

Università degli studi di Padova



Facoltà di Scienze Statistiche
Corso di Laurea in Statistica e Gestione delle Imprese

Tesi di Laurea

**Modellazione e previsione di serie storiche delle vendite:
Il caso DAB PUMPS S.p.a.**

Relatore: Prof.ssa Luisa Bisaglia

*Laureanda: Dashevici Silvia
Matricola: 557664- GEI*

Anno Accademico 2008-2009

*Un sincero ringraziamento a tutti coloro
che mi sono stati vicini in questa
tappa importante
della mia vita.*

INDICE

Sommario	1
Capitolo 1. Mercato italiano delle elettropompe	3
1.1 Mondo delle elettropompe.....	3
1.2 L'industria italiana delle elettropompe.....	5
1.2.1 Tendenze a medio termine.....	7
1.3 Breve storia dell'azienda DAB PUMPS S.p.a.....	8
1.3.1 Le strategie di produzione dell'azienda.....	13
1.3.2 Ruolo delle previsione nella pianificazione della produzione.....	14
Capitolo 2. Metodologie utilizzate	15
2.1 Introduzione.....	15
2.2 Analisi delle serie storiche.....	16
2.3 Lisciamiento esponenziale.....	18
2.3.1 I metodi di Holt-Winters stagionali.....	19
2.4 L'approccio di Box-Jenkins.....	21
2.5 Valutazione delle previsioni.....	26
Capitolo 3. Il caso DAB PUMPS S.p.a	29
3.1 Dati.....	29
3.2 Modellazione delle serie storiche.....	29
3.3 Risultati finali.....	47
3.4 Valutazione economica delle previsioni.....	49
Bibliografia	53
Pacchetti statistici	53

Elenco delle figure

Figura 1.1. Stabilimento di Mestrino (Pd).....	8
--	---

Elenco delle tabelle

Tabella 1.1. Esportazioni italiani di pompe 2005-2007 (milioni di euro).....	6
Tabella 1.2. Produzione italiana di pompe 2005-2007 (milioni di euro).....	6
Tabella 1.3. Mercato industriale italiano delle pompe 2005-2007 (milioni di euro).....	6
Tabella 3.1. Confronto tra modello additivo e quello moltiplicativo.....	32
Tabella 3.2. Previsioni con il modello moltiplicativo.....	32
Tabella 3.3. Indici di bontà di adattamento dei modelli stimati.....	44
Tabella 3.4. Confronto tra il modello moltiplicativo e SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.....	45
Tabella 3.5. Previsioni elaborate con il modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche	45
Tabella 3.6. Previsioni del modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche aggiornate per ogni mese.....	46
Tabella 3.7. Modelli scelti per le categorie di prodotti e gli indici di bontà di adattamento delle previsioni.....	47
Tabella 3.8. Modelli scelti per i codici della categoria CP e gli indici di bontà di adattamento delle previsioni.....	48
Tabella 3.9. Modelli scelti per i codici della categoria FE e gli indici di bontà di adattamento delle previsioni.....	49
Tabella 3.10. Valutazione economica delle previsioni elaborate per le categorie di prodotti modellati.....	50
Tabella 3.11. Valutazione economica delle previsioni elaborate per i codici della categoria CP.....	50
Tabella 3.12 Valutazione economica delle previsioni elaborate per i codici della categoria FE.....	51

Elenco dei grafici

Grafico 1.1. Fatturato totale (milioni di euro).....	11
Grafico 1.2 Investimenti totali (milla euro).....	12
Grafico 1.3 Totale delle risorse umane.....	12
Grafico 3.1. Serie storica FD e serie lisciata. Modello additivo.....	30
Grafico 3.2. <i>Previsioni-in-sample</i> . Modello additivo.....	30
Grafico 3.3. Serie storica FD e serie lisciata. Modello moltiplicativo.....	31
Grafico 3.4. <i>Previsioni-in-sample</i> . Modello moltiplicativo.....	31
Grafico 3.5. <i>Previsioni-in-sample</i> del modello additivo e quello moltiplicativo	32
Grafico 3.6. Serie storica delle vendite per la categoria FD.....	33
Grafico 3.7. Funzioni di autocorrelazione globale (ACF) e parziale (PACF) della serie FD.....	34
Grafico 3.8. Correlogramma dei residui del modello SARIMA (1,0,1)(1,1,0)...	35
Grafico 3.9. <i>Previsioni-in sample</i> del modello SARIMA(1,0,1)(1,1,0).....	37
Grafico 3.10. Correlogramma dei residui del modello SARIMAX (2,0,0)(0,0,0) con variabili dummy.....	38
Grafico 3.11. <i>Previsioni-in-sample</i> del modello SARIMAX (2,0,0)(0,0,0) con variabili dummy.....	39
Grafico 3.12. Correlogramma dei residui del modello SARIMAX (1,0,1)(1,1,0) con la variabile prezzi.....	41
Grafico 3.13. <i>Previsioni-in sample</i> del modello SARIMAX(1,0,1)(1,1,0) con la variabile prezzi.....	41
Grafico 3.14. Correlogramma dei residui del modello SARIMAX (2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.....	43
Grafico 3.15. <i>Previsioni-in sample</i> del modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.....	43
Grafico 3.16. <i>Previsioni-in sample</i> dei modelli stimati.....	44

SOMMARIO

Il presente lavoro prende il via da uno *stage* della durata di tre mesi, svolto all'interno dell'ufficio di pianificazione della produzione presso la DAB PUMPS S.p.a, azienda multinazionale leader nella produzione di elettropompe idrauliche e circolatori.

L'obiettivo principale dello *stage* è stato quello di elaborare previsioni delle vendite, attraverso strumenti e tecniche statistiche atti ad ottenere delle previsioni più accurate, allo scopo di pianificare (possibilmente diminuendo) le scorte nei magazzini mantenendo, comunque, costante il livello di servizio.

Con questo fine sono state analizzate le serie storiche mensili delle vendite per categorie di prodotti e singoli codici. I metodi utilizzati per la modellazione delle serie storiche considerate, sono l'approccio Box-Jenkins (modelli SARIMA e SARIMAX) e i metodi di HOLT-WINTERS stagionali. Le ultime dodici osservazioni sono state utilizzate allo scopo di quantificare l'accuratezza delle previsioni dei modelli selezionati. In base a vari indici di bontà delle previsioni (EQM, EMA, indice di Theil), è stato scelto il modello finale per produrre le previsioni *out-of-sample*.

Nel capitolo 1 di questo lavoro viene fatta una breve presentazione del mercato italiano delle elettropompe, con particolare riferimento all'azienda DAB PUMPS S.p.a. Nell'ultimo sottoparagrafo di questo capitolo viene esposto il ruolo delle previsioni per l'area di pianificazione della produzione.

Nel capitolo 2 vengono brevemente descritti i metodi sopra menzionati per l'analisi delle serie storiche.

Nel capitolo 3 viene presentata l'analisi delle serie storiche considerate ed i risultati relativi ad esse. Il capitolo si conclude con la valutazione economica delle previsioni.

Capitolo 1

Mercato italiano delle elettropompe.

L'azienda DAB PUMPS S.p.a

1.1 Mondo delle elettropompe.

Si intende normalmente per pompa, il dispositivo usato per spostare fluidi. La pompa è una delle "macchine" più diffuse ed essenziali, oltre che tra le più antiche storicamente (la prima descrizione di una pompa si deve ad Archimede nel III sec a.C, *La vite di Archimede*) ma in genere non è mai visibile perchè situata all'interno dell'impianto idraulico.

L'industria delle elettropompe rappresenta il baricentro vitale dell'economia mondiale, essendo, le pompe, il "nocciolo" di ogni attività industriale. Esse vengono utilizzate ogni qualvolta il fluido deve essere sollevato, trasportato o rimosso. Ovunque, in ambienti domestici o industriali, ci sono elettropompe, siano esse per acqua calda (sistemi di riscaldamento) o per l'acqua fredda (movimentazione dell'acqua, irrigazione, sistemi di raffreddamento /condizionamento).

Le pompe sono disponibili in varie forme con caratteristiche diverse.

Queste possono essere classificate sostanzialmente in tre gruppi: pompe centrifughe, pompe alternative a pistone e pompe rotative. Esse sono disponibili in tutte le forme e dimensioni: da pompe che riesci a tenere in una mano a pompe molto grosse.

Le pompe centrifughe funzionano trasferendo l'energia cinetica di una rotazione della girante al fluido pompato dall'azione accelerata della girante sul fluido. In generale, esse offrono elevati tassi di flusso e funzionano meglio con fluidi puliti a bassa viscosità. Le pompe centrifughe sono le più ampiamente utilizzate e sono disponibili in tutte le forme e dimensioni, a seconda del loro uso.

Le pompe sommergibili hanno motori elettrici sigillati all'interno di un dispositivo a tenuta d'acqua, permettendo l'utilizzo delle pompe nei pozzi o nei serbatoi delle acque di fognatura.

Le pompe per drenaggio consentono di pompare liquidi contenenti particelle abrasive. Esse sono provviste di rivestimenti speciali (di gomma o di materiali molto resistenti) per prevenire danni erosivi ai componenti.

Le pompe "igieniche" sono realizzate in acciaio inossidabile e sono progettate per essere facilmente pulite senza smontaggio.

Le pompe alternative a pistone spostano i fluidi attraverso la reciprocità di un pistone. Il pistone può essere messo in funzione meccanicamente dall'azione idraulica oppure pneumatica. Esso può agire direttamente sul liquido o attraverso una membrana flessibile. La quantità di fluido pompato è pari a quello dello spostamento del pistone. Queste pompe sono in grado di effettuare pompaggi ad alta pressione, ma le portate sono generalmente inferiori a quelle di pompe centrifughe. Con ogni "corsa" del pistone viene pompato un volume preciso e ripetitivo; per questo motivo, queste pompe sono ideali per l'utilizzo in applicazioni di misurazione in cui quantità precise di un fluido sono aggiunte a un processo.

Le pompe rotative funzionano attraverso lo spostamento dei fluidi da una cavità. La cavità è creata dall'interazione di un rotore e il corpo della pompa. Il fluido viene spostato con il movimento della cavità da parte del rotore verso l'uscita. Le pompe rotative sono utili per il pompaggio di fluidi viscosi, ossia fluidi con alto contenuto di solidi e materiali delicati.

1.2 L'industria italiana delle elettropompe

L'industria globale delle pompe è dominata da poche aziende di grande dimensione. I *leader* di questo settore sono ITT Corporation, Ebara, Grundfos, Wilo, KSB e Pentair. Tutte le aziende nominate (con l'eccezione di Wilo) possiedono degli stabilimenti produttivi sul territorio nazionale italiano tramite filiali o aziende acquisite, come nel caso DAB S.p.a, che è stata acquisita da Grundfos.

Negli ultimi vent'anni, in tutto il mondo, si sono verificate importanti fusioni tra i produttori di pompe. Per fronteggiare la concorrenza dei prodotti a basso prezzo provenienti dall'Asia, infatti, molti piccoli produttori sono entrati a far parte di gruppi più grandi, con lo scopo di conseguire le massime economie di scala. L'Italia ha cercato di resistere a questa tendenza, continuando a competere sul prezzo, anche se molte aziende operano su piccola scala.

L'Italia è un produttore ed esportatore importante nel mercato mondiale delle elettropompe (vedi tabella 1.1). La maggior parte delle aziende produttrici è concentrata nel Nord-Italia, in particolare 47 stabilimenti si trovano in Lombardia, 42 in Emilia Romagna e 34 in Veneto. La produzione industriale totale di pompe e componenti nel 2007, in termini monetari, raggiunge 1,63 miliardi di euro (vedi tabella 1.2). Questo risultato è ottenuto dalle circa 160 aziende autonome di piccola e media dimensione e dalle aziende che appartengono alle grandi multinazionali citate in precedenza.

Il settore italiano delle pompe è specializzato in piccole e medie pompe centrifughe per acque pulite, elettropompe sommerse¹ ed elettropompe ad uso domestico.

¹ Elettropompe sommerse dette anche pompe sommergibili descritti nel paragrafo 1.1

Di seguito nelle tabelle 1.1, 1.2 e 1.3 vengono presentati alcuni dati sulla produzione italiana delle elettropompe.

	2005	2006	2007
Rotative	47,4	39,4	35,6
Alternative a pistone	173,0	198,3	222,7
Centrifughe	310,1	357,8	398,3
Componenti	488,2	524,2	586,3
TOTALE	1018,7	1119,7	1242,6

Tabella 1.1. Esportazioni italiani di pompe 2005-2007 (milioni di euro)

Fonte: Eurostat

	2005	2006	2007
Rotative	197,2	203,5	214,1
Alternative a pistone	217,8	232,8	272,2
Centrifughe	722,8	793,7	922,5
Componenti	203,8	219,9	223,6
TOTALE	1341,6	1449,9	1632,4

Tabella 1.2. Produzione italiana di pompe 2005-2007 (milioni di euro)

Fonte: Eurostat

	2005	2006	2007
Rotative	179,5	193,8	209,7
Alternative a pistone	66,9	66,0	78,2
Centrifughe	592,7	632,8	725,7
Componenti	59,6	49,1	15,6
TOTALE	898,7	941,7	1029,2

Tabella 1.3. Mercato industriale italiano delle pompe 2005-2007 (milioni di euro)

Fonte: Eurostat

1.2.1 Tendenze a medio termine

Le previsioni di sviluppo a medio termine sono positive (la gestione dell'acqua calda e la produzione di energia saranno temi rilevanti nei prossimi anni), anche se condizionate dai mutamenti in atto nell'assetto competitivo del settore, ovvero:

- la crescita, anche tecnologica, dei produttori dei Paesi asiatici (Cina, India);
- l'aumento delle iniziative imprenditoriali all'estero degli operatori nazionali;
- la crescita della concentrazione nel settore, con l'acquisizione di piccole e medie imprese specializzate, da parte degli operatori maggiori, nazionali e internazionali;
- le incognite relative alle scelte strategiche di localizzazione produttiva dei gruppi esteri *leader* di settore attualmente presenti in Italia con propri stabilimenti.

Le storiche peculiarità delle imprese nazionali, tra cui flessibilità, velocità d'innovazione e ampiezza di gamma, caratteristiche vincenti nel recente passato, potrebbero rivelarsi insufficienti per operare nel quadro competitivo che si sta delineando, nel quale l'innovazione tecnologica e la dimensione finanziaria e organizzativa, necessarie per reggere strutture commerciali e di servizio adeguate al mercato globale, saranno requisiti indispensabili.

Inoltre, il trend della domanda è condizionato dal quadro legislativo e, in particolare, dall'emanazione di incentivi specifici a sostegno delle tecnologie per la riduzione dei consumi energetici e per il controllo dell'ecosistema (trattamento acque).

L'inasprimento della concorrenza giocata sui prezzi determinerà una contrazione della redditività delle imprese in un contesto competitivo che richiede sempre maggiore risorse da dedicare all'attività commerciale, in particolare sui mercati esteri.

Fonte : Databank. Pompe (giugno 2007).

1.3 Breve storia dell'azienda DAB PUMPS S.p.a

DAB PUMPS S.p.a è un'azienda che dal 1975 progetta e costruisce elettropompe per rispondere alle sempre più esigenti richieste del mercato non solo con prodotti che si attestano ai più elevati livelli di qualità, ma anche con un servizio di assistenza efficiente e versatile.

Dal primo stabilimento produttivo a Mestrino (vedi fig. 1.1), oggi sede centrale, agli altri due stabilimenti produttivi a Brendola (VI) ed a Castello di Godego (TV), DAB, dalla nascita ad oggi, si è evoluta fino ad avere più di 500 dipendenti e 60.000 mq di superficie.

Il numero di elettropompe vendute ogni anno supera i 1.500.000 pezzi, con 800 modelli a catalogo.



Fig. 1.1: Stabilimento di Mestrino (Padova)

I prodotti DAB si possono suddividere in quattro tipi:

- Domestici

In questa tipologia entrano a far parte i prodotti ad utilizzo domestico che rispondono alle esigenze legate al giardinaggio, alla circolazione dell'acqua, alla cucina e al bagno.



- Civili



Sono prodotti disegnati per l'utilizzo civile con lo scopo di rispondere alle esigenze legate alle abitazioni (portare l'acqua ai diversi piani di un condominio), riscaldamento e condizionamento.

- Industriali

Sono prodotti progettati per l'ambito industriale che si adattano alle esigenze legate a installazioni di dimensioni considerevoli, che comportano elevati quantitativi di acqua da movimentare .



- Motori ed elettronica



A questa categoria appartengono i controlli elettronici e motori elettrici di sistemi di pompaggio e motori elettrici.

Nel 1996 DAB PUMPS viene acquistata da parte della multinazionale danese GRUNDFOS, azienda nota come *leader* mondiale nella produzione e vendita di elettropompe, la quale, con una produzione annuale di circa 10 milioni di unità, assorbe circa il 50% del mercato mondiale.

Poco tempo dopo inizia a far parte del gruppo anche un'altra giovane azienda italiana, nata a Bientina (Pisa) nel 1991: la LEADERS PUMPS S.p.a. Tale azienda produce una vasta gamma di elettropompe domestiche, ed è specializzata nella produzione di pompe da giardino, pompe sommerse e booster set² per il settore DIY (“fai da te”).

IL GRUPPO DWT

Nel 2008 entrano a far parte del gruppo le aziende WaCS, TESLA e ALMA. WaCS è l'azienda specializzata nella produzione di sistemi elettronici con tecnologie avanzate per il controllo idraulico (quadri elettronici³, inverter⁴, sistemi on/off⁵ etc.); TESLA si occupa della produzione di pompe sommerse e motori sommersi incapsulati; ALMA produce pompe per drenaggio di acque non chiare.

Nel 2009 con l'entrata nel gruppo di BRISAN (Sud-Africa), specializzata in pompe per il settore agricolo, nasce il gruppo DWT.

All'interno del gruppo DWT inizia una proficua collaborazione, che porta le aziende non solo a lavorare insieme per il miglioramento reciproco e l'ottimizzazione, ma anche a dividersi le nuove realtà e filiali che nascono nel mondo.

² I *booster set* sono delle pompe già associate ad un serbatoio per il liquido e/o ad un quadro di comando.

³ Le pompe infatti, recentemente, vengono spesso comandate da un *quadro* che può essere meccanico o elettronico.

⁴ *Inverter* è un apparato elettronico in grado di convertire corrente continua in corrente alternata con notevoli vantaggi a livello di consumi.

⁵ I controlli possono essere di tipo *on/off*, quindi un sistema di autospegnimento che blocca la pompa quando non è più richiesta pressione nell'impianto.

Dopo la continua crescita degli ultimi anni, DAB viene oggi riconosciuta come sinonimo di azienda capace di essere ai massimi vertici nel proprio settore, in ambito nazionale e internazionale.

Attualmente DAB conta otto filiali: sei nel territorio Europeo (Belgio, Olanda, Germania, Inghilterra, Spagna e Russia), una nel continente americano (Carolina del Sud) e una in Cina (Qingdao). Quest'ultima riveste anche un ruolo di produzione, essendo nata per soddisfare la sempre crescente domanda di mercato attraverso una maggiore capacità produttiva del gruppo.

Le vendite totali, gli investimenti e le risorse umane del gruppo dal 1998 fino al 2009 vengono rappresentate nei seguenti diagrammi a barre (grafici 1.1, 1.2 e 1.3).

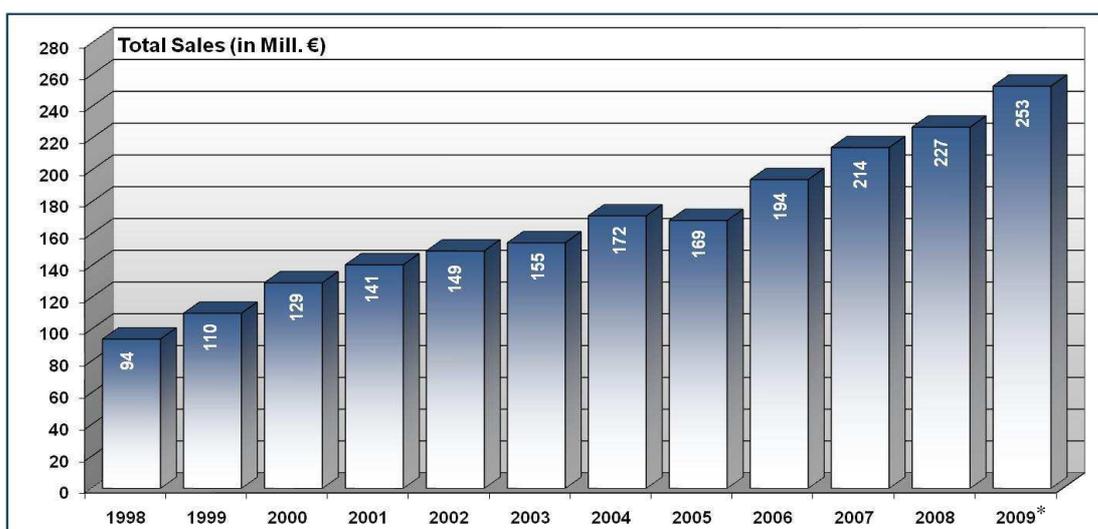


Grafico 1.1 Fatturato totale (milioni di euro)

Fonte: Azienda DAB PUMPS S.p.a.

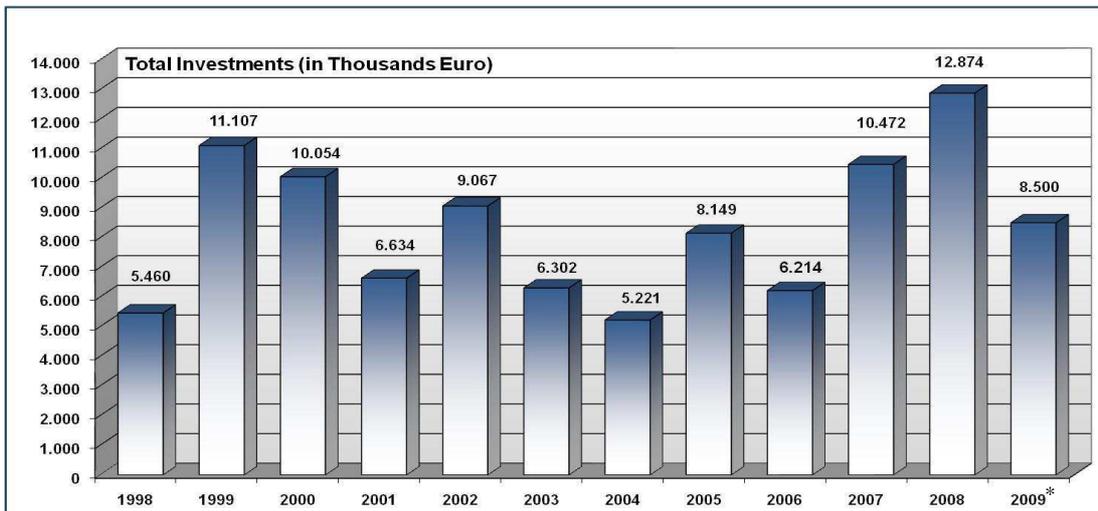


Grafico 1.2 Investimenti totali (milla euro)

Fonte: Azienda DAB PUMPS S.p.a.

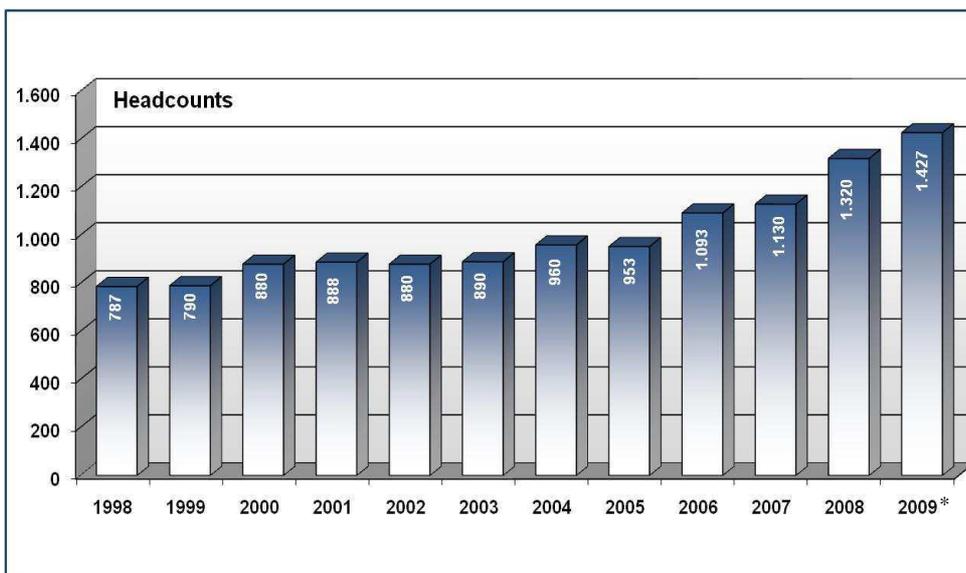


Grafico 1.3 Totale delle risorse umane

Fonte: Azienda DAB PUMPS S.p.a.

2009* - valori stimati

1.3.1 Le strategie di produzione dell'azienda

La DAB PUMPS S.p.a trova risposta alle proprie esigenze organizzative e di efficienza adottando le seguenti strategie di produzione :

- Produzione su ordine (*Make-to-order strategy*)
- Produzione per il magazzino (*Make-to-stock strategy*)
- Progettazione su ordine (*Engineer-to-order strategy*)

Secondo la strategia *make-to-order* i prodotti finiti e i semilavorati vengono realizzati in base agli ordini cliente acquisiti, mentre gli acquisti dei materiali vengono realizzati per il magazzino. Questa strategia viene applicata per bassi volumi di produzione e per prodotti personalizzati che richiedono un processo di produzione flessibile.

La strategia *make-to-stock* presuppone la produzione in *stock* in base a previsioni di vendita. Gli ordini di vendita devono essere coperti o dalla quantità prodotta secondo le previsioni o dagli stock attuali. Ciò significa che è importante elaborare previsioni accurate, in quanto l'obiettivo principale è quello di rispondere tempestivamente ai fabbisogni dei clienti, gestendo nel modo più efficiente le scorte dei magazzini e dei depositi; è ben noto, infatti, che lo stoccaggio dei prodotti comporta dei costi per l'azienda. Questa strategia viene adottata per prodotti standard con alti volumi di produzione.

Secondo la strategia *engineer-to-order* la produzione viene realizzata a fronte di uno specifico ordine, includendo varianti e personalizzazioni che richiedono la riprogettazione completa o parziale di un determinato prodotto.

1.3.2 Ruolo delle previsioni nella pianificazione della produzione

La pianificazione è la funzione aziendale che ha il ruolo di produrre un piano delle azioni future.

Nel caso DAB, l'ufficio pianificazione elabora il budget annuale e semestrale di produzione partendo dal budget commerciale, conciliandolo con le capacità produttive e il livello di stock desiderato.

Ogni mese viene sviluppato il piano previsionale per i successivi tre mesi dei prodotti *make-to-stock* e successivamente il *master production plan*⁶ mensile.

Il piano previsionale contiene i dati sui prodotti e le quantità necessarie da produrre.

Attualmente viene sviluppato partendo dalle previsioni di vendita, in gran parte fondate su informazioni qualitative derivante dai dati storici unite alle informazioni provenienti dall'ufficio commerciale sugli ordini ricevuti e sulle offerte in corso per i mesi successivi.

Le previsioni così ottenute vengono incrociate con la disponibilità di prodotti finiti presente a stock e con la disponibilità desiderata per i periodi successivi, il tutto adattato alle reali capacità produttive dei diversi stabilimenti.

Una volta elaborato il piano, i dati vengono comunicati agli uffici acquisti, *procurement*⁷ e programmazione dove vengono acquisite e programmate le risorse necessarie per la produzione.

⁶ *Master production plan* è un piano di produzione che indica quando e quanto di ciascun prodotto sarà richiesto.

⁷ Il *procurement* è l'attività di acquisto da parte degli addetti di un'azienda di beni non legati al ciclo produttivo dell'azienda stessa (che vengono trattati dall'Ufficio Approvvigionamenti).

Capitolo 2

Metodologie utilizzate

2.1 Introduzione

| “L’aumento della dimensione e complessità delle organizzazioni pubbliche e private, l’ampliamento dei mercati e della competizione, sono alcuni fattori alla base della crescente domanda di previsione di grandezze economiche. Con tale termine ci si riferisce comunemente alla prefigurazione di una realtà futura ma, più in generale, prevedere significa indurre un fatto (ignoto) da un altro fatto (noto) attraverso l’intermediazione di una legge. Quanto “noto” è costituito sia da misure della realtà osservata sia da informazioni qualitative e conoscenze teoriche; l’ “ignoto” può essere indifferentemente rappresentato da fatti futuri, da grandezze già realizzate ma non misurate, o persino da un passato incognito” (Guizzardi, A. 2001)

Al giorno d’oggi, per un’azienda che produce e commercializza prodotti è importante avere a disposizione previsioni affidabili sulle vendite, in quanto l’informazione sull’andamento delle stesse viene impiegata direttamente o indirettamente nelle aree di produzione, pianificazione, vendite, amministrazione.

Per questo motivo l’azienda DAB PUMPS S.p.a ha deciso di effettuare un’analisi dettagliata sulle serie di vendite delle varie categorie e dei singoli codici dei prodotti avvalendosi di strumenti e tecniche statistiche di previsione.

2.1 Analisi delle serie storiche

Si definisce serie storica una successione di dati numerici nella quale ogni dato è associato ad un particolare istante o intervallo di tempo (Vianelli, 1983). L'ordine temporale è essenziale, poiché si presume che vi sia una struttura di dipendenza temporale tra le osservazioni della serie. Lo scopo dell'analisi statistica di una serie storica è proprio quello di modellare tale dipendenza e cercare di spiegare il meccanismo casuale che l'ha generata.

Le realizzazioni di un fenomeno economico Y_t possono essere espresse mediante la somma di una parte prevedibile (f_t) e di una casuale (ε_t) che rappresenta la parte stocastica della serie :

$$Y_t = f_t + \varepsilon_t$$

L'analisi statistica delle serie storiche economiche scompone f_t in tre componenti : il trend (T), il ciclo (C) e la stagionalità (S).

Il *trend* rappresenta la tendenza di fondo del fenomeno considerato, determinata dall'evoluzione strutturale del sistema economico.

Con il termine *ciclo* si intendono invece fluttuazioni pluriennali a periodicità non casuale legate alle fasi di espansione e contrazione del sistema economico. A un periodo di espansione segue sempre un periodo di contrazione e viceversa, in modo che la dinamica descritta da questa componente è di tipo oscillatorio attorno ad un valore medio. Nel caso si utilizzino dati a frequenza sub-annuale (mensili, trimestri etc.), se i movimenti oscillatori si ripetono con una certa regolarità all'interno di un anno si parla di *stagionalità*.

I vari modelli di combinazione delle tre componenti della parte sistematica (f_t) sono (vedi Di Fonzo T., Lisi F., 2005):

1. Modello additivo:

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

In questo modello le quattro componenti sono tutte espresse nella stessa unità di misura di Y_t .

2. Modello moltiplicativo:

$$Y_t = T_t * C_t * S_t * \varepsilon_t$$

Il modello moltiplicativo può essere ricondotto a quello additivo mediante trasformazione logaritmica. In questo modello di solito la componente T_t è espressa nella stessa unità di misura di Y_t e le altre due sono espresse mediante coefficienti.

3. Modello misto :

$$Y_t = T_t * C_t * S_t + \varepsilon_t$$

In questo modello le componenti T_t e ε_t sono espresse nella stessa metrica di Y_t .

L'approccio decompositivo ha il pregio di essere concettualmente molto semplice ma soprattutto può essere un'utile guida nella specificazione di un modello di riferimento per la previsione anche quando si dispone di serie di lunghezza limitata.

Quest'approccio viene chiamato *classico*, perchè l'analisi si concentra sulla parte deterministica (f_t), invece si ipotizza che la parte stocastica (ε_t) sia stata generata da un processo *white noise*, ossia da una successione di variabili casuali indipendenti di media nulla e varianza costante.

A volte questo approccio non porta a risultati soddisfacenti, perchè non è sempre possibile individuare le tre componenti dall'andamento della serie. Per

questo motivo si preferisce *l'approccio moderno*. Esso si focalizza sulla modellazione della parte stocastica, ipotizzando che l'andamento della serie si basi su una certa distribuzione di probabilità.

Lo studio di tale successione trova la propria ragion d'essere nel fatto che la conoscenza di quanto è avvenuto determina ciò che avverrà secondo certe “regole probalistiche”. Questo accade per il presunto fatto che vi sia una struttura di dipendenza temporale tra le osservazioni della serie. Perciò \mathcal{E}_t viene considerato come un processo a componenti correlate.

Inoltre si può ipotizzare che nel campione considerato vi siano unità più informative di altre come, ad esempio, le ultime realizzazioni campionarie della serie.

Infatti i dati della serie sono quelli che “guidano” l'analista; il modello si deve adattare ai dati e non viceversa, come accade nell'approccio deterministico.

Si definisce processo stocastico una collezione di variabili casuali indicizzate dal tempo

$$\{Y_t, t = t_1, t_2, \dots\} \text{ (vedi Di Fonzo T., Lisi F., 2005).}$$

In altre parole, si può affermare che una data serie temporale è una particolare realizzazione di un processo stocastico.

2.1 Il lisciamiento esponenziale

Il lisciamiento esponenziale (*l'Exponential Smoothing*) rappresenta un utile strumento di previsione puntuale, soprattutto quando si hanno a disposizione pochi dati. Tale procedura si basa sull'idea che una ragionevole previsione del valore di una serie X al tempo t possa essere costituita da una combinazione lineare della previsione fatta sulla stessa serie nell'istante precedente. Tale combinazione lineare deve però tener conto della variazione registrata nell'unità temporale precedente tra l'effettivo valore della serie e la previsione realizzata.

Quindi si ha:

$$F_{n,1} = \delta \cdot F_{n-1,1} + (1 - \delta) \cdot y_n$$

ossia la nuova previsione (al tempo $n+1$) può essere pensata come la media pesata fra l'osservazione al tempo n (ultima osservazione disponibile) e la vecchia previsione (relativa al tempo n formulata al tempo $n-1$).

Il valore assegnato al parametro δ è la chiave dell'analisi. Se si desidera che la previsione sia stabile e che le variazioni casuali della serie siano smussate, allora si sceglie un valore per δ vicino ad uno e viceversa se si vuole dare maggior peso alle osservazioni più recenti, si opta per un valore vicino a zero.

2.3.1 I metodi di Holt-Winters stagionali

Una tecnica di previsione, utile quando si devono modellare le serie storiche composte da una componente di trend ed eventualmente da un fattore stagionale, è la tecnica Holt-Winters . Questa costituisce la naturale generalizzazione della tecnica del lisciamento esponenziale presentato nel paragrafo 2.3. In questo sotto paragrafo saranno presi in considerazione i metodi stagionali.

Secondo tale approccio la serie storica è il risultato di tre componenti: il livello al tempo n (\bar{y}_n), il trend (T_n) e la stagionalità (S_n). Ci sono due metodi stagionali proposti da Holt-Winters: additivo e moltiplicativo.

Metodo stagionale additivo.

L'ipotesi di base è che in prossimità di n la previsione con orizzonte di previsione k può essere fornita dalla retta :

$$F_{n,k} = \bar{y}_n + T_n \cdot k + S_n , \text{ con } S_n \text{ fattore stagionale di periodo pari a } s.$$

Le stime di \hat{y}_n , \hat{T}_n , \hat{S}_n sono date dalle formule di aggiornamento :

$$\hat{y}_n = \alpha \cdot (\hat{y}_{n-1} + \hat{T}_{n-1}) + (1-\alpha) \cdot (y_n - \hat{S}_{n-s}), \quad 0 < \alpha < 1$$

$$\hat{T}_n = \beta \cdot \hat{T}_{n-1} + (1-\beta) \cdot (\hat{y}_n - \hat{y}_{n-1}), \quad 0 < \beta < 1$$

$$\hat{S}_n = \gamma \cdot \hat{S}_{n-s} + (1-\gamma) \cdot (y_n - \hat{y}_n), \quad 0 < \gamma < 1$$

Previsione k periodi in avanti :

$$F_{n,k} = \hat{y}_n + k \cdot \hat{T}_n + \hat{S}_{n+k-s}, \quad 1 \leq k \leq s$$

$$F_{n,k} = \hat{y}_n + k \cdot \hat{T}_n + \hat{S}_{n+k-2s}, \quad s+1 \leq k \leq 2s$$

Metodo stagionale moltiplicativo.

L'ipotesi di base è che in prossimità di n la previsione con orizzonte di previsione k può essere fornita dalla retta :

$$F_{n,k} = (\bar{y}_n + T_n \cdot k) \cdot S_n$$

Le stime di \hat{y}_n , \hat{T}_n , \hat{S}_n sono date dalle formule di aggiornamento :

$$\hat{y}_n = \alpha (\hat{y}_{n-1} + \hat{T}_{n-1}) + (1-\alpha) \cdot \frac{y_n}{\hat{S}_{n-s}}, \quad 0 < \alpha < 1$$

$$\hat{T}_n = \beta \cdot \hat{T}_{n-1} + (1-\beta) \cdot (\hat{y}_n - \hat{y}_{n-1}), \quad 0 < \beta < 1$$

$$\hat{S}_n = \gamma \cdot \hat{S}_{n-s} + (1-\gamma) \cdot \frac{y_n}{\hat{y}_n}, \quad 0 < \gamma < 1$$

Previsione k periodi in avanti :

$$F_{n,k} = (\hat{y}_n + k \cdot \hat{T}_n) \cdot \hat{S}_{n+k-s}, \quad 1 \leq k \leq s$$

$$F_{n,k} = (\hat{y}_n + k \cdot \hat{T}_n) \cdot \hat{S}_{n+k-2s}, \quad s+1 \leq k \leq 2s$$

I parametri α, β, γ possono essere fissati arbitrariamente sulla base delle caratteristiche della serie (per serie molto lisce si scelgono valori dei parametri prossimi a zero, per serie con osservazioni imprevedibili molto rilevanti, valori

dei parametri prossimi a uno) o seguendo il criterio oggettivo della minimizzazione di qualche funzione di costo sul periodo di osservazione. Per le serie analizzate è stato usato il criterio di minimizzazione dell'errore quadratico medio entro il campione di stima.

2.4 L'approccio di Box-Jenkins

Quando si analizzano fenomeni economici (non controllabili sperimentalmente) come il consumo, la produzione, il reddito etc., il modello per il processo generatore dei dati, utilizzato per inferire l'incognito, deve comprendere necessariamente una parte probabilistica ε_t per generalizzare la parte deterministica. Come è stato accennato nel paragrafo 2.2, la componente ε_t è la parte di maggior interesse nell'approccio moderno.

I modelli più utilizzati nell'analisi moderna delle serie storiche univariate sono quelli introdotti da Box e Jenkins (Box G. E., Jenkins G. M, 1976).

I due modelli fondamentali sono: il modello autoregressivo (AR) che consiste in una somma pesata di valori passati e di uno shock casuale contemporaneo ed il modello a media mobile (MA) che è il risultato di una serie di impulsi casuali.

Dalla combinazione di questi due modelli discende il modello ARMA (p,q) (*AutoRegressive Moving Average*). Esso assume che la variabile Y_t dipenda linearmente sia da p tempi precedenti della variabile stessa (parte autoregressiva), sia dai q valori passati del termine di errore ε (parte a media mobile). Quindi il modello così descritto può essere formalizzato nell'espressione:

$\phi(B) \cdot Y_t = \theta(B) \cdot \varepsilon_t$, dove

- B è l'operatore ritardo t.c. $B \cdot y_t = y_{t-1}$;
- $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, è l'operatore autoregressivo non stagionale di ordine p ;
- $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$, è l'operatore a media mobile non stagionale di ordine q ;
- ε_t è un processo *white noise* ($\varepsilon_t \sim WN(0, \sigma^2)$).

Per la stazionarietà e l'invertibilità, le radici delle equazioni caratteristiche associate ai polinomi $\phi(B)$ e $\theta(B)$ devono essere in modulo maggiori di 1.

L'interesse di unire un modello AR(p) e un modello MA(q) in un modello ARMA(p, q) è nato dall'esigenza di descrivere mediante un esiguo numero di parametri una serie storica, cosa non molto semplice utilizzando solamente un processo autoregressivo oppure a medie mobili.

La classe dei modelli ARMA viene ampliata dai modelli ARIMA(p, d, q) (*Autoregressive Integrated Moving Average*) che consentono di modellare i processi non stazionari, ovvero i processi che non mantengono la media o la variabilità costante. Una serie non stazionaria può essere resa stazionaria tramite d differenziazioni oppure tramite trasformazione, per esempio quella logaritmica, nel caso la serie non sia stazionaria in varianza.

Dato quindi un intero non negativo d , il processo ARIMA(p,d,q) può essere espresso come:

$$\phi(B) \cdot (1 - B)^d \cdot Y_t = \theta(B) \cdot \varepsilon_t$$

Se d è uguale a zero la serie non necessita di essere differenziata e si ritorna alla classe ARMA(p,q).

Per le serie storiche stagionali un notevole apporto è rappresentato dai modelli SARIMA o ARIMA stagionali che permettono di modellare la natura stocastica anche della stagionalità.

I modelli ARIMA stagionali sono stati introdotti da Box e Jenkins (1976) e sono dei modelli ARIMA (p,d,q) che, a loro volta, possiedono dei residui ε_t modellabili tramite una struttura ARIMA (P,D,Q)_s.

Nella sua forma più generale un processo SARIMA(p,d,q)*(P,D,Q)_s può essere scritto come:

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Y_t = \theta(B)\Theta(B^S)\varepsilon_t, \text{ dove}$$

- $\Phi(B^S) = (1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS})$ è l'operatore autoregressivo stagionale di ordine P;

- $\Theta(B^S) = (1 + \Theta_1 B^S + \Theta_2 B^{2S} + \dots + \Theta_Q B^{QS})$ è l'operatore a media mobile stagionale di ordine Q;

- $(1 - B^S)^D$ è l'operatore differenza di ordine D stagionale.

Con uso dell'operatore autoregressivo non stagionale di ordine p e con l'operatore a media mobile non stagionale di ordine q , si stima la correlazione tra i mesi contigui; invece con gli operatori autoregressivo stagionale di ordine P e con l'operatore a media mobile stagionale di ordine Q si riesce a modellare la correlazione tra gli stessi mesi di anni diversi.

Per le serie storiche di natura economica è utile considerare i modelli ARMAX (*AutoRegressive Moving Average with eXogenous variable*). Il modello ARMAX è un raffinamento di un ARMA in quanto riduce la varianza dell'ARMA con l'introduzione nel modello di una o più variabili esogene. Questo accade per il motivo che le variabili esterne contengono informazioni aggiuntive rispetto al "passato" della serie in grado di contribuire nella spiegazione della variabilità del modello.

La struttura del modello ARMAX in un forma compatta può essere espressa come:

$$\phi(B) \cdot Y_t = \theta(B) \cdot u_t + \varepsilon_t,$$

dove u_t è la variabile esogena e ε_t un processo *white noise*.

Invece per le serie storiche che presentano la stagionalità si possono applicare i modelli della classe SARMAX. Questi modelli non sono altro che un'estensione dei modelli ARMA per includere la stagionalità e l'informazione delle variabili esogene.

Invece se le serie storiche non sono stazionarie si utilizzano i modelli ARIMAX e SARIMAX, nel caso in cui le serie storiche presentino stagionalità.

La procedura di Box-Jenkins per costruire, a partire dai dati, un modello atto ad approssimare il processo generatore della serie storica consiste nelle seguenti fasi:

1. Analisi preliminare
2. Identificazione
3. Stima dei parametri
4. Controllo diagnostico

► Analisi preliminare

Si analizza l'andamento della serie tramite il grafico e si verifica se sono presenti dei valori anomali.

► Identificazione

Si specificano gli ordini $(p,d,q)*(P,D,Q)$ del modello. Ciò viene effettuato principalmente mediante analisi successive dei grafici delle funzioni autocorrelazione globale e parziale. Se le autocorrelazioni tendono ad annullarsi molto lentamente (o non si annullano affatto) è probabile che il processo generatore sia non stazionario. In questo caso la serie deve essere resa stazionaria. Se la serie presenta non stazionarietà in media si applica la differenziazione di ordine d (per i modelli ARIMA) oppure D (per i modelli SARIMA). Invece se la serie presenta non stazionarietà in varianza si effettua un'opportuna trasformazione della serie (la trasformazione deve avvenire prima della differenziazione).

► Stima dei parametri

Una volta fissato l'ordine dei parametri, si procede con la stima degli stessi, mediante la massimizzazione della funzione di verosimiglianza, nell'ipotesi che \mathcal{E}_t sia un processo *white noise*.

► Controllo diagnostico

Questa fase consiste nella verifica dell'adeguatezza del modello stimato tramite l'analisi dei residui. Il diagramma di dispersione dei residui al tempo t contro i residui al tempo $t+1$ non dovrebbe mostrare nessun andamento particolare. Dal correlogramma si verifica se c'è correlazione tra i residui; se i residui sono incorrelati, i valori delle funzioni di autocorrelazione si trovano entro le bande di confidenza al 5%. Per una verifica complessiva di assenza di correlazione si può utilizzare la statistica sviluppata da Ljung e Box (Ljung G. M., Box G. E., 1978).

2.1 Valutazione della bontà di adattamento delle previsioni.

La misura della qualità della previsione è generalmente data dal confronto con la realtà. Nella pratica, il campione viene suddiviso in due insiemi: l'insieme di stima e l'insieme di previsione. L'insieme di stima viene impiegato nella specificazione e stima del modello e l'insieme di previsione viene utilizzato nella valutazione della capacità previsiva *in sample*⁸ del modello.

Se viene assunta "l'universalità" nel tempo delle relazioni specificate nel modello, si può considerare che le *performances* delle *previsioni-out-of-sample*⁹ dipenderanno da quelli delle *previsioni-in-sample*.

Di seguito verranno presentati vari indici che quantificano la capacità previsiva del modello specificato.

⁸ Le *previsioni-in-sample* sono le previsioni simulate per il campione di previsione

⁹ Si definiscono *previsioni-out-of-sample* le previsioni ottenute al di fuori del campione considerato per la stima.

In primo luogo si definisce l'errore di previsione al tempo t come:

$$e_t = r_t - p_t, \text{ dove}$$

r_t è il valore realizzato (osservato) al tempo t e p_t è il valore previsto per il tempo t .

$$EQM = \sqrt{\frac{1}{n} \sum_{i=1}^n e_t^2} \text{ è l'errore quadratico medio,}$$

$$EMA = \frac{1}{n} \sum_{i=1}^n |e_t| \text{ è l'errore medio assoluto,}$$

$$U = \frac{EQM}{\sqrt{\frac{1}{n} \sum_{i=1}^n r_t^2}}, \text{ è l'indice di Theil (Theil H. 1965).}$$

L'indice di Theil varia tra 0 e infinito. Esso risulta nullo nel caso in cui le previsioni coincidono perfettamente con i valori osservati. Se i valori di U sono prossimi o maggiore di 1 significa che il modello specificato non si adatta affatto bene ai dati.

Inoltre, attraverso questo indice si può fare il confronto tra modelli diversi

Un altro indice è l'errore medio percentuale:

$$EMP = \frac{1}{n} \sum_{i=1}^n \frac{e_t}{r_t} \cdot 100 .$$

Esso indica la percentuale di sovrastima (se il valore è negativo) oppure di sottostima (se il valore è positivo) dei valori veri.

Capitolo 3

Caso aziendale DAB PUMPS S.p.a

3.1 Dati

L'analisi delle serie storiche è stata condotta sulle serie storiche mensili delle vendite per categoria di prodotti e i singoli codici. I dati presi in considerazione riguardano il periodo gennaio 1996 – febbraio 2009. Per i modelli SARIMAX, oltre alle serie storiche delle vendite, sono state utilizzate le serie storiche dei prezzi; all'inizio si intuiva, poi si è anche verificato, che la variabile prezzi contiene informazione significativa nella spiegazione della variabilità del modello.

L'insieme di previsione è stato costituito dalle ultime dodici osservazioni.

Di seguito verranno analizzate le serie storiche per categoria di prodotti con i modelli di Holt-Winters e SARIMA/SARIMAX.

3.2 Modellazione delle serie storiche

Durante lo *stage* sono state analizzate in totale una quarantina di serie storiche delle vendite per varie categorie di prodotti (circolatori, linea d'acqua fredda, domestic booster) e per singoli codici. In questo paragrafo verrà elaborata la modellazione solo della serie storica delle vendite della categoria FD descritta nel seguito. Nel paragrafo 3.3 verranno presentati i risultati di tutte le altre serie storiche modellate.

La categoria FD contiene dei prodotti che fanno parte della classe denominata Jet (linea d'acqua fredda).

- ✓ Metodo Holt-Winters stagionale

Nelle serie storiche delle vendite analizzate è presente la componente stagionale. Come è stato accennato nel paragrafo 2.3.1, il metodo Holt-Winters stagionale

stima attraverso il modello additivo e il modello moltiplicativo anche la componente stagionale.

Stima del modello additivo.

La stima del livello al tempo n (\hat{y}_n), del trend (\hat{T}_n) e della stagionalità (\hat{S}_n) sono state calcolate come indicato nel paragrafo 2.3.1.

In questo caso sono stati usati i seguenti valori iniziali :

per \hat{y}_{11} e \hat{y}_{12} sono state applicate le medie mobili semplici centrate di ordine 13;

$$\hat{T}_{12} = \hat{y}_{12} - \hat{y}_{11};$$

$$\hat{S}_i = y_i - \hat{y}_{12}, i=1, \dots, 12$$

invece α, β, γ sono stati scelti in modo da minimizzare l'EQM.

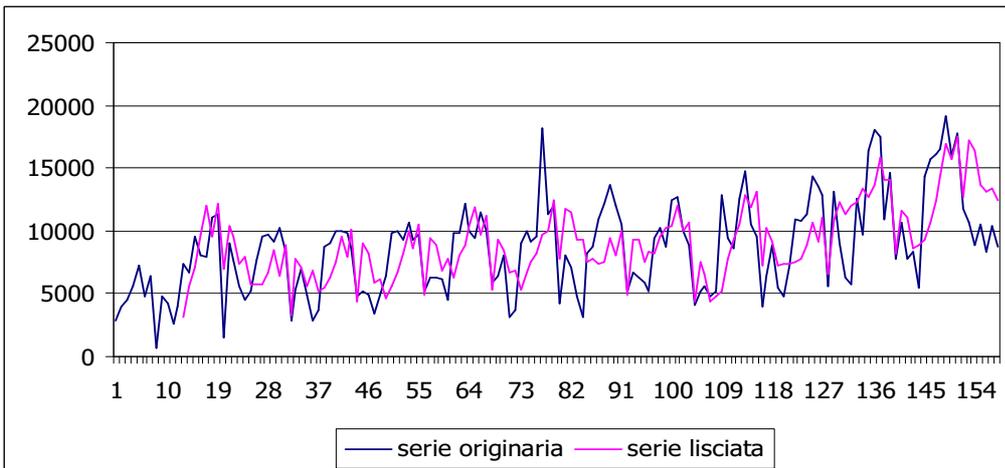


Grafico 3.1. Serie storica FD e serie lisciata. Modello additivo.

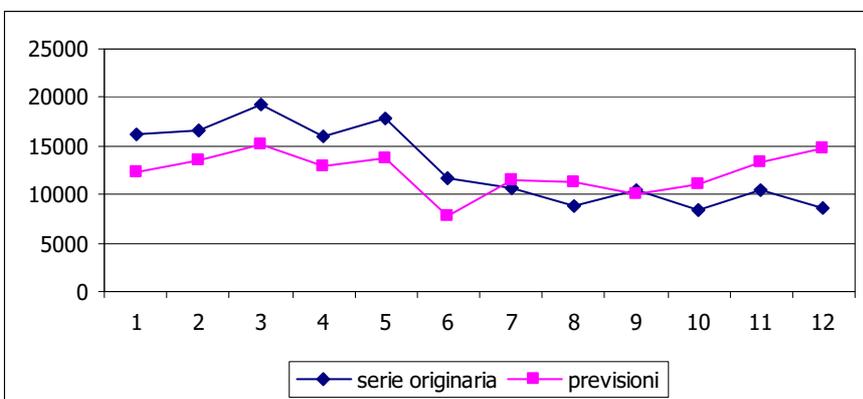


Grafico 3.2. Previsioni-in-sample. Modello additivo.

Stima del modello moltiplicativo

Per il modello moltiplicativo i valori iniziali delle stime del livello e trend sono gli stessi usati nel modello additivo, cambia solo la stima della stagionalità, che

$$\text{diventa } \hat{S}_i = \frac{y_i}{\hat{y}_{12}}, \quad i=1, \dots, 12.$$

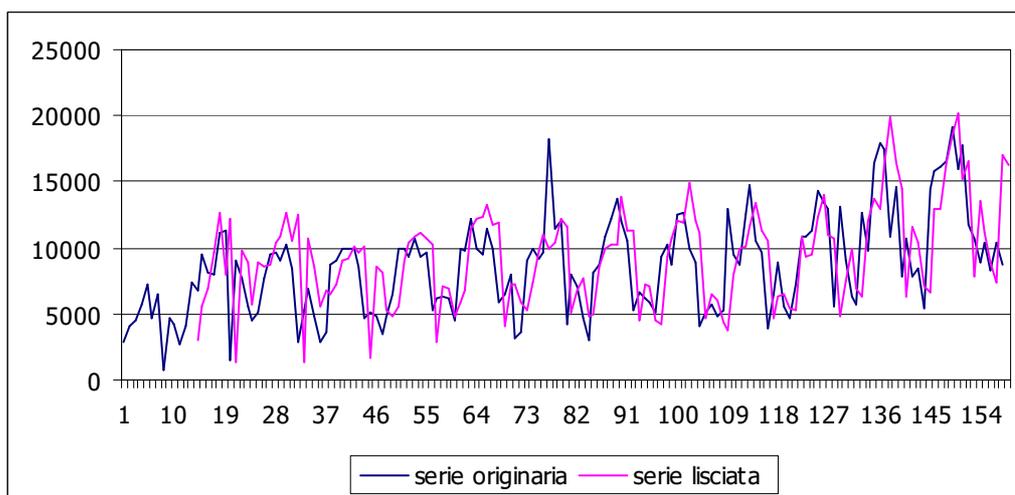


Grafico 3.3. Serie storica FD e serie lisciata. Modello moltiplicativo.

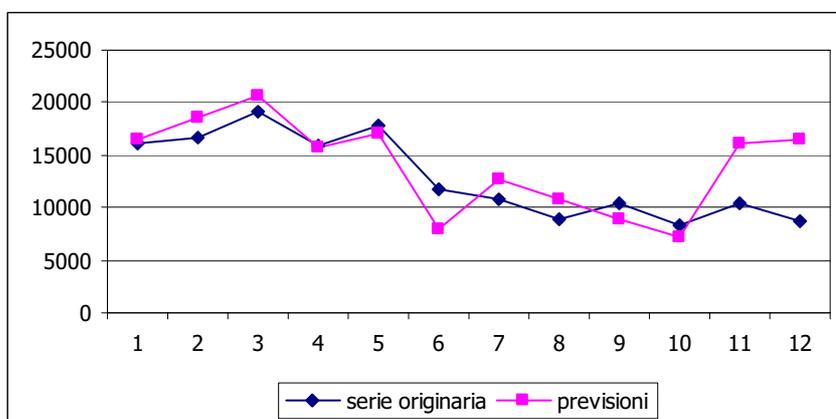


Grafico 3.4. *Previsioni-in-sample*. Modello moltiplicativo.

Si confrontano i due modelli utilizzando gli indici di bontà di adattamento descritti nel paragrafo 2.4.

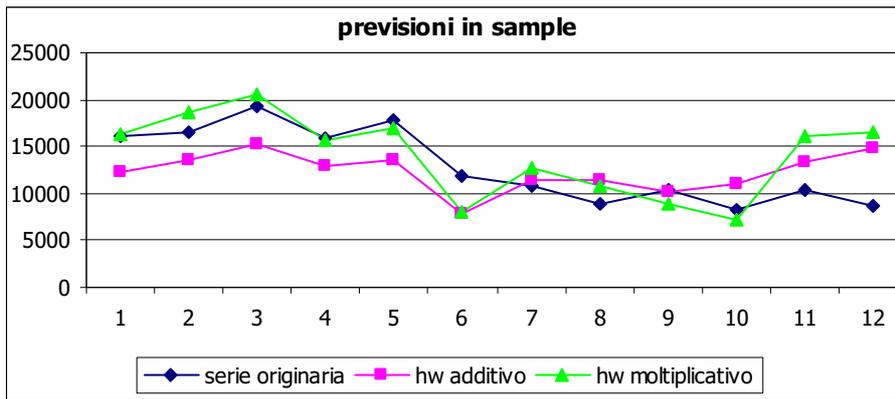


Grafico 3.5. *Previsioni-in-sample* del modello additivo e quello moltiplicativo.

	Modello additivo		Modello moltiplicativo	
	Insieme di stima	Insieme di previsione	Insieme di stima	Insieme di previsione
EQM	8997,01	3433,20	2057,23	3226,94
EMA	2312,57	3115,51	1539,09	2361,9
Indice di Theil		0,26		0,24

Tabella 3.1. Confronto tra modello additivo e quello moltiplicativo.

Dalla tabella 3.1 si osserva che il modello moltiplicativo presenta gli indici di bontà di adattamento minori e quindi, da un punto di vista previsivo, risulta essere il modello migliore fra i due.

Previsioni

Le previsioni sono state elaborate per i successivi tre mesi.

Anno	Mese	Previsione
2009	Marzo	16.673
	Aprile	18.074
	Maggio	20.249

Tabella 3.2. Previsioni con il modello moltiplicativo.

✓ L'approccio Box-Jenkins.

La modellazione delle serie storiche tramite i modelli SARIMA/SARIMAX è stata eseguita mediante il software di libero dominio *Gretl*.

Analisi preliminare.

La modellazione di una serie storica viene generalmente iniziata con l'analisi del suo andamento.

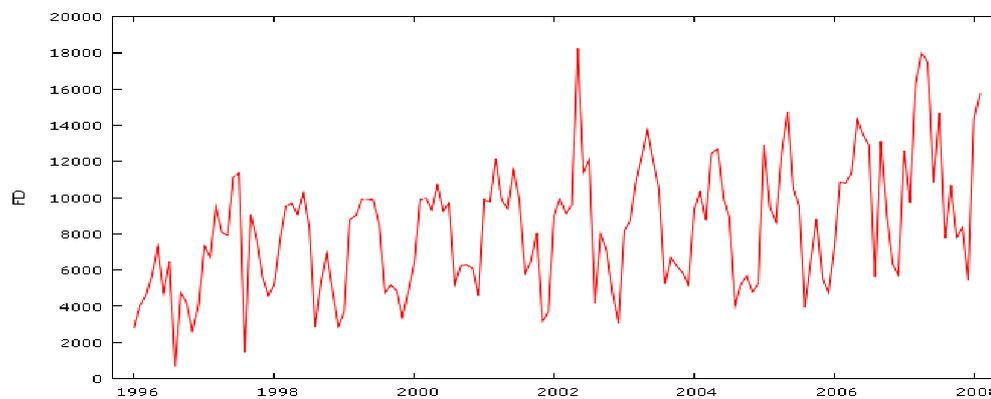


Grafico 3.6. Serie storica delle vendite per la categoria FD.

Dal grafico 3.6 si osserva che la serie storica presenta un trend leggermente crescente. La serie storica sembra essere non stazionaria in media.

Identificazione

In questa fase viene analizzata il correlogramma della serie storica.

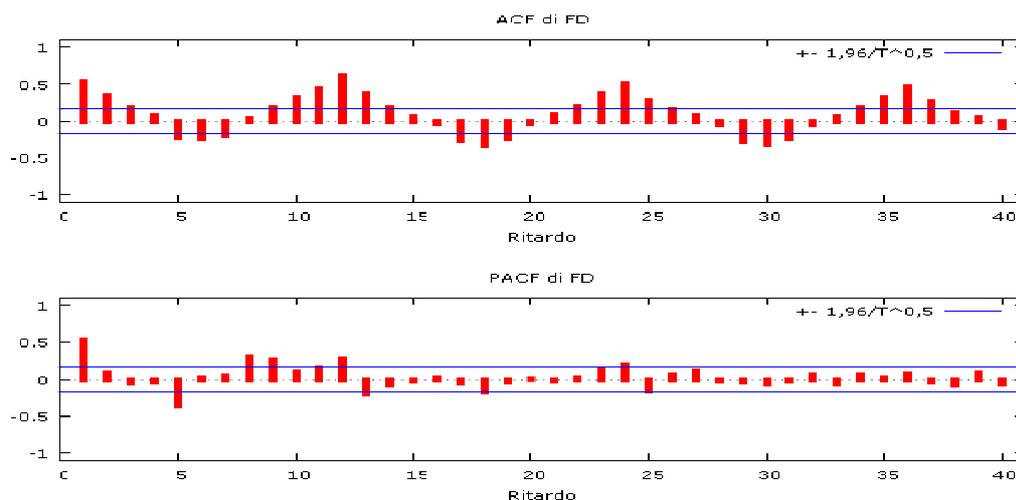


Grafico 3.7. Funzioni di autocorrelazione globale (ACF) e parziale (PACF) della serie FD.

La funzione di autocorrelazione globale (ACF) conferma che la serie storica non è stazionaria (l'ACF ai ritardi stagionali tende molto lentamente a zero); quindi la serie storica verrà resa stazionaria tramite una differenziazione stagionale.

Dall'analisi delle funzioni di autocorrelazione e dopo aver provato diversi modelli, si è giunto alla conclusione che un buon modello potrebbe essere il modello SARIMA (1,0,1)(1,1,0) di cui si riportano i risultati relativi alla stima dei parametri.

Stima dei parametri

	coefficiente	errore std.	rapporo t	p-value	
const	752,716	352,805	2,134	0,0329	**
phi_1	0,900856	0,0905308	9,951	2,50E-23	***
Phi_1	-0,447609	0,0861159	-5,198	2,02E-07	***
theta_1	-0,725201	0,131942	-5,496	3,88E-08	***

Reale Immaginario Modulo

AR			
Radice 1	1,1101	0,0000	1,1101
AR (stagionale)			
Radice 1	-2,2341	0,0000	2,2341
MA			
Radice 1	1,3789	0,0000	1,3789

Controllo diagnostico

Si può notare che tutti i parametri sono significativi al 5% e le radici delle equazioni caratteristiche associate ai polinomi $\phi(B)$, $\Phi(B)$ e $\Theta(B)$ sono in modulo maggiori di 1, quindi le condizioni di stazionarietà e invertibilità sono soddisfatte.

Per quanto riguarda l'analisi dei residui sono state considerate le funzioni di autocorrelazione dei residui e il test di Ljung-Box.

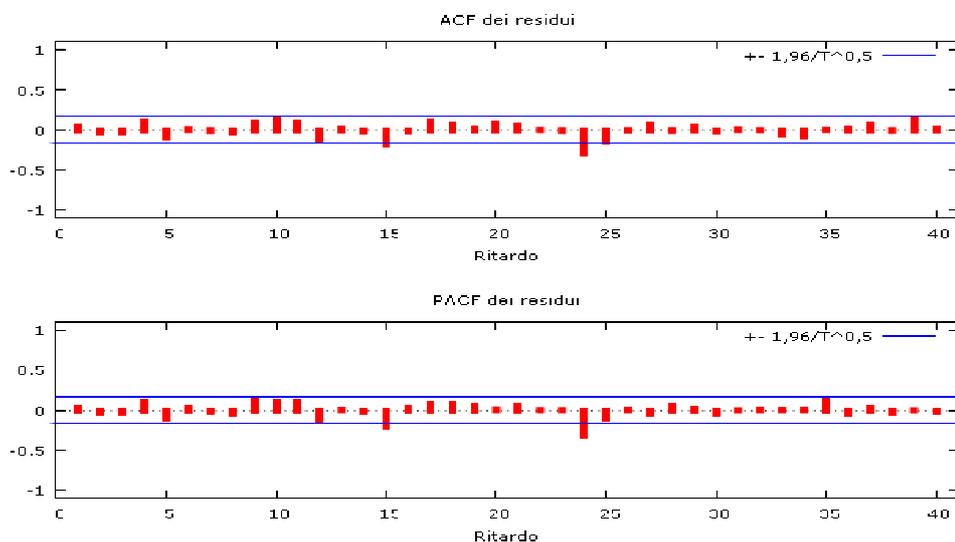


Grafico 3.8. Correlogramma dei residui del modello SARIMA (1,0,1)(1,1,0).

Nonostante i valori della funzione di autocorrelazione ai ritardi 15 e 24 (ritardo stagionale) siano esterni alle fasce di Bartlett (bande di confidenza al 5%) si può accettare l'ipotesi che i residui siano incorrelati ad un livello di significatività del 5%.

I valori della statistica di Ljung-Box:

LAG	ACF		PACF		Q-stat.	[p-value]
1	0,0334		0,0334		0,153	[0,696]
2	-0,0422		-0,0434		0,3991	[0,819]

3	-0,0353		-0,0325		0,5723	[0,903]
4	0,1076		0,1085		2,1962	[0,700]
5	-0,1021		-0,1144		3,6688	[0,598]
6	0,0149		0,0325		3,7002	[0,717]
7	-0,0138		-0,0175		3,7277	[0,811]
8	-0,0364		-0,0547		3,919	[0,864]
9	0,0983		0,1319		5,3281	[0,805]
10	0,1381		0,1083		8,1292	[0,616]
11	0,095		0,1028		9,4666	[0,579]
12	-0,1298		-0,1186		11,9833	[0,447]
13	0,0153		0,008		12,0185	[0,526]
14	-0,0197		-0,0283		12,0776	[0,600]
15	-0,1857	**	-0,2117	**	17,361	[0,298]
16	-0,0276		0,0421		17,4784	[0,355]
17	0,1116		0,0829		19,4194	[0,305]
18	0,0731		0,0763		20,2588	[0,318]
19	0,0268		0,0575		20,3723	[0,372]
20	0,086		0,0121		21,5557	[0,365]
21	0,0486		0,049		21,9371	[0,403]
22	0,0041		0,0034		21,9399	[0,463]
23	-0,013		0,002		21,9678	[0,522]
24	-0,2922	***	-0,3091	***	36,1173	[0,053]
25	-0,1597	*	-0,1101		40,3808	[0,027]
26	-0,0087		0,009		40,3935	[0,036]
27	0,0658		-0,0502		41,1305	[0,040]
28	-0,0129		0,0443		41,1593	[0,052]
....						
37	0,0629		0,0363		44,2033	[0,194]
38	-0,0107		-0,0347		44,2251	[0,225]
39	0,1332		-0,0014		47,6279	[0,162]
40	0,0173		-0,026		47,6861	[0,189]

I p-value della statistica di Ljung-Box sono quasi tutti maggiori del 5%, quindi l'ipotesi nulla di incorrelazione tra i residui non viene rifiutata.

Adesso si può passare alla fase di previsione.

Le *previsioni-in sample* risultano:

Oss	FD	Previsione
2008.03	16.134	16.161
2008.04	16.593	17.174
2008.05	19.183	18.122
2008.06	15.931	14.002
2008.07	17.760	15.751
2008.08	11.771	8.601
2008.09	10.704	13.488
2008.10	8.813	10.040
2008.11	10.466	9.068
2008.12	8.335	7.123
2009.01	10.406	15.091
2009.02	8.693	14.534

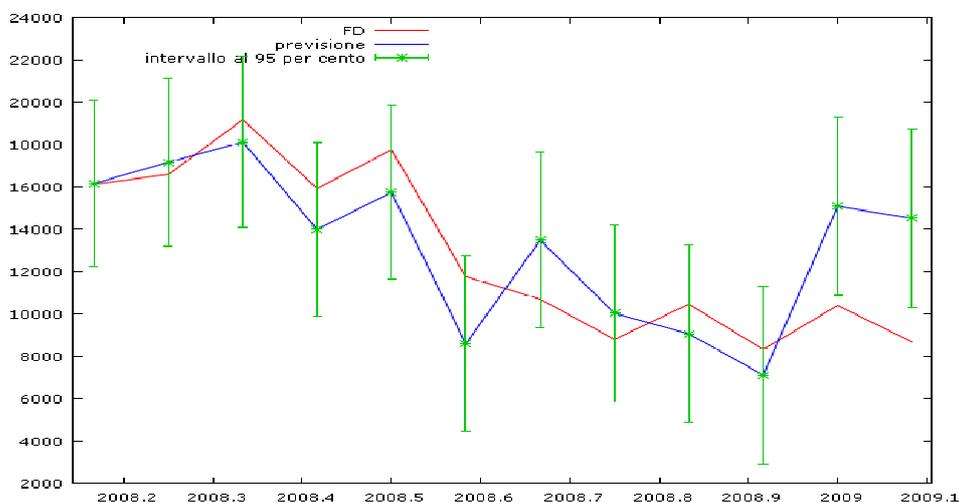


Grafico 3.9. *Previsioni-in sample* del modello SARIMA(1,0,1)(1,1,0).

Dal grafico 3.9 si osserva che il modello è in grado di prevedere l'andamento della serie originaria abbastanza bene; tutte le osservazioni con l'eccezione delle ultime due sono interne all'intervallo di confidenza al 95%.

In aggiunta ai modelli SARIMA vengono utilizzati i modelli SARIMAX. Come variabili esterne sono state impiegate le variabili dummy mensili (per cogliere la stagionalità) e la variabile prezzo medio unitario per categoria.

Stima dei parametri del modello SARIMAX (2,0,0)(0,0,0).

In questo modello sono state usate le variabili dummy mensili per modellare la parte stagionale.

	coefficiente	errore std.	rapporto t.	p-value	
phi_1	0,453042	0,0796	5,689	1,28E-08	***
phi_2	0,327088	0,0832	3,932	8,44E-05	***
dm1	8777,88	832,663	10,54	5,53E-26	***
dm2	9725,17	829,196	11,73	9,11E-32	***
dm3	10730,8	851,121	12,61	1,91E-36	***
dm4	11386,5	855,104	13,32	1,87E-40	***
dm5	12635,8	859,745	14,7	6,72E-49	***
dm6	10822,3	860,129	12,58	2,65E-36	***
dm7	10574,6	858,978	12,31	7,94E-35	***
dm8	4561,01	856,622	5,324	1,01E-07	***
dm9	7467,47	853,42	8,75	2,13E-18	***
dm10	7054,48	849,972	8,3	1,04E-16	***
dm11	5249,25	845,135	6,211	5,26E-10	***
dm12	4602,69	841,694	5,468	4,54E-08	***

Reale Immaginario Modulo

AR

Radice 1	1,1881	0,0000	1,1881
Radice 2	-2,5732	0,0000	2,5732

Controllo diagnostico.

Si osserva che tutti parametri stimati sono significativi all'1 %. Tramite i correlogrammi i residui si verifica che questi non siano correlati.

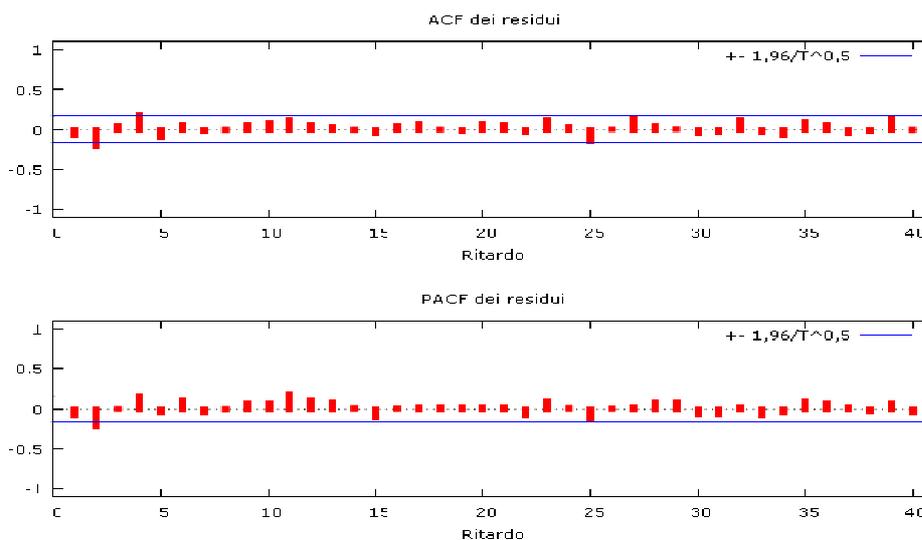


Grafico 3.10. Correlogramma dei residui del modello SARIMAX (2,0,0)(0,0,0) con variabili dummy.

Sebbene i valori della funzione di autocorrelazione al ritardo 2 e 4 siano esterni alle fasce di Bartlett ad un livello di significatività del 5%, si può accettare l'ipotesi nulla di incorrelazione dei residui.

Le previsioni-in-sample:

Oss	FD	Previsione
2008.03	16134	15.302
2008.04	16593	15.433
2008.05	19183	15.964
2008.06	15931	13.654
2008.07	17760	12.946
2008.08	11771	6.562
2008.09	10704	9.149
2008.10	8813	8.471
2008.11	10466	6.441
2008.12	8335	5.606
2009.01	10406	9.622
2009.02	8693	10.436

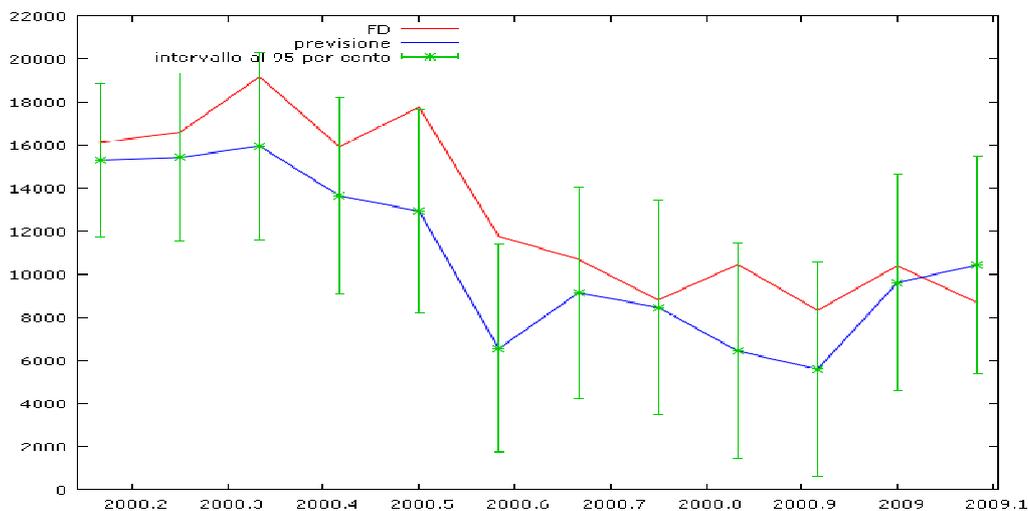


Grafico 3.11. *Previsioni-in-sample* del modello SARIMAX (2,0,0)(0,0,0) con variabili dummy.

Si può notare che il modello porta a sottostimare i valori reali, inoltre un'osservazione non appartiene all'intervallo di confidenza della previsione e un'altra si trova al limite dell'intervallo.

Stima dei parametri del modello SARIMAX (1,0,1)(1,1,0) con la variabile prezzi.

	coefficiente	errore std.	rapporto t.	p-value	
const	4030,7	2870,38	1,404	0,1602	
phi_1	0,921602	0,0714157	12,9	4,23E-38	***
Phi_1	-0,441835	0,0871138	-5,072	3,94E-07	***
theta_1	-0,738871	0,109805	-6,729	1,71E-11	***
prezzi_FD	-52,6137	45,3762	-1,159	0,2463	

Reale Immaginario Modulo

AR			
Radice 1	1,0851	0,0000	1,0851
AR (stagionale)			
Radice 1	-2,2633	0,0000	2,2633
MA			
Radice 1	1,3534	0,0000	1,3534

La costante e la variabile prezzi non sono significative. Si prova di stimare il modello senza la costante.

	coefficiente	errore std.	rapporto t.	p-value	
phi_1	0,892249	0,0961586	9,279	1,71E-20	***
Phi_1	-0,448206	0,0860641	-5,208	1,91E-07	***
theta_1	-0,717401	0,138914	-5,164	2,41E-07	***
prezzi_FD	11,2202	5,15261	2,178	0,0294	**

Reale Immaginario Modulo

AR			
Radice 1	1,1208	0,0000	1,1208
AR (stagionale)			
Radice 1	-2,2311	0,0000	2,2311
MA			
Radice 1	1,3939	0,0000	1,3939

In questo modello tutti i parametri risultano significativi al 5%.

Controllo diagnostico.

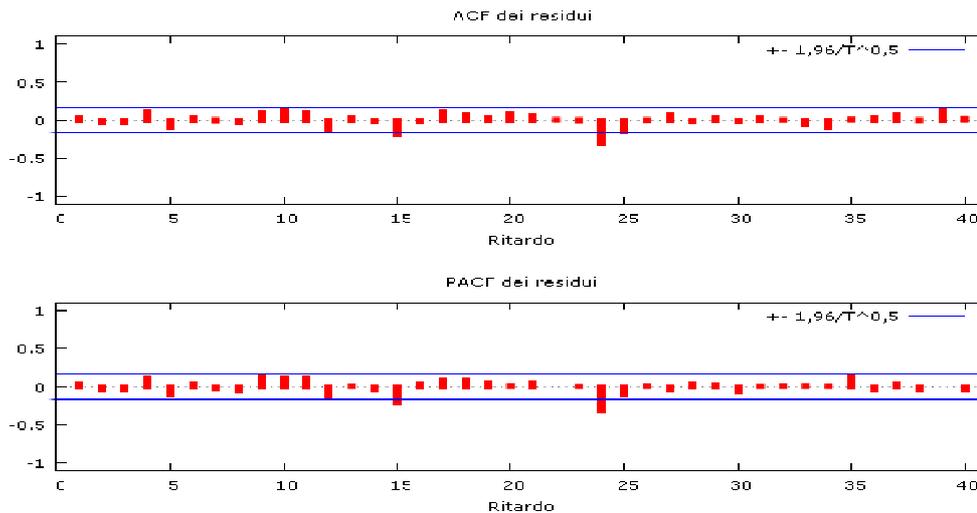


Grafico 3.12. Correlogramma dei residui del modello SARIMAX (1,0,1)(1,1,0) con la variabile prezzi.

I valori della funzione di autocorrelazione sono interni all'intervallo di confidenza, con l'eccezione dei valori al ritardo 15 e 24, ma con un livello di fiducia al 5% si può accettare l'ipotesi nulla di incorrelazione dei residui.

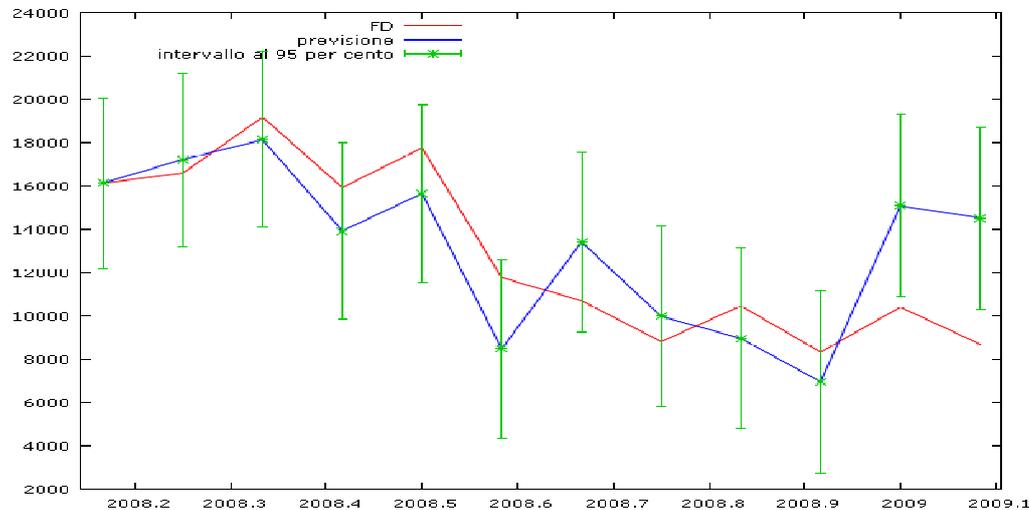


Grafico 3.13. Previsioni-in sample del modello SARIMAX(1,0,1)(1,1,0) con la variabile prezzi.

L'andamento delle previsioni assomiglia con quelle fornite dal modello SARIMA(1,0,1)(1,1,0).

Si prova di modellare la serie storica con un modello SARIMAX che include sia la variabile prezzi che le variabili dummy mensili.

Stima dei parametri del modello SARIMAX (2,0,0)(0,0,0) con le variabili dummy per la modellazione della stagionalità e la variabile prezzi.

	coefficiente	errore std.	rapporto t.	p-value	
phi_1	0,475673	0,0759138	6,266	3,71E-10	***
phi_2	0,402335	0,0799917	5,03	4,91E-07	***
prezzi_FD	-175,052	42,1044	-4,158	3,22E-05	***
dm1	19150,4	2832,18	6,762	1,36E-11	***
dm2	20168,6	2846,21	7,086	1,38E-12	***
dm3	21370,4	2898,55	7,373	1,67E-13	***
dm4	21838,7	2854,47	7,651	2,00E-14	***
dm5	23058,5	2847,59	8,098	5,61E-16	***
dm6	21455,5	2892,77	7,417	1,20E-13	***
dm7	21106,3	2872,07	7,349	2,00E-13	***
dm8	14978,3	2848,88	5,258	1,46E-07	***
dm9	18101,6	2899,71	6,243	4,30E-10	***
dm10	17373,9	2831,64	6,136	8,48E-10	***
dm11	16156,8	2962,24	5,454	4,92E-08	***
dm12	14955,1	2840,99	5,264	1,41E-07	***

Reale Immaginario Modulo

AR

Radice 1 1,0926 0,0000 1,0926

Radice 2 -2,2749 0,0000 2,2749

Tutti parametri sono altamente significativi. In questo modello il segno della variabile prezzi è negativo.

Controllo diagnostico.

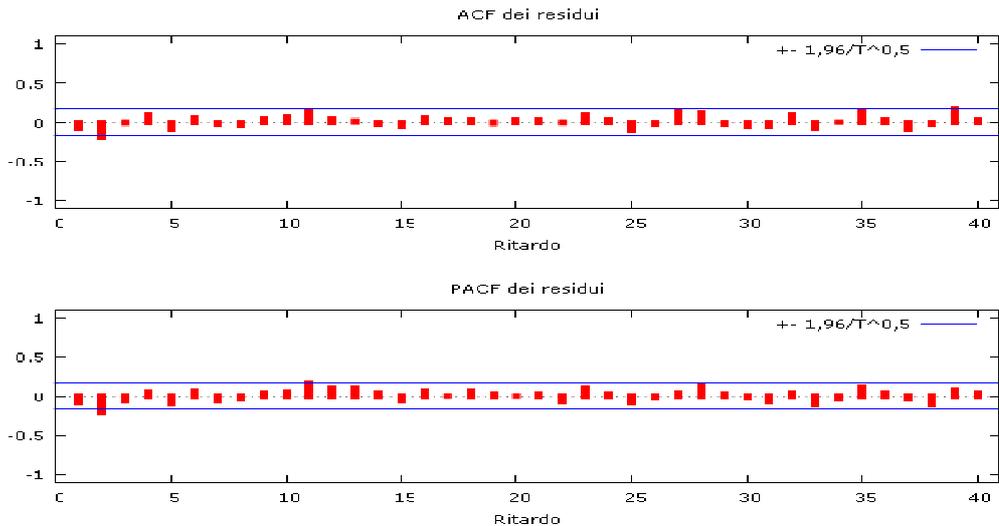


Grafico 3.14. Correlogramma dei residui del modello SARIMAX (2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.

L'ipotesi di incorrelazione dei residui non viene rifiutata; i valori della funzione di autocorrelazione sono compresi all'interno delle bande di confidenza al 5%.

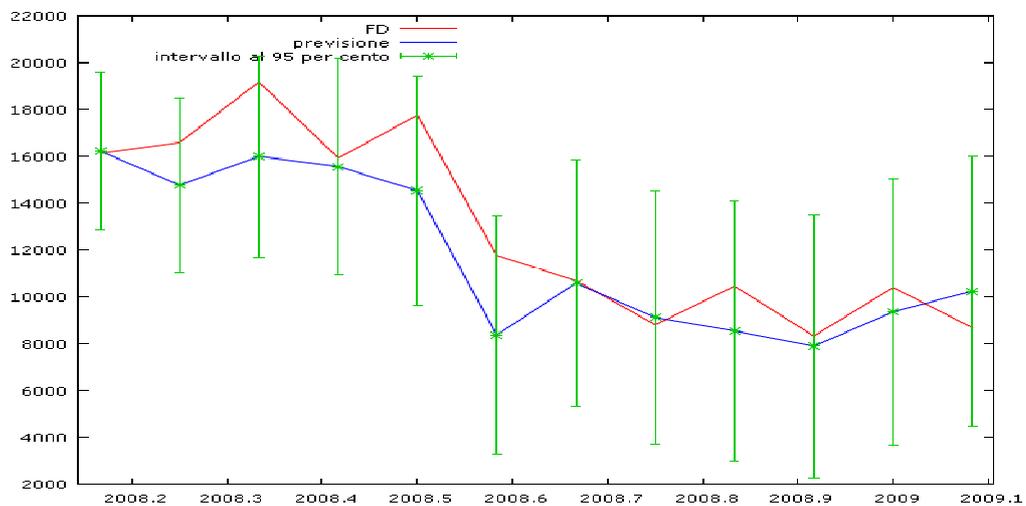


Grafico 3.15. Previsioni-in sample del modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.

Il modello sottostima leggermente i valori veri, ma in generale è in grado di stimare l'andamento della serie di partenza.

Confronto tra i modelli stimati.

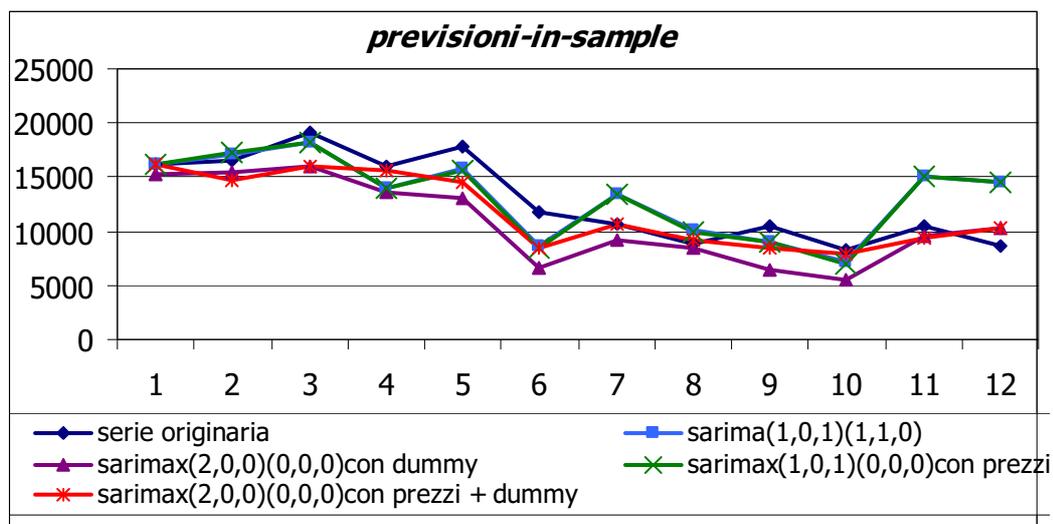


Grafico 3.16. *Previsioni-in sample* dei modelli stimati.

	<i>sarima(1,0)</i> <i>(1,1,0)12</i>	<i>sarimax(2,0,0)</i> <i>(0,0,0)12con</i> <i>dummy</i>	<i>sarimax(1,0,1)</i> <i>(1,1,0)12 con</i> <i>prezzi</i>	<i>sarimax(2,0,0)(0,0,0)</i> <i>12 con</i> <i>dummy+prezzi</i>
	Insieme di previsione			
EQM	2708,25	2852,10	2731,31	1882,08
EMA	2160,4	2396,2	2196,08	1445,30
Indice di Theil	0,20	0,21	0,20	0,14

Tabella 3.3. Indici di bontà di adattamento dei modelli stimati.

Tra i modelli stimati con l'approccio di Box-Jenkins, il modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche presenta gli indici minori, quindi dal punto di vista previsivo può essere considerato il migliore tra i modelli stimati. Si possono confrontare questi risultati con il modello moltiplicativo del metodo di Holt-Winters.

	Modello moltiplicativo	<i>sarimax(2,0,0)(0,0,0)12 con dummy+prezzi</i>
	Insieme di previsione	
EQM	3226,94	1882,08
EMA	2361,9	1445,30
Indice di Theil	0,24	0,14

Tabella 3.4. Confronto tra il modello moltiplicativo e SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.

Gli indici di bontà di adattamento del modello SARIMAX sono decisamente inferiori a quelli del modello moltiplicativo.

Quindi per prevedere le vendite della categoria FD è stato utilizzato il modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.

Previsioni

Anno	Mese	Previsione
2009	Marzo	11.054
	Aprile	10.874
	Maggio	12.187

Tabella 3.5. Previsioni elaborate con il modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche.

Le previsioni sono state aggiornate ogni mese e calcolato l'errore percentuale mensile ($\%errore = \frac{e_t}{r_t} \cdot 100$).

		<i>Previsione</i>	% errore	realizzato	<i>Previsione</i>	% errore	<i>Previsione</i>
2009	3	11.054	-45,49	7.598			
	4	10.874	16,25	12.984	9.579	26,2	
	5	12.187			10.577		12.194
	6				8.971		11.067
	7						10.468

Tabella 3.6. Previsioni del modello SARIMAX(2,0,0)(0,0,0) con la variabile prezzi e le dummy periodiche aggiornate ogni mese.

3.2 Risultati finali

categoria	modello scelto	EMA	Indice di Theil	EMP	% errore marzo	% errore aprile
CC non elett	sarimax(2,0,rit14)(0,0,0) con dummy+prezzi	214,2	0,13	13,60	-83,03	37,2
CD	sarimax(1,0,1)(1,0,0) con dummy+prezzi	7074,01	0,21	17,64	-34,38	-26,15
CP	sarimax(1,0,1)(0,0,0) con dummy	785,8	0,17	15,98	-51,57	3,36
FA non KPS-KPF	sarimax(2,0,0)(1,0,0)12 con dummy	1409,5	0,19	20,45	-42,56	-10,0
FD	sarimax(2,0,0)(0,0,0)12 con dummy+prezzi	1445,30	0,14	11,11	-45,49	26,2
FE	sarimax(1,0,1)(0,0,0)12 con dummy+prezzi	758,9	0,16	16,82	-33,83	-62,1
FS	sarima(1,0,1)(1,1,1)	1014,66	0,33	36,09	23,54	-39,38
FL	sarimax(2,0,0)(0,0,0) con dummy+prezzi	806,4	0,23	21,41	-9,80	-16,6
FF	sarimax(rit6,0,rit6)(2,1,0) con prezzi	164,08	0,25	18,3	-45,03	-119,56
SUB 5"	sarimax(2,0,0)(1,0,0) con dummy+prezzi	344,5	0,20	18,91	-98,7	3,54

Tabella 3.7. Modelli scelti per le categorie di prodotti e gli indici di bontà di adattamento delle previsioni.

Per la categoria CP (categorie che contiene i circolatori piccoli) e FE (categorie che contiene i prodotti per la linea d'acqua fredda) sono state analizzate anche le serie per i singoli codici. Ci sono codici che non presentano correlazione come: K1, K2, P6, quindi le serie storiche di questi prodotti non possono essere modellate con la procedura di Box-Jenkins. Per le serie storiche date sono stati utilizzati solo i metodi di Holt-Winters stagionali.

codice	modello scelto	EMA	Indice di Theil	EMP	% errore marzo	% errore aprile
F3	sarima(1,0,0)(0,1,1)	401,98	0,30	26,2	-70,9	-20,2
F4	sarima(1,0,1)(1,0,0)	19,58	0,24	18,4	-13,9	-48,1
F5	Modello additivo	481,0	0,20	20,2	-63,1	-43,6
F6	sarimax(2,0,1)(0,0,0) dummy+prezzi	316,23	0,56	137,51	-1829,7	62,9
F7	sarimax(1,0,1)(2,0,0) con prezzi	155,03	0,50	36,24	8,4	68,5
F8	sarimax(1,0,1)(2,0,0) con prezzi	13,05	0,30	21,00	14,8	7,7
F9	sarimax(0,0,rit24)(1,1,0) con prezzi	10,37	0,32	29,83	-158,2	-3,3
G0	sarimax(1,0,1)(0,0,0) con dummy	8,99	0,27	26,74	-99,2	9,4
G1	Modello additivo	14,7	0,19	14,0	-63,1	-7,6

Tabella 3.8. Modelli scelti per i codici della categoria CP e gli indici di bontà di adattamento delle previsioni.

codice	modello scelto	EMA	Indici di Theil	EMP	% errore marzo	% errore aprile
J7	Modello moltiplicativo	39,01	0,15	14,5	6,5	-128,7
J8	sarimax(0,0,0)(2,0,0) con prezzi	19,46	0,41	24,02	18,6	-23,1
J9	sarima(0,0,0)(2,0,0)	19,49	0,21	18,6	13,5	-274,0
P6	Modello additivo	35,0	0,24	43,2	-4,8	-132,9
P7	sarimax(rit4,0,0)(1,0,0) con prezzi	235,6	0,53	40,7	0,7	-226,9
P8	Modello moltiplicativo	14,02	0,50	221,7	-63,7	-10,0
P9	sarimax(rit27,0,0)(1,0,0)	50,64	0,54	94,63	81,1	-35,0
K0	sarima(0,0,0)(1,1,0)	51,10	0,39	47,4	-1,4	-581,2
K1	modello moltiplicativo	73,65	0,46	42,8	-138,5	-353,1
K2	Modello additivo	92,7	0,44	82,7	7,2	-202,1
K3	Modello additivo	184,7	0,26	24,9	10,33	13,74
K4	sarimax(1,0,0)(2,0,0) ₁₂ con prezzi	40,7	0,36	34,31	-1,89	62,4
K5	sarima(1,0,1)(1,0,0) ₁₂ con dummy	51,53	0,33	34,2	-27,32	-27,7
K6	sarimax(1,0,1)(2,0,0) ₁₂ con prezzi	156,34	0,45	32,03	-48,50	-240,19
K7	sarimax(1,0,1)(0,0,0) ₁₂ con dummy e prezzi	52,91	0,26	24,71	-26,61	18,8
K8	Modello additivo	17,7	0,50	37,71	36,01	-474,59
K9	sarima(rit24,0,0)(1,1,0)	74,6	0,42	61,84	-3,61	-19,12
L0	Modello additivo	21,8	0,30	41,07	25,2	21,18

Tabella 3.9. Modelli scelti per i codici della categoria FE e gli indici di bontà di adattamento delle previsioni.

3.4 Valutazione economica delle previsioni.

Una volta elaborate le previsioni di vendita per le varie categorie di prodotti, l'azienda era interessata nella valutazione economica delle previsioni, cioè di stimare i costi e la superficie di stoccaggio dei prodotti sovrastimati e il mancato guadagno dei prodotti sottostimati.

Le informazioni riguardanti l'aspetto economico sono state riepilogate nelle tabelle 3.9, 3.10 e 3.11.

Categoria	Somma degli errori nel periodo 05.08-04.09	Costi di produzione	Numero di palette ¹¹⁰	Superficie necessaria (m ²)	Costi di stoccaggio	Mancato guadagno
CC non elett	1.556					211.530
CD	-7.479	73.384	43	41	82	
CP	-2.144	60.034	42	41	81	
FA non KPS-KPF	-12.090	301.452	448	430	860	
FD	-1.483	55.032	68	66	131	
FE						
FS	6.817					345.183
FL	-27.625	1.561.060	2.029	1.948	3.895	
FF	-1.158	125.160	71	68	136	
SUB 5"	1.531					164.003
TOTALE	-42.075	2.176.121	2.701	2.593	5.185	720.717

Tabella 3.10. Valutazione economica delle previsioni elaborate per le categorie di prodotti modellati.

Codice	Somma degli errori nel periodo 05.08-04.09	Costi di produzione	Numero di palette	Superficie necessaria (m ²)	Costi di stoccaggio	Mancato guadagno
F3	-2.492	40.412	22	21	42	
F4	-133	1.234	1	1	2	
F5	-5.988	106.297	58	56	112	
F6	781					26.787
F7	1.818					76.113
F8	23					946
F9	35					2.236
G0	-77	5.059	3	3	5	
G1	14					1.235
TOTALE	-6.019	153.002	84	80	161	107.317

Tabella 3.11. Valutazione economica delle previsioni elaborate per i codici della categoria CP.

¹⁰ Le palette sono delle piattaforme d'appoggio sopraelevate mediante dei piedini alla base e fornite di apposite feritoie per poter essere sollevate con i carrelli elevatori.

codice	Somma degli errori nel periodo 05.08-04.09	Costi di produzione	Numero di palette	Superficie neccessaria (m²)	Costi di stoccaggio	Mancato guadagno
J7	-171	12.835	10	9	18	
J8	17					1.086
J9	-61	4.126	3	3	6	
P6	-123	7.091	6	6	11	
P7	2.319					184.977
P8	-41	2.419	2	2	4	
P9	487					32.334
K0	25					1.918
K1	-225	10.445	11	10	21	
K2	363					14.016
K3	-658	46.739	37	35	70	
K4	449					46.408
K5	-123	8.234	7	7	13	
K6	-325	28.795	22	21	42	
K7	-108	7.219	8	7	14	
K8	-226	11.066	13	12	24	
K9	-690	59.047	50	48	96	
L0	5					401
TOTALE	914	198.017	166	159	319	281.139

Tabella 3.12. Valutazione economica delle previsioni elaborate per i codici della categoria FE.

I costi di produzione e il numero di palette costituiscono dei valori medi per categoria di prodotti. La superficie neccessaria è stata calcolata moltiplicando il numero delle palete per la superficie che occupa una paleta (1,20 m * 0,8 m = 0,96 m²). Anche i costi di stoccaggio sono stati stimati da un valore medio. Il mancato guadagno è stato stimato come un valore aggiunto del 35% al costo di produzione.

Conclusioni

La modellazione delle serie storiche delle vendite di elettropompe ha consentito di stimare alcuni modelli per le varie categorie di prodotti e successivamente elaborare delle previsioni.

Per le serie storiche modellate si è verificato che i modelli SARIMA e SARIMAX si adattano meglio ai dati e prevedono in modo più accurato. A differenza dei metodi di Holt-Winters, i modelli SARIMA riescano a modellare la natura stocastica della stagionalità. Inoltre, la variabile esogena (i prezzi medi dei prodotti), utilizzata nei modelli SARIMAX, risulta significativa per quasi tutte le serie; quindi essa ha contribuito alla spiegazione della variabilità del modello.

Per quanto riguarda la modellazione delle serie storiche dei singoli codici dei prodotti, si osserva un aumento dell'indice di Theil, ciò significa che i modelli stimati per i singoli codici si adattano peggio ai valori osservati rispetto ai modelli stimati per l'intera categoria. In ogni caso, date le circostanze attuali, per la gran parte dei prodotti, le previsioni delle vendite risultano sistematicamente superiori ai valori reali a causa della forte controtendenza delle serie storiche rispetto alla storia passata. Gli errori si possono parzialmente ridurre mediante l'aggiornamento delle serie.

Dall'analisi della parte riguardante la valutazione economica delle previsioni si può concludere che le previsioni fornite per le categorie di prodotti possono essere utilizzate come linea guida nella pianificazione semestrale della produzione.

Bibliografia

- Box, G. E., Jenkins, G. M. (1976): *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.
- Di Fonzo, T. e Lisi, F. (2005): *Serie storiche economiche*, Carocci, Roma.
- Guizzardi, A. (2001): *La previsione economica: tecniche statistiche e strategie di valutazione*, Guaraldi.
- Ljung, G. M., Box, G. E. (1978): *On a measure of a lack of fit in time series models*, "Biometrika".
- Mark, Purvis (2008): *Guide to the Italian Pump Industry*, November 7.
- Theil, H. (1965): *Economic forecast and policy*, North-Holland, Amsterdam.
- Vianelli, S. (1983): *L'analisi delle serie temporali nello sviluppo storico e metodologico della statistica*, Milano
- Wei, W. W. S. (2006): *Time series analysis: univariate and multivariate methods*, Pearson Education.

Pacchetti statistici

Gretl open source, URL http://gretl.sourceforge.net/gretl_italiano.html