

COMBINAREA PROGNOZELOR ÎN FUNCȚIE DE EFICIENȚA LOR PE TERMEN MEDIU ȘI LUNG PRIN UTILIZAREA FACTORILOR DE PONDERARE. ANALIZĂ COMPARATIVĂ

Andrei Dospinescu *

Rezumat

Lucrarea de față dorește să vadă în ce situații utilizarea combinată a metodelor de prognoză este mai eficientă decât utilizarea lor separată.

Dorim în același timp să analizăm în ce context metoda pe care o propunem: combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung prin utilizarea factorilor de ponderare este mai eficientă decât combinarea prognozelor prin regresie și combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză Wenzel (2001). Propunem o metodă de calcul a coeficientului γ din modelul lui Wenzel și utilizăm modelul pentru a identifica diferența de eficiență pe termen mediu și lung a prognozelor combinate.

Cuvinte cheie: combinarea prognozelor, eficiența pe termen mediu și lung, media și abaterea standard.

Clasificarea JEL: C13, C49, C51, C52, C53

1. Introducere

Combinarea metodelor de prognoză are la bază ipoteza conform căreia prognoza rezultată în urma combinării va genera performanțe mai bune (de exemplu erori mai mici) decât fiecare metodă de prognoză în parte. Astfel, într-un sistem format din mai multe metode de prognoză, fiecare prognoză în parte va putea oferi informații în plus care nu erau oferite de celelalte prognoze. Cheia este să găsim modul de combinare care să permită adăugarea acestor informații în plus în modelul final.

Lucrarea care stă la baza dezvoltării acestui tip de analiză este lucrarea lui Bates și Granger (1969), în care autorii au dezvoltat și combinat diferite tehnici pentru combinarea prognozelor. Studii ulterioare au adus contribuții în identificarea de metode de combinare a prognozelor. Newbold și Granger (1973) au testat combinarea metodelor diferite de prognoză pe baza mediilor ponderate. Calculul avea la bază coeficientul de corelație r al fiecărei metode de prognoză în parte. Bunn (1975, 1977) a propus ideea combinării prognozelor pe baza analizei bayesiene în care se

* Institutul de Prognoză Economică, Academia Română, București, e-mail: andreidospinescu@yahoo.com

analiza cât de probabil era ca o prognoză să fie mai eficientă decât alta. Crane și Coty (1967) au propus combinarea prognozelor prin regresie. Metoda a mai fost analizată, printre alții, de către Granger și Ramanathan (1984), Guerard și Beidleman (1987). Lippoletti și Webb (1986) au sugerat utilizarea unui model VAR pentru evaluarea prognozelor macroeconomice care compuneau o combinație de prognoze. Poulos, Kvaneli, Pavar (1987) au descris un sistem de prognoză automat care utiliza metoda Box-Jankins și Holt-Winters.

2. Obiective

Lucrarea are următoarele obiective:

- 1) Să identifice în ce situații utilizarea combinată a metodelor de prognoză este mai eficientă decât utilizarea lor separată.
- 2) Să analizeze în ce context metoda pe care o propunem: combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare este mai eficientă decât combinarea prognozelor prin regresie și combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză Wenzel (2001).
- 3) Să propună o metodă de calcul a coeficientului γ din modelul lui Wenzel și să utilizeze modelul pentru identificarea diferenței de eficiență pe termen mediu și lung a prognozelor combinate.

3. Metodologie

În analiză vom utiliza trei metode diferite de combinare:

- 1) Combinarea prognozelor prin regresie.
- 2) Combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză (Wenzel 2001)
- 3) Combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare.

Combinarea prognozelor prin regresie

Fie Y seria analizată, $P1$ și $P2$ prognoze ale seriei Y , $w1$ și $w2$ ponderi ale acestor prognoze în modelul final. Avem

$$Y = cI + w1*PI + w2*P2 + u$$

unde parametrii cI , $w1$ și $w2$ sunt estimați prin MCMMP (metoda celor mai mici pătrate).

Combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză (Wenzel 2001) (titlul original Hits-and-misses for the evaluation and combination of forecasts).

Metoda este propusă de Thomas Wenzel de la Universitatea Dortmund, Germania. Metoda are la bază identificarea frecvenței cu care o prognoză se încadrează în intervalul $[Y_{T+1} - \gamma|Y_{T+1}|, Y_{T+1} + \gamma|Y_{T+1}|]$.

Fie F_{T+1} o prognoză a lui Y_{T+1} la un moment $t+1$. F_{T+1} este „hitt”, $\gamma \in [0, \infty]$, dacă $F_{T+1} \in [Y_{T+1} - \gamma|Y_{T+1}|, Y_{T+1} + \gamma|Y_{T+1}|]$

Dacă restricționăm γ la intervalul $[0,1]$, atunci vom denumi prognoza într-un interval de $100(1-\gamma)$ procent de succes. Putem să folosim criteriul de încadrare a prognozei în interval pentru a combina prognozele.

Fie $F_{1,T+1}, \dots, F_{n,T+1}, n \geq 2$, prognoze ale lui Y_{T+1} și fie $F_{iT}, Y_T, i = 1, \dots, n, t = 1, \dots, T$ observații trecute. Vom defini ponderile după cum urmează:

$$w_{i,\gamma} = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma|Y_t|, Y_t + \gamma|Y_t|]}(F_{it})}{\sum_{j=1}^n \sum_{t=T-v+1}^T I_{[Y_t - \gamma|Y_t|, Y_t + \gamma|Y_t|]}(F_{jt})}, i = 1, \dots, n, \quad (1)$$

unde $\sum_{i=1}^n w_{i,\gamma} = 1$. După cum afirma Wenzel (2001, p.13) „For each weight we divide the number of hits of the special individual forecast by the number of hits of all forecasts. By definition

$$\sum_{i=1}^n w_{i,\gamma} = 1 ”.$$

După cum vom vedea când vom realiza o analiză comparativă între cele trei metode de combinare, metoda propusă de Wenzel (2001) nu reușește să identifice ponderi care să surprindă exact diferența de performanță între prognoze. ¹

Combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung prin utilizarea factorilor de ponderare

Este propusă ca un răspuns la metoda lui Wenzel și vrea să arate o anumită rigiditate a metodei, rigiditate care se observă în cazul combinării unor prognoze complementare în care una dintre prognoze este mai eficientă pe termen mediu și alta mai eficientă pe termen lung.

Fie Y seria analizată, P_1 și P_2 prognoze ale seriei Y , n numărul de perioade din segmentul martor și m perioada prognozată din segmentul martor. Avem

$$\hat{Y}_{t+m} = \frac{m}{n} \times P_{1,t+m} + \frac{n-m}{n} \times P_{2,t+m} \quad (2), \text{ unde } P_1 \text{ este prognoza mai eficientă pe}$$

termen lung, iar P_2 este prognoza mai eficientă pe termen mediu.

Prin eficiență înțelegem erori mai mici, pe care le măsurăm prin MAPE (mean absolute percentage error) și RMSE (root mean square error). (Armstrong și Collopy 1992).

Wenzel nu propune o metodă de calculare a lui γ , sugerează doar că valoarea lui ar trebui să se găsească în intervalul $[0, 1]$. γ poate fi fixat în funcție de gradul de eroare care se vrea a fi acceptat. Această valoare poate fi discutată cu directorul de marketing dacă avem de exemplu o prognoză a vânzărilor. Din aceasta perspectivă, în alegerea lui γ poate fi înglobată și cunoașterea expertului (domain knowledge), fapt care va duce la o îmbunătățire a prognozei și la o adaptare mai bună la situația analizată. Deși înglobarea cunoașterii expertului este utilă, ea nu este suficientă, alegerea lui γ trebuie să se facă și în funcție de metodele de prognoză care sunt utilizate în generarea prognozelor pe care le combinăm. În cazul în care γ este prea mare, atunci va fi dificil să fie identificată o pondere corectă a prognozelor în model, deoarece intervalul în care poate varia valoarea prognozată este prea mare (intervalul $[Y_{T+1} - \gamma|Y_{T+1}|, Y_{T+1} + \gamma|Y_{T+1}|]$ va fi prea

¹ Wenzel propune în articolul său și alte metode asemănătoare, ne-am oprit la aceasta pentru că am propus o modificare a metodei în ceea ce privește alegerea lui γ și am utilizat-o într-o anumită măsură în metoda pe care am propus-o.

mare). În această situație nu se va face o disociere clară între prognozele mai eficiente și cele mai puțin eficiente.

În cazul în care γ este prea mic atunci va fi dificil să fie identificată ponderea prognozelor în modelul final, deoarece intervalul în care poate varia valoarea prognozată este prea mic (intervalul $\left[Y_{T+1} - \gamma|Y_{T+1}|, Y_{T+1} + \gamma|Y_{T+1}| \right]$ va fi prea mic). În această situație prognoza mai eficientă (cea care va genera erori mai mici) nu „va intra” în intervalul ales de un număr suficient de ori, număr care să reflecte diferența de eficiență dintre ea și prognoza mai puțin eficientă.

Pentru a surprinde eficiența prognozelor utilizate am folosit media și abaterea standard din segmentul martor pentru a defini pe γ .

Fie Y seria prognozată, P1 și P2 prognozele seriei Y și S un segment martor de lungime n.

Definim mediile:
$$\bar{Y} = \frac{\sum_{t=1}^n Y_t}{n}, \bar{P}_1 = \frac{\sum_{t=1}^n P_{1,t}}{n}, \bar{P}_2 = \frac{\sum_{t=1}^n P_{2,t}}{n}$$

Definim abaterile standard:
$$\sigma_Y = \sqrt{\frac{\sum_{t=1}^n (Y_t - \bar{Y})^2}{n}}, \sigma_{P_1} = \sqrt{\frac{\sum_{t=1}^n (P_{1,t} - \bar{P}_1)^2}{n}}$$

$$\sigma_{P_2} = \sqrt{\frac{\sum_{t=1}^n (P_{2,t} - \bar{P}_2)^2}{n}}$$

Valoarea lui γ va fi:

$$\gamma \in \left[1 - \frac{\bar{P} - 2\sigma_P}{\bar{Y}}, 1 - \frac{\bar{P} + 2\sigma_P}{\bar{Y}} \right] \text{ dacă } \bar{P} \pm \sigma_P \leq \bar{Y}$$

$$\gamma \in \left[\frac{\bar{P} - 2\sigma_P}{\bar{Y}} - 1, \frac{\bar{P} + 2\sigma_P}{\bar{Y}} - 1 \right] \text{ dacă } \bar{P} \pm \sigma_P \geq \bar{Y}$$

$$\gamma \in \left[1 - \frac{\bar{P} - 2\sigma_P}{\bar{Y}}, \frac{\bar{P} + 2\sigma_P}{\bar{Y}} - 1 \right] \text{ dacă } \bar{P} + \sigma_P \geq \bar{Y}, \bar{P} - \sigma_P \leq \bar{Y}$$

unde $\bar{P} = \bar{P}_1$ dacă $\bar{P}_1 \in [\bar{Y} - 2\sigma_Y, \bar{Y} + 2\sigma_Y]$ și $\bar{P} = \bar{P}_2$ dacă $\bar{P}_2 \in [\bar{Y} - 2\sigma_Y, \bar{Y} + 2\sigma_Y]$.

În cazul în care \bar{P}_1 și \bar{P}_2 se găsesc în interval definit mai sus, atunci vom alege acel \bar{P} care se apropie mai mult de \bar{Y} .

Dacă definim $\bar{\gamma} = \frac{\bar{P}}{\bar{Y}}$, atunci $\bar{\gamma} \in \left[\frac{\bar{P} - 2\sigma_P}{\bar{Y}}, \frac{\bar{P} + 2\sigma_P}{\bar{Y}} \right]$

Dorim să definim γ în funcție de media și abaterea standard a seriei analizate Y și a

prognozelor P_1 și P_2 pentru a defini un interval $\left[1 - \frac{\bar{P} - 2\sigma_P}{\bar{Y}}, 1 - \frac{\bar{P} + 2\sigma_P}{\bar{Y}} \right]$ care să

facă o discriminare între cele două prognoze pe baza performanței lor. De fapt, modelul realizează o discriminare vizibilă între cele două prognoze în cazul în care există o diferență de performanță între cele două prognoze, diferență care se va reflecta în mediile și abaterile standard. Avem trei cazuri:

Cazul 1

Prognoza P_1 și prognoza P_2 sunt asemănătoare din punct de vedere al performanței. În acest caz

vom avea $\left[\frac{\bar{P}_1 - 2\sigma_{P_1}}{\bar{Y}}, \frac{\bar{P}_1 + 2\sigma_{P_1}}{\bar{Y}} \right] \approx \left[\frac{\bar{P}_2 - 2\sigma_{P_2}}{\bar{Y}}, \frac{\bar{P}_2 + 2\sigma_{P_2}}{\bar{Y}} \right]$

În acest caz, alegerea lui γ nu va influența prea mult valorile pe care le vor lua W_1 și W_2 , unde W_1 , W_2 sunt ponderile în care „se vor găsi” prognozele P_1 și P_2 în modelul final $Y = W_1 \cdot P_1 + W_2 \cdot P_2$.

Cazul 2

Prognoza P_1 este mai performantă decât P_2 pe tot segmentul martor pe care se calculează W_1 și W_2 . În acest caz vom avea

$\left[\frac{\bar{P}_1 - 2\sigma_{P_1}}{\bar{Y}}, \frac{\bar{P}_1 + 2\sigma_{P_1}}{\bar{Y}} \right] \neq \left[\frac{\bar{P}_2 - 2\sigma_{P_2}}{\bar{Y}}, \frac{\bar{P}_2 + 2\sigma_{P_2}}{\bar{Y}} \right]$

În acest caz, alegerea lui γ va influența valorile pe care le vor lua W1 și W2. O valoare prea mare sau prea mică a lui γ va modifica ponderea corectă a prognozelor. O valoare prea mare va permite prognozei mai ineficiente să „intre” de mai multe ori în intervalul

$\left[Y_{T+1} - \gamma |Y_{T+1}|, Y_{T+1} + \gamma |Y_{T+1}| \right]$. O valoare prea mică a lui γ va face ca prognoza mai

eficientă să nu intre în intervalul $\left[Y_{T+1} - \gamma |Y_{T+1}|, Y_{T+1} + \gamma |Y_{T+1}| \right]$ de un număr de ori care să reflecte diferența de performanță între ea și prognoza mai ineficientă.

Cazul 3

Împărțim segmentul martor de lungime n în două: unul de lungime n_1 (pe termen mediu) și unul de lungime n_2 (pe termen lung), unde $n_1 + n_2 = n$.

P1 este mai eficientă pe termen mediu, iar P2 este mai eficientă pe termen lung. În funcție de lungimea segmentului martor există tendința ca ineficiența prognozei P2 pe termen mediu să se compenseze cu eficiența ei pe termen lung și ca ponderile W1 și W2 să fie aproximativ egale.

Adică $[n_1 \approx n_2 \Rightarrow W_1 \approx W_2, n_1 \gg n_2 \Rightarrow W_1 \gg W_2, n_1 \ll n_2 \Rightarrow W_1 \ll W_2]$

În acest caz este recomandat să se calculeze ponderi diferite pe termen mediu și lung.

Alegerea valorii lui γ din intervalul identificat este și ea importantă. Fie P1 și P2 prognoze ale seriei Y. Dacă P1, P2 se aseamănă, din punct de vedere al performanței lor (aproximativ același MAPE), dar dorim să propunem o diferențiere între ele, diferențiere care să se reflecte în factorii de ponderare, atunci definim γ apropiat de limita inferioară a intervalului. Dacă prognozele nu au o performanță asemănătoare, dar nu dorim să facem o discriminare puternică între ele în funcție de eficiența lor pe segmentul martor, alegem γ apropiat de limita superioară a intervalului¹.

Identificarea diferenței de eficiență a prognozelor pe termen mediu și lung

Pentru a vedea cât de mare este diferența de performanță între P1 și P2 pe termen mediu și lung, calculăm mediile și abaterile standard pentru P1, P2, Y. Vom împărți segmentul martor de lungime n

¹ O întrebare care se ridică este dacă avem în cazul analizei segmentului martor o distribuție normală a valorilor. Intervalul în care se află γ este definit în funcție de medie și abaterea standard, prin urmare este important să știm dacă avem o distribuție normală. În același timp, este important să știm dacă probleme ca posibila asimetrie a distribuției valorilor din segmentul martor afectează definirea lui γ . Având în vedere amploarea problemei ea va fi abordată într-o lucrare ulterioară care va continua analiza realizată în această lucrare.

în două: unul de lungime n_1 (pe termen mediu) și unul de lungime n_2 (pe termen lung), unde

$n_1+n_2 = n$. În acest caz vom avea γ_1 și γ_2 care se calculează după metoda prezentată mai sus.

Calculăm pe termen mediu

$$C_1 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t]}(P_{1,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t]}(P_{2,t})} \quad (3), \text{ unde } P1 \text{ este mai eficientă pe termen mediu și } P2$$

mai eficientă pe termen lung

Calculăm pe termen lung

$$C_2 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t]}(P_{2,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t]}(P_{1,t})} \quad (4), \text{ unde } P1 \text{ este mai eficientă pe termen mediu și } P2$$

mai eficientă pe termen lung.

Pentru a vedea cât de mare este diferența de performanță pe termen mediu și lung a celor două prognoze analizăm C_1 și C_2 . Diferența de performanță între P1 și P2 crește cu cât $C_1 \gg 1$ și $C_2 \gg 1$. In cazul in care numărătorul este egal cu 0 considerăm C_1 și $C_2 \gg 0$

De ce alegem această metodă și nu calculul erorilor ?

Calculul erorilor a fost folosit pentru a se identifica ponderi care să fie utilizate în combinarea prognozelor Wenzel (2001). Metoda este criticabilă din perspectiva faptului că un calcul al erorilor pentru o perioadă nu este suficient de reprezentativ în contextul în care într-o altă perioadă prognozele nu se vor comporta exact în același mod Lупpoletti și Webb (1986), Wenzel (2001).

Dacă în loc de a calcula erorile stabilim un interval în care să se situeze prognozele și calculăm de câte ori prognozele se încadrează în acest interval evităm punctul slab al analizei pe bază de erori Wenzel (2001).

4. Aplicație

Aplicația își propune să analizeze comparativ eficiența celor trei metode de combinare a prognozelor și să vadă în ce măsură metoda pe care am propus-o este mai eficientă. Analiza a fost realizată pe baza a două serii.

(1) Seria 1, serie cu volatilitate medie - serie în care are loc o modificare de trend, dar această modificare nu este contrară evoluției de până atunci a seriei (Anexa 2 fig. 1).

(2) Seria 2, serie cu volatilitate mică - serie în care nu are loc o modificare a trendului (Anexa 2 fig. 2).

Seriile sunt propuse de Robert F. Nau profesor la Duke University în notele de curs ale unui curs de prognoză, pentru examenul final. Cele două serii au frecvență lunară.

Ne vom referi în mod generic la seriile analizate ca fiind seria Y, metodele de analiză prezentate se aplică în același mod seriei 1 și seriei 2. Seriile pe care le vom utiliza în analiză vor fi serii pe care deja le-am deflatizat și desezonalizat.

Vom nota metodele care vor fi combinate cu P1 și P2. P1 reprezintă prognoza seriei analizate prin metoda Holt-Winters (Eviews 2002). P2 reprezintă prognoza prin regresie liniară cu variabila explicativă venituri. Variabila explicativă venituri este propusă ca variabilă explicativă tot în cursul profesorului Robert F. Nau.

Combinarea prognozelor prin regresie

Pentru a identifica valorile lui W1 și W2 din ecuația $Y = W1*P1 + W2*P2$ am procedat după cum urmează:

Cazul 1

Avem $\hat{y} = W1*P1 + W2*P2$ (1)

Lungimea seriei Y este 1972:01 -1996:05. Oprim perioada 1992:06-1996:05 ca segment martor.

Parametrii W1 și W2 sunt evaluați prin MCMMP (metoda celor mai mici pătrate).

\hat{y} - este seria prognozată prin ecuația 1.

Cazul 2

În cazul 2 estimăm parametrii asociați prognozelor P1 și P2 luând în considerare și segmentul martor.

P1 și P2 sunt estimate pe baza seriei Y de lungime 1972:01 – 1988:05.

Parametrii W1 și W2 asociați prognozelor P1 și P2 sunt estimați pe baza seriei Y' de lungime 1972:01- 1992:05.

Avem: $\hat{y} = W1*P1 + W2*P2$ (2).

Datele au fost prelucrate în Eviews. Pentru a analiza eficiența metodei de combinare am calculat RMSE (root mean square error) și MAPE (mean absolute percentage error) (a se vedea Anexa 1 Tabelul 1)

Rezultatele analizei:

1) Dacă luăm în considerare întreaga perioadă de analiză, metoda combinată a avut rezultate mai bune (valori mai mici ale RMSE și MAPE) decât extrapolarea pe baza seriei analizate (metoda Holt-Winters) și regresiei pe baza variabilei explicative.

2) Dacă luăm în considerare doar perioada de validare ajungem la următoarele concluzii:

2.1.) Pe termen scurt și mediu utilizarea metodei combinate este mai avantajoasă, în sensul în care erorile sunt mai mici.

2.2.) Pe termen lung utilizarea variabilelor explicative duce la rezultate mai bune, erorile sunt mai mici. Acest lucru a fost observat și de Armstrong și Collopy (1998) (apud Armstrong (1999)).

2.3) În cazul în care seriile au volatilitate medie, metoda combinată are rezultate mai bune în comparație cu regresia pe baza variabilei explicative; cu cât volatilitatea scade cu atât diferența de acuratețe scade și ea. Acest lucru este explicabil prin faptul că seria cu volatilitate medie este mai puternic afectată de variația de tip „random” fapt care micșorează acuratețea regresiei cu variabila explicativă venituri.

3) Dacă evoluția seriilor nu susține ipoteza conform căreia "asupra" seriei acționează forțe care întăresc trendul existent, atunci prognozele prin extrapolare nu ar trebui să fie utilizate ca predicții pe termen lung. Utilizarea combinată a metodelor de extrapolare nu se dovedește mult mai eficientă pe termen lung. De exemplu în cazul seriei 2 comparația între valorile MAPE obținute pe perioada de validare 92 -94 și 92-96 pentru Holt-Winters arată aproape o dublare a erorilor 4.88 respectiv 8.71, iar pentru metoda combinată rezultatele nu sunt mult mai bune 4.39 respectiv 7.7.

4) Un posibil avantaj al utilizării combinate a metodelor este și calitatea variabilei explicative; în cazul regresiiilor pe baza variabilei explicative venituri s-a obținut o corelare a erorilor, fapt

sugerat de valorile testului Durbin-Watson. Metoda combinată care a utilizat regresia pe baza variabilei explicative venituri și metoda Holt-Winters a obținut o valoare mult mai bună a testului Durbin –Watson. (a se vedea Anexa 3)

Pentru a întări rezultatele obținute sunt necesare analize pe un număr mai mare de serii.

Concluzii parțiale

1) Pe termen mediu, dar mai ales scurt, prognozele pe baza variabilelor exogene sunt afectate de variația de tip „random” făcându-le mai puțin exacte decât prognozele prin extrapolare.

2) Lipsa eficienței metodei combinate pe termen lung poate fi explicată prin utilizarea MCMMP în perioada de analiză în care metodele de extrapolare generează mai puține erori decât regresiiile cu variabile exogene. În această situație, ponderea regresiei cu variabilele exogene în modelul final este mai mică (a se vedea Anexa 3).

3) Metoda combinată reduce din punctele slabe ale celor două metode, dar are tendința de a reduce și din avantajele. Acest lucru poate fi explicat și prin faptul că perioada de analiză nu este destul de relevantă pentru calitatea unei prognoze și că ar fi necesare reactualizări succesive.

Combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză(Wenzel 2001)

Fie Y seria analizată, P1 și P2 prognoze ale acestei serii, W1 și W2 factorii de ponderare, S segmentul martor de lungime n = 48 (perioada 1988:06-1992:05).

Stabilim valoarea lui γ pe baza metodei prezentate la pagina 5-6. Avem cazul general:

$$\gamma \in \left[1 - \frac{\bar{P} - 2\sigma_P}{\bar{Y}}, 1 - \frac{\bar{P} + 2\sigma_P}{\bar{Y}} \right] \text{ dacă } \bar{P} \pm \sigma_P \leq \bar{Y}$$

$$\gamma \in \left[\frac{\bar{P} - 2\sigma_P}{\bar{Y}} - 1, \frac{\bar{P} + 2\sigma_P}{\bar{Y}} - 1 \right] \text{ dacă } \bar{P} \pm \sigma_P \geq \bar{Y}$$

$$\gamma \in \left[1 - \frac{\bar{P} - 2\sigma_P}{\bar{Y}}, \frac{\bar{P} + 2\sigma_P}{\bar{Y}} - 1 \right] \text{ dacă } \bar{P} + \sigma_P \geq \bar{Y}, \bar{P} - \sigma_P \leq \bar{Y}$$

Pentru seria 1 avem:

$$\gamma \in [0.026, 0.039]$$

Alegem $\gamma = 0.0325$

Identificăm frecvența cu care cele două prognoze se încadrează în intervalul:

$$\left[Y_{T+1} - 0.0325|Y_{T+1}|, Y_{T+1} + 0.0325|Y_{T+1}| \right]$$

Calculăm W1 și W2 pe baza formulei (1). Vom avea W1 = 0.27, W2 = 0.73

$$\text{Avem } Y = 0.27*P1 + 0.73*P2$$

Pentru seria 2 avem:

$$\gamma \in [0.004, 0.094]$$

Alegem $\gamma = 0.049$

Identificăm frecvența cu care cele două prognoze se încadrează în intervalul:

$$\left[Y_{T+1} - 0.049|Y_{T+1}|, Y_{T+1} + 0.049|Y_{T+1}| \right]$$

Calculăm W1 și W2 pe baza formulei (1). Vom avea W1 = 0.47, W2 = 0.53

$$\text{Avem } Y = 0.39*P1 + 0.61*P2$$

Se observă că în cazul seriei 1 există o diferență mai evidentă între W1 și W2.

Calculăm C1 și C2 pentru cele două serii, pe baza formulelor (3) și (4).

Pentru seria 1 avem:

Pe termen mediu 1988:06-1990:05

$$C_1 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t})}$$

$$\text{Unde } \sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t}) = 6$$

$$\text{și } \sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t}) = 9$$

Pe termen lung 1990:06-1992:05

$$C_2 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t})}$$

$$\text{unde } \sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t}) = 7$$

$$\text{\c{si}} \sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t}) = 0$$

Pentru seria 2 avem:

Pe termen mediu 1988:06-1990:05

$$C_1 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t})}$$

$$\text{unde } \sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t}) = 15$$

$$\text{\c{si}} \sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t}) = 19$$

Pe termen lung 1990:06-1992:05

$$C_2 = \frac{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t})}{\sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t})}$$

$$\text{unde } \sum_{t=T-v+1}^T I_{[Y_t - \gamma_2 | Y_t |, Y_t + \gamma_2 | Y_t |]}(P_{2,t}) = 10$$

$$\text{\c{si}} \sum_{t=T-v+1}^T I_{[Y_t - \gamma_1 | Y_t |, Y_t + \gamma_1 | Y_t |]}(P_{1,t}) = 20$$

Analiza lui C1 și C2 pentru cele două serii susține concluziile la care am ajuns cu privire la calculul lui γ în funcție de medie și abaterea standard. Analiza lui C1 și C2 sugerează că există o diferență mai mare de performanță între prognozele P1 și P2 pe termen lung în cazul seriei 1, decât în cazul seriei 2, iar diferența de performanță pe termen mediu nu diferă semnificativ.

Calculăm γ pentru cele două serii. Valoarea pe care o alegem este limita superioară a intervalului definit de media prognozei P1 și a seriei Y și de abaterea standard a prognozei P1, unde P1 este prognoza în care diferența dintre media ei și media seriei analizate este mai mare.

În cazul seriei 1 alegem γ în intervalul $\gamma \in [0.035, 0.167]$.

Pentru $\gamma = 0.167$ avem $W1 = 0.43$, $W2 = 0.57$

Pentru $\gamma = 0.0325$ am avut $W1 = 0.27$, $W2 = 0.73$

În cazul seriei 2 alegem γ în intervalul $\gamma \in [0.096, 0.122]$

Pentru $\gamma = 0.122$ avem $W1 = 0.5$, $W2 = 0.5$

Pentru $\gamma = 0.049$ am avut $W1 = 0.39$, $W2 = 0.61$

Modificarea mai mare a ponderilor ($W1$, $W2$) în cazul seriei 1 se explică pe baza modului în care calculăm intervalul în care se încadrează γ . Pentru a calcula acest interval am utilizat media lui P1, P2 și Y și abaterile standard ale lui P1 și P2. Dacă performanța prognozei P1 este comparabilă cu performanța prognozei P2, atunci intervalul definit de media lui P1 și Y și abaterea standard a lui

P1, adică $\left[1 - \frac{\bar{P}_1 - 2\sigma_{P_1}}{\bar{Y}}, 1 - \frac{\bar{P}_1 + 2\sigma_{P_1}}{\bar{Y}} \right]$ va avea o lungime aproximativ egală cu

intervalul definit de mediile lui P2 și Y și dispersia lui P2, adică

$\left[1 - \frac{\bar{P}_2 - 2\sigma_{P_2}}{\bar{Y}}, 1 - \frac{\bar{P}_2 + 2\sigma_{P_2}}{\bar{Y}} \right]$.

În acest caz, alegerea unor valori prea mari sau prea mici ale lui γ nu va afecta în mod *discreționar* una dintre prognoze, adică ponderea celor două prognoze în modelul final nu se va modifica semnificativ¹.

¹ prin semnificativ înțelegem: diferența de pondere a celor două prognoze în modelul final nu se va traduce în erori mai mari, un MAPE mai mare.

Cu cât performanța prognozelor este mai diferită, cu atât valorile prea mari sau prea mici ale lui γ vor afecta în mod *discreționar* una dintre prognoze, adică ponderea celor două prognoze în modelul final se va modifica semnificativ.

Analiza erorilor RMSE (root mean square error) și MAPE (mean absolute percentage error)

Analiza scoate în evidență limitele metodei. Metoda nu face o discriminare eficientă între performanțele prognozelor pe termen mediu și lung, astfel încât ponderea celor două prognoze P1 și P2 în modelul final ($Y = W1*P1+W2*P2$) nu reflectă exact perioadele în care o prognoză este mai eficientă decât cealaltă. Această situație se observă bine în cazul seriei 1, în care avem $C2 \gg 1$, fapt care scoate în evidență performanța mai bună a prognozei P2 în perioada 1994-1996. Această performanță nu este reflectată suficient de bine în valoarea MAPE (mean absolute percentage error) din perioada 1992-1996 (Anexa 1 tabelul 2). În condițiile în care nu se aleg ponderi diferite pentru perioadele în care performanța prognozelor P1 și P2 este diferită, este greu să obținem o performanță mai bună.

Utilizarea unor factori de ponderare diferiți pentru perioada medie și lungă va genera o prognoză mai eficientă.

Aplicăm factori de ponderare diferiți pentru seria 1.

Fie Y seria prognozată, P1 și P2 prognozele seriei Y și S un segment martor de lungime n.

Vom împărți segmentul martor de lungime n în două: unul de lungime n1 (pe termen mediu) și unul de lungime n2 (pe termen lung), unde $n1+n2 = n$. În acest caz vom avea γ_1 și γ_2 care se calculează după metoda prezentată la pagina 7.

Pentru n1 verificăm de câte ori cele două prognoze se încadrează în intervalul:

$$\left[Y_{T+1} - \gamma_1 |Y_{T+1}|, Y_{T+1} + \gamma_1 |Y_{T+1}| \right]$$

Pentru n2 verificăm de câte ori cele două prognoze se încadrează în intervalul:

$$\left[Y_{T+1} - \gamma_2 |Y_{T+1}|, Y_{T+1} + \gamma_2 |Y_{T+1}| \right].$$

Vom avea $W1,1 = 0.4$ și $W1,2 = 0.6$ pentru n1 și $W2,1 = 0$, $W2,2 = 1$ pentru n2.

Avem:

$$Y_{t+i} = W_{1,1} \times P_1 + W_{1,2} \times P_2, i = 1 \dots n_1$$

$$Y_{t+z} = W_{2,1} \times P_1 + W_{2,2} \times P_2, z = n_1 \dots n_2$$

Calculul erorilor MAPE ne arată superioritatea utilizării unor ponderi diferite pentru perioade diferite în funcție de performanța celor două prognoze pe acele perioade. (a se vedea Anexa 1 Tabelul 5).

Combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung prin utilizarea factorilor de ponderare

Fie Y seria analizată, de lungime 1972:01 – 1996:05, P1 și P2 prognoze ale acestei serii. După cum am precizat la începutul aplicației P1 reprezintă prognoza seriei analizate prin metoda Holt-Winters (Eviews 2002), P2 reprezintă prognoza prin regresie liniară cu variabila explicativă venituri.

Intuitiv P1 va fi mai eficientă pe termen scurt și mediu decât P2 (Armstrong și Collopy (1998) (apud Armstrong (1999)). Acest lucru este explicabil prin faptul că P2 este mai puternic afectată pe termen scurt de variația de tip „random” decât P1, această variație micșorând acuratețea regresiei cu variabila exogenă venituri.

Utilizăm ca perioadă de analiză a diferenței de performanță între cele două prognoze pe termen mediu și lung, perioada 1988:06-1992:05.

Metoda presupune calcularea lui C1 și C2 pentru a se vedea cât de mare este diferența de performanță între cele două prognoze. Calculele au fost făcute deja (pag. 13-14).

Pentru prima serie am obținut $C_1 = 0.66$, $C_2 \gg 1$

Prin urmare, prognoza P2 este mult mai eficientă pe termen lung decât prognoza P1, iar prognoza P1 este mai ineficientă pe termen mediu decât P2. Cu toate că prognoza P1 nu este mai eficientă pe

termen mediu, prognoza seriei Y prin formula $\hat{y}_{t+m} = \frac{m}{n} \times P_{2,t+m} + \frac{n-m}{n} \times P_{1,t+m}$,

unde P2 este prognoza mai eficientă pe termen lung, iar P1 este prognoza care trebuie să fie mai eficientă pe termen mediu, va da rezultate bune după cum vom vedea în analiza erorilor (Anexa 1 tabelul 3). Aceste rezultate bune se pot explica prin: 1) prognoza P2 este mult mai eficientă decât

P1 pe termen lung , 2) prognoza P1 a fost generată prin utilizarea unui model adaptativ (Holt-Winters) care este mai eficient pe termen scurt decât prognoza prin regresie pe baza unei variabile explicative ((Armstrong si Collopy (1998) (apud Armostrong (1999)). Acest lucru este explicabil prin faptul că P2 este mai puternic afectată de variația de tip “random” decât P1, această variație micșorând acuratețea regresiei cu variabila exogenă venituri. Performanța mai bună a seriei P1 pe termen scurt va compensa tendința de pierdere graduală din performanță pe termen mediu.

Pentru prima serie calculul erorilor prin *RMSE* (*root mean square error*) și *MAPE* (*mean absolute percentage error*) (*a se vedea Anexa 1, Tabelul 3*) ne arată superioritatea metodei combinate față de prognozele P1 și P2 în toate cele trei perioadele analizate: 1972-1992, 1992-1996, 1992-1994.

Interesant este modul în care se completează pe termen mediu P1 și P2 pentru a genera erori mai mici (calculate prin MAPE), în cazul lui P1 avem MAPE = 4.88, în cazul lui P2 avem MAPE = 5.76, iar in cazul metodei combinate avem MAPE = 3.55. Îmbunătățirea este mult mai bună decât cea generată prin utilizarea metodei Wenzel sau cea prin utilizarea combinării prin regresie (Anexa 1 , tabelul 1 și tabelul 2) și se explică prin faptul că metoda propusă surprinde creșterea graduală în performanță a prognozei P2 și scăderea graduală în performanță a prognozei P1.

Pentru a doua serie am obținut $C_1 = 0.78$, $C_2 = 2$

Prognoza P2 este puțin mai eficientă decât prognoza P1 pe termen mediu, iar pe termen lung prognoza P2 este evident mai eficientă decât prognoza P1.

Diferența mai mică de performanță între cele două prognoze P1 și P2 pe termen lung, față de cazul seriei 1, va genera o prognoză mai puțin bună.

Rezultate mai bune, un MAPE mai mic am obținut doar pe perioada 1972-1992. Pe intervalul 1992-1996, 1992-1994 am obținut rezultate îndoielnice, adică metoda combinată era mai eficientă decât una dintre prognoze, dar mai ineficientă decât cealaltă (Anexa 1 tabelul 3).

Analiza comparativă a celor trei metode de combinare a prognozelor

Cea mai eficientă prognoză este CFP (Combinarea prognozelor în funcție de eficiența lor pe termen scurt și lung prin utilizarea factorilor de ponderare) cu o performanță evident mai bună pentru seria 1. Situația poate fi explicată prin scăderea graduală în performanță a prognozei P1 și creșterea

graduală în performanță a prognozei P2. Acest fenomen nu este surprins de combinarea prognozelor pe baza regresiei și nici de metoda Wenzel.

Analiza seriei 2 scoate în evidență următorul fenomen: cu cât performanța pe segmentul martor a prognozelor este mai apropiată, cu atât diferența de performanță între CFP și metoda Wenzel scade, iar performanța combinării prin regresie scade și ea. Acest lucru poate fi explicat prin analiza metodelor la care fac apel cele trei modele.

Combinarea prognozelor prin regresie estimează $W1$ și $W2$ din ecuația $Y = W1*P1 + W2*P2$ prin MCMMP. Pentru a obține parametri stabili avem nevoie de o perioadă mai lungă de analiză (în comparație metodele Wenzel și CFP nu necesită o perioadă lungă de analiză). În cazul combinării prognozelor P1 și P2 prin regresie, ponderea lor în model este afectată de comportamentul lui P1. Prognoza P1 (prognoza rezultată în urma utilizării unui model adaptativ Holt-Winters) este mai eficientă în perioada de analiză decât P2 (regresia liniară pe baza variabilei explicative venituri). Problema este că în general, modelele adaptative au o performanță mult mai bună în perioada de analiză în comparație cu performanța înregistrată pe segmentul martor. Performanța mai bună a prognozei P1 în perioada de analiză și lungimea mai mare a perioadei de analiză față de lungimea segmentului martor va influența valoarea parametrului $W1$, care va reflecta mai puțin performanța mai slabă a lui P1 pe segmentul martor. Din aceasta cauză, combinarea prognozelor prin regresie este mai ineficientă pe segmentul martor.

În cazul seriei S2, diferența de performanță între metoda Wenzel și CFP scade deoarece CFP are ca ipoteză de bază diferența de performanță între prognozele P1 și P2 pe termen mediu și lung. Cu cât această diferență scade cu atât performanța metodei scade.

Cu toate că performanța metodei CFP scade, fapt atestat și de valorile lui $C1$ și $C2$, prognoza CFP rămâne cea mai eficientă pe perioada 1992-1996, iar performanța ei în perioada 1992-1994 se apropie de performanța celorlalte două metode utilizate (a se vedea Anexa 1 tabelul 4).

Anexa 1
Sumarul erorilor de prognoză pentru cele două serii analizate

	1972-1996							
	S1				S2			
	P1	P2	Rcomb	Cregresie	P1	P2	Rcomb	Cregresie
MAPE (1972-1996)	3.83	4.33	3.62	3.88	5.26	5.81	4.91	5.16
MAPE (1992-1996)	8.71	5.79	7.66	6.19	7.97	4.76	7.03	7.94
MAPE 1992-1994	4.88	5.76	4.38	4.42	6.13	4.72	5.66	6.67
RMSE (1972-1996)	0.0517	0.0486	0.0475	0.0473	0.0261	0.0262	0.0239	0.0255
RMSE (1992-1996)	0.1074	0.0744	0.0961	0.0961	0.0482	0.0309	0.0428	0.0472
RMSE (1992-1994)	0.0586	0.0652	0.0509	0.0506	0.0366	0.0283	0.0339	0.0386

Tabel 1 Sumar al erorilor de prognoză pentru cele două serii analizate. Combinarea prognozelor prin regresie

Legenda: rezultat mai bun rezultat îndoielnic rezultat mai slab

P1 si P2 – prognozele seriilor 1 si 2, Rcomb – combinarea prognozelor prin regresie fără utilizarea valorilor din segmentul martor, Cregresie combinarea prognozelor prin regresie, se utilizează valorile din segmentul martor.

	1972-1996					
	S1			S2		
	P1	P2	WENZEL	P1	P2	WENZEL
MAPE (1972-1996)	3.83	4.33	3.81	5.26	5.81	5.0
MAPE (1992-1996)	8.71	5.79	5.97	7.97	4.76	5.94
MAPE 1992-1994	4.88	5.76	4.61	6.13	4.72	5.27
RMSE (1972-1996)	0.0517	0.0486	0.0451	0.0261	0.0262	0.0236
RMSE (1992-1996)	0.1074	0.0744	0.0778	0.0482	0.0309	0.0367
RMSE (1992-1994)	0.0586	0.0652	0.0532	0.0366	0.0283	0.0314

Tabel 2 Sumar al erorilor de prognoză pentru cele două serii analizate. Metoda „Wenzel”

Legenda: rezultat mai bun rezultat îndoielnic rezultat mai slab
P1 și P2 – prognozele seriilor 1 și 2, Wenzel - combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză

	1972-1996					
	S1			S2		
	P1	P2	CFP	P1	P2	CFP
MAPE (1972-1996)	3.83	4.33	3.25	5.26	5.81	4.69
MAPE (1992-1996)	8.71	5.79	5.36	7.97	4.76	5.71
MAPE 1992-1994	4.88	5.76	3.55	6.13	4.72	5.55
RMSE (1972-1996)	0.0517	0.0486	0.0401	0.0261	0.0262	0.0239
RMSE (1992-1996)	0.1074	0.0744	0.0711	0.0482	0.0309	0.0428
RMSE (1992-1994)	0.0586	0.0652	0.0443	0.0366	0.0283	0.0339

Tabel 3 Sumar al erorilor de prognoză pentru cele două serii analizate. Metoda „CFP”

Legenda: rezultat mai bun rezultat îndoielnic rezultat mai slab

P1 si P2 – prognozele seriilor 1 si 2, CFP combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare

	1972-1996					
	S1			S2		
	Rcomb	WENZEL	CFP	Rcomb	WENZEL	CFP
MAPE (1972-1996)	3.62	3.81	3.25	4.91	5.0	4.69
MAPE (1992-1996)	7.66	5.97	5.36	7.03	5.94	5.71
MAPE 1992-1994	4.38	4.61	3.55	5.66	5.27	5.55
RMSE (1972-1996)	0.0475	0.0449	0.0401	0.0239	0.0234	0.0239
RMSE (1992-1996)	0.0961	0.0785	0.0711	0.0428	0.0381	0.0428
RMSE (1992-1994)	0.0509	0.0522	0.0443	0.0339	0.0321	0.0339

Tabel 4 Sumar al erorilor de prognoză pentru cele două serii analizate. Analiza comparată a celor trei metode de combinare

Legenda: rezultatul cel mai bun rezultatul cel mai slab

Rcomb – combinarea prognozelor prin regresie fără utilizarea valorilor din segmentul martor , Wenzel - combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză, CFP - combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare

	1972-1996	
	S1	
	WENZEL ALTERNATIV	WENZEL
MAPE (1972-1996)	3.21	3.81
MAPE (1992-1996)	5.1	5.97
MAPE 1992-1994	4.4	4.61

Tabel 5 Sumar al erorilor de prognoză pentru seria 1.

Legenda: rezultat mai bun

Anexa 2

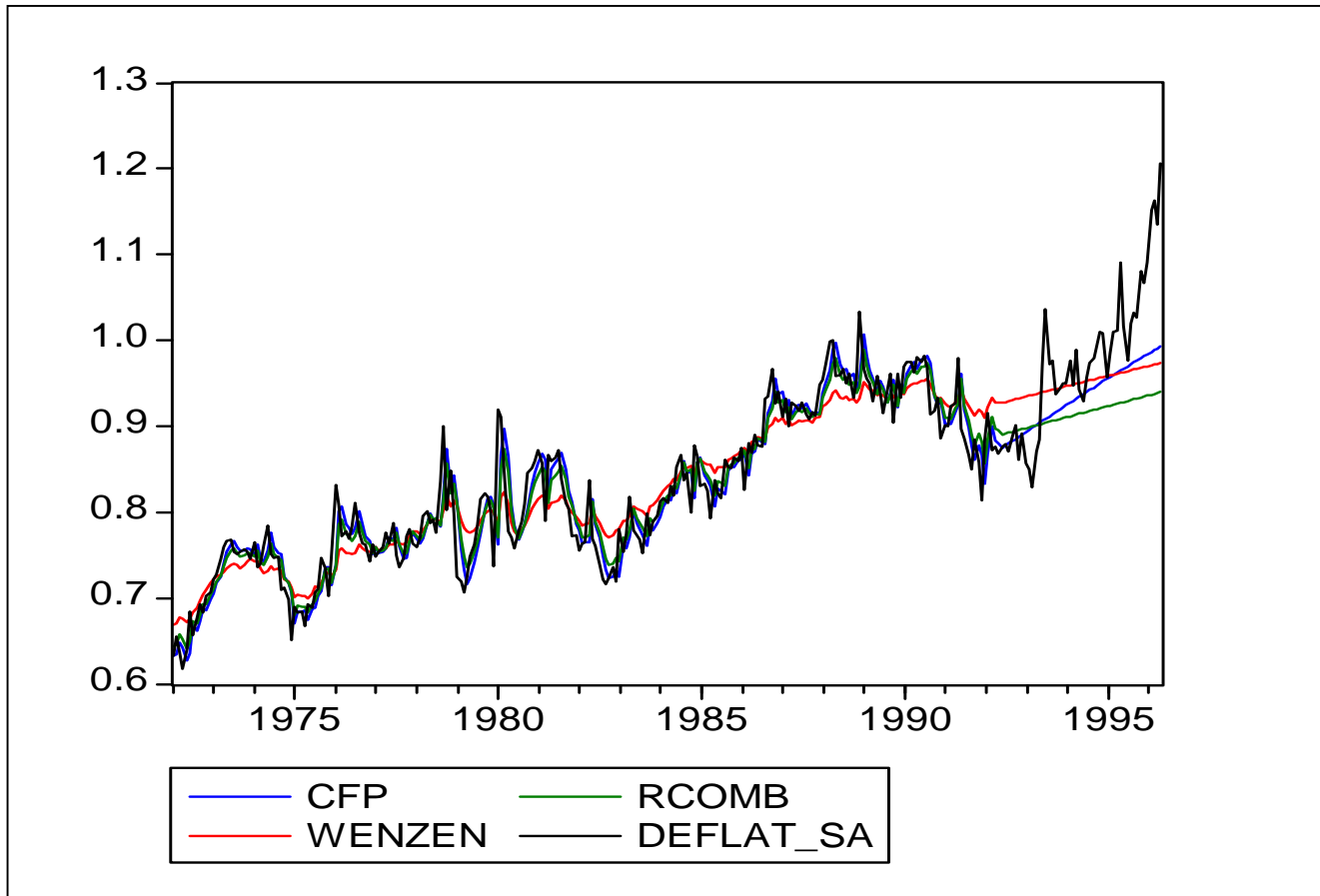


Fig. 1 Prognoza seriei 1 prin cele trei metode de combinare

CFP = combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare

RCOMB = combinarea prognozelor prin regresie

WENZEL = combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză

DEFLAT_SA = seria 1

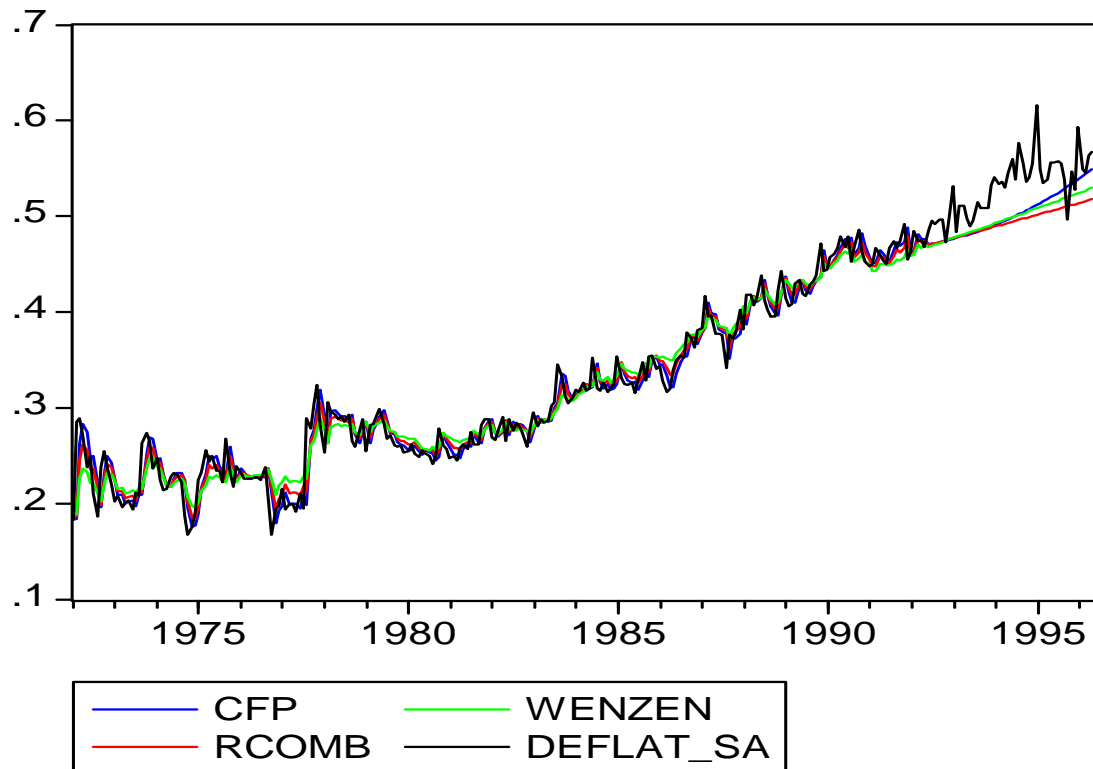


Fig. 2 Prognoza seriei 2 prin cele trei metode de combinare

CFP = combinarea prognozelor în funcție de eficiența lor pe termen mediu și lung, prin utilizarea factorilor de ponderare

RCOMB = combinarea prognozelor prin regresie

WENZEL = combinarea prognozelor pe baza încadrării prognozelor într-un interval de analiză

DEFLAT_SA = seria 2

Anexa 3

Dependent Variable: DEFLAT_SA

Method: Least Squares

Date: 12/08/04 Time: 19:12

Sample: 1972:01 1992:05

Included observations: 245

DEFLAT_SA = C(2)*EXPLICSM+C(3)*EXPLICSM^2

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	0.000379	6.47E-06	58.61317	0.0000
C(3)	-3.28E-08	2.10E-09	-15.56744	0.0000
R-squared	0.782260	Mean dependent var		0.821348
Adjusted R-squared	0.781364	S.D. dependent var		0.089349
S.E. of regression	0.041778	Akaike info criterion		-3.504764
Sum squared resid	0.424133	Schwarz criterion		-3.476182
Log likelihood	431.3335	Durbin-Watson stat		0.575245

Regresie liniară cu variabila exogenă venituri - seria 1

DEFLAT_SA – seria prognozată (seria 1)

EXPLICSM – prognoza cu Holt-Winters a variabilei exogene venituri

Dependent Variable: DEFLAT_SA

Method: Least Squares

Date: 12/08/04 Time: 19:03

Sample: 1972:01 1992:05

Included observations: 245

DEFLAT_SA = C(2)*REXPPLIC + C(3)*DEFLATSM

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	0.215203	0.049924	4.310639	0.0000
C(3)	0.784536	0.049855	15.73632	0.0000
R-squared	0.892158	Mean dependent var		0.821348
Adjusted R-squared	0.891714	S.D. dependent var		0.089349
S.E. of regression	0.029402	Akaike info criterion		-4.207396
Sum squared resid	0.210064	Schwarz criterion		-4.178814
Log likelihood	517.4060	Durbin-Watson stat		1.805793

Combinarea prognozelor prin regresie –seria 1

DEFLAT_SA – seria prognozată (seria 1)

REXPPLIC – prognoza seriei 1 pe baza regresiei liniare cu variabila exogenă venituri

DEFLATSM - prognoza cu Holt-Winters a seriei 1

Dependent Variable: DEFLAT_SA

Method: Least Squares

Date: 12/07/04 Time: 17:12

Sample: 1972:01 1992:05

Included observations: 245

DEFLAT_SA = C(2)*EXPLICSM+C(3)*EXPLICSM^2

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	2.15E-05	3.90E-06	5.502423	0.0000
C(3)	2.87E-08	1.27E-09	22.57452	0.0000
R-squared	0.913585	Mean dependent var		0.312340
Adjusted R-squared	0.913230	S.D. dependent var		0.085621
S.E. of regression	0.025221	Akaike info criterion		-4.514142
Sum squared resid	0.154573	Schwarz criterion		-4.485560
Log likelihood	554.9823	Durbin-Watson stat		0.609515

Regresie liniară cu variabila exogenă venituri - seria 2

DEFLAT_SA – seria prognozată (seria 2)

EXPLICSM – prognoza cu Holt-Winters a variabilei exogene venituri

Dependent Variable: DEFLAT_SA

Method: Least Squares

Date: 12/07/04 Time: 18:01

Sample: 1972:01 1992:05

Included observations: 245

DEFLAT_SA = C(2)*REXPPLIC + C(3)*DEFLATSM

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	0.246186	0.049954	4.928244	0.0000
C(3)	0.755172	0.049915	15.12906	0.0000
R-squared	0.955501	Mean dependent var		0.312340
Adjusted R-squared	0.955317	S.D. dependent var		0.085621
S.E. of regression	0.018099	Akaike info criterion		-5.177823
Sum squared resid	0.079598	Schwarz criterion		-5.149241
Log likelihood	636.2833	Durbin-Watson stat		1.779353

Combinarea prognozelor prin regresie –seria 2

DEFLAT_SA – seria prognozată (seria 1)

REXPPLIC – prognoza seriei 1 pe baza regresiei liniare cu variabila exogenă venituri

DEFLATSM - prognoza cu Holt-Winters a seriei 1

Bibliografie

1. Armstrong (1999), *Forecasting for Marketing*, International Thompson Press, 92-119
2. Armstrong J. S., Collopy F. (1992), *Error measures for generalizing about forecasting. Methods: Empirical comparison*, International Journal of Forecasting, 8, 69-80
3. Bates si Granger (1969), *The combination of forecast*, Operational Research Quarterly, 20, 451-468
4. Bunn (1975), *A Bayesian approach to the linear combination of forecasts*, Operational Research Quarterly, 26, 653-660
5. Bunn (1977), *A comparative evolution of the outperformance and minimum variance procedures for linear syntheses of forecasting*, Operational Research Quarterly, 28, 653-660
6. Crane si Coty (1967), *A two-stage forecasting model Exponential smoothing and multiple regression*, Management Science, 13, B501-B507
7. Clemen T. Robert (1989), *Combining forecasts. A review and adnnotated bibliography*, International Journal of Forecasting, 5, 559-583
8. Granger si Ramanathan (1984), *Improved methods of forecasting*, Journal of Forecasting, 3, 197-204
9. Granger si Newbold (1973), *Some coments on the evolution of forecasting*, Applied Economics, 5, 34-37
10. Guerard si Beidleman (1987), *Composite earnings forecasting efficiency*, Interface, 17, 103-113
11. Lippoletti si Webb (1986), *Defining and improving the accuracy of macroeconomic forecasting. Contribution from a VAR model*, Journal of Business, 59, 263-285
12. Poulos, Kvaneli, Pavar (1987), *A comparation of the accuracy of the Box-Jankins method with that of automated forecasting methods*, International Journal of Forecasting, 3, 261-267
13. Wenzel (2001), *Hits-and-misses for the evaluation and combination of forecasts*, Journal of Applied Statistics, 28, 6, 759-773
14. Eviews (2002), *Eviews 4 User's Guide*, Quantitative Macro Software, LLC
15. www.duke.edu/~rnau/411home.htm