

INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES

Autarquia associada à Universidade de São Paulo

**ANÁLISE DAS VARIÁVEIS DE ENTRADA DE UMA REDE NEURAL
USANDO TESTE DE CORRELAÇÃO E ANÁLISE DE
CORRELAÇÃO CANÔNICA**

VALTER MAGALHÃES COSTA

Dissertação apresentada como parte dos requisitos para a obtenção do Grau de Mestre em Ciências na Área de Tecnologia Nuclear – Reatores.

Orientadora:

Profa Dra Iraci Martinez Pereira

São Paulo

2011

Dedico este trabalho: **à minha mãe, Dona Maria.
ao meu pai, Sr. Manoel
aos meus irmãos e esposa.**

Agradecimentos

À professora Iraci Martinez P. Gonçalves pela orientação deste trabalho e pela paciência.

À minha esposa Elaine Maria Costa, por tudo.

Aos funcionários da biblioteca e professores do IPEN que sempre estiveram dispostos a ajudar.

Eu não falhei, encontrei 10 mil soluções que não davam certo.
[Thomas A. Edison]

ANÁLISE DAS VARIÁVEIS DE ENTRADA DE UMA REDE NEURAL USANDO TESTE DE CORRELAÇÃO E ANÁLISE DE CORRELAÇÃO CANÔNICA

Valter Magalhães Costa

RESUMO

A monitoração de variáveis e o diagnóstico de falhas é um aspecto importante a se considerar seja em plantas nucleares ou indústrias de processos, pois um diagnóstico precoce de falha permite a correção do problema proporcionando a não interrupção da produção e a segurança do operador e, assim, não causando perdas econômicas. O objetivo deste trabalho é, dentro do universo de todas as variáveis monitoradas de um processo, construir um conjunto de variáveis, não necessariamente mínimo, que será a entrada de uma rede neural e, com isso, conseguir monitorar, o maior número possível de variáveis. Esta metodologia foi aplicada ao reator de pesquisas IEA-R1 do IPEN.

Para isso, as variáveis Potência do reator, Vazão do primário, Posição de barras de controle/segurança e Diferença de pressão no núcleo do reator ΔP , foram agrupadas, pois por hipótese quase todas as variáveis monitoradas em um reator nuclear tem relação com alguma dessas ou pode ser resultado da interação de duas ou mais. Por exemplo, a Potência está relacionada ao aumento e diminuição de algumas temperaturas bem como à quantidade de radiação devido à fissão do urânio; as Barras são reguladoras de potência e, por conseqüência podem influenciar na quantidade de radiação e/ou temperaturas; a Vazão do Circuito Primário, responsável pelo transporte de energia e pela conseqüente retirada de calor do núcleo. Assim, tomando o grupo de variáveis mencionadas, calculamos a correlação existente entre este conjunto B e todas as outras variáveis monitoradas (coeficiente de correlação múltipla), isto é, através do cálculo da correlação múltipla, que é uma ferramenta proposta pela teoria das Correlações Canônicas, foi possível calcular o quanto o conjunto B pode prever cada uma das variáveis monitoradas. Uma vez que não seja possível uma boa qualidade de predição com o conjunto B, é acrescentada uma ou mais variáveis que possuam alta correlação com a variável melhorando a qualidade de predição. Finalmente, uma rede pode ser treinada com o novo conjunto e os resultados quanto a monitoração foram bastante satisfatórios quanto às 64 variáveis monitoradas pelo sistema de aquisição de dados do reator IEA-R1 através de sensores e atuadores, pois com um conjunto de 9 variáveis foi possível monitorar 51 variáveis.

ANALYSIS OF INPUT VARIABLES OF AN ARTIFICIAL NEURAL NETWORK USING BIVARIATE CORRELATION AND CANONICAL CORRELATION

Valter Magalhães Costa

ABSTRACT

The monitoring of variables and diagnosis of sensor fault in nuclear power plants or processes industries is very important because an early diagnosis allows the correction of the fault and, like this, do not cause the production interruption, improving operator's security and it's not provoking economics losses. The objective of this work is, in the whole of all variables monitor of a nuclear power plant, to build a set, not necessary minimum, which will be the set of input variables of an artificial neural network and, like way, to monitor the biggest number of variables. This methodology was applied to the IEA-R1 Research Reactor at IPEN.

For this, the variables Power, Rate of flow of primary circuit, Rod of control/security and Difference in pressure in the core of the reactor (ΔP) was grouped, because, for hypothesis, almost whole of monitoring variables have relation with the variables early described or its effect can be result of the interaction of two or more. The Power is related to the increasing and decreasing of temperatures as well as the amount radiation due fission of the uranium; the Rods are controls of power and influence in the amount of radiation and increasing and decreasing of temperatures and the Rate of flow of primary circuit has function of the transport of energy by removing of heat of the nucleus. Like this, labeling $B = \{\text{Power, Rate of flow of Primary Circuit, Rod of Control/Security and } \Delta P\}$ was computed the correlation between B and all another variables monitoring (coefficient of multiple correlation), that is, by the computer of the multiple correlation, that is tool of Theory of Canonical Correlations, was possible to computer how much the set B can predict each variable.

Due the impossibility of a satisfactory approximation by B in the prediction of some variables, it was included one or more variables that have high correlation with this variable to improve the quality of prediction. In this work an artificial neural network was trained and the results were satisfactory since the IEA-R1 Data Acquisition System reactor monitors 64 variables and, with a set of 9 input variables resulting from the correlation analysis, it was possible to monitor 51 variables using neural networks.

Sumário

1. INTRODUÇÃO.....	1
1.1. Motivação	6
1.2. Objetivo	7
1.3. Contribuição do trabalho.....	7
1.4. Organização da dissertação	8
2. REVISÃO BIBLIOGRÁFICA	9
3. REATORES NUCLEARES.....	14
3.1. O Reator PWR	17
3.1.1.O Circuito Primário.....	18
3.1.2.O Circuito Secundário.....	20
3.2. O reator IEA-R1	20
3.3. Considerações Físicas dos Reatores	24
4. CORRELAÇÃO LINEAR.....	27
4.1. Equações lineares e o modelo de regressão linear	27
4.2. Coeficiente de Correlação de Pearson	30
4.2.1.Fórmula alternativa para o coeficiente de correlação	33
5. CORRELAÇÃO MÚLTIPLA.....	35
5.1. Análise de Correlação Canônica	35
5.1.1.A Matemática da correlação canônica.....	35
6. REDES NEURAIS	43
6.1. Modelagem de um neurônio	44
6.1.1.Função de ativação.....	46
6.1.2.Arquitetura de Redes Neurais Artificiais	48
6.2. O processo de aprendizado	50
6.2.1.Funções de aprendizado	50

6.2.2. Tipos de Aprendizado.....	51
6.2.2.1. Método do Gradiente Descendente	52
6.2.2.2. Aprendizado por correção do erro.....	53
6.3. Algoritmo Backpropagation	54
6.4. Validação de modelos em redes neurais	57
7. MODELAGEM MATEMÁTICA DE SISTEMAS E REDES NEURAIAS COMO APROXIMAÇÃO DE FUNÇÕES	60
7.1. Linearização de modelos	60
7.2. Redes Neurais e Aproximação de funções	60
8. METODOLOGIA DE TRABALHO	62
8.1. Arquitetura das redes MLP utilizadas neste trabalho	62
8.2. Desenvolvimento do Trabalho	63
9. DISCUSSÃO DOS RESULTADOS	75
9.1. Conclusões	76
10. Anexos.....	78
Referências Bibliográficas.....	98

1. INTRODUÇÃO

O crescente avanço tecnológico verificado nas últimas décadas exige das máquinas e, em geral, em grandes sistemas de processos (usinas químicas, indústria manufatureira, indústria aeroespacial, indústria automotiva etc.) maiores capacidades de trabalho e velocidade de operação e, neste contexto, o controle automático tem desempenhado um papel fundamental e se tornou parte integrante dos modernos processos industriais e de produção.

Do ponto de vista teórico e prático, os avanços conseguidos na área de controle automático, vêm produzindo meios para aperfeiçoar o desempenho dos sistemas dinâmicos, aumentar a produtividade, diminuir o trabalho árduo de várias rotinas de operações manuais repetitivas, entre outros.

Por outro lado, a complexidade das configurações físicas em grandes sistemas de processo em adição aos diversos componentes de hardware e software existentes nesses ambientes sinaliza para a importância de um sistema de monitoração e diagnóstico eficaz. Torna-se evidente a necessidade de estabelecer não apenas a automatização das operações, mas, também, o monitoramento sobre os processos nas suas várias etapas como forma de garantir a identificação de problemas.

Além disso, atualmente, uma das maiores preocupações da indústria é a de manter seus equipamentos em funcionamento sem que ocorram paradas não programadas. Assim, um dos fatores de interesse das indústrias no desenvolvimento de novas técnicas de detecção de falhas é a preocupação com a segurança de seus sistemas, tendo-se a necessidade de supervisão e monitoramento de modo que a falha seja detectada e corrigida o mais rápido possível.

Qualquer sistema de potência deve ser controlado. Ele deve ser iniciado e desenvolver-se de forma segura até um nível de potência pré-determinado e, além disso, deve responder a qualquer mudança nas condições de operação que são impostas sobre ele [16] [43] [15].

Em se tratando de reatores nucleares, seu ciclo de operações consiste de três estágios [43]:

- (a) a aproximação à criticalidade, que é caracterizado por um baixo fluxo de nêutrons sendo, nessa primeira etapa, necessário a utilização de detectores para observar os

efeitos dessas partículas tal que uma elevação lenta e gradual deve ser garantida pela limitação da velocidade de retirada das barras de controle.

- (b) a elevação da potência até à de operação, onde o fluxo de nêutrons torna-se elevado.
- (c) a operação em regime estacionário.

Quer o reator seja usado para produção de energia ou como fonte de nêutrons para experimentos, é conveniente provê-lo com controle automático para eliminar uma excessiva atenção dos operadores e para manter níveis constantes e estacionários. [43]

Falhas em instrumentos de monitoração em um sistema de controle podem causar sérios danos ao ambiente. Uma falha que ocorre gradualmente pode causar uma deterioração no desempenho do sistema e, se a mesma for detectada no seu estágio inicial, o instrumento falho pode ser reparado antes que problemas mais sérios ocorram. Problemas esses que envolvem desde questões de segurança da instalação e humanas até perdas econômicas.

As tarefas de um sistema de Monitoração e Diagnóstico consistem em supervisionar variáveis, verificando se os seus valores estão dentro de um intervalo considerado seguro ou a detecção de falhas. Para isso utilizam-se, em sistemas de monitoração e diagnóstico, a redundância física e a redundância analítica.

Na redundância física (ou por hardware), alguns componentes do sistema de controle (sensores, atuadores e controladores) são duplicados, ou seja, existem dois componentes para o desempenho da mesma função. O maior problema encontrado com este método são os equipamentos extras necessários, o custo de manutenção e instalação e o espaço físico requerido para acomodação [8].

Uma derivação de tal método é a redundância física em paralelo, na qual dois ou mais conjuntos de sensores, atuadores e controladores, cada qual com capacidade individual satisfatória de controle, são instalados para a execução da mesma tarefa. Um sistema gerenciador compara os sinais de controle para detectar o conjunto faltoso e, com dois sinais redundantes, um sistema votante pode detectar a existência de um conjunto faltoso. Com três sinais redundantes, um sistema votante pode detectar e isolar o conjunto faltoso, selecionando, assim aquele que não deve ser utilizado para o controle do sistema [16].

Redundância física pode proteger o sistema contra falhas nos componentes do sistema de controle e, a menos que todos os sensores redundantes indiquem a mesma leitura

ou sinal de falha, a redundância física não pode proteger os componentes da usina. Além disso, em se tratando de centrais nucleares que necessitam de equipamentos mais caros e robustos, redundância física é quase sempre limitada pelos fatores custo de instalação e disponibilidade de espaço para a instalação de sensores.

Em oposição à redundância física (ou por hardware) tem-se a redundância analítica (ou redundância por software) que faz uso da predição de sinais gerados por um modelo matemático do sistema considerado. Este modelo matemático pode ser elaborado a partir das equações matemáticas que descrevem o fenômeno real (conhecido como “first principle model”). Estas predições são comparadas com as medidas reais dos sensores do sistema. A principal vantagem do emprego de redundância analítica é que não há necessidade de acrescentar novos componentes ao sistema existente para implementar o algoritmo de monitoração e detecção de falhas [30].

No caso específico de um reator nuclear, muitas de suas características operacionais, que são pertinentes a um projeto de um sistema de controle, são baseadas em análise matemática. Estas características operacionais são tão complicadas que não é possível escrever equações matemáticas para descrevê-las completamente e as equações simplificadas contêm muitas hipóteses referentes ao comportamento do reator que representam apenas uma aproximação bastante restrita do verdadeiro comportamento, mas que podem ser justificados apenas empiricamente.

Os sistemas de processo, em geral, são caracterizados por serem não-lineares e apresentarem multivariáveis tornando a operação complexa. O projeto de um sistema de monitoração consiste na sua modelagem tal que descreva o processo dinâmico com precisão suficiente para trabalhar off-line com o sistema controlador da usina, ou seja, o projeto de sistemas de segurança deve ser completamente separado do projeto de controle [48] [45] [65].

Neste contexto surgem os Sistemas Especialistas ou Sistemas Baseados em Conhecimento, que desempenham um importante papel no suporte de operação, por exemplo, em sistemas nucleares, e fornecem ajuda na manutenção das atividades da usina. O operador do reator deve ser hábil em reagir a indicações de condições anormais e o sistema especialista deve fornecer a informação necessária para a ação efetiva [11].

Um Sistema Especialista é um programa de computador onde é incorporado o conhecimento do especialista humano como ferramenta de decisão. Um Sistema Especialista típico inclui um conhecimento base, regras de inferência e uma interface humana. O conhecimento base consiste em disponibilizar informações relevantes de uma

base de dados bem como várias regras próprias desenvolvidas da experiência pelos chamados processos heurísticos. As regras de inferência que manipulam o conhecimento base por regras da Lógica para chegar a conclusões. A tradução das entradas em linguagem de computador e então, dos resultados calculados, dispor a saída numa interface característica [36].

A maior vantagem deste tipo de programa é a habilidade para aplicar critérios do especialista sistematicamente para solucionar problemas sem o mesmo estar presente [16] [36].

O final da década de 80 marcou o ressurgimento da área de Redes Neurais Artificiais (RNA), também conhecidas como conexionismo ou sistemas de processamento paralelo distribuído. Esta forma de computação não algorítmica é caracterizada por sistemas que, em algum nível, relembram a estrutura do cérebro humano. Por não ser baseada em regras ou programas, a computação neural se constitui em uma alternativa à computação algorítmica convencional [6] [20].

Em processos nos quais há grande volume de dados e não existe um conhecimento estruturado sobre os mesmos ou, mais especificamente, em processos onde existem muitas variáveis a serem monitoradas e que são de difícil inter-relacionamento, se torna especialmente útil o uso de *Redes Neurais Artificiais* no monitoramento e controle destes processos [39].

A solução de problemas através de RNA é bastante atrativa, já que a forma como estes são representados internamente pela rede e o paralelismo natural inerente à arquitetura das RNA criam a possibilidade de um desempenho superior ao dos modelos convencionais. Em RNA, o procedimento usual na solução de problemas passa inicialmente por uma fase de aprendizagem, onde um conjunto de exemplos é apresentado para a rede, a qual extrai as características necessárias para gerar, posteriormente, respostas para o problema.

A capacidade de *aprender* através de exemplos e de generalizar a informação aprendida são os atrativos principais da solução de problemas através de RNA. Não obstante, as RNA, através do aprendizado, são capazes de atuar como mapeadores universais de funções multivariáveis, com custo computacional que cresce apenas linearmente com o número de variáveis. Outras características importantes são a capacidade de auto-organização e de processamento temporal que, aliadas àquelas citadas anteriormente, fazem das RNA uma ferramenta computacional extremamente poderosa e atrativa para a solução de problemas complexos.

A definição da arquitetura de RNA é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser tratado pela rede. Redes com camada única de neurônios, por exemplo, só conseguem resolver problemas linearmente separáveis e redes recorrentes, por sua vez, são mais apropriadas para resolver problemas que envolvem processamento temporal [32].

Portanto, RNA apresentam problemas e limitações que não permitem seu uso exclusivo para a solução de uma quantidade significativa de problemas. Assim, integrar técnicas baseadas em dados (RNA) com técnicas que se utilizam de conhecimentos foi a idéia para se obter soluções mais robustas e eficientes, pois o uso de uma única técnica, devido às suas limitações e/ou deficiências, pode não ser capaz, por si só, de resolver um dado problema [6] [7].

Algumas abordagens têm sido propostas para integrar RNA com:

- Estatística;
- Lógica Fuzzy;
- Algoritmos Genéticos;
- Sistemas Baseados em Conhecimento;
- Raciocínio Baseado em casos;
- Lógica Matemática;

Em qualquer decisão que tomamos sempre levamos em conta um grande número de fatores e, obviamente, nem todos estes pesam da mesma maneira na hora de uma escolha. Às vezes, não identificamos de maneira sistemática esses fatores, ou seja, não identificamos quais as variáveis que afetaram a nossa decisão.

As diversas ciências, cuja pretensão é conhecer a realidade e interpretar os acontecimentos, baseiam-se no conhecimento das variáveis intervenientes consideradas importantes nos eventos estudados. Assim, estabelecer relações, encontrar ou propor leis explicativas é o papel próprio da ciência e, para isso, é necessário controlar, manipular e medir as variáveis que são consideradas relevantes ao entendimento do fenômeno analisado.

Seja no dia-a-dia ou em ciência, podemos sempre analisar as variáveis isoladamente e a partir desta análise fazer inferências sobre a realidade. Esta simplificação tem vantagens e desvantagens. Quando um fenômeno depende de muitas variáveis, geralmente

este tipo de análise falha, pois não basta conhecer informações estatísticas isoladas, mas é necessário também conhecer a totalidade destas informações fornecida pelo conjunto das variáveis. As relações existentes entre as variáveis não são percebidas e assim efeitos antagônicos ou sinérgicos de efeito mútuo entre variáveis complicam a interpretação do fenômeno a partir das variáveis consideradas. Porém, no caso restrito de variáveis independentes entre si é possível, com razoável segurança, interpretar um fenômeno complexo usando as informações estatísticas de poucas variáveis.

Essa preocupação em reduzir a poucas variáveis o estudo de um determinado fenômeno que dependa de várias variáveis desenvolveu a área da estatística chamada Análise Multivariada. A denominação “Análise Multivariada” corresponde a um grande número de métodos e técnicas que utilizam simultaneamente todas as variáveis na interpretação do conjunto de dados obtidos.

1.1. Motivação

Nas aplicações onde se utilizam RNA, ou em outras áreas onde se tem um grande número de variáveis para analisar, um problema fundamental é encontrar uma conveniente representação dos dados multivariados como, por exemplo, em uma sala de controle onde existem centenas de variáveis monitoradas que indicam o *status* de operação da instalação [49]. Por razões de simplicidade conceitual e computacional, a representação é freqüentemente pensada como uma transformação linear dos dados originais, ou seja, cada componente da representação é uma combinação linear das variáveis originais. Métodos conhecidos de transformações lineares incluem Análise de Componentes Principais, Análise de Fatores, entre outros [24] [46].

De qualquer forma, a seleção adequada das variáveis de entrada é importante para se obter o menor número possível de variáveis, que contenham as informações necessárias para as Redes Neurais realizarem a monitoração da instalação. Normalmente deve-se recorrer ao conhecimento de um especialista na área para fazer a seleção das variáveis de entrada. Assim mesmo, podem ser testadas várias combinações destas variáveis pré-selecionadas até que seja atingido um resultado ótimo. Por esse motivo, seria muito interessante que houvesse um método de seleção automática das variáveis de entrada que seriam utilizadas numa rede neural, ou seja, uma seleção de variáveis que independa do prévio conhecimento do sistema em questão. Os resultados obtidos com uma seleção automática das variáveis será a utilização de RNA menores e com um tempo de

treinamento menor, e a eliminação da necessidade do conhecimento especialista para realizar esta tarefa.

1.2. Objetivo

O objetivo deste trabalho é o desenvolvimento de uma metodologia para seleção de variáveis de entrada de uma Rede Neural utilizando o Teste de Correlação Simples ou bivariado e o Teste de Correlação Múltipla que são metodologias da Estatística de Análise Multivariada.

Com as metodologias empregadas queremos construir, dentro do universo de todas as variáveis monitoradas de uma instalação nuclear, um conjunto de variáveis, não necessariamente mínimo, que será o conjunto de entrada numa Rede Neural. Tal conjunto será avaliado, quanto à eficiência na monitoração de sensores do Reator de Pesquisas IEA-R1 do IPEN e será utilizada uma base de dados contendo os valores das variáveis de processo da instrumentação coletados e armazenados pelo Sistema de Aquisição de Dados (SAD) do reator [41].

O Teste de Correlação Múltipla tem como objetivo principal o estudo das relações lineares existentes entre dois conjuntos de variáveis e é uma metodologia proposta por Hotelling conhecida como “Análise de correlações canônicas”. A idéia básica do método é resumir a informação de cada conjunto de variáveis-resposta em combinações lineares, sendo que a escolha dos coeficientes dessas combinações é feita tendo-se como critério a maximização da correlação entre os conjuntos de variáveis-resposta. Essa correlação mede basicamente o grau de associação existente entre os dois conjuntos de variáveis [19] [22] [23] [24] [53].

1.3. Contribuição do trabalho

Este trabalho propõe uma metodologia para a escolha de variáveis de entrada para uma rede neural visando a monitoração e diagnóstico de falhas nos sensores e atuadores do reator IEA-R1.

A contribuição do trabalho está na proposição de uma metodologia para a escolha conveniente das variáveis que permitirá o monitoramento de um número maior de

variáveis pelas redes neurais. O uso de redes neurais artificiais no monitoramento e diagnóstico de falhas em sensores e atuadores do reator IEA-R1 foi uma metodologia proposta em [8] demonstrando a eficiência da aplicação desta metodologia.

1.4. Organização da dissertação

Capítulo 1: Introdução teórica situando o tema do trabalho e objetivos.

Capítulo 2: Revisão bibliográfica mostrando algumas técnicas atualmente utilizadas em controle e monitoramento de sistemas.

Capítulo 3: Descrição do reator PWR e do Reator IEA-R1. São tratadas também questões físicas relativas ao seu funcionamento.

Capítulo 4: Trata dos aspectos formais da correlação linear bem como sua relação com o conceito de regressão linear.

Capítulo 5: Trata da correlação múltipla (aspectos formais e teóricos). O objetivo é generalizar o conceito de correlação linear.

Capítulo 6: Descrição das redes neurais artificiais.

Capítulo 7: Breve discussão das redes neurais como aproximadores universais de funções.

Capítulo 8: Descrição da metodologia utilizada e resultados.

Capítulo 9: Discussão dos resultados.

2. REVISÃO BIBLIOGRÁFICA

Apesar da utilização dos métodos de Monitoração e Detecção de Falhas terem iniciado na década de 70, ainda nos dias atuais existe a preocupação com seu desenvolvimento principalmente devido ao aumento e complexidade das configurações físicas de instalações seja em indústrias manufatureiras ou de processos [8] [11].

Monitoração e Diagnóstico consistem em supervisionar variáveis, verificando se os seus valores estão dentro de um intervalo considerado seguro e, portanto, a necessidade de se eliminar falhas em algum componente do sistema ou evitar acidentes torna necessário o aprimoramento de métodos de controle e diagnóstico de falhas.

Seja por redundância física ou analítica, um sistema de monitoramento e diagnóstico de falhas é fundamental para se prevenir acidentes ou evitar paradas não programadas na produção.

Os trabalhos constantes da revisão bibliográfica mostram uma tendência na utilização de sistemas de controle e diagnóstico através da redundância analítica, principalmente pela melhoria dos métodos de computação e processamento e, neste contexto, surgem as Redes Neurais Artificiais que são uma metodologia da área de Inteligência Artificial. As redes neurais são consideradas aproximadores universais de funções e são muito utilizadas na monitoração e diagnóstico de falhas [8] [20] [32].

A seguir, são estão descritos alguns trabalhos atuais situando a área de Monitoração e Diagnóstico.

Hwang, et al. (2009) [25] avaliaram o desempenho do sistema de controle principal e desenvolveram uma estratégia para aprimorar a eficiência e segurança do projeto da interface de controle-digital numa sala de controle. Foram analisados potenciais erros humanos onde a avaliação do fator humano foi verificado primeiramente durante a operação da usina nuclear e posteriormente durante a observação do treinamento. Foram identificados problemas devido à interface de operação da sala de controle principal, procedimentos empregados e layout da sala de controle principal. Os potenciais erros humanos e sugestões foram listados e os três problemas detectados foram analisados.

Savva, et al. (2010) [54] apresentaram duas diferentes aproximações, uma determinística e outra estocástica, para a determinação da dependência das barras de controle em relação ao nível de queima de combustível e nível de concentração de Xe no

núcleo do reator. Eles mostraram que a posição da barra é afetado pelos parâmetros: combustível queimado no entorno da barra, a concentração de Xe no núcleo, o tempo de operação da barra e sua posição no núcleo.

Hossam A. Gabbar 2010 [14] apresentou uma estrutura integrada para análise de controle de segurança e projeto para usinas nucleares. Este trabalho mostra o uso da metodologia do processo de modelagem orientado a objeto e modelagem de falhas para integrar segurança necessária, identificar perigos e cenários de propagação de falhas.

Controladores de processo são responsáveis por executar procedimentos de operação da plataforma do sistema para produzir um produto alvo ou serviço estável, seguro e otimizado. Isto significa que os sistema de controle de processo deve incluir aspectos de segurança de processo. Em plantas nucleares, sistemas de segurança são representados na forma de camadas independentes de proteção ou barreiras.

Correntemente, controle e segurança são práticas fragmentadas e o vão entre eles causa aumento nos riscos, custo e interrupção na produção.

Yun, et al (2010) [65] Este trabalho inicia a discussão sobre a estratégia de utilizar multi-modelos e a eficiência comprovada no que se refere a aproximações não lineares e, apesar dos parâmetros variarem com a mudança na potência, o sistema de potência pode ser pensado como um sistema linear desde que exista baixa variação na mesma. Desta forma, é compreendido que os processos de mudança de potência podem ser divididos em vários modelos lineares tal que um controlador local seja projetado.

Neste trabalho, a metodologia MPC (Model Predictive Control) é integrada a metodologia de multi-modelos que foi nomeada IIMPC (Improved Implicit Multiple Model Predictive Control), que primeiramente constrói o modelo de espaços de estados conforme as equações dinâmicas do núcleo do reator PWR e analisa a influência dos parâmetros durante o processo de mudança da potência para assim construir modelos locais.

O resultado mostrou que o IIMPC é eficiente para solucionar problemas de controle de sistemas não lineares variáveis no tempo.

Rodrigues, et al. (2006) [52] desenvolveram um método de diagnóstico de falhas para sistemas descritos por multi-modelos que consiste no projeto de um novo esquema de detecção e isolamento de falhas através de filtros adaptativos para tais sistemas.

Baseado na afirmação que o comportamento dinâmico do processo é descrito por uma aproximação multi-modelo em torno de diferentes pontos de operação, um conjunto de resíduos é estabelecido para gerar funções para falhas que estão diretamente ligadas

com a efetividade do filtro adaptativo que fornece estimativa de múltiplas magnitudes de falhas para todo o limite de operação do sistema. Neste trabalho, as condições de estabilidade do filtro adaptativo são estudadas e seu desempenho é testado usando um sistema hidráulico.

A metodologia de modelos multi-lineares é baseada na partição dos limites de operação de um sistema em regiões distintas e aplicando localmente modelos lineares a cada região e a função do filtro adaptativo é detectar, isolar e estimar falhas através da metodologia multi-modelos tal que o objetivo é a descrição ou reprodução do comportamento dinâmico do sistema.

A utilização de filtros adaptativos demonstrou sua efetividade em um sistema hidráulico de tratamento de água sobre múltiplos regimes de operação.

Sjöberg, et al. (2008) [58] Apresentaram um algoritmo para controle de sintonia para plantas. Sua contribuição consiste em não apenas identificar um modelo global não linear da usina e sim identificar modelos lineares variando no tempo durante a trajetória de operação da usina. O modelo variando no tempo é usado para calcular as variações das saídas da usina com relação aos parâmetros de controle e, como o modelo é válido somente para uma trajetória particular, a identificação foi repetida a cada iteração, isto é, a cada mudança nos parâmetros de controle. Foi assumido que as variáveis de processo a serem controladas podem ser descritas localmente por um modelo linear e para obter o modelo foi considerado sinais de referência que mudam vagarosamente entre dois regimes de operação.

Wen Tan (2011) [61] propôs um sistema de controle do nível de água para um gerador de vapor nuclear, pois 25% dos desligamentos de emergência numa usina nuclear do tipo PWR são causados por um deficiente controle do nível de água do gerador de vapor. O sistema de controle consiste de um controle de retroalimentação e um controlador que reutilize o dado de retroalimentação onde os parâmetros dos dois controladores estão diretamente relacionados aos parâmetros do modelo da usina.

A metodologia proposta mostrou ter bom desempenho com baixos e altos níveis de potência.

Papadokonstantakis, et al. (2005) [47] compararam métodos para testes de influência de variáveis em redes neurais. Os métodos são: Teoria da Informação (ITSS), Estrutura Bayesiana (ARD), Análise dos Pesos das Redes (GIM) e Omissão Sequencial das Variáveis (SZW). A comparação é baseada em conjuntos de dados teóricos e reais de diferentes tamanhos, complexidade e níveis de ruídos além da influência do tamanho da

rede. Este trabalho mostrou que: ITSS produz resultados similares a SZW e GIM, apesar de ser mais afetados pela dimensionalidade; SZW e GIM diferem de ARD com relação influência da variável, apesar de aplicada a rede neural com similar precisão na modelagem.

Funahashi (1989) [13] Neste trabalho provou que qualquer aplicação contínua pode ser realizada por redes neurais multicamadas com, ao menos, uma camada escondida cujas funções de saída são funções sigmóides. O ponto inicial da prova para o caso de uma camada escondida é a fórmula integral recentemente proposta por Irie-Miyake e disto, o caso geral (para qualquer número de camadas escondidas) pode ser provado por indução.

Attali e Pagès (1997) [2] demonstraram a propriedade da aproximação universal de funções do perceptron com uma camada escondida baseado na expansão de Taylor e no determinante de Vandermonde. Este resultado produz limitações para o projeto de camadas escondidas e resultados de convergência.

Minin, et al. (2010) [40] Mostraram que em aplicações de redes neurais não somente a melhor aproximação de função deve ser garantida, mas também a homogeneidade e monotonicidade das relações de entrada e saída devem ser garantidas, pois de outra forma a estabilidade das leis de controle podem se perder. Neste artigo foram comparadas duas redes neurais incorporando a restrição de monotonicidade na estrutura de trabalho das redes: redes monótonas do tipo Perceptron multicamadas e do tipo Min-Max. Foi mostrado que ambos são aproximadores universais de funções monotônicas onde o primeiro tipo se comportou bem com relação a tarefa de aproximação de funções e o segundo tipo de redes são mais convenientes para problemas de classificação.

Marcua, et al. (2008) [37] Investigaram o projeto de um sistema de detecção de falha para um circuito hidráulico de uma fábrica. Uma rede neural foi usada para identificar as diferentes relações das variáveis do processo e aproximar diferentes modelos não lineares do processo investigado (relações entre múltiplas entradas e uma única saída), sinais residuais em termos de predição de erros e uma decisão lógica baseado nos valores de *threshold*. Foram usados dados reais e simulados caracterizando o comportamento real do processo e posteriormente, diferentes falhas foram detectadas pelo sistema projetado e, desta forma, as investigações mostraram que o uso de redes neurais para um projeto de um sistema de detecção de falhas fornece uma ferramenta eficiente para modelagem de sistemas e identificação de falhas.

Bueno (2006) [8] Utilizou redes neurais artificiais na monitoração e detecção de falhas em sensores e atuadores do reator IEA-R1. Neste trabalho, escolheu arbitrariamente um conjunto de variáveis e desenvolveu um Sistema de Monitoração e Detecção de Falhas usando a metodologia de Redes Neurais Artificiais que foi aplicado ao reator de pesquisas IEA-R1. O desenvolvimento deste sistema foi dividido em três etapas: na primeira etapa foi feita a monitoração da rede utilizando RNA; na segunda etapa foi dedicado à detecção de falhas também utilizando as RNA com uma base de dados contendo falhas; e na terceira fase foi utilizado um sistema Neuro-Fuzzy para realizar o diagnóstico de falhas. O trabalho mostrou a viabilidade do emprego da metodologia de Redes Neurais Artificiais na monitoração e diagnóstico de falhas no reator IEA-R1, embora não tenha conseguido monitorar todas as variáveis do conjunto escolhido.

3. REATORES NUCLEARES

A fonte de energia de um reator é a fissão nuclear de átomos pesados, tais como urânio, isótopo U^{235} , por nêutrons e dentre os produtos desta reação de fissão estão a produção de dois ou mais nêutrons. A geração de mais nêutrons permite a continuidade da reação, pois, cada novo nêutron pode causar outra fissão [43].

Os fragmentos da fissão separam-se com velocidades muito grandes e cedem sua energia sob a forma de calor ao meio no qual são formados e, com efeito da fissão de vários núcleos, e com conseqüente emissão de nêutrons a cada fissão, é gerada uma reação em cadeia que, controlada, pode ser utilizada para várias finalidades como, por exemplo: servir de fonte de nêutrons para irradiação e ativação de materiais; servir para gerar potência para diversas utilizações; servir para produzir materiais transurânicos (plutônio principalmente), etc.

Os reatores nucleares possuem os meios para manter a reação em cadeia de forma controlada e podem ser entendidos como uma fonte de calor que substitui o processo de combustão nas plantas a gás ou a vapor pela fissão nuclear [63].

A energia liberada no núcleo pela fissão aparece em várias formas, mas principalmente como energia cinética dos fragmentos da fissão, nêutrons e partículas betas resultante do decaimento radioativo dos produtos da fissão tal que o calor gerado por estas três fontes compreende cerca de 90% ou mais da energia gerada e, o restante 10%, ou menos, representa os raios gama produzido na fissão [15] [43].

A remoção do calor efetivo é um aspecto importante no projeto de um reator, pois é a taxa com o que o calor é removido do núcleo que determina seu nível de potência. É uma função de muitos parâmetros que inclui a geometria do combustível, características do fluido refrigerante, e propriedades dos materiais, bem como questões relacionadas com o comportamento dos neutrons. Em muitos aspectos do projeto, princípios convencionais da engenharia de transferência de calor e mecânica dos fluidos são aplicáveis [31] [64].

A Termo-hidráulica descreve a relação envolvendo a integração da transferência de calor e os princípios da mecânica dos fluidos para remoção de calor do núcleo de um reator e como fundo para os aspectos termos-hidráulicos do projeto do reator é necessário considerar o caminho do transporte de energia do ponto de vista termodinâmico. Em uma usina nuclear, o calor flui do elemento combustível para o líquido refrigerante, e depois para um refrigerante secundário, como em um PWR, para formar vapor onde, em um

núcleo de um reator típico, o transporte de calor do combustível ao refrigerante circulante envolve o tradicional processo de condução, convecção e radiação.

Este calor deve ser removido continuamente. Em geral, um fluido que transfere calor deve circular pelo reator e então, por um trocador de calor, onde o vapor é produzido. Um bom refrigerante, desta forma, deve ser um fluido que, além de ter boas propriedades de transferência de calor, deve ser estável a altas temperaturas e a ação da radiação e ter pouca tendência para captura de nêutrons. Assim, o refrigerante que circula através do reator e troca calor deve ser não corrosivo e deve requerer baixa energia para seu bombeamento [15].

Uma primeira identificação do reator pode ser dada pela citação de materiais particulares usados.

O moderador deve possuir duas propriedades: um valor grande para a perda logarítmica de energia por colisão e uma baixa seção de choque para absorção de nêutrons térmicos, pois, de modo geral, o nêutron tem uma probabilidade maior de interagir com outro núcleo quando sua velocidade é menor. Os principais moderadores em uso são grafite, água leve e água pesada. Berílio ou óxido de berílio, embora caros, também podem ser usados e têm a função de reduzir a velocidade dos nêutrons produzidos na fissão para que possam atingir outros átomos fissionáveis mantendo a reação.

O refletor é uma camada de material moderador, usualmente não fissionável, que envolve a região ativa ou núcleo. Serve para duas finalidades que são: fazer retornar alguns nêutrons e que poderiam, de outra forma, escapar; e auxiliar a atenuação de nêutrons para energias térmicas. Assim, os moderadores acima citados podem servir, também, como refletores, com a mesma ou diferente escolha de material para o núcleo. Os reatores rápidos são também envolvidos com uma camada refletora, usualmente urânio natural. Este cobertor conserva os nêutrons e constitui uma fonte de plutônio.

Um refrigerante é necessário em todos os reatores exceto os de muito pequena potência, para extrair o calor de fissão continuamente. São indicados para tal fim fluidos com boas características de transmissão de calor e seção de choque de absorção pequena para nêutrons. Os líquidos usados são água leve e pesada; metais líquidos, ligas metálicas de sódio e potássio (NaK) e bismuto. O mercúrio foi usado num reator rápido. Entre os gases: o ar, o dióxido de carbono e o hélio. Os líquidos possuem uma vantagem sobre os gases, devido à maior densidade e capacidade calorífica, mas constituem uma limitação, pois existe o problema da ebulição.

Um reator é classificado em rápido, térmico ou intermediário baseado na energia dos nêutrons causadores da fissão.

Um reator rápido é aquele que tem pouco ou nenhum material moderador e, assim sendo, a reação em cadeia é provocada por nêutrons com energias da ordem de MeVs.

Nos reatores térmicos, nos quais a fissão é devida, principalmente, a nêutrons de baixa energia, o número relativo de átomos de combustível e de moderador é típico 1:400 e montagens críticas são possíveis até 0,5 Kg de U^{235} . A maioria das unidades atualmente em operação se inclui nesta categoria, por causa da probabilidade de uso de U natural e da segurança inerente associada com a dilatação térmica do moderador.

Os reatores intermediários possuem menor moderador e mais combustível do que os reatores térmicos e dependem da fissão na gama epitérmica, na região dos eV para keV.

Desta forma, podem-se classificar os reatores nucleares sob os seguintes itens:

i. Energia dos nêutrons utilizados para fissão

- a) Reatores rápidos - os nêutrons rápidos, ou seja, de alta energia, ($E_n > 100$ keV) causam a maioria das fissões.
- b) Reatores intermediários ou epitérmicos - nêutrons epitérmicos, ou seja, de média energia (0.3 keV $< E_n < 10$ keV) causam a maioria das fissões.
- c) Reatores térmicos - nêutrons térmicos, ou seja, de baixa energia ($E_n < 0.3$ keV) causam a maioria das fissões.

ii. Propósito de funcionamento do reator

- a) Reatores de potência - têm a finalidade de gerar energia. Podem gerar energia elétrica, servir para aquecimento industrial e residencial, servir para propulsão de embarcações, etc.
- b) Reatores de pesquisa - têm a finalidade de pesquisa, irradiação e teste de materiais, produção de radioisótopos, etc.
- c) Reatores de ensino - tem a finalidade de treinamento e ensino.

iii. Tipo de combustível e/ou refrigerante e/ou moderador

- a) reatores refrigerados a gás (moderados a grafite) (GCR), reatores avançados refrigerados a gás (AGR), reatores de alta temperatura refrigerados a gás (HTGR) e reatores rápidos refrigerados a gás (GCFBR) usando Urânio natural ou levemente enriquecido.

- b) reatores refrigerados e moderados a água leve (PWR, BWR), usando urânio enriquecido.
- c) reatores de água pesada (HWR) usando Urânio natural, água pesada como moderador e água leve ou pesada como refrigerante.
- d) reatores rápidos usando Plutônio e Urânio como combustível e metal líquido (Sódio) como refrigerante.

3.1. O Reator PWR

O primeiro reator do tipo PWR (Pressurized Water Reactor) comercial foi desenvolvido pela Westinghouse. Com potência de 250 MW, o Yankee Rowe começou sua operação em 1960 e permaneceu até 1992. No final da década de 60 começou o surgimento das unidades PWR com mais de 100 MW de potência. A França destaca-se na produção de reatores deste tipo, começando em 1956 com modelos gás-grafite, e já entrou na terceira geração desses reatores, cuja estratégia é unir eficiência e custo [56].

Resumidamente segue uma descrição do funcionamento de uma usina nuclear em suas várias etapas [51]:

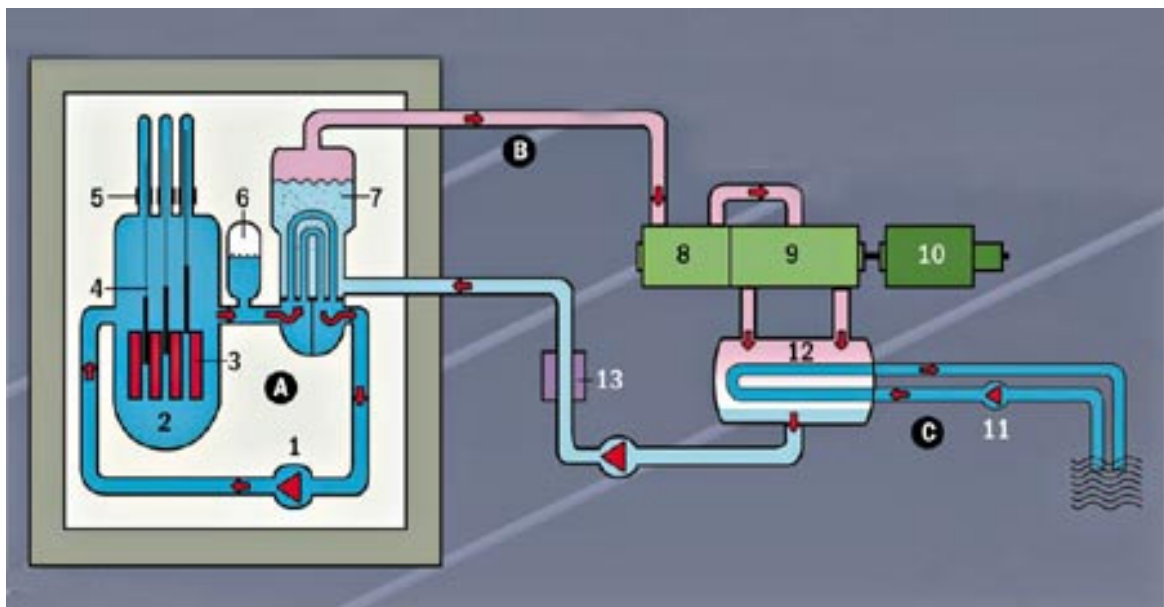


Figura 1. Descrição do funcionamento de um reator do tipo PWR

A bomba de refrigeração, (1) movimentada a água do circuito primário (A) constantemente. O vaso de pressão (2) contém o núcleo do reator com os elementos combustíveis (3). As barras de controle (4 e 5) controlam a taxa de fissão do urânio. A água passa pelo núcleo do reator, e o calor liberado pela fissão do urânio a aquece até

aproximadamente 325 °C, mas o pressurizador (6) impede sua evaporação. A água passa por tubos dentro do gerador de vapor (7), no circuito secundário (B), evaporando o líquido dentro dele. O vapor entra nas turbinas (8 e 9). Lá a energia cinética do vapor transforma-se em energia mecânica pela rotação da turbina, que está acoplada ao gerador elétrico. No gerador (10), essa energia é convertida em eletricidade. No terceiro circuito (C), a água do mar ou rio entra e sai do sistema (11), esfriando o vapor do condensador (12) e transformando-o em água. O líquido é aquecido (13) e volta ao gerador.

Projetado dentro do conceito de "Defesa em Profundidade", o PWR possui um conjunto de barreiras que impedem o escape do material radioativo para o meio ambiente [44].

1. A primeira barreira é a própria natureza cerâmica do combustível - dióxido de urânio - que somente se funde em temperatura superior a 2800°C;
2. A segunda é o revestimento do combustível - uma liga de zircônio que na presença de água resiste à temperatura de 1400°C;
3. A terceira é constituída pelo vaso de pressão, cujas paredes de aço têm espessura de 20 cm em Angra 1 e 25 cm em Angra 2;
4. A quarta barreira é a blindagem radiobiológica;
5. A quinta é constituída pelo vaso de contenção de aço, com 3 cm de espessura, que retém completamente qualquer material radioativo que eventualmente venha a escapar em um acidente, impedindo assim que atinja o meio ambiente.

Externamente, existe uma sexta barreira - o edifício de concreto reforçado, que protege o reator de agentes externos e também contribui para reter o material radioativo.

3.1.1. O Circuito Primário

O circuito primário é responsável pela remoção do calor gerado no núcleo do reator pelo processo de fissão. Em geral, nos sistemas com potência de 1300MW, existem 4 circuitos em paralelo que compõem o circuito primário e estão ligados ao núcleo do reator, que é o local onde está localizado o combustível e ocorre a reação em cadeia. Simplificadamente, existem 4 geradores de vapor e 4 bombas refrigerantes em operação no funcionamento normal do reator.

No núcleo do reator, o combustível é acondicionado em varetas combustíveis que são agrupadas em arranjos formando um elemento combustível. O núcleo é composto por

vários elementos apresentando arranjos chamados de barras de controle, que são constituídas de venenos neutrônicos, como por exemplo, carbeto de boro, que tem como objetivo controlar o processo de fissão. Essas barras podem ser movimentadas verticalmente, existindo vários grupos com movimentação independente; dependendo do número de barras de controle inseridas tem-se maior ou menor potência, sendo a inserção de todos os grupos a metodologia para “desligamento” do reator.

No PWR, a água é usada como refrigerante do núcleo e como moderador, sendo por isso o reator do tipo térmico. Existem também no núcleo materiais que se destinam a minimizar a saída de nêutrons do reator. Estes materiais, chamados de refletores servem para diminuir o valor da massa crítica para o sistema.

Os geradores de vapor são trocadores de calor multi-tubulares tendo o feixe de tubos a forma de U invertido . O refrigerante do circuito primário flui internamente aos tubos e o fluido de trabalho, que também é a água, externamente a estes. Os tubos são projetados para minimizar a possibilidade de contaminação da água do circuito secundário com a água radioativa que serve como refrigerante do circuito primário.

O circuito primário fica contido dentro de uma estrutura de contenção que é projetada para resistir a sinistros, como terremotos de média e grande magnitude e até mesmo o choque de aeronaves de grande porte, além de pressões internas. Tais precauções objetivam minimizar a possibilidade de vazamento e contaminação do meio ambiente periférico com material radioativo.

Para manter a pressão no circuito primário e controlar as alterações volumétricas ocorridas no refrigerante existe um equipamento ou componente chamado pressurizador, ligado a um dos 4 circuitos. O pressurizador é um vaso de pressão que possui cerca de 60% de seu conteúdo na forma líquida e 40% na forma de vapor, contendo um volume que representa 7% a 10% do circuito primário. Na sua parte inferior existem resistências elétricas e na parte superior um conjunto de “sprays”. As resistências elétricas mantêm a água em estado saturado. Se o volume de refrigerante aumentar (devido a um aumento na potência do reator ou queda de tensão nas bombas de refrigeração do núcleo, por exemplo) os “sprays” automaticamente borrifam água fria no pressurizador, provocando uma condensação do vapor nele existente, diminuindo assim a pressão. Se ocorrer uma contração de volume do refrigerante o conteúdo de líquido no pressurizador é reduzido, causando uma diminuição na pressão dentro do vaso, o que causa a vaporização instantânea de água de modo a estabilizar a pressão ou diminuir a magnitude de sua variação.

3.1.2. O Circuito Secundário

Em reatores de grande porte existem também 4 circuitos secundários. Simplificadamente, cada um dos 4 circuitos possui, além dos geradores de vapor, uma bomba de circulação, um condensador e uma turbina a vapor com respectivo gerador de energia acoplado.

O vapor produzido no gerador de vapor faz girar a turbina a vapor que aciona o gerador elétrico acoplado, produzindo energia elétrica. O vapor de saída da turbina é condensado e o líquido bombeado para o gerador de vapor, completando o ciclo.

3.2. O reator IEA-R1

O Reator de Pesquisas IEA-R1 do IPEN é um reator do tipo piscina, refrigerado e moderado a água leve, que usa berílio e grafite como refletor. Sua primeira criticalidade foi atingida em 16 de setembro de 1957. Desde esta data ele foi operado a 2 MW de potência, e a partir de 1997 sofreu um processo de modernização para poder operar a níveis mais altos de potência.

Para propiciar o novo ciclo de operação do reator IEA-R1 foram necessárias alterações no sistema de refrigeração do reator, sistemas de combate a incêndio, sistema de ventilação e ar-condicionado, sistemas elétricos, e sistemas de monitoração da radiação. Um sistema de resfriamento de emergência baseado em spray passivo também foi instalado. Por esse motivo, a implantação de um sistema de monitoração e detecção de falhas no reator IEA-R1 torna-se importante como parte de uma estratégia da gestão do envelhecimento da instalação.

As finalidades do reator são [26] [27]:

- Produção de radioisótopos para aplicações na medicina, indústria, agricultura e para pesquisas;
- Servir como uma fonte intensa de nêutrons em experimentos científicos de física nuclear, química, engenharia e biologia;
- Treinamento do corpo científico do IPEN-CNEN/SP em física de reatores, projetos, desenvolvimento de instrumentação nuclear e segurança de reatores;
- Treinamento e formação de operadores de reator;
- Análise radioquímica de amostras através do método de análise por ativação.

O prédio do reator IEA-R1 é composto por cinco pavimentos, assim distribuído [18]:

- Subsolo: onde se localiza a Casa das Máquinas;
- 1º andar: Salão de Experimentos;
- 2º andar: Sala de Ventilação e Ar-Condicionado, Almoxxarifados e o acesso principal do prédio;
- 3º andar: Saguão da Piscina do Reator, Sala de Controle e Oficinas de Apoio;
- 4º andar: Sala de Exaustão, com filtros e chaminé do Sistema de Exaustão do Reator.

O núcleo do reator encontra-se dentro da piscina a 6,89 metros da superfície da água (da altura do ladrão à parte superior dos elementos combustíveis) e está montado em uma placa matriz com oitenta orifícios, na qual é possível a disposição dos combustíveis em vários arranjos experimentais. Esta placa é sustentada por uma treliça conectada a uma plataforma móvel, onde estão montados os quatro mecanismos de acionamento das barras e detectores que enviam sinais para a mesa de controle.

Oito tubos colimadores com diâmetros de 6 e 8 polegadas dispostos radialmente ao núcleo do reator e um tubo tangencial à face sul do núcleo são usados para experimentos com nêutrons. Outros dois tubos colimadores estão colocados em frente à coluna térmica.

A piscina possui um volume de água de 272 m^3 , sendo dividida em dois compartimentos. O primeiro destina-se à estocagem, manuseio de material radioativo e elementos combustíveis, e o segundo contém o núcleo do reator e é destinado à operação. A piscina é revestida internamente por chapas de aço inoxidável, sendo que a blindagem radiológica que envolve o núcleo é feita por uma camada de água de 2 metros de espessura na região lateral, e até 2,4 metros de concreto com barita nas paredes da piscina.

Os circuitos de resfriamento do núcleo foram construídos com redundância, sendo que apenas um circuito é suficiente para remover a potência nominal de operação do reator a 5 MW. A forma com que foram projetados permite alternar circuitos e alinhamentos dos seus diversos componentes.

Cada circuito de resfriamento consiste de:

- Uma bomba do circuito primário;
- Uma bomba do circuito secundário;
- Um trocador de calor;
- Uma torre de resfriamento do circuito secundário;

- Válvulas de fechamento automático posicionadas na saída e no retorno da água isolam a piscina do circuito primário em caso de perda do refrigerante. O sistema de resfriamento de emergência permite que haja resfriamento do núcleo em caso de exposição do núcleo.

Um tanque de decaimento localizado no circuito primário de resfriamento funciona como um retardo para o retorno da água para a piscina, possibilitando o decaimento do ^{16}N formado na água através da reação $^{16}\text{O}(n,p)^{16}\text{N}$.

O elemento combustível usado é do tipo MTR (Material Testing Reactor) com 18 placas planas contendo urânio enriquecido a 20%.

O controle de reatividade do reator é feito através de três barras de segurança e uma barra de controle, do tipo placa, as quais são movimentadas no núcleo por um mecanismo de acionamento constituído de motores síncronos. As barras são sustentadas por magnetos que são desligados no caso de algum evento não usual, fazendo-as cair por ação da gravidade, desligando o reator.

Dois câmaras de ionização não compensadas, uma câmara de ionização compensada e uma câmara de fissão fazem parte da instrumentação nuclear, que enviam sinais para a mesa de controle localizada na sala de controle. Também na sala de controle estão instalados os instrumentos de apoio à operação, abaixo relacionados:

- Monitores de radiação de área, dutos e contaminação de ar;
- Sistema de Alarmes de Radiação;
- Comando das bombas dos circuitos primário, secundário e ventiladores das torres de resfriamento;
- Comandos do Sistema de Ventilação e Exaustão de Ar-Condicionado;
- Alarme de incêndio;
- Comando do Sistema de Resfriamento de Emergência;
- Comando das válvulas de isolamento do circuito primário;
- Iluminação de emergência;
- Indicação de condutividade de água da piscina e da água de reposição;
- Sistema de análise de vibração dos mancais dos volantes de inércia;
- Comunicação interna e externa.

O Sistema de Ventilação e Ar-Condicionado (SVAC) mantém o prédio com pressão ligeiramente negativa nas áreas com maior risco de contaminação, tais como:

Saguão da Piscina, Casa de Máquinas e Salão de Experimentos. Estas áreas estão isoladas do restante do prédio por antecâmaras.

A água da piscina do reator é mantida em boa qualidade por sistemas de resinas trocadoras de íons, consistindo nos Sistemas de Tratamento e Retratamento da Água. O Sistema de Tratamento é utilizado para completar o nível de água da piscina com a pureza necessária aos níveis operacionais. O Sistema de retratamento opera continuamente e tem como finalidade a manutenção da qualidade da água da piscina de maneira a minimizar os efeitos de corrosão e os níveis de dose de radiação na superfície da piscina.

O fornecimento de energia elétrica é feito por meio da rede elétrica da concessionária local. No caso de uma eventual interrupção no fornecimento, um conjunto de quatro geradores fornece energia elétrica aos sistemas vitais e essenciais para manter o reator em operação.

Os principais sistemas do reator enviam sinais à mesa de controle, onde por meio de uma cadeia de relés, compõem o Sistema de intertravamento do circuito de segurança. Este circuito está ligado diretamente ao Sistema de Desligamento Automático do Reator (SCRAM), de maneira que, se alguma anormalidade ocorrer durante o funcionamento do reator, haverá a abertura do respectivo contato do relé e conseqüente interrupção da corrente elétrica que alimenta os magnetos que sustentam as barras de controle e de segurança, provocando a queda destas por gravidade e o conseqüente desligamento do reator.

Na Figura 2, tem-se um fluxograma de processo do reator IEA-R1.

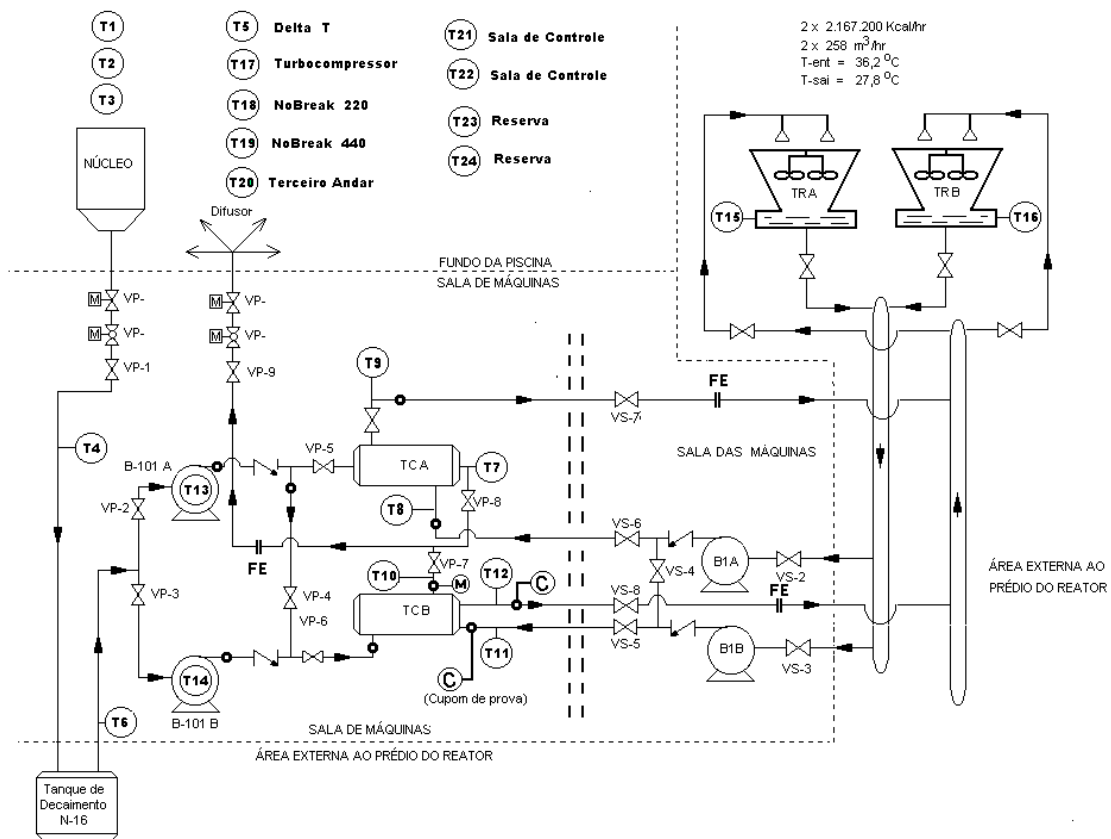


Figura 2. Fluxograma de processo do reator IEA-R1.

3.3. Considerações Físicas dos Reatores

O reator de potência é termodinamicamente acoplado ao restante do sistema de potência e este acoplamento envolve as características termodinâmicas do fluido refrigerante e o desempenho dos componentes do sistema. O fluxo de nêutrons está intimamente ligado à produção de calor ou potência, isto é, para uma taxa constante de geração de energia, a taxa de fissão deve ser constante requerendo, para isso, que um nêutron de cada fissão produza uma nova fissão, ou seja, a potência gerada pelo reator é proporcional à taxa de fissões [43] [63].

Em um dado reator a taxa de fissão é proporcional a quantidade de material fissionável presente e à densidade neutrônica média. Assim, uma mudança na densidade neutrônica média resulta diretamente em uma mudança no nível de potência do reator [60].

A mais direta e imediata quantidade que indica o desempenho do reator é a densidade de nêutrons e, então, o controle do reator é executado pela regulação desta quantidade. Tal controle da quantidade de nêutrons pode ser feito:

(1) adicionando material absorvedor, tal como boro em áreas estratégicas do reator. Em geral, as barras (compostas de háfnio, cuja secção de choque para absorção é relativamente elevada) podem ser usadas para capturar nêutrons bem como elevar e baixar a potência, conforme se deseje;

(2) utilizando material refletor;

(3) adicionando ou retirando material fissionável do reator.

A transferência do calor gerado durante a fissão pode ser feita para um fluido refrigerante e o princípio físico empregado nesta transferência são as leis da condução e convecção.

Em consequência da lei da conservação de energia, o calor que é produzido nos elementos combustíveis faz crescer a temperatura do mesmo e a transferência de calor da superfície do combustível (ou de sua blindagem) para o refrigerante ocorre por combinação de convecção e condução conforme a natureza do refrigerante, ou seja, conforme calor específico, condutividade e velocidade [43].

As principais fontes de radiação num reator são:

1 – Nêutrons prontos provenientes da fissão;

2 – Raios gama prontos são consequentes diretos da fissão;

3 - Raios gama provenientes da desintegração dos produtos da fissão que são formados depois da fissão;

4 – A atividade induzida no refrigerante pelos nêutrons do reator que, também deve ser considerada para a blindagem do núcleo, porém é mais importante na parte externa do sistema de refrigeração do próprio reator;

5 – Raios gama decorrentes da captura de nêutrons, pois possuem energias nas proximidades da energia de ligação de um nêutron de 6 a 8 MeV sendo altamente penetrantes;

6 – A atividade gama dos radioisótopos que é oriunda de vários elementos de um reator bem como de qualquer material usado para irradiação.

Levando em consideração que pequenas camadas de metal ou de plástico são suficientes para barrar as partículas alfa e beta, das diversas radiações que são produtos diretos ou indiretos da fissão, os nêutrons e os raios gama são os de maior significado sob o ponto de vista da blindagem, pois o nível das radiações gama e de nêutrons próximo ao núcleo de um reator de potência é bastante alto para desenvolver um considerável calor nos materiais da estrutura, tais como o vaso de pressão de um reator moderado a água. Os raios gama são absorvidos e os nêutrons cedem energia por colisões elásticas e inelásticas. Além

disso, os raios gama proveniente da captura de nêutrons térmicos contribuem para o aquecimento.

A composição inicial de um reator é a reunião de certa quantidade de combustível e moderador dentro da região do núcleo, de modo a tornar o sistema super crítico. Durante a utilização prática, existem várias operações e efeitos naturais que fazem com que a constante de multiplicação de nêutrons (K) difira da unidade, o que exige que o K seja alterado sem um ajuste laborioso do combustível e do moderador como, por exemplo, a fim de se alcançar uma elevada potência e operação, o valor de K deve, temporariamente, tornar-se maior do que 1, permitindo que a densidade e fluxo de nêutrons desejados sejam estabelecidos. O método de controle do fator de multiplicação de um reator mais freqüentemente encontrado é o das barras de controle móveis constituídas de elementos (tais como cádmio, boro ou háfnio) que absorvem nêutrons térmicos. Sua função, quando inseridas no núcleo, é a retirada contínua de nêutrons que ajudariam na continuação da reação.

O reator é controlado por um conjunto de barras, compostas de háfnio, cuja secção de choque para absorção é relativamente elevada. O ajuste da profundidade de inserção no núcleo do reator modifica a área exposta à absorção de nêutrons, o que proporciona compensação para o consumo de combustível e para a elevação da temperatura na partida. Em geral, as barras podem ser usadas para elevar e baixar a potência, conforme se deseje.

São três tipos de barras, segundo as suas principais funções:

- (1) Barras de segurança que são, geralmente, mantidas fora do núcleo, prontas para serem inseridas rapidamente para estancar uma elevação súbita de potência devida a outras alterações no sistema do reator.
- (2) Barras de controle, que são leves e pouco absorvedoras e utilizadas para pequenos ajustes manuais ou automáticos da potência.
- (3) Barras de controle grosseiro, que são usadas na partida de um reator para compensar variações da reatividade a longo termo devido ao acúmulo de produtos de fissão e ao consumo de combustível.

4. CORRELAÇÃO LINEAR

4.1. Equações lineares e o modelo de regressão linear

Dados os números reais $\alpha_1, \alpha_2, \dots, \alpha_n, \beta$ ($n \geq 1$), à equação

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = \beta$$

onde os x_i são variáveis reais, damos o nome de equação linear nas incógnitas x_1, \dots, x_n .

Suponha que dadas duas variáveis X e Y cada uma com n observações. Cada par (X_i, Y_i) , $1 \leq i \leq n$, pode ser representado por um ponto no gráfico e plotado num diagrama de dispersão [34].

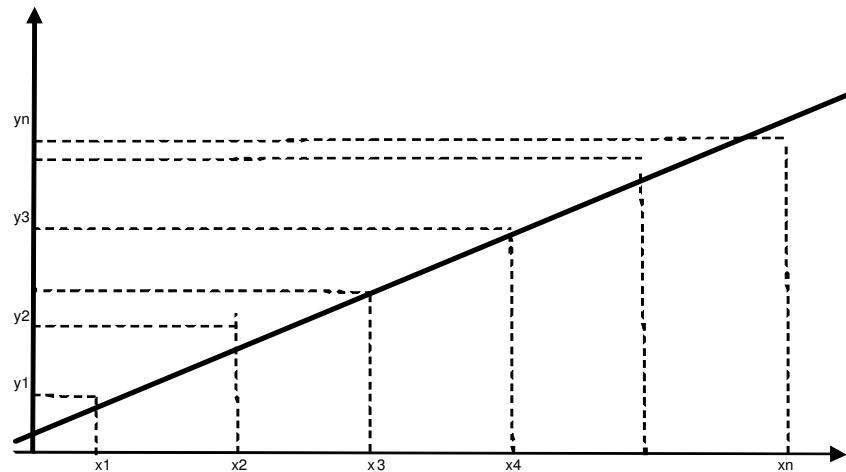


Figura 3: Gráfico de dispersão

A disposição dos pontos sugere um comportamento linear e, portanto, podemos traçar uma reta que imite o comportamento dos pontos.

Segue do gráfico acima que para um dado X_i temos os pares (X_i, Y_i) e (X_i, \hat{Y}_i) onde Y_i e \hat{Y}_i representam o valor observado (valor real) e o valor predito (valor pertencente à reta) respectivamente.

Chamando $\hat{Y} = a + bX$ como reta de regressão, nosso objetivo é encontrar os parâmetros a e b tal que tal que o $|Y_i - \hat{Y}_i|$ seja minimizado.

Portanto, tomando a somatória dos quadrados das diferenças entre o valor observado e o valor predito pela equação, temos:

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Substituindo \hat{Y}_i desta equação por $\hat{Y}_i = a + bX_i$ e chamando de G esta diferença, temos:

$$G = \sum_{i=1}^n [Y_i - (a + bX_i)]^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2 \quad (4.1.1)$$

Agora, tomando o quadrado, à direita, desta expressão:

$$G = \sum_{i=1}^n [Y_i^2 + a^2 + b^2 X_i^2 - 2aY_i - 2bX_i Y_i + 2abX_i]$$



$$G = \sum_{i=1}^n Y_i^2 + na^2 + b^2 \sum_{i=1}^n X_i^2 - 2a \sum_{i=1}^n Y_i - 2b \sum_{i=1}^n X_i Y_i + 2ab \sum_{i=1}^n X_i \quad (4.1.2)$$

Para encontrar os valores de a e b que minimizam G tomamos suas derivadas parciais em relação aos parâmetros a e b de 4.1.2:

$$\frac{\partial G}{\partial a} = 2na - 2 \sum_{i=1}^n Y_i + 2b \sum_{i=1}^n X_i$$

$$\frac{\partial G}{\partial b} = 2b \sum_{i=1}^n X_i^2 - 2 \sum_{i=1}^n X_i Y_i + 2a \sum_{i=1}^n X_i$$

Igualando cada uma das expressões anteriores a zero e rearranjando-as segue que:

$$na + b \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \quad (4.1.3)$$

$$a \sum_{i=1}^n X_i + b \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \quad (4.1.4)$$

As equações 4.1.3 e 4.1.4 devem ser resolvidas simultaneamente como um sistema de equações cujo objetivo é encontrar os valores de a e b.

Multiplicando 4.1.3 por $\sum X$ e 4.1.4 por n,

$$na \sum_{i=1}^n X_i + b \left(\sum_{i=1}^n X_i \right)^2 = \sum_{i=1}^n X_i \sum_{i=1}^n Y_i \quad (*)$$

$$na \sum_{i=1}^n X_i + nb \sum_{i=1}^n X_i^2 = n \sum_{i=1}^n X_i Y_i \quad (**)$$

Subtraindo (*) de (**) temos:

$$nb \sum_{i=1}^n X_i^2 - b \left(\sum_{i=1}^n X_i \right)^2 = n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i$$

E isolando a incógnita b, temos:

$$b = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{\sum_{i=1}^n X_i Y_i - \frac{\sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n}}{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i \right)^2}{n}} \quad 4.1.5$$

A expressão para “a” pode ser obtida substituindo “b” em 4.1.3 para obter:

$$a = \bar{Y} + b \bar{X}$$

Por outro lado, adotando a seguinte notação:

$$S_{xx} = \sum (X - \bar{X})^2 = \frac{n \sum X^2 - \left(\sum X \right)^2}{n} = \sum X^2 - \frac{\left(\sum X \right)^2}{n}$$

$$S_{yy} = \sum (Y_i - \bar{Y})^2 = \frac{n \sum Y^2 - (\sum Y)^2}{n} = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$S_{xy} = \sum (X_i - \bar{X})(Y_i - \bar{Y}) = \frac{n \sum XY - \sum X \sum Y}{n} = \sum XY - \frac{\sum X \sum Y}{n}$$

Temos:

$$b = \frac{S_{xy}}{S_{xx}}$$

A equação desenvolvida para prever Y de X é chamada equação de regressão, reta de regressão ou reta de mínimos quadrados. Tem a forma geral $Y = a + bX$ e as constantes a e b na equação de regressão são chamados coeficientes de regressão [34] [50].

4.2. Coeficiente de Correlação de Pearson

O coeficiente de correlação mede o quanto duas variáveis estão relacionadas, ou seja, o grau de dependência entre as mesmas.

Aqui mostraremos como o coeficiente de correlação está relacionado à qualidade do ajuste linear de pontos por uma reta de regressão.

Considerando a identidade $Y_i - \bar{Y} = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y})$ e tomando o quadrado de

ambos os lados da equação temos:

$$(Y_i - \bar{Y})^2 = (Y_i - \hat{Y}_i)^2 + (\hat{Y}_i - \bar{Y})^2 + 2(Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})$$

Tomando o somatório do último termo do lado direito da equação e fazendo

$$\hat{Y}_i = \bar{Y} + b(X_i - \bar{X})$$

temos:

$$\sum (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y}) = \sum (Y_i - \hat{Y}_i)b(X_i - \bar{X})$$

Segue que,

$$\sum (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y}) = b \sum [(Y_i - \hat{Y}_i)b(X_i - \bar{X})](X_i - \bar{X}) \Leftrightarrow b \sum [(Y_i - \bar{Y})(X_i - \bar{X}) - b(X_i - \bar{X})^2]$$



$$bS_{xy} - b^2S_{xx} = \frac{S_{xy}^2}{S_{xx}} - \frac{S_{xy}^2}{S_{xx}} = 0$$

Assim,

$$\sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

Como cada um dos termos desta equação é maior ou igual a zero, temos que:

$$\sum (Y_i - \bar{Y})^2 \geq \sum (Y_i - \hat{Y}_i)^2$$

e

$$\sum (Y_i - \bar{Y})^2 \geq \sum (\hat{Y}_i - \bar{Y})^2$$

Dividimos ambos os lados da equação por $\sum (Y_i - \bar{Y})^2$ obtemos:

$$1 = \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} + \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

e, daí ,

$$1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

Considerando o lado esquerdo da equação anterior temos que, [34] [50]:

- Se $\sum (Y_i - \hat{Y}_i)^2 = 0$ (se não houvesse erro de estimativa) então o lado esquerdo da equação seria igual a 1 e, em tal caso, deveríamos ter uma perfeita relação entre X e Y. Por outro lado, se não existisse relação linear entre X e Y, então a melhor avaliação disponível para Y deveria ser \bar{Y} , isto é, $\hat{Y} = \bar{Y}$ e o lado esquerdo da igualdade seria igual a zero.

- Consideremos agora o membro da direita. Se não existisse relação linear entre X e Y e cada estimativa de Y fosse igual a \bar{Y} , então este termo deveria ser igual a zero. Se cada estimativa fosse igual ao valor real de Y, isto é, igual a Y_i , então o numerador deste termo deveria ser igual a $\sum (Y_i - \bar{Y})^2$ e a expressão da direita deveria ser igual a 1 indicando uma perfeita relação linear entre X e Y.

Cada lado da expressão dada define o quadrado do coeficiente de correlação de Pearson, r. Assim,

$$r^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (4.2.1)$$

ou

$$r = \pm \sqrt{1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2}} = \pm \sqrt{\frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}}$$

As equações de r e r^2 têm as seguintes propriedades:

1. O valor de r é igual a zero quando não existe relação linear entre X e Y.

Neste caso $\sum (Y_i - \hat{Y}_i)^2$ tem valor maximizado, $\sum (Y_i - \bar{Y})^2$, e $\sum (\hat{Y}_i - \bar{Y})^2$ tem valor mínimo igual a zero.

2. Os possíveis valores de r variam de -1 a +1. Os valores de $r = 1$ ou $r = -1$

indicam perfeita relação linear entre X e Y. Ou seja, $\sum (Y_i - \hat{Y}_i)^2 = 0$ e

$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2$. Assim, $r^2 = 1$ e, então $r = 1$ ou $r = -1$

Desta forma, temos que o coeficiente de correlação de Pearson é uma medida de dependência linear entre duas variáveis e pode, também, ser entendido como uma medida do ajuste linear de pontos a uma reta, ou seja, quanto mais próximo dos valores 1 ou -1 melhor o ajuste.

4.2.1. Fórmula alternativa para o coeficiente de correlação

Alternativamente temos a seguinte fórmula para r:

Como $\hat{Y}_i = a + bX_i$, podemos escrever

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - a - bX_i)^2$$

Substituindo a por $\bar{Y} - b\bar{X}$

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y} + b\bar{X} - bX_i)^2 = \sum_{i=1}^n [(Y_i - \bar{Y}) - b(X_i - \bar{X})]^2$$

Desenvolvendo o quadrado da última igualdade

$$\begin{aligned} \sum (Y_i - \hat{Y}_i)^2 &= \sum_{i=1}^n [(Y_i - \bar{Y})^2 - b^2 (X_i - \bar{X})^2 - 2b(X_i - \bar{X})(Y_i - \bar{Y})] = \\ &= \sum_{i=1}^n [(Y_i - \bar{Y})^2 + b^2 \sum (X_i - \bar{X})^2 - 2b \sum (X_i - \bar{X})(Y_i - \bar{Y})] \end{aligned}$$

Temos:

$$\sum (Y_i - \hat{Y}_i)^2 = S_{yy} + b^2 S_{xx} - 2b S_{xy}$$

Como $b = \frac{S_{xy}}{S_{xx}}$, segue que:

$$\sum (Y_i - \hat{Y}_i)^2 = S_{yy} + \frac{S_{xy}^2 S_{xx}}{S_{xx}^2} - 2 \frac{S_{xy}^2}{S_{xx}} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

Usando agora esta última igualdade e o fato que $S_{yy} = \sum (Y_i - \bar{Y})^2$ podemos substituir $\sum (Y_i - \hat{Y}_i)^2$ e S_{yy} em 4.2.1 para obter

$$r = \pm \sqrt{1 - \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{S_{yy}}} = \pm \sqrt{\frac{S_{xy}^2}{S_{xx} S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sqrt{S_{xx}}} \right) \left(\frac{y_i - \bar{y}}{\sqrt{S_{yy}}} \right) \quad 4.2.2$$

5. CORRELAÇÃO MÚLTIPLA

5.1. Análise de Correlação Canônica

A Análise de Correlação Canônica é uma técnica para a identificação e quantificação da associação entre dois grupos de variáveis e pode ser entendida como uma extensão da Correlação Linear entre duas variáveis [34] [50].

5.1.1. A Matemática da correlação canônica

Tem-se interesse em medir a associação entre os dois grupos de variáveis. O primeiro grupo de p variáveis é representado pelo vetor aleatório X de dimensão $(p \times 1)$ e o segundo de q variáveis Y , $(q \times 1)$, sendo $p \leq q$, ou seja,

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_p \end{pmatrix} \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_q \end{pmatrix}$$

Mostremos agora como esta correlação pode ser calculada e para isso será utilizada a técnica mostrada em [22] [23] [34].

Consideremos duas variáveis Z_x e Z_y onde a primeira é uma combinação linear de p e a outra de q variáveis. Isto é:

$$Z_x = u_1 X_1 + \dots + u_p X_p$$

e

$$Z_y = v_1 Y_1 + \dots + v_q Y_q$$

em que os X 's são as p variáveis em um conjunto e os Y 's são a q variáveis no outro. Os u 's e v 's são pesos nas combinações.

Considerando Z_x , podemos mostrar que a média \bar{Z}_x é igual a:

$$\bar{Z}_x = u_1 \bar{X}_1 + \dots + u_p \bar{X}_p$$

Definimos agora

$$\begin{aligned} z_x &= Z_x - \bar{Z}_x = u_1 X_1 + \dots + u_p X_p - u_1 \bar{X}_1 + \dots + u_p \bar{X}_p = \\ &u_1 (X_1 - \bar{X}_1) + \dots + u_p (X_p - \bar{X}_p) \end{aligned}$$

Tomando o quadrado de z_x :

$$\begin{aligned} \sum z_x^2 &= u_1^2 \sum (X_1 - \bar{X}_1)^2 + \dots + u_p^2 \sum (X_p - \bar{X}_p)^2 + \\ &+ 2u_1 u_2 \sum (X_1 - \bar{X}_1)(X_2 - \bar{X}_2) + \dots + 2u_{p-1} u_p \sum (X_{p-1} - \bar{X}_{p-1})(X_p - \bar{X}_p) = \\ &= u_1^2 S_{11} + \dots + u_p^2 S_{pp} + 2u_1 u_2 S_{12} + \dots + 2u_{p-1} u_p S_{p-1,p} \end{aligned}$$

Em que

$$S_{jk} = \sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k) \quad \text{tal que } j, k = 1, \dots, p$$

Agora, fazendo

$$u = \begin{pmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_p \end{pmatrix} \quad \text{e} \quad S_{xx} = \begin{pmatrix} S_{11} & S_{12} & \cdot & \cdot & \cdot & S_{1p} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ S_{p1} & S_{p2} & & & & S_{pp} \end{pmatrix}$$

Temos que a expressão para $\sum z_x^2$ pode ser escrita como

$$\sum z_x^2 = u' S_{xx} u \quad (5.1.1.1)$$

E de forma similar

$$\sum z_y^2 = v' S_{yy} v \quad (5.1.1.2)$$

e

$$\sum z_x z_y = u' S_{xy} v \quad (5.1.1.3)$$

onde

$$v = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_p \end{pmatrix} \quad S_{yy} = \begin{pmatrix} S_{11} & S_{12} & \cdot & \cdot & \cdot & S_{1p} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ S_{q1} & S_{q2} & \cdot & \cdot & \cdot & S_{pp} \end{pmatrix} \quad S_{xy} = \begin{pmatrix} S_{x_1 y_1} & S_{x_1 y_2} & \cdot & \cdot & \cdot & S_{x_1 y_q} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ S_{x_p y_1} & S_{x_p y_2} & \cdot & \cdot & \cdot & S_{x_p y_q} \end{pmatrix}$$

Na matriz S_{yy} , $S_{jk} = \sum_{i=1}^n (Y_{ij} - \bar{Y}_j)(Y_{ik} - \bar{Y}_k)$; $j, k = 1, \dots, q$. Em S_{xy} ,

$$S_{x_j y_k} = \sum_{i=1}^n (X_{ij} - \bar{X}_j)(Y_{ik} - \bar{Y}_k) \quad j=1, \dots, p \quad e \quad k=1, \dots, q. \text{ Agora usando as equações}$$

5.1.1.1, 5.1.1.2 e 5.1.1.3, o coeficiente de correlação $r_{z_x z_y}$ pode ser escrito como

$$r_{z_x z_y} = \frac{\sum z_x z_y}{\sqrt{\sum z_x^2 \sum z_y^2}} = \frac{u' S_{xy} v}{\sqrt{(u' S_{xx} u)(v' S_{yy} v)}} \quad (5.1.1.4)$$

No caso $p = 2$, o coeficiente de correlação entre duas variáveis X_1 e X_2 é:

$$r_{12} = \frac{S_{12}}{\sqrt{S_{11} S_{22}}}$$

E, além disso,

$$r_{11} = \frac{S_{11}}{\sqrt{S_{11}S_{11}}} = r_{22} = \frac{S_{22}}{\sqrt{S_{22}S_{22}}} = 1$$

Queremos agora um método computacional para o cálculo do coeficiente de correlação que não dependa de u e v .

$$\text{Seja } r^2_{z_x z_y} = \frac{(u' S_{xy} v)^2}{(u' S_{xx} u)(v' S_{yy} v)} \quad (*)$$

Diferenciando (*) em relação a u e igualando a zero temos:

$$\frac{\partial r^2_{z_x z_y}}{\partial u} = \frac{(u' S_{xx} u)(v' S_{yy} v)(2u' S_{xy} v)(S_{xy} v) - (u' S_{xy} v)^2 (v' S_{yy} v)(2S_{xx} u)}{(u' S_{xx} u)^2 (v' S_{yy} v)^2} = 0 \quad (5.1.1.5)$$

Multiplicando 5.1.1.5 por $(u' S_{xx} u)(v' S_{yy} v)$ e dividindo por 2:

$$(u' S_{xy} v) S_{yx} v - \frac{(u' S_{xy} v)^2}{(u' S_{xx} u)(v' S_{yy} v)} (v' S_{yy} v) S_{xx} u = 0 \Leftrightarrow (u' S_{xy} v) S_{yx} v - r^2 (v' S_{yy} v) S_{xx} u = 0$$

Diferenciando (*) em relação a v , de modo análogo teremos:

$$(u' S_{xy} v) S_{xy} u - r^2 (u' S_{xx} u) S_{yy} v = 0$$

Resolvendo v em termos de u nesta equação:

$$v = \frac{(u' S_{xy} v)}{r^2 (u' S_{xx} u)} S_{yy}^{-1} S_{yx} u$$

E substituindo em (*) resulta:

$$(S_{xy} S_{yy}^{-1} S_{yx} - r^2 S_{xx}) u = 0 \Leftrightarrow r^2 = S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx}$$

Considerando as p variáveis no conjunto X e q variáveis no conjunto Y , definimos:

$$D_x = \begin{pmatrix} \sqrt{S_{x_1x_1}} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sqrt{S_{x_px_p}} \end{pmatrix} \quad D_y = \begin{pmatrix} \sqrt{S_{y_1y_1}} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sqrt{S_{y_qy_q}} \end{pmatrix}$$

$$R_{xx} = \begin{pmatrix} 1 & r_{x_1x_2} & r_{x_1x_3} & \dots & r_{x_1x_p} \\ r_{x_2x_1} & 1 & & & r_{x_2x_p} \\ & & \cdot & & \\ & & & \cdot & \\ r_{x_px_1} & r_{x_px_2} & & & 1 \end{pmatrix}$$

$$R_{yy} = \begin{pmatrix} 1 & r_{y_1y_2} & r_{y_1y_3} & \dots & r_{y_1y_q} \\ r_{y_2y_1} & 1 & & & r_{y_2y_p} \\ & & \cdot & & \\ & & & \cdot & \\ r_{y_py_1} & r_{y_py_2} & & & 1 \end{pmatrix}$$

$$R_{xy} = \begin{pmatrix} r_{X_1Y_1} & r_{X_1Y_2} & r_{X_1Y_3} & \cdots & r_{X_1Y_q} \\ r_{X_2Y_1} & r_{X_2Y_2} & & & r_{X_2Y_q} \\ & & \cdot & & \\ & & & \cdot & \\ r_{X_pY_1} & r_{X_pY_2} & & & r_{X_pY_q} \end{pmatrix}$$

Então podemos reescrever 5.1.1.1, 5.1.1.2 e 5.1.1.3 como

$$\sum z_x^2 = u' D_x R_{xx} D_x u \quad 5.1.1.6$$

$$\sum z_y^2 = v' D_y R_{yy} D_y v \quad 5.1.1.7$$

$$\sum z_x z_y = u' D_x R_{xy} D_y v \quad 5.1.1.8$$

Daí, fazendo:

$$S_{xx} = D_x R_{xx} D_x \quad , \quad S_{yy} = D_y R_{yy} D_y \quad , \quad S_{xy} = D_x R_{xy} D_y$$

Desta forma pode ser mostrado, conforme [40] [41] [42] que o quadrado do coeficiente de correlação entre Z_x e Z_y pode ser determinado pela equação

$$R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx} - r^2 I = 0 \Rightarrow r^2 = R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx} \quad (5.1.1.9)$$

onde I é a matriz identidade e λ é o quadrado do coeficiente de correlação entre Z_x e Z_y e o número de raízes não nulas (valores de r^2) obtidas na equação é igual a p se $p \leq q$ ou q se $q \leq p$ tal que o maior valor encontrado é a máxima correlação obtida entre as variáveis Z_x e Z_y .

Em particular, se

$$Y = (Y_1) \quad \text{e} \quad X = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_p \end{pmatrix} \quad (5.1.1.10)$$

temos:

$$Z_y = v_1 Y_1 \quad (\text{com } v_1=1) \quad \text{e} \quad Z_x = u_1 X_1 + \dots + u_p X_p$$

onde queremos encontrar a máxima correlação entre o vetor de variáveis Y e o vetor de variáveis X. Assim, temos:

$$R_{xy} = \begin{pmatrix} r_{X_1 Y_1} \\ r_{X_2 Y_1} \\ \cdot \\ \cdot \\ r_{X_p Y_1} \end{pmatrix} \quad R_{yx} = \begin{pmatrix} r_{Y_1 X_1} & r_{Y_1 X_2} & \dots & r_{Y_1 X_p} \end{pmatrix}$$

Daí, temos agora uma forma de calcular r^2 em termos, somente, de correlações bivariadas. Ou seja:

$$(1) \begin{pmatrix} r_{Y_1 X_1} & r_{Y_1 X_2} & \dots & r_{Y_1 X_p} \end{pmatrix} \begin{pmatrix} 1 & r_{x_1 x_2} & r_{x_1 x_3} & \dots & r_{x_1 x_p} \\ r_{x_2 x_1} & 1 & & & r_{x_2 x_p} \\ & & \cdot & & \\ & & & \cdot & \\ r_{x_p x_1} & r_{x_p x_2} & & & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_{X_1 Y_1} \\ r_{X_2 Y_1} \\ \cdot \\ \cdot \\ r_{X_p Y_1} \end{pmatrix} = r^2 I = r^2$$

Neste caso, o número de raízes será 1, pois $q = 1$.

O número r^2 é o quadrado do coeficiente de correlação entre Y e a combinação linear das variáveis do conjunto X e, além disso, expressa a proporção do total da variância calculada pela relação linear entre Y e X. Outra propriedade importante é que tal número é maior do que qualquer correlação linear obtida em R_{xy} e aumentando o tamanho do vetor X (variáveis independentes) não há uma diminuição do coeficiente de correlação bem como não há garantias de um aumento no seu valor [22] [56].

6. REDES NEURAIS

Redes neurais, também referidas na literatura como neuro-computadores, assemelham-se ao cérebro quanto à habilidade de aprender e generalizar. Estas duas capacidades de processamento de informação tornam possível a rede neural resolver problemas complexos que são correntemente intratáveis.

O cérebro humano contém em torno de 10^{11} neurônios, sua célula fundamental. Cada um destes neurônios processa e se comunica com milhares de outros neurônios continuamente e em paralelo. A estrutura individual dos neurônios, a topologia de suas conexões e o comportamento conjunto destes neurônios naturais constituem a base para o estudo das RNA.

As RNA tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento básico e sua dinâmica. No entanto, do ponto de vista físico, no momento, as redes artificiais se diferem bastante das redes biológicas.

Os neurônios são divididos em três seções: o corpo da célula, os dendritos e o axônio, cada um com funções específicas, porém complementares. O corpo do neurônio mede apenas alguns milésimos de milímetros e os dendritos apresentam poucos milímetros de comprimento. O axônio, contudo, pode ser muito longo e, em geral, tem calibre uniforme. Os dendritos recebem as informações, ou impulsos nervosos, oriundos de outros neurônios, e conduzem-nas até o corpo celular. Ali, a informação é processada e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurônios, passando através do Axônio até os dendritos dos neurônios seguintes. O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado de sinapse. É pelas sinapses que os neurônios se unem funcionalmente, formando redes neurais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos – isto é, o fluxo da informação – entre os neurônios na rede neural [3].

A Figura 4 ilustra, de forma simplificada, os componentes do neurônio.

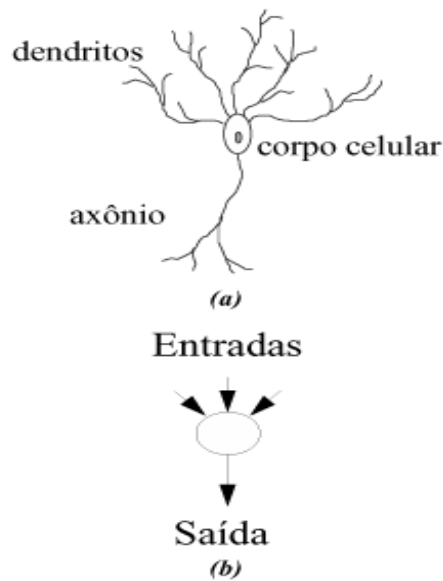


Figura 4. (A) Neurônio Biológico; (B) Neurônio Artificial.

6.1. Modelagem de um neurônio

O modelo de neurônio proposto por McCulloch e Pitts é uma simplificação do que se sabia a respeito do neurônio biológico naquela época [38]. A sua descrição matemática resultou em um modelo com n terminais de entrada que representam os dendritos, e apenas um terminal de saída representada pelo axônio. Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados, cujos valores podem ser positivos ou negativos, dependendo das sinapses correspondentes serem inibitórias ou excitatórias.

O estudo de Redes Neurais Artificiais (RNA, computação neural, conexionismo ou processamento paralelo distribuído) fornece um enfoque alternativo a ser aplicado em problemas onde o enfoque numérico e simbólico não são julgados muito adequados.

Redes neurais artificiais são apenas inspiradas no nosso conhecimento atual sobre os sistemas nervosos biológicos da natureza, e não buscam ser realísticas em todos os detalhes, isto é, a modelagem de sistemas nervosos biológicos não é o ponto principal de interesse.

Um neurônio é uma unidade de processamento da informação que é fundamental para a operação de uma rede neural. Identificam-se três elementos básicos do modelo de um neurônio:

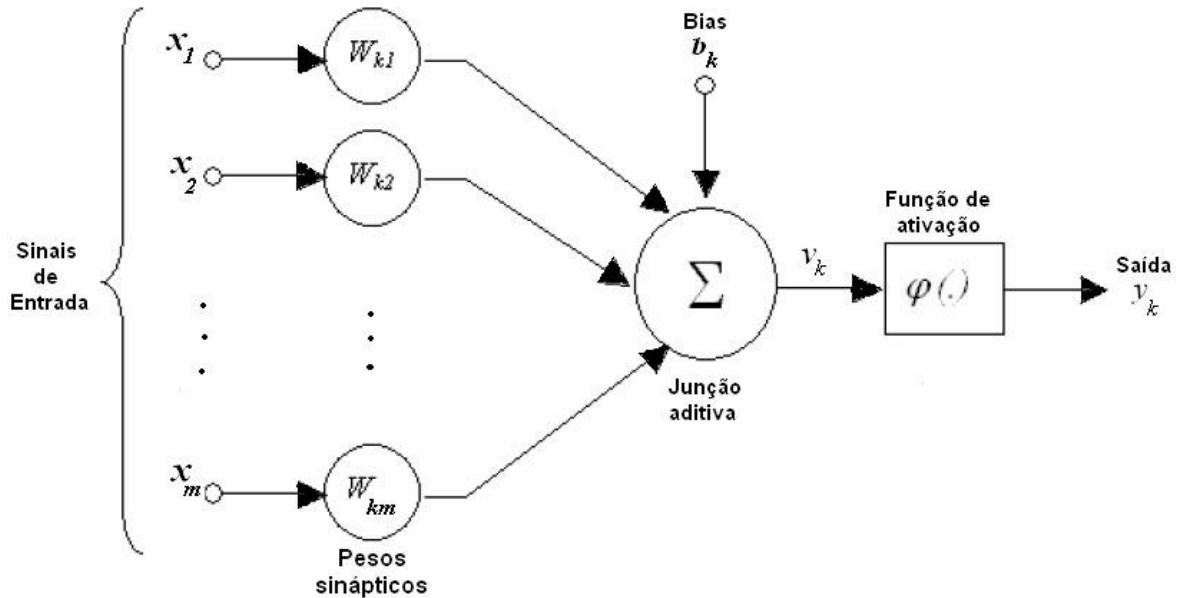


Figura 5. Modelo de neurônio base para projetos de RNA. **Fonte:** Haykin (1999)

1. Um conjunto de sinapses onde cada uma é caracterizada por um peso. Especificamente, um sinal de uma entrada x_j da sinapse j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} .
2. Uma soma para adicionar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio. A operação descrita constitui uma combinação linear.
3. Uma função de ativação (chamamos de φ) para limitar a amplitude da saída do neurônio. A função de ativação tem a função de restringir (ou limitar) a amplitude permitida do sinal de saída para algum valor finito.
4. O modelo de um neurônio inclui o bias, denotado por b_k . O bias tem efeito de aumentar ou diminuir a entrada da rede na função de ativação, dependendo se for positivo ou negativo.

Em termos matemáticos podemos descrever um neurônio como:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad y_k = \varphi(u_k + b_k) \quad (6.1.1)$$

onde x_1, x_2, \dots, x_m são os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos do neurônio k ; u_k é o sinal de saída; b_k é o bias; φ é a função de ativação; y_k é o sinal de saída do neurônio. O uso do bias b_k tem o efeito de transformar u_k em uma função afim, ou seja, $v_k = u_k + b_k$

6.1.1. Função de ativação

A função de ativação, denotada por $\varphi(v)$, define a saída de um neurônio em termos de v . Existem três tipos básicos de função de ativação [20]:

1. Função Threshold.

Função definida por

$$y_k = \varphi(v_k) = \begin{cases} 1 & , \text{ se } v_k \geq 0 \\ 0 & , \text{ se } v_k < 0 \end{cases}$$

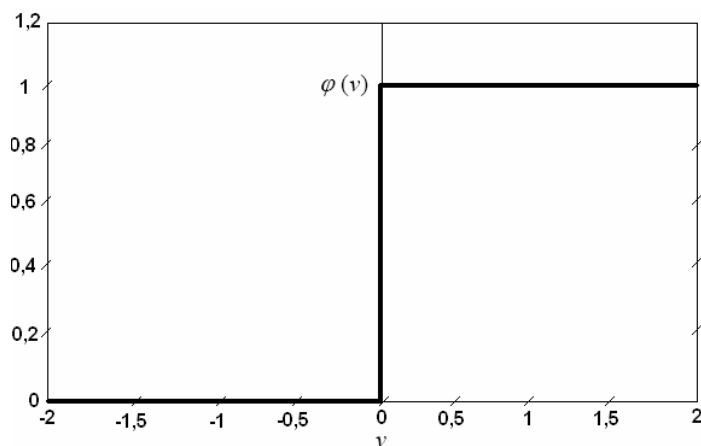


Figura 6. Gráfico função de limiar

2. Função linear por partes

Esta forma de função de ativação pode ser visualizada como uma aproximação para um amplificador não-linear. Tal função tem as seguintes propriedades:

$$\varphi(v_k) = \begin{cases} 1 & , \text{ se } v_k \geq +\frac{1}{2} \\ v_k & , \text{ se } -\frac{1}{2} < v_k < +\frac{1}{2} \\ 0 & , \text{ se } v_k \leq -\frac{1}{2} \end{cases}$$

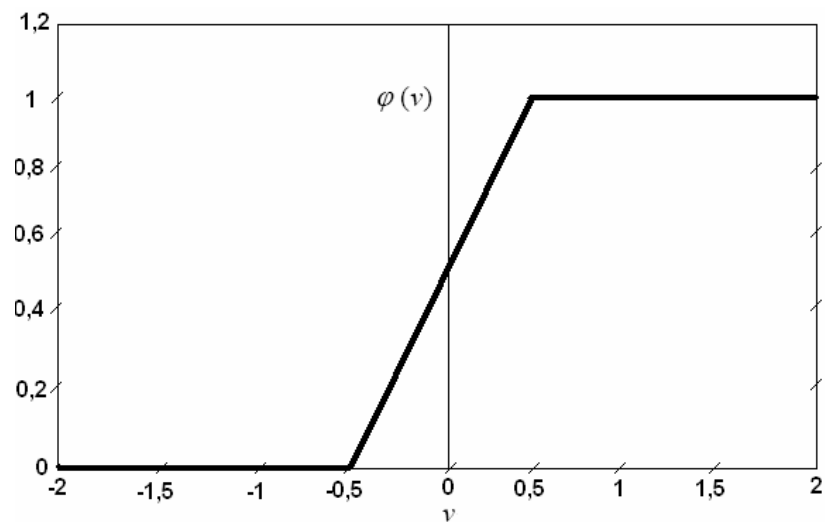


Figura 7. Função linear por partes

3. Função Sigmóide

A função sigmóide é a forma mais comum de função de ativação usada na construção de redes neurais artificiais. É definida como uma função estritamente crescente que exibe um comportamento, alternadamente, linear e não-linear. Um exemplo de função sigmóide é mostrado abaixo:

$$\varphi(v_k) = \frac{1}{1 + \exp(-av_k)} \quad , \text{ onde } a \text{ é um parâmetro da função.}$$

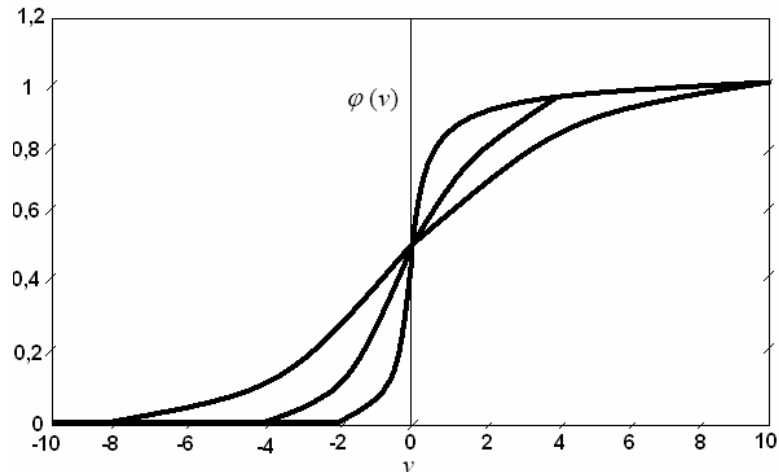


Figura 8. Gráfico função sigmóide

6.1.2. Arquitetura de Redes Neurais Artificiais

As redes neurais artificiais executam suas tarefas usando simultaneamente um grande número de processadores, ou seja, esses processadores operam em paralelo. A representação do conhecimento é distribuída pelas conexões e o aprendizado é feito alterando-se os valores associados com as conexões, não via programação “tradicional”. Todavia, os métodos de aprendizado ainda precisam ser programados e para cada problema específico um método de aprendizado apropriado deve ser escolhido [32].

Uma RNA pode ser projetada através de:

- 1 – Uma definição ou escolha de um conjunto de neurônios artificiais;
- 2 – Definição ou escolha de um padrão de conectividade entre os neurônios, ou seja, uma arquitetura para a rede;
- 3 – Definição de um método de determinação dos parâmetros livres da rede, denominado de algoritmo de aprendizagem ou treinamento.

A definição da arquitetura de uma RNA é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser tratado pela rede. Fazem parte da definição da arquitetura os seguintes parâmetros: número de camadas da rede, número de neurônios em cada camada, tipo de conexão entre os neurônios e topologia da rede [6] [17].

Quanto ao número de camadas, pode-se ter:

- a) redes de camada única – existe apenas um nó entre qualquer entrada e qualquer saída da rede (Figura 9 a,e);
- b) redes de múltiplas camadas – existem mais de um neurônio entre alguma entrada e alguma saída da rede (Figura 9 b, c, d).

Os neurônios podem ter conexões do tipo:

- a) Feedforward, ou acíclica – a saída de um neurônio na i -ésima camada da rede não pode ser usada como entrada de neurônios em camadas de índice menor ou igual a i (Figura 9 a, b, c);
- b) Feedback ou cíclica – a saída de algum neurônio na i -ésima camada da rede usada como entrada de neurônios em camadas de índice menor ou igual a i (Figura 9 d,e).

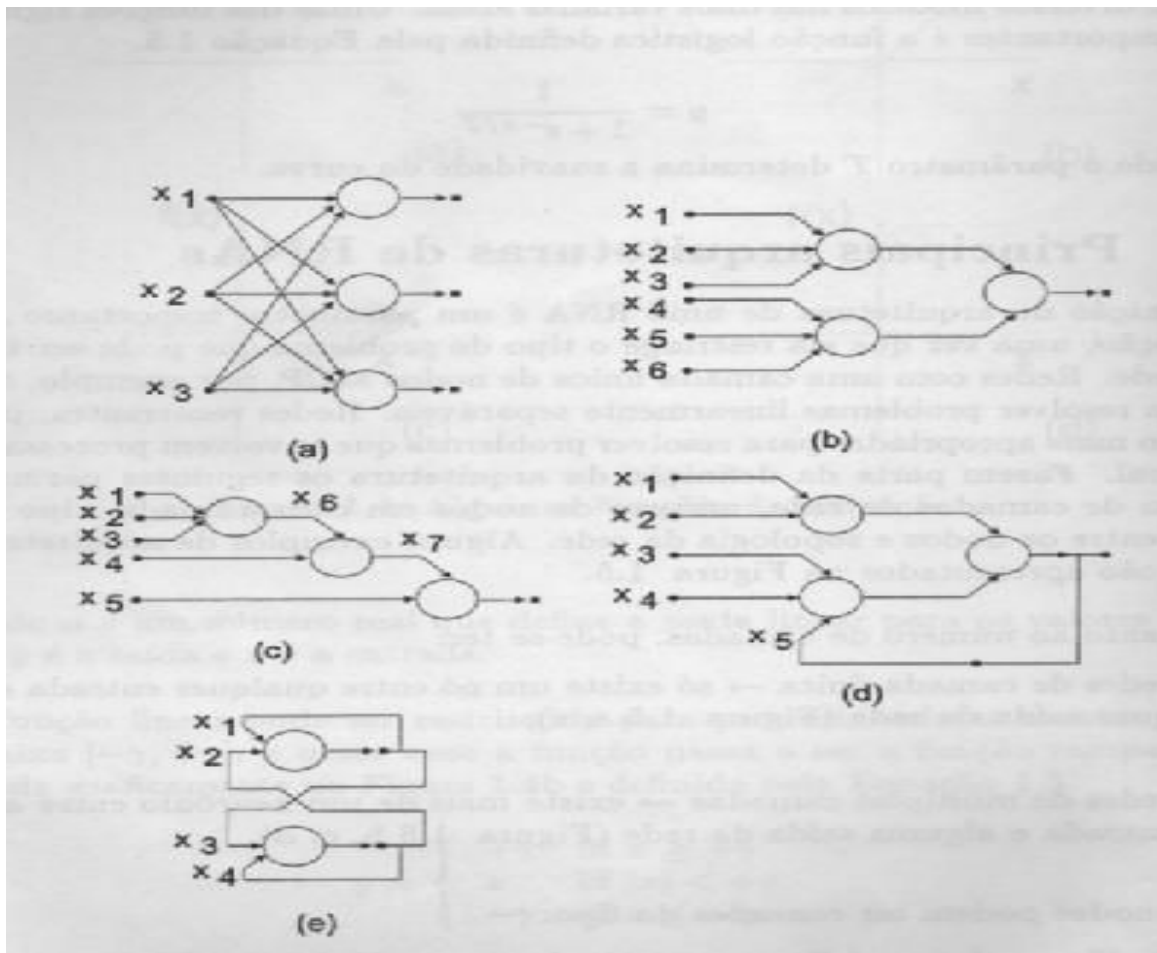


Figura 9. Exemplos de arquiteturas de RNA.

6.2. O processo de aprendizado

O procedimento usado para executar o processo de aprendizagem é chamado algoritmo de aprendizagem cuja função é modificar os pesos sinápticos (força de conexão dos neurônios) da rede de forma adequada a fim de que a rede neural retenha ou armazene o conhecimento. A capacidade de generalização refere-se a produção de saídas para entradas não encontradas durante o treinamento (aprendizado).

Uma rede neural aprende sobre o ambiente através de um processo iterativo de ajustamento aplicado aos pesos sinápticos e bias. Assim, pode-se definir aprendizado no contexto de redes neurais como [32]:

“Aprendizado é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente na qual a rede está inserida”.

6.2.1. Funções de aprendizado

Conforme o tipo de emprego das RNA pode-se ter diferentes funções de aprendizado. Os tipos mais comuns são:

Reconhecimento de padrões

Reconhecimento de padrão é formalmente definido como o processo pelo qual um padrão/sinal recebido é assimilado de um número prescrito de classes (categorias). Uma rede neural desempenha o reconhecimento de padrões primeiramente através de um treinamento, durante o qual à rede é repetidamente apresentado um conjunto de entradas padrões juntamente com a categoria da qual cada padrão particular pertence. Posteriormente, um novo conjunto é apresentado à rede que não tenha sido visto anteriormente, mas que pertence a mesma população de padrões usados para treinar a rede. A rede consegue identificar a classe daquele padrão particular, pois a informação é extraída durante o treinamento.

Aproximação de funções

Considerando uma aplicação não linear entrada-saída descrita pela relação $d = f(x)$ onde o vetor x é a entrada e o vetor d é a saída. O vetor de valores da função não é conhecido, ou seja, $f(.)$ é desconhecido e para termos um conhecimento sobre a mesma é apresentado um conjunto de exemplos: $F = \{(x_i, d_i)\}_{i=1}^n$

A rede neural aproxima a função desconhecida $f(.)$ para $F(.)$ descrevendo a relação entrada-saída. Tal aproximação realizada pela rede é suficientemente fechada para $f(.)$ em um senso Euclidiano sobre todas as entradas, ou seja, $\|F(x) - f(x)\| < \varepsilon$ para todo x onde ε é um número pequeno positivo. Dado o tamanho N do conjunto de treinamento seja grande o suficiente e a rede equipada com o número adequado de parâmetros livres, então o erro de aproximação ε pode ser pequeno o suficiente para a tarefa. [21]

Controle

O controle de processos é outra função do aprendizado que pode ser feito pela rede neural; por um processo entendemos como uma parte crítica de um sistema que deve ser mantido em condições controladas. O sistema envolve o uso de unidades de retroalimentação em torno de uma usina a ser controlada, isto é, a saída da usina é retornada diretamente para a entrada. Assim, a saída y é subtraída do sinal de referência d e o erro e produzido é aplicado a um controlador neural para propósito de ajustar os parâmetros [3] [35] [45].

6.2.2. Tipos de Aprendizado

As RNA têm sido empregadas, na maioria das suas aplicações, como classificadores baseados nas variáveis medidas do processo. Em tais soluções, as RNA podem ser empregadas com aprendizado supervisionado ou não-supervisionado.

Aprendizado Supervisionado: Método mais comum de treinamento das RNA sendo chamado aprendizado supervisionado porque a entrada e saída desejadas para a rede são fornecidas. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma

ligação entre os pares de entrada e saída fornecidos. A rede tem sua saída corrente (calculada) comparada com a saída desejada, recebendo informações do supervisor sobre o erro da resposta atual. A cada padrão de entrada submetido à rede, compara-se a resposta desejada com a resposta calculada, e ajustando-se os pesos das conexões para diminuir o erro.

Os exemplos mais conhecidos de algoritmos para aprendizado supervisionado são a regra delta e a sua generalização para redes de múltiplas camadas, o algoritmo Backpropagation [9].

Aprendizado não-supervisionado: Como o próprio nome sugere, não há um supervisor para acompanhar o processo de aprendizado ou não há exemplos da função a ser aprendida pela rede. O aprendizado é feito através de uma medida independente da qualidade da representação como, por exemplo, a regularidade estatística dos dados de entrada, ou seja, os pesos da rede são modificados em função de critérios internos, tais como, a repetição de padrões de ativação em paralelo de vários neurônios e o comportamento resultante deste tipo de aprendizado é usualmente comparado com técnicas de análise de dados empregadas em estatística.

6.2.2.1. Método do Gradiente Descendente

O método do gradiente descendente é um método de aprendizado supervisionado e será descrito a seguir [20]:

Considere um sistema dinâmico, na qual caracterização matemática é desconhecida e temos disponível apenas um conjunto de entradas e suas respectivas saídas geradas pelo sistema em instantes discretos de tempo a uma taxa uniforme. O modelo de neurônio opera sob a influência de um algoritmo que controla necessariamente o ajuste dos pesos sinápticos. Tal ajuste é determinado pela função erro.

Considerando a função erro $E(w)$ que é continuamente diferenciável em relação ao vetor w de pesos desconhecidos. O objetivo é encontrar uma solução ótima w^* que satisfaz a condição: $E(w^*) \leq E(w)$, isto é, minimizar a função erro em relação aos pesos do vetor w .

Para isso, a condição necessária para otimização é:

$$\nabla E(w^*) = 0 \quad (6.2.2.1.1)$$

onde ∇ é o operador gradiente.

Denotando $\nabla E(w)$ por:

$$\nabla E(w) = \left[\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_m} \right]^t \quad (6.2.2.1.2)$$

e, por conveniência fazer $g = \nabla E(w)$ o objetivo é reduzir a cada iteração do algoritmo a energia do erro, ou seja,

$$E[w(n+1)] < E[w(n)]$$

No método do gradiente descendente, os sucessivos ajustes aplicados ao vetor w de pesos estão na direção oposta ao vetor gradiente $\nabla E(w)$ e, então a correção dos pesos é formalmente descrito por:

$$\Delta w(n) = w(n+1) - w(n) = -\eta g(n) \quad (6.2.2.1.3)$$

onde $g(n)$ é vetor gradiente aplicado ao ponto $w(n)$.

Para mostrar que o algoritmo satisfaz $E[w(n+1)] < E[w(n)]$ a cada iteração usaremos a expansão da série de Taylor em torno de $w(n)$ para aproximar $E[w(n+1)]$ como

$$E[w(n+1)] \approx E[w(n)] + g^t \Delta w(n) \quad (6.2.2.1.4)$$

Substituindo 6.2.2.1.1 em 6.2.2.1.2, temos:

$$E[w(n+1)] \approx E[w(n)] - g^t(n) \eta g(n) = E[w(n)] - \eta \|g(n)\|^2$$

o que mostra, do primeiro e último membro da igualdade acima, que o algoritmo decresce com o decorrer das iterações.

6.2.2.2. Aprendizado por correção do erro

Considere um vetor $x(n)$ produzido por uma ou mais camadas escondidas de neurônios e uma saída $y(n)$. Este sinal de saída é comparado a uma resposta desejada ou alvo, denotado por $d(n)$ e, então, um erro $e(n)$ é produzido. Ou seja,

$$e(n) = d(n) - y(n) \quad (6.2.2.2.1)$$

onde o erro e atua como um mecanismo de controle. O propósito é aplicar uma correção ou ajuste dos pesos sinápticos do neurônio k tal que o sinal de saída y se torne mais

próximo da saída d . Tal objetivo é realizado minimizando a função do erro ou índice de desempenho $E(n)$, definido em termos do sinal de erro $e(n)$ como:

$$E(n) = \frac{1}{2} e_k^2(n) \quad (6.2.2.2)$$

onde $E(n)$ é o valor instantâneo da energia do erro. Onde a correção dos pesos é feita da seguinte forma:

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n) \quad (6.2.2.3)$$

onde η é uma constante positiva que determina a taxa de aprendizado e, finalmente a correção dos pesos sinápticos:

$$\Delta w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (6.2.2.4)$$

O ajuste dos pesos sinápticos do neurônio k continua, até que o sistema atinja a estabilidade e, então, o processo de aprendizado é interrompido.

6.3. Algoritmo Backpropagation

As redes multicamadas feedforward são uma importante classe de redes neurais. Consiste de um conjunto de unidades de sensoriais (neurônios fontes) que constituem a camada de entrada, uma ou mais camadas escondidas de neurônios computacionais e uma camada de saída de neurônios computacionais. O sinal de entrada propaga através da rede em uma direção camada por camada. Finalmente um conjunto de saídas é produzido como resposta atual da rede. Neste passo, os pesos sinápticos da rede são todos fixos. Durante o passo de retorno, por outro lado, os pesos sinápticos são ajustados de acordo com a regra de correção de erro. Especificamente, a resposta atual é subtraída da resposta desejada (alvo) para produzir o sinal erro. Este sinal erro é então propagado para trás através da rede até a direção dos pesos sinápticos. Os pesos sinápticos são ajustados de forma a atualizar a resposta da rede aproximando-a da resposta desejada [20] [57].

Estas redes neurais são comumente referidas como perceptrons multicamadas (MLP), que representam uma generalização do perceptron de camada simples (modelo de

neurônio cuja arquitetura consiste na entrada e uma camada de saída, com função limiar e aprendizado supervisionado).

Em termos matemáticos temos:

O sinal de erro do neurônio de saída j na iteração n é definido por:

$$e_j(n) = d_j(n) - y_j(n) \quad (6.3.1)$$

Definimos o valor instantâneo da energia do erro para o neurônio j como $\frac{1}{2}e_j^2(n)$.

Correspondentemente, o valor instantâneo $E(n)$ da energia total do erro é obtido somando $\frac{1}{2}e_j^2(n)$ sobre todos os neurônios na camada de saída. Assim, podemos escrever:

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (6.3.2)$$

onde o conjunto C inclui todos os neurônios na camada de saída da rede. Tome N denotando o número total de padrões (exemplos) contidos no conjunto de treinamento. A energia média do erro quadrático é obtida somando $E(n)$ sobre todo n e então normalizando com relação ao tamanho de N , conforme:

$$E_{m\u00e9dia} = \frac{1}{N} \sum_{i=1}^N E(n) \quad (6.3.3)$$

A energia do erro instantâneo $E(n)$ e, também, a energia média do erro $E_{m\u00e9dia}$, é uma função de todos os parâmetros livres (isto é, pesos sinápticos e nível de bias) da rede. Para um dado conjunto de treinamento, $E_{m\u00e9dia}$ representa a função custo como medida de desempenho de aprendizado. O objetivo do processo de aprendizado é ajustar os parâmetros livres da rede para minimizar $E_{m\u00e9dia}$. Especificamente, consideramos um método simples de treinamento que os pesos são carregados padrão por padrão até uma época, isto é, uma apresentação completa da entrada do conjunto de treinamento. O ajustamento dos pesos é feito em acordo com o respectivo erro computado para cada padrão apresentado à rede.

Assim o sinal da função $y_j(n)$ na saída do neurônio j após a iteração n é:

$$y_j(n) = \varphi_j(v_j(n)) \quad (6.3.4)$$

onde $v_j = \sum_{i=0}^m w_{ji}(n) y_i(n)$ tal que m é o número total de entradas (excluindo o bias) aplicado ao neurônio.

Agora, o algoritmo Backpropagation aplica a correção $\Delta w_{ij}(n)$ ao peso sináptico $w_{ij}(n)$ onde, de acordo com a regra da cadeia do cálculo temos:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (6.3.5)$$

Diferenciando ambos os lados de 6.3.2 com relação à $e_j(n)$, temos:

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n) \quad (6.3.6)$$

Diferenciando ambos os lados de 6.3.1 em relação à $y_j(n)$ temos:

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (6.3.7)$$

Diferenciando 6.3.4 em relação a $v_j(n)$ temos:

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)) \quad (6.3.8)$$

E, finalmente, diferenciando “ $v_j = \sum_{i=0}^m w_{ji}(n) y_i(n)$ ” com relação à $w_{ji}(n)$ temos:

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \quad (6.3.9)$$

Substituindo (6.3.6) à (6.3.9) em (6.3.5) temos:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n) \phi_j'(v_j(n)) y_i(n) \quad (6.3.10)$$

A correção $\Delta w_{ji}(n)$ aplicada a $w_{ji}(n)$ é definido pela regra delta:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (6.3.11)$$

onde η é o parâmetro da taxa de aprendizagem do algoritmo Backpropagation e o sinal negativo indica que o gradiente é descendente em relação ao peso.

E, finalmente, substituindo 6.3.5 em 6.3.11 temos:

$$\Delta w_{ji}(n) = \eta e_j(n) \phi_j'(v_j(n)) y_i(n) \quad (6.3.12)$$

6.4. Validação de modelos em redes neurais

As Redes Neurais constituem uma classe de modelos bastante flexíveis que podem ser usados em uma variedade de aplicações diferentes e um dos problemas centrais no estudo desta metodologia é a seleção de modelos, pois a questão de qual arquitetura de rede escolher para resolver determinado problema é uma questão não resolvida. Neste contexto, existem métodos que permitem a avaliação do modelo de rede proposto, tais como validação cruzada, regularização ou métodos estatísticos [1] [12] [17].

Devido a essa variedade de tipos e modelos de redes é necessário uma metodologia para selecionar um tipo apropriado para um dado problema e, em particular, neste trabalho, no que se refere ao algoritmo Backpropagation, é necessário um método que avalie o desempenho da rede nesta tarefa.

É durante a fase de aprendizado (treinamento) que a rede, através da apresentação de exemplos, mapeia o conjunto que lhe é apresentado criando uma função de aproximação para esses dados e na fase de teste é verificado o quanto a função criada consegue aproximá-los. Em geral, temos [40]:

Seja X um conjunto de entradas representado por I características da rede tal que um ponto $x \in X$ é definido por $x = (x_1, x_2, \dots, x_I)$ e um conjunto de alvos T_x definido sobre X tal que, ao fim do processo de treinamento e teste da rede, devemos ter:

$$T_x = f(x) + \varepsilon \quad (6.4.1)$$

onde f é a função aproximada pela rede e ε é o erro cometido nessa aproximação.

Observemos que se o erro na aproximação da função é nulo teremos $T_x = f(x)$, ou seja, os valores do conjunto alvo serão idênticos aos da função f .

Neste trabalho, para a validação do modelo de rede, utilizaremos a equação (6.3.3) que é o valor da Energia do Erro:

$$E_{média} = \frac{1}{N} \sum_{i=1}^N E(n)$$

E o Teste de Correlação dado pela equação (4.2.1):

$$r^2 = \sqrt{1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}}$$

O Teste de Correlação envolve a saída da rede f e o valor alvo T que, quanto mais próximo do valor 1 melhor será o modelo proposto quanto a homogeneidade da aproximação [17] [40].

Onde y e \hat{y} correspondem aos valores do alvo e saída da rede respectivamente e \bar{y} é o valor médio de y .

Os valores para r estão entre 0.0 e 1.0 e não tem unidades. O valor 0.0 significa que um conjunto de variáveis X não pode ajudar a prever Y ou não existe relação linear entre X e Y . Quando r é igual a 1.0, todos os pontos estão exatamente sobre uma linha reta sem nenhuma dispersão (ver figura 10).

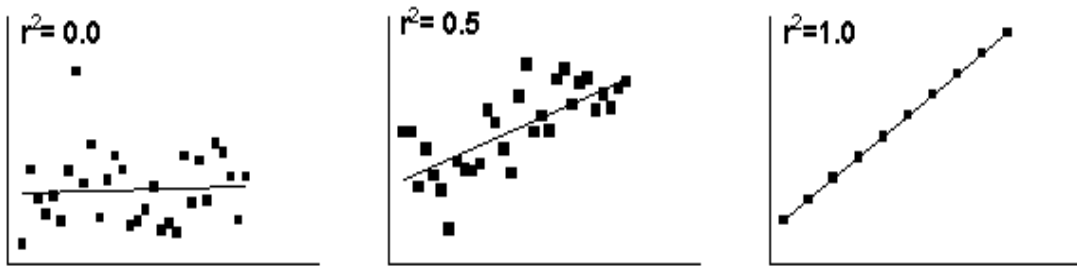


Figura 10. Dispersão dos pontos em relação ao grau de correlação

Com o Teste de Correlação (Qualidade do Ajuste) e a Energia do Erro temos dois números que medem o quão próximo a função calculada pela rede está da função alvo (ou dos valores que se quer ajustar pela função) e o quanto a função segue o padrão de comportamento dos pontos [40].

7. MODELAGEM MATEMÁTICA DE SISTEMAS E REDES NEURAIS COMO APROXIMAÇÃO DE FUNÇÕES

O modelo matemático de um sistema é definido como um conjunto de equações que representa com precisão ou, pelo menos, razoavelmente bem o comportamento do sistema.

Na obtenção de um modelo matemático devemos estabelecer uma conciliação entre a simplicidade do modelo e a precisão dos resultados da análise sendo necessário, com frequência, ignorar certas propriedades físicas inerentes ao sistema como, por exemplo, a característica altamente não linear de alguns processos a fim de melhorar o desempenho de controladores lineares. Se os efeitos que essas propriedades ignoradas têm na resposta forem pequenos, pode-se obter boa aproximação entre os resultados da análise de um modelo matemático e os resultados do estudo experimental do sistema físico. [5] [29] [47]

Em geral, na solução de um novo problema, é conveniente construir um modelo simplificado para que possamos ter uma percepção geral em relação à solução. Um modelo matemático mais completo pode, então, ser construído e utilizado para que sejam obtidas soluções mais precisas.

7.1. Linearização de modelos

Um sistema é dito linear se o princípio da superposição se aplicar a ele. O princípio da superposição afirma que a resposta produzida pela aplicação simultânea de duas funções diversas é a soma das duas respostas individuais. Então, para um sistema linear, a resposta a diversas entradas pode ser calculada tratando uma entrada de cada vez e somando os resultados. Esse é o princípio que permite construir soluções complicadas para equações diferenciais lineares a partir de soluções simples. [10] [45]

7.2. Redes Neurais e Aproximação de funções

Redes Neurais Artificiais do tipo feedforward são usadas em uma variedade de aplicações com grande sucesso. RNA não requer a programação algorítmica que resolve problemas específicos, pois aprende de exemplos e outra vantagem é sua habilidade de

generalização, isto é, podem identificar e responder a padrões que são similares, porém não idênticos aos quais ela foi treinada [4].

Redes neurais são aproximadores universais de funções e um perceptron multicamadas treinado com o algoritmo Backpropagation pode ser visto como um modo de se apresentar uma aplicação não linear de entrada-saída [20] [32].

Porém, existem restrições para a função que se quer aproximar [13] [21] [33] [59].

O seguinte teorema mostra as restrições que garantem a aproximação de funções contínuas por redes com três camadas (uma camada escondida) cuja função de saída é uma função sigmóide.

Segundo o teorema demonstrado em [13]:

Seja $\phi(x)$ função contínua não constante, crescente, monótona e limitada. Seja K um subconjunto compacto (fechado e limitado) e $f(x_1, \dots, x_n)$ uma função contínua de valores reais sobre K . Então, para qualquer valor arbitrário ε , existem inteiros N e constantes reais $c_i, \theta_i (i = 1, \dots, N), w_{ij} (i = 1, \dots, N, j = 1, \dots, n)$ tal que satisfazendo:

$$\tilde{f}(x_1, \dots, x_n) = \sum_{i=1}^N c_i \phi\left(\sum_{j=1}^n w_{ij} x_j - \theta_i\right) \quad (7.2.1)$$

ou seja, para qualquer $\varepsilon > 0$ existe uma rede com três camadas cujas funções de saída para a camada oculta são $\phi(x)$, cujas funções para entrada e saída das camadas são lineares e

que tem uma função de entrada-saída $\tilde{f}(x_1, \dots, x_n)$ tal que

$$\max_{x \in K} |f(x_1, \dots, x_n) - \tilde{f}(x_1, \dots, x_n)| < \varepsilon \quad (7.2.2)$$

Observação: O teorema é diretamente aplicável ao perceptron multicamadas.

Primeiramente notamos que a função $\phi(x) = \frac{1}{1 + e^{-x}}$ usada como não linearidade em um modelo de neurônio para a construção de um perceptron multicamadas é não constante, contínua e monótona crescente e, portanto, satisfaz as condições impostas para função ϕ .

O teorema da aproximação de funções é importante de um ponto de vista teórico, pois fornece a ferramenta matemática necessária para a viabilidade das redes feedforward com uma camada escondida como uma classe de soluções aproximadas. Sem tal teorema poderíamos procurar por soluções inexistentes [20] [32].

8. METODOLOGIA DE TRABALHO

O objetivo deste trabalho é, dentro do universo de todas as variáveis monitoradas de um processo, construir um conjunto de variáveis, não necessariamente mínimo, que será o conjunto de entrada numa rede neural e, com isso, conseguir monitorar, via própria rede, o maior número possível de variáveis. Esta metodologia foi aplicada ao reator de pesquisas IEA-R1 do IPEN.

Para isso, foi tomado um conjunto de variáveis formado por Potência do reator (N2), Vazão do circuito primário (F1M3), Posição das barras de controle/segurança (Z1, Z2, Z3 e Z4) e Delta P, pois pela análise do comportamento físico das variáveis do reator quase todas as variáveis monitoradas têm relação com alguma dessas ou pode ser resultado da interação de duas ou mais. Por exemplo, a potência está relacionada à quantidade de radiação devido à fissão do Urânio e, também, está relacionada a algumas temperaturas; as barras de controle/segurança são reguladoras de potência e, por consequência podem influenciar na quantidade de radiação e/ou temperaturas; a vazão do circuito primário, responsável pelo transporte de energia pela consequente retirada de calor do núcleo somando-se a isso ao provável transporte de radiação, etc.

Assim, tomando o grupo de variáveis $B = \{\text{Potencia, Vazão do primário, Posição das barras de controle/segurança, Delta P}\}$ calcula-se a correlação existente entre o conjunto B e todas as outras variáveis monitoradas (coeficiente de correlação múltipla), isto é, através do cálculo da correlação múltipla, que é uma ferramenta proposta pela teoria das Correlações Canônicas, calculamos o quanto o conjunto B pode predizer cada uma das variáveis monitoradas e, além disso, será medida a força da relação linear entre B e a variável que se pretende monitorar.

8.1. Arquitetura das redes MLP utilizadas neste trabalho

Neste trabalho foram desenvolvidas redes neurais do tipo Multi-Layer Perceptron ou perceptron de multi-camadas (MLP) através da ferramenta de Redes Neurais existente no software Matlab.

As redes constam de três camadas (uma de entrada, uma oculta e uma de saída) e foram treinadas com o algoritmo de retropropagação (Backpropagation). Esta arquitetura com três camadas foi escolhida por ter apresentado bons resultados na monitoração de variáveis no trabalho mostrado em Bueno [8] e a escolha do algoritmo de treinamento

Backpropagation deve-se, conforme já comentado, ao fato de que uma rede do tipo MLP treinada com este algoritmo ser um aproximador universal de funções.

O número de neurônios na camada oculta será igual ao número de variáveis de entrada e a função de ativação será a função sigmóide.

Uma vez que para cada variável monitorada pela rede o algoritmo converge para valores distintos, o parâmetro para interrupção do treinamento (Energia do erro) não foi fixado, porém fixamos o número de iterações (ciclos ou épocas) do algoritmo num valor igual a 1000. Para validação da rede usamos o resultado da energia do erro e a qualidade do ajuste.

8.2. Desenvolvimento do Trabalho

Os dados de operação do reator utilizados correspondem ao período de 19 a 22 de janeiro de 2009 e, inicialmente, para efeito deste trabalho, serão consideradas todas variáveis monitoradas e receberão uma numeração de 1 a 64 conforme abaixo [62]:

Tabela 1 : Variáveis Monitoradas pelo SAD. Potência do reator: 3,5MW. (*)Equipamentos não pertencentes à configuração de operação no período de 19 a 22 de janeiro de 2009

1	T1:	Temperatura da água na superfície da piscina [°C].
2	T2:	Temperatura da água na superfície da piscina [°C].
3	T3:	Temperatura da água sobre o núcleo do reator [°C].
4	T4:	Temperatura da água na entrada do tanque de decaimento [°C].
5	T4-T3:	Diferença de temperaturas
6	T6:	Temperatura da água na saída do tanque de decaimento [°C].
7	T7:	Temperatura da água na saída do circuito primário (trocador A) [°C].
8	T8:	Temperatura da água na entrada do circuito secundário [°C].
9	T9:	Temperatura da água na saída do circuito secundário [°C].
10	T9-T8:	Diferença de temperaturas
11	T13:	Temperatura na carcaça do motor da bomba B101-A [°C].
12	T15:	Temperatura na bacia da torre de refrigeração A [°C].
13	T16:	Temperatura na bacia da torre de refrigeração B [°C].
14	T16-T15:	Diferença de temperaturas
15	T19:	Temperatura no No-Break-440V [°C].
16	T20:	Temperatura ambiente (3° andar) [°C].
17	T20-T7:	Diferença de temperaturas
18	T22:	Temperatura ambiente (sala de controle) [°C].
19	F1M3:	Vazão do circuito primário [gpm].
20	F23:	Vazão da água de emergência [m ³ /h]
21	F2M3:	Vazão do circuito secundário (trocador A) [gpm]

22	F3M3*:	Vazão do circuito secundário (trocaador B) [gpm]
23	C1:	Condutividade da água da piscina após tratamento [μ mho].
24	C2:	Condutividade da água da piscina antes da entrada no sistema de retratamento [μ mho].
25	L1:	Nível da piscina [%]
26	F1:	Constante de entrada para cálculo da vazão do circuito primário.
27	F2:	Constante de entrada para cálculo da vazão do circuito primário
28	F3:	Constante de entrada para cálculo da vazão do circuito primário
29	F4 *:	Constante de entrada para cálculo da vazão do circuito primário
30	delta P:	Sinal elétrico relativo à diferença de pressão do núcleo do reator [V]
31	Preator:	Cálculo da potência do reator [MWatts]
32	PPtcA :	Cálculo da potência do reator, circuito primário, trocaador de calor A [MWatts]
33	PPtcB:	Cálculo da potência do reator, circuito primário, trocaador de calor B [MWatts]
34	PStcA:	Cálculo da potência do reator, circuito secundário, trocaador de calor A [MWatts]
35	PStcB*:	Cálculo da potência do reator, circuito secundário, trocaador de calor B [MWatts]
36	NBPtcA:	Cálculo da potência do reator, circuito secundário, trocaador de calor A [MWatts]
37	NBPtcB:	Cálculo da potência do reator, novo balanço, trocaador de calor B [MWatts]
38	N1:	Período do reator, câmara de fissão [s]
39	N2 :	% de potência (canal de segurança 1) [%].
40	N3:	% de potência (canal de segurança 2) [%].
41	N4:	% de potência (canal de segurança 2) [%].
42	N5:	Potência logarítmica (canal log) [%].
43	N6:	% de potência (canal linear) [%].
44	N7:	% de manda (modo automático) [%].
45	N8:	Potência do N16 (canal N16) [%]
46	Z1:	Posição da barra de controle [0 a 1000 mm]
47	Z2:	Posição da barra de segurança 1 [0 a 999 mm]
48	Z3:	Posição da barra de segurança 2 [0 a 999 mm]
49	Z4 :	Posição da barra de segurança 3 [0 a 999 mm]
50	R1M3:	Dose junto à ponte de sustentação do núcleo, lado esquerdo [mR/h].
51	R2M3:	Dose próximo da ponte de sustentação do núcleo, lado direito [mR/h].
52	R3M3:	Dose próximo da parede do prédio do saguão da piscina, face norte [mR/h].
53	R4M3:	Dose próximo da parede da piscina, face sul, sobre o tubo de irradiação nº8 [mR/h]
54	R5M3:	Dose próximo da parede da piscina, face oeste, sobre os tubos de irradiação nº3 e nº4 [mR/h].
55	R6M3:	Dose próximo da parede do reator, face leste, sobre os tubos de armazenamento de material radioativo [mR/h].
56	R7M3:	Dose próximo do poço de coleta de água da piscina (poço SUMP) [mR/h].

57	R8M3:	Dose próximo do sistema de retratamento de água – coluna de resinas [mR/h].
58	R9M3:	Dose próximo do trocador de calor A [mR/h].
59	R10M3:	Contagem próximo da parede do fundo do porão, encostada ao duto de exaustão entre os trocadores de calor A e B [cps].
60	R11M3:	Contagem próximo à sala do ar condicionado, junto ao duto de exaustão de ar do 1º andar [cps].
61	R12M3:	Contagem próximo ao saguão da piscina, junto ao duto de exaustão de ar [cps].
62	R13M3:	Contagem próximo ao duto de exaustão de ar da chaminé [cps].
63	R14M3:	Contagem próximo ao tanque de retenção [cps].
64	NOBRES:	Gases nobres.

Analisando o gráfico de cada uma destas variáveis temos que o padrão das curvas a eles associados tem as características ou propriedades de funções sendo então razoável a aproximação destas por redes neurais. Em anexo têm-se os gráficos destas variáveis com seus valores (normalizadas entre 0 e 1) onde podemos observar seus comportamentos variando no tempo.

Tomamos o conjunto $B = \{\text{Potencia (39), Vazão do primário (19), Barras de controle/segurança (46, 47, 48 e 49) e Delta P (30)}\}$ como base.

Primeiramente vamos verificar se cada uma das variáveis de B pode ser aproximada pela rede neural tendo como entrada as outras variáveis deste mesmo conjunto bem como o grau de correlação de cada conjunto com a respectiva variável monitorada.

Tabela 2: Energia do erro e qualidade de ajuste

Variável	Energia do Erro (Eq. 6.3.3)	Qualidade do Ajuste (Eq. 4.2.1)
F1M3 (19)	129,7635	0,0479
Delta P (30)	5,7459e-004	0,7356
N2 (39)	0,0608	0,9965
Posição de Barra (46)	88,0056	0,9755
Posição de Barra (47)	0,1351	0,9999
Posição de Barra (48)	0,0598	1,0000
Posição de Barra (49)	0,2819	0,9999

Pela teoria das Correlações Canônicas podemos aumentar o grau de correlação linear entre o conjunto de entrada e a variável monitorada adicionando outras que possuam alto grau de correlação com a variável monitorada e, além disso, seu grau de correlação deve ser maior do que as correlações bivariadas entre as variáveis já existentes e variável de entrada, pois, caso contrário, não há garantias do aumento do coeficiente de correlação linear.

Comparando a energia do erro e a qualidade do ajuste observa-se que a rede não obteve êxito na aproximação da variável Vazão do Circuito Primário (19) e Posição de Barra de Controle (46). Desta forma, acrescentaremos variáveis ao conjunto primeiramente para obter um melhor ajuste linear para posteriormente verificar a capacidade de aproximação pela rede neural tendo como entrada as novas variáveis do conjunto.

A seguir tem-se a correlação bivariada entre a posição de barra (46) e as variáveis monitoradas pelo SAD e, analogamente, entre a vazão do primário (19) e as variáveis do SAD.

Tabela 3: Correlação bivariada entre 19, 46 e as variáveis do SAD

Variável	Correlação com 19	Variável	Correlação com 19	Variável	Correlação com 46	Variável	Correlação com 46
1	0,4132	33	-0,2396	1	0,5112	33	-0,0632
2	0,3773	34	-0,8677	2	0,5150	34	0,1487
3	0,3635	35	-0,0776	3	0,5138	35	-0,0006
4	0,3808	36	-0,8841	4	0,5416	36	0,0927
5	-0,1195	37	-0,8410	5	0,2998	37	0,2660
6	0,3551	38	-0,9574	6	0,5284	38	-0,1563
7	0,3388	39	0,1011	7	0,4575	39	0,8143
8	0,2723	40	0,1237	8	0,1489	40	0,7961
9	0,3102	41	0,0792	9	0,3501	41	0,8011
10	-0,2376	42	-0,9439	10	0,6406	42	-0,0620
11	0,3160	43	0,0815	11	0,7538	43	0,8012
12	0,2774	44	0,0806	12	0,3470	44	0,7187
13	0,2073	45	-0,0315	13	0,1734	45	0,2134
14	-0,3850	46	0,1514	14	0,5037	46	1,000
15	0,0377	47	0,1463	15	-0,4324	47	0,8937
16	0,2485	48	0,1457	16	0,5366	48	0,8947
17	-0,6426	49	0,1456	17	0,3653	49	0,8946
18	0,1384	50	-0,4285	18	0,5349	50	0,6765
19	1,0000	51	-0,4559	19	0,1514	51	0,6664
20	0,2509	52	-0,6983	20	0,1330	52	0,4456
21	0,4223	53	-0,9497	21	0,2430	53	-0,0274
22	0,0022	54	-0,9575	22	-0,0117	54	-0,1707
23	-0,9613	55	-0,9608	23	-0,1684	55	-0,1664
24	-0,9530	56	-0,5792	24	-0,2620	56	0,0975
25	-0,9602	57	-0,0354	25	-0,1945	57	0,6388
26	0,9920	58	-0,6247	26	0,1401	58	0,4935
27	0,4915	59	0,2147	27	0,1886	59	0,3761
28	0,5762	60	0,0394	28	0,1255	60	0,8519
29	-0,4546	61	0,0917	29	0,2796	61	0,2021
30	-0,9495	62	0,0554	30	-0,2382	62	0,0023
31	-0,8920	63	0,0607	31	-0,0452	63	0,6063
32	-0,8508	64	-0,9617	32	0,1358	64	-0,1644

Tabela 4: Correlação bivariada entre 19 e as variáveis de B

19	30	39	46	47	48	49
Correlação	-0,9495	0,1011	0,1514	0,1463	0,1457	0,1456

Tabela 5: Correlação bivariada entre 46 e as variáveis de B

46	19	30	39	47	48	49
Correlação	0,1514	-0,2382	0,8143	0,8937	0,8947	0,8946

Da Tabela 4 vemos que a maior correlação bivariada é (em módulo) entre 19 e 30 e, da Tabela 5 a maior correlação bivariada é entre 46 e 48. Portanto, para aumentar o grau de relação linear entre 19 e seu conjunto podemos acrescentar a variável 26 (correlação entre 26 e 19 é 0,9920) e não acrescentaremos variáveis ao conjunto para a predição de 46 uma vez que todas as correlações bivariadas encontradas são menores do que a de maior valor da tabela 5. Assim, obtemos um progresso quanto à aproximação pela rede neural somente da variável 19 observada abaixo pela baixa energia do erro (0,0018) e pela alta qualidade do ajuste (1,0000).

Testamos o desempenho da rede neural primeiramente com um conjunto no qual consta a variável 46 e depois testaremos o desempenho de outra rede da qual a excluimos do conjunto de entrada da rede neural e desta forma avaliaremos a importância desta variável para o monitoramento das variáveis do SAD. Outro fato importante do conjunto B é que excluiremos a variável 49 devido à sua redundância com 48, pois apresenta grau de correlação igual a 1,0000 com a mesma além do fato de ambas representarem leituras de posição de barras de segurança do reator. Chamamos o novo conjunto formado por 19 26 30 39 46 47 e 48 de B_1 e o conjunto formado pelo mesmo conjunto excluído a variável 46 de B_2 .

Verificamos, posteriormente, o quanto uma rede neural com o algoritmo Backpropagation pode fazer a aproximação entre B_1 e B_2 e as variáveis monitoradas do reator IEA-R1, porém, antes foi calculado o coeficiente de correlação múltipla. Com isso iremos estabelecer uma comparação entre o grau de aproximação linear de variáveis e a respectiva aproximação via redes neurais.

A Tabela 6 mostra o coeficiente de correlação entre B_1 e B_2 e cada uma das variáveis preditas.

Tabela 6: Coeficiente de correlação entre B_1 e B_2 e as variáveis do SAD.

Variável	B_1	B_2
1	0,7116	0,5467
2	0,6830	0,5256
3	0,6713	0,5181
4	0,6936	0,5583
5	0,7656	0,7656
6	0,6656	0,5303
7	0,6362	0,4708
8	0,5560	0,3771
9	0,5874	0,4094
10	0,7609	0,7471
11	0,7954	0,7495
12	0,5715	0,3840
13	0,5691	0,3439
14	0,7432	0,7399
15	0,6867	0,6391
16	0,6340	0,4693
17	0,8326	0,8267
18	0,6370	0,4142
19	1,0000	1,000
20	0,3941	0,3909
21	0,6575	0,6488
22	0,0305	0,0264
23	0,9997	0,9997
24	0,9991	0,9990
25	0,9993	0,9992
26	1,0000	1,0000
27	0,6088	0,6054
28	0,5989	0,5989
29	0,6149	0,6131
30	1,0000	1,0000
31	0,9743	0,9743
32	0,9408	0,9398

Variável	B_1	B_2
33	0,6284	0,4342
34	0,9545	0,9525
35	0,0858	0,0844
36	0,9579	0,9564
37	0,9720	0,9646
38	0,9971	0,9971
39	1,0000	1,0000
40	0,9996	0,9994
41	0,9999	0,9974
42	0,9975	0,9974
43	0,9998	0,9990
44	0,8788	0,8732
45	0,9731	0,9667
46	1,0000	0,9014
47	1,0000	1,0000
48	1,0000	1,0000
49	1,0000	1,0000
50	0,9772	0,9772
51	0,9797	0,9797
52	0,9815	0,9814
53	0,9966	0,9964
54	0,9962	0,9961
55	0,9993	0,9993
56	0,7987	0,7850
57	0,9121	0,8823
58	0,9623	0,9608
59	0,4237	0,4177
60	0,9683	0,9677
61	0,2468	0,1647
62	0,0595	0,0591
63	0,6231	0,5874
64	1,0000	1,0000

Na Tabela 7 tem-se um resumo do desempenho da rede na aproximação das variáveis usando B_1 e B_2 como conjunto de entrada.

Tabela 7: Aproximação de variáveis por B₁ e B₂

Aproximação por B ₁		
Variável	Energia do Erro	Qualidade do Ajuste
1	0,0028	0,9969
2	0,0029	0,9938
3	0,0032	0,9875
4	0,0350	0,9636
5	0,0031	0,47011
6	0,0030	0,9965
7	0,0168	0,9754
8	0,0147	0,9660
9	0,0066	0,8017
10	0,0061	0,8865
11	0,0034	0,9962
12	0,0183	0,9682
13	0,0072	0,9512
14	0,0055	0,8822
15	0,6989	0,8958
16	0,0037	0,6265
17	0,0067	0,8251
18	0,0044	0,9949
20	8,32e+03	0,0227
21	9,244,984	0,0124
22	,2902e+03	0,0095
23	1,29e-05	0,54789
24	3,68e-01	0,9725
25	1,17e-01	0,8258
27	2,14e+03	0,0450
28	331,096	0,0304
29	0,2824	0,4739
31	0,0024	0,8663
32	0,0045	0,8460
33	0,0443	0,9340
34	0,0019	0,9262
35	15,845	0,3723
36	0,0027	0,8992
37	0,0031	0,9629
38	1,7982	0,3796

Aproximação por B ₂		
Variável	Energia do erro	Qualidade do Ajuste
1	0,1116	0,8781
2	0,1132	0,8732
3	0,2128	0,7434
4	0,0826	0,905
5	0,0031	0,457
6	0,0842	0,9046
7	0,0918	0,8819
8	0,0674	0,9083
9	0,6058	0,8977
10	0,0058	0,8958
11	0,0381	0,9412
12	0,0801	0,8859
13	0,0748	0,8967
14	0,083	0,8359
15	0,9423	0,8757
16	0,0861	0,9029
17	0,0061	0,7814
18	0,0963	0,9049
20	8,31e+03	0,0408
21	9,247,954	0,0458
22	7,31e+03	-0,0027
23	3,02e-05	0,2361
24	5,32e-05	0,975
25	2,16e-05	0,896
27	2,17e+03	0,0298
28	3,34e+01	0,0234
29	0,2913	0,4465
31	0,0044	0,8228
32	0,1192	0,9271
33	0,0043	0,8372
34	0,0044	0,8344
35	0,9843	0,3015
36	0,0048	0,9502
37	1,78e-04	0,4026
38	1,7879	0,4036

40	0,0017	0,9999
41	0,0013	0,9999
42	2,26e-01	0,9766
43	0,0014	0,9997
44	0,1915	0,9902
45	0,0398	0,8574
49	0,0012	10,000
50	0,0120	0,9930
51	0,0115	0,9928
52	0,0076	0,9823
53	0,0022	0,8891
54	0,000081	0,4832
55	5,42e-04	0,0048
56	0,0568	0,5187
57	0,1613	0,9987
58	0,0311	0,9520
59	4,13e+02	0,3516
60	1,249	0,9863
61	2,88e+02	0,3458
62	1,80e+03	1894
63	5,30e+02	0,6587
64	3,16e-06	0,9855

40	0,0232	0,9988
41	2,40e-05	0,9779
42	0,0126	0,9993
43	0,0105	0,9991
44	0,0535	0,804
45	107	0,9676
49	0,1052	10,000
50	0,0348	0,9832
51	0,0194	0,9894
52	0,0131	0,9749
53	0,0023	0,8745
54	0,0026	0,2983
55	6,18e-04	-0,0039
56	0,2238	0,4826
57	0,3582	0,9871
58	0,0292	0,9643
59	6,37e+02	0,299
60	8,649	0,969
61	3,98e+02	0,2644
62	2,02e+03	0,0813
63	7,05e+02	0,6378
64	1,45e-06	0,925

Uma vez que tanto B_1 quanto B_2 não apresentaram bom desempenho para aproximação linear das variáveis destacadas (Tabela 7) acrescentaremos mais variáveis ao conjunto para elevar o coeficiente de correlação. Por outro lado, temos que 38, 54, 55, 56 e 60 já têm boa correlação (ver Tabela 6) com os dois conjuntos sendo, então, excluídas desta análise.

Observando a Figura 11 podemos visualizar a correlação conjunta entre as variáveis destacadas onde as cores fornecem o grau de associação das mesmas.

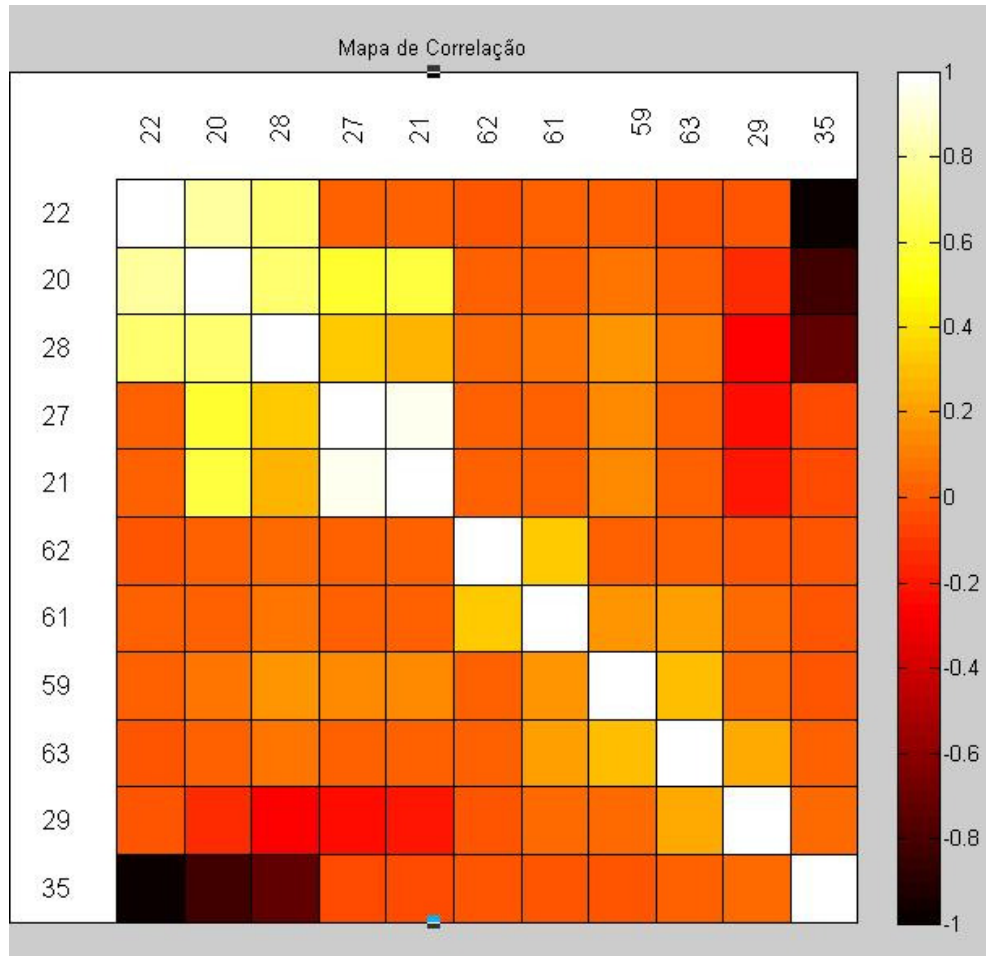


Figura 11. Mapa de Correlação

A figura 11 (Mapa de Correlação) é gerado pelo software Matlab e permite visualizar o grau de correlação entre duas variáveis através da gradação de cores (conforme coluna à direita) além de fazer uma hierarquização e ordenação destes grupos também por meio da variação de cores (do branco à cor negra).

Desta forma, temos os seguintes grupos com alta correlação entre si:

- a) 20, 22, 28 e 35;
- b) 20, 21 e 27;
- c) 20, 21 e 22

Podemos ver na Tabela 8 correlações entre cada par variável do conjunto 20, 21, 22, 27, 28 e 35.

Tabela 8. Correlação Bivariada

	20	21	22	27	28	35
20	1,0000	0,6003	0,8108	0,5763	0,7109	-0,8283
21	0,6003	1,0000	0,0185	0,9542	0,2706	-0,0525
22	0,8108	0,0185	1,0000	0,0218	0,6891	-0,9969
27	0,5763	0,9542	0,0218	1,0000	0,3159	-0,0615
28	0,7109	0,2706	0,6891	0,3159	1,0000	-0,7338
35	-0,8283	-0,0525	-0,9969	-0,0615	-0,7338	1,0000

Resumidamente, mostramos, para cada item acima, como cada variável se comportou na aproximação pela rede neural. Novamente, para cada variável monitorada foi usado como conjunto de entrada as variáveis restantes do conjunto.

Tabela 9: Aproximação por rede neural das variáveis 20, 22, 28 e 35

Variável monitorada	Energia do Erro	Qualidade do ajuste
20	2,5662e+03	0,8296
22	183,7862	0,9866
28	8,8241	0,7994
35	0,3684	0,9929

Tabela 10: Aproximação por rede neural das variáveis 20, 21 e 27

Variável Monitorada	Energia do Erro	Qualidade do ajuste
20	8,1905 e3	0,1470
21	0,0323	1,0000
27	0,0027	1,0000

Tabela 11: Aproximação por rede neural das variáveis 20, 21 e 22

Variável Monitorada	Energia do erro	Qualidade do Ajuste
20	0,5832	1,0000
21	0,0153	1,0000
22	0,1645	1,0000

Dentre os grupos testados somente o último obteve êxito na monitoração das variáveis do próprio conjunto. Utilizando esta tripla como entrada da rede neural vamos verificar seu desempenho na aproximação de 27, 28 e 35 (tabela a seguir).

Tabela 12: Monitoramento de 27, 28 e 35 usando 20, 21 e 22

Variável	Energia do erro	Qualidade do Ajuste
27	0.0017	1.0000
28	8.2805	0.8662
35	2.4475e-004	1.0000

Conforme mostrado, a rede tendo 20, 21 e 22 como entradas apresentou bom desempenho tanto na monitoração das variáveis do próprio conjunto quanto na monitoração de 27 e 35. Esta tripla será anexada ao conjunto B_2 que chamaremos de B_3 , ou seja, $B_3 = \{19.26\ 30\ 39\ 47\ 48\ 20\ 21\ 22\}$.

A seguir, tem-se as tabelas de correlação múltipla entre B_1 , B_2 , B_3 e as variáveis 27 e 35 bem como os resultados do desempenho da rede usando B_3 como entrada.

Tabela 13: Correlação Múltipla

variável	B_1	B_2	B_3
27	0.3679	0.3709	0.9389
35	0.0070	0.0074	1.000

Tabela 14: Monitoramento de: 27 e 35 usando B_3

Variável	Energia do erro	Qualidade do Ajuste
27	0.0017	1.0000
35	2.4475e-004	1.0000

9. DISCUSSÃO DOS RESULTADOS

O objetivo do trabalho foi aplicar a metodologia do teste de correlação na análise de variáveis de entrada de uma rede. Para isso, foi feito um estudo das considerações físicas dos reatores que mostrou que poderíamos tomar um conjunto base de variáveis (conjunto B) tal que a maior parte das variáveis monitoradas pelo SAD tivessem dependência (ou alguma relação) com alguma variável ou, ainda, fosse resultado da interação de mais de uma variável de B.

A correlação múltipla mede o grau de dependência entre dois conjuntos de variáveis e, em particular, entre um conjunto de variáveis (variáveis independentes) e uma variável resposta (variável dependente). A correlação múltipla permite que sejam acrescentadas variáveis a um conjunto sem prejuízo do grau de correlação, ou seja, o valor do coeficiente previamente calculado não diminui se acrescentarmos mais variáveis ao conjunto de variáveis independentes, embora não seja garantido o aumento de seu valor.

A segunda parte consistiu em verificar se todas variáveis de B poderiam ser monitoradas, por uma rede neural, individualmente utilizando como conjunto de entrada as variáveis restantes do conjunto. Como não houve êxito na monitoração de F1M3 acrescentamos ao conjunto a variável F1, pois a correlação entre as mesmas é 0,99.

O conjunto B₂ apresentou uma correlação linear menor que B₁ com relação às variáveis do SAD (Tabela 6) e os resultados quanto à monitoração também foram melhores para B' mostrando que a variável 46 é importante para obtermos melhores resultados ara a rede neural, embora ambas as redes tenham obtido o mesmo desempenho na tarefa de monitoração com relação à quantidade de variáveis, ou seja, ambas as redes tiveram bom desempenho na monitoração do mesmo conjunto de 39 variáveis (Tabela 7).

Para a avaliação do desempenho da rede foram considerados os valores da energia do erro e o teste de correlação onde os dois números têm importância, pois alto grau de correlação e erro pequeno são as condições ideais.

As situações encontradas com relação ao teste empregado são:

- energia do erro pequeno e qualidade do ajuste pequeno;
- energia do erro pequeno e qualidade do ajuste alto;
- energia do erro grande e qualidade do ajuste alto.

Conforme dito em 6.4, para a validação do desempenho da rede a situação que deve ocorrer é um comportamento linear representado pela qualidade do ajuste entre a saída da

rede e o alvo, ou seja, devemos ter a igualdade $T_x = f(x)$ onde T_x é o alvo e $f(x)$ é a saída da rede e uma baixa energia do erro.

Um alto grau de correlação linear entre o conjunto de entrada e variável monitorada não é determinante para termos uma boa aproximação via rede neural. Um exemplo desta situação é a variável 60 que tem correlação linear com B_1 igual 0.9377, boa qualidade de ajuste (0.9889), porém energia do erro bastante alta (da ordem de $E+03$).

Embora B' apresente resultados melhores que B_2 (que diferem somente pela presença e ausência da variável 46 respectivamente), ambos os conjuntos não produziram o aumento do número de variáveis aproximadas tanto pela rede neural quanto linearmente e acréscimo à B_2 da tripla (20, 21, 22) mostrou a viabilidade de se acrescentar variáveis ao conjunto tal que o grau de correlação com a variável que se quer monitorar seja alto. Outro ponto é que o grau de correlação está associado ao sucesso ou não da rede neural, pois pelos resultados obtidos (Tabela 6) temos que todas as variáveis com êxito na aproximação via rede neural apresentam alto grau de correlação com o conjunto de entrada.

9.1. Conclusões

A prévia análise das variáveis monitoradas do reator IEA-R1 primeiramente do ponto de vista fenomenológico e posteriormente aplicando o teste de correlação mostrou ser um método bastante eficiente, pois o estudo dos fenômenos físicos que ocorrem num reator juntamente com o teste de correlação permitiu relacionar variáveis monitoradas do SAD que antes estavam desconexas.

A construção do conjunto inicial através da análise das relações físicas existentes em um reator permitiu construir um conjunto inicial tal que as demais variáveis monitoradas fossem, em sua maior parte, dependentes das primeiras. Essa primeira análise permitiu obter um conjunto com amplo alcance no que se refere ao monitoramento via redes neurais.

Acrescentar variáveis para aumentar o grau de correlação linear entre a variável predita e um conjunto de entrada possibilitou obter um conjunto de 9 variáveis (19, 20, 21, 22, 26, 30, 39, 47, 48) tal que fosse possível aproximar pela rede neural um total de 51 variáveis. As variáveis em que não houve êxito na monitoração foram: 5, 23, 29, 38, 46, 54, 55, 56, 59, 60, 61, 62 e 63.

O prévio cálculo da dependência entre conjunto de variáveis através da metodologia de Análise de Correlações Canônicas permitiu construir esse conjunto revelou ser uma ferramenta efetiva na aproximação das variáveis, pois quanto maior o coeficiente de correlação maior é a probabilidade de êxito das redes neurais. Esta maior probabilidade pode ser justificada pelo fato de quanto maior o coeficiente de correlação mais semelhante será o padrão de comportamento das variáveis ou conjunto de variáveis em questão facilitando o trabalho da rede em mapeá-los.

Recomendações para trabalhos futuros

- . Monitorar as 13 variáveis do reator não conseguidas neste trabalho usando redes neurais;
- . Construir uma rede neural mais geral que aproxime todas as variáveis do SAD;
- . Minimizar o conjunto de variáveis de entrada que, ainda assim, efetue o trabalho de monitoração de todas as variáveis do reator IEA-R1.
- . Testar métodos de análise de correlação não lineares para análise de variáveis de entrada numa rede neural.

10. Anexos

A seguir, têm-se os gráficos das variáveis de operação do reator do IPEN IEA-R1 referentes à semana de operação de 19 a 22 de janeiro de 2009 (dados normalizados).

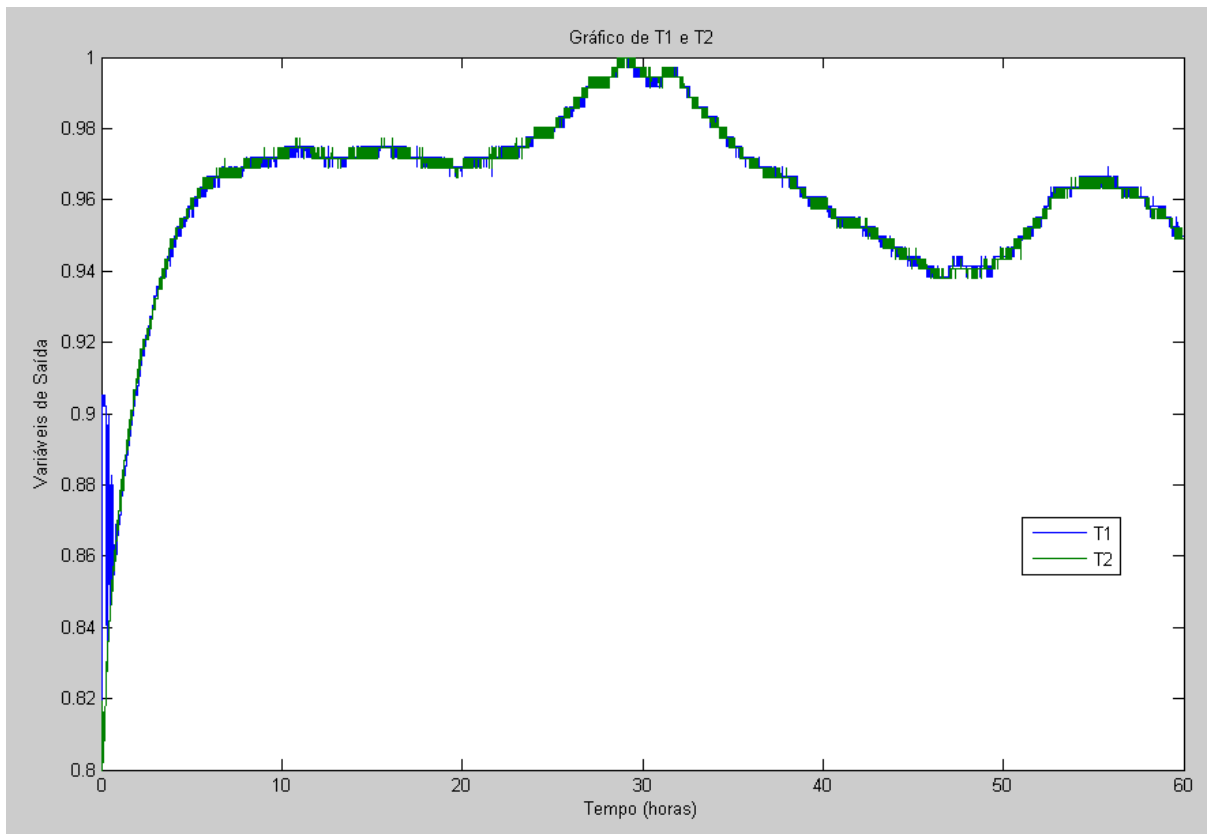


Figura 1: Dados de operação das variáveis T1 e T2.

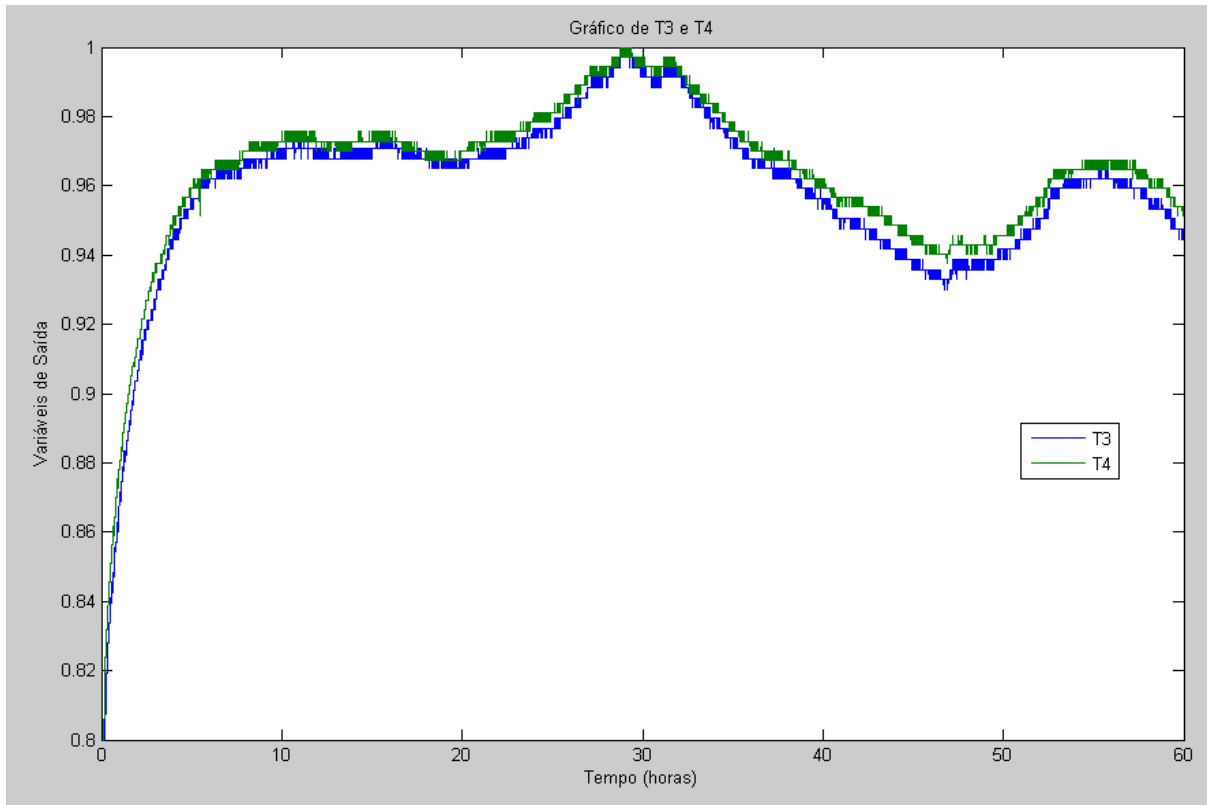


Figura 2: Dados de operação das variáveis T3 e T4.

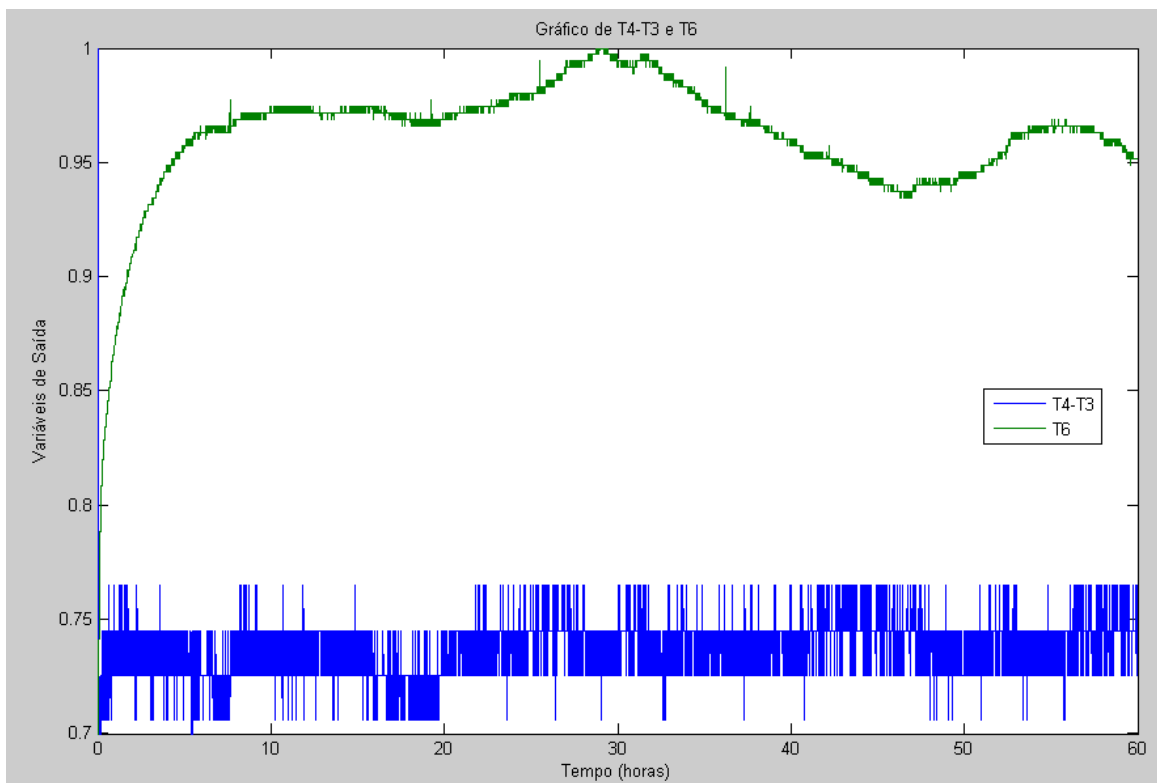


Figura 3: Dados de operação das variáveis T4-T3 e T6.

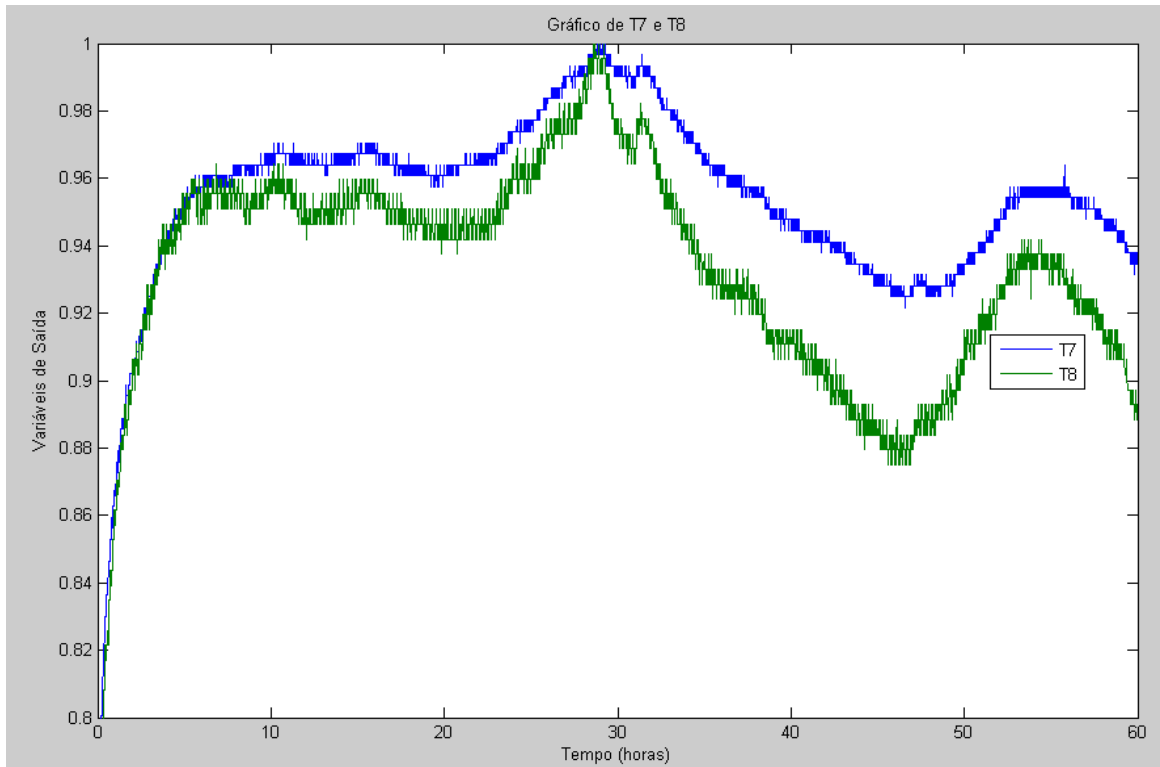


Figura 4: Dados de operação das variáveis T7 e T8.

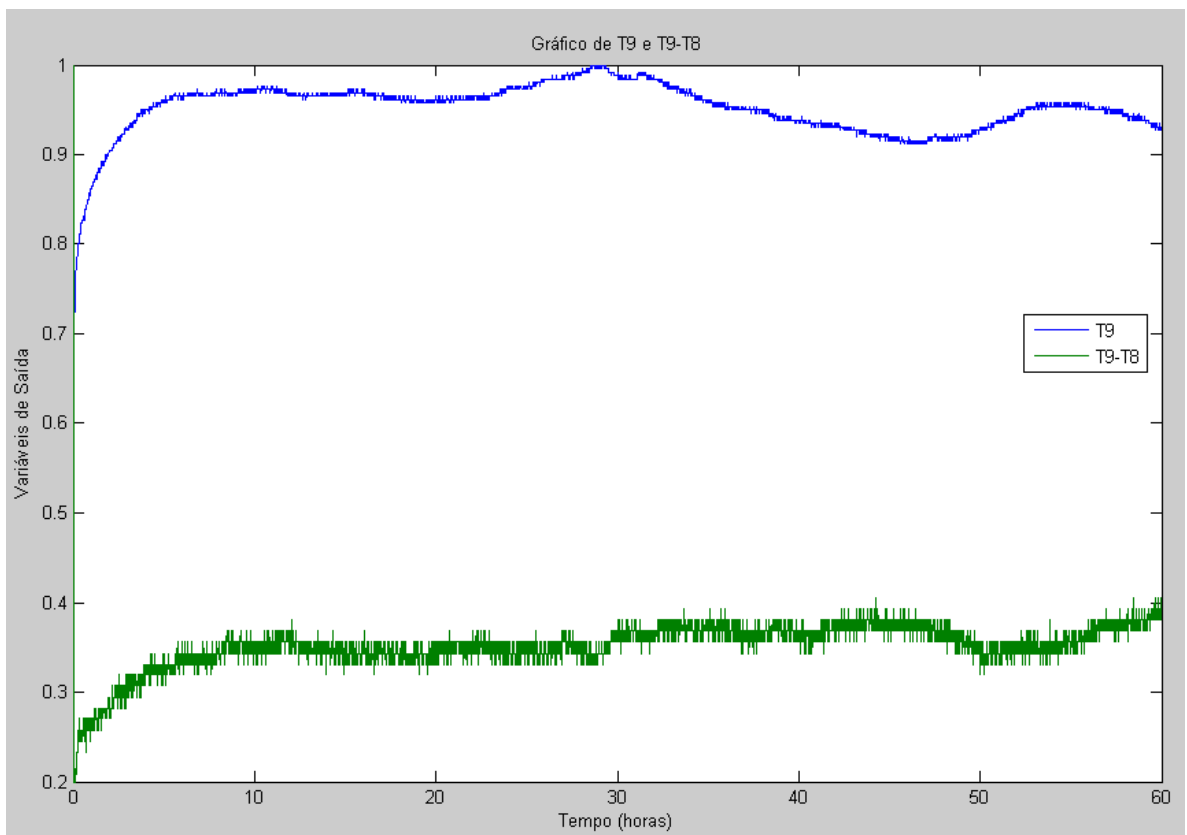


Figura 5: Dados de operação das variáveis T9 e T9-T8.

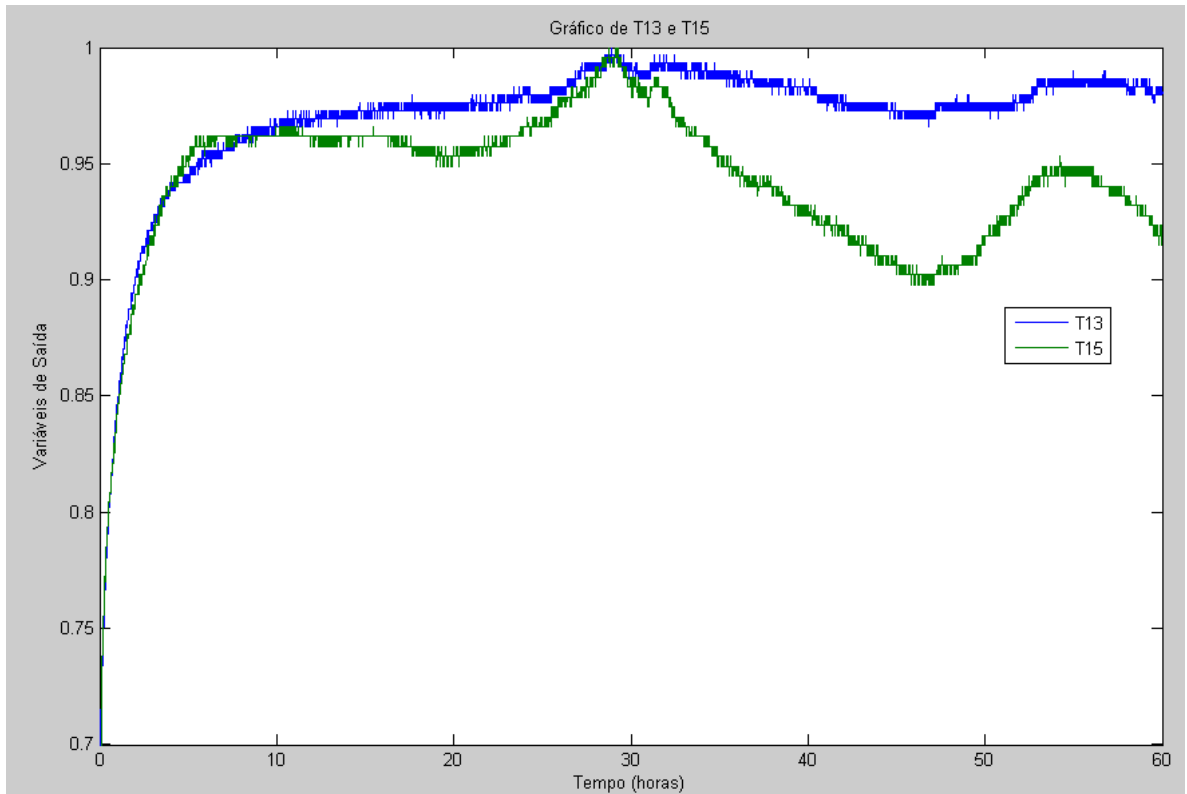


Figura 6: Dados de operação das variáveis T13 e T15.

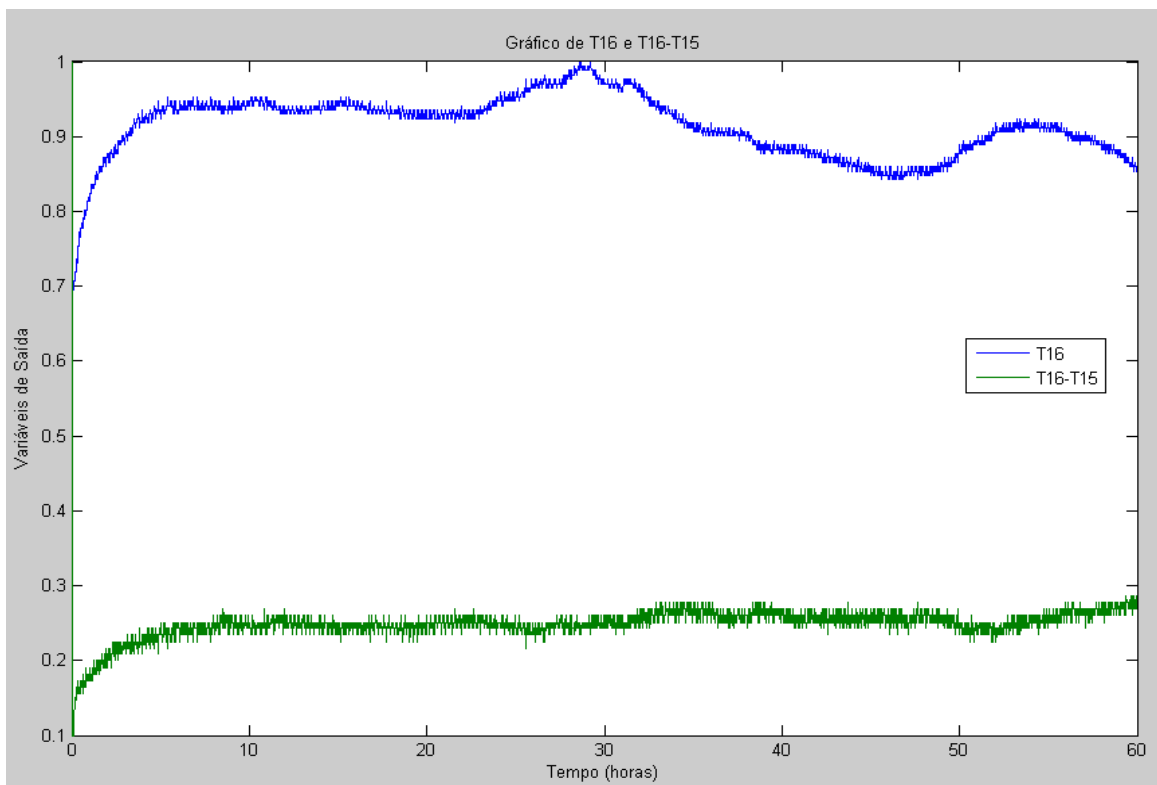


Figura 7: Dados de operação das variáveis T16 e T16-T15.

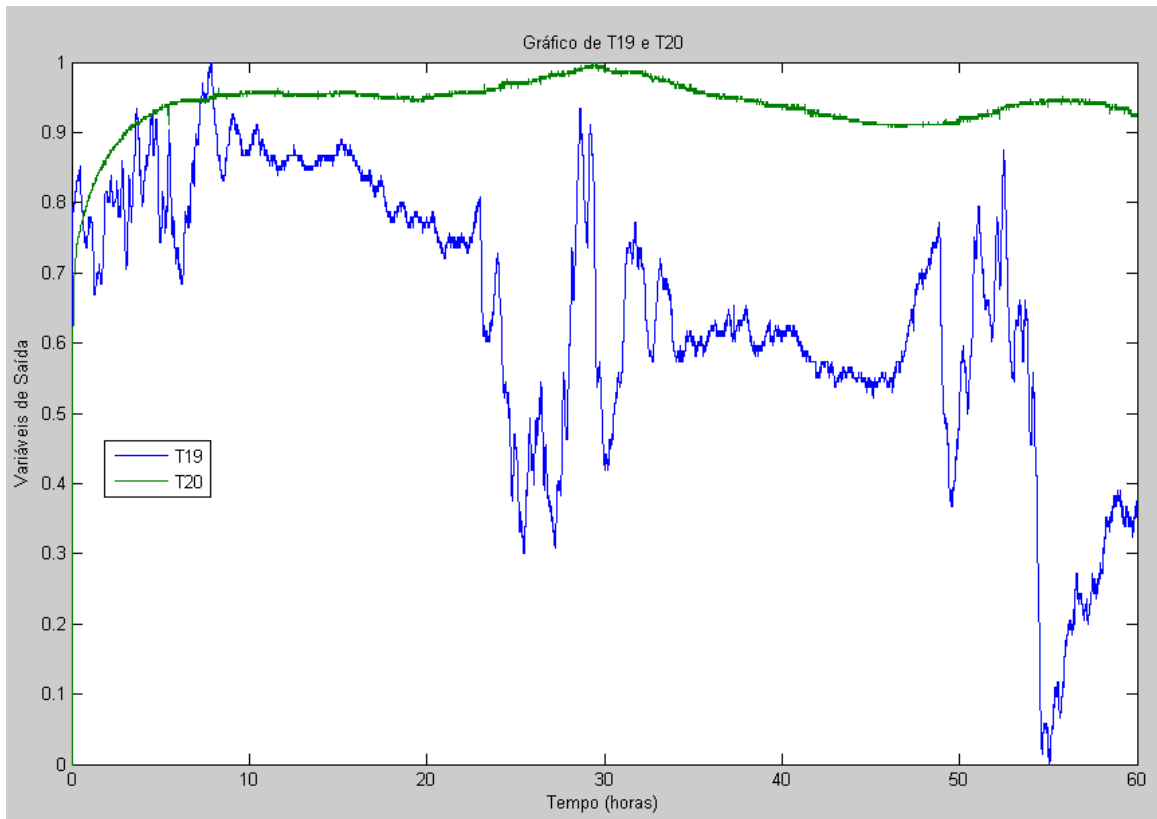


Figura 8: Dados de operação das variáveis T19 e T20.

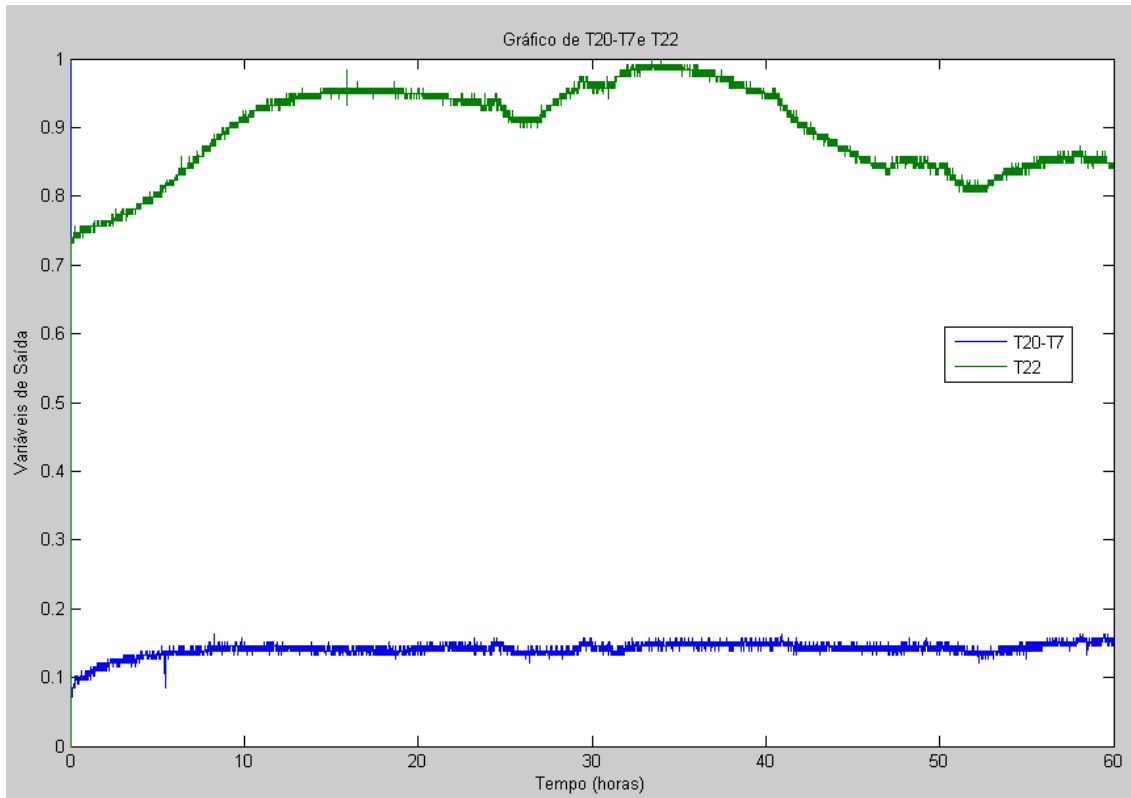


Figura 9: Dados de operação das variáveis T20-T7 e T22.

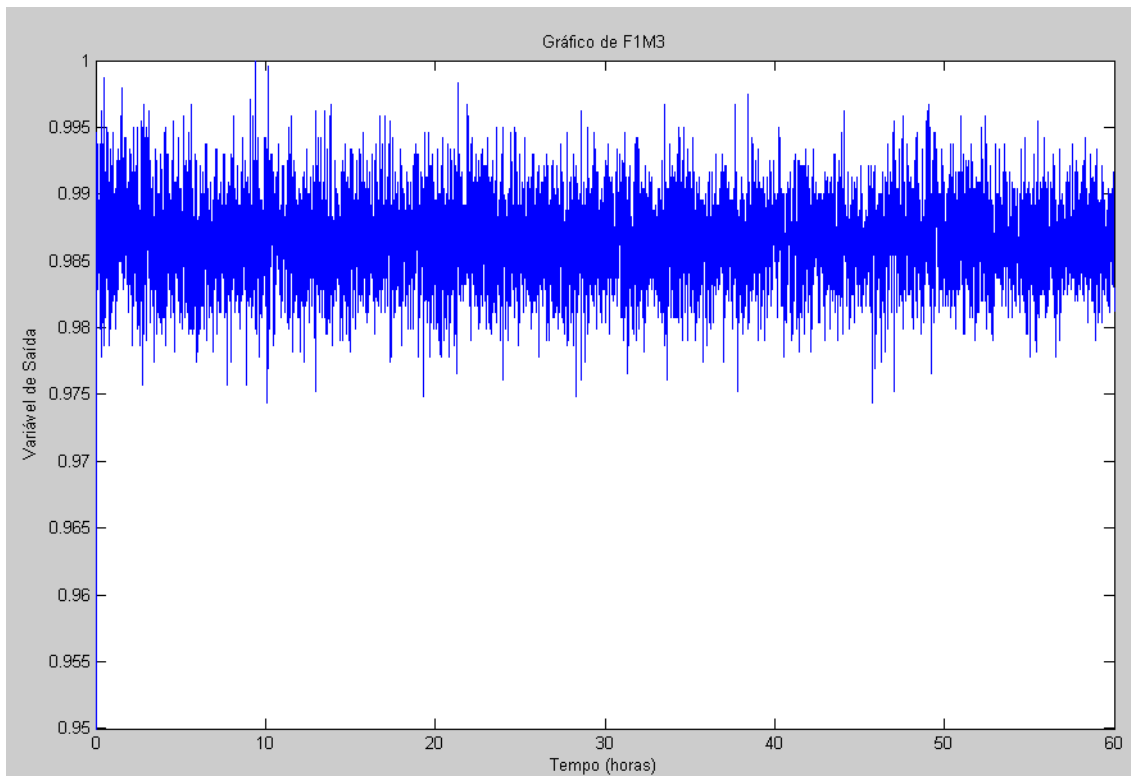


Figura 10: Dados de operação da variável F1M3.

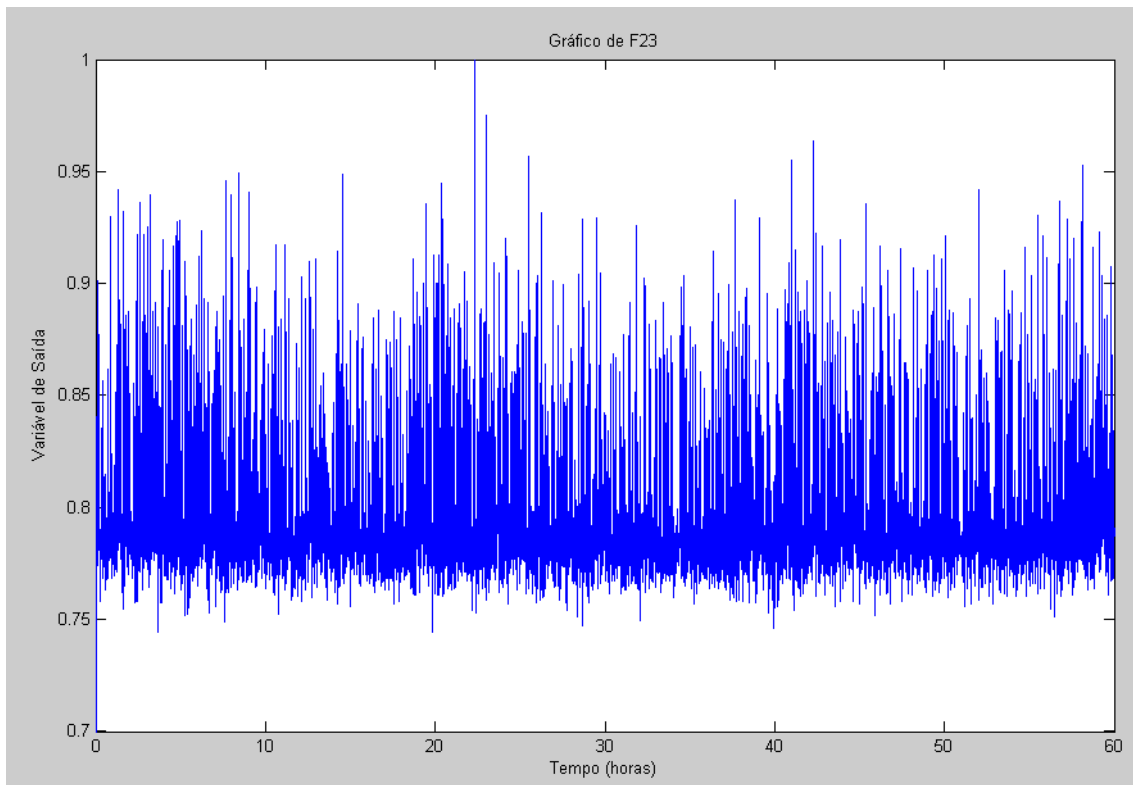


Figura 11: Dados de operação da variável F23.

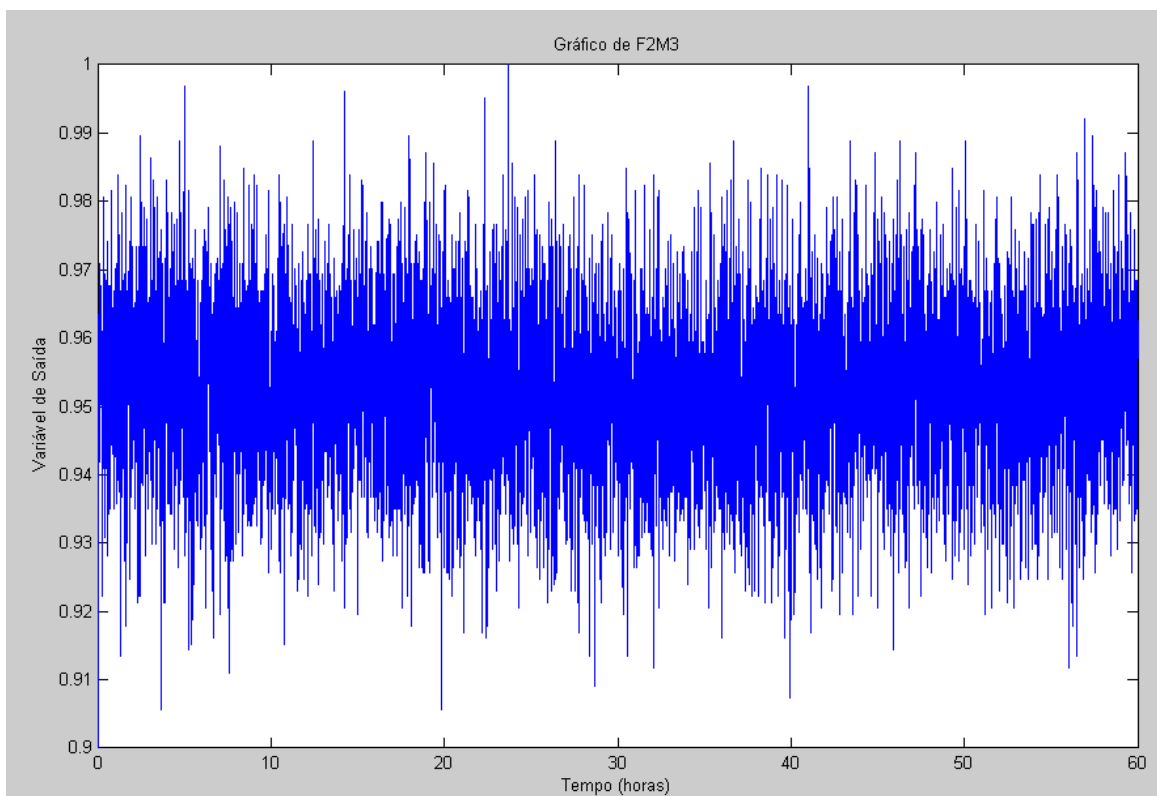


Figura 12: Dados de operação das variável F2M3.

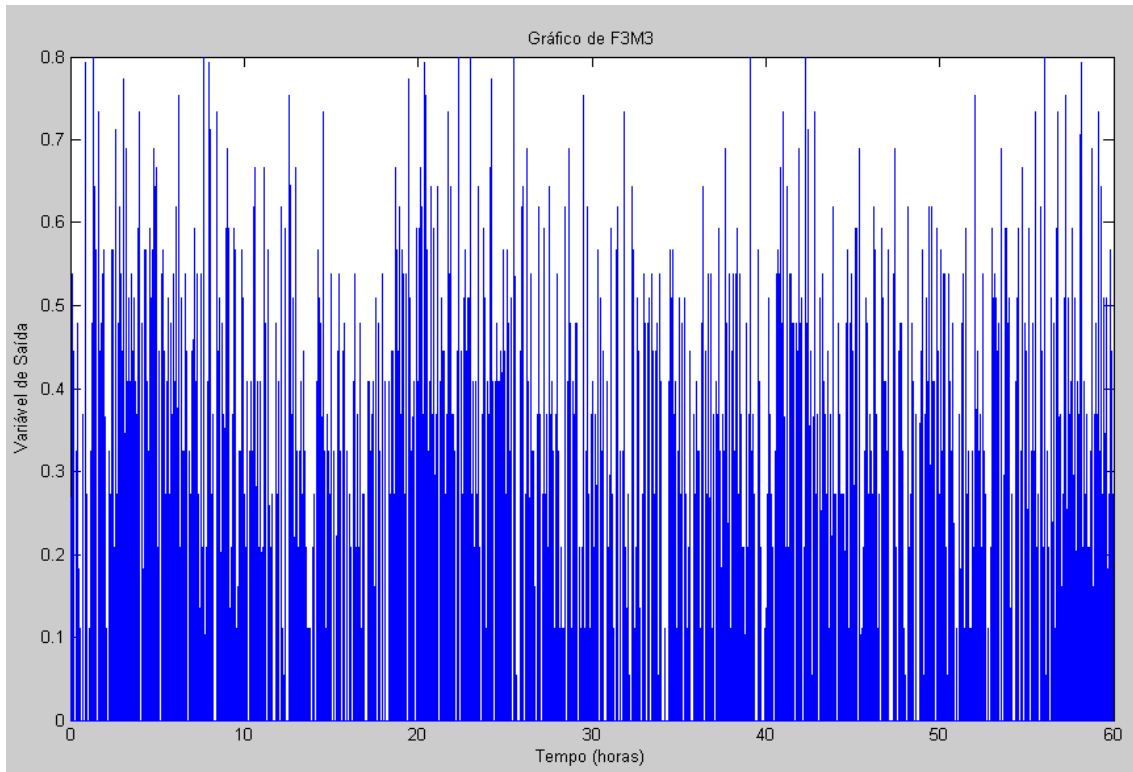


Figura 13: Dados de operação da variável F3M3.

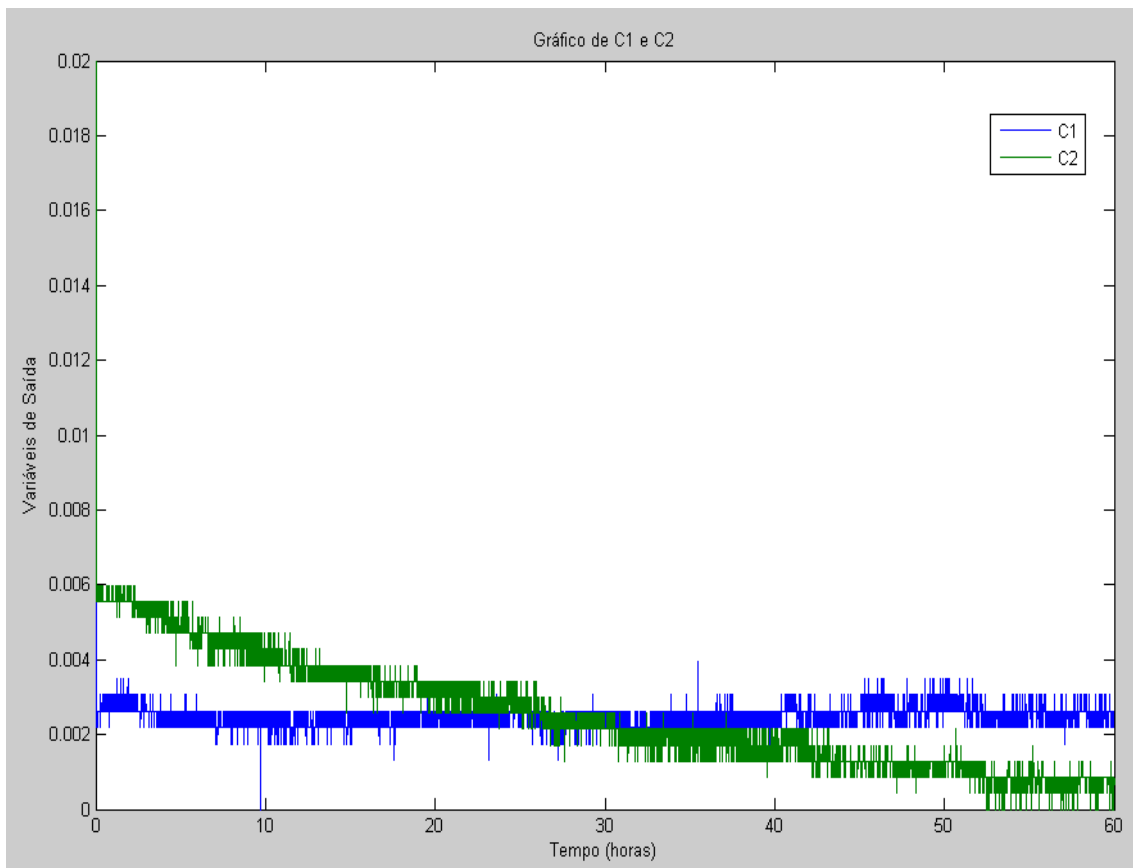


Figura 14: Dados de operação das variáveis C1 e C2.

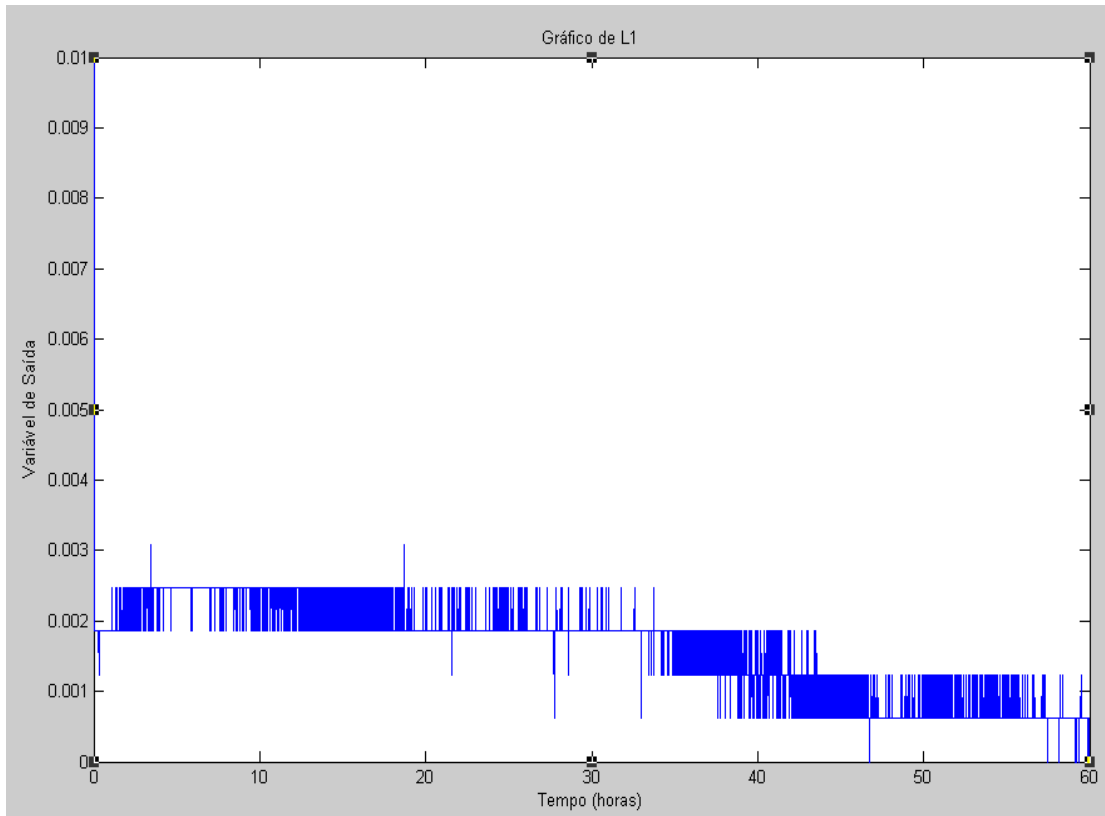


Figura 15: Dados de operação da variável L1.

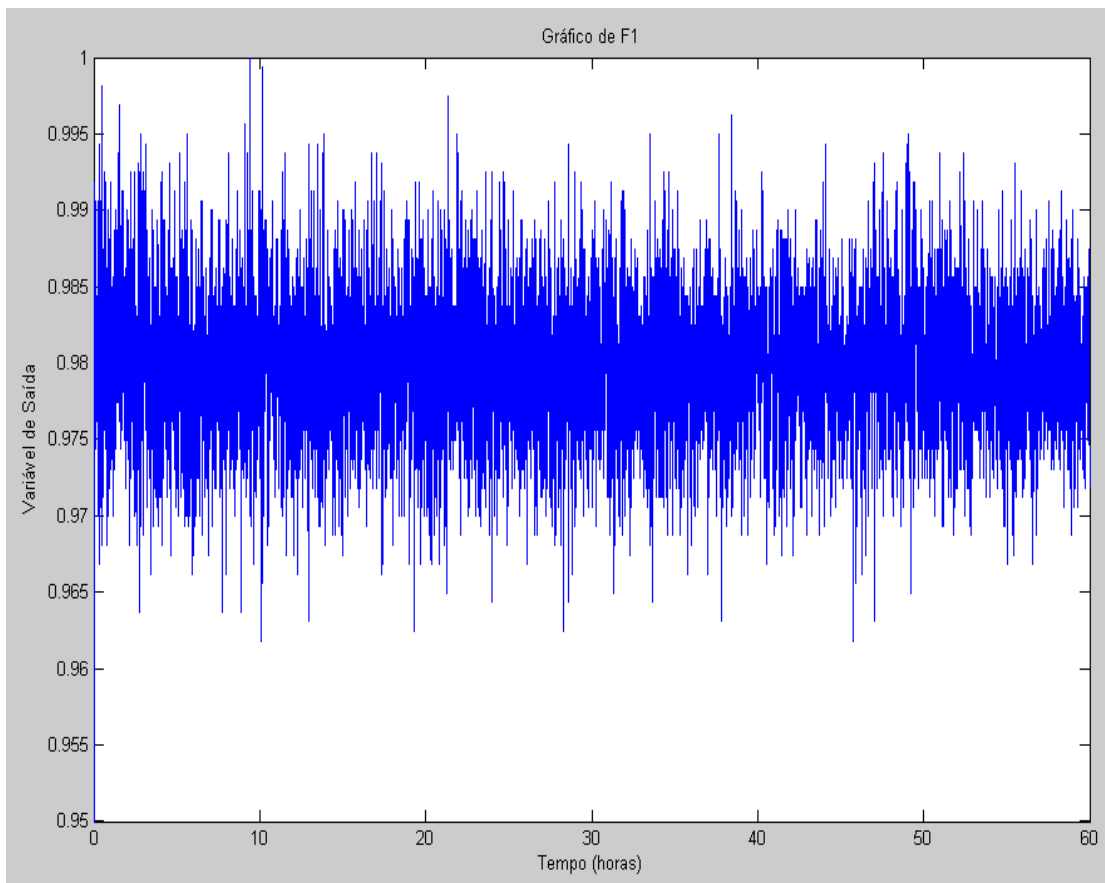


Figura 16: Dados de operação da variável F1.

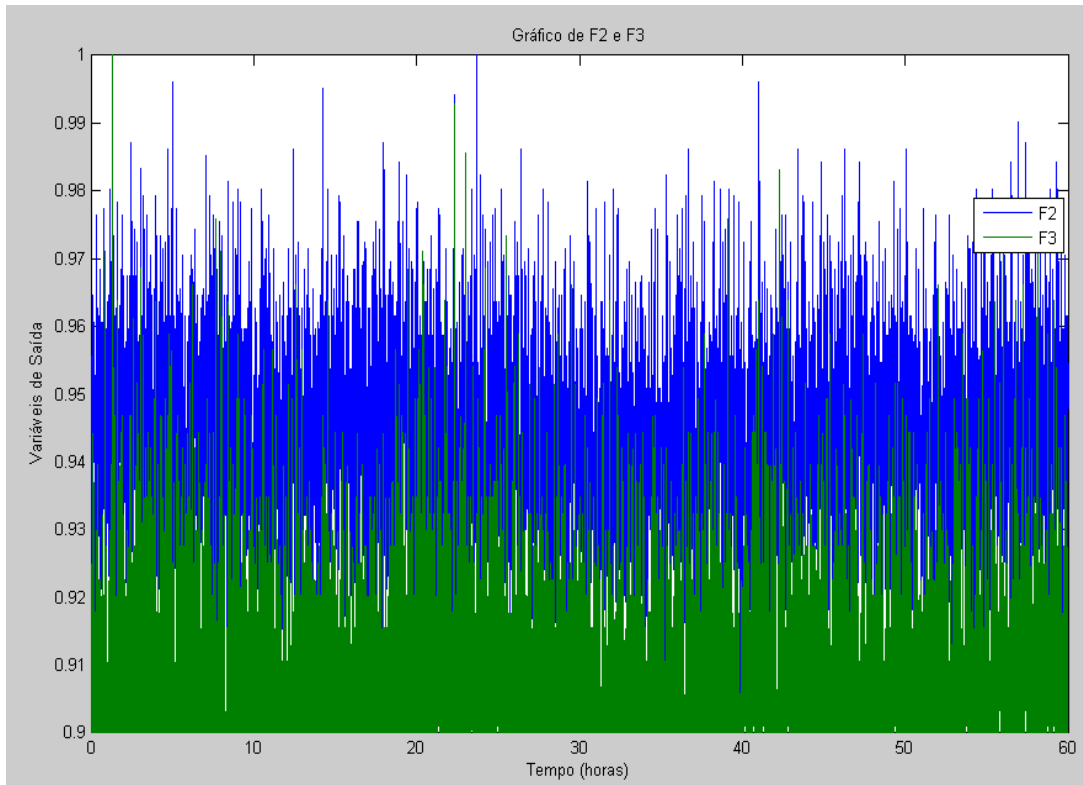


Figura 17: Dados de operação das variáveis F2 e F3.

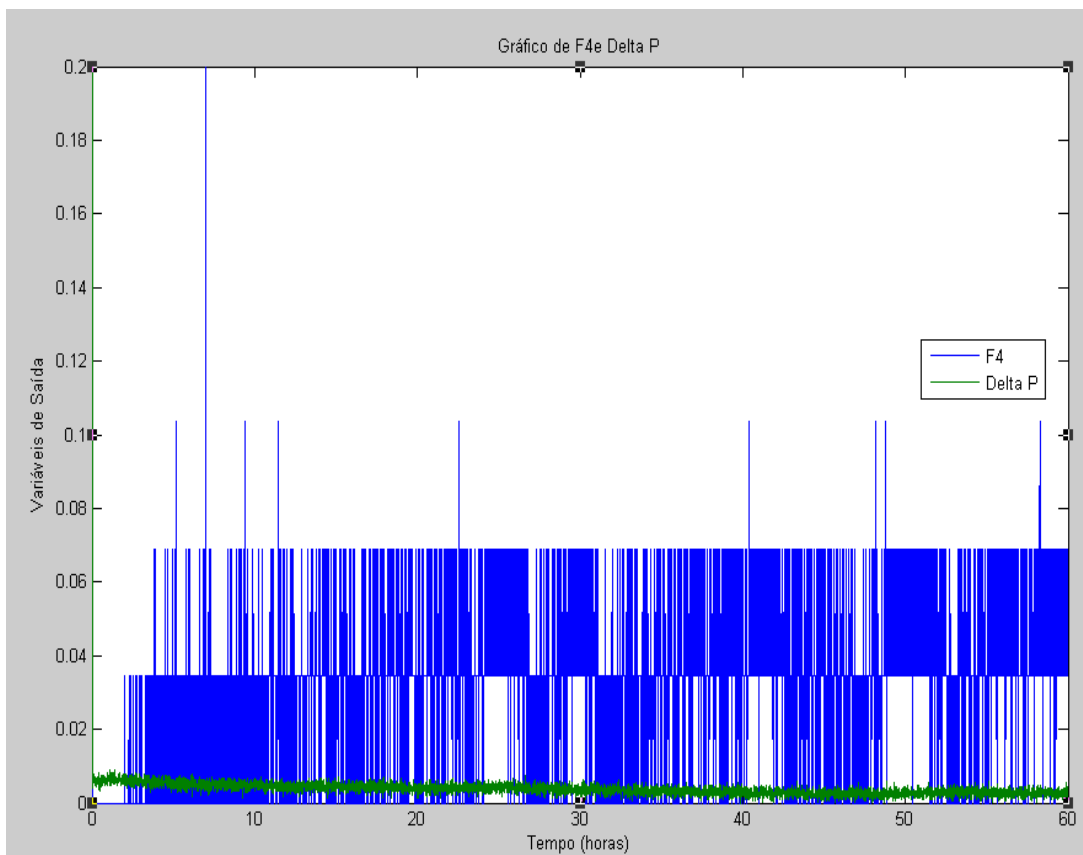


Figura 18: Dados de operação das variáveis F4 e Delta P.

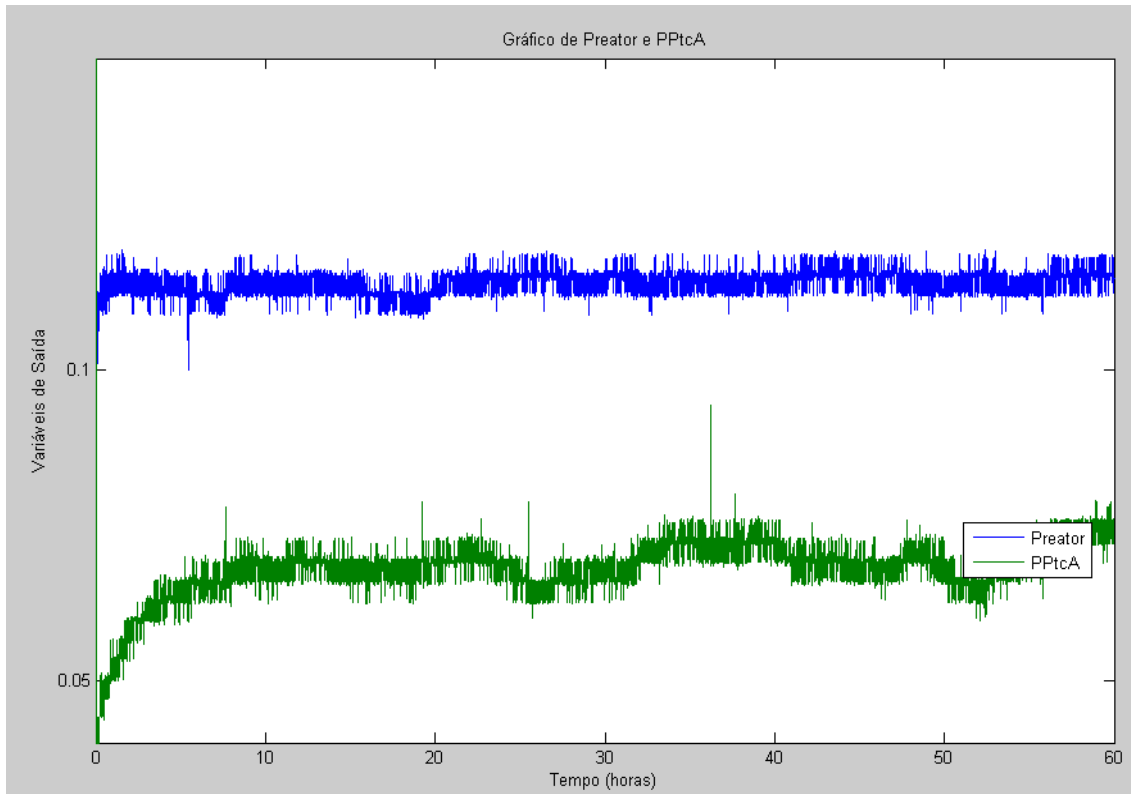


Figura 19: Dados de operação das variáveis Preator e PPtcA

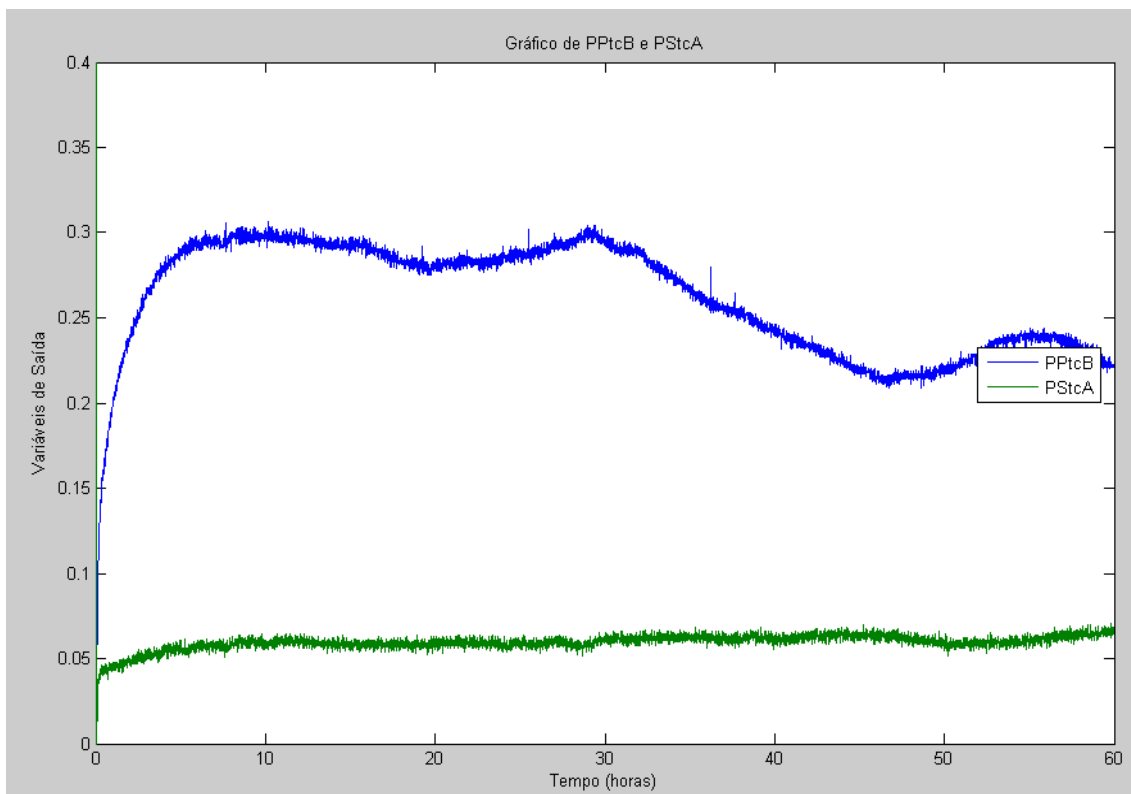


Figura 20: Dados de operação das variáveis PPtcB e PStcA.

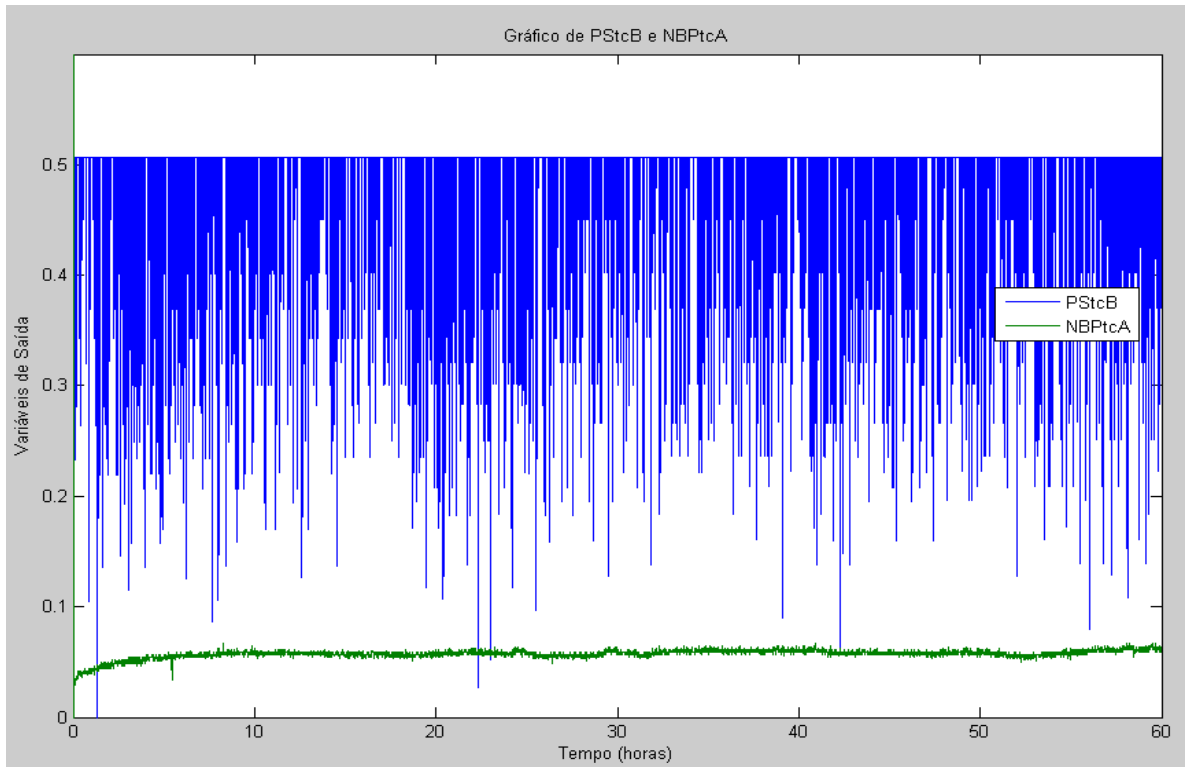


Figura 21: Dados de operação das variáveis PStcB e NBPtCA.

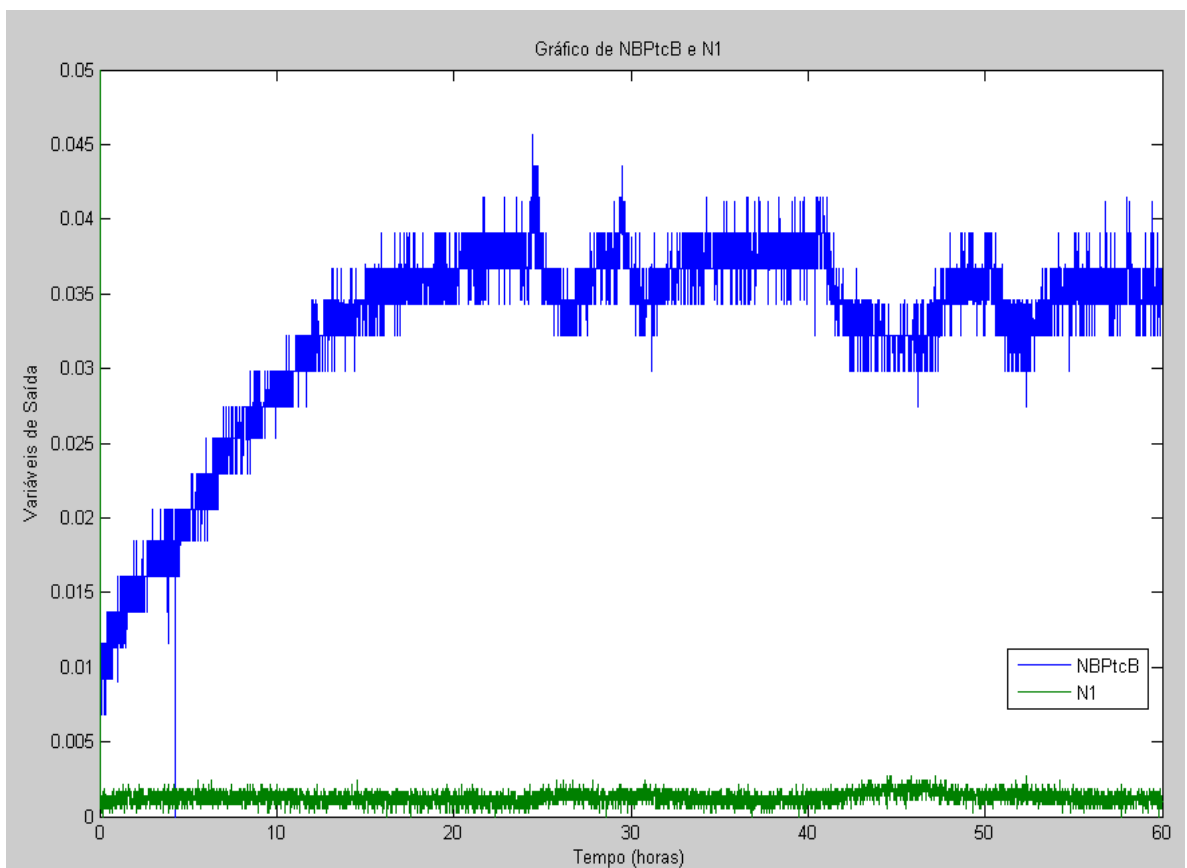


Figura 22: Dados de operação das variáveis NBPtCB e N1.

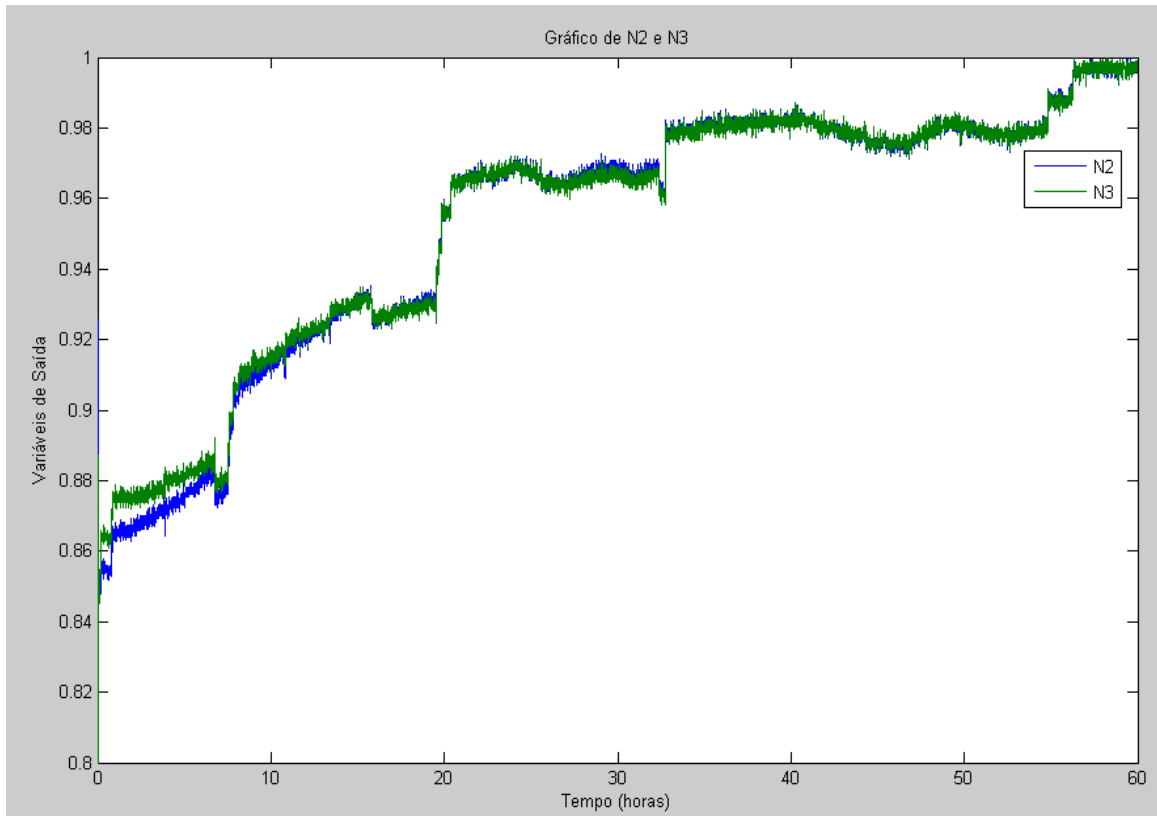


Figura 23: Dados de operação das variáveis N2 e N3.

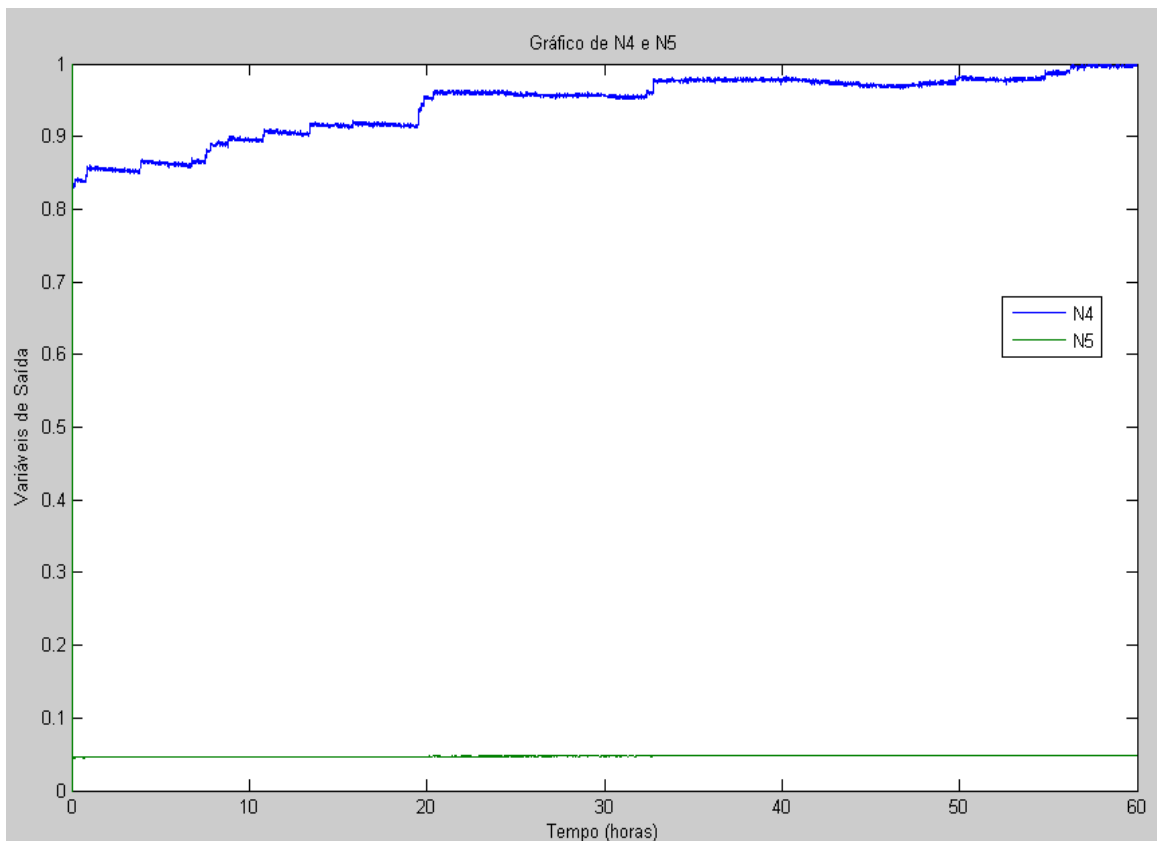


Figura 24: Dados de operação das variáveis N4 e N5.

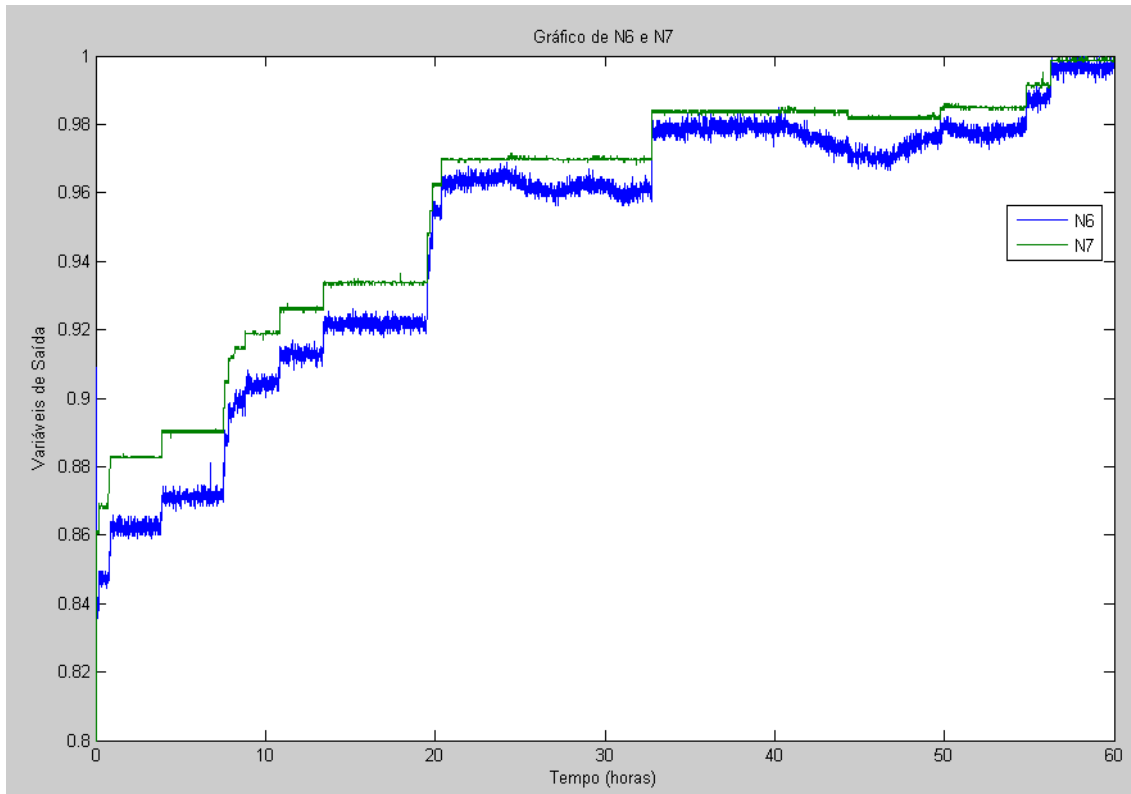


Figura 25: Dados de operação das variáveis N6 e N7.

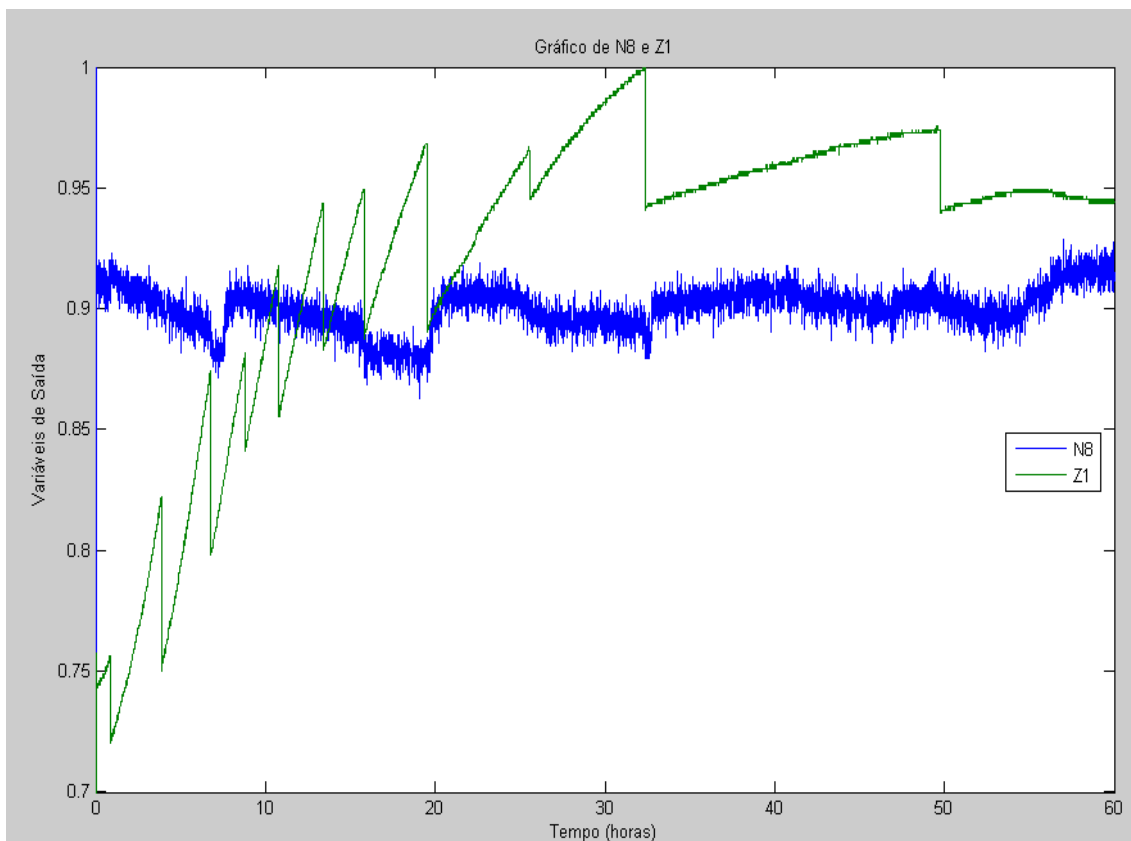


Figura 26: Dados de operação das variáveis N8 e Z1.

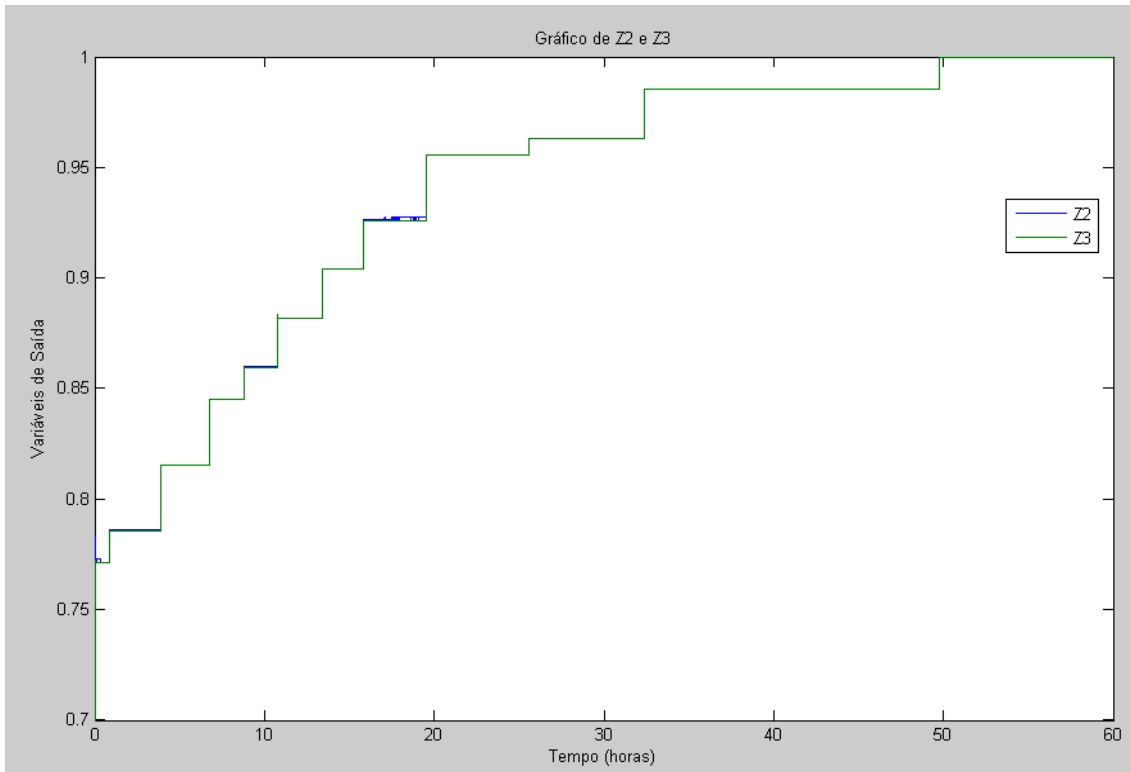


Figura 27: Dados de operação das variáveis Z2 e Z3.

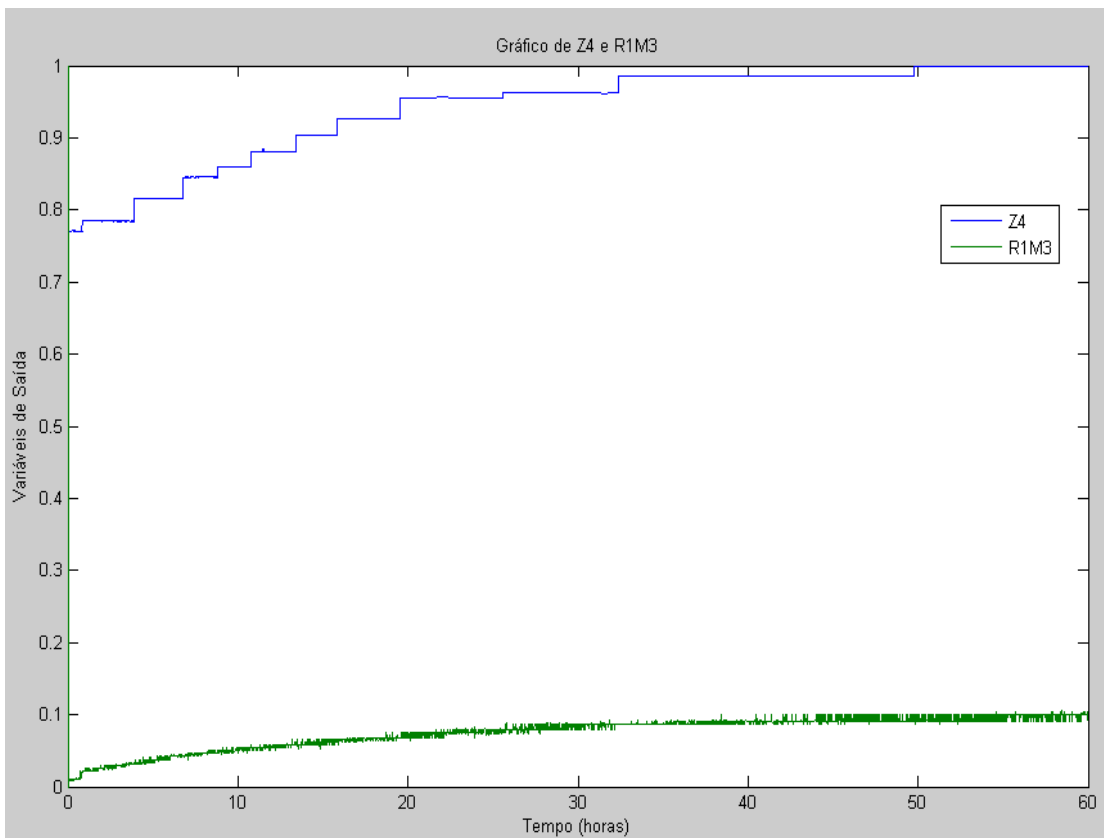


Figura 28: Dados de operação das variáveis Z4 e R1M3.

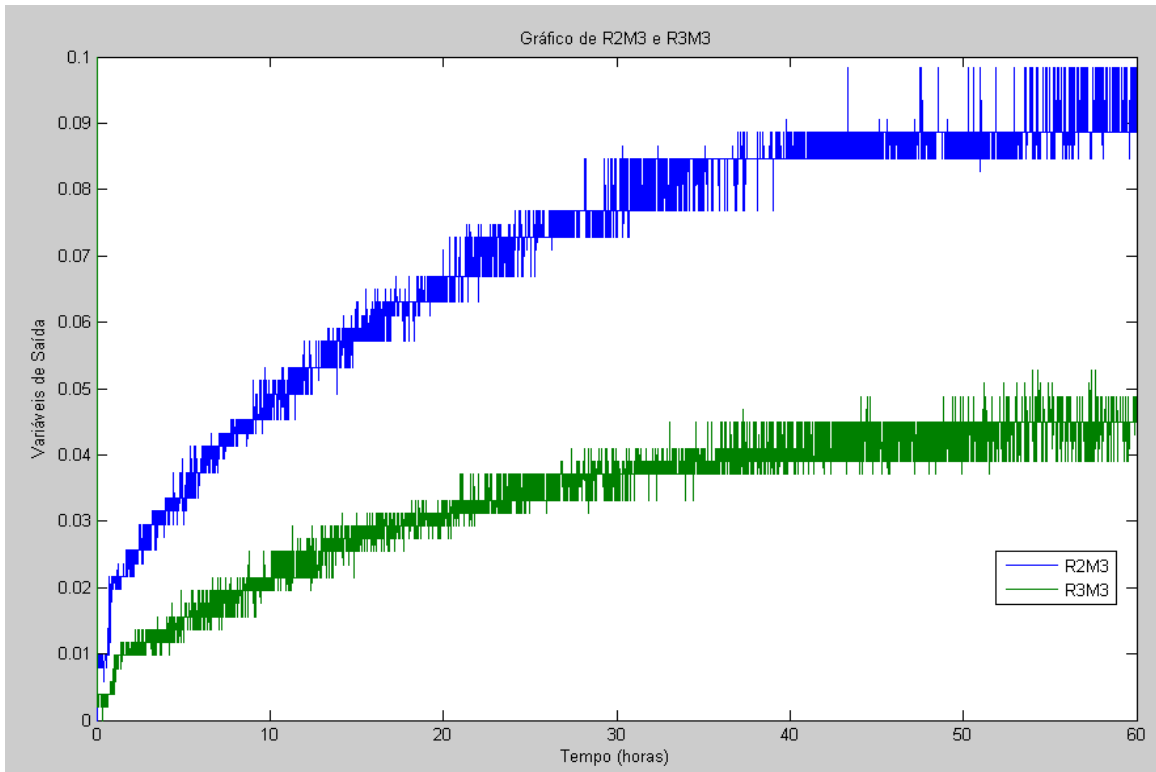


Figura 29: Dados de operação das variáveis R2M3 e R3M3.

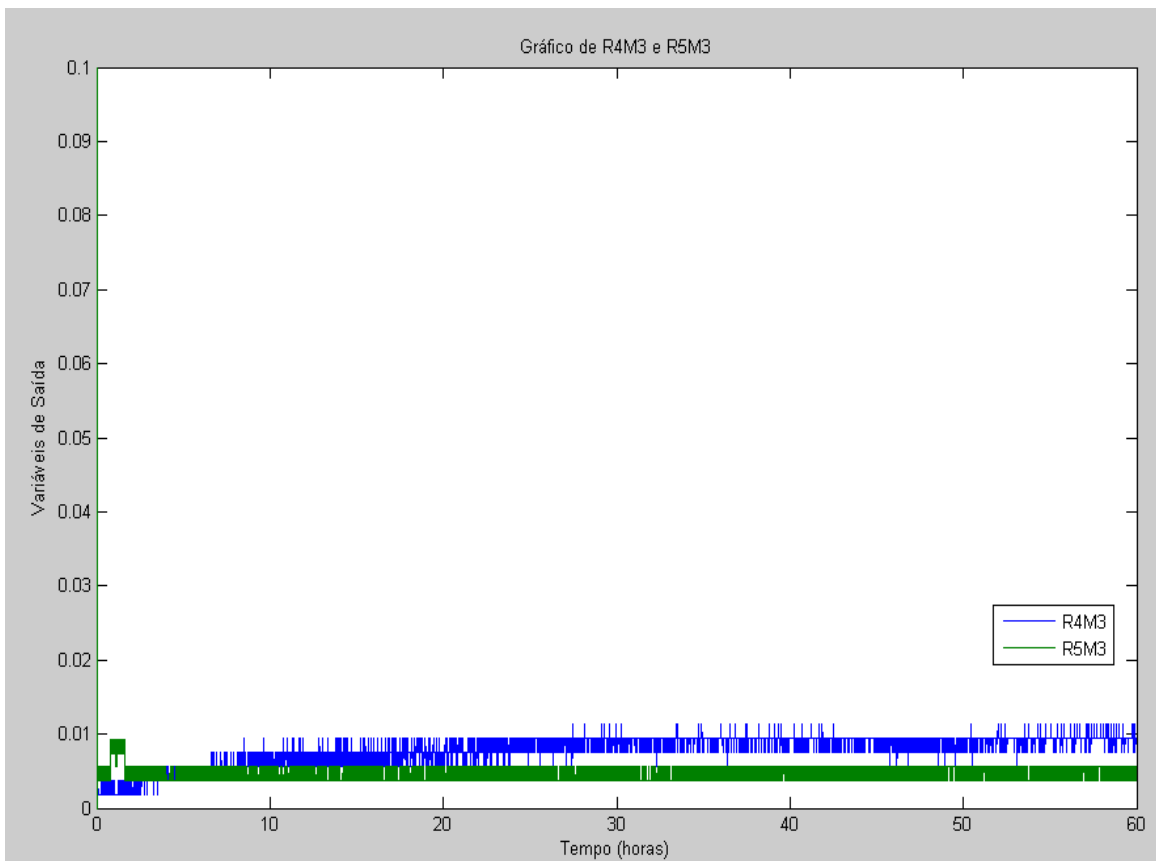


Figura 30: Dados de operação das variáveis R4M3 e R5M3.

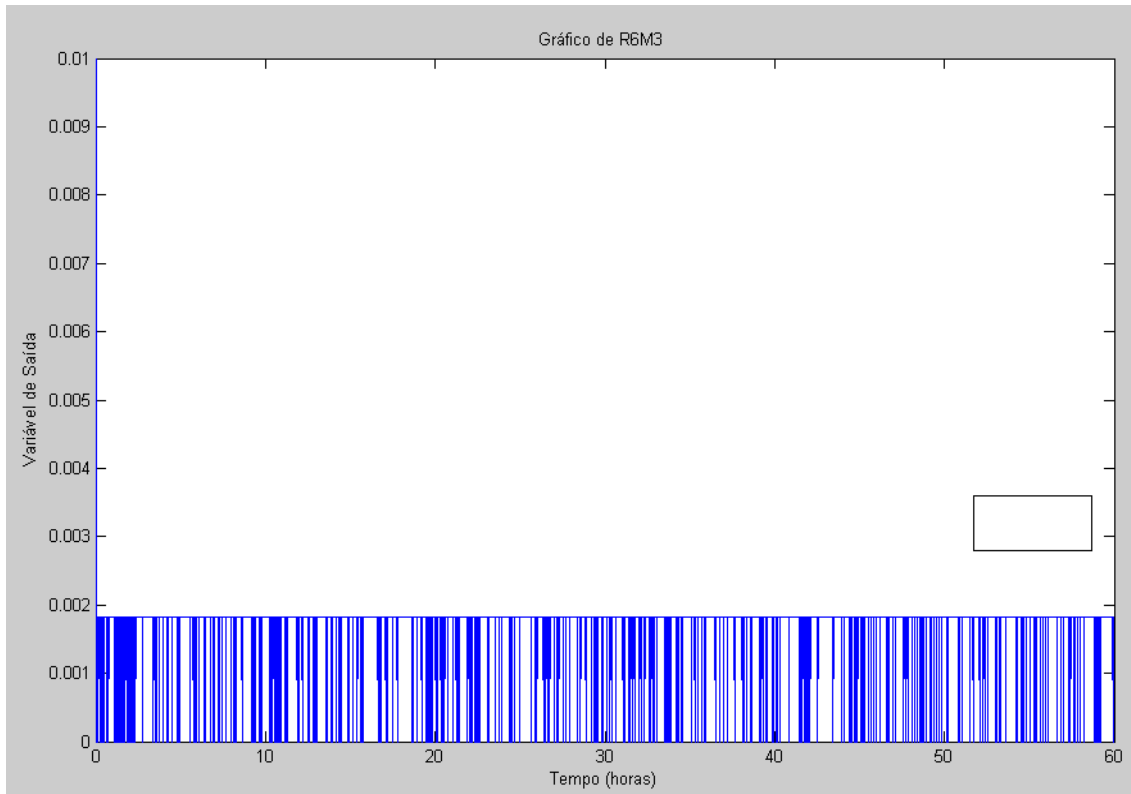


Figura 31: Dados de operação da variável R6M3.

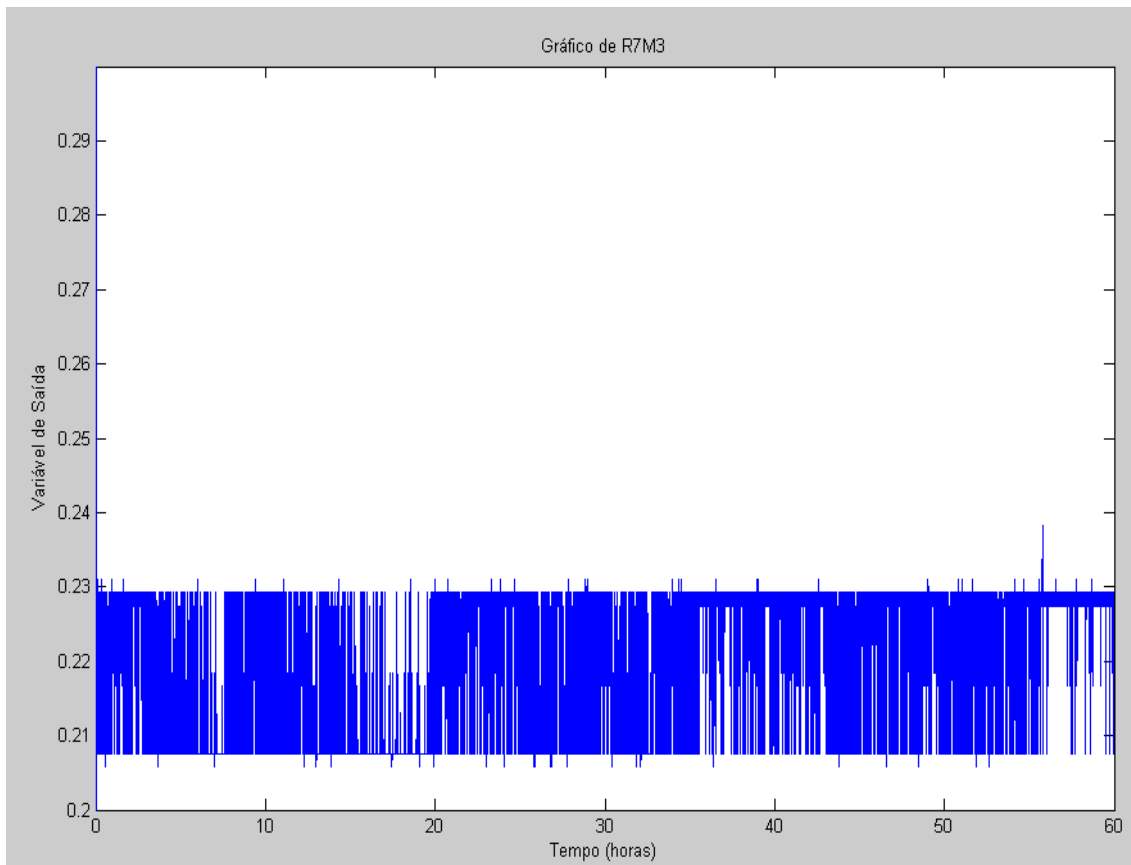


Figura 32: Dados de operação das variáveis R7M3.

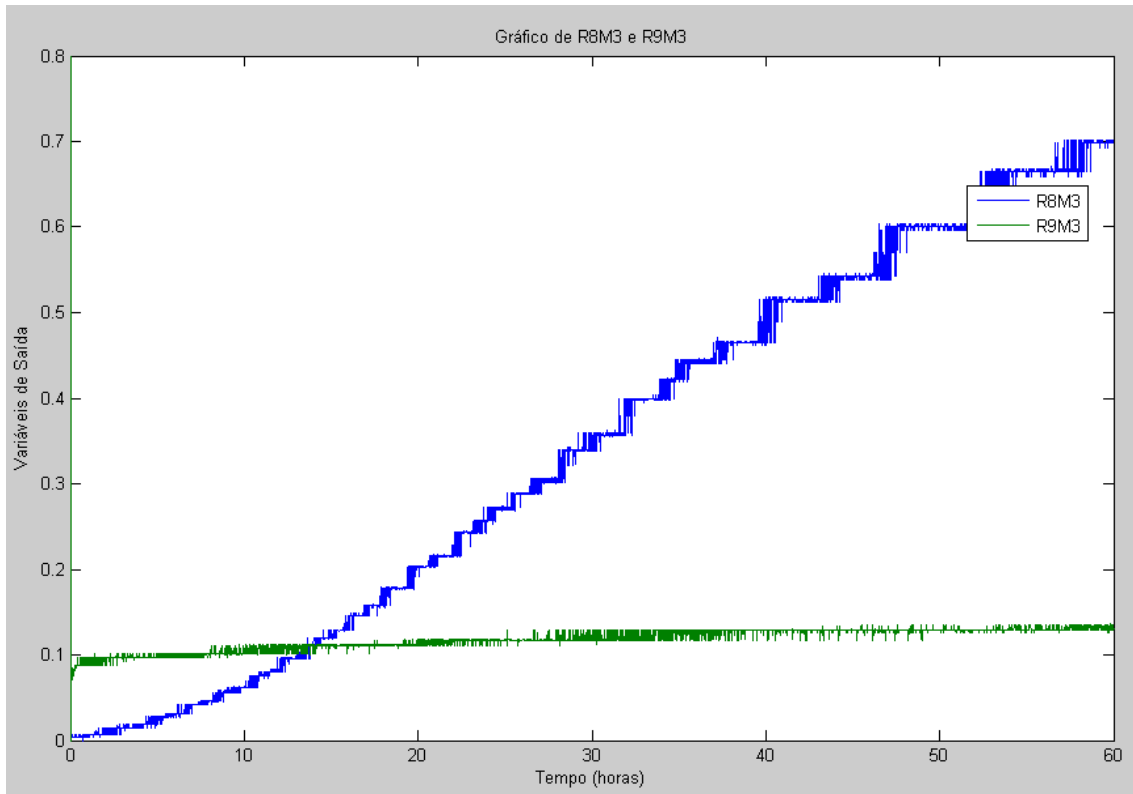


Figura 33: Dados de operação das variáveis R8M3 e R9M3.

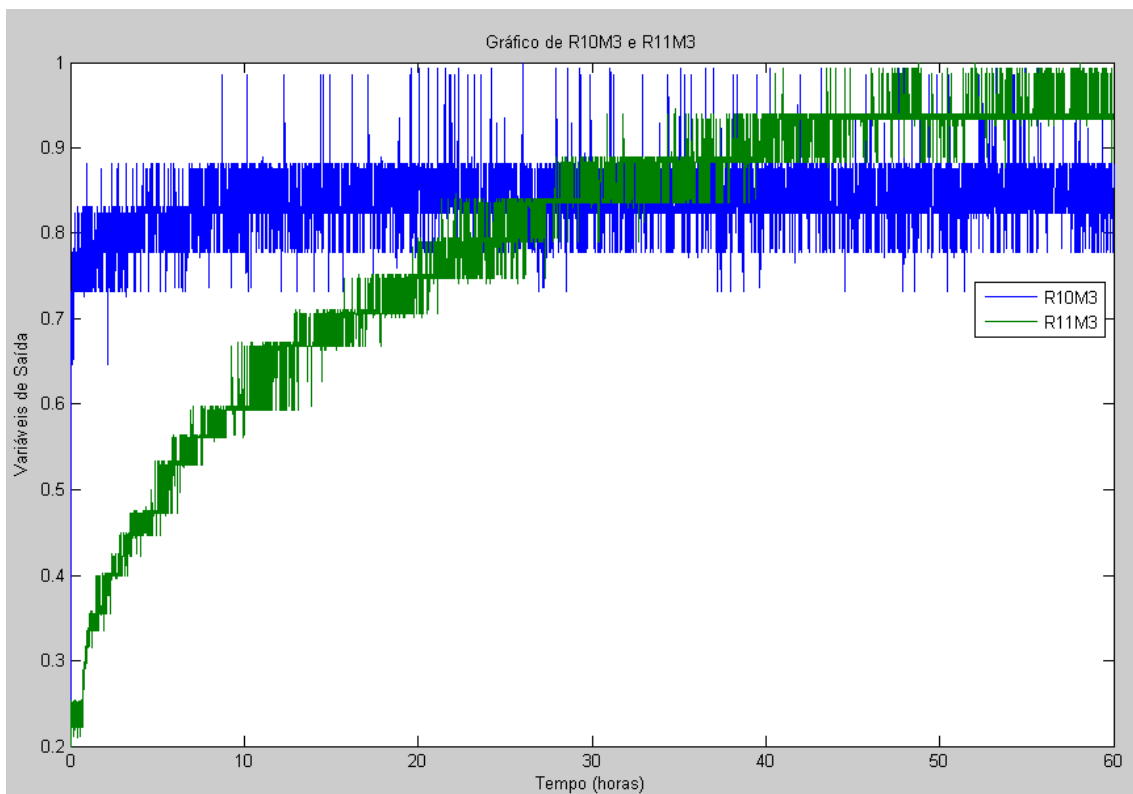


Figura 34: Dados de operação das variáveis R10M3 e R11M3.

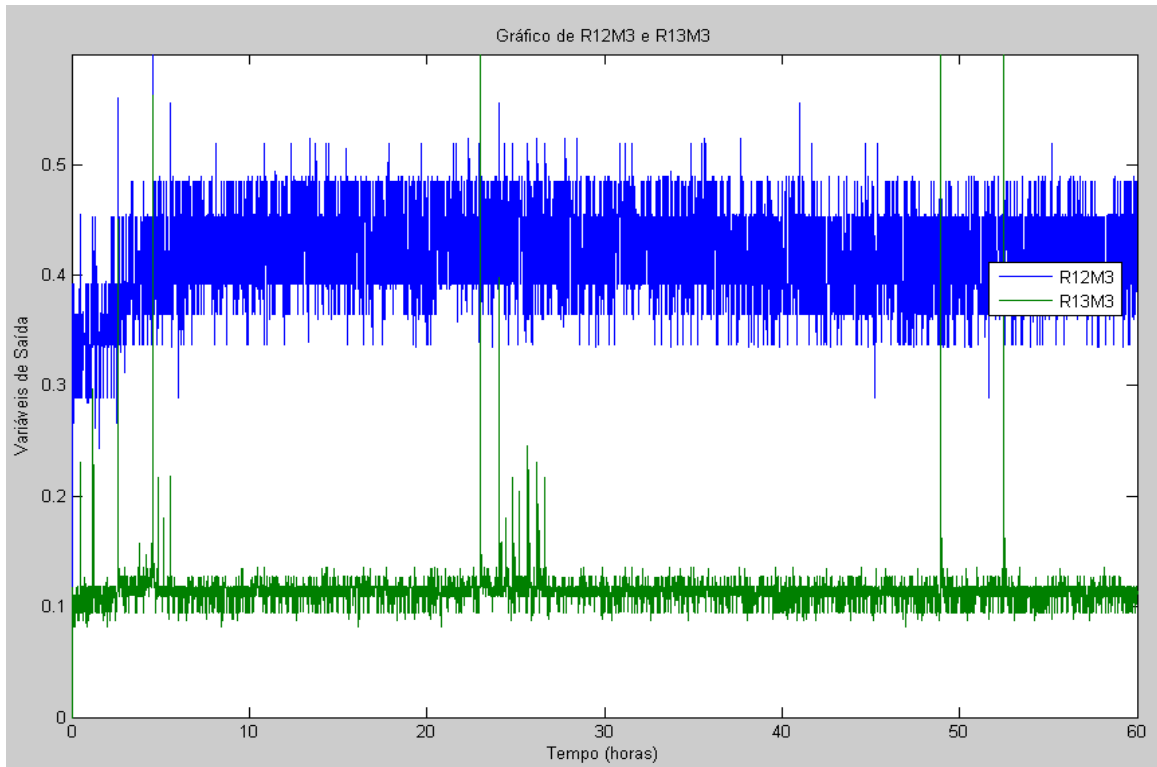


Figura 35: Dados de operação das variáveis R12M3 e R13M3.

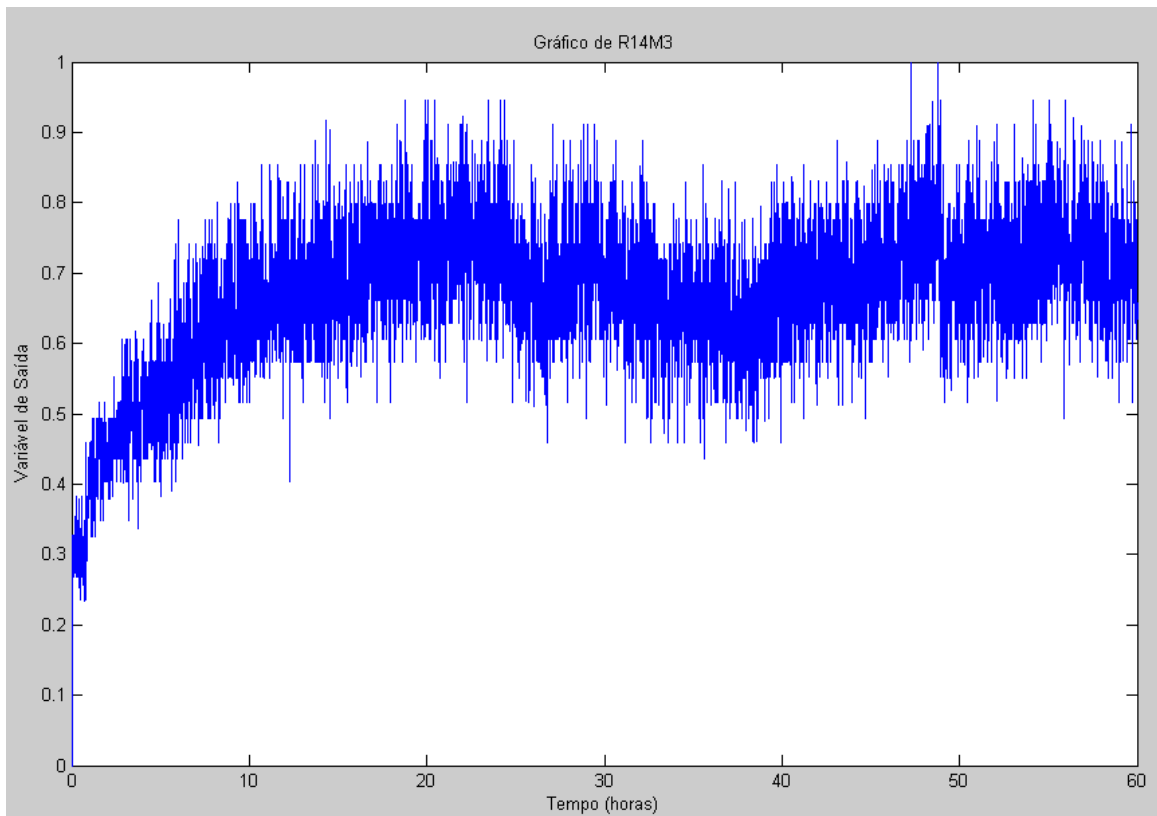


Figura 36: Dados de operação da variável R14M3.

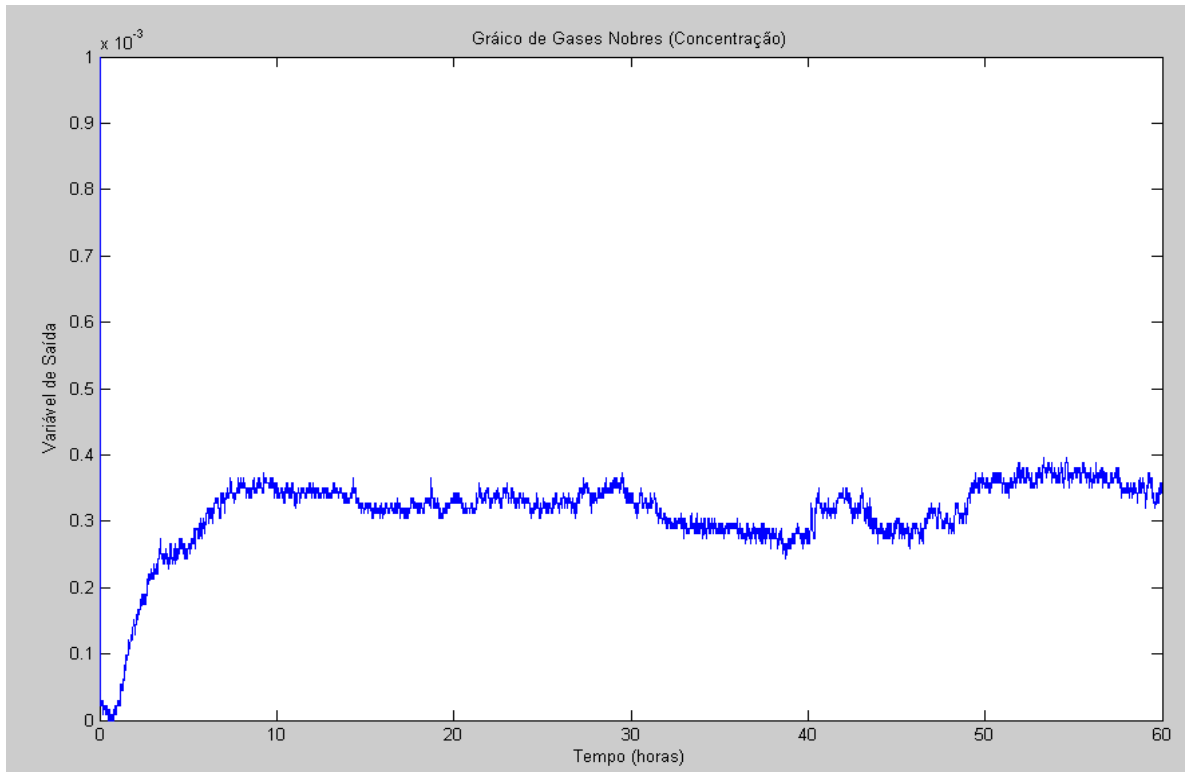


Figura 37: Dados de operação da variável Gases Nobres.

Referências Bibliográficas

- [1] Anders, U.; Korn, O. **Model selection in neural networks** Neural Network, v 12 , pp 309-323, 1999.
- [2] Attali, J. G.; Pagès, G. **Approximations of Functions by a Multilayer Perceptron: a new approach** Neural Networks, v 10, pp 1069-1081, 1997.
- [3] Bear, M.F.; Connors, B.W.; Paradiso, M. A. **Neurociências – Desvendando o sistema nervoso** 3ª ed., 2008.
- [4] Benardos, P. G., Vosniakos, C. **Optimizing feedforward artificial neural network architecture** Engineering Applications of Artificial Intelligence, v 20, Issue 3, April 2007, Pages 365-382
- [5] Boukhezzar, B.; Siguerdidjane, H., **Comparison between linear and nonlinear control strategies for variable speed wind turbines** Control Engineering Practice, v 18, pp 1357-1368, 2010.
- [6] Braga, A. P.; Carvalho, A. P. L. F.; Ludemir, T. B., **Fundamentos De Redes Neurais Artificiais**, Rio De Janeiro: 11º Escola De Computação/Universidade Federal do Rio de Janeiro, 1998.
- [7] Braga, A. P.; Ludemir, T. B.; Carvalho, A., **Redes Neurais Artificiais - Teoria E Aplicações**, Livros Técnicos E Científicos, São Paulo, 1999.
- [8] Bueno, Elaine I., **Utilização de Redes Neurais Artificiais na Monitoração e Detecção de Falhas em sensores do Reator IEA-R1**. Dissertação apresentada como parte dos requisitos para obtenção do Grau de Mestre em Ciências na Área de Tecnológica Nuclear-Reatores, 2006.
- [9] Caminhas, W. M.; Tavares, H. M. F.; E Gomide, F., **"Rede Lógica Neurofuzzy: Aplicação Em Diagnóstico De Falhas Em Sistemas Dinâmicos"**, Anais do XI Congresso Brasileiro de Automática, São Paulo, Brasil, P.459-464, 1996.
- [10] Cirrea, C.M.; Loreto, V.; Berenguelb, M.; Camachod, E. F., **Feedback linearization control for a distributed solar collector field** , Control Engineering Practice, v 15 , pp 1533-1544, 2007.
- [11] Feldman, J. A. **Neural networks, artificial intelligence and computational reality** Computers in Industry, v 14, Issues 1-3, pp 145-148, 1990.
- [12] Finnof, W.; Hergert, R.; Zimmermann, G. **Improving Model Selection by Nonconvergent Methods** Neural Networks, v 6, pp 771-783, 1996.
- [13] Funahashi, K. I. **On the Approximate Realization of Continuous Mappings by Neural Networks** Neural Networks, v. 2, pp. 183-192, 1989.

- [14] Gabbar, H. A. **Integrated framework for safety control design of nuclear power plants** Nuclear Engineering and Design, v 240, pp 3550-3558, 2010.
- [15] Glasstone, S. **Principles of Nuclear Reactor Engineering**, Publisher: D. Van Nostrand, Princeton, 5^a ed. 1955.
- [16] Glasstone, S.; Sesonske, A. **Nuclear reactor engineering**, Van Nostrand Reinhold Company, 1981 4^a Ed. Vol. II.
- [17] Gomm, J.B. ; Evans, J.T. ; Williams, D. **Deveolpmen and Performance of a Neural-Network Predictive Controller**, Control Eng. Practice, Vol. 5, No. 1, p. 49-59, 1997.
- [18] GONÇALVES, I. M. P. **Monitoração e diagnóstico para detecção de falhas de sensores utilizando a metodologia GMDH**. Tese (Doutorado), IPEN, São Paulo, 2005
- [19] Gou, Z.; Fyfe, C. **A canonical correlation neural network for multicollinearity and functional data** Neural Network, v 17, pp 285-293, 2004.
- [20] Haykin, S. **Neural Networks: A Comprehensive Foundation**, 2^a ed. Prentice Hall, 1999.
- [21] Hornik, K. **Multilayer Feedforward Networks are Universal Approximators** Neural Networks, Vol. 2, pp. 359-366, 1986.
- [22] Hoteling, H. **The most predictable criterion** Journal of Educational Psychology, v 26, pp 139-142, 1935.
- [23] Hotelling, H. **Analysis of a complex of statistical variables into principal components** Journal of Educational Psychology, v 24 pp 417-441, 1933.
- [24] Hsieh, W. W. **Nonlinear canonical correlation analysis by neural networks** Neural Network, v 13, pp 1095-1105, 2000.
- [25] Hwang , S.L.; Liang, S. F. M.; Liu, T. Y. Y.; Yang, Y. J. ; **Evaluation of human factors in interface design in main control rooms** Nuclear Engineering and Design v 239,p 3069-3075, 2009.
- [26] INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES (IPEN). **Dados Operacionais do Reator IEA-R1 Arranjo 214 e 215 Mês Abril Ano 2004**. Abril 2004.
- [27] INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES (IPEN). **Relatório de Análise de Segurança do reator de pesquisas do Ipen IEA-R1 (RAS)**, 2005.
- [28] Kaufmann, A. **“Introduction to the Theory of Fuzzy Subsets”**. v.1, 1973.
- [29] Kern, G.; Przycłuski, M., **On perturbations of linear controllable infinite dimensional time-varying discrete-time systems**, Systems & Control Letters, v 15, pp 61-66, 1990.

- [30] Kinnaert, M. **Fault diagnosis based on analytical models for linear and nonlinear systems. A tutorial**, In Preprints of the fifth IFAC symposium on fault detection, supervision and safety for technical processes, SAFEPROCESS 2003 pg 37 a 50, 2003.
- [31] Kjaerheim, G., **Heat transfer as limiting factor in water-cooled nuclear reactors**, Nuclear Engineering and Design, v. 6, p 301-324, 1972.
- [32] Kosko, B. Neural Networks and Fuzzy Systems. Prentice Hall, Englewood Cliffs – NJ, 1992.
- [33] Leshno, M.; Ya, V. I.; Pinkus A.; Schocken, S., **Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function** Neural Networks, Vol. 6, pp. 861-867, 1993.
- [34] Lindeman, Ricchard **Introduction to Bivariate and Multivariate Analysis**, 1985
- [35] Linsker, R. **Neural network learning of optimal Kalman prediction and control** Neural Networks, v 21, pp 1328-1343, 2008.
- [36] Lugger, G. F. **Artificial Intelligence: structure and strategies for complex problem solution**. 4^a ed., 2002.
- [37] Marcu, T.; Selinger, B. K.; Stücher, R. **Design of fault detection for a hydraulic looper using dynamic neural networks** Control Engineering Practice, v 16, pp 192-213, 2008.
- [38] McCulloch, W. S.; Pitts, W. **A logical calculus of the ideas immanent in nervous activity**, Bulletin of Mathematical Biophysics, v. 5, p. 115 – 133, 1943.
- [39] Mendonça, L.F.; Sousa, J.M.C.; Costa, J.M.G. **“Na architecture for fault detection and isolation based on fuzzy methods”** Expert Systems with Applications vol. 36, p 1092 a 1104, 2009
- [40] Minin, A.; Velikova, M.; Lang, B.; Daniels, H. **Comparision of universal approximators incorporating partial monotonicity by structure** Neural Networks, v 23, pp 471-475, 2010.
- [41] Monteiro, P. R. B.; Gonçalves, I. M. P. M., **Sistema De Aquisição De Dados Do Reator IEA-R1**, Inac 2002 – International Nuclear Atlantic Conference Proceedings (Cdroom), Outubro 2002.
- [42] Moore, D. S.; MacCabe, G. P., **Introduction to the Practice Statistics”** 3rd edition. New York W. H. Freeman, 1998
- [43] Murray, R. L. **“Energia Nuclear”**, São Paulo, Hemus,2004.
- [44] NUTEC. Diagrama Esquemático de uma Central Nuclear PWR. Disponível em: <http://www.nuctec.com.br:80/educacional/usina.htm>. Acesso em 19 Julho de 2009.

- [45] Ogata, K. **Modern Control Engineering**, 4^a ed. Prentice Hall, Inc,2002.
- [46] Oja, H. E. **Independent component analysis: algorithms and applications**, Neural Networks, v 13, pp 411-430, 2000.
- [47] Papadokonstantakis, S.; Lygeros, A.; Jacobsson, S. P. **Comparison of recent methods for inference of variable influence in neural networks** Neural Networks, v 19 , pp 500-513, 2006.
- [48] Patton, R. J.; Chen, J.; Nielsen, S. B., **Model-Based Methods For Fault Diagnosis: Some Guidelines**, IEEE Transactions Of The Institute Of Measurement And Control, Vol. 17, No. 2, 1995.
- [49] Qina, S. J.; Badgwellb, T. A. **A survey of industrial model predictive control technology** Control Engineering Practice, v 11 pp 733-764, 2003.
- [50] Retherford, R. D. **Statistical Models for Causal Analysis**
- [51] Revista Galileu – Os segredos de uma usina Edição 161 - Dez/04).
- [52] Rodrigues, M.; Theilliol, D.; Adam-Medina, M.; Sauter, D., **A fault detection and isolation scheme for industrial systems based on multiple operating models** Control Engineering Practice, v 16, pp 225-239, 2008.
- [53] Santamaría, I.; Vía, J.; Pérez, J. **A learning algorithm for adaptive canonical correlation analysis of several data** Neural Networks, v 20, pp 139-152, 2007.
- [54] Savva , P.; Varvayanni, M.; Catsaros, N., **Dependence of control rod worth on fuel burnup** Nuclear Engineering and Design, v 241, pp 492-497, 2011.
- [55] Scaglia, G.; Rosales, A.; Quintero, L.; Mut, V.; Agarwal, R. **A linear-interpolation-based controller design for trajectory tracking of mobile robots** Control Engineering Practice, v 18, pp 318-329, 2010.
- [56] SCHAFFER, M. B. **Power the world with nuclear energy - past, present, and inevitable future**. Foresight, v. 7, n. 2, pp. 41-53, 2005.
- [57] Schiffmann, W.; Geffers W. H. **Adaptive Control of Dynamic Systems by Back Propagation Networks** Neural Networks, v 6, pp 517-524, 1993.
- [58] Sjöberg, J.; Gutman, P.; Agarwal, M.; Bax, M. **Nonlinear controller tuning based on a sequence of identifications of linearized time-varying models**, Control Engineering Practice, v 17, pp 311-321, 2009.
- [59] Stinchcombe, M.B **Neural network approximation of continuous functionals and continuous functions on compactifications** Neural Networks, v. 12 ,pp 467–477, 1999.

- [60] Tachibana, Y.; Sawahata, H.; Yyoku, T.; Nakazawa, T., **Reactivity control system of the high temperature engineering test reactor** Nuclear Engineering and Design, v 233, pp 89-101, 2004.
- [61] Tan, W. **Water level control for a nuclear steam generator** School Nuclear Engineering and Design, v 241, pp 1873-1880, 2011.
- [62] Tanomaru, N.; Hiromoto, M. Y. C., Manual De Instalação E Operação Do SAD IEA-R1, 1998.
- [63] Thompson, A. S.;Rodgers, O. E., **Thermal Power from Nuclear Reactors**, 1956.
- [64] Tong, L. S., **Heat transfer in water-cooled nuclear reactors** Nuclear Engineering and Design, v 6, p. 301-324,1967.
- [65]Y, T., Su-xia; H., Chong, L.; Fu-yu, Z. **An improved implicit multiple model predictive control used for movable nuclear power plant**, Nuclear Engineering and Design v 240 ,p 3582–3585, 2010.