

# Introduzione

Il modello **SARIMA** (Seasonal AutoRegressive Integrated Moving Average) è un'estensione dei modelli ARIMA progettata per gestire efficacemente dati con componenti stagionali. In ambito aziendale, SARIMA è ampiamente utilizzato per effettuare previsioni su serie temporali (ad esempio vendite mensili, domanda stagionale di prodotti, ecc.), integrando trend, stagionalità e dipendenze temporali. Nel seguito forniremo un approfondimento teorico sul modello SARIMA e sulle sue applicazioni business-critical, con un linguaggio chiaro pensato per professionisti senza competenze avanzate di statistica.

## 1. Fondamenti teorici del modello SARIMA

Un modello ARIMA cerca di spiegare l'evoluzione di una serie temporale attraverso tre componenti: **autoregressione (AR)**, **integrazione (I)** e **media mobile (MA)**. In pratica: (a) la parte AR indica che il valore attuale dipende in modo lineare dai valori passati (fino a  $p$  ritardi); (b) la parte I indica che la serie è stata differenziata  $d$  volte per rimuovere trend e renderla stazionaria; (c) la parte MA indica che il valore attuale dipende anche dagli errori di previsione commessi nei  $q$  passi temporali precedenti. Un modello ARIMA è quindi notato come  $ARIMA(p, d, q)$  dove  $p, d, q$  sono interi  $\geq 0$  che rappresentano rispettivamente l'ordine autoregressivo, il grado di differenziazione (trend) e l'ordine della media mobile.

Tuttavia, un semplice ARIMA non contempla esplicitamente la **stagionalità**, cioè variazioni periodiche ricorrenti. Quando i dati presentano pattern stagionali (ad esempio picchi annuali delle vendite durante le festività), si ricorre al modello **SARIMA**, indicato come  $ARIMA(p, d, q)(P, D, Q)_m$ . In questa notazione, oltre ai parametri  $p, d, q$  già descritti, compaiono i parametri stagionali  $P, D, Q$  (anch'essi interi  $\geq 0$ ) che rappresentano rispettivamente l'ordine autoregressivo stagionale, il grado di differenziazione stagionale e l'ordine della media mobile stagionale. Il parametro  $m$  indica la lunghezza del periodo stagionale (es.  $m = 12$  per dati mensili con stagionalità annuale,  $m = 4$  per dati trimestrali con ciclo annuale, ecc.). In sostanza, il SARIMA aggiunge al modello ARIMA tradizionale termini che considerano i valori e gli errori con ritardo di  $m$  passi temporali e consente di differenziare la serie ogni  $m$  osservazioni per rimuovere la componente periodica. Ad esempio, se  $m = 12$  (stagionalità annuale su dati mensili), SARIMA può includere un termine autoregressivo stagionale che usa il valore di 12 mesi prima e un differenziamento stagionale (ordine  $D = 1$ ) che sottrae il valore di un anno prima per eliminare sistematicamente gli effetti annuali ricorrenti.

Il ruolo della **stagionalità** in SARIMA è quindi fondamentale: il modello cerca di "catturare" pattern che si ripetono regolarmente ogni  $m$  periodi. Senza un componente stagionale esplicito, un normale ARIMA faticerebbe a modellare questi cicli periodici e potrebbe richiedere un ordine AR o MA molto elevato per approssimarli. SARIMA risolve il problema integrando direttamente nel modello sia la stagionalità sia eventuali trend, dopo aver reso stazionaria la serie tramite differenziazioni. In sintesi, ARIMA e SARIMA appartengono alla famiglia dei modelli Box-Jenkins per serie storiche: SARIMA non è altro che un ARIMA arricchito che **rimuove le componenti stagionali tramite differenziazione stagionale e**

**aggiunge termini autoregressivi e di media mobile dedicati alla stagionalità**, migliorando le previsioni quando esistono oscillazioni periodiche regolari .

## 2. Vantaggi di SARIMA rispetto ad altri modelli di previsione

**Rispetto ad ARIMA tradizionale:** Il vantaggio principale di SARIMA è la capacità di gestire esplicitamente la stagionalità. Un ARIMA tradizionale (senza termini stagionali) assume che la serie, una volta differenziata per eliminarne il trend, non presenti più pattern periodici fissi. Se invece esistono fluttuazioni stagionali (es. picchi annuali, effetti settimanali, etc.), è generalmente preferibile utilizzare un modello SARIMA dedicato piuttosto che forzare un ARIMA ad adattarsi aumentando arbitrariamente  $p$  o  $q$  . In altre parole, **SARIMA può modellare meglio i picchi ricorrenti** (ad esempio l'aumento delle vendite a dicembre rispetto a gennaio) proprio perché include termini che guardano a un anno prima, mentre un ARIMA puro dovrebbe compensare introducendo molti lag nel componente AR o MA . Il risultato è che SARIMA tende a fornire previsioni più accurate e parametri più interpretabili in presenza di stagionalità, dove ARIMA altrimenti rischierebbe di sottostimare o sovrastimare i cicli stagionali.

**Rispetto a Prophet (Facebook Prophet):** Prophet è un modello di forecasting sviluppato da Meta (Facebook) concepito per essere facile da usare su dati business, con componenti additive di trend e stagionalità definite in modo flessibile. Prophet integra automaticamente effetti stagionali (giornalieri, settimanali, annuali) e festività, ed è noto per funzionare bene “out-of-the-box” con minimo tuning. Il modello SARIMA, essendo statistico-classico, richiede in genere una maggiore analisi per la selezione dei parametri, ma offre alcuni vantaggi: *i)* una maggiore **interpretabilità** parametrica (ogni coefficiente AR/MA ha un significato preciso in termini di dipendenza temporale); *ii)* un approccio parametrico rigoroso che, se ben specificato, può essere molto preciso. Prophet dal canto suo brilla per semplicità d'uso e robustezza: ad esempio, **richiede meno ottimizzazioni manuali dei parametri** poiché individua automaticamente pattern di trend e stagioni nei dati business . In contesti aziendali standard (es. vendite retail con forti effetti di calendario) Prophet fornisce buoni risultati con sforzo minimo; SARIMA può però risultare più accurato se la struttura stagionale è ben compresa e modellata, e tende a fallire meno **“in modo spettacolare” su domini non standard** – è riportato infatti che Prophet può avere difficoltà su serie temporali dove la stagionalità del calendario non è rilevante . In sostanza, SARIMA offre un approccio più tradizionale ma affidabile, mentre Prophet offre una soluzione pronta all'uso che privilegia la praticità.

**Rispetto ai modelli ETS (Error-Trend-Seasonality, es. Holt-Winters):** I modelli ETS, come le varie forme di **smorzamento esponenziale** (SES, Holt, Holt-Winters), adottano un approccio diverso, modellando esplicitamente un livello (o base), un trend e una stagionalità della serie, aggiornati ad ogni nuovo dato con decrescente peso ai valori più vecchi. Il confronto tra ETS e SARIMA può essere riassunto così: *un modello ETS si concentra sul catturare direttamente trend e stagionalità nei dati*, mentre *un ARIMA/SARIMA si concentra sul modellare le autocorrelazioni nella serie* . In pratica, **ETS** è preferibile se la struttura della

serie è ben descrivibile da componenti stabili di trend + stagione (additiva o moltiplicativa) e se si desidera reagire velocemente agli ultimi valori (grazie al maggior peso ai dati più recenti) . Ad esempio, per serie con trend lineare chiaro e stagionalità definita, un modello Holt-Winters può essere semplice ed efficace. **SARIMA**, invece, è vantaggioso quando occorre un maggiore dettaglio nel modellare la dipendenza da vari ritardi specifici o quando la componente stocastica residua (dopo aver rimosso trend e stagione) ha una struttura di autocorrelazione non trascurabile. In molti casi pratici entrambi gli approcci producono risultati simili , e la scelta può dipendere da quale modello si adatta meglio (ad esempio tramite criteri AIC/BIC o test statistici) o da preferenze in termini di interpretabilità. Va notato inoltre che SARIMA può essere esteso ad **ARIMAX/SARIMAX** includendo covariate esogene, cosa non prevista dai modelli ETS standard.

**Rispetto alle reti neurali e ad altri metodi di Machine Learning:** Negli ultimi anni, metodi di *Machine Learning* e *Deep Learning* (come le reti neurali ricorrenti LSTM, modelli Sequence-to-Sequence, alberi decisionali e XGBoost, ecc.) sono stati applicati con successo al forecasting di serie temporali. Il principale vantaggio di questi modelli è la capacità di catturare anche **relazioni non lineari e pattern complessi** nei dati, spesso integrando facilmente molte variabili esplicative. Ad esempio, una rete neurale può potenzialmente modellare interazioni complesse e adattarsi anche se la serie non è stazionaria o se esistono più stagionalità sovrapposte. Inoltre, algoritmi come le reti LSTM non richiedono esplicitamente che i dati siano trasformati per stazionarietà o che esista un campo data (possono imparare direttamente dalle sequenze) . D'altro canto, **i modelli ARIMA/SARIMA restano competitive soprattutto con serie storiche univariate di dimensioni moderate**, dove spesso forniscono accuratezza paragonabile a modelli ML più complessi ma con minore costo computazionale e maggiore interpretabilità. Le reti neurali richiedono molti dati per allenarsi efficacemente e possono risultare “*black box*” difficili da interpretare, il che in ambito aziendale può essere uno svantaggio se c'è bisogno di giustificare le previsioni. Inoltre, anche il tuning delle reti ML può essere oneroso e una rete mal configurata può sovradattarsi. In sintesi, SARIMA offre una **soluzione parsimoniosa e interpretabile** che in molti casi reali risulta **altrettanto accurata** di modelli di apprendimento più complessi, soprattutto quando la struttura stagionale e di trend è forte e chiara . Naturalmente, in presenza di **big data** o di pattern estremamente complessi/non lineari (ad es. segnali con molte influenze esterne), i modelli di ML avanzati possono superare SARIMA; ma spesso una buona strategia in ambito business è usarli in combinazione o confronto – ad esempio partendo da un SARIMA come baseline e validando se modelli più sofisticati apportano miglioramenti significativi.

### 3. Applicazioni aziendali di SARIMA

I modelli SARIMA trovano ampia applicazione in vari domini aziendali in cui si hanno dati storici con trend e stagionalità. Di seguito alcune aree chiave e come SARIMA può essere utilizzato:

- **Previsione della domanda e delle vendite:** Anticipare la domanda futura è cruciale per molte decisioni aziendali. SARIMA permette di sfruttare i dati storici di vendita/domanda, identificando trend di crescita o declino e pattern stagionali (ad esempio incrementi di domanda nei weekend o in determinati mesi). *Esempio:* un'azienda retail può

usare SARIMA per prevedere le vendite di un prodotto durante la stagione natalizia, analizzando i dati degli anni precedenti e catturando i picchi di dicembre seguiti dal calo di gennaio . Questo aiuta a **identificare i picchi stagionali di domanda e pianificare di conseguenza le scorte**, le campagne di marketing o le risorse di personale. In generale, la **forecast della domanda** con modelli SARIMA consente ai manager commerciali di prendere decisioni proattive – come adeguare i livelli di produzione o offrire sconti mirati – basandosi su previsioni quantitative affidabili invece che su sole intuizioni.

- **Pianificazione del budget e gestione finanziaria:** Le serie temporali finanziarie (fatturato mensile, costi operativi trimestrali, flussi di cassa, ecc.) spesso presentano sia trend di fondo sia effetti stagionali (es. trimestri fiscali, stagionalità del settore). SARIMA può essere impiegato per *forecast* finanziari a breve-medio termine, dando al management uno strumento per la pianificazione del budget. *Esempio:* un'azienda può modellare con SARIMA i ricavi mensili degli ultimi 5 anni, evidenziando la crescita annuale (trend positivo) e le variazioni stagionali (es. calo estivo e picco a fine anno). La previsione dei prossimi 12 mesi fornirà una base quantitativa per allocare budget ai vari reparti, pianificare investimenti o individuare in anticipo possibili carenze di liquidità. In pratica, **SARIMA aiuta la funzione finanziaria a creare forecast “rolling” aggiornati** (es. forecast ogni mese per i successivi 12 mesi), integrando i risultati effettivi man mano che sono disponibili e migliorando la precisione delle stime di fine anno. Ciò rende più solida la gestione di cassa e la definizione di obiettivi trimestrali, riducendo sorprese dovute a stagionalità non considerate.

- **Ottimizzazione della produzione e della logistica:** Nella gestione operativa, prevedere i volumi futuri è essenziale per ottimizzare l'uso degli impianti, il personale e la catena logistica. SARIMA viene utilizzato per **prevedere i carichi di produzione** o la domanda di spedizione, tenendo conto di cicli stagionali (es. maggiore produzione prima dell'estate se le vendite estive sono alte) e trend (crescita o calo della domanda generale). *Esempio:* un produttore automobilistico potrebbe impiegare un modello SARIMA per stimare il fabbisogno mensile di auto di un certo modello per l'anno seguente, sulla base dei dati degli anni precedenti . Se le previsioni indicano un forte aumento stagionale in primavera, l'azienda può pianificare turni aggiuntivi o accumulare scorte di componenti chiave in anticipo. Allo stesso modo, **nella logistica**, previsioni SARIMA delle spedizioni settimanali aiutano a ottimizzare rotte e mezzi: ad esempio una società di trasporti potrà prevedere un aumento di consegne nel periodo natalizio e noleggiare automezzi extra o assumere personale temporaneo. In sintesi, le **decisioni operative** – come livelli di produzione, manutenzioni programmate, gestione dei trasporti – beneficiano di previsioni affidabili, riducendo sia il rischio di sotto-produzione (che porta a mancate vendite) sia di sovrapproduzione (che porta a costi di magazzino elevati).

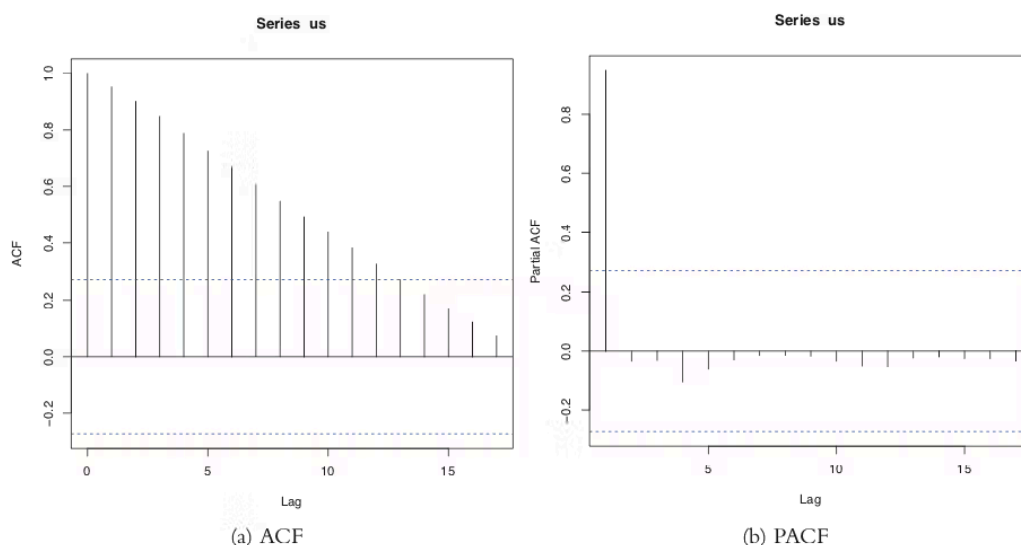
- **Gestione delle scorte e supply chain:** Un'accurata previsione della domanda attraverso SARIMA si riflette direttamente in una migliore **gestione dell'inventario**. Stimando in anticipo i futuri fabbisogni, l'azienda può impostare politiche di riordino *just-in-time* più efficaci, mantenendo scorte sufficienti a soddisfare la domanda ma evitando surplus costosi. *Esempio:* un rivenditore che nota, grazie al modello SARIMA, un picco di domanda per climatizzatori tra giugno e luglio, potrà rifornire i magazzini già a maggio, mentre ridurrà le scorte in inverno quando le vendite previste sono minime. Come citato sopra, un negozio al dettaglio potrebbe identificare **spike stagionali durante le**

**festività** e adeguare l'inventario di conseguenza , assicurando di poter soddisfare i clienti senza trovarsi con scorte eccessive post-festività. Nella supply chain estesa, i fornitori possono condividere previsioni SARIMA con i sub-fornitori per sincronizzare la produzione di materie prime: ad esempio, un produttore alimentare prevede con SARIMA le vendite stagionali di un succo di frutta e allerta con anticipo i fornitori di bottiglie e tappi per farsi trovare pronti al picco. In generale, **l'integrazione di previsioni SARIMA nella pianificazione della supply chain riduce rotture di stock e sprechi**, migliorando il livello di servizio al cliente e l'efficienza dei costi di magazzino.

## 4. Scelta dei parametri e tuning del modello

La costruzione di un modello SARIMA richiede di specificare gli ordini  $(p, d, q)$  e  $(P, D, Q)$  appropriati per i dati in esame. La determinazione di questi parametri può avvenire sostanzialmente in due modi (spesso complementari): **analisi grafica ACF/PACF** e **selezione automatizzata via criteri d'informazione**.

- **Approccio manuale – Analisi di ACF e PACF:** Una tecnica classica per scegliere  $p$  e  $q$  è analizzare i grafici dell'**Autocorrelation Function (ACF)** e della **Partial Autocorrelation Function (PACF)** della serie (eventualmente dopo differenziazione per renderla stazionaria). In breve, l'ACF misura la correlazione tra la serie e i suoi ritardi, mentre la PACF misura la correlazione *residua* di un ritardo quando si sono già spiegati i ritardi precedenti. Questi grafici presentano spesso dei “tagli netti” (cut-off) caratteristici: ad esempio, se la PACF mostra un calo brusco dopo il lag  $p$  mentre l'ACF decresce gradualmente, suggerisce una componente AR di ordine  $p$  (firma AR) . Viceversa, un taglio netto nell'ACF dopo il lag  $q$  accompagnato da PACF che decresce lentamente indica una componente MA di ordine  $q$  . In pratica, l'esperto osserva i primi picchi significativi fuori dalle bande di confidenza nei correlogrammi:



- Se l'ACF *si annulla* (cade a zero) oltre un certo lag  $q$  (mentre la PACF decresce lentamente), possiamo ipotizzare un termine MA( $q$ ).

- Se la PACF *si annulla* oltre un certo lag  $p$  (mentre l'ACF decresce lentamente), ipotizziamo un termine AR( $p$ ).
- Se entrambe ACF e PACF decrescono lentamente (senza cut-off chiari), potrebbe essere presente sia una parte AR che MA (un ARMA), o indicare che la serie necessita ulteriori differenziazioni (non è ancora stazionaria al 100%).

Lo stesso ragionamento si estende ai pattern **stagionali** nei correlogrammi: ad esempio, picchi significativi all'esatto lag 12 nell'ACF di una serie mensile indicano potenziale autocorrelazione annuale, suggerendo l'inclusione di un termine stagionale ( $P$  o  $Q$ ). In genere, un picco alla lag 12 nella PACF implica un AR stagionale di ordine 1 ( $P=1$ ), mentre un picco alla lag 12 nell'ACF implica un MA stagionale ( $Q=1$ ). Naturalmente queste sono indicazioni euristiche: il metodo Box-Jenkins prevede di provare diversi modelli coerenti con ACF/PACF e valutare quello con performance migliore. È anche prassi utilizzare test statistici di stazionarietà (come Dickey-Fuller) per determinare  $d$  (e  $D$ ): se la serie ha un trend deterministico si usa  $d = 1$ , se mostra stagionalità si prova  $D = 1$ , assicurandosi di non sovra-differenziare. L'obiettivo è ottenere residui (dopo il fit del modello) privi di autocorrelazione significativa, segno che  $p, d, q, P, D, Q$  scelti hanno catturato tutta la struttura correlativa presente nei dati.

- **Approccio automatico – Ricerca via AIC/BIC:** In parallelo all'analisi grafica, oggi si fa spesso ricorso a metodi automatizzati per selezionare i parametri di un ARIMA/SARIMA ottimale. I criteri informativi come **AIC (Akaike Information Criterion)** o **BIC (Bayesian Information Criterion)** forniscono una misura della qualità del modello penalizzata dalla complessità: in breve, un modello ha migliore trade-off accuratezza/semplificata se produce un AIC/BIC più basso. Si può quindi condurre una **ricerca su griglia** degli iper-parametri, valutando AIC/BIC di ogni combinazione ( $p, d, q, P, D, Q$ ) plausibile e scegliendo la combinazione con AIC/BIC minimo. Esistono librerie e funzioni che automatizzano questo processo, ad esempio la funzione `auto.arima()` in R o `pmdarima.auto_arima()` in Python, che testano sistematicamente vari modelli (sottoponendo anche la serie a test di stazionarietà per decidere  $d$  e  $D$ ) e restituiscono il modello ottimale secondo AIC/BIC. Questi approcci sfruttano algoritmi intelligenti di riduzione del search space e spesso includono logiche per evitare combinazioni non stazionarie o incoerenti. Il vantaggio è una selezione rapida e oggettiva, sebbene sia sempre buona norma validare poi i risultati (controllando i residui del modello selezionato e verificando che abbiano un senso rispetto al contesto). In sintesi, **il tuning del modello SARIMA avviene iterativamente**: si prova un set di parametri, si guarda la bontà del fit e le diagnostiche sui residui, eventualmente si modifica il modello. Criteri come AIC/BIC aiutano a confrontare modelli diversi: un AIC più basso indica che un modello spiega i dati in modo più efficiente (miglior log-verosimiglianza a parità di parametri). Vale la pena notare che BIC penalizza di più la complessità (tende a scegliere modelli più semplici rispetto ad AIC), quindi in alcuni casi AIC e BIC possono indicare ordini differenti; la scelta può dipendere dal contesto (AIC per massimizzare accuratezza, BIC se si preferisce un modello più parsimonioso). Spesso si esamina anche l'**AICc** (Akaike corretto) per serie molto corte. Nel complesso, unendo l'intuito derivato da ACF/PACF e la selezione via criterio informativo si arriva a determinare una specifica SARIMA che **bilancia buon fit e semplicità**.

## 5. Valutazione delle prestazioni del modello

Una volta stimato un modello SARIMA, è fondamentale valutarne la capacità predittiva e l'affidabilità, soprattutto in un contesto aziendale dove le decisioni prese in base alle previsioni hanno impatti economici. Gli aspetti chiave della valutazione sono: le **metriche di errore** utilizzate e la **strategia di validazione temporale** adottata.

**Metriche di valutazione (errore di previsione):** Per misurare l'accuratezza di un modello di forecast si usano varie metriche, le più comuni essendo:

- **MAE (Mean Absolute Error)** – errore assoluto medio: è la media dei valori assoluti degli errori  $MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}$ . Indica di quanto, in media, le previsioni si discostano dal valore reale, senza considerare la direzione dell'errore (sovra o sotto-stima). Essendo lineare, è facile da interpretare (ad es. MAE = 500 unità significa che in media sbagliamo di 500 unità) ed è poco sensibile ai grandi errori isolati, il che può essere un vantaggio se si vuole una misura “robusta” agli outlier.

- **RMSE (Root Mean Squared Error)** – scarto quadratico medio: è la radice quadrata della media dei quadrati degli errori. A differenza del MAE, penalizza maggiormente gli errori grandi perché il quadrato dell'errore amplifica le deviazioni più elevate. In concreto, RMSE dà più peso ai casi in cui il modello “sbaglia di molto”; è utile se si vuole essere severi con grossi errori (ad es. in certi contesti un errore molto grande è particolarmente critico). Spesso RMSE viene paragonato alla deviazione standard dei dati per capire se l'errore è accettabile. Ad esempio, se le vendite mensili tipicamente fluttuano con uno scarto di  $\pm 100$  unità, un RMSE di 80 può considerarsi buono.

- **MAPE (Mean Absolute Percentage Error)** – errore percentuale assoluto medio:  $MAPE = 100 \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$ . In pratica indica in media di che percentuale ci si sbaglia. Ad esempio un MAPE del 5% significa che le previsioni differiscono dai valori reali del 5% in media. Il MAPE è molto popolare in ambito business perché normalizza l'errore sulla grandezza della serie, consentendo di confrontare performance su dataset diversi e dando un'idea immediata (“errore del 5%” è più intuitivo da comunicare). Occorre però attenzione: se  $y$  può assumere valori zero o molto piccoli, il MAPE diventa problematico (divide per valori vicini a 0). In questi casi si preferiscono alternative come il **SMAPE** o il **MASE**.

Oltre a queste, esistono molte altre metriche (MSE, MASE, SMAPE, etc.), ma MAE, RMSE e MAPE coprono bene sia interpretabilità sia sensibilità agli outlier. In un contesto di previsione vendite, ad esempio, il management potrebbe fissare un obiettivo di MAPE < 10% per considerare il modello accettabile. Durante la fase di valutazione, si calcolano queste metriche sulle previsioni out-of-sample del modello SARIMA e le si confronta con modelli di benchmark (es. una previsione naive o il semplice ultimo valore noto) per assicurarsi che il SARIMA offra un miglioramento sostanziale.

**Validazione del modello con walk-forward (rolling forecast):** A differenza della validazione incrociata standard usata in altri ambiti di Machine Learning, per le serie temporali bisogna rispettare l'ordine cronologico dei dati. Non si può mescolare liberamente training e test, perché il modello non deve mai “vedere” informazioni dal futuro. La tecnica

più diffusa è la **validazione a finestra temporale scorrevole**, nota anche come *walk-forward validation* o *rolling origin*. Consiste nel procedere così: si dividono i dati in una prima parte per stimare il modello e una parte successiva per test; si stima il modello sulla finestra iniziale e si fanno previsioni per il periodo immediatamente seguente (es. predire il mese successivo). Si calcolano gli errori, poi si “avanza” la finestra temporale includendo il periodo appena predetto nel training e predicendo il successivo, iterando questo processo su tutto l’orizzonte di test. In altre parole, **si simula l’uso reale del modello nel tempo**, aggiornando ad ogni passo la conoscenza con i dati nuovi man mano che diventano disponibili. Questo metodo garantisce che ad ogni predizione il modello usi solo informazioni del passato (come avverrebbe operativamente). Ad esempio, su 5 anni di dati mensili, si potrebbe usare un modello SARIMA sui primi 4 anni per predire il mese 49; poi si aggiunge il mese 49 al set e si ripredice il 50, e così via fino al 60. In questo modo si ottiene una sequenza di previsioni *out-of-sample* che possono essere confrontate coi valori reali corrispondenti. Le metriche di errore calcolate su queste predizioni simulano la performance che il modello avrebbe avuto “sul campo”. La cross-validation tradizionale k-fold non è adatta alle serie temporali proprio perché romperebbe la sequenzialità (dando al modello occhi sul futuro); il walk-forward invece rispetta il principio causale. Un ulteriore beneficio del walk-forward è che permette di valutare **la stabilità del modello nel tempo**: si possono notare, ad esempio, se gli errori crescono mano a mano che ci si allontana dall’inizio (forse perché la struttura del modello invecchia) o se certe sottoperiodi (es. ogni anno a gennaio) hanno errori sistematicamente maggiori (forse serve includere un effetto esterno, come il post-festività). In sintesi, una valida valutazione delle prestazioni di un SARIMA prevede: calcolo di metriche appropriate (MAE, RMSE, MAPE) su un insieme di test significativo e una procedura di **backtesting temporale** rigorosa, assicurando che le previsioni siano sempre generate senza guardare avanti nel futuro.

## 6. Limitazioni e sfide di SARIMA e possibili alternative

Benché il modello SARIMA sia uno strumento potente e diffuso, non è adatto a tutte le situazioni. Di seguito alcune limitazioni da tenere presenti e strategie alternative da considerare:

- **Serie con componenti stagionali multiple o irregolari:** SARIMA, per costruzione, gestisce *una* singola stagionalità di periodo fisso  $m$ . Se una serie presenta più cicli stagionali (es. un dato economico con pattern sia annuali che trimestrali, oppure dati giornalieri con stagionalità settimanale e annuale insieme), un classico SARIMA non può incorporarle tutte contemporaneamente. In questi casi si può ricorrere a modelli più avanzati come **TBATS** o **Prophet**, che supportano molteplici stagionalità, oppure modellare la serie con **modelli moltiplicativi** o con approcci a base di decomposizione (ad esempio, decomporre e modellare separatamente le diverse frequenze). La limitazione di SARIMA vale anche per stagionalità **non perfettamente stabili** nel tempo: se il periodo cambia o la forma della stagionalità evolve (ad es. il picco di vendita si sposta da un mese all’altro negli anni), i termini stagionali statici di SARIMA potrebbero non catturare bene il fenomeno. In tali situazioni può aiutare introdurre variabili dummy per i mesi, oppure optare per modelli di apprendimento automatico capaci di intercettare variazioni più complesse.



- **Incapacità di modellare eventi esterni o variabili esplicative (modello univariato puro):** Un SARIMA standard considera solo la serie storica oggetto di previsione. Ciò significa che **non incorpora esplicitamente l'effetto di variabili esterne** (prezzi, promozioni, fattori macroeconomici, meteo, ecc.) che spesso influenzano le serie di business. Se tali fattori sono rilevanti, il SARIMA puro può risultare sottoperformante perché attribuisce tutto ai pattern interni della serie. Un rimedio è passare a un modello **SARIMAX** (ARIMA stagionale con eXogene), includendo regressori aggiuntivi nel modello, oppure usare modelli di regressione multivariata o approcci di machine learning che integrino queste variabili. Ad esempio, se le vendite sono fortemente influenzate dalle campagne pubblicitarie, un SARIMA ignorerebbe il calendario delle promozioni e i suoi residui mostrerebbero errori sistematici in corrispondenza delle promo. SARIMAX invece permetterebbe di aggiungere una variabile binaria "promo attiva" migliorando la previsione. In generale, la **limitazione univariata** di ARIMA/SARIMA è stata uno dei motivi per cui si sono sviluppati modelli estesi: *"SARIMA aveva come limite l'impossibilità di incorporare variabili esogene, poi superato da SARIMAX"*.

- **Presenza di trend o cambiamenti strutturali non lineari:** SARIMA gestisce bene trend lineari (rimossi con differenziazione) e stagionalità stazionarie, ma se la serie ha **comportamenti non lineari o discontinuità** (cambi di regime, soglie, effetti saturazione, ecc.), la sua struttura lineare potrebbe non bastare. Ad esempio, durante la pandemia COVID molti modelli ARIMA/SARIMA addestrati sui dati pre-2020 hanno fallito nel prevedere l'andamento, perché è sopraggiunto un cambiamento strutturale enorme non inferibile dai pattern passati. In tali casi, occorre aggiornare frequentemente il modello (riaddestrandolo man mano che arrivano nuovi dati post-cambiamento) o considerare modelli più flessibili. Alternative possibili: modelli a **breakpoints** (che stimano diversi parametri prima/dopo un punto di cambio), modelli **ARIMA con coefficienti variabili nel tempo**, oppure passare a tecniche machine learning capaci di riaddestrarsi continuamente e captare nuove dinamiche. Anche modelli a rete neurale possono adattarsi meglio se opportunamente allenati su scenari con cambiamenti (pur richiedendo dati sufficienti per apprendere). In assenza di molti dati, approcci più semplici come la metodologia di **forecasting robusto** (es. usare sempre la parte più recente dei dati per prevedere il futuro prossimo, concettualmente simile a un *rolling window* molto corta) possono talvolta funzionare meglio di un SARIMA globale che dà peso a pattern storici ormai sorpassati.

- **Dati scarsi o altamente volatili:** Quando si dispone di poche osservazioni storiche, soprattutto meno di due periodi stagionali completi, stimare un modello SARIMA affidabile è difficile – i parametri potrebbero essere instabili e le stime poco robuste. In questi casi, può essere preferibile usare modelli più semplici (ad es. un modello stagionale naïve che ripete l'ultimo anno noto, o modelli di smoothing semplice) finché non si accumulano dati sufficienti. Analogamente, se la serie è dominata da forte rumore aleatorio (ad es. vendite di un prodotto nuovo con variazioni erratiche), un modello lineare come SARIMA potrebbe non distinguere segnale e rumore efficacemente. In tali situazioni i **modelli di media mobile semplice o naïve** a volte sono difficili da battere. Ad esempio, se le vendite oscillano in modo imprevedibile del  $\pm 30\%$  ogni mese senza un chiaro pattern, un SARIMA cercherà pattern che forse non esistono, introducendo errore; una strategia banale come "prevedi uguale all'ultimo mese" potrebbe paradossalmente fare meglio in termini di MAE. Pertanto, è sempre importante valutare la **sostenibilità statistica** di un SARIMA: se i dati non supportano un certo complesso di parametri, l'overfitting è dietro l'angolo.

- **Altre alternative modellistiche:** Oltre a quelli già citati (Prophet, ETS, modelli ML), vale la pena menzionare altre tecniche utili quando SARIMA non è l'ideale. I **modelli a spazio di stato** (Kalman filter e affini) forniscono un framework più generale che include ARIMA e ETS come casi particolari, e possono facilmente estendersi a più stagionalità o dinamiche più complesse. I **modelli ARIMA misti con rete neurale (hybrid)** sono stati proposti per catturare sia la parte lineare che quella non lineare di una serie. I **modelli di autoregressione vettoriale (VAR)** diventano rilevanti se si hanno più serie correlate (es. domanda di prodotti correlati): invece di un SARIMA separato per ogni serie, un VAR (eventualmente con stagionalità, chiamato VARIMA) coglie le interdipendenze tra serie, cosa che un SARIMA univariato ignora. Infine, modelli di **forecasting gerarchico** sono importanti se le previsioni devono essere coerenti a diversi livelli di aggregazione (es. previsioni di vendita per negozio e totale nazionale): in tal caso si usano procedure per aggiustare le previsioni SARIMA a ogni livello in modo consistente.

**Quando SARIMA non è ideale?** Riassumendo quanto sopra: se la serie presenta pattern stagionali chiari e stabili, un SARIMA ben calibrato è spesso una scelta eccellente per accuratezza e interpretabilità. Ma **se la serie è influenzata pesantemente da fattori esterni non inclusi, ha più stagionalità, o subisce svolte improvvise**, allora SARIMA puro potrebbe non performare al meglio. In tali scenari si valutano modelli alternativi o aggiuntivi: ad esempio, Prophet per includere facilmente effetti di calendario multipli; modelli a regressione multipla o alberi di decisione per tener conto di variabili esplicative; reti neurali o gradient boosting per pattern complessi; oppure un approccio ensemble che combini previsioni di SARIMA con quelle di altri modelli (sperando di catturare il meglio di entrambi). È importante sottolineare che nel problem solving reale spesso **si confrontano diverse soluzioni di forecasting**: SARIMA diventa uno dei candidati da confrontare, sapendo che rappresenta un solido benchmark "lineare". Se modelli più avanzati non battono significativamente il SARIMA (ad esempio in termini di MAPE out-of-sample), si può tranquillamente adottare il SARIMA per la sua semplicità. Viceversa, se esistono limiti intrinseci (come quelli discussi sopra) che ne degradano la qualità predittiva, allora quelle sono bandiere rosse per esplorare alternative. In definitiva, l'analista dovrebbe conoscere bene i **presupposti di SARIMA** – stazionarietà dopo differenziazione, una stagionalità fissa, relazioni lineari – e monitorare quando questi presupposti sono violati: è in tali circostanze che SARIMA "non è il modello ideale" e bisogna rivolgersi altrove per migliorare le previsioni.