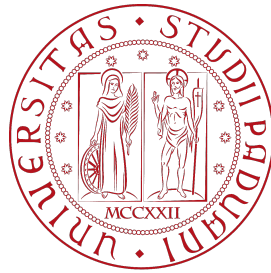


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in

Scienze Statistiche



Bootstrap ibrido e combinazioni bootstrap per la stima del parametro di memoria lunga

Relatrice: Prof.ssa Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureanda: Margherita Palomba
Matricola n. 2073400

Anno Accademico 2023/2024

Indice

Elenco degli acronimi	4
Introduzione	7
1 Combinazioni di previsioni	9
1.1 Stato dell'arte	12
1.1.1 Combinazioni semplici	13
1.1.2 Combinazioni lineari	15
1.1.3 Combinazioni non lineari	17
1.2 Come combinare le previsioni	18
2 Bootstrap per serie storiche a memoria lunga	21
2.1 Serie storiche a memoria lunga	22
2.2 Metodi bootstrap nel dominio del tempo	23
2.2.1 Bootstap Parametrico	23
2.2.2 Wild bootstrap	24
2.2.3 Sieve Bootstrap	26
2.2.4 Block Bootstrap	27
2.2.5 Pre-filtered bootstrap	29
3 Le combinazioni bootstrap	31
3.1 Bootstrap ibrido	31
3.1.1 Combinazioni semplici	32
3.1.2 Combinazioni lineari	34
3.2 Combinazione di stime bootstrap	37
3.2.1 Combinazioni semplici	38

3.2.2	Combinazioni lineari	39
4	Studio Monte Carlo	43
4.1	Bootstrap ibrido	46
4.2	Combinazione di stime bootstrap	57
5	Applicazione su dataset reali	69
5.1	Livelli minimi del fiume Nilo	69
5.2	Rendimenti assoluti dei prezzi azionari di Microsoft	72
	Conclusioni	81
	Bibliografia	83

Elenco degli acronimi

ACF	Funzione di autocorrelazione
ADF	Test aumentato di Dickey-Fuller
CBB	Circular Block Bootstrap
CI	Intervallo di confidenza
GPH	Stimatore del log-periodogramma
LWE	Stimatore di Whittle locale
MBB	Moving Block Bootstrap
MLE	Stimatore di massima verosimiglianza
MAE	Errore medio assoluto
MASE	Errore medio assoluto acalato
MSE	Errore quadratico medio
NBB	Non-overlapping Block Bootstrap
OLS	Minimi quadrati ordinari
PACF	Funzione di autocorrelazione parziale
PB	Parametric Bootstrap
PCR	Regressione sulle componenti principali
PFSB	Pre-Filtered Sieve Bootstrap

PFMBB	Pre-Filtered Moving Block Bootstrap
PGD	Processo generatore dei dati
RMSD	Radice della deviazione quadratica media
RMSE	Radice dell'errore quadratico medio
SB	Sieve Bootstrap
sMAPE	Errore percentuale medio assoluto simmetrico
WB	Wild Bootstrap

Introduzione

In ambito previsivo è ormai ampiamente diffusa la prassi di combinare più previsioni per una serie storica, al fine di aumentare l'accuratezza della previsione finale. Questo approccio, integrando informazioni provenienti da diverse fonti, elimina la necessità di selezionare una singola previsione, diminuendo il rischio associato a questa scelta.

Ispirandosi alle metodologie e ai progressi derivati dalla letteratura nell'ambito della combinazione di previsioni, questo lavoro ha l'obiettivo di estendere tali conoscenze alle strategie bootstrap ed esaminare se la combinazione di diversi metodi bootstrap possa migliorare le prestazioni degli stimatori del parametro di memoria lunga per serie storiche. A nostra conoscenza, finora in letteratura non esiste nulla a tal proposito. Applicare le strategie bootstrap alle serie storiche a memoria lunga implica la necessità di gestire la dipendenza presente all'interno del campione e, purtroppo, la forte persistenza della memoria nel tempo può rendere inefficaci le tecniche convenzionali. Inoltre, la scelta del metodo bootstrap più adatto per una specifica serie temporale può essere complicata a causa della varietà di tecniche disponibili e delle diverse caratteristiche delle serie stesse. Per questo motivo, la combinazione di più metodi bootstrap, analogamente a quanto avviene nella combinazione di previsioni, può rappresentare una soluzione efficace, in grado di sfruttare al meglio le diverse informazioni contenute nei dati. In particolare, vengono considerati e discussi alcuni metodi bootstrap nel dominio del tempo quali: bootstrap parametrico, *wild* bootstrap, *sieve* bootstrap, *moving block* bootstrap, *circular block* bootstrap, *pre-filtered sieve* bootstrap e *pre-filtered moving block* bootstrap.

Questo lavoro si concentra sulla proposta originale di due strategie di combinazione: la prima prevede la combinazione di campioni bootstrap ottenuti da differenti metodi, in una sorta di bootstrap ibrido, mentre la seconda si basa sulla combinazione diretta delle stime prodotte dai diversi approcci bootstrap tradizionali. Seguendo la letteratura sulla combinazione di previsioni, inoltre, vengono considerati, per la combinazione dei campioni o delle stime bootstrap, sia schemi di combinazione semplici, che combinazioni lineari i cui pesi sono basati

su diversi schemi di ponderazione.

L'efficacia delle metodologie proposte viene valutata mediante studi di simulazione Monte Carlo, utilizzando diverse numerosità campionarie e diversi valori per il parametro di memoria lunga, e due applicazioni su dati reali.

La trattazione si sviluppa come segue. Nel Capitolo 1 viene presentata una rassegna della letteratura esistente nell'ambito delle combinazioni di previsioni, di cui vengono discussi vantaggi e criticità. Nel Capitolo 2 vengono presentati i vari metodi bootstrap presi in considerazione per le serie storiche a memoria lunga e ne vengono approfonditi gli algoritmi. Il Capitolo 3 descrive le due strategie di combinazione proposte, bootstrap ibrido e combinazione di stime bootstrap, esplorando sia le combinazioni semplici che quelle lineari. Nel Capitolo 4 vengono presentati i risultati degli esperimenti Monte Carlo e, nel Capitolo 5, quelli delle applicazioni su serie storiche reali. Nel Capitolo finale vengono, quindi, fornite delle osservazioni conclusive per evidenziare i risultati ottenuti nel corso del lavoro e presentati i possibili sviluppi futuri.

Capitolo 1

Combinazioni di previsioni

Negli ultimi anni, la letteratura riguardante le combinazioni di previsioni per serie storiche ha conosciuto un notevole sviluppo nella comunità degli analisti previsivi e questa pratica è diventata una parte fondamentale della ricerca e dell'attività previsiva. Tramite la combinazione di più previsioni generate da modelli diversi per una serie temporale specifica è, infatti, possibile migliorare l'accuratezza della previsione stessa, grazie all'integrazione di informazioni provenienti dalle previsioni individuali, evitando la necessità di identificare una singola previsione "migliore".

I vantaggi derivanti dalla combinazione di previsioni, infatti, derivano principalmente da due considerazioni. In primo luogo, il processo generatore di dati (PGD) di una serie storica difficilmente è noto a priori a causa della possibile presenza di trend varianti nel tempo, variazioni nella stagionalità e *breaks* strutturali (Clements e Hendry, 1998). Determinare quale metodo di previsione sia il migliore tra quelli disponibili può essere, quindi, fuorviante e può portare a previsioni con varianza elevata, a causa di almeno tre fonti di incertezza: incertezza nei dati, incertezza nei parametri e incertezza nel modello (Kourentzes et al., 2019; Petropoulos et al., 2018). Dato questo aspetto, è intuibile che, nella pratica, il vantaggio di combinare le previsioni è quello di avere un rischio meno elevato rispetto a selezionare un metodo di previsione individuale, nonostante la migliore tra le combinazioni possa non ottenere *performance* superiori a quelle relative alla migliore previsione individuale (Hibon e Evgeniou, 2005). In secondo luogo, i diversi predittori potrebbero portare all'integrazione di informazioni predittive diverse tra loro e non completamente sovrapponibili, e questa condivisione può migliorare le *performance* previsive anche nel caso in cui alcuni di questi singoli metodi previsivi (definiti

come “*weak learners*”) non riescano a funzionare adeguatamente da soli. Ad esempio, nel caso di *breaks* strutturali o altre instabilità, integrare informazioni provenienti da modelli con gradi di adattabilità variabili può mitigare l’incertezza dovuta a queste caratteristiche della serie ed aiutare a comprendere il successo empirico delle combinazioni predittive (Timmermann, 2006; Rossi, 2021).

Nell’ambito delle combinazioni di previsioni, gli aspetti principali di cui tenere conto sono quindi l’accuratezza e la diversità dei singoli metodi previsivi ed il rischio e la robustezza delle combinazioni (Lichtendahl e Winkler, 2020). Nel contesto delle previsioni, l’accuratezza riveste un ruolo fondamentale. Integrare in una combinazione un metodo che, di per sé, non è molto preciso potrebbe non migliorare la previsione complessiva. Di conseguenza, quando si dispone di previsioni ottenute con diversi metodi, ha senso concentrarsi su quelli più accurati ed escludere quelli meno performanti. In letteratura, il processo di selezione di un sottoinsieme di previsioni è noto come *pooling* ed è dimostrato che lo stesso può migliorare l’accuratezza complessiva della combinazione (Kourentzes et al., 2019). In generale, aggiungere ulteriori previsioni a questo sottoinsieme potrebbe portare a risultati peggiori, anche quando le previsioni sono accurate. Hora (2004) suggerisce di utilizzare, per le combinazioni, tra le cinque e le dieci previsioni individuali.

Anche la diversità, come anticipato, riveste un ruolo importante. Un aumento della diversità tra i metodi di previsione può migliorare l’accuratezza della loro combinazione. La diversità tra i metodi è di solito misurata in termini di correlazione tra i loro errori di previsione, dove correlazioni più basse indicano una maggiore diversità delle previsioni individuali. Le alte correlazioni indicano che le previsioni cadono generalmente dalla stessa parte rispetto alla realizzazione, mentre i guadagni in accuratezza ricavati dalla combinazione sono massimi quando le previsioni cadono nelle parti opposte rispetto alla realizzazione, in questo caso, infatti, i loro errori di previsione tendono a cancellarsi l’un l’altro. Nella pratica, ciò accade raramente, poiché le previsioni vengono solitamente stimate basandosi sugli stessi insiemi di informazione e vengono utilizzati metodi previsivi simili. Un’accuratezza maggiore si può ottenere quando le correlazioni tra gli errori di previsione dei vari metodi presentano valori bassi o, meglio ancora, negativi (Larrick e Soll, 2006). Pertanto, la ricerca di metodi con distorsioni di segno opposto può essere considerata un modo alternativo per identificare metodi le cui previsioni cadano nelle parti opposte rispetto alla realizzazione. Quando due o più metodi previsivi sono molto correlati, includerne più di uno nella combinazione crea ridondanza, pertanto ha più senso

tenere solo il metodo con accuratezza più elevata per evitare che si possano creare problemi di multicollinearità.

Misure di accuratezza delle previsioni usate nell'ambito delle serie storiche sono, ad esempio, l'errore percentuale medio assoluto simmetrico (sMAPE) e l'errore medio assoluto scalato (MASE), utilizzati per stimare l'accuratezza media (Lichtendahl e Winkler, 2020). Nella scelta dei modelli previsivi da combinare è, però, necessario bilanciare il *trade-off* tra l'accuratezza media e la varianza dell'accuratezza. Questa varianza fornisce, infatti, un'indicazione del rischio dell'accuratezza ed è correlata alla probabilità di ottenere delle previsioni poco accurate per una determinata serie. Utilizzando i dati dalla prima *M-competition*, competizione con lo scopo di imparare dall'evidenza empirica sia come migliorare l'accuratezza delle previsioni, sia come l'utilizzo di tale apprendimento possa far avanzare la teoria e la pratica della previsione, Makridakis e Winkler (1983) hanno scoperto che i benefici delle combinazioni si manifestano più in termini di riduzione del rischio previsivo che in termini di miglioramento dell'accuratezza. Per comprendere meglio l'importanza della riduzione del rischio nelle previsioni, è fondamentale considerare come vengono utilizzati gli insiemi di osservazioni per sviluppare e valutare i modelli predittivi nelle serie storiche. L'insieme di *training* è la finestra temporale iniziale della serie storica, utilizzato per addestrare il modello, mentre l'insieme di validazione è la porzione successiva della serie, usata per testare le prestazioni del modello sulle osservazioni non incluse nell'insieme di *training*. Per garantire una valutazione accurata dei modelli previsivi, l'insieme di *training* e quello di validazione non devono sovrapporsi temporalmente: il periodo di validazione inizia solo dopo che il periodo di training è concluso. Questo approccio sequenziale implica che un modello possa rilevare un *pattern*, come un trend o una componente stagionale, nell'insieme di *training*, ma tale *pattern* potrebbe non mantenere la sua validità o rilevanza nell'insieme di validazione, con il rischio di ridurre l'accuratezza delle previsioni. Per cui, quando si ha un'instabilità nel processo, è necessario che un metodo sia robusto all'evoluzione di questo *pattern*. I metodi di combinazione semplici e poco sensibili rispetto ai cambiamenti sono stati spesso utilizzati nella pratica poiché, in media, si comportano molto bene e, in termini di riduzione del rischio, dimostrano una maggiore robustezza (Winkler et al., 2019). Metodi più sofisticati, o anche metodi leggermente meno semplici come le medie ponderate, possono ottenere risultati migliori rispetto ai metodi più semplici, ma possono anche produrre previsioni che si comportano in maniera inadeguata (come è stato il caso in alcune delle *M-competition*).

L'idea originale di combinare diverse previsioni individuali risale a Francis Galton, studioso britannico e, tra le tante cose, statistico che, nel 1906, partecipò ad una competizione in

cui veniva chiesto di indovinare il peso di un bue e notò che la media delle 787 stime del peso si avvicinava notevolmente al reale peso del bue (Surowiecki, 2005). Il concetto della combinazione previsiva è stato poi reso popolare da Bates e Granger (1969). Nel loro lavoro, gli Autori proponevano di impiegare il criterio di minimizzazione dell'errore quadratico medio (MSE) per stimare i pesi ottimali per la combinazione tramite la matrice di varianze e covarianze degli errori di previsione. Dopo il loro contributo, sono state numerose le proposte che sono seguite, progressivamente più sofisticate. Il guadagno dato dalla combinazione delle previsioni, infatti, non si basa solo sulla qualità delle previsioni individuali che sono tra loro combinate, ma anche sulla stima dei pesi della combinazione assegnati a ciascuna previsione (Cang e Yu, 2014). Le combinazioni che possono essere utilizzate sono molteplici e possono andare da metodi più semplici che evitano la stima dei pesi (Petropoulos e Svetunkov, 2020) a metodi più sofisticati che utilizzano pesi diversi per ogni modello individuale (Wang et al., 2022). Di conseguenza, le combinazioni di previsioni possono essere lineari o non lineari, statiche o variabili nel tempo, specifiche della serie o basate sul trasferimento di conoscenze tra serie diverse, e possono ignorare o tenere conto delle correlazioni tra le singole previsioni. Tuttavia, con l'aumentare della popolarità dei metodi di combinazione basati sulle *performance* previsive, è emerso che i pesi stimati con approcci più sofisticati possono spesso deviare dall'obiettivo di trovare pesi teoricamente ottimali a causa degli errori di stima e dell'incertezza a questi associata (e.g., Claeskens et al., 2016; Smith e Wallis, 2009). Questo effetto, insieme ad altri fattori come la discontinuità strutturale e l'aggiunta di nuove informazioni (Lahiri et al., 2017; Qian et al., 2019), contribuisce al fenomeno noto come "*forecast combination puzzle*" (Hendry e Clements, 2004; Stock e Watson, 2004), secondo il quale è possibile osservare, nella pratica, come la semplice media aritmetica ed alcune sue varianti superino le alternative più complesse in termini di *performance*.

1.1 Stato dell'arte

Nel seguito del capitolo si vuole approfondire lo stato dell'arte riguardo le combinazioni di previsioni puntuali.

1.1.1 Combinazioni semplici

Come scritto in precedenza, in letteratura è stato più volte rimarcato come gli schemi di combinazione semplici siano spesso difficili da superare in termini di *performance* (Clemen, 1989, Fischer e Harvey, 1999). Le medie aritmetiche semplici, che ignorano le informazioni passate sulla precisione delle singole previsioni e sulle correlazioni tra gli errori di previsione, funzionano ragionevolmente bene rispetto a schemi di combinazione più sofisticati. La semplice media aritmetica, che assegna a tutte le previsioni pesi uguali, si distingue come la regola di combinazione più popolare e sorprendentemente robusta (Genre et al., 2013), oltre ad essere di facile implementazione. Un'altra importante osservazione, fatta in seguito alla già citata *M-competition* da Makridakis e Winkler (1983), fu che l'inclusione di più metodi previsivi individuali all'interno di uno schema di media semplice aiutasse a migliorare l'accuratezza della previsione e riducesse la variabilità associata alla selezione del singolo metodo.

Altre strategie di combinazione, derivate direttamente dall'utilizzo della media e basate sulla distribuzione empirica delle previsioni individuali ottenute da diversi modelli previsivi, sono la mediana, la moda, la media troncata e la media *winsorizzata* (vedi Chan et al., 1999; Genre et al., 2013; Stock e Watson, 2004). Tali combinazioni sono apprezzate per la loro robustezza e risultano essere meno sensibili alle previsioni estreme rispetto alla media semplice (Lichtendahl e Winkler, 2020). Tuttavia, in letteratura, non esiste un parere unanime su quale sia la migliore strategia per la combinazione di previsioni puntuali da utilizzare tra la media e la mediana. Nello specifico, lo studio di McNees (1992) non ha evidenziato alcuna differenza significativa tra le due strategie, mentre i risultati di Stock e Watson (2004) supportano la media e quelli di Agnew (1985) la mediana.

Nonostante, in un'epoca dominata dai *big data*, le combinazioni semplici possano sembrare superate da approcci ed algoritmi di *machine learning* più complessi, i risultati recenti della *M4-competition* (Makridakis et al., 2020) dimostrano che queste metodologie continuano ad ottenere risultati previsivi competitivi. Innanzitutto, i risultati della competizione hanno (ri)confermato i vantaggi della combinazione di più previsioni rispetto all'utilizzo di previsioni individuali, infatti dei 17 metodi più accurati, 12 erano combinazioni di approcci prevalentemente statistici. In particolare, una combinazione semplice con pesi uguali si è classificata al terzo posto per le serie temporali annuali (Shaub, 2020), mentre una combinazione basata sulla mediana di quattro modelli di previsione ha ottenuto il sesto posto per le previsioni puntuali (Petropoulos e Svetunkov, 2020). In generale, la media semplice sembra essere un *benchmark* che pochi schemi di combinazione riescono a superare in termini di *performance*, senza considerare anche

il fatto che una combinazione così semplice ha un costo computazionale minore e può essere implementata in modo più efficiente rispetto alle combinazioni alternative.

In uno studio recente, Gastinger et al. (2021) hanno esaminato le prestazioni previsive di un insieme di metodi di combinazione su molte serie temporali provenienti da fonti diverse e hanno scoperto che i metodi di combinazione vincenti differiscono a seconda del diverso contesto da cui provengono i dati, concludendo che non sia possibile determinare un metodo che sia il “migliore” per tutte le serie storiche. Allo stesso tempo, le strategie basate sulla media semplice mostrano, mediamente, un maggiore guadagno in termini di miglioramento dell’accuratezza rispetto ad altri metodi più complessi. Stock e Watson (2004), per questo fenomeno, hanno coniato il termine “*forecast combination puzzle*”: in teoria, i sistemi di ponderazione sofisticati dovrebbero fornire maggiori benefici rispetto alla media aritmetica per la combinazione delle previsioni, mentre empiricamente la media semplice è riuscita costantemente a dominare approcci più complicati. Timmermann (2006) ha osservato che il successo delle combinazioni semplici è attribuibile all’aumento dell’errore di stima dei pesi nelle combinazioni ponderate, poiché i metodi di combinazione semplice non richiedono la stima di tali pesi per la combinazione. Spesso, infatti, questo errore di stima arriva a sovrastare i benefici derivanti da una strategia di combinazione di previsioni potenzialmente ottimale. Elliott (2011) ha delineato i limiti potenziali dei guadagni derivanti dall’utilizzo dei pesi ottimali rispetto a quelli derivati dall’utilizzo di pesi uguali e ha dimostrato che tali guadagni sono spesso troppo esigui per poter bilanciare l’errore di stima, fornendo così una spiegazione supplementare del fenomeno dal punto di vista degli errori di stima. Zischke et al. (2022), invece di concentrarsi sull’impatto della stima dei pesi di combinazione, hanno esplorato le implicazioni dovute alla variabilità campionaria nelle combinazioni di previsioni. Tali Autori hanno dimostrato che, asintoticamente, la variabilità campionaria delle misure di *performance* delle combinazioni è dovuta interamente alla variabilità delle stime dei modelli previsivi individuali e che, quindi, le stime dei pesi non aggiungono alcuna distorsione o variabilità additiva alle misure di *performance* della combinazione. Questa scoperta è opposta a quella di Claeskens et al. (2016), che hanno invece dimostrato che la stima dei pesi per la ponderazione introduce sempre, anche se a volte in minore misura, varianza additiva nella combinazione di previsioni. Questi risultati implicano che, quando i pesi della combinazione sono teoricamente equivalenti, cioè non differiscono tra loro in maniera significativa, ci sarà una piccola differenza tra le *performance* di una combinazione più sofisticata e quelle di una combinazione con pesi ugualmente ponderati.

1.1.2 Combinazioni lineari

Nonostante la semplicità e le buone *performance* delle combinazioni semplici, può avere più senso assegnare un peso maggiore ai metodi previsivi più accurati. Data l'informazione fino al tempo T , il vettore N -dimensionale di previsioni h passi in avanti è, dove N è il numero di previsioni combinate, $\hat{\mathbf{y}}_{T+h|T} = (\hat{y}_{T+h|T,1}, \hat{y}_{T+h|T,2}, \dots, \hat{y}_{T+h|T,N})'$, mentre la combinazione di queste previsioni h passi in avanti è $C(\hat{\mathbf{y}}_{T+h|T}; \mathbf{w}_{T+h|T})$, dove $\mathbf{w}_{T+h|T}$ è il vettore N -dimensionale dei pesi della combinazione. Tipicamente, la combinazione di previsioni è costruita come una combinazione lineare delle previsioni individuali, e può essere quindi scritta come $\tilde{y}_{T+h|T} = \mathbf{w}'_{T+h|T} \hat{\mathbf{y}}_{T+h|T}$, dove $\mathbf{w}_{T+h|T} = (w_{T+h|T,1}, w_{T+h|T,2}, \dots, w_{T+h|T,N})'$ è un vettore N -dimensionale dei pesi della combinazione lineare assegnati alle N previsioni individuali. I pesi delle combinazioni lineari in ambito previsivo possono essere divisi nelle seguenti categorie:

- **Pesi ottimali.** Bates e Granger (1969) proposero per primi un metodo per trovare i pesi "ottimali", minimizzando la varianza dell'errore della combinazione di due metodi previsivi. Newbold e Granger (1974) hanno poi esteso il metodo alla combinazione di più di due metodi previsivi. La stima dei pesi così determinati è tale che la combinazione di previsioni $\tilde{y}_{T+h|T}$ non può avere una varianza dell'errore maggiore della minima varianza degli errori delle N previsioni individuali. La perdita assunta da queste combinazioni è quadratica e simmetrica, mentre Elliott e Timmermann (2004) hanno dimostrato che i pesi ottimali della combinazione dipendono fortemente dal grado di asimmetria della funzione di perdita e della reale distribuzione dell'errore di previsione. Successivamente, Patton e Timmermann (2007) hanno generalizzato le proprietà delle previsioni ottimali così da considerare funzioni di perdita asimmetriche e PGD non lineari.
- **Pesi basati sulla regressione.** Granger e Ramanathan (1984) sono stati i primi a suggerire di stimare i pesi per la combinazione di previsioni tramite minimi quadrati ordinari (OLS). Nel *framework* di regressione lineare da loro proposto, la variabile risposta è costituita dal vettore delle osservazioni passate ed ogni vettore di previsioni individuali, stimate con un determinato modello previsivo, costituisce una variabile esplicativa. In particolare, è possibile considerare tre diversi modelli di regressione: (i) modello di regressione lineare con intercetta; (ii) modello di regressione lineare senza intercetta; (iii) modello di regressione lineare senza intercetta con il vincolo che i pesi della regressione sommino a 1 (in questo caso i risultati sono uguali a quelli ottenuti da Bates e Granger (1969)). Un'altra possibile alternativa, che si è dimostrata essere almeno tanto robusta

e accurata quanto la media semplice e che porta a risultati migliori rispetto a quelli di altre combinazioni basate sulla regressione (Aksu e Gunter, 1992), è quella di usare una regressione OLS dove i pesi sono vincolati ad essere non negativi. Imponendo entrambe le restrizioni (cioè che i pesi sommino a 1 e siano non negativi) si ottiene inoltre un caso speciale della regressione LASSO (Conflitti et al., 2015). Sono state sviluppate anche combinazioni di previsioni con pesi che non rimangono costanti ma variano, così da seguire i cambiamenti strutturali delle previsioni. I pesi in questo caso vengono descritti come funzioni (polinomiali) deterministiche non lineari oppure è possibile che comprendano variazioni casuali (vedi Diebold e Pauly, 1987; Deutsch et al., 1994; Terui e van Dijk, 2002). Chan et al. (1999) hanno però trovato nelle combinazioni OLS un limite, infatti queste combinazioni hanno delle pessime *performance* quando il numero di previsioni da combinare è elevato. Tra le soluzioni a questo problema, quelle più comunemente usate sono i metodi basati sui fattori, come ad esempio la regressione sulle componenti principali (PCR), che dovrebbero eliminare la multicollinearità (Stock e Watson, 2004). Approcci alternativi sono stati invece proposti da Diebold e Shin (2019), i quali hanno introdotto una procedura basata sulla regressione LASSO, e da Aiolfi e Timmermann (2006), i quali hanno utilizzato l'algoritmo di clustering *k-means* basato sulla *performance* storica delle previsioni individuali.

- **Pesi basati sulla *performance*.** I lavori di Newbold e Granger (1974) e di Makridakis e Winkler (1983) hanno dimostrato che le combinazioni che ignorano la correlazione tra le previsioni individuali hanno più successo rispetto a quelle che la considerano nel calcolo dei pesi e hanno esteso le cinque procedure suggerite da Bates e Granger (1969) per stimare le combinazioni dei pesi ai casi con più serie temporali, più previsioni individuali e orizzonti previsivi multipli. Tuttavia, gli schemi proposti tendono a sovrastimare la varianza delle stime dei pesi di ponderazione e a non performare adeguatamente quando il PGD è stazionario in covarianza (Timmermann, 2006). Un più ampio insieme di schemi di pesi per le combinazioni previsive è stato esaminato e sviluppato in una serie di studi. Ad esempio, Stock e Watson (1999) hanno proposto dei pesi per le previsioni individuali che fossero proporzionali alla *k*-esima potenza del loro MSE, mentre Stock e Watson (2004) hanno usato uno schema con i pesi inversamente proporzionali agli errori ridotti esponenzialmente. Molti di questi schemi di ponderazione, per studiare la *performance* pseudo-*out-of-sample* e stimare i pesi adeguatamente, adottano le finestre temporali mobili o ricorsive, al fine di gestire i cambiamenti strutturali nei dati. In questo

caso, è importante che la lunghezza della finestra non sia troppo breve, poiché ciò potrebbe rendere le stime dei pesi troppo instabili o imprecise (Baumeister e Kilian, 2015). Un'altra alternativa, più robusta alla presenza di valori anomali rispetto all'utilizzo di pesi basati direttamente sugli errori delle previsioni passate, è quella di basare i pesi sul "ranking" delle previsioni individuali (vedi Aiolfi e Timmermann, 2006 e Andrawis et al., 2011), tuttavia questa classe di metodi di combinazione ha il limite che i pesi possono appartenere solo ad un insieme limitato di valori.

- **Pesi basati sui criteri d'informazione automatica.** Burnham e Anderson (2002) hanno proposto un approccio nel quale la combinazione di previsioni ottenute da diversi modelli previsivi è basata su criteri d'informazione, quali AIC (Akaike, 1974), AIC_c (Sugiura, 1978) e BIC (Schwarz, 1978), per mitigare il rischio di selezione di un singolo modello. Nello specifico, poiché il criterio AIC stima la distanza di Kullback-Leibler tra un modello ed il vero PGD, i pesi basati su questa misura possono essere interpretati come la probabilità che un dato modello performi meglio nell'approssimare il PGD, date le informazioni storiche possedute (Kolassa, 2011). Le ottime *performance* delle combinazioni i cui pesi sono basati sui criteri d'informazione sono state supportate in numerosi studi.
- **Pesi bayesiani.** Bunn (1975) suggerì l'idea di combinazioni bayesiane basate sulla probabilità che ciascun modello previsionale performasse al meglio in ogni occasione data, mentre Öller (1978) presentò un altro approccio basato su pesi proporzionali alla valutazione soggettiva, data dall'analista, dell'abilità previsiva dei vari modelli. Un filone di ricerca differente ha anche sostenuto l'integrazione di informazioni precedenti nella stima dei pesi della combinazione, ma con i pesi che vengono ridotti verso una media a priori in un *framework* di combinazione basato sulla regressione (Newbold e Harvey, 2004). Questi metodi possono essere considerati delle transizioni razionali tra una stima soggettiva ed una basata sui dati per i pesi della combinazione.

1.1.3 Combinazioni non lineari

Gli approcci basati su combinazioni lineari assumono implicitamente una dipendenza lineare tra le previsioni individuali e la variabile di interesse, ma ciò potrebbe non portare alla miglior previsione combinata (Ming Shi et al., 1999). Ciò avviene specialmente quando le previsioni individuali sono generate da modelli non lineari o se la relazione tra i membri della combinazione

e la miglior previsione è non lineare (Babikir e Mwambi, 2016). Sono due i tipi di non-linearità che possono essere incorporati in una combinazione previsiva: funzioni non lineari delle previsioni individuali e pesi della combinazione non lineari (Timmermann, 2006).

Un approccio particolarmente utilizzato per la stima delle componenti non lineari è quello delle reti neurali, nonostante sia computazionalmente oneroso e possa portare a scarse *performance* predittive (vedi Donaldson e Kamstra, 1996; Harrald e Kamstra, 1997; Krasnopolsky e Lin, 2012). I risultati empirici derivati da questi studi dimostrano che, generalmente, combinazioni non lineari hanno *performance* migliori rispetto a quelle delle tradizionali strategie di combinazione lineari. Tuttavia, l'evidenza empirica si basa su un numero ridotto di serie temporali, solitamente meno di dieci, potenzialmente selezionate per favorire questo risultato. Inoltre, questi metodi di combinazione non lineare presentano altri svantaggi, inclusa la mancata considerazione delle correlazioni tra gli errori di previsione, l'instabilità della stima dei pesi e la multicollinearità derivante dalla sovrapposizione negli insiemi di informazioni utilizzati per generare le previsioni individuali. Di conseguenza, è ancora necessario molto lavoro per valutare le prestazioni delle combinazioni non lineari rispetto a quelle lineari.

Alcuni ricercatori hanno cercato di costruire combinazioni non lineari includendo in una regressione lineare un termine non lineare aggiuntivo, così da gestire i casi in cui gli errori di previsione individuali sono correlati. In questo modo, il *framework* generale per le combinazioni lineari viene esteso per affrontare le non-linearità (vedi Freitas e Rodrigues, 2006 e Adhikari e Agrawal, 2012), tuttavia anche questi schemi hanno bisogno di essere ulteriormente approfonditi.

1.2 Come combinare le previsioni

Alla luce delle numerose metodologie presenti in letteratura riguardanti la combinazione di previsioni, è necessario tenere presenti alcune considerazioni (Wang et al., 2023).

Nella vastità degli approcci proposti, la media semplice continua ad ottenere *performance* più che competitive rispetto a metodologie più sofisticate per ottenere i pesi della combinazione, diventando così un *benchmark* difficile da superare. In particolare, le combinazioni semplici risultano essere l'alternativa migliore nel caso in cui si abbia accesso ad un limitato numero di dati storici, poiché alternative più complesse risultano essere poco robuste per la stima dei pesi; in questo caso la media troncata e *winsorizzata* possono essere adottate per eliminare previsioni estreme, poiché risultano essere ancora meno sensibili a queste rispetto alla media. Un altro caso in cui queste combinazioni semplici sembrano costituire una buona regola è quello

in cui ci siano cambiamenti strutturali che possono impattare considerevolmente su approcci di combinazioni più sofisticati.

I modelli di previsione ed i metodi di combinazione delle previsioni hanno conosciuto una rapida crescita in termini di dimensione e complessità. Tuttavia, i risultati empirici rimangono ambigui e non vi è una chiara evidenza che la complessità migliori in modo sistematico l'accuratezza delle previsioni. Come suggeriscono Zellner et al. (2002), sarebbe meglio adottare combinazioni sofisticate, ma semplici, al fine di bilanciare il *trade-off* tra i vantaggi derivati dall'adattamento dei pesi per diversi modelli individuali e l'instabilità dei pesi stimati in questi schemi di pesatura più complessi. Inoltre, è consigliabile che gli analisti previsivi analizzino bene i dati prima di identificare la strategia di combinazione e che scelgano quindi la regola di combinazione più adatta allo specifico problema previsivo.

La scelta delle previsioni individuali costituisce la base per una combinazione efficace. Queste previsioni possono provenire da modelli statistici o di *machine learning* basati su dati osservati, oppure possono essere fornite da esperti. Le evidenze suggeriscono che il futuro della previsione risiede nella combinazione di previsioni statistiche, integrando anche il giudizio umano (Makridakis et al., 2020). Con una vasta gamma di previsioni disponibili, selezionare il sottoinsieme più rilevante diventa cruciale per migliorare l'accuratezza delle previsioni e ridurre i costi computazionali. Considerazioni chiave nella selezione delle previsioni includono accuratezza, affidabilità, diversità e calibrazione. Se si dispone di molte previsioni individuali, la regressione sulle componenti principali (PCR) e la strategia di *clustering* potrebbero essere utili per ridurre gli errori di stima e risolvere il problema della multicollinearità riducendo il numero di parametri da stimare. La correlazione presente tra le previsioni individuali, inoltre, risulta ancora essere una criticità, in quanto può avere un impatto sull'efficacia della combinazione.

Mentre i benefici delle combinazioni lineari delle previsioni sono ben apprezzati nella letteratura previsiva, minore attenzione è stata data agli schemi di combinazione non lineari per modellare la dipendenza non lineare tra le previsioni individuali. Ciò è probabilmente dovuto alla mancanza di fondamenta teoriche a sostegno di tali metodologie e ad un limitato numero di studi degli stessi che hanno ottenuto successo. Recentemente, invece, molti studi hanno confermato i benefici dell'utilizzo di informazioni da serie multiple per studiare *pattern* comuni tra le serie, così da facilitare la determinazione dei pesi delle combinazioni e sfruttare i vantaggi del *cross-learning*. Le prove del potenziale del *cross-learning* sono state largamente osservate nelle competizioni (Makridakis et al., 2020, 2022) e negli studi empirici e si crede che in questo risieda il futuro delle combinazioni previsive.

Capitolo 2

Bootstrap per serie storiche a memoria lunga

Il bootstrap è un metodo di ricampionamento, computazionalmente intensivo, che fornisce soluzioni a una vasta gamma di problemi di inferenza statistica senza fare ipotesi strutturali rigorose sul processo casuale che genera i dati (Lahiri, 2013). Dopo la sua introduzione da parte di Efron (1979), il bootstrap è stato utilizzato per affrontare numerosi problemi statistici di differente natura. In molti casi, il bootstrap si è dimostrato essere più efficace rispetto alle metodologie tradizionali, sia per problemi standard che per quelli più complessi, dove gli approcci convenzionali non hanno fornito soluzioni adeguate. Infatti, i metodi bootstrap riescono a fornire approssimazioni accurate delle distribuzioni delle statistiche, delle probabilità di copertura degli intervalli di confidenza e delle probabilità di rifiuto dei test. In alcuni casi, tali approssimazioni risultano essere anche più accurate rispetto a quelle determinate dalla teoria della distribuzione asintotica di primo ordine degli stimatori.

Originariamente sviluppato per osservazioni indipendenti (Efron, 1979), il bootstrap è stato successivamente adattato per gestire la dipendenza temporale tipica delle serie storiche (vedi, per esempio, Politis e Romano, 1994). Tuttavia, la forte dipendenza nelle serie a memoria lunga richiede, a volte, l'adozione di tecniche specifiche per riuscire ad utilizzare una strategia di ricampionamento valida.

Le strategie bootstrap tradizionali per serie storiche possono essere suddivise in due categorie. La prima categoria include i metodi che campionano i residui ottenuti da un modello parametrico, che si assume essere correttamente specificato e i cui parametri sono stimati attraverso i dati osservati. In questo caso si assume, quindi, che i residui siano tra loro indipendenti

ed identicamente distribuiti (Andrews et al., 2006). La seconda categoria di strategie bootstrap include, invece, le tecniche bootstrap che non si basano fortemente su restrizione parametriche, le più usate sono il *sieve* bootstrap (Bühlmann, 1997) ed il *block* bootstrap (Kunsch, 1989). In entrambe le categorie, gli approcci hanno come elemento comune che il campionamento finale viene implementato su quantità approssimativamente indipendenti, che possono essere residui, come nel caso della prima categoria e del *sieve* bootstrap, o blocchi, come nel caso del *block* bootstrap. L'assunzione che le quantità campionate siano indipendenti è soddisfatta nelle serie storiche a memoria corta, mentre, nelle serie a memoria lunga, dove la dipendenza rimane significativa anche quando le osservazioni sono distanti, ciò difficilmente avviene. Un metodo alternativo per gestire la dipendenza seriale consiste nell'applicare un filtro preliminare alle osservazioni appartenenti alla serie storica, così da ottenere residui o blocchi approssimativamente indipendenti da campionare successivamente (Poskitt et al., 2015).

Le tecniche bootstrap menzionate finora sono sviluppate nel dominio del tempo, ma in letteratura sono di particolare interesse anche quelle appartenenti al dominio delle frequenze, basate sul ricampionamento delle ordinate del periodogramma, studentizzato con uno stimatore consistente della densità spettrale (Franke e Hardle, 1992). Tuttavia, nonostante la stima della funzione di densità spettrale sia stata ampiamente studiata per le serie storiche a memoria corta, questa risulta essere un argomento meno esplorato per quanto riguarda i processi a memoria lunga, dove le tecniche tradizionali solitamente falliscono (vedi Arteche, 2015).

Una panoramica delle metodologie bootstrap maggiormente utilizzate per le serie storiche a memoria lunga è presente nel lavoro di Arteche (2024).

2.1 Serie storiche a memoria lunga

Si consideri il processo stocastico lineare della forma:

$$y_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j},$$

dove ϵ_t sono variabili i.i.d. di media 0 e varianza finita e a_j sono costanti con valore reale e tali che $\sum_{j=0}^{\infty} a_j^2 < \infty$, in tal caso il processo è stazionario. Le serie storiche a memoria lunga sono caratterizzate da una dipendenza significativa tra osservazioni molto distanti tra loro, in particolare le autocovarianze sono tali per cui:

$$\gamma_y(k) = \text{cov}(y_t, y_{t-k}) \sim Gk^{2d-1} \quad \text{per } k \rightarrow \infty, \quad (2.1)$$

dove G è una costante finita e d è il parametro di memoria lunga che governa la persistenza di y_t (Robinson, 2003). Se $0 < d < 1/2$ la serie è stazionaria e presenta memoria lunga, in questo caso le autocovarianze non sono sommabili. Equivalentemente la funzione di densità spettrale, si comporta come:

$$f_y(\lambda) \sim C\lambda^{-2d} \quad \text{per } \lambda \rightarrow 0, \quad (2.2)$$

con C costante positiva. Il concetto di memoria lunga è solitamente collegato al concetto di integrazione frazionaria, nonostante tali concetti non siano necessariamente equivalenti. Si dice che il processo è integrato di ordine d (possibilmente frazionario), scritto come $Y_t \sim I(d)$, se $(1-L)^d Y_t = u_t$, dove L è l'operatore ritardo ($L^k Y_t = Y_{t-k}$) e $u_t \sim I(0)$. Se $Y_t \sim I(d)$ allora f_y soddisfa la 2.2. Tra i processi $I(d)$, i più utilizzati sono gli *ARMA* Frazionalmente Integrati (*ARFIMA*), che assumono che u_t sia un processo *ARMA* stazionario e invertibile. I processi *ARFIMA*(p, d, q) sono quindi definiti come

$$\Phi_p(L)(1-L)^d y_t = \Theta_q(L)\epsilon_t, \quad (2.3)$$

dove $\Phi_p(L)$ e $\Theta_q(L)$ sono polinomi, rispettivamente di ordine p e q , con tutte le radici esterne alla circonferenza di raggio unitario e nessuna radice in comune. In questo caso $Y_t \sim I(d)$ con $u_t \sim \text{ARMA}(p, q)$. Nel caso di processi a memoria lunga non stazionari ($d \geq 1/2$) la parametrizzazione ed il trattamento matematico dei processi $I(d)$ sono più complicati.

Nel seguito, si considerano la serie storica a memoria lunga $y_t = \{y_1, \dots, y_T\}$, dove T è la numerosità campionaria, ed il corrispondente campione bootstrap $y_t^* = \{y_1^*, \dots, y_T^*\}$. Inoltre, viene notata come $S = S(y_1, \dots, y_T)$ la statistica di interesse, funzione di y_t , e come $S^* = S(y_1^*, \dots, y_T^*)$ la statistica calcolata sul campione bootstrap.

2.2 Metodi bootstrap nel dominio del tempo

In questo lavoro saranno considerati, e di seguito presentati, i metodi bootstrap per serie storiche a memoria lunga nel dominio del tempo la cui validità è supportata dalla teoria (per validità, in questo contesto, si intende che la funzione di distribuzione empirica dello stimatore ottenuta tramite bootstrap converge alla funzione di distribuzione asintotica basata sulle osservazioni originali; vedi Cavaliere e Georgiev, 2020).

2.2.1 Bootstrap Parametrico

I vantaggi teorici dei metodi bootstrap per le serie storiche a memoria lunga furono per primi illustrati da Andrews et al. (2006). L'assunzione alla base del modello proposto nel loro lavoro

implica che la serie storica a memoria lunga sia stazionaria e Gaussiana con media $\mu_0 \in R$ ignota e con funzione di densità spettrale che appartiene ad una famiglia parametrica e soddisfa:

$$f_y(\lambda) = O(|\lambda|^{-2d-\delta}) \quad \text{per } \lambda \rightarrow 0, \forall \delta > 0 \text{ e } d \in (0, 1/2).$$

La specificazione è leggermente più debole dell'usuale $f(\lambda) = O(|\lambda|^{-2d})$ ed è soddisfatta dai modelli *ARFIMA* dell'Equazione 2.3. Andrews et al. (2006) assumono, quindi, che la serie storica y_t sia la realizzazione di un processo *ARFIMA*(p, d, q) (dove gli ordini p e q sono stati precedentemente selezionati) e, tramite massima verosimiglianza (ML) o stima parametrica di Whittle, vengono stimati i parametri, la media e la varianza del modello parametrico. Il campione bootstrap y_t^* viene, quindi, calcolato a partire dal modello stimato usando, come residui, variabili normali simulate con media e varianza ottenute dal modello parametrico, e, come parametri, i valori stimati. Gli intervalli di confidenza empirici per il parametro di memoria lunga ricavati da questo metodo bootstrap risultano avere una copertura migliore rispetto agli intervalli di confidenza basati sulla distribuzione asintotica dello stimatore di massima verosimiglianza (MLE) o dello stimatore di Whittle dello stesso parametro. L'errore della probabilità di copertura degli intervalli di confidenza del parametro di memoria lunga ricavati tramite il bootstrap parametrico (PB), infatti, risulta essere più piccolo rispetto allo stesso errore basato sulla distribuzione asintotica di un fattore moltiplicativo di ordine $o(T^{-1/2} \ln T)$. Tuttavia, il metodo bootstrap parametrico presenta alcune limitazioni. In particolare, esso si basa sulla corretta e completa specificazione del modello, richiede l'assunzione di gaussianità e trascura altre possibili strategie basate sui residui. Inoltre, nel caso in cui il modello non fosse correttamente specificato, è possibile che anche le *performance* degli intervalli di confidenza bootstrap peggiorino severamente.

2.2.2 Wild bootstrap

Il *wild* bootstrap è una tecnica di ricampionamento proposta da White (1980) per affrontare l'eteroschedasticità nei modelli di regressione. In particolare, il *wild* bootstrap (WB) è stato concepito per mantenere la struttura di eteroschedasticità presente nei dati originali, consentendo, così, di ottenere stime più robuste delle varianze e, di conseguenza, inferenze più affidabili. Questo metodo si differenzia dagli altri metodi bootstrap, come il bootstrap parametrico, proprio per la sua capacità di adattarsi alla variabilità non costante degli errori.

Cavaliere et al. (2015, 2017) hanno esteso, poi, il *wild* bootstrap per l'analisi di serie storiche a

memoria lunga. Nel metodo bootstrap proposto da Cavaliere et al., vengono utilizzati i residui centrati di un modello $ARFIMA(p, d, q)$ stimato sui dati osservati, minimizzando la somma dei quadrati condizionale, senza che sia necessaria l'assunzione di gaussianità. Il vantaggio del *wild* bootstrap rispetto agli altri approcci per il ricampionamento è che, tramite questo metodo bootstrap, è possibile gestire sia l'eteroschedasticità condizionale che quella non condizionale. Inoltre, gli stessi Autori sono riusciti, tramite questo metodo, ad implementare un test per ipotesi lineari sui parametri di memoria corta e lunga del modello. Il *wild* bootstrap è stato, infatti, costruito con lo scopo di migliorare l'accuratezza dei test statistici nelle situazioni in cui la variabilità degli errori non è costante, superando i limiti delle tecniche asintotiche tradizionali. L'algoritmo implementato da Cavaliere et al. (2015) consiste in:

Step 1: Si consideri il processo stocastico frazionalmente integrato $\Delta_+^d Y_t = u_t$, con $u_t \sim ARMA(p, q)$ tale che $u_t = \Theta_q(L)\epsilon_t$, con $\epsilon_t \sim WN$, per $t = 1, \dots, T$. Si selezionino gli ordini p e q del modello $ARFIMA(p, d, q)$ e se ne stimino i coefficienti, tramite quasi massima verosimiglianza (QML), utilizzando le osservazioni della serie storica y_t . Si calcolino, quindi, i residui del modello $\hat{\epsilon}_t$, $t = 1, \dots, T$.

Step 2: Si calcolino i residui centrati

$$\tilde{\epsilon}_t = \hat{\epsilon}_t - \bar{\epsilon}, \quad \text{dove} \quad \bar{\epsilon} = \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t.$$

Si costruiscano, poi, gli errori bootstrap $\epsilon_t^* = \tilde{\epsilon}_t w_t$, dove w_t , $t = 1, \dots, T$, è una sequenza di variabili casuali i.i.d. con $E(w_t) = 0$, $E(w_t^2) = 1$ e $E(w_t^4) < \infty$.

Step 3: Si costruisca il campione bootstrap $y_t^* = \{y_1^*, \dots, y_t^*\}$ come

$$y_t^* = \Delta_+^{-d} u_t^*, \quad t = 1, \dots, T,$$

dove u_t^* è tale per cui $u_t^* = \Theta_q(L)\epsilon_t^*$, con $\epsilon_t^* = 0$ per $t \leq 0$.

Step 4: Si ottenga S^* come funzione di y_t^* , utilizzando la stessa procedura che si utilizzerebbe per ottenere la statistica di interesse S come funzione del campione originario y_t .

Step 5: Ripetere i passi 2, 3 e 4 B volte per generare B campioni bootstrap indipendenti $y_{t,b}^*$ e calcolare S_b^* per $b = 1, \dots, B$. Si approssima la distribuzione di S con la funzione di distribuzione bootstrap

$$FW_{S_B^*}(s) = \frac{1}{B} \sum_{b=1}^B I(S_b^* \leq s)$$

dove I è la funzione indicatrice.

2.2.3 Sieve Bootstrap

I metodi bootstrap semi-parametrici o non parametrici non richiedono una completa parametrizzazione del modello e la gaussianità non è un'assunzione essenziale dei modelli da loro usati. Il *sieve* bootstrap (SB), proposto da Bühlmann (1997), è una tecnica semi-parametrica basata sull'approssimazione del processo generatore di dati (PGD) con un modello autoregressivo di ordine p , $AR(p)$, dove p generalmente aumenta con l'aumentare della dimensione campionaria. In questo metodo bootstrap, vengono ricampionati i residui standardizzati ottenuti dal modello autoregressivo precedentemente stimato. Poskitt (2008) ha approfondito lo studio di questo metodo di ricampionamento mostrando che, tramite appropriati adattamenti, i risultati riguardanti le proprietà del *sieve* bootstrap possono essere estesi anche ai processi frazionalmente integrati e non-invertibili. La procedura proposta da Poskitt (2008) per i processi a memoria lunga stazionari è la seguente:

Step 1: Si consideri un processo stocastico $AR(p)$ e se ne stimino i coefficienti adattandolo alla serie storica osservata y_t , $t = 1, \dots, T$. Per motivi pratici, viene suggerito di usare l'algoritmo di Burg (Burg, 1975) per la stima dei coefficienti, mentre per la selezione di p si consiglia di procedere tramite minimizzazione di un criterio d'informazione, come ad esempio l'AIC (Sakamoto et al., 1986). Altre scelte sono possibili e disponibili in letteratura, sia per la selezione dell'ordine p che per la stima dei coefficienti autoregressivi. Si calcolino, quindi, i residui del modello nel seguente modo:

$$\hat{\epsilon}_t = y_t - \hat{\phi}_0 - \sum_{j=1}^p \hat{\phi}_j y_{t-j}, \quad t = 1, \dots, T$$

dove $\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p$ sono le stime dei coefficienti di autoregressione, utilizzando $y_{1-j} = y_{T-j+1}$, $j = 1, \dots, p$ come valori iniziali. Si costruiscano, poi, i residui standardizzati, per $t = 1, \dots, T$, come segue:

$$\hat{\epsilon}_{s,t} = \frac{\hat{\epsilon}_t - \bar{\epsilon}}{s_{\hat{\epsilon}}}, \quad \text{dove } \bar{\epsilon} = \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t \quad \text{e} \quad s_{\hat{\epsilon}} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{\epsilon}_t - \bar{\epsilon})^2}.$$

Step 2: Si consideri un campione casuale di valori i.i.d ϵ_t^+ , $t = 1, \dots, T$, derivati da una distribuzione la cui massa di probabilità è uguale a $1/T$ per ogni ϵ_t^+ , $t = 1, \dots, T$.

Step 3: Si costruisca il campione bootstrap $y_t^* = \{y_1^*, \dots, y_T^*\}$ come:

$$y_t^* - \hat{\phi}_0 - \sum_{j=1}^p \hat{\phi}_j y_{t-j}^* = \epsilon_t^+, \quad t = 1, \dots, T,$$

con $y_{1-j} = y_{\tau-j+1}$, $j = 1, \dots, p$, dove τ ha una distribuzione discreta uniforme sugli interi p, \dots, T , e $\epsilon_t^* = \hat{\sigma}\epsilon_t^+$, dove $\hat{\sigma}^2$ è la stima della varianza dell'errore di previsione un passo in avanti, ottenuto con i residui calcolati nello Step 1.

Step 4: Si ottenga S^* come funzione di y_t^* , utilizzando la stessa procedura che si utilizzerebbe per ottenere la statistica di interesse S come funzione del campione originario y_t .

Step 5: Ripetere i passi 2, 3 e 4 B volte per generare B campioni bootstrap indipendenti $Y_{t,b}^*$ e calcolare S_b^* per $b = 1, \dots, B$. Si approssima la distribuzione di S con la funzione di distribuzione bootstrap

$$FS_{S_B^*}^*(s) = \frac{1}{B} \sum_{b=1}^B \mathbf{I}(S_b^* \leq s)$$

dove \mathbf{I} è la funzione indicatrice.

Poskitt (2008) ha inoltre dimostrato che $FS_{S_B^*}^*(s)$ converge in distribuzione a S con probabilità pari a 1, giustificando, così, la validità teorica del *sieve* bootstrap nelle serie stazionarie a memoria lunga.

2.2.4 Block Bootstrap

Le principali alternative al *sieve* bootstrap per le serie storiche si basano sul campionamento di blocchi di osservazioni, una tecnica completamente non parametrica. I blocchi possono essere tra loro sia sovrapposti (Kunsch, 1989) che non (Carlstein, 1986).

Sia $l < T$ un numero intero che denota la lunghezza dei blocchi e può essere fissa oppure casuale. Sia $B_i = (y_i, \dots, y_{i+l-1})$ un blocco di osservazioni che inizia dall'osservazione i -esima tale per cui $1 \leq i \leq T - l + 1$. La serie ricampionata tramite questo metodo bootstrap, di lunghezza bl , dove b è il numero di blocchi utilizzati, è definita come $(B_{I_1}, \dots, B_{I_b})$, dove I_1, \dots, I_b sono variabili casuali i.i.d. con distribuzione uniforme su $\{1, \dots, T - l + 1\}$ nel caso del metodo bootstrap a blocchi sovrapposti (*moving block* bootstrap, MBB), e su $\{1 + l(i - 1); i = 1, \dots, b\}$ nel caso del metodo bootstrap a blocchi non sovrapposti (*non-overlapping block* bootstrap, NBB). Una volta che la serie ricampionata y_t^* è stata ottenuta, la statistica di interesse e la sua distribuzione sono stimate alla stessa maniera degli Step 4 e 5 visti per il *wild* e per il *sieve* bootstrap. Inoltre, per $l \rightarrow \infty$, è stato dimostrato che la distribuzione di probabilità congiunta finito-dimensionale può essere riprodotta dalle osservazioni ricampionate. Il *moving block* bootstrap, quindi, è capace di approssimare efficientemente le caratteristiche del processo

per l'intera popolazione (Radovanov e Marcikić, 2014).

Un ulteriore metodo bootstrap basato sul ricampionamento dei blocchi è il *circular block* bootstrap, CBB, proposto da Politis e Romano (1991). Questo metodo rappresenta un modo semplice e automatico per ottenere una distribuzione bootstrap non distorta “avvolgendo” le osservazioni Y_i attorno ad un “cerchio”, in cui, per $i > T$, $Y_i \equiv Y_{i_T}$, dove $i_T = i(\text{mod}T)$, e $Y_0 \equiv Y_T$, creando, così, una continuità artificiale nei dati. Vengono poi selezionati casualmente i blocchi da questo “cerchio” per costruire la nuova sequenza di dati bootstrap y_i^* . Questa idea è associata con la definizione della sequenza circolare di autocovarianza dei modelli di serie storiche (Lahiri, 2013).

Affinchè questi schemi di ricampionamento abbiano delle prestazioni desiderabili, è necessario che i blocchi siano, tra loro, approssimativamente indipendenti ma, purtroppo, nel caso di serie storiche a memoria lunga, questa assunzione potrebbe non essere soddisfatta. Lahiri (1993) è stato il primo ad analizzare l'applicabilità dei metodi *block* bootstrap per serie storiche a memoria lunga. Nel suo lavoro, l'Autore si è concentrato sui processi la cui funzione di autocorrelazione si comporta, asintoticamente, come indicato nell'Equazione 2.1. Lahiri ha dimostrato che il MBB riesce ad approssimare accuratamente la distribuzione della media campionaria normalizzata, ma solo quando questa statistica è asintoticamente normale. Questo limite è dovuto al fatto che, per costruzione, i blocchi del campione bootstrap sono indipendenti tra loro, ma ciò annulla la forte dipendenza presente nelle osservazioni originali. Inoltre, le *performance* di questi metodi dipendono fortemente dalla scelta dei blocchi e dalla loro lunghezza ed è stato dimostrato da Kim e Nordman (2011) che il numero di blocchi ottimale (in termini di minimizzazione della varianza della stima) decresce quando la forza della dipendenza nel processo sottostante aumenta. Gli stessi Autori hanno, quindi, raccomandato, nel caso di processi a memoria lunga, l'utilizzo dei blocchi sovrapposti (MBB) ed hanno suggerito, come “regola del pollice” per determinare la lunghezza dei blocchi, che questa sia dell'ordine $O(T^{1/2})$, dove T è la lunghezza delle serie storica originale.

Un altro limite dei metodi *block* bootstrap è che la varianza della media bootstrap, condizionata ai dati originali, ha un tasso di crescita minore rispetto a quello della varianza della media delle osservazioni originali, proprio a causa dell'eliminazione della dipendenza tra le osservazioni. Una possibile soluzione a questo problema implica una diversa normalizzazione della media bootstrap, che avviene moltiplicando la statistica per b^d , dove b è il numero di blocchi e d è il parametro di memoria lunga. Questa normalizzazione richiede, tuttavia, una stima a priori di

d , non può essere generalizzata a statistiche diverse dalla media campionaria e la sua validità è stata dimostrata solo per serie storiche a memoria lunga (Butka e Puka, 2014).

2.2.5 Pre-filtered bootstrap

Con il fine di gestire il problema della persistenza della memoria nelle serie storiche a memoria lunga, una strategia che si è dimostrata avere successo è quella di eliminare questa persistenza a priori, così da ottenere una serie storica a memoria corta, ed implementare, quindi, delle tecniche bootstrap valide per serie storiche debolmente persistenti.

Partendo, quindi, dall'idea che l'applicazione di questo "filtro" preliminare possa portare dei vantaggi in termini di convergenza dello stimatore, Poskitt et al. (2015) hanno proposto una versione modificata del *sieve* bootstrap, nella quale viene applicato un pre-filtraggio alla serie, e, solo successivamente, viene applicato il metodo *sieve* bootstrap alla serie storica pre-filtrata. La procedura proposta da Poskitt et al. (2015) per il metodo *pre-filtered sieve* bootstrap (PFSB) è, quindi, la seguente:

Step 0: Si consideri il filtro $(1 - L)^d = \sum_{j=0}^{\infty} a_j(d)L^j$ e si calcoli \hat{d} , tramite uno stimatore consistente del parametro di memoria lunga d . A questo punto, ottenere la serie filtrata:

$$\hat{u}_t = \sum_{j=0}^{t-1} a_j(\hat{d})y_{t-j}, \quad t = 1, \dots, T.$$

Si applicano poi gli Step 1, 2 e 3 del *sieve* bootstrap descritto nella Sezione 2.2.3, utilizzando \hat{u}_t anziché y_t , ottenendo così il campione *sieve* bootstrap \hat{u}_t^* .

Step 4: Si ottenga il campione *pre-filtered sieve* bootstrap:

$$y_t^* = \sum_{j=0}^{t-1} a_j(-\hat{d})\hat{u}_{t-j}^*, \quad t = 1, \dots, T,$$

e si calcoli S^* come funzione di y_t^* , utilizzando la stessa procedura che si utilizzerebbe per ottenere la statistica di interesse S come funzione del campione originario y_t .

Step 5: Ripetere i passi 1, 2, 3 e 4 B volte per generare B campioni bootstrap indipendenti $y_{t,b}^*$ e calcolare S_b^* per $b = 1, \dots, B$. Si approssima la distribuzione di S con la funzione di distribuzione bootstrap

$$FPF_{S_B^*}(s) = \frac{1}{B} \sum_{b=1}^B I(S_b^* \leq s)$$

dove I è la funzione indicatrice.

L'analisi di Poskitt et al. (2015), ristretta al caso di serie a memoria lunga stazionarie ($-1/2 < d < 1/2$), mostra che l'errore della distribuzione bootstrap del metodo PFSB, usata come approssimazione della distribuzione di S , è minore rispetto a quello relativo al metodo SB, per $d > 0$.

Kapetanios et al. (2019) hanno esteso i lavori di Kapetanios e Papailias (2011) e di Poskitt et al. (2015) dimostrando la validità asintotica di questi metodi di ricampionamento basati sul pre-filtraggio, quando abbinati a diversi approcci bootstrap (tra cui il *block* bootstrap, ottenendo così il metodo *pre-filtered moving block* bootstrap, PFMBB), per serie frazionalmente integrate sia stazionarie che non. Questi Autori hanno, infatti, mostrato che la distanza di Mallows (Bickel e Freedman, 1981) tra la distribuzione di S e l'approssimazione bootstrap di questi metodi tende a zero quando la dimensione campionaria aumenta, fintanto che si consideri uno stimatore consistente di d , provando, quindi, la validità dei metodi bootstrap basati sul pre-filtraggio nel caso di serie storiche a memoria lunga.

Capitolo 3

Le combinazioni bootstrap

Seguendo l'intuizione alla base delle combinazioni di previsioni trattate nel Capitolo 1, in questo lavoro vengono presentati due diversi approcci originali, tramite i quali estendere le metodologie di combinazione ai metodi bootstrap, con il fine di ottenere una migliore stima del parametro di memoria lunga per le serie storiche stazionarie a memoria lunga. In entrambi gli approcci, i metodi bootstrap utilizzati sono quelli approfonditi nel Capitolo 2. In questo Capitolo, la Sezione 3.1 approfondisce l'approccio basato sulla combinazione di ciascun campione bootstrap, con il fine di ottenere una singola serie campionaria da utilizzare per la stima del parametro di memoria lunga, usando, quindi, una metodologia di tipo bootstrap ibrido. La Sezione 3.2 approfondisce il secondo approccio, nel quale la combinazione è direttamente costruita con le stime del parametro di memoria lunga ottenute da ciascun metodo bootstrap precedentemente utilizzato per il ricampionamento dei dati.

3.1 Bootstrap ibrido

In letteratura, con “bootstrap ibrido” si intende una tecnica di ricampionamento che integra più approcci bootstrap in un'unica procedura. Queste metodologie, per le quali esistono ancora pochi studi in letteratura, sono state sviluppate per migliorare l'accuratezza delle stime statistiche o per affrontare specifiche problematiche legate al campionamento dei dati. In generale, l'idea alla base del bootstrap ibrido è quella di sfruttare le caratteristiche di diversi metodi bootstrap, per ottenere risultati migliori rispetto a quelli che si otterrebbero utilizzando un singolo metodo di ricampionamento (Kreiss e Paparoditis, 2011).

L'approccio di bootstrap ibrido proposto in questa tesi rappresenta, invece, un'innovazione significativa rispetto a quanto presente in letteratura. A differenza degli approcci esistenti, che integrano i metodi bootstrap a livello di algoritmo, il nostro lavoro introduce un concetto nuovo, basato sulla combinazione diretta dei campioni bootstrap generati da diversi metodi. Questa metodologia, per nostra conoscenza, non ha precedenti nella letteratura esistente e offre una nuova prospettiva per migliorare le *performance* degli stimatori. Al contrario, l'approccio di combinare le stime bootstrap, presentato nella Sezione 3.2, pur essendo innovativo, si ispira al concetto consolidato della combinazione di previsioni, ma applicato nel contesto del ricampionamento.

Questo lavoro si concentra sullo studio di serie storiche stazionarie a memoria lunga, delle quali si vuole stimare il parametro di memoria lunga d . Le metodologie proposte in questo Capitolo sono estendibili anche allo studio di altre statistiche, definite, in generale, come $S = S(y_1, \dots, y_T)$, la cui distribuzione può essere approssimata tramite le distribuzioni bootstrap. L'idea alla base del bootstrap ibrido proposto in questa Sezione è quella di generare un campione bootstrap finale, y_t^{**} , $t = 1, \dots, T$, ottenuto combinando i diversi campioni bootstrap, ottenuti dai metodi approfonditi nel Capitolo 2. Ogni metodo bootstrap considerato viene implementato usando, per i parametri specificati, le raccomandazioni provenienti dalla letteratura correlata. In particolare, saranno combinati i campioni bootstrap ottenuti a partire da: bootstrap parametrico (PB), *wild* bootstrap (WB), *sieve* bootstrap (SB), *moving block* bootstrap (MBB), *circular block* bootstrap (CBB), *pre-filtered sieve* bootstrap (PFSB) e *pre-filtered moving block* bootstrap (PFMBB). Successivamente, la statistica S viene stimata usando il campione bootstrap combinato.

Con il fine di avere una panoramica più ampia riguardo le *performance* di questo possibile approccio, si analizzano diverse proposte di bootstrap ibrido, utilizzando delle combinazioni simili a quelle maggiormente studiate in ambito previsivo, con una particolare attenzione alle combinazioni semplici ed alle combinazioni lineari.

3.1.1 Combinazioni semplici

Come evidenziato nel Capitolo 1, nella letteratura previsiva è noto che la combinazione tramite media semplice conduce spesso a risultati difficili da superare sia in termini di *performance* che di robustezza. Inoltre, altre strategie di combinazione, come la mediana e la media troncata, possono offrire il vantaggio di essere meno sensibili ai valori estremi rispetto alla media, ottenendo talvolta risultati più robusti rispetto a quelli ottenuti tramite la media stessa. Per questa

analisi, quindi, le combinazioni semplici selezionate per costruire i metodi bootstrap ibridi sono la media, la mediana e la media troncata, di cui in seguito verrà illustrata l'operatività.

Sia y_t , $t = 1, \dots, T$, una serie storica a memoria lunga stazionaria e sia S la statistica di interesse, allora, per un numero B di campioni bootstrap, l'algoritmo dei metodi bootstrap ibridi che utilizzano combinazioni semplici è il seguente:

- Step 1. Si generano, a partire dalla serie originale y_t , le serie bootstrap y_t^{*boot} utilizzando i metodi bootstrap descritti in precedenza, in particolare $boot \in \{PB, WB, SB, MBB, CBB, PFSB, PFMBB\}$.
- Step 2. Si combinano i campioni bootstrap utilizzando uno schema di combinazione semplice, così da ottenere il campione bootstrap ibrido y_t^{**} . Si calcola S^* come funzione di y_t^{**} , utilizzando la stessa procedura che si utilizzerebbe per ottenere la statistica di interesse S come funzione del campione originario y_t .
- Step 3. Ripetere i passi 1 e 2 B volte per generare B campioni bootstrap ibridi $y_{t,b}^{**}$, $b = 1, \dots, B$, e calcolare i B valori della desiderata statistica S_b^* , $b = 1, \dots, B$. Si approssima la distribuzione della statistica S con la distribuzione bootstrap

$$FS^*(s) = \frac{1}{B} \sum_{b=1}^B I(S_b^* \leq s)$$

dove I è la funzione indicatrice.

La serie bootstrap ibrida y_t^{**} , $t = 1, \dots, T$, ottenuta al passo 2, dipende dallo schema di combinazione scelto.

In particolare, nel caso dello schema di combinazione che utilizza la media semplice, la serie bootstrap combinata è ottenuta tramite la formula:

$$y_t^{**} = \frac{y_t^{*PB} + y_t^{*WB} + y_t^{*SB} + y_t^{*MBB} + y_t^{*CBB} + y_t^{*PFSB} + y_t^{*PFMBB}}{7}$$

per $t = 1, \dots, T$.

Nel caso dello schema di combinazione che utilizza la mediana, invece, nel passo 2 è, innanzitutto, necessario ordinare, per ogni $t = 1, \dots, T$, i valori $y_t^{*PB}, y_t^{*WB}, y_t^{*SB}, y_t^{*MBB}, y_t^{*CBB}, y_t^{*PFSB}, y_t^{*PFMBB}$ in ordine non decrescente, costruendo con questi valori i vettori

\mathbf{y}_t^* , $t = 1, \dots, T$. I valori della serie bootstrap ibrida y_t^{**} , $t = 1, \dots, T$, corrispondono ai termini che occupano la posizione centrale dei vettori ordinati \mathbf{y}_t^* , per $t = 1, \dots, T$, cioè $y_t^{**} = y_{t,4}^*$, dove il pedice 4 indica il valore che si trova in quarta posizione nel vettore ordinato.

Infine, nel caso dello schema di combinazione che utilizza la media troncata, nel passo 2 è necessario ordinare, per ogni $t = 1, \dots, T$, i valori $y_t^{*PB}, y_t^{*WB}, y_t^{*SB}, y_t^{*MBB}, y_t^{*CBB}, y_t^{*PFSB}, y_t^{*PFMBB}$ in ordine non decrescente, costruendo con questi valori i vettori \mathbf{y}_t^* , $t = 1, \dots, T$. Si procede, poi, eliminando in ogni vettore ordinato \mathbf{y}_t^* il primo e l'ultimo termine, ossia, rispettivamente, il valore minore e maggiore tra quelli presenti. Infine, y_t^{**} , $t = 1, \dots, T$, viene calcolato come media aritmetica dei cinque valori rimanenti nel vettore \mathbf{y}_t^* . La formula per ottenere la serie bootstrap combinata, in questo caso, è, quindi, la seguente:

$$y_t^{**} = \frac{y_t^{*PB} + y_t^{*WB} + y_t^{*SB} + y_t^{*MBB} + y_t^{*CBB} + y_t^{*PFSB} + y_t^{*PFMBB} - \min \mathbf{y}_t^* - \max \mathbf{y}_t^*}{5}$$

per ogni $t = 1, \dots, T$.

3.1.2 Combinazioni lineari

Vengono ora considerate alcune combinazioni più sofisticate, concentrandosi su quattro schemi di combinazione lineare facilmente intuibili. In particolare, come schemi di combinazione lineare per costruire i metodi bootstrap ibridi, ne ho considerati uno derivato dalla regressione lineare semplice senza intercetta, uno basato sulla correlazione tra la serie originale e le serie bootstrap e due basati, rispettivamente, sull'errore quadratico medio (MSE) e sull'errore medio assoluto (MAE) delle serie bootstrap (riprendendo l'idea dei pesi basati sulla *performance*).

Sia, nuovamente, y_t , $t = 1, \dots, T$, una serie a memoria lunga stazionaria, e sia S la statistica di interesse, allora, per un numero B di campioni bootstrap, l'algoritmo dei metodi bootstrap ibridi che utilizzano combinazioni lineari è il seguente:

- Step 1. Si generano, a partire dalla serie originale y_t , le serie bootstrap y_t^{*boot} utilizzando i metodi bootstrap descritti in precedenza, in particolare $boot \in \{PB, WB, SB, MBB, CBB, PFSB, PFMBB\}$.
- Step 2. (a) Si calcolano le quantità di interesse per il calcolo dei pesi per la combinazione dei campioni bootstrap; (b) si calcola il peso relativo a ciascun campione bootstrap, w^{boot} ; (c) si combinano i campioni bootstrap con uno schema di combinazione lineare che utilizza i pesi precedentemente calcolati, così da ottenere il campione bootstrap ibrido y_t^{**} ; (d) si

calcola S^* come funzione di y_t^{**} , utilizzando la stessa procedura che si utilizzerebbe per ottenere la statistica di interesse S come funzione del campione originario y_t .

Step 3. Ripetere i passi 1 e 2 B volte per generare B campioni bootstrap ibridi $y_{t,b}^{**}$, $b = 1, \dots, B$, e calcolare i B valori della desiderata statistica S_b^* , $b = 1, \dots, B$. Si approssima la distribuzione della statistica S con la distribuzione bootstrap

$$FS^*(s) = \frac{1}{B} \sum_{b=1}^B I(S_b^* \leq s)$$

dove I è la funzione indicatrice.

La serie bootstrap ibrida y_t^{**} , ottenuta al passo 2, dipende dallo schema di combinazione scelto.

In particolare, nel caso dello schema di combinazione basato sulla regressione, si considera il modello di regressione lineare semplice senza intercetta in cui la serie originale corrisponde alla variabile risposta e ciascun campione bootstrap calcolato al passo 1 corrisponde ad una variabile esplicativa. Il modello lineare risulta quindi essere:

$$y_t = \beta_{PB} y_t^{*PB} + \beta_{WB} y_t^{*WB} + \beta_{SB} y_t^{*SB} + \beta_{MBB} y_t^{*MBB} + \beta_{CBB} y_t^{*CBB} + \\ + \beta_{PFSB} y_t^{*PFSB} + \beta_{PFMBB} y_t^{*PFMBB} + \epsilon_t$$

per $t = 1, \dots, T$, dove $\epsilon_t \sim N(0, \sigma^2)$ i.i.d. per ogni $t = 1, \dots, T$.

Il vettore di coefficienti $\hat{\beta} = (\hat{\beta}_{PB}, \hat{\beta}_{WB}, \hat{\beta}_{SB}, \hat{\beta}_{MBB}, \hat{\beta}_{CBB}, \hat{\beta}_{PFSB}, \hat{\beta}_{PFMBB})'$, stimato tramite OLS, corrisponde al vettore dei pesi da usare per la serie bootstrap combinata. La serie bootstrap ibrida y_t^{**} , per $t = 1, \dots, T$, viene calcolata, quindi, come:

$$y_t^{**} = \hat{\beta}_{PB} y_t^{*PB} + \hat{\beta}_{WB} y_t^{*WB} + \hat{\beta}_{SB} y_t^{*SB} + \hat{\beta}_{MBB} y_t^{*MBB} + \hat{\beta}_{CBB} y_t^{*CBB} + \\ + \hat{\beta}_{PFSB} y_t^{*PFSB} + \hat{\beta}_{PFMBB} y_t^{*PFMBB} .$$

È possibile imporre restrizioni aggiuntive sui coefficienti di regressione $\hat{\beta}$, come la non negatività e la condizione che questi sommino ad 1, similmente a quanto avviene nel contesto della combinazione di previsioni (Capitolo 1). Tuttavia, nel resto del presente lavoro si adotta un modello di regressione senza intercetta ma senza imporre ulteriori vincoli, con la possibilità di esplorare altre opzioni in studi futuri.

Nel caso dello schema di combinazione basato sulla correlazione tra la serie originale e i campioni bootstrap, la serie bootstrap combinata y_t^{**} è ottenuta a sua volta in tre passi:

Step 2a. Per ogni campione bootstrap ricavato al passo 1, si calcola il quadrato della correlazione tra il campione bootstrap y_t^{*boot} e la serie originale y_t .

Step 2b. Si calcola il peso di ciascun metodo bootstrap come:

$$w^{boot} = \frac{\rho^{2,boot}}{\rho^{2,PB} + \rho^{2,WB} + \rho^{2,SB} + \rho^{2,MBB} + \rho^{2,CBB} + \rho^{2,PFSB} + \rho^{2,PFMBB}},$$

dove $\rho^{2,boot} = cor(y_t, y_t^{*boot})^2$, in modo che la somma di tali pesi sia pari a 1.

Step 2c. Si combinano i campioni bootstrap, ottenuti al passo 1, tramite uno schema di combinazione lineare i cui pesi sono stati ricavati al passo 2b. La serie bootstrap ibrida y_t^{**} viene calcolata, quindi, come:

$$y_t^{**} = w^{PB} y_t^{*PB} + w^{WB} y_t^{*WB} + w^{SB} y_t^{*SB} + w^{MBB} y_t^{*MBB} + w^{CBB} y_t^{*CBB} + w^{PFSB} y_t^{*PFSB} + w^{PFMBB} y_t^{*PFMBB},$$

per $t = 1, \dots, T$.

Nel caso dello schema di combinazione basato sull'errore quadratico medio (MSE), gli errori sono calcolati come differenza tra le osservazioni della serie originale e i valori delle serie bootstrap, ed i pesi di ponderazione sono proporzionali al reciproco dello stesso MSE. Più in particolare, la serie bootstrap combinata y_t^{**} è ottenuta in tre passi:

Step 2a. Per ogni campione bootstrap ricavato al passo 1, si calcola MSE come $MSE^{boot} = \frac{1}{T} \sum_{i=1}^T (y_t - y_t^{*boot})^2$. Inoltre, per il successivo calcolo dei pesi, è necessario calcolare il reciproco di MSE per ogni metodo bootstrap come:

$$\overline{rMSE}^{boot} = \frac{1}{MSE^{boot}}.$$

Step 2b. Si calcola il peso di ciascun metodo bootstrap come:

$$w^{boot} = \frac{\overline{rMSE}^{boot}}{\overline{rMSE}^{PB} + \overline{rMSE}^{WB} + \overline{rMSE}^{SB} + \overline{rMSE}^{MBB} + \overline{rMSE}^{CBB} + \overline{rMSE}^{PFSB} + \overline{rMSE}^{PFMBB}},$$

in modo che la somma di tali pesi sia pari a 1.

Step 2c. Si combinano i campioni bootstrap, ottenuti al passo 1, tramite uno schema di combinazione lineare i cui pesi sono stati ricavati al passo 2b. La serie bootstrap ibrida y_t^{**} viene calcolata, quindi, come:

$$y_t^{**} = w^{PB} y_t^{*PB} + w^{WB} y_t^{*WB} + w^{SB} y_t^{*SB} + w^{MBB} y_t^{*MBB} + w^{CBB} y_t^{*CBB} + w^{PFSB} y_t^{*PFSB} + w^{PFMBB} y_t^{*PFMBB},$$

per $t = 1, \dots, T$.

Si procede similmente anche nel caso dello schema di combinazione basato sull'errore medio assoluto (MAE), in cui i pesi di ponderazione sono proporzionali al reciproco dello stesso. La procedura per ottenere la serie bootstrap combinata y_t^{**} è:

Step 2a. Per ogni campione bootstrap ricavato al passo 1, si calcola MAE come $MAE^{boot} = \frac{1}{T} \sum_{i=1}^T |y_t - y_t^{*boot}|$. Inoltre, per il successivo calcolo dei pesi, è necessario calcolare il reciproco di MAE per ogni metodo bootstrap come:

$$\overline{rMAE}^{boot} = \frac{1}{MAE^{boot}}.$$

Step 2b. Si calcola il peso di ciascun metodo bootstrap come:

$$w^{boot} = \frac{\overline{rMAE}^{boot}}{\overline{rMAE}^{PB} + \overline{rMAE}^{WB} + \overline{rMAE}^{SB} + \overline{rMAE}^{MBB} + \overline{rMAE}^{CBB} + \overline{rMAE}^{PFSB} + \overline{rMAE}^{PFMBB}},$$

in modo che la somma di tali pesi sia pari a 1.

Step 2c. Si combinano i campioni bootstrap, ottenuti al passo 1, tramite uno schema di combinazione lineare i cui pesi sono stati ricavati al passo 2b. La serie bootstrap ibrida y_t^{**} viene calcolata, quindi, come:

$$y_t^{**} = w^{PB} y_t^{*PB} + w^{WB} y_t^{*WB} + w^{SB} y_t^{*SB} + w^{MBB} y_t^{*MBB} + w^{CBB} y_t^{*CBB} + w^{PFSB} y_t^{*PFSB} + w^{PFMBB} y_t^{*PFMBB},$$

per $t = 1, \dots, T$.

3.2 Combinazione di stime bootstrap

Il secondo approccio che proponiamo si basa, invece, sulla combinazione delle stime del parametro di memoria lunga d . Ogni stima del parametro, usata in seguito per la combinazione, è ottenuta a partire da un metodo bootstrap diverso. Pur essendo una proposta originale nel campo, questo approccio si avvicina maggiormente a quello utilizzato nella combinazione di previsioni descritto nel Capitolo 1. Anche in questo caso, le metodologie sono estendibili anche allo studio di altre statistiche, la cui distribuzione può essere approssimata tramite la distribuzione bootstrap, per cui si fa riferimento ad una generale statistica $S = S(y_1, \dots, y_T)$. I metodi bootstrap utilizzati per le combinazioni di stime bootstrap sono gli stessi presi in considerazione nella Sezione 3.1 (PB, WB, SB, MBB, CBB, PFSB, PFMBB), implementati secondo le

raccomandazioni provenienti dalla letteratura correlata. Traendo nuovamente spunto dalla letteratura in ambito previsivo, le combinazioni analizzate in questa Sezione sono intuitivamente simili a quelle trattate nella Sezione precedente (con l'eccezione delle combinazioni basate su regressione lineare e correlazione), ma gli algoritmi implementati seguono una logica diversa. Infatti, in questo approccio, l'attenzione è maggiormente rivolta verso le stime bootstrap della statistica, mentre nei metodi bootstrap ibridi si concentra sulle serie ricampionate.

3.2.1 Combinazioni semplici

Le strategie di combinazione semplice selezionate per costruire le combinazioni di stime bootstrap sono, nuovamente, la media e le sue alternative più robuste, la mediana e la media troncata.

Sia y_t , $t = 1, \dots, T$, una serie storica a memoria lunga stazionaria, e sia S la statistica di interesse, allora, per un numero B di campioni bootstrap, l'algoritmo per combinare le stime bootstrap attraverso combinazioni semplici è il seguente:

Step 1. Si generano, a partire dalla serie originale y_t , le serie bootstrap y_t^{*boot} utilizzando i metodi bootstrap descritti in precedenza, in particolare $boot \in \{PB, WB, SB, MBB, CBB, PFSB, PFMBB\}$. Si ricavano, a partire dai campioni bootstrap, le stime della statistica desiderata S^{boot} .

Step 2. Ripetere il passo 1 B volte, così da ottenere, per ogni metodo bootstrap, il vettore B -dimensionale \mathbf{S}^{boot} che contiene tutti i valori delle stime S_b^{boot} , $b = 1, \dots, B$, dove $boot$ è uno dei metodi bootstrap utilizzati.

Step 3. Si calcola, per ogni metodo bootstrap, la media delle stime contenute nel vettore \mathbf{S}^{boot} come:

$$S^{*boot} = \frac{1}{B} \sum_{b=1}^B S_b^{boot}.$$

Questo valore è una stima bootstrap di S .

Step 4. Si combinano le stime bootstrap S^{*boot} , ottenute al passo precedente, tramite uno schema di combinazione semplice, così da ottenere la stima finale della statistica, S^{**} .

Il valore della stima finale della statistica S^{**} , ottenuta al passo 4, dipende dallo schema di combinazione scelto.

In particolare, nel caso dello schema che utilizza la media semplice, la stima finale della statistica è ottenuta tramite la formula:

$$S^{**} = \frac{S^{*PB} + S^{*WB} + S^{*SB} + S^{*MBB} + S^{*CBB} + S^{*PFSB} + S^{*PFMBB}}{7}.$$

Nel caso dello schema di combinazione che utilizza la mediana, al passo 4 è, innanzitutto, necessario ordinare i valori $S^{*PB}, S^{*WB}, S^{*SB}, S^{*MBB}, S^{*CBB}, S^{*PFSB}, S^{*PFMBB}$ in ordine non decrescente nel vettore \mathbf{S}^* . Il valore della stima finale S^{**} è, in questa successione ordinata di valori, uguale al termine che occupa la posizione centrale, cioè $S^{**} = S_4^*$, dove il pedice 4 indica il valore che si trova in quarta posizione nel vettore ordinato \mathbf{S}^* .

Infine, nel caso dello schema di combinazione che utilizza la media troncata, al passo 4 è necessario ordinare i valori $S^{*PB}, S^{*WB}, S^{*SB}, S^{*MBB}, S^{*CBB}, S^{*PFSB}, S^{*PFMBB}$ in ordine non decrescente nel vettore \mathbf{S}^* . Si procede, poi, eliminando dal vettore ordinato \mathbf{S}^* il primo e l'ultimo termine, ossia, rispettivamente, il valore minore e maggiore tra quelli presenti. Infine, S^{**} viene calcolato come media aritmetica dei cinque valori rimanenti nel vettore \mathbf{S}^* . La formula per ottenere la stima finale della statistica, in questo caso, è, quindi, la seguente:

$$S^{**} = \frac{S^{*PB} + S^{*WB} + S^{*SB} + S^{*MBB} + S^{*CBB} + S^{*PFSB} + S^{*PFMBB} - \min \mathbf{S}^* - \max \mathbf{S}^*}{5}.$$

3.2.2 Combinazioni lineari

Come schemi di combinazione lineare per costruire le combinazioni di stime bootstrap, vengono considerati solamente quelli che si basano sull'idea dei pesi basati sulle *performance* degli stimatori, misurate in termini di errore. In particolare, si adattano a questo approccio gli schemi di combinazione basati sui due indicatori MSE e MAE. Vengono escluse, invece, le combinazioni basate su regressione lineare e su correlazione poiché, come verrà illustrato, il calcolo dei pesi, per questo approccio, dipende dalla stima della statistica S sulla serie storica originale e dalle stime bootstrap, calcolate durante le B replicazioni. Tuttavia, la stima derivata dalla serie originale rimane costante per ogni replicazione bootstrap, rendendo l'applicazione di questi due metodi di combinazione di stime bootstrap non praticabile o inefficace.

Sia y_t , $t = 1, \dots, T$, una serie storica a memoria lunga stazionaria e sia S la statistica di interesse, allora, per un numero B di campioni bootstrap, l'algoritmo per combinare le stime ottenute dai diversi metodi bootstrap è il seguente:

Step 1. Si generano, a partire dalla serie originale y_t , le serie bootstrap y_t^{*boot} utilizzando i metodi bootstrap descritti in precedenza, in particolare $boot \in \{PB, WB, SB, MBB, CBB, PFSB, PFMBB\}$. Si ricavano, a partire dai campioni bootstrap, le stime della statistica desiderata S^{boot} .

Step 2. Ripetere il passo 1 B volte, così da ottenere, per ogni metodo bootstrap, il vettore B -dimensionale \mathbf{S}^{boot} che contiene tutti i valori delle stime S_b^{boot} , $b = 1, \dots, B$, dove $boot$ è uno dei metodi bootstrap utilizzati. Si ricava, inoltre, il valore della statistica di interesse S^* come funzione del campione originale y_t .

Step 3. Si calcola, per ogni metodo bootstrap, la media delle stime contenute nel vettore \mathbf{S}^{boot} come:

$$S^{*boot} = \frac{1}{B} \sum_{b=1}^B S_b^{boot}.$$

Questo valore è una stima bootstrap di S .

Step 4. (a) Si calcolano le quantità di interesse per il calcolo dei pesi e per la combinazione delle stime bootstrap; (b) si calcola il peso relativo a ciascun metodo bootstrap, w^{boot} ; (c) si combinano le medie delle stime bootstrap con uno schema di combinazione lineare che utilizza i pesi precedentemente calcolati, così da ottenere la stima finale combinata S^{**} .

Anche per queste strategie di combinazione, è possibile vedere come la stima della statistica S^{**} , ottenuta al passo 4, dipenda dallo schema di combinazione scelto.

In particolare, nel caso dello schema di combinazione basato su MSE, i cui errori sono calcolati come differenza tra la stima della statistica sulla serie originale e i valori delle stime bootstrap, i pesi di ponderazione sono proporzionali al reciproco dello stesso. In particolare, la stima finale della statistica, S^{**} , è ottenuta in tre passi:

Step 4a. Per ogni metodo bootstrap, si calcola MSE come $MSE^{boot} = \frac{1}{B} \sum_{b=1}^B (S^* - S_b^{boot})^2$. Inoltre, per il successivo calcolo dei pesi, è necessario calcolare il reciproco di MSE per ogni metodo bootstrap come:

$$\overline{rMSE}^{boot} = \frac{1}{MSE^{boot}}.$$

Step 4b. Si calcola il peso di ciascun metodo bootstrap come:

$$w^{boot} = \frac{\overline{rMSE}^{boot}}{\overline{rMSE}^{PB} + \overline{rMSE}^{WB} + \overline{rMSE}^{SB} + \overline{rMSE}^{MBB} + \overline{rMSE}^{CBB} + \overline{rMSE}^{PFSB} + \overline{rMSE}^{PFMBB}},$$

in modo che la somma di tali pesi sia pari a 1.

Step 4c. Si combinano le medie delle stime bootstrap, ottenute al passo 3, tramite uno schema di combinazione lineare i cui pesi sono stati ricavati allo Step 4b. La stima finale della statistica S^{**} viene calcolata, quindi, come:

$$S^{**} = w^{PB} S^{*PB} + w^{WB} S^{*WB} + w^{SB} S^{*SB} + w^{MBB} S^{*MBB} + w^{CBB} S^{*CBB} + \\ + w^{PFSB} S^{*PFSB} + w^{PFMBB} S^{*PFMBB}.$$

Si procede similmente anche nel caso dello schema di combinazione basato su MAE, in cui i pesi di ponderazione sono proporzionali al reciproco dello stesso. La procedura per ottenere la stima finale della statistica, S^{**} , è:

Step 4a. Per ogni metodo bootstrap, si calcola MAE come $MAE^{boot} = \frac{1}{B} \sum_{i=1}^B |S^* - S_b^{boot}|$. Inoltre, per il successivo calcolo dei pesi, è necessario calcolare il reciproco del MAE per ogni metodo bootstrap come:

$$\overline{rMAE}^{boot} = \frac{1}{MAE^{boot}}.$$

Step 4b. Si calcola il peso di ciascun metodo bootstrap come:

$$w^{boot} = \frac{\overline{rMAE}^{boot}}{\overline{rMAE}^{PB} + \overline{rMAE}^{WB} + \overline{rMAE}^{SB} + \overline{rMAE}^{MBB} + \overline{rMAE}^{CBB} + \overline{rMAE}^{PFSB} + \overline{rMAE}^{PFMBB}},$$

in modo che la somma di tali pesi sia pari a 1.

Step 4c. Si combinano le medie delle stime bootstrap, ottenute al passo 3, tramite uno schema di combinazione lineare i cui pesi sono stati ricavati al passo 4b. La stima finale della statistica S^{**} viene calcolata, quindi, come:

$$S^{**} = w^{PB} S^{*PB} + w^{WB} S^{*WB} + w^{SB} S^{*SB} + w^{MBB} S^{*MBB} + w^{CBB} S^{*CBB} + \\ + w^{PFSB} S^{*PFSB} + w^{PFMBB} S^{*PFMBB}.$$

Dopo aver definito i metodi di combinazione proposti in questo lavoro, nel Capitolo 4 sono presentati degli studi di simulazione Monte Carlo per valutare le *performance* degli approcci proposti rispetto a quelle dei metodi bootstrap tradizionali, mentre, nel Capitolo 5, le stesse procedure sono applicate a due serie storiche reali.

Capitolo 4

Studio Monte Carlo

L'obiettivo di questo Capitolo è quello di valutare le *performance* sia dell'approccio bootstrap ibrido che di quello riguardante la combinazione di stime bootstrap e di confrontarne i risultati con quelli ottenuti tramite l'applicazione dei metodi bootstrap tradizionali usati singolarmente. A tal fine, sono stati svolti degli studi Monte Carlo. In particolare, nella Sezione 4.1 verranno presentati i risultati dell'esperimento Monte Carlo relativo al bootstrap ibrido, mentre nella Sezione 4.2 quelli relativi alla combinazione di stime bootstrap. Questi approcci rappresentano una novità nel campo della stima dei parametri di serie storiche, e offrono una prospettiva inedita su come migliorare la robustezza e l'affidabilità delle stime. Questo capitolo, dunque, non solo valuta le *performance* di tali metodi, ma ne esplora anche le potenzialità innovative rispetto alle pratiche consolidate.

Sono stati considerati, in tutte le simulazioni, i metodi bootstrap precedentemente presentati nel Capitolo 2 ed utilizzati per lo sviluppo delle procedure nel Capitolo 3: bootstrap parametrico (PB), *wild* bootstrap (WB), *sieve* bootstrap (SB), *moving block* bootstrap (MBB), *circular block* bootstrap (CBB), *pre-filtered sieve* bootstrap (PFSB) e *pre-filtered moving block* bootstrap (PFMBB).

Il disegno Monte Carlo implementato in questo lavoro segue quello utilizzato da Arteche (2024) per analizzare l'applicabilità delle strategie bootstrap nel contesto delle serie storiche a memoria lunga. In particolare, vengono considerate $S = 1000$ repliche indipendenti da diversi modelli $ARFIMA(0, d, 0)$ stazionari della forma:

$$(1 - L)^d X_t = u_t, \quad t = 1, 2, \dots, T, \quad (4.1)$$

dove u_t segue un processo *white noise*. Vengono considerate come diverse numerosità campionarie, $T = 250$ e $T = 500$, e due diversi valori per il parametro di memoria lunga, $d = 0.2$ e $d = 0.4$, per i quali il PGD è stazionario. Il numero di replicazioni bootstrap è pari a $B = 500$.

Per stimare il parametro di memoria lunga d vengono utilizzati tre noti stimatori: lo stimatore del log-periodogramma (GPH), lo stimatore di Whittle locale (LWE) e lo stimatore di massima verosimiglianza (MLE).

Lo stimatore del log-periodogramma (GPH), \hat{d}_{GPH} , è ottenuto tramite una regressione semplice con i minimi quadrati, facendo riferimento al modello lineare:

$$\log I_y(\lambda_j) = a + dX_j + u_t, \quad j = 1, 2, \dots, m$$

dove $X_j = -2 \log \lambda_j$ e $I_y(\lambda_j) = |W_y(\lambda_j)|^2$ è il periodogramma definito sulle frequenze di Fourier λ_j , $\lambda_j = 2\pi j/T$, $j = 1, \dots, [T/2]$ (Geweke e Porter-Hudak, 1983). La larghezza di banda m indica il numero di frequenze usate nella stima ed è vincolata a soddisfare: $m^{-1} + mT^{-1} \rightarrow 0$ per $T \rightarrow \infty$; questo vincolo impone che, anche quando $m \rightarrow \infty$, la banda delle frequenze usate nella stima converga a zero. Sotto determinate condizioni, le proprietà di \hat{d}_{GPH} sono ben stabilite (Robinson, 1995; Hurvich et al., 1998), in particolare \hat{d}_{GPH} è consistente per $-1/2 < d \leq 1$ ed asintoticamente normale per $-1/2 < d < 3/4$. Inoltre, per $-1/2 < d < 3/4$, è noto che:

$$\sqrt{m}(\hat{d}_{GPH} - d) \rightarrow N\left(0, \frac{\pi^2}{24}\right) \quad \text{per } T \rightarrow \infty. \quad (4.2)$$

Arteche e Orbe (2005) hanno proposto un metodo bootstrap basato sui residui ed hanno mostrato, usando un'analisi Monte Carlo, che, nel caso di piccoli campioni, l'approssimazione bootstrap della distribuzione dello GPH è più accurata della sua distribuzione asintotica. Tuttavia, questo risultato non è stato ancora dimostrato teoricamente. Nelle simulazioni svolte in questo Capitolo, si utilizza, come valore per l'ampiezza di banda, $m_{GPH} = 1 + T^{0.8}$, coerentemente con la relativa letteratura.

Lo stimatore di Whittle locale (LWE) del parametro di memoria lunga, \hat{d}_{LWE} , è ottenuto minimizzando la funzione:

$$\log(m^{-1} \sum \lambda_j^{2d} I_y(\lambda_j)) - m^{-1} 2d \sum \log \lambda_j, \quad j = 1, 2, \dots, m,$$

dove m indica la larghezza di banda (Fox e Taqqu, 2003). Le proprietà asintotiche dello stimatore LWE sono molto simili a quelle dello stimatore GPH (Robinson, 1995): se m soddisfa

la condizione per cui $m^{-1} + mT^{-1} \rightarrow 0$ per $T \rightarrow \infty$, allora \hat{d}_{LWE} è consistente per $-1/2 < d \leq 1$ ed asintoticamente normale per $-1/2 < d < 3/4$. Inoltre, per $-1/2 < d < 3/4$, è noto che:

$$\sqrt{m}(\hat{d}_{LWE} - d) \rightarrow N(0, \frac{1}{4}) \quad \text{per } T \rightarrow \infty. \quad (4.3)$$

La validità delle strategie bootstrap considerate in questo studio per approssimare la distribuzione dello stimatore LWE è, inoltre, giustificata anche teoricamente (Poskitt, 2008; Poskitt et al., 2015; Kapetanios et al., 2019). Nelle simulazioni svolte in questo Capitolo, si utilizza, come valore per l'ampiezza di banda, $m_{LWE} = 1 + T^{0.65}$, seguendo la relativa letteratura.

Lo stimatore di massima verosimiglianza (MLE), \hat{d}_{MLE} , (Beran, 1995), rispetto ai precedenti stimatori semiparametrici, è uno stimatore parametrico e necessita, quindi, della specificazione del modello *ARFIMA*. Nelle simulazioni svolte in questo Capitolo, la verosimiglianza è approssimata utilizzando il metodo proposto da Haslett e Raftery (1989), in quanto risulta essere veloce ed accurato. Dalla teoria asintotica (Beran, 1995), inoltre, è noto che, nel caso di processi *ARFIMA*(0, d , 0), sia stazionari che non, e di processi *ARFIMA*(p , d , q) con parametri *ARMA* fissati, vale la seguente equazione:

$$\lim_{T \rightarrow \infty} \text{var}(\hat{d}_{MLE}) = \frac{4\pi}{\int_{-\pi}^{\pi} \log(2 - 2 \cos \lambda)^2 d\lambda} = \frac{6}{\pi^2}. \quad (4.4)$$

Per quanto riguarda l'implementazione dei metodi bootstrap e la selezione dei parametri a loro relativi, sono di seguito riportate le scelte effettuate. Per l'implementazione del PB, il modello *ARFIMA*(0, d , 0) e i parametri a questo associati vengono stimati attraverso la massima verosimiglianza. Per quanto riguarda il WB, vengono utilizzati, come residui $\hat{\epsilon}_t$, i residui derivati dal modello *ARFIMA*(0, d , 0) stimato dal PB, mentre le variabili che costituiscono i fattori casuali w_t seguono una distribuzione normale standard $N(0, 1)$. Anche per il SB, il modello *AR*(p), e i parametri a questo associati, vengono stimati tramite massima verosimiglianza, mentre l'ordine del modello p è selezionato attraverso minimizzazione dell'AIC. Per quanto riguarda il MBB e il CBB, la lunghezza dei blocchi k è uguale a $k = [T^{1/2}]$, come viene suggerito da Kim e Nordman (2011). Infine, per quanto riguarda i metodi PFSB e PFMBB, il pre-filtraggio avviene sempre attraverso la stima preliminare di un modello *ARFIMA*(0, d , 0) attraverso la massima verosimiglianza.

Le simulazioni sono state eseguite con il software R (Team, 2013). Il codice utilizzato per le analisi è disponibile su richiesta, contattando me o la Professoressa Luisa Bisaglia.

4.1 Bootstrap ibrido

In questa Sezione viene descritto l'esperimento Monte Carlo svolto con il fine di valutare le *performance* delle metodologie di bootstrap ibrido proposte, e confrontare i risultati da queste ottenuti con quelli relativi ai metodi bootstrap tradizionali, usati singolarmente. Le metodologie di bootstrap ibrido analizzate sono quelle trattate nella Sezione 3.1, alle quali si fa riferimento come: *mean*, *median* e *trimmean* per indicare, rispettivamente, i metodi che utilizzano, per combinare le serie bootstrap, gli schemi di combinazione basati su media, mediana e media troncata, e *reglin*, *corr*, *rMSE* e *rMAE* per indicare, rispettivamente, i metodi che utilizzano, per combinare le serie bootstrap, gli schemi di combinazione i cui pesi sono basati su regressione lineare, correlazione, reciproco di MSE e reciproco di MAE. Le *performance* delle diverse strategie bootstrap vengono analizzate basandosi su due diverse misure comparative (Arteche, 2024): la radice della deviazione quadratica media della funzione di densità di probabilità bootstrap rispetto alla densità reale, e la frequenza di copertura degli intervalli di confidenza (ottenuti o tramite le stime bootstrap ricavate dai diversi modelli, o utilizzando le distribuzioni asintotiche). Entrambe le misure sono ottenute con $S = 1000$ replicazioni indipendenti delle serie ottenute dai PGD descritti all'inizio del Capitolo.

La radice della deviazione quadratica media (RMSD) è ottenuta come:

$$RMSD(\text{boot}) = \sqrt{\frac{1}{S} \sum_{j=1}^S (p_{\text{boot}}(\hat{d}_j - d_0) - p_{\text{mc}}(\hat{d}_j - d_0))^2},$$

dove *boot* rappresenta una delle strategie bootstrap tradizionale o una delle strategie bootstrap ibride discusse nella Sezione 3.1, d_0 rappresenta il vero valore del parametro di memoria lunga, $p_{\text{mc}}(\hat{d}_j - d_0)$ è l'ordinata della stima di densità basata sul kernel, ottenuta con le S stime, $\hat{d}_k - d_0$, $k = 1, \dots, S$ (valutate su $\hat{d}_j - d_0$), mentre $p_{\text{boot}}(\hat{d}_j - \hat{d}_0)$ è la media sulle S simulazioni delle ordinate delle stime della densità kernel ottenuta usando B stime bootstrap $\hat{d}_j^b - \hat{d}_j$, $b = 1, \dots, B$ (valutate su $\hat{d}_j - d_0$) (Poskitt et al., 2015). Più questo valore è vicino allo zero, migliore è considerata la *performance* del metodo.

Gli intervalli di confidenza (CI) per il parametro di memoria lunga ottenuti con le diverse strategie bootstrap e bootstrap ibride sono, invece, costruiti come:

$$CI_{1-\alpha}^*(\hat{d}_a) = (\hat{d}_{a|\alpha/2}^*; \hat{d}_{a|1-\alpha/2}^*)$$

dove \hat{d}_α^* denota il $B\alpha$ -esimo valore ordinato delle stime bootstrap di d , per uno degli stimatori $a = LWE, GPH, MLE$.

Vengono considerate, inoltre, per completezza, anche le *performance* relative alle distribuzioni normali asintotiche dei tre stimatori. Gli intervalli di confidenza (CI) per il parametro di memoria lunga, basati sulla distribuzione asintotica di uno dei tre stimatori considerati, per un livello di confidenza $1 - \alpha$, sono della forma:

$$CI_{1-\alpha}(\hat{d}_a) = (\hat{d}_a - \sqrt{\widehat{var}(\hat{d}_a)}z_{\alpha/2}; \hat{d}_a + \sqrt{\widehat{var}(\hat{d}_a)}z_{1-\alpha/2})$$

per $a = LWE, GPH, MLE$, dove z_α è il 100α -esimo percentile della distribuzione normale standard $N(0, 1)$. Come valori per la stima delle varianze asintotiche dei tre stimatori, si fa riferimento alla letteratura correlata, in particolare alle varianze asintotiche riportate in 4.2, 4.3 e 4.4.

Sono di interesse le discrepanze di copertura rispetto ad un livello di confidenza nominale, come suggerito da de Peretti e Siani (2010). Le discrepanze di copertura sono definite, per ogni stimatore, come segue. Per una copertura nominale x ed un totale di S simulazioni, la discrepanza di copertura è calcolata come $\hat{F}(x) - x$, dove $\hat{F}(x) = S^{-1} \sum_{s=1}^S I(x_s \leq x)$ per $x_s = 1 - 2 \min(pv, 1 - pv)$, con pv funzione di distribuzione di probabilità stimata per il vero parametro di memoria lunga d_0 . La funzione di distribuzione di probabilità, pv , può essere calcolata o tramite la distribuzione asintotica dello stimatore a cui si fa riferimento, o tramite la distribuzione empirica delle stime del parametro di memoria lunga, ottenute mediante un determinato metodo bootstrap individuale o bootstrap ibrido. Pertanto, $\hat{F}(x)$ può essere interpretata come la frequenza di copertura del vero valore del parametro di memoria lunga d_0 degli intervalli di confidenza, aventi un livello di confidenza nominale x , costruiti a partire dalle distribuzioni empiriche delle stime. Una maggiore vicinanza della discrepanza di copertura a zero indica una maggiore precisione della strategia utilizzata per approssimare la distribuzione di d_0 . Viene posta particolare attenzione all'intervallo di confidenza di livello 0.95, spesso usato, in letteratura, per l'inferenza. Per analizzare le *performance* di un intervallo di confidenza, non è importante solo la frequenza di copertura, ma anche l'ampiezza è una misura rilevante della sua efficacia (de Peretti e Siani, 2010). Una copertura più elevata è solitamente associata ad intervalli più ampi ma, in alcuni casi, intervalli più stretti sono accompagnati da elevata copertura, indicando una precisione più alta della stima. Di conseguenza, se due intervalli di confidenza hanno lo stesso livello di confidenza, quello meno ampio è preferibile. Per poter comprendere a pieno i risultati, inoltre, è importante chiarire che il PGD considerato negli esperimenti è lo stesso del modello stimato dall'algoritmo del bootstrap parametrico; ciò implica che il PB è implementato in quello che si può considerare il miglior scenario per lo stesso e dovrebbe, quindi, offrire le migliori *performance*. La motivazione dietro a questa scelta è di

avere un limite superiore delle *performance* ideali con le quali poter comparare gli altri metodi bootstrap. Poiché lo scenario esaminato risulta essere particolarmente favorevole al PB, ma meno ai metodi bootstrap ibridi proposti, sono necessari, in futuro, ulteriori studi Monte Carlo che considerino anche situazioni meno svantaggiose per questi ultimi, al fine di valutarne appieno le *performance* in condizioni più realistiche.

La Tabella 4.1 mostra RMSD, frequenza di copertura (*Cov Freq*) e ampiezza degli intervalli di confidenza (*Ampl*) ottenuti dall'esperimento Monte Carlo in cui sono state simulate $S = 1000$ serie storiche da un modello $ARFIMA(0, 0.2, 0)$, utilizzando, rispettivamente, $T = 250$ per la tabella sopra, e $T = 500$ per la tabella sotto. Nelle colonne relative al RMSD si vedono evidenziati in grassetto i valori relativi al metodo che, per quel determinato stimatore, ha avuto la *performance* migliore in termini di RMSD. Le Figure 4.1 e 4.2 riportano una rappresentazione grafica di RMSD e della discrepanza della copertura con un livello di confidenza del 95% (*excess Cov Freq*), per ogni metodo, così che sia possibile apprezzare maggiormente le *performance* e le differenze dei vari metodi bootstrap e bootstrap ibridi, sia per $T = 250$ che per $T = 500$.

Per quanto riguarda i risultati relativi allo stimatore MLE, il PB risulta ottenere la RMSD minore tra tutti i metodi considerati, ed offre, inoltre, un'ottima copertura: ciò è dovuto allo scenario favorevole in cui si trova questo metodo. Gli altri metodi bootstrap tradizionali, infatti, non si comportano altrettanto bene. I metodi bootstrap ibridi proposti (non considerando quello relativo alla combinazione tramite mediana (*median*)) mostrano delle *performance* generalmente superiori, sia in termini di RMSD che in termini di copertura, rispetto a quelle dei metodi bootstrap tradizionali, in particolare quando aumenta la lunghezza T della serie storica.

Concentrandoci, invece, sui risultati relativi allo stimatore LWE, i metodi bootstrap non parametrici (MBB, CBB e PFMBB) risultano avere una RMSD minore rispetto agli altri metodi bootstrap. Allo stesso tempo, però, questi stessi metodi offrono una copertura poco soddisfacente rispetto a quella generata dalle altre metodologie. Alcuni dei metodi bootstrap ibridi mostrano, per questo stimatore, dei risultati molto simili a quelli dei metodi bootstrap non-parametrici, anche se portano a valori della RMSD leggermente superiori, ed offrono una copertura decisamente migliore. In particolar modo, i metodi bootstrap ibridi i cui schemi di combinazioni sono basati su correlazione (*corr*), regressione lineare (*reglin*) e media semplice (*mean*) offrono delle ottime *performance*.

Infine, studiando i risultati relativi allo stimatore GPH, è possibile notare che, quando

$T = 250$, il metodo che offre la *performance* migliore in termini di RMSD è il metodo bootstrap ibrido *corr* seguito dal PFMBB, mentre, per $T = 500$, le posizioni in classifica occupate da questi due metodi risultano invertite. Si può, inoltre, notare che, in termini di copertura, in entrambi i casi il metodo bootstrap ibrido riesce a dare dei risultati migliori, risultando quindi preferibile. In generale, anche in questo caso, i metodi bootstrap ibridi sembrano offrire una garanzia migliore in termini di *performance* generali rispetto ai metodi bootstrap tradizionali, i cui risultati sono caratterizzati da un'incertezza maggiore.

Similmente, la Tabella 4.2 mostra RMSD, frequenza di copertura (*Cov Freq*) e ampiezza degli intervalli di confidenza (*Ampl*) ottenuti dall'esperimento Monte Carlo in cui sono state simulate $S = 1000$ serie storiche da un modello $ARFIMA(0, 0.4, 0)$, utilizzando sempre $T = 250$ e $T = 500$. I grafici riportati nelle Figure 4.3 e 4.4 rappresentano, quindi, RMSD e discrepanza della copertura con un livello di confidenza del 95% (*excess Cov Freq*) per gli stessi metodi.

In queste simulazioni, gli ottimi risultati raggiunti dai metodi bootstrap ibridi proposti vengono maggiormente messi in evidenza. Infatti, nonostante il miglior metodo bootstrap in termini di RMSD, per lo stimatore MLE, risulti sempre essere il PB, sia per lo stimatore LWE che GPH, la *performance* migliore in termini di RMSD, quando $T = 250$, viene raggiunta dal metodo bootstrap ibrido la cui combinazione si basa sulla correlazione (*corr*). Per $T = 500$ è possibile osservare che sono numerosi i metodi bootstrap ibridi che offrono degli ottimi risultati in termini di RMSD, sia per lo stimatore LWE che GPH, come i metodi le cui combinazioni sono basate su media semplice (*mean*), correlazione (*corr*), reciproco di MSE (*rMSE*) e reciproco di MAE (*rMAE*). Tra questi metodi, in termini di copertura, spicca sempre il metodo bootstrap ibrido *corr*, che ottiene, però, intervalli di confidenza di ampiezza leggermente maggiore rispetto agli altri.

I risultati appena analizzati mostrano come, anche in questo ambito, le combinazioni di più metodi possono portare ad un effettivo vantaggio. Il beneficio principale delle combinazioni non è tanto quello di ottenere sempre la migliore stima del parametro, ma quello di poter contare, con una fiducia maggiore, su dei risultati che a questo si avvicinano, riuscendo a compensare l'incertezza dovuta all'ignoranza del PGD (e di conseguenza, l'incertezza riguardo quale possa essere il miglior metodo bootstrap per il caso specifico). Inoltre, le analisi evidenziano che, all'aumentare della numerosità campionaria e, soprattutto, con l'incremento del valore del

parametro di memoria lunga, i metodi bootstrap ibridi presentano un netto miglioramento delle loro *performance*. Questo suggerisce che, in condizioni di maggiore complessità e dipendenza a lungo termine, i metodi proposti offrono vantaggi significativi rispetto ai singoli approcci bootstrap.

Negli studi di simulazione, è stato, inoltre, possibile riconoscere quali, tra i metodi bootstrap ibridi analizzati, offrono *performance* migliori in termini di RMSD e di copertura degli intervalli di confidenza. Le combinazioni lineari più “s sofisticate” (*corr*, *rMSE*, *rMAE*, *reglin*) tendono a raggiungere dei risultati migliori, tuttavia, anche in questo contesto, la combinazione basata sulla media aritmetica (*mean*), nonostante la sua semplicità, riesce ad ottenere dei risultati simili a quelli degli altri metodi più complessi. Questo ragionamento, invece, non sembra essere valido per il metodo bootstrap ibrido basato sulla combinazione tramite mediana (*median*). Infatti, in tutti gli scenari considerati, il metodo bootstrap ibrido basato sulla combinazione tramite mediana (*median*) porta a risultati poco desiderabili, sia in termini di RMSD che in termini di copertura. Un’ultima osservazione, infine, riguarda l’ampiezza degli intervalli di confidenza dei metodi bootstrap ibridi. Gli intervalli di confidenza di questi metodi risultano essere, generalmente, leggermente più ampi rispetto a quelli dei metodi bootstrap tradizionali, tuttavia, questa debolezza viene largamente superata dal vantaggio generale che ne trae la copertura.

Tabella 4.1: RMSD, frequenza di copertura (*Cov Freq*) e ampiezza del CI (*Ampl*) per un $ARFIMA(0,0.2,0)$ con $T = 250$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.

T=250									
	LWE			GPH			MLE		
<i>Metodo</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>
Mean	0,869	0,929	0,391	1,250	0,985	0,315	0,808	0,868	0,202
Median	1,244	0,832	0,390	1,970	0,952	0,315	3,162	0,557	0,191
Trimmean	0,943	0,915	0,392	1,381	0,977	0,316	0,993	0,808	0,201
Reglin	0,814	0,941	0,389	1,244	0,983	0,311	0,791	0,854	0,198
Corr	0,789	0,950	0,409	1,219	0,985	0,319	0,906	0,866	0,204
rMSE	0,876	0,928	0,391	1,261	0,982	0,315	0,771	0,866	0,202
rMAE	0,872	0,930	0,391	1,253	0,981	0,315	0,787	0,870	0,202
PB	1,120	0,981	0,392	1,317	0,990	0,316	0,543	0,85	0,202
WB	3,954	0,912	0,253	5,461	0,957	0,234	10,769	0,649	0,108
SB	1,550	0,508	0,394	1,569	0,952	0,315	1,130	0,846	0,204
MBB	0,766	0,780	0,382	1,372	0,966	0,306	1,213	0,812	0,195
CBB	0,771	0,777	0,382	1,371	0,965	0,307	1,196	0,810	0,194
PFSB	1,153	0,981	0,392	1,422	0,989	0,315	1,749	0,864	0,202
PFMBB	0,791	0,907	0,384	1,235	0,980	0,309	1,445	0,852	0,194
T=500									
	LWE			GPH			MLE		
<i>Metodo</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>
Mean	1,174	0,945	0,300	1,898	0,985	0,228	0,914	0,905	0,143
Median	1,811	0,796	0,299	3,233	0,904	0,228	6,220	0,405	0,143
Trimmean	1,299	0,923	0,300	2,196	0,981	0,228	2,015	0,847	0,143
Reglin	1,100	0,966	0,301	1,872	0,990	0,227	0,948	0,898	0,140
Corr	1,047	0,965	0,321	1,863	0,992	0,232	0,979	0,898	0,144
rMSE	1,180	0,945	0,300	1,921	0,983	0,229	0,900	0,901	0,143
rMAE	1,178	0,946	0,300	1,910	0,984	0,228	0,903	0,902	0,143
PB	1,558	0,994	0,300	2,022	0,995	0,228	0,702	0,902	0,142
WB	5,605	0,913	0,181	8,526	0,947	0,151	18,306	0,590	0,062
SB	2,399	0,423	0,302	2,361	0,879	0,228	1,439	0,869	0,147
MBB	1,028	0,805	0,294	2,093	0,973	0,224	1,551	0,869	0,139
CBB	1,047	0,810	0,293	2,086	0,970	0,224	1,551	0,869	0,139
PFSB	1,577	0,994	0,300	2,073	0,993	0,229	2,189	0,892	0,142
PFMBB	1,040	0,949	0,294	1,791	0,987	0,225	1,619	0,903	0,138

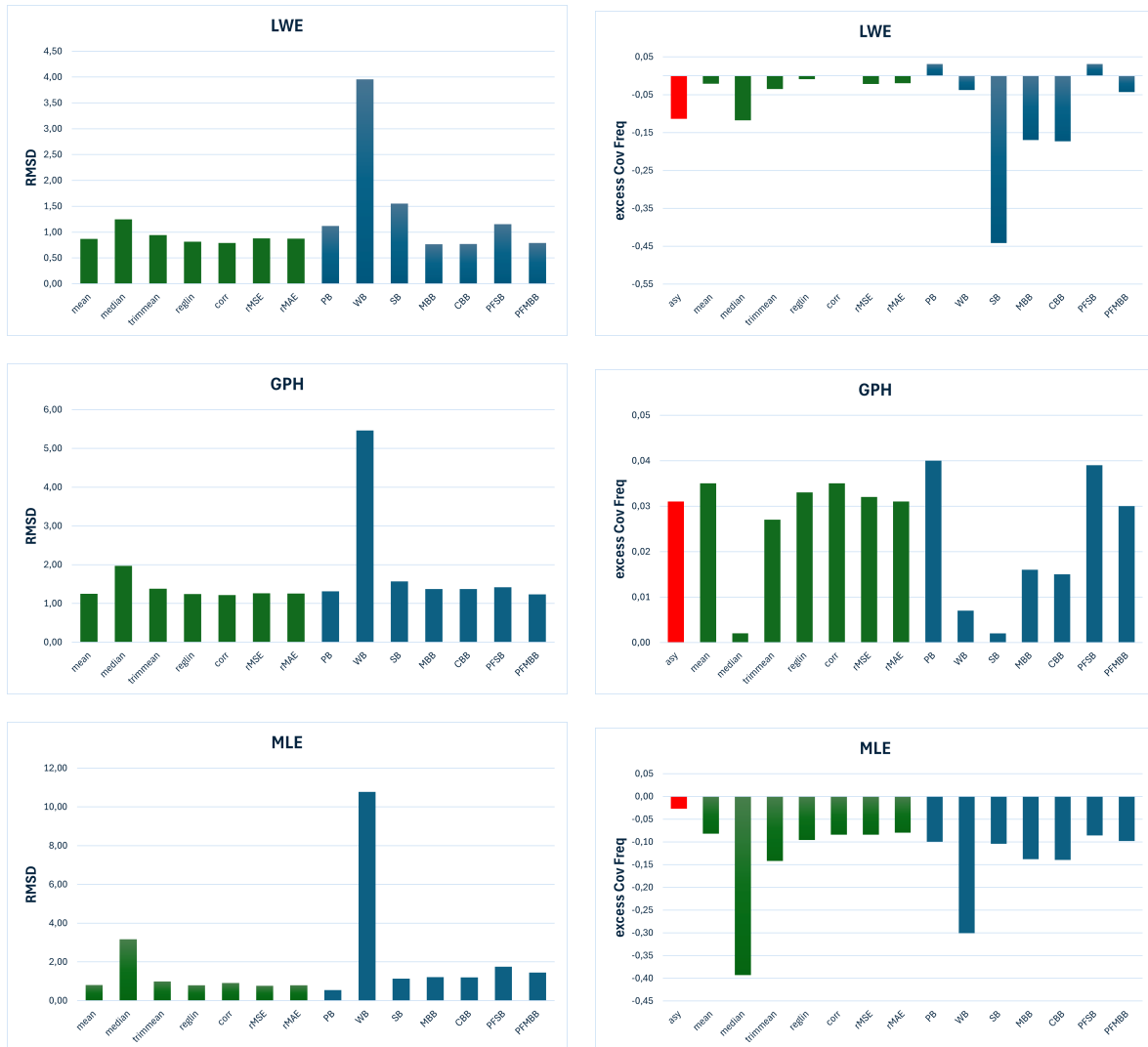


Figura 4.1: RMSD e discrepanza dalla copertura con un livello di confidenza del 95% (*excess Cov Freq*) per un $ARFIMA(0, 0.2, 0)$ e $T = 250$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.

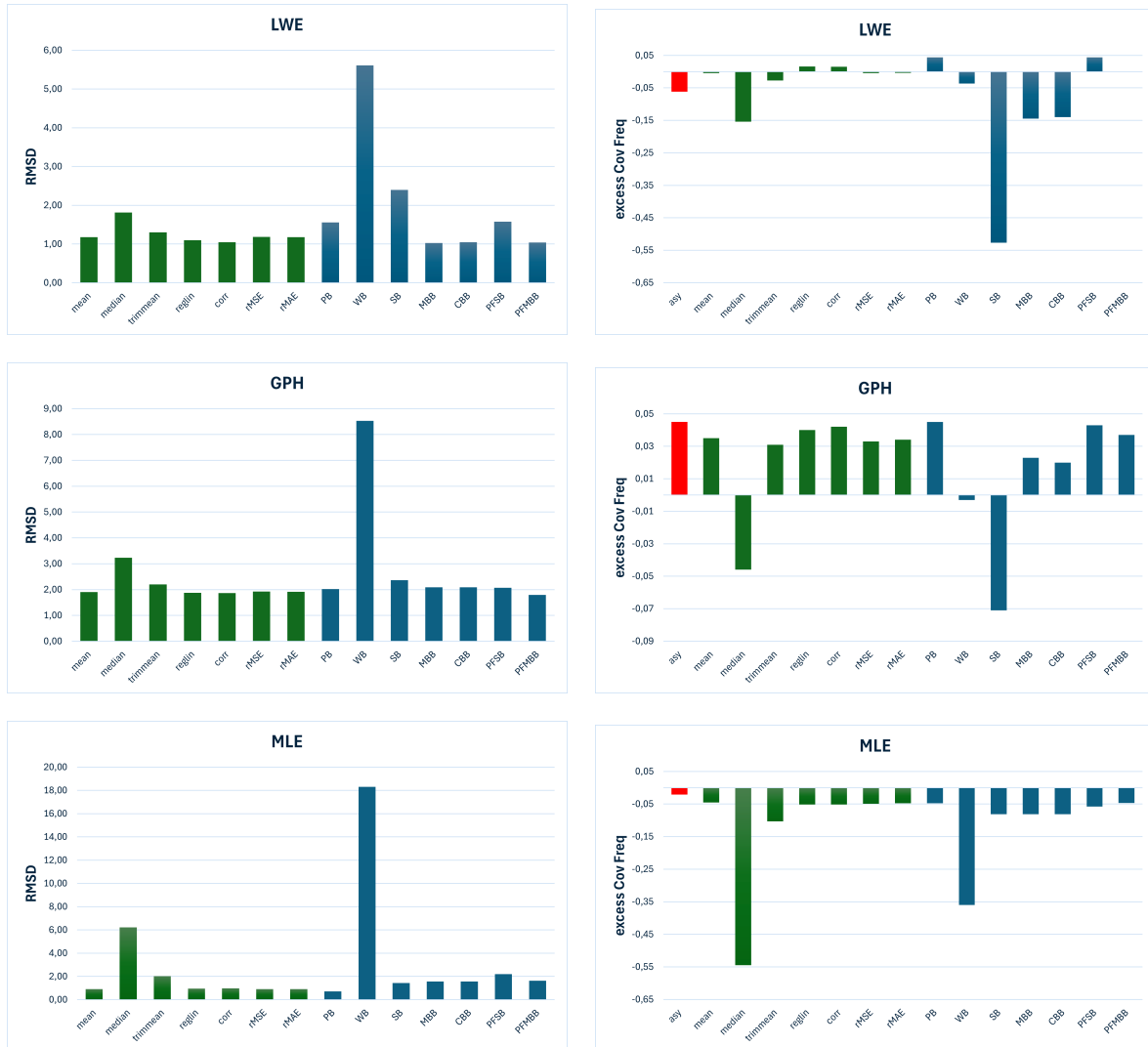


Figura 4.2: RMSD e discrepanza dalla copertura con un livello di confidenza del 95% (*excess Cov Freq*) per un $ARFIMA(0, 0.2, 0)$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.

Tabella 4.2: RMSD, frequenza di copertura (*Cov Freq*) e ampiezza del CI (*Ampl*) per un $ARFIMA(0,0.4,0)$ con $T = 250$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.

T=250									
	LWE			GPH			MLE		
<i>Metodo</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>
Mean	0,726	0,896	0,392	1,461	0,979	0,314	1,595	0,860	0,183
Median	1,550	0,589	0,394	3,072	0,758	0,320	5,133	0,117	0,210
Trimmean	0,851	0,848	0,393	1,836	0,970	0,315	1,119	0,730	0,192
Reglin	0,853	0,885	0,380	1,613	0,971	0,310	1,332	0,804	0,183
Corr	0,671	0,956	0,413	1,352	0,987	0,321	1,814	0,880	0,185
rMSE	0,738	0,873	0,393	1,539	0,974	0,315	1,442	0,844	0,185
rMAE	0,727	0,884	0,392	1,496	0,974	0,315	1,515	0,854	0,184
PB	1,078	0,977	0,396	1,620	0,992	0,317	0,717	0,809	0,190
WB	2,771	0,925	0,308	4,110	0,962	0,249	9,759	0,650	0,122
SB	1,111	0,723	0,404	1,407	0,955	0,315	3,141	0,903	0,178
MBB	0,955	0,713	0,384	1,856	0,946	0,307	2,129	0,799	0,183
CBB	0,962	0,708	0,384	1,863	0,947	0,307	2,101	0,791	0,183
PFSB	1,121	0,972	0,395	1,596	0,989	0,317	2,162	0,854	0,184
PFMBB	0,806	0,919	0,385	1,397	0,977	0,311	1,929	0,826	0,178
T=500									
	LWE			GPH			MLE		
<i>Metodo</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>	<i>RMSD</i>	<i>Cov Freq</i>	<i>Ampl</i>
Mean	0,897	0,907	0,301	1,883	0,985	0,229	2,582	0,906	0,132
Median	2,081	0,562	0,303	4,815	0,589	0,232	9,039	0,030	0,147
Trimmean	0,997	0,854	0,302	2,484	0,969	0,229	1,995	0,783	0,137
Reglin	1,147	0,920	0,297	2,097	0,973	0,227	1,771	0,882	0,132
Corr	0,982	0,975	0,325	1,815	0,994	0,234	2,659	0,918	0,134
rMSE	0,853	0,870	0,301	1,979	0,983	0,229	2,406	0,900	0,133
rMAE	0,861	0,883	0,301	1,926	0,982	0,229	2,497	0,902	0,133
PB	1,545	0,992	0,303	2,119	0,995	0,231	0,919	0,872	0,135
WB	3,830	0,956	0,226	5,697	0,967	0,164	17,627	0,629	0,073
SB	1,585	0,700	0,313	1,966	0,954	0,228	5,133	0,925	0,131
MBB	1,131	0,715	0,299	2,462	0,947	0,223	3,544	0,860	0,132
CBB	1,134	0,727	0,299	2,445	0,949	0,223	3,616	0,872	0,133
PFSB	1,570	0,990	0,303	2,173	0,990	0,230	2,900	0,881	0,132
PFMBB	1,120	0,940	0,297	1,957	0,986	0,225	2,430	0,869	0,128

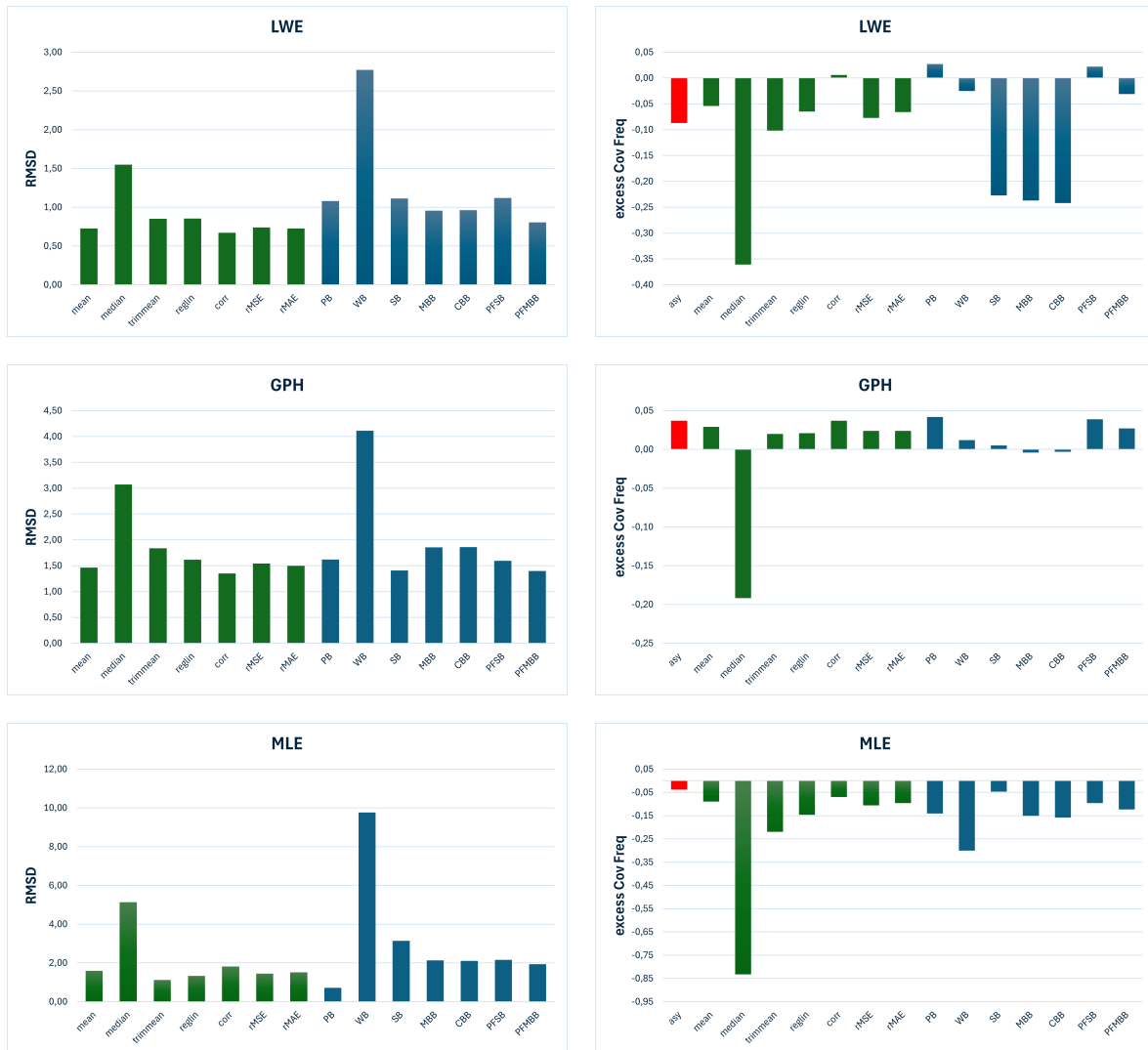


Figura 4.3: RMSD e discrepanza dalla copertura con un livello di confidenza del 95% (*excess Cov Freq*) per un $ARFIMA(0, 0.4, 0)$ e $T = 250$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.



Figura 4.4: RMSD e discrepanza dalla copertura con un livello di confidenza del 95% (*excess Cov Freq*) per un $ARFIMA(0, 0.4, 0)$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi bootstrap ibridi.

4.2 Combinazione di stime bootstrap

In questa Sezione viene descritto l'esperimento Monte Carlo svolto con il fine di valutare la *performance* dei metodi di combinazione delle stime bootstrap proposti, e confrontare i risultati da questi ottenuti con quelli relativi ai metodi bootstrap tradizionali, precedentemente trattati. Le metodologie di combinazione di stime bootstrap analizzate sono, quindi, quelle trattate nella Sezione 3.2, alle quali si fa riferimento come: *mean*, *median* e *trimmean* per, rispettivamente, i metodi che utilizzano gli schemi di combinazione basati su media, mediana e media troncata, e *rMSE* e *rMAE* per, rispettivamente, i metodi che utilizzano gli schemi di combinazione i cui pesi sono basati sul reciproco di MSE e reciproco di MAE. In questo caso, le *performance* degli stimatori vengono valutate con misure più tradizionali, quali la radice dell'errore quadratico medio (RMSE) e la distorsione. Entrambe le misure sono ottenute con $S = 1000$ repliche indipendenti delle serie ottenute dal PGD descritto all'inizio del Capitolo.

La radice dell'errore quadratico medio (RMSE) è ottenuta come:

$$RMSE(\text{boot}) = \sqrt{\frac{1}{S} \sum_{j=1}^S (\hat{d}_j - d_0)^2},$$

mentre la distorsione è ottenuta come:

$$Dist(\text{boot}) = \frac{1}{S} \sum_{j=1}^S (\hat{d}_j - d_0),$$

dove *boot* rappresenta una delle strategie bootstrap tradizionali o una delle strategie di combinazione delle stime bootstrap discusse nella sezione 3.2, d_0 rappresenta il vero valore del parametro di memoria lunga e \hat{d}_j la stima finale della statistica combinata S^{**} calcolata alla j -esima replicazione Monte Carlo. Più i valori di RMSE e distorsione sono vicini allo zero, migliore è considerata la *performance* del metodo.

La Tabella 4.3 mostra RMSE e distorsione (*Dist*) ottenuti dall'esperimento Monte Carlo in cui sono state simulate $S = 1000$ serie storiche da un modello $ARFIMA(0, 0.2, 0)$, utilizzando, rispettivamente, le numerosità campionarie $T = 250$ per la tabella sopra, e $T = 500$ per la tabella sotto. Nelle colonne relative al RMSE si vedono evidenziati in grassetto i valori relativi al metodo che, per quel determinato stimatore, ha avuto la *performance* migliore in termini di RMSE. Similmente a quanto fatto per gli studi di simulazione per i metodi bootstrap ibridi, viene riportata la rappresentazione grafica di RMSE e distorsione nelle Figure 4.5 e 4.6, così

da poter confrontare meglio i metodi bootstrap e le combinazioni di stime bootstrap, sia per $T = 250$ che per $T = 500$.

Sono interessanti i risultati riguardo la distorsione delle stime dei vari metodi. Infatti, la distorsione risulta essere, in ogni scenario, per ogni metodo e per ogni stimatore, sempre negativa, cioè tutti i metodi bootstrap e tutte le combinazioni di stime bootstrap portano ad una sottostima del parametro di memoria lunga. Sia per $T = 250$ che per $T = 500$, è possibile notare che, tra i metodi bootstrap individuali, quelli pre-filtrati portano generalmente ad una distorsione minore. Le strategie basate sulle combinazioni delle stime bootstrap sembrano comunque portare, in termini di distorsione, a buoni risultati, offrendo un compromesso efficace tra i metodi individuali più performanti e quelli meno, come il MBB ed il CBB.

Contrariamente a quanto visto nei risultati relativi allo studio Monte Carlo descritto nella Sezione precedente, il metodo PB non ottiene buone *performance*, in termini di RMSE, per lo stimatore MLE. Concentrandosi sui risultati relativi allo stimatore MLE, infatti, la strategia bootstrap che ottiene un RMSE minore è quella relativa al metodo PFMBB. Per $T = 250$, inoltre, seguono direttamente, in questa classifica, il metodo WB ed i metodi di combinazione di stime bootstrap che si basano sul reciproco di MSE e di MAE (*rMSE* e *rMAE*). Per gli altri due stimatori (LWE e GPH), il metodo migliore, in termini di RMSE, risulta essere il metodo bootstrap WB, che nel precedente caso forniva le prestazioni peggiori. In generale, e soprattutto con l'aumentare della lunghezza della serie storica, i metodi basati sulle combinazioni delle stime bootstrap hanno dei buoni risultati, in termini di RMSE, anche se non sempre occupano il primo posto nella classifica, mentre la *performance* dei metodi bootstrap tradizionali è molto variabile ed incerta. Infatti, nonostante alcuni dei metodi bootstrap originali portino a dei risultati preferibili ed altri offrano prestazioni significativamente inferiori, i metodi di combinazione di stime bootstrap riescono a bilanciare efficacemente le prestazioni complessive di questi metodi individuali. Un'altra differenza rispetto a quanto visto precedentemente è che, in questo caso, non risultano esserci combinazioni che spiccano per *performance* rispetto ad altre, ed anche la strategia di combinazione delle stime bootstrap tramite mediana (*median*) risulta essere una valida alternativa, in alcuni casi offrendo anche risultati migliori rispetto a quelli dati dalla strategia di combinazione di stime bootstrap tramite media (*mean*).

La Tabella 4.4 mostra, RMSE e distorsione (*Dist*) ottenuti dall'esperimento Monte Carlo in cui sono state simulate $S = 1000$ serie storiche da un modello $ARFIMA(0, 0.4, 0)$, utilizzando sempre $T = 250$ e $T = 500$. I grafici nelle Figure 4.7 e 4.8 rappresentano, quindi, RMSE e

distorsione ($Dist$) per gli stessi metodi.

Anche per questo valore del parametro di memoria lunga, con cui la persistenza della memoria aumenta, i risultati portano a conclusioni simili a quelle viste per $d = 0.2$. In questo caso, però, le strategie di combinazioni delle stime bootstrap risultano ottenere delle *performance* particolarmente buone, sia in termini di RMSE che in termini di distorsione. Le strategie di combinazione, infatti, portano a dei risultati migliori, specialmente rispetto ai metodi bootstrap non-parametrici (MBB e CBB). Per lo stimatore MLE, in particolare, i metodi di combinazione delle stime bootstrap basati su media (*mean*), reciproco di MSE (*rMSE*) e reciproco di MAE (*rMAE*) ottengono delle ottime *performance*, soprattutto con l'aumentare della numerosità campionaria, e sono secondi, in termini di RMSE, soltanto al SB. Inoltre, in questi scenari, le stime relative al metodo bootstrap SB sono le uniche che, in alcuni casi, presentano una distorsione positiva, e generalmente più bassa, in valore assoluto, rispetto a quella delle altre stime.

Le Figure 4.9 e 4.10 riportano, per ogni metodo di combinazione di stime bootstrap e metodo bootstrap individuale, e per ciascuno dei tre stimatori considerati, la media delle S stime del parametro di memoria lunga d , ottenute dalle S replicazioni Monte Carlo, rappresentata dal corrispondente punto. Gli intervalli di confidenza di Monte Carlo di livello 0.95 per queste stime, costruiti a partire dalla distribuzione empirica delle stime stesse, sono indicati con le linee verticali, i cui estremi rappresentano i quantili 0.025 e 0.975 della distribuzione empirica. Viene riportato inoltre, con la linea rossa orizzontale, il vero valore per parametro di memoria lunga d , pari, rispettivamente, a 0.2 per la Figura 4.9 e a 0.4 per la Figura 4.10. In entrambe le Figure, i grafici a sinistra si riferiscono alle simulazioni in cui $T = 250$, mentre i grafici a destra a quelle in cui $T = 500$.

Per lo stimatore MLE, le differenze tra i vari metodi di combinazione e i metodi bootstrap considerati, in termini di intervalli di confidenza Monte Carlo, non risultano essere particolarmente evidenti. Tuttavia, per gli stimatori LWE e GPH, tali differenze diventano significative, specialmente quando il parametro di memoria lunga d assume un valore più elevato. Per quanto riguarda i metodi bootstrap originali, si osserva che i metodi bootstrap parametrici (PB e WB) forniscono degli intervalli di confidenza più stretti e, di conseguenza, stime più precise, mentre ciò non accade con i metodi non parametrici (MBB, CBB e PFMBB) ed il SB. Infatti, confrontando gli intervalli di confidenza Monte Carlo di questi ultimi metodi con quelli ottenuti dalle combinazioni di stime bootstrap, le combinazioni sembrano portare ad ampiezze

più desiderabili, in quanto offrono una maggiore precisione. Inoltre, per $d = 0.4$, alcune stime del parametro di memoria lunga ottenute con il metodo SB risultano essere maggiori di $1/2$, oltrepassando così la soglia di stazionarietà delle serie storiche.

Nonostante non si siano evidenziate particolari differenze di *performance* tra le varie strategie di combinazioni delle stime bootstrap, anche in questo caso i risultati mostrano come combinare le stime dei metodi bootstrap individuali possa essere effettivamente vantaggioso, in particolare all'aumentare della numerosità campionaria e del valore del parametro di memoria lunga. I risultati ottenuti, infatti, suggeriscano che questo nuovo approccio possa portare a risultati non solo generalmente accettabili ma, eliminando il rischio di scegliere un metodo bootstrap non appropriato per il contesto, spesso anche preferibili, poiché consentono di ottenere una maggiore robustezza delle stime.

Tabella 4.3: RMSE e distorsione (*Dist*) per un $ARFIMA(0, 0.2, 0)$ con $T = 250$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

T=250						
	LWE		GPH		MLE	
<i>Metodo</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>
Mean	0,1076	-0,0894	0,0600	-0,0222	0,0643	-0,0368
Median	0,1057	-0,0834	0,0606	-0,0203	0,0644	-0,0375
Trimmean	0,1038	-0,0847	0,0596	-0,0216	0,0644	-0,0374
rMSE	0,1048	-0,0812	0,0596	-0,0200	0,0637	-0,0356
rMAE	0,1053	-0,0839	0,0596	-0,0207	0,0639	-0,0361
PB	0,0873	-0,0687	0,0569	-0,0176	0,0644	-0,0386
WB	0,0804	-0,0596	0,0567	-0,0162	0,0629	-0,0352
SB	0,1700	-0,1449	0,0791	-0,0338	0,0684	-0,0355
MBB	0,1340	-0,1086	0,0665	-0,0325	0,0678	-0,0437
CBB	0,1338	-0,1087	0,0661	-0,0323	0,0677	-0,0435
PFSB	0,0871	-0,0675	0,0591	-0,0126	0,0644	-0,0304
PFMBB	0,1033	-0,0677	0,0625	-0,0126	0,0622	-0,0309
T=500						
	LWE		GPH		MLE	
<i>Metodo</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>
Mean	0,0845	-0,0715	0,0432	-0,0158	0,0424	-0,0198
Median	0,0795	-0,0628	0,0431	-0,0138	0,0423	-0,0201
Trimmean	0,0789	-0,0649	0,0420	-0,0148	0,0423	-0,0201
rMSE	0,0811	-0,0629	0,0427	-0,0140	0,0418	-0,0191
rMAE	0,0815	-0,0653	0,0426	-0,0144	0,0420	-0,0193
PB	0,0593	-0,0462	0,0383	-0,0092	0,0419	-0,0207
WB	0,0548	-0,0402	0,0380	-0,0074	0,0409	-0,0187
SB	0,1600	-0,1412	0,0668	-0,0313	0,0470	-0,0193
MBB	0,1083	-0,0884	0,0497	-0,0247	0,0446	-0,0244
CBB	0,1084	-0,0887	0,0496	-0,0248	0,0446	-0,0244
PFSB	0,0588	-0,0454	0,0398	-0,0061	0,0425	-0,0149
PFMBB	0,0773	-0,0503	0,0442	-0,0061	0,0407	-0,0161

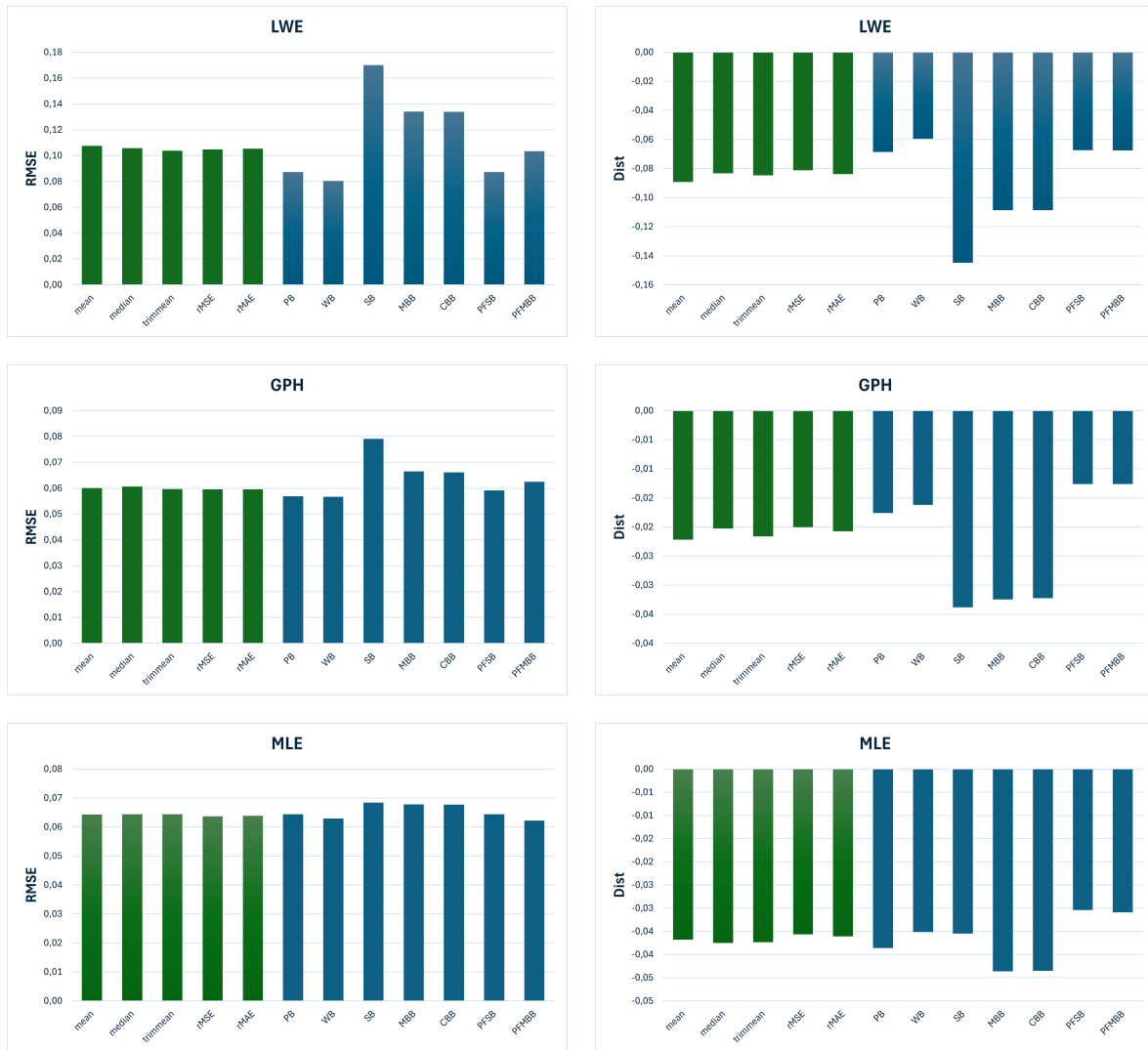


Figura 4.5: RMSE e distorsione ($Dist$) per un $ARFIMA(0,0.2,0)$ e $T = 250$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

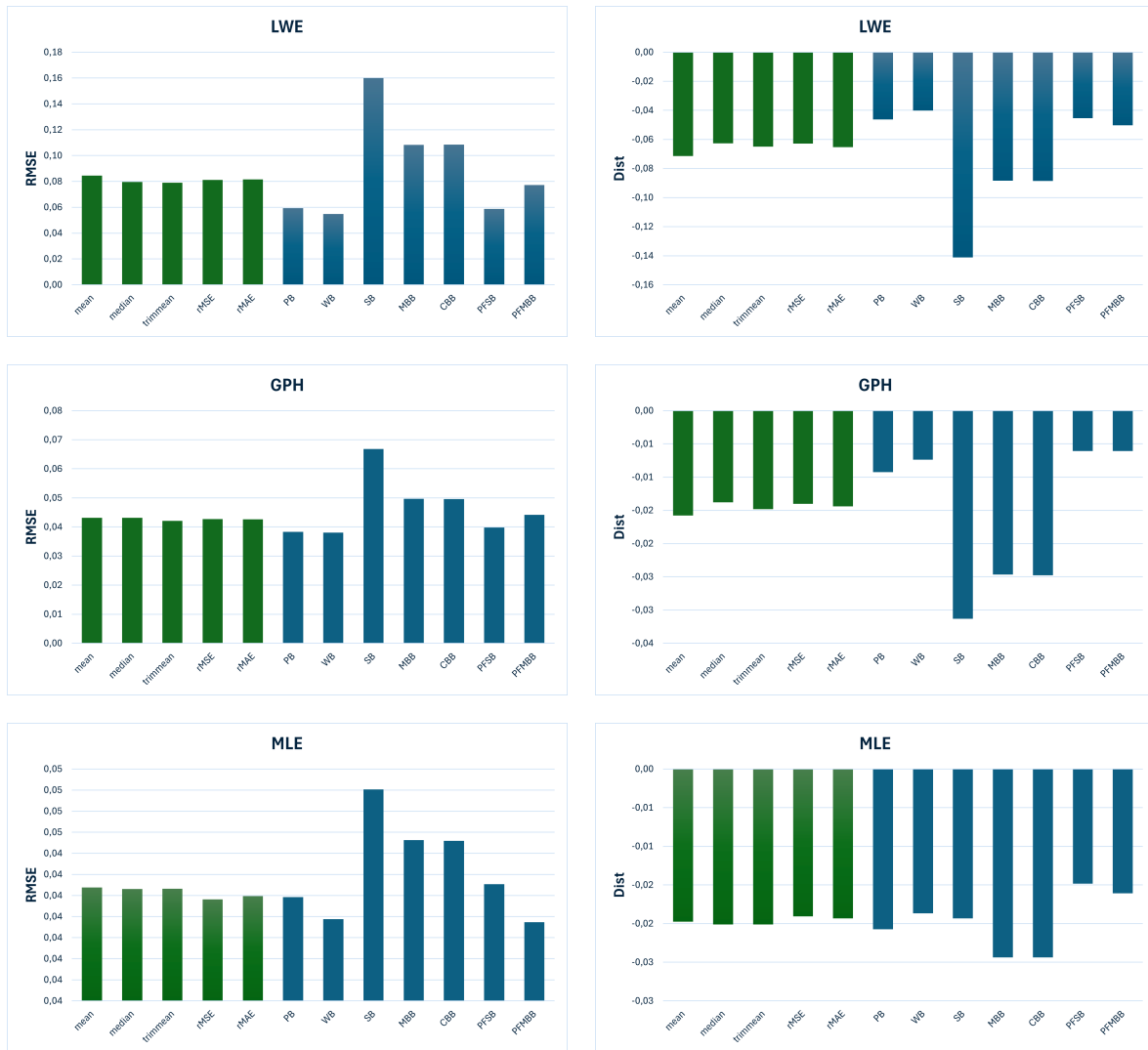


Figura 4.6: RMSE e distorsione ($Dist$) per un $ARFIMA(0,0.2,0)$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

Tabella 4.4: RMSE e distorsione ($Dist$) per un $ARFIMA(0, 0.4, 0)$ con $T = 250$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

T=250						
	LWE		GPH		MLE	
<i>Metodo</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>
Mean	0,1157	-0,0922	0,0602	-0,0200	0,0649	-0,0439
Median	0,1079	-0,0834	0,0591	-0,0184	0,0665	-0,0468
Trimmean	0,1126	-0,0913	0,0601	-0,0215	0,0658	-0,0453
rMSE	0,1125	-0,0835	0,0604	-0,0177	0,0644	-0,0420
rMAE	0,1133	-0,0867	0,0601	-0,0187	0,0646	-0,0431
PB	0,0888	-0,0740	0,0541	-0,0224	0,0703	-0,0548
WB	0,0807	-0,0634	0,0533	-0,0195	0,0677	-0,0507
SB	0,1751	-0,1065	0,0760	-0,0047	0,0562	-0,0230
MBB	0,1605	-0,1306	0,0759	-0,0371	0,0717	-0,0479
CBB	0,1605	-0,1305	0,0758	-0,0371	0,0715	-0,0478
PFSB	0,0873	-0,0703	0,0573	-0,0112	0,0650	-0,0416
PFMBB	0,1047	-0,0698	0,0625	-0,0112	0,0642	-0,0418
T=500						
	LWE		GPH		MLE	
<i>Metodo</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>	<i>RMSE</i>	<i>Dist</i>
Mean	0,0830	-0,0626	0,0418	-0,0086	0,0395	-0,0181
Median	0,0746	-0,0547	0,0410	-0,0076	0,0410	-0,0214
Trimmean	0,0801	-0,0622	0,0415	-0,0096	0,0402	-0,0195
rMSE	0,0801	-0,0556	0,0422	-0,0077	0,0401	-0,0193
rMAE	0,0807	-0,0581	0,0418	-0,0081	0,0399	-0,0192
PB	0,0586	-0,0462	0,0370	-0,0096	0,0435	-0,0286
WB	0,0530	-0,0386	0,0364	-0,0069	0,0425	-0,0266
SB	0,1390	-0,0731	0,0552	0,0044	0,0378	0,0018
MBB	0,1210	-0,0951	0,0545	-0,0204	0,0428	-0,0162
CBB	0,1215	-0,0954	0,0545	-0,0206	0,0430	-0,0164
PFSB	0,0579	-0,0451	0,0391	-0,0049	0,0421	-0,0206
PFMBB	0,0744	-0,0451	0,0441	-0,0049	0,0404	-0,0204



Figura 4.7: RMSE e distorsione ($Dist$) per un $ARFIMA(0,0.4,0)$ e $T = 250$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

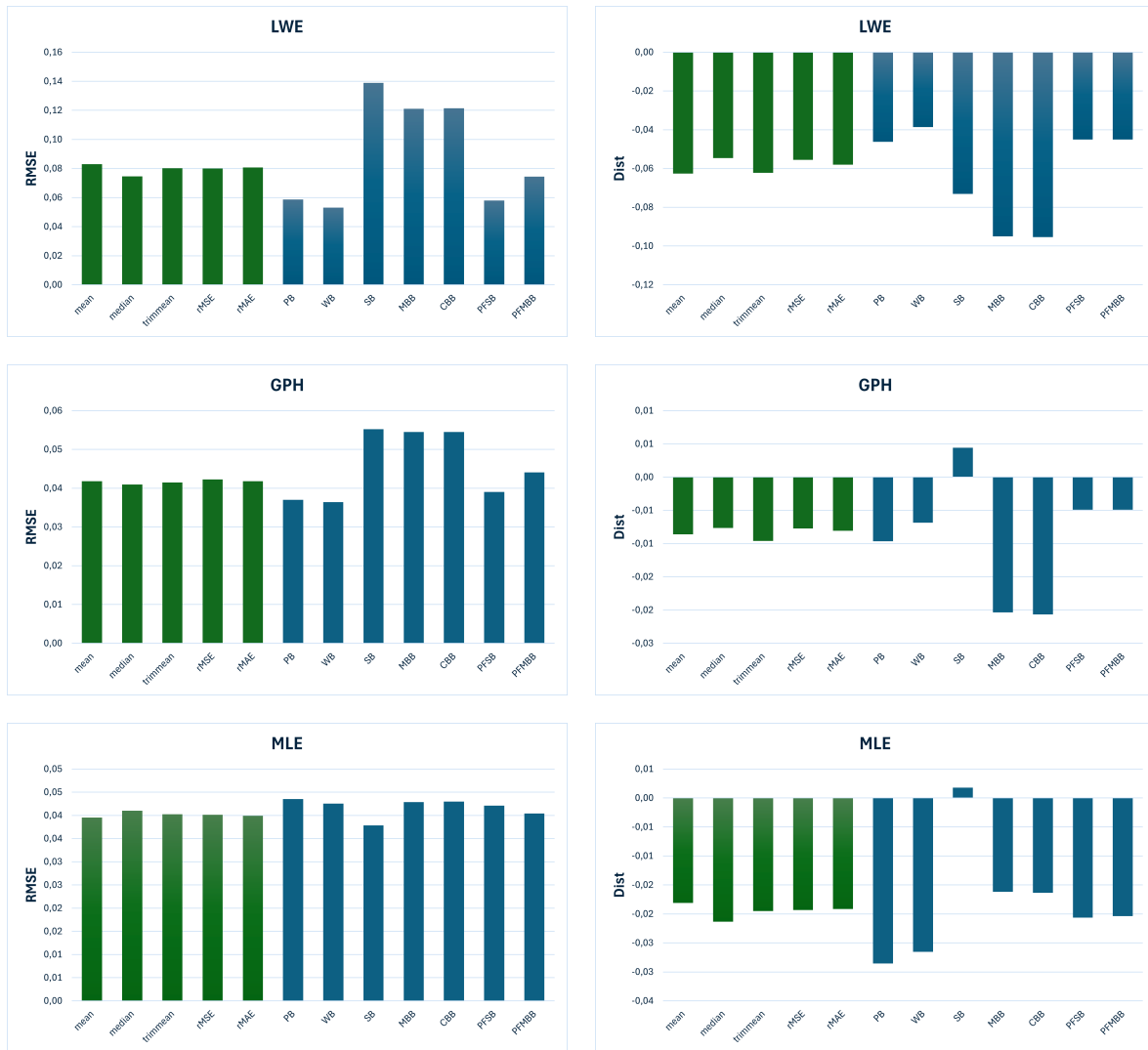


Figura 4.8: RMSE e distorsione ($Dist$) per un $ARFIMA(0,0.4,0)$ e $T = 500$. Esperimenti Monte Carlo relativi ai metodi di combinazione di stime bootstrap.

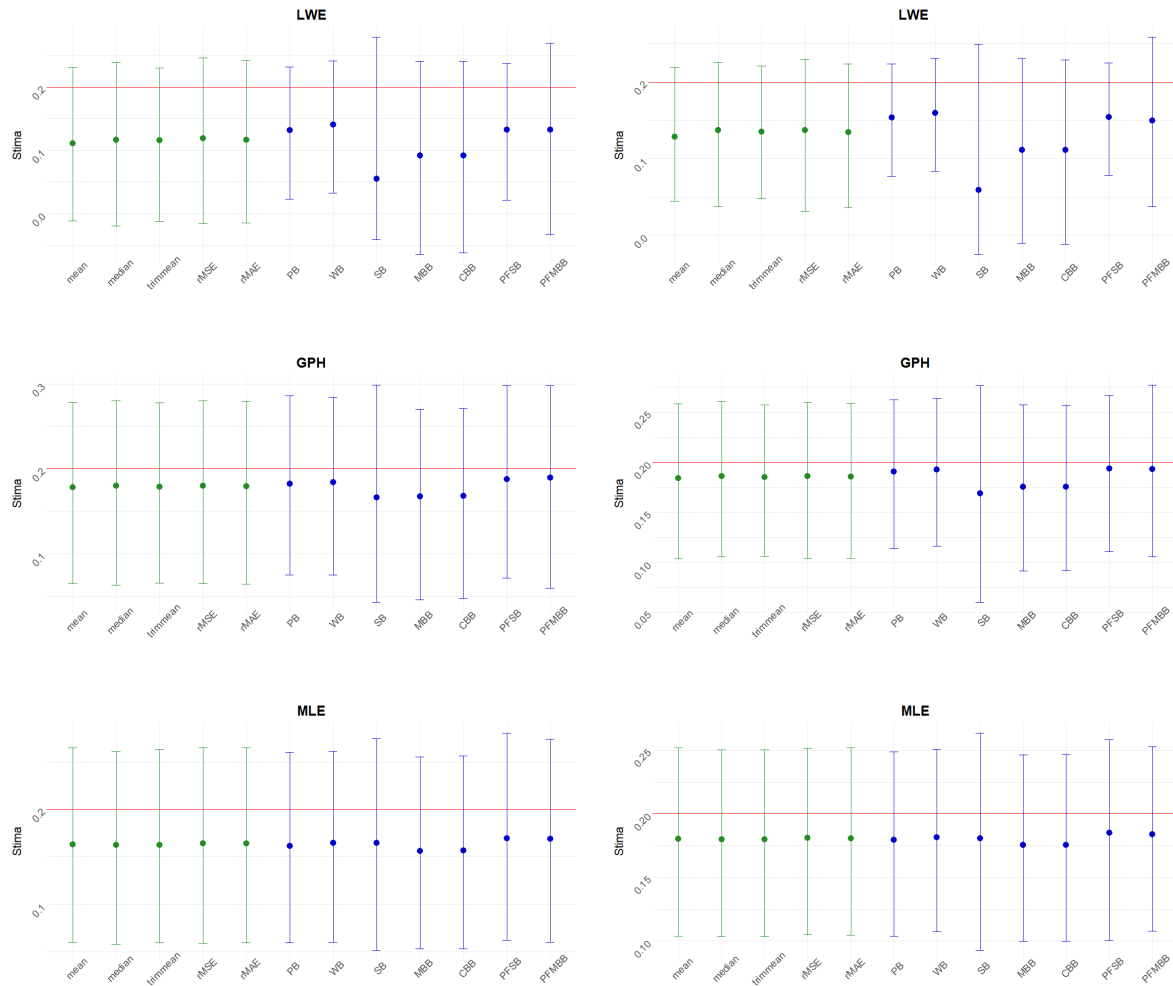


Figura 4.9: Medie delle stime bootstrap del parametro di memoria lunga, ottenute su $S = 1000$ replicazioni Monte Carlo, per ogni metodo di combinazione di stime bootstrap e metodo bootstrap originale (punti), e relativi intervalli di confidenza empirici di livello 0.95 (linee verticali). Esperimenti relativi ai metodi di combinazione di stime bootstrap a partire da un processo $ARFIMA(0, 0.2, 0)$, con $T = 250$ (grafici a sinistra) e $T = 500$ (grafici a destra). La linea rossa orizzontale corrisponde al vero valore del parametro di memoria lunga, $d = 0.2$.

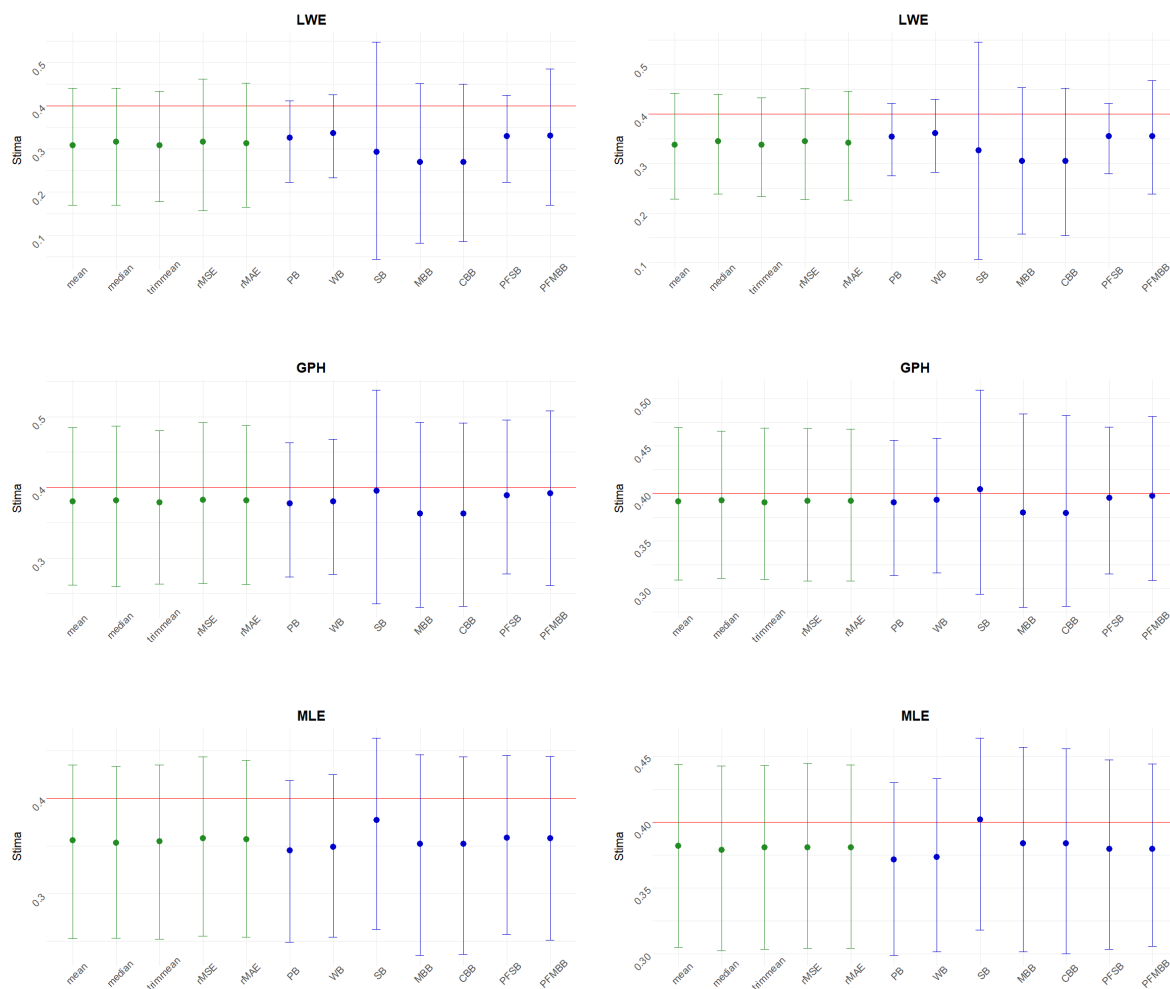


Figura 4.10: Medie delle stime bootstrap del parametro di memoria lunga, ottenute su $S = 1000$ replicazioni Monte Carlo, per ogni metodo di combinazione di stime bootstrap e metodo bootstrap originale (punti), e relativi intervalli di confidenza empirici di livello 0.95 (linee verticali). Esperimenti relativi ai metodi di combinazione di stime bootstrap a partire da un processo $ARFIMA(0, 0.4, 0)$, con $T = 250$ (grafici a sinistra) e $T = 500$ (grafici a destra). La linea rossa orizzontale corrisponde al vero valore del parametro di memoria lunga, $d = 0.4$.

Capitolo 5

Applicazione su dataset reali

In questo Capitolo vengono presentati i risultati relativi all'applicazione dei metodi proposti a due diverse serie storiche reali, con il fine di stimarne il parametro di memoria lunga. Le serie storiche scelte per le applicazioni presentano caratteristiche tipiche delle serie stazionarie a memoria lunga. In ciascuna Sezione sono descritti i dati utilizzati e, successivamente, vengono presentati i risultati relativi alle stime ottenute attraverso le diverse metodologie di bootstrap ibrido e di combinazione di stime bootstrap proposte nel Capitolo 3.

5.1 Livelli minimi del fiume Nilo

I primi dati ad essere analizzati sono quelli relativi alla serie storica annuale dei livelli minimi del fiume Nilo. Si tratta di un dataset classico, spesso utilizzato in statistica per l'analisi di serie temporali, ed è disponibile direttamente sul software R (Team, 2013). Questa serie contiene i livelli minimi annuali del fiume Nilo misurati presso Roda gauge, vicino il Cairo, tra il 622 e il 1284 d.C. (Ṭūsūn, 1925), per un totale di $T = 633$ osservazioni. I dati sono stati analizzati da Sir Francis Galton e Karl Pearson nei primi anni del ventesimo secolo. Ciò che rende interessante questa serie storica, ai fini del presente lavoro, è la presenza di caratteristiche tipiche delle serie storiche a memoria lunga stazionarie. Queste caratteristiche rendono possibile l'applicazione dei metodi bootstrap approfonditi nel Capitolo 2, con la certezza che il loro utilizzo, in questo contesto, sia teoricamente giustificato.

In Figura 5.1 vengono riportati, in alto, il grafico della serie oggetto di studio e, in basso, i grafici delle funzioni di autocorrelazione (ACF) e di autocorrelazione parziale (PACF) della serie stessa. Da questi grafici è evidente, osservando il lento decrescimento della funzione di

autocorrelazione verso lo zero, che la serie presenta un comportamento tipico delle serie a memoria lunga. Il test aumentato di Dickey-Fuller (ADF; Said e Dickey, 1984), applicato alla serie, respinge l'ipotesi nulla di radice unitaria in favore dell'ipotesi alternativa di stazionarietà; inoltre, le stime del parametro di memoria lunga di Whittle locale (LWE), del log-periodogramma (GPH) e di massima verosimiglianza (MLE) calcolate per la serie sono, rispettivamente, pari a circa 0.391, 0.399 e 0.394, valori alti ma inferiori a 0.5.

Per l'applicazione dei metodi bootstrap tradizionali, bootstrap ibridi e delle combinazioni di stime bootstrap, viene sempre utilizzato un numero di replicazioni bootstrap pari a $B = 500$.

La Figura 5.2 offre una rappresentazione grafica dei risultati ottenuti dall'applicazione dei metodi bootstrap ibridi e bootstrap originali sulla serie storica dei livelli minimi annuali del fiume Nilo, relativi al periodo compreso tra il 622 e il 1284. In particolare, per ciascun metodo considerato e ciascuno stimatore (LWE, GPH e MLE), è riportata la media delle B stime del parametro di memoria lunga d , ottenute dalle B replicazioni bootstrap, rappresentata dal corrispondente punto. Gli intervalli di confidenza di livello 0.95 per queste stime, costruiti a partire dalla distribuzione empirica delle stime stesse, sono indicati con linee verticali, i cui estremi rappresentano i quantili 0.025 e 0.975 della distribuzione empirica. Si riportano, inoltre, per ognuno dei tre stimatori, la stima del parametro di memoria lunga sulla serie originale, indicata con una linea rossa continua orizzontale, ed i quantili 0.025 e 0.975 della sua distribuzione asintotica, indicati con le linee rosse tratteggiate orizzontali. Dall'osservazione dei grafici emerge, innanzitutto, che gli intervalli di confidenza empirici delle stime ottenute con i metodi bootstrap ibridi, in generale, sono leggermente più ampi rispetto a quelli relativi ai metodi bootstrap parametrici (PB e WB), ma presentano ampiezze comparabili o minori rispetto agli altri metodi bootstrap tradizionali. Inoltre, la media delle stime ottenute con la maggior parte dei metodi non si discosta significativamente dalla stima calcolata sulla serie originale. Questi risultati suggeriscono che i metodi bootstrap ibridi possano essere preferibili, in quanto le stime che producono sono accompagnate, in media, da una maggiore precisione. Tuttavia, il metodo bootstrap ibrido basato sulla mediana (*median*) sembra essere un'eccezione in questa applicazione, poiché fornisce risultati meno affidabili rispetto agli altri metodi, confermando le criticità già discusse nel Capitolo precedente.

La Figura 5.3 presenta, invece, la sintesi grafica dei risultati ottenuti dall'applicazione dei metodi di combinazione di stime bootstrap e dei metodi bootstrap originali sulla medesima serie storica. Per ogni metodo e per ogni stimatore, la stima del parametro di memoria lunga derivata dalle B replicazioni bootstrap è rappresentata da un punto. A differenza dei metodi

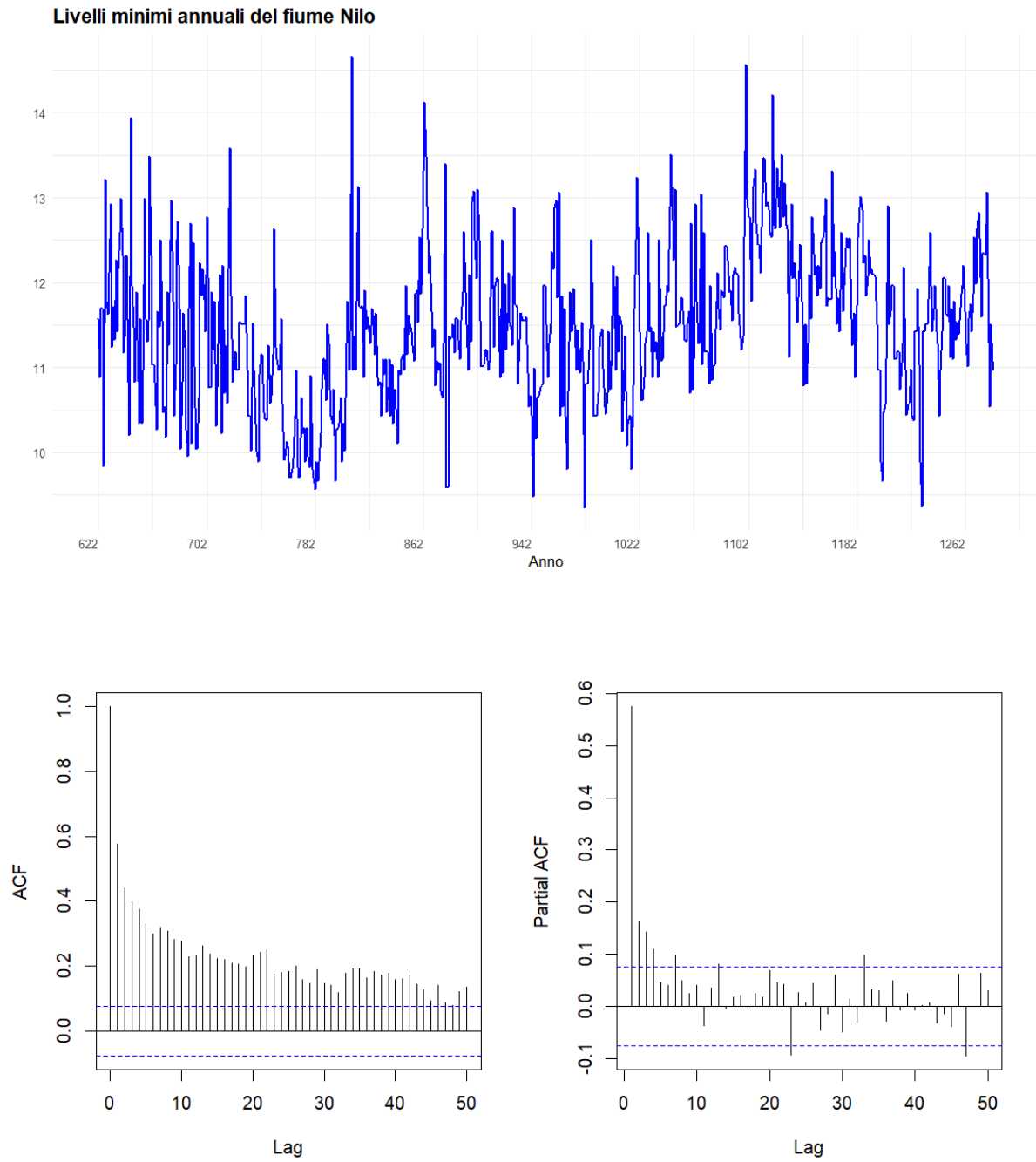


Figura 5.1: Livelli minimi annuali del fiume Nilo, 622-1284. Grafico della serie e di ACF e PACF.

bootstrap ibridi, per questi metodi non è possibile ottenere un intervallo di confidenza empirico a causa della diversa modalità di costruzione. Anche in questo caso, vengono riportati la stima del parametro di memoria lunga sulla serie originale (linea rossa continua orizzontale) ed i quantili 0.025 e 0.975 della sua distribuzione asintotica (linee rosse tratteggiate orizzontali). Dai grafici emerge chiaramente che, mentre le stime ottenute tramite la combinazione di stime bootstrap risultano essere, tra loro, piuttosto omogenee, le stime derivanti dai metodi bootstrap originali mostrano una variabilità significativamente maggiore, specialmente per lo stimatore MLE. Questa variabilità porta a risultati anche molto differenti tra loro, a seconda che si consideri una stima ottenuta con un metodo parametrico, semiparametrico o non parametrico. Tale diversità nei risultati rappresenta una criticità, poiché, in assenza di conoscenza del PGD, non è possibile determinare con certezza a quale metodo bootstrap originale affidarsi. Al contrario, le combinazioni di stime bootstrap si dimostrano più affidabili, riducendo il rischio di errori derivanti dalla scelta di un singolo metodo e migliorando la robustezza delle stime.

5.2 Rendimenti assoluti dei prezzi azionari di Microsoft

La seconda serie analizzata in questo studio è di natura finanziaria e riguarda i rendimenti giornalieri in valore assoluto dei prezzi azionari di Microsoft Corporation. I dati coprono il periodo compreso tra il 2 gennaio 2020 e il 29 dicembre 2023, per un totale di $T = 1006$ osservazioni. Il dataset relativo ai prezzi azionari analizzati è gratuitamente disponibile sul sito di Yahoo Finance (2024). I rendimenti giornalieri sono calcolati come $x_t = \ln p_t - \ln p_{t-1}$, $t = 2, \dots, T$, dove p_t è la quotazione di chiusura giornaliera dell'azienda Microsoft Corporation. La serie di interesse è costituita dai rendimenti assoluti $r_t = |x_t|$ poiché questa trasformazione possiede alcune proprietà statistiche desiderabili. Infatti, come sottolineato da Cotter (2011), i rendimenti assoluti sono robusti rispetto alla presenza di valori estremi. Inoltre, essi consentono di ottenere, asintoticamente, stime accurate della volatilità latente non osservabile. Infine, la volatilità dei rendimenti assoluti offre proprietà campionarie finite particolarmente utili, nella pratica, per la gestione del rischio finanziario.

In Figura 5.4 sono riportati, in alto, il grafico della serie dei rendimenti assoluti oggetto di studio assieme ai grafici delle funzioni di autocorrelazione (ACF) e di autocorrelazione parziale (PACF) della serie stessa. Sebbene il comportamento di questa serie differisca notevolmente da quello della serie analizzata nella Sezione precedente, è possibile osservare nuovamente le caratteristiche tipiche dei processi a memoria lunga, come il lento decrescimento della funzione di autocorrelazione verso lo zero. Anche in questo caso, il test ADF rifiuta l'ipotesi nulla di

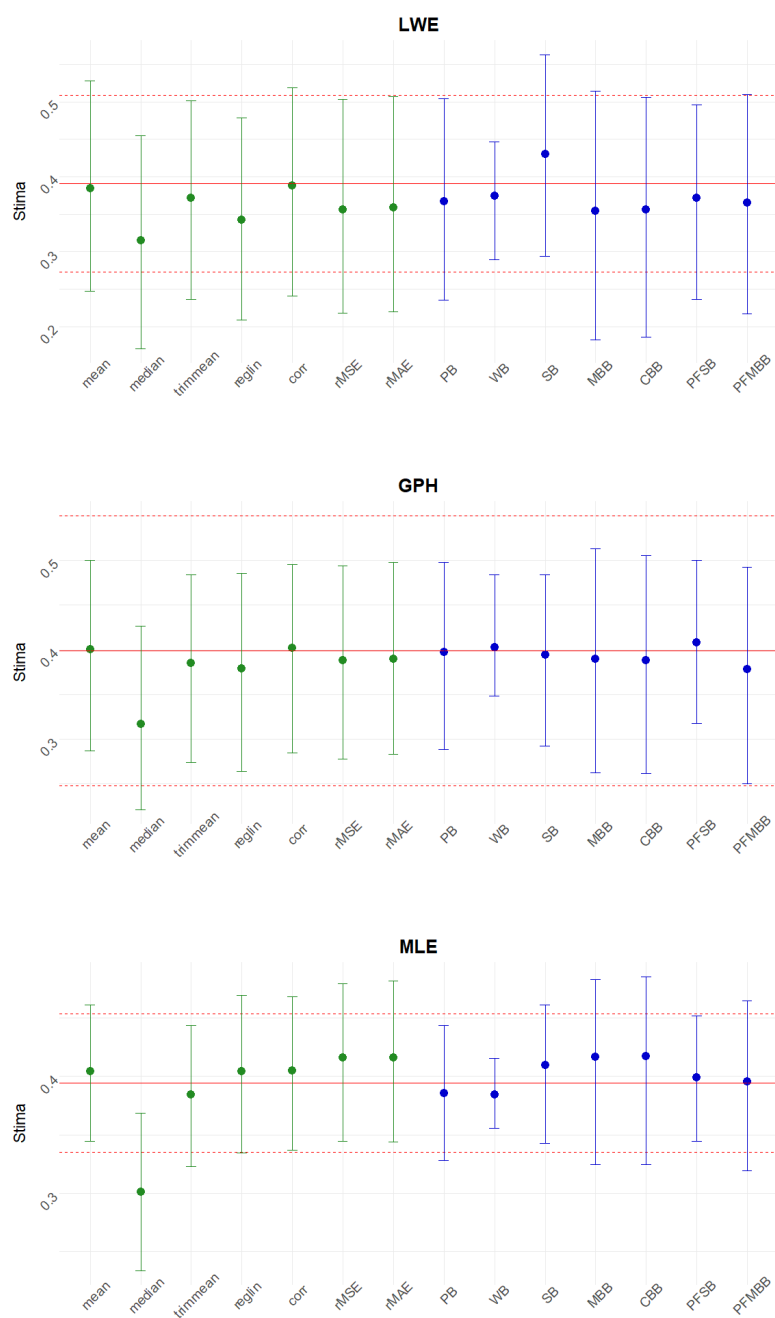


Figura 5.2: Medie di $B = 500$ stime bootstrap del parametro di memoria lunga della serie storica dei livelli minimi annuali del fiume Nilo, per ogni metodo bootstrap ibrido e bootstrap originale (punti), e relativi intervalli di confidenza empirici di livello 0.95 (linee verticali). Confronto con la stima del parametro sulla serie storica originale (linea rossa continua) e relativo intervallo di confidenza asintotico di livello 0.95 (linea rossa tratteggiata).

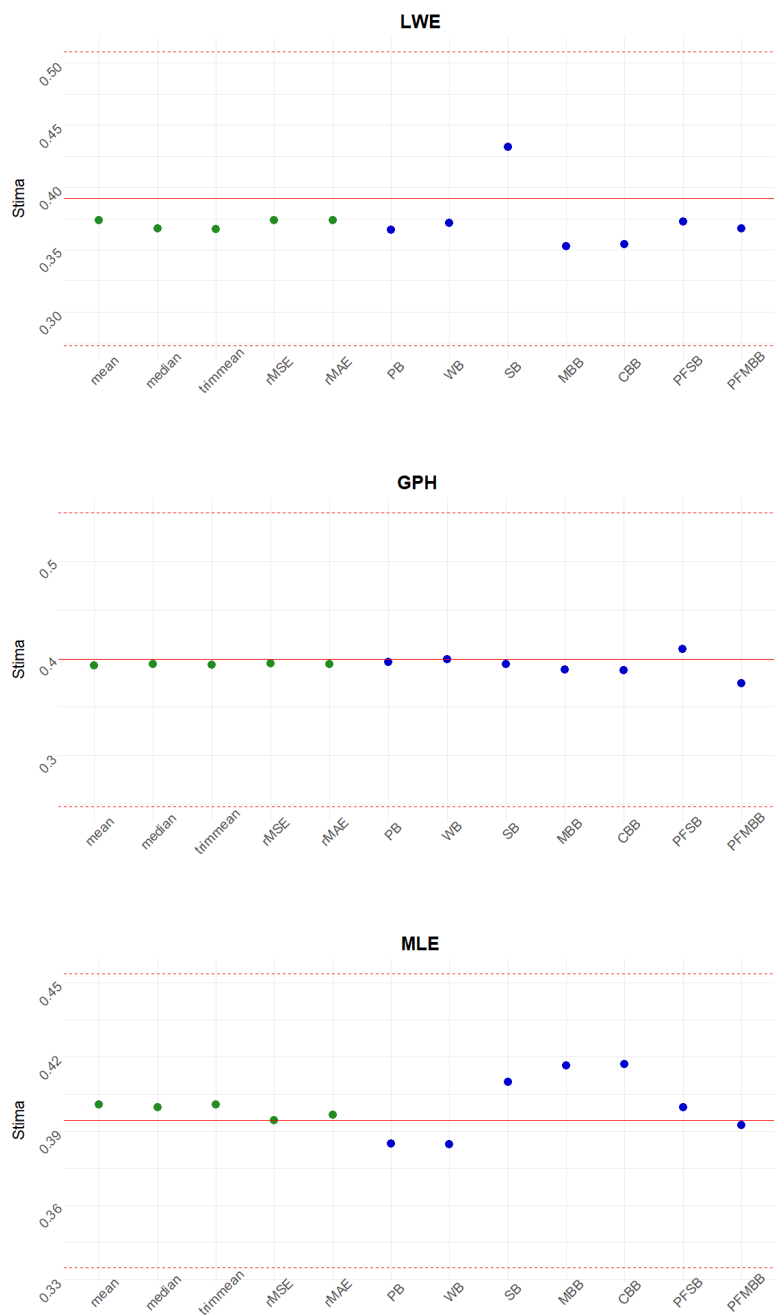


Figura 5.3: Stime del parametro di memoria lunga della serie storica dei livelli minimi annuali del fiume Nilo, ricavate da $B = 500$ replicazioni bootstrap, per ogni metodo di combinazione di stime bootstrap e metodo bootstrap originale (punti). Confronto con la stima sulla serie storica originale (linea rossa continua) e relativo intervallo di confidenza asintotico di livello 0.95 (linea rossa tratteggiata).

radice unitaria. Le stime del parametro di memoria lunga ottenute con gli stimatori di Whittle locale (LWE), del log-periodogramma (GPH) e di massima verosimiglianza (MLE) risultano, rispettivamente, pari a circa 0.399, 0.327 e 0.255, evidenziando una variabilità nei valori stimati per la persistenza della memoria nel tempo, ma rimanendo comunque all'interno del *range* che caratterizza la stazionarietà della serie.

Come nelle Simulazioni del Capitolo 4 e nella precedente applicazione, viene utilizzato un numero di replicazioni bootstrap pari a $B = 500$.

In Figura 5.5 è presentata una sintesi grafica dei risultati ottenuti dall'applicazione dei metodi bootstrap ibridi e bootstrap originali sulla serie storica dei rendimenti assoluti giornalieri dei prezzi azionari di Microsoft Corporation, analogamente a quanto visto in precedenza nella Sezione 5.1. Per ciascun metodo considerato e ciascuno stimatore (LWE, GPH e MLE), viene mostrata la media delle B stime del parametro di memoria lunga d ottenute dalle B replicazioni, rappresentata da un punto, insieme al corrispondente intervallo di confidenza di livello 0.95, costruito a partire dalla distribuzione empirica delle stime, indicato da una linea verticale. Le linee rosse rappresentano, quindi, la stima del parametro di memoria lunga sulla serie originale (linea continua) ed i quantili 0.025 e 0.975 della sua distribuzione asintotica (linee tratteggiate). In questi grafici risalta l'ampiezza degli intervalli di confidenza empirici associati alle stime ottenute con i metodi bootstrap non parametrici, quali MBB, CBB e PFMBB, suggerendo una bassa precisione di tali stime. Tuttavia, in un contesto in cui non è noto il PGD, affidarsi esclusivamente a metodi bootstrap parametrici può essere rischioso, anche se sembrano fornire stime più affidabili. È inoltre interessante osservare come le stime del parametro di memoria lunga ottenute tramite metodi bootstrap e bootstrap ibridi si discostino, generalmente, dalla stima ottenuta dalla serie originale, che sembra sovrastimare la persistenza della memoria in tutti e tre gli stimatori considerati. Per quanto riguarda i metodi bootstrap ibridi considerati, le osservazioni relative ai metodi che sfruttano schemi di combinazione basati su mediana e media rimangono invariate: la combinazione tramite mediana (*median*) mostra risultati che si discostano in modo poco desiderabile da quelli degli altri metodi bootstrap ibridi, mentre la combinazione tramite media (*mean*), nella sua semplicità, fornisce risultati molto simili a quelli ottenuti da metodi bootstrap ibridi più complessi e la sua stima sembra essere anche più precisa.

In Figura 5.6 è illustrata la sintesi grafica dei risultati ottenuti dall'applicazione dei metodi di combinazione di stime bootstrap e dei metodi bootstrap originali sulla stessa serie storica. Analogamente a quanto visto in precedenza, per ciascun metodo e per ciascuno stimatore

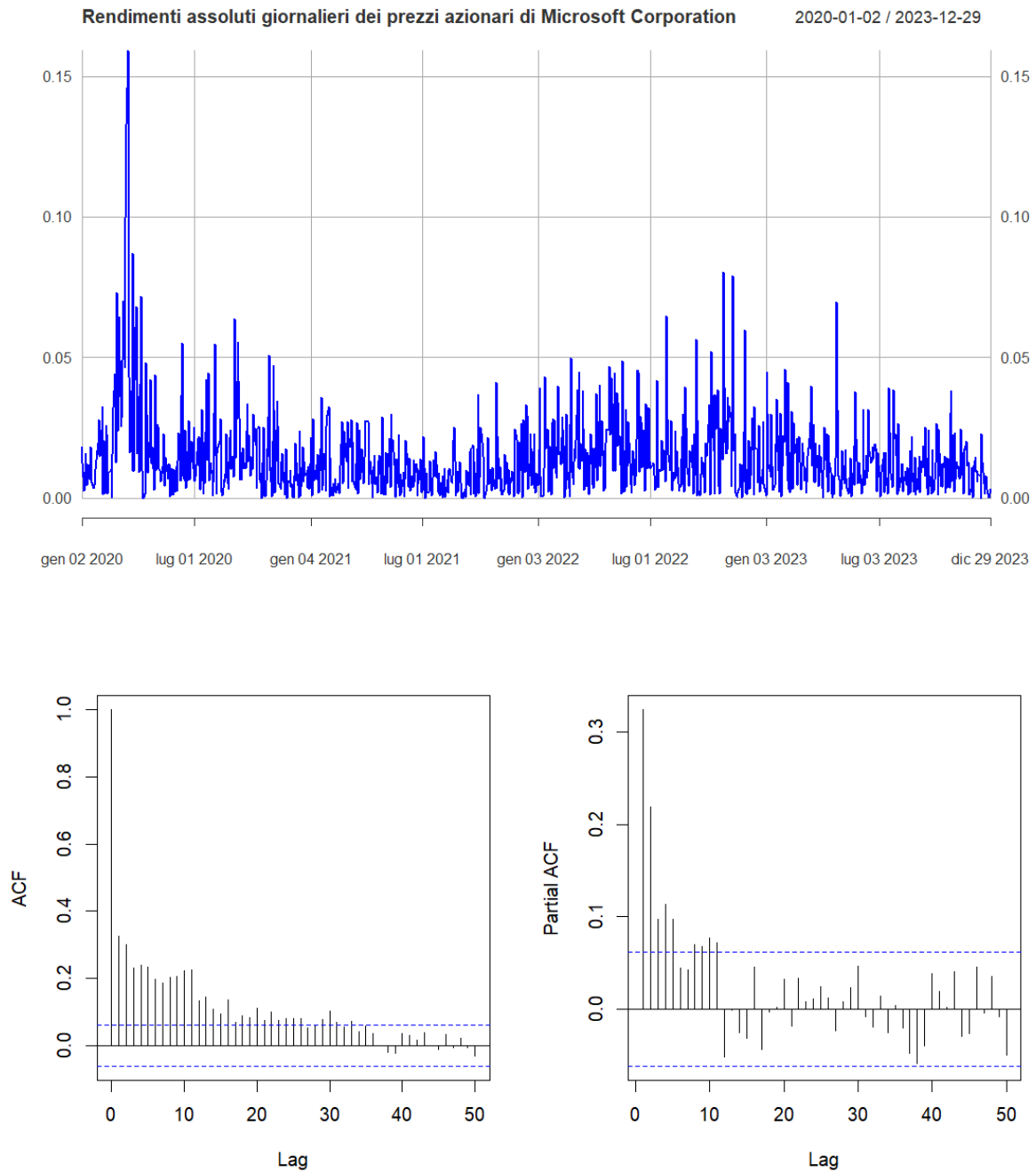


Figura 5.4: Rendimenti assoluti giornalieri dei prezzi azionari di Microsoft Corporation dal 2 gennaio 2020 al 29 dicembre 2023. Grafico della serie e di ACF e PACF.

(LWE, GPH e MLE), la stima del parametro di memoria lunga ottenuta dalle B repliche bootstrap è rappresentata un punto. Tuttavia, a causa della diversa costruzione di questi metodi rispetto ai bootstrap ibridi, non è possibile ottenere un intervallo di confidenza empirico per tali stime. Le linee orizzontali rosse indicano la stima del parametro di memoria lunga calcolata dalla serie originale (linea continua) e i quantili 0.025 e 0.975 della sua distribuzione asintotica (linee tratteggiate). Le stime del parametro di memoria lunga derivate dai diversi metodi bootstrap e dalle combinazioni di stime bootstrap, in particolare quelle relative allo LWE, sembrano discostarsi molto dalla stima ottenuta sulla serie originale. Per quanto riguarda le altre stime, invece, emerge nuovamente che i metodi bootstrap originali possono generare risultati molto diversi tra loro, evidenziando l'incertezza associata a queste stime quando considerate singolarmente.

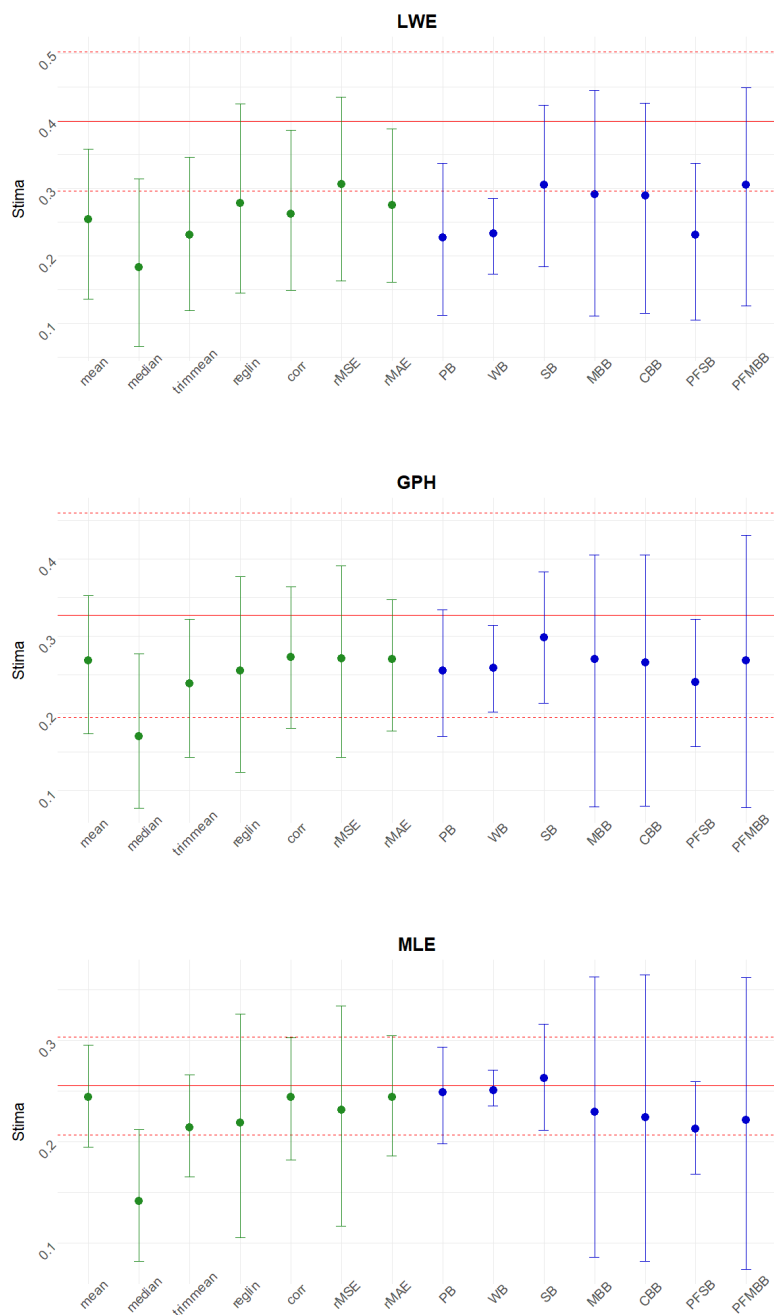


Figura 5.5: Medie di $B = 500$ stime bootstrap del parametro di memoria lunga della serie storica dei rendimenti assoluti giornalieri dei prezzi azionari di Microsoft Corporation, per ogni metodo bootstrap ibrido e bootstrap originale (punti), e relativi intervalli di confidenza empirici di livello 0.95 (linee verticali). Confronto con la stima del parametro sulla serie storica originale (linea rossa continua) e relativo intervallo di confidenza asintotico di livello 0.95 (linea rossa tratteggiata).

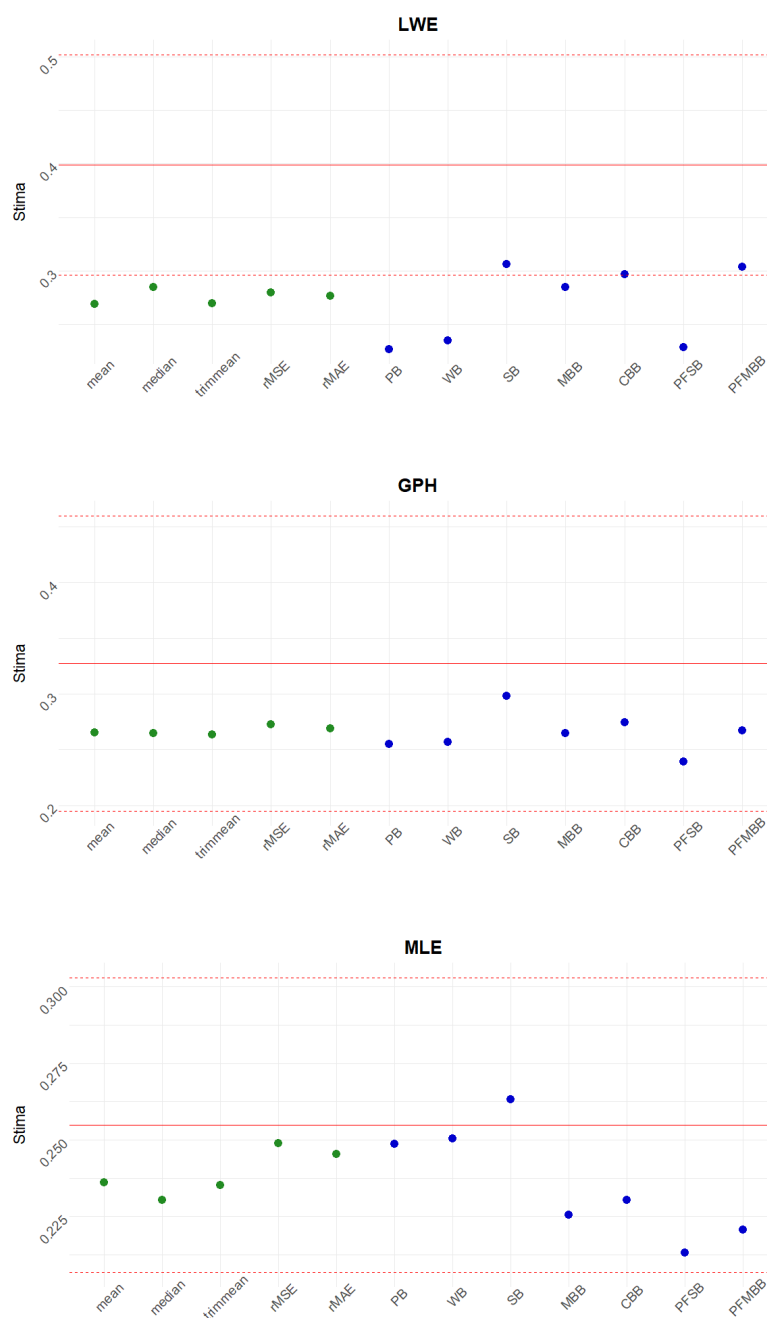


Figura 5.6: Stime del parametro di memoria lunga della serie storica dei rendimenti assoluti giornalieri dei prezzi azionari di Microsoft Corporation, ricavate da $B = 500$ replicazioni bootstrap, per ogni metodo di combinazione di stime bootstrap e metodo bootstrap originale (punti). Confronto con la stima sulla serie storica originale (linea rossa continua) e relativo intervallo di confidenza asintotico di livello 0.95 (linea rossa tratteggiata).

Conclusioni

In questa tesi, partendo dall'intuizione, basata sulle conoscenze in ambito previsivo, che la combinazione di diverse fonti di informazione può mitigare le limitazioni dei singoli metodi, sono stati proposti ed analizzati due nuovi approcci per integrare i risultati ottenuti da diversi metodi bootstrap. Questo lavoro si distingue per l'originalità nella proposta di nuove strategie di combinazione, mirate a migliorare la stima del parametro di memoria lunga in serie storiche stazionarie a memoria lunga. L'obiettivo principale di questo lavoro era quello di esplorare e dimostrare l'efficacia di questi nuovi approcci, offrendo così un contributo innovativo al campo.

L'analisi empirica dei metodi proposti, condotta attraverso esperimenti Monte Carlo e applicazioni a serie storiche reali, ha evidenziato che le combinazioni di metodi bootstrap sembrano effettivamente portare a stimatori del parametro di memoria lunga che presentano una minore variabilità e una maggiore accuratezza. In particolare, i risultati ottenuti indicano che i metodi di combinazione basati su schemi di combinazione lineari, nei quali vengono stimati pesi diversi per ogni metodo bootstrap individuale, offrono prestazioni superiori rispetto alle combinazioni semplici e ai singoli metodi bootstrap, fornendo stime che sono più vicine ai valori attesi e con intervalli di confidenza più affidabili. Tuttavia, proprio come in ambito previsivo, gli schemi di combinazione basati sulla media si sono dimostrati molto competitivi.

Nonostante i risultati siano promettenti, la ricerca presenta alcune limitazioni. In primo luogo, l'analisi è stata condotta su un numero limitato di PGD e valori del parametro di memoria lunga, concentrandosi su processi a memoria lunga stazionaria. Un possibile sviluppo di questo lavoro riguarda l'estensione di questi metodi a serie storiche non stazionarie, dove la complessità aggiuntiva della non stazionarietà potrebbe influenzare significativamente le stime del parametro di memoria lunga. Un'altra limitazione riguarda i metodi bootstrap considerati, i quali sono sviluppati nel dominio del tempo. L'integrazione di metodi bootstrap operanti nel dominio delle frequenze all'interno delle strategie di combinazione rappresenta una promettente estensione metodologica, che potrebbe portare all'integrazione di informazioni utili sulle serie

temporali e migliorare la qualità delle stime. Inoltre, l'esplorazione di ulteriori schemi di combinazione, inclusi quelli più sofisticati e basati su combinazioni non lineari, potrebbe offrire nuovi strumenti per ottimizzare la scelta dei pesi e migliorare ulteriormente la *performance* degli stimatori.

Lo studio delle prestazioni dei metodi bootstrap ibridi e di combinazione di stime bootstrap proposti si è svolto, in questo lavoro, unicamente dal punto di vista empirico. Sarebbe di grande interesse un'analisi teorica approfondita delle proprietà degli stimatori ottenuti attraverso la combinazione di metodi bootstrap, con l'obiettivo di formalizzare i vantaggi osservati empiricamente e fornire una base teorica solida per l'applicazione di queste tecniche, consolidando il loro valore nel contesto delle serie temporali a memoria lunga.

In conclusione, il lavoro presentato in questa tesi ha mostrato che la proposta di combinare diversi metodi bootstrap offre un'alternativa innovativa e promettente agli approcci tradizionali per la stima del parametro di memoria lunga. Questi metodi originali hanno permesso di ottenere stimatori più robusti e affidabili, offrendo una soluzione efficace per affrontare le sfide associate alle serie storiche. I risultati ottenuti suggeriscono che gli approcci combinati meritano ulteriore esplorazione e sviluppo, sia dal punto di vista teorico che applicativo, per affinare le tecniche di stima anche in contesti complessi e variabili, tipici nell'ambito delle serie storiche.

Bibliografia

- Adhikari, R., & Agrawal, R. K. (2012). A Novel Weighted Ensemble Technique for Time Series Forecasting. In P.-N. Tan, S. Chawla, C. K. Ho & J. Bailey (Cur.), *Advances in Knowledge Discovery and Data Mining* (pp. 38–49). Springer Berlin Heidelberg.
- Agnew, C. E. (1985). Bayesian consensus forecasts of macroeconomic variables. *Journal of Forecasting*, 4(4), 363–376.
- Aiolfi, M., & Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, 135(1), 31–53.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Aksu, C., & Gunter, S. I. (1992). An empirical analysis of the accuracy of SA, OLS, ERLS and NRLS combination forecasts. *International Journal of Forecasting*, 8(1), 27–43.
- Andrawis, R. R., Atiya, A. F., & El-Shishiny, H. (2011). Combination of long term and short term forecasts, with application to tourism demand forecasting. *International Journal of Forecasting*, 27(3), 870–886.
- Andrews, D. W., Lieberman, O., & Marmar, V. (2006). Higher-order improvements of the parametric bootstrap for long-memory Gaussian processes. *Journal of Econometrics*, 133(2), 673–702.
- Arteche, J., & Orbe, J. (2005). Bootstrapping the log-periodogram regression. *Economics Letters*, 86(1), 79–85.
- Arteche, J. (2015). Signal extraction in long memory stochastic volatility. *Econometric Theory*, 31(6), 1382–1402.
- Arteche, J. (2024). Bootstrapping long memory time series: Application in low frequency estimators. *Econometrics and Statistics*, 29, 1–15.

- Babikir, A., & Mwambi, H. (2016). Evaluating the combined forecasts of the dynamic factor model and the artificial neural network model using linear and nonlinear combining methods. *Empir Econ*, 51, 1541–1556.
- Bates, J., & Granger, C. (1969). The Combination of Forecasts. *Journal of the Operational Research Society*, 20(4), 451–468.
- Baumeister, C., & Kilian, L. (2015). Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach. *Journal of Business and Economic Statistics*, 33(3), 338–351.
- Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(4), 659–672.
- Bickel, P. J., & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The annals of statistics*, 9(6), 1196–1217.
- Bühlmann, P. (1997). Sieve Bootstrap for Time Series. *Bernoulli*, 3(2), 123–148.
- Bunn, D. W. (1975). A Bayesian Approach to the Linear Combination of Forecasts. *Journal of the Operational Research Society*, 26(2), 325–329.
- Burg, J. P. (1975). *Maximum entropy spectral analysis*. Stanford University.
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2^a ed.). Springer New York, NY.
- Butka, A., & Puka, L. (2014). A block bootstrap procedure for long memory processes. *International Journal of Mathematics Trends and Technology-IJMTT*, 14.
- Cang, S., & Yu, H. (2014). A combination selection algorithm on forecasting. *European Journal of Operational Research*, 234(1), 127–139.
- Carlstein, E. (1986). The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence. *The Annals of Statistics*, 14(3), 1171–1179.
- Cavaliere, G., & Georgiev, I. (2020). Inference Under Random Limit Bootstrap Measures. *Econometrica*, 88(6), 2547–2574.
- Cavaliere, G., Nielsen, M. Ø., & Taylor, A. R. (2015). Bootstrap score tests for fractional integration in heteroskedastic ARFIMA models, with an application to price dynamics in commodity spot and futures markets [Econometric Analysis of Financial Derivatives]. *Journal of Econometrics*, 187(2), 557–579.

- Cavaliere, G., Nielsen, M. Ø., & Taylor, A. R. (2017). Quasi-maximum likelihood estimation and bootstrap inference in fractional time series models with heteroskedasticity of unknown form. *Journal of Econometrics*, 198(1), 165–188.
- Chan, Y., Stock, J., & Watson, M. (1999). A dynamic factor model framework for forecast combination. *Span Econ Rev*, 1(2), 91–121.
- Claeskens, G., Magnus, J., Vasnev, A., & Wang, W. (2016). The forecast combination puzzle: A simple theoretical explanation. *International Journal of Forecasting*, 32(3), 754–762.
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559–583.
- Clements, M., & Hendry, D. (1998). *Forecasting Economic Time Series*. Cambridge University Press.
- Conflitti, C., De Mol, C., & Giannone, D. (2015). Optimal combination of survey forecasts. *International Journal of Forecasting*, 31(4), 1096–1103.
- Cotter, J. (2011). Absolute return volatility. *arXiv preprint arXiv:1103.5976*.
- de Peretti, C., & Siani, C. (2010). Graphical methods for investigating the finite-sample properties of confidence regions. *Computational Statistics & Data Analysis*, 54(2), 262–271.
- Deutsch, M., Granger, C. W., & Teräsvirta, T. (1994). The combination of forecasts using changing weights. *International Journal of Forecasting*, 10(1), 47–57.
- Diebold, F. X., & Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting*, 6(1), 21–40.
- Diebold, F. X., & Shin, M. (2019). Machine learning for regularized survey forecast combination: Partially-egalitarian LASSO and its derivatives. *International Journal of Forecasting*, 35(4), 1679–1691.
- Donaldson, R. G., & Kamstra, M. (1996). Forecast combining with neural networks. *Journal of Forecasting*, 15(1), 49–61.
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1–26.
- Elliott, G. (2011). Averaging and the optimal combination of forecasts. *University of California, San Diego*.
- Elliott, G., & Timmermann, A. (2004). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics*, 122(1), 47–79.

- Fischer, I., & Harvey, N. (1999). Combining forecasts: What information do judges need to outperform the simple average? *International Journal of Forecasting*, 15(3), 227–246.
- Fox, R., & Taqqu, M. (2003, giugno). Large-sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series.
- Franke, J., & Hardle, W. (1992). On Bootstrapping Kernel Spectral Estimates. *The Annals of Statistics*, 20(1), 121–145.
- Freitas, P. S., & Rodrigues, A. J. (2006). Model combination in neural-based forecasting. *European Journal of Operational Research*, 173(3), 801–814.
- Gastinger, J., Nicolas, S., Stepić, D., Schmidt, M., & Schülke, A. (2021). A study on Ensemble Learning for Time Series Forecasting and the need for Meta-Learning, 1–8.
- Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, 29(1), 108–121.
- Geweke, J., & Porter-Hudak, S. (1983). The Estimation And Application Of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4(4), 221–238.
- Granger, C., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2), 197–204.
- Harrald, P., & Kamstra, M. (1997). Evolving artificial neural networks to combine financial forecasts. *IEEE Transactions on Evolutionary Computation*, 1(1), 40–52.
- Haslett, J., & Raftery, A. E. (1989). Space-Time Modelling with Long-Memory Dependence: Assessing Ireland’s Wind Power Resource. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 38(1), 1–21.
- Hendry, D., & Clements, M. (2004). Pooling of forecasts. *Econometrics Journal*, 7(1), 1–31.
- Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, 21(1), 15–24.
- Hora, S. C. (2004). Probability judgments for continuous quantities: Linear combinations and calibration [Cited by: 86]. *Management Science*, 50(5), 597–604.
- Hurvich, C. M., Deo, R., & Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak’s estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19(1), 19–46.
- Kapetanios, G., & Papailias, F. (2011). *Block bootstrap and long memory* (rapp. tecn.). Working paper.

- Kapetanios, G., Papailias, F., & Taylor, A. M. R. (2019). A Generalised Fractional Differencing Bootstrap for Long Memory Processes. *Journal of Time Series Analysis*, 40(4), 467–492.
- Kim, Y., & Nordman, D. (2011). Properties of a block bootstrap under long-range dependence. *Sankhya A*, 73, 79–109.
- Kolassa, S. (2011). Combining exponential smoothing forecasts using Akaike weights. *International Journal of Forecasting*, 27(2), 238–251.
- Kourentzes, N., Barrow, D., & Petropoulos, F. (2019). Another look at forecast selection and combination: Evidence from forecast pooling. *International Journal of Production Economics*, 209, 226–235.
- Krasnopolsky, V. M., & Lin, Y. (2012). A Neural Network Nonlinear Multimodel Ensemble to Improve Precipitation Forecasts over Continental US. *Advances in Meteorology*, 2012(1).
- Kreiss, J.-P., & Paparoditis, E. (2011). Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4), 357–378.
- Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*, 17(3), 1217–1241.
- Lahiri, K., Peng, H., & Y., Z. (2017). Online learning and forecast combination in unbalanced panels. *Econometric Reviews*, 36(1-3), 257–288.
- Lahiri, S. N. (1993). On the moving block bootstrap under long range dependence. *Statistics & Probability Letters*, 18(5), 405–413.
- Lahiri, S. N. (2013). *Resampling methods for dependent data*. Springer Science & Business Media.
- Larrick, R. P., & Soll, J. B. (2006). Intuitions About Combining Opinions: Misappreciation of the Averaging Principle. *Management Science*, 52(1), 111–127.
- Lichtendahl, K. C., & Winkler, R. L. (2020). Why do some combinations perform better than others? [M4 Competition]. *International Journal of Forecasting*, 36(1), 142–149.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1), 54–74.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions [Special Issue: M5 competition]. *International Journal of Forecasting*, 38(4), 1346–1364.

- Makridakis, S., & Winkler, R. L. (1983). Averages of Forecasts: Some Empirical Results. *Management Science*, 29(9), 987–996.
- McNees, S. K. (1992). The uses and abuses of ‘consensus’ forecasts. *Journal of Forecasting*, 11(8), 703–710.
- Ming Shi, S., Da Xu, L., & Liu, B. (1999). Improving the accuracy of nonlinear combined forecasting using neural networks. *Expert Systems with Applications*, 16(1), 49–54.
- Newbold, P., & Granger, C. W. J. (1974). Experience with Forecasting Univariate Time Series and the Combination of Forecasts. *Journal of the Royal Statistical Society. Series A (General)*, 137(2), 131–165.
- Newbold, P., & Harvey, D. I. (2004). *Forecast Combination and Encompassing*. John Wiley & Sons, Ltd.
- Öller, L.-E. (1978). A Method for Pooling Forecasts. *Journal of the Operational Research Society*, 29(1), 55–63.
- Patton, A. J., & Timmermann, A. (2007). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics*, 140(2), 884–918.
- Petropoulos, F., Hyndman, R. J., & Bergmeir, C. (2018). Exploring the sources of uncertainty: Why does bagging for time series forecasting work? *European Journal of Operational Research*, 268(2), 545–554.
- Petropoulos, F., & Svetunkov, I. (2020). A simple combination of univariate models [M4 Competition]. *International Journal of Forecasting*, 36(1), 110–115.
- Politis, D. N., & Romano, J. P. (1991). *A circular block-resampling procedure for stationary data*. Purdue University. Department of Statistics.
- Politis, D. N., & Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313.
- Poskitt, D. S. (2008). Properties of the Sieve Bootstrap for Fractionally Integrated and Non-Invertible Processes. *Journal of Time Series Analysis*, 29(2), 224–250.
- Poskitt, D., Grose, S. D., & Martin, G. M. (2015). Higher-order improvements of the sieve bootstrap for fractionally integrated processes. *Journal of Econometrics*, 188(1), 94–110.
- Qian, W., Rolling, C., Cheng, G., & Yang, Y. (2019). On the Forecast Combination Puzzle. *Econometrics*, 7(3).
- Radovanov, B., & Marcikić, A. (2014). A comparison of four different block bootstrap methods. *Croatian Operational Research Review*, 189–202.

- Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence. *The annals of Statistics*, 1048–1072.
- Robinson, P. M. (2003). *Time series with long memory*. Advanced Texts in Econometrics.
- Rossi, B. (2021). Forecasting in the Presence of Instabilities: How We Know Whether Models Predict Well and How to Improve Them. *Journal of Economic Literature*, 59(4), 1135–90.
- Said, S. E., & Dickey, D. A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71(3), 599–607.
- Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel*, 81(10.5555), 26853.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464.
- Shaub, D. (2020). Fast and accurate yearly time series forecasting with forecast combinations [M4 Competition]. *International Journal of Forecasting*, 36(1), 116–120.
- Smith, J., & Wallis, K. (2009). A Simple Explanation of the Forecast Combination Puzzle. *Oxford Bulletin of Economics and Statistics*, 71(3), 331–355.
- Stock, J., & Watson, M. (1999). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. In R. Engle & H. White (Cur.), *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger* (pp. 1–44). Oxford University Press.
- Stock, J., & Watson, M. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6), 405–430.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, 7(1), 13–26.
- Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.
- Team, R. C. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. (*No Title*).
- Terui, N., & van Dijk, H. K. (2002). Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting*, 18(3), 421–438.
- Timmermann, A. (2006). Chapter 4 Forecast Combinations. In G. Elliott, C. Granger & A. Timmermann (Cur.). Elsevier.
- Ṭūsūn, U. (1925). *Mémoire sur l'histoire du Nil ; Par le Prince Omar Toussoun: ...* Inst. franç. d'archéol. orient.

- Wang, X., Hyndman, R. J., Li, F., & Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, *39*(4), 1518–1547.
- Wang, X., Kang, Y., Petropoulos, F., & Li, F. (2022). The uncertainty estimation of feature-based forecast combinations. *Journal of the Operational Research Society*, *73*(5), 979–993.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838.
- Winkler, R. L., Grushka-Cockayne, Y., Jr., K. C. L., & Jose, V. R. R. (2019). Probability Forecasts and Their Combination: A Research Perspective. *Decision Analysis*, *16*(4), 239–260.
- Yahoo Finance. (2024). Historical Data for Microsoft Corporation (MSFT). <https://finance.yahoo.com/quote/MSFT/history/>
- Zellner, A., Keuzenkamp, H., & McAleer, M. (2002). *Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple*. Cambridge University Press.
- Zischke, R., Martin, G. M., Frazier, D. T., & Poskitt, D. S. (2022). The Impact of Sampling Variability on Estimated Combinations of Distributional Forecasts.