



DEPARTMENT OF ECONOMICS,  
MANAGEMENT AND STATISTICS  
UNIVERSITY OF MILAN – BICOCCA

## DEMS WORKING PAPER SERIES

### **Modelli di previsione per l'aggiornamento di serie mensili del turismo russo in Italia**

Giovanni Tonini

No. 314 – November 2015

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa  
Università degli Studi di Milano - Bicocca  
<http://dems.unimib.it/>

# Modelli di previsione per l'aggiornamento di serie mensili del turismo russo in Italia

*Giovanni Tonini*<sup>\*</sup>

## **Abstract**

Il presente contributo affronta il problema di come aggiornare le serie della domanda turistica russa in Italia utilizzando modelli di previsione di tipo sia univariato che multivariato. Tale aggiornamento si rende necessario perché i dati definitivi del movimento mensile dei clienti negli esercizi ricettivi (arrivi e presenze) vengono pubblicati dall'ISTAT con un consistente ritardo, il che impedisce di effettuare analisi tempestive di detto movimento. Un modo per ovviare a questo inconveniente è quello di adottare soluzioni di stima dei valori futuri degli arrivi e delle presenze. Per ottenere tali valori si è fatto ricorso a opportuni modelli di previsione che, andando dai più semplici ai più complessi, sono costituiti da: 1) i modelli univariati della classe SARIMA; 2) i modelli dinamici di regressione multipla; 3) i modelli con funzioni di trasferimento. Ciascuno di tali modelli è stato proposto in due versioni, una che non incorpora il trattamento degli outliers e l'altra, invece, che lo comprende; ciò ha reso possibile valutare se è in che misura vi fosse un impatto significativo degli outliers sulle previsioni. Inoltre, l'applicazione ha riguardato diversi intervalli temporali e segmenti turistici. In particolare, si è voluto verificare l'effetto sulle previsioni della diversa durata dell'intervallo di osservazione, cioè passando da un intervallo più lungo (undicennio 2001-2011) ad uno più breve e recente (quinquennio 2007-2011). Per quanto riguarda invece il confronto tra le previsioni dei diversi segmenti turistici considerati, si è valutato come variano tali previsioni al variare del tipo di esercizio ricettivo (alberghiero e complementare) e della nazionalità dei turisti (russa, tedesca, straniera, italiana). Infine, l'accuratezza delle previsioni di orizzonte da 1 a 12 mesi, valutata con due opportuni indici relativi, è stata posta a confronto anche con la bontà delle stime, verificando così le differenze tra la capacità previsiva dei modelli proposti e la loro bontà di adattamento ai valori passati.

**Key words:** Modelli per l'aggiornamento di serie del turismo; previsori univariati e multivariati; previsioni della domanda turistica; presenze russe in Italia; confronti intersettoriali, intertemporali e internazionali

**JEL codes:** C32, L83

<sup>\*</sup>Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa, University of Milan Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy. E-mail: [giovanni.tonini@unimib.it](mailto:giovanni.tonini@unimib.it)

## 1.1 Introduzione

Com'è noto, i dati definitivi del movimento mensile dei clienti negli esercizi ricettivi (arrivi e presenze) vengono pubblicati dall'ISTAT<sup>1</sup> con un considerevole ritardo. Ad esempio, i dati disaggregati relativi al 2013 hanno visto la luce alla fine del 2014. Ciò impedisce di effettuare analisi dettagliate tempestive di tale movimento e orienta verso soluzioni di stima dei valori futuri degli arrivi e delle presenze. A tal fine si possono utilizzare modelli di previsione delle serie storiche sia di tipo univariato che di tipo multivariato. Nel primo caso, la scelta del modello è giustificata dalla sua relativa semplicità, che comunque non va a scapito dell'accuratezza delle previsioni. Nel secondo caso, invece, è possibile sfruttare il legame tra i predittori e la serie da prevedere per ottenere i valori futuri di quest'ultima in funzione dei valori futuri dei predittori. Questi ultimi valori possono essere noti in anticipo, e ciò è preferibile, oppure dovranno essere a loro volta previsti, e ciò introduce un ulteriore elemento di incertezza nelle previsioni.

Nel seguito, dopo una breve rassegna della letteratura sulle previsioni della domanda turistica, illustreremo anzitutto i modelli di previsione delle presenze turistiche mensili, formalizzando dapprima i modelli univariati SARIMA e successivamente i modelli dinamici di regressione multipla e i modelli con funzioni razionali di trasferimento. Si passerà poi ad esporre gli indici utilizzati per sintetizzare l'accuratezza delle previsioni, uno dei quali presenta degli elementi di novità rispetto agli indici solitamente impiegati in letteratura. Infine, si descriveranno i risultati ottenuti in termini di accuratezza delle previsioni della domanda turistica russa, confrontando poi tali risultati sia con la bontà delle stime sia con le previsioni delle presenze mensili di altre nazionalità.

## 1.2 Breve rassegna della letteratura sulle previsioni turistiche

A partire dalla metà degli anni '80 sono stati pubblicati diversi articoli in cui si valutano le performance dei modelli di previsione della domanda turistica di tipo sia univariato che multivariato. Alcuni di tali articoli sono delle rassegne di precedenti studi (Calantone et al., 1987; Sheldon e Var, 1985; Li et al., 2005; Song e Li, 2008; Uysal e Crompton, 1985; Witt e Witt, 1995); altri, invece, sono delle applicazioni al turismo di svariati modelli con il fine di individuare i previsori più accurati nei vari contesti in esame (Athanasopoulos et al., 2011; Fritz et al., 1984; Kulendran e King, 1997; Song et al., 2003; Tonini, 1994).

---

<sup>1</sup> I dati utilizzati in questo contributo sono perlopiù di fonte ISTAT e riguardano le presenze turistiche alberghiere, complementari e totali. Essi sono stati ricavati da: ISTAT (2002 e 2004), per gli anni 2001 e 2002; ISTAT (2005-2011), per gli anni 2003-2010; ISTAT (2012-2014), per gli anni 2011-2012. Oltre a questi dati, si è fatto ricorso anche alle serie mensili dei pernottamenti di fonte Banca d'Italia (2014) che, essendo note con almeno 6 mesi di anticipo rispetto alle serie ISTAT, risultano utili per effettuare gli opportuni aggiornamenti di queste ultime, limitatamente alle serie alberghiere, come si vedrà meglio in seguito.

Il più recente e completo lavoro (Athanasopoulos et al., 2011), tra quelli in cui si confrontano le performance dei modelli di previsione della domanda turistica, valuta comparativamente, su centinaia di serie mensili, trimestrali e annuali, l'accuratezza previsiva di numerosi modelli sia univariati (ARIMA, exponential smoothing, ecc.) che multivariati (regressione statica e dinamica, modelli autoregressivi a ritardi distribuiti, modelli con parametri variabili nel tempo e autoregressione vettoriale). I risultati di questa notevole mole di confronti portano gli Autori a concludere che, nella previsione della domanda turistica basata su serie mensili, i modelli univariati funzionano meglio di quelli multivariati che utilizzano variabili esplicative. In tale contesto, si sono dimostrati particolarmente performanti i modelli SARIMA, in quanto hanno dato luogo a previsioni generalmente più accurate di quelle degli altri modelli, anche più complessi, presi in esame. Questi risultati sono in sintonia con quanto riscontrato da altri Autori (Kulendran e King, 1997; Newbold e Granger, 1974; Witt e Witt, 1995) e portano a privilegiare i previsori univariati, anche in virtù del principio di parsimonia (Ledolter e Abraham, 1981) che, a parità di accuratezza delle previsioni, fa preferire il modello più semplice a quello più complesso. Le precedenti conclusioni, quindi, giustificano ampiamente la scelta compiuta in questo lavoro di includere la modellistica univariata SARIMA tra i modelli utilizzati per prevedere le presenze mensili dei turisti russi e delle altre nazionalità considerate.

Nel seguito, oltre ai previsori univariati, si è deciso di impiegare anche previsori multivariati, sia perché i primi non sempre risultano superiori ai secondi, almeno operando con serie annuali (Allen e Fildes, 2001; Song et al., 2003), sia perché è interessante verificare se e in che misura la maggior capacità esplicativa dei modelli multivariati si traduca anche in una maggior capacità previsiva. Inoltre, la struttura multivariata consente di incorporare nel modello anche variabili anticipatrici che possono risultare particolarmente utili in chiave previsiva.

In un precedente contributo (Tonini, 2015) si sono utilizzati i modelli dinamici di regressione multipla per stimare il trend e la stagionalità delle presenze mensili dei turisti russi e di altre nazionalità; tali modelli hanno mostrato una apprezzabile bontà di adattamento e capacità esplicativa della domanda turistica. Sembra perciò naturale utilizzare tali modelli anche in sede previsiva, se non altro per valutare se a un buon adattamento ai dati del passato corrisponda pure una soddisfacente capacità di previsione dei valori futuri.

Nel caso dei modelli multivariati, come si è detto, per prevedere la variabile risposta è necessario prevedere anzitutto le variabili esplicative e ciò introduce un ulteriore elemento di incertezza nel processo previsivo. Tale incertezza, tuttavia, può essere ridotta utilizzando predittori di cui si conoscono già i valori futuri come, ad esempio, le funzioni del tempo  $t$  e le variabili stagionali dummy incorporate nei suddetti modelli dinamici di regressione multipla (Tonini, 2015).

Altri predittori utili a fini previsivi sono le variabili anticipatrici, che sono legate alla variabile dipendente da prevedere ma i cui valori sono noti in anticipo rispetto a quest'ultima; in questo caso le previsioni si configurano come un'operazione di aggiornamento dei dati, in quanto si sfruttano i valori più aggiornati di una o più serie esplicative per prevedere (aggiornare) i valori della serie dipendente. Per eseguire detta operazione sembrano adeguati i modelli con funzioni di trasferimento in forma

razionale i quali, pertanto, saranno utilizzati nel seguito in aggiunta e come generalizzazione degli altri previsori, univariati e multivariati, presi in considerazione.

### 1.3 Modelli di previsione delle presenze turistiche mensili e relativi indici di accuratezza

Come anticipato in apertura, per prevedere i valori futuri delle presenze turistiche mensili si è fatto ricorso a modelli sia univariati che multivariati. Nel primo caso si sono utilizzati modelli della classe SARIMA, mentre nel secondo caso i modelli impiegati sono stati di due tipi: i modelli dinamici di regressione multipla e i modelli con funzioni razionali di trasferimento. Questi ultimi, com'è noto, rappresentano una classe molto ampia di modelli, che comprende come casi particolari anche i modelli univariati SARIMA e i modelli dinamici di regressione multipla utilizzati in Tonini (2015).

Nel seguito, si riporta una breve formalizzazione della modellistica citata, partendo dal caso particolare per arrivare a quello più generale e facendo riferimento, senza perdita di generalità, a serie mensili (stagionalità  $s = 12$ ). In ogni caso, le versioni del modello considerate sono due, la prima delle quali non opera alcun aggiustamento degli outliers, mentre la seconda li aggiusta.

#### 1.3.1 Modelli di previsione SARIMA

Pur essendo di natura univariata, i modelli SARIMA (Seasonal AutoRegressive Integrated Moving Average) si sono rivelati dei previsori accurati della domanda turistica mensile (Athanasopoulos et al., 2011; Tonini, 1994). Essi inoltre possono costituire un utile termine di riferimento rispetto al quale confrontare le previsioni fornite da modelli più complessi di tipo multivariato.

In generale, e con riferimento a serie storiche mensili, i modelli SARIMA si possono formalizzare come segue:

$$\Delta \cdot Y_t = c_1 + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.1)$$

dove:  $\Delta = \Delta^d \cdot \Delta_{12}^D = (1 - B)^d \cdot (1 - B^{12})^D$  è l'operatore differenza con componente non stagionale di ordine  $d$  e componente stagionale di periodo 12 e grado  $D$ ;  $B$  è l'operatore di ritardo tale per cui  $B \cdot Y_t = Y_{t-1}$ ;  $Y_t$  è la variabile da prevedere, rappresentata da una serie storica mensile di presenze turistiche;  $c_1$  è un termine costante;  $\theta(B) \cdot \theta_{12}(B^{12}) = (1 - \theta_1 B - \dots - \theta_q B^q) \cdot (1 - \theta_{12} B^{12} - \dots - \theta_{12Q} B^{12Q})$  è l'operatore media mobile invertibile con componente non stagionale di ordine  $q$  e componente stagionale di periodo 12 e grado  $Q$ ; inoltre,  $\phi(B) \cdot \phi_{12}(B^{12}) = (1 - \phi_1 B - \dots - \phi_p B^p) \cdot (1 - \phi_{12} B^{12} - \dots - \phi_{12P} B^{12P})$  è l'operatore autore-

gressivo stazionario con componente non stagionale di ordine  $p$  e componente stagionale di periodo 12 e grado  $P$ . Entrambi questi operatori soddisfano le consuete condizioni di ammissibilità dei modelli SARMA (Piccolo, 1990, pp. 163-166). Infine,  $a_t \approx WN(0, \sigma^2)$  è una successione di variabili casuali con media nulla, omoschedastiche e incorrelate, nota come processo White Noise.

Se nel modello SARIMA introduciamo il trattamento degli outliers, la sua struttura diventa inevitabilmente multivariata e può essere formalizzata come segue:

$$Y_t = c_2 + \sum_{k=1}^K O_{k,t} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.2)$$

Nell'espressione (1.2), il termine  $O_{k,t}$  rappresenta il  $k$ -esimo outlier o valore irregolare eccezionale che altera l'andamento della serie osservata e la cui causa può essere nota oppure no. Per il trattamento dell'impatto del  $k$ -esimo outlier  $O_{k,t}$  si è fatto ricorso al seguente modello:

$$O_{k,t} = \omega_k \cdot L_k(B) \cdot I_t(T_k) \quad (1.3)$$

dove  $\omega_k$  misura l'intensità dell'impatto iniziale del  $k$ -esimo outlier, mentre  $I_t(T_k)$  è la variabile indicatore che ne segnala la data di accadimento  $T_k$ , ed è tale per cui:

$$I_t(T_k) \begin{cases} = 1, & \text{per } t = T_k; \\ = 0, & \text{per } t \neq T_k \text{ e } t = 1, 2, \dots, n. \end{cases} \quad (1.4)$$

Infine,  $L_k(B)$  rappresenta la modalità di trasmissione dinamica dell'effetto in esame, la quale può essere formalizzata nel modo seguente:

$$L_k(B) = \frac{1}{1 - \delta_k B^{s_k}} \quad (1.5)$$

La precedente espressione assume forme diverse a seconda dei valori assegnati a  $\delta_k$  e a  $s_k$ , e comprende tutti i casi di interesse per le serie del turismo allo studio. Nello specifico:

a) se  $\delta_k = 0$ , allora  $L_k(B) = 1$ ; ciò significa che in  $t = T_k$  siamo in presenza di una anomalia di tipo additivo (Additive Outlier, AO);

b) se  $0 < \delta_k < 1$  e  $s_k = 1$ , allora  $L_k(B) = 1/(1 - \delta_k B)$ , cioè a partire da  $t = T_k$  si ha una variazione temporanea del livello di  $Y_t$ , che si smorza nel tempo al tasso  $\delta_k$  (Transient Change, TC);

c) per  $\delta_k = s_k = 1$  si ha che  $L_k(B) = 1/(1 - B)$ ; quindi, da  $t = T_k$  in poi vi è un cambiamento permanente del livello di  $Y_t$ , la cui intensità rimane costante nel tempo in corrispondenza di  $t = T_k, T_k + 1, T_k + 2, \dots$  (Level Shift, LS);

d) per  $\delta_k = 1$  e  $s_k = 12$  si ha che  $L_k(B) = 1/(1 - B^{12})$ ; in questo caso vi è una variazione permanente stagionale del livello di  $Y_t$ , la cui intensità rimane costante per lo stesso mese di anni diversi, cioè per  $t = T_k, T_k + 12, T_k + 24, \dots$  (Seasonal Level Shift, SLS);

e) da ultimo si considera il cosiddetto Innovational Outlier (IO), il cui modello appartiene alla classe SARIMA specificata sopra e che, a seconda dei casi, può approssimare l'effetto di ciascuno dei precedenti tipi di valori anomali considerati (Chen e Liu, 1993, p. 16).

### 1.3.2 Modelli dinamici di regressione multipla DR (Dynamic Regression)

Nei modelli dinamici di regressione multipla utilizzati, il trend è rappresentato da una funzione polinomiale del tempo  $t$ , gli effetti stagionali da variabili dummy e il residuo da una struttura SARIMA. In particolare, nel caso di serie con trend lineare la scelta dell'interpolante è ovviamente caduta sulla retta; nel caso invece di serie con trend non lineare, si è fatto ricorso alla parabola, in virtù del buon adattamento ai dati mostrato da tale funzione quadratica. Per quanto riguarda poi le variabili stagionali dummy, esse assumono una forma diversa a seconda che la stagionalità presenti ampiezza costante oppure variabile.

Il vantaggio di operare con un trend funzione del tempo  $t$  e con una stagionalità funzione di variabili dummy è che, in entrambi i casi, i valori di tali predittori sono noti anche per il futuro e quindi, non dovendo essere previsti, non introducono ulteriori elementi di incertezza nelle previsioni. Infine, nella seconda versione dei modelli, oltre alle suddette componenti si è tenuto conto anche degli outliers, cioè dei valori anomali o eccezionali che possono perturbare l'andamento delle serie mensili delle presenze turistiche.

Nel seguito, vengono formalizzati in successione due dei quattro modelli che si possono ottenere combinando tra loro i due diversi tipi di trend (lineare e quadratico) con i due diversi tipi di stagionalità considerati (costante e variabile); si tratta del modello DR con trend lineare e stagionalità costante (1.6) e del modello DR con trend quadratico e stagionalità variabile (1.7)<sup>2</sup>, dapprima nella versione senza aggiustamento degli outliers e poi in quella con il trattamento dei valori anomali:

<sup>2</sup> Oltre ai modelli (1.6) e (1.7), solo il modello con trend non lineare e stagionalità costante ha mostrato un buon adattamento ad almeno una delle serie in esame. Ciò non vale per il restante modello, con trend lineare e stagionalità variabile, per cui quest'ultimo non verrà utilizzato nel seguito, a differenza degli altri tre.

$$Y_t = \sum_{i=0}^1 \alpha_i \cdot t^i + \sum_{j=1}^{12} \beta_j \cdot D_{j,t}^c + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.6)$$

$$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{12} \beta_j \cdot D_{j,t}^v + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.7)$$

$$Y_t = \sum_{i=0}^1 \alpha_i \cdot t^i + \sum_{j=1}^{12} \beta_j \cdot D_{j,t}^c + \sum_{k=1}^K O_{k,t} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.8)$$

$$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{12} \beta_j \cdot D_{j,t}^v + \sum_{k=1}^K O_{k,t} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.9)$$

dove  $D_{j,t}^c$  e  $D_{j,t}^v$  sono due variabili stagionali dummy tali per cui, fissato il mese  $j$ , per  $j = 1, 2, \dots, 12$ :

$$D_{j,t}^c \begin{cases} = 1, & \text{per il mese } t = j + 12 \cdot \gamma, \text{ con } \gamma = 0, 1, 2, \dots \text{ (anni);} \\ = 0, & \text{per gli altri mesi di ciascun anno.} \end{cases}$$

$$D_{j,t}^v \begin{cases} = 1 + \gamma, & \text{per il mese } t = j + 12 \cdot \gamma, \text{ con } \gamma = 0, 1, 2, \dots \text{ (anni);} \\ = 0, & \text{per gli altri mesi di ciascun anno.} \end{cases}$$

Nel primo caso, la costanza della stagionalità è catturata dal fatto che l'indicatore di ciascun mese rimane sempre pari a 1 per ogni anno considerato; nel secondo caso, invece, la stagionalità variabile viene colta dal fatto che l'indicatore di ciascun mese aumenta di 1 unità all'anno, a partire dal secondo. In fase di stima, per evitare perfetta multicollinearità tra i regressori  $D_{j,t}^c$  e  $D_{j,t}^v$ , si è imposto il vincolo  $\beta_{12} = 0$ ; in tal modo il coefficiente  $\beta_j$  misura l'effetto del  $j$ -esimo mese, per  $j = 1, 2, \dots, 11$ , rispetto a dicembre.

### 1.3.3 Modelli TF con funzioni razionali di trasferimento

In generale, un modello razionale TF (Transfer Function), che rappresenta la relazione tra la variabile risposta (serie dipendente)  $Y_t$  e un insieme di predittori (serie esplicative)  $X_{1t}, X_{2t}, \dots, X_{mt}$ , assume la seguente forma:

$$Y_t = c_3 + \sum_{i=1}^m \frac{U_i}{L_t} \cdot B^{b_i} \cdot X_{it} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.10)$$

dove, per  $i = 1, 2, \dots, m$ ,  $X_{it}$  è l' $i$ -esima variabile esplicativa considerata;  $B^{b_i}$  è l'operatore di ritardo tale per cui  $B^{b_i} \cdot X_{it} = X_{it-b_i}$ ; infine,  $U_i = (\omega_{i0} - \omega_{i1}B - \dots - \omega_{iu}B^u) \cdot (1 - \Omega_{i1}B^{12} - \dots - \Omega_{iv}B^{12v})$  e  $L_t = (1 - \delta_{i1}B - \dots - \delta_{iv}B^v) \cdot (1 - \Delta_{i1}B^{12} - \dots)$  sono, rispettivamente, il numeratore e il denominatore della funzione razionale di trasferimento degli effetti dei predittori  $X_{it}$  sulla variabile dipendente  $Y_t$ .

Tenendo conto degli outliers il modello TF assume la seguente forma:

$$Y_t = c_4 + \sum_{i=1}^m \frac{U_i}{L_t} \cdot B^{b_i} \cdot X_{it} + \sum_{k=1}^K O_{k,t} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.11)$$

Come si è detto in precedenza, l'utilizzo a fini previsivi dei modelli TF presenta un serio inconveniente legato al fatto che per ottenere le previsioni della variabile dipendente  $Y_t$  è preliminarmente necessario prevedere i valori futuri dell'insieme dei predittori  $X_{it}$ , il che comporta generalmente una perdita di accuratezza delle previsioni stesse. Un modo per ovviare a questo limite è quello di ricorrere a variabili il cui valore sia noto non solo per il passato ma anche per il futuro, come ad esempio il tempo  $t$  e/o variabili stagionali dummy. Un'altra possibilità consiste nell'utilizzare serie affini più aggiornate per prevedere le serie di interesse meno aggiornate, come si è fatto impiegando i modelli TF. Con questi ultimi, infatti, per prevedere le serie mensili delle presenze turistiche di fonte ISTAT, si è fatto ricorso alle corrispondenti serie mensili dei pernottamenti di fonte Banca d'Italia, note con almeno 6 mesi di anticipo rispetto alle prime<sup>3</sup>.

Operando con una sola variabile esplicativa, i precedenti modelli generali TF si riducono ai seguenti due, il primo dei quali non incorpora il trattamento degli outliers mentre il secondo sì:

<sup>3</sup> I dati di fonte ISTAT e Banca d'Italia presentano un'omogeneità accettabile per il settore alberghiero e molto scarsa per il settore extralberghiero. Nel seguito, pertanto, l'applicazione dei modelli TF riguarderà solo le presenze e i pernottamenti negli alberghi e non quelli effettuati negli esercizi complementari e totali.

$$Y_t = c_5 + \frac{U}{L} \cdot B^b \cdot X_t + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.12)$$

$$Y_t = c_6 + \frac{U}{L} \cdot B^b \cdot X_t + \sum_{k=1}^K O_{k,t} + \frac{\theta(B) \cdot \theta_{12}(B^{12})}{\Delta \cdot \phi(B) \cdot \phi_{12}(B^{12})} \cdot a_t \quad (1.13)$$

dove  $X_t$  rappresenta la serie anticipatrice dei pernottamenti alberghieri rilevati dalla Banca d'Italia, utile per prevedere la serie  $Y_t$  delle presenze alberghiere rilevate dall'ISTAT, nota con un ritardo di almeno 6 mesi rispetto alla prima.

### 1.3.4 Modelli di previsione identificati per le serie turistiche mensili del 2001-2011 e del 2007-2011

Per ciascuna serie in esame si è cercato di specificare il modello di previsione ottimale ricorrendo a una procedura iterativa e interattiva di identificazione, stima e verifica sulla falsariga di quella proposta da Box e Jenkins (1970), utilizzando a tal fine il software IBM SPSS Statistics, versione 21. In particolare, per le serie delle presenze alberghiere i modelli specificati sono stati tre, uno per ciascuno dei diversi tipi di modelli formalizzati nei paragrafi 1.3.1-1.3.3; al contrario, nel caso delle presenze complementari e totali i modelli identificati si riducono a due, non essendosi potuto specificare, per quanto detto in nota 3, il modello TF con variabile anticipatrice.

Nel seguito si riportano solo i modelli di previsione specificati per le presenze turistiche russe, senza trattamento degli outliers (s.t.o.), distinguendoli secondo l'intervallo di osservazione sul quale sono stati identificati: l'undicennio 2001-2011 da una parte (Tabella 1.1) e il quinquennio 2007-2011 dall'altra (Tabella 1.2). Per tutti gli altri casi considerati (senza e con aggiustamento degli outliers) si è proceduto nello stesso modo, ma i relativi modelli non vengono riportati per motivi di spazio<sup>4</sup>.

### 1.3.5 Indici di accuratezza delle previsioni

Per valutare l'accuratezza delle previsioni sembra opportuno ricorrere a indici sintetici espressi in termini relativi; questi ultimi infatti, essendo dei puri numeri, rendono

<sup>4</sup> I modelli di cui sopra, disponibili presso l'Autore, riguardano le presenze mensili alberghiere, complementari e totali, dei periodi 2001-2011 e 2007-2011, relative sia ai russi (con trattamento degli outliers) sia ai turisti tedeschi, stranieri e italiani (senza e con aggiustamento degli outliers). I modelli prescelti soddisfano, a parità di numero di parametri, sia il criterio di casualità dei residui, secondo i relativi correlogrammi e il test di Ljung-Box (1979), sia il criterio di massimizzare la bontà di adattamento alla serie osservata, in base al coefficiente di variazione dei residui, all'indice BIC (Bayesian Information Criterion) di Akaike (1978) e all'indice di determinazione corretto.

**Tabella 1.1** Modelli utilizzati per prevedere le serie mensili 2001-2011 delle presenze russe alberghiere, complementari e totali (modelli senza trattamento degli outliers, s.t.o., di tipo: SARIMA, dinamico di regressione multipla DR e con funzioni razionali di trasferimento TF)

	Modello per le serie mensili 2001-2011 delle presenze russe alberghiere
SARIMA s.t.o.	$(1 - B^{12}) \cdot Y_t = c_1 + \frac{(1 - \theta_3 B^3)}{(1 - \phi_1 B) \cdot (1 - \phi_{24} B^{24})} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{1}{(1 - \phi_1 B) \cdot (1 - \phi_{12} B^{12})} \cdot a_t$
Modello TF s.t.o.	$Y_t = c_5 + \frac{\omega_0}{(1 - \delta_1 B)} \cdot X_t + \frac{(1 - \theta_3 B^3 - \theta_7 B^7)}{(1 - \phi_1 B) \cdot (1 - B^{12})} \cdot a_t$
	Modello per le serie mensili 2001-11 delle presenze russe complementari
SARIMA s.t.o.	$(1 - B^{12}) \cdot Y_t = c_1 + \frac{(1 - \theta_{13} B^{13})}{(1 - \phi_1 B) \cdot (1 - \phi_{24} B^{24})} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{(1 - \theta_{13} B^{13})}{(1 - \phi_1 B) \cdot (1 - \phi_{12} B^{12})} \cdot a_t$
	Modello per le serie mensili 2001-2011 delle presenze russe totali
SARIMA s.t.o.	$(1 - B) \cdot (1 - B^{12}) \cdot Y_t = \frac{(1 - \theta_3 B^3)}{(1 - \phi_{24} B^{24} - \phi_{48} B^{48})} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{(1 - \theta_{16} B^{16})}{(1 - \phi_1 B) \cdot (1 - \phi_{12} B^{12})} \cdot a_t$

Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2002 e 2004), (2005-2011) e (2012-2014).

**Tabella 1.2** Modelli utilizzati per prevedere le serie mensili 2007-2011 delle presenze russe alberghiere, complementari e totali (modelli senza trattamento degli outliers, s.t.o., di tipo: SARIMA, dinamico di regressione multipla DR e con funzioni razionali di trasferimento TF)

	Modello per le serie mensili 2007-2011 delle presenze russe alberghiere
SARIMA s.t.o.	$(1-B) \cdot (1-B^{12}) \cdot Y_t = \frac{1}{(1-\phi_3 B^3)} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{1}{(1-\phi_1 B - \phi_2 B^2)} \cdot a_t$
Modello TF s.t.o.	$Y_t = \frac{\omega_0}{(1-\delta_{12} B^{12})} \cdot X_t + \frac{1}{(1-\phi_1 B) \cdot (1-B^{12})} \cdot a_t$
	Modello per le serie mensili 2007-11 delle presenze russe complementari
SARIMA s.t.o.	$(1-B^{12}) \cdot Y_t = c_1 + \frac{(1-\theta_{13} B^{13})}{(1-\phi_1 B)} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{(1-\theta_{13} B^{13})}{(1-\phi_1 B - \phi_3 B^3)} \cdot a_t$
	Modello per le serie mensili 2007-2011 delle presenze russe totali
SARIMA s.t.o.	$(1-B) \cdot (1-B^{12}) \cdot Y_t = \frac{1}{(1-\phi_3 B^3)} \cdot a_t$
Modello DR s.t.o.	$Y_t = \sum_{i=0}^2 \alpha_i \cdot t^i + \sum_{j=1}^{11} \beta_j \cdot D_{j,t}^v + \frac{1}{(1-\phi_1 B - \phi_2 B^2)} \cdot a_t$

Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2005-2011) e (2012-2014).

possibile il confronto tra situazioni diverse (obiettivo prioritario di questo contributo). L'indice relativo che, oltre ad essere facilmente interpretabile, sembra il più adeguato allo scopo è l'indice MAPE (Mean Absolute Percentage Error). Oltre a questo indice, in letteratura ne vengono proposti altri, tra cui l'indice RMSE (Root Mean Squared Error), che però non si presta ad essere utilizzato per effettuare confronti tra situazioni diverse. Il modo qui adottato per ovviare a questo inconveniente è quello di rapportare l'indice RMSE alla corrispondente media dei valori osservati, ottenendo così una misura relativa utile anche per i confronti. Tale indice, data la sua struttura, può essere denominato coefficiente di variazione CV dell'errore di previsione.

In formula, i due suddetti indici relativi sono dati rispettivamente da:

$$MAPE = \frac{1}{12} \cdot \sum_{j=1}^{12} \left| \frac{Y_{n+j} - \hat{Y}_{n+j}}{Y_{n+j}} \right| \cdot 100$$

dove  $Y_{n+j}$  è il valore osservato nel mese  $n+j$  e  $\hat{Y}_{n+j}$  è la corrispondente previsione di origine  $n$  e orizzonte previsivo  $j$ , per  $j = 1, 2, \dots, 12$  mesi.

$$CV = \frac{RMSE}{MEAN} \cdot 100$$

dove  $RMSE = \sqrt{\frac{1}{12} \cdot \sum_{j=1}^{12} (Y_{n+j} - \hat{Y}_{n+j})^2}$  e  $MEAN = \frac{1}{12} \cdot \sum_{j=1}^{12} Y_{n+j}$ . La denominazione di coefficiente di variazione data a questo indice deriva dal fatto che la sua struttura richiama quella del rapporto tra lo scarto quadratico medio e la media dei valori, noto appunto come coefficiente di variazione.

## 1.4 Valutazione delle previsioni della domanda turistica mensile

Per poter valutare l'accuratezza previsiva si è operato con previsioni ex-post, stimando i diversi modelli fino al dicembre 2011 e riservando gli ultimi 12 dati definitivi disponibili per il confronto con i corrispondenti valori previsti. Pertanto, l'origine  $n$  delle previsioni coincide con il mese di dicembre 2011 e l'orizzonte previsivo  $j = 1, 2, \dots, 12$  va dal gennaio al dicembre 2012 (ultimo anno con dati definitivi al momento delle elaborazioni). Inoltre, volendo verificare se e in che modo le previsioni variano in funzione dell'intervallo di stima, tutti i modelli sono stati identificati e stimati su due intervalli di diversa ampiezza: il primo, di 132 osservazioni, va dal gennaio 2001 al dicembre 2011 e il secondo, di 60 osservazioni, va dal gennaio 2007 al dicembre 2011. Nella Tabella 1.3 sono riportati i valori degli indici di accuratezza delle previ-

sioni (MAPE e CV)<sup>5</sup> relativi alle serie mensili delle presenze turistiche russe. Questi indici sono stati calcolati anche per le altre nazionalità considerate a titolo di confronto (presenze tedesche<sup>6</sup>, straniere<sup>7</sup> e italiane), ma di essi vengono riportate solo le differenze dell'indice MAPE rispetto ai corrispondenti valori delle presenze russe (Tabella 1.5). In ogni caso, gli indici sono stati calcolati distintamente per le presenze alberghiere, complementari e totali, come pure per ognuno dei diversi tipi di modelli considerati (SARIMA, DR e TF), sia nella versione senza trattamento degli outliers (in sigla s.t.o.) sia in quella con aggiustamento degli outliers (in sigla c.t.o.). Inoltre, al fine di verificare se e come le previsioni siano influenzate dall'ampiezza dell'intervallo di stima, tutti i calcoli sono stati replicati sui due diversi intervalli indicati sopra. Infine, per avere informazioni anche sulla difformità tra la capacità di previsione e la bontà di stima dei modelli, valutate con gli indici MAPE e CV, la Tabella 1.4 contiene le differenze tra i valori relativi al periodo di previsione e quelli riguardanti il corrispondente intervallo di stima.

Nel seguito si valuterà anzitutto l'accuratezza delle previsioni ottenute nel caso delle presenze turistiche russe, passando poi al confronto tra tale accuratezza e la bontà delle corrispondenti stime. Da ultimo, si analizzeranno le differenze tra i risultati previsionali riguardanti le presenze russe e quelli relativi alle altre nazionalità di rilievo considerate.

#### 1.4.1 Accuratezza dei previsori delle presenze turistiche russe

Nel caso delle presenze russe totali 2001-2011, il previsore migliore è quello associato al modello SARIMA s.t.o., che dà luogo a un valore di MAPE e di CV pari, rispettivamente, al 4,07% e al 5,54% (Tabella 1.3). Si tratta di valori piuttosto bassi, simili o inferiori a quelli di altre applicazioni (Athanasopoulos et al., 2011; Tonini, 1994), che confermano la capacità previsiva dei modelli SARIMA riscontrata nei lavori citati. Questa capacità si manifesta anche con riferimento ai due sottoinsiemi delle presenze alberghiere e complementari, seppur in misura diversa. Nell'ultimo caso, infatti, i valori degli indici sono sensibilmente superiori a quelli dell'altro settore ricettivo, in conseguenza della maggior variabilità che caratterizza il segmento complementare rispetto a quello alberghiero, come documenta il coefficiente di variazione delle relative presenze pari, rispettivamente, al 101,23% e al 68,78%.

---

<sup>5</sup> Data la loro natura di sintesi degli errori di previsione, gli indici MAPE e CV variano in funzione diretta di tali errori. Di conseguenza, quanto minore è il valore di detti indici, tanto maggiore è l'accuratezza delle previsioni, e viceversa.

<sup>6</sup> Nel caso dei tedeschi ci si riferisce non a tutti i turisti tedescofoni ma solo a quelli provenienti dalla Germania che, comunque, costituisce il principale Paese di origine delle presenze turistiche straniere che soggiornano in Italia. Nel seguito, pertanto, "tedesco" è da intendere come "germanico".

<sup>7</sup> Le presenze straniere considerate riguardano i pernottamenti mensili di tutti i turisti stranieri che hanno soggiornato in Italia, compresi i tedeschi. Questa scelta consente di operare adeguati confronti con il sottoinsieme complementare delle presenze mensili "non straniere", cioè italiane, rilevate dall'ISTAT.

**Tabella 1.3** Indici di accuratezza relativa delle previsioni dal gennaio al dicembre 2012. Serie mensili 2001-2011 e 2007-2011 delle presenze russe alberghiere, complementari e totali

Modello	Presenze russe 2001-2011			Presenze russe 2007-2011		
	alberghi.	complem.	totali	alberghi.	complem.	totali
	Indice di accuratezza MAPE (valori percentuali)					
SARIMA s.t.o.	4,49	6,79	4,07	6,02	11,52	6,46
Mod. DR s.t.o.	4,79	12,97	6,07	11,36	32,91	14,15
Mod. TF s.t.o.	5,23	—	—	5,27	—	—
SARIMA c.t.o.	6,23	15,74	9,93	7,97	14,78	6,46
Mod. DR c.t.o.	9,24	12,86	10,44	11,36	17,51	14,15
Mod. TF c.t.o.	4,63	—	—	4,48	—	—
	Indice di accuratezza CV (valori percentuali)					
SARIMA s.t.o.	4,93	9,37	5,54	9,22	12,11	9,40
Mod. DR s.t.o.	8,42	11,17	8,61	10,84	21,64	12,19
Mod. TF s.t.o.	6,23	—	—	5,17	—	—
SARIMA c.t.o.	7,42	38,88	10,12	10,78	28,90	9,40
Mod. DR c.t.o.	10,66	31,11	10,69	10,84	35,56	12,19
Mod. TF c.t.o.	6,28	—	—	5,84	—	—

Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2002 e 2004), (2005-2011) e (2012-2014).

Passando poi ai confronti tra i modelli senza trattamento degli outliers (s.t.o.) e quelli con tale trattamento (c.t.o.), si può notare che i primi presentano generalmente una capacità previsiva superiore a quella dei secondi, soprattutto secondo l'indice CV. Infatti, con qualche eccezione per l'indice MAPE, sia quest'ultimo che l'indice CV danno luogo a valori generalmente più bassi nei casi in cui non si aggiustano gli outliers rispetto a quelli in cui li si aggiusta. Siccome questo risultato è diverso da quello riscontrato altrove (Tonini, 1994), per individuarne le cause si è estesa la stima dei modelli anche al 2012, effettuando una ricerca degli outliers riferita pure a tale anno di "previsione". Questa ricerca ha portato alla scoperta di outliers significativi in corrispondenza del mese di luglio 2012, nel caso delle presenze russe alberghiere, e dei mesi di giugno e luglio 2012, nel caso delle presenze russe complementari e totali, in

analogia con quanto si registra nel 2011. Nello specifico, il fatto che i modelli s.t.o., a differenza di quelli c.t.o., non eliminino gli outliers dell'ultimo anno di osservazione (2011), consente loro di intercettare gli outliers del 2012 e, quindi, di dar luogo a previsioni più accurate, almeno nel caso dei modelli SARIMA e DR relativi al 2001-2011.

Un'eccezione parziale ai precedenti risultati riguarda le previsioni da modelli TF, la cui accuratezza, almeno secondo l'indice MAPE, aumenta passando dai modelli che non effettuano il trattamento degli outliers a quelli che invece lo effettuano. Ciò dipende dal fatto che, in presenza di variabili esplicative con outliers negli stessi mesi della variabile dipendente, è possibile tener conto meglio degli effetti dei valori anomali, con ricadute positive anche sulle previsioni.

Basare le previsioni solo sugli anni più recenti non sembra migliorarne sempre l'accuratezza; infatti, almeno nel caso dei modelli senza trattamento degli outliers, i valori degli indici MAPE e CV crescono sistematicamente passando dalle previsioni basate su 11 anni (2001-2011) a quelle basate sugli ultimi 5 anni (2007-2011), rivelando in questo secondo caso una minore capacità previsiva dei modelli. Tale ridotta capacità sembra legata, in particolare, al fatto che la frequenza relativa dei valori anomali, che alterano l'andamento delle serie, è nettamente più elevata nel quinquennio 2007-2011 che nell'undicennio 2001-2011<sup>8</sup>. Una riprova, seppur parziale, del suddetto risultato si ha considerando i modelli che effettuano il trattamento degli outliers i quali, almeno nel caso dei modelli TF e SARIMA, danno luogo a previsioni generalmente più accurate quando si basano sull'intervallo 2007-2011 anziché sul 2001-2011<sup>9</sup>.

### 1.4.2 Confronti tra l'accuratezza delle previsioni e la bontà delle stime

Il confronto tra gli indici di accuratezza delle previsioni e quelli delle stime risulta piuttosto istruttivo perché rivela le differenze tra la variabilità degli errori di previsione (con orizzonte da 1 a 12 mesi) e la variabilità degli errori di stima<sup>10</sup>.

Per confrontare l'accuratezza delle previsioni con la bontà delle stime si sono calcolate le differenze tra i valori dell'indice MAPE basati sugli errori di previsione e quelli basati sugli errori di stima. Lo stesso si è fatto anche per l'indice CV. Un segno negativo delle differenze significa che l'accuratezza delle previsioni è maggiore della bontà delle stime, in quanto l'indice che sintetizza gli errori di previsione è minore di

<sup>8</sup> Nel caso ad esempio delle presenze russe totali, passando dall'intervallo di osservazione più lungo (2001-2011) a quello più breve (2007-2011), la percentuale di outliers identificati quasi raddoppia, essendo rispettivamente pari al 4,55% e all'8,33% del numero di osservazioni.

<sup>9</sup> La difformità di risultati forniti dai modelli DR rispetto agli altri modelli utilizzati merita un supplemento di indagine che sarà oggetto di un prossimo contributo.

<sup>10</sup> Secondo la teoria delle previsioni, il processo *White Noise*  $a_t$  che "genera"  $Y_t$  si può anche interpretare come una successione di errori di previsione con orizzonte 1. Vi è pertanto uno stretto legame tra le fasi di stima e di previsione; in esse, infatti, si cerca di minimizzare, rispettivamente, la varianza del *WN* e la varianza dell'errore di previsione (Piccolo, 1990, pp. 241-242).

quello che sintetizza gli errori di stima. L'interpretazione è ovviamente opposta nel caso di segno positivo. Inoltre, quanto maggiore è in modulo il valore delle differenze, tanto più forte è la discrepanza tra accuratezza di previsione e bontà di stima.

Con riferimento agli indici MAPE e CV, nel caso dei modelli senza trattamento degli outliers l'accuratezza delle previsioni è generalmente maggiore di quella delle stime, essendo negativo il segno delle corrispondenti differenze degli indici; l'unica eccezione riguarda i modelli DR s.t.o. relativi al quinquennio 2007-2011 che, registrando segno positivo delle differenze, presentano valori di MAPE e di CV delle stime inferiori a quelli delle previsioni (Tabella 1.4).

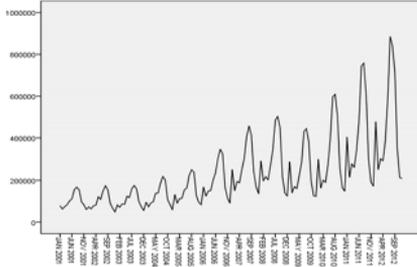
Il precedente risultato generale rispecchia il fatto che, nel corso dell'intervallo 2001-2011, le serie delle presenze turistiche russe hanno sperimentato una variabilità molto forte, con un trend decisamente accelerato e una stagionalità di ampiezza nettamente crescente nel tempo (vedasi Figura 1.1).

**Tabella 1.4** Differenze tra gli indici di accuratezza delle previsioni e quelli delle stime. Serie mensili 2001-2011 e 2007-2011 delle presenze russe alberghiere, complementari e totali

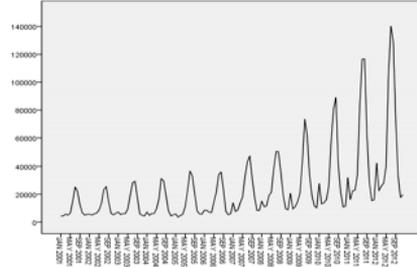
Modello	Presenze russe 2001-2011			Presenze russe 2007-2011		
	alberghi.	complem.	totali	alberghi.	complem.	totali
	MAPE previsioni – MAPE stime					
SARIMA s.t.o.	-3,22	-9,74	-3,25	-3,19	-4,92	-2,76
Mod. DR s.t.o.	-2,99	-3,28	-1,50	6,59	20,02	9,20
Mod. TF s.t.o.	-2,76	—	—	-1,68	—	—
SARIMA c.t.o.	-0,31	0,95	3,67	0,77	2,78	-2,76
Mod. DR c.t.o.	3,52	-1,56	4,84	6,59	8,18	9,20
Mod. TF c.t.o.	-2,20	—	—	-3,18	—	—
	CV previsioni – CV stime					
SARIMA s.t.o.	-5,17	-10,83	-3,80	-0,67	-6,98	-0,71
Mod. DR s.t.o.	-,14	-6,50	0,22	3,94	6,45	4,97
Mod. TF s.t.o.	-3,26	—	—	-2,78	—	—
SARIMA c.t.o.	0,28	26,75	2,98	2,28	18,47	-0,71
Mod. DR c.t.o.	4,36	17,96	4,49	3,94	25,46	4,97
Mod. TF c.t.o.	-1,53	—	—	-2,76	—	—

Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2002 e 2004), (2005-2011) e (2012-2014).

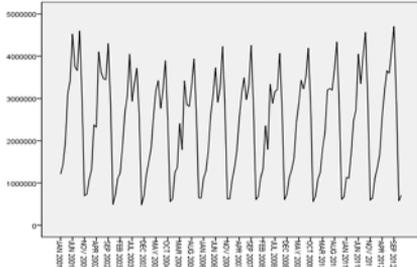
Presenze russe alberghiere 2001-2012



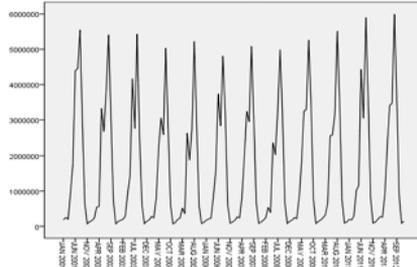
Presenze russe complementari 2001-12



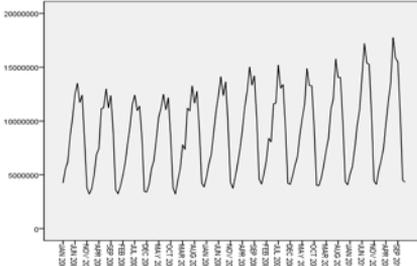
Presenze tedesche alberghiere 2001-12



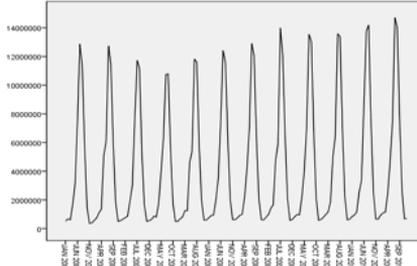
Presenze tedesche complement. 2001-12



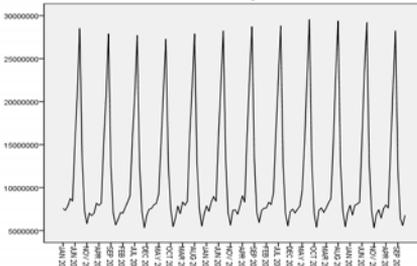
Presenze straniere alberghiere 2001-12



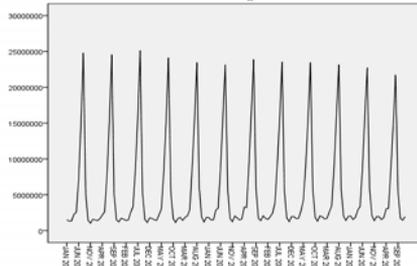
Presenze straniere complement. 2001-12



Presenze italiane alberghiere 2001-2012



Presenze italiane complement. 2001-12



**Figura 1.1** Presenze mensili russe, tedesche, straniere e italiane. Gennaio 2001-dicembre 2012. Ciò ha fatto sì che la stima dei modelli fosse più difficoltosa e meno precisa delle previsioni di orizzonte da 1 a 12 mesi.

Nel caso delle stime, a differenza che in quello delle previsioni, l'aggiustamento degli outliers produce un miglioramento anziché un peggioramento dei risultati; infatti, passando dai modelli s.t.o. a quelli c.t.o. il segno delle differenze degli indici MAPE e CV da negativo diventa spesso positivo, segnalando così un'accuratezza delle stime superiore a quella delle previsioni. Ciò dipende dal fatto che nel corso del processo di stima il trattamento degli outliers è continuo, il che avvicina i valori stimati a quelli osservati; viceversa, durante l'intervallo di previsione non si può ovviamente procedere alla stima degli outliers, per cui l'effetto negativo di questi ultimi sull'errore di previsione è in generale maggiore del loro effetto sull'errore di stima.

Le differenze tra l'accuratezza relativa delle previsioni e quella delle stime variano al variare dell'orizzonte di osservazione considerato. Infatti, quando quest'ultimo è il 2001-2011, il segno delle differenze è prevalentemente negativo, almeno secondo l'indice MAPE, segnalando una maggiore accuratezza relativa delle previsioni rispetto alle stime. Viceversa, quando l'intervallo è più breve (2007-2011), vi è una maggior presenza del segno positivo, in quanto le stime superano più frequentemente in accuratezza le previsioni. Questi risultati dipendono dall'andamento temporale contrapposto dell'accuratezza delle previsioni rispetto alla bontà delle stime; in generale, infatti, passando dal 2001-2011 al 2007-2011, l'accuratezza previsiva peggiora mentre la bontà delle stime migliora, anche per effetto della minor variabilità relativa che caratterizza la serie più breve rispetto a quella più lunga, come documenta il coefficiente di variazione delle presenze russe totali pari, rispettivamente, al 54,11% e al 71,10%.

### **1.4.3 L'accuratezza delle previsioni dei turisti russi rispetto a quella di altre nazionalità**

Ulteriori elementi di analisi sono dati dal confronto tra l'accuratezza delle previsioni delle presenze russe e quella relativa ad altre nazionalità, costituite dalle presenze tedesche, straniere e italiane. Si è voluto cioè raffrontare la domanda russa con i principali sottoinsiemi del mercato turistico nazionale, rappresentati appunto dai turisti italiani, stranieri e tedeschi (tradizionale segmento di punta del turismo straniero in Italia).

Il confronto si è basato sulle differenze tra gli indici MAPE di accuratezza delle previsioni dei russi e quelli delle previsioni rispettivamente dei tedeschi, degli stranieri e degli italiani. Un segno negativo di tali differenze significa che le previsioni delle presenze russe sono più accurate di quelle della nazionalità posta a confronto, in quanto i rispettivi indici MAPE sono il primo inferiore al secondo. Un'interpretazione opposta si ha ovviamente nel caso di segno positivo delle differenze.

Dai risultati riportati nella Tabella 1.5 emerge che, nel caso dei russi, l'indice MAPE assume valori quasi sempre inferiori a quelli dei tedeschi, evidenziando così che le previsioni delle presenze russe sono in generale migliori, e in diversi casi nettamente migliori, rispetto alle previsioni delle presenze tedesche.

**Tabella 1.5** Differenze tra russi e altre nazionalità dell'indice MAPE di accuratezza delle previsioni dal gennaio al dicembre 2012. Serie mensili 2001-2011 e 2007-2011 delle presenze alberghiere, complementari e totali

Modello	Presenze 2001-2011			Presenze 2007-2011		
	alberghi.	complem.	totali	alberghi.	complem.	totali
	MAPE russi - MAPE tedeschi					
SARIMA s.t.o.	-2,94	-8,72	-7,60	-0,90	-16,81	-2,15
Mod. DR s.t.o.	-5,87	-54,58	-13,28	1,33	-37,17	-5,21
Mod. TF s.t.o.	-3,03	—	—	-1,81	—	—
SARIMA c.t.o.	-1,20	-21,53	-4,48	-0,78	-28,03	-0,65
Mod. DR c.t.o.	-1,42	-66,30	-2,61	-1,17	-25,68	1,56
Mod. TF c.t.o.	-10,24	—	—	-2,60	—	—
	MAPE russi - MAPE stranieri					
SARIMA s.t.o.	1,08	-0,43	0,55	2,40	4,39	1,90
Mod. DR s.t.o.	-0,46	-2,15	1,16	5,78	25,17	9,49
Mod. TF s.t.o.	1,36	—	—	0,23	—	—
SARIMA c.t.o.	-1,89	6,27	2,89	0,59	7,36	1,14
Mod. DR c.t.o.	2,03	0,87	4,96	4,08	-4,58	7,72
Mod. TF c.t.o.	0,76	—	—	-2,81	—	—
	MAPE russi - MAPE italiani					
SARIMA s.t.o.	-1,32	0,98	-1,60	-0,36	5,30	0,19
Mod. DR s.t.o.	-2,43	5,87	-1,68	5,94	26,61	8,34

SARIMA c.t.o.	0,42	10,01	4,26	1,59	9,10	0,27
Mod. DR c.t.o.	2,36	-3,79	3,24	5,89	11,76	8,29

Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2002 e 2004), (2005-2011) e (2012-2014).

Ciò si verifica indipendentemente sia dal tipo di esercizio ricettivo e intervallo temporale considerato, sia dal tipo di modello utilizzato e dal trattamento o meno degli outliers. Le uniche due eccezioni, di valore piuttosto contenuto, si hanno in corrispondenza delle previsioni basate sui dati 2007-2011, quando si utilizza il modello DR s.t.o. per prevedere le presenze alberghiere e il modello DR c.t.o. per prevedere le presenze totali.

All'origine di queste differenze a favore delle previsioni della domanda russa sta il fatto che, nonostante quest'ultima registri un trend più sostenuto e una stagionalità più accentuata, essa risulta caratterizzata da una maggiore regolarità di andamento e dall'assenza degli effetti perturbatori delle festività mobili (Pasqua e Pentecoste) che influenzano invece il segmento tedesco<sup>11</sup> (vedasi Fig. 1.1 e Tab. 1.6, dove si considerano sia gli effetti di tali festività sia il numero degli outliers).

Passando dalle presenze tedesche all'insieme delle presenze straniere, le differenze di accuratezza delle previsioni rispetto al segmento russo assumono per lo più segno positivo, soprattutto in corrispondenza dell'intervallo di osservazione più breve (2007-2011). Ciò significa che, tranne alcune eccezioni, le previsioni della domanda turistica russa sono generalmente peggiori di quelle della domanda straniera totale.

Un risultato analogo si registra anche per i confronti con il segmento dei turisti italiani. Pure in questo caso, infatti, si hanno segni prevalentemente positivi delle differenze tra gli indici MAPE dei russi e quelli degli italiani. Inoltre, come per le presenze straniere totali, anche per le presenze italiane le differenze positive sono in modulo generalmente maggiori delle differenze negative. In altri termini, le previsioni delle presenze russe risultano quasi sempre meno accurate, spesso in misura sensibile, rispetto a quelle delle presenze italiane, soprattutto quando si basano sul periodo di osservazione più recente (2007-2011).

A spiegare queste differenze a sfavore delle previsioni della domanda russa concorre certamente il fatto che tale domanda manifesta un andamento meno regolare rispetto sia ai turisti stranieri totali sia ai turisti italiani, specialmente nell'intervallo di osservazione 2007-2011, come si può evincere dalla Figura 1.1 e come documenta la

<sup>11</sup> Essendo festività mobili legate al calendario lunare, la Pasqua e la Pentecoste possono cadere, a seconda degli anni, in due mesi diversi che sono, rispettivamente, marzo e aprile ovvero maggio e giugno. Ciò fa lievitare le presenze tedesche nel mese di ricorrenza delle suddette festività, soprattutto nel caso della Pentecoste, provocando tra un anno e l'altro delle perturbazioni nell'andamento delle serie storiche. Di tali alterazioni, visibili anche nella Figura 1.1, si è tenuto conto nella Tabella 1.6, conteggiando il numero di volte in cui è cambiato il mese di ricorrenza della Pasqua e della Pentecoste nel corso sia dell'undicennio 2001-2011 sia del quinquennio 2007-2011. Si fa infine notare che il mese di ricorrenza della Pentecoste varia anche tra l'ultimo anno di osservazione (2011) e l'anno di previsione (2012), rendendo così ulteriormente difficoltose e inaccurate le previsioni delle presenze tedesche.

più elevata frequenza relativa di outliers che caratterizza il segmento russo rispetto a quello straniero e italiano<sup>12</sup>.

In definitiva, le previsioni delle presenze russe mostrano un'accuratezza soddisfacente, a volte superiore e a volte inferiore a quella di altre nazionalità, a seconda della maggiore o minore regolarità delle serie mensili russe rispetto alle altre poste a confronto e a seconda che tale regolarità si mantenga o meno anche nel futuro.

**Tabella 1.6** Numero di outliers e degli effetti di Pasqua e Pentecoste nelle serie mensili 2001-2011 e 2007-2011. Presenze alberghiere, complementari e totali relative ai turisti russi, tedeschi, stranieri e italiani

Modello	Presenze 2001-2011			Presenze 2007-2011		
	alberghi.	complem.	totali	alberghi.	complem.	totali
	Serie mensili delle presenze russe					
SARIMA c.t.o.	7	9	6	5	5	5
Mod. DR c.t.o.	6	5	5	4	4	4
Mod. TF c.t.o.	6	—	—	0	—	—
	Serie mensili delle presenze tedesche					
SARIMA c.t.o.	12	25	15	9	7	11
Mod. DR c.t.o.	12	18	17	7	8	7
Mod. TF c.t.o.	13	—	—	4	—	—
	Serie mensili delle presenze straniere					
SARIMA c.t.o.	1	4	1	2	4	2
Mod. DR c.t.o.	1	3	2	2	4	3
Mod. TF c.t.o.	0	—	—	2	—	—
	Serie mensili delle presenze italiane					
SARIMA c.t.o.	0	1	0	0	3	3
Mod. DR c.t.o.	1	3	1	1	2	3

<sup>12</sup> Con riferimento al quinquennio 2007-2011 e ai modelli SARIMA, la percentuale di outliers rilevati nel caso delle presenze totali straniere e italiane è pari, rispettivamente, al 3,33% e al 5,00% del numero di osservazioni, mentre sale all'8,33% nel caso delle presenze russe.

*Fonte: Nostre elaborazioni di dati della Banca d'Italia (2014) e dell'ISTAT (2002 e 2004), (2005-2011) e (2012-2014).*

## 1.5 Conclusioni

Con il presente lavoro si è inteso affrontare il problema di aggiornare le serie delle presenze turistiche russe in Italia utilizzando diversi modelli di previsione, sia univariati che multivariati, e applicando tali modelli a diversi settori ricettivi, intervalli temporali e nazionalità. Da questo contributo si può anzitutto concludere che, in sintonia con altri studi sull'argomento (Athanasopoulos et al., 2011; Kulendran e King, 1997; Witt e Witt, 1995), non sempre i modelli più complessi danno luogo a previsioni più accurate. Infatti, nel nostro caso, per prevedere le presenze russe in Italia i modelli che hanno funzionato generalmente meglio non sono stati quelli multivariati bensì i più semplici modelli univariati della classe SARIMA. In questo senso, i due indici relativi di accuratezza delle previsioni (MAPE e CV) forniscono informazioni coerenti, indicando nel modello SARIMA il miglior previsore delle presenze turistiche in esame, almeno quando si opera senza trattamento degli outliers (con l'unica eccezione delle presenze russe alberghiere relative all'intervallo 2007-2011). Viceversa, quando si aggiustano i valori anomali, le previsioni più accurate sono fornite dal modello multivariato TF con funzioni razionali di trasferimento, cioè quello che sfrutta la conoscenza della serie anticipatrice dei pernottamenti (di fonte Banca d'Italia) per prevedere la corrispondente serie delle presenze (di fonte ISTAT). In ogni caso, le performance peggiori sono fornite, quasi sempre, dal modello dinamico di regressione multipla DR, in corrispondenza del quale (e con poche eccezioni) gli indici MAPE e CV presentano valori superiori a quelli sia dell'altro modello multivariato in esame, sia del processo univariato SARIMA. Non sempre, dunque, i modelli che funzionano bene in fase di stima, come i modelli DR (Tonini, 2015), riescono poi ad essere altrettanto performanti in sede di previsione, a causa di alcune difficoltà incontrate dalle variabili polinomiali e dummy stagionali a catturare al meglio, rispettivamente, il trend e la stagionalità futura. Viceversa, le buone prestazioni del modello univariato SARIMA sono dovute tanto alla sua natura adattativa quanto al fatto che, per effettuare le previsioni, esso si basa soprattutto sulle ultime osservazioni, le più importanti a fini previsionali.

Passando dai modelli senza trattamento degli outliers a quelli che lo effettuano si ha in generale un peggioramento delle previsioni, tranne che nel caso del modello TF con funzioni di trasferimento, almeno secondo l'indice MAPE. Ciò dipende dal fatto che l'aggiustamento dei valori anomali dell'ultimo anno di osservazione dà luogo a previsioni migliori solo se i valori futuri non sono contaminati da outliers, come invece capita nei casi in esame.

Un ulteriore confronto di interesse è quello tra l'accuratezza delle previsioni da 1 a 12 orizzonti mensili e la bontà delle stime; a questo proposito si può notare che non sempre le stime risultano più accurate delle previsioni, specialmente quando si fa riferimento all'intervallo di osservazione più lungo (undicennio 2001-2011). Tale intervallo, infatti, è caratterizzato da una variabilità dei dati sensibilmente superiore a

quella sia dell'intervallo di previsione sia del periodo di osservazione più breve e recente (quinquennio 2007-2011), il che contribuisce a rendere le stime meno precise.

Altri aspetti rilevanti emersi da questo lavoro riguardano i risultati dei confronti sviluppati a livello intertemporale, intersettoriale e internazionale. Con riferimento ai due diversi intervalli di osservazione considerati (2001-2011 e 2007-2011) si può notare che, passando dal primo al secondo, l'accuratezza delle previsioni generalmente diminuisce, anche per effetto della maggior frequenza relativa di outliers che si registrano nell'intervallo più breve rispetto a quello più lungo, i quali rendono più difficoltosa l'operazione previsiva. Nel caso poi dei confronti intersettoriali si ha che le prestazioni previsionali peggiorano nettamente passando dal settore alberghiero a quello complementare; questo risultato non presenta eccezione alcuna, valendo indipendentemente dall'intervallo di osservazione, dal tipo di modello e dal trattamento o meno degli outliers. Anche in questo caso è la maggior variabilità delle presenze complementari rispetto a quelle alberghiere a determinare il peggioramento dei risultati previsionali segnalato sopra. Infine, dai confronti internazionali si può rilevare che le presenze russe registrano performance previsionali a volte superiori e a volte inferiori a quelle delle altre nazionalità considerate; in particolare, si hanno prestazioni migliori rispetto al segmento tedesco e peggiori rispetto sia alle presenze italiane che alle presenze straniere totali. Questi due ultimi segmenti sono caratterizzati da una maggiore regolarità di andamento sia durante il periodo di osservazione che durante l'intervallo di previsione, e ciò favorisce l'accuratezza relativa delle previsioni; viceversa, nel caso delle presenze tedesche, i valori previsti risentono negativamente della maggior quantità di outliers e di effetti perturbatori dovuti alle festività mobili della Pasqua e della Pentecoste, i quali influenzano particolarmente il turismo tedesco.

In definitiva, il problema dell'aggiornamento di serie della domanda turistica, e di quella russa in particolare, può essere adeguatamente risolto effettuando previsioni di tale domanda mediante opportuni modelli univariati o multivariati. Nello specifico, i modelli della classe SARIMA, relativamente semplici da identificare e applicare, si sono rivelati utili soprattutto nel caso in cui non si effettui il trattamento degli outliers, confermando la soddisfacente capacità previsionale dei modelli univariati già evidenziata in altri lavori (Athanasopoulos et al., 2011; Kulendran e King, 1997; Witt e Witt, 1995). Nel caso invece si proceda all'aggiustamento dei valori anomali, allora sembra preferibile ricorrere, almeno secondo l'indice MAPE, al modello TF con funzioni razionali di trasferimento, in cui si sfrutta in chiave previsionale il legame tra la serie anticipatrice dei pernottamenti e la serie delle presenze da prevedere. Questo risultato è in sintonia con quello ottenuto da alcuni Autori (Allen e Fildes, 2001; Song et al., 2003), i quali tuttavia testano su serie quasi esclusivamente annuali i diversi previsori multivariati proposti.

In conclusione, pur non essendoci piena univocità di risultati, come evidenzia anche la breve rassegna iniziale della letteratura, sembra tuttavia che il modello di previsione da preferire sia quello SARIMA, in base sia alla capacità previsionale dimostrata nel presente contesto, sia al criterio di parsimonia, secondo il quale, a parità o quasi di accuratezza delle previsioni, il modello da privilegiare è quello più semplice. In prospettiva, quindi, sarà interessante verificare, con ulteriori applicazioni, se e in

che misura la precedente conclusione sia generalizzabile anche a periodi e contesti diversi.

## Riferimenti bibliografici

- ABRAHAM B. - CHUANG A. (1989), *Outlier detection and time series modeling*, in *Technometrics*, vol. 31, n. 2, pp. 241-248.
- AKAIKE H. (1978), *A bayesian extension of the minimum AIC procedure*, *Annals of the Institute of Statistical Mathematics*, vol. 30, A, pp. 9-14.
- ALGIERI B. (2006), *An econometric estimation of the demand for tourism: The case of Russia*, in *Tourism Economics*, vol. 12, pp. 5-20.
- ALLEN P. G. - FILDES R. (2001), *Econometric forecasting*. In J. Armstrong (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*, pp. 271-327. Boston: Kluwer Academic Publishing.
- ATHANASOPOULOS G. - HYNDMAN R.J. - SONG H. - WU D.C. (2011), *The tourism forecasting competition*, in *International Journal of Forecasting*, vol. 27, pp. 822-844.
- BANCA D'ITALIA (2014), *Turismo internazionale dell'Italia*, accessed December 2014, [https://www.bancaditalia.it/statistiche/rapp\\_estero/turismo-int](https://www.bancaditalia.it/statistiche/rapp_estero/turismo-int)
- BELL W.R. - HILLMER S.C. (1983), *Modeling time series with calendar variation*, in *Journal of the American Statistical Association*, vol. 78, pp. 526-534.
- BOX G.E.P. - JENKINS G.M. (1970), *Time Series Analysis, Forecasting and Control* (rev. ed., 1976), Holden-Day, San Francisco.
- BOX G.E.P. - JENKINS G.M. - REINSEL G. (2008), *Time series analysis* (4th ed.). New Jersey: Wiley.
- CALANTONE R.J. - DI BENEDETTO C.A. - BOJANIC D. (1987), *A comprehensive review of the tourism forecasting literature*, in *Journal of Travel Research*, vol. 26, pp. 28-39.
- CHANG I. - TIAO G.C. - CHEN C. (1988), *Estimation of time series parameters in the presence of outliers*, in *Technometrics*, vol. 30, pp. 193-204.
- CHEN C. - LIU L.-M. (1993a), *Forecasting time series with outliers*, in *Journal of Forecasting*, vol. 12, pp. 13-35.
- CHEN C. - LIU L.-M. (1993b), *Joint estimation of model parameters and outliers effects in time series*, in *Journal of the American Statistical Association*, vol. 88, pp. 284-297.
- CHO V. (2001), *Tourism forecasting and its relationship with leading economic indicators*, in *Journal of Hospitality and Tourism Research*, vol. 25, pp. 399-420.
- CHO V. (2003), *A comparison of three different approaches to tourist arrival forecasting*, in *Tourism Management*, vol. 24, pp. 323-330.
- CROUCH G. I. (1994), *The study of international tourism demand: A review of practice*, in *Journal of Travel Research*, vol. 33, n. 41-54.
- DU PREEZ J. - WITT S.F. (2003), *Univariate versus multivariate time series forecasting: An application to international tourism demand*, in *International Journal of Forecasting*, vol. 19, pp. 435-451.
- ENIT MOSCA (2012, a cura di), *Mercato turistico russo*.
- FILDES R. - WEI Y. - ISMAIL S. (2011), *Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures*, in *International Journal of Forecasting*, vol. 27, n. 3, pp. 902-922.
- FOX A.J. (1972), *Outliers in time series*, in *Journal of the Royal Statistical Society, Series B*, vol. 34, pp. 350-363.

- FRITZ R.G. - BRANDON C. - XANDER J. (1984), *Combining time-series and econometric forecast of tourism activity*, in *Annals of Tourism Research*, vol. 11, pp. 219-229.
- FURMANOV K. - BALAEVA O. - PREDVODITELEVA M. (2012), *Tourism flows from the Russian Federation to the European Union*, in *Anatolia: An International Journal of Tourism and Hospitality Research*, vol. 23, n. 1, pp. 17-31.
- GOH C. - LAW R. (2002). *Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention*, in *Tourism Management*, vol. 23, pp. 499-510.
- GUPTA M. - GAO J. - AGGARWAL C. C. - HAN J. (2014), *Outlier detection for temporal data: A survey*, *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, n. 1, pp. 1-20.
- GUSTAVSSON P. - NORDSTROM J. (2001), *The impact of seasonal unit roots and vector ARMA modelling on forecasting monthly tourism flows*, in *Tourism Economics*, vol. 7, pp. 117-133.
- HILLMER S.C. (1984), *Monitoring and adjusting forecasts in the presence of additive outliers*, in *Journal of Forecasting*, vol. 3, pp. 205-215.
- ISNART-ISTITUTO NAZIONALE RICERCHE TURISTICHE - ONT-OSSERVATORIO NAZIONALE DEL TURISMO (2011, a cura di), *Indagine sul turismo organizzato internazionale. Rapporto annuale*, accessed August 2014, [http://www.assolombarda.it/fs/2011728115736\\_57.pdf](http://www.assolombarda.it/fs/2011728115736_57.pdf).
- ISTAT (2002 e 2004), *Statistiche del turismo*, accessed June 2014, *Anno 2001*, [http://www3.istat.it/dati/catalogo/20021009\\_00/](http://www3.istat.it/dati/catalogo/20021009_00/); *Anno 2002*, [http://www3.istat.it/dati/catalogo/20041116\\_00/](http://www3.istat.it/dati/catalogo/20041116_00/).
- ISTAT (2005-2011), *Movimento dei clienti negli esercizi ricettivi. Dati definitivi*, accessed July 2014, *Anno 2003*, <http://www.istat.it/it/archivio/10409>; *Anno 2004*, <http://www.istat.it/it/archivio/10390>; *Anno 2005*, <http://www.istat.it/it/archivio/10293>; *Anno 2006*, <http://www.istat.it/it/archivio/10263>; *Anno 2007*, <http://www.istat.it/it/archivio/10228>; *Anno 2008*, <http://www.istat.it/it/archivio/13620>; *Anno 2009*, <http://www.istat.it/it/archivio/4014>; *Anno 2010*, <http://www.istat.it/it/archivio/48501>.
- ISTAT (2012-2014), *I.stat, Servizi-Turismo-Capacità degli esercizi ricettivi e movimento dei clienti negli esercizi ricettivi-Movimento dei clienti per Paese d'origine, Dati mensili, Anno 2011 e 2012*, accessed June 2014, <http://dati.istat.it/>.
- KIM J.H. - WONG K.K.F. - ATHANASOPOULOS G. - LIU S. (2011), *Beyond point forecasting: evaluation of alternative prediction intervals for tourist arrivals*, in *International Journal of Forecasting*, vol. 27, n. 3, pp. 887-901.
- KULENDRAN K. - KING M.L. (1997). *Forecasting international quarterly tourist flows using error-correction and time-series models*, in *International Journal of Forecasting*, vol. 13, pp. 319-327.
- LEDOLTER J. (1989), *The effect of additive outliers on the forecasts from ARIMA models*, in *International Journal of Forecasting*, vol. 5, pp. 231-240.
- LEDOLTER J. - ABRAHAM B. (1981), *Parsimony and its importance in time series forecasting*, in *Technometrics*, vol. 23, pp. 411-414.
- LI G. - SONG H. - WITT S.F. (2005), *Recent developments in econometric modeling and forecasting*, in *Journal of Travel Research*, vol. 44, pp. 82-99.
- LIM C. (1997a), *Review of international tourism demand models*, in *Annals of Tourism Research*, vol. 24, pp. 835-849.
- LIM C. (1997b), *An econometric classification and review of international tourism demand models*, in *Tourism Economics*, vol. 3, pp. 69-81.

- LIM C. (1999), *A meta analysis review of international tourism demand*, in *Journal of Travel Research*, vol. 37, pp. 273–284.
- LIU L.-M. (1986), *Identification of time series models in the presence of calendar variation*, in *International Journal of Forecasting*, vol. 2, pp. 357–372.
- LIU L.-M. (2009), *Time Series Analysis and Forecasting*, (rel. 2.2), Scientific Computing Associates, River Forest, IL.
- LJUNG G.M. (1993), *On outlier detection in time series*, in *Journal of the Royal Statistical Society, Series B*, vol. 55, n. 2, pp. 559-567.
- LJUNG G.M. - BOX G.E.P. (1978), *On a measure of lack of fit in time series models*, *Biometrika*, vol. 65, pp. 297-303.
- MAHMOUD E. (1984), *Accuracy in forecasting: A survey*, in *Journal of Forecasting*, vol. 3, pp. 139–159.
- MAKRIDAKIS S. (1993), *Accuracy measures: Theoretical and practical concerns*, in *International Journal of Forecasting*, vol. 9, pp. 527–529.
- MINISTERO DEGLI AFFARI ESTERI - ENIT (2014), *Federazione Russa. Rapporto congiunto Ambasciate, Consolati, ENIT, 2014*, accessed August 2014, [http://www.esteri.it/MAE/pdf\\_paesi/EUROPA/Federazione\\_Russa.pdf](http://www.esteri.it/MAE/pdf_paesi/EUROPA/Federazione_Russa.pdf).
- NEWBOLD P. - GRANGER C. W. J. (1974), *Experience with forecasting univariate time series and the combination of forecasts*, *Journal of the Royal Statistical Society, Series A (General)*, vol. 137, n. 2, pp. 131–165.
- OH C.O. - MORZUCH B.J. (2005), *Evaluating time-series models to forecast the demand for tourism in Singapore: Comparing within-sample and post-sample results*, in *Journal of Travel Research*, vol. 43, pp. 404–413.
- ONT-OSSERVATORIO NAZIONALE DEL TURISMO (2011), *Schede mercato Russia*, accessed August 2014, <http://www.ontit.it/opencms/export/sites/default/ont/it/documenti/files/Russia.pdf>.
- PANKRATZ A. (1991), *Forecasting with Dynamic Regression Models*, Wiley, New York.
- PICCOLO D. (1990), *Introduzione all'analisi delle serie storiche*, La Nuova Italia Scientifica, Roma.
- SHELDON P.J. - VAR T. (1985), *Tourism forecasting: A review of empirical research*, in *Journal of Forecasting*, vol. 4, pp. 183-195.
- SHEN S. - LI G. - SONG H. (2010), *Combination forecasts of international tourism demand*, in *Annals of Tourism Research*, vol. 38, n. 1, pp. 72–89.
- SMERAL E. - WUGER M. (2005), *Does complexity matter? Methods for improving forecasting accuracy in tourism: The case of Australia*, in *Journal of Travel Research*, vol. 44, pp. 100–110.
- SONG H. - HYNDMAN R.J., (2011), *Tourism forecasting: An introduction*, in *International Journal of Forecasting*, vol. 27, pp. 817–821.
- SONG H. - LI G. (2008), *Tourism demand modelling and forecasting - A review of recent research*, in *Tourism Management*, Vol. 29, n. 2, pp. 203–220.
- SONG H. - TURNER L. (2006), *Tourism demand forecasting*, in Dwyer L. - Forsyth P. (Eds.), *International handbook on the economics of tourism*, Cheltenham: Edward Elgar.
- SONG H. - WITT, S.F. (2000), *Tourism demand modelling and forecasting: modern econometric approaches*, Oxford: Pergamon.

- SONG H. - WITT, S.F. (2003), *Tourism forecasting: the general-to specific approach*, in *Journal of Travel Research*, vol. 42, pp. 65–74.
- SONG H. - WITT, S.F. - JENSEN T. (2003), *Tourism forecasting: Accuracy of alternative econometric models*, *International Journal of Forecasting*, vol. 19, pp. 123–141.
- TONINI G. (1994), *Modelli stocastici e previsioni a breve termine della domanda turistica. Il caso italiano dagli anni ottanta in poi*, in *Giornale degli Economisti e Annali di Economia*, vol. 53 (N.S.), n. 10-12, pp. 525-546.
- TONINI G. (2013), *Tourist exchange between Italy and Russia*, in *BECTHK PMAT*, vol. 9, n. 3, pp. 78-82.
- TONINI G. (2015), *Trend e stagionalità delle presenze turistiche russe in Italia*, DEMS Working Paper Series No. 313, Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa, Università degli Studi di Milano-Bicocca.
- TSAY R.S. (1988), *Outliers, level shifts, and variance changes in time series*, in *Journal of Forecasting*, vol. 7, pp. 1-20.
- UYSAL M. - CROMPTON J.L. (1985), *An overview of approaches used to forecast tourism demand*, in *Journal of Travel Research*, vol. 23, pp. 7-15.
- WITT C.A. - WITT S.F. - WILSON N. (1994), *Forecasting international tourist flows*, in *Annals of Tourism Research*, vol. 21, pp. 612-628.
- WITT S.F. - SONG H. (2000), *Forecasting future tourism flows*, in S. Medlik - A. Lockwood (Eds.), *Tourism and hospitality in the 21st century*, pp. 106–118, Oxford: Butterworth-Heinemann.
- WITT S.F.; WITT, C.A. (1992), *Modeling and Forecasting Demand in Tourism*, London: Academic Press.
- WITT S.F. - WITT, C.A. (1995), *Forecasting tourism demand: A review of empirical research*, in *International Journal of Forecasting*, vol. 11, n. 3, pp. 447-475.
- WONG K.K.F.- SONG H. - WITT S.F. - WU D.C. (2007), *Tourism forecasting: to combine or not to combine*, in *Tourism Management*, vol. 28, pp. 1068–1078.