

Ajuste de um modelo de previsão de séries temporais para um processo de fabricação de cerâmicas

Fernando de Jesus Moreira Junior <fmjunior@smail.ufsm.br>
Robert Wayne Samohyl <samohyl@deps.ufsc.br>
Enio Júnior Seidel <ejrseidel@hotmail.com>
Luis Felipe Dias Lopes <lflopes@smail.ufsm.br>

Resumo: Neste artigo desenvolveu-se um estudo para melhorar o gerenciamento da produção de uma indústria de cerâmica utilizando ferramentas de Análise de Séries Temporais. Ajustou-se uma equação de regressão múltipla, com variáveis dummies altamente significativas, que mostrou a presença da sazonalidade, obtendo um valor para o R-quadrado igual a 0,84, e para o R-quadrado ajustado igual a 0,79. O coeficiente de correlação entre os valores estimados pelo modelo e os valores observados foi igual a 0,91. O modelo foi considerado adequado para fazer previsões.

Palavras-chave: Análise de Séries Temporais; Fabricação de Cerâmicas; Sazonalidade.

Adjusting a model for predicting time series for a manufacturing process of ceramics

Abstract: In this article one study was developed to improve the management of the production for one ceramic industry using tools of Time Series Analysis. An equation of multiple regression was adjusted, with variables dummies highly significant, what showed the presence of the seasonality, obtaining a value for the R-square equal to 0,84, and for the adjusted R-square equal to 0,79. The coefficient of correlation between the values estimated by the model and the observed values was equal to 0 91. The model was considered appropriate to make predictions.

Keywords: Time series analysis; Manufacture of ceramics; Seasonality.

1. Introdução

A demanda por produtos cerâmicos sempre esteve atrelada ao potencial do mercado regional que se pretende atingir, de modo a considerar-se quais as necessidades deste mercado com relação à qualidade dos produtos, quantidade e preços praticados. Tendo em vista a rapidez com que ocorrem mudanças nesse mercado e a grande concorrência, as empresas cerâmicas são forçadas a buscarem a evolução dos seus processos produtivos, monitorando e otimizando a produção.

Dantas *et al* (2001) destaca que o processo de previsão é uma das etapas mais importantes dentro do planejamento agregado. Mesmo assim, um dos maiores problemas enfrentados pelas empresas é a dificuldade de prever, argumenta Ramos *et al* (2001). Muitas vezes, as previsões são feitas por um gerente de vendas ou uma equipe através de estimativas de vendas baseadas nas previsões dos próprios vendedores e representantes no mercado, sem

uma metodologia explícita e consistente, com forte tendência de superestimar as vendas observadas.

Conforme Gonçalves (2007), a utilização de modelos de previsões cresce à medida em que gestores buscam métodos científicos para tentar diminuir a dependência da sorte. Hoje, os modelos de previsões fazem parte do processo decisório da gestão empresarial nas áreas financeira, de recursos humanos, de marketing, de produção, etc. Segundo Balestrassi *et al* (1998), a previsão de séries temporais é um dos fatores mais importantes na engenharia de produção.

A análise de séries temporais é uma área de pesquisa relevante em diversos campos do conhecimento, tendo como principal objetivo em suas pesquisas, providenciar uma previsão, quando o modelo matemático de um fenômeno é desconhecido ou incompleto (SAMOHYL, 2001b).

Com o objetivo de estimar e prever a produção de cerâmica para melhorar o gerenciamento da produção, buscou-se utilizar as ferramentas de análise de Séries Temporais.

2. Modelos de Previsão de Séries Temporais

Segundo Samohyl *et al* (2001a), as técnicas de análise de séries temporais são extremamente importantes, pois permitem identificar, em dados temporais, padrões de comportamento históricos não-aleatórios, o que possibilita não só a previsão de ocorrências futuras, como também a organização de ações preventivas em relação às suas conseqüências. Estas previsões são feitas por meio de técnicas científicas que projetam dados passados para o futuro, após a identificação de um padrão de comportamento relacionado com o tempo.

De acordo com Gonçalves (2007) os métodos de previsão podem ser classificados como qualitativos ou quantitativos. Os métodos qualitativos, utilizados quando os dados históricos são inexistentes ou escassos, são baseados na opinião de especialistas e referem-se ao uso de técnicas de grupos focais e de consenso, como o método Delphi. Por outro lado, os métodos quantitativos, referentes a os modelos matemáticos, são utilizados quando existe quantidade suficiente de dados históricos.

Entre os métodos quantitativos de previsão, destacam-se os modelos ingênuos, de média móvel, de regressão (linear simples, múltipla, não-linear, logística, etc.), de alisamento exponencial, auto-regressivo, integrado e de média móvel (ARIMA), auto-regressão de vetor (VAR), entre outros. Cada uma dessas classes de modelo é utilizada de acordo com as características dos dados (tendência, sazonalidade, autocorrelação, ciclos, etc.) e dos pressupostos de cada modelo. Também é possível que os mesmos dados possam ser ajustados por mais de um modelo da mesma classe ou de diferentes classes. Para a escolha do melhor modelo, devem ser observados determinados critérios, de acordo com o modelo (coeficientes, critérios de informação, análise de resíduos, parcimônia, etc.). Makridakis *et al* (1998) destaca que a previsão deve ser executada considerando-se alguns importantes passos: i) definição do problema; ii) coleta das informações; iii) análise preliminar dos dados; iv) escolha e ajuste de modelos; e v) uso e avaliação do modelo de previsão.

Segundo Samohyl *et al* (2001b), uma série temporal pode ser definida como uma função de uma variável independente (tempo) vinculada a um processo em que uma descrição matemática é desconhecida. O comportamento futuro de uma série temporal não pode ser previsto exatamente, como pode ser previsto de uma função determinística. Entretanto, o comportamento de uma série temporal pode algumas vezes ser antecipado através de

procedimentos estocásticos. Quatro comportamentos associados a uma série temporal podem ser identificados: o efeito de tendência, o efeito sazonal, os ciclos de negócios e as variações irregulares ao acaso.

3. Materiais e Método

Os dados utilizados neste estudo foram coletados em uma indústria cerâmica localizada no município de Pinheirinho do Vale, norte do Rio Grande do Sul, o qual se caracteriza pelas estações climáticas bem definidas, isto é, temperaturas altas no verão e baixas no inverno. Foram coletados dados referente a produção de tijolos no período de janeiro de 2003 a dezembro de 2005, totalizando 36 observações.

3.1 Processo de produção de tijolos

Na cerâmica em estudo tem-se o processo de produção de tijolos como segue. Na caixa de alimentação, a matéria-prima (argila) segue por uma esteira até chegar ao misturador, onde é controlada a umidade e é acrescida água, ficando com aspecto de “massa”. Do misturador, a matéria-prima é conduzida até o laminador, que a reduz a uma forma pastosa em lâminas finas, fazendo-a passar entre dois cilindros de ferro fundido que trituram por esmagamento todas as pedrinhas ou torrões ainda não desfeitos. O material laminado cai na maromba (máquina de fabricar tijolos) a vácuo, e vai para o parafuso-sem-fim, onde a matéria-prima é impelida para a frente, passa através da câmara de vácuo e toma a forma de tijolo. O bloco de argila (já na forma do tijolo), saindo da boquilha, corre sobre os rolos da máquina cortadora e é automaticamente cortado no tamanho especificado.

Após o corte, os tijolos são transportados manualmente até os estaleiros de secagem, onde são colocados em prateleiras e permanecem para secagem natural por um período médio de 7 (sete) dias com tempo bom.

Após a secagem, os tijolos são transportados até os fornos e empilhados a fim de que a queima se processe de forma homogênea em todos os tijolos. No lado interno do forno, próximo às fornalhas, é colocada uma pilha de tijolos para que contenha o fogo que provem das fornalhas, protegendo assim os demais tijolos. A indústria possui 6 (seis) fornos, e cada forno armazena aproximadamente 13.000 (treze mil) tijolos. A energia utilizada na queima é obtida através de madeiras de reflorestamentos. O processo de queima dos tijolos do forno é de cinco dias. No primeiro dia, a temperatura do forno gira em torno de 50°C. Nos segundo e terceiro dias, a temperatura vai de 100°C a 300°C. Nestes três primeiros dias, a temperatura é mantida baixa (entre 50 e 300°C) para finalizar a secagem dos tijolos. No quarto dia, a temperatura do forno é elevada a 500°C, para iniciar o processo de queima do tijolo. Este processo é feito no quinto dia, aonde a temperatura vai de 500°C a 1000°C (temperatura final do processo) aproximadamente.

Após completo o processo, quando se passaram os cinco dias, ou seja, 120 horas de queima dos tijolos, o forno é então fechado, sem que haja entrada de ar, por um período de cinco horas. Depois deste período o forno é aberto para resfriamento dos tijolos (período que dura em torno de dois dias).

4. Resultados e discussões

O gráfico da Figura 1 apresenta a evolução da produção ao longo dos meses durante os três anos de coleta.

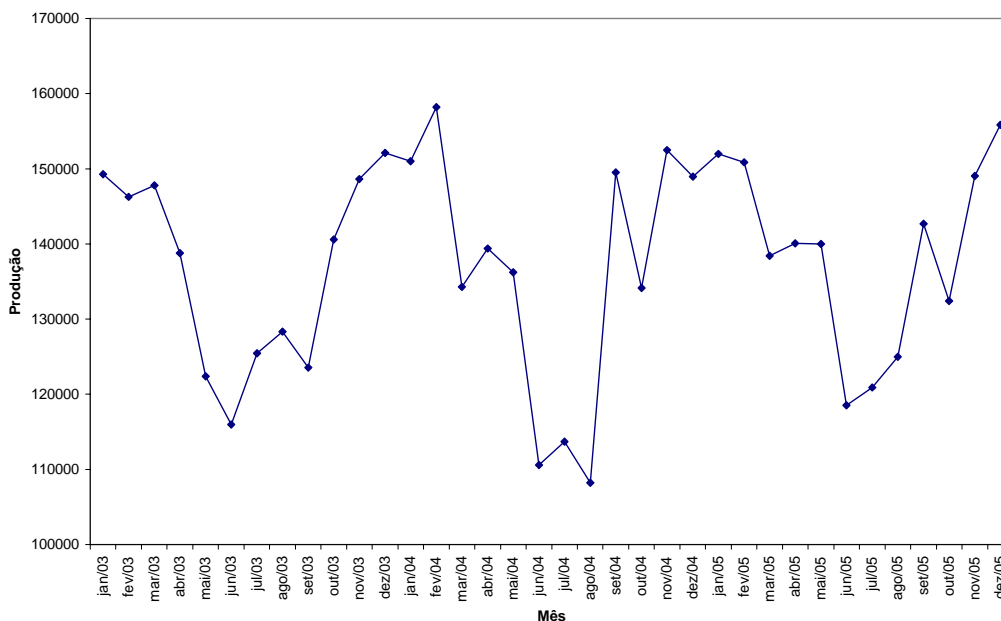


Figura 1 – Produção de Cerâmicas

Através do gráfico da Figura 1 pode-se observar um comportamento sazonal em relação aos meses dos anos, onde a produção tende a ser maior nos meses mais quentes e, menor nos meses mais frios. Por outro lado, não é possível verificar se existe uma tendência crescente ou decrescente na produção durante o período examinado.

A fim de ajustar uma equação de regressão aos dados da produção, optou-se por uma equação que pudesse detectar a presença de sazonalidade e tendência. Segundo Gonçalves (2007), existem várias formas de captar o efeito sazonal. Para captar um padrão sazonal que se repete a cada 12 meses, juntamente com a tendência, Moore (2006) sugere a seguinte equação:

$$Y = \beta_0 + \beta X + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{11} X_{11} + \varepsilon, \tag{1}$$

onde

β_0 é a constante,

X é a variável de regressão relacionada a tendência,

β é o coeficiente da variável de tendência X ,

$$X_1 = \begin{cases} 1 & \text{se o mês é janeiro} \\ 0 & \text{caso contrário} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{se o mês é fevereiro} \\ 0 & \text{caso contrário} \end{cases}$$

⋮

$$X_{11} = \begin{cases} 1 & \text{se o mês é novembro} \\ 0 & \text{caso contrário} \end{cases}$$

$\beta_1, \beta_2, \dots, \beta_{11}$ são os coeficientes das variáveis sazonais X_1, X_2, \dots, X_{11} , respectivamente, e

ε é o erro.

Os dados de dezembro são indicados quando todas as 11 variáveis indicadoras forem iguais a zero.

Na equação 1, as variáveis X_1, X_2, \dots, X_{11} são variáveis *dummies*, que servem para verificar a presença ou ausência de um atributo (ou qualidade), nesse caso, para identificar o mês. Segundo Gujarati (2000), as variáveis *dummies* são binárias, ou seja, assumem os valores 1, se o atributo está presente, ou 0, se o atributo está ausente. Segundo Samohyl (2009), uma equação de Regressão múltipla proporciona que o lado direito da equação está aberto para receber qualquer número de variáveis independentes.

Assim, aplicando-se o modelo da equação 1, obteve-se os resultados apresentados nas Tabelas 1 e 2. A Tabela 1 mostra que a regressão foi significativa ($p = 0,00000$), enquanto que a Tabela 2 apresenta os parâmetros estimados para a regressão. Essa mostra que 7 variáveis relacionadas à sazonalidade possuem efeito significativo ($p < 0,05$) no modelo, porém a variável relacionada à tendência não foi significativa ($p = 0,4395$). O valor obtido para o R-quadrado foi 0,84 e para o R-quadrado ajustado foi 0,757, indicando bom ajuste do modelo.

Tabela 1 – Teste de Significância para a Regressão

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	12	5,7E+09	4,75E+08	10,08625	0,00000
Resíduo	23	1,08E+09	47107001		
Total	35	6,79E+09			

Tabela 2 – Estimação dos Parâmetros da Regressão

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	150103	4853	30,93	0,0000
X	92	117	0,79	0,4395
X1	-530	5749	-0,09	0,9274
X2	385	5724	0,07	0,9469
X3	-11317	5702	-1,98	0,0592
X4	-12152	5681	-2,14	0,0433
X5	-18790	5663	-3,32	0,0030
X6	-36712	5648	-6,50	0,0000
X7	-31831	5634	-5,65	0,0000
X8	-31416	5623	-5,59	0,0000
X9	-13441	5615	-2,39	0,0252
X10	-16403	5609	-2,92	0,0076
X11	-2155	5605	-0,38	0,7042

Verificando-se que a tendência não foi significativa, o modelo será ajustado novamente, sem o componente de tendência, conforme a equação 2:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{11} X_{11} + \varepsilon, \quad (2)$$

Aplicando-se o modelo da equação 2, obteve-se os resultados apresentados nas Tabelas 3 e 4. A Tabela 3 mostra que a regressão foi significativa ($p = 0,00000$), enquanto que a Tabela 4 apresenta os parâmetros estimados para a regressão. Essa mostra que 8 variáveis relacionadas à sazonalidade possuem efeito significativo ($p < 0,05$) no modelo. A remoção da variável relacionada à tendência fez com que outra variável sazonal (X_3) se tornasse significativa no modelo, e aumentou a significância das variáveis que já eram significativas.

O valor obtido para o R-quadrado foi 0,84 e para o R-quadrado ajustado foi 0,761, o que indica um bom ajuste do modelo.

Tabela 3 – Teste de Significância para a Regressão

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	11	5,67E+09	5,16E+08	11,1236	0,00000
Resíduo	24	1,11E+09	46358797		
Total	35	6,79E+09			

Tabela 4 – Estimação dos Parâmetros da Regressão

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	152307	3931	38,74	0,0000
X1	-1540	5559	-0,28	0,7841
X2	-533	5559	-0,10	0,9244
X3	-12143	5559	-2,18	0,0389
X4	-12887	5559	-2,32	0,0293
X5	-19433	5559	-3,50	0,0019
X6	-37263	5559	-6,70	0,0000
X7	-32290	5559	-5,81	0,0000
X8	-31783	5559	-5,72	0,0000
X9	-13717	5559	-2,47	0,0211
X10	-16587	5559	-2,98	0,0065
X11	-2247	5559	-0,40	0,6897

As variáveis sazonais significativas foram: X_3 , X_4 , X_5 , X_6 , X_7 , X_8 , X_9 e X_{10} , que representam os meses de março, abril, maio, junho, julho, agosto, setembro, outubro e novembro, respectivamente. Por outro lado, a Tabela 4 mostra que as variáveis X_1 , X_2 e X_{11} , que representam os meses de janeiro, fevereiro e novembro, respectivamente, não foram significativas nesse modelo. Dessa forma, os parâmetros do modelo serão novamente estimados sem essas variáveis.

Fazendo-se isso, obteve-se os resultados apresentados nas Tabelas 5 e 6. A Tabela 5 mostra que a regressão foi significativa ($p = 0,00000$), enquanto que a Tabela 6 apresenta os parâmetros estimados para a regressão. Essa mostra que todas as variáveis têm efeito significativo no modelo ($p < 0,05$).

O valor obtido para o R-quadrado foi 0,83, o que indica um bom ajuste do modelo, porém é menor que o valor do modelo 2 (0,84). Segundo Sartoris (2003), o valor de R-quadrado nunca cairá quando se acrescenta qualquer variável explicativa, mesmo que não seja significativa. Portanto, para se comparar modelos que possuem número de variáveis explicativas diferentes, deve-se utilizar o R-quadrado ajustado. Dessa forma, tem-se que o valor do R-quadrado ajustado para esse modelo foi de 0,79 indicando que esse é melhor que o modelo da equação 1 (0,76).

Tabela 5 – Teste de Significância para a Regressão

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	8	5,66E+09	7,08E+08	17,03959	0,00000
Resíduo	27	1,12E+09	41545370		
Total	35	6,79E+09			

Tabela 6 – Estimação dos Parâmetros da Regressão

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	151227	1861	81,28	0,0000
X3	-11063	4161	-2,66	0,0130
X4	-11807	4161	-2,84	0,0085
X5	-18353	4161	-4,41	0,0001
X6	-36183	4161	-8,70	0,0000
X7	-31210	4161	-7,50	0,0000
X8	-30703	4161	-7,38	0,0000
X9	-12637	4161	-3,04	0,0052
X10	-15507	4161	-3,73	0,0009

A exclusão das variáveis X_1 , X_2 e X_{11} indica que a produção entre os meses de novembro, dezembro, janeiro e fevereiro não varia de forma significativa, o que pode ser observado no gráfico da Figura 2, que apresenta os valores observados e previstos da produção. Essa exclusão também contribuiu para aumentar o valor do R-quadrado ajustado, melhorar o ajuste da equação ($F = 17,04$) e aumentar a significância das variáveis do modelo, como observa-se na Tabela 6. Assim, o modelo final ajustado foi:

$$Y = 151227 - 11063 X_3 - 11807 X_4 - 18353 X_5 - 36183 X_6 - 31210 X_7 - 30703 X_8 - 12637 X_9 - 15507 X_{10}$$

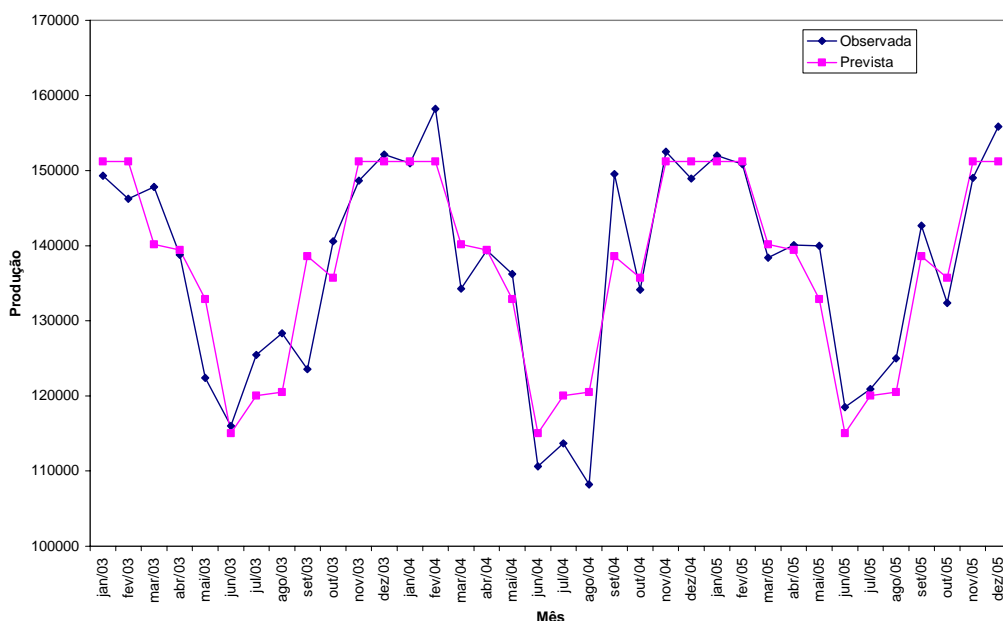


Figura 2 – Valores Observados e Estimados da Produção de Cerâmicas

A correlação entre os valores observados e estimados da produção de cerâmica, foi de 0,91, revelando um alto grau de associação positiva entre eles, conforme mostra o gráfico de dispersão da Figura 3.

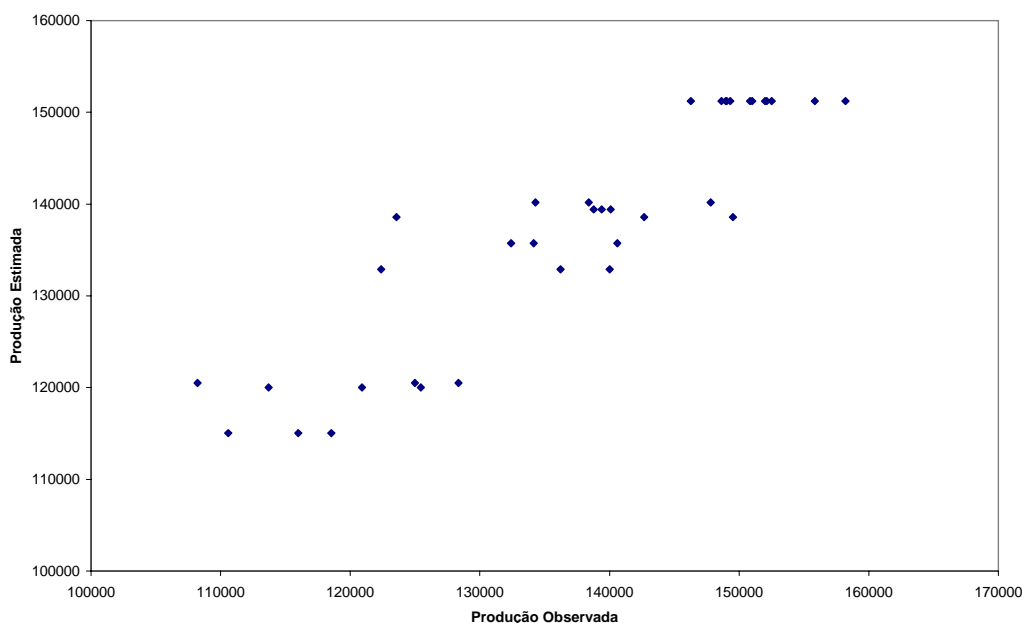


Figura 3 – Diagrama de Dispersão entre os Valores Observados e Estimados

Na sequência, utilizou-se o modelo obtido para fazer previsões. Segundo Gonçalves (2007), a previsão deve ser no máximo um terço do total dos dados. Assim, com base no modelo final ajustado, fez-se a previsão da produção para o ano seguinte (doze meses de

2006), que é mesma para os anos seguintes, uma vez que não existe tendência nos dados. O resultado da previsão é apresentado na Figura 4.

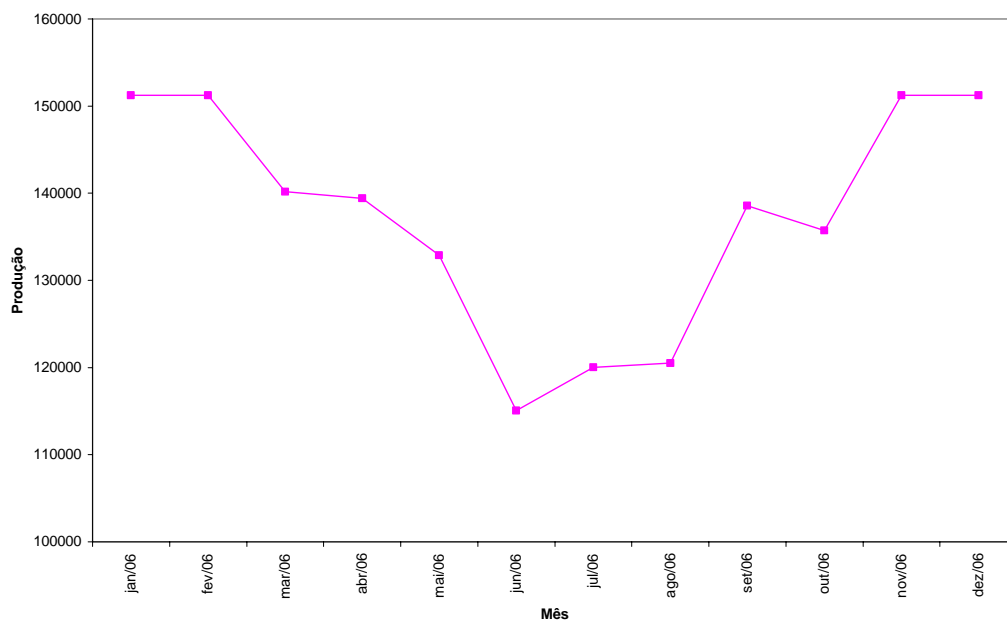


Figura 4 – Valores Previstos para a Produção de Cerâmicas

Observa-se também que nos meses de novembro, dezembro, janeiro e fevereiro, a produção não varia de forma significativa, por isso o modelo ajustado prevê o mesmo valor nesses meses.

5. Conclusões

Esse trabalho procurou ajustar um modelo de séries temporais aos dados referentes a produção de cerâmicas, com o objetivo de estimar e prever a produção de cerâmica para melhorar o gerenciamento da produção.

Foi ajustada uma equação de regressão múltipla com variáveis *dummies* que captou a presença da sazonalidade, obtendo um valor para o R-quadrado igual a 0,84, e para o R-quadrado ajustado igual a 0,79. Os resultados mostraram que: o modelo obteve um bom ajuste e que as suas variáveis foram altamente significativas.

Dessa forma, o modelo ajustado foi considerado adequado e pode ser utilizado para a previsão da produção desde que não haja mudança na tendência da produção.

Referências

BALESTRASSI, P. P.; SAMOHYL, R. W.; MEURER, R. *Métodos de Previsão usando a Série Econométrica de Nelson-Plosser: Um Estudo Comparativo*. XVIII Encontro Nacional de Engenharia de Produção (ENEGEP), Niterói, RJ, 1998.

DANTAS, A. B.; SAMOHYL, R. W.; PEREIRA, A. F. O, A. *Previsão de Custo de Hora-Extra com a Utilização do Modelo Tobit*. XXI Encontro Nacional de Engenharia de Produção (ENEGEP), Salvador, BA, 2001.

GONÇALVES, F. *Excel Avançado 2003/2007: Análise e Previsão de Demanda*. Rio de Janeiro: Editora Ciência Moderna Ltda., 2007.

GUJARATI, D. N. *Econometria Básica*. 3ª. Edição. São Paulo: Pearson Makron Books, 2000.

MAKRIDAKIS, S. G.; WHEELWRIGHT, S. C.; HYNDMAN, R.J. *Forecasting: methods and applications*. 3 ed. New York: John Wiley & Sons, 1998.

MOORE, D. S.; MCCABE, G. P., ;DUCKWORTH, W. M.; SCLOVE, S. L. *A Prática da Estatística Empresarial: Como Usar Dados para Tomar Decisões*. Rio de Janeiro: LTC, 2006.

RAMOS, E. M. L. S.; KONRATH, A. C.; SAMOHYL, R. W. *Previsão em MRP usando a Transformação de Box-Cox através do Aplicativo Glim com aplicação*. XXI Encontro Nacional de Engenharia de Produção (ENEGEP), Salvador, BA, 2001.

SAMOHYL, R. W. *Controle Estatístico de Qualidade*. São Paulo: Editora Elsevier/Campus, no prelo, a sair em 2009.

SAMOHYL, R. W.; MATTOS, V. L. D.; ROCHA, R. *Técnica de Amortecimento Exponencial Simples com Taxa de Resposta Adaptativa: Uma Reflexão a Respeito do Comportamento do Coeficiente Alfa*. XXI Encontro Nacional de Engenharia de Produção (ENEGEP), Salvador, BA, 2001a.

SAMOHYL, R. W.; ROCHA, R.; MATTOS, V. L. D. *Utilização do Modelo de Holt-Winters para a Previsão do Leite Entregue às Indústrias Catarinenses*. XXI Encontro Nacional de Engenharia de Produção (ENEGEP), Salvador, BA, 2001b.

SARTORIS, A. *Estatística e Introdução à Econometria*. São Paulo: Saraiva, 2003.