EKONOMETRIA ECONOMETRICS 1(39) • 2013
ISSN 1507-3866

Joanna Perzyńska
West Pomeranian University of Technology, Szczecin, Poland
joanna.perzynska@zut.edu.pl

COMPARATIVE ANALYSIS OF ACCURACY OF SELECTED METHODS OF BUILDING OF COMBINED FORECASTS AND META-FORECAST

Abstract: In this paper the author presents a method of building a meta-forecast as an arithmetic mean of the combined forecasts set by various methods. The empirical example, in which the forecasts (individual, combined and meta-forecasts) are determined for the microeconomic variable with seasonal fluctuations, is the illustration of theoretical considerations. The accuracy of meta-forecasts is compared with the accuracy of their component combined forecasts and individual forecasts. The empirical studies confirm the usefulness of meta-forecasts. In most cases, they have lower errors than their component combined forecasts, also they are more accurate than individual forecasts.

Keywords: individual forecasts, combined forecasts, meta-forecasts, forecasts errors, forecasts weights.

1. Introduction

In cases when there are different forecasts of the same variable, rather than making an arbitrary choice from the best of them, we can build a combined forecast which is a linear or non-linear combination of the available individual forecasts. Combined forecasts can be built using simple methods (such as arithmetic mean or median), also their weight can be chosen subjectively or estimated so as to minimize the combined forecast error [see Armstrong 2001].

As we can see, also this time we have to choose an analytical form of the combined forecast and a method of its building. Some attempt to solve this problem would be the construction of the meta-forecast, which is a simple arithmetic mean of the combined forecasts built using different methods.

The presented procedure of constructing the meta-forecasts is based on twice averaging the individual forecasts. In the empirical research we will verify the hypothesis that the meta-forecasts obtained in this way will be more accurate than the combined forecasts that are of their components.

2. Research material

For the modeling and forecasting, a variable was selected which describes the total costs of heat and electricity energy production (KCE) in power plant B. The time series of KCE variable includes 60 months. Figure 1 shows the shaping of the forecasted variable.

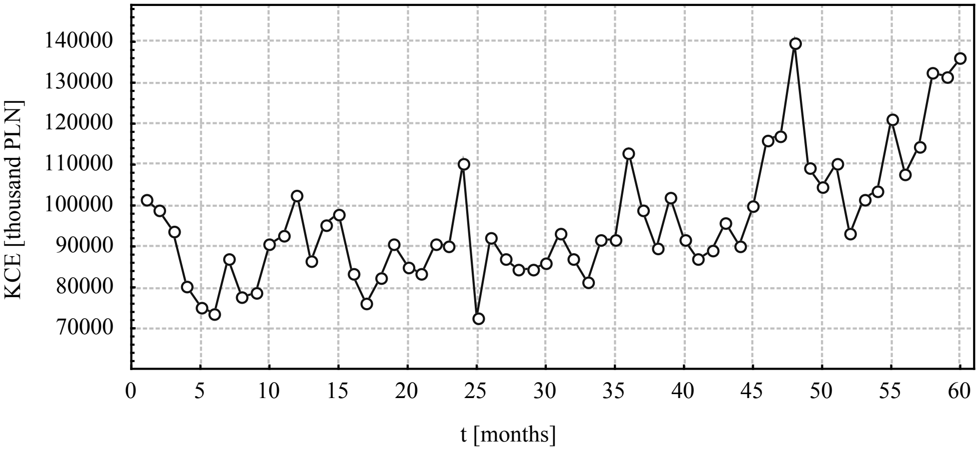

Figure 1. Total costs of heat and electricity energy production (KCE) in power plant B

Source: own study.

Explanatory variables in the causal-descriptive models are:

PC – total production of heat energy in power plant B (in GJ),

PE – total production of electricity energy in power plant B (in MWh),

PCE – total production of heat and electricity energy in power plant B (in MWh),

SC – income from sales of heat energy in power plant B (in thou. PLN),

SE – income from sales of electricity energy in power plant B (in thou. PLN),

SCE – income from sales of heat and electricity energy in power plant B (in thou. PLN).

In the forecasted variable and its explanatory variables we can find seasonal fluctuations. Table 1 shows the multiplicative seasonal indicators of the analyzed variables obtained for estimation period $t = 1, 2, ..., 36$.

Values of the seasonal indicators of the forecasted variable and its explanatory variables are varied. The intensity of seasonal fluctuations in KCE variable is moderate – the difference between the maximum and minimum values of seasonal factors is 31.79 percentage points. The explanatory variables have a moderate (PE, SE and SCE), strong (PCE) and very strong (PC and SC) intensity of seasonality.

Table 1. Seasonal indicators of KCE variable and its explanatory variables for estimation period $t = 1, 2, ..., 36$

Month	KCE	PC	PE	PCE	SC	SE	SCE
I	95.93	196.25	110.43	128.23	188.09	108.22	116.70
II	103.14	169.60	98.57	113.29	161.35	98.30	104.96
III	106.07	162.30	104.96	116.66	147.65	103.01	107.65
IV	95.15	95.87	99.95	99.09	94.84	93.49	93.62
V	90.14	46.47	94.15	84.69	55.38	90.32	86.75
VI	92.90	29.84	95.37	82.48	42.64	91.99	86.97
VII	102.17	27.57	97.04	82.19	40.36	94.75	88.93
VIII	94.42	26.75	97.55	82.47	41.71	96.68	90.77
IX	92.01	32.16	90.00	77.72	46.56	94.58	89.39
X	102.86	102.29	101.30	101.57	96.97	100.04	99.72
XI	103.28	139.85	103.56	111.17	126.56	100.71	103.46
XII	121.93	171.05	107.11	120.45	157.90	127.91	131.09
max – min	31.79	169.50	20.43	50.51	147.73	37.59	44.34

Source: own study.

The lowest amplitude of seasonal fluctuation was observed for explanatory variables PE and SE (20.43 per cent and 37.59 per cent), the highest for PC and SC (169.50 per cent and 147.73 per cent).

3. Description of research procedure

The process of modeling and forecasting of KCE variable was divided into 5 stages.

In the first step, for estimation period $t = 1, 2, ..., 3$, six types of models were estimated: classical and hierarchical time series models with seasonal fluctuations, causal-descriptive classical and hierarchical models with seasonal changing parameters, Holt-Winters models and artificial neural networks. Depending on the class, differences between estimated models equations occurred in: the independent variables, the analytical form of the trend, types of seasonal fluctuations, exponential smoothing parameters or structure of neural networks. Based on the estimated models, the expired forecasts of KCE variable were constructed for the next 12-month period ($t = 37, 38, ..., 48$).

In the second stage, one form from each class of model was selected, based on the mean absolute percentage error of forecast ($MAPE_{37\text{-}48}$), analysis of goodness-of-fit and significance test of model parameters.

In the third stage, after the estimation time was increased to 48 observations ($t = 1, 2, ..., 48$), six selected earlier models were estimated once more. Based on them, individual *ex post* forecasts for KCE variable were constructed for the next 12 months ($T = 49, 50, ..., 60$) – they were labeled: $f_{1T}, f_{2T}, f_{3T}, f_{4T}, f_{5T}, f_{6T}$.

In the fourth stage there were constructed linear and non-linear combined forecasts of KCE variable on period T, for 57 combination of individual forecasts, which contained from 2 to 6 component forecasts. Linear combined forecasts were calculated as a weighted average of the individual forecasts:

$$f_{CT,mk} = \sum_{i=1}^{m} \lambda_{ik} f_{ikT}, \tag{1}$$

assuming:

$$\sum_{i=1}^{m} \lambda_{ik} = 1, \tag{2}$$

where: $f_{CT,mk}$ – combined forecast of KCE variable on period T,
m – number of component forecasts of combined forecasts ($m = 2,3,\dots,6$),
k – number of combination of component forecast ($k = 1,2,\dots, 57$),
f_{ikT} – i-th component forecast in k-th combination ($i = 1, 2,..., m$),
λ_{ik} – weight of i-th component forecast in k-th combination.

Non-linear forecasts were constructed as:

$$f_{CT,mk} = \psi\left(f_{1kT}, f_{2kT}, \dots, f_{mkT}\right), \tag{3}$$

where ψ – nonlinear function.

In the construction of the non-linear function artificial neural networks (nANN) were used [Liu et al. 1996). Linear weights of combined forecasts were counted using the methods:

– simple arithmetic mean (AM):

$$\lambda_i = \frac{1}{m}; \tag{4}$$

– variance-covariance (VC) [Bates, Granger 1969; Granger, Newbold 1974]:

$$\lambda = \frac{\Sigma^{-1}1}{1'\Sigma^{-1}1}, \tag{5}$$

where Σ – matrix of variance-covariance of component forecasts errors;

- Bates-Granger's (BG) which is a special case of the VC method assuming zero correlation between the errors of component forecasts;
- regression assuming non-negative weights (NERLS) [Aksu, Gunter 1997];
- minimization of module of errors autocorrelation coefficient (Au);
- minimization of MAPE (M);
- multi-criteria optimization [Kaźmierska-Zatoń, Zatoń 2010]:

– simultaneous minimization of module of errors autocorrelation coefficient and MAPE (AuM),
– simultaneous minimization of module of errors autocorrelation coefficient and Theil's coefficient (AuT),
– simultaneous minimization of MAPE and Theil's coefficient (MT),
– simultaneous minimization of module of errors autocorrelation coefficient, MAPE and Theil's coefficient (AuMT);
– linear artificial neural networks (lANN) [Perzyńska 2010].

The variance-covariance method and linear artificial neural networks can give negative and greater than zero combination coefficients (weights). Using these or non-linear artificial neural networks, we can obtain the value of the combined forecast exceeding the scope of individual forecasts.

In the construction of the weights of combined forecasts (excluding the AM method) we used two sources of information: values of expired ex-post forecast, their errors (variant W1) and theoretical values of individual models (variant W2).

A division of the methods on single-criterion (Jk) or multi-criteria (Wk) was caused by the numbers of criteria used for the evaluation of quality of forecast, which were used to construct the weights of the combined forecasts. Table 2 shows the specification of the methods which were applied in both variants.

Table 2. Specification of methods of building of combined forecasts

No.	Variant W1			No.	Variant W2		
	method	weights	criteria		method	weights	criteria
1	VC1	+/-	Jk	10	VC2	+/-	Jk
2	BG1	+	Jk	11	BG2	+	Jk
3	NERLS1	+	Jk	12	NERLS2	+	Jk
4	M1	+	Jk	13	lANN2	+/-	Jk
5	Au1	+	Jk	14	nANN2	+/-	Jk
6	AuM1	+	Wk				
7	AuT1	+	Wk				
8	AuMT1	+	Wk				
9	MT1	+	Wk				

1/2 – number of variant; Jk/Wk – number of criteria; + positive weights; – negative weights.

Source: own study.

In the fifth stage, meta-forecasts of the KCE variable were built on period *T*. The meta-forecasts were a simple mean of the combined forecasts, which were constructed earlier by different methods, for the same combination of component forecasts and the same period:

$$f_{MT,mkJ} = \frac{1}{J}\sum_{j=1}^{J} f_{CT,mkj}, \tag{6}$$

where: $f_{MT,mkJ}$ – meta-forecast of KCE variable on period T,
$f_{CT,mkj}$ – combined forecast $f_{CT,mk}$ built by j-th method ($j = 1, 2, \ldots, J$),
J – numbers of combined forecasts which were components of meta-forecast ($2 \leq J \leq 15$).

4. Presentation and analysis of research results

Depending on the presented procedure, six individual models were selected and ex-post forecasts of the KCE variable were built for 12 months ($t = 49, 50, ..., 60$). Table 3 shows the mean absolute percentage error of individual forecasts.

Table 3. Mean absolute percentage error of individual forecasts

Component forecast	$MAPE_{49\text{-}60}$
f_1	11.80
f_2	11.80
f_3	10.15
f_4	10.50
f_5	9.84
f_6	8.26

Source: own study.

The lowest errors were received for individual forecasts calculated for artificial neural network (8.26%) and the exponential smoothing model (9.84%). Forecasts which were constructed on descriptive models had errors only slightly lower than the time series models. This means that the decrease of forecasts' accuracy was determined by the high amplitude of seasonal fluctuations.

Received individual forecasts were the component of the combined forecasts. For each m combination of component forecasts (m = 2, 3, ..., 6), based on the methods specified in Table 2, linear and non-linear combined forecast were built. In Table 4 averages of mean absolute percentage errors ($MAPE_{49\text{-}60}$) of combined forecasts are presented, which contain the same number of component forecasts. For $m = 6$ the values of $MAPE_{49\text{-}60}$ (this case was marked as "*") were presented.

An analysis of the information contained in the last column of Table 4 shows that the lowest value of average of $MAPE_{49\text{-}60}$ (5.95%) was obtained by method NERLS1. Only a slightly higher error (5.98%) was obtained for the multi-criteria optimization method AuMT1. For other methods the average values of $MAPE_{49\text{-}60}$ were in the 6.00% to 6.46% range. The maximum value was obtained by method VC1.

Forecasts constructed with the AM method had a lower accuracy than the combined forecasts, which were obtained in variant W1 – the average value of their $MAPE_{49\text{-}60}$ was 6.76%. This is a higher value than for method NERLS1 (by 0.81 per

Table 4. Averages of $MAPE_{49\text{-}60}$ of combined forecasts

Method	*m* = 2	*m* = 3	*m* = 4	*m* = 5	*m* = 6*	*m* = 2 – 6
AM	7.42	6.77	6.35	6.18	6.12	6.76
VC1	7.67	6.07	6.31	5.08	6.75	6.46
BG1	7.57	6.19	5.50	4.93	4.46	6.21
NERLS1	7.71	5.88	4.98	4.48	4.34	5.95
M1	8.00	6.27	5.48	5.06	4.53	6.36
Au1	8.09	6.01	4.83	4.10	4.04	6.01
AuM1	7.82	6.03	4.98	4.53	4.06	6.03
AuT1	8.00	6.08	5.09	4.54	4.05	6.13
AuMT1	7.78	5.92	4.97	4.55	4.06	5.98
wMT1	7.79	5.96	4.98	4.51	4.50	6.00
VC2	10.22	10.60	10.97	10.82	10.16	10.61
BG2	9.14	8.76	8.81	8.90	8.94	8.89
NERLS2	10.18	9.70	10.02	9.71	9.20	9.90
lANN2	9.13	7.05	8.36	7.67	7.00	8.01
nANN2	8.01	7.58	6.72	5.28	5.97	7.20

Source: own study.

cent) and VC1 method (by 0.30 per cent). The AM method was better than all other methods in variant W2.

In variant W2, the lowest value of average of $MAPE_{49\text{-}60}$ (7.20%) was obtained for the combined forecasts constructed by non-linear artificial neural networks (nANN). This is a higher value than for the methods AM and NERLS1 (by 0.44 per cent) and the VC1 method (by 1.25 per cent). Simultaneously, this value is less (by 1.06 per cent) from the lowest error of individual forecast (8.26%). Average values of $MAPE_{49\text{-}60}$ less than 8.26% were also obtained for the method lANN2 (8.01%).

Comparing the average accuracy of the combined forecasts for $m = 2, 3, ..., 6$ we can see that the lowest average values of $MAPE_{49\text{-}60}$ were obtained for different methods: for $m = 2$ and $m = 3$ the best forecast were constructed by methods AM and NERLS1; the best method of forecasting for $m = 4, 5, 6$ was Au1. In other variants, these methods were not as accurate: for the $m > 2$ methods in variant W1 were more accurate than AM, for the $m = 2$ method AU1 was the least accurate from all of the methods in variant W1.

The accuracy of the selected combined forecasts was also compared with the accuracy of the component forecasts. Table 5 shows the percentage of the combined forecasts whose errors were lower than the lowest error of their component forecasts.

Analysis of the information contained in the last column of Table 5 shows that in about 80% of cases, for all methods in variant W1 and methods nANN2 and AM, the combined forecasts were better than the component forecasts. In fewer than 80% but in more than 50% of cases in variant W2, better forecasts were obtained for the

Table 5. Percentage of combined forecasts with errors lower than the lowest error of their component forecasts

Method	*m* = 2	*m* = 3	*m* = 4	*m* = 5	*m* = 6	*m* = 26
AM	66.7	80.0	93.3	100.0	100.0	82.5
VC1	73.3	95.0	93.3	100.0	100.0	89.5
BG1	66.7	80.0	93.3	100.0	100.0	82.5
NERLS1	73.3	90.0	100.0	100.0	100.0	89.5
M1	73.3	85.0	93.3	100.0	100.0	86.0
Au1	73.3	95.0	100.0	100.0	100.0	91.2
AuM1	80.0	95.0	100.0	100.0	100.0	93.0
AuT1	80.0	95.0	100.0	100.0	100.0	93.0
AuMT1	80.0	95.0	100.0	100.0	100.0	93.0
MT1	73.3	90.0	100.0	100.0	100.0	89.5
VC2	40.0	20.0	6.7	0.0	0.0	19.3
BG2	60.0	55.0	60.0	16.7	0.0	52.6
NERLS2	33.3	35.0	6.7	0.0	0.0	22.8
lANN2	66.7	75.0	66.7	83.3	100.0	71.9
nANN2	93.3	85.0	86.7	100.0	100.0	89.5

Source: own study.

methods BG2 and 1ANN2. For two other methods in variant W2 this percentage was lower (below 23%).

In the cases $m = 5, 6$ for the methods nANN2 and AM, and all methods of variant W1, the percentages were 100% – this means that all of the combined forecasts determined using these methods were more accurate than their component forecasts.

In the last stage of the research, meta-forecasts were built as a simple mean of combined forecasts constructed by different methods. Table 4 presents the averages of the mean absolute percentage errors of meta-forecast (MF) and their component combined forecasts (CF) for different groups of methods. For comparison, Table 6 also contains averages of $MAPE_{49\text{-}60}$ of combined forecasts obtained by AM.

Table 6. Averages of $MAPE_{49\text{-}60}$ of meta-forecasts and combined forecasts for different groups of methods

Method	*m* = 2		*m* = 3		*m* = 4		*m* = 5		*m* = 6		*m* = 2–6	
	MF	CF	MF	CF	MF	CF	MF	CF	MF	CF	MF	CF
AM	-	7.42	-	6.77	-	6.35	-	6.18	-	*6.12	-	6.76
AM+W1+W2	8.11	8.30	6.54	6.99	6.32	6.56	5.72	6.02	5.63	5.88	6.42	7.10
W1	7.51	7.83	5.64	6.05	4.80	5.24	4.42	4.64	4.47	4.53	5.80	6.13
W2	8.48	9.34	7.61	8.74	7.83	8.98	7.19	8.48	7.49	8.25	7.81	8.92
Jk	7.64	8.57	6.20	7.41	5.86	7.20	5.12	6.60	5.31	6.54	6.30	7.56
Wk	7.50	7.85	5.61	6.00	4.67	5.01	4.19	4.53	4.07	4.17	5.59	6.04

Source: own study.

Analysis of the information contained in Table 6 shows that the averages of $MAPE_{49\text{-}60}$ of meta-forecasts were lower than the averages' errors of their component combined forecasts. In most cases (for $m > 2$ and excluding W2), meta-forecasts were more accurate than the combined forecasts obtained by the AM method.

The lowest errors of meta-forecasts were obtained for a group of multi-criteria methods and methods of variant W1, and the highest for variant W2 – but in most cases they were lower than the smallest individual forecast error.

Note that the meta-forecasts determined as the arithmetic mean of all combined forecasts (AM + W1 + W2), and therefore the least accurate forecasts from W2 variant, were more accurate than their component forecasts and individual forecasts.

5. Conclusions

Empirical studies in most cases confirmed that the methods used to construct the combined forecasts and meta-forecasts were useful.

The highest accuracy of combined forecasts was achieved by the application of the multi-criteria method and single-criterion method in variant W1. A relatively low error of forecast was obtained with the AM method and nonlinear artificial neural networks (multilayer perceptrons) – in most cases the combined forecasts determined by these methods were more accurate than their individual component forecasts.

The empirical study shows that it is difficult to identify only one method which we should use to build combined forecasts. In this situation, the construction of meta-forecasts, as a simple arithmetic means of combined forecasts, which were obtained using different methods, allows us to avoid the problem of the selection of one method and also increases the accuracy of the forecasts. In most cases, meta-forecasts had errors lower than their component combined forecasts. They were also more accurate than individual forecasts. This can be a good solution when all individual forecasts are overestimated (or underestimated) because it appears that in this way we might obtain some overestimated combined forecasts and the rest of them underestimated, and the meta-forecast as their average could have a small error.

The presented procedure of the construction of meta-forecasts can be extended to other classes of individual models and variables, which differ by the intensity of seasonal fluctuation. It appears that in cases when there are different forecasts of the same variable, a good solution is the construction of one meta-forecast as a simple arithmetic mean of the combined forecasts built using selected methods of varying values of combination coefficients and an analytical form of combination.

Literature

Aksu C., Gunter S. (1997), The usefulness of heuristic N(E)RLS algorithms for combining forecasts, *Journal of Forecasting* 16: 439–463.

Armstrong J. (2001), *Principles of Forecasting: a Handbook for Researchers and Practitioners*, Kluwer Academic Publishers, Boston.

Bates J., Granger C. (1969), The combination of forecasts, *Operational Research Quarterly* 20.
Granger C., Newbold P. (1974), Experience with forecasting univariate time series and the combination of forecasts, *Journal of the Royal Statistical Society*, A, 137.
Kaźmierska-Zatoń M., Zatoń W. (2010), Multi-criteria combined forecasts, *Econometrics* 28 (91), 59–75, Wrocław.
Liu B., Shi S., Xu L. (1996), Improving the accuracy of nonlinear combined forecasting using neural networks, *Expert Systems with Applications* 16: 49–54.
Perzyńska J. (2010), Budowa prognoz kombinowanych z wykorzystaniem sztucznych sieci neuronowych, [in:] P. Dittmann, E. Szabela-Pasierbińska (eds.), *Prognozowanie w zarządzaniu firmą*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 103, 133–145.

ANALIZA PORÓWNAWCZA DOKŁADNOŚCI WYBRANYCH METOD BUDOWY PROGNOZ KOMBINOWANYCH I METAPROGNOZ

Streszczenie: W artykule przedstawiono propozycję metody budowy metaprognoz jako średnich arytmetycznych prognoz kombinowanych wyznaczonych za pomocą różnych metod. Ilustracją rozważań o charakterze teoretycznym jest przykład empiryczny, w którym prognozy (indywidualne, kombinowane oraz metaprognozy) wyznaczono dla zmiennej mikroekonomicznej wykazującej wahania sezonowe. Dokładność metaprognoz porównano z dokładnością ich składowych prognoz kombinowanych oraz prognoz indywidualnych. Przeprowadzone badania empiryczne potwierdziły użyteczność metaprognoz. W większości przypadków były one obarczone niższymi błędami niż ich składowe prognozy kombinowane, okazały się one również bardziej trafne niż prognozy indywidualne.

Słowa kluczowe: prognozy indywidualne, prognozy kombinowane, metaprognozy, wagi prognoz, błędy prognoz.