

Research Article

Selecting Single Model in Combination Forecasting Based on Cointegration Test and Encompassing Test

Chuanjin Jiang,^{1,2} Jing Zhang,² and Fugen Song¹

¹ *Glorious Sun School of Business and Management, Donghua University, Shanghai 200051, China*

² *Shanghai Business School, Shanghai 200233, China*

Correspondence should be addressed to Chuanjin Jiang; jiang@sbs.edu.cn

Received 8 March 2014; Accepted 29 March 2014; Published 22 April 2014

Academic Editor: Xiaofei Zhao

Copyright © 2014 Chuanjin Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Combination forecasting takes all characters of each single forecasting method into consideration, and combines them to form a composite, which increases forecasting accuracy. The existing researches on combination forecasting select single model randomly, neglecting the internal characters of the forecasting object. After discussing the function of cointegration test and encompassing test in the selection of single model, supplemented by empirical analysis, the paper gives the single model selection guidance: no more than five suitable single models can be selected from many alternative single models for a certain forecasting target, which increases accuracy and stability.

1. Introduction

In 1969, Bates and Granger [1] put forward the idea of “the combination of forecasts” for the first time; that is to say, after the characters of each single forecasting method are considered, combine different single forecasting. This method aroused the attention of forecasting scholars as soon as it was put forward.

The combination forecasting model is the progress of selecting and utilizing the information of single forecasting model, but there are fewer special researches on the selection problem of single model when forming a combination forecasting model. In previous combination forecasting researches and applications, often, there is no screening of single forecasting models involved in combination forecasting, and the forecasting model is chosen according to its obvious characters presented by particular progress and existing knowledge and experience.

During the regression analysis of economic measurement, an important index to evaluate the fitting degree is R^2 , but one of the most direct methods to raise R^2 is to increase explanatory variables. So a contradiction occurs: after the number of explanatory variables increases, problems such as

t -tests and colinearity may not be explained. That is to say, it is not the case that the more explanatory variables the better.

The researches of Armstrong [2] prove that the best combinations, most of the time, only use several (no more than 5) methods. When more and more single models are combined, the probability to get the optimum performance decreases.

So, this paper will discuss how to select suitable model from alternative single models to form combination for a certain forecasting target; according to some selecting principles, the number of single models to form combination is no more than five.

2. Cointegration Test of Single Model in Combination Forecasting

From the end of the 1980s to the 1990s, a great breakthrough of the modeling theory in econometrics is the cointegration relationship research of the time series. Cointegration actually presents the long-run equilibrium relationship of different time series, which is a key basic thought and theory in the current econometric field and is also an important

theoretical cornerstone in current researches on combination forecasting launched by time series.

If forecasting series is an objective performance of sample series, in the long term, they should have an equilibrium relationship, even though, in the short term, these variables may deviate from mean value due to random disturbance and other reasons. Therefore, we can calculate the forecasting series of all alternative single forecasting models and reject the alternative single forecasting models that do not conform to cointegration relationship through the cointegration test with sample series, and the goal of selecting single model can be achieved.

2.1. Brief Introduction to Cointegration. Suppose x_t is stationary time series, the mean is $E(x_t)$, $\text{var}(x_t)$ is the variance, covariance $\text{cov}(x_t, x_{t+k})$ is finite, and these numbers can fluctuate according to the changes of time t ; that is,

$$\begin{aligned} E(x_t) &= E(x_{t+h}) = \mu, \\ \text{var}(x_t) &= \text{var}(x_{t+h}) = \sigma^2, \\ \text{cov}(x_t, x_{t+k}) &= \text{cov}(x_{t+h}, x_{t+k+h}) = \rho. \end{aligned} \tag{1}$$

The three mathematical equations above hold true for all t, k, h , and the μ, σ^2, ρ here are constants.

Suppose some series requires difference transformation for d times before changing into stationary series; this series is integrated of order d , denoted by $I(d)$. Often, economic series is integrated of order one or two.

Suppose $X_{1t}, X_{2t}, \dots, X_{kt}$ series are all integrated of order d and a vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ exists, satisfying

$$Z_t = \alpha X_t' \sim I(d - b). \tag{2}$$

Let $b > 0$, $X_t^1 = (X_{1t}, X_{2t}, \dots, X_{kt})$; it can be considered that $(X_{1t}, X_{2t}, \dots, X_{kt})$ is cointegrated of order (d, b) [3].

The key of this result is to be able to explain that if some linear combination of two variables is $I(0)$, they must present the character of tracking each other, and these two variables will not deviate from each other too far in the long term. Therefore, in the long term, one of the necessary conditions for two variables to keep stationary is cointegration.

2.2. Cointegration Test of Single Models in Combination Forecasting. Generally speaking, economic series is unstable, so as a relative reasonable forecasting series. Supposing F_t and Y_t are forecasting series and real series, respectively, and both of them are integrated of order d , namely, $I(d)$, only when F_t and Y_t are cointegrated can forecasting error reach $I(0)$; otherwise, forecasting error would be unstable; that is, it will augment with time prolonging. The conclusion results in the 2 following theorems.

Theorem 1. *Given that both predicted series Y_t and forecasting series F_t belong to $I(d)$ and only when F_t and Y_t are cointegrated can forecasting error reach $I(0)$, thus the forecasting error is unstable.*

Theorem 2. *Only when every single forecasting in series Y_t 's combination forecasting in $I(d)$ is cointegrated with Y_t can forecasting error be $I(0)$; otherwise, combination forecasting error is unstable.*

The former parts in Theorems 1 and 2 are sufficient conditions while the latter parts are necessary conditions. This proposition can be learned that if the combination forecasting or single forecasting in unstable series is in inconformity with the requirements specified in the above proposition, the unstable forecasting error would exist.

Attestation of Theorem 1 can be directly obtained through related definitions of cointegration.

For different combination forecasting models, Theorem 2 has many proven ways, and only the proof process of linear regression method for combination forecasting is given below.

Assuming the forecasting series Y_t is of $I(d)$ type, then two single forecasting models of F_{1t} and F_{2t} in the series are combined to form C_t :

$$C_t = \beta_0 + \beta_1 F_{1t} + \beta_2 F_{2t}. \tag{3}$$

And $\beta_1 + \beta_2 = 1$. Assuming F_{1t} has a cointegration relation with Y_t , F_{1t} and Y_t are not cointegrated. From the definition of cointegration, it can be known that $(Y_t - F_{1t})$ belongs to $I(0)$, while, for any real number γ , $(Y_t - \gamma F_{2t})$ belongs to $I(d)$. The error of combination forecasting C_t is recorded into η_t ; namely,

$$\eta_t = -\beta_0 + \beta_1 (Y_t - F_{1t}) + \beta_2 (Y_t - F_{2t}). \tag{4}$$

It is obvious in the above equation that η_t belongs to $I(d)$, meaning that forecasting model has unstable forecasting errors, thus proving Theorem 2, which can easily apply to the combination of N kinds of single forecasting methods.

If the single model in combination forecasting is in inconformity with the cointegration conditions in Theorem 2, even creating a combination forecasting model cannot guarantee the forecasting accuracy because the established model forecasting error in this way will be unstable.

2.3. Steps of Cointegration Test. Step one of single model cointegration test in combination forecasting is to detect the integration order d of predicted series and the integration order d' of single forecasting series as Theorem 1. If the single forecasting series is cointegrated with predicted series, then $d = d'$. When the predicted sample series has the same integration order d as single forecasting series, the cointegration test can be made to the predicted sample series and single forecasting series, and the specific steps are as follows.

2.3.1. Integration Order Testing. The most widely used method is unit root test for integration order calculation of sample series or forecasting series. Standard unit root test method is Dicker-Fuller testing (DF), but the frequently used method in reality is generalized DF testing (ADF) [4]. ADF testing is actually making the following regression:

$$\Delta x_t = \beta_1 x_{t-1} + \beta_2 \Delta x_{t-1} + \beta_3 \Delta x_{t-2} + \beta_4 + \beta_5 t. \tag{5}$$

TABLE 1: N999 fitting sample series.

Quarter	Year										
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
1	3944.5	3969	4389	4244	4497	4938.5	5031	4895	5120.5	5790	6302.5
2	3901	4068	4311.5	4244	4475.5	5012.5	4994	5034	5262.5	5944.5	6207
3	3802.5	4250	4274.5	4185	4697.5	5071	4984.5	5018.5	5355	5984.5	6805.5
4	3737.5	4373.5	4222	4370.5	4913.5	4978.5	4975.5	5145	5475.5	6009.5	7185

The above formula adopts the lagged term $\beta_1 x_{t-1}$ of first degree and two differential lagged terms $\beta_2 \Delta x_{t-1}$ and $\beta_3 \Delta x_{t-2}$ of time series with the purpose to transform residue term to the white noise, and the number of terms is determined according to actual situations; constant term β_4 and time trend term $\beta_5 t$ represent the first difference of time series.

Unit root test is mainly performed to test the coefficient β_2 of the regression expression x_{t-1} ; if the coefficient significance was unequal to 0, the hypothesis that x has the unit root is rejected and x is a smooth series. Then the following hypothesis testing can be made:

$$H_0: \beta_2 = 0, \quad H_1: \beta_2 < 0. \quad (6)$$

The output result of ADF test includes the statistical magnitude t of coefficient β_2 in differential lagged terms and the tested critical value of the assumption whether it equals 0. If the output result is a negative number far smaller than 0, then H_1 is accepted and H_0 is rejected, which also indicates that the time series is stable with no unit root. The stability of first difference is also needed to be tested for unstable time series. If stable, then the series can be deemed as $I(1)$ type. If unstable, the first difference will continue the difference till the series becomes stable; when the series becomes stable after n -time differences, the integration order of the series is n .

2.3.2. *Cointegration Test.* To know whether two variables X and Y have cointegration correlation, EG two-step method [5] can be used for verification.

Step 1. Evaluate the following equation by ordinary least squares (OLS):

$$Y_t = \alpha_t X_t + \varepsilon_t. \quad (7)$$

So it can lead to

$$\hat{Y}_t = \hat{\alpha}_t X_t. \quad (8)$$

It is called cointegration regression.

Step 2. Test the integrity of $\hat{\varepsilon}_t$. If it is a stable series, then variables X, Y can be considered as cointegrated of order $(1, 1)$; if $\hat{\varepsilon}_t$ is integrated of order, then variables X, Y can be considered as cointegrated of order $(1, 2)$.

The testing methods for $\hat{\varepsilon}_t$'s integrity are the above DF testing or ADF testing.

Of course, in the specific application process, the following manner can also be carried out.

Expand the ADF regression of y_t to x_t through OLS method:

$$y_t = \text{const} + \beta x_t + \varepsilon_t. \quad (9)$$

Study the stability of the obtained ε_t regression residuals from this equation; if stable, then it indicates the presence of cointegration process, the equation can also be called cointegration equation, and the equation expresses the cointegration relationship between variables. The stability of the obtained ε_t regression residuals can be tested by ADF testing method; that is, it makes the following regression:

$$\Delta u_t = \gamma_0 u_{t-1} + \sum_{i=1}^k \gamma_i \Delta u_{t-i} + v_t. \quad (10)$$

In formula (10), the new error term is v_t ; k is the optimal lagged order to transform residue term v_t to white noise. Lagrange multiplication testing method or Durbin-Watson (DW) testing can be employed to determine whether the residual terms are white noise.

t testing in $\hat{\gamma}_0$ is ADF statistics; if it exceeds the critical value when the significance level is 0.01, 0.05, or 0.10, then the original hypothesis of noncointegration is rejected.

2.4. *Empirical Analysis.* M3-Competition sample series [6] is used below for empirical analysis of single model cointegration validation. The paper selects the 999th series in M3 (denoted by N999) as verification series; N999 is a macroeconomic series and quarterly data with totally 52 sample points, of which the first 44 samples are fitting samples while the last 8 are forecasting samples, (see Tables 1 and 2).

In the paper, 12 single models are selected to fit and predict N999 series, which are, respectively, moving average (Naïve), single exponential smoothing (Single), linear exponential smoothing (Holts), dampen trend exponential smoothing (Dampen), seasonal exponential smoothing (Winter), time series decomposition model (Decomposition), ARIMA, ARARMA, BP neural network (BP), NARX neural network (NARX), grey forecasting (GM), and support vector machine (SVM). Real things contain two kinds of information: the nonlinear information and linear information; therefore, some classical linear and nonlinear single models are considered in the paper.

The evaluation indexes of forecasting are of a large number, and Goodwin and Lawton [7] have listed many indicators used in forecasting evaluation currently; the paper selects "symmetric mean absolute percentage error" as the

TABLE 2: N999 forecasting sample series.

Quarter	Year	
	2001	2002
1	7373.5	7608
2	7259.5	7592.5
3	7558.5	7722
4	7660.5	7611

forecasting evaluation index. Y_t is the observation value at time t , and F_t represents the forecasting value of Y_t ; then

$$sMAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{(Y_t - F_t)}{(Y_t + F_t)} \right| \times 200. \tag{11}$$

sMAPE is proposed for the greater punishment shortcomings of MAPE to positive error. For example, if $Y_t = 100$, $F_t = 50$, then $MAPE = 50$; if $Y_t = 50$, $F_t = 100$, then $MAPE = 100$. But in both cases, sMAPE are 66.7. Furthermore, fluctuation range of sMAPE is 0–200, while MAPE has no upper limit.

Next, we conduct unit root test to N999 sample series through the test tool of Eviews6.0 software.

It can be observed from Table 3 that after a first difference to N999 sample series, ADF statistic magnitude is -5.025 , less than the 1% critical level. Therefore, N999 sample series can be determined as integrated of order; namely, it belongs to $I(1)$.

Next, unit root test is conducted for the fitting data of twelve single models, and only when the fitting sample of single model is also integrated of order can they go into cointegration test phase (see Table 4).

Through the unit root test, we can see that the Naïve and Single are stable series, and the integrated order of Auto-ANN is greater than 1, so these three single models are removed first.

The cointegration test is conducted to fitting series and sample series of the remaining nine single models; according to (9) and (10), stability test is actually made to the residues of nine single models. Table 5 shows the ADF testing of fitting sample residues of nine single forecasting models.

It can be observed from Table 5 that the forecasting of DAMPEN and GM(1,1) to fitting sample residue series of two single models does not pass ADF testing, so it is unstable; therefore these two single forecasting models do not pass the cointegration verification.

Finally, the remaining seven single models are predicted through the combination of simple average, weighted optimal, and ANN forecasting model again, and the result is shown in Table 6.

It can be observed from Table 6 that the combination forecasting accuracy after single model cointegration screening has improved to a certain extent, though not obvious. More importantly, forecasting accuracy of simple average combination (4.90) exceeds the Naïve (4.91) with little advantage, but it at least proves the cointegration verification validity of single forecasting model.

TABLE 3: Unit root test of N999 sample series after first difference.

(a)				
		t -statistic	Prob.*	
Augmented Dickey-Fuller test equation				
Dependent variable: $D(M, 2)$				
Method: least squares				
Date: 09/25/13, time: 00:18				
Sample (adjusted): 1990Q3 2000Q4				
Included observations: 42 after adjustments				
Variable	Coefficient	Std. error	t -statistic	Prob.
$D(M(-1))$	-0.822358	0.163652	-5.025046	0.0000
C	66.08964	24.93664	2.650303	0.0115
R -squared	0.386983	Mean dependent var.		10.07143
Adjusted R -squared	0.371658	S.D. dependent var.		182.3687
S.E. of regression	144.5601	Akaike info. criterion		12.83172
Sum squared resid.	835905.0	Schwarz criterion		12.91446
Log likelihood	-267.4660	Hannan-Quinn criter.		12.86205
F -statistic	25.25109	Durbin-Watson stat.		2.023140
Prob. (F -statistic)	0.000011			

*MacKinnon (1996) one-sided P values.

(b)				
Augmented Dickey-Fuller test equation				
Dependent variable: $D(M, 2)$				
Method: least squares				
Date: 09/25/13, time: 00:18				
Sample (adjusted): 1990Q3 2000Q4				
Included observations: 42 after adjustments				
Variable	Coefficient	Std. error	t -statistic	Prob.
$D(M(-1))$	-0.822358	0.163652	-5.025046	0.0000
C	66.08964	24.93664	2.650303	0.0115
R -squared	0.386983	Mean dependent var.		10.07143
Adjusted R -squared	0.371658	S.D. dependent var.		182.3687
S.E. of regression	144.5601	Akaike info. criterion		12.83172
Sum squared resid.	835905.0	Schwarz criterion		12.91446
Log likelihood	-267.4660	Hannan-Quinn criter.		12.86205
F -statistic	25.25109	Durbin-Watson stat.		2.023140
Prob. (F -statistic)	0.000011			

3. Encompassing Test of Single Model in Combination Forecasting

Numerous research literatures on forecasting accuracy mostly focus on the forecasting accuracy and stability evaluation and cannot directly compare the differences between competing forecastings of the same object. To achieve such discriminative comparison, forecasting encompassing theory came naturally, which was proposed by Meese and Rogoff [8, 9]: if a forecasting model contains or includes forecasting model B-related information, we believe that model A includes model B.

3.1. Significance of Encompassing Test of Single Model in Combination Forecasting. Kisinbay [10] establishes a statistical magnitude of approximate pivot in line with t distribution to apply encompassing test to single model selection in combination forecasting so as to filter the included single models and reduce the single model quantity involved in the combination. He believes that adding the included single models into combination models has not any sense and even increases the complexity of the algorithm.

Tomiyuki and Ryoji [11] point out that finding the single models of high predictability and small correlation degree with other models is important to improve the combination forecasting effects. Encompassing test is the answer to the competing models in combination forecasting. Therefore, it

TABLE 4: Unit root test on fitting data of 12 single forecasts.

Test sample	ADF statistics	1% critical value	IS stable	ADF statistics after first difference	1% critical value	IS stable after first difference	Result
Forecast sample	2.66	-3.59	No	-5.03	-3.60	Yes	$I(1)$
Naive	-6.02	-3.60	Yes	—	—	—	$I(0)$
SINGLE	-6.02	-3.60	Yes	—	—	—	$I(0)$
Holts	1.65	-3.60	No	-7.53	-3.59	Yes	$I(1)$
DAMPEN	0.15	-3.60	No	-16.02	-3.61	Yes	$I(1)$
WINTE	4.20	-3.59	No	-15.31	-3.61	Yes	$I(1)$
THETA	5.12	-3.59	No	-9.01	-3.60	Yes	$I(1)$
ARIMA	3.31	-3.60	No	-7.59	-3.59	Yes	$I(1)$
ARARMA	2.08	-3.60	No	-11.4	-3.60	Yes	$I(1)$
Auto-ANN	10.8	-3.61	No	1.55	-3.59	No	$I(2)$ or more
NARX	5.21	-3.60	No	-19.74	-3.60	Yes	$I(1)$
GM(1, 1)	3.11	-3.59	No	-12.56	-3.59	Yes	$I(1)$
SVM	4.10	-3.60	No	-4.21	-3.60	Yes	$I(1)$

TABLE 5: ADF test on residuals of 9 single forecasts' fitting sample.

The single forecast model residual sequence	ADF statistics	1% critical value	IS stable	IS cointegrated with sample series
Holts	-5.47	-4.21	Yes	Yes
DAMPEN	-2.56	-4.20	No	No
WINTER	-7.12	-4.22	Yes	Yes
THETA	-7.33	-4.21	Yes	Yes
ARIMA	-9.35	-4.23	Yes	Yes
ARARMA	-12.38	-4.22	Yes	Yes
NARX	-15.21	-4.23	Yes	Yes
GM(1, 1)	1.21	-4.22	No	No
SVM	-4.89	-4.21	Yes	Yes

TABLE 6: Result of combined forecasts of single models by cointegration test.

	Simple average	Weighted optimal	ANN
12 single models' combination without cointegration test	5.00	5.42	5.88
7 single models' combination with cointegration test	4.90	5.21	5.44

seems extremely important to create encompassing methods and principles of combination forecasting on the basis of statistical approaches, so as to improve combination forecasting efficiency.

Encompassing test can identify the origins of accuracy differences between competing models; help predictors distinguish whether the difference is caused by sample variability or the significance of information set in construction models. But we cannot ensure that the forecasting model with the best forecasting accuracy performance must be able to explain all competing models. The research of Ericsson [12] shows that taking advantages in forecasting accuracy out of samples

is not a sufficient condition for forecasting encompassing. In our empirical analysis, we further find that when the forecasting model A includes model B, the model A's accuracy may be lower than B's for containing more useless or inferior information.

Combination forecasting theory holds that the information in different forecasting models may vary due to the differences of models, even leading to different forecasting results. We can obtain better results than the forecasting values of originally two single models through combination forecasting. However, if no additional information is included in the constructed combination forecasting, this combination will be not able to get higher efficiency, and thus the combination forecasting loses its significance.

Based on the above analysis, besides cointegration test to single models before combination forecasting, we should also adopt forecasting encompassing test to identify the containment relationships between alternative single models to select single modes through certain heuristic strategies.

3.2. Basic Principle of Encompassing Test in Single Model Selection for Combination Forecasting. We first consider the two-two encompassing between single forecasting models

by making $f_1^{(t)} f_2^{(t)}$ the forecasting values of two single forecasting models at time t , and they have the following regression equation:

$$y_t = \beta_1 f_1^{(t)} + \beta_2 f_2^{(t)} + \varepsilon_t. \quad (12)$$

That model $f_1^{(t)}$ encompasses model $f_2^{(t)}$ can be written as $(\beta_1, \beta_2) = (1, 0)$, and in turn, it can also be that model $f_2^{(t)}$ includes model $f_1^{(t)}$. If (β_1, β_2) has other values, then the two models are intolerant and each model contains independent y_t -related information. If the covariance is relatively stable, then some standard methods can be used for encompassing verification test. If no any predictive model includes other models, it can be shown that all models are wrong (i.e., no one model can reflect the true data generating process); and only by relying on the combination forecasting can combination forecasting be made to useful data information.

Ericsson [12] constructs an encompassing test model:

$$y_t = \beta_1 f_1^{(t-s,t)} + \beta_2 f_2^{(t-s,t)} + \varepsilon_t. \quad (13)$$

$f_1^{(t-s,t)}$ means the forecasting value of the information at time $t-s$ used by $f_1^{(t)}$ to y_t , and $f_2^{(t-s,t)}$ means the forecasting value of the information at time $t-s$ used by $f_2^{(t)}$ to y_t . Under the restriction condition that $\beta_1 + \beta_2 = 1$, it is tested that $\beta_1 = 1$ or $\beta_2 = 0$.

Fair and Shiller [13] propose the following encompassing test model:

$$\Delta_s y_t = a_0 + \beta_1 (f_1^{(t-s,t)} - y_{t-s}) + \beta_2 (f_2^{(t-s,t)} - y_{t-s}) + \varepsilon_t. \quad (14)$$

If $\Delta_s y_t = y_t - y_{t-s}$, $\beta_1 = 0$, and $\beta_2 \neq 0$, model $f_1^{(t)}$ is included by model $f_2^{(t)}$; otherwise, model $f_2^{(t)}$ is included by model $f_1^{(t)}$, and β_1 and β_2 are unconstrained.

In testing process through (14), if the two models have the same information, the forecasting results will show a high correlation degree, and thus β_1 and β_2 are difficult to distinguish. Although this problem can obtain the evaluation of (14) through 0 regressions, the test results are not obvious.

Harvey, Leybourne, and Newbold provide HLN encompassing testing method for the above model defects [14]. HLN testing can verify and test whether two competing models have similar performances, and considering one simple c_t of loss difference sequence, it has

$$c_t = L(e_{1t}) - L(e_{2t}), \quad (15)$$

where $L(\cdot)$ represents any loss function, such as sMAPE or MSE; e_{1t} is the step forward forecasting error of model i to t ; $i = 1, 2$; $t = 1, \dots, T$. And the forecasting performance shows that the expected value of c_t is 0, namely, $E(c_t) = 0$; besides, the test adopts the mean of observation sample c_t . Assumed that the loss difference sequence has the stability of covariance, then, under the null hypothesis of the same forecasting accuracy, the test statistical magnitude of HLN

will also be more consistent with normal distribution, and the specific formula of the statistical test is as follows:

$$\text{HLN} = \frac{\bar{c}}{\sqrt{\widehat{V}(\bar{c})}}. \quad (16)$$

Assuming that the t step forward forecasting depends on the ordinal number $t-1$ and $\widehat{V}(\bar{c})$ is the consistent evaluation of asymptotic variance \bar{c} , we can calculate $\widehat{V}(\bar{c})$ through the following formula:

$$\widehat{V}(\bar{c}) \approx \frac{1}{T} \left(\gamma_0 + 2 \sum_{i=1}^{T-1} \gamma_i \right). \quad (17)$$

Here γ_i is the i th autocovariance of \bar{c} , whose size can be estimated by $\widehat{r}_i = T^{-1} \sum_{t=i+1}^T (c_t - \bar{c})(c_{t-i} - \bar{c})$.

Through the Monte Carlo method, we simulate and evaluate the testing capacity of HLN and find that it has poor performance in dealing with small samples.

Therefore, the paper amends on HLN testing so that its performance in testing small samples improves (after amendments, it can be referred to as MHLN testing). Specific amendments are as follows: first, compare and test statistical magnitude through t distribution critical value of $T-1$ degree of freedom, rather than a normal distribution; second, modify test statistical magnitude as follows:

$$\text{MHLN} = T^{-1/2} [T + 1 - 2t + T^{-1}t(t-1)]^{1/2} \text{HLN}. \quad (18)$$

Testing for MHLN statistic magnitude is very simple, and only by calling correlation test function in Matlab7.0 statistical toolbox can test results of various significant levels be obtained.

3.3. Steps of Encompassing Test of Single Models in Combination Forecasting. This section chooses N999 series in M3 as the sample series of encompassing test; through the cointegration selection in Section 2.2, the original twelve single forecasting models are only chosen to seven: Holts linear exponential smoothing, WINTER seasonal exponential smoothing, THETA time series decomposition model, BJ auto ARIMA, ARARMA, NARX neural networks, and SVM support vector machine. Here we will use MHLN encompassing testing described above for N999 sample series to conduct encompassing testing for the alternative 7 single models.

Step 1. Calculate the fitting sMAPE in the samples of 7 single models for N999 series and sort all models according to that numerical number and the performance of each model. It should be noted that the fitting sMAPE in the samples of single models is different from the forecasting sMAPE out of samples.

Step 2. Select the model with the lowest sMAPE value as the optimal model, and for optimal model NARX neural network, the adopted MHLN encompassing test (confidence level of 0.01) and t distribution table show that the remaining six single models are all not included by the optimal model.

TABLE 7: sMAPE of combined forecast of single models by inclusive test.

	Simple average	Weighted optimal	ANN
12 single models' combination without cointegration test	5.00	5.42	5.88
7 single models' combination with cointegration test	4.90	5.21	5.44
6 single models' combination with encompassing test (0.01 confidence level)	5.31	5.14	5.69
5 single models' combination with encompassing test (0.05 confidence level)	4.62	4.88	5.32

Step 3. Select the second optimal model ARARMA and operate as step two, and the result shows that ARIMA is included by ARARMA, so we remove the ARIMA model among alternative single models.

Step 4. Select the third optimal model of Decomposition and operate as step two again until the end of encompassing test and no included single forecasting model is found.

Since the number of single models in combination forecasting is better to be no more than 5, we repeat the above encompassing test procedures again and increase confidence level to 0.05. And the result shows that the Decomposition includes SVM.

Through encompassing test, seven single models are finally selected into five models. These five single models are used to make combination forecasting again, and further analysis is made on the basis of Table 6. For comparison, six single models that are selected the first time are compared together (confidence level of 0.01) (see Table 7).

We find that ARARMA includes ARIMA; but in fact, the real forecasting accuracy out of ARIMA samples is higher than that of ARARMA, which also proves that "taking advantages in forecasting accuracy out of samples is not a sufficient condition for forecasting encompassing" mentioned at the beginning of this chapter.

Furthermore, we find that when confidence level is 0.01, the simple average combination accuracy of the remaining six single models in conclusiveness test decreases while the accuracy of the weighted optimal combination upgrades. The main reason is that the eliminated ARIMA has a high forecasting accuracy itself, whose exclusion will necessarily decrease the accuracy of simple average combination; but for weighted optimal combination, the exclusion of ARIMA leads to the significant increase of weights of ARARMA and NARX neural network, so the combination forecasting accuracy improves.

When the confidence level is 0.05, SVM is excluded; the accuracy of the remaining five single model combinations has significantly improved. In particular, the sMAPE value of weighted optimal combination is lower than the Naïve single forecasting accuracy, proving the effectiveness of combination forecasting.

Therefore, when the single models are under encompassing tests, if the remaining single model number is no more than 5 when under the test at confidence level of 0.01, the test directly ends; if greater than 5, then confidence level should be adjusted to 0.05 or 0.1.

4. Conclusion

This thesis discusses single forecasting model selection in combination forecasting through cointegration test first and encompassing test method then. The result shows that the forecasting accuracy has improved to a certain extent after single model selection.

Cointegration verification aims to keep the fitting sample and real sample of single models have a consistent fluctuation trend, thus guaranteeing the combination forecasting accuracy of samples to the largest extent, especially the accuracy of middle and long-term forecasting. And the encompassing test is to eliminate the inclusive models; the start point of single model forecasting lies in the fact that each model establishes on different information set and owns different model type. If one single model is included by another single model, it will lose its meaning in the combination model and only increase the burden of combination.

According to the empirical analysis about N999 sample series, the paper screens 12 alternative single models to the final 5; and finally the forecasting accuracy is calculated to prove the efficiency of single model selection.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. M. Bates and C. W. J. Granger, "The combination of forecasts," *Operational Research Quarterly*, vol. 20, no. 4, pp. 451–468, 1969.
- [2] S. Armstrong, *Principles of Forecasting. A Handbook for Researchers and Practitioners*, KAP, Dordrecht, The Netherlands, 2001.
- [3] C. R. Nelson and C. R. Plosser, "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of Monetary Economics*, vol. 10, no. 4, pp. 139–162, 1982.
- [4] J. D. Hamilton, *Time Serial Analysis*, Social Science Press of China, Beijing, China, 1999.
- [5] R. F. Enger and C. W. J. Geanger, "Co-integration and error correction representation estimation and testing," *Econometric*, vol. 55, no. 2, pp. 251–276, 1987.
- [6] S. Makridakis and M. Hibon, "The M3-competition: results, conclusions and implications," *International Journal of Forecasting*, vol. 16, no. 4, pp. 451–476, 2000.
- [7] P. Goodwin and R. Lawton, "On the asymmetry of the symmetric MAPE," *International Journal of Forecasting*, vol. 15, no. 4, pp. 405–408, 1999.
- [8] R. A. Meese and K. Rogoff, "Empirical exchange rate models of the seventies: do they fit out of sample?" *Journal of International Economics*, vol. 14, no. 1-2, pp. 3–24, 1983.

- [9] R. Meese and K. Rogoff, "Was it real? The exchange rate-interest differential relation over the modern floating-rate period," *Journal of Finance*, vol. 43, pp. 933–948, 1988.
- [10] T. Kisinbay, "The use of encompassing tests for forecast combinations," IMF Working Paper WP/264, International Monetary Fund, 2007.
- [11] K. Tomiyuki and K. Ryoji, "The effectiveness of forecasting methods using multiple information variables," IMES Discussion Paper Series 2002-E-20, Institute for Monetary and Economic Studies, 2002.
- [12] N. R. Ericsson, "Parameter constancy, mean square forecast errors, and measuring forecast performance: an exposition, extensions, and illustration," *Journal of Policy Modeling*, vol. 14, no. 4, pp. 465–495, 1992.
- [13] R. Fair and R. Shiller, "Comparing information in forecasts from econometric models," *The American Economic Review*, vol. 80, pp. 375–389, 1990.
- [14] D. I. Harvey, S. J. Leybourne, and P. Newbold, "Tests for forecast encompassing," *Journal of Business and Economic Statistics*, vol. 16, no. 2, pp. 254–259, 1998.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

