



# **Analisi Indice di produzione industriale manifatturiera 2003-2023**

## **Gruppo 5**

Adamo Lorenzo, Baldan Giorgia, Del Giudice Xavier, Lipari Federica

4 Maggio 2023

# 1 Introduzione

”L’indice generale della produzione industriale misura la variazione nel tempo del volume fisico della produzione effettuata nel settore dell’industria in senso stretto (ovvero con esclusione delle costruzioni). Esso si basa su una rilevazione statistica condotta mensilmente presso le imprese, che forniscono informazioni dettagliate riguardo alla produzione di specifici prodotti.” [Fonte: Istat, 19 marzo 2018]

Si considera quindi la serie storica mensile destagionalizzata dell’indice di produzione industriale manifatturiera in Italia negli ultimi 20 anni, a partire da Febbraio 2003, con base di riferimento l’anno 2015, al fine di prevedere i valori dell’indice per l’anno 2023.

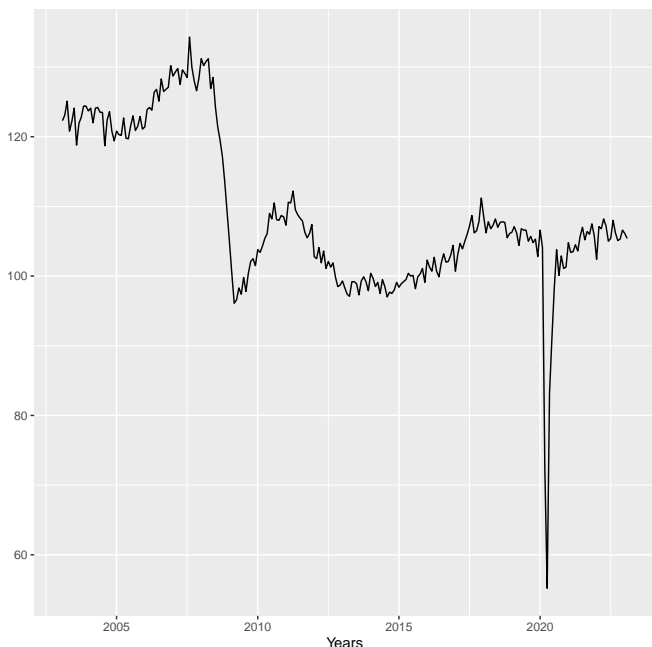
## 2 Analisi

Osservando la serie rappresentata in Figura 1(a), saltano subito all’occhio due periodi di intensa flessione e uno più lieve. Il primo ha avuto luogo negli anni 2008-2009, dovuto alla crisi economica che ha colpito l’Italia in quel periodo. Successivamente si verifica un moderato recupero fino al 2011, anno in cui si verifica il periodo di lieve flessione che si esaurisce nei due anni successivi. L’ultimo picco negativo coincide con la fase più drastica ed ha luogo nel 2020 a causa delle misure sanitarie connesse alla pandemia. Si nota inoltre che la serie non è generata da un processo stocastico stazionario in media in quanto i valori osservati non oscillano intorno al valore medio.

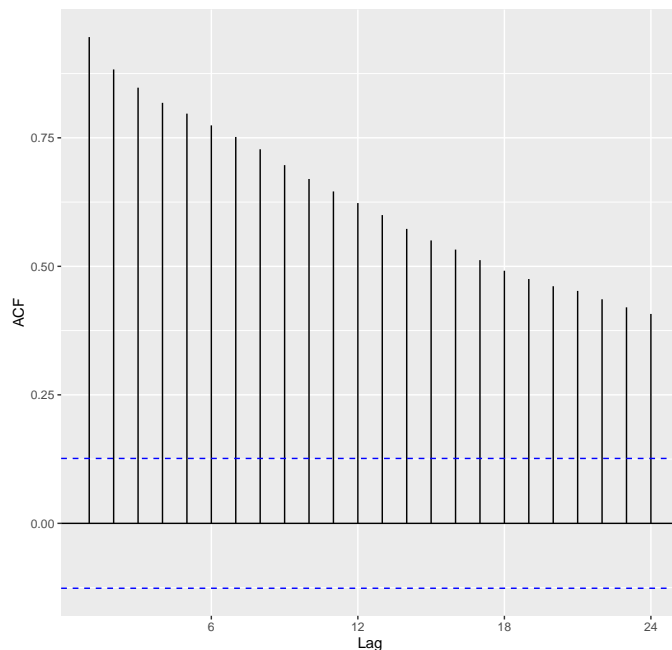
Ulteriore conferma la si può avere osservando l’autocorrelogramma che rappresenta graficamente le autocorrelazioni ossia correlazioni tra la variabile e la stessa ritardata di  $K$  periodi. Infatti come si può notare in Figura 1(b) le autocorrelazioni decrescono lentamente verso lo 0 al crescere di  $K$ , in caso contrario le autocorrelazioni sarebbero decresciute molto più velocemente.

Si pone l’obiettivo di individuare un processo stocastico stazionario, dal quale la serie storica integrata è stata generata; infatti la stazionarietà è una condizione necessaria per essere in grado di risalire al processo stocastico generatore e quindi identificare il miglior modello ARIMA, che ci permette di studiare la componente non deterministica del processo derivata dal Teorema di scomposizione di Wold in base al quale un processo stocastico stazionario ( $Y_t$ ) può essere scomposto nella somma di due componenti incorrelate: la componente deterministica ( $V_t$ ) e non deterministica ( $Z_t$ ). Escludendo che la non stazionarietà sia dovuta alla presenza di un processo trend-stazionario, in quanto non risulta esserci un trend rilevante, si suppone la presenza di un processo stazionario alle differenze. Di conseguenza si provvede ad una differenziazione della serie per rendere il processo stocastico stazionario:

$$Y_t = Y_{t-1} + c_t, \quad \text{con} \quad c_t \sim \text{ARMA}(p,q)$$

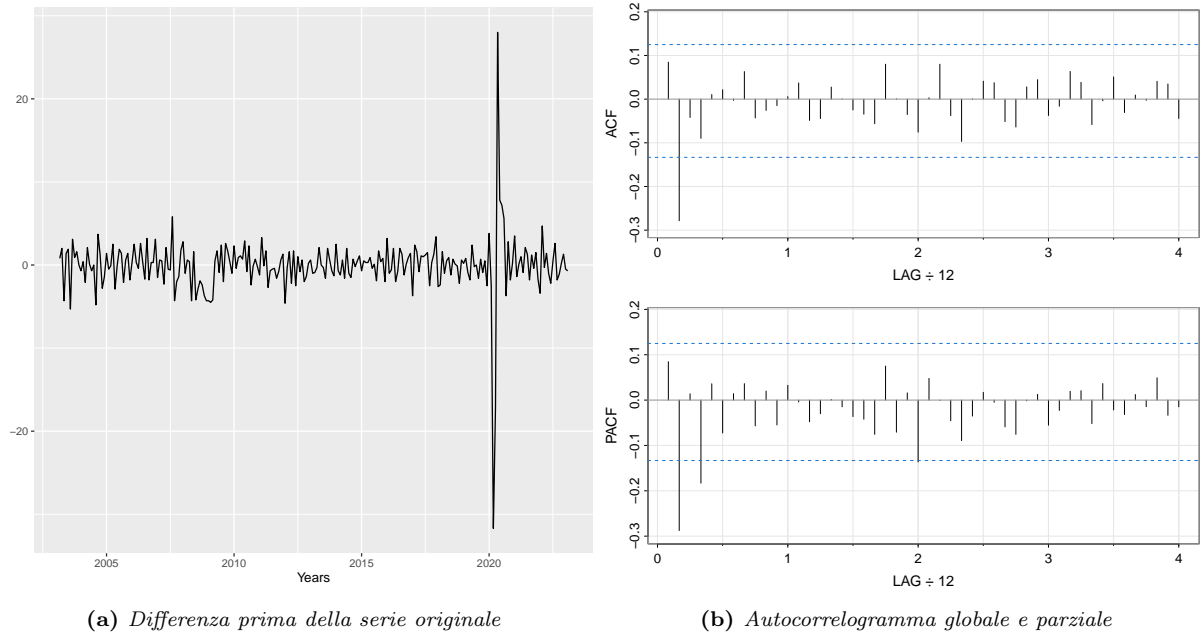


(a) Serie originale



(b) Autocorrelogramma globale e parziale

**Figura 1:** Serie storica mensile dell’indice di produzione industriale dell’industria manifatturiera



**Figura 2:** Serie storica integrata di ordine 1

Si può notare in Figura 2 come il processo integrato di ordine 1 sia stazionario. Osservando, quindi l'autocorrelogramma globale e parziale si può pensare che sia un ARIMA(0,0,2) in quanto nell'autocorrelogramma globale l'unica autocorrelazione al di fuori della banda di Bartlett (rappresentata in figura dall'intervallo delimitato dalle linee blu tratteggiate) e quindi significativamente diversa da 0, è quella di ordine 2.

Inoltre nell'autocorrelogramma parziale, al crescere di  $k$  le autocorrelazioni tendono esponenzialmente a 0. Di seguito la rappresentazione di un modello ARIMA(0,0,2):

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2}$$

Si stimano ora i coefficienti del modello:

ARIMA(0,0,2) with zero mean

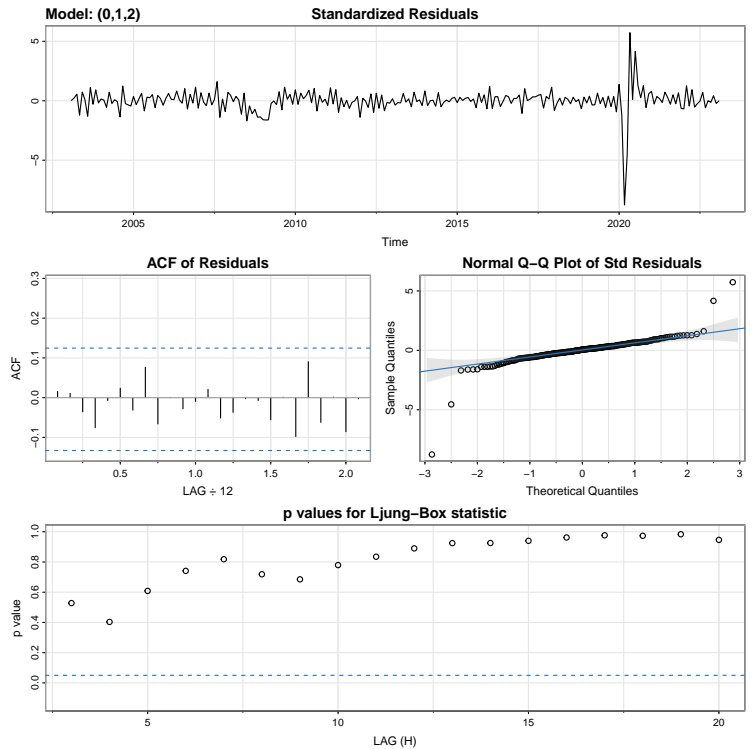
Coefficients:

	ma1	ma2
	0.1102	-0.3508
s.e.	0.0619	0.0623

sigma^2 = 11.49: log likelihood = -632.64  
AIC=1271.28 AICc=1271.38 BIC=1281.72

Il risultato ottenuto conferma che il processo stocastico generatore dei dati è effettivamente un ARIMA(0,0,2). Si può notare che il coefficiente  $\psi_1$  risulta statisticamente non significativo, come supposto precedentemente nell'interpretazione dell'autocorrelogramma. Si può quindi concludere che il miglior modello sulla serie iniziale destagionalizzata è un ARIMA(0,1,2).

Viene valutata, adesso, la validità del modello stimato. Si definisce una serie storica dei residui, rappresentata in Figura 3, e si verifica che sia generata da un processo stocastico White Noise, ossia una successione di variabili casuali incorrelate, con media nulla e varianza costante nel tempo. Il modello risulta valido, infatti come si può notare in Figura 3, il processo è stazionario in media con media pari a 0 e le autocorrelazioni sono tutte non significativamente diverse da 0.



**Figura 3:** Diagnostica modello

### 3 Previsione

A questo punto sono stati stimati i valori dell'indice della produzione industriale dell'industria manifatturiera dal mese di Febbraio 2023 fino a Dicembre 2023. Dato il modello ARIMA(0,1,2) come migliore rappresentazione del processo stocastico generatore dei dati si provvede ad una previsione puntuale del tipo:

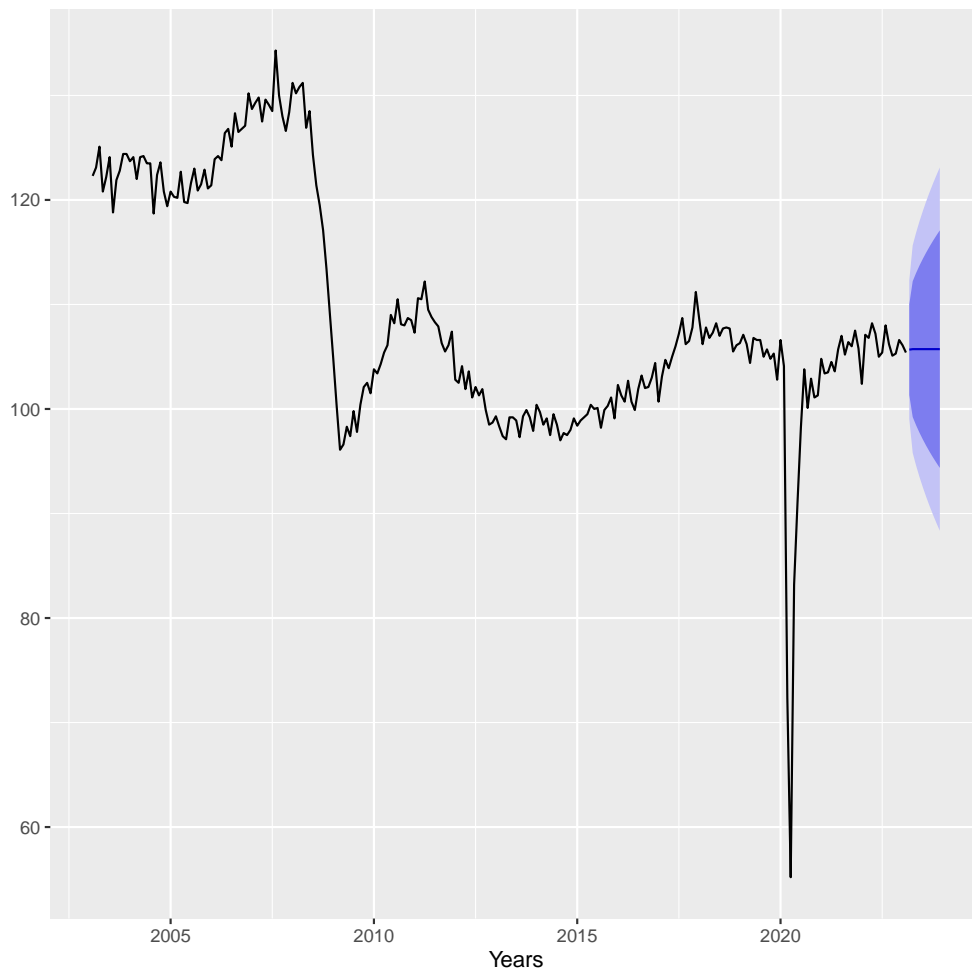
$$E(Z_{T+h}|I_T), \quad \text{dove} \quad Z_{T+h} = a_{T+h} + \psi_1 a_{T-1+h} + \psi_2 a_{T-2+h}$$

Di seguito sono riportati i valori delle differenze previste:

	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2023	0.2774	0.0487	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

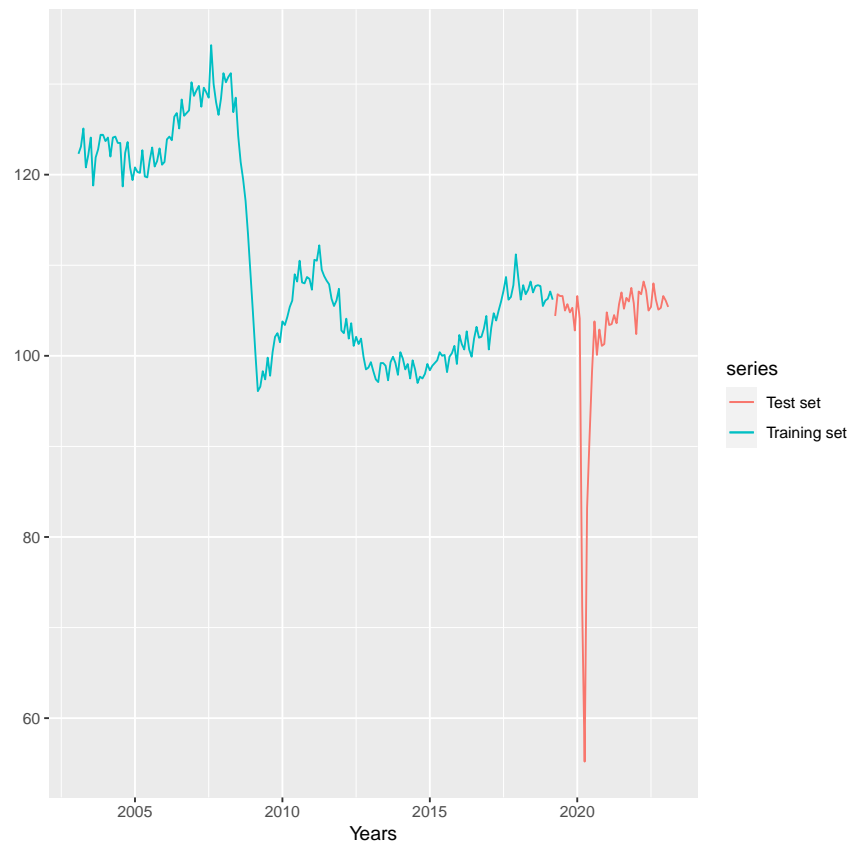
Essendo un modello ARIMA(0,1,2) con media nulla, solamente le previsioni fino a 2 passi in avanti sono diverse da 0, mentre le restanti sono pari alla media. Quindi per ottenere i valori previsti per ogni mese  $t$ , è necessario sommare il valore corrispondente al mese  $t-1$  con il valore delle differenze previsto al mese  $t$ . Le previsioni ottenute, rappresentate in colore blu in Figura 4, sono le seguenti:

	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
2023	105.6774	105.7262	105.7262	105.7262	105.7262	105.7262	105.7262	105.7262
	Nov	Dec						
2023	105.7262	105.7262						



**Figura 4:** *Previsione per Marzo-Dicembre 2023 dell'indice di produzione industriale manifatturiera*

## 4 Valutazione modello previsivo



**Figura 5:** *Partizione della serie in training e test set*

Per la valutazione del modello previsivo è stata fatta una partizione della serie in due parti, come rappresentato in Figura 5: training set e test set. Il training set è composto dall'80% dei valori osservati della serie iniziale, quindi da Febbraio 2003 fino a Marzo 2019, mentre il test set è formato dai restanti valori osservati.

Viene simulata la prestazione del modello senza dover aspettare il verificarsi della previsione stessa. Dunque, si impone il modello stimato precedentemente, quindi  $ARIMA(0,1,2)$ , ai dati presenti nel training test. Tramite questo vengono effettuate delle previsioni che saranno poi confrontate con i valori realmente osservati presenti nel test set. Vengono quindi calcolati diversi indici di accuratezza, basati sugli errori di previsione, ovvero la differenza tra i valori osservati nella finestra temporale del test set e quelli previsti dal training set. Di conseguenza, l'ideale sarebbe ottenere valori molto piccoli. I valori di questi indici sono riportati di seguito:

	RMSE	MAE	MAPE
Training set	1.8850	1.4720	1.3257
Test set	2.0271	2.0271	1.9417

L'indice RMSE è la radice dell'errore quadratico medio, il MAE invece rappresenta l'errore medio assoluto. Entrambi questi indici di accuratezza dipendono dalla scala metrica utilizzata. Il MAPE, l'errore percentuale, è invece adimensionale, cioè non dipende dalla scala metrica utilizzata (scale-free). Si può notare che i valori degli indici di accuratezza relativi al Test set sono di poco superiori di quelli relativi al Training set. Ciò è tipico in quanto il modello utilizza il training set per fare le previsioni, di conseguenza è normale che performi leggermente meglio. La porzione dei dati utile per valutare la bontà del modello previsivo è il Test set. Di conseguenza, siccome i valori degli indici di accuratezza relativi al Test set sono piccoli, il modello  $ARIMA(0,1,2)$  calcola delle previsioni accurate.

## 5 Conclusione

Le previsioni dell'indice di produzione manifatturiera per l'anno 2023 in Italia risultano essere affidabili, a meno di shock dovuti a eventi imprevisi, impossibili da pronosticare per il modello di previsione. Stando alle previsioni effettuate, si può concludere che l'indice di produzione industriale rimarrà sugli stessi livelli raggiunti nel periodo post pandemia.