}) \end{cases} \quad \text{eq. (10)}$$

Da equação da amplitude A(t) é obtida a solução de:

$$-\beta_2 A^2 + 2(\beta_1 - \beta_3)A + \beta_2 = 0 \quad \text{eq. (11)}$$

$$\beta_1 = \frac{1}{\Sigma_f} \left[4D \left(\frac{\pi}{H} \right)^2 + \frac{1}{2}(\Sigma_{a1} + \Sigma_{a2}) + \frac{32}{15\pi}(\sigma_X X_0 + 3\alpha_F \phi_0 \bar{\Sigma}_a) \right] \quad \text{eq. (12)}$$

$$\beta_2 = \frac{1}{\Sigma_f} \left[\frac{8}{3\pi}(-\Sigma_{a1} + \Sigma_{a2}) + \frac{64}{15\pi} \sigma_X X_0 B(t) \right] \quad \text{eq. (13)}$$

$$\beta_3 = \frac{1}{\Sigma_f} \left[D \left(\frac{\pi}{H} \right)^2 + \frac{1}{2}(\Sigma_{a1} + \Sigma_{a2}) + \frac{8}{3\pi}(\sigma_X X_0 + \alpha_F \phi_0 \bar{\Sigma}_a) \right] \quad \text{eq. (14)}$$

$$I_0 = \frac{\gamma_I \Sigma_f \phi_0}{\lambda_I} \quad \text{eq. (15)}$$

$$I_0 = \frac{(\gamma_I + \gamma_X) \Sigma_f \phi_0}{\lambda_X + \frac{\pi}{4} \sigma_X \phi_0} \quad \text{eq. (16)}$$

As equações das Amplitudes $B(t)$ e $C(t)$

$$\frac{dC(t)}{dt} = \left(\gamma_I \Sigma_f \frac{\phi_0}{I_0} \right) A(t) - \lambda_I C(t) \quad \text{eq. (17)}$$

$$\frac{dB(t)}{dt} = \left(\gamma_I \Sigma_f \frac{\phi_0}{X_0} \right) A(t) + \left(\lambda_I \frac{I_0}{X_0} \right) C(t) - \lambda_X B(t) - \frac{2}{3} \sigma_X \phi_0 [A(t) + B(t)] - \frac{\pi}{4} \sigma_X \phi_0 [A(t)B(t)] \quad \text{eq. (18)}$$

A média espacial do fluxo nas metades inferior e superior é dada, respectivamente por:

$$\bar{\psi}_1(t) = \frac{\int_{-H/20}^0 \psi(z,t) dz}{\int_{-H/2}^0 dz} = \frac{2}{\pi} [1 - A(t)] \quad \text{eq. (19)}$$

$$\bar{\psi}_2(t) = \frac{\int_0^{H/2} \psi(z,t) dz}{\int_0^{H/2} dz} = \frac{2}{\pi} [1 + A(t)] \quad \text{eq. (20)}$$

A média espacial das concentrações de xenônio $x(t)$ e iodo $y(t)$ nas metades inferior e superior é dada por:

$$\bar{x}_1(t) = \frac{2}{\pi} [1 - B(t)] \quad \text{e} \quad \bar{x}_2(t) = \frac{2}{\pi} [1 + B(t)] \quad \text{eq. (21)}$$

$$\bar{y}_1(t) = \frac{2}{\pi} [1 - C(t)] \quad \text{e} \quad \bar{y}_2(t) = \frac{2}{\pi} [1 - C(t)] \quad \text{eq. (22)}$$

A2.2. Simulação da oscilação axial da potência

A simulação do reator foi realizada, considerando o reator a 100% e a concentração de xenônio em regime permanente, através de uma perturbação da seção de choque provocando assim oscilações de fluxo xenônio e iodo.

Tabela A2.1 - Parâmetros do reator

Parâmetro	Valor
ϕ_0 [$\text{cm}^{-2} * \text{sec}^{-1}$]	2.1×10^{13}
σ_x [cm^2]	2.6×10^{-18}
α_f [$\text{cm}^2 * \text{sec}$]	3.4×10^{-16}
γ_I	0.061
γ_X	0.003
λ_I [sec^{-1}]	2.87×10^{-5}
λ_X [sec^{-1}]	2.09×10^{-5}
D [cm]	0.375
H [cm]	365.81
Σ_f [cm^{-1}]	0.651
$\nu\Sigma_f$ [cm^{-1}]	1.56
$\bar{\Sigma}_a$ [cm^{-1}]	1.523

Os parâmetros do reator estão mostrados na tabela A2.1. onde:

ϕ_0	fluxo nominal do reator
σ_x	seção de choque microscópica do xenônio
α_f	constante de decaimento de fissão
γ_I	fração de iodo
γ_X	fração de xenônio
λ_I	constante de decaimento do iodo 135
λ_X	constante de decaimento do xenônio 135

D	diâmetro do reator
H	altura do reator
Σ_f	seção de choque macroscópica de fissão
$\nu\Sigma_f$	seção de choque macroscópica de fissão
$\bar{\Sigma}_a$	seção de choque macroscópica de absorção

Como pode ser visto pela figura A2.2, na ausência de uma ação de controle, a amplitude das oscilações aumenta com o tempo.

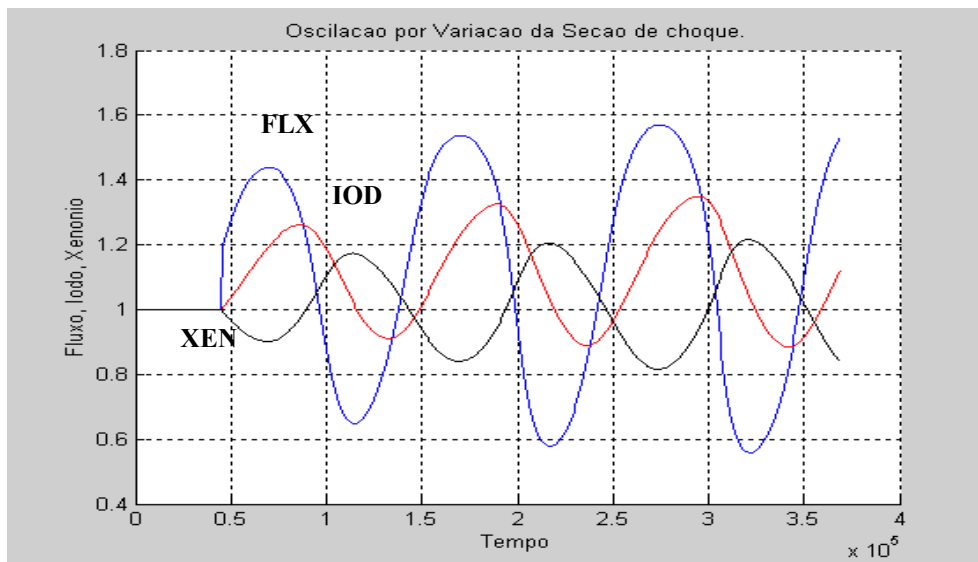


Figura .A2.2 – Oscilação de potência do reator

As oscilações de fluxo, xenônio e iodo podem ser expressas por meio das diferenças axiais de fluxo, iodo e xenônio definidas por:

$$AOP = \frac{(P1 - P2)}{(P1 + P2)} \quad \text{eq. (23)}$$

$$AOI = \frac{(I1 - I2)}{(I1 + I2)} \quad \text{eq. (24)}$$

$$AOX = \frac{(I + X)\Sigma_f(X_1 - X_2)}{(I + X) - 2\sigma_x(X_1 - X_2)} \quad \text{eq. (25)}$$

Na condição de equilíbrio, em regime permanente, devemos ter:

$$AOP = AOI - AOX \quad \text{eq. (26)}$$

No espaço de estados definido por: $daoix = AOI - AOX$ e $daopx = AOP - AOX$, ver figura A2.3, o ponto de equilíbrio está no centro. A oscilação do reator pode ser vista graficamente onde se pode ver o comportamento divergente a partir do centro do espaço de estado.

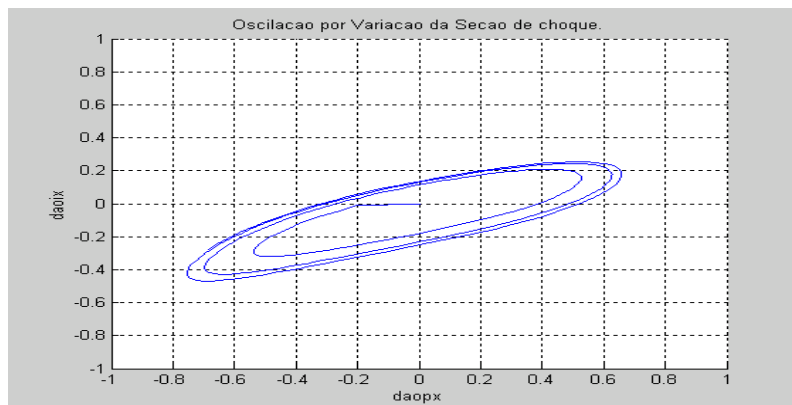


Figura A2.3 – Diferenças axiais

ANEXO 3

REDES NEURAIIS ARTIFICIAIS

A3.1 – Neurônios Artificiais

As redes neurais artificiais são modelos de computação inspirados no funcionamento do processamento de informação pelo cérebro. O neurônio artificial, ver figura A3.1, é a unidade de processamento de informação, constituinte fundamental das redes neurais artificiais.

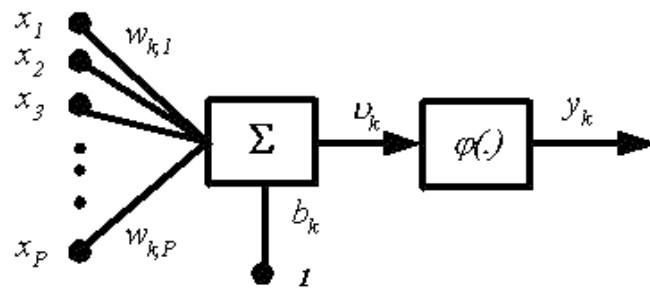


Figura A3.1 – Neurônio básico

Cada neurônio recebe em sua entrada o vetor de sinais \mathbf{x} e aplica um peso $w_{k,i}$ a cada entrada x_i do vetor de entrada. As entradas do neurônio, assim ponderadas, são somadas constituindo assim o argumento v_k de entrada de função de ativação $\varphi(\cdot)$ que produz a saída y_k do neurônio :

$$v_k = \mathbf{W}\mathbf{x} + b_k \quad \text{eq. (1)}$$

$$y_k = \varphi(v_k) = \varphi(\mathbf{W}\mathbf{x} + b_k) \quad \text{eq. (2)}$$

onde

$$\mathbf{W} = [w_{k,1} \ w_{k,2} \ , \dots, \ w_{k,p}] \quad \text{eq. (3)}$$

$$\mathbf{x} = [x_1 \ x_2 \ , \dots, \ x_p]^T \quad \text{eq. (4)}$$

A3.2 – Funções de Ativação

A função de transferência da função de ativação $\varphi(\cdot)$ pode ser qualquer função. A função produz em sua saída um valor y_k resultante da aplicação de uma função de transferência ao argumento v_k recebido como entrada.

A seguir são exemplificados alguns tipos de funções de transferência comumente utilizadas. As funções de transferência mais utilizadas são as funções linear, função degrau e função sigmóide.

Função Limitador

$$\varphi(v_k) = \begin{cases} 1 & \text{se } v_k \geq 0 \\ -1 & \text{se } v_k < 0 \end{cases} \quad \text{eq. (5)}$$

Função Degrau

$$\varphi(v_k) = \begin{cases} 1 & \text{se } v_k \geq 0 \\ 0 & \text{se } v_k < 0 \end{cases} \quad \text{eq. (6)}$$

Função Linear

$$\varphi(v_k) = m \ v_k, \forall v_k \quad \text{eq. (7)}$$

Função Sigmóide (logística)

$$\varphi(v_k) = \frac{1}{1 + \exp(-\alpha v_k)}, \quad \forall v_k, \alpha \in \mathfrak{R} \quad \text{eq. (8)}$$

A3.3 – Camada de Neurônios

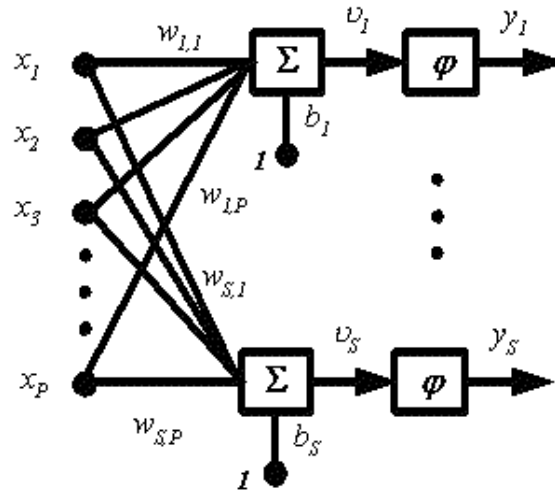


Figura A3.2 – Camada de neurônios

Os neurônios, conforme pode ser visto na figura A3.2, podem ser organizados formando uma rede neural de uma camada de neurônios. As entradas da rede são aplicadas a todos os neurônios e cada neurônio produz a sua própria saída:

$$\mathbf{y} = \varphi(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad \text{eq. (9)}$$

onde

$$\mathbf{y} = [y_1, y_2, \dots, y_s]^T \quad \text{eq. (10)}$$

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,p} \\ w_{2,1} & w_{2,2} & \dots & w_{2,p} \\ \dots & \dots & \dots & \dots \\ w_{s,1} & w_{s,2} & \dots & w_{s,p} \end{bmatrix} \quad \text{eq. (11)}$$

$$\mathbf{x} = [x_1, x_2, \dots, x_p]^T \quad \text{eq. (12)}$$

$$\mathbf{b} = [b_1, b_2, \dots, b_s]^T \quad \text{eq. (13)}$$

A3.4 – Redes de Múltiplas Camadas de Neurônios

As redes neurais artificiais podem ser constituídas por várias camadas de neurônios conectadas em cascata em que as saídas de uma camada constituem a entrada da próxima camada. Cada camadas pode ter número de neurônios diferente das outras camadas bem como funções de ativação diferentes. As redes assim constituídas são conhecidas como redes multicamadas feed-forward.

Assim para uma rede constituída por três camadas, teríamos para:

$$\text{primeira camada: } \mathbf{y}^1 = \varphi^1(\mathbf{IW}_{11}\mathbf{x} + \mathbf{b}^1) \quad \text{eq. (14)}$$

$$\text{segunda camada: } \mathbf{y}^2 = \varphi^2(\mathbf{LW}_{21}\mathbf{y}^1 + \mathbf{b}^2) \quad \text{eq. (15)}$$

$$\text{terceira camada: } \mathbf{y}^3 = \varphi^3(\mathbf{LW}_{31}\mathbf{y}^2 + \mathbf{b}^3) \quad \text{eq. (16)}$$

$$\text{ou } \mathbf{y}^3 = \varphi^3(\mathbf{LW}_{31}\varphi^2(\mathbf{LW}_{21}\varphi^1(\mathbf{IW}_{11}\mathbf{x} + \mathbf{b}^1) + \mathbf{b}^2) + \mathbf{b}^3) \quad \text{eq. (17)}$$

onde:

\mathbf{IW}_{11} matriz de pesos da entrada para a primeira camada

\mathbf{LW}_{21} matriz de pesos da primeira para a segunda camada

\mathbf{LW}_{32} matriz de pesos da segunda para a terceira camada

A3.5 – Aprendizado

Redes neurais podem realizar o mapeamento $S(.)$ de um conjunto de entrada x em um conjunto de saída y :

$$y = S(x) \quad \text{eq. (18)}$$

As redes aprendem o mapeamento $S(.)$ por meio de um processo de treinamento supervisionado, ver figura A3.3, no qual os pesos da matriz de pesos \mathbf{W} da rede são ajustados a partir da apresentação para a rede dos conjuntos de dados de entrada e saída (x, y) .

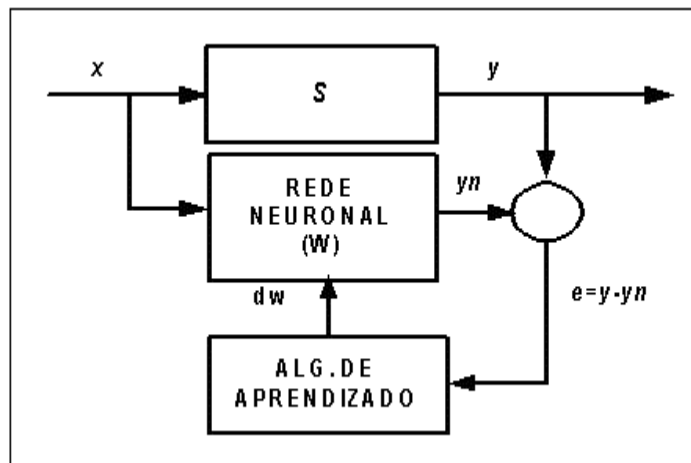


Figura A3.3. Treinamento supervisionado

A3.6 – Algoritmo de Aprendizado Backpropagation

O algoritmo de aprendizado backpropagation, utilizado em redes neurais multicamadas feed-forward, é um algoritmo de treinamento supervisionado em que pares de valores de entrada e saída (x, y) são apresentados à rede neural.

Para cada par de valores, começando da primeira camada, a rede calcula os valores de saída das diversas camadas até atingir a camada de saída, a rede neural avalia o valor de saída da rede y_n . O erro $e = y - y_n$ entre o valor esperado e o valor determinado pela rede é calculado e usado por um algoritmo de ajuste dos pesos para ajustar os pesos da matriz de pesos:

$$W = W + dw \quad \text{eq. (19)}$$

O algoritmo de correção retropropaga o erro, camada por camada, até atingir a camada de entrada. Em cada passo da retropropagação, os pesos de uma camada são ajustados de modo a minimizar a função de erro definida pela superfície:

$$J = \frac{1}{2} E[y - y_n] \quad \text{eq. (20)}$$

Vários métodos de minimização, ver ANEXO 1, podem ser usados, como por exemplo o método gradiente descendente. A cada passo do algoritmo, os pesos da rede são atualizados no sentido contrário ao do gradiente $\text{grad}(J) = \nabla(J)$:

$$W = W - \alpha \nabla(J) \quad \text{eq. (21)}$$

até atingir a condição ótima dada por $\nabla(J) = 0$. A constante de aprendizado α regula a intensidade da atualização dos valores da matriz de pesos, acelerando ou retardando a convergência.

A3.7 – Predição com Redes Neurais

Usando-se uma linha de retardo com derivações na entrada de uma rede neural, pode-se construir um modelo de predição com redes neurais. Uma linha de retardo com derivações, consiste de um conjunto de elementos de retardo em que se a entrada do elemento é $x(t)$ então a sua saída é $x(t-1)$.

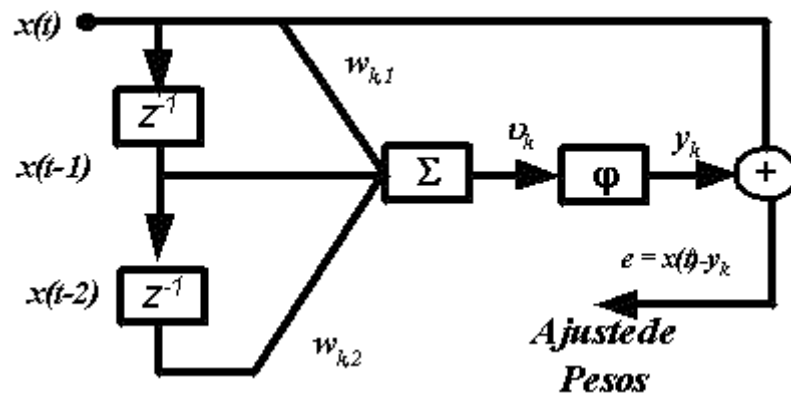


Figura A4.4 – Filtro linear com rede neural

A figura A3.4 apresenta uma rede neural de uma única camada com uma linha de retardo com uma entrada e três derivações, correspondendo a diferentes estágios de retardo do sinal, aplicadas à rede neural. Usando-se uma função de ativação linear e ajustando-se os pesos da rede de modo a minimizar o valor da saída da rede quando o valor de $x(t)$ for aplicado, teremos construído um filtro preditor capaz de calcular o valor de $x(t)$ em função dos valores de $x(t-2)$ e $x(t-1)$.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, J. C. S., 2001, Método de Identificação de Transientes com Abordagem Possibilística, Otimizado por Algoritmo Genético, Tese de Mestrado, Universidade Federal do Rio de Janeiro, COPPE/UFRJ, Engenharia Nuclear.

ALVARENGA M.A.B., 1998, Diagnóstico do Desligamento de um Reator Nuclear Através de Técnicas Avançadas de Inteligência Artificial. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, Brasil.

BARTLETT, E. B., AND UHRIG R. E., 1992, “Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network”, Nuclear Technology, Vol. 97, (Mar), pp. 272-281.

BASU, A, AND BARTLETT E. B., 1994, “Detecting Faults in a Nuclear Power Plant by Using a Dynamic Node Architecture Artificial Neural Network”, Nuclear Science and Engineering, Vol. 116, pp. 313-325.

CADZOW, J.A, MARTENS, H.R., 1970, Discrete Time and Computer Controls Systems, Prentice-Hall, Computer Applications in Electrical Engineering Systems.

CAMACHO, E. F., BORDONS, C., 1995, Model Predictive Control in The Process Industry, Springer, London.

CLARKE, D. W., MOHTADI, C., TUFSS, P.S., 1987, Generalized Predictive Control, Part 1, Part 2. Automatica 24, 137-169.

DOMINGOS, R.P., 2003, Sistemas de Computação Flexível Aplicados ao Controle e Identificação de Processos em Engenharia Nuclear[Rio de Janeiro], vol 140, (COPPE/UFRJ, d. Sc., Engenharia Nuclear), Tese, Universidade Federal do Rio de Janeiro.

DOMINGOS, R.P., CALDAS, G.H.F., PEREIRA, C.M.N.A, SCHIRRU, R., 2003, PWR Xenon Oscillation Control through a Fuzzy Expert System Automatically Designed by Means of Genetic Programming, Applied Soft Computing, pp 317-323.

DOMINGUES, R.P., 1997, Modelos Não Lineares da Engenharia Nuclear como uma Aplicação de Programação Genética, Tese de Doutorado, Universidade Federal do Rio de Janeiro, COPPE/UFRJ, Engenharia Nuclear.

DUDERSTAD, J., HAMILTON, L. H., 1976, Nuclear Reactor Analysis, U.S.A., John Wiley & Sons, Inc.

FURUKAWA H., UEDA T. AND KITAMURA M., 1995, “Use of Self-Organizing Neural Networks for Rational Definition of Plant Diagnostic Symptoms”, Proceedings of the International Topical Meeting on Computer-Based Human Support Systems, pp. 441-448.

GLASSTONE, S., 1960, Principles o Nuclear Reactor Engineering, Mc Milan & CO Ltd, London.

GOLDBERG, D., E., 1989, Genetic Algorithms in Search Optimization and Machine Learning, Addison Wesley.

HAYKIN, S., 1994, Neural Networks A Compreensive Foundation, Prentice Hall.

HAYKIN, S., 1996, Adaptive Filter Theory, Third Edition, Prentice Hall Information And Sciences Series.

HOLLAND, J. H., 1975, Adaptation in Natural and Artificial Systems, University of Michigan Press.

JEONG E., FURUTA K., KONDO S., 1996, "Identification of Transient in Nuclear Power Plant Using Adaptive Template Matching with Neural Network", Proceedings of the International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies, pp.243-250.

JEONG E., FURUTA K., KONDO S., 1996, "Identification of Transient in Nuclear Power Plant Using Adaptive Template Matching with Neural Network", Proceedings of the International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies, pp.243-250.

KAYLATH, T., 1980, Linear Systems, Prentice-Hall Information and System Sciences Series.

KENNEDY, J., EBERHART, R. C., 1995, Particle Swarm Optimization, Proceedings IEEE, IV Conference on Neural Networks, 1995, pp. 1942-1948.

KENNEDY, J., EBERHART, R. C., SHI, Y., 2001, Swarm Intelligence, Morgan Kaufmann Publishers, CA, USA.

LAMARSH, J. R., 1966, Introduction to Nuclear Reactor Theory, Addison-Wesley Publishing Company, Reading, Massachusetts, U.S.A.

LUENBERGER, D. G., 1979, Introduction to Dynamic Systems Theory, Models and Applications, John Wiley and Sons.

LUENBERGER, D. G., 1986, Linear and Non Linear Programming, 2nd Edition, New York Addison Wesley Publishing Company.

MOL, A.C.A., 2002, Um Sistema de Identificação de Transientes com Inclusão de Ruídos e Indicação de Eventos Desconhecidos, Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, Brasil.

MOL, A.C.A., MARTINEZ, A.S., SCHIRRU, R., 2002, A New Approach for Transient Identification with “D’ont Know” Response Using Neural Networks. Power Plant Surveillance and Diagnostics _ Applied Research with Artificial Intelligence, Springer Verlag.

MOL, A.C.A., MARTINEZ, A.S., SCHIRRU, R., 2002, Sistema de Identificação de Transientes com Inclusão de Ruídos e Indicação de Eventos Desconhecidos, XIII ENFIR – National Meeting of Reactor Physics and Thermal Hydraulics, Rio de Janeiro, Agosto, 2002.

NA, G. M., UDADHYAYA., B.R., CHOI, J.I., 2000, Adaptive Control for Axial Power Distribution in Nuclear Reactors, magyna@ic21a.chosun.ac.kr.

PEREIRA, C. M. N. A, SCHIRRU R, MARTINEZ A.S, 1998, Basic Investigations Related to Genetic Algorithms in Core Designs, Annals of Nuclear Energy, 26 173-179, Pergamon Press.

PEREIRA, C.M.N.A, SCHIRRU, R., 2002, Genetic Based Identification System Design With Automatic Selection of Meaningful Variables, World Scientific.

ROJAS, R., 1996, Neural Networks A Systematic Introduction, Springer Verlag.

ROVERSO, D., 1998, A Neural Model for Event Identification in Dynamic Processes. In: Report HWR-516, OECD HALDEN REACTOR PROJECT, Institutt for Energiteknikk, Norway.

ROVERSO, D., 1998, A Neural Model for Event Identification in Dynamic Processes. In: Report HWR-516, OECD HALDEN REACTOR PROJECT, Institutt for Energiteknikk, Norway.

RUSSELL, S., NORVIG, P., 1995, Artificial Intelligence A Modern Approach, Prentice Hall Series in Artificial Intelligence.

SCHIRRU, R., MARTINEZ, A. S., PEREIRA, C. M. N. A., 1999, DOMINGOS R. P., MACHADO, M. D. and MACHADO, L., 1999, Intelligent Soft Computing In Nuclear Energy in Brazil, Progress in Nuclear Energy, Vol 35, No 3-4, pp 367-391, Pergamon Press Ltd.

SHIMAZU, Y., 1992, “Direct Method of Search for Optimal Xenon Oscillation Control Based On New Concept of Axial Offsets”, Journal of Nuclear Science and Tecnology, 29[10], pp. 966-971.

SHIMAZU, Y., 1995, “Continuous Guidance Procedure for Xenon Oscillation”, Journal of Nuclear Science and Tecnology, 29[10], pp. 966-971.

SILVENNOINEN, P., 1976, Reactor Core Fuel Management, Pergamon Press Ltd., First Edition.

SUICH, J. E. and HONEC, H. C, 1967, The Hammer System Heterogeneous Analysis by Multigroup Methods of Exponentials and Reactors, Savannah River Laboratory, Aiken South Carolina.

TRELEA, I. C., 2002, The Particle Swarm Optimization Algorithm: Convergence Analysis and Parameter Selection, INA P-G, UMR Gémie et Microbiologie des Procédès Alimentaries, PB01, 78850 Thiverval-Grignon, France, September, ELSEVIER, www.ComputerScienceWeb.com.

TSOUKALAS, L., H., UHRIG, R., E., 1997, Fuzzy and Neural Approaches in Engineering, John Wiley & Sons, Inc., 1997