

# MODELOS DE PREVISÃO APLICADOS NO CONTROLE ESTATÍSTICO DE PROCESSO NA PRESENÇA DE DADOS AUTOCORRELACIONADOS

**Mário William Pessoa de Lima**

Instituto de Desenvolvimento Gerencial - INDG  
Alameda da Serra, 500 34000-000 - Nova Lima/MG

**Reinaldo Charnet**

Departamento de Estatística  
Instituto de Matemática, Estatística e Computação Científica-Unicamp  
Caixa Postal 6065, CEP 13083-970, Campinas, SP.  
e-mail:charnet@ime.unicamp.Br

## **Resumo**

O controle estatístico de processos, quando aplicado a dados provenientes de processos químicos, em geral, pode identificar uma série muito grande de alarmes falsos, devido à interdependência das observações. Uma alternativa para solucionar esse problema, é ajustar um modelo de séries temporais aos dados e fazer o controle do processo nos resíduos do modelo ajustado. Neste trabalho é analisada uma série de observações provenientes de uma indústria de fabricação de alumínio, para o qual, o controle estatístico do processo é feito por meio da análise dos resíduos de um modelo de séries temporais, devido à presença de autocorrelações significativas nos dados interferindo no controle usual do processo. Mostra-se também que os modelos de séries temporais reduzem significativamente a presença de alarmes falsos no processo.

## **1. Introdução**

Em muitas situações, em processos de produção, quando o controle é feito a partir de técnicas estatísticas usuais de controle estatístico do processo, é comum deparar-se com uma seqüência de ocorrências que podem ser, em sua maioria, classificadas como alarmes falsos, ou seja, há indicações de que o processo está fora de controle de modo que a produção necessita ser interrompida e investigada. Entretanto, o processo não se encontra fora de controle, e o aparecimento desses alarmes falsos é comum quando as observações são autocorrelacionadas.

O objetivo deste trabalho é analisar e apresentar uma alternativa de controle estatístico de processo, quando o mesmo é estruturado por componentes de interdependência nos seus dados, bem como alertar para as conseqüências, em geral, da existência dessa interdependência dos

dados coletados. Nesse sentido é proposta uma técnica de controle estatístico do processo a partir dos resíduos de um modelo de séries temporais ajustado aos dados após ser identificada a existência de autocorrelação nos dados.

Para empreender os objetivos descritos, na Seção 2 será apresentada uma breve noção de modelos de séries temporais, em particular os modelos denominados Box-Jenkins com ênfase na escolha de um determinado modelo, no ajuste e na verificação da adequação aos dados, bem como nas principais definições e conceitos, utilizados neste trabalho. Na Seção 3, serão apresentados conceitos gerais sobre controle de processos e a utilização dos resíduos do modelo ajustado, no controle do processo, bem como a construção dos gráficos de controle a partir dos resíduos do modelo de série temporal escolhido. Na Seção 4, o processo de fabricação e a coleta dos dados serão descritos e a análise das observações será feita considerando os conceitos apresentados nas seções anteriores. As conclusões serão apresentadas na Seção 5.

## **2. Modelos de Séries Temporais**

O estudo da previsibilidade envolve dados históricos, e tem por objetivo detectar a existência de tendências e padrões. O seu conhecimento é utilizado na projeção de dados para períodos futuros.

Mais recentemente os modelos de Séries Temporais, particularmente os procedimentos denominados Box-Jenkins e Redes Neurais, têm recebido atenção especial com uma aplicação cada vez mais crescente.

Uma Série Temporal é uma seqüência de observações de certa variável aleatória de interesse,  $X$ , a qual é observada em certos intervalos de tempo, usualmente igualmente espaçados, e denota-se por  $x_t$ , o valor observado da variável de interesse  $X$ , no tempo  $t$ .

Os modelos mais simples de Séries Temporais são aqueles em que a variável  $X$  é expressa através de um modelo de regressão linear, no tempo  $t$ . Esta classe de modelos pode ser analisada usando-se as técnicas de Regressão Linear, vide Charnet, Freire, Charnet e Bonvino (1999).

Nesses modelos a pressuposição de independência dos erros  $\epsilon_t$ , e, conseqüentemente, a independência entre as observações  $x_t$ , nem sempre podem ser garantidas. É freqüente se deparar com situações em que as observações são fortemente dependentes uma das outras. Nesses casos o uso dos modelos descritos é inapropriado para se fazer previsões. Os modelos de Box-Jenkins

consideram especificamente a situação em que as observações são autocorrelacionadas, justificando a sua importância como um modelo de previsão para séries temporais.

Os modelos de Box-Jenkins podem ser gerados através do modelo abaixo, chamado de filtro linear, que consiste em considerar-se que a observação  $x_t$  possa ser modelada como uma função dos erros anteriores,  $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ , da seguinte forma:

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad (1)$$

onde,  $\varepsilon_i$  são variáveis aleatórias normais independentes com média zero e variância constante  $\sigma^2$ . Para todo  $i = t, t-1, t-2, \dots$ ,  $\psi_j, j = 0, 1, 2, \dots$ , são os parâmetros do modelo, usualmente chamados de peso e  $\mu$  é a constante que determina o nível do processo. Em geral adota-se o valor 1 para  $\psi_0$ .

Segundo Box e Jenkins (1976), uma série temporal é estacionária se existe equilíbrio em torno de uma constante e como consequência tem-se que  $E(x_t) = \mu$ , finita.

Em termos do modelo definido acima se a série for estacionária tem-se que:

$$E(x_t) = E\left(\mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right) = \mu. \quad (2)$$

Conseqüentemente  $\sum_{j=0}^{\infty} \psi_j = 0$ , visto que  $E(\varepsilon_t) = 0$ .

Pode-se mostrar que para uma série estacionária a variância de  $x_t$  é dada por:

$$\gamma_0 = V(x_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 \quad (3)$$

Define-se a autocovariância de período  $k$  da série como:

$$\gamma_k = Cov(x_t, x_{t-k}) \quad (4)$$

No caso da série ser estacionária, pode-se mostrar que a autocovariância de período  $k$  é dada por:

$$\gamma_k = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \quad (5)$$

A autocorrelação de período  $k$  é definida a partir de (3) e (4), ou seja:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (6)$$

Se for considerado o gráfico de  $\rho_k$  em função de  $k$ , tem-se a função de autocorrelação do processo.

Outra medida importante para se estudar a relação entre as componentes da série é a autocorrelação parcial de período  $k$ . Esta é a correlação entre  $x_t$  e  $x_{t+k}$ , eliminado os efeitos das componentes entre  $x_t$  e  $x_{t+k}$ , ou seja,  $x_{t+1}, x_{t+2}, \dots, x_{t+k-1}$ . Essa medida é também chamada de autocorrelação parcial de ordem  $k$ , e denotada por  $\Phi_{kk}$ . O gráfico de  $\Phi_{kk}$  em termos de  $k$  é chamado de função de autocorrelação parcial.

O modelo, definido em (1), apresenta uma complicação na sua utilização visto que o número de componentes do modelo é infinito. Um caso particular mais simples de ser utilizado é o modelo dado por:

$$x_t = \zeta + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + \varepsilon_t \quad (7)$$

os quais são denominados de modelos autoregressivos de ordem  $p$ , visto que  $x_t$  é expressa em termos das  $p$  componentes anteriores a ela e  $\zeta, \varphi_1, \varphi_2, \dots, \varphi_p$  são os parâmetros desconhecidos do modelo. A notação usada para esses modelos é AR( $p$ ).

Um outro caso especial a ser destacado do modelo de Box-Jenkins é o modelo de Médias Móveis, o qual é dado por:

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (8)$$

onde  $\theta_1, \theta_2, \dots, \theta_q$ , formam um conjunto finito de pesos.

O modelo acima é denotado por MA( $q$ ), e tem a média, variância e autocovariância de período  $k$ , dados por:

$$\begin{aligned} E(x_t) &= \mu \\ \gamma_0 = V(x_t) &= \sigma^2 \sum_{i=0}^q \theta_i^2 \\ \gamma_k &= \sigma^2 (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q), k = 1, 2, \dots, q \end{aligned} \quad (9)$$

onde  $\theta_0 = 1$  e  $\gamma_k = 0$  para  $k > q$ .

De maneira análoga ao modelo Ar( $p$ ) pode-se calcular a função de autocorrelação de um modelo MA( $q$ ), como sendo:

$$\rho_k = \frac{-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, k = 1, 2, \dots, q, \quad (10)$$

de forma que esta expressão é igual a zero se  $k > q$ .

Os modelos ARMA que envolvem processos autoregressivos e de média móvel, são dados por modelos do tipo:

$$x_t = \zeta + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (11)$$

cuja notação utilizada é ARMA(p,q), de acordo com Box e Jenkins (1976).

Uma outra classe muito importante a ser destacada são os modelos ARIMA(p,d,q) que englobam todos os modelos vistos anteriormente bem como os processos não estacionários.

A principal ferramenta para se identificar o modelo é a função de autocorrelação, que depende diretamente dos parâmetros desconhecidos do modelo. Deve-se estimar essa função através das observações consideradas na série histórica. Qualquer que seja o modelo a ser estimado a função de autocorrelação será estimada pela autocorrelação amostral expressa como:

$$\hat{\rho}_k = \frac{\frac{1}{N-k} \sum_{i=1}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}, \quad k = 0,1,2,\dots,K, \quad (12)$$

onde N é o número total de observações considerada na série e  $\bar{x}$  é a média amostral das N observações. Como regra geral calcula-se as primeiras  $K < N/4$  autocorrelações amostrais.

Outra informação importante, para a identificação do modelo, é fornecida pela função de autocorrelação parcial. Pode-se mostrar que em um modelo autoregressivo, a autocorrelação parcial,  $\phi_{kk}$ , é exatamente o coeficiente de ordem k do modelo, e que satisfazem a equação:

$$\rho_j = \varphi_{k1} \rho_{j-1} + \varphi_{k2} \rho_{j-2} + \dots + \varphi_{kk} \rho_{j-k}, \quad j = 1,2,\dots,k. \quad (13)$$

Considerando, portanto, a equação (13) é possível estimar as autocorrelações parciais, através da equação:

$$\hat{\rho}_j = \varphi_{k1} \hat{\rho}_{j-1} + \varphi_{k2} \hat{\rho}_{j-2} + \dots + \varphi_{kk} \hat{\rho}_{j-k}, \quad j = 1,2,\dots,k, \quad (14)$$

resolvendo o sistema para todo  $k = 1,2,\dots,K$ , e obtendo as estimativas  $\hat{\varphi}_{11}, \hat{\varphi}_{22}, \dots, \hat{\varphi}_{KK}$  dos respectivos,  $\varphi_{kk}$ ,  $k=1,2,\dots,K$ .

Para facilitar a análise das autocorrelações é usual apresentar o gráfico das estimativas das autocorrelações em função dos valores de k acrescentando-se dois limites, um inferior e outro superior, de forma que valores fora desses limites podem ser considerados como autocorrelações significativamente diferente de zero.

De maneira análoga os gráficos são construídos para as estimativas das autocorrelações parciais com os respectivos intervalos de significância.

No caso das autocorrelações os limites de significância são dados pela expressão apresentada por Bartlett (1946):

$$\pm 2S(\hat{\rho}_k) = \pm 2\sqrt{\frac{1}{N} \left[ 1 + 2 \sum_{j=1}^{k-1} \hat{\rho}_j \right]} . \quad (15)$$

Para as autocorrelações parciais os limites usando os resultados apresentados por Quenouille (1949) são:

$$\pm 2S(\hat{\varphi}_{kk}) = \pm 2\sqrt{\frac{1}{N}} \quad (16)$$

Esses limites colocados em um gráfico e anotando-se as autocorrelações, nos fornece uma melhor visualização do processo.

A Figura 1 apresenta, o gráfico da função de autocorrelação e pode-se observar que somente a autocorrelação de ordem 1 está fora dos limites, ou seja, podemos, então, considerar que essa autocorrelação é estatisticamente diferente de zero.

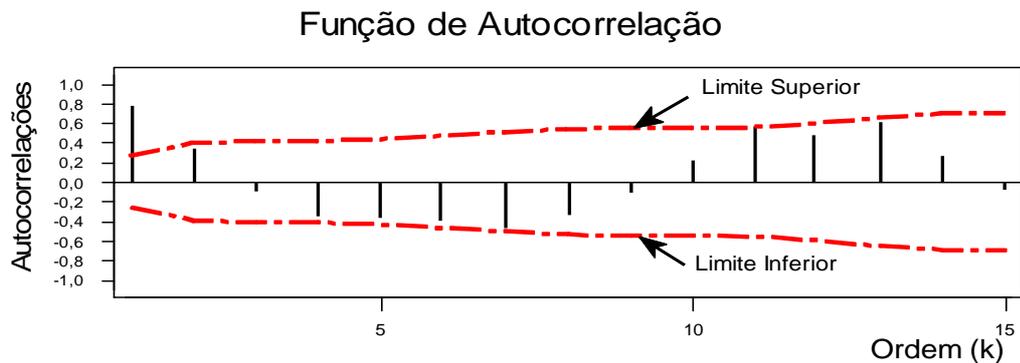


Figura 1: Gráfico das Estimativas das Autocorrelações

Considerando agora o gráfico para as autocorrelações parciais na Figura 2, nota-se que as autocorrelações parciais que estão fora dos limites, são as de ordem 1 e 2, todas as demais podem ser consideradas nulas. A partir dessa análise pode-se ajustar um modelo AR(2), visto que só as duas primeiras correlações parciais se mostram estatisticamente diferentes de zero.

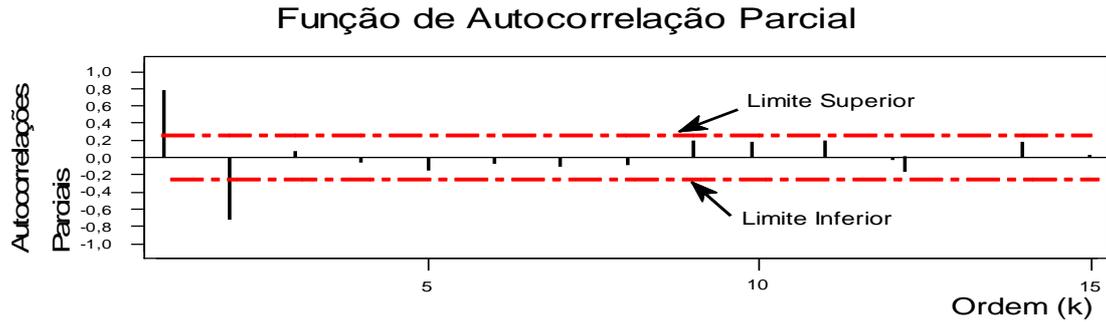


Figura 2: Estimativas das Autocorrelações Parciais

Uma vez identificado qual modelo deve ser considerado, o passo seguinte é estimar os parâmetros (coeficientes) do modelo. O método utilizado para estimação é denominado de Quadrados Mínimo, que é o mesmo utilizado para Modelos de Regressão. Este método consiste em encontrar as estimativas dos parâmetros que minimizem a soma de quadrados das diferenças entre os valores observados  $x_t$  e os respectivos valores fornecidos através do modelo proposto. Essas estimativas são denominadas de Estimativas de Quadrados Mínimo (EQM) dos parâmetros, ver Montgomery e Johnson (1976). Essas estimativas são obtidas através de processos iterativos, por não haver uma expressão fechada para elas é necessário a utilização de softwares específicos, como por exemplo, o MINITAB.

As medidas de ajuste do modelo estimado serão baseadas nos resíduos, os quais são dados por:

$$r_t = x_t - \hat{x}_t \quad (17)$$

onde  $\hat{x}_t$  é o valor estimado pelo modelo ajustado.

Para examinar o ajuste do modelo calcula-se a função de autocorrelação dos resíduos,  $r_t$ , denotada por  $\hat{\rho}_k(r_t)$ , de forma que, se o modelo estiver adequado, espera-se que essas autocorrelações não sejam diferentes de zero, caso contrário, é necessário mudar o modelo escolhido inicialmente.

Além das autocorrelações dos resíduos, pode-se, para modelos ARIMA(p,d,q), em geral, utilizar a estatística:

$$Q = (N - d) \sum_{k=1}^K \hat{\rho}_k^2(r_t) \quad (18)$$

que indicará se as  $K$  primeiras autocorrelações dos resíduos são significantes, ou não, sugerindo a adequação ou não do modelo, uma vez que se o modelo for adequado a estatística  $Q$  tem distribuição aproximada Qui-Quadrado com  $(N-p-q)$  graus de liberdade. Portanto rejeita-se que o modelo é adequado se  $Q$  for maior que o valor crítico de uma distribuição Qui-Quadrado com  $N-p-q$  graus de liberdade, para um nível de significância pré-determinado,  $\alpha$ .

### **3. Controle de Processos Autocorrelacionados**

A aplicação das ferramentas estatísticas como base no controle e determinação da estabilidade dos processos vem crescendo muito nos últimos anos. Dentre essas ferramentas, os gráficos de controle que compõem o controle estatístico do processo se destacam como uma das mais conhecidas e aplicadas.

Apesar da grande difusão e ampla utilização dos gráficos de controle nos mais diversos tipos de processos, ainda existem algumas dificuldades ou até mesmo barreiras, quando essa ferramenta é empregada para controlar processos que apresentam autocorrelação e, ou, sazonalidade. Nesse sentido, Runger e Willemain (1995), concluíram que processos com autocorrelação positiva desqualificam os gráficos de controle tradicionais. Se um modelo de série temporal é adequadamente ajustado a um processo autocorrelacionado, o uso de gráficos de controle dos resíduos desse modelo é recomendado, como se pode ver em Kenneth et al. (1997), Montgomery e Mastrangelo (1991) e Zhang (1998).

Os limites dos gráficos de controle convencionais têm como uma de suas premissas a independência dos dados do processo. Em processos autocorrelacionados essa premissa não é satisfeita o que inviabiliza a utilização dos gráficos de controle convencionais especialmente em certos tipos de processos, como concluíram Atienza et al (1998).

As presenças de autocorrelações significativas causam um profundo impacto nos gráficos de controle. Um dos principais impactos é o aumento significativo de falsos alarmes no gráfico de controle. Mesmo quando as observações são fracamente correlacionadas, a influência nos gráficos de controle pode ser grande conforme afirmam Hu e Roan (1996). Quando processos envolvem interdependência das observações, mesmo os procedimentos mais simples de análise de dados correlacionados fornecem resultados extremamente satisfatórios, identificando de forma simples e objetiva a existência de falsos alarmes.

Neste trabalho após seleção do modelo Box-Jenkins adequado, é considerada a construção de gráficos de controle para os resíduos do modelo ajustado, com a finalidade de se controlar o processo de interesse e identificar os possíveis falsos alarmes.

Após a etapa de modelagem já será possível saber onde os dados futuros do processo deverão se localizar. Percebe-se que sem a modelagem de um processo autocorrelacionado seria praticamente impossível estabelecer o controle estatístico do processo, pois nesses processos o comportamento natural dos dados adquire formas que, em princípio, poderiam ser atribuídas a causas especiais atuando no processo, mas na verdade seriam falsos alarmes.

Uma vez determinado o modelo adequado os resíduos,  $r_t$ , devem ser calculados como definido em (17) e, em seguida, as hipóteses de normalidade e independência dos resíduos, devem ser validadas.

Considerando-se finalmente os resíduos do modelo ajustado, os limites de controle do processo são calculados como se segue:

$$LSC = \bar{r} + 3 \hat{\sigma}_r \quad (19)$$

$$LM = \bar{r} \quad (20)$$

$$LIC = \bar{r} - 3 \hat{\sigma}_r \quad (21)$$

No gráfico para a amplitude móvel temos que o limite superior de controle é dado por  $3 \bar{R}_r$ . Nessas expressões temos:

$$\bar{r} = \frac{\sum_{t=1}^n r_t}{n} \quad (22)$$

$$\hat{\sigma}_r = \frac{\bar{R}_r}{d_2} \quad (23)$$

$$\bar{R}_r = \frac{\sum_{j=1}^{n-1} R_j}{(n-1)} \quad (24)$$

e

$$R_j = |r_{j+1} - r_j|, j=1,2,\dots,n-1 \quad (25)$$

É importante observar que se o processo está sob controle estatístico, os resíduos do modelo deverão estar entre os limites de controle apresentados acima flutuando em torno da média,  $\bar{r}$ , de maneira similar os gráficos de controle usuais.

#### 4. Estudo de Caso

A empresa na qual foi aplicada a técnica apresentada neste trabalho é do ramo siderúrgico e produz alumínio, utilizando como matéria prima o minério de bauxita.

O processo de produção do alumínio inicia-se pela etapa que se refere à clarificação do minério de bauxita. Nessa etapa, a bauxita passa pela fase de lavagem do minério em filtros, com o objetivo de retirar as impurezas nela contidas. Uma das impurezas mais críticas é o ferro, pois para aplicação do produto final (alumina) essa impureza precisa atender às especificações rígidas do mercado.

Com a finalidade de controlar a contaminação, por ferro, na bauxita, optou-se por controlar estatisticamente essa variável, teor de ferro na bauxita, através da técnica de gráficos de controle.

Como o processo é contínuo e não há interrupções de produção, foi estabelecida a frequência de amostragem a cada 2 horas. Essa frequência de amostragem possibilitou reduzir o erro de alterar o processo sem necessidade. Justifica-se essa frequência de amostragem pelo fato de que em um período superior a 2 horas tem-se um grande volume de produção sem avaliação. Também foi considerado para a frequência da amostragem o tempo entre a coleta da amostra, a análise do laboratório e a tomada de decisão de alterações nos parâmetros de controle do processo.

O fluxo do processo da clarificação é apresentado na figura seguinte.

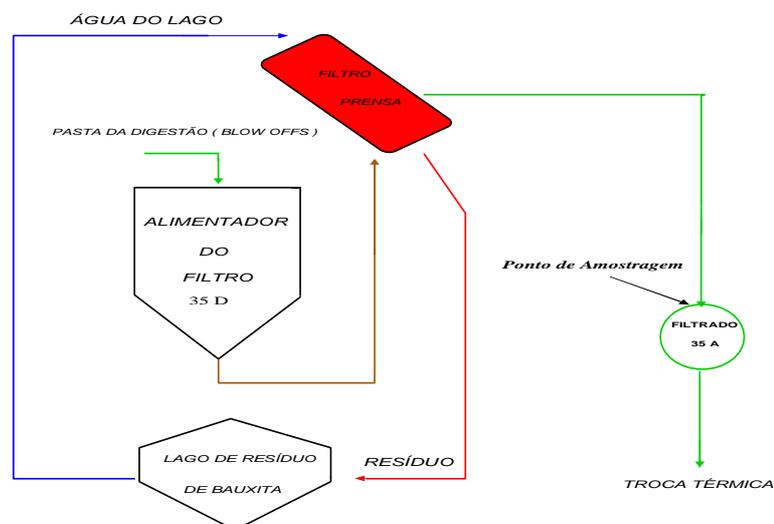


Figura 3: Fluxograma do Processo

No fluxograma do processo é possível observar que a entrada do processo é representada pela pasta da digestão, que é a bauxita dissolvida em solução caustica. O alimentador do filtro 35 D tem a função de estocar o material a ser filtrado. O filtro prensa faz a separação do produto das impurezas (principalmente o ferro). Após a filtração, o produto vai para o tanque 35 A onde fica estocado e segue para a troca térmica. A parte rejeitada no filtro prensa é chamada de resíduo de bauxita, sendo o mesmo enviado para o lago de resíduo de bauxita. A fim de recuperar a soda presente no resíduo de bauxita, a água do lago retorna ao processo pelo filtro prensa.

É importante ressaltar que a tomada de decisão para se alterar os parâmetros do processo é baseada em dados amostrais e, portanto, existe a possibilidade de se cometer dois tipos de erros:

- O erro chamado tipo 1, que ocorre quando se altera o processo sem motivos suficientes (causas especiais) para fazê-lo. Esse erro é o mais grave pois insere variações adicionais ao processo.
- O erro chamado tipo 2, que ocorre quando não se altera o processo quando de fato deveria ser alterado. Esse erro é minimizado pela adequada frequência da amostragem.

O erro tipo 1 provoca alarmes falsos no processo, ou seja, indica que algo especial ocorreu sem que isso realmente tivesse acontecido. Geralmente, quando o alarme falso acontece, a operação é mobilizada para corrigir o processo e alterações são realizadas.

Ao se considerar o processo de amostragem de duas em duas horas são gerados 12 pontos por dia. Para avaliar o processo foi necessário coletar dados referentes a três meses de produção, pois segundo os engenheiros de processo esse período é suficiente para obtenção de todas as fontes de variação do processo.

Com o objetivo de avaliar o controle estatístico do processo, a variável teor de ferro na bauxita é que será estudada através de um gráfico de controle.

O processo em estudo apresenta um plano de amostragem no qual cada amostra é representada por um único resultado. Dessa forma o gráfico de controle indicado é o de dados individuais e amplitude móvel ( $\bar{X} - AM$ ). Mais detalhes sobre gráfico de controle com amplitudes móveis, pode ser encontrado em Montgomery (2004).

Os gráficos de controle para a análise da estabilidade do teor de ferro, são apresentados a seguir.

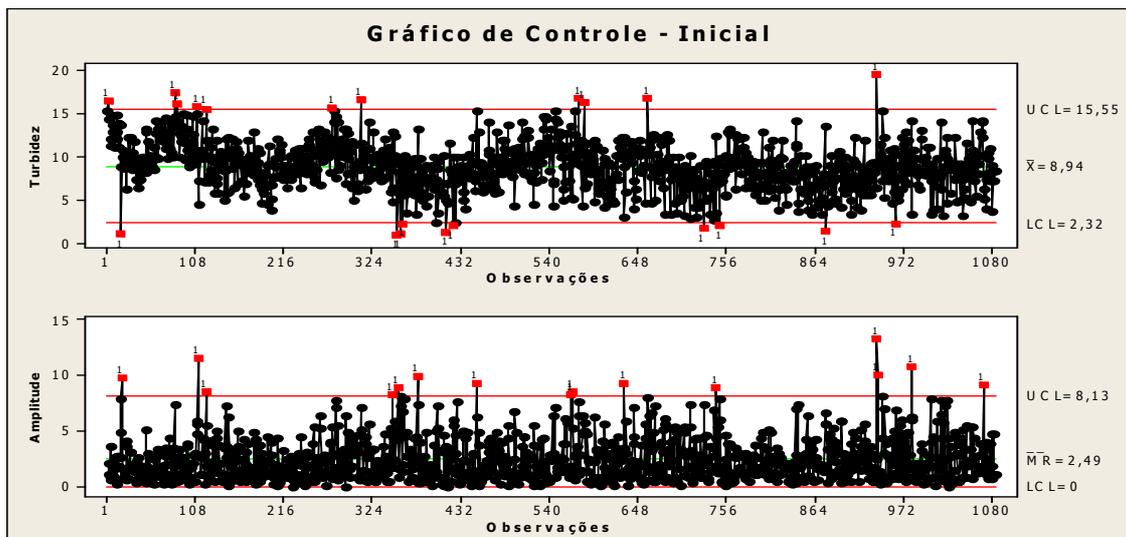


Figura 4: Gráficos de Controle - Individual e Amplitude Móvel

No gráfico de valores individuais da Figura 4 foram sinalizadas causas especiais de variação em 11 pontos acima do limite superior e 9 pontos abaixo do limite inferior de controle. Pode-se ainda verificar um comportamento cíclico.

No gráfico de amplitudes móveis, também são observadas várias sinalizações de situações de fora de controle.

Para avaliar a estabilidade do processo se considerou apenas o critério ponto fora de controle que codifica como 1 os pontos que ultrapassam os limites de controle.

O próximo passo da análise é a identificação das causas especiais de variação que estão provocando a instabilidade no processo. Algumas causas foram encontradas, porém, a grande maioria dos pontos não apresentava nenhum motivo que justificasse uma atuação no processo de forma pontual, pois se tratava de causas comuns de variação, ou seja, todo o sistema estava afetado.

As sinalizações apontadas nas cartas de controle nada mais eram que “alarmes falsos” e caso alguma ação fosse implementada pontualmente, a variabilidade do processo aumentaria.

Durante as discussões sobre o problema foi mencionado que os resultados da variável considerada não eram independentes, ou seja, o ponto amostrado, por exemplo, às 10:00hs sofria influência da amostragem das 08:00hs e influenciava no resultado da amostra das 12:00hs, o que indica um processo autocorrelacionado sugerindo a necessidade de uma análise de autocorrelação do processo.

O modelo inicial identificado após a análise das autocorrelações parciais foi o de média móvel 1, MA(1), que se mostrou não suficiente para eliminar a autocorrelação dos resíduos e, portanto, deveria ser melhorado.

A análise das autocorrelações do modelo MA(1) sugeriu a necessidade da inclusão do efeito autoregressivo no modelo e, portanto, um modelo ARMA(5,1) foi ajustado.

A análise estatística do modelo ARMA(5,1), mostrou que, apesar de haver uma melhora no ajuste com relação ao modelo MA(1), os parâmetros AR3, AR4 e AR5, não eram significativos, e, além disso, os resíduos continuavam autocorrelacionados. Com essas informações simplificou-se o modelo, eliminando-se os parâmetros não significativos e ajustando-se um modelo ARMA(2,1).

O modelo ARMA(2,1) ajustado, além de ser mais simples, mostrou que todos os seus parâmetros eram significativos.

O gráfico de autocorrelação dos resíduos do modelo ARMA(2,1) mostra que, praticamente todos os pontos se encontram dentro dos limites de significância indicando apenas que alguns valores altos estão igualmente espaçados de 12 pontos sugerindo, assim, um possível efeito de sazonalidade de três períodos, como se pode verificar na Figura 5.

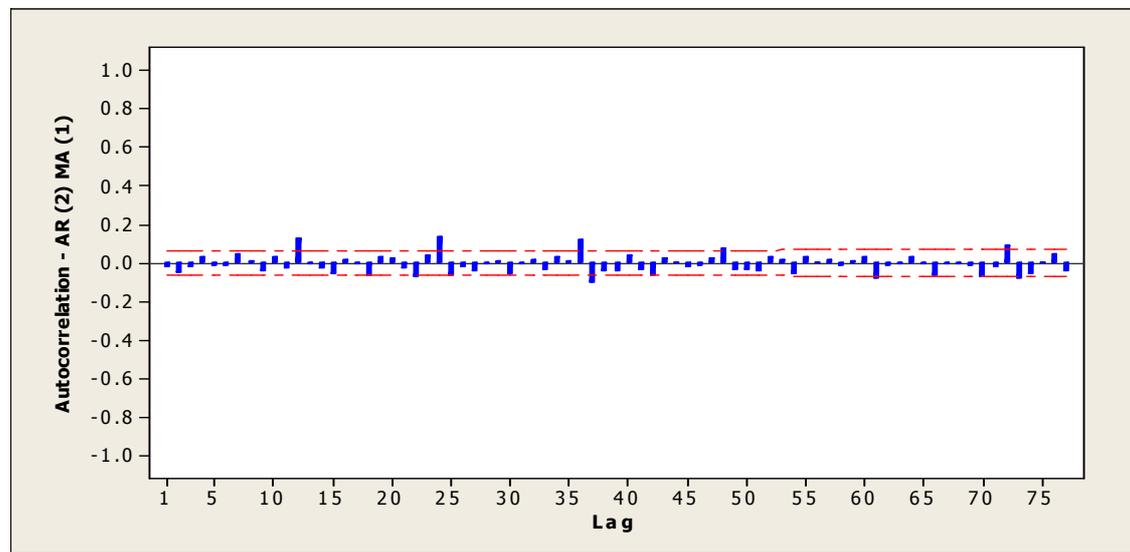


Figura 5: Gráfico da Autocorrelação dos Resíduos - Segundo Modelo

Para confirmar a necessidade de incluir o efeito da sazonalidade foi considerado o gráfico da autocorrelação parcial desse modelo, apresentado na Figura 6.

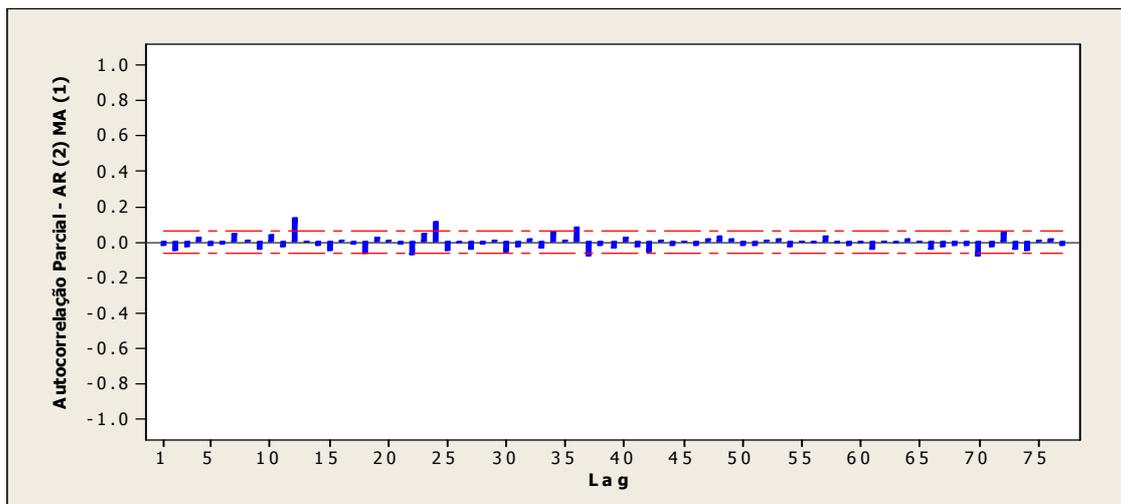


Figura 6: Gráfico da Autocorrelação Parcial dos Resíduos - Segundo Modelo

No gráfico da Figura 6, observa-se a necessidade de inclusão do efeito da sazonalidade nos pontos 12, 24 e 36, pois os mesmos ultrapassaram o limite superior de confiança.

A sazonalidade detectada no gráfico acima era esperada, pois, tendo em vista que o processo é contínuo e não há interrupção da produção, como foi mencionado anteriormente, e que estabeleceu-se a frequência de amostragem a cada duas horas, a presença de sazonalidade é inerente ao processo devida ao plano de amostragem utilizado.

Com o objetivo de “eliminar” o efeito da sazonalidade do modelo ajustou-se um modelo ARMA(2;1) Sazonal 12 por três períodos. A análise de variância desse modelo encontra-se na Tabela 1.

Tabela 1: Análise de Variância – Modelo ARMA(2;1) Sazonal 12

Estimativa Final dos Parâmetros				
Parâmetro	Coef.	Erro Padrão do Coef.	T	P
<b>AR 1</b>	1,0760	0,0430	25,03	<b>0,000</b>
<b>AR 2</b>	-0,0993	0,0360	-2,76	<b>0,006</b>
<b>SAR 12</b>	0,1112	0,0311	3,58	<b>0,000</b>
<b>SAR 24</b>	0,1250	0,0312	4,01	<b>0,000</b>
<b>SAR 36</b>	0,1055	0,0311	3,39	<b>0,001</b>
<b>MA 1</b>	0,8900	0,0282	31,51	<b>0,000</b>
<b>Constant</b>	0,138515	0,007940	17,44	<b>0,000</b>
Média	9,0390	0,5182		
Número de observações:1084				
Resíduos:	SS = 6061,23			
	MS = 5,63	GL = 1077		
Lag	12	24	36	48
Q	10,4	22,6	32,5	44,9
GL	5	17	29	41
<b>Valor P</b>	<b>0,064</b>	<b>0,163</b>	<b>0,300</b>	<b>0,314</b>

Na Tabela 1 observa-se que os valores p da estatística Q são todos maiores ou iguais ao valor usual 0,05 indicando que os resíduos não se encontram autocorrelacionados e, portanto, pode-se concluir que o modelo é adequado. Além disso, todos os parâmetros do modelo são significantes (valores p são todos menores que 0,006).

A partir das análises realizadas, o modelo final ajustado foi:

$$\hat{x}_t = 0,138515 + 1,076x_{t-1} - 0,0993x_{t-2} + 0,1112x_{t-12} + 0,125x_{t-24} + 0,1055x_{t-36} + 0,89\varepsilon_{t-1}$$

Para verificar se os efeitos das autocorrelações foram totalmente inseridos no modelo ajustado é necessário avaliar a autocorrelação dos resíduos do modelo ajustado, através dos gráficos de autocorrelação e autocorrelação parcial, Figura 7 e Figura 8.

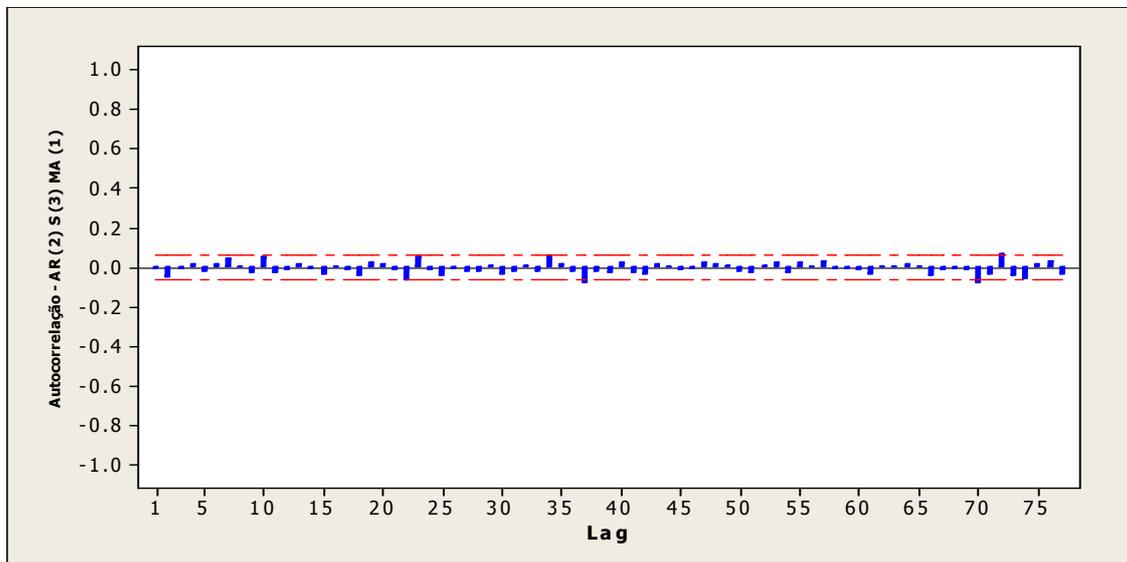


Figura 7: Gráfico da Autocorrelação dos Resíduos – Modelo Final

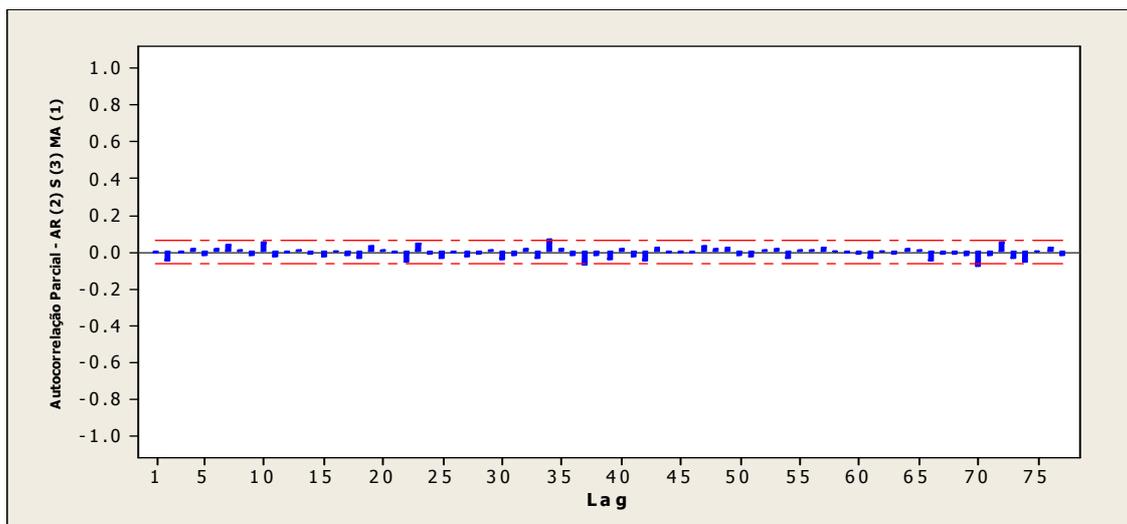


Figura 8: Gráfico da Autocorrelação Parcial dos Resíduos – Modelo Final

Nas Figuras 7 e 8 não são observadas autocorrelações significativas. Isso indica que o modelo considerado foi adequado para “eliminar” os efeitos da autocorrelação e sazonalidade que estavam presentes nos dados.

Para se validar o modelo ajustado os resíduos devem satisfazer as seguintes condições:

- 1- Variância constante – Gráfico dos resíduos versus valor ajustado;
- 2- Aleatoriedade – Gráfico dos resíduos versus a ordem da coleta dos dados; e,
- 3- Normalidade – Histograma dos Resíduos

Os dois primeiros critérios podem ser avaliados de forma visual através do gráfico dos resíduos. O gráfico dos resíduos versus os valores ajustado mostrou um padrão aleatório, indicando assim a homogeneidade de variância. Da mesma forma o gráfico dos resíduos versus a ordem das observações também mostrou aleatoriedade. Já o terceiro critério pode ser avaliado através do histograma, o qual no caso confirmou a normalidade dos resíduos.

Para avaliar a estabilidade do processo construiu-se um gráfico de controle (Indivíduos e Amplitude Móvel) para os resíduos do modelo ajustado, já que os mesmos refletem a real situação do processo e possuem as características necessárias (independência e normalidade) para poderem ser avaliados.

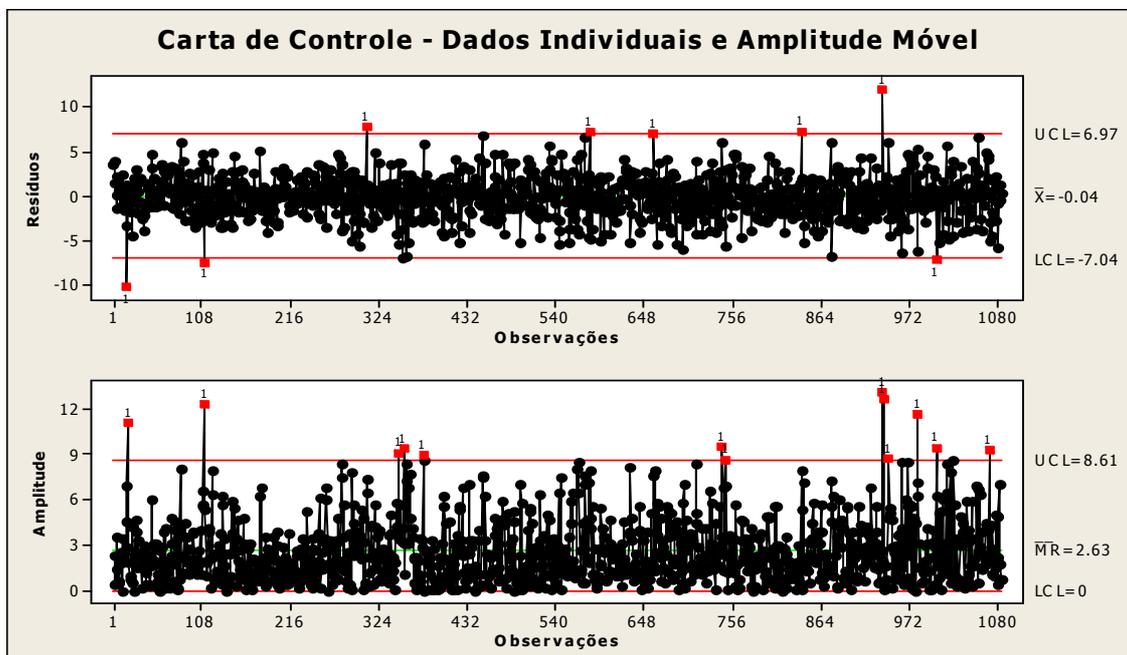


Figura 9: Gráficos de Controle dos Resíduos - Individual e Amplitude Móvel

É importante salientar que os limites do gráfico de controle na Figura 9, foram obtidos a partir de (19) e (21).

No primeiro gráfico de controle de dados individuais, aplicado na variável teor de ferro, identifica-se 11 pontos acima do limite superior de controle sendo que, a grande maioria dessas sinalizações é caracterizada como “alarmes falsos”. Com o ajuste do modelo foi possível definir o comportamento natural do processo e então se pode avaliar a estabilidade do processo através do gráfico de controle para os resíduos do modelo.

Na Figura 9 observa-se que o gráfico de dados individuais dos resíduos mostra apenas 5 pontos acima do limite superior de controle. Esses pontos são as reais sinalizações de causas especiais. A partir destas análises, foi possível, então, verificar se o processo contém ou não causas especiais de variação.

Conforme observado nos gráficos de controle para os resíduos do modelo, constatou-se que existem situações especiais que provocam uma variação não natural ao processo.

Para a identificação das causas de variação sugeridas no gráfico de controle, foi realizado um *brainstorming* e elaborado um diagrama de causa e efeito, apresentado a seguir:

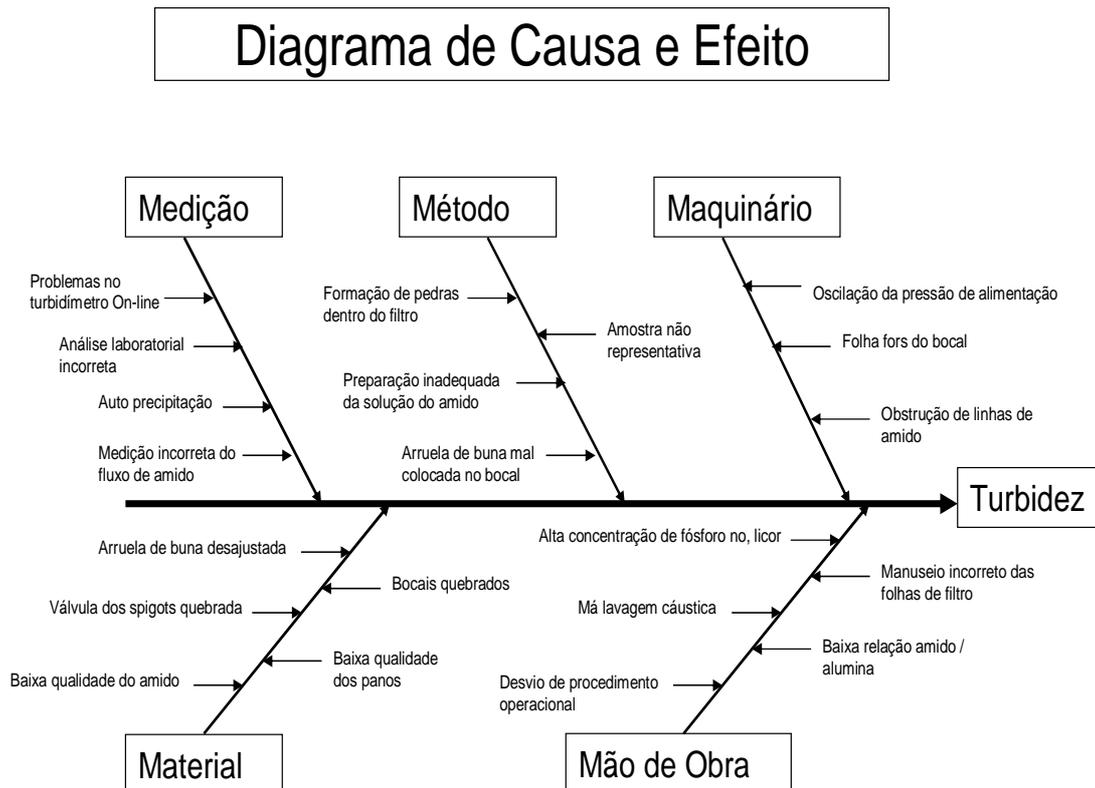


Figura 10: Diagrama de Causa e Efeito

Dentre as causas sugeridas no diagrama, quatro foram selecionadas para serem “atacadas”: “Folha fora do bocal”, “Obstrução da linha de amido”, “Manuseio incorreto das folhas de filtro” e “Bocais quebrados”.

#### **4.1 Capacidade do Processo**

Após a modelagem o controle do processo continua utilizando a variável teor de ferro e incorporando as informações sobre a variabilidade obtida pelo modelo ajustado.

Os limites de especificação estabelecidos na área de produção são construídos da seguinte forma, uma vez que o valor do teor de ferro deve ser o menor possível, estabelece-se apenas o limite superior. A faixa natural do processo foi definida pelo gráfico de controle dos resíduos, Figura 9, onde o limite superior foi 7, ou seja, o máximo permitido naturalmente pelo processo.

Realizando essa adaptação estabelece-se o seguinte limite de controle: Média dos valores do teor de ferro + 7,00, ou seja,  $LSC = 8,94 + 7,00 = 15,94$ .

Para avaliação da capacidade do processo foi necessário levantar a especificação do cliente, uma vez que o limite de especificação do processo é definido pelo máximo de contaminação de ferro permitido, esse valor deverá ser igual a 15.

Para avaliar a capacidade de atendimento à especificação do processo são comparados os resultados antes e depois da implantação do plano de ação definido para bloquear as causas especiais identificadas na carta de controle para os resíduos do modelo ajustado.

No período inicial das análises foram coletadas observações durante o período de 90 dias. A análise de capacidade do processo mostrou que a cada um milhão de resultados de análises do teor de ferro 16.605 superavam o máximo especificado, ou seja, provocavam não conformidades no processo.

Na Figura 11, a curva normal de linha contínua representa uma simulação do processo na qual estão inclusas apenas variações provocadas por causas comuns, enquanto a curva normal de linha pontilhada representa uma simulação do processo na qual estão inclusas variações em decorrência de causas comuns e especiais. Com essas duas curvas normais pode-se verificar que resolver os problemas devido a causas especiais proporcionará ao processo uma melhoria na capacidade em atender a especificação do cliente.

As curvas normais de linha contínua e pontilhada utilizam os desvios padrão 2,20504 e 2,71813 nos quais estão contidas respectivamente: causas comuns e causas comuns e especiais.

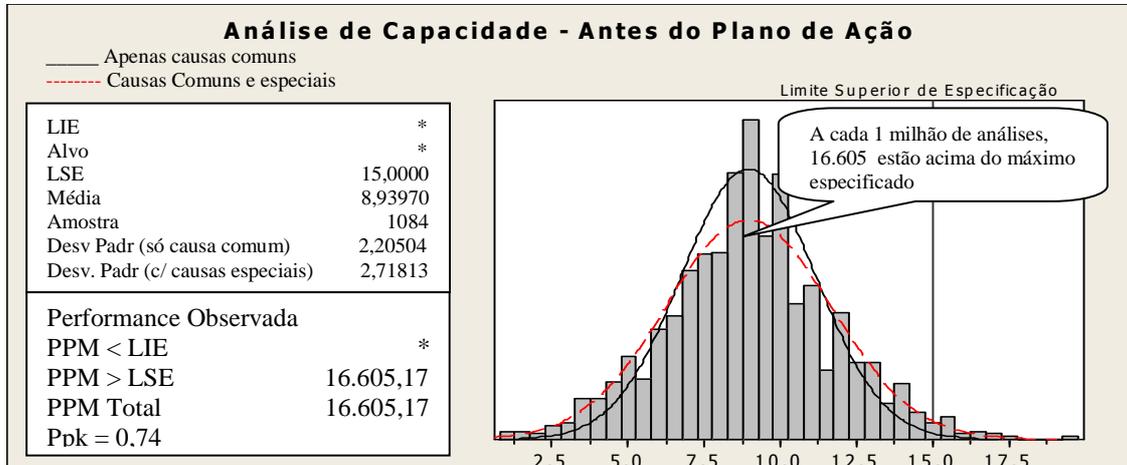


Figura 11: Atendimento ao Cliente – Antes do Plano de Ação

Ao finalizar a execução do plano de ação descrito, foram coletadas observações referentes a 30 dias de operação, para em seguida, avaliar da estabilidade do processo.

Na Figura 12 observa-se que o comportamento cíclico apresentado na primeira carta de controle não existe mais, já que a correta aplicação da ferramenta reduziu as intervenções operacionais no processo devido à ocorrência de alarmes falsos.

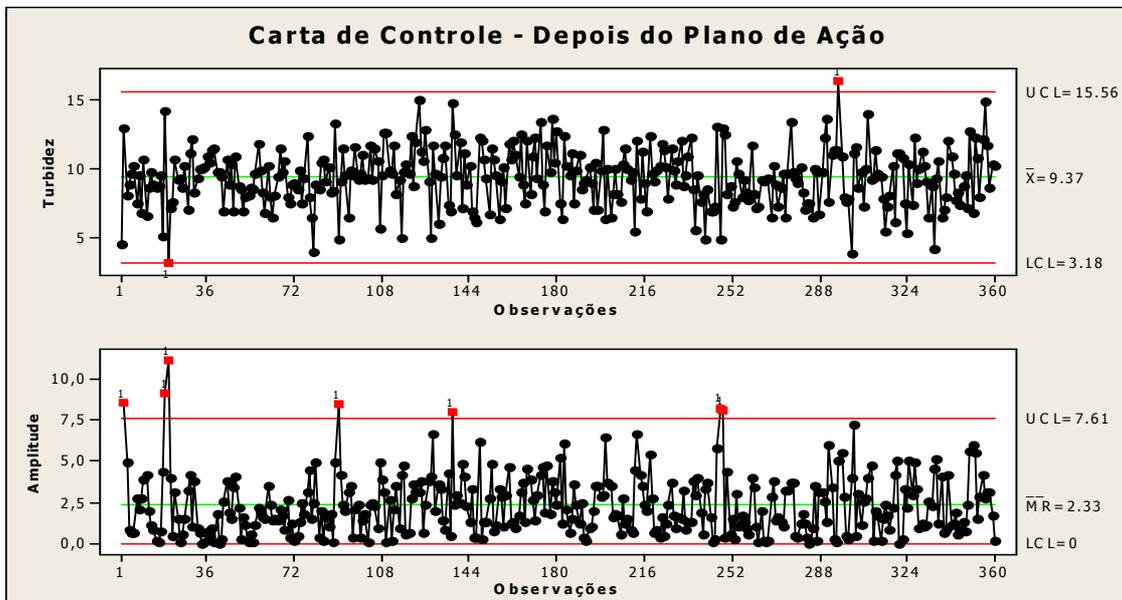


Figura 12: Gráficos de Controle – Depois do Plano de Ação

Após eliminação das causas especiais a faixa de variação natural do processo é de 3,18 a 15,56 com média igual a 9,37.

Avaliou-se a capacidade depois do plano de ação, e foram obtidos os resultados apresentados na Figura 13.

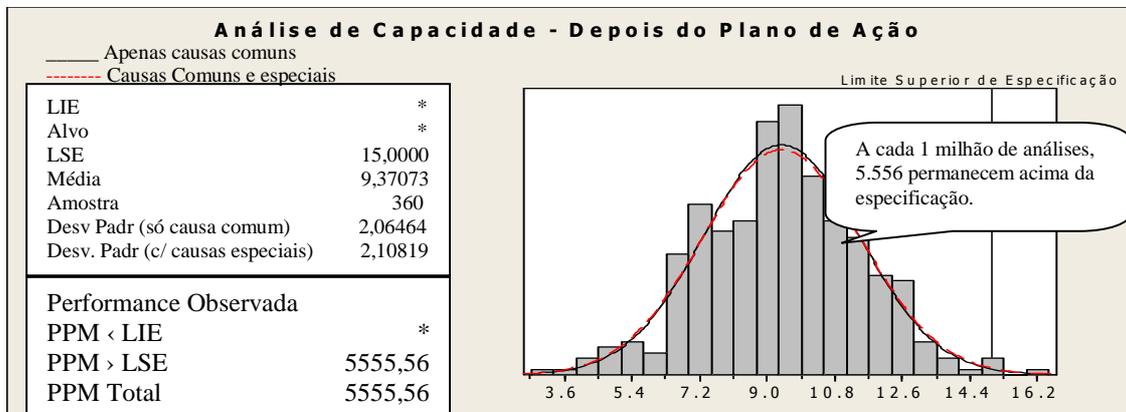


Figura 13: Atendimento ao Cliente – Depois do Plano de Ação

Observando a Figura 13 é possível constatar que após a implantação do plano de ação o desempenho do processo apresenta a cada um milhão de análises do teor de ferro, 5.556 acima do limite de especificação, o que corresponde a um índice de desempenho  $Ppk = 0,89$ . Com a redução alcançada de 16.605 ppm para 5.556 ppm a melhoria foi de 66,5%.

Para melhor atender as necessidades dos clientes outros estudos deverão ser conduzidos e assim reduzir as causas de variação inerentes (causas comuns) ao processo e as novas causas especiais que ainda geram alguns pontos fora do especificado.

## 5. Conclusões

Quando o controle de um processo é feito utilizando-se o CEP, a presença de dados correlacionados causa sério impacto com o aumento significativo de falsos alarmes, levando a paradas desnecessárias no processo de produção. Uma alternativa para controlar processos que tenham interdependência nos dados é a utilização de Modelos de Séries Temporais, em particular Modelo Box-Jenkins.

Uma vez detectada a autocorrelação a utilização de modelos Box-Jenkins constitui uma ferramenta importante para que através da análise dos resíduos do modelo ajustado, seja possível identificar a presença ou não de falsos alarmes no processo. A construção de cartas de controles para os resíduos apresenta-se extremamente eficiente e simples de ser implementada, conduzindo a resultados plenamente satisfatórios, quanto à identificação dos falsos alarmes, quando se tem um processo com dados autocorrelacionados.

Os resultados apresentados neste trabalho mostram a eficiência na utilização dos modelos Box-Jenkins no controle de processos autocorrelacionados quando os resíduos do modelo são utilizados no controle do processo. No estudo de caso o gráfico de controle utilizado nos dados originais mostrava onze pontos acima do limite superior e, após a identificação da autocorrelação dos dados e o ajuste do modelo adequado, o gráfico de controle aplicado aos resíduos, apresenta apenas cinco pontos fora de controle. Além disso, o estudo da capacidade do processo mostra que antes da aplicação do modelo de Séries Temporais tinha-se 16.605 pontos acima do limite máximo de especificação, enquanto que após o ajuste do modelo sugerido obteve-se 5.556 pontos acima do limite máximo de especificação, ou seja, houve uma melhoria de 66,5%, com o índice Ppk que foi 0,74, passando para 0,89 após ter sido feito o controle dos resíduos do modelo ajustado aos dados

Como visto no estudo de caso, pode-se afirmar que se houver autocorrelação significativa nos dados, existe uma grande chance de que as causas especiais de variação sejam detectadas se as técnicas descritas neste trabalho forem utilizadas.

### **Referências Bibliográficas**

- ATIENZA, O. O.; TANG, L.C. and ANG, B.W. Simultaneous Monitoring of Sample and Group Autocorrelations. **Quality Engineering**, Vol 14, No. 3, pp. 489-499, 2002.
- BARTLETT, M.S. On the theoretical specification and sampling properties of auto correlated time-series. **Suppl. J. Royal Statist. Soc.** 8, 27-41, 1946.
- BOX, G. E. P. and JENKINS, G. M. (1976). **Time Series Analysis, Forecasting, and Control**. Horden Day, Oakland , CA, 1976.
- CHARNET, R.; FREIRE, L. A. C.; CHARNET, E. M.; BONVINO H. **Análise de Modelos de Regressão Linear com Aplicações**. Editora da UNICAMP, 1999.
- HU, J. S. and ROAN, C. Change Patterns of Time Series-Based Control Charts. **Journal of Quality Technology** Vol 28, No. 3, pp. 302-312, 1996.
- KENNETH, G.C.; KENNETH, K. and CHERYL, H.R. Charting Autocorrelated Data: Guidelines for Practitioners. **Quality Engineering**, Vol.9, No.3, pp.367-382, March 1997.
- MONTGOMERY, D. C.; JOHNSON, L.A. **Forecasting and Time Series Analysis**. Mcgraw-Hill, NY, 1976.

MONTGOMERY, D. C and MASTRANGELO, C.M. Some Statistical Process Control Methods for Autocorrelated Data. **Journal of Quality Technology**, Vol.23, No.3, July 1991

MONTGOMERY, D.C. **Introdução ao Controle Estatístico da Qualidade**. LTC Editora, 2004

QUENOUILLE, M.H. Approximate tests of correlation in time-series. J. Roy. Statist. Soc. Ser. B. 11, 68-84, 1949.

RUNGER, G. C. and WILLEMAIN T. R. Model-Based and Model-Free Control of Auto correlated Process. **Journal of Quality Technology** Vol 27, 4 pp. 283-292, 1995.

ZHANG, N.F. Statistical Control Chart for Stationary Process Data. **Technometrics**, Vol.40, No.1 February 1998.

### **FORECASTING MODELS APPLIED IN STATISTICAL CONTROL PROCESS OF AUTOCORRELATED DATA**

#### **Abstract**

This work presents an alternative model for statistical control of the process when we have autocorrelated data. The traditional methods suggest the control chart as the most appropriate tool to be used for the identification of the two different sources of variation for all types of processes.

The prediction models based on time series play an important role when the main purpose is to control those processes that produce series of autocorrelated data. Box -Jenkins models deal specifically with those situations of autocorrelated data and this brings up its importance as a prediction model in time series analysis.

A case study conducted in a chemical industry is presented as an example of one application of the suggested model, once it deals with an autocorrelated process in which the initial data extend its influence on subsequent data for a certain period of time. The main variable of interest in this case study is the contamination of iron in the final product. This variable has been evaluated since the beginning of the process attempting to control the contamination of iron in the product in a level such that its utilization does not cause any damage to the customer. With this adjusted model the purpose is to control the residuals (predicted – observed) which should remain inside the interval determined by the control limits with mean zero.

*Key Words: Time Series Forecasting, Statistics Control, Capability*