

***GUIDA ALL'UTILIZZO DI
TRAMO-SEATS
PER LA DESTAGIONALIZZAZIONE
DELLE SERIE STORICHE***

*P. Anitori
F. Bacchini
C. Baldi
G. Bruno
M. Cammarota
V. De Vita
F. Di Iorio
R. Gatto
E. Otranto
A. Pallara
F. Polidoro
M. Poli
U. Triacca*

CAPITOLO 1

ASPETTI GENERALI

1.1 Introduzione

L'analisi della congiuntura economica viene svolta correntemente sulla base di numerosi indicatori, molti dei quali prodotti dall'Istituto nazionale di statistica. Negli ultimi anni gli sviluppi metodologici e l'acquisizione di strumenti informatici innovativi hanno dato nuovo impulso allo studio della congiuntura, contribuendo a migliorare la comprensione delle dinamiche economiche. D'altra parte, l'avvio dell'Unione Monetaria Europea ed i processi di internazionalizzazione in atto, allargando la platea degli utilizzatori delle informazioni statistiche prodotte in ciascun Paese, richiedono che esse siano sempre più agevolmente interpretabili e comparabili. In particolare, è cresciuta la richiesta di dati di carattere congiunturale presentati in forma "destagionalizzata", cioè depurati dalle fluttuazioni di carattere stagionale che impediscono di cogliere, in modo adeguato, l'evoluzione di breve termine dei fenomeni considerati.

In questo contesto l'Istat, agli inizi del 1997, ha istituito una commissione scientifica con il compito di formulare proposte relative alle strategie da utilizzare per la destagionalizzazione di serie storiche prodotte dall'Istituto: la commissione SARA (*Seasonal Adjustment Research Appraisal*), alla quale hanno partecipato esperti del mondo accademico, dell'Istat, della Banca d'Italia e di altre istituzioni pubbliche e private. Tale commissione ha esaminato e confrontato le caratteristiche delle principali procedure di destagionalizzazione attualmente disponibili (in particolare, delle procedure TRAMO-SEATS e X-12-ARIMA). I risultati, presentati nell'ambito di un convegno, hanno mostrato come la procedura TRAMO-SEATS (**T**ime **S**eries **R**egression with **A**rima **N**oise, **M**issing **O**bservations and **O**utliers **X**E "Outliers" - **S**ignal **E**xtraction in **A**rima **T**imes **S**eries) consenta una stima XE "stima" più accurata delle diverse componenti e, in definitiva, una più corretta informazione congiunturale in favore degli utilizzatori.

Sulla base di queste conclusioni l'Istat ha successivamente deciso di adottare la procedura TRAMO-SEATS, abbandonando quella denominata X-11-ARIMA, impiegata a partire dalla metà degli anni Ottanta.

La sperimentazione della procedura TRAMO-SEATS è stata condotta dall'Istat nell'ambito del progetto di ricerca "Destagionalizzazione delle serie storiche prodotte dall'ISTAT". La presente guida prende spunto da questo periodo di sperimentazione e cerca di rispondere all'esigenza di rendere trasparente l'informazione prodotta e facilitarne l'interpretazione. L'obiettivo fondamentale è, quindi, quello di fornire agli operatori le linee guida per una corretta utilizzazione della procedura TRAMO-SEATS per la destagionalizzazione delle serie temporali. In particolare, la guida è stata concepita allo scopo di guidare un utente, privo di esperienza specifica, che voglia sperimentare la procedura TRAMO-SEATS per depurare una o più serie storiche da effetti legati alla stagionalità del fenomeno oggetto di studio. Esso si configura, dunque, come uno strumento essenzialmente operativo che, pur non dimenticando di fornire – laddove necessario – gli opportuni chiarimenti metodologici illustra le singole fasi della procedura e le diverse possibilità di scelta nell'impostazione dei parametri, arricchendo le spiegazioni con esempi concreti. La guida è inoltre corredata da un *floppy-disk* contenente i dati necessari a replicare gli esempi presentati.

1.2 Che cosa è la destagionalizzazione e a cosa serve una serie destagionalizzata XE "serie destagionalizzata"

La costruzione di una qualsiasi procedura di destagionalizzazione si basa sull'ipotesi che ogni serie storica, y_t , a cadenza infrannuale sia esprimibile come combinazione delle seguenti componenti, non osservabili:

- la componente XE "componente" di *trend*, T_t , che rappresenta la tendenza di lungo periodo;
- la componente XE "componente" *ciclica*, C_t , che rappresenta le fluttuazioni intorno alla tendenza di lungo termine;
- la componente XE "componente" stagionale XE "componente stagionale", S_t , caratterizzata da oscillazioni ricorrenti con periodo inferiore all'anno;

- la componente XE "componente" ciclica, C_t , che rappresenta le fluttuazioni intorno alla tendenza a lungo termine;
- la componente XE "componente" stagionale XE "componente stagionale", S_t , caratterizzata da oscillazioni ricorrenti con periodo inferiore all'anno;
- la componente XE "componente" irregolare XE "componente irregolare", I_t , dovuta a fattori erratici.

Il legame tra le componenti può essere di tipo additivo:

$$Y_t = T_t + C_t + S_t + I_t$$

oppure moltiplicativo:

$$Y_t = T_t C_t S_t I_t$$

Si noti che, spesso, le componenti C e T vengono considerate come un'unica componente XE "componente" denominata ciclo-trend XE "ciclo-trend", data la difficoltà teorica ed empirica di identificarle separatamente.

Le procedure di destagionalizzazione sono finalizzate all'eliminazione della componente XE "componente" stagionale XE "componente stagionale". Quasi tutte prevedono un trattamento preliminare XE "preliminare" dei dati, inteso a correggere l'influenza esercitata dal diverso numero di giorni lavorativi XE "giorni lavorativi" contenuti nei periodi considerati (in genere, mesi o trimestri), dalla presenza di festività fisse o "mobili" come la Pasqua XE "Pasqua" o, infine, da valori anomali (*outlier*), legati a fenomeni accidentali o straordinari (scioperi, calamità naturali, ecc.). Alcune procedure, come TRAMO-SEATS, consentono inoltre di trattare serie con valori mancanti.

Una volta che i dati siano stati sottoposti a questo trattamento preliminare XE "preliminare" è necessario stimare le componenti non osservate, precedentemente ricordate (ciclo-trend XE "ciclo-trend", stagionalità e componente XE "componente" irregolare XE "componente irregolare"). A questo fine, gli approcci utilizzati sono essenzialmente due. Il primo (approccio *filter-based*) consente di trovare stimatori senza considerare l'esistenza di un modello statistico sottostante la serie analizzata; tali procedure, come ad esempio le procedure X-11, X-11 ARIMA XE "ARIMA" e, successivamente, X-12 ARIMA, si basano sull'applicazione ripetuta di una serie di filtri lineari costituiti da medie mobili centrate di diversa lunghezza. Queste procedure sono dette *ad hoc* in quanto i filtri adottati per stimare le componenti non osservate non derivano dalla struttura probabilistica del processo stocastico che ha generato la serie.

Il secondo approccio (*model-based*), al contrario, ipotizza l'esistenza di un modello statistico parametrico, concependo la serie storica osservata come la parte finita della realizzazione di un processo stocastico, la cui struttura probabilistica è adeguatamente descritta dal suddetto modello. I filtri lineari utilizzati nell'approccio *model-based* sono, di solito, filtri simmetrici, la cui espressione analitica dipende dal modello seguito dalla serie storica considerata e quindi risultano legati alle caratteristiche stocastiche della serie stessa. Questo tipo di approccio è adottato dalla procedura TRAMO-SEATS; in particolare, tale procedura si basa su modelli di tipo ARIMA XE "ARIMA" (*AutoRegressive Integrated Moving Average*).

Sottraendo dalla serie storica originaria la componente XE "componente" relativa alla stagionalità (o dividendo, nel caso di un modello moltiplicativo, per il fattore stagionale) è possibile ottenere la serie "destagionalizzata", la quale presenta normalmente un andamento meno oscillante della serie "grezza", pur mantenendo al suo interno la componente irregolare XE "componente irregolare". Se si elimina anche quest'ultima, si ottiene la stima XE "stima" del ciclo-trend XE "ciclo-trend", il quale rappresenta la tendenza di medio-lungo termine della serie storica.

Per comprendere l'utilità da un punto di vista informativo della destagionalizzazione, è utile fare riferimento all'esempio della produzione industriale (figura 1.1).

La serie "grezza" dell'indice della produzione industriale presenta una brusca e profonda caduta nel mese di agosto, a causa della chiusura di molti impianti industriali per le ferie estive. Ciò rende difficile un confronto corretto tra mesi contigui (nello specifico tra agosto e luglio e tra settembre e agosto). Pertanto, per valutare l'evoluzione nel tempo di una serie storica "grezza", come quella della produzione industriale in cui è presente una componente XE "componente" stagionale XE "componente stagionale", si deve ricorrere alle variazioni percentuali "tendenziali" (calcolate cioè rispetto allo stesso periodo dell'anno precedente). La corrispondente serie "destagionalizzata" mostra, invece, una dinamica molto più omogenea nel corso dell'anno, pur presentando oscillazioni significative, dovute sia a fattori di natura ciclica, sia alla presenza di una componente puramente irregolare. La destagionalizzazione, quindi, consente di interpretare correttamente anche le variazioni percentuali "congiunturali", calcolate rispetto al periodo immediatamente precedente (mese su mese). In questo modo, si può disporre di una misura delle tendenze di breve periodo del fenomeno analizzato, non influenzata dagli elementi di natura stagionale.

stagionale.

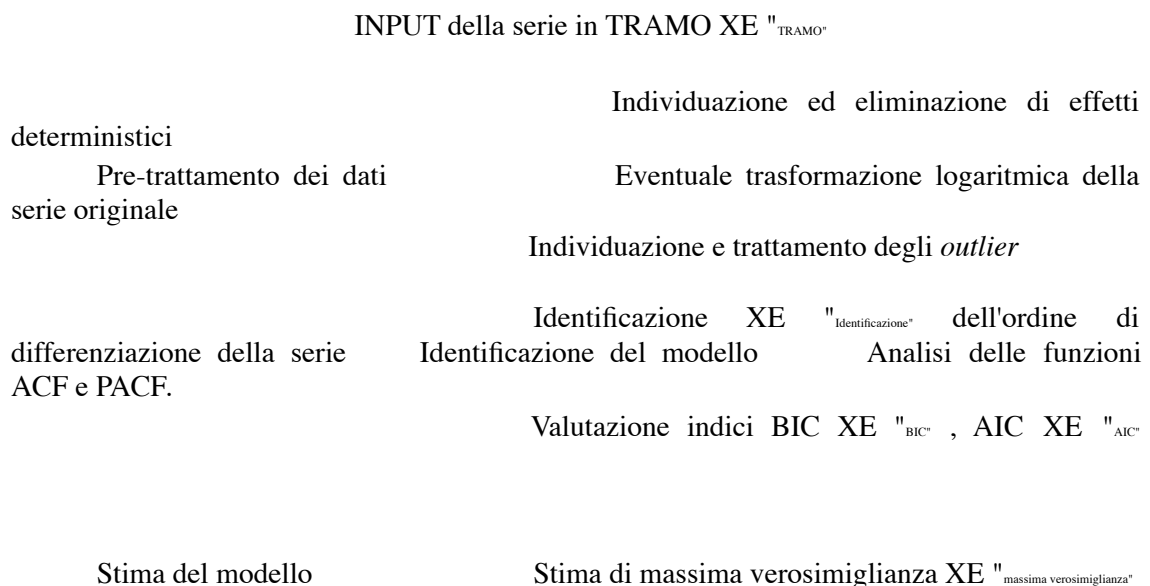
Figura 1.1 – Serie grezza (linea continua) e destagionalizzata (linea tratteggiata) dell'indice generale della produzione industriale (base 1995=100)

1.3 La procedura TRAMO-SEATS

La procedura TRAMO-SEATS si compone di due parti. La prima, costituita da TRAMO XE "TRAMO", è dedicata ad eliminare dalla serie storica d'interesse i cosiddetti effetti deterministici dovuti al diverso numero di giorni lavorativi XE "giorni lavorativi" nei vari periodi di riferimento, alla presenza di festività "mobili" (ad esempio, la Pasqua XE "Pasqua") e di valori anomali. Inoltre, TRAMO identifica e stima XE "stima" il modello ARIMA XE "ARIMA" per la serie storica osservata. La seconda parte della procedura, costituita da SEATS XE "SEATS", effettua la vera e propria destagionalizzazione della serie originaria utilizzando il modello ARIMA identificato in TRAMO.

TRAMO-SEATS è una procedura di tipo *model-based*, cioè basata sulla costruzione di un particolare modello statistico per ciascuna serie storica analizzata; essa incorpora gli avanzamenti compiuti negli ultimi anni nell'ambito della cosiddetta "analisi moderna delle serie storiche" ed offre un ampio spettro di strumenti di carattere statistico per valutare la qualità della destagionalizzazione effettuata. Lo schema riportato in figura 1.2 riassume le fasi essenziali della procedura e descrive le funzioni principali dei due programmi integrati TRAMO XE "TRAMO" e SEATS XE "SEATS".

Figura 1.2. Fasi principali della procedura TRAMO-SEATS



Le serie trattate vengono sottoposte a SEATS XE "SEATS"

Le serie trattate vengono sottoposte a SEATS XE "^{SEATS}"

Verifica del modello individuato " stima" del modello	Eventuale nuova identificazione e stima XE da TRAMO XE " ^{TRAMO} "
" Wiener-Kolmogorov"	Applicazione dei filtri di Wiener-Kolmogorov XE
Stima delle componenti	Analisi degli errori
Attribuzione degli effetti deterministici originale	Scomposizione finale della serie alle diverse componenti della serie

Dal punto di vista operativo, la procedura TRAMO XE "^{TRAMO}" effettua nell'ordine:

1. l'analisi preliminare XE "^{preliminare}" della serie allo scopo di individuare eventuali effetti deterministici (ad esempio effetti di calendario, gli *outlier*) ed effettuare opportune trasformazioni dei valori originari;
2. la fase di identificazione del modello che consente di specificarne gli ordini dei polinomi *AR* e *MA* XE "^{MA}" e l'eventuale grado di integrazione del modello *ARIMA* XE "^{ARIMA}";
3. la stima XE "^{stima}" del modello *ARIMA* XE "^{ARIMA}" identificato, utilizzando il metodo della massima verosimiglianza XE "^{massima verosimiglianza}";
4. l'elaborazione di test statistici per verificare se il modello individuato sia una buona rappresentazione del processo stocastico che ha generato la serie osservata.

Una volta che la serie originaria è stata depurata dagli effetti deterministici ottenendo la cosiddetta serie "linearizzata", quest'ultima viene suddivisa nelle sue componenti non osservabili dalla procedura SEATS XE "^{SEATS}". La scomposizione XE "^{scomposizione}" della serie si basa sull'ipotesi che ogni componente XE "^{componente}" segua un particolare modello *ARIMA* XE "^{ARIMA}" coerente con quello individuato per la serie originaria. Va ricordato che, oltre alla scomposizione della parte stocastica, SEATS assegna i vari effetti deterministici individuati da TRAMO XE "^{TRAMO}" alle diverse componenti, ottenendo così le componenti finali. In estrema sintesi, le operazioni compiute da SEATS prevedono tre passi:

1. l'identificazione dei modelli *ARIMA* XE "^{ARIMA}" per le componenti della serie linearizzata XE "^{serie linearizzata}";
2. la stima XE "^{stima}" delle componenti e l'analisi degli errori;
3. l'allocazione tra le varie componenti delle parti deterministiche individuate da TRAMO XE "^{TRAMO}".

In seguito, SEATS XE "^{SEATS}" offre l'opportunità di rivalutare i risultati raggiunti sia attraverso una diagnostica molto accurata sia tramite un'analisi grafica delle componenti non osservabili.

1.4 La struttura della guida

La guida si articola nel seguente modo. Nel secondo capitolo viene illustrato l'utilizzo del programma TRAMO XE "^{TRAMO}", si approfondiscono i suoi presupposti metodologici, vengono spiegate le opzioni disponibili, i loro effetti, e le procedure informatiche per la loro applicazione.

Nel terzo capitolo viene presentato il programma SEATS XE "^{SEATS}"; la trattazione sull'uso del programma segue la stessa impostazione data al capitolo precedente.

Il quarto capitolo è dedicato all'utilizzo congiunto di TRAMO-SEATS, con alcune esemplificazioni.

Nel quinto capitolo vengono trattate alcune problematiche di carattere sia pratico sia metodologico (stabilità del modello e dei coefficienti, revisioni, problema dell'aggregazione temporale e settoriale).

Infine, un indice analitico guida alla consultazione del manuale, mentre in appendice è riportato l'elenco dei parametri per funzione utilizzati da TRAMO-SEATS con una breve descrizione del loro significato.

Le parti del testo in carattere più piccolo vengono considerate come degli approfondimenti di particolari aspetti metodologici ed operativi, che non sono essenziali per la comprensione degli

significato.

Le parti del testo in carattere più piccolo vengono considerate come degli approfondimenti di particolari aspetti metodologici ed operativi, che non sono essenziali per la comprensione degli argomenti esposti.

La versione di TRAMO-SEATS, a cui fa riferimento questa guida, è quella di giugno 1998.

CAPITOLO 2

TRAMO XE " TRAMO "

2.1 Introduzione

Il programma TRAMO XE " TRAMO " effettua l'identificazione e la stima XE " stima " di un modello di regressione XE " regressione " con errori autocorrelati. Supponendo che la serie y_t da analizzare possa essere scomposta in una parte deterministica w_t e in una parte stocastica z_t , il modello considerato dal programma è:

$$y_t = w_t + z_t = \sum_{i=1}^k \beta_i x_{it} + z_t \quad (2.1)$$

dove le x_{it} sono k variabili non stocastiche e i β_i i corrispondenti coefficienti di regressione XE " regressione ". In particolare, i termini x_{it} misurano gli effetti deterministici imputabili a fenomeni quali la presenza di festività fisse o "mobili" (ad esempio la Pasqua XE " Pasqua), l'esistenza di valori anomali (*outlier*) legati a fenomeni accidentali o straordinari (scioperi, calamità naturali, ecc.), il diverso numero di giorni lavorativi XE " giorni lavorativi " nei mesi o trimestri.

Si ipotizza inoltre che la parte stocastica z_t segua, a sua volta, un modello ARIMA XE " ARIMA " stagionale moltiplicativo, cioè:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D z_t = \theta(B)\Theta(B^s)a_t \quad (2.2)$$

dove B è l'operatore *ritardo* (*backward*), $\phi(B)$, $\Phi(B)$ sono polinomi nell'operatore ritardo B di ordine p e q rispettivamente, $\phi(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_p B^{ps}$, $\Phi(B^s) = 1 + \phi_1 B^s + \phi_2 B^{2s} + \dots + \phi_q B^{qs}$, e le variabili a_t rappresentano disturbi incorrelati con media nulla (*white noise* XE " white noise ") per i quali, ai fini della stima XE " stima " e della verifica delle ipotesi, viene generalmente assunta l'ipotesi di normalità. Combinando le espressioni (2.1) e (2.2) otteniamo la seguente equazione:

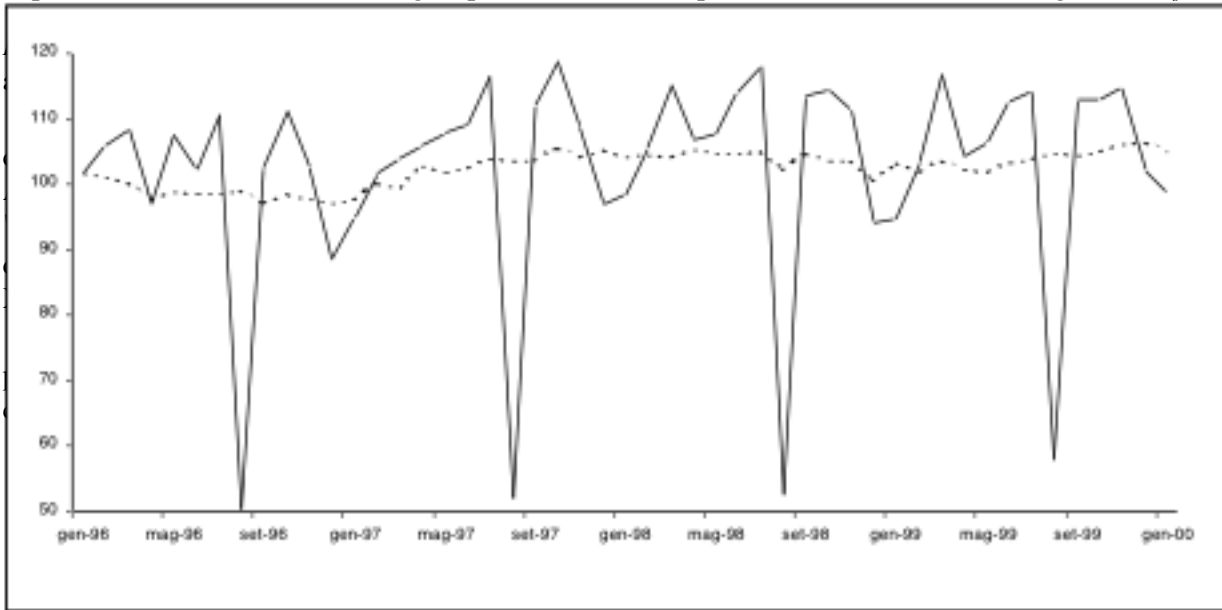
$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D (y_t - \sum_{i=1}^k \beta_i x_{it}) = \theta(B)\Theta(B^s)a_t \quad (2.3)$$

La sequenza operativa per l'identificazione e la stima XE " stima " del modello (2.3) mediante il programma TRAMO XE " TRAMO " può essere descritta nel seguente modo. In prima istanza la serie y_t

viene depurata dagli effetti deterministici (stimati) pervenendo alla serie $z_t = (y_t - \sum_{i=1}^k \beta_i x_{it})$. Successivamente la serie z_t , qualora si renda necessario, viene resa stazionaria applicando,

$$z_t = (y_t - \sum_{i=1}^d \beta_i x_{it})$$

viene depurata dagli effetti deterministici (stimati) pervenendo alla serie z_t . Successivamente la serie z_t , qualora si renda necessario, viene resa stazionaria applicando, rispettivamente, d volte e D volte gli operatori differenza prima $D=I-B$ e differenza stagionale $D_s=I-B_s$.



Alla fine del capitolo è riportata un'appendice che contiene un file di output XE "file di output" di TRAMO XE "TRAMO" commentato.

2.2 L'input di TRAMO XE "TRAMO"

2.2.1 Il formato del file di input XE "file di input"

Il programma TRAMO XE "TRAMO" richiede un file di input XE "file di input" in formato ASCII, così costruito:

prima riga: *nome della serie*

seconda riga: *numero di osservazioni; anno di partenza della serie di riferimento; periodo (ad es. mese o trimestre) di partenza;*

righe successive: *osservazioni*

ultime righe: *sequenza dei parametri di input separati da spazio o da virgola; la sequenza deve essere preceduta dal comando \$INPUT e conclusa con \$.*

Come esempio si riporta il file di input XE "file di input" riferito alla serie dell'indice grezzo della produzione industriale:

Indice grezzo produzione industriale
220 1981 1

```
80.7
81.9
90.0
.....
.....
94.6
102.6
$INPUT LAM=1 IEAST=1 IDUR=6 ITRAD=2 SEATS XE "SEATS" =2$
```

La serie è dunque composta da 220 osservazioni con cadenza mensile, ed ha inizio con il mese di gennaio 1981. Il significato dei parametri compresi tra \$INPUT e \$ - che, comunque, possono occupare più righe - sarà illustrato nelle pagine successive. Al file con l'impostazione appena descritta va assegnato il nome *serie* e deve essere inserito nella directory TRAMO XE "TRAMO" .

Alternativamente, è possibile procedere come segue:

- creare un file con un nome qualunque su un massimo di otto posizioni, che contenga il nome della serie (prima riga) il numero di osservazioni, l'anno di partenza della serie, il periodo di partenza della serie (seconda riga) e le osservazioni (righe successive), escludendo la sequenza dei parametri. Tale file andrà collocato nella subdirectory di TRAMO XE "TRAMO", chiamata SERIES;
- digitare dalla directory TRAMO XE "TRAMO" il comando **INPUT** e premere **ENTER**: apparirà l'elenco delle serie contenute nella directory SERIES;
- selezionare il file appena creato, contenente la serie che si desidera analizzare, e premere **ENTER**: apparirà una schermata con tutti i parametri di TRAMO XE "TRAMO" con i propri valori di default;
- qualora si desiderino conservare i valori di default digitate **END** e premere **ENTER**;
- qualora si desideri modificare i valori di default bisognerà digitare, per ogni parametro da modificare, *nome*

- selezionare il file appena creato, contenente la serie che si desidera analizzare, e premere **ENTER**: apparirà una schermata con tutti i parametri di TRAMO XE "TRAMO" con i propri valori di default;
- qualora si desiderino conservare i valori di default digitare **END** e premere **ENTER**;
- qualora si desideri modificare i valori di default bisognerà digitare, per ogni parametro da modificare, *nome parametro* = *nuovo valore* e premere **ENTER**;
- terminata l'impostazione dei parametri, digitare **END** e premere **ENTER**. A questo punto il file *serie* è creato secondo il tracciato descritto ed è quindi pronto per essere elaborato da TRAMO XE "TRAMO".

Digitando dalla directory di TRAMO XE "TRAMO" il comando **TRAMO** viene effettuata l'analisi della serie storica. I risultati dell'elaborazione saranno collocati in un file il cui nome sarà uguale alla prima parola della prima riga del file di input XE "file di input" (o alle sue prime otto lettere) ed avrà estensione *out*. Tale file viene creato automaticamente, al termine di ogni elaborazione, all'interno della subdirectory OUTPUT, della directory TRAMO. Nel caso dell'esempio 1 i risultati saranno inseriti nel file *indice.out*.

2.2.2 Alcuni parametri iniziali

Nel programma TRAMO XE "TRAMO" ogni parametro ha un valore di default per cui tutti i parametri che non vengono specificati nel file di input XE "file di input" sono assunti con tale valore.

Per quanto riguarda la frequenza della serie, in assenza di diversa specificazione, essa viene supposta mensile. Per specificare una frequenza diversa è necessario inserire tra i parametri il comando:

MQ = numero di osservazioni nell'anno (es. 4 se la serie è trimestrale, 6 se bimestrale, ecc.).

TRAMO XE "TRAMO" consente poi di analizzare più serie contemporaneamente attraverso l'utilizzo del parametro *ITER*; nello stesso file *serie* dovranno essere inserite in sequenza (senza alcuna riga bianca di separazione) tutte le serie storiche da analizzare seguendo le specifiche precedentemente illustrate nel paragrafo 2.2.1 e aggiungendo *ITER* tra i parametri della prima serie in input. In particolare si imposterà:

- ITER*=2 se le serie sono trattate dal programma con lo stesso modello (in questo caso solamente la prima serie avrà la stringa di istruzioni $\$INPUT... \$$ con i parametri);
- ITER*=3 se le serie sono trattate seguendo ognuna un modello diverso specificato dalle corrispondenti righe dei parametri di input (in questo caso tutte le serie nel file di input XE "file di input" avranno la stringa di istruzioni $\$INPUT... \$$ con i parametri).

In ogni caso verranno prodotti diversi file di output XE "file di output", ognuno in corrispondenza di una serie secondo le modalità precedentemente definite (è necessario che le prime otto lettere della denominazione delle serie diano luogo a nomi diversi per ciascuna di esse).

Il valore del parametro *ITER* = 0 (che è il valore di default) indica a TRAMO XE "TRAMO" di analizzare una sola serie con un solo modello opportunamente specificato, mentre il valore *ITER* = 1 consente di analizzare una sola serie utilizzando diversi modelli a loro volta debitamente specificati.

TRAMO XE "TRAMO" viene quasi sempre utilizzato unitamente al programma SEATS XE "SEATS" (descritto nel terzo capitolo), nel senso che SEATS utilizza un gruppo di file di output XE "file di output" prodotti da TRAMO per procedere alla scomposizione XE "scomposizione" della serie considerata. Affinché ciò sia possibile è necessario definire, tra i parametri contenuti nella stringa di istruzioni ($\$INPUT... \$$), il parametro *SEATS* ponendolo pari a 2 (il valore di default è 0).

La tavola seguente riassume quanto specificato finora per i parametri *MQ*, *ITER* e *SEATS* XE "SEATS" (ricordiamo che i valori in grassetto sono quelli di default):

PARAMETRI	VALORI ASSUMIBILI	SIGNIFICATO
<i>MQ</i>	numero intero positivo (es. 1, 4, 6, 12)	Numero di osservazioni per anno
<i>ITER</i>	0 , 1, 2, 3	Opzioni per selezionare una o più serie con uno o più modelli
<i>SEATS</i> XE "SEATS"	0 , 1, 2	Utilizzo di TRAMO XE "TRAMO" insieme con <i>SEATS</i> XE "SEATS"

2.2.3 L'analisi preliminare XE "preliminare" della serie: correzione per effetti di regressione XE "regressione".

Come più volte sottolineato, la serie da analizzare tramite il modello ARIMA XE "ARIMA" deve essere puramente stocastica; pertanto è necessario depurarla preliminarmente da tutti gli effetti deterministici, quali gli effetti di calendario (legati al numero di giorni lavorativi XE "giorni lavorativi" e alle feste pasquali), l'effetto indotto da eventuali valori anomali (*outlier*) e altri introdotti dall'utente come regressori esterni. Di seguito vengono riportati gli effetti che il programma TRAMO XE "TRAMO" individua e provvede ad eliminare.

quali gli effetti di calendario (legati al numero di giorni lavorativi XE "giorni lavorativi" e alle feste pasquali), l'effetto indotto da eventuali valori anomali (*outlier*) e altri introdotti dall'utente come regressori esterni. Di seguito vengono riportati gli effetti che il programma TRAMO XE "TRAMO" individua e provvede ad eliminare.

2.2.3.1 Gli effetti di calendario dovuti ai giorni lavorativi XE "giorni lavorativi"

L'andamento di molte serie economiche di flusso è influenzato dalla composizione del calendario nei singoli periodi. Ad esempio, la produzione industriale sarà più alta, a parità di altri fattori, nei mesi che contengono un numero maggiore di giorni lavorativi XE "giorni lavorativi". Anche il numero di giorni specifici della settimana (numero di lunedì, di martedì, ecc.) presenti in un mese o in un trimestre può avere una influenza sulle serie; ciò accade, ad esempio, nel caso di serie riguardanti vendite di particolari beni. Il programma TRAMO XE "TRAMO" prevede allora alcune opzioni per affrontare e risolvere questo problema attraverso il parametro *ITRAD*. Ponendo $ITRAD=0$, questo effetto non viene considerato, mentre impostando $ITRAD=6$ vengono definiti 6 regressori, uno per ogni giorno lavorativo della settimana (compreso il sabato). Se indichiamo con $X1t, X2t, X3t, X4t, X5t, X6t, X7t$, rispettivamente il numero di lunedì, martedì, ..., domeniche nel periodo di riferimento t (mese, trimestre, ...), allora i sei regressori saranno costituiti da $(X1t-X7t), (X2t-X7t), (X3t-X7t), (X4t-X7t), (X5t-X7t), (X6t-X7t)$. Eventualmente, sarà possibile aggiungere un ulteriore regressore, che rappresenti l'effetto degli anni bisestili, ponendo $ITRAD=7$. Quest'ultimo regressore assume valore 0,75 in corrispondenza dei mesi di febbraio degli anni bisestili, -0,25 in corrispondenza degli altri mesi di febbraio e zero in tutti gli altri casi.

Con $ITRAD=1$ il numero dei regressori viene ridotto ad uno solo, considerando uguali da una parte l'effetto di $X1t, X2t, X3t, X4t, X5t$, e dall'altra quello di $X6t$ e $X7t$; la nuova parametrizzazione porterà alla costruzione del regressore $(X1t+X2t+X3t+X4t+X5t)-5/2(X6t+X7t)$. L'inclusione dell'effetto degli anni bisestili in quest'ultimo caso, si otterrà impostando $ITRAD=2$.

E' tuttavia possibile effettuare un test per valutare la presenza o meno di questi effetti: ponendo $ITRAD=-1, -2, -6, -7$, si può verificare la significatività dell'effetto giorni lavorativi XE "giorni lavorativi" secondo le specifiche di *ITRAD* uguale, rispettivamente, a 1, 2, 6, 7. L'output di TRAMO XE "TRAMO" conterrà il valore della stima XE "stima", gli errori standard ed il valore della statistica t in corrispondenza delle variabili denotate da *TRAD* nella sezione "ESTIMATES OF REGRESSION PARAMETERS CONCENTRATED OUT OF THE LIKELIHOOD".

PARAMETRI VALORI ASSUMIBILI SIGNIFICATO

<i>ITRAD</i>	0, 1, 2, 6, 7, -1, -2, -6, -7	Presenza o meno dei giorni lavorativi XE "giorni lavorativi" e dell'effetto degli anni bisestili
--------------	-------------------------------	--

2.2.3.2 Effetto Pasqua XE "Pasqua"

Un effetto deterministico particolare è rappresentato dalle festività mobili, in particolare dalla Pasqua XE "Pasqua" (*Easter Effect*). Infatti, mentre le festività natalizie riguardano sempre il mese di dicembre, le festività pasquali, essendo variabili, possono avere un effetto sia nel mese di marzo che in quello di aprile, creando qualche instabilità nelle dinamiche stagionali. Per questo motivo è importante rimuovere il loro effetto dalla serie originale. L'effetto Pasqua riguarda il periodo di tempo che intercorre tra n giorni prima della Pasqua ed il sabato, vigilia di Pasqua. Per tener conto di questo effetto, viene introdotta una variabile (*dummy*) che assumerà valore 0 per tutti i mesi che non risentono dell'effetto Pasqua, e valore j/n per quei mesi che contengono j giorni del periodo che risente dell'effetto Pasqua. Per esempio, se il giorno di Pasqua di un dato anno è il 5 aprile e la durata dell'effetto Pasqua è $n=6$, la variabile in questione assumerà (per quell'anno) valore $2/6$ per il mese di marzo, $4/6$ per il mese di aprile e 0 per tutti gli altri mesi. Il programma TRAMO XE "TRAMO" contiene un calendario interno in grado di individuare le date della Pasqua, che vengono segnalate nell'output del programma, sia relativamente al periodo di riferimento della serie analizzata, sia relativamente alle date future, sotto la voce *DATES OF EASTER DURING THE REQUESTED TIME SPAN*. I parametri di input di TRAMO di riferimento sono due: *IEAST* e *IDUR*. *IEAST* è il parametro che consente di considerare o meno l'effetto Pasqua. Il suo valore di default è 0; in tal caso l'effetto Pasqua non viene considerato.

Affinché TRAMO XE "TRAMO" consideri tale effetto è necessario definire $IEAST=1$. Se invece si desidera testare la presenza o meno dell'effetto Pasqua XE "Pasqua" è necessario definire $IEAST=-1$. Qualora la presenza di tale effetto venga rilevata si procederà alla depurazione della serie.

IDUR è il parametro che permette di stabilire la durata, in termini di numero di giorni, dell'effetto Pasqua XE "Pasqua". Il valore di default di tale parametro è $IDUR=6$.

L'output di TRAMO XE "TRAMO" conterrà il valore della stima XE "stima", il suo errore standard ed il valore della statistica t in corrispondenza della variabile *EAST* nella sezione *ESTIMATES OF*

Pasqua XE "Pasqua". Il valore di default di tale parametro è $IDUR=6$.

L'output di TRAMO XE "TRAMO" conterrà il valore della stima XE "stima", il suo errore standard ed il valore della statistica t in corrispondenza della variabile EAST nella sezione *ESTIMATES OF REGRESSION PARAMETERS CONCENTRATED OUT OF THE LIKELIHOOD*.

Nel caso in cui si ritenga che la serie in questione debba presentare necessariamente un effetto di questo tipo, ma il pre-test non lo ha individuato, è bene modificare IDUR per tentativi (riducendolo). In quest'ultimo caso è preferibile utilizzare il valore di IDUR più significativo in base al valore del test t .

PARAMETRI VALORI ASSUMIBILI SIGNIFICATO

IEAST	0, 1, -1	Presenza o meno dell'effetto Pasqua XE "Pasqua"
IDUR	0, 1, 2, ..., 6, ...	Lunghezza dell'effetto Pasqua XE "Pasqua"

2.2.3.3 La correzione dei valori anomali

E' noto che molte serie economiche sono soggette a variazioni improvvise (temporanee o permanenti) dovute sia all'influenza di eventi eccezionali, quali scioperi e/o calamità naturali, sia a cambiamenti legislativi o di politica economica, sia alla presenza di errori di diverso tipo nei dati osservati. I loro effetti sulla dinamica della serie storica si manifestano sotto forma di singole occorrenze o sequenze di valori anomali (*outlier*), o di cambi di livello della serie a partire da uno specifico istante di tempo. La presenza di questi valori anomali può avere un impatto diretto sull'adeguatezza del modello ARIMA XE "ARIMA" per la serie osservata e sull'efficienza della sua stima XE "stima"; pertanto è molto importante eliminare il loro effetto prima di procedere all'identificazione del modello.

TRAMO XE "TRAMO" considera e tratta separatamente quattro tipi di outlier:

- outlier additivo* XE "outlier additivo" (*AO XE "AO"*): si manifesta con una brusca variazione della serie in un dato istante temporale, il cui effetto è però immediatamente riassorbito nell'istante successivo;
- cambiamento temporaneo* XE "cambiamento temporaneo" (*TC XE "TC"*): è un cambiamento brusco in un dato istante temporale, che viene riassorbito gradualmente nel tempo;
- cambiamento di livello* XE "cambiamento di livello" (*LS XE "LS"*): è un cambiamento brusco in un dato istante temporale che perdura nel tempo;
- innovation outlier* (*IO XE "IO"*): è un outlier il cui effetto dipende dal processo ARIMA XE "ARIMA" che ha generato la serie e che si può manifestare in modo anche molto diverso a seconda che la serie sia stazionaria o meno (si veda Chen e Liu, 1993). L'utilizzo di questo tipo di outlier è alquanto controverso ai fini della destagionalizzazione e, di conseguenza, non ne verrà illustrato l'impiego nel seguito.

L'individuazione dei valori anomali si basa sulla procedura descritta in Chen e Liu (1993) L'identificazione automatica XE "identificazione automatica" è regolata tramite i parametri *IATIP* ed *AIO*; ponendo il parametro *IATIP=1* viene effettuata la ricerca automatica degli outlier, mentre il parametro *AIO* determina quali tipi si vogliono individuare. Per default *AIO* è pari a 2 e ricerca outlier di tipo AO XE "AO", LS XE "LS" e TC XE "TC"; ponendolo pari a 0 li ricerca tutti e quattro, mentre ponendolo pari ad 1 ricerca solo i tipi AO e TC; infine ponendolo pari a 3 ricerca i tipi AO ed LS.

La procedura richiede la specificazione di un valore soglia dei test, definito mediante il parametro *VA* il cui valore di default è pari a 3,5. Aumentando il valore critico diminuisce la sensibilità della procedura di identificazione degli outlier ed eventualmente il numero di osservazioni trattate come anomale. E' importante comunque avere presente che, data la natura iterativa della procedura di identificazione ed aggiustamento per i diversi outlier, il numero ed il tipo di valori anomali identificati può variare a seconda del modello ARIMA XE "ARIMA" usato e della presenza di altri effetti deterministici.

Alternativamente, o in aggiunta a quelli identificati automaticamente, i valori anomali possono essere imposti dall'utente, specificandone la tipologia. In questo caso assumono la natura di veri e propri regressori esterni e, pertanto, nella stringa di istruzioni (*\$INPUT.....\$*) andrà inserito il parametro *IREG*. Esso definisce il numero di regressori esterni forniti a TRAMO XE "TRAMO" per il trattamento delle serie analizzata. In questo caso *IREG* deve essere uguale al numero di outlier imposti. Successivamente alla stringa (*\$INPUT.....\$*) dovrà essere inserita un'altra stringa di istruzioni, aperta da *\$REG* e chiusa da *\$*, secondo il seguente schema:

```
$INPUT.....IREG=numero regressori esterni .....$  
$REG IUSER=2 NSER=numero di outlier imposti dall'utente $
```

Nella riga successiva andranno poi specificati l'istante temporale in cui si manifesta l'osservazione anomala ed il tipo, secondo le sigle che li contraddistinguono (*AO XE "AO"*, *TC XE "TC"*, *LS XE "LS"*, *IO XE "IO"*).

Per esempio, se si volessero imporre tre *outlier*, due di tipo additivo in corrispondenza della

in una riga successiva andranno poi specificati l'istante temporale in cui si manifesta l'osservazione anomala ed il tipo, secondo le sigle che li contraddistinguono (*AO XE* "AO", *TC XE* "TC", *LS XE* "LS", *IO XE* "IO").

Per esempio, se si volessero imporre tre *outlier*, due di tipo additivo in corrispondenza della trentesima e cinquantunesima osservazione ed un cambiamento di livello *XE* "cambiamento di livello" in corrispondenza della ottantaduesima osservazione, bisognerà specificare quanto segue:

```
$INPUT.....IREG=3...$
$REG IUSER=2 NSER=3$
30 AO XE "AO" 51 AO 82 LS XE "LS"
```

PARAMETRI	VALORI ASSUMIBILI	SIGNIFICATO
IATIP	0, 1	Ricerca e correzione automatica degli outlier
AIO	0, 1, 2, 3	Selezione del tipo di outlier
VA	..., 3.5, ...	Valore soglia per la significatività degli outlier
IREG	0, numero intero positivo	Numero di regressori esterni
IUSER	0, 2	Imposizione degli outlier
NSER	0, numero intero positivo	Numero di regressori esterni relativo ad un singolo comando REG

2.2.3.4 Regressori esterni *XE* "Regressori esterni"

TRAMO *XE* "TRAMO" contempla la possibilità di tenere conto di altri effetti deterministici che influenzano la serie, utilizzando variabili individuate direttamente dall'utente (i cosiddetti regressori esterni). In particolare, ai fini della destagionalizzazione, è interessante valutare l'effetto delle festività nazionali che, essendo peculiari di ciascun Paese, non sono state codificate nel programma.

Le serie relative ai regressori esterni possono essere contenute in un file autonomo in formato ASCII, oppure possono essere inserite nello stesso file *serie*, contenente i dati da destagionalizzare ed i relativi parametri.

Nel primo caso, il file relativo ai regressori esterni dovrà contenere le sole sequenze dei dati (senza alcun titolo iniziale) in colonne affiancate (una colonna per ogni regressore). Nella stringa di istruzioni (*\$INPUT...\$*) andrà quindi inserito il parametro *IREG* che, come abbiamo visto, definisce il numero di regressori esterni. Quindi, come nel caso degli outlier imposti dall'utente, successivamente alla stringa di istruzioni (*\$INPUT...\$*) andrà inserita un'altra stringa di istruzioni (*\$REG...\$*) ed una riga con il nome dei file dei regressori, secondo il seguente schema:

```
$INPUT.....$
$REG IUSER=-1 NSER=numero di regressori contenuti nel file ILONG=lunghezza delle serie dei regressori REGEFF=componente XE "componente" alla quale viene assegnato l'effetto del regressore $ nomefile dei regressori
```

E' importante notare che se l'obiettivo dell'analisi è la suddivisione della serie nelle diverse componenti, gli effetti deterministici individuati nell'analisi preliminare *XE* "preliminare" devono essere attribuiti ad una o più di queste componenti. Ciò può essere fatto utilizzando TRAMO *XE* "TRAMO" insieme con *SEATS XE* "SEATS" (mediante il parametro *SEATS=2*, cfr. paragrafo 2.2.2) e il parametro *REGEFF* che compare nella stringa (*\$REG...\$*) del file di input *XE* "file di input" di TRAMO. Infatti, in base al valore definito per *REGEFF*, *SEATS* attribuisce ad una delle componenti in cui scompone la serie analizzata, gli effetti dovuti al regressore esterno utilizzato. Ad esempio, se si introduce un regressore che ha un effetto di lungo periodo, allora occorrerà allocare l'effetto al trend, ponendo *REGEFF=1*.

Qualora nell'analisi si desideri sia imporre il trattamento di determinati valori anomali, sia considerare l'influenza di regressori definiti dall'utente, il parametro *IREG* dovrà essere posto uguale al numero complessivo dei regressori esterni (outlier imposti più i regressori definiti dall'utente).

A questo punto andranno aggiunte alla stringa di istruzioni (*\$INPUT...\$*) due stringhe di istruzioni (*\$REG...\$*) relative, rispettivamente, ai valori anomali imposti ed ai regressori definiti dall'utente e seguite la prima dalla specificazione dell'istante temporale e del tipo di valore anomalo e la seconda dalla dichiarazione del nome del file dove sono contenuti gli altri regressori esterni.

Ad esempio, se si volesse imporre un outlier additivo *XE* "outlier additivo" nella trentesima osservazione ed un regressore esterno di 180 osservazioni contenuto nel file *nomefile*, si dovrebbe specificare quanto segue:

```
$INPUT.....IREG=2...$
$REG IUSER=2 NSER=1$
30 AO XE "AO"
$REG IUSER=-1 NSER=1 ILONG=180$
```

```

$INPUT...$ IREG=2...$
$REG IUSER=2 NSER=1$
30 AO XE "AO"
$REG IUSER=-1 NSER=1 ILONG=180$
nomefile

```

Poiché TRAMO XE "^{TRAMO}", quando utilizzato insieme con SEATS XE "^{SEATS}", effettua una previsione a 24 mesi della serie analizzata, il file dei regressori esterni deve contenere, per ciascun regressore, i dati dei 24 mesi successivi all'ultimo periodo osservato nella serie considerata. Qualora si desiderino effettuare previsioni per un intervallo temporale superiore ai 24 mesi, è necessario che il file dei regressori contenga i dati relativi a tutto il periodo per il quale si desidera effettuare la previsione.

Alternativamente, nel caso di un solo regressore esterno, i dati possono essere inseriti direttamente nel file *serie* subito dopo la stringa di istruzioni ($\$REG... \$$). L'unica differenza rispetto alla situazione precedente è che bisognerà definire IUSER=1. Considerando il precedente esempio, si avrebbe:

```

$INPUT...$ IREG=2...$
$REG IUSER=1 NSER=1 ILONG=180$
prima osservazione regressore
seconda osservazione regressore
...
...
ultima osservazione regressore

```

Particolare attenzione deve essere posta al caso in cui si voglia considerare l'effetto dovuto alle festività nazionali pubbliche e religiose. Questo effetto può essere considerato mediante l'utilizzo di un regressore esterno. L'influenza di tale regressore può essere analizzata separatamente ovvero considerando le modifiche apportate all'effetto giorni lavorativi XE "^{giorni lavorativi}" individuati con il parametro ITRAD. Nel primo caso il parametro IUSER nella stringa di istruzioni $\$REG... \$$ deve essere pari a -1. Nel secondo caso il parametro IUSER deve essere pari a -2. In questa eventualità è necessario distinguere il caso in cui nella stringa ($\$INPUT... \$$) si è posto ITRAD=1 o 2 dal caso in cui si è posto ITRAD=6 o 7:

- ITRAD=1 o 2; i giorni lavorativi XE "^{giorni lavorativi}" sono rappresentati da un regressore unico e pertanto avremo bisogno di un file (chiamiamolo *holidays*) con una unica colonna, che contenga in corrispondenza di ogni mese (trimestre) il numero di festività che cadono tra i lunedì ed i venerdì del mese (trimestre) stesso. I comandi da impiegare saranno gli stessi di prima seguendo le stesse modalità, ad eccezione del parametro IUSER nella stringa di istruzioni ($\$REG... \$$), che dovrà essere posto uguale a -2. In pratica con questo comando si modifica il regressore $(X1t+X2t+X3t+X4t+X5t)-5/2(X6t+X7t)$, sottraendo al termine $(X1t+X2t+X3t+X4t+X5t)$ i giorni festivi indicati nel file *holidays*, che verranno invece aggiunti al termine $(X6t+X7t)$;
- ITRAD=6 o 7; i giorni lavorativi XE "^{giorni lavorativi}" sono rappresentati da sei regressori e pertanto avremo bisogno di un file *holidays* con sei colonne, che contengano in corrispondenza di ogni mese (trimestre) rispettivamente il numero di lunedì, martedì, ...,sabati festivi. Le istruzioni da usare sono identiche al caso precedente. Nel caso in questione, i regressori $(X1t-X7t)$, $(X2t-X7t)$, $(X3t-X7t)$, $(X4t-X7t)$, $(X5t-X7t)$, $(X6t-X7t)$ vengono modificati sottraendo a Xit ($i=1, \dots, 6$) il numero espresso nella colonna i , riga t , del file *holidays*, che sarà invece aggiunto a $X7t$.

PARAMETRI	VALORI ASSUMIBILI	SIGNIFICATO
IREG	0, numero intero positivo	Numero di regressori esterni
IUSER	-2, -1, 0, 1	Modalità di immissione dei regressori esterni
NSER	0, numero intero positivo	Numero di regressori esterni relativo ad un singolo comando REG
ILONG	0, numero intero positivo	Numero di osservazioni per i regressori esterni
REGEFF	0, 1,2,3,4,5,	Assegnazione degli effetti deterministici alle componenti della serie

2.2.4 L'analisi preliminare XE "^{preliminare} della serie: trasformazioni

Per giungere alla rappresentazione della parte stocastica z_t della serie tramite un modello ARIMA XE "^{ARIMA}", può essere necessario trasformare la serie iniziale y_t in modo che essa soddisfi alcuni requisiti di omogeneità temporale, che riguardano le proprietà di stazionarietà in media ed in varianza XE "^{varianza}" del processo stocastico di cui la serie costituisce una particolare realizzazione.

In termini intuitivi, la stazionarietà di una serie richiede che la media e la varianza XE "^{varianza}" calcolate in un qualsiasi sottoperiodo risultino approssimativamente costanti. Poiché in pratica questo

requisiti di omogeneità temporale, che riguardano le proprietà di stazionarietà in media ed in varianza XE "varianza" del processo stocastico di cui la serie costituisce una particolare realizzazione.

In termini intuitivi, la stazionarietà di una serie richiede che la media e la varianza XE "varianza" calcolate, in un qualsiasi sottoperiodo, risultino approssimativamente costanti. Poiché in pratica queste condizioni sono raramente verificate, è spesso necessario operare alcune trasformazioni che consentano di soddisfare questi requisiti.

2.2.4.1 Trasformazione logaritmica XE "Trasformazione logaritmica"

Il problema della non stazionarietà in varianza XE "varianza" può essere risolto, almeno parzialmente, ricorrendo alla trasformazione logaritmica della serie. In particolare la trasformazione porta ad una soluzione ottimale nel caso in cui:

- la serie è ipotizzabile come la somma di un livello medio mt non stocastico e variabile nel tempo ed una componente XE "componente" di errore casuale e_t ;
- la serie ha una varianza XE "varianza" proporzionale a mt .

Sotto le ipotesi a) e b) la trasformazione logaritmica produce una serie con varianza XE "varianza" costante (si veda Mills, 1990, pagg. 92 e 93). Naturalmente, per effettuare tale trasformazione è necessario che i valori della serie siano tutti positivi.

Il parametro che gestisce la trasformazione logaritmica è LAM . Se posto uguale a 0 (default) $TRAMO$ XE "TRAMO" effettua la trasformazione; se posto uguale ad 1 mantiene la serie nei livelli e se posto pari a -1, effettua un test per verificare l'opportunità di procedere o meno alla trasformazione logaritmica. Da notare, comunque, che anche se il test di $TRAMO$ indica l'opportunità della trasformazione logaritmica, non sempre è appropriato utilizzarla; per maggiori approfondimenti su questo punto rimandiamo il lettore interessato al quinto capitolo di questa guida (sezione 5.4).

2.2.4.2 Differenziazione

La differenziazione della serie consente di risolvere il problema della non stazionarietà in media. In particolare, la differenza prima $y_t - y_{t-1} = (I-B)y_t$, consente di eliminare un trend lineare. Nel caso in cui dopo questa trasformazione la serie risulti ancora non stazionaria in media, la si potrà differenziare ulteriormente. In generale, per eliminare un trend polinomiale di grado d , si ricorrerà alla differenza d -esima. Infine, se nella serie sono presenti anche onde periodiche deterministiche, annuali o di sottomultipli dell'anno (come è solito accadere nelle serie stagionali), sarà necessario ricorrere anche a differenze stagionali; ad esempio, per una serie mensile abbiamo: $y_t - y_{t-12} = (I-B^{12})y_t$.

$TRAMO$ XE "TRAMO" consente di identificare automaticamente l'ordine di differenziazione da applicare alla serie analizzata mediante il parametro $IDIF$; ponendo $IDIF$ uguale a 3, il programma individua l'ordine della differenza non stagionale (fino ad un massimo di 2) e quello della differenza stagionale (fino ad un massimo di 1).

L'utente può anche rinunciare all'identificazione automatica XE "identificazione automatica" dell'ordine di differenziazione, imponendolo mediante i parametri D (per specificare l'ordine delle differenze non stagionali) e BD (per specificare l'ordine delle differenze stagionali) e lasciando il valore di default $IDIF=0$.

In figura 2.1 è riportata la serie grezza dell'indice generale della produzione industriale; è immediato osservare la chiara non stazionarietà in media ed i picchi stagionali in corrispondenza dei mesi di agosto, mentre non c'è una netta evidenza per la non stazionarietà in varianza XE "varianza".

Nella figura 2.2, in cui è riportata la serie trasformata attraverso le differenze prime e dodicesime, possiamo notare come la stazionarietà in media sia stata raggiunta. Infine, nella figura 2.3 è riportata la serie preliminarmente trasformata tramite logaritmi e successivamente differenziata; l'evidenza grafica suggerisce come non esistano sostanziali differenze rispetto alla figura 2.2, a testimonianza dell'irrelevanza della trasformazione logaritmica.

Figura 2.1 - Indice grezzo della produzione industriale

Figura 2.2 - Indice della produzione industriale. Differenze prime e dodicesime

Figura 2.3 - Indice della produzione industriale. Logaritmi e differenze prime e dodicesime

Può, tuttavia, accadere che le trasformazioni adottate non permettano di ottenere una serie stazionaria; in questi casi può essere utile restringere il periodo temporale di analisi. Tale argomento sarà affrontato più approfonditamente nel paragrafo 5.4. La seguente tavola sintetizza, al solito, i valori dei parametri discussi in precedenza.

PARAMETRI VALORI ASSUMIBILI SIGNIFICATO

LAM	0, 1, -1	Trasformazione logaritmica XE " ^{Trasformazione logaritmica} " o meno, pre-test
IDIF	0, 1, 2, 3	Identificazione XE " ^{Identificazione} " automatica dell'ordine di differenziazione
D	0, 1, 2	Ordine delle differenze non stagionali
BD	0, 1	Ordine delle differenze stagionali

2.2.5 L'identificazione del modello

2.2.5 L'identificazione del modello

La procedura di Box e Jenkins XE "Box e Jenkins" per la costruzione di un modello ARIMA XE "ARIMA" consta di tre fasi: identificazione, stima XE "stima" e verifica delle ipotesi. Queste tre fasi possono essere iterate finché non si perviene al modello ritenuto migliore.

La fase di identificazione consiste nello specificare gli ordini del modello, che abbiamo indicato con (p, d, q) per la parte non stagionale e con (P, D, Q) per la parte stagionale. Per quanto riguarda la scelta degli ordini d e D , si rimanda al sottoparagrafo 2.2.4.2. Relativamente all'individuazione degli ordini dei polinomi AR ed MA XE "MA" stagionali e non stagionali si possono seguire due strade:

- modello identificato automaticamente; questo modo di procedere si basa sul confronto di vari modelli utilizzando criteri standard sviluppati in letteratura; in particolare viene considerato il cosiddetto criterio BIC XE "BIC" (*Bayesian Information Criterion*) (si veda Schwarz –1979 e Planas-1997, pag. 56). In questo contesto ci limitiamo a notare che tale criterio consiste nello scegliere il modello che riporta il minor valore della statistica espressa dal BIC. Per seguire questa opzione, si impiega, oltre al parametro IDIF=3, descritto sempre nel paragrafo 2.2.4.2, anche il parametro INIC=3. In questo caso nel file di output XE "file di output" saranno indicati gli ordini del modello scelto, unitamente al valore assunto dal BIC. Da notare che il parametro INIC stabilisce anche se inserire o meno una costante nella regressione XE "regressione". Se presente, ciò sarà segnalato nel file di output.
- modello identificato dall'utente; in questo caso l'identificazione degli ordini del modello avviene sulla base delle funzioni di autocorrelazione, riportate anche esse nel file di output XE "file di output"; questo modo di procedere richiede una conoscenza più approfondita della modellistica ARIMA XE "ARIMA", che esula dagli scopi di questa guida; si rimanda perciò a testi più specifici come Piccolo e Vitale (1981), Piccolo (1990), Mills (1990). Per imporre gli ordini del modello, si utilizzano i parametri P, D, Q, BP, BD, BQ , che rappresentano rispettivamente gli ordini del modello p, d, q, P, D, Q . In questa circostanza si assume $INIC=0$ che è il valore di default del parametro INIC e che istruisce TRAMO XE "TRAMO" affinché non effettui l'identificazione automatica XE "identificazione automatica" del modello. I valori di default per questi parametri sono $P=0, D=1, Q=1, BP=0, BD=1, BQ=1$, che danno luogo al modello ARIMA $(0, 1, 1)(0, 1, 1)_s$:

$$\tilde{N}\tilde{N}_s Z_t = (1 - qB)(1 - qB^s)a_t$$

noto in letteratura come modello *airline*, in quanto introdotto da Box e Jenkins XE "Box e Jenkins" (1970) per lo studio delle serie del numero dei passeggeri delle linee aeree internazionali. L'esperienza empirica dimostra che questo modello si adatta bene a diverse serie storiche economiche. Per quanto riguarda l'impiego della costante, esso è gestibile tramite il parametro IMEAN; ponendo tale parametro uguale ad 1 la costante viene inserita, ponendolo uguale a 0 non viene considerata.

PARAMETRI VALORI ASSUMIBILI SIGNIFICATO

IDIF	0, 2, 3, 5	Identificazione XE "Identificazione" automatica o meno dell'ordine di differenziazione delle componenti regolare e stagionale
INIC	0, 2, 3, 4, -1	Identificazione XE "Identificazione" automatica o meno dei polinomi regolari e stagionali
IMEAN	0, 1	Inserimento della costante di regressione XE "regressione"

2.2.6 Il parametro RSA

Un parametro che facilita il trattamento completamente automatico dei dati è il parametro RSA. Si tratta di un parametro di sintesi, nel senso che il suo utilizzo implica automaticamente la definizione di alcuni dei parametri sin qui illustrati ed in particolare quelli che consentono di effettuare il pre-test per la trasformazione logaritmica e per gli effetti di calendario, di rimuovere gli outlier, di interpolare le osservazioni mancanti e di identificare in maniera completamente automatica i modelli ARIMA XE "ARIMA" sulla serie depurata dagli effetti di natura deterministica.

All'utente sono offerte diverse possibilità:

- ponendo $RSA = 1$ il programma sottopone a verifica l'opportunità della trasformazione della serie in logaritmi, interpola i dati mancanti, effettua la correzione per i tre principali tipi di outlier, e stima XE "stima" il modello di default (*airline*).
- Ponendo $RSA = 2$ il programma effettua anche un pretest per i giorni lavorativi XE "giorni lavorativi" ($ITRAD=-1$) e effetto Pasqua XE "Pasqua" ($IEAST=-1$).
- Ponendo $RSA=3$ il programma automaticamente considera i parametri previsti per $RSA=1$ aggiungendo $INIC=3$ e $IDIF=3$, per cui procede anche all'identificazione

Il programma accetta anche un pre-test per i giorni lavorativi XE "giorni lavorativi" ($ITRAD=-1$) e effetto Pasqua XE "Pasqua" ($IEAST=-1$).

Ponendo $RSA=3$ il programma automaticamente considera i parametri previsti per $RSA=1$ aggiungendo $INIC=3$ e $IDIF=3$, per cui procede anche all'identificazione automatica XE "identificazione automatica" del modello ARIMA XE "ARIMA".

Impostando $RSA=4$ si seguono le stesse opzioni di $RSA=3$ con l'aggiunta del pre-test per i giorni lavorativi XE "giorni lavorativi" ($ITRAD=-1$) e per l'effetto Pasqua XE "Pasqua" ($IEAST=-1$).

$RSA=5$ è equivalente a $RSA=2$, ma con $ITRAD=-6$.

$RSA=6$ è equivalente a $RSA=4$, ma con $ITRAD=-6$.

$RSA=7$ è equivalente a $RSA=2$, ma con $ITRAD=-7$.

$RSA=8$ è equivalente a $RSA=4$, ma con $ITRAD=-7$.

Per una prima analisi della serie si suggerisce l'uso di questo parametro, esaminando la diagnostica di output e, quindi, effettuando, sulla base di questa, eventuali modifiche ai parametri di input.

Bisogna tener presente che l'utilizzo del parametro RSA e lo studio della serie mediante l'assegnazione dei valori per i parametri di interesse possono portare a risultati leggermente diversi anche se si identificano gli stessi modelli, in quanto si utilizzano valori di inizializzazione diversi per la stima XE "stima" dei coefficienti.

2.3 L'output di TRAMO XE "TRAMO"

2.3.1 Introduzione

Una volta immessi gli opportuni parametri nel file di input XE "file di input", si può mandare in esecuzione il programma con il comando TRAMO XE "TRAMO". Come evidenziato nel paragrafo 2.2.1, i risultati dell'elaborazione vengono immessi in un file il cui nome è uguale alla prima parola della prima riga del file di input (o alle sue prime otto lettere) con l'estensione out e che viene salvato in una subdirectory di TRAMO, denominata OUTPUT. Inoltre TRAMO, a conclusione dell'elaborazione, genera un set di grafici che possono essere visualizzati dall'utente e possono agevolare l'interpretazione dei risultati ottenuti.

Nelle successive pagine vengono descritti i contenuti principali del file di output XE "file di output" ed il set di grafici prodotto da TRAMO XE "TRAMO".

2.3.2 Descrizione ed interpretazione del file di OUTPUT.

Il file riporta in apertura la serie originale sottoposta ad analisi e le date relative alla Pasqua XE "Pasqua", di cui si è parlato nel sottoparagrafo 2.2.3.2. In aggiunta, vengono elencati i parametri di input usati dal programma ed i valori iniziali dei coefficienti del modello ARIMA XE "ARIMA", usati per la stima XE "stima". Successivamente, viene segnalato se è stata effettuata la trasformazione logaritmica e, in caso affermativo, viene riportata la serie storica trasformata. Viene poi riportata la serie differenziata XE "serie differenziata" (eventualmente calcolata sulla serie trasformata in logaritmi), le sue autocorrelazioni XE "autocorrelazioni" totali e parziali. Analogamente vengono segnalate le eventuali correzioni per gli effetti deterministici e, se il modello ARIMA è scelto automaticamente, ne vengono indicati gli ordini.

Terminato l'elenco degli elementi costitutivi del modello, sono riportati gli aspetti relativi alla stima XE "stima" ed alla diagnostica. In particolare, vengono elencate le stime dei coefficienti del modello ARIMA XE "ARIMA" con i relativi errori standard e le statistiche t (t-ratio XE "t-ratio"). Queste ultime sono calcolate dividendo il valore del parametro stimato per il suo errore standard e seguono asintoticamente una distribuzione Normale standardizzata. Come regola generale, si può ritenere significativamente diverso da zero un parametro quando il t-ratio è maggiore di 2 in valore assoluto.

I valori in corrispondenza di AR1 e di MA1 indicano le stime dei coefficienti della parte autoregressiva AR e della parte a media mobile MA XE "MA" non stagionale, mentre quelli in corrispondenza di AR2 e MA2 indicano le stime dei coefficienti AR e MA stagionali.

Successivamente viene riportata la matrice di correlazione degli stimatori dei coefficienti (sotto la voce "CORRELATIONS OF THE ESTIMATES"). Si noti come un'elevata correlazione possa indicare una errata specificazione del modello evidenziando, di fatto, un legame lineare tra i coefficienti e, di conseguenza, la necessità di provvedere ad una semplificazione, anche se tutti i coefficienti risultano significativi.

L'output riporta poi i valori dei due criteri di selezione automatica, ossia l'AIC XE "AIC" (Akaike Information Criterion) e il BIC XE "BIC" (Bayesian Information Criterion). Il primo è un ulteriore criterio di identificazione degli ordini del modello, e può essere utilizzato per valutarne la bontà in concomitanza con la valutazione del BIC. Tuttavia, è quest'ultimo criterio quello che prevalentemente viene impiegato nella letteratura più recente. In generale, un miglioramento nella selezione del modello sarà indicata da valori di AIC e BIC inferiori.

di identificazione degli ordini del modello, e può essere utilizzato per valutarne la bontà in concomitanza con la valutazione del BIC. Tuttavia, è quest'ultimo criterio quello che prevalentemente viene impiegato nella letteratura più recente. In generale, un miglioramento nella selezione del modello sarà indicata da valori di AIC e BIC inferiori.

Sotto la voce ESTIMATES OF REGRESSION PARAMETERS CONCENTRATED OUT OF THE LIKELIHOOD, sono riportate le stime dei coefficienti relativi alla parte deterministica. In particolare MU indica l'eventuale presenza di una costante, TRAD individua la presenza dei coefficienti dei regressori relativi ai giorni lavorativi $XE^{giorni\ lavorativi}$, EAST il coefficiente relativo all'effetto Pasqua XE^{Pasqua} , mentre OUT fa riferimento ai coefficienti relativi agli outlier. Infine, REG indica i coefficienti relativi a regressori imposti dall'utente.

In corrispondenza di ogni coefficiente stimato vengono riportati l'errore standard ed il t-ratio $XE^{t-ratio}$; dopo è illustrata la matrice di covarianze dei coefficienti stimati.

Una parte fondamentale dell'output è rappresentata dalla diagnostica sui residui, in quanto essi, affinché il modello rappresenti adeguatamente la serie analizzata, devono rispettare le ipotesi di partenza, ossia devono avere media zero, essere incorrelati e distribuiti normalmente.

Nell'output viene dunque riportata la serie dei residui ed immediatamente dopo le relative statistiche test ("TEST-STATISTICS ON RESIDUALS"). Innanzitutto, si verifica se essi hanno la media non significativamente diversa da zero utilizzando il T-VALUE (=MEAN/ST.DEV OF MEAN). Al solito, questa ipotesi è rifiutata se esso è maggiore di 2 in valore assoluto.

Per la verifica dell'ipotesi di normalità si considera il test parametrico di Jarque-Bera, indicato come "NORMALITY TEST" (si veda Planas, 1997, pagg. 58-59), che combina i momenti terzo (SKEWNESS, che nel caso della distribuzione normale è uguale a 0) e quarto (KURTOSIS, che nel caso della distribuzione normale è pari a 3). Sotto l'ipotesi nulla che l'errore si distribuisca normalmente, la statistica test di Jarque-Bera si distribuisce come una variabile casuale chi-quadro con 2 gradi di libertà; in pratica, si rifiuta l'ipotesi di normalità al livello di confidenza del 5% se la statistica test è superiore a 5,99.

Un'ulteriore indicazione può venire dall'errore quadratico medio $XE^{errore\ quadratico\ medio}$ dei residui ("MSE OF RESID"); tanto più esso è piccolo, tanto meglio il modello si adatta ai dati; può dunque essere un utile strumento quando si confrontano modelli alternativi. Può inoltre risultare utile considerare anche i valori di SKEWNESS e KURTOSIS assieme ai loro errori standard. Una regola empirica consiste nel controllare che il corrispondente valore "desiderato" (0 per la SKEWNESS e 3 per la KURTOSIS) cada nell'intervallo [valore stimato ± 2 errore standard]. Va ricordato che un valore elevato della KURTOSIS segnala, generalmente, la presenza di valori anomali e può quindi suggerire di rivedere le specifiche relative al controllo degli outlier.

Per la verifica dell'assenza di autocorrelazione dei residui, è di fondamentale importanza l'analisi delle funzioni di autocorrelazione empirica totale e parziale (rispettivamente sotto le voci "AUTOCORRELATIONS" e "PARTIAL AUTOCORRELATIONS"). Per un appropriato uso delle stesse si rimanda a Piccolo (1990). Un'ulteriore test viene utilizzato per verificare l'ipotesi che le prime m autocorrelazioni $XE^{autocorrelazioni}$ siano nulle; il test in questione è quello di Ljung-Box $XE^{Ljung-Box}$. La corrispondente statistica, quando le autocorrelazioni sono calcolate per i residui di un modello ARMA, si distribuisce come una variabile casuale chi-quadro con (m-q-p) gradi di libertà. Per default m è pari a 24 per le serie mensili, a 4f per le altre serie infrannuali (dove f=2, 3, 4, 6) e a 8 per le serie annuali. Il valore assunto dalla statistica test e la sua distribuzione vengono riportate nell'output di TRAMO XE^{TRAMO} di seguito alla funzione di autocorrelazione totale. Se il valore della statistica è superiore al valore critico derivato dalle tavole della distribuzione chi-quadro con (m-q-p) gradi di libertà, si rifiuta l'ipotesi nulla di assenza di autocorrelazione dei residui.

L'assenza di autocorrelazione nei residui non esclude che questi possano presentare un legame di tipo non lineare; a questo proposito utili indicazioni possono derivare dall'analisi dei residui al quadrato. Infatti, se la statistica test di Ljung-Box $XE^{Ljung-Box}$ calcolata su questi ultimi è maggiore della medesima statistica calcolata sui residui non trasformati, si può sospettare la presenza di un legame non lineare (si veda Planas, 1997, pagg. 57-58). Il programma TRAMO XE^{TRAMO} riporta quindi la serie dei residui al quadrato con il relativo errore standard e il test di Ljung-Box. L'eventuale differenza tra le due statistiche test, quella sui residui originali e quella sui residui al quadrato, costituirà effettivamente un segnale di non linearità $XE^{non\ linearità}$ quando la prima determina l'accettazione dell'ipotesi nulla di incorrelazione e la seconda la rifiuta.

Le successive righe dell'output riportano le seguenti serie:

- a) "INTERPOLATED SERIES"; è la serie originale in cui eventuali valori mancanti sono stati stimati mediante il modello adottato. Dunque, il programma TRAMO XE^{TRAMO} è anche in grado di utilizzare serie che presentano dati mancanti;
- b) "LINEAR SERIES"; è la serie depurata dagli effetti di regressione $XE^{regressione}$, ossia

di utilizzare serie che presentano dati mancanti;

b) "LINEAR SERIES"; è la serie depurata dagli effetti di regressione XE "regressione", ossia

$$z_t = (y_t - \sum_{-1}^k \beta_i x_{it})$$

, che verrà usata dal programma SEATS XE "SEATS" per procedere alla scomposizione XE "scomposizione" del modello;

c) eventuali serie rappresentanti gli effetti di regressione XE "regressione" (outlier, effetti di calendario ed eventuali effetti di regressori esterni).

2.3.3 L'analisi dei grafici

I grafici delle diverse serie riportate nel file di output XE "file di output" possono essere visualizzati digitando dal prompt di DOS (sempre partendo dalla directory TRAMO XE "TRAMO") il comando **GRAPH**. Appairà una schermata con i gruppi di grafici disponibili in TRAMO, del tipo:

TRAMO XE "TRAMO" GRAPHS	
AVAILABLE GROUPS	
1 SERIES	2 ACF
3 REGOUTSE	99 END

Selezionando il gruppo di interesse, saranno disponibili i grafici riportati nella seguente tabella:

Grafici disponibili in SERIES	
1 ORIGINAL SERIES	2 RESIDUALS
3 INTERPOLATED SERIES	4 FORECAST OF THE SERIES
5 LINEAR SERIES	6 REG. CORR. SERIES
7 S.E. CORR. SERIES	77 COMPARE WITH OLD MODELS
88 OVERLAY	99 END
Grafici disponibili in ACF	
1 ACF RESIDUALS	2 PACF RESIDUALS
3 ACF DIFF. SERIES	77 COMPARE WITH OLD MODELS
88 OVERLAY	99 END
Grafici disponibili in REGOUTSE	
1 REG1	2 TRAD1
3 TRAD2	4 EAST
5 TOTAL TRAD-DAY	6 TOTAL SPEC. EFF.
7 TOTAL REG.	8 TOTAL REG-O-S.E.
99 END	88 OVERLAY

Tra i numerosi grafici disponibili, risulta interessante l'analisi dei grafici relativi ai residui. Essi infatti possono evidenziare valori anomali o bruschi cambiamenti di livello che potrebbero segnalare una errata specificazione del modello; il grafico delle autocorrelazioni XE "autocorrelazioni" globali e parziali ("ACF" e "PACF") può suggerire in corrispondenza di quali ritardi si manifesta autocorrelazione significativa. A questo scopo, ossia per individuare residui o autocorrelazioni significativamente diverse da zero, è utile fare riferimento alle bande di confidenza riportate per questi grafici.

Un'opzione utile è rappresentata da "COMPARE WITH OLD MODELS"; con essa è possibile confrontare i grafici attuali con quelli relativi ad eventuali modelli precedentemente stimati. Per rendere disponibile questa opzione, una volta usciti dalle schermate dei grafici, è necessario digitare **Y** alla domanda: "DO YOU WANT TO CREATE A BACKUP COPY OF THESE GRAPHICS? (Y/N)"; così i grafici verranno memorizzati dal programma e resi disponibili per i successivi confronti.

E' anche possibile stampare i grafici visualizzati; per fare ciò bisognerà eseguire il comando **EGALASER** dal prompt prima del comando **GRAPH**; successivamente bisognerà digitare **GRAPH - P**. Così facendo, tutti i grafici che si visualizzeranno saranno salvati nel file *graph.lhj* e potranno essere stampati. Da notare che il file *graph.lhj* non è un file temporaneo; pertanto in esso verranno cumulati i grafici relativi a diverse sessioni di lavoro. Dunque, se non si è più interessati ad essi, si consiglia di cancellare questo file alla fine di ogni sessione o di salvarlo con altro nome.

cancellare questo file alla fine di ogni sessione o di salvarlo con altro nome.

2.4 Esempio: analisi dell'indice generale della produzione industriale

Consideriamo come esempio la serie mensile dell'indice generale della produzione industriale per il periodo gennaio 1981-aprile 1999. Tale serie è riportata nel floppy disk allegato, nel file *esempio 1*, per consentire all'utente di ripercorrere i vari passi dell'esercizio e poter così analizzare l'intero output.

Al fine di acquisire indicazioni preliminari, possiamo procedere in maniera completamente automatica usando il solo parametro $RSA=4$ (**Modello 1**). In tabella 2.1 sono riportate alcune delle statistiche derivanti dal file di output XE "file di output" che, in questo caso, si chiama *ipigengt.out*. È immediato notare che la serie è stata trasformata mediante logaritmo, e che il modello prescelto è un ARIMA XE "ARIMA" (0,1,1)(0,1,1) con un regressore per i giorni lavorativi XE "giorni lavorativi" ed uno per l'effetto Pasqua XE "Pasqua"; non è presente la costante e sono stati identificati due outlier di tipo additivo in corrispondenza di agosto 1984 ed agosto 1995. Il test di normalità dei residui è soddisfacente (statistica di Jarque-Bera pari a 2,176), mentre i test di Ljung-Box XE "Ljung-Box" evidenziano problemi di autocorrelazione nei residui e nei residui al quadrato (sono infatti pari rispettivamente a 36,02 e 43,65, perciò cadono entrambi nella regione di rifiuto). La differenza tra le due statistiche di Ljung-Box segnala anche problemi di non linearità XE "non linearità".

L'analisi della serie può essere raffinata considerando un ulteriore regressore che tenga conto delle festività nazionali e religiose, ossia il regressore denominato in precedenza *holidays* (cfr. sottoparagrafo 2.2.3.4). Aggiungendo le istruzioni di input descritte nella casella all'incrocio fra la riga **Input** e la colonna **Modello 2** della tabella 2.1, si ottiene un modello analogo, con un notevole miglioramento in termini di BIC XE "BIC" ed AIC XE "AIC"; l'ipotesi di normalità è però rifiutata e, anche se non è presente l'autocorrelazione nei residui, emerge una evidente non linearità XE "non linearità" degli stessi.

Data la natura della serie, che come si è visto risente molto dell'effetto dei giorni lavorativi XE "giorni lavorativi" effettivi, può risultare significativa la presenza dell'effetto della diversa lunghezza del mese per gli anni bisestili. Con il **Modello 3** viene esplicitata questa ultima opzione ($ITRAD=2$); e vengono identificati automaticamente gli ordini del modello ARIMA XE "ARIMA" (con $INIC=3$ ed $IDIF=3$) e la presenza di eventuali outlier ($IATIP=1$) si impone la presenza dell'effetto Pasqua XE "Pasqua" ($IEAST=1$) e del regressore *holidays* ($IREG=1$ e successivo comando REG); mentre non viene inserita la costante ($IMEAN=0$); al contrario si verifica di nuovo l'opportunità della trasformazione logaritmica ($LAM=-1$). I risultati dimostrano la significatività del nuovo regressore senza, però, sostanziali cambiamenti nel valore delle altre statistiche rispetto al modello 2. Una ulteriore indicazione può trarsi dall'analisi grafica della serie grezza (figura 2.1); essa evidenzia una sostanziale stabilità nella varianza XE "varianza" della stessa e pertanto, nonostante i risultati del test di TRAMO XE "TRAMO" sull'opportunità della trasformazione logaritmica, si può tentare di utilizzare un modello, indicato in tabella 2.1 come **Modello 4**, senza questa trasformazione (per maggiori approfondimenti sull'uso della trasformazione logaritmica si rimanda al capitolo 5 di questa guida). In pratica, viene usato lo stesso input precedente, imponendo però $LAM=1$. In questo caso non vengono più identificati outlier, mentre risultano superati i problemi di non normalità e non linearità XE "non linearità" dei residui. Si osservi inoltre che i criteri AIC XE "AIC" e BIC XE "BIC" non consentono ora confronti con i precedenti modelli, poiché in questo caso i dati non sono trasformati.

In figura 2.4 sono riportati i grafici dei residui ottenuti con i 4 diversi modelli (i dati sono reperibili nel file *resid.out*, contenuto nella directory TRAMO XE "TRAMO" \GRAPH\SERIES). Come si può notare, il quarto modello riesce a rappresentare meglio i dati rispetto ai modelli 1, 2 e 3, per i quali alcuni valori superano notevolmente la banda di confidenza.

In definitiva, il modello 4 appare essere quello che meglio rispetta le ipotesi di partenza e meglio si adatta alle osservazioni.

Tabella 2.1: Sintesi dell'output di TRAMO XE "TRAMO" per la serie IPIGENGT

	MODELLO 1	MODELLO 2	MODELLO 3	MODELLO 4
Input	\$INPUT RSA=4\$	\$INPUT RSA=4 IREG=1\$ \$REG IUSER=-1 NSER=1 ILONG=244 REGEFF=2\$ HOL_MF81	\$INPUT LAM=-1 INIC=3 IDIF=3 IATIP=1 ITRAD=2 IEAST=1 IMEAN=0 IREG=1\$ \$REG IUSER=-1 NSER=1 ILONG=244 REGEFF=2\$ HOL_MF81	\$INPUT LAM=1 INIC=3 IDIF=3 IATIP=1 ITRAD=2 IEAST=1 IMEAN=0 IREG=1\$ \$REG IUSER=-1 NSER=1 ILONG=244 REGEFF=2\$ HOL_MF81

		REGEFF=2\$ HOL_MF81	NSER=1 I LONG=244 REGEFF=2\$ HOL_MF81	NSER=1 I LONG=244 REGEFF=2\$ HOL_MF81
Logaritmo	Si	Si	Si	No
ARIMA XE " ^{ARIMA*}	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)
Giorni lav.	1 regressore	1 regressore+holidays	1 regressore+effetto anno bisestile+holidays	1 regressore+effetto anno bisestile+holidays
Effetto Pasqua XE " ^{Pasqua*}	Si	Si	Si	Si
Outlier XE " ^{Outlier*}	Ago 84 AO XE " ^{AO*} Ago 95 AO XE " ^{AO*}	Ago 84 AO XE " ^{AO*} Ago 95 AO XE " ^{AO*}	Ago 84 AO XE " ^{AO*} Ago 95 AO XE " ^{AO*}	Nessuno
AIC XE " ^{AIC*}	-889,69	-918,4	-922,27	881,70
BIC XE " ^{BIC*}	-7,063	-7,183	-7,185	1,493
Jarque-Bera	2,18	12,80	14,60	1,08
Ljung-Box XE " ^{Ljung-Box*} res.	36,02	21,70	18,57	22,81
Ljung-Box XE " ^{Ljung-Box*} res. ^2	43,65	69,26	58,68	23,31

Figura 2.4 – Residui XE "^{Residui*} dei modelli sperimentati sull'indice generale della produzione industriale

segue : **Figura 2.4 – Residui XE "Residui" dei modelli sperimentati sull'indice generale della produzione industriale**

APPENDICE: Analisi dettagliata del file IPIGENG.T.OUT

In questa Appendice viene descritto più nel dettaglio il file di output XE "file di output" di TRAMO XE "TRAMO" relativo all'analisi dell'indice grezzo della produzione industriale; in particolare, si tratta dell'output ottenuto con il modello 4 visto nel paragrafo 2.4. Lo scopo è quello di facilitare l'utente nella ricerca dei vari strumenti diagnostici e degli indicatori descritti in questo secondo capitolo.

A fini espositivi, il file è stato suddiviso in varie parti numerate, ognuna seguita da un commento (in corsivo). In alcuni punti dell'output sono riportate delle serie storiche in sequenza; per comodità di lettura vengono riportate solo le prime e le ultime osservazioni delle serie stesse, inserendo puntini sospensivi al posto dei valori intermedi.

1

TIME SERIES REGRESSION MODELS WITH ARIMA XE "ARIMA" ERRORS, MISSING VALUES AND OUTLIERS.
BETA VERSION (*)

BY

VICTOR GOMEZ & AGUSTIN MARAVALL

with the programming assistance of G. CAPORELLO

(*) Copyright : V. GOMEZ, A. MARAVALL (1994,1996)

SERIES TITLE=ipigengt

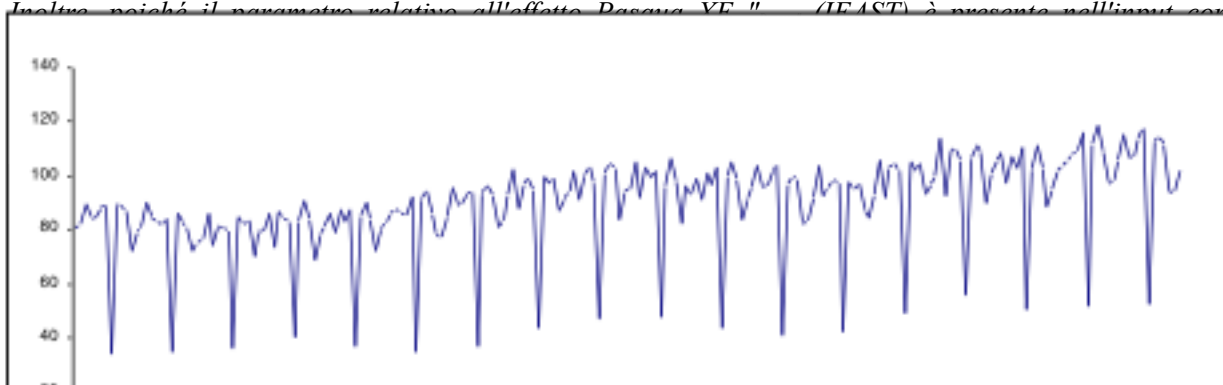
Dopo l'intestazione viene riportato il titolo della serie, indicato nella prima riga del file SERIE.

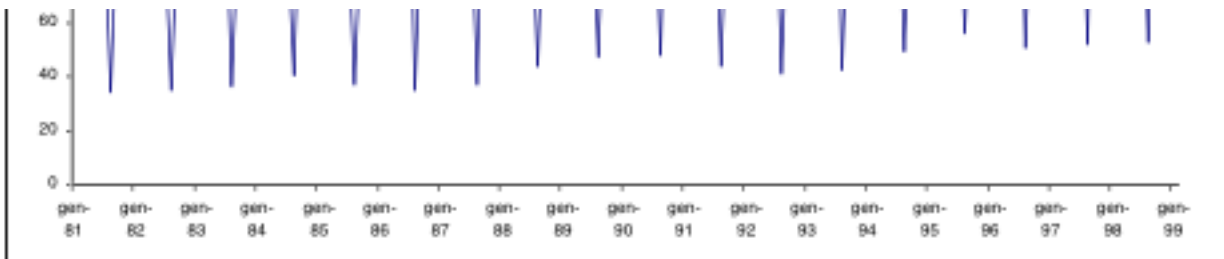
2

SINCE LONGER FORECAST FUNCTION IS REQUIRED
BY SEATS XE "SEATS", NPRED CHANGED TO (24)

IF IEAST NOT = 0, IDUR SHOULD BE > 0
IDUR CHANGED TO: 6

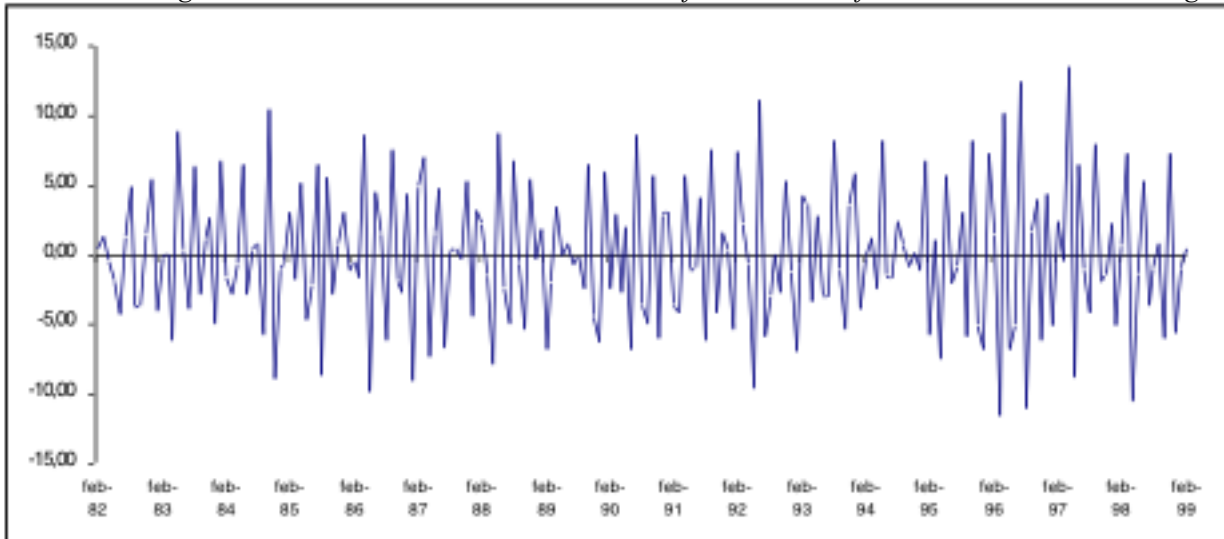
Nell'input non era presente il parametro NPRED, che perciò per default è pari a 0. Poiché SEATS XE "SEATS" sfrutta per le sue analisi anche le previsioni, questo parametro è automaticamente posto pari a 24. Inoltre, poiché il parametro relativo all'effetto Pasqua XE "IEAST" è presente nell'input con





1999 94.600 103.300 116.800 104.300

E' la serie originale di 220 osservazioni contenuta nel file SERIE in forma matriciale, con le righe



1987	APRIL	19
1988	APRIL	3
1989	MARCH	26
1990	APRIL	15
1991	MARCH	31
1992	APRIL	19
1993	APRIL	11
1994	APRIL	3
1995	APRIL	16
1996	APRIL	7
1997	MARCH	30
1998	APRIL	12
1999	APRIL	4
2000	APRIL	23
2001	APRIL	15

Sono le date della Pasqua XE "Pasqua" per gli anni a cui fa riferimento la serie (1981-1999) e per gli anni a cui si riferisce il periodo di previsione (essendo di 24 mesi, esso sarà relativo al 2000 ed al 2001). Queste date servono al programma per valutare l'effetto Pasqua.

5

MODEL PARAMETERS				
MQ= 12	IMEAN= 0	LAM= 1	D= 1	BD= 1
P= 0	BP= 0	Q= 1	BQ= 1	IREG= 4
ITRAD= 2	IEAST= 1	IDUR= 6	M= 36	QM= 24
INCON= 0	NBACK= 0	NPRED= 24	INTERP= 2	INIT= 0
IFILT= 2	IDENSC= 1	IROOT= 2	INIC= 3	ICONCE= 1
ICDET= 1	IATIP= 1	IMVX= 0	IDIF= 3	PG= 0
AIO= 2	INT1= 1	NT2= 220	RSA= 0	SEATS= XE "SEATS"
				2
VA= 3.50	TOL= 0.100E-03		PC= 0.143E+00	
NOADMISS= 1	BIAS= 1	SMTR= 0		

2
 VA= 3.50 TOL= 0.100E-03 PC= 0.143E+00
 NOADMISS= 1 BIAS= 1 SMTR= 0
 THTR= -0.400 RMOD= 0.500 MAXBIAS= 0.500

Sono i principali parametri di input con il corrispondente valore; quelli non impostati dall'utente nel file SERIE sono riportati con il loro valore di default.

6

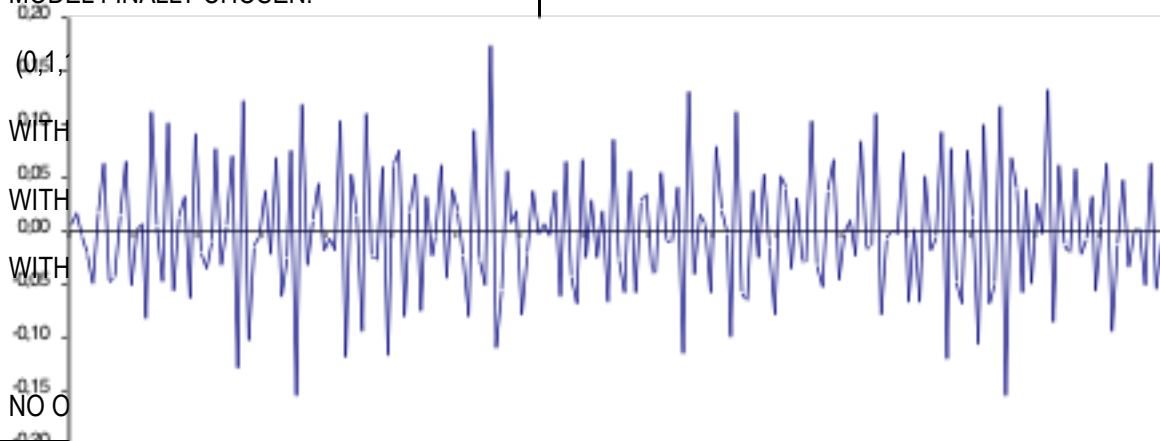
TH = -0.10
 BTH = -0.10
 NUMBER OF INITIAL OBS. = 13

Sono indicati i valori iniziali della stima XE "stima" dei parametri (per default sono posti pari a -0.10; TH è riferito alla componente XE "componente" MA XE "MA" non stagionale, BTH alla componente MA stagionale XE "MA stagionale") ed il numero di osservazioni iniziali a cui condizionarsi per la stima del modello (in questo caso 13 essendo 13 i ritardi considerati dal modello identificato da TRAMO XE "TRAMO" , come segnalato nel successivo punto 7).

7

AUTOMATIC MODEL IDENTIFICATION BEGINS

MODEL FINALLY CHOSEN:



Il programma ha effettuato l'identificazione automatica XE "identificazione automatica" del modello. Come si può notare, il modello scelto è l'airline privo della costante; sono stati inoltre effettuati i test per la verifica della significatività degli effetti di calendario e sono stati inseriti nel modello sia i regressori relativi ai giorni lavorativi XE "giorni lavorativi" , sia quello relativo all'effetto Pasqua XE "Pasqua" . La ricerca automatica degli outlier ha dato esito negativo.

8

METHOD OF ESTIMATION: EXACT MAXIMUM LIKELIHOOD

PARAMETER	ESTIMATE	STD ERROR	T RATIO	LAG
MA1 1	-0.51160	0.61071E-01	-8.38	1
MA2 1	-0.58029	0.63538E-01	-9.13	12

Viene segnalato il metodo di stima XE "stima" usato (la massima verosimiglianza XE "massima verosimiglianza") e riportate le stime dei coefficienti della parte stocastica con i relativi errore standard t-ratio XE "t-ratio" . Si può notare come entrambi i coefficienti siano significativi (il t-ratio in valore assoluto è nettamente superiore a 2).

9

REGULAR MA XE "MA" INVERSE ROOTS ARE
 NO. REAL P. IMAG.P. MODULUS ARGUMENT PERIOD
 1 0.51160 0.00000E+00 0.51160 0.00000E+00 -

REGULAR MA XE " MA^* INVERSE ROOTS ARE				
NO.	REAL P.	IMAG.P.	MODULUS	ARGUMENT PERIOD
1	0.51160	0.00000E+00	0.51160	0.00000E+00 -

SEASONAL MA XE " MA^* INVERSE ROOTS ARE				
NO.	REAL P.	IMAG.P.	MODULUS	ARGUMENT PERIOD
1	0.58029	0.00000E+00	0.58029	0.00000E+00 -

Vengono riportate le radici inverse dei polinomi MA XE " MA^* , sia per la parte regolare (non stagionale) che per quella stagionale.

10

CORRELATIONS OF THE ESTIMATES	
1.0000	0.0322
0.0322	1.0000

E' la matrice delle correlazioni degli stimatori; chiaramente sulla diagonale principale il valore è pari ad 1 (correlazione tra il coefficiente e se stesso), mentre gli altri due valori si riferiscono alla correlazione tra il coefficiente MA1 e quello MA2. Essa è abbastanza vicina a 0, a testimonianza di assenza di correlazione tra gli stimatori.

11

AIC XE " AIC^*	881.7916
BIC XE " BIC^*	1.4933
FINAL VALUE OF OBJECTIVE FUNCTION:	809.76
ITERATIONS:	2
NUMBER OF FUNCTION EVALUATIONS:	7

Sono riportati i valori dei criteri AIC XE " AIC^* e BIC XE " BIC^* . Si ricorda che essi sono utili ai fini del confronto con altri modelli; il modello che presenta il valore AIC e BIC minore, è da considerarsi il migliore come compromesso tra parsimonia e adattamento ai dati. Inoltre è riportato il valore la funzione obiettivo, ossia la funzione di verosimiglianza, ed il numero di iterazioni per massimizzare la funzione.

12

ESTIMATES OF REGRESSION PARAMETERS CONCENTRATED OUT OF THE LIKELIHOOD				
PARAMETER	VALUE	ST. ERROR	T VALUE	
REG 1	-2.2274	0.26625	-8.37	
TRAD 1	0.78221	0.03452	22.66	
TRAD 2	2.2952	0.79505	2.89	
EAST 1	-1.5788	0.50145	-3.15	

E' riportata la stima XE " $stima$ dei coefficienti relativi ai regressori non stocastici: REG 1 si riferisce alle holidays, TRAD 1 al regressore relativo ai giorni lavorativi XE " $giorni\ lavorativi$, TRAD 2 al regressore relativo alla diversa lunghezza dell'anno bisestile, EAST 1 al regressore relativo all'effetto Pasqua XE " $Pasqua$. Inoltre sono riportati i corrispondenti errore standard e t-ratio XE " $t-ratio$: si può notare come tutti i regressori siano significativi.

13

COVARIANCE MATRIX OF ESTIMATORS	
---------------------------------	--

13

COVARIANCE MATRIX OF ESTIMATORS			
0.709E-01	-0.663E-03	0.274E-03	-0.219E-01
-0.663E-03	0.119E-02	-0.227E-03	-0.460E-04
0.274E-03	-0.227E-03	0.632E+00	0.733E-02
-0.219E-01	-0.460E-04	-0.733E-02	0.251E+00

E' la matrice di varianze e covarianze dei soli regressori non stocastici (la diagonale principale contiene le varianze, che si possono ottenere elevando al quadrato gli errori standard di cui al punto 12).

14

NUMBER OF WHITE NOISE RESIDUALS							
							203
WHITE NOISE RESIDUALS							
-3.3088	1.7920	0.6947	-0.6580	-1.1771	-1.8047	2.0137	-1.9237
0.3317	1.2743	-2.6778	-0.3904	1.8058	1.3818	3.9696	-0.8943
...							
...							
0.0521	0.4555	-4.2911	1.7907	-1.1976	0.1417	-5.1468	0.5986
-1.0076	1.8666	-2.1145					

Sono riportati i 203 residui (il numero di osservazioni 220 meno le 13 osservazioni iniziali, a cui la stima XE " stima " è condizionata, e i quattro regressori esterni).

15

TEST-STATISTICS ON RESIDUALS	

MEAN=	0.0516737
ST.DEV.=	0.1383709
OF MEAN	
T-VALUE=	0.3734
NORMALITY TEST=	1.087 (CHI-SQUARED(2))
SKEWNESS=	-0.1502 (SE = 0.1719)
KURTOSIS=	2.8042 (SE = 0.3438)
SUM OF SQUARES=	789.5505
DURBIN-WATSON=	2.0919
STANDARD ERROR=	1.981946
OF RESID.	
MSE OF RESID.=	3.928112

Sono i test statistici sui residui per verificare le ipotesi di partenza relative ai disturbi del modello. Si ricorda che essi dovrebbero essere distribuiti normalmente, indipendentemente ed identicamente, con media zero e varianza XE " varianza " costante. Viene calcolata la media dei residui e la sua deviazione standard, ottenendo (per divisione) il t-value. Esso è inferiore a 2 in valore assoluto, dunque è verificata l'ipotesi che i residui abbiano media nulla. Viene riportato il test di normalità di Jarque-Bera; la cui statistica è pari a 1.087 e dunque ben al di sotto del valore critico (corrispondente ad un livello di significatività del 5%) della distribuzione chi-quadro con 2 gradi di libertà (5.99), per cui anche l'ipotesi di distribuzione normale può essere accettata. Sono inoltre riportati il valore del test di asimmetria e della curtosi con il relativo errore standard. La somma dei quadrati dei residui (SUM OF SQUARES) così come il MSE possono essere utili per confronti con altri modelli (tanto più piccoli sono questi indicatori, tanto più i valori stimati sono vicini a quelli reali). E' infine riportato sia il test

asimmetria e della curtosi con il relativo errore standard. La somma dei quadrati dei residui (SUM OF SQUARES) così come il MSE possono essere utili per confronti con altri modelli (tanto più piccoli sono questi indicatori, tanto più i valori stimati sono vicini a quelli reali). E' infine riportato sia il test di Durbin-Watson XE "Durbin-Watson" per la verifica di autocorrelazione di ordine 1 (i residui sono incorrelati se questo valore è vicino a 2) sia l'errore standard dei residui, che ci fornisce una misura della variabilità degli stessi.

16

AUTOCORRELATIONS												

	-0.0557	0.0282	0.0620	0.0469	0.0556	0.0429	-0.0944	0.0804	0.1402	0.0045	-0.0324	0.0502
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	0.64	0.80	1.61	2.07	2.72	3.11	5.00	6.38	10.59	10.59	10.82	11.37
PV	-1.00	-1.00	0.21	0.36	0.44	0.54	0.42	0.38	0.16	0.23	0.29	0.33
	-0.0236	-0.1563	-0.0259	-0.0812	-0.0961	-0.0512	-0.0135	0.0366	-0.0279	-0.0011	-0.0624	-0.0342
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	11.49	16.87	17.02	18.49	20.55	21.14	21.18	21.49	21.66	21.66	22.56	22.84
PV	0.40	0.15	0.20	0.19	0.15	0.17	0.22	0.26	0.30	0.36	0.37	0.41
	-0.0077	-0.0213	-0.0637	0.0398	-0.0890	0.0529	0.1790	0.0318	0.0165	0.0135	0.1054	-0.0498
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	22.85	22.96	23.92	24.29	26.19	26.86	34.62	34.86	34.93	34.98	37.73	38.34
PV	0.47	0.52	0.52	0.56	0.51	0.53	0.22	0.25	0.29	0.33	0.26	0.28

Viene fornita la serie delle autocorrelazioni XE "autocorrelazioni" dei residui per i primi 36 ritardi. Dopo ogni sequenza di 12 ritardi sono riportati i corrispondenti errori standard e test di Ljung-Box XE "Ljung-Box" (Q) con m posto pari al numero dei ritardi di riferimento (cfr. paragrafo 2.3.2). Per ogni test Q è riportato il p -value (PV), ossia l'area, sottesa alla distribuzione della statistica Q , alla destra del valore assunto dalla statistica stessa. Un $PV > 0.05$ porta a non rifiutare l'ipotesi nulla di incorrelazione dei residui ad un livello di significatività del 5%. Come si può notare, l'ipotesi di incorrelazione viene accettata per tutti i ritardi. Si noti anche che per i primi due ritardi è riportato un PV pari a -1.00; essendo negativo è chiaramente un valore simbolico che indica l'impossibilità di applicare il test in questione perché il numero di gradi di libertà sarebbe inferiore a zero.

17

LJUNG-BOX Q VALUE OF ORDER 24 IS 22.84 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED(22)
--

Viene riportato il valore della statistica Q ritenuto più importante ai fini della verifica, ossia quello con $m=24$, insieme alla corrispondente distribuzione; chiaramente questo valore corrisponde al Q in corrispondenza del ritardo 24 (si veda il punto 16), per cui è possibile controllare immediatamente quale sia il suo PV .

18

PARTIAL AUTOCORRELATIONS												

	-0.0557	0.0251	0.0652	0.0537	0.0584	0.0435	-0.1001	0.0577	0.1461	0.0235	-0.0456	0.0302
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
	-0.0335	-0.2002	-0.0495	-0.0494	-0.1125	-0.0808	0.0319	0.0819	-0.0401	0.0486	0.0109	-0.0406
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
	-0.0025	0.0522	-0.0395	-0.0191	-0.1021	0.0200	0.1612	0.0537	0.0300	-0.0041	0.0997	-0.0725
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702

Sono riportate le autocorrelazioni XE "autocorrelazioni" parziali con i rispettivi errori standard per i primi 36 ritardi.

19

APPROXIMATE TEST OF RUNS ON RESIDUALS

NUM.DATA= 203
NUM.(+)= 102
NUM.(-)= 101
NUM DI RUNS= 112

NUM.DATA= 203
 NUM.(+)= 102
 NUM.(–)= 101
 NUM.RUNS= 113
 T-VALUE= 1.5480

APPROXIMATE TEST OF RUNS ON AUTOCORRELATION FUNCTION

NUM.DATA= 36
 NUM.(+)= 18
 NUM.(–)= 18
 NUM.RUNS= 15
 T-VALUE= -1.3528

Vengono riportati i test dei run sui residui e sulle autocorrelazioni XE "autocorrelazioni" dei residui; è questo un test di tipo non parametrico che sfrutta i cosiddetti run, ossia sequenze di numeri con lo stesso segno. Se non ci fosse autocorrelazione il numero di run con segno positivo dovrebbe essere approssimativamente uguale a quello dei run con segno negativo. Nell'output è riportato il numero dei residui ed autocorrelazioni con segno positivo, NUM. (+), e di quelli con segno negativo, NUM. (–), il numero dei run ed il valore della statistica test (T-VALUE) che è distribuita asintoticamente come una normale standard. I valori trovati di 1.5480 e -1.3528 consentono di accettare l'ipotesi di incorrelazione.

20

SQUARED RESIDUALS:

AUTOCORRELATIONS

	-0.0750	-0.0701	0.0318	0.0178	0.0430	-0.1119	0.0205	0.0664	0.0010	-0.0021	-0.1492	-0.0016
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	1.16	2.18	2.39	2.45	2.84	5.49	5.57	6.52	6.52	6.52	11.34	11.34
PV	-1.00	-1.00	0.12	0.29	0.42	0.24	0.35	0.37	0.48	0.59	0.25	0.33
	0.0046	-0.0967	-0.0480	-0.0350	0.0480	-0.0459	-0.0322	0.0299	-0.1110	0.0087	-0.0584	0.1339
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	11.34	13.41	13.92	14.19	14.70	15.18	15.41	15.62	18.43	18.45	19.24	23.41
PV	0.41	0.34	0.38	0.44	0.47	0.51	0.57	0.62	0.49	0.56	0.57	0.38
	-0.0806	-0.0506	0.0481	-0.0103	-0.0195	-0.0582	0.0853	0.0773	-0.0522	-0.0376	-0.0063	0.1214
SE	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702	0.0702
Q	24.92	25.53	26.07	26.10	26.19	27.01	28.77	30.22	30.89	31.24	31.24	34.92
PV	0.35	0.38	0.40	0.46	0.51	0.52	0.48	0.45	0.47	0.51	0.55	0.42

E' riportata la serie delle autocorrelazioni XE "autocorrelazioni" relative ai residui al quadrato per 36 ritardi, con relativo errore standard, statistica Q e PV . Per la valutazione della loro significatività valgono le stesse considerazioni svolte al punto 16; anche in questo caso notiamo che nessuno dei test rifiuta l'ipotesi di incorrelazione dei residui al quadrato, per cui possiamo escludere la presenza di strutture non lineari nei residui.

21

LJUNG-BOX Q VALUE OF ORDER 24 IS 23.41 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED(22)

Viene riportato il valore della statistica Q per i residui al quadrato ritenuto più importante ai fini della verifica della non linearità XE "non linearità", ossia quello in cui $m=24$, insieme alla corrispondente distribuzione; anche in questo caso questo valore corrisponde al Q in corrispondenza del ritardo 24 (si veda il punto 20), per cui è possibile controllare immediatamente quale sia il suo PV .

22

FORECASTS:

FORECASTS:					
ORIGIN:	220	NUMBER:	24		
OBS FORECAST STD ERROR ACTUAL RESIDUAL FORECAST STD ERROR (TR. SERIES) (ORIGINAL SERIES)					
221	108.236	1.99527		108.236	1.99527
222	112.290	2.20660		112.290	2.20660
...					
...					
243	114.979	6.03300		114.979	6.03300
244	100.188	6.19698		100.188	6.19698

Vengono riportate le previsioni per i successivi 24 periodi (osservazioni 221-244) ed i corrispondenti errori standard. Da notare che se si fosse effettuata la trasformazione logaritmica, le previsioni e gli errori standard sarebbero stati riportati sia per la serie trasformata che per quella originale.

23

INTERPOLATED SERIES												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	80.700	81.900	90.000	84.500	84.500	88.200	89.000	34.200	89.800	89.300	86.700	72.200
1982	79.500	81.100	90.600	84.600	82.700	82.100	84.300	34.500	86.300	82.200	81.100	72.100
...												
...												
1998	98.400	105.800	115.300	106.800	107.600	114.200	117.900	52.500	113.500	114.400	111.400	94.100
1999	94.600	103.300	16.800	104.300								

Viene riportata la serie interpolata, ossia quella originale con l'interpolazione, effettuata in base al modello, degli eventuali valori mancanti. In questo caso, poiché non erano presenti valori mancanti, la serie interpolata coinciderà con quella originale.

24

LINEAR SERIES												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	83.319	82.474	89.602	85.952	89.856	86.636	86.653	37.329	88.236	89.691	87.873	74.308
1982	84.856	81.674	87.464	86.052	85.829	80.536	84.691	34.891	84.736	85.329	81.763	71.981
...												
...												
1998	103.246	106.374	114.902	108.252	112.956	112.636	115.553	55.629	111.936	114.791	112.573	96.208
1999	102.184	103.874	114.453	104.963								

E' riportata la serie linearizzata XE "serie linearizzata", ossia la serie originale depurata dall'effetto dei regressori deterministici (in questo caso: l'effetto dei giorni di calendario e l'effetto holidays). Questa è la serie che sarà analizzata da SEATS XE "SEATS" per la scomposizione XE "scomposizione".

25

TOTAL TRADING DAY - EASTER EFFECT FACTORS												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC

TOTAL TRADING DAY - EASTER EFFECT FACTORS

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-0.391	-0.574	0.398	0.775	-3.129	1.564	2.347	-3.129	1.564	-0.391	-1.173	2.347
1982	-3.129	-0.574	3.136	0.775	-3.129	1.564	0.391	-0.391	1.564	-3.129	1.564	2.347
...												
...												
1998	-0.391	-0.574	0.398	0.775	-3.129	1.564	2.347	-3.129	1.564	-0.391	-1.173	2.347
1999	-3.129	-0.574	2.347	1.564								

E' riportata la serie degli effetti dei giorni di calendario (giorni lavorativi XE "giorni lavorativi", anno bisestile, Pasqua XE "Pasqua"). Una volta effettuata la scomposizione XE "scomposizione", questo effetto sarà assegnato da SEATS XE "SEATS" alla componente XE "componente stagionale XE "componente stagionale".

26

TOTAL REGRESSION EFFECT

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-2.227	0.000	0.000	-2.227	-2.227	0.000	0.000	0.000	0.000	0.000	0.000	-4.455
1982	-2.227	0.000	0.000	-2.227	0.000	0.000	0.000	0.000	0.000	0.000	-2.227	-2.227
...												
...												
1998	-4.455	0.000	0.000	-2.227	-2.227	0.000	0.000	0.000	0.000	0.000	0.000	-4.455
1999	-4.455	0.000	0.000	-2.227								

E' riportata la serie degli effetti di regressione XE "regressione" totale (in questo caso le sole holidays). Poiché nell'input è stato usato il parametro REGEFF=2, questo effetto sarà allocato da SEATS XE "SEATS" alla componente XE "componente stagionale XE "componente stagionale". Chiaramente la somma delle serie dei punti 24, 25, 26 darà la serie originaria.

27

ELAPSED TIME IS 1.9200 "

E' il tempo impiegato da TRAMO XE "TRAMO" per le analisi effettuate (solo 1,92 secondi).

CAPITOLO 3

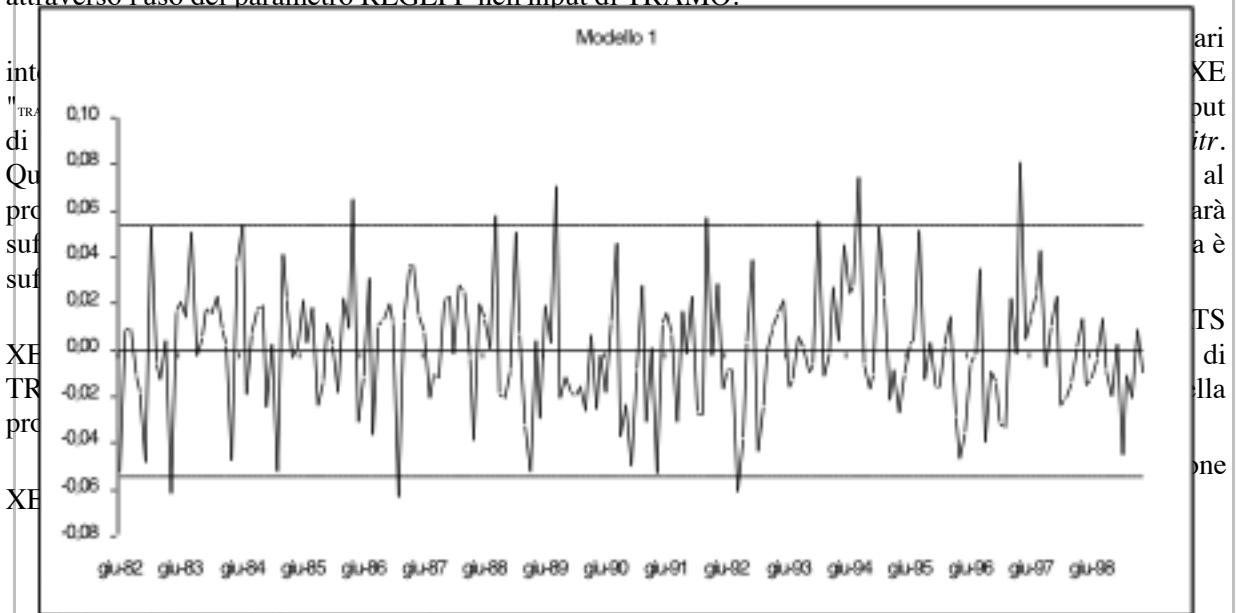
SEATS XE "SEATS"

3.1 Introduzione

Una volta depurata la serie di interesse dagli effetti deterministici, si ottiene la cosiddetta serie linearizzata XE "serie linearizzata" z_t , che si ipotizza essere il risultato dell'interazione di tre componenti non osservabili: il ciclo-trend XE "ciclo-trend" T_t , la componente XE "componente" stagionale XE "componente stagionale" S_t e la componente irregolare XE "componente irregolare" I_t . La prima rappresenta la dinamica di medio-lungo periodo del fenomeno oggetto di studio, la seconda la componente stagionale le sue fluttuazioni periodiche regolari infra-annuali e la terza la componente irregolare tiene conto dei movimenti erratici. Il loro legame può essere di tipo additivo ($z_t = T_t + S_t + I_t$) o moltiplicativo ($z_t = T_t S_t I_t$). Dato che la scomposizione XE "scomposizione" moltiplicativa XE "scomposizione moltiplicativa" è additiva nei logaritmi, d'ora in avanti si farà riferimento esclusivamente a quella additiva.

Il programma SEATS XE "SEATS" realizza la scomposizione XE "scomposizione" della serie z_t , seguendo il cosiddetto approccio *model-based*, ossia ipotizzando che ogni componente XE "componente" segua un modello ARIMA XE "ARIMA", deducibile da quello seguito dalla serie z_t . In particolare, la componente irregolare XE "componente irregolare" segue sempre un processo *white-noise*.

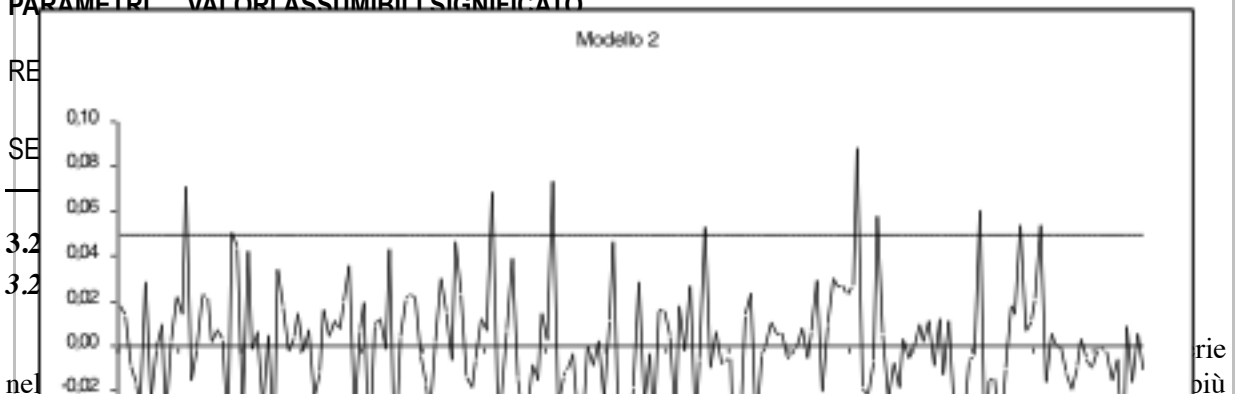
La scomposizione XE "scomposizione" effettuata riguarda solo la parte stocastica della serie, z_t , quindi è necessario allocare gli effetti deterministici individuati da TRAMO XE "TRAMO" tra le varie componenti. Tale operazione può essere effettuata automaticamente dal programma attraverso una serie di regole specifiche che saranno descritte nel paragrafo 3.3.1.4, oppure, come illustrato nel paragrafo 2.2.3.4 attraverso l'uso del parametro REGEFF nell'input di TRAMO.

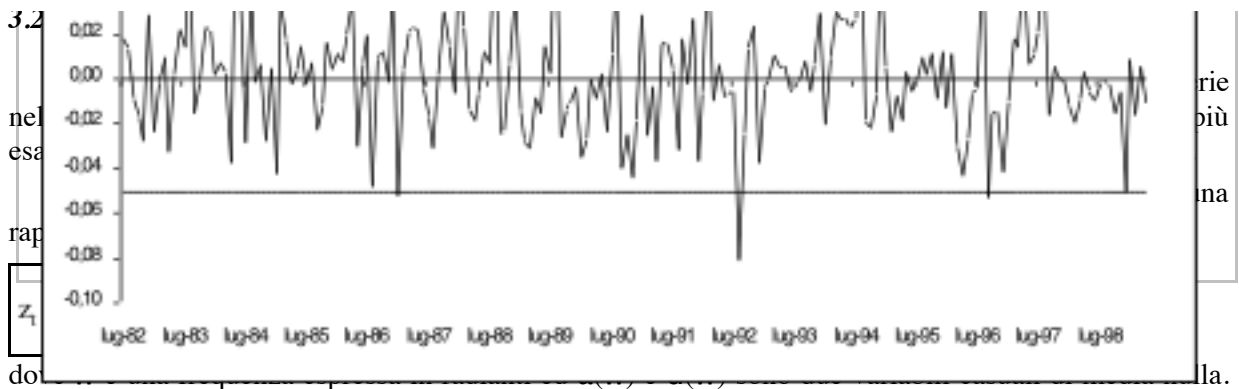


- 3) stima XE "stima" delle componenti,
- 4) analisi degli errori.

Questa descrizione sarà preceduta da brevi cenni teorici necessari per la sua comprensione.

PARAMETRI VALORI ASSUMIBILI SIGNIFICATO





In altri termini, il processo z_t può essere visto come somma di infinite componenti sinusoidali (seni e coseni) di frequenza compresa tra 0 e ρ (per dettagli, si veda Hamilton, 1995, cap. 6). Per semplicità, supponiamo nel prosieguo che $m=0$.

La funzione di autocovarianza del processo z_t può essere espressa nel seguente modo:

$$\gamma_k = E(z_t, z_{t-k}) = \int_{-\pi}^{\pi} e^{i\omega k} f(\omega) d\omega \quad (3.2)$$

dove:

$$e^{i\omega k} = \cos(\omega k) + i \sin(\omega k)$$

$$i = \sqrt{-1}$$

e:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k}$$

La funzione $f(\omega)$ è nota in letteratura come funzione di densità spettrale. Nel seguito ci riferiremo ad essa chiamandola semplicemente spettro. In particolare, ponendo $k=0$ nella (3.2), si ottiene la varianza XE "varianza" γ_0 del processo stocastico:

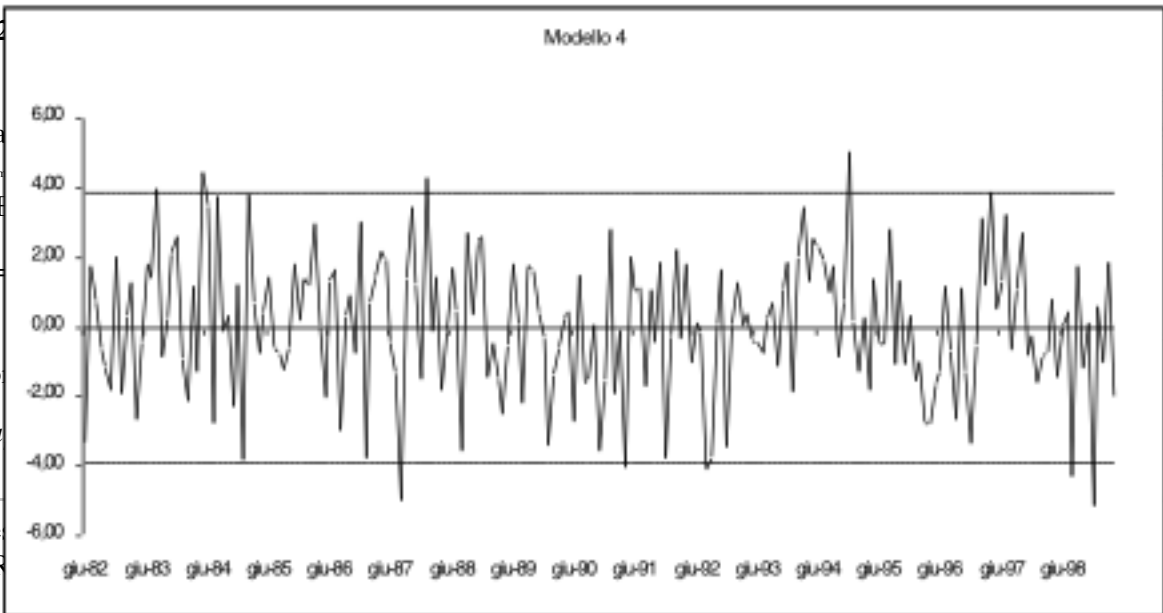
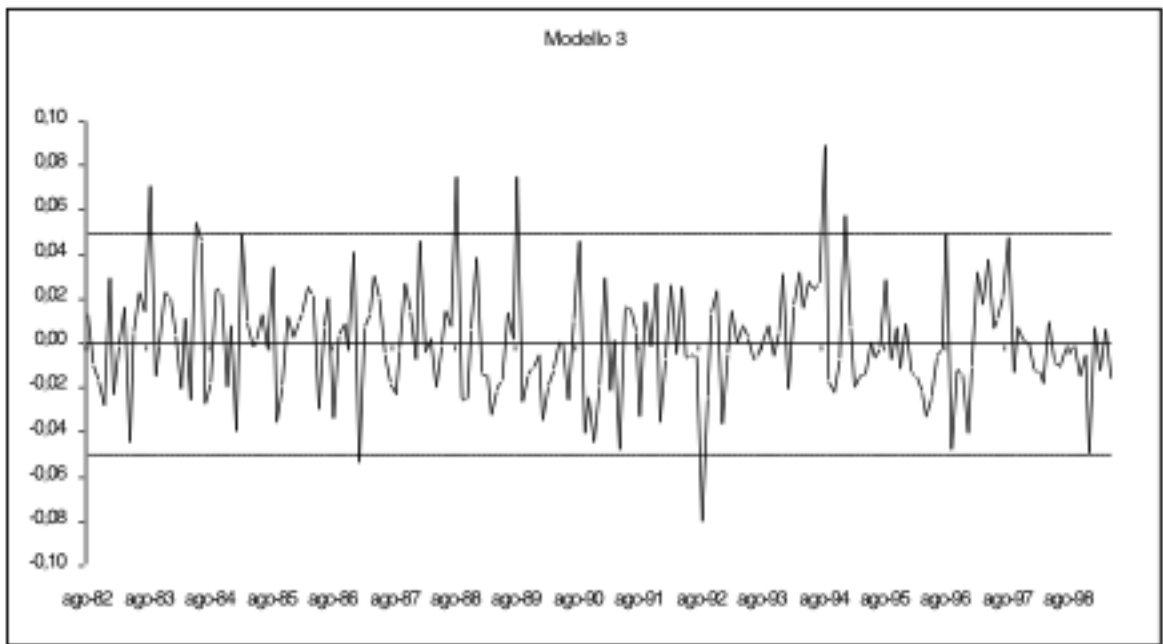
$$\gamma_0 = \int_{-\pi}^{\pi} f(\omega) d\omega$$

In base a questa espressione, l'interpretazione dello spettro diventa chiara: l'area sottesa alla funzione di densità spettrale f nell'intervallo $[-\rho, \rho]$ è pari alla varianza XE "varianza" del processo, mentre $f(\omega)d\omega$ rappresenta la parte della varianza dell'intero processo dovuta alle componenti con frequenza compresa nell'intervallo $[\omega, \omega+d\omega]$. Pertanto, un elevato valore dello spettro in corrispondenza della frequenza ω significa che la componente XE "componente" del processo z_t caratterizzata da quella particolare frequenza contribuisce in larga parte a spiegare la varianza del processo, mentre un basso valore di $f(\omega)$ indica che il contributo di quella componente è piccolo.

Tali elementi, in questo contesto notevolmente semplificati, si rivelano fondamentali per effettuare la destagionalizzazione delle serie storiche in quanto la presenza di componenti stagionali nel processo viene evidenziata da valori elevati nello spettro in corrispondenza delle frequenze stagionali. Se con s si indica il periodo della stagionalità (ossia $s=12$ per serie mensili, $s=4$ per serie trimestrali, ecc.), allora le frequenze stagionali saranno date da $\omega_j = 2\pi j/s$, per $j=1, 2, \dots, s/2$. Analogamente, la presenza di un eventuale trend è rivelata da un picco nella funzione di densità spettrale in corrispondenza della frequenza $\omega = 0$. Una funzione di densità spettrale parallela all'asse delle ascisse è propria di un processo di tipo *white noise* XE "white noise".

A titolo di esempio, si riportano in figura 3.1 gli spettri relativi alla serie osservata, al ciclo-trend XE "ciclo-trend" ed alla componente XE "componente" stagionale XE "componente stagionale" dell'indice generale della produzione industriale, analizzato secondo i parametri esplicitati nel modello 4 del capitolo 2 di questa guida.

Figura 3.1 – Spettro XE "Spettro" della serie dell'indice generale della produzione industriale e spettri teorici del ciclo-trend e della componente XE "componente" stagionale XE "componente stagionale"



3.2

sia
"com
XE
 $z_t =$
co
sa
de
TR

z_t
KE
are
lla
rie
da

$$\varphi(B)z_t = \theta(B)a_t$$

$$\varphi(B)z_t = \vartheta(B)a_t,$$

dove:

$$\varphi(B) = \phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D$$

e:

$$\vartheta(B) = \theta(B)\Theta(B^s).$$

e a_t è un white noise XE "white noise".

SEATS XE "SEATS" scompone la serie z_t in maniera tale che ogni componente XE "componente" segua un modello ARIMA XE "ARIMA":

$$\varphi_T(B)T_t = \vartheta_T(B)a_{Tt}, \quad \varphi_S(B)S_t = \vartheta_S(B)a_{St} \quad e \quad I_t = u_t$$

dove a_{Tt}, a_{St}, u_t sono errori *white noise* XE "white noise" e i due polinomi $\varphi_T(B)$ e $\varphi_S(B)$ non hanno radici in comune.

Pertanto la serie z_t viene espressa da SEATS XE "SEATS" come:

$$z_t = \frac{\vartheta(B)}{\varphi(B)} a_t = \frac{\vartheta_T(B)}{\varphi_T(B)} a_{Tt} + \frac{\vartheta_S(B)}{\varphi_S(B)} a_{St} + u_t \quad (3.3)$$

In altri termini, si ipotizza che le componenti seguano modelli ARIMA XE "ARIMA" coerenti con il modello ARIMA seguito dalla serie z_t . Se lo spettro di tutte le componenti è non negativo, allora la scomposizione XE "scomposizione" è detta ammissibile. Poiché ad ogni modello ARIMA per la serie z_t sono associate infinite scomposizioni ammissibili del tipo (3.3), esiste cioè un problema di sovraidentificazione, è necessario identificare la sola scomposizione ammissibile XE "scomposizione ammissibile" utile per l'identificazione e stima XE "stima" delle componenti. E' possibile dimostrare che le scomposizioni ammissibili sono uguali tra loro a meno di quote della varianza XE "varianza" totale imputabili alle diverse componenti. Quindi una delle possibili procedure di identificazione consiste nello scegliere quella scomposizione ammissibile che assegna la maggior quota di varianza alla componente XE "componente" irregolare XE "componente irregolare". Tale scomposizione viene detta *canonica* ed è quella effettivamente utilizzata da SEATS XE "SEATS".

Si noti che se il modello identificato in TRAMO XE "TRAMO" non ammette alcuna scomposizione XE "scomposizione" ammissibile XE "scomposizione ammissibile", SEATS XE "SEATS" modifica opportunamente tale modello ottenendone uno con tale proprietà.

3.2.3 Stima delle componenti

Una volta identificati i modelli ARIMA XE "ARIMA" per le singole componenti, SEATS XE "SEATS" effettua la loro stima XE "stima" mediante l'impiego di opportuni filtri lineari, derivati dai modelli ARIMA stessi. Ricordiamo che un filtro lineare $l(B)$ invariante rispetto al tempo può essere espresso come:

$$\lambda(B) = \sum_{k=-\infty}^{\infty} \lambda_k B^k \quad (3.4)$$

dove i pesi λ_k sono numeri reali tali che $\sum_{k=-\infty}^{\infty} \lambda_k^2 < \infty$. L'uso dei filtri consente di esplicitare gli stimatori di ogni singola componente XE "componente".

In particolare SEATS XE "SEATS" utilizza il filtro di Wiener-Kolmogorov XE "Wiener-Kolmogorov" (WK), che è simmetrico, infinito e convergente (si veda Planas, 1997, cap. 6) e consente di ottenere lo stimatore ottimale delle componenti. Poiché la proprietà della convergenza implica che il valore assoluto dei pesi tende a zero al tendere di k a $\pm \infty$, nelle applicazioni vengono utilizzati filtri "troncati": il punto in cui viene troncato il filtro dipende dal numero di osservazioni di cui si dispone. Le principali caratteristiche dello stimatore di WK sono essenzialmente le seguenti:

- a) minimizza l'errore quadratico medio XE "errore quadratico medio" della stima XE "stima" della componente XE "componente", ossia è un *MMSE (Minimum Mean Squared Error) Estimator*;
- b) rappresenta bene il segnale in corrispondenza delle frequenze per le quali la densità spettrale della componente XE "componente" è alta mentre, al contrario, lo sottostima in corrispondenza delle frequenze

- a) minimizza l'errore quadratico medio (errore quadratico medio) della stima della componente XE "componente", ossia è un *MMSE (Minimum Mean Squared Error) Estimator*;
- b) rappresenta bene il segnale in corrispondenza delle frequenze per le quali la densità spettrale della componente XE "componente" è alta mentre, al contrario, lo sottostima in corrispondenza delle frequenze per cui la densità è bassa. Questo peculiare funzionamento dello stimatore fa sì che la sua variabilità sia sempre inferiore a quella della componente;
- c) gli stimatori delle componenti possono essere correlati tra loro nonostante le componenti non lo siano.

Per una generica componente XE "componente" C_t , il corrispondente stimatore è dato da:

$$E(C_t | z_1, K, z_T) = \alpha_C(B) z_t$$

dove $\alpha_C(B)$ è un polinomio nell'operatore ritardo B .

Gli stimatori ottenuti possono essere di vario tipo a seconda dell'istante temporale t per il quale si vuole ottenere una stima XE "stima". Indicando con T l'ultimo istante temporale per il quale si dispone di osservazioni, gli stimatori possono essere suddivisi in:

- a) stimatori preliminari se $t < T$ ($C_{t|T}$);
- b) stimatori *concurrent* XE "concurrent" se $t = T$ ($C_{T|T}$);
- c) stimatori finali se $T < t < \infty$ (C_t).

Inoltre, quando $t > T$ lo stimatore rappresenta la previsione della componente XE "componente" rispetto a $t - T$ periodi in avanti.

3.2.4 Errori di stima XE "stima"

In corrispondenza di ogni stimatore, possono essere individuati diversi tipi di errore:

- 1) errore di stima XE "errore di stima" XE "stima" finale: $\delta_{C_t} = C_t - C_t$;
- 2) errore di stima XE "errore di stima" XE "stima" preliminare XE "preliminare": $\delta_{C_{t|T}} = C_t - C_{t|T}$;
- 3) errore di revisione XE "errore di revisione": $r_{C_{t|T}} = C_t - C_{t|T}$;
- 4) errore di revisione XE "errore di revisione" totale XE "errore di revisione totale": $r_{C_{T|T}} = C_t - C_{T|T}$.

Si noti che una generica componente XE "componente" C_t può essere espressa come:

$$C_t = \alpha_C(B) z_t = \alpha_C(B) \frac{\vartheta(B)}{\varphi(B)} a_t = \psi_C(B) a_t = \sum_{i=-\infty}^{\infty} \psi_{C_i} a_{t+i} \quad (3.5)$$

La (3.5) si ottiene applicando il filtro di WK alla serie osservata z_t e sostituendo quest'ultima con l'espressione (3.3). Il relativo stimatore finale XE "stimatore finale" C_t si ottiene sostituendo ai polinomi $J(B)$, $j(B)$, $a_C(B)$ le corrispondenti stime, mentre gli stimatori preliminare XE "preliminare" e *concurrent* XE "concurrent" si ottengono troncando opportunamente il termine superiore della sommatoria. Le stime dei coefficienti ψ_{C_i} sono dette *psi-weights*.

L'errore di revisione XE "errore di revisione" può anche essere espresso come segue:

$$r_{C_{t|T}} = \sum_{j=-\infty}^{\infty} \psi_{C_j} a_{t+j} - \sum_{j=-\infty}^{T-t} \psi_{C_j} a_{t+j} = \sum_{j=T-t+1}^{\infty} \psi_{C_j} a_{t-j}$$

Supponendo, senza perdita di generalità, che la varianza XE "varianza" dei disturbi a_t sia unitaria, la varianza dell'errore di revisione XE "errore di revisione" $r_{C_{t|T}}$ può essere espressa come:

$$\text{var}(r_{C_{t|T}}) = \sum_{j=T-t+1}^{\infty} \psi_{C_j}^2$$

In aggiunta, è possibile dimostrare che (si veda Pierce XE "Pierce", 1980):

$$\lim_{T \rightarrow \infty} \text{var}(r_{C_{t|T}}) = \lim_{T \rightarrow \infty} \sum_{j=T-t+1}^{\infty} \psi_{C_j}^2 = 0 \quad (3.6)$$

In virtù della (3.6) l'errore di revisione XE "errore di revisione" diventa trascurabile dopo un certo numero di periodi. che equivale a dire che lo stimatore preliminare XE "preliminare" converge a quello finale.

$$\lim_{T \rightarrow \infty} \text{var}(r_{C|T}) = \lim_{T \rightarrow \infty} \sum_{j=T-1+1}^{\infty} \Psi_{Cj} = 0 \quad (3.0)$$

In virtù della (3.6) l'errore di revisione XE "errore di revisione" diventa trascurabile dopo un certo numero di periodi, che equivale a dire che lo stimatore preliminare XE "preliminare" converge a quello finale.

3.3 L'output di SEATS XE "SEATS"

Analogamente al programma TRAMO XE "TRAMO", l'output di SEATS XE "SEATS" è contenuto in un apposito file nella subdirectory di SEATS denominata OUTPUT (anche in questo caso il nome del file sarà uguale alla prima parola della prima riga del file di input XE "file di input", o alle sue prime otto lettere, con l'estensione *out*).

L'output di SEATS XE "SEATS" è molto ricco di informazioni, anche se per l'interpretazione di alcune di esse sono necessarie particolari conoscenze statistiche. Naturalmente, considerando le finalità della presente trattazione, non sarà possibile dare indicazioni sull'interpretazione dell'intero output, per cui ci si limiterà a rilevare gli aspetti salienti e più importanti relativamente alla sola destagionalizzazione. In aggiunta, il commento completo di un file di output XE "file di output" è rimandato al capitolo 4, dove verranno illustrati alcuni esempi relativi a serie reali.

3.3.1 Descrizione ed interpretazione del file di OUTPUT

3.3.1.1 Parte prima: stima XE "stima" del modello ARIMA XE "ARIMA"

Nella prima parte dell'output (denominata "ARIMA XE "ARIMA" ESTIMATION") vengono richiamati i risultati della stima XE "stima" del modello ottenuti con la procedura TRAMO XE "TRAMO", per cui vengono riportate:

- la serie grezza;
- la serie relativa agli effetti deterministici;
- la serie depurata dagli effetti deterministici, data dalla serie in (a) meno la serie in (b). Nel seguito dell'output tale serie viene indicata come serie Z;
- i parametri di input;
- eventuali serie trasformate tramite logaritmi e differenze;

Inizialmente, SEATS XE "SEATS" provvede a "centrare" la serie da analizzare, nel senso che la trasforma sottraendo da ogni osservazione la media, ottenendo così una nuova serie con media nulla. Su questa serie (ovviamente stazionaria) vengono calcolati i valori delle autocorrelazioni XE "autocorrelazioni" totali e parziali ed i relativi *errori standard*. Viene quindi richiamato il modello ARIMA XE "ARIMA" scelto da TRAMO XE "TRAMO" con i parametri, da esso stimati, che serviranno per la scomposizione XE "scomposizione" della serie. Questo modello viene utilizzato per filtrare la serie e calcolare nuovamente i residui relativi alle prime osservazioni perse a seguito dell'utilizzo degli operatori di differenziazione. Poiché il calcolo è reimpostato *ex-novo*, i residui differiranno leggermente rispetto a quelli stimati da TRAMO.

A questo punto, SEATS XE "SEATS" produce le stesse diagnostiche sui residui fornite da TRAMO XE "TRAMO", aggiungendovi il test di Pierce XE "Pierce" (si veda Planas, 1997, pag. 57) che considera esclusivamente i ritardi stagionali (nel nostro caso il ritardo 12 ed il ritardo 24 per dati mensili, 4 e 8 per dati trimestrali e così via). La statistica relativa al test di Pierce segue approssimativamente la distribuzione chi-quadro con 2 gradi di libertà, per cui si rifiuterà l'ipotesi nulla di incorrelazione tra i residui ai ritardi stagionali se questa statistica supera il valore 5,99.

Data l'operazione di ricostruzione dei residui, descritta in precedenza, effettuata da SEATS XE "SEATS", può accadere che i test sui residui di TRAMO XE "TRAMO" e SEATS diano indicazioni discordanti; in tali casi si suggerisce di dare maggior peso ai test di TRAMO.

3.3.1.2 Parte seconda: derivazione dei modelli per le componenti

Nella seconda parte dell'output ("DERIVATION OF THE MODELS FOR THE COMPONENTS") vengono esplicitati i risultati della scomposizione XE "scomposizione" canonica XE "scomposizione canonica" in termini di modelli per ogni singola componente XE "componente". Nei casi in cui il modello ARIMA XE "ARIMA" identificato e stimato da TRAMO XE "TRAMO" non ammetta alcuna scomposizione, SEATS XE "SEATS" modifica opportunamente il modello finché non trova una sua approssimazione che permetta tale scomposizione.

In ogni caso, il modello importato da TRAMO XE "TRAMO" per l'intera serie viene riscritto secondo la formulazione (3.3). Dopo la derivazione dei modelli per le singole componenti in base alla scomposizione XE "scomposizione" canonica XE "scomposizione canonica", SEATS XE "SEATS" provvede al calcolo della varianza XE "varianza" dei disturbi relativi a ciascuna componente XE "componente" espressa come frazione della varianza dei disturbi del modello aggregato. Dai valori delle varianze è possibile stabilire quanta parte dell'irregolarità della serie totale è spiegata dalle singole componenti (si noti che la componente ciclo-

scomposizione XE "scomposizione canonica", SEATS XE "SEATS" provvede al calcolo della varianza XE "varianza" dei disturbi relativi a ciascuna componente XE "componente" espressa come frazione della varianza dei disturbi del modello aggregato. Dai valori delle varianze è possibile stabilire quanta parte dell'irregolarità della serie totale è spiegata dalle singole componenti (si noti che la componente ciclo-trend XE "ciclo-trend" in SEATS è denominata semplicemente TREND). Vengono poi forniti i coefficienti *psi-weights* relativi alle componenti ciclo-trend, stagionale, transitoria, destagionalizzata e irregolare, denominati rispettivamente PSIEP, PSIES, PSIEC, PSIEA, PSIUE. Si tenga conto che i ritardi con segno negativo sono riferiti ai periodi successivi al tempo t (che corrisponde al ritardo 0) mentre quelli con segno positivo ai periodi precedenti (cfr. Appendice, punto 25). Va notato che questa rappresentazione del filtro non è, in generale, simmetrica e che la somma degli *psi-weights* per ogni singola componente deve essere pari a:

- zero per i ritardi negativi;
- uno per il ritardo zero;
- i pesi della rappresentazione MA XE "MA" (¥) di Box-Jenkins del processo generatore per i ritardi positivi.

Di seguito nell'output sono riportati i coefficienti del filtro WK per ogni componente XE "componente" e, successivamente, le autocorrelazioni XE "autocorrelazioni" per le componenti teoriche ("COMPONENT"), per lo stimatore ("ESTIMATOR") e per la serie stimata ottenuta applicando il filtro troncato alla serie Z ("ESTIMATE"). Il confronto tra i diversi valori ottenuti per le componenti permette di comprendere le differenze indotte dal processo di stima XE "stima" sia nel caso teorico, in cui si potesse disporre di una serie infinita, sia nel caso effettivo. In particolare, la differenza tra le varianze delle diverse componenti COMPONENT, ESTIMATOR ed ESTIMATE, deve soddisfare quanto è atteso *ex-ante* dalla teoria, ossia che la varianza XE "varianza" della componente ESTIMATOR sia inferiore a quella di COMPONENT. Come si è visto, infatti, lo stimatore sottostima nel complesso la variabilità della componente; la varianza di ESTIMATE deve essere vicina a quella di ESTIMATOR in quanto tale vicinanza è una garanzia che il "taglio" del filtro non ha comportato ulteriore perdita di informazione.

Dopo le informazioni sulle autocorrelazioni XE "autocorrelazioni" e sulle componenti, l'output riporta le correlazioni incrociate, per ciascuna coppia di componenti, per gli stimatori e per le stime. Anche in questo caso, come regola generale, ci si deve attendere che le due misure siano sufficientemente vicine.

3.3.1.3 Parte terza: analisi degli errori

Nella terza parte dell'output ("ERROR ANALYSIS") è riportata l'analisi dei vari tipi di errore: l'errore finale XE "errore finale", l'errore di revisione XE "errore di revisione" e l'errore totale (cfr. par. 3.2.4). SEATS XE "SEATS" provvede a confrontare tra loro le varianze di tali errori. In particolare, è interessante verificare la velocità di convergenza della stima XE "stima" preliminare XE "preliminare" alla stima finale per TREND e SEASONAL ADJUSTED (alla voce "VARIANCE OF THE REVISION ERROR") in termini di riduzione dell'errore standard di revisione all'aumentare delle osservazioni. Chiaramente, la convergenza sarà avvenuta quando la riduzione percentuale è vicina al 100%.

Di seguito, viene illustrata una tavola contenente la varianza XE "varianza" dell'errore di revisione XE "errore di revisione" della componente XE "componente" stagionale XE "componente stagionale" ottenuta con proiezioni in avanti per un anno e la riduzione percentuale media del RMSE (Root Mean Squared Error) che si sarebbe ottenuta utilizzando la stima XE "stima" *concurrent* XE "concurrent". Occorre, tuttavia, fare una precisazione: quest'ultima, se da un lato sfrutta una maggiore quantità di informazione rispetto alle proiezioni, per cui risulta sicuramente più affidabile, dall'altro richiede che il modello venga stimato ogni volta che un nuovo dato si rende disponibile ed implica revisioni per tutti i periodi precedenti. Ovviamente, se la riduzione percentuale del RMSE che si avrebbe utilizzando la stima *concurrent* è di modesta entità, potrebbe essere conveniente sfruttare le previsioni della componente stagionale nel corso dell'anno. Questo argomento sarà trattato con maggiore precisione nel capitolo 5 di questa guida.

Un'altra utile indicazione può scaturire dal confronto tra le medie della serie originale, del ciclo-trend XE "ciclo-trend", della serie destagionalizzata XE "serie destagionalizzata" e della componente XE "componente" stagionale XE "componente stagionale", sia per l'intero periodo di analisi sia per gli ultimi tre anni. Ovviamente le medie del ciclo-trend e della serie destagionalizzata dovrebbero essere pari alla media della serie originaria, mentre la media della componente stagionale dovrebbe essere pari a zero.

3.3.1.4 Parte quarta: risultati finali

Nella quarta ed ultima parte di SEATS XE "SEATS", "ESTIMATES OF THE COMPONENTS (LEVELS)", sono riportati i risultati finali della scomposizione XE "scomposizione". Inizialmente vengono illustrate le serie puramente stocastiche ("ORIGINAL SERIES"), ossia la serie originaria corretta da TRAMO XE "TRAMO" per gli effetti deterministici, la componente XE "componente" stagionale XE "componente stagionale", il ciclo-trend XE "ciclo-

(LEVELS) , SONO

riportati i risultati finali della scomposizione XE "scomposizione". Inizialmente vengono illustrate le serie puramente stocastiche ("ORIGINAL SERIES"), ossia la serie originaria corretta da TRAMO XE "TRAMO" per gli effetti deterministici, la componente XE "componente" stagionale XE "componente stagionale", il ciclo-trend XE "ciclo-trend", la serie destagionalizzata XE "serie destagionalizzata" (tutte con i relativi *errori standard*) e la componente irregolare XE "componente irregolare". Successivamente, sotto la voce "DETERMINISTIC COMPONENT (from TRAMO)", sono riportate le serie degli eventuali effetti deterministici, precedentemente isolati dal programma TRAMO, ossia l'effetto Pasqua XE "Pasqua", l'effetto giorni lavorativi XE "giorni lavorativi", eventualmente, l'effetto dei regressori imposti dall'utente. Infine, i vari effetti deterministici vengono assegnati alle singole componenti, ottenendo le componenti finali (alla voce "FINAL COMPONENT"), costituite da "FINAL SEASONAL ADJUSTED SERIES", "FINAL TREND" e "FINAL SEASONAL" nonché le corrispondenti previsioni per i due anni successivi. L'assegnazione delle parti deterministiche alle varie componenti avviene in base ad alcuni criteri fissati da SEATS XE "SEATS", e cioè:

- 1) gli outlier di tipo LS XE "LS" ed i regressori imposti dall'utente con parametro $REGEFF=1$ sono assegnati alla componente XE "componente" ciclo-trend XE "ciclo-trend";
- 2) gli effetti di calendario (giorni lavorativi XE "giorni lavorativi", anni bisestili e Pasqua XE "Pasqua") ed i regressori imposti dall'utente con parametro $REGEFF=2$ sono assegnati alla componente XE "componente" stagionale XE "componente stagionale";
- 3) gli outlier di tipo AO XE "AO" e TC XE "TC" ed i regressori imposti dall'utente con parametro $REGEFF=3$ sono assegnati alla componente XE "componente" irregolare XE "componente irregolare";
- 4) i regressori imposti dall'utente con parametro $REGEFF=5$ sono assegnati alla componente XE "componente" transitoria XE "componente transitoria";
- 5) i regressori imposti dall'utente con parametro $REGEFF=0$ non sono assegnati ad alcuna componente XE "componente" (sono cioè componenti di regressione XE "regressione" separati). Si ricorda che per default $REGEFF=0$.

3.3.2 L'analisi dei grafici

I grafici delle varie serie riportate nel file di output XE "file di output" possono essere visualizzati digitando dal prompt di DOS (sempre partendo dalla directory SEATS XE "SEATS") il comando GRAPH. Apparirà una schermata con i gruppi di grafici disponibili in SEATS, strutturata come segue:

**SEATS XE "SEATS" GRAPHS
AVAILABLE GROUPS**

1 SERIES	2 ACF
3 SPECTRA	4 FILTERS
5 FORECAST	99 END

Digitando il numero relativo al gruppo di serie di interesse sarà possibile ottenere la loro visualizzazione. Dal gruppo SERIES si potranno così visualizzare i grafici di tutte le componenti; dal gruppo ACF i grafici delle funzioni di autocorrelazione delle varie componenti; dal gruppo SPECTRA i loro spettri; dal gruppo FILTERS i grafici dei filtri WK e degli *psi-weights*; dal gruppo FORECAST le previsioni per tutte le serie disponibili.

Chiaramente, per il produttore di dati, il maggior interesse risiede nel grafico delle serie destagionalizzate finali. Esse sono disponibili anche su file nella subdirectory di SEATS XE "SEATS" denominata GRAPHSERIES; in particolare:

- il file *safin.t* contiene la serie destagionalizzata XE "serie destagionalizzata" finale;
- il file *trfin.t* contiene la serie ciclo-trend XE "ciclo-trend" finale;
- il file *sfin.t* contiene la serie relativa alla componente XE "componente" stagionale XE "componente stagionale" finale;
- il file *irfin.t* contiene la componente XE "componente" irregolare XE "componente irregolare" finale.

Per quanto riguarda tutte le altre serie, si rimanda a Gomez e Maravall (1997), pagg. 88-92.

3.4 Esempio: l'analisi dell'indice generale della produzione industriale

In questa sezione si riprende l'esempio precedentemente analizzato nella parte seconda di questa guida, relativo all'indice generale della produzione industriale. Nella tabella 3.1 sono riportati i principali risultati derivanti dall'output di SEATS XE "SEATS" relativi ai quattro modelli stimati con TRAMO XE "TRAMO". In termini di diagnostica dei residui, si può notare come i risultati rispecchino quelli già osservati in TRAMO, ossia la presenza di una forte autocorrelazione e non linearità XE "non linearità" dei residui per i primi tre modelli ed il rifiuto dell'ipotesi di normalità con i modelli 2 e 3. In aggiunta, il test di Pierce XE "Pierce" conferma la presenza di autocorrelazione anche per i ritardi stagionali.

La seconda parte della tabella mette in luce la volatilità delle singole componenti; si noti come tutti

aggiunta, il test di Pierce XE "Pierce" conferma la presenza di autocorrelazione anche per i ritardi stagionali.

La seconda parte della tabella mette in luce la volatilità delle singole componenti; si noti come tutti i modelli attribuiscono una forte variabilità alla componente XE "componente irregolare XE "componente irregolare", mentre il ciclo-trend XE "ciclo-trend" appare meno variabile.

Il confronto tra le varianze di COMPONENT, ESTIMATOR ed ESTIMATE per ogni singola componente XE "componente" rispettano in tutti i casi le relazioni attese: la varianza XE "varianza" di COMPONENT è maggiore di quella di ESTIMATOR, che è a sua volta simile a quella di ESTIMATE.

Nella sezione "Riduzione nell'errore standard di revisione" della tabella 3.1 si evince che il processo di convergenza verso lo stimatore finale XE "stimatore finale" è molto lento per la serie destagionalizzata XE "serie destagionalizzata", mentre è abbastanza veloce per il ciclo-trend XE "ciclo-trend". In termini di riduzione del RMSE, il modello 4 sembra essere quello che produce minori revisioni.

La figura 3.2 illustra le serie destagionalizzate che si ottengono con i modelli 3 e 4; le diversità più degne di nota derivano dalla presenza dell'outlier additivo XE "outlier additivo" per il modello 3 in corrispondenza dei mesi di agosto 1984 ed agosto 1995 e dalla minore volatilità della prima serie rispetto alla seconda.

Tabella 3.1: Sintesi dell'output di SEATS XE "SEATS" per la serie IPIGENGT

Diagnostica sui residui									
	Modello 1		Modello 2		Modello 3		Modello 4		
Jarque-Bera	1,444		9,782		12,12		0,536		
Ljung-Box XE "Ljung-Box"	34,94		27,93		25,98		24,74		
Pierce XE "Pierce"	11,96		7,79		7,31		2,10		
Ljung-Box XE "Ljung-Box" res.^2	45,99		62,32		64,93		22,85		
Pierce XE "Pierce" res.^2	12,53		37,19		38,30		4,83		
Varianza innovazioni dei modelli delle componenti									
Ciclo-Trend	0,024		0,030		0,029		0,037		
Stagionale	0,049		0,050		0,052		0,044		
Irregolare	0,385		0,355		0,351		0,357		
Destagionalizzata	0,625		0,614		0,604		0,645		
Varianza component-estimator-estimate									
	0,046	0,004	0,057	0,005	0,055	0,005	0,071	0,008	
		0,004		0,005		0,005		0,007	
Trend	2,359	1,824	2,186	1,679	2,160	1,645	2,211	1,745	
Destagionalizzata		1,662		1,510		1,527		1,754	
Irregolare	0,385	0,239	0,355	0,211	0,351	0,208	0,357	0,213	
Stagionale		0,225		0,199		0,195		0,198	
	0,199	0,045	0,285	0,054	0,292	0,059	0,300	0,043	
		0,033		0,038		0,041		0,023	
Riduzione nell'errore standard di revisione per ciclo-trend XE "ciclo-trend" e serie destagionalizzata XE "serie destagionalizzata"									
	86,00		83,84		84,35		81,10		
		42,98		44,40		45,60		40,52	
dopo 1 anno	92,23		91,31		91,79		89,03		
dopo 2 anni		68,32		70,10		71,45		65,48	
dopo 3 anni	85,68		95,33		95,69		93,64		
dopo 4 anni		82,40		83,92		85,02		79,97	
dopo 5 anni	97,60		97,49		97,74		96,31		
		90,22		91,35		92,14		88,38	
	98,67		98,65		98,81		97,86		
		94,57		95,35		95,88		93,26	
Riduzione percentuale media nel RMSE dopo la revisione concurrent XE "concurrent"	27,71		27,29		28,03		24,47		

Figura 3.2 – Serie destagionalizzate dell'indice generale della produzione industriale ottenute con il modello 3 (linea trattata) e con il modello 4 (linea continua)

Figura 3.2 – Serie destagionalizzate dell'indice generale della produzione industriale ottenute con il modello 3 (linea tratteggiata) e con il modello 4 (linea continua)

APPENDICE: Analisi dettagliata del file IPIGNGT.OUT

Analogamente al capitolo 2, anche in questa appendice si presenta un esempio completo dell'output di SEATS XE "SEATS", riferito alla serie della produzione industriale, diviso in parti numerate, con commenti miranti a chiarirne maggiormente i contenuti.

1

```

SIGNAL EXTRACTION IN 'ARIMA XE "ARIMA" ' TIME SERIES (BETA VERSION) (*)

BY

V. GOMEZ & A. MARAVALL,

with the programming assistance of G. CAPORELLO

Thanks are due to G. FIORENTINI and C. PLANAS for their research assistance

(Based on an original program developed by J.P.BURMAN at the Bank of England, version 1982)

(*) Copyright : V. GOMEZ, A. MARAVALL (1994,1996)

FIRST PART:
ARIMA XE "ARIMA" ESTIMATION

SERIES TITLE: GEN81
PREADJUSTED WITH TRAMO XE "TRAMO" : YES
METHOD: MAXIMUM LIKELIHOOD

NO OF OBSERVATIONS =220
    
```

Dopo l'intestazione del programma, viene indicato il titolo della serie, l'eventuale linearizzazione effettuata da TRAMO XE "TRAMO", il metodo di stima XE "stima" impiegato ed il numero di osservazioni totali.

2

```

ORIGINAL UNCORRECTED SERIES (from TRAMO) XE "TRAMO"

YEAR   JAN   FEB   MAR   APR   MAY   JUN   JUL   AUG   SEP   OCT   NOV   DEC
1981   80.700 81.900 90.000 84.500 84.500 88.200 89.000 34.200 89.800 89.300 86.700 72.200
1982   79.500 81.100 90.600 84.600 82.700 82.100 84.300 34.500 86.300 82.200 81.100 72.100
...
...
1998   98.400 105.800 115.300 106.800 107.600 114.200 117.900 52.500 113.500 114.400 111.400 94.100
1999   94.600 103.300 116.800 104.300
    
```

E' la serie originale non corretta per gli effetti deterministici.

E' la serie originale non corretta per gli effetti deterministici.

3

PREADJUSTMENT COMPONENT												
Outliers XE "Outliers" and Other Deterministic Effects												
(from TRAMO) XE "TRAMO"												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-1.282	0.763	1.736	-0.116	-4.019	2.901	3.683	-1.792	2.901	0.946	0.164	-0.771
1982	-4.019	0.763	4.474	-0.116	-1.792	2.901	0.946	0.946	2.901	-1.792	0.674	1.456
...												
...												
1998	-3.509	0.763	1.736	-0.116	-4.019	2.901	3.683	-1.792	2.901	0.946	0.164	-0.771
1999	-6.247	0.763	3.683	0.674								

E' l'effetto delle componenti deterministiche individuate da TRAMO XE "TRAMO".

4

ORIGINAL SERIES (Corrected by TRAMO) XE "TRAMO"												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	81.982	81.137	88.264	84.616	88.519	85.299	85.317	35.992	86.899	88.354	86.536	72.971
1982	83.519	80.337	86.126	84.716	84.492	79.199	83.354	33.554	83.399	83.992	80.426	70.644
...												
...												
1998	101.909	105.037	113.564	106.916	111.619	111.299	114.217	54.292	110.599	113.454	111.236	94.871
1999	100.847	102.537	113.117	103.726								

E' la serie originaria corretta da TRAMO XE "TRAMO" per i precedenti effetti deterministici, ossia la serie linearizzata XE "serie linearizzata" ottenuta sottraendo dalla serie indicata al punto 2 quella espressa al punto 3.

5

INPUT PARAMETERS			

LAM= 1	IMEAN= 1	RSA= 0	MQ=12
P= 0	BP= 0	Q= 1	BQ= 1
D= 1	BD= 1	NOADMISS= 1	RMOD= 0.500
M=36	QMAX=50	BIAS= 1	SMTR= 0
THTR= -0.400			

Riepilogo dei parametri di input, analogamente a TRAMO XE "TRAMO".

6

TRANSFORMATION: Z -> Z

6

TRANSFORMATION: Z -> Z

Viene indicato se è stata effettuata la trasformazione logaritmica (non è questo il caso, essendo $LAM=1$). Nel caso in cui essa fosse stata usata, sarebbe apparsa la dizione "transformation :z log z con la serie trasformata di seguito (transformed series)".

7

NONSEASONAL DIFFERENCING D= 1												
SEASONAL DIFFERENCING BD= 1												
DIFFERENCED SERIES												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1982		-2.338	-1.338	2.238	-4.127	-2.073	4.138	-0.475	-1.062	-0.862	-1.748	3.783
1983	-4.000	2.127	1.154	-3.342	2.251	3.038	-1.162	3.152	-2.290	1.200	2.700	0.065
...												
...												
1998	-1.852	-2.038	0.754	-2.374	-0.043	-0.075	-0.962	-3.027	3.027	-3.262	1.825	-5.090
1999	4.065	-1.438	2.053	-2.743								

Vengono indicati gli ordini di differenziazione (sia stagionali sia non stagionali) e, di seguito, la serie così trasformata.

8

SERIES HAS BEEN MEAN CORRECTED												
DIFFERENCED AND CENTERED SERIES												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1982	-2.315	-1.315	2.260	-4.105	-2.050	4.160	-0.452	-1.040	-0.840	-1.725	3.805	
1983	-3.977	2.150	1.177	-3.320	2.274	3.060	-1.140	3.175	-2.267	1.223	2.723	0.088
...												
...												
1998	-1.830	-2.015	0.777	-2.351	-0.020	-0.052	-0.940	-3.005	3.050	-3.240	1.848	-5.067
1999	4.088	-1.415	2.076	-2.720								

La serie è stata corretta per una costante additiva, per cui risulta centrata; di seguito è riportata, al solito, la serie così trasformata.

9

.....

solito, la serie così trasformata.

9

```
MEAN OF DIFFERENCED SERIES -0.2284D-01

VARIANCE OF Z SERIES = 0.2875D+03

VARIANCE OF DIFFERENCED SERIES = 0.6136D+01
```

Viene indicata la media della serie differenziata XE "serie differenziata", la varianza XE "varianza" della serie Z e quella della serie differenziata

10

```
AUTOCORRELATIONS OF STATIONARY SERIES
-----

-0.4414 0.0504 0.0398 -0.0775 0.1363 0.0308 -0.1873 0.1308 0.0236 -0.0189 0.1325 -0.3240
SE 0.0695 0.0819 0.0821 0.0822 0.0825 0.0836 0.0837 0.0857 0.0866 0.0867 0.0867 0.0876

0.1740 -0.1084 0.0404 -0.0297 -0.0595 0.0119 -0.0205 0.0204 -0.0565 0.0467 -0.0530 -0.0284
SE 0.0933 0.0948 0.0954 0.0955 0.0955 0.0957 0.0957 0.0957 0.0958 0.0959 0.0960 0.0962

0.0093 0.1042 -0.1043 0.0404 -0.0604 0.0102 0.1155 -0.0356 0.0012 -0.0523 0.1792 -0.0891
SE 0.0962 0.0962 0.0968 0.0973 0.0974 0.0976 0.0976 0.0982 0.0983 0.0983 0.0984 0.1000

PARTIAL AUTOCORRELATIONS
-----

-0.4414 -0.1794 -0.0169 -0.0702 0.0972 0.1712 -0.1088 -0.0222 0.0860 0.0487 0.1580 -0.2254
SE 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695

-0.0877 -0.1983 -0.0687 -0.0755 -0.0588 -0.0098 -0.1396 0.0010 -0.0396 0.0872 0.0709 -0.1372
SE 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695

-0.0883 0.0091 -0.0055 -0.0503 -0.1145 -0.1122 -0.0278 0.1029 0.0850 -0.0163 0.1867 -0.0178
SE 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695 0.0695
```

Autocorrelazioni totali e parziali (con relativi errori standard) della serie centrata e differenziata.

11

```
MODEL FITTED

NONSEASONAL P= 0 D= 1 Q= 1
SEASONAL BP= 0 BD= 1 BQ= 1
PERIODICITY MQ= 12

MEAN = -0.228397E-01

SE = *****

ARIMA XE "ARIMA" PARAMETERS

TH = -0.5112
SE = *****
BTH = -0.5804
CF = *****
```

```

TH = -0.5112
SE = *****
BTH = -0.5804
SE = *****

```

Il modello ARIMA XE "ARIMA" scelto è del tipo $(P,D,Q)(BP,BD,BQ)$. Vengono riportate anche le stime dei coefficienti effettuate in TRAMO XE "TRAMO", che saranno impiegate da SEATS XE "SEATS" per la scomposizione XE "scomposizione". In definitiva, il modello che SEATS usa per la scomposizione è $(1-B)(1-B^{12})z_t^{-m} = (1-0.5112B)(1-0.5804B^{12})a_t$.

12

RESIDUALS												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-0.259	0.255	0.197	0.128	1.034	0.487	-1.002	-1.299	0.428	0.273	0.593	-2.124
1982	2.261	-0.934	-1.753	1.380	-2.837	-3.524	1.633	-0.074	-0.443	-1.035	-1.991	1.379
...												
1998	-1.439	-0.755	-0.126	0.386	-1.091	0.048	0.555	-3.915	1.778	-1.043	0.277	-5.085
1999	0.792	-1.021	1.704	-1.587								

Serie dei residui XE "Serie dei residui" ricostruita da SEATS XE "SEATS", tramite il modello prescelto. I residui calcolati sono diversi da quelli ottenuti da TRAMO XE "TRAMO".

13

TEST-STATISTICS ON RESIDUALS	

MEAN=	0.1140D+00
ST.DEV.=	0.1277D+00
OF MEAN	
T-VALUE=	0.8925
NORMALITY TEST=	0.5419 (CHI-SQUARED(2))
SKEWNESS=	-0.1199 (SE = 0.1651)
KURTOSIS=	3.0396 (SE = 0.3303)
SUM OF SQUARES=	0.7927D+03
DURBIN-WATSON=	2.0750
STANDARD DEVI.=	0.1971D+01
OF RESID.	
VARIANCE=	0.3886D+01
OF RESID.	

Test sui residui di SEATS XE "SEATS", analoghi a quelli calcolati in TRAMO XE "TRAMO".

14

AUTOCORRELATIONS OF RESIDUAL	

AUTOCORRELATIONS OF RESIDUAL

```

-----
-0.0432 0.0206 0.0906 0.0303 0.0560 0.0582 -0.1018 0.1080 0.1039 0.0051 -0.0148 0.0294
SE 0.0674 0.0675 0.0676 0.0681 0.0682 0.0684 0.0686 0.0693 0.0701 0.0708 0.0708 0.0708

-0.0231 -0.1430 -0.0591 -0.0776 -0.1224 -0.0440 -0.0112 0.0222 -0.0232 -0.0097 -0.0501 -0.0891
SE 0.0708 0.0709 0.0722 0.0724 0.0728 0.0737 0.0738 0.0738 0.0738 0.0739 0.0739 0.0740

0.0033 -0.0127 -0.1007 0.0578 -0.0793 0.0210 0.1936 0.0431 -0.0035 0.0136 0.1057 -0.0766
SE 0.0745 0.0745 0.0745 0.0752 0.0754 0.0757 0.0758 0.0780 0.0781 0.0781 0.0781 0.0787

```

THE LJUNG-BOX Q VALUE IS 24.67 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED (21)

THE PIERCE QS VALUE IS 2.07 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED (2)

Autocorrelazioni dei residui e relativi test di Ljung-Box XE "Ljung-Box" (valgono considerazioni analoghe a quelle fatte per TRAMO XE "TRAMO"). Inoltre, viene calcolato il test di Pierce XE "Pierce" , di cui viene indicata anche la distribuzione della relativa statistica (chi quadro con 2 gradi di libertà). Entrambi i test accettano l'ipotesi di incorrelazione.

PARTIAL AUTOCORRELATIONS

```

-----
-0.0432 0.0188 0.0925 0.0382 0.0558 0.0544 -0.1066 0.0869 0.1060 0.0225 -0.0350 0.0101
SE 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674

-0.0311 -0.1835 -0.0723 -0.0652 -0.1299 -0.0619 0.0309 0.0741 -0.0196 0.0445 0.0084 -0.0892
SE 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674

0.0175 0.0505 -0.0792 0.0051 -0.0875 -0.0119 0.1441 0.0905 0.0195 -0.0410 0.1269 -0.0936
SE 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674 0.0674

```

Autocorrelazioni parziali dei residui e relativi errori standard dal ritardo 1 al ritardo 36.

APPROXIMATE TEST OF RUNS ON AUTOCORRELATION FUNCTION

```

-----
NUM.DATA= 36
NUM.(+)= 19
NUM.(-)= 17
T-VALUE= -0.660

```

APPROXIMATE TEST OF RUNS ON RESIDUALS

```

-----
NUM.DATA= 220
NUM.(+)= 111
NUM.(-)= 109
T-VALUE= 0.812

```

Test dei run approssimato per la funzione di autocorrelazione dei residui e per i residui (analoghi a quelli visti per TRAMO XE "TRAMO").

AUTOCORRELATIONS OF SQUARED RESIDUAL

AUTOCORRELATIONS OF SQUARED RESIDUAL

-0.0253 -0.0180 0.0605 -0.0190 0.0407 -0.0793 0.0287 0.0315 -0.0244 0.0198 -0.1423 0.0207
 SE 0.0674 0.0675 0.0675 0.0677 0.0678 0.0679 0.0683 0.0683 0.0684 0.0684 0.0685 0.0698

-0.0076 -0.1165 -0.0603 -0.0229 0.0174 -0.0417 -0.0393 0.0829 -0.1214 -0.0259 -0.0293 0.1406
 SE 0.0698 0.0698 0.0707 0.0709 0.0710 0.0710 0.0711 0.0712 0.0716 0.0726 0.0726 0.0727

-0.0593 -0.0311 0.0297 0.0628 -0.0221 -0.0791 0.0810 0.0672 -0.0458 -0.0231 -0.0333 0.1136
 SE 0.0739 0.0741 0.0742 0.0742 0.0745 0.0745 0.0749 0.0753 0.0755 0.0757 0.0757 0.0758

THE LJUNG-BOX Q VALUE IS 22.80 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED (21)

THE PIERCE QS VALUE IS 4.77 AND IF RESIDUALS ARE RANDOM IT SHOULD BE DISTRIBUTED AS CHI-SQUARED (2)

Residui XE "Residui" al quadrato con relativi errori standard; in fondo sono riportati i test di Ljung-Box XE "Ljung-Box" e di Pierce XE "Pierce", che accettano entrambi l'ipotesi di incorrelazione.

18

BACKWARD RESIDUALS

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-0.874	-0.404	-0.564	-1.011	1.286	2.376	0.430	-1.809	0.144	0.463	2.435	-3.418
1982	3.597	1.485	-1.137	5.554	2.655	-1.956	2.929	-0.362	2.132	1.656	0.059	-1.155
...												
...												
1998	-0.584	0.686	-1.252	1.252	-0.066	0.559	1.695	-1.450	1.455	0.630	2.152	-1.725
1999	0.477	-0.686	1.186	-0.471								

Residui XE "Residui" calcolati "all'indietro".

19

SECOND PART:
 DERIVATION OF THE MODELS FOR THE COMPONENTS

SERIES TITLE: GEN81

MODEL PARAMETERS
 (0,1,1)(0,1,1)

PARAMETER VALUES PASSED FROM ARIMA XE "ARIMA" ESTIMATION (TRUE SIGNS)

THETA PARAMETERS
 1.00 -0.51

BTHETA PARAMETERS
 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 -0.58

PHI PARAMETERS
 1.00

BPHI PARAMETERS
 1.00

1.00

BPHI PARAMETERS

1.00

Comincia la seconda parte del programma, relativa alla derivazione dei modelli per le singole componenti. Il modello impiegato è quello identificato e stimato da TRAMO XE "TRAMO", espresso in termini di polinomi regolari e stagionali. In particolare, i parametri theta, Btheta si riferiscono ai polinomi MA XE "MA" regolare e stagionale mentre i parametri phi e Bphi si riferiscono ai polinomi AR regolare e stagionale. In questo caso il polinomio MA regolare è pari a (1-0.51B); il polinomio MA è pari a (1+0B+0B^2+0B^3+...-0.58B^12)=(1-0.58B^12); il polinomio AR regolare è pari ad 1; Il polinomio AR stagionale è pari ad 1.

20

NUMERATOR OF THE MODEL

1.0000 -0.5112 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-0.5804 0.2967

Dai risultati della parte 19 ne deriva che il numeratore del modello è pari a: (1-0.5112B)(1+0B+0B^2+0B^3+...-0.5804B^12)=(1-0.5112B+0B^2+0B^3+...-0.5804B^12+0.2967B^13).

21

STATIONARY AUTOREGRESSIVE TREND

1.0000

NON-STATIONARY AUTOREGRESSIVE TREND

1.0000 -2.0000 1.0000

AUTOREGRESSIVE TREND

1.0000 -2.0000 1.0000

STATIONARY AUTOREGRESSIVE TRANSITORY COMP.

1.0000

NON-STATIONARY AUTOREGRESSIVE TRANSITORY COMP.

1.0000

AUTOREGRESSIVE TRANSITORY COMP.

1.0000

STATIONARY AUTOREGRESSIVE SEASONAL COMPONENT

1.0000

NON-STATIONARY AUTOREGRESSIVE SEASONAL COMPONENT

1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000

AUTOREGRESSIVE SEASONAL COMPONENT

1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000

STATIONARY AUTOREGRESSIVE SEASONALLY ADJUSTED COMPONENT

1.0000

NON-STATIONARY AUTOREGRESSIVE SEASONALLY ADJUSTED COMPONENT

1.0000 -2.0000 1.0000

AUTOREGRESSIVE SEASONALLY ADJUSTED COMPONENT

1.0000 -2.0000 1.0000

AUTOREGRESSIVE SEASONALLY ADJUSTED COMPONENT		

1.0000	-2.0000	1.0000

Per ogni componente XE "componente" (TREND, TRANSITORY, SEASONAL, SEASONAL ADJUSTED) vengono elencati i polinomi AR (parte stazionaria, parte non stazionaria e loro prodotto totale $\square_q(B)$, $\square_s(B)$ etc. illustrati nel par. 3.2.2). Nel TREND, per esempio, il polinomio stazionario è pari ad 1, quello non stazionario è pari a $(1-B)$ $(1-B)=(1-2B+B^2)$: in totale il polinomio AR del trend, prodotto delle due parti, è $(1-2B+B^2)$.

22

TOTAL DENOMINATOR											

1.0000	-1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.0000	1.0000										

Il prodotto del polinomio AR del TREND e della SEASONAL costituisce il denominatore del modello $\square(B)$: $(1-2B+B^2)*1*(1+B+B^2+...+B^{11})=(1-B+0B^2 +...+0B^{11}-B^{12}+B^{13})=(1-B-B^{12}+B^{13})$.

23

MA XE "MA" ROOTS OF TREND					

REAL PART	IMAGINARY PART	MODULUS	ARGUMENT	PERIOD	
				(DEG.)	
0.956	0.000	0.956	0.000	0.000	-
-1.000	0.000	1.000	180.000		2.0
TOTAL SQUARED ERROR= 0.2802927D-32					
MA XE "MA" ROOTS OF SEAS.					

REAL PART	IMAGINARY PART	MODULUS	ARGUMENT	PERIOD	
				(DEG.)	
-0.966	0.259	1.000	165.007		2.182
0.648	0.656	0.921	45.353		7.938
-0.696	0.695	0.984	135.023		2.666
0.244	0.919	0.951	75.135		4.791
-0.252	0.936	0.970	105.055		3.427
0.790	0.000	0.790	0.000		-
TOTAL SQUARED ERROR= 0.4473941D-14					
MA XE "MA" ROOTS OF SEASONALLY ADJUSTED SERIES					

REAL PART	IMAGINARY PART	MODULUS	ARGUMENT	PERIOD	
				(DEG.)	
0.520	0.000	0.520	0.000	0.000	-
0.956	0.000	0.956	0.000	0.000	-
TOTAL SQUARED ERROR= 0.1068141D-30					

Per ogni componente XE "componente" vengono illustrate le radici dei polinomi MA XE "MA". Poichè, in generale, queste radici sono numeri complessi, di esse vengono riportate la rappresentazione vettoriale (parte reale e immaginaria), radiale (modulo e argomento in gradi) e il periodo. Quest'ultimo si ottiene dividendo 2π per l'argomento espresso in radianti. Infine è riportato l'Errore Quadratico Totale.

24

MODEL S FOR THE COMPONENTS

MODELS FOR THE COMPONENTS										

TREND NUMERATOR										
1.0000	0.0442	-0.9558								
TREND DENOMINATOR										
1.0000	-2.0000	1.0000								
INNOV. VAR. (*) 0.03737										
SEAS. NUMERATOR										
1.0000	1.2548	1.2170	1.0738	0.8457	0.5966	0.3497	0.1228	-0.0572	-0.2167	-0.3174
-0.5520										
SEAS. DENOMINATOR										
1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0000										
INNOV. VAR. (*) 0.04369										
IRREGULAR										
VAR. 0.35651										
SEASONALLY ADJUSTED NUMERATOR										
1.0000	-1.4757	0.4970								
SEASONALLY ADJUSTED DENOMINATOR										
1.0000	-2.0000	1.0000								
INNOV. VAR. (*) 0.64550										
(*) IN UNITS OF VAR(A)										

E' riportata la rappresentazione ARIMA XE "ARIMA" delle singole componenti (cfr. 3.2.2). Ad esempio, il modello completo per il TREND è dato da:

$$T_t = [(1 + 0.0442B - 0.9558B^2) / (1 - 2B + B^2)] a_{T_t}$$

con $var(a_{T_t}) = 0.03737 var(a_t)$, ovvero un IMA(2,2) potendosi scrivere come:

$$(1-B)(1-B) T_t = (1 + 0.0442B - 0.9558B^2) a_{T_t}$$

L'IRREGULAR, I_T risulta essere un White Noise con $var(I_t) = 0.35651 var(a_t)$

MOVING AVERAGE REPRESENTATION OF ESTIMATORS (NONSTATIONARY)

The last column (the sum of the Psi-Weights) should be zero for negative lags, 1 for lag=0, and equal to the Box-Jenkins Psi-Weights for positive lags.

PSIEP(LAG), for example, represents the effect of the overall innovation $a(t-lag)$ on the estimator of the trend for period t . Similarly for the other components.

PSIEP(LAG), for example, represents the effect of the overall innovation $a(t\text{-lag})$ on the estimator of the trend for period t . Similarly for the other components.

LAG	PSIEP	PSIES	PSIEC	PSIEA	PSIUE	PSIEP + PSIES + PSIUE
-24	-0.0211	0.1079	0.0000	-0.1079	-0.0868	0.0000
-23	-0.0234	0.0233	0.0000	-0.0233	0.0001	0.0000
...						
...						
-1	0.2703	-0.0960	0.0000	0.0960	-0.1743	0.0000
0	0.3770	0.2665	0.0000	0.7335	0.3565	1.0000
1	0.4298	0.0590	0.0000	0.4298	0.0000	0.4888
...						
...						
24	0.8229	0.2906	0.0000	0.8229	0.0000	1.1136

Sono riportati gli PSI-Weights. Gli PSI-Weights del TREND, SEASONAL, TRANSITORY, SEASONAL ADJUSTED e IRREGULAR sono rispettivamente denominati: *psiep*, *psies*, *psiec*, *psiea*, *psiue* (cfr. par.3.3.1.2).

Nell'esempio la rappresentazione dello stimatore del trend è:

$$T_t = -0.0211a_{t+24} + \dots + 0.0506 a_{t+5} + 0.0768 a_{t+4} + 0.1139 a_{t+3} + 0.1719 a_{t+2} + 0.2703 a_{t+1} + 0.3770 a_t + 0.4298 a_{t-1} + 0.4469 a_{t-2} + 0.4640 a_{t-3} + 0.4810 a_{t-4} + 0.4981 a_{t-5} + \dots$$

26

WIENER-KOLMOGOROV FILTERS (ONE SIDE)

TREND COMPONENT

0.2058	0.1604	0.0916	0.0560	0.0373	0.0272	0.0213	0.0172	0.0133	0.0082	-0.0002	-0.0154
-0.0256	-0.0168	-0.0030	0.0038	0.0070	0.0083	0.0085	0.0080	0.0067	0.0042	-0.0004	-0.0091
-0.0149	-0.0098	-0.0018	0.0022	0.0041	0.0048	0.0050	0.0047	0.0039	0.0025	-0.0002	-0.0053
-0.0087	-0.0057	-0.0010	0.0013	0.0024	0.0028	0.0029	0.0027	0.0023	0.0014	-0.0001	-0.0031
-0.0050	-0.0033	-0.0006	0.0007	0.0014	0.0016	0.0017	0.0016	0.0013	0.0008	-0.0001	-0.0018

SA SERIES COMPONENT

0.8030	0.0145	0.0171	0.0180	0.0181	0.0178	0.0173	0.0167	0.0161	0.0155	0.0151	0.0150
-0.1510	0.0138	0.0127	0.0119	0.0112	0.0107	0.0102	0.0098	0.0094	0.0090	0.0088	0.0087
-0.0876	0.0080	0.0074	0.0069	0.0065	0.0062	0.0059	0.0057	0.0054	0.0052	0.0051	0.0051
-0.0509	0.0047	0.0043	0.0040	0.0038	0.0036	0.0034	0.0033	0.0032	0.0030	0.0030	0.0029
-0.0295	0.0027	0.0025	0.0023	0.0022	0.0021	0.0020	0.0019	0.0018	0.0018	0.0017	0.0017

SEASONAL COMPONENT

0.1970	-0.0145	-0.0171	-0.0180	-0.0181	-0.0178	-0.0173	-0.0167	-0.0161	-0.0155	-0.0151	-0.0150
0.1510	-0.0138	-0.0127	-0.0119	-0.0112	-0.0107	-0.0102	-0.0098	-0.0094	-0.0090	-0.0088	-0.0087
0.0876	-0.0080	-0.0074	-0.0069	-0.0065	-0.0062	-0.0059	-0.0057	-0.0054	-0.0052	-0.0051	-0.0051
0.0509	-0.0047	-0.0043	-0.0040	-0.0038	-0.0036	-0.0034	-0.0033	-0.0032	-0.0030	-0.0030	-0.0029
0.0295	-0.0027	-0.0025	-0.0023	-0.0022	-0.0021	-0.0020	-0.0019	-0.0018	-0.0018	-0.0017	-0.0017

0.0876	-0.0080	-0.0074	-0.0069	-0.0065	-0.0062	-0.0059	-0.0057	-0.0054	-0.0052	-0.0051	-0.0051
0.0509	-0.0047	-0.0043	-0.0040	-0.0038	-0.0036	-0.0034	-0.0033	-0.0032	-0.0030	-0.0030	-0.0029
0.0295	-0.0027	-0.0025	-0.0023	-0.0022	-0.0021	-0.0020	-0.0019	-0.0018	-0.0018	-0.0017	-0.0017
IRREGULAR COMPONENT											
0.5971	-0.1459	-0.0745	-0.0380	-0.0192	-0.0094	-0.0040	-0.0005	0.0028	0.0073	0.0153	0.0304
-0.1254	0.0306	0.0157	0.0081	0.0042	0.0024	0.0017	0.0018	0.0027	0.0048	0.0092	0.0178
-0.0727	0.0178	0.0091	0.0047	0.0025	0.0014	0.0010	0.0010	0.0015	0.0028	0.0053	0.0103
-0.0422	0.0103	0.0053	0.0027	0.0014	0.0008	0.0006	0.0006	0.0009	0.0016	0.0031	0.0060
-0.0245	0.0060	0.0031	0.0016	0.0008	0.0005	0.0003	0.0003	0.0005	0.0009	0.0018	0.0035

Sono riportati i pesi dei coefficienti del filtro WK applicato a z_t . Il filtro è simmetrico e vengono quindi elencati i coefficienti solo per i ritardi positivi. Nell'esempio, la rappresentazione dello stimatore del trend, relativo alle 59 osservazioni, è:

$$T_t = -0.0018z_{t-59} + \dots + 0.0272z_{t-5} + 0.0373z_{t-4} + 0.0560z_{t-3} + 0.0916z_{t-2} +$$

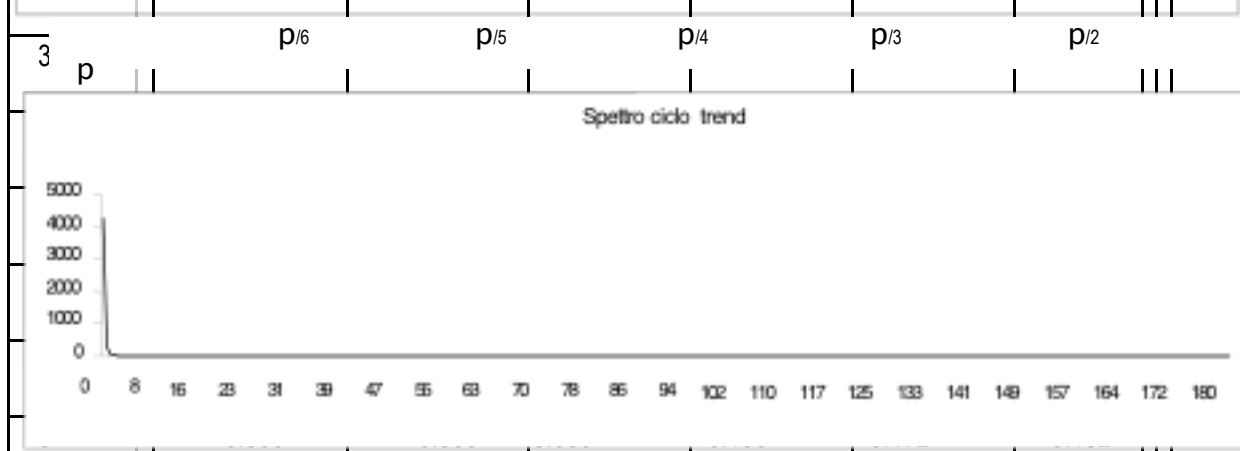
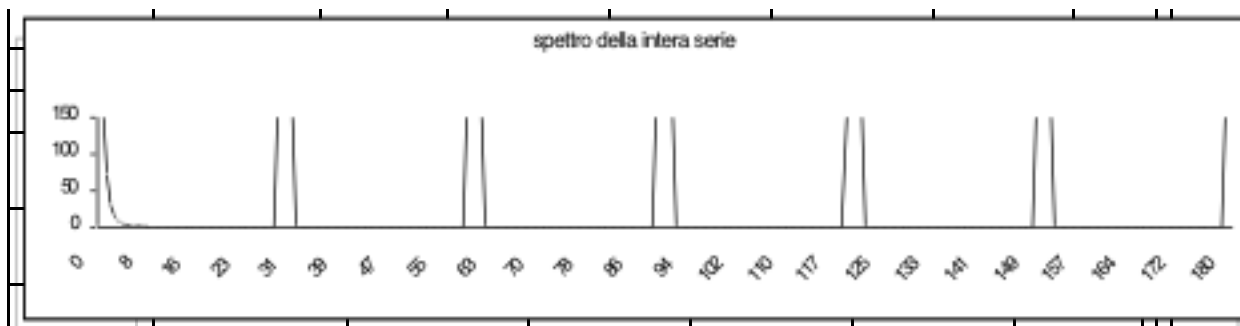
$$0.1604z_{t-1} + 0.2050z_t + 0.1604z_{t+1} + 0.0916z_{t+2} + 0.0560z_{t+3} + 0.0373z_{t+4} + 0.0272z_{t+5} + \dots - 0.0018z_{t+59}$$

27

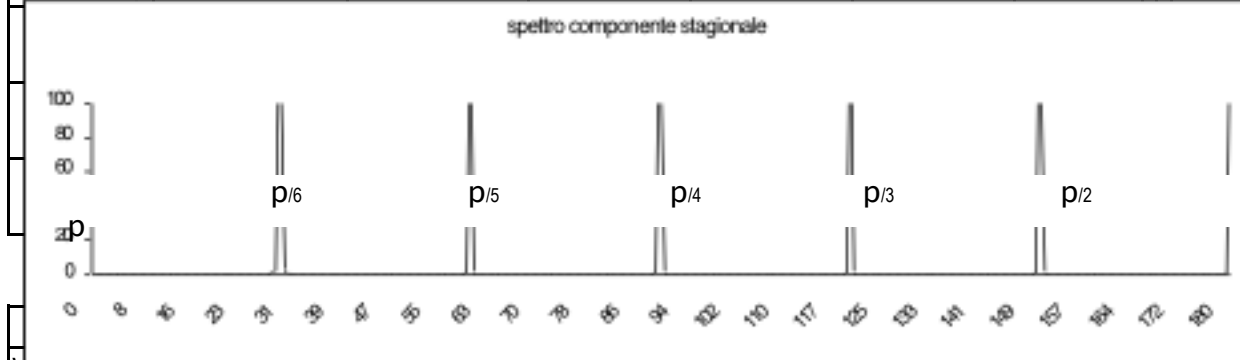
AUTOCORRELATION FUNCTION OF COMPONENTS (STATIONARY TRANSFORMATION)						
	TREND			ADJUSTED		
LAG	COMPONEN T	ESTIMATOR	ESTIMATE	COMPONEN T	ESTIMATOR	ESTIMATE
1	0.001	0.460	0.441	-0.645	-0.646	-0.666
2	-0.499	-0.257	-0.253	0.145	0.146	0.153
3	0.000	-0.327	-0.280	0.000	0.000	0.066
4	0.000	-0.165	-0.107	0.000	0.000	-0.105
5	0.000	-0.081	-0.023	0.000	0.000	0.051
6	0.000	-0.034	-0.097	0.000	0.000	0.087
7	0.000	-0.004	-0.153	0.000	0.000	-0.176
8	0.000	0.024	0.051	0.000	0.000	0.100
9	0.000	0.064	0.249	0.000	0.000	0.016
10	0.000	0.052	0.174	0.000	-0.031	-0.044
11	0.000	-0.097	-0.030	0.000	0.136	0.078
12	0.000	-0.210	-0.128	0.000	-0.210	-0.134
VAR.(*)	0.072	0.008	0.007	2.211	1.745	1.755

(*) IN UNITS OF VAR(A)

spettro della intera serie



9	0.000	0.012	0.097	-0.187	-0.214	-0.010
---	-------	-------	-------	--------	--------	--------



VAR.()	0.357	0.213	0.196	0.300	0.045	0.023
---------	-------	-------	-------	-------	-------	-------

(*) IN UNITS OF VAR(A)
 For all components it should happen that :
 - Var(Component) > Var(Estimator)
 - Var(Estimator) close to Var(Estimate)

Sono riportate le autocorrelazioni XE "autocorrelazioni" per le trasformazioni stazionarie delle componenti teoriche, degli stimatori e delle stime (applicazione del filtro troncato alla serie z). Tale confronto permette di comprendere le differenze indotte dal processo di stima XE "stima", sia nel caso teorico in cui si disponesse di una serie infinita, che nel caso effettivo. In particolare la differenza tra le varianze di componente XE "componente", stimatore e stima deve soddisfare quanto è atteso ex-ante dalla teoria della stima. La varianza XE "varianza" dello stimatore deve essere inferiore a quella della componente. Come abbiamo visto, infatti, il primo sottostima la variabilità della componente. La varianza della stima deve essere vicina a quella dello stimatore. Questa vicinanza è una garanzia che il "taglio" del filtro non ha comportato ulteriore perdita di informazione.

CROSSCORRELATION BETWEEN STATIONARY TRANSFORMATION OF ESTIMATORS		
	ESTIMATOR	ESTIMATE
TREND-SEASONAL	-0.119	-0.303
SEASONAL-IRREGULAR	0.583E-01	0.167
TREND-IRREGULAR	0.200	-0.154

SEASONAL-IRREGULAR

0.583E-01

0.167

TREND-IRREGULAR

0.200

-0.154

Sono riportate le correlazioni incrociate, per ciascuna coppia delle trasformazioni stazionarie degli stimatori e delle stime. Anche in questo caso ci si deve attendere che le due misure siano sufficientemente vicine.

29

PSEUDO-INNOVATIONS IN THE COMPONENTS

PSEUDO INNOVATIONS IN TREND

YEAR JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

1981 0.06 0.02 0.02 0.00 0.15 0.24 0.13 0.02 0.07 0.14 0.02 0.02

1982 -0.01 -0.04 -0.01 0.04 -0.01 -0.08 -0.17 -0.23 0.19 -0.27 -0.30 -0.24

...

...

1998 0.28 0.16 0.23 0.03 0.02 0.11 0.00 -0.04 0.03 0.16 0.10 -0.05

1999 -0.10 -0.07 -0.07 -0.03

PSEUDO INNOVATIONS IN SEASONAL

YEAR JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

1981 0.04 -0.08 0.05 -0.04 0.14 -0.12 0.01 -0.09 0.11 -0.11 -0.01 0.02

1982 0.02 -0.12 0.12 0.03 0.27 -0.19 -0.10 -0.09 0.27 -0.24 0.02 0.09

...

...

1998 0.16 -0.11 0.05 0.10 -0.18 0.10 0.01 -0.02 -0.13 0.00 0.13 0.09

1999 -0.02 0.00 -0.01 -0.04

PSEUDO INNOVATIONS IN SEASONALLY ADJUSTED SERIES

YEAR JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

1981 -0.51 1.29 -0.77 0.51 -1.88 2.34 0.61 1.48 -1.71 1.72 0.43 -0.26

1982 0.79 -0.65 -0.01 -0.48 1.29 -0.70 0.86 -1.61 -0.70 -0.26 -1.66 -1.43

...

...

1998 3.96 -0.47 1.14 2.54 -2.15 1.57 0.29 0.08 -1.14 0.29 1.49 0.76

1999 -0.70 -0.40 -0.28 -0.56

E' la stima XE "stima" delle pseudo innovazioni delle componenti a_{Tt} e a_{St} (cfr. par. 3.2.2).

30

THIRD PART:				
ERROR ANALYSIS				
FINAL ESTIMATION ERROR		REVISION IN CONCURRENT ESTIMATOR		
ACF (LAG)	TREND	ADJUSTED	TREND	ADJUSTED
1	0.773	0.157	0.656	0.291
2	0.410	0.024	0.410	0.148

	TREND	ADJUSTED	TREND	ADJUSTED
1	0.773	0.157	0.656	0.291
2	0.410	0.024	0.419	0.148
3	0.187	-0.078	0.243	0.026
4	0.051	-0.151	0.109	-0.073
5	-0.030	-0.196	0.007	-0.148
6	-0.070	-0.214	-0.068	-0.196
7	-0.079	-0.205	-0.118	-0.217
8	-0.062	-0.172	-0.141	-0.209
9	-0.028	-0.115	-0.135	-0.171
10	0.018	-0.036	-0.096	-0.102
11	0.061	0.064	-0.014	-0.002
12	0.084	0.734	0.081	0.594
VAR.(*)	0.100	0.097	0.134	0.098
TOTAL ESTIMATION ERROR (CONCURRENT ESTIMATOR)				

ACF (LAG)	TREND	ADJUSTED		
1	0.706	0.224		
2	0.415	0.086		
3	0.219	-0.026		
4	0.084	-0.112		
5	-0.009	-0.172		
6	-0.069	-0.205		
7	-0.101	-0.211		
8	-0.107	-0.190		
9	-0.089	-0.143		
10	-0.047	-0.069		
11	0.018	0.031		
12	0.082	0.664		
VAR.(*)	0.234	0.195		
(*) IN UNITS OF VAR(A)				

La terza parte dell'output, relativa all'analisi degli errori, inizia presentando le autocorrelazioni XE "autocorrelazioni" e la varianza XE "varianza" delle serie degli errori finali, di revisione e totali per il TREND e per la serie SEASONAL ADJUSTED. La varianza è espressa in termini di unità di varianza dell'innovazione del processo generatore.

Situazioni come questa, in cui la varianza XE "varianza" dell'errore di stima XE "errore di stima" totale per lo stimatore concurrent XE "concurrent" del trend è maggiore della varianza dell'errore di stima XE "stima" totale per lo stimatore concurrent della componente XE "componente" stagionale XE "componente stagionale", riflettono, generalmente la presenza di una componente stagionale stabile.

VARIANCE OF THE REVISION ERROR (*)				

ADDITIONAL	TREND	ADJUSTED		

VARIANCE OF THE REVISION ERROR (*)		
ADDITIONAL PERIODS	TREND	ADJUSTED
0	0.1339	0.9761E-01
12	0.4795E-02	0.3454E-01
24	0.1615E-02	0.1164E-01
36	0.5438E-03	0.3919E-02
48	0.1832E-03	0.1320E-02
60	0.6170E-04	0.4446E-03

PERCENTAGE REDUCTION IN THE STANDARD ERROR OF THE REVISION AFTER ADDITIONAL YEARS (COMPARISON WITH CONCURRENT ESTIMATORS)		
AFTER 1 YEAR	81.07	40.51
AFTER 2 YEAR	89.02	65.47
AFTER 3 YEAR	93.63	79.96
AFTER 4 YEAR	96.30	88.37
AFTER 5 YEAR	97.85	93.25

Vengono presentate diagnostiche relative alla velocità di convergenza della stima XE "stima" preliminare XE "preliminare" alla stima finale per TREND e SEASONAL ADJUSTED, espressa in termini di riduzione dell'errore standard di revisione all'aumentare delle osservazioni. Nell'esempio l'errore standard di revisione del trend si è ridotto dell'81,07% dopo un anno rispetto all'errore dello stimatore concurrent XE "concurrent", mentre per la SEASONAL ADJUSTED la riduzione è del 40,51%.

32

VARIANCE OF THE REVISION ERROR FOR THE SEASONAL COMPONENT (ONE YEAR AHEAD ADJUSTMENT)	
PERIODS AHEAD VARIANCE (*)	
0	0.9761E-01
1	0.1687
2	0.1721
3	0.1739
4	0.1745
5	0.1746
6	0.1747
7	0.1754
8	0.1773
9	0.1809
10	0.1870
11	0.1960
12	0.2085

AVERAGE PERCENTAGE REDUCTION IN RMSE FROM CONCURRENT ADJUSTMENT 24.46

(*) IN UNITS OF VAR(A)

Questa tavola esprime la varianza XE "varianza" dell'errore di revisione XE "errore di revisione" (totale), per la componente XE "componente" stagionale XE "componente stagionale" S , nello stimatore concurrent XE "concurrent" ($r_{ST+i} = S_{ST} - S_{ST+i}$) e le varianze degli errori di revisione ottenuti proiettando in avanti, per 12 periodi, i fattori stagionali dell'anno precedente, cioè $S_{ST+i} - S_{ST-12+i}$, per $i=1, \dots, 12$, dove gli stimatori $S_{ST-12+i}$ si ottengono applicando i fattori stagionali stimati al tempo T ai dati osservati per i periodi successivi $T+i$ per $i=1, \dots, 12$. In fondo alla tavola viene riportata la riduzione percentuale dell'RMSE che si sarebbe ottenuta utilizzando la stima XE "stima" concurrent.

$S^*_{ST-12+i|T}$ si ottengono applicando i fattori stagionali stimati al tempo T ai dati osservati per i periodi successivi $T+i$ per $i=1, \dots, 12$. In fondo alla tavola viene riportata la riduzione percentuale dell'RMSE che si sarebbe ottenuta utilizzando la stima XE "stima" concurrent.

33

DECOMPOSITION OF THE SERIES: RECENT ESTIMATES						
PERIOD	SERIES	TREND			ADJUSTED	
		ESTIMATE	STANDARD ERROR	TOTAL OF REVISION	ESTIMATE	STANDARD ERROR
-24	104.4	101.4	0.6293	0.7921E-01	102.8	0.6503
	0.2126					
-23	109.2	102.0	0.6307	0.8949E-01	101.8	
	0.6842	0.3008				
...						
-1	113.1	102.8	0.7913	0.4861	103.7	0.8492
	0.5861					
0	103.7	102.7	0.9539	0.7212	102.2	0.8700
	0.6159					
STANDARD ERROR OF FINAL ESTIMATOR			0.6243		0.6145	
DECOMPOSITION OF THE SERIES: SEASONAL						
PERIOD	SEASONAL		STANDARD ERROR			
	ESTIMATE	TOTAL OF REVISION	TOTAL OF REVISION	OF REVISION		
-24	1.609	0.6503	0.6503	0.2126		
-23	7.409	0.6842	0.6842	0.3008		
...						
-1	9.430	0.8492	0.8492	0.5861		
0	1.567	0.8700	0.8700	0.6159		
STANDARD ERROR OF FINAL ESTIMATOR			0.6145			

Per gli ultimi dati della serie si riportano: la serie originale (z); la stima XE "stima" del TREND, della SEASONAL ADJUSTED e della componente XE "componente stagionale" con i relativi errori standard sia totali che di revisione. Nell'esempio si riscontra la maggiore velocità di convergenza da parte del TREND (a -24 periodi l'errore di revisione XE "errore di revisione" è praticamente nullo, pari a $0.7921E-01$) rispetto alla SEASONAL ADJUSTED (a -24 periodi l'errore di revisione è ancora pari a circa un terzo dell'errore totale, ossia 0.2126).

34

DECOMPOSITION OF THE SERIES: FORECAST						
PERIOD	SERIES FORECAST ERROR	TREND			ADJUSTED	
		S.E.	FORECAST	STANDARD ERROR	FORECAST	STANDARD ERROR
			TOTAL OF REVISION	TOTAL OF REVISION	TOTAL OF REVISION	TOTAL OF REVISION

PERIOD	FORECAST ERROR		TOTAL OF REVISION		FORECAST		STANDARD ERROR OF REVISION	
	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION
1	110.0	1.971	102.7	1.209	1.035	102.7	1.687	1.571
2	109.3	2.194	102.7	1.476	1.338	102.7	1.888	1.785
...								
23	111.7	6.001	102.2	5.848	5.814	102.2	5.965	5.933
24	103.7	6.155	102.2	6.060	6.027	102.2	6.173	6.142

PERIOD	ESTIMATE		SEASONAL STANDARD ERROR OF REVISION	
	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION	ESTIMATE	SEASONAL STANDARD ERROR OF REVISION
1	7.295	1.016	0.8095	
2	6.601	1.023	0.8178	
...				
23	9.496	1.274	1.116	
24	1.529	1.293	1.137	

Le valutazioni espresse nel caso precedente vengono riferite alle previsioni includendo anche l'errore standard di previsione per la serie originaria.

35

CONFIDENCE INTERVAL AROUND A SEASONAL COMPONENT OF 0				
	FINAL ESTIMATOR		CONCURRENT ESTIMATOR	
95% CONFIDENCE INTERVAL	-1.204	1.204	-1.705	1.705
70% CONFIDENCE INTERVAL	-0.6373	0.6373	-0.9022	0.9022

Si riportano gli intervalli di confidenza per lo stimatore finale XE "stimatore finale" e per lo stimatore concurrent XE "concurrent" al 95% e al 70%.

36

SAMPLE MEANS		
	COMPLETE PERIOD	LAST THREE YEARS
SERIES	89.93	102.2
TREND	89.88	102.2
ADJUSTED SEASONAL	0.4992E-01	-0.1700E-01

Si riportano la media della serie, del TREND, della SEASONAL ADJUSTED e della componente XE "componente stagionale" sia per tutto il periodo della serie che per gli ultimi tre anni. Ovviamente le medie del TREND e della SEASONAL ADJUSTED dovrebbero essere pari alla media della serie originaria mentre la media della componente stagionale dovrebbe essere pari a zero.

37

STANDARD ERROR OF ALTERNATIVE MEASURES OF GROWTH (NONANNUALIZED GROWTH)	
1	PERIOD TO PERIOD GROWTH OF THE SERIES

STANDARD ERROR OF ALTERNATIVE MEASURES OF GROWTH
(NONANNUALIZED GROWTH)

1. PERIOD TO PERIOD GROWTH OF THE SERIES

	TREND	SEASONALLY ADJ. SERIES
CONCURRENT ESTIMATOR	0.501	1.067
1 - PERIOD REVISION	0.462	1.067
2 - PERIOD REVISION	0.447	1.066
3 - PERIOD REVISION	0.441	1.066
4 - PERIOD REVISION	0.438	1.066
5 - PERIOD REVISION	0.436	1.065
6 - PERIOD REVISION	0.435	1.065
7 - PERIOD REVISION	0.434	1.065
8 - PERIOD REVISION	0.433	1.064
9 - PERIOD REVISION	0.433	1.064
10 - PERIOD REVISION	0.432	1.064
11 - PERIOD REVISION	0.432	1.024
12 - PERIOD REVISION	0.429	0.898
FINAL ESTIMATOR	0.420	0.798

Stime dell'errore standard dei tassi di crescita del TREND e della SEASONAL ADJUSTED. In particolare, si riporta la stima XE "stima" concurrent XE "concurrent", un anno di stime preliminari e il relativo stimatore finale XE "stimatore finale".

38

2. GROWTH OF A 3 - PERIOD (CENTERED) MOVING AVERAGE

	TREND	SEASONALLY ADJ. SERIES
CONCURRENT ESTIMATOR	1.114	1.787
1 - PERIOD REVISION	0.986	1.214
2 - PERIOD REVISION	0.909	1.211
3 - PERIOD REVISION	0.877	1.208
4 - PERIOD REVISION	0.861	1.204
5 - PERIOD REVISION	0.852	1.201
6 - PERIOD REVISION	0.846	1.199
7 - PERIOD REVISION	0.841	1.196
8 - PERIOD REVISION	0.837	1.193
9 - PERIOD REVISION	0.834	1.191
10 - PERIOD REVISION	0.833	1.143
11 - PERIOD REVISION	0.832	1.129
12 - PERIOD REVISION	0.825	1.114
FINAL ESTIMATOR	0.796	0.902

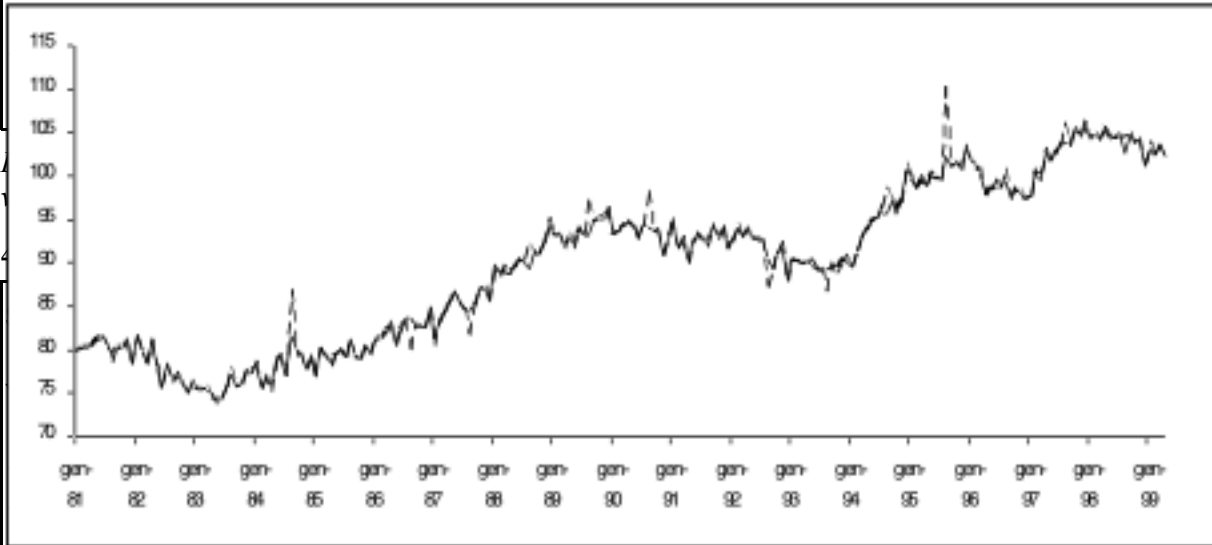
Stime dell'errore standard dei tassi di crescita di una media mobile centrata di 3 periodi.

39

3. ACCUMULATED GROWTH OVER THE LAST DECEMBER

	CONCURRENT ESTIMATOR		FINAL ESTIMATOR	
	TREND	SEASONALLY ADJ. SERIES	TREND	SEASONALLY ADJ. SERIES
JANUARY	6.009	12.806	5.043	9.578
FEBRUARY	4.897	6.896	4.071	5.152
MARCH	3.943	4.854	3.184	3.610
APRIL	3.279	3.785	2.581	2.798
MAY	2.793	3.107	2.150	2.281
JUNE	2.447	2.626	1.827	1.945

1998 101.909 105.037 113.564 106.916 111.619 112.299 114.217 54.292 110.599 113.454 111.236 94.871



1998 -2.297 0.396 9.280 1.650 7.317 6.571 9.112 -48.234 5.681 9.725 6.967 -6.165
 1999 -2.325 0.248 9.430 1.567

STANDARD ERROR OF SEASONAL

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	0.870	0.849	0.833	0.822	0.816	0.812	0.811	0.810	0.810	0.810	0.808	0.804
1982	0.715	0.704	0.697	0.692	0.689	0.687	0.687	0.687	0.687	0.686	0.686	0.684
...												
1998	0.692	0.697	0.704	0.715	0.804	0.808	0.810	0.810	0.810	0.811	0.812	0.816
1999	0.822	0.833	0.849	0.870								

TREND

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	80.520	80.603	80.718	80.860	80.969	80.872	80.579	80.365	80.350	80.364	80.230	80.076
1982	79.995	79.716	79.326	78.901	78.195	77.559	77.257	76.986	76.638	76.238	75.887	75.713
...												
1998	104.721	104.641	104.643	104.647	104.586	104.489	104.285	104.038	103.907	103.719	103.329	102.937
1999	102.791	102.809	102.802	102.725								

STANDARD ERROR OF TREND

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	0.954	0.791	0.715	0.679	0.662	0.654	0.651	0.651	0.651	0.650	0.648	0.643

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	0.000	0.000	0.790	-0.790	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1982	0.000	0.000	0.790	-0.790	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
...												
...												
1998	0.000	0.000	0.790	-0.790	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1999	0.000	0.000	0.000	0.000								
TRADING DAY EFFECT												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-0.381	-0.564	-0.381	1.574	-3.119	1.574	2.356	-3.119	1.574	-0.381	-1.163	2.356
1982	-3.119	-0.564	2.356	1.574	-3.119	1.574	-0.381	-0.381	1.574	-3.119	1.574	2.356
...												
...												
1998	-0.381	-0.564	-0.381	1.574	-3.119	1.574	2.356	-3.119	1.574	-0.381	-1.163	2.356
1999	-3.119	-0.564	.2.356	1.574								
SEASONAL REGRESSION EFFECT												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	-0.900	1.327	1.327	-0.900	-0.900	.1.327	1.327	1.327	1.327	1.327	1.327	-3.128
1982	-0.900	1.327	1.327	-0.900	1.327	1.32	1.327	1.327	1.327	1.327	-0.900	-0.900
...												
...												
1998	-3.128	1.327	1.327	-0.900	-0.900	1.327	1.327	1.327	1.327	1.327	1.327	-3.128
1999	-3.128	1.327	1.327	-0.900								

Si riportano le serie relative a i singoli effetti deterministici (in questo caso effetto Pasqua XE "Pasqua" , effetto giorni lavorativi XE "giorni lavorativi" ed effetto holidays).

44

FINAL DECOMPOSITION												

ORIGINAL UNCORRECTED SERIES (from TRAMO) XE "TRAMO"												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	80.700	81.900	90.000	84.500	84.500	88.200	89.000	34.200	89.800	89.300	86.700	72.200
1982	79.500	81.100	90.600	84.600	82.700	82.100	84.300	34.500	86.300	82.200	81.100	72.100
...												
...												
1998	98.400	105.800	115.300	106.800	107.600	114.200	117.900	52.500	113.500	114.400	111.400	94.100
1999	94.600	103.300	116.800	104.300								

E' la serie originaria.

E' la serie originaria.

45

FINAL COMPONENT												

FINAL SEASONALLY ADJUSTED SERIES												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	80.208	80.611	80.665	80.673	81.692	81.631	80.273	79.410	80.558	80.585	81.046	78.407
1982	81.774	79.801	78.459	80.910	77.845	75.573	78.275	76.668	77.072	76.243	75.025	75.715
...												
...												
1998	104.206	104.642	104.284	105.266	104.302	104.727	105.105	102.526	104.918	103.729	104.269	101.036
1999	103.172	102.290	103.687	102.159								
FINAL TREND												

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	80.520	80.603	80.718	80.860	80.969	80.872	80.579	80.365	80.350	80.364	80.230	80.076
1982	79.995	79.716	79.326	78.901	78.195	77.559	77.257	76.986	76.638	76.238	75.887	75.713
...												
...												
1998	104.721	104.641	104.643	104.647	104.586	104.489	104.285	104.038	103.907	103.719	103.329	102.937
1999	102.791	102.809	102.802	102.725								
FINAL SEASONAL												
YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1981	0.492	1.289	9.335	3.827	2.808	6.569	8.727	-45.210	9.242	8.715	5.654	-6.207
1982	-2.274	1.299	12.141	3.690	4.855	6.527	6.025	-42.168	9.228	5.957	6.075	-3.615
...												
...												
1998	-5.806	1.158	11.016	1.534	3.298	9.473	12.795	-50.026	8.582	10.671	7.131	-6.936
1999	-8.572	1.010	13.113	2.241								
FORECAST OF FINAL COMPONENT												

ORIGINAL (UNCORRECTED)	SEASONALLY ADJUSTED SERIES		TREND				SEASONAL COMPONENT					

ORIGINAL (UNCORRECTED) SERIES	SEASONALLY ADJUSTED SERIES	TREND	SEASONAL COMPONENT				
108.3	102.8	102.8	5.504				
112.4	102.9	102.9	9.502				
...							
...							
115.5	104.3	104.3	11.23				
100.8	104.4	104.4	-3.552				

Vengono riportate le componenti finali, la cui somma deve dare la serie originaria data in 44. Queste componenti si ottengono aggiungendo alle componenti stocastiche gli effetti deterministici ad esse assegnati (di cui al punto 43).

46

ELAPSED TIME : 1.1000 "

E' il tempo impiegato da SEATS XE "SEATS" per le elaborazioni.

62

CAPITOLO 4

UTILIZZO CONGIUNTO DI TRAMO XE "TRAMO" E SEATS XE "SEATS"

4.1 Introduzione

Nei due capitoli precedenti sono stati descritti separatamente i programmi TRAMO XE "TRAMO" e SEATS XE "SEATS". Già nel paragrafo 2.2.2 veniva specificato che, generalmente, i due programmi vengono usati congiuntamente impostando il parametro SEATS=2 nella stringa di istruzioni \$INPUT...\$ per il programma TRAMO. In questo capitolo si farà riferimento all'utilizzo congiunto dei due programmi; verrà inoltre definito uno schema logico che un utilizzatore potrebbe seguire per procedere alla destagionalizzazione, basato sugli strumenti di stima XE "stima" ed identificazione

vengono usati congiuntamente impostando il parametro SEATS=2 nella stringa di istruzioni \$INPUT...\$ per il programma TRAMO. In questo capitolo si farà riferimento all'utilizzo congiunto dei due programmi; verrà inoltre definito uno schema logico che un utilizzatore potrebbe seguire per procedere alla destagionalizzazione, basato sugli strumenti di stima XE "stima" ed identificazione automatica messi a disposizione da TRAMO-SEATS. Tale schema sarà poi illustrato mediante alcuni esempi relativi a serie storiche prodotte dall'ISTAT.

La scelta degli esempi segue criteri di natura didattica, per cui sono state selezionate quelle serie che possiedono contenuti esplicativi ritenuti rilevanti. Tali serie dovrebbero permettere di individuare facilmente gli aspetti salienti dello schema proposto e, in particolare, la sua flessibilità.

Più nel dettaglio, il primo esempio è relativo alle serie dei nuovi ordinativi esteri ed evidenzia l'utilità dei parametri contenuti in TRAMO-SEATS per l'identificazione automatica XE "identificazione automatica" dei modelli. Il secondo esempio fa riferimento a due serie degli indici degli occupati nelle grandi imprese dell'industria e consente di porre l'accento sia sull'utilizzo del parametro *ITER* (cfr. paragrafo 2.2.2) per l'analisi contemporanea di più serie, sia sull'imposizione di alcune scelte *a priori* da parte dell'utente.

Il terzo esempio, relativo alla serie in valore delle esportazioni italiane verso i paesi extracomunitari, tratta un caso in cui le valutazioni a priori prevalgono sulle scelte automatiche effettuate dal programma.

Infine, nel quarto esempio è illustrata la destagionalizzazione degli indici per destinazione economica della produzione industriale (beni di consumo, beni di investimento, beni intermedi). In esso si evidenzierà l'opportunità di utilizzare per le serie modelli omogenei con il modello identificato per la serie dell'indice generale della produzione industriale.

Nel floppy disk allegato sono riportati i quattro file contenenti le serie relative agli esempi.

4.2 Schema logico per procedere alla destagionalizzazione mediante TRAMO-SEATS

In questa sezione viene presentato uno schema logico che dovrebbe essere di aiuto soprattutto all'utente non esperto che si accinge ad affrontare la destagionalizzazione di una serie. La sequenza che viene proposta mira a favorire da una parte le scelte effettuate automaticamente da TRAMO-SEATS, dall'altra il rispetto di quelli che sono i canoni dell'analisi delle serie storiche.

In riferimento al primo aspetto, si può osservare come i vari passi dello schema presentino sempre i parametri INIC=3, IDIF=3, IATIP=1, LAM=-1 che consentono l'identificazione automatica XE "identificazione automatica" degli ordini del modello, degli *outlier* nonché l'opportunità della trasformazione logaritmica (cfr. paragrafi 2.2.3.3 e 2.2.4.1).

Per quanto riguarda il secondo aspetto, ogni riferimento è sempre legato alla prassi ormai consolidata dell'analisi delle serie storiche, sintetizzabile nei seguenti passaggi fondamentali:

- 1) formulazione di ipotesi sul fenomeno da analizzare e test preliminari;
- 2) impostazione del modello in base alle risultanze dei test preliminari.

In particolare, il primo passaggio è legato alle conoscenze empiriche e/o teoriche che l'utilizzatore ha del fenomeno rappresentato nella serie. Si tratta, quindi, di un passaggio esclusivamente concettuale. Le conoscenze empiriche e/o teoriche del fenomeno permettono infatti di formulare ipotesi su eventuali effetti di regressione XE "regressione", di natura deterministica, specialmente quelli dovuti agli effetti di calendario (giorni lavorativi XE "giorni lavorativi", lunghezza del mese, effetto Pasqua XE "Pasqua", festività nazionali). L'assenza di tali conoscenze rende invece necessario procedere con l'ausilio delle opzioni di default messe a disposizione da TRAMO-SEATS.

Ad esempio, è noto che la produzione industriale è influenzata dal numero di sabati e domeniche presenti nel calendario di ogni mese. In tal caso, è bene sottoporre a test la presenza dell'effetto giorni lavorativi XE "giorni lavorativi" mediante l'utilizzo del parametro ITRAD. Viceversa, il numero degli occupati rappresenta una grandezza di stock non influenzata da tale effetto, per cui ci si attende che il relativo coefficiente di regressione XE "regressione" non sia significativamente diverso da zero; può essere quindi consigliabile escluderne a priori la presenza tra gli effetti deterministici. In sostanza, si possono verificare due casi:

CASO A: esistenza di conoscenze a priori sul fenomeno analizzato.

Poiché, come accennato, le conoscenze a priori riguardano principalmente gli effetti di calendario, si possono distinguere tre diverse situazioni:

- A1: gli effetti di calendario sono sicuramente presenti e si ha intenzione di considerarli, anche se risulteranno non significativi. In tal caso, bisognerà stabilire anche se, per l'effetto giorni lavorativi XE "giorni lavorativi", il sabato sia da equiparare alla domenica (ITRAD=2), oppure se è da considerarsi alla stregua degli altri giorni lavorativi (ITRAD=7);
- A2: gli effetti di calendario potrebbero essere presenti ma si dovrà verificarne la significatività (per cui si adotteranno le specifiche per ITRAD ed IEAST con il segno negativo davanti al valore del

lavorativi XE "giorni lavorativi", il sabato sia da equiparare alla domenica (ITRAD=2), oppure se e da considerarsi alla stregua degli altri giorni lavorativi (ITRAD=7);

A2: gli effetti di calendario potrebbero essere presenti ma si dovrà verificarne la significatività (per cui si adotteranno le specifiche per ITRAD ed IEAST con il segno negativo davanti al valore del parametro). Anche in questo caso è importante stabilire se, per l'effetto giorni lavorativi XE "giorni lavorativi", il sabato sia da equiparare alla domenica (ITRAD=-1, -2), oppure da considerarsi alla stregua degli altri giorni lavorativi (ITRAD=-6, -7);

A3: gli effetti di calendario sicuramente non sono presenti. In questo caso sarà opportuno non inserire i parametri ITRAD ed IEAST, che saranno perciò posti pari al loro valore default, ossia 0.

Eventuali informazioni *a priori* potrebbero riguardare anche la scelta di utilizzare la trasformazione logaritmica o l'inserimento (o l'esclusione) di particolari *outlier*; questi argomenti saranno trattati più dettagliatamente nel capitolo 5.

CASO B: assenza di conoscenze a priori sul fenomeno analizzato.

L'assenza di informazioni a priori implica una situazione simile a quella esposta per il caso A2; sarà dunque opportuno sottoporre ancora a test la presenza degli effetti di calendario.

Si noti che sia nel caso A2 che nel caso B può risultare utile procedere in modo automatico con l'utilizzo del parametro RSA, già descritto nel paragrafo 2.2.6.

Il secondo passaggio si riferisce invece al controllo della diagnostica, descritta nei paragrafi 2.3.2 e 3.3.1 e contenuta nei file di output XE "file di output" ottenuti dall'esecuzione di TRAMO-SEATS. Se i risultati della diagnostica sono tutti accettabili, l'operazione di destagionalizzazione può ritenersi conclusa. Se, al contrario, alcuni di essi fanno ritenere che sia opportuno apportare alcune modifiche all'input utilizzato, si dovranno impostare le operazioni di identificazione e stima XE "stima" di un nuovo modello. In questo secondo caso, potrà comunque accadere che l'utente sia indeciso fra modelli alternativi; è bene allora fissare una gerarchia tra le varie diagnostiche a disposizione, in modo tale da prediligere il modello che soddisfi più degli altri le ipotesi ritenute maggiormente importanti. Nello schema seguente sarà illustrata questa possibile scala gerarchica attraverso la formalizzazione in *step* successivi dello schema finora descritto per sommi capi.

4.2.1 Schema di procedura per la destagionalizzazione

PASSO 1

Formulare, se possibile, delle ipotesi sugli effetti di regressione XE "regressione" dovuti ai giorni di calendario, ossia se il tipo di fenomeno studiato può presentare l'effetto giorni lavorativi XE "giorni lavorativi", quello dovuto alla diversa lunghezza del mese, l'effetto Pasqua XE "Pasqua" o quello relativo alle festività nazionali. In particolare, se si ritiene che possa esservi l'effetto dei giorni lavorativi, stabilire se il sabato sia da considerarsi alla stregua degli altri giorni lavorativi o se sia da equiparare alla domenica. Come precedentemente fatto, sarà definito A il caso in cui si hanno conoscenze a priori sul fenomeno, distinguendo tra il caso di certezza della presenza degli effetti di calendario (A1), il caso di incertezza (A2) e quello di certezza di assenza degli effetti di calendario (A3). Sarà definito come B il caso di assenza di informazioni a priori, che è equiparabile al caso A2.

PASSO 2

CASO A1:

porre

```
$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=2 (7) IEAST=1 IATIP=1 SEATS XE "SEATS" =2 IREG=1$
```

```
$REG ...$
```

IREG si riferisce al file *holidays*.

Eseguire il modello specificato.

CASI A2 e B:

porre

```
$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=-2 (-7) IEAST=-1 IATIP=1 SEATS XE "SEATS" =2 IREG=1$
```

```
$REG ...$
```

IREG si riferisce al file *holidays*.

Eseguire il modello. Se il pre-test inserisce nel modello i giorni lavorativi XE "giorni lavorativi", verificare nell'output di TRAMO XE "TRAMO" se il parametro relativo al regressore TRAD 2 non è significativo; in quest'ultima evenienza, sostituire ITRAD=-2 (oppure -7) con ITRAD=-1 (-6) ed eseguire il modello.

CASO A3:

porre

```
$INPUT INIC=3 IDIF=3 LAM=-1 IATIP=1 SEATS XE "SEATS" =2 $
```

Eseguire il modello

CASO A3:

porre

$\$INPUT\ INIC=3\ IDIF=3\ LAM=-1\ IATIP=1\ SEATS\ XE\ "SEATS" =2\ \$$

Eseguire il modello.

PASSO 3 (solo per i CASI A2 e B)

Verificare la significatività del regressore *holiday* e, se questo risulta significativo, andare al passo 4; altrimenti, usare l'input:

$\$INPUT\ INIC=3\ IDIF=3\ LAM=-1\ ITRAD=-2\ (-7)\ IEAST=-1\ IATIP=1\ SEATS\ XE\ "SEATS" =2\ \$$

Eseguire il modello. Se il *pre-test* inserisce nel modello i giorni lavorativi $XE\ "giorni\ lavorativi"$, verificare nell'output di TRAMO $XE\ "TRAMO"$ se il parametro relativo al regressore TRAD 2 non è significativo; in quest'ultima evenienza, sostituire ITRAD=-2 (oppure -7) con ITRAD=-1 (-6) ed eseguire il modello (questa operazione, già effettuata nel Passo 2, viene riproposta in questo passo poiché la presenza delle *holiday* potrebbe aver influito sulla stima $XE\ "stima"$ del parametro relativo a TRAD 2).

PASSO 4

Controllare la diagnostica dei residui sia di TRAMO $XE\ "TRAMO"$ che di SEATS $XE\ "SEATS"$. Se il modello supera i test, la procedura termina. Se non li supera, andare al Passo 5.

PASSO 5

Se il problema riguarda il test di normalità, ridurre progressivamente il parametro VA (prima pari a 3.4, poi a 3.3, ecc.) ed eseguire il modello, che dovrebbe individuare nuovi *outlier*. Se anche questa operazione non migliora la diagnostica, provare modelli ARIMA $XE\ "ARIMA"$ alternativi, lasciando fisso l'ordine di differenziazione ed i regressori esterni seguendo la metodologia proposta di Box e Jenkins $XE\ "Box\ e\ Jenkins"$, basata sull'analisi delle autocorrelazioni $XE\ "autocorrelazioni"$. Se non si ottengono miglioramenti, mantenere il modello del punto 4.

Se si è indecisi tra modelli alternativi, scegliere in base a quello che possiede la migliore diagnostica relativa ai seguenti aspetti, elencati in ordine di importanza:

autocorrelazioni $XE\ "autocorrelazioni"$ in TRAMO $XE\ "TRAMO"$ (test di Ljung-Box $XE\ "Ljung-Box"$);

test di normalità in TRAMO $XE\ "TRAMO"$;

BIC $XE\ "BIC"$ in TRAMO $XE\ "TRAMO"$;

autocorrelazioni $XE\ "autocorrelazioni"$ in SEATS $XE\ "SEATS"$ (test di Ljung-Box $XE\ "Ljung-Box"$ e di Pierce $XE\ "Pierce"$);

test di normalità in SEATS $XE\ "SEATS"$;

autocorrelazioni $XE\ "autocorrelazioni"$ sui residui al quadrato in TRAMO $XE\ "TRAMO"$ (test di Ljung-Box $XE\ "Ljung-Box"$ sui residui al quadrato);

autocorrelazioni $XE\ "autocorrelazioni"$ sui residui al quadrato in SEATS $XE\ "SEATS"$ (test di Ljung-Box $XE\ "Ljung-Box"$ e di Pierce $XE\ "Pierce"$ sui residui al quadrato).

4.3 Esempi

4.3.1 La serie dei nuovi ordinativi esteri

Si consideri la serie dell'indice dei nuovi ordinativi esteri (ORDEST) per il periodo gennaio 1990-aprile 1999. Il file contenente i dati per replicare l'esercizio ha il nome *esempio 2*.

E' presumibile che la serie degli ordinativi contenga effetti di calendario, legati ai cinque giorni lavorativi $XE\ "giorni\ lavorativi"$ settimanali, per cui viene considerato il caso A2; per quanto riguarda l'uso del parametro ITRAD, esso è posto pari a -1 e -2. Dovrebbero essere presenti anche l'effetto *holidays* e l'effetto Pasqua $XE\ "Pasqua"$. Tuttavia, non essendo una serie riferita esplicitamente a fenomeni di produzione, queste ipotesi a priori non sono da ritenersi vincolanti.

Il primo modello esplicitato è dunque caratterizzato dai seguenti parametri:

$\$INPUT\ INIC=3\ IDIF=3\ LAM=-1\ ITRAD=-2\ IEAST=-1\ IATIP=1\ SEATS\ XE\ "SEATS" =2\ IREG=1\ \$$

$\$REG\ IUSER=-1\ NSER=1\ ILONG=136\ REGEFF=2\ \$$

hol_90

dove il file allegato *hol_90* contiene i giorni festivi non relativi ai sabati e alle domeniche, a partire da gennaio 1990.

I risultati (sintetizzati nella tabella 4.1, alla voce **Modello 1**) evidenziano la non significatività dell'effetto Pasqua $XE\ "Pasqua"$ (che non viene perciò inserito nel modello) e la scarsa significatività dell'effetto *holidays*, mentre viene inserito l'effetto giorni lavorativi $XE\ "giorni\ lavorativi"$; il regressore relativo alla lunghezza del mese (TRAD 2) non è però significativo. Viene individuata la presenza di un *outlier* di tipo additivo in corrispondenza di agosto 1994. Il modello ARIMA $XE\ "ARIMA"$ prescelto è un *airline* moltiplicativo (è stata effettuata infatti la trasformazione logaritmica) e, in generale, la diagnostica sui residui risulta buona.

Il successivo passo consiste nell'eliminazione del regressore relativo alle *holidays* che non risulta significativo; l'input per creare il **Modello 2** è dunque:

$\$INPUT\ INIC=3\ IDIF=3\ LAM=-1\ ITRAD=-2\ IEAST=-1\ IATIP=1\ SEATS\ XE\ "SEATS" =2\ \$$

Il successivo passo consiste nell'eliminazione del regressore relativo alle *holidays* che non risulta significativo; l'input per creare il **Modello 2** è dunque:

$\$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=-2 IEAST=-1 IATIP=1 SEATS XE "SEATS" =2 \$$

Il modello ARIMA XE "ARIMA" identificato in precedenza non cambia e viene identificato lo stesso *outlier* additivo XE "outlier additivo". Anche in questo caso la diagnostica non presenta particolari problemi anche se il regressore TRAD 2 continua a non essere significativo. Di conseguenza, in base allo schema presentato in precedenza, il passo successivo comporta l'eliminazione anche di questo regressore tramite l'input:

$\$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=-1 IEAST=-1 IATIP=1 SEATS XE "SEATS" =2 IREG=1 \$$

Il **Modello 3** può ritenersi il modello finale, in quanto presenta una buona diagnostica per i residui, non contiene regressori non significativi e non contraddice le ipotesi di partenza, che, comunque, non erano particolarmente restrittive.

Tabella 4.1 - Sintesi dell'output di TRAMO XE "TRAMO" (prime 8 righe) e SEATS XE "SEATS" (ultime 5 righe) per la serie ORDEST (i regressori non significativi e le statistiche test che rifiutano l'ipotesi nulla sono evidenziati in grassetto).

	MODELLO 1	MODELLO 2	MODELLO 3
Logaritmo	Sì	Sì	Sì
ARIMA XE "ARIMA"	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)
Regressori	TRAD 1 TRAD 2 HOLIDAYS	TRAD 1 TRAD 2	TRAD 1
Outlier XE "Outlier"	Ago 94 AO XE "AO"	Ago 94 AO XE "AO"	Ago 94 AO XE "AO"
BIC XE "BIC"	-5,2625	-5,2991	-5,3354
Jarque-Bera	1,868	2,148	2,451
Ljung-Box XE "Ljung-Box" res.	12,92	10,63	11,17
Ljung-Box XE "Ljung-Box" res. ^2	20,52	20,64	23,31
Jarque-Bera	1,060	18,71	1,789
Ljung-Box XE "Ljung-Box" res.	14,45	12,74	13,12
Pierce XE "Pierce" res.	3,22	2,99	2,79
Ljung-Box XE "Ljung-Box" res. ^2	17,89	14,06	14,20
Pierce XE "Pierce" res. ^2	1,64	2,16	1,95

4.3.2 Le serie dell'occupazione nelle grandi imprese

Si considerino le serie relative al numero indice degli occupati nella grande industria al netto (TOTOCC) ed al lordo della cassa integrazione (TOTOCCCL) da gennaio 1993 a novembre 1999. Le serie sono allegate nel file *esempio3*.

Le serie relative agli occupati non possono risentire degli effetti di calendario in maniera significativa per cui ci si trova nella situazione descritta dal CASO A3. Data la presenza di due serie, potrà essere utilizzato il parametro *ITER* per analizzarle contemporaneamente (cfr. paragrafo 2.2.2). In altri termini, il **Modello 1** (i cui risultati sono sintetizzati nelle tabelle 4.2 e 4.3) per entrambe le serie si otterrà con il seguente input (posto al termine della prima serie):

$\$INPUT ITER=2 INIC=3 IDIF=3 LAM=-1 IATIP=1 SEATS XE "SEATS" =2 \$$

Per la serie TOTOCC viene identificato un modello ARIMA XE "ARIMA" (0,1,0)(0,1,1) moltiplicativo, senza alcun *outlier*, che presenta nel complesso una buona diagnostica. Per la serie TOTOCCCL, invece, viene selezionato un modello *airline* moltiplicativo, con un *outlier* additivo XE "outlier additivo", che presenta problemi nella diagnostica dei residui, in particolare nel test di normalità sia in TRAMO XE "TRAMO" che in SEATS XE "SEATS". Mentre per la prima serie si può accettare il modello, per la seconda è necessario cercarne uno alternativo che presenti una diagnostica migliore. Data la presenza di un elevato valore del test di normalità, il tentativo seguente mira a ridurre progressivamente il parametro VA finché non si individua un nuovo *outlier*.

A questo punto, si potrebbe inserire l'input per la sola serie TOTOCCCL, ma per esemplificare l'uso del parametro *ITER* vengono esplicitati i modelli per entrambe le serie.

Il nuovo *outlier* viene individuato attraverso l'input al termine della prima serie:

A questo punto, si potrebbe inserire l'input per la sola serie TOTOCCL, ma per esemplificare l'uso del parametro *ITER* vengono esplicitati i modelli per entrambe le serie.

Il nuovo outlier viene individuato attraverso l'input al termine della prima serie:

$\$INPUT\ ITER=3\ INIC=3\ IDIF=3\ LAM=-1\ IATIP=1\ SEATS\ XE\ "SEATS" =2\ \$$

ed al termine della seconda:

$\$INPUT\ INIC=3\ IDIF=3\ LAM=-1\ IATIP=1\ VA=3.2\ SEATS\ XE\ "SEATS" =2\ \$$

L'outlier individuato in dicembre 1997 è di tipo *level shift* (**Modello 2**); come si può notare, si verifica un miglioramento nel test di normalità in TRAMO XE "TRAMO", ma non in SEATS XE "SEATS". Nel capitolo 5 verrà evidenziato come questa situazione possa essere dovuta proprio all'introduzione del *level shift* e si suggerirà un'alternativa per migliorare il modello. In questa fase, invece, in base alle gerarchie citate al Passo 5 dello schema proposto nel paragrafo 4.2.1, il Modello 2 è da preferirsi e viene scelto per destagionalizzare la serie TOTOCCL.

Tabella 4.2 - Sintesi dell'output di TRAMO XE "TRAMO" (prime 8 righe) e SEATS XE "SEATS" (ultime 5 righe) per la serie TOTOCCL (i regressori non significativi e le statistiche test che rifiutano l'ipotesi nulla sono evidenziati in grassetto).

	MODELLO 1
Logaritmo	Sì
ARIMA XE "ARIMA"	(0,1,0)(0,1,1)
Regressori	
Outlier XE "Outlier"	
BIC XE "BIC"	-10,7053
Jarque-Bera	1,291
Ljung-Box XE "Ljung-Box" res.	32,34
Ljung-Box XE "Ljung-Box" res. ^2	16,51
Jarque-Bera	0,3868
Ljung-Box XE "Ljung-Box" res.	25,21
Pierce XE "Pierce" res.	3,05
Ljung-Box XE "Ljung-Box" res. ^2	26,71
Pierce XE "Pierce" res. ^2	3,69

Tabella 4.3 - Sintesi dell'output di TRAMO XE "TRAMO" (prime 8 righe) e SEATS XE "SEATS" (ultime 5 righe) per la serie TOTOCCL (i regressori non significativi e le statistiche test che rifiutano l'ipotesi nulla sono evidenziati in grassetto).

	MODELLO 1	MODELLO 2
Logaritmo	Sì	Sì
ARIMA XE "ARIMA"	(0,1,1)(0,1,1)	(0,1,1)(0,1,1)
Regressori		
Outlier XE "Outlier"	Dic 96 AO XE "AO"	Dic 96 AO XE "AO" Dic 97 LS XE "LS"
BIC XE "BIC"	-11,7210	-11,8160
Jarque-Bera	10,29	5,711
Ljung-Box XE "Ljung-Box" res.	20,16	17,72
Ljung-Box XE "Ljung-Box" res. ^2	23,10	12,64
Jarque-Bera	20,96	12,99
Ljung-Box XE "Ljung-Box" res.	19,77	17,67
Pierce XE "Pierce" res.	4,57	4,50
Ljung-Box XE "Ljung-Box" res. ^2	17,66	15,09
Pierce XE "Pierce" res. ^2	5,61	0,21

4.3.3 La serie delle esportazioni verso i paesi extra-UE

Si consideri la serie in valori relativa alle esportazioni italiane verso i paesi extracomunitari da gennaio 1988 a dicembre 1998. La serie è contenuta nel file *esempio 4*.

Essendo una serie relativa al commercio, essa è presumibilmente legata in maniera abbastanza significativa a tutti gli effetti di calendario per cui i regressori relativi a giorni lavorativi $XE^{*}_{\text{giorni lavorativi}}$, lunghezza del mese, *holidays* ed effetto Pasqua XE^{*}_{Pasqua} , dovrebbero essere tenuti in considerazione nel modello finale. La situazione coincide con quella che nel paragrafo 4.2.1 è stata definita CASO A1.

Nel modello, quindi, imporremo a priori i vari effetti deterministici tramite il seguente input (**Modello 1** in tabella 4.4):

```
$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=2 IEAST=1 IATIP=1 SEATS XE "SEATS" =2 IREG=1$
$REG IUSER=-1 NSER=1 ILONG=154 REGEFF=2$
hol_88
```

dove *hol_88* è il file allegato contenente le *holidays* a partire da gennaio 1988.

Il modello selezionato è un ARIMA $XE^{*}_{\text{ARIMA}} (2,1,0)(0,1,1)$ moltiplicativo senza *outlier* che evidenzia la non significatività di tutti gli effetti di calendario (ad eccezione di TRAD 1) per cui è in parte in contraddizione con le ipotesi iniziali; i risultati della diagnostica sono, comunque, complessivamente buoni e ciò induce ad accettare il modello. Può essere utile, come esercizio, confrontare questi risultati con quelli che si sarebbero ottenuti utilizzando una procedura automatica, ossia ipotizzando una situazione di mancanza di informazioni a priori (CASO B). In altri termini, l'esercizio verte a stimare anche i modelli che potrebbero ottenersi impostando i seguenti input (diversi dai precedenti soltanto per la presenza o meno del regressore TRAD 2):

Modello 2:

```
$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=-2 IEAST=-1 IATIP=1 SEATS XE "SEATS" =2 IREG=1$
$REG IUSER=-1 NSER=1 ILONG=154 REGEFF=2$
hol_88
```

Modello 3:

```
$INPUT INIC=3 IDIF=3 LAM=-1 ITRAD=-1 IEAST=-1 IATIP=1 SEATS XE "SEATS" =2 IREG=1$
$REG IUSER=-1 NSER=1 ILONG=154 REGEFF=2$
hol_88
```

Come si nota agevolmente dalla tabella 4.4, non si riscontrano sostanziali differenze rispetto al **Modello 1** in termini di identificazione e diagnostica. Ciò evidenzia, di conseguenza, come mentre dal punto di vista puramente statistico bisognerebbe eliminare dal modello sia le *holidays* sia l'effetto degli anni bisestili, sulla base delle ipotesi precedentemente considerate si dà grosso peso alla presenza degli effetti di calendario che vengono dunque mantenuti. I risultati in termini di dati destagionalizzati sono comunque equivalenti; ciò è ben messo in risalto dalla figura 4.1, in cui sono riportate le variazioni congiunturali per il 1998 distintamente per i vari modelli e in cui risulta immediato constatare come le discrepanze siano, in realtà, minime.

L'utilizzo di un modello meno efficiente dal punto di vista statistico, ma coerente con le ipotesi a priori, garantisce anche una maggiore stabilità. Infatti, nell'esempio, EAST ha un t-ratio $XE^{*}_{\text{t-ratio}}$ pari a -1,32, per cui è possibile che l'aggiunta di nuove osservazioni renda tale regressore significativo al punto da rendere necessaria in tempi brevi una revisione del modello.

Tabella 4.4: Sintesi dell'output di TRAMO XE^{*}_{TRAMO} (prime 8 righe) e SEATS XE^{*}_{SEATS} (ultime 5 righe) per la serie EXP_XUE (i regressori non significativi e le statistiche test che rifiutano l'ipotesi nulla sono evidenziati in grassetto).

	MODELLO 1	MODELLO 2	MODELLO 3
Logaritmo	Sì	Sì	Sì
ARIMA XE^{*}_{ARIMA}	(2,1,0)(0,1,1)	(2,1,0)(0,1,1)	(2,1,0)(0,1,1)
Regressori	TRAD 1 TRAD 2 EAST HOLIDAYS	TRAD 1 TRAD 2 HOLIDAYS	TRAD 1 HOLIDAYS
Outlier XE^{*}_{Outlier}			
BIC XE^{*}_{BIC}	-5,0897	-5,1166	-5,1559
Jarque-Bera	2,838	2,578	2,753
Ljung-Box $XE^{*}_{\text{Ljung-Box}}$ res.	14,45	15,43	15,65

BIC XE "BIC"	-5,0697	-5,1100	-5,1559
Jarque-Bera	2,838	2,578	2,753
Ljung-Box XE "Ljung-Box" res.	14,45	15,43	15,65
Ljung-Box XE "Ljung-Box" res. ^2	27,69	29,63	30,12
Jarque-Bera	5,248	4,449	4,443
Ljung-Box XE "Ljung-Box" res.	14,08	15,46	15,74
Pierce XE "Pierce" res.	0,61	0,91	0,94
Ljung-Box XE "Ljung-Box" res. ^2	24,32	25,21	25,35
Pierce XE "Pierce" res. ^2	3,48	3,25	3,36

Figura 4.1 – Esportazioni extra-UE: variazioni congiunturali 1998 su dati destagionalizzati con modello 1 (barre grigie), modello 2 (barre bianche), modello 3 (barre nere)

4.3.4 Le serie dei settori della produzione industriale

Quest'ultimo esempio presenta l'analisi dei settori della produzione industriale da gennaio 1981 ad aprile 1999, la cui aggregazione ponderata fornisce la serie dell'indice generale della produzione industriale già trattata, più nel dettaglio, nei precedenti capitoli. In questo paragrafo, tuttavia, non sarà applicata la procedura fin qui proposta, ma un tipo di approccio diverso, poiché nella pratica è frequente il caso in cui un Ente produttore di dati si trovi nella necessità di destagionalizzare un gran numero di serie tra loro omogenee, come accade, appunto, nella produzione industriale in cui, anche supponendo di limitare la destagionalizzazione solo alle sezioni e sottosezioni, si devono diffondere contemporaneamente ben 22 serie. In casi come questo, può essere conveniente, anziché applicare l'intera procedura, sfruttare il modello già ottenuto per l'indice generale e verificare se si adatta bene anche alle altre serie. Per semplicità espositiva, l'approccio scelto in questo contesto sarà limitato solo ai tre settori della produzione industriale, ossia beni di consumo (CON81), beni di investimento (INV81) e beni intermedi (INT81).

Nel file *esempio 5* sono riportate in sequenza le tre serie; ricordando che nel **Modello 4** della tabella 2.1 (cfr. paragrafo 2.4), la soluzione prescelta utilizzava un modello *airline* additivo, senza costante e con regressori costituiti da *holidays*, TRAD 1, TRAD 2, EAST, l'input che si inserisce alla fine della prima serie è dato da:

```
$INPUT ITER=3 LAM=1 ITRAD=2 IEAST=1 IMEAN=0 SEATS XE "SEATS" =2 IREG=1$
$REG IUSER=-1 NSER=1 ILONG=244 REGEFF=2$
hol_81
```

e per le altre due serie:

```
$REG IUSER=-1 NSER=1 ILONG=244 REGEFF=2$
hol_81
```

Nella tabella 4.5 sono riportati i principali risultati dell'output. Si noti come tutte e tre le serie possano essere ben rappresentate dal modello in questione; l'unico elemento della diagnostica che non viene accettato al 95% di significatività è il test di Ljung-Box XE "Ljung-Box" di SEATS XE "SEATS" per la serie INT81, nonostante il valore della statistica sia abbastanza vicino al valore critico (infatti verrebbe accettato al 99% di significatività). La serie CON81 presenta un regressore non significativo (TRAD 2), ma il suo t-value è pari a 1,66, dunque abbastanza ai limiti della zona di significatività.

In definitiva, il modello usato per l'indice generale si adatta bene anche agli indici relativi ai tre settori, per cui può essere adottato per essi senza particolari controindicazioni.

ma il suo t-value è pari a 1,66, dunque abbastanza ai limiti della zona di significatività.

In definitiva, il modello usato per l'indice generale si adatta bene anche agli indici relativi ai tre settori, per cui può essere adottato per essi senza particolari controindicazioni.

Tabella 4.5 - Sintesi dell'output di TRAMO XE "^{TRAMO} (prime 8 righe) e SEATS XE "^{SEATS} (ultime 5 righe) per le serie CON81, INV81, INT81 (nella riga dei regressori sono riportati solo quelli non significativi, mentre le statistiche test che rifiutano l'ipotesi nulla sono evidenziate in grassetto).

	CON81	INV81	INT81
Logaritmo	No	No	No
ARIMA XE " ^{ARIMA}	(0,1,1)	(0,1,1)	(0,1,1)
Regressori	TRAD 1	TRAD 1	TRAD 1
	TRAD 2	TRAD 2	TRAD 2
	EAST	EAST	EAST
	HOLIDAYS	HOLIDAYS	HOLIDAYS
Jarque-Bera	0,0478	3,042	1,171
Ljung-Box XE " ^{Ljung-Box} res.	25,84	14,49	30,33
Ljung-Box XE " ^{Ljung-Box} res. ^2	20,07	24,15	14,42
Jarque-Bera	0,006	2,868	1,211
Ljung-Box XE " ^{Ljung-Box} res.	25,94	14,54	34,87
Pierce XE " ^{Pierce} res.	3,60	0,95	5,69
Ljung-Box XE " ^{Ljung-Box} res. ^2	20,27	20,76	18,07
Pierce XE " ^{Pierce} res. ^2	0,68	1,03	4,63

CAPITOLO 5

PROBLEMI APERTI E CASI DI STUDIO

In questo quinto capitolo vengono affrontate alcune problematiche legate alla destagionalizzazione, per le quali non è stata ancora proposta in letteratura una soluzione definitiva. Tali problematiche sono essenzialmente quella delle revisioni, della stabilità del modello, dell'aggregazione di serie ed altre di varia natura che frequentemente si incontrano nelle applicazioni pratiche. La trattazione si limiterà ad una breve illustrazione e ad alcuni consigli pratici per l'utente. I temi esaminati in questo capitolo hanno ovviamente un carattere generale che pur prescindendo dall'impiego di TRAMO-SEATS, rivestono spesso un'importanza significativa negli aspetti empirici dell'attività di destagionalizzazione.

5.1 Il problema delle revisioni

La problematica relativa alle revisioni nasce dal fatto che l'operazione di destagionalizzazione sfrutta tutta l'informazione disponibile relativa alla serie storica analizzata; ogni dato che si aggiunge modifica tale informazione e produce un conseguente cambiamento (revisione) della serie destagionalizzata XE "serie destagionalizzata". In particolare, si può verificare una delle seguenti situazioni.

- a) identificazione dello stesso modello precedente con gli stessi effetti deterministici ed approssimativamente la stessa stima XE "stima" dei coefficienti incogniti; in questo caso potranno comunque esserci cambiamenti nella parte finale della serie, dovuti alle differenti previsioni che si potrebbero ottenere con l'aggiunta di un nuovo dato;
- b) identificazione dello stesso modello con gli stessi effetti deterministici ma una diversa stima XE "stima" dei coefficienti incogniti, dovuta ad esempio al fatto che alcuni dei nuovi dati possono configurarsi come valori anomali; in questo caso l'intera serie destagionalizzata XE "serie destagionalizzata" potrebbe subire una profonda revisione;
- c) identificazione dello stesso modello ma con diversi effetti deterministici, quali, ad esempio, l'identificazione di nuovi outlier o la perdita di significatività di altri effetti deterministici; anche in questo caso si verifica una modifica delle stime dei coefficienti ed una revisione della serie destagionalizzata XE "serie destagionalizzata";
- d) identificazione di un modello diverso; in questo caso è probabile che la serie destagionalizzata XE "serie destagionalizzata" cambi radicalmente, a meno che modelli diversi portino ad inferenze simili, in termini di stima XE "stima" delle componenti.

La scelta di una metodologia e di criteri per il trattamento delle revisioni deve, perciò, tenere conto di tutti gli aspetti appena descritti, ma anche di alcune considerazioni di carattere pratico, legate al fatto che modifiche radicali in una serie destagionalizzata XE "serie destagionalizzata" possono essere dovute anche ad aspetti temporanei, non ancora analizzabili per insufficienza di dati. La possibilità di eventuali shock che in futuro il modello è in grado di assorbire può rendere, quindi, preferibile mantenere fisso il modello nel corso di un periodo prefissato (circa un anno) in termini di ordini dell'ARIMA XE "ARIMA" e di effetti deterministici, e provvedere ad una sua eventuale nuova identificazione successivamente (come consigliato anche da Gomez e Maravall, 1998, in Appendice, pagg. 39-40). Considerando tali problemi, risulta pertanto di fondamentale importanza che nella fase di identificazione del modello sia verificata la sua stabilità; questa operazione potrà essere effettuata anche mediante simulazioni XE "simulazioni", come si vedrà nel prossimo paragrafo.

Dal punto di vista pratico, è possibile utilizzare quattro approcci diversi, la cui efficacia va valutata in relazione all'entità delle revisioni, alla luce di considerazioni sia pratiche che metodologiche:

- 1) *Utilizzo delle proiezioni del fattore (componente XE "componente") stagionale*: Consiste nell'ottenere i valori destagionalizzati sottraendo (dividendo) ai valori grezzi la proiezione della componente (fattore) stagionale. Nel momento in cui si effettua la scomposizione XE "output di SEATS YE "rendrà disponibili per H periodi successivi al tempo T

- 1) *Utilizzo delle proiezioni del fattore (componente XE "componente") stagionale*: Consiste nell'ottenere i valori destagionalizzati sottraendo (dividendo) ai valori grezzi la proiezione della componente (fattore) stagionale. Nel momento in cui si effettua la scomposizione XE "scomposizione", l'output di SEATS XE "SEATS" renderà disponibili, per H periodi successivi al tempo T , le previsioni per il fattore (componente) stagionale nella sezione "FORECAST OF FINAL COMPONENT" oppure direttamente nel file *fsfn.t5* contenuto nella directory di SEATS denominata GRAPH\FORECAST. Per default, in questo file le previsioni sono gli ultimi 24 valori della serie se i dati sono mensili. Di conseguenza, se il modello è stato stimato con i dati fino al tempo T , si avranno a disposizione gli S fattori (componenti) stagionali per H periodi successivi.

Per ottenere la serie destagionalizzata XE "serie destagionalizzata" sa per gli H periodi successivi a T , sarà sufficiente effettuare una delle seguenti operazioni su i nuovi dati y :

$sa(T+h)=y(T+h)/S(T+h)$ ($h=1, \dots, H$) se il modello è moltiplicativo

$sa(T+h)=y(T+h)-S(T+h)$ ($h=1, \dots, H$) se il modello è additivo

questo metodo è semplice dal punto di vista pratico, in quanto richiede soltanto la memorizzazione dei fattori (componenti) stagionali per l'intero periodo in cui il modello è tenuto fisso. D'altra parte, però, poiché esso si basa solo sulle previsioni dei fattori stagionali, può determinare forti revisioni nella serie destagionalizzata XE "serie destagionalizzata" quando si stima XE "stima" nuovamente il fattore (componente XE "componente") stagionale. In linea di massima, con un metodo di questo tipo, è bene non tenere fisso il modello per periodi superiori ai 4-6 mesi.

- 2) *Utilizzo del medesimo modello nel corso dell'anno*. Questo metodo consiste nel fissare oltre al modello identificato anche i coefficienti stimati al tempo T , e nell'applicarlo ogni volta che nuovi dati si rendono disponibili. In altri termini la fase di identificazione e di stima XE "stima" non viene effettuata nel periodo in cui il modello viene tenuto fisso e durante il quale si possono verificare revisioni di entità contenuta imputabile all'errore di previsione o a cambiamenti nella parte deterministica del modello (dovuti, ad esempio, al sopraggiungere di valori anomali o di cambiamenti strutturali nei coefficienti del modello per il periodo considerato). Ciò si può verificare perché l'attuale versione di TRAMO-SEATS consente di fissare solo i coefficienti della parte ARIMA XE "ARIMA" (con il parametro INIT=2 e gli altri ad esso connessi, (cfr. la descrizione completa dei parametri nell'appendice finale), mentre questa operazione non è prevista per la parte deterministica. E' probabile, tuttavia, che si verifichino revisioni consistenti quando si procede ad identificare un nuovo modello per la serie in oggetto.
- 3) *Utilizzo del concurrent XE "concurrent"*. Questo metodo consiste nel fissare il modello al tempo T , compresi gli outlier e gli altri effetti deterministici, stimando però i coefficienti ogni volta che un nuovo dato è aggiunto alla serie (stima XE "stima concurrent"). Rispetto al caso 2), l'approccio è più flessibile in quanto consente ai coefficienti di variare in modo tale da adattarsi quanto più possibile ai nuovi dati. I dati destagionalizzati sono aggiornati ogni volta che una nuova osservazione si rende disponibile, dunque la serie destagionalizzata XE "serie destagionalizzata" cambierà ad ogni aggiornamento. Questo metodo ha l'indubbio vantaggio di comportare revisioni di minore entità rispetto ai precedenti ogni volta che l'intero modello viene identificato e stimato.
- 4) *Identificazione XE "identificazione" del modello per ogni nuova informazione*. Questo metodo consiste nel procedere all'identificazione e stima XE "stima" del modello non appena un nuovo dato si rende disponibile. In questo caso sussiste, ad esempio, il rischio di cambiare modello soltanto a causa di dati anomali, la cui importanza può invece ridimensionarsi nel corso del tempo.

Ovviamente, se la serie è sufficientemente regolare e la procedura di identificazione ha individuato un valido modello, sia in termini di adattamento ai dati che di flessibilità, i quattro approcci non dovrebbero portare a grosse revisioni né a risultati molto differenti. In caso contrario, in base a quanto o in precedenza, il primo tipo di approccio non è consigliabile in quanto presenta maggiori possibilità di errori. Tenendo conto che la stima XE "stima" del modello, una volta stabiliti i parametri di input, non richiede calcoli onerosi, l'approccio 3) sembra preferibile, anche per le informazioni che ne possono scaturire in termini di diagnostica e valutazione della stabilità del modello stesso. In altri termini, con tale approccio è possibile verificare, oltre alle consuete ipotesi statistiche descritte nel secondo capitolo di questa guida, anche la stabilità delle stime nel periodo considerato e da esse valutare l'opportunità di procedere ad una nuova identificazione del modello o meno. A questo proposito, si ricorda che nell'output di SEATS XE "SEATS" sono riportate diverse informazioni utili all'utente per valutare l'entità dei vari tipi di errore; in particolare, nella sezione relativa agli errori di stima, è riportata la riduzione dell'errore di previsione XE " " ottenuta con l'impiego della

valutare l'opportunità di procedere ad una nuova identificazione del modello o meno. A questo proposito, si ricorda che nell'output di SEATS XE "SEATS" sono riportate diverse informazioni utili all'utente per valutare l'entità dei vari tipi di errore; in particolare, nella sezione relativa agli errori di stima, è riportata la riduzione dell'errore di revisione XE "errore di revisione" ottenuta con l'impiego dello stimatore concurrent XE "concurrent" (cfr. paragrafo 3.3.1.3 e punto 32 dell'Appendice al capitolo 3).

In definitiva, grosse revisioni dovrebbero spingere l'utente ad analizzare nuovamente la serie aggiornata con maggiore attenzione e valutare l'opportunità di una eventuale nuova identificazione del modello.

A titolo di esempio analizziamo l'andamento dei dati destagionalizzati dell'indice generale della produzione industriale per il 1998 con i quattro diversi approcci, avendo identificato e stimato il modello per il periodo 1981-1997.

Le istruzioni di input per la serie in questione sono le seguenti:

```
$INPUT IMEAN=0 LAM=1 IEAST=1 ITRAD=2 IREG=1 SEATS XE "SEATS" =2$
$REG IUSER=-1 NSER=1 ILONG=240 REGEFF=2 $
hol_81
```

relative, cioè, ad un modello airline additivo, con effetto Pasqua XE "Pasqua" (di durata 6 giorni), effetto giorni lavorativi XE "giorni lavorativi" con due regressori ed effetto holidays come regressore esterno.

In tabella 5.1 sono riportati i risultati con i primi due approcci, mentre in tabella 5.2 si evidenzia la dinamica delle revisioni durante il 1998 (l'approccio 3 si riferisce alla penultima colonna di quest'ultima tabella); i dati finali si riferiscono alla stima XE "stima" ottenuta per il periodo 1981-1998 dopo aver introdotto nella serie il dato relativo a dicembre 1998. Anche con l'ausilio grafico della figura 5.1, si nota come i tre metodi non pervengano a grosse differenze, grazie alla stabilità del modello ed alla robustezza delle stime ottenute. La stabilità può essere ulteriormente messa in evidenza dalla totale uguaglianza del secondo e del terzo approccio; ciò equivale ad affermare che lasciare fissi i coefficienti o stimarli nel corso del 1998 è, di fatto, indifferente grazie alla costanza delle stime nel tempo. Leggendo la tabella 5.2 per riga si osserva come, rivedendo con l'approccio 3) i dati ogni mese, le modifiche effettuate siano molto piccole e graduali.

La serie della produzione industriale è molto regolare e pertanto non si evidenziano eccessive differenze tra i tre approcci.

Problemi notevoli, invece, si possono avere quando cambiano i coefficienti o addirittura il modello nel corso dell'anno. A tale proposito, consideriamo come esempio, la serie dell'indice generale del totale degli ordinativi (file *esempio 6*). L'analisi dei dati da gennaio 1990 al dicembre 1997 con l'opzione RSA=4, individua un modello ARIMA XE "ARIMA" (3,1,1)(0,1,1) senza costante, con un regressore per i giorni lavorativi XE "giorni lavorativi" ed un outlier additivo XE "outlier additivo" in agosto 1994; i primi tre approcci portano ai risultati delle tabelle 5.3 e 5.4. Quando si identifica di nuovo il modello con i dati fino a dicembre 1998, si ottiene un modello airline con gli stessi effetti deterministici. Si può notare come le revisioni siano notevoli per tutti gli approcci anche se la differenza fondamentale risiede nel fatto che solo l'approccio 3) permette un graduale avvicinamento alla serie finale; ad esempio, il dato di gennaio passa da una stima XE "stima" iniziale di 104,5 a quella di 106,9, abbastanza vicina al dato finale di 107,2. La revisione dei dati a dicembre 1998 subirebbe il cambiamento illustrato in figura 5.2. Il quarto approccio è illustrato in tabella 5.5. In questo caso, il modello originale ARIMA(3,1,1)(0,1,1) è sostituito da un ARIMA(3,1,0)(0,1,1) nei mesi di febbraio e marzo, ritorna ad un ARIMA(3,1,1)(0,1,1) ad aprile, per poi passare all'airline da maggio in poi. Dalla tabella si può notare come l'evoluzione dei dati destagionalizzati non sia molto diversa da quella ottenuta con l'approccio 3), tenendo fisso il modello. Dunque, in questo secondo esempio l'approccio 3) risulta preferibile (pur non identificando sempre il modello che trova automaticamente TRAMO XE "TRAMO") per evitare analisi che possono generare continui cambiamenti di modello.

Per ulteriori consigli sulle revisioni, si può consultare Feldmann e Mazzi (1998). Per riferimenti maggiormente tecnici si veda invece Pierce XE "Pierce" (1980), Dagum (1982), Maravall (1986), mentre alcuni studi empirici sono illustrati in Dossé e Planas (1996) e Barbieri e Savio (2000).

Tabella 5.1 – Risultati dei primi due approcci alle revisioni per l'indice della produzione industriale

	APPROCCIO 1	APPROCCIO 2	SERIE FINALE
gen-98	104.0	104.5	104.4
feb-98	104.1	104.5	104.5
mar-98	104.6	104.8	104.7
apr-98	105.3	105.3	105.0
mag-98	104.2	104.6	104.3
giu-98	105.0	105.1	104.7

apr-98	105.3	105.3	105.0
mag-98	104.2	104.6	104.3
giu-98	105.0	105.1	104.7
lug-98	105.7	105.6	105.0
ago-98	101.7	102.7	102.5
set-98	105.7	105.4	104.9
ott-98	104.0	104.1	103.7
nov-98	105.0	104.8	104.2
dic-98	100.0	101.1	101.0

Tabella 5.2 – Dinamica delle revisioni con l'approccio 3 per l'indice della produzione industriale

	GEN-98	FEB-98	MAR-98	APR-98	MAG-98	GIU-98	LUG-98	AGO-98	SET-98	OTT-98	NOV-98	DIC-98	SERIE FINALE
gen-98	104.5	104.3	104.3	104.3	104.3	104.3	104.3	104.2	104.2	104.2	104.2	104.5	104.4
feb-98		104.5	104.5	104.5	104.4	104.4	104.4	104.3	104.3	104.3	104.3	104.5	104.5
mar-98			104.8	104.8	104.7	104.7	104.7	104.6	104.6	104.6	104.6	104.8	104.7
apr-98				105.3	105.2	105.2	105.2	105.0	105.1	105.0	105.0	105.1	105.0
mag-98					104.5	104.5	104.6	104.3	104.4	104.4	104.4	104.4	104.3
giu-98						105.1	105.1	104.8	104.9	104.8	104.8	104.8	104.7
lug-98							105.6	105.2	105.4	105.3	105.3	105.1	105.0
ago-98								102.7	102.9	102.8	102.8	102.6	102.5
set-98									105.4	105.2	105.3	105.0	104.9
ott-98										104.1	104.2	103.8	103.7
nov-98											104.8	104.3	104.2
dic-98												101.1	101.0

Tabella 5.3 – Risultati dei primo due approcci alle revisioni per gli ordinativi totali

	APPROCCIO 1	APPROCCIO 2	SERIE FINALE
gen-98	105.6	105.6	107.2
feb-98	107.3	107.3	107.2
mar-98	110.9	110.8	110.3
apr-98	103.2	103.5	104.0
mag-98	107.8	107.6	106.9
giu-98	103.6	103.8	104.0
lug-98	104.9	105.1	104.2
ago-98	111.7	111.2	108.5
set-98	103.4	103.8	103.5
ott-98	103.8	103.8	102.9
nov-98	101.4	101.4	101.3
dic-98	93.8	94.6	96.8

Tabella 5.4 – Dinamica delle revisioni con l'approccio 3 per gli ordinativi

	GEN-98	FEB-98	MAR-98	APR-98	MAG-98	GIU-98	LUG-98	AGO-98	SET-98	OTT-98	NOV-98	DIC-98	SERIE FINALE
gen-98	104.5	105.2	106.3	105.5	105.6	105.6	105.6	105.7	106.0	106.1	106.3	106.9	107.2
feb-98		106.5	107.5	107.1	107.1	107.2	107.1	107.0	106.9	107.0	107.1	107.3	107.2
mar-98			110.9	108.6	108.6	110.4	110.2	110.2	110.2	110.3	110.3	110.6	110.3
apr-98				105.7	105.9	103.4	103.6	103.7	103.8	103.7	103.8	104.0	104.0
mag-					107.0	107.4	107.5	107.4	107.2	107.1	107.1	106.9	106.9

mar-98	110.9	108.6	108.6	110.4	110.2	110.2	110.2	110.3	110.3	110.6	110.3
apr-98		105.7	105.9	103.4	103.6	103.7	103.8	103.7	103.8	104.0	104.0
mag-98			107.0	107.4	107.5	107.4	107.2	107.1	107.1	106.9	106.9
giu-98				103.8	103.9	104.2	104.1	104.1	104.0	104.1	104.0
lug-98					105.1	105.3	105.0	104.9	104.7	104.3	104.2
ago-98						110.9	110.2	110.1	109.6	108.6	108.5
set-98							104.2	104.0	103.7	103.5	103.5
ott-98								103.8	103.5	102.9	102.9
nov-98									101.6	101.1	101.3
dic-98										96.5	96.8

Tabella 5.5 – Dinamica delle revisioni con l'approccio 4 per gli ordinativi totali

	GEN-98	FEB-98	MAR-98	APR-98	MAG-98	GIU-98	LUG-98	AGO-98	SET-98	OTT-98	NOV-98	DIC-98	SERIE FINALE
gen-98	104.5	105.9	106.0	105.5	105.8	105.6	105.5	105.8	106.1	106.1	106.2	107.2	107.2
feb-98		107.2	107.4	107.1	107.2	107.1	107.1	107.0	106.9	107.0	107.0	107.2	107.2
mar-98			110.7	108.6	110.4	110.3	110.3	110.3	110.2	110.2	110.2	110.3	110.3
apr-98				105.7	103.8	103.6	103.6	103.8	103.8	103.7	103.7	104.0	104.0
mag-98					107.8	107.6	107.6	107.5	107.2	107.2	107.1	106.9	106.9
giu-98						104.0	104.0	104.2	104.1	104.0	104.0	104.0	104.0
lug-98							105.1	105.3	104.9	104.8	104.7	104.2	104.2
ago-98								110.9	110.0	110.0	109.8	108.5	108.5
set-98									104.2	104.0	103.8	103.5	103.5
ott-98										103.9	103.7	102.9	102.9
nov-98											101.8	101.3	101.3
dic-98												96.8	96.8

Figura 5.1 - Confronto tra revisioni (linea tratteggiata) e serie finale (linea continua) per l'indice della produzione industriale

Figura 5.2 - Confronto tra revisioni (linea tratteggiata) e serie finale (linea continua) per gli ordinativi totali

5.2 L'analisi di stabilità dei coefficienti XE "^{stabilità dei coefficienti}"

La scelta di un modello non è scindibile dalla garanzia di una certa stabilità dello stesso nel tempo; è l'instabilità del modello, infatti, che causa il maggior impatto sulle revisioni.

Il termine stabilità può, in realtà, fare riferimento a diversi aspetti: da un lato può riguardare il problema dell'identificazione del medesimo modello nel tempo, ossia la determinazione degli stessi effetti deterministici e dello stesso ordine dell'ARIMA XE "^{ARIMA}". Dall'altro può riferirsi alle stime dei coefficienti incogniti che, di norma, non dovrebbero variare eccessivamente, soprattutto nel breve periodo.

Un modo efficace per verificare la stabilità del modello è quello di effettuare simulazioni XE "^{simulazioni}" storiche rispetto alla variabile analizzata; in pratica, ciò significa eliminare le osservazioni più recenti per un intervallo di uno o più anni, a seconda della lunghezza della serie, ed effettuare l'identificazione e la stima XE "^{stima}" del modello ARIMA XE "^{ARIMA}" che la rappresenta. Successivamente, si aggiunge un dato per volta e si effettuano due tipi di analisi:

- a) si procede, via via che si aggiunge un dato, all'identificazione e stima XE "^{stima}" del modello confrontando i risultati;
- b) si mantiene fisso il modello (come avviene nelle revisioni) e, aggiungendo dati, si registrano le stime dei coefficienti e se ne analizza la loro stabilità, ricorrendo anche ad una misura dell'errore di stima XE "^{errore di stima}" XE "^{stima}" con gli intervalli di confidenza.

Per illustrare meglio le due analisi, si possono riconsiderare le serie usate per gli esempi nel paragrafo precedente. Si analizzi dapprima il caso dell'indice generale della produzione industriale; in precedenza era già stato osservato come le revisioni durante il 1998 risultassero abbastanza limitate grazie alla stabilità del modello. In effetti, mantenendo lo stesso input impiegato per il precedente esperimento, si nota dai grafici in figura 5.3 come le stime dei parametri siano molto stabili durante l'anno considerato; soltanto il parametro relativo alle *holidays* subisce un lieve incremento con il subentrare del dato relativo ad agosto 1998 anche se di entità irrilevante (il valore cambia da -2,321 a -2,256). In definitiva, dunque, questo tipo di approccio fornisce un ulteriore supporto per poter valutare in maniera positiva la fase di identificazione precedentemente effettuata.

Si consideri invece il caso dell'indice generale del totale degli ordinativi; come già detto, per la serie fino al 1997 il modello identificato automaticamente da TRAMO XE "^{TRAMO}" è un ARIMA XE "^{ARIMA}" (3,1,1)(0,1,1) con un regressore per i giorni lavorativi XE "^{giorni lavorativi}" ed un altro per un outlier di tipo additivo in agosto 1994, mentre per la serie fino al 1998 si identifica un *airline* con gli stessi effetti deterministici. Questo cambio nel modello comporta notevoli revisioni nella serie destagionalizzata XE "^{serie destagionalizzata}" durante il 1998. Se ogni mese si identifica automaticamente il modello, è immediato notare che l'ARIMA(3,1,1)(0,1,1) viene confermato con il dato di gennaio 1998, in febbraio e marzo viene identificato un ARIMA(3,1,0)(0,1,1), ad aprile si ritorna al modello precedente e da maggio in poi viene selezionato un *airline*. C'è dunque una instabilità nella identificazione, ma anche nella stima XE "^{stima}" dei parametri, come si può notare dai grafici in figura 5.4. Le componenti autoregressive, nei cinque mesi in cui queste sono presenti sono molto instabili con intervalli di confidenza molto ampi (si noti in particolare il parametro AR1); la componente XE "^{componente}" MA XE "^{MA}" non stagionale sembra stabilizzarsi da maggio in poi (ossia da quando viene selezionato il modello *airline*), mentre quella stagionale subisce notevoli cambiamenti con misure di errore anche elevate. Gli unici parametri che denotano una certa stabilità sono quelli relativi agli effetti deterministici. In questo secondo caso, sicuramente sarà preferibile la scelta dell'*airline*, ma durante il periodo in cui il modello verrà tenuto fisso è consigliabile controllare il parametro MA stagionale XE "^{MA stagionale}" ed eventuali nuovi possibili cambiamenti di modello.

Figura 5.3 – Serie Indice generale della produzione industriale: evoluzione dei parametri nel periodo dic. 97 – dic 98

fisso è consigliabile controllare il parametro MA stagionale $XE^{MA \text{ stagionale}}$ ed eventuali nuovi possibili cambiamenti di modello.

Figura 5.3 – Serie Indice generale della produzione industriale: evoluzione dei parametri nel periodo dic. 97 – dic 98

Figura 5.4: Serie Indice ordinativi generale totale: evoluzione dei parametri nel periodo dic 97-dic 98

5.3 Il problema dell'aggregazione

Accade spesso che una data serie storica Y sia il risultato dell'aggregazione di due o più serie storiche. Situazioni di questo tipo possono verificarsi quando la serie Y che, ad esempio, registra i valori di una data variabile relativamente ad una certa regione geografica, scaturisce dall'aggregazione di n serie Y_1, Y_2, \dots, Y_n che misurano la medesima variabile rilevata nelle n sub-regioni dell'area geografica considerata. Ad esempio, la serie mensile dei matrimoni in Italia può essere ottenuta come la somma delle corrispondenti 20 serie regionali; le forze lavoro totali possono essere viste come la somma delle forze lavoro relative alle varie aree geografiche, e così via.

L'aggregazione può avvenire anche tra settori o branche di attività economica oppure tra differenze di genere. Ad esempio, l'indice generale della produzione industriale si può ottenere come media aritmetica ponderata della produzione delle industrie estrattive, delle attività manifatturiere e della produzione e distribuzione di energia elettrica, gas e acqua; gli occupati totali in Italia possono ottenersi come la somma degli occupati totali femmine e degli occupati totali maschi.

Per fissare le idee ed esplicitare quello che in letteratura è definito come il problema dell'aggregazione, si supponga che

$$Y=Y_1+Y_2.$$

Si indichi con D una data procedura di destagionalizzazione (TRAMO-SEATS, X12-RegARIMA, ecc.) e la si applichi alla serie Y ottenendo la serie destagionalizzata $XE^{serie \text{ destagionalizzata}} sa=D(Y)$. Per destagionalizzare la serie aggregata Y si possono seguire due strade: procedere direttamente alla destagionalizzazione della serie composta Y oppure destagionalizzare le serie Y_1 ed Y_2 e poi sommarle. Nel primo caso si parla di *metodo diretto* $XE^{metodo \text{ diretto}}$, nel secondo di *metodo indiretto* $XE^{metodo \text{ indiretto}}$. In generale,

$$D(Y)^1D(Y_1)+D(Y_2),$$

nel senso che l'applicazione della procedura D alla serie composta fornisce un risultato diverso da quello che si ottiene sommando le serie componenti destagionalizzate: in altre parole, il metodo diretto $XE^{metodo \text{ diretto}}$ ed indiretto non conducono allo stesso risultato.

In letteratura non esiste una produzione su questo tema che metta in risalto l'importanza del problema né è stata ancora individuata una soluzione definitiva. Alcuni risultati sono stati raggiunti recentemente da Gomez (2000) e Planas e Campolongo (2000), a cui si rimanda anche per citazioni bibliografiche sull'argomento. In questa sede verranno svolte soltanto alcune considerazioni di carattere prevalentemente pratico.

Si supponga di dover destagionalizzare $n+1$ serie, Y_1, Y_2, \dots, Y_n, Y , tali che

$$Y= Y_1+Y_2+\dots+Y_n.$$

In generale, l'applicazione di una data procedura di destagionalizzazione D a queste $n+1$ serie dà luogo a ciò che in precedenza abbiamo denominato problema della aggregazione, cioè:

$$D(Y)-[D(Y_1)+D(Y_2)+\dots+D(Y_n)]^10 \quad (5.1)$$

La differenza a sinistra della disuguaglianza (5.1) viene indicata in letteratura con il termine *discrepanza*. La scelta fra metodo diretto $XE^{metodo \text{ diretto}}$ ed indiretto dipende dall'entità di questa

La differenza a sinistra della disuguaglianza (5.1) viene indicata in letteratura con il termine *discrepanza*. La scelta fra metodo diretto XE "metodo diretto" ed indiretto dipende dall'entità di questa discrepanza; se essa è trascurabile, è preferibile destagionalizzare la serie aggregata Y ricorrendo al metodo diretto, in quanto esso non dà luogo a problemi di stagionalità spuria che potrebbero, invece, manifestarsi nel caso in cui si adottasse il metodo indiretto XE "metodo indiretto". Infatti, l'applicazione di quest'ultimo metodo fa sì che la serie aggregata risenta di tutti gli errori di stima XE "stima" della stagionalità presente nelle serie componenti. Inoltre, il metodo diretto è coerente con l'impostazione *model-based* adottata in TRAMO-SEATS.

Se la discrepanza è notevole si hanno a disposizione due possibili soluzioni:

- 1) applicare il metodo indiretto XE "metodo indiretto";
- 2) applicare il metodo diretto XE "metodo diretto" e riallocare la discrepanza fra le serie componenti destagionalizzate.

L'applicazione del metodo indiretto XE "metodo indiretto" è senza dubbio la maniera più semplice di risolvere il problema dell'aggregazione, in quanto la serie destagionalizzata XE "serie destagionalizzata" totale è ottenuta come somma delle serie destagionalizzate componenti; bisogna però tenere presente che l'utilizzo del metodo indiretto, come è stato precedentemente evidenziato, può indurre stagionalità spuria nella serie aggregata. La soluzione 2) consiste nell'applicare la procedura di destagionalizzazione a tutte le $n+1$ serie per poi attribuire, seguendo un criterio appropriato, una porzione della discrepanza (5.1) ad ognuna delle serie componenti destagionalizzate.

Per fissare le idee si consideri il caso di una serie trimestrale Y , osservata per n anni, esprimibile come somma di tre serie $Y^{(1)}, Y^{(2)}, Y^{(3)}$, cioè:

$$Y_{ij} = Y_{ij}^{(1)} + Y_{ij}^{(2)} + Y_{ij}^{(3)} \quad i=1,2,3,4 \quad j=1,2,\dots,n.$$

dove l'indice i si riferisce al trimestre e l'indice j all'anno. Si supponga inoltre che ogni serie segua un modello additivo del tipo:

$$Y_{ij}^{(h)} = T_{ij}^{(h)} + S_{ij}^{(h)} + I_{ij}^{(h)} \quad h=1,2,3$$

dove $T_{ij}^{(h)}$ è la componente XE "componente" di ciclo-trend XE "ciclo-trend", $S_{ij}^{(h)}$ è la componente stagionale XE "componente stagionale", $I_{ij}^{(h)}$ è la componente irregolare XE "componente irregolare".

Applicando la procedura di destagionalizzazione D alle quattro serie $Y^{(1)}, Y^{(2)}, Y^{(3)}$ ed Y , otteniamo le serie destagionalizzate $D(Y^{(1)}), D(Y^{(2)}), D(Y^{(3)})$ e $D(Y)$.

Si supponga di aver constatato che:

$$D(Y) - [D(Y^{(1)}) + D(Y^{(2)}) + D(Y^{(3)})] \neq 0$$

ovvero:

$$D(Y_{ij}) - [D(Y_{ij}^{(1)}) + D(Y_{ij}^{(2)}) + D(Y_{ij}^{(3)})] \neq 0 \quad i=1,2,3,4 \quad j=1,2,\dots,n.$$

Sia R_{ij} la discrepanza tra la serie destagionalizzata XE "serie destagionalizzata" e la \hat{a} delle serie componenti destagionalizzate tale che:

$$D(Y_{ij}) - [D(Y_{ij}^{(1)}) + D(Y_{ij}^{(2)}) + D(Y_{ij}^{(3)})] = R_{ij} \quad i=1,2,3,4 \quad j=1,2,\dots,n.$$

L'obiettivo consiste nell'apportare una correzione, $c_{ij}^{(h)}$, ai valori delle serie componenti destagionalizzate, $D(Y_{ij}^{(h)})$, al fine di ottenere l'uguaglianza tra la serie aggregata destagionalizzata e la somma delle componenti destagionalizzate e corrette. Più precisamente, si tratta di determinare un insieme di termini di aggiustamento, $\{c_{ij}^{(1)}, c_{ij}^{(2)}, c_{ij}^{(3)}, i=1,2,3,4 \quad j=1,2,\dots,n\}$, tale che

$$D(Y_{ij}) = [D(Y_{ij}^{(1)}) + c_{ij}^{(1)}] + [D(Y_{ij}^{(2)}) + c_{ij}^{(2)}] + [D(Y_{ij}^{(3)}) + c_{ij}^{(3)}] \quad i=1,2,3,4 \quad j=1,\dots,n \quad (5.2)$$

e:

$$[D(Y_{1j}^{(h)}) + c_{1j}^{(h)}] + [D(Y_{2j}^{(h)}) + c_{2j}^{(h)}] + [D(Y_{3j}^{(h)}) + c_{3j}^{(h)}] + [D(Y_{4j}^{(h)}) + c_{4j}^{(h)}] = Y_j^{(h)} \quad h=1,2,3, \quad j=1,\dots,n. \quad (5.3)$$

in altre parole, si richiede che tali termini di aggiustamento soddisfino l'uguaglianza tra la serie aggregata destagionalizzata e la somma delle serie componenti destagionalizzate corrette e che la somma di queste ultime all'interno di ogni anno sia uguale alla corrispondente serie annuale.

Dalla (5.2) segue che:

$$R_{ij} = c_{ij}^{(1)} + c_{ij}^{(2)} + c_{ij}^{(3)} \quad i=1,2,3,4 \quad j=1,2,\dots,n \quad (5.4)$$

Inoltre, ricordando che per definizione

$$D(Y_{1j}^{(h)}) + D(Y_{2j}^{(h)}) + D(Y_{3j}^{(h)}) + D(Y_{4j}^{(h)}) = Y_j^{(h)} \quad h=1,2,3, \quad j=1,\dots,n,$$

è evidente che dalla (5.3) segue che:

$$c_{1j}^{(h)} + c_{2j}^{(h)} + c_{3j}^{(h)} + c_{4j}^{(h)} = 0 \quad h=1,2,3, \quad j=1,\dots,n \quad (5.5)$$

$$D(Y_{1j}^{(h)}) + D(Y_{2j}^{(h)}) + D(Y_{3j}^{(h)}) + D(Y_{4j}^{(h)}) = Y_j^{(h)} \quad h=1,2,3, j=1,\dots,n,$$

è evidente che dalla (5.3) segue che:

$$c_{1j}^{(h)} + c_{2j}^{(h)} + c_{3j}^{(h)} + c_{4j}^{(h)} = 0 \quad h=1,2,3, j=1,\dots,n \quad (5.5)$$

Una volta esplicitati i vincoli (5.4) e (5.5) possiamo considerare la seguente funzione obiettivo per l'anno j-esimo:

$$a_{1j}^{(1)}[c_{1j}^{(1)}]^2 + a_{2j}^{(1)}[c_{2j}^{(1)}]^2 + a_{3j}^{(1)}[c_{3j}^{(1)}]^2 + a_{4j}^{(1)}[c_{4j}^{(1)}]^2 + \dots + a_{1j}^{(3)}[c_{1j}^{(3)}]^2 + a_{2j}^{(3)}[c_{2j}^{(3)}]^2 + a_{3j}^{(3)}[c_{3j}^{(3)}]^2 + a_{4j}^{(3)}[c_{4j}^{(3)}]^2,$$

dove $\{a_{1j}^{(1)}, a_{2j}^{(1)}, a_{3j}^{(1)}, a_{4j}^{(1)}, \dots, a_{1j}^{(3)}, a_{2j}^{(3)}, a_{3j}^{(3)}, a_{4j}^{(3)}\}$ è un dato sistema di pesi.

A questo punto i termini di aggiustamento, per il j-esimo anno, $\{c_{ij}^{(1)}, c_{ij}^{(2)}, c_{ij}^{(3)}; i=1,2,3,4\}$, si possono ottenere risolvendo il seguente problema di minimo vincolato:

$$\min \{ a_{1j}^{(1)}[c_{1j}^{(1)}]^2 + a_{2j}^{(1)}[c_{2j}^{(1)}]^2 + a_{3j}^{(1)}[c_{3j}^{(1)}]^2 + a_{4j}^{(1)}[c_{4j}^{(1)}]^2 + \dots + a_{1j}^{(3)}[c_{1j}^{(3)}]^2 + a_{2j}^{(3)}[c_{2j}^{(3)}]^2 + a_{3j}^{(3)}[c_{3j}^{(3)}]^2 + a_{4j}^{(3)}[c_{4j}^{(3)}]^2 \}$$

con i vincoli:

$$R_{ij} = c_{ij}^{(1)} + c_{ij}^{(2)} + c_{ij}^{(3)} \quad i=1,2,3,4,$$

$$c_{1j}^{(h)} + c_{2j}^{(h)} + c_{3j}^{(h)} + c_{4j}^{(h)} = 0 \quad h=1,2,3.$$

Il problema a questo punto è la scelta dei pesi. Se indichiamo con $S_{si}^{(h)}$ e $S_{ii}^{(h)}$, rispettivamente, la deviazione standard delle componenti XE "componente stagionale" ed irregolare della serie $Y^{(h)}$, relativa all'i-esimo trimestre, possiamo porre:

$$a_{ij}^{(h)} = 1/[S_{si}^{(h)} S_{ii}^{(h)}], \quad (6)$$

secondo cui i pesi sono inversamente proporzionali al prodotto della deviazione standard della componente XE "componente stagionale" ed irregolare. Una scelta di questo tipo fa sì che i termini di correzione siano positivamente legati alla variabilità delle componenti stagionale ed irregolare. Il termine di correzione assegnato ad una data serie in un dato trimestre sarà tanto più ampio quanto più è elevata la variabilità della componente stagionale ed irregolare di quella serie relativa a quel dato trimestre. Di solito questa scelta viene giustificata sostenendo che la stima XE "stima" della componente stagionale è tanto meno accurata quanto più elevata è la sua variabilità e quella della componente irregolare XE "componente irregolare". Da questo punto di vista sembra quindi appropriato assegnare una quota maggiore della discrepanza R_{ij} a quelle serie che presentano una stagionalità ed una componente irregolare più instabili (si veda den Butter e Fase, 1991).

Chiaramente, le esemplificazioni sono state fatte per semplicità espositiva solo per serie trimestrali, ma l'estensione al caso mensile è immediata.

Come esempio, consideriamo la serie mensile delle presenze turistiche in Italia presso esercizi complementari da gennaio 1987 a dicembre 1998, data dalla somma della serie relativa agli italiani e quella relativa agli stranieri (file *esempio 7*). Analizziamo le tre serie immettendo i seguenti comandi di input:

```
$INPUT LAM=0 INIC=3 IDIF=3 SEATS XE "SEATS" =2 BIAS=-1$
```

Il comando $BIAS=-1$ è introdotto per permettere che la (5.3) sia verificata (cfr. la descrizione dei parametri di TRAMO-SEATS nel capitolo 4).

In figura 5.5 sono rappresentate le serie destagionalizzate dei totali col metodo diretto XE "metodo diretto" ed indiretto; le discrepanze sono in alcuni casi notevoli, dunque la scelta dell'uno o dell'altro comporta risultati diversi.

L'applicazione del metodo di den Butter-Fase modifica invece le serie destagionalizzate disaggregate, dunque quella relativa agli italiani e quella relativa agli stranieri. Come si può notare dalle figure 5.6 e 5.7, mentre la serie relativa agli italiani non subisce grosse modifiche, le differenze sono notevoli per gli stranieri. Questo è dovuto al fatto che la varianza XE "varianza" della seconda serie è notevolmente superiore a quella della prima e perciò ad essa vengono assegnate la maggior parte delle correzioni. Con entrambi i metodi, dunque, almeno una serie viene notevolmente corretta. Un ulteriore limite del metodo di den Butter-Fase è la possibilità di apportare correzioni ai soli dati relativi ad anni completi.

In conclusione, la discrepanza tra la serie aggregata destagionalizzata e la somma delle serie componenti destagionalizzate è senza dubbio un problema che può creare un qualche imbarazzo nell'utilizzatore.

Si tenga presente che quelle appena descritte sono solo alcune delle strategie che possono essere seguite al fine di eliminare o ridurre tale inconsistenza. Ciò nonostante occorre sottolineare che nessuna di queste possiede un fondamento teorico. Avendo un carattere empirico, vanno, quindi, utilizzate con le dovute cautele.

Figura 5.5 – Serie delle presenze turistiche totali in esercizi complementari: metodo diretto XE "metodo diretto" (linea tratteggiata) ed indiretto (linea continua)

Figura 5.5 – Serie delle presenze turistiche totali in esercizi complementari: metodo diretto XE "metodo diretto" (linea tratteggiata) ed indiretto (linea continua)

Figura 5.6 – Serie delle presenze turistiche italiane in esercizi complementari metodo diretto XE "metodo diretto" (linea tratteggiata) e correzioni di den Butter-Fase (linea continua)

Figura 5.7 – Serie delle presenze turistiche straniere in esercizi complementari metodo diretto XE "metodo diretto" (linea tratteggiata) e correzioni di den Butter-Fase (linea continua)

5.4 Casi di studio

In questo paragrafo vengono affrontati alcuni problemi pratici che sorgono frequentemente nell'uso del programma TRAMO-SEATS. In particolare, saranno esaminati quattro problemi:

- a) ampiezza dell'intervallo temporale da analizzare;
- b) opportunità di usare la trasformazione logaritmica o meno (a prescindere dal risultato del test, effettuato imponendo $LAM=-1$);
- c) superamento di eventuali problemi nella diagnostica;
- d) cautele da assumere con alcuni tipi di valori anomali.

Naturalmente, le soluzioni prospettate sono basate su considerazioni di natura empirica; in ogni caso, è opportuno usare inizialmente tutti i dati a disposizione e le procedure automatiche di TRAMO-SEATS per poi procedere ad eventuali correzioni.

5.4.1 Scelta dell'intervallo temporale

Il problema della scelta del periodo temporale su cui basare l'analisi deriva dal fatto che una serie di dati poco omogenea può portare ad errate specificazioni del modello ed all'identificazione di un elevato numero di outlier e quindi può produrre instabilità nei coefficienti e nella serie destagionalizzata XE "serie destagionalizzata". Per altro verso, esiste la necessità di utilizzare un adeguato numero di dati per consentire una più efficiente stima XE "stima" dei coefficienti.

La non omogeneità dei dati può essere causata da eventuali cambiamenti di base e concatenamenti (se si tratta di numeri indici), cambiamenti nella rilevazione, ricostruzioni dei dati o anomalie simili. In termini statistici, questo problema si manifesta in *break* strutturali XE "break strutturali" e varianza XE "varianza" non costante nel tempo (eteroschedasticità XE "eteroschedasticità"); a sua volta, la presenza di *break* strutturali e di eteroschedasticità si risolve, generalmente, in non linearità XE "non linearità" della serie. La non omogeneità può quindi essere segnalata da test sui residui al quadrato significativi (cfr. paragrafo 2.3.2) che persistono anche con interventi *ad hoc* dell'utente; inoltre, nella ricerca di un adattamento forzato della serie al modello lineare, il programma potrebbe riuscire a superare i problemi di diagnostica con l'individuazione di un gran numero di outlier (si veda Planas, 1998).

Se la serie è sufficientemente lunga, è possibile "tagliare" una prima parte di essa ritenuta non omogenea rispetto al resto della serie; ciò può essere fatto con l'ausilio dell'analisi grafica, che mette in rilievo visivamente le disomogeneità in questione, o tramite tentativi, ossia eliminando un anno per volta finché la diagnostica non diventa accettabile. E' chiaro che i problemi di diagnostica potrebbero essere dovuti ad altri motivi, perciò è preferibile effettuare i tagli se c'è un riscontro grafico delle disomogeneità, possibilmente giustificato da conoscenze a priori.

Ad esempio, consideriamo la serie dell'indice degli ordinativi nazionali (in base 1995=100) da gennaio 1981 ad agosto 1999 (file *esempio 8*); in questo periodo sono avvenuti due cambi base (nel 1985 e nel 1990). Il programma TRAMO XE "TRAMO" opera automaticamente la trasformazione logaritmica e le differenze prima e dodicesima; in figura 5.8 è rappresentata la serie dopo le trasformazioni. Già dal semplice andamento grafico possiamo notare una certa eteroschedasticità XE "eteroschedasticità", che sembra attenuarsi solo dal 1990 in poi. Un opportuno taglio dovrebbe perciò riguardare i primi nove anni. Questa scelta è comunque confortata dall'analisi della serie, riducendo iterativamente di un anno il periodo in questione. In tabella 5.6 sono riportati i risultati, analizzati per ogni anno con l'opzione $RSA=4$; possiamo notare come la non linearità XE "non linearità" della serie venga corretta nell'intero periodo da numerosi outlier. Le difficoltà di adattamento sono inoltre evidenziate dalla presenza di alcuni residui standardizzati elevati. Quando il numero di outlier diminuisce, si riscontra qualche problema in termini di diagnostica (generalmente un rifiuto dell'ipotesi di normalità). Il periodo gennaio 86-agosto 99 sembra essere un adeguato periodo di analisi, anche se la instabilità del modello è evidenziata dal fatto che un'ulteriore riduzione di un anno porta al cambiamento degli ordini dell'ARIMA XE "ARIMA" (si veda nota in tabella 5.6). Il periodo gennaio 90-agosto 99 è l'unico che non presenta outlier né problemi di diagnostica; dunque, l'intuizione iniziale appare giustificata.

Figura 5.8 - Ordinativi nazionali (logaritmi e differenze prime e dodicesime)

Tabella 5.6 - serie Ordinativi Nazionali: destagionalizzazione per vari periodi di analisi*

PERIODO	OUTLIER AO XE "AO"	OUTLIER TC XE "TC"	OUTLIER LS XE "LS"	RESIDUI STANDARDIZZA TI	RIFIUTO TRAMO XE "TRAMO"	RIFIUTO SEATS XE "SEATS"
gen 81-ago 99	12 81	12 91	11 94	6 94		
	8 83			4 97		
	8 85					
	8 87					
	12 88					
	8 89					
	12 89					
	8 93					
gen 82-ago 99	8 87			8 89	Normalità XE "Normalità"	Normalità XE "Normalità"
	12 88					
	12 89					
	8 93					
gen 83-ago 99	8 83					Ljung-Box XE "Ljung-Box"
	8 85					
	8 87					
	12 88					
	12 89					
gen 84-ago 99	3 84	12 91	2 84	6 94		
	8 85		11 94			
	8 87		4 97			
	8 88					
	12 88					
	12 89					
	8 93					
gen 85-ago 99	8 85				Normalità XE "Normalità"	
	8 87					
	8 88					
	12 88					
	12 89					
gen 86-ago 99	8 87					

	12 88		
	12 89		
gen 86-ago 99	8 87		
	12 88		
	12 89		
gen 87-ago 99	8 87		
	12 88		
	12 89		
gen 88-ago 99	8 88	12 92	Normalità XE "Normalità"
	12 90		Normalità XE "Normalità"
gen 89-ago 99	12 89		L-B quadrati
gen 90-ago 99			

* Per tutti i casi è stato identificato un modello airline, tranne che per gen 87-ago 99 per cui è stato usato un ARIMA XE "ARIMA" (1,1,0)(0,1,1)

5.4.2 La scelta della trasformazione logaritmica

Come già detto nel secondo capitolo di questa guida, il programma TRAMO XE "TRAMO" contiene un parametro per verificare la validità dell'operazione di trasformazione logaritmica (basta porre LAM=-1). Questo test favorisce però la trasformazione logaritmica in maniera eccessiva, ossia può indurre a sottoporre a trasformazione serie palesemente additive (quest'ultima considerazione può derivare dall'evoluzione grafica della serie, dalla quale si può notare una sostanziale stazionarietà in media e varianza XE "varianza", che implica un modello additivo). L'esperienza porta a sostenere che spesso il criterio BIC XE "BIC" conduce a risultati più plausibili rispetto al criterio adottato in TRAMO. Si ritiene perciò opportuno, prima di destagionalizzare, calcolare il BIC sia per il modello additivo che per quello moltiplicativo e scegliere quello con valore minore. Purtroppo, l'attuale versione di TRAMO non consente questo confronto in quanto sia il BIC che l'AIC XE "AIC" sono calcolati sulle serie trasformate, dunque è necessario ricorrere ad un altro programma (ad esempio X12).

Come esempio si osservi in figura 5.9 il grafico dell'indice generale della produzione industriale; come si può notare, l'andamento non è di tipo esponenziale per cui non c'è evidenza grafica per un modello moltiplicativo. Tuttavia, analizzando la serie con i seguenti parametri:

```
$INPUT LAM=-1 IATIP=1 INIC=3 IDIF=3 IEAST=1 ITRAD=2 IREG=1 SEATS XE "SEATS" =2$
$REG IUSER=-1 NSER=1 ILONG=247 REGEFF=2 $
hol_81
```

il programma TRAMO XE "TRAMO" individua un modello airline moltiplicativo, le cui valutazioni diagnostiche (tabella 5.7) mettono in rilievo una forte non linearità XE "non linearità", nonostante l'individuazione di due outlier, ed evidenziano il rifiuto dell'ipotesi di normalità. Con la verifica, tramite X12, del valore del BIC XE "BIC" sia per il modello additivo che moltiplicativo, viene scelto il modello additivo. Imponendo quest'ultimo nell'input di TRAMO tramite:

```
$INPUT LAM=1 IATIP=1 INIC=3 IDIF=3 IEAST=1 ITRAD=2 IREG=1 SEATS XE "SEATS" =2$
$REG IUSER=-1 NSER=1 ILONG=247 REGEFF=2 $
hol_81
```

si notano miglioramenti anche in termini di diagnostica (si veda tabella 5.7).

Figura 5.9: Indice generale della produzione industriale gen 1981-lug 1999

Tabella 5.7 – Indice generale della produzione industriale: diagnostica destagionalizzazione con modello additivo e moltiplicativo (in neretto i test rifiutati al 95% di significatività)

	MOD. MOLTIPLICATIVO	MOD. ADDITIVO
BIC XE " ^{BIC} (X12)	1188,73	1148,57
Outlier XE " ^{Outliers} XE " ^{Outlier}	AO: XE " ^{AO} (8 84) (8 95)	nessuno
Norm	14,73	1,14
L-B	18,68	23,49
L-B quad	59,25	24,14
Norm SEATS XE " ^{SEATS}	12,06	0,57
L-B SEATS XE " ^{SEATS}	25,88	25,11
Pierce XE " ^{Pierce} SEATS XE " ^{SEATS}	7,14	1,92
L-B quad SEATS XE " ^{SEATS}	65,93	24,22
Pierce XE " ^{Pierce} quad SEATS XE " ^{SEATS}	39,91	5,59

5.4.3 Problemi nella diagnostica

Spesso la procedura automatica non individua modelli ottimali, nel senso che la diagnostica mette in rilievo solo problemi relativi ad alcune ipotesi; in effetti, in alcuni casi è possibile che il modello scelto automaticamente non sia il migliore. Infatti, la scelta automatica del modello avviene tramite il confronto tra i BIC XE " ^{BIC} , da cui la scelta del migliore rispetto a questo criterio. Può darsi però che esista qualche modello, più rappresentativo di quello selezionato automaticamente da TRAMO XE " ^{TRAMO} , che riesca a "catturare" eventuali strutture ancora presenti nei residui. Gli accorgimenti che di seguito verranno proposti non hanno la valenza di regole, ma sono frequentemente utilizzati nell'analisi di serie storiche. Resta salvo il fatto che non sempre accade che si sia in grado di ottenere un modello che rispetti tutte le ipotesi di partenza; pertanto bisognerà accontentarsi di soluzioni meno rigorose.

In generale, possiamo dire che un valore del test di normalità superiore alla soglia critica può essere indice della presenza di outlier non identificati automaticamente dalla procedura. In questi casi è opportuno da una parte osservare il grafico dei residui per verificare che non ci siano valori abbondantemente fuori dalle soglie di significatività; dall'altra accertarsi che non vi siano residui standardizzati molto elevati. La procedura SEATS XE " ^{SEATS} segnala nella prima parte dell'output, subito dopo la serie dei residui, gli eventuali residui standardizzati elevati con la posizione dell'osservazione a cui si riferisce e la data di riferimento. In questi casi, si può provare ad imporre nel programma un outlier di tipo additivo (con l'opzione IUSER=2 del comando REG) in corrispondenza di quelle date. Si riconsideri allora l'esempio illustrato in tabella 5.6 relativo alla serie degli indici degli ordinativi nazionali e si supponga di voler analizzare i dati da gennaio 1988 ad agosto 1999. Come si può notare dalla tabella, TRAMO XE " ^{TRAMO} individua automaticamente 2 outlier additivi (agosto '88 e dicembre '90), ma le ipotesi di normalità sono rifiutate sia da TRAMO che da SEATS (i valori dei test sono rispettivamente 6,496 e 13,14). Dalla tabella si nota anche che esiste un residuo standardizzato elevato in corrispondenza di dicembre '92 (osservazione numero 60). Utilizzando il seguente input:

```
$INPUT RSA=4 IREG=1$
$REG IUSER=2 NSER=1$
60 AO XE " AO
```

si ottiene lo stesso modello precedente, con gli stessi effetti deterministici (più l'outlier imposto) e con tutte le ipotesi accettate (in particolare, le statistiche di normalità sono pari a 2,486 e 4,360).

Se i problemi riguardano la diagnostica relativa ai test di Ljung-Box XE " ^{Ljung-Box} , è bene controllare direttamente le autocorrelazioni XE " ^{autocorrelazioni} per verificare se persiste qualche struttura nei residui.

Ad esempio, riprendendo ancora i risultati di tabella 5.6, consideriamo la serie dell'indice per gli ordinativi nazionali da gennaio 1983 ad agosto 1999. In questo caso, il test di Ljung-Box XE " ^{Ljung-Box} in SEATS XE " ^{SEATS} rifiuta l'ipotesi di incorrelazione, essendo pari a 35,88. In effetti il valore è abbastanza vicino alla zona di accettazione, dunque è possibile che un intervento sul modello sia superfluo; dal grafico (figura 5.10) delle ACF e PACF infatti risulta che solo due coefficienti (segnalati in nero) sono significativamente diversi da zero. In ogni caso si può provare a migliorare il modello imponendo un ARIMA XE " ^{ARIMA} (1,1,1)(0,1,1) con gli stessi effetti deterministici del precedente modello automatico; le istruzioni saranno perciò le seguenti:

significativamente diversi da zero. In ogni caso si può provare a migliorare il modello imponendo un ARIMA XE "ARIMA" (1,1,1)(0,1,1) con gli stessi effetti deterministici del precedente modello automatico; le istruzioni saranno perciò le seguenti:

```
$INPUT IMEAN=0 ITRAD=1 IREG=5 P=1 SEATS XE "SEATS" =2$
$REG IUSER=2 NSER=5$
8 AO XE "AO" 56 AO 84 AO 72 AO 32 AO
```

Il nuovo modello presenta un *t-value* significativo per il coefficiente autoregressivo ed un deciso miglioramento nel test di Ljung-Box XE "Ljung-Box" in SEATS XE "SEATS", che ora è pari a 26,46.

Figura 5.10 – Indice del fatturato nazionale (gen.83 – ago.99): funzioni di autocorrelazione globale e parziale per i residui del modello automatico

5.4.4 Dati anomali problematici

Particolare attenzione deve essere posta al caso in cui il modello identifichi outlier di tipo *level shift* in corrispondenza degli ultimi anni della serie; un cambiamento di livello XE "cambiamento di livello", infatti, per essere consolidato in una serie storica, ha bisogno di un certo numero di anni a seguire prima che possa essere considerato come tale. Può darsi infatti che, in seguito, il modello lo consideri come outlier di altro tipo o addirittura che lo assorba nella serie come dato "normale". Mantenendo il cambiamento di livello, l'informazione che ne scaturirebbe in termini congiunturali potrebbe essere dunque distorta. E' opportuno perciò che in questi casi l'utente provi anche ad usare le opzioni *IATIP=1* e *AIO=1* per effettuare una ricerca di outlier che escluda quelli di tipo LS XE "LS".

Come esempio, si consideri la serie dell'occupazione nelle grandi imprese al lordo dei cassa-integrati nel complesso dei servizi da gennaio 1993 a maggio 1999 (file *esempio 9*). Procedendo in automatico (RSA=4), il programma individua un modello airline con due outlier di tipo LS XE "LS" in dicembre 93 e gennaio 95. Dato lo scarso numero di dati a disposizione, un modello di questo tipo può essere fuorviante; procedendo come sopra si può fornire il seguente input:

```
$INPUT LAM=-1 INIC=3 IDIF=3 IATIP=1 AIO=1 IEAST=-1 ITRAD=-2 SEATS XE "SEATS" =2$
```

I due precedenti outlier LS XE "LS" vengono trasformati in TC XE "TC" ed il modello migliora anche dal punto di vista della diagnostica, convalidando l'ipotesi di normalità che prima veniva rifiutata.

Si consideri invece il caso della serie delle esportazioni mensili verso i Paesi UE dal gennaio 1988 a luglio 1999 (file *esempio 10*). Procedendo in automatico (RSA=4) si identifica un modello moltiplicativo ARIMA XE "ARIMA" (0,1,2)(0,1,0) con effetti deterministici costituiti da un regressore per i giorni lavorativi XE "giorni lavorativi", due outlier di tipo additivo (dicembre 1992 ed agosto 1993) ed un outlier di tipo LS XE "LS" (in marzo 1996). Aggiungendo un solo dato alla serie (agosto 1999) e procedendo sempre in automatico, il programma identifica un modello moltiplicativo ARIMA (1,1,0)(0,1,1) con effetti deterministici costituiti da un regressore per i giorni lavorativi ed un outlier di tipo additivo (dicembre 92).

Il risultato testimonia dunque la possibilità di notevoli cambiamenti ottenuti aggiungendo nuovi dati alla serie, che devono cautelare l'utente nell'uso degli outlier, in particolare di quelli tipo LS XE "LS".