

UNIVERSITA' DEGLI STUDI DI PADOVA

FACOLTA' DI SCIENZE STATISTICHE

CORSO DI LAUREA IN STATISTICA, ECONOMIA E

FINANZA

Tesi di laurea

LA PREVISIONE DEI PREZZI DELL'ELETTRICITA'

CON MODELLI LINEARI PER SERIE STORICHE

Relatore: Ch.mo prof. LUISA BISAGLIA

Laureanda: Destro Ilaria

Matricola: 537531-SEF

Anno Accademico 2007/2008

Alla mia famiglia e ai miei amici

INDICE

| | |
|--------------------------|----------|
| INTRODUZIONE..... | 7 |
|--------------------------|----------|

CAPITOLO PRIMO

| | |
|--|-----------|
| CARATTERISTICHE EMPIRICHE DEI PREZZI DEL MERCATO ELETTRICO..... | 11 |
|--|-----------|

| | |
|---|----|
| 1.1 CARATTERISTICHE..... | 11 |
| 1.1.1 Alta frequenza di osservazione..... | 12 |
| 1.1.2 Periodicità..... | 12 |
| 1.1.3 Distribuzione empirica non approssimabile da una distribuzione normale..... | 14 |
| 1.1.4 Volatilità, presenza dei “ <i>jumps</i> ” e della “ <i>mean-reversion</i> ” | 16 |

CAPITOLO SECONDO

| | |
|---|-----------|
| PREVISIONE DEI PREZZI DELL’ELETTRICITA’..... | 21 |
| 2.1 CLASSIFICAZIONE DEI MODELLI PER LE PREVISIONI..... | 21 |
| 2.1.1 Modelli basati sui costi di produzione..... | 22 |
| 2.1.2 Modelli di equilibrio..... | 23 |
| 2.1.3 Modelli econometrici e modelli statistici per serie storiche..... | 23 |

| | |
|--|----|
| 2.1.4 Modelli non parametrici..... | 24 |
| 2.2 MODELLISTICA..... | 25 |
| 2.2.1 Modelli ARMA-ARIMA..... | 25 |
| 2.2.2 Modelli SARIMA..... | 27 |
| 2.2.3 Modelli ARCH-GARCH..... | 29 |
| 2.2.4 Modelli ARMAX..... | 31 |
| 2.2.5 Modelli <i>regime-switching</i> | 33 |
| 2.2.6 Modelli <i>jump-diffusion</i> | 36 |
| 2.3 MISURE PER VERIFICARE LA BONTA' DELLE PREVISIONI..... | 38 |

CAPITOLO TERZO

| | |
|--|-----------|
| BREVE PRESENTAZIONE DEL MERCATO IPEX ED ANALISI DEI DATI..... | 41 |
| 3.1 IL MERCATO ELETTRICO ITALIANO..... | 41 |
| 3.2 ANALISI EMPIRICA DI UN GRUPPO DI DATI APPARTENENTI AL MERCATO IPEX..... | 45 |
| 3.2.1 Presentazione dei dati..... | 45 |
| 3.2.2 Analisi descrittive..... | 47 |
| 3.2.3 Stima dei modelli..... | 53 |
| 3.2.4 Previsioni..... | 60 |

CONCLUSIONI.....65

APPENDICE 1.....67

APPENDICE 2.....72

BIBLIOGRAFIA.....97

INTRODUZIONE

Nel corso degli ultimi vent'anni una serie di Paesi, tra cui anche quelli dell'Europa occidentale, hanno conosciuto un periodo di profonde trasformazioni caratterizzato dall'affermarsi del cosiddetto "Liberalismo economico". L'idea alla base di questo movimento è quella di libera concorrenza tra gli agenti economici e di pari opportunità di accesso al mercato. Tale ideologia, nata principalmente nell'ambito economico-giuridico, si è poi espansa anche ad attività che tradizionalmente erano caratterizzate da un monopolio naturale dello Stato, tra le quali il settore dell'energia elettrica. I punti chiave su cui poggiava la tradizione monopolistica erano la centralizzazione ed una struttura rigorosamente verticalizzata del mercato elettrico in tutte le fasi della filiera produttiva, vale a dire che un'unica impresa controllava e possedeva più fasi successive della catena del valore; come si può immaginare, in un tale contesto veniva chiaramente lasciato gran poco spazio alla concorrenza. Quest'ultima rappresenta invece l'idea centrale della nuova situazione di liberalismo: all'elevata chiusura economica è stata sostituita una struttura

di mercato molto più aperta, decentralizzata, in cui molteplici produttori, grossisti e consumatori interagiscono con gli obiettivi comuni di garantire un'ottimizzazione rispettivamente dei propri profitti e della propria utilità e di riuscire a raggiungere quindi uno stato di benessere generale.

In tale nuova atmosfera iniziano a sorgere due tipi di strumenti atti a facilitare gli scambi tra i produttori ed i consumatori: un contesto che permette di stipulare contratti fisici bilaterali ed il cosiddetto *pool*. Il primo tipo di strumento prevede accordi tra produttori e grossisti riguardo all'offerta di determinate quantità di elettricità per un certo orizzonte temporale; questi contratti sono poi perfezionati con l'aiuto di un Operatore Indipendente di Sistema, il cui compito è quello di controllare i requisiti necessari a compiere le transazioni in maniera sicura. Il secondo è invece un luogo di commercio virtuale (chiamato *power exchange*) nel quale i produttori di energia elettrica presentano offerte di vendita con i relativi prezzi minimi e le quantità a cui essi sono disposti a vendere, ed allo stesso tempo i grossisti ed i consumatori presentano offerte d'acquisto con i relativi prezzi massimi e le quantità a cui sono disposti ad acquistare il bene in questione. Il prezzo finale (o *spot*) è determinato dall'operatore di mercato attraverso un'asta condotta incrociando il sistema di offerta e di domanda aggregata. Tale prezzo viene determinato ogni giorno per qualsiasi ora (o mezz'ora) del giorno successivo.

Proprio per questo motivo è diventata importante la questione delle previsioni dei prezzi dell'elettricità: i produttori, i grossisti ed i consumatori hanno bisogno di informazioni il più accurate possibile per ottimizzare le proprie strategie di offerta necessarie a determinare il prezzo *spot*. Tali informazioni possono essere fornite loro tramite un'analisi statistica delle caratteristiche dei prezzi del mercato elettrico – che, come vedremo in seguito, sono abbastanza anomale-, una successiva individuazione di modelli adatti a stimarle e che permettano anche di effettuare delle buone previsioni non solo a breve termine (orizzonte giornaliero) ma anche a medio-lungo termine (orizzonte mensile).

Questo lavoro ha come oggetto principale proprio la trattazione di modelli adatti a prevedere il meglio possibile i prezzi del mercato elettrico.

Nei vari capitoli articoleremo la questione nei seguenti punti:

- presentazione delle caratteristiche principali che identificano i prezzi dell'elettricità;
- esposizione di modelli utili per fare previsioni;
- breve resoconto della situazione attuale del mercato elettrico italiano, con particolare riferimento al funzionamento della Borsa elettrica italiana, nata il 1° aprile 2004, ed analisi empirica di una serie storica di prezzi relativi al mercato elettrico italiano, effettuata applicando alcuni dei modelli menzionati sopra.

CAPITOLO PRIMO

CARATTERISTICHE EMPIRICHE DEI PREZZI DEL MERCATO ELETTRICO

1.1 Caratteristiche

Come è stato già accennato nella parte introduttiva, i prezzi dell'elettricità, a dispetto di alcuni aspetti riguardanti la loro distribuzione - come ad esempio la presenza di code pesanti e la persistenza nei quadrati delle serie - presentano delle caratteristiche molto differenti da quelle tipiche dei mercati finanziari.

Tali caratteristiche sono comuni alla maggior parte dei mercati elettrici mondiali, e molte di esse sono sorte in seguito alle nuove modalità di formazione dei prezzi *spot* introdotte con la riforma di tipo liberale avvenuta recentemente.

Specificatamente esse sono:

- ❖ alta frequenza di osservazione;
- ❖ periodicità;
- ❖ distribuzione empirica non approssimabile con una distribuzione normale;

- ❖ volatilità molto elevata, presenza dei cosiddetti “jumps” e del fenomeno chiamato “*mean-reversion*”.

1.1.1. Alta frequenza di osservazione

I prezzi dell'elettricità presentano una frequenza molto elevata, oraria o addirittura ogni mezz'ora. Essi hanno la particolarità di poter essere forniti sempre ad intervalli di tempo costanti, vale a dire per tutte le 24 ore di ogni giorno dell'anno (anche per i giorni festivi).

1.1.2. Periodicità

Analizzando le serie storiche dei prezzi elettrici si nota che possono essere riscontrati fondamentalmente tre tipi di periodicità: mensile, settimanale ed intragiornaliera.

Il primo tipo di periodicità è causato dalla grossa influenza sulla domanda di energia elettrica delle condizioni climatiche. In alcuni paesi (come ad esempio i paesi scandinavi) il periodo dell'anno in cui la domanda è più elevata sono i mesi invernali, durante i quali la temperatura è molto bassa ed il numero di ore di luce è al minimo, e quindi vengono utilizzati molto

gli impianti di riscaldamento e l'illuminazione elettrica. In altri paesi (come ad esempio la parte occidentale degli USA) quelli in cui la domanda è più elevata sono i mesi estivi, in cui la temperatura è molto alta, c'è una grande umidità e quindi si sente il bisogno di accendere gli impianti di condizionamento. Di conseguenza, all'aumento della domanda, i prezzi *spot* in questi periodi si innalzano notevolmente.

Il secondo tipo di periodicità è dovuto alla differenza della quantità di domanda di energia elettrica tra i giorni lavorativi ed i week-end: durante questi ultimi la maggior parte delle attività sono chiuse, quindi si verifica una diminuzione della domanda di energia, e dunque dei prezzi.

Il terzo tipo di periodicità è invece legato ai ritmi giornalieri delle persone: infatti i prezzi massimi, corrispondenti ai consumi massimi, solitamente si riscontrano nelle ore mattutine, quando ci si alza dal letto e ci si reca al lavoro o a scuola, e verso le 17-19, quando abitualmente si rientra dal lavoro e si inizia a preparare la cena, a guardare la tv,...

La periodicità – di qualunque tipo essa sia- può essere modellata attraverso l'uso di variabili *dummy* (che assumono valore 1 se si è nel periodo j , 0 altrimenti), di medie mobili o di funzioni sinusoidali, oppure considerando

modelli stagionali come, per esempio, gli ARIMA stagionali (che verranno trattati in seguito).

1.1.3. Distribuzione empirica non approssimabile con una distribuzione normale

Una caratteristica che accomuna i prezzi dell'elettricità a quelli delle attività finanziarie è che non è possibile approssimare la loro distribuzione con quella normale a causa della presenza di asimmetria (in genere positiva, vale a dire che la distribuzione appare con una coda più lunga verso destra) e, soprattutto, a causa della presenza di code molto più pesanti di quelle tipiche di una distribuzione normale: quest'ultimo fenomeno è noto con il nome di *leptocurtosi*, e la distribuzione che ne deriva è più appuntita della normale. La Figura 1.1, in cui è raffigurato l'istogramma dei prezzi orari (espressi in Euro per Megawatt, Euro/MWh) del mercato elettrico tedesco *Leipzig Power Exchange* (LPX) nel periodo 16 Giugno 2000 – 15 Ottobre 2001, fornisce un'idea più precisa di questi due aspetti dei prezzi dell'elettricità: si nota infatti una coda abbastanza pesante verso destra (non a caso l'indice di asimmetria è positivo e pari a 2,37 e l'indice di curtosi molto elevato e pari a 23,5). Secondo gli studiosi Knittel e Roberts (2005) queste caratteristiche suggeriscono la presenza di un

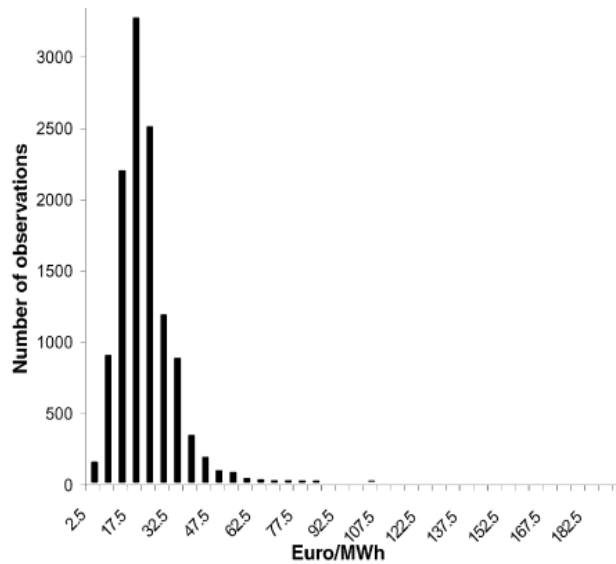


Figura 1.1: Istogramma dei prezzi *spot* del mercato LPX dal 16 Giugno 2000 al 15 Ottobre 2001.

“*effetto leverage*” inverso: mentre nel caso dei mercati finanziari sembra che la volatilità futura sia maggiormente influenzata da *shocks* negativi nel sistema dei prezzi, nel caso dei mercati elettrici sembra che la stessa sia maggiormente influenzata da *shocks* positivi, e molto meno da quelli negativi.

Per modellare l’asimmetria e la leptocurtosi esistono svariati modelli: i più semplici e più diffusi sono i modelli del tipo ARCH/GARCH (che saranno illustrati in maniera più approfondita nel prossimo capitolo).

1.1.4. Volatilità, presenza dei “jumps” e della “mean-reversion”

Queste sono probabilmente le caratteristiche più studiate e quelle che maggiormente contraddistinguono le serie dei prezzi dei mercati elettrici.

Per cominciare ad avere un'idea della grande rilevanza di questi aspetti nelle serie in questione, compariamo i valori della volatilità attesa (calcolata tradizionalmente come la deviazione standard dei rendimenti) per i prezzi dell'elettricità con quella che tradizionalmente ci si aspetta per alcuni tipi di attività finanziarie a livello giornaliero:

- Buoni del Tesoro e cambiali: *meno di 0.5%*;
- azioni ed obbligazioni: *tra 1% e 1.5%*;
- petrolio greggio e gas naturale: *tra 1.5% e 4%*;
- azioni ed obbligazioni considerate molto volatili: *al più 4%*;
- prezzi dell'elettricità: *più del 50%!*

Come si può notare, la volatilità attesa delle serie che ci interessano per la nostra analisi è straordinariamente più alta di quella di altre attività finanziarie, anche nel caso in cui esse siano generalmente abbastanza rischiose.

Quali sono le cause alla base della presenza della grande variabilità nei prezzi dell'elettricità? E perché essa è così prepotentemente elevata?

Esistono diverse risposte. La prima, e probabilmente la principale, è il fatto che non si può immagazzinare energia elettrica in maniera diretta, perciò essa deve essere utilizzata nel momento stesso in cui viene prodotta. Ciò comporta la necessità di bilanciare in continuazione la produzione con il consumo, ma allo stesso tempo l'impossibilità di equilibrare all'istante offerta e domanda, causando così degli *shocks* che avranno influenza sui prezzi.

La seconda motivazione è una conseguenza del meccanismo domanda-offerta. La domanda di elettricità è altamente inelastica, dal momento che essa è un bene necessario (cioè l'elasticità della domanda rispetto al reddito è <1) e particolarmente dipendente dalle condizioni atmosferiche; ad una domanda di questo tipo non corrisponde però un'offerta soddisfacente, in quanto i generatori riescono ad offrire solamente determinate quantità di energia, a seconda delle loro capacità, con costi marginali diversi l'uno dall'altro. Inoltre, un'improvvisa crisi operativa (per esempio, la rottura di una linea di trasmissione) può creare un deficit di offerta non sostituibile con altre fonti, e pertanto un aumento subitaneo e improvviso del prezzo dell'energia.

Queste circostanze conducono, nei momenti di maggior richiesta, ad un incremento notevole del prezzo in poco tempo.

Entrambi i motivi analizzati trovano una loro concretizzazione nel fenomeno noto come *jump* o *spike*, il quale consiste in un cambiamento repentino dei prezzi verso l'alto o verso il basso: se ne può trovare una rappresentazione nella Figura 1.2, che raffigura il grafico dei prezzi *spot* provenienti dal mercato LPX dei quali è stato analizzato l'istogramma nel paragrafo 1.1.3.

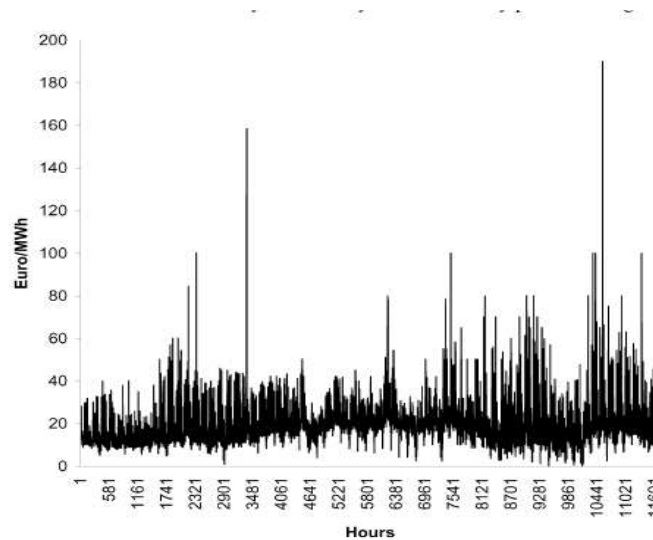


Figura 1. 2: Grafico dei prezzi *spot* del mercato LPX nel periodo 16 Giugno 2000 – 15 Ottobre 2001.

Alcuni studi - specialmente riguardo alle serie orarie dei prezzi elettrici californiani e scandinavi- hanno dimostrato che gli *spikes* si manifestano soprattutto nelle ore di punta dei giorni lavorativi (09:00 e 18:00) e nei mesi di maggior consumo di energia elettrica (mesi rispettivamente estivi

18

ed invernali), e questi cambiamenti repentini solitamente sono verso l'alto. Questo risultato non sorprende, poiché i periodi menzionati sono esattamente quelli in cui nei rispettivi paesi la domanda è più alta tra tutti i momenti dell'anno.

Gli *spikes* sono solitamente non equispaziati tra di loro e di dimensioni che possono variare abbastanza (anche se parliamo sempre di livelli molto superiori alla media della serie che stiamo analizzando); essi sono ben visibili a livello orario o giornaliero, mentre man mano che si estende l'orizzonte temporale (settimane o mesi) lo diventano sempre meno. Un'altra loro caratteristica è la durata brevissima: non appena gli shocks economici che ne stanno alla base lasciano spazio ad un nuovo stato di equilibrio, i prezzi si riassetano immediatamente tornando ai livelli precedenti l'avvento degli *spikes*. Per questo motivo i prezzi nei mercati elettrici sono definiti non-persistenti o *mean-reverting* (espressione che propriamente significa "che ritornano al livello medio"). In letteratura - più in particolare secondo un'analisi effettuata da Escibano A., Pena J.I., Villaplana P. nel 2002 - si trovano due possibili spiegazioni al fenomeno di *mean-reversion*: da un lato, dal momento che l'aumento della domanda fa aumentare i prezzi, con un aumento degli incentivi economici dei generatori più costosi aumenta l'offerta, e quindi ci si aspetta un complessivo ritorno alla media; dall'altro, invece, si considera il fatto che

le serie dei prezzi dell'elettricità sono periodiche e perciò caratterizzate per definizione dal ritorno alla media, e questo influenza la domanda ed i prezzi di equilibrio.

L'esistenza dei *jumps* nei prezzi dell'elettricità ha ripercussioni notevoli a livello finanziario: chi decide di investire nei mercati elettrici, a causa di essi può veder svanire i propri guadagni tutto d'un tratto, anche nel giro di poche ore. Ma c'è anche chi decide di approfittare dell'occasione per speculare: essendo l'elettricità un bene necessario ed essendoci quindi agenti disposti ad acquistarla a qualunque prezzo, alcuni venditori applicano a questo bene il prezzo massimo consentito. E' proprio per questa ragione che gli *spikes* sono così tanto elevati.

Modelli particolarmente adatti a cogliere la caratteristica dei *jumps* sono i modelli *jump-diffusion* (che verranno anch'essi analizzati nel capitolo successivo).

CAPITOLO SECONDO

PREVISIONE DEI PREZZI DELL'ELETTRICITA'

2.1 Classificazione dei modelli per le previsioni

Nell'introduzione abbiamo parlato del ruolo di grande importanza svolto dalle previsioni dei prezzi dell'elettricità nel recente contesto di liberalismo.

In questo paragrafo presenteremo i principali modelli studiati al fine di riuscire ad effettuare delle previsioni soddisfacenti, classificandoli sia in termini di pianificazione dell'orizzonte di durata sia in termini di metodologia applicata.

Nell'orizzonte temporale a lungo termine l'obiettivo è l'analisi e la pianificazione della profittabilità degli investimenti, in modo da determinare i siti futuri di centrali elettriche o le piene risorse delle industrie di potere.

Gli scopi principali perseguiti negli orizzonti temporali a medio termine sono quelli di bilanciare calcoli pesanti, di organizzare il rischio ed i prezzi derivati; in molti casi essi si concentrano non sulle attuali previsioni

puntuali, bensì sulla distribuzione di prezzi futuri durante determinati periodi di tempo.

Infine, le previsioni a breve termine sono particolarmente interessanti per coloro che partecipano ai mercati *spot* con asta ai quali viene richiesto giornalmente di esprimere un'offerta in termini di prezzi e di quantità (per capirne meglio il meccanismo si rimanda alla parte introduttiva).

Per quanto riguarda la classificazione secondo la metodologia applicata si possono distinguere quattro classi:

- modelli basati sui *costi di produzione*;
- modelli di *equilibrio*;
- modelli *econometrici* e modelli *statistici per serie storiche*;
- modelli *non parametrici*.

2.1.1. Modelli basati sui costi di produzione

Tali modelli simulano l'operazione di generazione di energia elettrica in grado di soddisfare la domanda esistente al minimo costo; le previsioni dei prezzi possono arrivare ad essere fatte anche a livello orario. Tuttavia questo tipo di modelli era più adatto al contesto economico presente prima del liberalismo, e quindi essi non riescono a tenere conto dell'importanza delle strategie competitive.

2.1.2. Modelli di equilibrio

Essi possono essere considerati come una sorta di ammodernamento dei modelli basati sui costi di produzione, nel senso che a differenza di questi ultimi tengono conto anche dell'importanza della competizione e delle strategie riguardo ad essa. Tramite questi approcci si può comprendere come i prezzi possano essere superiori ai costi marginali, e come questo possa influenzare i risultati delle scelte degli agenti economici. Però c'è da dire che è presente tra i concorrenti un sostanziale rischio di non corretta specificazione del modello nelle loro strategie potenziali e nell'insieme di profitti che devono essere definiti subito.

2.1.3. Modelli econometrici e modelli statistici per serie storiche

Tramite i modelli econometrici gli agenti economici cercano di riprodurre le proprietà statistiche dei prezzi dell'elettricità con il fine ultimo di una valutazione derivata e di una gestione del rischio; questo approccio ha portato all'introduzione di logiche strategiche finanziarie anche nel settore elettrico.

La classe di modelli in questione non è particolarmente utile per prevedere accuratamente i prezzi a livello orario, bensì per illustrare le caratteristiche

dei prezzi dell'elettricità su una scala più larga, solitamente giornaliera o mensile.

Alcuni modelli più strettamente statistici che risultano maggiormente utili al fine di effettuare previsioni anche a livello orario sono i tradizionali modelli per serie storiche univariati (del tipo ARIMA /SARIMA), oppure le tecniche di lisciamiento.

Per accrescere la loro efficienza, molti di questi approcci incorporano anche l'eventuale impatto sui prezzi dell'elettricità di altri fattori fisici ed economici *esogeni*, come ad esempio il prezzo di caricamento, fattori legati alle condizioni climatiche, il prezzo del combustibile, ed altri.

2.1.4. Modelli non parametrici

Questa classe comprende tecniche come il modello del prezzo elaborato attraverso strumenti non parametrici (come ad esempio reti neurali artificiali o ANN_s), sistemi esperti, logica incoerente e vettori macchine di supporto. Tali modelli tendono ad essere flessibili e possono occuparsi di complessità e non linearità; queste caratteristiche li rendono promettenti per previsioni a breve termine. Tuttavia, si è notato che i modelli in questione vengono quasi sempre confrontati solo tra di loro; il confronto con modelli

statistici più semplici è abbastanza arduo, e rappresenta tuttora un argomento di ricerche per il futuro.

2.2 Modellistica

2.2.1. Modelli ARMA-ARIMA

Uno dei modelli per serie storiche maggiormente utilizzati per tenere in considerazione la natura casuale e la struttura di correlazione temporale di un certo fenomeno è il modello Auto Regressivo a Media Mobile, abbreviato con ARMA; nel modello $ARMA(p,q)$ il prezzo attuale P_t è espresso linearmente in termini dei suoi p valori passati (parte autoregressiva) e dei q precedenti valori del *noise* (parte a media mobile):

$$\phi(\mathbf{B})P_t = \theta(\mathbf{B})\varepsilon_t \quad (1)$$

Nell'equazione (1) \mathbf{B} è l'operatore ritardo, tale che $\mathbf{B}^h P_t = P_{t-h}$; inoltre:

$$\phi(\mathbf{B}) = 1 - \phi_1 \mathbf{B} - \dots - \phi_p \mathbf{B}^p$$

$$\theta(\mathbf{B}) = 1 + \theta_1 \mathbf{B} + \dots + \theta_q \mathbf{B}^q$$

dove $\phi(\mathbf{B})$ e $\theta(\mathbf{B})$ sono polinomi di grado rispettivamente p e q in \mathbf{B} . Per la stazionarietà e l'invertibilità le radici delle equazioni caratteristiche

associate a tali polinomi ($\phi(B) = 0$ e $\theta(B) = 0$) devono essere maggiori di 1 in modulo.

ϕ_1, \dots, ϕ_p sono i coefficienti del polinomio autoregressivo, mentre $\theta_1, \dots, \theta_q$ sono quelli della parte a media mobile; ε_t è indipendentemente ed identicamente distribuito (i.i.d.) come un *white noise* di media 0 e varianza finita (generalmente con distribuzione gaussiana).

Per $q=0$ si ottiene il modello auto regressivo AR(p), mentre per $p=0$ otteniamo il processo a media mobile MA(q). L'approccio di modellazione ARMA ha come assunto che la serie storica che si sta analizzando sia debolmente stazionaria¹; se non lo è, per prima cosa si deve effettuare una trasformazione della serie per renderla tale. Il tipo di trasformazione più diffuso è la differenziazione della serie (Box e Jenkins, 1976). Il risultato è noto come modello auto regressivo a media mobile differenziato (o *ARIMA*).

¹ Una serie (e più in generale un processo stocastico) ha *stazionarietà debole* se rispetta le seguenti tre condizioni:

- $E[Y_t] = \mu$ per ogni t
- $\text{Var}[Y_t] = \sigma^2$ per ogni t
- $\text{Cov}[Y_t, Y_{t-k}] = \gamma_k$ per ogni t e per ogni k

dove Y_t e Y_{t-k} rappresentano i valori della serie (o del processo) ai momenti t e t-k.

2.2.2. Modelli SARIMA

Questo tipo di modelli è un'estensione dei processi ARIMA proposta da Box e Jenkins nel 1976 e consente di modellare andamenti periodici di tipo stagionale. La componente stagionale può essere deterministica ed indipendente dalle altre componenti oppure stocastica e correlata con le componenti non stagionali. Inoltre grazie a questo approccio può venire trattata anche una non stazionarietà di tipo periodico.

Nella sua forma più generale un processo $SARIMA(p,d,q) \times (P,D,Q)_s$ può essere scritto nel modo seguente:

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^DP_t = \theta(B)\Theta(B^S)\epsilon_t \quad (2)$$

Nell'equazione (2):

- ✓ P_t è il prezzo al tempo t ;
- ✓ S rappresenta il periodo della stagionalità;
- ✓ $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ è l'operatore autoregressivo non stagionale di ordine p stazionario;
- ✓ $\Phi(B^S) = (1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS})$ è l'operatore autoregressivo stagionale di ordine P stazionario;

- ✓ $\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)$ è l'operatore a media mobile non stagionale di ordine q invertibile;
- ✓ $\Theta(B^S) = (1 + \Theta_1 B^S + \Theta_2 B^{2S} + \dots + \Theta_Q B^{QS})$ è l'operatore a media mobile stagionale di ordine Q invertibile;
- ✓ $(1-B)^d$ è l'operatore differenza di ordine d non stagionale;
- ✓ $(1-B^S)^D$ è l'operatore differenza di ordine D stagionale;
- ✓ $\varepsilon_t \sim \text{WN}(0, \sigma^2)$.

Il concetto che sta alla base di tale modello è il fatto che osservazioni che distano S periodi, come $P_t, P_{t-S}, P_{t-2S}, \dots$ dovrebbero essere simili e fortemente correlate tra loro. Il periodo S assume il valore 4 nel caso di serie trimestrali, 12 nel caso di serie mensili,...

Nell'equazione (2) l'eventuale non stazionarietà viene modellata secondo la stessa logica di modellazione dei processi ARIMA: eventuali non stazionarietà stagionali possono essere eliminate tramite l'operatore $(1-B^S)^D$, mentre gli operatori $\Phi(B^S)$ e $\Theta(B^S)$ trattano la dipendenza tra osservazioni distanti $S, 2S, 3S, \dots$ istanti temporali.

Si può notare che, scomponendo $(1-B^S)$ in $(1-B)(1+B+B^2+\dots+B^{S-1})$, l'operazione di differenziazione stagionale comporta necessariamente anche una differenziazione del primo ordine, che ha lo scopo di rimuovere eventuali trend lineari.

2.2.3. Modelli ARCH-GARCH

I modelli ARMA ed ARIMA presuppongono l'omoschedasticità, cioè funzioni di varianza e di covarianza costanti. In realtà, da un punto di vista empirico, i prezzi *spot* dell'elettricità presentano varie forme di non linearità, tra cui in particolare proprio una forte dipendenza della variabilità delle serie dal proprio passato.

Ciò implica l'esistenza di una varianza condizionale non costante e, quindi, di eteroschedasticità.

Uno dei primi approcci allo studio di questo fenomeno è stato la modellistica *ARCH* (*AutoRegressive Conditional Heteroschedasticity*), introdotta da Engle nel 1982, la quale parte proprio dall'idea di varianza condizionata all'informazione disponibile. In tale modello si assume che il prezzo corretto per la media, h_t , sia incorrelato ma dipendente, e che la dipendenza di h_t possa venire descritta da una funzione quadratica dei suoi valori ritardati.

Più specificatamente un modello ARCH di ordine m può essere scritto nel modo seguente:

$$\mathbf{h}_t = \boldsymbol{\varepsilon}_t \boldsymbol{\sigma}_t, \text{ con}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 h_{t-1}^2 + \dots + \alpha_m h_{t-m}^2 \quad (3)$$

con ε_t variabile casuale iid(0,1), avente distribuzione Normale o t di Student; inoltre dev'essere $\alpha_0 > 0$ e $\alpha_i \geq 0$ per assicurare la positività di σ^2 . Infine, i coefficienti α_i devono soddisfare dei vincoli perché la varianza non condizionata di h_t sia finita.

Nel 1986 Bollerslev estese il modello ARCH sviluppando il modello *GARCH (Generalized AutoRegressive Conditional Heteroschedasticity)*.

Una possibile espressione per rappresentare un modello GARCH(m,s) è:

$$\mathbf{h}_t = \varepsilon_t \sigma_t, \text{ con}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i h_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (4)$$

dove ε_t è sempre una variabile casuale iid (0,1), $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$,

$$\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1.$$

Le prime tre condizioni garantiscono la non negatività della varianza condizionata, l'ultima garantisce che la varianza non condizionata sia finita. Ovviamente, osservando l'espressione del modello, nel caso in cui s fosse =0, si cadrebbe nel caso di un ARCH(m).

2.2.4. Modelli ARMAX

I modelli ARMA relazionano il segnale sotto osservazione al proprio passato, senza utilizzare esplicitamente informazioni contenute in altre serie storiche. Tuttavia, in molti casi, un certo segnale non è riferito solamente al proprio passato, ma può essere influenzato anche da valori presenti e passati di variabili esogene ad esso. Questa è una situazione particolarmente diffusa nei prezzi dell'elettricità. Oltre a variazioni stagionali, tali prezzi sono infatti generalmente governati da fattori quali ad esempio i profili di carico, le condizioni atmosferiche, la temperatura. Per riuscire a catturare adeguatamente la relazione esistente con questi fattori fondamentali, si utilizza un modello di tipo ARMAX (cioè Auto Regressivo a Media Mobile con variabili esogene). Un modello $ARMAX(p,q,r_1,\dots,r_k)$ può essere scritto in modo compatto come (Ljung,1999):

$$\phi(\mathbf{B})P_t = \theta(\mathbf{B})\varepsilon_t + \sum_{i=1}^k \psi^i(\mathbf{B}) v_t^i \quad (5)$$

Nell'equazione (5) P_t è il prezzo *spot* al tempo t , r_i sono gli ordini dei k fattori esogeni, rappresentati come v^1, \dots, v^k , e $\psi^i(B)$ è una notazione abbreviata per la funzione relativa al fattore i : $\psi^i(B) = \psi_0^i + \psi_1^i B + \dots +$

$\psi_{r_i}^i B^{r_i}$, dove ψ_j^i sono i corrispondenti coefficienti. In alternativa, il modello ARMAX è talvolta definito nella forma di una “funzione di trasferimento” nel modo espresso nell’equazione (6):

$$P_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t + \sum_{i=1}^k \tilde{\psi}^i(B) v_t^i \quad (6)$$

dove $\tilde{\psi}^i$ rappresentano gli appropriati coefficienti polinomiali.

In un lavoro effettuato da F. J. Nogales, J. Contreras, A. J. Conejo e R. Espinola nel 2002, la “funzione di trasferimento” in questione è scritta come:

$$P_t = \omega^d(B) d_t + N_t \quad (6)$$

dove $\frac{\theta(B)}{\phi(B)} \varepsilon_t = N_t$ e la variabile esogena è la domanda di energia elettrica, rappresentata come d_t , con $\omega^d(B)$ relativa funzione polinomiale dell’operatore ritardo B contenente i coefficienti relativi a d_t . In tale analisi la funzione di trasferimento sembra risolvere efficacemente il problema della presenza di correlazione seriale (la quale verrà spiegata e trattata meglio nella parte riguardante l’analisi empirica) nei modelli di regressione lineare tradizionalmente usati per i prezzi dell’elettricità .

Nei vari studi effettuati a riguardo, i modelli di serie storiche con variabili esogene sono applicati soprattutto per calcolare le previsioni di prezzo a breve termine.

2.2.5. Modelli regime-switching

Tutti i modelli trattati finora sono lineari. Tuttavia, la presenza di *spikes* nelle serie dei prezzi dell'elettricità suggerisce l'esistenza di un meccanismo non lineare tra la situazione "normale" e quella con *jumps*. Ciò implica una distinzione tra questi stati, la quale può essere trattata attraverso i cosiddetti modelli "*regime-switching*".

L'idea alla base di essi è proprio quella di modellare il comportamento osservato di una determinata serie storica tramite due o più fasi (o "*regimi*") aventi differenti processi sottostanti.

Le specificazioni di tali modelli differiscono a seconda del modo in cui il regime si evolve nel tempo.

Tradizionalmente, tali modelli possono essere distinti in due classi: quelli nei quali il regime è determinabile attraverso una variabile osservabile (e, di conseguenza, i regimi che si sono verificati nel passato e nel presente sono conosciuti con certezza) e quelli nei quali il regime è determinabile attraverso una variabile latente, non osservabile. In quest'ultimo caso non

si riesce mai ad essere sicuri del fatto che un particolare regime si sia verificato in un particolare momento, ma si può solamente stimare la probabilità del suo accadimento.

Il più importante modello tra quelli appartenenti alla prima classe citata è il modello *TAR* (*Threshold Autoregressive*), il quale assume che il regime sia specificato dal valore di una variabile osservabile v_t relativa ad un valore soglia (*threshold* appunto) T :

$$f(x) = \begin{cases} \phi_1(B)P_t = \varepsilon_t, & v_t \geq T \\ \phi_2(B)P_t = \varepsilon_t, & v_t < T \end{cases} \quad (7)$$

Nella funzione (7) $\phi_i(B)$ è un'annotazione abbreviata per $\phi_i(B) = 1 - \phi_{i,1}B - \dots - \phi_{i,p}B^p$, $i=1,2$, e B è l'operatore ritardo. Per semplificare l'esposizione, è stato specificato un modello con due soli regimi, ma, naturalmente, è possibile – ed anzi spesso opportuna – una specificazione che ne preveda più di due.

Nel modello TAR è possibile includere anche delle variabili esogene, che portano al modello *TARX*.

Altri approcci sono poi il modello *SETAR* (*Self Exciting TAR*), il quale presume che la variabile soglia sia il valore ritardato della serie stessa dei prezzi, vale a dire $v_t = P_{t-d}$; inoltre esiste il modello *STAR* (*Smooth*

Transition AR), dove la variabile v_t è fissata permettendo una transizione graduale da un regime all'altro. Una scelta popolare per la funzione di transizione è la funzione logistica, che porta al modello conosciuto come *LSTAR (Logistic STAR)*. Una scelta piuttosto diffusa per esprimere la funzione logistica è:

$$G(P_{t-d}; \gamma, T) = \frac{1}{1 + \exp\{-\gamma (P_{t-d} - T)\}}$$

Tali modelli sembrano aver condotto a risultati abbastanza soddisfacenti; tuttavia, si ritiene che i prezzi *spot* dell'elettricità siano il prodotto di un vasto numero di fattori, sia esogeni – come ad esempio i profili di carico – sia di tipo psicologico o sociologico, non quantificabili, che possono causare acquisti inaspettati ed irrazionali del bene in questione che conducono a forti *spikes*.

In un tale contesto i modelli *Markov-switching*, dove il regime è determinato da una variabile casuale latente, sembrano interessanti.

L'utilità dei modelli *Markov-switching* nelle applicazioni sui mercati dell'elettricità, in particolare la loro capacità di modellare diversi *jumps* consecutivi –in opposizione al concetto alla base dei modelli *jump-diffusion*, che vedremo in seguito- è stata riconosciuta, ed è stato proposto un numero significativo di modelli per i prezzi *spot* dell'elettricità. Però, la loro adeguatezza per fare previsioni finora non è ancora stata testata approfonditamente; questo d'altra parte è un punto che solo recentemente è

stato affrontato nella letteratura, dunque probabilmente ci si dovranno aspettare dei risvolti.

2.2.6. Modelli jump-diffusion

I modelli *jump-diffusion* sono i più conosciuti modelli di tipo quantitativo utilizzati al fine di trattare la questione dei *jumps* nelle serie dei prezzi dell'elettricità.

Essi sono nati come possibile estensione ai mercati elettrici del *moto geometrico browniano*, il principale modello stocastico applicato nei mercati finanziari, il quale tuttavia non prende in considerazione caratteristiche peculiari dei prezzi dell'elettricità come i *jumps* e la *mean-reversion*.

Un modello *jump-diffusion* può essere espresso nel modo seguente (Kaminski,1997):

$$\frac{dP_t}{P_t} = \mu dt + \sigma dW_t + dq(t) \quad (8)$$

dove μ è lo spostamento, σ è la volatilità, W_t è il processo di moto browniano e $q(t)$ rappresenta un processo di Poisson omogeneo con intensità fissata e dimensione degli *spikes* avente distribuzione log-normale. Il modello *jump-diffusion* è essenzialmente costruito

addizionando una componente di Poisson, $dq(t)$, ad un moto geometrico browniano standard. Il suo difetto principale è il fatto che esso ignora un'altra caratteristica molto importante dei prezzi dell'elettricità: la mean-reversion allo stato “normale”, cioè senza *jumps*; infatti, il modello in esame assume che, se si verifica un *jump*, il nuovo livello raggiunto dal prezzo sia un evento “normale”, e che la probabilità di ritornare al livello precedente sia molto bassa (che, come sappiamo, è il contrario di quanto accade realmente nelle serie dei mercati elettrici).

In uno studio comparativo, Johnson e Barz (1999) valutano l'efficacia di vari modelli di diffusione nella descrizione dell'evoluzione dei prezzi *spot*; essi, tralasciando i moti browniani, sperimentano dei processi *mean-reverting diffusion* (altrimenti conosciuti come *processi aritmetici di Ornstein-Uhlenbeck*, proposti da Vasiček nel 1977 per modellare le dinamiche dei tassi d'interesse):

$$dP_t = (\alpha - \beta P_t)dt + \sigma dW_t \quad (9)$$

Inoltre essi sperimentano dei processi *mean-reverting diffusion geometrici*, con e senza *jumps*, nella forma di un processo di Poisson $q(t)$:

$$\frac{dP_t}{P_t} = (\alpha - \beta P_t)dt + \sigma dW_t \quad (10)$$

Gli studiosi in questione arrivano alle conclusioni che il modello (10) offra la *performance* migliore e che tutti i modelli che non includano *jumps* siano inappropriati per trattare i prezzi dell'elettricità.

2.3 Misure per verificare la bontà delle previsioni

Analizzare l'accuratezza delle previsioni è di importanza fondamentale; a tale scopo si confrontano i valori previsti con quelli osservati realmente, e si valuta l'entità degli errori di previsione, definiti nel modo seguente:

$$e_t = \text{valore vero-valore previsto}$$

Nella pratica vengono utilizzate diverse misure statistiche; supponendo di lavorare con serie orarie, le principali sono:

- l'errore medio assoluto giornaliero, calcolato come

$$\text{MAE} = \frac{1}{24} \sum_{h=1}^{24} |P_h - \bar{P}_h|$$

dove P_h è il prezzo reale per una determinata ora, mentre \bar{P}_h rappresenta il prezzo previsto per quella stessa ora;

- l'errore medio giornaliero, calcolato come:

$$\text{MDE} = \frac{1}{24} \sum_{h=1}^{24} \frac{|P_h - \bar{P}_h|}{\bar{P}_{24}}$$

dove \bar{P}_{24} è il prezzo medio per un certo giorno;

- l'errore medio giornaliero in percentuale assoluta:

$$\text{MAPE} = \frac{1}{24} \sum_{h=1}^{24} \frac{|P_h - \bar{P}_h|}{P_h}$$

- la radice dell'errore quadratico medio giornaliero:

$$\text{DRMSE} = \sqrt{\frac{1}{24} \sum_{h=1}^{24} (P_h - \bar{P}_h)^2}$$

- l'errore medio settimanale:

$$\text{MWE} = \frac{1}{168} \sum_{h=1}^{168} \frac{|P_h - \bar{P}_h|}{P_{168}}$$

dove $168=24*7$ ed è il numero di ore totali in una settimana, e \bar{P}_{168}

è il prezzo medio per una data settimana;

- la radice dell'errore quadratico medio settimanale, definita come:

$$\text{WRMSE} = \sqrt{\frac{1}{168} \sum_{h=1}^{168} (P_h - \bar{P}_h)^2}$$

Per avere un punto di riferimento nel testare l'efficacia di un certo metodo, si possono sottoporre le misure menzionate sopra al cosiddetto “*test naïve*”: in quest'ultimo le previsioni dei prezzi in ciascuna ora della giornata sono comparate a quelle di un giorno 'similare' a quello che si sta prendendo in considerazione. Un giorno 'similare' viene di solito costruito nel modo seguente: un lunedì è considerato simile al lunedì della settimana precedente; un martedì, un mercoledì, un giovedì ed un venerdì sono ritenuti simili al lunedì di quella medesima settimana; infine, un sabato o

una domenica sono pensati come simili rispettivamente al sabato o alla domenica precedente. Il *test naïve* viene superato nei casi in cui gli errori misurati utilizzando un qualunque test siano minori di quelli misurati per i prezzi dei giorni ‘similari’.

CAPITOLO TERZO

BREVE PRESENTAZIONE DEL MERCATO IPEX ED ANALISI DEI DATI

3.1 Il mercato elettrico italiano

L'avvio operativo del mercato elettrico, avvenuto il 31 marzo 2004 per effetto del decreto legislativo 16 marzo 1999, n. 79 (d.lgs. n.79/99), nell'ambito del processo di recepimento della direttiva comunitaria sulla creazione di un mercato interno dell'energia (96/92/CE), ha determinato la nascita del primo mercato organizzato dell'elettricità in Italia, analogamente alle esperienze già maturate in altri paesi industrializzati (conseguentemente all'avvento del liberalismo economico).

La creazione di un mercato elettrico risponde a due esigenze ben precise:

- stimolare la concorrenza nelle attività, potenzialmente competitive, di produzione e vendita di energia elettrica;
- favorire la massima efficienza nella gestione del dispacciamento² dell'energia elettrica, attraverso la creazione di un mercato per l'acquisto delle risorse per il servizio di dispacciamento.

² Il dispacciamento è quella fase della filiera produttiva che segue la generazione di energia elettrica. Esso consiste nella determinazione istante per istante dell'insieme di impianti di generazione che devono immettere energia nelle reti, per poi trasmetterla.

Il **Gestore del mercato elettrico** S.p.A. (GME) è la società a cui è affidata l'organizzazione e la gestione economica del mercato elettrico, secondo criteri di neutralità, trasparenza, obiettività e concorrenza tra produttori; esso assicura, inoltre, la gestione economica di un'adeguata disponibilità della riserva di potenza.

Questo mercato, comunemente indicato come “**borsa elettrica italiana**” (o IPEX = “Italian Power Exchange”), consente a produttori, consumatori e grossisti di stipulare contratti orari di acquisto e vendita di energia elettrica per il giorno successivo. Le transazioni si svolgono su una piattaforma telematica alla quale gli operatori si connettono attraverso la rete Internet, con procedure di accesso sicuro, tramite certificati digitali, per riuscire a concludere on-line contratti di acquisto e di vendita di energia elettrica. Il prezzo e la quantità di equilibrio di mercato si ottengono dall'incontro tra le domande e le offerte degli operatori che vi partecipano (Figura 3.1). Il mercato elettrico è anche un vero e proprio mercato fisico, dove si definiscono programmi di immissione e di prelievo sulla rete secondo il criterio di merito economico.

L'incontro tra domanda e offerta

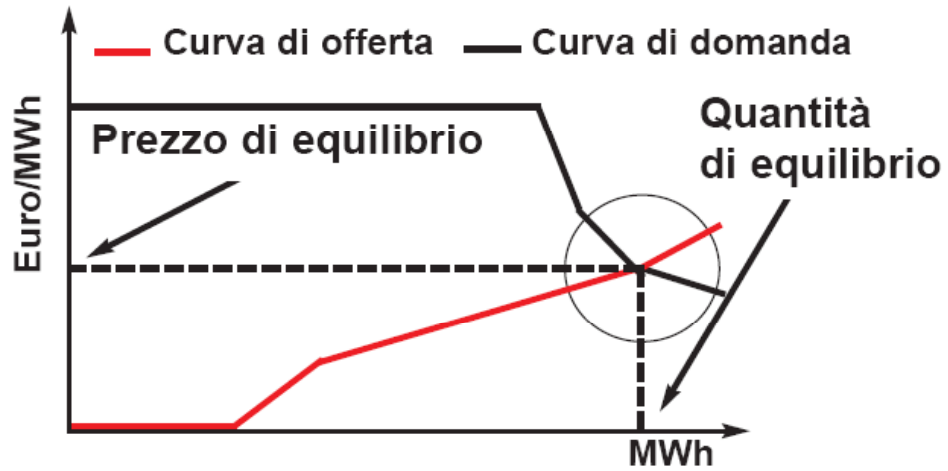


Figura 3.1: La formazione del prezzo di equilibrio

Il mercato elettrico si articola in molteplici mercati dell'energia:

a) **mercato del giorno prima** (MGP) dove i produttori, i grossisti ed i clienti finali idonei possono vendere o acquistare energia elettrica per il giorno successivo; tale mercato si chiude alle ore 9.00 del giorno precedente al giorno di consegna;

b) **mercato di aggiustamento** (MA) dove i produttori e i grossisti possono modificare i programmi di immissione determinati sul mercato del giorno prima presentando ulteriori offerte di vendita o di acquisto; tale mercato si apre alle 10.30, dopo la comunicazione degli esiti del MGP, e si chiude alle 14.00;

c) un **mercato per il servizio di dispacciamento** (MSD), sul quale gli operatori presentano offerte di disponibilità di aumento o di riduzione della potenza immessa o prelevata ad ogni ora; si apre alle 14.30, dopo la comunicazione degli esiti del MA e si chiude alle 16.00.

| BORSA ELETTRICA | | |
|--|----------------|---|
| Le sessioni della borsa elettrica | | |
| ↓ | entro le 8.00 | il GME diffonde le informazioni preliminari |
| ↓ | fino alle 9.00 | contrattazioni MGP |
| ↓ | 10.30-14.00 | contrattazioni MA |
| ↓ | 14.30-16.00 | contrattazioni MSD |
| ↓ | ore 21.00 | pubblicazioni esiti: prezzi e quantità di energia del giorno dopo |

3.2 Analisi empirica di un gruppo di dati appartenenti al mercato

IPEX

3.2.1. Presentazione dei dati

In questa parte utilizzeremo alcuni dei modelli citati nel capitolo precedente allo scopo di effettuare delle stime e delle previsioni sulla serie storica dei prezzi orari giornalieri del mercato elettrico italiano a partire dalla data della sua apertura, il 01 Aprile 2004, fino al 31 Marzo 2008, per un totale di 35064 osservazioni; abbiamo preso il Prezzo Unico Nazionale, espresso in euro per megawatt (Euro/MWh).

Dal momento che lavoriamo con serie orarie rilevate su più anni, dobbiamo trattare la questione del cambio dell'ora, da solare a legale e viceversa: nel primo caso abbiamo tra i nostri dati un giorno con 23 ore, e per portarle a 24 possiamo aggiungere l'ora del medesimo giorno della settimana precedente corrispondente a quella in cui avviene il cambio; nel secondo caso, al contrario, ci ritroviamo ad avere un giorno con 25 ore, e per ridurle a 24 possiamo calcolare la media tra l'ora in cui viene effettuato il cambio e l'ora 25. Solitamente in entrambi i casi il cambio dell'ora avviene alle 2.00.

Iniziamo il nostro lavoro con un'analisi grafica preliminare della serie in questione: la Figura 3.2 riporta il grafico della serie dei prezzi orari:

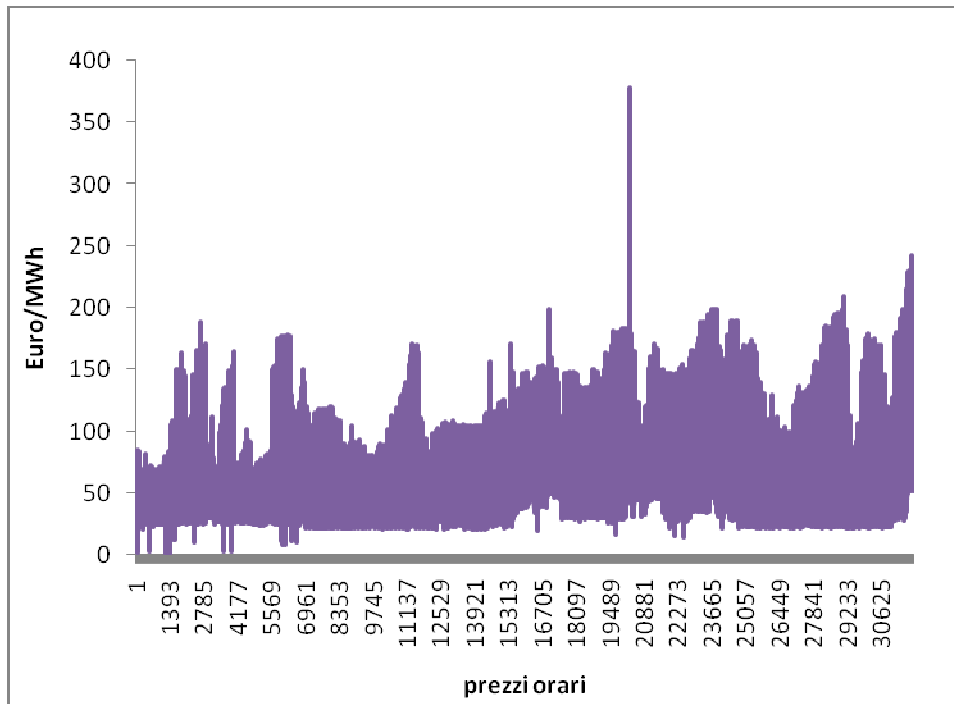


Figura 3.2: Prezzi orari del mercato elettrico IPEX nel periodo 1 aprile 2004 – 31 marzo 2008

Da questo grafico possiamo cogliere alcune caratteristiche peculiari delle serie storiche dei prezzi dell'elettricità: innanzitutto una variabilità non costante nel tempo (fenomeno conosciuto anche come *volatility clustering*), e la presenza di numerosi *jumps*, più verso l'alto che verso il basso, il più evidente dei quali si verifica in corrispondenza della fascia oraria 10-17 del giorno 26 Luglio 2006 (questo fatto è conforme alla teoria, la quale sostiene che le ore in cui si manifestano maggiormente gli *spikes* sono quelle di punta dei giorni lavorativi, solitamente comprese tra le 09.00 e le 18.00); la periodicità, invece, di qualunque tipo essa sia, è più difficile da

osservare, dal momento che le osservazioni , essendo in numero molto elevato, risultano vicinissime le une alle altre.

3.2.2. Analisi descrittive

Un modo interessante per evidenziare la periodicità, sia a livello intragiornaliero che a livello settimanale, è la suddivisione dei dati rilevati in quelli appartenenti ai giorni feriali da quelli appartenenti ai week-end; questi ultimi possono poi venire divisi ulteriormente in sabati e domeniche. La Figura 3.3 riporta il grafico delle medie orarie dei prezzi per i giorni lavorativi, per i sabati e per le domeniche del nostro campione:

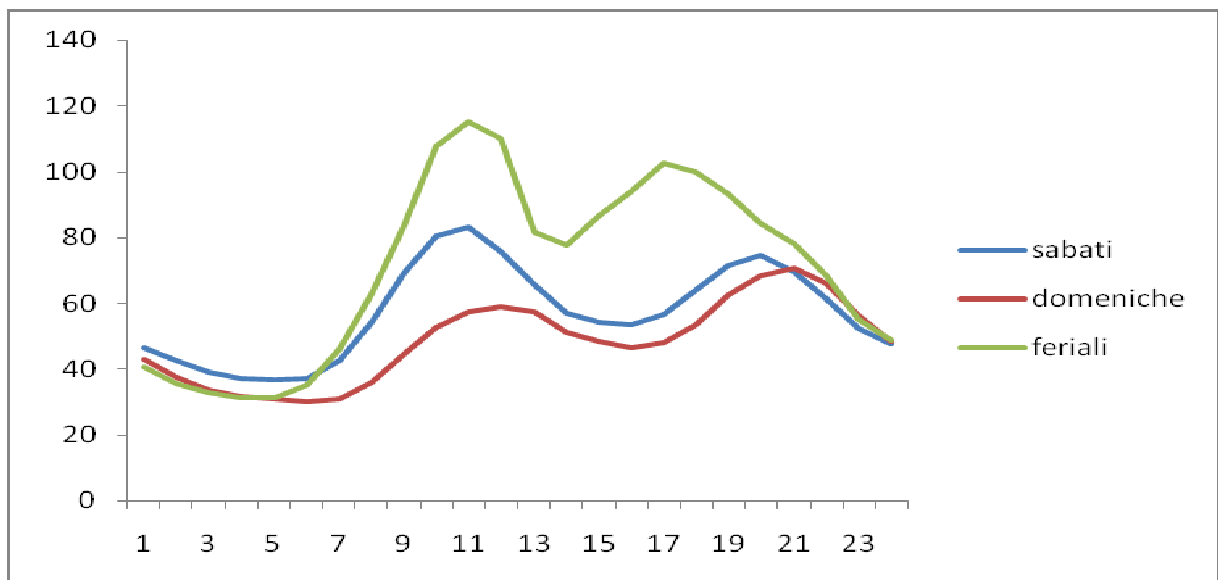


Figura3.3: Medie orarie dei prezzi nei diversi giorni

Per quanto riguarda la periodicità settimanale, come potevamo immaginare, notiamo che le medie nei giorni feriali risultano essere significativamente più elevate di quelle nel week-end; esse diminuiscono leggermente in corrispondenza dei sabati ed ancora di più in corrispondenza delle domeniche. Questo fenomeno è legato al fatto che durante il week-end (ed in particolare di domenica) molte attività sono chiuse, e quindi diminuisce il consumo di energia elettrica.

Per quanto riguarda invece la periodicità intragiornaliera, possiamo notare che al sabato ed alla domenica i prezzi presentano un andamento abbastanza simile nell'arco della giornata, mentre durante i giorni lavorativi subentrano dei cambiamenti.

Nelle primissime ore del giorno (1.00-4.00) i tre gruppi di dati presentano comportamenti simili, anche se le medie dei sabati e delle domeniche sono leggermente superiori a quelle dei giorni feriali: questo può essere legato al fatto che durante il fine settimana solitamente di sera si resta fuori di più rispetto agli altri giorni e si rientra più tardi. Verso le 5.00/6.00 del mattino si verifica una decisa impennata verso l'alto dei prezzi nei giorni feriali, mentre nel week-end le medie iniziano ad aumentare verso le 7.00/8.00: questo comportamento è legato al fatto che nei giorni lavorativi ci si alza più presto per recarsi al lavoro.

I prezzi continuano ad aumentare fino a raggiungere il primo picco della giornata, il quale si verifica in corrispondenza delle ore 11.00 per i giorni feriali e per i sabati, e, con un leggero slittamento, in corrispondenza delle ore 12.00 per le domeniche. Per i primi due gruppi questo è il picco più alto e più marcato della giornata. Verso le ore 13.00 i prezzi di tutti e tre i gruppi iniziano a scendere, per poi raggiungere un minimo locale intorno alle 14.00 nei giorni lavorativi ed intorno alle 15.00 nel fine settimana: probabilmente ciò è legato al momento della pausa dal lavoro nel primo caso e a momenti di relax o di uscita con la famiglia negli altri due casi. Dopodiché le medie ricominciano a salire fino ad arrivare al secondo picco della giornata, che si riscontra intorno alle 17.00 nei giorni feriali, intorno alle 19.00 nei sabati, ed intorno alle 21.00 nelle domeniche: tale comportamento può essere ricondotto al momento del rientro dal lavoro nel primo gruppo, ed al momento del rientro a casa e di un'eventuale preparazione per un'uscita serale negli altri due gruppi. Questo è il picco più alto della giornata per le medie domenicali. Dopo di esso, le medie di tutti i gruppi iniziano a diminuire costantemente fino alle ore 24.00.

Per comprendere ancora meglio le caratteristiche dei prezzi dell'elettricità ed il modo in cui esse cambiano nelle diverse ore della giornata, abbiamo calcolato le principali statistiche descrittive ed il test di Jarque-Bera, con il

relativo *p-value*, della serie storica intera e delle 24 serie relative ad ogni ora del giorno, ed abbiamo riassunto il tutto nella tabella 1. Abbiamo preso in considerazione gli indici principali di posizione, la media e la mediana, di variabilità, la deviazione standard, il valore massimo e quello minimo, e di forma, indici di asimmetria e di curtosi. L'indice di asimmetria, altresì noto come *skewness*, serve a stabilire se la distribuzione è simmetrica intorno allo 0; se tale indice risulta maggiore di 0 si parla di asimmetria positiva (come già accennato nel Capitolo Primo, presentando le caratteristiche dei prezzi) e la distribuzione appare con una coda più lunga verso destra; se al contrario il valore ottenuto è minore di 0, si parla di asimmetria negativa, e la distribuzione presenta una coda più lunga verso sinistra. L'indice di curtosi (anch'esso già introdotto parlando delle caratteristiche dei prezzi) determina invece la pesantezza delle code della distribuzione che si sta analizzando, ed è un ottimo indicatore di normalità: infatti in caso di dati generati dalla distribuzione normale tale indice risulta essere pari a 3, mentre se è maggiore di 3 si parla di distribuzione *leptocurtica* – più appuntita e con code più pesanti di quelle della normale-, infine se è minore di 3 si parla di distribuzione *platicurtica* –più appiattita e con code meno pesanti di quelle della normale -.

Un altro modo efficace che sfrutta congiuntamente gli indici di asimmetria e di curtosi al fine di valutare la normalità/gaussianità del processo

generatore dei dati è il test di Jarque-Bera. La statistica test corrispondente ad esso è riportata nella penultima colonna della Tabella 3.1, ed è calcolata nel modo seguente:

$$JB = \frac{T-1}{6} (S^2 + \frac{1}{4} (K-3)) \sim \chi_2^2 \quad \text{sotto } H_0$$

dove S rappresenta l'indice di *skewness*, K l'indice di curtosi e T rappresenta il numero di osservazioni del campione che si sta analizzando.

| ora | media | mediana | sd | max | min | asimm | curt | jb | p-value |
|--------------|--------|---------|-------|--------|------|-------|-------|----------|----------|
| Ora1 | 41,79 | 40,26 | 10,60 | 73,12 | 1,12 | 0,25 | 2,72 | 19,89 | 0,000048 |
| Ora2 | 36,76 | 35,02 | 10,32 | 71,95 | 1,17 | 0,44 | 2,57 | 59,03 | 0,000000 |
| Ora3 | 33,65 | 30,55 | 9,82 | 65,98 | 1,17 | 0,64 | 2,69 | 105,86 | 0,000000 |
| Ora4 | 32,14 | 29,94 | 9,13 | 63,6 | 1,2 | 0,79 | 3,15 | 153,59 | 0,000000 |
| Ora5 | 31,95 | 29,88 | 9,001 | 62,6 | 1,21 | 0,79 | 3,22 | 156,08 | 0,000000 |
| Ora6 | 34,68 | 32,29 | 9,86 | 68,05 | 1,25 | 0,55 | 2,67 | 79,66 | 0,000000 |
| Ora7 | 43,59 | 44,79 | 12,81 | 79,97 | 1,27 | 0,11 | 2,72 | 7,73 | 0,020907 |
| Ora8 | 57,89 | 57,63 | 18,05 | 137,6 | 1,24 | 0,24 | 3,65 | 40,48 | 0,000000 |
| Ora9 | 75,91 | 76,95 | 24,02 | 176,94 | 1,18 | 0,22 | 3,75 | 46,29 | 0,000000 |
| Ora10 | 95,93 | 95,11 | 33,89 | 286,3 | 1,14 | 0,28 | 3,19 | 21,18 | 0,000025 |
| Ora11 | 102,37 | 100,22 | 37,45 | 288,51 | 1,13 | 0,34 | 2,79 | 30,55 | 0,000000 |
| Ora12 | 97,84 | 93,95 | 36,58 | 378,47 | 1,12 | 0,75 | 4,89 | 354,46 | 0,000000 |
| Ora13 | 75,97 | 75,2 | 21,93 | 248,63 | 1,13 | 1,22 | 8,38 | 2125,66 | 0,000000 |
| Ora14 | 70,98 | 70,36 | 21,23 | 288,31 | 1,13 | 1,38 | 12,75 | 6263,19 | 0,000000 |
| Ora15 | 76,65 | 75,22 | 27,64 | 288,63 | 1,15 | 1,13 | 7,18 | 1373,67 | 0,000000 |
| Ora16 | 81,57 | 79,7 | 32,59 | 288,75 | 1,17 | 0,92 | 4,92 | 432,74 | 0,000000 |
| Ora17 | 88,34 | 84,5 | 37,69 | 287,99 | 1,18 | 0,78 | 3,61 | 172,50 | 0,000000 |
| Ora18 | 88,32 | 79,68 | 39,93 | 239,52 | 1,18 | 1,12 | 4,05 | 37,61 | 0,000000 |
| Ora19 | 85,88 | 74,97 | 37,71 | 242,42 | 1,14 | 1,16 | 4,13 | 404,72 | 0,000000 |
| Ora20 | 80,67 | 75,6 | 27,43 | 190,01 | 1,11 | 0,69 | 3,49 | 131,11 | 0,000000 |
| Ora21 | 75,87 | 75,23 | 21,21 | 162,63 | 1,1 | 0,49 | 4,12 | 134,67 | 0,000000 |
| Ora22 | 66,91 | 67,66 | 15,73 | 145,56 | 2,26 | 0,13 | 4,10 | 78,483 | 0,000000 |
| Ora23 | 54,72 | 53,63 | 12,70 | 99,51 | 1,13 | 0,15 | 3,54 | 23,74 | 0,000007 |
| Ora24 | 48,55 | 49,39 | 10,67 | 81,69 | 1,17 | -0,11 | 3,20 | 5,53 | 0,062811 |
| Serie intera | 65,79 | 59,79 | 33,37 | 378,47 | 1,1 | 1,19 | 5,01 | 14155,77 | 0,000000 |

Tabella 3.1: Statistiche descrittive, indicatori di sintesi e test di Jarque-Bera con relativo p-value.

Come potevamo immaginare, la media più alta tra tutte le ore si registra in corrispondenza dell'ora 11, ora in cui si verifica il primo picco della giornata.

Da notare anche il range particolarmente elevato della serie intera, nella quale il prezzo minimo (che si manifesta in corrispondenza della 21° ora) arriva a toccare 1,1 €/MWh , mentre il prezzo massimo (che si registra nella 12° ora) raggiunge addirittura gli 378,47 €/MWh . Questo conferma l'esistenza di una volatilità molto elevata .

Conformemente a quanto sostenuto dalla teoria, in quasi tutte le serie – eccetto in quella relativa all'ora 24 – l'indice di asimmetria è positivo, e predomina la leptocurtosi –tranne che nelle serie relative ad alcune tra le prime ore della giornata, più precisamente le ore 1, 2, 3 e l'ora 6, l'ora 7 e l'ora 11-. Inoltre, sempre conformemente alle caratteristiche peculiari delle serie dei prezzi dell'elettricità, l'ipotesi di normalità del test di Jarque-Bera viene abbondantemente rifiutata ad un livello di significatività del 5% in tutti i casi, tranne che nella serie relativa all'ora 24: questo significa che, intuitivamente, per la serie riguardante tale ora possiamo considerare modelli basati sull'assunzione di normalità per rappresentare il processo generatore dei dati. Tutte le caratteristiche osservate sono evidenziate nella Tabella 3.1 in colore rosso.

3.2.3. Stima dei modelli

In questa sezione presentiamo i risultati relativi all'applicazione di alcuni modelli atti a stimare (e successivamente prevedere) i prezzi dell'elettricità.

Nel nostro lavoro ci limitiamo ad utilizzare modelli lineari del tipo SARIMA-GARCH per la serie intera dei prezzi del mercato IPEX e per le serie riguardanti le singole ore del giorno.

L'approccio di modellare separatamente ora per ora è sorto grazie a studi di J.C.Cuaresma, J. Hlouskova, S. Kossmeier e M. Obersteiner, i quali in un lavoro del 2004 hanno utilizzato per primi questa tecnica con semplici modelli lineari univariati (vale a dire "relativi ad un solo fenomeno") allo scopo di modellare, e successivamente prevedere, i prezzi *spot* del mercato *Leipzig Power Exchange* (LPX): tali studiosi sono arrivati alla conclusione che modellare questi prezzi considerando serie separate per ciascuna ora del giorno incrementasse le *performance* previsive, dal momento che gli errori misurati in questo caso risultavano minori di quelli ottenuti considerando la serie intera.

In un articolo del 2007 R. Huisman, C. Hurman e R. Mahieu hanno fornito una spiegazione più approfondita della valenza dell'uso del metodo in questione: essi hanno affermato che modellare separatamente le ore della

giornata era logico poiché le diverse ore presentano caratteristiche diverse le une dalle altre; i prezzi in ore con abbondanza di *spikes* sembrano essere molto correlati al loro interno, e lo stesso vale per i prezzi in ore con meno o prive di *spikes*; sembra che esista invece una correlazione molto minore *tra* questi due gruppi.

Questo fatto può essere spiegato dalle differenze nella capacità di riserva tra questi due blocchi: questa capacità è più bassa nelle ore ricche di *jumps*, e ciò implica che i prezzi in tali ore mostrino una volatilità più elevata.

Ora verificheremo se la tecnica di considerare 24 serie separate per le varie ore è efficace anche per i prezzi *spot* del mercato IPEX.³

Dal momento che le serie dei prezzi da noi analizzate non erano stazionarie né in media né in varianza, abbiamo stimato i modelli considerando i rendimenti logaritmici, vale a dire $r_t = \log(P_t) - \log(P_{t-1})$, dove P_t e P_{t-1} indicano i prezzi rispettivamente al tempo t ed al tempo $t-1$. A partire da tali rendimenti, abbiamo stimato i modelli seguenti:

³ I grafici delle 24 serie dei prezzi sono presenti nell'Appendice 1: in essi si riscontrano caratteristiche simili a quelle osservate nell'analisi dell'intera serie oraria, effettuata all'inizio di questo capitolo. In particolar modo, i *jumps* notati nella serie intera in corrispondenza della fascia oraria 10-17 del 26/07/2006 ricompaiono anche nelle singole serie corrispondenti a queste ore.

| SERIE | MODELLO |
|---------|--|
| serie1 | SARIMA(1,0,1)(1,0,2) ₇ +d1+d2+d3+d4 |
| serie2 | SARIMA(1,0,1)(2,0,1) ₇ +d1+d2+d3+d4 |
| serie3 | SARIMA(1,0,1)(2,0,1) ₇ +d1+d2+d3+d4 |
| serie4 | SARIMA(1,0,1)(2,0,1) ₇ +d1+d2+d3+d4 |
| serie5 | SARIMA(1,0,1)(2,0,1) ₇ + GARCH(1,1)+d1+d2+d3+d4 |
| serie6 | SARIMA(1,0,1)(1,0,1) ₇ + GARCH(1,1)+d1+d2+d3+d4 |
| serie7 | SARIMA(1,0,1)(2,0,1) ₇ +d1+d2+d3+d4 |
| serie8 | SARIMA(1,0,1)(2,0,1) ₇ |
| serie9 | SARIMA(1,0,1)(2,0,1) ₇ |
| serie10 | SARIMA(2,0,1)(2,0,1) ₇ |
| serie11 | SARIMA(0,0,2)(2,0,1) ₇ |
| serie12 | SARIMA(0,0,2)(2,0,1) ₇ |
| serie13 | SARIMA(0,0,2)(1,0,3) ₇ |
| serie14 | SARIMA(0,0,2)(2,0,1) ₇ |
| serie15 | SARIMA(0,0,2)(2,0,1) ₇ |
| serie16 | SARIMA(0,0,2)(1,0,2) ₇ |
| serie17 | SARIMA(0,0,2)(1,0,2) ₇ |
| serie18 | SARIMA(3,0,1)(2,0,1) ₇ |
| serie19 | SARIMA(1,0,1)(2,0,1) ₇ |
| serie20 | SARIMA(3,0,1)(2,0,1) ₇ |
| serie21 | SARIMA(0,0,2)(1,0,2) ₇ |
| serie22 | SARIMA(3,0,1)(1,0,1) ₇ |
| serie23 | SARIMA(3,0,1)(1,0,0) ₇ |
| serie24 | SARIMA(1,0,1)(1,0,0) ₇ +d1+d2+d3+d4 |

| | |
|-----------------|---|
| Serie intera | $(1-B^{168})(1-\phi_1 B-\phi_2 B^{168})(1-\Phi B^{24}) Y_t = (1-\theta_1 B-\theta_2 B^6) \varepsilon_t +$ GARCH(1,1) |
|-----------------|---|

Tabella 3.2: Modelli utilizzati

Le stime dei parametri sono state ottenute tramite la procedura di massima verosimiglianza.

I tipi di periodicità che riscontriamo in questi modelli sono quella settimanale – con $S=7$ nei modelli stimati per le 24 serie (giornaliere) corrispondenti alle diverse ore e con $S=168$ nel modello riguardante la serie

intera – e quella giornaliera – presente nel solo modello utilizzato per stimare la serie intera con $S=24$ -.

Dal lavoro di stima e dalla Tabella 3.2, che sintetizza i modelli utilizzati, è possibile fare alcune considerazioni:

- la modellazione dell'eteroschedasticità tramite i modelli GARCH è efficace solamente in 3 casi;
- per quanto riguarda la parte ARMA non stagionale, notiamo che nelle prime ore della giornata sembrano essere significative in egual misura sia la componente autoregressiva (AR) sia la componente a media mobile (MA); diversamente, nella tarda mattinata e nelle ore centrali del pomeriggio prevale in genere una componente MA che arriva fino al ritardo 2; infine, nelle ultime ore (a partire dalle 18) sembra prevalere nella maggior parte dei casi la parte AR, la quale arriva fino al ritardo 3;
- per quanto riguarda la parte ARMA stagionale, c'è in genere una prevalenza della componente SAR (componente autoregressiva stagionale); l'ordine massimo raggiunto è il secondo, eccetto nel caso della 13^o ora, in cui compare una componente SMA (componente a media mobile stagionale) di ordine 3;
- dal momento che, osservando i grafici dei rendimenti delle singole serie (presenti nell'Appendice 1), notiamo in tutte la presenza di

spikes particolarmente elevati in corrispondenza delle date 23,24,30 e 31 Maggio 2004, abbiamo provato ad introdurre nei modelli 4 variabili *dummy*, ciascuna delle quali assumeva valore 1 solo in corrispondenza del rispettivo *outlier*; l'aggiunta di tali *dummies* ha migliorato i risultati finali solamente in alcuni casi – precisamente nelle prime 7 ore della giornata e nella 24° ora -. Queste 4 variabili sono indicate nella tabella con d1, d2, d3 e d4. Probabilmente la presenza di questi *spikes* è dovuta al fatto che il periodo in cui essi compaiono è quello di apertura del mercato IPEX, quindi un periodo ancora “di aggiustamento” della situazione;

- per quanto riguarda la modellazione della serie intera dei rendimenti – indicata nella tabella con Y -, abbiamo provato ad adattare diversi modelli (introducendo anche in questo caso delle variabili *dummy* in corrispondenza degli *outlier* più evidenti), non riuscendo però ad ottenere risultati soddisfacenti: abbiamo così deciso di adattare il modello della classe SARIMA-GARCH che ci sembrava maggiormente rappresentativo dei dati e che riusciva al contempo ad essere più parsimonioso in termini di parametri.

Per modello “buono” intendiamo un modello i cui residui presentino *autocorrelazione nulla* o quasi tale. La definizione di “funzione di autocorrelazione” è la seguente:

$$\rho_{t_1,t_2} = \frac{\gamma_{t_1,t_2}}{\sigma_{t_1}\sigma_{t_2}}$$

dove γ_{t_1,t_2} rappresenta la “funzione di autocovarianza” (vale a dire la covarianza tra variabili casuali di un medesimo processo stocastico spaziate tra di loro di uno *sfasamento temporale* pari a $k = |t_2 - t_1|$), mentre σ_{t_1} e σ_{t_2} sono gli scarti quadratici medi rispettivamente nei momenti t_1 e t_2 . Il grafico della funzione di autocorrelazione, rappresentante i valori ρ_k , con $k = 0,1,2,\dots$, è noto con il nome di “*correlogramma*”.

Per assicurarsi del fatto che i residui siano incorrelati tra di loro, bisogna controllare che nel correlogramma essi rientrino tutti nelle bande di confidenza al 5% (corrispondenti all'intervallo $[-1,96/\sqrt{n}, 1,96/\sqrt{n}]$), le quali nel programma *EViews* - che stiamo utilizzando per la nostra analisi empirica - compaiono automaticamente nel grafico; inoltre deve essere accettata l'ipotesi nulla del test di Ljung-Box⁴ e del test di Breusch-

⁴ La statistica test del test di Ljung-Box è: $Q(k) = T(T+2) \sum_{i=1}^k \frac{\hat{\rho}_i^2}{T-i}$ ($\sim \chi_k^2$ sotto H_0), dove ρ_i rappresenta

la funzione di autocorrelazione al tempo i (noi la stimiamo), mentre T sono le osservazioni totali presenti nel campione.

Godfrey⁵, tramite i quali si vuole verificare proprio l'assenza di correlazione seriale.

Oltre ai correlogrammi dei residui, poi, si possono osservare anche i correlogrammi dei loro quadrati: la presenza di correlazione tra di essi può essere un sintomo di eteroschedasticità. L'eventuale esistenza di quest'ultima può comunque essere poi appurata anche attraverso molteplici test, come il test di White⁶, il test per effetti ARCH⁷, ecc.

Il correlogramma è solitamente affiancato al grafico della “funzione di autocorrelazione parziale”, che misura l'autocorrelazione tra due variabili distanziate di k ritardi Y_t e Y_{t-k} al netto delle variabili intermedie; perché un modello sia valido, anche le autocorrelazioni parziali devono rientrare nelle bande di confidenza al 5%.

Nella nostra procedura di stima notiamo che, utilizzando modelli SARIMA con ordini elevati e/o la modellistica ARCH/GARCH, non riusciamo

⁵ La statistica test in questo caso è invece basata sull' R^2 di una regressione ausiliaria avente come variabile dipendente i residui al tempo t e come variabili esplicative le componenti SARIMA-GARCH individuate nel modello preso in considerazione ed i residui ritardati; sotto H_0 tale statistica test si distribuisce asintoticamente come un χ^2 con gradi di libertà pari al numero dei ritardi considerati per i residui nella regressione.

⁶ Questo test utilizza l' R^2 della regressione dei quadrati dei residui su una costante e su tutti i momenti del primo ordine, del secondo ordine ed i momenti incrociati dei regressori originari; la statistica test ha sotto H_0 distribuzione asintotica χ^2 con gradi di libertà pari al numero dei regressori nella regressione ausiliaria esclusa l'intercetta.

⁷ Il test per verificare la presenza di eventuali effetti ARCH prende come statistica test l' R^2 di una regressione ausiliaria dei quadrati dei residui al tempo t sui quadrati dei residui ritardati, e sotto l'ipotesi nulla ha distribuzione asintotica χ^2 con gradi di libertà pari al numero dei ritardi considerati per i residui nella regressione.

comunque ad ottenere residui e residui al quadrato del tutto incorrelati in nessuna serie; inoltre molte volte siamo indotti a rifiutare le ipotesi di incorrelazione seriale e di omoschedasticità dei residui anche dai risultati dei test citati sopra. Per ottenere stime più consistenti in caso di autocorrelazione e di eteroschedasticità dei residui si possono utilizzare gli *standard error di Newey-West*, ed è proprio quello che abbiamo fatto noi.

Supponiamo comunque che per migliorare i risultati sia necessario prendere in considerazione anche modelli più complicati, per esempio del tipo ARMAX, o TAR/TARX, o *Markov-switching*, o ancora *jump-diffusion*, al fine di modellare più efficacemente anche gli *spikes*.

3.2.4. Previsioni

Nella parte finale del nostro lavoro riportiamo i risultati ottenuti dalle previsioni effettuate sulle singole serie stimate in precedenza. Queste previsioni sono di tipo *ex-post*, cioè teniamo a disposizione un certo numero di osservazioni del campione in esame per il periodo di previsione, in modo che i valori previsti dai modelli possono poi venire direttamente confrontati con i valori reali.

Procediamo quindi dividendo l'insieme dei nostri dati in due sottoinsiemi: il primo, da 1 a T, utilizzato per la costruzione dei modelli; il secondo, da T

+ 1 a T^* , utilizzato per il confronto tra i valori previsti e quelli veri. Nel nostro caso poniamo T uguale a 34320, $T + 1$ naturalmente pari a 34321 e T^* pari al totale delle osservazioni (35064); in termini temporali, ciò significa che stimiamo i parametri delle serie fino al 29/02/2008 per poi prevedere l'ultimo mese, effettuando così 744 previsioni in totale.

Vogliamo verificare, se, lavorando con i nostri dati, arriviamo alla medesima conclusione a cui sono arrivati Cuaresma *et al.*, cioè se la previsione effettuata con serie separate sia migliore di quella effettuata con la serie intera.

Le misure utilizzate per verificare la bontà delle nostre previsioni sono:

- a) l'errore medio assoluto giornaliero (MAE);
- b) l'errore medio giornaliero in percentuale assoluta (MAPE);
- c) la radice dell'errore quadratico medio giornaliero (RMSE).

Per quanto riguarda le serie separate, per trovare questi indici abbiamo calcolato la media degli indici riguardanti le singole ore del giorno.

I risultati che abbiamo ottenuto sono:

| indice | serie intera | serie separate |
|--------|--------------|----------------|
| MAE | 11,90411 | 11,06781 |
| MAPE | 16,31968 | 15,78541 |
| RMSE | 16,38177 | 13,93473 |

Tabella 3.3: Indici di bontà delle previsioni

Di seguito sono invece riportati dei grafici con il confronto delle previsioni tra i due metodi:

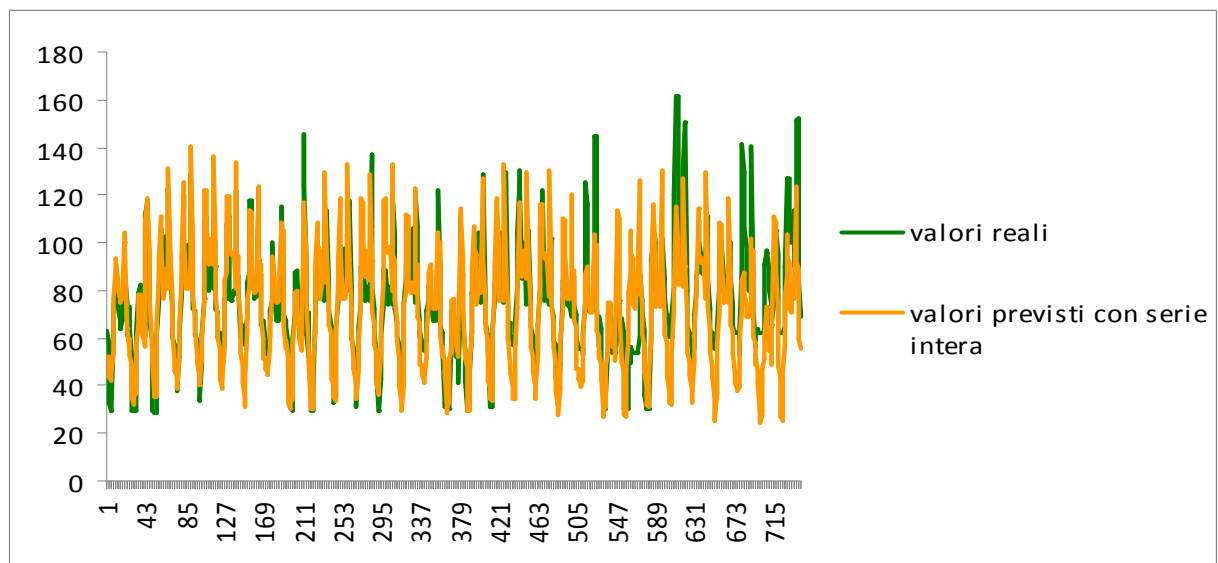


Figura 3.4

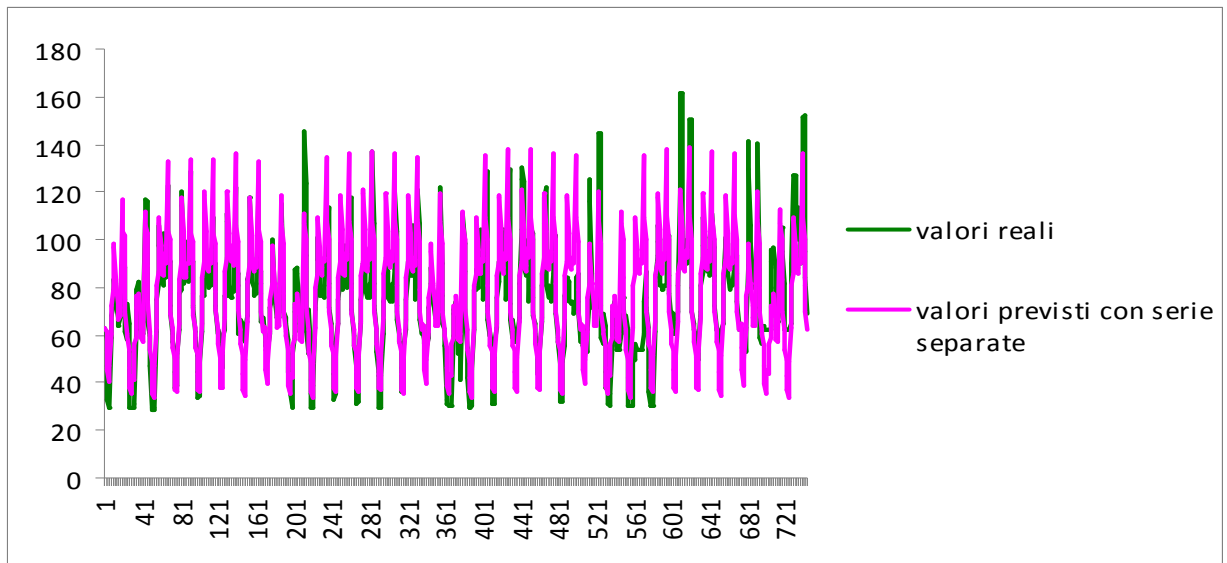


Figura 3.5

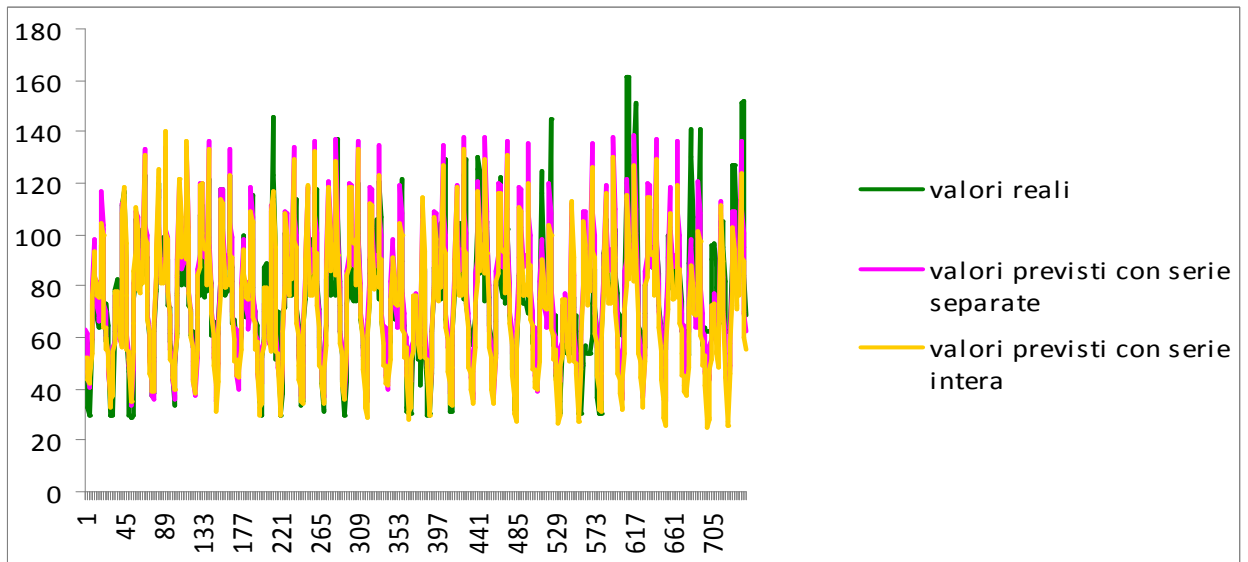


Figura 3.6

Figure 3.4,3.5,3.6: Grafici delle previsioni

Come si può notare osservando i risultati riportati nella Tabella 3.3 e nei tre grafici, i risultati finali ottenuti con le due tecniche di modellazione non sono significativamente differenti, però sembra che l'idea di separare le

diverse ore del giorno conduca a *performance* previsive leggermente migliori (infatti tutti e tre i tipi di errore calcolati risultano strettamente minori di quelli ottenuti considerando la serie intera).

Osservando le Figure 3.4,3.5 e 3.6, constatiamo in maniera piuttosto evidente che, come immaginavamo, i modelli lineari da noi utilizzati non riescono a comprendere i *jumps* più elevati presenti all'interno della serie.

CONCLUSIONI

Lo scopo principale di questa tesi è stato quello di passare in rassegna le caratteristiche salienti e vari modelli utili per la stima e la previsione dei prezzi *spot* nei mercati elettrici; la peculiarità di alcune di tali caratteristiche e l'importanza di effettuare delle buone previsioni – specialmente a breve termine – sono conseguenze del processo di riforma liberale che è avvenuto negli ultimi due decenni in numerosi Paesi in tutto il mondo (tra cui anche l'Italia), e che ha innescato un nuovo modo di determinazione del prezzo *spot*.

Nell'ultimo capitolo alcuni dei modelli citati in precedenza – precisamente i modelli più semplici, del tipo SARIMA-GARCH – sono stati utilizzati in un'analisi empirica riguardante l'intera serie oraria dei prezzi della borsa elettrica italiana al fine di stimare e soprattutto prevedere il più accuratamente possibile questi ultimi. Sulla base di un lavoro simile effettuato da Cuaresma *et al.* nel 2004 la serie oraria dei prezzi è stata suddivisa in 24 serie giornaliere, ognuna relativa ad ogni ora del giorno; è stato quindi identificato tramite il pacchetto statistico *Eviews* un modello

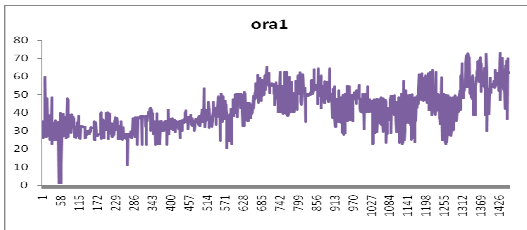
della classe SARIMA-GARCH (in alcuni casi arricchito con delle variabili *dummy* in corrispondenza di certi *outliers*) sia per ogni singola serie che per la serie completa, ed infine sono state confrontate le previsioni ottenute con le due differenti procedure. I risultati indicano che la strategia di modellare separatamente le diverse ore della giornata aumenta effettivamente la bontà delle previsioni, confermando perciò la tesi di Cuaresma *et al.*

Siamo comunque consapevoli del fatto che i modelli utilizzati in questa analisi empirica non tengono in considerazione molte caratteristiche dei prezzi dell'elettricità, in particolar modo quelle di non linearità, per stimare le quali probabilmente sarebbe necessario utilizzare *software* più elaborati.

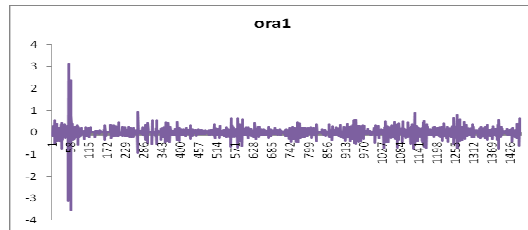
APPENDICE 1

I GRAFICI DEI PREZZI E DEI RENDIMENTI PER LE SINGOLE ORE

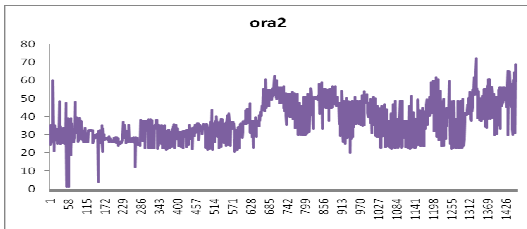
Prezzi ora 1



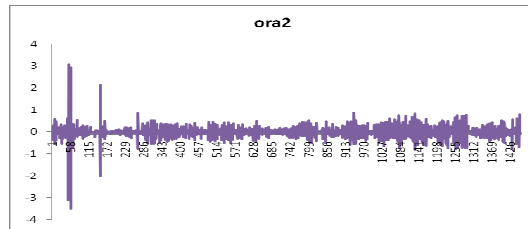
Rendimenti ora 1



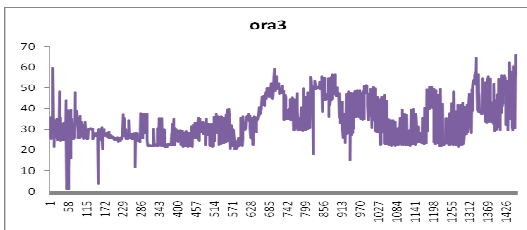
Prezzi ora 2



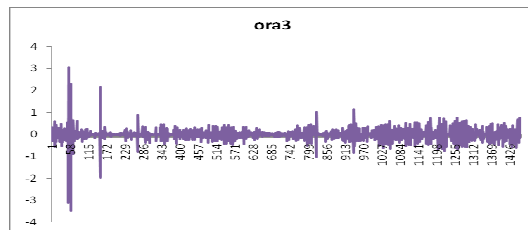
Rendimenti ora 2



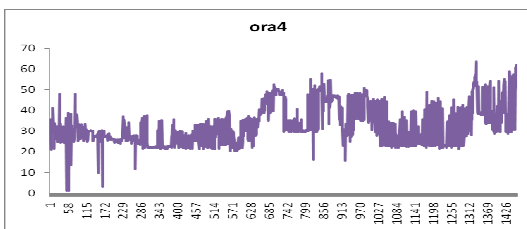
Prezzi ora 3



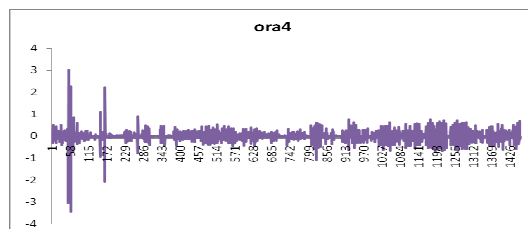
Rendimenti ora 3



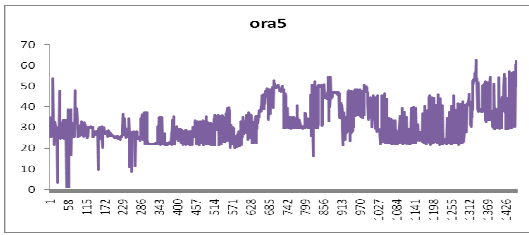
Prezzi ora 4



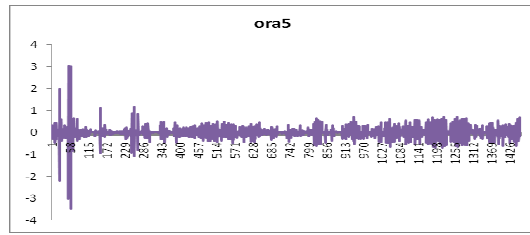
Rendimenti ora 4



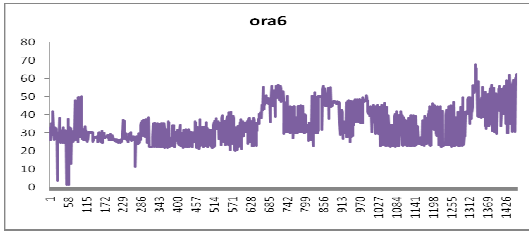
Prezzi ora 5



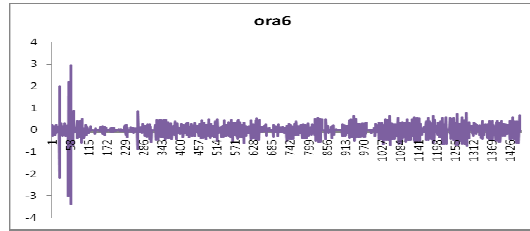
Rendimenti ora 5



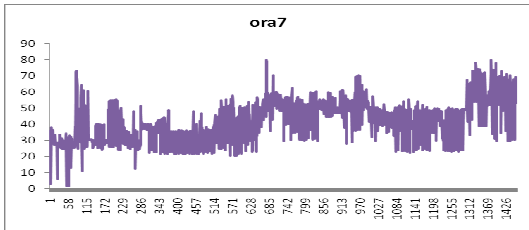
Prezzi ora 6



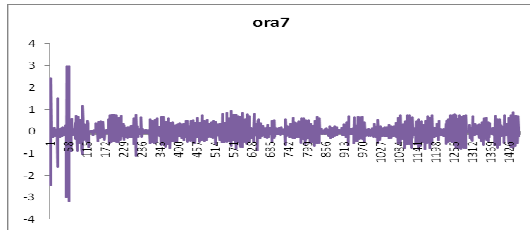
Rendimenti ora 6



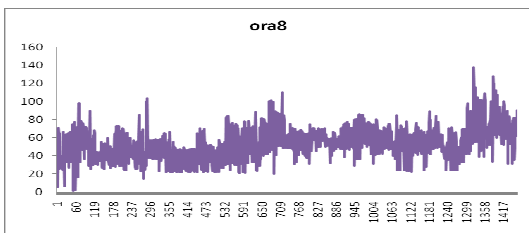
Prezzi ora 7



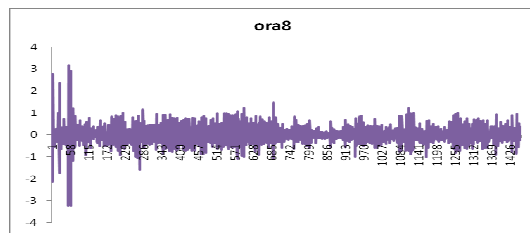
Rendimenti ora 7



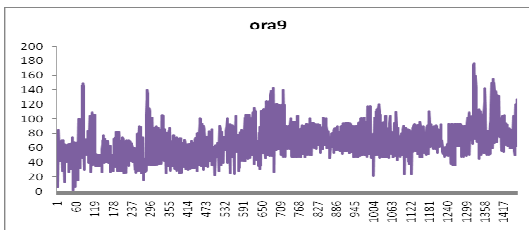
Prezzi ora 8



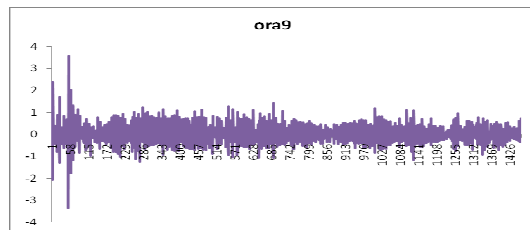
Rendimenti ora 8



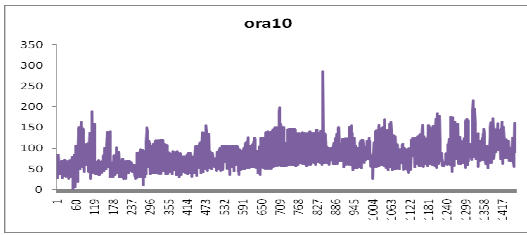
Prezzi ora 9



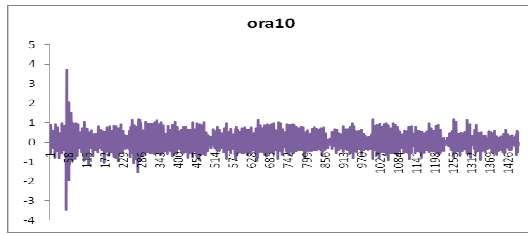
Rendimenti ora 9



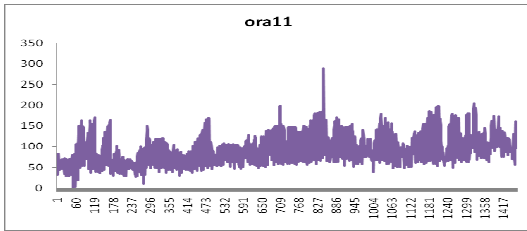
Prezzi ora 10



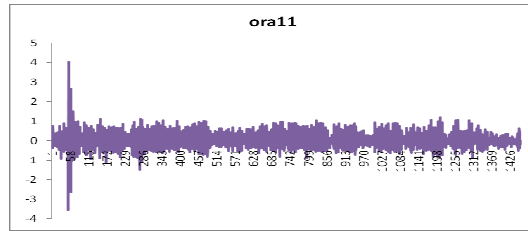
Rendimenti ora 10



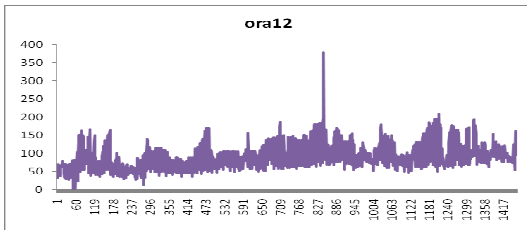
Prezzi ora 11



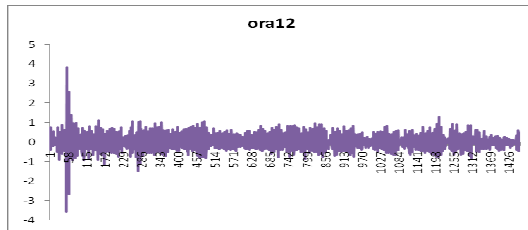
Rendimenti ora 11



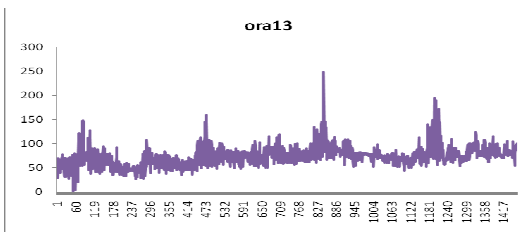
Prezzi ora 12



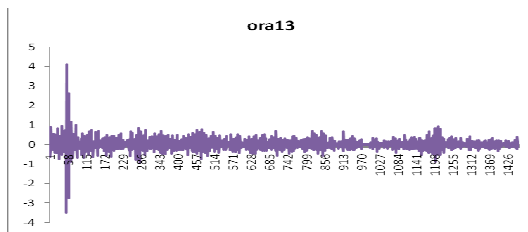
Rendimenti ora 12



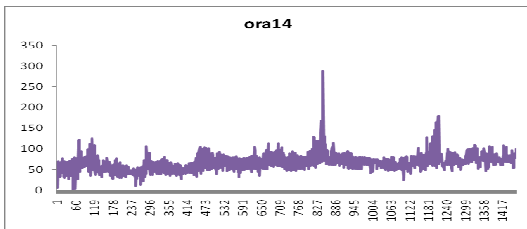
Prezzi ora 13



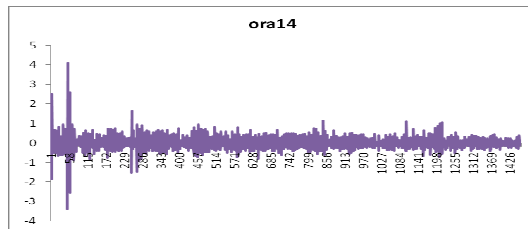
Rendimenti ora 13



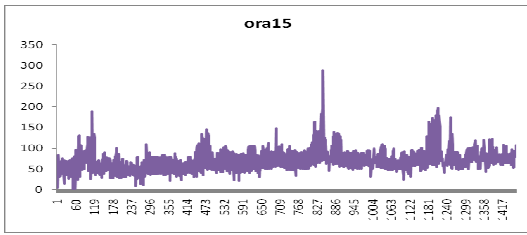
Prezzi ora 14



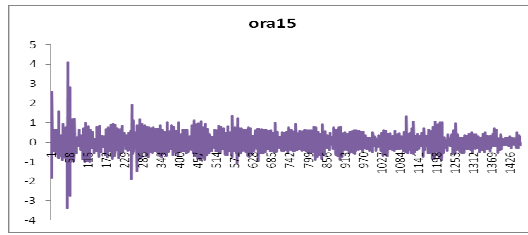
Rendimenti ora 14



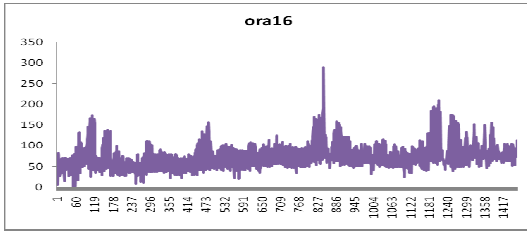
Prezzi ora 15



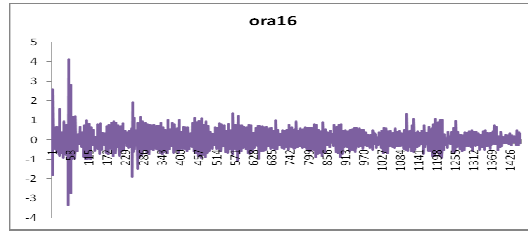
Rendimenti ora 15



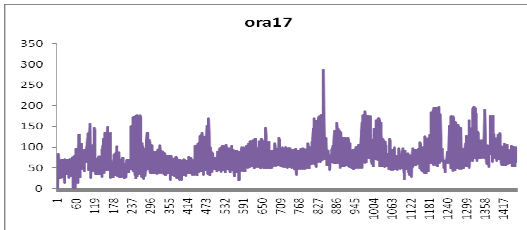
Prezzi ora 16



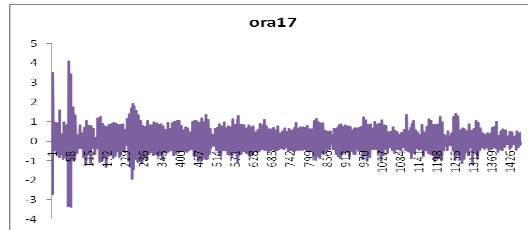
Rendimenti ora 16



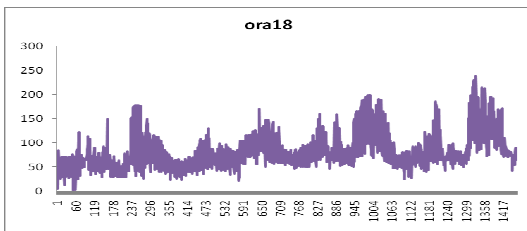
Prezzi ora 17



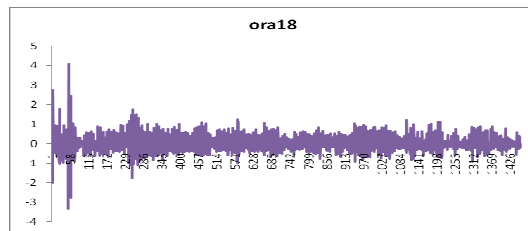
Rendimenti ora 17



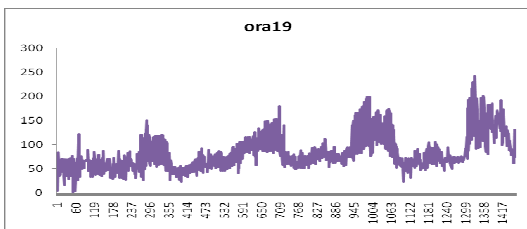
Prezzi ora 18



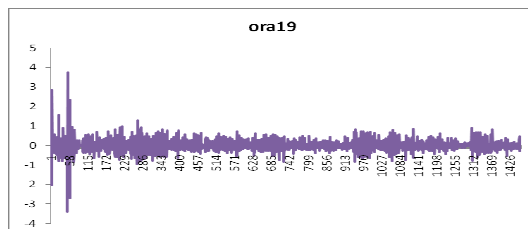
Rendimenti ora 18



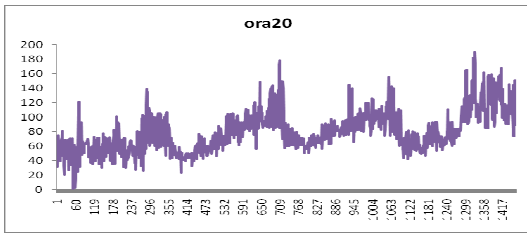
Prezzi ora 19



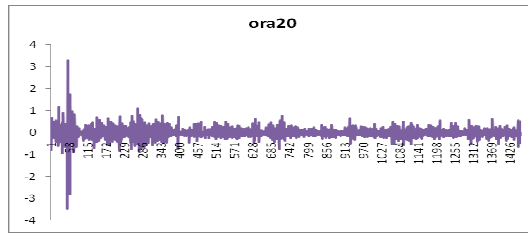
Rendimenti ora 19



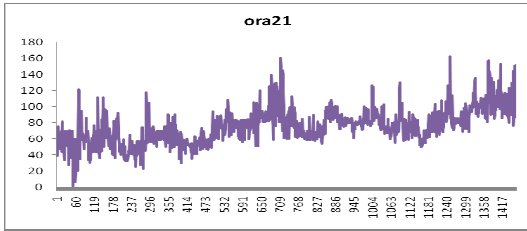
Prezzi ora 20



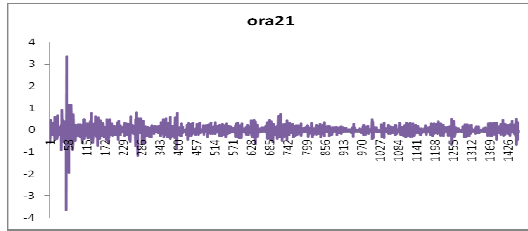
Rendimenti ora 20



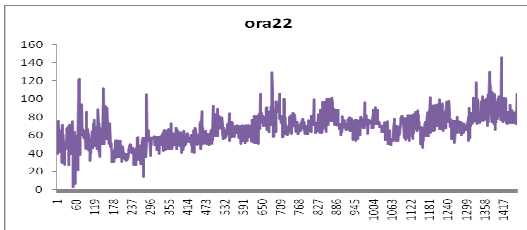
Prezzi ora 21



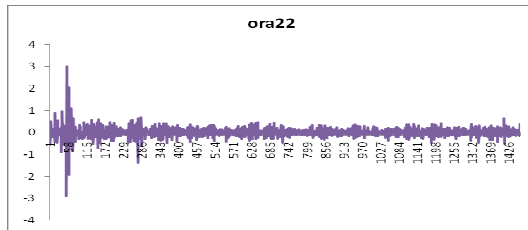
Rendimenti ora 21



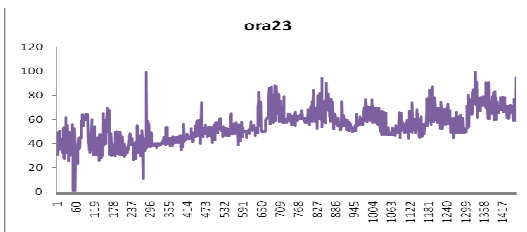
Prezzi ora 22



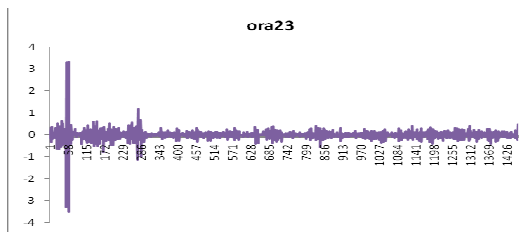
Rendimenti ora 22



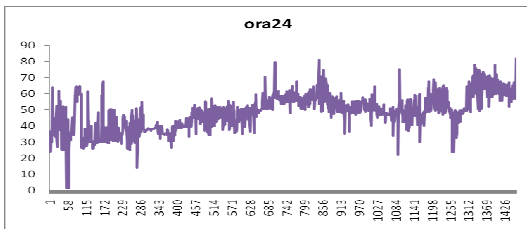
Prezzi ora 23



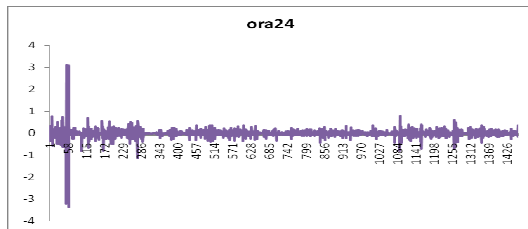
Rendimenti ora 23



Prezzi ora 24



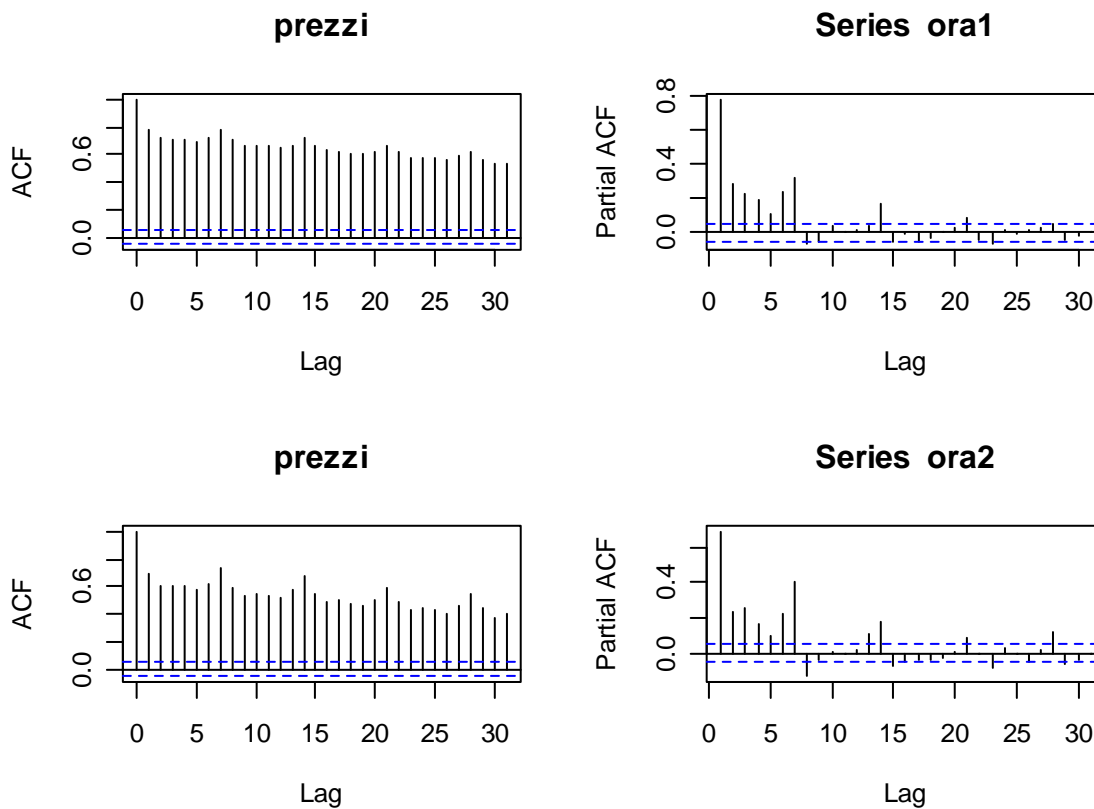
Rendimenti ora 24



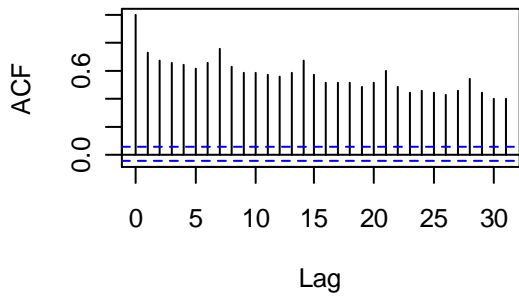
APPENDICE 2

CORRELOGRAMMI E MODELLI STIMATI

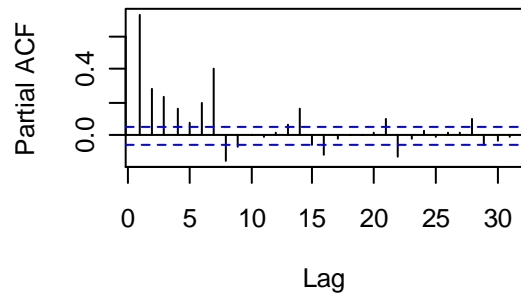
a) Correlogrammi dei prezzi (serie separate ed intera)



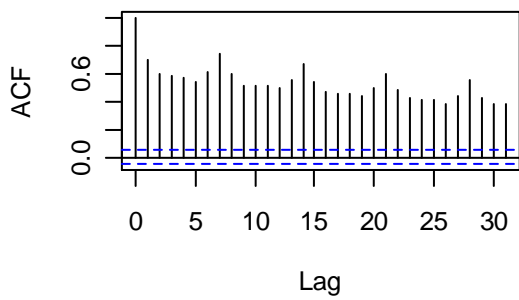
prezzi



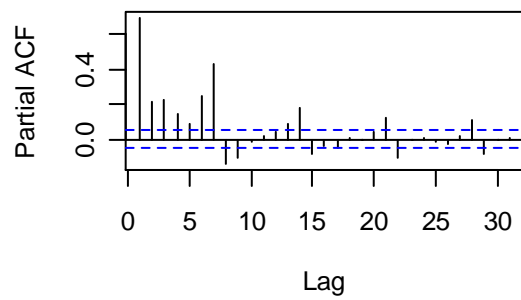
Series ora3



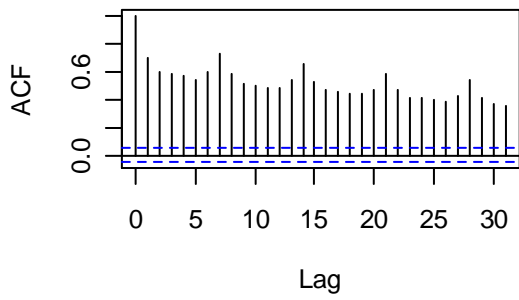
prezzi



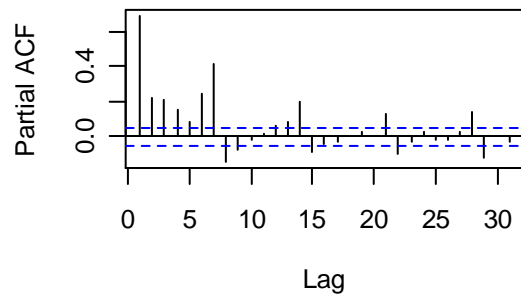
Series ora4



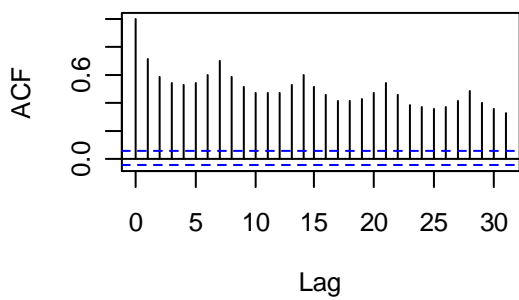
prezzi



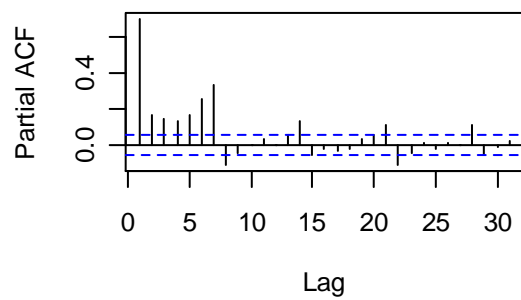
Series ora5



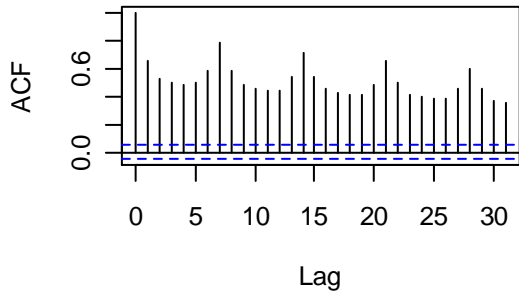
prezzi



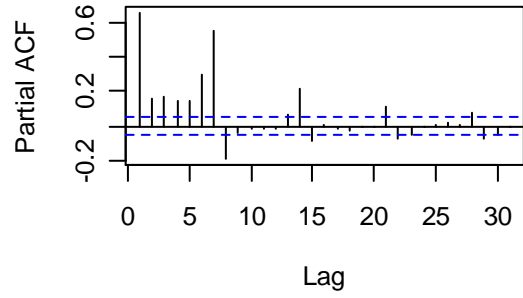
Series ora6



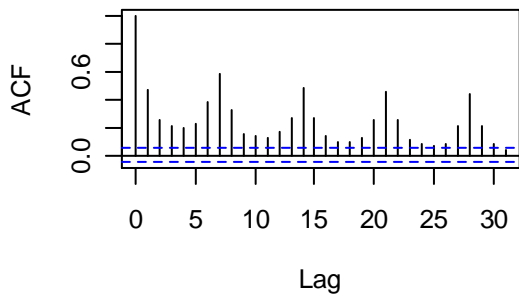
prezzi



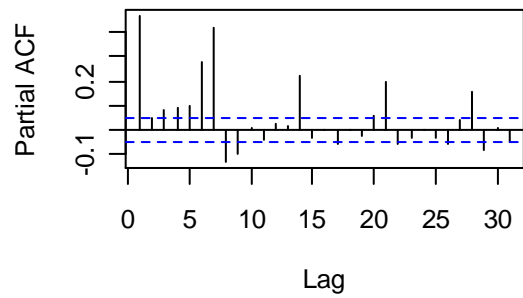
Series ora7



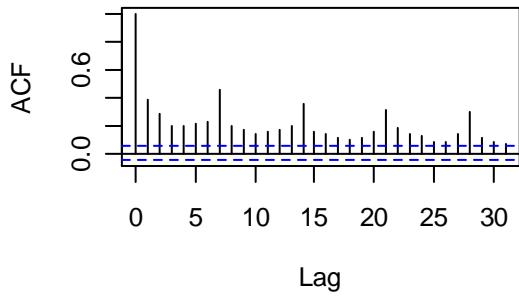
prezzi



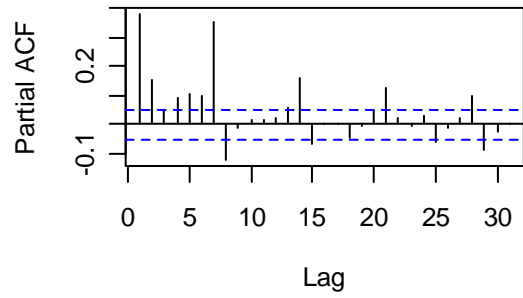
Series ora8



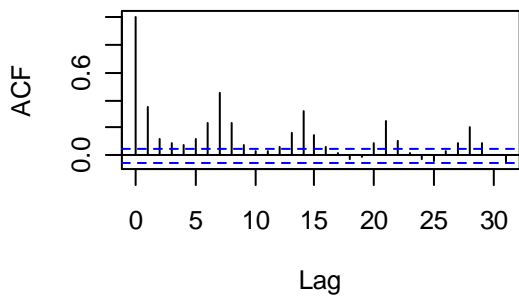
prezzi



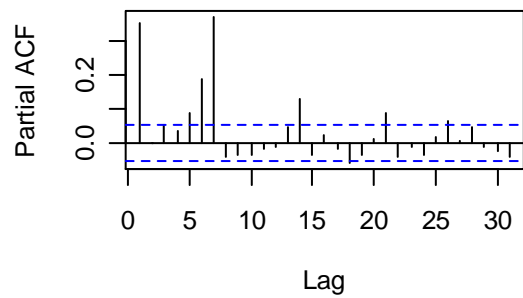
Series ora9



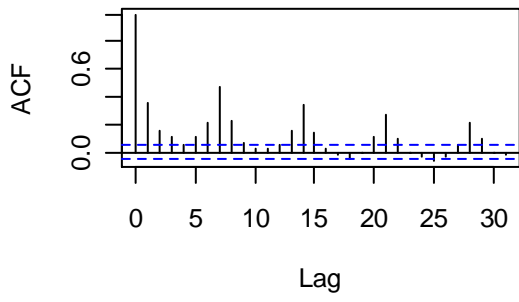
prezzi



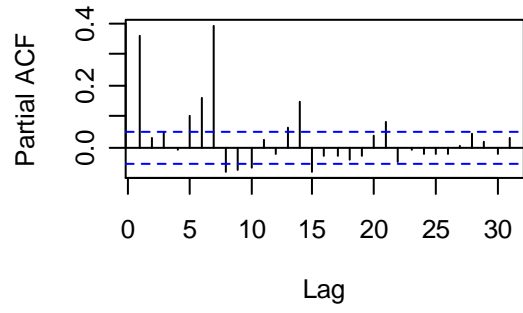
Series ora10



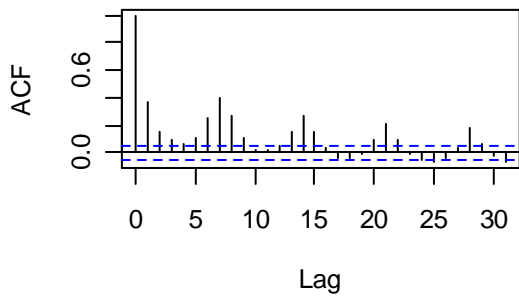
prezzi



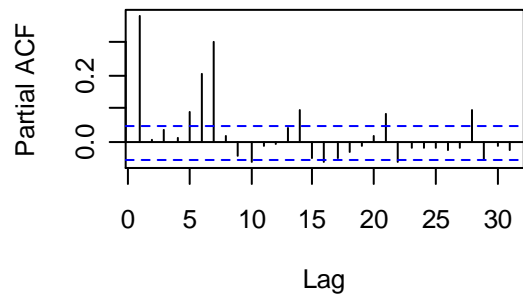
Series ora11



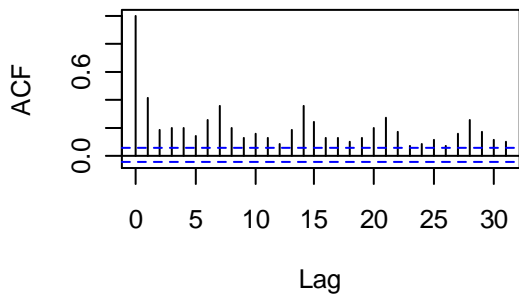
prezzi



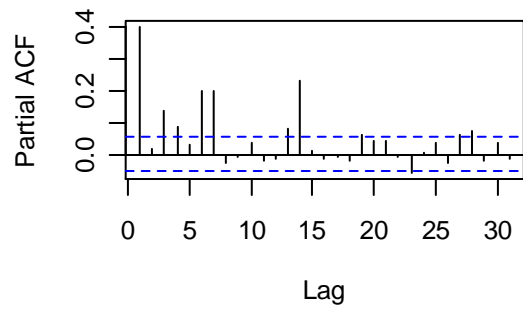
Series ora12



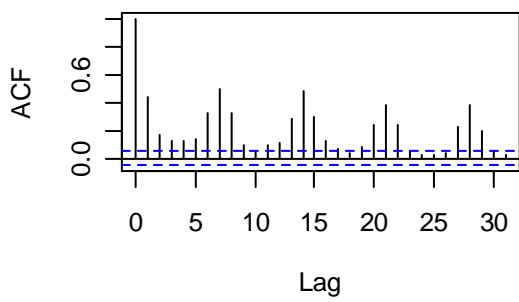
prezzi



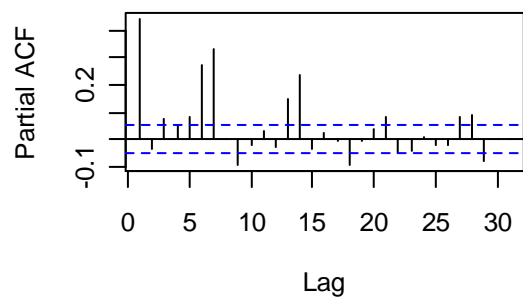
Series ora13



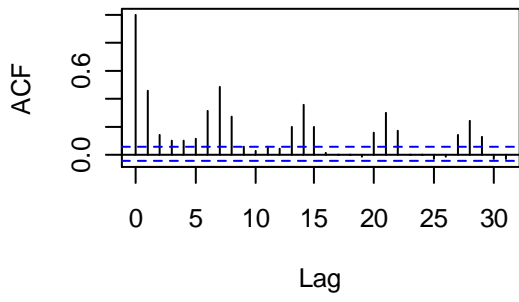
prezzi



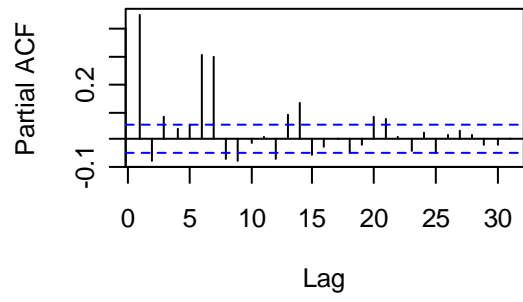
Series ora14



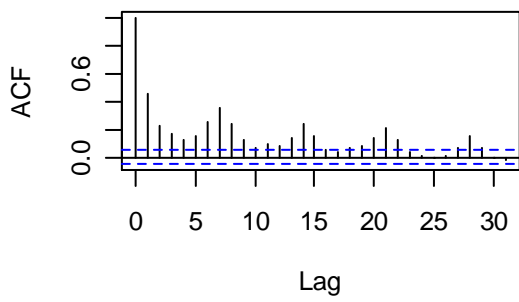
prezzi



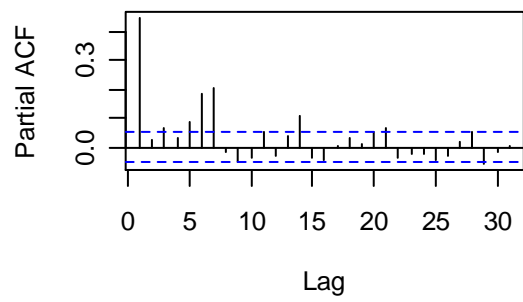
Series ora15



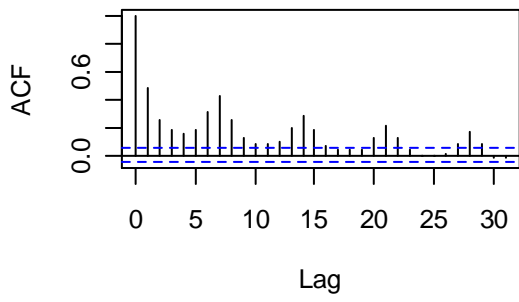
prezzi



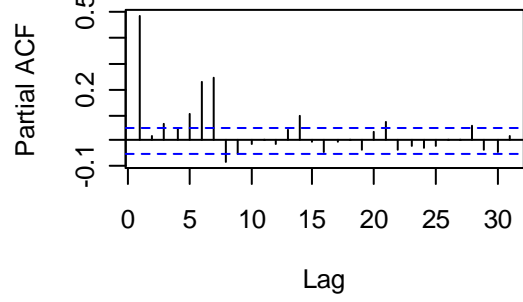
Series ora16



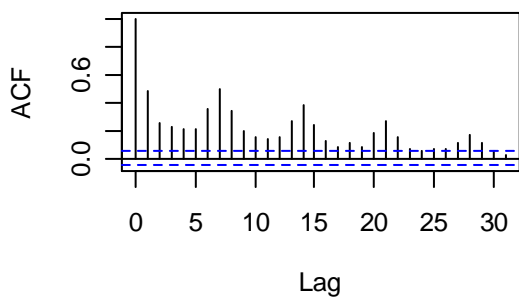
prezzi



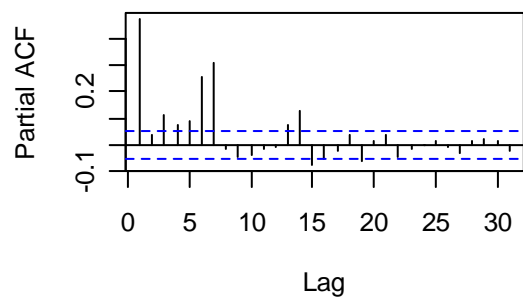
Series ora17



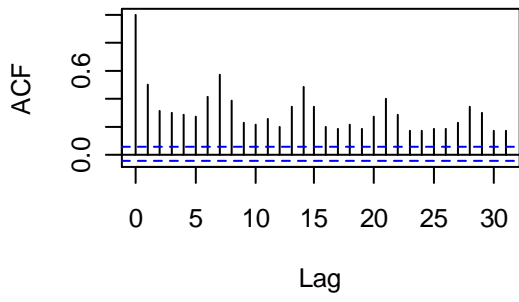
prezzi



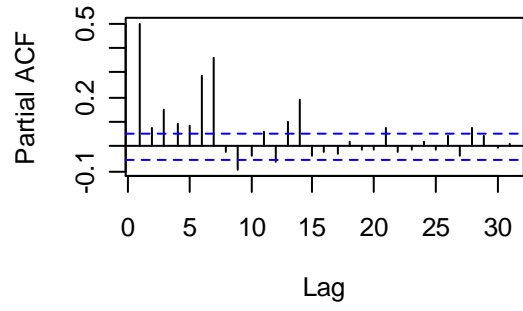
Series ora18



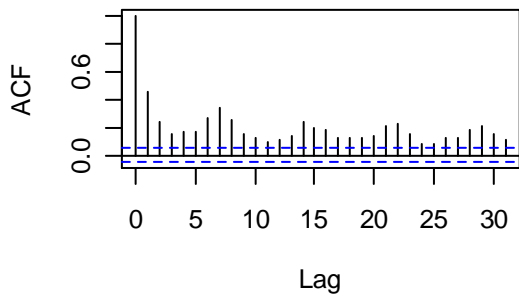
prezzi



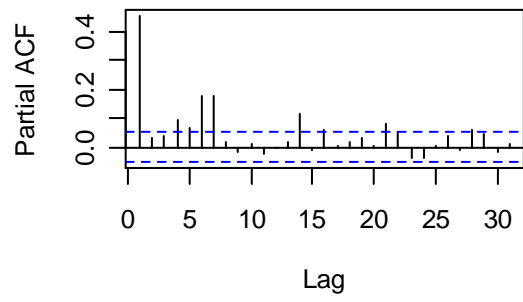
Series ora19



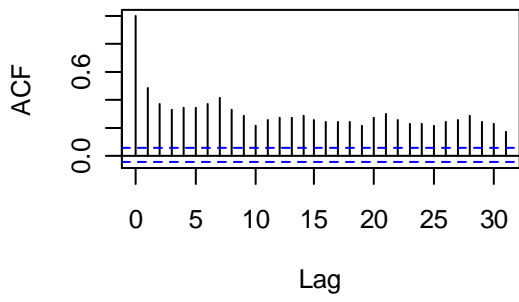
prezzi



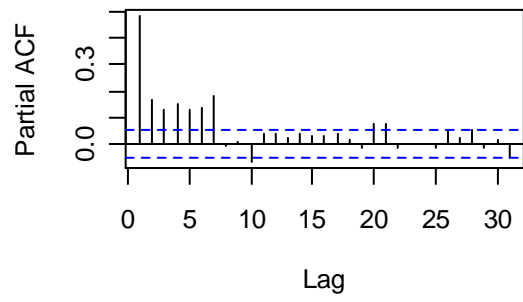
Series ora20



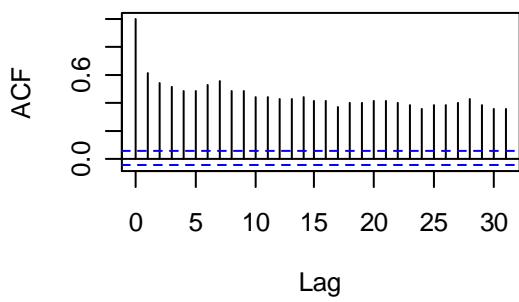
prezzi



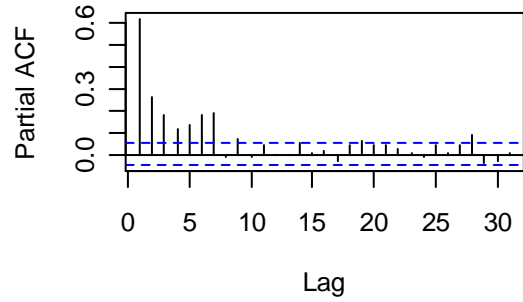
Series ora21



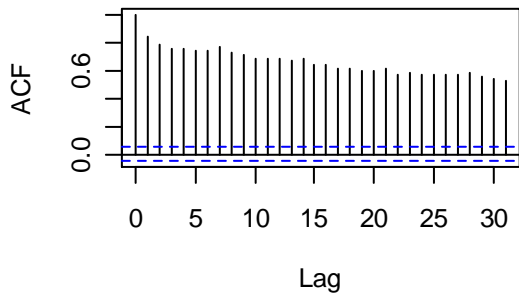
prezzi



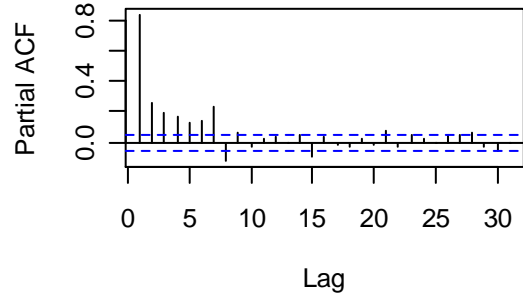
Series ora22



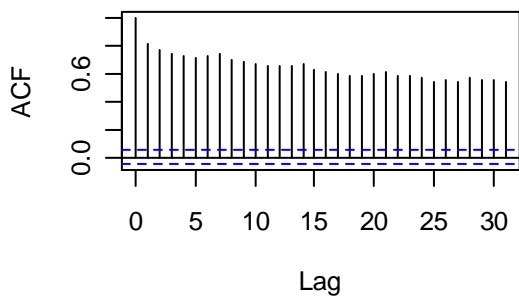
prezzi



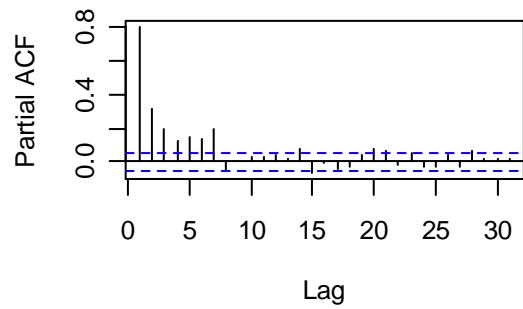
Series ora23



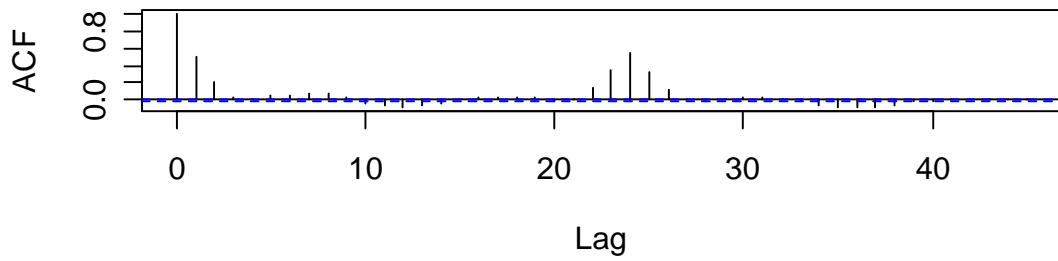
prezzi



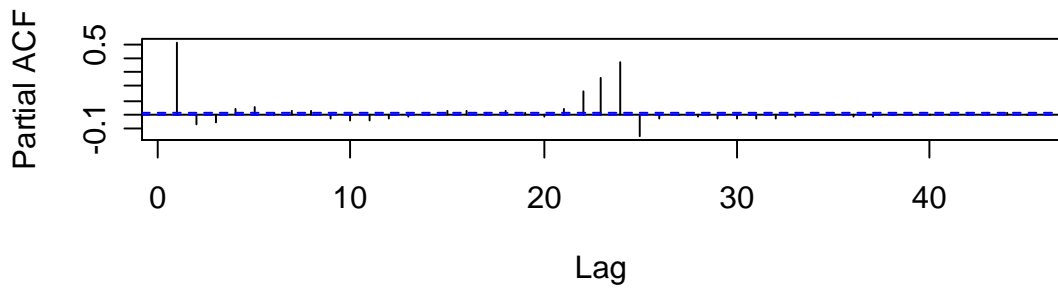
Series ora24



Prezzi

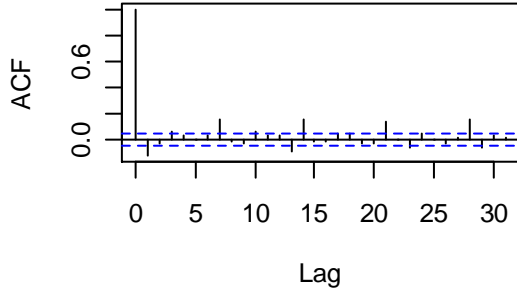


Series prezzi

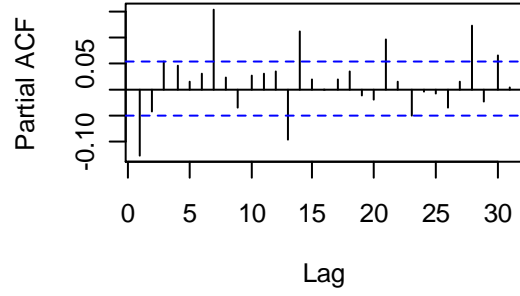


b) Correlogrammi dei rendimenti (serie separate ed intera)

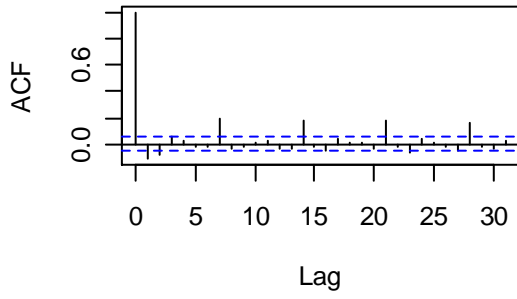
Rendimenti



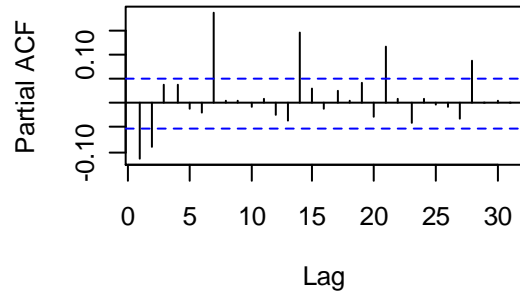
Series r1



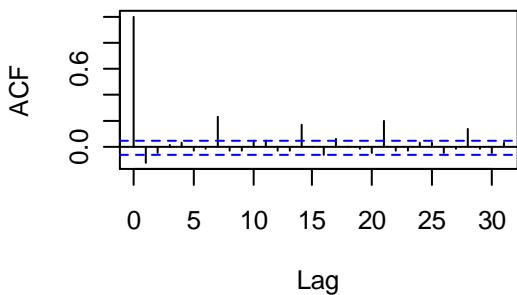
Rendimenti



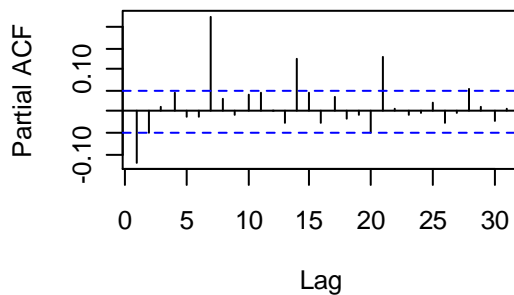
Series r2



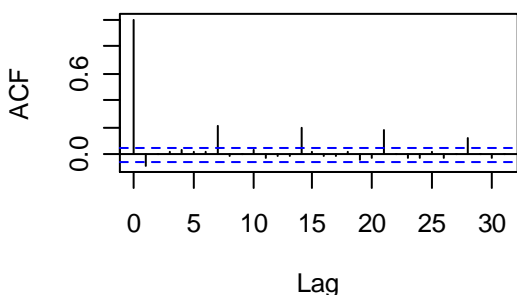
Rendimenti



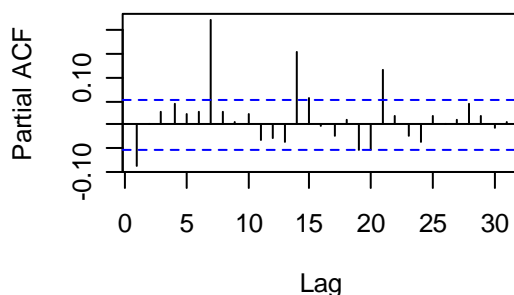
Series r3



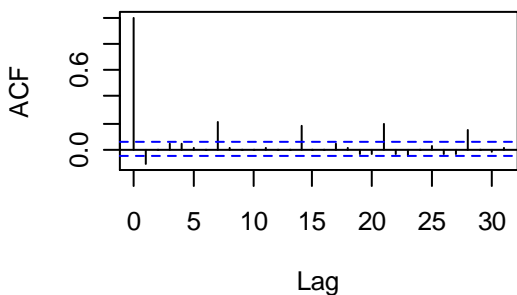
Rendimenti



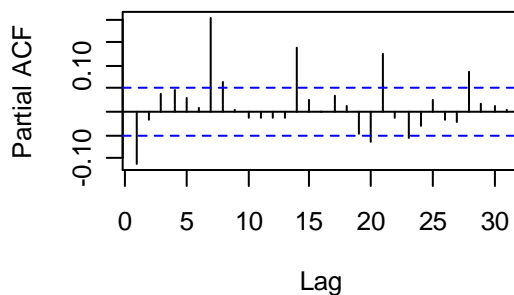
Series r4



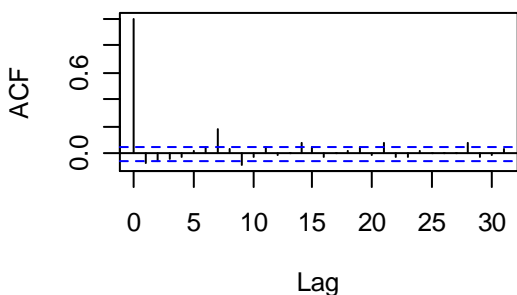
Rendimenti



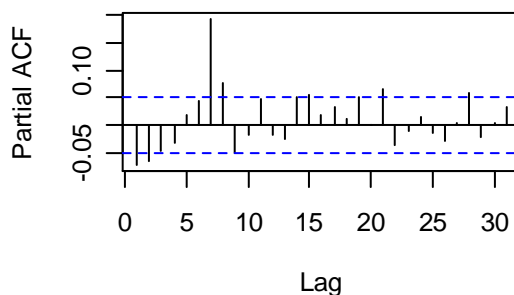
Series r5



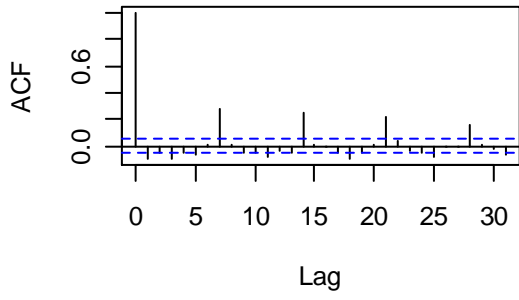
Rendimenti



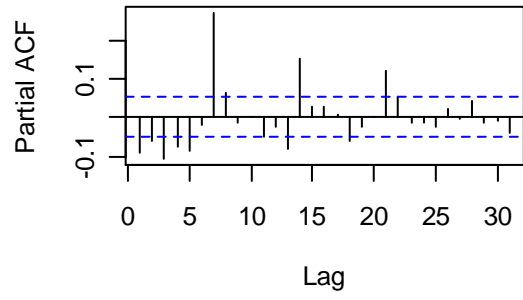
Series r6



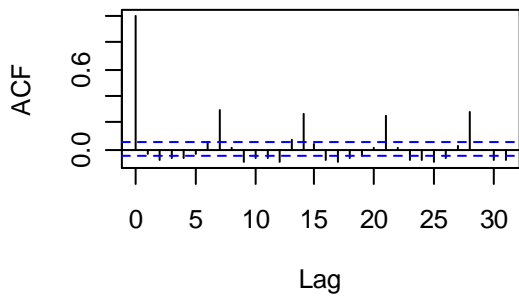
Rendimenti



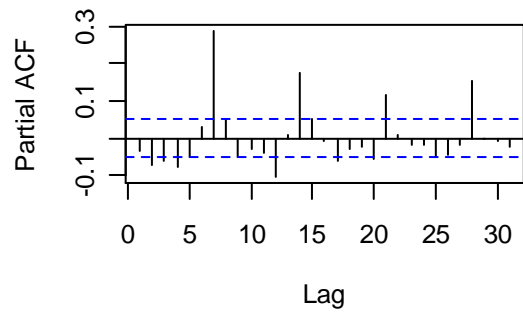
Series r7



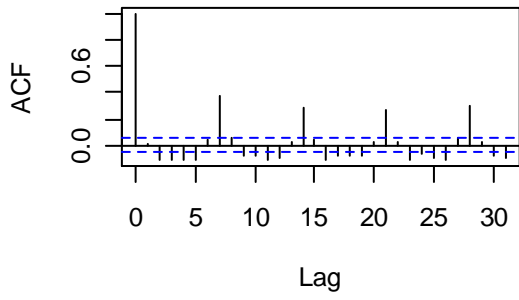
Rendimenti



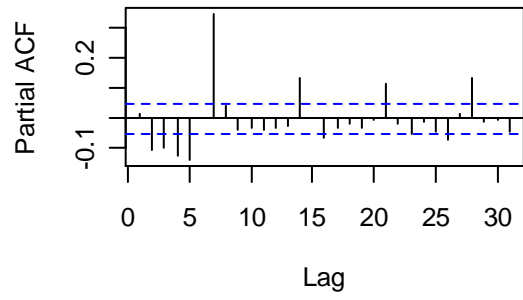
Series r8



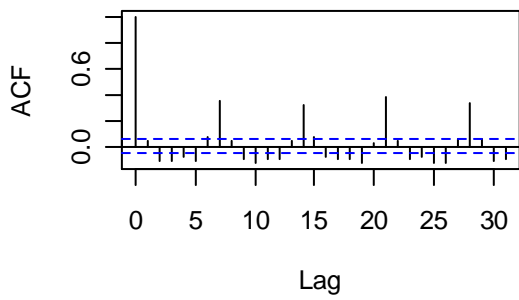
Rendimenti



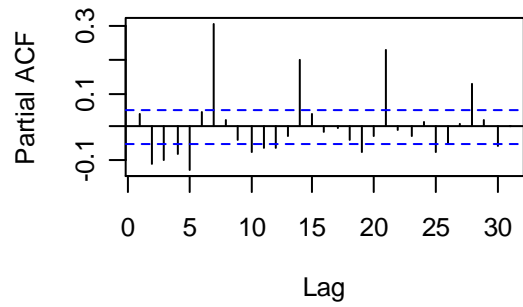
Series r9



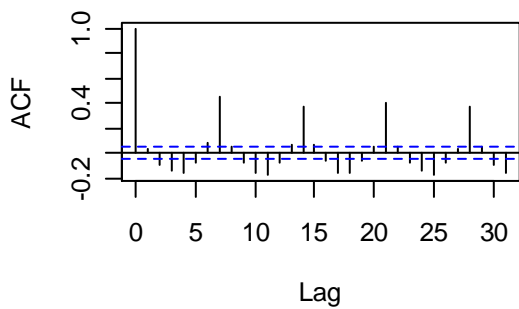
Rendimenti



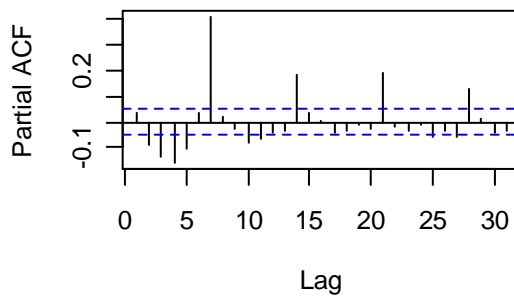
Series r10



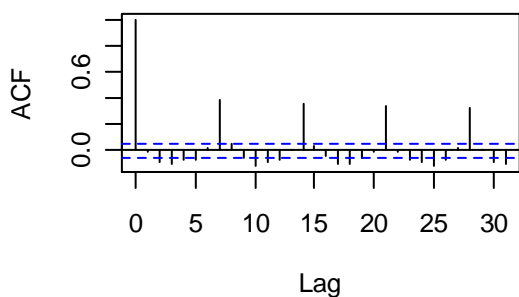
Rendimenti



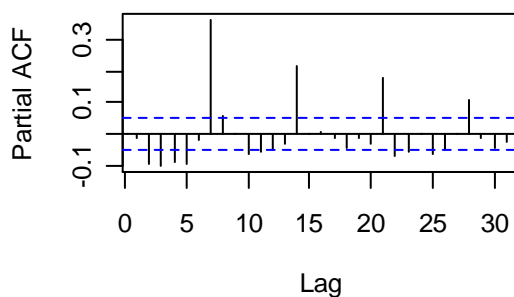
Series r11



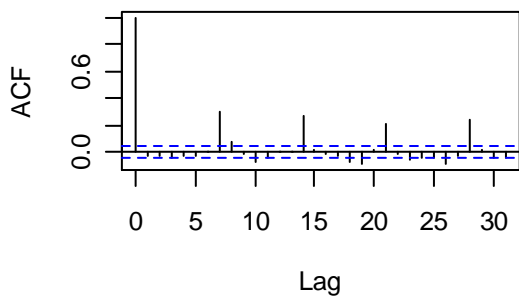
Rendimenti



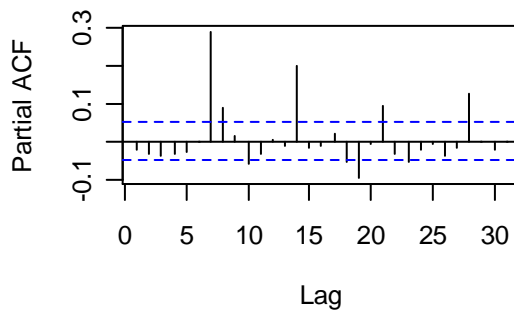
Series r12



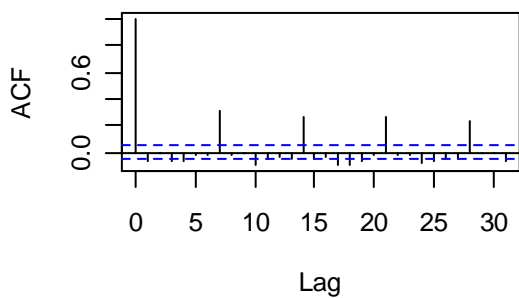
Rendimenti



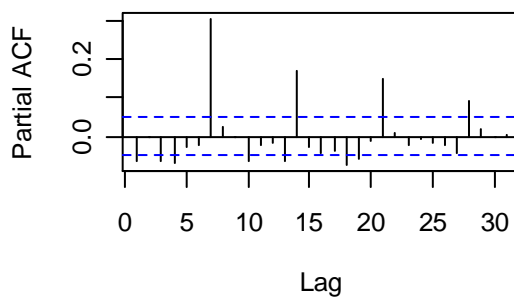
Series r13



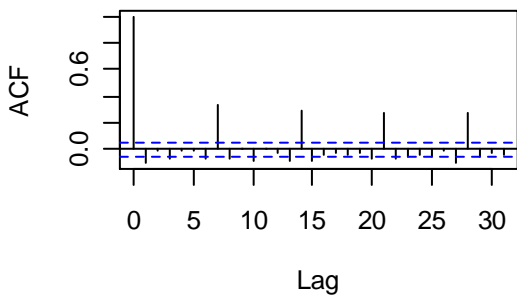
Rendimenti



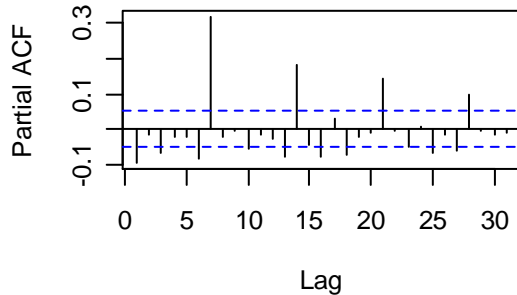
Series r14



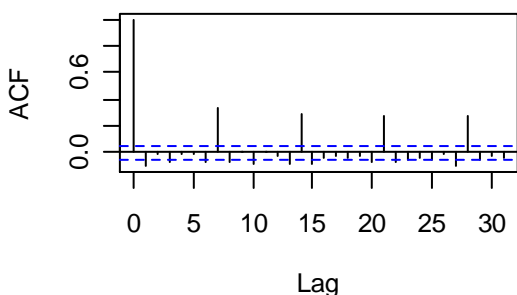
Rendimenti



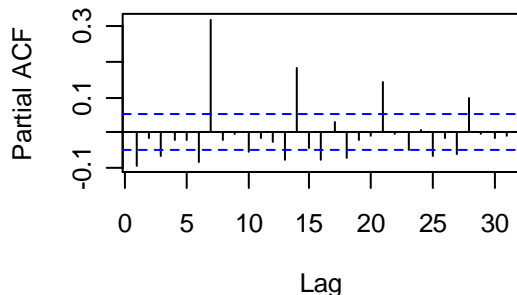
Series r15



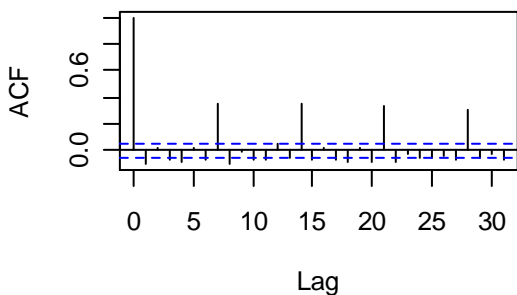
Rendimenti



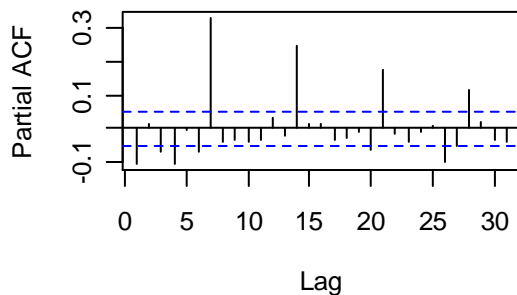
Series r16



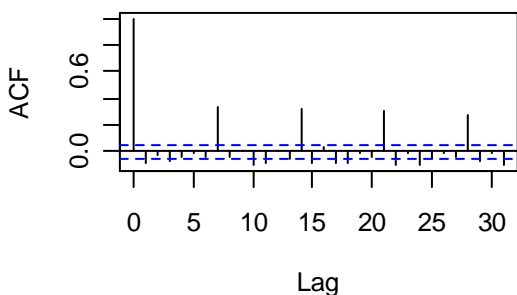
Rendimenti



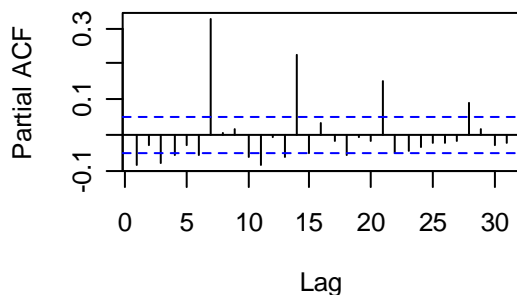
Series r17



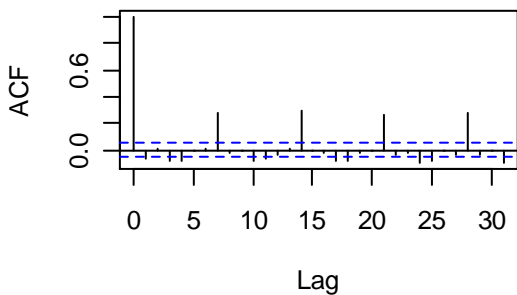
Rendimenti



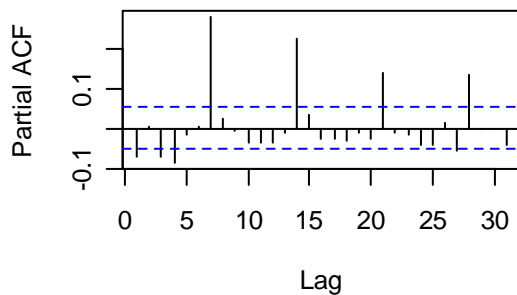
Series r18



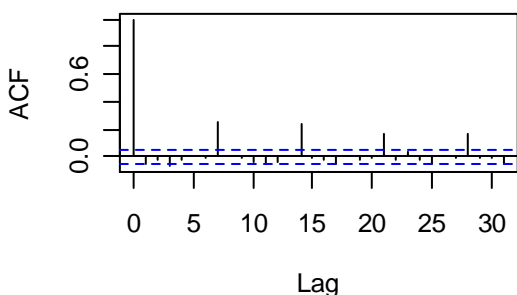
Rendimenti



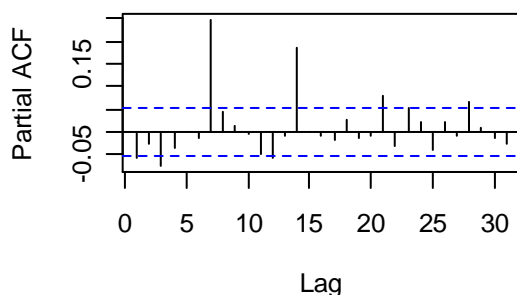
Series r19



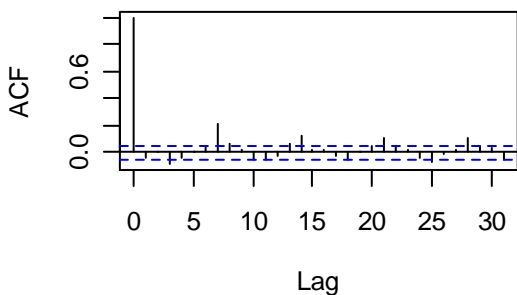
Rendimenti



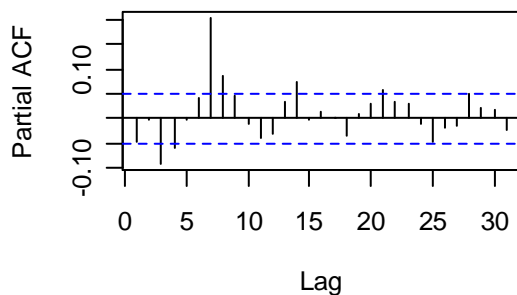
Series r20



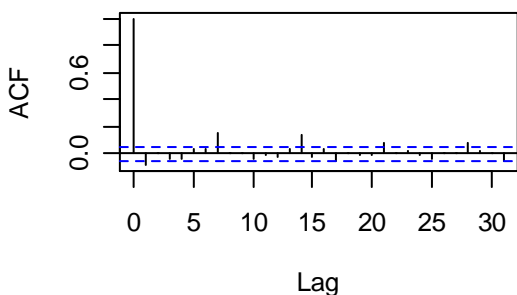
Rendimenti



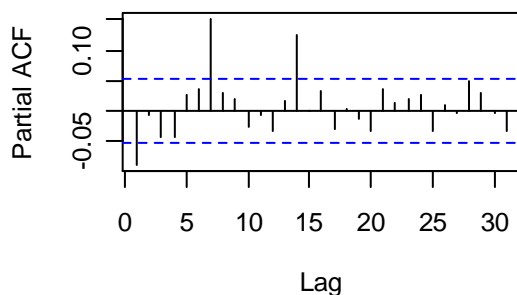
Series r21



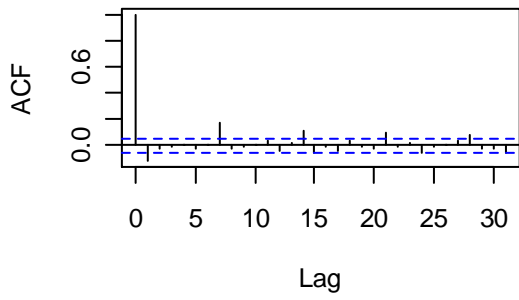
Rendimenti



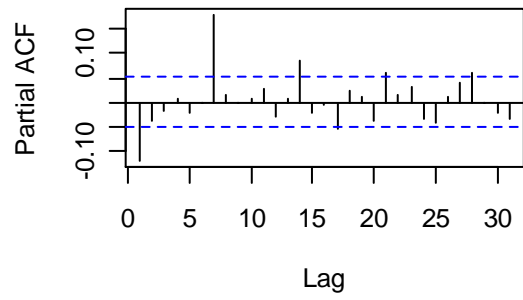
Series r22



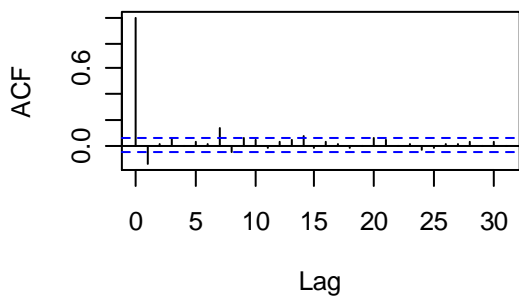
Rendimenti



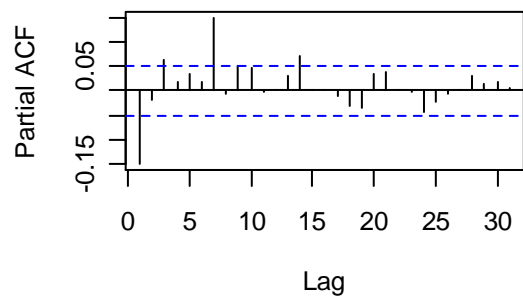
Series r23



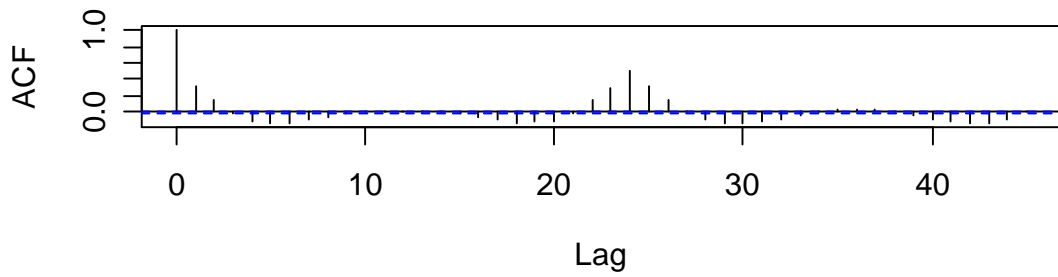
Rendimenti



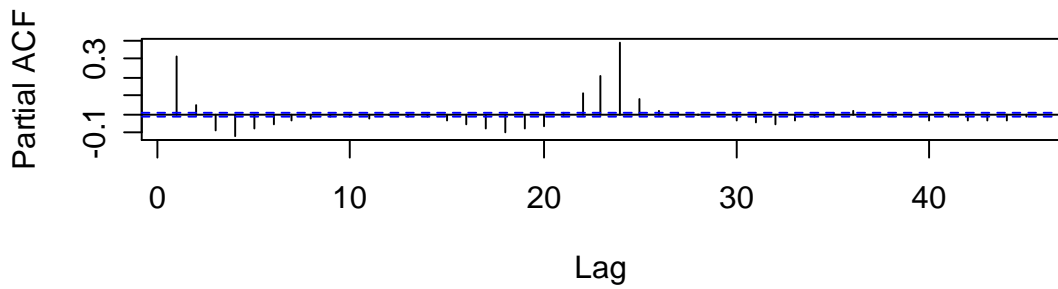
Series r24



Rendimenti



Series rendimenti



MODELLI STIMATI:

Serie intera:

Dependent Variable: D(Y,1,168)
 Method: ML – ARCH (Marquardt)
 Date: 06/23/08 Time: 11:03
 Sample(adjusted): 362 35064
 Included observations: 34703 after adjusting endpoints
 Convergence achieved after 48 iterations
 MA backcast: 356 361, Variance backcast: ON

| | Coefficient | Std. Error | z-Statistic | Prob. |
|-------------------|-------------|------------|-------------|--------|
| AR(1) | 0.383615 | 0.004884 | 78.54338 | 0.0000 |
| AR(168) | -0.303575 | 0.002076 | -146.2291 | 0.0000 |
| SAR(24) | 0.429026 | 0.001988 | 215.7657 | 0.0000 |
| MA(1) | -0.916571 | 0.002710 | -338.1872 | 0.0000 |
| MA(6) | -0.078250 | 0.002668 | -29.32847 | 0.0000 |
| Variance Equation | | | | |
| C | 21.08435 | 0.219018 | 96.26781 | 0.0000 |
| ARCH(1) | 0.430924 | 0.005132 | 83.97474 | 0.0000 |
| GARCH(1) | 0.448992 | 0.003675 | 122.1904 | 0.0000 |

Serie separate:

Dependent Variable: D(LOG(ORA1))
 Method: Least Squares
 Date: 06/23/08 Time: 00:07
 Sample (adjusted): 4/10/2004 3/31/2008
 Included observations: 1452 after adjustments
 Convergence achieved after 18 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 3/26/2004 4/09/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -3.145502 | 0.042588 | -73.85829 | 0.0000 |
| D2 | 3.175410 | 0.057169 | 55.54464 | 0.0000 |
| D3 | -3.142004 | 0.053541 | -58.68461 | 0.0000 |
| D4 | 3.006489 | 0.161563 | 18.60876 | 0.0000 |
| AR(1) | 0.153819 | 0.038713 | 3.973267 | 0.0001 |
| SAR(7) | 0.995825 | 0.003783 | 263.2073 | 0.0000 |
| MA(1) | -0.839788 | 0.019650 | -42.73710 | 0.0000 |
| SMA(7) | -0.863623 | 0.031785 | -27.17055 | 0.0000 |
| SMA(14) | -0.098991 | 0.031391 | -3.153498 | 0.0016 |

Dependent Variable: D(LOG(ORA2))
 Method: Least Squares
 Date: 06/23/08 Time: 00:08
 Sample (adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjustments
 Convergence achieved after 12 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -3.176701 | 0.055020 | -57.73731 | 0.0000 |
| D2 | 3.135846 | 0.065511 | 47.86766 | 0.0000 |
| D3 | -3.121729 | 0.111411 | -28.01996 | 0.0000 |
| D4 | 3.078737 | 0.099697 | 30.88104 | 0.0000 |
| AR(1) | 0.142214 | 0.045316 | 3.138256 | 0.0017 |
| SAR(7) | 1.138757 | 0.038528 | 29.55662 | 0.0000 |
| SAR(14) | -0.143272 | 0.037262 | -3.844998 | 0.0001 |
| MA(1) | -0.845684 | 0.019800 | -42.71069 | 0.0000 |
| SMA(7) | -0.947519 | 0.014188 | -66.78514 | 0.0000 |

Dependent Variable: D(LOG(ORA3))
 Method: Least Squares
 Date: 06/23/08 Time: 00:08
 Sample (adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjustments
 Convergence achieved after 12 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -3.160754 | 0.060211 | -52.49482 | 0.0000 |
| D2 | 3.126614 | 0.079741 | 39.20948 | 0.0000 |
| D3 | -3.047123 | 0.151690 | -20.08783 | 0.0000 |
| D4 | 2.815113 | 0.131294 | 21.44122 | 0.0000 |
| AR(1) | 0.137080 | 0.044930 | 3.050970 | 0.0023 |
| SAR(7) | 1.116092 | 0.041140 | 27.12929 | 0.0000 |
| SAR(14) | -0.119956 | 0.039657 | -3.024814 | 0.0025 |
| MA(1) | -0.798566 | 0.023895 | -33.41952 | 0.0000 |
| SMA(7) | -0.948207 | 0.015178 | -62.47162 | 0.0000 |

Dependent Variable: D(LOG(ORA4))
 Method: Least Squares
 Date: 06/23/08 Time: 00:08
 Sample (adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjustments
 Convergence achieved after 11 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -3.054608 | 0.045429 | -67.23890 | 0.0000 |
| D2 | 3.097529 | 0.058356 | 53.08024 | 0.0000 |
| D3 | -3.019127 | 0.120418 | -25.07199 | 0.0000 |
| D4 | 2.815940 | 0.113712 | 24.76381 | 0.0000 |
| AR(1) | 0.220027 | 0.047362 | 4.645677 | 0.0000 |
| SAR(7) | 1.127124 | 0.054890 | 20.53435 | 0.0000 |
| SAR(14) | -0.135405 | 0.052471 | -2.580573 | 0.0100 |
| MA(1) | -0.849605 | 0.021322 | -39.84674 | 0.0000 |
| SMA(7) | -0.926725 | 0.018362 | -50.46921 | 0.0000 |

Dependent Variable: D(LOG(ORA5))
 Method: ML - ARCH (Marquardt)
 Date: 06/23/08 Time: 11:20
 Sample(adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjusting endpoints
 Convergence achieved after 94 iterations
 MA backcast: 4/09/2004 4/16/2004, Variance backcast: ON

| | Coefficient | Std. Error | z-Statistic | Prob. |
|-------------------|-------------|------------|-------------|--------|
| D1 | -2.919558 | 0.764352 | -3.819649 | 0.0001 |
| D2 | 2.912359 | 0.658064 | 4.425649 | 0.0000 |
| D3 | -2.997026 | 0.416179 | -7.201296 | 0.0000 |
| D4 | 2.940831 | 0.307107 | 9.575926 | 0.0000 |
| AR(1) | 0.289339 | 0.032237 | 8.975259 | 0.0000 |
| SAR(7) | 1.119511 | 0.030153 | 37.12718 | 0.0000 |
| SAR(14) | -0.143561 | 0.028038 | -5.120285 | 0.0000 |
| MA(1) | -0.848288 | 0.018031 | -47.04525 | 0.0000 |
| SMA(7) | -0.864460 | 0.016916 | -51.10404 | 0.0000 |
| Variance Equation | | | | |
| C | 0.001123 | 0.000151 | 7.434791 | 0.0000 |
| ARCH(1) | 0.069399 | 0.008943 | 7.759812 | 0.0000 |
| GARCH(1) | 0.881700 | 0.012857 | 68.57677 | 0.0000 |

Dependent Variable: D(LOG(ORA6))
 Method: ML - ARCH (Marquardt)
 Date: 06/23/08 Time: 11:21
 Sample(adjusted): 4/10/2004 3/31/2008
 Included observations: 1452 after adjusting endpoints
 Convergence achieved after 103 iterations
 MA backcast: 4/02/2004 4/09/2004, Variance backcast: ON

| | Coefficient | Std. Error | z-Statistic | Prob. |
|-------------------|-------------|------------|-------------|--------|
| D1 | -2.563654 | 0.560792 | -4.571489 | 0.0000 |
| D2 | 2.487017 | 0.386500 | 6.434720 | 0.0000 |
| D3 | -2.821716 | 0.545620 | -5.171578 | 0.0000 |
| D4 | 2.982696 | 0.524563 | 5.686063 | 0.0000 |
| AR(1) | 0.270535 | 0.031945 | 8.468667 | 0.0000 |
| SAR(7) | 0.956539 | 0.009466 | 101.0493 | 0.0000 |
| MA(1) | -0.841368 | 0.017580 | -47.85943 | 0.0000 |
| SMA(7) | -0.822966 | 0.020387 | -40.36673 | 0.0000 |
| Variance Equation | | | | |
| C | 0.000590 | 0.000127 | 4.656156 | 0.0000 |
| ARCH(1) | 0.053419 | 0.009326 | 5.728110 | 0.0000 |
| GARCH(1) | 0.923647 | 0.012822 | 72.03809 | 0.0000 |

Dependent Variable: D(LOG(ORA7))
 Method: Least Squares
 Date: 06/23/08 Time: 00:09
 Sample (adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjustments
 Convergence achieved after 17 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -2.739695 | 0.085308 | -32.11527 | 0.0000 |
| D2 | 2.719941 | 0.083743 | 32.47970 | 0.0000 |
| D3 | -2.637220 | 0.146876 | -17.95539 | 0.0000 |
| D4 | 2.662127 | 0.149298 | 17.83093 | 0.0000 |
| AR(1) | 0.267050 | 0.039666 | 6.732455 | 0.0000 |
| SAR(7) | 1.210444 | 0.044583 | 27.15055 | 0.0000 |
| SAR(14) | -0.210314 | 0.044659 | -4.709342 | 0.0000 |
| MA(1) | -0.906238 | 0.015017 | -60.34805 | 0.0000 |
| SMA(7) | -0.979623 | 0.004497 | -217.8171 | 0.0000 |

Dependent Variable: D(LOG(ORA8))
Method: Least Squares
Date: 06/22/08 Time: 10:26
Sample (adjusted): 4/17/2004 3/31/2008
Included observations: 1445 after adjustments
Convergence achieved after 21 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.257851 | 0.037837 | 6.814839 | 0.0000 |
| SAR(7) | 1.263095 | 0.082772 | 15.25990 | 0.0000 |
| SAR(14) | -0.266446 | 0.082812 | -3.217472 | 0.0013 |
| MA(1) | -0.928243 | 0.014358 | -64.64952 | 0.0000 |
| SMA(7) | -0.979917 | 0.005105 | -191.9376 | 0.0000 |

Dependent Variable: D(LOG(ORA9))
Method: Least Squares
Date: 06/22/08 Time: 10:26
Sample (adjusted): 4/17/2004 3/31/2008
Included observations: 1445 after adjustments
Convergence achieved after 23 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.251194 | 0.049728 | 5.051315 | 0.0000 |
| SAR(7) | 1.136073 | 0.051492 | 22.06322 | 0.0000 |
| SAR(14) | -0.140186 | 0.051312 | -2.732038 | 0.0064 |
| MA(1) | -0.857420 | 0.017689 | -48.47085 | 0.0000 |
| SMA(7) | -0.981647 | 0.003645 | -269.3118 | 0.0000 |

Dependent Variable: D(LOG(ORA10))
Method: Least Squares
Date: 06/22/08 Time: 10:27
Sample (adjusted): 4/18/2004 3/31/2008
Included observations: 1444 after adjustments
Convergence achieved after 18 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/10/2004 4/17/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.189091 | 0.041779 | 4.525944 | 0.0000 |
| AR(2) | -0.080045 | 0.039707 | -2.015885 | 0.0440 |
| SAR(7) | 1.205995 | 0.049723 | 24.25427 | 0.0000 |
| SAR(14) | -0.209156 | 0.049700 | -4.208384 | 0.0000 |
| MA(1) | -0.792257 | 0.033521 | -23.63470 | 0.0000 |
| SMA(7) | -0.982591 | 0.003693 | -266.0985 | 0.0000 |

Dependent Variable: D(LOG(ORA11))
Method: Least Squares
Date: 06/22/08 Time: 10:28
Sample (adjusted): 4/16/2004 3/31/2008
Included observations: 1446 after adjustments
Convergence achieved after 9 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/07/2004 4/15/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 1.306244 | 0.066471 | 19.65124 | 0.0000 |
| AR(14) | -0.309073 | 0.066473 | -4.649596 | 0.0000 |
| MA(1) | -0.574324 | 0.060713 | -9.459681 | 0.0000 |
| MA(2) | -0.195166 | 0.041786 | -4.670565 | 0.0000 |
| SMA(7) | -0.982592 | 0.003909 | -251.3504 | 0.0000 |

Dependent Variable: D(LOG(ORA12))
Method: Least Squares
Date: 06/22/08 Time: 10:28
Sample (adjusted): 4/16/2004 3/31/2008
Included observations: 1446 after adjustments
Convergence achieved after 13 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/07/2004 4/15/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 1.268131 | 0.074137 | 17.10517 | 0.0000 |
| AR(14) | -0.271392 | 0.074119 | -3.661564 | 0.0003 |
| MA(1) | -0.557967 | 0.055614 | -10.03284 | 0.0000 |
| MA(2) | -0.217596 | 0.038235 | -5.690953 | 0.0000 |
| SMA(7) | -0.982530 | 0.003900 | -251.9440 | 0.0000 |

Dependent Variable: D(LOG(ORA13))
Method: Least Squares
Date: 06/22/08 Time: 10:29
Sample (adjusted): 4/09/2004 3/31/2008
Included observations: 1453 after adjustments
Convergence achieved after 8 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 3/17/2004 4/08/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 0.990442 | 0.002955 | 335.1751 | 0.0000 |
| MA(1) | -0.628756 | 0.065576 | -9.588245 | 0.0000 |
| MA(2) | -0.220329 | 0.042049 | -5.239780 | 0.0000 |
| SMA(7) | -0.627361 | 0.086128 | -7.284059 | 0.0000 |
| SMA(14) | -0.203840 | 0.074152 | -2.748966 | 0.0061 |
| SMA(21) | -0.142021 | 0.061725 | -2.300876 | 0.0215 |

Dependent Variable: D(LOG(ORA14))
Method: Least Squares
Date: 06/22/08 Time: 10:30
Sample (adjusted): 4/16/2004 3/31/2008
Included observations: 1446 after adjustments
Convergence achieved after 18 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/07/2004 4/15/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 1.287274 | 0.081008 | 15.89069 | 0.0000 |
| AR(14) | -0.291707 | 0.080972 | -3.602558 | 0.0003 |
| MA(1) | -0.671066 | 0.058301 | -11.51036 | 0.0000 |
| MA(2) | -0.185120 | 0.039483 | -4.688574 | 0.0000 |
| SMA(7) | -0.982112 | 0.004899 | -200.4703 | 0.0000 |

Dependent Variable: D(LOG(ORA15))
Method: Least Squares
Date: 06/22/08 Time: 10:31
Sample (adjusted): 4/16/2004 3/31/2008
Included observations: 1446 after adjustments
Convergence achieved after 25 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/07/2004 4/15/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 1.213684 | 0.082448 | 14.72051 | 0.0000 |
| AR(14) | -0.217596 | 0.082318 | -2.643347 | 0.0083 |
| MA(1) | -0.644581 | 0.052496 | -12.27874 | 0.0000 |
| MA(2) | -0.190394 | 0.038646 | -4.926551 | 0.0000 |
| SMA(7) | -0.982608 | 8.93E-07 | -1100476. | 0.0000 |

Dependent Variable: D(LOG(ORA16))
Method: Least Squares
Date: 06/22/08 Time: 10:31
Sample (adjusted): 4/09/2004 3/31/2008
Included observations: 1453 after adjustments
Convergence achieved after 10 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 3/24/2004 4/08/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 0.996086 | 0.001022 | 974.5511 | 0.0000 |
| MA(1) | -0.622008 | 0.052734 | -11.79517 | 0.0000 |
| MA(2) | -0.184404 | 0.037430 | -4.926662 | 0.0000 |
| SMA(7) | -0.740626 | 0.072928 | -10.15564 | 0.0000 |
| SMA(14) | -0.237699 | 0.068910 | -3.449414 | 0.0006 |

Dependent Variable: D(LOG(ORA17))
 Method: Least Squares
 Date: 06/22/08 Time: 10:31
 Sample (adjusted): 4/09/2004 3/31/2008
 Included observations: 1453 after adjustments
 Convergence achieved after 19 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 3/24/2004 4/08/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 0.995105 | 0.001150 | 865.2687 | 0.0000 |
| MA(1) | -0.628211 | 0.050764 | -12.37522 | 0.0000 |
| MA(2) | -0.169415 | 0.035405 | -4.785067 | 0.0000 |
| SMA(7) | -0.710056 | 0.079288 | -8.955384 | 0.0000 |
| SMA(14) | -0.267544 | 0.074074 | -3.611845 | 0.0003 |

Dependent Variable: D(LOG(ORA18))
 Method: Least Squares
 Date: 06/22/08 Time: 10:32
 Sample (adjusted): 4/19/2004 3/31/2008
 Included observations: 1443 after adjustments
 Convergence achieved after 20 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/11/2004 4/18/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.216154 | 0.058125 | 3.718789 | 0.0002 |
| AR(3) | 0.096369 | 0.042994 | 2.241460 | 0.0251 |
| SAR(7) | 1.256085 | 0.067031 | 18.73884 | 0.0000 |
| SAR(14) | -0.259549 | 0.066881 | -3.880785 | 0.0001 |
| MA(1) | -0.875364 | 0.027744 | -31.55186 | 0.0000 |
| SMA(7) | -0.982117 | 0.004732 | -207.5654 | 0.0000 |

Dependent Variable: D(LOG(ORA19))
 Method: Least Squares
 Date: 06/22/08 Time: 10:33
 Sample (adjusted): 4/17/2004 3/31/2008
 Included observations: 1445 after adjustments
 Convergence achieved after 13 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004 4/16/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.211862 | 0.065510 | 3.234054 | 0.0012 |
| SAR(7) | 1.244232 | 0.076720 | 16.21791 | 0.0000 |
| SAR(14) | -0.250378 | 0.076333 | -3.280073 | 0.0011 |
| MA(1) | -0.844327 | 0.025917 | -32.57866 | 0.0000 |
| SMA(7) | -0.982464 | 0.004565 | -215.2251 | 0.0000 |

Dependent Variable: D(LOG(ORA20))
Method: Least Squares
Date: 06/22/08 Time: 10:34
Sample (adjusted): 4/19/2004 3/31/2008
Included observations: 1443 after adjustments
Convergence achieved after 14 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/11/2004 4/18/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.354431 | 0.059275 | 5.979458 | 0.0000 |
| AR(3) | 0.136316 | 0.051783 | 2.632464 | 0.0086 |
| SAR(7) | 1.289769 | 0.078721 | 16.38403 | 0.0000 |
| SAR(14) | -0.302462 | 0.078221 | -3.866782 | 0.0001 |
| MA(1) | -0.937328 | 0.017712 | -52.92153 | 0.0000 |
| SMA(7) | -0.978096 | 0.004061 | -240.8269 | 0.0000 |

Dependent Variable: D(LOG(ORA21))
Method: Least Squares
Date: 06/22/08 Time: 10:34
Sample (adjusted): 4/09/2004 3/31/2008
Included observations: 1453 after adjustments
Convergence achieved after 12 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 3/24/2004 4/08/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(7) | 0.975089 | 0.008361 | 116.6216 | 0.0000 |
| MA(1) | -0.601765 | 0.048280 | -12.46394 | 0.0000 |
| MA(2) | -0.214412 | 0.033135 | -6.470827 | 0.0000 |
| SMA(7) | -0.767659 | 0.054193 | -14.16534 | 0.0000 |
| SMA(14) | -0.194758 | 0.052717 | -3.694380 | 0.0002 |

Dependent Variable: D(LOG(ORA22))
Method: Least Squares
Date: 06/22/08 Time: 10:35
Sample (adjusted): 4/12/2004 3/31/2008
Included observations: 1450 after adjustments
Convergence achieved after 10 iterations
Newey-West HAC Standard Errors & Covariance (lag truncation=7)
MA Backcast: 4/04/2004 4/11/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.357777 | 0.066878 | 5.349674 | 0.0000 |
| AR(3) | 0.167563 | 0.034230 | 4.895184 | 0.0000 |
| SAR(7) | 0.696017 | 0.065753 | 10.58534 | 0.0000 |
| MA(1) | -0.988639 | 0.005399 | -183.1129 | 0.0000 |
| SMA(7) | -0.341383 | 0.108214 | -3.154721 | 0.0016 |

Dependent Variable: D(LOG(ORA23))
 Method: Least Squares
 Date: 06/22/08 Time: 10:36
 Sample (adjusted): 4/12/2004 3/31/2008
 Included observations: 1450 after adjustments
 Convergence achieved after 9 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/11/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| AR(1) | 0.328017 | 0.076499 | 4.287844 | 0.0000 |
| AR(3) | 0.126043 | 0.031002 | 4.065672 | 0.0001 |
| SAR(7) | 0.461588 | 0.153851 | 3.000226 | 0.0027 |
| MA(1) | -0.983301 | 0.007225 | -136.1017 | 0.0000 |

Dependent Variable: D(LOG(ORA24))
 Method: Least Squares
 Date: 06/23/08 Time: 00:11
 Sample (adjusted): 4/10/2004 3/31/2008
 Included observations: 1452 after adjustments
 Convergence achieved after 11 iterations
 Newey-West HAC Standard Errors & Covariance (lag truncation=7)
 MA Backcast: 4/09/2004

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| D1 | -3.341048 | 0.040427 | -82.64310 | 0.0000 |
| D2 | 3.169352 | 0.061676 | 51.38708 | 0.0000 |
| D3 | -3.231465 | 0.064870 | -49.81445 | 0.0000 |
| D4 | 3.268483 | 0.059497 | 54.93537 | 0.0000 |
| AR(1) | 0.340293 | 0.058848 | 5.782529 | 0.0000 |
| SAR(7) | 0.205783 | 0.039509 | 5.208504 | 0.0000 |
| MA(1) | -0.896471 | 0.022703 | -39.48763 | 0.0000 |

BIBLIOGRAFIA:

- [1] Conejo A. J., Contreras J., Espinola R., Plazas M. A. (2005a), "Forecasting electricity prices for a day-ahead pool-based electric energy market" *International Journal of Forecasting*, 21(3), 435-462
- [2] Crespo Cuaresma J., Hlouskova J., Kossmeier S., Obersteiner M. (2004), "Forecasting electricity spot-prices using linear uni variate time-series models" *Applied Energy*, 77, 87-106
- [3] Escribano A., Peña J. I., Villaplana P. (2002), "Modeling electricity prices: international evidence" Working paper 02-27, University Carlos III, Madrid
- [4] Huisman R., Hurman C., Mahieu R. (2007), "Hourly electricity prices in day-ahead markets" *Energy Economics*, 29, 240-248

[5] Misiorek A., Trueck S., Weron R. (2006), “Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models” *Studies in Nonlinear Dynamics & Econometrics*, 10(3), Article 2

[6] Nogales F. J., Contreras J., Conejo A. J., Espinola R., Plazas M. A. (2002), “Forecasting Next-Day Electricity Prices by Time Series Models” *IEEE Transactions on Power Systems*, 17, 342-348

[7] Sito Web www.mercatoelettrico.org

[8] Verbeek M., “Econometria”, Zanichelli (2006)

[9] Weron R. (2006), “Modeling and Forecasting Electricity Loads and Prices. A Statistical Approach” John Wiley & Sons, Ltd.