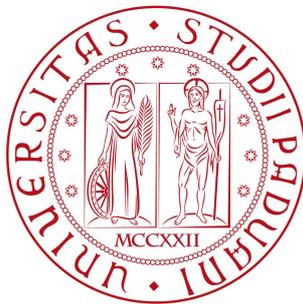


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in

Statistica per l'Economia e l'Impresa



**Analisi dell'erogato giornaliero di carburante nelle stazioni
di servizio: il caso Vega**

Relatore: prof. Francesco Lisi

Dipartimento di Scienze Statistiche

Laureando: Federica Piergallini

Matricola n. 1220661

Anno Accademico 2021/2022

Indice

1	Introduzione	2
2	Il dataset	4
2.1	Presentazione dei dati	4
2.2	Analisi descrittiva	5
3	Il modello statistico	10
3.1	Modelli per intervento	10
3.2	Costruzione del modello	12
4	Analisi delle serie	15
4.1	Arino di Dolo	15
4.2	Bacone	17
4.3	Castel Maggiore	19
4.4	Paccagnella	22
4.5	Carpi	24
4.6	Monselice	26
4.7	Forte Tomba	28
4.8	Treviso Sud	30
5	Conclusione	33
A	Appendice	35

Capitolo 1

Introduzione

Questo elaborato nasce a seguito di una collaborazione con la Società Vega Carburanti S.p.A, la quale si occupa della distribuzione di carburante.

L'azienda è presente nel settore del commercio di prodotti petroliferi in alcune regioni italiane, contando un centinaio di punti vendita totali.

Grazie al contributo dell'azienda, sono stati recuperati i dati necessari per lo studio dell'erogato dei distributori facenti parte di Vega.

Si dispone, quindi, delle serie giornaliere di ogni singolo punto vendita relative ad un periodo temporale nel quale si è verificato un evento inatteso: la diffusione in Italia del Covid-19.

Quest'ultima ha cambiato profondamente lo scenario economico italiano, colpito da uno shock imprevedibile e di grande portata, innescando così un crollo della domanda di beni e servizi.

La rapida circolazione del virus ha costretto l'Italia ad adottare misure drastiche ma necessarie per arginare l'epidemia: la messa in atto di un lockdown totale delle persone e delle attività economiche, ad eccezione di quelle essenziali, ha provocato effetti ingenti e persistenti sul contesto macroeconomico del Paese.

Di conseguenza, anche il settore della distribuzione del carburante ne ha risentito, registrando un crollo dei volumi di vendita, dovuto alle limitazioni alla mobilità delle persone per il rispetto delle rigide misure di contenimento della pandemia.

Partendo da questa premessa, l'obiettivo di questo lavoro è lo studio dell'impatto del Covid sugli erogati di carburante nelle stazioni di servizio dell'azienda Vega.

In particolare si vuole stabilire se l'inizio della diffusione del virus e le sue conseguenze negative abbiano avuto un impatto statisticamente significativo sulla variabile relativa all'erogato giornaliero e, in caso affermativo, cercare di stimare la diminuzione verificatasi nel periodo successivo all'arrivo in Italia del Covid-19.

L'analisi risultata migliore per spiegare l'evoluzione del fenomeno, in questo studio, è l'analisi dell'intervento.

Considerando "intervento" un evento esogeno che in un certo periodo influenza la serie, esso viene introdotto nella costruzione del modello per stimare l'effetto che ha avuto sulla dinamica di base.

Introdurre nel modello ulteriori variabili relative all'andamento del Covid, inoltre, può essere utile per eseguire un'analisi più corretta e che tenga conto dell'effetto di lungo periodo del virus in modo da determinare la diminuzione registrata rispetto al periodo pre-Covid in termini di erogato medio giornaliero.

Il lavoro è strutturato come segue.

Nel capitolo 2 vengono descritte le caratteristiche generali dell'azienda e viene effettuata l'analisi descrittiva delle serie degli erogati nel periodo che va dal 2018 al 2021.

Il capitolo 3 si concentra sulla spiegazione e descrizione del modello di riferimento, in particolare questo capitolo viene dedicato alla presentazione della successiva analisi dell'intervento.

Nel capitolo 4 si passa poi ad analizzare le 8 serie di interesse, volendo valutare la significatività statistica dell'intervento del Covid-19 sugli erogati giornalieri di carburante.

Il capitolo 5, infine, presenta i risultati ottenuti in questa analisi.

Capitolo 2

Il dataset

2.1 Presentazione dei dati

I dati trattati ai fini delle analisi sono stati forniti dall'azienda Vega, la quale svolge la propria attività nel settore petrolifero, dove esercita il commercio di carburanti e prodotti petroliferi in genere.

La Società Vega Carburanti S.p.A., con sede legale a Venezia Mestre, è attiva nella distribuzione di carburanti per autotrazione attraverso una rete di 110 punti vendita, ubicati in 8 diverse regioni: Piemonte, Lombardia, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Toscana, Marche, Puglia.

L'azienda opera sul mercato attraverso 3 marchi:

- Vega, i cui punti vendita sono situati su strade urbane ed extraurbane, adottando una politica di prezzi mirata principalmente alla marginalità;
- Enercoop, marchio caratterizzato da una collocazione dei suoi punti vendita nei pressi dei centri commerciali e da una strategia dei prezzi più competitiva rispetto al marchio Vega e con marginalità inferiore;
- Tatanka, presente in un solo punto vendita e avente come unico mezzo di pagamento la monetica, tuttavia esso non viene analizzato in questo elaborato.

Le serie disponibili sono relative all'erogato di carburante per un periodo che va dal 2018 al 2021 e si riferiscono ad ogni singolo punto vendita.

Una prima divisione delle informazioni viene dunque fatta sulla base della tipologia di carburante: benzina, gasolio, dieselmax, gpl, metano e gnl.

Altra informazione fornita dal dataset è la divisione per marchio: Vega o Enercoop. Per ciascun punto vendita, infine, si dispone dell'informazione relativa alla posizione della singola stazione di servizio:

- tipologia di strada: urbana o extra-urbana secondaria;
- vicinanza centro commerciale: un distributore è considerato vicino ad un centro commerciale se si trova ad un raggio minore di 2km.

Per questioni di praticità, in questo elaborato, vengono analizzate soltanto 8 delle 70 stazioni di servizio del Veneto e dell'Emilia-Romagna relative all'erogato della benzina. Le serie che saranno analizzate sono relative all'erogato della benzina del:

- punto vendita di Arino di Dolo (VE);
- punto vendita di Bacone (RE);
- punto vendita di Castel Maggiore (BO);
- punto vendita di Paccagnella (VE);
- punto vendita di Carpi (MO);
- punto vendita di Monselice (PD);
- punto vendita di Forte Tomba (VR);
- punto vendita di Treviso Sud (TV);

	Vicino al centro comm.	Non vicino al centro comm.
Strada urbana	Castel Maggiore (<i>Emilia</i>) Paccagnella (<i>Veneto</i>)	Bacone (<i>Emilia</i>) Arino di Dolo (<i>Veneto</i>)
Strada extra-urbana secondaria	Forte Tomba (<i>Veneto</i>) Treviso Sud (<i>Veneto</i>)	Monselice Volta (<i>Veneto</i>) Carpi (<i>Emilia</i>)

Tabella 2.1: Stazioni di servizio analizzate divise per tipo di strada e vicinanza centro commerciale

2.2 Analisi descrittiva

Le serie relative agli erogati di carburante, nonostante mostrino andamenti lievemente diversi per alcune variabili, presentano caratteristiche generali simili.

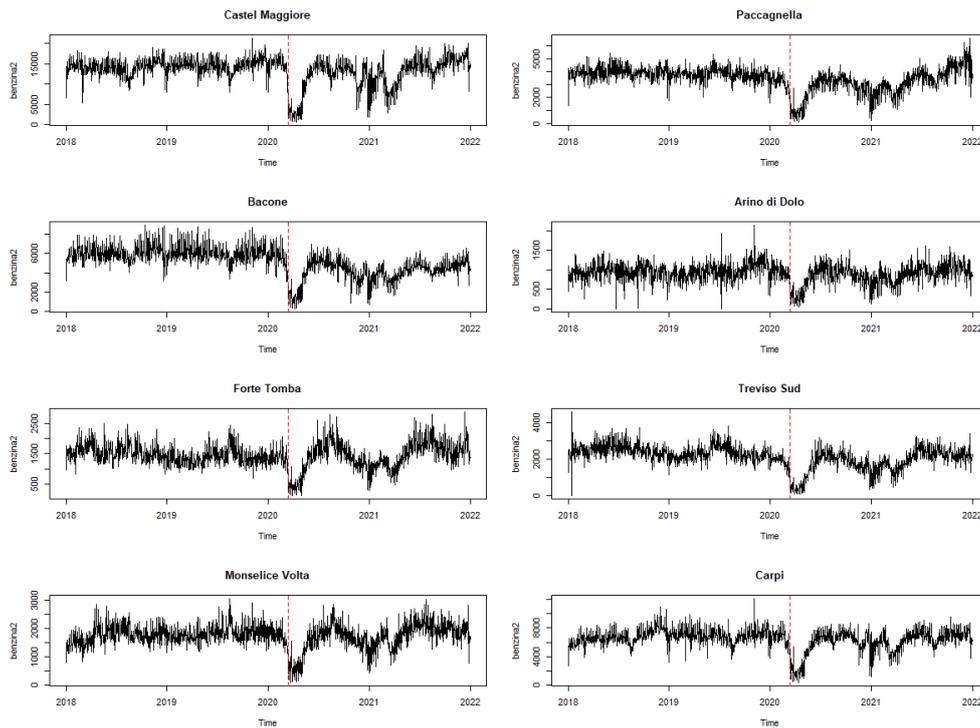


Figura 2.1: Serie storiche degli erogati di benzina delle stazioni di servizio dal 2018 al 2021. La linea rossa indica l'inizio della diffusione del virus.

Dall'analisi visiva emerge la presenza di un picco negativo nei mesi di febbraio e marzo 2020, causato dall'arrivo in Italia del Covid-19, il cui impatto verrà analizzato e approfondito successivamente.

Dal grafico, in Figura 2.1, si possono individuare tre periodi che corrispondono alle tre fasi dell'epidemia Covid: il primo picco negativo corrisponde alla prima ondata, verificatasi a partire da fine febbraio 2020 fino alla fine di maggio 2020 e caratterizzata da una rapidissima diffusione dei casi e dei decessi. Successivamente, il grafico sembra evidenziare una leggera crescita seguita però da un ulteriore calo corrispondente alla seconda ondata relativa a settembre-ottobre 2020. Infine, si nota nei primi mesi del 2021 una riduzione di entità minore alle precedenti, la quale identifica la terza fase del Covid.

I grafici in Figura 2.2 e in Figura 2.3 sembrano evidenziare un andamento settimanale delle serie di interesse relative sia a strade urbane che a strade extra-urbane secondarie.

Nei boxplot, inoltre, in Figura 2.4 e in Figura 2.5 appare evidente la presenza di una

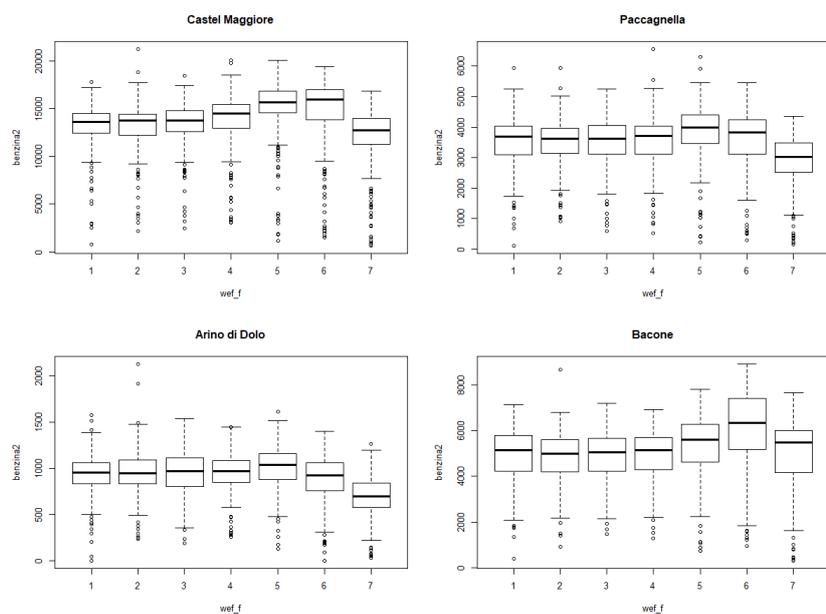


Figura 2.2: Boxplot degli erogati relativi alle stazioni di servizio situate nelle strade urbane per giorni della settimana. La prima riga si riferisce a distributori collocati vicino a centri commerciali, la seconda riga a distributori non vicini a questi ultimi.

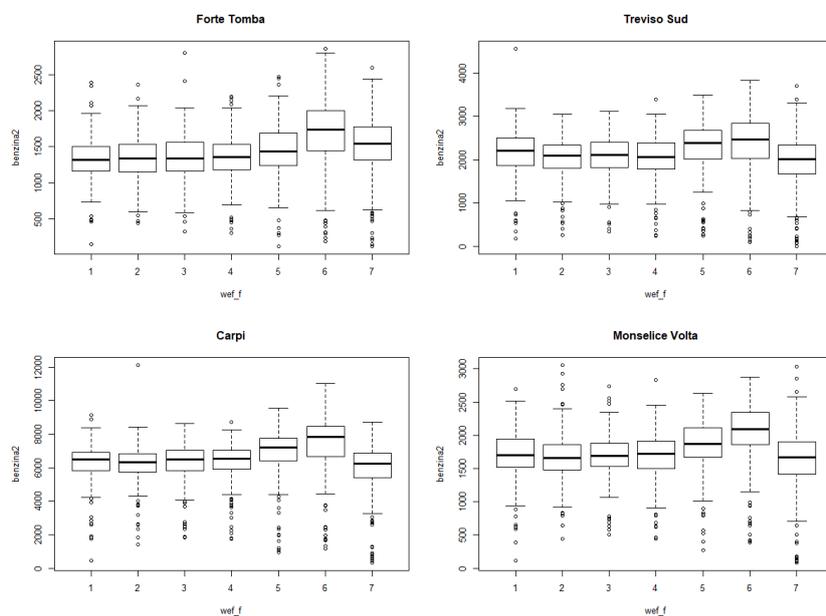


Figura 2.3: Boxplot degli erogati relativi alle stazioni di servizio situate nelle strade extra-urbane secondarie per giorni della settimana. La prima riga si riferisce a distributori collocati vicino a centri commerciali, la seconda riga a distributori non vicini a questi ultimi.

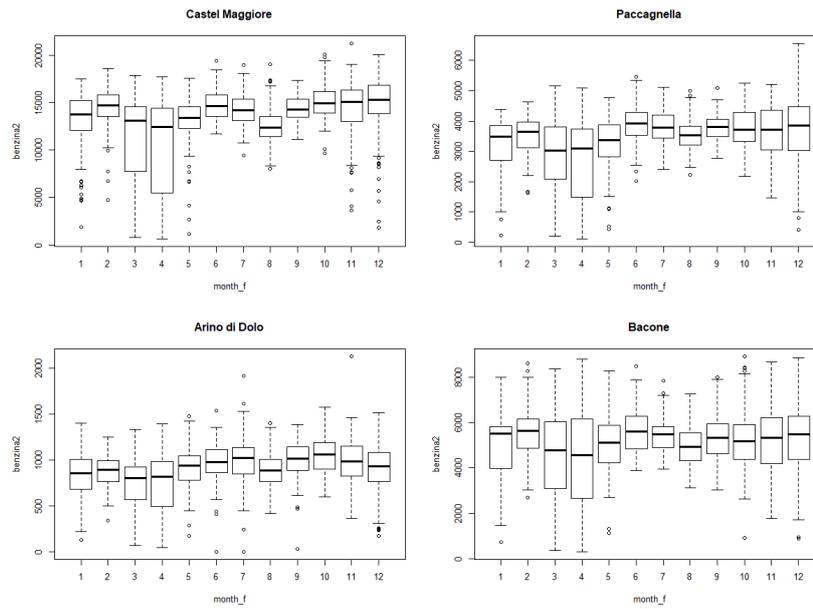


Figura 2.4: Boxplot degli erogati relativi alle stazioni di servizio situate nelle strade urbane per mesi dell'anno. La prima riga si riferisce a distributori collocati vicino a centri commerciali, la seconda riga a distributori non vicini a questi ultimi.

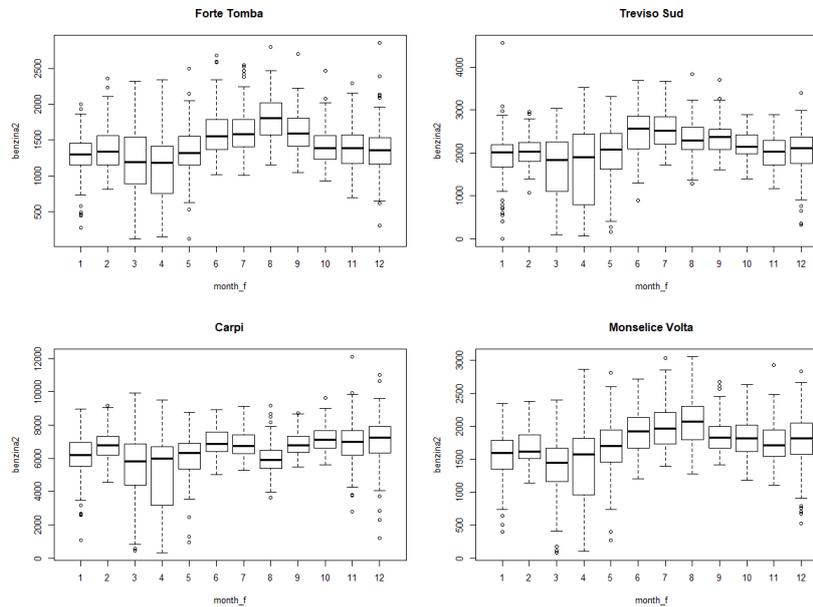


Figura 2.5: Boxplot degli erogati relativi alle stazioni di servizio situate nelle strade extra-urbane secondarie per mesi dell'anno. La prima riga si riferisce a distributori collocati vicino a centri commerciali, la seconda riga a distributori non vicini a questi ultimi.

componente annuale: si osserva come spesso per le strade extra-urbane secondarie nei mesi estivi gli erogati siano maggiori (la mediana è più elevata) rispetto ai mesi invernali. Per le strade urbane, invece, il mese di agosto registra un erogato minore rispetto agli altri mesi (la mediana risulta più bassa).

Capitolo 3

Il modello statistico

Nel capitolo precedente sono state osservate caratteristiche che hanno suggerito l'analisi dell'intervento come strumento per modellare in modo opportuno l'andamento delle serie degli erogati di carburante.

3.1 Modelli per intervento

L'analisi dell'intervento viene utilizzata quando l'evoluzione di un fenomeno è influenzata da fattori che non fanno parte della dinamica intrinseca del fenomeno stesso.

I fattori che possono impattare sull'andamento della serie sono collegati a:

- composizione del calendario;
- interventi esogeni che influenzano l'evoluzione del fenomeno;
- cause esogene di varia natura;
- cause non note o errori.

Le cause della presenza di questi fattori nella serie possono essere note o sconosciute, l'impatto di essi, inoltre, può essere temporaneo o duraturo nel tempo.

Per questo motivo è giusto analizzarli tramite opportuni modelli per interpretare adeguatamente la dinamica della serie.

Nella costruzione di un modello atto a descrivere una serie storica Y_t , essa può essere modellata come:

$$Y_t = X_t + Z_t + \varepsilon_t$$

in cui X_t rappresenta la parte ARIMA e Z_t è l'effetto dell'intervento.

Il procedimento dell'analisi prevede come primo passo la stima del modello ARIMA sulla serie Y_t e il successivo calcolo dei residui.

Sulla serie di questi ultimi si stima l'effetto dell'intervento, cercando di individuare il tipo di impatto che l'intervento ha prodotto sull'evoluzione del fenomeno e verificando il suo livello di significatività. Si controlla, infine, l'adeguatezza dell'impatto al modello osservando la serie finale dei residui.

Esistono diversi tipi di impatto: brusco e monoperiodale, brusco e permanente, brusco e temporaneo oppure un impatto misto, come ad esempio un intervento che produce un impatto che cresce gradualmente e poi tende a smorzarsi fino ad annullarsi.

In generale, supponendo che l'intervento si verifichi al tempo t^* l'effetto dell'intervento Z_t viene costruito tramite una funzione $S_t(t^*)$ che indica la presenza o assenza di un evento.

$$S_t(t^*) = \begin{cases} 0 & \text{se } t < (t^*) \\ 1 & \text{se } t \geq (t^*) \end{cases} \quad (3.1)$$

Considerando ω come impatto iniziale e δ come tasso di estinzione dell'effetto con $0 \leq |\delta| \leq 1$, si definisce:

$$Z_t = \frac{\omega}{1 - \delta B} * (1 - B) * S_t(t^*)$$

In particolare:

- se $\delta = 0$, l'impatto è brusco e monoperiodale;
- se $0 < \delta < 1$ l'impatto è brusco e temporaneo;
- se $\delta = 1$, l'impatto è brusco e permanente;

L'impatto brusco e monoperiodale, in Figura 3.1, prevede che l'intervento produca un impatto immediato che dura un solo periodo corrispondente all'osservazione in cui l'intervento è stato effettuato.

Se l'intervento produce un impatto immediato che dura per più periodi ma con un effetto che diminuisce gradualmente, l'impatto è definito brusco e temporaneo.

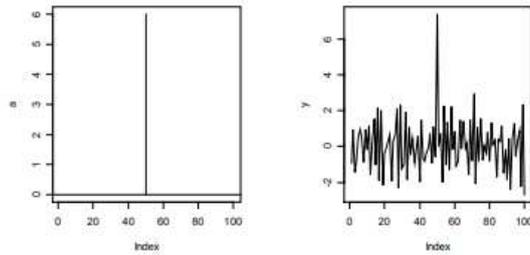


Figura 3.1: Rappresentazione grafica di un impatto brusco e monoperiodale in $t = 50$ e $\omega = 6$

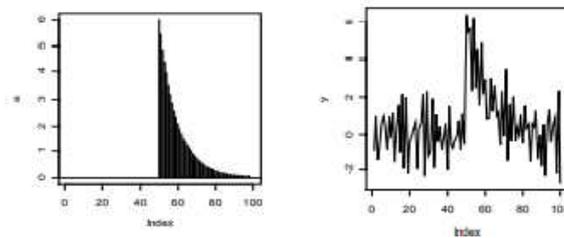


Figura 3.2: Rappresentazione grafica di un impatto brusco e temporaneo in $t = 50$, $\omega = 6$ e $\delta = 0.9$

L'impatto brusco e permanente presuppone che l'intervento produce un effetto immediato che dura per sempre.

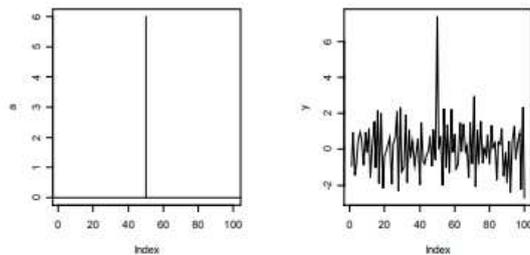


Figura 3.3: Rappresentazione grafica di un impatto brusco e permanente in $t = 50$ e $\omega = 6$

3.2 Costruzione del modello

La dinamica delle serie di interesse suggerisce di indicare come intervento l'inizio della diffusione del Covid-19 in Italia.

Lo studio parte da gennaio 2018 fino ad arrivare a dicembre 2021 per un totale di 1462 osservazioni giornaliere. L'inizio dell'intervento viene posto alla fine di febbraio 2020 ($t = 792$).

In generale, seguendo la procedura descritta nel paragrafo precedente, si cerca di costruire un modello che meglio si adatta alla serie di interesse.

Si considera quindi inizialmente un modello SARIMA $(p, 0, 0)(P, 0, 0)_7$, includendo anche delle variabili esogene quali:

- *yef*: assume valori interi da 1 a 366 e indica il giorno dell'anno;
- *bank*: : variabile dummy che indica se il giorno considerato è festivo;
- *ago*: : variabile dummy che indica se il giorno considerato è nel mese di agosto;

Successivamente, attraverso l'analisi dei residui, si cerca di individuare che tipo di impatto ha avuto la diffusione del virus sulla dinamica dei dati rilevati dopo l'intervento.

Il tipo di impatto che sembra il più adeguato, per questa particolare analisi, è quello misto. Infatti, come si vede anche nelle analisi descrittive, l'intervento provoca nel primo periodo un impatto che inizialmente decresce in modo più o meno rapido e poi tende ad aumentare gradualmente nel tempo.

Per modellare, quindi, l'effetto di breve e lungo periodo dell'intervento del Covid, si introduce ulteriori variabili, le quali assumono valore 0 nel tempo pre-Covid.

Si considera come "periodo Covid" il lasso di tempo che va da fine febbraio 2020 a dicembre 2021.

Le variabili sono:

- *inter3*: variabile quantitativa definita in (3.2);
- *ricoverati*: : variabile quantitativa che indica il numero di ricoverati;
- *terapia*: : variabile quantitativa che indica il numero di ricoverati in terapia intensiva;

Valutando in questa analisi un impatto misto, la variabile *inter3* viene definita nel seguente modo, indicando con t_1 il tempo in cui si raggiunge il valore minimo, di solito posto uguale a 805.

$$Z_t = \begin{cases} \frac{\omega}{1-\delta B} * (1 - B) * S_t(t^*) & \text{se } t < t_1 \\ \delta & \text{se } t = t_1 \\ \frac{\omega}{1-\delta B} * (1 - B) * S_t(t^*) & \text{se } t > t_1 \end{cases} \quad (3.2)$$

Si assume nella prima parte un δ negativo poichè l'effetto in questo periodo considerato è decrescente, successivamente si raggiunge il punto minimo ed infine nella seconda parte si registra una graduale ripresa, ponendo, quindi, un δ positivo.

La variabile descritta sopra entra nella stima del modello per spiegare principalmente l'effetto a breve termine del Covid relativo al primo periodo febbraio 2020-maggio 2020, mentre le variabili *ricoverati* e *terapia* sono state introdotte per modellare le conseguenze dell'intervento a lungo termine.

Il modello finale considerato sarà quindi:

$$Y_t = \begin{cases} \phi_1 * Y_{t-1} + \dots + \phi_p * Y_{t-p} + \Phi_1 * Y_{t-1} + \dots + \Phi_P * Y_{t-P} + \beta_1 * yef + \\ \beta_2 * bank + \beta_3 * ago & \text{se } t < 792 \\ \phi_1 * Y_{t-1} + \dots + \phi_p * Y_{t-p} + \Phi_1 * Y_{t-1} + \dots + \Phi_P * Y_{t-P} + \beta_1 * yef + \\ \beta_2 * bank + \beta_3 * ago + \beta_4 * inter3 + \beta_5 * ricoverati + \beta_6 * terapia & \text{se } t \geq 792 \end{cases} \quad (3.3)$$

Si passa ora ad analizzare le serie di interesse.

Capitolo 4

Analisi delle serie

4.1 Arino di Dolo

La prima serie analizzata è relativa alla stazione di Arino di Dolo, situata in una strada urbana, non vicina ad un centro commerciale.

Seguendo il procedimento descritto nel capitolo precedente, si costruisce un modello SARIMA e si analizza poi la serie dei residui, cercando di evidenziare l'effetto dell'intervento.

Il modello risultato più adatto è un SARIMA $(3, 0, 0)(1, 0, 0)_7$, le cui stime sono riportate in tabella A.1 in Appendice A per una semplice questione di spazio.

La tabella mostra come tutte le stime risultano significative ad eccezione della variabile *ago*, la quale viene rimossa dall'analisi.

Il grafico dei residui del modello adottato è riportato nella Figura 4.1.

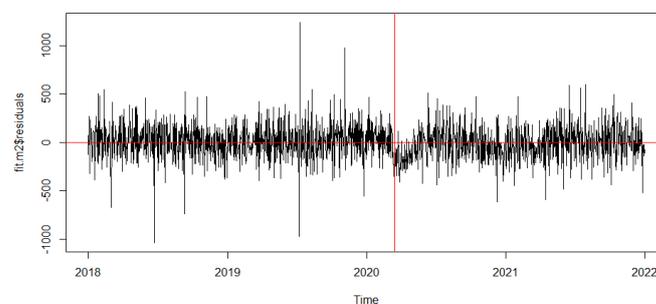


Figura 4.1: Serie dei residui del modello stimato.

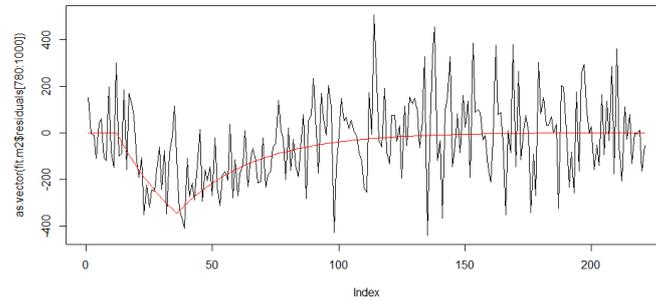


Figura 4.2: Effetto dell'intervento.

Dall'analisi visiva, esso sembra presentare un andamento anomalo nel periodo corrispondente alla diffusione del virus in Italia.

Infatti, concentrandosi su una ridotta finestra temporale, in cui viene messo in evidenza il periodo Covid si nota un picco negativo degli erogati, il quale viene stimato tramite opportuna variabile intervento chiamata *inter3*.

L'impatto risulta significativo: guardando la stima dell'effetto, infatti, l'inizio della diffusione del virus inizialmente porta ad una diminuzione degli erogati giornalieri, mostrando poi successivamente una graduale ripresa.

In particolare, esso ha fatto registrare una diminuzione dell'erogato medio del 38,4% circa rispetto a quello medio pre-intervento.

Modellando, infine, l'effetto del lungo periodo del Covid, la variabile *terapia* non risulta significativa, mentre la variabile *ricoverati* sembra influenzare la dinamica del fenomeno.

	Coefficiente	Standard error	Test t	p-value
ar1	0.1555	0.0266	5.8296	< 0.0001
ar2	0.1239	0.0265	4.6651	< 0.0001
ar3	0.0824	0.0262	3.1409	< 0.0001
sar1	0.1038	0.0266	3.8944	< 0.0001
intercetta	938.3040	17.0071	55.1713	< 0.0001
bank2	-276.5606	13.1912	-20.9654	< 0.0001
yef2	0.4241	0.0761	5.5689	< 0.0001
inter3	1.8200	0.1944	9.3608	< 0.0001
ricoverati	-0.1244	0.0131	-9.4786	< 0.0001

Tabella 4.1: Stima del modello finale SARIMA $(3, 0, 0)(1, 0, 0)_7$ per Arino

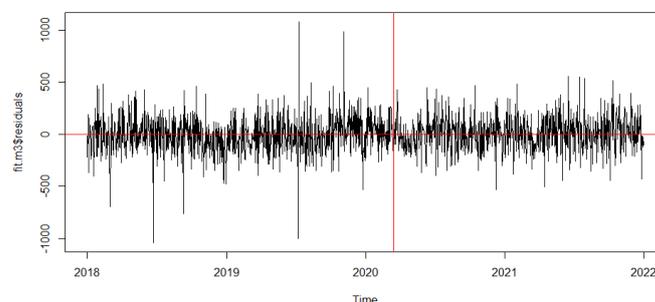


Figura 4.3: Serie dei residui del modello finale.

Complessivamente, rispetto al periodo pre-Covid, l'erogato medio giornaliero relativo al periodo Covid risulta essere diminuito del 15%.

La serie dei residui, al netto dell'impatto stimato, non sembra presentare valori anomali, confermando la conformità dell'effetto individuato.

4.2 Bacone

Si passa ora ad analizzare la serie relativa al punto vendita di Bacone, a Reggio Emilia, posizionato in una strada urbana, non vicino ad un centro commerciale.

Il modello che sembra meglio adattarsi all'intero periodo è un SARIMA $(2, 0, 0)(3, 0, 0)_7$.

Le stime dei parametri, mostrate in tabella A.2 in Appendice, risultano significative ad esclusione della variabile *ago*.

La Figura 4.4 mostra l'andamento dei residui nel quale si può notare uno scostamento nel periodo corrispondente all'inizio della diffusione del virus.

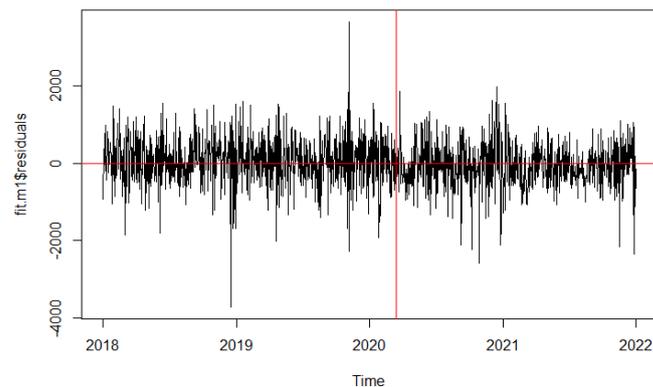


Figura 4.4: Serie dei residui del modello stimato.

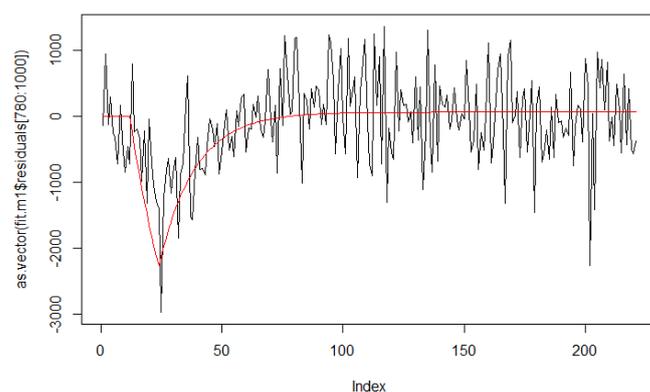


Figura 4.5: Effetto dell'intervento.

Perciò, focalizzando l'attenzione in questo periodo temporale, si studia l'effetto dell'intervento sulla serie dei residui.

Osservando il grafico in Figura 4.5, l'effetto dell'intervento sembra portare ad una diminuzione sempre più contenuta dopo un drastico calo avvenuto nei primi mesi del 2020.

L'intervento ha impattato significativamente sull'andamento degli erogati, facendo registrare relativamente al periodo della prima fase del Covid un calo del 40.3% circa dell'erogato medio giornaliero rispetto alla media pre-intervento.

Considerando, inoltre, la variabile *ricoverati* per descrivere l'effetto relativo al lungo periodo, tutte le stime dei parametri risultano significative come mostrato in tabella

4.2.

	Coefficiente	Standard error	Test t	p-value
ar1	0.3088	0.0271	11.3758	< 0.0001
ar2	0.0563	0.0264	2.1295	0.0332
sar1	0.2911	0.0263	11.0332	< 0.0001
sar2	0.2653	0.0262	10.1153	< 0.0001
sar3	0.225	0.02588	8.6930	< 0.0001
intercetta	5595.3100	144.0092	38.8538	< 0.0001
bank2	-920.9488	85.2582	-10.80186	< 0.0001
yef2	0.9799	0.4458	2.1980	0.0279
ago2	-335.9910	105.8845	-3.1731	< 0.0001
inter3	0.9076	0.1551	5.8488	< 0.0001
ricoverati	-0.8323	0.0643	-12.9363	< 0.0001

Tabella 4.2: Stima del modello finale SARIMA $(2, 0, 0)(3, 0, 0)_7$ per Bacone

In generale, la diminuzione registrata nel periodo Covid corrisponde ad un calo medio giornaliero del 21.9% rispetto alla media pre-Covid.

Il grafico dei residui conferma il buon adattamento dell'effetto individuato.

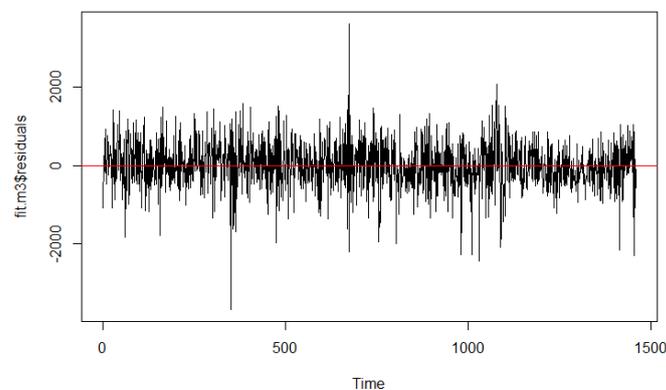


Figura 4.6: Serie dei residui del modello finale.

4.3 Castel Maggiore

Si esamina la serie relativa alla stazione di servizio di Castel Maggiore, situata in una strada urbana e vicina ad un centro commerciale.

Dall'analisi descrittiva del capitolo precedente e da ulteriori analisi, il modello risultato adatto per descrivere la serie è un SARIMA $(2, 0, 0)(3, 0, 0)_7$. Come mostra la tabella A.3 presente in Appendice tutte le variabili risultano significative.

Il grafico dei residui riportato in Figura 4.7 non evidenzia problemi particolari, ad eccezione del periodo relativo ai primi mesi del 2020, il quale viene studiato tramite un'analisi dell'intervento.

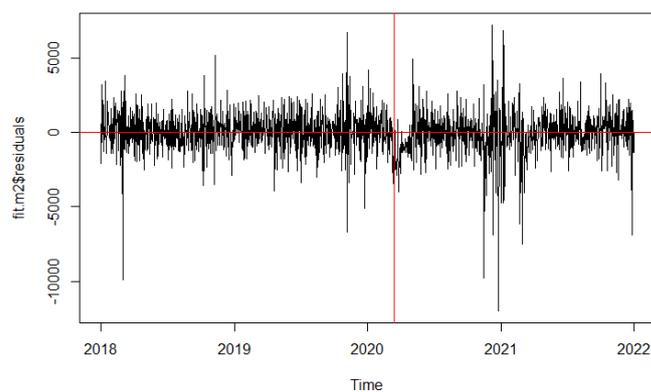


Figura 4.7: Serie dei residui del modello stimato.

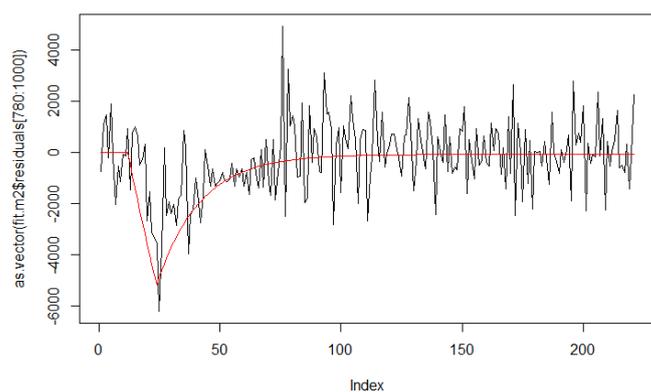


Figura 4.8: Effetto dell'intervento.

Come si può notare dal grafico in Figura 4.8 l'intervento provoca un calo significativo nei primi mesi con una diminuzione complessiva dell'erogato medio giornaliero del 22.4% rispetto alla media pre-Covid.

Successivamente, si introduce nella stima le variabili *ricoverati* e *terapia*, le quali permettono di modellare al meglio l'effetto di lungo periodo relativo al Covid. Tuttavia, soltanto la variabile *ricoverati* risulta significativa, in particolare la media giornaliera media dell'erogato risulta diminuita del 18.6% rispetto alla media pre-covid.

	Coefficiente	Standard error	Test t	p-value
ar1	0.4891	0.0264	18.4968	< 0.0001
ar2	0.0502	0.0259	1.9340	0.053
sar1	0.2998	0.0261	11.4678	< 0.0001
sar2	0.1282	0.0276	4.6386	< 0.0001
sar3	0.1715	0.0267	6.4149	< 0.0001
intercept	0.00014	323.8958	45.0810	< 0.0001
bank2	-2785.486	166.0235	-16.7776	< 0.0001
yef2	4.4702	1.2321	3.6279	< 0.0001
ago2	-1448.064	304.6505	-4.7531	< 0.0001
inter3	1.3139	0.2322	5.6570	< 0.0001
ricoverati	-2.1126	0.1740	-12.14055	< 0.0001

Tabella 4.3: Stima del modello finale SARIMA $(2, 0, 0)(3, 0, 0)_7$ per Castel Maggiore

Il grafico dei residui in Figura 4.9 confermano la conformità dell'effetto individuato, non mostrando alcuna apparente irregolarità.

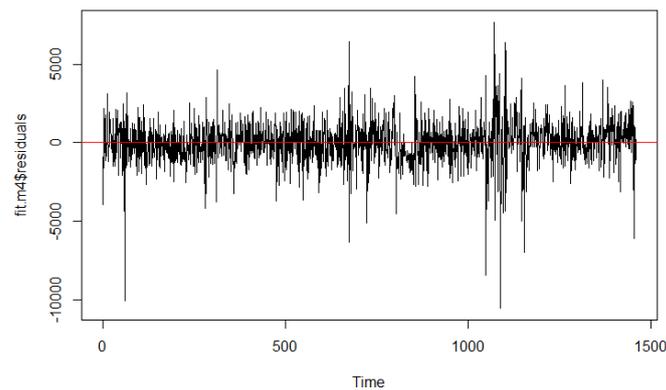


Figura 4.9: Serie dei residui del modello finale.

4.4 Paccagnella

Si analizza la serie giornaliera del punto vendita Paccagnella, a Venezia, collocato in una strada urbana e vicina ad un centro commerciale.

Il modello che sembra adattarsi meglio è un SARIMA $(6, 0, 0)(1, 0, 0)_7$, le cui stime dei parametri, tutte significative ad eccezione della variabile *ago*, sono riportate in tabella A.4 in Appendice A.

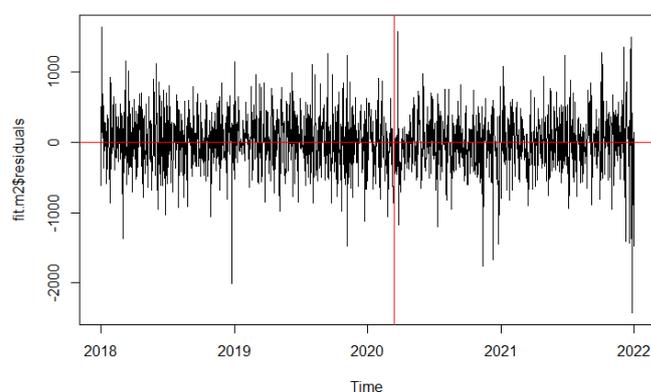


Figura 4.10: Serie dei residui del modello stimato.

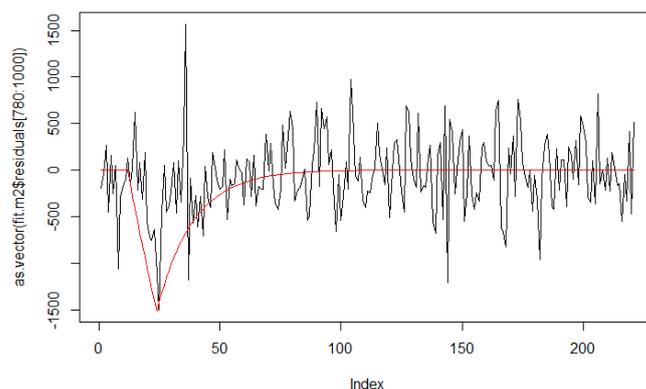


Figura 4.11: Effetto dell'intervento.

Il grafico dei residui riportato in Figura 4.12 non sembra mostrare comportamenti anomali, tuttavia, focalizzando l'attenzione sul periodo relativo ai primi mesi dell'anno 2020, si nota un calo significativo.

Come mostra la Figura 4.11, l'intervento provoca una diminuzione degli erogati e successivamente una lenta ripresa.

Se si guarda, inoltre, la stima dell'effetto, la diffusione del virus in Italia porta a registrare un erogato medio giornaliero pari al 18.5% in meno rispetto al periodo pre-intervento.

	Coefficiente	Standard error	Test t	p-value
ar1	0.31438457	0.02634328	11.934146	< 0.0001
ar2	0.10923327	0.02764839	3.950801	< 0.0001
ar3	0.08643565	0.02751536	3.141359	< 0.0001
ar4	0.10420879	0.02741543	3.801100	< 0.0001
ar5	0.09761669	0.02777088	3.515073	< 0.0001
ar6	0.14263092	0.02720083	5.243623	< 0.0001
sar1	0.15855446	0.02864005	5.536110	< 0.0001
intercetta	3565.719606	126.38383273	28.213416	< 0.0001
bank2	-805.44061985	31.23468156	-25.786740	< 0.0001
yef2	1.59871229	0.43050110	3.713608	< 0.0001
inter3	0.86656314	0.30491526	2.841980	0.0048
terapia	-4.31203311	0.79774471	-5.405279	< 0.0001

Tabella 4.4: Stima del modello finale SARIMA $(6, 0, 0)(1, 0, 0)_7$ per Paccagnella

L'effetto di lungo periodo del covid, infine, viene modellato attraverso l'introduzione della variabile *terapia*, la quale risulta l'unica significativa.

Il coefficiente relativo a quest'ultima entra nel modello con segno negativo, confermando la diminuzione degli erogati anche nel lungo periodo.

In particolare, dalla stima del modello finale si evince che nel periodo Covid si registra un calo di vendite del 18,4% rispetto al periodo precedente.

Il modello finale stimato si adatta adeguatamente; conferma di ciò è il grafico dei residui, il quale oscilla intorno allo zero senza apparente irregolarità.

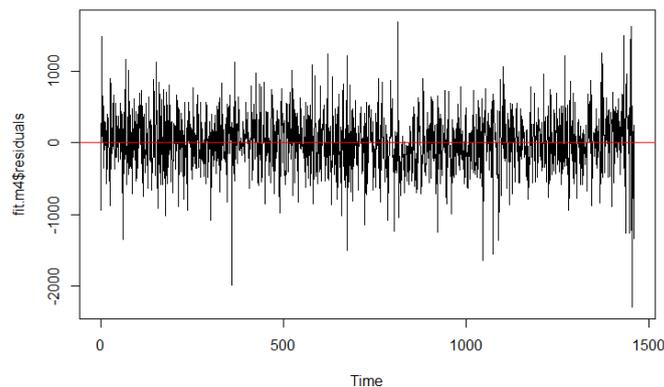


Figura 4.12: Serie dei residui del modello finale.

4.5 Carpi

Analizzando ora la serie storica relativa al punto vendita di Carpi, situata in una strada extra-urbana secondaria e non vicina al centro commerciale, si individua come modello ottimale un SARIMA $(1, 0, 0)(4, 0, 0)_7$, le cui stime sono riportate in tabella A.5 in Appendice A.

Il grafico dei residui in Figura 4.13 del modello adottato mostra un comportamento anomalo: è necessaria l'introduzione della variabile *intervento* che possa spiegare questa "anomalia".

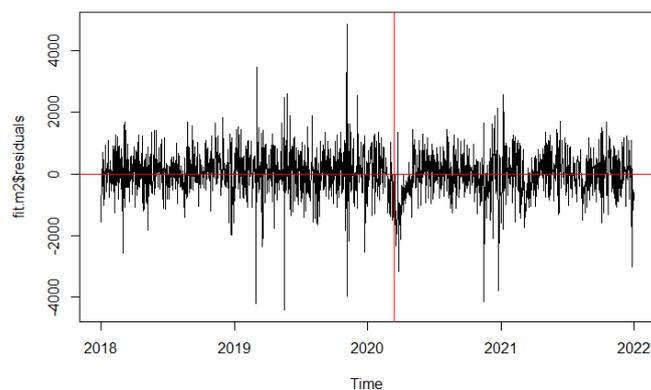


Figura 4.13: Serie dei residui del modello stimato.

L'impatto risulta significativo: guardando il grafico in Figura 4.13 e la stima dell'effetto, l'inizio della diffusione del virus porta nei primi mesi iniziali un calo significa-

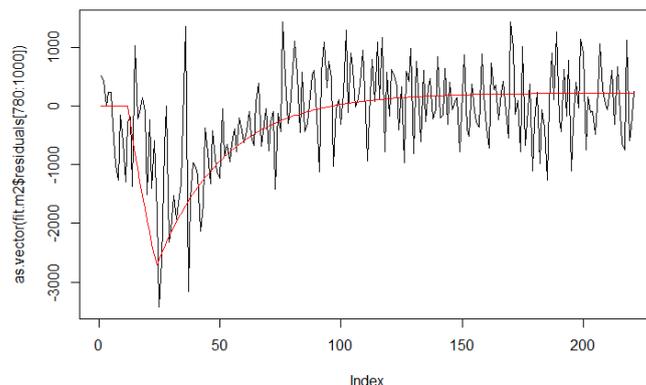


Figura 4.14: Effetto dell'intervento.

tivo e successivamente una progressiva ripresa.

Complessivamente, la diminuzione registrata nel periodo relativo all'intervento corrisponde mediamente ad un calo del 20.7% dell'erogato giornaliero rispetto a quello pre-Covid.

Successivamente, a ciò si aggiunge l'effetto della variabile *ricoverati*, la quale aiuta a spiegare meglio l'effetto del Covid nel lungo periodo.

	Coefficiente	Standard error	Test t	p-value
ar1	0.3571154	0.02553671	13.984392	< 0.0001
sar1	0.2911268	0.02692697	10.811716	< 0.0001
sar2	0.1515068	0.02747150	5.515055	< 0.0001
sar3	0.1355746	0.02733422	4.959887	< 0.0001
sar4	0.1100754	0.02661825	4.135336	< 0.0001
intercept	6742.4164190	141.72262501	47.574736	< 0.0001
bank2	-1356.3485703	100.05356296	-13.556225	< 0.0001
yef2	2.3679458	0.51270818	4.618506	< 0.0001
ago2	-699.2852444	123.02470079	-5.684104	< 0.0001
inter3	0.9043335	0.15630772	5.785597	< 0.0001
ricoverati	-0.8759010	0.07035800	-12.449202	< 0.0001

Tabella 4.5: Stima del modello finale SARIMA $(1, 0, 0)(4, 0, 0)_7$ per Carpi

Essa risulta altamente significativa: nel periodo Covid l'erogato giornaliero mediamente registra una diminuzione del 15,5% circa rispetto a quello precedente.

La serie dei residui, una volta stimato l'effetto dell'intervento, non sembra presentare altri cali significativi nel periodo successivo all'inizio della diffusione del virus, conferma della corretta individuazione dell'impatto.

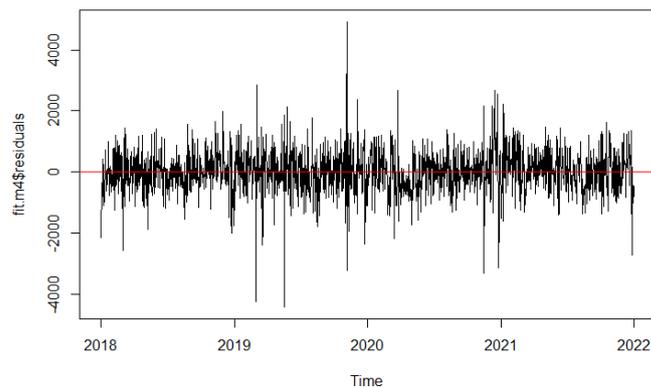


Figura 4.15: Serie dei residui del modello finale.

4.6 Monselice

Si esamina ora la stazione di servizio di Monselice, situata in una strada extra-urbana secondaria e non vicina ad un centro commerciale.

Il modello che sembra meglio adattarsi alla serie è un SARIMA $(3, 0, 0)(2, 0, 0)$, le cui stime sono riportate in tabella A.6 in Appendice A.

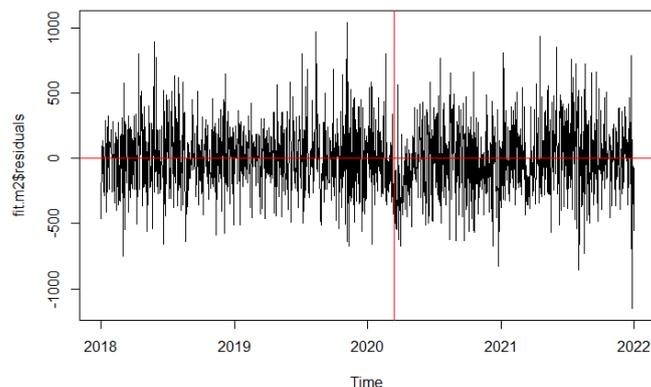


Figura 4.16: Serie dei residui del modello stimato.

La Figura 4.16 mostra l'andamento dei residui nel quale si può evidenziare un notevole picco negativo in corrispondenza dell'inizio della circolazione del virus in Italia.

Si introduce quindi la variabile *intervento* che possa tenere conto di questa "anomalia", come viene mostrata in Figura A.6.

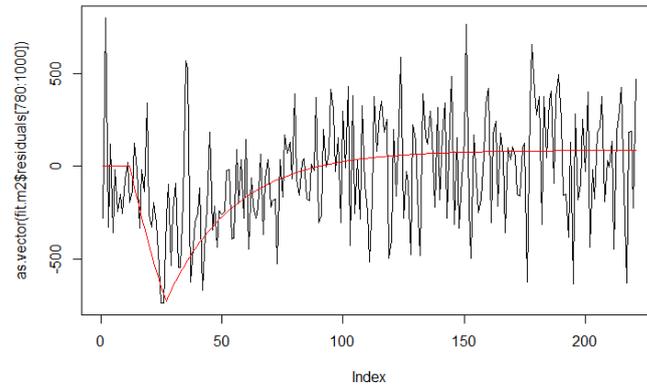


Figura 4.17: Effetto dell'intervento.

La stima dell'effetto dell'intervento sottolinea l'impatto significativo che ha avuto l'inizio della diffusione del virus, in particolare nel periodo considerato si registra una diminuzione media dell'erogato giornaliero pari al 24% in meno rispetto alla media pre-intervento.

Per quanto riguarda, inoltre, l'effetto che ha avuto il Covid nel lungo periodo si aggiunge al modello la sola variabile *ricoverati*, la quale risulta ampiamente significativa.

	Coefficiente	Standard error	Test t	p-value
ar1	0.25302944	0.02678508	9.446656	< 0.0001
ar2	0.08637430	0.02697560	3.201942	0.0013
ar3	0.08170477	0.02660054	3.071546	0.0021
sar1	0.28883279	0.02623702	11.008595	< 0.0001
sar2	0.22735544	0.02651058	8.576027	< 0.0001
intercetta	1719.37218758	43.70445923	39.340887	< 0.0001
bank2	-250.23887676	29.28961237	-8.543605	< 0.0001
yef2	0.64262945	0.18373666	3.497557	< 0.0001
ago2	106.64638094	49.47881611	2.155395	0.0031
inter3	1.50312348	0.21796142	6.896282	< 0.0001
ricoverati	-0.26328387	0.03610068	-7.293045	< 0.0001

Tabella 4.6: Stima del modello finale SARIMA $(3, 0, 0)(2, 0, 0)_7$ per Monselice

Perciò l'erogato medio nell'intero periodo del Covid corrisponde all'11% circa in meno rispetto a quello pre-Covid.

In Figura 4.18 viene riportato il grafico dei residui: esso conferma la conformità dell'effetto individuato e la buona adeguatezza del modello finale.

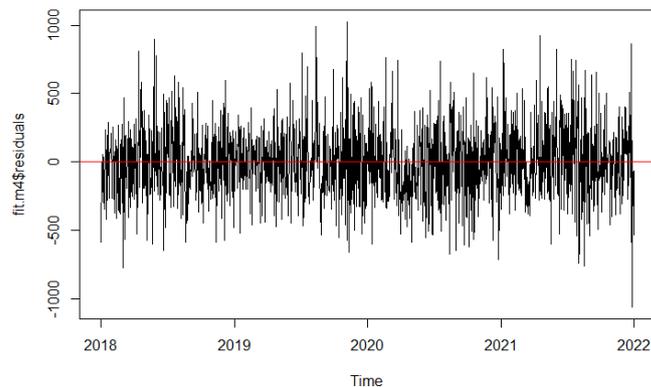


Figura 4.18: Serie dei residui del modello finale.

4.7 Forte Tomba

Si analizza ora la serie relativa al punto vendita Forte Tomba, a Verona, collocato in una strada extra-urbana secondaria e vicino ad un centro commerciale.

Il modello risultato più adatto è un SARIMA $(2, 0, 0)(4, 0, 0)_7$, le cui stime dei parametri, presenti in Appendice A, sono risultate tutte significative ad eccezione della variabile *ago*.

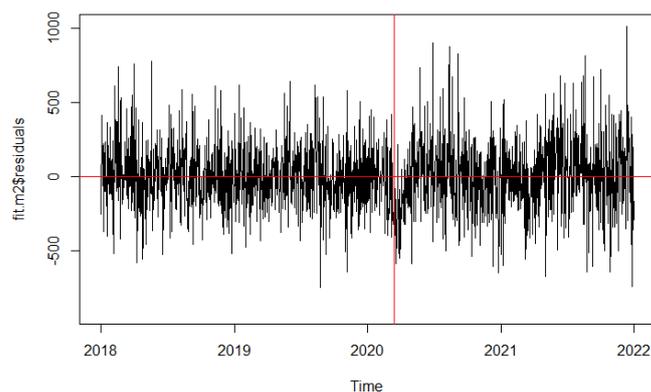


Figura 4.19: Serie dei residui del modello stimato.

Il grafico dei residui, in Figura 4.19, evidenzia uno scostamento dallo zero nel periodo corrispondente all'inizio della diffusione del virus.

Per migliorare la modellazione, si effettua un'analisi dell'intervento che permetta di individuarne l'impatto dell'effetto relativo al covid.

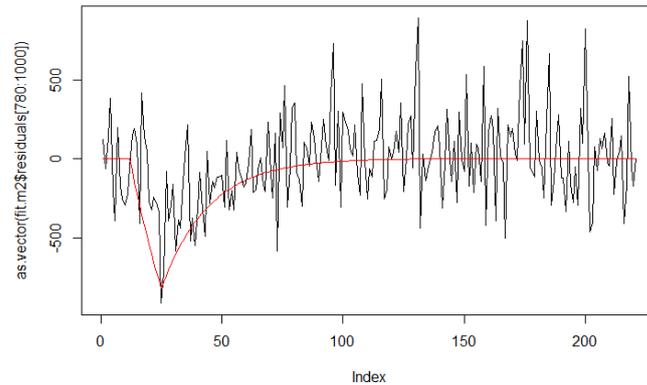


Figura 4.20: Effetto dell'intervento.

Si considera quindi la variabile *intervento*, la quale mostra un'elevata significatività: essa mostra un forte calo nei primi mesi per poi attenuarsi gradualmente.

Pertanto, l'andamento anomalo in questo periodo provoca un calo dell'erogato medio giornaliero del 16,6% circa rispetto al periodo pre-intervento.

Successivamente, introducendo nel modello la variabile *ricoverati*, si riesce a modellare l'effetto relativo al lungo periodo.

	Coefficiente	Standard error	Test t	p-value
ar1	0.2801974	0.02685206	10.434857	< 0.0001
ar2	0.1459405	0.02606450	5.599207	< 0.0001
sar1	0.2857444	0.02695933	10.599091	< 0.0001
sar2	0.1582501	0.02713491	5.831973	< 0.0001
sar3	0.1599372	0.02729106	5.860425	< 0.0001
sar4	0.1035336	0.02665894	3.883634	< 0.0001
intercetta	1479.0556501	51.26087683	28.853499	< 0.0001
bank2	-74.2763001	31.27796011	-2.374717	0.0175
yef2	0.3948106	0.17953908	2.199023	0.0272
inter3	0.7519767	0.18764049	4.007539	< 0.0001
ricoverati	-0.2812260	0.03907841	-7.196454	< 0.0001

Tabella 4.7: Stima del modello finale SARIMA $(2, 0, 0)(4, 0, 0)_7$ per Forte Tomba

In particolare, la diminuzione complessiva registrata nel periodo Covid (febbraio 2020 -dicembre 2021) corrisponde ad un calo dell'erogato medio del 13.7% rispetto a quello pre-Covid. Il grafico finale dei residui non sembra evidenziare valori anomali, evidenza del buon adattamento del modello stimato.

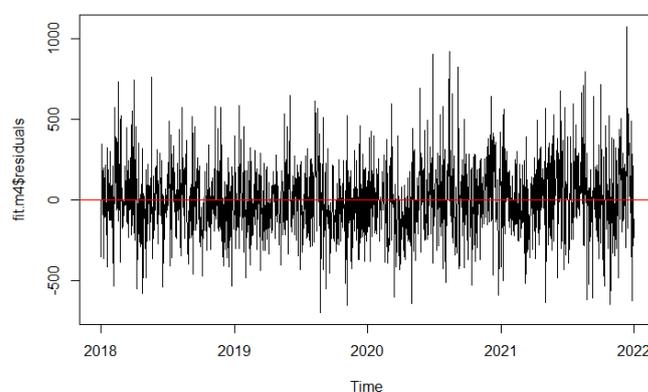


Figura 4.21: Serie dei residui del modello finale.

4.8 Treviso Sud

L'ultima serie analizzata è il punto vendita Treviso Sud, situata in una strada extra-urbana secondaria e vicina al centro commerciale.

Il modello che sembra adattarsi meglio è un SARIMA $(4, 0, 0)(3, 0, 0)_7$, le cui stime sono riportate nella tabella A.8 in Appendice e risultano tutte significative ad eccezione della variabile *ago*.

Nella Figura 4.22 il grafico dei residui mostra un comportamento anomalo da modellare tramite un'analisi dell'intervento. Infatti, si evidenzia qui un drastico calo seguito poi da una lenta ripresa.

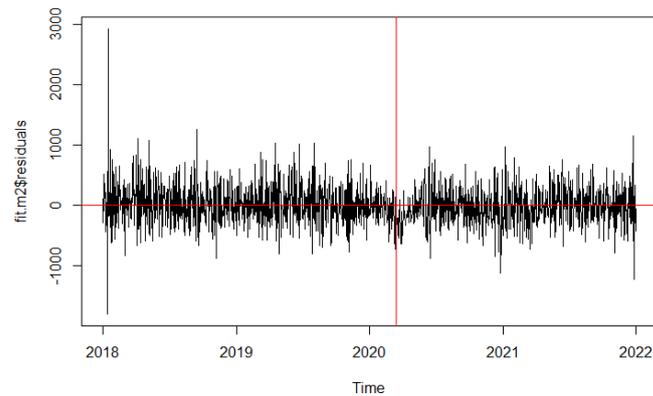


Figura 4.22: Serie dei residui del modello stimato.

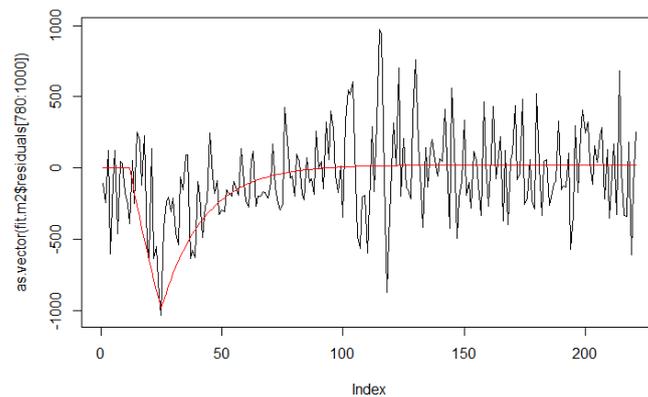


Figura 4.23: Effetto dell'intervento.

La stima dell'intervento risulta significativa: l'inizio della circolazione del virus in Italia provoca nel periodo di interesse una diminuzione dell'erogato medio giornaliero pari al 22,9% rispetto a quello del periodo pre-intervento.

Per modellare, inoltre, l'effetto del lungo periodo del covid si introduce nella stima la variabile *ricoverati*, la quale risulta ampiamente significativa.

	Coefficiente	Standard error	Test t	p-value
ar1	0.2111	0.0276	7.6468	< 0.0001
ar2	0.1085	0.0268	4.0439	< 0.0001
ar3	0.0744	0.0266	2.7905	< 0.0001
ar4	0.1150	0.0263	4.3678	< 0.0001
sar1	0.2220	0.0271	8.1948	< 0.0001
sar2	0.1700	0.0274	6.2097	< 0.0001
sar3	0.2375	0.0271	8.7653	< 0.0001
intercetta	2204.7763	68.0792	32.3854	< 0.0001
bank2	-342.9411	39.6354	-8.6524	< 0.0001
yef2	0.7274	0.2506	2.9028	< 0.0001
inter3	1.0691	0.2309	4.6301	< 0.0001
ricoverati	-0.4883	0.05614	-8.6981	< 0.0001

Tabella 4.8: Stima del modello finale SARIMA $(4, 0, 0)(3, 0, 0)_7$ per Treviso

Complessivamente, la stima complessiva del modello mostra un calo nel periodo Covid dell'erogato medio pari al 22,3% in meno rispetto al periodo pre-Covid.

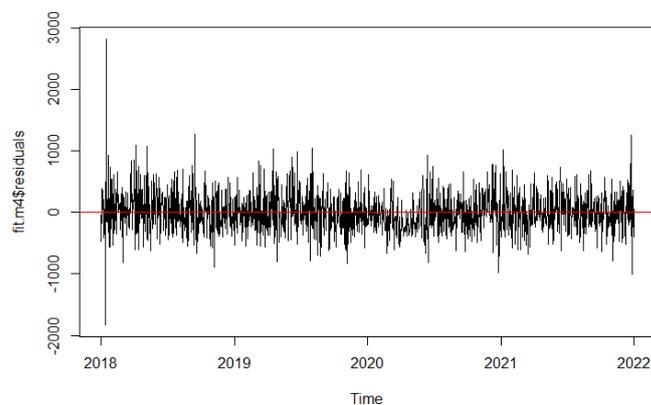


Figura 4.24: Serie dei residui del modello finale.

Il grafico dei residui, mostrato in Figura 4.24, non presenta evidenze anomali e conferma, quindi, il buon adattamento del modello.

Capitolo 5

Conclusione

Lo scopo di questo elaborato è stato quello di studiare l'effetto del Covid-19 sull'andamento degli erogati giornalieri delle stazioni di servizio dell'azienda Vega, attraverso l'analisi di 8 serie relative agli erogati di benzina da gennaio 2018 a dicembre 2021.

Il lavoro è iniziato partendo dall'analisi esplorativa, con la quale sono state evidenziate le proprietà salienti dell'andamento complessivo degli erogati ed è stato individuato un comportamento anomalo nella serie dei residui di ogni modello.

Si è ritenuto opportuno, quindi, andare a studiare, attraverso l'analisi dell'intervento, l'evento che ha inciso sulla dinamica del fenomeno.

Come si è visto, la diffusione del Covid in Italia ha avuto un impatto significativo in tutte le serie analizzate: l'effetto dell'intervento, infatti, è risultato influente nel periodo in cui si è verificato, continuando ad avere rilevanza nei mesi successivi.

In generale, nel breve periodo si registra una diminuzione consistente dell'erogato medio giornaliero rispetto al periodo pre-intervento.

Nella prima fase del Covid, esso ha impattato maggiormente nei punti vendita situati nelle strade urbane, distanti dai centri commerciali con una diminuzione dell'erogato medio giornaliero pari al 40% circa rispetto a quello pre-intervento.

Meno impattante ma pur sempre significativo è l'effetto del Covid nel lungo periodo, nel quale si verifica una vendita media giornaliera che oscilla tra l'11% e il 22% circa in meno rispetto a quello pre-Covid.

Le conseguenze dell'intervento a lungo termine hanno influito in maggior misura

sugli erogati di benzina delle stazioni di servizio collocati vicino a centri commerciali, colpiti duramente dai provvedimenti governativi che ne hanno imposto la chiusura per diversi mesi.

Complessivamente, la diffusione del virus non sembra aver avuto un diverso tipo di effetto sull'andamento degli erogati nelle 8 serie esaminate, tuttavia, avendo analizzato un campione relativamente ristretto, non si è potuto sviluppare in modo approfondito un confronto tra i diversi punti vendita che differiscono tra loro per alcune variabili, quali la vicinanza ad un centro commerciale oppure la tipologia di strada in cui essi sono situati.

Queste conclusioni, perciò, andranno verificate in seguito ad un'analisi comprendente un maggior numero di stazioni, la quale consentirà di stimare accuratamente l'effetto della diffusione del Covid sull'erogato giornaliero di carburante.

Appendice A

Appendice

	Coefficiente	Standard error	Test t	p-value
ar1	0.2484	0.02651	9.371	< 0.0001
ar2	0.2134	0.02672	7.9840	< 0.0001
ar3	0.1761	0.02618	6.7275	< 0.0001
sar1	0.1271	0.02858	4.4460	< 0.0001
intercetta	864.5819	28.2786	30.5737	< 0.0001
bank2	-275.9084	13.2317	-20.8521	< 0.0001
yef2	0.5032	0.1287	3.9080	< 0.0001

Tabella A.1: Stima del modello SARIMA $(3, 0, 0)(1, 0, 0)_7$ per Arino di Dolo

	Coefficiente	Standard error	Test t	p-value
ar1	0.3965	0.02761	14.3596	< 0.0001
ar2	0.1605	0.02611	6.1492	< 0.0001
sar1	0.3304	0.02768	11.9377	< 0.0001
sar2	0.2907	0.02635	11.0313	< 0.0001
sar3	0.1965	0.02609	7.5308	< 0.0001
intercetta	5011.9538	229.2876	21.8588	< 0.0001
bank2	-894.0181	86.0740	-10.3866	< 0.0001
yef2	1.4562	0.6150	2.3676	0.0180

Tabella A.2: Stima del modello SARIMA $(2, 0, 0)(3, 0, 0)_7$ per Bacone

	Coefficiente	Standard error	Test t	p-value
ar1	0.5704	0.02737	20.8367	< 0.0001
ar2	0.1423	0.02613	5.4439	< 0.0001
sar1	0.32904	0.02801	11.7444	< 0.0001
sar2	0.1449	0.02786	5.2027	< 0.0001
sar3	0.1712	0.02627	6.5187	< 0.0001
intercetta	12968.1131	496.2966	26.1298	< 0.0001
bank2	-2811.7348	171.8224	-16.3642	< 0.0001
yef2	5.7947	1.7351	3.3394	< 0.0001
ago2	-1030.3155	381.0102	-2.7042	0.0068

Tabella A.3: Stima del modello SARIMA (2, 0, 0)(3, 0, 0)₇ per Castel Maggiore

	Coefficiente	Standard error	Test t	p-value
ar1	0.33203	0.02624	12.6490	< 0.0001
ar2	0.1229	0.02773	4.4310	< 0.0001
ar3	0.09906	0.02763	3.5848	< 0.0001
ar4	0.1146	0.02751	4.1661	< 0.0001
ar5	0.1057	0.02781	3.8016	< 0.0001
ar6	0.1513	0.0267	5.6490	< 0.0001
sar1	0.1559	0.0280	5.5667	< 0.0001
intercetta	3338.3989	194.5178	17.1624	< 0.0001
bank2	-802.1019	31.01604	-25.8608	< 0.0001
yef2	1.54821	0.4731	3.2723	< 0.0001

Tabella A.4: Stima del modello SARIMA (6, 0, 0)(1, 0, 0)₇ per Paccagnella

	Coefficiente	Standard error	Test t	p-value
ar1	0.4906	0.02610	18.7989	< 0.0001
sar1	0.3564	0.02874	12.40089	< 0.0001
sar2	0.1821	0.02810	6.4793	< 0.0001
sar3	0.119	0.02783	4.2762	< 0.0001
sar4	0.0608	0.02662	2.2844	< 0.0001
intercetta	6231.911	191.3889	32.5615	< 0.0001
bank2	-1352.579	100.2088	-13.4976	< 0.0001
yef2	2.6077	0.6905	3.7763	< 0.0001
ago2	-508.1343	157.9617	-3.2168	< 0.0001

Tabella A.5: Stima del modello SARIMA (1, 0, 0)(4, 0, 0)₇ per Carpi

	Coefficiente	Standard error	Test t	p-value
ar1	0.2894	0.02727	10.6093	< 0.0001
ar2	0.1306	0.02694	4.8493	< 0.0001
ar3	0.1343	0.02640	5.0870	< 0.0001
sar1	0.3191	0.02681	11.9018	< 0.0001
sar2	0.2646	0.02593	10.2043	< 0.0001
intercetta	1646.2889	58.5147	28.1346	< 0.0001
bank2	-259.19022	31.3637	-8.2640	< 0.0001
yef2	0.7358	0.2363	3.1128	< 0.0001
ago2	97.4642	57.4395	1.6968	0.0897

Tabella A.6: Stima del modello SARIMA $(3, 0, 0)(2, 0, 0)_7$ per Monselice

	Coefficiente	Standard error	Test t	p-value
ar1	0.3216	0.0270	11.9071	< 0.0001
ar2	0.1903	0.0260	7.3021	< 0.0001
sar1	0.3038	0.0276	10.9745	< 0.0001
sar2	0.1703	0.0273	6.2279	< 0.0001
sar3	0.1610	0.0274	5.873	< 0.0001
sar4	0.0905	0.0267	3.3843	< 0.0001
intercetta	1380.3732	61.0540	22.6090	< 0.0001
bank2	-78.9854	31.7354	-2.4889	0.0128
yef2	0.3543	0.2098	1.6886	0.0912

Tabella A.7: Stima del modello SARIMA $(2, 0, 0)(4, 0, 0)_7$ per Forte Tomba

	Coefficiente	Standard error	Test t	p-value
ar1	0.2533	0.0278	9.0812	< 0.0001
ar2	0.1485	0.0269	5.5198	< 0.0001
ar3	0.1157	0.0267	4.3350	< 0.0001
ar4	0.1607	0.0263	6.1064	< 0.0001
sar1	0.2238	0.0278	8.0429	< 0.0001
sar2	0.1837	0.0276	6.6524	< 0.0001
sar3	0.2450	0.02701	9.0722	< 0.0001
intercetta	2038.9767	97.1362	20.9901	< 0.0001
bank2	-341.8235	40.4983	-8.4404	< 0.0001
yef2	0.6842	0.3194	2.1422	0.0321

Tabella A.8: Stima del modello SARIMA $(4, 0, 0)(3, 0, 0)_7$ per Treviso Sud

Bibliografia

- [1] Di Fonzo Tommaso & Lisi Francesco. (2007). Serie storiche economiche. Carocci
- [2] William W.S. Wei. (2006). Time Series Analysis. Pearson