

Uniwersytet Mikołaja Kopernika w Toruniu  
Katedra Ekonometrii i Statystyki

*Mariola Piłatowska*

## PROGNOZY KOMBINOWANE Z WYKORZYSTANIEM WAG AKAIKE’A

**Z a r y s t r e ś c i.** W artykule uwaga jest skupiona na podejściu wykorzystującym kryteria informacyjne, a w szczególności kryterium Akaike’a, które jest wykorzystywane do wyznaczenia wag Akaike’a. Podejście to umożliwia otrzymanie nie jednego, a kilku wiarygodnych modeli, dla których można stworzyć ranking stosując wagi Akaike’a. Modele te stanowią podstawę obliczenia prognoz indywidualnych, a te z kolei służą do wyznaczenia ostatecznej prognozy kombinowanej, przy formułowaniu której wykorzystuje się wagi Akaike’a.

**S ł o w a k l u c z o w e:** prognozy kombinowane, systemy wag, kryteria informacyjne.

### 1. WSTĘP

Dzięki rozwojowi metod analizy szeregów czasowych oraz zdolności obliczeniowych komputerów, prognozując tę samą zmienną ekonomiczną na podstawie różnych metod można łatwo otrzymać wiele różnych prognoz. Do wyboru najlepszej prognozy wykorzystuje się wiele kryteriów wyboru bazujących na przebiegu prognoz *ex post* (Armstrong, 2001). Alternatywnym podejściem w stosunku do szukania najlepszej prognozy indywidualnej jest wyznaczenie prognoz kombinowanych. Już od czasu pionierskiej pracy Batesa i Grangera (1969) wiadomo, że prognozy kombinowane mogą mieć mniejszy średni błąd prognozy (MSFE, *mean square forecast error*) niż każda z indywidualnych prognoz<sup>1</sup>. Stąd zamiast szukać jednego najlepszego modelu prognostycznego wyznacza się prognozy kombinowane spośród wszystkich dostępnych prognoz lub pewnej ich grupy.

Uzasadnieniem dla prognozowania kombinowanego (czy uśredniania modeli, ang. *model averaging*) jest fakt, że nie jest znany model generujący dane

---

<sup>1</sup> Praca Batesa i Grangera (1969) spowodowała rozwój badań dotyczących prognozowania kombinowanego (por. przeglądową pracę Timmermana, 2006).

(prawdziwy model), a zatem każdy model stanowi aproksymację nieznanego mechanizmu generującego. Modele te mogą być niekompletne (nie do specyfikowane) pod różnymi względami, prognozy wyznaczane na ich podstawie mogą być obciążone, a jeśli nawet prognozy są nieobciążone, to kowariancje między nimi mogą być różne od zera. W takiej sytuacji wyznaczenie prognoz kombinowanych na podstawie nie do specyfikowanych modeli (czy nawet błędnie wyspecyfikowanych) może polepszyć prognozy w stosunku do prognoz indywidualnych otrzymanych z poszczególnych modeli.

W związku z tym pojawia się pytanie dotyczące schematu wyboru indywidualnych prognoz spośród dostępnych prognoz oraz systemu wag wykorzystanego przy wyznaczaniu prognoz kombinowanych. Szczególne znaczenie ma wybór systemu wag. Najczęściej wyróżnia się następujące schematy wag: wagi jednakowe (Stock, Watson, 2004, 2006; Marcellino, 2004), wagi Akaike'a (Atkinson, 1980; Swanson, Zeng, 2001; Kapetanios i inni, 2008), wagi  $R^2$  (Kichen, Monaco, 2003), kowariancyjne podejście i zoptymalizowane wagi ograniczone (Jagannathan, Ma, 2003), wagi bayesowskie (Min, Zellner, 1993; Diebold, Pauly, 1980; Wright, 2003).

W artykule uwaga jest skupiona na podejściu wykorzystującym kryteria informacyjne, a w szczególności kryterium Akaike'a, które posłuży do wyznaczenia wag Akaike'a. Podejście to umożliwia otrzymanie nie jednego, a kilku wiarygodnych modeli, dla których można stworzyć ranking stosując wagi Akaike'a. Modele te stanowią podstawę obliczenia prognoz indywidualnych, a te z kolei służą do wyznaczenia ostatecznej prognozy kombinowanej, przy formułowaniu której wykorzystuje się wagi Akaike'a. Artykuł stanowi popularyzację nieznanych wcześniej w polskiej literaturze wag Akaike'a oraz ilorazu szans do zagadnienia wyboru modeli spośród określonego zestawu modeli, a także wyznaczania prognoz kombinowanych.

Celem artykułu jest zaproponowanie procedury wyznaczania prognoz kombinowanych według wag Akaike'a, a także porównanie prognoz kombinowanych otrzymanych według wag Akaike'a i wag jednakowych w odniesieniu do prognoz indywidualnych wyznaczonych z jednego, „najlepszego” modelu wybranego według kryterium Akaike'a oraz według tradycyjnego podejścia opartego na testowaniu hipotez statystycznych. Analiza zostanie przeprowadzona w ramach eksperymentu symulacyjnego, w którym jako modele aproksymacyjne przyjmuje się modele autoregresyjne oraz przyczynowo-skutkowe przy założeniu, że nie jest znany model generujący.

Struktura artykułu jest następująca. W podrozdziale drugim zostanie przedstawiona konstrukcja wag Akaike'a, w trzecim – zostanie zaproponowana procedura wyznaczania prognoz kombinowanych z wykorzystaniem wag Akaike'a. Następnie zostaną przedstawione wyniki dwóch eksperymentów numerycznych. W ostatnim podrozdziale zawarte będą wnioski.

## 2. WAGI AKAIKE'A

Kryterium informacyjne Akaike'a (AIC) służy do wyboru najlepszego modelu spośród zbioru modeli-kandydatów, przy czym celem wyboru modelu według AIC jest oszacowanie straty informacji, gdy rozkład prawdopodobieństwa  $f$  związany z prawdziwym (generującym) modelem jest aproksymowany rozkładem prawdopodobieństwa  $g$  związanym z modelem, który ma być oszacowany. Miarą rozbieżności między prawdziwym i aproksymacyjnym modelem jest informacja (odległość) Kullbacka-Leiblera (K-L). Akaike (1973) pokazał, że wybór modelu z najniższą oczekiwaną stratą informacji (tj. modelu, który minimalizuje oczekiwaną miarę K-L) jest asymptotycznie równoważne wyborowi modelu  $M_i$  ( $i = 1, 2, \dots, R$ ), który ma najniższą wartość kryterium AIC.

Akaike zdefiniował kryterium informacyjne (AIC) następująco<sup>2</sup>:

$$AIC = -2 \ln L(\hat{\theta}) + 2K, \quad (1)$$

gdzie  $L(\hat{\theta})$  – funkcja wiarygodności dla modelu-kandydata, która jest korygowana o liczbę szacowanych parametrów  $K$ . Pierwszy składnik kryterium AIC zmniejsza się w miarę jak więcej parametrów zawiera model, natomiast drugi składnik ( $2K$ ) staje się większy w miarę jak rośnie liczba parametrów modelu.

Przy wyznaczaniu wag Akaike'a korzysta się nie z bezwzględnych wartości AIC otrzymanych dla poszczególnych modeli-kandydatów, a z różnic kryteriów<sup>3</sup> AIC dla modeli-kandydatów w stosunku do wartości AIC dla najlepszego modelu ( $AIC_{\min}$ ), tj.

$$\Delta_i = AIC_i - AIC_{\min}. \quad (2)$$

Wartości  $\Delta_i$  są łatwe do interpretacji i pozwalają na szybkie porównanie i stworzenie rankingu modeli-kandydatów. Najlepszy model spośród całego zestawu modeli ma  $\Delta_i \equiv \Delta_{\min} = 0$ . Im większa jest różnica  $\Delta_i$ , tym mniej prawdopodobne jest, że dany model jest dobrym modelem w sensie kryterium K-L przy założeniu określonego zbioru danych. Dla modeli zagnieżdżonych podaje się następującą praktyczną zasadę wyboru modelu (Burnham, Anderson, 2002): dla  $\Delta_i < 2$  poparcie dla  $i$ -tego modelu jest znaczne, dla  $4 < \Delta_i < 7$  – znacznie mniejsze, wartości  $\Delta_i > 10$  oznaczają brak poparcia i modele takie powinny być pominięte, ponieważ nie wyjaśniają znacznej zmienności badanego zjawiska. W praktyce akceptuje się modele, dla których  $\Delta_i < 4$ .

<sup>2</sup> Kryterium AIC jest równoważne estymatorowi oczekiwanej informacji K-L, tj.  $\ln L(\hat{\theta}) - K$ , przemnożonemu przez  $-2$ , przy czym  $K$  stanowi obciążenie funkcji wiarygodności, które na ogół interpretowane jest w kategoriach kary (por. Akaike, 1973; Burnham, Anderson, 2002).

<sup>3</sup> Por. Akaike, 1978; Burnham, Anderson, 2002.

Na podstawie różnic  $\Delta_i$  można uzyskać ocenę relatywnej wiarygodności  $i$ -tego modelu (jego relatywnej przewagi w zbiorze modeli-kandydatów) obliczając wiarygodność modelu przy założeniu zbioru danych, tj.  $L(g_i | x)$ , którą Akaike (1983) określił następująco:

$$L(g_i | x) \propto \exp(-0.5\Delta_i), \quad (3)$$

gdzie  $\propto$  oznacza „jest proporcjonalny”.

Po normalizacji relatywnych wiarygodności modelu otrzymuje się wagi Akaike'a (Burnham, Anderson, 2002)

$$w_i = \frac{\exp(-0.5\Delta_i)}{\sum_{r=1}^R \exp(-0.5\Delta_r)}, \quad \sum_{i=1}^R w_i = 1. \quad (4)$$

Dana waga  $w_i$  jest interpretowana jako prawdopodobieństwo, że model  $M_i$  jest najlepszym modelem (w sensie *AIC*, tj. modelem, który minimalizuje informację K-L) spośród  $R$  modeli-kandydatów przy założeniu danego zbioru danych.

Dodatkowo wagi  $w_i$  są przydatne do oceny relatywnej przewagi modelu najlepszego (o największej wadze  $w_i$ ) nad pozostałymi modelami ze zbioru  $R$ . W tym celu oblicza się ilorazy szans (evidence ratios)  $w_i / w_j$ , a szczególnie stosunki  $w_1 / w_j$ , gdzie  $w_1$  jest wagą dla najlepszego modelu, a  $w_j$  – reprezentuje wagę  $j$ -tego modelu ze zbioru  $R$ . Stosunki wag informują o skali przewagi najlepszego w sensie K-L modelu nad pozostałymi modelami. Warto zwrócić uwagę, że w podejściu tym prawdziwy model generujący nie musi znajdować się w zbiorze  $R$  modeli. Rozważa się tylko ranking modeli w kontekście informacji (odległości) K-L<sup>4</sup>.

### 3. PROCEDURA WYNYCZANIA PROGNOZ KOMBINOWANYCH Z WAGAMI AKAIKE'A

Przystępując do wyznaczenia prognoz kombinowanych na bazie wag Akaike'a powinny być spełnione pewne warunki<sup>5</sup>. Mianowicie, wszystkie modele w zbiorze modeli-kandydatów muszą być oszacowane na podstawie tej samej liczby obserwacji, ponieważ wnioskowanie na bazie kryteriów informacyjnych jest przeprowadzane warunkowo względem danego zbioru danych. Ponadto, wszystkie modele w zbiorze  $R$  modeli muszą mieć tę samą zmienną objaśnianą. Powszechnym błędem jest porównywanie modeli dla zmiennej  $y_i$  z modelami dla zmiennej po transformacji np.  $\ln y_i$  czy  $\Delta y_i$ .

<sup>4</sup> Jest to zasadnicza różnica w stosunku do bayesowskiego podejścia do uśredniania modeli, w którym zakłada się że prawdziwy model znajduje się w zbiorze  $R$  modeli lub jest nim model otrzymany w wyniku uśredniania modeli.

<sup>5</sup> Warunki te dotyczą też sytuacji wyboru modelu za pomocą kryteriów informacyjnych.

Procedurę wyznaczania prognoz kombinowanych z wagami Akaike'a przedstawiają następujące etapy.

**Etap 1.** Ustalenie wyjściowego zestawu  $R$  modeli opisujących dane zjawisko i ich specyfikacji. Należy podkreślić, że w przypadku modeli przyczynowo-skutkowych przy określaniu specyfikacji należy kierować się informacjami z teorii ekonomicznej związanej z danym zjawiskiem. W przypadku dużej liczby zmiennych nie zaleca się postępowania zmierzającego do stworzenia wszystkich możliwych regresji. Zestaw modeli powinien być wiarygodny pod względem ekonomicznym, a nie pochodzić z automatycznego wyboru. Podejście to nie wymaga specyfikacji prawdziwego modelu.

**Etap 2.** Szacowanie indywidualnych modeli ( $i=1,2,\dots,R$ ), wyznaczenie kryterium informacyjnego AIC oraz wyznaczenie różnic  $\Delta_i$ . Modele powinny być poprawne pod względem statystycznym i ekonomicznym.

**Etap 3.** Stworzenie zredukowanego zestawu modeli ( $i=1,2,\dots,R^*$ ), dla którego  $\Delta_i < 4$ , czyli zestawu modeli, dla których jest duże prawdopodobieństwo, że są to modele dobre w sensie kryterium K-L.

**Etap 4.** Wyznaczenie wag Akaike'a i prognoz kombinowanych.

Przy wyznaczaniu prognoz kombinowanych stosuje się następującą formułę:

$$\hat{y}_{t,t+h} = \sum_{i=1}^{R^*} w_{i,h} f_{i,t,t+h}, \quad \sum_{i=1}^{R^*} w_{i,h} = 1,$$

gdzie  $\hat{y}_{t,t+h}$  – prognoza kombinowana,  $w_{i,h}$  – waga przypisana prognozie  $f_{i,t,t+h}$  bazującej na  $i$ -tym indywidualnym modelu.

Przy wyznaczaniu wag chodzi o takie wagi, które będą minimalizować funkcję straty zależną od błędów prognozy. Na ogół jako funkcję straty przyjmuje się średni błąd prognoz ex post (*MSFE*, *mean square forecast error*).

#### 4. WYNIKI EKSPERYMENTU SYMULACYJNEGO

Celem eksperymentu numerycznego jest porównanie prognoz kombinowanych otrzymanych według wag Akaike'a i wag jednakowych w odniesieniu do prognoz indywidualnych wyznaczonych z jednego, „najlepszego” modelu wybranego według kryterium Akaike'a oraz według tradycyjnego podejścia opartego na testowaniu hipotez statystycznych. Analiza zostanie przeprowadzona w ramach eksperymentu symulacyjnego, w którym jako modele aproksymacyjne przyjmuje się modele autoregresyjne oraz przyczynowo-skutkowe przy założeniu, że nie jest znany model generujący (prawdziwy).

Scenariusz eksperymentu jest następujący:

Model generujący  $Y_t$  miał postać:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon), \quad \sigma_\varepsilon = 1, 2, 3,$$

z parametrami:  $\beta_0 = 10$ ,  $\beta_1 = 1.5$ ,  $\beta_2 = 1.2$ ,  $\beta_3 = 2$ , dla liczby obserwacji:  $n = 100, 50$  (liczba replikacji  $m = 1000$ ). Procesy  $X_{1,t}$ ,  $X_{2,t}$  i  $X_{3,t}$  miały następującą strukturę:

$$X_{1,t} = 12 + 0.8X_{1,t-1} + \zeta_t + 0.6\zeta_{t-1}, \quad \zeta_t \sim N(0, 1),$$

$$X_{2,t} = 14 + 0.7X_{2,t-1} + \eta_t + 0.8\eta_{t-1}, \quad \eta_t \sim N(0, 1),$$

$$X_{3,t} = 8 + 1.2X_{3,t-1} - 0.4X_{3,t-2} + \xi_t, \quad \xi_t \sim N(0, 1).$$

Jako modele aproksymacyjne przyjęto modele autoregresyjne

$$Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \dots + \gamma_q Y_{t-q}, \quad q = 1, 2, 3, 4,$$

oraz modele przyczynowo-skutkowe, przy czym założono, że nie jest znany model prawdziwy (generujący), stąd w specyfikacji modeli przyczynowo-skutkowych pominięto zmienną  $X_{2,t}$  i  $X_{3,t}$ . Modele te miały postać<sup>6</sup>:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \alpha_4 Y_{t-4} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \quad (\text{M1})$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \quad (\text{M2})$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \quad (\text{M3})$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \quad (\text{M4})$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \gamma_1 X_{1,t} + v_t, \quad (\text{M5})$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 X_{1,t} + v_t, \quad (\text{M6})$$

Analiza była prowadzona oddzielnie dla modeli autoregresyjnych i modeli przyczynowo-skutkowych. W każdej replikacji do dalszej analizy były brane modele, dla których  $\Delta_i < 4$ . Dla tych modeli wyznaczano wagi Akaike'a i wagi jednakowe ( $1/R^*$ ), a następnie po obliczeniu prognoz indywidualnych<sup>7</sup> wyznaczano prognozy kombinowane na  $h$ -okresów naprzód ( $h = 1, 2, \dots, 10$ ). Prognozy porównywano za pomocą średniego błędu prognoz ( $MSFE$ ). Błędy te wyznaczono również dla prognoz indywidualnych otrzymanych z modeli wybranych jako najlepsze (według kryterium AIC i według metody selekcji a posteriori stosowanej do wersji M1 modelu przyczynowo-skutkowego<sup>8</sup>) w celu porównania prognoz kombinowanych i prognoz indywidualnych. Wyniki prezentuje tabela 1 (dla  $n = 100$ ) oraz tabela 2 (dla  $n = 50$ ).

<sup>6</sup> Po analizie na podstawie wstępnych symulacji modele te zostały uznane za wiarygodne.

<sup>7</sup> Prognozy indywidualne były prognozami dynamicznymi, przy czym za wartości zmiennej  $X_{1,t}$  w okresie prognozowanym zostały wzięte wartości generowane.

<sup>8</sup> Eliminacja w metodzie selekcji a posteriori przebiegała przy założeniu poziomu istotności  $\alpha = 0,05$ .

Tabela 1. Średnie błędy prognoz ( $MSFE$ ) dla liczby obserwacji  $n = 100$ 

h	Modele przyczynowo-skutkowe				Modele autoregresyjne		
	WAIC	WJEDN.	minAIC	mSEL	WAIC	WJEDN.	minAIC
$\sigma_\varepsilon = 1$							
1	2,950	2,973	2,935	3,050	3,102	3,117	3,096
2	2,290	2,301	2,287	2,357	2,530	2,547	2,523
3	2,471	2,500	2,456	2,604	2,568	2,595	2,548
4	3,280	3,320	3,254	3,436	2,569	2,602	2,537
5	4,064	4,097	4,043	4,192	3,706	3,748	3,653
6	3,949	3,985	3,926	4,092	3,413	3,452	3,367
7	3,807	3,842	3,785	3,944	3,192	3,225	3,156
8	3,589	3,622	3,568	3,717	3,038	3,073	3,001
9	3,503	3,528	3,488	3,607	2,931	2,960	2,906
10	3,903	3,915	3,900	3,961	3,266	3,280	3,265
$\sigma_\varepsilon = 2$							
1	1,958	1,944	2,008	2,035	2,220	2,217	2,229
2	2,200	2,175	2,280	2,207	3,111	3,105	3,121
3	3,236	3,221	3,284	3,266	2,890	2,885	2,899
4	4,795	4,794	4,823	4,901	3,200	3,198	3,204
5	6,334	6,342	6,350	6,497	4,696	4,698	4,697
6	6,333	6,341	6,347	6,484	4,836	4,838	4,835
7	5,971	5,978	5,984	6,107	4,596	4,598	4,595
8	5,675	5,681	5,689	5,792	4,359	4,361	4,358
9	5,397	5,403	5,410	5,508	4,172	4,173	4,171
10	5,224	5,229	5,237	5,324	4,087	4,088	4,086
$\sigma_\varepsilon = 3$							
1	3,172	3,177	3,168	3,222	4,430	4,415	4,459
2	3,456	3,455	3,459	3,477	5,061	5,048	5,087
3	5,283	5,251	5,321	5,211	6,934	6,921	6,961
4	7,320	7,320	7,331	7,361	9,907	9,897	9,933
5	8,008	8,014	8,010	8,066	11,199	11,189	11,225
6	7,621	7,624	7,625	7,670	10,966	10,957	10,992
7	7,171	7,172	7,178	7,215	10,380	10,371	10,405
8	6,794	6,796	6,802	6,836	9,846	9,837	9,870
9	6,514	6,518	6,520	6,563	9,339	9,331	9,362
10	6,260	6,262	6,267	6,305	8,993	8,985	9,015

W kolumnach  $w_{AIC}$ ,  $w_{JEDN.}$  znajdują się średnie błędy prognoz kombinowanych według odpowiednio wag AIC i wag jednakowych, a w kolumnach  $min_{AIC}$ ,  $m_{SEL}$  – prognozy otrzymane z modelu wybranego według minimum AIC oraz modelu otrzymanego w wyniku stosowania metody selekcji a posteriori w odniesieniu do modelu M1, przy czym selekcji dokonywano przy poziomie istotności  $\alpha = 0,05$ .

Źródło: obliczenia własne.

Tabela 2. Średnie błędy prognoz (*MSFE*) dla liczby obserwacji  $n = 50$ 

h	Modele przyczynowo-skutkowe				Modele autoregresyjne		
	WAIC	WJEDN.	minAIC	mSEL	WAIC	WJEDN.	minAIC
$\sigma_{\varepsilon} = 1$							
1	1,019	1,018	1,018	1,023	1,069	1,069	1,076
2	1,502	1,587	1,409	1,888	2,712	2,705	2,721
3	1,564	1,576	1,570	2,002	2,804	2,796	2,815
4	2,440	2,393	2,493	2,494	2,489	2,482	2,498
5	3,599	3,486	3,722	3,251	2,288	2,283	2,296
6	3,685	3,609	3,761	3,274	2,136	2,132	2,144
7	3,455	3,391	3,518	3,144	2,219	2,212	2,229
8	3,279	3,212	3,352	3,093	2,443	2,435	2,456
9	3,164	3,110	3,224	2,988	2,342	2,334	2,355
10	3,669	3,642	3,699	3,412	2,569	2,565	2,576
$\sigma_{\varepsilon} = 2$							
1	3,522	3,508	3,539	4,230	5,419	5,399	5,414
2	3,854	3,803	3,918	4,644	5,739	5,711	5,744
3	4,199	4,100	4,318	4,900	5,669	5,641	5,677
4	4,832	4,671	5,018	5,261	5,483	5,458	5,489
5	5,379	5,222	5,555	5,668	5,679	5,660	5,685
6	5,493	5,324	5,683	5,616	5,356	5,340	5,362
7	6,133	5,955	6,329	5,949	5,273	5,262	5,280
8	6,171	6,016	6,346	5,890	5,139	5,130	5,146
9	6,311	6,195	6,456	6,048	5,362	5,357	5,369
10	6,286	6,189	6,411	6,034	5,380	5,376	5,387
$\sigma_{\varepsilon} = 3$							
1	5,088	5,096	5,085	5,335	7,026	6,997	7,064
2	6,378	6,371	6,389	6,498	8,233	8,203	8,276
3	6,968	6,952	6,987	6,993	8,696	8,683	8,720
4	6,578	6,562	6,597	6,599	8,389	8,386	8,399
5	6,193	6,175	6,213	6,205	7,944	7,945	7,948
6	5,803	5,788	5,823	5,821	7,391	7,393	7,396
7	5,599	5,579	5,627	5,593	6,938	6,939	6,942
8	5,504	5,484	5,532	5,492	6,593	6,594	6,597
9	5,325	5,308	5,351	5,334	6,456	6,458	6,459
10	5,181	5,166	5,206	5,228	6,601	6,603	6,603

Objaśnienia jak w tabeli 1.

Źródło: obliczenia własne.

Wyniki przedstawione w tabeli 1 i 2 pokazują, że różnice między średnimi błędami prognozy (*MSFE*, *mean square forecast error*) otrzymanymi dla prognoz kombinowanych (według wag *AIC* i wag jednakowych) oraz prognoz indywidualnych (z najlepszego modelu wybranego według minimum kryterium *AIC* oraz według metody selekcji *a posteriori*) są nieznaczne. Mogłoby to wskazywać, że prognozy kombinowane nie mają przewagi w sensie niższego błędu *MSFE* nad prognozami indywidualnymi, jednak w ramach otrzymanych wyników można zauważyć pewne prawidłowości świadczące o przydatności prognoz kombinowanych.

W przypadku  $\sigma_\varepsilon = 1$  prognozy kombinowane według wag *AIC* ( $w_{AIC}$ ) na podstawie modeli przyczynowo-skutkowych mają nieznacznie mniejsze błędy *MSFE* niż prognozy kombinowane według wag jednakowych ( $w_{JEDN.}$ ) w całym horyzoncie prognozowania (por. tabela 1). Ta nieznaczna przewaga jest zachowana dla  $\sigma_\varepsilon = 2$  i 3 dla dłuższego horyzontu prognozy ( $h > 5$ ), natomiast dla krótszego horyzontu prognozy niższe *MSFE* otrzymuje się dla prognoz kombinowanych według wag jednakowych.

Można zauważyć, że błędy *MSFE* dla prognoz kombinowanych według wag *AIC* i jednakowych są już wyraźnie mniejsze niż *MSFE* dla prognoz indywidualnych otrzymanych po zastosowaniu metody selekcji *a posteriori* ( $m_{SEL.}$ ); dotyczy to wszystkich poziomów zakłócenia  $\sigma_\varepsilon$  (por. tabela 1). Wskazywałoby to na przewagę prognoz kombinowanych nad prognozami indywidualnymi. Najniższe błędy *MSFE* otrzymuje się dla prognoz z modelu wybranego według minimum kryterium *AIC* (dla  $\sigma_\varepsilon = 1$ ), natomiast dla  $\sigma_\varepsilon = 2$  i 3 niższe błędy *MSFE* mają prognozy kombinowane według wag jednakowych dla horyzontu  $h \leq 5$ , a dla  $h \geq 6$  – prognozy kombinowane według wag *AIC*. Te niższe błędy *MSFE* dla prognoz z modeli wybranych według minimum *AIC* dla zakłócenia  $\sigma_\varepsilon = 1$  odpowiadały sytuacji, w której zbiór modeli-kandydatów był niewielki, a dodatkowo jeden z modeli miał wysoką wagę *AIC* ( $w_{AIC} > 0,7$ ). W tym przypadku najczęściej zbiór modeli wykorzystywanych przy wyznaczaniu prognoz kombinowanych składał się z modelu M3 i M4, w których jeden miał wysoką wagę *AIC*. W przypadku natomiast większego zakłócenia  $\sigma_\varepsilon = 2$  i 3 zbiór modeli brany do obliczenia prognoz kombinowanych składał się najczęściej z modeli M3, M4, M5, M6, z których żaden nie miał wysokiej wagi *AIC*. Wtedy mniejsze błędy *MSFE* miały prognozy kombinowane (według wag jednakowych czy wag *AIC*), a nie prognozy wyznaczone z modelu wybranego według minimum kryterium *AIC*.

W przypadku modeli autoregresyjnych najniższe błędy *MSFE* otrzymano dla prognoz obliczonych na podstawie modeli wybranych według minimum kryterium *AIC*, a błędy *MSFE* dla prognoz kombinowanych według wag *AIC* są niższe niż według wag jednakowych (dla  $\sigma_\varepsilon = 1$ ) – por. tabela 1. Takie wyniki związane były z sytuacją, w której w większości przypadków do zbioru modeli wchodziły modele AR(2), AR(3) i AR(4), przy czym model AR(2) dominował

na pozostałymi, bo miał wyższą wagę  $AIC$  (w granicach 0,5–0,6), stąd lepsze wyniki w sensie błędów  $MSFE$  dawały prognozy z modelu pojedynczego, wybranego według minimum kryterium  $AIC$ . Dla większego zakłócenia, tj.  $\sigma_\varepsilon = 2$  (czyli większej niepewności) i krótkiego horyzontu prognozowania ( $h \leq 5$ ) przewagę mają prognozy kombinowane według wag jednakowych, a dla horyzontu  $h \geq 6$  niższe błędy otrzymuje się dla prognoz kombinowanych według wag  $AIC$ . Natomiast prognozy wyznaczone z modeli wybranych według minimum kryterium  $AIC$  mają błędy  $MSFE$  wyższe niż błędy dla prognoz kombinowanych. Zaznaczyć jednak należy, że różnice te nie są duże. Dla zakłócenia  $\sigma_\varepsilon = 3$  przewagę w całym horyzoncie prognozowania mają prognozy kombinowane wyznaczone według wag jednakowych. Błędy  $MSFE$  dla prognoz kombinowanych według wag  $AIC$  ( $W_{AIC}$ ) są nieznacznie większe niż według wag jednakowych ( $W_{JEDN.}$ ) i nieznacznie mniejsze niż błędy prognoz otrzymane dla modeli wybranych według minimum kryterium  $AIC$  ( $\min_{AIC}$ ) – por. tabela 1. Ta relatywna przewaga prognoz kombinowanych nad prognozami indywidualnymi związana była z sytuacją, w której do zbioru modeli wykorzystywanych do obliczenia prognoz kombinowanych najczęściej wchodziły wszystkie modele AR i żaden z nich nie miał przewagi wyraźnej nad pozostałymi (w sensie wag  $AIC$ ). Wtedy niższe błędy  $MSFE$  otrzymuje się dla prognoz kombinowanych.

Wyniki zawarte w tabeli 2, dla liczby obserwacji  $n = 50$ , pokazują, że błędy  $MSFE$  na podstawie modeli przyczynowo-skutkowych są niższe dla prognoz kombinowanych według wag jednakowych niż według wag  $AIC$  – dotyczy to wszystkich poziomów zakłócenia  $\sigma_\varepsilon$  (z wyjątkiem  $\sigma_\varepsilon = 1$  i  $h = 2, 3$ , gdy niższe są  $MSFE$  dla prognoz kombinowanych według wag  $AIC$ ). Przewaga prognoz kombinowanych czy prognoz indywidualnych ( $\min_{AIC}$ ,  $m_{SEL.}$ ) zależy od długości horyzontu prognozowania i wielkości zakłócenia  $\sigma_\varepsilon$ . Dla krótszego horyzontu  $h \leq 4$  i zakłócenia  $\sigma_\varepsilon = 1$  (dla  $\sigma_\varepsilon = 2$  i  $h \geq 7$ ) błędy  $MSFE$  dla prognoz kombinowanych ( $W_{AIC}$ ,  $W_{JEDN.}$ ) są niższe niż dla prognoz z modelu wybranego według metody *a posteriori* ( $m_{SEL.}$ ), a dla  $h \geq 5$  błędy  $MSFE$  są niższe dla prognoz z  $m_{SEL.}$ , jednak przewaga  $m_{SEL.}$  nie występuje dla  $\sigma_\varepsilon = 3$  (por. tabela 2). Prognozy z modelu wybranego według kryterium minimum  $AIC$  ( $\min_{AIC}$ ) mają większe błędy  $MSFE$  niż prognozy kombinowane, a także są na ogół wyższe niż błędy  $MSFE$  dla prognoz z  $m_{SEL.}$

W przypadku prognoz obliczanych na podstawie modeli autoregresyjnych błędy  $MSFE$  dla prognoz kombinowanych są zawsze niższe niż dla prognoz z modeli wybranych według  $\min_{AIC}$  (por. tabela 2). Natomiast prognozy kombinowane według wag jednakowych ( $W_{JEDN.}$ ) mają przewagę nad prognozami otrzymanymi według wag  $AIC$  ( $W_{AIC}$ ).

## 5. PODSUMOWANIE

Z porównania prognoz kombinowanych według wag  $AIC$  i wag jednakowych ( $W_{AIC}$ ,  $W_{JEDN.}$ ) wynika, że prognoz kombinowanych według  $W_{AIC}$  nie opłaca

się wyznaczać, jeżeli waga  $AIC$  jednego z modeli wchodzących do zbioru modeli-kandydatów jest wyraźnie większa, tj.  $w_{AIC} > 0,7$ , niż dla pozostałych modeli, co oznacza, że relatywna przewaga danego modelu jako modelu najlepszego w sensie kryterium  $AIC$  spośród zbioru modeli-kandydatów jest duża. Powoduje to, że błędy prognoz  $MSFE$  będą większe niż przy założeniu prostych wag jednakowych. W takiej sytuacji wagi  $AIC$  mogą służyć do stworzenia rankingu modeli, a także obliczenia ilorazów szans informujących o skali przewagi najlepszego w sensie  $AIC$  modelu (tj. modelu, który minimalizuje informację K–L) nad pozostałymi modelami. Korzyści ze stosowania wag  $AIC$  pojawiają się, gdy w zbiorze modeli-kandydatów jest większa liczba modeli i żaden nie ma przewagi na pozostałymi, tj. brak modelu z dominującą wagą  $w_{AIC}$ .

Warto podkreślić, że wyniki otrzymane w eksperymencie wskazały na przewagę prognoz kombinowanych nad prognozami indywidualnymi wyznaczanymi na podstawie modeli autoregresyjnych. Przewaga ta dla prognoz z modeli przyczynowo-skutkowych wybranych według metody a posteriori ( $m_{SEL}$ ) i minimum kryterium  $AIC$  ( $\min_{AIC}$ ) występuje szczególnie dla dużego zakłócenia  $\sigma_\varepsilon$ , które wiąże się z większą niepewnością co do wyboru modelu.

Wykorzystanie zatem podejścia bazującego na kryteriach informacyjnych, w tym szczególnie zastosowanie wag  $AIC$  do tworzenia rankingu modeli i wyznaczania prognoz kombinowanych można traktować jako alternatywne w stosunku do podejścia bazującego na testowaniu hipotez statystycznych i zmierzającego do wyboru jednego, najlepszego modelu i na jego podstawie wyznaczania indywidualnej prognozy.

## LITERATURA

- Akaike H. (1973), *Information Theory as an Extension of the Maximum Likelihood Principle*, [w:] Petrov B. N., Csaki F., *Second International Symposium on Information Theory, Akademia Kiado, Budapest*.
- Akaike H. (1978), *On the Likelihood of a Time Series Model*, „The Statistician”, 27, 217–235.
- Armstrong J. S. (2001), *Principles of Forecasting*, Springer.
- Atkinson A. C. (1980), *A Note on the Generalized Information Criteria for Choice of a Model*, „Biometrika”, 67 (2), 413–418.
- Bates J. M., Granger C. W. J. (1969), *The Combinations of Forecasts*, „Operations Research Quarterly”, 20, 415–468.
- Burnham K. P., Anderson D. R. (2002), *Model Selection and Multimodel Inference*, Springer.
- Jagannathan R. Ma T. (2003), *Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps*, „The Journal of Finance”, 58 (4), 1651–1684.
- Kapetanios G., Labhard V., Price S. (2008), *Forecasting using Bayesian and Information-theoretic Model Averaging: an Application to U.K. Inflation*, „Journal of Business and Economics Statistics”, 26 (1), 33–41.
- Kitchen J., Monaco R. (2003), *Real-Time Forecasting in Practice*, „Business Economics”, 38 (4), 10–19.
- Marcellino M. (2004), *Forecast Pooling for Short Time Series of Macroeconomic Variables*, „Oxford Bulletin of Economic and Statistics”, 66, 91–112.

- Min C. K., Zellner A. (1993), *Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates*, „Journal of Econometrics”, 53 (1–2), 89–118.
- Stock J. H., Watson M. (2004), *Combination Forecasts of Output Growth in a Seven-Country Data Set*, „Journal of Forecasting”, 8, 230–251.
- Stock J. H., Watson M. (2006), *Forecasting with Many Predictors*, [w:] Elliott G., Granger C. W. J., Timmermann A. (ed.), *Handbook of Economic Forecasting*, Elsevier.
- Swanson N. R., Zeng T. (2001), *Choosing Among Competing Econometric Forecasts: Regression-based Forecast Combination using Model Selection*, „Journal of Forecasting”, 20, 425–440.
- Timmermann A. (2006), *Forecast Combinations*, [w:] Elliott G., Granger C. W. J., Timmermann A. (ed.), *Handbook of Economic Forecasting*, Elsevier.

### COMBINING FORECASTS USING THE AKAIKE WEIGHTS

**A b s t r a c t.** The focus in the paper is on the information criteria approach and especially the Akaike information criterion which is used to obtain the Akaike weights. This approach enables to receive not one best model, but several plausible models for which the ranking can be built using the Akaike weights. This set of candidate models is the basis of calculating individual forecasts, and then for combining forecasts using the Akaike weights. The procedure of obtaining the combined forecasts using the AIC weights is proposed. The performance of combining forecasts with the AIC weights and equal weights with regard to individual forecasts obtained from models selected by the AIC criterion and the *a posteriori* selection method is compared in simulation experiment. The conditions when the Akaike weights are worth to use in combining forecasts were indicated. The use of the information criteria approach to obtain combined forecasts as an alternative to formal hypothesis testing was recommended.

**K e y w o r d s:** combining forecasts, weighting schemes, information criteria.