

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE
CORSO DI LAUREA SPECIALISTICA IN SCIENZE STATISTICHE,
ECONOMICHE, FINANZIARIE ED AZIENDALI



TESI DI LAUREA

MODELLAZIONE E PREVISIONE DEI PREZZI DEL
MERCATO DELL'ENERGIA ELETTRICA: UN CONFRONTO
FRA DIVERSE STRATEGIE

Relatore: Ch.ma Prof.ssa Luisa Bisaglia

Correlatore: Ch.mo Prof. Silvano Bordignon

Laureanda: Marina Marzovilli

Matricola: 584647

ANNO ACCADEMICO 2008-2009

Indice

Introduzione.....	5
Capitolo 1	9
1.1 La liberalizzazione del mercato elettrico in Italia e nel mondo.....	9
1.2 <i>Power Pool</i> e <i>Power Exchange</i>	10
1.2.1 <i>Power pools</i>	11
1.2.2 <i>Power Exchanges</i>	11
1.3 La struttura del mercato	12
1.4 Il mercato dell'energia elettrica in Europa	13
1.4.1 Inghilterra e Galles	13
1.4.2 Il Nord Pool	14
1.4.3 L'Europa continentale	14
1.4.4 L'Italia	15
1.4.5 La Francia	16
1.5 Il mercato dell'energia elettrica in America del Nord.....	17
1.6 Conclusioni	17
Capitolo 2	19
2.1 <i>Jumps o Spikes</i>	19
2.2 Stagionalità	20
2.3 <i>Mean Reversion</i>	22
2.4 Distribuzioni di riferimento per i prezzi dell'elettricità	23
2.4.1 Distribuzioni stabili	23
2.4.2 Distribuzioni iperboliche	24

2.5 Conclusioni	24
Capitolo 3	27
3.1 Modelli per serie storiche	27
3.1.1 Metodi statistici per la previsione dei prezzi dell'energia elettrica	28
3.1.1.1 Modelli ARMA e ARMAX.....	29
3.1.1.2 Modelli GARCH	30
3.1.1.3 Modelli <i>Regime Switching</i>	31
3.1.2 Metodi quantitativi per il <i>pricing</i> dei derivati	34
3.1.2.1 Modelli <i>Jump-Diffusion</i>	35
3.1.2.2 Modelli Ibridi	36
3.2 Modelli per dati di <i>panel</i> e multivariati	36
3.2.1 Modelli SUR e loro stima.....	37
3.2.2 Modello VAR	39
3.3 Conclusioni	41
Capitolo 4	43
4.1 La serie storica dei prezzi	44
4.2 Analisi descrittiva dei dati giornalieri	50
4.3 Le serie orarie	52
4.4 Analisi dei dati per le previsioni.....	53
Capitolo 5	57
5.1 Stima della componente deterministica.....	57
5.2 Stima della componente stocastica	61
5.2.1 ARMA - GARCH.....	62
5.2.2 <i>Markov-Switching</i>	64
5.2.3 SUR	66
5.2.4 VAR.....	68
5.3 Analisi delle serie depurate dai picchi	70
5.3.1 ARMA - GARCH.....	71
5.3.2 SUR	73
5.3.3 VAR.....	75

5.4 Conclusioni	75
Capitolo 6	77
6.1 Previsioni	78
6.2 Previsioni un mese in avanti	84
6.3 Conclusioni	86
Conclusioni.....	89
Appendice A.....	91
Appendice B.....	99
Appendice C.....	105
Bibliografia	111

Introduzione

Il mercato dell'energia elettrica è stato liberalizzato negli ultimi due decenni, a partire dagli inizi degli anni '90. La precedente situazione monopolistica ha lasciato spazio ad un mercato competitivo nel quale i consumatori hanno la libertà di decidere l'operatore dal quale ricevere il servizio.

L'energia elettrica, però, ha caratteristiche che la distinguono da tutte le altre fonti energetiche, poiché le possibilità di immagazzinarla sono davvero poche: è possibile costruirne riserve solo indirettamente, attraverso altre fonti energetiche, come acqua, gas o carbone. Tra le conseguenze principali dell'impossibilità di immagazzinare energia ci sono il forte comportamento stagionale delle serie dei prezzi elettrici e la presenza di numerosi valori anomali, i cosiddetti *spikes*, che possono essere legati, per esempio, all'improvvisa chiusura di uno stabilimento nucleare o ad un'improvvisa ondata di calore. A causa di ciò, la volatilità che si rileva nei prezzi dell'energia elettrica è decisamente maggiore di quella che si riscontra nei mercati azionari ordinari. L'impossibilità di creare riserve energetiche, inoltre, implica che ci siano numerose differenze tra i diversi mercati regionali, ma questo non determina l'esistenza di opportunità di arbitraggio: non è possibile, infatti, comprare energia per trasportarla e venderla all'estero.

La domanda di elettricità è, inoltre, fortemente legata alla temperatura: l'energia è necessaria per la refrigerazione in aree particolarmente calde e per il riscaldamento laddove le temperature risultano molto rigide. Nei periodi caldi, infatti, gli impianti nucleari sono costretti a ridurre la loro attività a causa della scarsità di acqua fredda; allo stesso tempo, il caldo riduce la produzione di energia elettrica che deriva dalle fonti

idriche a causa dell'eccessiva evaporazione dell'acqua. Queste caratteristiche spiegano la forte presenza di stagionalità nelle serie storiche dei prezzi elettrici.

Proprio per l'impossibilità di creare scorte energetiche, i prezzi dell'energia elettrica vengono stabiliti simultaneamente per tutte le ore del giorno successivo in un prestabilito momento della giornata (le 11:00 o le 12:00, in genere). Per le sue caratteristiche peculiari, che la distinguono da tutti gli altri *assets* regolarmente negoziati, l'energia elettrica è stata oggetto di numerosi studi, nei quali sono stati proposti differenti modelli che permettano di tener conto di stagionalità, *mean reversion* e presenza di picchi. Alcuni studiosi (p.e. De Jong, 2006) hanno confrontato modelli classici per serie storiche, del tipo ARMA-GARCH, con modelli più complessi che permettessero di considerare la presenza di più regimi. Altri (p.e. Huisman et al, 2006) hanno proposto modelli innovativi, che, trattando le serie infragiornaliere simultaneamente, permettessero di considerare la presenza di correlazione tra le diverse ore del giorno.

In questo lavoro ho applicato alcuni dei modelli proposti in letteratura e loro evoluzioni a dati che ho ritenuto interessanti da analizzare, quali quelli del mercato dell'energia elettrica francese, dall'inizio delle contrattazioni ad Aprile 2009, al fine di ottenere previsioni. Ho scelto di utilizzare questi dati perché mostrano le caratteristiche proprie dei prezzi elettrici, presentando picchi molto evidenti in corrispondenza delle ondate di calore del 2003 e del 2006 e delle agitazioni dei lavoratori nell'autunno 2007.

Il mio lavoro si sviluppa nel seguente modo: nel Capitolo 1 descriverò le caratteristiche dei mercati dell'energia elettrica in Europa e nel Mondo, spiegandone l'evoluzione nel tempo; nel Capitolo 2 discuterò in modo approfondito le caratteristiche peculiari dei prezzi dell'elettricità, quali la *mean reversion*, la stagionalità e la presenza dei picchi; nel Capitolo 3 presenterò alcuni modelli utilizzati in letteratura per l'analisi di serie di prezzi dell'energia elettrica, mettendo a confronto metodi univariati e multivariati, cosa che ancora non era stata fatta in letteratura, almeno a mia conoscenza; nel Capitolo 4 riporterò le principali statistiche descrittive e le analisi grafiche sulle serie dei prezzi, sia orarie sia giornaliere del mercato francese, *Paris Powernext*; nel Capitolo 5 mostrerò i risultati che ho ottenuto, applicando quattro modelli, due univariati (ARMA-GARCH e *Markov Switching*) e due multivariati (SUR e VAR), che ho ritenuto interessanti, ai dati orari presentati nel Capitolo 4, mettendo in luce in particolare le differenze di

comportamento delle diverse ore del giorno e lavorando sia sui dati originari, che su quelli ripuliti dai picchi secondo l'approccio proposto da Weron (2008); nel Capitolo 6, infine, riporterò le previsioni che ho costruito applicando ai dati i modelli proposti, seguendo due strade: in primo luogo ho considerato tutte le previsioni k passi in avanti, $k=1, \dots, 28$, per quattro mesi, uno per ogni stagione e ho confrontato, graficamente e attraverso indicatori di bontà previsiva, i risultati ottenuti con la serie originaria, passo per passo; in un secondo momento, considerando di avere informazioni disponibili fino al tempo t , ho costruito le serie dei valori previsti fino a 28 passi in avanti, confrontandole con i primi 28 giorni del mese al quale erano riferite le previsioni. Per fare questo ho lavorato con i software S plus e R, costruendo manualmente le routine necessarie.

La mia analisi mi ha portato a concludere che, per quanto riguarda la scelta del modello da applicare ai prezzi dell'energia elettrica, tutti i modelli utilizzati sono in grado di cogliere alcune delle peculiarità proprie della serie dei dati, ma nessuno sembra modellare bene tutte le caratteristiche tipiche dei prezzi dell'elettricità.

Per quanto riguarda la previsioni, invece, ho confrontato tra loro i risultati ottenuti con i diversi modelli, attraverso indicatori di bontà previsiva e ho valutato se le differenze osservate fossero o meno significative. Da questa analisi ho concluso che, nel prevedere un passo in avanti, quindi a breve termine, per le 24 ore del giorno successivo, modelli univariati, di tipo ARMA-GARCH, permettono di ottenere i risultati migliori. Nel caso in cui si lavori con un orizzonte temporale più ampio, invece, i modelli multivariati, ed in particolare i VAR, tenendo conto delle correlazioni tra errori di equazioni successive e dell'effetto di un'ora su quelle contigue, producono le previsioni migliori.

Capitolo 1

Il mercato dell'energia elettrica e le sue caratteristiche

1.1 La liberalizzazione del mercato elettrico in Italia e nel mondo

Il paese che si può considerare pioniere nell'introduzione di un mercato competitivo per la distribuzione dell'energia elettrica è il Cile: nel 1982 una riforma, basata sull'idea che varie compagnie di produzione e distribuzione potessero competere per fornire il servizio alla comunità, diede il via alla privatizzazione che, su scala sempre maggiore, iniziò nel 1986 e portò alla nascita di una parziale disintegrazione verticale del settore.

Negli anni '90 molti paesi europei seguirono l'esempio del Cile: nel 1990 il settore dell'elettricità fu riorganizzato in Gran Bretagna, nel 1992 aprì il Nordic Market, prima in Norvegia, poi in Svezia, Finlandia e Danimarca. Nel resto del mondo, anche Australia e Nuova Zelanda riformarono il settore dell'energia tra il 1996 e il 1998.

Nel Nord Est dell'America un numero consistente di mercati cominciò ad operare negli anni '90 e alla fine del decennio anche California, Canada e Alberta seguirono la tendenza del resto del mondo.

In Italia, invece, la nascita di un mercato regolamentato è avvenuta solo successivamente al decreto legislativo del marzo 1999.

In alcuni casi la liberalizzazione ha operato con successo per decenni, in altri casi, invece, alcune riforme hanno condotto a veri e propri fallimenti. Un esempio su tutti è il crollo del mercato Californiano tra il 2000 e il 2001 alla quale seguirono la spettacolare bancarotta di Enron e i *blackouts* in Nord America e in tutta Europa nel 2003.

D'altra parte la liberalizzazione ha avuto un impatto positivo sull'economia, rendendo più efficiente la rete di distribuzione dell'elettricità e riducendone i prezzi.

In realtà molti contestano questo punto sostenendo che la liberalizzazione abbia addirittura portato ad un aumento dei prezzi dell'energia elettrica perché, pur istituendo la concorrenza, ha introdotto nuove tasse, che gravano specialmente sulle piccole imprese e sui singoli consumatori.

In base alle loro caratteristiche organizzative si possono individuare tre tipologie di mercati:

1. *capacity payments*, nel quale si assegna ad ogni produttore di elettricità un pagamento giornaliero proporzionale al suo contributo alla disponibilità di energia;
2. *capacity market*, il cui obiettivo principale è garantire un'adeguata capacità di fornitura in modo che domanda e offerta si incontrino;
3. *'energy only' market* in cui si provvede a compensare costi fissi e variabili con la vendita di energia. I forti movimenti verso l'alto dei prezzi segnalano agli investitori la necessità di generare nuova capacità.

È necessario sottolineare che non esiste un modello migliore di mercato, ma che in ogni contesto bisogna considerare la situazione economica e le caratteristiche tecniche per valutare il modello più adatto.

1.2 Power Pool e Power Exchange

Con la liberalizzazione del settore dell'energia elettrica è stato necessario dare vita a mercati organizzati; a seconda del grado di iniziativa e di partecipazione è possibile

distinguere, spesso in modo non banale, tra due tipi di mercati: *power pools* e *power exchanges*.

1.2.1 Power pools

I mercati di tipo *power pool* possono a loro volta essere ulteriormente classificati in tecnici ed economici. I *technical pools* sono sempre esistiti e si basano sull'utilizzo di un sistema di cooperazione per ottimizzare l'erogazione di energia minimizzando i costi. Questo tipo di mercato è caratterizzato da un basso livello di scambi internazionali a causa della scarsa capacità di collegamento: gli unici scambi che si verificano sono quelli di servizi tra le differenti aree.

Gli *economic pools*, o più semplicemente *power pools*, sono stati creati attraverso iniziative dei governi al fine di facilitare la competizione tra i diversi fornitori. In questo contesto nessuna trattazione è consentita al di fuori del *pool* e al mercato possono partecipare solo i generatori di energia elettrica. Il prezzo di mercato è determinato attraverso un'asta, individuando l'intersezione tra la curva di offerta e la domanda stimata. Questo garantisce un certo livello di trasparenza delle transazioni.

1.2.2 Power Exchanges

I mercati di tipo *Power Exchanges* sono basati principalmente sull'iniziativa privata. Al mercato possono partecipare generatori, compagnie di distribuzione, commercianti e grandi utenti e la partecipazione è volontaria.

L'obiettivo di questo tipo di mercati è quello di favorire l'incontro tra domanda e offerta al fine di determinare il prezzo di mercato che sarà poi reso pubblico. Questo prezzo non è stabilito con transazioni continue ma ogni giorno, attraverso un'asta bilaterale, acquirenti e offerenti formulano le loro richieste per le 24 ore del giorno successivo ed ogni prezzo orario si ottiene dall'equilibrio domanda-offerta. Nelle aste con prezzo uniforme gli acquirenti con richieste maggiori o uguali al prezzo di equilibrio (*clearing*

price) pagano esattamente il *clearing price* e i fornitori che offrono meno o quanto il prezzo stabilito pagano la stessa cifra.

Al contrario nelle aste *pay-as-bid* gli offerenti pagheranno esattamente il prezzo che offrono per la quantità scambiata. In questo contesto si crea il problema dell'*extra money*, ossia della quantità pagata dagli acquirenti ma non pagata agli offerenti, a differenza di ciò che accade nelle aste uniformi nelle quali la quantità di denaro pagata dagli acquirenti eguaglia quella ricevuta dagli offerenti.

1.3 La struttura del mercato

Parlando di *asset* finanziari, si indica con prezzo *spot* il prezzo istantaneo di vendita dello strumento in esame. Nei mercati elettrici, però, il prezzo viene determinato per tutte le ore del giorno successivo ad una data ora del giorno precedente, il che significa che non è possibile in questo contesto parlare di prezzo *spot* nel senso classico del termine.

Inoltre, la nascita di un mercato elettrico autonomo a portato lo sviluppo di numerosi tipi di contratti innovativi, che vengono scambiati sia nei mercati regolamentati che *Over the Counter* (OTC).¹

È possibile distinguere tra contratti fisici e finanziari, ognuno dei quali è caratterizzato da una dimensione, un prezzo, un periodo e un luogo di consegna.

I contratti cosiddetti fisici possono essere ulteriormente classificati a seconda che siano a breve o a lungo termine. In particolare, tra questi ultimi, si distinguono i *Power Purchase Agreements* (PPA), che in molti paesi rappresentano la parte più consistente del mercato elettrico. In Polonia, ad esempio, costituiscono il 40% del totale della produzione.

Altri contratti a lungo termine di particolare interesse sono quelli relativi alle emissioni di CO₂. Già dal 2005 nel Nord Pool e nell'EEX (*European Energy Exchange*) hanno

¹ L'espressione Mercato *Over the Counter* (OTC) indica il complesso di operazioni di compravendita di titoli che non figurano nei listini ufficiali di borsa e che vengono effettuate da intermediari ed investitori al di fuori del mercato regolamentato.

cominciato ad essere diffusi questi contratti, scambiati prevalentemente OTC. Probabilmente in futuro questo mercato entrerà in una fase successiva di sviluppo che permetterà ai partecipanti di avere maggiori informazioni e più trasparenza, dato che, al momento attuale, coloro che operano in questi mercati devono accontentarsi di informazioni scarse e imperfette.

1.4 Il mercato dell'energia elettrica in Europa

1.4.1 Inghilterra e Galles

Il mercato dell'energia elettrica in Europa nasce proprio in seguito all'UK *Electricity Act* del 1989. La caratteristica che contraddistingue questo mercato, dominato da due delle compagnie fondatrici, è la stima dei prezzi per la giornata successiva a frequenza semi-oraria anziché oraria, come nel caso degli altri mercati.

Fino al 2001 il metodo per la determinazione del prezzo era il “contratto per le differenze” (CfD, *contracts for differences*), che specificava un prezzo *strike*² e un volume per gli scambi: se il prezzo stabilito dal mercato era più alto di quello del CfD il produttore pagava la differenza al consumatore; viceversa il consumatore corrispondeva la differenza al produttore.

Questo tipo di contrattazioni mostrava parecchi inconvenienti: innanzitutto scarsa trasparenza, poi inadeguatezza e ritardi nei pagamenti.

Nel marzo 2001 è stato introdotto il *New Electricity Trading Arrangements* che ha sostituito il vecchio metodo con un sistema bilaterale di scambi; immediatamente dopo si è anche avuta la nascita delle contrattazioni OTC degli *assets* legati al mercato elettrico.

² Prezzo *strike* o di esercizio: prezzo al quale l'acquirente o il venditore si riserva la facoltà di comperare o vendere un determinata quantità di attività finanziaria sottostante.

1.4.2 Il Nord Pool

Il mercato dei paesi del Nord Europa è stato il primo nel mondo ed è stato creato nel 1992. La formazione del prezzo avviene attraverso un'asta bilaterale a prezzo uniforme ad intervalli di tempo orari: alle 12 p.m. i partecipanti trasmettono all'amministratore del mercato i loro prezzi *ask* e *bid*³ per le successive 24 ore, dall'1 a.m. del giorno successivo. I partecipanti al mercato sono liberi di comprare o vendere l'intera giornata o singole ore.

1.4.3 L'Europa continentale

Nell'Europa continentale, la liberalizzazione ha avuto inizio nel 1997 con una direttiva che ha definito regole ben precise al fine di evitare il monopolio.

La Spagna per prima ha dato vita ad un mercato autonomo, nel quale le due maggiori compagnie producono e forniscono oltre l'80% dell'energia.

La Germania ha il mercato più vasto, ma caratterizzato da due problemi: la mancanza di una vera integrazione verticale e il fallimento del tentativo di creare un organo autonomo che funga da regolatore del sistema.

Nell'Europa dell'Est i mercati elettrici sono ancora in fase di sviluppo ed esistono tuttora barriere e difficoltà nei collegamenti. In questo contesto, la Polonia ha il più importante mercato dell'energia elettrica (PPA) che sta vivendo una fase di forte liberalizzazione, iniziata nel 1997 e che ha portato alla nascita di un mercato operativo nel 2000; al momento attuale, però, le strutture sono inappropriate e i costi molto elevati.

³ Con la dicitura di prezzo *Ask* si indica il valore al quale viene offerta una valuta o un qualunque altro strumento finanziario. È il prezzo al quale si può comprare inviando un ordine a mercato. L'espressione *Bid price*, invece, indica il prezzo proposto dall'acquirente di valute o di altri strumenti finanziari. Rappresenta il valore al quale è possibile vendere lo strumento specificato.

1.4.4 L'Italia

Per quanto riguarda l'Italia, dopo la direttiva comunitaria per la creazione di un mercato interno dell'energia, con il decreto legislativo n. 79 del 16 marzo 1999 è nato il mercato elettrico italiano, a sua volta articolato in un Mercato a Pronti (MPE) ed un Mercato Elettrico a termine (MTE).

Il MPE si articola a sua volta in:

- Mercato del giorno prima – MGP;
- Mercato dell'aggiustamento – MA ;
- Mercato del servizio di dispacciamento – MSD.

Il Mercato del Giorno Prima ospita la maggior parte delle transazioni di compravendita di energia elettrica e rappresenta il mercato in cui vengono scambiati i blocchi di energia per il giorno successivo. Sul MGP gli operatori presentano le offerte, indicando quantità e prezzo ai quali acquistare o vendere, secondo una logica di negoziazione ad asta e non continua. Il prezzo che viene determinato per ogni ora si ottiene dall'incontro tra domanda e offerta e differisce da zona a zona.

Il Mercato di Aggiustamento consente agli operatori di apportare modifiche ai programmi definiti nel MGP attraverso ulteriori offerte di acquisto o vendita. Anche sul MA le offerte di acquisto e vendita sono accettate sulla base dei criteri utilizzati nel MGP.

Il Mercato per il Servizio di Dispacciamento, infine, è lo strumento attraverso il quale avviene l'approvvigionamento delle risorse necessarie alla gestione e al controllo del sistema. Mentre nei primi due mercati è il GME (Gestore del Mercato Elettrico) a svolgere il ruolo di controparte centrale, nel MSD è Terna S.p.a a svolgere questo compito.

Il Mercato elettrico a termine, invece, è la sede per la negoziazione di contratti a termine dell'energia elettrica con obbligo di consegna e ritiro. A differenza del MGP, col quale ha in comune alcuni operatori, sul mercato a termine la negoziazione avviene in modalità continua; si possono negoziare sia contratti *base-load* che *peak-load* e gli operatori presentano le loro proposte indicando tipologia, numero e periodo di consegna dei contratti stessi e il prezzo al quale acquistare o vendere energia. Sul MTE si possono registrare anche contratti conclusi OTC e il GME svolge il ruolo di controparte centrale.

1.4.5 La Francia

La distribuzione dell'energia elettrica in Francia è stata monopolio esclusivo dell'azienda *Électricité de France* (EDF) fino al 1999, quando venne recepita la prima direttiva della Commissione Europea per la regolamentazione del mercato dell'elettricità.

La decisione di aprire il mercato alla concorrenza a livello europeo avvenne attraverso due direttive, quella del 19 Dicembre del 1996 e quella del 26 Giugno 2003. I passaggi fondamentali della nascita del mercato in Francia furono i seguenti:

- 19 febbraio 1999: 20% del mercato : clienti che impiegano più di 100 GWh.
- 29 maggio 2000: 30% del mercato : clienti che impiegano più di 16 GWh.
- 10 febbraio 2003: 37% del mercato: clienti che impiegano più di 7 GWh.
- 1 luglio 2004: 70% del mercato: aziende e industrie (1,3 milioni di clienti).
- 1 luglio 2007: apertura totale.

Il 26 Luglio 2001 nacque ufficialmente il *Paris Powernext* (PPX). Il mercato dell'energia elettrica in Francia si compone di diversi mercati energetici complementari, trasparenti ed anonimi:

- *Powernext Day-Ahead* sul quale ogni giorno, alle 11:00 vengono stabiliti i prezzi per le 24 ore del giorno successivo;
- *Powernext Futures* sul quale vengono contrattati i *futures*⁴ con scadenze comprese tra uno e tre mesi;
- *Powernext Balancing GRTgaz* che permette a tutti gli operatori del sistema di coprire il loro fabbisogno energetico giornaliero.

Dalla fine del 2008 sono stati introdotti anche due mercati, *Powernext Gas Spot* e *Powernext Gas Futures*, sui quali avvengono le contrattazioni dei prezzi spot e *futures* per il gas.

Per quanto riguarda le fonti energetiche, attualmente l'energia elettrica viene prodotta in Francia attraverso la combustione di carbone, gas naturale o petrolio per il 30%, attraverso l'utilizzo di energia nucleare per il 51% e sfruttando l'energia idrica per il restante 19%.

⁴ *Futures*: il termine indica contratti a termine attraverso i quali due controparti si obbligano a scambiarsi a scadenza una quantità di una determinata attività finanziaria ad un prezzo stabilito al momento della stipula.

1.5 Il mercato dell'energia elettrica in America del Nord

Il sistema elettrico negli Stati Uniti si è evoluto in tre reti principali: l'*Eastern Interconnected System* (che copre la zona centro-orientale del paese), il *Western Interconnected System* (che rifornisce la zona del Pacifico) e il *Texas Interconnected System*. Nella maggior parte delle aree, però, non si sono creati mercati competitivi indipendenti: ciò è accaduto solo in New England, New York, Midwest, Pjm(Pennsylvania-New Jersey-Maryland) e California che hanno una struttura di trasporti collegata, consistente in un mercato orario e/o giornaliero e di un mercato in tempo reale. In seguito al *blackout* dell'Agosto 2003, molti stati hanno deciso di non effettuare più sforzi di deregolamentazione e il volume di elettricità coinvolta nei mercati è cresciuto rapidamente.

In Canada ci sono differenze consistenti tra le varie province, ognuna delle quali ha un proprio regolatore; solo due province, Alberta ed Ontario, hanno stabilito un mercato con quotazioni in tempo reale.

1.6 Conclusioni

La complessità del mercato dell'energia elettrica oggi è enorme e questo, ovviamente, rende impossibile definire il tipo migliore di mercato. Infatti, le caratteristiche tecniche ed economiche, così come le regolamentazioni giuridiche e politiche, che contribuiscono alla complessità del mercato, fanno sì che non ci sia una formula unica per ottenere un mercato funzionante. L'unica caratteristica necessaria per garantire trasparenza a tutti coloro che partecipano al mercato, quale che sia il tipo di regolamentazione applicata, è la determinazione di un criterio unico e formale per la definizione del prezzo dell'energia.

La principale motivazione, infatti, che ha portato alla nascita e alla regolamentazione del mercato, è stata la necessità di garantire forniture di energia elettrica, in condizioni di neutralità e trasparenza, garantendo la concorrenza tra competitori.

Per maggiori informazioni riguardo all'evoluzione e alla struttura dei mercati si rimanda a Weron (2006).

Capitolo 2

Le principali caratteristiche di *loads*⁵ e prezzi dell'energia elettrica

I prezzi e i *loads* dell'energia elettrica vengono prodotti 24 ore al giorno, 365 giorni all'anno e questo permette di applicare a questi dati metodi statistici che non possono essere utilizzati in nessun altro contesto. Nel seguito vengono descritte le principali caratteristiche che rendono questo tipo di serie così particolari.

2.1 *Jumps o Spikes*

La caratteristica più evidente dei prezzi dell'energia elettrica è la presenza di bruschi e repentini movimenti verso l'alto; questi “salti” vengono detti *jumps* o *spikes*. Un *jump* o *spike* si verifica quando il prezzo supera una specifica soglia per un breve periodo di tempo. Questa ed altre peculiarità rendono i prezzi associati all'energia elettrica estremamente volatili rispetto a quelli degli altri beni. Questi picchi, inoltre, non sono omogeneamente distribuiti nel tempo: sono, infatti, maggiormente diffusi tra le 09:00 e le 18:00 dei giorni lavorativi, nelle cosiddette *on-peak hours*.

⁵ Qui e nel seguito si preferisce utilizzare il termine *load* invece del corrispettivo in italiano, per indicare l'ammontare di energia elettrica destinata alla distribuzione.

La presenza di picchi nei prezzi dell'elettricità è dovuta al fatto che l'elettricità non è un bene accumulabile e deve essere utilizzata nel momento in cui viene prodotta.

In genere, comunque, i picchi che si riscontrano sono di breve durata e i prezzi tornano rapidamente alla normalità.

In alcuni mercati, inoltre, la presenza di picchi è praticamente irrilevante, come ad esempio in Polonia.

Nonostante non siano molto frequenti, i *jumps* rappresentano una ragione valida per assicurarsi contro i loro movimenti improvvisi e anche una ragione per inserire componenti discontinue per modellare realisticamente la dinamica dei prezzi.

L'impatto che le fluttuazioni nei *loads* di energia hanno sui picchi dipende dall'incontro tra domanda e offerta: l'offerta dipende dalla disponibilità di risorse e dai costi marginali, mentre la domanda ha un comportamento stagionale dovuto soprattutto alle condizioni meteorologiche e che si mostra non uniforme durante la settimana. Poiché il prezzo *spot* si ottiene dall'intersezione tra domanda e offerta, è molto sensibile alle loro fluttuazioni. In particolare, se la domanda è alta, piccole variazioni del consumo, provocheranno repentini cambiamenti nel livello di prezzo.

Nella modellazione, è necessario considerare la presenza di questi valori anomali, poiché hanno impatto significativo sia sulla volatilità, che sul livello medio di prezzi. Proprio per questo motivo, nel seguito di questo lavoro si è deciso di sottoporre ad analisi sia i dati originari, che quelli ripuliti dei valori che superano una soglia stabilita.

2.2 Stagionalità

Le serie dei prezzi dell'energia evidenziano una forte componente stagionale legata ai cambiamenti del clima ed in particolare al numero di ore di luce e alle temperature registrate. È, quindi, necessario tener conto, oltre che dell'usuale stagionalità mensile, anche di periodicità settimanali e giornaliere nel costruire modelli per i prezzi. All'interno della giornata si osservano, infatti, prezzi molto elevati alla mattina e alla sera (ossia quando la gente si sveglia e va al lavoro e quando rientra a casa) ed inferiori

alla media a metà giornata e durante la notte. La variabilità infra-settimanale è legata soprattutto alle differenze nei consumi tra i giorni lavorativi e i fine settimana, in cui le aziende sono chiuse.

Le serie dei prezzi evidenziano, ovviamente, caratteristiche differenti a seconda delle stagioni: durante l'inverno le aziende lavorano a pieno regime e la gente passa molto tempo in casa; al contrario, durante l'estate; in particolare nel mese di Agosto, i consumi delle imprese si riducono a causa della chiusura estiva, mentre quelli dei privati sono legati all'aumento della temperatura e all'utilizzo dei condizionatori. Quindi, è necessario cercare di capire che forma di correlazione mostrino i dati, partendo dall'analisi delle funzioni di autocorrelazione.

Quello che si nota nei prezzi dell'energia è la forte dipendenza soprattutto per osservazioni che distano 7 giorni. Inoltre, utilizzando dati orari, è possibile osservare nei dati anche una componente stagionale di periodo 24. Per approfondimenti e grafici si veda il Capitolo 4, relativo alle analisi preliminari dei dati.

A partire da una serie storica, Y_t , si può considerare la sua decomposizione in tre componenti, che si possono combinare in modo additivo o moltiplicativo: trend o trend-ciclo, T_t , stagionalità vera e propria, S_t e componente erratica residua, ε_t .

Una delle operazioni che spesso si applicano ad una serie storica al fine di ripulirla da trend o stagionalità è la differenziazione. In particolare si applica la differenza prima la fine di eliminare un trend lineare e di rendere stazionaria la serie. Più in generale, l'applicazione di una differenza di ordine k elimina un trend polinomiale del medesimo grado. Spesso la differenziazione viene applicata in combinazione con la procedura di lisciamiento attraverso medie mobili. Anche dopo l'applicazione delle procedure citate è possibile che i dati mostrino ancora una componente ciclica residua, legata alla periodicità del mercato. Spesso si preferisce eliminarla, utilizzando sinusoidi o *splines*, prima di procedere ad ulteriori esami dei dati.

Nel seguito si sceglierà di tenere in considerazione la periodicità delle serie dei prezzi elettrici utilizzando variabili *dummy* che indichino l'ora, il giorno della settimana e che individuino anche particolari festività in corrispondenza delle quali si potrebbero osservare valori più bassi del livello dei prezzi. Inoltre, al fine di ripulire i dati della componente ciclica si inserirà una *splines* cubica. Per maggiori approfondimenti sull'argomento si vedano le sezioni successive o si consulti Weron (2006).

2.3 Mean Reversion

I prezzi *spot* dell'energia sono caratterizzati dalla cosiddetta *mean reversion*, letteralmente "ritorno alla media". Si osserva, infatti, di frequente che, in seguito a picchi improvvisi, i prezzi ritornano piuttosto rapidamente al precedente livello. La velocità con cui ci ritornano dipende da diversi fattori, quali, ad esempio, il tipo di risorsa energetica considerata o i mezzi attraverso i quali avviene la distribuzione dell'energia. Inoltre, i prezzi dell'energia elettrica sono influenzati dal clima, che è caratterizzato da un andamento ciclico e *mean-reverting*, quindi la tendenza a tornare alla media influenzerà anche il livello di domanda e di equilibrio dei prezzi. In economia e finanza lo studio della presenza di memoria lunga nei dati è tuttora molto attivo, poiché molti dati mostrano caratteristiche indicative della presenza di una forte persistenza, non compatibile con la non stazionarietà. La presenza di memoria lunga, però, non permette di applicare alcuni degli strumenti standard utilizzati per l'analisi dei dati. Per cercare di misurare il livello di *mean reversion* presente nei dati sono state sviluppate quattro tecniche:

- *R/S analysis*, o metodo dei *range* riscaldati, che permette di calcolare il parametro di *self similarity* H , che misura l'intensità della dipendenza;
- *Detrended Fluctuation Analysis*, che rispetto alla *R/S analysis* riesce ad evitare di cogliere correlazioni di lungo termine spurie, causate dalla presenza di non stazionarietà;
- *Periodogram Regression*, procedura semi-parametrica che consente la stima del parametro di integrazione d ;
- *Average Wavelet Coefficient*, che utilizza la trasformazione *wavelet* (ondata) per la stima dell'esponente di Hurst H .

Per maggiori dettagli sulle procedure citate si veda Weron (2006).

2.4 Distribuzioni di riferimento per i prezzi dell'elettricità

I rendimenti di *assets* finanziari non risultano praticamente in nessun caso compatibili con le caratteristiche di una distribuzione normale. Tenzialmente, infatti, mostrano un eccesso di curtosi: si dice che provengano da una distribuzione a code pesanti o leptocurtica. È, quindi, necessario individuare una distribuzione che possa cogliere al meglio queste caratteristiche dei rendimenti.

Ci si chiede quindi se anche per i prezzi dell'energia elettrica possa essere opportuno utilizzare distribuzioni a code pesanti e quale di queste distribuzioni si adatti meglio ai dati.

Di seguito si propongono le due famiglie di distribuzioni a code pesanti più utilizzate in letteratura: le distribuzioni stabili e le distribuzioni iperboliche generalizzate. Nell'analisi pratica dei dati che si presenterà nei capitoli successivi, tuttavia, si manterrà l'ipotesi di compatibilità dei prezzi con una distribuzione normale.

2.4.1 Distribuzioni stabili

Le distribuzioni stabili hanno caratteristiche che le rendono ideali per modellare i rendimenti di titoli finanziari: innanzitutto sono leptocurtiche; in secondo luogo possono essere utilizzate come distribuzione di riferimento per il teorema centrale del limite generalizzato. La legge stabile o α -stabile richiede 4 parametri per una completa descrizione. Il parametro α indica la velocità con cui si assottigliano le code della distribuzione. Se $\alpha=2$ si ottiene la distribuzione normale. Se $\alpha>2$, invece, la varianza risulta infinita e le code della distribuzione sono asintoticamente equivalenti a quelle di una distribuzione di Pareto. Il parametro β , invece, indica la simmetria della distribuzione ed assume valore positivo se c'è asimmetria positiva. Infine $\sigma>0$ e μ rappresentano, rispettivamente i parametri di scala e locazione della distribuzione. In genere è conveniente descrivere questa famiglia di distribuzioni attraverso la funzione

caratteristica perché la funzione di densità non ha una forma chiusa per tutti i valori dei parametri.

2.4.2 Distribuzioni iperboliche

La famiglia delle distribuzioni iperboliche deve il suo nome al fatto che il logaritmo della funzione di densità forma un'iperbole. La distribuzione iperbolica si ottiene da una mistura di normali; in particolare una variabile Z ha distribuzione iperbolica se

$$(Z/Y) \sim N(\mu + \beta Y, Y)$$

con Y che indica una variabile casuale Gaussiana inversa generalizzata. La distribuzione iperbolica è descritta da quattro parametri, μ il parametro di locazione, δ parametro di scala, α e β parametri che descrivono la forma della distribuzione. Facendo variare i parametri è possibile anche ottenere distribuzioni a code pesanti che ben si adattano a dati con caratteristiche descritte in precedenza. In particolare, in genere, si utilizza il metodo della massima verosimiglianza per ottenere stime dei parametri. La distribuzione iperbolica fa parte di una classe più estesa di distribuzioni a cui appartengono anche la distribuzione gaussiana inversa e le distribuzioni gamma.

2.5 Conclusioni

Le caratteristiche specifiche dei prezzi dell'energia elettrica, che la contraddistinguono dalle altre *commodities*, ed in particolare la diffusa presenza di valori anomali, portano a costruire modelli basati su distribuzioni a code pesanti e che possano prendere in considerazione l'elevata volatilità dei dati, spesso decisamente maggiore che negli altri *assets* finanziari. Inoltre, i prezzi dell'energia elettrica mostrano un'evidente stagionalità sia settimanale che giornaliera, fortemente persistente nel tempo.

Infine è possibile notare che anche quando i prezzi dell'energia elettrica mostrano valori anomali, il ritorno al livello medio avviene più rapidamente di quanto accada per i prezzi di altre fonti energetiche, quali il gas naturale o il petrolio.

L'analisi e la comprensione delle particolarità dei prezzi dell'energia elettrica è fondamentale per stabilire il tipo di modello o la tecnica statistica più adatta a rappresentare i dati. A tale scopo sono state presentate le caratteristiche che contraddistinguono le serie dei *loads* e degli *spot*, prima di introdurre gli strumenti più utilizzati per modellare dati reali.

Capitolo 3

Modelli per i prezzi dell'energia elettrica

La caratteristica che contraddistingue il mercato dei titoli elettrici è la simultaneità della quotazione dei prezzi per le 24 ore della giornata successiva. In questo contesto è possibile pensare a due approcci distinti per modellare i dati relativi ai prezzi dell'energia elettrica e poter, quindi, fare previsioni. Il primo, più tradizionale, considera i dati in forma di serie storica e si basa sull'applicazione di modelli per le serie storiche, introducendo varianti che tengano in considerazione le caratteristiche proprie dei prezzi dell'energia elettrica.

Il secondo, invece, innovativo e più recente, considera i dati sotto forma di *panel*, supponendo che ogni ora rappresenti un'unità statistica osservata per diversi istanti temporali. Questo tipo di approccio è stato considerato da Huisman et al. (2006).

Di seguito si presentano i due approcci citati, considerando che per l'approccio classico si concentrerà l'attenzione su metodi statistici e quantitativi.

3.1 Modelli per serie storiche

Negli ultimi anni si sono sviluppati differenti approcci per l'analisi del comportamento dei mercati dell'energia, che tenessero in considerazione tutte le caratteristiche fin qui osservate.

In particolare si ricordano:

- modelli basati sui costi che puntano a minimizzare i costi nella generazione dell'energia;
- approcci di equilibrio, in cui i prezzi si producono cercando di ottenere l'equilibrio sui mercati;
- metodi strutturali, che descrivono le dinamiche dei prezzi modellando l'impatto di eventi fisici o economici sui prezzi dell'elettricità;
- modelli quantitativi (o stocastici o econometrici), utilizzati per il *pricing* degli strumenti finanziari derivati e per la gestione del rischio;
- approcci statistici, che consistono nell'applicazione di tecniche statistiche alla previsione dei prezzi elettrici;
- tecniche basate sull'"intelligenza artificiale", che modellano i prezzi usando reti neurali e modelli *fuzzy logic*.

Nei paragrafi successivi si concentra l'attenzione sugli approcci di tipo statistico, che hanno avuto largo sviluppo, e sui metodi di tipo quantitativo, utilizzati nel mercato dei derivati.

3.1.1 Metodi statistici per la previsione dei prezzi dell'energia elettrica

Per costruire modelli statistici utili ai fini previsivi nel contesto di mercati elettrici, è necessario considerare tre problemi legati a questa tematica.

In primo luogo, l'accuratezza previsiva è strettamente legata non solo all'efficienza degli algoritmi utilizzati, ma anche alla qualità dei dati analizzati e alla capacità di considerare correttamente i fattori esogeni che entrano nel modello. In particolare, le fluttuazioni nei *loads* di energia, il giorno della settimana e l'ora di riferimento e il costo variabile della generazione dell'energia sono fattori esterni che hanno un impatto significativo sulle variazioni di prezzi.

In secondo luogo, come si è evidenziato nel paragrafo 2.1, i prezzi dell'energia elettrica sono spesso caratterizzati dalla presenza di valori anomali, i cosiddetti *spikes*, che condizionano pesantemente l'accuratezza previsiva del modello.

Infine, una volta effettuate le previsioni, è necessario capire se queste siano accurate o meno. A tale scopo è possibile calcolare diversi indicatori della bontà previsiva, come ad esempio il MAE (*mean absolute error*), che calcola la media delle differenze in valore assoluto tra i valori effettivamente realizzati e quelli previsti, il MSE (*Mean Squared Error*), che calcola la media degli errori al quadrato fra le stesse differenze, oppure il MAPE (*Mean Absolute Percentage Error*) che calcola la media delle differenze in valore assoluto divise per il prezzo realizzato.

Fatte queste considerazioni, è opportuno notare che per le serie dei prezzi è interessante stimare tanto l'andamento medio quanto la volatilità. A tale proposito si introducono modelli ARMA (con la variante ARMAX con variabili esogene) per la media, e modelli GARCH per la varianza.

3.1.1.1 Modelli ARMA e ARMAX

I modelli più utilizzati per tenere in considerazione la presenza di dipendenza nei dati, sono i cosiddetti modelli ARMA (*AutoRegressive Moving Average*), che esprimono il prezzo corrente P_t in funzione dei suoi p valori passati e di q valori di una componente erratica. Si può, dunque, scrivere:

$$\phi(B)P_t = \mu + \theta(B)\varepsilon_t.$$

con $\varepsilon_t \sim \text{WN}(0, \sigma^2)$. Se $q=0$ si ottengono modelli autoregressivi AR(p).

I modelli ARMA si basano sull'ipotesi che le serie storiche alle quali vengono applicate siano debolmente stazionarie; nel caso in cui questa ipotesi non valga, è possibile, con una semplice trasformazione, ricondursi al caso di stazionarietà debole. Si ottengono così i modelli ARIMA o SARIMA (I sta per *Integrated*, integrati).

Al fine di considerare la possibilità che sul livello medio del prezzo dell'elettricità abbiano effetto anche variabili esogene è possibile costruire modelli ARMAX, nei quali P_t è funzione anche di input legati al clima o alla relazione prezzi-loads di energia.

La forma compatta di un modello ARMAX è la seguente:

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

dove v^1, v^2, \dots, v^k rappresentano le k variabili esogene. Anche in questo contesto, è possibile considerare modelli da applicare a serie non stazionarie e che presentino una componente di stagionalità nella media, ottenendo modelli ARIMAX semplici e stagionali.

3.1.1.2 Modelli GARCH

I modelli finora descritti per spiegare la presenza di autocorrelazione nelle serie dei prezzi ipotizzano omoschedasticità condizionale; questa ipotesi, però, non è supportata dall'evidenza empirica dei dati, che spesso mostrano una dipendenza non lineare dal loro passato. In particolare, spesso si rileva la presenza di *cluster* di volatilità che rendono necessario inserire una componente che tenga conto dell'eteroschedasticità.

I modelli che sono stati costruiti al fine di modellare l'eteroschedasticità condizionale sono stati proposti da Engle nel 1982 e sono noti come modelli ARCH (*AutoRegressive Conditional Heteroskedastic*); questi in seguito sono stati generalizzati da Bollerslev nel 1986 ai modelli GARCH (*Generalized AutoRegressive Conditional Heteroskedastic*).

Supponendo di avere una serie y_t , è possibile scomporla nel seguente modo:

$$y_t = \mu_t + v_t = \mu_t + \sigma_t \varepsilon_t$$

con μ_t che rappresenta la media e $\varepsilon_t \sim N(0,1)$.

Nei modelli di tipo ARCH la varianza assume la seguente forma:

$$\sigma_t^2 = \alpha_0 + \sum_{i=0}^q \alpha_i h_{t-i}^2,$$

con $h_t = \varepsilon_t \sigma_t$.

Per generalizzare modelli di questo tipo, si può far dipendere σ_t^2 anche da suoi ritardi passati, ottenendo l'espressione seguente:

$$\sigma_t^2 = \alpha_0 + \sum_{i=0}^q \alpha_i h_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

Al fine di garantire la positività della varianza è necessario imporre dei vincoli sui parametri del modello, per cui $\alpha_i, \beta_j > 0$ e $\alpha_0 > 0$.

È immediato comprendere che nel caso in cui si voglia modellare sia la media che la varianza condizionale per la serie di prezzi è possibile applicare modelli misti ARMAX-GARCH, che combinano tutte le specificazioni viste finora.

Molto spesso, soprattutto nel caso di dati finanziari, è possibile che siano presenti processi che mostrano una persistenza forte nel tempo, ma che sono stazionari, sia in media che in varianza. Si tratta dei cosiddetti modelli frazionalmente integrati (*fractionally integrated*) a memoria lunga, dei quali fanno parte i modelli ARFIMA(p,d,q), introdotti Hosking (1981) e da Granger and Joyeux (1980).

In questi modelli il parametro di integrazione d può assumere qualsiasi valore reale e prende il nome di parametro a memoria lunga. Analogamente, anche per la varianza condizionata si possono costruire modelli FIGARCH quando la persistenza della dipendenza nei quadrati o nei valori assoluti dei prezzi è alta, ma non tale da giustificare non stazionarietà, che si ha nel caso in cui

$$\sum_{i=0}^q \alpha_i + \sum_{j=1}^p \beta_j = 1$$

Molto spesso, in letteratura, si veda Weron (2006) e Spagnolo ed Uddin (2006), prima di applicare modelli come quelli descritti ai prezzi dell'energia elettrica, si è ritenuto opportuno eliminare i dai anomali (i picchi o *spikes*) al fine di migliorare le capacità previsive del modello. Ciò che si evidenzia dalle analisi dei risultati ottenuti è che lavorare su serie depurate da *outliers* dà risultati soddisfacenti solo in periodi in cui l'andamento dei prezzi è caratterizzato da bassa volatilità. Quando la volatilità cresce, l'eliminazione degli *spikes* non è ottimale. Nel seguito si presenteranno tre approcci diversi. Si lavorerà sulla serie originaria dei prezzi, sulla serie depurata dei valori che superano una soglia prestabilita e sulla serie originaria per la quale la componente deterministica è stata stimata senza *spikes*.

3.1.1.3 Modelli *Regime Switching*

La presenza di picchi che si osservano analizzando i prezzi dell'energia elettrica suggerisce la possibilità che esista un meccanismo di *switching* tra due regimi, uno

normale ed uno a prezzi più elevati. Su questa considerazione si basa la costruzione di modelli *regime-switching*.

Si possono distinguere due classi di modelli di questo tipo: quelli per i quali il regime può essere determinato da una variabile osservabile (e quindi è possibile sapere con certezza quali regimi si sono realizzati nel passato e nel presente) e quelli per i quali il regime è determinato da una variabile non osservabile, per i quali si può solo stimare la probabilità che si sia in un regime piuttosto che in un altro.

I due modelli, uno della prima classe e l'altro della seconda, che vengono maggiormente applicati ai dati dei prezzi dell'energia sono i modelli a soglia TAR (*Threshold AutoRegressive*) e i modelli *Markov-Switching*, che verranno approfonditi ulteriormente nel seguito.

Per quanto riguarda i modelli TAR, introdotti per la prima volta da Tong nel 1978, i due regimi sono specificati dal valore di una variabile osservabile v_t e da una relativa soglia T .

Per il prezzo P_t si può quindi scrivere:

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

con $\phi(B)$ che indica un polinomio di grado p nell'operatore ritardo B . Il modello così scritto è nella forma più semplice, in cui si considerano solo due regimi e un semplice modello $AR(p)$. È possibile complicare il modello introducendo variabili esplicative esogene (TARX) oppure aumentando il numero di regimi. Nel caso in cui la soglia sia un valore ritardato del prezzo stesso, si ottengono i modelli SETAR (*Self-Exciting TAR*).

Del secondo tipo di modelli, in cui il regime è determinato da una variabile non osservabile, fanno parte i modelli MRS (*Markov regime-switching*), introdotti da Hamilton nel 1989. L'idea di fondo è quella di modellare il comportamento di una serie storica caratterizzata da due o più regimi provenienti da un differente processo generatore dei dati.

Il regime di provenienza dei dati è determinato da una variabile latente non osservabile. Si assume che i parametri del processo generatore dei dati possano cambiare per un certo periodo di tempo per poi tornare ai loro valori originari. Se si ipotizza che ci siano solo due regimi, si potrà costruire una matrice di transizione, contenente le probabilità di

passaggio da un regime all'altro e di permanenza nel regime di partenza. Tale matrice, denominata \mathbf{P} , sarà di dimensione 2×2 e avrà una forma del tipo:

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}$$

Ovviamente, nel caso di k regimi la matrice \mathbf{P} sarà di dimensione $k \times k$.

Il regime di provenienza dei dati dipende dal regime dell'istante precedente:

$$P(R_t = j | R_{t-1} = i, R_{t-2} = k, \dots) = P(R_t = j | R_{t-1} = i).$$

Quindi la probabilità di stare nello stato j in $t+m$, se in t si era nello stato i si ottiene come:

$$P(R_{t+m} = j | R_{t-1} = i) = (\mathbf{P}')^m \cdot e_i$$

dove \mathbf{P}' indica la trasposizione della matrice \mathbf{P} ed e_i è l' i -esima colonna della matrice identità 2×2 .

Ethier e Mount, nel 1998, applicarono per la prima volta i modelli *Markov Switching* alle serie dei prezzi dell'energia elettrica, proponendo un modello a due regimi, entrambi governati da processi AR(1), con media e varianza diversa.

Una possibile estensione venne proposta da Huisman e Mahieu (2003) che costruirono un modello a tre regimi, il primo di base, il secondo di picco, per gli improvvisi cambi di livello, e un terzo regime per descrivere il modo in cui i prezzi ritornano al livello base dopo un *jump*. La matrice di transizione associata a questo modello ha la seguente forma:

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

che non permette la presenza di *jumps* consecutivi, come invece si osserva spesso nelle serie dei prezzi dell'energia elettrica.

Per superare questo limite Huisman e Jong (2003) proposero un modello con due regimi, uno rappresentato da un processo AR(1) stabile con ritorno alla media e un regime per i picchi, assumendo che quest'ultimo potesse semplicemente essere modellato da una distribuzione normale con media e varianza più alte di quelle del regime base. In formule il modello è il seguente:

- regime *mean-reverting*: $\ln S_{M,t} = \ln S_{M,t-1} + \alpha(\mu - \ln S_{M,t-1}) + \varepsilon_{M,t}, \varepsilon_{M,t} \sim N(0, \sigma_M)$

- regime di picco: $\ln S_{S,t} = \mu_S + \varepsilon_{S,t}$ $\varepsilon_{S,t} \sim N(0, \sigma_S)$

dove S_t è il prezzo spot che assume il livelli $S_{S,t}$ e $S_{M,t}$ nei due regimi.

Weron e Bierbrauer (2004) estesero il modello presentato introducendo l'uso di distribuzioni log-normali e di Pareto per il regime dei picchi, mentre Mount (2006) rese ulteriormente flessibile il modello esprimendo i parametri come funzioni di variabili come i *loads* o i margini di riserva, che evolvono nel tempo.

Confrontando questo tipo di modelli con quelli tradizionali per l'analisi di serie storiche, come hanno fatto Kosater e Mosler (2006) si nota come gli MRS non siano ottimali in termini previsivi nei periodi di calma, in assenza di *outliers*. Nei periodi i cui si concentrano maggiormente gli *spikes*, invece, i modelli *Markov-Switching* non lineari danno risultati significativamente migliori e permettono di ottenere previsioni adeguate. Nel contesto in esame, in particolare, a causa della presenza di valori anomali, si potrebbe pensare di utilizzare distribuzioni a code pesanti, scegliendo un numero di regimi adeguato alle caratteristiche evidenziate dai dati.

3.1.2 Metodi quantitativi per il *pricing* dei derivati

Spesso, in ambito finanziario si applicano ai dati modelli quantitativi (definiti anche come stocastici o econometrici) ai dati, cercando di replicare le proprietà statistiche dei prezzi dell'energia, al fine di fare valutazioni sui prodotti finanziari derivati.

In questo contesto assume importanza rilevante il premio al rischio, poiché i derivati sono nati proprio al fine di tutelarsi dai rischi legati al mercato e alla volatilità dei sottostanti.

A tale scopo è possibile valutare due tipologie di modelli: i modelli *jump-diffusion* e la loro estensione, rappresentata dai modelli ibridi, che tuttavia non saranno considerati in questo lavoro. In questo contesto vengono spesso utilizzati anche i modelli MRS discussi nel paragrafo 3.1.1.3

3.1.2.1 Modelli *Jump-Diffusion*

I modelli *jump-diffusion* si basano sugli studi pubblicati da Merton nel 1976, nei quali si assume che i prezzi seguano la dinamica così definita:

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

dove μ è il *drift*, σ la volatilità, W_t è un processo Browniano (o di Wiener) e $q(t)$ è un processo di Poisson omogeneo (HPP), che definisce le caratteristiche della distribuzione dei picchi.

È evidente come questo tipo di modelli siano ottenuti semplicemente aggiungendo la componente $dq(t)$ ad un processo Browniano Geometrico standard.

In realtà così come è stato definito il modello non prende in considerazione la *mean reversion*. Al fine di considerare questa caratteristica è possibile fare due estensioni:

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

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

La seconda specificazione (il cosiddetto modello *mean reverting* geometrico) sembra dare i risultati migliori in termini di adattamento ai dati. Nella forma in cui è stato espresso il modello, il parametro β rappresenta il tasso di *mean reversion* ed indica la velocità con cui i dati ritornano al livello base dopo un picco.

Questo tipo di modello presenta un inconveniente: assume che il processo di diffusione *mean-reversion* sia indipendente dalla componente di Poisson. Ma questa caratteristica non è supportata dall'evidenza empirica per quanto riguarda i dati dell'energia elettrica. Per risolvere questo inconveniente è stato proposto un modello in cui la dimensione del *jump* è proporzionale al livello di prezzo corrente.

Inoltre, la presenza di stagionalità nei dati suggerisce la possibilità di utilizzare processi di Poisson non-omogenei.

3.1.2.2 Modelli Ibridi

I modelli descritti nella sezione precedente non prendono in considerazione l'impatto di variabili esogene sulla formazione dei prezzi. A tale scopo si può pensare di introdurre nel modello parametri, legati all'ordine di grandezza e all'intensità dei salti, che siano funzione dei *loads*, della capacità di generazione di energia o del margine di riserva. I modelli di questo tipo vengono definiti ibridi e sono assimilabili ad i modelli con variabili esogene discussi in precedenza.

Un modello ibrido molto semplice è stato proposto da Barlow nel 2002: si tratta di un modello con ritorno alla media per la domanda e una funzione fissata per l'offerta.

A causa del comportamento non lineare dell'offerta, il modello *mean reverting e jump diffusion* permette di produrre picchi nella traiettoria simulata. A causa di ciò questo tipo di modello dà risultati migliori dei modelli più classici con ritorno alla media.

È possibile complicare la struttura incorporando anche modelli SARIMA per i *loads* o costruendo più modelli per le diverse caratteristiche da considerare e come la temperatura, i prezzi del carburante.

Al fine di rendere il modello più realistico si incorporano sempre più fattori e ciò lo rende sempre più complesso soprattutto a livello computazionale. Questi inconvenienti limitano l'utilizzo di modelli ibridi nel contesto dei prezzi dell'energia elettrica, poiché è sempre necessario cercare un compromesso tra parsimonia e realismo del modello.

3.2 Modelli per dati di *panel* e multivariati

Quando si parla di dati di *panel* ci si riferisce alla situazione in cui si hanno più unità statistiche, $i=1, \dots, N$, osservate ripetutamente nel tempo in t istanti differenti, $t=1, \dots, T$.

In genere la dimensione sezionale predomina su quella temporale, ossia si hanno molte unità statistiche (N grande) osservate in pochi istanti temporali.

Nel contesto che si sta esaminando, invece, si considerano 24 (o 48) unità statistiche, che rappresentano i prezzi per le ore (mezz'ore) della giornata. Le unità statistiche, poi,

vengono osservate per periodi di tempo piuttosto lunghi, dal momento che per ogni ora si hanno 365 osservazioni all'anno. È possibile considerare le 24 ore del giorno come unità statistiche che variano nel tempo poiché i prezzi per una giornata vengono stabiliti simultaneamente il giorno precedente.

L'approccio che si può considerare in questo contesto è quello dei modelli SUR (*Seemingly Unrelated Regression*), che si possono applicare solo nel caso in cui N sia minore di T .

I modelli SUR permettono di stimare sistemi di equazioni, nel caso in esame saranno N , assumendo che i termini d'errore di equazioni consecutive possano essere tra loro correlati, mentre sono incorrelati temporalmente.

Nel caso dell'energia elettrica ha senso fare un'assunzione di questo tipo in quanto, come già specificato in precedenza, i prezzi vengono stabiliti simultaneamente ad una data ora del giorno per le 24 ore del giorno successivo.

Un ulteriore approccio che si può prendere in considerazione per considerare le 24 serie orarie simultaneamente è quello dell'applicazione di modelli multivariati di tipo VAR (*Vector Autoregression*). Il VAR è un modello econometrico che rappresenta la generalizzazione dei modelli autoregressivi al caso in cui si vogliono considerare le interdipendenze tra più serie storiche. Il modello VAR(p), in realtà, può essere anche visto semplicemente come un'estensione del SUR al caso in cui tutte le variabili ritardate siano regressori comuni alle diverse equazioni.

3.2.1 Modelli SUR e loro stima

Come detto nel paragrafo precedente, i modelli *Seemingly Unrelated Regression*, introdotti già negli anni '60 da Theil e Zellner, sono utilizzati quando è necessario stimare sistemi di N equazioni tenendo conto della possibilità che errori di equazioni successive siano correlati.

In particolare, è possibile scrivere il modello nel seguente modo:

$$y_i = X_i \beta_i + \varepsilon_i$$

dove y_i ed ε_i sono vettori di dimensione $T \times 1$, β_i ha dimensioni $k \times 1$ e X_i è una matrice $T \times k$. Riguardo agli errori, le condizioni sopra presentate possono essere formalizzate assumendo che $E[\varepsilon_{it}, \varepsilon_{js}] = \sigma_{ij}$ se $i \neq j$, 0 altrimenti. Inoltre, $E[\varepsilon\varepsilon'] = V = \Sigma \otimes I_T$, dove Σ , è la matrice diagonale di dimensione $N \times N$, che contiene gli elementi σ_{ij} .

Se la matrice V non è diagonale, una stima dei parametri attraverso il metodo dei minimi quadrati ordinari risulta non efficiente, poiché non considera la presenza di correlazione tra errori di equazioni successive.

I parametri del modello vengono stimati attraverso il metodo dei minimi quadrati generalizzati (GLS).

In particolare:

$$\hat{\beta}_{GLS} = (X'V^{-1}X)^{-1}X'V^{-1}y$$

rappresenta una stima efficiente del vettore di parametri β .

Nella maggior parte dei casi, però, la matrice Σ , e di conseguenza V , non è nota. Quindi, per ottenere una stima del vettore dei parametri di interesse, è necessario in primo luogo stimare gli elementi della matrice Σ nel modo seguente:

$$\hat{\sigma}_{ij} = df_{ij}^{-1} \hat{\varepsilon}_i' \hat{\varepsilon}_j$$

con df_{ij} che indica il numero di gradi di libertà utilizzati per stimare $i \hat{\sigma}_{ij}$. Di conseguenza, la stima dei parametri, calcolata col metodo dei *Feasible Generalized Least Squares* (FGLS), è la seguente:

$$\hat{\beta}_{GLS} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$$

con $\hat{V} = \hat{\Sigma} \otimes I_T$ e con matrice di varianza e covarianza:

$$\hat{v}ar(\hat{\beta}_{GLS}) = X'V^{-1}X$$

Lo stimatore FGLS è equivalente, asintoticamente, a quello GLS.

Una possibile generalizzazione dei modelli SUR si può ottenere ipotizzando l'esistenza di una relazione non lineare tra esplicative e dipendenti. In formule:

$$y_i = f_i(\beta_i, X_i) + \varepsilon_i$$

con $f_i(\beta_i, X_i)$ che indica un vettore $T \times 1$ di funzioni non lineari. Le assunzioni relative al vettore dei termini d'errore sono le medesime di quelle fatte nel caso si modello SUR lineare. Per ottenere la stima dei parametri per il modello non lineare è necessario minimizzare la seguente espressione:

$$\min_{\beta} \sum_{i=1}^N \sum_{j=1}^N \hat{\sigma}^{ij} (y_i - \mathbf{f}_i(\beta, \mathbf{X})) (y_j - \mathbf{f}_j(\beta, \mathbf{X}))$$

con $\hat{\sigma}^{ij}$ che indica l'elemento di posto ij della matrice Σ^{-1} .

3.2.2 Modello VAR

I modelli VAR (*Vector Autoregression*) generalizzano modelli AR a serie storiche multivariate.

Tali modelli descrivono l'evoluzione di un insieme di k serie storiche, tenendo in considerazione le interdipendenze esistenti tra esse, attraverso una funzione lineare del loro passato. Con \mathbf{Y}_t si indica il vettore relativo alle osservazioni delle k serie storiche al tempo t , di dimensioni $k \times 1$, allora il modello VAR(p) assume la seguente forma:

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t \quad t=1, \dots, T$$

Con \mathbf{A}_i matrice $k \times k$ di coefficienti e $\boldsymbol{\varepsilon}_t$ vettore $k \times 1$ di termini d'errore, che in genere si assume provengano da un processo White Noise, a media zero e con matrice di varianza e covarianza Σ . In notazione matriciale più compatta, utilizzando l'operatore ritardo B , è possibile riscrivere il modello come

$$A(B)\mathbf{Y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t$$

Dove $A(B) = \mathbf{I}_k - \mathbf{A}_1 B - \dots - \mathbf{A}_p B^p$. Affinché il modello sia stabile è necessario che le radici dell'equazione $\det(\mathbf{I}_k - \mathbf{A}_1 z - \dots - \mathbf{A}_p z^p) = 0$ siano tutte in modulo maggiori di 1.

In alcuni contesti il modello VAR(p) può essere inadeguato a cogliere le caratteristiche dei dati; in particolare potrebbe essere necessario inserire una componente deterministica, come ad esempio un trend lineare o un gruppo di *dummy* per modellare la stagionalità. Inoltre, è possibile estendere il modello aggiungendo anche l'effetto di variabili esogene. La forma più completa che si ottiene è la seguente:

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \boldsymbol{\Phi} \mathbf{D}_t + \mathbf{G} \mathbf{X}_t + \boldsymbol{\varepsilon}_t$$

dove \mathbf{D}_t rappresenta la componente deterministica $l \times 1$, \mathbf{X}_t è la matrice $m \times 1$ di variabili esogene e $\boldsymbol{\Phi}$ e \mathbf{G} sono matrici di parametri.

Per quanto riguarda la stima dei parametri del modello, assumendo la stazionarietà in covarianza e che non ci siano restrizioni sui parametri, è possibile pensare che ogni equazione di un modello VAR(p) possa essere scritta, in notazione SUR, nel seguente modo:

$$y_i = \mathbf{Z}a_i + e_i, i = 1, \dots, k$$

dove y_i è il vettore $T \times 1$ di osservazioni dell' i -esima equazione, \mathbf{Z} è una matrice $T \times q$, la cui t -esima riga è data da $\mathbf{Z}'_t = (1, \mathbf{Y}'_{t-1}, \dots, \mathbf{Y}'_{t-p})$, con $q = kp + 1$, a_i è un vettore $q \times 1$ di parametri ed e_i è un vettore $T \times 1$ di termini d'errore a media zero e matrice di varianza e covarianza $\sigma_i^2 \mathbf{I}_T$. Poiché, come detto in precedenza, il VAR(p) è solo una complicazione del SUR, allora è possibile stimare le equazioni utilizzando il metodo dei minimi quadrati o dei minimi quadrati generalizzati, come indicato nel paragrafo 3.2.1. Se $\hat{\mathbf{A}} = [\hat{a}_1, \dots, \hat{a}_k]$ è la matrice $q \times k$ di coefficienti stimati e $vec(\)$ indica l'operatore che trasforma le colonne di una matrice $q \times k$ in un vettore di lunghezza $qk \times 1$, allora, sotto assunzioni di stazionarietà ed ergodicità, $vec(\hat{\mathbf{A}})$ è consistente e asintoticamente distribuito come una normale con matrice di covarianza asintotica

$$a \hat{v}ar(vec(\hat{\mathbf{A}})) = \hat{\Sigma} \otimes (\mathbf{Z}'\mathbf{Z})^{-1}$$

dove

$$\hat{\Sigma} = \frac{1}{T-k} \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}'_t \hat{\boldsymbol{\varepsilon}}_t$$

e $\hat{\boldsymbol{\varepsilon}}_t = \mathbf{Y}_t - \hat{\mathbf{A}}'\mathbf{Z}_t$ è il residuo al tempo t .

Infine, l' i -esimo elemento di $vec(\hat{\mathbf{A}})$, \hat{a}_i , è distribuito normalmente, con errore standard dato dalla radice quadrata dell' i -esimo elemento della matrice $\hat{\Sigma} \otimes (\mathbf{Z}'\mathbf{Z})^{-1}$. Quindi, per verificare ipotesi sui singoli coefficienti si costruiscono test t nel modo usuale. Inoltre, è possibile verificare ipotesi lineari più generali, del tipo $\mathbf{R} \cdot vec(\mathbf{A}) = r$, utilizzando una statistica di Wald del seguente tipo:

$$W = (\mathbf{R} \cdot vec(\hat{\mathbf{A}}))' \{ \mathbf{R} [a \hat{v}ar(vec(\hat{\mathbf{A}}))] \mathbf{R}' \}^{-1} (\mathbf{R} \cdot vec(\hat{\mathbf{A}}))$$

che sotto l'ipotesi nulla segue una distribuzione χ_m^2 , con $m = \text{rango}(\mathbf{R})$, che indica il massimo numero di restrizioni possibili.

Per ulteriori approfondimenti ed applicazioni in S plus dei modelli SUR e VAR si rimanda alla consultazione del manuale Modelling Financial Time Series with S-PLUS di Erich Zivot e Jiahui Wang, 2006.

3.3 Conclusioni

Nel corso dell'ultimo decennio sono stati proposti numerosi tipi di modelli per riuscire a cogliere tutte le caratteristiche che contraddistinguono i prezzi dell'energia elettrica da quelli di altre fonti energetiche o di altri strumenti finanziari. L'obiettivo di coloro che hanno proposto i diversi metodi è fondamentalmente quello di riuscire a produrre previsioni quanto più possibile adeguate che permettano alle compagnie che gestiscono l'energia elettrica di pianificare strategie e prendere decisioni in merito alla produzione. Ciò che emerge dall'applicazione dei modelli ai dati è che il potere previsivo di un modello è strettamente legato alla presenza di stagionalità e di valori anomali nei dati. Inoltre, come in ogni contesto, la capacità del modello di cogliere le peculiarità dei dati deve essere accompagnata da un certo livello di parsimonia.

Nelle sezioni successive si proporrà l'applicazione di alcuni dei modelli descritti a dati reali del mercato dell'energia elettrica francese e si valuteranno i diversi approcci proposti proprio sulla base delle previsioni ad uno e più passi.

Capitolo 4

I dati

Per effettuare la analisi descritte in precedenza si è deciso di studiare la serie di dati relativi ai prezzi dell'energia elettrica per la borsa elettrica francese per il periodo che va dal 27 Novembre 2001, data di inizio delle contrattazioni, al 21 Aprile 2009, per un totale di 56112 osservazioni. Si è, inoltre, deciso di tenere l'ultimo anno di osservazioni a fini previsivi.

Nel periodo considerato è compresa anche la fortissima ondata di caldo che colpì l'intera Europa nell'Estate del 2003. Si trattò di un fenomeno eccezionale per durata ed intensità, che colpì soprattutto paesi come Francia e Portogallo, che subirono una crisi politica conseguente ai danni su ecosistema, infrastrutture e popolazione.

In particolare, la Francia raggiunse temperature record in quella che è stata l'ondata di calore più intensa e duratura dal 1873. Nel paese, nelle prime due decadi di Agosto del 2003, si registrarono circa 15000 morti in più rispetto alla media dello stesso periodo di anni precedenti. L'aumento di decessi riguardò soprattutto la popolazione di età superiore ai 75 anni di età e le donne.

I giorni più tragici furono l'11 e il 12 Agosto, a causa di un effetto di accumulazione del caldo e della quasi totale mancanza di ventilazione.

La grande ondata di calore ha, ovviamente, avuto un impatto significativo anche sui prezzi dell'energia, soprattutto nel mercato francese, il *Paris Power Next*. I dati, quindi, risultano particolarmente adatti per effettuare le analisi descritte perché evidenziano, soprattutto nel periodo del grande caldo, il cosiddetto fenomeno degli *spikes*.

Di seguito si presentano analisi descrittive e grafiche della serie dei prezzi elettrici relativi al periodo considerato.

Si analizza poi il comportamento delle serie relative alle 24 ore del giorno.

4.1 La serie storica dei prezzi

Per valutare le caratteristiche della serie storica dei prezzi dell'energia elettrica è necessario costruire alcuni grafici che permettano di individuare gli *spikes* e la loro durata.

In prima battuta si può considerare il semplice grafico della serie (Fig:4.1)

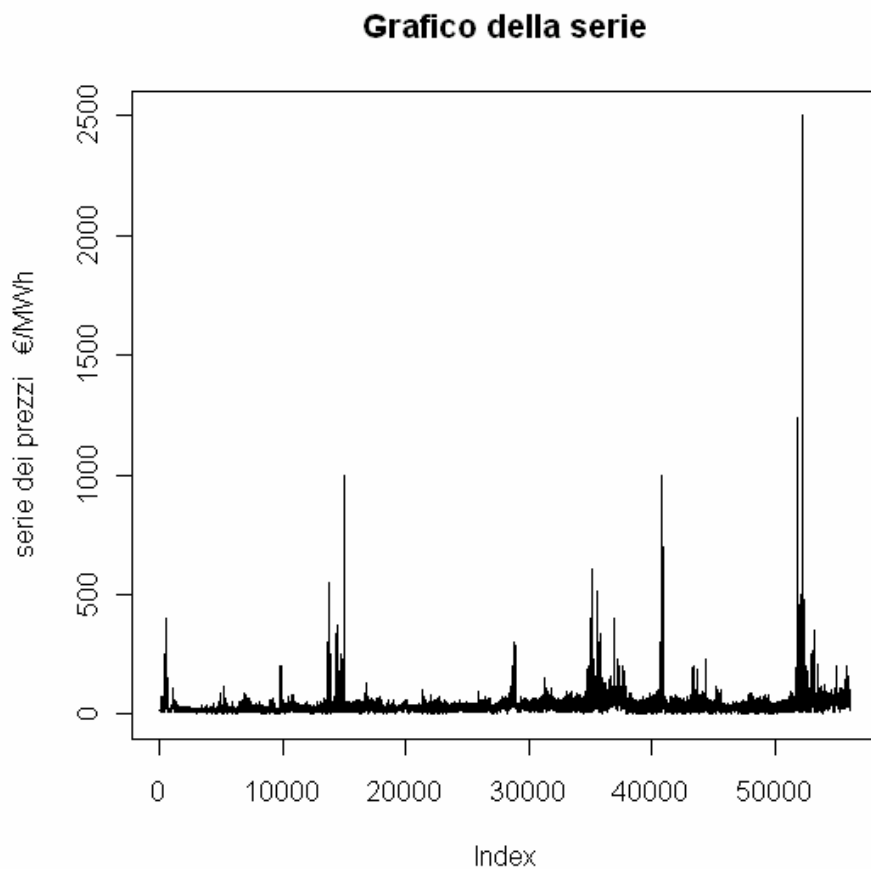


Figura 4.1:Grafico della serie storica dei prezzi dal 27/11/2001 al 21/04/2008

Dall'analisi del grafico si osservano numerosi picchi. In particolare, è possibile osservare valori anomali in corrispondenza della grande ondata di caldo del 2003

(osservazioni da 13500 a 15000), nel mese di Luglio del 2006 (osservazioni da 40000 a 41000) e nell'autunno del 2007 (osservazioni da 51900 a 52400). Il picco più consistente è quello osservato il 12 Novembre del 2007, quando si è toccata la soglia di 2500 euro per MWh. Nel Novembre del 2007, infatti, il freddo, gli scioperi in Francia e la fermata delle centrali a carbone in Germania provocarono una crescita smisurata dei prezzi dell'energia elettrica in tutta l'Europa. A Parigi, in alcuni istanti si superarono i duemila euro per chilowattora a causa del massiccio utilizzo degli impianti di riscaldamento che in Francia sono alimentati ad elettricità.

Sulla serie si calcolano poi le principali statistiche descrittive, relative alla serie oraria dei prezzi elettrici, vengono riportate nella Tabella 4.1:

Statistiche descrittive	
Media	37.36262
Mediana	30.94
Varianza	1119.562
Minimo	0.004
Massimo	2500
Simmetria	17.41407
Curtosi	372.9632

Tabella 4.1: Statistiche descrittive per la serie dei dati orari

Quello che emerge chiaramente da un esame delle statistiche riportate nella Tabella 4.1 è che la serie ha una varianza molto alta a causa della presenza di dati anomali.

La distribuzione mostra una forte asimmetria positiva e un valore della curtosi decisamente superiore a 3, indicazione di code più pesanti di quelle di una normale standard.

Le caratteristiche che emergono dall'esame delle statistiche descrittive risultano evidenti anche dall'analisi dell'istogramma e del qqplot, ossia il grafico che mette a confronto i quantili teorici di una distribuzione normale con quelli della distribuzione empirica dei dati.

I due grafici sono riportati nella Figura 4.2:

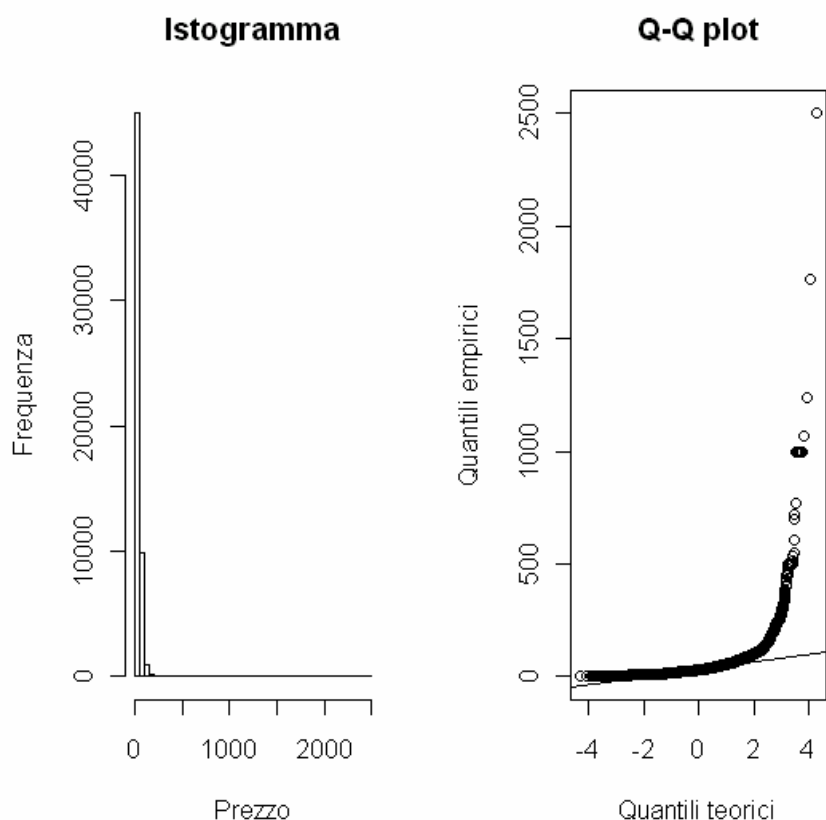


Figura 4.2: Istogramma e qq-plot per la serie dei prezzi

L'istogramma mostra che i dati hanno una coda destra molto pesante a causa della presenza degli *spikes*; questa evidenza spiega sia la presenza di un coefficiente di simmetria positivo e grande, sia un indice di curtosi molto maggiore di 3.

Inoltre, dall'analisi del qqplot emerge chiaramente che la distribuzione dei dati si discosta da quella di una variabile normale. In particolare, la coda sinistra e i valori centrali sembrerebbero compatibili con la normalità, mentre la parte destra della distribuzione è decisamente distante dalla retta che rappresenta i quantili di una distribuzione gaussiana.

Al fine di comprendere meglio le caratteristiche della serie è possibile prendere in esame i grafici relativi alle funzioni di autocorrelazione globale e parziale ed il periodogramma.

La funzione di autocorrelazione permette di valutare se ci sia dipendenza temporale nei dati, mentre l'autocorrelazione parziale permette di valutare l'esistenza di un legame diretto tra due variabili al netto di quelle intermedie.

Il periodogramma, invece, è la rappresentazione grafica della funzione di densità spettrale e aiuta ad identificare le periodicità presenti nelle serie storiche.

Per la serie considerata i tre grafici che si ottengono sono rappresentati in Figura 4.3:

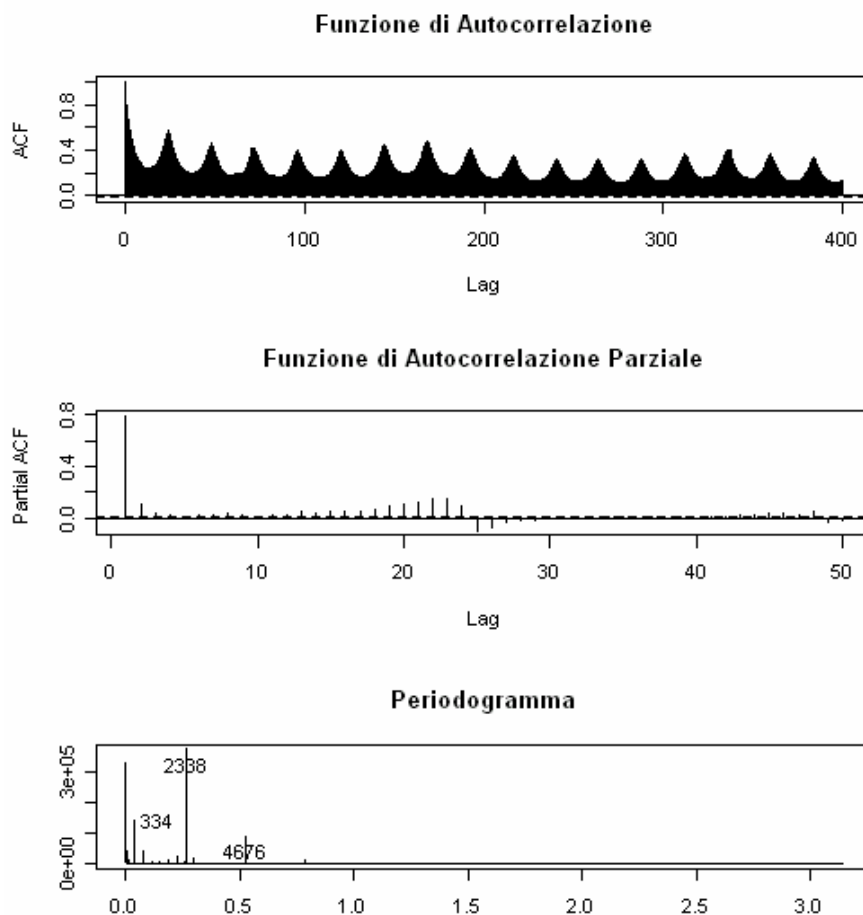


Figura 4.3: Autocorrelogramma della serie; correlogramma parziale; periodogramma

Sia per la funzione di correlazione globale che per quella parziale nel grafico vengono rappresentate le bande di confidenza al 5%. Se la serie storica fosse generata da un White Noise i valori delle due funzioni sarebbero interne alle bande.

Nel caso della serie analizzata si nota che la funzione di autocorrelazione globale è fortemente significativa anche dopo 400 ritardi. Questa caratteristica è indicazione che nei dati c'è una forte persistenza temporale. Sembra, inoltre, presente una certa ciclicità che si giustifica con l'esistenza di una componente stagionale, sia di periodo 24 (le ore del giorno) che di periodo 7 (settimanale), nei dati. Infatti, si evidenzia anche la presenza di un picco più alto ai ritardi 168 e 336, che si ottengono per l'effetto della

stagionalità di periodo 7 in dati orari. ($168=7\times 24$). Per quanto riguarda il periodogramma, invece, si osservano picchi ai valori 334, 2338 e 4676, in corrispondenza dei quali il periodo è pari a 12, 24 e 168 e questo è coerente con la presenza di stagionalità semigiornaliere, giornaliera e settimanali.

Data l'evidenza di una stagionalità settimanale è possibile pensare di rappresentare in un grafico i boxplot dei dati relativi ai prezzi suddivisi per giorno della settimana.

Per rendere i grafici più leggibili si è deciso di applicare la trasformata logaritmica ai dati al fine di ridurre la varianza della serie.

I boxplot sono riportati nella Figura 4.4:

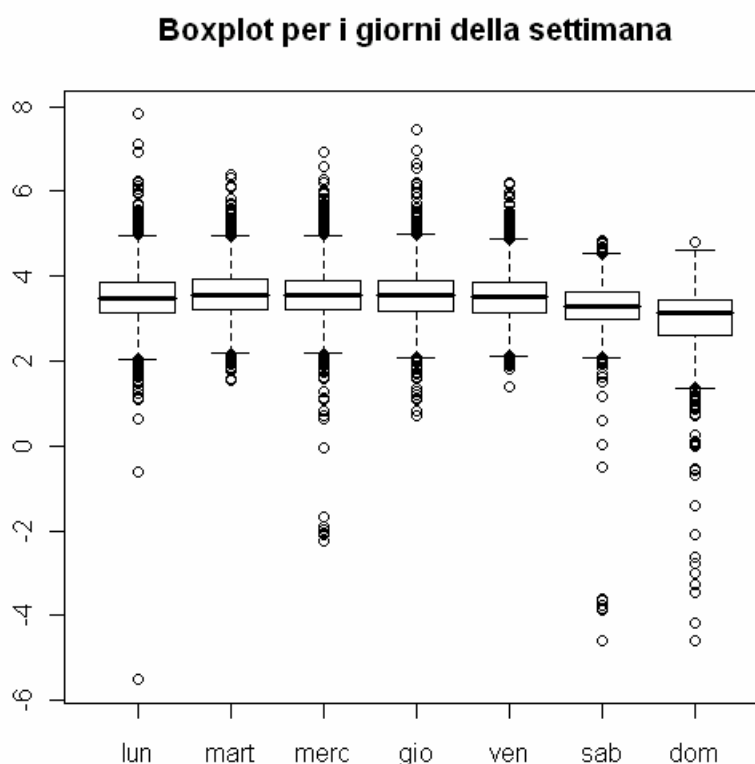


Figura 4.4: Boxplot del logaritmo della serie dei prezzi per i diversi giorni della settimana

L'analisi del grafico evidenzia la presenza di numerosi *outliers*, ossia valori che eccedono le bande poste ad una distanza pari ad 1,5 volte il *range* interquartile dal primo e terzo quartile. I giorni lavorativi hanno grafici molto simili; il lunedì e il venerdì hanno un livello medio leggermente inferiore ai giorni centrali poiché sono in prossimità del weekend. Il boxplot relativo al sabato risulta spostato verso il basso rispetto a quello degli altri giorni e questa caratteristica è ulteriormente accentuata per quanto riguarda la domenica. Questa caratteristica si motiva con la riduzione dei

consumi nel weekend per le chiusure di negozi ed industrie; una domanda inferiore spinge il prezzo verso il basso, soprattutto nei mesi invernali.

Analogamente, è possibile valutare se esistano differenze di livello tra i giorni lavorativi e le festività del calendario francese. Il grafico che si ottiene, (Figura 4.5) costruito anche in questo caso sui logaritmi dei dati per rendere più leggibile il risultato, è il seguente:

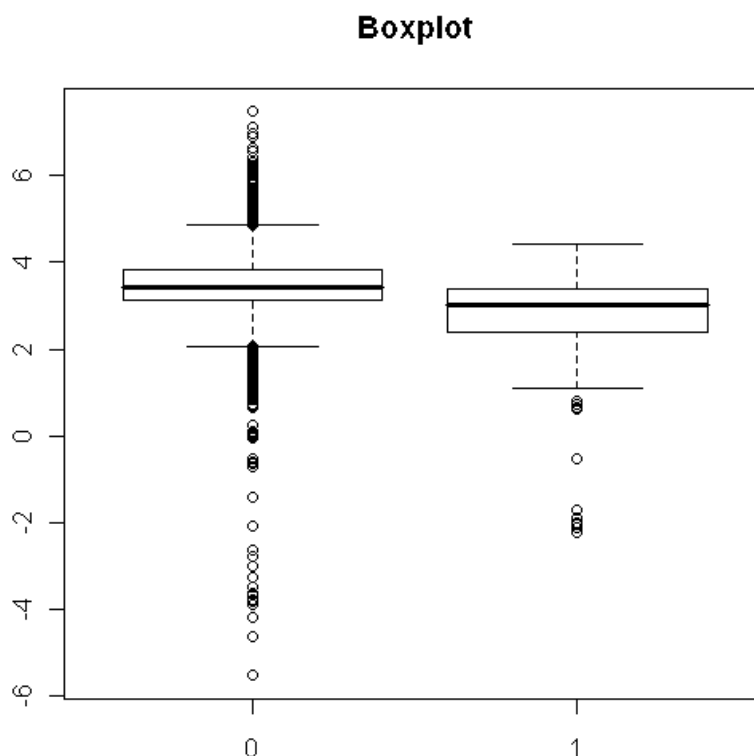


Figura 4.5: Boxplot per la serie dei logaritmi, per giorni festivi (1) ed lavorativi(0)

dove il boxplot a destra è relativo ai giorni festivi.

Ciò che emerge dall'analisi del grafico è che in corrispondenza delle festività, come era logico attendersi, il livello medio dei prezzi è inferiore rispetto ai restanti giorni dell'anno. Il grafico relativo ai giorni festivi è spostato in basso rispetto a quello per i giorni ordinari, mentre quest'ultimo mostra la presenza di numerosi valori anomali. Anche in questo caso l'interpretazione del grafico è legata alla riduzione della domanda di energia elettrica nei giorni festivi, durante i quali la maggior parte delle imprese rimangono chiuse.

4.2 Analisi descrittiva dei dati giornalieri

In molti lavori presenti in letteratura sono stati analizzati dati giornalieri piuttosto che dati orari (si veda, ad esempio Escrivano et al,2002, Weron, 2008 e Spagnolo e Uddin, 2006). A partire dalle serie storiche orarie dei prezzi dell'energia elettrica, i prezzi giornalieri possono essere costruiti come media dei 24 dati orari. Inoltre, è utile distinguere i dati in *baseload* e *peakload*. Nel primo caso si considerano tutti i dati a disposizione, nel secondo ci si sofferma solo sulle ore cosiddette di picco, in cui i prezzi dovrebbero essere più alti. Nel seguito si considerano come ore caratterizzate da picchi quelle comprese tra le 09:00 e le 21:00. Si ottengono così due serie storiche di dati giornalieri, composte di 2338 osservazioni, su cui si possono effettuare tutte le analisi descrittive già eseguite per le serie di dati orari. Innanzitutto si possono rappresentare graficamente le due serie (Figura 4.6):

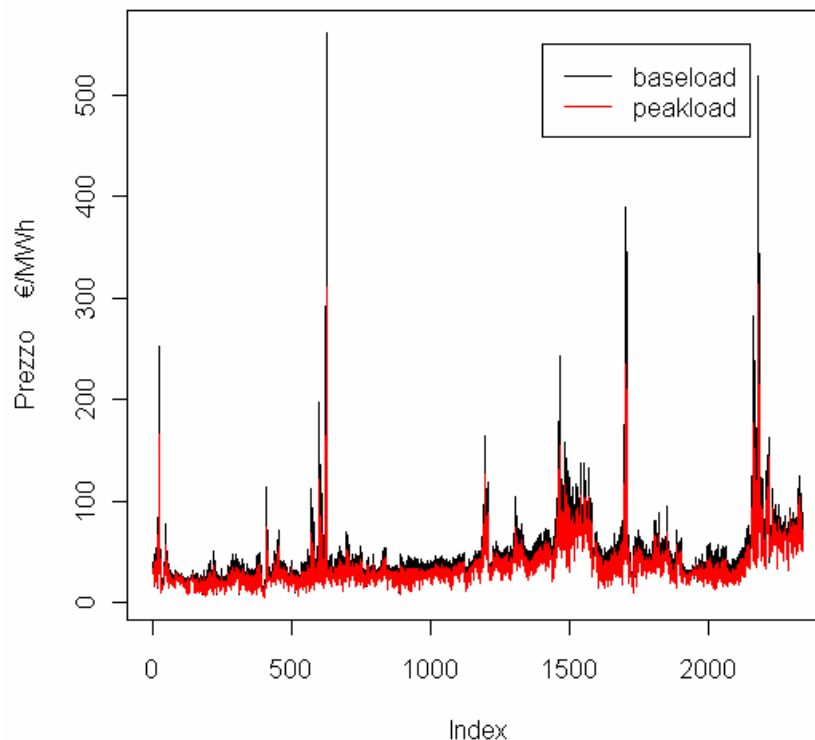


Figura 4.6: Grafico per la serie dei prezzi giornalieri calcolati sulle 24 ore (*baseload*) e dei prezzi giornalieri calcolati solo sulle ore di picco, dalle 09:00 alle 21:00 (*peakload*)

In rosso è rappresentato il grafico della serie giornaliera calcolata prendendo la media dei prezzi sulle 24 ore (*baseload*).

Dall'analisi del grafico si nota che il picco più elevato è quello corrispondente all'osservazione 623, ossia all'11 Agosto del 2003, che rappresenta il giorno più caldo del periodo di grande afa.

Ovviamente, come ci si attendeva, la media calcolata nelle ore di picco presenta valori sistematicamente maggiori di quelli corrispondenti al *baseload*.

Per approfondire l'analisi, si possono rappresentare anche in questo caso gli istogrammi delle due serie (Figura 4.7) e confrontare le statistiche descrittive (Tabella 4.2).

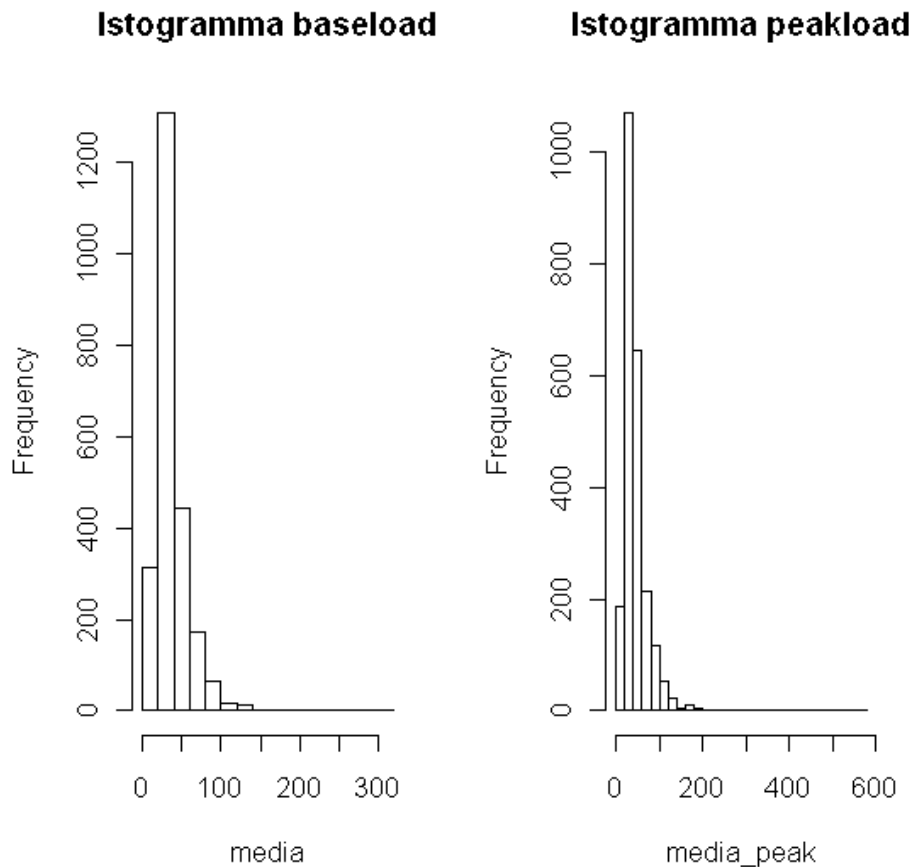


Figura 4.7: Istogramma della serie dei prezzi giornalieri calcolati sulle 24 ore (*baseload*) e dei prezzi giornalieri calcolati nelle ore di picco (*peakload*)

Statistiche descrittive		
	Baseload	Peakload
Media	37.32039	45.67556
Mediana	31.73952	38.59346
Varianza	496.2205	1030.582
Minimo	4.929	5.292
Massimo	314.2690	561.2863
Simmetria	3.398008	5.377535
Curtosi	28.91261	63.14337

Tabella 4.2: Statistiche descrittive per la serie dei prezzi giornalieri calcolati sulle 24 ore (*baseload*) e dei prezzi giornalieri calcolati nelle ore di picco (*peakload*)

Dall'analisi degli istogrammi emerge chiaramente come la distribuzione della media calcolata solo sulle ore di picco abbia una coda destra più pesante e mostri una maggiore asimmetria.

Quest'evidenza è confermata anche dall'esame delle statistiche descrittive che mostrano che nelle ore di picco il prezzo medio è più elevato, c'è una maggiore variabilità nei dati, si evidenzia un'asimmetria positiva più consistente e l'indice di curtosi è più alto.

4.3 Le serie orarie

Una volta evidenziate le caratteristiche dell'intera serie si può procedere con l'esame delle 24 serie orarie, al fine di capire in che ore si concentrino i picchi e quale sia la loro durata.

Per non appesantire l'analisi, i grafici e gli autocorrelogrammi delle 24 serie orarie sono riportati in Appendice A.

Dall'analisi dei grafici emerge immediatamente che i picchi sono più evidenti nelle ore comprese tra le 10:00 e le 21:00, che corrispondono con le ore lavorative.

Gli autocorrelogrammi, invece, mettono in luce la forte persistenza temporale presente nei dati; è evidente che le autocorrelazioni sono ancora molto significative anche dopo molti ritardi.

L'analisi dei correlogrammi, inoltre, mette in luce la presenza di una stagionalità di periodo sette piuttosto persistente, che risulta particolarmente evidente nelle ore comprese tra le 06:00 e le 20:00.

4.4 Analisi dei dati per le previsioni

I dati che vengono tenuti da parte ai fini previsivi sono relativi alle osservazioni per il periodo che va dal 22 Aprile del 2008 fino al 21 Aprile del 2009. La rappresentazione grafica di questi dati è riportata in Figura 4.8:

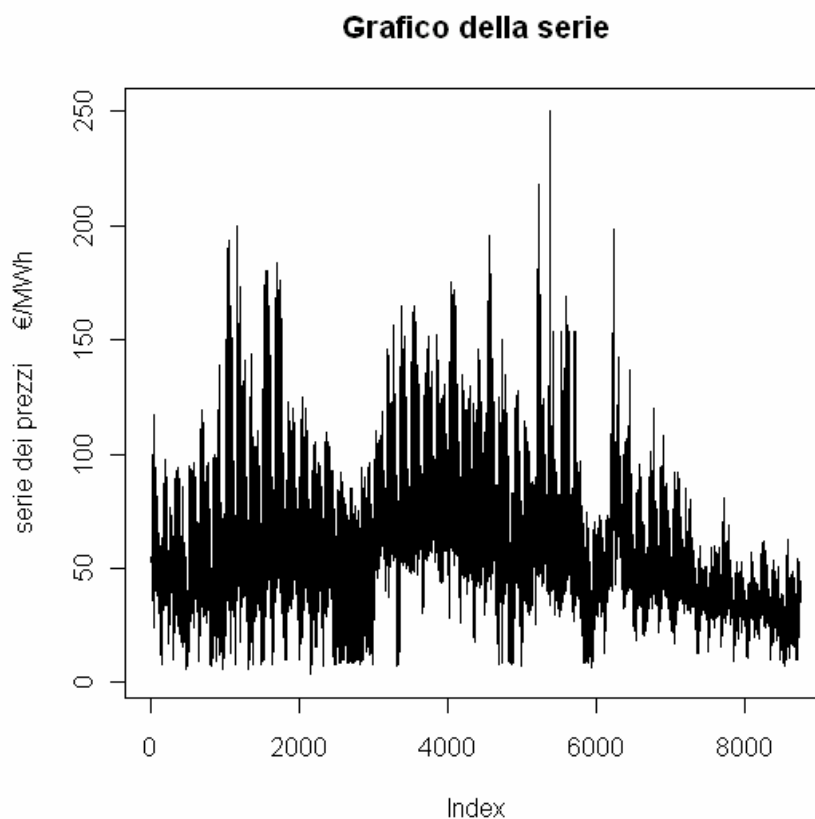


Figura 4.8: Grafico della serie dei prezzi utilizzati a fini previsivi, che vanno dal 22/04/2008 al 21/04/2009

Dall'esame del grafico della serie si nota come nell'anno considerato la serie non evidenzia particolari picchi. Il valore più elevato si osserva in corrispondenza delle ore 20:00 dell'1 Dicembre 2008 (osservazione 5371), ed è pari a 250 €/MWh.

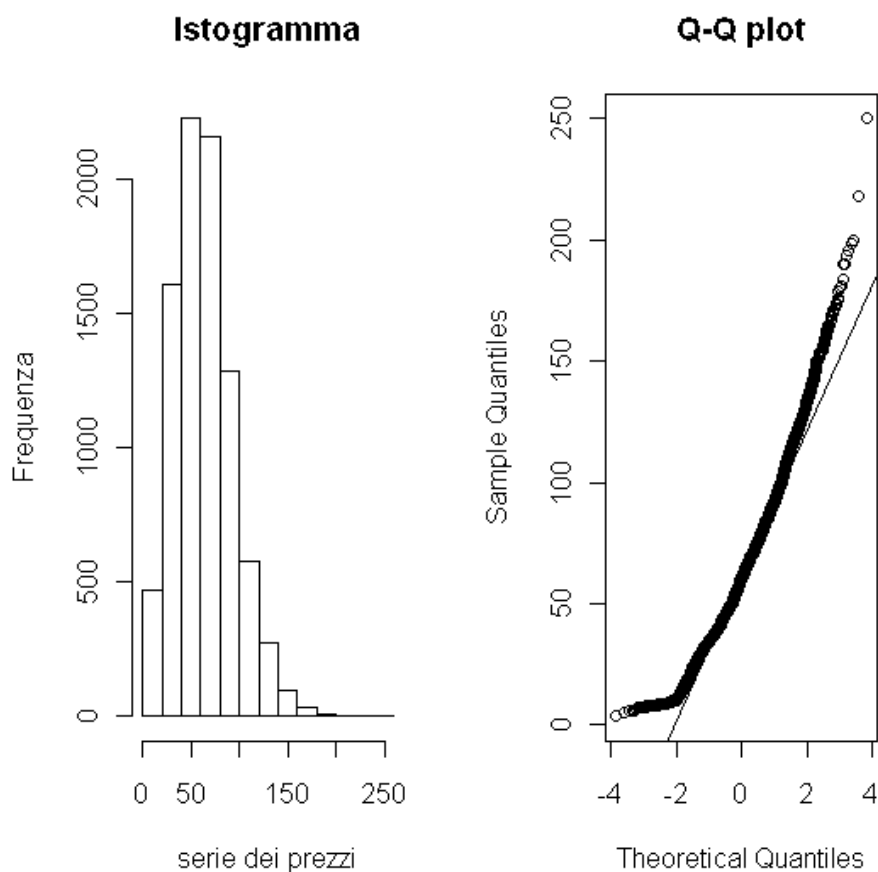


Figura 4.9: Istogramma e qq-plot della serie dei prezzi orari utilizzati a fini previsivi

L'istogramma, riportato in Figura 4.9, evidenzia come la serie abbia una leggera asimmetria positiva, confermata, tra l'altro dall'indice riportato nella Tabella 4.3. Inoltre, l'indice di curtosi, di poco superiore a 3, segnala che le code della distribuzione hanno peso simile a quelle di una distribuzione normale. Il qqplot (Figura 4.9), infine, evidenzia come la parte centrale della distribuzione abbia quantili prossimi a quelli di una variabile gaussiana. I quantili teorici e quelli empirici si discostano maggiormente in corrispondenza delle code.

Statistiche descrittive	
Media	63.44526
Mediana	60.8715
Varianza	910.737
Minimo	3.491
Massimo	250
Simmetria	0.6892429
Curtosi	3.7137183

Figura 4.9: Statistiche descrittive della serie dei prezzi orari utilizzati a fini previsivi

I boxplot costruiti per ogni giorno della settimana, riportati in Figura 4.10, mostrano come le distribuzioni dei giorni centrali siano tra loro molto simili, mentre quelle relative al weekend risultano spostate verso il basso. In particolare, il grafico per la domenica mostra una distribuzione ulteriormente spostata in basso rispetto a quella del sabato.

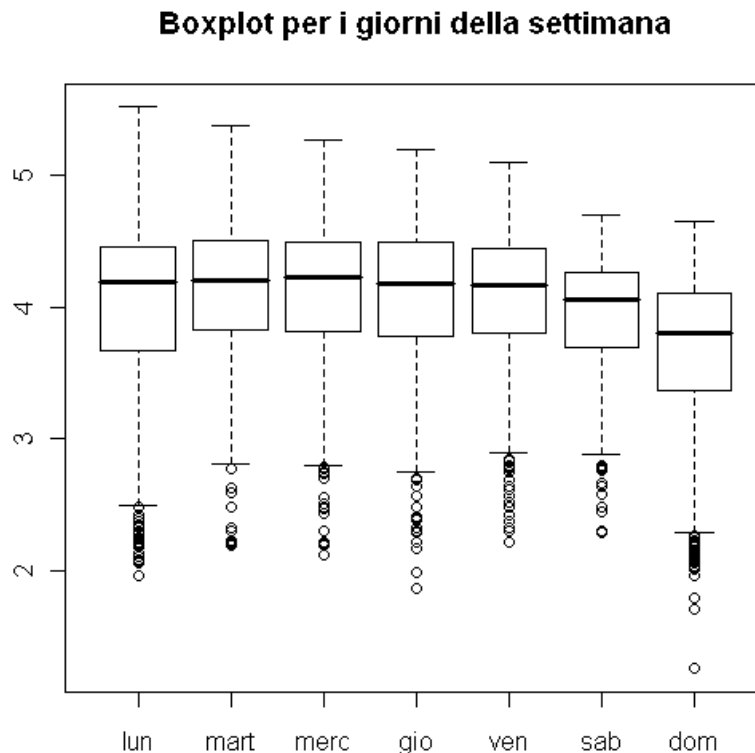


Figura 4.10: Boxplot della serie dei logaritmi dei prezzi orari utilizzati a fini previsivi, considerati per i diversi giorni della settimana.

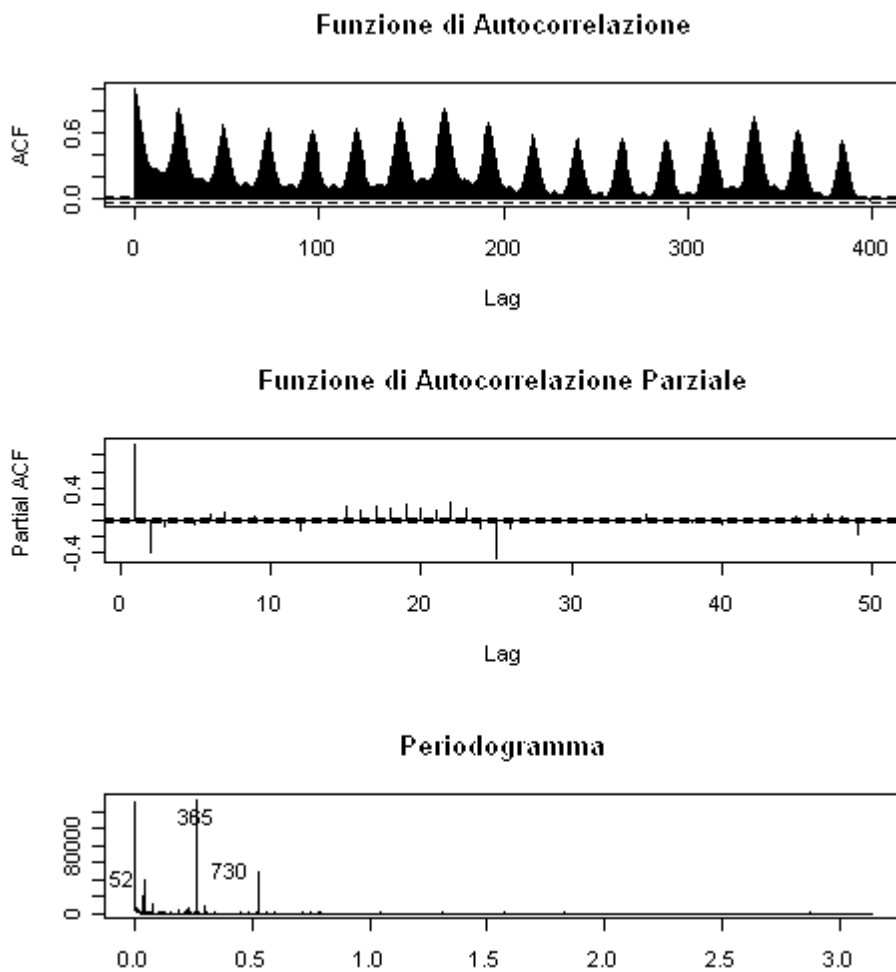


Figura 4.11: Autocorrelogramma dei prezzi orari utilizzati a fini previsivi, correlogramma parziale e periodogramma

L'analisi del correlogramma e dello spettro, rappresentati in Figura 4.11, portano a conclusioni analoghe a quelle per la serie complessiva. Infatti la funzione di autocorrelazione mostra una periodicità dovuta alla stagionalità settimanale e oraria ed una forte persistenza temporale e il periodogramma evidenzia picchi che corrispondono a periodicità semigiornaliere, giornaliera e settimanali.

Capitolo 5

Applicazione dei modelli ai dati

Nelle sezioni precedenti si sono presentati alcuni dei modelli più utilizzati per l'analisi delle serie dei prezzi dell'energia elettrica e si sono descritte le caratteristiche dei dati presi in esame. Nel seguito si applicheranno alcuni dei modelli indicati, valutandone la bontà di adattamento ai dati e confrontandoli in termini previsivi.

In particolare, si è scelto di applicare ai dati del mercato francese presentati nel Capitolo 4 il modello SUR, che permette di stimare i parametri tenendo in considerazione la correlazione tra i termini di errore di ore successive, la sua evoluzione, rappresentata dal modello VAR; modelli ARMA-GARCH ora per ora e modelli *Markov Switching*, anche in questo caso per le 24 ore considerate separatamente.

Tutte le analisi sono state compiute utilizzando i software R e S plus, costruendo le opportune routine quando necessario.⁶

5.1 Stima della componente deterministica

Sia P_t la serie storica dei prezzi dell'energia elettrica e $p_t = \log(P_t)$. È possibile suddividere la serie in due componenti, una stocastica, $f_i(t)$ e l'altra deterministica, cosicché:

⁶ Per motivi di spazio in questo lavoro non sono stati riportati le routine utilizzate. Sono comunque disponibili su richiesta.

$$p_h(t) = \log(P_h(t)) = f_h(t) + X_h(t)$$

dove $p_h(t)$ è il logaritmo naturale del prezzo. La componente deterministica, $f_h(t)$, tiene in considerazione componenti regolari periodiche predicibili, come il trend o la stagionalità, e generalmente è una funzione deterministica del tempo. La componente stocastica $X_h(t)$ è la più interessante e su di essa si concentreranno maggiormente le successive analisi. Nel seguito di questo lavoro si applicheranno i modelli proposti alla serie storica dei logaritmi, che per semplicità continueranno ad essere indicati come prezzi.

Prima di procedere all'applicazione dei modelli sopraelencati è necessario depurare le serie dei prezzi della componente deterministica, che tiene conto di alcune regolarità come il livello medio dei prezzi e la loro periodicità. Seguendo l'approccio proposto da Huisman et al. (2006), si può modellare la componente deterministica inserendo un livello medio di prezzo μ_0 e una quantità μ_h che rappresenta la deviazione della media dell'ora h dal livello medio totale del prezzo. A queste quantità si aggiungono delle *dummy* specifiche per ogni giorno della settimana: la variabile $I^d(t)$ vale 1 se il giorno di riferimento per il prezzo($t+1$) è il giorno d , $d=1,2,\dots,7$. La formula per $f_h(t)$ è quindi:

$$f_h(t) = \mu_0 + \mu_h + \sum_d \beta_d I^d(t) + \gamma F(t), \quad h=1, 2, \dots, 24$$

Per quanto riguarda i dati in esame si è anche deciso di inserire nella componente deterministica una variabile *dummy*, $F(t)$, che tenga conto dei giorni di festività in cui si potrebbe registrare una riduzione della domanda di energia elettrica. Nel caso dei dati in esame, relativi al mercato francese, le festività considerate sono la Pasqua, il Natale, la Presa della Bastiglia (il 14 Luglio), l'Ascensione, l'Assunzione della Vergine (il 15 Agosto) e il giorno dell'armistizio (l'11 Novembre).

Oltre a quello proposto, ci sono altri modi in cui è possibile modellare la componente deterministica. Si possono, ad esempio, inserire come variabili esplicative il tempo t , una variabile *dummy* che indichi se il giorno della settimana fa parte del weekend o meno e una funzione sinusoidale del tempo che permetta di tenere in considerazione la stagionalità. Questo tipo di modellazione risulta più adatta a dati giornalieri e non orari ed è stata utilizzata da Spagnolo e Uddin (2006).

Anche nel modello descritto in precedenza e che verrà applicato ai dati in esame per eliminare la componente deterministica, è possibile inserire una funzione del tempo, in modo da tenere in considerazione la ciclicità, dovuta alle caratteristiche del mercato, che

spesso si nota nelle serie dei prezzi dell'elettricità. Si può ad esempio costruire una *spline* del tempo. La *spline* è una funzione, costituita da un insieme di polinomi raccordati tra loro, il cui scopo è interpolare un insieme di punti, i nodi, in modo da essere continua ed ammettere derivate almeno fino ad un dato ordine, in ogni punto dell'intervallo. Affinché sia garantita la continuità nei nodi, in genere indicati con $\xi_1, \xi_2, \dots, \xi_k$, è necessario imporre dei vincoli. L'interpolazione attraverso *spline* rappresenta uno dei metodi non parametrici per stimare una funzione $f(x)$ in un modello di regressione del tipo:

$$Y(t) = f(x) + a(t)$$

Con $a_t \sim \text{i.i.d.}(0, \sigma^2)$.

Le *splines* più utilizzate sono quelle cubiche (grado tre). In generale una *spline* di grado k assume la seguente forma:

$$S(x) = \beta_0 + \beta_1 x + \dots + \beta_p x^p + b_1 (x - \xi_1)_+^p + \dots + b_k (x - \xi_k)_+^p$$

$$\text{con } (x-t)_+^p = \begin{cases} 0 & x < t \\ (x-t)^p & x \geq t \end{cases}$$

In forma più compatta si può scrivere:

$$S(x) = \sum_{j=1}^{k+p+1} \beta_j S_j(x)$$

con

$$S_j(x) = x^{j-1} \quad j = 1, \dots, p+1$$

$$S_j(x) = (x - t_j)_+^p \quad j = p+2, \dots, k+p+1$$

Per stimare i parametri del modello si utilizza il metodo dei minimi quadrati.

Nel contesto in esame la variabile temporale è costruita in modo da assumere valore i nell' i -esimo giorno dell'anno. Per la stima della *spline* si è utilizzata la libreria `splines` di R.

In questo lavoro si è deciso di stimare la componente deterministica sia sulla serie originaria, che su quella ripulita dei valori anomali. Nel secondo caso, poi, si possono seguire due approcci: si può continuare a lavorare sulla serie depurata degli *spikes* oppure si può eliminare la componente deterministica stimata sulla serie ripulita dalla serie originaria.

Le stime dei parametri della componente deterministica relativa ai logaritmi dei prezzi, con e senza *spikes*, sono riportate nella Tabella 5.1:

	Serie originaria	Serie senza picchi
μ_0	3.518977 (0.00)	3.517395 (0.00)
μ_1	-0.189512 (0.00)	-0.189376 (0.00)
μ_2	-0.296045 (0.00)	-0.295910 (0.00)
μ_3	-0.389942 (0.00)	-0.389806 (0.00)
μ_4	-0.539894 (0.00)	-0.539759 (0.00)
μ_5	-0.640596 (0.00)	-0.640461 (0.00)
μ_6	-0.414044 (0.00)	-0.413909 (0.00)
μ_7	-0.237619 (0.00)	-0.237483 (0.00)
μ_8	0.028218 (0.06)	0.027514 (0.07)
μ_9	0.209699 (0.00)	0.207731 (0.00)
μ_{10}	0.316245 (0.00)	0.311662 (0.00)
μ_{11}	0.365735 (0.00)	0.360458 (0.00)
μ_{12}	0.412268 (0.00)	0.404227 (0.00)
μ_{13}	0.370052 (0.00)	0.366536 (0.00)
μ_{14}	0.318303 (0.00)	0.315289 (0.00)
μ_{15}	0.262727 (0.00)	0.259364 (0.00)
μ_{16}	0.202153 (0.00)	0.199298 (0.00)
μ_{17}	0.172389 (0.00)	0.170156 (0.00)
μ_{18}	0.235978 (0.00)	0.230647 (0.00)
μ_{19}	0.324785 (0.00)	0.314807 (0.00)
μ_{20}	0.302437 (0.00)	0.299008 (0.00)
μ_{21}	0.212988 (0.00)	0.212614 (0.00)
μ_{22}	0.127816 (0.00)	0.127581 (0.00)
μ_{23}	0.110790 (0.00)	0.110849 (0.00)
β_1	-0.002128 (0.80)	-0.001495 (0.85)
β_2	-0.064984 (0.00)	-0.063349 (0.00)
β_3	-0.264361 (0.00)	-0.260589 (0.00)
β_4	-0.538804 (0.00)	-0.535086 (0.00)

β_5	-0.097003 (0.00)	-0.097548 (0.00)
β_6	0.013812 (0.09)	0.014123 (0.08)
γ	-0.490953 (0.00)	-0.488337 (0.00)

Tabelle 5.1: Stima della componente deterministica calcolata sulla serie originaria e sulla serie depurata dei picchi

I risultati che si ottengono nelle due situazioni sono tra loro molto simili. I coefficienti μ_i , $i = 1, \dots, 23$, rappresentano gli scarti tra il livello medio μ_0 e le medie delle singole ore. Ciò che si osserva è che per le ore dalla 1:00 alla 7:00 il livello medio risulta significativamente più basso del livello medio complessivo. Il coefficiente dell'ora 8:00 non è significativo, considerando una soglia del 5%. Le ore dalle 9 alle 23:00 hanno un livello medio più alto di quello complessivo e i coefficienti sono tutti significativi. L'ora 24:00 è esclusa per evitare collinearità. Per quanto riguarda i giorni della settimana la *baseline* nel modello stimato è il mercoledì ed il coefficiente β_d corrisponde al d -esimo giorno della settimana, con $d=1$ che corrisponde a giovedì, $d=2$ a venerdì e così via. In entrambe le situazioni i coefficienti relativi al giovedì e al martedì non sono significativi. Sono invece, significativi e negativi i coefficienti β_3 e β_4 , che corrispondono al weekend. Questo risultato è coerente con l'ipotesi che nel fine settimana il livello medio dei prezzi si riduca a causa della diminuzione della domanda di energia elettrica. Anche il coefficiente per la *dummy* che distingue i giorni festivi è negativo e significativo. L'interpretazione è analoga a quella data per i coefficienti β_3 e β_4 .

5.2 Stima della componente stocastica

Alla serie storica dei prezzi del mercato elettrico francese, relativa al periodo che va dall'inizio delle contrattazioni al 21 Aprile del 2009, si è deciso di applicare modelli ARMA-GARCH e *Markov-Switching* univariato, alle 24 singole serie orarie e i modelli SUR e VAR, che, invece, considerano le 24 ore simultaneamente. Nel seguito si mostreranno i risultati ottenuti, utilizzando i software S+ e R; nella sezione successiva si utilizzeranno i metodi proposti a fini previsivi. Per tutti i modelli proposti si sono

analizzati i residui ottenuti, ma i risultati di tale analisi non sono stati riportati per evitare di appesantire il lavoro. In tutti i casi l'analisi dei residui è abbastanza buona; tuttavia le funzioni di autocorrelazione sono significative ai ritardi stagionali, segnale della presenza di stagionalità residua nei dati.

5.2.1 ARMA - GARCH

Dall'analisi dei correlogrammi delle 24 serie orarie, (si veda l'Appendice A) è emersa la presenza di una forte dipendenza nei dati. Al fine di tener conto di questa caratteristica si è deciso di provare ad applicare ai dati dei modelli AR(1); si è eliminata l'intercetta perché è risultata non significativa per tutte le serie orarie.⁷ Ai residui dei modelli è stato, quindi, applicato il test ARCH-lm che verifica l'ipotesi nulla di omoschedasticità contro l'alternativa di presenza di eteroschedasticità di forma ARCH nei dati. Per tutte e 24 le serie storiche il *p-value* del test è risultato molto basso, portando alla conclusione che siano presenti effetti ARCH. A seguito di questo risultato si è deciso di complicare il modello stimato, aggiungendo una componente che tenga conto dell'eteroschedasticità.

Il modello stimato in definitiva è un modello AR(1)-GARCH(1,1):

$$y_t = \phi y_{t-1} + v_t, v_t = \sigma_t \varepsilon_t, \text{ dove } \sigma_t^2 = \alpha_0 + \alpha_1 y_t^2 + \beta \sigma_{t-1}^2 \text{ e } \varepsilon_t \sim N(0,1)$$

I risultati delle stime dei 24 modelli orari sono riportati nella Tabella 5.2:

	ϕ	α_0	α_1	β
Ora 1	0.7589 (0.00)	0.1057 (0.00)	0.1473 (0.00)	
Ora 2	0.7614 (0.00)	0.1174 (0.00)	0.2254 (0.00)	
Ora 3	0.7742 (0.00)	0.1152 (0.00)	0.2570 (0.00)	
Ora 4	0.67006 (0.00)	0.02673 (0.00)	0.06781 (0.00)	0.80495 (0.00)
Ora 5	0.58157 (0.00)	0.03139 (0.00)	0.22025 (0.00)	0.68348 (0.00)
Ora 6	0.7812 (0.00)	0.1092 (0.00)	0.3137 (0.00)	

⁷ Per la stima dei modelli si è utilizzata la libreria `finmetrics` del software S plus

Ora 7	0.67932 (0.00)	0.01256 (0.00)	0.08120 (0.00)	0.86842 (0.00)
Ora 8	0.66605 (0.00)	0.01334 (0.00)	0.09496 (0.00)	0.85460 (0.00)
Ora 9	0.727776 (0.00)	0.006134 (0.00)	0.080269 (0.00)	0.875122 (0.00)
Ora 10	0.86015 (0.00)	0.00812 (0.00)	0.24424 (0.00)	0.70692 (0.00)
Ora 11	0.889014 (0.00)	0.003145 (0.00)	0.150259 (0.00)	0.820818 (0.00)
Ora 12	0.884713 (0.00)	0.002546 (0.00)	0.135985 (0.00)	0.839056 (0.00)
Ora 13	0.895650 (0.00)	0.002009 (0.00)	0.130767 (0.00)	0.843507 (0.00)
Ora 14	0.879295 (0.00)	0.002099 (0.00)	0.135302 (0.00)	0.841469 (0.00)
Ora 15	0.87952 (0.00)	0.00421 (0.00)	0.17917 (0.00)	0.77573 (0.00)
Ora 16	0.864125 (0.00)	0.004492 (0.00)	0.167155 (0.00)	0.795133 (0.00)
Ora 17	0.837170 (0.00)	0.006585 (0.00)	0.148422 (0.00)	0.786420 (0.00)
Ora 18	0.871235 (0.00)	0.00441 (0.00)	0.114336 (0.00)	0.822097 (0.00)
Ora 19	0.866539 (0.00)	0.002541 (0.00)	0.113393 (0.00)	0.860514 (0.00)
Ora 20	0.862000 (0.00)	0.002697 (0.00)	0.082541 (0.00)	0.872399 (0.00)
Ora 21	0.90889 (0.00)	0.03325 (0.00)	0.43605 (0.00)	0.05098 (0.00)
Ora 22	0.88885 (0.00)	0.04015 (0.00)	0.39400 (0.00)	
Ora 23	0.8551 (0.00)	0.06215 (0.00)	0.20986 (0.00)	
Ora 24	0.86486 (0.00)	0.04495 (0.00)	0.30274 (0.00)	

Tabella 5.2: Stima dei modelli AR(1)-GARCH(1,1) sulle 24 serie orarie; tra parentesi sono riportati i pvalue dei test t per la significatività dei parametri.

I parametri autoregressivi sono tutti significativi, mentre per le ore 1, 2, 3, 6 e 22, 23 e 24 il coefficiente β per la parte GARCH non è risultato significativo ai livelli usualmente considerati; per queste ore si è rispecificato il modello, stimando un AR(1)-ARCH(1). In tutte le serie dei residui non risultano presenti effetti GARCH, a seguito dell'introduzione della componente per l'eteroschedasticità.

5.2.2 Markov-Switching

Il secondo approccio considerato per la stima della componente stocastica è quello dei modelli *Markov-Switching*, particolarmente indicato per serie che presentino valori anomali. Nel caso dei dati in esame, come si è detto nelle sezioni precedenti, si evidenzia la presenza di *spikes* in particolari periodi dell'anno, che rendono interessante valutare la presenza di due regimi tra i quali si muovono i dati, uno di base e uno di picco. A tale scopo si sono applicati modelli *Markov-Switching* univariati, considerando per ogni serie un modello AR(1), con parametro autoregressivo che nel tempo può assumere valori diversi a seconda del regime di provenienza. Si è ipotizzato, inoltre, che anche la varianza possa assumere due diversi valori, dipendenti dai due diversi stati. Per fare ciò si è utilizzato il pacchetto `fMarkovSwitching` di R. L'applicazione del modello ai dati ha portato alle stime riprodotte in Tabella 5.3; per non appesantire la presentazione dei risultati le stime delle varianze sono riportate nell'Appendice C.

	<u>Stato 1</u>	<u>Stato 2</u>	<u>Durata</u>	<u>Matrice di transizione</u>
Ora 1	0.9587 (0.00)	0.4541 (0.00)	1:3.49 2:2.68	$\begin{bmatrix} 0.71 & 0.37 \\ 0.29 & 0.63 \end{bmatrix}$
Ora 2	0.7962 (0.00)	0.2109 (0.01)	1:17.90 2:2.01	$\begin{bmatrix} 0.94 & 0.50 \\ 0.06 & 0.50 \end{bmatrix}$
Ora 3	0.7883 (0.00)	0.1850 (0.02)	1:24.77 2:2.08	$\begin{bmatrix} 0.96 & 0.48 \\ 0.04 & 0.52 \end{bmatrix}$
Ora 4	0.6840 (0.00)	0.1023 (0.49)	1:80.45 2:1.71	$\begin{bmatrix} 0.99 & 0.58 \\ 0.01 & 0.42 \end{bmatrix}$
Ora 5	0.6735 (0.00)	0.1402 (0.12)	1:51.93 2:3.11	$\begin{bmatrix} 0.98 & 0.32 \\ 0.02 & 0.68 \end{bmatrix}$
Ora 6	0.8140 (0.00)	0.2056 (0.00)	1:21.51 2:2.21	$\begin{bmatrix} 0.95 & 0.45 \\ 0.05 & 0.55 \end{bmatrix}$
Ora 7	0.9232 (0.00)	0.2984 (0.00)	1:4.84 2:2.75	$\begin{bmatrix} 0.79 & 0.36 \\ 0.21 & 0.64 \end{bmatrix}$
Ora 8	0.9302 (0.00)	0.3793 (0.00)	1:3.62 2:3.03	$\begin{bmatrix} 0.72 & 0.33 \\ 0.28 & 0.67 \end{bmatrix}$
Ora 9	0.9300 (0.00)	0.5338 (0.00)	1:4.19 2:3.35	$\begin{bmatrix} 0.76 & 0.30 \\ 0.24 & 0.70 \end{bmatrix}$
Ora 10	0.9316 (0.00)	0.5658 (0.00)	1:7.72 2:3.38	$\begin{bmatrix} 0.87 & 0.30 \\ 0.13 & 0.70 \end{bmatrix}$

Ora 11	0.9420 (0.00)	0.6362 (0.00)	1:10.39 2:4.23	$\begin{bmatrix} 0.90 & 0.24 \\ 0.10 & 0.76 \end{bmatrix}$
Ora 12	0.9206 (0.00)	0.6207 (0.00)	1:16.62 2:5.29	$\begin{bmatrix} 0.94 & 0.19 \\ 0.06 & 0.81 \end{bmatrix}$
Ora 13	0.9347 (0.00)	0.6765 (0.00)	1:15.85 2:5.64	$\begin{bmatrix} 0.94 & 0.18 \\ 0.06 & 0.82 \end{bmatrix}$
Ora 14	0.9342 (0.00)	0.6635 (0.00)	1:11.73 2:4.79	$\begin{bmatrix} 0.91 & 0.21 \\ 0.09 & 0.79 \end{bmatrix}$
Ora 15	0.9343 (0.00)	0.6423 (0.00)	1:9.72 2:3.84	$\begin{bmatrix} 0.90 & 0.26 \\ 0.10 & 0.74 \end{bmatrix}$
Ora 16	0.9246 (0.00)	0.6370 (0.00)	1:6.78 2:3.27	$\begin{bmatrix} 0.85 & 0.31 \\ 0.15 & 0.69 \end{bmatrix}$
Ora 17	0.9223 (0.00)	0.6199 (0.00)	1:6.20 2:3.19	$\begin{bmatrix} 0.84 & 0.31 \\ 0.16 & 0.69 \end{bmatrix}$
Ora 18	0.9137 (0.00)	0.7577 (0.00)	1:10.10 2:4.44	$\begin{bmatrix} 0.90 & 0.23 \\ 0.10 & 0.77 \end{bmatrix}$
Ora 19	0.9231 (0.00)	0.7901 (0.00)	1:16.26 2:6.04	$\begin{bmatrix} 0.94 & 0.17 \\ 0.06 & 0.83 \end{bmatrix}$
Ora 20	0.9109 (0.00)	0.7467 (0.00)	1:16.47 2:3.74	$\begin{bmatrix} 0.94 & 0.27 \\ 0.06 & 0.73 \end{bmatrix}$
Ora 21	0.9013 (0.00)	0.6676 (0.00)	1:16.79 2:2.69	$\begin{bmatrix} 0.94 & 0.37 \\ 0.06 & 0.63 \end{bmatrix}$
Ora 22	0.9330 (0.00)	0.7648 (0.00)	1:3.20 2:3.64	$\begin{bmatrix} 0.69 & 0.27 \\ 0.31 & 0.73 \end{bmatrix}$
Ora 23	0.8281 (0.00)	0.0902 (0.66)	1:106.19 2:1.49	$\begin{bmatrix} 0.99 & 0.67 \\ 0.01 & 0.33 \end{bmatrix}$
Ora 24	0.8161 (0.00)	0.1616 (0.35)	1:125.35 2:2.16	$\begin{bmatrix} 0.99 & 0.46 \\ 0.01 & 0.54 \end{bmatrix}$

Tabella 5.3: Stima dei parametri del modello *Markov Switching*. In Colonna 1 e 2 si riportano i valori dei coefficienti autoregressivi dei due stati; tra parentesi i pvalue che indicano la significatività dei parametri ed in grassetto quelli maggiori di 0.01. In Colonna 3, si riportano le durate dei due regimi. In colonna 4 le stime della matrici di transizione tra i due regimi.

I coefficienti autoregressivi, relativi al primo regime (Colonna 1 della tabella) risultano sempre significativi, mentre per le ore 4, 5, 23 e 24 il coefficiente dello stato 2 non risulta significativamente diverso da zero, ad indicare che in queste ore non si verificano praticamente mai dei picchi. A conferma di ciò, si noti che, per queste ore, la probabilità di essere nello stato 1 e di rimanerci, p_{11} , è molto alta, pari a 0.99. Si osserva, inoltre, come in certe ore del giorno i due regimi abbiano durate molto simili, come accade dalle 07:00 alle 09:00. In altri momenti del giorno, invece, il regime 1 ha una durata

decisamente superiore rispetto a quella del regime 2. Solo in un'occasione, alle 22:00 il regime 2 ha una durata leggermente superiore a quella del regime 1.

Nel costruire modelli di questo tipo si potrebbe pensare di raggruppare le ore che evidenzino omogeneità di comportamento; nel contesto in esame non si otterrebbero miglioramenti significativi e per questo si è deciso di mantenere l'approccio univariato.

5.2.3 SUR

Nell'ambito dell'utilizzo di modelli multivariati, ai dati oggetto di studio si applica, per la stima della parte stocastica il modello SUR, che lavorando sulle 24 serie simultaneamente, tiene conto della correlazione esistente tra i termini di errore in equazioni successive nella fase di stima dei parametri. Anche in questo contesto si effettuano regressioni aventi come esplicative le serie ritardate, escludendo i termini d'intercetta che non risultano significativi. Le stime che si ottengono sui dati in esame sono riportate nella Tabella 5.4:

	α_h	$1-\alpha_h$
Ora 1	0.3513 (0.00)	0.6487
Ora 2	0.3274 (0.00)	0.6726
Ora 3	0.3279 (0.00)	0.6721
Ora 4	0.2620 (0.00)	0.7380
Ora 5	0.2311 (0.00)	0.7689
Ora 6	0.2505 (0.00)	0.7495
Ora 7	0.2517 (0.00)	0.7483
Ora 8	0.2985 (0.00)	0.7015
Ora 9	0.3464 (0.00)	0.6532
Ora 10	0.3471 (0.00)	0.6529
Ora 11	0.3817 (0.00)	0.6183
Ora 12	0.3886 (0.00)	0.6114

Ora 13	0.4124 (0.00)	0.5876
Ora 14	0.3864 (0.00)	0.6136
Ora 15	0.3856 (0.00)	0.6144
Ora 16	0.3660 (0.00)	0.6340
Ora 17	0.4199 (0.00)	0.5801
Ora 18	0.4930 (0.00)	0.507
Ora 19	0.5349 (0.00)	0.4651
Ora 20	0.5304 (0.00)	0.4696
Ora 21	0.5388 (0.00)	0.4612
Ora 22	0.5295 (0.00)	0.4705
Ora 23	0.4243 (0.00)	0.5757
Ora 24	0.3790 (0.00)	0.6232

Tabella 5.4: Stime dei coefficienti del modello SUR sulla serie originaria, tra parentesi i pvalue dei test per la significatività dei parametri.

I coefficienti sono tutti significativi. Nella seconda colonna della Tabella, si riporta la stima del parametro che indica la *mean reversion*, che si ottiene sottraendo ad uno i coefficienti considerati. Quello che si nota è che tutti i coefficienti di *mean reversion* relativi alle ore dalla 1:00 alle 16:00 sono maggiori di 0.6. Nelle ore successive, il valore dei coefficienti tende a ridursi fino alle 22:00. Nelle ultime due ore del giorno, i coefficienti tornano a livelli simili a quelli assunti dalla 1:00 alla 16:00. Questo tipo di comportamento è dovuto al fatto che il fenomeno della *mean reversion* non è stabile nel corso della giornata. In particolare, nelle ore definite di *super peak*, (come sono indicate da Huisman et al, 2006) dalle 18:00 alle 22:00, il fenomeno del ritorno alla media sembra meno evidente che nel resto della giornata e i dati mostrano una persistenza più forte. Questo importante risultato si spiega con il fatto che ad una domanda maggiore in queste ore, corrisponde una minore capacità di generare scorte e quindi una probabilità più elevata che ci siano picchi. Il risultato che si ottiene è che, generalmente, i prezzi in queste ore saranno meno facilmente predicibili.

5.2.4 VAR

L'ultimo modello che è stato di applicato ai dati è il modello VAR(p), che come spiegato in precedenza, rappresenta una semplice estensione del modello SUR. In questo contesto si è deciso di costruire un modello VAR(1) completo, in cui ognuna delle 24 ore fosse regredita sul proprio ritardo e sui ritardi delle restanti 23. il modello è stato quindi semplificato, mantenendo solo le variabili statisticamente significative. Il vantaggio di questo approccio, rispetto al più semplice modello SUR, è rappresentato dalla possibilità di valutare l'effetto che ha un'ora su quelle contigue.

La matrice di coefficienti, stimata sulla serie originaria è riportata nella Tabella 5.5:

Ora 1	1:00 0.32	3:00 0.17	6:00 -0.18	7:00 0.11	08:00 0.10	9:00 0.21	11:00 0.30	12:00 -0.20	21:00 0.07		
Ora 2	2:00 0.33	3:00 0.12	6:00 -0.15	7:00 0.13	08:00 0.11	9:00 0.17	10:00 0.28	12:00 -0.16	20:00 0.06		
Ora 3	3:00 0.37	7:00 0.05	08:00 0.15	Ora10 0.33	Ora11 0.25	12:00 -0.31	22:00 0.20	23:00 -0.15			
Ora 4	3:00 0.22	4:00 0.08	5:00 0.10	7:00 0.08	8:00 0.16	11:00 0.10	17:00 0.16	22:00 0.18	23:00 0.52	24:00 0.29	
Ora 5	2:00 -0.17	3:00 0.28	5:00 0.25	9:00 0.22	10:00 0.25	11:00 0.46	12:00 -0.41	16:00 -0.25	17:00 0.15	23:00 -0.07	
Ora 6	3:00 0.05	6:00 0.17	9:00 0.35	13:00 -0.09	21:00 0.19	22:00 0.17					
Ora 7	1:00 -0.19	6:00 0.09	7:00 0.11	9:00 0.16	13:00 -0.42	14:00 0.43	21:00 0.27	22:00 0.36			
Ora 8	1:00 -0.15	4:00 -0.11	6:00 0.09	8:00 0.07	9:00 0.18	13:00 -0.09	19:00 0.26	21:00 0.16	22:00 0.47		
Ora 9	2:00 -0.21	3:00 0.09	6:00 0.07	7:00 -0.09	9:00 0.20	12:00 0.14	13:00 -0.20	15:00 0.27	19:00 0.17	22:00 0.44	
Ora 10	2:00 -0.14	3:00 0.14	7:00 -0.08	9:00 0.15	10:00* -0.08	12:00 0.24	15:00 0.17	17:00 0.11	18:00 -0.11	19:00 0.16	22:00 0.35
Ora 11	2:00 -0.09	3:00 0.12	7:00 -0.11	9:00 0.11	10:00 -0.11	11:00 0.14	12:00 0.32	17:00 0.20	18:00 -0.10	19:00 0.10	22:00 0.32
Ora 12	3:00 0.05	7:00 -0.13	9:00 0.08	10:00 -0.08	12:00 0.54	17:00 0.22	18:00 -0.09	19:00 0.02	22:00 0.30		
Ora 13	3:00 0.08	5:00 -0.05	7:00 -0.08	12:00 0.20	13:00 0.31	17:00 0.20	18:00 -0.06	19:00 0.01	22:00 0.15	23:00 0.13	
Ora 14	3:00 0.07	7:00 -0.11	9:00 0.07	10:00 -0.15	12:00 0.19	13:00 0.17	14:00* 0.08	16:00 0.17	17:00 0.11	19:00 0.01	22:00 0.27

Ora 15	3:00 0.05	7:00 -0.11	9:00 0.07	10:00 -0.20	12:00 0.24	15:00 0.24	16:00 0.17	17:00 0.13	18:00 -0.08	19:00 0.07	22:00 0.31
Ora 16	3:00 0.02	7:00 -0.09	10:00 -0.14	12:00 0.26	16:00 0.30	17:00 0.15	19:00 0.04	22:00 0.34			
Ora 17	4:00 -0.04	7:00 -0.06	10:00 -0.08	12:00 0.21	17:00 0.31	19:00 0.26	20:00 -0.10	22:00 0.25	23:00 0.10		
Ora 18	3:00 0.07	5:00 -0.05	7:00 -0.07	18:00 0.40	19:00 0.31	22:00 0.21					
Ora 19	3:00 0.13	5:00 -0.05	7:00 -0.05	18:00 0.11	19:00 0.58	21:00 0.19					
Ora 20	1:00 0.04	3:00 0.14	5:00 -0.06	11:00 0.01	18:00 0.11	20:00 0.51	21:00 0.28	22:00 -0.14			
Ora 21	1:00 0.17	3:00 0.09	5:00 -0.09	6:00 -0.03	7:00 0.08	11:00 0.23	12:00 -0.19	20:00 0.22	21:00 0.27	22:00 0.24	24:00 -0.10
Ora 22	1:00 0.18	3:00 0.11	5:00 -0.08	6:00 -0.06	7:00 0.11	13:00 0.11	21:00 0.14	22:00 0.40	24:00 -0.05		
Ora 23	1:00 0.17	3:00 0.13	6:00 -0.09	11:00 0.25	12:00 -0.20	13:00 0.14	17:00 0.10	22:00 0.34	23:00* 0.04		
Ora 24	1:00 0.15	3:00 0.17	6:00 -0.10	11:00 0.38	12:00 -0.25	13:00 0.13	22:00 0.29	24:00* 0.04			

Tabella 5.5: Stime dei coefficienti del modello VAR sulla serie originaria, in grassetto i coefficienti relativi al ritardo della variabile dipendente. Con l'asterisco (*) sono indicati i coefficienti non significativi.

Si è deciso, per tutte le ore di inserire nell'equazione come esplicativa la variabile risposta ritardata, anche qualora non fosse risultata significativa, per continuità rispetto agli altri modelli. Ciò accade per le ore 10:00, 14:00, 23:00 e 24:00. Proprio per queste ore, però, risultano significativi i coefficienti delle ore immediatamente precedenti o successive. Inoltre si osserva che per ore consecutive risultano significative le stesse variabili, i cui coefficienti assumono valori simili. In particolare si nota che i ritardi delle ore 3:00, 12:00, 19:00 e 22:00 risultano significativi nella maggior parte delle equazioni. Per le ore per le quali risulta significativo il coefficiente del proprio ritardo, i risultati sono simili a quelli che si ottengono con il SUR. Anche per il VAR se si calcola il coefficiente di *mean reversion*, si nota che si ottiene un valore maggiore di 0.7 fino alle 17:00 e poi inferiore fino alle 22:00. il risultato è coerente con quanto evidenziato nel paragrafo precedente.

5.3 Analisi delle serie depurate dai picchi

Quando si stimano modelli per serie che presentano numerosi valori anomali si possono ottenere risultati distorti dai picchi. In molti degli studi citati, come ad esempio in Weron (2008) e Spagnolo e Uddin(2006) si propongono approcci alternativi al fine di risolvere questi problemi. Si può pensare di lavorare sulle serie dopo averle ripulite dai valori che superano una determinata soglia. In alternativa, poiché, come fa notare Weron, (2008), è la stima della componente deterministica a risentire della presenza di *jumps*, si può pensare di stimare tale componente sulla serie ripulita e poi sottrarla alla serie originaria. Seguendo Weron, (2008), si considerano in questo contesto entrambe le soluzioni. Si fissa la soglia T , come $T = \bar{\mu} + 3\sigma$, dopodichè si sostituiscono tutti i valori che eccedono T con $T + T \log_{10}(y_t / T)$.

La rappresentazione grafica della serie ripulita dei picchi, confrontata con quella originaria è riportata in Figura 5.1:

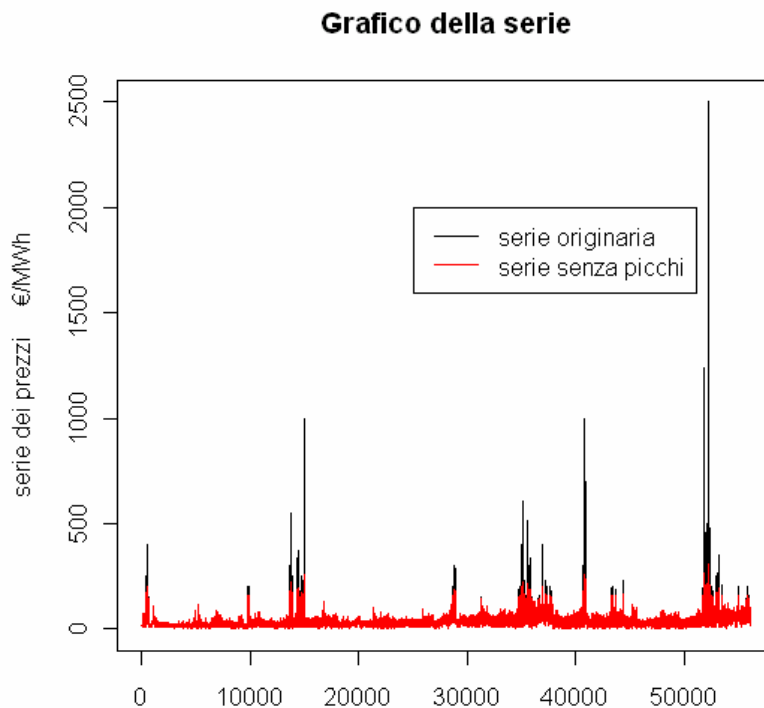


Figura 5.1: Grafico della serie ripulita dei picchi (in rosso), confrontata con la serie originale (in nero)

Sulla serie ottenuta si stima la componente deterministica. Si seguono quindi i due approcci: i) si lavora sulla serie così ripulita, ii) si stima la componente deterministica sulla serie ripulita e la si sottrae alla serie originaria.

Per la componente stocastica si possono considerare tre dei modelli proposti in precedenza: gli ARMA-GARCH, il VAR e il SUR.

5.3.1 ARMA - GARCH

Alla serie originaria, depurata della parte deterministica, stimata senza picchi, si applicano, come fatto in precedenza, modelli GARCH univariati alle 24 serie orarie. I risultati che si ottengono sono riportati nella Tabella 6.6:

	ϕ	α_0	α_1	β
Ora 1	0.7593 (0.00)	0.1093 (0.00)	0.1374 (0.00)	
Ora 2	0.7733 (0.00)	0.1165 (0.00)	0.2243 (0.00)	
Ora 3	0.7765 (0.00)	0.1154 (0.00)	0.2552 (0.00)	
Ora 4	0.67492 (0.00)	0.02810 (0.00)	0.06433 (0.00)	0.80095 (0.00)
Ora 5	0.58260 (0.00)	0.03147 (0.00)	0.22134 (0.00)	0.68239 (0.00)
Ora 6	0.7769 (0.00)	0.1070 (0.00)	0.3335 (0.00)	
Ora 7	0.67862 (0.00)	0.01255 (0.00)	0.08142 (0.00)	0.86844 (0.00)
Ora 8	0.66470 (0.00)	0.01335 (0.00)	0.09448 (0.00)	0.85524 (0.00)
Ora 9	0.726266 (0.00)	0.006143 (0.00)	0.079905 (0.00)	0.875587 (0.00)
Ora 10	0.86076 (0.00)	0.00808 (0.00)	0.24271 (0.00)	0.70909 (0.00)
Ora 11	0.887666 (0.00)	0.003109 (0.00)	0.149014 (0.00)	0.822451 (0.00)
Ora 12	0.884867 (0.00)	0.002463 (0.00)	0.130802 (0.00)	0.842943 (0.00)
Ora 13	0.895640 (0.00)	0.001972 (0.00)	0.129550 (0.00)	0.845230 (0.00)
Ora 14	0.878995 (0.00)	0.002058 (0.00)	0.133891 (0.00)	0.843404 (0.00)
Ora 15	0.881499 (0.00)	0.003899 (0.00)	0.181875 (0.00)	0.778435 (0.00)
Ora 16	0.862938 (0.00)	0.004494 (0.00)	0.165714 (0.00)	0.796474 (0.00)

Ora 17	0.836214 (0.00)	0.006586 (0.00)	0.147602 (0.00)	0.787327 (0.00)
Ora 18	0.872450 (0.00)	0.004459 (0.00)	0.112341 (0.009)	0.824889 (0.00)
Ora 19	0.865430 (0.00)	0.002519 (0.00)	0.113042 (0.00)	0.861117 (0.00)
Ora 20	0.866891 (0.00)	0.002525 (0.00)	0.082864 (0.00)	0.876118 (0.00)
Ora 21	0.90863 (0.00)	0.03312 (0.00)	0.43589 (0.00)	0.05239 (0.00)
Ora 22	0.88939 (0.00)	0.04002 (0.00)	0.39230 (0.00)	
Ora 23	0.85235 (0.00)	0.06094 (0.00)	0.22821 (0.00)	
Ora 24	0.84415 (0.00)	0.05372 (0.00)	0.24887 (0.00)	

Tabella 5.6: Stima dei modelli AR(1)-GARCH(1,1) sulle 24 serie orarie, per le quali la componente deterministica è stata stimata senza picchi; tra parentesi sono riportati i pvalue dei test t per la significatività dei parametri.

Anche in questo caso tutti i coefficienti autoregressivi sono significativi, mentre per alcune ore risulta non significativo il parametro β della parte GARCH. In particolare questo vale per le ore 1, 2, 3, 6, 22, 23 e 24. Per queste ore, analogamente a quanto fatto per la serie originaria, si è rispecificato il modello, stimando ARCH(1) per la varianza. Sulle serie dei residui si può applicare il test ARCH-lm, che porta a concludere che non ci sia eteroschedasticità di forma ARCH residua, dopo l'applicazione dei modelli GARCH.

Lo stesso tipo di modello si può applicare alla serie ripulita dei picchi nel modo descritto in precedenza. Si ottengono risultati riportati nella Tabella 5.7:

	ϕ	α_0	α_1	β
Ora 1	0.7597 (0.00)	0.1082 (0.00)	0.1426 (0.00)	
Ora 2	0.7630 (0.00)	0.1167 (0.00)	0.2268 (0.00)	
Ora 3	0.7681 (0.00)	0.1153 (0.00)	0.2600 (0.00)	
Ora 4	0.67610 (0.00)	0.02678 (0.00)	0.06645 (0.00)	0.80552 (0.00)
Ora 5	0.58224 (0.00)	0.03128 (0.00)	0.22113 (0.00)	0.68369 (0.00)
Ora 6	0.7766 (0.00)	0.1070 (0.00)	0.3306 (0.00)	
Ora 7	0.68104 (0.00)	0.01348 (0.00)	0.08038 (0.00)	0.86426 (0.00)
Ora 8	0.67408 (0.00)	0.01268 (0.00)	0.09571 (0.00)	0.85561 (0.00)
Ora 9	0.728464 (0.00)	0.005096 (0.00)	0.076660 (0.00)	0.885890 (0.00)

Ora 10	0.869125 (0.00)	0.009719 (0.00)	0.241005 (0.00)	0.677282 (0.00)
Ora 11	0.887402 (0.00)	0.002586 (0.00)	0.116556 (0.00)	0.851281 (0.00)
Ora 12	0.884564 (0.00)	0.001866 (0.00)	0.101956 (0.00)	0.874202 (0.00)
Ora 13	0.887610 (0.00)	0.001428 (0.00)	0.096376 (0.00)	0.880762 (0.00)
Ora 14	0.880138 (0.00)	0.001679 (0.00)	0.100593 (0.00)	0.874148 (0.00)
Ora 15	0.876080 (0.00)	0.003835 (0.00)	0.154619 (0.00)	0.795271 (0.00)
Ora 16	0.854182 (0.00)	0.004243 (0.00)	0.134616 (0.00)	0.816098 (0.00)
Ora 17	0.843610 (0.00)	0.006094 (0.00)	0.140348 (0.00)	0.792552 (0.00)
Ora 18	0.877835 (0.00)	0.003188 (0.00)	0.099351 (0.009)	0.851046 (0.00)
Ora 19	0.896518 (0.00)	0.002191 (0.00)	0.102443 (0.00)	0.860116 (0.00)
Ora 20	0.875891 (0.00)	0.002532 (0.00)	0.066093 (0.00)	0.883905 (0.00)
Ora 21	0.91081 (0.00)	0.03356 (0.00)	0.36148 (0.00)	0.05972 (0.00)
Ora 22	0.88844 (0.00)	0.04004 (0.00)	0.37660 (0.00)	
Ora 23	0.85049 (0.00)	0.06052 (0.00)	0.23482 (0.00)	
Ora 24	0.86710 (0.00)	0.04027 (0.00)	0.35601 (0.00)	

Tabella 5.7: Stima dei modelli AR(1)-GARCH(1,1) sulle 24 serie orarie; ripulite dei picchi. Tra parentesi sono riportati i pvalue dei test t per la significatività dei parametri.

Ancora una volta, i coefficienti autoregressivi sono tutti significativi, mentre per le ore 1, 2, 3, 6 e 22, 23 e 24 i coefficienti β per la stima della varianza non sono significativamente diversi da zero. Anche in questo contesto si è deciso di procedere, come visto in precedenza, stimando un modello AR(1)-ARCH(1) nei casi in cui il parametro β sia non significativo.

5.3.2 SUR

Lavorando sulla serie per la quale si è stimata la componente deterministica dopo aver eliminato i picchi (Colonna 1) e sulla serie ripulita dai *jumps* (Colonna 2) si ottengono i risultati riportati nella tabella 5.8:

	1	2
Ora 1	0.3525 (0.00)	0.3542 (0.00)
Ora 2	0.3285 (0.00)	0.3290 (0.00)
Ora 3	0.3289 (0.00)	0.3296 (0.00)
Ora 4	0.2627 (0.00)	0.2638 (0.00)
Ora 5	0.2315 (0.00)	0.2323 (0.00)
Ora 6	0.2506 (0.00)	0.2494 (0.00)
Ora 7	0.2515 (0.00)	0.2512 (0.00)
Ora 8	0.2979 (0.00)	0.3004 (0.00)
Ora 9	0.3458 (0.00)	0.3479 (0.00)
Ora 10	0.3469 (0.00)	0.3692 (0.00)
Ora 11	0.3817 (0.00)	0.4041 (0.00)
Ora 12	0.3886 (0.00)	0.4081 (0.00)
Ora 13	0.4125 (0.00)	0.4295 (0.00)
Ora 14	0.3863 (0.00)	0.4041 (0.00)
Ora 15	0.3853 (0.00)	0.4053 (0.00)
Ora 16	0.3657 (0.00)	0.3906 (0.00)
Ora 17	0.4194 (0.00)	0.4264 (0.00)
Ora 18	0.4928 (0.00)	0.4896 (0.00)
Ora 19	0.5349 (0.00)	0.5443 (0.00)
Ora 20	0.5308 (0.00)	0.5462 (0.00)
Ora 21	0.5395 (0.00)	0.5579 (0.00)
Ora 22	0.5305 (0.00)	0.5392 (0.00)
Ora 23	0.4245 (0.00)	0.4295 (0.00)
Ora 24	0.3770 (0.00)	0.3817 (0.00)

Tabella 5.8: Stima dei coefficienti del modello SUR sulla serie originaria, per la quale la componente deterministica è stata stimata senza picchi (Colonna 1) e sulla serie ripulita dei picchi (Colonna 2); tra parentesi i pvalue del test per la significatività dei parametri.

I coefficienti sono significativi per tutte e 24 le ore, in entrambe le situazioni. Anche in questo caso come per la serie originaria, si osserva che i coefficienti sono inferiori a 0.4 e quindi i loro complementi ad uno saranno maggiori di 0.6, fino all'ora 16:00 e per l'ora 24:00. Dalle 17:00 alle 23:00 il fenomeno della *mean reversion* è meno rilevante e la spiegazione di questo risultato è analoga a quella fornita per i dati originali.

5.3.3 VAR

Anche sulle serie dei dati ripulite dai picchi si è applicato il modello VAR(1), seguendo lo stesso criterio adottato per la serie originaria. I risultati che si ottengono sono coerenti con quelli presentati per i dati di partenza, quindi per una maggiore scorrevolezza del testo si è deciso di presentarli in appendice. Sia sulla serie senza *spikes* che su quella originaria, depurata della componente deterministica stimata senza picchi risultano non significativi i coefficienti relativi al proprio ritardo per le ore 10:00, 14:00, 23:00 e 24:00. Inoltre, per le ore per le quali i coefficienti sulla diagonale della matrice di parametri sono significativi, si ottengono risultati coerenti con quelli prodotti dal SUR. Per i commenti e le interpretazioni si veda il paragrafo 3.2.4.

5.4 Conclusioni

L'applicazione dei diversi modelli alle serie storiche dell'energia elettrica mette in luce le differenze di comportamento delle diverse ore del giorno: la *mean reversion* e la probabilità che si verifichino picchi sono fortemente legati alla domanda di energia, che varia nel corso della giornate, risultando inferiore nel corso della notte.

Inoltre, l'analisi delle serie ripulite dei picchi o per le quali la componente deterministica è stata stimata senza picchi porta a conclusioni simili a quelle che si ottengono per la serie originaria, per tutti i modelli proposti.

Capitolo 6

Previsioni

Nelle sezioni precedenti sono stati presentati i dati, i possibili modelli applicabili e i risultati che si sono ottenuti dalla stima degli stessi. Nel seguito si utilizzeranno i modelli proposti a fini previsivi. In particolare si è deciso di fare previsioni per quattro mesi dell'anno, uno per ogni stagione: Gennaio, Maggio, Agosto e Novembre, per tenere in considerazione periodi completamente diversi dal punto di vista climatico. Inoltre, si è deciso di costruire previsioni fino a 28 passi in avanti. Nel contesto dei prezzi dell'energia elettrica le previsioni sono particolarmente importanti sia perché il prezzo del giorno successivo viene stabilito in anticipo, sia perché sull'elettricità si costruiscono strumenti finanziari derivati, come *forward* e *futures*. Le previsioni ottenute con i diversi modelli sono state confrontate costruendo indicatori come l'errore medio assoluto (MAE) e l'errore quadratico medio (MSE).

Le previsioni ottenute sono state riorganizzate in due modi diversi. In primo luogo si sono considerate tutte le previsioni k passi in avanti, $k=1, \dots, 28$, e si sono confrontati i risultati ottenuti con la serie originaria, passo per passo. In un secondo momento poi, si è considerato di avere informazioni disponibili fino al tempo t e si sono costruite le serie dei valori previsti fino a 28 passi in avanti, che sono state confrontate con i primi 28 giorni del mese al quale erano riferite le previsioni.

6.1 Previsioni

Sui dati oggetto di studio sono stati stimati quattro diversi modelli, descritti ed analizzati in dettaglio nel Capitolo precedente. Con i modelli considerati sono state effettuate previsioni da 1 a 28 passi in avanti. Una volta ottenute le previsioni si sono costruiti indicatori che permettessero di effettuare confronti tra i diversi metodi previsivi. Si sono calcolati MSE e MAE, che rappresentano, rispettivamente la media degli errori di previsione al quadrato e la media dei loro valori assoluti. Per il modo in cui sono costruiti, un valore più basso di MSE o MAE indica un miglior adattamento del modello ai dati. Quindi, per valutare se le previsioni tendano a sovrastimare o sottostimare i veri valori, si sono calcolate anche le medie semplici degli errori di previsione. Inoltre, è opportuno verificare che le differenze che si riscontrano tra i diversi modelli, in termini previsivi, siano statisticamente significative. A tale fine si può calcolare il test di Diebold e Mariano, che confronta le serie di errori di previsione che si ottengono con modelli diversi. Se $\{y_t\}$ è la serie per la quale si vogliono calcolare le previsioni e $y_{t+k|t}^1$ e $y_{t+k|t}^2$ sono le previsioni ottenute con due metodi diversi per il valore di y_{t+k} , data l'informazione disponibile fino al tempo t , allora gli errori di previsione si ottengono nel seguente modo

$$\mathcal{E}_{t+k|t}^1 = y_{t+k} - y_{t+k|t}^1$$

$$\mathcal{E}_{t+k|t}^2 = y_{t+k} - y_{t+k|t}^2$$

Supponendo di costruire un totale di T_0 previsioni, allora si otterranno due serie di errori di previsione, $\{\mathcal{E}_{t+k|t}^1\}_{t_0}^T$ e $\{\mathcal{E}_{t+k|t}^2\}_{t_0}^T$, che saranno tra loro correlate. Definendo $L(\cdot)$ la funzione di perdita associata al modello, si vuole verificare l'ipotesi:

$$H_0 : E[L(\mathcal{E}_{t+k|t}^1)] = E[L(\mathcal{E}_{t+k|t}^2)]$$

contro l'alternativa che

$$H_1 : E[L(\mathcal{E}_{t+k|t}^1)] \neq E[L(\mathcal{E}_{t+k|t}^2)]$$

Il test di Diebold e Mariano si basa sulla quantità

$$d_t = L(\mathcal{E}_{t+k|t}^1) - L(\mathcal{E}_{t+k|t}^2)$$

ed è pari a

$$S = \frac{\bar{d}}{(\widehat{av\hat{r}(\hat{d})})^{1/2}}$$

con \bar{d} che indica la media dei d_t . Sotto l'ipotesi nulla S si distribuisce asintoticamente come una variabile normale standard, $S \sim N(0,1)$.

Di seguito sono riportati i grafici relativi alle previsioni ad 1 e 28 passi e la tabelle che riassumono gli indicatori calcolati per i quattro modelli proposti e relativi ai mesi di Gennaio, Maggio, Agosto e Novembre, per previsioni 1, 2, 3, 7, 14, 21 e 28 passi in avanti. Nell'Appendice B sono riportati i risultati delle previsioni ottenute sulla serie ripulita dei picchi e sulla serie per la quale la componente deterministica è stata stimata senza *spikes*. In entrambe le situazioni si ottengono risultati coerenti con quelli relativi alla serie originaria. Nell'Appendice C, invece, sono inseriti i grafici relativi alle previsioni 2, 3, 7, 14 e 21 passi in avanti. Nei grafici seguenti, si riporta in nero la serie originaria, in rosso la serie di previsioni ottenute con il modello GARCH, in verde quelle ottenuto con il SUR, in blu quelle calcolate con il *Markov Switching* ed in azzurro quelle calcolate con il VAR.

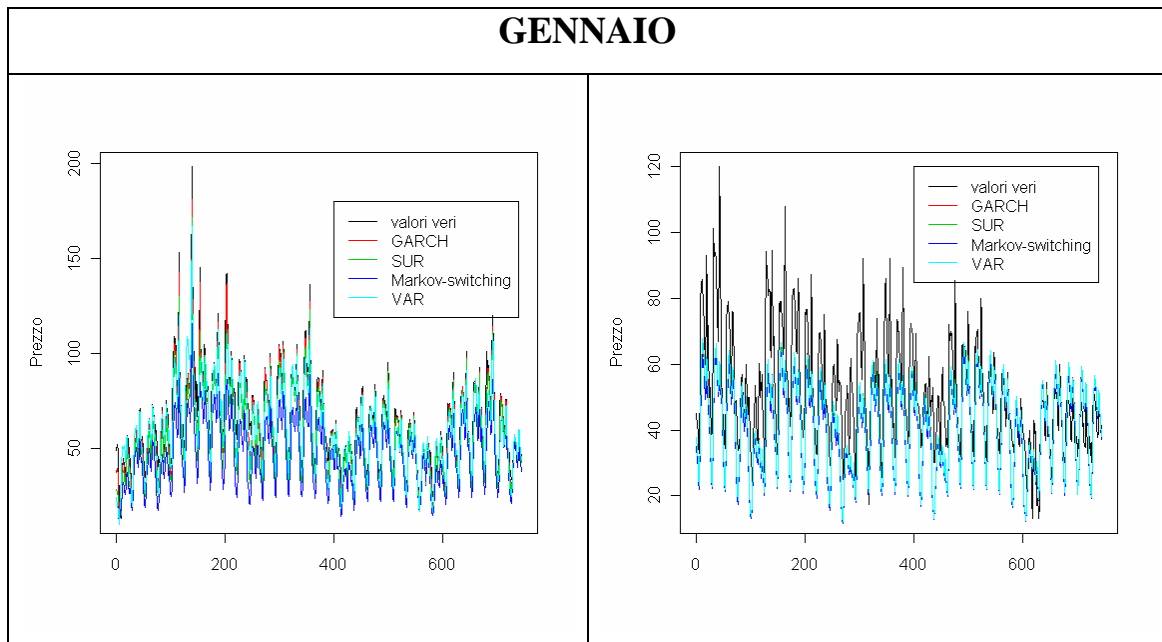


Figura 6.1: Grafico delle previsioni 1 e 28 passi in avanti, confrontate con al serie di prezzi rilevati per il mese di Gennaio

	ARMA-GARCH			SUR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	29.55	3.56	3.46	367.21	16.12	16.08
2	239.96	11.69	6.34	635.03	20.97	20.90
3	310.10	13.55	9.51	760.51	23.39	23.35
7	389.99	15.22	14.38	703.05	22.67	22.65
14	323.09	14.79	14.49	440.49	18.00	17.97
21	254.49	13.04	12.74	303.82	14.60	14.45
28	229.45	12.02	10.41	249.76	12.65	11.27

	MARKOV			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	65.95	6.12	6.10	87.75	6.70	3.30
2	267.35	12.33	9.50	210.20	10.56	4.41
3	370.94	14.82	12.83	252.40	12.00	6.31
7	481.59	17.36	17.13	263.79	12.29	10.49
14	371.16	16.13	16.02	253.82	12.61	12.61
21	278.50	13.85	13.64	216.02	11.71	11.21
28	242.69	12.42	10.92	208.56	11.33	9.53

Tabella 6.1: Indicatori di bontà previsiva per i quattro modelli proposti, relativi al mese di Gennaio. In grassetto gli indicatori er il modello migliore.

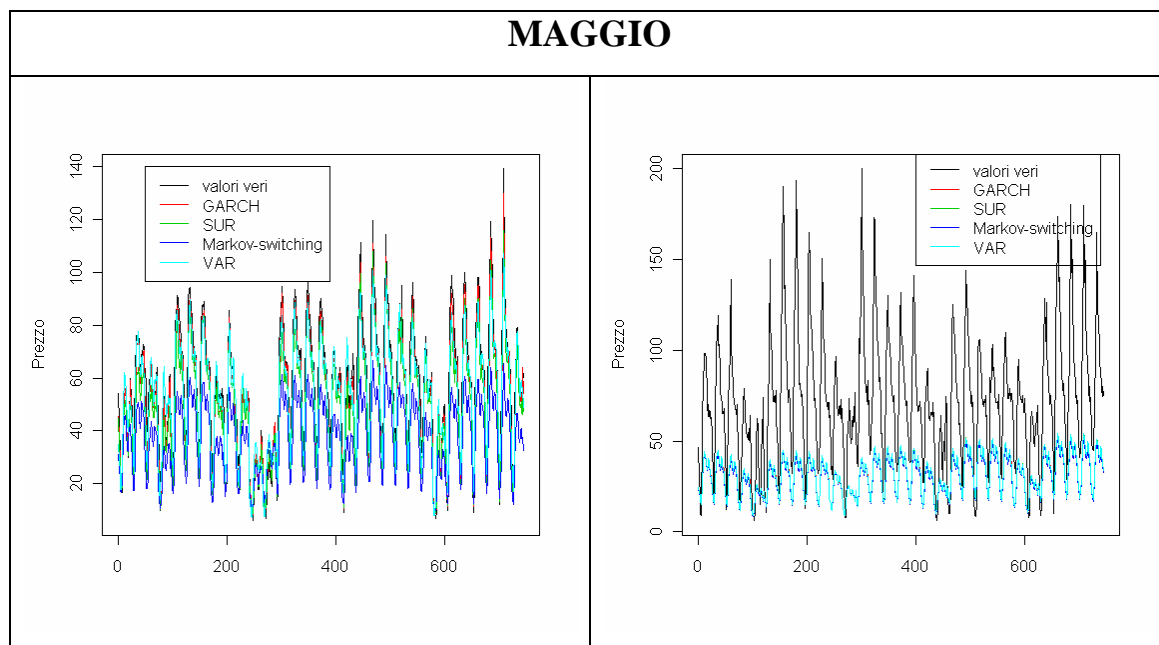


Figura 6.2: Grafico delle previsioni 1 e 28 passi in avanti, confrontate con al serie di prezzi rilevati per il mese di Maggio

	ARMA-GARCH			SUR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	35.11	4.50	4.11	518.07	19.79	19.20
2	278.61	13.02	9.63	874.05	25.96	24.79
3	427.35	16.89	13.36	1032.38	28.09	26.90
7	983.78	25.95	23.85	1515.11	32.39	31.31
14	1761.50	35.60	35.32	2026.61	38.27	38.00
21	1973.39	37.40	36.86	2084.30	38.49	37.95
28	2410.69	41.66	41.04	2464.93	42.13	41.51

	MARKOV			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	79.22	7.52	7.24	70.24	6.82	2.8
2	346.26	15.19	12.66	216.55	11.64	6.77
3	513.98	19.01	16.23	327.51	14.6	9.8
7	1083.99	27.41	25.7	855.64	23.33	20.24
14	1832.75	36.49	36.2	1621.02	33.54	33.11
21	2009.38	37.8	37.26	1889.02	36.31	35.65
28	2430.01	41.84	41.22	2357.89	41.07	40.38

Tabella 6.2: Indicatori di bontà previsiva per i quattro modelli proposti, relativi al mese di Maggio. In grassetto gli indicatori per il modello migliore.

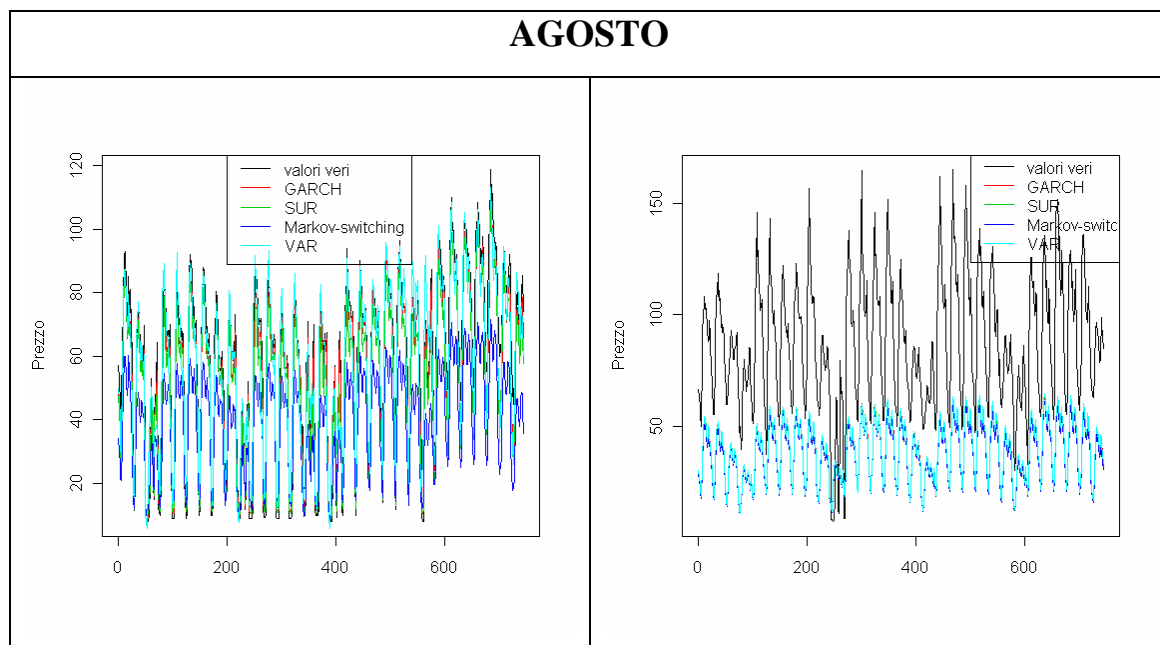


Figura 6.3: Grafico delle previsioni 1 e 28 passi in avanti, confrontate con la serie di prezzi rilevati per il mese di Agosto

Passi	ARMA-GARCH			SUR		
	MSE	MAE	Media	MSE	MAE	Media
1	41.88	4.86	4.31	549.30	20.67	19.54
2	292.17	13.43	9.45	1000.58	27.90	25.96
3	405.56	16.74	13.19	1209.33	30.75	28.87
7	887.38	26.51	24.85	1577.65	35.67	34.24
14	1618.18	36.81	35.55	1981.75	40.67	39.51
21	2300.28	45.06	44.45	2496.18	46.87	46.29
28	2686.00	48.97	48.45	2784.56	49.83	49.32

Passi	MARKOV			VAR		
	MSE	MAE	Media	MSE	MAE	Media
1	79.11	7.39	6.85	98.42	7.43	0.57
2	334.52	15.02	12.16	215.62	11.91	4.76
3	465.48	18.51	15.9	282.36	14.16	8.12
7	986.63	28.2	26.66	656.24	23.02	20.22
14	1704.49	37.82	36.57	1412.17	34.3	32.57
21	2352.81	45.59	45	2143.79	43.37	42.67
28	2715.73	49.24	48.73	2586.15	47.98	47.44

Tabella 6.3: Indicatori di bontà previsiva per i quattro modelli proposti, relativi al mese di Agosto. In grassetto gli indicatori per il modello migliore.

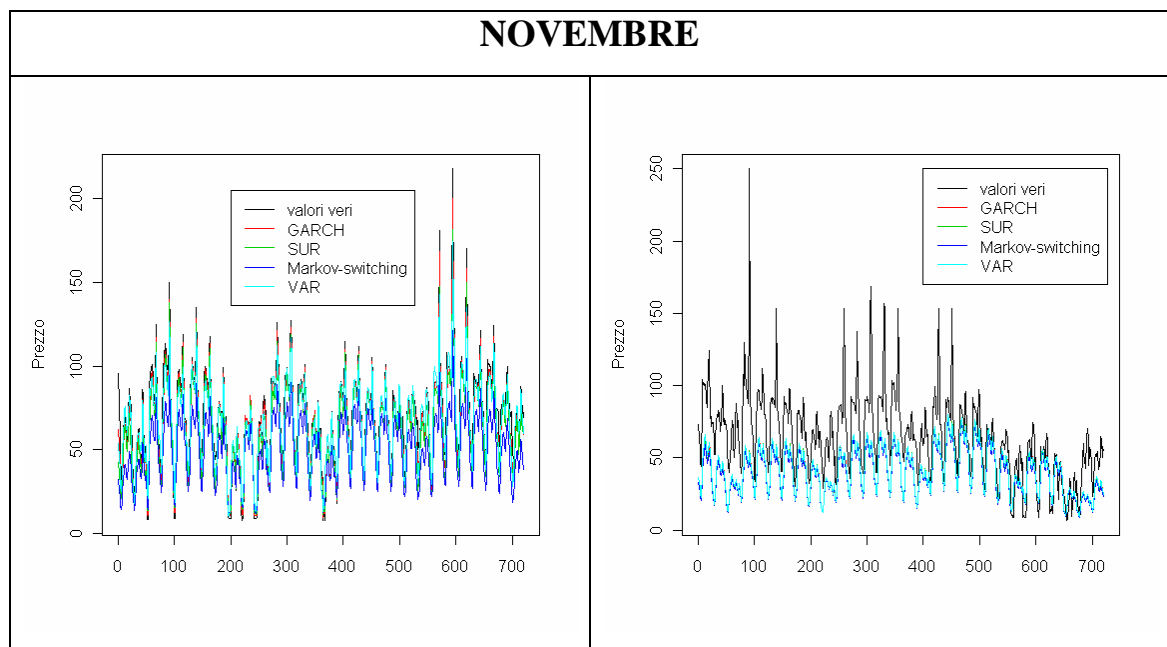


Figura 6.4: Grafico delle previsioni 1 e 28 passi in avanti, confrontate con al serie di prezzi rilevati per il mese di Novembre

	ARMA-GARCH			SUR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	38.25	4.26	3.95	499.43	19.39	18.89
2	299.62	13.05	7.78	930.53	25.93	25.05
3	440.04	15.96	10.91	1121.48	28.68	27.83
7	654.27	20.32	19.01	1195.91	29.83	29.22
14	1036.51	27.15	26.96	1320.89	31.65	31.50
21	1169.00	28.64	28.56	1308.76	30.77	30.80
28	936.87	24.89	23.79	995.01	25.80	24.77

	MARKOV			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media
1	85.31	7.18	6.86	80.73	6.41	2.22
2	338.4	14.02	11.16	234.38	11.04	3.95
3	487.76	17.24	14.5	331.82	13.38	5.85
7	752.6	22.54	21.69	497.32	16.37	13.75
14	1124.55	28.74	28.59	878.73	23.78	23.43
21	1221.38	29.58	29.52	1068.79	26.65	26.46
28	961.46	25.34	24.27	888.98	24	22.74

Tabella 6.4: Indicatori di bontà previsiva per i quattro modelli proposti, relativi al mese di Novembre. In grassetto gli indicatori per il modello migliore.

La prima osservazione che si può fare sulla base delle statistiche riportate è che per ogni modello la media degli errori di previsione è positiva. Poiché gli errori sono stati calcolati, come detto in precedenza, sottraendo il valore previsto a quello vero, allora tutte le previsioni tendono a sottostimare i valori che si sono effettivamente realizzati. In secondo luogo si può notare che per tutti e quattro i mesi scelti per i confronti previsivi, ad un passo, il modello ARMA-GARCH sembra prevedere meglio i valori futuri, sia in termini di MSE che di MAE. Considerando le previsioni a più passi, invece, le *performance* migliori si ottengono applicando il modello VAR(1). In tutti i casi, inoltre, le previsioni prodotte con il SUR sembrano essere le più distanti dai valori effettivamente realizzati. Aumentando il numero di passi per le previsioni, si nota, inoltre, che i quattro modelli proposti portano ad indicatori sempre più simili tra loro. Per valutare se le differenze riscontrate fra le previsioni prodotte da modelli diversi siano statisticamente significative, si sono calcolati i test di Diebold e Mariano, confrontando i modelli a coppie. L'applicazione del test ai dati porta ad accettare l'ipotesi nulla che i due modelli considerati siano equivalenti ai fini del calcolo delle previsioni solo in pochi casi. Considerando la soglia di significatività dell'1%, per il mese di Gennaio, a due passi, gli errori di previsione ottenuti con il modello *Markov*

Switching e con il modello ARMA-GARCH non risultano significativamente diversi. Il p-value del test di Diebold e Mariano, con alternativa bilaterale, è pari a 0.011. Sempre per la previsione a due passi relativa a Gennaio, il modello ARMA-GARCH risulta equivalente, in termini previsivi, al modello VAR; il p-value, in questo caso, è pari a 0.030. Per il mese di Maggio, invece, il test porta ad accettare l'ipotesi nulla per la previsione ad un passo effettuata con i modelli VAR e *Markov Switching*, riportando un p-value di 0.011. Infine, a Novembre gli errori di previsione ad un passo, ottenuti con i VAR e il *Markov Switching* non sono significativamente diversi: in questo caso il p-value del test è pari a 0.39 e l'ipotesi nulla sarebbe accettata anche se si scegliesse come soglia il livello di 0.05.

In tutte le altre situazioni il calcolo del test di Diebold e Mariano porta a rifiutare l'ipotesi che i due modelli siano equivalenti, contro un'alternativa bilaterale, a livello di significatività dell'1%. Si osserva che il modello SUR è significativamente peggiore degli altri in tutti i casi analizzati. Se si calcola, infatti, il test, con ipotesi alternativa che la funzione di perdita calcolata per gli errori di previsione ottenuti con il SUR sia maggiore di quella del modello alternativo, si rifiuta in tutti i casi l'ipotesi nulla.

6.2 Previsioni un mese in avanti

Per confrontare in altro modo le previsioni provenienti dai diversi modelli proposti, si è pensato di considerare di avere informazione disponibile fino all'istante t , ad esempio fino al 31 Dicembre del 2008 e di voler fare previsioni da quell'istante per le quattro settimane (28 giorni) successive. Le previsioni così ottenute sono state confrontate con i prezzi effettivamente registrati nei 28 giorni successivi all'istante t . Anche in questo caso si sono presi in considerazione quattro mesi differenti, uno per ogni stagione e si sono calcolati, a partire dagli errori di previsione, MSE, MAE e media.

Di seguito si riportano i grafici relativi ai quattro mesi considerati e le tabelle contenenti le statistiche riassuntive. Per quanto riguarda i grafici seguenti la rappresentazione dei risultati è la stessa utilizzata nel paragrafo precedente.

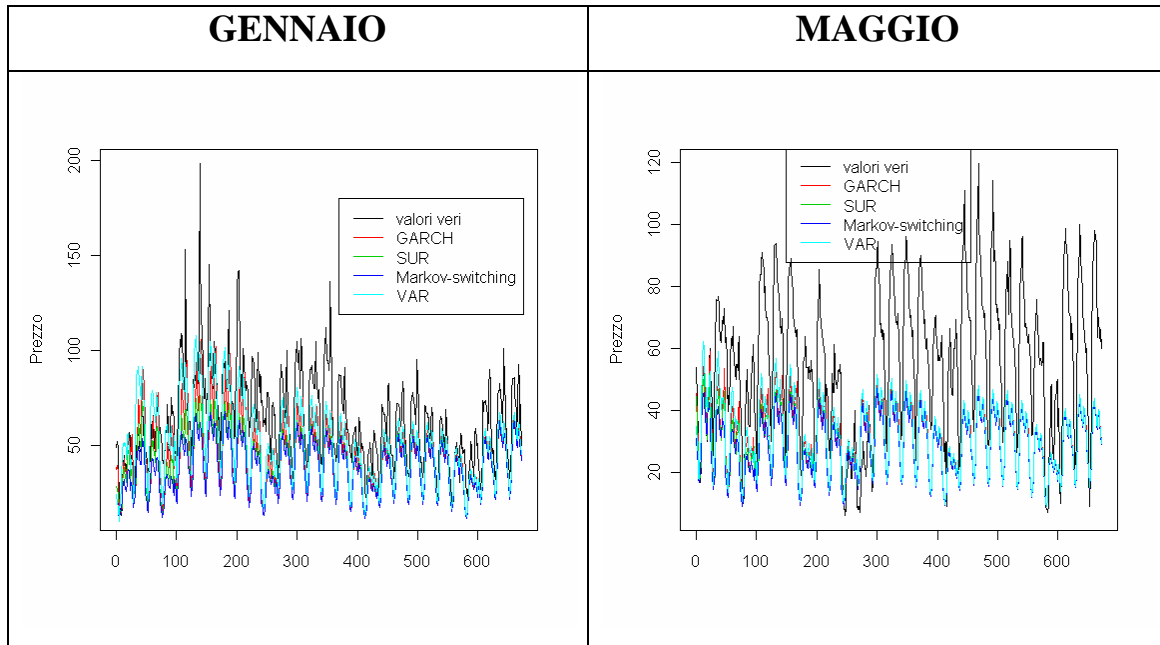


Figura 6.4: Grafico dei previsioni per i 28 giorni successivi all'ultima informazione disponibile, relativi ai mesi di Gennaio (a sinistra) e Maggio (a destra)

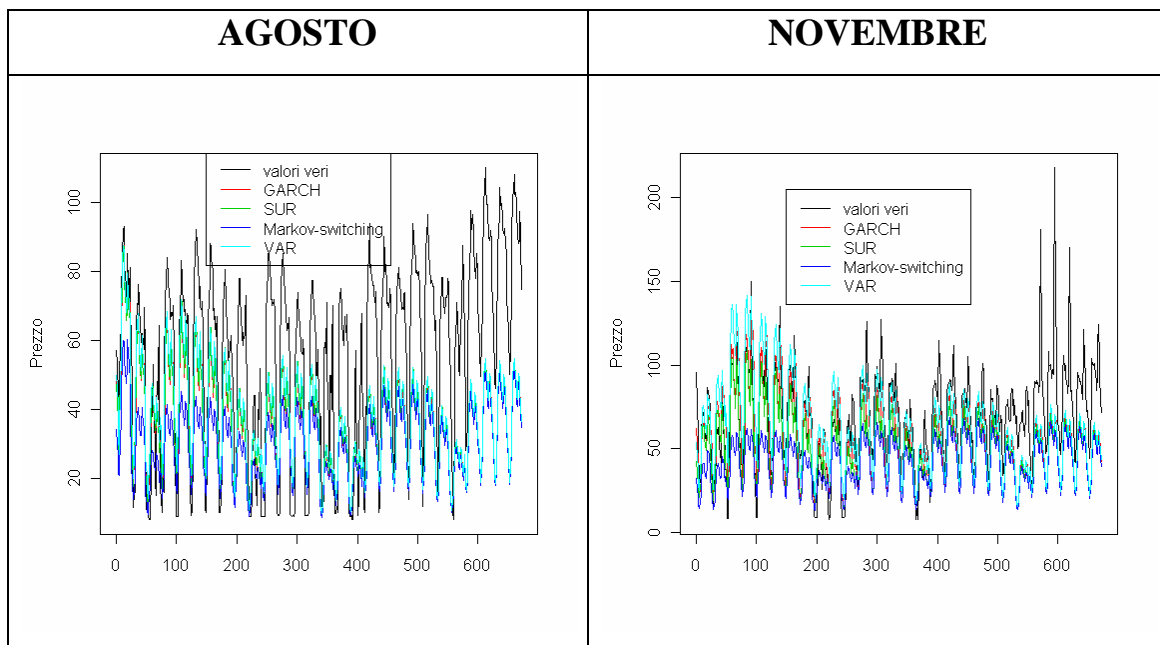


Figura 6.5: Grafico dei previsioni per i 28 giorni successivi all'ultima informazione disponibile, relativi ai mesi di Agosto (a sinistra) e Novembre (a destra)

	ARMA-GARCH			SUR			MARKOV			VAR		
	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
Genn	583.13	19.51	18.22	865.51	25.13	25.09	768.65	23.41	23.23	366.17	15.35	12.34
Mag	898.57	26.09	24.89	995.23	27.85	26.74	955.37	27.06	25.96	837.02	25.3	23.55
Ago	799.87	24.12	21.28	1038.19	28.47	26.39	793.85	24.06	21.09	724.29	22.87	18.63
Nov	655.27	19.86	14.53	1094.99	28.50	27.63	765.19	22.18	20.35	552.43	17.95	6.06

Tabella 6.5: Indicatori di bontà previsiva per i quattro modelli proposti, per i quattro mesi considerati. In grassetto gli indicatori per il modello migliore.

Dall'analisi dei dati riportati in tabella emerge che, sia confrontando le previsioni sulla base del MSE che del MAE per tutti e quattro i mesi considerati il modello che meglio riesce a prevedere i prezzi è il VAR(1). Anche da queste analisi emerge che il SUR sembra essere il peggiore dei modelli, in termini previsivi, poiché in tutti i casi gli indicatori considerati assumono i valori massimi. Come nel paragrafo precedente, anche in questo caso si osserva che i metodi considerati tendono a sottostimare i valori effettivamente osservati: in tutti i mesi e per tutti i modelli la media degli errori di previsione è maggiore di zero. Anche in questo contesto si è deciso di valutare se le differenze esistenti tra i modelli fossero o meno significative. L'applicazione del test di Diebold e Mariano porta a rifiutare per tutti e quattro i mesi considerati, che ci siano due modelli equivalenti in termini di efficacia previsiva.

6.3 Conclusioni

L'applicazione dei quattro modelli proposti ai dati analizzati ha evidenziato che per le previsioni ad un passo dei modelli che lavorano sulle 24 serie orarie, considerate singolarmente, come gli ARMA GARCH sembrano prevedere bene i prezzi dell'energia elettrica e fare meglio di modelli multivariati, quando si vogliono previsioni di brevissimo periodo. Aumentando l'orizzonte previsivo, però i modelli di tipo VAR, che permettono di considerare le 24 serie simultaneamente, tenendo conto della correlazione esistente tra termini d'errore di equazioni consecutive e dell'effetto che alcune ore possono avere sulle ore contigue, prevedono meglio di quelli univariati. Inoltre, si è osservato che nel confrontare i due modelli univariati proposti, gli ARMA-GARCH tendono a fare meglio dei modelli *Markov Switching*, che considerano due diversi regimi. Queste evidenze suggeriscono un possibile sviluppo dei modelli proposti, che permetta di tener conto sia della correlazione tra le ore che della possibilità che i dati si muovano tra due regimi, uno di base e uno di picco. A tale scopo sarebbe interessante stimare modelli *Markov Switching* multivariati, con la possibilità di considerare sottogruppi di ore con comportamento omogeneo.

Un ulteriore sviluppo che si potrebbe prendere in considerazione è la combinazione di previsioni generate da modelli diversi, per valutare se sia possibile ottenere risultati migliori. Questo tipo di approccio è stato sviluppato negli ultimi decenni; ad esempio Granger e Newbold nel 1986 mostrarono come la combinazione di due previsori non distorti ad un passo in avanti, ottenuta attraverso una semplice media ponderata, portasse ad errori di previsione con varianza minore di quella degli errori ottenuti utilizzando i due metodi separatamente. Per maggiori informazioni sulla combinazione di previsioni si vedano Clemen (1989) e Granger e Ramanathan (1984).

Conclusioni

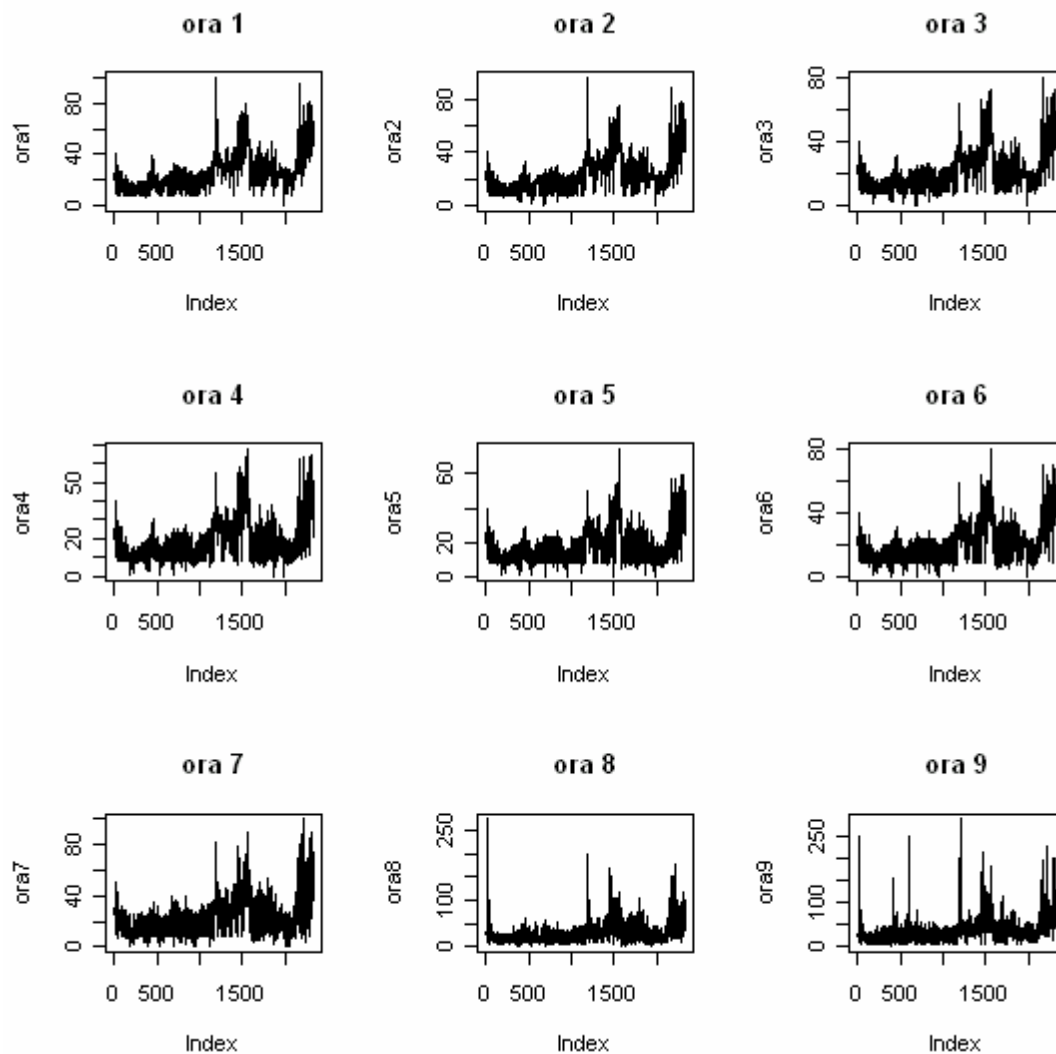
Negli ultimi decenni ha avuto luogo la liberalizzazione della fornitura di energia elettrica e conseguentemente sono nati in tutta Europa e nel resto del mondo, mercati sui quali vengono negoziati i prezzi dell'elettricità, ottenuti dall'incrocio tra domanda e offerta, e gli strumenti finanziari derivati, che hanno come sottostante l'energia. Dopo la liberalizzazione, ha assunto un interesse sempre maggiore lo studio dei prezzi dell'energia elettrica, che mostrano caratteristiche che li contraddistinguono da tutti gli *assets* finanziari. Molti studiosi negli ultimi decenni hanno cercato di individuare un modello che potesse cogliere tutte le peculiarità delle serie storiche dei prezzi dell'elettricità, come la *mean reversion*, la stagionalità e la presenza di *spikes*.

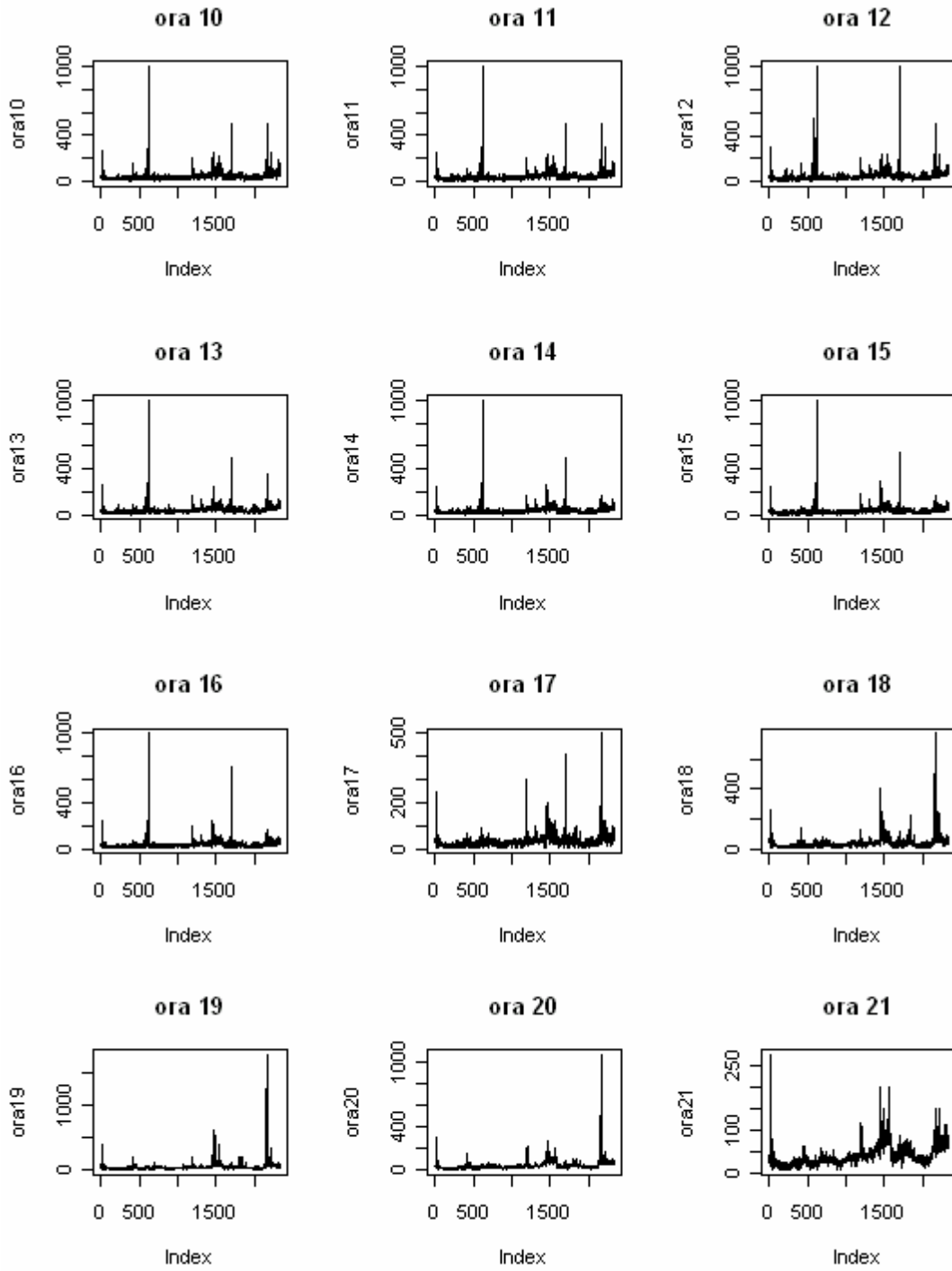
In questo lavoro si è proposta l'applicazione di alcuni modelli univariati, come gli ARMA-GARCH e i *Markov-Switching*, e multivariati, come i SUR e i VAR, ai dati reali del mercato dell'energia elettrica francese. Inoltre si è concentrata l'attenzione sulla capacità previsiva dei modelli proposti, perché nel mercato dell'energia elettrica i prezzi per le 24 ore del giorno successivo vengono stabiliti simultaneamente ad una data ora del giorno precedente. L'applicazione dei modelli a fini previsivi ha indicato che nel prevedere un passo in avanti, quindi per le 24 ore del giorno successivo, modelli univariati, di tipo ARMA-GARCH, permettono di ottenere i risultati migliori. Se l'interesse, invece, è quello di effettuare previsioni con un orizzonte più ampio, al fine di stabilire strategie di sviluppo e di pianificare la gestione, da parte delle aziende coinvolte nel servizio di fornitura, è più adeguato utilizzare modelli multivariati, che considerando le 24 ore simultaneamente, permettano di tener conto delle correlazioni esistenti tra di loro.

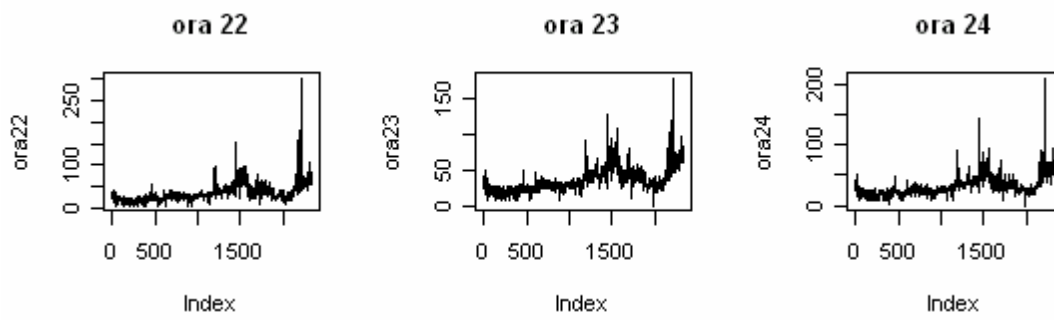
Appendice A

Analisi grafiche delle serie orarie

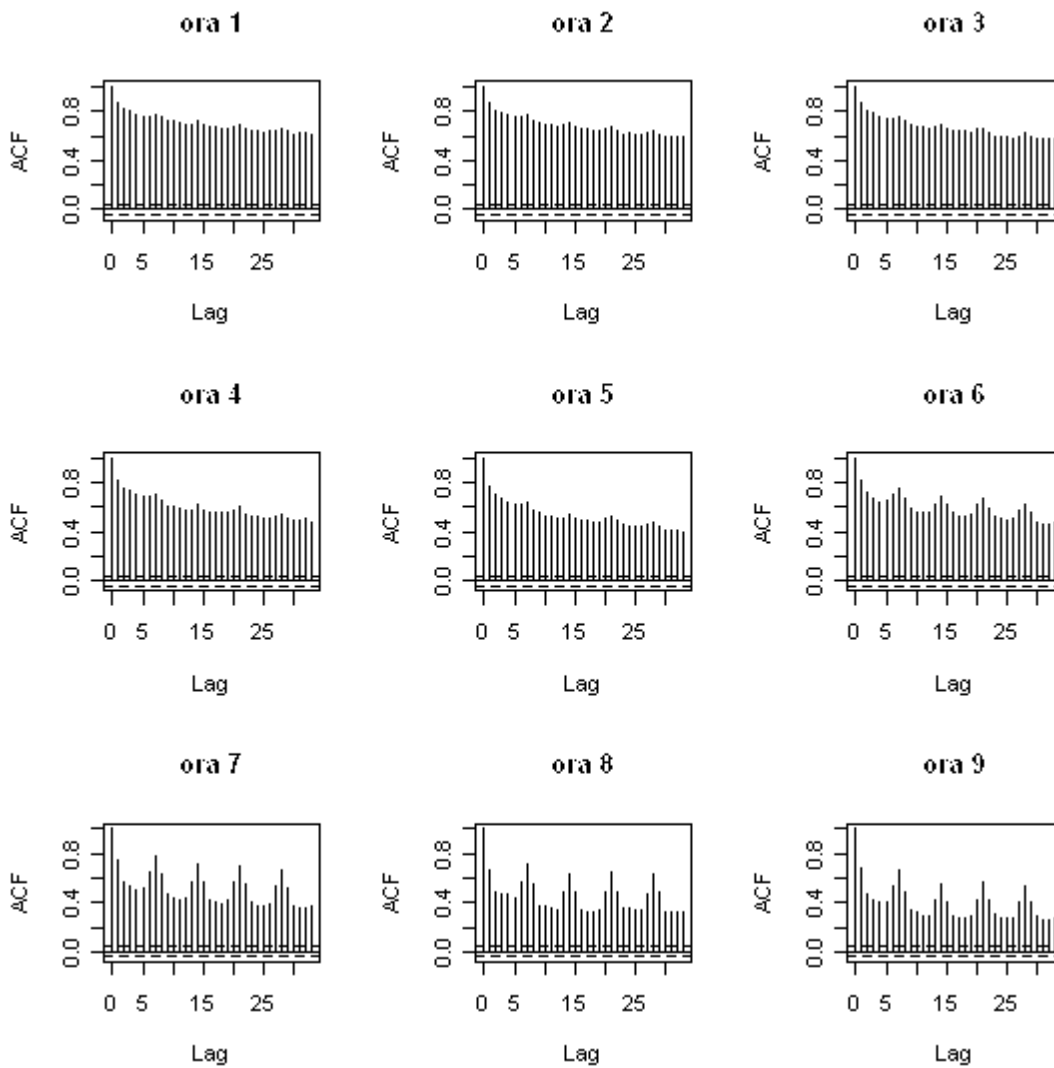
Grafici delle singole serie orarie

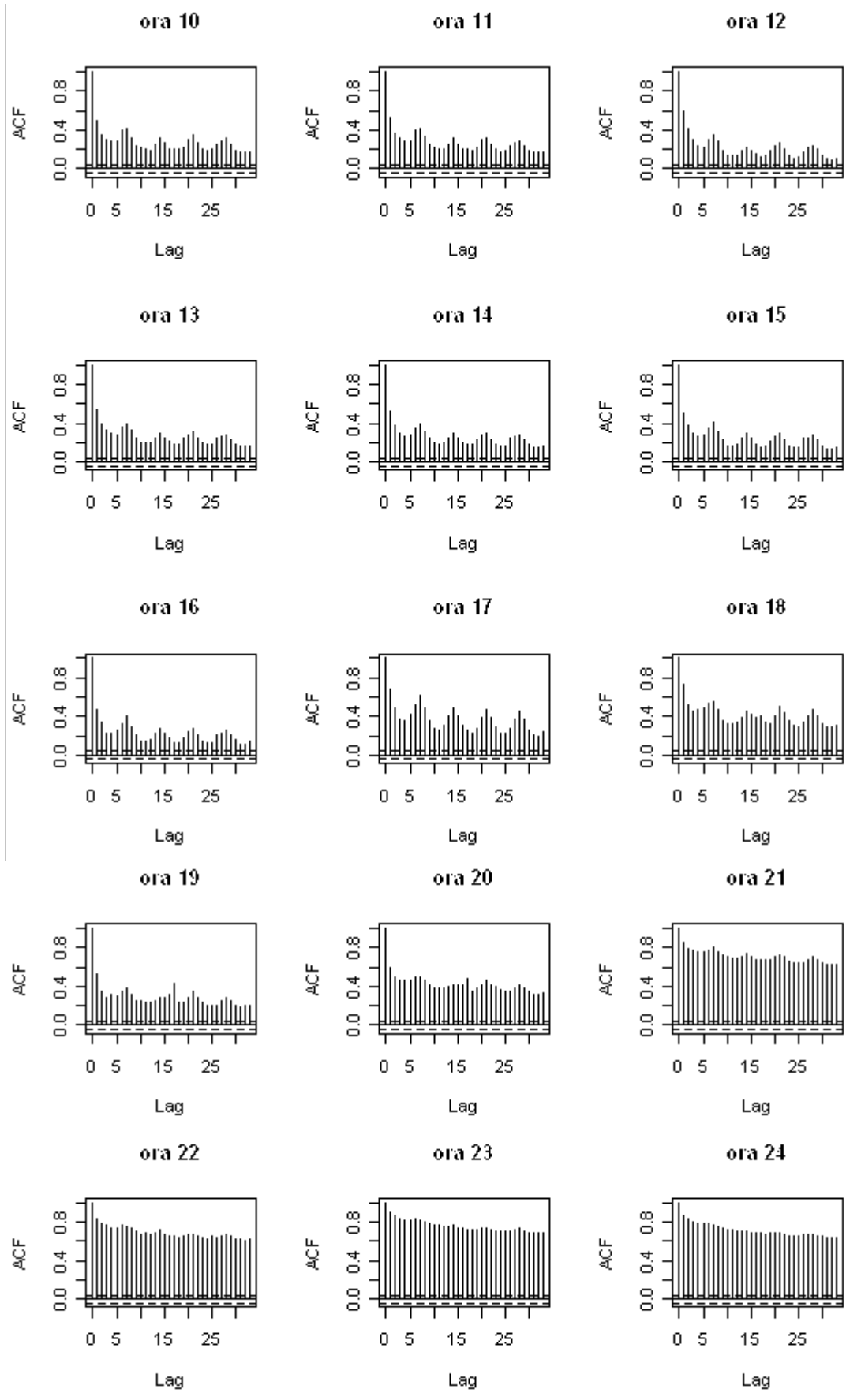




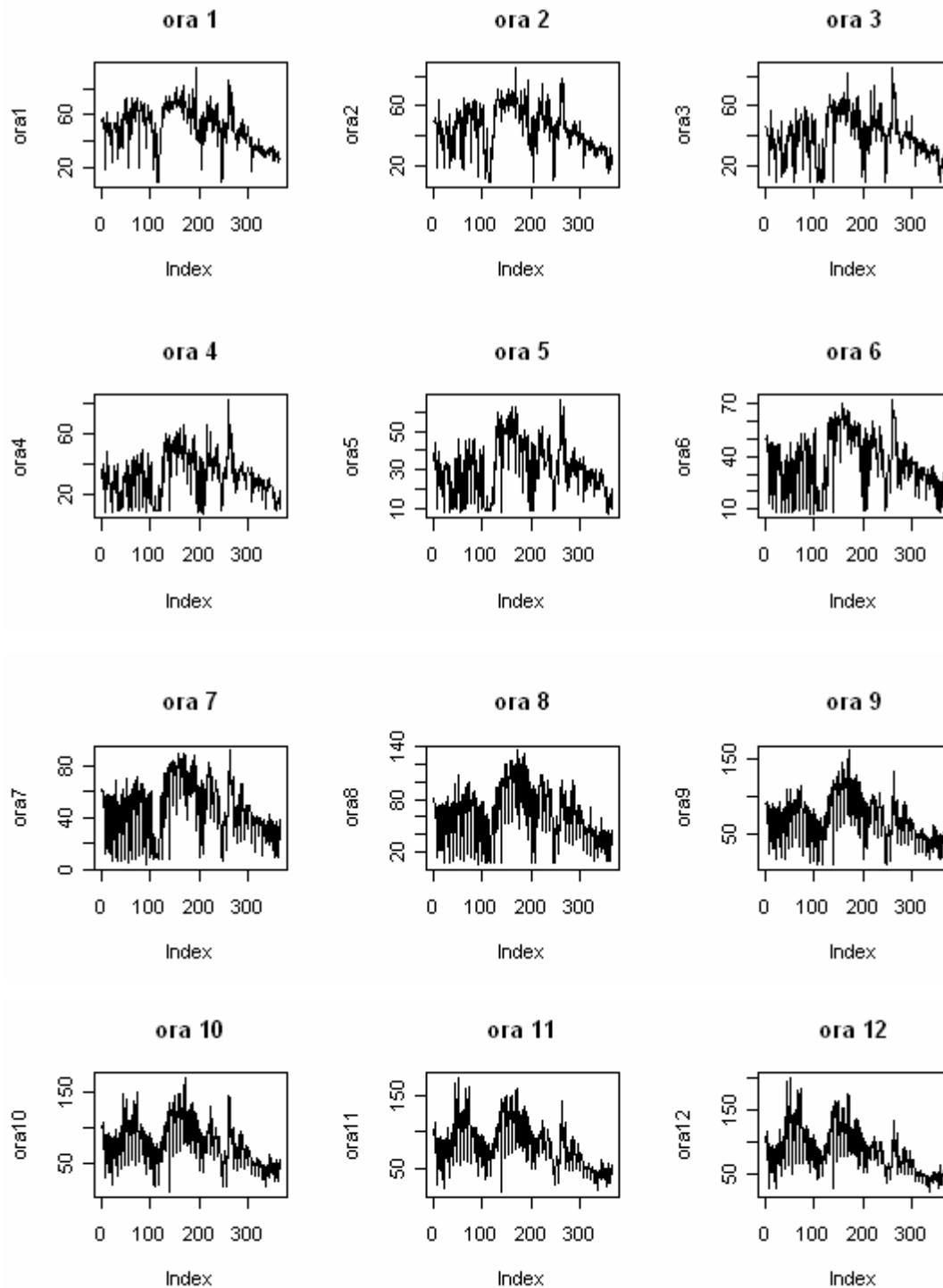


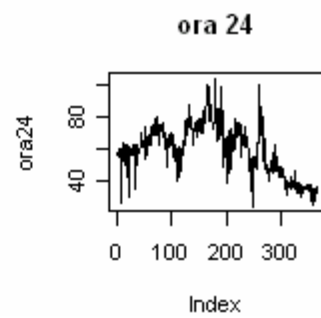
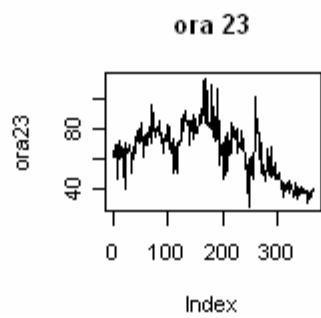
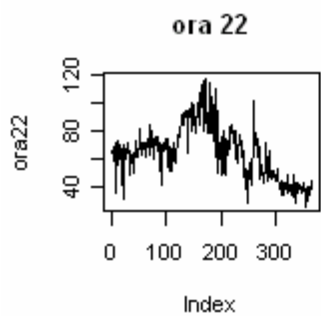
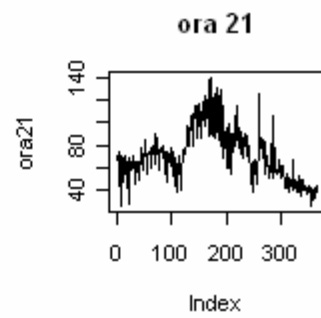
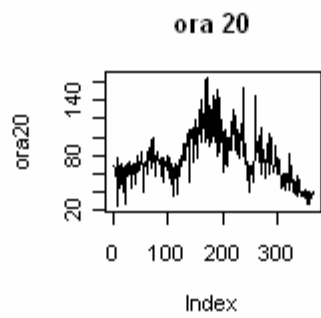
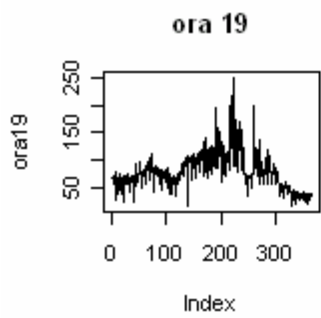
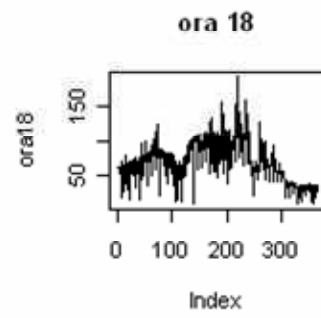
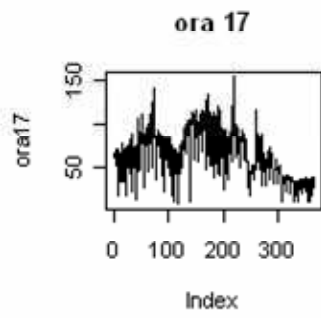
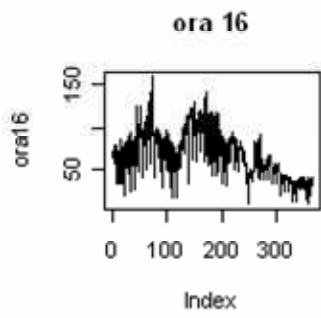
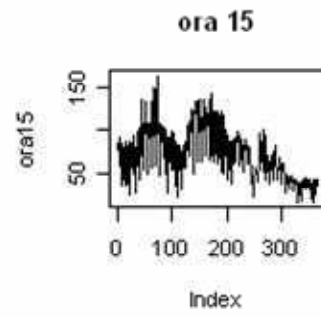
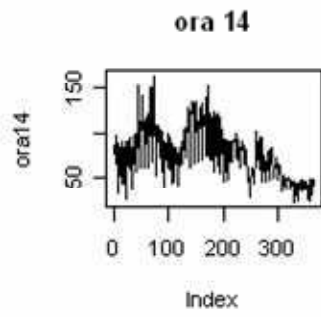
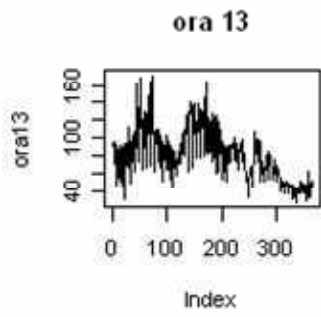
Funzioni di autocorrelazione



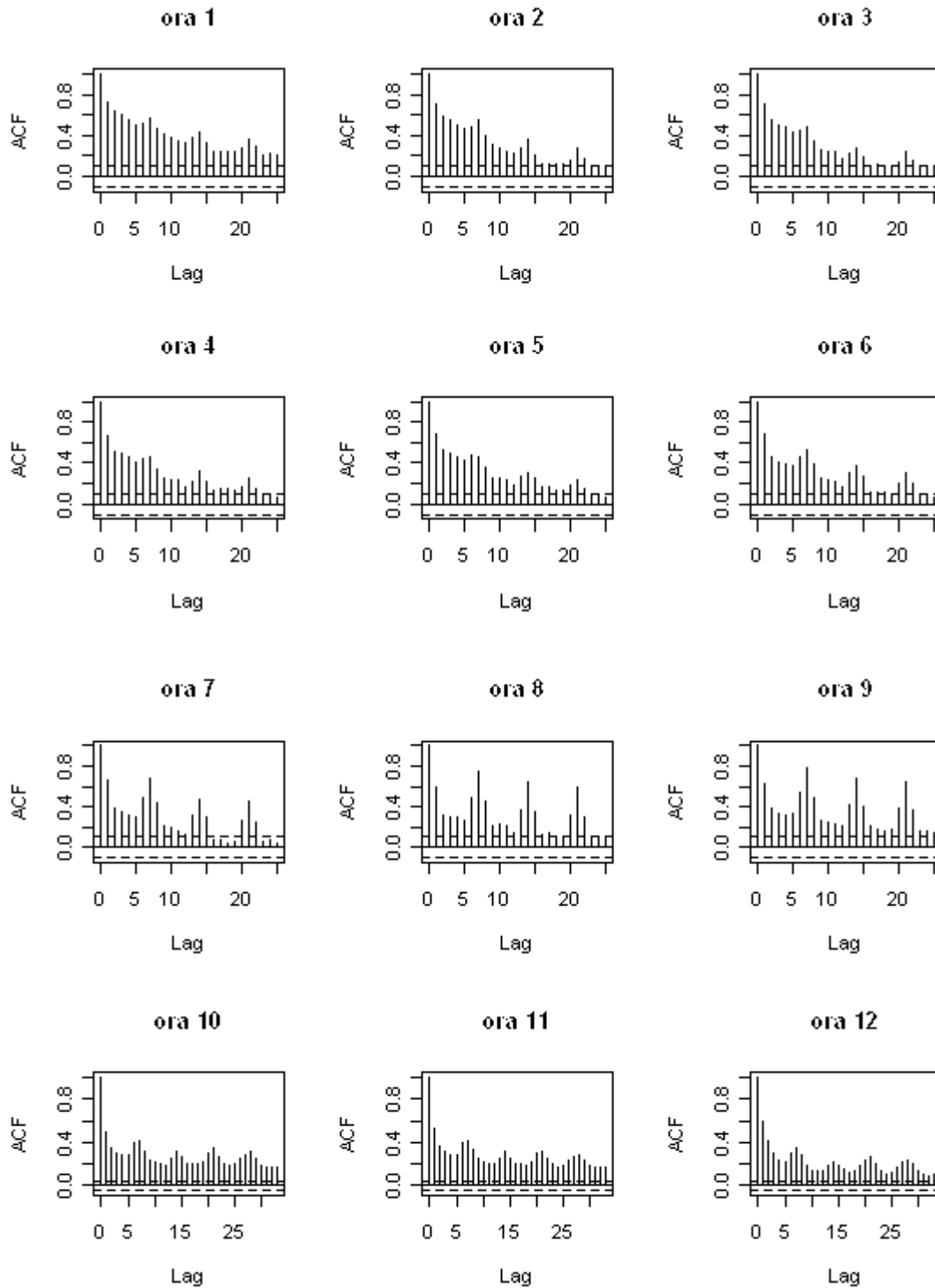


Grafici delle serie orarie per i dati usati a fini previsivi





Correlogrammi delle serie orarie per i dati usati a fini previsivi



Appendice B

Analisi delle serie ripulite dai picchi

Modello VAR per la serie originaria con componente deterministica stimata senza picchi

Ora 1	1:00 0.31	3:00 0.17	6:00 -0.18	7:00 0.11	08:00 0.09	9:00 0.21	11:00 0.29	12:00 -0.20	21:00 0.08		
Ora 2	2:00 0.33	3:00 0.12	6:00 -0.15	7:00 0.13	08:00 0.11	9:00 0.17	10:00 0.28	12:00 -0.16	20:00 0.06		
Ora 3	3:00 0.37	7:00 0.05	08:00 0.15	Ora10 0.33	Ora11 0.25	12:00 -0.31	22:00 0.20	23:00 -0.15			
Ora 4	3:00 0.22	4:00 0.08	5:00 0.10	7:00 0.08	8:00 0.16	11:00 0.10	17:00 0.15	22:00 0.18	23:00 -0.52	24:00 0.29	
Ora 5	2:00 -0.17	3:00 0.28	5:00 0.25	9:00 0.22	10:00 0.25	11:00 0.46	12:00 -0.41	16:00 -0.25	17:00 0.15	23:00 -0.07	
Ora 6	3:00 0.05	6:00 0.17	9:00 0.34	13:00 -0.08	21:00 0.25	22:00 0.12					
Ora 7	1:00 -0.20	6:00 0.09	7:00 0.11	9:00 0.16	13:00 -0.42	14:00 0.42	21:00 0.36	22:00 0.30			
Ora 8	1:00 -0.15	4:00 -0.11	6:00 0.09	8:00 0.06	9:00 0.18	13:00 -0.08	19:00 0.24	21:00 0.26	22:00 0.40		
Ora 9	2:00 -0.21	3:00 0.09	6:00 0.07	7:00 -0.09	9:00 0.20	12:00 0.14	13:00 -0.20	15:00 0.27	19:00 0.17	22:00 0.44	
Ora 10	2:00 -0.14	3:00 0.14	7:00 -0.08	9:00 0.14	10:00* -0.08	12:00 0.24	15:00 0.17	17:00 0.11	18:00 -0.11	19:00 0.16	22:00 0.35
Ora 11	2:00 -0.10	3:00 0.13	7:00 -0.11	9:00 0.11	10:00 -0.10	11:00 0.14	12:00 0.32	17:00 0.20	18:00 -0.10	19:00 0.10	22:00 0.32
Ora 12	3:00 0.05	7:00 -0.13	9:00 0.08	10:00 -0.08	12:00 0.54	17:00 0.22	18:00 -0.09	19:00 0.02	22:00 0.30		

Ora 13	3:00 0.08	5:00 -0.05	7:00 -0.08	12:00 0.20	13:00 0.31	17:00 0.20	18:00 -0.06	19:00 0.01	22:00 0.15	23:00 0.13	
Ora 14	3:00 0.07	7:00 -0.10	9:00 0.07	10:00 -0.15	12:00 0.19	13:00 0.17	14:00* 0.08	16:00 0.17	17:00 0.10	19:00 0.01	22:00 0.28
Ora 15	3:00 0.05	7:00 -0.11	9:00 0.07	10:00 -0.20	12:00 0.24	15:00 0.24	16:00 0.17	17:00 0.13	18:00 -0.08	19:00 0.07	22:00 0.31
Ora 16	3:00 0.02	7:00 -0.08	10:00 -0.14	12:00 0.26	16:00 0.30	17:00 0.15	19:00 0.04	22:00 0.34			
Ora 17	4:00 -0.04	7:00 -0.06	10:00 -0.09	12:00 0.21	17:00 0.31	19:00 0.26	20:00 -0.10	22:00 0.25	23:00 0.10		
Ora 18	3:00 0.07	5:00 -0.05	7:00 -0.07	18:00 0.40	19:00 0.31	22:00 0.21					
Ora 19	3:00 0.13	5:00 -0.05	7:00 -0.05	18:00 0.11	19:00 0.58	21:00 0.19					
Ora 20	1:00 0.04	3:00 0.13	5:00 -0.06	11:00 0.01	18:00 0.11	20:00 0.48	21:00 0.35	22:00 -0.18			
Ora 21	1:00 0.16	3:00 0.09	5:00 -0.08	6:00 -0.04	7:00 0.08	11:00 0.21	12:00 -0.17	20:00 0.18	21:00 0.40	22:00 0.15	24:00 -0.09
Ora 22	1:00 0.18	3:00 0.11	5:00 -0.07	6:00 -0.06	7:00 0.10	13:00 0.11	21:00 0.20	22:00 0.35	24:00 -0.05		
Ora 23	1:00 0.17	3:00 0.13	6:00 -0.09	11:00 0.24	12:00 -0.20	13:00 0.14	17:00 0.10	22:00 0.34	23:00* 0.05		
Ora 24	1:00 0.15	3:00 0.17	6:00 -0.10	11:00 0.38	12:00 -0.25	13:00 0.13	22:00 0.29	24:00* 0.04			

Modello VAR per la serie senza picchi

Ora 1	1:00 0.30	3:00 0.17	6:00 -0.18	7:00 0.11	08:00 0.09	9:00 0.21	11:00 0.34	12:00 -0.25	21:00 0.08		
Ora 2	2:00 0.31	3:00 0.13	6:00 -0.15	7:00 0.13	08:00 0.11	9:00 0.17	10:00 0.30	12:00 -0.16	20:00 0.08		
Ora 3	3:00 0.37	7:00 0.04	08:00 0.13	Ora10 0.35	Ora11 0.33	12:00 -0.39	22:00 0.21	23:00 -0.15			
Ora 4	3:00 0.22	4:00 0.09	5:00 0.10	7:00 0.07	8:00 0.15	11:00 0.11	17:00 0.17	22:00 0.18	23:00 -0.53	24:00 0.29	
Ora 5	2:00 -0.17	3:00 0.28	5:00 0.25	9:00 0.21	10:00 0.26	11:00 0.62	12:00 -0.53	16:00 -0.32	17:00 0.18	23:00 -0.06	
Ora 6	3:00 0.05	6:00 0.16	9:00 0.36	13:00 -0.10	21:00 0.25	22:00 0.11					
Ora 7	1:00 -0.19	6:00 0.09	7:00 0.10	9:00 0.16	13:00 -0.44	14:00 0.44	21:00 0.36	22:00 0.30			

Ora 8	1:00 -0.16	4:00 -0.11	6:00 0.09	8:00 0.06	9:00 0.18	13:00 -0.10	19:00 0.29	21:00 0.24	22:00 0.41		
Ora 9	2:00 -0.23	3:00 0.10	6:00 0.08	7:00 -0.09	9:00 0.18	12:00 0.19	13:00 -0.22	15:00 0.28	19:00 0.18	22:00 0.43	
Ora 10	2:00 -0.14	3:00 0.14	7:00 -0.08	9:00 0.10	10:00* -0.01	12:00 0.23	15:00 0.19	17:00 0.11	18:00 -0.16	19:00 0.18	22:00 0.34
Ora 11	2:00 -0.09	3:00 0.12	7:00 -0.11	9:00 0.07	10:00 -0.06	11:00 0.21	12:00 0.26	17:00 0.20	18:00 -0.14	19:00 0.11	22:00 0.32
Ora 12	3:00 0.04	7:00 -0.12	9:00 0.03	10:00 -0.01	12:00 0.50	17:00 0.22	18:00 -0.13	19:00 0.05	22:00 0.30		
Ora 13	3:00 0.07	5:00 -0.05	7:00 -0.07	12:00 0.16	13:00 0.37	17:00 0.19	18:00 -0.10	19:00 0.04	22:00 0.13	23:00 0.14	
Ora 14	3:00 0.06	7:00 -0.10	9:00 0.02	10:00 -0.08	12:00 0.10	13:00 0.17	14:00* 0.09	16:00 0.21	17:00 0.07	19:00 0.02	22:00 0.26
Ora 15	3:00 0.05	7:00 -0.11	9:00 0.02	10:00 -0.11	12:00 0.21	15:00 0.22	16:00 0.21	17:00 0.12	18:00 -0.13	19:00 0.11	22:00 0.29
Ora 16	3:00 0.02	7:00 -0.10	10:00 -0.11	12:00 0.26	16:00 0.33	17:00 0.11	19:00 0.05	22:00 0.32			
Ora 17	4:00 -0.04	7:00 -0.07	10:00 -0.08	12:00 0.23	17:00 0.29	19:00 0.31	20:00 -0.12	22:00 0.24	23:00 0.10		
Ora 18	3:00 0.06	5:00 -0.04	7:00 -0.07	18:00 0.38	19:00 0.34	22:00 0.20					
Ora 19	3:00 0.12	5:00 -0.05	7:00 -0.04	18:00 0.05	19:00 0.58	21:00 0.18					
Ora 20	1:00 0.04	3:00 0.12	5:00 -0.06	11:00 0.01	18:00 0.09	20:00 0.52	21:00 0.31	22:00 -0.17			
Ora 21	1:00 0.17	3:00 0.08	5:00 -0.08	6:00 -0.03	7:00 0.07	11:00 0.27	12:00 -0.22	20:00 0.18	21:00 0.41	22:00 0.14	24:00 -0.09
Ora 22	1:00 0.17	3:00 0.11	5:00 -0.07	6:00 -0.06	7:00 0.10	13:00 0.12	21:00 0.19	22:00 0.35	24:00 -0.05		
Ora 23	1:00 0.17	3:00 0.13	6:00 -0.10	11:00 0.30	12:00 -0.22	13:00 0.12	17:00 0.09	22:00 0.33	23:00 0.05		
Ora 24	1:00 0.15	3:00 0.17	6:00 -0.10	11:00 0.46	12:00 -0.30	13:00 0.13	22:00 0.28	24:00 0.04			

Previsioni per la serie originaria con componente deterministica stimata senza picchi

GENNAIO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	29.50	3.56	3.47	367.95	16.12	16.09	88.22	6.74	3.33
2	239.56	11.67	6.35	636.94	20.99	20.92	209.91	10.55	4.44
3	310.55	13.57	9.54	763.48	23.42	23.38	252.70	12.00	6.34
7	391.95	15.26	14.43	706.00	22.70	22.68	265.11	12.33	10.52
14	325.14	14.83	14.54	443.37	18.04	18.01	255.88	12.65	11.96
21	256.11	13.08	12.78	305.84	14.63	14.48	217.62	11.74	11.24
28	230.69	12.05	10.44	251.01	12.66	11.29	209.60	11.34	9.56

MAGGIO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	35.00	4.49	4.11	519.74	19.81	19.21	70.38	6.82	2.78
2	280.05	13.04	9.61	876.89	25.98	24.81	216.12	11.62	6.74
3	428.78	16.89	13.36	1036.49	28.12	26.94	327.96	14.60	9.78
7	986.76	25.96	23.88	1522.33	32.43	31.36	859.82	23.36	20.26
14	1769.78	35.66	35.37	2036.60	38.33	38.06	1628.63	33.58	33.14
21	1982.92	37.46	36.92	2094.77	38.55	38.01	1898.04	36.36	35.70
28	2422.84	41.72	41.09	2477.44	42.19	41.57	2369.45	41.12	40.43

AGOSTO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	41.61	4.86	4.30	549.87	20.68	19.55	98.35	7.43	0.51
2	292.20	13.41	9.44	1002.93	27.92	25.99	214.64	11.88	4.71
3	405.84	16.75	13.20	1212.85	30.78	28.90	281.63	14.16	8.10
7	888.88	26.52	24.86	1581.76	35.71	34.28	656.87	23.03	20.22
14	1621.97	36.83	35.57	1987.46	40.71	39.55	1415.11	34.32	32.59
21	2306.14	45.09	44.48	2503.72	46.91	46.33	2149.16	43.39	42.69
28	2693.35	49.00	48.48	2793.09	49.87	49.36	2593.21	48.01	47.47

NOVEMBRE									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	38.28	4.27	3.96	500.94	19.40	18.91	81.02	6.41	2.21
2	299.15	13.03	7.81	936.02	25.97	25.09	235.21	11.03	3.96
3	439.23	15.93	10.97	1128.49	28.73	27.88	333.03	13.37	5.88
7	655.53	20.33	19.07	1202.33	29.86	29.25	500.65	16.39	13.77
14	1041.32	27.20	27.01	1328.41	31.69	31.54	884.06	23.81	23.45
21	1175.03	28.69	28.61	1316.30	30.84	30.80	1074.92	26.67	26.48
28	942.00	24.93	23.83	1000.38	25.84	24.80	893.62	24.02	22.76

	ARMA-GARCH			SUR			VAR		
	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
Genn	584.00	19.52	18.22	867.60	25.16	25.11	367.27	15.37	12.37
Mag	899.79	26.10	24.90	997.05	27.86	26.76	836.64	25.28	23.52
Ago	798.03	24.08	21.24	1041.76	28.52	26.43	726.36	22.89	18.65
Nov	664.04	20.00	15.16	1101.50	28.55	27.67	553.94	17.94	6.09

Previsioni per la serie depurata dei picchi

GENNAIO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	29.16	3.53	3.44	353.93	15.87	15.84	86.37	6.69	3.29
2	234.82	11.61	6.32	618.13	20.77	20.70	203.84	10.50	4.41
3	304.63	13.56	9.46	744.97	23.26	23.21	246.68	11.96	6.30
7	390.79	15.27	14.38	704.69	22.69	22.66	265.19	12.34	10.53
14	322.71	14.79	14.49	442.38	18.02	18.00	255.37	12.64	11.96
21	254.52	13.06	12.76	305.43	14.63	14.48	217.39	11.75	11.25
28	229.35	12.02	10.43	250.81	12.66	11.30	209.46	11.35	9.58

MAGGIO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	40.58	4.76	4.19	534.32	20.41	19.28	99.07	4.76	0.52
2	281.62	13.06	9.58	862.99	25.79	24.62	215.86	11.61	6.83
3	430.48	16.94	13.31	1025.67	28.00	26.81	328.52	14.64	9.86
7	953.14	25.72	23.63	1473.36	32.22	31.14	818.39	23.15	20.04

14	1696.79	35.30	35.00	1955.77	38.00	37.73	1545.73	33.17	32.73
21	1906.01	37.12	36.58	2014.20	38.24	37.69	1812.42	35.96	35.29
28	2313.62	41.27	40.64	2365.85	41.76	41.13	2253.21	40.62	39.93

AGOSTO									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	40.57	4.76	4.19	534.31	20.41	19.28	99.07	7.43	0.52
2	291.40	13.36	9.29	987.68	27.73	25.79	214.13	11.87	4.65
3	403.84	16.68	13.03	1201.70	30.67	28.79	280.71	14.12	8.00
7	880.04	26.33	24.67	1579.19	35.72	34.29	648.76	22.92	20.08
14	1606.18	36.65	35.38	1978.97	40.68	39.52	1392.80	34.11	32.35
21	2276.86	44.86	44.24	2480.70	46.78	46.20	2111.08	43.09	42.39
28	2664.45	48.81	48.29	2769.21	49.75	49.24	2556.58	47.76	47.21

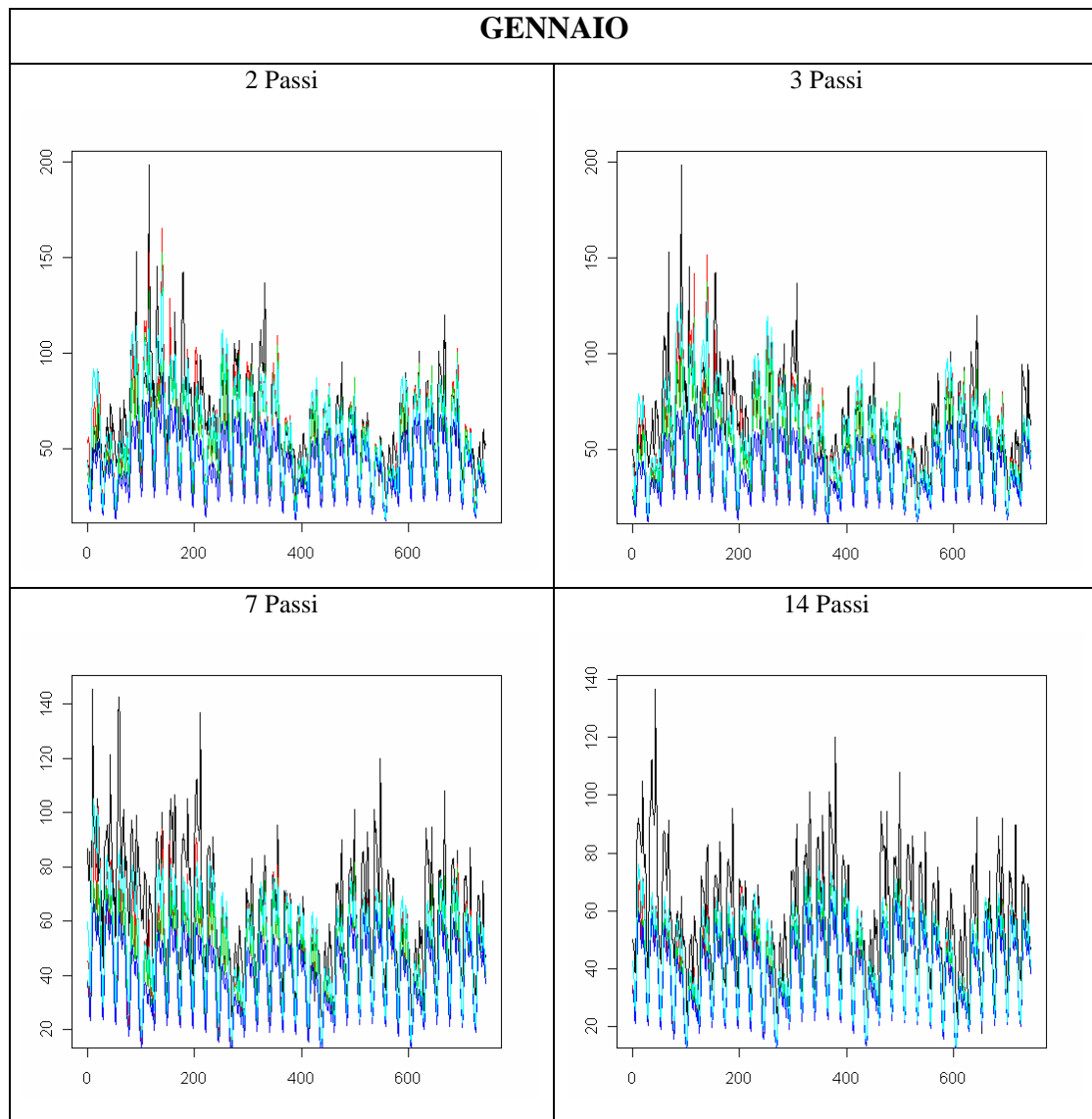
NOVEMBRE									
	ARMA-GARCH			SUR			VAR		
Passi	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
1	37.77	4.23	3.93	485.18	19.35	18.86	75.67	6.31	2.07
2	269.23	12.91	7.68	884.73	25.92	25.05	202.52	10.86	3.64
3	397.97	15.82	10.84	1077.57	28.79	27.95	291.08	13.17	5.55
7	607.08	20.22	18.98	1157.07	30.05	29.45	438.40	16.14	13.39
14	968.55	27.04	26.86	1265.23	31.73	31.59	800.27	23.38	23.01
21	1101.11	28.63	28.57	1250.26	30.91	30.88	990.54	26.40	26.22
28	902.17	24.97	23.88	965.21	25.94	24.92	847.83	23.94	22.68

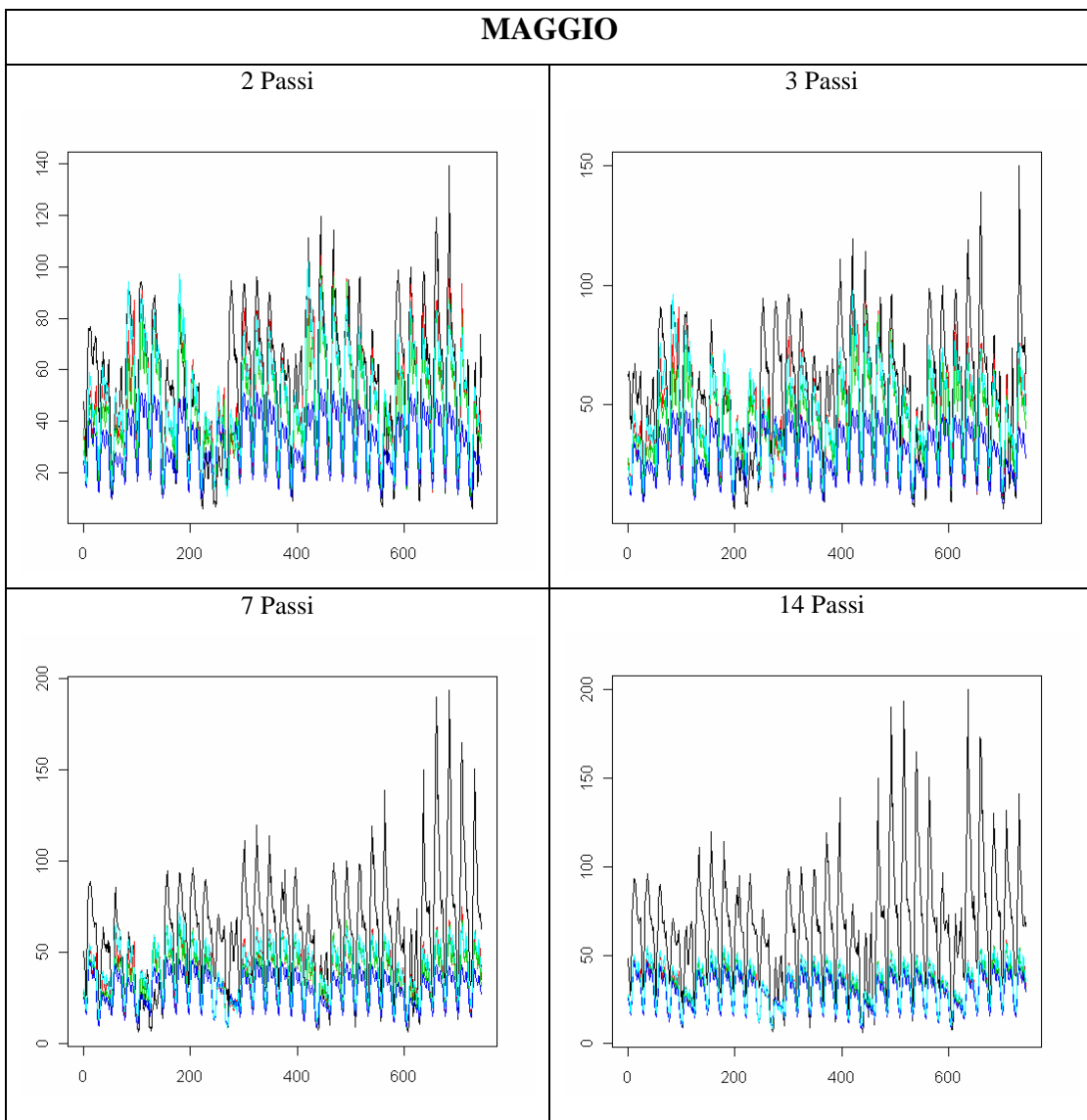
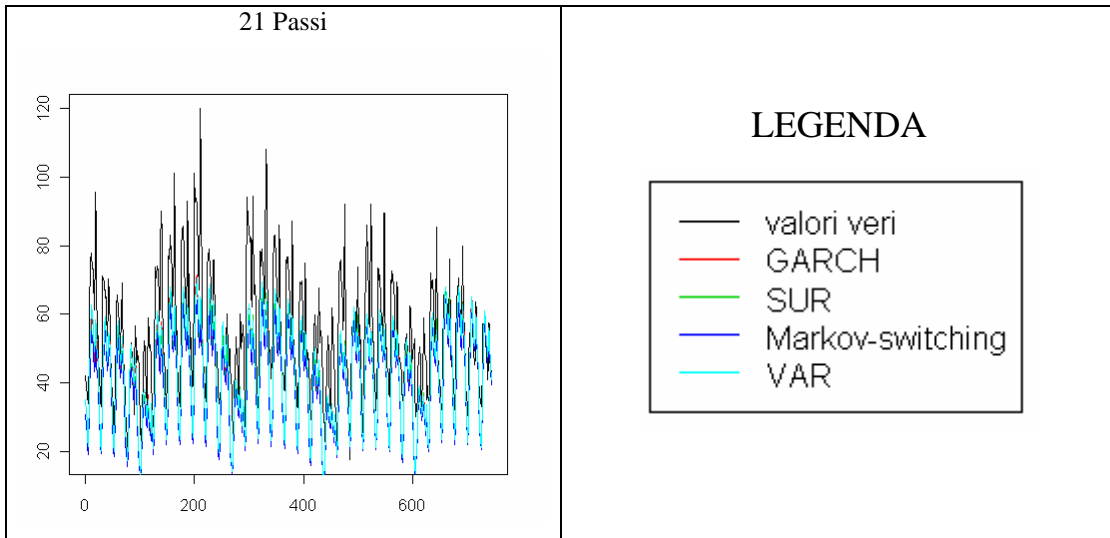
	ARMA-GARCH			SUR			VAR		
	MSE	MAE	Media	MSE	MAE	Media	MSE	MAE	Media
Genn	575.33	19.55	18.28	853.48	25.08	25.04	358.62	15.32	12.33
Mag	898.99	26.08	24.88	996.55	27.85	26.75	838.53	25.32	23.57
Ago	785.09	23.80	20.94	1043.91	28.55	26.48	721.10	22.84	18.55
Nov	633.96	20.06	15.03	1086.69	28.82	27.95	523.10	17.93	5.83

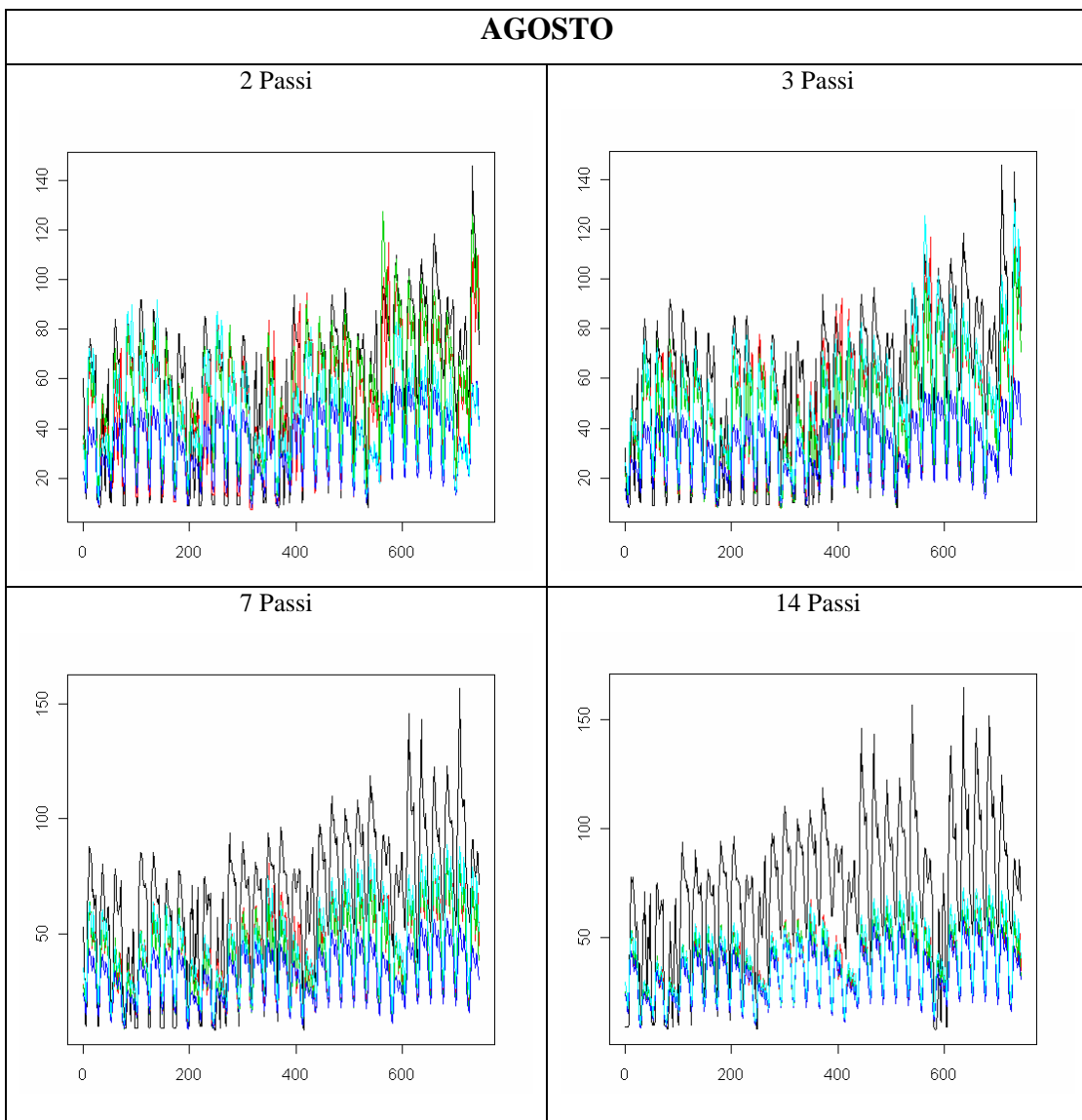
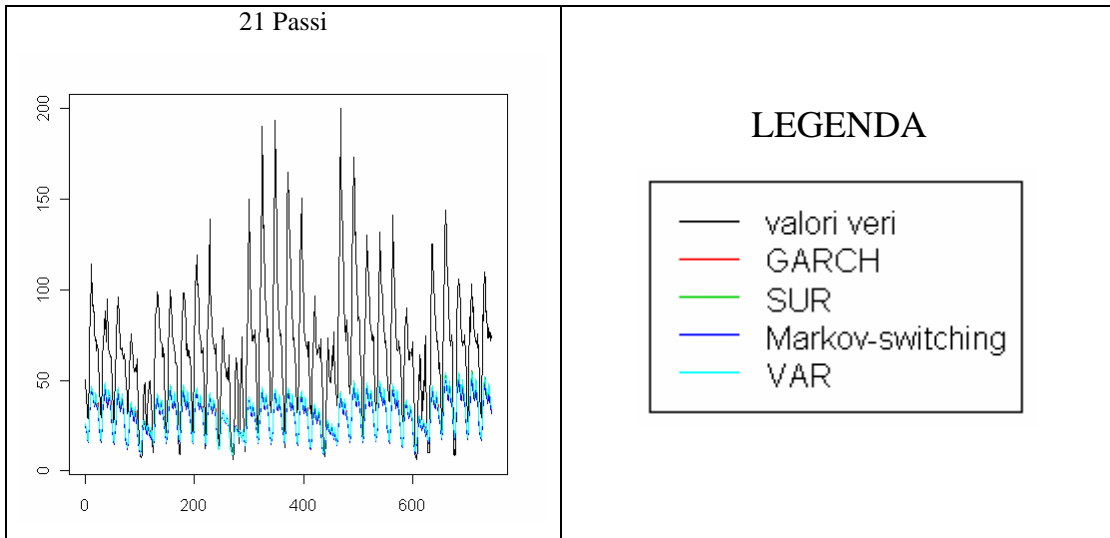
Appendice C

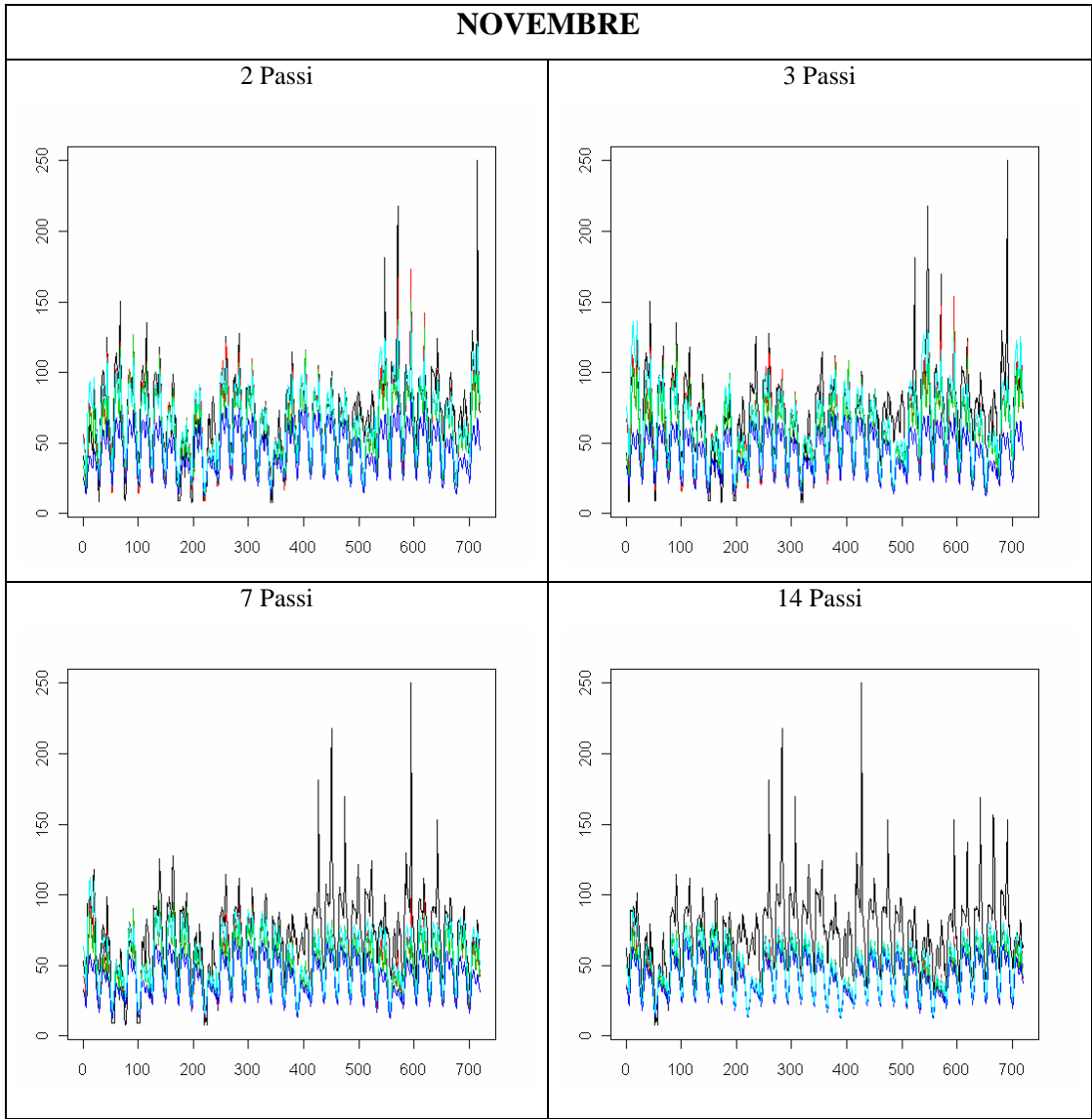
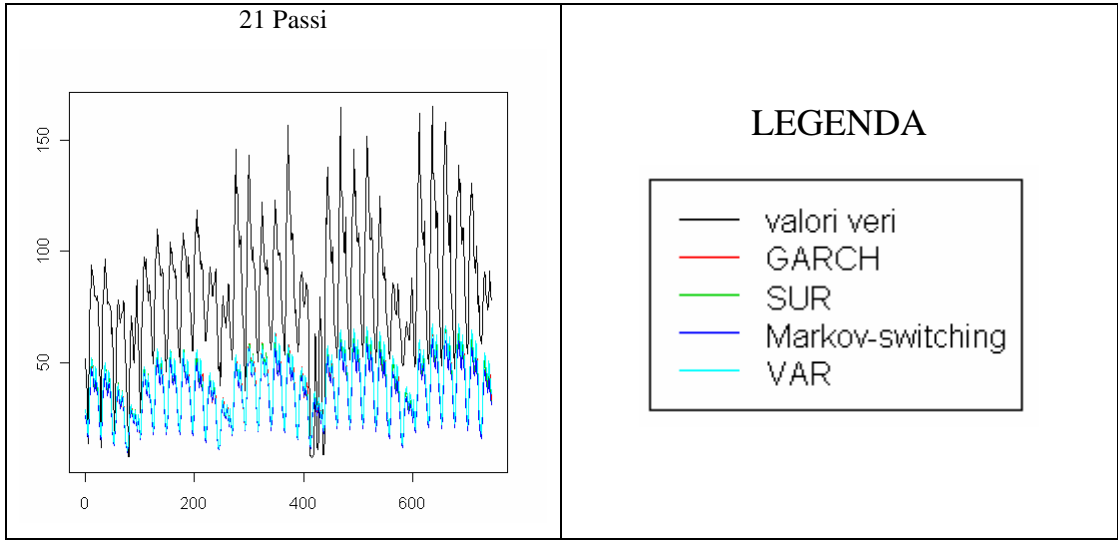
Ulteriori analisi sulla serie originaria

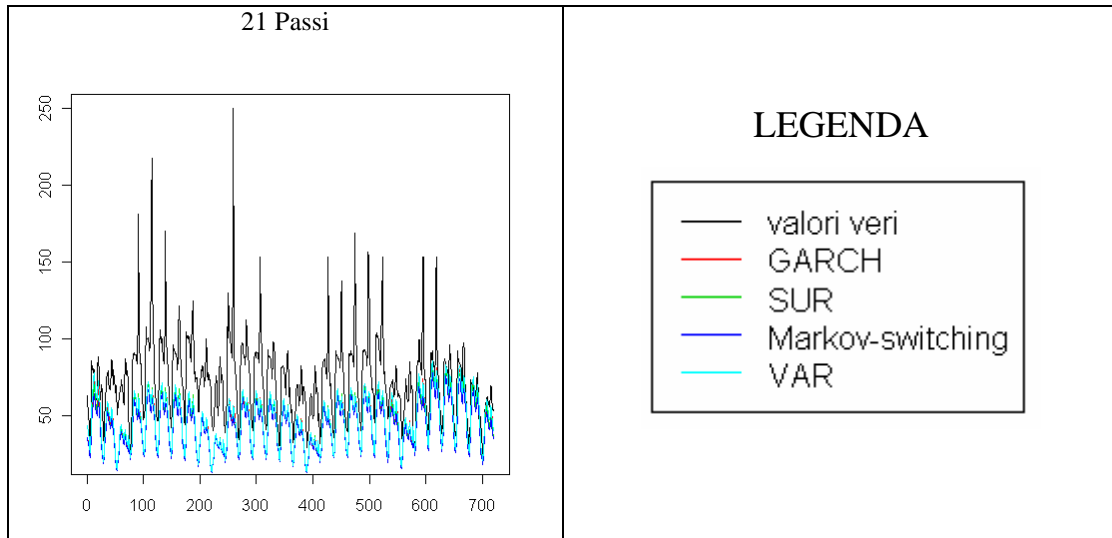
Grafici per le previsioni a 2, 3, 7, 14 e 21 passi











Varianze del modello *Markov-Switching*

	<u>Stato 1</u>	<u>Stato 2</u>
Ora 1	0.1692 (0.00)	0.4988 (0.00)
Ora 2	0.2882 (0.00)	0.8329 (0.00)
Ora 3	0.2988 (0.00)	0.9110 (0.00)
Ora 4	0.3546 (0.00)	1.9914 (0.00)
Ora 5	0.3452 (0.00)	1.3917 (0.00)
Ora 6	0.2487 (0.00)	0.9721 (0.00)
Ora 7	0.1470 (0.00)	0.6680 (0.00)
Ora 8	0.1404 (0.00)	0.6142 (0.00)
Ora 9	0.1288 (0.00)	0.5015 (0.00)
Ora 10	0.1394 (0.00)	0.5203 (0.00)
Ora 11	0.1315 (0.00)	0.4771 (0.00)

Ora 12	0.1543 (0.00)	0.5389 (0.00)
Ora 13	0.1385 (0.00)	0.4472 (0.00)
Ora 14	0.1332 (0.00)	0.4384 (0.00)
Ora 15	0.1355 (0.00)	0.4596 (0.00)
Ora 16	0.1280 (0.00)	0.4597 (0.00)
Ora 17	0.1279 (0.00)	0.4530 (0.00)
Ora 18	0.1374 (0.00)	0.4199 (0.00)
Ora 19	0.1433 (0.00)	0.4479 (0.00)
Ora 20	0.1611 (0.00)	0.4390 (0.00)
Ora 21	0.1739 (0.00)	0.5208 (0.00)
Ora 22	0.1080 (0.00)	0.3264 (0.00)
Ora 23	0.2391 (0.00)	1.4523 (0.00)
Ora 24	0.2386 (0.00)	1.4798 (0.00)

Bibliografia

- [1] Barlow, M. (2002), A diffusion model for electricity prices. *Mathematical Finance*
- [2] Benth et al. (2008) *A Survey of Electricity and Related Markets*.
- [3] Bollerslev, T.(1986) Generalize autoregressive conditional heteroscedasticity, *Journal of Econometrics*
- [4] Bunn D.W. e Karakatsani N. (2003) *Forecasting Electricity Prices*. London Business School
- [5] Clemen, R. T. (1989), Combining Forecasts. *International Journal of Forecasting*
- [6] De Jong C. (2006) *The Nature of Power Spikes: A Regime-Switch Approach*
- [7] Diebold, F.X. and Mariano, R.S. (1995) Comparing predictive accuracy. *Journal of Business and Economic Statistics*
- [8] Dordonnat, V., Koopman, S.J., Ooms, M., Dessertaine, A. e Collet, J. (2008) An hourly periodic state space model for modelling French national electricity load
- [9] Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*
- [10] Escribano, A., Peña, J.I. e Villaplana, P. (2002) Modeling electricity prices: international evidence.
- [11] Ethier, R. e Mount, T. (1998) Estimating volatility of spot prices in restructured electricity markets and the implications for option values.
- [12] Granger, C. W. J. e Joyeux, R. (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*
- [13] Granger, C.W. J. e Newbold, P. (1986), *Forecasting economic time series*, London Academic Press
- [14] Granger, C.W. J. e Ramanathan (1984), Improved Methods of Combining Forecasts, *Journal of Forecasting*
- [15] Hosking, J. (1981). Fractional differencing. *Biometrika*

- [16] Huisman, R. e Mahieu, R. (2003) Regime jumps in electricity prices.
- [17] Huisman, R., Huurman, R. e Mahieu, R. (2006) Hourly electricity prices in day-ahead markets.
- [18] Huisman, R.e De Jong, C. (2003) Option pricing for power prices with spikes.
- [19] Kosater, P. e Mosler, K. (2006) Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices.
- [20] Mount , T., Ning, Y., Cai, X. (2005) Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters
- [21] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [22] Schindlmayr, G. (2005) A Regime-Switching Model for Electricity Spot Prices.
- [23] Spagnolo, N. e Uddin, M.B. (2006) Modelling and Forecasting Spot Electricity Prices. Brunel Business School, London.
- [24] S-PLUS : Copyright (c) 1988, 2007 Insightful Corp. S : Copyright Insightful Corp. Enterprise Developer Version 8.0.4 for Microsoft Windows : 2007
- [25] Tong , H. (1978) On a threshold model. In: Pattern Recognition and Signal Processing. C.H. Chen (ed.) Sijthoff & Noordhoff, Amsterdam
- [26] Weron, R. (2008) Heavy-tails and regime-switching in electricity prices.
- [27] Weron, R.(2006) Modelling and Forecasting Electricity Loads and Prices. A statistical approach. John Wiley and Sons.