



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA QUÍMICA

DESENVOLVIMENTO DE PROCESSOS QUÍMICOS

**MODELAGEM E SIMULAÇÃO DO PROCESSO DE
EXTRAÇÃO SELETIVA DE COBALTO E NÍQUEL POR
SOLUÇÃO ORGÂNICA ATRAVÉS DA APLICAÇÃO DE
TÉCNICAS DE REDES NEURAIIS**

Wesley Marinho Ferreira

Autor

Prof. Dr. Milton Mori

Orientador

Dr. Reinaldo Krause Spitzner Júnior

Co-Orientador

**UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE**

**Dissertação de Mestrado apresentada à Faculdade de Engenharia Química como parte
dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Química**

Campinas - SP

Novembro / 2002

| | |
|------------|-------------------------------------|
| UNIDADE | B.E. |
| Nº CHAMADA | UNICAMP F413m |
| V | EX |
| TOMBO BC/ | 55079 |
| PROC. | 16-124103 |
| C | <input type="checkbox"/> |
| D | <input checked="" type="checkbox"/> |
| PREÇO | R\$ 11,00 |
| DATA | 07/08/03 |
| Nº CPD | |

CM00188039-B

BIB ID 296339

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

~~F414m~~
F413m

Ferreira, Wesley Marinho

Modelagem e simulação do processo de extração seletiva de cobalto e níquel por solução orgânica através da aplicação de técnicas de redes neurais / Wesley Marinho Ferreira.--Campinas, SP: [s.n.], 2002.

Orientador: Milton Mori e Reinaldo Krause Spitzner Júnior.

Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Redes neurais (Computação). 2. Inteligência artificial. 3. Simulação (Computadores). 4. Metais. I. Mori, Milton. II. Spitzner Júnior, Reinaldo Krause. III. Universidade Estadual de Campinas. Faculdade de Engenharia Química. IV. Título.

Dissertação de Mestrado defendida por Wesley Marinho Ferreira e aprovada em 29 de Novembro de 2002 pela banca examinadora constituída pelos doutores:



Prof. Dr. Milton Mori - Orientador



Prof.ª Dr.ª Liliane Maria Ferrareso Lona



Prof. Dr. Carlos Alberto Ubirajara Gontarski

200326735

Este exemplar corresponde à versão final da Dissertação de Mestrado em Engenharia Química.



Prof. Dr. Milton Mori

Orientador

Dedico este trabalho ao meu pai João Eudes Ferreira, minha mãe Sonia Marinho , meu irmão Wonder Marinho, meu sobrinho Matheus e todos os meus familiares e amigos que me apoiaram na imersão e conquista deste sonho, que muito mais que família são minha estrutura espiritual, emocional e intelectual.

AGRADECIMENTOS

“Feliz sou eu que tenho muito e muitas pessoas para agradecer. Feliz sou eu que consigo agradecer. Feliz sou eu que não sei por onde começar, afinal foram tantas as pessoas que passaram por mim e que passei por elas, graças a Deus...Epa...agora sei por onde começar...”

“Senhor eu sei que tu me sondas, Senhor eu sei que tu me amas....Senhor eu sei que tu me sondas, Senhor eu sei que tu me amas.” Agradeço ao Senhor, meu Deus, pela vida e pelo ser humano que é brilhante, filhos de Ti Perfeito, capazes de crescer, se adaptar e de brilhar. Não posso deixar de agradecer por tudo que aprendi, cresci e pelas pessoas que conheci. Sejam elas tendo ligações passageiras, sejam elas com ligações fortes e duradouras com meu coração. Não ligações químicas, mas sim ligações de afeto que nem o mais “forte” cabo produzido pela ciência tem resistência igual.

“Abençoa Senhor as famílias Amém, Abençoa Senhor a minha também...” Agradeço de coração aos meus pais, João Eudes Ferreira e minha mãe Sonia Marinho que tanto investiram em mim, me mostrando os valores realmente importantes da vida para o crescimento e acreditam que o amanhã sempre será melhor. Ao meu irmão Wonder Marinho, que de maneira muito pessoal, transmite o seu carinho e apoio às minhas conquistas. Agradeço pelo incentivo dos meus tios, tias e primos, em especial: Aloizio Marinho, Amália Marinho, Edméia Marinho, Maria Aparecida Marinho, Maria Aparecida Acioli Vilela, Lourdes Marinho, Eduardo Marinho (*in memoriam*), e Maria Amália Almeida Marinho, minha querida avó, que onde quer que esteja, está olhando por toda a família .

“Ensinar é uma arte, deixar aprender é uma dádiva.” Agradeço ao amigo Professor Dr. Milton Mori pela estrutura de estudo e por muitas oportunidades em vários ramos da Engenharia, além da pesquisa aqui apresentada, seja quando convidado a participar da diretoria da SBEQ – Sociedade Bandeirante de Engenharia Química, seja como monitor das disciplinas do Curso de Engenharia Ambiental. Agradeço por partilhar sua sabedoria acadêmica, social e administrativa visando sempre a melhoria e a busca de resultados.

Agradeço ao professor Dr. Reinaldo Krause Spitzner Júnior pelo apoio e por possibilitar a união, que leva ao desenvolvimento, universidade-indústria, e por humildemente partilhar sua sabedoria para enriquecimento deste trabalho.

“Amigo é coisa pra se guardar do lado esquerdo do peito, dentro do coração assim falava a canção...” Agradeço pelo carinho, amizade e paciência das pessoas que me permitiram fazer parte do seu “hall” de sentimento no coração: “Faça uma lista de grandes amigos, quem você mais via há 10 anos atrás. Quantos você ainda vê todo dia e quantos você já não encontra mais...faça uma lista dos sonhos que tinha, quantos você desistiu de sonhar...Faça uma lista...” Agradeço pela amizade que completam a minha vida para realizar trabalhos como este e adicionam nas necessidades básicas do meu ser, os meus amigos: Wanda Batista, Sérgio Bernardo, Érica Maróstica, Flávia Glória, Caio Luminatti, João Paulo Barros Machado, Paulo Limeira, Fabiola Andrea, Karla Patrícia, Juliana Ferreira, Juliana Acioli, Edson Fujita, Miriam, Patrícia, Samira Leão, Édler Lins, Elaine, Luciana, Vinicius, Mirla, Graça, Marcel, Patrícia, Joseane, Josilene, Nagel, Elver Radke, Leonardo Rosa, Vanessa Ito, Maria Cristina, Edvaldo, Roziléia Paloschi (*im memoriam*), Renato Botinni & Cristina, Amanda, Alexandre Antunes, João Marcelino, Miguel, Breno, Alberto Laham, amigos do LMSPQ e todos que pelo olhar me fazem amigo.

De forma especial também, agradeço a atenção, a amizade e o carinho da Anna Ritta, pessoa de coração extraordinário, que com sua vivência, sempre me apoiou e aconselhou em muitos momentos importantes.

Agradeço a Prof. Meuris Gurgel Carlos da Silva pela amizade e pela permissão em utilizar as dependências do seu laboratório para parte da execução deste trabalho. Bem como seu esposo, Laerte, pessoa de humor apurado que muito participou de vários momentos alegres.

Agradeço à Votorantim – CNT – Companhia Níquel Tocantins pelo auxílio na pesquisa do processo de extração e análises químicas para obtenção dos dados, em especial o Químico Marco A. Ângora , o Técnico Sérgio e o Engenheiro de Processo Anderson V. Ramicelli.

Ao CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, pelo apoio financeiro.

Meu muito obrigado a todos os funcionários e professores que de alguma forma me orientaram e ajudaram para que este trabalho fosse feito da melhor maneira .

“Nunca ande pelo caminho traçado,
pois ele conduz somente até onde
os outros foram.”

Alexandre Graham Bell

RESUMO

A indústria química sofreu e tem sofrido mudanças constantes durante os últimos anos, principalmente devido à progressiva transformação dos recursos tecnológicos, aumento do custo de energia, às restrições ambientais e à crescente competitividade mundial, tendo como causa principal a globalização do mercado.

Em função deste panorama, vê-se a necessidade do contínuo desenvolvimento de competências essenciais, que irão sustentar esta competitividade para assim diferenciá-la estrategicamente, aumentando a necessidade de conhecimento detalhado dos processos e maior demanda de desenvolvimento tecnológico.

A inteligência artificial vem ao encontro deste interesse por ser atrativa e importante na compreensão de vários processos principalmente químicos, sendo um conjunto de técnicas promissoras na modelagem e simulação de processos industriais que apresentam não linearidades, como por exemplo, as técnicas de Redes Neurais Artificiais (RNA).

Este trabalho apresenta a aplicação da metodologia de redes neurais na modelagem, simulação e aplicação de uma parte do processo industrial da VOTORANTIM – CNT (Companhia Níquel Tocantins) de obtenção de níquel e cobalto eletrolíticos. Esta parte do processo consiste na extração de níquel e cobalto de uma solução de sulfatos através de uma solução orgânica fosfínica, em que os parâmetros característicos ainda não são bem conhecidos por apresentarem complexa modelagem fenomenológica.

Foram gerados dados através de experimentos (removendo erros grosseiros), variando os valores dos seguintes parâmetros: pH, temperatura e volume de orgânico. O Planejamento experimental foi realizado após estudo do processo e verificação das possibilidades e necessidades industriais para simular a extração dos metais e podendo assim mapear uma solução (otimização) através de redes neurais.

Os resultados da modelagem via RNA foram muito satisfatórios quando comparados a muitos da literatura, pois mostra a necessidade da qualidade dos dados e seu tratamento antes da alimentação à rede e a possibilidade de encontrar redes simples para modelagens complexas. Desta forma, as RNA são apresentadas como importante ferramenta na otimização e controle de processos não lineares.

ABSTRACT

The chemical industry changed and has been changing constantly during the last years, mainly due to the progressive transformation of the technological resources, energy increasing price, environmental restrictions and world increasing competition, being the market globalization the main cause.

Thus, there is a necessity of a continuous development of essential competences that will support this competition to make it strategically different, increasing the necessity of a detailed knowledge of the processes and a higher request of technological development.

The neural networks techniques come together with this interest because they are attractive and important to the several processes comprehension, mainly chemicals, being a part of a set of promising techniques in modeling and simulation of industrial processes that are non linear.

This task presents the application of neural networks methodology for modeling, simulation and application of a part of VOTORANTIM - CNT (Companhia Níquel Tocantins) industrial processes, obtaining nickel and cobalt from a sulphate mixture through a phosphinic organic mixture where the characteristic parameters are not well known yet because they present a phenomenological modeling compound.

Data were available from the changing the values of the following parameters, pH, temperature and organic volume. The experimental planning was accomplished after a study of the process and examination of the possibilities and industrial necessities to simulate the metals extration and in this manner, being able to map a mixture (optimization) from neural networks.

The results of the modeling via ANN were very satisfactory when compared to those of the literature, because they show the necessity of data quality and its treatment before net supplying and the possibility of finding ordinary nets to complex modeling. In this way the ANN are presented as important tools for the optimization and control of non linear processes.

SUMÁRIO

| | |
|---|-----------|
| 1- Introdução | 12 |
| 2- Revisão Bibliográfica | |
| 2.1 - Redes Neurais | 16 |
| 2.1.1- Introdução | 16 |
| 2.1.2- Histórico | 17 |
| 2.1.3- Neurônios Biológicos | 20 |
| 2.1.4- Neurônios Artificiais | 23 |
| 2.1.5- Tipos de Redes Neurais | 26 |
| 2.1.5.1- Tipos de Redes Neurais, quanto à sua estrutura. | 26 |
| 2.1.5.2- Tipos de Redes quanto ao algoritmo de treinamento. | 30 |
| 2.1.5.2.1- O Algoritmo Back-Propagation | 33 |
| 3- Descrição do Processo | |
| 3.1- Histórico da Companhia Níquel Tocantins | 41 |
| 3.2 – Processo Industrial | 42 |
| 3.2.1- Níquel e Cobalto eletrolítico | 43 |
| 3.2.2- Processo de Extração de Cobalto e Níquel | 44 |
| 4- Materiais e Métodos | |
| 4.1- Equipamentos, Reagentes e Ferramentas Computacionais | 48 |
| 4.2- Metodologia Utilizada | 49 |

| | |
|--|-----------|
| 4.2.1- Revisão Bibliográfica. | 50 |
| 4.2.2- Estudo Do Problema a ser Modelado e Aquisição dos Dados. | 50 |
| 4.2.2.1- Obtenção das Curvas de Extração de Níquel e Cobalto por Solução Orgânica de Cyanex 272 | 50 |
| 4.2.3- Análise Estatística dos Dados para o Treinamento da Rede Neural. | 53 |
| 4.2.4- Verificação do Domínio das Variáveis da Rede Neural | 54 |
| 4.2.5- Normalização dos Dados | 54 |
| 4.2.6- Treinamento da Rede Neural | 54 |
| 4.2.7- Validação da Rede Neural Treinada | 55 |
| 4.2.8- Utilização da Rede. | 56 |
| | |
| 5 – Resultados e Discussões | |
| 5.1- Análise dos Dados Experimentais | 57 |
| 5.1.1- Extrações de Cobalto por Solução Orgânica Cyanex 272 | 57 |
| 5.1.2- Extração de Níquel por Solução Orgânica Cyanex 272. | 61 |
| 5.1.3- Determinação do Coeficiente de Partição através de um planejamento fatorial a três níveis e três fatores (Planejamento 3 ³). | 67 |
| 5.2- Redes Neurais Artificiais | 73 |
| | |
| 6 – Conclusões e Sugestões | 78 |
| 6.1- Conclusões | 78 |
| 6.2- Sugestões | 79 |
| | |
| Referências Bibliográficas | 80 |

Anexos 83

Capítulo 1 – INTRODUÇÃO

A indústria química sofreu e tem sofrido mudanças constantes durante os últimos anos, principalmente devido à progressiva transformação dos recursos tecnológicos, aumento do custo de energia, às restrições ambientais e à crescente competitividade mundial, tendo como causa principal a globalização do mercado.

Em função deste panorama, vê-se a necessidade de contínuo desenvolvimento de competências essenciais que irão sustentar esta competitividade para assim diferenciá-la estrategicamente.

Vários processos químicos, que ainda não foram modelados ou simulados, precisam de novas técnicas para serem dominados, e assim capazes de serem um diferencial para a melhoria do seu sistema e se tornar mais produtiva otimizando seu processo e assim mais competitiva perante o mercado pois objetiva maior produção e menores custos.

Esta necessidade de maior conhecimento detalhado e otimização de processos faz com que a procura por novas tecnologias cresça tornando o processo de automação muito importante e atrativo para as empresas.

A automatização industrial faz-se necessária pela confiabilidade, segurança e pela busca da qualidade permanente e de melhorias contínuas que os sistemas exigem para a sobrevivência da empresa e aumento de sua lucratividade pela redução dos custos, pelo aumento da produção e seu controle.

A Inteligência artificial (IA) é uma outra fonte de tecnologia, que segundo CROWE e VASSILIADIS (1995) está pagando dividendos em produtividade e qualidade de produtos após um período de expectativas irreais.

De forma muito simples, a própria aplicação de tecnologias de IA pode gerar indústrias muito mais competitivas oferecendo aprimoramentos para os sistemas de controle de processo, resultando em aumento consistente de qualidade e de produção.

Para manter sua posição no mercado, as empresas devem atender basicamente duas metas: *fazer um produto com uma qualidade requerida com o menor custo possível.*

Casos industriais têm sido estudados, tais como processos de produção de cimento através de modelos via Redes Neurais por FUJITA (2001), processos de tratamento de efluentes através de Análise de Componentes Principais e Redes Neurais por OLIVEIRA (2000), processos de absorção e separação por SONG et al. (1999), processos de polimerização via modelos híbridos por CHAN & NASCIMENTO (1994), mostrando enormes desenvolvimentos, aplicações e perspectivas de emprego desta técnica em solução de problemas reais.

Já se percebeu que a otimização de processos e de seus subsistemas é uma das ferramentas mais importantes para redução de custos. Com o desenvolvimento e aumento da potencialidade dos computadores, a complexidade dos problemas que podem ser resolvidos pelas técnicas tradicionais de otimização têm aumentado.

O emprego de modelos fenomenológicos e técnicas em Redes Neurais, chamados de sistemas híbridos, atualmente possui também boa aceitação na solução de problemas de otimização. Pode-se utilizar simuladores comerciais existentes capazes de simular rigorosamente os mais variados equipamentos e processos. Com um conjunto de dados consistentes, dentro de um domínio determinado, gerados por estes modelos, é possível mapear soluções pelas técnicas de redes neurais.

O trabalho aqui desenvolvido apresenta o estudo a aplicação da metodologia de redes neurais na modelagem, simulação e na otimização de um processo industrial de extração de Cobalto e Níquel. As motivações, que podem ser citadas como de grande importância deste trabalho são:

- Modelagem de sistemas complexos;
- Obtenção de soluções rápidas a partir de simulações para este processo complexo;
- Conhecimento e redefinição das variáveis que realmente influenciam no processo;
- Busca de regiões ótimas de operação e busca de novas tecnologias para sustentar a vantagem competitiva da empresa.

O capítulo 2, revisão bibliográfica, abrange o processo industrial de extração de Níquel e Cobalto, a otimização de processos e redes neurais. O capítulo 3, descrição do processo da VOTORANTIM, apresenta a descrição do processo de obtenção de Níquel e Cobalto eletrolíticos, suas principais características, bem como o alvo deste trabalho que é a extração dos metais (Ni e Co) através de solução orgânica. O Capítulo 4, Materiais e Métodos, trata da metodologia utilizada para obtenção dos dados e para a modelagem das Redes Neurais que representem o processo industrial estudado. O Capítulo 5, Resultados e Discussões, segue com a apresentação do comportamento experimental da extração e dos modelos gerados, tendo como resultado desta etapa os pesos associados às redes neurais. O capítulo 6, Conclusões e Sugestões, apresenta observações pertinentes ao processo atribuindo a todo estudo as conclusões gerais e recomendações referentes a modelagem, simulação e otimização deste processo de extração de Níquel e Cobalto.

Capítulo 2 – REVISÃO BIBLIOGRÁFICA

2.1 - Redes Neurais

2.1.1- Introdução

As Redes Neurais Artificiais são identificadas freqüentemente como subespecialidade da Inteligência Artificial e também são conhecidas como RNA – Redes Neurais Artificiais (ou ANN – Artificial Neural Network), Conexionismo ou ainda Sistemas de Processamento Paralelo e Distribuído.

Foram inspiradas a partir do trabalho que o neuro-biologista, Warren McCulloch e o estatístico, Walter Pitts, publicaram, “A Logical Calculus of Ideas Imminent in Nervous Activity” em 1943.

Redes Neurais são sistemas paralelos distribuídos compostos por unidade de processamento simples (nodos) que calculam determinadas funções matemáticas (normalmente não-lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. O funcionamento destas redes é inspirado em uma estrutura física concebida pela natureza: o cérebro humano.

A solução de problemas através de RNAs é bastante atrativa, pois não dependem de um modelo matemático que relacione a entrada do processo com a saída, e já que a forma como estes são representados internamente pela rede e o paralelismo natural inerente à arquitetura das RNAs criam a possibilidade de um desempenho superior ao dos modelos convencionais. Elas ‘aprendem’ essa relação a partir de um processo de “treinamento” semelhante ao aprendizado de um cérebro humano, uma boa leitura sobre o assunto é o livro KÓVACS (1996) e o artigo WILLIS et. al. (1991).

2.1.2- Histórico

A era moderna das Redes Neurais começou com o trabalho pioneiro de McCulloch e Pitts em 1943 como já citado. McCulloch foi um psiquiatra e neuroanatomista por treinamento; passou cerca de 20 anos refletindo sobre a representação de um evento no sistema nervoso. Pitts foi um prodígio matemático que se associou a McCulloch em 1942. De acordo com HALL (1990), o artigo de McCulloch e Pitts surgiu dentro de uma comunidade de modelagem neural que tinha estado em atividade na Universidade de Chicago por pelo menos cinco anos antes de 1943, sob a liderança de Rashevsky.

No trabalho que publicaram juntos - “ A Logical Calculus of the Ideas Immanent in Nervous Activity” - é apresentada uma discussão sofisticada de redes lógicas de nodos (neurônios) e idéias inovadoras sobre máquinas de estados finitos, elementos de decisão de limiar lineares e representações lógicas de várias formas de comportamento e memória. Este trabalho de McCulloch e Pitts descreve um modelo artificial de um neurônio e apresenta suas capacidades computacionais e não se preocupa tanto com a apresentação de técnicas de aprendizado como é comum em discussões de RNAs, ou seja descreve um cálculo lógico das redes neurais que unificava os estudos de neurofisiologia e da lógica matemática. Eles assumiam que o seu modelo formal de um neurônio seguia uma lei “tudo ou nada”. Com um número suficiente dessas unidades simples e com conexões sinápticas ajustadas apropriadamente e operando de forma síncrona, McCulloch e Pitts mostraram que uma rede assim construída realizaria, a princípio, a computação de qualquer função computável. Este era um resultado muito significativo e com ele é geralmente aceito o nascimento das disciplinas de redes neurais e inteligência artificial.

Frank Rosenblat, em 1957, publicou em Cornell o primeiro grande projeto em computação neural, o desenvolvimento de um elemento chamado “perceptron”, um método inovador de aprendizagem supervisionada, que desencadeou uma série de pesquisas em computação neural.

Este tipo de rede despertou grande interesse tanto no início como em meados dos

anos 60. Durante este período clássico do perceptron, parecia que as redes neurais poderiam realizar qualquer coisa. Mas então em 1969, Marvin Minsky e Seymour Papert, que utilizaram a matemática para demonstrar que existem limites fundamentais para aquilo que os perceptrons de camada única podem calcular, ambos do Laboratório de Pesquisas Eletrônicas do MIT, publicaram um livro, “Perceptrons”, criticando profundamente o perceptron, e praticamente liquidaram este tópico de pesquisa ao mostrar as limitações do perceptron em termos de poder de expressão e habilidade de aprendizagem.

No anos 70, a abordagem conexionista ficou adormecida, em grande parte por causa do trabalho publicado por Minsky e Papert, apesar de alguns poucos pesquisadores continuarem trabalhando na área.

Em 1982, Hopfield publicou um artigo que chamou a atenção para as propriedades associativas das RNAs, sendo responsável por parte da retomada das pesquisas na área. Hopfield mostrou a relação entre redes recorrentes auto-associativas e sistemas físicos, o que também abriu espaço para a utilização de teorias correntes da Física para estudar tais modelos. Não obstante, a descrição do algoritmo de treinamento back-propagation alguns anos mais tarde mostrou que Minsky e Papert estavam com uma visão muito pessimista sobre o Perceptron. A partir da década de 80, houve nova explosão de interesse na área: em primeiro lugar, o avanço da tecnologia, sobretudo da microeletrônica, que vem permitindo a realização física de modelos de neurônios e sua interconexão de um modo antes impensável; em segundo, o fato de a escola simbolista, a despeito de seu sucesso na solução de determinados tipos de problemas, não ter conseguido avanços significativos na resolução de alguns problemas simples para um ser humano.

Em 1984, foi publicado o livro de Braitenberg, *Vehicles: Experiments in Synthetic Psychology*. Neste livro Braitenberg defende o *princípio do desempenho auto-organizado, direcionado a objetivo*: obtém-se o melhor entendimento de um problema complexo pela síntese de mecanismos elementares putativos do que por uma análise de cima para baixo (*top-down*).

Em 1986, Rumelhart, Hinton e Williams desenvolveram o algoritmo de retropropagação (back-propagation). Naquele ano foi publicado o célebre livro em dois volumes, *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*,

editado por Rumelhart e McClelland exercendo forte influência na utilização da aprendizagem por retropropagação, que emergiu como o algoritmo de aprendizagem mais popular para o treinamento de perceptrons de múltiplas camadas. Na verdade, segundo HAYKIN (2001) a aprendizagem por retropropagação foi descoberta de modo independente em outros dois lugares, na mesma época (Parker, 1985; LeCun, 1985). Após a descoberta do algoritmo de retropropagação em meados dos anos 80, revelou-se que o algoritmo havia sido descrito anteriormente por Werbos na sua tese de doutorado em Agosto de 1974 em Universidade de Harvard.

Em 1988, Linkser descreveu um novo princípio para a auto-organização em uma rede perceptiva (LINKSER, 1988a), que concebe o princípio para preservar o máximo de informação sobre os padrões de atividade de entrada, sujeito a limitações como as conexões sinápticas e o intervalo dinâmico das sinápses.

No início dos anos 90, Vapnik e co-autores inventaram uma classe de redes de aprendizagem supervisionada poderosa do ponto de vista computacional, chamada de máquinas de vetor de suporte, para ser utilizada em reconhecimento de padrões, regressão e problemas de estimação de densidade (BOSER *Et al.*, 1998).

Hoje, pela habilidade em identificar sistemas complexos, fortemente não lineares, as Redes Neurais têm sido utilizadas, com muito êxito, em aplicações na área da Engenharia Química. Há vários trabalhos publicados hoje em dia, como, por exemplo, na revista *Computers & Chemical Engineering*¹ com uma edição exclusiva, apresentando treze artigos sobre redes neurais.

Segundo BRAGA et al. (2000), hoje determinados tipos de redes podem resolver qualquer problema. E ainda segundo BISHOP (1994), atualmente aplicações de redes neurais podem ser encontradas em produtos eletrônicos de consumo, desde câmeras de vídeo até máquinas de lavar “inteligentes”.

Talvez mais do que qualquer outra publicação, o artigo de 1982 de Hopfield e o livro em dois volumes de 1986 de Rumelhart e McClelland foram as publicações mais

¹ *Computers & Chemical Engineering*. Oxford, v. 16, n. 4, 423p, April 1992. (Neural Network Applications in Chemical Engineering).

influentes, responsáveis pelo ressurgimento do interesse em redes neurais nos anos 80. As redes neurais certamente trilharam um longo caminho desde os dias iniciais de McCulloch e Pitts. De fato, elas se estabeleceram como um tema interdisciplinar com raízes profundas em neurociências, psicologia, matemática, ciências físicas e engenharia.

2.1.3- Neurônios Biológicos

O modelo das Redes Neurais foi desenvolvido a partir do cérebro humano e portanto uma breve discussão sobre o funcionamento do cérebro humano é necessária para se entender a associação à rede neural artificial.

O neurônio é a unidade biológica fundamental do cérebro e são divididas em três seções: o corpo da célula, os dendritos e o axônio, cada uma com funções específicas, porém complementares. O Corpo do neurônio mede apenas alguns milésimos de milímetros, e os dendritos apresentam poucos milímetros de comprimento. O Axônio, contudo, pode ser mais longo e, em geral, tem calibre uniforme.

O cérebro humano contém 10^{11} neurônios, sua célula fundamental. Cada um destes neurônios processa e se comunica com milhares de outros continuamente e em paralelo.

O cérebro humano é responsável pelo que se chama de emoção, pensamento, percepção e cognição, assim como pela execução de funções sensoriomotoras e autônomas. Além disso, sua rede de nodos tem a capacidade de reconhecer padrões e relacioná-lo, usar e armazenar conhecimento por experiência além de interpretar observações.

No cérebro, bilhões de células nervosas interligam-se através de dendritos, que se comunicam com os demais pelas junções conhecidas como sinapses. A representação esquemática de um neurônio biológico está na Figura 2.1. Quando um estímulo, ou uma combinação de estímulos atinge um neurônio, através dos dendritos de outros neurônios, conectados neste por sinapses, dependendo da qualidade, ou intensidade desse estímulo o neurônio receptor pode tornar-se excitado, e, por seu axônio, repassar esse estímulo a seus dendritos, que irão estimular os milhares de neurônios a eles conectados. As sinapses funcionam como válvulas, e são capazes de controlar a transmissão de impulsos – isto é, o

fluxo da informação – entre os neurônios na rede neural. O efeito das sinapses é variável, e é esta variação que dá ao neurônio capacidade de adaptação. Os sinais oriundos dos neurônios pré-sinápticos são passados para o corpo do neurônio pós-sináptico, onde são comparados com os outros sinais recebidos pelo mesmo. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nodos pós-sinápticos). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro.

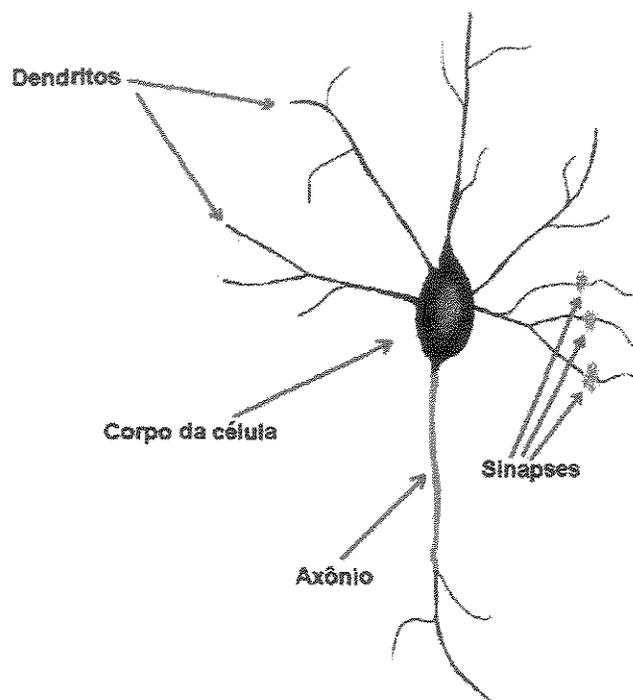


Figura 2.1 - Esquema de um neurônio

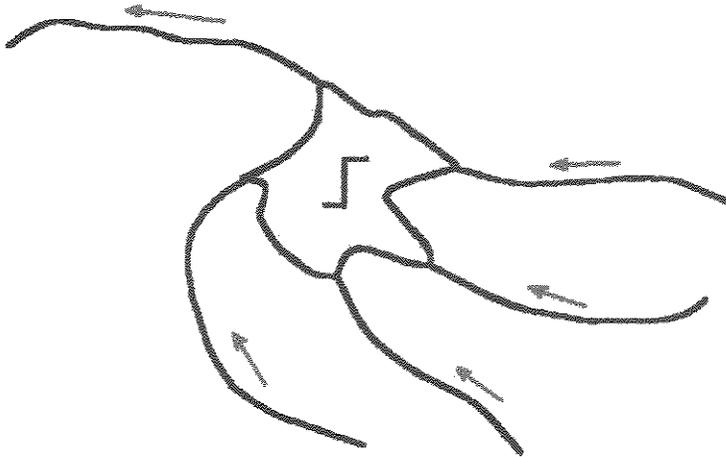


Figura 2.2 - Se a soma dos sinais de entrada excederem a um valor de excitação o neurônio "dispara"

Um neurônio pode ser considerado como uma unidade limiar – um elemento de processamento que recolhe as entradas e produz uma saída somente se o somatório das entradas exceder a um valor de limite interno, como ilustrado na Figura 2.2.

A comunicação no cérebro pode ser de dois tipos: sinais químicos através das sinapses e sinais elétricos dentro do neurônio. A ação da membrana dos neurônios é que cria a habilidade de produzir e transmitir esse dois tipos de sinais. Esta membrana contém proteínas como funções específicas de, por exemplo, mover íons, apressar reações químicas, etc. A membrana que envolve o exterior do corpo do neurônio tem a capacidade de gerar impulsos nervosos (elétricos), uma função vital do sistema nervoso e central à sua capacidade computacional. O corpo, por sua vez, combina os sinais recebidos, e, se o valor resultante for acima do limiar de excitação do neurônio, um impulso elétrico é produzido e propagado através do axônio para os neurônios seguintes.

Há uma diferença de potencial (em volts) entre o interior e o exterior do neurônio, ocasionado pela diferença entre concentração de potássio (interna à célula) e de sódio (externa à célula). A concentração de íons de potássio dentro da célula cria um potencial elétrico de -70 mv (potencial de repouso) em relação ao exterior. Para que a célula dispare, produzindo um potencial de ação (impulso nervoso), é preciso que os impulsos das sinapses reduzam este nível para cerca de -50 mv. Nesse momento, o fluxo de sódio e de potássio é invertido, e o interior da célula torna-se, subitamente, positivo em relação ao exterior. Esta inversão de polaridade faz com que o impulso nervoso se propague pelo axônio até sua

conexões sinápticas. Quando o impulso chega ao terminal de um axônio, os canais controlados por tensão de abrem. Isto libera as moléculas neurotransmissoras para dentro da clave sináptica (região entre a membrana pré-sináptica e a membrana pós sináptica, e o processo continua no neurônio seguinte. O tipo de neurotransmissor (há em torno de 100 tipos diferentes de neurotransmissores) liberado determinará a polarização ou a despolarização do corpo do neurônio seguinte. De acordo com o tipo de neurotransmissor liberado, a sinapse poderá ser inibitória ou excitatória. A contribuição de todos os neurônios pré-sinápticos na polarização do neurônio pós-sináptico determinará se este irá ou não gerar um impulso nervoso. Portanto, o percentual de disparo de um neurônio é determinado pelo acúmulo de um número grande de entradas inibitórias e excitatórias, medido pelo corpo da célula em um pequeno intervalo de tempo. Depois de gerar um impulso, o neurônio entra em um período de refração (período em que o axônio não pode ser novamente estimulado), durante o qual retorna ao seu potencial de repouso enquanto se prepara para a geração de um novo impulso.

2.1.4- Neurônios Artificiais

O modelo de neurônio proposto por McCulloch e Pitts é uma simplificação do que se sabia então a respeito do neurônio biológico. Sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n (que representam os dendritos) e apenas um terminal de saída y (representando o axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados $w_1, w_2, w_3, \dots, w_n$, cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por $x_i w_i$.

O Neurônio computacional funciona de maneira idêntica ao biológico, possuindo várias entradas e saídas. Saídas estas que são conectadas ao elemento adjacente através e conexões ponderadas de maneira similar às ramificações sinápticas. Cada conexão possui um peso correspondente que modifica os sinais de entrada. Os sinais ponderados são somados, modificados por uma função de ativação e enviados da saída para a entrada do próximo neurônio.

O modelo de neurônio mostrado na Figura 2.3 corresponde a um modelo matemático de um neurônio humano e possui como sinal de entrada o vetor:

$$X = (x_1, x_2, \dots, x_i, \dots, x_n),$$

onde x_i representa a ativação do i -ésimo neurônio.

Associado a cada neurônio existe um valor ajustável, w_{ij} (peso). Os pesos ligados ao j -ésimo neurônio (j), formam o vetor: $w_j = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{nj})$, onde o peso w_{ij} representa o valor da ligação do neurônio i para o neurônio j . Ocasionalmente existe um parâmetro adicional (bias) θ_j modulado pelo peso w_{0j} compondo as entradas dos neurônios, que serve para dar estabilidade aos cálculos, especialmente durante a fase de treinamento da rede. Este modelo é, também, referenciado como Elemento de Processamento – PE.

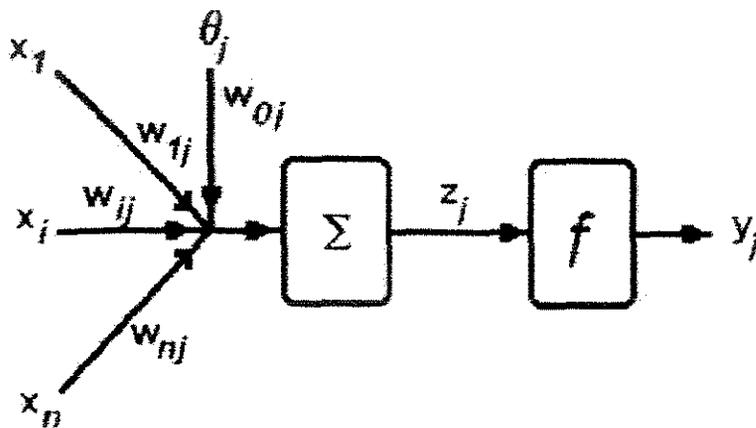


Figura 2.3 - Representação matemática de um neurônio

A somatória das entradas de um neurônio (que também é referenciado como nó da rede, ou elemento de processamento) balanceadas por seus respectivos pesos e mais um bias, também ponderado, resulta no valor de entrada do j -ésimo neurônio, z_j . Este valor sofre uma função de ativação, gerando a saída efetiva do neurônio, y_i , como representado a seguir na representação matemática do processamento realizado por um nó (elemento de processamento), equação 2.1:

$$z_j = \left(\sum x_i \cdot w_{ij} \right) + \theta_j \cdot w_{0j} \quad y_j = f(z_j) \quad (2.1)$$

Existem várias funções de ativação propostas e em uso [PHAM,1995], como na Tabela 2.1, que apresenta as funções comumente mais utilizadas.

Tabela 2.1 - Funções de Ativação

| Tipo de Funções | Funções |
|----------------------|--|
| Linear | $f(s) = s$ |
| De Ativação | $f(s) = +1$ se $s > s_t$ $f(s) = -1$ nos demais casos |
| Sigmóide | $f(s) = 1/(1 + \exp(-s))$ |
| Tangente Hiperbólica | $f(s) = (1 - \exp(-2s))/(1 + \exp(2s))$ |
| Função base radial | $f(s) = \exp(-s^2/\beta^2)$ |

Um neurônio biológico dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação (threshold). O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_{ij}$ recebidos pelo neurônio (soma ponderada) e decide se o neurônio deve ou não disparar (saída igual a 1 ou a 0) comparando a soma obtida ao limiar ou threshold do neurônio. No modelo MCP (McCulloch e Pitts), a ativação do neurônio é obtida através da aplicação de uma “função ativação”, que ativa ou não a saída, dependendo do valor da soma ponderada das suas entradas.

As Redes Neurais colocam-se como uma técnica muito interessante de ser empregada em processos químicos industriais, que apresentam forte não linearidade, dificultando a obtenção de modelos fenomenológicos, a ponto de, em alguns casos, conseguir-se apenas um modelo extremamente simplificado, válido em condições muito

específicas e pela habilidade em contornar essa dificuldade em obter um modelo fenomenológico representativo do processo. O sucesso na aplicação dependerá, entretanto, do adequado conhecimento do processo e de suas variáveis significativas, além da disponibilidade de uma boa base de dados, em quantidade e qualidade, representando corretamente o domínio onde a solução será modelada, uma vez que as redes não possuem habilidade para extrapolar adequadamente para fora do domínio dos dados a ela apresentados para treinamento.

2.1.5- Tipos de Redes Neurais

Como existem diversas maneiras de interligar os nós das redes, inúmeras estruturas têm sido estudadas e publicadas, porém, somente algumas realmente têm tido aplicações práticas em processos químicos industriais, e é destas que falaremos com mais detalhes.

A Rede Neural pode ser classificada quanto à *estrutura* da ligação entre seus nós, e também ser diferenciada pelo *tipo de algoritmo de treinamento* utilizado.

2.1.5.1- Tipos de Redes Neurais, quanto à sua estrutura

Cada elemento de processamento (nó/neurônio) executa um processo de cálculo em suas entradas para gerar sua saída. As entradas de um neurônio geralmente são tratadas uniformemente, e muitas vezes, as saídas de um elemento de processamento são idênticas. Cada conexão pode, também, executar um processo de cálculo simples, mas opera somente um sinal.

Fica claro que o resultado calculado por uma Rede Neural irá depender da topologia (estrutura) dessa rede, da natureza dos cálculos efetuados em cada neurônio, da natureza dos cálculos realizados em cada conexão e dos valores de cada parâmetro ajustável nos neurônios e nas conexões. Assim, as Redes Neurais podem ser totalmente projetadas, desde seu início, para executar um cálculo muito específico.

O tipo de uma determinada RNA é um importante aspecto, uma vez que pode definir quais as classes de problemas que poderão ser abordados pela rede. Um exemplo disto é a rede MCP (McCulloch Pitts) com uma única camada de nodos, que são hábeis para resolver problemas linearmente separáveis apenas. Como uma rede neural é composta por nodos ou neurônios, a maneira pela qual eles estão interconectados, bem como a quantidade dos mesmos, define a estrutura da rede. O modo como são ajustados os pesos entre as conexões é definido pelo algoritmo de treinamento. Assim sendo, os parâmetros importantes que participam da definição de uma estrutura são, segundo BRAGA et al. (2000), os seguintes:

- número de camadas da rede;
- número de nodos em cada camada;
- tipo de conexão entre os nodos e topologia da rede.

Alguns tipos de arquiteturas de RNAs são apresentados a seguir na Figura 2.4:

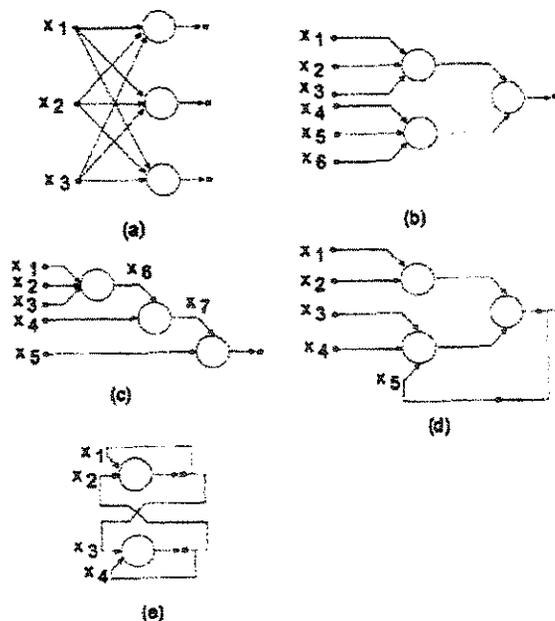


Figura 2.4 - Exemplos de Arquiteturas de RNAs

De acordo com o que pode ser observado da Figura 2.4, as redes podem ser de uma única camada, ou seja, apenas com um nó entre qualquer entrada e saída rede (Figura 2.4 a,

e), ou ainda, de múltiplas camadas, quando existem mais de um neurônio entre alguma entrada e saída da rede (Figura 2.4 b,c,d).

Estruturalmente as redes podem ser divididas em Acíclicas, Feedforward, ou Cíclicas, Feedback. Na primeira, em nenhum momento a saída de um neurônio na i -ésima camada é utilizada para realimentar nodos em camadas de índice menor ou igual a i , como pode ser observado da Figura 2.4 a,b,c. Quando a rede é cíclica, a saída de algum neurônio da i -ésima camada é retornada como entrada para os nodos das camadas inferiores (Figura 2.4 d,e).

Considerando uma estrutura em camadas, caso todos os neurônios de uma camada, de entrada, por exemplo, que fazem a interface com o ambiente externo, recebendo os dados, estiverem ligados unicamente com os neurônios da camada seguinte, ou seja, a camada intermediária, que alguns autores preferem denominar de camada escondida (*hidden*), e todos os neurônios desta camada intermediária somente estiverem ligados aos neurônios da camada de saída, e estes só ao ambiente externo, essa estrutura é conhecida como rede *feedforward*, ou seja, as ligações só permitem a passagem de estímulos aos neurônios da camada seguinte.

No caso de existirem ligações que permitam que os estímulos possam ser distribuídos para os neurônios da mesma camada, ou retornando para os neurônios da camada anterior, este tipo de estrutura é conhecido como rede *recorrente*.

- Redes Recorrentes

As redes recorrentes possuem conexões que permitem que o fluxo de sinais possa ir tanto em direção aos neurônios de saída, quanto em direção aos neurônios de entrada, ou lateralmente, em direção aos neurônios da mesma camada.

Esse tipo de estrutura faz com que a rede tenha uma memória dinâmica: suas saídas, em determinado instante, refletem a entrada atual da rede, assim como entradas e saídas anteriores. Por isso, este tipo de rede é aplicável onde são necessárias suas propriedades de “memória”, de curto ou longo prazo. Dentre as diversas possibilidades de estruturas de rede recorrente, as mais conhecidas e utilizadas são: rede Hopfield, rede Elman, rede Jordan e rede Kohonen.

- Redes Feedforward

Nestas redes os neurônios geralmente são agrupados em camadas, e os sinais de entrada evoluem dentro da rede, camada por camada, até a camada de saída, de forma unidirecional. As redes feedforward podem, muito naturalmente, executar o mapeamento entre o espaço de entrada e o espaço de saída. Desta maneira, num dado instante, a saída é somente função da entrada naquele instante. Entre os tipos de rede feedforward estão as MLP – Multilayer Perceptron, as redes LVQ – Learning Vector Quantization e as redes GMDH – Group Method of Data Handling. Entre essas redes, a que melhor se aplica aos processos químicos industriais é a MLP, que será utilizada na modelagem, objeto desta dissertação. FRANCISCO (2000), indica mais algumas aplicações industriais utilizando redes neurais.

Di Massimo et al. (1992) informam que a rede neural, *feedforward*, embora possua uma estrutura muito elementar, demonstra que qualquer função não linear, contínua, pode ser representada por uma rede neural *feedforward* com uma topologia consistindo de duas camadas intermediárias.

- Rede MLP

A MLP é, talvez, o tipo de rede *feedforward* melhor conhecido. A principal razão para isso é a sua habilidade tanto em modelar relações funcionais simples, como relações muito complexas. Os neurônios da camada de entrada agem somente como estocagem para distribuição dos sinais de entrada x_i aos neurônios da camada intermediária. Na estrutura MLP cada neurônio j da camada intermediária efetua o somatório de seus sinais de entrada, após serem ponderados com o peso de suas respectivas conexões w_{ij} com a camada de entrada, e calcula a saída y_j , como uma função f desse somatório, como pode ser visto na figura 2.5.

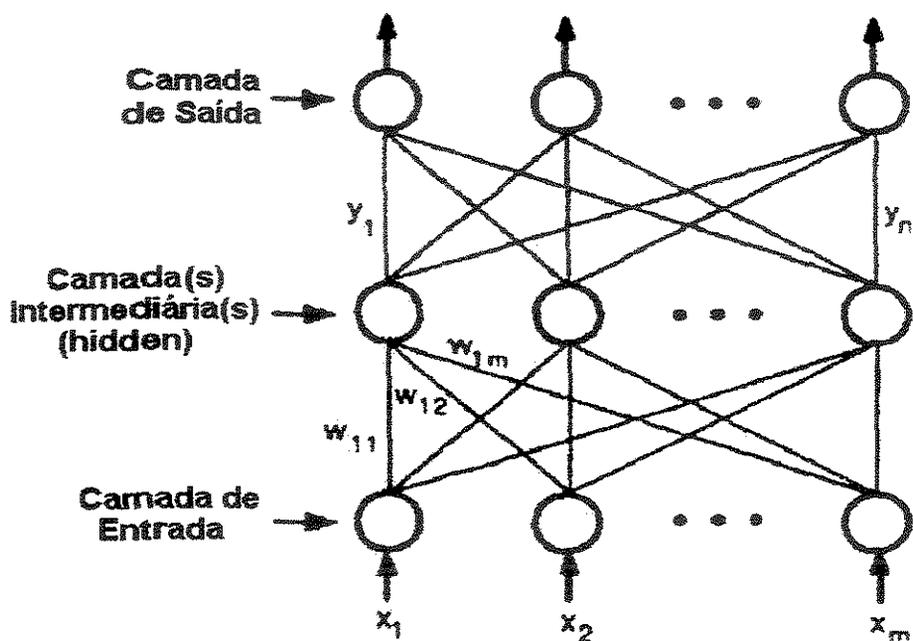


Figura 2.5 - Estrutura de uma Rede Neural tipo MLP

Assim como apresentado na Equação 2.1, a saída dos neurônios da camada de saída pode ser calculada pela equação 2.1(a) seguinte:

$$y_i = f\left(\sum w_{ji} x_j\right) + w_{0i} \quad 2.1(a)$$

2.1.5.2- Tipos de Redes quanto ao algoritmo de treinamento

As redes neurais artificiais possuem a capacidade de aprender por exemplos e fazer interpolações e extrapolações do que aprenderam. Um conjunto de procedimentos bem definidos para adaptar os parâmetros de uma RNA para que a mesma possa “aprender” uma determinada função é chamado de algoritmo de aprendizado ou treinamento.

Existem vários tipos de algoritmos de treinamento. Cada algoritmo ou método apresenta suas vantagens e desvantagens que diferem entre si, basicamente, pela maneira pela

qual é feito o ajuste de pesos. Portanto a utilização de um método ou outro depende exclusivamente do caso em estudo.

Segundo BRAGA et al. (2000): “Aprendizagem é o processo pelo qual os parâmetros de uma rede neural são ajustados através de uma forma continuada de estímulo pelo ambiente no qual a rede está operando, sendo o tipo específico de aprendizagem realizado definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros”.

Segundo PHAN (1995) há dois métodos de aprendizado básicos de uma RNA: aprendizado supervisionado e aprendizado não-supervisionado.

O método de aprendizado supervisionado é o mais utilizado para o treinamento de redes neurais. Este método pode ser implementado basicamente de duas formas: on-line e off-line. No primeiro, a rede deve estar em processo contínuo de treinamento, pois os dados mudam continuamente. No sistema, off-line o conjunto de treinamento é fixo, e uma vez obtida a solução para a rede esta deve permanecer fixa (BRAGA et al., 2000).

No aprendizado supervisionado ocorre a interação de um supervisor externo (professor), daí o nome do método. Os valores de entrada e saída são fornecidos por este supervisor, conforme ilustra a Figura 2.6. O objetivo é encontrar uma ligação entre os pares de entrada e saída, e para tanto é feita a comparação entre a resposta calculada pela rede e a desejada. A informação sobre o erro da resposta atual é complementada pelo supervisor, indicando o comportamento da rede. Em seguida os pesos das conexões são ajustados. A soma dos erros quadráticos de todas as saídas é geralmente utilizada tanto para medir o desempenho da rede, como função a ser minimizada, quanto como medida de desempenho da mesma.

A regra delta e a sua generalização, o algoritmo Backpropagation, para redes de múltiplas camadas são os mais conhecidos algoritmos de aprendizado supervisionado tendo a desvantagem da impossibilidade da rede aprender novas estratégias para situações não apresentadas dentro dos exemplos de treinamento.

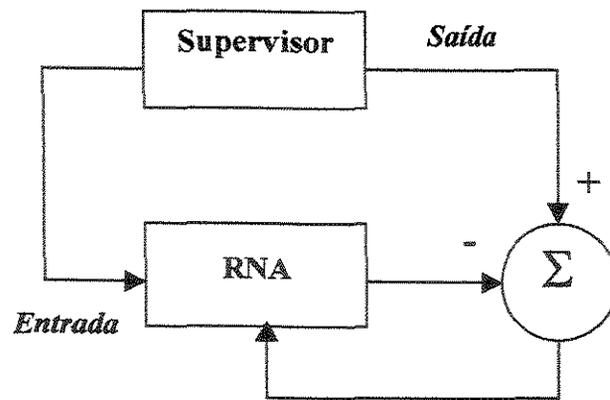


Figura 2.6 - Aprendizado Supervisionado.

Já o aprendizado não-supervisionado, como o próprio nome diz, não possui um supervisor para acompanhar o processo de treinamento. Este tipo de aprendizado necessita, de forma geral, somente padrões ou dados de entrada, conforme ilustrado na Figura 2.7 respostas.

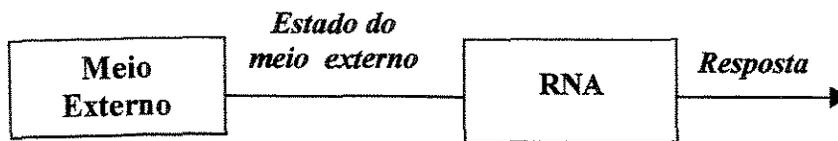


Figura 2.7 - Aprendizado não-supervisionado.

As principais formas para implementação desta metodologia são as seguintes: Aprendizado hebbiano; Modelo de Linsker; Regra de Oja; Regra de Yuille; Aprendizado por competição.

Existe ainda um terceiro sistema de aprendizado, aprendizado por reforço, trata-se apenas de um caso particular do aprendizado supervisionado. Segundo BRAGA (2000), o que diferencia esse método daqueles discutidos é a medida de desempenho utilizado em cada um. Enquanto no primeiro esta medida é baseada em um conjunto de respostas desejadas

usando um critério de erro conhecido, no aprendizado por reforço é realimentado à rede somente um simples sinal de reforço (sim/não), sem qualquer informação instrutiva que melhore o desempenho da rede. De uma maneira geral, se uma ação tomada pelo sistema de aprendizagem não é seguida de estados satisfatórios, a tendência do sistema de produzir esta ação é enfraquecida.

O aprendizado por reforço ainda pode ser “associativo” ou “não-associativo”. No primeiro, um mapeamento, na forma estímulo/ação, deve ser aprendido com as informações adicionais fornecidas pelo meio. Já no segundo, o sistema não recebe informações adicionais, funcionando apenas com a avaliação do desempenho da rede (sim/não).

2.1.5.2.1- O Algoritmo Back-Propagation

O Algoritmo back-propagation, também conhecido como algoritmo do *gradiente descendente* é o algoritmo mais comumente utilizado para o treinamento de redes MLP.

Historicamente derivado da regra delta padrão proposta por Widrow e Hoff, este algoritmo foi um dos principais responsáveis pela volta do interesse das RNAs durante a década de 80. Isto porque combinado com uma topologia adequada de rede, fornece bons resultados para qualquer problema implementado.

O treinamento com o *backpropagation*, utilizando pares de exemplos (entrada, saída desejada), ocorre em duas fases. Na primeira, conhecida como fase *forward*, os padrões de entrada são apresentados à rede, ou melhor, são apresentados a camada de entrada, seguindo pelas camadas intermediárias, até que uma resposta seja produzida pela camada de saída. Na segunda, fase *backward*, a resposta obtida da rede é comparada com a saída desejada e caso não esteja correta, o erro é calculado e retropropagado da camada de saída à entrada, ajustando assim os pesos das conexões nas unidades das camadas internas tendo continuidade até que a rede atinja um critério de parada pré-definido.

Veja algoritmo back-propagation como segue:

O Treinamento da Rede é realizado procurando-se por valores de pesos que minimizem a função erro, que é geralmente a soma dos erros quadráticos:

$$E = \frac{1}{2} \sum_p \sum_{i=1}^k (d_i^p - y_i^p)^2 \quad (2.2)$$

Sendo:

E a medida de erro total,

k o número de unidades de saída,

p é o número de padrões,

y é a i -ésima saída gerada e

d a i -ésima saída desejada.

A função erro corresponde à superfície de erro sobre as coordenadas de peso, geometricamente (ver Figura 2.8). O treinamento da rede consiste em se procurar um ponto mínimo na superfície do erro. Porém, existem pontos de mínimo que não correspondem ao mínimo global, são pontos conhecidos como mínimos locais, e podem em alguns casos fornecer a resposta de treinamento para a rede. Entretanto, o objetivo principal do treinamento é a procura pelo ponto de mínimo global.

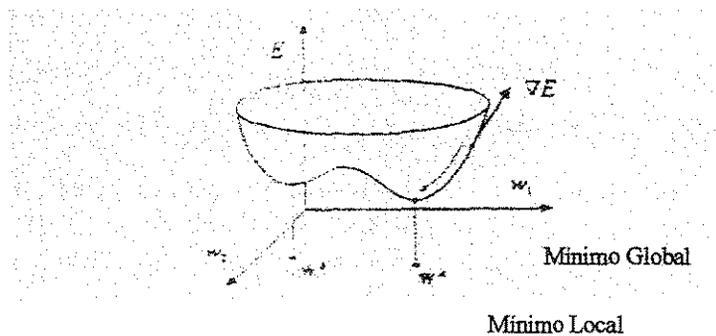


Figura 2.8 - Superfície da função erro $E(w)$

BUFO (2000), lista uma série de técnicas de minimização do erro global, aplicadas às redes neurais tipo *backpropagation*.

Nas redes multicamadas a função erro é uma função altamente não-linear dos pesos e o ponto de mínimo é obtido iterativamente, tendo um peso aleatório inicial.

O algoritmo *backpropagation* é muito eficiente para realizar a distribuição dos pesos entre as conexões das funções não lineares utilizadas na rede MLP. Este método é obtido a partir da função soma dos erros quadráticos.

O erro descrito na equação 2.2 pode ser dividido em:

$$E = \sum_p E^p, \quad E^p = \frac{1}{2} \sum_{j=1}^k (d_j^p - y_j^p)^2 \quad (2.3)$$

(2.3a)

(2.3b)

Simplificando a notação, será omitido o índice p .

É sugerido pela regra delta que a variação dos pesos seja definida de acordo com o gradiente descendente, ou seja, o treinamento da rede começa com um peso aleatório, como já foi dito, e procede com pequenas variações no vetor peso com o objetivo de mover-se na direção do gradiente negativo da função erro, até que o vetor peso atinja um mínimo local ou global. Desta forma,

$$\Delta w_{ji} \propto - \frac{\partial E}{\partial w_{ji}} \quad (2.4)$$

Para definir como cada um dos pesos de cada nodo da rede deve ser ajustado de forma a diminuir o erro total gerado pela rede, a derivada encontrada na última equação deve ser dividida como o produto de duas partes, a primeira refletindo as mudanças no erro como

uma função de mudanças em ξ , definido como $net_j = \xi = \sum_{i=1}^n x_i \cdot w_{ji}$, e a segunda parte

representando os efeitos da mudança de um peso particular em ξ . Utilizando a regra da cadeia,

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial \text{net}_j} \cdot \frac{\partial \text{net}_j}{\partial w_{ji}} \quad (2.5)$$

A segunda parte da equação 2.5 pode ser calculada facilmente:

$$\frac{\partial \text{net}_j}{w_{ji}} = \frac{\partial \sum_{i=1}^n x_i \cdot w_{ji}}{w_{ji}} = x_i \quad (2.6)$$

A primeira parte do lado direito da equação 2.5 mede o erro do nodo j , e geralmente é abreviada para δ_j .

$$\delta_j = \frac{\partial E}{\partial \text{net}_j} \quad (2.7)$$

Novamente utilizando a regra da cadeia;

$$\delta_j = \frac{\partial E}{\partial \text{net}_j} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial \text{net}_j} \quad (2.8)$$

Onde,

$$\frac{\partial y_j}{\partial \text{net}_j} = \frac{\partial f(\text{net}_j)}{\partial \text{net}_j} = f'(\text{net}_j) \quad (2.9)$$

O cálculo da 1ª derivada no lado direito da equação 2.8 que utiliza o erro, depende da camada onde o nodo j se encontra. Quando o nodo estiver na última camada, o erro pode ser definido utilizando a equação 2.3-b.

$$\frac{\partial E}{\partial y_j} = \frac{\partial \left(1/2 \sum_{i=1}^k (d_i - y_i)^2 \right)}{\partial y_j} = (d_i - y_i) \quad (2.10)$$

A equação 2.10 é a mesma fórmula da regra delta original. Agora, substituindo os dois termos do lado direito da equação 2.8:

$$\delta_j = (d_j - y_j)f'(net_j) \quad (2.11)$$

Quando o nodo não for de saída, a regra da cadeia é utilizada para escrever,

$$\frac{\partial E}{\partial y_j} = \sum_{l=1}^M \frac{\partial E}{\partial net_l} \frac{\partial net_l}{\partial y_j} = \sum_{l=1}^M \frac{\partial E}{\partial net_l} \frac{\partial \sum_{i=1}^n w_{li} y_i}{\partial y_j} = \sum_{l=1}^M \frac{\partial E}{\partial net_l} w_{jl} \quad (2.12)$$

Onde,

$$\sum_{l=1}^M \frac{\partial E}{\partial net_l} w_{jl} = \sum_{l=1}^M \delta_l w_{jl} \quad (2.13)$$

Substituindo novamente os dois termos do lado direito da Equação 2.8, é obtido o erro definido para os nodos das camadas intermediárias.

$$\delta_j = f'(net_j) \sum_l \delta_l w_{lj} \quad (2.14)$$

Assim, generalizando a fórmula de ajuste de pesos proposta na Equação 2.4,

$$\Delta w_{ji} = \eta \delta_j x_i \quad (a) \quad (2.15)$$

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j(t) x_i(t) \quad (b)$$

Portanto se o nodo for de saída, o erro será definido pela Equação 2.11, caso contrário, δ_j será dado pela Equação 2.14.

Muitos algoritmos de aprendizado supervisionado foram desenvolvidos para treinar redes do tipo MLP (Multilayer Perceptron), o mais conhecido entre todos é o algoritmo backpropagation, pois a maioria dos métodos de aprendizado usa variações deste.

O treinamento com o *backpropagation*, utilizando pares de exemplos (entrada, saída desejada), ocorre em duas fases. Na primeira, conhecida como fase *forward*, os padrões de entrada são apresentados à rede, ou melhor, são apresentados a camada de entrada, fluindo pelas camadas intermediárias, até que uma resposta seja produzida pela camada de saída. Na segunda, fase *backward*, a resposta obtida da rede é comparada com a saída desejada. Caso não esteja correta, o erro é calculado e retropropagado da camada de saída à entrada, ajustando assim os pesos das conexões nas unidades das camadas internas. Este procedimento tem continuidade até que a rede atinja um critério de parada pré-definido.

Segundo OLIVEIRA (2000), na modelagem de processos químicos o uso de uma única camada interna tem-se mostrado suficiente. Quando há necessidade de modelos mais complexos, o ajuste do número de neurônios da camada intermediária geralmente é suficiente.

Normalmente os dados de entrada, usados para treinamento da rede, são gerados com ruídos inerentes do próprio processo. No caso da rede ser sobre treinada, os pesos identificam e armazenam esses ruídos, e assim, com a introdução de novos dados, ela não consegue representa-los adequadamente.

Segundo BHAT & McAVOY (1992) o sobreajuste é uma condição que ocorre quando a razão entre os parâmetros do modelo e os dados do conjunto de testes é grande. Com um grande número de parâmetros, o erro da predição no conjunto de treinamento pode ser grandemente minimizado porque cada parâmetro é ajustado para reproduzir parte dos dados de treinamento. Dessa maneira, o modelo reproduz os dados de treinamento com precisão, mas não necessariamente captura o comportamento global do processo. Estes modelos sobreajustados devem ser evitados a todo custo. Atualmente as redes *backpropagation* são os paradigmas de redes mais utilizados para a resolução de problemas de engenharia química. Apesar de sua ampla aplicação não há uma metodologia para a determinação da estrutura da rede, para uma particular aplicação de mapeamento. A falta de um procedimento tem resultado numa tendência do uso de redes muito maiores que o necessário. Esses modelos possuem parâmetros excessivos, ou pesos, e muitas vezes memorizam os dados de treinamento, causando dificuldades no manuseio de dados não apresentados à rede. Desta maneira, é importante o uso de redes que tenham a mais simples

estrutura, ou seja, utilize o mínimo de pesos e nós, e indicando que as redes BP com menor complexidade podem ser obtidas de 2 maneiras:

1. utilizando redes com o menor número possível de nós;
2. utilizando redes com conexões esparsas.

HASSIB & STORK (1993) investigaram o uso de todas as derivadas de segunda ordem da função erro para redução da rede neural (ou seja, removendo os pesos sem importância, de uma rede treinada) de modo a melhorar a generalização, simplificar as redes, reduzir as necessidades computacionais, de memória e de armazenamento, aumentar a velocidade de retreinamentos das redes, e em alguns casos, permitir a extração de regras. Para isso propuseram o método “*Optimal Brain Surgeon*” – OBS, afirmando ser melhor que os métodos baseados em magnitude e que o método “*Optimal Brain Damage*” (FUJITA, 2001).

Para GONTARSKI (2000), na aplicação da análise de significância a rede deve estar devidamente treinada e gerando bons resultados mesmo para condições não utilizadas no treinamento. Ele utilizou o critério de acrescentar e descontar 10% ao valor da média para se ter os níveis, inferior e superior, ao aplicar a metodologia de Planejamento de Experimentos, considerando um dos procedimentos fundamentais para o sucesso da análise a escolha dos fatores a analisar e os respectivos níveis de variação. Aplicou a metodologia de experimento fatorial fracionado, para reduzir o número de experimentos necessários à metodologia de Planejamento de Experimentos, uma vez que estudou 9 fatores a dois níveis, requerendo 2^9 experimentos, e conseguiu reduzir para 2^{9-2} experimentos, utilizando o programa MINITAB para auxiliar na execução da análise.

GONTARSKI et al. (2000) propuseram uma análise de sensibilidade baseada no índice de correlação do modelo, R_j , entre os valores experimentais e os valores calculados para a variável j , conforme equação 2.16 a seguir:

$$R_j = \frac{N \left[\sum_1^N (x_i - \mu_x)(y_i - \mu_y) \right]}{\sqrt{\left[N \sum_1^N x_i^2 - \left(\sum_1^N x_i \right)^2 \right] \left[N \sum_1^N y_i^2 - \left(\sum_1^N y_i \right)^2 \right]}} \quad (2.16)$$

$(j = 1, N_v)$

onde N é o número de dados utilizados, x_i é o valor calculado para a j ésima variável, y_i é o valor experimental da j ésima variável, μ_x é a média dos valores calculados pela rede neural, μ_y é a média dos valores experimentais, e, N_V o número de nós de entrada da rede.

A análise de sensibilidade é efetuada pelo estudo do efeito da eliminação de uma variável de entrada no índice de correlação do modelo, em relação ao índice de correlação da rede treinada com todas as variáveis de entrada. Veja Equação 2.17:

$$(\% \text{ influência})_j = \left(1 - \frac{R_j}{R_{CB}}\right) \cdot 100, \quad (j = 1, N_V) \quad (2.17)$$

onde R_{CB} é o índice de correlação entre os valores experimentais e os valores calculados, para o caso base. Como essa análise não permitiu uma conclusão bem definida sobre a importância das variáveis de entrada no resultado desejado, foi efetuada uma nova análise de sensibilidade, agora baseada em pequenos incrementos dados a cada variável de entrada, conforme a equação 2.18 a seguir:

$$V_j^a = V_j^0 \cdot w_j, \quad (j = 1, N_V) \quad (2.18)$$

onde V_j^a é o valor da j ésima variável utilizado no modelo para obter o índice R_j , V_j^0 é o valor da j ésima variável utilizado para o caso base e w_j é o fator aplicado a cada j ésima variável.

Capítulo 3 - DESCRIÇÃO DO PROCESSO

3.1- Histórico da Companhia Níquel Tocantins

As Atividades do Grupo Votorantim, fundado em 1918, iniciaram-se a partir da fábrica de tecidos, localizada em Sorocaba, hoje município de Votorantim, e gradualmente avançaram para outros setores: cimento, alumínio, zinco, aço, produtos químicos, papel transparente, níquel, equipamentos pesados, cal, refratários, papel, açúcar e álcool e agricultura entre outros.

As empresas do Grupo, em sua maioria são voltadas para a produção industrial, essencialmente em função do aproveitamento de recursos minerais.

Este trabalho foi feito na Cia Níquel Tocantins, da qual sua história é muito anterior ao seu próprio nome. Começou em 1929, quando um farmacêutico da pequena localidade de São José do Tocantins, hoje Niquelândia, situada a 400km de Goiânia, no estado de Goiás, intrigado com os torrões verdes que os tatus arrancavam do solo ao cavar suas tocas, decidiu mandar analisá-los. E o laudo do laboratório caracterizou Níquel. Três anos depois, em 1932, foi constituída a Empresa Comercial de Goiás, que em 1942 passaria a se chamar Cia Níquel Tocantins. Embora a empresa detivesse o manifesto para funcionamento da mina desde 1935, nunca chegou a explorá-la, até que em 1957 o seu controle acionário passou para o Grupo Votorantim. O grupo, nessa época, concentrava esforços para conclusão de seu para produção de alumínio, e na década seguinte, investiu na produção de Zinco, mesmo assim, nesse período, desenvolveu paralelamente todos os estudos geológicos e de pesquisa para conhecer profundamente suas reservas minerais de Níquel.

Ao final da década de 60, após concluídos os estudos geológicos, o grupo partiu definitivamente para o desenvolvimento do projeto de níquel. Foram desenvolvido estudos para escolha do processo metalúrgico, seguidos dos testes em escala de laboratório, onde conclui-se que o mais indicado para o tratamento dos minérios, seria através da secagem, redução e lixiviação (separação dos sais através da lixiviação com amônia).

Uma fábrica piloto foi montada em São Paulo, no bairro do Jaguaré, para se conhecer em escala industrial, todos os parâmetros de transformação do minério em carbonato de níquel, com uma tecnologia própria, cujo desenvolvimento, teve como diretriz o

aproveitamento de insumos nacionais e a auto suficiência energética com a utilização de lenha e do carvão vegetal, inclusive para redução de minério.

Para a fase industrial seguinte de transformação de carbonato em níquel metálico, foi adquirido um know-how da empresa finlandesa Outokumpu-oy, que utiliza tecnologia de anodos insolúveis de chumbo (eletrodos, positivos que atraem os íons negativos).

Em meados de 1973 foi aprovado pelo CDI (Conselho de Desenvolvimento Industrial) o projeto para a produção de 5000 toneladas por ano de Níquel Eletrolítico, cujo cronograma previa o início da produção em 1976.

Este cronograma, porém não pode ser cumprido devido ao não recebimento de energia elétrica da linha local, conforme fora inicialmente prometido. Afinal após três anos de discussões a CNT decidiu transferir a etapa de eletrólise para São Miguel Paulista e decidiu gerar sua própria energia.

Outras dificuldades se sucederam, como os constantes atrasos nas entregas de equipamentos por parte dos fornecedores e muitas outras, decorrentes da localização da fábrica numa região distante e totalmente desprovida de infra-estrutura.

Todas as barreiras foram superadas em 1981 e assim sendo a Companhia Níquel Tocantins produziu o primeiro níquel eletrolítico brasileiro.

3.2- Processo Industrial

O Processo industrial de obtenção de níquel e cobalto eletrolíticos da CNT – Companhia Níquel Tocantins – é bem representado pela figura 3.1 (anexo 1).

A etapa de produção de níquel eletrolítico e cobalto eletrolítico inicia-se com o recebimento de carbonato de níquel vindo de Niquelândia – Goiás, como pode-se verificar na figura 3.2 (Anexo 2).

3.2.1- Níquel e Cobalto eletrolítico

O carbonato de níquel é dissolvido no circuito de lixiviação utilizando-se soluções contendo ácido sulfúrico (anolito), obtendo-se uma solução aquosa denominada solução de processo, que iniciará suas etapas de purificação.

No circuito de lixiviação a solução de processo recebe adição de hidróxido de bário para a purificação de chumbo, posteriormente a solução é filtrada para eliminação do chumbo precipitado e todos os materiais insolúveis do carbonato de níquel. Veja Figura 3.3. - Figura do Processo de Lixiviação (Anexo 3).

A solução filtrada da lixiviação recebe no estágio de neutralização nova adição de carbonato de níquel para correção de pH a 5,8, condição favorável para assim haver a precipitação de cobre e ferro e tornar a solução adequada para o circuito de extração de cobalto. Esta solução de processo a pH de 5,8 da neutralização irá alimentar um tanque decantador onde haverá a precipitação de cobre e ferro. A solução de topo do decantador (over-flow) contendo solução de processo parcialmente purificada em cobre e ferro será filtrada para tanque de estocagem.

Esta solução de processo irá alimentar o circuito de extração de Cobalto. Já a solução de fundo do decantador (under-flow) irá alimentar o circuito de tratamento para retirada de cobre e ferro. Veja figura 3.4 - Figura do Processo de Neutralização (Anexo 4).

No circuito de extração de cobalto a solução de níquel contendo vários contaminantes como cobalto, cobre, ferro, magnésio e manganês é contactada com um extratante orgânico (Cyanex 272) seletivo para a retirada de impurezas (cobalto, cobre, zinco, ferro, magnésio, manganês). Nesta etapa a solução de níquel é purificada (mas recebe um incremento de sódio proveniente da reação de extração). Esta solução de níquel purificada passará pelas etapas de acidulação – correção de pH; clarificação – eliminação de material orgânico e Polimento – etapa de segurança para retirada de traço de material orgânico que é realizado com adição de carvão ativado e água oxigenada. Veja Figuras 3.5 – Figura do Processo de Extração de Cobalto e 3.6 – Figura do Processo de Acidificação e Clarificação (Anexos 5 e 6, respectivamente).

A solução de níquel purificada e polida passa a ser chamada de catolito de níquel, estando pronta para alimentar as células de eletrólise de níquel. Figuras 3.7 – Figura do

Processo de Eletrólise de Níquel e 3.8 – Figura do Processo de Eletrólise de Cobalto (Anexos 7 e 8, respectivamente).

Nas células de eletrólise de níquel em função de passagem de corrente elétrica “contínua” o níquel é depositado no catodo (polo negativo). Este níquel eletrolítico é retirado da célula de eletrólise e encaminhado para expedição para ser embalado e vendido.

No anodo (polo positivo) é produzido o ácido sulfúrico; esta solução de níquel contendo ácido sulfúrico que sai da célula é chamada de anolito de níquel, a solução retorna para o início do processo, etapa de lixiviação, onde recebe mais carbonato de níquel para enriquecer a solução em níquel e consumir a acidez .

Parte do anolito também passa pelo circuito de cristalização para a retirada de sódio, voltando posteriormente para a lixiviação. Veja Figura 3.9 – Figura do processo de Cristalização (Anexo 9).

A solução orgânica da extração de cobalto contendo os metais Co, Zn, Fe, Mn, Cu e Mg é lavada e estripada para retirada dos metais, obtendo-se uma solução de cobalto. Esta solução de cobalto é purificada de forma semelhante a solução de níquel, estando pronta para o circuito de eletrólise de cobalto. O anolito de cobalto retorna para o circuito de estripagem. Veja Figura 3.10 – Figura do Processo de Estripagem de Cobalto (Anexo 10).

3.2.2- Processo de Extração de Cobalto e Níquel

Faz-se necessária a abordagem do processo de extração devido ao objetivo deste trabalho ser sua modelagem, para assim criar condições de simulação, haja vista que esta é a função objetivo que poder ser crucial para maximizar e otimizar todos os processo de produção e assim o lucro da empresa .

O Processo de extração de Cobalto e níquel basicamente segue o princípio da extração líquido-líquido. O objetivo desta etapa é a purificação da solução aquosa com níquel contido retirando as impurezas (Co, Cu, Zn, Mn, Mg), fornecendo uma solução mais purificada para a seção de polimento.

O pH é o principal parâmetro desta etapa, pois toda a separação baseia-se na extração seletiva por pH. O pH ideal de trabalho é de 5,0 a 5,5, pois neste pH garantimos que a maior parte do cobalto e também os demais metais passaram para a solução orgânica. A solução orgânica é composta por 87% de Exxsol 13% de Cyanex 272. O Cyanex é um ácido organofosfínico e foi especificamente desenvolvido para separar metais como Cobalto do Níquel pela extração por solvente. Sua introdução no mercado foi em 1982 e usado em planta industrial pela primeira vez em 1985. Em 1995 já era empregado em mais de 10 instalações pelo mundo localizado na Europa, América do Sul, Canadá e África, e hoje é responsável pela produção de 40% do cobalto produzido mundialmente. O Cyanex, ou o Di- 2,4,4 trimetil pentil ácido fosfínico, de fórmula molecular apresentada na Figura 3.11, funciona como um extratante seqüestrando os metais que estão na forma de sulfatos e assim formando ácido sulfúrico, daí a necessidade da adição de soda cáustica. A adição de soda cáustica ocorre automaticamente nos reatores conforme variações de pH. A solução aquosa entra no processo pelo RE-06 com pH de 4.5 a 5.0 e a solução orgânica entra em contra-fluxo pelo reator RE-01 como mostrado na Figura 3.5 (Anexo 7).

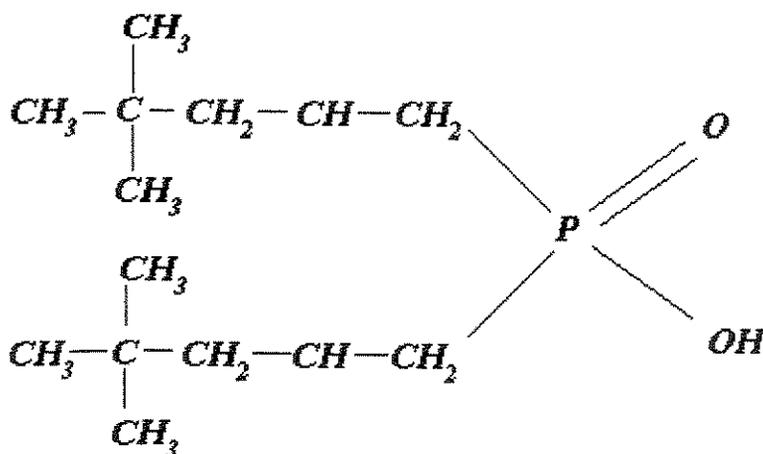


Figura 3.11 – Fórmula Molecular do Cyanex , ou Di- 2,4,4 trimetil pentil ácido fosfínico.

A reação de extração de dá resumidamente da seguinte forma:



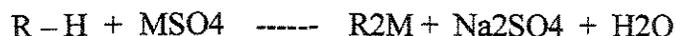
Onde: M.....Metal (pode ser Ni, Co, Zn, Mg, Cu, Fe, etc...)

R...Composto Orgânico - H

O objetivo da extração é deslocarmos o sentido da reação para a direita e isso acontece em função do pH. Durante a reação há liberação de ácido sulfúrico, causando diminuição do pH que desloca o equilíbrio para a esquerda e para evitarmos que isso ocorra adicionamos soda para eliminar ou seja neutralizar o ácido, conforme a reação:



Reação Global:



RICKELTON e McRAE (1996) estudaram o Cyanex comparando-o a outros solventes extratantes, organofosfóricos e organofosfônicos, mostrando a capacidade de utilização e a cinética da extração, bem como a velocidade de extração em operação nas indústrias químicas, provando assim a superioridade do solvente diante outros solventes para extração de Cobalto do Níquel.

Tendo descrito o processo verifica-se a importância da modelagem do processo de extração, pois este precisa de superior refinamento para que os produtos finais sejam de máxima qualidade bem como o consumo de todas as matérias primas, principalmente do

solvente Cyanex, seja tal que a empresa possa maximizar seu lucro e desprender cada vez menos de recursos.

Capítulo 4 - MATERIAIS E MÉTODOS

Este trabalho foi realizado no Laboratório de Modelagem e Simulação de Processos Químicos – LMSPQ, no Laboratório de Sistemas Particulados, da Faculdade de Engenharia Química – FEQ, da Universidade Estadual de Campinas – UNICAMP, em Campinas e na Companhia Níquel Tocantins – CNT do Grupo Votorantim.

4.1- Equipamentos, Reagentes e Ferramentas Computacionais

Para a parte experimental usou-se os seguintes equipamentos e reagentes:

- Espectrofotômetro de Absorção Atômica,
- Cuba de banho eletrostático,
- PHmetro,
- Funil de separação,
- Solvente orgânico Cyanex 272,
- Solução aquosa de Sulfatos (matéria prima para aquisição dos metais) e
- Hidróxido de Sódio

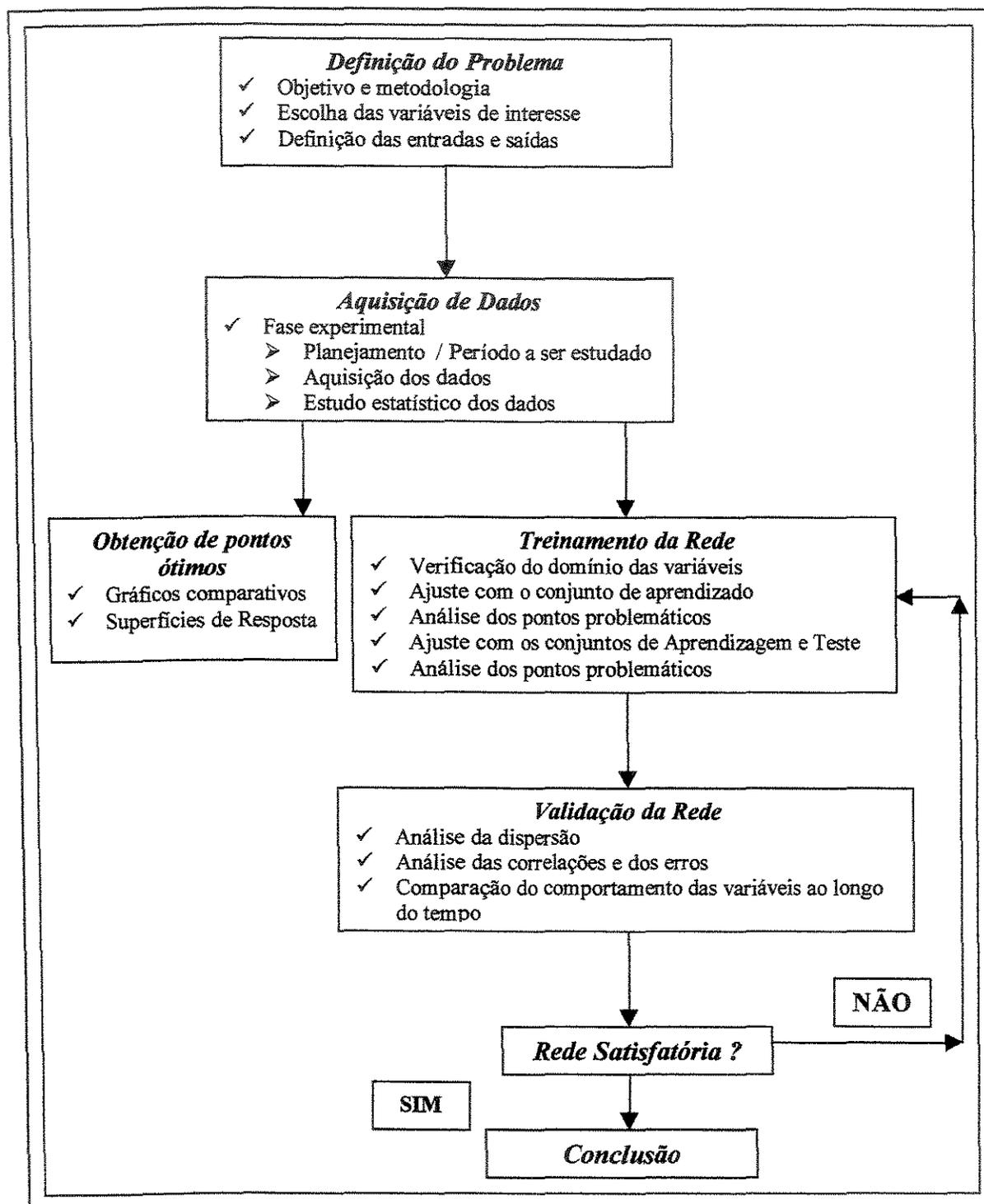
E para a parte computacional foram usados os seguintes pacotes:

- *Minitab Statistical Software*, para o planejamento, otimização dos experimentos e para o pré tratamento dos dados para modelagem.
- *NeuroSolutions Consultants*, para a modelagem aplicando as técnicas de redes neurais artificiais.
- *Statistica Industrial/SPC Solution Software*, para estudo das variáveis e obtenção de pontos ótimos do processo através dos pontos experimentais.

4.2- Metodologia Utilizada

A metodologia desenvolvida neste trabalho é apresentada na figura 4.1 abaixo:

Figura 4.1 – Esquema da Metodologia desenvolvida para estudo



4.2.1- Revisão Bibliográfica

Com a definição do projeto, a modelagem de um processo de extração de Níquel e Cobalto através de solução orgânica, planejou-se e executou-se um estudo bibliográfico sobre o processo e definição de pontos importantes para aquisição de um melhor processo de extração e a melhor forma da repetibilidade deste em laboratório. Também realizou-se estudo sobre as técnicas de redes neurais e suas aplicações.

4.2.2- Estudo do Problema a ser modelado e aquisição dos Dados

As variáveis que caracterizam o processo devem ser consideradas, pois como o procedimento de treinamento se assemelha com o aprendizado do cérebro humano, na falta de informações importantes a rede vai determinar relações não lineares também de forma incompleta, empobrecendo sua capacidade preditiva. É importante também, ter muita atenção para não se utilizar informações redundantes.

A aquisição dos dados em laboratório consistiu na obtenção da quantidade de Níquel e Cobalto extraído da solução aquosa de sulfatos pela solução orgânica de Cyanex 272 simulando parte do processo de obtenção de Níquel e Cobalto metálicos como já descrito no capítulo 3, *Descrição do Processo – Extração, da VOTORANTIM*. Esta simulação laboratorial baseou-se na obtenção de curvas de extração dos metais de maior interesse (níquel e cobalto) variando-se a concentração de solução orgânica, a temperatura e o pH pois estas são as variáveis que influenciam no processo. Após a extração e separação líquido-líquido utilizou-se o espectrofotômetro de absorção atômica para aquisição das quantidades extraídas dos metais.

4.2.2.1- Obtenção das Curvas de Extração de Níquel e Cobalto por Solução Orgânica de Cyanex 272

Para a execução das análises foi utilizada técnica do Planejamento Fatorial, a qual consiste na realização de experimentos com todas as combinações possíveis das variáveis em estudo (NETO et al., 1995). Este método fornece resultados que permitem avaliar a

influência das variáveis e suas interações na resposta. Esta é uma técnica mais eficaz que o método univariado, na qual uma variável é mantida em determinado valor enquanto as outras são alteradas, assumindo desta forma que o máximo valor de uma variável é independente do nível da outra, levando muitas vezes a conclusões não fundamentadas (BERNARDO, 2000 in NETO et al., 1995).

Neste trabalho, foram utilizados três planejamentos fatoriais, todos em três níveis, isto é, cada variável foi avaliada em três valores, um chamado nível inferior (-), outro nível zero e o terceiro chamado nível superior (+). Os ensaios foram realizados em três planejamentos fatoriais representado por três níveis e três variáveis, correspondendo a um fatorial 3^3 , compreendendo 27 análises por planejamento, totalizando 81 ensaios com análise duplicada para aquisição de ruídos, obtendo 162 dados.

Variáveis analisadas:

- Temperatura;
- pH e
- Volume de Orgânico

Níveis de Estudo das variáveis:

- Nível Inferior(-);
- Nível Zero e
- Nível Superior(+)

Resposta:

- Concentração de Níquel e Cobalto em solução aquosa de sulfatos.

Os Quadros a seguir esquematizam as variáveis e seus níveis:

Planejamento 1

| Variável / Nível | (-) | 0 | (+) |
|-----------------------|------|-------|-------|
| pH | 5.0 | 5.4 | 5.8 |
| Temperatura (°C) | 50°C | 60°C | 70°C |
| Vol. de Orgânico (ml) | 2ml | 150ml | 300ml |

Planejamento 2

| Variável / Nível | (-) | 0 | (+) |
|-----------------------|------|-------|-------|
| pH | 5.0 | 5.4 | 5.8 |
| Temperatura (°C) | 50°C | 60°C | 70°C |
| Vol. de Orgânico (ml) | 3ml | 100ml | 200ml |

Planejamento 3

| Variável / Nível | (-) | 0 | (+) |
|-----------------------|------|------|------|
| pH | 5.0 | 5.4 | 5.8 |
| Temperatura (°C) | 50°C | 60°C | 70°C |
| Vol. de Orgânico (ml) | 20ml | 45ml | 70ml |

Os ensaios foram realizados mantendo constante o volume de solução aquosa de sulfatos, que contém os metais a serem extraídos pela solução orgânica de Cyanex, a 100ml. Adicionou-se o volume da solução orgânica quando na temperatura desejada sob agitação constante. Esperou-se atingir o ponto de equilíbrio, verificado através da constância do pH e então ajuste deste para o planejado. E então, análise quantitativa dos metais foi realizada em espectrofotômetro de absorção atômica.

A qualidade e quantidade de dados relativos ao processo são de importância fundamental. A confiabilidade discutível dos valores pode atrapalhar o treinamento da rede, assim como poderá levar a previsões futuras erradas ou de baixa qualidade;

4.2.3- Análise Estatística dos Dados para o Treinamento da Rede Neural

A qualidade dos dados influi diretamente na qualidade do ajuste realizado pelas redes neurais artificiais, o que faz esta fase ser de extrema importância.

Este tratamento pode seguir os seguintes aspectos:

- Identificação de outliers, de erros grosseiros, por exemplo: valores nulos, negativos, ordens de grandeza diferentes, descontinuidade brusca do comportamento das variáveis ao longo do tempo, possivelmente devido a erros de medição do equipamento ou erro analítico. Esta análise foi feita plotando-se as variáveis e estudando o comportamento delas segundo uma pré-determinação ou não de coerência. A experiência do processo permitiu que alguns pontos fossem eliminados por estarem fora de uma faixa aceitável, ou possível, de operação.
- Tratamento de dados utilizando a própria rede neural: a partir dos dados definidos na etapa anterior faz-se o treinamento com o conjunto completo de dados, verificando o erro residual de ajuste e o gráfico da variável experimental em função da variável calculada pela rede neural. Dessa forma, pode-se identificar vários dados discrepantes em relação à tendência da maioria dos

pontos, que caracteriza o comportamento daquela determinada variável. Uma vez o ponto sendo falso, ele foi descartado e o novo conjunto de dados alimentado ao treinamento da rede novamente e repetindo este procedimento até que não haja pontos ou grupos de pontos discrepantes.

4.2.4- Verificação do Domínio das Variáveis da Rede Neural

As redes neurais apresentam uma ótima característica de interpolação de resultados, ou seja, a simulação de casos contidos dentro do domínio do conjunto de aprendizagem, no entanto extrapolações devem ser acompanhadas com muito cuidado. Assim, o procedimento sistemático em que o domínio do conjunto de aprendizagem contenha o domínio do conjunto de teste torna-se importante.

4.2.5- Normalização dos Dados

A faixa de normalização dos dados pode ser fixada entre [0 ; 1]. Aconselha-se, no entanto, a utilizar o intervalo de normalização entre [0,1 ; 0,9] para evitar que os pesos w_{ij} tenham valores muito altos (PELLICCI, 2001).

4.2.6- Treinamento da Rede Neural

Como as variáveis que são controladas no processo já são definidas, sabe-se com antecedência o número de neurônios nas camadas de entrada e saída da rede neural. Juntamente com o software específico, no caso o *Neurosolutions Consultants*, arbitra-se o número de neurônios na camada oculta, o número de iterações para o treinamento, o valor do resíduo e o conjunto de dados todo como sendo de aprendizagem.

Faz-se, primeiramente, o treinamento da rede neural observando os possíveis pontos com problemas. Mudando-se o número de neurônios da camada oculta em uma unidade, gerando-se várias redes com diferente erros. Com essas várias redes neurais,

escolhe-se aquela que apresentar menor erro residual para aquele conjunto de dados. Após análises dos pontos e dos erros, faz-se uma “limpeza” dos dados e separação em dois conjuntos, o de Aprendizado e o de Teste, com as características já citadas. Faz-se novamente o treinamento, variando-se o número de neurônios na camada intermediária e aquela que obtiver menor valor residual de erro para o conjunto de Teste e obtiver um bom ajuste e resultado dos testes de consistência é a escolhida. Os pesos das redes neurais são escolhidos em função do erro mínimo, este tipo de avaliação pode gerar duas situações: erro obtido a partir dos dados de aprendizagem e de teste diminuem em função do número de iterações. Com isto, o número de iterações correspondente ao erro mínimo é o mesmo para os dois conjuntos, definindo os pesos da rede; e o erro dos dados do conjunto de teste pode começar a aumentar a partir de um determinado número de iterações, neste caso, os pesos são os correspondentes ao menor erro deste conjunto de teste.

A fase de aprendizado ou treinamento é basicamente um problema de estimação de parâmetros e freqüentemente exige a apresentação da base de dados à rede por milhares de vezes. Dessa forma, depara-se com as mesmas dificuldades associadas à otimização de processos, tais como: convergência, existência de mínimos locais, esforço computacional, entre outras.

4.2.7- Validação da Rede Neural Treinada

Passa-se aos teste de consistência, após a escolha da rede neural que melhor representa as variáveis de saída. Primeiro, avalia-se a dispersão dos valores das variáveis de saída calculados comparando-os com os experimentais. Após esta avaliação passa-se para o teste da função distância (definida como a diferença entre o valor experimental e o calculado pela rede neural). Esta informação nos fornece a distribuição da diferença de erro absoluto. Uma distribuição equitativa em torno de zero mostra a boa qualidade do ajuste realizado. Uma outra forma de apresentar a distribuição dos erros é a verificação de distribuição normal dos erros entre o valor das variáveis experimentais e das calculadas pela rede neural. Esta informação reafirma a boa qualidade do ajuste realizado à medida que a distribuição normal demonstra o equilíbrio de frequência dos valores encontrados em torno de zero, quanto mais normal a distribuição, melhor ajustadas as variáveis. O outro

teste de consistência a ser realizado é a verificação do comportamento das variáveis experimentais em conjunto com as calculadas pela rede neural ao longo do tempo, que também nos fornece uma boa informação da qualidade do ajuste realizado.

4.2.8- Utilização da Rede

Após o treinamento, a rede neural pode ser simulada com grande rapidez. Milhares de simulações podem ser processadas em segundos, o que possibilita o emprego do modelo de redes neurais em procedimentos de simulação, otimização, controle de processo em linha, análise de restrições do processo, análise de sensibilidade de variáveis, analisadores virtuais, entre outras aplicações.

A metodologia para este trabalho foi adquirir a curva de seletividade dos metais variando algumas propriedades em laboratório para obter maior confiabilidade do processo de extração, bem como, ampliar o intervalo de aplicação para possibilitar maior generalização das simulações no processo industrial. Em seguida, foram analisadas as variáveis de acordo com a disponibilidade dos dados obtidos no laboratório e importância das mesmas e por fim foi feita a validação da rede.

Capítulo 5 - RESULTADOS E DISCUSSÕES

5.1- Análise dos Dados Experimentais

A extração de Cobalto através da solução orgânica Cyanex 272 tem forte influência pela temperatura, pH e volume de orgânico adicionado como se pode verificar neste capítulo. Fez-se variações segundo o planejamento experimental apresentado no capítulo “Materiais e Métodos”, com o objetivo de obter dados que dêem a capacidade e possibilidade de modelar e simular o processo de extração de níquel e cobalto da VOTORANTIM visando sua otimização e melhor controle.

5.1.1- Extrações de Cobalto por Solução Orgânica Cyanex 272

Os gráficos a seguir apresentam o perfil de extração do cobalto a pH's 5,0; 5,4 e 5,8 nas temperaturas de 50°C, 60°C e 70°C variando o volume de solução orgânica Cyanex 272 (solvente).

- Temperatura de 50°C

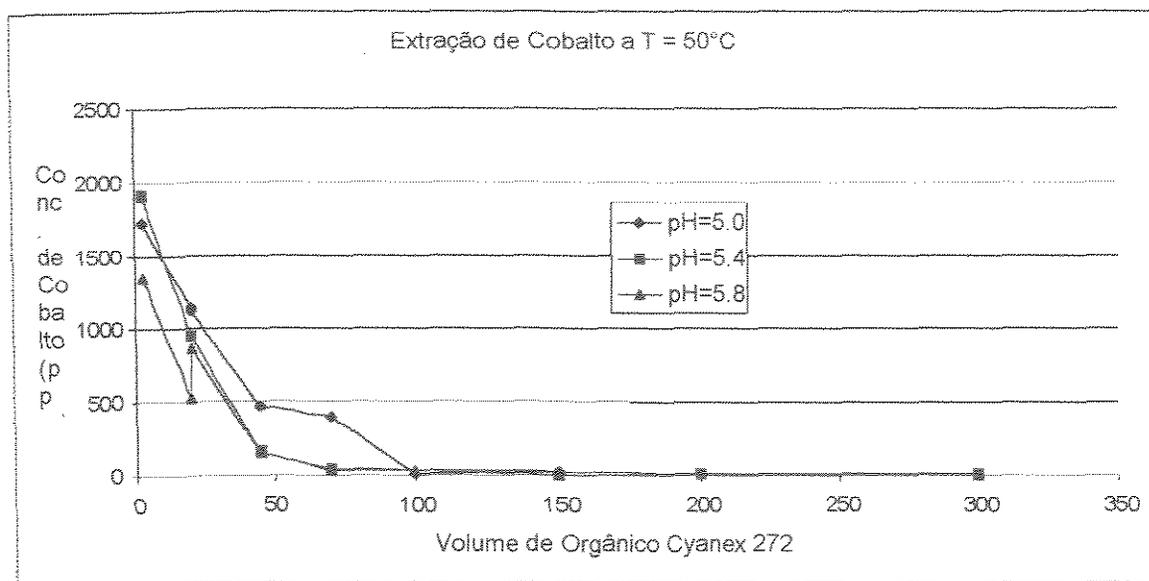


Figura 5.1: Comparação da concentração do cobalto restante em solução aquosa a T= 50°C a pH = 5,0; 5,4 e 5,8

Pode-se observar que o efeito do pH está relacionado diretamente à extração de cobalto. A concentração de cobalto na solução aquosa de sulfatos após o equilíbrio das fases (com solução orgânica – Cyanex 272) é menor em pH mais alto, cujo valor no caso estudado à temperatura de 50°C é 5,8.

- Temperatura de 60°C

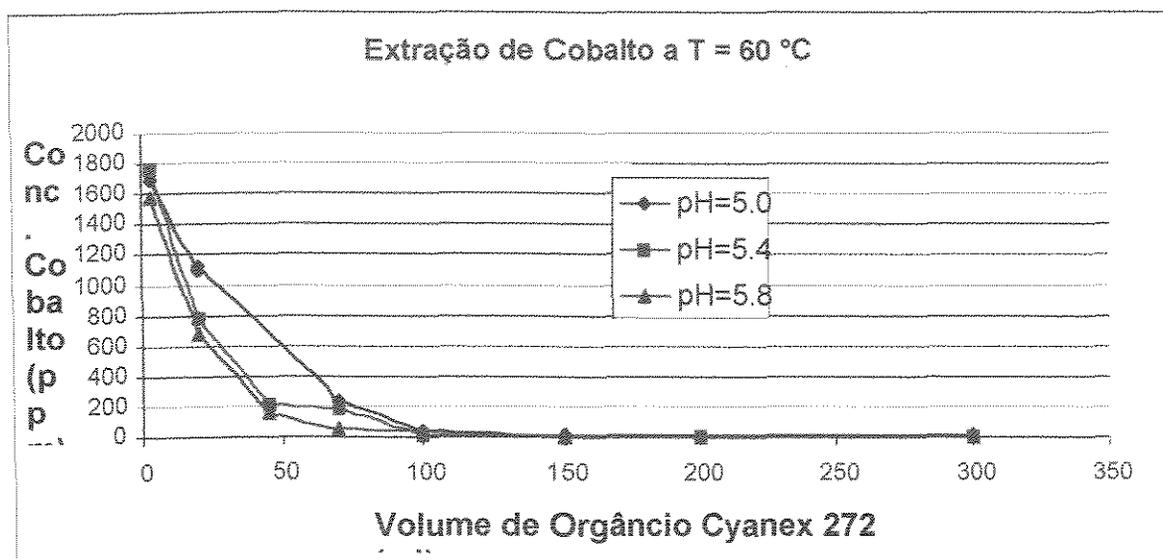


Figura 5.2: Comparação da concentração restante de cobalto em solução aquosa a T= 60°C a pH = 5,0; 5,4 e 5,8

Pode-se observar que o efeito do pH está relacionado diretamente à extração de cobalto. A concentração de cobalto na solução aquosa de sulfatos após o equilíbrio das fases (com solução orgânica – Cyanex 272) é menor em pH mais alto, cujo valor no caso estudado a temperatura de 60°C é igual a 5,8.

- Temperatura = 70°C

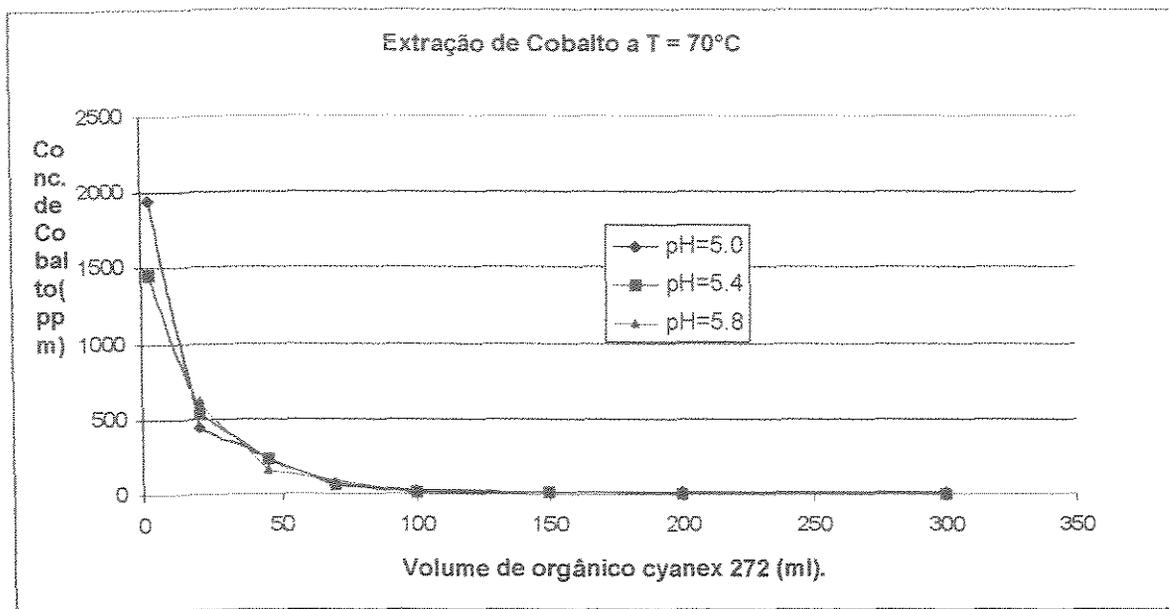


Figura 5.3: Comparação da Concentração restante de cobalto em solução aquosa a T= 70°C a pH = 5,0; 5,4 e 5,8

Pode-se observar que o efeito do pH está relacionado diretamente à extração de cobalto. A concentração de cobalto na solução aquosa de sulfatos após o equilíbrio das fases (com solução orgânica – Cyanex 272) é menor em pH mais alto, cujo valor no caso estudado à temperatura de 70°C é 5,8.

Como podemos verificar em todos os gráficos a quantidade de cobalto nas soluções aquosas de sulfatos após sua extração através da solução orgânica de Cyanex 272, como era de se esperar, diminui com o aumento do volume adicionado de orgânico. Isso mostra que a extração de cobalto está diretamente ligada a quantidade do volume adicionada de solução orgânica.

- pH = 5,0

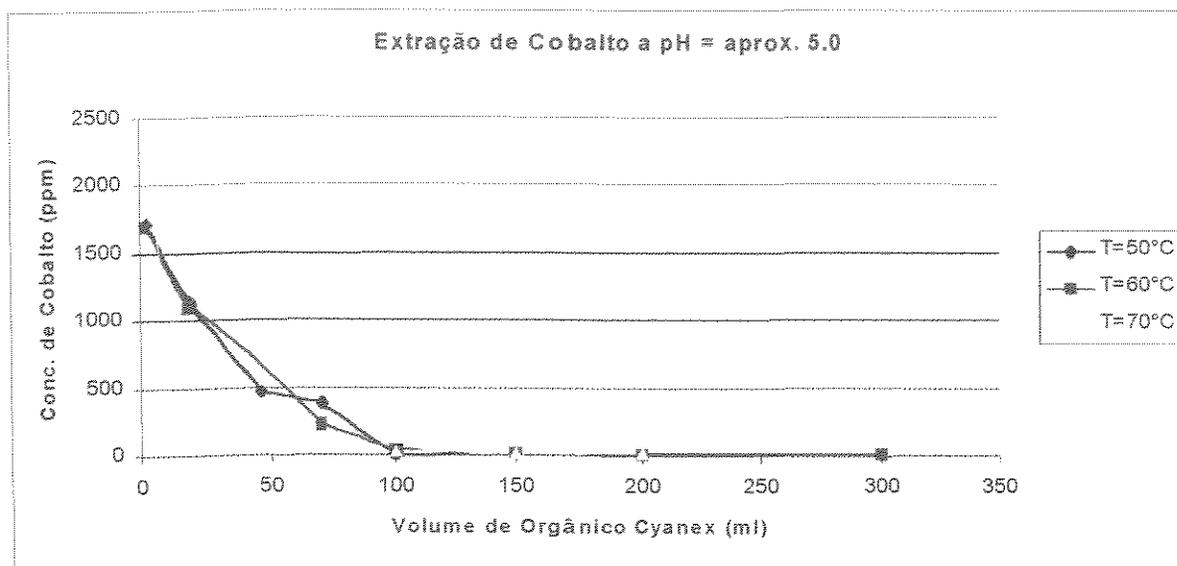


Figura 5.4: Comparação da Concentração Restante de Cobalto em solução aquosa a pH=5.0

- pH = 5,4

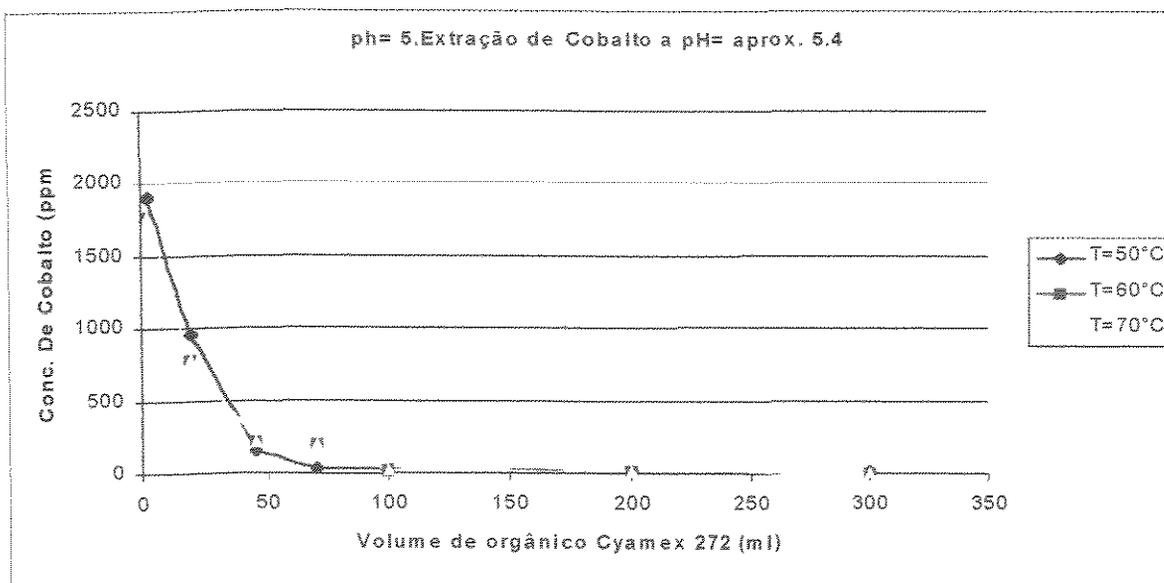


Figura 5.5 - Comparação da Concentração Restante de Cobalto em solução aquosa a pH=5.4

- pH = 5,8

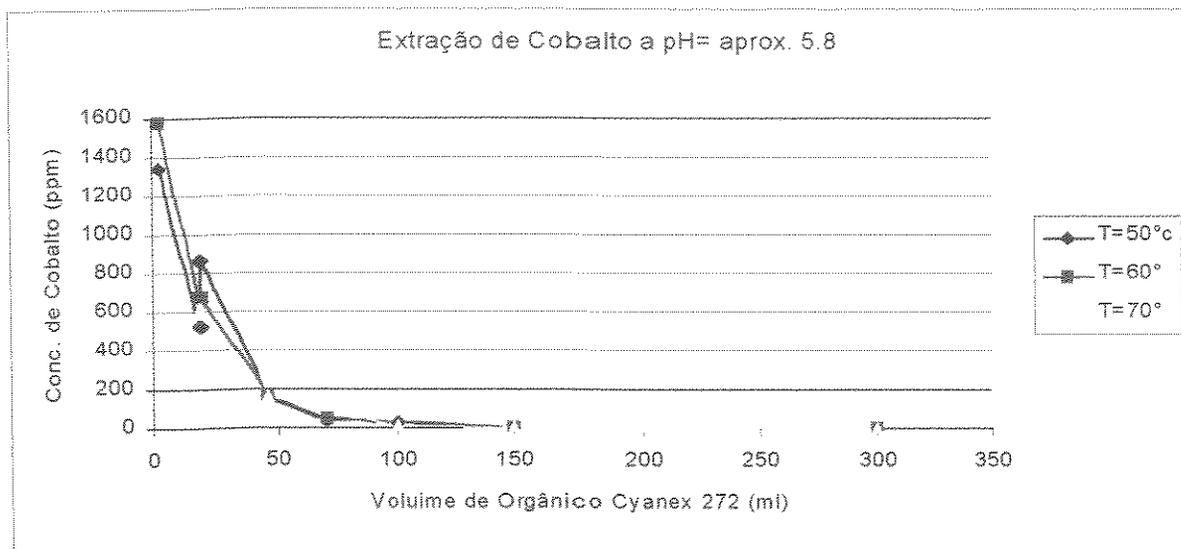


Figura 5.6- Comparação da Concentração Restante de Cobalto em solução aquosa a pH=5,8

Verifica-se que a temperaturas mais altas a extração de cobalto tem maior eficiência, como visto nas figuras 5.4 a 5.6, pois apresenta maior extração de cobalto pela solução orgânica a volumes próximos de 200 ml. Porém faz-se necessárias análise através de superfície de resposta (*Surface plot*), para que se possa concluir as melhores condições de extração, como será apresentada no item 5.2.

5.1.2- Extração de Níquel por Solução Orgânica Cyanex 272

As figuras 5.7 a 5.12, apresentam o comportamento das extrações de níquel nos pHs e temperaturas indicados variando o volume de orgânico.

- Temperatura = 50°C

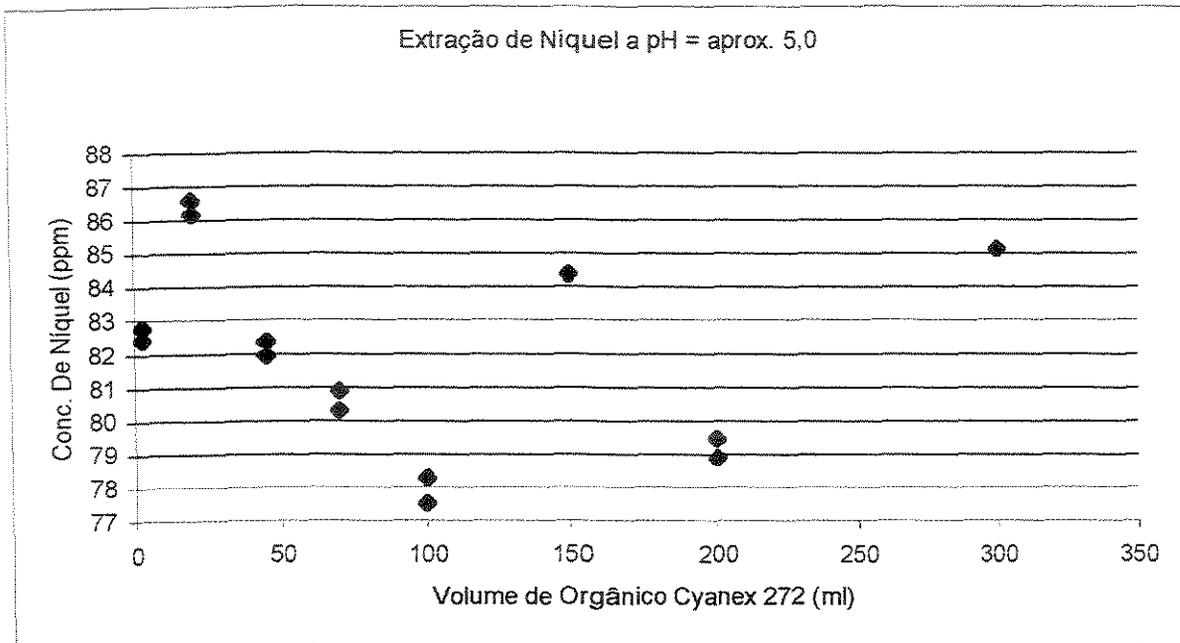


Figura 5.7: Concentração restante de Níquel em solução aquosa a T= 50°C e pH = 5,0

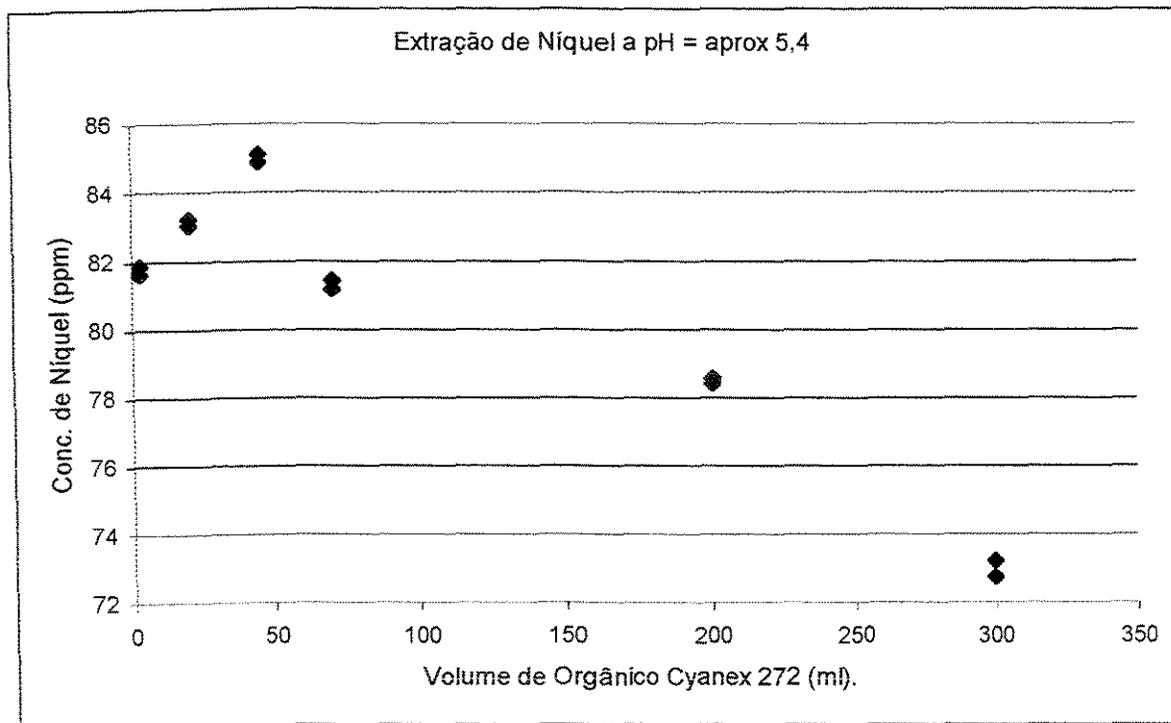


Figura 5.8: Concentração restante de Níquel em solução aquosa a T= 50°C e pH = 5,4

Figura 5.8: Concentração restante de Níquel em solução aquosa a T= 50°C e pH = 5,4

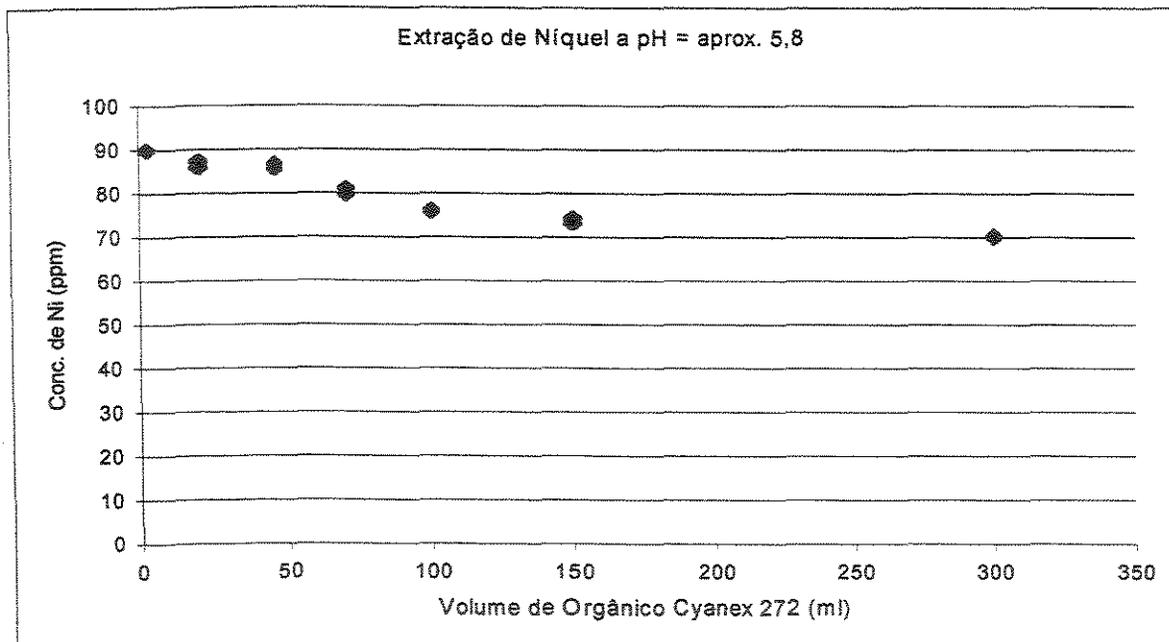


Figura 5.9: Concentração restante de Níquel em solução aquosa a T= 50°C e pH = 5,8.

Analisando a extração de níquel à temperatura de 50° observa-se que não há uma tendência a pH's menores, apenas a pH igual a 5,8 observa-se uma suave tendência de a extração aumentar com maior presença de solvente Cyanex 272. Isso mostra a complexidade do sistema de extração e a variação do comportamento relativa à extração do níquel.

- Temperatura de 60°C

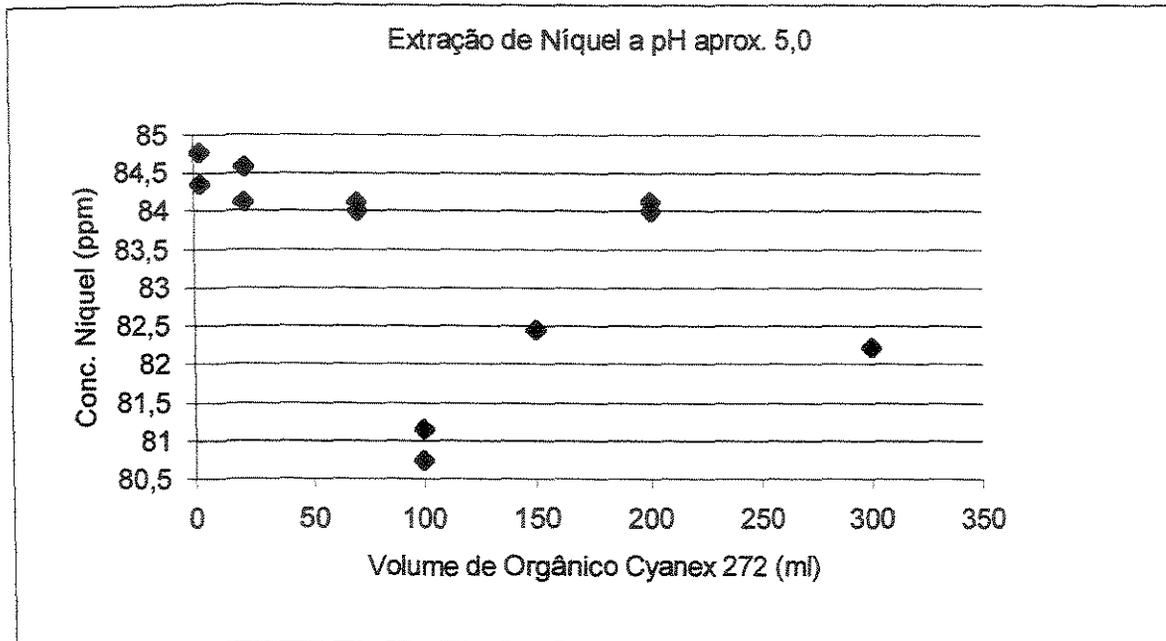


Figura 5.10: Concentração restante de Níquel em solução aquosa a T= 60°C e pH = 5,0

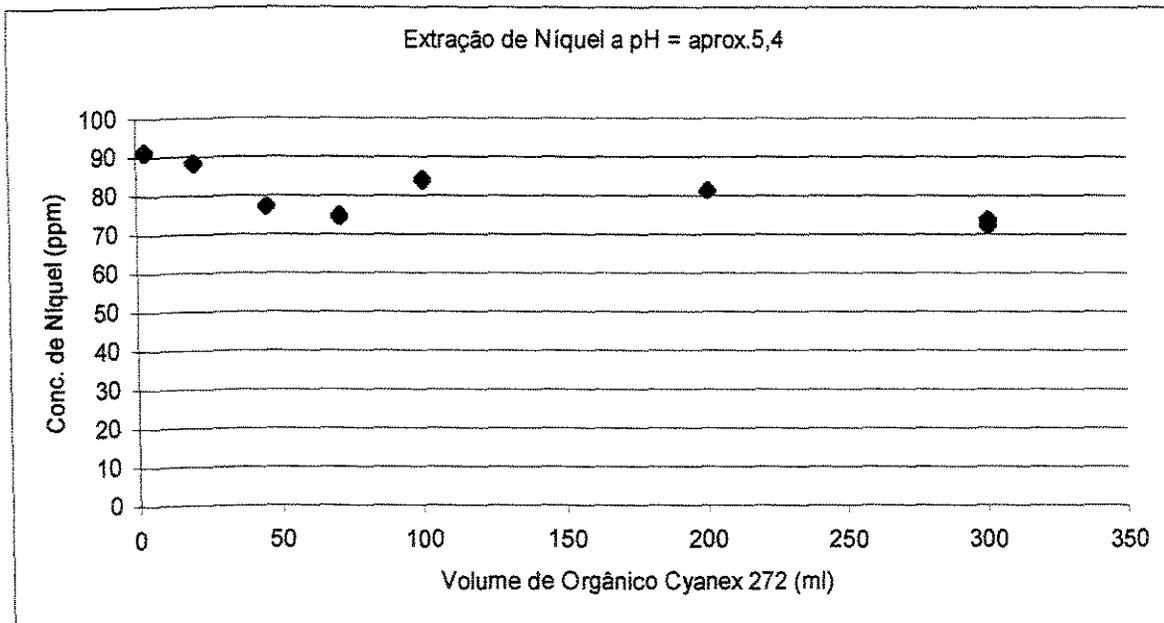


Figura 5.11: Concentração restante de Níquel em solução aquosa a T= 60°C e pH = 5,4

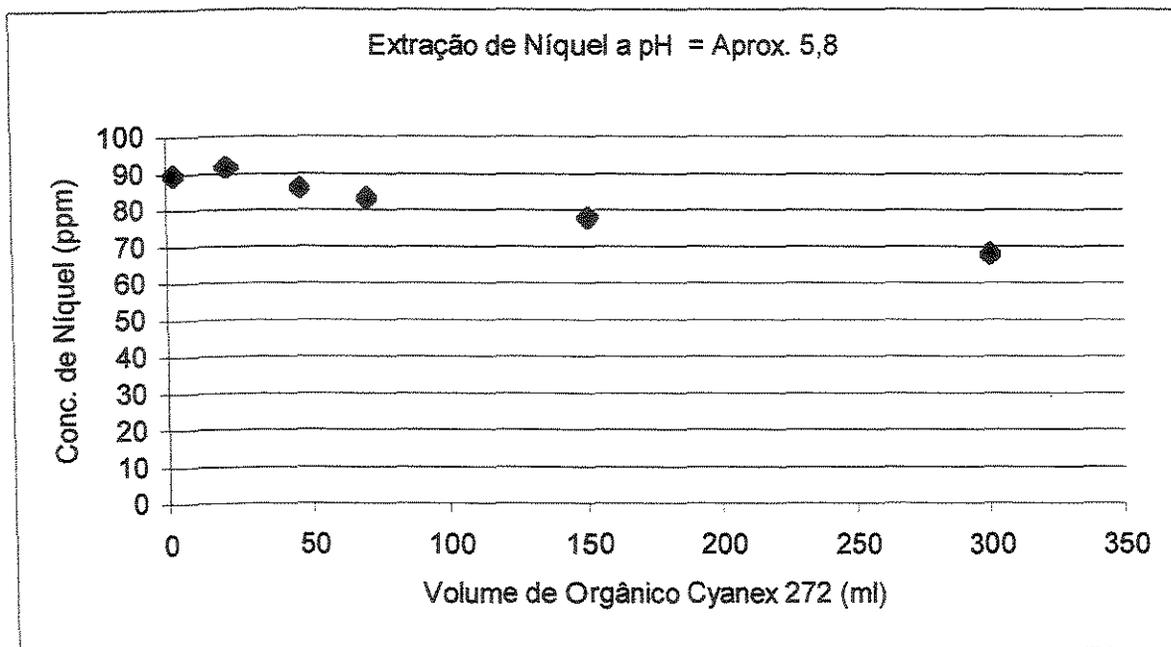


Figura 5.12: Concentração restante de Níquel em solução aquosa a $T = 60^{\circ}\text{C}$ e $\text{pH} = 5,8$

- Temperatura = 70°C

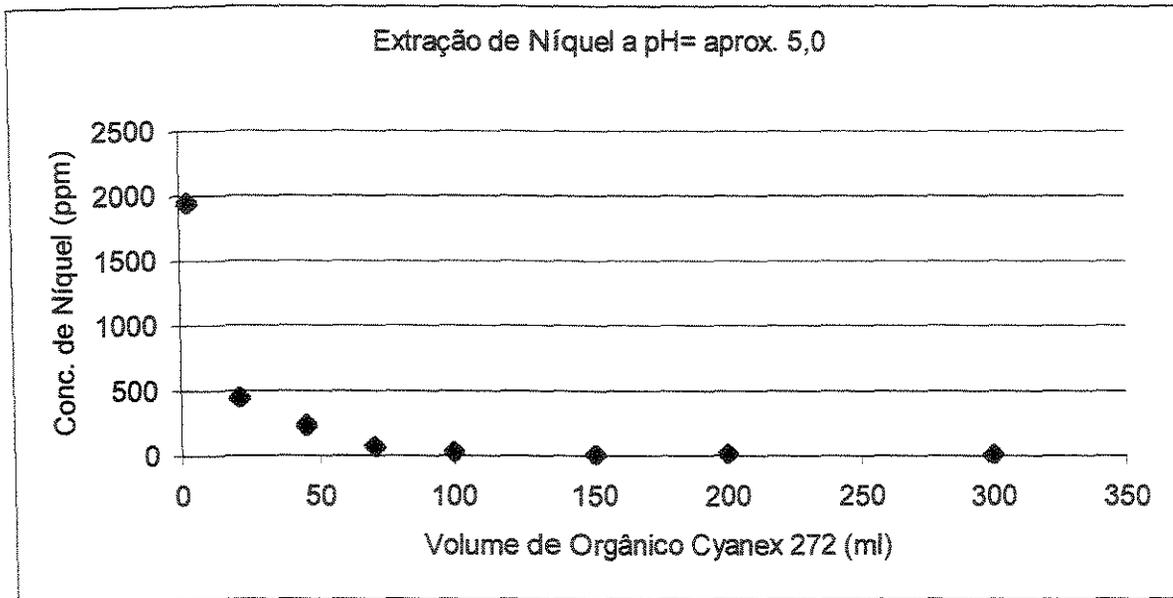


Figura 5.13: Concentração restante de Níquel em solução aquosa a T= 70°C e pH = 5,0

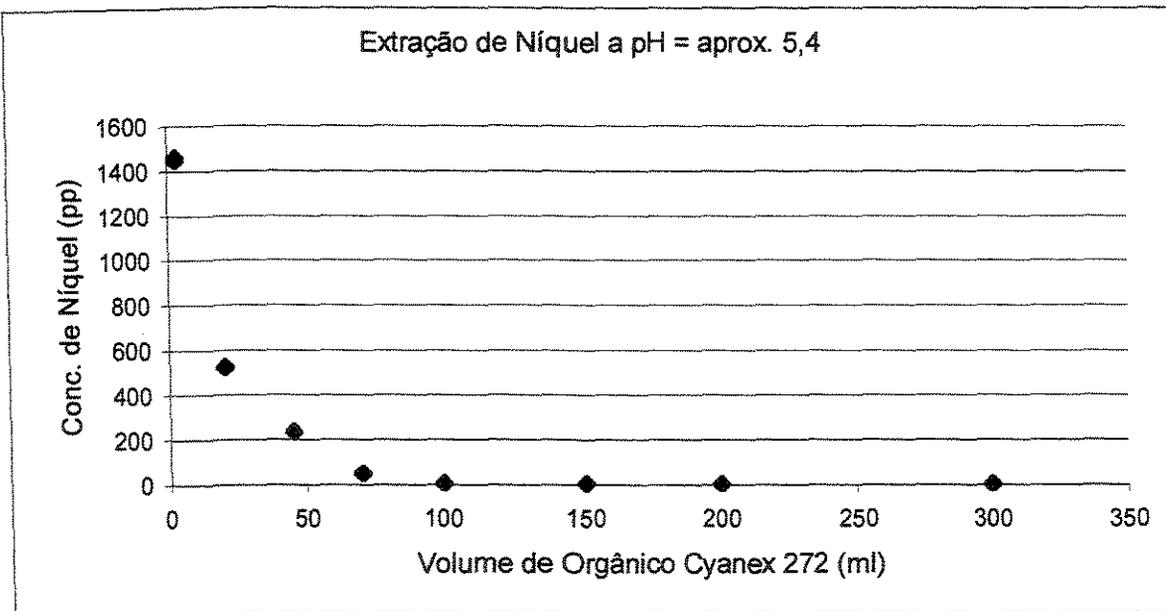


Figura 5.14: Concentração restante de Níquel em solução aquosa a T= 70°C e pH = 5,4

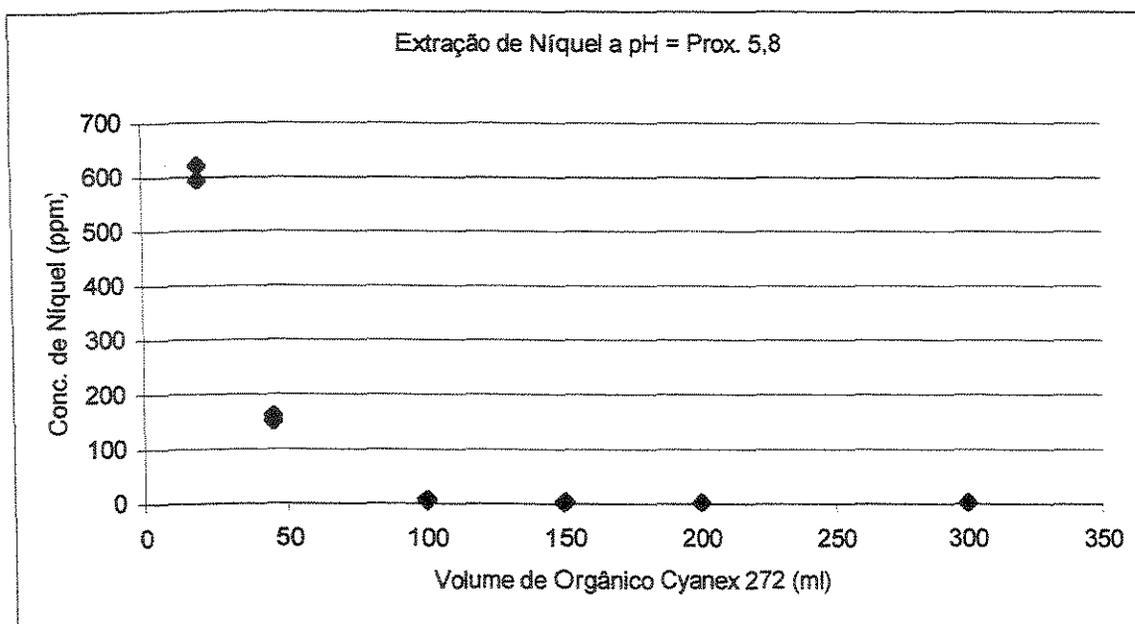


Figura 5.15: Concentração restante de Níquel em solução aquosa a $T = 70^{\circ}\text{C}$ e $\text{pH} = 5,8$

Como já esperado, estas análises apresentadas acima, mostram que a quantidade de níquel extraída também é influenciada pelo pH e temperaturas. Porém não é possível afirmar que a quantidade de orgânico adicionada tem relação com a quantidade de níquel extraída, o que nos faz concluir uma maior dependência das variáveis temperatura e pH. Estas dependências e relações entre variáveis pode ser verificada mais detalhadamente nas superfícies de resposta no próximo item.

5.1.3- Determinação do Coeficiente de Partição através de um Planejamento Fatorial a Três Níveis e Três Fatores (Planejamento 3^3)

O Coeficiente de partição é uma função complexa de uma variedade de diferentes fatores, desde a concentração dos componentes envolvidos nas fases à escolha dos componentes que compõem o sistema.

Em geral, a partição em sistemas bifásicos é definida em termos do seu coeficiente de partição (K).

$$K=[C]_1/[C]_2 \quad (5.1)$$

Onde $[C]_1$ é a concentração do composto de interesse no topo e $[C]_2$ é a concentração na fase de fundo.

O Coeficiente de partição K é a medida tomada como padrão comparativo para a eficiência da separação do soluto nos sistemas bifásicos (BERNARDO, 2000).

O Coeficiente de partição é igual a 1 quando a concentração de soluto em ambas as fases for igual, ou seja, indica que o soluto particionou igualmente entre as duas fases. Por exemplo, quando o coeficiente de partição for igual a 10, isto significa que o soluto particiona 10 vezes mais na fase de topo que na fase de fundo, mostrando o potencial de sua purificação neste sistema.

A Tabela 5.1 a seguir apresenta os resultados obtidos para este ensaio:

Tabela 5.1 – Valores de Coeficiente de Partição

| V(ml) | T(°C) | pH | K p/ Co | K p/ Ni |
|-------|-------|------|---------|---------|
| 100 | 50 | 5,04 | 307,64 | 0,16 |
| 100 | 50 | 5,40 | 99,00 | 0,12 |
| 100 | 50 | 6,00 | 73,29 | 0,18 |
| 100 | 60 | 5,03 | 54,90 | 0,11 |
| 100 | 60 | 5,40 | 100,99 | 0,07 |
| 100 | 60 | 5,80 | 283,90 | 0,07 |
| 100 | 70 | 5,02 | 74,56 | 0,03 |
| 100 | 70 | 5,44 | 275,63 | 0,07 |
| 100 | 70 | 5,75 | 323,15 | 0,08 |
| 200 | 50 | 5,09 | 362,64 | 0,13 |
| 200 | 50 | 5,44 | 149,94 | 0,15 |
| 200 | 50 | 5,80 | 324,20 | 0,11 |
| 200 | 60 | 5,00 | 432,84 | 0,07 |
| 200 | 60 | 5,45 | 324,20 | 0,11 |
| 200 | 60 | 5,80 | 652,59 | 0,22 |
| 200 | 70 | 5,00 | 150,06 | 0,17 |
| 200 | 70 | 5,39 | 683,93 | 0,05 |
| 200 | 70 | 5,85 | 762,36 | 0,17 |
| 300 | 50 | 5,06 | 204,34 | 0,06 |
| 300 | 50 | 5,55 | 370,75 | 0,24 |
| 300 | 50 | 5,88 | 1189,48 | 0,29 |
| 300 | 60 | 4,98 | 279,11 | 0,09 |
| 300 | 60 | 5,37 | 424,53 | 0,22 |
| 300 | 60 | 5,77 | 1502,76 | 0,33 |
| 300 | 70 | 5,02 | 349,88 | 0,16 |
| 300 | 70 | 5,43 | 443,44 | 0,21 |
| 300 | 70 | 5,78 | 774,19 | 0,43 |

A influência das variáveis estudadas (Volume de orgânico, Temperatura e pH) sobre o coeficiente de partição do cobalto pode ser avaliada de acordo com a metodologia das superfícies de resposta, as quais se encontram representadas nas figuras 5.16 a 5.18 a seguir:

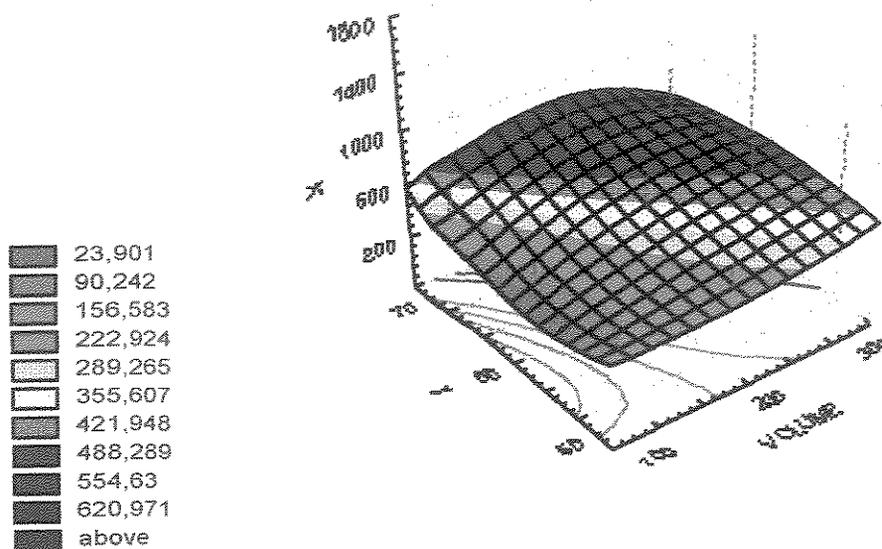


Figura 5.16 – Superfície de Resposta para o Coeficiente de Partição do Processo de Extração do Cobalto. Condição: pH= 5,4

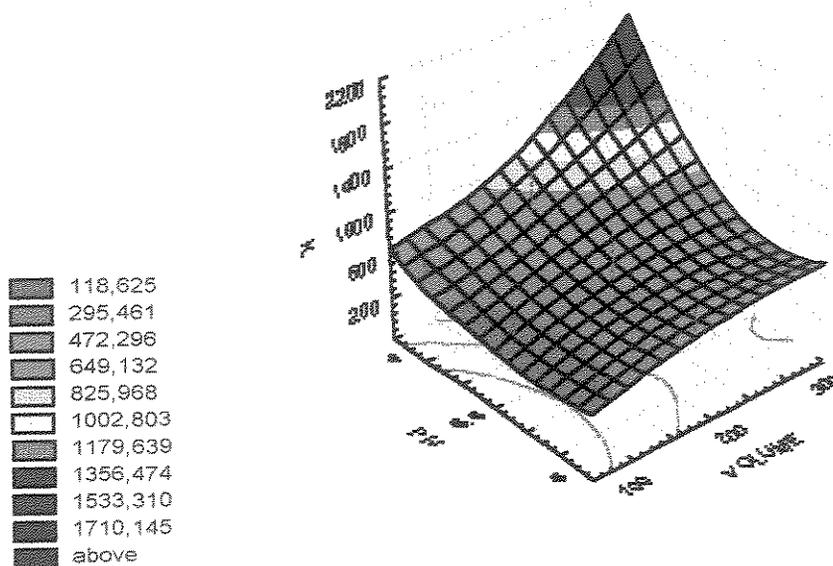


Figura 5.17 – Superfície de Resposta para o Coeficiente de Partição do Processo de Extração do Cobalto. Condição: T = 60°C

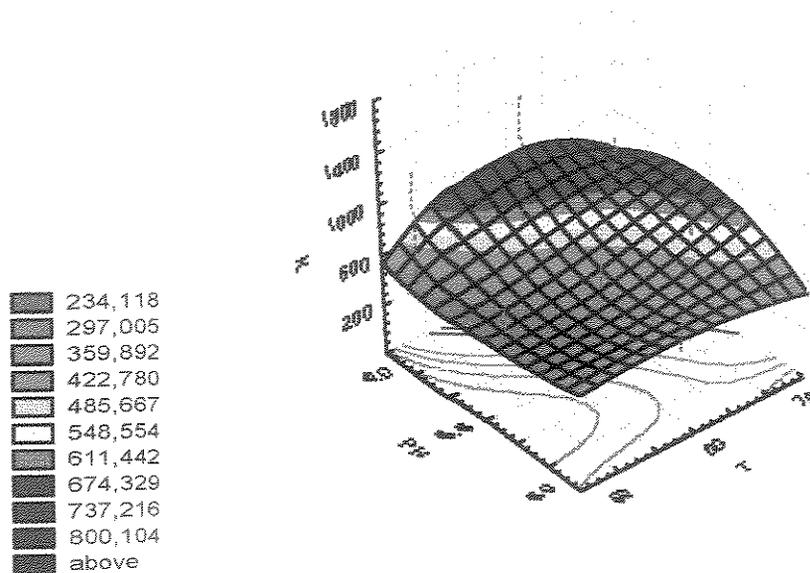


Figura 5.18 – Superfície de Resposta para o Coeficiente de Partição do Processo de Extração do Cobalto. Condição: Volume = 200 ml.

Esta metodologia nos permite avaliar a tendência de determinadas variáveis sobre a resposta procurada, em nosso caso o coeficiente de partição (K) do cobalto, representado nas figuras 5.16 a 5.18.

Observa-se na figura 5.16 que a tendência de obtenção dos maiores do coeficientes de partição encontra-se na faixa situada entre 60°C e 70°C e para volumes em torno de 200 ml. Quando se avalia a influência do pH sobre K observa-se que para regiões em torno de 5,4 a temperatura ótima situa-se na faixa de 60°C e 70°C para volumes próximos a 200 ml. Como pode-se observar nas figuras 5.17 e 5.18.

No caso da análise da extração do níquel pela solução de Cyanex 272, a resposta esperada será tão válida quanto menores forem os coeficientes de partição, já que não é objeto deste trabalho a extração do mesmo. As figuras 5.19 a 5.21 a seguir mostram as superfícies de resposta para este caso.

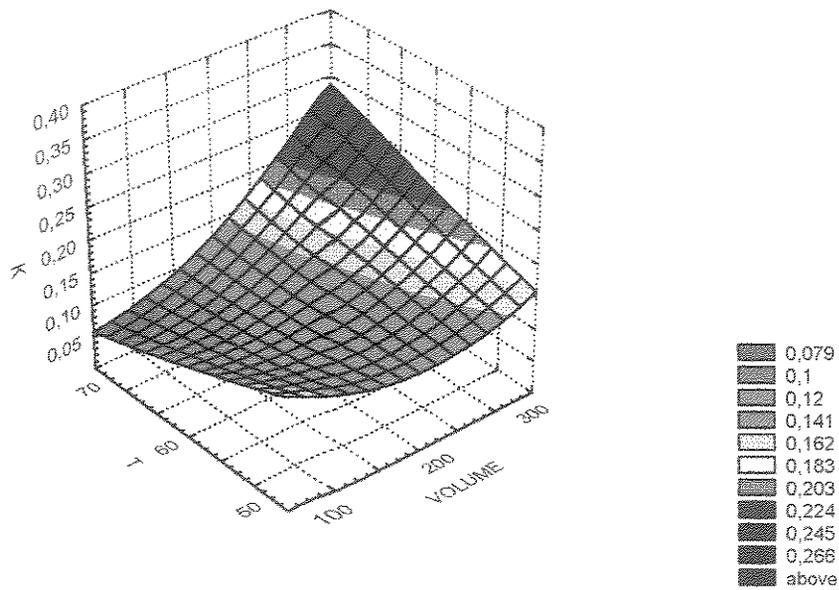


Figura 5.19– Superfície de Resposta para o Coeficiente de Partição do Processo de Extração do Níquel. Condição: pH= 5,4

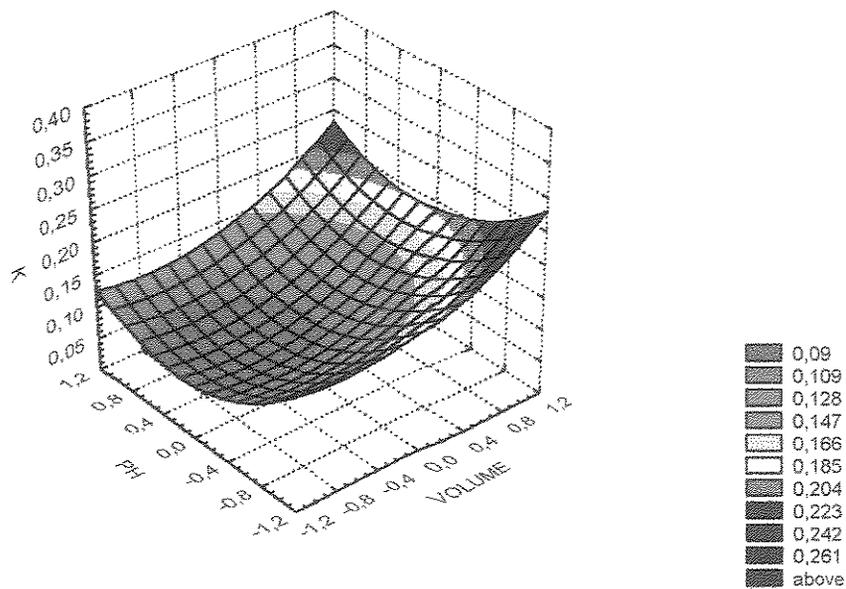


Figura 5.20– Superfície de Resposta para o Coeficiente de Partição do Processo de Extração do Níquel. Condição: T=60°C

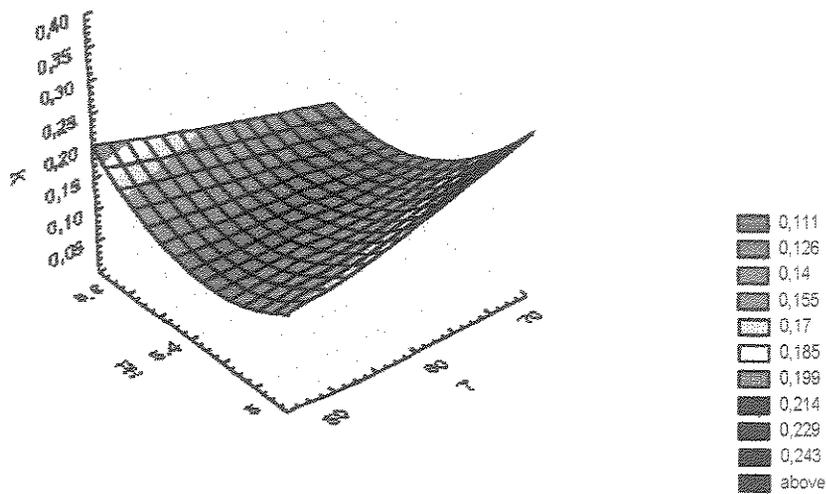


Figura 5.21 – Superfície de Resposta para o Coeficiente de Partição do Processo de Extração de Níquel. Condição: Volume = 200 ml.

Neste caso, os menores coeficientes de partição (K) são encontrados nas regiões assinaladas pela coloração verde nos gráficos. Da mesma forma, comparativamente à extração do cobalto a tendência do resultado desejado aponta para temperaturas entre 60°C e 70°C e volume de solução orgânica Cyanex 272 na faixa de 200 ml.

5.2- Redes Neurais Artificiais

Para a modelagem do processo de extração de cobalto obteve-se modelos com as seguintes estruturas:

- Número de Camadas Intermediárias: 1 e 2;
- Variação de neurônios na camada intermediária: 2 a 12 neurônios;
- Funções de Ativação: Sigmoidal e Tangente hiperbólica;
- Arquitetura da Rede: Feedforward - Redes Multi Layer Perceptron e Redes Generalizadas;
- Saídas: 1Co; 1 Ni e 2NiCo.

• Algoritmo de treinamento: back-propagation/Quickpropagation e Delta-bar-delta.

• Iterações: 100; 500 e 10000

Porém, destas estruturas, os melhores modelos foram os que, quando em seu treinamento da rede, utilizou-se o algoritmo backpropagation com algoritmo Quickpropagation, tendo como função de transferência a tangente hiperbólica tanto na camada intermediária quanto na camada de saída.

Após estudo estatístico dos dados, com detecção de outliers, verificação dos dados experimentais, correlação dos parâmetros e formação do conjunto de dados, obteve-se 126 dados para modelagem. Destes destinou-se 15% destes para validação, 10% para teste e o restante para o treinamento da rede.

Os melhores resultados obtidos foram para as redes que seguem:

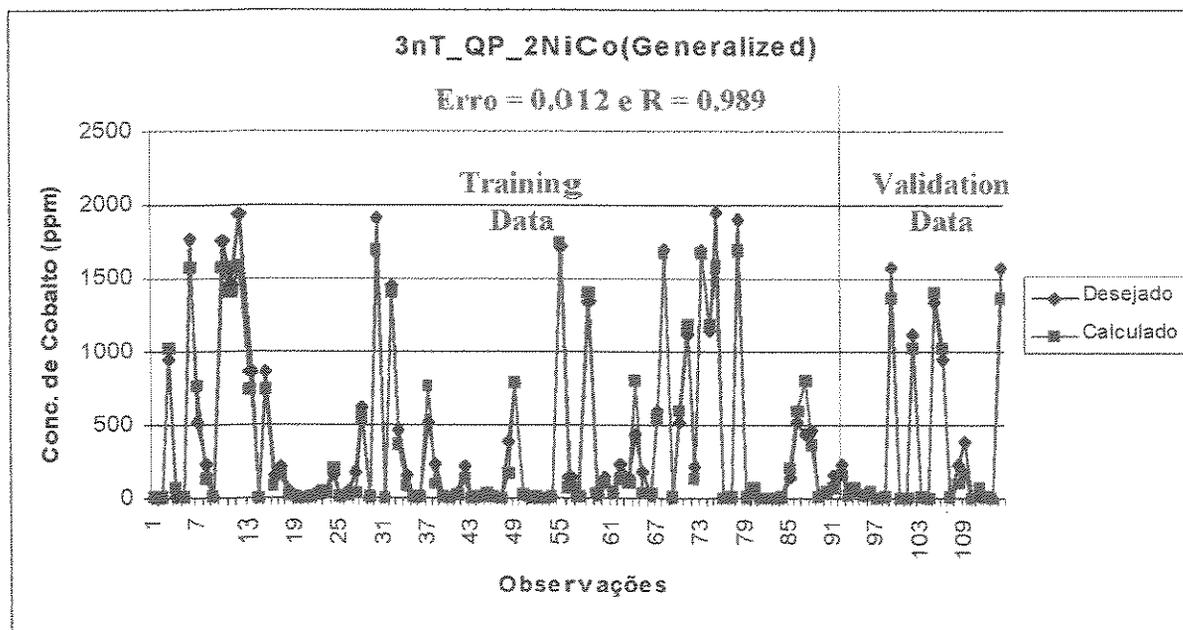


Figura 5.22 – Concentração de Cobalto estimada usando Redes neurais artificiais com 3 neurônios na camada intermediária (hidden) e dois neurônios de saída (Output).

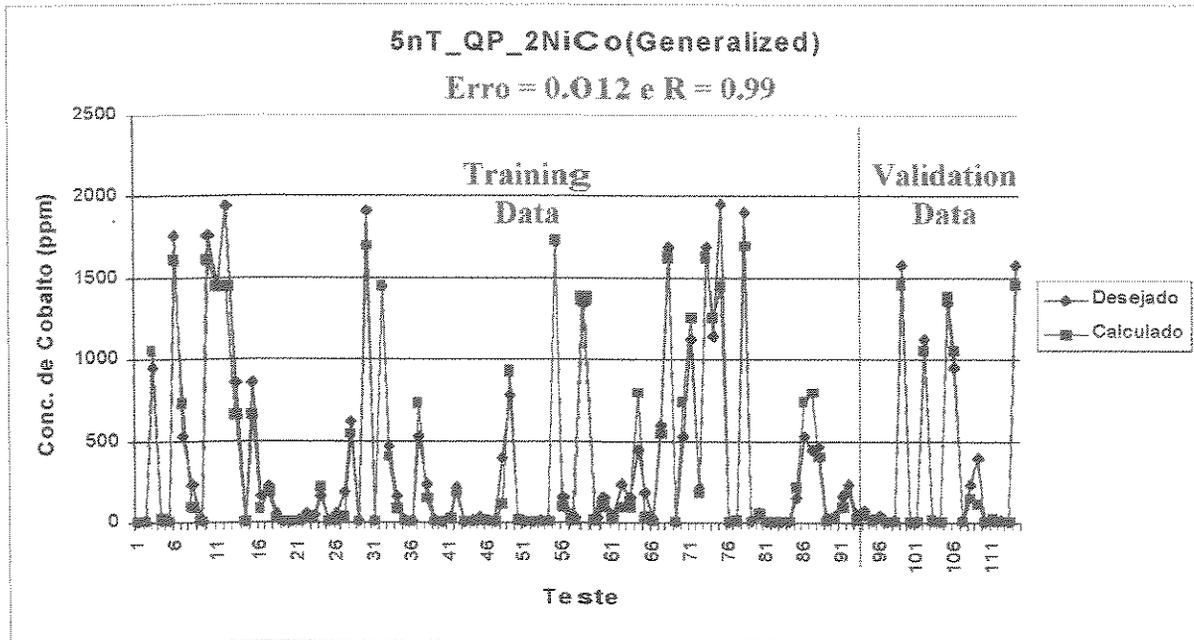


Figura 5.23 – Concentração de Cobalto estimada usando Redes neurais artificiais com 5 neurônios na camada intermediária e 2 neurônios de saída

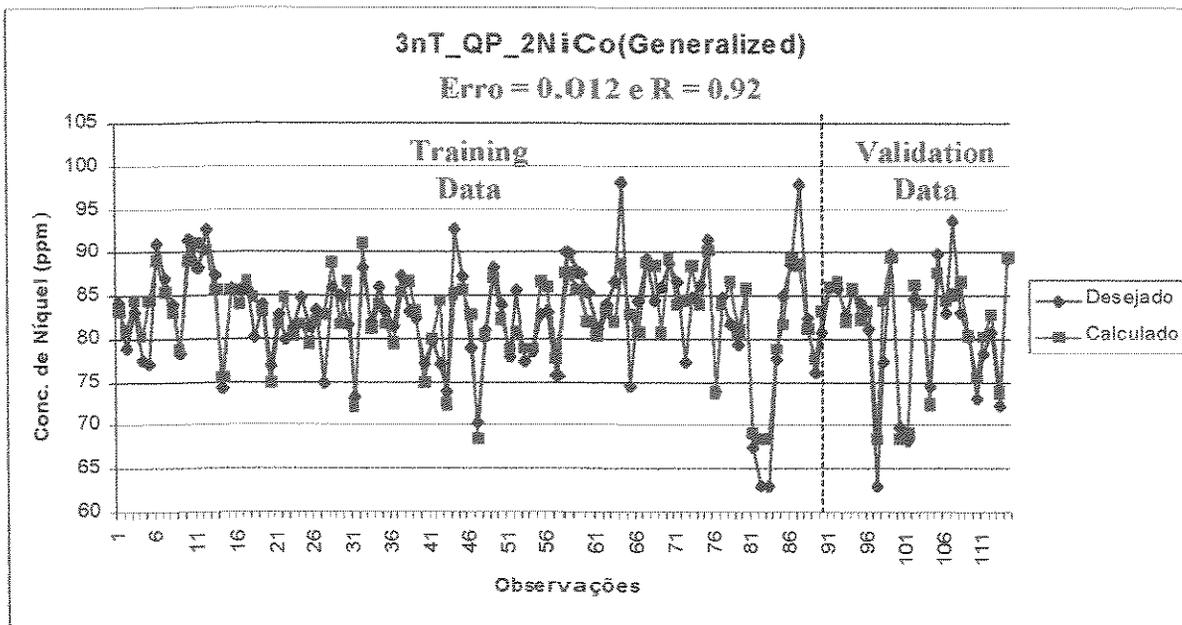


Figura 5.24 – Concentração de Níquel estimada usando Redes neurais artificiais com 3 neurônios na camada intermediária e 2 neurônios de saída.

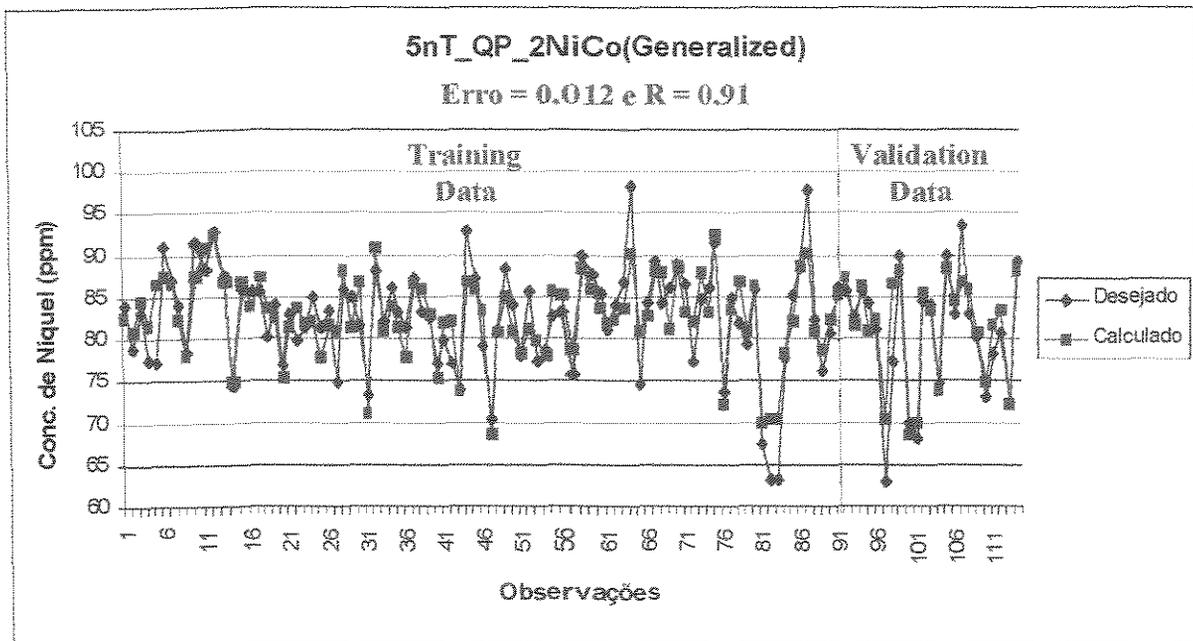


Figura 5.25 – Concentração de Níquel estimada usando Redes neurais artificiais com 5 neurônios na camada de intermediária e 2 neurônios na saída.

Com base no erro tanto do conjunto de teste, como no de treino, nos gráficos acima mostrados, verificou-se que a melhor performance para predição da concentração restante de níquel e cobalto em solução aquosa de sulfatos é obtida com uma rede neural de 3 neurônios na camada intermediária e com 2 saídas. As variáveis de entrada são temperatura, volume de orgânico e pH. Pode-se verificar na Tabela 5.2 abaixo, para 100 iterações, os erros e suas respectivas correlações:

Tabela 5.2 – Erro quadrático médio e índice de correlação para teste

| Número de Neurônios | Erro - E (Teste) | Correlação Níquel (R) | Correlação Cobalto (R) |
|---------------------|--------------------|---------------------------|----------------------------|
| 3 | 0.012 | 0.92 | 0.99 |
| 4 | 0.030 | 0.89 | 0.90 |
| 5 | 0.012 | 0.91 | 0.99 |

Como podemos verificar, as correlações (R) e erros (E) com 3 e 5 neurônios na camada intermediária são as mesmas, porém, a rede com 3 neurônios na camada intermediária é a rede que melhor representa o processo aqui estudado, pelo fato de ser a mais simples.

Realizou-se também treinamento e teste para 80, 200, 500, 10000 iterações para observar a rede quanto ao sobre-treinamento (*over-training*), e verificou-se que para menores iterações o E é maior e R menor. Porém para as iterações realizadas iguais ou maiores que 200 o erro (E) praticamente permaneceu constante, assim como suas correlações. Pode-se afirmar que não ocorreu sobre-treinamento.

Concluindo toda a configuração da rede, com melhor performance para a extração de níquel e cobalto do processo da Votorantim, utilizando solução orgânica cyanex 272 é:

- 3 Neurônios na Camada Intermediária - Rede mais simples
- Rede MLP Generalizada
- Função de Ativação: Tangente Hiperbólica
- Algoritmo: Back-propagation - QuickPropagation
- 2 Saídas - 2NiCo

Capítulo 6 - CONCLUSÕES E SUGESTÕES

6.1 – Conclusões

Experimentos de extração de cobalto e níquel para obtenção de seletividade destes metais, em função da temperatura, volume de orgânico e pH foram realizados em intervalo que permita a otimização do processo industrial.

As superfícies de resposta do coeficiente de partição mostraram que através dos experimentos realizados é possível tendenciar condições ótimas para o processo. Estas tendências não permitem estabelecer um intervalo ótimo de maior eficiência de extração do cobalto e menor afinidade com o níquel pela solução orgânica Cyanex 272, com a precisão que a tecnologia de controle hoje possibilita ter. Os resultados obtidos pelos dados experimentais apontam a região ótima de trabalho na proporção de volume de orgânico: volume de solução aquosa igual a 2:1, em pH em torno do intervalo de 5,4 a 5,8 e entre as temperaturas de 60 e 70°C.

As técnicas de RNAs foram aplicadas na modelagem deste processo de extração. Várias arquiteturas foram testadas em busca daquela que gerasse um menor erro para o conjunto de dados disponíveis.

Os baixos erros encontrados e as correlações altas mostram que as redes foram satisfatoriamente treinadas e, portanto, são capazes de prever valores a partir de dados gerados ou para otimizar o processo com valores no intervalo estudado.

Uma vez que as RNAs necessitam de um número e qualidade consideráveis de dados, o motivo de tais resultados pode estar na representatividade não limitada dos dados gerados e sua confiabilidade pelas minuciosas análises e tratamento.

Foi verificado que as redes neurais para as condições aqui apresentadas são uma importante ferramenta na predição de condições de processos, que possibilita a otimização e também melhor controle do processo.

A Rede neural obtida com menor erro e que melhor representa o processo, é de grande simplicidade quanto ao número de neurônios na camada intermediária e apresenta 2 neurônios na camada de saída, ou seja duas variáveis para predição.

Apesar dos resultados apresentarem boa concordância, a integração do modelo ainda será realizada pela VOTORANTIM – CNT (Companhia Níquel Tocantins), que poderá agregar este trabalho ao simulador da planta global e, fazer os ajustes às melhores condições de processo e obter melhor produto pelo menor custo possível.

6.2- Sugestões

- Obtenção de novo banco de dados;
- Verificar a capacidade de generalização da rede através da apresentação de um novo conjunto de dados ;
- Simular a Rede Neural obtida em Software Aspen para obtenção de valores ótimos de processo;
- Levantar outras possíveis variáveis e fazer análise de significância das mesmas;
- Aplicar técnicas estatísticas como PCR (Principal Component Regression), PLS (Partial Least Square) e PCA (Principal Component Analysis) para análise estatística dos dados para depois aplicar as técnicas de RNAs;

REFERÊNCIAS BIBLIOGRÁFICAS

- ✓ BHAT, N. & McAVOY, T.J. Determining Model Structure for Neural Models by Network Stripping. *Computers chem. Engng.* Vol. 16, n. 4, p. 271-281, 1992.
- ✓ BISHOP, C.M. Neural Networks and Their Applications. *Rev. Sci. Instrum.*; Vol. 65, n. 6, USA, June 1994.
- ✓ BRAGA, A.P., LUDEMIR, T. B., CARVALHO, A. C. P. de L. F. *Redes Neurais Artificiais – Teoria e Aplicações*. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2000, 262p.
- ✓ BRUNS, R. E.; SCARMINIO I.S. & NETO, B.B. *Planejamento e Otimização de Experimento*. São Paulo: Editora da Unicamp, 1996, 299 p.
- ✓ BUFO, M. J. *Aplicação de Redes Neural Artificial como auxiliar na Predição do Desempenho de um Landfarming*. 2000. 94f. Dissertação (Mestrado em Engenharia Química) - Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas.
- ✓ BERNARDO. SÉRGIO. Partição dos Aminoácidos L-Fenilalanina, L-Tirosina e L-Triptofano, e da Proteína Insulina Humana, em Sistemas Aquosos Bifásicos Contendo Polímeros, Copolímero e Sais. 130f. Dissertação (Mestrado em Engenharia Química) - Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas, 2000.
- ✓ BHAT, N. & McAVOY, T.J. Determining Model Structure for Neural Models by Network Stripping. *Computers chem. Engng.* Vol. 16, n. 4, p. 271-281, 1992.
- ✓ CHAN, W.M; NASCIMENTO, C.A.O. Use of Neural Networks for modeling of Olefin Polymerization in High Pressure tubular Reactors. 1994
- ✓ CROWE, E. R. VASSILIADIS, C. A. Artificial Intelligence: Starting to Realize Its Practical Promise. *Chemical Engineering Progress*, p.22-31, 1995.

- ✓ CYTEC, Apostila de Cyanex 272 Extractant da Cytec Industries. New Jersey.1996
- ✓ DI MASSIMO, C.; MONTAGUE, G.A.; WILLIS, M.J.; THAM, M.T. & MORRIS, A.J. Towards Improved Penicillin Fermentation via Artificial Neural Networks. *Computers chem. Engng.* Vol. 16, n. 4, p. 283-291, 1992.
- ✓ FUJITA, EDSON G.L., Modelagem de Forno para Produção de Cimento através da Aplicação de Técnicas de Redes Neurais. 120f. Dissertação (Mestrado em Engenharia Química) - Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas,2001.
- ✓ GARSON, G.D. Interpreting Neural-Network Connection Weights. *AI Expert.* Apr, p. 47-51,1991.
- ✓ GONTARSKI, C. A. U. – *Avaliação da utilização de Redes Neurais aplicadas a processos químicos.*2000. 158 f. . Tese (Doutorado em Engenharia Química) - Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas.
- ✓ GONTARSKI, C. A.; RODRIGUES, P.R.; MORI, M. & PRENEM, L.F. Simulation of an Industrial Wastewater Treatment Plant using Artificial Neural Networks. *Computers and Chemical Engineering.* V. 24, p.1719-1723, 2000.
- ✓ HASSIB, B. & STORK, D.G. *Second Order Derivatives for Network Pruning: Optimal Brain Surgeon.* NIPS 5, Ed. S.J.Hanson et al., 164, San Mateo, Morgan Kaufmann, 1993.
- ✓ HYKIN, SIMON. *Redes Neurais: Princípios e Práticas.*2.ed. Porto Alegre: Bookman, 2001, 900 p.
- ✓ KIM, Y.I.; MOON, K.C.; KANG, B.S.; HAN, C. & CHANG, K.S. Application of Neural Network to the Supervisory Control of a Reheating Furnace in the Steel Industry. *Control Engineering Practice.* V.6, p.1009-1014, 1998.
- ✓ KOVÁCS, Z. L. *Redes Neurais Artificiais: fundamentos e aplicações: um texto básico.* 2ª. ed. São Paulo: Edição Acadêmica, 1996.

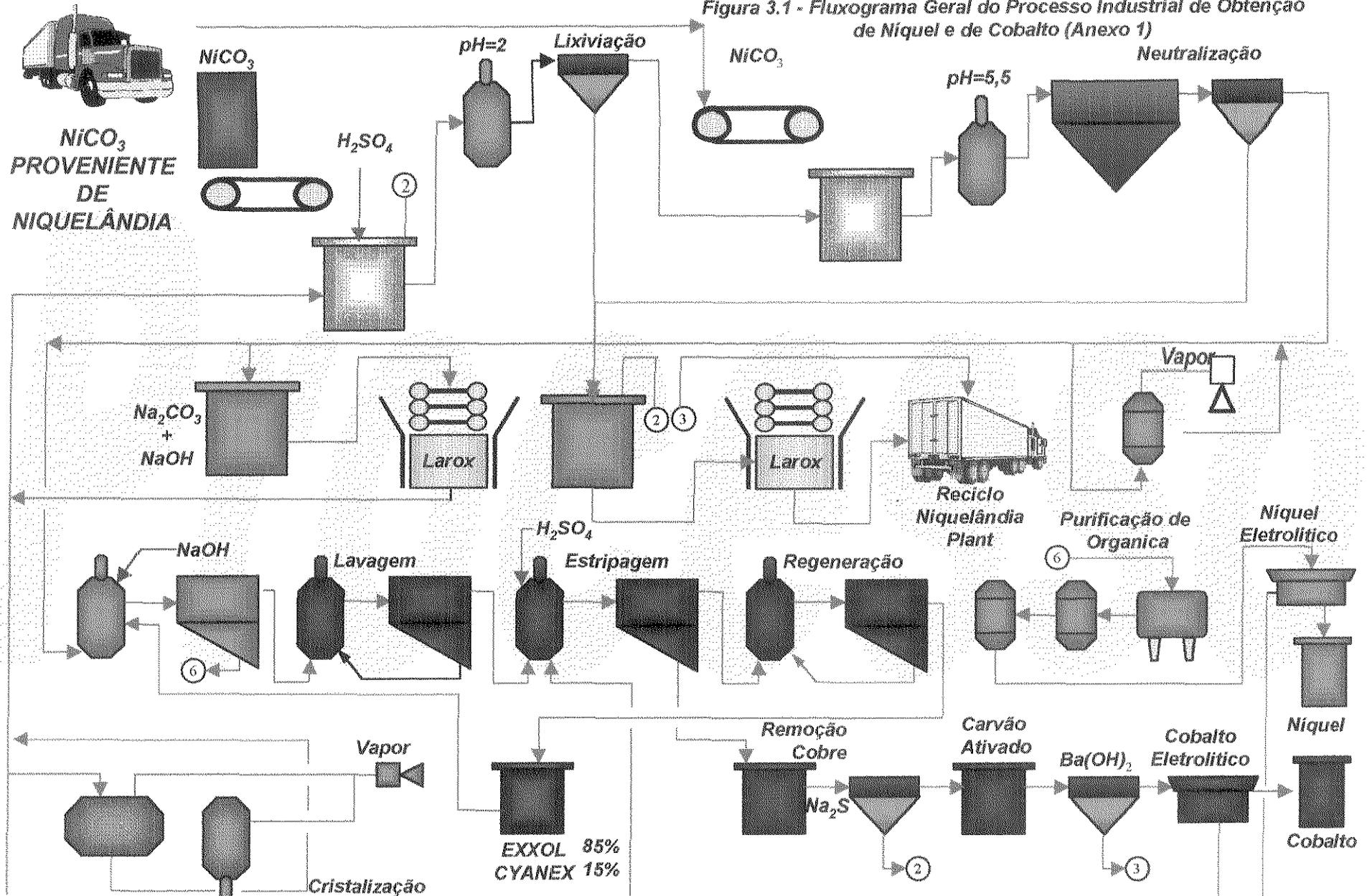
- ✓ LAI, S. & SERRA, M. Concrete Strength Prediction by means of Neural Network. *Construction and Building Materials*. V. 11, n^o. 2, p. 93-98, 1997.
- ✓ MUKESH, D. *Hate Statistics? Try Neural Networks* – Neural networks offer an alternative to mathematical regression for analysing complex data sets. *Chemical Engineering*, p.96-104, 1997.
- ✓ MURKHERJEE, A. & BISWAS, S.N. Artificial Neural Networks in Prediction of Mechanical Behavior of Concrete at High Temperature. *Nuclear Engineering and Design*. V.178, p. 1-11, 1997.
- ✓ NASCIMENTO, C.A.^o; PELLICCI, R.L. Modelling of Industrial Nylon-6,6 Polymerization Process in a Twin-Screw Extruder Reactor. II. Neural Networks and Hybrid Models. *Journal Of Applied of polymer Science*, v. n, p., 1999
- ✓ OLIVEIRA, K.P.S., MORI, M., BRUNS, R.E., Industrial Effluent Treatment System Modeling using PCA and Neural Networks; DPQ - FEQ – UNICAMP; 1999.
- ✓ OLIVEIRA, K. P. S. - *Aplicação das Técnicas de Redes Neurais e de Análise de Componentes Principais na Modelagem de uma Lagoa Aerada da RIPASA S/A*. 2000. 64 f. . Dissertação (Mestrado em Engenharia Química) - Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas.
- ✓ PHAM, D.T. – *An introduction to Artificial Neural Networks*. In: *Neural Networks for Chemical Engineering*. Great Britain. Elsevier Science B. V., 1995. Cap I.
- ✓ POLLARD, J.F; BROUSSARD, M.R.; GARRISON, D.B. & SAN, K.Y. Process Identification using Neural Networks. *Computers chem. Engng*. Vol. 16, n. 4, p. 253-270, 1992.
- ✓ VOTORANTIM. Apostila de laboratório CNT – Companhia Níquel Tocantins. São Paulo-Sp, 2002.
- ✓ WILLIS, M.J.; DI MASSIMO, C.; MONTAQUE, G.A.; THAN, M.T.; MORRIS, A.J. Inferential Measurement via Artificial Neural Networks. *IFAC Intelligent Tuning and Adaptive Control*, Singapore, 1991.

ANEXOS



Fluxograma Geral

Figura 3.1 - Fluxograma Geral do Processo Industrial de Obtenção de Níquel e de Cobalto (Anexo 1)



Recebimento de Carbonato

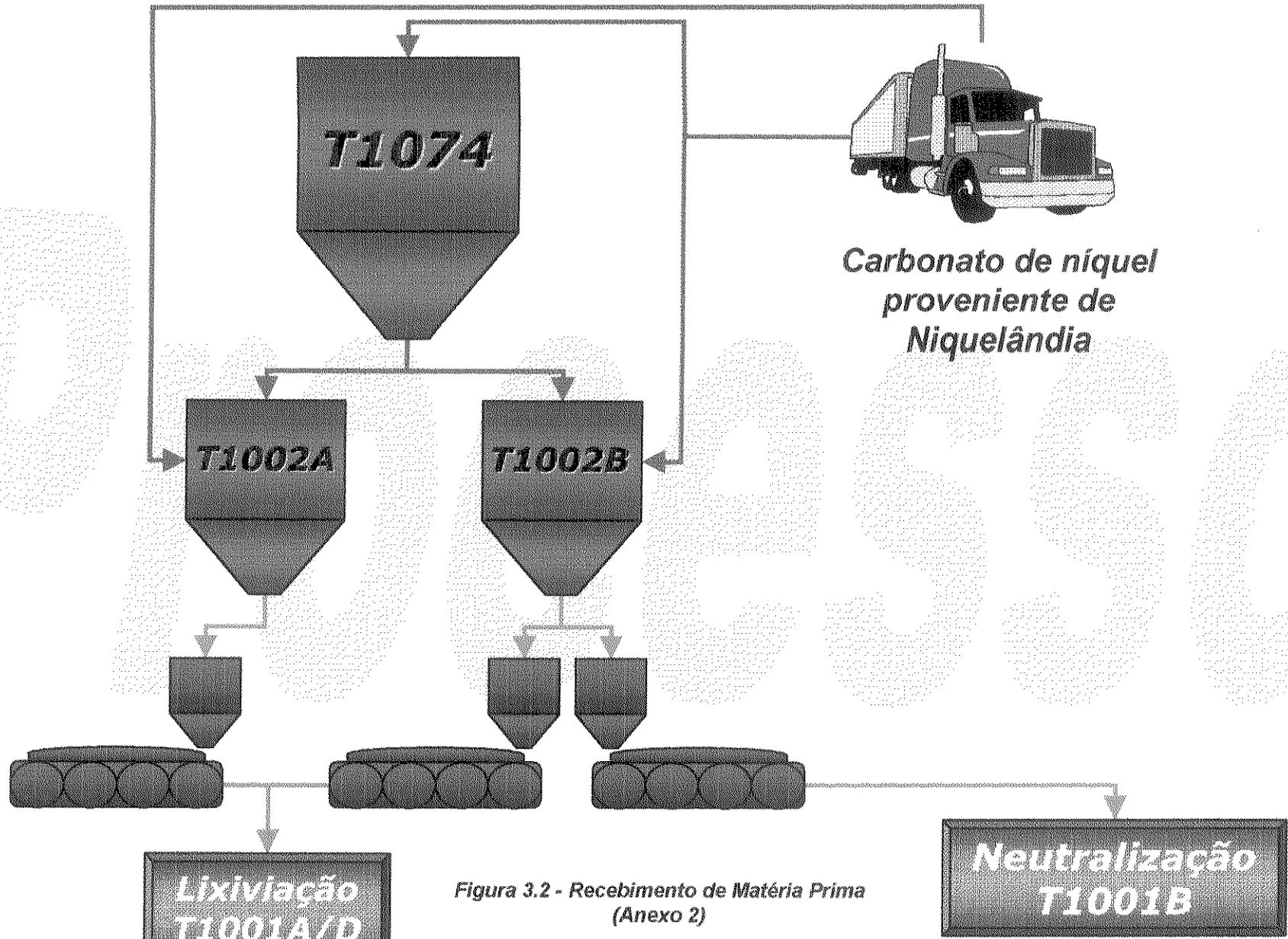


Figura 3.2 - Recebimento de Matéria Prima (Anexo 2)

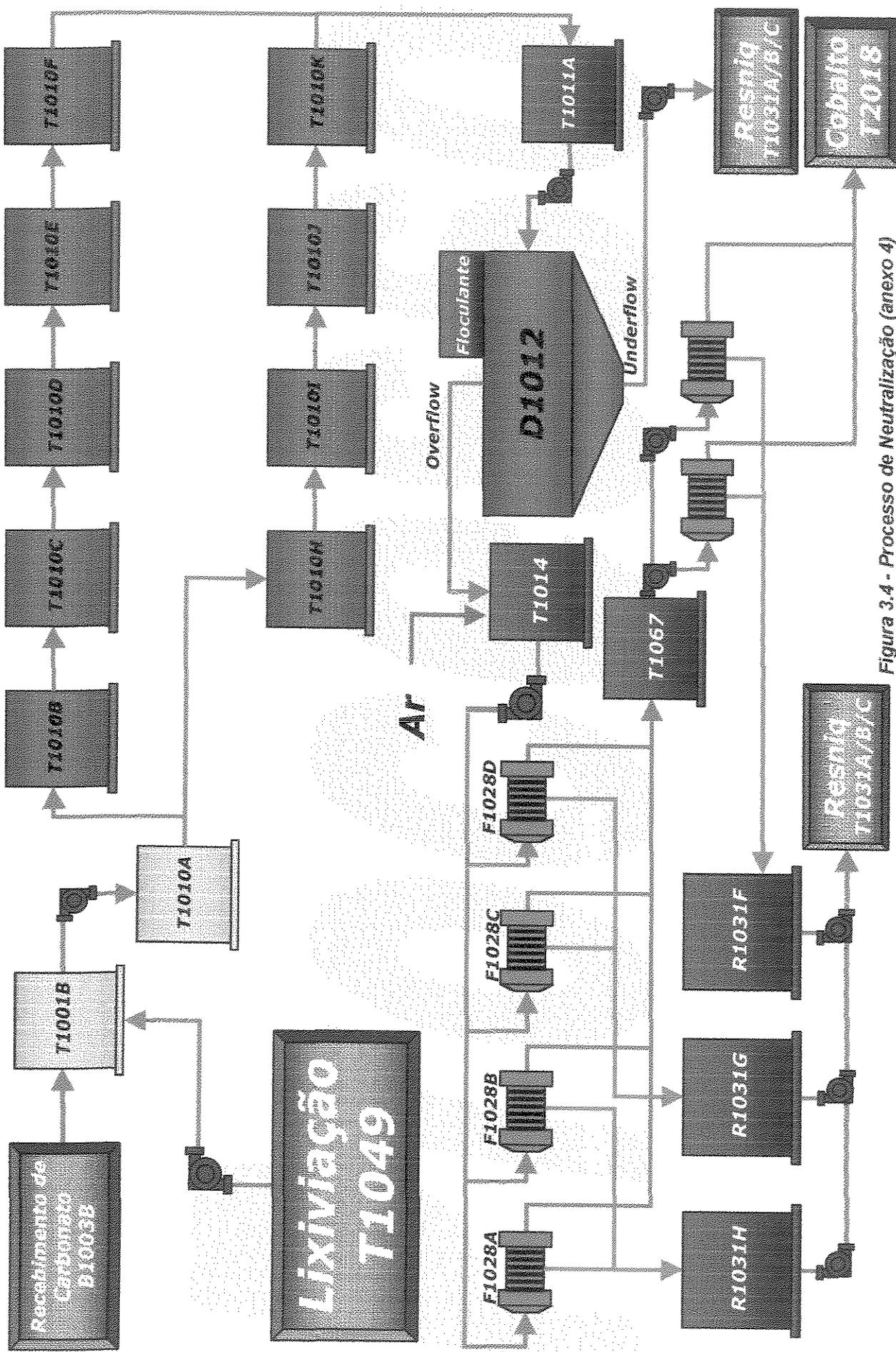


Figura 3.4 - Processo de Neutralização (anexo 4)

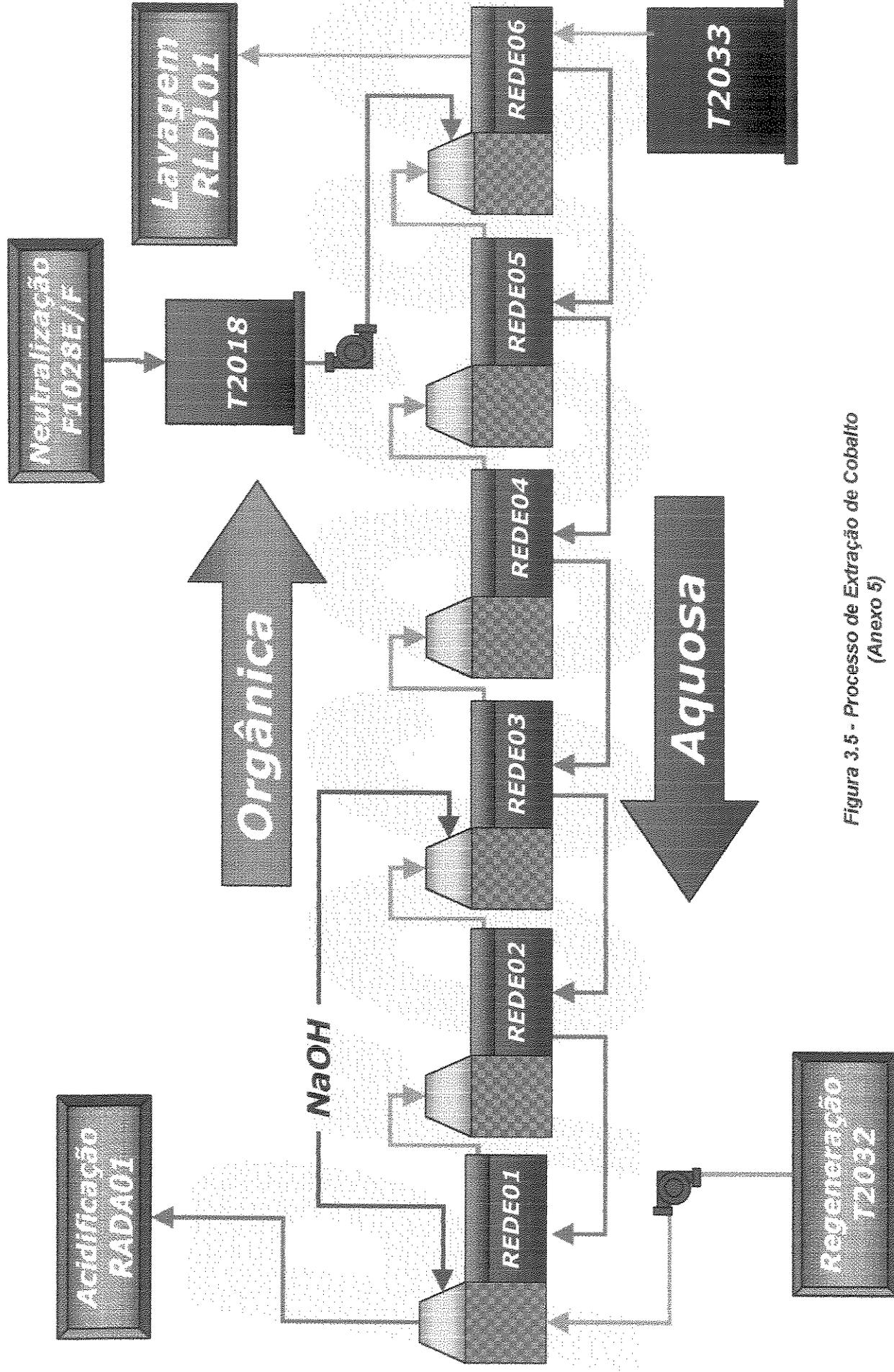


Figura 3.5 - Processo de Extração de Cobalto (Anexo 5)

Cobalto - Acidificação/ Clarificação

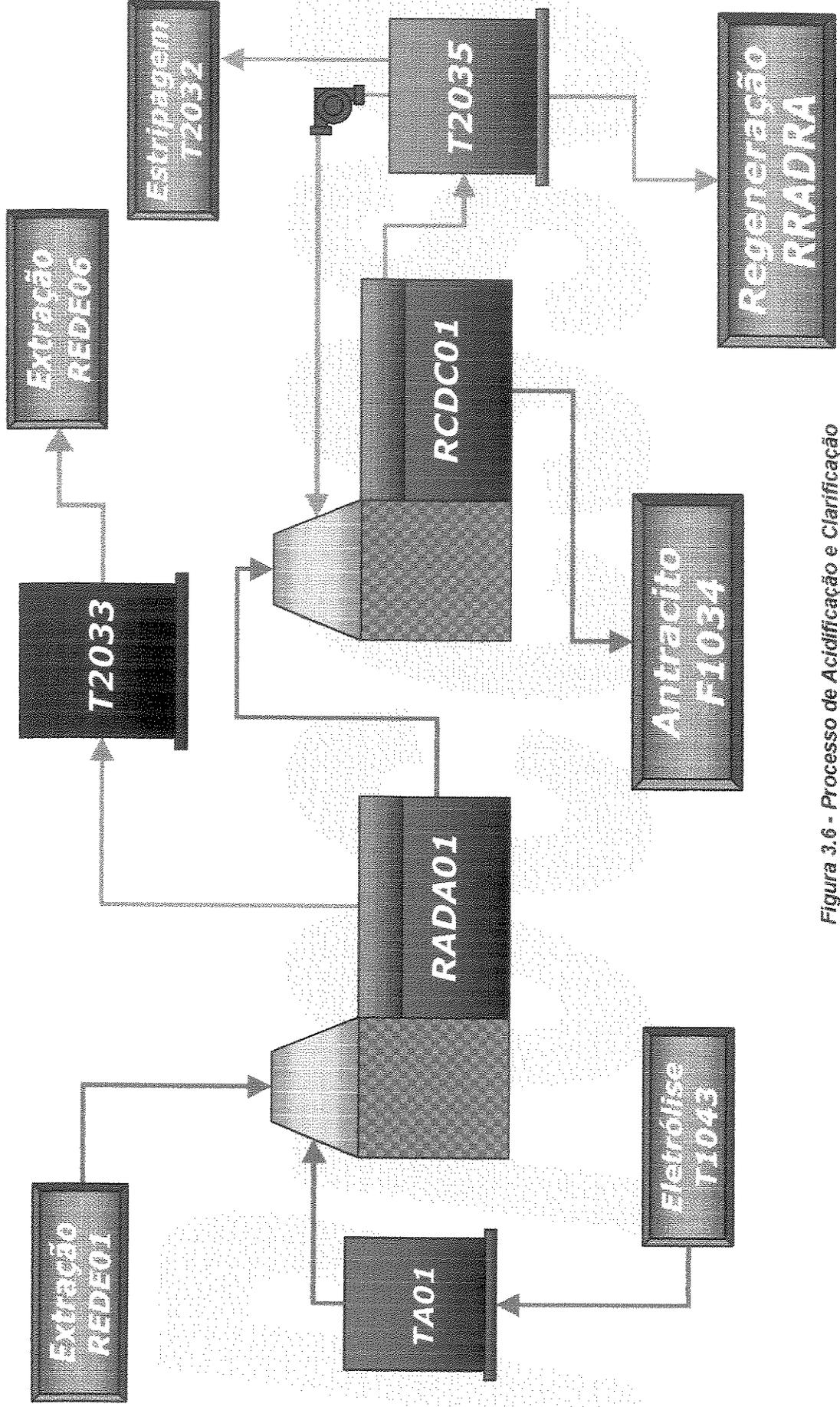


Figura 3.6 - Processo de Acidificação e Clarificação
(Anexo 6)

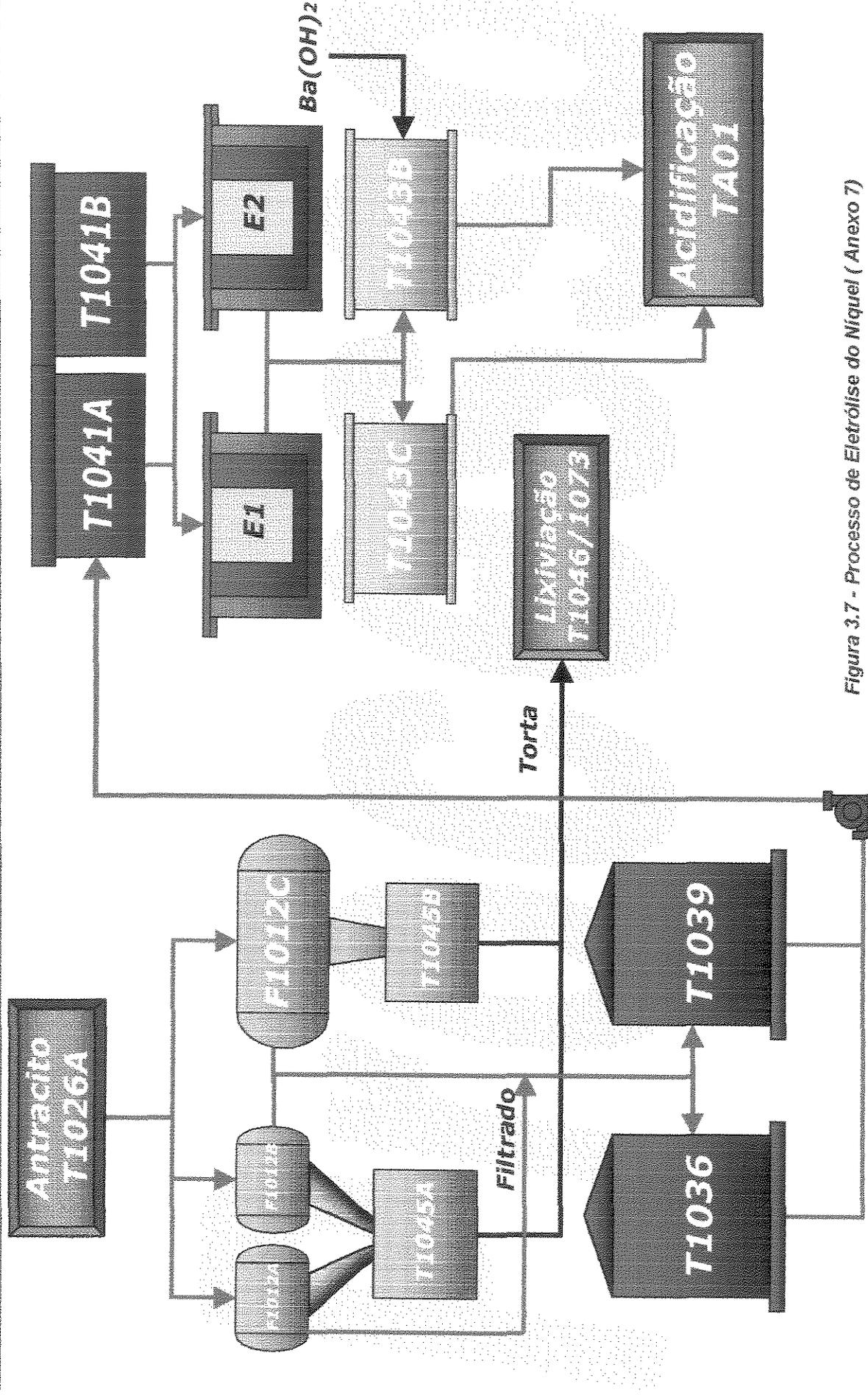


Figura 3.7 - Processo de Eletrolise do Níquel (Anexo 7)

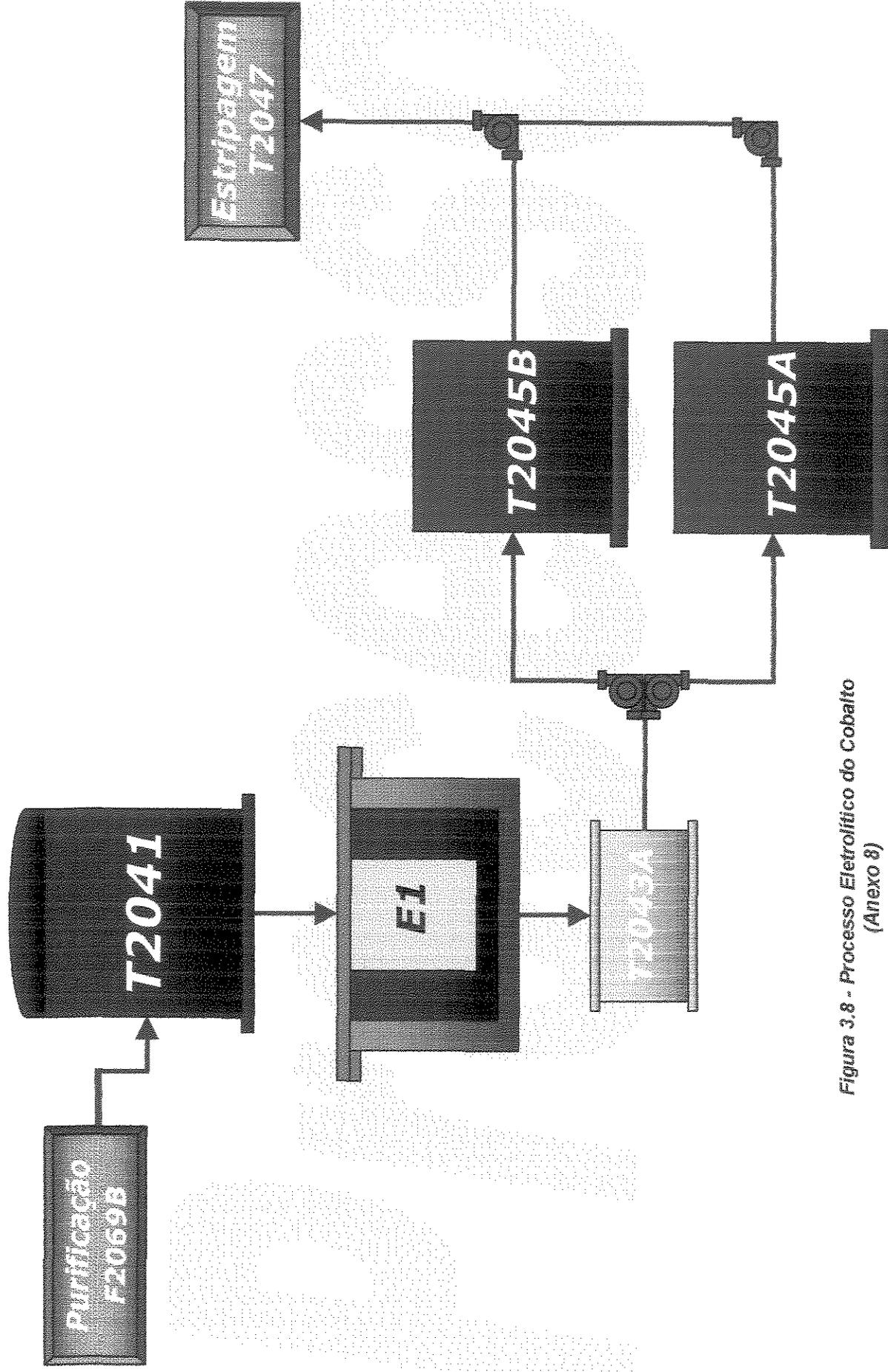


Figura 3.8 - Processo Eletrolítico do Cobalto (Anexo 8)

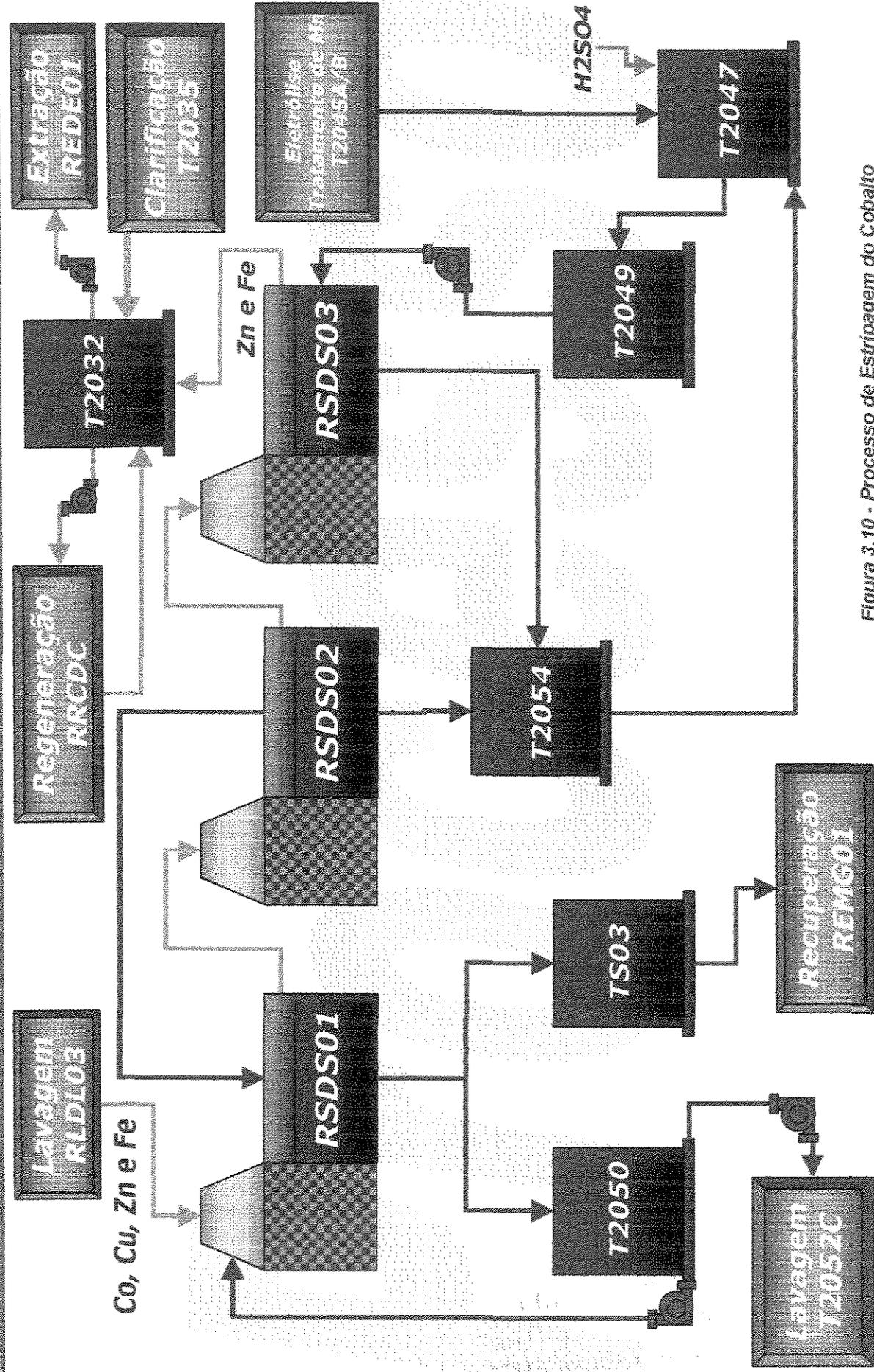
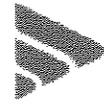


Figura 3.10 - Processo de Estripagem do Cobalto (Anexo 10)