

# Progetto SAS per Modulo di Previsione delle Quantità Prodotte

Analisi e Implementazione di Modelli di Forecasting

Andrea Roberto Benvenuti

3 settembre 2025

## Indice

<b>1</b>	<b>Introduzione e Obiettivi del Progetto</b>	<b>4</b>
1.1	Contesto Aziendale . . . . .	4
1.2	Architettura del Sistema . . . . .	4
<b>2</b>	<b>Portfolio di Soluzioni Implementate</b>	<b>4</b>
2.1	ARIMA_AutoSelection_7M.R - Approccio di Selezione Automatica . . . . .	4
2.1.1	Caratteristiche Principali . . . . .	4
2.1.2	Portfolio di Modelli . . . . .	5
2.1.3	Vantaggi . . . . .	5
2.1.4	Limiti . . . . .	5
2.1.5	Caso d'uso ideale . . . . .	5
2.2	HybridModel_Advanced_Forecast.R - Modello Ibrido Personalizzato . . . . .	5
2.2.1	Architettura Multi-Componente . . . . .	5
2.2.2	Innovazioni Metodologiche . . . . .	6
2.2.3	Vantaggi Specifici . . . . .	6
2.2.4	Limiti . . . . .	6
2.2.5	Caso d'uso ideale . . . . .	6
2.3	RobustForecasting_Production.R - Framework Enterprise . . . . .	6
2.3.1	Architettura Robusta . . . . .	6
2.3.2	Modelli Implementati . . . . .	6
2.3.3	Robustezza Operativa . . . . .	7
2.3.4	Vantaggi . . . . .	7
2.3.5	Limiti . . . . .	7
2.3.6	Caso d'uso ideale . . . . .	7
<b>3</b>	<b>Strategia di Modellazione</b>	<b>7</b>
3.1	Decisioni Architettureali . . . . .	7
3.1.1	Scelta degli Algoritmi . . . . .	7
3.1.2	Evoluzione Future . . . . .	7

3.2	Matrice di Selezione del Modello . . . . .	8
3.3	Integrazione Dati Esterni . . . . .	8
3.3.1	Fonti Gratuite . . . . .	8
3.3.2	Fonti Commerciali . . . . .	8
<b>4</b>	<b>Implementazione e Testing</b>	<b>8</b>
4.1	Lavoro Svolto . . . . .	8
4.1.1	Caricamento e Pulizia Dati . . . . .	8
4.1.2	Trasformazione e Validazione . . . . .	9
4.1.3	Aggregazione e Creazione Serie Temporalì . . . . .	9
4.2	Analisi Esplorativa . . . . .	9
4.3	Confronto Performance tra Approcci . . . . .	10
4.3.1	Testing ARIMA_AutoSelection_7M.R . . . . .	10
4.3.2	Testing HybridModel_Advanced_Forecast.R . . . . .	10
4.3.3	Testing RobustForecasting_Production.R . . . . .	10
4.4	Gestione Regressori Futuri . . . . .	10
4.4.1	Problema delle Variabili Esogene . . . . .	10
4.4.2	Soluzione Implementata . . . . .	10
<b>5</b>	<b>Risultati dei Test</b>	<b>11</b>
5.1	Performance Comparativa . . . . .	11
5.2	Osservazioni sui Tre Approcci . . . . .	11
5.2.1	ARIMA_AutoSelection_7M.R . . . . .	11
5.2.2	HybridModel_Advanced_Forecast.R . . . . .	12
5.2.3	RobustForecasting_Production.R . . . . .	13
<b>6</b>	<b>Criticità e Problemi Identificati</b>	<b>14</b>
6.1	Problemi nei Dati . . . . .	14
6.2	Limitazioni Tecniche . . . . .	14
6.3	Sfide Metodologiche . . . . .	14
<b>7</b>	<b>Raccomandazioni d'Uso</b>	<b>15</b>
7.1	Matrice Decisionale . . . . .	15
7.2	Workflow Consigliato . . . . .	15
<b>8</b>	<b>Roadmap e Sviluppi Futuri</b>	<b>15</b>
8.1	Fase 1: Consolidamento . . . . .	15
8.2	Fase 2: Enhancement . . . . .	15
8.3	Fase 3: Operazionalizzazione . . . . .	16
<b>9</b>	<b>Conclusioni</b>	<b>16</b>
<b>10</b>	<b>Confronto dei Modelli</b>	<b>16</b>
10.1	Analisi Comparativa delle Architetture . . . . .	16
10.1.1	Modello Personalizzato Ibrido (HybridModel_Advanced_Forecast) . . . . .	16
10.1.2	Modello Ibrido Avanzato con Machine Learning (HybridMachine-Learning) . . . . .	17
10.2	Matrice Comparativa delle Performance . . . . .	17
10.3	Criteri di Selezione del Modello . . . . .	17

10.3.1	Utilizzo Modello Base . . . . .	17
10.3.2	Utilizzo Modello Avanzato . . . . .	18
10.4	Framework di Implementazione Consigliato . . . . .	18
10.4.1	Approccio Graduale . . . . .	18
10.4.2	Metriche di Decisione . . . . .	18

# 1 Introduzione e Obiettivi del Progetto

Il progetto si propone di sviluppare un sistema di previsione accurato per le quantità prodotte mensili, utilizzando dati storici di produzione e commesse aziendali. L'obiettivo principale è creare un modulo predittivo che possa essere integrato nel sistema gestionale esistente per ottimizzare la pianificazione della produzione e la gestione delle scorte.

## 1.1 Contesto Aziendale

Il progetto nasce dall'esigenza di:

- Prevenire situazioni di stock-out
- Ottimizzare i livelli di scorte
- Migliorare la pianificazione della produzione
- Ridurre i costi operativi attraverso previsioni accurate

## 1.2 Architettura del Sistema

Il flusso di lavoro prevede:

1. Estrazione dati dal Data Warehouse aziendale
2. Elaborazione e analisi in RStudio
3. Applicazione di modelli di forecasting in R
4. Esportazione risultati verso Qlik Sense per visualizzazione

# 2 Portfolio di Soluzioni Implementate

Il progetto comprende tre script R specializzati, ciascuno con un approccio metodologico distinto per soddisfare diverse esigenze operative:

## 2.1 ARIMA\_AutoSelection\_7M.R - Approccio di Selezione Automatica

### 2.1.1 Caratteristiche Principali

- **Filosofia:** Selezione automatica del miglior modello singolo
- **Orizzonte temporale:** 7 mesi (marzo-settembre 2025)
- **Criterio selezione:** AIC (Akaike Information Criterion)
- **Complessità:** Media - approccio sistematico ma non eccessivamente elaborato

### 2.1.2 Portfolio di Modelli

Lo script implementa un confronto strutturato tra:

- **Baseline:** MEAN, NAIVE, DRIFT
- **Automatici:** ETS automatico, ARIMA automatico
- **Stagionali:** SNAIVE, ARIMA con componenti sin/cos (se stagionalità rilevata)
- **Con regressori:** ARIMA con n.commesse come variabile esogena

### 2.1.3 Vantaggi

- Approccio "textbook" affidabile e standardizzato
- Selezione oggettiva basata su criteri statistici
- Implementazione robusta della gestione regressori futuri
- Output chiaro e interpretabile

### 2.1.4 Limiti

- Rigidità nella selezione (solo miglior modello singolo)
- Non considera approcci ensemble
- Orizzonte limitato a 7 mesi

### 2.1.5 Caso d'uso ideale

Analisi rapide e confronti standardizzati quando si necessita di un approccio metodologicamente solido ma non eccessivamente complesso.

## 2.2 HybridModel\_Advanced\_Forecast.R - Modello Ibrido Personalizzato

### 2.2.1 Architettura Multi-Componente

Il modello implementa un approccio ibrido proprietario che combina:

- **Componente deterministica:** Regressione quadratica per relazione quantità/numero commesse
- **Stagionalità adattiva:** Decomposizione STL robusta con aggiornamento dinamico
- **Modellazione residui:** ARIMA sui residui per catturare dinamiche stocastiche non spiegate
- **Trend breaks detection:** Algoritmo proprietario per identificazione cambiamenti strutturali
- **Ensemble pesato:** Combinazione componenti con pesi basati su performance storica

### 2.2.2 Innovazioni Metodologiche

- **Weighted moving average:** Per stima robusta n\_commesse future
- **Adaptive seasonal component:** Pattern stagionali che si adattano ai dati recenti
- **Performance-based weighting:** Pesi ensemble basati su validazione rolling window
- **Realistic constraints:** Vincoli per evitare previsioni irrealistiche

### 2.2.3 Vantaggi Specifici

- Sfruttamento intelligente delle relazioni quantità-commesse
- Gestione automatica di outlier e trend breaks
- Intervalli di confidenza dinamici crescenti con orizzonte temporale
- Robustezza contro overfitting tramite regolarizzazione implicita

### 2.2.4 Limiti

- Complessità elevata - può essere over-engineered per dati semplici
- Richiede comprensione approfondita per manutenzione
- Tempo di esecuzione maggiore per ottimizzazione componenti

### 2.2.5 Caso d'uso ideale

Quando i dati presentano relazioni specifiche (quantità-commesse) da sfruttare e si necessita di un modello su misura per il dominio applicativo.

## 2.3 RobustForecasting\_Production.R - Framework Enterprise

### 2.3.1 Architettura Robusta

- **Gestione errori estensiva:** Fallback multipli per ogni componente
- **Portfolio completo:** Include Prophet, Random Forest, XGBoost (se disponibili)
- **Orizzonte esteso:** 12 mesi per pianificazione long-term
- **Controlli qualità:** Validazione automatica di dati e risultati

### 2.3.2 Modelli Implementati

- **Tradizionali:** ARIMA, ETS con tutte le varianti
- **Prophet:** Modello Facebook per trend non-lineari e festività
- **Machine Learning:** Random Forest, XGBoost (caricamento condizionale)
- **Ensemble:** Combinazioni weighted e stacked

### 2.3.3 Robustezza Operativa

- **Safe loading:** Caricamento condizionale dei package opzionali
- **Error handling:** Gestione graceful di fallimenti con alternative
- **Data validation:** Controlli automatici di consistenza e qualità
- **Fallback strategies:** Approcci semplificati quando modelli complessi falliscono

### 2.3.4 Vantaggi

- Adatto per implementazioni in produzione
- Universalmente applicabile a diverse tipologie di dati
- Analisi esplorativa completa integrata
- Output enterprise-ready con export automatico

### 2.3.5 Limiti

- Complessità esecutiva elevata
- Controlli ridondanti che rallentano l'esecuzione
- Può essere eccessivo per analisi esplorative rapide

### 2.3.6 Caso d'uso ideale

Sistemi in produzione che richiedono massima robustezza, gestione errori sofisticata e capacità di adattamento a situazioni impreviste.

## 3 Strategia di Modellazione

### 3.1 Decisioni Architettureali

#### 3.1.1 Scelta degli Algoritmi

Sono stati identificati tre approcci principali:

- **ETS (Error-Trend-Seasonality):** per serie con trend e stagionalità regolari
- **ARIMA:** per analisi di serie temporali stazionarie
- **Prophet:** per gestione di trend non lineari e festività

#### 3.1.2 Evoluzione Future

- **XGBoost:** per modelli multivariati più accurati
- **Reinforcement Learning:** per ottimizzazione dinamica dei prezzi
- **Feature Engineering:** automatico con explainability (SHAP)

Criterio	ARIMA Auto	Hybrid Model	Robust Framework
Complessità implementazione	Bassa	Alta	Media
Accuratezza attesa	Media	Alta	Alta
Robustezza operativa	Media	Media	Molto Alta
Interpretabilità	Alta	Media	Media
Tempo di esecuzione	Veloce	Medio	Lento
Manutenibilità	Alta	Bassa	Alta
Adattabilità domini	Media	Specifica	Universale

Tabella 1: Confronto caratteristiche dei tre approcci implementati

## 3.2 Matrice di Selezione del Modello

## 3.3 Integrazione Dati Esterni

### 3.3.1 Fonti Gratuite

- **ISTAT**: produzione mensile, indici prezzi, import/export
- **Eurostat**: PRODCOM, External Trade, PPI UE
- **Assocarta**: report trimestrali, prezzi carta da macero
- **CEPI**: statistiche UE, sustainability reports

### 3.3.2 Fonti Commerciali

- **FastMarkets RISI**: prezzi PPI settimanali, capacity tracker, trade flows
- Costi variabili: €1.500-15.000/mese
- Dashboard, API, alert personalizzabili

# 4 Implementazione e Testing

## 4.1 Lavoro Svolto

### 4.1.1 Caricamento e Pulizia Dati

Il primo step ha riguardato la gestione dei problemi di caricamento dati dal file `MargineCommessa.xlsx`:

- **Problema**: Valori mancanti rappresentati come "-" o "NA" non venivano interpretati correttamente
- **Soluzione**: Utilizzo di `read_excel` con parametro `na = c("-", "NA", "")`
- **Standardizzazione**: Applicazione di `clean_names()` per uniformare i nomi delle colonne

```
dati_raw <- read_excel(file_path, na = c("-", "NA", "")) %>%
  clean_names()
```

### 4.1.2 Trasformazione e Validazione

La pulizia dei dati ha incluso:

- Conversione dei mesi italiani in numeri mediante mapping
- Creazione della colonna `data_mese` come data completa
- Filtraggio di righe con quantità prodotta nulla o negativa
- Validazione dei range temporali (anni 2020-2025)

```
mesi_it <- c("gen", "feb", "mar", "apr", "mag", "giu",
            "lug", "ago", "set", "ott", "nov", "dic")

dati_clean <- dati_raw %>%
  mutate(
    mese_num = match(tolower(mese), mesi_it),
    anno = as.numeric(anno),
    data_mese = as.Date(paste(anno, sprintf("%02d", mese_num),
                              "01", sep = "-")),
    qta_prodotta = as.numeric(qta_prodotta)
  ) %>%
  filter(!is.na(data_mese), !is.na(qta_prodotta),
         qta_prodotta > 0)
```

### 4.1.3 Aggregazione e Creazione Serie Temporal

L'aggregazione mensile ha prodotto metriche comprehensive:

- `qta_prodotta_tot`: totale quantità prodotta per mese
- `n_commesse`: numero di commesse per mese
- `commessa_tipo_A/C`: conteggio per tipologia
- `perc_tipo_A/C`: percentuali per analisi composizione

La conversione in `tsibble` ha incluso variabili ausiliarie per modellazione avanzata:

- Componenti stagionali: `sin_mese`, `cos_mese`
- Trend temporale: `trend` in anni dal primo periodo
- Classificazioni: trimestre, stagione

## 4.2 Analisi Esplorativa

L'analisi descrittiva ha rivelato:

- Calcolo di statistiche robuste (media, mediana, CV)
- Visualizzazione con trend LOESS per pattern identification
- Decomposizione STL per separazione componenti

## 4.3 Confronto Performance tra Approcci

### 4.3.1 Testing ARIMA\_AutoSelection\_7M.R

- **Split:** Training fino a febbraio 2025, test su 7 mesi
- **Selezione:** Automatica basata su AIC più basso
- **Regressori:** Gestione n\_commesse con media stagionale
- **Output:** Previsioni puntuali con intervalli  $\pm 15\%$  stimati

### 4.3.2 Testing HybridModel\_Advanced\_Forecast.R

- **Componenti:** Regressione quadratica + STL + ARIMA residui
- **Breakpoints:** Rilevazione automatica cambiamenti strutturali
- **Ensemble:** Pesi adattivi basati su performance rolling
- **Intervalli:** Dinamici crescenti con volatilità storica

### 4.3.3 Testing RobustForecasting\_Production.R

- **Portfolio:** ARIMA, ETS, Prophet, ML (se disponibili)
- **Validation:** Rolling window con split 80/20
- **Fallback:** Gestione graceful di errori con alternative
- **Orizzonte:** 12 mesi per pianificazione estesa

## 4.4 Gestione Regressori Futuri

### 4.4.1 Problema delle Variabili Esogene

Per i modelli con regressori (es. ARIMA con n\_commesse), è necessario stimare i valori futuri delle variabili esplicative.

### 4.4.2 Soluzione Implementata

Stima di n\_commesse future tramite media storica stagionale:

```
media_commesse_mese <- train_data %>%  
  mutate(mese_num = month(year_month)) %>%  
  group_by(mese_num) %>%  
  summarise(n_commesse_media = mean(n_commesse, na.rm = TRUE))
```

## 5 Risultati dei Test

### 5.1 Performance Comparativa

I test implementati nei file `testDatiQuantitaProduzioneARIMAETS.R` e `confrontoARIMAETSPROHET.R` hanno mostrato:

- **Variabilità nei risultati:** Nessun modello è consistentemente superiore
- **ARIMA:** Migliori performance su anni recenti (1975-1980) con RMSE e MAE più bassi
- **ETS:** Competitive e robusto, particolarmente efficace su trend stabili
- **Prophet:** Performance inferiore, probabilmente necessita di variabili aggiuntive

### 5.2 Osservazioni sui Tre Approcci

#### 5.2.1 ARIMA\_AutoSelection\_7M.R

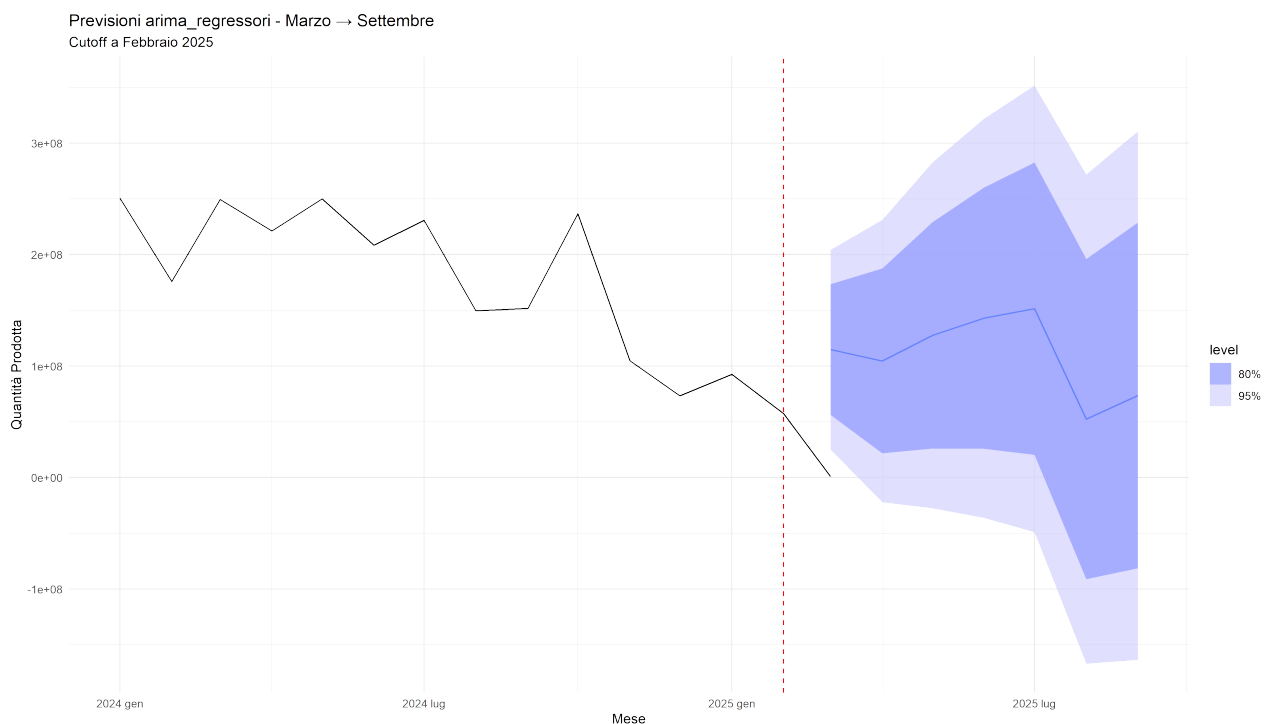


Figura 1: Previsioni ARIMA\_AutoSelection\_7M.R da marzo a settembre 2025 con intervalli di confidenza.

- **Descrizione:** Questo script implementa un modello ARIMA automatico, che seleziona in maniera ottimale l'ordine dei parametri AR, I e MA tramite criteri statistici (AIC, BIC). È pensato per serie temporali mensili fino a 7 mesi di previsione.
- **Punti di forza:**
  - Estremamente veloce e leggero in termini computazionali.

- Output facile da interpretare grazie alla natura lineare del modello ARIMA.
- Stabile su serie temporali regolari e senza grandi outlier.

- **Limiti:**

- Non gestisce bene dati mancanti o valori anomali senza preprocessing.
- Performance ridotta se la serie ha pattern complessi o relazioni tra variabili esterne (ad esempio numero di commesse).

- **Utilizzo ottimale:**

- Analisi esplorative rapide.
- Baseline per confrontare altri modelli più complessi.
- Serie temporali con trend e stagionalità semplici.

### 5.2.2 HybridModel\_Advanced\_Forecast.R

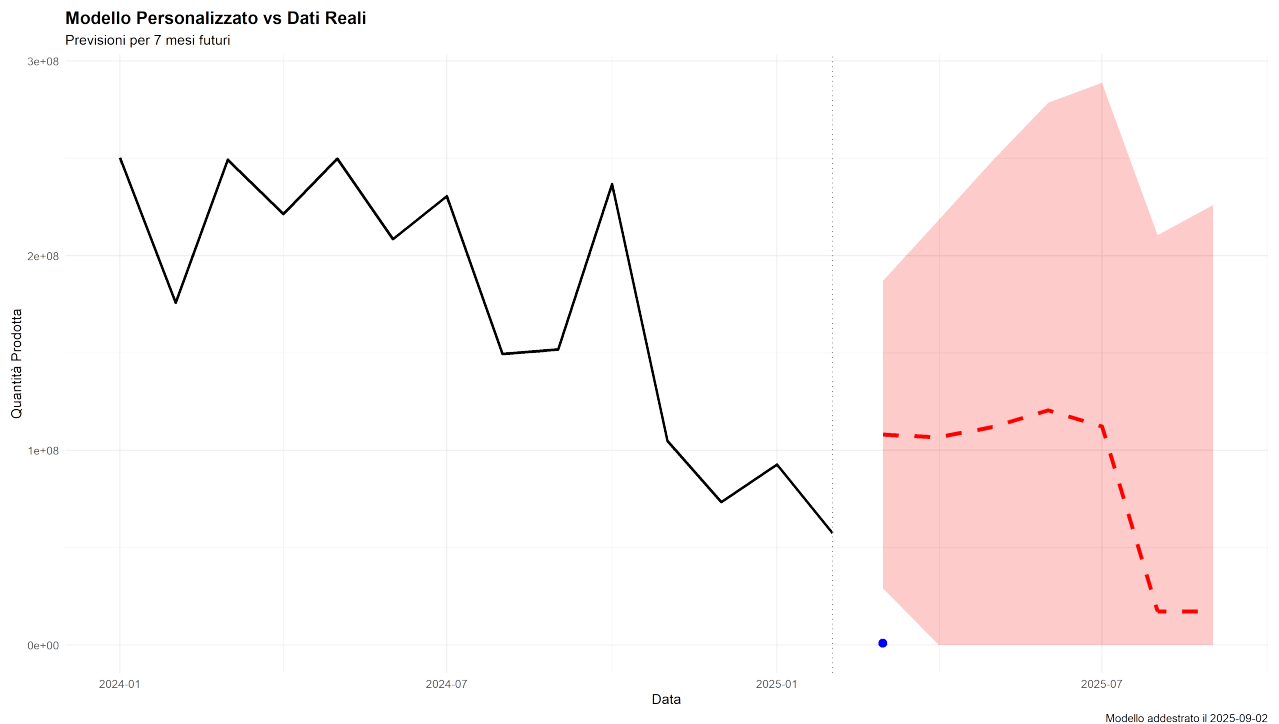


Figura 2: Confronto tra previsioni del HybridModel\_Advanced\_Forecast.R e dati reali.

- **Descrizione:** Modello ibrido avanzato che combina più approcci: regressione quadratica per la relazione quantità-commesse, decomposizione STL per la stagionalità adattiva, ARIMA sui residui e rilevamento automatico di trend break. Include anche un ensemble pesato per migliorare la precisione delle previsioni.

- **Punti di forza:**

- Riesce a catturare pattern complessi e non lineari tra variabili (es. quantità prodotta e numero di commesse).
- Gestione adattiva della stagionalità e dei cambiamenti strutturali nel trend.
- Intervalli di confidenza dinamici che aumentano con l'orizzonte di previsione.

- **Limiti:**

- Richiede più risorse computazionali rispetto a un ARIMA standard.
- Implementazione più complessa e meno immediata da interpretare per utenti non esperti.

- **Utilizzo ottimale:**

- Dati con pattern multivariati evidenti.
- Situazioni in cui è necessario considerare relazioni tra quantità prodotte e numero/tipo di commesse.
- Previsioni di medio termine (fino a 7 mesi) con maggiore accuratezza rispetto a modelli standard.

### 5.2.3 RobustForecasting\_Production.R

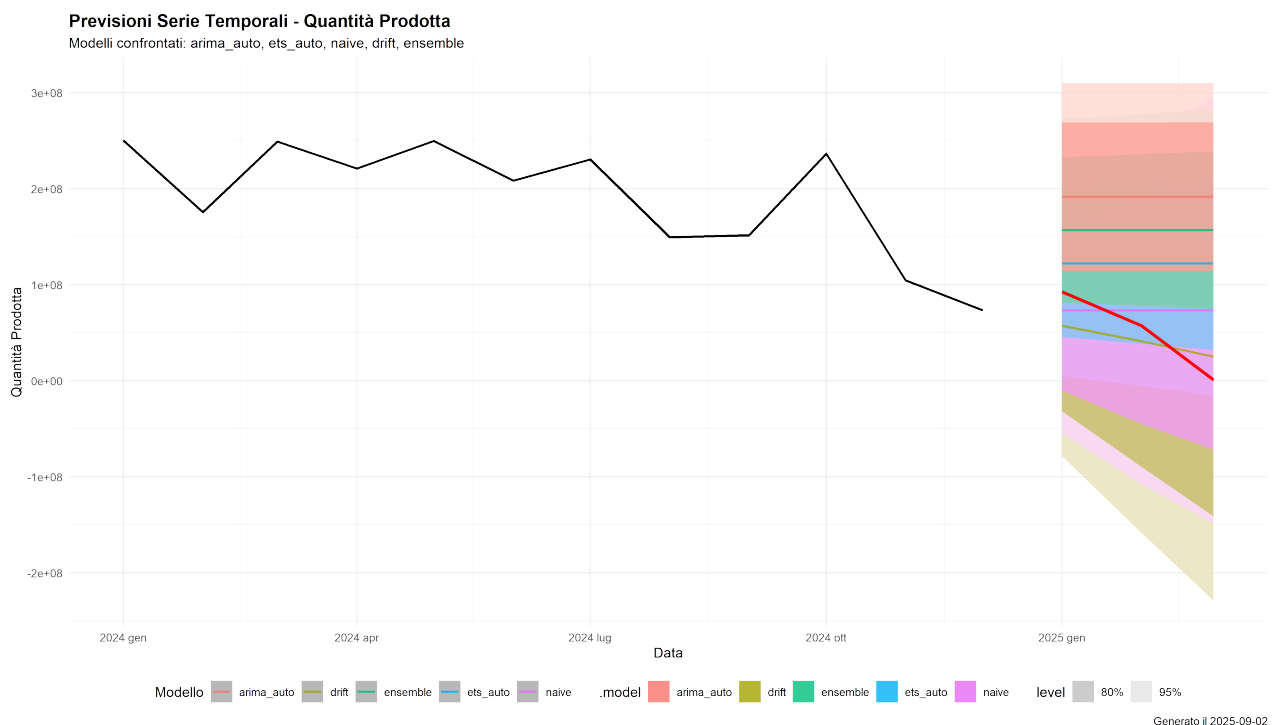


Figura 3: Confronto tra i modelli più promettenti e i dati reali (in rosso).

- **Descrizione:** Framework progettato per la produzione reale. È focalizzato sulla robustezza e sulla gestione completa degli errori, inclusi outlier, valori mancanti e dati sporchi. Può integrare modelli standard o ibridi e applicare vincoli realistici alle previsioni.

- **Punti di forza:**

- Massima stabilità anche in presenza di dati incompleti o rumorosi.
- Output pronto per l'uso operativo con formati di esportazione standard (CSV, grafici, modelli salvati).
- Intervalli di confidenza realistici e adattivi per la pianificazione operativa.

- **Limiti:**

- Minor precisione rispetto all'Hybrid Model in contesti con pattern complessi ma dati puliti.
- Richiede una configurazione iniziale più complessa per i parametri di robustezza.

- **Utilizzo ottimale:**

- Implementazioni in produzione, dove la continuità e la gestione automatica degli errori sono cruciali.
- Dati problematici, incompleti o soggetti a outlier frequenti.
- Pianificazione operativa con previsioni affidabili e vincoli realistici.

## 6 Criticità e Problemi Identificati

### 6.1 Problemi nei Dati

- **Inconsistenza formato:** Valori mancanti con encoding diversi
- **Qualità dati:** Presenza di valori nulli, negativi o anomali
- **Stagionalità:** Periodi insufficienti per modelli stagionali complessi
- **Regressori:** Disponibilità inconsistente della variabile `n.commesse`

### 6.2 Limitazioni Tecniche

- **Compatibilità package:** Conflitti tra `fpp3`, `forecast` e `prophet`
- **Gestione errori:** Necessità di approcci robusti per fallback
- **Scalabilità:** Performance su dataset di grandi dimensioni
- **Complessità:** Balance tra sofisticazione e manutenibilità

### 6.3 Sfide Metodologiche

- **Model Selection:** Variabilità delle performance tra modelli
- **Forecast Horizon:** Accuratezza decrescente con l'orizzonte temporale
- **External Regressors:** Stima dei valori futuri per variabili esogene
- **Ensemble strategies:** Combinazione ottimale di modelli diversi

## 7 Raccomandazioni d'Uso

### 7.1 Matrice Decisionale

Scenario	ARIMA Auto	Hybrid Model	Robust Framework
Analisi esplorativa rapida	Raccomandato	-	-
Dati con relazioni quantità-commesse	-	Raccomandato	Alternativa
Implementazione produzione	-	-	Raccomandato
Risorse computazionali limitate	Raccomandato	-	-
Necessità interpretabilità	Raccomandato	Parziale	Parziale
Dati problematici/mancanti	-	Gestisce	Raccomandato

Tabella 2: Matrice decisionale per selezione approccio

### 7.2 Workflow Consigliato

1. **Fase Esplorativa:** Iniziare con ARIMA\_AutoSelection\_7M.R per baseline
2. **Fase Sviluppo:** Se relazioni quantità-commesse evidenti, testare HybridModel\_Advanced\_Forecas
3. **Fase Produzione:** Implementare RobustForecasting\_Production.R per robustezza operativa
4. **Monitoraggio:** Confronto periodico performance tra approcci

## 8 Roadmap e Sviluppi Futuri

### 8.1 Fase 1: Consolidamento

- Stabilizzazione pipeline dati per tutti e tre gli approcci
- Automazione controlli qualità integrati
- Implementazione monitoring performance comparativo
- Documentazione operativa per ciascun script

### 8.2 Fase 2: Enhancement

- Integrazione dati esterni (ISTAT, ASSOCARTA) in framework robusto
- Ottimizzazione HybridModel per performance computational
- Implementazione ensemble dinamico che combina i tre approcci
- Feature engineering automatico cross-model

### 8.3 Fase 3: Operazionalizzazione

- Integrazione con Qlik Sense via SSE per tutti gli approcci
- Dashboard comparativo real-time per decision support
- Sistema di alert e recommendation basato su ensemble
- API unificata per accesso ai tre modelli

## 9 Conclusioni

Il progetto ha sviluppato un ecosistema completo di soluzioni di forecasting che copre diverse esigenze operative, dall'analisi esplorativa rapida all'implementazione enterprise robusta. I tre approcci implementati offrono flessibilità nella scelta della metodologia più appropriata per ciascun contesto d'uso.

L'approccio `ARIMA_AutoSelection_7M.R` fornisce una baseline solida e interpretabile, il `HybridModel_Advanced_Forecast.R` sfrutta le specificità del dominio per accuratezza superiore, mentre il `RobustForecasting_Production.R` garantisce affidabilità operativa in ambienti produttivi.

La strategia multi-approccio permette di bilanciare accuratezza, robustezza e complessità operativa, fornendo agli utenti finali la flessibilità di scegliere la soluzione più adatta alle proprie esigenze specifiche. Il sistema è pronto per il deployment in ambiente di test con monitoraggio comparativo delle performance tra i diversi approcci.

## 10 Confronto dei Modelli

### 10.1 Analisi Comparativa delle Architetture

I due modelli ibridi sviluppati rappresentano approcci complementari alla previsione delle serie temporali di produzione, con livelli crescenti di sofisticazione e complessità computazionale.

#### 10.1.1 Modello Personalizzato Ibrido (`HybridModel_Advanced_Forecast`)

Il primo modello implementa un approccio ibrido fondamentale che combina:

- **Componente deterministico:** Modellazione della relazione quantità/numero commesse tramite regressione quadratica
- **Stagionalità adattiva:** Decomposizione STL robusta per pattern stagionali
- **Dinamiche stocastiche:** ARIMA sui residui per catturare variazioni non spiegate
- **Rilevamento breakpoint:** Detection semplice dei cambiamenti strutturali
- **Ensemble pesato:** Combinazione con pesi fissi basati su performance storica

Le caratteristiche tecniche includono circa 15-20 variabili base, validazione rolling window semplificata e change point detection tramite sliding window t-test. Gli intervalli di confidenza sono basati su volatilità storica con crescita temporale lineare.

### 10.1.2 Modello Ibrido Avanzato con Machine Learning (HybridMachineLearning)

Il secondo modello estende significativamente l'architettura base integrando:

- **Machine Learning Ensemble:** Random Forest e XGBoost per pattern non-lineari complessi
- **Feature Engineering automatico:** Oltre 40 variabili derivate includendo lag features, moving averages, encoding ciclico, ratios e volatilità rolling
- **Change Point Detection avanzato:** Algoritmo PELT per rilevamento simultaneo di cambiamenti in media, varianza e trend
- **Cross-validation temporale:** Framework robusto a 5-fold per validazione performance
- **Pesi ensemble adattivi:** Aggiornamento dinamico basato su performance recente
- **Sistema di alert:** Rilevamento automatico anomalie e monitoraggio qualità previsioni

## 10.2 Matrice Comparativa delle Performance

Criterio	Modello Base	Modello Avanzato
Tempo di training	Veloce (2-5 minuti)	Medio (10-20 minuti)
Accuratezza	Buona per pattern lineari	Superiore per pattern complessi
Interpretabilità	Alta - componenti chiari	Media-bassa - black-box ML
Robustezza	Media - gestione errori base	Alta - validazione completa
Manutenzione	Semplice - poche dipendenze	Complessa - librerie ML
Requisiti dati	Minimo 24 osservazioni	Ottimale 30+ osservazioni
Gestione outlier	Standard - STL robust	Avanzata - ML ensemble
Adattabilità	Media - pesi fissi	Alta - pesi dinamici

Tabella 3: Confronto prestazioni e caratteristiche operative

## 10.3 Criteri di Selezione del Modello

### 10.3.1 Utilizzo Modello Base

Il modello personalizzato ibrido base è raccomandato quando:

- Dataset limitato ( $< 30$  osservazioni mensili)
- Necessità di interpretabilità completa delle previsioni
- Pattern di produzione relativamente stabili e lineari
- Risorse computazionali o competenze tecniche limitate
- Ambiente operativo che richiede semplicità di deployment

### 10.3.2 Utilizzo Modello Avanzato

Il modello ibrido avanzato con machine learning è preferibile per:

- Dataset consistente (30+ osservazioni con buona qualità)
- Priorità assoluta all'accuratezza delle previsioni
- Presenza di pattern complessi, non-lineari o multiple stagionalità
- Disponibilità di risorse per manutenzione e aggiornamento
- Contesto operativo che tollera complessità per performance superiori
- Necessità di feature importance analysis e insight automatici

## 10.4 Framework di Implementazione Consigliato

### 10.4.1 Approccio Graduale

Si raccomanda un approccio di implementazione progressiva:

1. **Baseline Assessment:** Iniziare con il modello base per stabilire performance di riferimento e validare qualità dati
2. **Complexity Evaluation:** Se i risultati base sono insufficienti o i pattern mostrano alta non-linearità, procedere con il modello avanzato
3. **A/B Testing:** Mantenere entrambi i modelli in parallelo per periodo di validazione comparativa
4. **Production Selection:** Scegliere il modello finale basandosi su trade-off accuracy/complexity per il contesto specifico

### 10.4.2 Metriche di Decisione

La selezione finale deve considerare:

- **Performance quantitative:** MAE, RMSE, MAPE su validation set
- **Robustezza operativa:** Stabilità previsioni, gestione missing data
- **Business value:** Impatto delle previsioni su decisioni operative
- **Sustainability:** Costi computazionali e di manutenzione nel lungo termine

Il modello base garantisce solidità e trasparenza, mentre il modello avanzato massimizza l'accuratezza a costo di maggiore complessità. La scelta ottimale dipende dalle specifiche esigenze operative e dalla maturità dell'organizzazione nell'adozione di soluzioni analytics avanzate.