

UNIVERSITA' DEGLI STUDI DI PADOVA

FACOLTA' DI SCIENZE STATISTICHE

CORSO DI LAUREA IN STATISTICA E GESTIONE DELLE IMPRESE

RELAZIONE FINALE:

ANALISI E PREVISIONI DELLE SERIE DI TRAFFICO VOCE DI UN'AZIENDA
DI TELEFONIA MOBILE

RELATORE: PROF. LUISA BISAGLIA

LAUREANDO: GIACOMO CASSOL

ANNO ACCADEMICO: 2005-2006

INDICE

CAPITOLO 1 – PRESENTAZIONE

- 1.1 INTRODUZIONE
- 1.2 ANALISI DEL SETTORE DELLA TELEFONIA MOBILE PER L'ANNO 2002
- 1.3 BREVE STORIA DELL'AZIENDA H3G

CAPITOLO 2 – METODOLOGIE UTILIZZATE

- 2.1 PREMESSA
- 2.2 PROCESSI STOCASTICI
 - 2.2.1 MODELLI SARIMA
- 2.3 METODO DI HOLT-WINTERS

CAPITOLO 3 – ANALISI DELLE SERIE STORICHE CON I MODELLI SARIMA

- 3.1 SERIE DEI MINUTI VOCE ABBONAMENTO
- 3.2 PREVISIONI DEL TRAFFICO VOCE ABBONAMENTO
- 3.3 SERIE DEI MINUTI VOCE RICARICABILE
- 3.4 PREVISIONI DEL TRAFFICO VOCE RICARICABILE

CAPITOLO 4 – ANALISI DELLE SERIE STORICHE CON IL METODO DI HOLT-WINTERS

- 4.1 PREVISIONI DEL TRAFFICO ABBONAMENTI E DEL TRAFFICO RICARICABILE
- 4.2 CONFRONTO FRA LE DUE METODOLOGIE UTILIZZATE

CAPITOLO 5 – ANALISI DESCRITTIVA DELLE SERIE COMPONENTI IL TRAFFICO VOCE

- 5.1 SERIE ABBONATI
- 5.2 SERIE RICARICABILI

CAPITOLO 6 – CONCLUSIONI

- 6.1 CONSIDERAZIONI FINALI

1. PRESENTAZIONE

1.1 INTRODUZIONE

In questo lavoro si analizzeranno i dati relativi ai minuti di traffico voce di una nota azienda di telefonia mobile mediante lo strumento delle serie storiche. In particolare, si considereranno due serie classificate a seconda che il consumatore possieda una carta ricaricabile o un contratto di abbonamento.

I due casi verranno studiati separatamente, per il periodo di un anno, con lo scopo di verificare se i due diversi tipi di contratto influenzano il comportamento dei consumatori.

I metodi utilizzati per la modellazione dell'andamento delle due serie storiche, sono l'approccio Box-Jenkins, e quindi i processi ARIMA stagionali, e il metodo di Holt Winters stagionale.

Infine, dopo aver verificato la bontà di adattamento di questi modelli alle serie prese in esame, sono state effettuate delle previsioni per un periodo di quattro settimane, allo scopo di quantificare l'accuratezza in termini previsivi dei modelli selezionati.

1.2 ANALISI DEL SETTORE DELLA TELEFONIA MOBILE PER L'ANNO 2002

Fino al 2002 il settore della telefonia mobile presentava un oligopolio composto da 3 compagnie: TIM, Vodafone e Wind. L'accesa concorrenza fra questi operatori ha provocato una continua riduzione dei prezzi per cellulari, SMS e chiamate, compensata necessariamente da ingenti investimenti. Tali spese dovevano contribuire alla realizzazione di guadagni nel medio-lungo periodo e invece hanno provocato soltanto un alto indebitamento e una perdita di fiducia da parte degli investitori. Per questo si è cercato, attraverso fusioni e condivisione degli investimenti strutturali (es. rete), di abbassare i costi. Anche gli organici sono stati ridimensionati, tuttavia la riduzione dei costi operativi non è stato un mezzo sufficiente a contribuire alla crescita di settore. Gli operatori hanno dovuto puntare alla massimizzazione del ricavo sulla base clienti esistente. Si è cercato di aumentare il ricavo medio per utente (ARPU: Average Revenue Per User) anche se gli analisti di mercato prevedevano una graduale riduzione di questo indice.

L'unica strada per conseguire un vantaggio competitivo sui concorrenti era di offrire i servizi standard (SMS, chiamate) a prezzi sempre più convenienti.

1.3 BREVE STORIA DELL'AZIENDA

La società di telefonia mobile H3G, facente parte di una multinazionale con sede ad Hong Kong, nasce con l'intento di cercare un vantaggio competitivo tramite la differenziazione del prodotto.

3 Italia (questo è il nome originario, risalente al 2002) è il primo operatore di telefonia mobile in Italia ad offrire i servizi di terza generazione. La sigla UMTS sta per “*Universal Mobile Telecommunications System*”, ovvero lo standard di comunicazione *wireless* a banda larga. Grazie al trasferimento di dati ad alta velocità, l'UMTS permette di fare video telefonate, video conferenze, ascoltare musica, guardare la TV ecc.

Questi servizi finora mai offerti da nessun'altra compagnia telefonica sono frutto della continua ricerca e innovazione tecnologica, per riuscire ad offrire soluzioni sempre all'avanguardia. 3 Italia vuol essere un passo avanti rispetto al passato, come è successo per Internet, i messaggi SMS, la posta elettronica, anche i servizi 3 sono destinati a diventare uno standard.

I terminali offerti dalla compagnia sono *dual mode*, ovvero dei telefonini che funzionano sia sulla rete UMTS che sulla tradizionale linea di seconda generazione GSM/DCS. Questa tipologia di cellulare permette all'azienda di offrire i servizi GSM nelle aree non ancora coperte dall'UMTS. Proprio per questo motivo è stato stipulato durante i primi anni di attività un accordo di collaborazione con TIM, l'operatore che possiede la rete più capillare in Italia.

Secondo il *management* dell'azienda, l'elemento chiave nel futuro mercato delle comunicazioni è costituito dalla convergenza fra telecomunicazioni, internet e media in un unico terminale multimediale mobile. Questo strumento diventerà indispensabile nella vita di tutti i giorni e sarà oggetto di continui aggiornamenti; la società punta a diventare leader nell'innovazione tecnologica, rivolgendosi ad un mercato in continua crescita. È quest'ultimo fattore, che ha permesso a 3 di entrare in un settore già saturo come quello della telefonia. Essendo un nuovo entrante, ha dovuto investire pesantemente oltre che in infrastrutture, anche in pubblicità e promozione. Per farsi conoscere ha scelto un logo semplice, in modo

da essere riconosciuto in tutto il mondo. Il numero 3 esiste in ognuna delle seimila lingue e dei quattrocento paesi del pianeta.

Lo scopo principale dell'azienda è creare una solida base clienti, attraverso promozioni molto convenienti. Fra queste, c'è da ricordare l'offerta di cellulari a basso costo con la *USIM LOCK*, un "blocco" che impedisce ai possessori del cellulare di terza generazione, di utilizzarlo con altre schede SIM. Questa limitazione ha una durata di 9 mesi dall'acquisto, dopodiché è possibile sbloccare il telefono pagando una somma non superiore al 50% del valore del contributo economico fornito dall'operatore all'atto dell'acquisto. Tale "forzatura" ha permesso di costruire una esigua base clienti, in modo da poter concorrere con le compagnie già affermate. Esse non hanno potuto difendersi dalla nuova entrante a causa della nuova tecnologia offerta.

A partire dal 2002 fino ad oggi i fruitori dei telefonini di terza generazione hanno avuto una crescita esponenziale. Partendo da poche migliaia di clienti, si è passati in pochi anni a quasi 5 milioni (dati aggiornati al 31 dicembre 2005), e il trend è ancora oggi crescente.

2. METODOLOGIE UTILIZZATE

2.1 PREMESSA

Il periodo preso in esame va dal primo luglio del 2004 fino al 31 agosto 2005. Per la stima del modello si utilizzano i dati fino al 31 luglio 2005, mentre l'ultimo mese viene considerato per poter effettuare dei confronti a fini previsivi. In un mercato così difficile come quello della telefonia mobile, saper prevedere "correttamente" è fondamentale, ma devono essere utilizzati i dovuti accorgimenti.

Dopo la fase di pianificazione (allocazione delle risorse e investimenti), si passa alla stima dei possibili guadagni futuri. È in questo passaggio che la previsione permette al *management* aziendale di costruire degli scenari prospettici.

A intervalli di tempo prestabiliti, i calcoli vengono aggiornati e si quantificano gli errori di previsione. Nel seguito questa metodologia verrà attuata nel caso in cui la serie presenti una improvvisa crescita o un forte calo dei minuti di traffico attesi. Questi, infatti, rappresentano indirettamente il fatturato della società, anche se non sono espressi in termini monetari. Il loro andamento complessivo, dipende, oltre che dalla variazione della spesa del singolo, anche dal numero totale degli utenti.

In questo lavoro i dati verranno prima analizzati con i modelli ARIMA stagionali e poi con il metodo di Holt-Winters stagionale, per poter valutare quale dei due fornisce le previsioni migliori.

2.2 PROCESSI STOCASTICI

L'analisi statistica di una serie storica si propone di chiarire il meccanismo causale che l'ha generata. Indicando con y_t la realizzazione di un fenomeno al tempo t , si può scrivere:

$$y_t = f(t) + \varepsilon_t$$

dove $f(t)$ rappresenta la parte deterministica della serie, mentre ε_t quella stocastica o casuale.

$f(t)$ è a sua volta composta da tre elementi: il trend, il ciclo e la stagionalità. Lo studio di queste componenti costituisce l'approccio classico all'analisi delle serie storiche. Tuttavia, quando si studiano serie reali, quali ad esempio l'andamento del prodotto interno lordo di un paese, o la quotazione dei titoli in borsa, risulta difficile individuare con chiarezza una sequenza deterministica attendibile per spiegare l'oggetto di studio. Per questo non si ricorre a formule matematiche complesse, bensì si cerca di studiare un modello sulla base della sua storia passata. Ciò significa che è meglio concentrarsi sulla parte stocastica, ε_t , ipotizzando che il meccanismo che genera i dati sia governato da regole probabilistiche. Per questo motivo ε_t viene considerato come un processo a componenti correlate, per il quale cioè $Cov(\varepsilon_t, \varepsilon_s) \neq 0$ per qualche $t \neq s$. Da quanto appena detto, deriva anche che l'obiettivo non è più quello di arrivare ad una stima delle componenti di una serie, ma piuttosto quello di individuare un modello probabilistico che descriva l'evoluzione del fenomeno in esame, modello che può essere utilizzato sia a fini descrittivi che previsivi.

2.2.1 MODELLI SARIMA

I modelli più comunemente usati nell'analisi di serie storiche sono gli *ARIMA* stagionali o *SARIMA* $(p, d, q)x(P, D, Q)_s$ (Box, Jenkins, Reinsel, 1994).

Questi modelli cercano di spiegare l'andamento di una serie storica sulla base della sua storia passata.

L'idea che sta alla base di tale metodo è la descrizione del fenomeno attraverso l'adattamento sia della parte stagionale, che di quella non stagionale. Ciò comporta che le serie studiate siano caratterizzate da una forte correlazione seriale ai ritardi stagionali (settimana, mese, anno), oltre all'usuale autocorrelazione di breve periodo. La parte stagionale non viene più considerata deterministica e indipendente dalle altre componenti.

La procedura che permette di ottenere un modello *ARIMA* che rappresenti adeguatamente il processo generatore dei dati, è stata proposta da Box e Jenkins e consta di tre fasi fondamentali: identificazione, stima dei parametri, controllo diagnostico.

Identificazione

Dopo aver reso stazionaria la serie attraverso opportune trasformazioni o differenziazioni, si identifica l'ordine dei parametri a memoria corta (p,d,q) e stagionali (P,D,Q) . Ciò viene effettuato principalmente mediante analisi successive dei grafici di autocorrelazione globale e parziale.

Stima dei parametri

Una volta fissato l'ordine dei parametri, si passa alla fase di stima degli stessi, mediante la massimizzazione della funzione di verosimiglianza.

Controllo diagnostico

Successivi accertamenti verranno effettuati per valutare la bontà del modello, tra i quali l'analisi dei residui e alcuni indici che misurano l'adattamento e la capacità previsiva del modello. Fra questi i criteri di Akaike e Schwarz che assegnano un "costo" all'introduzione di ogni parametro addizionale, oltre a tenere in considerazione la varianza residua.

2.3 METODO DI HOLT-WINTERS

La tecnica di *Holt-Winters* è una evoluzione del metodo del liscio esponenziale. Quest'ultimo prevede l'assegnazione di pesi via via decrescenti man mano che le osservazioni si allontanano nel passato.

Tali metodi risultano utili quando la serie non varia occasionalmente, ma presenta un marcato andamento di fondo (trend) ed una eventuale componente stagionale.

A differenza dei SARIMA, questi modelli hanno una struttura più rigida e vanno aggiornati di volta in volta per ottenere delle previsioni attendibili. Sono abbastanza semplici da utilizzare, in quanto considerano la serie come il risultato di tre componenti: la stima del livello al tempo n (\hat{y}_n), il trend (\hat{T}_n) e la stagionalità (\hat{S}_n).

METODO STAGIONALE ADD.

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

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

METODO STAGIONALE MOLT.

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

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

Le formule di aggiornamento per ottenere \hat{y} , \hat{T} e \hat{S} sono le seguenti:

METODO STAGIONALE ADD.

$$\hat{y} = \alpha \cdot (\hat{y}_{n-1} + T_{n-1}) + (1 - \alpha) \cdot (y_n - S_{n-s})$$

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

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

METODO STAGIONALE MOLT.

$$\hat{y} = \alpha \cdot (\hat{y}_{n-1} + T_{n-1}) + (1 - \alpha) \cdot \frac{y_n}{\hat{S}_{n-s}}$$

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

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

La stima del livello al tempo t tiene conto del trend del periodo precedente, aggiungendolo all'ultimo valore liscio (\hat{y}_{n-1}).

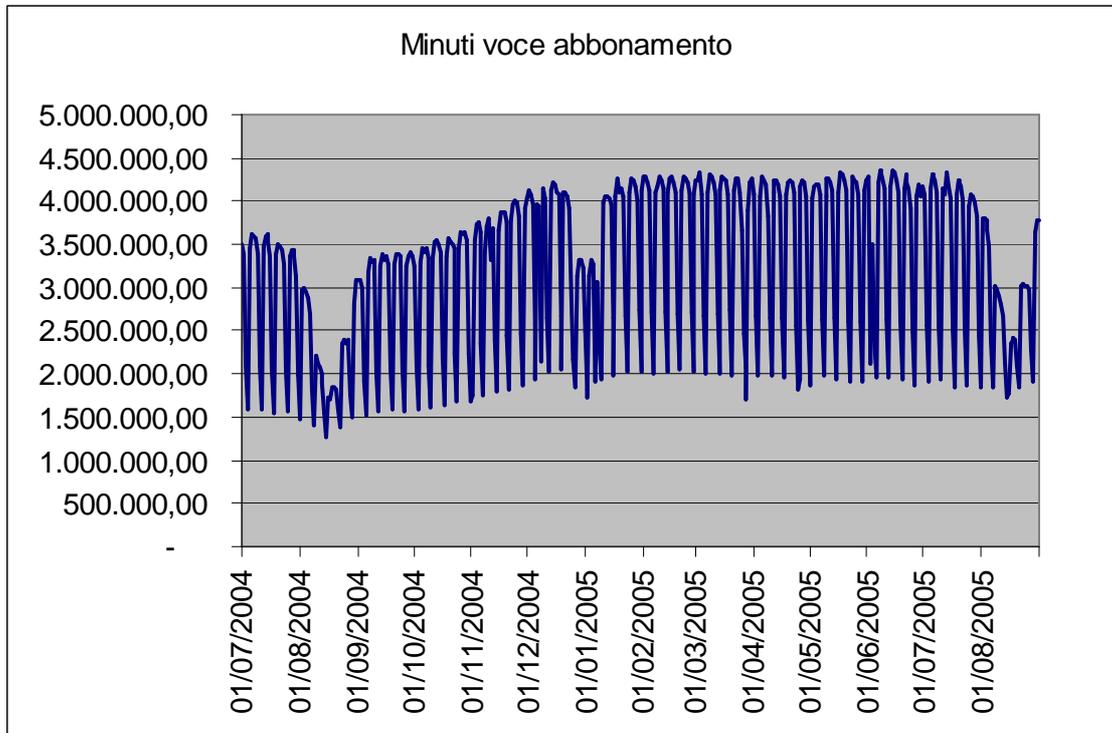
Il trend è la media pesata della differenza fra i due livelli più l'ultimo valore stimato del trend. In questo modo si ottiene un trend liscio.

La stagionalità si ottiene come differenza tra il valore vero della serie (y_n) e il livello stimato al tempo n (\hat{y}), più la stagionalità del periodo precedente (\hat{S}_{n-s}).

Tutte queste misure vengono perequate da 3 coefficienti (α , β , γ) che permettono di minimizzare gli indici per la valutazione delle previsioni (EQM, EMA, ecc.).

3. ANALISI DELLE SERIE STORICHE CON I MODELLI SARIMA

3.1 SERIE DEI MINUTI VOCE ABBONAMENTO



Analisi preliminare della serie

Partendo dal traffico voce degli abbonati, escluse le video chiamate, si nota immediatamente la stagionalità settimanale. In corrispondenza dei fine settimana i minuti totali di chiamate diminuiscono significativamente. Tale fenomeno è probabilmente dovuto al fatto che buona parte dei contratti di abbonamento vengono stipulati con le aziende, è quindi logico aspettarsi un calo di telefonate durante i giorni festivi. Per lo stesso motivo in corrispondenza delle feste natalizie e delle vacanze estive c'è un vistoso calo del traffico voce.

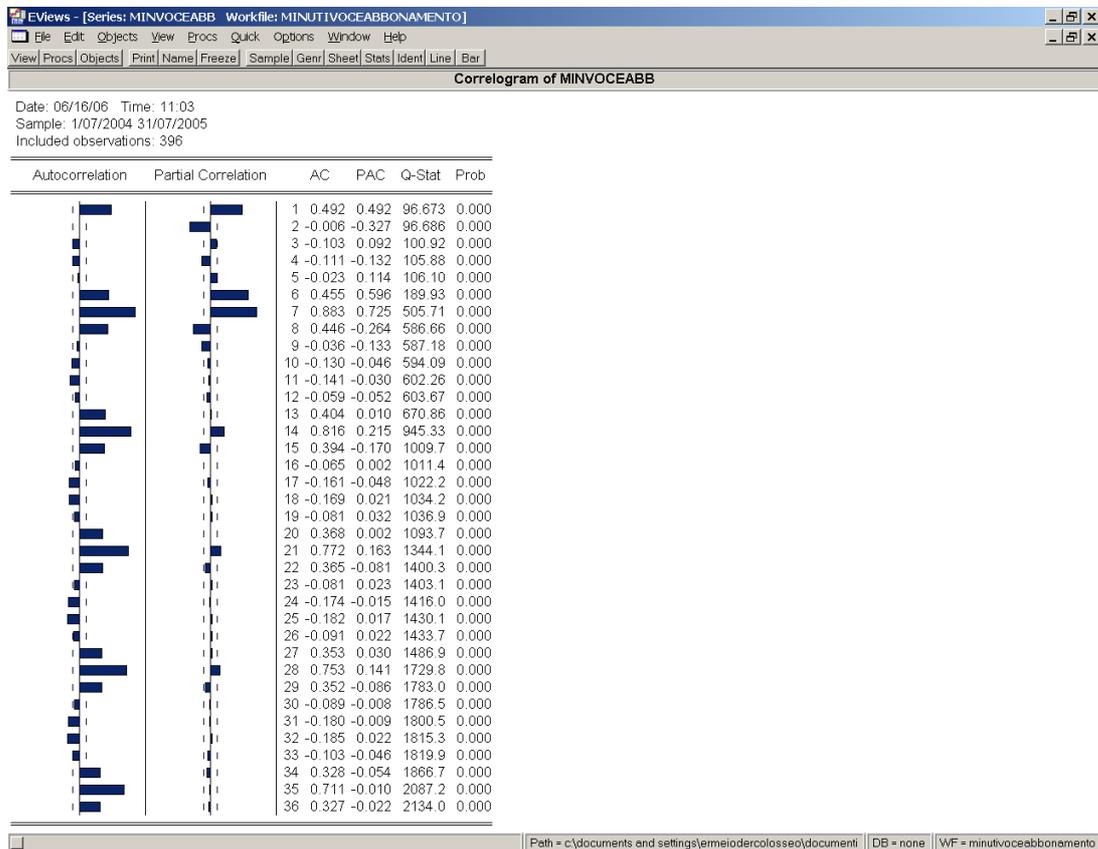
L'andamento di fondo è leggermente crescente, ciò sta a significare un continuo aumento della base clienti. Al momento del lancio dei telefonini di terza generazione la crescita era quasi esponenziale, ora si è di fronte a una stabilizzazione delle vendite e del traffico.

PROCEDURA DI BOX-JENKINS

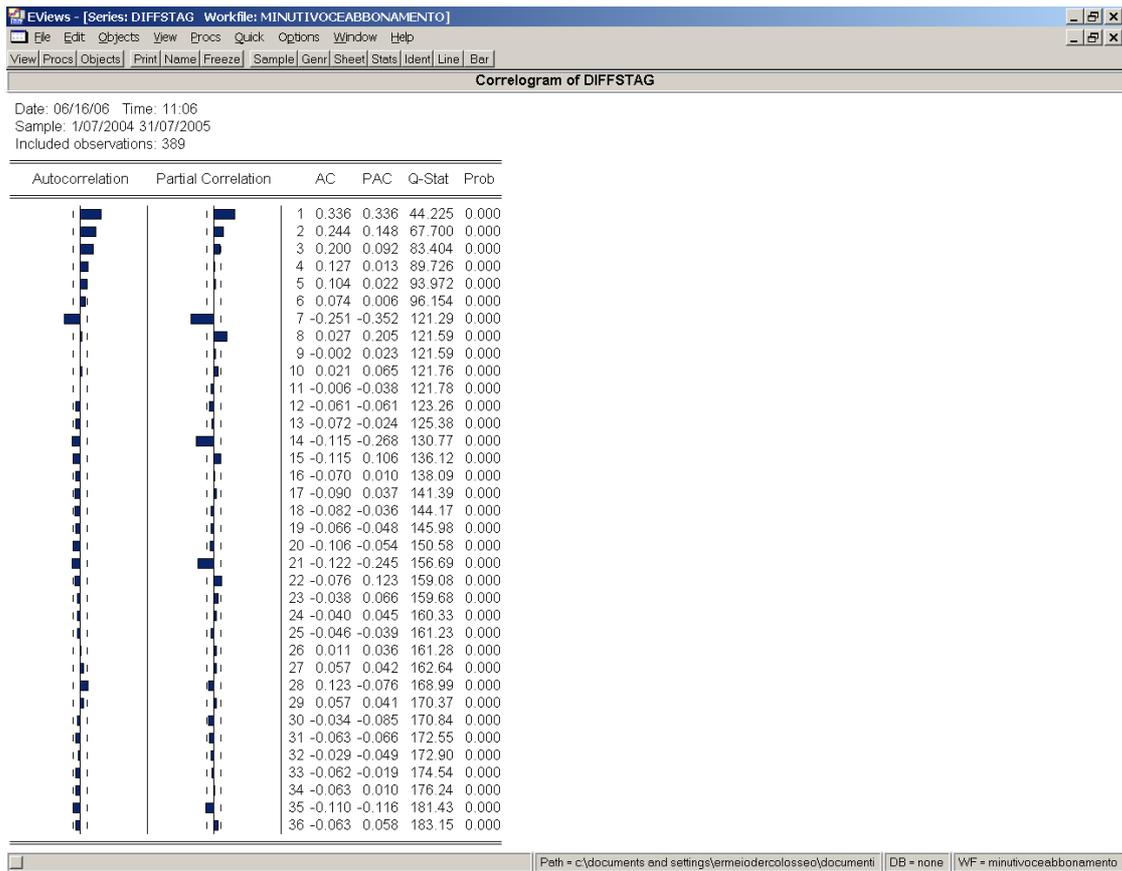
Identificazione

Nella prima fase si cerca di identificare l'ordine dei parametri p, d, q del modello. Lo strumento principale per raggiungere questo scopo è costituito dai grafici delle funzioni di autocorrelazione globale (ACF) e autocorrelazione parziale (PACF) della serie.

I grafici seguenti riportano le funzioni di autocorrelazione empiriche del processo.



Come si può osservare, l'ACF ai ritardi stagionali tende a zero molto lentamente, il che suggerisce una differenziazione di tipo stagionale. I correlogrammi della serie differenziata stagionalmente sono riportati nel grafico che segue.



Dall'osservazione degli autocorrelogrammi sembra che la serie sia stazionaria. Si cerca ora di identificare l'ordine del modello che potrebbe aver generato i dati.

Stima dei parametri e controllo diagnostico

Per quanto riguarda la componente stagionale, l'ACF ha un unico ritardo significativo per $k=7$ mentre la PACF si azzera lentamente fino al ritardo 35. Questo comportamento è compatibile con una componente a media mobile stagionale. Un punto di partenza è quindi il modello SARIMA(0,0,0)x(0,1,1).

La seguente tabella, ottenuta tramite E-Views, riporta i risultati della stima.

Dependent Variable: D(MINVOCEABB,0,7)

Method: Least Squares

Date: 06/16/06 Time: 11:11

Sample(adjusted): 8/07/2004 31/07/2005

Included observations: 389 after adjusting endpoints

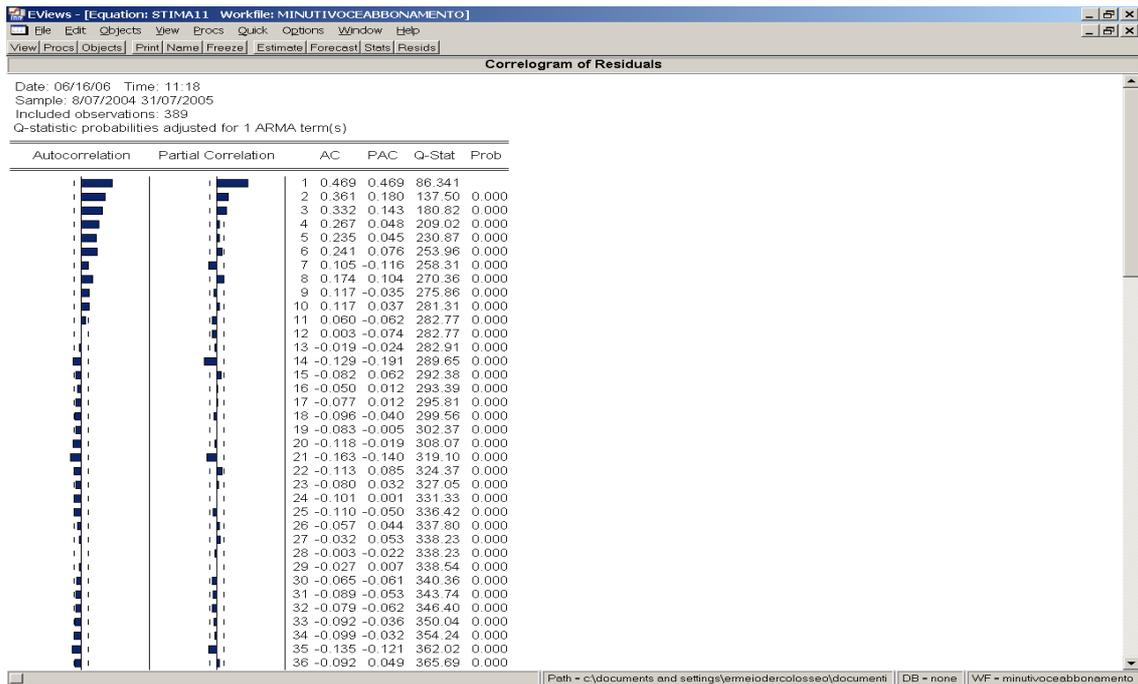
Convergence achieved after 13 iterations

Backcast: 1/07/2004 7/07/2004

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(7)	-0.472612	0.044682	-10.57715	0.0000
R-squared	0.108554	Mean dependent var		7586.909
Adjusted R-squared	0.108554	S.D. dependent var		426487.6
S.E. of regression	402674.3	Akaike info criterion		28.65221
Sum squared resid	6.29E+13	Schwarz criterion		28.66240
Log likelihood	-5571.855	Durbin-Watson stat		1.059336

Come si può osservare, il coefficiente a media mobile stagionale risulta significativamente diverso da zero.

Dall'analisi dei residui, in particolare dall'analisi delle funzioni di autocorrelazione, appare evidente che è necessario aggiungere altre componenti di breve periodo.



L'autocorrelazione globale tende a zero lentamente, quella parziale ha soltanto i primi due/tre ritardi significativi; due modelli alternativi possono pertanto essere il SARIMA (1,0,1)x(0,1,1) e il SARIMA (2,0,0)x(0,1,1). Di seguito vengono confrontati questi due modelli in base ai criteri di *Akaike e Schwarz*:

MODELLO SARIMA (1,0,1)x(0,1,1)

Akaike info criterion= 28,233

Schwarz criterion= 28,263

MODELLO SARIMA (2,0,0)x(0,1,1)

Akaike info criterion= 28,283

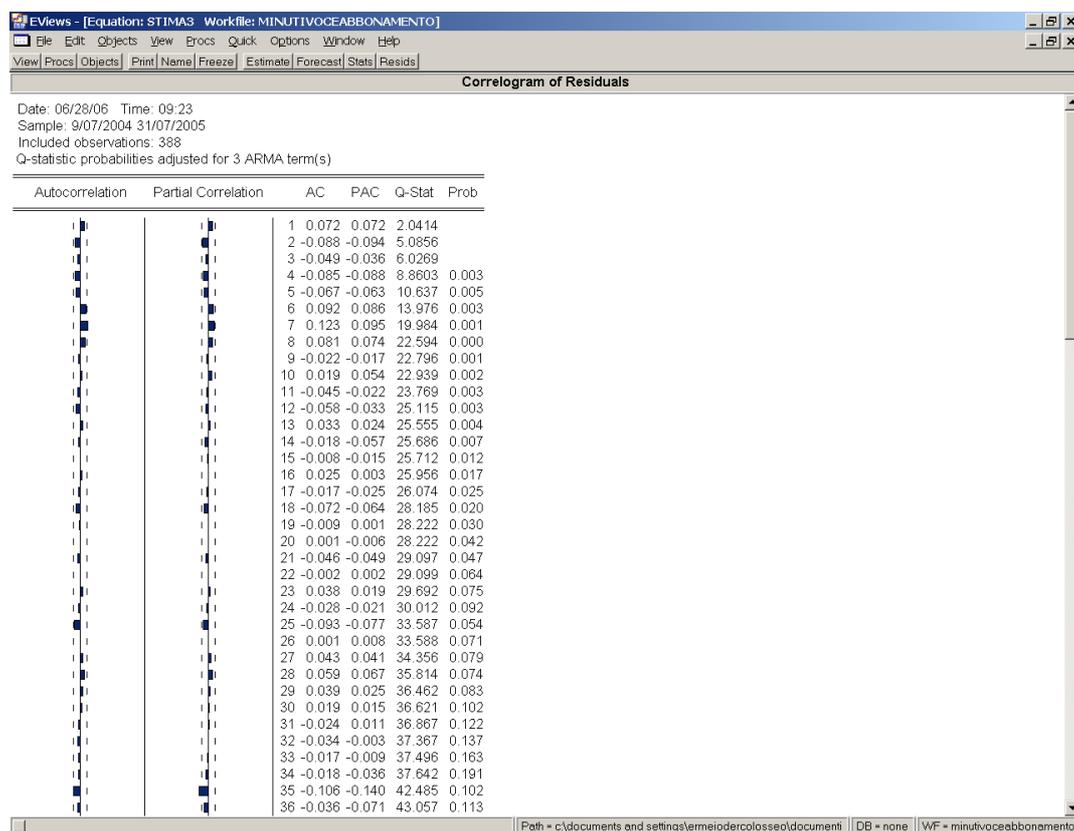
Schwarz criterion= 28,313

Non c'è una differenza troppo alta fra i due valori, ma il SARIMA (1,0,1)x(0,1,1) sembrerebbe essere quello che meglio rappresenta la serie e quindi può essere utilizzato per effettuare le previsioni del mese di agosto.

Di seguito vengono riportati l'*output* di stima del modello scelto e i grafici di autocorrelazione dei residui ad esso relativi.

Dependent Variable: D(MINVOCEABB,0,7)
 Method: Least Squares
 Date: 05/25/02 Time: 17:15
 Sample(adjusted): 9/07/2004 31/07/2005
 Included observations: 388 after adjusting endpoints
 Convergence achieved after 11 iterations
 Backcast: 1/07/2004 8/07/2004

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.929520	0.026134	35.56809	0.0000
MA(1)	-0.596109	0.055329	-10.77397	0.0000
SMA(7)	-0.886521	0.024904	-35.59724	0.0000
R-squared	0.421221	Mean dependent var		7422.929
Adjusted R-squared	0.418215	S.D. dependent var		427026.0
S.E. of regression	325713.4	Akaike info criterion		28.23313
Sum squared resid	4.08E+13	Schwarz criterion		28.26375
Log likelihood	-5474.226	Durbin-Watson stat		1.847739



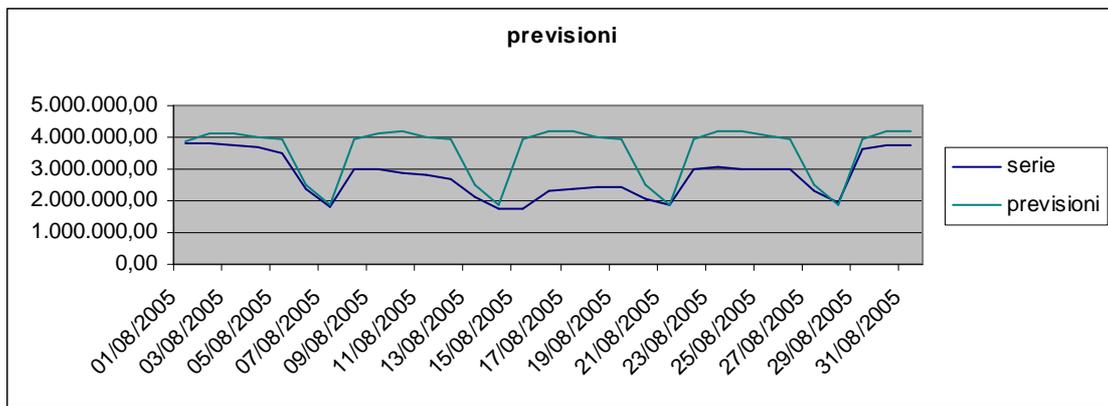
I coefficienti del SARIMA (1,0,1)x(0,1,1) risultano significativi, come dimostra la statistica t di Student; inoltre dai grafici di autocorrelazione dei residui si accetta l'ipotesi nulla di incorrelazione degli stessi.

3.2 PREVISIONI DEL TRAFFICO VOCE ABBONAMENTO

Dal grafico precedentemente prodotto con Excel si evince che il mese di agosto presenta non poche difficoltà per ottenere delle previsioni attendibili. Avendo a disposizione soltanto i dati riferiti ad un anno, non è possibile trattare la stagionalità di tipo annuale, ma solo quella dovuta ai giorni della settimana.

Osservando il grafico, è evidente il calo delle telefonate nei mesi di luglio e agosto. Per avere una spiegazione del fenomeno, sarebbe necessario disporre dei dati relativi agli anni precedenti il 2004.

Senza considerare questi problemi, le previsioni col modello SARIMA (1,0,1)x(0,1,1) danno i seguenti risultati.



EM= -770.576

EQM= 980.642,73

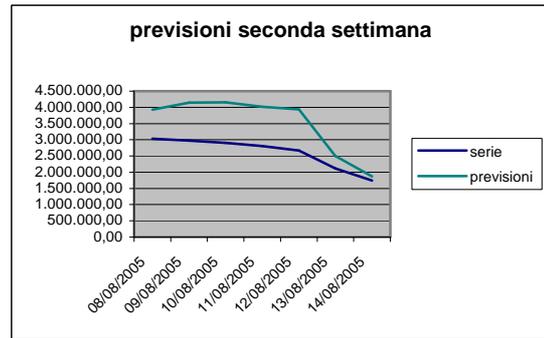
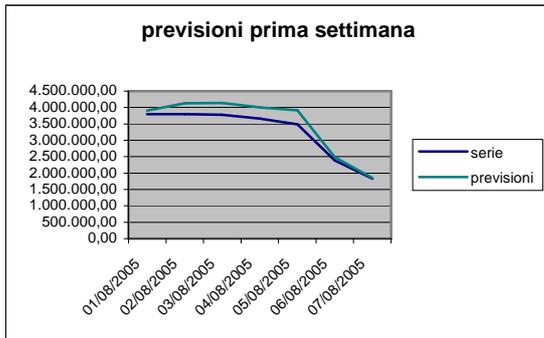
EMA= 773.212,75

L'EM, l'EQM e l'EMA misurano rispettivamente l'errore medio, l'errore quadratico medio e l'errore medio assoluto:

$$EM = \frac{1}{n} \sum_{t=1}^n e_t \quad EQM = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad EMA = \sqrt{\frac{1}{n} \sum_{t=1}^n |e_t|} \quad \text{dove } e_t = y_t - \hat{y}$$

Poiché risulta evidente una sovrastima sistematica dei valori reali, si prova ad effettuare delle previsioni settimana dopo settimana. Il metodo è piuttosto ripetitivo, tuttavia con questo procedimento si ottengono delle stime più accurate.

Previsioni senza aggiornamento settimanale

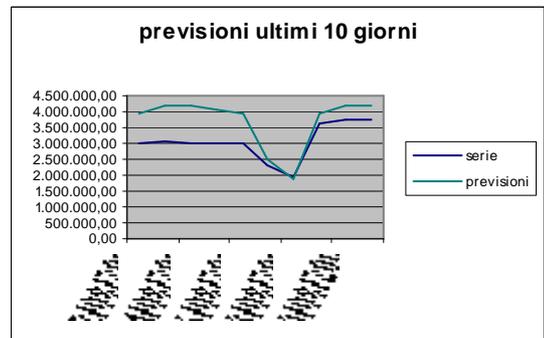
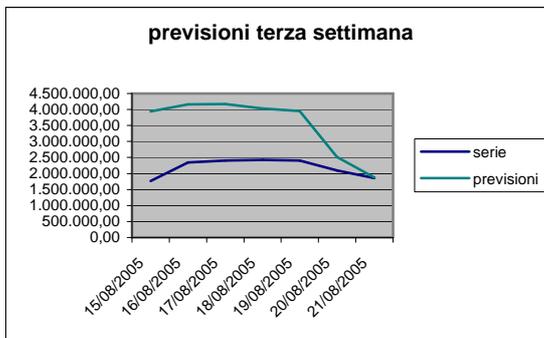


EQM= 282.309

EQM= 1.001.802

EMA= 240.290

EMA= 904.711



EQM= 1.525.966

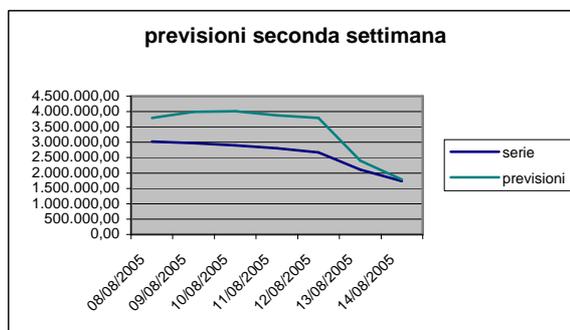
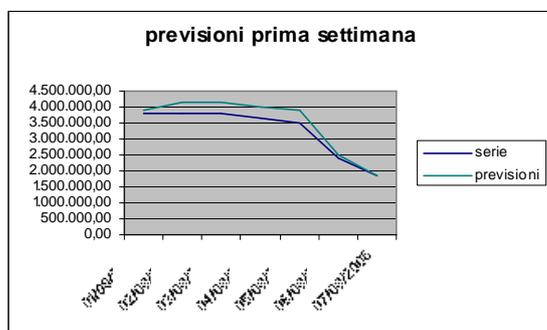
EQM= 769.956

EMA= 1.336.068

EMA= 660.211

Questa volta gli indici di bontà delle previsioni si riferiscono solo alla settimane considerate.

Previsioni con aggiornamento settimanale

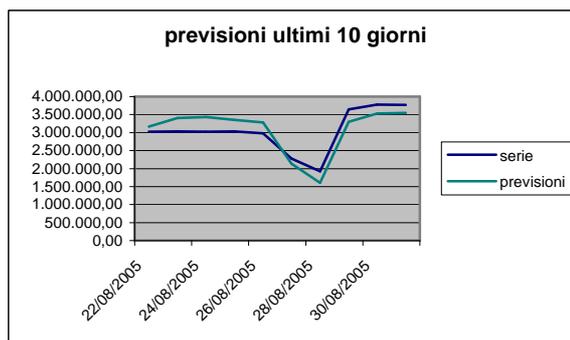
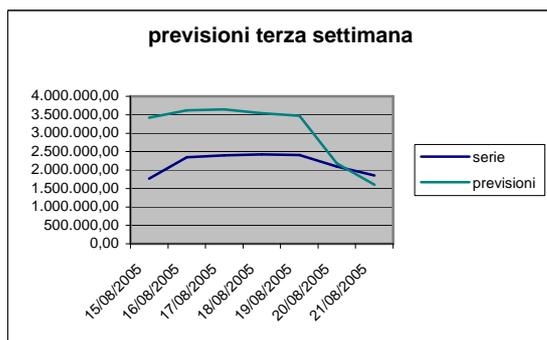


EQM= 282.309

EMA= 240.290

EQM= 874.491,06

EMA= 774.686,25



EQM= 1.091.673,72

EMA= 955.847,28

EQM= 295.607,85

EMA= 282.349,64

Come si può osservare anche solo dai grafici, le previsioni sono migliorate, inoltre, a conferma di ciò, gli indici EQM e EMA sono notevolmente diminuiti, soprattutto nella quarta settimana. È opportuno che il modello si adatti di volta in volta alla serie reale, dal momento che i minuti di chiamate nell'arco dell'anno erano mediamente maggiori.

3.3 SERIE DEI MINUTI VOCE RICARICABILE

Analisi preliminare della serie

La maggior parte degli utenti (3,5 milioni su 4, dati aggiornati al 31 agosto 2005) utilizza schede ricaricabili. Secondo delle indagini condotte dalla divisione marketing di 3, i clienti sono in prevalenza di età compresa tra i 14 e i 35 anni, una popolazione piuttosto giovane, dato che i ragazzi, non percependo un reddito, hanno minori disponibilità di spesa. Le schede prepagate agiscono sulla psicologia del consumatore. Pagando un esiguo supplemento (generalmente il 20% del traffico totale) si ottengono gli stessi benefici di un abbonamento, con la differenza che il credito a disposizione è limitato (si tende a spendere di meno). Un genitore, attraverso questo tipo di pressione, può influenzare la smania di spendere del figlio.

Oltre a questo, le offerte promozionali che la società propone mirano all'incremento delle vendite, per riuscire a concorrere con le altre tre compagnie già affermate.

Le tipologie promozionali maggiormente utilizzate sono:

- 1- traffico bonus durante le settimane centrali del mese;
- 2- vendita di cellulari a basso costo nel periodo estivo.

Entrambi questi fenomeni sono visibili nella serie presa in esame.

Per quanto concerne il traffico aggiuntivo, si osserva, per esempio, la tabella seguente che riporta i dati relativi a settembre 2004:

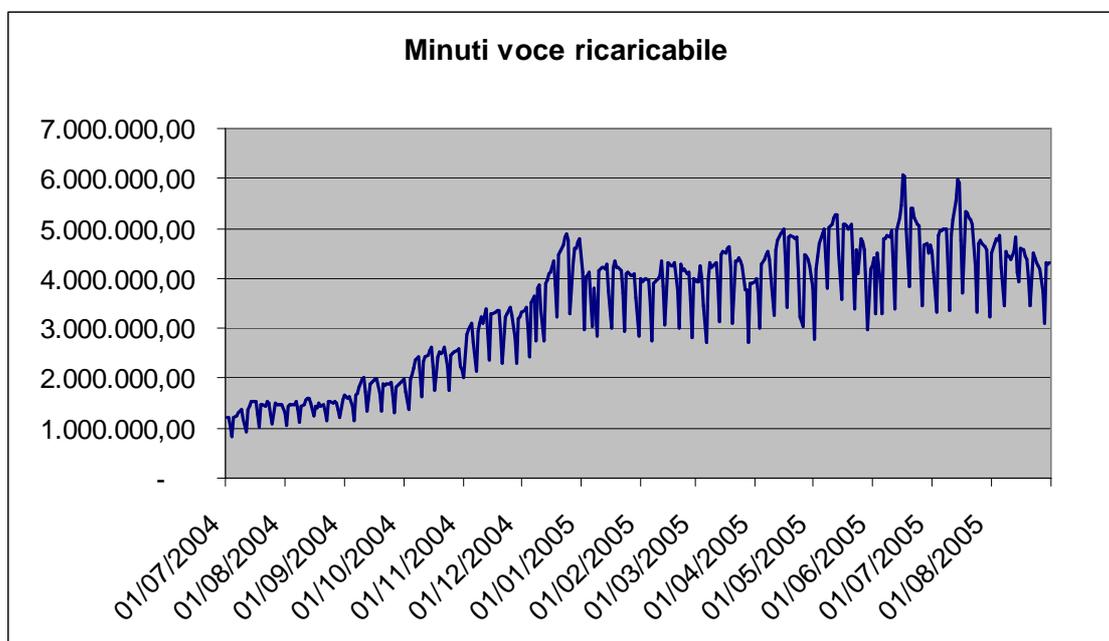
SETTEMBRE 2004	<i>Prima settimana</i>	<i>Seconda settimana</i>	<i>Terza settimana</i>	<i>Quarta settimana</i>
<i>Mercoledì</i>	1.629.084,60	1.818.964,73	1.948.913,32	1.881.043,23
<i>Giovedì</i>	1.600.563,39	1.967.308,78	1.986.448,45	1.891.310,73
<i>Venerdì</i>	1.645.661,03	2.029.273,49	1.968.968,18	1.902.420,92
<i>Sabato</i>	1.424.915,86	1.727.114,39	1.697.104,11	1.661.095,95
<i>Domenica</i>	1.155.115,36	1.354.270,42	1.337.397,83	1.296.335,77
<i>Lunedì</i>	1.646.564,77	1.885.119,80	1.876.934,96	1.837.406,58
<i>Martedì</i>	1.707.337,01	1.920.751,48	1.866.866,06	1.878.778,47

Si nota che nella prima settimana del mese di settembre, le chiamate sono leggermente inferiori rispetto alla media mensile. Nella parte centrale vengono concessi dei minuti gratuiti, accumulati nel mese precedente (tramite il raggiungimento di una soglia stabilita di traffico). Ecco perché c'è un leggero aumento dei minuti di telefonate a partire dal giorno 7.

Nell'ultima settimana invece si ripresenta la situazione iniziale. Ciò è dovuto, oltre che a quest'offerta, alla tendenza ad aspettare la fine del mese prima di effettuare un'ulteriore ricarica (fatto causato più dall'abitudine e non dipendente dall'azienda). Queste considerazioni sono valide in generale per ogni mese.

La vendita di cellulari a basso costo nel periodo estivo, permette di supplire alla modesta diminuzione del traffico. Si verificano perciò due effetti: calo della spesa unitaria (presente anche negli abbonati) compensato da un aumento del numero di cellulari in circolazione.

Il grafico della serie è il seguente:



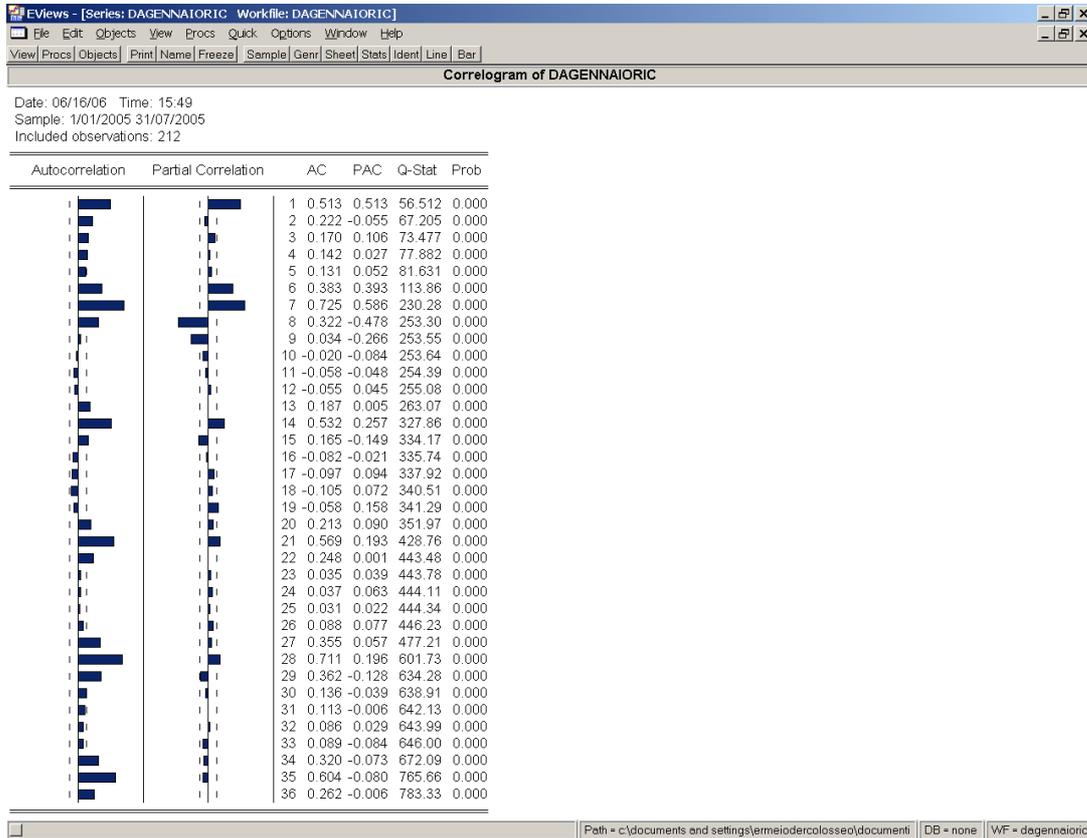
I dati relativi alla prima parte della serie, potrebbero creare dei problemi per l'analisi della stessa, a causa del diverso comportamento rispetto alla seconda metà della serie. Questo primo periodo è infatti caratterizzato da un trend crescente che poi si stabilizza nel secondo periodo.

Durante gli ultimi mesi del 2004 il traffico aumenta considerevolmente, in particolare nel mese di gennaio. Probabilmente in corrispondenza delle feste natalizie, la compagnia telefonica ha effettuato delle promozioni per assicurarsi un maggior numero di clienti o per farli consumare di più. Questa scelta non ha prodotto soltanto effetti nel breve periodo, ma ha innalzato la media di telefonate giornaliere anche nel lungo periodo.

PROCEDURA DI BOX-JENKINS

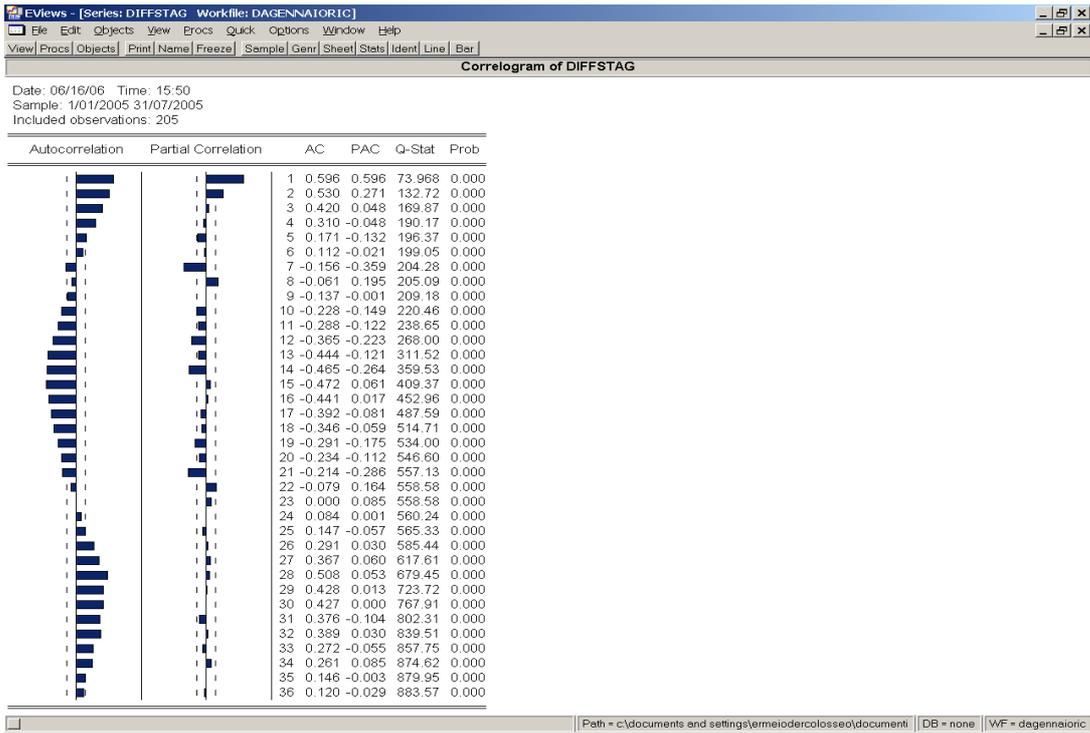
Identificazione

Il grafico seguente riporta gli autocorrelogrammi della serie analizzata.



I correlogrammi suggeriscono di differenziare la serie stagionalmente, visto che l'autocorrelazione globale a cadenza settimanale decresce molto lentamente verso lo zero.

Il grafico degli autocorrelogrammi della serie differenziata stagionalmente è il seguente.



Stima dei parametri e controllo diagnostico

Il grafico precedente, suggerisce una componente autoregressiva di ordine 2. Stimando un SARIMA(2,0,0)x(0,1,0) si ottengono i seguenti risultati.

Dependent Variable: D(DAGENNAIORIC,0,7)

Method: Least Squares

Date: 06/03/02 Time: 11:36

Sample(adjusted): 10/01/2005 31/07/2005

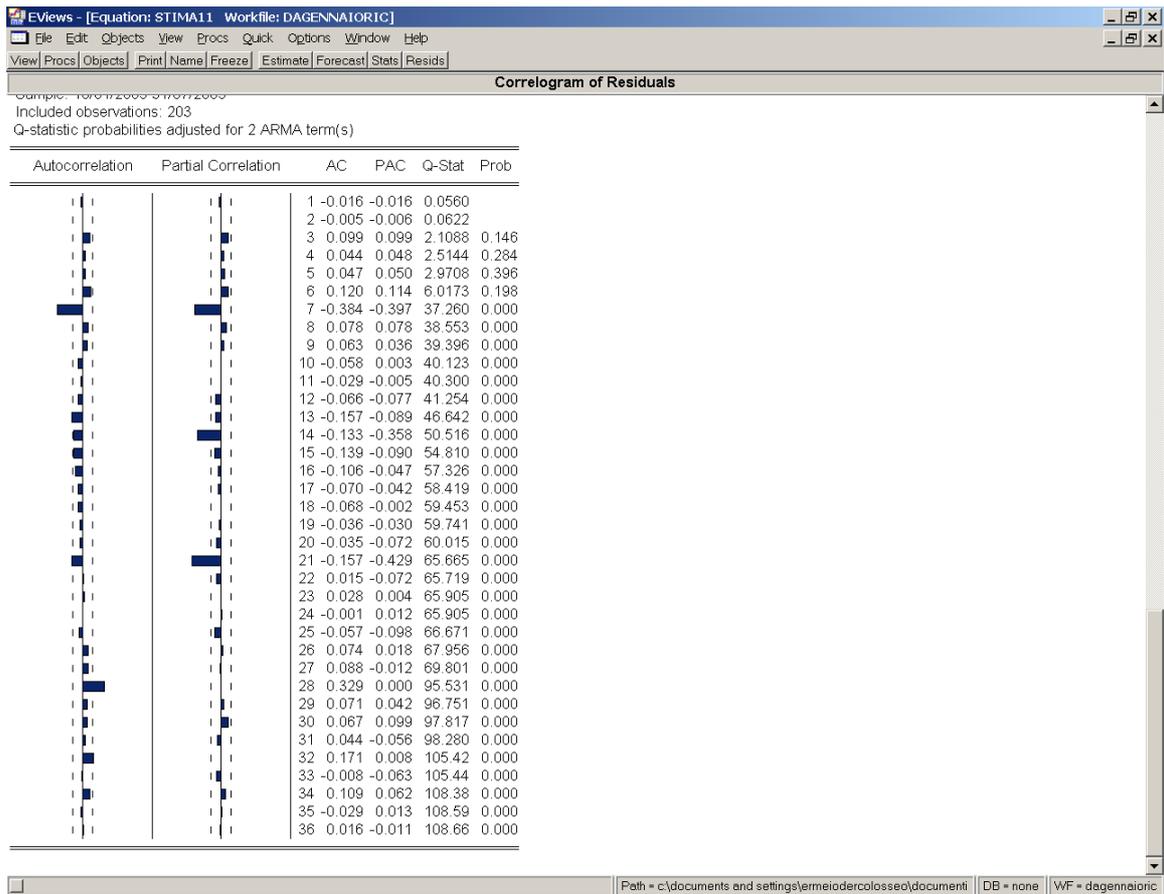
Included observations: 203 after adjusting endpoints

Convergence achieved after 2 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.439251	0.067785	6.480047	0.0000
AR(2)	0.270154	0.067487	4.003036	0.0001
R-squared	0.408268	Mean dependent var		25510.42
Adjusted R-squared	0.405324	S.D. dependent var		486848.2
S.E. of regression	375434.3	Akaike info criterion		28.51936
Sum squared resid	2.83E+13	Schwarz criterion		28.55200
Log likelihood	-2892.715	Durbin-Watson stat		2.025309

Entrambi i coefficienti autoregressivi di breve periodo sono significativi, come dimostra la statistica t di student e il p -value molto vicino allo zero. Si passa all'analisi dei residui.

Il seguente è il grafico delle funzioni di autocorrelazione dei residui.



Si vede come la componente stagionale sia ancora evidente. In particolare la PACF si azzera lentamente e l'ACF ha il settimo e ventottesimo ritardo significativi. Dopo aver provato diverse modelli, si riportano i criteri di Akaike e Schwarz relativi a quelli migliori.

Modelli	Akaike info criterion	Schwarz criterion
SARIMA (2,0,0)x(0,1,4)	27,91	27,98
SARIMA (2,0,0)x(4,1,1)	27,89	27,97
SARIMA (2,0,0)x(4,1,0)	28,35	28,42
SARIMA (2,0,0)x(1,1,4)	28,16	28,22

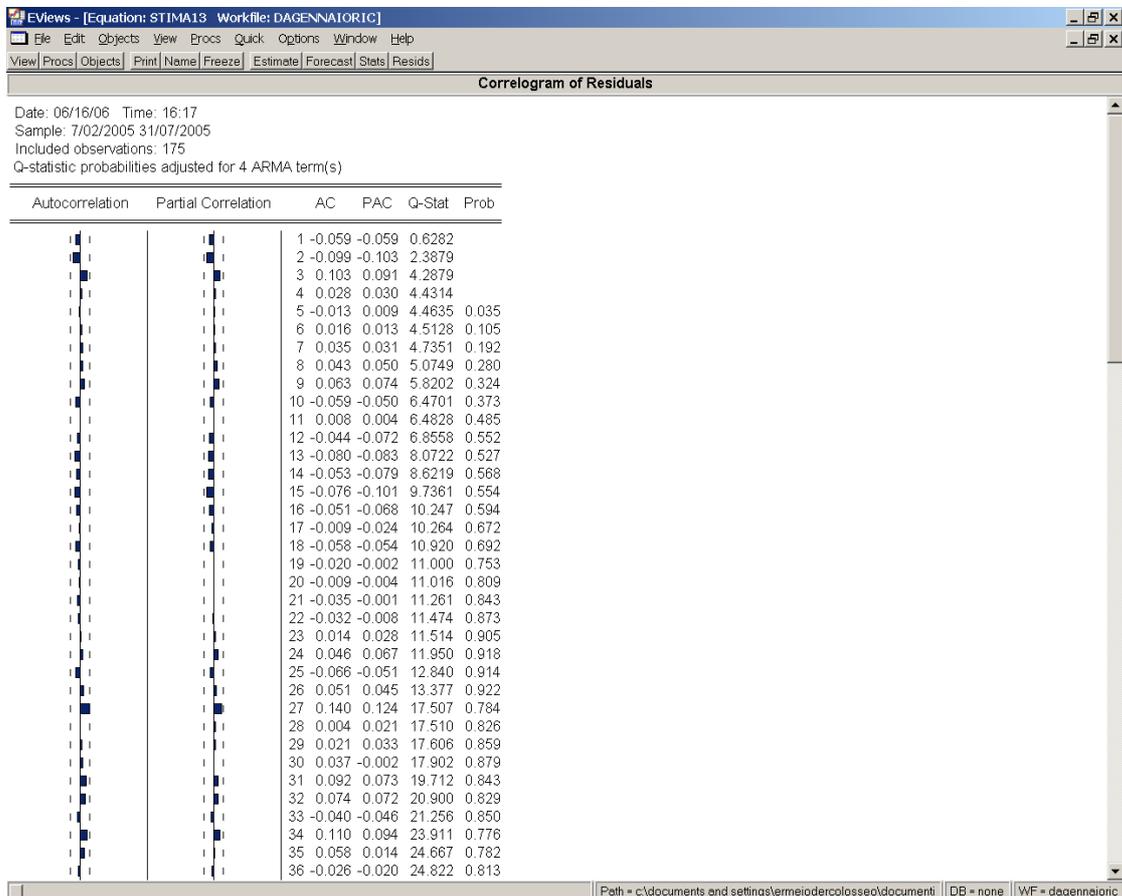
Sulla base di tali risultati si sceglie il secondo modello.

I risultati della stima sono i seguenti:

Dependent Variable: D(DAGENNAIORIC,0,7)
 Method: Least Squares
 Date: 06/03/02 Time: 11:40
 Sample(adjusted): 7/02/2005 31/07/2005
 Included observations: 175 after adjusting endpoints
 Convergence achieved after 10 iterations
 Backcast: 31/01/2005 6/02/2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.423998	0.072864	5.818995	0.0000
AR(2)	0.337158	0.072238	4.667300	0.0000
SAR(28)	0.417796	0.075274	5.550346	0.0000
MA(7)	-0.963801	0.016112	-59.82054	0.0000
R-squared	0.720588	Mean dependent var		25154.74
Adjusted R-squared	0.715686	S.D. dependent var		513272.6
S.E. of regression	273682.8	Akaike info criterion		27.89992
Sum squared resid	1.28E+13	Schwarz criterion		27.97226
Log likelihood	-2437.243	Durbin-Watson stat		2.100218

Come si può vedere dalle funzioni di autocorrelazione dei residui, il modello sembra essere buono.



La statistica di Ljung e Box Q(36) è pari a 24,822 (p-value= 0,813). Pertanto, non essendovi evidenza di comportamenti anomali o di correlazione nei residui, si può ritenere che il modello ben si adatti ai dati.

Resta il dubbio se un'ulteriore differenziazione avrebbe migliorato i risultati finali. Ripetendo l'analoga procedura e stimando un SARIMA (1,1,1)x(0,1,1), i correlogrammi dei residui ottenuti sono peggiori, mentre i criteri di Akaike e Schwarz sono più bassi.

Di seguito si riportano l'output di stima e i correlogrammi dei residui.

Dependent Variable: D(LOG(DAGENNAIORIC),1,7)

Method: Least Squares

Date: 05/25/02 Time: 16:00

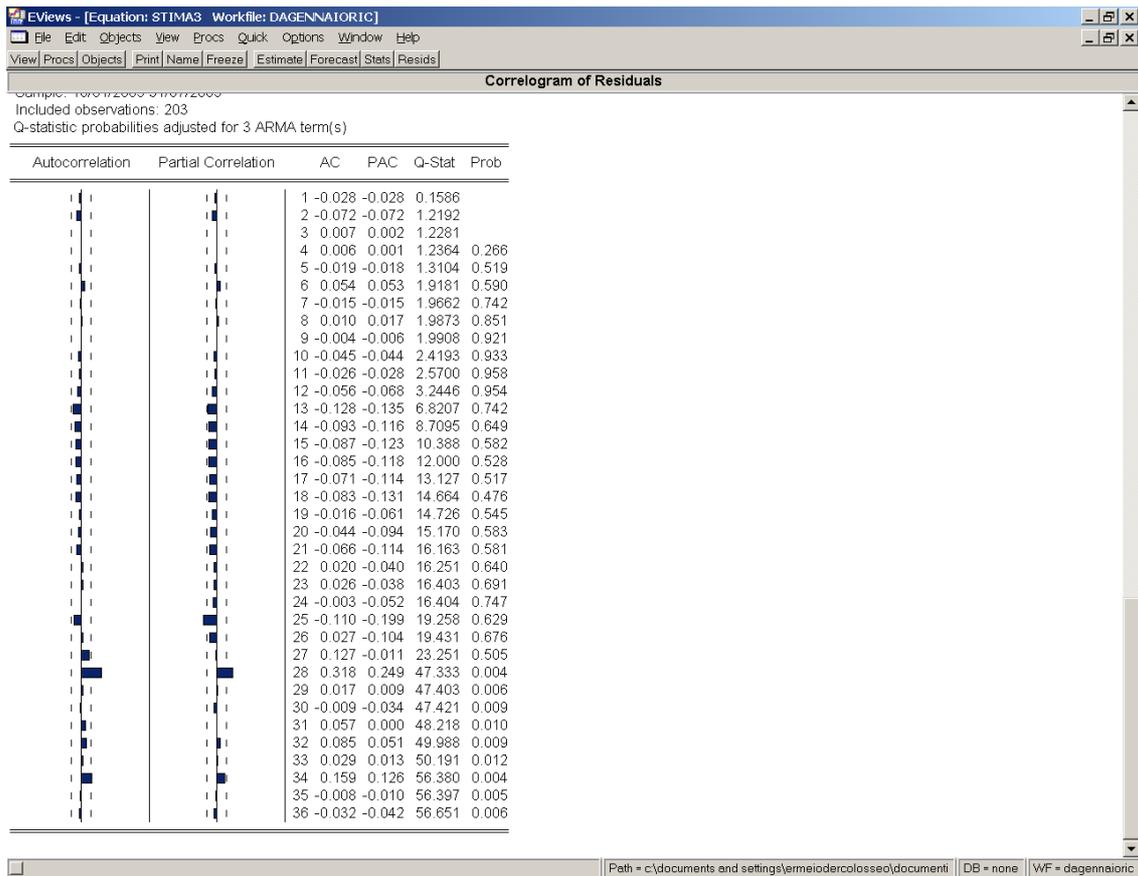
Sample(adjusted): 10/01/2005 31/07/2005

Included observations: 203 after adjusting endpoints

Convergence achieved after 10 iterations

Backcast: 2/01/2005 9/01/2005

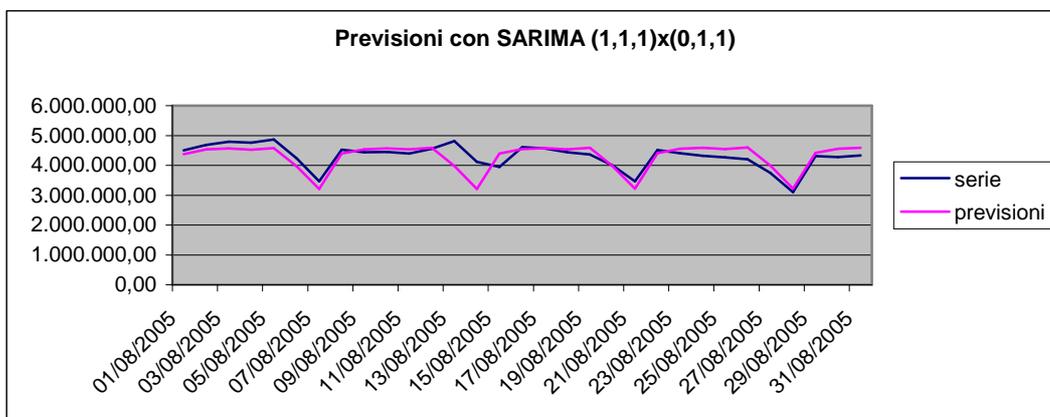
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.285779	0.122124	-2.340066	0.0203
MA(1)	-0.260918	0.120557	-2.164275	0.0316
SMA(7)	-0.963511	0.015937	-60.45824	0.0000
R-squared	0.645707	Mean dependent var		7.84E-05
Adjusted R-squared	0.642165	S.D. dependent var		0.117272
S.E. of regression	0.070152	Akaike info criterion		-2.461648
Sum squared resid	0.984250	Schwarz criterion		-2.412684
Log likelihood	252.8573	Durbin-Watson stat		2.048252



Anche se i residui non sono “ottimali”, il modello viene lo stesso considerato ai fini previsivi.

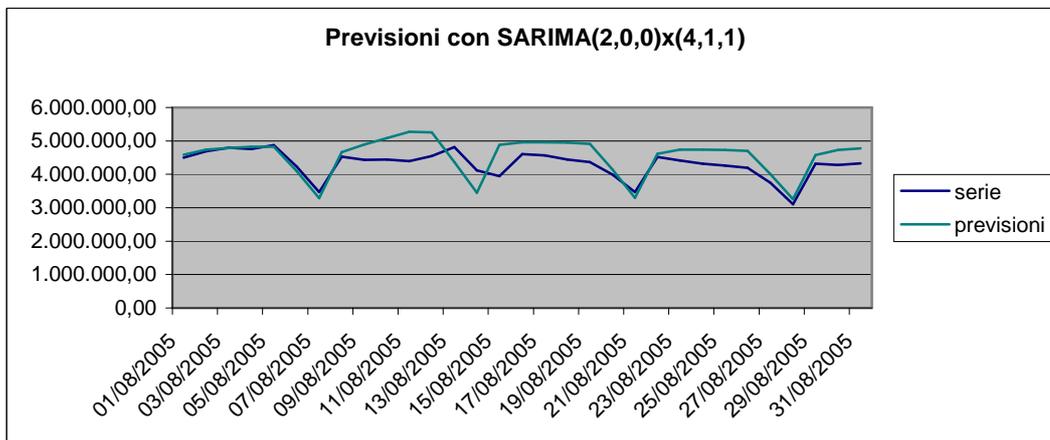
3.4 PREVISIONI DEL TRAFFICO VOCE RICARICABILE

Come prima citato nell'analisi preliminare della serie, il calo estivo delle chiamate è meno accentuato rispetto agli abbonati, a causa del contemporaneo aumento della vendita di cellulari. Proprio per questa ragione non è indispensabile aggiornare le informazioni settimanalmente. Ponendo a confronto la serie reale con le previsioni di un SARIMA(2,0,0)x(4,1,1) e di un SARIMA(1,1,1)x(0,1,1) si riscontrano queste differenze.



EQM= 304.285,35

EMA= 230.737,02



EQM= 428.367,79

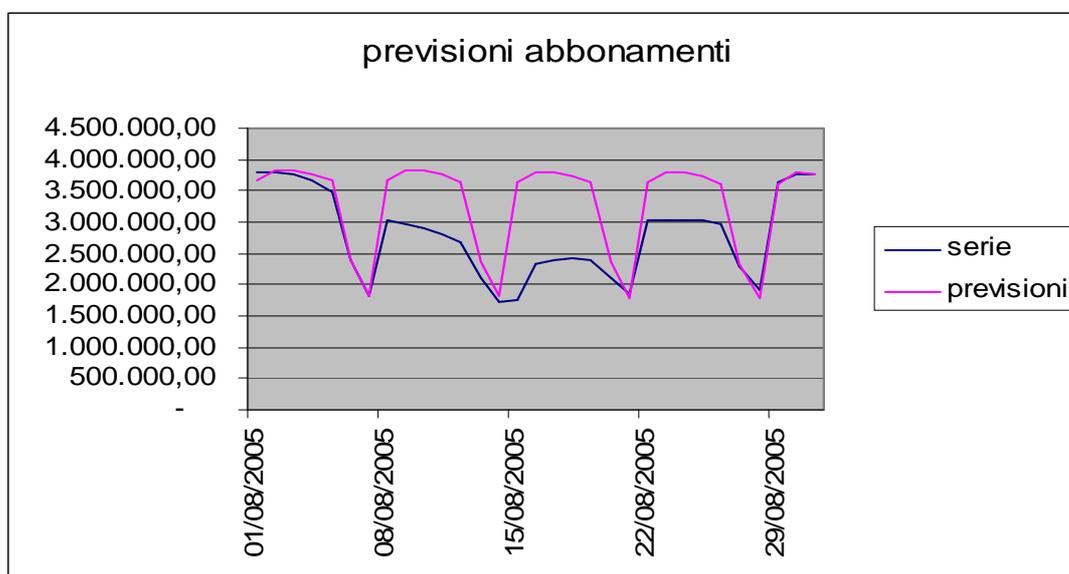
EMA= 352.483,22

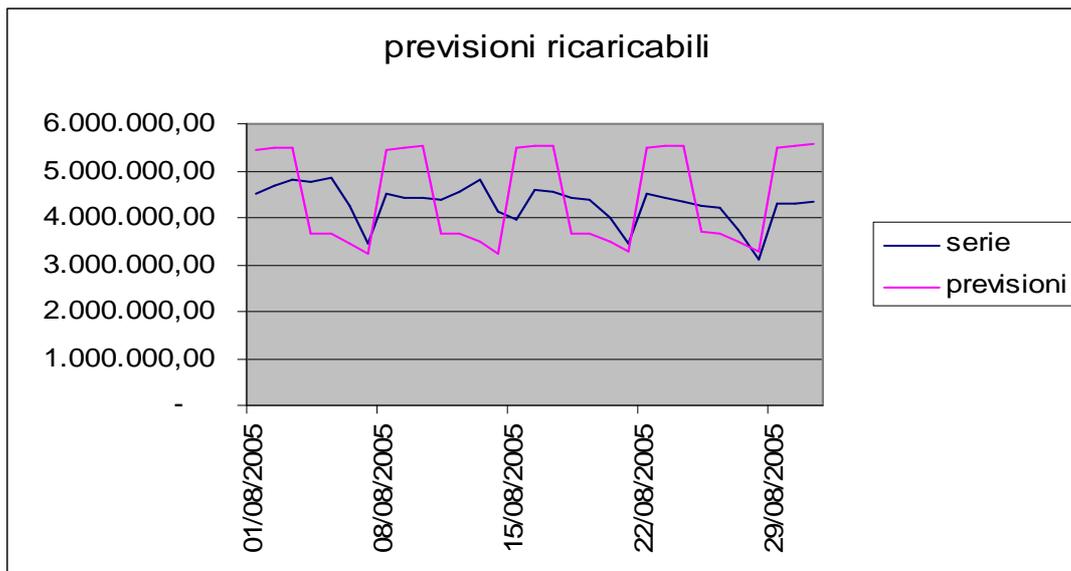
Nonostante il modello SARIMA(1,1,1)x(0,1,1) sia peggiore in fase di adattamento, gli indici di bontà delle previsioni considerati, portano alla scelta di questo modello a fini previsionali.

4. ANALISI DELLE SERIE STORICHE CON IL METODO DI HOLT-WINTERS

4.1 PREVISIONE DEL TRAFFICO ABBONAMENTI E DEL TRAFFICO RICARICABILE

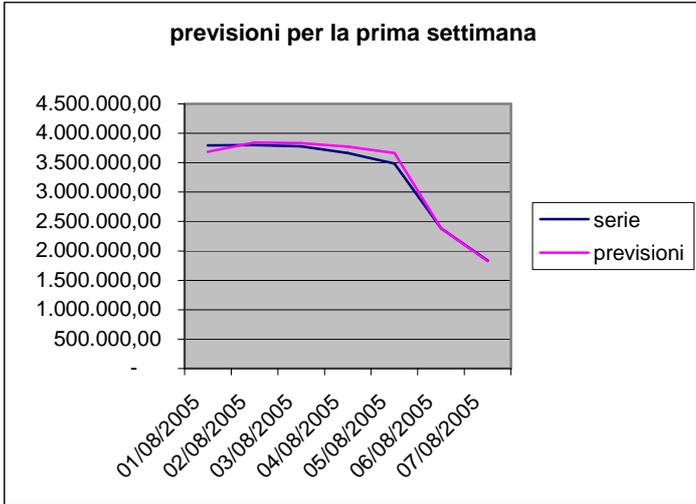
Le procedure (additiva e moltiplicativa) per il lisciamento della serie con il metodo di Holt-Winters sono state costruite in Excel. Le costanti di lisciamento sono state scelte in modo tale da minimizzare l'errore quadratico medio. Poiché dall'analisi degli indici EQM e EMA il metodo additivo risulta essere il migliore in fase di adattamento, nel seguito si utilizza il metodo di H-W stagionale additivo per fare previsioni. Le costanti di lisciamento trovate sono: $\alpha = 0,01$, $\beta = 0,99$, $\gamma = 0,99$. Le previsioni per il mese di agosto riguardanti il traffico voce suddiviso tra abbonamenti e ricaricabili sono le seguenti.



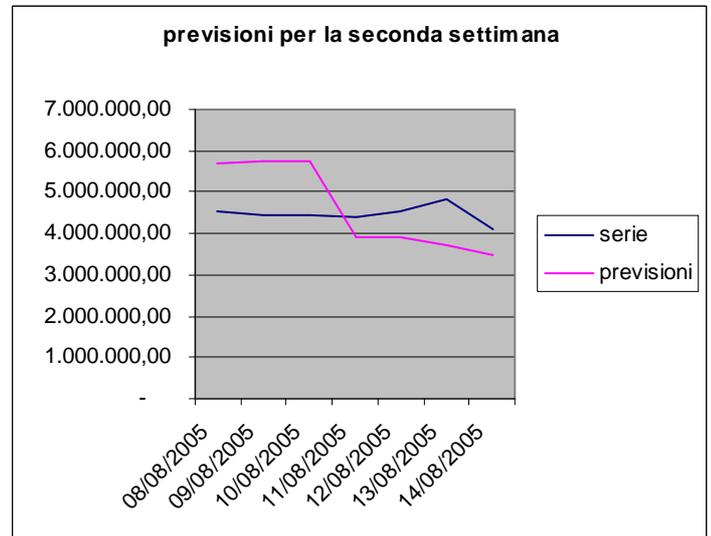
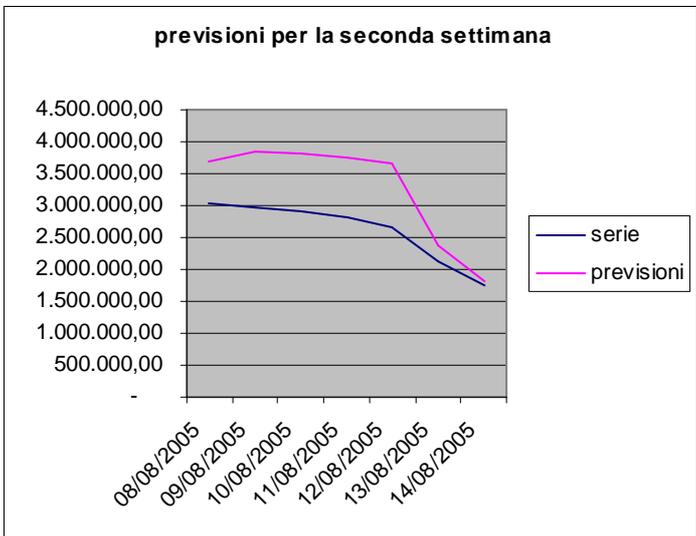
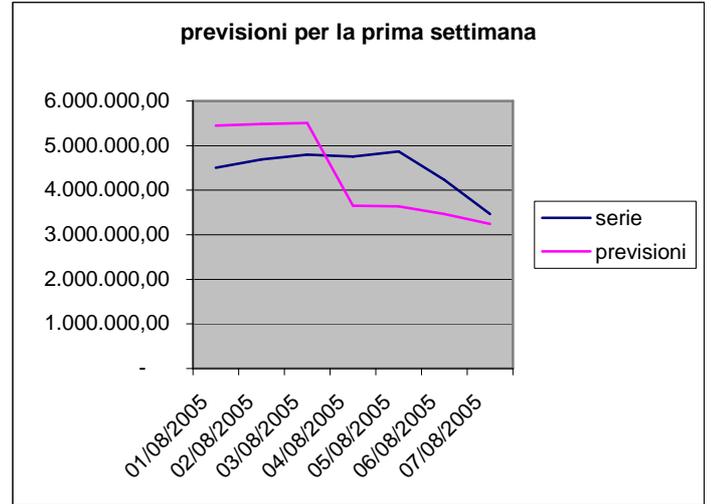


Come già si è detto, il mese di agosto presenta non poche difficoltà per le previsioni, soprattutto a causa della scarsità dei dati che non permettono di trattare la stagionalità annuale. Per come sono costruiti, i modelli di Holt-Winters possono essere utili nel caso di previsioni a breve termine, dal momento che i coefficienti stagionali derivano dalle informazioni rilevate nell'ultima settimana. Nel caso in cui ci sia un evidente abbassamento del trend, che comporta necessariamente un cambiamento anche nella stagionalità, il modello sovrastima (e non di poco) i dati attesi. Aggiornando settimana dopo settimana, questo fenomeno è meno accentuato, come si può vedere dai grafici che confrontano la serie originale con le previsioni.

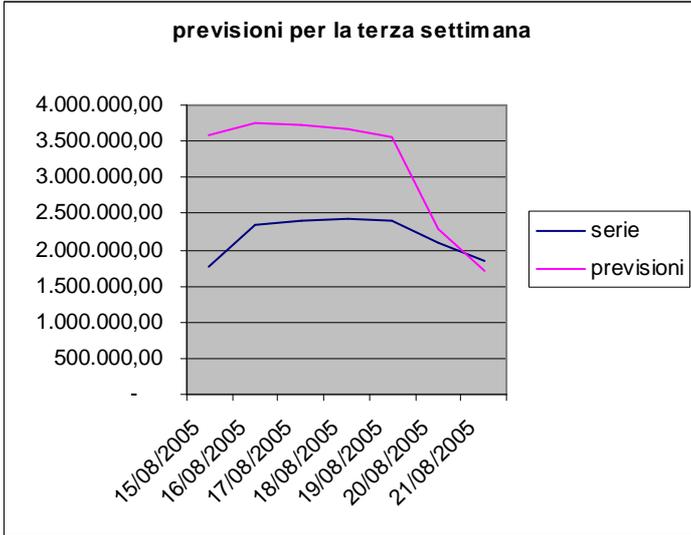
ABBONATI



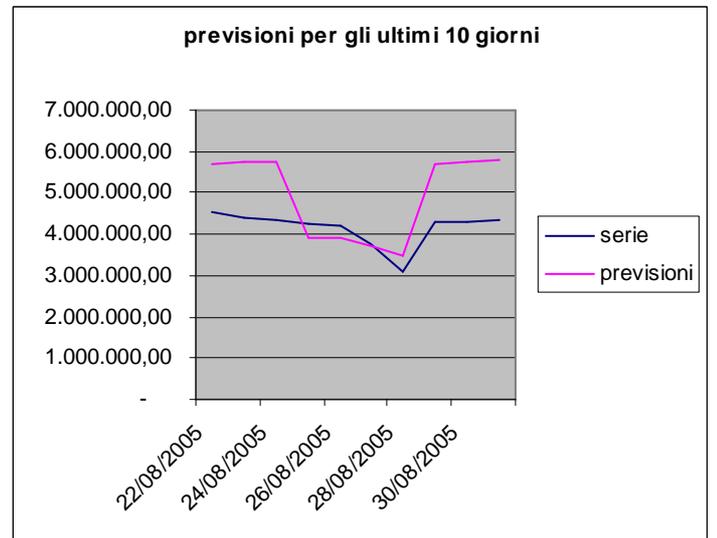
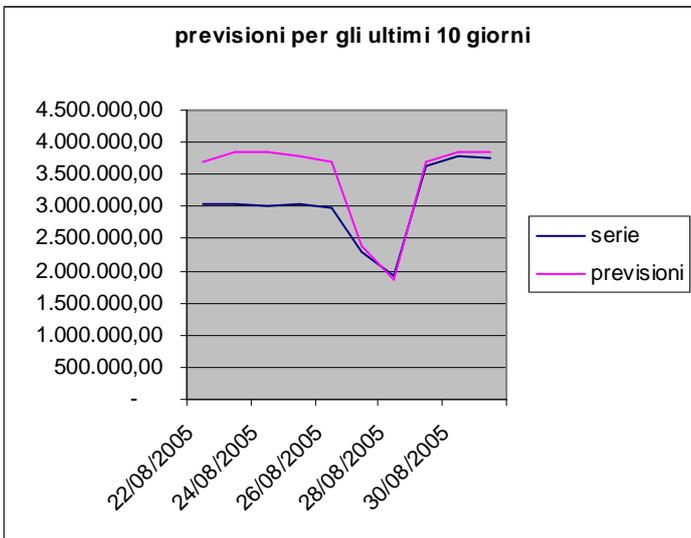
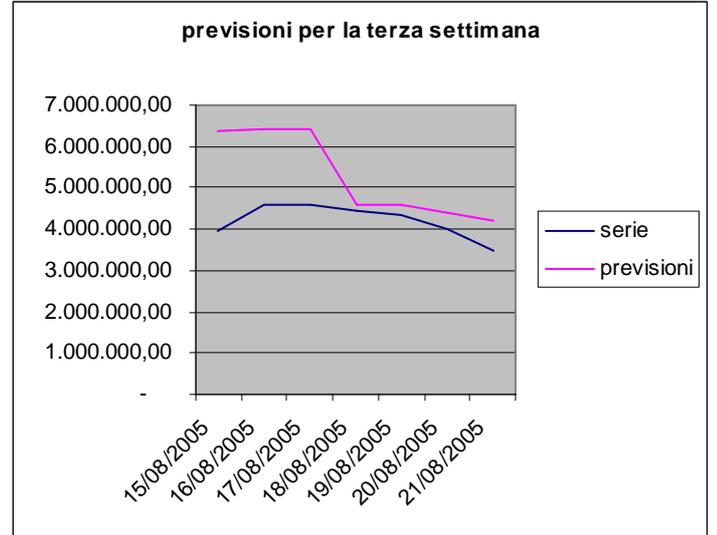
RICARICABILI



ABBONATI



RICARICABILI



4.2 CONFRONTO FRA LE DUE METODOLOGIE UTILIZZATE

Si mettono a confronto gli errori di un modello SARIMA (1,1,1)x(0,1,1) con il modello additivo di Holt Winters, avente come costanti di lisciamento $\alpha=0.01$, $\beta=0.99$, $\gamma=0.99$:

<i>ABBONAMENTI</i>		<i>RICARICABILI</i>
SARIMA(1,1,1)x(0,1,1)	HOLT-WINTERS ADD.	SARIMA(1,1,1)x(0,1,1)
		Intero mese
	Prima settimana	EQM= 304.285
EQM= 282.309	EQM= 93.644	EMA= 230.737
EMA= 240.290	EMA= 72.813	
		HOLT-WINTERS ADDITIVO
	Seconda settimana	Intero mese
EQM= 874.491	EQM= 755.042	EQM= 928.631
EMA= 774.686	EMA= 674.635	EMA= 862.360
	Terza settimana	Prima settimana
EQM= 1.091.673	EQM= 1.190.244	EQM= 878.490
EMA= 955.847	EMA= 1.035.526	EMA= 825.290
	Ultimi 10 giorni	Seconda settimana
EQM= 295.608	EQM= 539.134	EQM= 1.000.901
EMA= 282.350	EMA= 417.557	EMA= 948.438
		Terza settimana
		EQM= 1.377.733
		EMA= 1.082.931
		Ultimi 10 giorni
		EQM= 1.086.814
		EMA= 934.923

5. ANALISI DESCRITTIVA DELLE SERIE COMPONENTI IL TRAFFICO VOCE

5.1 SERIE ABBONATI

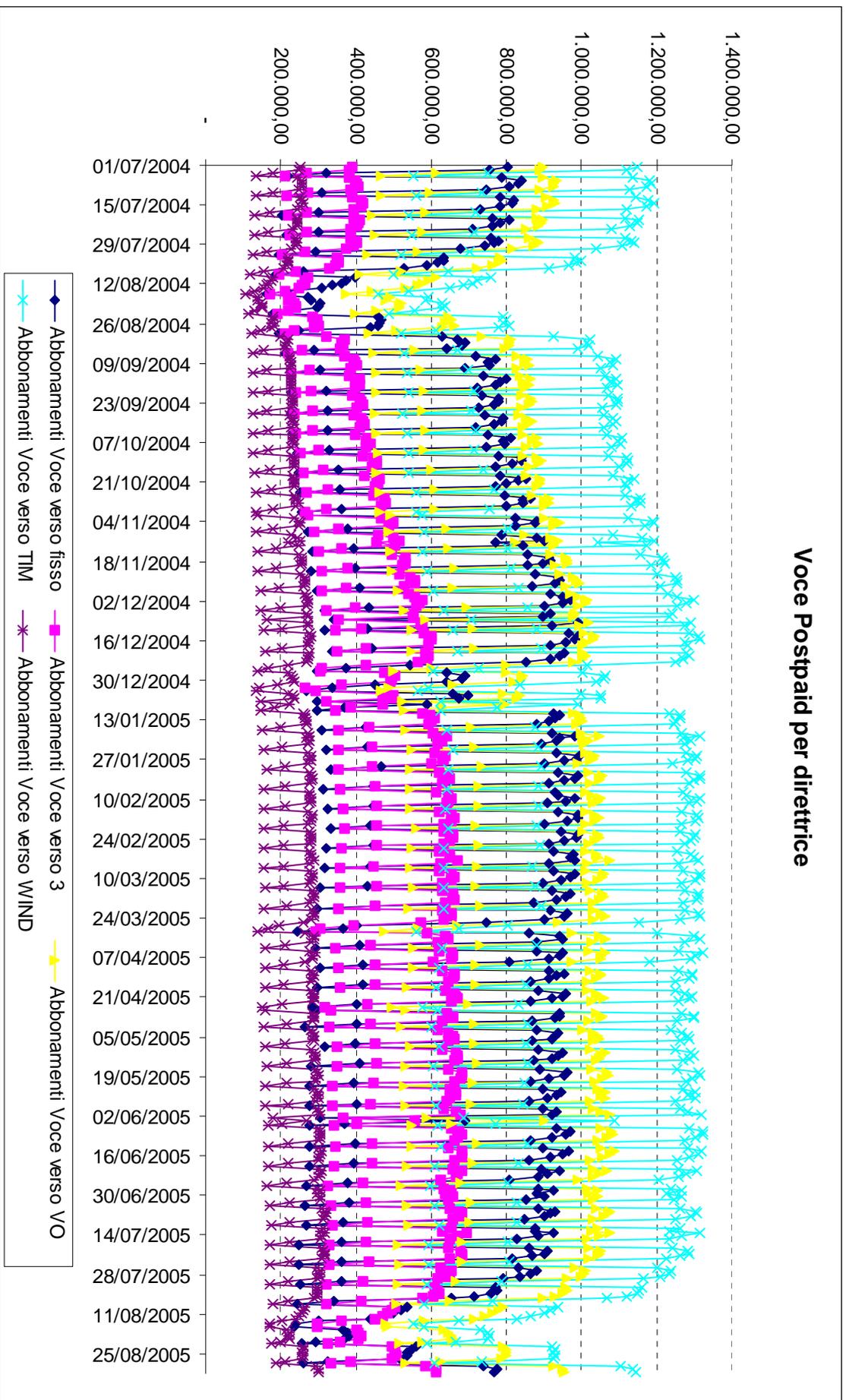
Come si vede dal grafico, le 5 serie che si ottengono sono abbastanza simili fra loro, e possono dare dei suggerimenti sulla mole di clienti per ciascuna compagnia. In particolare, TIM e Vodafone sono le due società che per prime sono entrate nel mercato della telefonia mobile, assicurandosi un gran numero di utenti. Questo fattore ha costretto le nuove entranti (Wind e 3), ad effettuare le promozioni “passa a...”, concedendo dei minuti gratuiti al momento dell’attivazione. In questo modo anche le due aziende hanno potuto concorrere all’interno del settore, sviando l’effetto marchio di TIM e Vodafone.

Il traffico voce verso i numeri di casa sta subendo una graduale riduzione. Probabilmente in un futuro non troppo lontano, il telefono fisso diventerà un apparecchio obsoleto, e favorirà la crescita di tutte e quattro le compagnie.

Comportamento inverso viene riscontrato osservando la serie delle chiamate verso 3. Si nota un trend in continua crescita nell’arco dell’intero anno. Questo può far ben sperare per i guadagni futuri.

Tornando invece ad una analisi più specifica con lo strumento delle serie storiche, non si notano processi differenti da operatore a operatore. In tutti i cinque casi c’è una forte diminuzione del traffico durante i fine settimana, nelle feste natalizie e nelle vacanze estive. Questa tendenza era già visibile nel grafico dei minuti totali, perciò non sembra necessario suddividere la serie.

Voce Postpaid per direttrice



5.2 SERIE RICARICABILI

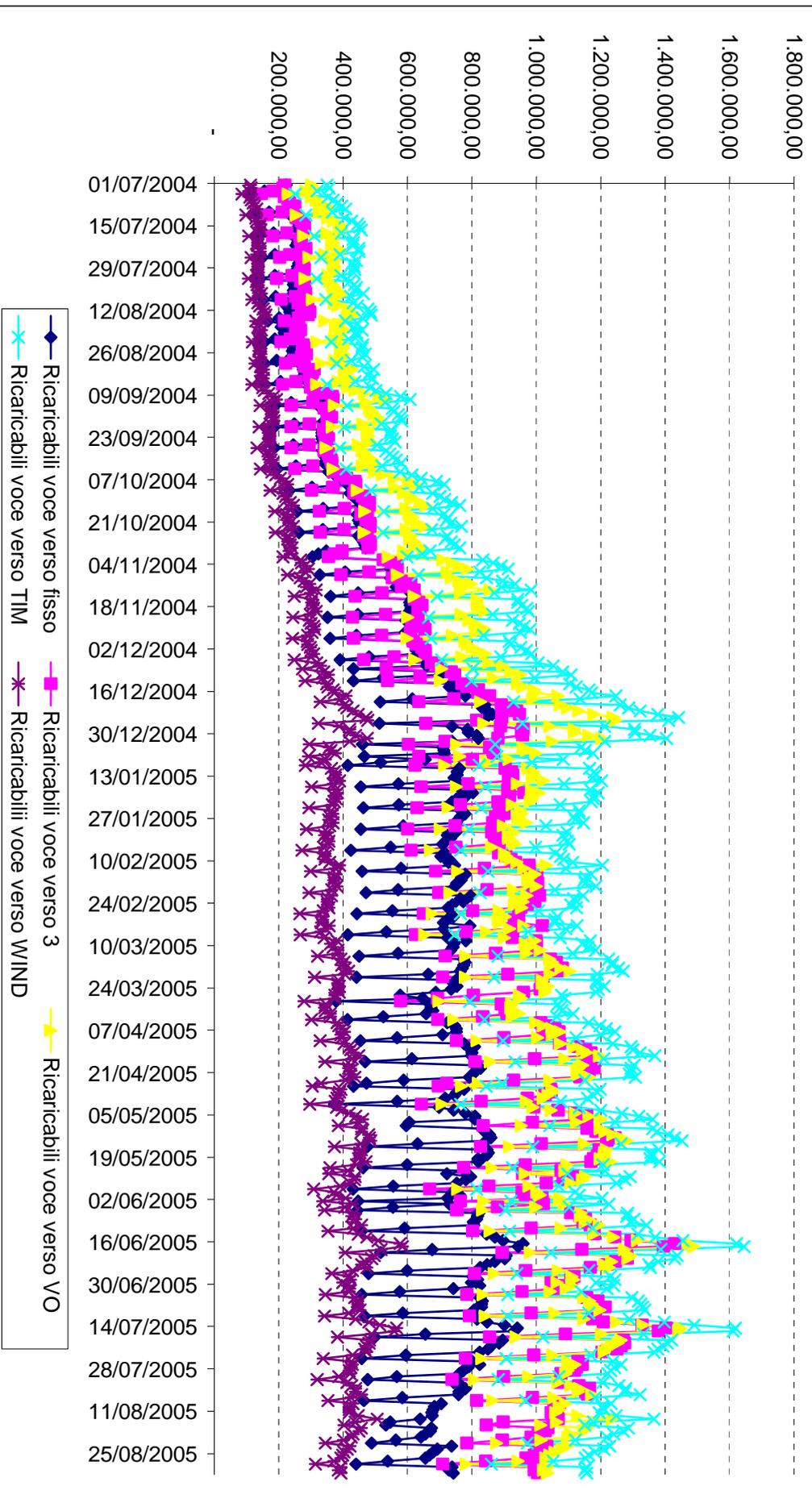
Anche nel caso delle ricaricabili, l'andamento delle 5 serie ricorda quello della serie complessiva, ovvero un trend crescente fino a fine anno e una crescita meno marcata fino a fine agosto. Rispetto agli abbonati i picchi di traffico sono molto più alti, arrivando ad oltre 1.600.000 minuti totali. La stagionalità settimanale provoca dei cali di chiamate molto meno evidenti a differenza del traffico abbonamenti.

È evidente il forte rialzo di chiamate per l'anno 2004 (vedi grafico). Per spiegare il motivo che ha generato questo fenomeno, si possono fare delle considerazioni riguardo l'intero mercato della telefonia mobile, dal momento che la maggior parte della clientela possiede una scheda ricaricabile (nel caso di H3G 3,5 milioni su 4).

Secondo alcune stime la crescita del settore è stata del 3,7%. I principali fattori di tale sviluppo sono stati la generale riduzione delle tariffe verso gli utenti dei servizi di telefonia mobile e l'espansione della banda larga, che ha registrato un notevole aumento di linee di accesso (crescita del 27% rispetto all'anno precedente). Si tratta di un incremento significativo, favorito dall'immissione sul mercato di prodotti innovativi, legati dalla convergenza di voce, dati e video e dei servizi di nuova generazione. Anche la domanda ha registrato un rialzo e questo ha favorito, da un lato, un aumento della competizione tra operatori, soprattutto per quanto riguarda la specializzazione e la capacità di offrire soluzioni personalizzate, dall'altro ha incoraggiato processi di concentrazione e accordi di partnership tra operatori, sia a livello nazionale che internazionale.

In questo senso H3G ha potuto sfruttare a pieno l'accordo stipulato nel 2002 con NEC e ALCATEL, le due aziende che hanno permesso di supportare l'introduzione dei servizi voce, dati e multimedia. La nuova tecnologia UMTS consiste proprio in questa convergenza nella trasmissione di uno o più segnali (dati, video, voce, ecc.) che condividono la stessa linea trasmissiva.

Voce Prepaid per direttrice



6. CONSIDERAZIONI FINALI

6.1 CONCLUSIONI

Avendo analizzato le serie riguardanti il traffico voce degli abbonati e delle ricaricabili, si evince che il comportamento dei due fenomeni è abbastanza simile. A conferma di ciò, i modelli che meglio si sono adattati ai due casi studio sono stati gli stessi sia con il metodo di Holt-Winters (stesse costanti di liscio per entrambe le serie), sia con i processi stocastici ARIMA stagionali (SARIMA(1,1,1)x(0,1,1)).

Per quanto concerne invece la fase previsiva, si notano dei comportamenti differenti fra la serie dei minuti abbonamento e quella dei minuti ricaricabile.

Nel primo caso, dal momento che gli indici per valutare la bontà di previsione dei metodi utilizzati non presentano differenze significative, si può optare sia per il modello SARIMA(1,1,1)x(0,1,1) che per il metodo di Holt-Winters stagionale additivo .

Nel secondo caso, basta confrontare l'EQM e l'EMA relativi all'intero mese di agosto per capire che il SARIMA(1,1,1)x(0,1,1) fornisce delle previsioni migliori e permette di "risparmiare informazioni", in quanto non è necessario l'aggiornamento delle stesse settimana dopo settimana.

BIBLIOGRAFIA

- (1) Box, Jenkins, Reinsel, (1994) “Time Series Analysis: Theory-Practice”
- (2) Di Fonzo T. , Lisi F. (2005) “Serie storiche economiche”, Carrocci

BIBLIOGRAFIA DEI SITI INTERNET

- (3) <http://www.tre.it/H3G/>
- (4) <http://www.wikipedia.org/>