

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO TECNOLÓGICO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

ALESSANDRO BERTOLANI OLIVEIRA

**MODELO DE PREDIÇÃO PARA ANÁLISE COMPARATIVA DE  
TÉCNICAS NEURO-*FUZZY* E DE REGRESSÃO**

VITÓRIA  
2010

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

ALESSANDRO BERTOLANI OLIVEIRA

**MODELO DE PREDIÇÃO PARA ANÁLISE COMPARATIVA DE  
TÉCNICAS NEURO-*FUZZY* E DE REGRESSÃO**

Tese apresentada ao Programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Orientador: Prof. Dr. Thomas Walter Rauber.

VITÓRIA  
2010

**ALESSANDRO BERTOLANI OLIVEIRA**

**MODELO DE PREDIÇÃO PARA ANÁLISE COMPARATIVA DE  
TÉCNICAS NEURO-FUZZY E DE REGRESSÃO**

Tese submetida ao programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Informática.

Aprovada em 12 de Fevereiro de 2010.

**COMISSÃO EXAMINADORA**

---

**Prof. Dr. Thomas Walter Rauber - Orientador**  
**Universidade Federal do Espírito Santo**

---

**Prof. Dr. Renato Antônio Krohling**  
**Universidade Federal do Espírito Santo**

---

**Prof. Dra. Maria Carolina Monard**  
**Universidade de São Paulo**

*“A vida da cidade moderna é praticamente toda determinada pela linha reta... A curva é cansativa, perigosa e funesta, possui um verdadeiro efeito paralisante... A estrada curva é um resultado arbitrário, fruto do acaso, do descuido, de uma ação puramente instintiva. A estrada retilínea é uma resposta a uma solicitação, é fruto de uma intervenção precisa, de um ato de vontade, um resultado atingido com plena consciência. É algo útil e belo.”*

*Le Corbusier*

*"Não é o ângulo reto que me atrai e nem mesmo a linha reta, dura, inflexível, criada pelo homem. O que me atrai é a curva livre e sensual, a curva que encontro nas montanhas de meu país, no curso sinuoso de seus rios, nas nuvens do céu, nas ondas do mar, no corpo da mulher preferida. De curvas é feito todo o universo, o universo curvo de Einstein"*

*Oscar Niemeyer*

*Qual é a sua Linha?*

*Domenico de Masi*

*Dedico esta conquista a minha mãe símbolo de força, dedicação e amor.*

# **Agradecimentos**

Gostaria de agradecer a todos os colegas de trabalho que supriram minhas atividades em períodos de ausência para dedicação a este trabalho.

Gostaria de agradecer a Samarco Mineração; primeiramente, na pessoa do Gerente Geral Reuber Koury, que prontamente dispôs, sem ressalvas, as permissões necessárias para conquista deste sonho profissional.

Posteriormente, também sou grato a todos os envolvidos da Samarco Mineração, que por este período de estudo e aprimoramento, apoiaram este investimento como co-idealizadores e colaboradores.

O agradecimento especial dedicado a minha mãe, Ana Bertolani Oliveira, como fonte de esperança e determinação e em instância superior, agradecer a Deus por abrir portas que não poderiam ser abertas somente por minhas forças, conciliar pensamentos contraditórios e apaziguar situações conflitantes que surgiram neste período de aprendizagem.

## **Resumo**

Neste trabalho são investigadas estratégias para a elaboração de Modelos de Predição que possam ser utilizados no monitoramento de uma variável de qualidade pertencente a um determinado Processo Produtivo Industrial. Neste cenário, a variável de qualidade é estimada por meio de técnicas da Inteligência Computacional e empiricamente avaliada na resolução de problemas de regressão.

A principal contribuição desta monografia é a análise comparativa de Técnicas da Inteligência Computacional associadas às estratégias heurísticas de treinamento para a construção dos Modelos de Predição. São propostas duas linhas de pesquisa investigadas a partir de uma pesquisa empírica dos dados, e analisados a partir de dois grandes ramos da Inteligência Computacional – Aprendizagem de Máquina e Redes Neurais Híbridas. Os Modelos de Predição desenvolvidos são protótipos conceituais para potencial implementação de Sistemas Inteligentes em tempo real de uma planta industrial.

O método de construção dos Modelos de Predição por técnicas de Regressão é comparado com o método de construção do Modelo de Predição por redes Neuro-Fuzzy e analisados por critérios estabelecidos a partir de ferramentas estatísticas que levam em consideração os níveis de adequação e generalização dos mesmos. Ao final, são apresentados resultados dos métodos implementados sobre a mesma base de dados bem como os pertinentes trabalhos futuros.



## **Abstract**

We investigate strategies to define prediction models for a quality parameter of an industrial process. We estimate this variable using computational intelligence and in special regression methods.

The main contribution of this paper is the comparative analysis of heuristic training models to create the prediction system. We propose two main paradigms to obtain the system, machine learning and hybrid artificial neural networks. The resulting system is a prototype for the intelligent supervision of a real-time production process.

Statistical tools are used to compare the performance of the regression based predictor and the neuro-fuzzy based predictor, considering the degree of adaptation of the system to the problem and its generalization ability.

# Sumário

Lista de Figuras .....	xi
Lista de Tabelas .....	xi
Nomenclatura .....	xiv
Capítulo 1: Introdução .....	22
1.1 A pirâmide do conhecimento .....	23
1.2 Monitoramento e controle de Processos Industriais .....	25
1.3 Trabalhos Relacionados .....	27
1.4 Objetivos do Trabalho .....	28
1.5 Organização deste Texto .....	29
Capítulo 2: O Processo Industrial de Pelotização .....	31
2.1 O ambiente Industrial de Pelotização .....	31
2.1.1 Os Estágios dos processos de Pelotização .....	31
2.2 Monitoramento e controle de planta Industrial .....	34
2.2.1 As variáveis de processo nos diversos níveis .....	35
2.3 Soluções em Automação para Modelos de Predição .....	37
2.3.1 Trabalhos correlatos .....	40
Capítulo 3: Técnicas de construção de Modelos de Predição .....	41
3.1 Introdução .....	41
3.2 Técnicas de Regressão clássicas e avançadas .....	45
3.2.1 Modelo de Predição clássico por Regressão Linear Múltipla .....	47
3.2.3 Máquina de Kernel Esparso .....	50
3.2.3.1 Regressão por Máquina de Vetor de Suporte .....	51
3.2.3.2 Regressão por Máquina de Vetor de Relevância .....	54
3.3 Técnica híbrida Neural-Adaptativa de regressão .....	58
3.3.1 O Modelo de Predição Neuro-Fuzzy .....	61
3.4 Síntese das técnicas de construção de Modelos de Predição .....	64
Capítulo 4: Medidas de adequação dos Modelos de Predição .....	68
4.1 Introdução .....	68
4.2 Análise Residual .....	71

Capítulo 5: Resultados dos Algoritmos de Aprendizagem .....	74
5.1 Configuração dos experimentos.....	77
5.1.1 Base de dados conceituais do problema.....	79
5.1.2 Redução da dimensionalidade por algoritmo de seleção .....	81
5.1.3 Heurística de formação dos conjuntos de dados.....	82
5.2 Resultado dos Modelos de Predição.....	83
5.2.1 Resultado experimental das Técnicas de Regressão.....	84
5.2.2 Resultado experimental da Técnica Neuro- <i>Fuzzy</i> .....	95
5.3 Conclusões .....	101
5.4 Trabalhos futuros .....	102
 Referências Bibliográficas.....	 104
 Apêndice A: Análise dos Dados .....	 108
Apêndice B: Resultado do algoritmo de seleção de característica .....	116

## Lista de Figuras

<b>Figura 1.1</b> Estrutura DIKW .....	23
<b>Figura 1.2</b> Visão da hierarquia DIKW .....	24
<b>Figura 1.3</b> Topologia da Automação e seus elementos.....	25
<b>Figura 2.1</b> Processo de formação da Pelota de Minério de Ferro.....	35
<b>Figura 2.2</b> Vista sistêmica da empresa e os estágios de formação da Pelota de Minério de Ferro .....	39
<b>Figura 2.3</b> Pirâmide da Automação .....	34
<b>Figura 2.4</b> Nível Um: Sensores / Atuadores digitais e analógicos .....	36
<b>Figura 2.5</b> Incremento de valor com o avanço no controle de processo .....	38
<b>Figura 3.1</b> Procedimento de identificação de Modelo de Predição .....	41
<b>Figura 3.2</b> Hierarquia dos Sistemas de Controle da Automação Industrial .....	45
<b>Figura 3.3</b> Exemplo de <i>Overfitting</i> e Generalização em problemas de regressão .....	49
<b>Figura 3.4</b> Exemplo de uma função de erro $\varepsilon$ -insensitive .....	51
<b>Figura 3.5</b> Ilustração da SVM para Regressão .....	52
<b>Figura 3.6</b> Ilustração da $\nu$ -SVM para Regressão .....	53
<b>Figura 3.7</b> O mesmo ambiente de teste da $\nu$ -SVM para Regressão aplicado ao RVM .....	56
<b>Figura 3.8</b> As Máquinas de <i>Kernel</i> Esparsos aplicadas a um problema de Regressão .....	57
<b>Figura 3.9</b> Relação dos conceitos da teoria <i>Fuzzy</i> .....	59
<b>Figura 3.10</b> Topologia do Sistema FIS .....	60
<b>Figura 3.11</b> Estrutura do Sistema ANFIS .....	62
<b>Figura 3.12</b> Exemplo de Regressão Linear Múltipla .....	64
<b>Figura 3.13</b> Exemplo de SVM para regressão .....	66
<b>Figura 3.14</b> Exemplo de RVM para regressão .....	67
<b>Figura 4.1</b> Gráficos da Análise Residual.....	72

<b>Figura 5.1</b> Gráficos da Compressão Mecânica da Pelota de Minério de Ferro a frio.....	77
<b>Figura 5.2</b> Exemplo hipotético de amostragem da variável de qualidade.....	78
<b>Figura 5.3</b> Visão sistêmica das variáveis de processo que impactam o parâmetro de qualidade da pelota de minério de ferro: Compressão Mecânica a frio .....	80
<b>Figura 5.4</b> Visão geral da formação dos conjuntos de dados utilizados no Treinamento, Validação e Testes dos Modelos de Predição.....	83
<b>Figura 5.5</b> Análise Residual para Regressão Linear Múltipla com trinta variáveis .....	92
<b>Figura 5.6</b> Análise Residual para Regressão Linear Múltipla com quatro variáveis.....	92
<b>Figura 5.7</b> Análise Residual para Máquina de Vetor de Suporte com trinta variáveis .....	93
<b>Figura 5.8</b> Análise Residual para Máquina de Vetor de Suporte com quatro variáveis .....	93
<b>Figura 5.9</b> Análise Residual para Máquina de Vetores Relevância com trinta variáveis .....	94
<b>Figura 5.10</b> Análise Residual para Máquina de Vetores Relevância com quatro variáveis ...	94
<b>Figura 5.11</b> Análise Residual para ANFIS com trinta variáveis .....	100
<b>Figura 5.12</b> Análise Residual para ANFIS com quatro variáveis .....	100
<b>Apêndice A-G1:</b> Distribuição da Probabilidade: Variável Dependente (Conjunto Treinar)	113
<b>Apêndice A-G2:</b> Distribuição da Probabilidade: Variável Dependente (Conjunto Validar)	114
<b>Apêndice A-G3:</b> Distribuição da Probabilidade: Variável Dependente (Conjunto Testar)..	115

## Lista de Tabelas

<b>Tabela 3.1</b> Variável Independente e Dependente .....	48
<b>Tabela 3.2</b> Etapas de aprendizagem – ANFIS .....	63
<b>Tabela 5.1</b> Resultado do Modelo de Predição por Regressão Linear Múltipla com trinta variáveis independentes.....	86
<b>Tabela 5.2</b> Resultado do Modelo de Predição por Regressão Linear Múltipla com quatro variáveis independentes.....	87
<b>Tabela 5.3</b> Resultado do Modelo de Predição por Máquina de Vetores de Suporte com trinta variáveis independentes.....	88
<b>Tabela 5.4</b> Resultado do Modelo de Predição por Máquina de Vetores de Suporte com quatro variáveis independentes.....	89
<b>Tabela 5.5</b> Resultado do Modelo de Predição por Máquina de Vetores de Relevância com trinta variáveis independentes.....	90
<b>Tabela 5.6</b> Resultado do Modelo de Predição por Máquina de Vetores de Relevância com quatro variáveis independentes.....	91
<b>Tabela 5.7</b> Resultado do Modelo de Predição por ANFIS com trinta variáveis independentes - Grupo 1 ao 5 .....	96
<b>Tabela 5.8</b> Resultado do Modelo de Predição por ANFIS com trinta variáveis independentes - Grupo 6 ao 10 .....	97
<b>Tabela 5.9</b> Resultado do Modelo de Predição por ANFIS com quatro variáveis independentes - Grupo 1 ao 5 .....	98
<b>Tabela 5.10</b> Resultado do Modelo de Predição por ANFIS com quatro variáveis independentes - Grupo 6 ao 10 .....	99
<b>Tabela A-T1</b> Variável de Processo .....	108
<b>Tabela B-T1</b> Resultado do seleção de característica com kernel RVM.....	116
<b>Tabela B-T2</b> Resultado do seleção de característica com kernel SVM .....	117
<b>Tabela B-T3</b> Resultado do seleção de característica com kernel ANFIS .....	118

# Nomenclatura

## Símbolos subscritos

Símbolo	Descrição	Exemplo
$k$	Iteração	$x_k$
$i, j$	Número de identificação ou índices	$x_{ij}$

## Outras Unidades

Unidade	Descrição	Nome
KWh	Demanda	Kilowatts hora
Ton	Peso	Tonelada
Kgf	Pressão	Kilograma força
mmH <sub>2</sub> O	Pressão	Milímetros de Coluna d'água
m/min	Velocidade Escalar	Metros por Minutos
KCAL/TMS	Capacidade Térmica	Kilocalorias por Toneladas Métricas Seca
TMS	Peso	Tonelada Métrica Seca

## Letras Gregas

Símbolo	Descrição	Unidade
$\epsilon$	Erro da função	Escalar
$\beta$	Coefficiente de regressão	Escalar
$\sigma$	Desvio Padrão	Escalar
$\Phi$	Função Canônica	

### Outros símbolos

Símbolo	Descrição	Exemplo
$\mathcal{X}$	Variável Independente	$x_1, x_2, x_k$
$f$	Função	
$y$	Variável Dependente	
$\beta$	Coefficiente de regressão	
$\hat{y}$	Valor Estimado	
$n$	Amostras	
$w$	Vetor de pesos	
$v$	Coefficiente Regulação	
$\xi$	Variável Slack	
$\lambda$	Coefficiente Regulação	
$C$	Parâmetro de Regulação	
$p$	Probabilidade	
$b$	Bias	
$a$	Hiperparâmetros	
$\Sigma$	Covariância	
$R^2$	Coefficiente de Determinação	

### Símbolos sobrepostos

Símbolo	Descrição	Exemplo
$\hat{\phantom{x}}$	Valor Estimado	$\hat{\beta}$
$'$	Matriz Transposta	$\mathcal{X}'$
$^\circ$	Graus Celsius	$^\circ\text{C}$

### Siglas

Símbolo	Descrição
DIKW	<i>Data – Information – Knowledge – Understanding – Wisdom</i>
DCS	<i>Distributed Control System</i>
APC	<i>Advanced Process Control</i>
MPC	<i>Model Predictive Control</i>
PID	Proporcional-Integral-Derivativo



SISO	<i>Single Input/Single Output</i>
PAC	<i>Probably Approximately Correct</i>
SVM	<i>Support Vector Machine</i>
RVM	<i>Relevance Vector Machine</i>
FIS	<i>Fuzzy Inference System</i>
MF	<i>Membership Function</i>
ANFIS	<i>Adaptive Neuro-Fuzzy Inference Systems</i>
LSE	<i>Least Squared Error</i>
RMS	<i>Root Mean Square</i>

# Capítulo 1: Introdução

Ao longo dos anos, o setor que cresce significativamente é o industrial. O aumento da população demanda um aumento encadeado nos setores produtivos, principalmente dos produtos industrializados. Com essa visão, o setor industrial tem buscado a melhoria contínua de seus processos e produtos utilizando-se para isso de recursos automatizados que possam mantê-lo alinhado com essa visão. Esses recursos automatizados estão presentes e perfazem todas as etapas na cadeia produtiva com intuito de otimizar o monitoramento e controle de todos os processos envolvidos na elaboração do produto final.

O monitoramento e controle do processo envolvem uma gama de variáveis que surge em diversos níveis e se destinam as mais variadas grandezas de ordem físico-química. Além disso, essas variáveis representam aspectos e características importantes em relação ao domínio do negócio a que estão destinadas, sendo imprescindíveis na obtenção de processos, materiais, insumos e mão-de-obra otimizados e produtos com alto grau de produtividade e qualidade. Neste contexto, o papel da Automação Industrial torna-se crucial para alcançar o nível de otimização desejado ao longo de todas essas etapas.

Diante do exposto, as pessoas responsáveis pela Automação Industrial devem buscar recursos e estruturar um ambiente de pesquisa e desenvolvimento que possa utilizar-se de técnicas da Inteligência Computacional visando à predição das variáveis de processo. Esse ambiente sistêmico englobaria interativamente pesquisas que alinhassem visões do processo do negócio bem como a utilização de mais avançadas técnicas na área de Automação de controle avançado de processos.

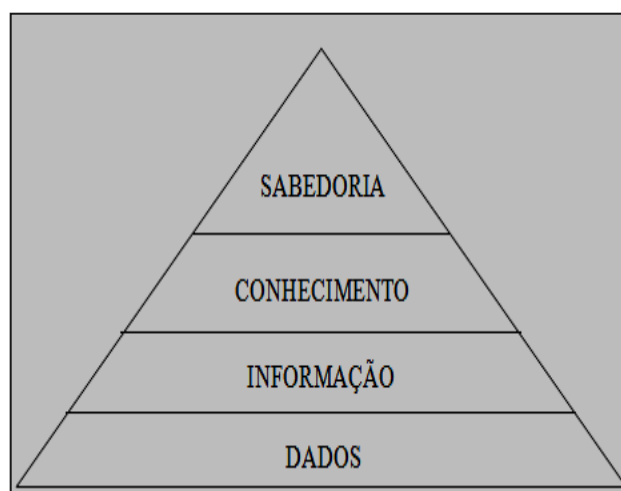
A diversidade de técnicas de monitoramento e controle que poderiam ser estudadas neste ambiente de pesquisa surtiria os resultados almejados na otimização do processo com uso de ferramentas estatísticas que consolidariam os mesmos na planta industrial. Em particular, neste trabalho é feita a pesquisa de algoritmos de Aprendizagem de Máquina destinados a problemas de Regressão alinhada a técnicas da Inteligência Computacional visando à predição de variáveis de processo de uma planta industrial.

Entretanto, a predição, no sentido da antecipação dos fatos e eventos, é conquistada por meio de inteligência advinda de técnicas estatísticas que se utilizam de dados, informações e até conhecimentos que devem estar bem organizados e hierarquicamente compreendidos dentro de um processo produtivo industrial.

## 1.1 A pirâmide do conhecimento

As variáveis presentes em um Processo Industrial fornecem dados que são tratados com diferentes graus de relevância. No contexto da Automação Industrial, *processos* são descritos como qualquer operação onde pelo menos uma propriedade física ou química possa variar ao longo do tempo (NATALE, 2007).

Essas variáveis, que são do tipo discreto ou contínuo, possuem diferentes graus de importância no processo de acordo com a qualidade da informação que está atrelada ao domínio do negócio. Esses graus de importância podem ser qualitativamente hierarquizados conforme proposto na Figura 1.1.



**Figura 1.1** Estrutura DIKW (FRICKÉ, 2007)

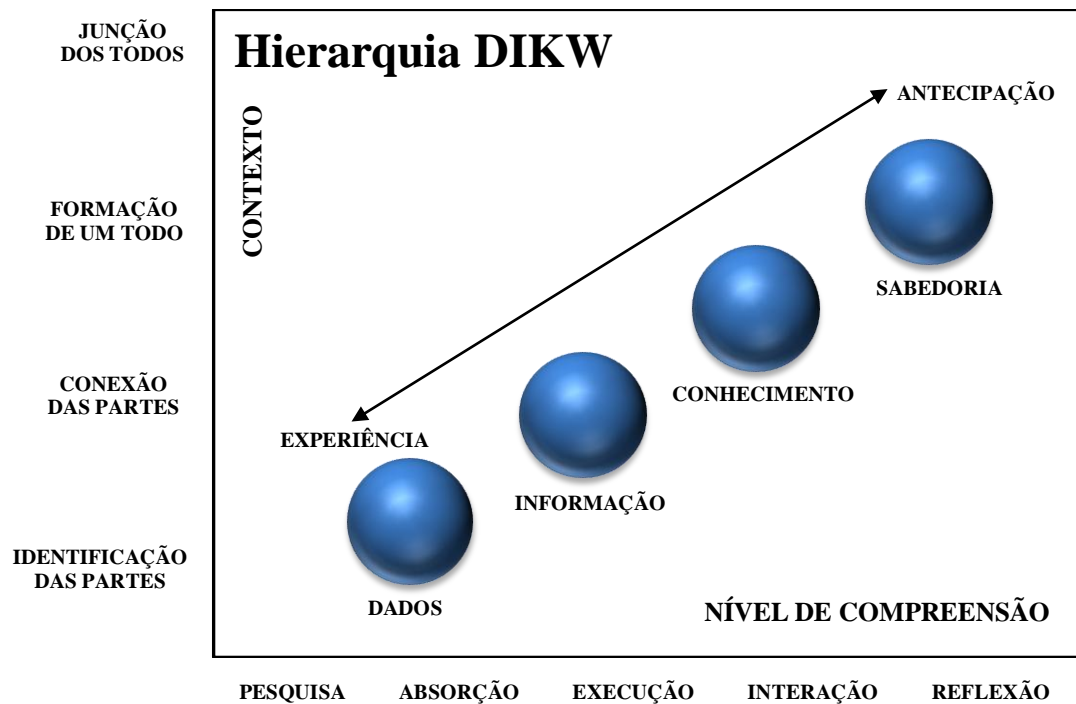
Na estrutura do conhecimento de Russel, do inglês: *Data – Information – Knowledge – Understanding – Wisdom Hierarchy*, como é conhecida perfaz também uma similar abstração de como as variáveis podem se comportar (ACKOFF, 1989). Entretanto, na prática, observamos também a aplicabilidade de Sistemas Inteligentes, linha de pesquisa da Inteligência Computacional, o que denota uma interação hierárquica destes conceitos com uma lacuna em suas ordens propostas originalmente (CAMPOS & SAITO, 2004).

Essa lacuna é preenchida pelo conceito de inteligência proposto e utilizado por esse tipo de sistema que interage com os sistemas de controle de processo buscando os níveis de aperfeiçoamento almejados em todas as etapas (MORAES & CASTRUCCI, 2007). Com a introdução do conceito de inteligência, em substituição ao conceito de

compreensão a hierarquia anteriormente mencionada torna-se mais equidistante e se aproxima das soluções de sistemas propostas na prática.

O conceito de sabedoria é mais avançado e requer pesquisas que possam melhor traduzi-lo em requisitos que pudessem ser efetivamente implementados na prática em projetos de sistemas. Essas questões fogem à priori ao escopo dos objetivos deste trabalho (ACKOFF, 2006).

Os conceitos supracitados se interagem como mostrado na Figura 1.2 com o propósito final de se obter a antecipação dos fatos. Essa antecipação é almejada pelos mais diversos sistemas e utilizam-se para isso de aspectos referentes a inteligência e assim conseqüentemente, pesquisas envolvendo a Inteligência Computacional aplicada a Aprendizagem de Máquina (GRANA, 2009).



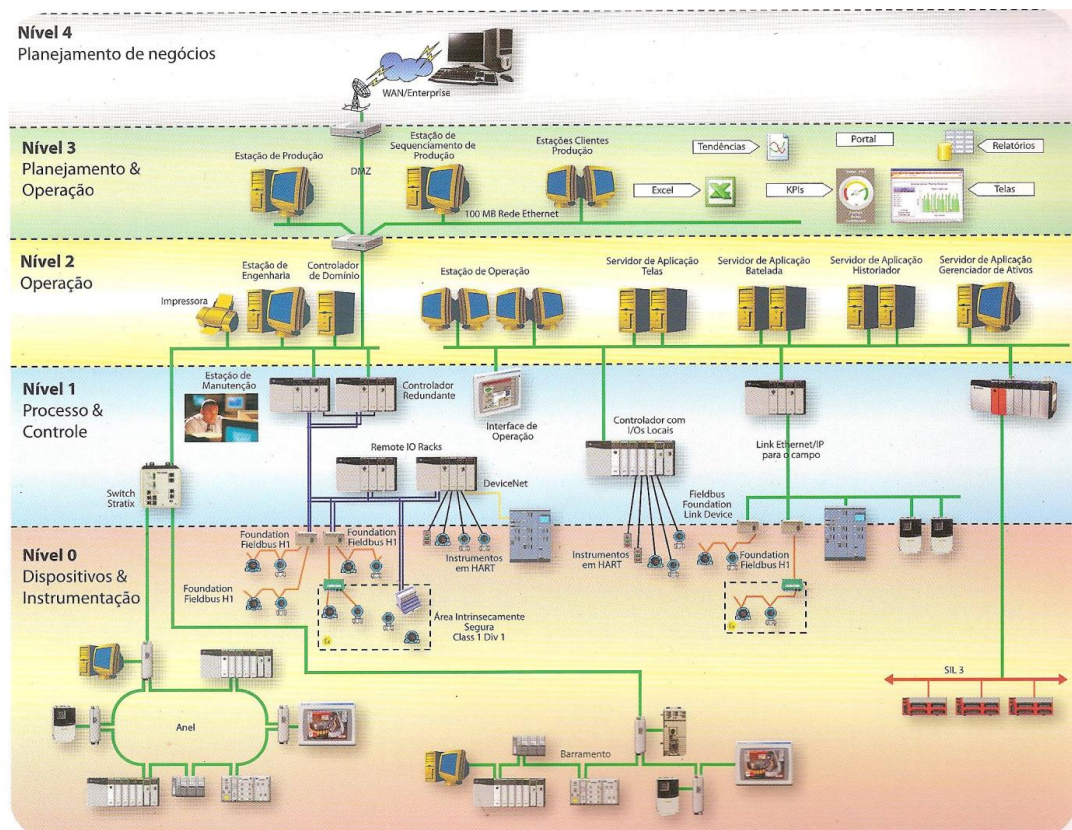
**Figura 1.2** Visão da Hierarquia DIKW (HEY, 2004)

Os sistemas de automação industrial, destinados ao monitoramento e controle de processos, estão buscando aperfeiçoar suas técnicas através de projetos de sistemas inteligentes que possam prover uma visão sistêmica do domínio do negócio. Essa visão, somente é completa até no nível da antecipação – Predição – objetivando alcançar processos com alto desempenho e qualidade.

Com isso, num ambiente de Engenharia de Automação moderno destinado a pesquisa do processo devem gerenciar o conhecimento industrial como um patrimônio valioso compreendido e armazenado de forma organizada, sistêmica e hierárquica, pois somente assim os resultados esperados com a utilização das técnicas estatísticas podem ser claramente evidenciados ao longo de todos os níveis da Pirâmide do Conhecimento.

## 1.2 Monitoramento e controle de Processos Industriais

Nesta parte do trabalho, mencionados previamente alguns conceitos pertinentes, podemos estratificar melhor o monitoramento e controle de Processos Industriais sob a óptica da Automação Industrial e contextualizar os Modelos de Predição neste cenário. A estrutura da Automação Industrial como mostrado na Figura 1.3 segue uma topologia que correlaciona alguns elementos que também perfazem por analogia a mesma abstração hierárquica do conhecimento citada previamente.



**Figura 1.3** Topologia da Automação Industrial e seus elementos (MORAES & CASTRUCCI, 2007)

No nível zero – Rede de Campo – encontramos os componentes elementares tais como: Sensores (Temperatura / Pressão / Velocidade / Entre outros), Atuadores, Sensores de Sinal Digital, Motores / Válvulas e outros, que constituem a identificação das partes (Dados) pormenorizadas dos processos.

Esses elementos são mencionados como dispositivos de Entrada/Saída conectando-se elemento a elemento aos dispositivos controladores. Contudo, apesar da evolução nas diversidades de medidas e grandezas, essa continua sendo uma área que deve desenvolver-se muito com elementos finais ainda mais significativos aos processos (BEGA, 2006). Daí também emerge a idéia de pesquisa por Modelos de Predição que funcionam como "Sensores Inferenciais" no monitoramento da planta (FISKE, 2006).

O nível um é responsável por uma camada denominada de Sistemas de Controle Distribuído (DCS) e iniciam o processo de conexão das partes (Informação) agregando assim mais valor aos dados (MORAES & CASTRUCCI, 2007). Nesta camada, podemos encontrar implementações destinadas ao controle das partes e suas conexões constituindo assim, ainda que de maneira simplória em relação ao todo, contribuições para o controle individualizado dos processos. Por Processo Contínuo entende-se que, o produto final é resultado da interação de diversos processos encadeados contribuindo individualmente na formação do todo (SEBORG, 2000).

No nível dois estão os responsáveis pelo monitoramento do processo (Operação) e incumbidos de analisar todas as etapas desta cadeia (Conhecimento) em tempo real. Neste nível, encontra-se a aplicação direta para os Modelos de Predição principalmente os destinados a inferência aos parâmetros de qualidade do produto final.

No nível de Supervisão / Engenharia (Nível Três) as conexões entre as informações tomam dimensões mais significativas tornando as decisões e assertivas em relação aos processos mais complexas. Neste nível, encontra-se o ambiente que deve ser destinado a pesquisa de técnicas da Inteligência Computacional visando a otimização dos processos.

Esses níveis formam uma camada, na topologia da Automação Industrial, que é denominada pela sigla SCADA (em inglês: *Supervisory Control And Data Acquisition*) e é responsável pelo controle da planta fabril como um todo (MORAES & CASTRUCCI, 2007). A princípio, o nível quatro foge do escopo deste trabalho.

A dimensionalidade e complexidade, alcançada nesta etapa pelas variáveis de processo, direcionam os responsáveis pela Engenharia de Processo e Automação alinharem seus pensamentos no sentido de tornar o processo de monitoramento e controle o mais

autônomo possível (Inteligência). Neste contexto, a pesquisa por técnicas da Inteligência Computacional torna-se fundamental na otimização da planta.

Sistemicamente, neste nível onde há a junção dos todos, surgem ainda dezenas de variáveis que pertencem verdadeiramente ao domínio do negócio e são estas que, trabalhadas podem resultar em processos mais otimizados alinhando qualidade e produção.

### 1.3 Trabalhos Relacionados

O intuito de mencionar os conceitos previamente citados foi uma introdução a dois problemas que surgem na camada de Supervisão e Engenharia: Variáveis mais complexas que estão fortemente interligadas a conceitos envolvendo o domínio do negócio e são impactadas por uma quantidade significativa de outras variáveis, mais simples, formadas por dados e informações, advindas dos processos, no nível de campo. Outro problema, a alta dimensionalidade na análise da supervisão e do controle da planta.

Essas variáveis mais complexas são apenas monitoradas ou são provenientes de ensaios realizados em laboratório com uma periodicidade, inerte ao próprio procedimento, que dificulta a análise e possíveis intervenções no processo que acontece em tempo real.

Outro problema, que os processos produtivos, acabam formando uma quantidade significativa de dados e informações, que se acumulam tomando uma dimensionalidade que se torna difícil sua análise e intervenção em tempo real.

Neste contexto, as soluções relacionadas ao nível de supervisão e controle, remetem a trabalhos relativos ao Controle Avançado de Processo (APC). Em outras citações, podemos encontrar com nome referente à Controle por Modelo de Predição (MPC – do inglês: *Model Predictive Control*) e outros nomes que são, na maioria das vezes, associados a aspectos proprietários da empresa que as desenvolveu (FISKE, 2006).

Os trabalhos relativos a Modelos de Predição para controle de processo passam, de maneira sistêmica, por duas linhas de pesquisa distintas: Pesquisa Formal e a Pesquisa empírica (WAZLAWICK, 2009). A pesquisa formal elabora uma teoria e a confirma através de prova formal utilizando-se para isso de, por exemplo, lógica formal. Suas contribuições não necessariamente estão implementadas em soluções computacionais que possam viabilizar testes em casos práticos no mundo real.

Contrapondo-se a isso, as pesquisas empíricas apresentam novas abordagens, com o desenvolvimento de modelos com técnicas previamente consolidadas, respaldadas por

análises de métodos estatísticos como ferramenta primordial de averiguação da adequação das mesmas.

As pesquisas empíricas correlatas que se utilizam de modelos de predição devem abordar de maneira sistêmica contribuições advindas da Engenharia de Processo e da Engenharia de Automação no intuito de elaborar soluções que agreguem conhecimento, proveniente do domínio do negócio e, técnicas de controle avançado que se utilize de ferramentas da Inteligência Computacional em suas implementações (CAMPOS & SAITO, 2004).

## 1.4 Objetivos do Trabalho

Segundo Nikolaou, todos os sistemas de controle por modelo de predição invocam a idéia de gerar valores para os processos como soluções de um problema de otimização em tempo-real (NIKOLAOU, 2007). Entretanto, a diversidade de soluções que emergem desta real necessidade torna complexa a estratégia de como alcançar esse objetivo, ainda mais que possuem intrinsecamente um caráter interdisciplinar no que se refere à interação de duas distintas engenharias: Automação e Processo.

Com relação à estratégia, podemos salientar que soluções ainda mais complexas tomariam a dimensão de Sistemas Especialistas (do inglês: *Expert Systems*) cujo escopo desta linha de pesquisa foge ao contexto deste trabalho.

O presente trabalho tem como objetivo o estudo de estratégias que levem a Modelos de Predição, baseadas em uma pesquisa empírica, no intuito de comparar, através de métodos heurísticos e estatísticos, os resultados alcançados. Nesse processo investigativo, os Modelos de Predição são destinados a inferência do valor de uma variável de qualidade do produto pertencente a um processo industrial contínuo.

A principal contribuição buscada através deste trabalho é a análise empírica dos modelos obtidos pela utilização de duas abordagens – Técnicas Neuro-*Fuzzy* e de Regressão – aplicada na inferência desta complexa variável de qualidade, intrinsecamente ligada ao domínio do negócio, servindo-se para isso, de uma mesma base de dados.

A princípio, esses Modelos de Predição elaborados, são protótipos conceituais (*Conceptual Frameworks*) para implementação de Modelos de Controle por Predição pertencente a um Controle Avançado de Processo. Seus resultados posteriormente fariam parte do monitoramento realizado em nível de supervisor e a sua consolidação pelos



especialistas de processo se daria analisando efetivamente seu desempenho em um ambiente de tempo real.

As técnicas advindas da Inteligência Computacional utilizadas neste trabalho visam justamente à elaboração desses Modelos de Predição conquistados através de pesquisa e investigação empírica dos dados perfazendo assim aspectos da Aprendizagem de Máquina. Essas técnicas englobam basicamente a Análise de Regressão dos dados por meio de uma base comum aplicada as mesmas com intuito de constituírem o mesmo ambiente de pesquisa no qual se pode determinar seus diferentes desempenhos sob uma mesma condição de teste.

Os diferentes desempenhos obtidos pelos Modelos de Predição extraídos das técnicas da Inteligência Computacional são mensurados por meio de ferramentas estatísticas que apontará para o modelo que melhor generaliza os dados através de uma função de regressão.

## **1.5 Organização deste Texto**

Neste capítulo, foram introduzidos aspectos relacionados aos conceitos macros para compreensão de um ambiente de Automação Industrial e bem como, contextualizá-los dentro dos problemas relativos à elaboração de Modelos de Predição destinados ao Controle Avançado de Processo.

A hierarquia dos conceitos mencionados como abstração da Pirâmide do Conhecimento conduz o trabalho na direção de investigar e propor soluções para predição (Antecipação) e conseqüentemente otimização, de variáveis mais complexas e por decorrência, mais significativas para o processo.

O capítulo dois contém uma explanação global sobre o Complexo Industrial de Pelotização e seus processos contextualizando e direcionando os conceitos citados as interações estabelecidas com a Automação Industrial e a formação da base de dados para treinamento dos Modelos de Predição.

O capítulo três aborda especificamente as técnicas para construção dos Modelos de Predição, suas teorias, definições, benefícios e limitações; bem como, conjunto de parâmetros de ajustes definidos e heurísticas de utilização destes ajustes. Complementarmente, o Capítulo quatro propõe medidas que possa servir de referência para comparação entre os Modelos de Predição, seu nível de adequação e alguns diagnósticos gráficos de avaliação.

Os dados utilizados na pesquisa e construção dos Modelos de Predição através das técnicas *Neuro-Fuzzy* e de Regressão são abordados no Capítulo cinco e formam a base de

dados utilizada em todas as etapas deste trabalho. Também são tratados, aspectos referentes à manipulação desta base de dados, para a sua utilização nas técnicas posteriormente explanadas. O problema da alta dimensionalidade, citado anteriormente, decorrente do nível de complexidade da variável estimada pelos Modelos de Predição também é tratado neste capítulo bem como técnica de resolução do mesmo.

O Capítulo seis disserta sobre a síntese dos métodos utilizados, aspectos referentes às suas comparações e a capacidade de predição alcançada com a sua utilização na referida base de dados. As perspectivas almejadas com a apresentação dos resultados também são explanadas neste capítulo de conclusões.

## Capítulo 2: O Processo Industrial de Pelotização

Neste capítulo é abordado o ambiente responsável por todos os processos extração, beneficiamento e preparação do minério de ferro, sua supervisão e controle, de onde foram extraídas as variáveis utilizadas neste trabalho. Este ambiente totaliza um Complexo Industrial de Mineração responsável pela Pelotização do minério de ferro sendo supervisionado e controlado por um conjunto de equipamentos automatizados que compõem o processo de Automação Industrial. Todo esse complexo compreende a Empresa Samarco Mineração S.A – Unidade Ponta Ubu – Anchieta, ES.

### 2.1 O ambiente Industrial de Pelotização

As Normas de Processo detalhadas no Apêndice A conduzem na prática os processos que formam o Complexo Industrial de Mineração e são responsáveis pelo beneficiamento e preparação do minério de ferro na formação da Pelota de Minério de Ferro (Figura 2.1). Segundo Meyer, Pelotas de Minério de Ferro são esferas ("bolas") produzidas a partir do Minério de Ferro natural ( $Fe_2O_3$ ) e concentrado formado por uma distinta composição química e mineralógica (MEYER, 1980).

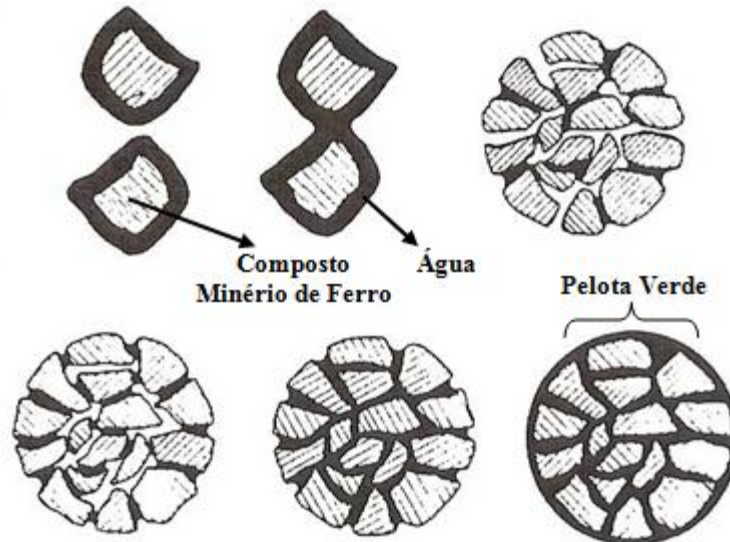
Os próximos tópicos descrevem resumidamente o ambiente industrial de mineração de onde foram extraídas por históricos as informações relevantes pertencentes à base de dados utilizadas pelas técnicas da Inteligência Computacional expostas neste trabalho.

#### 2.1.1 Os Estágios dos processos de Pelotização

No processo de formação da Pelota de Minério de Ferro as etapas são encadeadas em um processo de produção contínuo e a saída de cada etapa deve fornecer o insumo para a etapa seguinte com desvios não significativos em relação ao seu real prosseguimento. Segundo Meyer, para que esse objeto seja atingido, os estágios de processo devem ser definidos, conforme descrito a seguir:

- 1º Estágio: Preparação da matéria-prima

- 2º Estágio: Formação das pelotas verdes
- 3º Estágio: Queima das pelotas verdes



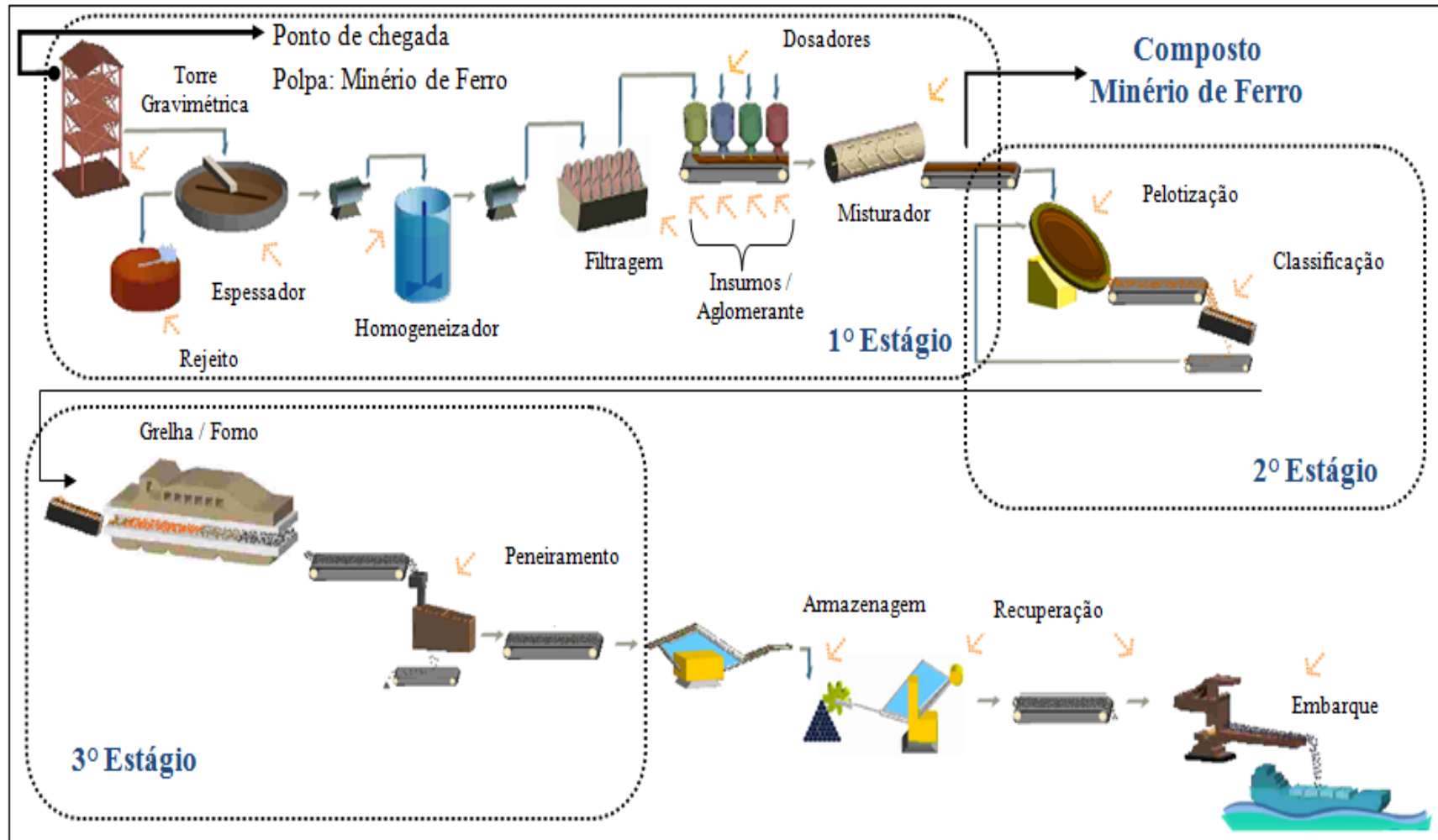
**Figura 2.1** Processo de formação da Pelota de Minério de Ferro (MEYER, 1980)

Os diferentes diâmetros das pelotas são categorizados em faixas (sete faixas) que são mensuradas a partir do processamento de imagem constituindo assim sua análise granulométrica. Esse segundo estágio bem como, todos os processos e equipamentos que o compõem, formam uma área que é comumente denominada de Pelotização.

A Grelha e o Forno perfazem o estágio referente ao processo de endurecimento da Pelota de Minério de Ferro crua. Seu monitoramento e controle ocorrem através de uma série de medições (Dados) realizadas principalmente envolvendo duas grandezas: Temperatura e pressão, mensuradas ao longo das etapas que o totalizam.

Ratificando o que foi escrito anteriormente, todas as informações acima descrevem resumidamente o ambiente industrial de produção contínuo de Pelotas Minério de Ferro de onde foram extraídos; por históricos, os dados utilizados ao longo de toda etapa de pesquisa dos Modelos de Predição obtidos por Aprendizagem de Máquina através de técnicas da Inteligência Computacional sob a linha de pesquisa da Análise de Regressão.

Todos esses macros processos resumidamente descritos formam os três estágios citados e são responsáveis pela formação de uma quantidade significativa de dados e informações que são monitorados e controlados pelos equipamentos que forma o ambiente da Automação Industrial. O conjunto desses elementos é associado formando uma visão sistêmica da empresa (Figura 2.2) Samarco Mineração S.A – Unidade Ponta Ubú – Anchieta.

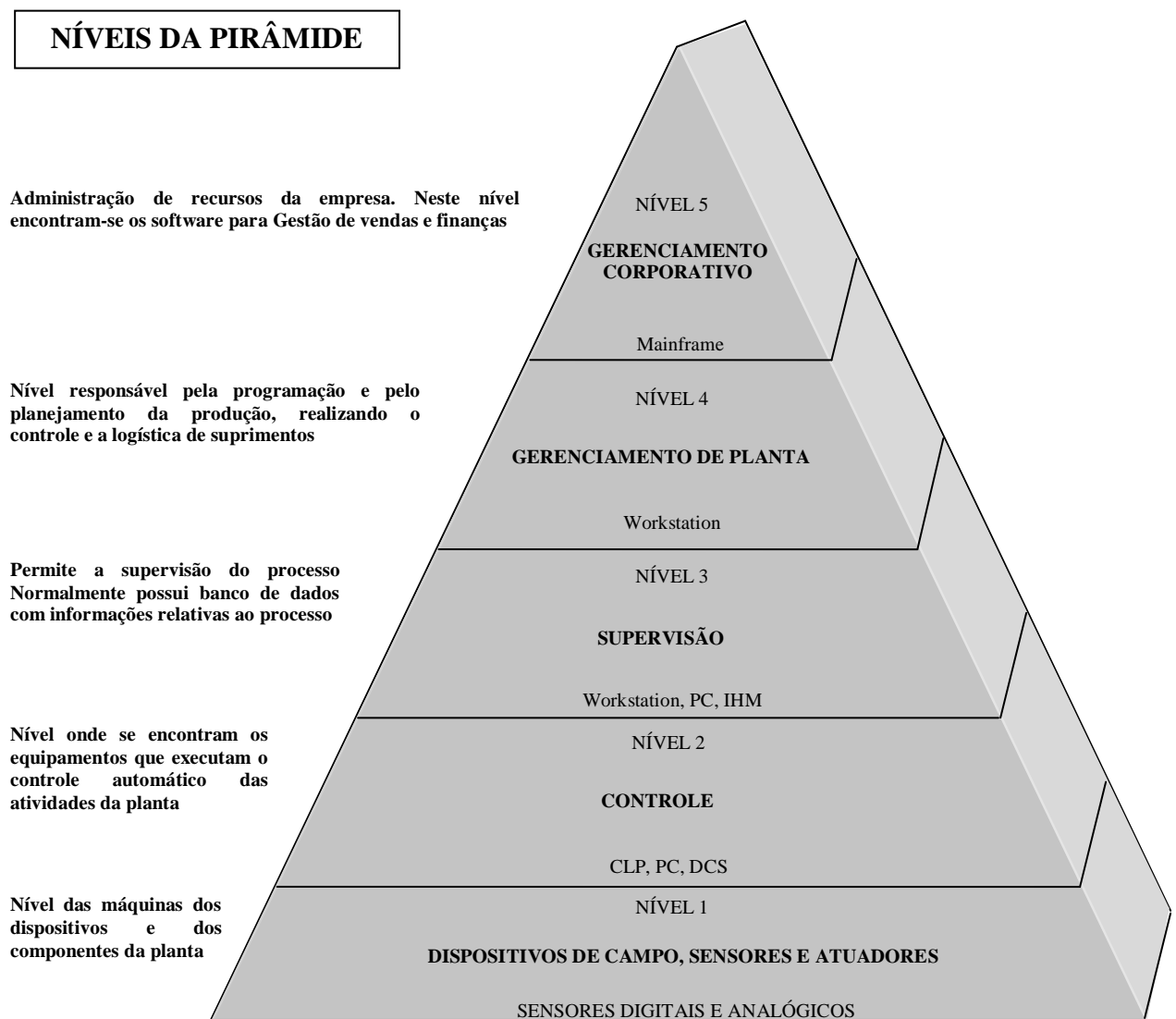


**Figura 2.2** Vista sistêmica da empresa e os estágios de formação da Pelota de Minério de Ferro

## 2.2 Monitoramento e controle de planta Industrial

A Automação Industrial exige a realização de muitas atividades dentro da Planta Industrial. Essas atividades podem ser hierarquizadas de acordo com o grau de importância e o nível de complexidade que suas soluções exigem dos envolvidos na automação dos processos.

Os diferentes níveis de automação encontrados em uma planta industrial formam a Pirâmide da Automação (Figura 2.3) e é análoga, em abstração, a Pirâmide do Conhecimento explanada no capítulo anterior (MORAES & CASTRUCCI, 2007).



**Figura 2.3** Pirâmide de Automação (MORAES & CASTRUCCI, 2007)

De maneira resumida, os níveis podem ser explicados, segundo Moraes, como:

- **Nível Um:** é o nível dos Transdutores, dispositivos e componentes de máquinas. Exemplo: Transdutores de temperatura, pressão e velocidade, acionamentos elétricos.
- **Nível Dois:** é o nível dos controladores digitais, dinâmicos e lógicos, e de algum tipo de supervisão associada ao processo. Aqui se encontram os concentradores de informações do Nível Um, e as Interfaces Homem-Máquina (IHM).
- **Nível Três:** este permite o controle operacional e produtivo da planta, normalmente é constituído por bancos de dados com informações dos índices de qualidade da produção, relatórios e estatísticas de processo, índices de produtividade e algoritmos de otimização da operação produtiva.
- **Nível Quatro:** é o nível responsável pela programação e planejamento da produção, realizando o controle e a logística de suprimentos
- **Nível Cinco:** é o nível responsável pela administração dos recursos da empresa, em que se encontram os softwares de gestão de vendas e gestão financeira e é onde se realizam também a decisão e o gerenciamento de todo o sistema.

Neste contexto, a Automação Industrial exerce um papel fundamental em todos os níveis da organização atuando de maneira direta para obtenção de resultados mais otimizados. Contudo, esses resultados somente alcançarão patamares mais expressivos com a evolução dos sistemas de automação em direção a técnicas estatísticas e da Inteligência Computacional mais elaboradas aplicadas as variáveis de processo.

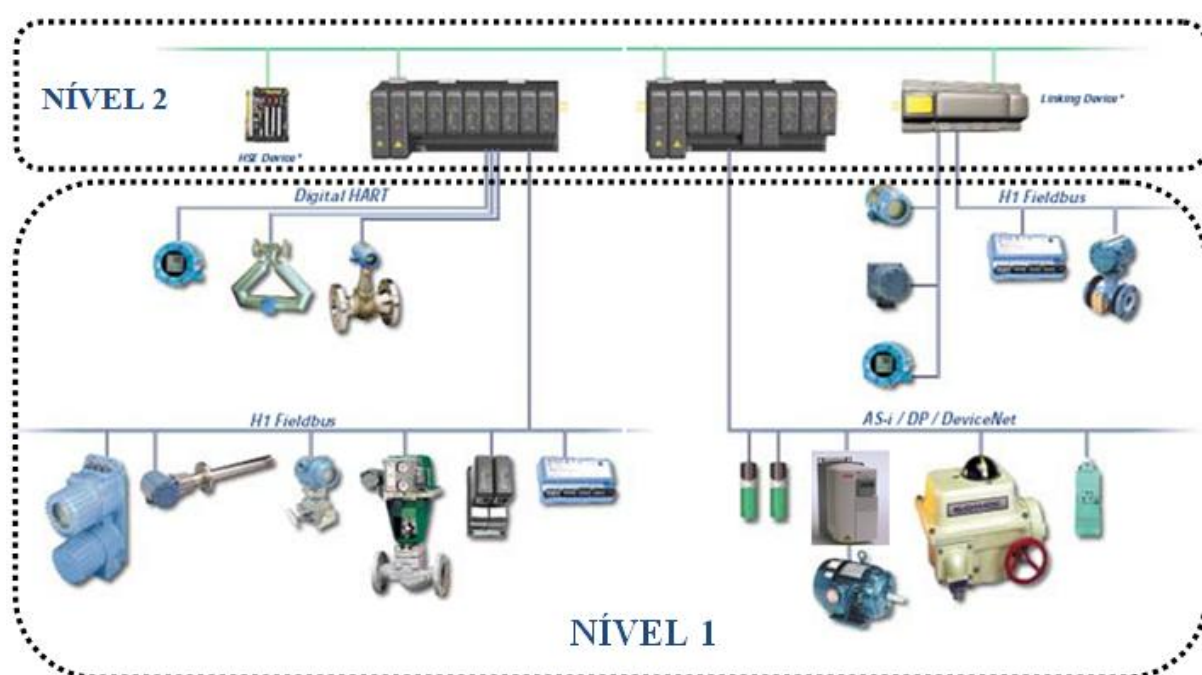
Esse ambiente de pesquisa e desenvolvimento de sistemas de automação mais avançados deve ser implementado no nível três da Pirâmide da Automação e é justamente neste ambiente onde se encontra a aplicação direta dos Modelos de Predição propostos neste trabalho.

### **2.2.1 As variáveis de processo nos diversos níveis**

As variáveis de processo estão presentes em todos os níveis da hierarquia da Pirâmide da Automação e o que as difere é o seu nível de complexidade e conseqüentemente importância na operação da planta Industrial. Esse nível de complexidade é acentuado pelo fato de que as variáveis mais significativas estão fortemente conectadas a conceitos relacionados ao processo de formação da Pelota de Minério de Ferro e seus estágios previamente citados.

Uma enorme gama de variáveis de processo menos significativas e diretamente ligadas aos equipamentos encontram-se localizadas nos Níveis Um e Dois da hierarquia da pirâmide. Essas variáveis de processo, majoritariamente são mensuradas e controladas, por medição direta, através de transdutores e dispositivos digitais (Figura 2.4) e são responsáveis pela operabilidade ininterrupta da planta e seus processos.

Essas variáveis respondem pormenorizadamente sobre algum aspecto, propriedade ou evento ocorrido nos diversos processos e se analisadas de maneira individual, pouquíssima informação agregam em relação ao domínio do negócio como um todo.



**Figura 2.4** Nível Um: Sensores / atuadores digitais e analógicos

Entretanto, nos últimos anos houve um avanço considerável em relação a alguns transdutores na medição de grandezas físico-químicas mais significativas como pH, umidade, vibração e o mais importante, todos em tempo real (BEGA, 2006). Apesar desses avanços, ainda há uma carência substancial em pesquisa e desenvolvimento nesta área a ponto de torná-los comercialmente e tecnicamente mais aplicáveis aos níveis Um e Dois da automação.

O processo de formação da pelota de minério de ferro é um processo contínuo onde todas as variáveis, nos níveis citados, são monitoradas, controladas ou ensaiadas em laboratório. As variáveis monitoradas são supervisionadas periodicamente por análise manual de seus resultados sendo que as ações de correção para eventuais desvios e alterações nos seus resultados esperados são realizadas de maneira corretiva e sem efeito retroativo no produto.



Por outro lado, as variáveis controladas, no nível Dois citado, são ainda aquém da necessidade operacional da planta por decorrência de transdutores indisponíveis para medição e controle de grandezas físico-químicas realmente significativas para o processo.

Por fim, as variáveis ensaiadas são mensuradas em laboratório por testes baseados em amostras colhidas periodicamente e sistemicamente do processo. Essas variáveis são mais complexas, comumente associadas à qualidade do produto e; conseqüentemente, fornecem informações mais qualitativas em relação ao domínio do negócio. Na totalidade, essas variáveis são apenas monitoradas, pois apresentam uma histerese entre seus resultados e sua análise, porque os testes de laboratório demandam um tempo significativo na sua realização.

Assim como as variáveis monitoradas, as variáveis ensaiadas em laboratório são também usadas no controle do processo; entretanto, de maneira manual, empírica e corretiva. Resultados indesejáveis dessas variáveis provocam ações tomadas no Nível Três alterando os ajustes estabelecidos para os níveis Um e Dois. Além disso, esses ajustes pré-estabelecidos não garantem necessariamente os resultados mais otimizados de produtividade e qualidade do produto almejado.

Diante de tudo isso, a idéia de Modelos de Predição que estimem os valores das variáveis monitoradas e principalmente, as ensaiadas em laboratório pode prover um monitoramento otimizado do processo e seus pré-ajustes estabelecidos aos níveis inferiores. Esses modelos são hierarquicamente desenvolvidos no nível Três da Pirâmide da Automação e representam efetivamente conhecimento e inteligência alinhados ao domínio do negócio.

## **2.3 Soluções em Automação para Modelos de Predição**

Segundo Fiske, os trabalhos envolvendo Controle Avançado de Processo (APC) datam da década de 60 e na verdade representavam quaisquer alterações propostas na trilogia – Proporcional-Integral-Derivada (PID) – clássica de controle (FISKE, 2006). Hoje, esses métodos incorporam uma série de técnicas envolvendo supervisão, inferência, multivariáveis, adaptativas, não-lineares e de predição; e ainda, controle *Fuzzy*, sistemas especialistas, redes neurais, estatísticas e modelos robustos; enfim, todos visando otimizações em tempo real.



**Figura 2.5** Incremento de valor com o avanço no controle de processo (FISKE, 2006)

Categorizando os métodos mencionados por estratégias de controle e níveis de utilização nas indústrias podemos descrevê-las como (SEBORG, 2000):

- **CATEGORIA I:** Estratégias de controle convencional

Exemplos:

- Controle manual
- Controle PID
- Controle Cascata

▪ **CATEGORIA II: Controle Avançado:** Estratégias clássicas

Exemplos:

- Compensação de tempo morto (do inglês: *Time Delay Compensation*)
- Desacoplamento de malhas (do inglês: *Decoupling Control*)
- Controladores Seletivos (do inglês: *Selective Controllers*)

▪ **CATEGORIA III: Controle Avançado:** Estratégia com larga aplicação industrial.

Exemplos:

- Controle Preditivo baseado em modelo (do inglês: *Model Predictive Control*)
- Controle Estatístico de Qualidade (do inglês: *Statistical Quality Control*)
- Controle Adaptativo (do inglês: *Adaptive Control*)

▪ **CATEGORIA IV: Controle Avançado:** Técnicas recentes com algumas aplicações industriais.

Exemplos:

- Controle Não-linear (do inglês: *Nonlinear Control*)
- Controladores baseados em Redes Neurais (do inglês: *Neural Network Controllers*)
- Controle *Fuzzy* (do inglês: *Fuzzy Control*)
- Sistemas Especialistas (do inglês: *Expert Systems*)

▪ **CATEGORIA V: Controle Avançado:** Estratégias propostas com rara (ou nenhuma) aplicação na indústria.

Explanando melhor a Categoria V, constituem métodos relacionados a Pesquisas Formais conceituadas no capítulo anterior. Contudo, existem uma série significativa dessas pesquisas, com nenhuma ou rara, aplicação implementada num ambiente industrial e formam até um importante potencial na resolução de problemas sem solução ou com resultados de otimização insatisfatórios relacionados aos processos.

O objetivo desse trabalho é investigar estratégias que estão classificadas entre as Categorias IV e V supracitadas na obtenção de Modelos de Predição. Esses modelos são protótipos conceituais (*Frameworks*), programáveis nos níveis Dois e Três da Pirâmide da Automação.

Neste contexto, surgem as técnicas de Regressão e Neuro-*Fuzzy* com significativo potencial utilizadas neste trabalho. Como pode ser observado pela descrição das categorias,

uma gama significativa de técnicas da Inteligência Computacional poderia ser testada na investigação de soluções de problema com resultados ainda insatisfatórios de performance.

### **2.3.1 Trabalhos correlatos**

A tecnologia predominante na área de Controle Avançado de Processo é o Controle Preditivo baseado em modelo (MPC) tanto que, os termos são utilizados de formas alternadas em pesquisas com o mesmo significado (FISKE, 2006). Segundo Seborg, o termo APC é subjetivo e designa abordagens diferentes de acordo com a formação dos envolvidos na pesquisa por decorrência de sua característica interdisciplinar.

Num relatório do grupo *ARC Consulting – ARC Advisory Group* – para o site *AutomationWorld* de maio de 2006 estavam cadastrados em seus registros mais de 10.000 solicitações em consultoria de MPC. Os trabalhos correlatos envolvem alternadamente esses conceitos – APC e MPC – em soluções híbridas mencionadas nas categorias acima e diversos estudos de caso.

Esses trabalhos correlatos justificam a implementação de um ambiente de engenharia de pesquisa e gestão do conhecimento dentro da hierarquia da Pirâmide da Automação e apontam para práticas modernas de controle que devem ser buscadas como evolução natural dos sistemas automatizados. Com isso, a pesquisa de técnicas da Inteligência Computacional alinhadas com técnicas estatísticas tornam-se ferramentas fundamentais neste avançado sistema de controle de processo.

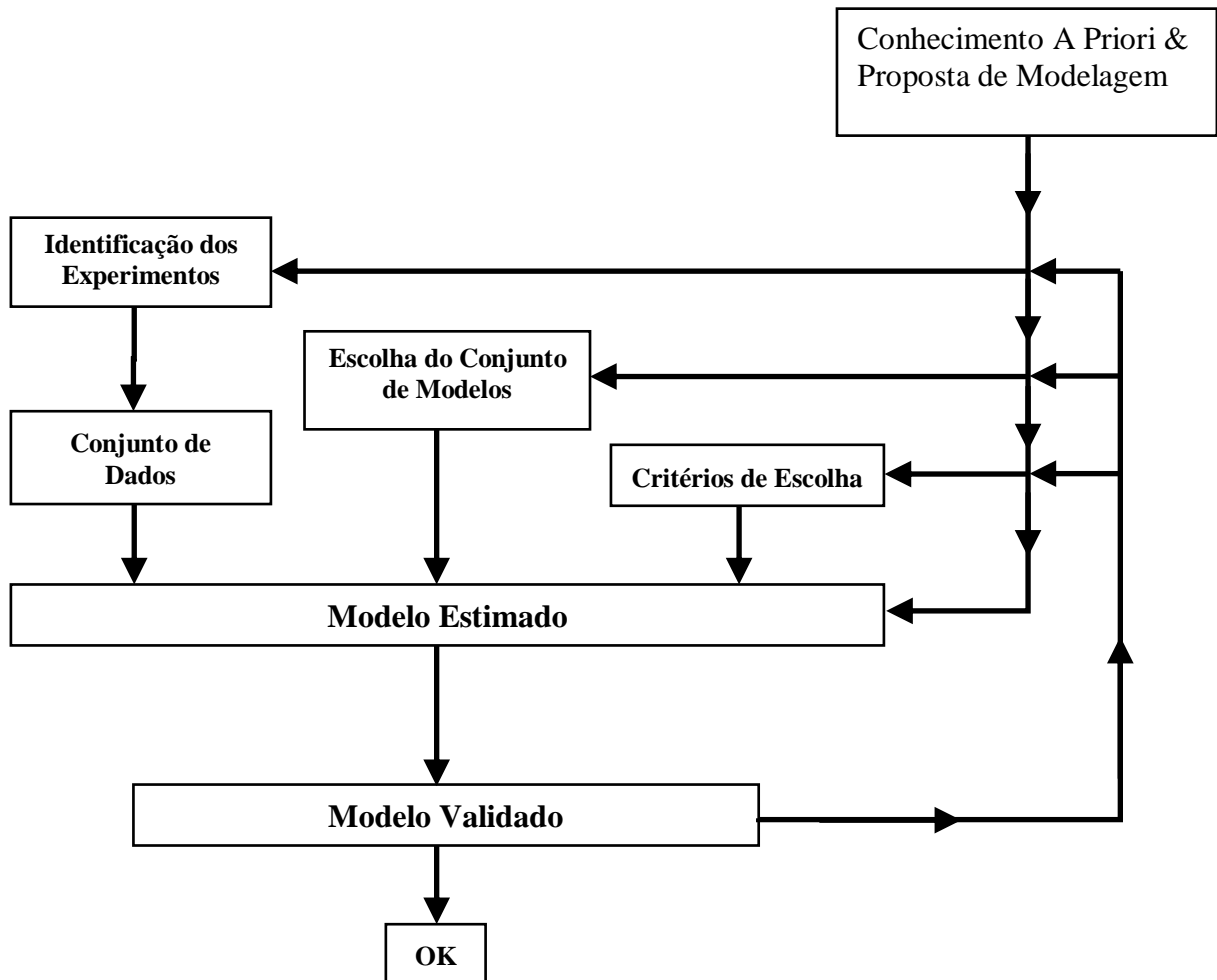
Essa evolução dos sistemas de automação é pertinente com a demanda cada vez mais acirrada por produtos competitivos e com qualidade atestada e ratificam os Modelos de Predição com uma das grandes potenciais ferramentas emergentes deste inovador ambiente de monitoramento e controle de plantas industriais.

## Capítulo 3: Técnicas de construção de Modelos de Predição

### 3.1 Introdução

A idéia de Modelos de Predição é buscar estratégias da Inteligência Computacional que possam resultar justamente em modelos que predizem o comportamento de uma variável de processo por meio de estatísticas. Contudo, essas estratégias formam uma gama de soluções, a ponto de ser categorizadas, conforme citado anteriormente, pelo seu nível de complexidade e aplicabilidade num ambiente industrial real.

Segundo Zhu, um modelo é uma representação dos aspectos essenciais de um sistema (processo), que apresenta conhecimento em uma forma utilizável (ZHU, 2001). A Figura 3.1, identifica os passos que devem ser adotados na identificação de um modelo.



**Figura 3.1** Procedimento de identificação de Modelo de Predição (ZHU, 2001)

Neste trabalho são adotados os procedimentos de identificação de Modelos de Predição como descritos onde podemos destacar três principais passos a seguir:

1. Formação do Conjunto de Dados: Consiste na investigação do conhecimento a priori, nas propostas de modelagem, na identificação dos experimentos envolvidos nas variáveis para estabelecer a base de dados utilizada na elaboração dos Modelos de Predição.
2. Escolha do Conjunto de Modelos: Envolve a pesquisa em relação aos métodos de solução conforme os categorizados anteriormente. Neste ponto de pesquisa, surge as técnicas da Inteligência Computacional como ferramentas para construção de Modelos de Predição.
3. Critérios de Escolha: envolve a pesquisa de ferramentas estatísticas que possam analisar o nível de adequação e generalização dos modelos construídos. Neste passo também são estabelecidas uma série de decisões heurísticas pertinentes as pesquisas de caráter empírico.

A etapa de escolha do Conjunto de Modelos passa pela análise do conhecimento a priori e principalmente em encontrar o relacionamento entre um conjunto de variáveis, quando pelo menos uma é aleatória e, eventualmente, sujeitas a erros de medição (SEBER & WILD, 1996).

Tipicamente, em problemas de regressão pelo menos uma variável deste processo apresenta esse comportamento, sendo estão denominadas de variável dependente ou variável resposta sendo denotada pela letra  $y$ . As outras variáveis  $X$ :  $x_1, x_2, \dots, x_k$  geralmente são chamadas de variáveis independentes e são primordialmente usadas para prever ou explicar o comportamento de  $y$ . A idéia do modelo é estabelecer uma relação entre  $y$  e os  $X$  através de alguma função  $f$ , simbolizada por:

$$y \approx f(x_1, x_2, \dots, x_k).$$

As predições justamente podem ser feitas através da utilização da função  $f$  de relacionamento aplicada a um Conjunto de Dados ( $X$ ). Essa é uma etapa do procedimento de identificação do modelo responsável pela estimativa das constantes ou coeficientes (Parâmetros) desconhecidos da função  $f$ . A função  $f$  pode ser redefinida matematicamente como:

$$y \approx f(x_1, x_2, \dots, x_k; \beta).$$

O objetivo é justamente estabelecer o modelo estimado através da determinação de seus coeficientes ( $\beta$ :  $\beta_1, \beta_2, \dots, \beta_k$ ) desconhecidos do sistema. Entretanto, uma importante especificação da função  $f$  é determinar se esse modelo é linear ou não-linear. Em geral, qualquer modelo de regressão que seja linear nos parâmetros  $\beta$  é um modelo de regressão linear, independentemente da forma da superfície que ele gere (MONTGOMERY & RUNGER, 2003).

Uma outra abordagem estabelece que em processos contínuos industriais, ou seja, aqueles que operam sob uma forma controlada por longos períodos de tempo e seus ajustes são mantidos dentro de uma zona limite de controle podem ser efetivamente controlados por modelos de processos lineares (ZHU, 2001). Isso ocorre porque qualquer função não-linear pode ser bem aproximada por uma função linear em torno de um equilíbrio, o que explica o sucesso do modelo linear baseado em tecnologias de controle para os processos contínuos.

Essa situação não se mantém em momentos de inicialização ou desligamento de equipamentos da planta industrial onde os comportamentos dos processos tornam-se predominantemente não-lineares, reduzindo significativamente a eficiência dos modelos lineares. Neste trabalho é utilizado Modelos de Predição lineares.

Num ambiente da Automação Industrial moderno, a estrutura de evolução de controle é hierárquica e também análoga em abstração aos conceitos relacionados à Pirâmide do Conhecimento podendo ser descritos em níveis (ZHU, 2001):

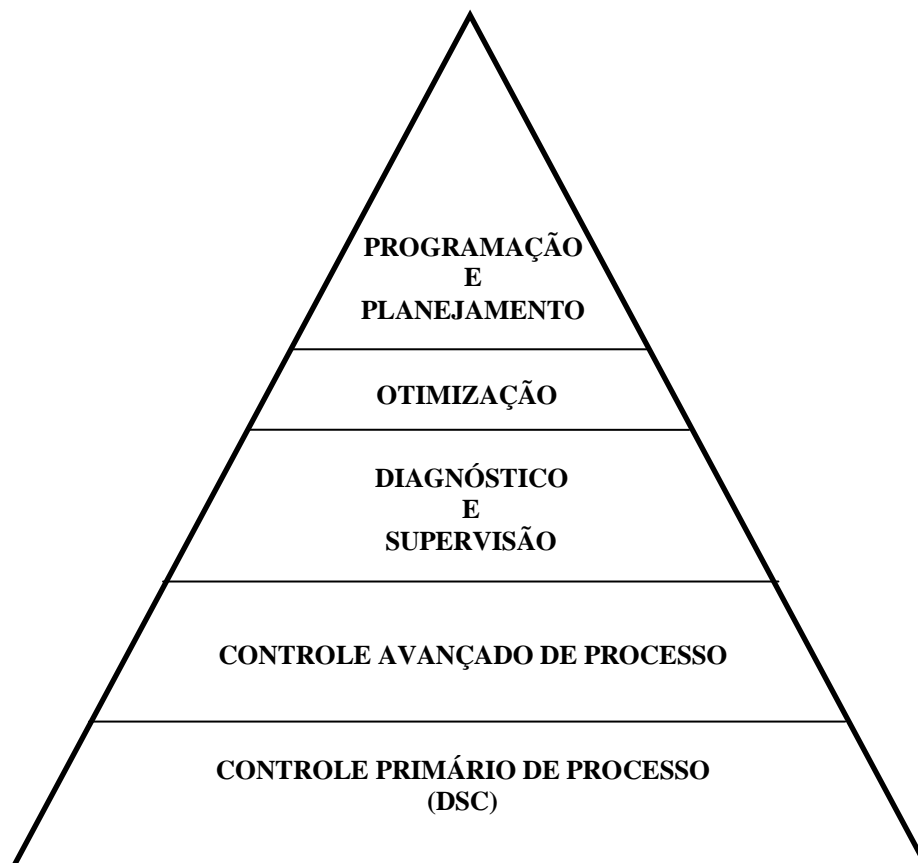
- **Instrumentação e controle primário:** Este é geralmente um Sistema de Controle Distribuído (DCS), que reúne as medições do processo e executa o simples monitoramento das mesmas. As medidas incluem as variáveis básicas de processo, tais como temperatura, vazão, pressão, nível e posição de válvulas. O DCS também realiza controles PID baseados em algumas das variáveis de processo. Normalmente apenas um controlador cuida de uma única variável primária, como o fluxo ou a temperatura.
- **Controle Avançado de processo:** Esta parte do sistema realiza o controle baseado em modelo multivariável que vai garantir a operação da unidade estável e conduzem o processo para os seus limites operacionais objetivando o máximo benefício

econômico. Neste nível, um controlador APC pode controlar todo um processo, como uma coluna de destilação, um reator. Em geral, a identificação de modelos dinâmicos (na maioria das vezes linear) é usada para os controladores APC. Esta camada é normalmente presente em um sistema de computadores.

- **Diagnóstico e Supervisão:** Esta parte do sistema é responsável por melhorar a segurança e a confiabilidade de operação da unidade (Planta). Um sistema de diagnóstico realiza a detecção e classificação das faltas e dá sugestões para a manutenção e soluções das mesmas. Os métodos mais simples baseiam-se principalmente na verificação do valor limite dos sinais e suas propriedades facilmente mensuráveis. Uma tendência recente é a utilização de modelos de processo mais precisos e de mais rápido diagnóstico. O sistema também pode avaliar o desempenho dos controladores em diferentes níveis. Esta camada é normalmente presente em um sistema de computadores.
- **Otimização:** Um sistema de otimização de processo manipula os graus de liberdade, que são ajustados após a obtenção dos requisitos de uma operação segura e estável, para atender objetivos econômicos das unidades, como a limitação de energia elétrica (exemplo: Demanda (kWh/TON)), insumos com alto custo operacional (exemplo: óleo combustível (LT/TON)) e / ou aumentar a taxa de transferência. Além disso, o sistema determina os melhores pontos de ajuste para os controladores APC. Normalmente a otimização é realizada com base nos rigorosos modelos estáticos não-lineares, obtidos anteriormente sob condição instável do processo, sendo utilizados como referência. Às vezes, esses modelos também são usados para otimização, pois o custo de utilização e manutenção desses sistemas pode ser muito alto. Normalmente, o sistema é executado em um ritmo lento, de modo que os controladores APC estejam em uma situação estável no que diz respeito à alteração anterior dos ajustes de processo (*set-point*). A otimização pode ser realizada em um único processo, bem como numa combinação deles. Um sistema otimizador deste porte geralmente está localizado em um sistema de computadores.
- **Planejamento e Escalonamento:** Esta parte pode abranger muitas unidades de processos e fornece apoio à decisão no planejamento da produção, alocação de matérias-primas e agendamento de operação da planta para a realização do programa



de metas da empresa e para a maximização do lucro. É utilizado para responder às mudanças do mercado, bem como alterações de regulação da produção. Um sistema deste porte geralmente está localizado em um sistema de computadores.



**Figura 3.2** Hierarquia dos Sistemas de Controle da Automação Industrial (ZHU, 2001)

Como podemos notar da hierarquia citada, os sistemas de controle modernos perfazem uma atuação sistêmica em relação à empresa e todos os seus processos, localizando assim a Automação Industrial numa posição estratégica deste integrador processo interdisciplinar. Os conceitos descritos até o momento serviram de base para nos tornamos aptos a situar os Modelos de Predição no processo industrial e em sua hierarquia.

### **3.2 Técnicas de Regressão clássicas e avançadas**

De maneira generalizada, as técnicas da Inteligência Computacional para resolução de problemas desta natureza envolvem a análise multivariada de dados. Esse fato é completamente abordado pela Análise de Regressão Múltipla (*Multiple Regression Analysis*)

que é uma técnica estatística que pode ser usada para analisar a relação entre uma única variável dependente e as muitas variáveis independentes (HAIR, 2009).

O objetivo da Análise de Regressão Múltipla é utilizar as variáveis independentes, cujo valor é conhecido, na predição de uma única variável dependente selecionada como objeto de investigação e pesquisa. A variável dependente estimada é simbolizada pela letra  $\hat{y}$ .

$$\hat{y} = E [\beta; X]$$

De maneira geral, a equação da Análise de Regressão Linear Múltipla (SEBER & WILD, 1989) é esboçada abaixo e inclui também a adição do erro ( $\varepsilon$ ); sendo que, a estimativa do erro é nula ( $E [\varepsilon] = 0$ ). O modelo de regressão é descrito como:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_{p-1} x_{k-1} + \varepsilon$$

Entretanto, reforçando os conceitos de modelos lineares e não-lineares temos que ressaltar que  $\beta$ :  $\beta_0, \beta_1, \dots, \beta_k$  e os  $x_i$ 's podem incluir termos quadráticos, multiplicação cruzada dos termos (*cross products*), polinômios e até transformações (logarítmicas) das medições originais ( $X$ ). Contudo, o requisito primordial é que a expressão deverá ser linear nos parâmetros  $\beta$  para confirmar a construção de modelos lineares. Por exemplo:

$$y \approx \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2$$

ou

$$y \approx \beta_0 + \beta_1 \text{sen} x_1 + \beta_2 \text{sen} x_2$$

são modelos de regressão linear. Entretanto, o modelo abaixo exemplifica um caso não-linear sendo-o no parâmetro  $\beta_2$ .

$$y \approx \beta_0 + \beta_1 e^{\beta_2 x}$$

Os Modelos de Predição descritos neste trabalho são modelos lineares conforme os conceitos descritos anteriormente. O conjunto de técnicas mostradas nos próximos subtítulos foi heurísticamente escolhido baseando-se nas propriedades físico-químicas da natureza da

variável dependente e as possíveis relações  $f$  que podem se estabelecer entre essa variável e as variáveis independentes envolvidas neste processo.

Complementarmente, a análise do comportamento da variável dependente a ser estimada mostra indícios da relação  $f$  que deve se estabelecer entre as variáveis independentes. Neste trabalho, como citado anteriormente, a variável dependente estimada é uma variável de qualidade do produto no processo de Pelotização e é detalhada no Capítulo 5.

### 3.2.1 Modelo de Predição clássico por Regressão Linear Múltipla

O Modelo de Predição resultante da Análise de Regressão Linear Múltipla é uma relação  $f$  estabelecida entre a variável dependente e as variáveis independentes a partir dos coeficientes de regressão  $\beta$  que podem ser estimados pelo método dos mínimos quadrados (MONTGOMERY, 2001). Esse método consiste em encontrar o valor dos coeficientes que minimizem a função  $L$  dos mínimos quadrados.

$$L = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2$$

Essa função é minimizada pelos coeficientes de regressão estimados ( $\hat{\beta}$ :  $\beta_0, \beta_1, \dots, \beta_k$ ) obtido partir da equação abaixo:

$$\hat{\beta} = (\mathcal{X}' \mathcal{X})^{-1} \mathcal{X}' \mathbf{y}$$

Sendo que os termos  $\mathcal{X}$  e  $\mathbf{y}$  são explicados pela Tabela 3.1. Os valores de  $k$  e  $n$  representam respectivamente o número de variáveis independentes e a quantidade de amostras do conjunto.

TABELA 3.1: Variável Independente e Dependente

$y$	$X$			
	$x_1$	$x_2$	...	$x_k$
$y_1$	$x_{11}$	$x_{12}$	...	$x_k$
$y_2$	$x_{21}$	$x_{22}$	...	$x_k$
⋮	⋮	⋮		⋮
$y_n$	$x_{n1}$	$x_{n2}$	...	$x_{nk}$

Este Modelo de Predição extraído deste método é a maneira mais clássica de regressão sendo que o plano de saída é uma reta que passa o mais próximo possível dos pontos observados (Amostras:  $n$ ) e são propositalmente utilizadas a fim de confirmar as características da variável dependente (MONTGOMERY & RUNGER, 2003).

Entretanto, as formas assumidas pelos planos da variável dependente a ser estimada nos proporcionam indícios de que este método não resulta em um modelo adequado e significativo para este tipo de variável mais complexa.

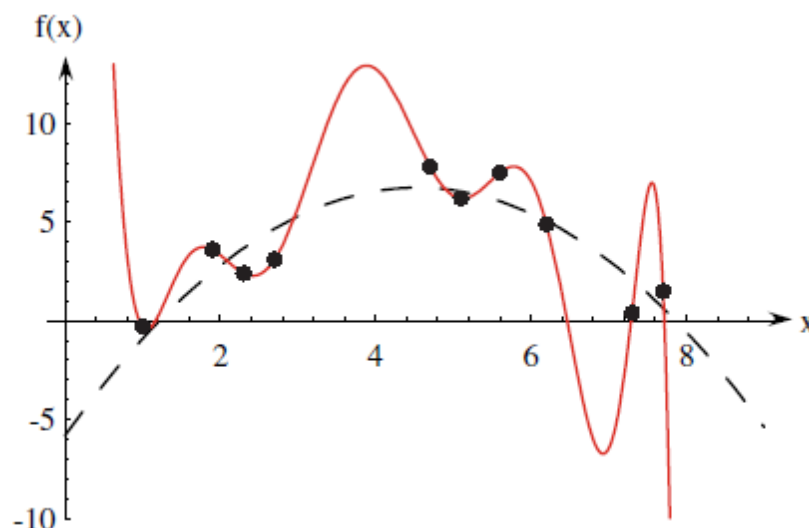
Essa variável complexa representa um parâmetro de qualidade do produto do processo de Pelotização e é a Compressão Mecânica da pelota de Minério de Ferro a frio que possui características curvilíneas (polinomiais) em relação ao plano formado pela mesma (RAWLINGS, 1998).

Contudo, as técnicas de Regressão Polinomial que poderia ser utilizada em decorrência destas características não podem ser diretamente aplicadas devido à (WEISBERG, 2005):

- ✓ Colinearidade – problemas de alta ordem (grau) das variáveis independentes, pois os termos de  $x^2$ ,  $x^3$ ,  $x^4$  e suas interações; assim por diante, tornam-se colineares, perdendo assim a sua característica de variáveis "independentes";
- ✓ O número de termos, para exemplificar, da função  $f$  quadrática cresce significativamente a uma lógica de: um termo intercessor,  $k$  termos lineares,  $k$  termos

quadráticos e  $\frac{k(k+1)}{2}$  termos de interações (Exemplo:  $k = 5$  resulta em uma função  $f$  quadrática de 26 termos,  $k = 10$  implica em 76 termos e assim sucessivamente);

- ✓ Essas funções  $f$  de mais alta ordem passam a representar a variação ajustada (*overfitting*) dos dados ( $n$ : amostras) do que perfazer uma representação generalizada e não sobre ajustada dos mesmos conforme Figura 3.3.



**Figura 3.3** Exemplo de *Overfitting* e Generalização em um problema de Regressão. O "Conjunto de Treinamento" (Pontos Pretos) foram artificialmente gerado a partir de uma função quadrática (Linha pontilhada) somada a um "ruído Gaussiano"; ou seja,  $f(\mathcal{X}) = ax^2 + bx + c + \varepsilon$ , onde  $p(\varepsilon) \sim N(0, \sigma^2)$ . Outra função Polinomial  $f(\mathcal{X})$  de 10º ordem (Linha vermelha) mostra essa função perfeitamente ajustada às amostras (DUDA, 2001).

Esses problemas citados com as Regressões Polinomiais podem ser contornados pela utilização de técnicas que passam a transformá-las em Regressões Ortogonais (WEISBERG, 2005). Contudo, sua utilização caracterizaria uma série de estratégias para adaptar os Modelos de Predição obtidos a partir de técnicas polinomiais o que; a princípio, foge do escopo deste trabalho.

Diante dos fatos expostos, técnicas da Inteligência Computacional mais moderna e avançada que possam contemplar essa característica intrínseca da variável dependente e, além disso, amenizar ou até mesmo contornar os problemas mencionados torna-se necessárias.

### 3.2.3 Máquina de *Kernel* Esparso

Os conceitos utilizados referentes à Máquina de *Kernel* Esparso (*Sparse Kernel Machine*) são extraídos da Teoria da Aprendizagem Computacional (*Computational Learning Theory*) (BISHOP, 2007). Segundo Bishop, essa teoria também denominada de Teoria de Aprendizagem Estatística (*Statistical Learning Theory*) foi inicialmente formulada pela teoria PAC (*Probably Approximately Correct*) na qual a idéia principal é compreender como um largo conjunto de dados precisa estar ordenado (relacionado) a fim de se obter uma boa generalização.

Neste contexto, existe uma classe de técnicas baseada em Reconhecimento de Padrões em que os pontos de dados treinados ou um subconjunto dos mesmos são mantidos e utilizados na fase de predição; ou seja, fazem parte do próprio Modelo de Predição.

Segundo Bishop, vários modelos lineares paramétricos (em função dos  $\beta$ ) podem ser remodelados em uma análoga representação em que as variáveis independentes  $\mathcal{X}$  são também baseadas na combinação linear da função *Kernel* evoluída a partir dos pontos de dados treinados.

Com isso, para modelos em que o espaço de características  $\phi(\mathcal{X})$  é não-linear, a função de *Kernel* é dada pela relação:

$$\mathcal{K}(\mathcal{X}, \mathcal{X}') = \phi(\mathcal{X})^T \phi(\mathcal{X}')$$

Desta formulação, podemos destacar que o conceito de *Kernel* é uma função simétrica de seus argumentos tal que  $\mathcal{K}(\mathcal{X}, \mathcal{X}') = \mathcal{K}(\mathcal{X}', \mathcal{X})$ . O objetivo destes Modelos de Predição é explorar algoritmos baseados em *Kernels* não-lineares de forma que predições para as novas amostras dependam somente da função de *Kernel* estabelecida e evoluída a partir do subconjunto dos dados treinados.

Esses algoritmos baseados em *Kernel* possuem propriedades de soluções "esparsas" (*Sparse*) em decorrência do espaço não-linear de características  $\phi(\mathcal{X})$ . Uma solução desta em especial é contemplada pela teoria da Máquina de Vetor de Suporte (*Support Vector Machine* – SVM).

Uma propriedade particularmente importante da SVM é que a determinação dos parâmetros do modelo corresponde a um problema de otimização convexo; implicando que,

qualquer solução local é também uma solução ótima global. Neste trabalho, a variável dependente a ser estimada assume valores contínuos o que é contemplado pela SVM para regressão.

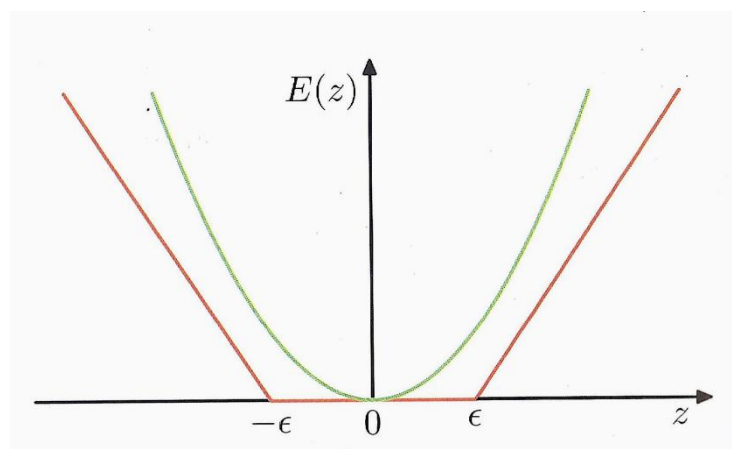
### 3. 2.3.1 Regressão por Máquina de Vetor de Suporte

Os conceitos mencionados sobre Máquina *Kernel* Esparsa aplicados a Máquina de Vetor de Suporte são estendidos aos problemas de regressão herdando suas propriedades inclusive a esparsa (BISHOP, 2007). Segundo Bishop, em problemas envolvendo regressão linear simples, é minimizada a função de erro dada por:

$$\frac{1}{2} \sum_{n=1}^N \{y_n - t_n\}^2 + \frac{\lambda}{2} \|w\|^2$$

Contudo, para garantir a propriedade de soluções esparsas, a função de erro quadrática é substituída por uma função de erro  $\epsilon$ -insensitive que resulta num valor nulo se a diferença absoluta entre o valor predito de  $y(x)$  e o valor real da amostra  $t$  seja menor do que  $\epsilon$ , com  $\epsilon > 0$ . Um exemplo de função de erro  $\epsilon$ -insensitive (Figura 3.4) é dada pela função:

$$E_{\epsilon}(y(x) - t) = \begin{cases} 0, & \text{se } |y(x) - t| < \epsilon; \\ |y(x) - t| - \epsilon, & \text{caso contrário} \end{cases}$$



**Figura 3.4** Exemplo de uma função de erro  $\epsilon$ -insensitive (Linha vermelha) em que o erro incrementa linearmente com uma distância além da região "insensível" comparada a função de erro quadrática (Linha verde) (BISHOP, 2007).

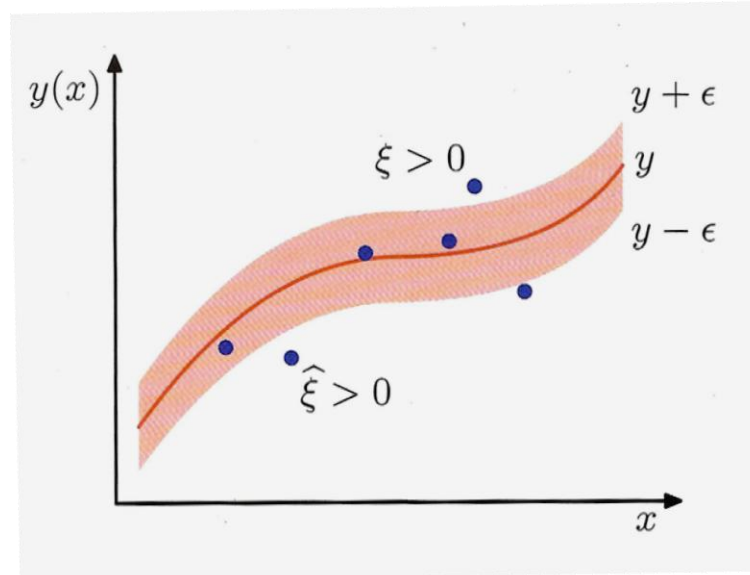
Contudo, a SVM para regressão minimiza a função de erro dada por:

$$C \sum_{n=1}^N E_{\varepsilon}(y(x_n) - t_n) + \frac{1}{2} \|w\|^2$$

Onde:

- $y(\mathcal{X}) = w^T \phi(\mathcal{X}) + b$ ;
- $C$  é o parâmetro de regulação (Inverso), que por convenção, aparece na frente à função de erro;
- $w$  é o vetor dos pesos.

Contudo, a função de erro  $\varepsilon$ -insensitive é reescrita como uma problema de otimização com a introdução da variável *Slack*. Para cada ponto  $x_n$ , é necessário duas variáveis *Slack*  $\xi_n \geq 0$  e  $\hat{\xi}_n \geq 0$ , onde  $\xi_n > 0$  correspondem aos pontos em que  $t_n > y(x_n) + \varepsilon$ , e  $\hat{\xi}_n > 0$  correspondem aos pontos em que  $t_n < y(x_n) - \varepsilon$  conforme Figura 3.5.



**Figura 3.5** Ilustração da SVM para Regressão, mostrando a curva de regressão (Linha vermelha) junto com o "tubo" da função de erro insensível  $\varepsilon$ -insensitive (Região sombreada –  $\varepsilon$ -tube). É também exibido as Variáveis *Slack*  $\xi$  e  $\hat{\xi}$ . Pontos acima do  $\varepsilon$ -tube têm  $\xi > 0$  e  $\hat{\xi} = 0$ , pontos abaixo do  $\varepsilon$ -tube têm  $\xi = 0$  e  $\hat{\xi} > 0$ , e pontos dentro do  $\varepsilon$ -tube têm  $\xi = \hat{\xi} = 0$  (BISHOP, 2007).

Nestas condições, a região sombreada da figura é denominada de  $\varepsilon$ -tube e englobam os pontos da amostra da Variável Dependente  $t_n$  sujeito a condição  $y_n - \varepsilon \leq t_n \leq y_n + \varepsilon$ ,

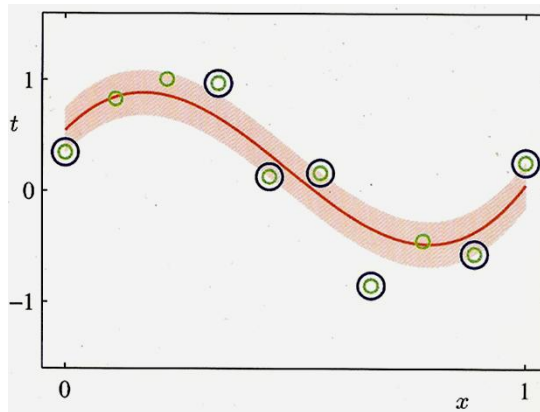


onde  $y_n = y(X_n)$ . Com isso, a função de erro  $\varepsilon$ -insensitive para SVM para Regressão pode ser reescrita como:

$$C \sum_{n=1}^N (\xi_n + \hat{\xi}_n) + \frac{1}{2} \|w\|^2$$

Entretanto, na prática, ao invés de ajustar simplesmente o erro  $\varepsilon$  do  $\varepsilon$ -tube é também ajustado um parâmetro  $\nu$  de modo que envolva também uma fração de pontos que estão fora do  $\varepsilon$ -tube.

Essa situação é ilustrada na Figura 3.6 do exemplo utilizando um conjunto de dados senoidal, onde pode ser visto que, muitos dos pontos de dados  $\nu N$  caem fora da região do  $\varepsilon$ -tube e os vetores de suporte selecionados são um conjunto de pontos formados pelos pontos localizados dentro e fora da região insensível. Na prática, os parâmetros  $\nu$  e  $C$  são tipicamente determinados através do método de *validação cruzada* dos dados de amostra.



**Figura 3.6** Ilustração da  $\nu$ -SVM para Regressão, aplicada a um conjunto sintético de dados senoidal usando *Kernels* Gaussiano. A curva de regressão predita é representada pela Linha vermelha e o  $\varepsilon$ -tube é representado pela região sombreada. Adicionalmente, os pontos de dados (amostras) são exibidos em verde e os vetores de suporte são destacados pelos círculos azuis (BISHOP, 2007).

O uso da técnica de Máquina de Vetor de Suporte para problemas de regressão é determinado pelo ajuste dos parâmetros  $\varepsilon$ ,  $\nu$  e  $C$  por estratégias aplicadas ao conjunto de treinamento ampliando a sua complexidade de utilização na prática. Com isso, métodos que possam aprimorar as SVM são propostos num intuito de reduzir tal complexidade.

### 3. 2.3.2 Regressão por Máquina de Vetor de Relevância

A Máquina de Vetor de Relevância (*Relevance Vector Machine*) para regressão é um modelo de Regressão linear, conforme conceituado anteriormente, mas com uma modificação a priori que resulta justamente em soluções "esparsas" (TIPPING, 2000). O Modelo de Predição preestabelece uma distribuição condicional dos dados representados pelos valores reais assumidos pela variável dependente  $t$ , dada pela equação:

$$p(t|x, w, \beta) = \mathcal{N}(t|y(x), \beta^{-1})$$

onde:

- A precisão do ruído  $\beta = \sigma^{-2}$ ; sendo este, o inverso da variância;
- O modelo linear  $y(x) = \sum_{n=1}^M w_n \phi_n(x) = W^T \phi(x)$ ;
- São funções de base não-lineares:  $\phi_n(x)$  que incluem um termo constante "bias"  $b$ ;

Complementarmente, a Máquina de Vetor de Relevância ou RVM (TIPPING, 2001) é uma técnica Bayesiana de *Kernel* esparso para regressão que herdaram muitas das características da SVM evitando as suas principais limitações (TIPPING & BISHOP, 2003). Além disso, possui a propriedade de soluções esparsas, proporcionando um desempenho mais rápido sob uma mesma base de teste e mantendo o poder de generalização dos dados comparável a SVM.

Generalizando, um conjunto de dados com  $n$  observações (amostras) possuindo a mesma matriz de variáveis independentes  $X$ ; entretanto, a variável dependente é expressa da forma  $t = (t_1, \dots, t_n)^T$ , possui a função densidade expressa por:

$$p(t|X, w, \beta) = \prod_{n=1}^N p(t_n | x_n, w, \beta^{-1})$$

Nesta expressão é introduzida uma distribuição a priori sobre o vetor de parâmetros  $w$  sendo uma função gaussiana com média zero (*zero-mean Gaussian prior distribution*). Entretanto, uma diferença estratégica nas RVM é que possuem hiper-parâmetros  $\alpha_i$  separadores para cada

parâmetro de peso  $w_i$  ao invés de um único hiper-parâmetro compartilhado no caso da SVM. Com isso, a função de densidade a priori tem a forma:

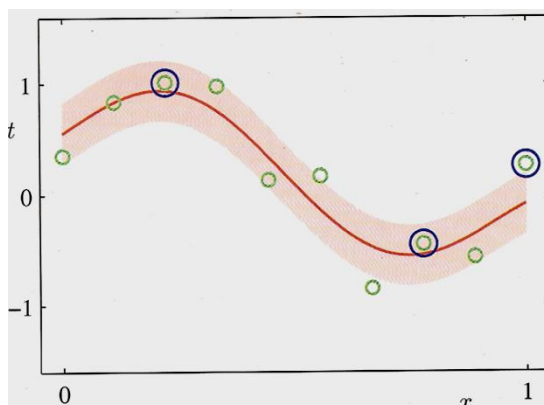
$$p(w|\alpha) = \prod_{i=1}^M \mathcal{N}(w_i | 0, \alpha_i^{-1})$$

onde:

- $\alpha = (\alpha_1, \dots, \alpha_M)^T$ ,
- $\alpha_i$ : representa a precisão do parâmetro correspondente a  $w_i$ ,
- $M = N + 1$  indica o número de parâmetros.

Segundo Bishop, quando a evidência destes respectivos hiper-parâmetros é maximizada, uma porção significativa dos mesmos tende a infinito e os seus correlacionados parâmetros de pesos tendem a zero (BISHOP, 2007). As funções de base associadas com esses parâmetros são parte do Modelo de Predição com exceção dos parâmetros eliminados pela maximização dos mesmos, efetivando assim a sua propriedade de modelo com soluções esparsas.

Com o processo de otimização (Maximização) dos parâmetros  $\alpha_i$  e  $w_i$  são obtidos largos valores de hiper-parâmetros e seus correspondentes pesos tendo distribuição a *posteriori* com média e variância nula. Os parâmetros infinitos e nulos resultante da maximização são eliminados do Modelo de Predição bem como suas respectivas funções de base  $\phi_i(x)$  associadas aos mesmos. Deste processo, os valores de entrada  $x_n$  correspondentes com os parâmetros de pesos remanescentes não nulos são denominados de Vetores de Relevância ou Relevantes (*Relevance Vectors*) e são análogos aos Vetores de Suporte (*Support Vectors*) das SVM.



**Figura 3.7** O mesmo ambiente de teste da v-SVM para Regressão aplicada ao RVM para regressão. O valor médio da distribuição do Modelo de Predição é esboçado pela linha vermelha seguido de um desvio-padrão ( $\pm 1\sigma$ ) representado pela região sombreada ( $\epsilon$ -tube). Novamente, os pontos de dados (amostras) estão representados pelos círculos verdes e os vetores de relevância circunscritos pelos azuis. Note que há somente três vetores de relevância em comparação aos sete vetores de suporte obtidos anteriormente (BISHOP, 2007).

A situação esboçada pela Figura 3.7 representa uma das principais vantagens da RVM sobre as SVM. Confirmando o processo de maximização dos parâmetros, a quantidade de vetores eliminados é bem significativa tornando o Modelo de Predição mais otimizado e com mais baixo custo computacional na fase de implementação sendo este um modelo potencialmente importante e valorizado neste trabalho.

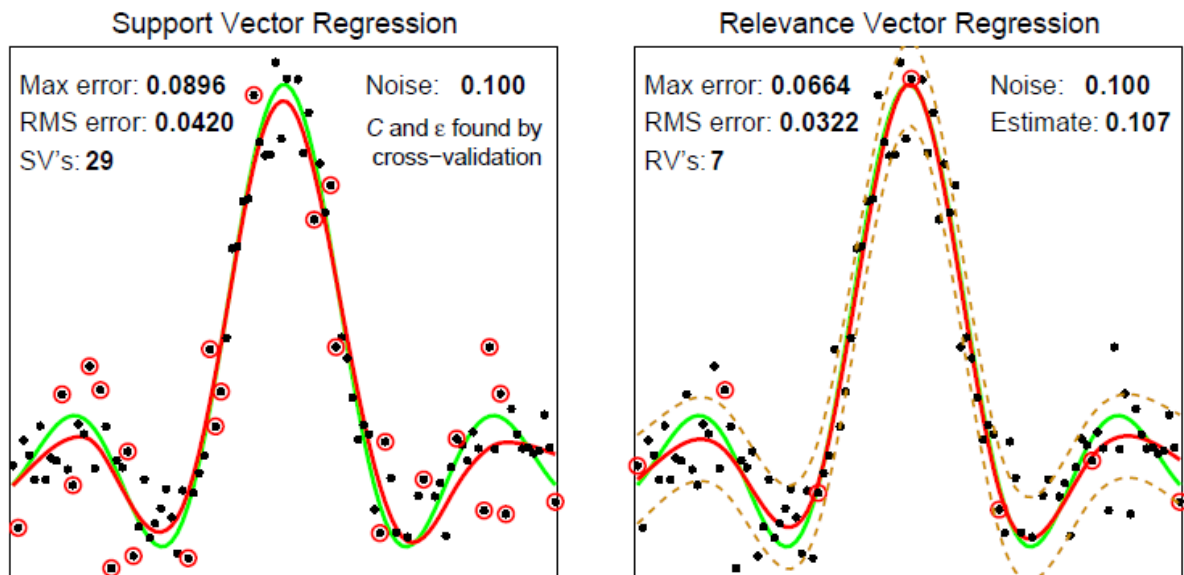
Além disso, como mencionado anteriormente, a RVM é uma técnica Bayesiana herdando assim suas principais propriedades. Segundo Tipping, essas propriedades correspondem a (TIPPING, 2006):

- Definir distribuições de probabilidades (Função de densidade) sobre todas as variáveis (Independente e Dependente) do Modelo de Predição;
- Eliminar por integrações variáveis que não são diretamente relevantes para fazer predições;
- Atualizar as distribuições baseado nos dados de treinamento do modelo (amostra) usando para isso a Regra de Bayes;
- Formas altamente eficazes para lidar com todos os tipos de incertezas;
- Modelos explícitos para representar o conhecimento prévio;
- Execução automática do princípio da "Navalha de Ockham"; ou seja, é o princípio de que "as entidades não devem ser multiplicadas além da necessidade" e a conclusão da mesma é que a explicação ou estratégia mais simples tende a ser a melhor. Este princípio é atribuído ao matemático inglês do século 14 e frade franciscano, Guilherme de Ockham.

Para o caso das RVM, a distribuição *a posteriori* é computada via Regra de *Bayes* da seguinte forma:

$$p(w|t, \alpha, \sigma^2) = \frac{\text{Verrossimilhança } x \text{ a priori}}{\text{fator normalizador}} = \frac{p(t|w, \sigma^2) x p(w| \alpha)}{p(t| \alpha, \sigma^2)}$$

Outro exemplo mais completo mostrado na Figura 3.8 formula uma comparação direta entre a SVM e a RVM aplicadas a uma mesma base de dados hipotética. Neste exemplo, os principais parâmetros ajustados para ambas as técnicas se tornaram notórias as vantagens obtidas em relação à aplicação otimizada dos vetores de relevância.



**Figura 3.8** As Máquinas de *Kernel* Esparsos aplicadas a um problema de regressão usando um *Kernel* Gaussiano, o que demonstra algumas das vantagens da técnica Bayesiana (TIPPING, 2006).

Com a introdução da técnica e do algoritmo de aprendizagem Bayesiana, a RVM utiliza em seu Modelo de Predição diferenciada ferramenta probabilística não presente nas SVM. Essa característica, reforçando uma das propriedades dos vetores de relevância, torna o modelo uma significativa técnica para resolução de problemas de regressão com conhecimento prévio do domínio do problema; entretanto, envolto em alguns tipos de incertezas em relação principalmente às variáveis (independente e dependente) presentes no processo.

### 3.3 Técnica híbrida neural-adaptativa de regressão

As variáveis de processos envolvidas nos estágios detalhados no capítulo 2 apresentam características dinâmicas bem peculiares em relação a suas propriedades ao longo de todo macro processo de Pelotização. Essas características das variáveis de processo são comumente referenciadas pelos especialistas do negócio em grupos que denominam aspectos referentes à qualidade das mesmas em tempo real.

As variáveis de processo mais significativas, presentes em todos os níveis da Pirâmide da Automação, assumem valores contínuos ao longo do tempo sendo quantitativamente mensuradas no decorrer de todas as etapas. Entretanto, essas mesmas variáveis são referenciadas qualitativamente de maneira não exata e precisa em relação aos seus valores absolutos assumidos ao longo do processo.

Esses aspectos de incertezas cotidianamente vivenciados no monitoramento e controle das variáveis de processo denotam características difusas e nebulosas em relação as suas propriedades qualitativas nem sempre de consenso e senso comum entre os especialistas do negócio. Tais aspectos, por decorrência de sua complexidade, nem sempre são obviamente distinguíveis principalmente em função do dinamismo presente em processos industriais contínuos.

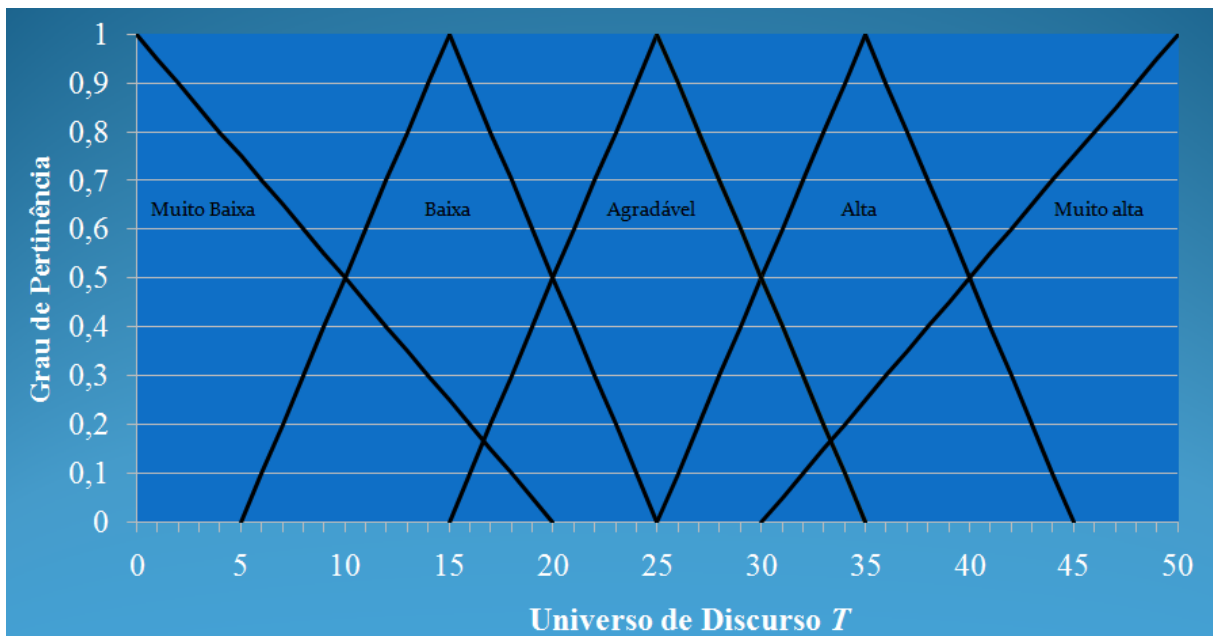
Uma linha de pesquisa da Inteligência Computacional que se atentou para esses aspectos originou-se inicialmente pelos conceitos fundamentais dos Conjuntos *Fuzzy* (ZADEH, 1965). A propriedade mais elementar do conjunto *Fuzzy* é a indistinguibilidade observada na variável através dos termos propostos pelos especialistas que a dominam. Um modelo que podemos destacar é o *Fuzzy Inference System* – FIS – que objetiva juntamente contemplar tal propriedade (JANG, 1997).

Este sistema é uma ferramenta da Inteligência Computacional que promove a junção de duas grandes linhas de pesquisa – Redes Neurais e Lógicas *Fuzzy* – na resolução de problemas de regressão. Deste ponto, podemos destacar o que individualmente essas linhas contribuem na resolução de problemas desta natureza e os principais conceitos pertinentes no desenvolvimento do Modelo de Predição.

Na outra vertente deste sistema, devemos destacar os conceitos referentes aos conjuntos *Fuzzy* (ZADEH, 1965). O principal conceito é representado pela própria palavra *Fuzzy* e pode ser compreendida como algo "difuso, indistinguível e / ou nebuloso" associado

às variáveis do sistema, características também observáveis nas variáveis de processo de um sistema industrial de regime contínuo de produção.

Essas variáveis, também comumente chamadas de variáveis lingüísticas, recebem este nome devido aos termos nominais difusos associados a seus valores (JANG 1997). Esses valores formam um conceito denominado de universo de *discurso* que representa os valores numéricos válidos atribuíveis a mesma e o conjunto formado por estes termos são denominados de Partição *Fuzzy*. A menção feita aos elementos individuais deste conjunto é denominada de Termos Primários. Os conceitos mencionados são correlacionados na Figura 3.9 a seguir.

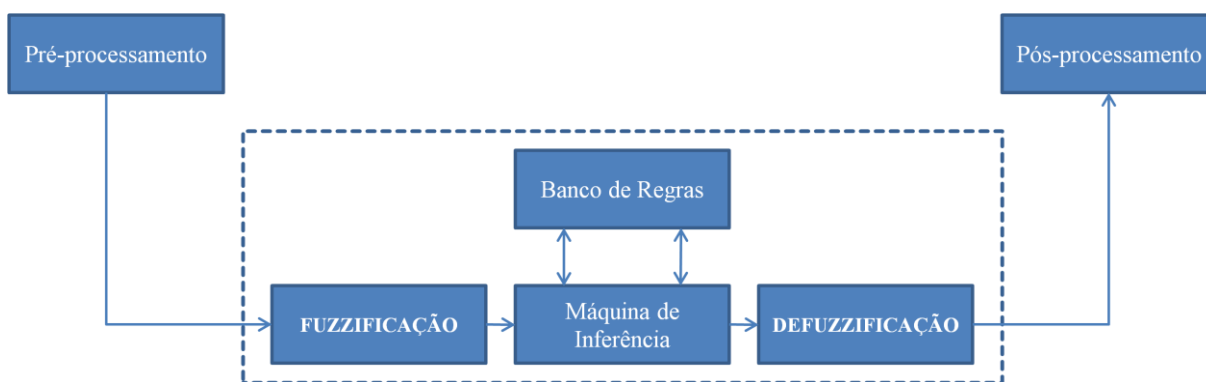


**Figura 3.9** Relação dos conceitos da teoria *Fuzzy*. Exemplo de Variável Lingüística (Temperatura do ar condicionado) estabelecida a partir do seu Universo de Discurso  $T: \{X \in \mathcal{R}: X = [0, 50]\}$  e nominada pelos Termos Primários {Muito Baixa, Baixa, Agradável, Alta, Muito Alta} onde este conjunto estabelece a Partição Fuzzy do sistema (REZENDE, 2003).

Além dos conceitos citados, outro conceito fundamental na teoria *Fuzzy* é a Função de Pertinência (em inglês: *Membership Function* – MF) aplicada individualmente aos termos primários da partição *Fuzzy* (KEVIN & YURKOVICK, 1997). Essa função estabelece uma associação entre os elementos do universo de discurso das variáveis lingüísticas do sistema formando assim subconjuntos (agrupamentos) do mesmo. Essa função deve ser individualmente estabelecida para cada variável lingüística do sistema de acordo com seu comportamento levantado em consideração aspectos do ambiente real.

Além disso, essa função determina o *grau de pertinência* dos termos primários. Esse grau estabelece o nível com o qual os termos nominais devem ser quantificados ao longo do sistema *Fuzzy*. Esse processo de quantificação dos termos das variáveis lingüísticas em graus de pertinência denomina-se Fuzzificação (*Fuzzification*).

Diante dos conceitos mencionados, podemos ilustrar na Figura 3.10 a topologia dos sistemas FIS. Esta topologia estabelece uma visão sistêmica do sistema e seus principais componentes.



**Figura 3.10** Topologia do sistema FIS (JANG 1993).

O sistema FIS associa todos os conceitos mencionados destas duas linhas de pesquisa em um Modelo de Predição híbrido que apresenta como importante característica a possibilidade de incluir esses aspectos indistinguíveis representados pelos termos primários individualmente nominados no conjunto da partição *Fuzzy*. Esses aspectos são rotineiramente referenciados pelos especialistas do domínio quando se reportam ao monitoramento e controle das variáveis do processo.

Contudo, o ajuste dos diversos parâmetros FIS e a determinação da partição *Fuzzy* responsável pelo conteúdo das tarefas presentes na topologia do sistema são etapas que, pelo nível de complexidade e pelo grande número de variáveis lingüísticas envolvidas, podem ser automaticamente descobertas através da aprendizagem de máquina. Essa possibilidade de automatismo propicia o descobrimento de informações referentes a essas variáveis que nem sempre condiz com os ajustes praticados pelos especialistas.

Essa estratégia de automatismo acaba confirmando a possibilidade de ajustes mais otimizados e agregando informação mais qualitativa em relação ao monitoramento e controle das variáveis de processo. Além disso, essa estratégia torna possível comparar automaticamente ajustes destinados às variáveis de processo viabilizando o amadurecimento e validação dos mesmos praticados no dia-a-dia.



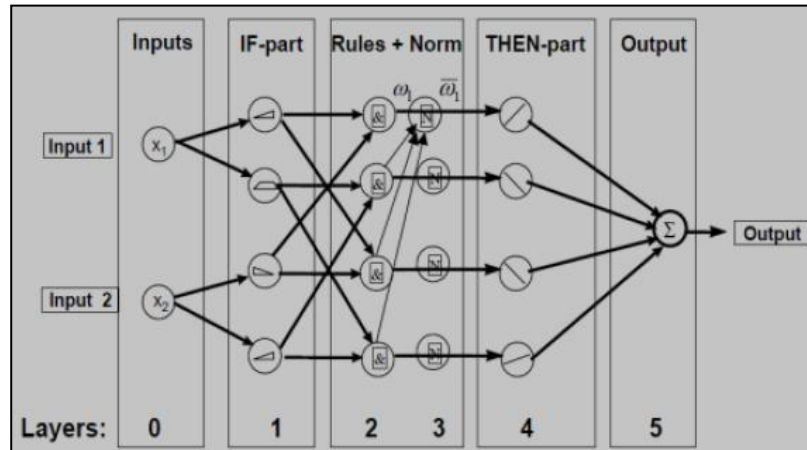
### 3.3.1 O Modelo de Predição Neuro-Fuzzy

Uma destacável instância dos modelos FIS é o Sistema de Inferência Adaptativa Neuro-Fuzzy (em inglês: *Adaptive Neuro-Fuzzy Inference Systems* – ANFIS) utilizada neste projeto (JANG 1993). Essa solução híbrida de Modelo de Predição viabiliza a aprendizagem de máquina dos parâmetros internos da topologia ANFIS que são adequados segundo o princípio de generalização a partir do conjunto de treinamento.

A topologia ANFIS é análoga a topologia do sistema FIS que por sua vez utiliza-se das estruturas consolidadas das redes neurais (JANG, 1993). A estrutura ANFIS mostrado na Figura 3.11 pode ser representada em camadas com as seguintes características estabelecidas:

- **Camada Zero:** Entrada do FIS onde cada nó  $i$  representa uma característica mapeada com a saída desejada;
- **Camada Um:** Cada nó  $i$  passa por uma Função de Pertinência (MF) onde são determinados os parâmetros de ajustes e os graus de pertinência para cada função. Nesta camada que são associados os termos lingüísticos de cada variável;
- **Camada Dois:** A Base de Regras é estabelecida nesta camada e esta tarefa se dá por duas etapas: Elaboração das regras e a interação dos resultados dos nós. A primeira etapa estabelece o grau de ativação de cada regra e seu posterior conseqüente sobre cada nó e a operação estabelecida é referenciada pela denominação de *T-norm* e geralmente é obtida pelo produto dos nós;
- **Camada Três:** A Normalização pondera o nível de ativação de cada nó em relação aos valores ativados em todos os outros nós desta camada.
- **Camada Quatro:** Segue o modelo estabelecido por e determina os ajustes de parâmetros necessários (TAKAGI & SUGENO, 1985);
- **Camada Cinco:** corresponde ao valor geral de saída obtido conforme os parâmetros ajustados na camada anterior.

A Função de pertinência é justificada determinantemente por dois critérios específicos: Tipo de função e ajuste de parâmetros (CHIU, 1996). O tipo de função adotado neste trabalho foi escolhido em relação ao teste da distribuição da probabilidade aplicado as variáveis lingüísticas do sistema e é detalhado no capítulo cinco.



**Figura 3.11** Estrutura do sistema ANFIS (JANG, 1993)

O segundo critério que altera a função de pertinência é o ajuste de parâmetros, que no caso do sistema ANFIS, são determinados por quatro parâmetros específicos, a saber, (CHIU, 1996):

1. Faixa de influência (*Influence Range* ou *Clusters Radius*) –  $r_a$ : Este fator é usado para multiplicar os raios que determinam os "vizinhos" do centróide; bem como, por consequência, anulam a possibilidade de "vizinhos distantes" pertencerem a este centróide.
2. Fator *Squash* (*Squash Factor*) –  $r_b$ : Este fator é uma constante positiva aplicada sobre a faixa de influência.
3. Coeficiente de Aceitação (*Accept Ratio*) –  $\bar{\epsilon}$ : Esse coeficiente ajusta o potencial, como fração do potencial do primeiro centróide com que outro dado seja aceito como um centróide.
4. Coeficiente de Rejeição (*Reject Ratio*) –  $\underline{\epsilon}$ : Esse coeficiente ajusta o potencial, como fração do potencial do primeiro centróide com que outro dado seja rejeitado como um centróide.

Esses parâmetros alteram a quantidade de MF e conseqüentemente o número de regras estabelecida na camada dois do sistema ANFIS. O ajuste mais otimizado destes parâmetros é proposto por Chiu e possuem o comportamento variando nos seguintes valores:  $0.15 \leq r_a \leq 0.30$ ,  $1.25 \leq r_b \leq 1.5$ ,  $\bar{\epsilon} = 0.5$  e  $\underline{\epsilon} = 0.15$  (Chiu 1996).

O algoritmo de aprendizagem da rede ANFIS é similar ao aplicado as redes neurais e se dá em duas etapas distintas, sendo que os ajustes dos parâmetros do sistema ocorrem nos dois sentidos da rede conforme a tabela abaixo.

**Tabela 3.2: Etapas de aprendizagem - ANFIS**

ETAPAS DE APREDIZAGEM		
ETAPA	PASSO	
	FORWARD	BACKWARD
Parâmetros preliminares	Fixos	Descida de Gradiente
Parâmetros conseqüentes	Estimador Least-Squares	Fixos
Sinal de Referência	Nós de Saída	Erro Sinal

O método utilizado na etapa *Forward* geralmente é a somatória do erro quadrático médio (LSE – *Least Squared Error*) e os valores ajustados são as saídas dos nós de cada camada descrita anteriormente. A etapa *Backward* é obtida pelo algoritmo de descida de gradiente, geralmente *Backpropagation*, sendo que nesta etapa a referência passa a ser o sinal de erro.

O Sistema de Inferência *Fuzzy* descrito aqui através da instância ANFIS é um Modelo de Predição que contempla um importante aspecto num ambiente industrial produtivo que é a referência feita às variáveis de maneira nominalmente indistinguíveis o que é contemplado pela partição *Fuzzy* elaborada automaticamente pelo sistema. Após todos os ajustes, os elementos associados pela função de pertinência agrupam-se formando os termos primários.

Esse agrupamento, no caso do sistema ANFIS, é feito automaticamente pelos ajustes dos parâmetros através de uma técnica denominada de *Subtractive Clustering* (JANTZEN 1998) e é utilizada neste trabalho como etapa de construção do Modelo de Predição resultante da técnica adaptativa *Neuro-Fuzzy*.

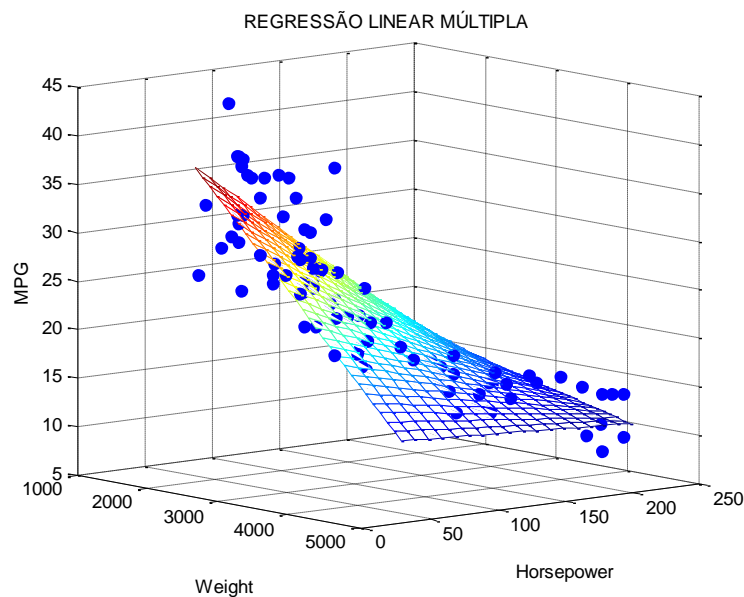
O termo adaptativo refere-se justamente ao princípio pelo qual os elementos são agrupados (*clusters*) automaticamente baseados no conjunto de aprendizado formando adaptativamente a partição *Fuzzy* do sistema.

### 3.4 Síntese das técnicas de construção de Modelos de Predição

Os Modelos de Predição são obtidos pelos ajustes dos algoritmos de aprendizagem através de seus parâmetros. Esses parâmetros de ajuste são também comumente referenciados como grau de liberdade dos algoritmos de aprendizagem (HAIR, 2009). A seguir, vamos explicar os parâmetros que podem ser manipulados na etapa de treinamento do algoritmo de aprendizagem.

O primeiro Modelo de Predição é extraído do algoritmo de Regressão Linear Múltipla que possui como parâmetro de ajuste o *Slop* (CHATTERJEE, 1986). Este parâmetro altera o intervalo de confiança dos coeficientes (pesos) estimados. A saída do algoritmo são os pesos e sua característica estritamente linear estabelece um plano com as mesmas propriedades também lineares como mostrado na Figura 3.12.

Na verdade, este é o modelo mais clássico e simples de análise de regressão apresentado aqui somente para confirmar as limitações de sua utilização em aplicações cujo plano de saída (da variável dependente) possui características não-lineares observadas do fenômeno físico-químico que a originou.

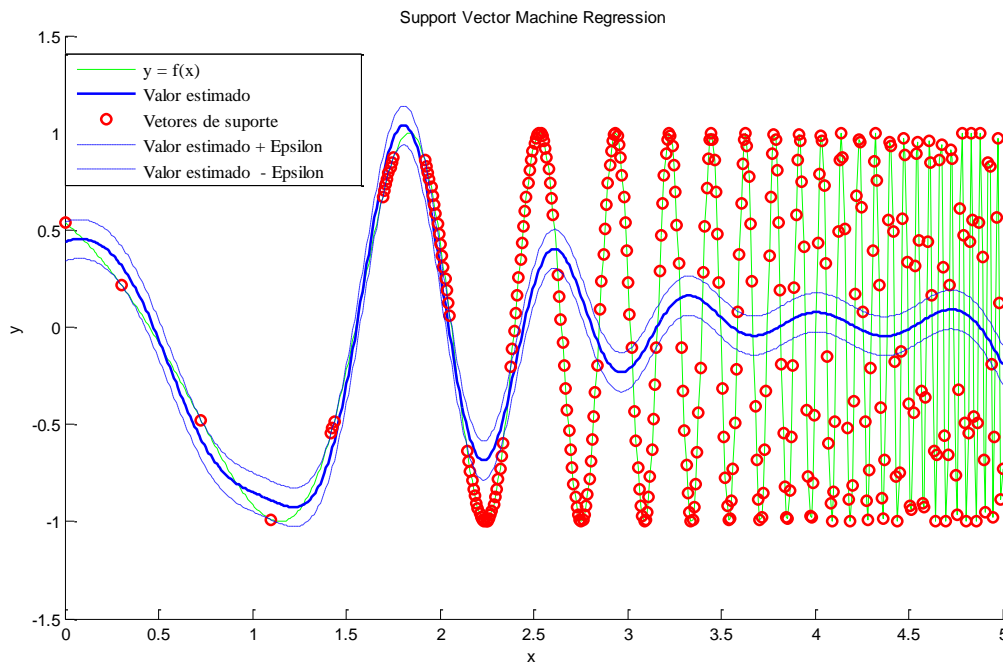


**Figura 3.12** Exemplo de Regressão Linear Múltipla confirmando as características lineares do plano ao modelo de saída com intervalo de confiança de 95%. (CHATTERJEE, 1986).

Os Modelos de Regressão avançados foram obtidos com as técnicas das Máquinas de *Kernel* Esperso implementadas pelos algoritmos da Máquina de Vetor de Suporte (SVM) e Máquina de Vetor de Relevância (RVM) apresentada anteriormente. O algoritmo SVM utilizado neste trabalho possui como graus de liberdade os seguintes parâmetros (RAKOTOMAMONJY, 2008):

1. Kernel: a função estabelecida pela variável dependente conforme testes realizados em seus dados descrevem o comportamento da função que a caracteriza;
2. Épsilon: Determina o comportamento da função  $\varepsilon$ -tube;
3. C: Este parâmetro controla a relação entre as folgas da função da variável dependente e as margens da função  $\varepsilon$ -tube;
4. Lambda: Parâmetro para os métodos de condicionamento da Programação Quadrática;
5. Largura do Kernel (*Kernel Option*): É o valor escalar ou um vetor dos valores Gamma idêntico a todas as variáveis de entrada.

Esses parâmetros determinam e otimizam a quantidade de Vetores de Suporte pertencentes ao modelo de saída conforme mostrado na Figura 3.13. Os outros parâmetros de saída do algoritmo são pesos e bias pertencentes também ao modelo. O conjunto de dados de treinamento foi aplicado a este algoritmo para desenvolvimento do Modelo de Predição pela SVM.

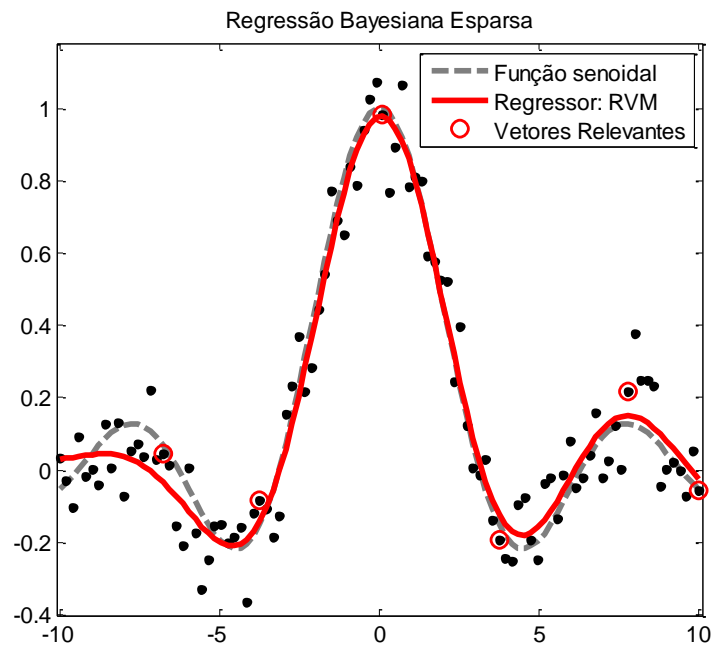


**Figura 3.13** Exemplo de SVM para regressão, com a função de Kernel gaussiana,  $C = 10$ ,  $\text{Épsilon} = 0.1$ ,  $\text{Lambda} = 0.000001$  e  $\text{Largura Kernel (Kernel Option)} = 0.4$ . (RAKOTOMAMONJY, 2008)

A segunda Máquina esparsa mencionada anteriormente é a RVM implementada com dois parâmetros de ajuste (Tipping, 2006):

1. Largura do Kernel (Escala): Largura escalar da função de Kernel (Cânon);
2. Alfa: Valor inicial escalar dos hiperparâmetros.

Esses parâmetros definidos previamente para o algoritmo de aprendizagem determinam a quantidade de Vetores Relevantes do modelo de saída. Esses vetores, assim como pesos e bias, formam o modelo de saída aplicado evidentemente na predição de novas amostras.



**Figura 3.14** Exemplo de RVM para regressão com a função de *kernel* gaussiana, Largura do Kernel = 100 e Alfa inicial = 3.

Em contrapartida, o Modelo de Predição advindo da técnica Neuro-Fuzzy estabelece quatro parâmetros de ajustes, como mencionados anteriormente e são responsáveis pela quantidade de funções de pertinência estabelecida como termos primários do sistema. A quantidade de termos primários deve ser determinada manualmente conforme características do processo para o qual está se desenvolvendo o sistema.

## Capítulo 4: Medidas de adequação dos Modelos de Predição

### 4.1 Introdução

Os Modelos estabelecidos a partir de técnicas (Regressão e Neuro-*Fuzzy*) de aprendizagem de máquina atribuídos a Inteligência Computacional conforme mencionados no capítulo anterior podem servir aos seguintes propósitos (MONTGOMERY, 2001):

- Descrição dos dados do sistema: especialistas podem usar esses modelos na descrição dos conjuntos de dados;
- Estimativas de parâmetros: determinar os melhores ajustes para otimização dos sistemas;
- Predição e estimativa: Aplicações que envolvam a estimativa da variável dependente do sistema.
- Controle: Modelos de Predição podem ser usados no controle das variáveis independentes do sistema.

Como citado anteriormente, os Modelos de Predição obtidos pelas técnicas de aprendizagem de máquina são partes do sistema destinado ao controle avançado de processo; neste trabalho, especificamente Industrial (APC – *Advanced Process Control*). Com isso, os modelos estabelecidos neste trabalho perfazem quase todos os propósitos citados acima exceto o de controle que será deixado para trabalho futuro.

Diante disso tudo, medidas que possam qualificar o nível de adequação (Generalização) dos Modelos de Predição devem ser adotadas. Essas medidas compreendem cálculos e análises comparativas que se utilizam de coeficientes estabelecidos a partir dos resultados dos Modelos de Predição obtidos na etapa de aprendizagem de máquina.

Um parâmetro utilizado neste trabalho com estimativa do erro é o RMS (do inglês: *Root Mean Square*); ou seja, raiz do erro quadrático médio, que é dado pela fórmula (MONTGOMERY & RUNGER, 2003):



$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f(X_i))^2}$$

A estimativa é dada pelo método de substituição dos valores utilizados no conjunto de aprendizagem (Treinamento) no cálculo do erro. Esses valores demonstram a diferença entre o valor desejado e o valor estimado da variável dependente pelo Modelo de Predição. Esse coeficiente deve ser minimizado para conotar uma técnica de construção significativa; entretanto, não sobre ajustado de modo a gerar modelos com baixo desempenho quando aplicados a novas amostras.

Para analisar de maneira comparativa o limiar entre generalização e o sobre ajuste (*Overfitting*) dos Modelos de Predição desenvolvidos neste trabalho que o método de substituição pode esconder erros nos ajuste dos parâmetros dos modelos que devem ser visualizados através de outras técnicas. Neste trabalho, foram separados dois conjuntos de dados de amostras que são aplicados os modelos e comparados os valores RMS destes conjuntos. Essa heurística é detalhadamente explicada no capítulo 5.

Outro índice, diz respeito ao quanto o Modelo de Predição desenvolvido consegue explicar a variável dependente estimada em distinção à média da mesma, esse índice é o Coeficiente de Determinação. Essa estratégia denomina-se Análise da Variância e quantifica o grau de influência das variáveis independentes  $\mathcal{X}$  sobre a variável dependente  $\mathcal{Y}$  (HAIR, 2009).

Essa análise compara-se os valores estimados pelo modelo em relação à própria média da variável dependente (BARBETTA, 2009). Com isso, espera-se que deve haver algum ganho na utilização do Modelo de Predição confirmando o grau de influência de  $\mathcal{X}$  sobre  $\mathcal{Y}$ ; caso contrário, o valor esperado de  $\mathcal{Y}$  poderia ser estimado simplesmente pela média aritmética das observações de  $\mathcal{Y}$ .

$$\underbrace{\sum_{i=1}^n (y_i - \bar{y})^2}_{\text{Variação Total}} = \underbrace{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}_{\text{Variação explicada pelo Modelo de Predição}} + \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)^2}_{\text{Variação não explicada}}$$

onde:

- $n$ : Número de amostras;

- $y_i$ : Variável dependente;
- $\bar{y}_i$ : Média aritmética da variável dependente;
- $\hat{y}_i$ : Valor estimado pelo Modelo de Predição.

Com isso, o Coeficiente de Determinação  $R^2$  é determinado pela fórmula abaixo:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{\text{Variação explicada}}{\text{Variação total}}$$

Com intuito de quantificar a influência do número de amostras e a quantidade de variáveis independentes utilizadas no Modelo de Predição desenvolvido na etapa de aprendizagem é determinado o coeficiente de determinação ajustado  $R^2_{Ajustado}$  para novas amostras aleatoriamente aplicadas ao modelo. A formulação deste coeficiente é calculado por:

$$R^2_{Ajustado} = 1 - \frac{(n - 1)}{(n - k - 1)} (1 - R^2)$$

Onde:

- $n$ : Número de amostras;
- $k$ : Quantidade de variáveis independentes  $\mathcal{X}$ ;
- $R^2$ : Coeficiente de determinação.

Esses coeficientes associados com os valores obtidos pela estimativa do erro RMS são ferramentas para auxiliar na análise comparativa dos Modelos de Predição propostos anteriormente. Essa análise consiste em um diagnóstico observado em relação a esses coeficientes quando os Modelos de Predição são aplicados a outros conjuntos que não foram utilizados na etapa de aprendizagem de máquina.

Novas amostras devem ter a análise do erro estimado cuidadosamente observado, pois podem indicar importantes propriedades e diagnósticos de falhas na adequação (generalização) estabelecidas pelos modelos. Com isso, uma ferramenta de análise fundamental é determinada pela observação residual do erro.

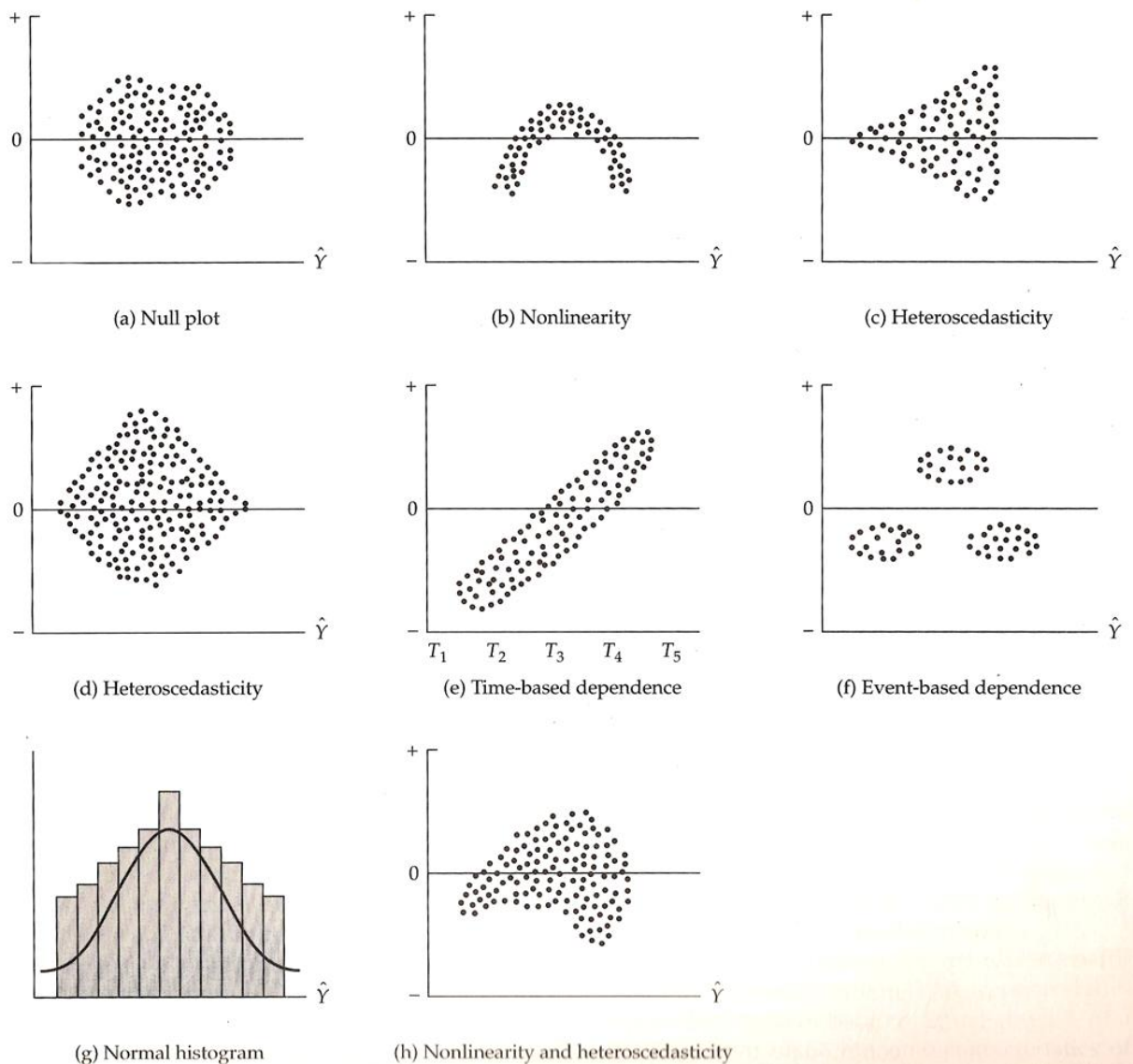
## 4.2 Análise residual

A Análise Residual do erro é uma ferramenta importante a ser associada em conjunto com os coeficientes de estimativa do erro e a análise da variância dos Modelos de Predição propostos (BOWERMAN & O'CONNELL 1997). A principal medida do erro de predição para a variável dependente é a Residual; ou seja, a diferença entre os valores observados (amostras) e os valores estimados pelos modelos.

Segundo Hair, os modelos de Predição propostos não devem violar as propriedades (Premissas) descritas a seguir (HAIR, 2009):

1. Linearidade do fenômeno (do inglês: *Linearity of the phenomenon measured*): A linearidade da relação entre as variáveis independentes e dependentes representa o grau com que modificações na variável dependente são associadas (explicadas) pelas variáveis independentes do modelo;
2. Variância constante dos termos do erro (do inglês: *Constant variance of the error terms*): A presença de variância desigual (Heterocedasticidade) é uma violação dos Modelos de Predição em relação às premissas assumidas pelas técnicas de aprendizagem. O diagnóstico é feito pela análise gráfica do resíduo normalizado com os valores estimados pelo modelo, tais valores devem ser uniformes e aleatórios;
3. Independência dos termos do erro (do inglês: *Independence of the error terms*): os Modelos de Predição assumem que cada valor posteriormente predito (estimado) é independente. Com isso, não existam influências sobre a variável dependente advindas de quaisquer outras variáveis independentes;
4. Distribuição normal dos termos de erro (do inglês: *Normality of the error term distribution*): as variações dos termos dos erros (independente / dependente) devem seguir um comportamento da distribuição normal.

Os gráficos da Análise Residual (Figura 4.1) fornecem uma visão sistêmica em relação ao comportamento que deve ser esperado do erro residual normalizado bem como os possíveis erros de violações observados nas amostras e nos Modelos de Predição propostos segundo os preceitos mencionados anteriormente.



**Figura 4.1** Gráficos da Análise Residual (HAIR, 2009)

O comportamento ideal almejado pelos resíduos normalizados observados a partir dos Modelos de Predição versus os valores estimados devem seguir o modelo do gráfico (a) - (g) confirmando que a técnica de aprendizagem utilizada não violou os preceitos esperados pelos algoritmos de aprendizagem. É notório o comportamento uniforme e aleatório dos termos do erro normalizado; tal comportamento denomina-se Homocedasticidade, oposto ao ilustrado pelos gráficos (c) (d) (h). Tais comportamentos violam as premissas abordadas nos itens 1) e 2) acima auxiliando no diagnóstico de adequação dos Modelos de Predição.

Outra violação comumente observada na Análise Residual é a influência de variáveis não previstas nos modelos. Esse comportamento resultaria em dependências baseadas no tempo ou em eventos não contemplados pelos Modelos de Predição observados

respectivamente nos gráficos (e) - (f). Essas situações apontariam necessidades de análise pormenorizadas das variáveis independentes utilizados nos modelos e evidenciam violações como estabelecida pelo item 3).

O comportamento dos gráficos (b) - (h) aponta para modelos não-lineares que devem ser analisados em última instância em relação às técnicas de aprendizagem lineares que podem generalizar os dados sob condições de controle estáveis conforme abordado anteriormente. Tal comportamento evidenciaria a necessidade de pesquisa de controle avançado de processo presentes no Nível Quatro (*Optimization*) da hierarquia de controle e é objeto de investigação de trabalhos futuros.

A análise residual é uma importante ferramenta de diagnóstico em problemas de regressão, pois possibilitam à análise gráfica dos resíduos resultantes da aplicação dos Modelos de Predição as amostras elucidando os possíveis problemas acima mencionados e qualificando assim, o modelo e a técnica de qual se originou. O gráfico da Análise Residual é obtido conforme algoritmo:

1. Aplique o Modelo de Predição extraído a cada amostra do conjunto;
2. Calcule a diferença residual entre o valor real da amostra e o estimado pelo Modelo de Predição.  $e_i = y_i - \hat{y}_i$ ;
3. Normalize cada valor residual  $e_i$  entre o valor mínimo e máximo conforme fórmula 
$$\|e_i\| = \left\{ \frac{(e_i - MenorValor(e_i)) * (Máximo - Mínimo)}{(MaiorValor(e_i) - MenorValor(e_i))} \right\} + Mínimo.$$
 O valor Mínimo é calculado por  $Mínimo = (-1) * Máximo$  por serem simétricos. Neste trabalho, o valor mínimo e máximo é pré-determinado no intervalo fechado:  $[-1, 1]$ ;
4. Reagrupe os valores normalizados  $\|e_i\|$  em ordem crescente e correlacione por par ordenado com o seu respectivo valor estimado  $\hat{y}_i^*$ ;
5. Construir o gráfico de dispersão:  $\hat{y}_i$  versus  $\|e_i\|$ .

As medidas e critérios que qualificam os níveis de adequação dos Modelos de Predição devem ser norteados pelos conceitos e premissas abordadas neste capítulo. Tais medidas são ferramentas estatísticas de investigação e análise comparativa entre os Modelos de Predição Neuro-*Fuzzy* e de regressão abordados anteriormente e é objetivo principal deste trabalho.

## Capítulo 5: Resultados dos Algoritmos de Aprendizagem

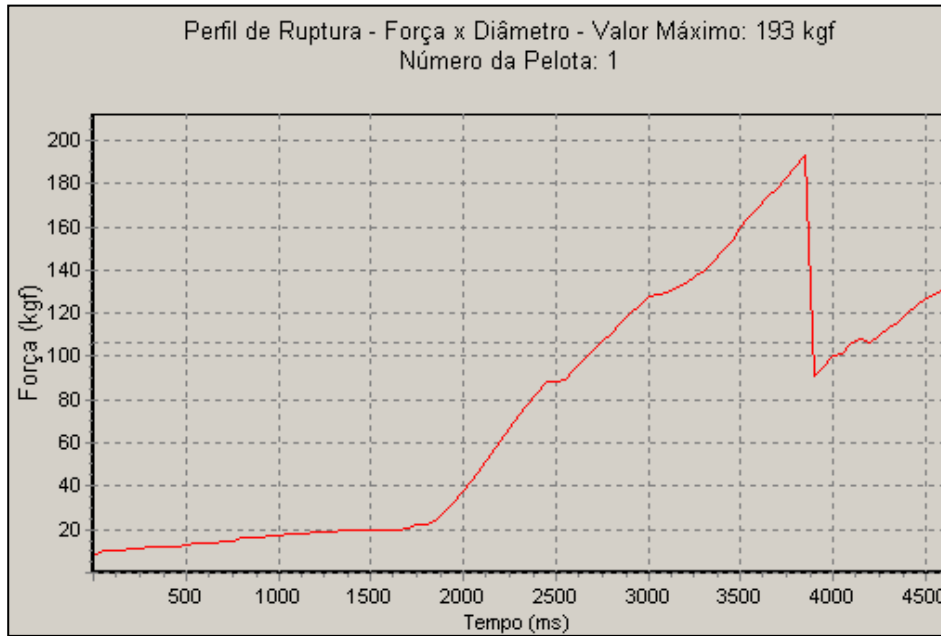
Os Modelos de Predição abordados neste trabalho são investigações e análises baseadas em dados do mundo real ou seja, é uma pesquisa empírica dos dados de processo de um ambiente industrial de mineração sob a óptica da predição visando posteriormente o controle das variáveis de processo envolvidas. Diante disso, é elaborado um ambiente de teste que possa conduzir os experimentos propostos associados aos algoritmos de aprendizagem.

Este ambiente utiliza-se de todos os conceitos e premissas abordadas até o momento e é alicerce para análise dos dados envolvidos, estimativas dos parâmetros necessários e predição por estimativa da variável de qualidade do produto neste processo produtivo contínuo industrial. Essa variável de qualidade do produto é a compressão mecânica da pelota de minério de ferro a frio quantificada em Kgf / pelota conforme norma de processo ensaiada em laboratório (ISO, 1996). Sua importância deve-se basicamente a (MEYER, 1980):

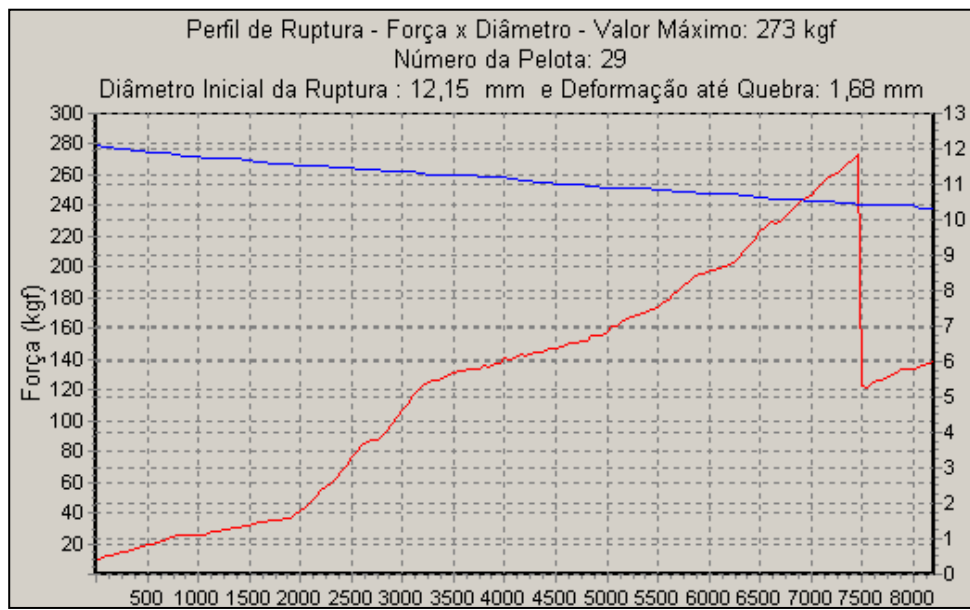
- Manter a integridade estrutural da pelota ao longo de toda manipulação e transporte desde sua produção até a utilização no cliente;
- Evitar desgastes de sua superfície no intuito de manter o diâmetro dentro das especificações desejadas.

Essa é uma variável complexa de processo que se encontra no Nível Três da hierarquia da Pirâmide da Automação, pois seu monitoramento é impactado por uma quantidade significativa de variáveis mais simples localizadas nos Níveis Um e Dois da mesma pirâmide. Além da alta dimensionalidade decorrente da quantidade de variáveis que a impactam, tal complexidade é acentuada pelo nível de conhecimento de processo que tal variável demanda na interpretação de seus valores.

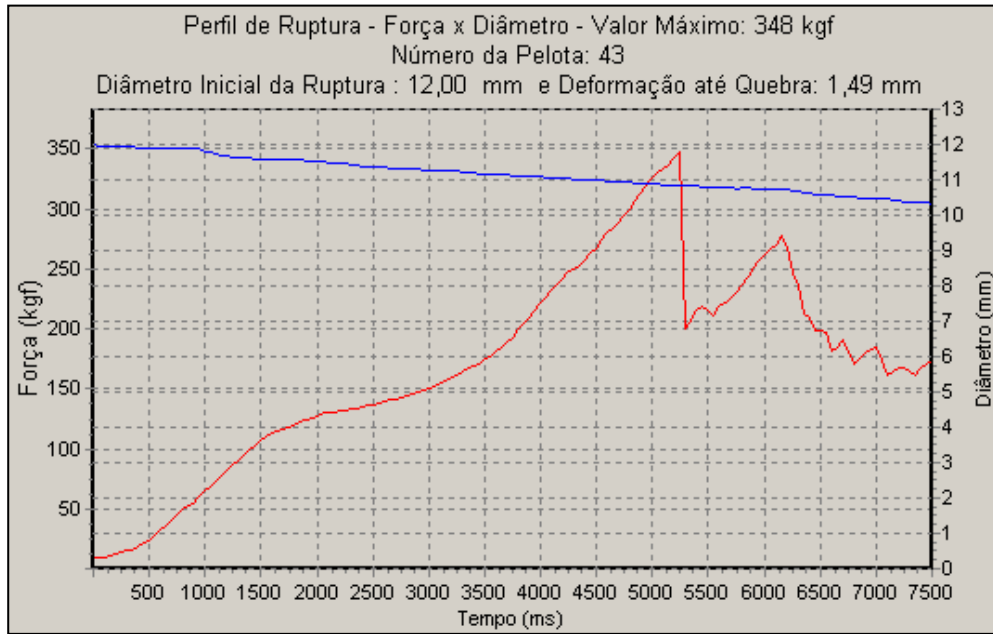
O objetivo principal dos Modelos de Predição é estimar essa variável de qualidade do produto baseado nas variáveis de processo que a impactam em tempo real. Essa variável apresenta um comportamento curvilíneo (polinomial) e distribuição normal (Apêndice A) conforme seus gráficos mostrados na Figura 5.1 observados empiricamente em laboratório.



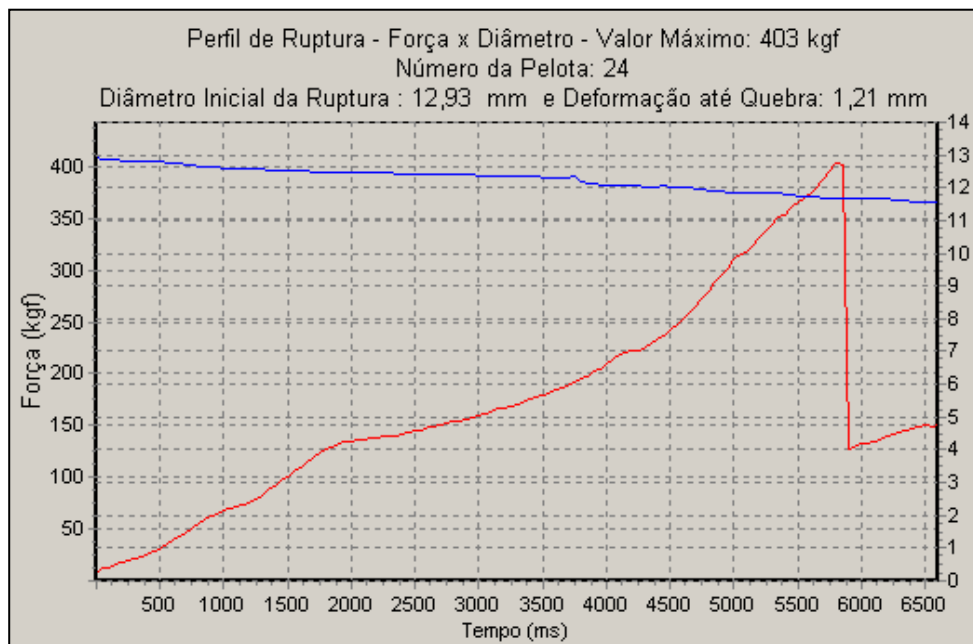
(a)



(b)

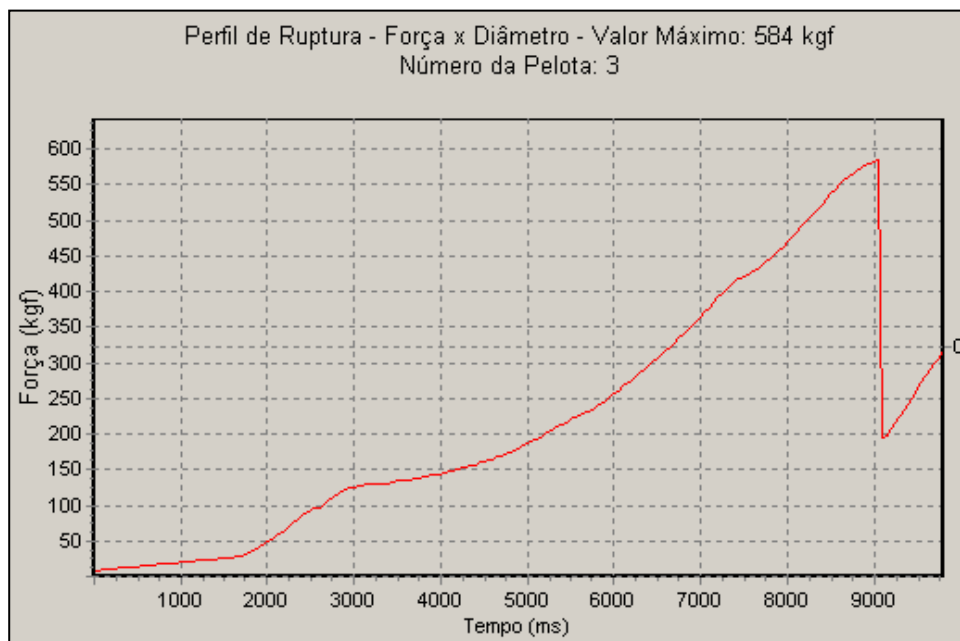


(c)



(d)





(e)

**Figura 5.1** Gráficos da compressão mecânica da pelota de minério de ferro a frio. De maneira difusa, (a) pode ser nominado como um resultado muito baixo de compressão; (b) um resultado baixo; (c) um resultado normal; (d) um resultado alto e o gráfico (e) representa um resultado muito alto de compressão. Esses gráficos foram obtidos aleatoriamente dos resultados de laboratório.

Na verdade, o resultado de compressão mecânica da pelota que é realmente levado em consideração é a média aritmética dos valores máximos (ruptura máxima) de uma quantidade específica de pelotas. Nas amostras utilizadas neste trabalho os ensaios de laboratório utilizavam 50 unidades de pelotas endurecidas recolhidas periodicamente no final do 3º estágio de Pelotização.

Os dados referentes à variável de qualidade em questão e todas as outras variáveis que impactam no resultado foram levantadas para que possam compor os conjuntos de testes utilizados pelos algoritmos de aprendizagem. Esses conjuntos são estratificados no próximo tópico.

## 5.1 Configurações do experimento

A variável de qualidade da pelota de minério de ferro compressão mecânica a frio é obtida em ensaio de laboratório conforme Norma ISO 4700 e possui as seguintes características, a saber (ISO, 1996):

- Universo de discurso: 0 – 800 Kgf / Pelota;

- Amostragem periódica:
  - Uma amostra / HH (a cada 2 horas);
  - 12 amostras / DIA
  - 360 amostras / MÊS
  - 4.320 amostras / ANO
- 50 pelotas de Minério de Ferro endurecidas amostradas e testadas conforme norma.

Entretanto, esse valor bi-horário de qualidade é resultado das diversas ações manuais e automáticas realizadas nos três estágios de processo pelo sistema de monitoramento e controle. Neste trabalho foram levantadas as ações de todo o ano produtivo de 2008 a cada um minuto e as amostras foram concatenadas pela média de cada variável independente que impacta a qualidade a cada duas horas conforme ilustrado pela Figura 5.2.

LEVANTAMENTO DOS DADOS															
PELOTAMENTO							FORNO				GRELHA				COMPRESSÃO
V1	V2	V3	V4	V5	V6	V7	V1	V2	V3	V4	V1	V2	V3	V4	S1
45,0	105,0	222,0	3550,0	3889,0	1449,0	687,0	203,7	148,9	121,2	122,6	5,8	245,4	45,2	26,1	
42,0	103,0	219,0	3803,0	3900,0	1292,0	624,0	203,7	149,9	121,2	122,3	5,7	245,4	44,8	26,1	
36,0	116,0	232,0	3785,0	3960,0	1407,0	688,0	203,7	149,9	122,0	122,3	5,7	239,1	45,0	26,1	
40,0	114,0	210,0	3795,0	3920,0	1383,0	731,0	203,7	149,9	122,0	122,3	5,8	239,1	45,1	26,1	
44,0	139,0	258,0	3677,0	3731,0	1250,0	534,0	202,9	155,7	126,6	129,2	5,9	242,2	44,8	26,7	323
47,0	132,0	260,0	3757,0	3764,0	1237,0	525,0	201,4	155,7	126,6	128,3	5,9	242,2	45,3	26,7	323
51,0	123,0	238,0	3657,0	3764,0	1317,0	620,0	201,4	153,9	126,6	128,3	5,9	239,3	45,1	26,7	323
51,0	136,0	231,0	3717,0	3837,0	1250,0	571,0	201,9	153,1	125,4	127,4	5,9	239,3	45,1	26,7	323
42,0	105,0	239,0	3499,0	3965,0	1377,0	666,0	205,6	151,3	124,2	127,4	5,8	244,0	45,2	26,5	340
45,0	118,0	217,0	3688,0	3768,0	1415,0	712,0	206,1	152,3	124,8	127,4	5,8	244,0	44,9	26,5	340
45,0	106,0	246,0	3648,0	3953,0	1373,0	705,0	206,9	153,3	125,4	127,4	5,9	244,0	44,9	26,5	340
33,0	89,0	190,0	3630,0	4023,0	1403,0	727,0	208,7	153,5	126,0	128,6	5,8	244,0	45,1	26,5	340
															294
															294
															294
															294

RESULTADO: 20:00

RESULTADO: 22:00

RESULTADO: 24:00

**Figura 5.2** Exemplo hipotético de amostragem da variável de qualidade. Neste exemplo, 15 variáveis independentes impactam a variável dependente: Compressão mecânica a frio compondo três amostras obtidas as 20, 22 e 24 horas respectivamente de um dia.

Esse procedimento ilustrado na Figura 5.2 foi recursivamente repetido para todas as amostras válidas obtidas ao longo do ano de 2008. Tal procedimento (*query SQL*) extraia todas as informações das ações tomadas pelo processo que diretamente impactam a variável de qualidade.

Noutra vertente, em relação aos métodos pesquisados e seus respectivos algoritmos de aprendizagem não foram implementadas melhorias no intuito de que perfizessem uma análise

comparativa como conceitualmente foram desenvolvidos por seus autores. Resumidamente, os algoritmos utilizados neste trabalho foram:

- Regressão Linear Múltipla:  
(<http://www.mathworks.com/access/helpdesk/help/toolbox/stats/regress.html>);
- ANFIS:  
(<http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/fp715dup12.html>)
- SVM: (<http://asi.insa-rouen.fr/enseignants/~arakotom/toolbox/index.html>)
- RVM: <http://www.miketipping.com/index.php?page=rvm>

Entretanto, esses algoritmos foram manipulados para serem utilizados em problemas de regressão de várias dimensões de entrada. Com isso, os algoritmos SVM e RVM apresentaram mensagens de advertências que devem ser tratadas como melhorias e perfazem objetos de investigação para trabalhos futuros, bem como a implementação por outras formas canônicas de *Kernel*.

Além disso, foi implementado uma série de métodos: algoritmos e funções que envolvessem os critérios ([Critério 1](#)) / [Critério 2](#)) / [Critério 3](#)) utilizados neste trabalho como ferramenta estatística para análise comparativa dos Modelos de Predição.

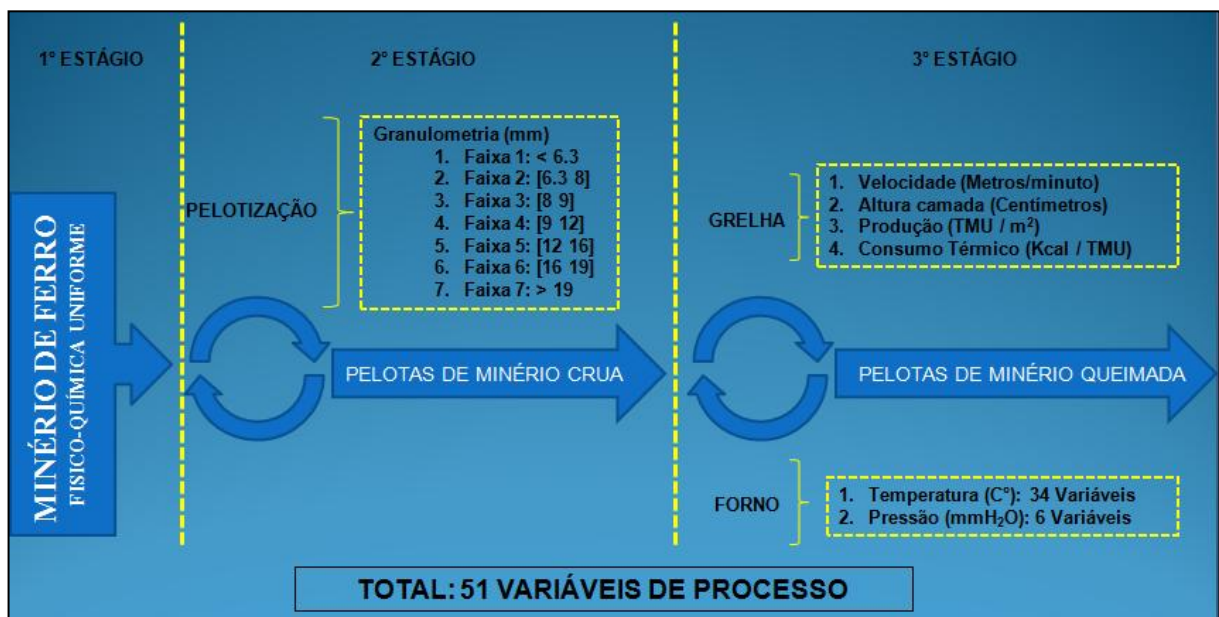
### 5.1.1 Base de dados conceitual do problema

Na verdade, a quantidade de variáveis de processo que impactam a variável de qualidade compressão mecânica engloba um número maior das que estão realmente disponíveis (mensuradas e controladas) ao longo dos três estágios produtivos. Essas variáveis de processo verdadeiramente, mensuradas e / ou controladas, são as variáveis independentes utilizadas nos Modelos de Predição.

Neste experimento foram levantadas conceitualmente 51 variáveis de processo, entre mensuradas e controladas, que impactam diretamente na variável de qualidade em questão conforme a amostragem e estratificação descritas anteriormente. Com isso, resultou em sete gigabytes de arquivos que foram submetidos a diversas operações de seleção e filtragem. Essas operações buscavam solucionar problemas como:

- Dados corrompidos do banco de dados (Servidor de Dados). Esses servidores de dados estão localizados a partir do nível três da Pirâmide da Automação;
- Dados fora da especificação (universo de discurso) de valores;
- Valores discrepantes em relação aos anteriormente estabilizados no curso do processo (*Outliers*); ou seja, são valores que, apesar de estarem na faixa válida do universo de discurso daquela variável, apresentaram discrepância em relação aos comumente praticados naquele mesmo período de análise (BOWERMAN & O'CONNELL, 1997);
- Períodos onde não houve registro das variáveis nos servidores decorrentes de paradas na planta industrial ou falha nos equipamentos de automação (em todos os níveis).

Essas variáveis independentes de processo (51 variáveis) e a variável dependente de qualidade do produto (compressão mecânica a frio) pertencem aos dois últimos estágios do macro processo produtivo da pelota de minério de ferro da empresa Samarco Mineração S.A conforme ilustrado na Figura 5.3:



**Figura 5.3** Visão sistêmica das variáveis de processo que impactam o parâmetro de qualidade da pelota de minério de ferro: Compressão mecânica a frio conforme os macros processos (estágios). Conforme mencionado no Capítulo 2, o 1º estágio não contribuiu com variáveis para os modelos e o 2º e 3º estágios contribuem com 51 variáveis independentes.

Contudo, um total de 21 variáveis de temperatura apresentaram colinearidade e foram concatenadas pela média dos valores. Com isso, na verdade, as variáveis de processo efetivamente totalizaram 30 variáveis independentes utilizadas pelos Modelos de Predição e seus respectivos algoritmos de aprendizagem que os originaram.

Essas 30 variáveis independentes  $\mathcal{X}$  associadas com a variável dependente  $\mathcal{y}$  de qualidade do produto (compressão mecânica a frio) formam a base de dado utilizada ao longo de todos os experimentos.

### 5.1.2 Redução da dimensionalidade por algoritmo de seleção

No item anterior explica detalhadamente a quantidade de variáveis de processo que impactam a variável de qualidade do produto (compressão mecânica a frio) alvo principal a ser estimado por este trabalho através dos algoritmos de aprendizagem de máquina. Para isso, uma base de dados com todas as amostras referentes a essas variáveis foram levantadas e filtradas do ano de 2008 e utilizadas nas etapas de treinamento dos algoritmos.

Essa base conceitual possui uma quantidade significativa de termos (variáveis independentes) que devem ser individualmente investigados e validados qualitativamente no intuito de determinar o seu verdadeiro nível de influência na variável de qualidade. Esse processo de investigação e validação das variáveis independentes (de processo) que efetivamente impactam na variável de qualidade é obtido pelo algoritmo de seleção de características (BENSCH, 2005).

Esse algoritmo de seleção automaticamente determina as variáveis de processo que mais diretamente impactam a variável dependente e a heurística desta seleção pode dar-se de diversas maneiras (BENSCH, 2005). A técnica exaustiva utilizada neste trabalho é o SFS (do inglês: *Sequential Forward Selection*) que consiste em um algoritmo de seleção de características baseado no erro estimado da variável dependente em relação às variáveis independentes individualmente testadas no Modelo de Predição (DUDA, 2001).

O subconjunto de variáveis independentes selecionadas forma uma base de dados que tem seu resultado confrontado com a base de dados conceitual do problema no intuito de quantificar a efetividade do algoritmo de seleção de características. Essa base selecionada e os resultados do algoritmo SFS estão postados no Apêndice B.

As características selecionadas utilizadas neste trabalho foram as estabelecidas pela máquina RVM que desempenhou o menor erro estimado RMS com uma quantidade de quatro variáveis independentes.

### 5.1.3 Heurística de formação dos conjuntos de dados

A base de dados conceitual e selecionada do problema foram heurísticamente subdivididas em três subconjuntos nominados de Treinar, Validar e Testar. Essa Heurística de subdividir a base de dados (amostras) pretende facilitar a investigação dos dados e possibilitar a análise de adequação dos modelos exposto anteriormente (DUDA, 2001).

Resumidamente, por validação cruzada, as amostras foram aleatoriamente redistribuídas nos subconjuntos Treinar, Validar e Testar para as duas bases de dados: a conceitual, com trinta variáveis independentes e a selecionada, com quatro variáveis independentes. Os conjuntos ficam assim distribuídos:

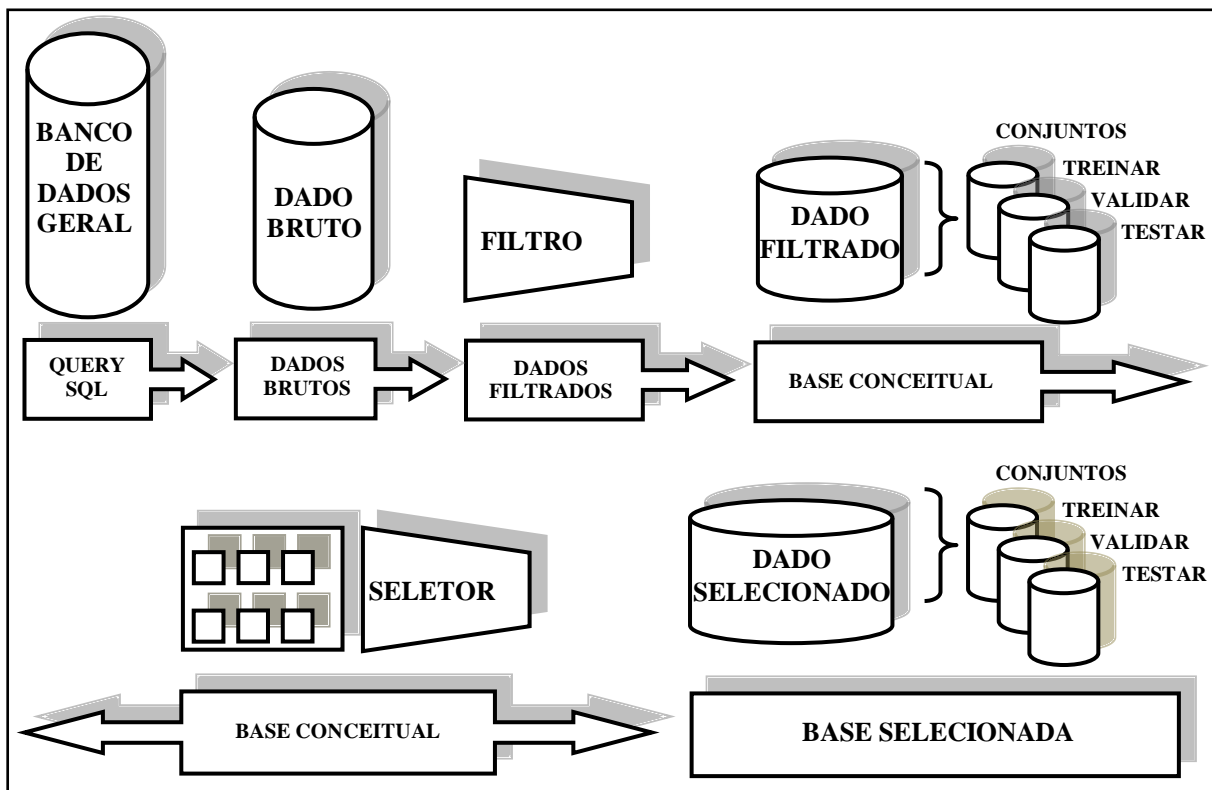
Base conceitual-

- Treinar: [493 amostras por 30 variáveis independentes]
- Validar: [493 amostras por 30 variáveis independentes]
- Testar: [493 amostras por 30 variáveis independentes]

Base selecionada-

- Treinar: [493 amostras por 4 variáveis independentes]
- Validar: [493 amostras por 4 variáveis independentes]
- Testar: [493 amostras por 4 variáveis independentes]

Em todos os casos a variável dependente é única e refere-se ao parâmetro de qualidade do produto que é a compressão mecânica a frio da pelota de minério de ferro detalhada anteriormente. A Figura 5.4 exemplifica todos os conjuntos de dados levantados e estratificados do problema.



**Figura 5.4** Visão geral da formação dos conjuntos de dados utilizados no treinamento, validação e testes dos Modelos de Predição.

A heurística proposta acima em relação aos conjuntos foi aplicada na etapa de construção dos Modelos de Predição, nas medidas de adequação e análise residual. A idéia principal é submeter às técnicas de aprendizagem (e conseqüentemente os Modelos de Predição) as mesmas condições de treinamento, validação e teste neste ambiente de pesquisa.

## 5.2 Resultado dos Modelos de Predição

Os resultados dos Modelos de Predição são analisados segundo três critérios conforme mencionados anteriormente e reforçados a seguir:

- Critério 1 – Estimativa de erro: Cálculo da estimativa do erro por substituição por valores RMS (*Root Mean Squared*) decorrentes dos conjuntos: Treinar, Validar e Testar. Relembrando que, os conjuntos Validar e Testar são amostras que não foram utilizadas na etapa de treinamento dos algoritmos;

- Critério 2 – Adequação do modelo: Esta adequação foi quantificada pelo coeficiente de determinação ( $R^2 / R_{\text{Ajustado}}^2$ ). Esse coeficiente determina o quanto o Modelo de Predição explica da variação do fenômeno em porcentagem da variação total;
- Critério 3 – Análise residual: A análise gráfica do erro normalizado aponta indícios de violações que podem ser cometidas pelos algoritmos. As minimizações dessas violações determinam a qualidade do Modelo de Predição desenvolvido pelos algoritmos de aprendizagem na etapa de treinamento.

Além disso, os parâmetros de ajuste dos algoritmos (grau de liberdade) foram manipulados no intuito de buscar os valores mais otimizados em relação aos três critérios acima mencionados. Os graus de liberdade manipulados possibilitaram determinar características dos Modelos de Predição num mesmo ambiente com dados reais de um processo industrial.

### **5.2.1 Resultado experimental das Técnicas de Regressão**

Os Modelos de Predição desenvolvidos pelas técnicas de regressão apresentaram os seguintes resultados conforme os critérios mencionados:

- Os modelos por Regressão Linear Múltipla e Máquina de Vetor de Relevância apresentaram resultados sem diferenças significativas em relação aos critérios um e dois de análise para trinta variáveis independentes (Tabela: 5.1 e 5.5). Sendo que, a análise residual apresentou baixa heterocedasticidade para o modelo RVM (Figura 5.9),
- Para quatro variáveis independentes, os modelos HAT e RVM não apresentaram diferenças significativa entre os valores RMS; entretanto, apresentaram redução significativa em relação aos modelos com trinta variáveis para os coeficientes de determinação o que não justifica a sua utilização (Tabela: 5.2 e 5.6),
- Os modelos com desempenho menos significativo foram desenvolvidos pela Máquina de Vetores de Suporte (Tabela: 5.3 e 5.4); o que se confirma, pelos valores consideravelmente baixos dos coeficientes de determinação e extremas heterocedasticidades apresentadas na análise residual (Figura 5.9),



- Os modelos desenvolvidos pela SVM apresentaram resultados mais significativos com quatro variáveis independente (Figura 5.10), o que aponta na direção que o algoritmo apresenta limitações com relação o aumento da dimensionalidade,
- Os resultados dos coeficientes de determinação dos modelos desenvolvidos pela SVM são ratificados pelo range apresentado em relação aos valores preditos (Figura 5.9 e 5.10). Esse range comportou-se dentro de uma faixa de valores mais estreita dos que os mesmos critérios observados nos modelos HAT e RVM.

Os resultados foram analisados conforme as tabelas e gráficos apresentados a seguir referentes aos modelos desenvolvidos pelas técnicas de regressão abordadas neste trabalho.

TABELA 5.1: Resultado do Modelo de Predição por Regressão Linear Múltipla com trinta variáveis independentes

HAT	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 30									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	29,3463	0,3877	0,3479	29,7922	0,3479	0,3056	30,5917	0,3226	0,2786
2	29,9783	0,3394	0,2965	31,3741	0,3281	0,2845	28,6724	0,3445	0,3019
3	29,6689	0,3151	0,2706	29,8788	0,3436	0,3010	30,1999	0,3183	0,2740
4	28,2977	0,4121	0,3739	30,7997	0,3504	0,3083	30,5120	0,3887	0,3490
5	27,9363	0,4395	0,4031	32,7293	0,3256	0,2818	29,6556	0,3909	0,3514
6	27,6565	0,4327	0,3959	30,5298	0,3928	0,3534	32,1521	0,3663	0,3251
7	27,5978	0,4208	0,3832	33,6191	0,3194	0,2753	30,0795	0,3752	0,3346
8	27,2938	0,3744	0,3338	32,4537	0,3306	0,2871	31,2145	0,3475	0,3051
9	29,4785	0,3592	0,3176	31,0816	0,3335	0,2902	29,1595	0,3405	0,2977
10	28,6353	0,4082	0,3698	30,5306	0,3686	0,3276	30,7156	0,3676	0,3265
MÍNIMO	27,2938	0,3151	0,2706	29,7922	0,3194	0,2753	28,6724	0,3183	0,2740
MÁXIMO	29,9783	0,4395	0,4031	33,6191	0,3928	0,3534	32,1521	0,3909	0,3514
MÉDIA	28,5889	0,3889	0,3492	31,2789	0,3441	0,3015	30,2953	0,3562	0,3144

TABELA 5.2: Resultado do Modelo de Predição por Regressão Linear Múltipla com quatro variáveis independentes

HAT	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 4									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	34,9254	0,1327	0,1256	33,8409	0,1645	0,1576	34,4214	0,1304	0,1233
2	34,6570	0,1171	0,1098	35,3567	0,1343	0,1272	32,9481	0,1391	0,1320
3	33,4578	0,1290	0,1218	35,1544	0,1292	0,1220	34,2230	0,1205	0,1133
4	33,6182	0,1702	0,1634	34,7412	0,1627	0,1554	34,5066	0,1641	0,1572
5	34,3757	0,1513	0,1444	34,8316	0,1501	0,1431	33,7629	0,1557	0,1487
6	33,3940	0,1729	0,1661	35,2888	0,1514	0,1444	34,3105	0,1632	0,1564
7	33,5130	0,1460	0,1390	35,9907	0,1366	0,1295	33,5280	0,1665	0,1597
8	31,7169	0,1553	0,1483	36,1709	0,1412	0,1342	34,9070	0,1327	0,1256
9	33,7588	0,1596	0,1527	35,8654	0,1399	0,1329	33,1877	0,1527	0,1458
10	34,7580	0,1290	0,1209	33,8712	0,1247	0,1176	34,1247	0,1452	0,1382
MÍNIMO	31,7169	0,1171	0,1098	33,8409	0,1247	0,1176	32,9481	0,1205	0,1133
MÁXIMO	34,9254	0,1729	0,1661	36,1709	0,1645	0,1576	34,9070	0,1665	0,1597
MÉDIA	33,8175	0,1463	0,1392	35,1112	0,1435	0,1364	33,9920	0,1470	0,1400

TABELA 5.3: Resultado do Modelo de Predição por Máquina de Vetor de Suporte com trinta variáveis independentes

SVM	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 30									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	39,2312	0,1261	0,0693	40,2187	0,1821	0,1290	39,7646	0,1246	0,0677
2	37,1371	0,0806	0,0208	39,5884	0,0914	0,0324	36,9786	0,0688	0,0084
3	34,7232	0,0206	0,0100	37,7479	0,0350	0,0100	37,0354	0,0076	0,0084
4	37,7268	0,0943	0,0356	37,7941	0,0948	0,0360	38,6024	0,0670	0,0100
5	39,2163	0,1351	0,0790	38,9399	0,1421	0,0864	40,7787	0,1496	0,0064
6	37,7043	0,1072	0,0493	39,6555	0,1144	0,0568	38,5005	0,1072	0,0944
7	36,4287	0,0678	0,0073	38,4551	0,0297	0,0100	36,1017	0,0243	0,0492
8	34,7545	0,0752	0,0152	39,4565	0,0416	0,0100	37,3010	0,0325	0,0100
9	36,3812	0,0505	0,0100	38,3390	0,0171	0,0100	36,0552	0,0568	0,0100
10	37,1487	0,0611	0,0100	37,3160	0,0416	0,0100	36,8647	0,0426	0,0100
MÍNIMO	34,7232	0,0206	0,0073	37,3160	0,0171	0,0100	36,0552	0,0076	0,0064
MÁXIMO	39,2312	0,1351	0,0790	40,2187	0,1821	0,1290	40,7787	0,1496	0,0944
MÉDIA	37,0452	0,0819	0,0307	38,7511	0,0790	0,0391	37,7983	0,0681	0,0275

TABELA 5.4: Resultado do Modelo de Predição por Máquina de Vetor de Suporte com quatro variáveis independentes

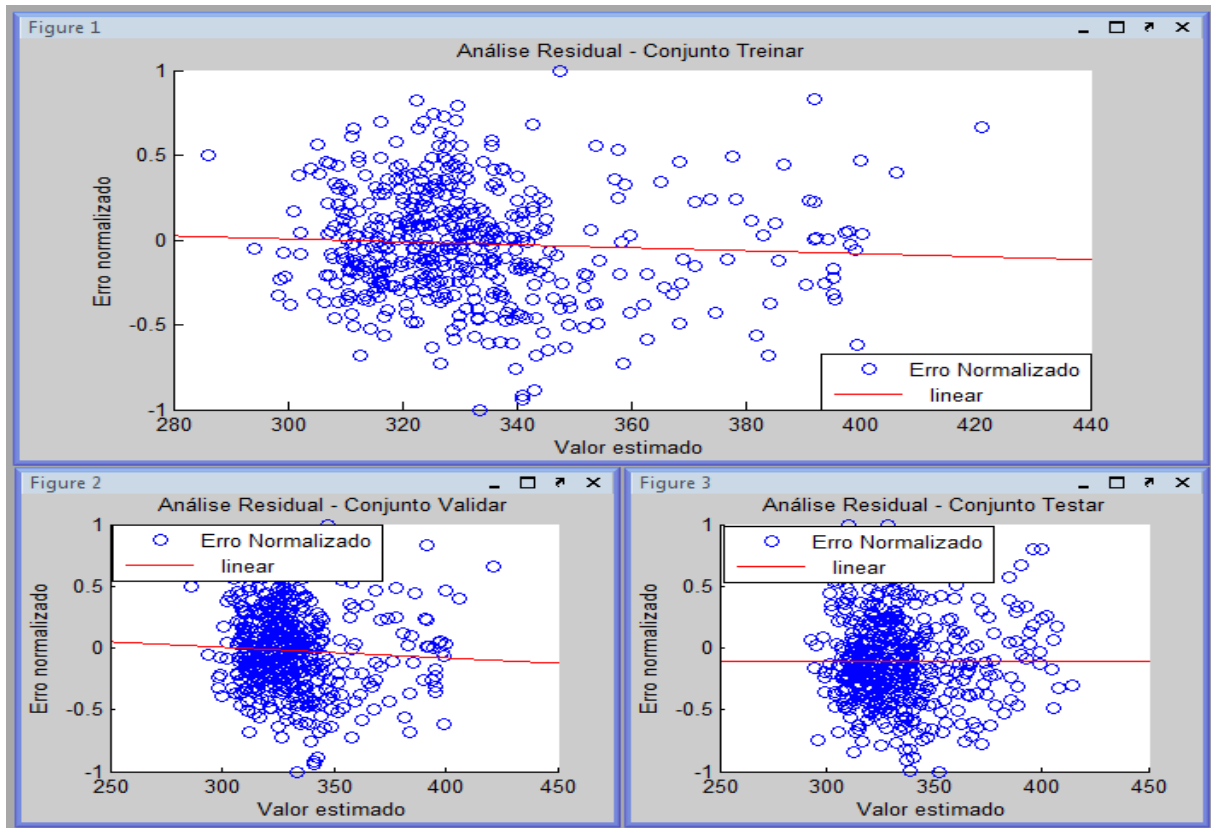
SVM	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 4									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	46,5740	0,2806	0,2747	48,3191	0,3260	0,3204	46,7906	0,2841	0,2782
2	42,1527	0,1902	0,1836	43,8331	0,2049	0,1984	40,8027	0,1888	0,1822
3	39,3135	0,1708	0,1640	42,0154	0,1941	0,1875	40,0665	0,1466	0,1396
4	48,7644	0,3135	0,3078	48,3519	0,3197	0,3142	48,2392	0,2952	0,2895
5	44,8515	0,2403	0,2341	43,9533	0,2558	0,2497	45,8941	0,2581	0,2520
6	44,6697	0,2729	0,2669	46,7003	0,2754	0,2695	45,0480	0,2755	0,2696
7	48,0809	0,3198	0,3142	47,5756	0,2797	0,2738	45,2776	0,2927	0,2869
8	36,2563	0,0980	0,0906	40,0730	0,0636	0,0559	37,8583	0,0540	0,0462
9	41,6726	0,2039	0,1973	41,4790	0,1548	0,1479	40,8619	0,2162	0,2098
10	38,7382	0,0960	0,0886	38,0305	0,0840	0,0765	37,4474	0,0826	0,0751
MÍNIMO	36,2563	0,0960	0,0886	38,0305	0,0636	0,0559	37,4474	0,0540	0,0462
MÁXIMO	48,7644	0,3198	0,3142	48,3519	0,3260	0,3204	48,2392	0,2952	0,2895
MÉDIA	43,1074	0,2186	0,2122	44,0331	0,2158	0,2094	42,8286	0,2094	0,2029

TABELA 5.5: Resultado do Modelo de Predição por Máquina de Vetor de Relevância com trinta variáveis independentes

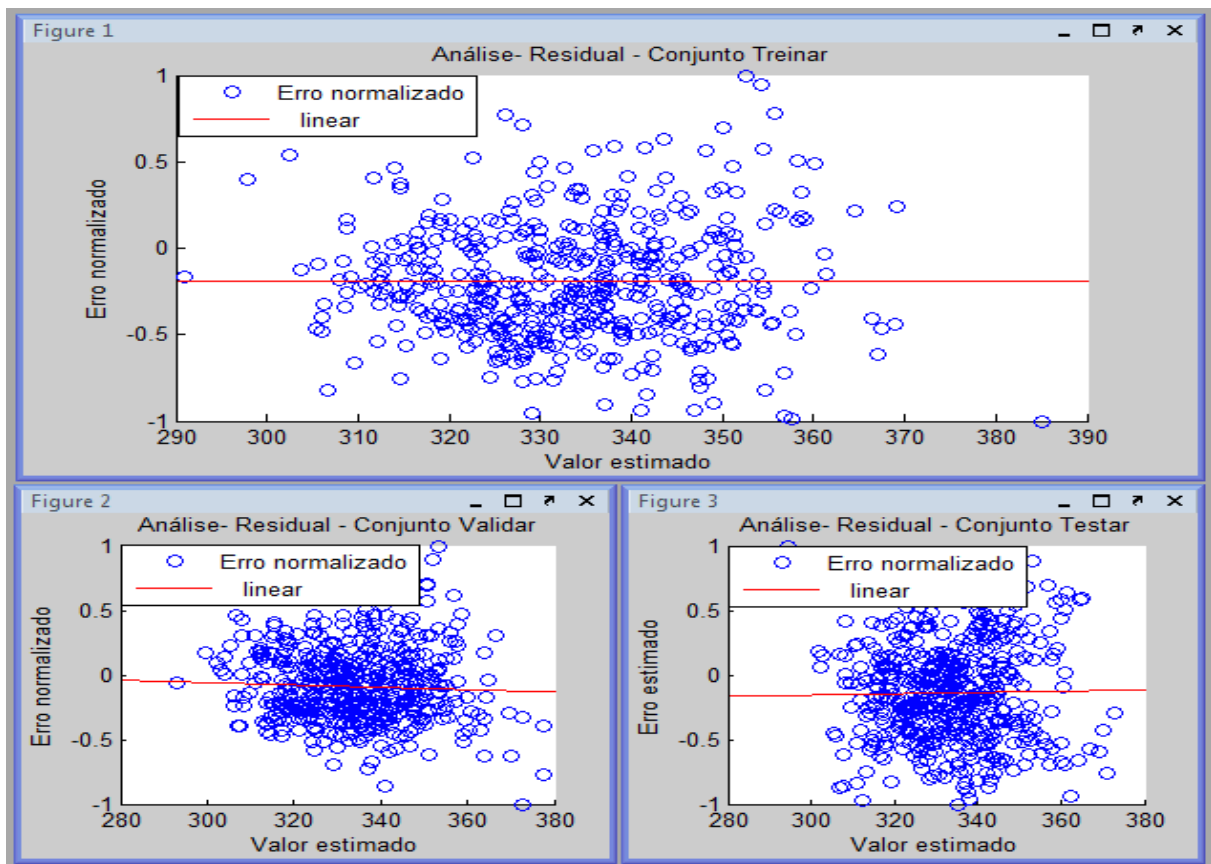
RVM	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 30									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	27,9068	0,4139	0,3759	31,3425	0,3412	0,2984	32,5402	0,3155	0,2710
2	27,6370	0,4035	0,3647	33,9819	0,2443	0,1952	31,7338	0,2780	0,2311
3	28,0092	0,3520	0,3100	33,7411	0,2476	0,1987	32,6025	0,2470	0,1981
4	26,7329	0,4462	0,4103	32,6849	0,2812	0,2346	33,4161	0,2928	0,2469
5	27,2074	0,4409	0,4046	33,3651	0,2664	0,2187	32,1448	0,2946	0,2488
6	26,4928	0,4520	0,4164	31,6040	0,3501	0,3079	32,2094	0,3149	0,2704
7	26,7185	0,4244	0,3871	34,0931	0,2869	0,2406	31,6535	0,3162	0,2718
8	25,6570	0,4127	0,3746	33,4746	0,2917	0,2457	32,9568	0,3100	0,2652
9	27,7701	0,3981	0,3590	32,3155	0,2824	0,2358	30,1608	0,2781	0,2312
10	28,6641	0,3705	0,3297	31,5920	0,2801	0,2334	30,5419	0,3006	0,2552
MÍNIMO	25,6570	0,3520	0,3100	31,3425	0,2443	0,1952	30,1608	0,2470	0,1981
MÁXIMO	28,6641	0,4520	0,4164	34,0931	0,3501	0,3079	33,4161	0,3162	0,2718
MÉDIA	27,2796	0,4114	0,3732	32,8195	0,2872	0,2409	31,9960	0,2948	0,2490

TABELA 5.6: Resultado do Modelo de Predição por Máquina de Vetor de Relevância com quatro variáveis independentes

RVM	TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 4									
VALIDAÇÃO CRUZADA	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	35,2293	0,1055	0,0982	34,4048	0,1279	0,1208	35,1142	0,0971	0,0897
2	34,7798	0,0997	0,0923	36,2414	0,1049	0,0976	33,4298	0,1112	0,1039
3	33,7653	0,1033	0,0959	35,6588	0,1000	0,0927	34,7423	0,1013	0,0940
4	35,1377	0,1468	0,1398	34,5781	0,1413	0,1343	33,8636	0,1370	0,1300
5	34,9262	0,1087	0,1014	35,2727	0,1171	0,1098	34,9889	0,1126	0,1054
6	33,8512	0,1389	0,1318	35,9128	0,1315	0,1243	34,6670	0,1398	0,1327
7	33,7298	0,1254	0,1182	36,9204	0,1208	0,1136	34,4512	0,1397	0,1326
8	32,1988	0,1173	0,1101	36,9414	0,1010	0,0936	35,4906	0,1087	0,1014
9	34,0884	0,1321	0,1250	36,0717	0,1161	0,1089	33,6019	0,1262	0,1190
10	35,2363	0,0936	0,0861	34,5964	0,0993	0,0919	34,1908	0,0952	0,0878
MÍNIMO	32,1988	0,0936	0,0861	34,4048	0,0993	0,0919	33,4298	0,0952	0,0878
MÁXIMO	35,2363	0,1468	0,1398	36,9414	0,1413	0,1343	35,4906	0,1398	0,1327
MÉDIA	34,2943	0,1171	0,1099	35,6599	0,1160	0,1088	34,4540	0,1169	0,1097

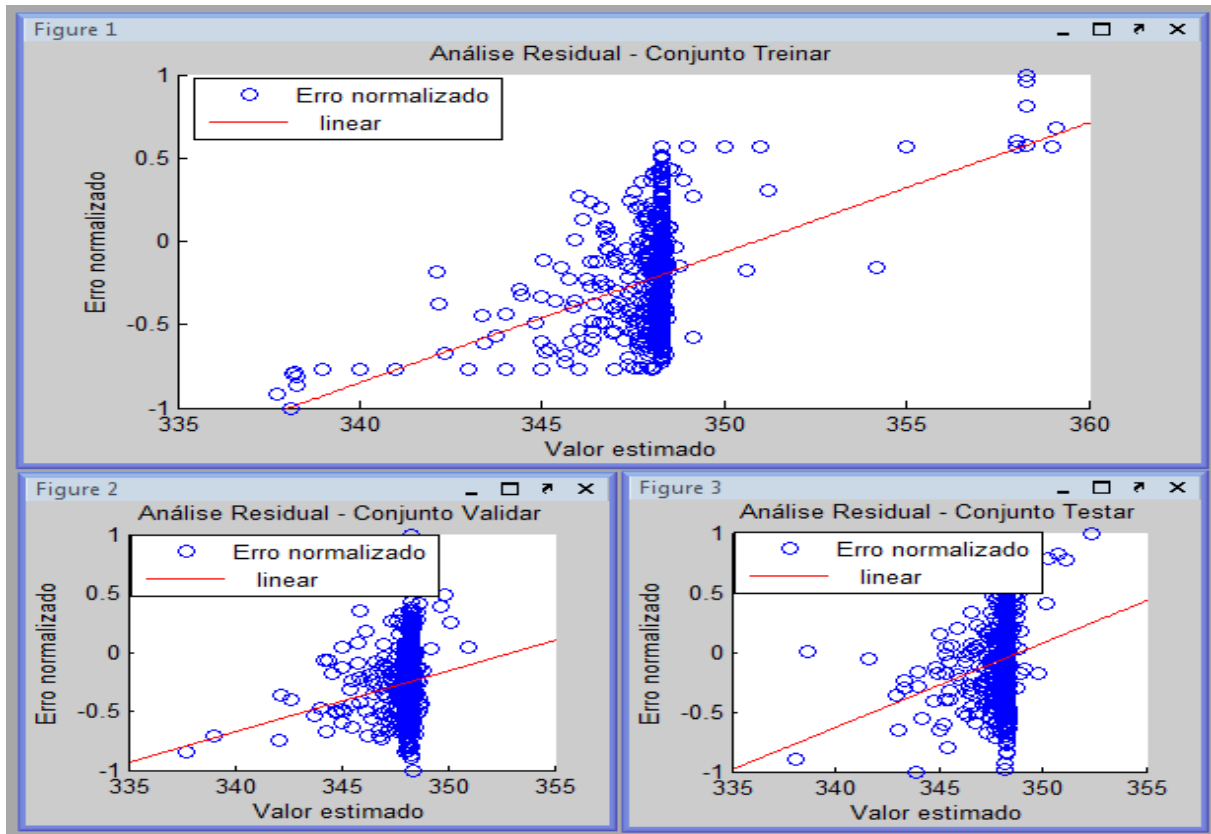


**Figura 5.5** - Análise Residual para Regressão Linear Múltipla com trinta variáveis

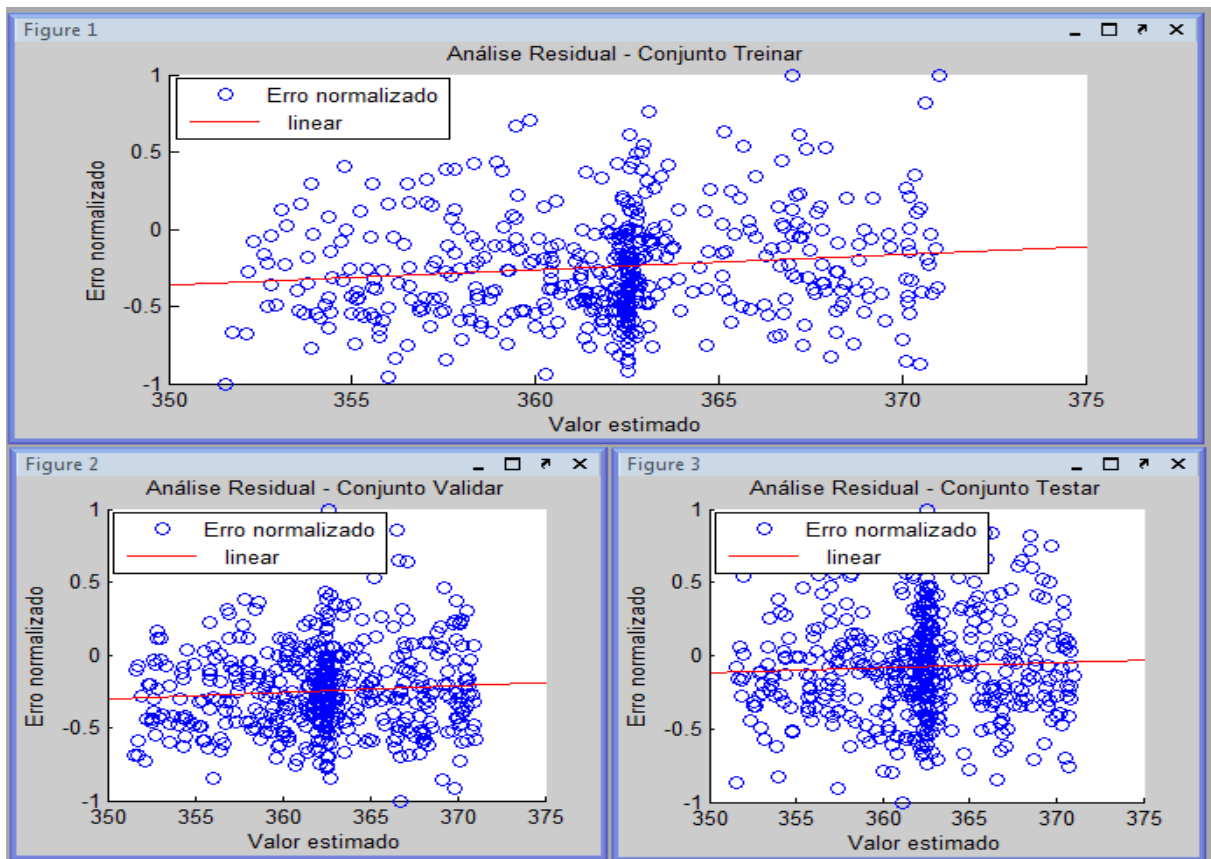


**Figura 5.6** - Análise Residual para Regressão Linear Múltipla com quatro variáveis

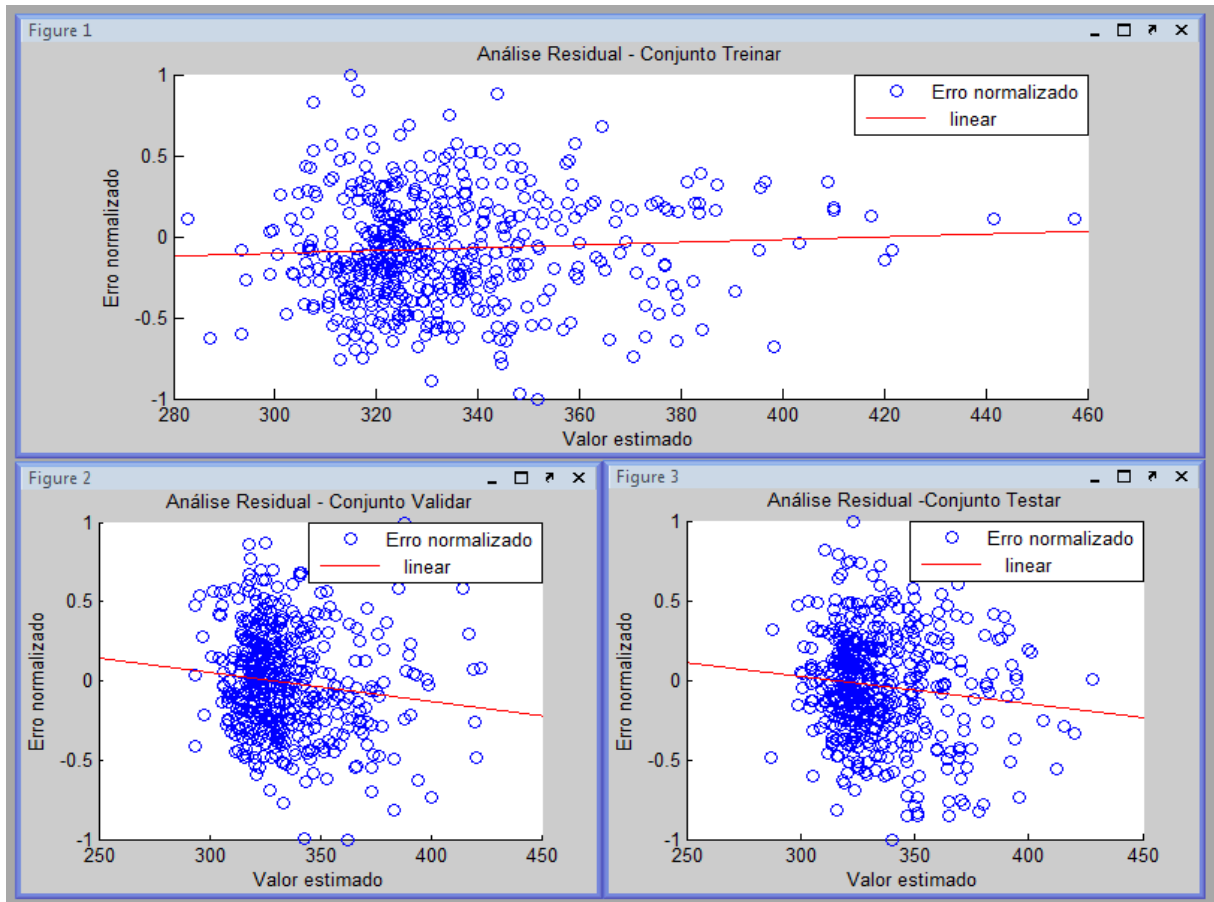




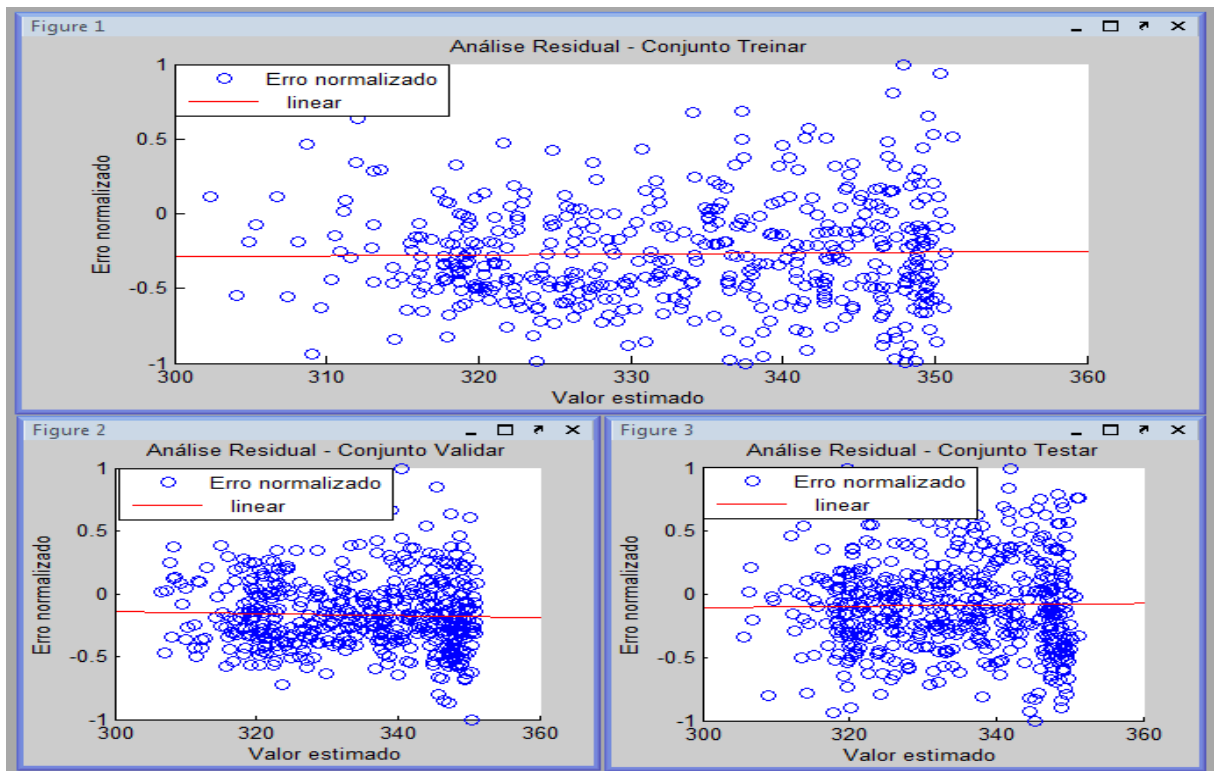
**Figura 5.7** - Análise Residual para Máquina de Vetores de Suporte com trinta variáveis



**Figura 5.8** - Análise Residual para Máquina de Vetores de Suporte com quatro variáveis



**Figura 5.9** - Análise Residual para Máquina de Vetores Relevância com trinta variáveis



**Figura 5.10** - Análise Residual para Máquina de Vetores Relevância com quatro variáveis

## 5.2.2 Resultado experimental da Técnica Neuro-Fuzzy

Os Modelos de Predição desenvolvidos pelo Sistema de Inferência Neuro-Fuzzy adaptativo apresentaram, conforme os critérios mencionados, os seguintes resultados:

- Menor valor RMS médio em relação aos demais modelos elaborados para trinta variáveis independentes (TABELA 5.7 e 5.8),
- Maior nível de adequação apontado pelo coeficiente de determinação para trinta variáveis independentes (TABELA 5.7 e 5.8),
- Baixo desempenho em relação aos critérios um e dois para a base com quatro variáveis independentes; o que não justifica também, a sua utilização (TABELA 5.9 e 5.10),
- Fenômeno de violação de critério pela análise residual denominado de dependência baseada no tempo. Esta violação é observada nos gráficos dos conjuntos Validar e Testar para trinta variáveis independentes (Figura 5.11).

Os resultados foram analisados conforme as tabelas e gráficos apresentados a seguir referentes aos modelos desenvolvidos pelas técnicas ANFIS abordada anteriormente.

TABELA 5.7: Resultado do Modelo de Predição por ANFIS com trinta variáveis independentes - Grupo 1 ao 5

ANFIS		TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 30										
VALIDAÇÃO CRUZADA	MF	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	3	25,1018	0,5520	0,5229	33,8436	0,4292	0,3921	34,3435	0,4039	0,3651
	5	21,9693	0,6570	0,6347	36,4687	0,4552	0,4198	35,0220	0,4285	0,3913
	7	18,2723	0,7645	0,7492	52,2104	0,5292	0,4986	43,5973	0,5392	0,5092
2	3	25,8460	0,5078	0,4758	34,3238	0,4057	0,3671	32,1280	0,4400	0,4079
	5	21,2105	0,6694	0,6479	38,0779	0,4320	0,3951	36,7712	0,4411	0,4048
	7	17,2480	0,7808	0,7666	75,7405	0,4811	0,4474	81,4790	0,4992	0,4667
3	3	24,9551	0,5147	0,4831	32,1085	0,4432	0,4070	32,4189	0,4498	0,4141
	5	22,3255	0,6119	0,5867	37,2604	0,5030	0,4707	35,8754	0,5088	0,4769
	7	18,3610	0,7368	0,7197	65,6364	0,5375	0,5074	64,2971	0,5053	0,4732
4	3	24,1684	0,5711	0,5433	32,6471	0,4272	0,3900	33,7086	0,4669	0,4323
	5	21,9200	0,6470	0,6241	36,1641	0,4694	0,4349	34,8723	0,4924	0,4594
	7	18,3692	0,7523	0,7362	40,0185	0,5010	0,4686	43,4524	0,4932	0,4603
5	3	24,1776	0,5802	0,5530	33,6105	0,3673	0,3262	30,1272	0,4718	0,4376
	5	21,6538	0,6632	0,6413	36,4299	0,3831	0,3430	33,3410	0,4776	0,4436
	7	18,5627	0,7520	0,7359	40,6492	0,4558	0,4205	37,4053	0,5248	0,4939
MÍNIMO		17,2480	0,5078	0,4758	32,1085	0,3673	0,3262	30,1272	0,4039	0,3651
MÁXIMO		25,8460	0,7808	0,7666	75,7405	0,5375	0,5074	81,4790	0,5392	0,5092
MÉDIA		21,6094	0,6507	0,6280	41,6793	0,4547	0,4192	40,5893	0,4762	0,4424

TABELA 5.8: Resultado do Modelo de Predição por ANFIS com trinta variáveis independentes - Grupo 6 ao 10

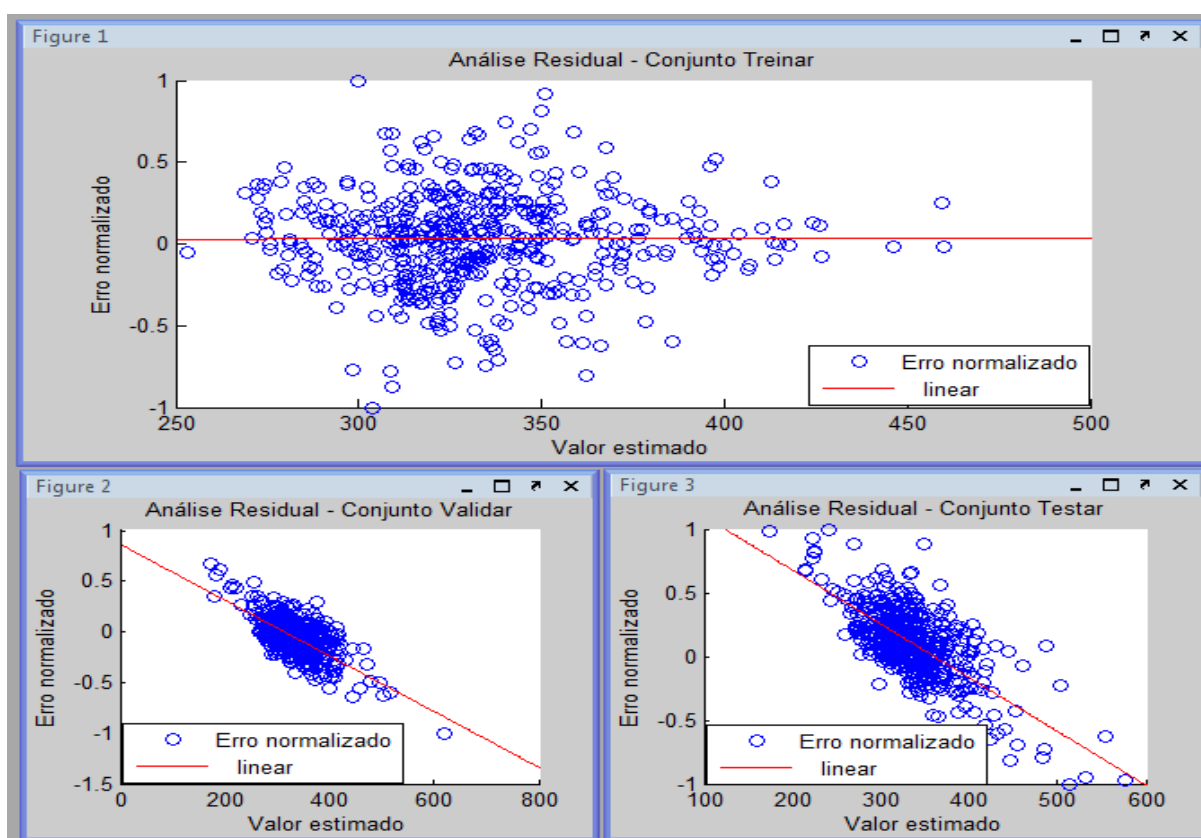
ANFIS		TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 30										
VALIDAÇÃO CRUZADA	MF	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
6	3	23,3658	0,5959	0,5687	33,0195	0,4103	0,3720	32,5051	0,4162	0,3783
	5	20,6663	0,6832	0,6626	32,9818	0,4623	0,4273	35,6400	0,4409	0,4046
	7	17,1929	0,7808	0,7665	42,9164	0,4780	0,4442	43,1848	0,4753	0,4413
7	3	24,2026	0,5547	0,5258	35,2726	0,3863	0,3464	31,1133	0,4484	0,4126
	5	19,9502	0,6974	0,6777	39,0496	0,4213	0,3838	35,3323	0,5000	0,4675
	7	17,5165	0,7665	0,7514	49,1690	0,4544	0,4189	50,1983	0,4403	0,4039
8	3	23,4764	0,5372	0,5072	34,6925	0,3965	0,3573	34,8948	0,4167	0,3788
	5	20,1506	0,6591	0,6369	38,9032	0,4726	0,4383	37,3063	0,4672	0,4326
	7	17,7284	0,7358	0,7187	58,5909	0,4774	0,4434	51,3054	0,5024	0,4701
9	3	25,6166	0,5162	0,4848	33,4266	0,4060	0,3674	32,1555	0,3871	0,3473
	5	22,8969	0,6134	0,5883	35,7550	0,4595	0,4244	34,4958	0,4456	0,4096
	7	19,6627	0,7149	0,6964	45,9176	0,4987	0,4661	42,9303	0,4946	0,4618
10	3	24,2928	0,5740	0,5463	35,1099	0,4508	0,4151	31,6710	0,4742	0,4401
	5	21,2230	0,6742	0,6530	37,3601	0,4677	0,4331	36,5449	0,5002	0,4678
	7	20,5778	0,6943	0,6745	39,5731	0,5023	0,4700	37,6632	0,4730	0,4388
MÍNIMO		17,1929	0,5162	0,4848	32,9818	0,3863	0,3464	31,1133	0,3871	0,3473
MÁXIMO		25,6166	0,7808	0,7665	58,5909	0,5023	0,4700	51,3054	0,5024	0,4701
MÉDIA		21,2346	0,6532	0,6306	39,4492	0,4496	0,4138	37,7961	0,4588	0,4237

TABELA 5.9: Resultado do Modelo de Predição por ANFIS com quatro variáveis independentes - Grupo 1 ao 5

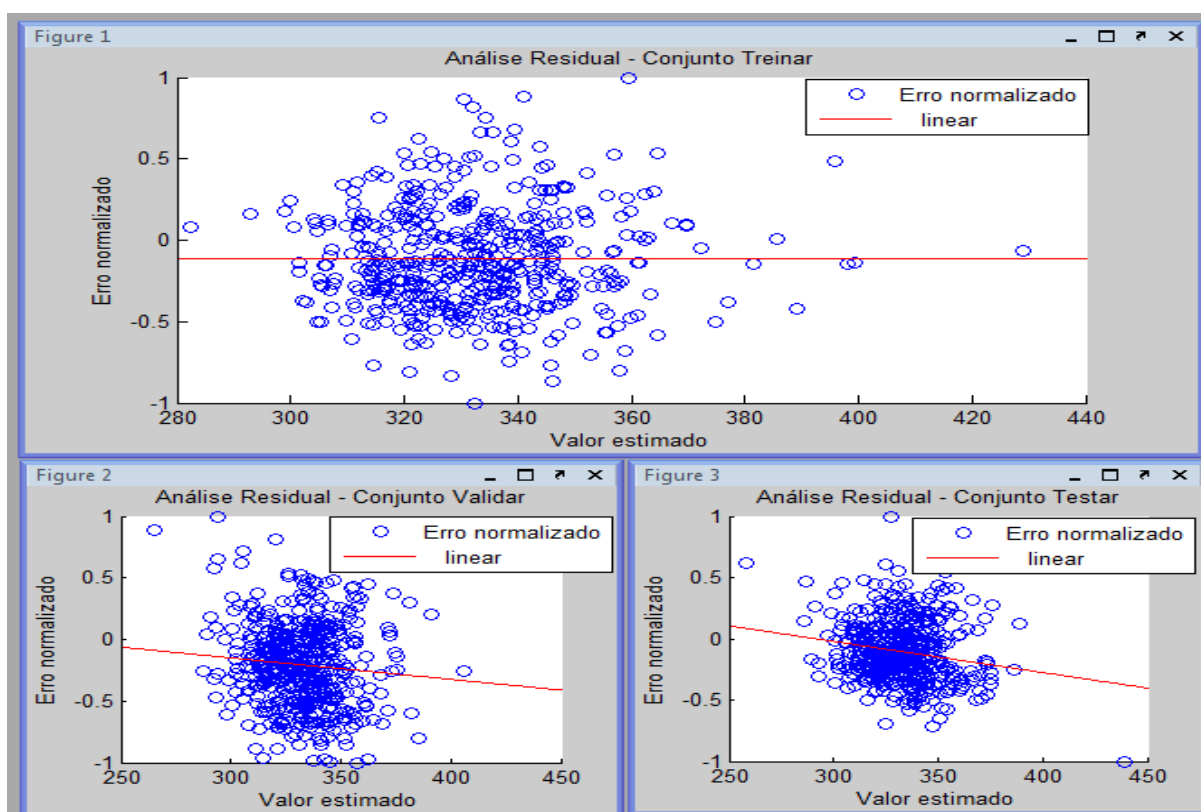
ANFIS		TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 4										
VALIDAÇÃO CRUZADA	MF	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
1	3	33,4748	0,2031	0,1514	34,6222	0,2259	0,1756	34,9219	0,1914	0,1176
	5	32,5948	0,2444	0,1953	41,3313	0,3282	0,2846	37,2968	0,2522	0,2036
	7	32,0956	0,2673	0,2128	37,6536	0,3133	0,2687	35,3994	0,2380	0,1886
2	3	33,6138	0,1692	0,1153	34,7536	0,1664	0,1123	33,0401	0,1612	0,1067
	5	33,5508	0,1723	0,1186	34,9365	0,1684	0,1144	33,3201	0,1674	0,1133
	7	32,7218	0,2124	0,1612	40,0253	0,2700	0,2226	34,4979	0,2063	0,1547
3	3	32,7070	0,1676	0,1135	36,0081	0,1592	0,1046	34,8263	0,1454	0,0899
	5	31,6249	0,2218	0,1712	36,7946	0,1954	0,1432	35,3384	0,1694	0,1155
	7	30,9329	0,2552	0,2068	35,6846	0,2111	0,1599	36,9566	0,2510	0,2024
4	3	32,8640	0,2070	0,1555	34,1289	0,1800	0,1268	34,7857	0,1871	0,1343
	5	31,8663	0,2540	0,2056	35,4946	0,2252	0,1749	34,3064	0,2344	0,1847
	7	30,7038	0,3069	0,2619	35,8339	0,2459	0,1969	35,0631	0,2475	0,1987
5	3	33,3623	0,2006	0,1487	35,8686	0,1960	0,1438	34,7758	0,1858	0,1329
	5	32,6439	0,2337	0,1840	35,6982	0,2158	0,1649	35,0653	0,2534	0,2050
	7	31,9647	0,2660	0,2183	36,4965	0,2518	0,2032	36,9832	0,2941	0,2482
MÍNIMO		30,7038	0,1676	0,1135	34,1289	0,1592	0,1046	33,0401	0,1454	0,0899
MÁXIMO		33,6138	0,3069	0,2619	41,3313	0,3282	0,2846	37,2968	0,2941	0,2482
MÉDIA		32,4481	0,2254	0,1747	36,3554	0,2235	0,1731	35,1051	0,2123	0,1597

TABELA 5.10: Resultado do Modelo de Predição por ANFIS com quatro variáveis independentes - Grupo 6 ao 10

ANFIS		TREINAR			VALIDAR			TESTAR		
VARIÁVEL: 4										
VALIDAÇÃO CRUZADA	MF	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>	RMS	R <sup>2</sup>	R <sup>2</sup> <sub>A</sub>
6	3	31,8663	0,2467	0,1978	36,8057	0,2008	0,1489	34,9900	0,2130	0,1619
	5	30,9766	0,2880	0,2418	37,0517	0,2397	0,1904	36,0653	0,2274	0,1772
	7	30,7657	0,2978	0,2521	38,3998	0,2370	0,1875	36,8385	0,2516	0,2030
7	3	32,5430	0,1945	0,1422	35,4775	0,1724	0,1186	33,7080	0,2028	0,1510
	5	31,9454	0,2240	0,1736	37,5971	0,2572	0,2089	35,1636	0,2302	0,1802
	7	31,2173	0,2583	0,2102	36,3029	0,2453	0,1963	35,2452	0,2636	0,2157
8	3	30,8199	0,2023	0,1505	37,1645	0,1867	0,1339	36,6349	0,1822	0,2191
	5	30,1718	0,2355	0,1859	37,4901	0,2057	0,1541	36,5372	0,2010	0,1491
	7	29,4553	0,2708	0,2235	37,5451	0,2405	0,1911	36,0932	0,2344	0,1847
9	3	31,8645	0,2500	0,2439	35,7774	0,1896	0,1829	33,1453	0,1908	0,1841
	5	31,9805	0,2458	0,1968	36,5184	0,2005	0,1486	34,2890	0,1921	0,1397
	7	30,6463	0,3069	0,3013	36,3172	0,2433	0,2371	33,6602	0,2113	0,2049
10	3	33,3376	0,1978	0,1912	35,1947	0,2243	0,2179	36,8131	0,2240	0,2176
	5	32,3717	0,2432	0,2370	35,0246	0,2599	0,2538	36,3123	0,2588	0,2527
	7	32,1223	0,2544	0,2483	35,2967	0,2686	0,2626	37,6793	0,2894	0,2836
MÍNIMO		29,4553	0,1945	0,1422	35,0246	0,1724	0,1186	33,1453	0,1822	0,1397
MÁXIMO		33,3376	0,3069	0,3013	38,3998	0,2686	0,2626	37,6793	0,2894	0,2836
MÉDIA		31,4723	0,2477	0,2131	36,5309	0,2248	0,1888	35,5450	0,2248	0,1950



**Figura 5.11** - Análise Residual para ANFIS com trinta variáveis



**Figura 5.12** - Análise Residual para ANFIS com quatro variáveis



### 5.3 Conclusões

A pesquisa empírica das Técnicas de Regressão e Neuro-*Fuzzy* que resultaram em todos os Modelos de Predição apresentados anteriormente serviram de subsídio para a análise comparativa das mesmas. O julgamento dos Modelos de Predição para análise comparativa dos métodos citados somente foi viável com a utilização de critérios ([Critério Um](#) / [Critério Dois](#) / [Critério Três](#)) bem definidos como os utilizados neste trabalho.

Outros pontos que podem ser ressaltados são:

- A utilização do algoritmo de seleção de características (SFS) que determinou as variáveis selecionadas não apresentou resultados significativos em relação a base de dados conceitual do problema,
- Não houve restrições em relação ao número de vetores para as máquinas de Kernel esparsas. A manipulação dos parâmetros dos algoritmos possibilitou a obtenção de qualquer ajuste para o número de vetores do modelo de saída,
- A heurística de dividir as amostras em subconjuntos: Treinar, Validar e Testar se apresentou significativamente interessante para o diagnóstico dos modelos,
- A heurística de redistribuir as amostras aleatoriamente nos conjuntos apresentou relativa uniformidade nos resultados repetidos dez vezes conforme exposto nas tabelas dos modelos.
- O coeficiente de determinação quantificou de maneira satisfatória a acurácia dos valores preditos pelos modelos de predição e salientou a capacidade de predição das diferentes técnicas abordadas neste trabalho,
- A análise residual expôs a qualidade dos valores preditos pelos modelos quando também submetidos a amostras não utilizadas na etapa de treinamento dos algoritmos.

Com isso, o Modelo de Predição que apresentou o desempenho mais significativo em relação aos critérios estabelecidos foi o desenvolvido pela técnica ANFIS. Por análise comparativa, o modelo de predição ANFIS conseguiu explicar, através do coeficiente de determinação calculado, até 78.08% da variação total da variável dependente.

## 5.4 Trabalhos futuros

Este trabalho tem por perspectiva a implantação de um ambiente de pesquisa e desenvolvimento de Modelos de Predição, situado no nível Três da Pirâmide da Automação ([Figura 2.6](#)), destinado aos seguintes propósitos:

- Descrição dos dados do sistema: especialistas podem usar esse ambiente para análise da descrição dos dados;
- Estimativas de parâmetros: determinar os melhores ajustes para otimização dos sistemas;
- Predição e estimativa: Aplicações que envolvam a estimativa da variável dependente do sistema.
- Controle: Modelos de Predição podem ser usados no controle das variáveis independentes do sistema.

Esse ambiente proporcionaria a investigação, desenvolvimento e implantação de técnicas da Inteligência Computacional destinadas à elaboração de Modelos de Predição lineares e até não-lineares podendo ser destinados inclusive ao controle do processo. Tal ambiente provê subsídios que podem perfazer todos os níveis do processo constituindo um moderno Sistema de Controle em Automação ([Figura 3.3](#)).

O processo de identificação dos Modelos de Predição ([Figura 3.1](#)) fica organizado e automatizado num ambiente de pesquisa como o citado acima e a sua implantação traria os benefícios esperados dos sistemas otimizados de controle. Todas as etapas de elaboração dos Modelos de Predição descritas neste trabalho estariam automaticamente presentes neste ambiente e a disposição dos especialistas do domínio do negócio.

Assim como os testes com outras funções canônicas de *Kernel* aplicadas aos algoritmos de aprendizagem, a perspectiva de investigação de outros modelos mais complexos e até não-lineares (Ver: [Item Um](#) / [Item Dois](#)) poderiam ser investigados neste ambiente. Além de prover um ambiente de simulação e testes para novos protótipos, possibilitaria também um ambiente de treinamento para os operadores de processos já direcionados a uma cultura de aperfeiçoamento dos controles da planta industrial.

Como mencionado anteriormente, há exemplo da variável dependente utilizada neste trabalho, uma série de outras variáveis igualmente complexas, na ordem de dezenas, são potenciais candidatas a pesquisa na busca por conhecimento. Esse ambiente proporcionaria a busca continuada por processos mais otimizados e conseqüentemente produtos mais competitivos e com alta qualidade assegurada em todas as etapas.

## Referências Bibliográficas

NATALE, Ferdinando. *Automação Industrial*. 6° ed. – São Paulo: Editora Érica, 2007.

FRICKÉ, Martin. *The Knowledge Pyramid: A Critique of the DIKW Hierarchy*. School of Information Resources and Library Science, The University of Arizona. Tucson, USA. 2007.

ACKOFF, Russell Lincoln. *From data to wisdom*. Journal of Applied Systems Analysis. Volume 16, 1989.

CAMPOS, Mario Massa de; SAITO, Kaku. *Sistemas Inteligentes em Controle e Automação de Processos*. Rio de Janeiro: Editora Ciência Moderna Ltda., 2004.

MORAES, Cícero Couto de; CASTRUCCI, Plínio de Lauro. *Engenharia de Automação Industrial – 2° edição – Rio de Janeiro: LTC, 2007.*

ACKOFF, Russell Lincoln. *Why few organizations adopt Systems Thinking*. 2006.

GRANA, Tom; OKE, Ameya; DENISON, Dan. *Cummins utilizes large scale automated data analysis for automotive diesel engine development*. MATLAB Virtual Conference. 2009.

HEY, Jonathan. *The Data, Information, Knowledge, Wisdom Chain: The Metaphorical link*. Berkeley Expert Systems Technology Lab. Berkeley, University California. 2004.

BEGA, Egídio Alberto. *Instrumentação Industrial*. 2° edição. Rio de Janeiro: Interciência: Instituto Brasileiro de Petróleo e Gás, 2006.

STAIR, Ralph M.; REYNOLDS, George W. *Principles of Information Systems – 4° edição – Rio de Janeiro: LTC, 2004.*

SEBORG, Dale E. *A Perspective on Advanced Strategies for Process Control (Revisited)*. Department of Chemical Engineering, University of California, Santa Barbara. 2000.

FISKE, Tom. *Improving Agility, Performance, and Profitability with MPC: A Decade of Success*. ARC Advisory Group. Dedham, USA. 2006.

WAZLAWICK, Raul Sidnei. *Metodologia de Pesquisa para Ciência da Computação*. Rio de Janeiro: Elsevier, 2008.

NIKOLAOU, Michael. *Model Predictive Controllers: A Critical Synthesis of Theory and Industrial Needs*. Department of Chemical Engineering. University of Houston. 2007.

MEYER, Kurt. *Pelletizing of Iron Ores*. Department of Chemical Engineering. University of Frankfurt. 1980.

ZHU, Yucui. *Multivariable System Identification for Process Control*. Measurement and Control Group, Department of Electrical Engineering, Eindhoven University of Technology, Netherlands. 2001.

SEBER, George A. Frederick and WILD, Christopher John. *Nonlinear regression*. Department of Mathematics and Statistics. University of Auckland. Auckland, New Zealand. 1989.

MONTGOMERY, Douglas C. and RUNGER, George C. *Applied Statistics and Probability for Engineers*. Arizona State University. 2nd edition. 2003.

MONTGOMERY, Douglas C. et al. *Introduction to Linear Regression Analysis*. American Statistical Association. Arizona State University. 3rd edition. 2001.

RAWLINGS, John O. et al. *Applied Regression Analysis: A Research Tool*. Department of Statistics North Carolina State University. 2nd Edition. 1998.

WEISBERG, Sanford. *Applied Linear Regression*. University of Minnesota. School of Statistics Minneapolis, Minnesota. 3rd Edition. 2005.

DUDA, Richard O. et al. *Pattern Classification*. Department of Electrical Engineering at San Jose State University. San Jose, California. 2001.

HAIR, JOSEPH F. et al. *Multivariate Data Analysis*. Louisiana State University. 7<sup>o</sup> edição. 2009.

BISHOP, Christopher M. *Pattern Recognition and Machine Learning*. Microsoft Research Ltd. Cambridge. U.K. 2007.

RAKOTOMAMONJY, A. et al. SVM and Kernel Methods Matlab Toolbox. Perception Systèmes et Information, INSA de Rouen, Rouen, France. 2008. Disponível em: <http://asi.insa-rouen.fr/enseignants/~arakotom/toolbox/index.html>. Acesso em: 10 de Agosto de 2009.

TIPPING, Michael E. *The Relevance Vector Machine*. Microsoft Research Ltd. Cambridge. U.K. 2000.

TIPPING, Michael E. *Sparse Bayesian learning and the relevance vector machine*. Journal of Machine Learning Research. 2001. Disponível em: <http://www.miketipping.com>. Acesso em: 20 de Maio de 2009.

BISHOP, C. M. and M. E. TIPPING. *Bayesian regression and classification*. Advances in Learning Theory: Methods, Models and Applications. Amsterdam, 2003.

TIPPING, Michael E. *Bayesian inference: An introduction to principles and practice in machine learning*. Advanced Lectures on Machine Learning. Microsoft Research Ltd. Cambridge. U.K. 2006.

ZADEH, L. A. *Fuzzy Sets*. Department of Electrical Engineering and Electronics Research Laboratory – University of California, Berkeley, California. 1965.

JANG, J.S.R. et al. *Neuro-Fuzzy and soft computing*. Upper Saddle River, NJ: Prentice Hall. 1997.

REZENDE, Solange O. *Sistemas Inteligentes - Fundamentos e Aplicações*. Editora Manole, 2003.

PASSINO, Kevin M. and Yurkovick, Stephen. *Fuzzy Control*. Department of Electrical Engineering. The Ohio State University. Columbus, Ohio. July.1997.

JANG, J.S.R. *ANFIS: adaptive-network-based fuzzy inference system*. IEEE Transactions on Systems, Man and Cybernetics, 1993.

TAKAGI, T. and SUGENO, M. *Fuzzy identification of systems and its applications to modeling and control*. IEEE Transactions on Systems, Man and Cybernetics, v. 15, n. 1, p. 116-132, 1985.

JATZEN, Jan. *Neurofuzzy Modelling*. Technical University of Denmark, Department of Automation, Bldg 326, DK-2800 Lyngby, DENMARK. October 1998.

CHATTERJEE, S., and A. S. Hadi. "Influential Observations, High Leverage Points, and Outliers in Linear Regression." *Statistical Science*. vol. 1, 1986, pp. 379–416. Disponível em: <http://www.mathworks.com/access/helpdesk/help/toolbox/stats/regress.html>. Acesso em: 10 de Agosto de 2009.

BARBETTA, Pedro Alberto et al. *Estatística para cursos de Engenharia e Informática*. Editora Atlas, São Paulo. 2º edição. 2009.

ISO – International Standard: Iron ore pellets – Determination of crushing strength. Reference Number: ISO4700:1996(E).

BOWERMAN, Bruce & O'CONNELL, Richard T. *Applied Statistics – Improving Business Processes*. University Miami. 1997.

BENSCH, Michael et al. *Feature Selection for High-Dimensional Industrial Data*. European Symposium on Artificial Neural Networks. Bruges, Belgium. April, 2005.

# Apêndices:

## Apêndice A: Análise dos Dados

Apêndice A-T1: Variável de processo

ESTATÍSTICA	LEVANTAMENTO DE DADOS						
	PELOTAMENTO						
	GRANULOMETRIA						
	F1	F2	F3	F4	F5	F6	F7
MÍNIMO	9.00	27.00	21.00	733.00	1058.00	217.00	72.00
MÁXIMO	516.00	828.00	852.00	4837.00	4728.00	2437.00	1640.00
MEDIA	67.42	186.02	296.35	3599.88	3472.67	1173.15	550.70
MEDIANA	58.00	176.00	290.00	3680.00	3515.50	1194.50	542.00
DESVIO	39.18	87.03	109.48	541.79	413.81	185.26	155.86

CARACTERÍSTICA	1	2	3	4	5	6	7
----------------	---	---	---	---	---	---	---

30 VARIÁVEIS

1	2	3
---	---	---

4 VARIÁVEIS

1
---



ESTATÍSTICA	LEVANTAMENTO DE DADOS				
	FORNO				
	CAIXA 1 (TEMPERATURA)	CAIXA 2 (TEMPERATURA)	CAIXA 3 (TEMPERATURA)	CAIXA 4 (TEMPERATURA)	CAIXA 5 (TEMPERATURA)
	°C	°C	°C	°C	°C
MÍNIMO	125.16	107.73	297.88	267.17	115.78
MÁXIMO	239.17	284.56	630.86	415.84	228.74
MEDIA	160.83	149.16	410.88	360.60	189.52
MEDIANA	159.01	144.12	404.09	355.97	188.68
DESVIO	13.50	24.32	47.81	21.55	15.86

CARACTERÍSTICA	8	9	10	11	12
----------------	---	---	----	----	----

30 VARIÁVEIS	4	5	6	7	8
--------------	---	---	---	---	---

4 VARIÁVEIS
-------------

2
---

ESTATÍSTICA	LEVANTAMENTO DE DADOS							
	FORNO							
	CAIXA DE VENTO (PRESSÃO)			COIFA (PRESSÃO)			COIFA (TEMPERATURA)	
	mmH <sub>2</sub> O	mmH <sub>2</sub> O	mmH <sub>2</sub> O	mmH <sub>2</sub> O	mmH <sub>2</sub> O	mmH <sub>2</sub> O	°C	°C
MÍNIMO	54.31	73.27	168.44	-49.79	-52.49	-19.93	470.42	178.41
MÁXIMO	170.43	224.02	332.26	-18.71	-19.17	5.33	1028.26	378.53
MEDIA	129.91	161.31	281.26	-36.74	-40.41	-10.23	913.46	244.32
MEDIANA	130.19	160.95	281.96	-37.15	-40.98	-10.40	948.58	240.10
DESVIO	9.99	14.90	21.28	5.46	6.33	4.03	115.37	29.55

CARACTERÍSTICA	18	19	20	21	22	23	24	25
----------------	----	----	----	----	----	----	----	----

30 VARIÁVEIS	9	10	11	12	13	14	15
--------------	---	----	----	----	----	----	----

4 VARIÁVEIS
-------------

ESTATÍSTICA	LEVANTAMENTO DE DADOS											
	FORNO											
	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	GRUPO 6	GRUPO 7	GRUPO 8	GRUPO 9	GRUPO 10	VENTILADOR (TEMPERATURA)	
	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	MÉDIA	°C	°C
MÍNIMO	609.91	883.09	1073.55	1128.47	1144.03	1060.16	1139.57	1033.36	825.76	1232.27	20.48	17.83
MÁXIMO	1029.67	1115.39	1236.97	1310.87	1344.53	1392.14	1397.33	1383.58	1387.52	1381.06	49.61	224.02
MEDIA	929.35	1069.54	1172.74	1269.34	1324.40	1286.65	1320.93	1315.21	1328.45	1334.74	33.49	162.31
MEDIANA	937.60	1074.91	1174.86	1273.46	1325.28	1280.40	1317.11	1327.82	1330.48	1330.24	33.51	160.95
DESVIO	65.97	16.42	9.53	12.16	12.18	35.55	28.16	35.68	42.67	17.83	4.16	13.52

CARACTERÍSTICA	25	30	33	36	39	42	45	48	51	54	46	47
----------------	----	----	----	----	----	----	----	----	----	----	----	----

30 VARIÁVEIS	16	17	18	19	20	21	22	23	24	25	26
--------------	----	----	----	----	----	----	----	----	----	----	----

4 VARIÁVEIS
-------------

3
---

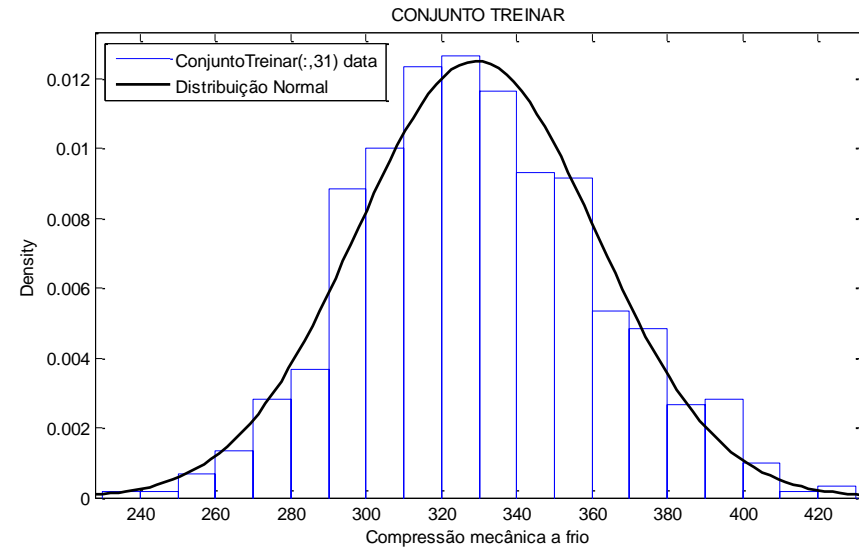
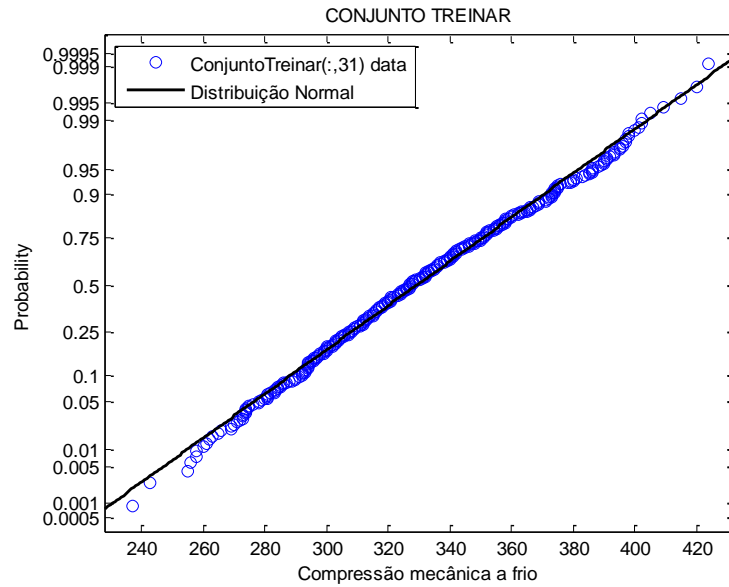
ESTATÍSTICA	LEVANTAMENTO DE DADOS				COMPRESSÃO
	GRELHA				
	VELOCIDADE	CAPACIDADE TÉRMICA	ALTURA CAMADA	PRODUÇÃO	
	m/min	KCAL/TMS	Cm	TMS/m <sup>2</sup>	
MÍNIMO	3.30	91.28	42.60	10.99	231.00
MÁXIMO	6.21	639.00	46.89	28.81	483.00
MEDIA	5.62	264.82	44.52	24.50	336.31
MEDIANA	5.71	251.17	44.50	25.15	332.00
DESVIO	0.39	36.39	0.64	2.54	37.85

CARACTERÍSTICA	48	49	50	51
----------------	----	----	----	----

30 VARIÁVEIS	27	28	29	30
--------------	----	----	----	----

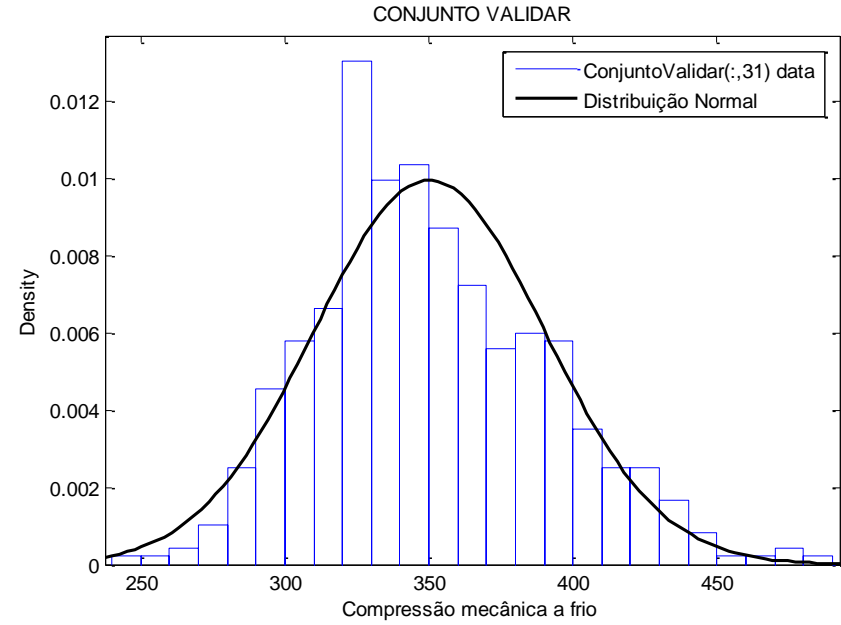
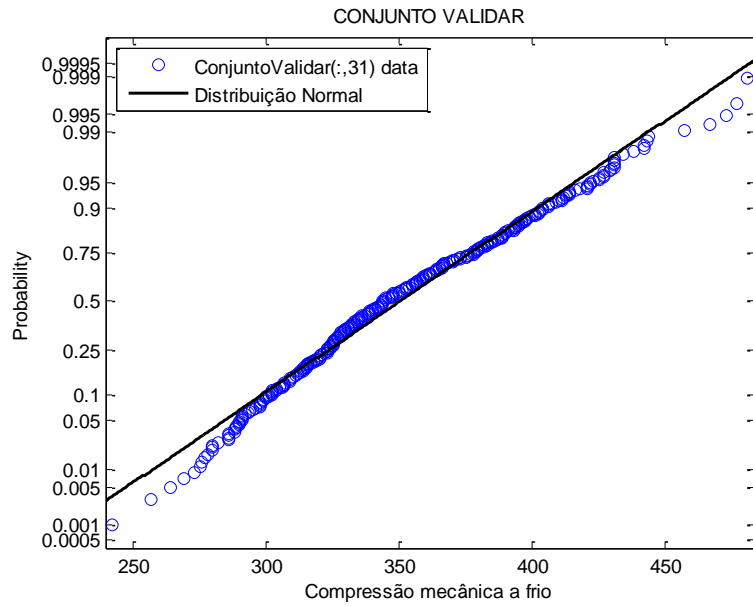
4 VARIÁVEIS	4
-------------	---

### Apêndice A-G1: Distribuição da Probabilidade: Variável Dependente (Conjunto Treinar)



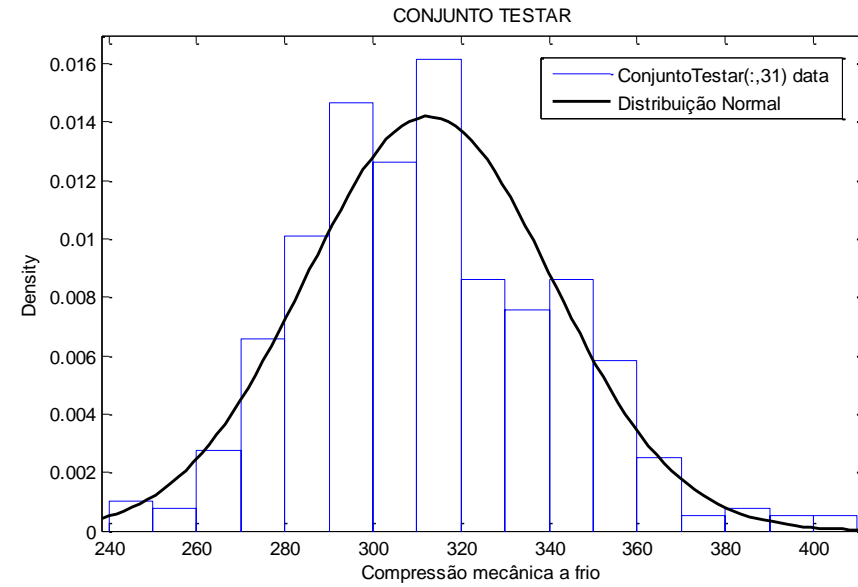
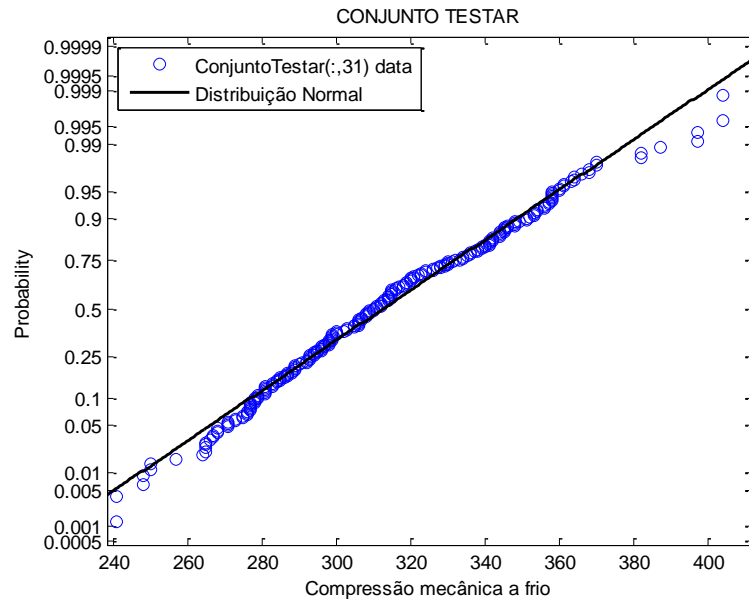
ESTATÍSTICA	COMPRESSÃO
	Kgf/PELOTA
MÍNIMO	237.00
MÁXIMO	424.00
MEDIA	329.30
MEDIANA	327.00
DESVIO	31.90

Apêndice A-G2: Distribuição da Probabilidade: Variável Dependente (Conjunto Validar)



ESTATÍSTICA	COMPRESSÃO
	Kgf/PELOTA
MÍNIMO	240.00
MÁXIMO	490.00
MEDIA	365.00
MEDIANA	365.00
DESVIO	72.65

### Apêndice A-G3: Distribuição da Probabilidade: Variável Dependente (Conjunto Testar)



ESTATÍSTICA	COMPRESSÃO
	Kgf/PELOTA
MÍNIMO	240.00
MÁXIMO	410.00
MEDIA	325.00
MEDIANA	325.00
DESVIO	49.61

## Apêndice B: Resultado do algoritmo de seleção de característica

### TABELA B-T1: Resultado do seletor de característica com kernel RVM

SFS		
SELEÇÃO	SEQÜÊNCIA DAS CARACTERÍSTICAS (Variável Independente)	RMS
1	2	18.0000
2	2 6	8.3205
3	2 6 18	8.1321
4	2 6 18 27	8.1318
5	2 6 18 27 29	8.1326
6	2 6 18 27 29 25	8.1401
7	2 6 18 27 29 25 30	8.1465
8	2 6 18 27 29 25 30 12	8.1490
9	2 6 18 27 29 25 30 12 15	8.1503
10	2 6 18 27 29 25 30 12 15 23	8.1516
11	2 6 18 27 29 25 30 12 15 23 13	8.1536
12	2 6 18 27 29 25 30 12 15 23 13 22	8.1555
13	2 6 18 27 29 25 30 12 15 23 13 22 3	8.1571
14	2 6 18 27 29 25 30 12 15 23 13 22 3 1	8.1571
15	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17	8.1571
16	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20	8.1572
17	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19	8.1574
18	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14	8.1591
19	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24	8.1628
20	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26	8.1705
21	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5	8.1727
22	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16	8.1731
23	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9	8.1731
24	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4	8.1731
25	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21	8.1731
26	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21 11	8.1731
27	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21 11 7	8.1731
28	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21 11 7 8	8.1731
29	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21 11 7 8 10	8.1731
30	2 6 18 27 29 25 30 12 15 23 13 22 3 1 17 20 19 14 24 26 5 16 9 4 21 11 7 8 10 28	8.1731



**TABELA B-T2: Resultado do seletor de característica com kernel SVM**

SFS		
SELEÇÃO	SEQÜÊNCIA DAS CARACTERÍSTICAS (Variável Independente)	RMS
1	24	30.6510
2	24 5	30.3943
3	24 5 28	30.1708
4	24 5 28 4	30.0200
5	24 5 28 4 10	29.8867
6	24 5 28 4 10 17	29.7584
7	24 5 28 4 10 17 19	29.7161
8	24 5 28 4 10 17 19 9	29.6713
9	24 5 28 4 10 17 19 9 27	29.6713
10	24 5 28 4 10 17 19 9 27 29	29.6717
11	24 5 28 4 10 17 19 9 27 29 30	29.6721
12	24 5 28 4 10 17 19 9 27 29 30 12	29.6739
13	24 5 28 4 10 17 19 9 27 29 30 12 20	29.6760
14	24 5 28 4 10 17 19 9 27 29 30 12 20 25	29.6501
15	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8	29.6131
16	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11	29.5272
17	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14	29.5265
18	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26	29.5291
19	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13	29.5354
20	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7	29.5487
21	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15	29.5491
22	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21	29.5090
23	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23	29.4874
24	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18	29.4872
25	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16	29.5080
26	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16 6	29.5289
27	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16 6 22	29.6210
28	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16 6 22 1	30.2443
29	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16 6 22 1 2	30.8538
30	24 5 28 4 10 17 19 9 27 29 30 12 20 25 8 11 14 26 13 7 15 21 23 18 16 6 22 1 2 3	30.9485

**TABELA B-T3: Resultado do seletor de característica com kernel ANFIS**

SFS		
SELEÇÃO	SEQÜÊNCIA DAS CARACTERÍSTICAS (Variável Independente)	RMS
1	25	29.7540
2	25 28	29.8017
3	25 28 24	29.8862
4	25 28 24 22	29.9662
5	25 28 24 22 5	30.2221
6	25 28 24 22 5 23	30.2891
7	25 28 24 22 5 23 6	30.3114
8	25 28 24 22 5 23 6 10	30.5331
9	25 28 24 22 5 23 6 10 4	30.5585
10	25 28 24 22 5 23 6 10 4 7	30.5732
11	25 28 24 22 5 23 6 10 4 7 29	30.8529
12	25 28 24 22 5 23 6 10 4 7 29 15	30.9001
13	25 28 24 22 5 23 6 10 4 7 29 15 9	30.9366
14	25 28 24 22 5 23 6 10 4 7 29 15 9 14	30.9741
15	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13	31.0746
16	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8	31.1648
17	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16	31.3360
18	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21	31.3508
19	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2	31.4114
20	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20	31.4497
21	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3	31.4797
22	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17	31.5713
23	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30	31.5739
24	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26	31.5991
25	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27	31.6761
26	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27 18	31.7260
27	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27 18 11	31.7770
28	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27 18 11 1	31.7934
29	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27 18 11 1 19	31.7970
30	25 28 24 22 5 23 6 10 4 7 29 15 9 14 13 8 16 21 2 20 3 17 30 26 27 18 11 1 19 12	31.9724

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)