

# Theory and Application of James-Stein Seasonal Forecasting Model for Short Time Series

Hui Mao

School of Economics and Management  
Wuhan Polytechnic University  
Wuhan, China  
e-mail: maomao520323@gmail.com

Kui Zhang<sup>1</sup>

School of Economics and Management  
Wuhan Polytechnic University  
City, Country  
e-mail: kuizzz@hotmail.com

**Abstract**—Accurate seasonal forecasting plays an important role in product demand forecast. This paper gives an insight into the theory of a James-Stein seasonal forecasting model. The conditions in which the method outperforms the classical decomposition method are then presented. The conditions show that James-Stein model has more accurate prediction results when dealing with large noise data. The conclusion is then examined through a set of real data from M-competition. The experimental results confirm the practical value of the theory.

**Keywords**- forecasting; seasonality; short time series; James-Stein; supply chain

## I. INTRODUCTION

In today's complex and volatile market environment, the ability of quickly adapting to market changes has become a key to an enterprise's success. But supply chain is a dynamic system, which involves many uncertainty factors. This often causes "bullwhip effect", a phenomenon that the demand order variability in the supply chain were amplified as they moved up the supply chain <sup>[1]</sup>. As a result, it will make the enterprise fail to identify the real market needs. The bullwhip phenomenon has been recognized in many diverse markets, such as Procter & Gamble (P&G) and Hewlett-Packard (HP) <sup>[2]</sup>. The common symptoms of the phenomenon could be excessive inventory, poor product forecasts, insufficient or excessive capacities, poor customer service due to unavailable products or long backlogs, uncertain production planning and high costs for corrections. Four major causes of the bullwhip effect have been identified by Lee <sup>[1]</sup>, which include: ① Demand forecast updating; ② batching; ③ Price fluctuation; ④ Rationing and shortage gaming. Among the four causes, the inaccuracy of demand forecasting ranks No. 1. Once an enterprise in supply chain makes the errors in demand forecast, the errors will be gradually exaggerated. Thus, accurate demand forecasting is very crucial for an enterprise to successfully response to market changes, reduce inventory costs and guarantee corporate profits. Moreover, every member of a supply chain conducts some sort of forecasting in connection with its planning (e.g., the manufacturer does the production planning; the wholesaler does the logistics planning, and so on). Thus, an effective demand forecast is a necessary tool to help all the departments in a supply chain, such as production,

sales, finance, marketing, and logistics, formulate plans and make quick decisions.

The demand of modern products tends to exhibit the following three key characteristics:

1) *Obvious seasonal variation*. In the field of forecasting, seasonal variations are the periodic pattern in a time series that complete themselves within a calendar year and are then repeated on a yearly basis <sup>[3]</sup>, for example, beverages are best-sellers in summer rather than winter, heater sales soar in winter etc.

2) *Few historical data*. Due to the rapid development of new products in modern society, their demand data often have very short history (for example, two or three years.) or no history available.

3) *Large amount of noise*. Noise is the irregular fluctuation in a time series that follow no recognizable or regular pattern.

Accurate seasonal forecasting plays an important role in product demand forecast. A lot of studies based on diffusion model, especially BASS model, have been done on the demand forecasting of new products with few historical data by many domestic and international scholars <sup>[4-10]</sup>. The BASS model deals with seasonal data only through inserting a seasonal parameter into the model. This solution can cause a large prediction error <sup>[4]</sup>. In addition, the relevant parameters in the BASS model are often set by experts for their subjective judgment, which severely prevents the practical application of the model.

Both a short data series and a need for highly efficient methods (for large numbers of items) limit the use of complex procedures, such as Box-Jenkins seasonal ARIMA models <sup>[11]</sup>. Instead, a lot of simple methods are proposed to deal with the noisy short time series with strong seasonality. It is common to depersonalize the data before applying forecasting methods and then to reseasonalize the forecasts.

Until now, the seasonal forecasting methods for the individual short time series include two categories: Individual Seasonal Indices (ISI) method and Shrinkage Seasonal Indices (SSI) method <sup>[12]</sup>.

<sup>1</sup> Corresponding author. Tel: 13871569983; Email address:

[kuizzz@hotmail.com](mailto:kuizzz@hotmail.com)

Funded by 2013 Innovation fund at Wuhan Polytechnic University (2013cx031)

The ISI method (or Classic Decomposition method) forecasts seasonal demand based on the individual item's data history. This prediction accuracy of this method is extremely low. Thus, it is not suitable for time series with big noise.

In order to improve the prediction accuracy, the SSI was then proposed by Miller and Williams (2003) [13] for the estimation of individual seasonal demand. The SSI method mainly employs James-Stein estimator (1961) [14], an empirical Bayes estimator, to shrink or dampen the noise. James-Stein estimator dominates the maximum likelihood estimator (MLE) in total Mean Square Error (MSE) [15]. The SSI method based on James-Stein estimator is called James-Stein Shrinkage Seasonal Indices (JSSSI) method. The method dampens or shrinks individual seasonal indices to the average of seasonal factors, i.e. multiplicative ISIs are adjusted toward 1.0 or additive ISIs are adjusted toward 0. Miller and Williams also applied this method to data in the M-competition data to verify the application value of the JSSSI method. The experimental results proved that the method is suitable for the data which have short history, large amount of samples and big noise.

However, Miller and Williams' conclusion has not been supported by any theoretical evidences. This paper will reveal the intrinsic characteristic of the JSSSI from the perspective of theory, and identify under what conditions JSSSI method is better than ISI method in terms of prediction accuracy. The theoretical results are then verified through an empirical experiment.

## II. THEORETICAL STUDY

### A. Model assumptions

This paper assumes that there is no trend in this model in order to concentrate on the seasonal component alone. Also, it is assumed that demand data  $Y_{th}$  only consists of the mean  $\mu$ , seasonal factors  $S_h$  and random noise factor  $\varepsilon_{th}$  by additive model [16-17] as follows:

$$Y_{th} = \mu + S_h + \varepsilon_{th} . \quad (1)$$

Where, suffix  $t$  represents the year and  $t=1,2,\dots,r$  (where  $r$  is the number of years' data history);

Suffix  $h$  represents the seasonal period and  $h=1,2,\dots,q$ ; (where  $q$  is the number of the quarter in a year);

$\varepsilon_{th}$  represents a random disturbance term which is assumed to be identically and independently normal distributed with mean zero and constant variance  $\sigma^2$ ;

The sum of seasonal factors in the additive model is zero, i.e.  $\sum_{h=1}^q S_h = 0$ .

### B. Mean squared error of ISI

The mean can be estimated by

$$\hat{\mu} = \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} . \quad (2)$$

Individual seasonal indices (ISI) of season  $h$  can be estimated by

$$\hat{S}_h^{ISI} = \frac{1}{r} \sum_{t=1}^r Y_{th} - \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} . \quad (3)$$

Given the number of years' data history  $r$ , the demand for season  $h$  in the next year (year  $r+1$ ) can be estimated through ISI method by

$$\begin{aligned} \hat{Y}_{(r+1)h}^{ISI} &= \hat{\mu} + \hat{S}_h^{ISI} \\ &= \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} + \frac{1}{r} \sum_{t=1}^r Y_{th} - \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} \\ &= \frac{1}{r} \sum_{t=1}^r Y_{th} . \end{aligned} \quad (4)$$

The mean squared error is:

$$\begin{aligned} \text{MSE}^{ISI} &= \text{E} \left( \hat{Y}_{(r+1)h}^{ISI} - Y_{(r+1)h} \right)^2 = \text{E} \left[ \frac{1}{r} \sum_{t=1}^r Y_{th} - Y_{(r+1)h} \right]^2 \\ &= \text{E} \left[ \mu + S_h + \frac{1}{r} \sum_{t=1}^r \varepsilon_{th} - \mu - S_h - \varepsilon_{(r+1)h} \right]^2 \\ &= \left( 1 + \frac{1}{r} \right) \sigma^2 . \end{aligned} \quad (5)$$

### C. Mean squared error of JSSSI

In additive model, JSSSI can be calculated as

$$\hat{S}_h^{JSSSI} = (1-W) \hat{S}_h^{ISI} . \quad (6)$$

$\hat{S}_h^{ISI}$  Can be estimated by equation (3).  $w = (q-2)D/Z$  In equation (6) is the shrinkage parameter of James-Stein, where  $D = (q-1)\sigma^2/qr$  is the average of the variances of different seasonal factors; and  $Z = \sum_{h=1}^q \left( \hat{S}_h^{ISI} \right)^2$ .

Similar to the ISI, the demand of the next year (that is year  $r+1$ ) can be estimated by using JSSSI method:

$$\begin{aligned} \hat{Y}_{(r+1)h}^{JSSSI} &= \hat{\mu} + \hat{S}_h^{JSSSI} \\ &= \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} + (1-W) \left( \frac{1}{r} \sum_{t=1}^r Y_{th} - \frac{1}{qr} \sum_{t=1}^r \sum_{h=1}^q Y_{th} \right) . \end{aligned} \quad (7)$$

The MSE of JSSSI also can be calculated in the way of equation (5):

$$\begin{aligned} \text{MSE}^{JSSSI} &= \text{E} \left( \hat{Y}_{(r+1)h}^{JSSSI} - Y_{(r+1)h} \right)^2 \\ &= \left( 1 + \frac{1}{qr} \right) \sigma^2 + (q-1)(1-W)^2 \sigma^2 / qr + W^2 S_h^2 \end{aligned} \quad (8)$$

### D. Compare ISI with JSSSI

Comparing MSE of the ISI with MSE of the JSSSI, the following equation can be obtained:

$$\begin{aligned} \text{MSE}^{ISI} - \text{MSE}^{JSSSI} &= \left( 1 + \frac{1}{r} \right) \sigma^2 - \left( 1 + \frac{1}{qr} \right) \sigma^2 - (q-1)(1-W)^2 \sigma^2 / qr - W^2 S_h^2 . \end{aligned} \quad (9)$$

From equation (9), a conclusion can be drawn that the JSSSI method is superior to the ISI method, i.e.  $MSE^{JSSSI} < MSE^{ISI}$  when  $W$  satisfy the condition of equation (10). It is found that the condition is easily achieved when noise is large ( $\sigma^2$  is small). That means that JSSSI method can perform better for the data with large noise.

$$0 < W < 1 - \frac{S_h^2 - (q-1)\sigma^2 / qr}{S_h^2 + (q-1)\sigma^2 / qr} . \quad (10)$$

### III. CASE STUDY

#### A. Data processing

The experiment is conducted based on 55 monthly time series that are included in the subset of 111 series used in the M-competition. The first 48 months data were used for forecast and the last 6 months data were withheld to assess the performance of the models.

The data is decomposed by using multiplicative mode in the experiment:  $Y_{th} = T_t \times S_h \times \varepsilon_{th}$ , where  $T_t$  is the trend component. Firstly, trend component and ISI can be obtained by using Class Decomposition method. Then, James-Stein shrinkage seasonal indices can be calculated by using the JSSSI method. Finally, the value of ISI or JSSSI is combined with the trend forecast to obtain the demand prediction results. The performance of two methods can be examined by comparing the forecasts with the remaining real data.

#### B. Prediction results

From Table 1, it is found that 35 series have large noise ( $\sigma^2 \geq 10^{-3}$ ) and 29 series nearly have no noise ( $\sigma^2 < 10^{-3}$ ) in the 55 series.

Since the results from MSE, MAPE and MAD are similar, they will discuss only based on the average of the three forecast measurements and for all of the three horizons.

Table 1 shows the ratio of JSSSI to ISI.  $JSSSI > ISI$  means that the JSSSI method is superior to the ISI method, Similarly,  $JSSSI \approx ISI$  means that JSSSI and ISI are similar, and  $JSSSI < ISI$  means that the JSSSI is inferior to ISI.

Table 1 indicates that: the JSSSI method is superior to or not inferior to the ISI method for nearly 77% series when the shrinkage parameter is large ( $W \geq 0.1$ ); and 42.86% series demonstrate obvious advantage. However, the number jumps from 77% to 98.29% as the shrinkage parameter turns to small ( $W < 0.1$ ). The results confirm the theoretical derivation that shrinkage seasonal model has great advantage in processing noisy data. As a matter of fact, whether large or small the shrinkage effect is, the shrinkage model can improve the prediction accuracy for most of series to some extent.

TABLE I. THE COMPARISON OF JSSSI METHOD AND ISI METHOD IN TERMS OF PREDICTION PERFORMANCE

category	JSSSI>ISI (%)			JSSSI=ISI (%)			JSSSI<ISI (%)		
	MSE	MAPE	MAD	MSE	MAPE	MAD	MSE	MAPE	MAD
14 series (Dependent) for $\sigma^2 \geq 10^{-3}$ and $W \geq 0.1$									
1 month	50.00	42.86	42.86	7.14	28.57	28.57	42.86	28.57	28.57
3 months	57.14	42.86	42.86	7.14	35.71	35.71	35.71	21.43	21.43
6 months	42.86	28.57	35.71	42.86	64.29	57.14	14.29	7.14	7.14
average	42.86				44.44	34.33			23.01
21 series (Dependent) for $\sigma^2 \geq 10^{-3}$ and $W < 0.1$									
1 month	10.77	38.46	38.77	89.23	61.54	69.23	8.00	0.00	8.00
3 months	11.38	11.38	11.38	84.62	84.62	84.62	8.00	0.00	8.00
6 months	58.48	35.95	100.00	33.83	38.48	33.03	7.69	7.69	8.89
average			57.61			68.68			1.71
10 series (Dependent) for $\sigma^2 < 10^{-3}$ and $W \geq 0.1$									
1 month	10.00	0.00	0.00	30.00	60.00	60.00	40.00	40.00	40.00
3 months	0.00	0.00	0.00	80.00	70.00	70.00	20.00	30.00	30.00
6 months	0.00	0.00	0.00	80.00	70.00	70.00	20.00	30.00	30.00
average			1.11			67.78			31.11
10 series (Dependent) for $\sigma^2 < 10^{-3}$ and $W < 0.1$									
1 month	20.00	20.00	20.00	80.00	80.00	80.00	20.00	20.00	20.00
3 months	10.00	0.00	0.00	80.00	100.00	100.00	10.00	0.00	8.89
6 months	10.00	0.00	0.00	80.00	100.00	100.00	10.00	0.00	8.89
average			8.89			82.22			8.89

Conversely, in the condition of little noise, when shrinkage parameter is large ( $W \geq 0.1$ ), using JSSSI method is obvious superior to ISI method for only 1.11% series and using JSSSI method is obvious inferior to ISI method for 31.11% series; while the shrinkage parameter is small ( $W < 0.1$ ), the percentage of series superior for JSSSI method is increase slightly to 8.89%. All in all, for the condition of small noise, JSSSI method has shown no obvious superiority for most of series.

Also, the experiment confirmed theoretical results from Miller and Williams [4] that JSSSI method will perform better for the data with large noise. For example, the 11th series, for which  $\sigma^2 = 0.018$  and with great shrinkage ( $W = 0.38$ ), its prediction accuracy can be improved by around 40% using JSSSI method compared to ISI method. For some data with large noise, for example, series No. 51, slight shrinkage ( $W = 0.02$ ) also can improve the accuracy by 23%. For most data with little noise, shrinkage method has no advantage, only for a very few data such as series No. 46 ( $\sigma^2 = 0.00019$ ), a little shrinkage may be achieved (by 17%).

### IV. CONCLUSIONS

This paper gives an insight into the characteristics of a James-Stein seasonal forecasting model. The conditions in which the JSSSI method outperforms the classical decomposition method are obtained. It is concluded that JSSSI model has better prediction performance when dealing with large noise data.

The conclusion is then examined through a set of real data from M-competition. The results are approximately the same for both multiplicative and additive models, and they were not sensitive to the choice of MSE or MAPE or MAD as an error statistic.

Furthermore, the conclusion is consistent with the previous research from Miller and Williams: James-Stein model is special suitable for forecasting time series with great random variation [13].

From the perspective of modeling, the JSSSI model is also considered as one of simple forecasting methods [13]. It can be easily implemented within a spreadsheet

and perform well. Therefore, this paper not only has a certain guiding significance for selecting reasonable prediction method, but also confirms a practical value of the theory.

#### REFERENCES

- [1] Lee, H.L., V. Padmanabhan, and S. Whang. The Bullwhip Effect In Supply Chains. *Sloan Management Review*, Spring 1997, 38(3): pp. 93-102.
- [2] Lee, H.L., V. Padmanabhan, and S. Whang. Information distortion in a supply chain: the bullwhip effect. *Management Science*, 1997. 43(4): p. 546-558.
- [3] Bowerman, B.L., R.T. O'Connell, and A.B. Koehler, *Forecasting, Time Series, and Regression: An Applied Approach*. 4 ed. 2005: Brooks/Cole.
- [4] Xu xianhao, Guo Xiaoyun. Forecasting the demand of products with short lifecycle based on Norton model. *Industrial Engineering and Management*, 2008. 2: p. 19-24.
- [5] Xu jiuping, Liao Zhigao. Technology innovation diffusion velocity model. *Chinese Journal of Management*, 2004. 1(3): p. 330-340.
- [6] Yang Jinghui, Wu chunyou. A study on upgrade products' diffusion trend by Norton model-- take the example of the upgrade online behaviour of China's Internet users. *Science Research*, 2005. 23(5): p. 682-687.
- [7] Meng Fandong, He Mingsheng. Research on IPv4 standard 's diffusion trend of China by using improved Bass model. *Journal of Dalian University of Technology*, 2008. 1: p. 137-142.
- [8] Ji Changhong, Ding Xihai. Forecasting the demand of products with short life cycle based on an improved Bass model. *Science technology and Engineering*, 2010. 10(10): p. 2577-2580.
- [9] Wu Jifeng. Research on short life cycle products demand forecasting and inventory control based on the bayesian updating, thesis in Nanjing University of Science & Technology, 2012.
- [10] Bass, F.M., A new product growth model for consumer durables. *Management Science*, 1969. 15(5): p. 215-227.
- [11] Ittig, P.T., Comparison of Efficient Seasonal Indexes. *Journal of Applied Mathematics and Decision Sciences*, 2004. 8(2): p. 87-105.
- [12] Zhang, K., et al., Generalised Estimators for Seasonal Forecasting by Combining Grouping with Shrinkage Approaches. *Journal of Forecasting*, 2013. 32(2): p. 137-150.
- [13] Miller, D.M. , D. Williams. Shrinkage estimators of time series seasonal factors and their effect on forecasting accuracy. *International Journal of Forecasting*, 2003. 19: p. 669-684.
- [14] James, W., C. Stein. Estimation with Quadratic Loss. in 4th Berkeley Symposium on Mathematical Statistics and Probability. 1961. Berkeley: University of California Press.
- [15] Efron, B., C. Morris, Stein's Estimation Rule and Its Competitors -- An Empirical Bayes Approach. *Journal of the American Statistical Association*, 1973. 68(341): p. 117-130.
- [16] Chen H, Boylan JE. 2007. Use of individual and group seasonal indices in subaggregate demand forecasting. *Journal of the Operational Research Society* 58: 1660-1671.
- [17] Chen H, Boylan JE. 2008. Empirical evidence on individual, group and shrinkage seasonal indices. *International Journal of Forecasting* 24: 525-534.