6. ADDING EMD PROCESS AND FILTERING ANALYSIS TO ENHANCE PERFORMANCES OF ARIMA MODEL WHEN TIME SERIES IS MEASUREMENT DATA

Feng-Jenq LIN[*]

Abstract

In this paper, one process that integratesthe Empirical Mode Decomposition with filtering analysis was proposed to reconstruct the de-noise data series when the original is measurement data. The ARIMA model was augmented with the above process (here from referred to as EF-ARIMA) to treat de-noise measurement data. Model fit and forecasting performance of EF-ARIMA, using de-noise data set, were compared to those of the traditional ARIMA, which used the original data set, in an empirical study. By examining the MAE, MAPE, RMSE and Theil's inequality coefficients, it was concluded that EF-ARIMA outperformed its traditional counterpart. It also shows that the proposed hybrid forecasting approach is feasible and reliable. The results suggest application implications for forecasting measurement data sets in other areas as well.

Keywords: Hilbert-Huang transform, empirical mode decomposition, filtering analysis, measurement data, ARIMA model

JEL Classification: C22

I. Introduction

The Autoregressive Integrated Moving Average (ARIMA) model (Box and Jenkins, 1976), a linear combination of time-lagged variables and error terms is one of the most widely used forecasting techniques. When the data series is of seasonality and trends in nature, the seasonal ARIMA is used to model and to forecast in many fields. Both

[*] Department of Applied Economics and Management, National I-Lan University, Yilan,Taiwan.E-mail:fjlin@niu.edu.tw.

modelsare limited because they assume linear relationships among time-lagged variables, making them inapplicable to non-linear relationships (Zhang *et al*., 1998).

For non-linear and non-stationary signal analysis, the Empirical Mode Decomposition (EMD), a stage of Hilbert-Huang Transform (HHT) proposed by Huang *et al*. (1998), is better suited. It adaptively represents the local characteristics of the given signal data (Huang *et al*., 1998). The concept of confidence limit for EMD, which can be applied to the analysis of nonlinear and non-stationary processes, was also introduced by using various adjustable stopping criteria in the sifting processes of the EMD step to generate a sample set of Intrinsic Mode Functions (IMFs) (Huang *et al*., 2003). By using EMD, any complicated signal can be decomposed into a number of Intrinsic Mode Functions (IMFs), which have simpler frequency components and high correlations, making forecasts easier and more accurate (Chen *et al*., 2012). However, in most empirical cases, when the meaningful IMFs are used as input for one forecasting model, it is difficult to decide how many meaningful IMFs should be kept (or summarized) and used for reconstructing the forecasting model, especially when the original is measurement data. There is still no consensus on the selection criterion in the studies just discussed.

To reduce the noise (error message or unimportant message) in time series data and to improve forecasting performance, some researchers have developed hybrid forecasting approaches by combining EMD with computer learning technologies such as the neural networks, the support vector machine, etc. (Yang *et al*., 2007; Zhu *et al*., 2007; Shen *et al*., 2008; Hamad *et al*., 2009). With the similarpurposes in mind, we want to augment the traditional ARIMA model for time series measurement data with IMFs derived from an EMD process. The significance test uses partial correlation and incremental R-square to filter IMFs. Meaningful IMFs are then summarized and used to reconstruct the de-noise data series for ARIMA model. It is our main motivation in this paper.

At the same time, the proposed hybrid forecasting approach combining the EMD process and filtering analysis with ARIMA model is used to forecast an empirical case, the monthly industrial productive index of Taiwan, to illustrate its forecasting performance. Final results show that our hybrid EF-ARIMA approach performs better than the traditional ARIMA do. It also presents the feasibility and reliability of our proposed hybrid forecasting approach.

2. Methodology

2.1. Literature Review for EMD

The essence of Empirical Mode Decomposition (EMD) is the sifting process which extracts a finite number of Intrinsic Mode Functions (IMFs) based on the local characteristic time scale and one residue (R; also can be regarded as an IMF), which generally represents the trend of data series in the original data. The extracted IMFs contain a range of frequencies, from high to low, and represent periodic patterns in the original data series. Individually, an IMF has simpler frequency components with high correlations (Chen *et al*., 2012). Hence, EMD can deal with non-linear and non-stationary data (Huang *et al*., 1998). More EMD related details can be found in the introduction of Huang *et al*. (1998).

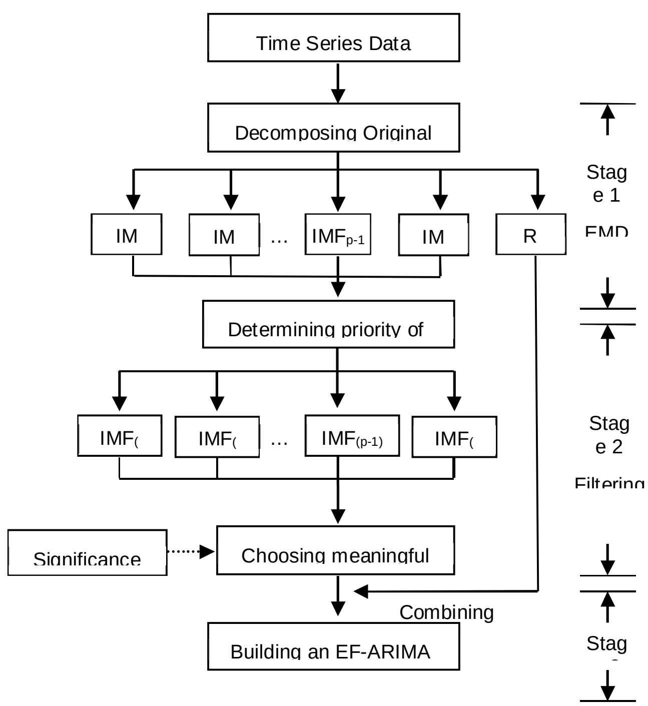
The EMD process has been successfully used to decompose the data series in many different fields, such as water wave (Huang *et al*., 1999), wind speed (Liu *et al*., 2009), structural health monitoring (Vincent *et al*., 1999), ocean waves (Veltcheva and Soares, 2004), fault diagnosis of roller bearings (Yu *et al*., 2005; Cheng *et al*., 2006), fault diagnosis ofrotating machinery(Wu and Qu, 2008), cardiorespiratory synchronization (Wu and Hu, 2006), financial fluctuation (Yang and Lin, 2012; Wang et al., 2009), oil price (Yu *et al*., 2008), tourism demand (Chen *et al*., 2012), metro passenger flow (Wei and Chen, 2012), etc. In addition to decomposing a data series into a finite number of IMFs and analyzing them, the last four studies also use meaningful IMFs as input data for forecasting models. All of these references show that using IMFs as data yields better performance than using the original data alone.

2.2. Constructing the EF-ARIMA Model

A hybrid EF-ARIMA model is proposed for better model fit and forecasting for time series measurement data.

Figure 1

The framework of a hybrid forecasting approach



It is made of a traditional ARIMA augmented with EMD and a filter. The EF-ARIMA is constructed by going through three stages (Figure1): the EMD stage (decomposing the data series into IMFs), the filtering stage (identifying the meaningful IMFs by significance test) and the hybrid model stage (reconstructing a de-noise data series to build EF-ARIMA). The three stages are described in detail next.

Stage 1: The EMD Stage

Firstly, the original data series is decomposed into a finite number of Intrinsic Mode Functions (IMFs) and one residue (R) by using the sifting process in an EMD.After the data series was decomposed,since meaningful IMFs need to be retained for the reconstruction of the de-noise data series the input data to ARIMA, non-contributing IMFs must be reduce or removed before forecasting models are built.

Stage 2: The Filtering Stage

The IMFs derived during the EMD stage aredecomposed from high to low frequencies according to their contributing priorities for the original data. However, it is not clear how many non-suitable IMFs should be treated as noise and how many suitable IMFs should be used along with R to reconstruct the de-noise data series. Most researchers simply resort to some subjective criterion. Note that the original data series is equal to the sum of all IMFs and R:

$$\text{Original data series } Y = IMF_1 + IMF_2 + \ldots + IMF_p + R$$

To identify meaningful IMFs, objectively, the following two steps will be taken:

Step 1: Determining the priority of all IMFs sequentially

Simple correlation and partial correlation are used to determine the priority set of all IMFs, $\{IMF_{(1)}, IMF_{(2)}, \ldots, IMF_{(P)}\}$, by following p number iterations in this step:

1. Calculate the simple correlations between(Y- R) and IMF_J, that is

   $$\text{Corr}(\text{Y-R}, \mathbf{IMF}_J) = \mathbf{r}_{(Y-R),IMF_J}, \; J \in \{1, 2, \ldots, p\}$$

   where the p simple correlations are calculated in this iteration. Now, if we suppose IMF_5 is the highest simple correlation with (Y-R), then let the symbol $\mathbf{IMF}_{(1)} = \mathbf{IMF_5}$.
2. Given the controlling variable $IMF_{(1)}$, the partial correlations between (Y-R) and IMF_J can be calculated and they are exactly the same as their simple correlation counterparts, between (Y-R-$IMF_{(1)}$) and IMF_J, i.e.,

   $$\mathbf{r}_{(Y-R), IMF_J \mid IMF_{(1)}} = \mathbf{r}_{(Y-R-IMF_{(1)}), IMF_J}, \; J \in \{1, 2, \ldots, p\} \text{ and } J \neq 5$$

   where the (p-1) partial correlations are calculated in this iteration. Suppose IMF_3 is the highest correlation with (Y-R-$IMF_{(1)}$) now, then let the symbol $\mathbf{IMF}_{(2)} = \mathbf{IMF_3}$.
3. Given the set of 2 controlling variables $\{\mathbf{IMF}_{(1)}, \mathbf{IMF}_{(2)}\}$, the partial correlations between (Y-R) and IMF_J can be calculated and they are exactly the same as their simple correlation counterparts (Y-R-$\mathbf{IMF}_{(1)}$ - $\mathbf{IMF}_{(2)}$) and IMF_J, i.e.:

$$\mathbf{r}_{(Y-R), IMF_J \mid IMF_{(1)}, IMF_{(2)}} = \mathbf{r}_{(Y-R-IMF_{(1)}-IMF_{(2)}), IMF_J \mid IMF_{(1)}, IMF_{(2)}}, \; J \in \{1, 2, \ldots, p\} \text{ and } J \neq 3,5$$

where the (p-2) partial correlations are calculated in this iteration. Suppose IMF_1 is the highest partial correlation with (Y-R-$IMF_{(1)}$ -$IMF_{(2)}$) now, then let the symbol $\mathbf{IMF}_{(3)} = \mathbf{IMF}_1$. Repeat for the remaining (p-3) iterations.

Step 2: Retaining the meaningful IMFs by using the incremental R-square
Note that in step 1, the contributing priorities of all IMFs for the original data are objectively determined. Now, start by substituting $\sum_{j=1}^{J}\mathbf{IMF}_{(j)}$ in Equation (1) with $IMF_{(1)}$. R-square $\mathbf{R}^2_{(1)}$ is obtained. Repeat the process and obtain $\mathbf{R}^2_{(2)}$ by adding $IMF_{(1)}$ and $IMF_{(2)}$.

$$(Y\text{-}R) = \hat{\boldsymbol{\beta}}_0 + \hat{\boldsymbol{\beta}}_1 \sum_{j=1}^{J}\mathbf{IMF}_{(j)} \text{ , } J \in \{1, 2, \ldots, p\} \quad (1)$$

Suppose $\mathbf{R}^2_{(0)} = 0$, then we now can define

$$\Delta\mathbf{R}^2_{(J)} = \mathbf{R}^2_{(J)} - \mathbf{R}^2_{(J-1)} \text{ ,} J \in \{1, 2, \ldots, p\} \quad (2)$$

where: $\mathbf{R}^2_{(J)}$ represents the coefficient of determination when one particular $\mathbf{IMF}_{(J)}$ is added into the regression model. The incremental R-square $\Delta\mathbf{R}^2_{(2)}$ is now obtained from Equation (2). As more iterations are performed, more $\Delta\mathbf{R}^2_{(J)}$ are obtained.

For most situations, the $\Delta\mathbf{R}^2_{(J)}$ will gradually die down in a damped exponential fashion with no oscillation when the index J is increased. Hence, in order to test whether a particular $\sqrt{\Delta\mathbf{R}^2_{(J)}}$ is zero or not, the standard error ($\mathbf{1}/\sqrt{\mathbf{n}}$)will be employed for evaluating the significance of the coefficients of partial autocorrelation function (PACF) based on the asymptotic results of Quenouille (1948). That is, we check whether $\sqrt{\Delta\mathbf{R}^2_{(J)}}$ is statistically significant at, say, the 5 percent level by determining whether it exceeds $\mathbf{2}/\sqrt{\mathbf{n}}$ in magnitude. A particular $\mathbf{IMF}_{(J)}$ can be used to reconstruct the de-noise data series for original data when its absolute value is larger than $\mathbf{2}/\sqrt{\mathbf{n}}$ in magnitude. Otherwise, it is treated as noise.

Stage 3: Building the EF-ARIMA

The hybrid EF-ARIMA is now complete with an ARIMA model equipped with an EMD with a filtering process that eliminates noise.

To put our EF-ARIMA into test, we use the industrial productive index of Taiwan as our original data series. We compare its model fit and forecasting performance to those of the traditional ARIMA.

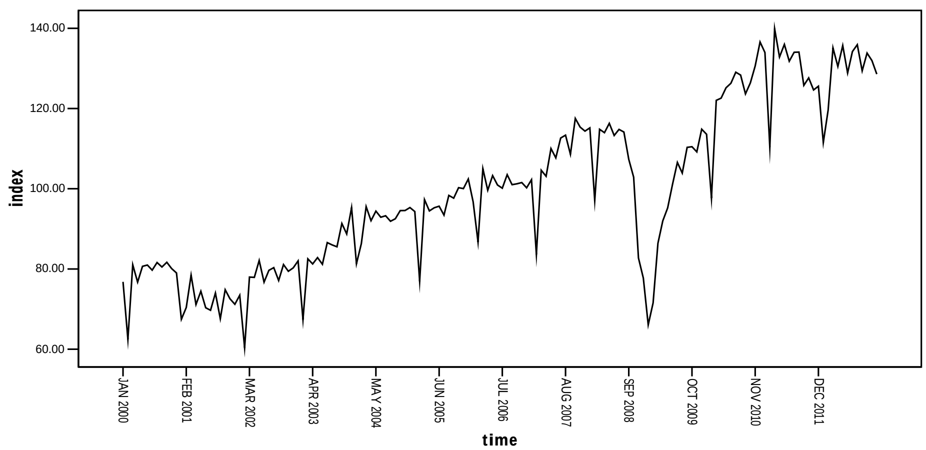
3. The Empirical Case

3.1. Data Sets

To illustrate the usefulness of the proposed EF-ARIMA and evaluate its noise reduction effect on measurement data, the official monthly industrial production index (from January 2000 to December 2012),obtained from the Taiwanese Ministry of Economic Affairs, is used (Figure 2). There are 156 monthly data points in total. The data series exhibits a long-term upward trend with short-term fluctuations that are independent from one time period to the next. The index appears to be non-stationary in that the mean increases over time.

Figure 2

The Monthly Industrial Production Index in Taiwan (Jan. 2000 - Dec. 2012)



A training period and a testing period are used to evaluate the performance of the proposed EF-ARIMA model. Since a longer training period gives more reliable and better results, it is decided that the first ten years (Jan. 2000 to Dec. 2010) serves as the training period and the next two years (Jan. 2011 to Dec. 2012) as the testing period. The iterative forecasting is performed on our empirical case.

Table 1

Training and Testing Data Sets

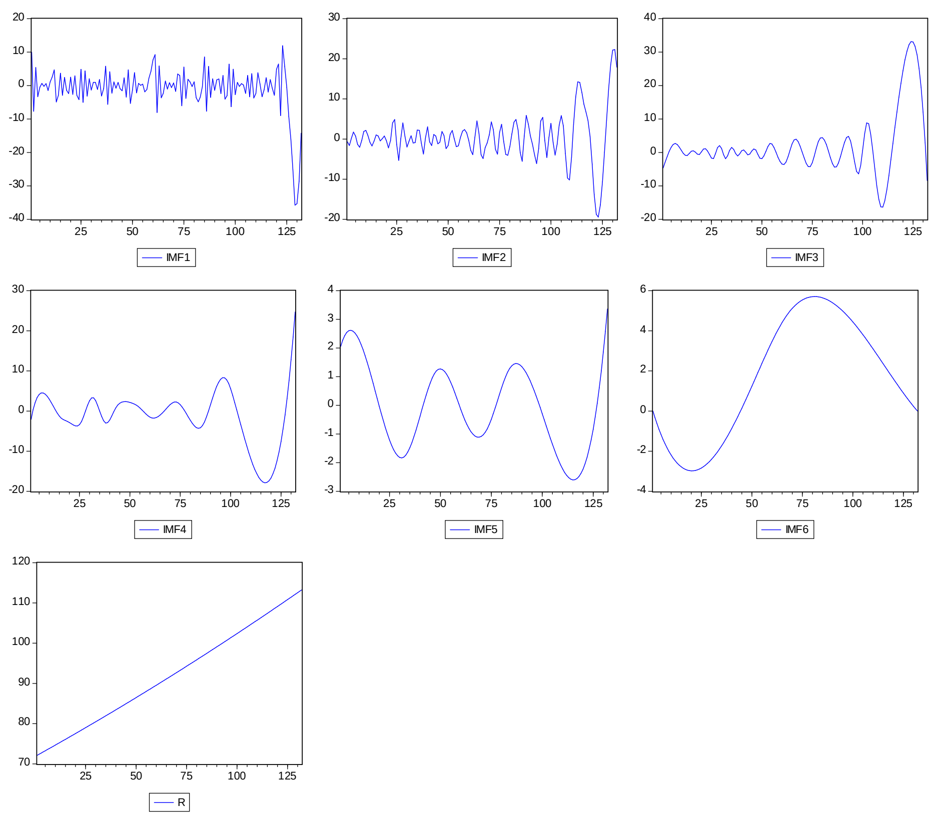
Data	Training set	Testing set
2000/01~2012/12 (156 data points)	2000/01~2010/12 (132 data points)	2011/01~2012/12 (24 data points)

3.2 The EMD process

Following the first stage in Figure 1, the monthly industrial production index of Taiwan is decomposed into six IMFs (named IMF_1, IMF_2, …, IMF_6) and one residue (called R), as shown in Figure 3. The IMFs obtained are graphically illustrated in the order they are extracted, indicative of the order of frequency (or period) from the highest frequency to the lowest one. IMFs with higher frequencies (or shorter periods) are extracted first and the ones with lower frequencies (or longer periods) are extracted later. The first few IMFs represent the high time variants or noise in the original data, while the last few IMFs represent the longer period components. The last component is the residue, which represents the long trend of monthly industrial production index. The IMFs are obtained using the Matlab software (Matlab 2008b).

Figure 3

The IMFs and R for Industrial Production Index in Taiwan



3.3. The Filtering Process

During Stage 2 in Figure 1, we prioritize IFMs and filter out the noise. In order toidentify and retain the meaningful IMFs for reconstructing de-noise data series, the following two filtering steps are executed:

Step 1: Determining the priority of all IMFs considered sequentially

Because we do not know which IMFs contribute more to (Y-R), the partial correlations are calculated to determine the priority of all IMFs by six iterations in this step.

The correlation coefficients between (Y-R) and all IMFs are calculated individually in the first iteration. Of all correlations between (Y-R) and IMFs, IMF_4 is found to have the highest correlations (0.4974 and 0.5121). It is therefore identified as $IMF_{(1)}$, the most important component.

Then, the correlation coefficients between (Y-R- $IMF_{(1)}$) and all remaining IMFs are calculated individually in the second iteration. Similarly, given the controlling variable IMF_4, IMF_3 is found to be the most contributing IMF (with correlations of 0.6803 and 0.4709) and thus identified as $IMF_{(2)}$ .the most important component.

After six iterations, all the IMF_s are identified in the order of their contribution. Table 2 shows our sequence (IMF_4, IMF_3, IMF_1, IMF_2, IMF_6, and IMF_5).

Table 2

The Correlation Coefficients between (with_var) and IMFs

Iter.	Correlation Coefficients	(with_var)	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6
1	Pearson	Y-R	-0.0533	0.2423	0.44794	***0.4974***	0.4234	0.2248
	Spearman		0.2096	0.1537	0.41448	***0.5121***	0.3985	0.3691
2	Pearson	Y-R-$IMF_{(1)}$	0.1879	0.0680	***0.6803***	--	-0.0251	0.2818
	Spearman		0.3568	0.1882	***0.4709***	--	-0.0084	0.4955
3	Pearson	Y-R- $\sum_{j=1}^{2} IMF_{(j)}$	***0.5000***	0.35876	--	--	0.11410	0.4473
	Spearman		***0.4842***	0.44784	--	--	0.12286	0.5029
4	Pearson	Y-R- $\sum_{j=1}^{3} IMF_{(j)}$	--	***0.88880***	--	--	0.31199	0.37468
	Spearman		--	***0.73482***	--	--	0.26107	0.56382
5	Pearson	Y-R- $\sum_{j=1}^{4} IMF_{(j)}$	--	--	--	--	0.37680	***0.8919***
	Spearman		--	--	--	--	0.37958	***0.9046***
6	Pearson	Y-R- $\sum_{j=1}^{5} IMF_{(j)}$	--	--	--	--	***1.0000***	--
	Spearman		--	--	--	--	***1.0000***	--

Note: (with_var) represents Y- R- $\sum_{j=1}^{K} IMF_{(j)}$, where K equals the number of iterations.

Step 2: Retaining the meaningful IMFs by examining incremental R-squares

Next, in order to identify the meaningful IMFs to reconstruct the de-noise data series for the original series, the incremental R-square (ΔR^2) is calculated when one particular

$IMF_{(J)}$ is added to the following regression model sequentially in each iteration, as shown in Table 3.

$$(\text{Y-R}) = \hat{\boldsymbol{\beta}}_0 + \hat{\boldsymbol{\beta}}_1 \sum_{i=1}^{J} \mathbf{IMF}_{(i)}, \qquad J \in \{1, 2, \ldots, 6\}$$

Results show that IMF_5 is the only IMF not to be used because its absolute value does not exceed 0.1705 ($2/\sqrt{132}$) under the 5 percent significance level. The other IMFs (IMF_4 , IMF_3 , IMF_1 , IMF_2 , IMF_6) are used to reconstruct the de-noise data series.

Table 3

R^2 and ΔR^2 when an IMF is Added to Simple Regression Formula

Iteration	Response Variable	Added IMF	Predictor Variable	R^2	ΔR^2	$\sqrt{\Delta R^2}$
1	Y- R	IMF_4	$IMF_{(1)}$	0.2474	0.2474	0.4974
2		IMF_3	$\sum_{i=1}^{2} IMF_{(i)}$	0.5797	0.3323	0.5765
3		IMF_1	$\sum_{i=1}^{3} IMF_{(i)}$	0.6737	0.0940	0.3066
4		IMF_2	$\sum_{i=1}^{4} IMF_{(i)}$	0.9011	0.2274	0.4769
5		IMF_6	$\sum_{i=1}^{5} IMF_{(i)}$	0.9874	0.0863	0.2938
6		IMF_5	$\sum_{i=1}^{6} IMF_{(i)}$	1.0000	0.0126	***0.1123***

3.4. Building the ARIMA and EF-ARIMA Models

After applying the EMD process and filtering analysis, the de-noise data series of the training set (Taiwan monthly industrial production index of Taiwan (from January 2000 to December 2010)) is obtained and then used to build the optimal EF-ARIMA model. The final parameter estimates, T values and related statistics of the model are as following:

$$\Delta^1\hat{Y}_t = \mathbf{0.9960}\Delta^1 Y_{t-12} - \mathbf{0.1470}\,\varepsilon_{t-1} + \mathbf{0.1251}\,\varepsilon_{t-11} - \mathbf{0.9108}\,\varepsilon_{t-12}$$

(36.4686*) (-3.7809*) (3.1215*) (-30.6758*)

R^2 =0.6029 $\bar{R}^2$ =0.5926 D.W.=2.1683

Q(6)=6.0472 Q(12)=13.926 Q(18)=15.746 Q(24)=21.968

where: * represents the estimate being significant at 0.05 level.

$\Delta^1 Y_t$ represents variable being first differencing transformation ($\Delta^1 Y_t = Y_t - Y_{t-1}$).

For benchmarking, the traditional ARIMA model is applied to the original industrial productive index of the same period. The final parameter estimates, T values and related statistics of optimal traditional ARIMA model are as following:

$$\mathbf{\Delta^1}\hat{Y}_t = \mathbf{0.9968\Delta^1}Y_{t-\mathbf{12}} - \mathbf{0.1403}\ \varepsilon_{t-\mathbf{1}} + \mathbf{0.1293}\ \varepsilon_{t-\mathbf{11}} - \mathbf{0.9082}\ \varepsilon_{t-\mathbf{12}}$$

(36.2267*) (-3.5958*) (3.2539*) (-30.8586*)

R^2=0.5972 $\bar{R}^2$=0.5867D.W.=2.1416

Q(6)=5.9086 Q(12)=12.938 Q(18)= 14.743 Q(24)=20.585

where: * represents the estimate being significant at 0.05 level

$\Delta^1 Y_t$ represents variable being first differencing transformation ($\Delta^1 Y_t = Y_t - Y_{t-1}$).

Both optimal models are suitable in the statistical tests of parameter significance and residual diagnosis (Q statistic) for the test of goodness-of-fit.

From the optimal traditional ARIMA model, two findings can bedescribed as following:

1. It chances that time-lagged variables and error terms of the model are same as the corresponding ones in EF-ARIMA model, although they are usually not the same.
2. Although data series are not the same, from values of R^2 and $\bar{R}^2$, the explanatory ability of the EF-ARIMA model is better than that of the traditional model. It implies the EMD and filtering process can indeed reduce or remove the noise from original measurement data to make better model fit.

3.5. Validating Forecasting Performances

After both optimal models have been built, the testing samples, industrial production indices from January 2011 to December 2012, are used to evaluate their forecasting performances. Three widely used performance indexes, mean absolute error (MAE), mean absolute percent error (MAPE), root mean square error (RMSE) and Theil's inequality coefficients (U1 and U2) are used to evaluate the residuals between actual and predicted values. Among these four indexes, MAEmeasures the average magnitude of prediction errors, MAPE measures the mean prediction accuracy, RMSE measures the prediction stability, and the Theil's inequality coefficient reflects the RMSE error in relative terms. In principle, the lower MAE, MAPE,RMSEandTheil's inequality coefficient values are, the better the model performances are.

The actual index of the testing set and the forecasted monthly values using EF-ARIMA and the traditional ARIMA are illustrated in Figure 4. Overall the forecasted values obtained using EF-ARIMA are closer to the actual values than those obtained using traditional ARIMA.

From Table 4, we can see that the four indexes (MAE, MAPE, RMSE and Theil's inequality coefficients) of the proposed hybrid ARIMA approach are 11.02, 8.64% and 12.22 respectively. The forecasting performance of EF-ARIMA is clearing better than that of the traditional ARIMA. EF-ARIMA provides more reliable forecasts and improved forecasting performance in our empirical case study. Since MAPE is less than 10%, it is said to have the excellent level of predictive ability (Lewis, 1982). Our design and results show that the EMD and filtering process indeed grasps the underlying information in the original data, and the hybrid EF- ARIMA provides a better forecasting result.

Figure 4

Actual Values and Forecasts of Industrial Production Index on Testing Set

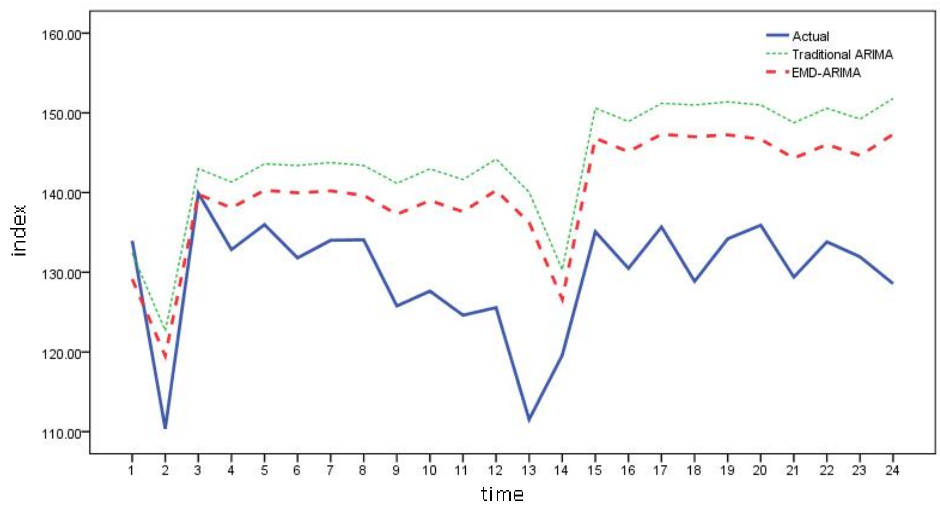


Table 4

The Forecasting Performances between Two Optimal Models

Models	MAE	MAPE	RMSE	U1	U2
EF-ARIMA	11.02	8.64%	12.22	0.0515	0.1070
Traditional ARIMA	14.57	11.40%	15.77	0.0576	0.1215

4. Conclusions

When the original data series is of measurement type, measurement errors are always unavoidable. In order to build **areliable** time series model with better forecasting performance, we proposed a process integrating the EMD process and filtering analysis in this study. It effectivelyreduces or removes the noise (error or unimportant message) from original data series and reconstructs the de-noise data series, which is expected to better grasp the underlying information.

To illustrate the goodness of model fit and forecasting performance, an EF-ARIMA model using the de-noise data was developed and its results were compared to those of the traditional ARIMA using the original measurement data. From the empirical case in this paper, three important findings are noted as following: (1) Integrating the EMD and filtering processes indeed can help reduce or remove the noise from original measurement; (2) Based on indexes of MAE, MAPE RMSE and Theil's inequality coefficients, the EF-ARIMA forecasts on testing samples are more reliable and accurate than forecasts using the traditional ARIMA model. (3) It also shows that the proposed hybrid forecasting approach is feasible and reliable.

Finally, in this paper, we just check whether $\sqrt{\Delta R^2_{(J)}}$ is statistically significant at the 5 percent level to identify the meaningful IMFs. Because it is up to the user to decide on the significance level, future researchers may also try to change this significance level to improve the de-noise data series.

References

Box, G.E.P. and Jenkins, G.M., 1976.*Time Series Analysis Forecasting And Control*. San Francisco: Holden Day.

Chen, C.F. Lai, M.C. and Yeh, C.C., 2012. Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge-Based Systems,* 26, pp.281-287.

Cheng, J.S. Yu, D.J., and Yang Y., 2006. A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration*, 294, pp. 269-277.

Hamad, K., Shourijeh, M.T., Lee, E. andFaghri, A., 2009. Near-term travel speed prediction utilizing Hilbert-Huang transform. *Computer-Aided Civil and Infrastructure Engineering,* 24(8), pp.551-576.

Huang, N.E. Shen, Z. and Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. *Annual Review of Fluid Mechanics,* 31, pp.417-457.

Huang, N.E. et al., 1998.The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.*The Royal Society,* A454, pp.903-995.

Huang, N.E. et al., 2003. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis.*Proceedings of the Royal Society of London*, A459, pp.2317-2345.

Lewis, C.D., 1982.*Industrial and Business Forecasting Method*. London: Butterworth Scientific.

Liu, X.J. Mi, Z.Q. Lu, B. and Wu, T., 2009. A novel approach for wind speed forecasting based on EMD and time-series analysis. IEEE, Asia-Pacific, pp. 1-4.

Quenouille, M.H., 1948. Some results in the testing of serial correlation coefficients. *Biometrika,* 35(3/4), pp.261-367.

Shen, Z.X. Huang, X.Y. andMa, X.X., 2008. An intelligent fault diagnosis method based on empirical mode decomposition and support vector machine. *Proceedings of the Third International Conference on Convergence and Hybrid Information Technology,* 1(11-13), pp.865-869.

Veltcheva, A.D. and Soares, C.G., 2004. Identification of the components of wave spectra by the Hilbert Huang transform method. *Applied Ocean Research,* 26(1-2), pp.1-12.

Vincent, H.T., Hu, S.-L.J. and Hou, Z., 1999.Damage detection using empirical mode decomposition method and a comparison with wavelet analysis. In:

Proceedings of the second international workshop on structural health monitoring, pp.891-900.

Wang, W., Zhao, H., Li, Q. and Liu, Z., 2009.A novel hybrid intelligent model for financial time series forecasting and its application.*International Conference on Business Intelligence and Financial Engineering,*pp.279-282.

Wei, Y. and Chen, M.C., 2012.Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks.*Transportation Research Part C: Emerging Technologies,* 21(1), pp.148-162.

Wu, M.C. and Hu, C.K., 2006.Empirical mode decomposition and synchrogram approach to cardiorespiratory synchronization. *Phys. Rev. E,* 73(051917), pp.1-11.

Wu, F.J. and Qu, L.S., 2008. An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery.*Journal Sound and Vibration*, 314, pp. 586-602.

Yang, H.L. and Lin, H.C., 2012.Combining artificial intelligence with non-linear data processing techniques for forecasting exchange rate time series, *International Journal of Digital Content and its Application*, 6(6), pp. 276 - 283.

Yang, J.Y. Zhang, Y.Y. and Zhu, Y.S., 2007. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension [J]. *Mechanical Systems and Signal Processing*, 21, pp. 2012-2024.

Yu, D.J., Cheng, J.S. and Yang, Y., 2005.Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings.*Mech. Syst. Signal Process,* 19(2), pp. 259-270.

Yu, L. Wang, S. and Lai, K.K., 2008. Forecasting crude oil price with and EMD-based neural network ensemble learning paradigm.*Energy Economics,*pp. 2623-2635.

Zhang, G. Patuwo, B.E. andHu, M.Y., 1998.Forecasting with artificial neural networks: the state of the art.*International Journal of Forecasting,* 14(1), pp.35-62.

Zhu, Z. Sun, Y. and Li, H., 2007. Hybrid of EMD and SVMs for short-term load forecasting. Proceedings of IEEE International Conference on Control and Automation, pp.1044-1047.