Forecast Fit vs. Forecast Error

- Clarifying the Concepts
- Understanding the Value

SPINNAKER

Spinnaker is a client advocate and is in no way sponsored by any software vendor.

Fit vs. Error
Anyone in a position to improve supply chain operations by influencing demand planning should understand the difference between fit and error…

After years of working in supply chain operations — first in industry, then as consultants helping corporations and clients better manage their demand planning processes and technologies — it's still surprising to hear so many demand planners misuse the terms *fit* and *error*.

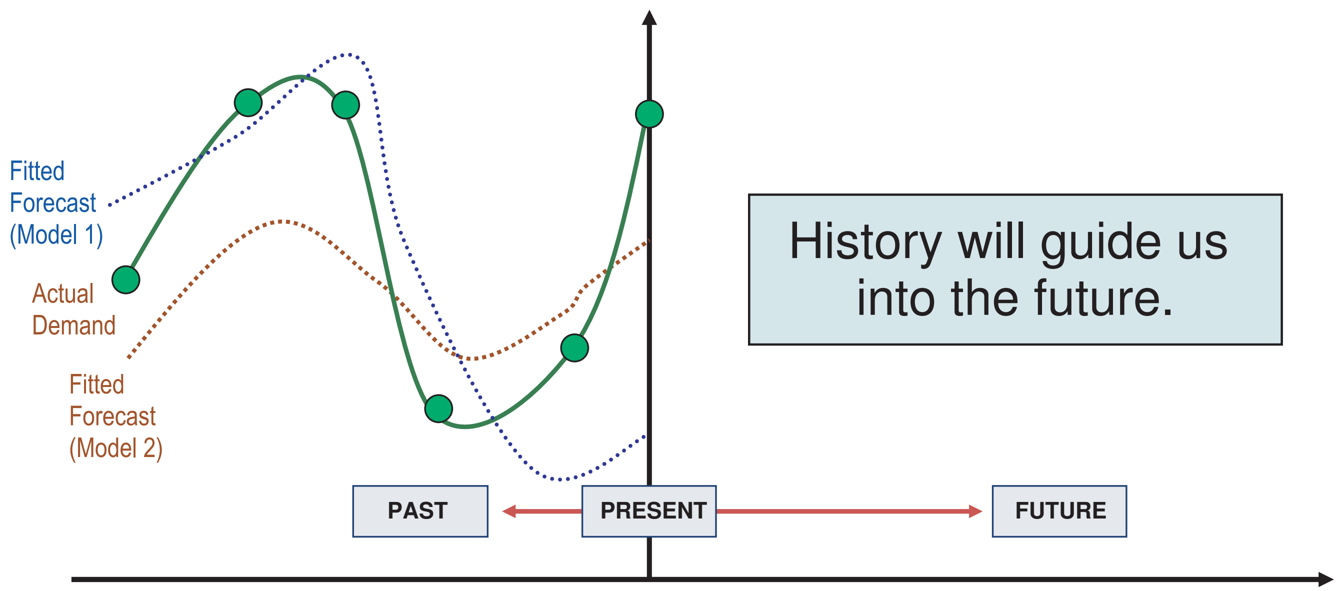
The concepts are relatively straightforward, yet they're quite distinct, and using one in place of the other is incorrect. Anyone in a position to improve supply chain operations by influencing demand planning should understand the difference between fit and error and be mindful of both the technical and functional implications these terms have in the world of forecasting.

By Definition…

Forecast fit describes the relative difference between actual historical data and a hypothetical forecast generated by a statistical model (or algorithm) using that same historical data as input. It's quite literally a backward-looking assessment of how closely a forecast created by any one of various statistical models would stack up against — or "fit" when compared to — actual historical demand.

Planners use forecast fit to project the suitability of one or more statistical forecasting algorithms to accurately forecast future demand (see **Figure 1**).

Figure 1: Forecast Fit



After running statistical forecasting models using actual historical demand data as input, planners analyze the "fit" of the resulting fitted forecasts to determine which model will likely yield forecasts that most closely match past demand patterns.

Forecast Error

Forecast error is defined by APICS as “the difference between actual and forecast demand, stated as an absolute value or as a percentage.” Forecast error is a postmortem benchmark of the variance between demand that was *projected* and *actual* demand that subsequently occurred (see **Figure 2**).

Opportunity: The Case for Using Fit and Error

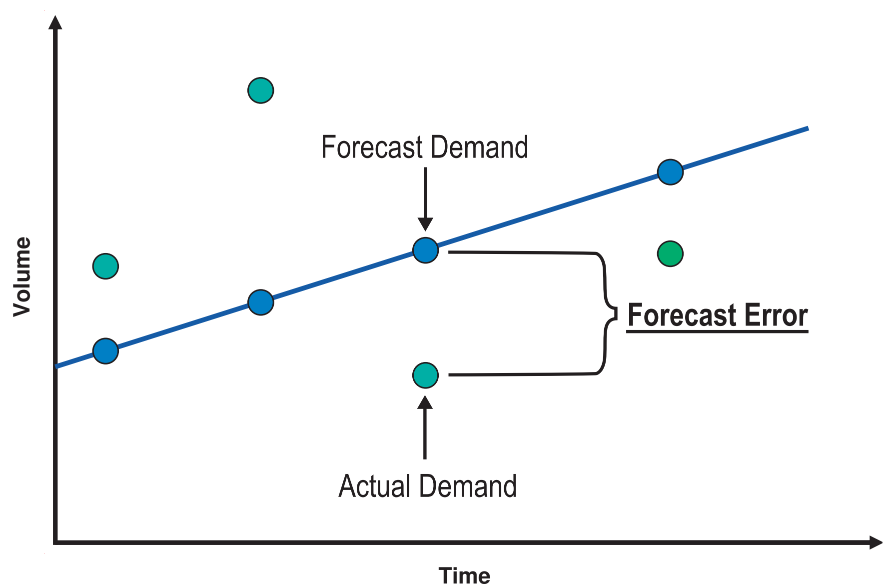
Consider the example of one Spinnaker client, a small business comprised of multiple shops that sell flowers, gift items, and packaging supplies. Business activity is highly seasonal, with peaks around Christmas, Valentine’s Day, Mother’s Day, and other signature occasions. At any given time, the manager of these stores oversees a total of 50,000 items in stock throughout 10 locations.

Until recently, he relied on historical sales data from prior years as a guide for how much product to buy for each upcoming season. As the company’s network of stores expanded over the years, however, planning in advance for expected demand became more complex, resulting in more stock outs or overstocking — problems directly related to forecasting.

Despite spending more time planning, the manager still found himself ordering either too much or too little. To streamline operations — inventory management, cash flow, etc. — company leaders decided to purchase a demand planning tool to help produce more accurate forecasts that the manager might then simply convert to purchase orders. After loading historical sales data into the forecasting tool, however, he ran into trouble configuring the package.

Figure 2: Forecast Error

The variance, expressed as an absolute value or a percentage, between forecast demand and actual demand.



Fitted Forecast

A hypothetical demand forecast that is created by a statistical forecasting algorithm using actual historical demand metrics as input data. Used for comparison's sake, fitted forecasts help planners select optimal forecasting algorithms to use in planning future demand.

Software documentation directed him simply to "choose a forecasting algorithm with the best fit." The software offered multiple statistical forecasting algorithms to choose from, as do most demand planning applications. Common algorithms include Holt-Winters, Fourier, and Box-Jenkins, but there are many other algorithms and many more that are proprietary to specific software vendors.

In addition to offering various algorithms to generate fitted forecasts based on historical demand data, as shown in **Figure 1**, demand planning software also calculates various metrics regarding the fit of each model's forecast. These metrics can be used to automatically identify the best model for forecasting future demand, or planners may choose a model themselves, based on their own judgment.

Once an algorithm is selected to produce a working forecast, and subsequent consumer demand is fulfilled over a period of time, the forecasting software can then be used to report on forecast error, using various statistical metrics to indicate how well *actual* demand stacked up against forecast demand.

A robust demand planning package can provide rich insights to help planners streamline operational efficiency, improve cost-effectiveness, and increase profitability, but **not all planners share the same level of expertise, and not all businesses have the luxury of employing skilled statisticians to drive planning using such tools**.

That was the situation our client's store manager faced, and his case was a perfect example of how **confusion about forecast fit and error can compound the task of projecting what products are likely to sell in the future**. This case also illustrates the importance of understanding the difference between fit and forecast, and their value in terms of demand forecasting.

Understanding Forecast Fit

When evaluating the fitted forecasts created by statistical forecasting algorithms, there are two fundamental questions a planner should consider:

1. Are the forecasts of good quality?

2. Considering the number of algorithms most packages offer, which will likely provide the most accurate forecasts going forward?

Analyzing the *fit* of a model forecast will help you answer both these questions, but it's also helpful to understand a bit more about how demand planning software calculates and assesses fit.

As previously described, most DP applications first apply multiple algorithms to create forecast projections based on historical

Why Forecast Error?
Clearly, forecast error is an after-the-fact measure of excellence that business leaders may use to drive process/technology improvements, or as a performance benchmark of a planner's activities — or lack thereof — to accurately forecast demand.

demand data and then assess the relative accuracy of the resulting fitted forecasts — the models "fitted" to the actual historical data.

Sometimes called *ex-post forecasting*, or forecasting after-the-fact, this exercise helps planners answer the question *"What would I have forecast had I used a particular algorithm?"* The next step — analyzing fit — answers the question *"How accurate would my forecast have been?"* Generally speaking, if one fitted forecast has lower error than another, it's reasonable to presume that the algorithm which produced the fitted forecast will yield actual forecasts with similarly low error in the future.

As shown in **Figure 1**, ex-post forecasting reveals at a glance that the forecast generated by Model 1 fits the actual demand history better than the forecast created by Model 2. To *precisely* determine the relative accuracy of fitted forecasts — especially when comparing more than one — demand planning packages automatically calculate the error of each fitted forecast. By contrasting the resulting fit measures of each algorithm, DP packages help demand planners select optimal forecasting models based on their organization's historical demand. Some demand planning applications define this automated selection capability as "best fit."

Evaluating forecast fit is a great way to determine the *potential* quality of future statistical forecasts *if* you believe that historical demand is a valid indicator of future demand for your organization.

If you have any reasonable expectation that future demand will significantly deviate from history, however — due to a projected surge in new customers, promotions, or increased competition, for example — then fit has only limited utility as a metric for helping you analyze and assess the potential value of a statistical forecasting algorithm.

Understanding Forecast Error

Once you select a preferred statistical algorithm to generate a forecast and start fulfilling subsequent demand, you'll begin accruing the data you need to calculate error and gauge the accuracy of your forecast compared to actual demand, as shown in **Figure 2**.

Clearly, forecast error is an after-the-fact measure that business leaders may use to drive process/technology improvements, or as a performance benchmark of a planner's abilities (or lack thereof) to accurately forecast demand.

Error revealed by this analysis is typically attributable to the degree of fit that a forecasting model is able to achieve with the historical data. Error may also reflect changes in business conditions that occur *after* creation of the actual forecast.

One best-practice measure often overlooked during implementation or optimization of a demand planning process is the review of forecast error. **Tracking error on a regular basis can reveal problems in your forecasting process or indicate opportunities to tune or adjust the statistical forecasting engine within your demand planning software** — to select different mathematical models, reconfigure forecasting time horizons, or change smoothing constant parameters — in an effort to improve your forecast.

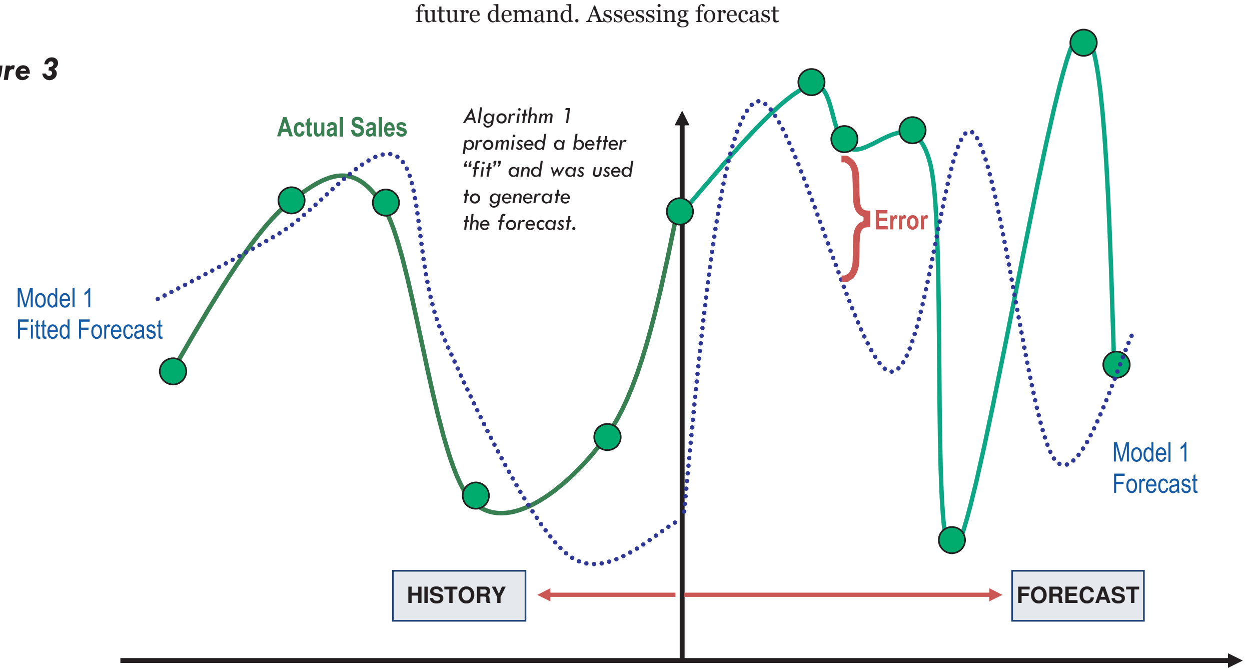
Putting it All Together: Applying Fit and Error in the Real World

Assessing forecast fit helps planners choose an optimal forecasting algorithm to use *prior* to forecasting future demand. Assessing forecast error suggests opportunities for tuning a forecast going forward, *after* actual demand has occurred.

Based on the analysis of forecast fit illustrated in **Figure 1**, which showed that Model 1 promised to provide a better fit for future forecasts than Model 2, **Figure 3** illustrates the natural outcome of that decision — how Model 1 was used to create a forecast for future demand, how subsequent demand actually stacked up in comparison to that forecast, and how the gaps between subsequent actual demand and forecast demand reveal the forecast error.

Whereas forecast fit initially provided guidance for projecting future demand, forecast error serves to guide ongoing forecast quality.

Figure 3



Forecast Fit looks at history to determine which algorithm to use.

Forecast error is the difference between actual and forecasted sales. It is used to provide a post-mortem analysis of forecast performance.

Common Measures of Error and Fit

One similarity between forecast fit and error is that most of the metrics used to measure them are the same. They're simply the mathematical differences between projected demand and actual demand. Most of the metrics are automatically calculated by the forecasting algorithm(s) during the process of generating a statistical forecast.

What determines the relevance of these metrics to either fit or error depends on when they are applied within the forecasting process — during the creation of fitted forecasts, or ex-post facto.

Some of the most common measures used to evaluate fit and error are:

- Mean Squared Error (MSE)
- Standard Error (SE)
- Coefficient of Determination or R-Squared (R^2)
- Mean Absolute Deviation (MAD) or Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MAPE)
- Symmetric Mean Absolute Percentage Error (SMAPE)

Note that there are some advanced emerging fit and error measures such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), but these are not covered as part of this paper.

Figure 4 shows a data set comprising actual historical sales for time periods 1-12 and forecasted sales for time periods 1-24 using two algorithms — Algorithm 1 and Algorithm 2. Periods 1-12 refer to *historical* time periods in the past; Periods 13-24 reference *future* time periods for which each of the two forecasting algorithms has generated a forecast but for which actual demand has yet to occur.

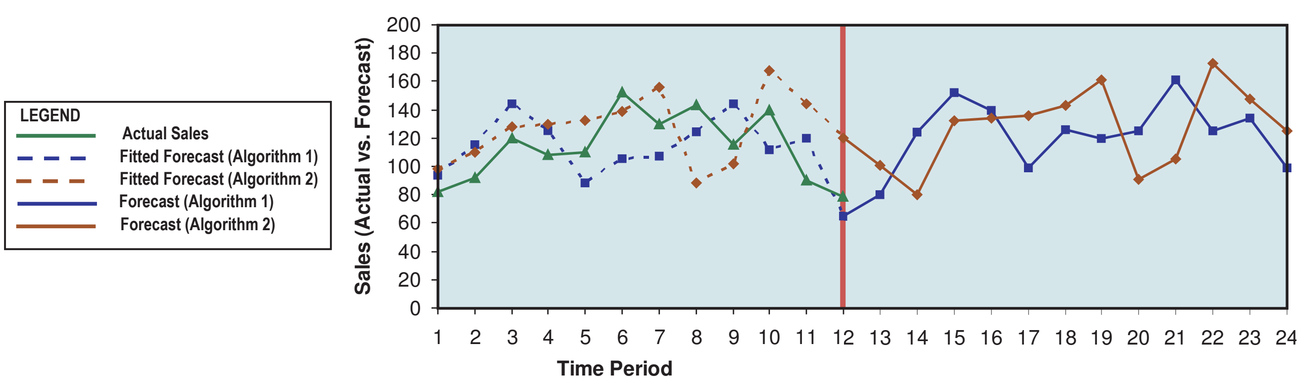
By using both the actual demand data and data from the two fitted forecasts — all from Periods 1-12 — the demand planning software package performs an ex-post forecasting analysis to evaluate the fit of the two algorithms.

***Figure 4* —**

Forecast Data Sample

Actual & Forecasted Sales Data

	Period	Actual Sales (X)		Algorithm 1 (Y)	Algorithm 2 (Y)
Historical Time Periods	1	82	Fitted Forecast	94	98
	2	92		115	110
	3	120		144	128
	4	108		125	130
	5	110		88	132
	6	152		105	139
	7	130		107	156
	8	143		124	88
	9	115		144	102
	10	140		112	168
	11	90		120	144
	12	78		65	121
Future	13	Future sales data not available yet	Forecast	80	101
	14			124	80
	15			152	132
	16			140	134
	17			99	136
	18			126	143
	19			120	161
	20			125	91
	21			161	105
	22			125	173
	23			134	148
	24			99	125



Forecast Fit

"Understanding forecast fit and its corresponding metrics helps planners tune their statistical models and improve forecast accuracy."

This analysis helps identify which of the algorithms better fits the historical data and which can potentially generate a better forecast.

The results of the mathematical analysis are summarized in the **Figure 5** table of fit and error metrics. For Algorithm 1, the detailed calculations used to determine each of those metrics are detailed in **Appendixes A**, **B**, **C**, **D**, and **E**, which are located at the end of this paper. For Algorithm 2, the results are calculated similarly but are not shown.

Improved Forecasting: The Benefits of Fit and Error

Understanding forecast fit and its corresponding metrics helps planners tune their statistical models and improve forecast accuracy. Some tools provide simulation capabilities, to help determine optimal configuration parameters, but clearly the ability to use such tools to make effective comparisons requires a reasonably clear understanding of fit and fit metrics in the first place — and how such insights serve to help make optimal decisions when it comes to forecasting.

Analyzing fit provides revealing insight into naturally occurring demand variation. It can also indicate the fundamental suitability of available historical data by revealing whether historical data really represents the future, or whether it is even statistically forecastable.

Error metrics provide the basis for analyzing forecast performance. They help planners pinpoint areas where there may be problems with the statistical forecast, or where assumptions that were valid

Figure 5 —

Fit and error metrics

	Ex-Post Forecasting Results comparison - Algorithm 1 vs. Algorithm 2		
Fit & Error Measure	**Algorithm 1**	**Algorithm 2**	**What does this result tell us about these two algorithms?**
Mean Squared Error (MSE)	**710.45**	**1018.18**	Algorithm 1 is better as it has a lower MSE i.e less dispersion of the fitted forecast from the actual sales
Standard Error (SE)	**26.65**	**31.91**	Algorithm 1 is better as it has a lower SE as compared to the average historical sales of 113 units /time period. SE of 23% vs. 28%
Coe fficient of Determination (R2)	**0.135**	**0.086**	Algorithm 1 is better as it has a higher R^2. Higher R^2 implies a better degree of "fit" between fitted forecast and actual sales data
Mean Absolute Deviation/Error (MAD/MAE)	**23.92**	**26.50**	Algorithm 1 is better as it has a lower MAD. Lower MAD demonstrates an overall lower error between fiitted forecast vs. actuals
Mean Absolute Percent Error (MAPE)	**21.04%**	**24.96%**	Algorithm 1 is better as it has a lower MAPE.
Symmetric MAPE (SMAPE)	**21.04%**	**22.75%**	Algorithm 1 is better as it has a lower SMAPE.

The final analysis — *Fit and error hold* ***great potential*** *to help planners more fully leverage the power of their demand forecasting software...*

at the time of forecast generation may have changed in the interim, warranting attention and possible corrective action. (For example, the business picked up a new customer, or a major customer was acquired, etc.)

Forecast error can also indicate possible time series problems, for example: history remapping to wrong customers, etc. Error is a key metric for exception-based forecasting; it enables planners to prioritize their efforts, work more efficiently, and rapidly drive business improvements by focusing on products with high error rates. And finally, forecast error is a powerful tool for evaluating the impact of forecaster bias.

Fit and error hold great potential to help planners more fully leverage the power of their demand forecasting software; and both provide baselines and benchmarks for improving this most fundamental aspect of supply chain management. Anyone in a position to impact demand planning and forecasting has a responsibility to understand the difference between these two concepts, and the value they hold.

Clearly.

To learn more, contact the supply chain specialists at Spinnaker. Call **877-476-05760** or visit **www.spinnakermgmt.com**.

To learn more about forecast fit and error…

Review the various calculation tables in the appendixes on the following pages.
Or call Spinnaker at ***877-476-0576.***

APPENDIX A:

Mean Squared Error (MSE) and Standard Error (SE) Calculations

Formulas:

$$\text{MSE} = \frac{\sum (Error)^2}{N-1} \qquad \text{SE} = \sqrt{\frac{\sum (Error)^2}{N-1}}$$

SE & MSE Calculations - Algorithm 1						
Period	Actual Sales (X)	Forecast (Y)	Error	Error^2		
1	82	94	12	144		
2	92	115	23	529		
3	120	144	24	576		
4	108	125	17	289		
5	110	88	-22	484		
6	152	105	-47	2209		
7	130	107	-23	529	Σ Error^2 =	7815
8	143	124	-19	361		
9	115	144	29	841	Σ Error^2 / (N-1) =	710.45
10	140	112	-28	784		
11	90	120	30	900		
12	78	65	-13	169	vΣ Error^2 / (N-1) =	26.65

MSE = 710.45 **SE = 26.65**

APPENDIX B:

Coefficient of Determination (R^2) Calculations

Formula:

$$R^2 = \left\{ \frac{\sum xy \quad \frac{(\sum x)(\sum y)}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n}\right]\left[\sum y^2 - \frac{(\sum y)^2}{n}\right]}} \right\}^2$$

Coefficient of Determination or R-Squared (R^2) calculations - Algorithm 1							
Period	Actual Sales (X)	Forecast (Y)	XY	X^2	Y^2		
1	82	94	7708	6724	8836		
2	92	115	10580	8464	13225		
3	120	144	17280	14400	20736		
4	108	125	13500	11664	15625		
5	110	88	9680	12100	7744	ΣXY - ΣX ΣY/N	2253.33
6	152	105	15960	23104	11025		
7	130	107	13910	16900	11449	ΣX^2 - (ΣX)^2/N	6680.67
8	143	124	17732	20449	15376		
9	115	144	16560	13225	20736	ΣY^2 - (ΣY)^2/N	5616.92
10	140	112	15680	19600	12544		
11	90	120	10800	8100	14400		
12	78	65	5070	6084	4225		
SUM	1360	1343	154460	160814	155921		
N=12	ΣX =1360	ΣY=1343	ΣXY =154460	ΣX^2=160814	ΣY^2=155921		

$$R^2 = \left\{ \frac{\Sigma XY - \Sigma X\, \Sigma Y/N}{\text{v}[\Sigma X^2 - (\Sigma X)^2/N]*[\Sigma Y^2 - (\Sigma Y)^2/N]} \right\}^2 = 0.1353$$

APPENDIX C:
Mean Absolute Deviation/Error (MAD/MAE) Calculations

Formula:

$$\textbf{MAD or MAE} = \frac{\sum |X - Y|}{N}$$

Mean Absolute Deviation (MAD) or Mean Absolute Error (MAE) Calculations - Algorithm 1				
Period	**Actual Sales (X)**	**Forecast (Y)**	**Absolute Deviation (AD) or Absolute error (AE) = |X-Y|**	
1	82	94	12	
2	92	115	23	
3	120	144	24	
4	108	125	17	
5	110	88	22	
6	152	105	47	
7	130	107	23	
8	143	124	19	
9	115	144	29	
10	140	112	28	
11	90	120	30	
12	78	65	13	**Σ|AE|/N 23.92**
			287 **Σ|AE|=287**	

MAD/MAE = 23.92

APPENDIX D:
Mean Absolute Percent Error (MAPE) Calculations

Formula:

Percent Error (PE) = $\frac{Actual - Forecast}{Actual} * 100$

MAPE = (ΣAbsolute Percent Error)/N = $\frac{\sum |PE|}{N}$

Mean Absolute Percent Error (MAPE) Calculations - Algorithm 1					
Period	Actual Sales (X)	Forecast (Y)	Percent Error (PE) = ((X-Y)/X)*100	Absolute PE (|PE|)	
1	82	94	-14.63%	14.63%	
2	92	115	-25.00%	25.00%	
3	120	144	-20.00%	20.00%	
4	108	125	-15.74%	15.74%	
5	110	88	20.00%	20.00%	
6	152	105	30.92%	30.92%	
7	130	107	17.69%	17.69%	
8	143	124	13.29%	13.29%	
9	115	144	-25.22%	25.22%	
10	140	112	20.00%	20.00%	
11	90	120	-33.33%	33.33%	
12	78	65	16.67%	16.67%	Σ|PE|/N 21.04%
				252.49% Σ|PE|=252.49%	

MAPE = 21.04%

APPENDIX E:
Symmetric Mean Absolute Percent Error (SMAPE) Calculations

Formula:

$$SMAPE = \frac{\sum (|X - Y| / [|X + Y| / 2])}{N} * 100$$

Symmetric MAPE Calculations - Algorithm 1							
Period	Actual Sales (X)	Forecast (Y)	Absolute Error AE=|X-Y|	Avg of Actuals & Forecast AVG=|X+Y|/2	(AE/AVG)*100		
1	82	94	12	88	13.64%		
2	92	115	23	104	22.22%		
3	120	144	24	132	18.18%		
4	108	125	17	117	14.59%		
5	110	88	22	99	22.22%		
6	152	105	47	129	36.58%		
7	130	107	23	119	19.41%		
8	143	124	19	134	14.23%	Σ(AE/AVG)*100	252.44%
9	115	144	29	130	22.39%		
10	140	112	28	126	22.22%	Σ(AE/AVG)*100)/12	21.04%
11	90	120	30	105	28.57%		
12	78	65	13	72	18.18%		

SMAPE = 21.04%