

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE STATISTICHE

*CORSO DI LAUREA TRIENNALE IN
STATISTICA PER L'ECONOMIA E L'IMPRESA*

STIMA DI MODELLI NON LINEARI SUI MERCATI DELL'ELETTRICITÀ ALLA LUCE DELLA RECENTE CRISI ENERGETICA

RELATORE:

PROF. LUIGI GROSSI

UNIVERSITÀ DEGLI STUDI DI PADOVA

LAUREANDA:

CHIARA ANGELETTI

MATRICOLA:

1224562

ANNO ACCADEMICO

2021-2022

Indice

1	Introduzione	5
2	Panoramica sul mercato elettrico italiano	7
2.1	Caratteristiche generali del mercato elettrico	7
2.2	Le sezioni del mercato elettrico	9
2.2.1	Mercato a pronti (MPE)	11
2.2.1.1	Mercato del giorno prima (MGP)	11
2.2.1.2	Mercato infragiornaliero(MI)	12
2.2.1.3	Mercato dei prodotti giornalieri (MPEG)	13
2.2.1.4	“Mercato del Servizio di Dispacciamento” (MSD)	14
2.3	Struttura zonale del mercato elettrico	15
2.4	Tecnologie produttive	19
3	Specificazione dei modelli non lineari per il PUN	21
3.1	Peculiarità delle serie storiche dei prezzi spot	21
3.2	Test di non linearità	24
3.2.1	BDS test	26
3.2.2	Test di McLeod e Li	27
3.2.3	RESET test	27
3.2.3.1	Keenan test	29
3.3	Modelli non lineari	29
3.3.1	Modelli autoregressivi a soglia (SETAR)	29
3.3.2	Markov switching model (MSM)	32

3.3.3	Smooth transition autoregressive model (STAR)	34
4	Stima di modelli e test di non linearità. Applicazione ai prezzi elettrici.	37
4.1	Applicazione test di non linearità	41
4.2	Stima dei modelli SETAR	41
4.3	Stima dei Markov switching model	48
4.4	Stima dei modelli STAR	54
5	Conclusione	59

Capitolo 1

Introduzione

La crisi energetica innescata dalla ripresa post COVID-19 ed aggravata dal conflitto Russia-Ucraina ha riportato all'ordine del giorno l'importanza di indagare e condurre ricerche nell'ambito della modellazione e previsione dei prezzi delle fonti energetiche e dell'elettricità.

In questo lavoro ci concentreremo nello specifico sulla modellazione del PUN (prezzo unico nazionale).

Nel capitolo 2 vengono descritti i cambiamenti che hanno interessato l'industria elettrica nell'ultimo ventennio, fino ad arrivare alla configurazione attuale del mercato elettrico italiano. Si entra nel dettaglio della sua articolazione, della configurazione zonale e del meccanismo che porta alla definizione del PUN. Viene inoltre delineato l'insieme delle tecnologie produttive in base alla percentuale di utilizzo.

Nel capitolo 3 è presente una caratterizzazione dei prezzi spot in cui si evidenziano le peculiarità dell'energia elettrica come non immagazzinabilità ed inelasticità della domanda, che si riflettono sulla serie storica del PUN, causando spikes ed elevata volatilità. Alla luce di queste evidenze si rende necessaria l'individuazione di modelli che siano in grado, almeno in parte, di tener conto delle caratteristiche dei prezzi spot. Si procede quindi alla descrizione di test di non linearità, utili per giustificare l'utilizzo di modelli non lineari. Le tipologie di modelli non lineari analizzate sono tre: modelli autoregressivi a soglia (SETAR), Markov Switching Model (MSM), Smooth Transition Autoregressive Model (STAR).

Nel capitolo 4 si passa all'applicazione dei modelli e dei test individuati, scegliendo dei regressori esogeni da includere nei MSM che siano in grado di migliorare l'adattamento nei tratti

in cui sono presenti picchi di prezzo. Si presentano, infine, i risultati ottenuti valutando le performance dei modelli con l'ausilio dell'analisi dei residui e dei criteri di confronto.

Capitolo 2

Panoramica sul mercato elettrico italiano

2.1 Caratteristiche generali del mercato elettrico

Negli ultimi due decenni, profondi e radicali cambiamenti hanno interessato l'impianto istituzionale, organizzativo e gestionale dell'industria elettrica di tutti i paesi europei. L'originario assetto monopolistico, introdotto in Italia dalla Legge n. 1643 del 6 Dicembre 1962 (legge di Nazionalizzazione), vedeva come suo principale attore l'Ente nazionale per l'energia elettrica (Enel), al quale era "riservato l'esercizio delle attività di: produzione, importazione/esportazione di energia elettrica, trasporto, trasformazione, distribuzione e vendita su tutto il territorio nazionale".

La volontà di perseguire un miglioramento dell'erogazione di servizi essenziali per lo sviluppo economico e il benessere sociale vede la sua realizzazione nel processo europeo di liberalizzazione del mercato dell'energia elettrica, avviato a livello comunitario con la direttiva 96/92/CEE, recepito e attuato a livello statale per effetto del Decreto Legislativo 16 Marzo 1999, n. 79 ("Decreto Bersani"). Si compie, in questo modo, una duplice transizione: da un lato si passa da un sistema nazionale a un sistema europeo, dall'altro da un sistema di monopolio pubblico ad un regime di concorrenza. Nell'ottica di introdurre nel mercato il principio di "workable competition" si sceglie di agire sotto il profilo dell'assetto proprietario con l'art. 8, co. 1, del d.lgs. n. 79 del 1999 il quale stabilisce che a decorrere dal 1° Gennaio 2003, a nessun soggetto è consentito produrre o importare, direttamente o indirettamente, più del 50% del totale dell'energia elettrica prodotta o importata in Italia. Sul fronte della domanda si individua un

doppio mercato: un mercato “vincolato” formato dai clienti “vincolati” e basato su un sistema di tariffa unica ed un parallelo mercato “libero” formato dai clienti “idonei”. Lo sviluppo concorrenziale è quindi legato al progressivo ampliamento della categoria dei clienti idonei che possono stipulare contratti di fornitura di energia elettrica con produttori da loro liberamente scelti. L’obbligo di passaggio al mercato libero è fissato per il 1 Gennaio 2024, data in cui cesserà l’esistenza del mercato tutelato.

Il decreto Bersani predispone strumenti finalizzati ad assicurare che gli interessi dei clienti vincolati non siano pregiudicati dall’apertura del mercato elettrico alla concorrenza. In tal senso viene prevista l’istituzione dell’Acquirente Unico(AU) che rende fruibile ai clienti vincolati una disponibilità di energia elettrica in condizioni di continuità sicurezza ed efficienza del servizio, nonché di parità di trattamento, anche dal punto di vista tariffario. Con questo scopo l’Acquirente Unico acquista energia elettrica dai produttori secondo piani previsionali della domanda e stipula con i distributori locali contratti di fornitura di elettricità, in base ai quali gli stessi distributori hanno l’obbligo di rifornire, nelle rispettive aree di competenza, i clienti vincolati a condizioni non discriminatorie e secondo una tariffa unica nazionale fissata dall’Autorità per l’energia elettrica e il gas.

Mentre la fase della produzione dell’energia elettrica è interessata da una liberalizzazione parziale, cioè limitata alla fornitura dei clienti idonei, le fasi di trasmissione e dispacciamento continuano ad essere sottoposte ad un regime di riserva. Tale scelta del legislatore nazionale si giustifica in base al carattere di monopolio naturale della rete di trasmissione elettrica. Infatti, non risultando conveniente la duplicazione della rete di trasmissione nazionale e comportando inefficienze sul piano tecnico la creazione di una pluralità di reti tra loro interconnesse, si è deliberato a favore del mantenimento di una sola rete di trasmissione tramite la creazione del Gestore della Rete di Trasmissione Nazionale (GRNT) il cui proprietario è il gruppo a partecipazione statale TERNA.

Si vengono così a delineare i tratti del mercato elettrico italiano odierno, la cui gestione è affidata al Gestore dei Mercati Energetici (GME), una società per azioni totalmente partecipata dal Gestore dei Servizi Energetici (GSE), a sua volta controllato dal Ministero di Economia e Finanze. Sono due le principali esigenze perseguite dal GME: promuovere, sulla base di criteri di neutralità, trasparenza e obiettività, la competizione nelle attività di produzione e di compravendita di energia elettrica tramite la creazione di una “piazza del mercato”; garantire

la gestione economica di un'adeguata disponibilità della riserva di potenza. Al GME è affidata la gestione del mercato dei Certificati Verdi e dei Titoli di Efficienza Energetica e del "mercato elettrico della borsa", in cui avvengono le transazioni che hanno per oggetto la compravendita all'ingrosso di energia elettrica. In questa sede si definiscono i prezzi di equilibrio ai quali viene valorizzata l'energia negoziata, come il risultato dell'incontro tra la sua domanda e la quantità offerta dagli operatori. La borsa elettrica si suddivide in Mercato elettrico a pronti (MPE) e Mercato elettrico a termine (MTE). L'articolazione ed il funzionamento di queste due sezioni verranno approfondite nei capitoli successivi.

2.2 Le sezioni del mercato elettrico

Il GME coordina e gestisce la domanda e l'offerta di energia in modo da disciplinare l'accesso dei produttori alla rete e l'immissione di energia elettrica secondo il "merito economico", cioè nel rispetto del canone di efficienza, valutato in termini di costi dei diversi impianti di generazione. Sulla base dell'incontro tra quantità e prezzi di domanda e di offerta, il GME determina il prezzo di equilibrio e l'ordine di attivazione delle centrali destinate ad immettere elettricità nella rete. La priorità di immissione nella rete è attribuita ai generatori che presentano offerte più vantaggiose. A differenza di altri mercati dell'energia, quello del GME, non è un mercato puramente finanziario finalizzato solo alla determinazione di prezzi e quantità, ma è un vero e proprio mercato fisico dove si definiscono i programmi di immissione e prelievo. In tutte le transazioni concluse il GME è controparte centrale dei produttori, grossisti e clienti finali che operano nel mercato. Come già detto il GME gestisce il mercato all'ingrosso organizzato dell'elettricità in Italia, comunemente indicato come "borsa elettrica", che consente a produttori, consumatori e grossisti di stipulare contratti orari di acquisto e vendita di energia elettrica. Il mercato dell'elettricità si articola, su base temporale, in quelle che è possibile definire sezioni e sottosezioni (figura 2.1). La prima suddivisione, nello specifico, è tra:

- **"Mercato a pronti" (MPE)**
- **"Mercato a termine" (MTE)**

Il "Mercato a pronti", poi, è suddiviso a sua volta in:

- **“Mercato del Giorno Prima”, (MGP)** dove i produttori, i grossisti e i clienti finali idonei possono vendere/acquistare energia elettrica per il giorno successivo, in un’unica seduta. Il GME agisce come controparte centrale.
- **“Mercato infragiornaliero”, (MI)** permette a produttori, grossisti e clienti finali idonei di modificare i programmi di immissione/prelievo determinati su MGP. Il mercato è strutturato in quattro sessioni: tre (MI-A₁, MI-A₂, MI-A₃) sono sessioni d’asta, mentre la rimanente è una sessione di negoziazione continua MI-XBID, articolata in tre fasi, una per ogni sessione d’asta.
- **“Mercato dei prodotti giornalieri”, (MPEG)** sede della la negoziazione dei prodotti giornalieri con obbligo di consegna dell’energia.
- **“Mercato del Servizio di Dispacciamento”, (MSD)** nel quale il TSO (Terna S.P.A) si approvvigiona dei servizi di dispacciamento necessari alla gestione e al controllo del sistema elettrico. Si articola in un sessione ex-ante (in sei sotto-fasi: MSD₁, MSD₂, MSD₃, MSD₄, MSD₅, MSD₆), finalizzata all’acquisto dei servizi di risoluzione delle congestioni e della riserva energetica, e in una seconda fase infragiornaliera di accettazione delle stesse offerte ai fini di bilanciamento (Mercato di Bilanciamento). Diversamente dagli altri due mercati, il MSD è aperto solo ad una categoria ristretta di operatori abilitati all’offerta di determinati servizi, chiamati “ancillari”, che vengono remunerati non ad un prezzo di equilibrio di mercato ma al prezzo offerto (pay-as-bid), qualora accettato dall’unica controparte, Terna.

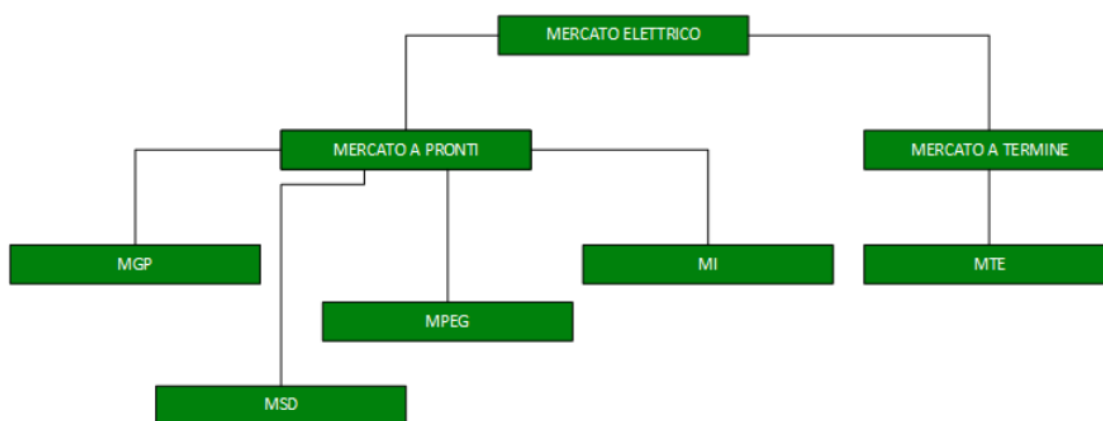


Figura 2.1: Sezioni Mercato elettrico

2.2.1 Mercato a pronti (MPE)

Il mercato a pronti, anche detto mercato spot, ospita la compravendita di energia nel breve periodo, da un giorno a poche ore prima del ritiro fisico, ed è composto da quattro sotto-mercati (MGP, MI, MPEG, MSD).

2.2.1.1 Mercato del giorno prima (MGP)

Il mercato del giorno prima è sede della maggior parte delle transazioni di compravendita di energia elettrica. Non è un mercato a contrattazione, bensì un mercato d'asta in cui si scambiano blocchi orari di energia per il giorno successivo. Ogni seduta viene aperta alle ore 8.00 del nono giorno antecedente il giorno di consegna e si chiude alle ore 12.00 del giorno precedente il giorno di consegna. Durante questa finestra temporale gli operatori possono partecipare presentando offerte in cui viene indicato il prezzo (minimo o massimo) che sono disposti ad accettare e la quantità che sono intenzionati a vendere o acquistare. I criteri sulla base dei quali si accettano le offerte prendono in considerazione il merito economico, cioè in ordine di prezzo crescente, sempre rispettando i limiti di transito tra le zone. Gli esiti della contrattazione vengono resi noti dopo la chiusura delle sedute ed entro le 12.58 del giorno precedente il giorno di consegna. Le offerte di acquisto accettate e riferite alle unità di consumo che appartengono alle zone geografiche d'Italia vengono valorizzate al cosiddetto PUN (Prezzo Unico Nazionale), che è uguale alla media dei prezzi delle zone geografiche (P_z) ponderata per le quantità acquistate in quelle zone (Q_z).

$$PUN = \frac{\sum_K E[P_z^{(k)} * Q_z^{(k)}]}{\sum_K E[Q_z^{(k)}]} \quad \text{con k zona di mercato} \quad (2.1)$$

Le offerte riferite a zone estere, ma accettate nel MGP vengono valorizzate al prezzo marginale di equilibrio (prezzo d'intersezione tra domanda e offerta, calcolato ogni ora) della zona di appartenenza. Il Gestore dei Mercati Energetici, nel MGP, opera in qualità di controparte centrale.

2.2.1.2 Mercato infragiornaliero(MI)

Il Mercato Infragiornaliero (MI) consente agli operatori di apportare modifiche ai programmi definiti nel MGP attraverso ulteriori offerte di acquisto o vendita. Questo mercato, definito mercato di aggiustamento prima della riforma del mercato elettrico del 2009, ha subito notevoli modifiche nel corso degli anni, l'ultima risalente al 21 Settembre 2021, che lo ha reso più flessibile ed adeguato alle esigenze del settore elettrico. Le tre sessioni MI-A sono organizzate nella forma di aste implicite di energia con orari di chiusura diversi ed in successione, attraverso le quali gli operatori possono sia effettuare un miglior controllo dello stato degli impianti di produzione, sia aggiornare i programmi di prelievo delle unità di consumo, tenendo conto di informazioni più aggiornate circa lo stato dei propri impianti di produzione, il fabbisogno di energia per il giorno successivo e le condizioni di mercato. Durante le sessioni d'asta MI-A, contestualmente alla negoziazione delle offerte di acquisto e vendita, viene allocata la capacità infragiornaliera di interconnessione tra tutte le zone del mercato italiano e le altre aree geografiche alle stesse interconnesse coinvolte nel Market Coupling.

Dopo la chiusura del MGP, le tre sedute d'asta si svolgono come segue:

- MI-A1 dalle ore 12.55 del giorno precedente il giorno di consegna alle ore 15.00 dello stesso giorno. Gli esiti della stessa vengono comunicati entro le ore 15.30 del giorno precedente il giorno di consegna.
- MI-A2 dalle ore 12.55 del giorno precedente il giorno di consegna alle ore 22.00 dello stesso giorno. Gli esiti del MI-A2 vengono comunicati entro le ore 22.30 del giorno precedente il giorno di consegna.
- MI-A3 dalle ore 12.55 del giorno precedente il giorno di consegna alle ore 10.00 del giorno di consegna. Gli esiti del MI-A3 vengono comunicati entro le ore 10.30 del giorno di consegna.

Le offerte di acquisto e vendita vengono selezionate sulla base dello stesso criterio descritto per MGP, tuttavia, in questo caso, non viene calcolato il PUN e tutti gli acquisti e le vendite sono valorizzate al prezzo zonale. Alla luce di ciò, le unità di consumo potrebbero compiere delle scelte arbitrarie tra il PUN stabilito dal MGP e il prezzo zonale del MI. Per questa ragione, è stato ideato un corrispettivo di non arbitraggio, pari alla differenza tra i due prezzi, che rende non profittevole tale comportamento.

La sessione continua MI-XBID prende il nome dall'iniziativa XBID Market Project, entrata in operatività nel centro Europa nel 2018 con lo scopo di creare un mercato infragiornaliero europeo basato sulla negoziazione continua che consente l'allocazione implicita delle capacità di trasmissione interzonale, fornite a XBID dai TSO. Grazie a questo progetto, le offerte di acquisto e vendita di energia presentati dai partecipanti al mercato di un paese/zona di mercato possono essere abbinati con quelli presentati da altri partecipanti di altri paesi/zona di mercato che partecipano a XBID, a condizione che vi sia capacità di interconnessione disponibile.

La sessione di negoziazione MI-XBID è articolata in tre fasi :

- la fase I MI-XBID si apre alle ore 15.30 del giorno precedente al giorno di consegna e si chiude alle ore 21.40 dello stesso
- la fase II MI-XBID si apre alle ore 22.30 del giorno precedente al giorno di consegna e si chiude: per i periodi rilevanti corrispondenti alle prime dodici ore del giorno di consegna, un'ora prima dell'inizio di ciascun periodo rilevante (h-1); per i periodi rilevanti corrispondenti alle seconde dodici ore del giorno di consegna, alle ore 09.40 del giorno stesso
- la fase III MI-XBID si apre dalle ore 10.30 del giorno di consegna e si chiude un'ora prima dell'inizio di ciascun periodo rilevante (h-1).

Le sessioni d'asta MI-A e le fasi di negoziazione della sessione MI-XBID si alternano in base al seguente ordine: 1) MI-A1; 2) prima fase MI-XBID; 3) MI-A2; 4) seconda fase MI-XBID; 5) MI-A3; 6) terza fase MI-XBID. Il GME agisce come controparte centrale.

2.2.1.3 Mercato dei prodotti giornalieri (MPEG)

Il Mercato dei Prodotti Giornalieri (MPEG), è la sede deputata alla negoziazione dei prodotti giornalieri con obbligo di consegna dell'energia. L'ammissione è concessa a tutti gli operatori del mercato elettrico e le negoziazioni avvengono in modalità continua.

Sul MPEP sono negoziabili due classificazioni di prodotti:

- “**differenziale unitario di prezzo**”, il cui prezzo di negoziazione è l’espressione del differenziale, rispetto al PUN, a cui gli operatori sono disposti a negoziare i prodotti;
- il “**prezzo unitario pieno**”, il cui prezzo è l’espressione del valore unitario di scambio dell’energia elettrica oggetto dei contratti che sono negoziati.

Ambedue le tipologie di prodotto hanno la possibilità di essere consegnate secondo i seguenti profili:

- il **Baseload**, quotato tutti i giorni, il cui sottostante è l’energia elettrica da consegnare in tutti i periodi rilevanti che appartengono al giorno di negoziazione;
- Il **Peak Load**, quotato da lunedì al venerdì, il cui sottostante è l’energia elettrica da consegnare nei periodi rilevanti dal nono al ventesimo che appartengono al giorno di negoziazione.

Le sessioni del MPEP si svolgono nei giorni feriali, secondo quanto di seguito specificato: dalle ore 8.00 alle ore 17.00 del giorno che precede di due giorni quello di consegna, dalle ore 8.00 alle ore 9.00 del giorno precedente al giorno di consegna.

2.2.1.4 “Mercato del Servizio di Dispacciamento” (MSD)

Il Mercato per il Servizio di Dispacciamento (MSD) è lo strumento attraverso il quale Terna S.p.A. si approvvigiona delle risorse necessarie alla gestione e al controllo del sistema (risoluzione delle congestioni intrazonali, creazione della riserva di energia, bilanciamento in tempo reale). Sul MSD Terna agisce come controparte centrale e le offerte accettate vengono remunerate al prezzo presentato (pay-as-bid).

Il MSD si articola in fase di programmazione (MSD ex-ante) e Mercato del Bilanciamento (MB).

Il MSD ex - ante si suddivide in sei sottofasce di programmazione: MSD₁, MSD₂, MSD₃, MSD₄, MSD₅ e MSD₆. (figura 2.2)

Sul MSD ex-ante Terna accetta offerte di acquisto e vendita di energia ai fini della risoluzione delle congestioni residue e della costituzione dei margini di riserva e le riferisce al GME il quale, a sua volta, comunica agli operatori gli esiti delle singole sessioni e i programmi post MSD ex-ante di immissione e prelievo agli utenti del dispacciamento.

Nel MB si presentano, in modo continuo, offerte di acquisto e vendita di energia al fine di svolgere il servizio di regolazione secondaria e mantenere il bilanciamento, nel tempo reale,

Giorno di riferimento	D-1	D				
	MSD1	MSD2	MSD3	MSD4	MSD5	MSD6
Apertura seduta	12.55	°	°	°	°	°
Chiusura seduta	17.30	°	°	°	°	°
Esiti definitivi	21.45	2.15	6.15	10.15	14.15	18.15

Figura 2.2: Orario sottofasi di programmazione MSD ex-ante

tra immissione e prelievi di energia sulla rete. Le offerte hanno dettaglio orario per le 24 ore del giorno di flusso di consegna. L'apertura della seduta per la presentazione delle offerte per il MB è alle 22.30 del giorno antecedente il giorno di flusso di consegna. Gli operatori potranno presentare offerte sino a 60' prima dell'inizio dell'ora H cui tali offerte si riferiscono.

2.3 Struttura zonale del mercato elettrico

Per semplificare la trasmissione, soprattutto allo scopo di aumentare la sicurezza e per evitare congestioni sulla rete, l'attuale disegno dei mercati elettrici in Italia prevede un approccio di tipo zonale.

Ognuna delle zone in cui è stata effettuata la suddivisione è definita "zona della rete rilevante", ossia una porzione di RTN costituita da un insieme di punti di offerta in immissione o prelievo per la quale esistono, ai fini della sicurezza del sistema elettrico, limiti di transito fisici dell'energia in entrata e in uscita con altre zone confinanti. Tali limiti di import e export, determinati ricorrendo ad un modello di calcolo basato sul bilancio tra generazione e consumi, sono essenziali per determinare il prezzo dell'energia all'ingrosso e per bilanciare eventuali squilibri che si vengono a creare quando una zona presenta carenza di produzione oppure ricorre a mix produttivi più costosi di altri.

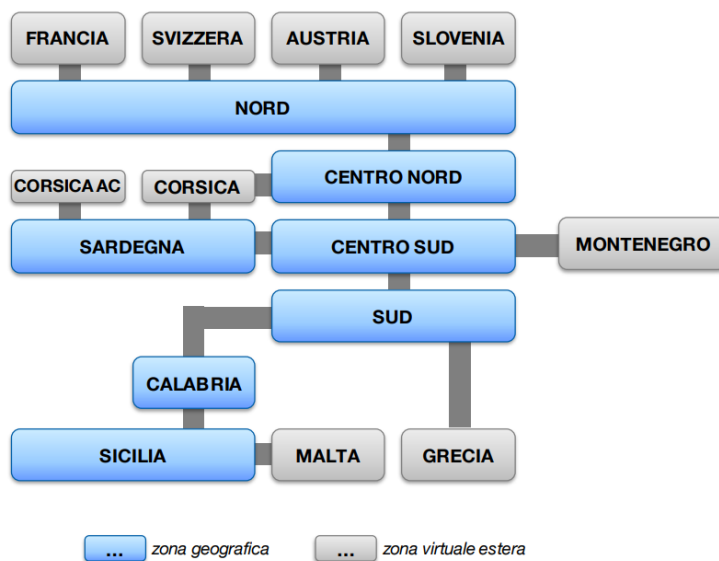


Figura 2.3: Topologia di interconnessione delle zone

Le zone della rete rilevante possono essere:

- aree geografiche fisiche;
- zone virtuali (ovvero senza un diretto corrispondente fisico);
- poli di produzione limitata (zone virtuali costituite da un insieme di unità di produzione connesse alla RTN per le quali esistono vincoli di rete sulla massima potenza complessiva erogabile). Nell'assetto attuale è stato eliminato l'unico polo di produzione limitata che era quello di Rossano .

La topologia di interconnessione tra le zone geografiche e/o virtuali della struttura zonale in vigore a partire dal 1° Gennaio 2021 mostrata in figura 2.3 è così costituita:

Zone virtuali:

- Zona Francia ;
- Zona Svizzera ;
- Zona Corsica ;
- Zona Corsica AC ;

- Zona Austria ;
- Zona Slovenia ;
- Zona Grecia ;
- Zona Malta ;
- Zona Montenegro.

Zone geografiche:

- Zona Nord, costituita dalle regioni Valle D'Aosta, Piemonte, Liguria, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna;
- Zona Centro Nord, costituita dalle regioni Toscana e Marche;
- Zona Centro Sud, costituita dalle regioni Lazio, Abruzzo, Umbria e Campania;
- Zona Sud, costituita dalle regioni Molise, Puglia e Basilicata;
- Zona Calabria;
- Zona Sicilia;
- Zona Sardegna.

La Zona Nord risulta interconnessa con le seguenti zone virtuali estere: Francia, Svizzera, Austria e Slovenia. Dal punto di vista delle interconnessioni interne, essa risulta collegata alla Zona Centro Nord. La Zona Centro Nord risulta interconnessa con la zona Nord e la zona Centro Sud. La Zona Centro Sud risulta interconnessa con la Zona Centro Nord e la zona Sud. La Zona Sud risulta interconnessa con la zona Centro Sud e con la zona Calabria. La zona Sicilia risulta interconnessa con la zona Calabria. La Sardegna è interconnessa al continente con la zona Nord, e con la zona Centro Sud .Inoltre è collegata alla zona virtuale della corsica AC.

La cooperazione tra zone italiane e zone virtuali estere, in particolare tra Italia e Slovenia, tra Italia e Francia ,tra Italia e Austria e tra Italia e Grecia è normata dal meccanismo del "market coupling " che attraverso un sistema di aste implicite consente di limitare i flussi antieconomici (flussi da zone più costose a meno, in opposizione al differenziale di prezzo).

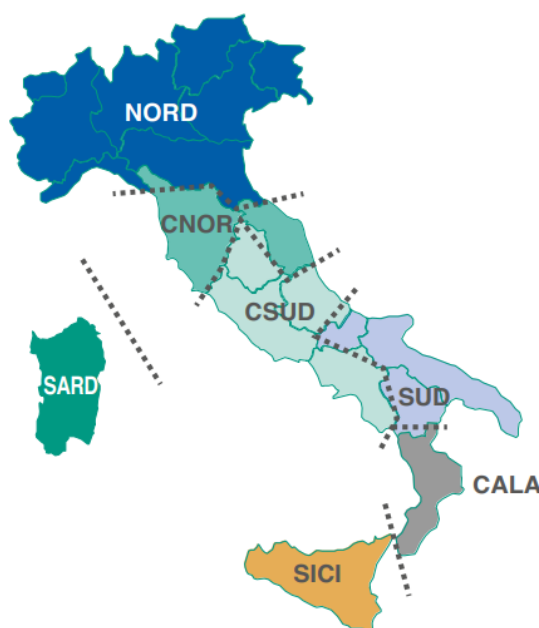


Figura 2.4: Sezioni di rete del territorio italiano

La struttura della RTN presenta naturalmente delle “sezioni” (intese come linee ideali che separano la rete in più parti, visibili in figura 2.4, lungo le quali in determinate condizioni di esercizio, è più probabile che possano verificarsi congestioni di rete.

In generale, una sezione di rete si dice:

- “strutturalmente critica” se esistono degli assetti di produzione interni alle zone che non permettono in condizioni di sicurezza il libero transito di energia tra zone adiacenti ;
- “operativamente critica” quando la potenza transitante su di essa è superiore al valore di soglia strutturale. Ogni qual volta si osserva una differenza di prezzo tra zone limitrofe siamo in presenza di una **congestione** .

Questa è conseguenza di una violazione dei limiti di transito che porta l’algoritmo che determina il prezzo di equilibrio dell’elettricità a separare il mercato in due zone : una in esportazione che comprende tutte le zone a monte del vincolo , e una in importazione che include tutte le zone a valle del vincolo. Si ridefinisce un prezzo di equilibrio differente nelle due zone di mercato P_z detto prezzo di equilibrio zonale che è maggiore nella zona di mercato importatrice e minore in quella esportatrice

2.4 Tecnologie produttive

Il primo anello della filiera elettrica è quello della produzione, che consiste nell'approvvigionamento delle fonti primarie disponibili in natura, nella trasformazione di queste attraverso impianti elettrici e, infine, nell'immissione del prodotto finito. I dati di TERNNA del primo trimestre del 2022 mostrano che circa il 29% del fabbisogno totale di energia elettrica in Italia, è stato coperto da fonti rinnovabili, con un aumento della produzione sia solare che eolica dell'11%. Tuttavia le fonti più utilizzate restano quelle tradizionali, come testimoniano i report forniti dal GME risalenti al 2021 (figura 2.5), che suddividono le vendite di energia per fonte.

Il gas risulta essere la fonte più utilizzata con un 49,2%, subito dopo troviamo solare e eolica con una quota rispettivamente del 9,2% e dell'8,4%. Le vendite da fonti rinnovabili diminuiscono del 10,2% rispetto all'anno precedente, con una riduzione della quota di mercato al 37,4%.

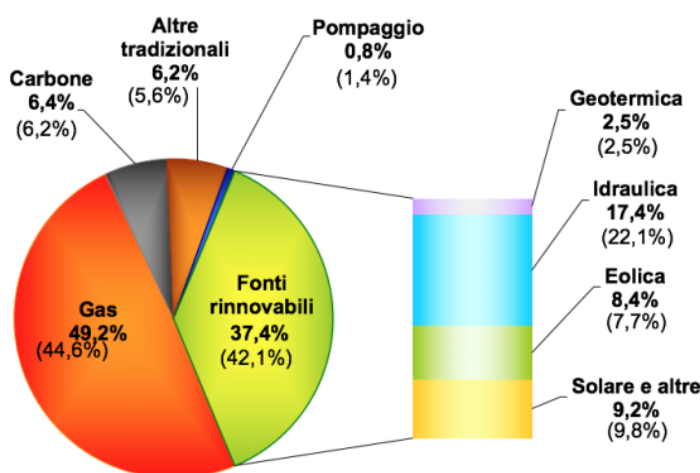


Figura 2.5: MGP, struttura delle vendite per fonte del Sistema Italia. (Tra parentesi sono indicati i valori dello stesso mese dell'anno precedente)

Poiché la domanda di energia elettrica è molto volatile nel breve periodo, ed essendo l'energia un bene difficilmente immagazzinabile, si rende necessaria un'offerta altrettanto flessibile. Per far fronte a questo bisogno si utilizzano, come abbiamo appena visto, fonti diversificate di produzione in modo da soddisfare la domanda al minor costo di generazione possibile. La

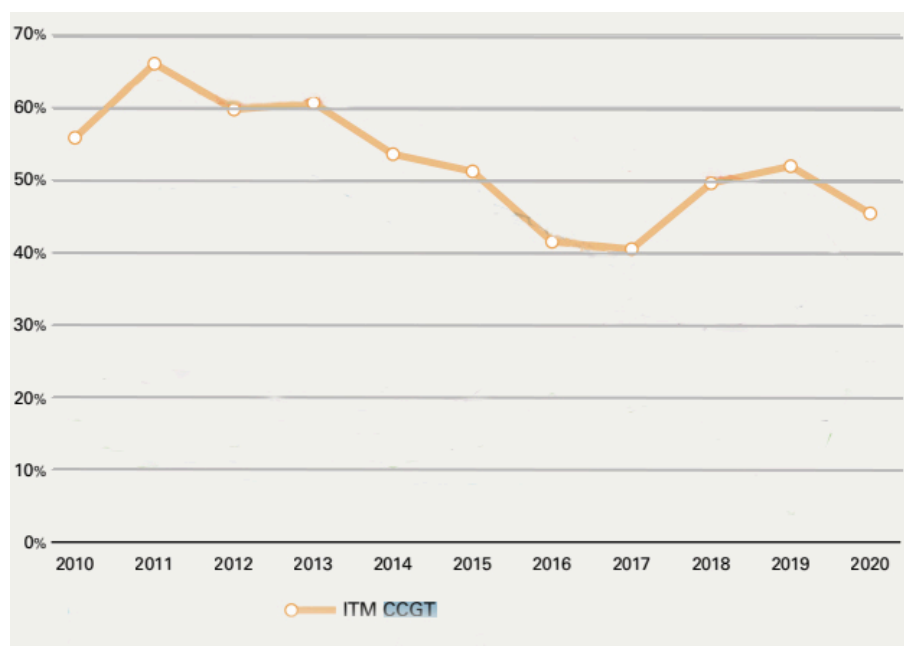


Figura 2.6: Indice ITM CCGT

domanda di base viene gestita ricorrendo ad impianti termoelettrici a gas o combustibile che sono efficienti se fatti lavorare per un numero prolungato di ore, quindi non in grado di rispondere a picchi improvvisi della domanda che possono essere fronteggiati da impianti più reattivi come quelli idroelettrici o a turbogas.

Gli impianti a ciclo combinato turbogas (ccgt) sono risultati i più influenti nella fissazione dei prezzi orari secondo l'indice **ITM**, ovvero l'indice che esprime per quante ore, percentualmente, gli impianti a ciclo combinato turbogas sono stati gli impianti marginali. Per impianti marginali si intende quelli che hanno fissato il prezzo dell'elettricità sul MGP. Come si può notare dall'ultimo report annuale della GME risalente al 2020 (figura 2.6), l'indice ITM ccgt è quasi sempre superiore al 50%.

L'analisi dei dati messi a disposizione conduce a considerare le tecnologie produttive che utilizzano il gas come le maggiormente marginali e quindi quelle che incidono in misura più significativa sulla formazione del PUN.

Capitolo 3

Specificazione dei modelli non lineari per il PUN

3.1 Peculiarità delle serie storiche dei prezzi spot

I prezzi spot di un determinato giorno, o prezzi a pronti, corrispondono al prezzo dei prodotti in consegna immediata. Nel caso del mercato elettrico il "prodotto" è l'energia elettrica e per consegna immediata si intende il giorno successivo (Day-ahead) rispetto alla chiusura delle sedute di compravendita. Il prezzo spot è quello che viene definito nel MGP, che, a livello nazionale, coincide quindi con il PUN.

Al fine di individuare dei modelli che siano in grado di prevedere l'andamento dei prezzi spot è utile richiamare alcune caratteristiche tipiche del settore elettrico che si riflettono sui prezzi dell'energia elettrica all'ingrosso. Queste sono:

- **Difficoltà di immagazzinare la commodity**, sia per insufficienze tecniche sia per gli elevati costi, che porta ad un delicato equilibrio tra domanda ed offerta. Eventuali shock non possono essere calmierati con il ricorso alle scorte di magazzino.
- **Elevata inelasticità della domanda**, che risulta poco sensibile a variazioni di prezzo, essendo l'energia elettrica un bene essenziale e non sostituibile.
- **Incidenza delle variabili meteo** come la temperatura (che, in particolare, influenza la domanda di energia per il condizionamento estivo o il riscaldamento invernale) o la ventosità e l'insolazione (che determinano il livello di generazione eolica o fotovoltaica).

- **Vincoli fisici del sistema elettrico.** La limitata capacità di trasporto della rete elettrica, insieme alla sua complessa gestione in tempo reale, rende efficiente dal punto di vista sistemico organizzare il mercato della commodity tenendo conto, ex ante, dei principali vincoli di rete.
- **La composizione dell'insieme delle fonti di generazione** caratterizzata da uno spettro di impianti che differiscono sia in termini di flessibilità che di costi marginali.

Viene illustrata di seguito la caratterizzazione che la letteratura è concorde ad attribuire alle serie storiche dei prezzi spot, in parte derivante dalle peculiarità del mercato elettrico precedentemente descritte.

Specificità delle serie storiche dei prezzi spot:

- **Elevata frequenza di osservazioni.**
Il clearing, di norma, avviene su base oraria (tramite meccanismo d'asta) con frequenza giornaliera.
- **Elevata volatilità.**
Nelle applicazioni finanziarie con l'espressione "volatilità del prezzo" si intende comunemente la deviazione standard della distribuzione dei rendimenti logaritmici del prezzo. Studiare la serie dei rendimenti logaritmici ($r_t = \log P_t - \log P_{t-1}$) permette di eliminare la non stazionarietà in media, perdendo informazioni sulle tendenze di lungo periodo ma rendendo più evidenti le caratteristiche della stagionalità. Dagli ultimi dati resi disponibili dal GME per l'anno 2022 si evince che il PUN è variato da un minimo di 10 €/MWh a un massimo di 689 €/MWh. La volatilità calcolata sui dati risulta del 10.5%. Questo è dovuto al limite di immagazzinamento (nonstorability) e la conseguente necessità di garantire istante per istante l'equilibrio tra domanda e offerta, che rendono il prezzo dell'energia elettrica eccezionalmente volatile, ben oltre gli standard dei mercati energetici. Per bassi livelli di domanda entrano nel mercato solo generatori che forniscono energia a prezzi marginali bassi. Quando la domanda aumenta non può essere allocata ricorrendo alle "scorte di magazzino", quindi entrano nel sistema anche i generatori con prezzi marginali più alti che determinano un aumento del PUN.
- **Volatilità non costante nel tempo.**
A periodi di bassa volatilità si alternano periodi di volatilità maggiore (volatility clustering).
- **Picchi di prezzo:** cioè variazioni del prezzo significative, improvvise e non prevedibili, detti spikes.
In questi casi, il prezzo si muove verso valori estremi molto velocemente per poi tornare con altrettanta rapidità a livelli considerati nella norma rispetto al suo andamento.

Di norma, gli spikes sembrano apparentemente isolati, ma in realtà sono accompagnati da salti del prezzo che, seppur di intensità molto inferiore, descrivono un cluster di alta volatilità. Se, ad esempio, condizioni atmosferiche avverse e inaspettate generano uno shock del prezzo, è plausibile che queste persistano per un periodo di tempo e non si esauriscano in un solo istante, comportando una maggiore incertezza nei mercati. Spikes e volatilità sono quindi interconnessi, e hanno cause comuni come la non storabilità e la natura essenziale del bene dell'elettricità. Molti acquirenti del mercato elettrico sono disposti a corrispondere qualsiasi tipo di prezzo pur di assicurarsi un approvvigionamento sufficiente e continuo della commodity, soprattutto nelle ore in cui la richiesta di energia è maggiore e la capacità di generazione dei produttori è portata ai limiti massimi.

- **Regressione verso la media (mean reversion):** cioè il fenomeno di ritorno rapido verso il valore medio della serie subito dopo uno spike di prezzo.
- **Fenomeni di stagionalità con frequenze intragiornaliera, settimanale e mensile.** Pur presentando una notevole volatilità, il prezzo spot mostra comunque segni di prevedibilità. Sono molteplici gli studi che individuano diverse forme di stagionalità dei prezzi, che dipendono fortemente dallo stato o dalla regione che si sta analizzando e dal mix energetico che la caratterizza. In Italia, il gas naturale copre più del 40% del parco di generazione, quindi i prezzi dell'energia tendono a seguire la stagionalità tipica del gas. Tuttavia, l'aumento dell'utilizzo di fonti rinnovabili chiama in causa anche i fattori meteorologici. Se un mercato elettrico serve un'area soggetta a clima temperato come l'Italia, caratterizzata da inverni abbastanza rigidi ed estati calde, sarà molto più intenso l'uso del riscaldamento piuttosto che dei condizionatori in estate, determinando prezzi maggiori durante l'inverno. La ragione della stagionalità del prezzo risiede principalmente nella variazione della domanda, per questo durante i giorni lavorativi si registrano prezzi mediamente più alti rispetto al weekend in cui molte attività rimangono chiuse causando una diminuzione della domanda. Similmente, anche la stagionalità intragiornaliera dei prezzi è dovuta ad una variazione della domanda, che nelle ore lavorative e diurne è più alta, mentre diminuisce notevolmente nelle ore notturne portando ad un decremento del PUN. Per questo motivo, come già specificato nel primo capitolo, i contratti a termine con obbligo di consegna e ritiro negoziati nel MTE possono essere di due tipologie: base load (fornitura costante in tutte le ore) e peak load (fornitura nelle sole ore comprese tra le 8.00 e le 20.00 dei giorni lavorativi).
- **Distribuzione non normale.** È noto da diversi studi che la distribuzione di probabilità dei rendimenti logaritmici del prezzo dell'elettricità si allontanano significativamente da quella normale, poiché le osservazioni presentano un'eccessiva curtosi. Siamo di fronte ad una distribuzione leptocurtica (figura 3.1): le code sono più pesanti rispetto alla distribuzione normale, la quale non è adatta a descrivere l'andamento dei dati perché sottostima la probabilità

di eventi estremi. Infatti, come abbiamo già evidenziato, variazioni giornaliere molto ampie dei prezzi sono relativamente frequenti nei mercati elettrici, lo stesso vale per gli improvvisi picchi di prezzo.

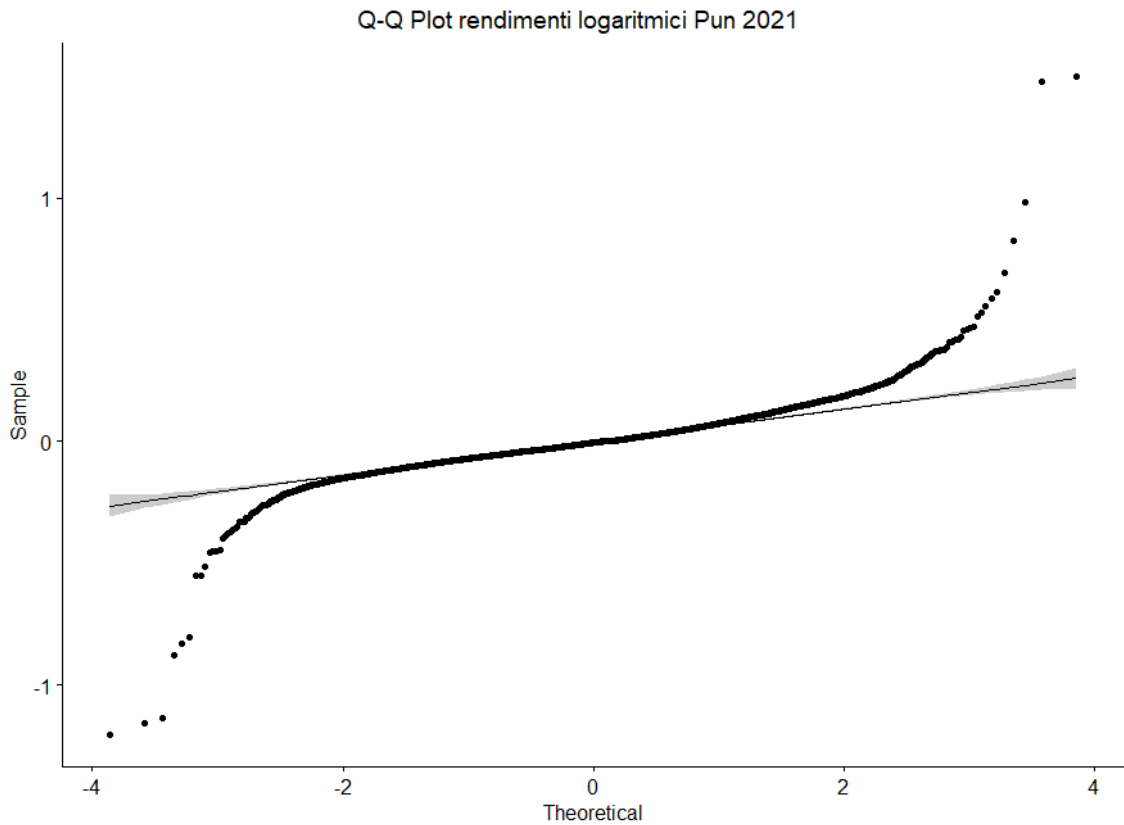


Figura 3.1: Q-Q plot rendimenti log PUN 2021

Alla luce delle caratteristiche evidenziate si procederà all'individuazione e alla descrizione di modelli adatti a trattare le serie storiche in questione.

3.2 Test di non linearità

Con l'obiettivo di individuare un modello che descriva l'andamento dei prezzi spot è necessario tenere in considerazione diversi aspetti.

Nel caso in cui si ritenga opportuno includere variabili esogene, la scelta di un modello complesso che comprenda tutte le caratteristiche della variabile risposta renderebbe più semplice

catturare il comportamento dei dati di un determinato campione, senza però alcun vantaggio esplicativo rischiando, inoltre, di incorrere in problemi di sovrapparametrizzazione. D'altra parte, un modello troppo semplice riuscirebbe a descrivere solo parzialmente la dinamica del fenomeno considerato. Una prima difficoltà si riscontra, quindi, nel trovare il giusto equilibrio tra parsimonia e bontà di adattamento, essendo molteplici le peculiarità dei prezzi spot: volatilità elevata e clusterizzata, stagionalità mensile, settimanale ed infragionaliera, spikes, mean reversion, distribuzione con code pesanti. Adunarle tutte in un unico modello va ben oltre gli intenti di questo lavoro, in cui ci concentreremo sulla natura non lineare del processo generatore dei dati.

Per introdurre il concetto di non linearità è opportuno fornire una definizione di processo stocastico lineare :

Definizione 1. Sia X_t un processo stocastico, esso si dice lineare se può essere scritto come

$$X_t = \mu + \sum_{i=0}^{\infty} \Psi_i a_{t-i} \quad (3.1)$$

dove μ è un numero reale, $\Psi_0 = 1$, Ψ_i sono numeri reali tali che $\sum_{i=0}^{\infty} |\Psi_i| < \infty$ e $\{a_t\}$ è una successione di variabili casuali i.i.d. con media zero e funzione di densità ben definita.

In altri termini, se un processo X_t è non lineare non può essere adeguatamente rappresentato da modelli lineari, perché non è possibile scriverlo come combinazione lineare di variabili indipendenti ed identicamente distribuite.

Per verificare l'ipotesi di non linearità delle serie storiche la letteratura mette a disposizione diversi test, sia parametrici che non, aventi il seguente sistema di verifica di ipotesi:

- **ipotesi nulla H_0 :** il processo generatore dei dati è lineare, perciò la serie storica considerata può essere adeguatamente descritta da un modello lineare.
- **ipotesi alternativa H_1 :** negazione ipotesi nulla, la serie considerata non è lineare.

Descriviamo in seguito alcuni tra i vari test fruibili.

3.2.1 BDS test

BDS test proposto da Brock e al. (1987), è uno dei test non parametrici più utilizzati per verificare l'ipotesi nulla, nella quale si assume che una data serie storica sia costituita da realizzazioni di variabili aleatorie i.i.d.. È stata dimostrata (Brock, 1991) una buona potenza del test anche nel discriminare tra ipotesi nulla di linearità e ipotesi alternativa di non linearità di una serie. Il test si basa sull'idea di integrale di correlazione che rappresenta una misura della frequenza con cui un pattern temporale si ripete nei dati.

Definizione 2. Sia $\{x_t | t = 1, \dots, T\}$ una serie storica dove T denota la grandezza del campione e sia $X_T^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$, per $t=m, \dots, T$, con m numero intero positivo. L'integrale di correlazione m -dimensionale è definito come

$$C_{(m,\varepsilon)} = \lim_{T_m \rightarrow \infty} \frac{2}{T_m(T_m - 1)} \sum_{m \leq s < t \leq T} I(X_t^m, X_s^m | \varepsilon) \quad (3.2)$$

dove $T_m = T - m + 1$ e $I(x_t^m, x_s^m | \varepsilon)$ è una funzione indicatrice che è pari a 1 quando $\|X_t^m - X_s^m\| < \varepsilon$ e 0 altrimenti.

Per un dato ε l'equazione (3.2) calcola la probabilità che due punti m -dimensionali siano a una distanza pari a ε . Intuitivamente, nel caso in cui le variabili fossero indipendenti e identicamente distribuite $C_{(m,\varepsilon)}$ e $C_{(1,\varepsilon)}^m$ dovrebbero essere uguali, ovvero la probabilità che due punti si trovino a distanza ε dovrebbe essere la stessa sia per punti m -dimensionali che per punti unidimensionali, poichè le m dimensioni incluse sono tra loro indipendenti. Sotto l'assunzione di indipendenza $P(A \cap B) = P(A)P(B)$, da qui :

$$C_{(m,\varepsilon)} = P(|X_t - X_s| < \varepsilon, |X_{t-1} - X_{s-1}| < \varepsilon, \dots, |X_{t-m+1} - X_{s-m+1}| < \varepsilon) = P(|X_t - X_s| < \varepsilon) = C_{(1,\varepsilon)}^m \quad (3.3)$$

La statistica test di BDS si basa sul confronto tra $C_{m,\varepsilon}$ e $C_{1,\varepsilon}$ ed è la seguente :

$$D_{(m,\varepsilon)} = \frac{\sqrt{T}[\hat{C}_{(m,\varepsilon)} - \hat{C}_{(1,\varepsilon)}^m]}{s(m, \varepsilon)} \quad (3.4)$$

dove

$$C_{(k,\varepsilon)} = \frac{2}{T_k(T_k - 1)} \sum_{k \leq s < t \leq T} I(X_t^k, X_s^k | \varepsilon) \quad (3.5)$$

3.2.2 Test di McLeod e Li

Un altro test non parametrico utilizzato per verificare la non linearità di una serie è quello di McLeod e Li (1983). Questo test si basa su una proprietà dimostrata da Granger and Andersen (1978) per cui se una serie è stazionaria e lineare allora: $corr(x_t^2, x_{t-k}^2) = (corr(x_t, x_{t-k}))^2$, l'assenza della suddetta proprietà indica non linearità.

Può accadere che i residui di un modello adattato ad una serie risultino incorrelati, ma i residui al quadrato presentino una struttura di correlazione: questo fenomeno ci suggerisce la presenza di non linearità nella serie oltre che eteroschedasticità condizionata.

Il test è condotto esaminando la statistica di ljung-box applicata ai residui al quadrato derivanti dall'adattamento di un modello ARMA alla serie storica interessata. La statistica test utilizzata è la seguente :

$$Q^*(m) = T(T+2) \sum_{i=1}^m \frac{\hat{\rho}_{aa}^2(i)}{T-i} \quad (3.6)$$

dove T è la numerosità campionaria , m indica il numero di autocorrelazioni considerate,

$$\hat{\rho}_{aa}^2(i) = \frac{\sum_{t=i+1}^T (\hat{a}_t^2 - \hat{\sigma}^2)(\hat{a}_{t-i}^2 - \hat{\sigma}^2)}{\sum_{t=1}^T (\hat{a}_t^2 - \hat{\sigma}^2)}$$

è la funzione di autocorrelazione dei residui al quadrato al lag i-esimo . Sotto ipotesi nulla, nella quale si assume che la serie sia lineare e conseguentemente che il modello ARMA adattato sia adeguato, la statistica test $Q^*(m)$ è asintoticamente distribuita come una χ_m^2

3.2.3 RESET test

Un test parametrico per verificare la proprietà di non linearità di una serie storica è il test RESET (Regression Specification Error Test) di Ramsey (1969).

Ramsey propone un test parametrico di errata specificazione del modello di regressione lineare. Allo stesso tempo, il test verifica la correttezza della forma funzionale del modello lineare e l'eventuale significatività di elementi non lineari omissi. Tale test è detto di linearità

perché regredisce la variabile dipendente y sulle esplicative e sulle potenze di \hat{y} , dopodiché effettua un test F di significatività dei parametri non lineari. Nell'ambito delle serie storiche il test è utilizzato per verificare la linearità del processo generatore dei dati. La procedura di verifica si articola in tre step e differisce leggermente da quella appena descritta, poichè si avvale della diagnostica dei residui del modello assunto.

Si consideri un modello lineare adattato ad una serie storica di cui si vuole verificare la linearità, come ad esempio un $AR(p)$:

$$x_t = \mathbf{X}'_{t-1} \Phi + a_t \quad (3.7)$$

dove $\mathbf{X}_{t-1} = (1, x_{t-1}, \dots, x_{t-p})'$ e $\Phi = (\Phi_0, \Phi_1, \dots, \Phi_p)'$.

Nel primo step si ottiene la stima del parametro Φ , indicata con $\hat{\Phi}$, attraverso il metodo dei minimi quadrati. Si calcolano $\hat{x}_t = \mathbf{X}'_{t-1} \hat{\Phi}$, i residui $\hat{a}_t = x_t - \hat{x}_t$ e la somma dei residui al quadrato:

$$SSR_0 = \sum_{t=p+1}^n \hat{a}_t^2 \quad (3.8)$$

, dove n indica la numerosità campionaria.

Nel secondo step si considera la regressione lineare

$$\hat{a}_t = \mathbf{X}'_{t-1} \hat{\alpha}_1 + \mathbf{M}'_{t-1} \hat{\alpha}_2 + v_t \quad (3.9)$$

dove $\mathbf{M}'_{t-1} = (x_t^2, \dots, x_t^s + 1)$, per $s > 1$.

Se il modello lineare è correttamente specificato ci si aspetta che i residui non presentino alcun tipo di correlazione con le variabili esplicative, quindi α_1 e α_2 dovrebbero essere uguali a zero. Per verificare questa ipotesi, ovvero la significatività dei parametri α_1 e α_2 , si ricorre al test-F con la seguente statistica test :

$$F = \frac{(SSR_0 - SSR_1)/g}{SSR_1/(n - p - g)} \quad \text{con } g = s + p + 1 \quad (3.10)$$

Sotto l'ipotesi nulla di linearità la 3.10 si distribuisce come una $F_{g, n-p-g}$.

Nel terzo step si applica il test-F dopo aver calcolato la devianza residua SSR_1

$$\hat{v} = \hat{a} - \mathbf{X}'_{t-1}\hat{\alpha}_1 - \mathbf{M}'_{t-1}\hat{\alpha}_2 \quad (3.11)$$

$$SSR_1 = \sum_{t=p+1}^n \hat{v}_t^2 \quad (3.12)$$

3.2.3.1 Keenan test

Keenan(1985) propone un test di non linearità che si basa sul test RESET di Ramsey(1969). Nello specifico viene modificato il secondo step del test RESET per evitare la multicollinearità tra \hat{x}_t^2 e \mathbf{X}_{t-1} . La regressione lineare 3.9 viene divisa in due step. Nello step 2.1 si rimuove la dipendenza lineare di \hat{x}_t^2 su \mathbf{X}_{t-1} applicando la regressione

$$\hat{x}_t^2 = \mathbf{X}'_{t-1}\beta + u_t \quad (3.13)$$

e si ottengono i residui $\hat{u}_t = \hat{x}_t^2 - \mathbf{X}'_{t-1}\hat{\beta}$. Nello step 2.2 si considera la regressione lineare

$$\hat{a}_t = \hat{u}_t\alpha + v_t \quad (3.14)$$

e si ottiene la somma dei quadrati dei residui $SSR_1 = \sum_{t=p+1}^n (\hat{a}_t - \hat{u}_t\hat{\alpha})^2 = \sum_{t=p+1}^n \hat{v}_t^2$ per testare l'ipotesi nulla $\alpha = 0$

3.3 Modelli non lineari

Esistono diverse classi di modelli adattabili alle serie storiche non lineari. Noi ci concentreremo sulla classe dei modelli regime-switching.

3.3.1 Modelli autoregressivi a soglia (SETAR)

Tra i modelli della classe regime-switching troviamo i modelli SETAR (*Self-Exciting Threshold AutoRegressive*)

Definizione 3. Una serie storica X_t segue un modello SETAR a m regimi se soddisfa :

$$X_t = \begin{cases} \Phi_{0,1} + \sum_{i=1}^{p_1} \Phi_{i,1} x_{t-i} + \sigma_1 \varepsilon_t, & \text{se } x_{t-d} \leq r_1, \\ \Phi_{0,2} + \sum_{i=1}^{p_2} \Phi_{i,2} x_{t-i} + \sigma_2 \varepsilon_t, & \text{se } r_1 < x_{t-d} \leq r_2, \\ \dots \\ \Phi_{0,m} + \sum_{i=1}^{p_m} \Phi_{i,m} x_{t-i} + \sigma_m \varepsilon_t, & \text{se } r_{m-1} < x_{t-d}, \end{cases} \quad (3.15)$$

dove $m > 1$ è un intero positivo e indica il numero di regimi, $d > 0$ è un intero positivo e indica il ritardo, r è il valore soglia e p indica l'ordine del regime. Si assume che $\varepsilon_t \sim WN(0, 1)$ e Φ_i , σ_i siano parametri ignoti.

Un caso particolare dei SETAR sono i modelli TAR (Threshold AutoRegressive) la cui forma differisce dall'equazione 3.15 poichè la variabile x_{t-d} presente nel modello SETAR viene sostituita da z_t che costituisce una variabile soglia esogena.

Noi ci concentreremo sui modelli SETAR a due regimi che presentano la seguente struttura

$$X_t = \begin{cases} \Phi_0 + \sum_{i=1}^{p_1} \Phi_i x_{t-i} + \sigma_1 \varepsilon_t, & \text{se } x_{t-d} \leq r, \\ \theta_0 + \sum_{i=1}^{p_2} \theta_i x_{t-i} + \sigma_2 \varepsilon_t, & \text{se } x_{t-d} > r \end{cases} \quad (3.16)$$

il modello nell'equazione 4.1 può essere riformulato in modo più compatto, imponendo $p_1 = p_2 = 2$, come :

$$X_t = \Phi_0 \sum_{i=1}^p \Phi_i x_{t-i} + \sigma_1 \varepsilon_t + I(x_{t-d} > r) (\beta_0 + \sum_{i=1}^p \beta_i x_{t-i} + \lambda \varepsilon_t) \quad (3.17)$$

dove $I(x_{t-d} > r)$ è una funzione indicatrice uguale a 1 se $x_{t-d} > r$ e uguale a 0 altrimenti, $\beta_i = \theta_i - \Phi_i$ per $i = 0, \dots, p$ e $\lambda = \sigma_2 - \sigma_1$. In questa formulazione il cambiamento strutturale del modello viene evidenziato da β_i che denota la modifica nell' i -esimo coefficiente.

La struttura della 3.17 è una "composizione" di modelli lineari autoregressivi AR(p). Quando la variabile X_{t-d} assume determinati valori avviene un cambiamento di regime che consiste in una modifica dei parametri del modello. L'alternarsi di diverse componenti lineari che diventano operative ad ogni cambio di regime conferisce ai modelli SETAR la capacità di cogliere le variazioni di livello dovute agli spikes presenti nella serie dei prezzi spot dell'elettricità.

Generalmente si ricorre ai modelli SETAR nel caso in cui i fenomeni oggetto di studio siano caratterizzati da asimmetrie, periodicità o cicli, tuttavia questi modelli presentano diverse criticità come la difficile derivazione di alcune proprietà teoriche (ergodicità) e le limitate proprietà di cui godono gli stimatori dei parametri ottenuti con il metodo dei minimi quadrati. Inoltre, è difficile stabilire le proprietà degli stimatori dei parametri del modello ottenuti con il metodo dei minimi quadrati.

Vediamo nel dettaglio la procedura di stima.

Siano $\Phi = (\Phi_0, \Phi_1, \dots, \Phi_p)'$ e $\theta = (\theta_0, \theta_1, \dots, \theta_p)'$ i vettori parametrici dei due regimi. Siano $\Theta = (\Phi', \theta, r, d)'$ il vettore dei coefficienti del modello dell'equazione 3.17 e Θ_0 il vero vettore dei coefficienti. Si supponga di avere a disposizione il campione $\{x_1, \dots, x_n\}$ così da definire la media condizionata del modello $E_{\Theta}(x_t | \mathbf{F}_{t-1})$ dati il vettore dei coefficienti Θ e $\mathbf{F}_{t-1} = (x_{t-1}, x_{t-2}, \dots, x_1)$.

La funzione dei minimi quadrati condizionati è la seguente:

$$L_n(\Theta) = \sum_{t=p+1}^n [x_t - E_{\Theta}(x_t | \mathbf{F}_{t-1})]^2 \quad (3.18)$$

La funzione obiettivo 3.18 viene minimizzata minimizzando a sua volta $\mathbf{S}(r,d)$, dove $\mathbf{S}(r,d)$ deve essere la più piccola somma dei quadrati dei residui del modello dato, fissati la soglia r e il ritardo d . Φ e θ vengono stimati il metodo dei minimi quadrati ordinari :

$$x_t = \Phi_0 + \sum_{i=1}^p \Phi_i x_{t-i} + \varepsilon_t, \quad \text{con } t \in S_1 \quad (3.19)$$

$$x_t = \Theta_0 + \sum_{i=1}^p \Theta_i x_{t-i} + \varepsilon_t, \quad \text{con } t \in S_2 \quad (3.20)$$

dove S_1 e S_2 costituiscono gli insiemi degli indici temporali t per i regimi 1 e 2 rispettivamente e ε_t indica il termine di errore.

Le varianze σ_1^2 e σ_2^2 vengono invece stimate dal MSE residuo delle due regressioni lineari 3.19 e 3.20. A questo punto la funzione obiettivo $L_n(\Theta)$ può essere minimizzata utilizzando $\mathbf{S}(r,d)$.

La procedura di costruzione e selezione del modello non è convenzionalmente definita ma

esistono più alternative percorribili.

Considerando un SETAR a due regimi, come primo passo del processo di costruzione si può ipotizzare un massimo per l'ordine p dei modelli autoregressivi assumendo che il ritardo d soddisfi $1 \leq d \leq p$. Inoltre si assume che la soglia r sia uguale a un valore della statistica ordinata dei dati; per statistica ordinata si intende $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(t)} \leq x_{(t+1)} \leq \dots \leq x_{(n)}$. Sia quindi $r \in \{x_l, x_{l+1}, \dots, x_u\}$, dove l e u sono la parte intera dei numeri reali αn e $(1 - \alpha)n$, rispettivamente. α deve essere compreso tra 0 e 0.3 e la sua scelta dipende anche dal bilanciamento del numero di osservazioni in ogni regime. Dopo aver fissato α e il massimo ordine autoregressivo p , vengono selezionati il ritardo d , l'ordine p_i per ogni regime, i coefficienti autoregressivi e il valore soglia r attraverso il criterio di informazione di Akaike (AIC).

Un'altra procedura di costruzione del modello viene illustrata in Tsay(1989) e può essere così riassunta: innanzitutto si fissa l'ordine p delle componenti AR e un insieme S di possibili valori per il ritardo. In seguito si ricorre al test di non linearità $F(d,p)$ la cui statistica si basa sui risultati ottenuti applicando il metodo dei minimi quadrati ai modelli autoregressivi stimati per il p selezionato e tutti gli elementi di S . Viene scelto il ritardo d_p dell'insieme S tale che

$F(p, d_p) = \max_{(d \in S)} F(p, d)$. Una volta definiti p e d si utilizzano strumenti grafici, cioè gli scatterplots, per allocare i valori soglia che vengono individuati come punti di cambiamento del modello.

Possono anche essere utilizzati metodi esclusivamente grafici e quindi non parametrici, come ad esempio lo scatterplot di x_t vs x_{t-d} , in cui la soglia corrisponde al punto di cambiamento del modello stimato attraverso la regressione polinomiale locale (generalizzazione della media mobile).

3.3.2 Markov switching model (MSM)

Un'altra tipologia di modelli non lineari ampiamente diffusa è quella dei *Markov switching model*. Anche in questo caso ci concentreremo sulla specificazione del modello a due regimi.

Definizione 4. Una serie storica X_t segue un modello MSM a 2 regimi se soddisfa:

$$X_t = \begin{cases} \sum_{i=1}^p \Phi_{1,i} x_{t-i} + \sigma_1 \varepsilon_t & \text{se } S_t = 1 \\ \sum_{i=1}^p \Phi_{2,i} x_{t-i} + \sigma_2 \varepsilon_t & \text{se } S_t = 2 \end{cases} \quad (3.21)$$

dove $\Phi_{m,i}$ sono i coefficienti del regime m al ritardo i , $\sigma_i > 0$, e $\varepsilon_t \sim N(0, 1)$. S_t rappresenta una variabile che indica lo stato del processo al tempo t , in questo caso può assumere solo due valori: 0 o 1.

Come i modelli SETAR anche la struttura dei MSD a due stati è costituita da due modelli autoregressivi che diventano operativi a seconda del valore assunto dalla variabile di stato S_t . Il comportamento di questa variabile è governato dalle probabilità di transizione

$$P(S_t = 2 | S_{t-1} = 1) = \eta_1, \quad P(S_t = 1 | S_{t-1} = 2) = \eta_2 \quad (3.22)$$

da cui deriva la matrice di transizione :

$$\begin{bmatrix} 1 - \eta_1 & \eta_1 \\ 1 - \eta_2 & \eta_2 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (3.23)$$

Che deve soddisfare la seguente proprietà $p_{11} + p_{12} = p_{21} + p_{22} = 1$. Nella precedente notazione p_{ij} rappresenta la probabilità di transizione dallo stato $S_{t-1} = i$ allo stato $S_t = j$. Essendo p_{ii} la probabilità che lo stato del modello rimanga costante e uguale a i da $t-1$ a t , la durata attesa della permanenza nello stato i viene calcolata come $1/(1 - p_{ii})$. Dalle definizioni, la struttura dei modelli MSM potrebbe sembrare molto simile a quella dei modelli SETAR, tuttavia una differenza significativa si riscontra nelle modalità di passaggio da un modello autoregressivo all'altro. Mentre nei modelli SETAR il cambiamento di regime avviene seguendo un disegno deterministico dopo che il valore soglia è stato definito, per i modelli MSM il cambiamento di stato è governato da leggi probabilistiche, quindi ha una struttura stocastica, cioè "casuale". Questa peculiarità dei MARKOV switching model è evidente nelle applicazioni, infatti raramente lo stato S_t di un dato MSM è certo ma si può fare inferenza sulle probabilità del tipo $P(S_t = 1 | \text{dati, modello})$

Anche i MSM presentano alcune criticità: le loro proprietà teoriche sono difficili da derivare e le condizioni necessarie per l'esistenza dei momenti di ordine superiore e per la stazionarietà in senso forte non sono di semplice definizione. Poiché la variabile di stato è latente, ovvero non direttamente osservabile, i modelli MSM appartengono alla famiglia statistica delle variabili latenti, ed in quanto tali, il processo di stima dei parametri risulta essere ostico e richiede calcoli complessi.

3.3.3 Smooth transition autoregressive model (STAR)

L'ultima classe di modelli non lineari che analizzeremo è quella dei STAR, la cui struttura è molto simile a quella dei SETAR. La differenza principale tra i due modelli risiede nel meccanismo che governa la transizione da un regime ad un altro .

Definizione 5. Una serie storica X_t segue un modello STAR a 2 regimi se soddisfa :

$$\begin{aligned} X_t = & (\Phi_{0,1} + \Phi_{1,1}x_{t-1} + \dots + \Phi_{p,1}x_{t-p} + [1 - G(s_t|\lambda, c)] + \\ & + (\Phi_{0,2} + \Phi_{1,2}x_{t-1} + \dots + \Phi_{p,2}x_{t-p} + [G(s_t|\lambda, c)] + a_t \end{aligned} \quad (3.24)$$

dove s_t è la variabile soglia, $[G(s_t|\lambda, c)]$ è la funzione di transizione, a_t è una sequenza di variabili i.i.d con media 0 e varianza $\sigma_a^2 > 0$, λ e c sono parametri di scale e di posizione della funzione di transizione.

Dalla struttura del modello si nota come il cambiamento di regime sia governato dalla funzione di transizione $[G(s_t|\lambda, c)]$ la quale, a differenza della funzione indicatrice presente nei modelli SETAR, è continua in un range che va da 0 a 1. La continuità della funzione permette di eliminare i punti di discontinuità presenti in corrispondenza delle soglie nei modelli SETAR, dove il passaggio da un regime all'altro avveniva in modo brusco, caratteristica non riscontrabile nelle serie storiche reali che si evolvono gradualmente nel tempo.

Le tipologie di funzione di transizione più comunemente utilizzate sono due: funzione esponenziale e funzione logistica.

La funzione logistica è definita come:

$$G(s_t|\lambda, c) = \frac{1}{1 + \exp[-\lambda(s_t - c)]} \quad (3.25)$$

ed il modello che ne fa utilizzo viene detto logistic STAR (LSTAR)

La funzione esponenziale è definita come:

$$G(s_t|\lambda, c) = 1 - \exp[-\lambda(s_t - c)^2] \quad (3.26)$$

ed il modello che ne fa utilizzo viene detto exponential STAR (ESTAR)

Le due funzioni si differenziano poichè quella esponenziale 3.26 tende a $\pm\infty$ quando la soglia $S_t \rightarrow \pm\infty$ mentre la logistica 3.25 tende a ± 1 quando $\lambda(s_t - c) \rightarrow \infty$

Capitolo 4

Stima di modelli e test di non linearità. Applicazione ai prezzi elettrici.

Dopo aver descritto alcune categorie di modelli non lineari e di test per verificare la non linearità di serie storiche, passiamo alla loro applicazione sulla serie storiche del PUN. L'approccio seguito in questo studio è ampiamente diffuso nell'analisi delle serie storiche dei prezzi dell'elettricità (Weron, 2006) e consiste nel modellare ogni ora separatamente. La motivazione di questa scelta risiede nelle modalità di formazione dei prezzi nel MGP: come spiegato nel 3 il prezzo di equilibrio viene fissato per ogni ora del giorno successivo attraverso un meccanismo d'asta. Durante l'asta le offerte di compravendita per ogni ora del giorno successivo non sono note, per questo i partecipanti si basano sui prezzi orari dei giorni precedenti per presentare le offerte. È quindi logico aspettarsi una relazione più forte tra i prezzi osservati nelle stesse ore di giorni successivi piuttosto che tra ore consecutive della stessa giornata.

In questa applicazione ci concentreremo esclusivamente sulle serie del PUN del Nord, della Sardegna e della Sicilia poiché sono le uniche aree geografiche non interessate dalle modificazioni dell'assetto zonale avvenute durante il nostro periodo di interesse, che va dal 1 Gennaio 2019 fino al 30 Giugno 2022.

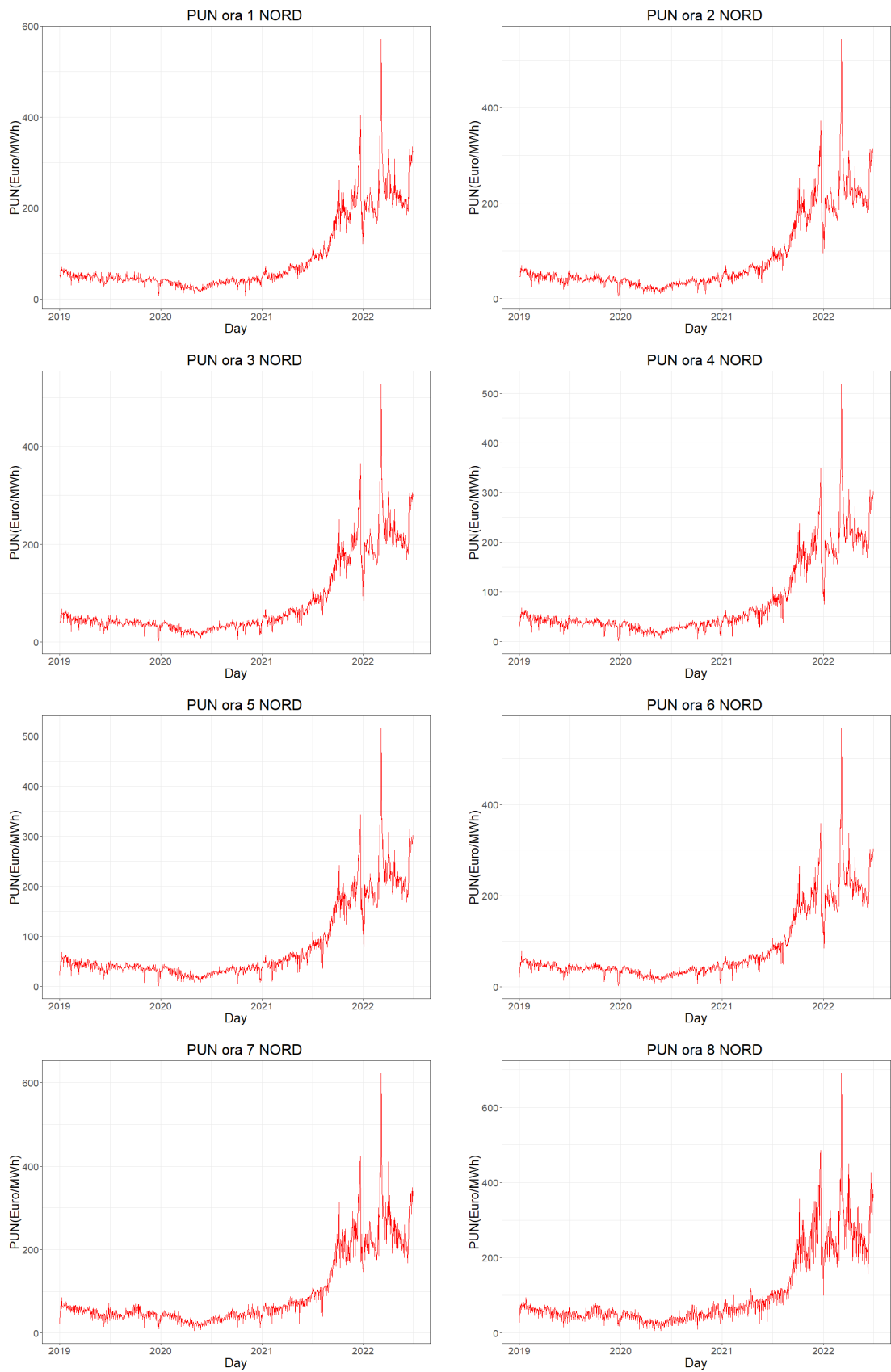


Figura 4.1: Serie storiche PUN nord da ora 1 a ora 8

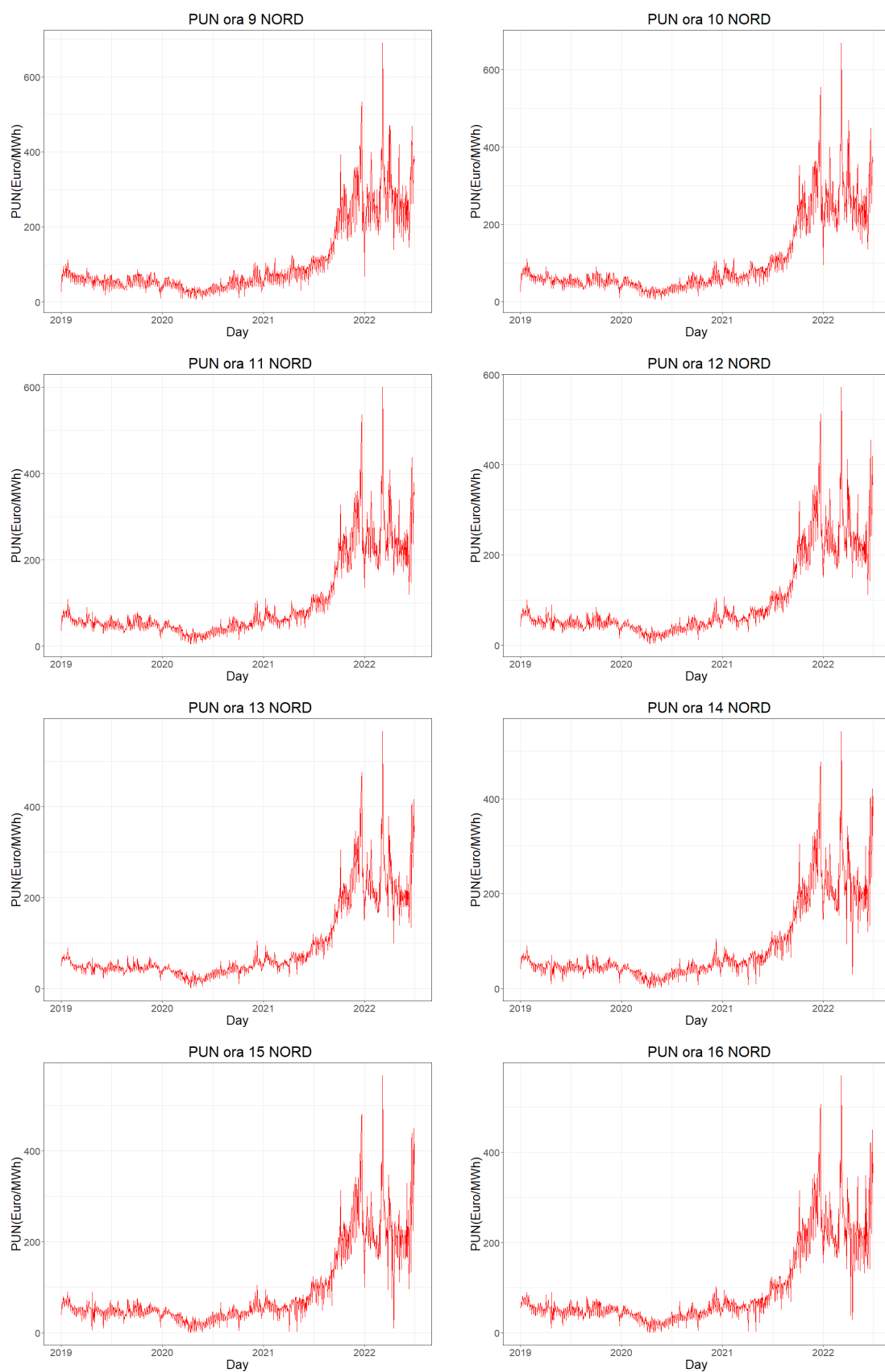


Figura 4.2: Serie storiche PUN Nord da ora 9 a ora 16

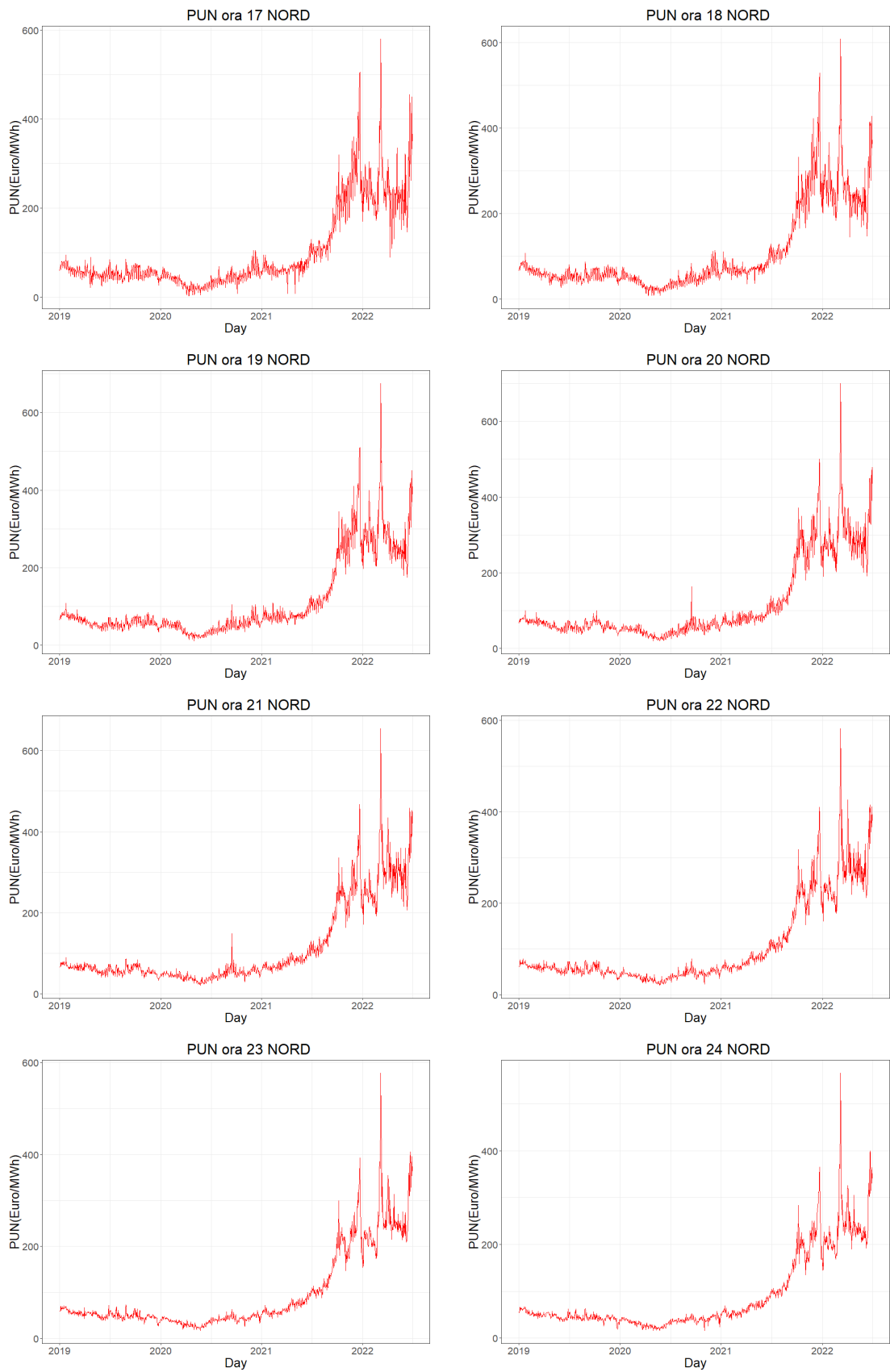


Figura 4.3: Serie storiche PUN Nord da ora 17 a ora 24

4.1 Applicazione test di non linearità

Dopo avere ricavato le 72 (24 ore * 3 aree geografiche) serie storiche di interesse (serie del Nord visibili in figura 4.1 4.2 4.3), applichiamo ad ognuna due test di non linearità per giustificare l'utilizzo di modelli non lineari. In particolare i test scelti sono il BDS test e il test di Keenan, che applichiamo ai rendimenti logaritmici delle serie.

Nel caso del BDS l'ipotesi nulla di linearità viene sempre rifiutata con p-value molto bassi, fatta eccezione per le serie dell'ora 19 e dell'ora 20 dell'area Nord. Per quanto riguarda il Keenan test l'ipotesi di linearità **non** viene rifiutata nelle serie dell'ora 2,6,20,21,22,23 del Nord e 21,22,23 della Sardegna.

4.2 Stima dei modelli SETAR

Ossevando le funzioni di autocorrelazione delle serie del PUN 4.4 4.5, dopo aver applicato una differenza prima, è visibile una componente periodica settimanale (di periodo 7) evidenziata dagli alti coefficienti di correlazione ai ritardi multipli di 7. Si sceglie quindi di applicare un SETAR(7,7,1) seguendo anche lo studio di L. Grossi e F. Nan (2019) dove si adattano alle 24 serie storiche del PUN, che vanno dal 1 Gennaio 2013 al 12 Dicembre 2014, 49 diversi modelli SETAR(p,d) risultanti dalle combinazioni dei valori dei parametri $p = \{1, 2, 3, 4, 5, 6, 7\}$ e $d = \{1, 2, 3, 4, 5, 6, 7\}$. Confrontando le 49 medie degli AIC calcolate sulle 24 ore i due modelli migliori sono risultati i seguenti: SETAR(6,1) e SETAR(7,1), rispettivamente primo e secondo classificato. Si è scelto di selezionare il SETAR(7,1) poiché la stagionalità settimanale è più facilmente catturabile con $p=7$.

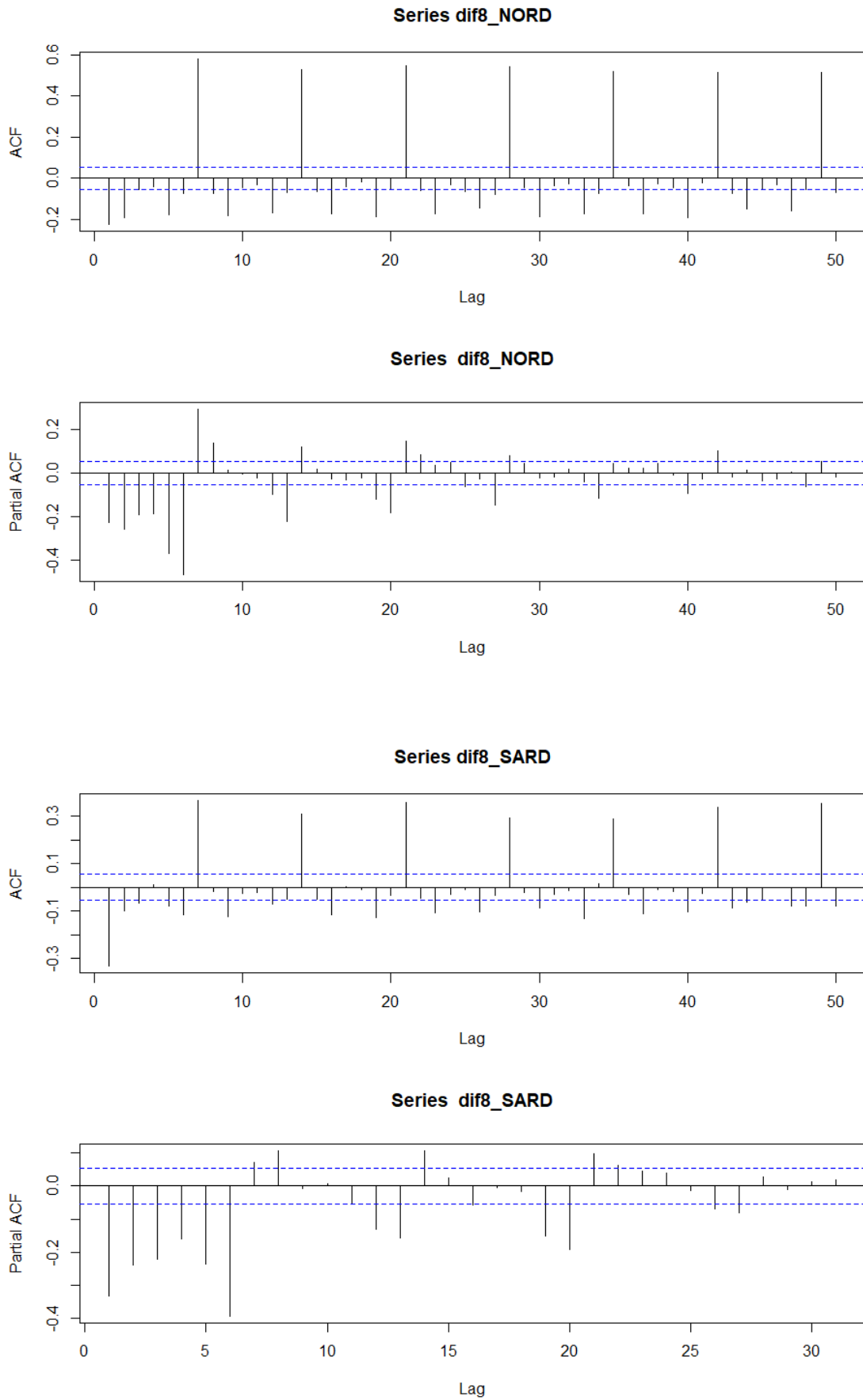


Figura 4.4: ACF e PACF delle serie del log PUN (ora 8 Nord e Sardegna) a cui è stata applicata una differenza prima

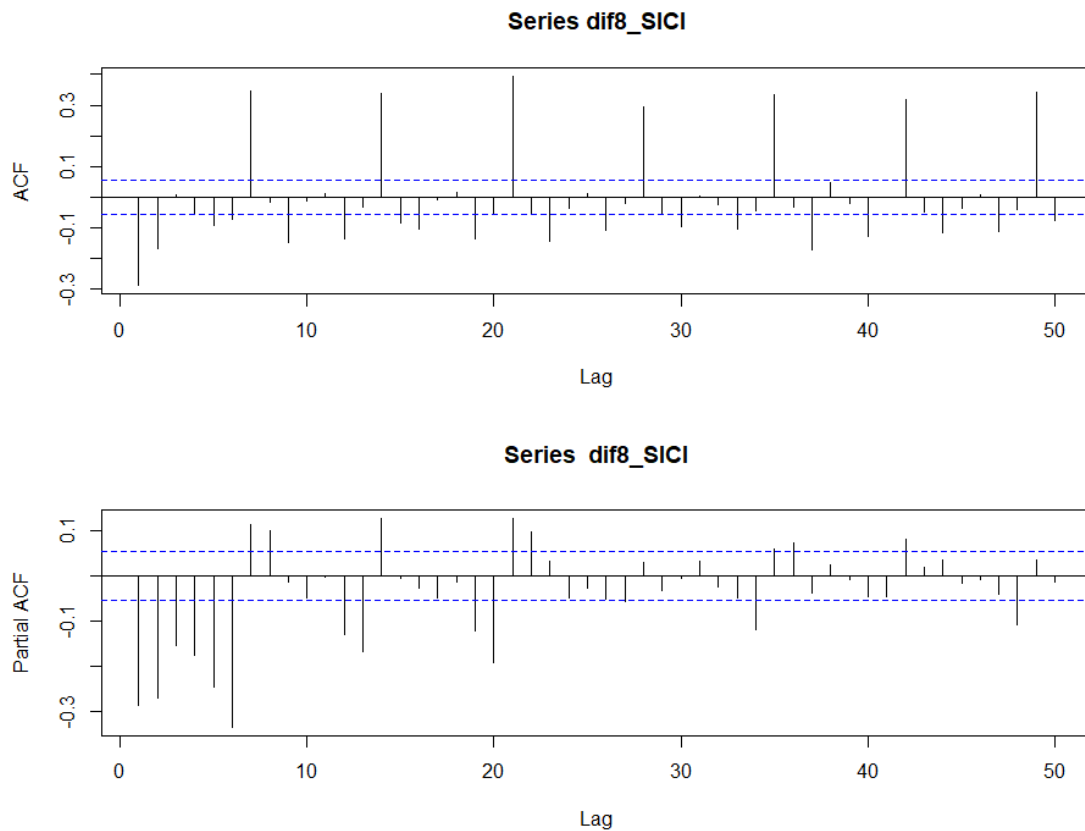


Figura 4.5: ACF e PACF delle serie del log PUN (ora 8 Sicilia) a cui è stata applicata una differenza prima

Il modello applicato risulta essere il seguente

$$X_t = \begin{cases} \Phi_0 + \sum_{i=1}^7 \Phi_i x_{t-i} + \sigma_1 \varepsilon_t, & \text{se } x_{t-1} \leq r, \\ \theta_0 + \sum_{i=1}^7 \theta_i x_{t-i} + \sigma_2 \varepsilon_t, & \text{se } x_{t-1} > r \end{cases} \quad (4.1)$$

Per valutare la bontà di adattamento si confronta l'AIC dei modelli SETAR(7,1) con quello di un modello lineare ARIMA(7,1,7). L'AIC dei modelli SETAR risulta sempre inferiore rispetto a quello dei modelli ARIMA fatta eccezione per le ore 7,8,9,10 19,20,21,22,23 del Nord, per l'ora 9 della Sardegna e l'ora 8 della Sicilia 4.1. Quindi complessivamente i modelli SETAR si adattano meglio alle serie del PUN rispetto agli ARIMA, soprattutto per Sardegna e Sicilia. Conducendo un'analisi dei residui 4.6 4.7 4.8 si nota una correlazione residua ai lag

44 *CAPITOLO 4. STIMA DI MODELLI E APPLICAZIONE AI PREZZI ELETTRICI*

stagionali di periodo 7. Inoltre la distribuzione dei residui si avvicina alla simmetria solo per alcune ore della giornata. Questo indica che nonostante i SETAR siano preferibili agli ARIMA il loro adattamento è ancora migliorabile.

Tabella 4.1: AIC dei modelli ARIMA(7,1,7) e SETAR(7,1) applicati alle 24 sier del PUN per ogni area geografica.

ORA	NORD		SARD		SICI	
	AIC ARIMA	AIC SETAR	AIC ARIMA	AIC SETAR	AIC ARIMA	AIC SETAR
1	-1452.8827	-1554.5213	98.8846	-462.03268	49.62165	-193.46047
2	-1281.5017	-1319.2315	671.2198	12.41807	310.79146	75.34300
3	-857.4299	-992.9429	1000.9819	190.74559	900.57881	547.30008
4	-509.0225	-656.9009	1227.5146	382.59724	1226.07099	905.76504
5	-488.1132	-648.8568	1226.0396	334.93598	1173.52389	867.65614
6	-835.7308	-1021.6000	903.2721	291.90008	578.98684	307.78717
7	-1095.5864	-1023.6110	514.8043	142.58157	-12.11992	-125.68898
8	-961.9370	-810.1409	-164.6096	-228.35173	-175.35096	-52.12573
9	-691.7195	-513.1127	252.3129	288.68244	494.60070	428.85079
10	-832.0300	-709.3886	956.9642	823.41362	916.51069	623.01512
11	-826.2321	-874.5942	1698.4875	1431.84882	1530.30459	1375.54016
12	-664.1743	-818.2855	2246.5211	1922.53858	2056.01857	1900.41732
13	-551.9544	-942.7679	2462.5357	2101.51969	2287.57080	2064.31575
14	139.9607	-241.7250	2751.6223	2373.00787	2657.94621	2455.45984
15	597.7519	327.5957	2949.1898	2604.27559	2622.27578	2499.70236
16	436.2778	150.9273	2622.1014	2261.39606	2208.27197	2046.21654
17	-443.5187	-479.5286	2102.7939	1617.86850	1336.63724	1254.87874
18	-1454.8827	-1554.5213	98.8846	-462.03268	49.62165	-193.46047
19	-1650.6383	-1449.9480	-226.3365	-866.44953	264.66687	62.39201
20	-1767.0916	-1595.7472	-1115.2556	-1279.01102	-360.83530	-377.67024
21	-1950.8306	-1841.0921	-1475.1521	-1515.30630	-312.02474	-452.48031
22	-2156.3248	-2081.4094	-1650.2009	-1724.45276	-146.53868	-375.65921
23	-2402.2584	-2339.8260	-1766.3324	-1897.40079	-26.47953	-266.36008
24	-2245.5190	-2275.4732	-859.3800	-1175.44331	106.01403	-45.56980

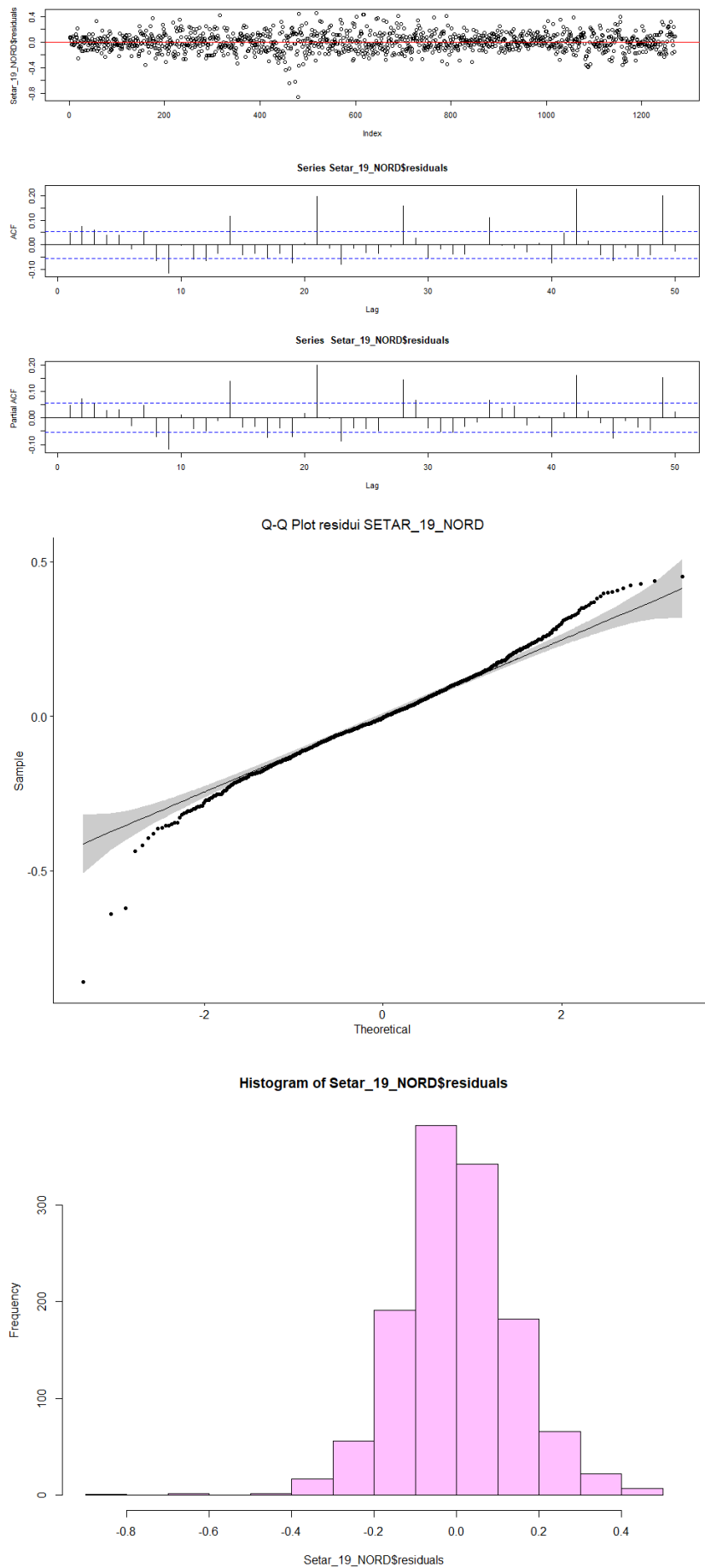


Figura 4.6: ACF , PACF , Q-Q Plot e istogramma dei residui del modello SETAR(7,1) (Nord ora 19)

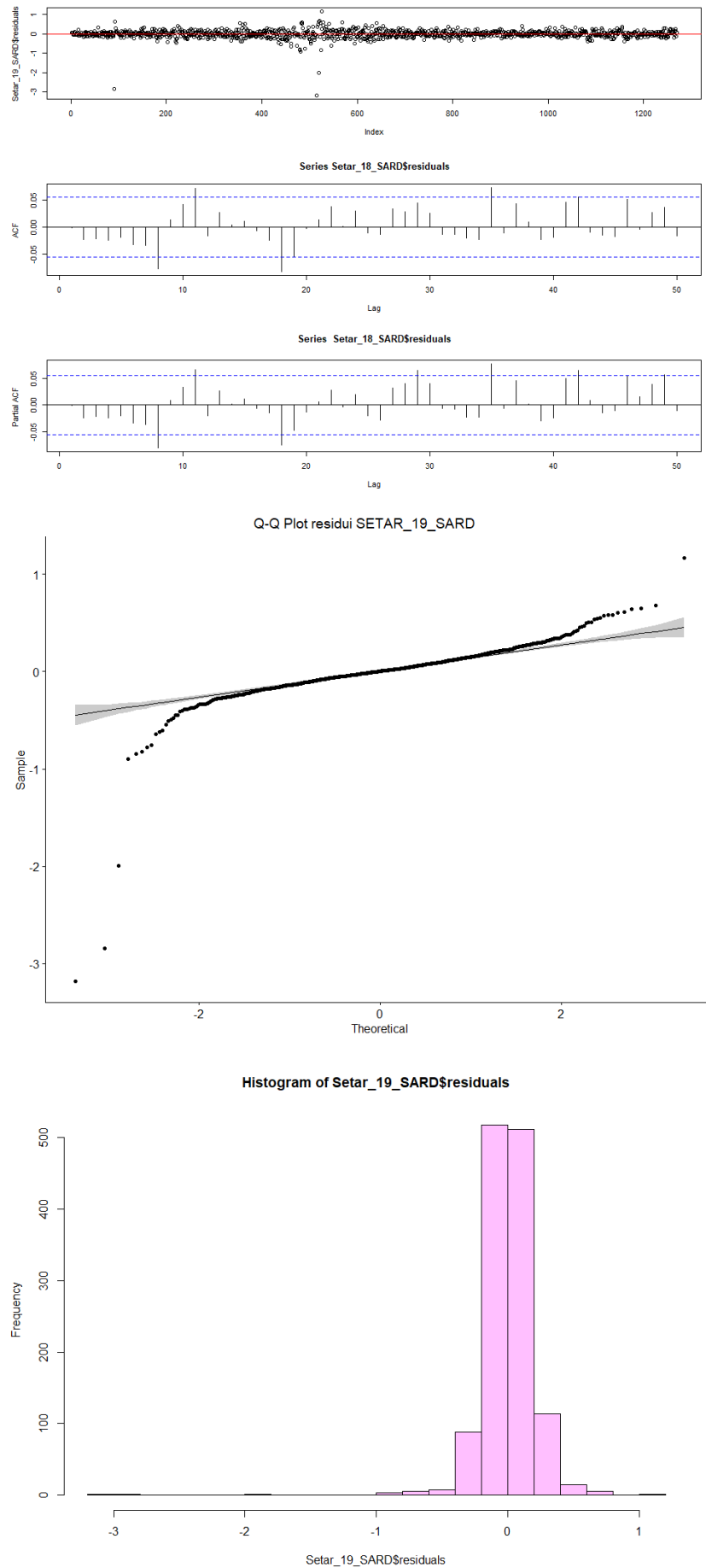


Figura 4.7: ACF , PACF , Q-Q Plot e istogramma dei residui del modello SETAR(7,1) (Sardegna ora 19)

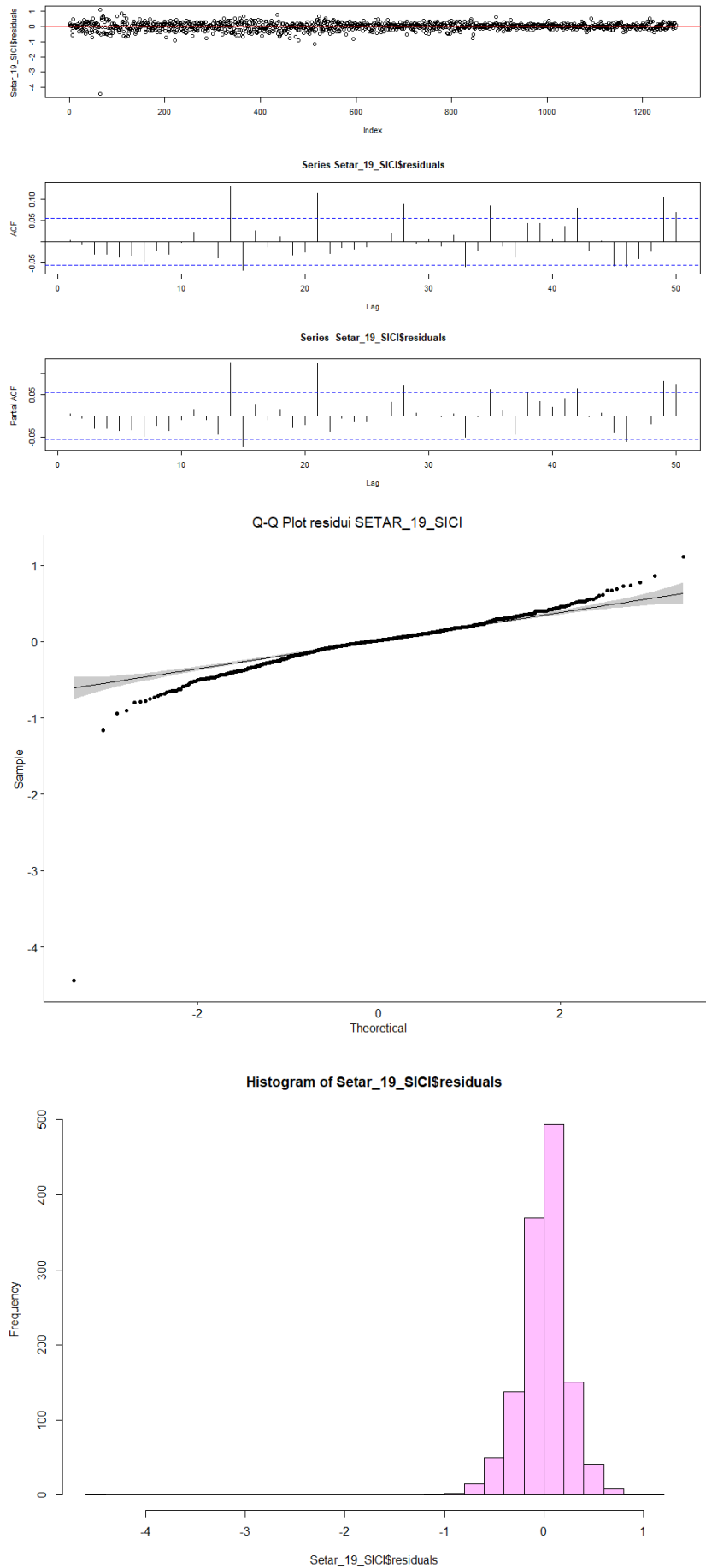


Figura 4.8: ACF , PACF , Q-Q Plot e istogramma dei residui del modello SETAR(7,1) (Sicilia ora 19)

4.3 Stima dei Markov switching model

Vista la natura dei prezzi spot, un modello non lineare a cambiamento di regime dovrebbe essere in grado di cogliere le dinamiche del processo in questione. Tuttavia è inverosimile ipotizzare che il cambiamento di regime sia esclusivamente regolato da una variabile soglia. Per serie storiche come quelle del PUN, che riflettono il consumo e la reperibilità di un bene essenziale, è importante valutare l'influenza di variabili esogene riferite a fattori economico-sociali. In questo senso vanno presi in considerazione due eventi le cui ripercussioni sono evidenti nel sistema energetico italiano: la pandemia causata dal Covid-19 e la guerra tra Russia e Ucraina. La ripresa della domanda di energia, dopo il trend decrescente dovuto alla pandemia, è stata troppo rapida per permettere all'offerta di adattarsi. Lo stato di carenza nella fornitura di fonti energetiche ha portato alla crisi energetica che persiste tutt'ora e che è stata aggravata dal conflitto Russo-Ucraino. Come abbiamo visto nel capitolo 2 (figura 2.5) la fonte energetica più utilizzata per la produzione di elettricità in Italia è il gas, con un 49,2% nel 2021. Secondo i dati forniti dal MITE (Ministero Transizione Ecologica) a Giugno 2022 la produzione nazionale di gas ammontava al 6,5% del consumo interno lordo. Questo dato indica che l'Italia è fortemente dipendente dall'importazione di fonti energetiche dall'estero, in particolare dalla Russia, da cui nel 2021 proveniva il 39,8% delle importazioni estere totali di gas. Dalla figura 4.12 si evince una relazione tra PUN e prezzi del gas le cui serie storiche presentano un andamento simile. Inoltre si nota un cambiamento nel trend a partire da Settembre/Ottobre 2021, il picco più alto viene raggiunto a Marzo in concomitanza dell'inizio del conflitto.

Alla luce di quanto detto fin'ora inseriremo nell'applicazione dei modelli MSM delle variabili esogene ipotizzando che abbiano un'influenza significativa sul PUN. Oltre al prezzo del gas, una dummy per la crisi energetica a partire da Settembre 2021 e una dummy per il conflitto Russia-Ucraina a partire da Marzo 2022, il modello stimato includerà anche il prezzo del carbone poiché in Sardegna sono presenti due centrali termoelettriche alimentate a carbone (Fiumesanto e Portoscuso). Inseriremo inoltre 6 dummy che indicheranno qual'è stata la tecnologia a fissare il prezzo marginale in una determinata ora. Le tecnologie marginali prese in considerazione sono: Carbone, Ciclo combinato a metano, Impianto idrico a pompaggio, Idrico fluente, Estero e MC che indica un prezzo formato nel Market Coupling in cui l'impianto al margine non è reso noto poiché nei mercati stranieri non c'è lo stesso livello di

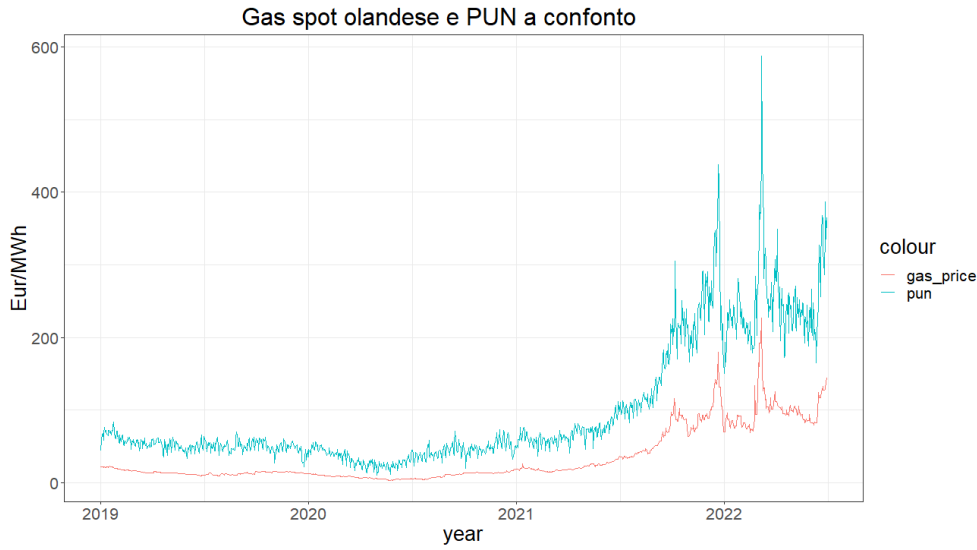


Figura 4.9: Serie storica del PUN e dei prezzi del gas olandese dal 1/1/2019 al 30/6/2022

trasparenza che caratterizza il mercato italiano. Tutte le altre tecnologie marginali (metano, olio combustibile, oliocarbene, oliometano, turbogas, idrico modulazione, fonti rinnovabili) saranno incluse in una variabile dummy "Altro" che rappresenterà la categoria di riferimento, quindi non sarà inserita implicitamente nel modello. Per evitare problemi di multicollinearità e ridondanza nelle informazioni si aggiungono sempre $k-1$ dummy in un modello lineare dove k indica il numero di modalità della variabile qualitativa. Il modello applicato risulterà quindi essere

$$X_t = \begin{cases} \sum_{i=1}^7 \Phi_{s_t,i} x_{t-i} + \beta_{s_t,0} + \sum_{i=1}^2 \beta_{s_t,i} z_i + \sum_{i=3}^{10} \gamma_{s_t,i} z_i + \sigma_{s_t} \varepsilon_t & \text{se } S_t = 1 \\ \sum_{i=1}^7 \Phi_{s_t,i} x_{t-i} + \beta_{s_t,0} + \sum_{i=1}^2 \beta_{s_t,i} z_i + \sum_{i=3}^{10} \gamma_{s_t,i} z_i + \sigma_{s_t} \varepsilon_t & \text{se } S_t = 2 \end{cases} \quad (4.2)$$

oppure in forma estesa

$$X_t = \sum_{i=1}^7 \Phi_{s_t,i} x_{t-i} + \beta_{s_t,0} + \beta_{s_t,1} z_1 + \beta_{s_t,2} z_2 + \gamma_{s_t,3} z_3 + \gamma_{s_t,4} z_4 + \gamma_{s_t,5} z_5 + \gamma_{s_t,6} z_6 + \gamma_{s_t,7} z_7 + \gamma_{s_t,8} z_8 + \gamma_{s_t,9} z_9 + \gamma_{s_t,10} z_{10} + \sigma_{s_t} \varepsilon_t \quad \text{con } S_t = 1, 2 \quad (4.3)$$

dove $\sum_{i=1}^7 \Phi_{s_t,i} x_{t-i}$ rappresenta la parte autoregressiva AR(7), z_1 e z_2 rappresentano rispetti-

vamente il prezzo del gas e il prezzo del carbone z_3, z_4, z_5, z_6, z_7 e z_8 sono le dummy relative alle tecnologie marginali, z_9 è la dummy per la crisi energetica e z_{10} è la dummy per la guerra.

In tutti i modelli il prezzo del gas risulta essere correlato positivamente con il PUN e altamente significativo, mentre per i prezzi del carbone si osserva una significatività meno costante nelle ore della giornata. Anche le dummy riferite alla guerra e alla crisi energetica sono significative solo per alcune ore della giornata e solo per uno dei due regimi del modello, le stesse considerazioni valgono per le dummy riferite alle tecnologie marginali. Gli R quadro indicano in tutti i modelli un adattamento soddisfacente tuttavia dall'analisi dei residui risulta presente dell'autocorrelazione ed una distribuzione spesso asimmetrica. Tutte le matrici di transizione indicano probabilità di passare da un regime all'altro $P(S_t = 2 | S_{t-1} = 1) = \eta_1$, $P(S_t = 1 | S_{t-1} = 2) = \eta_2$ molto basse.

```
> summary(lm22_SARD)
```

Call:

```
lm(formula = PUN22_SARD ~ GAS_PRICE + COAL_PRICE + MC22_SARD +
    Ccgt22_SARD + I.Pompaggio22_SARD + I.Fluyente22_SARD + Estero22_SARD +
    Carbone22_SARD + Crisi_eng + WAR)
```

Residuals:

Min	1Q	Median	3Q	Max
-135.186	-7.549	-1.050	6.549	125.236

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	31.98634	1.95506	16.361	< 2e-16 ***
GAS_PRICE	1.94165	0.04446	43.673	< 2e-16 ***
COAL_PRICE	0.10827	0.02271	4.769	2.07e-06 ***
MC22_SARD	-8.02964	1.99383	-4.027	5.98e-05 ***
Ccgt22_SARD	-11.11790	1.57218	-7.072	2.53e-12 ***
I.Pompaggio22_SARD	0.05267	3.97185	0.013	0.989421
I.Fluyente22_SARD	-7.46176	2.39069	-3.121	0.001842 **
Estero22_SARD	-12.79508	3.71963	-3.440	0.000601 ***
Carbone22_SARD	-3.50139	1.81922	-1.925	0.054494 .
Crisi_eng	13.33023	3.12164	4.270	2.10e-05 ***
WAR	17.31920	3.77775	4.585	5.00e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.89 on 1266 degrees of freedom

Multiple R-squared: 0.9609, Adjusted R-squared: 0.9605

F-statistic: 3108 on 10 and 1266 DF, p-value: < 2.2e-16

4.3. STIMA DEI MARKOV SWITCHING MODEL

```
> summary(msm22_SARD)
Markov Switching Model

Call: msmFit(object = lm22_SARD, k = 2, sw = c(T, T, T, T, T, T, T,
      T, T, T, T, T, T, T, T, T, T, T, T), p = 7)

      AIC      BIC    logLik
9308.158 9750.725 -4618.079
```

Coefficients:

```
Regime 1
-----
              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)(S)      9.2655     1.2945   7.1576 8.211e-13 ***
GAS_PRICE(S)        0.3615     0.0924   3.9123 9.142e-05 ***
COAL_PRICE(S)        0.0638     0.0412   1.5485  0.12150
MC22_SARD(S)        -7.4308     0.8866  -8.3812 < 2.2e-16 ***
Ccgt22_SARD(S)      -6.8061     0.6568 -10.3625 < 2.2e-16 ***
I.Pompaggio22_SARD(S) -3.4164     1.5659  -2.1817  0.02913 *
I.Fluente22_SARD(S) -5.2289     0.9791  -5.3405 9.269e-08 ***
Estero22_SARD(S)    -8.5345     1.6074  -5.3095 1.099e-07 ***
Carbone22_SARD(S)  -3.0354     0.7053  -4.3037 1.680e-05 ***
Crisi_enrg(S)       2.5223     1.9075   1.3223  0.18607
WAR(S)              -8.1160    21.9006  -0.3706  0.71094
PUN22_SARD_1(S)     0.4178     0.0301  13.8804 < 2.2e-16 ***
PUN22_SARD_2(S)     0.0688     0.0325   2.1169  0.03427 *
PUN22_SARD_3(S)     0.0568     0.0321   1.7695  0.07681 .
PUN22_SARD_4(S)     0.0176     0.0318   0.5535  0.57992
PUN22_SARD_5(S)     0.0405     0.0328   1.2348  0.21690
PUN22_SARD_6(S)     0.0466     0.0321   1.4517  0.14659
PUN22_SARD_7(S)     0.1246     0.0301   4.1395 3.481e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 6.546455
 Multiple R-squared: 0.942

```
Standardized Residuals:
      Min          Q1          Med          Q3          Max
-2.597373e+01 -2.761198e+00 -1.304579e-16  2.276489e+00  2.683955e+01
```

```
Regime 2
-----
              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)(S)      32.5334     10.1577   3.2028 0.0013610 **
GAS_PRICE(S)        1.0589     0.0982  10.7831 < 2.2e-16 ***
```

52 CAPITOLO 4. STIMA DI MODELLI E APPLICAZIONE AI PREZZI ELETTRICI

```

COAL_PRICE(S)          -0.0158      0.0436 -0.3624 0.7170531
MC22_SARD(S)           5.7989      5.0628  1.1454 0.2520435
Ccg22_SARD(S)          6.1040      4.6488  1.3130 0.1891829
I.Pompaggio22_SARD(S) 45.5174     13.7691  3.3058 0.0009471 ***
I.Flunte22_SARD(S)    16.0997     7.5632  2.1287 0.0332791 *
Estero22_SARD(S)      2.8429      8.0494  0.3532 0.7239385
Carbone22_SARD(S)     5.9407      8.5520  0.6947 0.4872433
Crisi_enrg(S)         -0.4243     1.8632 -0.2277 0.8198795
WAR(S)                 20.2743     7.7887  2.6030 0.0092412 **
PUN22_SARD_1(S)        0.3578      0.0542  6.6015 4.07e-11 ***
PUN22_SARD_2(S)        0.1737      0.0545  3.1872 0.0014366 **
PUN22_SARD_3(S)       -0.0554     0.0540 -1.0259 0.3049387
PUN22_SARD_4(S)       -0.0121     0.0167 -0.7246 0.4686975
PUN22_SARD_5(S)        0.0755      0.0460  1.6413 0.1007352
PUN22_SARD_6(S)        0.0168      0.0383  0.4386 0.6609514
PUN22_SARD_7(S)       -0.1195     0.0414 -2.8865 0.0038955 **

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 26.23453

Multiple R-squared: 0.8517

Standardized Residuals:

	Min	Q1	Med	Q3	Max
	-89.48539410	-0.09742443	-0.04434588	-0.03339940	107.06746060

Transition probabilities:

	Regime 1	Regime 2
Regime 1	0.998655688	0.007806558
Regime 2	0.001344312	0.992193442

Considerando il modello lineare adattato all'ora 22 della Sardegna tutte le variabili risultano altamente significative tranne le tecnologie marginali impianto idrico a pompaggio, che non è significativa, e carbone, che ha una significatività bassa con un p value di 0.054494. Crisi energetica, guerra e prezzo del gas presentano una relazione positiva con il PUN, con coefficienti di regressione rispettivamente pari a 13.3, 17.3 e 1.9. Per quanto riguarda il modello markow switching la variabile "war" risulta significativa solo nel secondo regime, il prezzo del gas è significativo in entrambi i regimi mentre il prezzo del carbone non è significativo in nessuno dei due. L'R quadro indica un adattamento buono, dall'analisi dei residui le funzioni di acf e pacf non indicano presenza di forte autocorrelazione e il Q-Q Plot evidenzia una di-

istribuzione non eccessivamente asimmetrica ma con code pesanti. L'analisi svolta indica un adattamento abbastanza buono ma migliorabile. Probabilmente un modello più complesso ottenuto con l'aggiunta di un maggiore numero di regimi potrebbe cogliere in modo migliore l'andamento della serie.

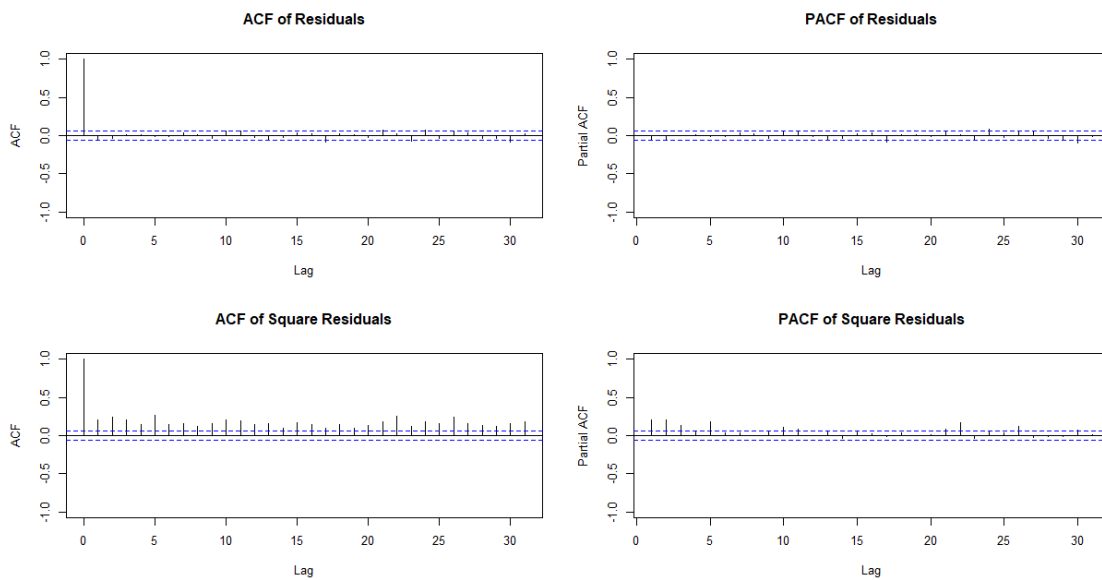


Figura 4.10: ACF e PACF dei residui del modello MSM applicato al PUN dell'ora 22 della Sardegna

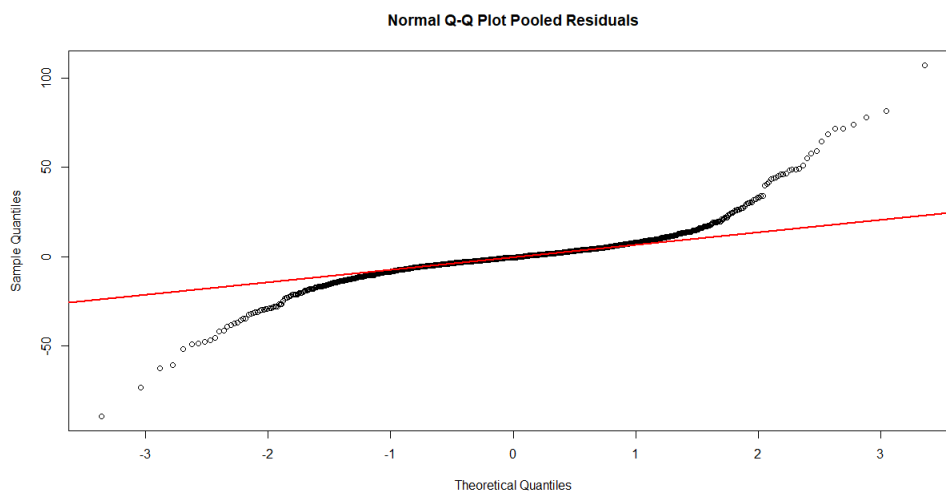


Figura 4.11: Q-Q Plot della distribuzione dei residui del modello MSM applicato al PUN dell'ora 22 della Sardegna

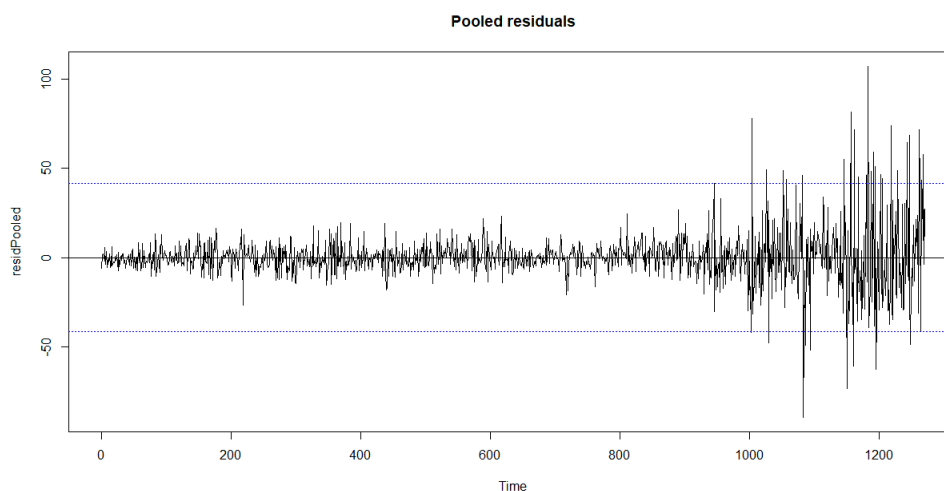


Figura 4.12: Grafico dei residui del modello MSM applicato al PUN dell'ora 22 della Sardegna

4.4 Stima dei modelli STAR

L'ultima classe di modelli che applichiamo alla serie storica del PUN è quella dei STAR la cui specificazione è la seguente :

$$\begin{aligned}
 X_t = & (\Phi_{0,1} + \Phi_{1,1}x_{t-1} + \dots + \Phi_{7,1}x_{t-7} + [1 - G(x_{t-1}|\lambda, c)] + \\
 & + (\Phi_{0,2} + \Phi_{1,2}x_{t-1} + \dots + \Phi_{7,2}x_{t-7} + [G(x_{t-1}|\lambda, c)] + a_t
 \end{aligned}
 \tag{4.4}$$

Le performance dei modelli STAR risultano simili a quelle dei modelli SETAR. Infatti, pur essendo presente autocorrelazione nei residui, il test di non linearità che confronta l'adattamento del modello STAR rispetto a quello del modello AR corrispondente rifiuta in tutte le 72 serie l'ipotesi nulla di linearità con p-value molto piccoli. Di seguito i grafici dell'analisi dei residui per il modello adattato all'ora 3 indicano distribuzioni abbastanza simmetriche ma con code pesanti (figure 4.13 4.14 4.15).

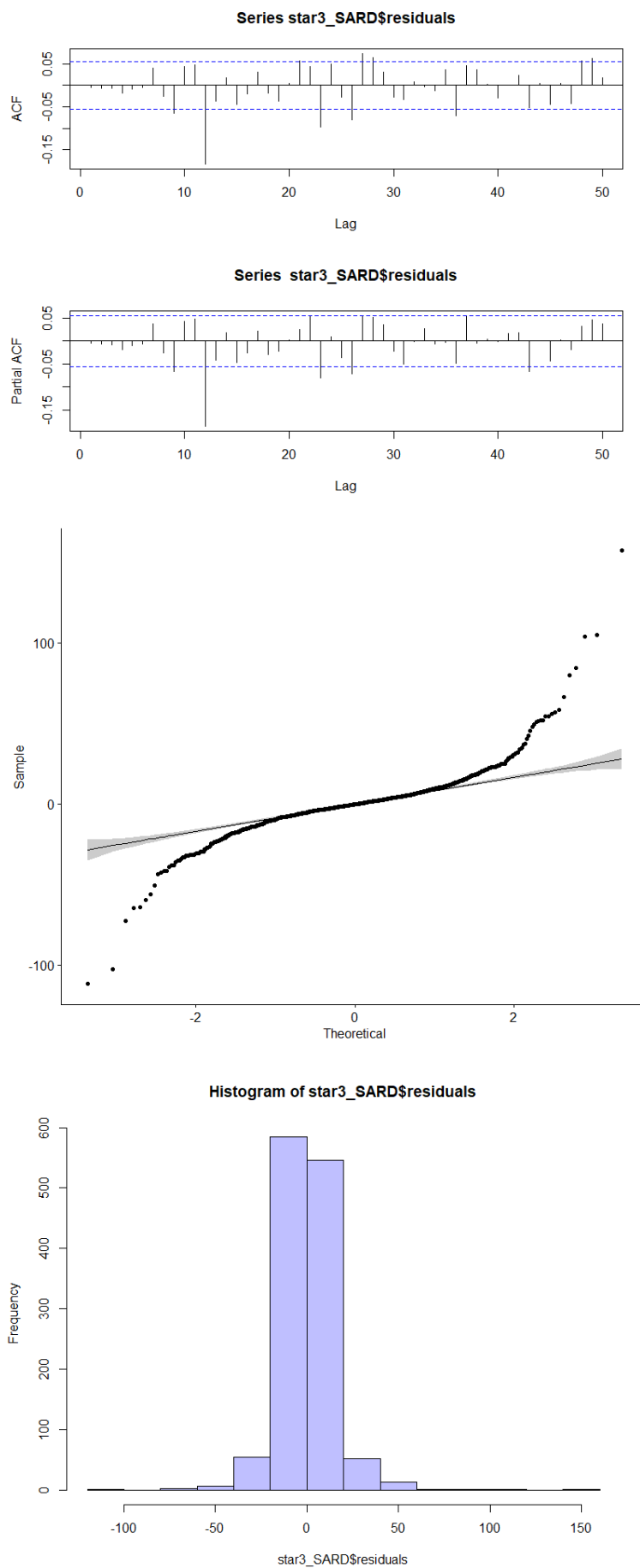


Figura 4.14: ACF , PACF , Q-Q Plot e istogramma dei residui del modello STAR (Sardegna ora 3)

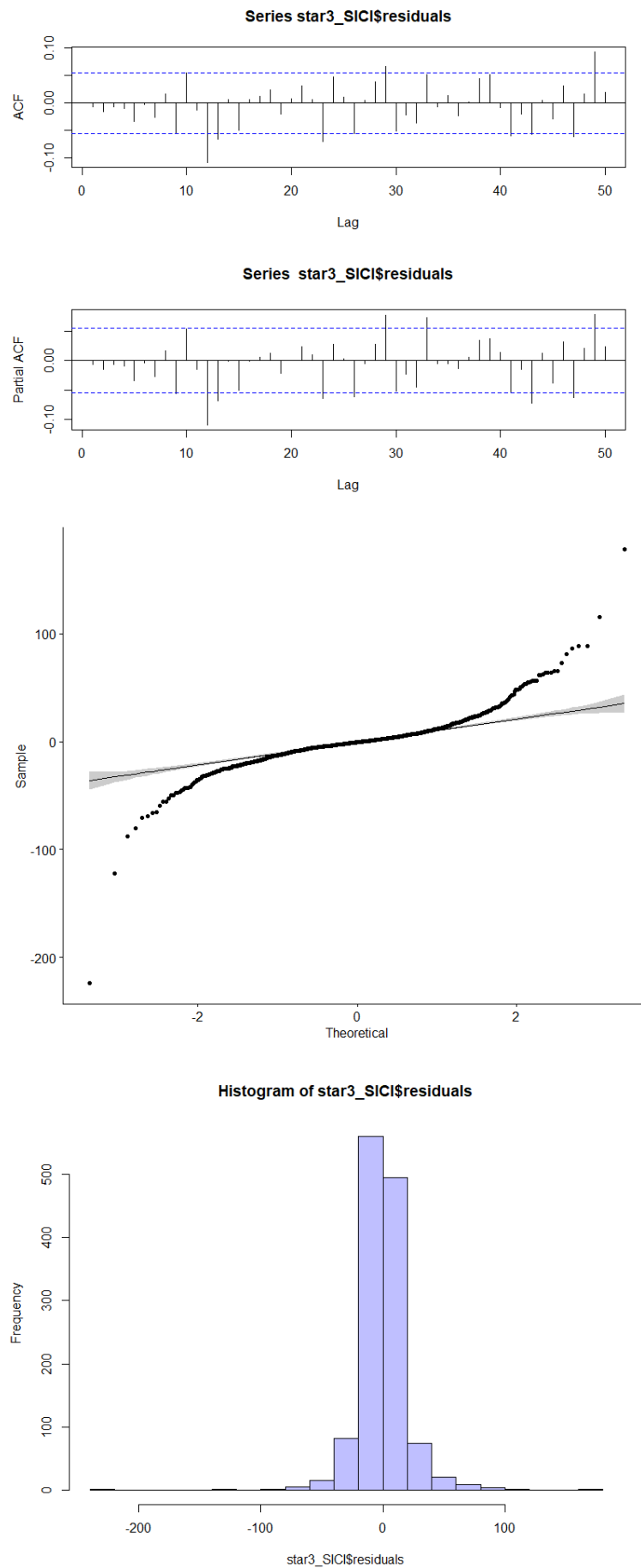


Figura 4.15: ACF , PACF , Q-Q Plot e istogramma dei residui del modello STAR (Sicilia ora 3)

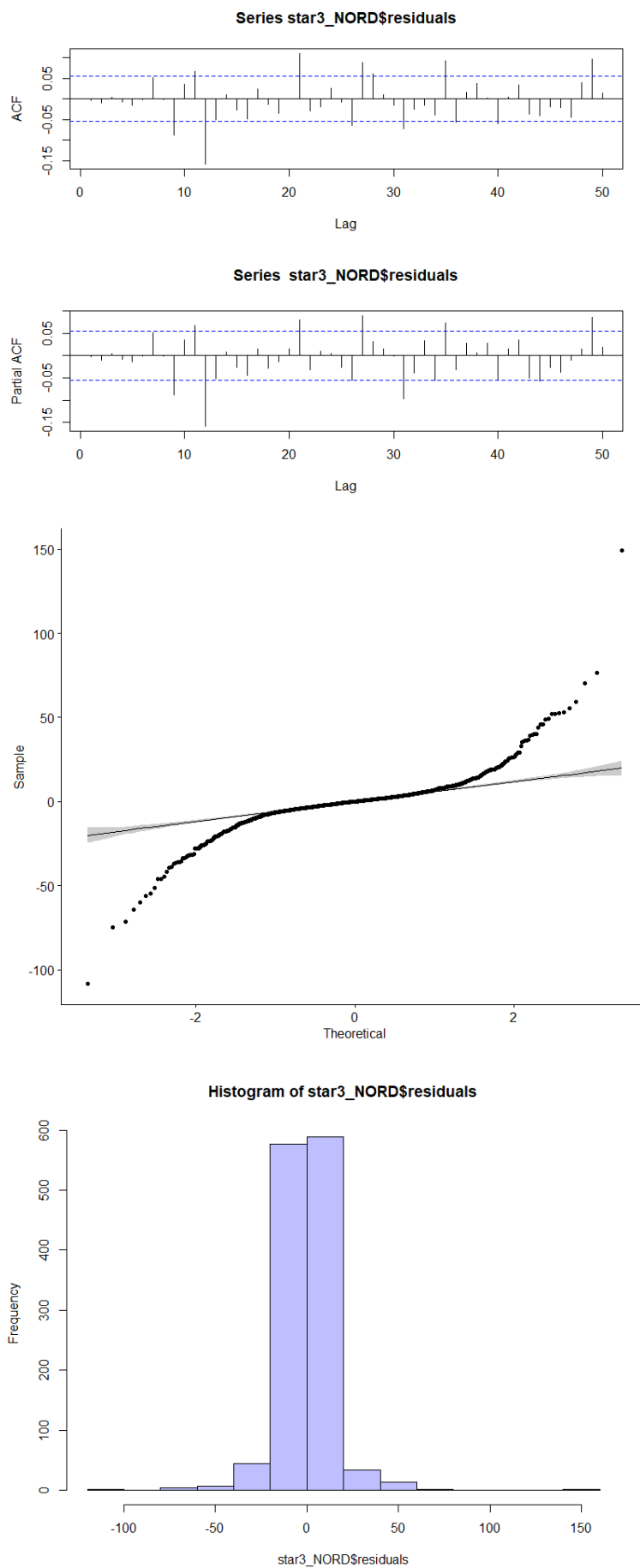


Figura 4.13: ACF , PACF , Q-Q Plot e istogramma dei residui del modello STAR (Nord ora 3)

Capitolo 5

Conclusione

Il presente studio si è posto l'obiettivo di valutare l'adattamento di diverse tipologie di modelli non lineari, tutte facenti parte della classe dei modelli a soglia, sulla serie storica del PUN di tre aree geografiche differenti: Nord, Sardegna e Sicilia. Ai fini dell'applicazione dei modelli non lineari sulle serie oggetto di interesse, si è proceduto all'individuazione di un $ARMA(7,1)$, che ha costituito la componente autoregressiva su cui implementare la costruzione di tre tipi di modelli: SETAR, MSM e STAR, tutti caratterizzati da una struttura a due regimi.

Si è scelto di suddividere la serie temporale in 24 sottoserie, una per ogni ora, ed applicare due diversi test per la verifica della non linearità del processo generatore dei dati. Per valutare la bontà di adattamento si è fatto ricorso alla diagnostica dei residui, ed in alcuni casi, al criterio di Akaike. I risultati derivanti dall'applicazione del modello SETAR $(7,1)$ sul logaritmo delle serie orarie del PUN indicano sicuramente una superiorità di quest'ultimo rispetto al modello lineare $ARIMA(7,1,7)$ dedotta attraverso il confronto dell'AIC dei due modelli, tenendo conto, quindi, non solo della bontà di adattamento ma anche della parsimonia. Tuttavia, l'analisi dei residui evidenzia una presenza di autocorrelazione ai ritardi stagionali ed una distribuzione più o meno asimmetrica a seconda delle ore. Sarebbe interessante provare ad adattare per ogni ora del giorno un modello costruito ad hoc in modo da riuscire a catturare nella misura maggiore possibile l'autocorrelazione presente nei residui. Nel caso dei modelli MSM è stato possibile inserire dei regressori esogeni come prezzo del gas, prezzo del carbone, dummy relative alle tecnologie marginali e dummy legate a fattori politici, economici e sociali come la crisi energetica post pandemia ed il conflitto Russia-Ucraina. Come atteso, il coefficiente

di regressione del prezzo del gas è positivo e risulta significativo in tutti i modelli, mentre per le dummy ed il prezzo del carbone si osservano comportamenti diversi a seconda delle ore e delle aree geografiche. L' R quadro risulta elevato in tutti i modelli ed in entrambi i regimi, tuttavia dall'analisi dei residui si evince una difficoltà del modello a catturare l'andamento dei prezzi nel secondo regime caratterizzato da una volatilità più alta e diversi spikes. Sarebbe opportuno provare ad implementare dei modelli a più di due regimi per verificare un ipotetico miglioramento nell'asimmetria dei residui. Le performance dei modelli STAR risultano migliori rispetto a quelle dei modelli lineari autoregressivi corrispondenti poichè in tutti i modelli stimati il test che confronta modello STAR e modello lineare rifiuta l'ipotesi nulla con p-value molto bassi a favore del modello STAR. Anche in questo caso, l'analisi dei residui porta alla luce la presenza di autocorrelazione, che lascia spazio alla possibilità di ulteriori miglioramenti, ad esempio inserendo regressori esogeni o provando ad aumentare il numero di regimi, tutto ciò cercando di non andare a discapito dell'interpretabilità dei modelli, rendendoli troppo complessi.

Bibliografia

- ARERA. (2018). Aggiornamento Trimestrale delle Condizioni Economiche per i Servizi di Tutela dell'Energia Elettrica e del Gas Naturale per il IV trimestre 2018. <https://www.arera.it/allegati/docs/18/491-18all.pdf>
- ARERA. (2022). Prezzi Finali Dell'Energia Elettrica Per i Consumatori Industriali - UE a Area Euro. <https://www.arera.it/it/dati/eepcfr2.htm>
- ARERA. (N/A). Andamento del prezzo dell'energia elettrica per il consumatore domestico tipo in maggior tutela. <https://www.arera.it/it/dati/eep35.htm>
- Armondi, S. (2021). Ripensare le politiche energetiche tra emergenze e transizioni. Una prospettiva geografica.
- Benth, F. E. & Koekebakker, S. (2005). Stochastic Modeling Of Financial Electricity Contracts. *Pure Mathematics*, 24.
- Chan, K. S. & Tsay, R. S. (1998). Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model. *Biometrika*, 85(2), 413–426.
- Chen, H. & Chong, J., Terence Tai-Leung Bai. (2012). Theory and Applications of TAR Model with Two Threshold Variables. *Econometric Reviews*, 31(2), 142–170.
- Corduas, M. (1994). Nonlinearity tests in time series analysis. *Journal of the Italian Statistical Society*, 3, 291–313.

- De Vincenzo, D. (2022). NEXTGENERATIONEU Tra Pandemia, Guerra e Transizione Energetica. *Documenti geografici*, (1), 23–36.
- Escribano, A., Ignacio Peña, J. & Villaplana, P. (2011). Modelling Electricity Prices: International Evidence. *Oxford Bulletin of Economics and Statistics*, 73(5), 622–650.
- Geman, H. & Roncoroni, A. (2006). Understanding the Fine Structure of Electricity Prices. *The Journal of Business*, 79(3), 1225–1261.
- Ghazi, S. & Panagiotis, M. (2004). Size and Power of the RESET Test as Applied to Systems of Equations: A Bootstrap Approach. *Journal of Modern Applied Statistical Methods*, 3(2), 370–385.
- Gianfreda, A. & Grossi, L. (2012). Forecasting Italian electricity zonal prices with exogenous variables. *Energy Economics*, 34(6), 2228–2239.
- GME. (N/A). GME - Gestore Mercati Elettrici. <https://www.mercatoelettrico.org/It/default.aspx>
- Grossi, L. & Nan, F. (2019). Robust forecasting of electricity prices: Simulations, models and the impact of renewable sources. *Technological Forecasting and Social Change*, 141(100), 305–318.
- Hansen, B. (1999). Testing for Linearity. *Journal of Economic Surveys*, 13(5), 551–576.
- Jianqing, F. & Qiwei, Y. (2003). Nonlinear Time Series: Nonparametric and Parametric Methods. <http://lib.uib.kz/wp-content/uploads/2019/04/Nonlinear-Time-Series..-Nonparametric-and-Parametric-Methods-Fan-J-2003.pdf>
- Mayer, K. & Trück, S. (2018). Understanding the Fine Structure of Electricity Prices. *Journal of Commodity Markets*, 9, 77–100.
- McLeod, A. I. & Li, W. K. (1983). Diagnostic Checking Arma Time Series Models Using Squared-Residual Autocorrelations. *Journal Of Time Series Analysis*, 4(4), 269–272.

- MISE. (N/A). Importazioni Gas Naturale - Analisi e statistiche energetiche e minerarie. <https://dgsaie.mise.gov.it/importazioni-gas-naturale>
- Natali, G. (2022). La crisi dei prezzi dell'energia, tra geopolitica ed emergenza climatica. https://www.treccani.it/magazine/atlante/geopolitica/La_crisi_prezzi_energia.html
- Neville, D. & Petruccioli, D. J. (1986). Detecting Non-Linearity in Time Series. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 35(2), 271–280.
- Potter, S. M. (1999). Nonlinear Time Series Modelling: An Introduction. https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr87.pdf
- Priestley, M. B. (1978). Non-Linear Models in Time Series Analysis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 27(3/4), 159–176.
- Prontera, A. (2020). Italy, Russia and the Great Reconfiguration in East-West Energy Relations. *Europe-Asia Studies*.
- Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2), 350–371.
- Renò, R. (2006). I prezzi dell'energia elettrica: proprietà e modelli. https://www.academia.edu/2795774/I_prezzi_dell_energia_elettrica_propriet%C3%A0_e_modelli
- Sensfuß, F., Ragwitz, M. & Genoese, M. (2008). The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. *Energy Policy*, 36, 3086–3094.
- Terna. (N/A). Zone Di Mercato. <https://www.terna.it/it/sistema-elettrico/mercato-elettrico/zone-mercato>
- Tong, H. & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(3), 245–292.

Tsay, R. S. (1986). Nonlinearity Tests for Time Series. *Biometrika*, 73(2), 461–466.

Tsay, R. S. (1989). Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association*, 84(405), 231–240.

Tsay, R. S. & Chen, R. (2018). Nonlinear Time Series Analysis.

Weron, R. (2006). Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach.

Weron, R. (2009). Heavy-tails and regime-switching in electricity prices. *Mathematical Methods of Operations Research*, 69, 457–473.