

LA COMBINAZIONE DI PREVISIONI

Monica Billio

Università Ca' Foscari e GRETA, Venezia

Domenico Sartore

Università Ca' Foscari e GRETA, Venezia

Introduzione. Negli ultimi decenni sono state sviluppate molteplici tecniche per la previsione di serie economiche ed in particolare finanziarie. Molta letteratura¹ si è interessata anche allo studio del modo ottimale di combinare tali diverse previsioni, qualora disponibili. Infatti la previsione combinata risulta spesso essere la migliore. Ad esempio, Granger e Newbold (1986) mostrano come combinando due previsori non distorti un passo avanti attraverso una media ponderata, sia facile calcolare i pesi che permettono di avere una varianza dell'errore di previsione della previsione combinata più piccola delle singole varianze dei previsori combinati. Ciò permette agli stessi Autori di affermare:

«...a priori, it is reasonable to expect in most practical situations that the best available combined forecast will outperform the better individual forecast - it cannot, in any case, do worse».

Lo scopo di questo contributo è di valutare l'effettiva possibilità di migliorare la capacità previsiva di serie finanziarie, utilizzando un opportuno combinatore, che permetta non solo di fondere i risultati di metodi di previsione diversi ma anche di legare previsioni ottenute su scadenze diverse. Infatti, al fine di utilizzare al meglio l'informazione disponibile su una scadenza più lunga (quale ad esempio la settimana nel caso di serie finanziarie), si intendono rivedere le previsioni giornaliere sulla base della previsione settimanale.

In particolare, saranno confrontate e combinate cinque diverse metodologie utilizzate per la previsione di serie finanziarie giornaliere. Tali metodologie possono essere raggruppate in due diverse classi:

- nella prima rientrano i metodi di tipo *adattivo*, nel senso che il modello migliore è definito sulla base del miglior risultato previsivo ottenuto su un sottocampione (*validation set*). Esse sono:
 - le reti neurali;
 - la logica *fuzzy*;
 - i modelli caotici.
- nella seconda sono riuniti i modelli di tipo *strutturale*, definiti sulla base di ipotesi circa il comportamento della serie ed il modello migliore è quindi selezionato sulla base del miglior adattamento sull'intero campione a disposizione. Considereremo:

- i modelli con varianza condizionalmente autoregressiva (GARCH);
- i modelli a cambiamenti di regime.

Sarà anche considerata una previsione settimanale ottenuta attraverso un modello di tipo strutturale a parametri variabili. Infatti sulla scadenza settimanale è possibile sfruttare informazioni che normalmente su base giornaliera non sono disponibili.

Le serie oggetto di studio è data dai valori minimi e massimi del tasso di cambio marco/dollaro e del *future* sui BTP, registrate al LIFFE ed osservate tra il maggio 1994 ed il luglio 1997.

Il combinatorio. Le previsioni giornaliere sono effettuate per l'intera settimana e riviste ogni giorno. Per combinare le previsioni si è scelto di utilizzare il metodo proposto da Granger e Ramanathan (1984), che consiste nel regredire i valori effettivi sulle previsioni e considerare i coefficienti di regressione (vincolati a sommare all'unità) come pesi delle diverse previsioni.

La serie settimanale consiste in una media aritmetica dei valori giornalieri ed è prevista sulla base di un modello a parametri variabili. Quest'ultimo utilizza, come si è detto, anche alcune informazioni disponibili su base settimanale.

Per combinare le previsioni giornaliere con la previsione settimanale si utilizza innanzitutto l'informazione derivante dalla relazione esistente tra media settimanale e singoli dati giornalieri (*condizione aggregante*):

$$[-1/5 \dots -1/5 \ 1] Y_{ts} = H Y_{ts} = 0$$

dove $Y_{ts} = (y_{ts1}, \dots, y_{ts5}, y_{ts})$ è il vettore che raccoglie le cinque previsioni giornaliere (dal lunedì al venerdì) e la previsione settimanale data t , per la settimana s .

L'interesse di rivedere le previsioni giornaliere sulla base di quella settimanale (e non solo il viceversa) deriva dal fatto che quest'ultima può essere effettuata sulla base di informazioni che non sono disponibili su frequenza giornaliera e quindi sarà potenzialmente migliore delle singole previsioni giornaliere.

Il processo di revisione è di tipo circolare, in quanto ogni nuova osservazione permetterà di aggiornare le previsioni giornaliere e la previsione settimanale e quindi attraverso l'operazione di ancoraggio, derivante dalla condizione aggregante, queste ultime saranno ulteriormente riviste. Chiameremo queste ultime *previsioni ancorate*.

Per l'applicazione di tale procedura consideriamo che la previsione settimanale sia effettuata per la sola settimana successiva e che le previsioni giornaliere siano fatte all'interno della settimana prevista (il lunedì quindi le previsioni sono anche a più passi avanti, mentre il giovedì si realizzano solo previsioni un passo avanti). La procedura di ancoraggio può essere formalizzata attraverso il seguente *modello spazio-stato*:

$$0 = H Y_{ts}$$

$$Y_{ts} = F_t Y_{t-1s} + F_t^{1/2} u_t$$

dove la prima equazione (detta *equazione di misura*) rappresenta la condizione aggregante, mentre la seconda (detta *equazione di transizione*) indica come le singole previsioni sono aggiornate, quando si rende disponibile una nuova osservazione (F_t è una matrice opportunamente stimata, così come la matrice di varianza e covarianza F_t , ed u_t è un termine di disturbo a media nulla). Tale rappresentazione è molto utile in quanto permette di scrivere la procedura di revisione² delle previsioni giornaliere utilizzando il *filtro di Kalman*. Quest'ultimo permette di calcolare facilmente la seguente proiezione lineare:

$$EL(Y_{ts} / H Y_{ts})$$

Tale proiezione consente di ottenere la migliore previsione di Y_{ts} sulla base dell'informazione fornita dalla condizione aggregante. La procedura di ancoraggio utilizza quindi la condizione aggregante e l'informazione contenuta nella matrice F_t . Quest'ultima infatti permette di cogliere le relazioni esistenti tra le varie previsioni ed è stimata utilizzando le serie storiche degli errori di previsione. Essa permette in particolare di sfruttare la presenza di eventuali comportamenti ciclici all'interno della settimana.

Applicazioni. Tale procedura è stata utilizzata per la previsione dei massimi e minimi del tasso di cambio marco/dollaro (chiamato DEM d'ora in poi) e del futuro sui BTP (chiamato FIB d'ora in poi). Il periodo campionario va dal 17 maggio 1994 al 14 luglio 1997.

La scelta di prevedere i massimi ed i minimi di una serie finanziaria è stata fatta partendo dall'osservazione che i *traders*, per pianificare le loro operazioni, hanno bisogno di prevedere l'escursione giornaliera della serie, piuttosto che il semplice valore di fine giornata. Inoltre, da un punto di vista statistico, si tratta di serie molto particolari, caratterizzate da distribuzioni fortemente non gaussiane e soprattutto con osservazioni non equi-distanziate, per le quali, quindi, i modelli non lineari utilizzati possono essere adatti.

Per valutare le capacità previsive dei vari modelli e del combinatore si è utilizzato il criterio dell'errore assoluto medio³ (MAE) e si è usata la semplice ipotesi di *random walk* (il domani è uguale all'oggi) come termine di paragone. Si sono utilizzate le 5 settimane intercorrenti tra il 14/07/97 ed il 28/08/97 per i vari confronti.

Le previsioni settimanali utilizzano come informazioni supplementari le richieste di lavoro ed il livello delle vendite al dettaglio negli Stati Uniti, nel caso del DEM, ed il tasso di interesse sui pronti contro termine, che è un ottimo indicatore della politica monetaria della banca d'Italia e l'offerta netta delle aste dei BTP, nel caso del FIB.

Le tabelle 1e 2 presentano i risultati ottenuti considerando la serie DEM. Ad eccezione della logica *fuzzy*, i vari modelli non riescono a migliorare la performance del *random walk*. Nonostante i risultati deludenti delle varie metodologie, le previsioni combinate ed in particolare quelle ancorate presentano i migliori risultati, anche rispetto al *random walk*. In particolare per la serie del minimo, il guadagno delle previsioni del combinatore ancorato rispetto al *random walk* è del 5,18%.

Tabella 1
MAE Serie del minimo del DEM

Passi di previsione	GARCH	Modelli a cambiamenti di regime	Reti neurali	Modelli caotici	Logica <i>fuzzy</i>	Combinatore semplice	Combinatore ancorato
1	0.0105	0.0109	0.0104	0.0107	0.0094	0.0099	0.0093
2	0.0153	0.0164	0.0155	0.0155	0.0148	0.0145	0.0129
3	0.0181	0.0195	0.0179	0.0180	0.0191	0.0191	0.0179
4	0.0223	0.0239	0.0225	0.0233	0.0232	0.0222	0.0205
5	0.0243	0.0263	0.0248	0.0252	0.0249	0.0219	0.0196

MAE *random walk* un passo avanti 0.0098

Tabella 2
MAE Serie del massimo del DEM

Passi di previsione	GARCH	Modelli a cambiamenti di regime	Reti neurali	Modelli caotici	Logica <i>fuzzy</i>	Combinatore semplice	Combinatore ancorato
1	0.0082	0.0083	0.0082	0.0083	0.0080	0.0081	0.0080
2	0.0133	0.0135	0.0129	0.0128	0.0133	0.0130	0.0123
3	0.0169	0.0171	0.0167	0.0169	0.0173	0.0169	0.0154
4	0.0198	0.0212	0.0200	0.0204	0.0205	0.0199	0.0164
5	0.0236	0.0243	0.0237	0.0238	0.0242	0.0239	0.0216

MAE *random walk* un passo avanti 0.0081

Le varie metodologie, ad eccezione della logica *fuzzy* e delle reti neurali, presentano migliori risultati previsivi nel caso delle serie del minimo e del massimo del FIB, come è possibile valutare dalle tabelle 3 e 4. In questo esempio, le previsioni ancorate, pur presentando un guadagno del 10,41% rispetto al *random walk*, non risultano la migliore previsione e, per i passi successivi al primo, in alcuni casi sono meno soddisfacenti delle previsioni combinate semplici. Ciò nonostante, la performance delle previsioni combinate può essere considerata molto interessante, in quanto rimane sempre posizionata tra le prime tre migliori, dimostrando perciò di essere robusta e quindi affidabile.

Tabella 3
MAE Serie del minimo del FIB

Passi di previsione	GARCH	Modelli a cambiamenti di regime	Reti neurali	Modelli caotici	Logica <i>fuzzy</i>	Combinatore semplice	Combinatore ancorato
1	0.380073	0.380740	0.500870	0.479857	0.523313	0.4343	0.4331
2	0.676028	0.644207	0.702200	0.712541	0.667707	0.7133	0.7337
3	0.735089	0.701182	0.775743	0.816193	0.760836	0.8147	0.7552
4	0.917967	0.895507	0.938167	0.998841	0.899400	0.9138	0.8374
5	1.043473	1.003442	1.043169	1.126731	1.025919	1.0313	1.0539

MAE *random walk* un passo avanti 0.483448

Tabella 4
MAE Serie del massimo del FIB

Passi di previsione	GARCH	Modelli a cambiamenti di regime	Reti neurali	Modelli caotici	Logica <i>fuzzy</i>	Combinatore semplice	Combinatore ancorato
1	0.339907	0.335260	0.381627	0.377343	0.489653	0.3565	0.3546
2	0.623438	0.596324	0.621703	0.632200	0.641083	0.6098	0.6160
3	0.681286	0.655182	0.688943	0.691871	0.734296	0.6569	0.6433
4	0.824467	0.778581	0.830456	0.845285	0.784841	0.7513	0.7224
5	0.984554	0.962419	0.933842	0.939031	0.871342	0.9189	0.9078

MAE *random walk* un passo avanti 0.377241

Infine la tabella 5 presenta la percentuale di segni corretti ottenuti nelle previsioni un passo avanti. In linea con i risultati precedenti, nessuno dei modelli presenta un comportamento uniformemente migliore: infatti la logica *fuzzy* ed i modelli caotici sono i migliori nel caso del DEM, mentre i modelli a cambiamenti di regime ed in parte i GARCH sono preferibili nel caso del FIB. Le previsioni combinate sono quindi la scelta più opportuna, in quanto presentano risultati uniformemente validi.

Tabella 5
Percentuale di segni corretti

	GARCH	Modelli a cambiamenti di regime	Reti neurali	Modelli caotici	Logica <i>fuzzy</i>	Combinatore semplice	Combinatore ancorato
Max FIB	0.5517	0.6207	0.3793	0.4828	0.4138	0.5517	0.5862
Min FIB	0.6897	0.7241	0.4828	0.4828	0.3793	0.5862	0.6552
Max DEM	0.4828	0.5172	0.5172	0.5517	0.5517	0.5172	0.5517
Min DEM	0.5862	0.4138	0.4828	0.5517	0.6552	0.6207	0.6207

Conclusioni. Le analisi condotte hanno mostrato come nessuna delle metodologie analizzate sia uniformemente superiore alle altre. Ciò è risultato vero sia rispetto al tipo di serie, sia rispetto all'orizzonte temporale. Per tale motivo, l'utilizzazione di algoritmi per la combinazione delle previsioni risulta molto interessante. Infatti, come si è potuto valutare nelle applicazioni, il combinatore presenta spesso i migliori risultati e quindi può fornire un punto di riferimento oggettivo per chi deve prendere delle decisioni operative.

Bibliografia

- Clemen, R. T. (1989), «Combining Forecasts: A Review and Annotated Bibliography», *International Journal of Forecasting*, 5, 559-83.
- Granger, C.W.J. (1989), «Invited review: Combining Forecasts - 20 years later», *Journal of Forecasting*, 8, 167-173.
- Granger, C.W.J. e Newbold(1989), *Forecasting Economic Time Series*, Academic Press, Inc.

Granger, C.W.J. e Ramanathan R. (1984), Improved Methods of Combining Forecasts, *Journal of Forecasting*, 3, 197-204.

NOTE

¹ Si vedano ad esempio Clemen (1989) e Granger (1989).

² La revisione riguarderà solamente i valori previsti. Infatti con il procedere della settimana, i valori previsti vengono sostituiti dai valori effettivi, che ovviamente non subiscono alcuna revisione.

³ Per la sua formulazione di veda l'Appendice.