



Manoela Rabello Kohler

**Redes Neurais Aplicadas à Construção de Aproximadores
para Simulação Integrada entre Reservatório e Sistema de
Produção**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para
obtenção do título de Mestre pelo Programa de Pós-
Graduação em Engenharia Elétrica da PUC-Rio.

Orientadora: Profa. Marley Maria Bernardes Rebuszi Vellasco
Co-Orientador: Prof. Eugênio da Silva

Rio de Janeiro
Setembro de 2013



Manoela Rabello Kohler

**Redes Neurais Aplicadas à Construção de Aproximadores
para Simulação Integrada entre Reservatório e Sistema de
Produção**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Profa. Marley Maria Bernardes Rebuszi Vellasco
Orientador
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Eugênio da Silva
Co-orientador
UEZO

Prof. Marco Aurélio Cavalcanti Pacheco
Departamento de Engenharia Elétrica - PUC-Rio

Prof. Alexandre Anozé Emerick
Petróleo Brasileiro – Rio de Janeiro - Matriz

Prof. Leandro Costa Reis
Petrobras

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 18 de Setembro de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

Manoela Rabello Kohler

Graduou-se em Engenharia de Computação pela Pontifícia Universidade Católica Do Rio de Janeiro (PUC-Rio) em 2011.

Ficha Catalográfica

Kohler, Manoela Rabello

Redes Neurais Aplicadas à Construção de Aproximadores para Simulação Integrada entre Reservatório e Sistema de Produção / Manoela Rabello Kohler; orientadora: Marley Maria Bernardes Rebuszi Vellasco; co-orientador: Eugênio da Silva – 2013.

96 f.: il. (color.) ; 30 cm

Dissertação (Mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2013.

Inclui referências bibliográficas.

1. Engenharia elétrica – Teses. 2. Aproximador. 3. Proxy. 4. Inteligência Computacional. 5. Redes Neurais. 6. Simulação de Reservatório. 7. Simulação de Sistema de Produção. 8. Simulação Integrada. 9. Hipercubo Latino. 10. Análise de Componentes Principais. I. Vellasco, Marley Maria Bernardes Rebuszi. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Agradecimentos

À Deus.

À minha orientadora Professora Marley, pelo apoio e parceria para a realização deste trabalho.

Ao meu co-orientador Eugênio, pela confiança, estímulo e importantes contribuições.

Aos meus pais, Zita e Helio, pela educação, atenção e carinho.

Aos meus irmãos amigos e amigos irmãos Marcella, Dedé, Roberta, JP, Davi, Leonardo, Mariana, Marco, Sarah, Marina, Mateus, por todo apoio, paciência e compreensão.

À todos os amigos da PUC-Rio.

À todos os amigos e familiares que de uma forma ou de outra me estimularam e me ajudaram.

Aos colegas da Petrobrás pela disponibilidade e interesse em contribuir.

Aos professores que participaram da Comissão examinadora.

Ao CNPq, à PUC-Rio, à Petrobrás e à Petroleum Experts, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Resumo

Kohler, Manoela Rabello; Vellasco, Marley Maria Bernardes Rebuzzi; Silva, Eugênio da. **Redes Neurais Aplicadas à Construção de Aproximadores para Simulação Integrada entre Reservatório e Sistema de Produção**. Rio de Janeiro, 2013. 96p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

O desenvolvimento de um reservatório de petróleo já conhecido e delimitado consiste em encontrar uma alternativa (configuração) de poços que contribua para maximizar a receita a ser obtida com o óleo recuperado do reservatório. A busca por esta alternativa frequentemente é baseada em processos de otimização que usam o valor presente líquido (VPL) do projeto como função de avaliação das alternativas encontradas durante a busca. Dentre outras variáveis, o cálculo do VPL é diretamente dependente dos dados de produção de óleo, gás e água durante a vida produtiva do reservatório, bem como de seus custos de desenvolvimento. Determinar a localização, os tipos (produtor ou injetor) e a trajetória de poços em um reservatório é um problema de otimização complexo que depende de uma grande quantidade de variáveis, dentre elas as propriedades do reservatório (tais como porosidade e permeabilidade) e os critérios econômicos. Os processos de otimização aplicados a este tipo de problema têm um alto custo computacional devido ao uso contínuo de simuladores que reproduzem as condições do reservatório e do sistema de superfície. O uso dos simuladores pode ser substituído por um aproximador, que neste trabalho, é um modelo que utiliza Redes Neurais Artificiais. Os aproximadores aqui apresentados são feitos para substituir a simulação integrada do reservatório, do poço e da superfície (linhas de produção e *riser*). As amostras para a construção do aproximador é feita utilizando os simuladores de reservatório e de superfície e para reduzir o número de amostras necessárias e tornar sua construção mais rápida, utiliza-se Hipercubo Latino e Análise de Componentes Principais. Os aproximadores foram testados em dois reservatórios petrolíferos: um reservatório sintético, e baseado em um caso real. Os resultados encontrados indicam que estes aproximadores conseguem bom desempenho na substituição dos simuladores no processo de otimização devido aos baixos erros encontrados e à substancial

diminuição do custo computacional.

Palavras-chave

Aproximador; Hiper cubo Latino; Análise de Componentes Principais;
Redes Neurais; Simulador de Reservatórios; Simulador do Sistema de Produção.

Abstract

Kohler, Manoela Rabello; Vellasco, Marley Maria Bernardes Rebuzzi (Advisor); Silva, Eugênio da (Co-Advisor). **Neural Networks Applied to Proxies for Reservoir and Surface Integrated Simulation**. Rio de Janeiro, 2013. 96p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The development of an oil reservoir consists in finding an alternative of wells that contributes to maximizing the revenue to be obtained from the recovered reservoir oil. The pursuit for this alternative is often based on optimization processes using the net present value (NPV) of the project as the evaluation function of the alternatives found during this pursuit. Among other variables, the NPV calculation is directly dependent on the oil, gas and water production data during the productive life of the reservoir, as well as their development costs. Determine the number, location, type (producer or injector) and the trajectory of wells in a reservoir is a complex optimization problem which depends on a lot of variables, including the reservoir properties (such as porosity and permeability) and economic criteria. The optimization processes applied to this type of problem has a high computational cost due to the continuous use of simulators that reproduce the conditions of the reservoir and the surface system. The use of simulators may be replaced by proxies. At the present work, proxies were constructed using artificial neural networks. The proxies presented here are meant to replace the integrated reservoir, well and surface (production lines and riser) simulation to reduce the computational cost of a decision support system. The samples for the construction of the proxies are produced using reservoir and surface simulators. To reduce the number of samples needed for the proxy construction, and, to reduce the dimension of the problem, Latin Hypercube and Principal Component Analysis are used. The approximators were tested in two oil reservoirs: a synthetic reservoir, and another with real features. The results indicate that these approximators can perform well in replacement of simulators in the optimization process due to low errors found and a substantial decrease in computational cost.

Keywords

Proxy; Latin Hypercube; Principal Component Analysis; Artificial Neural Networks; Reservoir Simulation; Surface Simulation; Integrated Simulation.

Sumário

1 . Introdução	15
1.1. Motivações	15
1.2. Objetivos	17
1.3. Descrição do Trabalho	17
1.4. Organização do Trabalho	19
2 . Fundamentação Teórica	21
2.1. Introdução	21
2.2. Hipercubo Latino	21
2.2.1. Metodologia	22
2.3. Análise de Componentes Principais	24
2.3.1. Desenvolvimento Matemático	25
2.4. Redes Neurais Artificiais	27
2.4.1. Arquitetura	28
2.4.2. Algoritmos de Aprendizado	29
2.5. Produção de Petróleo	35
2.6. Sistemas Marítimos de Produção	37
2.6.1. Linha de Produção e <i>Riser</i>	38
2.6.2. Manifolds	38
2.7. Fluxo Multifásico	39
2.7.1. Tabelas de Fluxo Vertical Multifásico	40
2.8. Simuladores	42
2.8.1. Simuladores de Reservatório	43
2.8.2. Simuladores de Superfície	44
3 . Modelagem do Aproximador	45
3.1. Introdução	45
3.2. Integração do Simulador de Reservatório ao Simulador de Superfície	46
3.2.1. Métodos de Acoplamento entre Simuladores	46

3.2.2. Acoplamento IMEX – GAP e PROSPER	47
3.2.3. Conexão entre as Tabelas de Fluxo Vertical Multifásico	52
3.2.4. Dados de Modelagem utilizando o Software PROSPER	52
3.3. Construção do Aproximador	53
3.3.1. Aquisição dos Dados de Treinamento	54
3.3.2. Seleção de Variáveis de Treinamento e Criação da Base	54
3.3.3. Geração dos Cenários de Poços	58
3.3.4. Transformação dos Dados de Treinamento	60
3.3.5. Redução de Dimensionalidade utilizando Análise de Componentes Principais	60
4 . Estudo de Casos	62
4.1. Parâmetros da Rede Neural	62
4.2. Métricas de Avaliação	62
4.3. Descrição dos Reservatórios	64
4.3.1. Reservatório Sintético	64
4.3.2. Reservatório Real	64
4.4. Avaliação dos Aproximadores	65
4.4.1. Testes 1 e 2	66
4.4.2. Testes 3 e 4	68
4.4.3. Testes 5 e 6	69
4.4.4. Testes 7 e 8	71
4.4.5. Testes 9 e 10	73
4.4.6. Teste 11	74
4.4.7. Testes 12 e 13	76
4.5. Avaliação dos Aproximadores no Sistema de Otimização de Localização de Poços	80
4.5.1. Comparação entre a Utilização dos Simuladores e dos Aproximadores em Processos de Otimização	81
5 . Conclusões e Trabalhos Futuros	86
5.1. Conclusões	86
5.2. Trabalhos Futuros	87

Referências Bibliográficas 89

Apêndice A – Criação de *Manifolds* Imaginários no Simulador de Reservatório IMEX 94

Lista de Figuras

Figura 1: Intervalos usados pela amostra hipercubo latina para variáveis de distribuição normal e uniforme.	23
Figura 2: Amostra hipercubo latina de duas dimensões.	24
Figura 3: Representação geral da arquitetura <i>Multilayer Perceptron</i> (<i>Feedforward</i>).....	29
Figura 4: Identificação do ponto de interrupção do treinamento com o método da parada antecipada.....	32
Figura 5: <i>Flowline</i> e <i>Riser</i>	38
Figura 6: <i>Manifold</i> (Cerqueira, 2005).....	39
Figura 7: Curva de Fluxo Vertical Multifásico	41
Figura 8: Ponto de operação do sistema de produção. (Ribeiro, 2012).....	48
Figura 9: Simulação Integrada.....	48
Figura 10: Diagrama do poço (IMEX)	50
Figura 11: Diagrama de Superfície (PROSPER).....	50
Figura 12: Uso das Tabelas e Fluxo Vertical Multifásico no Reservatório Real ..	51
Figura 13: Entradas da rede neural	55
Figura 14: Janela deslizante para previsão de produção de óleo por 5 anos	56
Figura 15: Esquema da distribuição das alternativas pelas Redes Neurais	57
Figura 16: Entrada das redes neurais para novo modelo	57
Figura 17: Reservatório sintético.....	64
Figura 18: Reservatório real	65
Figura 19: Curva de otimização com uso dos simuladores (20 gerações).....	84
Figura 20: Curva de otimização com uso dos aproximadores (20 gerações)	84
Figura 21: Curva de otimização com uso dos simuladores (30 gerações).....	85
Figura 22: Curva de otimização com uso dos aproximadores (30 gerações)	85

Lista de Tabelas

Tabela 1: Parâmetros das Redes Neurais Artificiais	62
Tabela 2: Estrutura de apresentação dos resultados para os aproximadores	66
Tabela 3: Configuração dos microcomputadores utilizados.....	66
Tabela 4: Testes 1 e 2 para o reservatório sintético.....	67
Tabela 5: Testes 3 e 4 para o reservatório sintético.....	69
Tabela 6: Melhores resultados dos testes 1, 2, 3 e 4.....	69
Tabela 7: Testes 5 e 6 para o reservatório sintético.....	70
Tabela 8: Melhores resultados dos testes 1, 2, 5 e 6.....	71
Tabela 9: Testes 7 e 8 para reservatório sintético.....	72
Tabela 10: Melhores resultados dos testes 5, 6, 7 e 8.....	72
Tabela 11: Melhores resultados dos testes 9 e 10 para previsão de óleo.....	74
Tabela 12: Melhores resultados do teste 10 para previsão de gás	74
Tabela 13: Melhores resultados do teste 10 para previsão de água.....	74
Tabela 14: Melhores resultados do teste 11 para previsão de óleo	75
Tabela 15: Melhores resultados do teste 11 para previsão de gás	75
Tabela 16: Melhores resultados do teste 11 para previsão de água.....	75
Tabela 17: Melhores resultados do teste 12 para previsão de óleo	76
Tabela 18: Melhores resultados do teste 12 para previsão de gás	76
Tabela 19: Melhores resultados do teste 12 para previsão de água.....	77
Tabela 20: Melhores resultados do teste 13 para previsão de óleo	77
Tabela 21: Melhores resultados do teste 13 para previsão de gás	77
Tabela 22: Melhores resultados do teste 13 para previsão de água.....	77
Tabela 23: Análise de variância das variáveis.....	78
Tabela 24: Análise da variância acumulada dos n primeiros componentes principais	79
Tabela 25: Tabela do tempo de construção dos aproximadores	79
Tabela 26: Comparação entre tempo de construção de um aproximador e uma otimização	80
Tabela 27: Parâmetros para cálculo do VPL	82
Tabela 28: Parâmetros do Algoritmo Genético	82

Tabela 29: Resultados das otimizações	83
Tabela 30: Média e desvio padrão dos dados obtidos com os aproximadores	83

1. Introdução

1.1. Motivações

Um simulador de reservatório é um programa que utiliza equações matemáticas a fim de descrever a dinâmica dos fluidos no interior de um reservatório em diferentes condições. Sendo assim, um simulador utiliza um conjunto de parâmetros como entrada e retorna informações de fluxo de fluido de saída (Túpac, 2005).

Para se resolver esse conjunto de equações, o simulador utiliza métodos de diferenças finitas, que é um método de resolução de equações diferenciais que se baseia na aproximação de derivadas por diferenças finitas (Mattax, 1990). A modelagem de um sistema de reservatório envolve inúmeras variáveis e inúmeras equações para descrever todo o escoamento do fluido dentro do reservatório. Por isso, o custo computacional envolvido na utilização de um simulador se torna extremamente alto, o que pode tornar certos processos inviáveis (Faletti, 2003).

Um simulador de superfície acoplado ao simulador de reservatório é fundamental para que o sistema de produção seja mais bem descrito como um todo. A integração da simulação das instalações de superfície permite que seja modelado, com mais precisão o sistema completo de produção de petróleo, já que se torna possível modelar a perda de carga no poço, nas linhas de produção e no *riser* (tubulação que faz a ligação da linha de produção à plataforma).

Alguns trabalhos vêm sendo desenvolvidos na área para fazer a simulação integrada de reservatório/sistema de produção utilizando diversos softwares de simulação de superfície e reservatório, como em Hayder (2006), Ursini (2010), Rodríguez (2007), Ageh (2010) e Hayder (2011).

Para reduzir o tempo necessário para realizar as simulações integradas, e até mesmo para que o número de simulações necessárias de um sistema não seja um fator limitante para a realização de otimizações e análises de incerteza, diversas

técnicas inteligentes podem ser usadas, como Redes Neurais Artificiais, sistemas Neuro-Fuzzy (Jang, 1993) e Programação Genética (Koza, 1992).

No presente trabalho é utilizada a técnica de Redes Neurais pelo fato delas terem se apresentado como uma importante ferramenta na solução desse tipo de problema. As Redes Neurais se destacam pela habilidade de aprendizado e capacidade de generalização, tendo mostrado serem capazes de identificar e assimilar as características mais marcantes das séries. As Redes Neurais têm a capacidade de resolver problemas de aproximação de funções contínuas (Haykin, 2001). Por serem aproximadores universais de funções (Hornik *et al.*, 1989), podem, a partir de dados numéricos obtidos através de simulações, mapear a superfície de resposta dos simuladores integrados. Com isso, é possível substituí-los em processos que exijam grandes quantidades de simulações como, por exemplo, processos de otimização de localização de poços.

Aproximadores podem ser utilizados para substituir a simulação integrada de reservatório e sistema de produção parcial ou totalmente para um determinado reservatório, tornando possível utilizar, de forma mais adequada e com custo computacional menor, processos que requerem grande número de simulações. Muitos trabalhos já foram desenvolvidos na área. Por exemplo, em Túpac (2005) propõe-se um aproximador para o simulador de reservatório utilizado no sistema de localização de poços que determina a alternativa de produção que maximiza o Valor Presente Líquido (VPL). Em Almeida (2008) e Almeida (2010) foram desenvolvidos aproximadores de reservatório mediante o uso de Rede Neurais e de modelos híbridos Neuro-Fuzzy considerando a produção de um reservatório com um número fixo de poços inteligentes. Em Talavera (2010) foram desenvolvidos aproximadores para o controle preditivo da produção de óleo em poços inteligentes. Em Yu (2007) é proposto um aproximador utilizando Programação Genética para ajuste de histórico e análises de incertezas em previsões de produção. E como último exemplo, em Villanueva (2010) foi construído um aproximador utilizando Redes Neurais para a substituição do simulador para um reservatório sintético específico onde os tipos e as localizações dos poços são variadas. Entretanto, em nenhum desses trabalhos foi construído um aproximador com a finalidade de aproximar o comportamento de uma simulação integrada entre o reservatório e o sistema de superfície, podendo-se variar a localização, a geometria e o tipo de cada poço.

1.2. Objetivos

Os objetivos principais deste trabalho são a modelagem, a implementação e a avaliação de um aproximador de função que seja capaz de substituir, eficientemente, o sistema de simulação integrada do reservatório (simulador IMEX (CMG, 2012)) e do sistema de superfície (simulador PROSPER (PETEX, 2012)), considerando um reservatório com um número fixo de poços. O modelo proposto deve considerar ainda a possibilidade de variações na localização, no tipo e na geometria de cada poço, e também deve respeitar um conjunto de restrições operacionais.

Os objetivos secundários são:

- Modelagem do sistema de superfície a partir de um modelo real de reservatório, para que a curva de produção leve em conta o escoamento do fluido e a perda de carga do mesmo nos poços, nas linhas de produção e nos *risers*.
- Geração de amostras com diversas localizações de poços e *layouts* submarinos diferentes, para serem usadas no treinamento do aproximador.
- Minimização do custo computacional, selecionando, de forma mais inteligente, as amostras a serem usadas para a construção do aproximador. Para isso, é utilizada a técnica de amostragem Hipercubo Latino.
- Redução da dimensão de entrada das Redes Neurais por meio da técnica de Análise de Componentes Principais.

1.3. Descrição do Trabalho

Este trabalho foi elaborado seguindo as seguintes etapas:

- Estudo sobre a área de exploração de petróleo, escoamento de fluidos, simulação integrada de reservatório e sistemas de superfície, e seus respectivos simuladores.

- Estudo sobre o modelo de Inteligência Computacional utilizado: Redes Neurais Artificiais.
- Estudo sobre amostragem por Hipercubo Latino.
- Estudo sobre Análise de Componentes Principais para a redução de dimensionalidade.
- Geração das tabelas de fluxo vertical multifásico para representar a perda de carga nos poços, nas linhas de produção e nos *risers*.
- Modelagem do problema e implementação do modelo utilizando Redes Neurais.
- Estudo de casos.

Na primeira etapa do trabalho, foi feito um levantamento bibliográfico referente à área de desenvolvimento de reservatórios de petróleo, escoamento de fluidos a partir do reservatório até a superfície, sobre o simulador de reservatório IMEX e o simulador de superfície PROSPER, assim como o estudo sobre as possibilidades de acoplamento entre esses simuladores.

Foi feito também um estudo sobre Redes Neurais Artificiais, que é a técnica de Inteligência Computacional utilizada na construção da solução proposta. Também foi feito um levantamento dos principais trabalhos existentes na área de Inteligência Computacional para desenvolvimento de aproximadores.

A partir do estudos desses trabalhos foram definidas características importantes a serem levadas em conta no desenvolvimento de aproximadores e, mais especificamente, no desenvolvimento de aproximadores de reservatório com sistema de produção integrado. Diferentes técnicas foram avaliadas e comparadas a fim de identificar a mais adequada para a solução do problema em questão.

Na etapa anterior à modelagem, foi feito um estudo sobre técnicas de amostragem e foi escolhida a técnica para obtenção das amostras: Hipercubo Latino.

Nessa etapa também se estudou a técnica de redução de dimensionalidade chamada Análise de Componentes Principais para ser usada, principalmente, nos casos em que há um número muito elevado de variáveis para a construção do aproximador.

Na etapa de modelagem do aproximador foi feita, primeiramente, uma análise do problema para identificar as variáveis relevantes. Em seguida, foi feita a aquisição de amostras, utilizando as técnicas de amostragem já mencionadas.

Na etapa anterior à implementação do modelo, foi construído o modelo de superfície utilizando o simulador PROSPER. Nesse simulador, todo o sistema de superfície foi modelado e inúmeras tabelas de fluxo vertical multifásico foram geradas para modelar a perda de carga.

Com os simuladores de reservatório e superfície acoplados e as amostras selecionadas, foram feitas as simulações para obtenção dos dados de saída da simulação: taxa de óleo, água e gás produzidas pelo reservatório durante um período determinado.

Na etapa de estudo de casos foram realizados diversos testes. Os testes foram feitos em dois modelos de reservatórios, sendo um sintético e um baseado em um reservatório real. Em ambos os modelos de reservatórios, as configurações de poços foram variadas de duas formas: a primeira utilizando somente poços verticais e a segunda utilizando livremente poços verticais e horizontais.

1.4. Organização do Trabalho

Esta dissertação contém mais quatro capítulos, cujos conteúdos estão descritos a seguir.

No Capítulo 2 faz-se uma introdução sobre todas as técnicas utilizadas para a conclusão do objetivo principal. Tais técnicas incluem Hipercubo Latino, Análise de Componentes Principais e Redes Neurais Artificiais.

O Capítulo 3 apresenta a modelagem da solução baseada em Redes Neurais. Nele é também tratada a metodologia de acoplamento entre os simuladores de reservatório e superfície, assim como a amostragem por Hipercubo Latino que contribuíram para a criação do aproximador, e a Análise de Componentes Principais que permitiu a redução da dimensionalidade do problema.

O Capítulo 4 descreve o estudo de casos realizados em dois modelos de reservatórios de petróleo. Um modelo é baseado no campo de Namorado e, por isso, possui grau de complexidade e características bem próximas a reservatórios reais. O outro modelo utilizado é sintético e mais simples que um modelo de um

reservatório real. Os resultados obtidos em ambos os casos são apresentados e discutidos.

Finalmente, no Capítulo 5, são apresentadas as conclusões e os possíveis direcionamentos para trabalhos futuros.

2. Fundamentação Teórica

2.1. Introdução

Neste capítulo são apresentados os conceitos fundamentais sobre Hipercubo Latino, Análise de Componentes Principais e sobre Redes Neurais do tipo Multilayer Perceptron (MLP), que é a técnica de Inteligência Computacional empregada na construção do aproximador de função.

2.2. Hipercubo Latino

Os modelos computacionais podem ser entendidos como uma série de comandos, ou procedimentos, que simulam um problema real e fornecem uma possível solução para este problema. No entanto, os problemas reais são muito complexos e difíceis de serem simulados, devido ao seu porte e à sua aleatoriedade ou imprevisibilidade.

Nos problemas de difícil previsibilidade, os modelos computacionais adotam técnicas para representar a aleatoriedade, tal como, a geração de valores aleatórios como dados de entrada. Na maioria das vezes esses valores são oriundos de um sorteio aleatório simples de um determinado espaço amostral. No entanto, existem técnicas que representam melhor este espaço amostral, através de seleções baseadas em procedimentos que percorrem todo o espaço amostral e selecionam valores de maneira que este espaço fique bem representado. A técnica do Hipercubo Latino é uma técnica de amostragem estratificada na qual os estratos selecionados da distribuição são representados na amostra através de um representante.

2.2.1. Metodologia

O método do Hipercubo Latino, utilizado para geração de amostras de distribuições multivariadas, foi desenvolvido por Mckay, Conover e Backman em 1979 (Mckay, 1979). Este método pode ser considerado um caso particular de amostragem estratificada, onde cada estrato possui a mesma probabilidade.

Esta metodologia consiste em selecionar n diferentes valores para cada variável $k(X_1, X_2, \dots, X_k)$ através do seguinte procedimento:

1 – Dividir o intervalo de cada variável k em N intervalos equiprováveis e disjuntos;

2 – Selecionar um valor, ou um representante, para cada intervalo de acordo com sua densidade de probabilidade;

3 – Após obtidos os N representantes de todas as variáveis, parear, ou seja, formar pares de maneira aleatória, dos N representantes de X_1 com os n representantes de X_2 ;

4 – Os N pares formados pela combinação de X_1 com X_2 são então combinados com os N representantes de X_3 , formando assim trios;

5 – Os N trios são então combinados com os N representantes de X_4 e assim sucessivamente, até que todas as variáveis façam parte da combinação.

Independentemente da distribuição que a variável aleatória possua, a divisão de seu intervalo será sempre equiprovável. Além disso, as variáveis de um problema podem ter distribuições diferentes.

O exemplo a seguir refere-se à geração de uma amostra do hipercubo latino de tamanho $N=5$ com duas variáveis de entrada $X = (X_1, X_2)$, sendo que a variável X_1 possui uma distribuição normal com média μ e variância σ^2 e a variável X_2 possui distribuição uniforme.

Ao dividir ambas as variáveis independentes em cinco intervalos equiprováveis tem-se os seguintes resultados segundo a Equação (2.1) e a Equação (2.1):

$$\begin{aligned} P(-\infty < X_1 < A) &= P(A < X_1 < B) = P(B < X_1 < C) = P(C < X_1 < K) = \\ &= P(K < X_1 < \infty) = 0,2 \end{aligned} \quad (2.1)$$

$$\begin{aligned}
 P(G < X_2 < H) &= P(H < X_2 < I) = P(I < X_2 < J) = P(J < X_2 < K) = \\
 &= P(K < X_2 < L) = 0,2
 \end{aligned}
 \tag{2.1}$$

A Figura 1 ilustra os correspondentes intervalos usados na amostra do hipercubo latino para ambas as variáveis em função da sua distribuição de probabilidade e da função de distribuição acumulada.

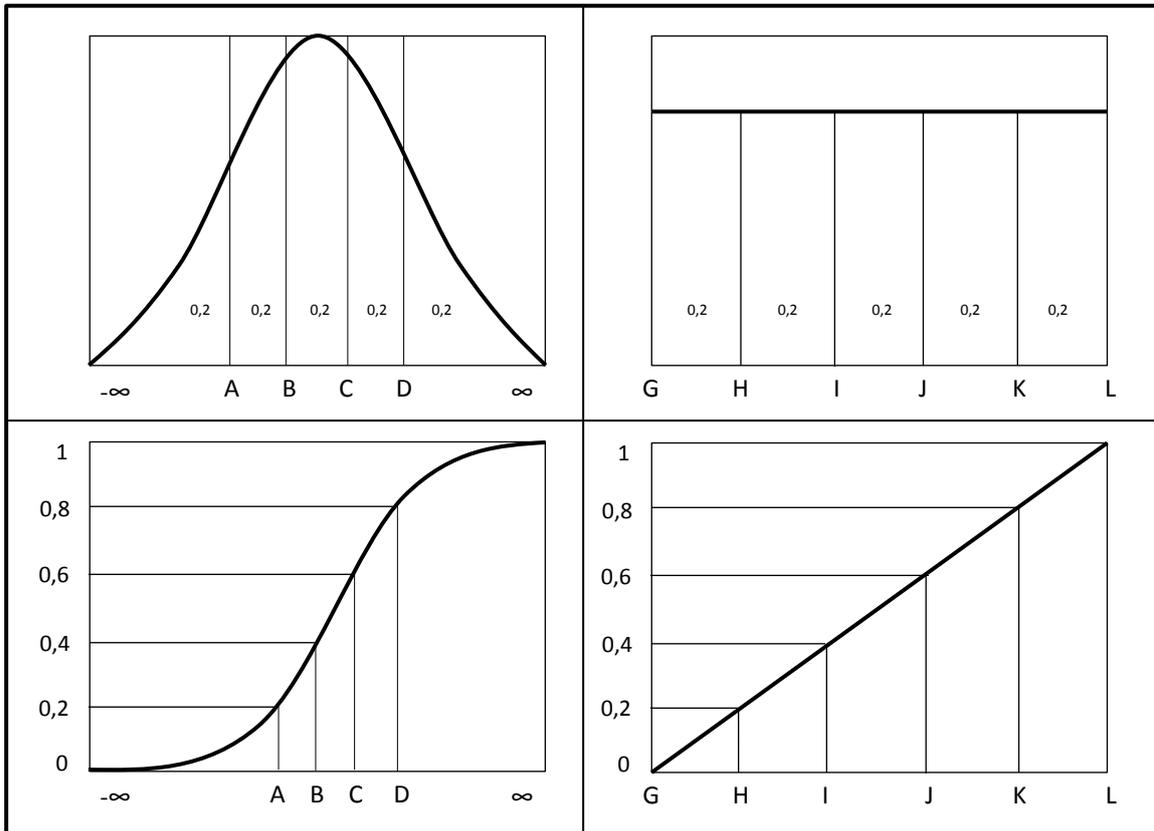


Figura 1: Intervalos usados pela amostra hipercubo latina para variáveis de distribuição normal e uniforme. (Wyss, 1998)

Sendo assim, a dimensão de cada componente de X é dividida em N partes, de iguais probabilidades, dando origem a N^k ($5^2 = 25$) intervalos no espaço amostral S de X , sendo cada um deles com uma probabilidade equivalente de N^{-k} ($5^{-2} = 1/25$).

Depois de feita a partição dos estratos, é realizada a seleção do representante de cada um deles. Esta seleção pode ser feita aleatoriamente ou segundo algum critério, como a escolha do ponto central, ou do ponto com maior probabilidade.

Uma vez feita a seleção dos representantes de cada um dos estratos é realizado então o pareamento entre os intervalos. Este pareamento é realizado

através da permutação dos números inteiros de intervalos de cada variável, ou seja, será feita uma permutação com os seguintes números (1,2,3,4,5) para a variável X_1 e outra permutação com os mesmos números para a variável X_2 , como por exemplo: permutação para $X_1 = (3,1,5,2,4)$ e permutação para $X_2 = (2,4,1,3,5)$. Essas permutações são combinadas, dando origem à amostra hipercubo latina conforme ilustrado na Figura 2.

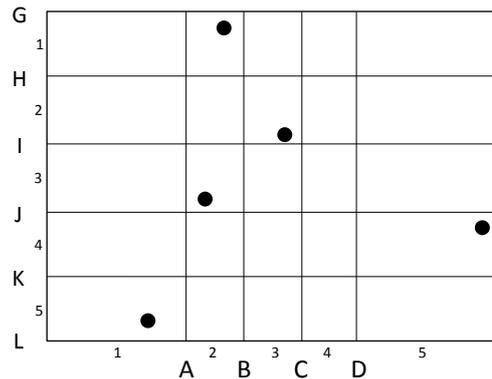


Figura 2: Amostra hipercubo latina de duas dimensões.

De acordo com a definição do cubo latino, numa amostra hipercúbica não pode haver mais de um representante na mesma linha e na mesma coluna.

2.3. Análise de Componentes Principais

A análise de componentes principais (ACP) é uma técnica da estatística multivariada que consiste em transformar um conjunto de variáveis originais em outro conjunto de variáveis de mesma dimensão denominadas de componentes principais. Os componentes principais apresentam propriedades importantes: cada componente principal é uma combinação linear de todas as variáveis originais, são independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo de informação, em termos da variação total contida nos dados. A análise de componentes principais é associada à ideia de redução de massa de dados, com menor perda possível da informação. Procura-se redistribuir a variação observada nos eixos originais de forma a se obter um conjunto de eixos ortogonais não correlacionados. Esta técnica pode ser utilizada para geração de

índices e agrupamento de indivíduos. A análise agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou seja, segundo seu comportamento dentro da população, representado pela variação do conjunto de características que define o indivíduo, ou seja, a técnica agrupa os indivíduos de uma população segundo a variação de suas características.

2.3.1. Densenvolvimento Matemático

A Análise de Componentes Principais consiste essencialmente em uma transformação de coordenadas de um conjunto de variáveis originais correlacionadas em um conjunto de variáveis sintéticas descorrelacionadas, os chamados componentes principais.

Cada componente é uma combinação linear das variáveis originais. Todos os componentes são ortogonais entre si, portanto não há informações redundantes. ACP tenta, simultaneamente, encontrar no espaço dimensional transformado a direção ao longo do qual os pontos se encontrem espalhados com variabilidade máxima e preservar a informação dos dados originais (Silva et al., 2005). De um ponto de vista prático, isto é feito através dos seguintes passos: obtenção de uma matriz que represente o conjunto de dados, centralização dos dados em torno da média, cálculo da matriz de covariância, cálculo dos autovalores e autovetores e diagonalização da matriz de covariância.

No primeiro passo, os dados são organizados em uma matriz $X_{n \times m}$, onde n representa o número de observações e m o número de variáveis independentes. O segundo passo, a centralização dos dados em torno da média, tem como objetivo prevenir que os pontos mais distantes do centro dos dados tenham maior influência que os mais próximos. Essa transformação é feita pela aplicação da Equação (2.3):

$$z_i = \frac{x_i - \mu}{\sigma} \quad (2.2)$$

Onde:

- z_i : valores transformados,
- x_i : amostras para $i = 1, 2, \dots, n$,
- μ : média amostral da variável aleatória x_i ,
- σ : desvio padrão amostral.

Desta forma, todas as variáveis aleatórias são distribuídas com média zero e desvio padrão unitário.

O terceiro passo consiste no cálculo da matriz de covariância C_z (Equação (2.4))

$$C_z = Z \cdot Z^T \quad (2.3)$$

Onde:

Z^T : transposta de Z .

Os elementos da diagonal principal desta matriz se referem às variâncias (λ) das colunas (variáveis independentes). Já os elementos fora da diagonal principal representam a covariância entre as variáveis. Quando a covariância é nula significa que as variáveis aleatórias são descorrelacionadas.

O método ACP permite a eliminação da covariância entre as coordenadas de um vetor de variáveis aleatórias por meio de uma mudança de base. A base formada pelos autovetores da matriz de covariância permite a diagonalização da matriz de covariância, ou seja, a eliminação da covariância entre as coordenadas do vetor de entrada. (Ludwig Jr. e Montgomery, 2007).

A matriz de covariância é real e simétrica, o que significa ser possível encontrar um conjunto de autovalores e correspondentes autovetores ortogonais.

O quarto passo consiste então na determinação dos autovalores λ e autovetores v_n correspondentes da matriz C_z . Os autovetores são arranjados de modo decrescentes de acordo com os valores dos autovalores.

Encontrados os autovetores v_n , estes formarão as colunas de uma matriz P (Equação (2.5)):

$$P = \{v_1, v_2, \dots, v_n\} \quad (2.4)$$

Em ACP, os autovalores da matriz de covariância correspondem à variância das variáveis transformadas. Assim, se um autovetor possui autovalor grande, significa que este fica em uma direção em que há uma grande variância dos padrões. A importância desse fato está na classificação, pois, em geral, é mais fácil distinguir padrões usando uma base em que seus vetores não são correlacionados e que apontam para a direção da maior variância dos dados.

Os elementos dos autovetores representam os cossenos diretores, ou seja, a contribuição com que cada um dos eixos originais entra na composição dos novos eixos, as componentes principais. Os autovalores, por sua vez, representam a quantidade de variância original descrita pelos respectivos autovetores, os *scores*. (Silva et al., 2005).

Os scores são as projeções das amostras na direção das componentes principais. Suas coordenadas são obtidas a partir do produto da matriz de dados pela matriz de autovetores. (Silva et al., 2005).

O último passo é a diagonalização. A matriz P é empregada para mudar a base de C_z obtendo uma matriz diagonal D de autovalores de C_z (Equação (2.6)).

$$D = P^{-1}C_zP \quad (2.5)$$

A matriz D apresenta elementos iguais aos autovalores na diagonal principal, ou seja, não apresenta covariância, conseqüentemente não tem nenhuma informação redundante. (Ludwig Jr. e Montgomery, 2007).

Na transformada de Hotelling diferente do processo de diagonalização de matrizes, os autovetores são arranjados de modo decrescente de acordo com os valores dos n autovalores. A importância disso está na representação ordenada dos componentes principais, onde o primeiro componente é representada no eixo de maior variância original dos dados, um eixo ao longo do qual os pontos se encontram espalhados de forma máxima. O segundo componente principal, ortogonal ao primeiro, representa o segundo eixo de maior variância, ou seja, o eixo com máxima quantidade de variância não explicada pelo primeiro autovetor, e assim sucessivamente, até o último componente principal. Resumindo, como cada componente captura o máximo de variância possível, o componente seguinte resta à variância residual, que se torna cada vez menor a cada componente sucessivo.

2.4. Redes Neurais Artificiais

Redes Neurais Artificiais (Haykin, 2001) são modelos computacionais não lineares, inspirados na estrutura de neurônios interconectados existente no cérebro humano, capazes de realizar operações de aprendizado, associação, generalização

e abstração. As redes neurais são compostas por diversos elementos processadores (neurônios artificiais), altamente interconectados, que efetuam operações simples, transmitindo seus resultados aos processadores vizinhos. A habilidade das redes neurais em realizar mapeamentos não-lineares entre suas entradas e saídas as tem tornado prósperas no reconhecimento de padrões (Bishop, 1995) e na modelagem de sistemas complexos.

Na literatura, podem-se encontrar muitos tipos de redes neurais, com diferentes arquiteturas e algoritmos de aprendizado. Embora exista uma grande quantidade de arquiteturas de redes neurais, a estrutura multicamada (MLP - *Multilayer Perceptron*) é a mais conhecida e utilizada (Hush, 1993), devido à capacidade de aproximação de funções e de generalização para uma ampla classe de problemas (Iyoda, 2000).

2.4.1. Arquitetura

A arquitetura multicamada (MLP) é a mais utilizada nas aplicações de engenharia. Conforme indicado pelo próprio nome, a arquitetura MLP é organizada em diversas camadas: uma camada de entrada, formada pelos neurônios que estão conectados às entradas globais da rede, isto é, recebem os atributos de entrada; uma camada de saída, contendo os neurônios que apresentam as saídas da rede neural ao ambiente externo; e uma ou mais camadas intermediárias (ou escondidas), compostas de neurônios cujas entradas e saídas estão conectadas somente a outros neurônios, não havendo interação com o ambiente externo à rede.

Portanto, de modo geral, esta arquitetura possui um vetor de entradas x de dimensão d , $x = [x_1, x_2, \dots, x_d]$; um vetor de c saídas $y = [y_1, y_2, \dots, y_c]$; e M neurônios na camada escondida (podendo possuir mais de uma camada escondida).

Todos os parâmetros adaptáveis (pesos e bias) desta arquitetura são agrupados convenientemente em um único vetor w -dimensional $w = [w_1, w_2, \dots, w_w]$; para facilitar os tratamentos analíticos.

Além disso, para o cálculo dos parâmetros adaptáveis (treinamento) $w = [w_1, w_2, \dots, w_w]$, utiliza-se um conjunto de observações (dados de

treinamento) $D = \{(x^{(1)}, t^{(1)}), (x^{(2)}, t^{(2)}), K, (x^{(N)}, t^{(N)})\}$ de algum processo a ser modelado, onde $x = \{(x^{(1)}), (x^{(2)}), K, (x^{(N)})\}$ é o conjunto de dados de entrada e $t = \{(t^{(1)}), (t^{(2)}), K, (t^{(N)})\}$ é o conjunto de dados contendo a saída desejada (Figura 3).

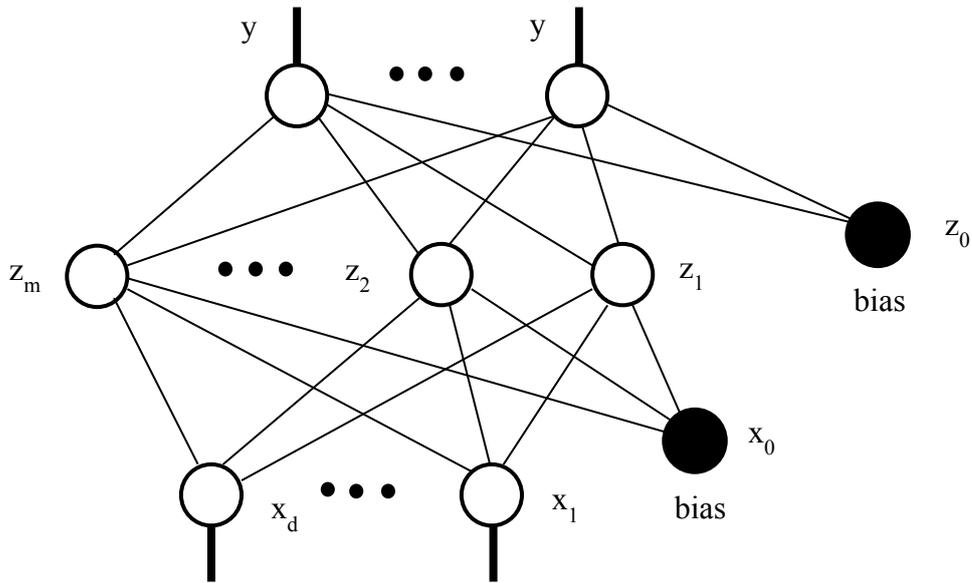


Figura 3: Representação geral da arquitetura *Multilayer Perceptron (Feedforward)*.

2.4.2. Algoritmos de Aprendizado

Redes neurais com treinamento por retropropagação do erro (*backpropagation*) (Haykin, 2001; Wasserman, 1993) são redes *feedforward*, de uma ou mais camadas escondidas, cujo algoritmo de aprendizado define uma maneira sistemática de atualização dos pesos das diversas camadas baseada na ideia de que os erros dos neurônios das camadas escondidas são determinados pela retropropagação reversa dos erros dos neurônios da camada de saída.

O treinamento supervisionado comumente usado é baseado no método do gradiente decrescente (*gradient descent*), buscando minimizar o erro global da camada de saída. Deste modo, a atualização do peso ($\Delta\omega_{ji}$) é proporcional ao negativo da derivada parcial do erro com relação ao próprio peso (Equação (2.6)):

$$\Delta\omega_{ji} = -\eta \frac{\partial E_{SSE}}{\partial \omega_{ji}} \quad (2.6)$$

Onde:

η é a taxa de aprendizado e E_{SSE} (*Sum of Squared Errors*) é a função erro definida conforme Equação (2.7):

$$E_{SSE} = \frac{1}{2} \sum_{p=1}^{N_p} \sum_{j=1}^{N_o} (t_j^p - s_j^p)^2 \quad (2.7)$$

Onde

N_o : número de processadores da camada de saída

N_p : número de padrões de treinamento

t_j^p : valor esperado na saída do processador j ao se apresentar o padrão

p .

Derivando a Equação (2.7), chega-se à Equação (2.8):

$$\Delta \omega_{ji} = -\eta s_i e_j \quad (2.8)$$

$$e_j = \begin{cases} (t_j - s_j) f'(net_j) & \text{se } j \in \text{camada de saída} \\ f'(net_j) \sum_{k=1}^N \omega_{kj} e_k & \text{se } j \in \text{camada escondida} \end{cases}$$

Onde:

η : taxa de aprendizado;

s_i : entrada associada ao peso w_{ji} ;

e_j : erro do j -ésimo processador;

t_j : valor desejado de saída do processador j ;

s_j : estado de ativação;

net_j : potencial interno;

f' : derivada da função de ativação;

N : número de processadores na camada seguinte à camada do processador j .

A taxa de aprendizado η é um parâmetro importante a ser definido no aprendizado. Esta não deve ser nem muito baixa, causando um treinamento muito lento, nem muito alta, gerando oscilações. Quando a taxa de aprendizado é baixa, e dependendo da inicialização dos pesos (feita de forma aleatória), a rede neural pode ficar presa em um mínimo local. Quando a taxa de aprendizado é alta, a rede neural pode nunca conseguir chegar ao mínimo global, pois os valores dos pesos são grandes. A solução para este problema é utilizar uma taxa de aprendizado

adaptativa. Além deste parâmetro, pode-se também utilizar um termo α de momento (Haykin, 2001), proporcional à variação no valor do peso sináptico no passo anterior. Deste modo, a atualização do peso sináptico w_{ji} é modificada conforme a Equação (2.9):

$$\Delta\omega_{ji}(t + 1) = \eta s_i e_j + \alpha \Delta\omega_{ji}(t) \quad (2.9)$$

A utilização do termo de momento tem a função de acelerar a convergência da rede, sem causar oscilações.

Como se pode verificar, o algoritmo de aprendizado do *backpropagation* tem duas fases, para cada padrão apresentado: *forward* e *backward*. Na primeira etapa, as entradas se propagam pela rede, da camada de entrada até a camada de saída, gerando a saída da rede em resposta ao padrão apresentado. Na segunda etapa, os erros se propagam na direção contrária ao fluxo de dados, indo da camada de saída até a primeira camada escondida, atualizando os pesos sinápticos. Este procedimento de aprendizado é repetido diversas vezes, até que, para todos os processadores da camada de saída e para todos os padrões de treinamento, o erro seja menor do que o especificado.

Foi demonstrado que o algoritmo *backpropagation* é um aproximador universal (Hornik, 1989), sendo capaz de aprender mapeamentos de entrada-saída. Entretanto, apesar do grande sucesso do *backpropagation* nas mais diferentes aplicações, existem alguns problemas básicos: a definição do tamanho da rede, o longo processo de treinamento, e fenômenos como paralisia da rede (contornado diminuindo o valor de η) e mínimo local (que pode ser solucionado utilizando-se métodos estatísticos).

A definição do tamanho da rede, isto é, o número de camadas escondidas e número de processadores em cada uma dessas camadas, é um compromisso entre convergência e generalização. Convergência é a capacidade da rede neural de aprender todos os padrões do conjunto de treinamento.

Generalização é a capacidade de responder corretamente aos padrões nunca vistos (conjunto de teste). O objetivo é utilizar a menor rede possível, de forma a se obter uma boa generalização, que seja capaz de aprender todos os padrões (Feitosa, 1999).

Para garantir a capacidade de generalização da rede, uma ferramenta padrão da estatística conhecida como validação cruzada fornece um princípio orientador atraente (Stone, 1974, 1978). A técnica consiste em dividir os dados disponíveis primeiramente em duas partes: conjunto de treinamento e conjunto de testes. O conjunto de treinamento é dividido adicionalmente em dois subconjuntos disjuntos: subconjunto de estimação e subconjunto de validação (Haykin, 2001). Com isto, é possível verificar se o modelo, com um conjunto de dados diferente daquele usado para estimar os parâmetros da rede, responde corretamente aos padrões apresentados.

O método de treinamento com parada antecipada (*early stopping*) faz uso dos subconjuntos de estimação e validação para determinar o momento de parada do treinamento, evitando que a rede se ajuste excessivamente ao subconjunto de estimação. O método consiste em apresentar o subconjunto de validação ao modelo ajustado após um determinado tempo de treinamento (Haykin, 2001). Quando o erro obtido a partir do subconjunto de validação – erro de validação – cresce de forma consistente, o treinamento da rede é interrompido (Figura 4).

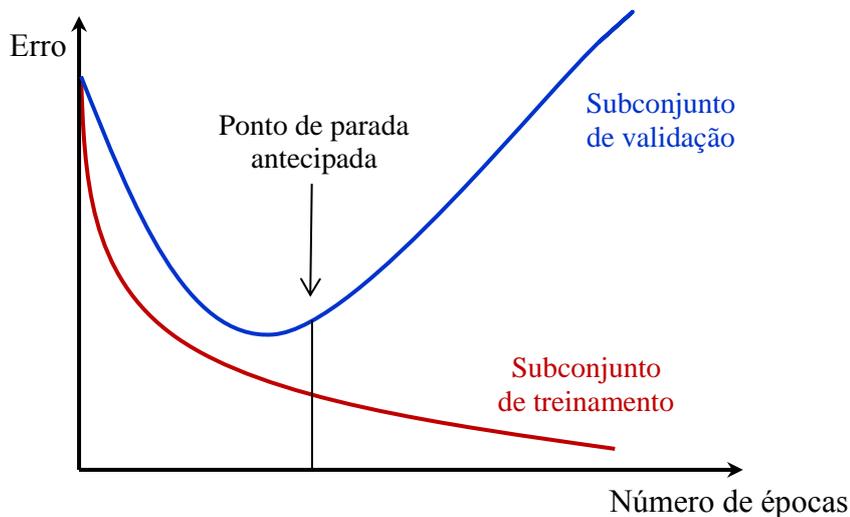


Figura 4: Identificação do ponto de interrupção do treinamento com o método da parada antecipada.

O treinamento supervisionado utilizado nesse trabalho foi o Levenberg-Marquardt. Esse algoritmo é considerado o método mais rápido para treinamento de redes *feedforward backpropagation*, que possui uma quantidade moderada de pesos sinápticos. Ele se baseia, para a aceleração do treinamento, na determinação das derivadas de segunda ordem do erro quadrático em relação aos pesos,

diferindo do algoritmo *backpropagation* tradicional que considera as derivadas de primeira ordem.

O algoritmo de Levenberg-Marquardt (Saini and Soni, 2002) se baseia no método de otimização de Newton, que faz uso da matriz Hessiana H . No método de Levenberg-Marquardt se faz uma aproximação para essa matriz, mostrada na Equação (2.10), determinada em função da matriz Jacobiana, que contém as primeiras derivadas dos pesos em função dos pesos sinápticos, expressa na Equação (2.11):

$$H = \frac{\partial^2 E_R(W)}{\partial W^2} \quad (2.10)$$

$$J = \frac{\partial e(W)}{\partial W} \quad (2.11)$$

onde $e(W)$ é definido conforme a Equação (2.12):

$$e(W) = \sum_{i=1}^n (y_i - y_{ei}) \quad (2.12)$$

A determinação da matriz Jacobiana é muito mais simples que a determinação da matriz Hessiana. Como, para uma rede neural, a performance de treinamento é expressa em função da soma dos erros quadráticos, a matriz Hessiana pode ser expressa pela Equação (2.13):

$$H = J^T(W) \cdot J(W) \quad (2.13)$$

O método de Newton atualiza os pesos segundo a Equação (2.14):

$$W(K + 1) = W(K) - H^{-1} \cdot g_k \quad (2.14)$$

onde g_k pode ser escrito conforme (2.15):

$$g_k = 2J^T(W) \cdot e(W) \quad (2.15)$$

O algoritmo de Levenberg-Marquardt procede a atualização dos pesos baseado na mesma expressão do método de Newton (Equação (2.14)), realizando

as modificações para a determinação da matriz Hessiana, mostrada na Equação (2.16):

$$W(K + 1) = W(K) - [J^T(W) \cdot J(W) + \mu K \cdot I]^{-1} \cdot J^T(W) \cdot e(W) \quad (2.16)$$

onde

- I: matriz identidade
 μK : constante do método de Levenberg-Marquardt.

O parâmetro μK funciona como um fator de estabilização do treinamento, ajustando a aproximação de forma a utilizar a rápida convergência do método de Newton e evitando passos muito grandes que possam levar a um erro de convergência.

Esse método apresenta convergência em menos iterações, mas requer mais cálculos por iteração devido ao cálculo de matrizes inversas. Apesar do grande esforço computacional, ele segue sendo o algoritmo de treinamento mais rápido para redes neurais, quando se trabalha com um número moderado de parâmetros na rede. Se esse número é elevado, a utilização desse algoritmo é pouco prática.

2.4.3. Redes Neurais *Multi-Step*

Quando o objetivo da rede neural é prever uma série temporal, isto é, prever valores futuros de uma determinada variável baseado em uma base de dados que contém valores passados da referida variável, deve-se escolher uma janela de previsão, isto é, o número de valores anteriores que serão usados como dados de entrada para a realização a previsão. Trata-se de um parâmetro muito importante neste caso, pois consiste em uma decisão a respeito do número de valores passados que é relevante para a previsão de um valor futuro.

Outro aspecto importante a ser destacado a respeito da previsão de séries temporais utilizando Redes Neurais é o chamado horizonte de previsão. O horizonte de previsão é o número de valores futuros que se deseja prever a partir da série de valores passados.

Uma Rede Neural tem uma única saída, que seria referente ao “próximo” valor da série temporal em questão. Trata-se da chamada previsão *single-step* (um passo a frente). No entanto, em certas aplicações, como no caso deste trabalho,

pode ser interessante prever resultados mais a frente, não se limitando ao próximo valor. Para atingir tal objetivo, é necessário partir de uma previsão single-step e, assumindo que seu valor é de fato o valor correto, utilizá-lo na entrada da rede para, então, prever o valor seguinte. É importante notar que o primeiro valor previsto não será, evidentemente, o valor exato, mas, como não se sabe qual será o número correto, a previsão é utilizada como estimativa. Desta maneira, quando um valor previsto e, portanto, incorreto é utilizado como entrada para uma nova previsão, o erro do valor seguinte será ainda maior. Este processo poder ser repetido para o número de valores à frente que se desejar prever; trata-se da previsão *multi-step* (vários passos à frente).

2.5. Produção de Petróleo

O petróleo tem sua origem a partir da matéria orgânica depositada junto com os sedimentos de rochas. A interação desses componentes sob condições termoquímicas apropriadas é fundamental para o início da cadeia de processos que leva à sua formação.

Para haver acumulação de petróleo é necessário que ocorra a migração desse material a outras rochas após o processo de geração na rocha geradora. Os fluidos devem ter os caminhos interrompidos pela existência de algum tipo de armadilha geológica, as trapas ou rocha selante, o que permite o seu aprisionamento nas rochas chamadas de reservatório. (Rosa, 2006)

Sem esse mecanismo geológico, o petróleo continuaria migrando em busca de zonas de menor pressão até se perder na superfície através de exsudações, oxidação ou degradação bacteriana.

Para uma rocha se constituir em um reservatório deve apresentar espaços vazios interconectados em seu volume. Define-se porosidade os espaços vazios e permeabilidade a interconexão entre estes espaços. Desta forma, as rochas que podem constituir um reservatório são geralmente os arenitos e calcarenitos, também sendo possíveis todas as demais rochas sedimentares que apresentem espaços vazios interconectados (Thomas, 2001).

Segundo Rosa *et al.* (2004), os fluidos contidos em uma rocha reservatório devem dispor de certa quantidade de energia para que possam ser produzidos, ou

seja, vencer toda a perda de carga oferecida pelos canais porosos na rocha reservatório, com suas tortuosidades e estrangulamentos, e se deslocar para os poços de produção. Essa energia é referida como energia natural ou primária, sendo manifestada por certa quantidade de pressão, como resultado de todos os fenômenos geológicos pelo qual a jazida passou até se formar completamente.

De um modo geral, a produção de fluidos é devida a dois efeitos principais. O primeiro é a descompressão, que causa a expansão dos fluidos contidos no reservatório e a contração do volume poroso. O segundo é o deslocamento de um fluido por outro fluido que como exemplo ocorre na injeção de água no reservatório para aumentar a recuperação. A razão entre o volume produzido de óleo e o volume total presente de óleo no reservatório é chamada fator de recuperação.

Deste modo, ao longo da produção de fluidos a pressão no reservatório cai até atingir a pressão de gaseificação dos componentes mais leves, a chamada pressão de saturação. Como o gás é mais expansível que o líquido, sua expansão ajuda no deslocamento do líquido para os poços. Quanto maior a queda de pressão devida à produção de fluidos pelo reservatório, maior é a dissociação de gás na zona de óleo, e com isso grande parte do gás passa a ser produzido junto com o óleo. A pressão do reservatório declina rápida e continuamente, reduzindo drasticamente a vazão de óleo produzida a valores antieconômicos.

A injeção de água em reservatórios traz ótimos resultados ao aumento da recuperação. O objetivo da injeção de água é aumentar a recuperação por meio do deslocamento de óleo aos poços produtores e, como consequência, manter a pressão original do reservatório evitando a depleção prematura. Através de um balanço de massas, determina-se o volume retirado do reservatório que deve ser o mesmo volume de água a ser injetado de modo a manter sua pressão próxima a original ao longo do tempo. A fração de gás produzida é relativamente constante devido à pressão no reservatório permanecer estável. Após um dado volume de produção de óleo, a água injetada chega aos poços produtores sendo produzida em conjunto com o óleo, aumentando consideravelmente a fração de água dos poços. Poços injetores são posicionados no reservatório de modo a maximizar o efeito da injeção de água na otimização da produção.

Em muitos casos, o aumento da fração de água derivada de injeção pode dificultar a produção de óleo. O incremento do volume de água no líquido

produzido pelo reservatório tem como consequência maior valor de pressão hidrostática sobre o fundo do poço. Tal fato é devido à maior densidade da água comparando-se com a do óleo. Para haver fluxo, a pressão do reservatório deve ser suficientemente alta para vencer a coluna hidrostática e as perdas de carga desde o meio poroso até a plataforma, envolvendo toda a tubulação.

Para evitar a queda de vazão de óleo e mantê-la a níveis econômicos, torna-se necessária a suplementação da energia do reservatório ou a redução de perdas de pressão no escoamento. Métodos de elevação artificial atuam neste sentido, por suplementação de energia através de bombeio ou reduzindo as perdas de pressão no escoamento por meio da redução da coluna hidrostática de fluido. Os métodos mais comuns de elevação artificial na indústria do petróleo são o *Gas Lift* Contínuo, e o Intermitente, o Bombeio Centrífugo Submerso, o Bombeio Mecânico com Haste e o Bombeio por Cavidade Progressiva. Na Bacia de Campos o método de elevação artificial mais utilizado é o *Gas Lift* Contínuo, que tem o objetivo de reduzir a pressão de fluxo no fundo do poço por meio da redução da pressão de coluna hidrostática com consequente aumento de vazão do poço. Certa vazão de gás é injetada nos poços, na coluna de produção, gaseificando o volume de líquido acima do ponto de injeção de forma a reduzir sua densidade, resultando em queda de pressão de coluna hidrostática (Rosa et al., 2004).

2.6. Sistemas Marítimos de Produção

Um sistema marítimo de produção consiste basicamente de uma ou mais unidades estacionárias de produção e de equipamentos posicionados tanto sobre esta unidade quanto no fundo do mar.

No presente trabalho, a modelagem do sistema de produção é feita através de um simulador de superfície para que o modelo de produção de petróleo seja mais próximo possível do real.

2.6.1. Linha de Produção e *Riser*

No caso de produção de óleo *offshore*, todo fluido extraído do reservatório é enviado à unidade estacionária de produção (UEP) por tubulações, chamadas de linhas de produção e *risers*. Estas tubulações fazem a interconexão entre os elementos submarinos e o sistema de produção. As linhas de produção ou *flowlines* interligam a árvore de natal molhada à base do *riser*, enquanto o *riser* liga a linha de produção no fundo do mar à superfície, onde se encontra posicionada uma UEP. A linha de produção apoia-se no leito marinho enquanto o *riser* é suspenso em catenária da plataforma até o fundo do mar, como mostra a Figura 5. O *riser* também pode estar perpendicular ao leito marinho, como no presente trabalho, em vez de formar uma catenária, quando UEPs fixas são utilizadas (Moreira, 2005).

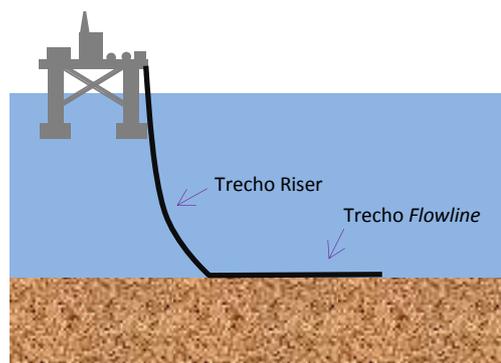


Figura 5: Flowline e Riser

2.6.2. Manifolds

Na prática, existem limitações quanto ao número de poços que podem ser ligados a uma UEP. Se há mais poços que o número possível de ligações, são utilizados *manifolds* (Figura 6), que organizam os poços em clusters. Um *manifold* é um conjunto de válvulas no fundo do mar que funciona como uma subestação de captação, ou seja, um concentrador do petróleo produzido por vários poços que envia essa produção para a plataforma por uma única tubulação. O objetivo de utilizar *manifolds* é reduzir o número das tubulações que interligam os poços à plataforma, reduzindo o comprimento total de linhas de produção. De

forma semelhante à UEP, os *manifolds* também têm limitações quanto ao número máximo de poços interligados (Cerqueira, 2005).

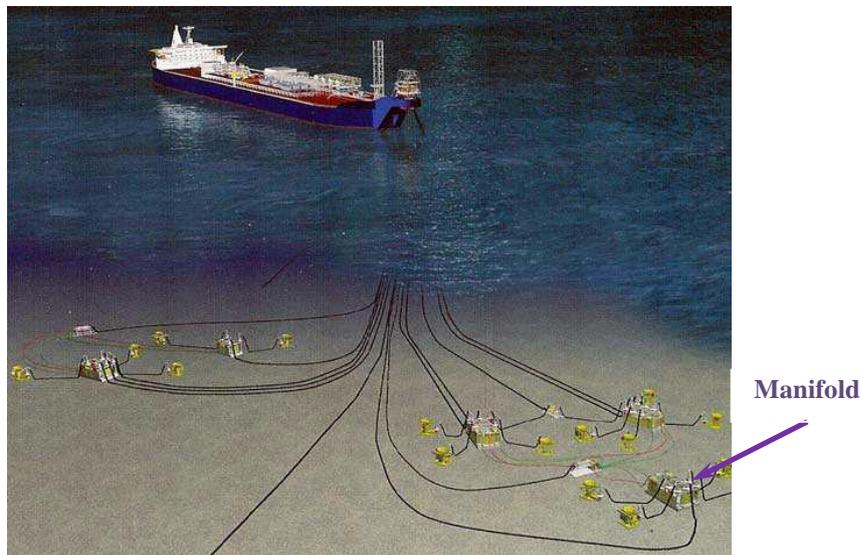


Figura 6: *Manifold* (Cerqueira, 2005)

A capacidade máxima da tubulação de coleta do óleo produzido pode também determinar o uso de *manifolds*. Poços com altas vazões podem ter o fluxo restringido no *manifold* para atender a capacidade de escoamento das tubulações e de processamento da plataforma (Franco, 2003).

No presente trabalho, *manifolds* virtuais são criados para a integração entre as tabelas de fluxo vertical multifásico do poço e da superfície, como é descrito no Capítulo 3.

2.7. Fluxo Multifásico

O fluxo multifásico compreende o escoamento simultâneo de óleo, gás e água. A origem do gás no fluxo dos poços à plataforma deve-se à queda de pressão ao longo do escoamento, chegando a valores inferiores à pressão de saturação com conseqüente liberação de gás livre. A liberação de gás pode iniciar no reservatório ou em qualquer outro ponto nas tubulações. À medida que a pressão torna-se cada vez menor que a pressão de saturação, maiores volumes de gás são liberados pelo óleo. Outra fonte de gás no escoamento é devida a métodos de elevação artificial (*Gas Lift*). No método de elevação artificial *Gas Lift*, gás

pode ser injetado na coluna de produção para auxiliar a elevação reduzindo a coluna hidrostática de fluido. É importante lembrar, como menciona Oliveira (2003), que o gás derivado da injeção e da liberação pelo óleo geralmente flui com velocidades superiores à do líquido causando o que se chama de escorregamento entre as fases.

A injeção de água ou a presença de um grande volume de água no reservatório, os chamados aquíferos, determinam a presença de água no fluido produzido. Ao longo do tempo de produção de um campo, a injeção contínua de água satura o reservatório de modo a atingir os poços e ser produzida com o óleo. Ao longo da vida produtiva do reservatório, a água pode alcançar os poços de produção e em alguns casos o poço pode passar a produzir mais água que óleo. A água e o óleo produzidos podem estar separados em duas fases por diferença de densidades, chamada de água livre, ou emulsionados em uma única fase líquida.

Desta forma, os fluidos produzidos pelo reservatório incluem óleo, gás e em muitos casos, água livre ou emulsionada ao óleo.

2.7.1. Tabelas de Fluxo Vertical Multifásico

Curvas de fluxo vertical multifásico (VLP - *Vertical Lift Performance*) descrevem a pressão de fundo de poço para várias taxas de um determinado conjunto de condições de fluxo (WHP – *Well Head Pressure*, GOR – *Gas Oil Rate*, *WaterCut* e *Gaslift: Gas Injection Rate*). (Petex, 2012)

Para cada taxa informada, um cálculo de gradiente é efetuado para descobrir o FBHP (*Flowing Bottom Hole Pressure*), ou pressão de fundo de poço, e, em seguida, essas pressões são juntadas para formar uma curva. Um exemplo de curva para um poço multifásico surgente é mostrado na Figura 7.



Figura 7: Curva de Fluxo Vertical Multifásico

A curva no gráfico da Figura 7 mostra o típico formato da curva VLP, que pode ser dividida em duas seções: direita e esquerda do ponto mínimo de inflexão.

O lado esquerdo é chamado de região instável da curva. Nessa seção, um aumento na taxa de líquido causa uma diminuição na FBHP requerida para elevar o fluido à superfície. Isso pode parecer contra intuitivo, mas é em parte causado pelo grande escorregamento entre fases que ocorre entre o gás e a fase líquida em função da baixa velocidade do líquido. O grande escorregamento leva a uma densidade maior da mistura e a um maior decaimento de pressão (Petex, 2012).

O lado direito é chamado de região estável. Nessa seção da curva, um aumento nas taxas leva a um aumento na FBHP requerida para elevar o fluido. O aumento de velocidade (e, em função disso, da fricção) causado pelo aumento da taxa de líquido age para estabilizar o poço e reduzir a probabilidade de instabilidades (Petex, 2012).

Em função de problemas com instabilidade, o ponto mínimo da VLP é comumente chamado de ponto mínimo de estabilidade já que qualquer valor de taxa menor do que esse ponto estará na região instável.

As tabelas de fluxo multifásico podem ser geradas calculando-se a pressão de fundo de poço para uma variedade de valores de:

- Taxa gás-óleo (GOR);
- Pressão de cabeça de poço (WHP);
- *Gaslift* – taxa de injeção de gás;
- *WaterCut*;

O sistema acoplado utiliza as tabelas VLP para determinar a pressão de fundo de poço para qualquer conjunto de condições, e também para ajustar a curva de IPR

(*Inflow Performance Relationship* - curva gerada pelo simulador de reservatório que representa a pressão disponível no reservatório) para cada poço. A curva é continuamente ajustada enquanto há mudanças de pressão no reservatório. O sistema utiliza o IPR ajustado e as tabelas VLP para determinar a interseção IPR-VLP.

2.8. Simuladores

Com a crescente demanda por recursos energéticos em todo o mundo, torna-se necessário para a sociedade otimizar o consumo de energia, buscar fontes alternativas, e sobretudo alcançar a autossuficiência em fontes de energias vitais como o petróleo e o gás-natural. Tal demanda fomentou a criação de áreas de atuação profissional tais como a engenharia de petróleo e a engenharia de reservatórios, que tratam de todos os ramos relacionados à indústria do petróleo e, em particular aos relacionados à exploração e produção.

Dentre as funções de um engenheiro de petróleo estão a quantificação de reservas, o dimensionamento de sistemas de produção, e a criação de projetos de extração de hidrocarbonetos. A execução de projetos de extração, que podem envolver investimentos de centenas de milhões de dólares, deve ter seus riscos calculados e minimizados. Entretanto, fatores físicos e químicos complexos como variações regionais das propriedades dos fluxos e características da permeabilidade relativa da rocha devem ser levados em consideração no cálculo do risco. A complexidade dessa tarefa levou à criação área de simulação de reservatórios (Oliveira, 2009).

Na engenharia de petróleo, uma das tarefas mais importantes é a previsão da produção de hidrocarbonetos de reservatórios. Essa tarefa permite quantificar reservas, avaliar e priorizar projetos de exploração e dimensionar sistemas de produção para otimizar a produção de um reservatório. Nesse sentido, a simulação computacional de reservatórios de petróleo vem sendo utilizada há décadas, e tem se tornado, cada dia mais, uma importante ferramenta para a indústria de exploração de petróleo e gás natural. Entretanto, esses simuladores requerem um grande poder de processamento para obter resultados confiáveis em um tempo

aceitável. Esse custo computacional aumenta ainda mais quando o simulador de reservatório é acoplado a um simulador de superfície.

2.8.1. Simuladores de Reservatório

Um simulador de reservatório é um modelo computacional capaz de gerar previsões do comportamento de um reservatório de petróleo sob diferentes condições de operação (Ertekin *et al*, 2001). Esse modelo tem como base a teoria de fluxo em meios porosos (Soares, 2002), que possui um campo de aplicação bastante abrangente, sendo empregada não só da engenharia de reservatórios, mas também em diversas outras áreas da ciência e engenharia como: modelagem de aquíferos, mecânica dos solos, fluxo de contaminantes e análises de fluxo sanguíneo no interior do corpo humano.

Alguns dos modelos de fluxo mais utilizados em simulações de reservatórios são: o modelo *black-oil*, o modelo composicional, os modelos térmicos e os modelos de fluxo miscível (Mattax *et al.*, 1990).

O modelo *black oil*, que é o modelo utilizado nesse trabalho, considera o sistema com apenas três componentes - óleo, água e gás - e três fases, também designadas por óleo, água e gás. Assume-se que o componente óleo só existe na fase óleo, o componente água só existe na fase água e o componente gás pode se encontrar como gás livre no reservatório, ou dissolvido no óleo. Esse modelo considera também que a temperatura do reservatório é constante e que não há reações químicas entre os componentes. A sua utilização é recomendada para reservatórios que possuem óleos pesados e com baixa volatilidade, como os encontrados no Brasil (Cordazzo, 2006).

Por obter resultados satisfatórios em espaços de tempo aceitáveis, o modelo *black oil*, mesmo não sendo o modelo mais completo existente, é o mais utilizado pelos simuladores comerciais de reservatórios (FEITOSA *et al.*, 1982; CMG, 2012; SCHLUMBERGER, 2013). Os demais modelos são mais completos, entretanto, para a maioria dos casos práticos, são altamente custosos do ponto de vista computacional. Este trabalho, por não pretender apresentar contribuição específica no que se refere à modelagem de reservatórios, abordará apenas o modelo *black oil*.

2.8.2. Simuladores de Superfície

Há vários simuladores de escoamento multifásicos na indústria do petróleo. O principal objetivo do simulador de escoamento multifásico é o cálculo do gradiente dinâmico de pressão-temperatura de escoamento numa tubulação, ou seja, calcular os valores de pressão e temperatura nos diferentes pontos do sistema de produção (Rosa, 2006).

Entretanto, para a determinação destes perfis é necessário conhecer o comportamento das fases e o cálculo das propriedades físicas do fluido em escoamento em vários estados termodinâmicos.

Os simuladores de escoamento multifásico comerciais utilizam alguns tipos de procedimentos para a determinação das propriedades dos fluidos e comportamento das fases. Estes procedimentos podem ser correlações empíricas do tipo *black oil* e, quando se sabe a composição e fração molar dos componentes do fluido, as equações de estado e as tabelas de propriedades físicas. Usualmente a composição e fração molar são utilizadas para modelagem de escoamento de gás e as correlações *black oil*, para escoamento de óleo.

Diversas correlações empíricas comumente utilizadas no cálculo do gradiente de pressão estão disponíveis nos simuladores de escoamento multifásicos comerciais. A utilização desses simuladores permite uma avaliação do desempenho dos modelos de poços, linhas de produção e *risers* frente a dados reais de campo. Da mesma forma é possível avaliar se as correlações utilizadas nestes modelos descrevem as condições reais de fluxo e realizar ajustes.

Após o modelo e suas correlações estarem em conformidade com os dados de campo, o simulador é utilizado para o controle dos parâmetros operacionais dos poços e do sistema de escoamento de fluidos. Além disso, o simulador permite o dimensionamento e análise de sistemas de produção a serem instalados.

3. Modelagem do Aproximador

3.1. Introdução

Um simulador de superfície acoplado ao simulador de reservatório é crítico para que o sistema de produção seja mais bem descrito como um todo. Assim, a integração entre as instalações de superfície e o reservatório permite que seja modelado, com mais precisão, as capacidades e limitações de produção. Para se fazer o acoplamento entre o simulador de reservatório e o simulador de superfície, deve-se ter uma ferramenta que faça a sincronização e a troca de dados entre os simuladores.

Para diminuir o tempo necessário para a realização de simulações integradas, e para que o número de simulações necessárias não seja um fator limitante em processos de otimização e de análise de incertezas, diversas técnicas inteligentes podem ser usadas para aproximar o resultado da simulação integrada original. Construir um aproximador para substituir parcialmente ou totalmente os simuladores de reservatório e superfície torna possível a utilização, de forma menos custosa (computacionalmente) e adequada, de processos que requerem grande quantidade de simulações.

Neste trabalho, utilizam-se Redes Neurais Artificiais do tipo Multilayer Perceptron (MLP) com algoritmo de aprendizado de minimização por mínimos quadrados (Levenberg-Marquardt) para a construção de um aproximador de função não linear para a simulação integrada entre o reservatório e a superfície. Os dados de treinamento, validação e teste são frutos de amostragens aleatórias, amostragem por Hipercubo Latino (Iman *et al.*, 1981), e Análise de Componentes Principais - PCA (Jolliffe, 2002).

3.2. Integração do Simulador de Reservatório ao Simulador de Superfície

A Modelagem Integrada de Produção (MIP) consiste na combinação dos modelos de engenharia de reservatórios, produção e superfície em uma ferramenta de gerenciamento que permita a simulação integrada de todo o sistema de produção de petróleo. Apesar de não serem recentes os primeiros estudos envolvendo acoplamento entre estas diferentes áreas, nos últimos anos o crescimento do número de artigos publicados que envolvem MIP é enorme e merece atenção especial.

A necessidade crescente de uma maior interação entre os sistemas de subsuperfície e superfície levou muitas companhias operadoras a desenvolverem soluções proprietárias. Em Rotondi *et al.* (2008) é feita uma apresentação de diversas destas iniciativas e são listados os respectivos trabalhos disponíveis na literatura.

3.2.1. Métodos de Acoplamento entre Simuladores

A companhia Petroleum Experts (PETEX) desenvolveu o aplicativo RESOLVE, uma interface para o acoplamento entre diversos simuladores de reservatórios e o GAP, simulador de sistemas de produção da própria PETEX.

O acoplamento de instalações de superfície a simuladores de reservatório pode ser realizada por diversos métodos, que podem ser divididos em dois grupos:

- métodos explícitos onde instalações de superfície e reservatórios são resolvidos em diferentes níveis de tempo.
- métodos implícitos em que os dois sistemas são resolvidos no mesmo nível de tempo.

No caso de acoplamento implícito, os cálculos envolvidos nos sistemas de produção estão geralmente inseridos no mesmo sistema de equações utilizado para obter as propriedades de cada bloco (célula) do modelo geológico que representa o reservatório. Trata-se de um método mais preciso, mas que exige alto custo

computacional e aplicativos desenvolvidos especificamente para este propósito, o que dificulta a sua utilização em grande escala.

Na metodologia de acoplamento explícito, os cálculos dos sistemas de produção são efetuados de forma separada do sistema de equações do meio-poroso, mas resolvidos no início de cada passo de tempo de sincronização entre os modelos de simulação de escoamento de reservatórios e sistemas de produção. De fato, tem precisão menor que o caso anterior, mas atende satisfatoriamente à grande maioria dos problemas práticos.

Devido à sua flexibilidade, o acoplamento explícito é adotado pela grande maioria das soluções que, utilizando um aplicativo externo, integram simuladores comerciais independentes, tanto de superfície quanto de subsuperfície.

3.2.2. Acoplamento IMEX – GAP e PROSPER

O simulador de reservatório IMEX, desenvolvido pela empresa CMG, é um simulador *black oil* capaz de modelar estruturas heterogêneas complexas, processos de recuperação primários e secundários, poços multilaterais, etc. Dentre as informações retornadas pelo simulador IMEX está uma curva que representa a pressão atuante no reservatório para cada poço, chamada de IPR (*Inflow Performance Relationship*). Já o GAP (simulador de superfície), auxiliado pelo PROSPER (modelagem de poços), gera uma curva que representa a pressão requerida pelo sistema para que o óleo escoe até a UEP, também para cada poço. Esta curva é chamada de VLP (*Vertical Lift Performance*), ou ainda, curva de fluxo vertical multifásico. A pressão na UEP e no fundo do poço são condições de contorno. O ponto de cruzamento dessas curvas, ilustrado pela Figura 8, representa o ponto de operação do sistema de produção (pressão de fundo e vazões de fluidos). Como as curvas são formadas por pontos de pressão e vazão, sabe-se então com qual vazão o sistema irá produzir no ponto de operação. O RESOLVE, software que integra o IMEX com GAP através de um acoplamento explícito, retorna uma curva de produção ao longo do tempo baseada na vazão correspondente ao ponto de operação. Sendo assim, é possível obter uma curva de produção diferente para cada alternativa de posicionamento de poços. A posição

dos poços influencia na construção da VLP e conseqüentemente na perda de carga do sistema de produção.

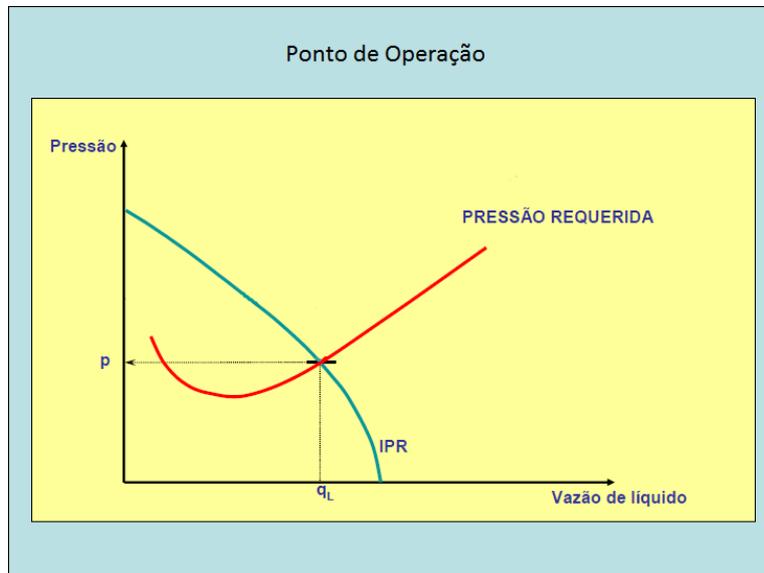


Figura 8: Ponto de operação do sistema de produção. (Ribeiro, 2012)

Para viabilizar a geração de inúmeras simulações para a construção do aproximador, há a necessidade de automatizar a integração entre os simuladores de reservatório e de superfície. O software OPENSERVER, também desenvolvido pela PETEX, é a ferramenta indicada para esta automatização. O esquema da integração é mostrado na Figura 9.

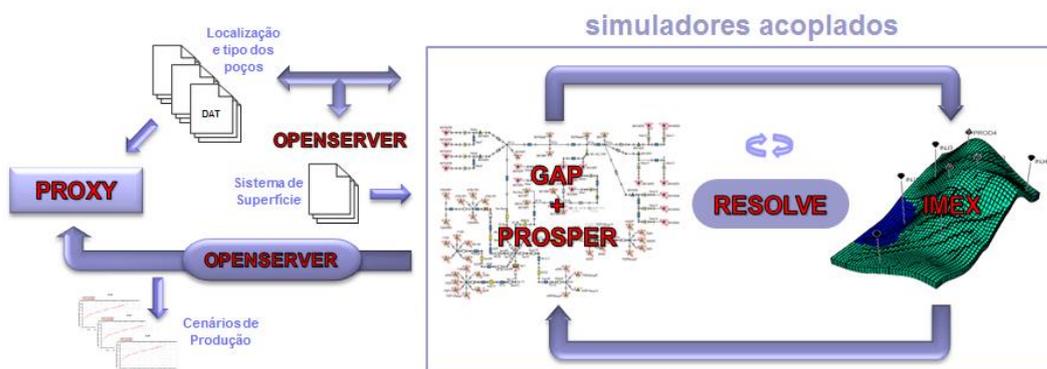


Figura 9: Simulação Integrada

Porém, devido ao fato de não haver licença acadêmica para o OPENSERVER, neste trabalho foi utilizado o software PROSPER (PETEX), para

modelar a perda de carga no sistema de superfície e subsuperfície (poço, linhas de produção e *riser*). Dessa forma, o software PROSPER substitui o GAP na modelagem de linhas de produção e *riser*.

Pelo PROSPER é possível gerar tabelas de fluxo vertical multifásico (VLP). Tais tabelas são geradas previamente à simulação do reservatório, e depois de geradas, são incluídas no arquivo de simulação do IMEX, o que torna desnecessário o uso dos softwares GAP e OPENSERVER.

A substituição do uso do OPENSERVER e do GAP pelo uso somente do PROSPER para modelar o poço e a superfície é uma aproximação, já que os poços compartilham somente uma tabela de fluxo vertical multifásico e, por isso, são modelados até o topo do reservatório, e as linhas de produção também são modeladas aproximadamente, de forma a cobrir variações de distâncias, como explicado posteriormente nesta mesma seção. O impacto da aproximação não é tão extenso pelo fato da perda de carga na linha de produção (trecho horizontal) não ser tão importante quanto a perda de carga em trechos verticais.

Para os aproximadores construídos para a simulação considerando a superfície, foram criados dois tipos de tabelas. Primeiramente foi criada uma única tabela (para todos os poços) para a perda de carga dentro do poço, começando aproximadamente no topo do reservatório e indo até a cabeça do poço, sendo esse trecho sempre vertical. A Figura 10 mostra o exemplo de um poço onde o topo do reservatório é representado pela linha tracejada azul. Em seguida foi criado um conjunto de tabelas que modelam a perda de carga nas linhas de produção e no *riser*, como exemplificado na Figura 11. A linha contínua vertical azul representa o *riser* e a horizontal representa a linha de produção.

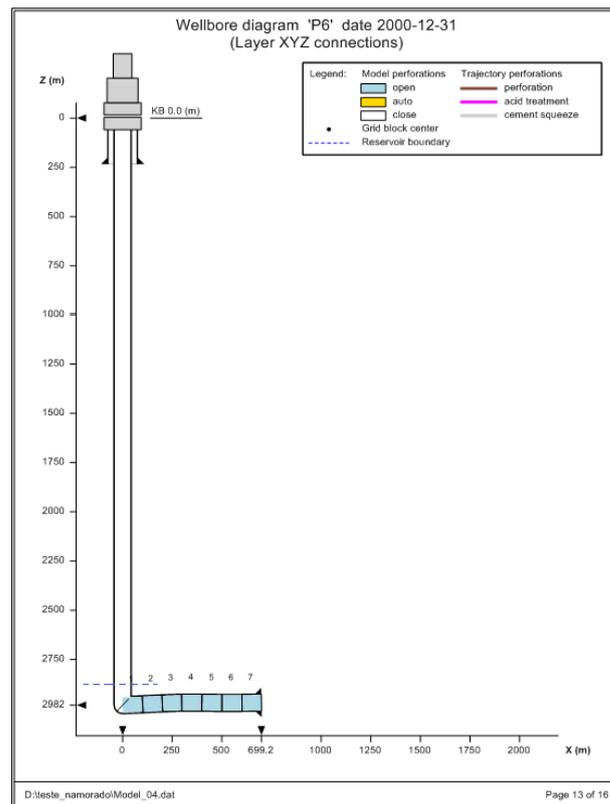


Figura 10: Diagrama do poço (IMEX)

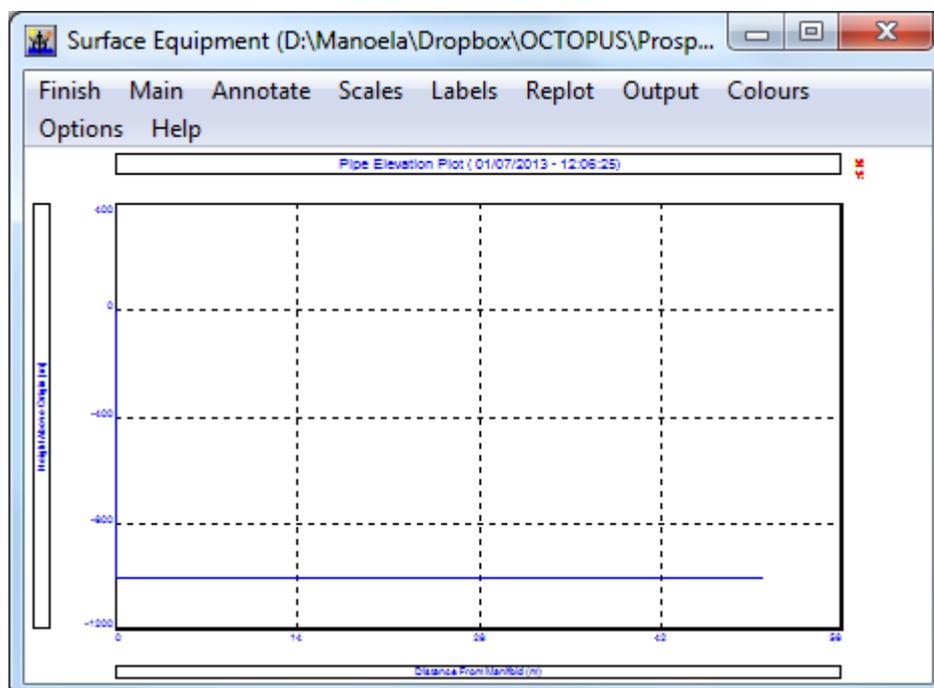


Figura 11: Diagrama de Superfície (PROSPER)

As tabelas das linhas de produção e do *riser* são utilizadas da seguinte forma:

- Poços que tenham a distância euclidiana entre eles e a plataforma compreendida entre 0 e 50 metros utilizam a tabela número 1, que não tem linhas de produção, e modela um *riser* de 1000 metros (lâmina d'água).
- Poços que tenham a distância euclidiana entre eles e a plataforma compreendida entre 50 e 100 metros utilizam a tabela número 2, que modela uma linha de produção com tamanho de 50 metros e um *riser* de 1000 metros (lâmina d'água).
- Poços que tenham a distância euclidiana entre eles e a plataforma compreendida entre 100 e 150 metros utilizam a tabela número 3, que modela uma linha de produção com tamanho de 100 metros e um *riser* de 1000 metros (lâmina d'água).

E assim por diante para todos os intervalos de distância possíveis, como ilustrado na Figura 12.

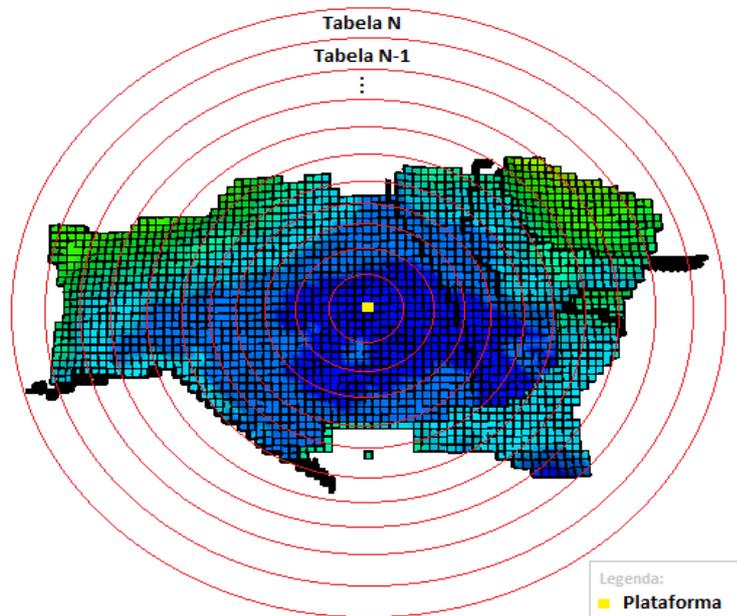


Figura 12: Uso das Tabelas e Fluxo Vertical Multifásico no Reservatório Real

3.2.3. Conexão entre as Tabelas de Fluxo Vertical Multifásico

A conexão entre as duas tabelas (poço + linhas de produção e *riser*) é feita utilizando uma funcionalidade do IMEX que possibilita a criação de *manifolds* submarinos imaginários¹ que fazem a ligação entre a cabeça do poço (tabela de fluxo vertical multifásico 1) e a linha de produção (tabela de fluxo vertical x correspondente à distância euclidiana entre o poço e a plataforma). Portanto, para cada poço criado no reservatório é criado um *manifold* submarino imaginário.

No IMEX, cada poço produtor pertence a um grupo diferente, sendo assim, se há, por exemplo, seis poços, ter-se-ão então seis grupos. Cada um desses poços tem uma tabela de VLP que modela o trecho vertical do poço.

Então, a cada poço é atrelada uma tabela de VLP (fixa de poço) e também um *manifold* imaginário, juntamente com uma tabela de VLP de superfície.

3.2.4. Dados de Modelagem utilizando o Software PROSPER

O PROSPER é um programa para modelagem da maioria dos tipos de configurações de poços encontrados na indústria de petróleo e gás em todo o mundo hoje. O PROSPER pode ajudar os engenheiros de reservatório e produção na previsão de tubos, tubulações hidráulicas e temperaturas com precisão e velocidade. O cálculo de sensibilidade do PROSPER permite que projetos de poços existentes sejam otimizados e os efeitos de futuras alterações nos parâmetros do sistema avaliados.

O software é projetado para permitir construção de modelos de poços confiáveis e consistentes, com a capacidade de lidar com cada aspecto do poço: PVT (caracterização de fluidos), as correlações de VLP (para cálculo de linha de fluxo e perda de pressão na tubulação) e IPR (entrada do reservatório). Ao modelar cada componente do sistema produtor do poço, o usuário pode verificar cada subsistema modelo por correspondência de desempenho. Uma vez que um modelo de sistema também foi ajustado aos dados reais de campo, o PROSPER

¹ O *manifold* não existe fisicamente, ele é somente criado para fazer a ligação entre as duas tabelas. Um tutorial de como foram incluídos os *manifolds* imaginários no IMEX está presente no Apêndice A.

pode ser seguramente usado para modelar o poço em diferentes cenários e fazer previsões de pressão do reservatório com base em dados de produção de superfície (Petex, 2012).

Os modelos gerados no PROSPER foram criados de acordo com o modelo do reservatório. Como já explicado, o modelo dos poços foi único e o mesmo é modelado da cabeça do poço até aproximadamente o topo do reservatório, sendo esse trecho sempre vertical.

Todos os poços produtores gerados são completados com válvulas de *gas lift* (gás é injetado na coluna de produção para auxiliar a elevação do fluido reduzindo a coluna hidrostática de fluido). Essa modelagem foi aplicada, pois ao aumentar o número de poços produtores, em um dado momento o BSW (*Basic Sediments and Water* – fração de água e sedimentos produzida comparada à produção total) de alguns poços aumenta e esses poços não conseguem mais produzir por surgência. A injeção de *gas lift* nos poços produtores foi fixada em 100.000 m³/dia. No PROSPER também foram modelados as linhas de produção (trecho horizontal) e o *riser* (trecho vertical).

3.3. Construção do Aproximador

O aproximador de função desenvolvido nesta dissertação visa a substituição de um sistema integrado de simulação de reservatório petrolífero e sistema de superfície. Sendo assim, a partir das mesmas entradas fornecidas ao simulador, nesse caso a localizações *I*, *J* e *K* iniciais e finais do poço, o aproximador fornece as curvas de produção anual de água, gás e óleo durante todo um período determinado, o que possibilita o cálculo do VPL (Valor Presente Líquido), utilizado como instrumento de avaliação no sistema de otimização.

A partir das entradas apresentadas, a rede realiza seu processamento e a saída obtida é comparada com a saída esperada. A partir do erro obtido, um processo de ajuste de pesos é aplicado, buscando-se minimizar o erro, conforme detalhado nas seções subsequentes. Como o aproximador deve fornecer as curvas para todo o período determinado, e como valores passados das próprias curvas são utilizados na entrada do modelo, o mesmo deve possuir a característica *multi-step*, onde os pontos previstos da série servem para prever pontos posteriores.

Deste modo, neste trabalho, optou-se por avaliar o desempenho da arquitetura *multilayer perceptron* com a seleção de padrões de treinamento através de técnicas de amostragens. O algoritmo de aprendizado utilizado foi o Levenberg-Marquardt (Saini and Soni, 2002). A escolha desse modelo se deve pela habilidade de aprendizado, a capacidade de generalização e a capacidade de resolver problemas de aproximação de funções contínuas (Haykin, 2001). Com isso, as redes neurais proporcionam as características necessárias para se construir um aproximador.

3.3.1. Aquisição dos Dados de Treinamento

Para a construção do aproximador, os dados foram obtidos usando os próprios simuladores que se quer substituir. Como o custo computacional da utilização dos simuladores integrados é alto, uma técnica que vise minimizar a quantidade de amostras é altamente recomendada. Deste modo, nesta dissertação utilizou-se o Hipercubo Latino, conforme será abordado mais adiante. Em casos em que se têm poços horizontais, o número de entradas se torna muito elevado, então também é utilizada a técnica de Análise de Componentes Principais para a redução de dimensionalidade.

3.3.2. Seleção de Variáveis de Treinamento e Criação da Base

Inicialmente, nos casos em que todos os poços são verticais, as entradas do aproximador são os valores das coordenadas I , J e K iniciais e a coordenada K final de cada poço - as coordenadas finais I e J não são entradas da rede. Isso porque nesses testes, na variação da localização os poços é imposta uma restrição para que todos os poços sejam verticais, o que faz com que as coordenadas I e J iniciais e finais sejam sempre as mesmas, o que as tornaria entradas redundantes para a rede neural.

Nos casos em que os poços podem ser verticais ou horizontais, as entradas do modelo são os valores das coordenadas I , J e K iniciais e finais de cada poço.

Além das entradas de localização dos poços, a rede neural tem também como entrada valores de vazão de óleo em três tempos anteriores ao que se quer prever. Para isso, é utilizada a técnica de janela deslizante de tamanho igual a três.

Também são entradas da rede neural os tipos de todos os poços, codificados da seguinte forma: entrada igual a 0, se o poço for do tipo injetor, entrada igual a 1, se o poço for do tipo produtor. Todas as entradas desse modelo inicial são mostradas na Figura 13.

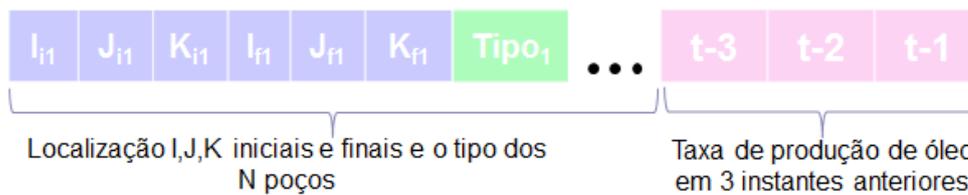


Figura 13: Entradas da rede neural

A base para a simulação de todos os testes foi obtida através da simulação dos modelos de reservatório apresentados, variando-se aleatoriamente (com distribuição uniforme) ou através de técnicas de amostragem (Hiper cubo Latino) a localização dos poços e os seus tipos, para que o aproximador possa se comportar como a simulação integrada para qualquer configuração de poços.

A saída aproximada pelo aproximador nos testes são valores de vazão de óleo total produzida pelo reservatório no decorrer de sua vida produtiva. Cada saída de uma simulação alimenta a entrada seguinte, já que os valores de vazão de óleo também são parâmetros de entrada. Por isso, a cada previsão, a rede é realimentada por valores previstos anteriormente por ela, para que seja previsto um próximo valor de vazão de óleo, como é exemplificado na Figura 14. Foram feitos testes modificando a entrada inicial da janela para que os mesmos valores não fiquem zerados, o que significa que tenho que ter um histórico de 3 anos para gerar os aproximadores.

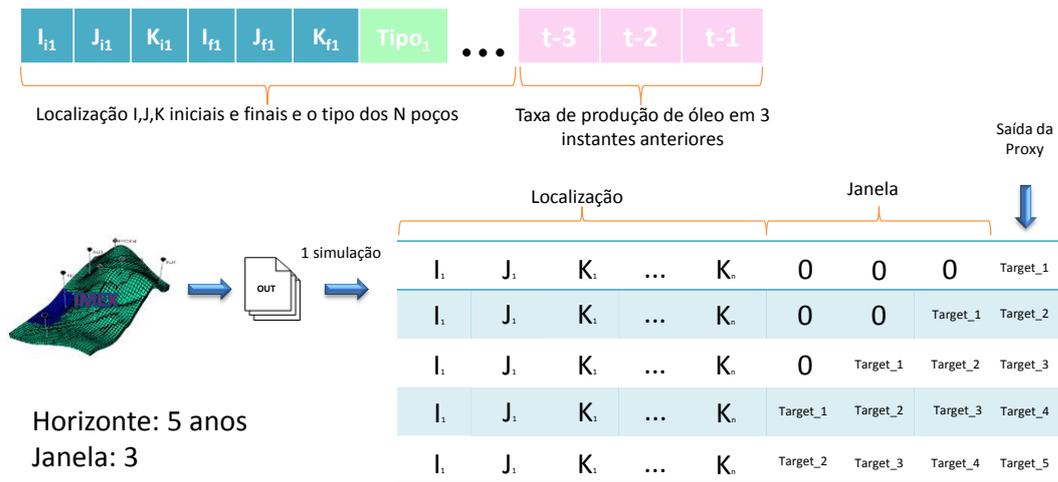


Figura 14: Janela deslizante para previsão de produção de óleo por 5 anos

Com a evolução dos testes, percebeu-se que os erros para o reservatório real – principalmente quando se utiliza poços horizontais e variam-se os tipos - se mostraram muito altos, então, a abordagem por Redes Neurais também foi modificada. Para os últimos testes, para aproximar o comportamento de um reservatório real onde são variadas as localizações e os tipos de, por exemplo, seis poços, são necessárias seis redes neurais que funciona da seguinte forma:

- Alternativas que tenham somente poços produtores, são aproximadas pela Rede Neural #1.
- Alternativas que tenham cinco poços produtores e um poço injetor são aproximadas pela Rede Neural #2.

E assim por diante para todas as alternativas possíveis, como demonstrado na Figura 15.

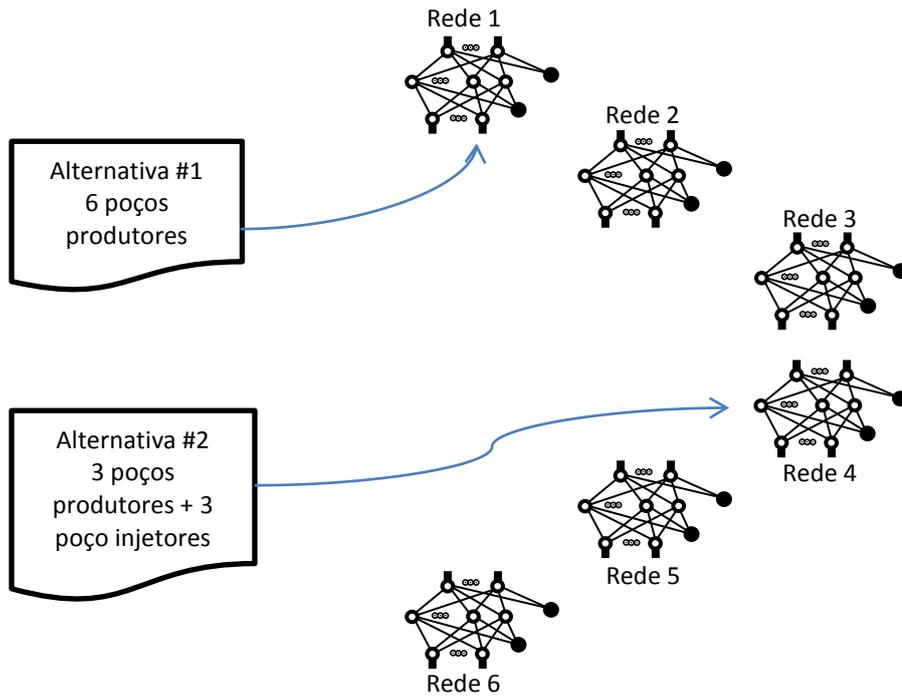


Figura 15: Esquema da distribuição das alternativas pelas Redes Neurais

Esse conjunto de redes prevê a produção de óleo. Para se ter também a previsão de produção de gás e de água, são necessárias mais doze redes, nesse caso: seis para a previsão de água e seis para a previsão de gás.

Nesse novo modelo, a entrada das redes passa a ser conforme a mostrada na Figura 16, que significa um aumento no número de redes para treinar, mas uma diminuição na quantidade de entradas das redes, o que resulta em erros menores, como será mostrado nos testes realizados.

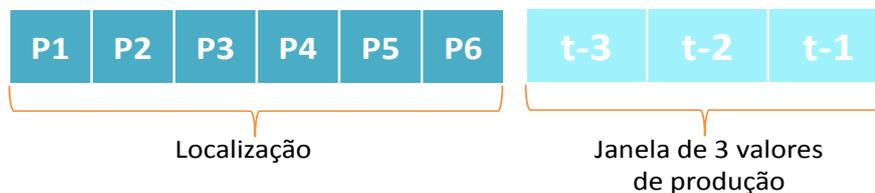


Figura 16: Entrada das redes neurais para novo modelo

A saída estimada pelo aproximador nos testes são valores de vazão de óleo como nos primeiros testes, e agora, nesse novo modelo, foram adicionadas também previsões da produção água e gás, para que se torne possível o cálculo do VPL, e a substituição dos simuladores em ferramentas de otimização de localização de poços onde se deseja maximizar o VPL.

Para os aproximadores de gás, as entradas seguem o mesmo padrão do aproximador de óleo: valores de I , J e K iniciais e finais e valores de vazão de gás em três tempos anteriores ao que se quer prever. O aproximador de água segue o mesmo padrão, porém necessita de mais uma entrada: a vazão de óleo em um instante anterior ao que se quer prever, que pode ser fornecido pelo aproximador de óleo. Essa metodologia foi adotada em função de testes já demonstrados em Almeida (2008).

3.3.3. Geração dos Cenários de Poços

A geração dos cenários de poços para posterior treinamento da rede foi feita de duas formas distintas:

- Aleatória com distribuição uniforme;
- Amostragem por Hipercubo Latino.

O fato de o Hipercubo Latino representar melhor a distribuição do espaço amostral do que a amostragem aleatória simples foi a principal motivação para aplicá-los na seleção dos valores utilizados na geração de configurações de poços, que inicialmente foram selecionadas aleatoriamente.

Todas as formas de geração de cenários de poços, porém, seguiram as mesmas metodologias e restrições:

1. O número total de poços sempre é fixo (seis poços).
2. Os tipos dos poços (produtor ou injetor) podem variar livremente nos modelos que possuem dezoito redes, e são fixos quando se utiliza somente três redes.
3. Os horizontes de produção são de 15, 20 e 25 anos, dependendo do teste.
4. São considerados dois tipos de testes em relação à geometria dos poços gerados: um teste utiliza somente poços verticais e o outro utiliza poços verticais e horizontais.
5. A lâmina d'água é de 1000 metros.
6. A distância mínima entre poços é de 200 metros.
7. O comprimento máximo dos poços é de 800 metros.

8. Não são consideradas alternativas que contenham somente poços injetores.
9. As simulações que incluem tabelas de fluxo multifásico têm válvulas de *gas lift* em todos os poços produtores, e a injeção de gás é fixa em todos os poços e igual a 100.000 m³/dia.
10. As restrições dos poços são as mesmas para cada tipo:
- a. No modelo sintético (primeiro modelo estudado que será apresentado no Capítulo 5):

i. Poços produtores:

Operate BHP min 55 Cont (*O poço opera normalmente até atingir a mínima pressão de fundo de poço igual a 55 kg/cm². Quando o poço viola essa restrição, a simulação continua, mas o poço muda para 'operando em modo de restrição violada'*)

Operate STO max 3000 Cont (*O poço opera normalmente até atingir a taxa de óleo na superfície igual a 300 m³/dia. Quando o poço viola essa restrição, a simulação continua, mas o poço muda para 'operando em modo de restrição violada'*)

Monitor WCUT 0.95 Shuttin (*Monitora a fração máxima de corte de água poço, que não deve passar de 95%. Quando o poço viola essa restrição, o poço é fechado*)

ii. Poços injetores:

Operate BHP max 450 Cont (*Restrição de máxima pressão de fundo de poço igual a 450 kg/cm²*)

Operate STW max 3000 Cont (*Restrição de máxima taxa de água na superfície igual a 3000 m³/dia*)

- b. No modelo baseado no campo de Namorado (segundo modelo estudado que será apresentado no Capítulo 5):

i. Poços produtores:

Operate BHP min 180 Cont (*Restrição de mínima pressão de fundo de poço igual a 180 kg/cm²*)

Operate STL max 1800 Cont (*Restrição de máxima taxa total de líquido (óleo + água) na superfície igual a 1800 m³/dia*)

ii. Poços injetores:

Operate BHP max 350 Cont (*Restrição de máxima pressão de fundo de poço igual a 350 kg/cm²*)

Operate STW max 3000 Cont (*Restrição de máxima taxa de água na superfície igual a 3000 m³/dia*)

3.3.4. Transformação dos Dados de Treinamento

De modo geral, a transformação de dados envolve a aplicação de alguma técnica de transformação aos valores das variáveis selecionadas, buscando obter os dados de forma mais apropriada para sua utilização na modelagem dos aproximadores, maximizando a informação. Entre as transformações mais comuns, está a normalização, que objetiva homogeneizar a variabilidade das variáveis que fazem parte do modelo.

Neste trabalho as variáveis foram normalizadas pela faixa de variação, conforme a Equação (3.1):

$$X_N = \frac{X_R - X_{min}}{X_{max} - X_{min}} \quad (3.1)$$

Onde:

- X_R : valor original para a variável
- X_N : valor normalizado para a variável
- X_{max} : maior valor para a variável
- X_{min} : menor valor para a variável

3.3.5. Redução de Dimensionalidade utilizando Análise de Componentes Principais

Para os casos em que o número de entradas aumenta muito, como por exemplo, o caso em que os poços são horizontais e verticais, o que resulta em um número de entradas igual a 45 para um cenário com seis poços (localização I, J, K

iniciais e finais e tipos de cada um dos seis poços mais a janela de produção em três instantes anteriores), é utilizada a técnica de redução de dimensionalidade chamada Análise de Componentes Principais. O objetivo é reduzir o número de entradas da rede neural e conseqüentemente obter resultados melhores de previsão de produção.

4. Estudo de Casos

4.1. Parâmetros da Rede Neural

Nos testes com Redes Neurais Artificiais, a topologia escolhida foi de apenas uma camada escondida, de tamanho variável, e uma única saída. A escolha do número de neurônios na camada escondida foi realizada através do conjunto de validação, buscando a configuração que forneça o melhor desempenho de generalização (*early stopping*), como descrito no Capítulo 2. Deste modo, para cada configuração de neurônios na camada escondida, quatro processos de treinamento, com inicialização aleatória dos pesos, são realizados, e escolhe-se a melhor rede, entre as quatro avaliadas, pelo desempenho no conjunto de validação. Os parâmetros básicos dos modelos de redes neurais avaliados neste estudo são apresentados na Tabela 1:

Parâmetros das Redes Neurais	Valor
Tipo	MLP
Topologia	Uma única camada escondida
Neurônios (camada escondida)	De 5 a 30 neurônios
Algoritmo de Aprendizado	Levenberg-Marquardt (trainlm)
Nº máximo de épocas	2000

*Critério de parada definido pelo erro do conjunto de validação ou pelo número máximo de épocas

Tabela 1: Parâmetros das Redes Neurais Artificiais

4.2. Métricas de Avaliação

Diversos indicadores foram desenvolvidos para se comparar os desempenhos dos modelos. Tais ferramentas estatísticas têm valor ao permitir o julgamento isento das vantagens dos modelos. Existem indicadores de precisão,

de adequação relativa e que tentam capturar a capacidade de acerto de tendência dos modelos em questão (Diebold, 1996).

Neste trabalho, para a avaliação do desempenho dos modelos, foram utilizadas as métricas de erro MAPE (*Mean Absolute Percentage Error*), RMSE (*Root Mean Square Error*) e SMAPE (*Symmetric Mean Absolute Percentage Error*).

O MAPE tenta capturar a importância do erro relativo, fornecendo um valor percentual, enquanto o RMSE fornece o erro na dimensão da variável e mede a magnitude média do erro. Os cálculos do MAPE e do RMSE são dados pelas Equações (4.1) e (4.2), respectivamente. No RMSE, uma vez que os erros são elevados ao quadrado, antes de ser calculada a média, a métrica RMSE atribui um peso relativamente elevado a erros maiores. Isto significa que essa métrica de erro é mais útil quando erros maiores são particularmente indesejáveis.

$$\text{MAPE} = \frac{\sum_{t=1}^N \frac{\text{real} - \text{previsto}}{\text{real}}}{N} \quad (4.1)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^N (\text{real} - \text{previsto})^2}{N}} \quad (4.2)$$

Onde:

real:	valor dado pelos simuladores
previsto:	valor dado pelos aproximadores
N:	número total de amostras

Como em alguns casos a produção chega a zero, e/ou valores muito próximos a zero, o valor MAPE acaba sendo muito alto e se torna uma medida inadequada para a estimação do modelo. Então foi adotada também outra medida: o SMAPE, dado pela Equação (4.3). Essa métrica de erro, medida em pontos percentuais, é uma métrica amplamente utilizada em trabalhos relacionados a séries temporais, principalmente por seu emprego em competições para modelos de previsão (Makridakis & Hibon, 2000; NN3, 2011).

$$\text{SMAPE} = \frac{1}{N} \sum_{t=1}^N \frac{|\text{previsto}_t - \text{real}_t|}{(\text{real}_t + \text{previsto}_t)/2} \quad (4.3)$$

4.3. Descrição dos Reservatórios

4.3.1. Reservatório Sintético

▪ Tutorial_01

Um dos modelos de reservatório utilizado nessa etapa foi um modelo sintético que contém no máximo 6 poços (podendo variar livremente entre poços produtores e poços injetores), como mostra a Figura 17. A geometria *corner point*, possui 33x57x3 blocos, nas direções *I*, *J* e *K* respectivamente (5.643 células ativas), com canais de alta permeabilidade e um pequeno aquífero pouco atuante.

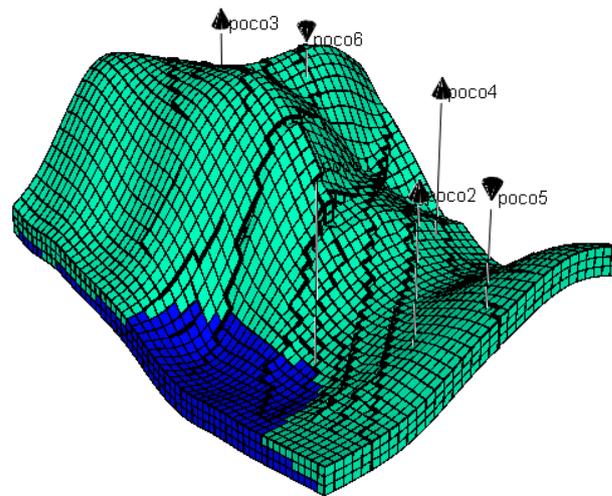


Figura 17: Reservatório sintético

4.3.2. Reservatório Real

▪ Campo Baseado em Namorado

O outro modelo de reservatório utilizado nessa etapa foi um modelo baseado em um caso real, o campo de Namorado, que contém no máximo 6 poços (podendo variar livremente entre poços produtores e poços injetores), como mostra a Figura 18. Trata-se de um reservatório de arenito consolidado,

representado por uma malha do tipo *corner point*, com 83x45x23 blocos nas direções *I, J, K* (31.486 ativas). A permeabilidade e a porosidade do reservatório são heterogêneas em toda a sua extensão e a saturação de água gira em torno de 0,25 em todo o reservatório.

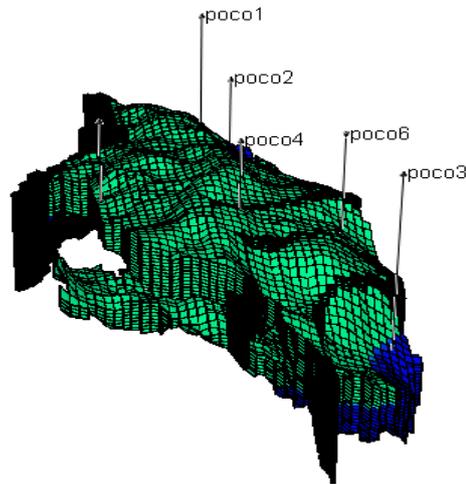


Figura 18: Reservatório real

4.4. Avaliação dos Aproximadores

Nesta seção são apresentados todos aproximadores desenvolvidos com os seus respectivos desempenhos. A Tabela 2 apresenta a estrutura desta seção e as principais características dos testes realizados. Na tabela, a coluna ‘Tipos’ diz respeito ao tipo do poço, que pode ser injetor ou produtor. Os tipos são considerados fixos quando é definida a quantidade de cada tipo de poço previamente. Os tipos são variáveis quando se pode ter qualquer quantidade de poços injetores e produtores.

Teste	Reservatório	Tipos	Redes	Amostras	Produção (anos)	Poços
Teste 1	Sintético	Fixos	1	Aleatório	15	Verticais
Teste 2	Sintético	Variáveis	1	Aleatório	15	Verticais
Teste 3	Sintético	Fixos	1	Aleatório	25	Verticais
Teste 4	Sintético	Variáveis	1	Aleatório	25	Verticais
Teste 5	Sintético	Fixos	1	Aleatório	15	Horizontais

Teste 6	Sintético	Variáveis	1	Aleatório	15	Horizontais
Teste 7	Sintético	Fixos	1	Aleatório	25	Horizontais
Teste 8	Sintético	Variáveis	1	Aleatório	25	Horizontais
Teste 9	Real	Variáveis	1	Aleatório	20	Verticais
Teste 10	Real	Variáveis	6x3	Aleatório	20	Verticais
Teste 11	Real	Variáveis	6x3	Hipercubo Latino	20	Verticais
Teste 12	Real	Variáveis	6x3	Aleatório	20	Horizontais
Teste 13	Real	Variáveis	6x3	Aleatório com PCA	20	Horizontais

Tabela 2: Estrutura de apresentação dos resultados para os aproximadores

Os testes 1 ao 8 são feitos com aproximadores considerando apenas o simulador de reservatório, desconsiderando o simulador de superfície. A partir do teste 9, todos os aproximadores são construídos a partir da integração dos simuladores de reservatório e de superfície.

A construção dos aproximadores foi realizada em sete microcomputadores com a configuração básica descrita na Tabela 3:

Configuração dos microcomputadores	
Processador	Intel Core i7-2600
Sistema Operacional	Windows 7 Professional 64 bits
Memória RAM	16 GB

Tabela 3: Configuração dos microcomputadores utilizados

4.4.1. Testes 1 e 2

Os primeiros testes foram realizados para o cenário do reservatório sintético com seis poços, todos eles verticais, e o treinamento foi conduzido a partir de um conjunto de amostras obtido de forma aleatória. Este teste foi dividido em duas partes: a primeira utilizando tipos de poços fixos, isto é, a localização dos poços muda, mas sempre se tem um cenário com 4 poços produtores e 2 poços injetores, a segunda parte os poços podem variar livremente entre produtor e injetor, o que significa mais entradas para a rede neural (cada poço representará uma nova entrada para a rede, totalizando seis novas entradas que codificam o tipo de cada

poço). O objetivo é verificar a diferença no erro para tipos fixos e variantes. Para ambos os casos foram geradas 8 mil simulações divididas em três conjuntos: 70% para treinamento, 20% para validação e 10% para teste. Para ambos os casos, o horizonte de produção é de 15 anos.

Os melhores resultados são apresentados na Tabela 4:

Poços Verticais	4 produtores e 2 injetores		6 poços (tipos variando)		
	Neurônios	RMSE	MAPE (%)	RMSE	MAPE (%)
5		116,37	5,10	257,87	27,20
6		119,52	4,99	256,31	22,82
7		112,00	4,74	246,94	19,61
8		114,96	4,89	244,96	22,58
9		134,93	6,04	246,82	21,48
10		105,44	3,68	253,80	21,44
11		128,27	4,87	227,07	19,57
12		115,85	3,84	259,94	20,44
13		143,14	4,65	246,53	21,82
14		133,04	5,59	239,38	20,51
15		149,44	5,75	255,69	22,44
16		143,79	4,29	282,47	23,26
17		125,91	4,69	253,15	22,29
18		140,46	5,26	228,09	17,55
19		186,84	7,70	278,66	20,89
20		156,98	4,65	239,31	19,29
21		151,75	5,37	261,56	22,15
22		177,82	6,83	265,55	21,91
23		163,47	5,23	256,48	20,29
24		138,18	4,79	294,99	22,66
25		197,34	6,63	311,20	22,57

Tabela 4: Testes 1 e 2 para o reservatório sintético

Pode-se perceber, ao analisar a Tabela 4 que há um significativo aumento no erro de previsão ao variar aleatoriamente os tipos de cada um dos seis poços,

mantendo o mesmo número de simulações para treinamento, validação e teste. Isso pode ser explicado pelo aumento no número de entradas da rede neural.

4.4.2. Testes 3 e 4

Os testes 3 e 4 são feitos como os testes 1 e 2, respectivamente. A diferença aqui é o horizonte de produção. O objetivo desse teste é analisar a diferença no erro de previsão variando o horizonte de previsão. Logo, nesses casos, o horizonte foi aumentado de 15 para 25 anos. Os resultados desses testes são apresentados na Tabela 5:

Poços Verticais	4 produtores e 2 injetores		6 poços (tipos variando)	
	Neurônios	RMSE	MAPE (%)	RMSE
5	105,66	4,97	205,93	25,30
6	303,08	11,32	239,38	27,46
7	141,84	7,15	213,95	23,35
8	98,17	4,66	220,27	32,04
9	88,93	4,35	183,42	28,05
10	107,92	5,37	182,44	23,30
11	136,47	5,90	214,69	27,75
12	132,61	7,17	208,67	26,45
13	112,11	5,65	202,92	26,43
14	145,33	6,59	238,82	33,93
15	90,75	4,88	191,80	23,20
16	84,90	4,18	293,73	36,76
17	104,61	5,11	244,22	30,56
18	101,04	4,15	229,59	31,01
19	132,68	6,53	22,25	27,23
20	266,16	5,83	209,87	25,20
21	212,39	5,15	264,72	33,24
22	158,36	7,56	219,85	24,80
23	151,52	7,58	596,78	53,04
24	133,93	6,62	241,88	27,74

25	175,24	7,24	201,41	22,48
----	--------	------	--------	-------

Tabela 5: Testes 3 e 4 para o reservatório sintético

O que se pode perceber ao analisar conjuntamente a Tabela 4 e a Tabela 5 é que o aumento no horizonte de produção não aumenta o erro da previsão, pois os erros se mantêm bem próximos.

A Tabela 6 apresenta uma compilação dos melhores resultados obtidos até aqui. Percebe-se um aumento no erro com a variação dos tipos e com o aumento do horizonte de produção. Na coluna tempo é mostrado o tempo requerido para se realizar todas as simulações somadas ao tempo de treinamento dos aproximadores.

Melhores Resultados	MAPE (%)	Horizonte	Tipos dos poços	Nº de Simulações	Tempo (horas)
Teste 1	3,68	15 anos	Fixos	8 mil	~42,6
Teste 2	17,55	15 anos	Variáveis	8 mil	~49,3
Teste 3	4,15	25 anos	Fixos	8 mil	~90
Teste 4	23,20	25 anos	Variáveis	8 mil	~95

Tabela 6: Melhores resultados dos testes 1, 2, 3 e 4.

4.4.3. Testes 5 e 6

Os testes 5 e 6 foram realizados para o cenário do reservatório sintético com seis poços, podendo ser verticais ou horizontais, e o treinamento foi conduzido a partir de um conjunto de amostras obtido de forma aleatória. Este teste foi dividido em duas partes: a primeira utilizando tipos de poços fixos, isto é, a localização dos poços muda, mas sempre se tem um cenário com 4 poços produtores e 2 poços injetores, a segunda parte os poços podem variar livremente entre produtor e injetor, o que significa mais entradas para a rede neural. O objetivo é verificar a diferença no erro para tipos fixos e variantes. Para ambos os casos foram geradas 15 mil simulações divididas em três conjuntos: 70% para treinamento, 20% para validação e 10% para teste. Para ambos os casos, o horizonte de produção é de 15 anos.

Os melhores resultados são apresentados na Tabela 7:

Poços Horizontais	4 produtores e 2 injetores		6 poços (tipos variando)	
	Neurônios	RMSE	MAPE (%)	RMSE
5	176,14	11,92	304,33	16,93
6	153,12	9,10	301,88	16,78
7	216,33	7,32	358,93	18,16
8	214,08	7,28	383,68	25,71
9	264,84	10,93	412,16	22,98
10	186,72	6,60	329,20	18,55
11	307,11	12,44	525,19	18,16
12	294,84	9,77	224,58	21,91
13	273,36	9,26	335,17	17,27
14	225,94	8,69	353,94	22,85
15	277,19	10,67	374,71	25,39
16	248,85	10,85	311,75	19,26
17	201,12	7,18	338,34	20,00
18	358,19	10,45	330,28	20,33
19	212,95	8,56	385,36	25,30
20	228,01	8,05	351,76	20,24
21	258,22	8,94	374,92	21,13
22	270,00	10,41	374,46	22,16
23	228,19	8,82	386,2	26,29
24	266,36	9,10	433,75	23,98
25	270,36	9,35	361,38	21,86

Tabela 7: Testes 5 e 6 para o reservatório sintético

Nesse teste é possível perceber que, assim como nos testes 1 e 2, há um significativo aumento no erro de previsão ao variar aleatoriamente os tipos de cada um dos seis poços mantendo o mesmo número de simulações para treinamento, validação e teste, o que pode ser explicado pelo aumento no número de entradas da rede neural, como mencionado anteriormente. Comparando os testes 1 e 2 com os testes 5 e 6, onde a diferença é que nos últimos têm-se tanto poços horizontais quanto verticais, percebe-se uma proximidade nos erros (erros

na mesma ordem de grandeza), que pode ser explicada pelo aumento do número de simulações para que os erros possam acompanhar o aumento no número de entradas do modelo para representar a localização dos poços. Como já explicado no Capítulo 3, a utilização de poços horizontais aumenta o número de entradas da rede. Cada poço no reservatório representa um aumento de 2 entradas para rede, resultando um total de 12 novas entradas. Uma comparação entre os melhores resultados dos testes 1, 2, 5 e 6 é feita na Tabela 8.

Melhores Resultados	MAPE (%)	Poços	Tipos dos poços	Nº de Simulações	Tempo (horas)
Teste 1	3,68	Verticais	Fixos	8 mil	~42,6
Teste 2	17,55	Verticais	Variáveis	8 mil	~49,3
Teste 5	6,6	Horizontais	Fixos	15 mil	~142,1
Teste 6	16,78	Horizontais	Variáveis	15 mil	~144,3

Tabela 8: Melhores resultados dos testes 1, 2, 5 e 6.

4.4.4. Testes 7 e 8

Os testes 7 e 8 são feitos como os testes 5 e 6, respectivamente. A diferença aqui é o horizonte de produção. O objetivo desse teste é analisar a diferença no erro de previsão variando o horizonte de previsão para os casos em que poços horizontais são utilizados. Logo, nesses casos, o horizonte foi aumentado de 15 para 25 anos. Os resultados desses testes são apresentados na Tabela 9:

Poços Horizontais	4 produtores e 2 injetores		6 poços (tipos variando)	
Neurônios	RMSE	MAPE (%)	RMSE	MAPE (%)
5	228,38	12,88	903,98	95,03
6	213,43	13,78	273,78	73,16
7	209,99	12,96	240,99	60,57
8	257,76	15,72	193,06	89,01
9	276,63	18,22	341,51	68,55
10	243,22	14,74	331,64	95,46
11	246,35	17,96	362,42	82,18

12	244,35	17,19	470,73	95,61
13	216,16	13,82	293,54	52,54
14	206,45	13,14	317,65	83,96
15	209,81	13,96	272,39	85,90
16	253,77	15,86	308,66	61,80
17	268,4	17,44	302,62	80,56
18	224,17	14,13	331,11	95,76
19	210,99	12,37	331,38	44,41
20	228,57	15,99	360,4	52,06
21	303,31	21,04	277,35	34,52
22	318,19	19,92	419,49	53,42
23	387,83	19,89	762,55	98,01
24	269,37	15,39	310,79	81,12
25	259,79	20,05	417,08	96,32

Tabela 9: Testes 7 e 8 para reservatório sintético

O que se pode perceber ao analisar conjuntamente a Tabela 7 e a Tabela 9 é que o aumento no horizonte de produção aumenta o erro da previsão, como se espera intuitivamente, e esse aumento no erro já é considerável nesse caso.

A Tabela 10 apresenta uma compilação dos melhores resultados alcançados até aqui. Nota-se um aumento no erro com a variação dos tipos e com o aumento do horizonte de produção:

Melhores Resultados	MAPE (%)	Horizonte	Tipos dos poços	Tempo (horas)
Teste 5	6,60	15 anos	Tipos fixos	~142,1
Teste 6	16,78	15 anos	Tipos variáveis	~144,3
Teste 7	12,37	25 anos	Tipos fixos	~269,3
Teste 8	44,41	25 anos	Tipos variáveis	~275

Tabela 10: Melhores resultados dos testes 5, 6, 7 e 8.

Esse aumento considerável no erro de aproximação não é desejável, o que provocou uma modificação na abordagem de construção do aproximador, como é feito do teste 10 em diante.

4.4.5. Testes 9 e 10

A partir do teste 9, é utilizado somente o reservatório com características reais inspirado no campo de Namorado. Também a partir desse teste, todos os aproximadores são da simulação integrada de simulador de reservatório IMEX e simulador de superfície (linhas de produção e *riser*) e subsuperfície (poço) PROSPER.

O teste número 9 foi realizado para o cenário do reservatório real com seis poços, todos eles verticais, e o treinamento foi conduzido a partir de um conjunto de amostras obtido de forma aleatória. Neste teste, os tipos e as localizações dos poços são variados como nos testes anteriores.

No teste 10, a abordagem do aproximador é modificada em função do aumento no erro de previsão observado nos testes 8 e 9. Nessa nova abordagem, cada aproximador passa a ter seis redes neurais para cada variável de saída do campo que se deseja analisar. Como se está avaliando três variáveis de saída (taxa de óleo, gás e água), um aproximador do sistema integrado é constituído por 18 redes neurais. O uso de cada rede é definido no Capítulo 3.

A partir desse modelo também foi empregado uma nova métrica de erro: SMAPE, já que o MAPE não é adequado para valores iguais ou próximos de zero, como explicado anteriormente.

O objetivo é verificar a diferença no erro para aproximadores com uma ou mais redes neurais. Para o teste 9 foram geradas 20 mil simulações divididas em três conjuntos: 70% para treinamento, 20% para validação e 10% para teste. Para o teste 10 são geradas 12 mil simulações para cada grupo de 3 redes neurais (resultando em um total de 72 mil simulações). Para ambos os casos, o horizonte de produção é de 20 anos.

Os melhores resultados são apresentados na Tabela 11:

Óleo	Testes	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
SMAPE (%)	Teste 9	26,96 (21)	-	-	-	-	-
	Teste 10	6,41 (25)	6,74 (16)	6,03 (25)	8,95 (18)	10,35 (21)	12,35 (30)
RMSE	Teste 9	235,85					

Teste 10	73,29	111,11	173,59	240,16	279,61	130,96
----------	-------	--------	--------	--------	--------	--------

Tabela 11: Melhores resultados dos testes 9 e 10 para previsão de óleo

Os valores entre parênteses representam o número de neurônios na camada escondida da rede que melhor aproximou o sistema integrado.

Com esse teste, pode-se perceber o ganho significativo obtido com o uso de seis redes neurais. Essa abordagem é utilizada em todos os testes a partir de agora.

A Tabela 12 e a Tabela 13 mostram os melhores resultados para o aproximador de gás e de água, respectivamente.

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Gás	SMAPE (%)	6,05 (16)	5,81 (16)	6,22 (24)	8,79 (22)	7,39 (17)	22,88(21)
	RMSE	10,67	16,06	31,92	25,86	27,53	19,02

Tabela 12: Melhores resultados do teste 10 para previsão de gás

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Água	SMAPE (%)	21,77(22)	8,05(26)	10,66(19)	17,29(17)	14,05(16)	21,77(22)
	RMSE	113,7	186,90	185,25	116,54	108,36	113,71

Tabela 13: Melhores resultados do teste 10 para previsão de água

4.4.6. Teste 11

No teste 11 é utilizada a técnica de amostragem Hipercubo Latino. O objetivo é reduzir o número de amostras necessárias para a aproximação do modelo, já que a necessidade de um número alto de simulações para o treinamento pode tornar a construção de um aproximador cara e inviável. O ideal é que o custo computacional de construção do aproximador não seja muito maior que o custo de executar todas as simulações necessárias em um processo de otimização.

A Tabela 14, a Tabela 15 e a Tabela 16 mostram os melhores resultados para o aproximador de gás e de água, respectivamente.

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Óleo	SMAPE						
	(%)	6,32 (15)	6,46 (23)	5,70 (29)	12,27(15)	20,96(25)	19,65(13)
	RMSE	150,14	209,48	476,89	275,91	309,44	284,97

Tabela 14: Melhores resultados do teste 11 para previsão de óleo

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Gás	SMAPE						
	(%)	17,69(27)	24,87(20)	19,79(28)	17,05(19)	18,30(19)	12,98(18)
	RMSE	36,16	31	91,04	29,72	31,56	32,81

Tabela 15: Melhores resultados do teste 11 para previsão de gás

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Água	SMAPE						
	(%)	15,94(12)	19,86(25)	18,24 (8)	20,77(23)	19,54(17)	20,23(12)
	RMSE	230,3	133,27	135,12	231,35	205,47	159,88

Tabela 16: Melhores resultados do teste 11 para previsão de água

Para o teste 11 foram geradas 5 mil simulações com o uso da técnica do Hipercubo Latino (para cada grupo de 3 redes neurais, resultando em um total de 30 mil simulações) divididas em três conjuntos: 80% para treinamento, 20% para validação. Para o conjunto de teste foram simuladas 500 alternativas (para cada rede neural) que tiveram suas localizações variadas aleatoriamente. O horizonte de produção continua sendo de 20 anos.

Comparando-se a geração de alternativas de forma aleatória (teste 10) e por Hipercubo Latino (teste 11), percebe-se uma diminuição significativa no tempo de simulação já que o número de simulações reduziu de 12 mil para 5 mil para cada grupo de 3 redes neurais e os erros se mantiveram bem próximos.

4.4.7. Testes 12 e 13

No teste 12 é utilizada a mesma abordagem do teste 10, onde são utilizadas seis redes para cada aproximador e o horizonte de previsão é de 20 anos e as amostras são obtidas de forma aleatória. A diferença é que nesse teste, os poços podem variar sua geometria livremente, respeitando algumas restrições já explicadas no Capítulo 3, como comprimento máximo de 800 metros e distância mínima entre poços de 200 metros.

O objetivo é verificar a viabilidade da criação de aproximadores quando se tem poços horizontais em um reservatório real e comparar essas abordagens.

O teste 13 tem as mesmas características do teste 12, porém, enquanto no teste 12 os valores I , J , K iniciais e finais são entradas da rede, no teste 13 é utilizada a técnica de Análise de Componentes Principais para a definição das entradas do modelo. Com isso, é possível reduzir o número de entradas necessárias, escolhendo somente um subconjunto dos atributos, que retém o máximo de informação.

Para os testes 12 e 13 foram geradas 15 mil simulações (resultando em um total de 90 mil simulações), divididas em três conjuntos: 70% para treinamento, 20% para validação e 10% para teste.

Os melhores resultados para o teste 12 são apresentados na Tabela 17, Tabela 18 e Tabela 19:

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Óleo	SMAPE	11,74	16,97	13,56	18,94	17,20	18,01
	(%)	(24)	(30)	(18)	(28)	(21)	(24)
	RMSE	208,72	243,41	192,63	290,85	295,51	170,73

Tabela 17: Melhores resultados do teste 12 para previsão de óleo

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Gás	SMAPE	16,07	17,21	14,79	18,88	7,95	13,94
	(%)	(21)	(18)	(24)	(30)	(22)	(19)
	RMSE	40,81	32,00	27,34	32,75	29,26	21,78

Tabela 18: Melhores resultados do teste 12 para previsão de gás

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Água	SMAPE	14,85	15,91	15,39	13,29	18,45	16,09
	(%)	(28)	(21)	(27)	(21)	(27)	(20)
	RMSE	115,29	159,33	199,18	270,97	169,74	187,17

Tabela 19: Melhores resultados do teste 12 para previsão de água

O aumento do erro não é significativo quando comparado ao do aproximador que considera somente poços verticais (teste 10). O pequeno aumento no erro deve-se ao aumento no número de entradas e, conseqüente, na dimensão do problema. O fato do aumento no erro ter sido pequeno se deve, em parte, pelo aumento no número de simulações para treinamento e validação no modelo.

Os melhores resultados para o teste 13, onde se diminuiu o número de entradas do modelo por meio da Análise de Componentes Principais, são apresentados na Tabela 20, Tabela 21 e Tabela 22:

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Óleo	SMAPE	19,36	18,32	13,84	17,98	7,12	8,94
	(%)	(17)	(18)	(25)	(20)	(11)	(13)
	RMSE	297,11	252,09	209,71	245,1	81,0	158,55

Tabela 20: Melhores resultados do teste 13 para previsão de óleo

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Gás	SMAPE	12,37	17,48	14,37	17,68	18,09	24,36
	(%)	(14)	(21)	(10)	(14)	(19)	(20)
	RMSE	30,82	30,43	26,20	30,39	32,66	20,15

Tabela 21: Melhores resultados do teste 13 para previsão de gás

Melhores Resultados	Erro	Rede 1	Rede 2	Rede 3	Rede 4	Rede 5	Rede 6
Água	SMAPE	12,38	16,93	14,65	9,44	16,29	11,65
	(%)	(27)	(22)	(19)	(15)	(23)	(15)
	RMSE	132,69	262,68	282,99	360,31	339,51	233,17

Tabela 22: Melhores resultados do teste 13 para previsão de água

A partir de uma tabela com todas as variáveis do problema, a técnica de Análise de Componentes Principais gera uma nova tabela com o mesmo número de variáveis, os chamados componentes principais. Cada componente principal é uma transformação linear de todo o conjunto de dados originais. Os coeficientes dos componentes principais são calculados de modo que o primeiro componente principal contém a variância máxima (que pode ser entendida como o "máximo de informação"). O segundo componente principal é calculado para ter a segunda maior variância, e, mais importante, não é correlacionado com o primeiro componente principal.

Analisando a Tabela 23, percebe-se que de fato a maior variância foi “colocada” no primeiro componente principal.

Variância das 36 entradas da rede				
6,166656	4,494104	3,904147	3,619027	3,161782
2,526182	2,28136	1,904756	1,801439	1,684057
1,412886	1,306203	1,104303	0,932319	0,813953
0,601887	0,500961	0,393865	0,300662	0,017977
0,014888	0,008923	0,00845	0,007062	0,006211
0,006033	0,003512	0,002964	0,002793	0,002602
0,00215	0,001893	0,001378	0,001048	0,00036
0,000349	-	-	-	-

Tabela 23: Análise de variância das variáveis

A variância acumulada nos n primeiros componentes principais também pode ser rapidamente obtida e analisada. Os valores para o caso em questão são mostrados na Tabela 24:

Variância acumulada				
0,158119	0,273353	0,373459	0,466255	0,547326
0,6121	0,670596	0,719436	0,765627	0,808808
0,845036	0,878528	0,906844	0,930749	0,95162
0,967053	0,979898	0,989997	0,997706	0,998167
0,998549	0,998778	0,998995	0,999176	0,999335
0,99949	0,99958	0,999656	0,999727	0,999794

0,999849	0,999898	0,999933	0,99996	0,999969
1	-	-	-	-

Tabela 24: Análise da variância acumulada dos n primeiros componentes principais

Interpreta-se a tabela acima da seguinte forma: o primeiro componente principal contém 15,81% da variância total dos dados originais; os dois primeiros componentes principais juntos contêm 27,33% da variância total dos dados originais. Assim sendo, escolhendo os dezoito primeiros componentes principais, tem-se 99% da variância total dos dados originais.

Dessa forma, diminuiu-se o número de entradas de trinta e nove (trinta e seis de localização e três entradas da janela de produção) para dezoito perdendo-se apenas 1% da variância dos dados.

O tempo dispendido para a construção de cada uma dos aproximadores (tempo de simulação somado ao tempo de treinamento dos aproximadores) está indicado na Tabela 25. Como já era esperado, percebe-se que há um ganho no tempo de construção dos aproximadores quando se utiliza Hipercubo Latino e Análise de Componentes Principais.

Melhores Resultados	Amostras	Poços	Tempo (horas)
Teste 9	Aleatórias	Verticais	~767,8
Teste 10	Aleatórias	Verticais	~2160
Teste 11	Hipercubo Latino	Verticais	~920
Teste 12	Aleatórias	Horizontais	~3360
Teste 13	Aleatórias / Análise de Componentes Principais	Horizontais	~3264

Tabela 25: Tabela do tempo de construção dos aproximadores

A utilização da técnica de Hipercubo Latino diminui tanto o tempo de simulação, já que diminui o número de simulações necessárias para a construção do aproximador, quanto o tempo de treinamento da rede, já que o tamanho da base influencia no tempo de treinamento da rede. A utilização da técnica de Análise de Componentes Principais não diminui o tempo de simulação (a base é a mesma do teste sem Análise de Componentes Principais), mas diminui o tempo de treinamento do aproximador reduzindo o número de variáveis de entrada do

mesmo. Logo, uma abordagem para reduzir ainda mais o tempo para construção dos aproximadores seria utilizar as duas técnicas conjuntamente.

Apesar do tempo necessário para a criação de um aproximador ser maior que o tempo necessário para a realização de uma otimização, a criação de aproximadores ainda vale a pena, já que uma vez construído, é possível realizar otimizações substituindo os simuladores sempre que for necessário. O ganho de velocidade na simulação viabiliza análises mais detalhadas do comportamento do processo produtivo como um todo. Em processos que necessitem de diversas otimizações, como em análises de incerteza, os aproximadores serão extremamente econômicos em termos de custo computacional, tornando seu custo inicial razoável. Como exemplo, na Tabela 26 é apresentado uma comparação entre os tempos de construção de um aproximador e de uma otimização.

		Simulações	Poços	Tempo (horas)
Construção do Aproximador	Hipercubo Latino	30 mil	Verticais	~920
Otimização	Algoritmo Genético	8 mil	Verticais	~213,3

Tabela 26: Comparação entre tempo de construção de um aproximador e uma otimização

Levando em conta que todos os tempos apresentados aqui são somados como se a construção dos aproximadores fosse feita em somente uma máquina e de forma serial, esse tempo pode diminuir ainda mais distribuindo-se as simulações necessárias e o treinamento da redes por várias máquinas em um *cluster*.

4.5. Avaliação dos Aproximadores no Sistema de Otimização de Localização de Poços

Nesta seção são apresentados os resultados relacionados ao desempenho dos aproximadores substituindo os simuladores no sistema de otimização apresentado em Emerick (2009). O sistema OCTOPUS é o resultado de uma parceria de pesquisa e desenvolvimento, entre a Petrobras e o Laboratório ICA. Trata-se de um sistema desenvolvido especificamente para a solução do problema de otimização de planos de drenagem. A partir da descrição de um reservatório, o

sistema busca encontrar, por meio de um algoritmo genético, uma configuração de poços que maximize o VPL do projeto. No processo de busca pela melhor configuração é levado em consideração a quantidade, o tipo (injetor ou produtor), a trajetória (vertical, horizontal ou direcional) e a localização dos poços.

O sistema utiliza o simulador IMEX para realizar inúmeras simulações para encontrar uma alternativa ótima de poços, o que decorre em um alto custo computacional. É então através do uso de aproximadores, que se propõe reduzir esse custo computacional. A substituição parcial ou total dos aproximadores pelo simulador proporcionaria ao sistema OCTOPUS uma ferramenta útil para acelerar o processo de otimização.

Os testes apresentados a seguir foram realizados exclusivamente para o reservatório baseado no campo de Namorado com poços verticais. Para isto, foram utilizados os aproximadores do teste 11 construídos para a predição das curvas de óleo, gás e água, onde a amostragem foi realizada pela técnica do Hipercubo Latino e a abordagem usada foi a que utiliza 18 redes neurais (6 redes neurais para previsão de cada uma das 3 curvas de produção).

Para avaliar o desempenho do sistema de otimização utilizando os aproximadores, dois testes foram realizados. No primeiro teste foi feita uma otimização utilizando o simulador IMEX (as tabelas de fluxo vertical multifásico foram geradas separadamente e, em seguida, inseridas no modelo do reservatório para que o simulador IMEX pudesse considerar a perda de carga nos poços, nas linhas de produção e no *riser*, como explicado no Capítulo 3), e um segundo teste, utilizando exclusivamente os aproximadores. O objetivo deste último teste é verificar o impacto causado nos resultados quando o simulador é substituído pelo aproximador em um processo de otimização. Todos os testes foram feitos com os aproximadores que contam com reservatórios com histórico de 3 anos para entrada da janela inicial.

4.5.1. Comparação entre a Utilização dos Simuladores e dos Aproximadores em Processos de Otimização

Nesta seção são apresentados os resultados do melhor indivíduo encontrado pelo sistema de otimização de localização de poços. No sistema de otimização, os parâmetros relacionados ao cálculo do VPL utilizados e os parâmetros do

algoritmo genético são apresentados na Tabela 27 e na Tabela 28, respectivamente:

Parâmetros Econômicos	Valor
Preço do petróleo (US\$/bbl)	100
Preço do gás (US\$/1000m ³)	1
Plataforma (MMUS\$)	1000
Perfuração e completção (MMUS\$/poço)	120 a 130
Custo de produção de óleo (US\$/m ³)	3
Custo de produção de gás (US\$/1000m ³)	1
Custo de produção de água (US\$/m ³)	3
Alíquota de impostos	0,34
Taxa mínima de atratividade	0,10
Royalties	0,10

Tabela 27: Parâmetros para cálculo do VPL

Parâmetros do Algoritmo Genético	Valor
Número de Gerações	20 e 30
Número de Indivíduos	30
Número de Rodadas	1
Steady State (inicial - final)	40% - 20%
Crossover (inicial - final)	65% - 8%
Mutação (inicial - final)	8% - 30%

Tabela 28: Parâmetros do Algoritmo Genético

Na Tabela 29 são apresentados os resultados do VPL do melhor indivíduo e o tempo de simulação quando se utiliza somente o simulador e quando se utiliza somente os aproximadores. No primeiro teste, a otimização foi feita utilizando um número de gerações igual a 20 e no segundo teste, igual a 30. Nos dois testes foram feitas três otimizações utilizando os aproximadores. O objetivo é verificar se a curva de evolução obtida na otimização com os aproximadores se aproxima da curva de evolução obtida na otimização com o simulador.

Observa-se que o sistema de otimização com a substituição do simulador conduz a valores de VPL próximos aos obtidos quando se utiliza somente o simulador. Observa-se também que o tempo necessário para uma otimização com

um mesmo número de avaliações é de aproximadamente 90% mais rápido quando se utiliza somente os aproximadores, tanto para o primeiro quanto para o segundo teste. Esse tempo poderia ser reduzido mais ainda se os teste fossem feitos com aproximadores que não necessitam de 3 anos de histórico, desse jeito não estaria embutido o tempo de simulação dos modelos pelos mesmos 3 anos.

		Tempo de Execução	VPL (10⁹US\$)
Teste 1	Simuladores	15,4 horas	2,478
	Simuladores	11,48 horas	2,124
	Simuladores	12,98 horas	2,493
	Aproximadores	1,51 horas	1,821
	Aproximadores	1,91 horas	1,628
	Aproximadores	1,83 horas	2,225
	Teste 2	Simuladores	22,6 horas
Simuladores		20,83 horas	3,010
Simuladores		19,58 horas	3,279
Aproximadores		2,10 horas	3,665
Aproximadores		2,43 horas	3,218
Aproximadores		2,75 horas	2,986

Tabela 29: Resultados das otimizações

	Aproximadores	Tempo de Execução	VPL (10⁹US\$)
Teste 1	Média	1,746	1,891
	Desvio Padrão	0,177	0,248
Teste 2	Média	2,425	3,290
	Desvio Padrão	0,265	0,282

Tabela 30: Média e desvio padrão dos dados obtidos com os aproximadores

	Simuladores	Tempo de Execução	VPL (10⁹US\$)
Teste 1	Média	13,28	2,365
	Desvio Padrão	1,615	0,171
Teste 2	Média	21,00	3,234
	Desvio Padrão	1,239	0,168

Tabela 31: Média e desvio padrão dos dados obtidos com os simuladores

A evolução do processo de otimização pode ser verificada na Figura 19, Figura 20, Figura 21, Figura 22, que apresentam as curvas de evolução do sistema de otimização onde se observa que os aproximadores aproximam bem a curva de evolução dos simuladores.

O sistema de otimização utiliza algoritmos genéticos (Mitchell, 1994; Koza, 1992) e uma das formas de avaliar o bom desempenho de um algoritmo genético é através das curvas de evolução:

- A curva *Offline* apresenta o valor médio das avaliações dos melhores indivíduos encontrados a cada passo da avaliação. Esta curva mostra a qualidade do algoritmo em encontrar soluções boas desde o início.
- A curva *Média* é obtida a partir da média das avaliações de todos os indivíduos. Esta curva permite verificar a rápida obtenção de boas soluções e também permite visualizar o grau de convergência dos indivíduos da população.
- A curva *Melhor* apresenta o melhor indivíduo encontrado em cada geração do algoritmo genético.

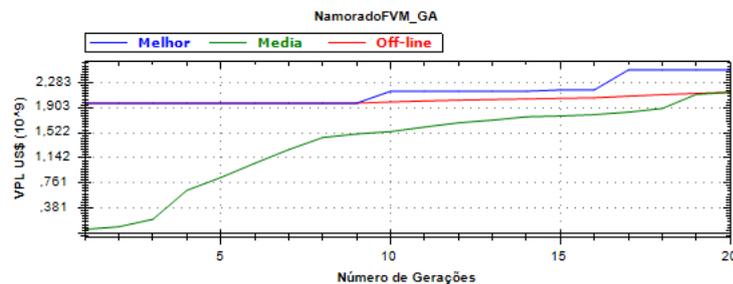


Figura 19: Curva de otimização com uso dos simuladores (20 gerações)

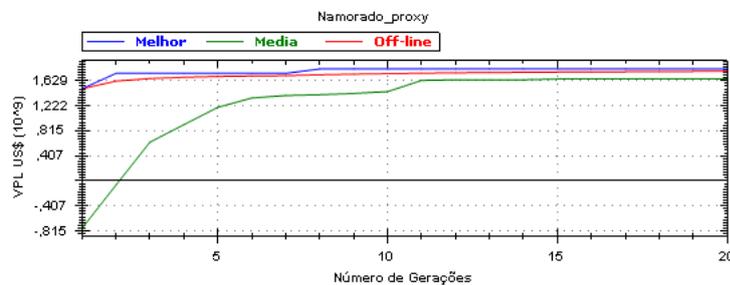


Figura 20: Curva de otimização com uso dos aproximadores (20 gerações)

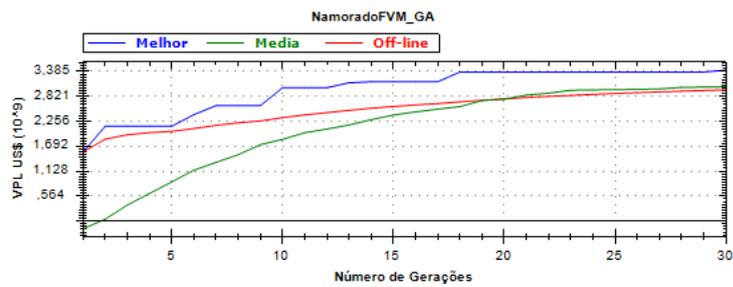


Figura 21: Curva de otimização com uso dos simuladores (30 gerações)

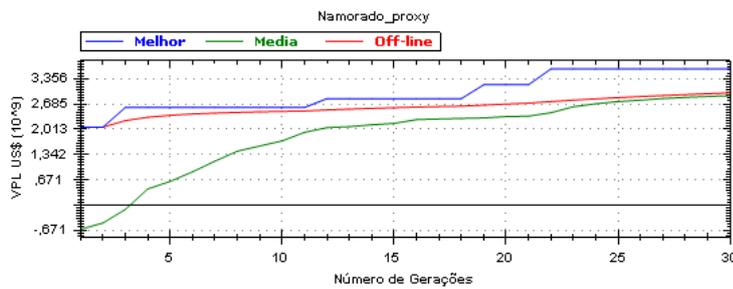


Figura 22: Curva de otimização com uso dos aproximadores (30 gerações)

5. Conclusões e Trabalhos Futuros

5.1. Conclusões

Este trabalho apresentou todas as etapas necessárias para a construção de um aproximador para a simulação integrada entre reservatório e sistema de produção com o objetivo de fornecer as curvas de produção de óleo, gás e água em um dado horizonte de tempo. O aproximador foi construído utilizando Redes Neurais do tipo MLP que foram treinadas com amostras fornecidas pelos próprios simuladores. A base de treinamento foi obtida por amostragem aleatória e por meio da técnica do Hipercubo Latino, visando minimizar a quantidade de amostras demandada aos simuladores. Além disso, em alguns dos testes realizados foi aplicada a técnica de Análise de Componentes Principais para a redução da quantidade de variáveis de entrada das redes neurais por trás dos modelos de aproximação.

Os modelos alcançaram bons resultados na aproximação das curvas de produção de óleo, gás e água, apresentando erros considerados aceitáveis quando essas curvas foram comparadas às curvas de produção geradas pela simulação. Os melhores resultados foram alcançados quando as amostras de treinamento foram obtidas por meio da técnica do Hipercubo Latino e com a abordagem que aproxima o simulador utilizando 3 grupos de 6 redes neurais, onde cada grupo é responsável pela aproximação de uma curva de produção. Os aproximadores construídos são limitados à um número fixo de poços definido antes da construção do mesmo. O tipo de cada um desses poços, entretanto, não é fixo. O algoritmo genético pode variar o tipo de cada um dos poços livremente.

Como se esperava, observou-se uma expressiva redução do custo computacional (média de redução de 89%) determinada pela utilização do aproximador. Uma preocupação constante foi para que o custo de criação dos aproximadores não fosse muito maior que o custo de se realizar uma otimização com algoritmos genéticos. Pelos resultados obtidos, observa-se que o custo para

construção dos aproximadores, quando se utiliza técnicas de amostragem como Hipercubo Latino, diminui bastante em relação ao uso de amostras geradas aleatoriamente. Apesar do custo de criação dos aproximadores, de forma serial, ainda ser maior que o de uma otimização que utiliza os simuladores, a criação dos aproximadores ainda é desejável, já que uma vez treinados, os mesmos são capazes de fornecer resultados rápidos todas as vezes que se desejar realizar uma otimização. Então, ao longo de sua utilização, o tempo de construção dos aproximadores pode se tornar menor que o custo de se realizar as otimizações.

Devido à rapidez na obtenção de bons resultados de VPL, associado ao baixo erro do aproximador, a substituição do simulador pelo aproximador se mostrou uma interessante estratégia a ser adotada em um modelo de otimização de localização de poços, disponibilizando ao especialista uma ferramenta com resultados rápidos para apoiar suas decisões.

5.2. Trabalhos Futuros

As seguintes oportunidades de trabalhos futuros podem ser destacadas a partir deste trabalho:

- Integração automática entre o simulador de reservatório IMEX, o simulador de superfície GAP e o simulador de poços PROSPER, a partir do uso do software de integração RESOLVE e o software OPENSEVER, que permite o uso dos simuladores por parte de uma aplicação externa.
- Integração do aproximador ao sistema de otimização OCTOPUS com o objetivo de diminuir ainda mais o custo computacional de otimizações. Desta forma se agregaria ao otimizador a oportunidade de se construir aproximadores disponibilizando ao especialista uma ferramenta com resultados rápidos para apoiar a tomada de decisões.
- Realização de testes de otimização com quantidades maiores de poços visando principalmente alternativas mais complexas e que demandem ainda mais recursos computacionais quando utilizado com os simuladores.

- Utilização da técnica de amostragem Hipercubo Latino em conjunto com a Análise de Componentes Principais.
- Desenvolvimento de políticas de substituição adaptativa dos simuladores utilizando percentuais de substituição que variem de acordo com o andamento do processo de otimização ou que privilegie a avaliação das melhores configurações de poços pelos simuladores.
- Utilização de outras técnicas inteligentes para a construção dos aproximadores, como Programação Genética e Neuro Fuzzy, com o objetivo de comparar modelos.
- Construção de aproximadores para atender às incertezas de natureza geológica, fazendo com que o aproximador seja capaz de aprender variações das características geológicas do reservatório.

Referências Bibliográficas

AGEH, E. A., **Using Integration Production Modeling (IPM) as an Optimization Tool for Field Development Planning and Management**. SPE 140625, 2010 – 34th Annual SPE International Conference and Exhibition, Calabar, Nigeria, 2010.

ALMEIDA, A. de C., **Aproximador de Função para Simulador de Reservatórios Petrolíferos Utilizando Técnicas de Inteligência Computacional e Projeto de Experimentos Fatoriais Fracionado**, Dissertação de Mestrado, Pontifícia Universidade Católica do Rio de Janeiro, 2008.

ALMEIDA, A. de C., **Proxies para Controle Pró-Ativo de Válvulas de Poços Inteligentes**, Congresso Brasileiro de Automática, Bonito-MS, 2010.

BISHOP, C., **Neural Network for Pattern Recognition**, Oxford University Press Inc. 1995.

CERQUEIRA, M., **Manifold Submarino**. Salvador, Universidade Petrobras, 2005.

CMG, IMEX - **Advanced Oil/Gas Reservoir Simulator - Version 2012 - User's Guide**. Calgary, Alberta, Canada: Computer Modelling Group LTD., 2012.

CORDAZZO, J. **Simulação de reservatórios de petróleo utilizando o método EbFVM e multigrid algébrico**. Tese (Doutorado em Engenharia Mecânica) - Programa de Pós-Graduação em Engenharia Mecânica. UFSC, Florianópolis. 2006.

DIEBOLD, F.X., LOPEZ, J., **Forecast Evaluation and Combination**, G.S. Maddala and C.R. Rao (eds.), Statistical Methods in Finance (Handbook of Statistics, Volume 14). Amsterdam: North-Holland, 241-268, 1996.

EMERICK, A. A., SILVA, E., MESSER, B., ALMEIDA, L. F., SZWARCMAN, D., PACHECO, M. A. P., VELLASCO, M. M. B. R. **Well Placement Optimization Using a Genetic Algorithm with Nonlinear Constraints**. SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA. SPE118808, 2009.

ERTEKIN, T., ABOU-KASSEM, J. H., KING, G. R. **Basic Applied Reservoir Simulation**. SPE Textbook Series, Vo1. 7. Richardson SPE, 2001.

FALETTI, L., **Otimização de Alternativas para Desenvolvimento de Campo de Petróleo utilizando Computação evolucionária**, Dissertação de Mestrado, Pontifícia Universidade Católica do Rio de Janeiro, 2003.

FANCHI, J. R., HARPOLE, K. J., BUJNOWSKI, S. W. **A Three-dimensional, Three-phase Black Oil Applied Simulation Tool**. Oklahoma, 1982.

FEITOSA, R. Q., Vellasco, M. M. B. R., Oliveira, D. T., Andrade, D. V., Maffra, S. A. R. S., **Classificação de Expressões Faciais Utilizando Redes Neurais Back Propagation e RBF**. Workshop de Computação (WORKCOMP'99), pp. 69-76, ITA, São José dos Campos, SP, 1999.

FRANCO, K., **Desenvolvimento de um Sistema Inteligente para Auxiliar a Escolha de Sistema para Produção no Mar**. Dissertação de Mestrado., Universidade Estadual de Campinas, Campinas, SP, Brasil, 2003.

HAYDER, M. E., **Facilities Planning Using Coupled Surface and Reservoir Simulation Models**, SPE 149037, 2011 SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition, Al-Khobar, Saudi Arabia, 2011.

HAYDER, M. E., **Production Optimization Through Coupled Facility/Reservoir Simulation**, SPE 100027, 2006 SPE Intelligent Energy Conference and Exhibition, Amsterdam, Netherlands, 2006.

HAYKIN, S., **Redes Neurais: princípios e prática**. Trad. Paulo Martins Engel, 2.ed. Porto Alegre: Bookman, 2001.

HORNIK, M., STINCHCOMBE, M., & WHITE, H., **Multilayer Feedforward Networks are Universal Approximators**, Neural Networks, 2, 359-366, 1989.

HUSH, D., HORNE, B., **Progress in Supervised Neural Networks**. IEEE Signal Processing Magazine pp. 8–39, 1993.

IMAN, R.L., HELTON, J.C., AND CAMPBELL, J.E., **An Approach to Sensitivity Analysis of Computer Models, Part 1. Introduction, Input Variable Selection and Preliminary Variable Assessment**, Journal of Quality Technology 13 (3): 174–183, 1981.

IYODA, E., **Inteligência computacional no projeto automático de redes neurais híbridas e redes neurofuzzy heterogêneas**. Tese de Mestrado, FEEC - Unicamp, Brasil, 2000.

JANG, J. S. R., **ANFIS: Adaptive-Network-Based Fuzzy Inference Systems**, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 03, pp. 665-685, 1993.

JOLLIFFE, I. T., **Principal Component Analysis**, Second ed. Springer Series in Statistics, New York: Springer-Verlag, New York, 2002.

KOZA, J. R., **Genetic Programming: On the Programming of Computers by Means of Natural Selection**, 6th edition. MIT Press, Cambridge, MA, EUA, 1992.

LOCHNER, R.H., MATAR, J.E., **Design for Quality**, Chapman and Hall, London, 1990.

LUDWIG JR., O., MONTGOMERY, E., **Redes Neurais: Fundamentos e Aplicações com Programas em C**. Ed. 1. Editora Ciência moderna. 2007. 186 f, 2007.

MAKRIDAKIS, S. G.; HIBON, M. **Exponential smoothing: The effect of initial values and loss functions on post-sample forecasting accuracy**. International Journal of Forecasting, v.7, p. 317–330, 1991.

MATTAX, C. C., DALTON, R. L. **Reservoir Simulation**. SPE Monograph Series Vo1. 13. Richardson SP, 1990.

MCKAY, M.D., BECKMAN, R.J., CONOVER, W.J., **A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code**. Technometrics, Vol. 21, n.2, p. 239-245, May 1979.

Mitchell, M., **An Introduction to Genetic Algorithms**, the MIT Press, USA, 1994.

MOREIRA, J., **Operações e Produção no Mar**. Salvador, Universidade Petrobras, 2005.

NN3. **NN3 Home**. Disponível em: <<http://www.neural-forecasting-competition.com/NN3/index.htm>>. Acesso em 02/07/2013.

OLIVEIRA, B. C., **Simulação de Reservatórios de Petróleo em Ambiente MPSoC**. Dissertação de Mestrado, Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Natal, 2009.

OLIVEIRA, M., **Análise da Aplicação de um Sistema de Bombeamento Multifásico Submarino na Produção de Petróleo**. Dissertação de Mestrado, Pontifícia Universidade Católica, Rio de Janeiro, RJ, Brasil, 2003.

PETEX, PROSPER - **Integrated Production Modeling** - Version 2012 - User's Guide. Edinburgh, Scotland: Petroleum Experts Limited, 2012.

RIBEIRO, S. G., **Notas de Aula**, Pontifícia Universidade Católica do Rio de Janeiro, 2012.

RODRÍGUEZ, R., **Integration of Subsurface, Surface and Economics Under Uncertainty in Orocual Field**, SPE 107259, 2007 SPE Latin America and Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina, 2007.

ROSA, A., J., CARVALHO, R. S., XAVIER., J. A. D., **Fluxo de Líquido em Meios Porosos e Mecanismo de Produção de Reservatórios**. Salvador, Universidade Petrobras, 2004.

ROSA, V. R., **Otimização em Localização de Plataformas de Produção**. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Brasil, 2006.

ROTONDI, M., COMINELLI, A., DI GIORGIO, C., ROSSI, R., VIGNATI, E., CARATI, B. **The Benefits of Integrated Asset Modeling: Lesson Learned from Field Cases**. Paper SPE 113831 presented at the SPE Europec/EAGE Annual Conference and Exhibition, Rome, Italy, 2008.

SAINI, L.M., SONI, M.K., **Artificial neural network based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods**, IEE Proceedings on Generation Transmission, Distribution, Vol.149, N.5, pp. 578- 584, 2002.

SCHLUMBERGER Ltd. **ECLIPSE Reservoir Engineering Software**. Disponível em: <<http://www.slb.com/>>. Acesso em 01/07/2013.

SILVA, A.P.D., **Efficient Variable Screening for Multivariate Analysis**. Journal of Multivariate Analysis, v. 76, pp. 35-62, 2001.

SOARES, A. A. M., **Simulação de Reservatórios de Petróleo em Arquiteturas Paralelas com Memória Distribuída**. Dissertação de Mestrado (Ciências e Engenharia Civil), Engenharia Civil, UFPE, Recife, 2002.

STONE, M., **Cross-validation: A Review**, Mathematische Operationsforschung Statistishen, Serie Statistics. v. 9, p. 127-139, 1978.

STONE, M., **Cross-validation Choice and Assessment of Statistical Prediction**. Journal of the Royal Statistical Society, v. B36, p. 111-133, 1974.

TALAVERA, A. L., TÚPAC, Y. J., VELLASCO, M. M. B. R., **Controlling Oil Production in Smart Wells by MPC Strategy with Reinforcement Learning**, Society of Petroleum Engineers LAPEC 2010. Lima, Peru: SPE 139299, 2010.

THOMAS, J. E., **Fundamentos de Engenharia de Petróleo**. Rio de Janeiro, Editora Interciência, 2001.

TÚPAC, Y. J. V., **Sistema Inteligente de Otimização de Alternativas de Desenvolvimento de Campos Petrolíferos**. Tese (Doutorado) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2005.

URSINI, F., **Forecasting Reservoir Management Through Integrated Asset Modelling**, SPE 128165, 2010 SPE North Africa Technical Conference and Exhibition, Cairo, Egypt, 2010.

VILLANUEVA, J. M. M., **Modelo de Aproximação de um Simulador de Produção de Petróleo Utilizando Redes Neurais Artificiais**, Revista de Inteligência Computacional Aplicada, 2010.

WASSERMAN, P. D., **Advanced Methods in Neural Computing**. Van Nostrand Reinhold, New York, N. Y., 1993.

YU, T., Wilkinson, D., Castellini, A., **Constructing Reservoir Flow Simulator Proxies Using Genetic Programming for History Matching and Production Forecast Uncertainty Analysis**, Journal of Artificial Evolution and Applications, 2007.

Apêndice A – Criação de *Manifolds* Imaginários no Simulador de Reservatório IMEX

Neste apêndice é mostrado como foram criados os *manifolds* virtuais no IMEX.

1. Primeiramente deve-se declarar um grupo para cada poço produtor que será ligado a um *manifold*. Supondo que somente dois poços tenham um *manifold* imaginário:

```
*GROUP 'GRUPO1' *ATTACHTO 'Field'
*GROUP 'GRUPO2' *ATTACHTO 'Field'
```

A linha acima deve ser inserida no arquivo .dat na parte de definição de poços, porém antes à definição do poço em questão. 'GRUPO1' é o nome do grupo filho ao qual irá se atrelar o poço e 'Field' é o nome do grupo ao qual os grupos filhos estarão atrelados.

2. Deve-se incluir também as tabelas de fluxo vertical multifásico dos poços (nesse caso, uma única tabela para todos os poços) e as tabelas de superfície (linhas de produção e *riser*) correspondentes às distâncias de cada poço à plataforma, como mostrado abaixo:

```
*INCLUDE 'PocoTipo.inc'
*include 'FVM_Surface_21.inc'
*include 'FVM_Surface_49.inc'
```

As tabelas de fluxo vertical devem estar em ordem, portanto dentro do arquivo 'PocoTipo.inc' deve estar declarado que se trata da tabela número 1 e deve também estar declarada a profundidade do trecho modelado pela tabela. O mesmo deve ser feito para as tabelas seguintes.

E após a definição dos poços deve-se especificar as taxas de *gas lift* para o conjunto de poços produtores:

```
*GLIFT 'poco1' 'poco2'
100000.0 100000.0
```

No caso, a taxa de injeção de gás é fixa e igual a 100 mil m³/dia para cada poço.

3. Na definição dos poços em questão deve-se atrelar os poços aos grupos anteriormente criados, como mostrado abaixo:

```
*WELL 'poco1' *ATTACHTO 'GRUPO1'
WELL 1 'poco1'
*PRODUCER 'poco1'
*PWELLBORE TABLE 2960.4 1
```

A última linha é a linha que define a tabela do poço, onde 2960.4 é a profundidade do poço e 1 é o número da tabela de fluxo vertical multifásico do poço (a tabela presente no arquivo 'PocoTipo.inc').

4. Após a declaração completa do poço, faz-se a declaração do *manifold* imaginário que conectará a tabela de cada poço à sua respectiva tabela de superfície:

```
*MANIFOLD *PROD 'GRUPO1' *ON
*PCON-MANIFOLD *PROD 'GRUPO1' *ON
*GPTABLE *PROD 'GRUPO1' 2
*GPHYDDEP *PROD 'GRUPO1' 1000
*GCONP 'GRUPO1'
*TARGET *CPP 15. *CONT
```

No grupo de comandos acima, 'GRUPO1' se refere ao grupo atrelado ao poço em questão. O número 2 no comando 'GPTABLE' é o número da tabela de superfície associado ao poço (no caso, o arquivo *include 'FVM_Surface_21.inc'). O número 1000 no comando 'GPHYDDEP' é a profundidade vertical do sistema

de superfície (linhas de produção e *riser*), que nesse caso é a medida da lâmina d'água, que é igual a 1000 metros. O valor 15 no comando ‘*TARGET *CPP’ define a pressão no compressor (plataforma) para o grupo em kg/cm^2 .