

Serie storiche consumi energetici Stati Uniti:

Energia Fossile vs Energia Rinnovabile

Aurora Musitelli

Matricola: 856741

Introduzione alle Serie Storiche - a.a. 2024/2025



Argomenti trattati

1. Introduzione	2
2. Analisi esplorativa delle serie storiche	4
3. Analisi di detrendizzazione, destagionalizzazione e decomposizione.....	17
4. Analisi di regressione e analisi dei residui.....	23
5. Stima e comparazione di modelli ARIMA, SARIMA e regARIMA.....	28
6. Conclusioni	33

1. Introduzione

L'energia rappresenta il centro di molte delle principali sfide ambientali e geopolitiche contemporanee. La disponibilità di abbondanti fonti di energia sostenibili a prezzi accessibili rimane fondamentale per sostenere l'attuale stile di vita, ma la domanda di energia esistente è soddisfatta prevalentemente da fonti non rinnovabili che generano notevoli effetti di inquinamento atmosferico e cambiamenti climatici. La concentrazione di combustibili fossili come petrolio e gas naturale in regioni geopolitiche rischiose aumenta i pericoli legati alla dipendenza energetica. La soluzione a questi problemi dipende dalla promozione di fonti di energia sostenibili e dal miglioramento dell'efficienza complessiva attraverso progressi tecnologici, come cambiamenti comportamentali e implementazione di strategie a lungo termine. La comprensione delle serie storiche di utilizzo dell'energia rinnovabile e fossile restituisce un quadro generale che porta verso un futuro energetico più consapevole. Il presente report utilizza i dati della [U.S. Energy Information Administration](#) (EIA), che coprono il periodo da gennaio 1973 a giugno 2024 e forniscono informazioni mensili sul consumo energetico statunitense, con una suddivisione per settore economico e tipo di fonte energetica. L'analisi si è concentrata su tre quesiti principali:

1. Come sono cambiati i consumi energetici derivanti da fonti fossili e rinnovabili negli Stati Uniti dal 1973 al 2024? Quali sono stati gli impatti di tali cambiamenti?
2. In che modo eventi storici come l'embargo petrolifero del 1973, la crisi finanziaria del 2008 o l'introduzione di politiche climatiche nei primi anni 2000 hanno influenzato il consumo di energia elettrica prodotta da fonti fossili e rinnovabili?
3. Qual è la relazione a lungo termine tra la quota di energia elettrica prodotta da fonti fossili e quella da fonti rinnovabili nei settori industriale, residenziale ed energia elettrica?

Attraverso queste analisi si è cercato di confrontare le fonti energetiche rinnovabili e fossili per valutare come l'adozione di risorse rinnovabili possa progressivamente sostituire l'uso delle fonti tradizionali. Tuttavia, si evidenzia anche un aspetto pratico: in molti settori, la transizione completa verso energie rinnovabili non è ancora tecnicamente o economicamente sostenibile. Questo implica la necessità di continuare a fare affidamento, seppur in modo più efficiente, sulle fonti fossili per coprire parte della domanda energetica.

Descrizione delle variabili del dataset analizzato

Sector: I dati considerano 5 settori economici principali, ma le analisi si concentrano su 3 settori: settore energia elettrica (settore più ampio e comprende la produzione, la trasmissione e la distribuzione di energia elettrica), settore industriale (comprende tutte le attività produttive a livello nazionale) e settore residenziale (comprende le abitazioni private e gli edifici pubblici).

Source: Tra le 32 fonti energetiche disponibili, si sono analizzate:

Totale Fonti fossili → *Carbone*: Utilizzato principalmente per produrre energia elettrica in ambito industriale. *Petrolio*: Miscela naturale di idrocarburi liquidi, viene estratto e raffinato per produrre carburanti (benzina, diesel) e materie prime per l'industria chimica. I principali produttori globali includono Stati Uniti, Arabia Saudita e Russia. L'uso del petrolio è una delle principali fonti di emissioni di CO₂ a livello mondiale. *Gas naturale*: Composto principalmente da metano, viene impiegato per riscaldamento, produzione di energia elettrica e come combustibile a minore impatto rispetto a petrolio e carbone.

Totale Fonti rinnovabili → *Energia solare*: Energia generata sfruttando la radiazione solare attraverso pannelli fotovoltaici, che trasformano la luce in elettricità, o collettori solari, che catturano il calore per usi termici. *Energia eolica*: Energia prodotta dal movimento del vento, che aziona le pale delle turbine eoliche, trasformando l'energia cinetica in elettricità mediante un generatore. *Energia idroelettrica*: Energia ottenuta sfruttando il movimento dell'acqua in impianti idroelettrici, che possono essere alimentati da dighe o da corsi d'acqua. *Energia geotermica*: Energia ricavata dal calore naturale della Terra, utilizzata per la produzione di elettricità in impianti geotermici ad alta temperatura. *Energia da biomasse*: Energia ottenuta dalla combustione diretta o dalla trasformazione di materiali organici (legno, rifiuti, colture energetiche) in biocarburanti o biogas per la produzione di elettricità.

Month: Dati mensili a livello nazionale USA che coprono periodo da gennaio 1973 a giugno 2024.

EnerCons_TrillBTU: Quantità fisica di energia consumata, espressa in Trillion BTU, dove trillion segue la scala corta e corrisponde a (10¹²) BTU. In italiano, questa unità dovrebbe essere indicata come bilione di BTU (nelle analisi si fa riferimento alla scala corta). Un BTU (British Thermal Unit) è l'energia necessaria per aumentare di 1°F fahrenheit la temperatura di una libbra d'acqua a pressione atmosferica standard. Il trillion BTU rappresenta quindi una misura significativa di energia consumata su scala nazionale.

2. Analisi esplorativa delle serie storiche

Per definire l'approccio delle analisi, sono state osservate le serie storiche originali attraverso tecniche di analisi esplorativa grafica e analitica. Considerando il lungo periodo temporale (1973-2024), è stata data particolare attenzione agli eventi economici nazionali che potrebbero aver influenzato le serie. Le analisi si sono concentrate sulle serie storiche aggregate del totale delle fonti fossili e del totale delle fonti rinnovabili a livello nazionale negli Stati Uniti. Questa scelta consente una visione sintetica e comparativa dei consumi energetici tra le due categorie, facilitando l'identificazione delle tendenze di lungo periodo.

Totale fonti fossili (petrolio, carbone, gas naturale)

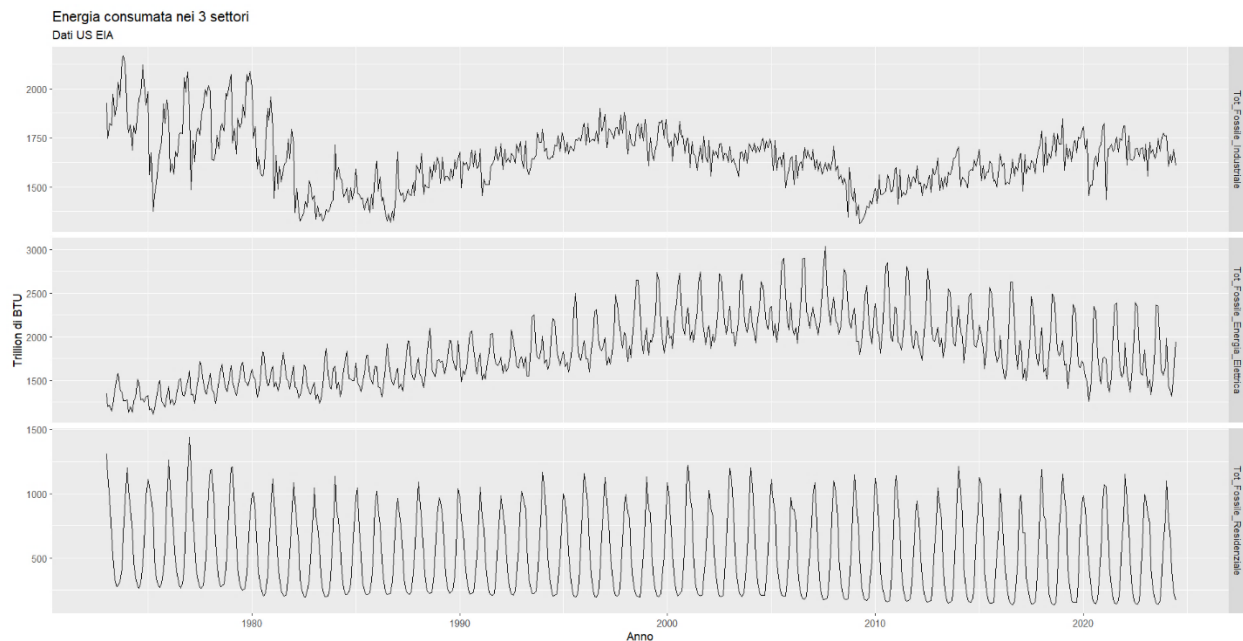


Grafico 1

Industriale: La serie storica mostra un consumo energetico a tendenza decrescente dal 1974 fino alla fine degli anni '80, a causa dell'embargo petrolifero imposto dai paesi arabi membri dell'OPEC contro gli Stati Uniti. Questo evento evidenziò la vulnerabilità economica legata alla dipendenza dalle importazioni di petrolio, spingendo gli Stati Uniti a puntare sull'indipendenza energetica. Dopo il 2000, i consumi si stabilizzano per gas naturale e petrolio, ma l'utilizzo del carbone prosegue ad una tendenza decrescente. Lievi fluttuazioni e una decrescita dovuta alla crisi economica globale del 2008, che ridusse temporaneamente la domanda industriale di energia. Il settore industriale, pur caratterizzato da processi produttivi costanti, mostra variazioni stagionali.

Energia Elettrica: La serie storica mostra una tendenza in crescita fino a poco dopo gli anni 2000, trainata dall'aumento del consumo di carbone e gas naturale. Questa crescita riflette l'aumento della domanda di elettricità dovuto all'espansione tecnologica e alla diffusione degli elettrodomestici. Successivamente, si osserva un declino del carbone e un aumento del gas naturale, in coincidenza con la transizione verso fonti di energia rinnovabile e un miglioramento dell'efficienza energetica. La crisi economica del 2008 e l'introduzione di politiche energetiche più restrittive hanno ridotto i consumi, con un calo significativo dell'uso del carbone e un incremento nell'utilizzo del gas naturale. L'impiego del petrolio in questo settore rimane molto basso, poiché non viene utilizzato come fonte energetica principale. La serie evidenzia una stagionalità marcata a causa della variazione della domanda energetica.

Residenziale: La serie storica mostra una stabilità a lungo termine, con consumi complessivamente stabili nel tempo, ma caratterizzati da forti variazioni stagionali dovute principalmente all'utilizzo del gas naturale. Questa marcata stagionalità, con picchi regolari nei periodi invernali, riflette l'elevata dipendenza dalle fonti fossili per il riscaldamento domestico. Tale caratteristica potrebbe influenzare la stazionarietà della serie, aspetto da verificare successivamente tramite il test di Dickey-Fuller.

Totale fonti rinnovabili (solare, eolico, idroelettrico, geotermico, biomasse)

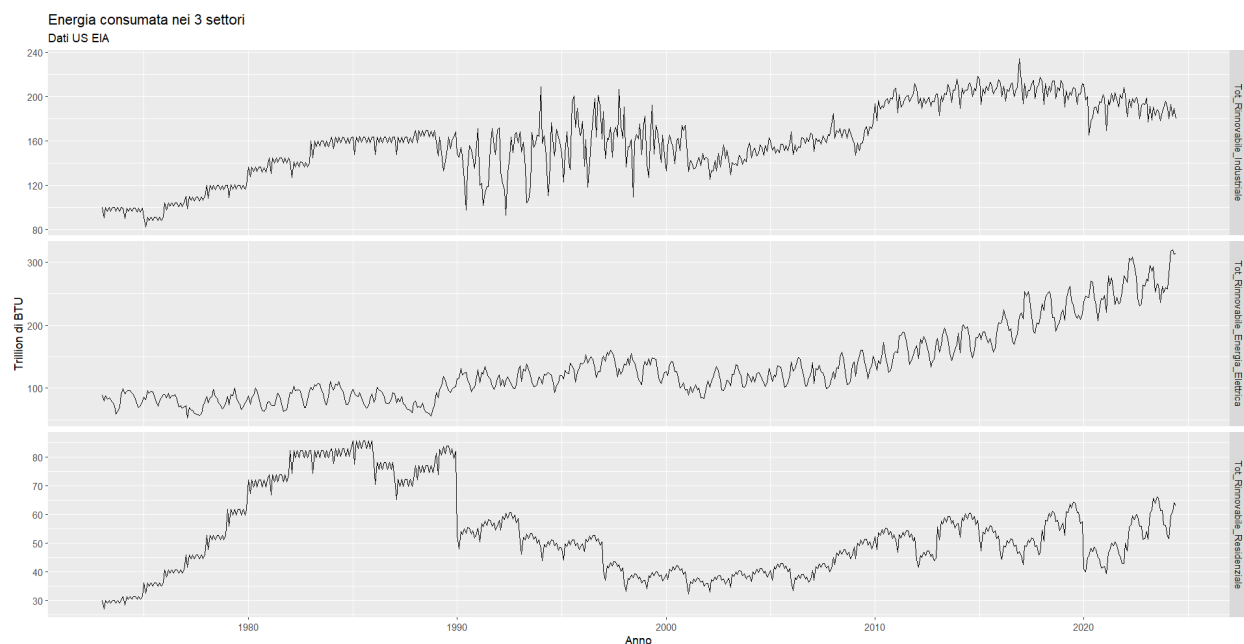


Grafico 2

Le fonti rinnovabili presentano valori nulli per solare, eolico e idroelettrico dal 1973 fino a poco prima degli anni 2000 poiché gli Stati Uniti dipendevano principalmente dalle fonti fossili, che erano più economiche e disponibili. Sebbene l'embargo petrolifero del 1973 abbia aumentato la consapevolezza energetica, le politiche per le rinnovabili erano ancora limitate. Solo a partire dalla fine degli anni '90 l'energia rinnovabile è diventata più rilevante e iniziava ad essere utilizzata.

Industriale: La serie storica mostra una crescita moderata fino agli anni '90, seguita da un andamento con evidenti fluttuazioni stagionali. Eventi storici significativi, come il rallentamento economico del 2008-2009 e l'aumento della consapevolezza sulle energie rinnovabili negli anni 2000, hanno probabilmente influenzato i consumi, che sono aumentati successivamente. Nel settore industriale, la serie è dominata dall'energia derivante dalla biomassa, mentre altre fonti rinnovabili come solare, eolico e geotermico, non sono praticamente utilizzate.

Energia Elettrica: È una serie storica caratterizzata da una forte crescita nel periodo osservato, con un incremento particolarmente marcato dagli anni '90 in avanti, evidenziando una transizione verso fonti rinnovabili per la generazione di energia elettrica (maggior aumento delle fonti eolico e solare). Gli eventi chiave sono stati l'introduzione di normative favorevoli alle energie rinnovabili negli anni 2000 e il calo dei costi tecnologici per l'energia solare ed eolica.

Residenziale: È una serie storica caratterizzata da una crescita dal 1974, con un rallentamento negli anni 2000 e una ripresa negli ultimi anni. La serie presenta fluttuazioni, principalmente legate all'uso delle biomasse nel settore residenziale, poiché questa fonte di energia è stata tradizionalmente più economica e disponibile per il riscaldamento domestico. Fattori storici che hanno influenzato i consumi includono il miglioramento dell'efficienza energetica, le politiche di incentivazione per le energie rinnovabili (ad esempio, i pannelli solari residenziali).

Aspetti generali rispetto alle due fonti energetiche totali: Si osserva che l'energia consumata dalle fonti fossili varia da 0 a 3000 trillion BTU, mentre quella delle fonti rinnovabili varia da 0 a 240 trillion BTU, mostrando una differenza significativa in termini di utilizzo. Questa discrepanza riflette la differenza di scala tra le due fonti energetiche: le fonti fossili, come carbone, petrolio e gas naturale, sono state storicamente dominanti e ampiamente utilizzate, con un'alta densità energetica che consente di soddisfare grandi quantità di domanda a livello industriale e residenziale. Al contrario, le fonti rinnovabili, pur in crescita, sono ancora in fase di espansione e non sono ancora in grado di competere, in termini di volume energetico, con le fonti fossili, che continuano a prevalere a causa dei costi più bassi, della disponibilità e della tecnologia consolidata.

Totale fonti fossili (petrolio, carbone, gas naturale)

Analisi distribuzione empirica dei dati: Dopo aver osservato le serie storiche originali si esplorano quanto i dati sono vicini/lontani dall'ipotesi di normalità, come prima cosa si visualizzano gli istogrammi attraverso la sovrapposizione della densità kernel e della curva gaussiana teorica.

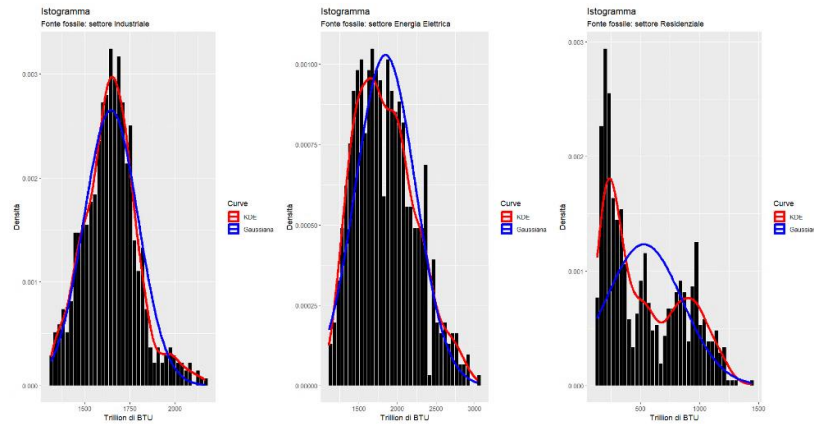


Grafico 3

Industriale: Il primo istogramma si osserva la curva in blu è la densità di una gaussiana con media dei consumi di fonti fossili e varianza di tali consumi; mentre in rosso è la Kernel Density Estimation (KDE), una densità parametrica applicata ai valori della serie storica. Dal grafico si può notare che la distribuzione è leggermente asimmetrica a destra dalla quale si notano la presenza di diversi outlier, si decide di risolvere il problema usando la funzione `tsclean()` di R:

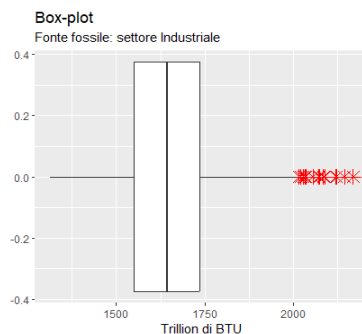


Grafico 4

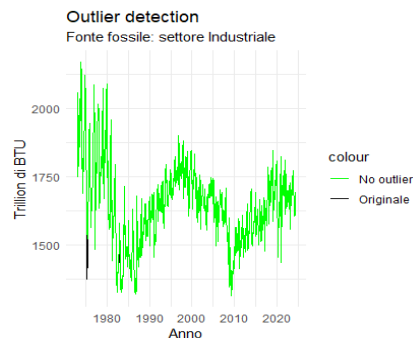


Grafico 5

Energia Elettrica: La distribuzione è leggermente asimmetrica con una coda destra più lunga. La deviazione dalla curva gaussiana indica una non perfetta normalità, ma potrebbe essere accettabile a seconda del livello di tolleranza.

Residenziale: La distribuzione in questo caso è chiaramente multimodale e si discosta significativamente dalla curva gaussiana, indicando una marcata non normalità, questo richiede trasformazioni o approcci alternativi per modellare i dati che vedremo di seguito.

→**Test di normalità:** Attraverso i grafici si possono osservare le distribuzioni dei dati, però esistono diversi test analitici per confermare le teorie osservate dai grafici. Si è applicato il test di Bera-Jarque che valuta l'ipotesi nulla di distribuzione normale dei dati contro l'ipotesi alternativa non siano normali, dai risultati essendo i p-value molto bassi, si rifiuta l'ipotesi nulla di normalità.

Test di Bera-Jarque	X-squared	df	p-value
Industriale	25.546	2	2.837e-06
Energia Elettrica	27.39	2	1.128e-06
Residenziale	59.343	2	1.3e-13

Analisi della persistenza: Una volta intuite le caratteristiche più evidenti delle serie storiche si passa all'analisi della distribuzione della persistenza. Si valuta la correlazione tra un'osservazione e le sue osservazioni passate a diversi ritardi temporali. Questi grafici permettono di ottenere informazioni riguardanti cicli temporali, stagionalità e dare un'idea della stazionarietà della serie.

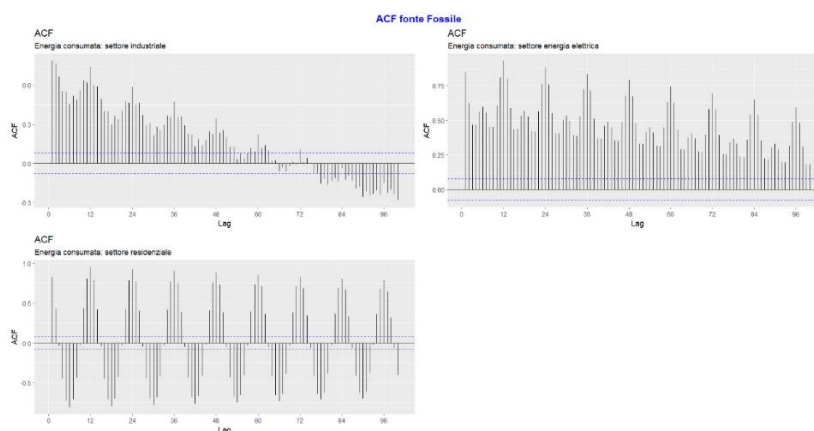


Grafico 6

Industriale: I valori di autocorrelazione iniziano alti (intorno a 0.8) e decrescono lentamente, mantenendosi significativi anche per lag elevati. È presente una forte persistenza nelle dinamiche del consumo energetico del settore industriale, suggerendo che i valori passati influenzano significativamente quelli futuri. Decadimento decrescente ma non monotono mostrando una certa stagionalità.

Energia Elettrica: L'ACF inizia con circa 0.75 e decresce lentamente nel tempo, rimanendo significativa fino a lag elevati, come per il settore industriale, è presente persistenza nei dati, suggerendo che i consumi di energia passati influenzano in modo marcato i consumi futuri. È visibile una stagionalità periodica con picchi ricorrenti ogni 12 lag indicando una frequenza annuale nei dati.

Residenziale: L'ACF inizia con un valore non troppo elevato al primo lag ma decresce rapidamente, diventando negativa a lag intermedi e alternando segni in modo regolare. Le oscillazioni con valori positivi e negativi alternati suggeriscono che la stagionalità è così potente da determinare l'andamento dell'ACF ben definita nel consumo energetico residenziale in base al periodo invernale o estivo.

Lo stesso ragionamento si ripete sul correlogramma parziale PACF, ovvero l'autocorrelazione parziale osservata al netto/ripulita da tutti i ritardi intermedi.

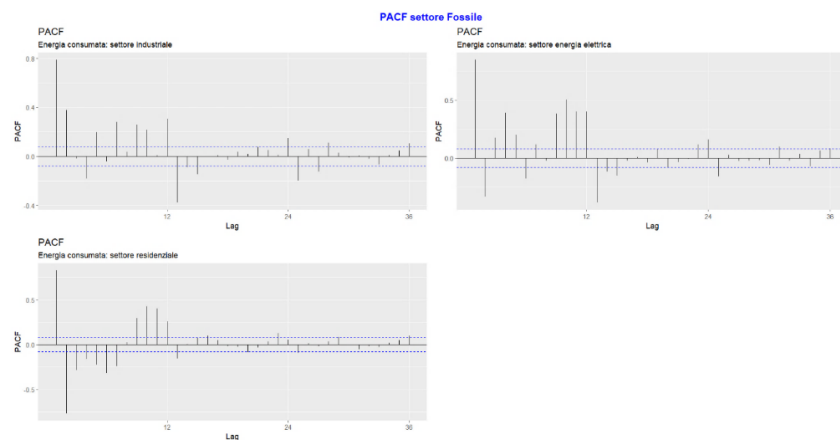


Grafico 7

Industriale: I picchi significativi nei primi lag suggeriscono una dipendenza a breve termine. Il consumo di energia industriale sembra influenzato principalmente dai valori immediatamente precedenti, senza effetti persistenti a lungo termine depurandola dai ritardi intermedi. Questo potrebbe riflettere un uso più legato a dinamiche di mercato o fattori di domanda a breve termine.

Energia elettrica: Il PACF mostra valori significativi nei primi lag (ad esempio, 1 e 12), suggerendo una forte correlazione a breve termine e probabilmente stagionale. Al lag 12 si osserva una forte correlazione, probabilmente dovuta alla stagionalità annuale. Il grafico suggerisce un modello con componenti stagionali.

Residenziale: Anche qui ci sono picchi significativi a breve termine, con un chiaro decadimento progressivo. Alcuni valori oltre il lag 12 sono significativi, ma più deboli, suggerendo la necessità di esplorare la stagionalità. Per questo settore, come l'energia elettrica, potrebbe esserci un'interazione significativa con fattori stagionali.

→ **Test Portmanteau Ljung e Box:** Il test chiamato Portmanteau nella versione di Ljung e Box viene utilizzato per verificare se esiste autocorrelazione significativa di una serie temporale a vari lag, in cui l'ipotesi alternativa afferma che almeno un coefficiente di autocorrelazione risulta

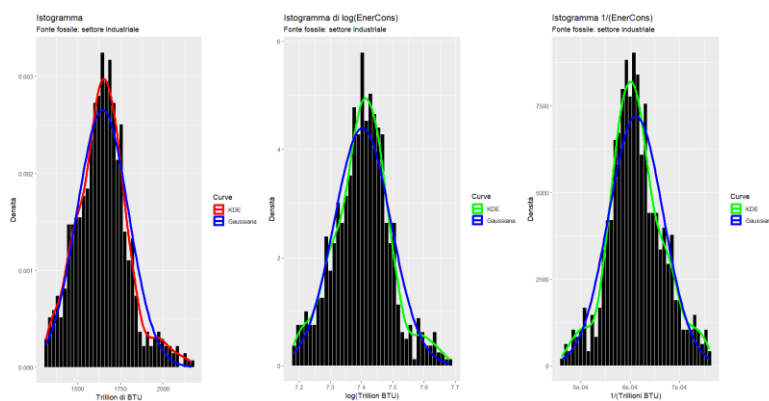
significativamente diverso da zero e quindi conduca al rifiuto dell'ipotesi che la serie sia la realizzazione di un processo White Noise. I p-value in tutti e 3 i casi sono uguali a zero, quindi si rifiuta l'ipotesi nulla.

→**Trasformazione Box-Cox:** A questo punto, dopo aver confermato la non normalità si osservano le trasformazioni ottimali e si valuta se conviene trasformare i dati per migliorarne la normalità. Si sono effettuate le trasformazioni Box-Cox con metodo Guerrero e massima verosimiglianza.

Fonte fossile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Industriale	-1.307185	0.15

Pertanto, si decide di trasformare la serie storica con il logaritmo (per $\lambda = 0.15$) e la trasformazione in scala inversa (per $\lambda = -1.30$) e mostrare gli istogrammi per il confronto.

Grafico 8



Sono stati condotti i test di Bera-Jarque e Shapiro per entrambe le trasformazioni, che mostrano un miglioramento rispetto ai dati non trasformati, ma non sufficiente a giustificare la trasformazione. Si è deciso di mantenere i dati nella loro forma originale per preservarne

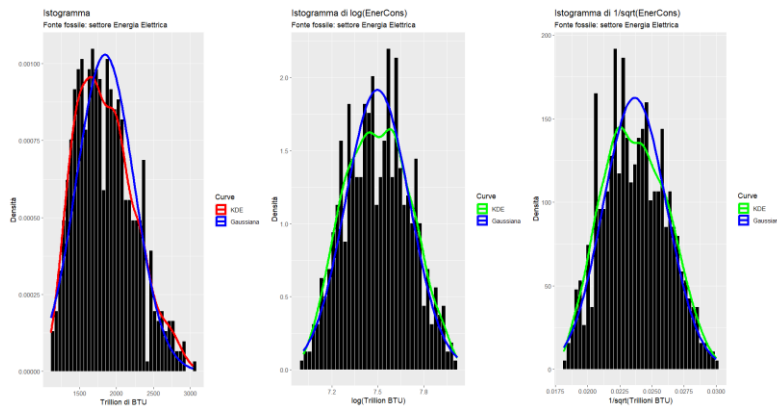
l'interpretabilità.

→**Test ADF:** Si è verificato con il test ADF (test di Dickey Fuller aumentato) la stazionarietà della serie. Si parte con type = “trend” ovvero con la presenza di un trend deterministico; dato che si rifiuta in tutti e tre i casi l'ipotesi nulla di non stazionarietà, in quanto il valore della statistica test è in modulo maggiore dei valori critici, la serie del consumo fossile per il settore industriale è stazionaria intorno al trend deterministico lineare. Non è quindi necessario continuare con gli altri due tipi di test ADF (type= “drift” e type= “None”).

Fonte fossile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Energia Elettrica	-0.6511392	-0.15

Trasformazione della serie con il logaritmo (per $\lambda = -0.15$) e trasformazione in radice quadrata inversa (per $\lambda = -0.65$) e mostrare gli istogrammi per confrontare le distribuzioni trasformate.

Grafico 9



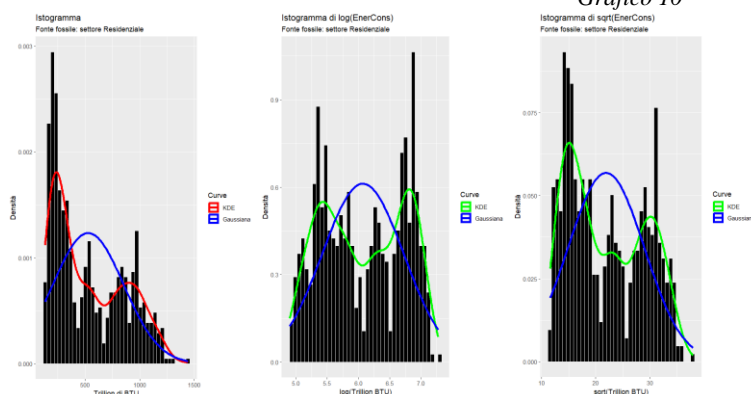
La trasformazione logaritmica migliora la simmetria e l'allineamento con la gaussiana, suggerendo una distribuzione più normale rispetto a quella originale. I test di Bera-Jarque e Shapiro per entrambe le trasformazioni mostrano un miglioramento

rispetto ai dati non trasformati, ma non sufficiente a giustificare l'applicazione di una trasformazione per preservarne l'interpretabilità.

→ **Test ADF:** Si parte anche in questo caso con $\text{type} = \text{"trend"}$ ovvero con la presenza di un trend deterministico; dato che si rifiuta in tutti e tre i casi l'ipotesi nulla di non stazionarietà in quanto il valore della statistica test è in modulo maggiore dei valori critici, la serie risulta essere stazionaria intorno ad un trend deterministico. Inoltre, il coefficiente associato alla variabile tt (trend temporale) è significativo e positivo suggerendo che la serie presenta un trend deterministico crescente nel tempo.

Fonte fossile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Residenziale	0.5811402	0.2

Grafico 10



Trasformazione della serie storica con il logaritmo (per $\lambda = 0.2$) e trasformazione in radice quadrata (per $\lambda = 0.58$). Sia con la trasformazione logaritmica sia quella a radice quadrata, la multimodalità della distribuzione persiste. Non vale

la pena applicare una trasformazione che comprometta l'interpretabilità finale.

→**Test ADF**: Osservando il grafico della serie originale del settore residenziale si decide di partire in questo caso dal test ADF con type = “none” ovvero senza la presenza di costante e di un trend; Dal test ADF si osserva un p-value praticamente nullo, si rifiuta l’ipotesi nulla per tutti 3 i livelli di significatività ciò indica che il processo è stazionario in varianza anche senza trend o senza includere una costante; quindi, la serie storica risulta stazionaria.

Totale fonti rinnovabili (solare, eolico, idroelettrico, geotermico, biomasse)

Analisi di distribuzione empirica dei dati

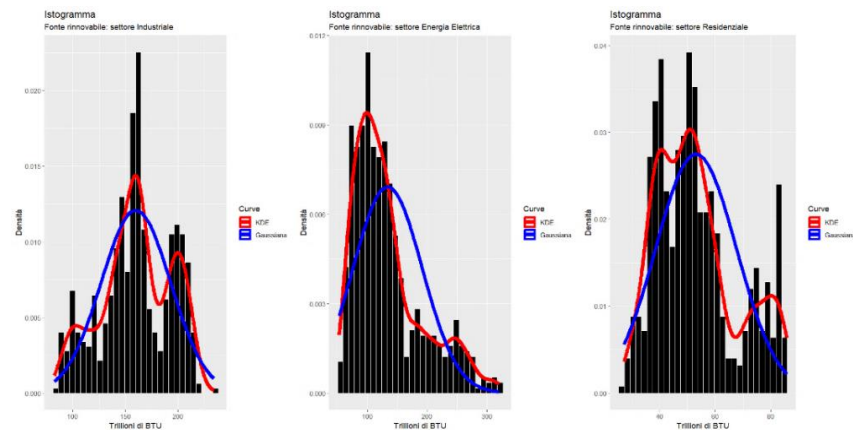


Grafico 11

Industriale: La distribuzione è bimodale suggerendo deviazioni importanti dalla normalità. Suggerisce l'esistenza di sottosettori con profili di consumo molto diversificati, potrebbe essere dovuto dalla varietà di processi produttivi e alle diverse tecnologie adottate.

Energia Elettrica: La distribuzione è asimmetrica verso destra e presenta una coda lunga con presenza di outlier (si decide di risolvere il problema usando la funzione `tsclean()` di R). Anche in questo caso, la curva gaussiana non rappresenta adeguatamente i dati.

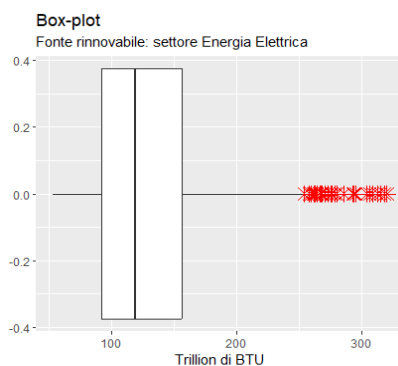


Grafico 12

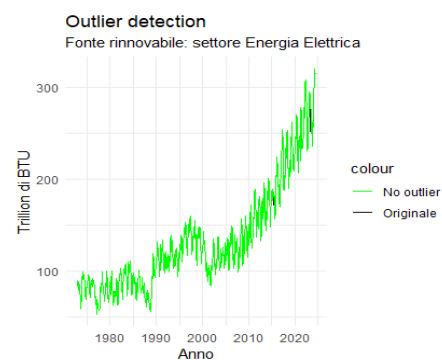


Grafico 13

Residenziale: La distribuzione dei consumi nel settore residenziale presenta caratteristiche simili a quelle del settore industriale, con una marcata asimmetria positiva e una curva KDE multimodale. Ciò indica una grande variabilità nei consumi tra le diverse unità abitative, dovuta a fattori come le dimensioni dell'abitazione, il clima e le abitudini di consumo.

→ **Test di normalità:** In tutti e tre i casi si rigetta l'ipotesi nulla di normalità, p-value molto piccoli.

Test di Bera-Jarque	X-squared	df	p-value
Industriale	17.25	2	0.0001795
Energia Elettrica	140.49	2	< 2.2e-16
Residenziale	48.172	2	3.463e-11

Analisi della persistenza

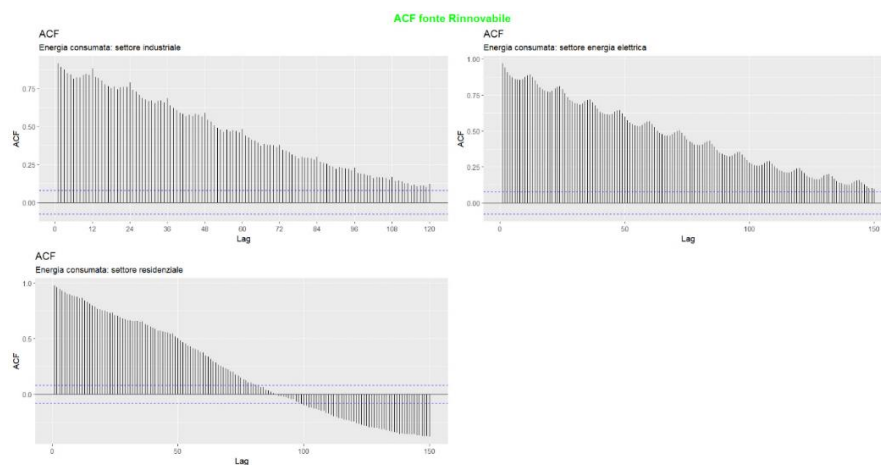


Grafico 14

Industriale: I valori più alti dell'ACF si concentrano nei primi lag, indicando una forte dipendenza dei consumi attuali dai consumi dei periodi immediatamente precedenti. Questo potrebbe essere dovuto a cicli produttivi stagionali. L'ACF diminuisce gradualmente all'aumentare del lag, suggerisce un'elevata persistenza della serie per elevati lag con una memoria a lungo termine.

Energia elettrica: Si osserva una forte autocorrelazione che porta ad affermare ad un'elevata persistenza, con un decadimento graduale all'aumentare del lag. È interessante notare come la serie storica sia soggetta a una certa stagionalità dato che i valori è come se seguissero una sinusoide.

Residenziale: In questo grafico i primi lag presentano valori di autocorrelazione molto elevati, vicini a 1. Questo indica una forte dipendenza dei consumi attuali dai consumi dei periodi precedenti (alta persistenza e caso di non stazionarietà). Si ha un decadimento graduale, quasi monotono indicando una memoria a lungo termine.

Lo stesso ragionamento si ripete sul correlogramma parziale PACF

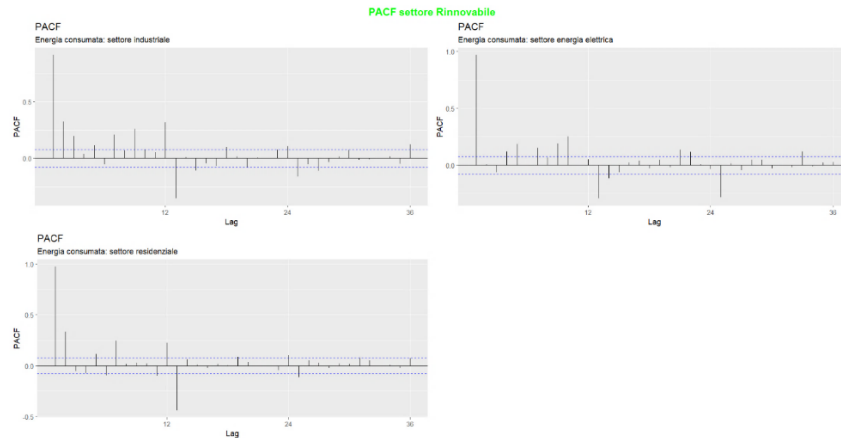


Grafico 15

Industriale: La PACF mostra un rapido decadimento, con poche spunte significative al di fuori dei primi lag. Questo suggerisce che la dipendenza tra i consumi attuali e quelli passati è principalmente dovuta ai consumi del periodo immediatamente precedente.

Energia elettrica: Anche in questo settore si osserva un rapido decadimento della PACF, con poche spunte significative. Tuttavia, rispetto al settore industriale, sembrano esserci alcune oscillazioni che potrebbero indicare una struttura di dipendenza leggermente più complessa.

Residenziale: La PACF presenta un andamento simile agli altri settori, con un rapido decadimento e poche spunte significative. Tuttavia, si osservano alcune oscillazioni negative che potrebbero indicare una struttura di dipendenza stagionale o ciclica.

→**Test Portmanteau Ljung e Box:** A conferma del fatto che le correlazioni sono statisticamente significative, test di Ljung e Box, anche in questo caso i valori dei p-value in tutti e tre i casi sono pari a zero; quindi, si rifiuta l'ipotesi nulla che la serie storiche siano la realizzazione di un processo White Noise.

→**Trasformazione Box-Cox:** Si effettuano le trasformazioni Box-Cox con il metodo Guerrero e massima verosimiglianza.

Fonte rinnovabile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Industriale	2.404008	1.45

Si trasforma la serie storica con il quadrato (per $\lambda = 2.40$) e nel caso di $\lambda = 1.45$ si decide di approssimare il valore a 1 e la trasformazione coincide con la variabile stessa.

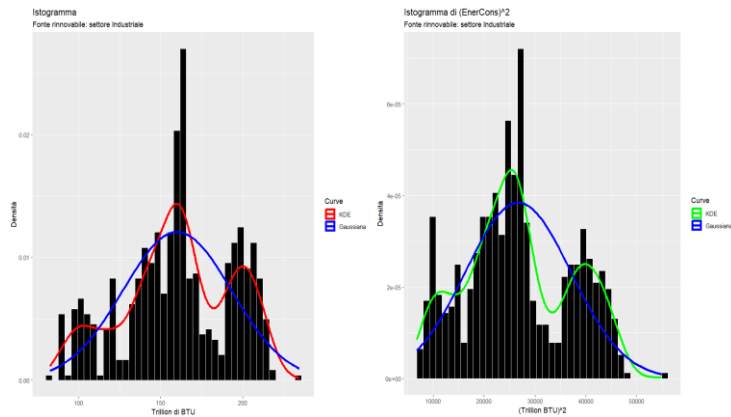


Grafico 16

La multimodalità della serie non si risolve quindi si decide di non trasformare per non perdere l'interpretabilità dei dati.

→**Test ADF:** Type = “trend” ovvero con la presenza di un trend deterministico; Il test ADF ha chiaramente rifiutato l'ipotesi nulla, indicando che la serie non ha una radice unitaria. Il fatto che il coefficiente associato alla variabile tt (trend temporale) sia significativo e positivo suggerisce che la serie presenta un trend deterministico crescente nel tempo.

Fonte rinnovabile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Energia Elettrica	0.3296667	-0.3

Si trasforma con radice quadrata (per $\lambda = 0.33$) e con radice quadrata inversa (per $\lambda = -0.3$).

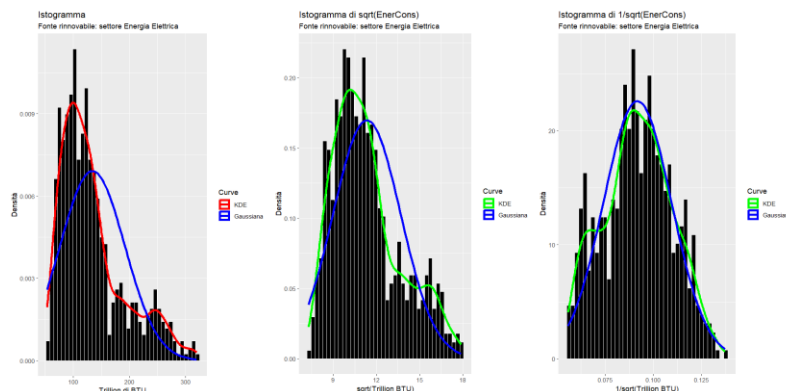


Grafico 17

Anche in questo caso la multimodalità non si risolve quindi si decide di non trasformare i dati per non perdere l'interpretabilità.

→**Test ADF:** Type = “trend” il test ADF ha chiaramente rifiutato l'ipotesi nulla, indicando che la serie non ha una radice unitaria. La serie non ha una tendenza a divergere sempre di più nel tempo, ma oscilla attorno a un valore medio che cambia nel tempo secondo il trend deterministico; quindi, risulta stazionaria attorno ad un trend deterministico.

Fonte rinnovabile	Metodo Guerrero: λ	Metodo massima verosimiglianza: λ
Residenziale	-0.9030261	-0.2

Si trasforma con logaritmo (per $\lambda = -0.2$) e con scala inversa (per $\lambda = -0.9$).

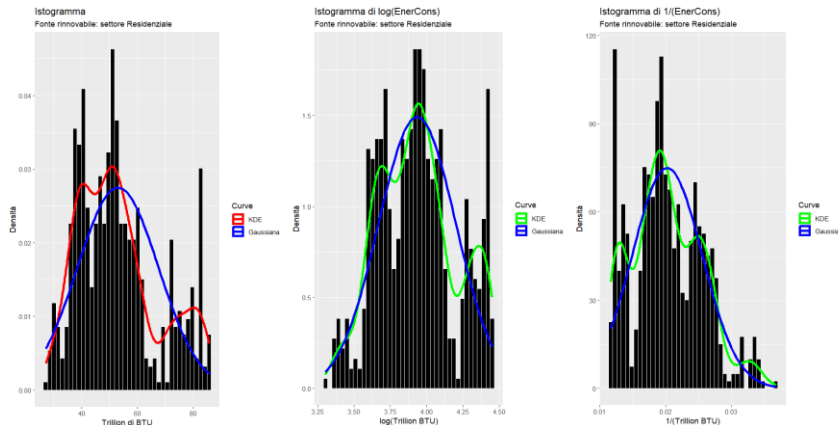


Grafico 18

Attraverso le trasformazioni migliora la distribuzione dei dati rispetto alla normalità ma anche in questo caso la multimodalità non si risolve quindi si decide di non trasformare per non perdere l'interpretabilità dei dati.

→**Test ADF:** Type = "trend" il test ADF con trend deterministico mostra che le statistiche test non superano i valori critici quindi, non si può rifiutare l'ipotesi nulla. La serie non è stazionaria intorno a un trend deterministico. Type = "drift" i risultati indicano che non si può affermare con certezza che la serie sia stazionaria, né intorno a una media costante né a un trend deterministico. Type = "none" l'ipotesi nulla viene accettata, suggerendo la presenza di una radice unitaria. La serie non è stazionaria e potrebbe richiedere trasformazioni come la differenziazione per diventarlo. Si osserva quindi che anche includendo una costante e un trend lineare deterministico, la serie storica dei consumi energetici da fonti rinnovabili nel settore residenziale non è stazionaria. La media e/o la varianza cambiano nel tempo, indicando la necessità di ulteriori trasformazioni per analisi affidabili.

Da queste analisi si conclude che, per tutte le serie storiche relative alle fonti fossili e rinnovabili analizzate nei vari settori, **non è stata applicata alcuna trasformazione**, in quanto o non migliorava significativamente la normalità dei dati oppure comprometteva l'interpretazione degli stessi; quindi, per le successive analisi si è preferito mantenere i dati nel formato originale senza trasformazione.

3. Analisi di detrendizzazione, destagionalizzazione e decomposizione

Trend (detrendizzazione)

I test ADF confermano il rigetto dell'ipotesi nulla di non stazionarietà (tranne per la fonte rinnovabile nel settore residenziale), consentendo di affermare che le serie storiche analizzate sono stazionarie attorno a un trend deterministico. L'analisi si concentrerà quindi sul confronto dei trend dei consumi energetici tra fonti fossili e fonti rinnovabili per esaminare le diverse dinamiche. A tal fine, per rendere le serie stazionarie intorno allo 0, si utilizza una detrendizzazione polinomiale.

Fonte Fossile: settore Industriale

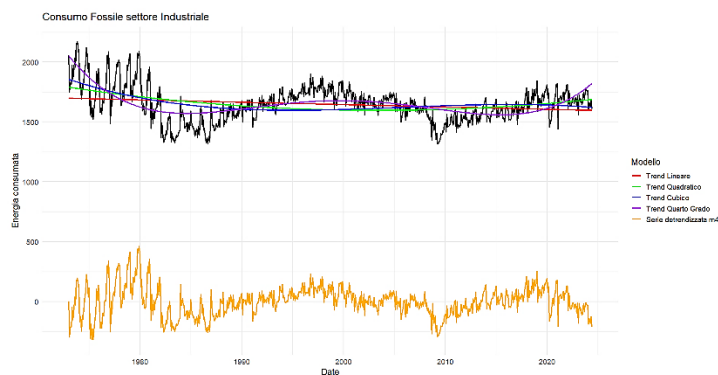


Grafico 19

Trend della serie storica del consumo di fonte fossile nel settore industriale, si è utilizzato un approccio di detrend polinomiale. Il polinomio scelto è quello di quarto grado, in quanto ha le metriche AIC, AICc, BIC più basse e R^2 adj più alto.

	AIC	AICc	BIC	R^2 adjusted
M4	7709.070	7709.207	7735.629	0.33440930

Fonte Rinnovabile: settore Industriale

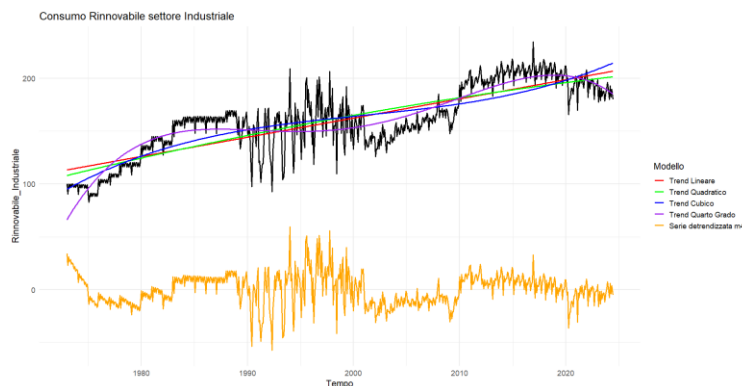


Grafico 20

Trend della serie storica del consumo di fonte rinnovabile nel settore industriale, con approccio di detrend polinomiale. Analisi performance e il polinomio scelto è quello di quarto grado, in quanto ha le metriche AIC, AICc, BIC più basse e R^2 adj più alto.

	AIC	AICc	BIC	R^2 adjusted
M4	5139.047	5139.184	5165.606	0.7834933

I grafici 19 e 20 evidenziano l'andamento opposto dei consumi energetici nel settore industriale. Il consumo di fonti fossili è leggermente in declino, con flessioni marcate durante le crisi petrolifere degli anni '70, la crisi finanziaria del 2008-2009 e la pandemia di COVID-19, accompagnate da una transizione verso energie pulite. Al contrario, il consumo di fonti rinnovabili cresce costantemente dagli anni '80, grazie a investimenti e alla transizione energetica.

Fonte Fossile: settore Energia Elettrica

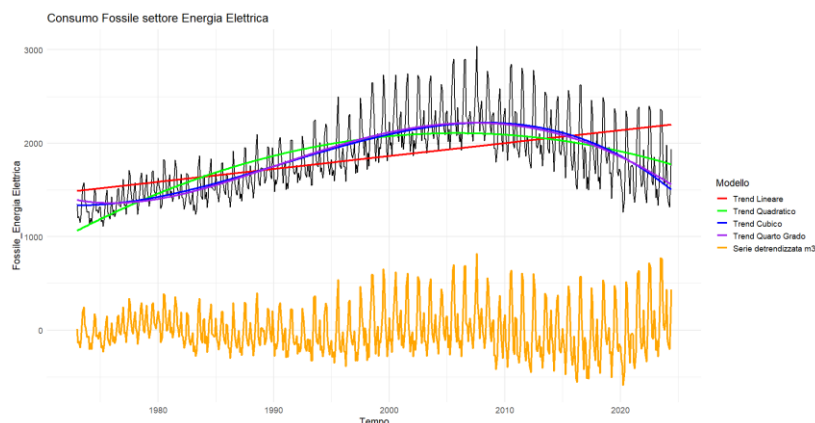


Grafico 21

	AIC	AICc	BIC	R ² adjusted
M3	8569.597	8569.695	8591.729	0.5937535

È stato scelto il modello 3, con trend cubico, poiché fornisce un adattamento molto simile a quello del modello di quarto grado. Pertanto, si è optato per il criterio della parsimonia.

Fonte Rinnovabile: settore Energia Elettrica

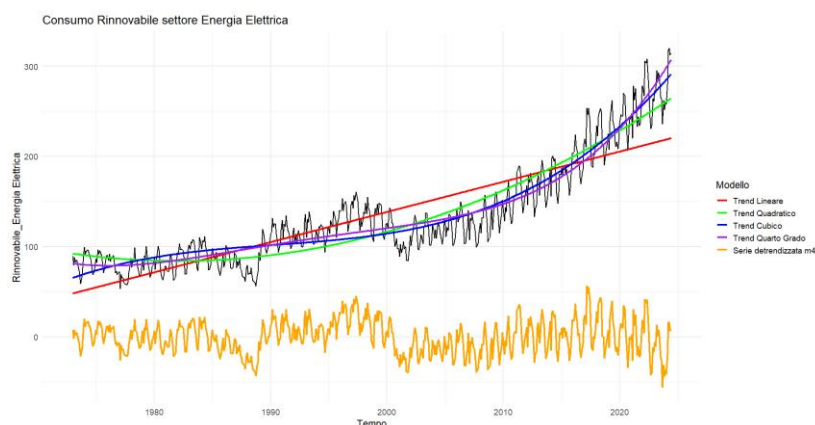


Grafico 22

	AIC	AICc	BIC	R ² adjusted
M4	5368.245	5368.382	5394.804	0.8976833

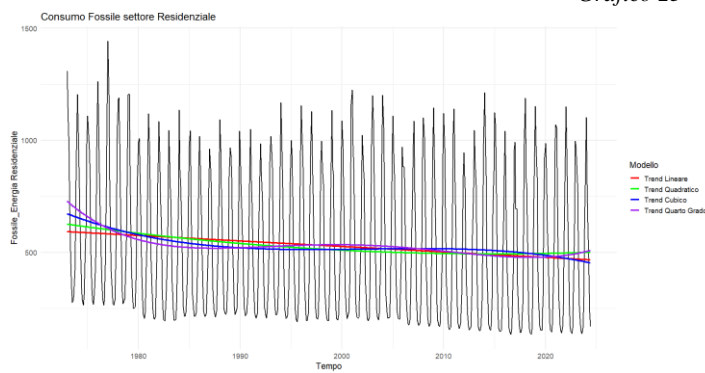
Trend della serie storica del consumo di fonte fossile nel settore energia elettrica, si è utilizzato un approccio di detrend polinomiale. Si è svolta un'analisi sulle performance.

Trend della serie storica del consumo di fonte fossile nel settore energia elettrica, si è utilizzato un approccio di detrend polinomiale. Svolta analisi sulle performance, scelta del trend di quarto grado.

Il consumo di energia derivante da fonte fossili nel settore elettrico statunitense è cresciuto fino agli anni 2000, per poi stabilizzarsi o calare leggermente, mostrando una marcata stagionalità (picchi in inverno, minimi in estate). Modelli di tendenza polinomiali cubici o di quarto grado risultano più adeguati. Il recente calo è attribuibile a prezzi elevati, crescita delle fonti rinnovabili e crisi economiche. Entrambe le fonti mostrano stagionalità e complessità temporale che verranno studiate tra poco, ma le fonti fossili hanno fluttuazioni più ampie, mentre le rinnovabili crescono costantemente con minori variazioni.

Fonte Fossile: settore Residenziale

Grafico 23



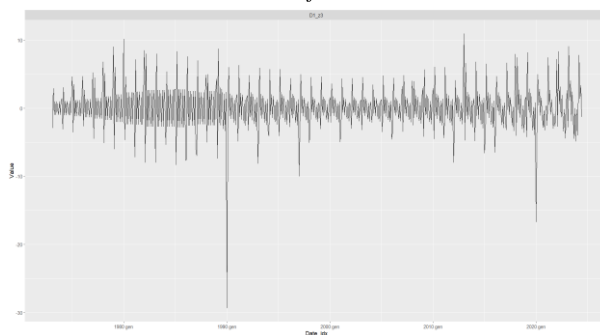
In questo caso la serie storica risulta stazionaria però si esegue la detrendizzazione per confermare il fatto che vi è un trend leggermente decrescente e non per correggere la non stazionarietà in media. All'aumento del polinomio la serie storica smussata tende

a interpolare più efficacemente i dati. Tuttavia, questa non è sempre la scelta ottimale poiché mantenendo basso il grado del polinomio consente di ottenere serie storiche più regolari.

Fonte Rinnovabile: settore Residenziale

In questo caso dai test ADF si è accettata l'ipotesi nulla di non stazionarietà; quindi, un processo non stazionario in varianza ma anche in media può essere reso stazionario applicando il concetto di **differenziazione** della serie storica, si decide di applicare la differenza prima. Differenziare una

Grafico 24



serie significa trasformarla in una serie che rappresenta le variazioni della variabile tra due periodi consecutivi. I valori ora rappresentano quanto la variabile è cambiata rispetto al periodo precedente, anziché il livello assoluto della variabile. Se un valore è positivo, significa che il consumo di energia rinnovabile è aumentato

rispetto al periodo precedente. Se invece è un valore negativo, il consumo è diminuito rispetto al periodo precedente. Una volta applicata la differenziazione e il test ADF la serie differenziata

risulta stazionaria poiché si rifiuta l'ipotesi nulla del test ADF. Nella serie differenziata si nota la presenza di due valori con picchi molto bassi; il primo nel 1990 perché ci sono state delle modifiche al “Clean Air Act”, mirate a ridurre l'inquinamento atmosferico, hanno introdotto normative sulle emissioni che potrebbero aver limitato l'uso delle biomasse (settore più colpito) nel settore residenziale, aumentando i costi per adeguare gli impianti di combustione, mentre il secondo nel 2020 a causa dell'effetto della pandemia da Covid 19.

Stagionalità (destagionalizzazione)

La destagionalizzazione è una procedura che serve per eliminare la stagionalità, ottenendo una serie storica destagionalizzata, che permette di distinguere tra cambiamenti causati da eventi naturali o da shock casuali. Nel caso del consumo energetico, l'obiettivo è verificare se le variazioni osservate siano effettivamente legate alla stagionalità. Questo processo aiuta a isolare le fluttuazioni non associate a fattori stagionali migliorando così la comprensione delle tendenze a lungo termine e delle anomalie nel consumo energetico. La destagionalizzazione è stata realizzata tramite una regressione armonica, utilizzando l'analisi di Fourier delle sinusoidi tramite seni e coseni con un polinomio di grado k , più è elevato il grado più il polinomio approssima bene l'andamento stagionale della serie storica, ma con un rischio di over fitting.

Fonte Fossile: settore Industriale

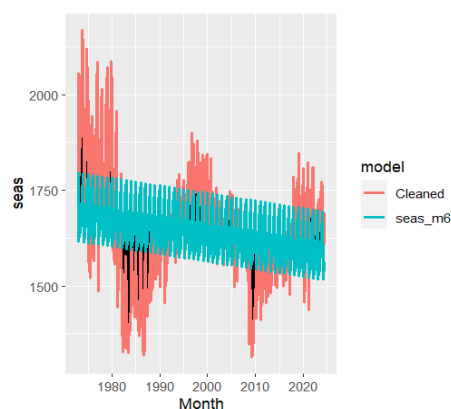


Grafico 25

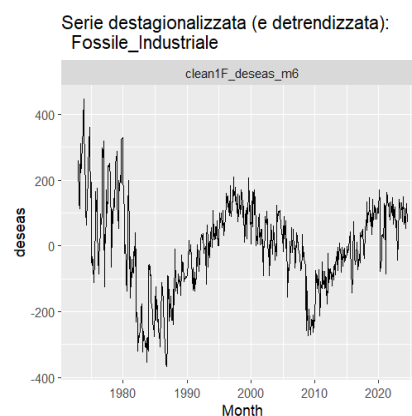


Grafico 26

Nel grafico 25 si può notare il trend e la stagionalità stimati dalla serie storica del consumo di fonte fossile nel settore industriale ottenuta con un numero di armoniche di Fourier pari a $K=6$, scelto attraverso criteri di fitting attraverso AIC e AICc. Si può notare, come già visto nell'analisi precedente un trend leggermente decrescente e una stagionalità piuttosto regolare e persistente. Nel grafico 26 invece si è riportata la serie destagionalizzata, la serie mostra le variazioni residue

una volta rimosse le componenti stagionali e trend. Si notano picchi significativi che possono essere collegati a eventi storici, come crisi energetiche (anni '70-'80), recessioni economiche o transizioni verso fonti energetiche più pulite. Dopo il 2000, si osservano oscillazioni più contenute, probabilmente legate alla maggiore stabilità nei consumi o a miglioramenti tecnologici.

Fonte Rinnovabile: settore Industriale

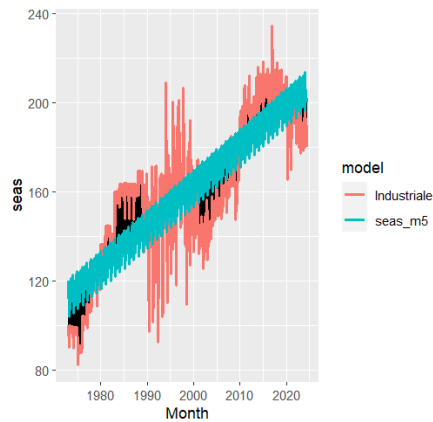


Grafico 27

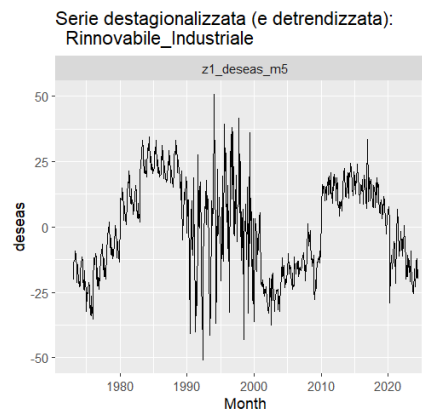


Grafico 28

Nel grafico 27 si può notare il trend e la stagionalità stimati dalla serie storica del consumo di energia proveniente da fonte rinnovabile nel settore industriale ottenuta con un numero di armoniche di Fourier pari a $K=5$, scelto attraverso criteri di fitting e parsimonia. Si nota un trend crescente riflettendo l'aumento graduale dei consumi di energia da fonti rinnovabili nel tempo e una stagionalità regolare che subisce fluttuazioni periodiche significative, probabilmente legate a fattori stagionali per la domanda industriale. La serie destagionalizzata risulta stazionaria intorno allo 0 questo è confermato dal rifiuto dell'ipotesi nulla del test ADF per Type = "none".

Fonte Fossile: settore Energia Elettrica

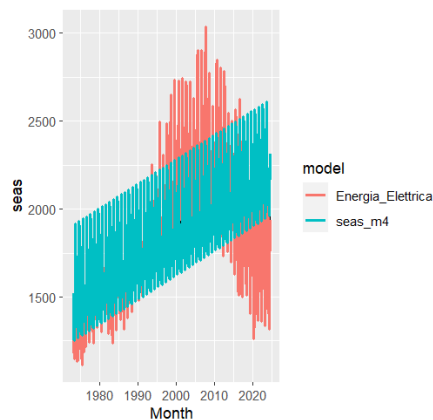


Grafico 29

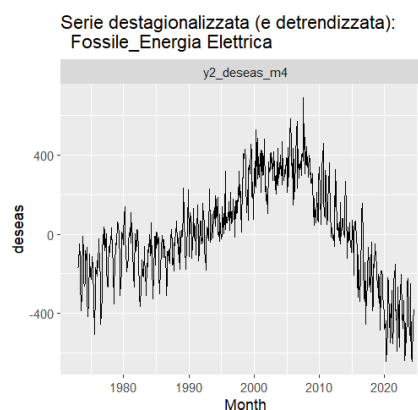


Grafico 30

Nel grafico 29 si osserva il trend e la stagionalità stimati dalla serie storica del consumo di energia proveniente da fonte fossile nel settore dell'energia elettrica ottenuta con un numero di armoniche

di Fourier pari a $K=4$, scelto attraverso criteri di fitting e parsimonia. Trend crescente e stagionalità regolare basata dalle oscillazioni relative al consumo di energia elettrica della fonte fossile. La serie destagionalizzata risulta stazionaria intorno allo 0 questo è confermato dal rifiuto dell'ipotesi nulla del test ADF per Type = "none", anche se il grafico risulta non centrato attorno allo zero.

Fonte Rinnovabile: settore Energia Elettrica

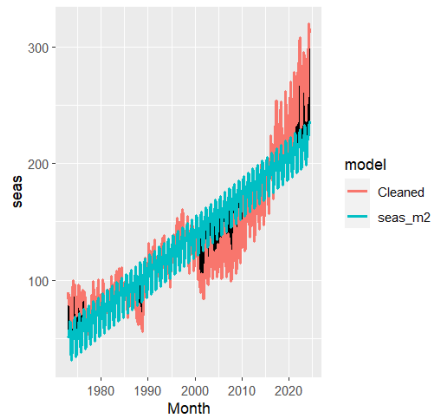


Grafico 31

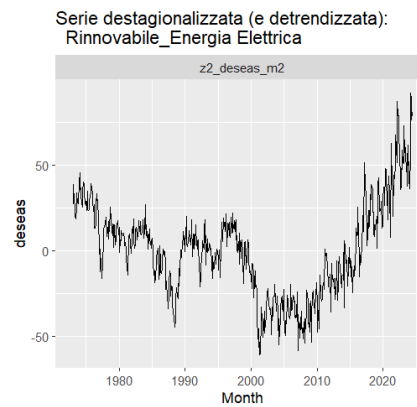


Grafico 32

Nel grafico 31 si osserva la serie storica del consumo di energia proveniente da fonte rinnovabile nel settore dell'energia elettrica ottenuta con un numero di armoniche di Fourier pari a $K=2$, scelto attraverso criteri di fitting con metriche di AIC, AICc, BIC le più basse e R^2 adj più alto. Trend crescente e stagionalità regolare, nel grafico 32 si osserva la serie destagionalizzata la serie appare stazionaria attorno allo zero, confermato dal rifiuto ipotesi nulla del test ADF per Type = "none".

Fonte Fossile: settore Residenziale

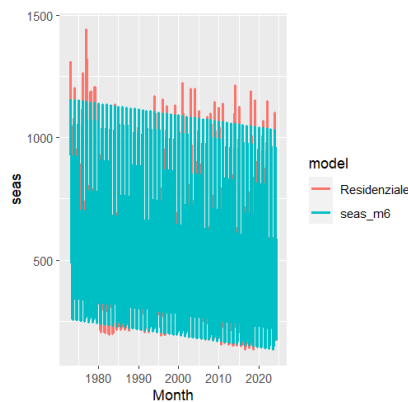


Grafico 33

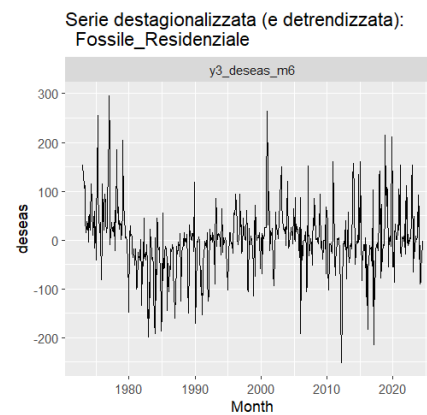


Grafico 34

Nel grafico 33 si osserva il trend e la stagionalità stimati dalla serie storica del consumo di energia proveniente da fonte fossile nel settore industriale ottenuta con un numero di armoniche di Fourier pari a $K=6$, scelto attraverso criteri di fitting con metriche di AIC, AIC più basse e BIC, R^2 adj più alte. Nel grafico 32 si osserva la serie destagionalizzata la serie appare stazionaria attorno allo

zero; questo significa che la serie originale era governata dalla stagionalità con andamenti oscillatori più elevati per periodi invernali e andamento più bassi per periodi estivi relativi al consumo di energia nel settore residenziale quindi legato a riscaldamento e raffreddamento.

Fonte Rinnovabile: settore Residenziale

Precedentemente per la fonte rinnovabile settore residenziale si è applicata la differenziazione della serie originale; pertanto, non è stata effettuata la destagionalizzazione della serie differenziata, poiché la differenziazione modifica la struttura della stagionalità. Nella serie originale, la stagionalità si manifesta come un andamento ripetitivo (ad esempio, consumi più alti in inverno e più bassi in estate). Nella serie differenziata, invece, la stagionalità emerge come una ripetizione di pattern nei cambiamenti tra i periodi (ad esempio, incrementi positivi o negativi in corrispondenza di determinati periodi temporali).

Decomposizione

La decomposizione è una tecnica fondamentale nell'analisi delle serie storiche, in quanto consente di suddividere la serie nelle sue componenti principali: il trend, la stagionalità e la parte residua. Questo processo permette di analizzare separatamente ciascun elemento, fornendo una visione più chiara delle dinamiche sottostanti ai dati osservati. Esistono diversi metodi per decomporre una serie storica, e in questo caso si è scelto di adottare la decomposizione additiva. Tale scelta è motivata dall'ipotesi che le fluttuazioni stagionali e quelle attorno al trend siano indipendenti dal livello complessivo della serie storica, rendendo il modello additivo più appropriato rispetto ad altri approcci, come quello moltiplicativo. I risultati ottenuti all'interno del codice in R, dall'estrazione delle componenti di trend e stagionalità, si sono rivelati coerenti con quelli già descritti nella sezione precedente. Per questo motivo ad analisi svolte, si è deciso di dare poca priorità a questa parte dell'analisi non riportando i risultati.

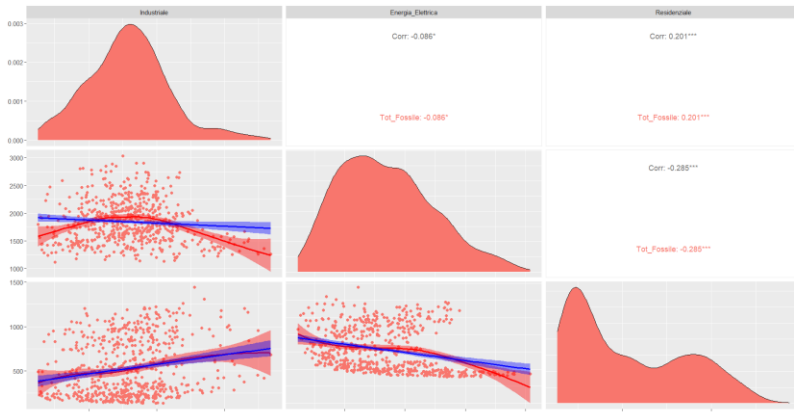
4. Analisi di regressione e analisi dei residui

Dato che le serie storiche sono stazionarie come confermato dai test ADF (tranne la serie storica per fonti rinnovabili settore residenziale dove verrà effettuata un'analisi in cui il modello di regressione è calcolato sulle serie differenziate), si procede all'analisi di regressione sulle serie dei dati grezzi per le fonti fossili.

Fonti fossili

Per prima cosa si passa ad uno studio approfondito degli scatterplot e delle correlazioni tra le variabili d'interesse dei consumi energetici nei vari settori:

Grafico 35



→Tra settore **residenziale** e **industriale** (0.201 correlazione poco forte), indicando che un aumento dei consumi fossili industriali tende ad associarsi a un incremento nei consumi residenziali, probabilmente per dinamiche macroeconomiche.

→Correlazione tra **residenziale** ed **energia elettrica** (-0.285 correlazione poco forte), suggerendo che un maggiore uso di energia fossile nel settore residenziale si accompagna a un minor consumo nel settore elettrico, indicando possibili diversificazioni nell'utilizzo delle fonti energetiche.

→Correlazione tra **industriale** ed **energia elettrica** è invece prossima allo zero, evidenziando l'assenza di una relazione significativa tra i due settori. Nel grafico, le linee blu rappresentano la regressione lineare, mentre le linee rosse indicano un modello loess; nel caso di residenziale e industriale, le linee si sovrappongono bene alla tendenza lineare. Si ipotizza che il settore elettrico possa rispondere a variazioni nella domanda residenziale e industriale attraverso economie di scala, influenzando i consumi. Inoltre, fattori come il prezzo del petrolio e regolamentazioni ambientali potrebbero incidere sui consumi complessivi. Si sono stimati tre modelli di regressione per le fonti fossili, per semplicità si indicano le variabili in questo modo: Energia Elettrica = EE, Industriale = I, Residenziale = R

1° modello → $EE = \alpha_0 + \alpha_1 I + \alpha_2 R + \varepsilon_t$

2° modello → $EE = \beta_0 + \beta_1 R + \varepsilon_t$

3° modello → $EE = \gamma_0 + \gamma_1 I + \varepsilon_t$

Model	AIC	AICc	BIC	R ² adj
1° modello	9074.411	9074.476	9092.117	0.079036271
2° modello	9072.993	9073.032	9086.273	0.079664544
3° modello	9120.703	9120.742	9133.982	0.005800308

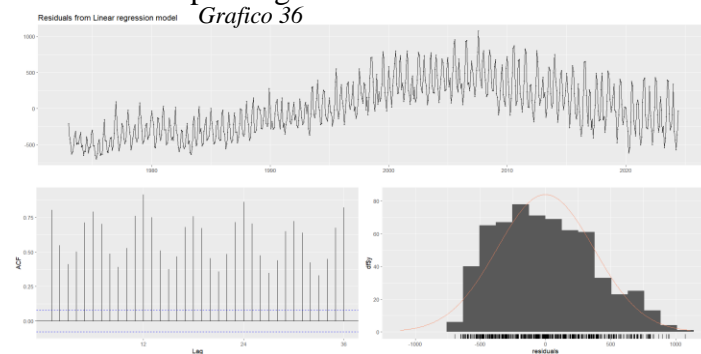
Il modello m2 sembra essere il modello migliore in base ai criteri informativi e al R^2_{adj} , mentre il modello m1 è una valida alternativa, con prestazioni molto simili a m2. Si decide quindi di utilizzare il modello completo per valutare una visione più dettagliata dei fattori che guidano il consumo di energia fonte fossile. Si osservano i parametri di questo modello attraverso il summary:

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2148.75637   164.88566   13.032 < 2e-16 ***
Tot_Fossile_Industriale -0.07744    0.10170   -0.761    0.447
Tot_Fossile_Residenziale -0.33410    0.04726   -7.070 4.22e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 372 on 615 degrees of freedom
Multiple R-squared:  0.08202,    Adjusted R-squared:  0.07904
F-statistic: 27.48 on 2 and 615 DF,  p-value: 3.724e-12
```

L'intercetta è circa 2149, che rappresenta il valore previsto del consumo di energia fossile nel settore energia elettrica quando il consumo di energia fossile nel settore industriale e residenziale è nullo; questo non ha senso da un punto di vista interpretativo in quanto non è possibile avere un consumo di energia nullo. Il coefficiente associato al consumo di energia fossile nel settore industriale è -0.07744, questo indica che, se il consumo di energia fossile settore energia elettrica cresce di una unità, questo è associato a una riduzione di -0.07744 unità nel consumo fossile settore industriale, mantenendo costanti le altre variabili. Questo concetto vale anche per il consumo fossile settore residenziale a cui è associato un valore stimato di -0.33410. Per quanto riguarda la significatività delle variabili si osserva che per energia fossile settore residenziale ha un p-value molto basso 4.22e-12, suggerendo che il coefficiente è statisticamente significativo mentre energia fossile settore industriale il p-value è pari a 0.447, che è relativamente alto indicando che questo coefficiente potrebbe non essere statisticamente significativo nell'ambito del modello stimato. L' R^2 è molto basso 0.08202, indicando che circa l'8.2% del consumo di energia fossile è spiegato dal modello 1 (adattamento molto basso che si osserva anche negli altri due modelli stimati).

Grafico 36: il primo grafico dei residui stimati con il modello 1, si può notare che l'andamento dei



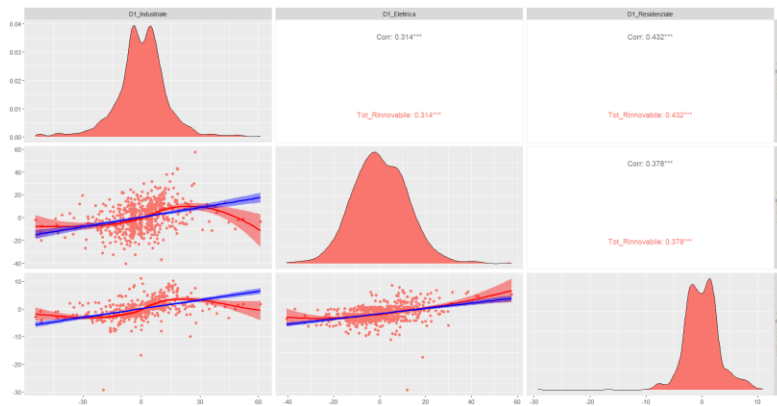
residui è abbastanza stazionario attorno allo zero. La distribuzione dei residui è simmetrica intorno allo zero anche se in maniera approssimativa. Il grafico delle ACF mostra come i residui sono significativi oltre i 30 lag, segno di

autocorrelazione e quindi cade l'ipotesi di distribuzione White Noise dei residui. Si conclude che i residui si distribuiscono solo normalmente con media zero e varianza costante.

Fonti rinnovabili

In questo caso la serie del settore residenziale è stata differenziata una volta, mentre gli altri due settori non necessitavano di differenziazione. Applicare un'ulteriore differenziazione a serie storiche già stazionarie potrebbe introdurre una perdita di informazione, ma l'obiettivo è garantire che le serie abbiano la stessa integrazione (ordine di differenziazione), per questo motivo sono state differenziate anche le altre due serie storiche.

Grafico 37



Dal grafico 37 si osservano gli scatter plot e le correlazioni: dal momento che le serie sono differenziate, l'analisi si concentra sulle variazioni rispetto al periodo precedente, e non sui valori assoluti del consumo energetico nei tre settori.

→ Tra settore **industriale** ed **energia elettrica** (0.314 correlazione poco forte), indicando che un incremento nel consumo rinnovabile dell'industria tende a essere accompagnato da un incremento nella stessa direzione nel settore dell'energia elettrica. → Tra settore **residenziale** e **industriale** (0.432, il più alto rispetto alle altre correlazioni), significa che le variazioni nel consumo rinnovabile dell'industria tende a essere collegato ad un incremento nella stessa direzione nel settore residenziale, può essere legato agli stessi utilizzati di energie rinnovabili nei settori residenziali e industriali. → La correlazione tra **residenziale** ed **energia elettrica** (0.378 correlazione poco forte), indica che il settore residenziale è moderatamente influenzato dalle variazioni nel consumo elettrico. I tre istogrammi con le curve di densità mostrano la distribuzione delle variazioni dei consumi nei tre settori, mentre ogni scatter plot mostra la relazione tra due settori con una regressione lineare (linee blu) mentre le linee rosse indicano un modello loess. La relazione tra le variabili è positiva anche se l'effetto non sembra molto forte in nessun caso. Stima dei modelli di regressione, come prima: Energia Elettrica = EE, Industriale = I, Residenziale = R

$$1^{\circ} \text{ modello } \rightarrow \Delta R_t = \alpha_0 + \alpha_1 \Delta I_t + \alpha_2 \Delta EE_t + \varepsilon_t$$

$$2^{\circ} \text{ modello } \rightarrow \Delta R_t = \beta_0 + \beta_1 \Delta I_t + \varepsilon_t$$

$$3^{\circ} \text{ modello } \rightarrow \Delta R_t = \gamma_0 + \gamma_1 \Delta EE_t + \varepsilon_t$$

Model	AIC	AICc	BIC	R ² adj
1° modello	3053.163	3053.228	3070.869	0.2498017
2° modello	3102.908	3102.947	3116.188	0.1856072
3° modello	3135.365	3135.404	3148.644	0.1416933

Il modello completo m1 sembra essere il miglior modello in base ai criteri informativi e al R²adj, si osservano i parametri di questo modello attraverso il summary:

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.015620   0.114684   0.136   0.892
Tot_Rinnovabile_D1_Industriale 0.084594  0.008928   9.475 < 2e-16 ***
Tot_Rinnovabile_D1_Elettrica  0.072782  0.009931   7.329 7.34e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.85 on 615 degrees of freedom
Multiple R-squared:  0.2522,    Adjusted R-squared:  0.2498
F-statistic: 103.7 on 2 and 615 DF,  p-value: < 2.2e-16

```

Valore intercetta pari a 0.015620, il p-value associato (0.892) è alto, indicando che l'intercetta non è statisticamente significativa. Rappresenta la variazione attesa nel consumo di energia rinnovabile settore residenziale quando le variazioni nei settori industriale ed elettrico sono entrambe zero (non ha senso da un punto di vista interpretativo). Il coefficiente associato al settore industriale è 0.084594 con p-value molto basso indicando una forte significatività statistica; significa che per ogni unità di aumento nella variazione del consumo di energia rinnovabile industriale, ci si aspetta un aumento di 0.084594 unità nella variazione del consumo nel settore residenziale (in altre parole, se il consumo industriale di rinnovabili aumenta più rapidamente rispetto al periodo precedente, anche il consumo residenziale tenderà ad aumentare più rapidamente). Il coefficiente associato al settore elettrico è 0.072782 con un p-value basso (7.34e-13), indicando significatività statistica. Significa che per ogni unità di aumento nella variazione del consumo di energia rinnovabile nel settore elettrico, ci si aspetta un aumento di 0.072782 unità nella variazione del consumo residenziale. Distribuzione dei residui del modello nel grafico 38: il primo grafico dei residui

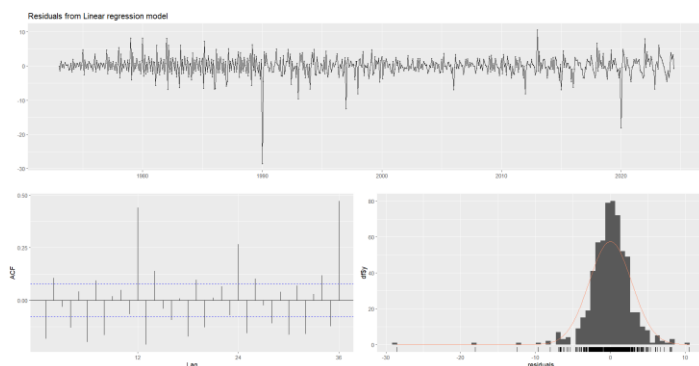


Grafico 38

stimati con il modello 1, si può notare che l'andamento dei residui è stazionario attorno allo zero. Distribuzione dell'istogramma è praticamente normale mentre il grafico delle ACF mostra come ci siano alcuni lag 12, 24 e 36 con autocorrelazioni significative, il che suggerisce che i residui sono autocorrelati, cade l'ipotesi di distribuzione White Noise dei residui.

5. Stima e comparazione di modelli ARIMA, SARIMA e regARIMA

In questa sezione si sono stimati i modelli ARIMA, SARIMA e regARIMA poiché l'obiettivo è quello di stimare il modello statistico che fitta meglio i dati delle serie storiche. I modelli ARIMA sono a due stadi: prima fase destagionalizzazione e detrendizzazione della serie tramite armoniche di Fourier K ; seconda fase applicazione del filtro ARIMA sui residui del modello con armoniche K . I modelli SARIMA invece si ottengono moltiplicando il modello ARIMA (p, d, q) sulla serie storica con il modello ARIMA che cattura la stagionalità. Mentre nei modelli regARIMA la stima avviene tramite una procedura multistadio basata sulla massima verosimiglianza in cui si passano uno o più filtri ARIMA sui residui di regressione fino a quando questi non risultano white noise.

Fonte Fossile: settore Industriale

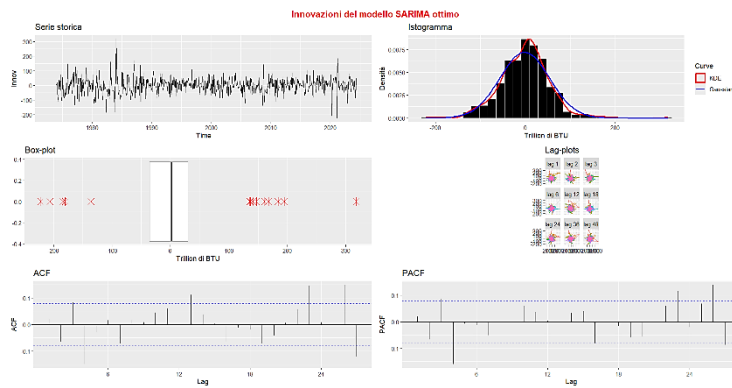
ARIMA: Dall'output si osserva che il modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(3,0,2) con ordine della parte auto regressiva $p=3$, con ordine di differenziazione $d=0$, quindi stazionario, e con ordine della parte a media mobile $q=2$. Quindi il processo è influenzato dal suo passato fino al ritardo $t-3$ e da shock esogeni fino al ritardo $t-2$.

SARIMA: Modello SARIMA che va a catturare al meglio l'andamento della serie è dato dalla moltiplicazione tra l'ARIMA(1,0,1) e il modello ARIMA(2,1,1)[12] che cattura la stagionalità della serie. La serie storica ha componente autoregressiva di ordine 1, ovvero che la serie storica dipende dalle sue osservazioni passate fino a $t-1$ e una componente a media mobile di ordine 1. Per la componente stagionale, il processo è differenziato una volta e ha una componente autoregressiva stagionale pari a 1 e media mobile pari a 1; con stagionalità pari a 12.

RegARIMA: Applicando la procedura regARIMA si ottiene un ARIMA (3,1,2), ovvero con una parte autoregressiva di ordine 3, il processo è differenziato una volta e ha una componente a media mobile pari a 2, ovvero che la serie storica dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 2$.

Il modello migliore tra ARIMA, SARIMA e regARIMA risulta essere il modello SARIMA. Nel grafico 39 vengono riportati diversi grafici, tra cui l'andamento dei residui, l'istogramma, il Box-plot, lag plot e i grafici delle autocorrelazioni. Dal primo grafico i residui hanno un andamento che sembra essere casuale attorno allo 0 e questo porta a pensare che potrebbero distribuirsi come dei WN a media 0 e varianza costante. Dall'istogramma si nota che la distribuzione dei residui è abbastanza normale, simmetrica attorno allo 0.

Grafico 39



Dal grafico delle ACF e PACF i residui sono incorrelati tra loro. Pertanto, si può concludere che i residui $\sim WN(0, \sigma^2)$.

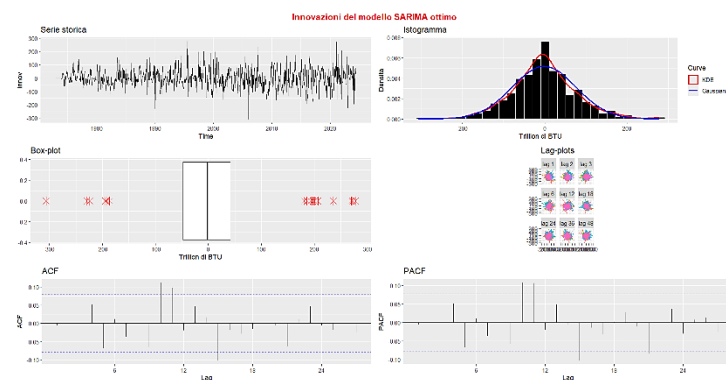
Fonte Fossile: settore Energia Elettrica

ARIMA: Modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(1,0,1) cioè con ordine della parte auto regressiva $p=1$, stazionario e con l'ordine della parte a media mobile $q=1$. Processo influenzato dal suo passato fino al ritardo $t-1$ e da shock esogeni fino al ritardo $t-1$.

SARIMA: Il modello SARIMA ottimale è dato dalla moltiplicazione tra l'ARIMA(1,1,1) e il modello ARIMA(2,1,1)[12] che cattura la stagionalità della serie. La serie storica ha componente autoregressiva di ordine 1 e una componente a media mobile di ordine 1. La stagionale, il processo è differenziato una volta rispetto alla stagionalità e ha una componente autoregressiva stagionale pari a 2 e media mobile stagionale pari a 1; con stagionalità pari a 12.

RegARIMA: regARIMA si ottiene un ARIMA (0,1,3), ovvero con una parte autoregressiva di ordine 0, il processo è differenziato una volta e ha una componente a media mobile pari a 3, la serie storica dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 3$.

Grafico 40



Il modello migliore risulta essere il modello SARIMA. Nel grafico 40 si nota che i residui hanno un andamento casuale attorno allo 0 e questo porta a pensare che potrebbero distribuirsi come dei WN a media 0 e varianza costante. Dall'istogramma si nota che

la distribuzione dei residui è abbastanza normale, simmetrica attorno alla media pari a 0 e questo è confermato anche nel box-plot. Dal grafico delle ACF e PACF i residui sono incorrelati tra loro. Si può concludere che i residui $\sim WN(0, \sigma^2)$.

Fonte Fossile: settore Residenziale

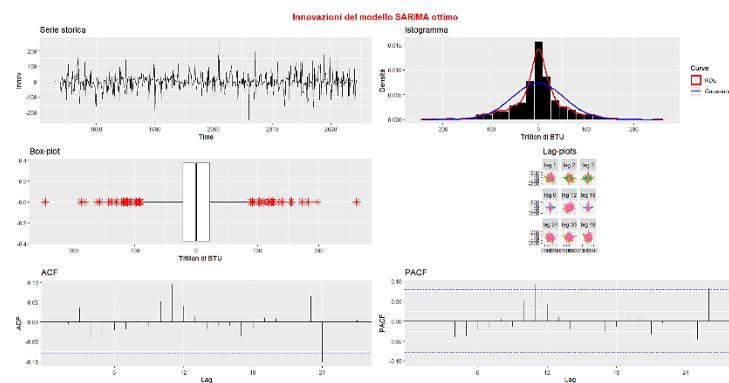
ARIMA: modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(2,0,1) cioè con ordine della parte auto regressiva $p=2$, stazionario e con l'ordine della parte a media mobile $q=1$. Processo influenzato dal suo passato fino al ritardo $t-2$ e da shock esogeni fino al ritardo $t-1$.

SARIMA: Modello ottimale è dato dalla moltiplicazione tra l'ARIMA(2,0,1) e il modello ARIMA(0,1,1)[12] che cattura la stagionalità. La serie storica ha componente autoregressiva di ordine 2, ovvero che la serie storica dipende dalle sue osservazioni passate fino a $t-2$ e una componente a media mobile di ordine 1. Nella componente stagionale, il processo è differenziato una volta rispetto alla stagionalità e ha una componente media mobile stagionale pari a 1.

RegARIMA: Procedura regARIMA si ottiene un ARIMA (2,1,3), ovvero con una parte autoregressiva di ordine 2, il processo è differenziato una volta e ha una componente a media mobile pari a 3, la serie dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 3$.

Il modello migliore tra ARIMA, SARIMA e regARIMA risulta essere il modello SARIMA.

Grafico 41



Dal grafico 41 si osserva che i residui hanno un andamento che sembra essere casuale attorno allo 0 e questo porta a pensare che potrebbero distribuirsi come dei WN a media 0 e varianza costante. Dall'istogramma si nota che la distribuzione dei residui non risulta

perfettamente normale. Dal grafico delle ACF e PACF i residui sono incorrelati tra loro, nel complesso si può concludere che i residui $\sim WN(0, \sigma^2)$.

Fonte Rinnovabile: settore Industriale

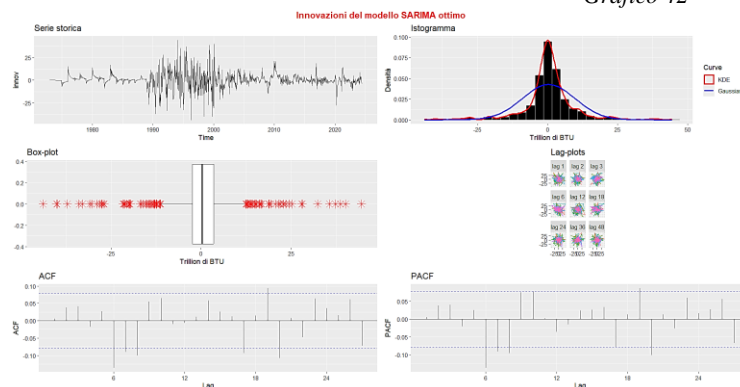
ARIMA: Modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(2,0,1) cioè con ordine della parte auto regressiva $p=2$, stazionario, e con ordine media mobile $q=1$. Quindi il processo è influenzato dal suo passato fino al ritardo $t-2$ e da shock esogeni fino al ritardo $t-1$.

SARIMA: Modello ottimale è dato dalla moltiplicazione tra l'ARIMA(1,0,2) e il modello ARIMA(0,1,2)[12] che cattura la stagionalità della serie. La serie storica ha componente autoregressiva di ordine 1, ovvero che la serie storica dipende dalle sue osservazioni passate fino

a $t-1$ e una componente a media mobile di ordine 2. La componente stagionale, il processo è differenziato una volta e ha una componente media mobile stagionale pari a 2; con stagionalità 12.

RegARIMA: Si ottiene un ARIMA (1,1,1), ovvero con una parte autoregressiva di ordine 1, il processo è differenziato una volta e ha una componente a media mobile pari a 1, ovvero la serie dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 1$.

Grafico 42



Il modello migliore risulta essere il modello SARIMA. Dal grafico 42 i residui hanno un andamento casuale attorno allo 0 e questo porta a dire che possono distribuirsi come dei WN a media 0 e varianza costante. Dall'istogramma si nota che la

distribuzione dei residui non risulta perfettamente normale. Dal grafico delle ACF e PACF i residui sono incorrelati tra loro, nel complesso si può concludere che i residui $\sim WN(0, \sigma^2)$.

Fonte Rinnovabile: settore Energia Elettrica

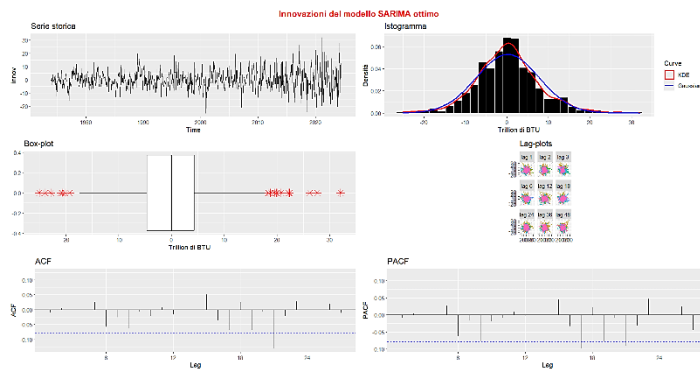
ARIMA: Dall'output si osserva che il modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(1,0,1) cioè con ordine della parte auto regressiva $p=1$, con ordine di differenziazione $d=0$, quindi stazionario, e con l'ordine della parte a media mobile $q=1$. Quindi il processo è influenzato dal suo passato fino al ritardo $t-1$ e da shock esogeni fino al ritardo $t-1$.

SARIMA: Modello ottimale è dato dalla moltiplicazione tra l'ARIMA(1,1,1) e il modello ARIMA(1,1,2)[12] che cattura la stagionalità della serie. Ha componente autoregressiva di ordine 1 e una componente a media mobile di ordine 1 con una differenziazione. Per quanto riguarda la componente stagionale, il processo è differenziato una volta rispetto alla stagionalità e ha una componente autoregressiva stagionale pari a 1 con media mobile pari a 2; con stagionalità 12.

RegARIMA: Applicando la procedura regARIMA si ottiene un ARIMA (0,1,3), il processo è differenziato una volta e ha una componente a media mobile pari a 3, ovvero che la serie storica dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 3$.

Modello migliore risulta essere il modello SARIMA. Dal grafico 43 si osserva che i residui hanno un andamento casuale attorno allo 0 e questo porta a pensare che potrebbero distribuirsi come dei WN a media 0 e varianza costante.

Grafico 43



Dall'istogramma si nota che la distribuzione dei residui risulta normale. Dal grafico delle ACF e PACF i residui sono incorrelati tra loro, nel complesso si può concludere che i residui $\sim WN(0, \sigma^2)$.

Fonte Rinnovabile: settore Residenziale

ARIMA: Modello ARIMA che meglio identifica l'andamento della serie è un ARIMA(3,1,1) cioè con ordine della parte auto regressiva $p=1$, non stazionario, e con ordine media mobile $q=1$. Quindi il processo è influenzato dal suo passato fino al ritardo $t-3$ e da shock esogeni fino al ritardo $t-1$.

SARIMA: Modello ottimale è dato dalla moltiplicazione tra l'ARIMA(0,1,2) e il modello ARIMA(2,0,0)[12] che cattura la stagionalità della serie. La serie storica ha componente a media mobile di ordine 2 con applicata una differenziazione. Per quanto riguarda la componente stagionale, il processo non è differenziato rispetto alla stagionalità e ha una componente autoregressiva stagionale pari a 2; con stagionalità pari a 12.

RegARIMA: Si ottiene un ARIMA (3,1,1), il processo è differenziato una volta e ha una componente autoregressiva pari a 3 e una media mobile pari a 1, ovvero che la serie storica dipende dalla media pesata lineare di shock esogeni casuali fino a $t - 1$.

Il modello migliore risulta essere anche in questo caso il modello SARIMA; nel grafico 44 si osserva che i residui hanno un andamento casuale attorno allo 0.

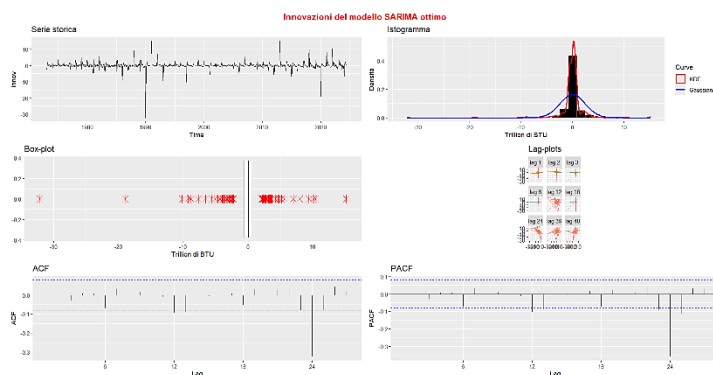


Grafico 44

Nell'istogramma si nota che la distribuzione dei residui risulta normale anche se molto sottile. Dal grafico delle ACF e PACF i residui sono incorrelati tra loro, nel complesso si può concludere che i residui $\sim WN(0, \sigma^2)$.

6. Conclusioni

Per concludere le analisi del seguente progetto si può dire che si è svolta una panoramica di quelli che sono i consumi energetici nei tre settori analizzati per le fonti fossili e rinnovabili. I risultati ottenuti in linea generale sono buoni e hanno permesso di rispondere alle domande poste all'inizio, in particolare si evidenzia un calo dell'uso di energia fossile e una crescita delle rinnovabili nei 3 settori anche se presentano ancora delle differenze a livello di quantità utilizzo. Negli ultimi decenni, il carbone è stato progressivamente sostituito dal gas naturale e dalle fonti rinnovabili, poiché rappresenta non più una fonte sostenibile nel lungo periodo. Il settore industriale, sebbene ancora dipendente dai combustibili fossili, ha migliorato l'efficienza energetica, mentre quello residenziale sta adottando sempre più fonti rinnovabili e impianti fotovoltaici, soprattutto su abitazioni private e aziende. Le politiche statunitensi, attraverso incentivi e normative ambientali, stanno accelerando la transizione energetica; tuttavia, restano aperte le sfide legate all'accumulo di energia rinnovabile, all'intermittenza della produzione e alla dipendenza dal gas naturale. Per quanto riguarda i modelli di regressione stimati non hanno dato risultati soddisfacenti a causa dell'adattamento molto basso (per fonti fossili) dovuto a correlazioni deboli e poco significative. Mentre per quanto riguarda i modelli ARIMA, SARIMA e regARIMA il modello che è risultato migliore per tutti i settori è stato il modello SARIMA perché presentava criteri di AIC, AICc, BIC migliori rispetto agli altri modelli, indicando una maggiore capacità di adattamento ai dati e una previsione più accurata; soprattutto perché SARIMA è un'estensione del modello ARIMA che incorpora componenti stagionali, permettendo di catturare meglio le variazioni periodiche nei dati che presentavano stagionalità essendo legati all'utilizzo di energia elettrica su base mensile. In conclusione, considerando l'ampiezza territoriale e la diversità politica ed economica degli Stati Uniti, la transizione alle fonti rinnovabili rimane una sfida complessa. Il sistema energetico statunitense è ancora fortemente dipendente dalle fonti fossili e, per garantire sicurezza energetica in un futuro, è necessario investire nella ricerca, diversificare le fonti e migliorare le infrastrutture di produzione e distribuzione dell'energia rinnovabile.

Riferimenti bibliografici

1. Richard S.J., Stephen W. Pacala, Robert H. Socolow, (2008). *Understanding Long-Term Energy Use and Carbon Dioxide Emissions in the USA*. [ScienceDirect].
2. Paul Sabin, (2012). *Crisis and Continuity in U.S. Oil Politics, 1965–1980*. [The Journal of American History].