

# Area B

---

Risorsa ambientale o limitazione per i cittadini

Luca De Simone

Matricola: 854294

# Indice

<b>1. Introduzione .....</b>	<b>2</b>
<b>2. Diossido di Azoto (NO<sub>2</sub>) .....</b>	<b>4</b>
2.2 Distribuzione della serie e test di Normalità .....	4
2.3 Trasformazione di Box-Cox .....	5
2.4 Outlier Detection .....	6
2.5 Analisi della Persistenza.....	7
2.6 Analisi delle componenti non osservabili.....	9
2.7 Modellistica .....	11
2.7.1 ARIMA: modellistica two steps .....	11
2.7.2 SARIMA .....	13
2.7.3 Confronto tra modelli .....	14
<b>3. Temperatura.....</b>	<b>14</b>
3.3 Analisi delle componenti non osservabili.....	15
3.4 Modellistica .....	16
<b>4. Precipitazioni .....</b>	<b>17</b>
4.1 Analisi delle componenti non osservabili.....	19
4.2 Modellistica .....	19
<b>5. Umidità relativa .....</b>	<b>20</b>
5.1 Analisi delle componenti non osservabili.....	22
5.2 Modellistica .....	22
<b>6. Velocità del Vento.....</b>	<b>23</b>
6.1 Analisi delle componenti non osservabili.....	25
6.2 Modellistica .....	25
<b>7. Radiazione Globale .....</b>	<b>26</b>
7.1 Analisi delle componenti non osservabili.....	27
7.2 Modellistica .....	28
<b>8. Policy Evaluation .....</b>	<b>29</b>
8.1 Regressione lineare e Analisi dei risultati .....	29
<b>9. Conclusioni .....</b>	<b>31</b>
<b>Bibliografia .....</b>	<b>31</b>

# 1. Introduzione

L'inquinamento atmosferico è uno dei principali fattori del rischio ambientale. Oltre a provocare malattie cardiovascolari e respiratorie, è responsabile di molteplici effetti negativi sul clima, sulla qualità dell'aria e sul suolo, danneggiando interi ecosistemi. Benché le fonti responsabili di tale fenomeno siano plurime, le principali sono rappresentate dal settore industriale, dalla produzione energetica, dall'agricoltura e dai trasporti. Gli agenti inquinanti presenti in atmosfera sono diversi, ma i più problematici sono le particelle fini (PM), gli ossidi d'azoto (NO<sub>x</sub>) e l'ozono troposferico (O<sub>3</sub>). Il piano d'azione dell'UE mira ad ottenere entro il 2050 una Neutralità Climatica, ovvero una società ed un'economia a zero emissioni. Questo obiettivo a lungo termine è stato suddiviso in obiettivi chiave per il 2030 tra i quali annoveriamo il cosiddetto pacchetto "Pronti per il 55%". Tali norme consistono nel ridurre le emissioni dei gas ad effetto serra dell'unione europea di almeno il 55% entro il 2030, da qua la denominazione.

I gas serra sono dei tipi di gas presenti in atmosfera che assorbono l'energia solare e il calore irradiato dalla superficie terrestre intrappolandolo nell'atmosfera. Questo processo è alla base dell'effetto serra che, mantenendo la temperatura più elevata sulla terra, rispetto a come sarebbe in sua assenza, crea le condizioni favorevoli alla vita sul pianeta. Molti gas ad effetto serra sono naturalmente presenti in atmosfera, tuttavia l'attività umana ne aggiunge una quantità così elevata da provocare il cosiddetto surriscaldamento globale. L'anidride carbonica (CO<sub>2</sub>) rappresenta circa l'80% dei gas serra rilasciati nell'atmosfera. Ad ogni modo, esistono anche altri tipi di gas che, seppur emessi in misura minore, hanno una capacità maggiore rispetto alla CO<sub>2</sub> di trattenere il calore. Ad esempio, il diossido di azoto (NO<sub>2</sub>), seppur non si tratti propriamente di un gas serra, ha un GWP (Global Warming Potential) 298 volte superiore di quello dell'anidride carbonica, su un periodo di 100 anni. Infatti, l'NO<sub>2</sub>, in presenza di luce solare e composti organici volatili (VOC), può reagire e formare l'ozono troposferico (O<sub>3</sub>), un gas serra. Inoltre, contribuisce alla formazione di particolato fine (PM) che anch'esso può avere effetti sul riscaldamento globale a seconda della sua composizione e interazione con la luce.

*Nota a margine: in questa introduzione si tratterà sempre e solamente di emissioni di CO<sub>2</sub>, questo perché, come avviene negli articoli scientifici e nella letteratura, per comparare i GWP delle sostanze presenti in atmosfera, essendo questi diversi, vengono convertiti in CO<sub>2</sub> equivalente.*

Il settore energetico è responsabile del 77.1% delle emissioni di gas serra, circa un terzo del quale è attribuibile ai trasporti. Secondo l'European Environment Agency (EEA), il 71.7% delle emissioni dei trasporti è prodotta da quello stradale, di cui il 60.7% derivante dalle autovetture. Di recente, il tasso di riduzione delle emissioni dei trasporti ha subito un rallentamento, l'incremento della mobilità delle persone ha causato un aumento di CO<sub>2</sub>. Infatti, nel 2017 si è registrata una controtendenza, le nuove auto hanno emesso una media di 0.4 grammi di anidride carbonica in più rispetto a quelle immatricolate l'anno precedente. Per frenare l'andamento, l'UE ha introdotto nuovi target sulle emissioni ed incentivato le alimentazioni alternative portando molte case automobilistiche a cessare la produzione del diesel in favore dell'elettrico o ibrido plug-in. Questo ha portato un aumento dell'immatricolazioni dei veicoli elettrificati passando dal 10%, registrato nel 2020, al 17.8% del 2021. Ad onor di cronaca è necessario riportare come le vetture elettriche non siano propriamente mezzi a zero emissioni in quanto le fasi di produzione e smaltimento dei veicoli contribuiscono al calcolo del tasso delle emissioni prodotte e, da questo punto di vista, quelle con motore a combustione interna sono di gran lunga più ecologiche. Anche sotto l'aspetto della ricarica, i mezzi non sono così green fintanto la maggior parte dell'elettricità utilizzata non è prodotta mediante fonti rinnovabili.

Oltre alle direttive europee, per combattere il fenomeno delle emissioni in aumento sono state emanate diverse normative nazionali e comunali. Tra queste si annovera l'introduzione dell'Area C da parte del Comune di Milano, avvenuta nel gennaio del 2012 per migliorare la qualità dell'area, ridurre il traffico ed incentivare i cittadini ad utilizzare il trasporto pubblico. Successivamente, nel febbraio del 2019, è stata introdotta, sempre dal comune meneghino, l'Area B, ovvero un'ulteriore zona a traffico limitato con un'estensione di circa il 72% dell'intera città rendendola così una delle più grandi ZTL europee. Nello specifico, si tratta di un'iniziativa dell'amministrazione comunale,

facente parte di un piano ambientale più ampio, avente come scopo quello di ridurre l'inquinamento atmosferico e migliorare la qualità dell'aria del capoluogo lombardo. Nel corso degli anni, in modo progressivo, si sono legate all'area diverse restrizioni riguardanti differenti categorie di veicoli. Rispettivamente:

- ) **25 febbraio 2019:** divieto d'accesso ai veicoli
  - Diesel Euro 0, 1, 2, 3
  - Benzina Euro 0
- ) **1° ottobre 2019:** divieto d'accesso ai veicoli
  - Diesel Euro 4 senza Filtro AntiParticolato (FAP)
- ) **1° ottobre 2020:** divieto d'accesso ai veicoli
  - Diesel Euro 4 con FAP
- ) **1° ottobre 2022:** divieto d'accesso ai veicoli
  - Diesel Euro 5
- ) **1° ottobre 2023:** divieto d'accesso ai veicoli
  - Benzina Euro 2

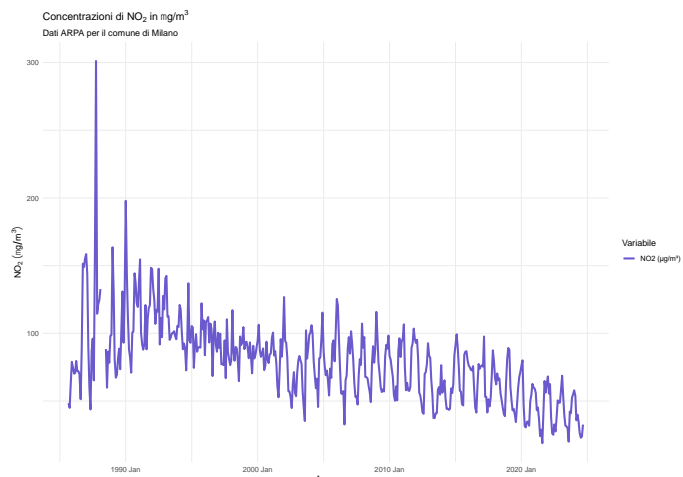
E molti altri in programma negli anni consultabili sulla pagina dedicata del Comune di Milano. Da una prospettiva ambientale, l'Area B rappresenta sicuramente una risorsa, d'altro canto, per i cittadini, ha rappresentato e rappresenterà sempre di più una limitazione ed un disagio, soprattutto dal punto di vista economico in quanto non tutti saranno in grado di adeguarsi e sostenere le spese che tali divieti comporteranno. Secondo le statistiche, circa un terzo dei residenti comunali dichiara di avere un reddito annuo inferiore ai 16.000 euro, mentre, solamente il 13.6%, supera i 55.000 euro, dati del Ministero dell'Economia e Finanze (MEF) del 2020. Dunque, da questa breve disamina la domanda d'analisi sorge spontanea: "L'Area B, e più specificatamente le restrizioni associate, rappresenta una risorsa per l'ambiente o una limitazione per i cittadini?". Lo scopo è quello di comprendere se tali divieti, applicati ad un perimetro così specifico e ristretto dell'area urbana milanese, hanno avuto un impatto ambientale e qual è stato il suo ordine di grandezza. Per rispondere a tutti questi quesiti si è provveduto ad analizzare la serie storica del diossido di azoto (NO<sub>2</sub>), inquinante maggiormente prodotto dal trasporto stradale secondo quanto riportato dal consiglio dell'Unione Europea. L'idea è quella di effettuare una policy evaluation sulla serie tenendo sotto controllo tutti i fattori meteorologici che potrebbero interferire nell'analisi. Nello specifico si è tenuto conto delle temperature medie, delle precipitazioni, della radiazione globale, dell'umidità relativa e della velocità del vento, dati facilmente reperibili, in maniera open, dall'Agenzia Regionale per la Protezione dell'Ambiente (ARPA).

Lo studio verrà articolato in due fasi, la prima riguardante le singole serie storiche per esplorare, identificare e comprendere al meglio le componenti di cui sono formate, mentre la seconda riguarderà nello specifico la policy evaluation dell'Area B e le restrizioni ad essa associate. L'analisi delle serie verrà effettuata mediante la procedura di Box-Jenkins, metodologia suddivisa su più punti che permetterà l'identificazione del modello ottimale in grado di approssimare i rispettivi andamenti temporali.

## 2. Diossido di Azoto (NO<sub>2</sub>)

Il diossido di azoto (NO<sub>2</sub>) è una delle sostanze primarie prodotte dalla combustione dei carburanti fossili, il che lo rende il principale inquinante nel settore dei trasporti. Le rilevazioni a disposizione sono espresse in microgrammi per metro cubo, hanno una frequenza mensile e ricoprono una finestra temporale che ha inizio nel settembre del 1985 e termina nel settembre del 2024.

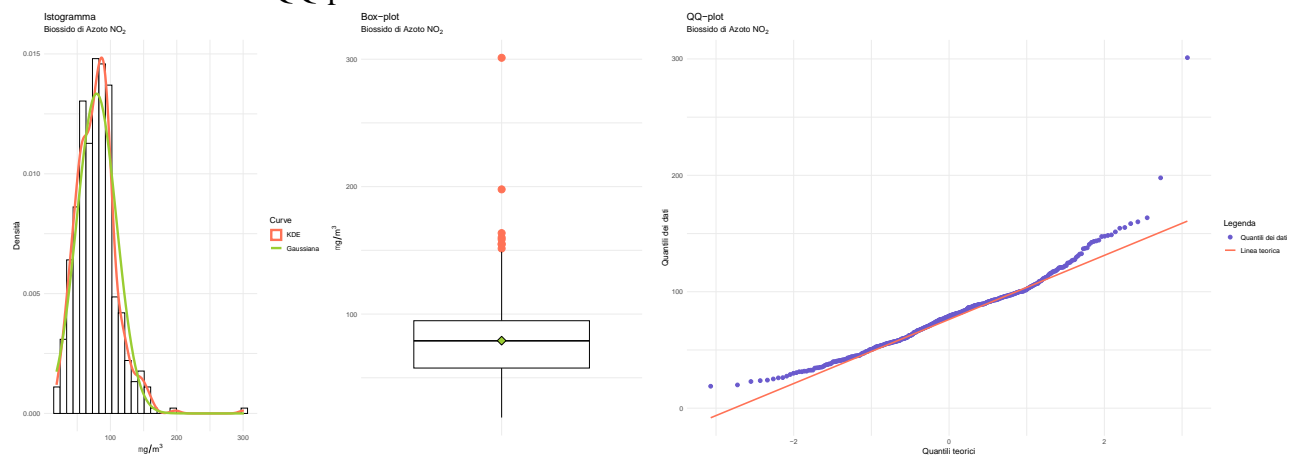
Da un'analisi grafica preliminare, la serie sembrerebbe assumere un comportamento stazionario attorno ad un trend decrescente accompagnato da una stagionalità di discreta importanza. Le oscillazioni parrebbero avere un andamento additivo, non si evidenzia un'influenza temporale, ma si ipotizzano essere affette da eteroschedasticità. La sequenza presenta, in un primo momento, una variabilità notevole che, intorno al 1995, si ritiene essersi contratta ed assestata ad un'oscillazione decisamente più contenuta. La fluttuazione, a partire dai primi 2000, torna a possedere una varianza più elevata che si



manterrà all'incirca costante fino alla conclusione della finestra d'osservazione. Tale variabilità potrebbe essere dovuta ad un cambio della stagionalità nel tempo, alla presenza di eventuali valori atipici o ad una modifica delle tecniche di rilevazione del fenomeno. Citando Pearson, si definisce valore anomalo o outlier un'osservazione *“incoerente con il resto dei dati”* cioè *“insolitamente al di sopra o al di sotto dei valori tipici dei dati”*. Nello specifico, si individuano solamente due rilevazioni anomale, ovvero quelle relative all'ottobre 1987 e al gennaio 1990.

### 2.2 Distribuzione della serie e test di Normalità

Gli stimatori di densità e gli strumenti grafici consentono di comprendere le proprietà distributive dei fenomeni in studio. Nello specifico, si sono utilizzati stimatori kernel, stimatori istogramma, il plot di Box-Whisker e il QQ-plot.



I grafici evidenziano come la serie si distribuisca approssimativamente come una Normale in quanto si rileva una buona sovrapposizione tra la curva empirica e quella teorica. Tramite l'istogramma e, in maniera ancora più accentuata, attraverso il Box-plot, è possibile denotare come la coda di destra risulti piuttosto pesante a causa di alcune osservazioni anomale. Il QQ-plot è uno strumento che permette di confrontare i quantili empirici con quelli teorici sotto ipotesi di normalità. Dal grafico non emerge alcuna criticità ad eccezione delle code, il che rafforza quanto affermato riguardo la loro pesantezza e l'asimmetria distributiva. Una soluzione a tale problematica potrebbe essere quella di attuare una trasformazione logaritmica.

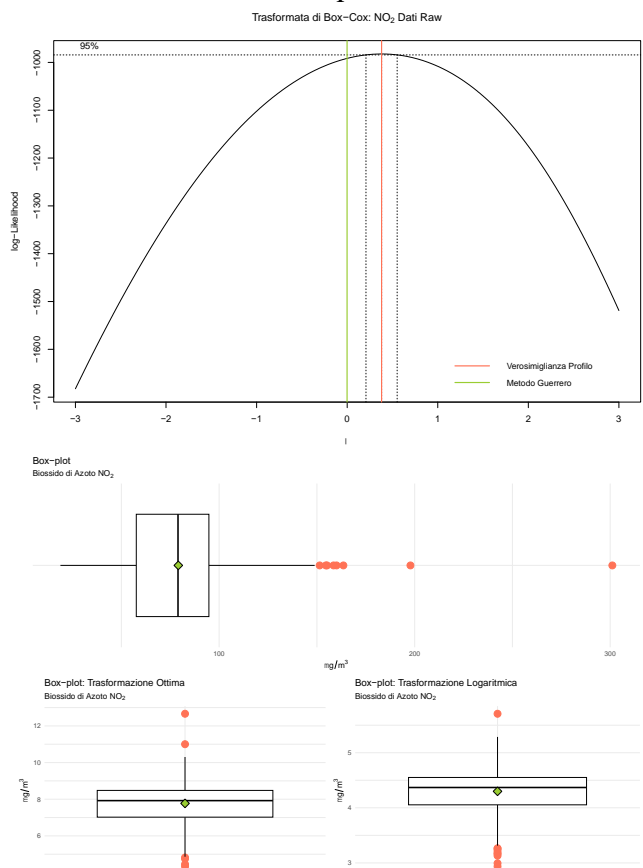
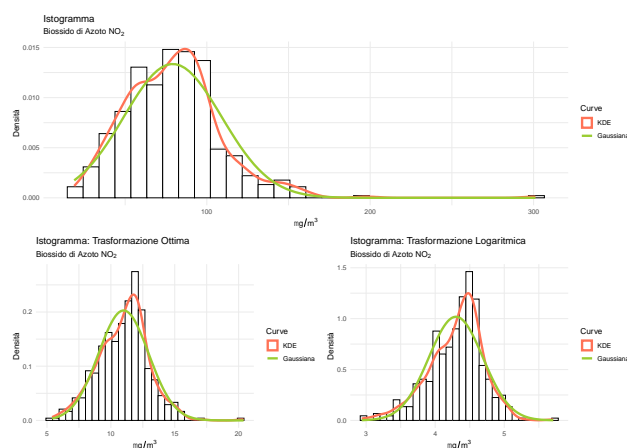
I test di Normalità effettuati sono tutti concordi nel rifiutare l'ipotesi nulla distributiva. Il test di Bera-Jarque, basato sulla stima campionaria degli indici di simmetria e curtosi, rifiuta aspramente l'assunto di normalità. Questo è comprensibile dato che le code, soprattutto per quanto riguarda il valore di curtosi, impattano fortemente sulla valutazione dell'assunto teorico. Allo stesso modo, i valori estremi della distribuzione hanno una forte influenza sul test di Shapiro-Wilk. Infatti, trattandosi di un indice sintetico di correlazione tra valori empirici e teorici, risulta maggiormente sensibile agli scostamenti limite. Al contrario, il test di Kolmogorov-Smirnov si basa sulla massima distanza tra la Funzione di Ripartizione Empirica (ECDF) e quella teorica (CDF), sotto ipotesi nulla. Questo tenderà a rifiutare  $H_0$  se la discrepanza tra le funzioni è molto ampia, e, osservando il valore del p-value, non sembrerebbe essere tale.

Test di normalità	Statistica test	P-value
Jarque Bera Test	906,39	< 2,2e-16
Shapiro-Wilk test	0,9381	5,49E-10
Kolmogorov-Smirnov test	0,0692	0,0231

## 2.3 Trasformazione di Box-Cox

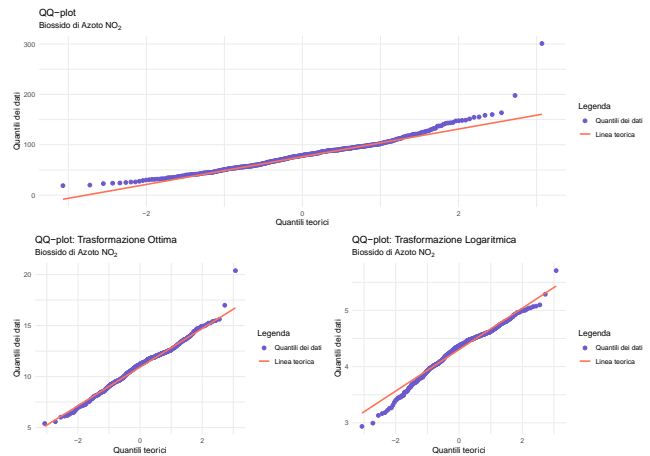
Uno strumento molto utile per risoluzione di molteplici problematiche, tra cui la gestione dei valori anomali, è la trasformata di Box-Cox. Si tratta di una funzione stabilizzatrice della varianza che, al variare del parametro  $\lambda$ , è in grado di ridurre l'eteroschedasticità e rendere la distribuzione della serie il più possibile prossima a quella di una Normale. Un modo per decretare il valore ottimo di  $\lambda$  è attraverso la massimizzazione della funzione di verosimiglianza di un modello di regressione lineare la cui risposta è regredita rispetto ad una costante. Questo perché, nella versione originale, la trasformata veniva utilizzata per ridurre il grado di non normalità degli errori di regressione e, utilizzare un modello vuoto, coincide con la riduzione del grado di non normalità della dipendente stessa, in questo caso la serie. Un secondo metodo per sancire il valore ottimo di  $\lambda$  è quello proposto da Guerrero nel 1993. L'idea consiste nel dividere la serie in H sottoserie e trovare il valore ideale per il parametro in grado di rendere minimo il coefficiente di variazione tra quest'ultime.

Per l'analisi è stato utilizzato sia il metodo della massimizzazione della verosimiglianza profilo che quello proposto da Guerrero. Rispettivamente, i  $\lambda$  ottimali sono risultati pari a 0.38 e 0, il che significa che entrambe le metodologie protendono per una contrazione della distribuzione. Nello specifico, il primo criterio suggerisce una trasformazione esponenziale, mentre il secondo una logaritmica, in linea con quanto ipotizzato nel paragrafo 2.2.



Le trasformazioni analizzate presentano caratteristiche distributive simili. Entrambe producono un'evidente asimmetria negativa, la quale risulta ben osservabile sia dagli istogrammi che dai Box-plot. Rispetto alla distribuzione normale, le mode si scostano sensibilmente verso destra, con un effetto particolarmente marcato nella trasformazione logaritmica. Questo comporta un incremento del numero di valori anomali, poiché le trasformazioni sono in grado di ridurre solo parzialmente la coda destra, mentre accentuano significativamente quella sinistra.

Osservando i QQ-plot, è possibile notare come la trasformazione logaritmica sia del tutto inappropriata, mentre i quantili della trasformazione esponenziale risultano ben aderenti a quelli teorici. Nonostante questo, non si evidenzia un significativo incremento delle performance distributive del fenomeno, il che potrebbe far ipotizzare che il problema principale sia la presenza dei due punti anomali sulla coda di destra. A fronte di un incremento prestazionale pressoché minimo, unito ad un'interpretazione alquanto complessa della serie, si è deciso di non compiere alcuna trasformazione.

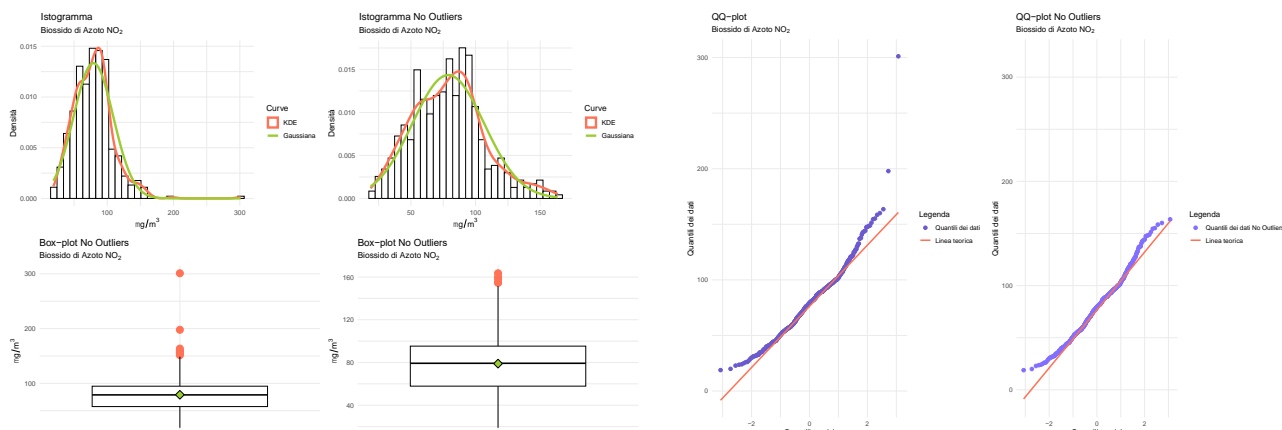


## 2.4 Outlier Detection

Con il termine outlier si indica un'osservazione che è incoerente o che si discosta particolarmente dal resto dei dati. Dalla definizione è immediato comprendere che per definire un valore atipico è necessario stabilire una soglia. Nel caso di osservazioni gaussiane, o approssimativamente tali, una delle regole più utilizzate è la  $3\sigma$  che sfrutta la relazione tra la deviazione standard dalla media e la probabilità nella normale. D'altra parte, la media, e di conseguenza la deviazione standard, è molto sensibile alla presenza di valori outlier e, per questo motivo, viene spesso sostituita dalla mediana. Quello che si ottiene è filtro di Hampel, una versione robusta della soglia  $3\sigma$  in cui il valore tipico è posto pari alla mediana e la dispersione è misurata dalla *median absolute deviation from median* (MADM). La regola rimane invariata, le osservazioni che si discostano tre volte la MADM sono considerate atipiche in quanto la loro probabilità di trovarsi a quella distanza è pari allo 0.3%.

Nel caso dell' $\text{NO}_2$ , il filtro ha rilevato la presenza di due valori anomali, in particolare quelli relativi all'ottobre 1987 e al gennaio 1990. Da una ricerca è emerso che il 15 ottobre del '87, nei pressi del comune di Milano, più precisamente a Conca di Crezzo, è avvenuto un'importante incidente aereo, il che potrebbe essere una delle cause della rilevazione atipica. Una volta identificati gli outlier, sono stati corretti mediante l'ausilio di un interpolatore lineare. Allo stesso tempo, sono stati imputati anche i quattro valori mancanti riferiti al periodo compreso tra il marzo e il giugno del 1988. In contemporanea, si è intrapresa un'analisi parallela nella quale i valori anomali non sono stati corretti per verificare se, in fase di decomposizione, queste rilevassero le medesime osservazioni.





I grafici hanno permesso di analizzare, in termini distributivi, l'impatto degli outlier rilevati mediante il filtro di Hampel. Sia dall'istogramma che dal Box-plot è possibile notare come il fenomeno corretto dalle rilevazioni atipiche risulti decisamente più Normale, la pesantezza della coda di destra si è totalmente riassorbita. A differenza delle trasformazioni di Box-Cox, la rettifica dei punti identificati non ha provocato alcuna asimmetria, media, moda e mediana del fenomeno sembrerebbero essere coincidenti. Anche attraverso il QQ-plot è possibile osservare come tale correzione abbia favorito una maggiore aderenza della coda di destra ai quantili teorici, come testimoniato dai p-value riportati nella tabella sottostante.

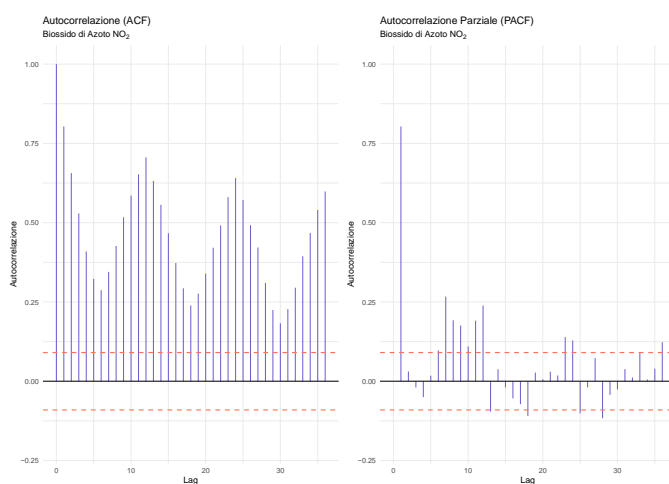
Dai test di normalità emerge come il test di KS non rifiuta più l'ipotesi nulla e anche i livelli di significatività osservati degli altri test sono decisamente aumentati.

Test di normalità	Statistica test	P-value
Jarque Bera Test	11,69	0,002901
Shapiro-Wilk test	0,9853	1,09E-04
Kolmogorov-Smirnov test	0,0458	0,2793

## 2.5 Analisi della Persistenza

La persistenza è una caratteristica delle serie storiche che si riferisce alla continuità degli shock nel tempo. La memoria, invece, descrive la capacità di una sequenza nel *ricordare* comportamenti passati nel lungo periodo. Per quanto possano sembrare concetti simili tra loro, in realtà sono molto differenti. La persistenza misura l'influenza che un evento o cambiamento improvviso ha sui valori futuri della serie, mentre la memoria riguarda, non solo i singoli shock, ma bensì l'effetto complessivo di una sequenza passata. Entrambi sono fenomeni che si quantificano in termini di autocorrelazioni, ma rispondono a quesiti differenti, rispettivamente "*Quanto dura l'effetto di un evento temporaneo sulla serie?*" e "*Quanto tempo occorre affinché un comportamento passato della serie smetta di influenzare i valori futuri?*". In sintesi, la persistenza è legata ad effetti di breve termine di uno shock singolo, mentre la memoria cattura una dipendenza di lungo termine rispetto al comportamento globale della serie.

Mediante l'analisi delle ACF è possibile notare come il fenomeno sia composto da correlazioni tra periodi e nei periodi. Le prime si riferiscono ad un legame tra osservazioni appartenenti al medesimo istante temporale ma riferite ad anni differenti, sintomo di un comportamento stagionale. Esaminandole, si denota un'elevata persistenza e una notevole memoria, il loro decadimento lento ed esponenziale riflette il fatto che vi sia una struttura autoregressiva a radici positive sottostante. Le seconde si riferiscono ad un legame tra osservazioni appartenenti alla medesima stagione ma a tempi diversi, ad





indicare come il livello del diossido del mese corrente influenzi anche quello successivo. Anche in questo caso si rilevano delle correlazioni importanti ed una marcata persistenza con un decadimento decisamente più rapido. Le PACF, al contrario, presentano un andamento pressoché nullo, se non per quanto riguarda alcuni lag, comportamento che conferma l'ipotesi autoregressiva della serie.

La stazionarietà è la proprietà di invarianza a traslazioni temporali delle serie, indica un comportamento omogeneo nel tempo, per cui ogni porzione della sequenza si comporta esattamente come le altre. La *mean reverting* è la caratteristica principale di un processo stazionario e si riferisce alla tendenza di questo a tornare verso un valore medio o di equilibrio. Si tratta, dunque, di una forza correttiva che permette alla serie, quando si allontana a causa di shock e cambiamenti repentini, di ritornare verso il proprio stato di equilibrio. Una serie storica a memoria infinita è una realizzazione di un processo stocastico il cui presente è influenzato da tutti i valori passati, indipendentemente da quanto lontani siano nel tempo. Si tratta di una dipendenza di lungo periodo che comporta una difficoltà della serie a *dimenticare* ciò che è avvenuto nel passato e tornare verso l'equilibrio di lungo periodo. Questo permette di comprendere come il concetto di memoria sia connesso a quello della stazionarietà e come un processo guidato da un trend stocastico, ovvero andamento di lungo periodo aleatorio, sia definito non stazionario.

I test di Dickey-Fuller permettono di verificare la non stazionarietà della serie contro l'ipotesi alternativa che questa risulti omogenea nel tempo. Ognuno è composto da due parti, la prima che testa la presenza di una radice unitaria all'interno del processo, mentre la seconda esamina la formulazione del modello utilizzato per analizzare l'ipotesi nulla. Nel caso specifico, avendo osservato una tendenza lineare decrescente all'interno della serie, si è deciso di applicare la specificazione che verifica l'ipotesi di stazionarietà attorno ad un trend deterministico. Il test di DF, mediante la statistica test  $\tau_3$ , verifica l'ipotesi di presenza di un trend stocastico, si suppone che il processo sottostante sia un Random Walk con un andamento lineare, contro l'alternativa di stazionarietà attorno ad una funzione deterministica del tempo. Conseguentemente,  $\phi_3$  e  $\phi_2$  esaminano le rispettive ipotesi congiunte, la serie presenta una radice unitaria e trend assente, la serie non è stazionaria e sia la tendenza che il drift sono nulli.

Dai risultati riportati in tabella è possibile rifiutare, per ogni livello di significatività, tutte le ipotesi nulle affermando come il processo risulti stazionario attorno ad una tendenza lineare avente termine costante diverso da zero. Successivamente, mediante una regressione lineare, è stato possibile detrendizzare la serie rimuovendo sia l'andamento lineare decrescente che il drift identificati dal test.

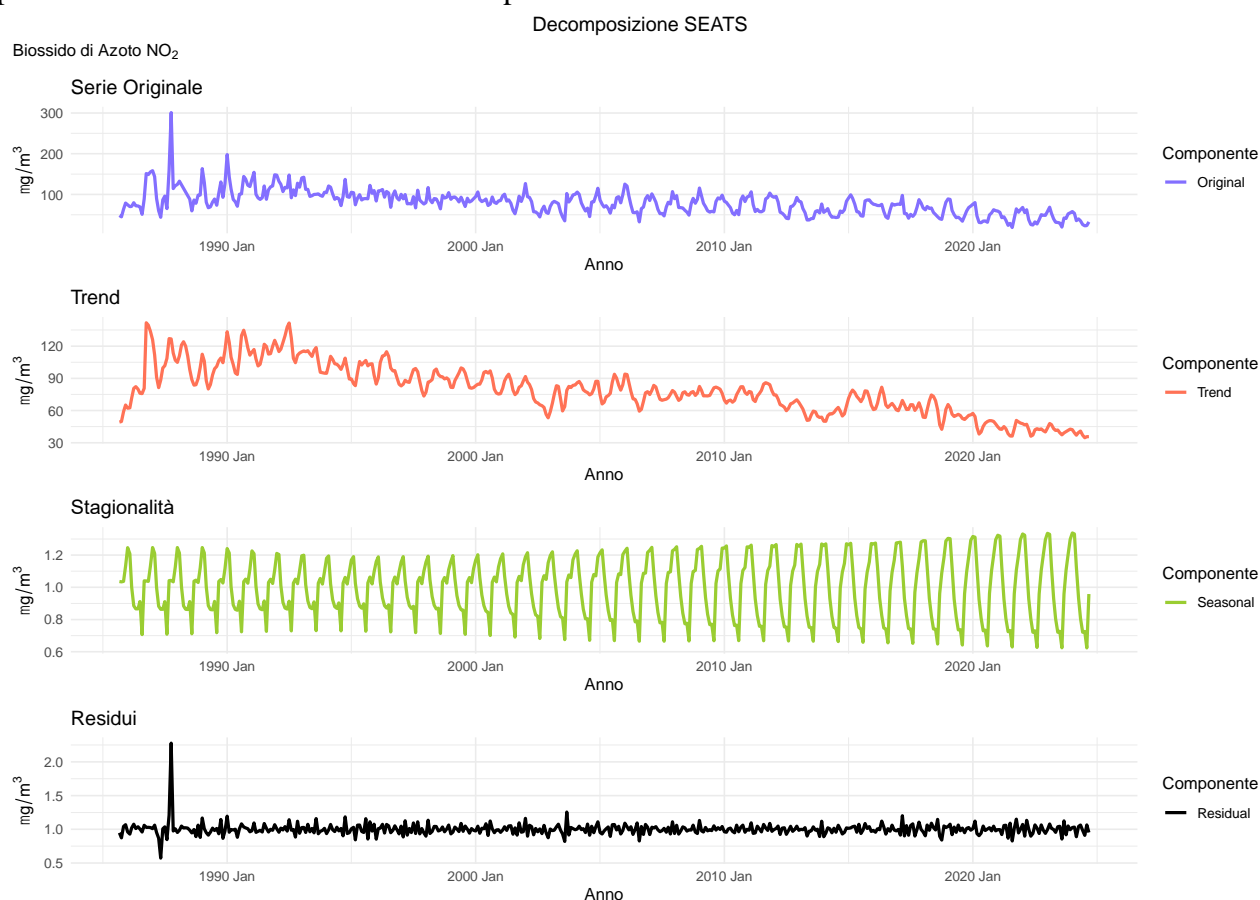
Da un'analisi preliminare della serie residuale emerge come questa non oscilli in modo totalmente casuale attorno al valore nullo. Il motivo potrebbe essere che la componente di lungo periodo non sia di tipo lineare bensì di un ordine polinomiale superiore, che in questo caso sembrerebbe essere il quarto. In generale, non si identifica un cattivo adattamento, se non per la parte iniziale della finestra temporale, il che giustifica il risultato emesso dal test di Dickey Fuller. Successivamente, le decomposizioni consentiranno di comprendere al meglio l'andamento di lungo periodo della serie.

Coefficients	Estimate	Std. Error	t value	Pr(> t )	Signif.
<b>Drift</b>	44,0639	4,5679	9,6460	< 2e-16	***
<b>z.lag.1</b>	-0,3898	0,0388	-10,0410	< 2e-16	***
<b>Trend</b>	-0,0564	0,0076	-7,4510	0,0000	***
$\tau_3$	$\tau_{3, 0.99}$	$\tau_{3, 0.95}$	$\tau_{3, 0.90}$		
-10,0406	-3,98	-3,42	-3,13		
$\phi_2$	$\phi_{2,0.99}$	$\phi_{2,0.95}$	$\phi_{2,0.90}$		
33,6804	6,15	4,71	4,05		
$\phi_3$	$\phi_{3,0.99}$	$\phi_{3,0.95}$	$\phi_{3,0.90}$		
50,5195	8,34	6,3	5,36		



## 2.6 Analisi delle componenti non osservabili

Le componenti non osservabili rappresentano strutture sottostanti alla serie storica che ne influenzano il comportamento. Queste risultano fondamentali per comprendere la dinamica del processo ed evidenziare i principali driver del fenomeno. Il trend rivela l'andamento di fondo della serie, la stagionalità le variazioni periodiche, la componente residuale la casualità delle oscillazioni e l'aggregazione di queste produce la sequenza osservata. Le decomposizioni consentono di identificare ed analizzare le componenti non osservabili delle serie storiche e, tra le più note, si annoverano la Classica, l' $X_{11}$ -ARIMA, la TRAMO-SEATS e la STL. Una decomposizione è definita ottimale quando la componente residuale ottenuta, depurando la serie dal trend e della stagionalità, è considerabile una realizzazione di un processo White Noise. Un WN è un fenomeno completamente aleatorio in cui ogni osservazione è indipendente dalle altre e perciò assolutamente non stimabile e prevedibile. Nel caso del  $\text{NO}_2$  la decomposizione identificata come ottimale è la TRAMO-SEATS.



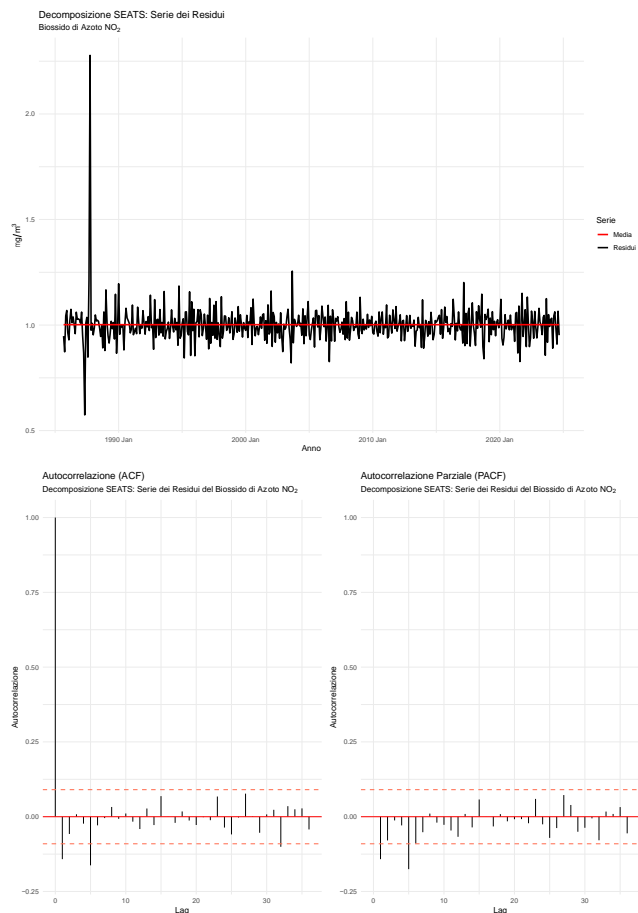
Alla decomposizione è stata data in pasto la serie originale poiché, essendo robusta alla presenza di outlier, l'obiettivo era esaminare se, rilevasse come atipiche, le medesime osservazioni identificate dal filtro di Hampel. La scomposizione ha determinato come outlier additivi le rilevazioni del maggio ed ottobre 1987, non segnalando come problematica quella del gennaio 1990. Dall'analisi delle componenti emerge come il trend sia decrescente ma non lineare, bensì sembrerebbe avere un andamento polinomiale di quarto grado, in linea con quanto evidenziato dalla detrendizzazione a seguito del test di Dickey Fuller. La stagionalità risulta essere di tipo moltiplicativo, nel tempo tende ad evolversi sia per forma che per variabilità diventando sempre più regolare ed oscillatoria. Inoltre, l'accostamento delle serie evidenzia un andamento opposto tra l'equilibrio di lungo periodo e le fluttuazioni stagionali; le emissioni, da un punto di vista dell'andamento medio tendono a diminuire, mentre le oscillazioni stagionali continuano ad amplificarsi. Negli ultimi anni si osserva che i mesi invernali, come dicembre, gennaio e febbraio, registrano concentrazioni più elevate di  $\text{NO}_2$ , mentre gli estivi, giugno, luglio e agosto, presentano livelli significativamente più bassi, risultando più "green". Questo fenomeno potrebbe essere attribuito a un maggiore utilizzo di combustibili fossili per

il riscaldamento domestico e a condizioni atmosferiche meno favorevoli alla dispersione degli inquinanti: in inverno, infatti, i venti sono più deboli e le precipitazioni meno frequenti. Tuttavia, esistono anche spiegazioni più complesse. Durante l'inverno, il suolo tende a essere più freddo rispetto agli strati superiori dell'atmosfera, causando il fenomeno dell'inversione termica. In queste condizioni, uno strato di aria calda si forma sopra quello freddo ed inquinato vicino al suolo agendo come una barriera che impedisce la dispersione degli inquinanti. Inoltre, l' $\text{NO}_2$  tende a essere scomposto dalla luce solare mediante reazioni fotochimiche producendo ozono troposferico e particolato. Ad ogni modo, durante l'inverno, le giornate più corte e la minore intensità dell'irraggiamento solare riducono l'efficacia di questa decomposizione, aumentando la persistenza dell' $\text{NO}_2$  nell'atmosfera.

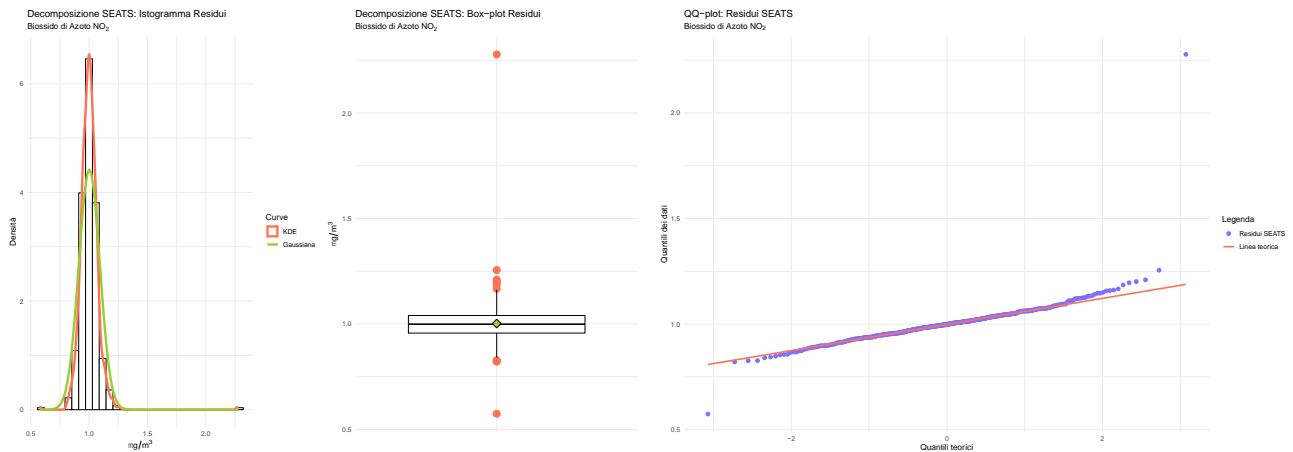
Un focus sulla componente residuale ne evidenzia un'oscillazione stazionaria attorno al valore unitario, dovuto probabilmente ad una decomposizione moltiplicativa data la struttura della stagionalità. L'ordine di variabilità della serie risulta contenuto con un andamento piuttosto casuale nonostante qualche oscillazione anomala. Malgrado l'identificazione degli outlier, la scomposizione non è stata in grado di modellarli completamente, infatti si rilevano due picchi atipici in corrispondenza delle date rilevate.

Dai correlogrammi non si evidenziano particolari pattern o valori che fuoriescano significativamente dalle bande di confidenza, il che conferma l'ipotesi di casualità delle osservazioni della serie residuale. Inoltre, non si identificano evidenti differenze tra ACF e PACF, sintomo che la dipendenza temporale sia stata completamente spiegata dalle componenti estratte. In conclusione, per verificare se la componente accidentale fosse una realizzazione di un processo White Noise, si è effettuata un'analisi mediante test di Portmanteau. In tabella vengono riportate le versioni di Ljung-Box (LB) e di Box-Pierce (BP), le cui differenze consistono nel modo in cui, ad ogni osservazione, viene assegnato un peso. Nonostante ciò, i test hanno come scopo quello di verificare se le autocorrelazioni dei residui siano congiuntamente nulle per una serie di lag specificati, contro l'alternativa che almeno una differisca in modo significativo. I p-value riportati consentono di affermare come le osservazioni della serie siano una realizzazione di un processo WN.

Da un punto di vista distributivo, la componente irregolare presenta un andamento prossimo a quello di una Normale. Le stime evidenziano una curva gaussiana leptocurtica in quanto si figurano code molto leggere ed una massa di densità molto accentuata in prossimità della moda. Tramite Box-plot e QQ-plot è possibile notare gli outlier precedentemente identificati e modellati solo in parte. Nonostante questi, si rileva un ottimo adattamento dei quantili empirici a quelli teorici che permette di affermare come la serie possa essere considerata una realizzazione di un processo White Noise gaussiano.



Lags	LB p-value	BP p-value
1	0,0021	0,0021
2	0,004	0,0041
3	0,0112	0,0116
6	0,0005	0,0005
12	0,0112	0,0124
24	0,1105	0,1248
36	0,1125	0,1453



## 2.7 Modellistica

I modelli ARIMA (AutoRegressive Integrated Moving Average) e SARIMA (Seasonal AutoRegressive Integrated Moving Average) sono tecniche che mirano a catturare le componenti strutturali delle serie storiche al fine di prevederne l'andamento futuro. A differenza delle decomposizioni, questi modelli costruiscono una rappresentazione matematica delle diverse componenti senza isolarle esplicitamente.

Un modello ARIMA si compone di tre elementi principali, come indicato dal suo acronimo:

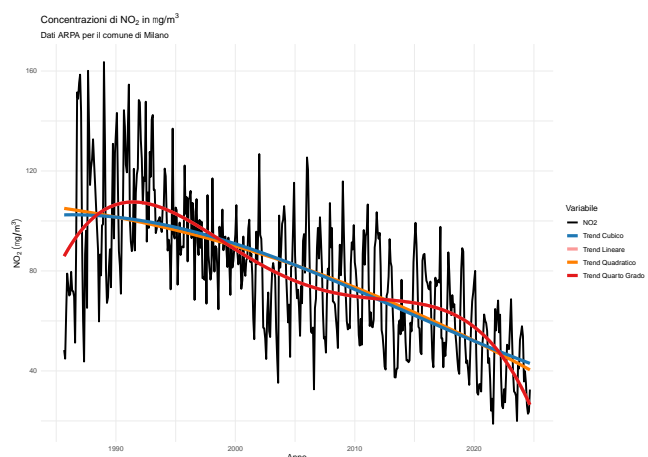
- **AR (AutoRegressive):** descrive la relazione tra i valori passati della serie e quelli correnti
- **I (Integrated):** si riferisce alla differenziazione della serie necessaria per rendere stazionarie le sequenze caratterizzate da un trend stocastico.
- **MA (Moving Average):** rappresenta le relazioni tra gli shock aleatori passati e i valori attuali della serie.

I modelli SARIMA estendono gli ARIMA includendo la capacità di gestire la stagionalità delle serie temporali. Oltre le componenti precedentemente descritte, aggiungono parametri specifici per la rappresentazione della componente periodica. Nel presente studio sono stati adottati entrambi i modelli: l'ARIMA attraverso un approccio multistadio e il SARIMA in modo diretto. I risultati ottenuti dai due metodi sono stati confrontati per determinare quale fosse più efficace nel descrivere il fenomeno analizzato.

### 2.7.1 ARIMA: modellistica two steps

La procedura si articola in due stadi, il primo consiste nella detrendizzazione e destagionalizzazione della serie mediante le informazioni ottenute nelle fasi precedenti dell'analisi, il secondo identifica il modello ARIMA che meglio si adatta alla sequenza residuale ottenuta.

Dall'analisi degli adattamenti dei trend polinomiali stimati non emergono sostanziali differenze tra un andamento lineare, quadratico o cubico. Al contrario, la tendenza di quarto grado sembrerebbe cogliere perfettamente il comportamento medio della serie. Tale impressione viene confermata dai criteri d'informazione e dagli indici di adattamento cross-validati, riportati nella tabella in alto a sinistra della pagina successiva. Infatti, la tendenza polinomiale è colei che meglio gestisce il trade off tra accuratezza e complessità adattandosi bene ai dati senza perdere la sua capacità di generalizzazione.

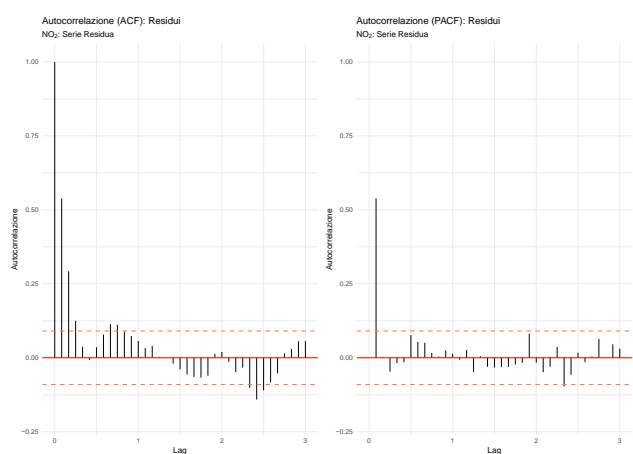


Modello	CV	AIC	AIC <sub>C</sub>	BIC	R <sup>2</sup> <sub>ADJ</sub>
Lineare	426,2380	2.841,0261	2.841,0777	2.853,4779	0,4518
Quadratico	420,8476	2.834,3034	2.834,3896	2.850,9058	0,4607
Cubico	422,1366	2.835,1898	2.835,3194	2.855,9428	0,4608
Quarto Grado	391,1944	2.799,8103	2.799,9921	2.824,7139	0,5011

Modello	CV	AIC	AIC <sub>C</sub>	BIC	R <sup>2</sup> <sub>ADJ</sub>
K1	241,3560	2.575,0447	2.575,1309	2.591,6471	0,3728
K2	242,7754	2.577,7639	2.577,9457	2.602,6675	0,3718
K3	233,7217	2.559,8886	2.560,2017	2.593,0934	0,3978
K4	231,5693	2.555,4819	2.555,9622	2.596,9879	0,4059
K5	227,1067	2.546,2719	2.546,9561	2.596,0791	0,4199
K6	227,4976	2.547,0307	2.547,8307	2.600,9886	0,4202

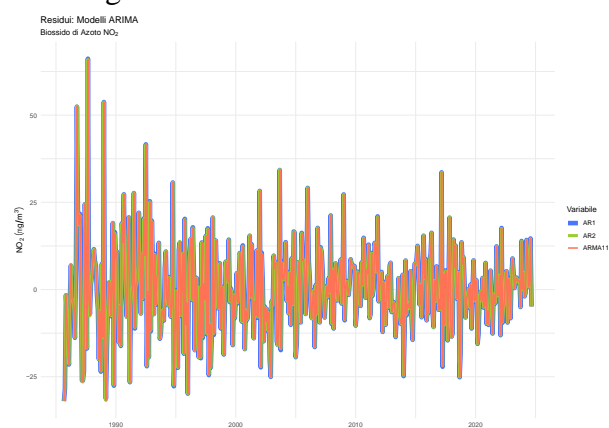
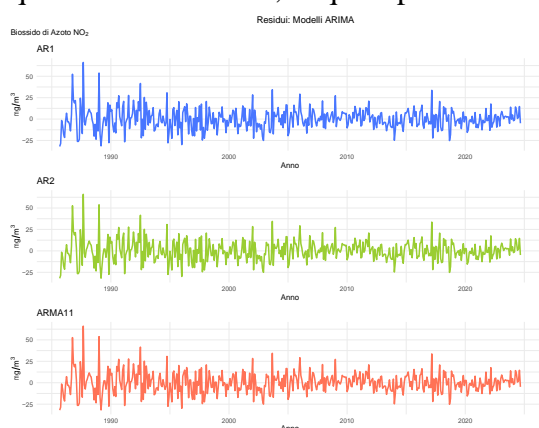
Successivamente, è stata effettuata una regressione armonica al fine di destagionalizzare la serie ottenuta dalla detrendizzazione. Dalla TRAMO-SEATS è emerso che la componente periodica è di tipo moltiplicativo, mentre le armoniche sono in grado di modellare solamente stagionalità costanti nel tempo. Tuttavia, non risulta essere un problema in quanto alla serie destagionalizzata verrà applicato un filtro ARMA in grado di catturare la struttura di dipendenza residua. Nella tabella in alto a destra vengono riportati i criteri che hanno consentito di selezionare il numero ottimale di armoniche per la destagionalizzazione. La specificazione, che per adattamento e generalizzazione è risultata essere la migliore, è stata quella composta da cinque armoniche, anche se le performance delle formulazioni sono tutte molto simili tra loro.

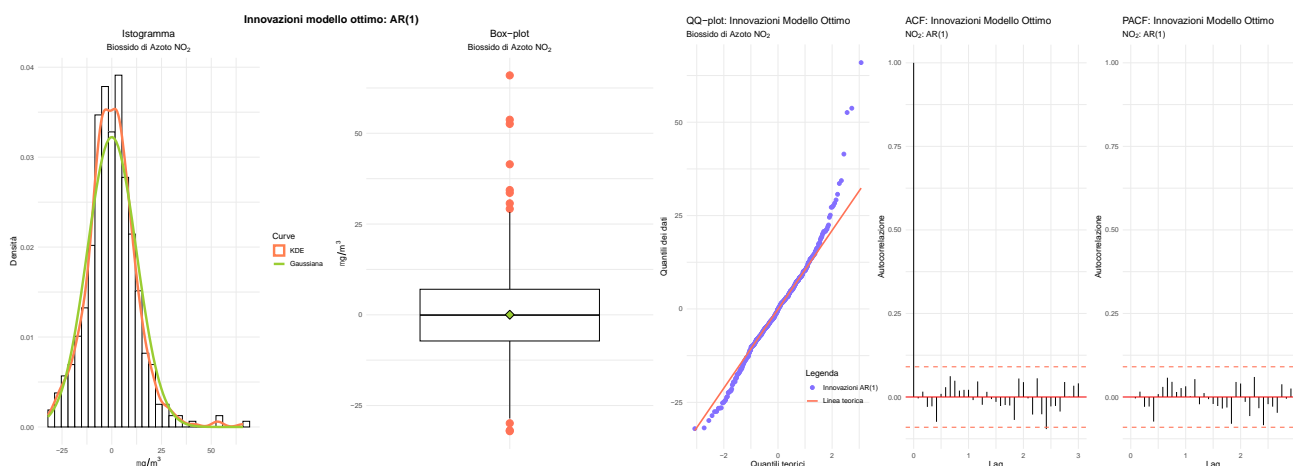
Le ACF della serie destagionalizzata confermano quanto riportato nel paragrafo precedente. Essendo la stagionalità moltiplicativa, l'utilizzo delle armoniche non ha permesso la completa rimozione della componente, il che si riflette sull'andamento sinusoidale delle autocorrelazioni. Il tutto, unito al comportamento delle PACF, porta ad affermare che il processo sottostante sia un AR (1) con radice positiva. Per confermare quanto dedotto dall'analisi e selezionare il modello ARMA che meglio si adatti alla serie residuale, sono state comparate diverse specificazioni. Tra le riportate in tabella, si sottolinea la presenza del modello nullo al fine di testare l'ipotesi che essa stessa sia una realizzazione di un processo White Noise (WN).



Modello	AIC	AIC <sub>C</sub>	BIC	R <sup>2</sup> <sub>ADJ</sub>	RMSE	MAE	ACF1
WN	3857,24	3857,26	3865,54	0	14,72	11,15	0,54
AR (1)	3695,99	3696,04	3708,44	0,29	12,36	9,26	-0,01
AR (2)	3697,92	3698,01	3714,52	0,29	12,36	9,25	0
MA (1)	3736,34	3736,39	3748,79	0,23	12,91	9,69	0,12
MA (2)	3697,92	3698,01	3714,52	0,29	12,47	9,38	0,03
ARMA (1,1)	3697,93	3698,02	3714,53	0,29	12,36	9,25	-0,01

I risultati confermano che il modello che meglio si adatta alla serie è un AR (1) con radice positiva pari a 0.55. Tuttavia, le performance delle formulazioni AR (2) e ARMA (1,1) risultano molto simili a quelle della vincente, la quasi perfetta sovrapposizione degli andamenti stimati ne è testimone.





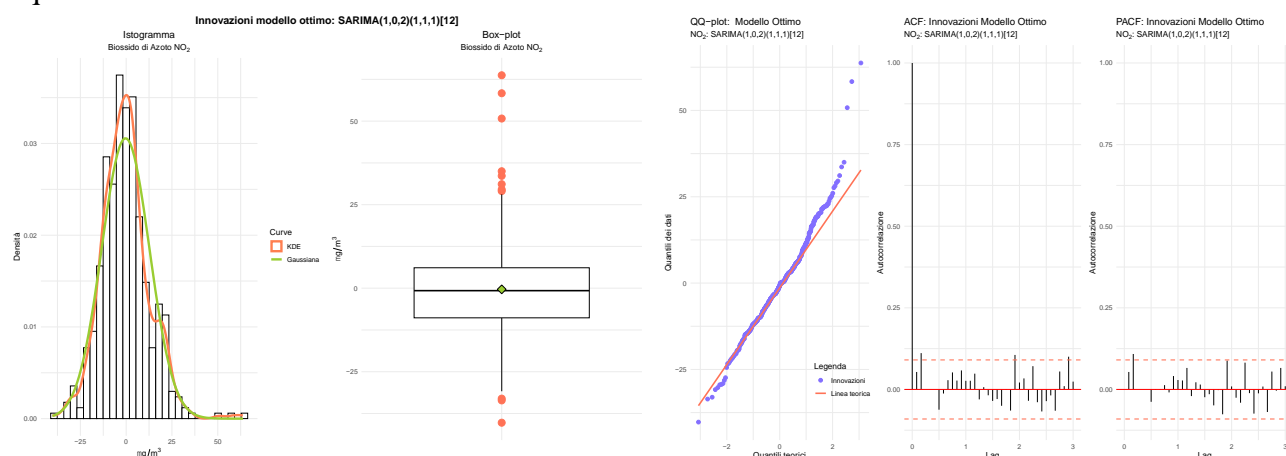
In conclusione, l'analisi dei residui AR (1) evidenzia come questi si distribuiscano approssimativamente come una Normale. Le stime indicano complessivamente una buona sovrapposizione della curva empirica a quella teorica, si nota solamente una maggiore densità in prossimità della moda e la presenza di alcuni outlier sulla coda di destra. I correlogrammi, non rilevando alcuna differenza tra le ACF e PACF e nessuna autocorrelazione significativa, permettono di affermare che si tratti di una realizzazione di un processo WN, la procedura multistadio è stata in grado di catturare completamente la dipendenza temporale del fenomeno.

## 2.7.2 SARIMA

A differenza dell'ARIMA two steps, la procedura consente di modellare l'andamento della serie in modo diretto senza la necessità di applicare una detrendizzazione e/o destagionalizzazione. Il modello ottimale identificato risulta essere il SARIMA (1,0,2)(1,1,1)[12], in tabella si riportano i dettagli.

SARIMA	AR (1)	MA (1)	MA (2)	SAR (1)	SMA (1)
Coefficiente	0,9784	-0,4305	-0,2780	-0,0278	-0,8827
Std. Error	0,0198	0,0514	0,0525	0,0620	0,0484

Il modello presenta valori di AIC, AIC<sub>c</sub> e BIC rispettivamente pari a 3685.63, 3685.82 e 3710.38 in linea con quanto performato dall'ARIMA two steps, sintomo che le procedure siano pressoché equivalenti.

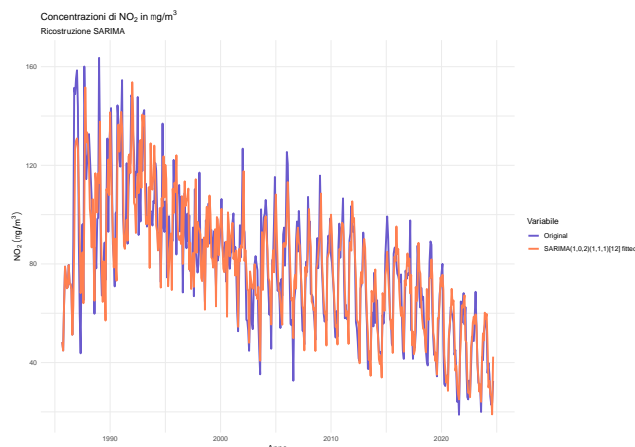
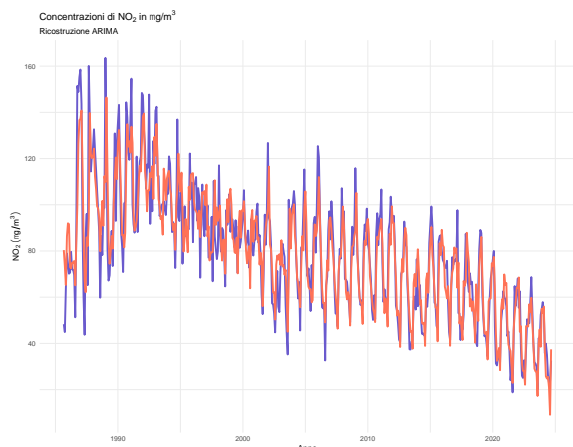


La serie residuale del modello SARIMA presenta un andamento prossimo ad una Normale nonostante la leggera asimmetria distributiva e la presenza di diversi punti anomali. La coda di destra ne è infucata particolarmente della problematica, si rileva un discreto discostamento tra i quantili empirici e la tendenza teorica. Inoltre, i correlogrammi, non evidenziando pattern stagionali e autocorrelazioni statisticamente significative, consentono di affermare che la componente accidentale sia una realizzazione di un processo WN gaussiano.



### 2.7.3 Confronto tra modelli

La ricostruzione della serie originale ha permesso di confrontare le performance delle metodologie applicate. Nel caso SARIMA, i valori fittati dal modello coincidono con la riproduzione della serie, mentre nell'ARIMA, i valori stimati, uniti a quelli previsti dal trend polinomiale e dalla stagionalità armonica, hanno consentito la ricostruzione della sequenza temporale.



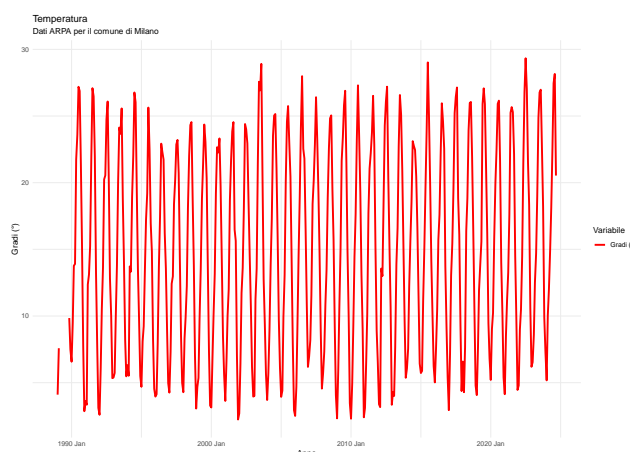
Da un'analisi grafica non si evidenziano particolari differenze negli adattamenti dei due modelli. Entrambe le formulazioni sono in grado di cogliere perfettamente l'andamento delle emissioni nell'ultima parte della finestra temporale, mentre, per quanto riguarda il quinquennio 1995 – 2000, il SARIMA, plot di destra, riporta delle difficoltà maggiori nella ricostruzione della sequenza. Al contrario, come osservabile nel grafico a sinistra, l'ARIMA riporta una capacità inferiore nel cogliere le oscillazioni repentine, ma nel complesso, ambedue le specificazioni posseggono delle ottime performance.

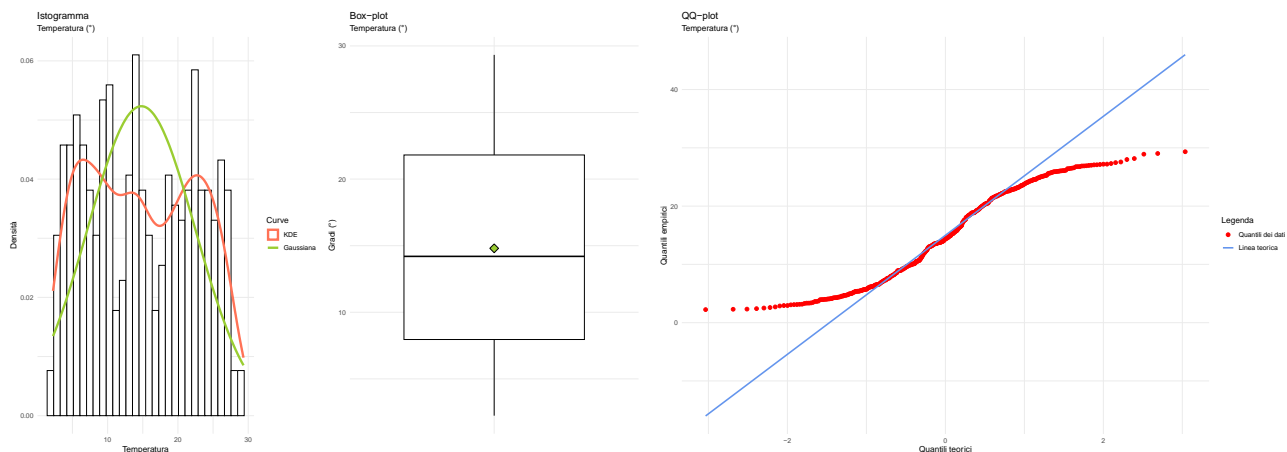
In conclusione, come mostrato in tabella, i modelli presentano delle performance pressoché equivalenti. Si osserva un lieve vantaggio a favore della metodologia multistadio, ma la differenza risulta trascurabile.

Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	12,37	13,05
MAE	9,25	9,82
R <sub>2</sub>	0,8	0,78

## 3. Temperatura

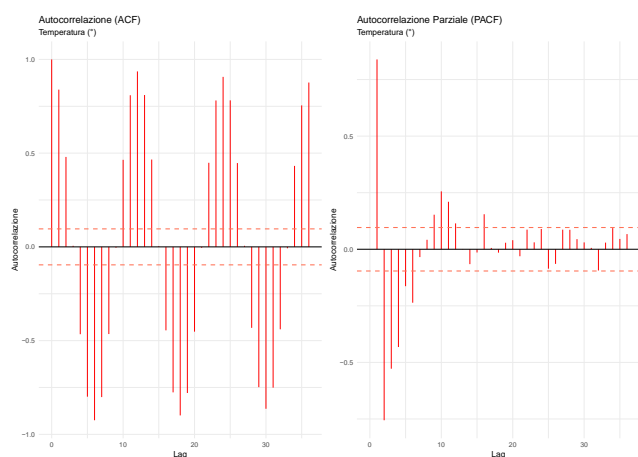
La temperatura influisce notevolmente sulle concentrazioni e sulle reazioni che coinvolgono l'NO<sub>2</sub>. A temperature elevate, si rileva una riduzione dell'inquinante, mentre in inverno, le concentrazioni aumentano a causa dell'inversione termica che ne limita la dispersione. La serie sembra presentare un andamento stazionario attorno ad una costante, caratterizzato da una marcata componente stagionale. Quest'ultima appare di tipo additivo, poiché le oscillazioni e il comportamento complessivo della serie si mantengono costanti nel tempo e, in aggiunta, non intaccate da valori anomali. Al contrario, la sequenza presenta una serie di valori mancanti nel primo anno d'osservazione. Per non creare problemi d'imputazione, si è optato per considerare il fenomeno dal primo gennaio 1990.





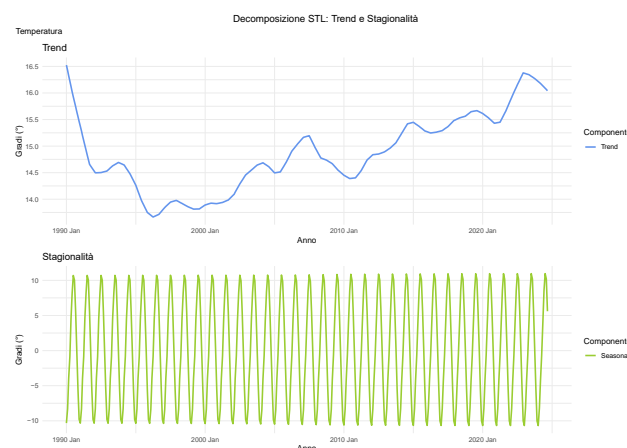
Le temperature riportano una distribuzione multimodale che si discosta totalmente dall'assunto di normalità e la causa potrebbe essere la forte componente stagionale rilevata precedentemente. Il fenomeno è caratterizzato da una leggera asimmetria positiva non affetta da punti outlier, né il box plot né il QQ-plot sembrerebbero rilevarne. Nel caso di una distribuzione multimodale, la trasformazione di Box-Cox risulta essere inefficace nell'intento di normalizzazione. La polimodalità riflette la presenza di più popolazioni o sottogruppi all'interno dei dati che, un solo parametro  $\lambda$ , non sarebbe in grado di gestire. Per tale ragione, alla serie non è stata applicata alcuna funzione stabilizzatrice. Infine, mediante l'ausilio del filtro di Hampel, è stata condotta un'analisi dei punti anomali che non ha riscontrato outlier.

I correlogrammi riportano un andamento fortemente influenzato dalla componente stagionale, il comportamento sinusoidale e l'elevata persistenza annuale ne è prova. Dalle ACF emerge, oltre una relazione tra periodi, anche un buon legame nei periodi, sintomo che le temperature, all'interno della medesima stagione, si influenzino tra loro. Analizzando le PACF, si rileva come, dopo i primi 12 lag, le correlazioni si annullino, sintomo di un processo AR (1) stagionale sottostante. I test di Dickey-Fuller confermano la stazionarietà del fenomeno rispetto ad un drift che, mediante regressione lineare, è stato identificato esser pari a 14.87 gradi celsius.



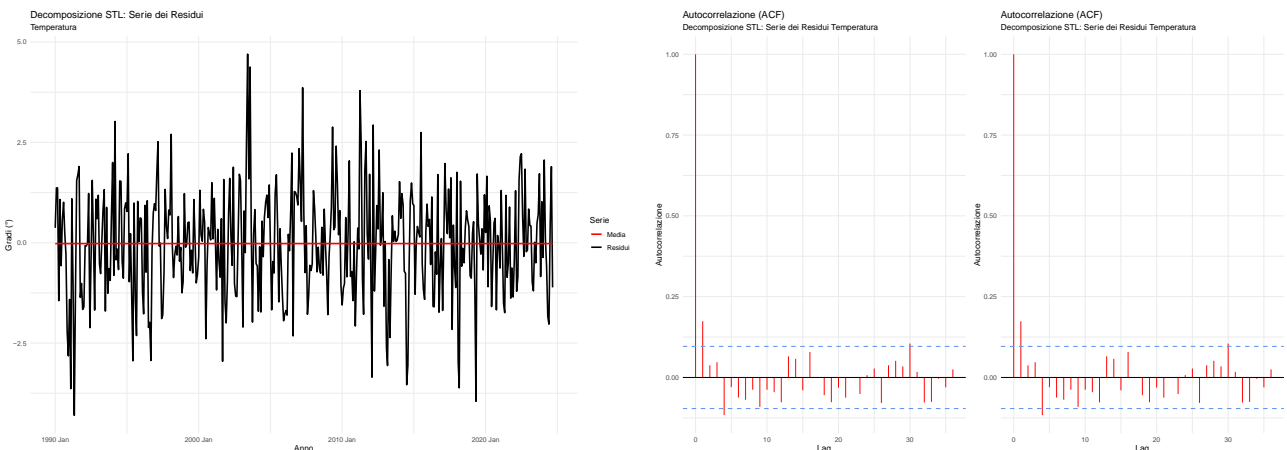
### 3.3 Analisi delle componenti non osservabili

La Seasonal and Trend decomposition using Loess (STL) ha consentito di estrarre in modo ottimale le componenti non osservabili della serie delle temperature. Come suggerisce il nome, la decomposizione si basa sull'utilizzo delle loess per modellare le variabili latenti. In particolare, per il trend e la stagionalità sono state impiegate finestre mobili con ampiezze rispettive di 42 e 34 lag, tunate in modo opportuno. L'analisi delle componenti evidenzia, per l'andamento di lungo periodo, un marcato crollo verticale all'inizio del periodo d'osservazione. Successivamente, a partire dal 1996, si rileva una ripresa lineare costante che lo riporta al livello originario. Per quanto riguarda la stagionalità, la decomposizione





mostra come questa rimanga pressoché invariante nel tempo. Le temperature seguono un andamento molto regolare, con valori massimi registrati a luglio e minimi a gennaio.



La componente residuale presenta un andamento casuale con un range di variazione contenuto nonostante la presenza di alcune osservazioni anomale. Dai correlogrammi non emergono differenze significative tra ACF e PACF, a testimonianza che l'intera la dipendenza temporale del fenomeno è stata estratta. La serie è considerabile come una realizzazione di un processo WN e, da un'analisi effettuata sulla forza delle componenti, si è rilevato un valore per la stagionalità pari a 0.96 e per il trend di 0.22, a rimarcare l'importanza della componente periodica.

### 3.4 Modellistica

In una prima fase è stata implementata una modellazione ARIMA two steps. Un'analisi cross-validata ha permesso di decretare come ottimo il polinomio lineare e tre il numero ideale di armoniche. A seguito della detrendizzazione e destagionalizzazione della serie, è stato eseguito un confronto tra specificazioni ARIMA che ha determinato i risultati in tabella.

Modello	AIC	AIC <sub>c</sub>	BIC	R <sup>2</sup> <sub>Adj</sub>	RMSE	MAE	ACF1
WN	1512,97	1513,00	1521,04	0,00	1,48	1,15	0,26
AR (1)	1486,08	1486,14	1498,18	0,06	1,43	1,13	-0,03
AR (2)	1483,42	1483,51	1499,55	0,07	1,42	1,12	-0,01
MA (1)	1491,21	1491,27	1503,31	0,05	1,44	1,13	0,02
MA (2)	1488,06	1488,16	1504,20	0,06	1,43	1,12	0,02
ARMA (1,1)	1481,58	1481,68	1497,71	0,08	1,42	1,12	0,00
ARMA (3,1)	1469,47	1469,48	1473,50	0,09	1,41	1,12	0,00

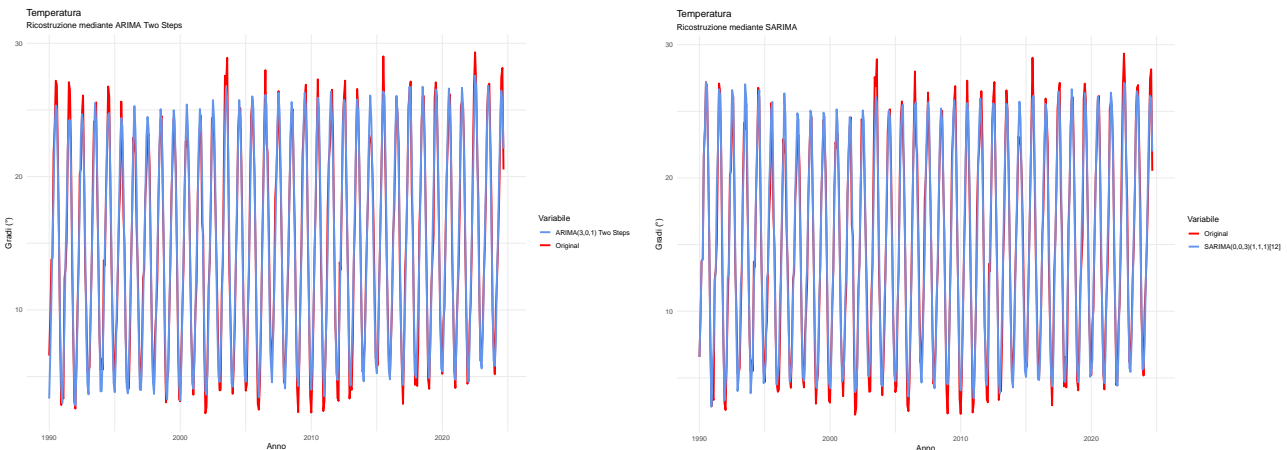
ARIMA	AR (1)	AR (2)	AR (3)	MA (1)
Coefficients	-0,4037	0,2312	0,1765	0,6395
Std. Error	0,1508	0,0637	0,0493	0,1475

Il modello che meglio si adatta alla serie, identificato dalla procedura automatica di Hyndman e Khandakar, è l'ARMA (3,1) aventi performance nettamente migliori di tutte le altre specificazioni.

Invece, per quanto riguarda la formulazione diretta, la specificazione che meglio si è adattata ai dati è risultata essere il SARIMA (0,0,3) (1,1,1) [12].

SARIMA	MA (1)	MA (2)	MA (3)	SAR (1)	SMA (1)
Coefficients	0,2701	0,1175	0,1668	-0,0673	-0,8672
Std. Error	0,0498	0,0526	0,0487	0,0639	0,0521

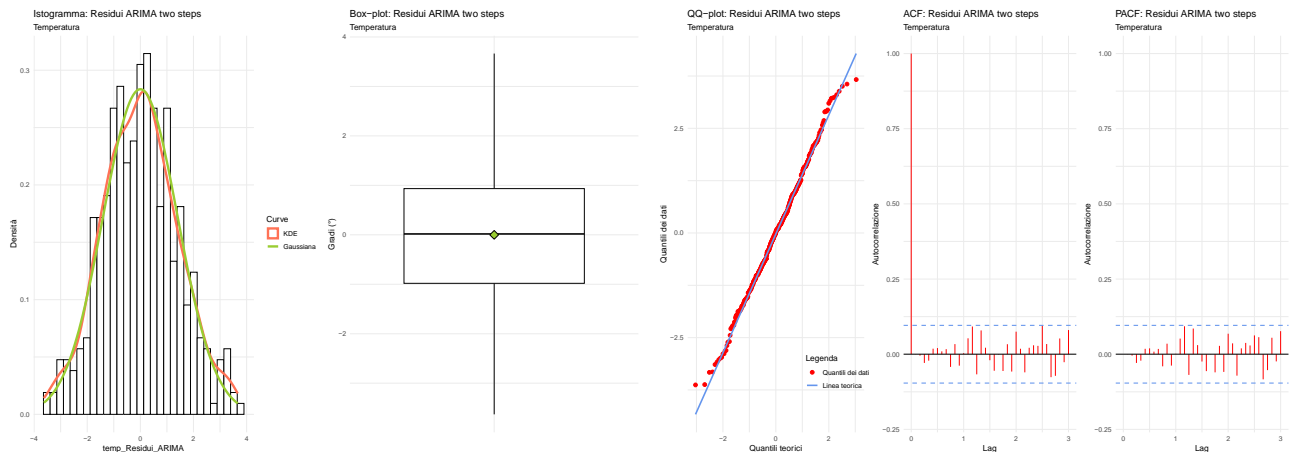
Le ricostruzioni della serie hanno permesso di confrontare le performance delle due metodologie



Il fitting dei due modelli è pressoché identico, entrambe le specificazioni ricostruiscono la serie in modo eccellente cogliendo al meglio le oscillazioni periodiche della sequenza temporale.

Gli indici riportati in tabella testimoniano come gli adattamenti dei due modelli siano assolutamente sovrapponibili. Per tale ragione si è deciso di riportare solamente le diagnostiche dei residui ARIMA per non creare ripetitività.

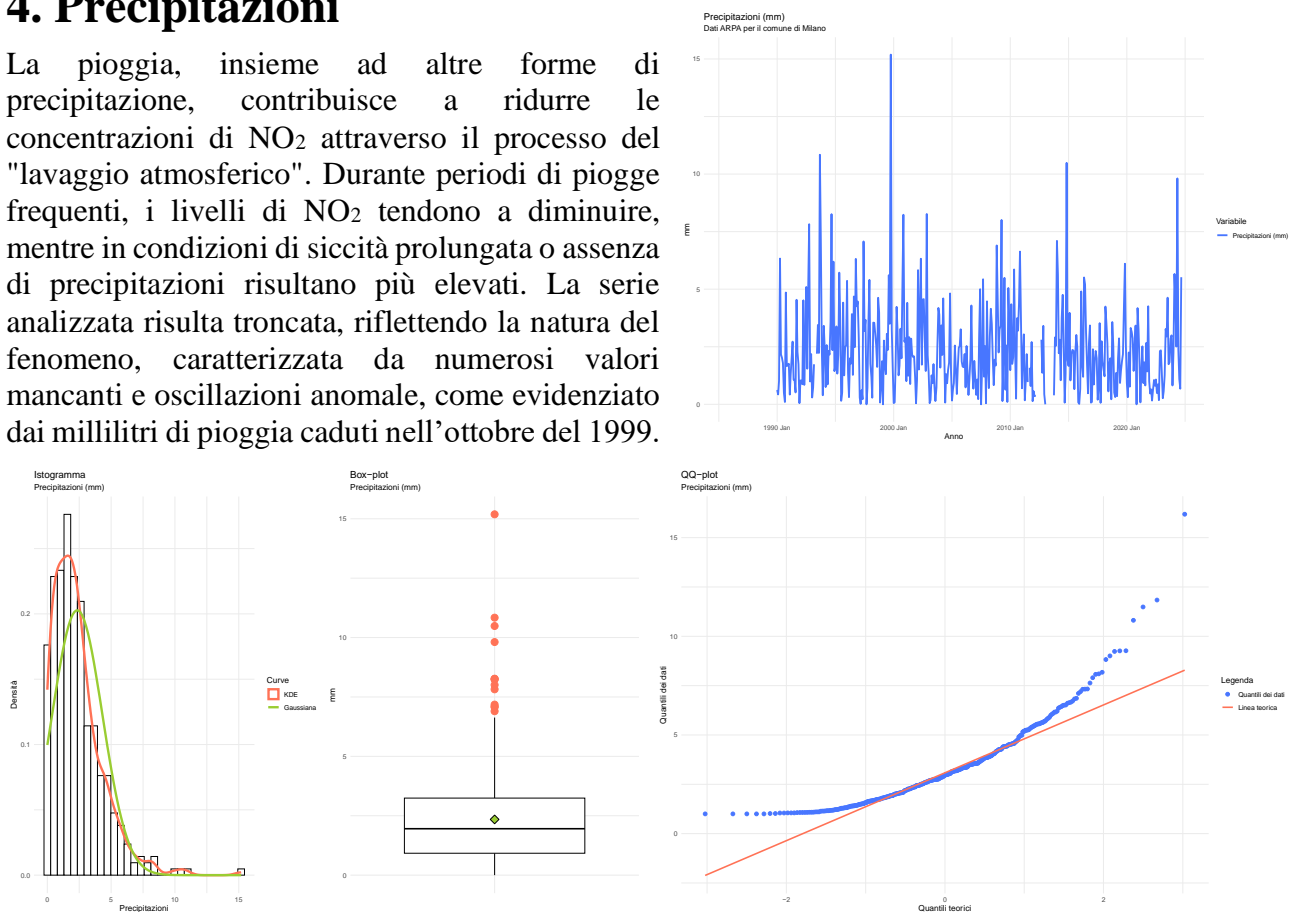
Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	1,4072	1,4237
MAE	1,1204	1,1181
R <sub>2</sub>	0,9659	0,9651



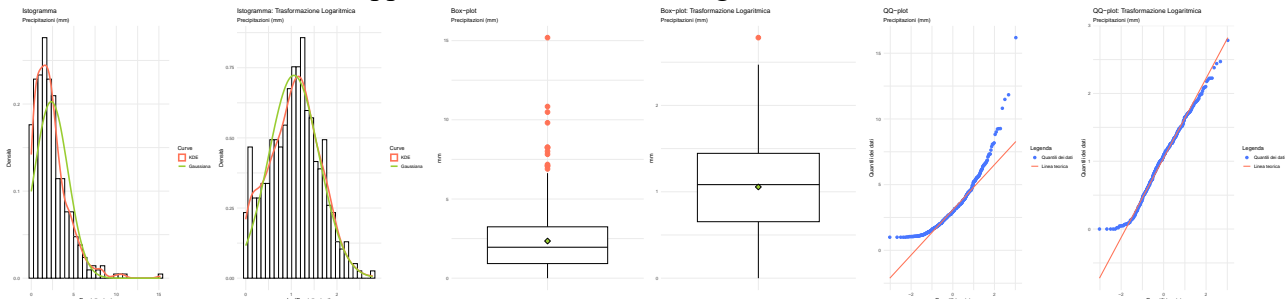
Dalla distribuzione dei residui e dai correlogrammi, è possibile osservare come questi si distribuiscano in modo *iid* da una Normale avente media nulla. In conclusione, entrambi i modelli sono stati in grado di spiegare completamente la dipendenza temporale della serie, la componente residuale è considerabile una realizzazione di un processo WN.

## 4. Precipitazioni

La pioggia, insieme ad altre forme di precipitazione, contribuisce a ridurre le concentrazioni di NO<sub>2</sub> attraverso il processo del "lavaggio atmosferico". Durante periodi di piogge frequenti, i livelli di NO<sub>2</sub> tendono a diminuire, mentre in condizioni di siccità prolungata o assenza di precipitazioni risultano più elevati. La serie analizzata risulta troncata, riflettendo la natura del fenomeno, caratterizzata da numerosi valori mancanti e oscillazioni anomale, come evidenziato dai millimetri di pioggia caduti nell'ottobre del 1999.

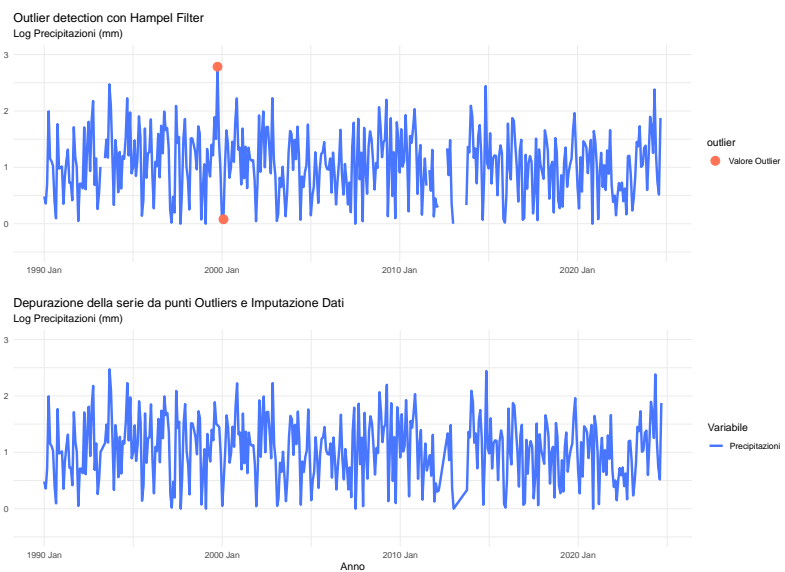


Il fenomeno, essendo troncato, risulta avere una maggiore massa di probabilità in prossimità del proprio valore limite. Questo ha comportato una traslazione della distribuzione verso l'estremità sinistra discostandola dall'assunto di Normalità. Inoltre, la presenza di alcuni valori outlier ha contribuito all'appesantimento della coda distributiva di destra. In queste situazioni, l'ausilio di una trasformazione logaritmica potrebbe essere la soluzione normalizzatrice dei dati. L'analisi di Box-Cox ha confermato quanto ipotizzato, sia il metodo Guerrero, che quello basato sulla verosimiglianza, sono risultati concordi nell'applicare una trasformata logaritmica.

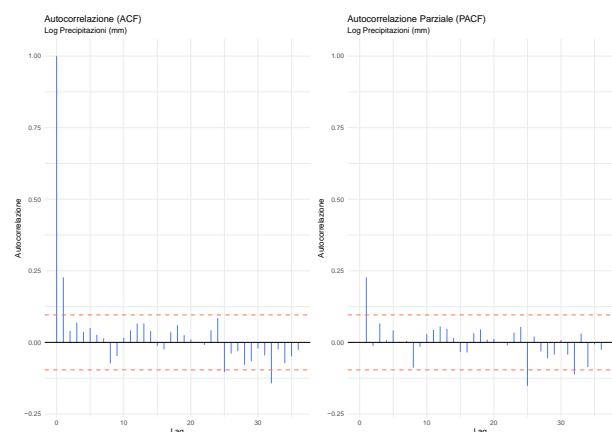


I grafici mostrano come la trasformazione logaritmica abbia gestito efficacemente la forte asimmetria negativa causata dal troncamento. Anche la coda destra ha beneficiato della trasformata, il problema dei valori anomali sembra essersi risolto. Tuttavia, il QQ-plot evidenzia che un residuo del troncamento persiste, come si può notare dall'estremità sinistra della distribuzione. Nonostante ciò, il fenomeno risulta significativamente migliorato e prossimo ad una Normale.

Successivamente, il filtro di Hampel ha permesso di estrarre e correggere i pochi outlier rimanenti. Nello specifico, sono risultate anomale le osservazioni di ottobre 1999 e febbraio 2000, due periodi rispettivamente di piogge abbondanti e di siccità atipica. L'ausilio di un interpolatore lineare ha consentito la correzione di questi valori identificati e l'imputazione dei missing values. Il grafico evidenzia come l'annata 2012 – 2013, pesantemente colpita da osservazioni mancanti, ha risentito di un assegnamento grossolano. In conclusione, è stata condotta un'analisi delle autocorrelazioni e della stazionarietà della serie.

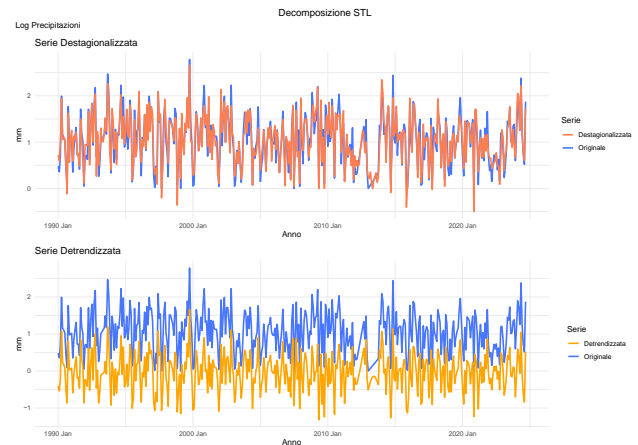
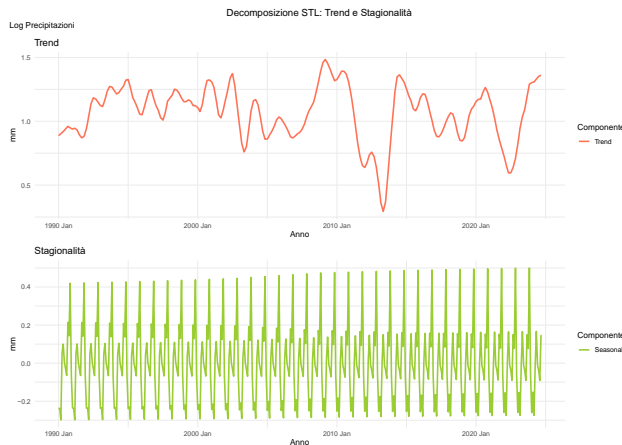


Dai correlogrammi non si evidenzia una struttura sottostante al fenomeno ben definita. La dipendenza temporale all'interno della serie è pressoché minima, se non fosse per il primo lag le correlazioni risulterebbero statisticamente non significative. Da questa breve disamina la serie non sembrerebbe essere affetta da componenti strutturali che ne condizionano l'andamento come trend e stagionalità. I test di stazionarietà confermano l'assenza di una tendenza temporale in quanto la sequenza risulta essere stazionaria attorno ad un drift di valore unitario.

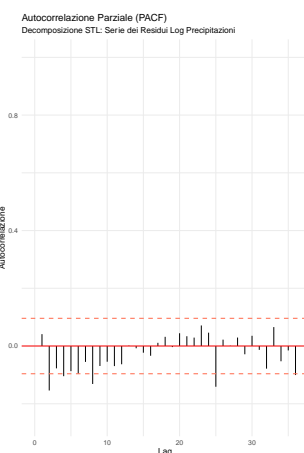
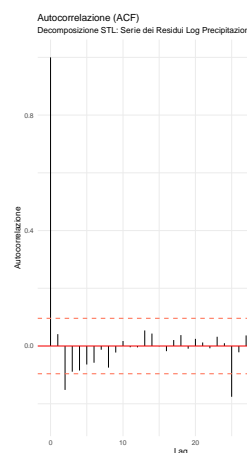
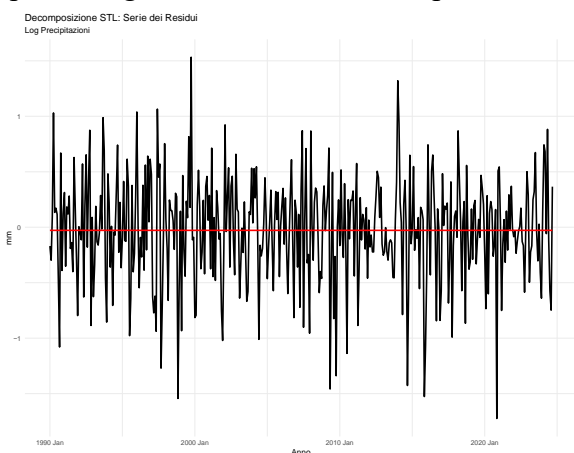


## 4.1 Analisi delle componenti non osservabili

Nel caso delle precipitazioni, la decomposizione in grado di estrarre al meglio le componenti non osservabili della serie risulta essere la STL.



La scomposizione rivela un andamento di lungo periodo costante nel tempo, le brusche oscillazioni che si osservano non sono dovute a break strutturali ma alla cattiva imputazione dei dati precedentemente evidenziata. La stagionalità è di tipo moltiplicativo, la componente si evolve nel tempo. Un'analisi delle fluttuazioni periodiche rileva i mesi di novembre e maggio come i più piovosi, mentre il contrario per gennaio, febbraio e marzo. Il grafico riportato a sinistra permette di osservare la forza delle componenti. Le serie destagionalizzate e detrendizzate tendono ad essere prossime a quella originale, il che rafforza l'ipotesi di casualità del fenomeno.



La componente residuale mostra un andamento casuale con un range di variazione molto contenuto. Sia il comportamento della serie che l'analisi dei correlogrammi non evidenziano pattern particolari o correlazioni significative, la sequenza sembrerebbe essere una realizzazione di un processo WN.

## 4.2 Modellistica

A seguito di una detrendizzazione mediante drift e destagionalizzazione tramite l'ausilio di due armoniche, si è proceduto ad applicare un filtro ARIMA alla serie.

Modello	AIC	AIC <sub>C</sub>	BIC	R <sup>2</sup> <sub>Adj</sub>	RMSE	MAE	ACF1
WN	658,75	658,78	666,82	0,00	0,53	0,42	0,20
AR (1)	643,43	643,49	655,53	0,04	0,52	0,42	0,00
AR (2)	645,39	645,48	661,52	0,04	0,52	0,42	0,00
MA (1)	639,79	639,80	643,82	0,04	0,52	0,42	0,00
MA (2)	645,73	645,83	661,86	0,04	0,52	0,42	0,00
ARMA (1,1)	644,64	644,74	660,77	0,04	0,52	0,41	0,02

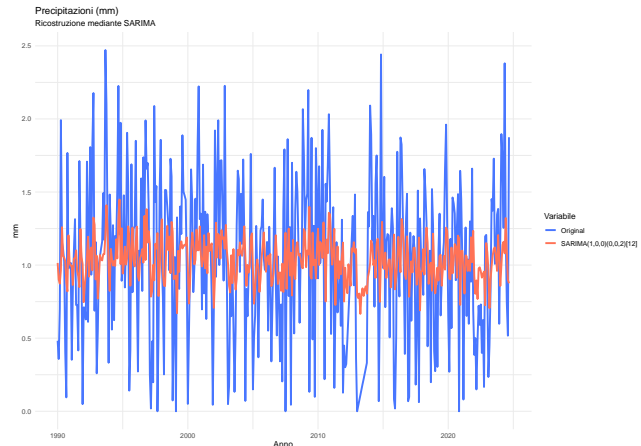
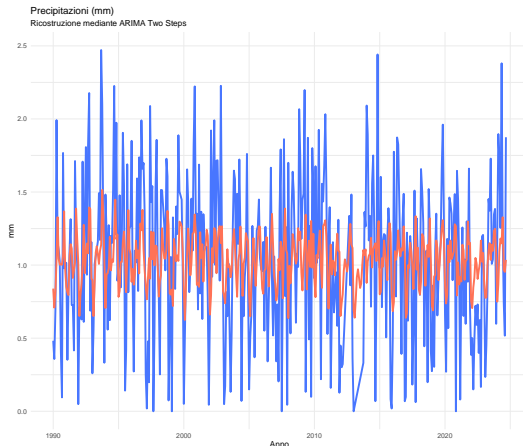
ARIMA	MA (1)
Coefficients	0,2018
Std. Error	0,0484

Tra le specificazioni analizzate quella che meglio si adattata alla serie, come osservabile dalla tabella, è l'MA (1). Da notare come il divario tra formulazioni e il WN è davvero minimo, a riprova del fatto che la componente residuale è considerabile come una realizzazione di un processo White Noise.

Per quanto riguarda il modello SARIMA, la specificazione migliore per fitting e generalizzazione è risultata essere la SARIMA (1,0,0) (0,0,2) [12].

SARIMA	AR (1)	SMA (1)	SMA (2)
Coefficients	0,2357	0,0511	0,1153
Std. Error	0,0481	0,0500	0,0499

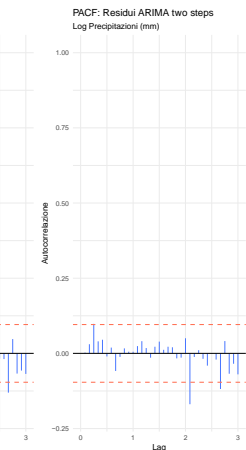
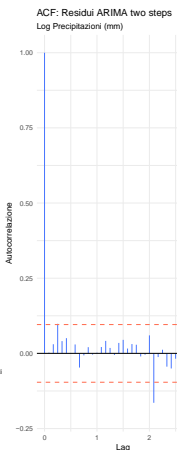
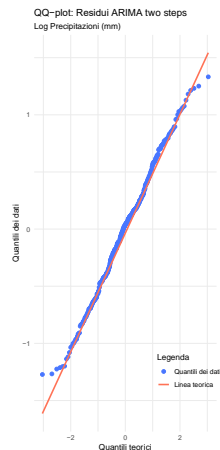
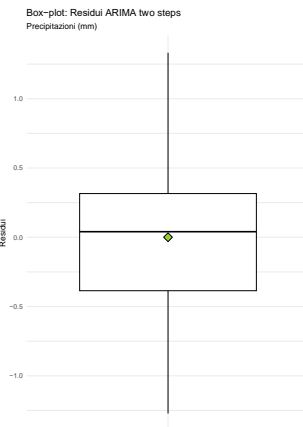
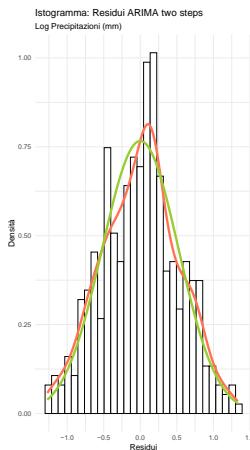
Per confrontare le performance dei modelli adattati una strategia è quella di analizzare la loro capacità di ricostruzione della serie originale.



Le precipitazioni mostrano un andamento prevalentemente casuale con una componente strutturale quasi trascurabile. Questo si riflette nella capacità dei modelli che riescono a cogliere solamente un leggero segnale di fondo ma non sono in grado di rappresentare le oscillazioni, il cui valore effettivo risulta imprevedibile. I dettagli delle performance sono riportati in tabella.

Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	0,5205	0,53
MAE	0,417	0,4288
R <sub>2</sub>	0,1001	0,0668

Siccome le formulazioni presentano risultati simili, anche da un punto di vista distributivo i residui mostrano le medesime caratteristiche. Per tale motivo vengono riportate le sole diagnostiche ARIMA.

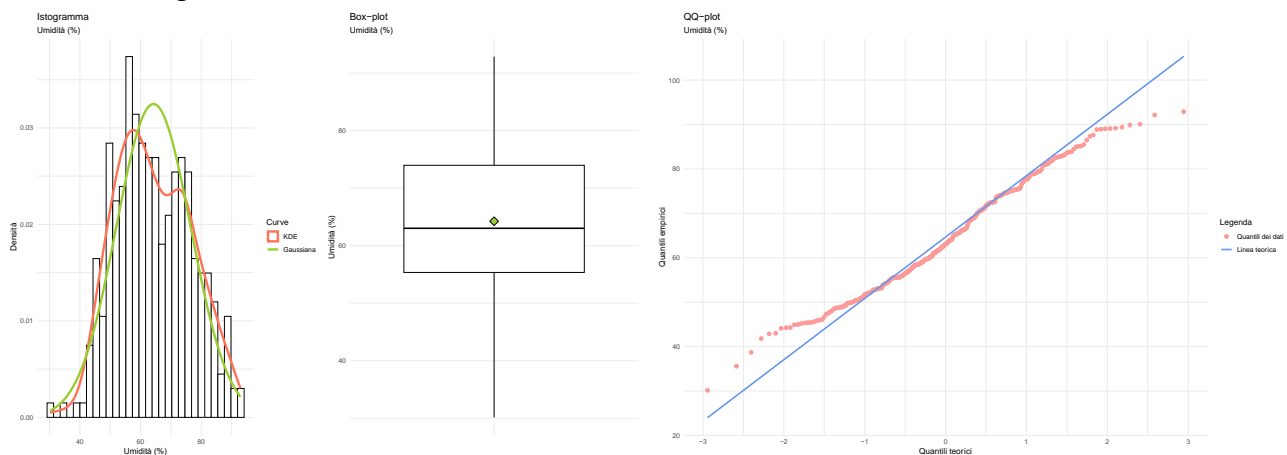


La componente residuale si distribuisce come una Normale, si rileva solamente una leggera asimmetria negativa trascurabile. I correlogrammi non evidenziano valori che fuoriescono ampiamente dalle bande di confidenza il che porta ad assumere che si tratti di una realizzazione di un processo WN gaussiano, i modelli sono stati in grado di estrarre completamente la componente strutturale della serie.

## 5. Umidità relativa

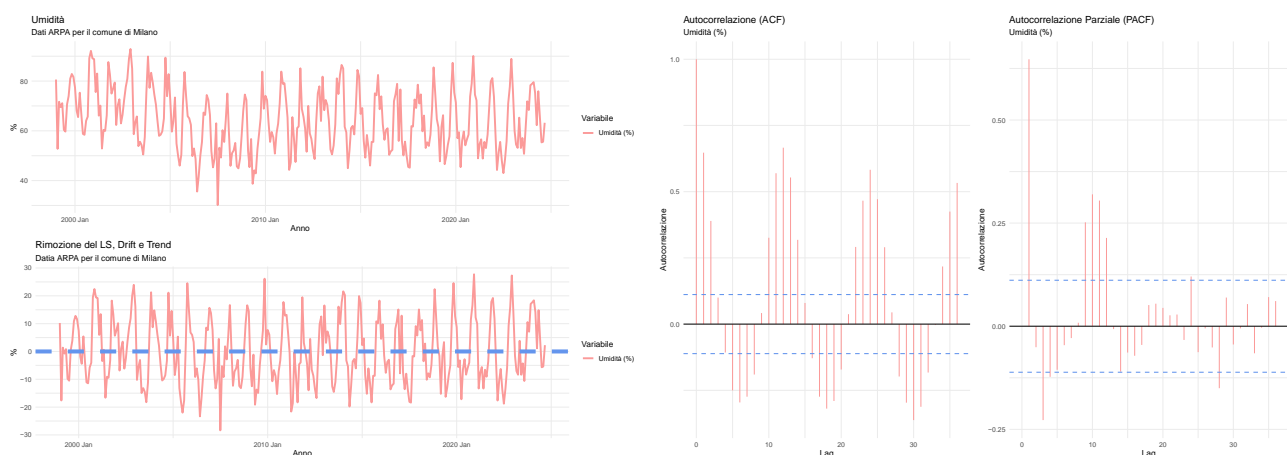
L'umidità influenza significativamente le reazioni chimiche e fisiche che coinvolgono l'NO<sub>2</sub>. Valori elevati del fattore facilitano il "lavaggio atmosferico" e le trasformazioni in acido nitrico o particelle sospese, riducendo la presenza dell'inquinante in atmosfera. Al contrario, in condizioni di bassa umidità, il tasso delle reazioni diminuisce favorendone un accumulo nell'aria.

La serie in analisi sembrerebbe riportare un level shift negativo in corrispondenza del periodo compreso tra ottobre 2005 e novembre 2009. Le oscillazioni appaiono indipendenti nel tempo e non influenzate dalla presenza di valori anomali. Infine, la sequenza riscontra una leggera tendenza crescente in prossimità del termine della finestra temporale. La distribuzione empirica presenta un'asimmetria negativa ed una leggera bimodalità probabilmente causata dal level shift temporaneo rilevato. Nonostante le criticità, la serie presenta una buona aderenza all'assunto di normalità e non si evidenziano valori anomali, ipotesi confermata anche dal filtro di Hampel. Per concludere, la sequenza non presenta missing values e, a seguito di un'analisi di Box-Cox, non necessita di alcuna trasformatrice stabilizzatrice.



Per verificare la presenza del level shift temporaneo, è stata effettuata una regressione lineare. La creazione di una variabile dummy, avente valore unitario in corrispondenza del periodo compreso tra ottobre 2005 e novembre 2009, ha permesso di replicare l'effetto del cambio di livello. Il coefficiente, essendo risultato fortemente significativo, ha permesso di confermare quanto ipotizzato e, dalla regressione, è inoltre emersa la presenza di una tendenza negativa che i test di Dickey-Fuller hanno confermato. La serie presenta, dunque, un andamento stazionario rispetto drift, trend e LS.

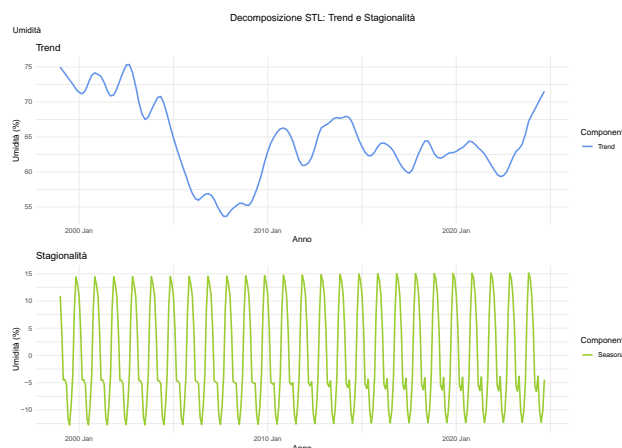
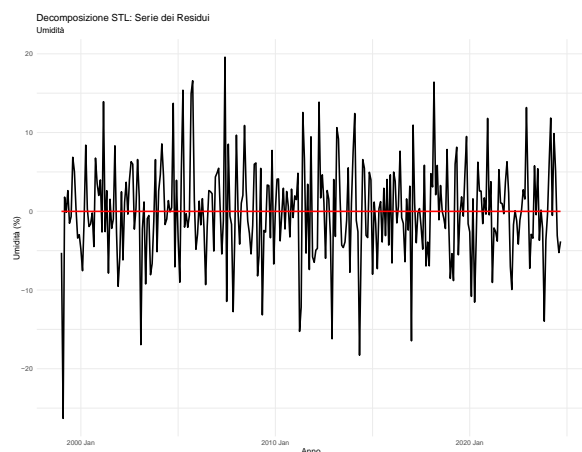
Dai correlogrammi, si evidenzia un comportamento fortemente stagionale, l'elevata persistenza delle ACF tra periodi ne è testimone. Si rileva in aggiunta un modesto legame tra periodi ma che, osservando le PACF, si annulla dopo una sola stagione. Tale risultato potrebbe essere sintomo di un processo AR (1) sottostante la serie.



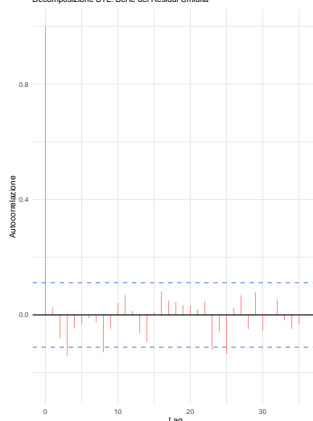


## 5.1 Analisi delle componenti non osservabili

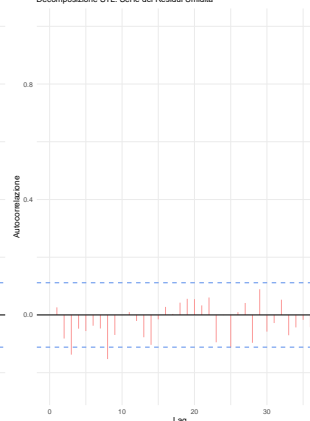
La decomposizione che meglio ha estratto le componenti non osservabili della serie è stata la STL. Dall'andamento di lungo periodo è ben evidente la presenza di un level shift temporaneo a cavallo tra gli anni 2005 e 2009. Prima del cambio di livello, la serie sembrerebbe avere un trend lineare decrescente, mentre, successivamente, una tendenza lievemente crescente. La scomposizione ha evidenziato la presenza di una stagionalità molto regolare nel tempo, sia per forma che per variabilità. La componente presenta valori massimi nei mesi di novembre-dicembre e minimi in giugno-luglio, con un leggero picco nel mese di maggio.



Autocorrelazione (ACF)  
Decomposizione STL: Serie dei Residui Umidità



Autocorrelazione Parziale (PACF)  
Decomposizione STL: Serie dei Residui Umidità

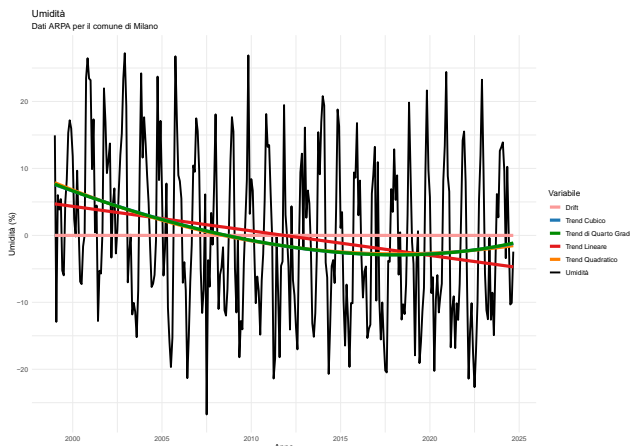


La componente residuale riporta un andamento casuale privo di evidenti problematiche, se non qualche oscillazione anomala. Dai correlogrammi non emergono differenze significative tra ACF e PACF, sintomo che la dipendenza temporale del fenomeno è stata estratta in modo esaustivo. Inoltre, non registrandosi valori al di fuori delle bande di confidenza è possibile affermare che si tratti di una realizzazione di un processo WN.

## 5.2 Modellistica

In prima analisi è stata applicata una procedura multistadio basata su una fase di detrendizzazione e destagionalizzazione della serie, seguita dall'applicazione di un filtro ARIMA. Siccome la serie risulta essere affetta da un level shift, prima di rimuovere il trend, si è proceduto a correggerlo.

Dal grafico emerge come l'andamento parabolico sia quello che meglio coglie il comportamento della serie, impressione confermata anche dai criteri informativi AIC, AIC<sub>C</sub> e BIC. Successivamente, la sequenza è stata destagionalizzata mediante l'ausilio di due armoniche e le sono state applicate una serie di specificazioni ARIMA. Come riportato nella tabella a pagina seguente, la formulazione che meglio gestisce il fitting e la generalizzazione risulta essere l'AR (1), nonostante non vi sia una sostanziale differenza dalle performance di un WN.

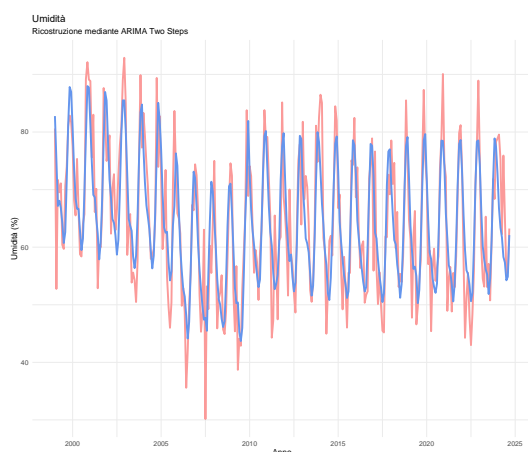


Modello	AIC	AIC <sub>c</sub>	BIC	R <sup>2</sup> <sub>Adj</sub>	RMSE	MAE	ACF1
WN	2066,36	2066,40	2073,83	0,00	6,81	5,36	0,15
AR (1)	2056,72	2056,73	2060,45	0,02	6,72	5,26	0,00
AR (2)	2062,72	2062,85	2077,65	0,02	6,72	5,26	0,00
MA (1)	2061,45	2061,53	2072,65	0,02	6,73	5,26	0,01
MA (2)	2062,69	2062,82	2077,62	0,02	6,72	5,26	0,00
ARMA (1,1)	2062,71	2062,84	2077,64	0,02	6,72	5,26	0,00

ARIMA	AR (1)
Coefficients	0,1540
Std. Error	0,0561

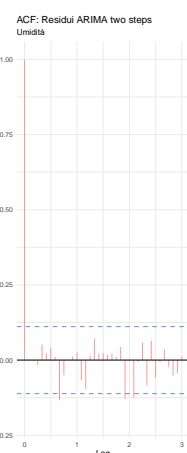
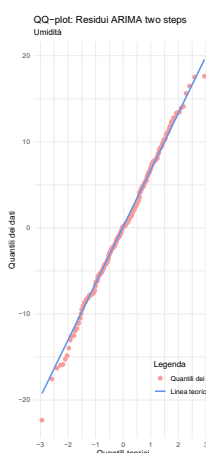
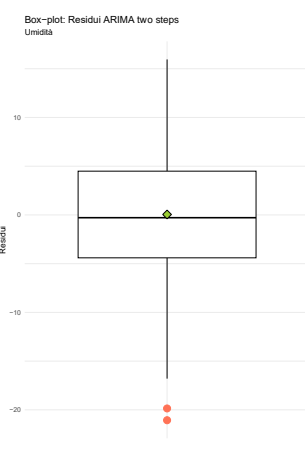
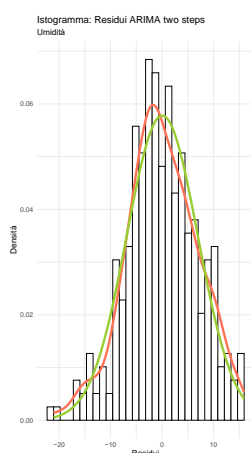
Per quanto riguarda la procedura diretta, il SARIMA (3,0,1) (0,1,1) [12] è il modello che meglio si è adattato alla serie, dettagli riportati in tabella.

SARIMA	AR (1)	AR (2)	AR (3)	MA (1)	SMA (1)
Coefficients	1,1013	-0,0987	-0,0245	-0,8529	-0,8833
Std. Error	0,0853	0,0876	0,0694	0,0609	0,0455



Tramite la ricostruzione della serie è possibile osservare come le diverse specificazioni producano risultati pressoché identici. Entrambe colgono molto bene la stagionalità, la tendenza sottostante la serie e soprattutto il cambio di livello temporaneo.

Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	6,7357	7,062
MAE	5,2593	5,5932
R <sub>2</sub>	0,6992	0,6699



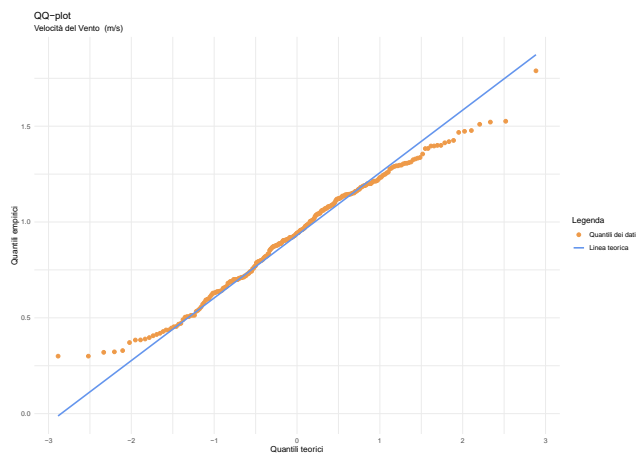
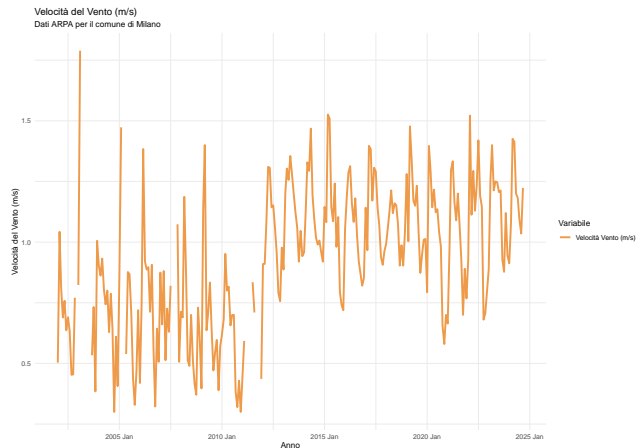
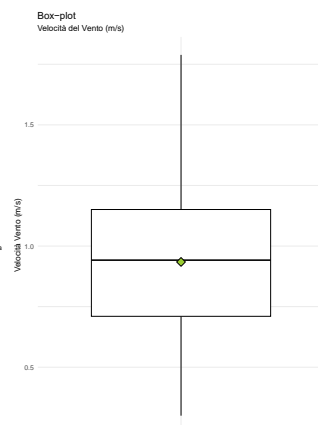
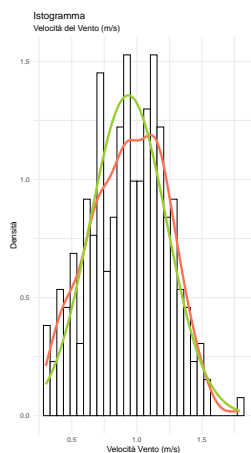
Essendo sia l'adattamento che le performance delle due procedure molto simili tra loro, sono state riportate le sole diagnostiche dei residui del modello ARIMA. La componente si distribuisce come una Normale, vi è una quasi perfetta sovrapposizione tra le curve di densità. I correlogrammi non mostrano pattern e valori che fuoriescono dalle bande di confidenza, dunque la serie è considerabile come una realizzazione di un processo WN gaussiano.

## 6. Velocità del Vento

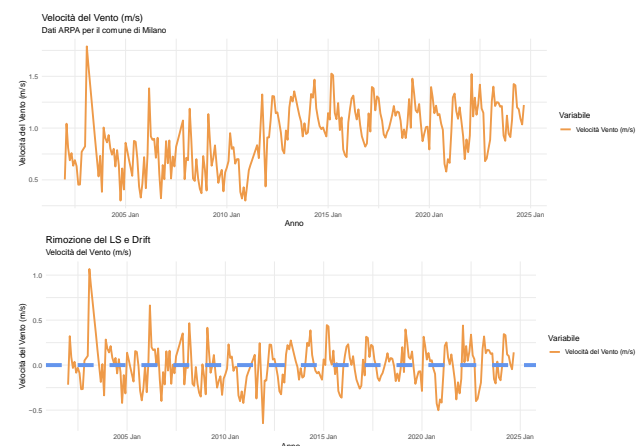
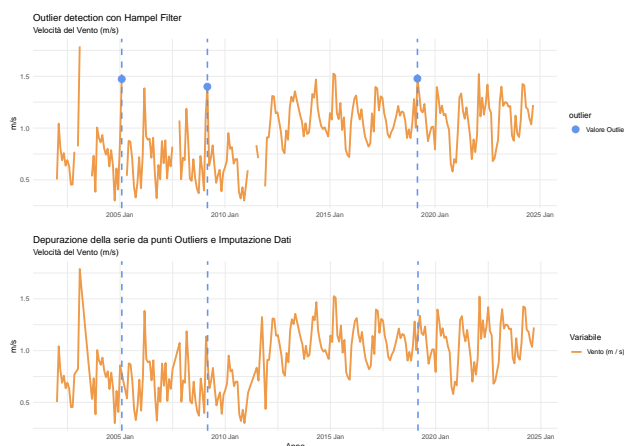
Il vento è uno dei fattori principali che contribuisce alla dispersione degli inquinanti atmosferici, maggiore è la sua velocità, più efficacemente si dissolvono. Al contrario, una sua mancanza favorisce l'accumulo delle sostanze, impattando negativamente sulla qualità dell'aria.



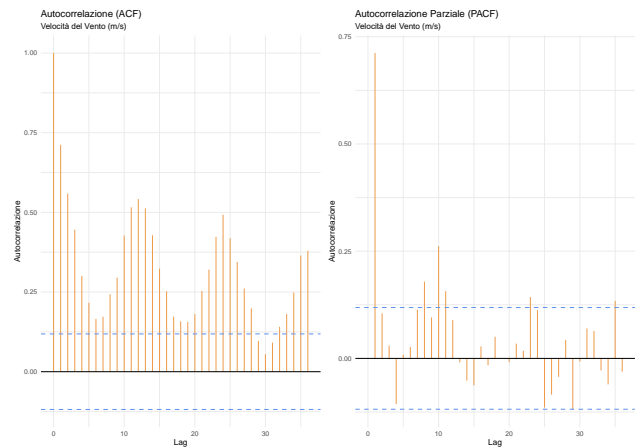
Dall'andamento della serie emerge la presenza di un level shift a cavallo tra il 2011 e il 2012, la sequenza passa da una velocità media di 0.75 m/s agli 1.1 m/s circa. Il cambiamento, essendo particolarmente repentino, potrebbe avere un'origine artificiale, come una variazione negli standard di rilevazione o la demolizione di un edificio vicino alla centralina che ne ha aumentato l'esposizione diretta. In generale, la serie appare stazionaria, ma nella prima parte presenta numerosi valori mancanti ed oscillazioni vertiginose, fenomeni assenti nella seconda finestra d'osservazione.



Nonostante le criticità rilevate, queste non sembrerebbero aver impattato sulle caratteristiche distributive della serie. La sequenza assume un comportamento prossimo a quello di una Normale, si rileva buona sovrapposizione tra le curve di densità, e non pare essere affetta da valori anomali. Tale osservazione verrà smentita dall'applicazione del filtro di Hampel, mediante il quale è stato possibile identificare tre rilevazioni outlier. Dai risultati riportati nella pagina successiva, solamente i valori di febbraio 2005 e marzo 2009 sembrerebbero essere anomali, al contrario, sorprendentemente, il dato di febbraio 2003 non è stato identificato come tale. Per questa ragione, è stata condotta un'analisi parallela per verificare, in fase di decomposizione, la natura delle osservazioni rilevate. Successivamente, è stata condotta un'analisi di Box-Cox dalla quale non è emersa la necessità di applicare alcuna trasformazione. In conclusione, è stato condotto uno studio per verificare l'effettiva presenza del level shift. Tramite regressione lineare si è rilevato come la serie abbia subito un cambio di livello a partire dal settembre 2011 e, mediante test di Dickey-Fuller, è stato possibile affermare come la sequenza riporti un andamento stazionario una volta depurato dalla variazione di livello.

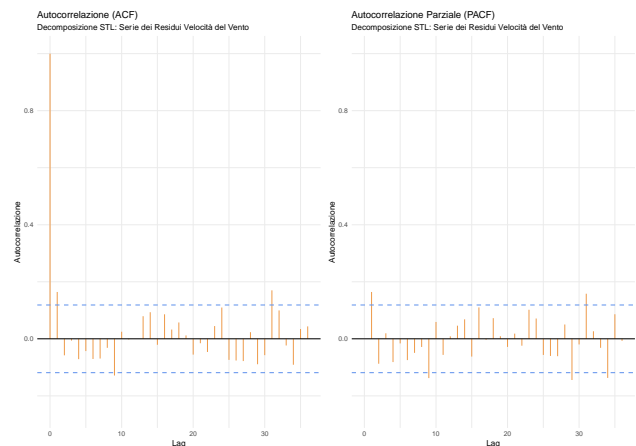
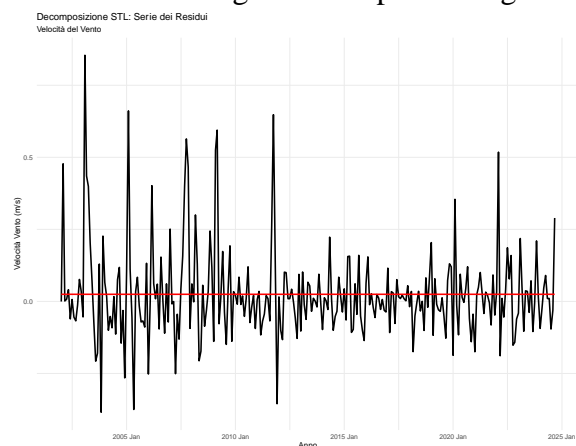
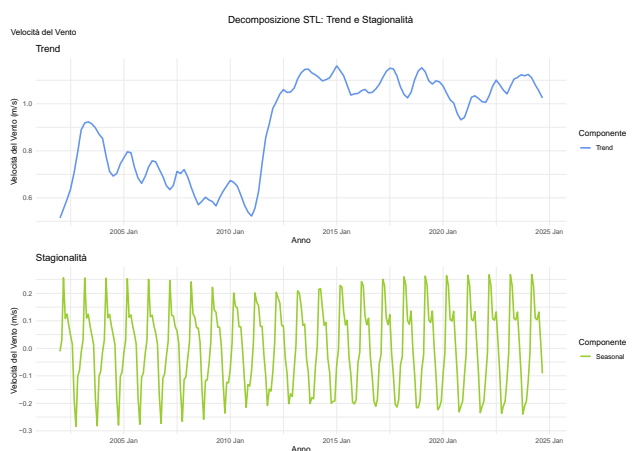


La fase esplorativa si è conclusa con l'analisi dei correlogrammi, dai quali emerge un andamento stagionale delle autocorrelazioni accompagnato da una persistenza moderatamente elevata. In particolare, il fenomeno mostra sia un legame tra periodi che nei periodi, suggerendo la presenza di un effetto annuale e di un'influenza reciproca tra le osservazioni adiacenti. Tramite le PACF è possibile notare come le autocorrelazioni tendono ad annullarsi dopo una stagione, indicando così l'ordine dell'autoregressività del fenomeno.



## 6.1 Analisi delle componenti non osservabili

La decomposizione che meglio si è adattata ai dati è la STL, la quale ha permesso di eliminare ogni dubbio circa la presenza di eventuali outliers e cambiamenti di livello. Tramite la componente di lungo periodo è evidente la presenza di un level shift in prossimità del settembre del 2011, mentre, dalle analitiche, emerge come anomala la sola osservazione relativa al febbraio 2003. In generale, la serie presenta un comportamento medio costante e una stagionalità moltiplicativa. Per tutta la finestra temporale, il mese di marzo è risultato il più ventoso, ottobre il meno, e negli ultimi anni si è registrato un picco a luglio.



I residui mostrano un andamento casuale cosparso, soprattutto nel primo periodo d'osservazione, da oscillazioni che fuoriescono dal range usuale di variazione dei dati. Questo è dovuto all'instabilità della serie e al fatto che la decomposizione, per caratteristiche intrinseche, non sia in grado di gestire sbalzi improvvisi. A testimonianza di quanto affermato, è possibile notare un residuo elevato in prossimità del level shift. Dai correlogrammi non si evidenziano pattern o valori significativi, il che potrebbe far intendere che la serie sia una realizzazione di un processo WN.

## 6.2 Modellistica

Dopo aver rimosso il cambiamento di livello e analizzato l'eventuale presenza di trend polinomiali, è emerso che la serie non riporta alcuna tendenza significativa. Per quanto riguarda la destagionalizzazione, è stato adottato un approccio basato sulla regressione armonica. Sebbene le diverse specificazioni abbiano mostrato performance quasi identiche, per modellare la componente

periodica sono state utilizzate tre armoniche. Infine, sono stati applicati vari filtri ARIMA che hanno decretato la formulazione AR (4) la migliore per fitting e per generalizzazione.

Modello	AIC	AIC <sub>c</sub>	BIC	R <sup>2</sup> <sub>Adj</sub>	RMSE	MAE	ACF1
WN	-172,88	-172,83	-165,66	0,00	0,18	0,13	0,37
AR (1)	-212,29	-212,20	-201,46	0,14	0,16	0,12	-0,05
AR (2)	-215,92	-215,77	-201,48	0,15	0,16	0,12	-0,02
AR (4)	-228,94	-228,93	-225,34	0,17	0,16	0,12	0,00
MA (1)	-202,62	-202,53	-191,79	0,11	0,17	0,12	0,04
MA (2)	-206,35	-206,20	-191,91	0,12	0,16	0,12	0,04
ARMA (1,1)	-217,19	-217,04	-202,75	0,16	0,16	0,12	-0,01

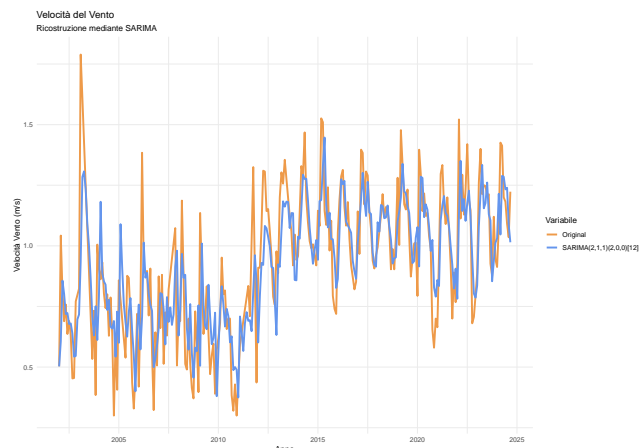
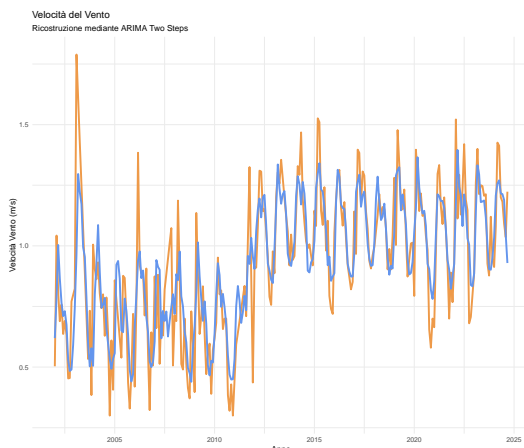
ARIMA	AR (1)	AR (2)	AR (3)	AR (4)
Coefficients	0,3165	0,1154	0,1544	-0,1073
Std. Error	0,0605	0,0633	0,0632	0,0608

Secondo un approccio diretto, la formulazione ottimale risulta essere la SARIMA (2,1,1) (2,0,0) [12].

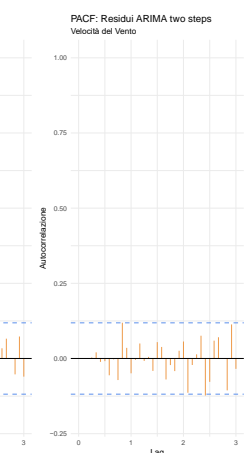
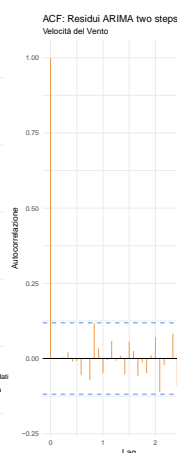
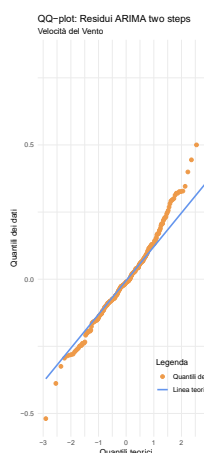
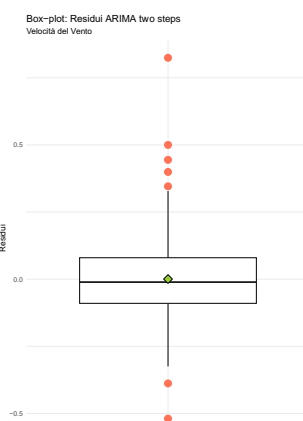
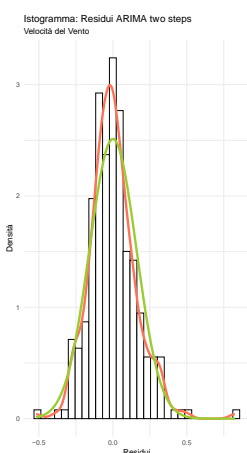
Nonostante le diverse procedure adottate, le specificazioni risultanti sono molto simili per performance adattamento come è ben visibile in tabella.

SARIMA	AR (1)	AR (2)	MA (1)	SAR (1)	SAR (2)
Coefficients	0,4699	0,1311	-0,9729	0,1623	0,3256
Std. Error	0,0655	0,0661	0,0206	0,0588	0,0637

Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	0,1587	0,1839
MAE	0,1181	0,1385
R <sub>2</sub>	0,6989	0,597



Entrambe le formulazioni godono di un'ottima capacità di ricostruzione della serie riuscendo a modellare, in maniera corretta, il cambio di livello e le oscillazioni improvvise della prima parte della finestra temporale. Siccome le prestazioni dei modelli sono molto simili tra loro, anche le analitiche residuali lo sono.

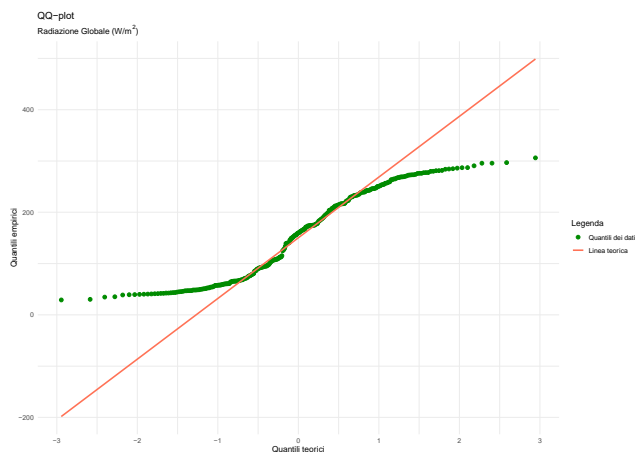
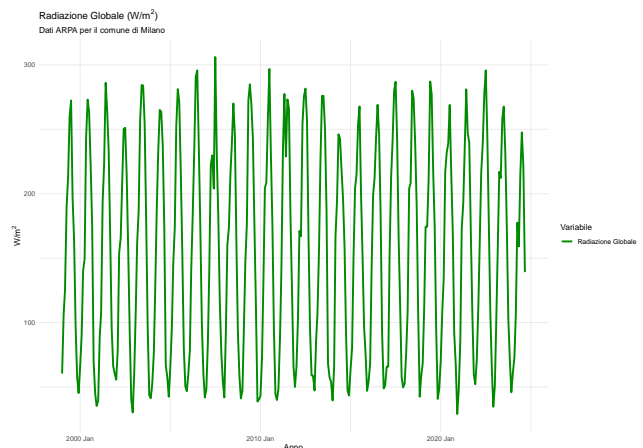
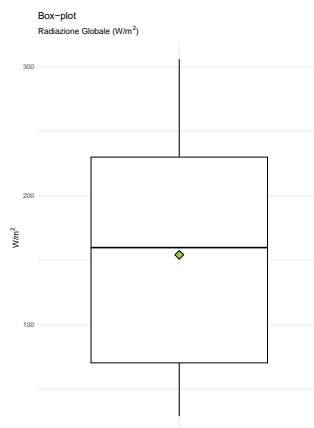
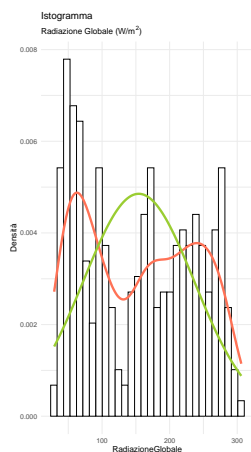


Nel complesso si nota una buona aderenza all'assunto di normalità intaccato solamente da una leggera pesantezza della coda destra a causa di qualche valore outlier. Entrambe le specificazioni sono state in grado di estrarre completamente la componente strutturale del fenomeno rendendo i residui una realizzazione di un processo WN, come testimoniato dai correlogrammi.

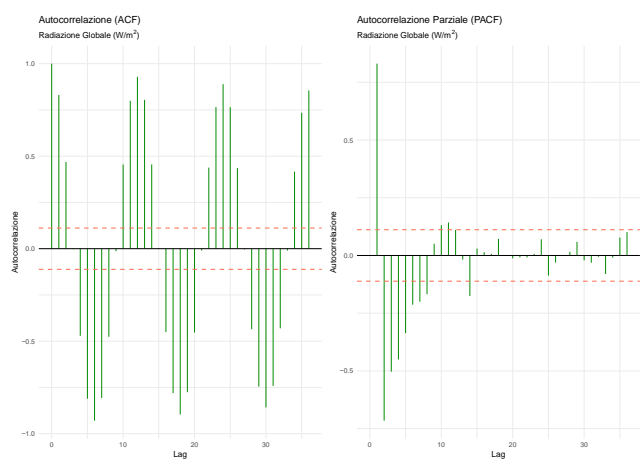
## 7. Radiazione Globale

In meteorologia, la radiazione globale è la somma di tutta l'energia solare che raggiunge la superficie terrestre, sia diretta che diffusa. La componente luminosa è fondamentale nei processi chimici che coinvolgono l'NO<sub>2</sub>. Quando la luce colpisce il diossido, tramite la fotolisi, lo decompone originando

l'ozono troposferico ( $O_3$ ). Nelle giornate particolarmente soleggiate, questo processo è molto intenso e contribuisce alla riduzione dell'inquinata, incrementando, al tempo stesso, le emissioni di  $O_3$ , gas serra altrettanto dannoso. La serie si mostra stazionaria attorno ad un drift e caratterizzata da un forte comportamento stagionale apparentemente invariante per tutta la finestra temporale. Non si evidenziano oscillazioni singolari, tendenze deterministiche e valori outlier. Da un punto di vista distributivo la serie appare multimodale, discostandosi totalmente dall'assunto di Normalità, la forte componente stagionale potrebbe esserne la causa. In conclusione, la sequenza non presenta punti anomali, assunto comprovato dal filtro di Hampel.



Nel caso della radiazione globale, non è stata effettuata alcuna trasformazione di Box-Cox, il parametro  $\lambda$ , essendo unico, non è in grado di gestire la multimodalità. L'analisi delle ACF evidenzia un forte legame tra periodi, caratterizzato da un andamento oscillatorio ed un'elevata persistenza. Il fenomeno, unito all'azzeramento delle PACF dopo il dodicesimo lag, indica la probabile presenza di un processo stagionale AR (1) con radice negativa sottostante. Inoltre, si osserva una moderata correlazione nei periodi. Infine, un'analisi della persistenza condotta mediante test di Dickey-Fuller ha confermato un comportamento stazionario della serie attorno a un drift di  $154 W/m^2$ .



## 7.1 Analisi delle componenti non osservabili

La decomposizione STL ha permesso d'identificare con precisione le componenti deterministiche del fenomeno. L'analisi del trend di lungo periodo non ha evidenziato la presenza di tendenze polinomiali, ma ha rilevato un calo significativo nell'ultima fase della serie. Altre scomposizioni, come la X11 e la SEATS, hanno segnalato un possibile cambiamento di livello nel marzo 2024. Tuttavia, essendo collocato ai margini della finestra temporale, non è stato possibile stabilire se si sia trattato di un effettivo level shift o di un decadimento casuale. Per quanto riguarda la stagionalità, la decomposizione ha evidenziato un andamento regolare e costante nel tempo, con un

massimo a luglio ed un minimo a dicembre. Complessivamente, si osservano valori più alti nei mesi estivi e più bassi in quelli invernali, in linea con la minore esposizione solare durante la stagione fredda. La componente residuale mostra un andamento casuale, con un range di variazione contenuto, affetto dalla presenza di alcuni valori atipici. Questi potrebbero derivare dalla difficoltà della decomposizione nel modellare le rilevazioni anomale, come quella del giugno 2007 avente residuo pari a  $-71 \text{ W/m}^2$ . I correlogrammi, non evidenziando valori significativamente diversi da zero, indicando come la decomposizione sia riuscita ad estrarre completamente la dipendenza temporale.



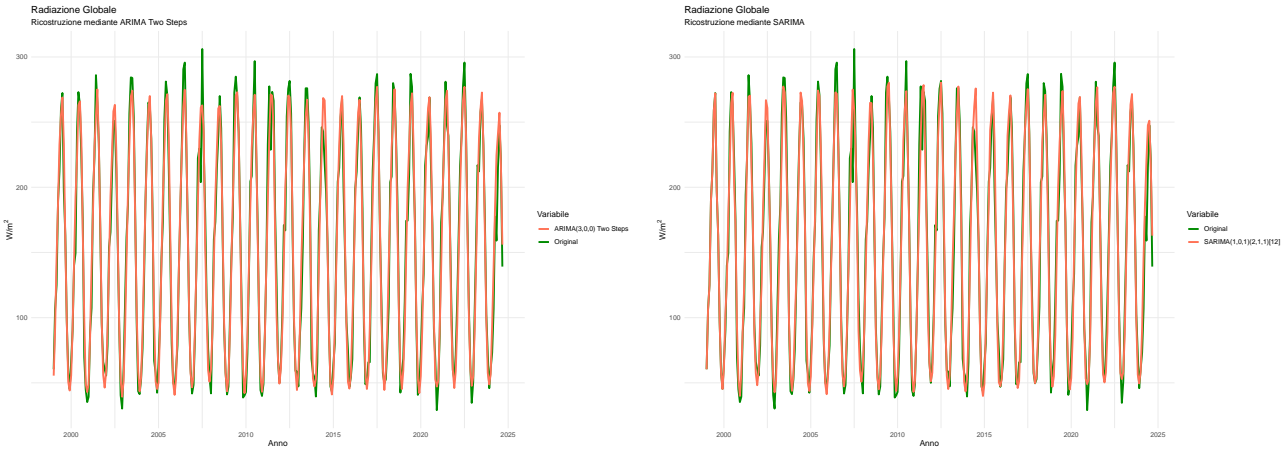
## 7.2 Modellistica

In una prima fase è stata eseguita una procedura ARIMA two steps. A seguito di una detrendizzazione mediante rimozione del drift ed una destagionalizzazione tramite l'ausilio di tre armoniche, la specificazione che meglio si è adattata alla serie è risultata essere l'AR (3). In seguito, una procedura diretta ha identificato come ottimo il modello SARIMA (1,0,1) (2,1,1) [12], dettagli in tabella.

Modello	AIC	AIC <sub>c</sub>	BIC	R <sup>2</sup> <sub>Adj</sub>	RMSE	MAE	ACF1
WN	2633,27	2633,31	2640,73	0,00	17,04	13,09	0,22
AR (1)	2619,88	2619,95	2631,08	0,05	16,62	12,72	-0,01
AR (2)	2621,20	2621,33	2636,13	0,04	16,60	12,69	-0,01
AR (3)	2607,87	2607,88	2611,60	0,07	16,40	12,51	0,00
MA (1)	2620,96	2621,04	2632,16	0,04	16,65	12,78	0,01
MA (2)	2622,72	2622,85	2637,66	0,04	16,64	12,75	0,00
ARMA (1,1)	2618,39	2618,52	2633,32	0,05	16,52	12,57	0,02

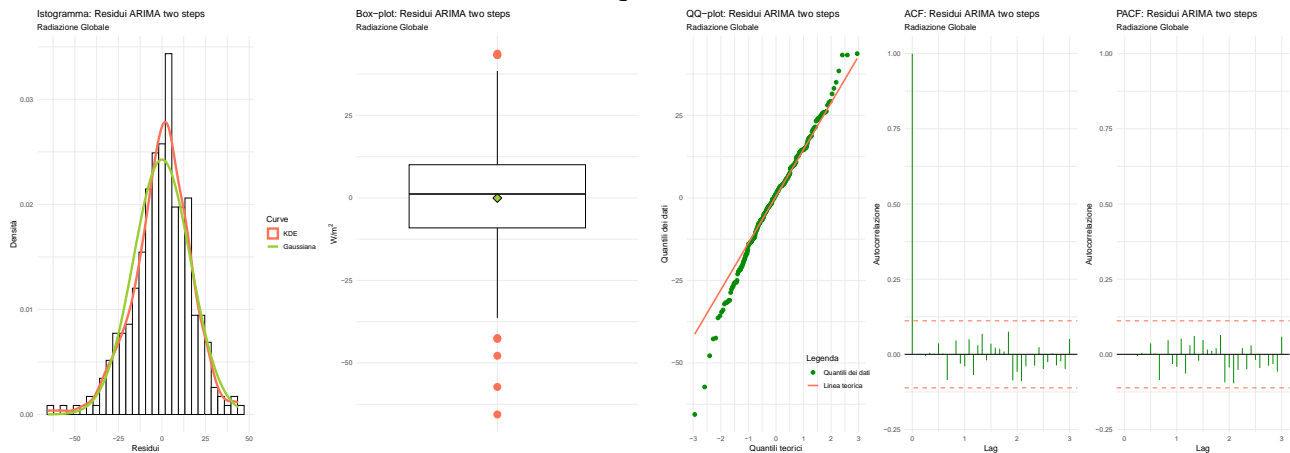
ARIMA	AR (1)	AR (2)	AR (3)
Coefficients	0,2043	0,0149	0,1540
Std. Error	0,0563	0,0575	0,0565

SARIMA	AR (1)	MA (1)	SAR (1)	SAR (2)	SMA (1)
Coefficients	0,7230	-0,5384	-0,0919	-0,1057	-0,8670
Std. Error	0,1418	0,1704	0,0762	0,0706	0,0628



Tramite la ricostruzione delle serie e gli indici di adattamento riportati in tabella, è stato possibile confrontare le performance delle due procedure. Entrambe le formulazioni godono di un eccellente adattamento ai dati, cogliendo perfettamente le oscillazioni stagionali ed evidenziando un fitting pressoché identico, confermato anche dalle statistiche riportate. Per tale ragione, si sono analizzati i residui di una sola formulazione in quanto le considerazioni risulterebbero le medesime.

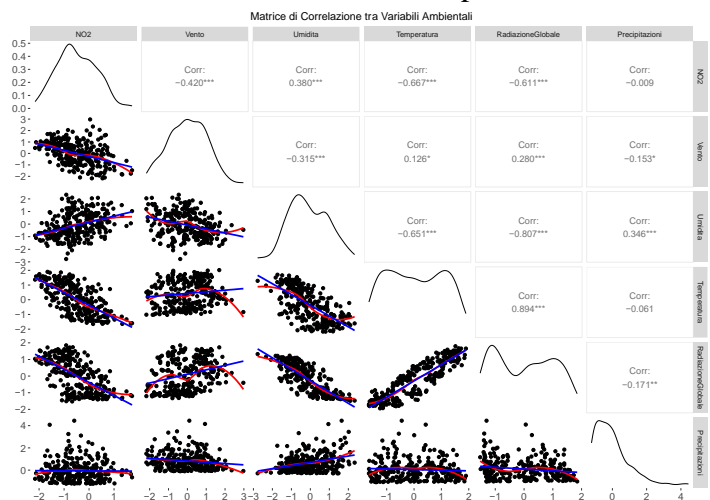
Indici di adattamento \ Modelli	ARIMA two step	SARIMA
RMSE	16,4296	16,8143
MAE	12,5124	12,4256
R <sub>2</sub>	0,96	0,9582



Dalle analitiche emerge come la componente residuale si distribuisca in modo prossimo a quello di una Normale. Nonostante la coda distributiva sinistra sia leggermente più pesante di quella teorica, non si nota uno scostamento significativo tra la curva empirica da quella gaussiana. I correlogrammi, non rilevando pattern o dipendenze temporali significative, consentono di affermare che la serie residuale non sia altro che una realizzazione di un processo WN.

## 8. Policy Evaluation

L'obiettivo della policy evaluation è quello di verificare se, al netto delle variabili ambientali, l'introduzione delle aree a traffico limitato abbia contribuito alla riduzione delle emissioni di NO<sub>2</sub>. A tal fine, è stata condotta una regressione lineare includendo la variabile categoriale "PoliticheAmbientali" per rappresentare i vincoli ecologici comunali introdotti nel tempo. Poiché i fenomeni atmosferici differiscono sia per unità di misura che per ordine di grandezza, per rendere confrontabili i loro effetti, si è proceduto a standardizzare le covariate e la dipendente. Prima di regredire il diossido d'azoto in funzione dell'esplicative identificate, si è svolta un'analisi delle correlazioni per valutare eventuali collinearità. Dal grafico non emergono criticità rilevanti, se non un forte legame positivo tra temperature e radiazione globale, dovuto al fatto che una maggiore esposizione solare comporta un incremento delle stesse. Analizzando la colonna di sinistra, si nota che la maggior parte delle variabili atmosferiche presenta una buona relazione lineare con l'inquinante, rendendo superflue eventuali trasformazioni per migliorarne la linearità.



### 8.1 Regressione lineare e Analisi dei risultati

Un confronto cross-validato tra diverse specificazioni ha individuato, sulla base dei criteri informativi AIC, AIC<sub>c</sub> e BIC, e degli indici di adattamento R<sup>2</sup><sub>ADJ</sub> e RMSE, la formulazione ottimale che segue.



NO<sub>2</sub> ~ Vento + Umidita + Temperatura + RadiazioneGlobale + Precipitazioni + PoliticheAmbientali

Coefficienti						
	Estimate	Std. Error	t-value	Pr(>   t   )	Signif.	
(Intercept)	-0,1442	0,0447	-3,2280	0,0014	**	
Vento	-0,0738	0,0369	-1,9980	0,0467	*	
Umidita	-0,1463	0,0451	-3,2470	0,0013	**	
Temperatura	-0,3215	0,0609	-5,2810	0,0000	***	
RadiazioneGlobale	-0,2852	0,0760	-3,7550	0,0002	***	
Precipitazioni	-0,0554	0,0271	-2,0420	0,0421	*	
politicheAmbientali: Area C	-0,3554	0,0740	-4,8050	0,0000	***	
politicheAmbientali: Area B	-0,5976	0,1500	-3,9850	0,0001	***	
politicheAmbientali: Euro 4 diesel No FAP	-0,7827	0,1276	-6,1360	0,0000	***	
politicheAmbientali: Euro 4 diesel FAP	-1,0329	0,0961	-10,7450	< 2e-16	***	
politicheAmbientali: Euro 5 diesel	-1,0814	0,1050	-10,3000	< 2e-16	***	
CV	AIC	AICc	BIC	AdjR2	RMSE	Sigma
0,1547	-505,6374	-504,4374	-462,3237	0,7534	0,3790	0,3870

La tabella evidenzia come tutte le variabili ambientali considerate nello studio siano risultate statisticamente significative nello spiegare le emissioni di NO<sub>2</sub>. I coefficienti di regressione mostrano una relazione negativa tra le condizioni atmosferiche e l'inquinante, indicando che all'aumentare dei fattori meteorologici è associata una riduzione delle concentrazioni di diossido. Nel caso specifico del comune di Milano, temperatura, radiazione globale e umidità emergono come le covariate più rilevanti nella regolazione del composto azotato, mentre velocità del vento e precipitazioni hanno un ruolo meno incisivo. I risultati dell'analisi suggeriscono che l'introduzione delle zone a traffico limitato, Area C ed Area B, e i divieti di circolazione sono state delle politiche efficaci. L'utilizzo della variabile categoriale "*PoliticheAmbientali*" ha permesso di replicare l'integrazione progressiva delle restrizioni ecologiche. Grazie al carattere cumulativo della variabile, è stato possibile evitare problemi di collinearità che sarebbero emersi con l'uso di singole variabili dummy. Questo spiega la monotonicità decrescente dei coefficienti associati ai livelli della variabile, poiché riflettono l'accumulo degli effetti delle misure rispetto alla situazione antecedente la loro introduzione. Per valutare il contributo delle politiche, è necessario riportare i coefficienti alla scala originale e calcolare la variazione percentuale secondo la formula indicata.

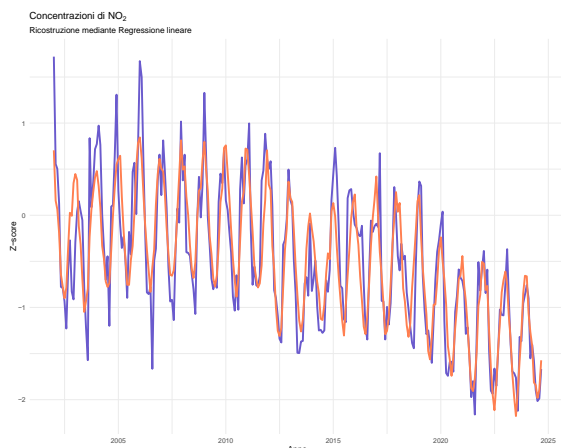
$$\hat{\beta}_{originale} = \hat{\beta}_{std} \cdot \sigma_Y; \quad \text{Variazione percentuale} = \frac{\hat{\beta}_{originale}}{\bar{Y}_0} * 100$$

dove  $\bar{Y}_0$  è la media della variabile dipendente originale prima dell'introduzione delle norme.

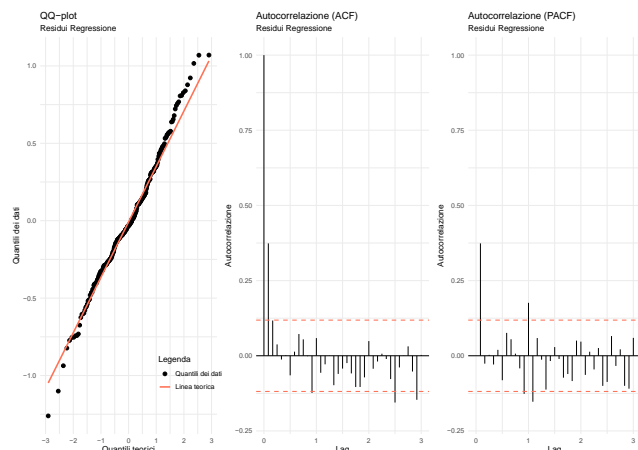
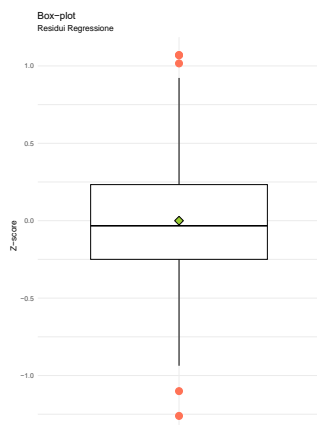
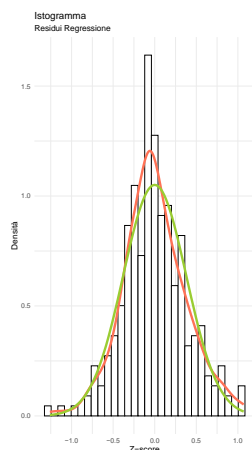
La tabella riporta i contributi individuali delle restrizioni ambientali considerate. L'introduzione delle zone a traffico limitato, Area C e Area B, ha determinato una riduzione complessiva delle emissioni di NO<sub>2</sub> pari al 16,86%, mentre i successivi divieti di circolazione hanno contribuito a una contrazione del 13,65%. L'analisi della serie ricostruita tramite regressione lineare, grafici riportati ad inizio pagina seguente, evidenzia un discreto adattamento con la sequenza temporale originale, soprattutto nell'ultimo periodo osservato.

Tuttavia, è interessante notare come le emissioni riportino un andamento decrescente già a partire dal 2008, suggerendo che il calo fosse in atto indipendentemente dalle restrizioni. Questo trend potrebbe essere legato a fattori esterni, come il potenziamento del trasporto pubblico e la graduale sostituzione dei veicoli Euro 0, 1 e 2, più inquinanti. La reale causa della flessione rimane incerta, ma l'analisi dei residui conferma l'omissione di un regressore non identificato, in quanto, l'andamento rilevato, non è compatibile con quello di un processo White Noise.

Politica Ambientale	Effetto Marginale (%)
Area C	-10,03
Area B	-6,83
<b>Limiti ZTL</b>	<b>-16,86</b>
Divieto Diesel Euro 4 no FAP	-5,22
Divieto Diesel Euro 4 FAP	-7,06
Divieto Diesel Euro 5 FAP	-1,37
<b>Divieti Circolazione Diesel</b>	<b>-13,65</b>



Esaminando i grafici sottostanti, i residui appaiono distribuirsi in modo gaussiano, la densità empirica è prossima a quella teorica e i quantili sono ben sovrapposti alla linea di tendenza. I correlogrammi, tuttavia, indicano l'esistenza di una dipendenza temporale, confermando il rifiuto dell'ipotesi di casualità delle osservazioni residue.



## 9. Conclusioni

Per completezza, è importante precisare che l'analisi è stata condotta sulla serie relativa alle concentrazioni totali di  $\text{NO}_2$  in atmosfera, e non specificamente sulle emissioni provenienti dal trasporto urbano, non direttamente osservabili. Le normative analizzate, riferendosi esclusivamente alla porzione di inquinante generata dalla viabilità cittadina, ma applicate alla serie totale, potrebbero aver sovrastimato il loro impatto effettivo. Secondo il Consiglio dell'Unione Europea, solo il 37% delle emissioni di ossidi di azoto è attribuibile al trasporto stradale. Pertanto, l'effetto reale delle normative comunali dovrebbe essere approssimativamente la metà di quanto osservato.

Dallo studio emerge che l'introduzione progressiva delle restrizioni urbane ha contribuito a una riduzione del 30% delle emissioni di  $\text{NO}_2$ . Tuttavia, rapportando questa stima alla quota di inquinante attribuibile al trasporto stradale, l'effetto si ridurrebbe a circa il 13-15% del totale atmosferico.

Nonostante questi risultati, le politiche ambientali presentano alcune criticità significative. Dati MEF evidenziano che circa un terzo dei residenti milanesi dichiara un reddito annuo inferiore ai 16.000 euro, rendendo difficile, se non impossibile, sostenere i costi per la sostituzione del proprio veicolo con uno più ecologico e conforme alle normative. A ciò si aggiunge l'inefficienza del trasporto pubblico, in particolare in alcune fasce orarie e aree periferiche, e l'incremento di infrastrutture decentrate, che trasformano queste restrizioni in un ostacolo più che in una soluzione.

In questo contesto, le politiche ambientali rischiano di ampliare le disuguaglianze sociali, rendendo gli spostamenti dei cittadini meno abbienti sempre più complessi, a fronte di un miglioramento ambientale che potrebbe risultare esiguo.



## Bibliografia

<https://www.consilium.europa.eu/it/infographics/air-pollution-in-the-eu/>

<https://www.europarl.europa.eu/topics/it/article/20190313STO31218/emissioni-di-co2-delle-auto-i-numeri-e-i-dati-infografica>

<https://www.europarl.europa.eu/topics/it/article/20180305STO99003/ridurre-le-emissioni-di-anidride-carbonica-obiettivi-e-azioni-dell-ue>

<https://www.europarl.europa.eu/topics/it/article/20230316STO77629/cambiamento-climatico-gas-a-effetto-serra-che-causano-il-riscaldamento-globale>

<https://www.europarl.europa.eu/topics/it/article/20180301STO98928/emissioni-di-gas-serra-per-paese-e-settore-infografica>

<https://www.europarl.europa.eu/topics/it/article/20180703STO07123/statistiche-sul-cambiamento-climatico-in-europa#co2-non--lunico-gas-serra-responsabile-del-riscaldamento-globale-1>

<https://www.europarl.europa.eu/topics/it/article/20180920STO14027/auto-furgoni-e-inquinamento-i-nuovi-obiettivi-per-le-emissioni>

[https://www.cittametropolitana.mi.it/Piano\\_Strategico\\_2022\\_2024/dati\\_territoriali/occupazione\\_red\\_dito](https://www.cittametropolitana.mi.it/Piano_Strategico_2022_2024/dati_territoriali/occupazione_red_dito)

<https://www.consilium.europa.eu/it/infographics/air-pollution-in-the-eu/>

<https://www.comune.milano.it/area-b-calendario-dei-divieti>