**DYNAMIC ECONOMETRIC MODELS**
Vol. 10 – Nicolaus Copernicus University – Toruń – 2010

*Mariola Piłatowska*
*Nicolaus Copernicus University in Toruń*

# Choosing a Model and Strategy of Model Selection by Accumulated Prediction Error

A b s t r a c t. The purpose of the paper is to present and apply the accumulative one-step-ahead prediction error (APE) not only as a method (strategy) of model selection, but also as a tool of model selection strategy (meta-selection). The APE method is compared with the information approach to model selection (AIC and BIC information criteria), supported by empirical examples. Obtained results indicated that the APE method may be of considerable practical importance.

K e y w o r d s: model selection, meta-selection, information criteria, accumulative prediction error.



## 1. Introduction

In the literature different methods (strategies) of model selection are available, among others: strategies based on sequences of tests (forward/backward selection), strategies related to information criteria of Akaike type, strategies based on predictive criteria (out-of-sample validation), which can be treated as mainstream directions in model selection. For the reason that the true generating model is unknown in practice, the focus in model selection strategies is being moved from the issue of selection the only one, true model to the issue of selection the best model among the set of candidate models fitted to the data or selection of several plausible models, where the best model may have relatively weak support against others models (Burnham, Anderson, 2002). Selection of the best model or multi-model inference assumes that the set of models has been well founded, because even the relatively best model in a set might be poor in an absolute sense.

Associated with each strategy is an algorithm to be specified which within given data enables to choose the best (in some sense) model among the candidate models (generally they may be nested or non-nested models, different models based on different scientific theories or modeling assumptions). However, the problem of model selection implies not only the choice of model in the framework of a given strategy but also the choice of model selection strategy. The focus in the literature is mainly on the choice of model or the comparison of different model selection strategies with regard to the choice of the best model, without touching the issue of model selection strategy.

The choice of model selection strategy and its suitability and properties may depend on the goals of an analysis (estimation, prediction), sample size (some strategies perform in different way in small and large samples), characteristics of the data generating model (DGM)[1]. In practice, there is a need to propose a data-driven framework which allows to help choosing a model selection strategy without making any reference to the actual DGM. This identification is called the meta-selection of a model (De Luna, Skouras, 2003). The meta-selection framework obeys the 'prequential' principle (Dawid, 1984)[2] which abandons the goal of selecting the true model in favor of seeking as small a predictive error as possible by comparing obtained predictions from each strategy and the actual values observed for the data independent on which model was used to forecast (Clarke, 2001). The essential point for this approach is that the adequacy of a model must be reflected in accurate prediction regardless of the goals of an analysis, i.e. if the goal of analysis is model estimation (model identification or hypothesis testing), then the best model should give the best predictions.

The purpose of the paper is to present and apply the accumulative one-step-ahead prediction error (APE) not only as a method (strategy) of model selection, but also as a tool of model selection strategy (meta-selection).

## 2. Accumulative One-Step-Ahead Prediction Error

The choice of model according to the accumulative prediction error (APE) consists in evaluating how well the models in the set are able to predict the next unseen data point $x_{n+1}$. In other words, according to the APE method the most useful model is the model with the smallest out-of-sample one-step-ahead prediction error. The prediction error cannot be calculated because $x_{n+1}$ has not been observed. What can be calculated, however, are the prediction error for $x_{i+1}$ based on the previous $x^i$ $(0 < i < n)$ by the sum of the previous one-step-ahead prediction errors for data that are available.

Let us consider a time series of $n$ observations, $x^n = (x_1, x_2, ..., x_n)$.

[1] Some strategies are optimal depending on whether the data generating model is one of the candidate models or not (Shao, 1997).

[2] 'Prequential' is from predictive sequential (Dawid, 1984).

The APE method proceeds by calculating sequential one-step-ahead forecasts based on a gradually increasing part of the data. For model $M_j$ the APE is calculated as follows (Wagenmaker, Grunwald, Steyvers, 2006):

1. Determine the smallest number $s$ of observations that makes the model identifiable. Set $i = s + 1$, so that $i - 1 = s$.
2. Based on the first $i - 1$ observations, calculate a prediction $\hat{p}_i$ for the next observation $i$.
3. Calculate the prediction error for observation $i$, e.g. squared difference between the predicted value $\hat{p}_i$ and the observed value $x_i$.
4. Increase $i$ by 1 and repeat steps 2 and 3 until $i = n$.
5. Sum all of the one-step-ahead prediction errors as calculated in step 3. The result is the APE.

For model $M_j$ the accumulative prediction terror is given by:

$$\mathrm{APE}(M_j) = \sum_{i=s+1}^{n} d[x_i, (\hat{p}_i \,|\, x^{i-1})],$$

where $d$ indicates the specific loss function that quantifies the discrepancy between observed and predicted values.

Applying the APE method the form of prediction should be considered: whether to predict using a single value (Skouras, Dawid, 1998) or a probability distribution (Aitchison, Dunsmore, 1975). In the first case, the predictions $\hat{p}_i$ are predictions for the mean value of *i*th outcome $x_i$. In the latter case, $\hat{p}_i$ is a distribution on the set of possible outcomes $x_i$.

The choice of the loss function should be considered in order to quantify the discrepancy between predicted values and observed values. This can be measured in a variety of different ways. For a single-value predictions, one typically uses the squared error $(x_i - \hat{p}_i)^2$. Another choice would be to compute the absolute value loss $|x_i - \hat{p}_i|$ or more generally an α-loss function $|x_i - \hat{p}_i|^{\alpha}$, where $\alpha \in [1, 2]$ (Rissanen, 2003). For probabilistic predictions, one typically uses the logarithmic loss function $-\ln \hat{p}_i(x_i)$, thus the loss depends on the probability mass or density that $\hat{p}_i$ assigns to the actually observed outcome $x_i$. The larger the probability, the smaller the loss[3].

[3] Taking the logarithmic loss function makes the APE method compatible with maximum likelihood, Bayesian inference and minimum description length (MDL) (amongst others: Wagenmaker, Grünwald, Steyvers, 2006).

The APE method can be also applied to select the model selection strategy (de Luna, Skouras, 2003). Let $S_1, S_2, ..., S_q$ $k = 1, 2, ..., q$ be $q$ potential model selection strategies applicable to a given set of model $P_p(\theta_p)$, $p = 1, 2, ..., M$ which approximate the data generating model. The parameters $\theta_p$ assigned to each model have to be estimated. If each strategy leads to an identical choice of model $p$, there is no real reason for selecting a given strategy. In the case of disagreement, however, a strategy $S_k$, $k = 1, 2, ..., q$, is selected for which the accumulated prediction error

$$\text{APE}(S_k) = \sum_{i=m}^{n} L(x_i, \hat{x}^{i-1}(S_k)),$$

reaches the minimum, where $\hat{x}^{i-1}(S_k)$ is the prediction $\hat{x}^{i-1}(p)$ resulting from the choice of model $p$ made by the $S_k$ strategy based on the sub-sample $x_1, x_2, ..., x_{i-1}$.

Hence the $\text{APE}(S_k)$ measures the predictive performance when strategy $S_k$ was used to form predictions sequentially, by updating not only the estimated parameters in each step but the choice of model as well (the meta-selection method computes APE for model selection method instead of models). The meta-selection should not just focus on the minimization of $\text{APE}(S_k)$, but also on its evolution for increasing sample sizes.

## 3. Empirical Example

To present the predictive performance of accumulated one-step-ahead prediction error (APE) in model selection and the choice of model selection strategy the data from Maddison base[4] have been taken. It includes annual time series of GDP for 36 countries. In the study, as an example, the GDP for France (1947-2003) and Poland (1952-2003) have been used. Data are expressed in millions of US dollars in constant prices from 1990 having taken into account purchasing power parity.

The essential point in model selection is the identification of initial set of candidate models. In this study the set of models consist of two models: ARIMA(1,1,0) and linear trend with autoregression of second order (T+AR(2)). This choice of models is justified by the traditional approach to the analysis of GDP fluctuations. During last thirty years this analysis focused on either the verification of unit root hypothesis (what means that GDP is nonstationary in variance or has stochastic trend and the ARIMA model is more appropriate) or testing hypothesis of stationary deviations around deterministic trend (what

[4] In the paper the updated Maddison base is used which is available on website www.ggdc.net, see also: Maddison (2001).

means that GDP is nonstationary in mean and model with deterministic trend is more appropriate). In spite of huge literature devoted to the distinguishing of these alternative hypothesizes, this dispute has not be settled upon yet[5].

Model ARIMA(1,1,0) was selected from different specification of ARIMA($p$, $d$, $q$) model, for $p$, $q$ = 0, 1, 2, $d$ = 0, 1, by the means of AIC differences[6], i.e. $\Delta_i = \text{AIC}_i - \text{AIC}_{\min}$, where $\text{AIC}_i$ denotes the AIC value for $i$-th model, $\text{AIC}_{\min}$ – AIC value for the best model. Models were estimated on the same sample length, i.e. 1947-2000 (GDP in France) and 1952-200 (GDP in Poland). The larger $\Delta_i$ is, the less plausible the fitted model is the good model in the K-L information sense[7], given the data. In practice, the models with $\Delta_i < 4$ are accepted (Burnham, Anderson, 2002). Having $\Delta_i$ the Akaike weights (evidence ratios) can be obtained which are useful in calculating the relative evidence for the best model (with the biggest weight) versus the rest of $R$-models in the set. The Akaike weights are given by (Burnham, Anderson, 2002; Piłatowska, 2009, 2010): $w_i = \exp(-0{,}5\Delta_i)/\sum_{r=1}^{R}\exp(-0{,}5\Delta_r)$, $\sum_{i=1}^{R} w_i = 1$. For ARIMA(1,1,0) the difference $\Delta_i$ was equal to zero, i.e. this model was the best, and for the rest of models $\Delta_i < 3$, so, they were plausible in the K-L information sense. However, the support for the ARIMA(1,1,0) was substantial (i.e. it had the dominating weight equal to 0.55).

[5] To papers concerning the choice of stochastic trend (nonstationarity in variance) versus deterministic trend (nonstationarity in mean) for GDP series belong among others: Nelson, Plosser, 1982; Stock, Watson, 1986; Quah, 1987; Perron, Phillips,1987; Christiano, Eichenbaum,1990; Rudebusch,1993; Diebold, Senhadji, 1996; Murray, Nelson, 1998. It is pointed out ((Haubrich, Lo, 2001) that the reason of no settlement in this dispute is the false assumption that one of the above hypothesizes is true. As a result, only the possibility of persistent fluctuations (shocks to GDP are persistent and there is no trend reversion at all) or transitory fluctuations (shocks are transitory and trend reversion occurs) is taken into account, but the indirect fluctuations, i.e. long memory dependence, are omitted, and the latter can be described by different model than previously, i.e. ARFIMA model.

[6] In the paper the modified AIC (second-order variant of AIC) was applied, i.e. $\text{AIC}_c = \text{AIC} + \frac{2K(K+1)}{n-K-1}$, where $\text{AIC} = -2\ln L + 2K$, $K$ denotes the number of estimated parameters, $n$ – sample size. Standard AIC may perform poorly (may indicate not parsimonious model), if there are too many parameter in relation of the size of the sample. The use of $\text{AIC}_c$ is advocated when the ratio $n/K$ is small, say < 40, (Sugiura, 1978). For the purposes of presentation further only 'AIC' notation is used.

[7] The Kullback-Leibler (K-L) distance or information is the measure of discrepancy between true (but unknown) model and fitted model. Akaike (1973) showed that the choice of model with minimum relative expected information loss (i.e. model with minimum K-L information) is asymptotically equivalent to the choice of model with minimum AIC.

In similar way the specification of an alternative model to ARIMA was chosen, i.e. model of linear trend with autoregression of second order T+AR(2), where maximum lag length was equal to 3.

To make a choice between ARIMA(1,1,0) model and T+AR(2) model three model selection strategies were used: information criteria: AIC and BIC, and also accumulated one-step-ahead prediction error (APE). In the latter case the squared error (APE_SE) and absolute error (APE_AE) were taken as a loss function[8]. The estimation[9] of models has been starting with minimum sample size equal to 11 observations, then the sample size has been increased by one until $n$ (until the year 2000) and the estimation was repeated. At each stage criteria: AIC and BIC, the forecasts from both types of models and accumulated one-step-ahead prediction error (APE_SE and APE_AE) were calculated. Results in form of differences among AIC, BIC and APE for both types of models depending on sample size are presented in Figures 1 (GDP in France) and 2 (GDP in Poland).

Figure 1 (panel A and B) shows that as the sample size increases the criteria AIC and BIC give a general support for the T+AR(2) model, because the difference of criteria: AIC(ARIMA)-AIC(T+AR(2)) and BIC(ARIMA)-BIC(T+AR(2)) is positive (what denotes smaller value of AIC and BIC for model T+AR(2)); only for a few periods: 18$^{th}$ (a year 1975), 28$^{th}$ and 29$^{th}$ (a year 1985 and 1986) the difference of criteria is negative, what gives a preference for the ARIMA(1,1,0) model in these periods.

However, observing the evolution of difference in APE (APE_SE and APE_AE) for both types of models no support for the T+AR(2) model as in the case of AIC and BIC is obtained – see panel C and D. Almost in the whole forecast period the difference in APE_SE for both models[10] is negative what leads to a general preference for the ARIMA(1,1,0) model when the GDP for France is to be forecast – see panel C (with exception of first 3 observations referring to 1958-1960 period, 12$^{th}$ and 13$^{th}$ observations referring to 1969-1970). Different performance shows the difference in APE_AE for both models (Figure 1, panel D), i.e. it favors the ARIMA(1,1,0) model from 4$^{th}$ observation up until the data set has increased to $n = 35$ (what refers to 1961-1993 period), after which it starts to prefer the T+AR(2) model[11]. This means that the

[8] Accumulated prediction error (APE) was calculated using gretl script written by author for that purpose.

[9] Model ARIMA(1,1,0) has been estimated by maximum likelihood method, and model T+AR(2) – least squares method.

[10] The notion APE_SE(ARIMA(1,1,0)-APE_SE(T+AR(2)) stands for the difference in APE_SE calculated for both models – see Figure 1.

[11] The negative difference in APE_AE denotes better predictive performance (smaller one-step-ahead prediction errors) of the ARIMA(1,1,0) model than the T+AR(2) model, and the positive difference in APE_AE – on the contrary.

choice of model will depend on the loss function taken to calculate the accumulated prediction error.

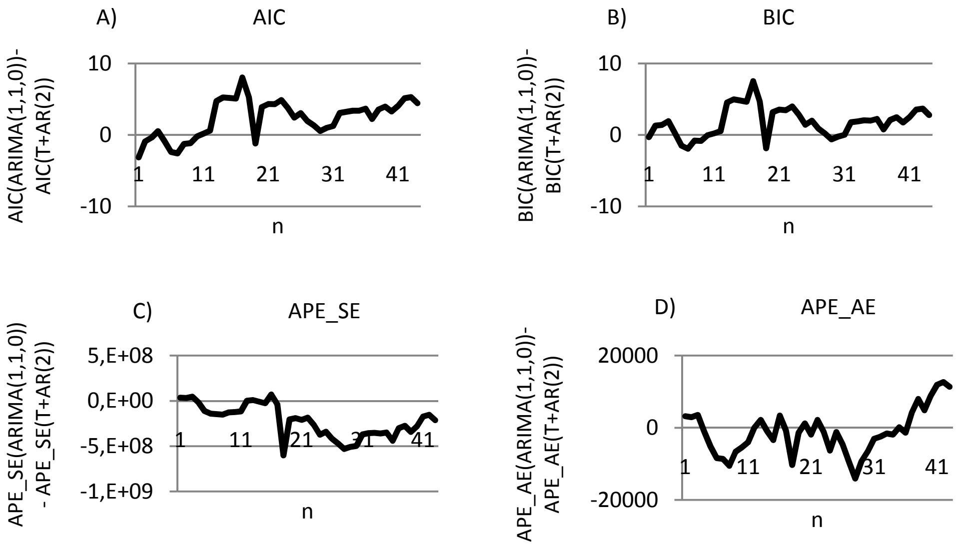


Figure 1. Difference between choice criteria for the ARIMA(1,1,0) model and the T+AR(2) model using to obtain forecasts of GDP in France. Panel A – AIC, panel B – BIC, panel C – APE_SE, panel D – APE_AE

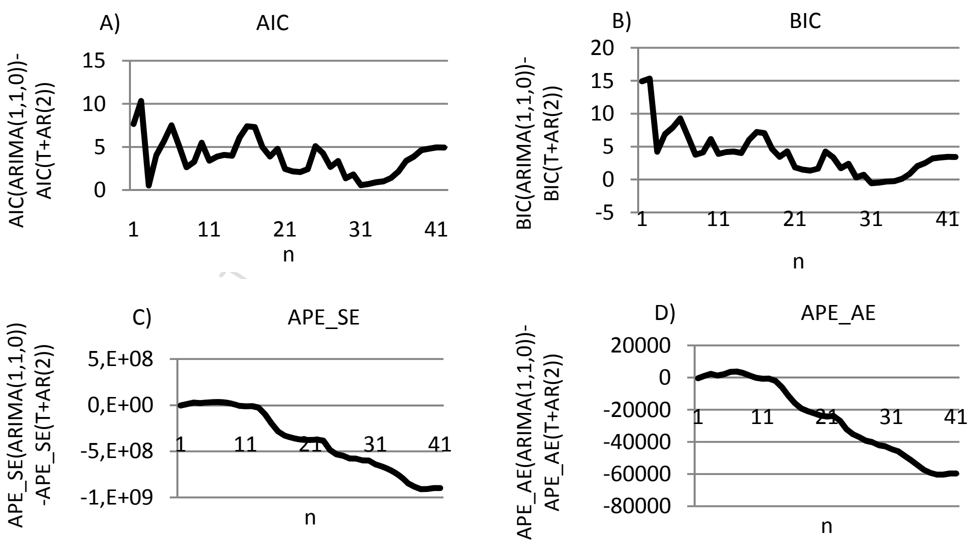


Figure 2. Difference between choice criteria for the ARIMA(1,1,0) model and the T+AR(2) model using to obtain forecasts of GDP in Poland. Panel A –AIC, panel B – BIC, panel C – APE_SE, panel D – APE_AE

When forecasting the GDP in Poland – see Figure 2, panel A and B – the positive difference of AIC and BIC criteria for alternative models indicates that the T+AR(2) model is to be preferred over the ARIMA(1,1,0) model. However, in the case of BIC the support for the T+AR(2) model decreases as the sample size increases what is seen in decreasing difference in BIC for both models. The opposite pattern shows the difference in APE for both models (APE_SE, APE_AE – see panel C and D), i.e. it indicates the substantial preference for the ARIMA(1,1,0) model (negative difference in APE_SE and also APE_AE for both models) and better predictive performance (smaller one-step-ahead prediction errors) almost in entire data set except the $2^{nd}$ and $9^{th}$ observations (1960 and 1968 periods).

An alternative method in assessing the performance for model selection methods is to quantify their predictive performance through a model meta-selection procedure. The aim of this procedure is to evaluate predictive value not of the models (e.g. ARIMA, ARMA), but the model selection methods (AIC, BIC, APE). Just as in the calculation of APE earlier, the meta-selection procedure requires to fit the ARIMA(1,1,0) and T+AR(2) models (in above case) for each of an increasing (by one) number of observations. The predictive value of, say AIC, is then quantified by the accumulative prediction error for the models chosen by AIC. For instance, suppose that for a particular time series, AIC prefers the ARIMA model up until the data set has increased to $n = 20$, after which AIC starts to prefer the T+AR(q) model. Then the accumulative prediction error for the AIC model selection procedure is a sum of the prediction errors made by the ARIMA and T+AR(q) models (for the first and second half of the time series respectively). Having calculated the difference in APE for different model selection procedures (strategies), the relative value of model selection tools as e.g. AIC is obtained. Figure 3 depicts the differences in accumulated prediction errors (APE) for various model selection procedures, i.e. AIC, BIC, APE_SE, APE_AE.

For particular time series (GDP in France) panel A in Figure 3 demonstrates that the use of AIC for model selection results in smaller one-step-ahead prediction error than the use of BIC (because the difference in APE_SE for AIC and BIC model selection methods (APE_SE(AIC)-APE_SE(BIC)) is negative)[12]. Note that horizontal stretches in Figure 3 indicate that the difference in accumulated prediction errors between two model selection strategies does not change (e.g. AIC and BIC, panel A). This occurs when two model selection strategies prefer the same model. The results are about the same when the absolute error (AE) was used as a loss function (panel D).

[12] The abbreviation, e.g. APE_SE(AIC) stands for the accumulated prediction error (with squared error, SE, as a loss function) calculated when the AIC procedure was used to select a model from two ones: ARIMA or T+AR(2) in the example at hand.

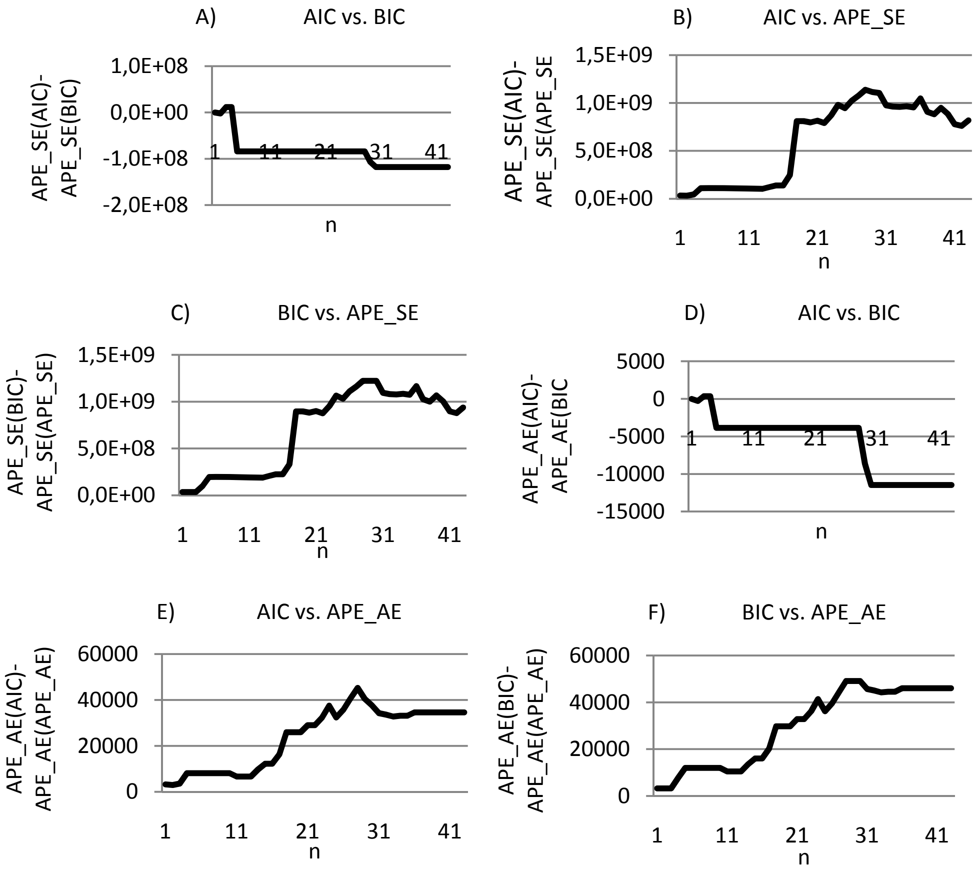


Figure 3. Model meta-selection as a function of the number of observations. Each panels shows the difference in APE for pairs of various model selection methods: AIC, BIC, APE_SE and APE_AE for GDP in France

Comparing the performance for pairs of model selection strategies, i.e. AIC and APE_SE, BIC and APE_SE (panel B and C, Figure 3) evidently smaller prediction error are obtained when the APE_SE strategy was used to select a model than AIC and BIC strategies[13]. Similar results are observed when the absolute error was taken as a loss function (panel E and F) except first ten periods when the difference in APE_AE is constant what denotes that both strategies (AIC vs. APE_AE and BIC vs. APE_AE) perform about the same. Generally, the use of APE_SE (or APE_AE) strategy leads to smaller accumulated prediction error than AIC or BIC strategy.

[13] The differences APE_SE(AIC)-APE_SE(APE_SE) are positive in entire data set what leads to a preference of APE_SE strategy.

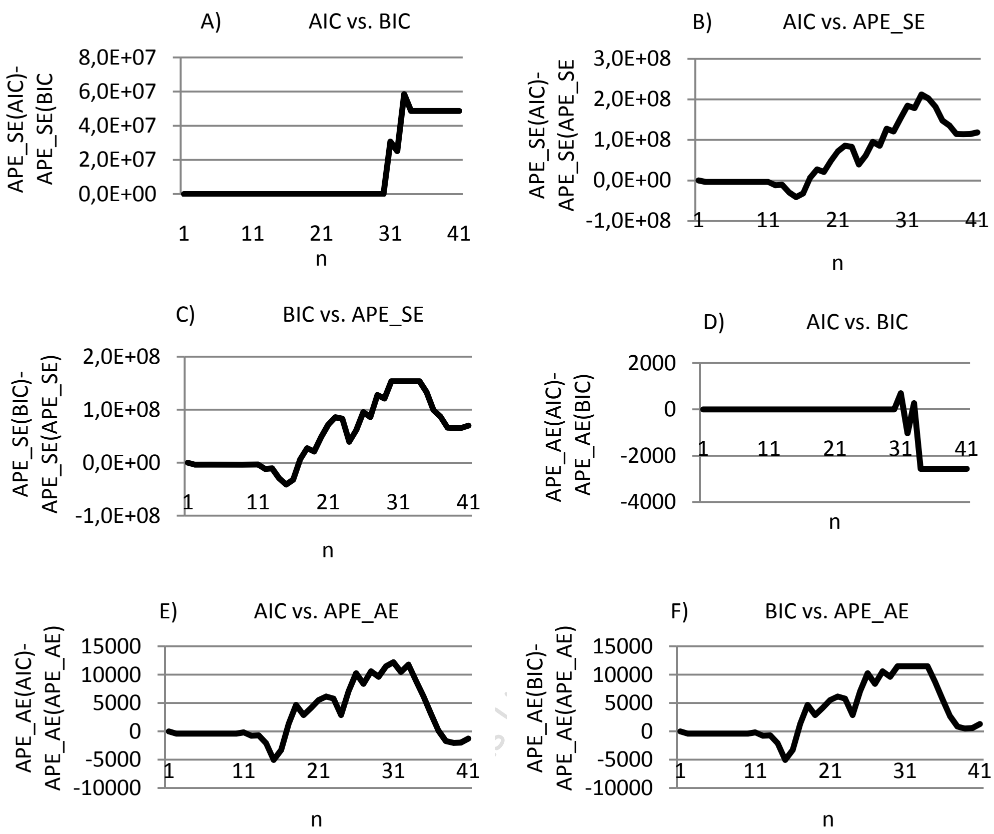


Figure 4. Model meta-selection for various model selection methods: AIC, BIC, APE_SE and APE_AE for GDP in Poland

For another series, GDP in Poland, the performance of AIC and BIC strategies is the same for the first 30 periods of data set (referring to 1960-1990 period) because the difference APE_SE(AIC)-APE_SE(BIC) is equal to zero (Figure 4) – but for the rest of data set the use of BIC strategy for model selection results in relatively smaller one-step-ahead prediction errors than in the case of AIC strategy (panel A, Figure 4), because the difference in APE_SE for AIC and BIC strategies is positive. However, when the absolute error (AE) is used as a loss function, the results are opposite, i.e. the strategy AIC is to be preferred (the difference in APE_AE for AIC and BIC strategies is negative, see panel D). This confirms earlier conclusion that the choice of model as well the choice of model selection strategy depends on the form of loss function.

Comparing the performance for pairs of model selection strategies, i.e. AIC and APE_SE, BIC and APE_SE (panel B and C, Figure 4) results that the APE_SE strategy performs better in model selection (i.e. gives smaller accumulated prediction errors) than AIC or BIC strategy almost in the entire data set except first 15 periods (1959-1965 period) when the difference in APE_SE for various pairs of strategies (AIC vs. APE_SE and BIC vs. APE_SE) are negative, and then the AIC and BIC strategies respectively are preferred. About the same results are obtained when the performance of AIC vs. APE_AE and BIC vs. APE_AE is compared (panel E and F, Figure 4) except the end of data set when the relative decrease for support of APE_AE strategy is noticed (the difference in APE_AE for various pairs of strategies is positive, but decreasing).

Table 1. One-step-ahead forecasts of GDP in France made by ARIMA(1,1,0) model and T+AR(2) model with prediction errors

| Forecast period | Realization | Model: ARIMA(1,1,0) forecast | $\delta_T$ | $\delta^*_T$ | Model: T+AR(2) Forecast | $\delta_T$ | $\delta^*_T$ |
|---|---|---|---|---|---|---|---|
| 2001 | 1289387 | 1297071 | -7684.3 | -0.60% | 1292864 | -3477.0 | -0.27% |
| 2002 | 1305136 | 1312186 | -7050.8 | -0.54% | 1309083 | -3947.3 | -0.30% |
| 2003 | 1315601 | 1323622 | -8021.1 | -0.61% | 1321281 | -5680.3 | -0.43% |

Table 2. One-step-ahead forecasts of GDP in Poland made by ARIMA(1,1,0) model and T+AR(2) model with prediction errors

| Forecast period | Realization | Model: ARIMA(1,1,0) forecast | $\delta_T$ | $\delta^*_T$ | Model: T+AR(2) Forecast | $\delta_T$ | $\delta^*_T$ |
|---|---|---|---|---|---|---|---|
| 2001 | 281508 | 286913 | -5406 | -1.92% | 286307 | -4798.8 | -1.70% |
| 2002 | 285365 | 284901 | 464 | 0.16% | 283789 | 1575.8 | 0.55% |
| 2003 | 296237 | 289382 | 6856 | 2.31% | 288394 | 7843.2 | 2.65% |

To check the choice of model (ARIMA or T+AR(2)) made by the accumulated prediction error (APE_SE and APE_AE) one-step-ahead forecasts of GDP in France and Poland were calculated in out-of-sample (i.e. 2001-2003 period). These forecasts with prediction errors (absolute $\delta_T$ and relative $\delta^*_T$) are showed in Table 1 and 2.

It is seen in Table 1 that one-step-ahead prediction errors are smaller when forecasts of GDP in France are made from T+AR(2) model what confirms the choice of model by the APE_SE method (see Figure 1, panel D). However, the prediction errors from the ARIMA(1,1,0) model are only slightly higher what would suggest the predictive value also for that model. This means that although the T+AR(2) model is preferred, the ARIMA model may be also useful in forecasting.

Forecasting the GDP in Poland the smaller one-step-ahead prediction errors are obtained when forecasts are made from ARIMA(1,1,0) model which was indicated by the APE method (see Figure 2, panel C and D).

## 4. Conclusions

The presented empirical example indicates the usefulness of one-step-ahead accumulated prediction error (APE) as a method of model selection. The APE method is conceptually straightforward, as it accumulates 'honest' one-step-ahead prediction errors, i.e. its predictions always concern unseen data. Additionally, observing the evolution of APE as the number of observations is increased, suggests that the choice of best model should be referred to the number of observations, that is, the best model in given sample size may be replace with another model which has better prediction value.

The APE method can be applied to nested and non-nested models alike and it is sensitive to the functional form of the model parameters (Myung, Pitt, 1997), and not just to their number as in AIC and BIC method. Also, the APE is a data-driven method that does not rely on the accuracy of asymptotic approximations. In particular, the use of APE does not require to include the true (data generating process) model into the set of candidate models. Unquestionable advantage of APE is that can be used not only for the selection of models, but also for the selection of model selection methods, and thus, the comparison of various model selection methods can be carried out. Hence the APE method enhances the issue of model selection and therefore may be of considerable practical importance.

## References

Aitchison, J., Dunsmore, I. R. (1975), *Statistical Prediction Analysis*, Cambridge University Press, Cambridge.

Akaike, H. (1973), *Information Theory and an Extension of the Maximum Likelihood Principle*, [in:] Petrov B. N., Csaki F., *Second International Symposium on Information Theory*, Kiado Academy, Budapest.

Burnham, K. P., Anderson, D. R. (2002), *Model Selection and Multimodel Inference*, Springer,

Christiano, L. J., Eichenbaum, M. (1990), Unit Roots in Real GNP: Do We Know and Do We Care?, *Carnegie-Rochester Conference Series on Public Policy*, no. 32, 7–61.

Clarke, B. (2001), Combining Model Selection Procedures for Online Prediction, *Sahkhya: The Indian Journal of Statistics*, 63, series A, 229–249.

Dawid, A. P. (1984), Statistical Theory: the Prequential Approach, *Journal of Royal Statistical Society Series B*, 147, 278–292.

De Luna, X., Skouras, K. (2003), Choosing a Model Selection Strategy, *Scandinavian Journal of Statistics*, 30, 113–128.

Diebold, F. X., Senhadji, A. (1996), Deterministic vs. Stochastic Trend in U.S. GNP. Yet again, NBER Working Papers, nr 5481.

Haubrich, J. G., Lo, A. W. (2001), The Source and Nature of Long-Term Memory in Aggregate Output, *Federal Reserve Bank of Cleveland „Economic Review"*, QII, 15–30.

Maddison, A. (2001), *The World Economy – a Millennial Perspective*, OECD Development Centre, Paris.

Murray, C., Nelson, C. (1998), The Uncertain Trend in U.S. GNP, Discussion Papers in Economics at the University of Washington, nr 0074.

Myung, I. J., Pitt, M. A. (1997), Applying Occam's Razor in Modeling Cognition: A Bayesian approach, *Psychonomic Bulletin and Review*, 4, 79–95.

Nelson, C. R. , Plosser, C. I. (1982), Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications, *Journal of Monetary Economics*, 10(2), 139–162.

Perron, P., Phillips, P. C. B. (1987), Does GNP Have a Unit Root?, *Economics Letters*, 23, 129–145.

Piłatowska, M. (2009), Prognozy kombinowane z wykorzystaniem wag Akaike'a (Combined Forecasts Using Akaike Weights), *Acta Universitatis Nicolai Copernici, Ekonomia*, XXXIX, 51–62.

Piłatowska, M. (2010), Kryteria informacyjne w wyborze modelu ekonometrycznego (Information Criteria in Model Selection), *Studia i Prace Uniwersytetu Ekonomicznego* w Krakowie, 25–37.

Quah, D. (1987), What do we Learn from Unit Roots in Macroeconomic Series?, NBER Working Papers, nr 2450.

Rissanen, J. (2003), Complexity of Simple Nonlogarithmic Loss Function, *IEEE Transactions on Information Theory*, 49, 476–484.

Rudebusch, G. D. (1993), The Uncertain Unit Root in Real GNP, *American Economic Review*, 83(1), 264–272.

Shao, J. (1997), An Asymptotic Theory for Linear Model Selection, *Statistica Sinica*, 7, 221–264.

Skouras, K., Dawid, A. P. (1998), On Efficient Point Prediction Systems, *Journal of Royal Statistical Society* B, 60, 765–780.

Sugiura, N. (1978), Further Analysis of the Data by Akaike's Information Criterion and the Finite Corrections, *Communications in Statistics, Theory and Methods*, A7, 13–26.

Stock, J., Watson, M. (1986), Does GNP Have a Unit Root?, *Economics Letters*, 22(2/3), 147–151.

Wagenmaker, E-J., Grünwald, P., Steyvers, M. (2006), Accumulative Prediction Error and the Selection of Time Series Models, *Journal of Mathematical Psychology*, 50, 149–166.

## Wybór modelu i strategii selekcji modelu za pomocą skumulowanego błędu predykcji

Z a r y s t r e ś c i. Celem artykułu jest prezentacja i wykorzystanie skumulowanego błędu prognoz na jeden okres naprzód (APE) nie tylko jako metody (strategii) wyboru modelu, ale również jako narzędzie do wyboru samej strategii (meta-wybór). Na przykładach empirycznych metoda APE jest porównywana z metodami wykorzystującymi kryteria informacyjne (AIC i BIC). Otrzymane wyniki wskazują na dużą praktyczną przydatność metody APE.

S ł o w a k l u c z o w e: wybór modelu, meta-wybór, kryteria informacyjne, skumulowany błąd prognoz