JMP Academic Case Study 054

Forecasting Copper Prices

Time Series, Exponential Smoothing, Forecasting

Produced by

M Ajoy Kumar, Siddaganga Institute of Technology
ajoy@sit.ac.in

Muralidhara A, JMP Global Academic Team
muralidhara.a@jmp.com

JMP STATISTICAL DISCOVERY

Forecasting Copper Prices

Time Series, Exponential Smoothing, Forecasting

Key Ideas

The case study deals with forecasting a univariate time series. Before the time series are modeled and forecasted, it is necessary to observe the patterns and trends in the data. It is better to smooth out the data before forecasting. Time-series data may exhibit a trend over the period or seasonality and sometimes both period and seasonality. Depending on the pattern of the data, different smoothing techniques are applied before forecasting. This case deals with different exponential smoothing techniques available in JMP.

Background

Copper is a base metal with a vast number of industrial applications. The copper prices in India were found to be highly volatile during 2017-18. The price was around ₹375 per kilogram at the beginning of 2017. During the two-year period, at one time the price touched a maximum of ₹490, plummeted to ₹350, and closed at ₹415 by end of 2018. Arun, a researcher at the Commodity Research Centre of a leading investment advisory firm in Mumbai, was trying to understand the behavior of copper prices. He wanted to forecast the price using an appropriate technique.

The Task

Arun performed the following tasks:

- Collected daily copper prices for two years
- Analyzed the pattern of price movements
- Applied various smoothing techniques
- Selected the most suitable model based on information criteria

The Data **copper.jmp**

Arun collected the daily spot price of copper from Bloomberg for two years, from Jan. 1, 2017, to Dec. 31, 2018. The data set is saved as Copper.jmp. There are 475 observations. The data set has two series: date and daily copper prices.

Date: Day on which the copper price is collected
Copper: Copper price on a specific date/day

Copper price is a continuous time-series variable, whereas Date is time-variable.

Analysis

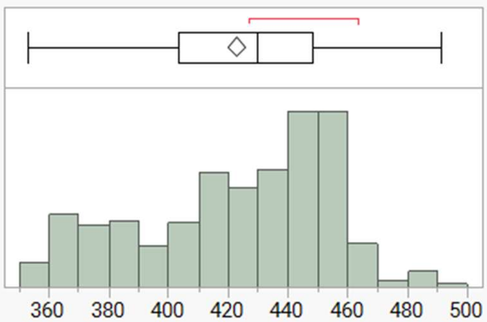
Descriptive Statistics

Arun explored the data using Distribution and Graph Builder.

Exhibit 1 Summary Statistics of Copper Prices

Distributions Student Edition

Copper



Quantiles

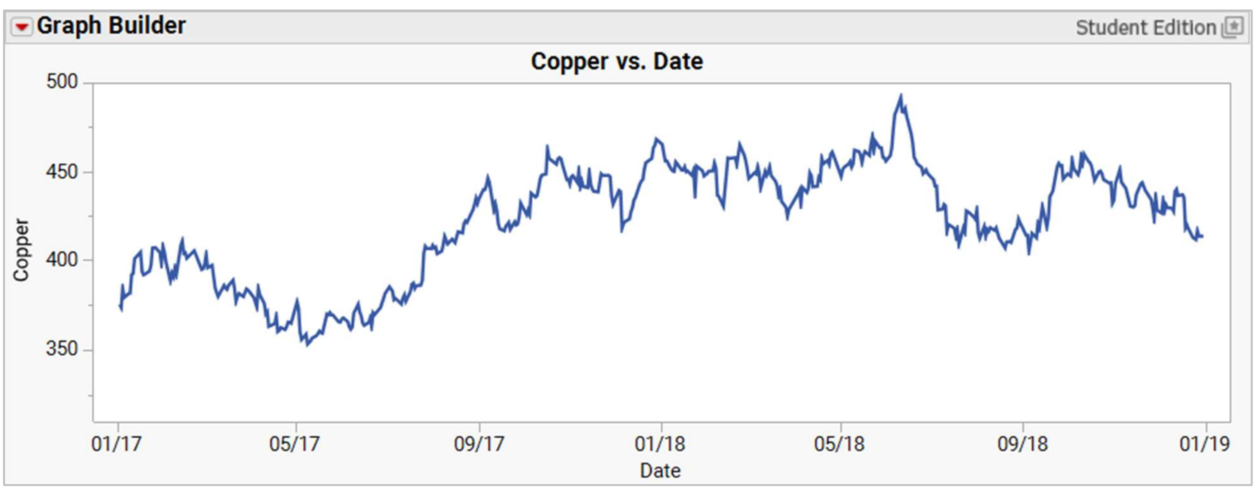
100.0%	maximum	491.25
99.5%		484.675
97.5%		466.925
90.0%		457.29
75.0%	quartile	448.65
50.0%	median	429.75
25.0%	quartile	403.2
10.0%		372.27
2.5%		360.27
0.5%		355.054
0.0%	minimum	353

Summary Statistics

Mean	422.93966
Std Dev	31.528757
Std Err Mean	1.4466386
Upper 95% Mean	425.78228
Lower 95% Mean	420.09705
N	475
Skewness	-0.503448
Kurtosis	-0.772115
N Missing	0
Minimum	353
Maximum	491.25

(Analyze > Distribution > Y = Copper > OK. Under the red triangle next to Distributions, select Stack. This will align the output horizontally. Under the red triangle next to Summary Statistics, select Customize Summary Statistics, and select Skewness, Kurtosis, Minimum and Maximum > OK.)

Exhibit 2 Movement of Daily Copper Price (2017-2018)



(Graph > Graph Builder > Y = Copper, X = Date, Select Line graph from the Chart options > Done.)

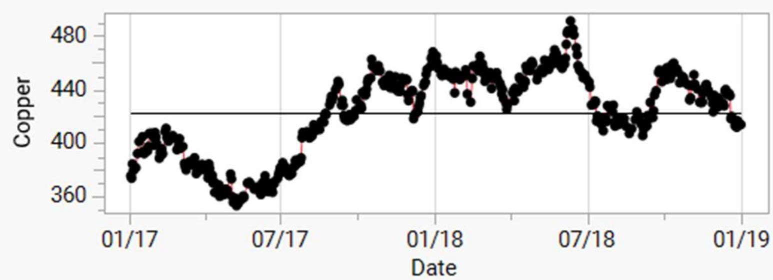
The basic descriptive characteristics of the data are presented in Exhibit 1, and the graph showing the movement of copper prices during the two years is given in Exhibit 2. The mean copper price was ₹422.94, with a maximum of ₹491.25 and a minimum of ₹353.00. The prices were negatively skewed with a negative kurtosis. It can be observed from Exhibit 2 that copper prices are highly volatile during the period. There is an overall increase in price during the period, but in between downward as well as upward movements are observed in the prices, which indicates the presence of some seasonality. The graph also indicates the need for applying smoothing techniques to the data before forecasting.

Time-Series Modeling

Before applying smoothing techniques to the data, Arun developed basic time-series plots. Time-series modeling is available under Analyze > Specialized Modeling. The default forecast period is 25, which can be changed as per the requirements.

Exhibit 3 Time-Series Output of Copper Prices

Time Series Copper Student Edition



Mean	422.93966
Std	31.495551
N	475
Zero Mean ADF	0.1741613
Single Mean ADF	-2.131604
Trend ADF	-2.06343

Time Series Basic Diagnostics

Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	0.9835		462.319	<.0001*	1	0.9835	
2	0.9691		912.179	<.0001*	2	0.0581	
3	0.9555		1350.43	<.0001*	3	0.0206	
4	0.9418		1777.14	<.0001*	4	-0.0058	
5	0.9282		2192.44	<.0001*	5	-0.0064	
6	0.9159		2597.66	<.0001*	6	0.0344	
7	0.9043		2993.58	<.0001*	7	0.0228	
8	0.8926		3380.18	<.0001*	8	-0.0047	
9	0.8831		3759.36	<.0001*	9	0.0601	
10	0.8755		4132.87	<.0001*	10	0.0645	
11	0.8651		4498.35	<.0001*	11	-0.0773	
12	0.8563		4857.14	<.0001*	12	0.0361	
13	0.8456		5207.80	<.0001*	13	-0.0570	
14	0.8342		5549.84	<.0001*	14	-0.0269	
15	0.8256		5885.59	<.0001*	15	0.0833	
16	0.8185		6216.28	<.0001*	16	0.0481	
17	0.8093		6540.33	<.0001*	17	-0.0533	
18	0.7999		6857.54	<.0001*	18	-0.0126	
19	0.7892		7167.04	<.0001*	19	-0.0577	
20	0.7791		7469.27	<.0001*	20	0.0032	
21	0.7700		7765.17	<.0001*	21	0.0449	
22	0.7604		8054.34	<.0001*	22	-0.0359	
23	0.7498		8336.13	<.0001*	23	-0.0219	
24	0.7392		8610.67	<.0001*	24	0.0053	
25	0.7283		8877.77	<.0001*	25	-0.0359	

(Analyze > Specialized Modeling > Y, Time Series = Copper, X, Time ID = Date, > OK.)

Exhibit 3 shows the descriptive statistics, autocorrelation function (ACF), and partial autocorrelation function (PACF) for copper prices. The ACF plot is not decaying and the PACF shows spikes up to the first lag. This indicates that the price series is not stationary. Now Arun could start to smooth the price curve.

Exponential Smoothing Techniques

One of the basic methods of smoothing a curve is applying a simple moving average, which attaches equal weight to all the observations in the averaging process. But exponential smoothing techniques give higher weights to more recent observations during averaging and are thus better.

Smoothing models are defined as follows:

$$Y_t = \mu_t + \beta_t t + s(t) + a_t$$

μ_t = time-varying mean term
β_t = time-varying slope term
$S(t)$ = one of the s time-varying seasonal terms
a_t = random shocks

Models without a trend have $\beta_t = 0$ and nonseasonal models have $S(t) = 0$. The estimators for these time-varying terms are defined as follows:

L_t is a smoothed level that estimates μ_t
T_t is a smoothed trend that estimates β_t
S_{t-j} for j = 0, 1,..., s - 1 are estimates of the $s(t)$

Each smoothing model defines a set of recursive smoothing equations that describe the evolution of these estimators. The smoothing equations are written in terms of model parameters called smoothing weights:

α is the level smoothing weight
ϒ is the trend smoothing weight
φ is the trend damping weight
δ is the seasonal smoothing weight

While these parameters enter each model in different ways, they have the common property that larger weights are assigned to more recent data. Apart from simple moving average and state-space smoothing, JMP offers six exponential smoothing models. The table below lists these models along with model specifications.

#	**Name**	**Model**
1	Simple Exponential Smoothing	$Y_t = \mu_t + a_t$
2	Double Exponential Smoothing	$Y_t = \mu_t + \beta_1 t + a_t$
3	Linear (Holt) Exponential Smoothing	$Y_t = \mu_t + \beta_t t + a_t$
4	Damped Trend Linear Exponential Smoothing	$Y_t = \mu_t + \beta_t t + a_t$
5	Seasonal Exponential Smoothing	$Y_t = \mu_t + s(t) + a_t$
6	Winter's Model (Additive)	$Y_t = \mu_t + \beta_t t + s(t) + a_t$

(For details on exponential smoothing, please refer the JMP documentation guide under Help menu or visit https://www.jmp.com/en_us/support/jmp-documentation.htmli)

Arun started with Simple Exponential Smoothing. Exhibit 4 shows the result of simple exponential smoothing and includes a model summary, parameter estimates, and the forecast graph.

Exhibit 4 Simple Exponential Smoothing

Model: Simple Exponential Smoothing(Zero to One)

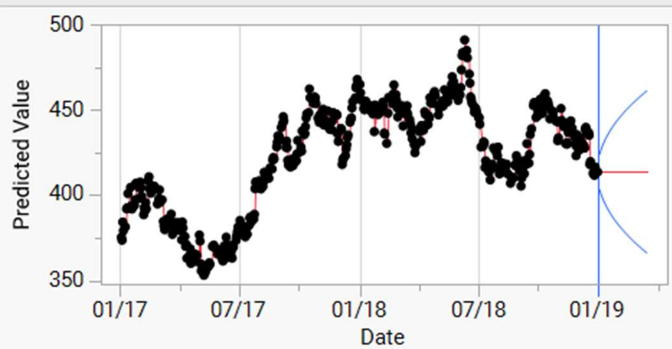
Model Summary

DF	473	Stable	Yes
Sum of Squared Innovations	13150.419	Invertible	Yes
Sum of Squared Residuals	13150.4448		
Variance Estimate	27.8021543		
Standard Deviation	5.27277482		
Akaike's 'A' Information Criterion	2922.26331		
Schwarz's Bayesian Criterion	2926.42452		
RSquare	0.97195565		
RSquare Adj	0.97195565		
MAPE	0.93080893		
MAE	3.93571812		
-2LogLikelihood	2920.26331		

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.91747523	0.0453475	20.23	<.0001*

Forecast



(Under the red triangle option, choose Smoothing Model > Simple Exponential Smoothing.)

Arun followed the same procedure for the remaining exponential smoothing techniques. Prediction Intervals and Constraints are two inputs required for exponential smoothing techniques. Arun maintained the default settings for these two parameters. Every time a new model is estimated, JMP provides a Model Comparison listing. Model comparison for all the models estimated by Arun is shown in Exhibit 5.

Exhibit 5 Model Comparison

Model Comparison

Report	Graph	Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	Weights	.2 .4 .6 .8	MAPE	MAE
		Seasonal Exponential Smoothing(12, Zero to One)	460	27.572602	2889.5889	2897.8600	0.969	2885.5889	0.937870		0.989641	4.170917
		Winters Method (Additive)	459	27.839033	2895.0176	2907.4243	0.969	2889.0176	0.062130		0.995273	4.194736
		Simple Exponential Smoothing(Zero to One)	473	27.802154	2922.2633	2926.4245	0.972	2920.2633	0.000000		0.930809	3.935718
		Linear (Holt) Exponential Smoothing	471	27.913236	2925.2994	2933.6176	0.971	2921.2994	0.000000		0.942849	3.983408
		Damped-Trend Linear Exponential Smoothing	471	27.92021	2926.2633	2938.7469	0.972	2920.2633	0.000000		0.930809	3.935718
		Double (Brown) Exponential Smoothing	472	32.681429	2993.9151	2998.0742	0.967	2991.9151	0.000000		1.036056	4.371689

Arun found that the model that best fits the data was Seasonal Exponential Smoothing. This can be identified by comparing the Akaike's information criterion (AIC) or Schwarz's Bayesian criterion (SBC) values. The model that minimizes the information criteria value is the most suitable one. The Seasonal Exponential Smoothing technique has both the lowest AIC and SBC values.

Forecasting Copper Prices

After this, Arun moved to forecasting the copper prices using Seasonal Exponential Smoothing. The parameter estimates and forecast graph are given in Exhibit 6.

Exhibit 6 Forecast with Seasonal Exponential Smoothing

Model: Seasonal Exponential Smoothing(12, Zero to One)

Model Summary

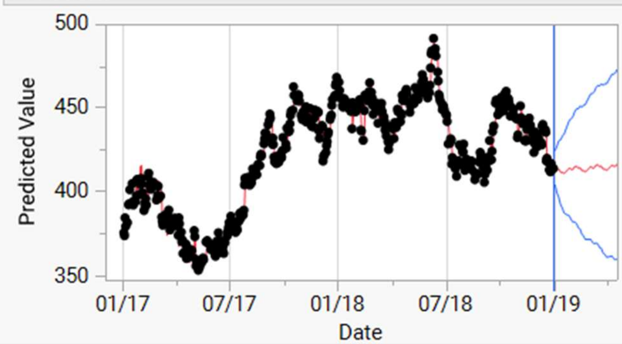
DF	460	Stable	Yes
Sum of Squared Innovations	12683.3969	Invertible	No
Sum of Squared Residuals	14160.3759		
Variance Estimate	27.572602		
Standard Deviation	5.250962		
Akaike's 'A' Information Criterion	2889.58886		
Schwarz's Bayesian Criterion	2897.85998		
RSquare	0.96879297		
RSquare Adj	0.96872513		
MAPE	0.98964082		
MAE	4.17091726		
-2LogLikelihood	2885.58886		

Failed: Cannot Decrease Objective Function Hessian is not positive definite.

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.91468328	0.0458612	19.94	<.0001*
Seasonal Smoothing Weight	3.8564e-12	0.0000000	.	.

Forecast



Once the model was estimated, Arun generated a new data table containing the actual observations, predicted values, and standard errors. The last 25 observations in the data table represent the forecast of copper prices for the first 25 days of 2019. This data is shown in Exhibit 7.

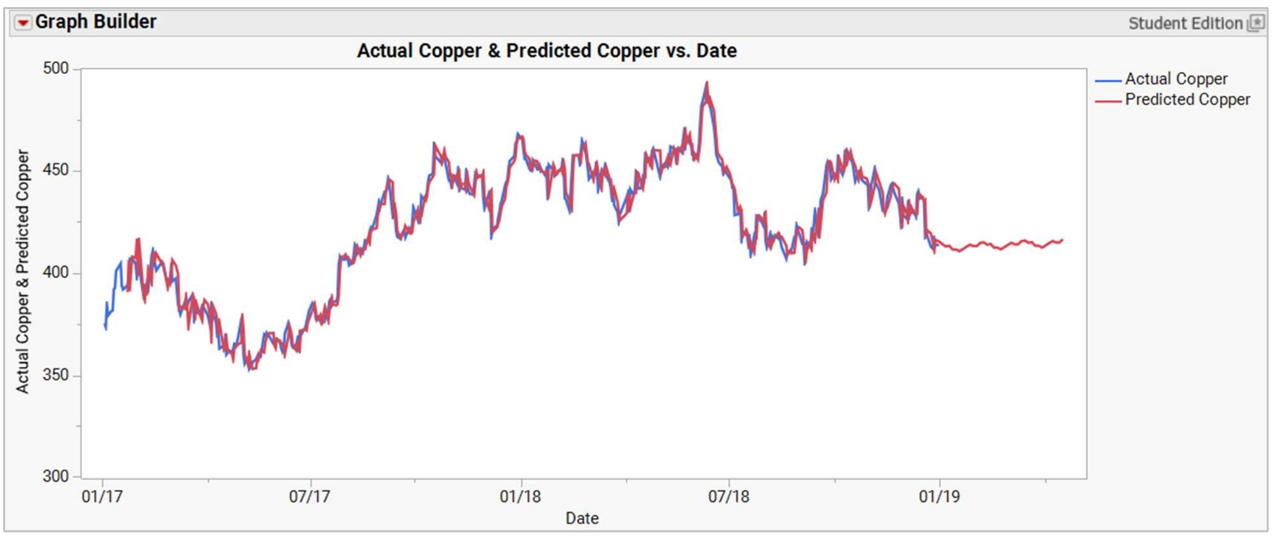
Exhibit 7 Forecast Values of Copper Prices

	Actual Copper	Date	Predicted Copper	Std Err Pred Copper	Residual Copper	Upper CL (0.95) Copper	Lower CL (0.95) Copper	Innovations Copper
496	•	04/03/19	414.76744699	22.112043596	•	458.10625606	371.42863791	•
497	•	07/03/19	414.04119699	22.627659302	•	458.39059427	369.6917997	•
498	•	10/03/19	414.07494699	23.131784605	•	459.41241171	368.73748226	•
499	•	13/03/19	415.62119699	23.625155079	•	461.92565007	369.3167439	•
500	•	16/03/19	415.89288996	24.108431015	•	463.14454647	368.64123344	•
501	•	19/03/19	414.76981303	24.582207783	•	462.95005495	366.58957112	•
502	•	22/03/19	415.0890438	25.047024434	•	464.18030961	365.99777799	•
503	•	25/03/19	413.27878739	25.503370885	•	463.26447581	363.29309897	•
504	•	28/03/19	413.28006944	25.951693972	•	464.14445497	362.41568392	•
505	•	31/03/19	412.4562826	26.392402572	•	464.1844411	360.72812409	•
506	•	03/04/19	413.4950326	26.825871971	•	466.07277551	360.91728968	•
507	•	06/04/19	414.6347826	27.252447606	•	468.04859839	361.2209668	•
508	•	09/04/19	415.6550326	27.672448284	•	469.8920346	361.41803059	•
509	•	12/04/19	414.9287826	28.086168972	•	469.97666224	359.88090295	•
510	•	15/04/19	414.9625326	28.493883221	•	470.80951749	359.1155477	•
511	•	18/04/19	416.5087826	28.895845281	•	473.14359865	359.87396654	•

(Under the red triangle of Model: Seasonal Exponential Smoothing (12, Zero to One), choose Save Columns. This will create a new data table containing the actual and predicted values together with standard errors, residuals, and 95% prediction intervals.)

Arun also generated a graph showing the actual observations and predicted values using Graph Builder, which is shown in Exhibit 8.

Exhibit 8 Graph of Actual and Predicted Values



Summary

Statistical Insights

To summarize, in this case, the scheme of analysis using JMP involved the following:

- Generating summary statistics of the data series
- Visual representation of the data using Graph Builder
- Time-series representation of the data
- Estimation of various smoothing models
- Selection of the best smoothing models using model comparison and information criteria
- Forecasting using the most suitable smoothing model
- Generating forecast values and visualizing actual and predicted values

Implications

Arun drew the following conclusions from the analysis:

- The copper prices in India exhibited seasonality during 2017-18.
- Seasonal Exponential Smoothing is the most suitable model for understanding the pattern and forecasting copper prices during 2017-18.

JMP® Features and Hints

This case used the Time Series capability of the Analyze platform for forecasting. It also used various exponential smoothing features of the time-series analysis. Graph Builder was explored to visualize the time-series data.

Exercise

The price of lead for 2017-18 is available in Zinc.jmp. Perform the scheme of analysis explained in this case. Identify the most suitable exponential smoothing technique and forecast lead prices for the first 25 days of 2019.