

4 Resultados

4.1. Escolha da metodologia Box & Jenkins

Como foi visto no ítem 2.1.3 deste trabalho, com a análise histórica dos preços do cobre, conseguimos entender um pouco melhor como é regido este mercado. Foi dito que a série de preços do cobre tem como algumas das principais variáveis a oferta, a demanda, o crescimento mundial, os estoques do minério, e expectativas futuras.

Todas essas variáveis tem em comum a sua dependência com os valores passados. Este fato nos leva a crer que existe alguma relação com dos preços futuros do cobre com os seus preços passados, pois os preços passados dependiam dos mesmos fatores.

Além da dependência da série com os preços passados, todas as outras variáveis citadas a cima podem ser consideradas constantes em curto prazo. Para aumentar a oferta, por exemplo, considerando todo o ciclo de produção do cobre, desde os primeiros estudos de viabilidade da mina até o início da sua produção, demora-se alguns anos.

A escolha do modelo Box&Jenkins se encaixa justamente nessa situação, em que os valores futuros sofrem influência dos passados e as outras variáveis dependentes não se alteram significativamente no curto prazo, fazendo com que a série tenda a sua média histórica a curto prazo.

4.2. Normalidade

Utilizando-se o teste de Jarque-Bera com 99% de confiança, não se encontra indícios de normalidade. Entretanto, quando usamos os dados para a função $\log(\text{retornos diários})$, uma melhora sensível foi sentida como pode-se observar no gráfico a seguir:

Figura 10: Histograma dos Preços da Commodity Cobre

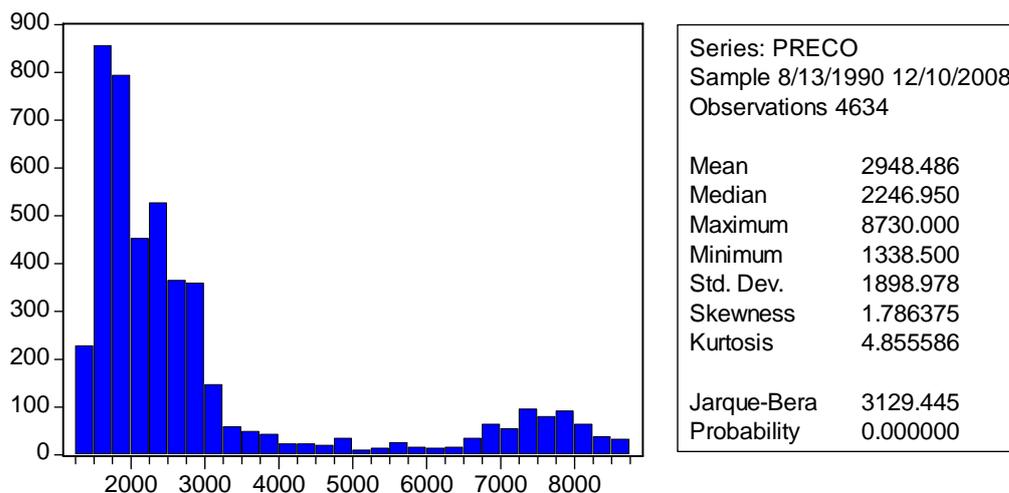
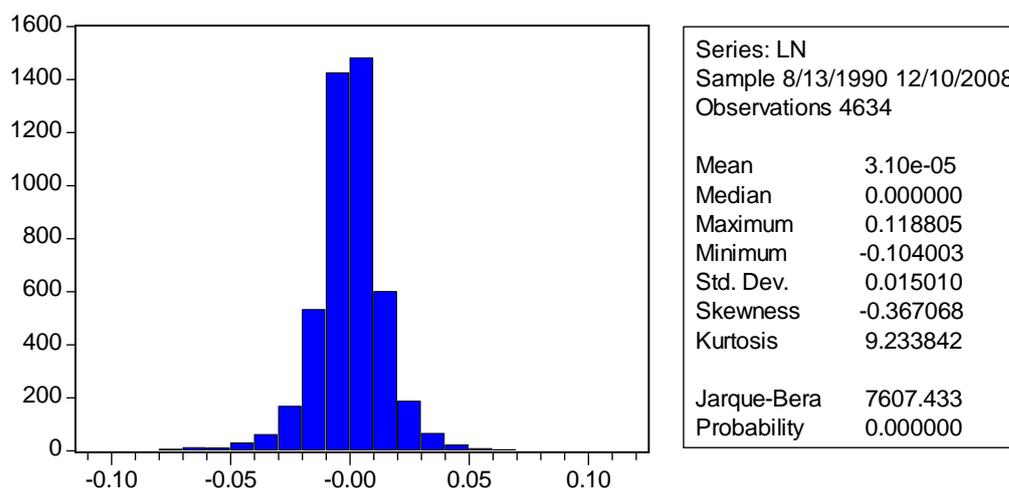


Figura 11: Histograma do Log dos Retornos Diários dos Preços do Cobre



Levando-se em consideração a probabilidade p muito próxima de zero, rejeitamos a hipótese nula de que a distribuição é normal em ambos os casos.

A aplicação da metodologia Box & Jenkins será feita, mesmo já tendo identificado que as séries não se comportam como uma normal. Como inconveniente, pode-se encontrar resultados de baixa qualidade. Entretanto, o impacto da violação desse pressuposto é mais intenso em modelos que levam em consideração o risco também, o que não acontece na metodologia Box & Jenkins. Além disso, o objetivo desse trabalho é testar a aplicabilidade do modelo Box & Jenkins, não sendo razoável interromper o estudo em função da não normalidade. Mesmo com essa restrição, algumas conclusões podem ser obtidas.

4.3. Estacionaridade

O teste de estacionaridade foi feito através do teste aumentado de Dickey-Fuller, também, acompanhado de seu p-valor, ou seja, a probabilidade da série apresentar uma raiz unitária. O intervalo de confiança adotado foi o mesmo do teste de normalidade, 99%.

Como no momento do teste, não é sabido qual o modelo auto-regressivo mais adequado para explicar o comportamento de cada série temporal, a defasagem utilizada foi escolhida automaticamente pelo Eviews 5.0 que utilizou o método SIC de escolha. Os resultados encontrados estão mostrados a seguir:

Figura 12: Teste Aumentado de Dickey-Fuller para a Série Ln dos Preços da Commodity Cobre

Null Hypothesis: LN has a unit root

Exogenous: Constant

Lag Length: 0 (Automatic based on SIC, MAXLAG=31)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-72.03138	0.0001
Test critical values:		
1% level	-3.431579	
5% level	-2.861968	
10% level	-2.567041	

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(LN)

Method: Least Squares

Date: 12/23/08 Time: 13:41

Sample (adjusted): 8/14/1990 12/10/2008
 Included observations: 4633 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN(-1)	-1.057275	0.014678	-72.03138	0.0000
C	3.24E-05	0.000220	0.147074	0.8831
R-squared	0.528388	Mean dependent var		6.97E-06
Adjusted R-squared	0.528286	S.D. dependent var		0.021824
S.E. of regression	0.014989	Akaike info criterion		-5.562575
Sum squared resid	1.040440	Schwarz criterion		-5.559795
Log likelihood	12887.71	F-statistic		5188.519
Durbin-Watson stat	2.001351	Prob(F-statistic)		0.000000

No teste ADF, a série analisada não apresentou a existência de uma raiz unitária, ou seja, a amostra pode ser considerada como estacionária. De acordo com Makridakis et al (1998), estacionaridade significa que não existe crescimento ou declínio da amostra ao longo do tempo, simplificando, os valores flutuam ao longo de uma média constante, independente do tempo.

4.4. Auto-correlação

Através da utilização do teste de Ljung-Box, analisamos a dependência entre os preços presentes e futuros com os seus valores passados. Foi considerado um intervalo de confiança de 95%. A seguir é mostrado o teste:

Figura 13: Teste de Ljung-Box

Date: 12/23/08 Time: 13:43

Sample: 8/13/1990 12/10/2008

Included observations: 4634

Autocorrelation		Partial Correlation		AC	PAC	Q-Stat	Prob	
				1	-0.056	-0.056	14.607	0.000
				2	-0.022	-0.025	16.804	0.000
				3	-0.015	-0.017	17.780	0.000
				4	0.049	0.047	29.068	0.000
				5	0.011	0.016	29.617	0.000
				6	0.020	0.024	31.512	0.000
				7	0.049	0.054	42.675	0.000
				8	0.011	0.016	43.230	0.000
				9	-0.031	-0.028	47.690	0.000
	*		*	10	0.075	0.072	73.707	0.000
				11	-0.001	0.001	73.711	0.000
				12	0.030	0.030	77.825	0.000
				13	0.002	0.008	77.840	0.000
				14	0.009	0.001	78.179	0.000
				15	0.007	0.007	78.410	0.000
				16	0.037	0.035	84.762	0.000
				17	-0.018	-0.022	86.308	0.000
				18	0.031	0.027	90.828	0.000
				19	0.003	0.006	90.858	0.000
				20	0.018	0.009	92.387	0.000
				21	0.013	0.018	93.170	0.000
				22	-0.019	-0.025	94.790	0.000
				23	0.001	-0.005	94.800	0.000
				24	-0.012	-0.016	95.520	0.000
				25	0.009	0.002	95.901	0.000
				26	-0.008	-0.015	96.217	0.000
				27	0.017	0.019	97.588	0.000
				28	0.006	0.002	97.771	0.000
				29	0.013	0.017	98.528	0.000
				30	-0.023	-0.022	101.10	0.000
				31	0.017	0.011	102.46	0.000
				32	0.004	0.005	102.52	0.000
				33	-0.007	-0.008	102.74	0.000
				34	0.008	0.008	103.00	0.000
				35	-0.007	-0.009	103.22	0.000
				36	0.016	0.016	104.48	0.000

Analisando o teste, percebe-se que existe auto-correlação a partir da primeira defasagem, ou seja, o valor imediatamente anterior influencia o preço seguinte. Em outras palavras, parte do preço de um determinado dia, depende, ou é explicado, pelo preço do cobre no dia anterior. A intensidade dessa correlação é de aproximadamente 6%, pequena, mas não desprezível.

O valor negativo da primeira defasagem explica a tendência de inversão da curva log-normal do preço no cobre. O valor futuro será positivo e irá se alternando na seqüência.

4.5. Escolha da melhor equação de regressão

Como na visualização do correlograma não conseguimos visualizar nada parecido com nenhum modelo. Não existem decaimentos exponenciais nem cortes bruscos as FAC e FACP, não sendo possível identificar o melhor modelo a ser aplicado. De acordo com a metodologia, deve-se diferenciar a série até que o correlograma apresente um comportamento que permita a identificação da regressão. Utilizaremos então, o correlograma da primeira derivada para esse fim, como é mostrado a seguir:

Figura 14: Correlograma da Primeira Derivada

Date: 12/23/08 Time: 13:43

Sample: 8/13/1990 12/10/2008

Included observations: 4633

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
****	****	1	-0.515	-0.515	1228.3	0.000
	***	2	0.012	-0.344	1228.9	0.000
	**	3	-0.028	-0.304	1232.7	0.000
	**	4	0.048	-0.212	1243.4	0.000
	*	5	-0.022	-0.183	1245.6	0.000
	*	6	-0.009	-0.182	1246.0	0.000
	*	7	0.031	-0.125	1250.4	0.000
	*	8	0.002	-0.073	1250.5	0.000
	*	9	-0.069	-0.160	1272.9	0.000
*	*	10	0.086	-0.076	1307.4	0.000
	*	11	-0.050	-0.096	1318.9	0.000
	*	12	0.030	-0.061	1323.2	0.000
		13	-0.019	-0.049	1324.9	0.000
		14	0.003	-0.052	1324.9	0.000
	*	15	-0.010	-0.066	1325.3	0.000
		16	0.038	-0.001	1332.1	0.000
		17	-0.050	-0.044	1343.5	0.000
		18	0.037	-0.014	1350.0	0.000
		19	-0.022	-0.014	1352.2	0.000
		20	0.010	-0.018	1352.7	0.000
		21	0.011	0.024	1353.2	0.000
		22	-0.027	-0.007	1356.6	0.000
		23	0.016	-0.004	1357.8	0.000
		24	-0.015	-0.026	1358.9	0.000
		25	0.018	-0.015	1360.4	0.000
		26	-0.018	-0.048	1361.9	0.000
		27	0.016	-0.032	1363.1	0.000
		28	-0.008	-0.048	1363.4	0.000
		29	0.021	-0.006	1365.5	0.000
		30	-0.037	-0.040	1371.9	0.000
		31	0.025	-0.034	1374.7	0.000
		32	-0.001	-0.020	1374.7	0.000
		33	-0.012	-0.037	1375.4	0.000
		34	0.013	-0.020	1376.2	0.000
		35	-0.016	-0.042	1377.5	0.000
		36	0.018	-0.026	1378.9	0.000

Analisando o correlograma, percebe-se que se trata de um modelo ARIMA, pois só é possível identificá-lo a partir da 1 derivada, logo o d é 1. Como na FAC acontece um corte brusco depois da defasagem q (1 também) e na FACP o comportamento decai exponencialmente ($p = 0$), concluímos que por essa análise o melhor modelo seria o ARIMA(0,1,1)

Para confirmar a afirmação anterior, faremos a regressão para os 6 modelos mais usuais e verificaremos o seu AIC. Os modelos são: AR(1); MA(1); ARMA(1,1); ARIMA(1,1,0); ARIMA(0,1,1); ARIMA(1,1,1). Os resultados completos estão expostos no anexo.

Para simplificar a análise, na tabela a seguir, encontram-se os valores de AIC de cada regressão:

Tabela 4: Valores de AIC e Raj^2 para os diferentes modelos

Modelo	AIC	Raj^2
AR(1)	-5,563	0,3%
MA(1)	-5,563	0,3%
ARMA(1,1)	-5,563	0,4%
ARIMA(1,1,0)	-6,551	62,9%
ARIMA(0,1,1)	-6,952	75,1%
ARIMA(1,1,1)	-6,953	75,2%

Com esta tabela concluímos que as regressões AR(1), MA(1) e ARMA(1,1) não são úteis para esse modelo. A regressão ARIMA (1,1,0) tem resultados medianos e os ARIMA(0,1,1) e ARIMA(1,1,1) conseguiram bons resultados de AIC e também de Raj^2 . Logo os resultados encontrados pelo correlograma e pelo AIC são semelhantes, pois ambos mostram o modelo ARIMA (0,1,1) como um dos mais indicados. Confirma-se assim, a teoria de Box & Jenkins no que diz respeito à seleção e identificação dos modelos. Concluímos que o AIC se apresentou como uma metodologia eficaz para a seleção do melhor modelo.

A identificação do modelo através dos correlogramas, também, se mostrou eficaz, entretanto, este processo pode ser bastante trabalhoso. Por exemplo, se estivéssemos analisando várias commodities, teríamos que fazer um correlograma para cada commodity com suas respectivas defasagens, se necessário.

Analisando o correlograma dos resíduos, não encontramos correlação serial entre eles, como é visto a seguir para o modelo ARIMA (0,1,1).

Figura 15: Correlograma dos Resíduos

Date: 12/23/08 Time: 14:15

Sample: 8/15/1990 12/10/2008

Included observations: 4632

Q-statistic
probabilities adjusted
for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.001	-0.001	0.0066	
		2	-0.038	-0.038	6.7516	
		3	-0.001	-0.001	6.7535	0.009
		4	0.038	0.037	13.421	0.001
		5	0.027	0.027	16.848	0.001
		6	0.010	0.013	17.352	0.002
	*	7	0.064	0.066	36.210	0.000
		8	-0.000	-0.001	36.211	0.000
		9	-0.016	-0.013	37.395	0.000
		10	0.063	0.062	55.941	0.000
		11	0.015	0.008	56.918	0.000
		12	0.020	0.021	58.754	0.000
		13	0.014	0.015	59.641	0.000
		14	-0.001	-0.007	59.644	0.000
		15	0.018	0.016	61.189	0.000
		16	0.029	0.028	65.210	0.000
		17	-0.009	-0.019	65.625	0.000
		18	0.026	0.026	68.748	0.000
		19	0.010	0.008	69.258	0.000
		20	0.014	0.008	70.208	0.000
		21	0.020	0.020	72.038	0.000
		22	-0.022	-0.028	74.394	0.000
		23	0.003	-0.003	74.448	0.000
		24	-0.015	-0.016	75.472	0.000
		25	0.009	0.002	75.880	0.000
		26	-0.009	-0.015	76.241	0.000
		27	0.019	0.021	77.957	0.000
		28	0.006	-0.000	78.117	0.000
		29	0.015	0.018	79.098	0.000
		30	-0.024	-0.025	81.864	0.000
		31	0.019	0.016	83.528	0.000
		32	0.002	0.000	83.541	0.000
		33	-0.004	-0.005	83.635	0.000
		34	0.005	0.004	83.742	0.000
		35	-0.004	-0.005	83.803	0.000
		36	0.014	0.013	84.764	0.000

4.6. Box & Jenkins como modelo de previsão

Utilizou-se o modelo de previsão dentro da própria amostra já ajustada, tentando de certa forma conseguir os melhores resultados possíveis do modelo.

Como era de se esperar, o modelo Box & Jenkins sozinho, não se mostrou um bom previsor. Sendo ele apenas um componente de um modelo com outras variáveis. Utilizamos apenas um modelo de previsão para curto prazo, para as 8 últimas observações da amostra. Um modelo de longo prazo não teria muito sentido, já que o modelo voltará para a sua média, ficando constante depois. Através da tabela dos diferentes modelos e do seu MAPE a seguir, consegue-se visualizar melhor:

Tabela 5: MAPE de cada modelo de regressão

Modelo	MAPE
AR(1)	102
MA(1)	102
ARMA(1,1)	103
ARIMA(1,1,0)	51
ARIMA(0,1,1)	48
ARIMA(1,1,1)	46

Por essa tabela, vemos que o modelo Box & Jenkins não é muito útil para se calcular preços futuros com precisão devido ao alto valor do MAPE. Percebe-se também, que os três modelos ARIMA seriam os mais razoáveis levando-se em consideração o seu menor MAPE, confirmando a escolha do modelo pelo correlograma: ARIMA(0,1,1); e pelo AIC: ARIMA (0,1,1) e ARIMA(1,1,1). O modelo ARIMA (1,1,0) foi o terceiro melhor considerando-se o AIC. Os resultados completos desse teste encontram-se em anexo. Para melhor entendimento podemos comparar os valores reais com os valores previstos pelos modelos ARIMA através da tabela a seguir:

Tabela 6: Valores de previsão x valores reais dos modelos ARIMA

Data	Real	ARIMA(0,1,1)	ARIMA(1,1,0)	ARIMA(1,1,1)
12/01/2008	(0.008322)	-0.002208	-0.002208	-0.010328
12/02/2008	-0.009797	-0.007688	-0.007448	0.001831
12/03/2008	-0.031141	-0.013528	-0.014855	-0.021830
12/04/2008	-0.052424	-0.020287	-0.020082	-0.019051
12/05/2008	-0.069648	-0.025077	-0.027878	-0.039924
12/08/2008	0.083316	0.054055	0.048392	0.049848
12/09/2008	-0.035307	-0.044790	-0.056220	-0.026720
12/10/2008	0.032286	0.038524	0.044902	0.024452

Isso demonstra que apesar dos preços futuros serem dependentes do passado, os preços e erros passados, tão somente, não são capazes de prever os preços futuros. Sendo necessário, então, adicionar outras variáveis e componentes aos modelos ARIMA, para que se tenha um resultado mais satisfatório.