

Componenti deterministiche e stocastiche nella destagionalizzazione delle serie storiche basata su modelli: una applicazione ed alcuni confronti.

Cosimo Vitale

Agostino Tarsitano(*)
Università degli studi della Calabria
Dipartimento di Economia e Statistica
87030 Arcavacata di Rende (Cs)
agotar@unical.it

Riassunto.

Il problema affrontato riguarda la decomposizione, ed in particolare la destagionalizzazione o depurazione stagionale di una serie storica. In questo lavoro non verrà analizzato il problema dell'esistenza o meno di componenti deterministiche, ma come e quando i processi rappresentabili con modelli della classe ARIMA invertibile possano contenere componenti deterministiche e come sia possibile, utilizzando la modellistica ARIMA, enucleare componenti, siano esse deterministiche o stocastiche, a cui attribuire i classici significati di ciclo-trend, stagionalità e residuo.

Keywords: Destagionalizzazione, decomposizione di una serie storica, modelli ARIMA

() Lavoro comparso nei quaderni di Statistica ed Econometria, vol VI, Aprile 1984, pp. 29-72. C. Vitale ha scritto i paragrafi 1-5 A. Tarsitano i paragrafi 6-8 ed il programma di calcolo per la determinazione dei filtri e delle componenti.*

1. Introduzione

Il problema che verrà affrontato nel seguito riguarda la decomposizione, ed in particolare la destagionalizzazione o depurazione stagionale di una serie storica Y_t , $t=1, \dots, N$. A tale proposito, supponiamo che Y_t sia il processo stocastico continuo a parametro discreto su cui è stata osservata la suddetta serie storica e che sia:

$$(1.1) \quad \begin{cases} Y = F(t, \Lambda) \\ F(t, \Lambda) = P_{1t} + S_{1t} \\ Z_t = P_{2t} + S_{2t} + e_t \end{cases}$$

ove P_{1t} ed S_{1t} sono rispettivamente la parte deterministica del ciclo-trend e della stagionalità, P_{2t} ed S_{2t} sono rispettivamente la parte stocastica del ciclo-trend e della stagionalità, mentre e_t è il residuo casuale non prevedibile e quindi possa considerarsi un *white noise*: $E(e_t) = 0$, $var(e_t) = \sigma_t^2$ e $cov(e_t, e_{t-k}) = 0$, per $k \neq 0$. In altri termini, $F(t, \Lambda)$ rappresenta la parte deterministica di Y_t , mentre Z_t è la parte stocastica.

I problemi che analizzeremo sono:

- (i) stabilire quando in Y_t , date determinate condizioni, è possibile individuare una componente deterministica del tipo $F(t, \Lambda)$;
- (ii) stabilire una o più metodologie che, a partire da date condizioni imposte su Y_t , permettano di derivare alcune o tutte le componenti presenti nella (1.1) ed in particolare $D_t = P_{1t} + P_{2t} + e_t$ ed $S_t = S_{1t} + S_{2t}$ cioè la componente destagionalizzata e quella stagionale, rispettivamente.

Nei lavori di Abraham-Box (1978) ed in quello di Vitale (1982a) è stato affrontato essenzialmente il punto (i) sotto l'ipotesi che Y_t sia un processo rappresentabile con un modello della classe ARIMA, mentre nei lavori di Pierce (1978) ed Havenner-Swamy (1981) l'attenzione è stata posta essenzialmente sul punto (ii).

In questo lavoro non verrà analizzato il problema dell'esistenza o meno di componenti deterministiche, ma come e quando i processi rappresentabili con modelli della classe ARIMA invertibile possano contenere componenti deterministiche e come sia possibile, utilizzando la modellistica ARIMA, enucleare componenti, siano esse deterministiche o stocastiche, a cui attribuire i classici significati di ciclo-trend, stagionalità e residuo. Più in particolare, nel paragrafo 2 illustreremo l'approccio proposto da Pierce (1978) commentandone le relative deficienze. Nel paragrafo 3 presenteremo la metodologia proposta da Havenner-Swamy

(1981), ne mostreremo i limiti e dimostreremo come il modello da loro proposto sia riconducibile, assieme alle componenti generate, a complicati modelli della classe ARIMA in

cui i parametri MA sono funzione del tempo. Nel paragrafo 4, seguendo il lavoro di Vitale (1982a), mostreremo come i processi rappresentabili con modelli ARIMA invertibili possano contenere, in casi abbastanza rilevanti, componenti deterministiche. Nel paragrafo 5 mostreremo come e quando un processo rappresentabile con un $ARIMA(0, 1, r + 1)(0, 1, 0)_r$ sia decomponibile nelle componenti correlate di stagionalità e destagionale, rappresentabili con particolari ARIMA. Nel paragrafo 6 adatteremo alla serie analizzata in Havenner-Swamy (1981) un modello ARIMA e lo decomporremo con la metodologia proposta nel paragrafo 5. Nel paragrafo 7 confronteremo i risultati ottenuti sull'applicazione con la procedura proposta da Pierce, da Havenner-Swamy, con il metodo X-11 e con quello da noi proposto. Le conclusioni, le prospettive ed i commenti saranno contenuti nel paragrafo 8.

2. L'approccio di Pierce

Nel lavoro di Pierce (1978) si suppone che sia

$$Y_t = D_t + S_t$$

con

$$D_t = P_t + e_t$$

e si osserva, giustamente, che il punto cruciale della problematica della depurazione stagionale è quello di stimare S_t . Infatti, se \hat{S}_t è una stima di S_t si otterrà $\hat{D}_t = Y_t - \hat{S}_t$. La stima di S_t , in generale, viene ottenuta con uno dei due seguenti metodi alternativi:

(a) Metodo che utilizza le medie mobili

È quello, in generale, più utilizzato nelle procedure standard di destagionalizzazione come l'X-11 e consiste nell'elaborare un filtro simmetrico del tipo media mobile come il seguente:

$$(2.1) \quad W_S(F, B) = w_{s0} + \sum_{j=1}^k w_{sj} (F^j + B^j)$$

ove, al solito, B è un operatore lineare per cui $B^j X_t = X_{t-j}$ e $F = B^{-1}$, tale che applicando il filtro (2.1) ad Y si abbia una stima di S_t :

$$\hat{S}_t = W_S(F, B)Y_t$$

e quindi, per ottenere una stima di D_t basta considerare:

$$\hat{D}_t = Y_t - \hat{S}_t = [1 - W_S(F, B)]Y_t = W_D(F, B)Y_t$$

Si può dimostrare (vedi per esempio Pierce (1979)) che se $W_S(F, B)$ è elaborato tenendo conto della struttura ARIMA di Y_t allora $W_S(F, B)$ è ottimale nel senso che minimizza gli errori quadratici medi della stima di S_t , cioè minimizza

$$E[(S_t - \hat{S}_t)^2]$$

ed il filtro (2.1) è derivabile se il modello ARIMA per Y_t è invertibile (vedi per esempio Vitale (1982b)). D'altra parte, osserva Pierce (1978), pag. 224, se Y_t contiene una componente non stazionaria e deterministica il modello per Y_t è non invertibile per cui questa procedura è valida se e solo se in Y_t vi è non stazionarietà stocastica. Nel paragrafo 5 mostreremo che questa affermazione non è sempre vera.

(b) Metodi che utilizzano modelli regressivi

Se Y_t contiene componenti non stazionarie esclusivamente di tipo deterministiche e quindi si ha:

$$P_t = \sum_{j=1}^h \alpha_j c_{jt}; \quad S_t = \sum_{j=1}^m \beta_j a_{jt}$$

che implicano per Y_t il seguente modello di regressione $Y = Ca + Db + e$ le componenti possono essere derivate con semplici metodi regressivi.

In conclusione, Pierce afferma che se in Y_t vi sono componenti solo stocastiche allora è necessario utilizzare le procedure descritte in (a), se vi sono componenti solo deterministiche bisogna ricorrere alle procedure (b). Il problema diviene più complesso se Y_t contiene sia componenti deterministiche che stocastiche. In tal caso, sostiene sempre Pierce, è necessario elaborare una procedura diversa da quelle attualmente in uso che in primo luogo verifichi la presenza di tali componenti.

La procedura proposta da Pierce è a stadi e consiste:

(i) nella stima preliminare di P_{1t} ed S_{1t} con metodi regressivi ottenendo \hat{P}_{1t} ed \hat{S}_{1t} e quindi

$$\hat{Z}_t = T_t - \hat{P}_{1t} - \hat{S}_{1t}$$

(ii) nel filtraggio di \hat{Z}_t con un operatore lineare tipo $\psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots - \psi_d B^d$ le cui d radici sono tutte sul cerchio unitario in modo da eliminare il trend stocastico, sotto l'ipotesi che non modifichi la stagionalità, e quindi ottenere:

$$\psi(B)\hat{Z}_t = \hat{X}_t = \hat{S}_{2t} + e_t$$

(iii) nel decomporre \hat{X}_t in \hat{S}_{2t} ed e_t , ove e_t è un white noise, ricorrendo al metodo della decomposizione canonica, cioè all'individuazione di e_t di modo che la sua varianza sia la massima possibile (vedi Box-Hillmer-Tiao (1978), Tiao-Hillmer (1978), Vitale (1980));
 (iv) nell'elaborazione di test che permettano di stabilire se le componenti deterministiche e stocastiche ottenute sono significative. Come si può notare, la procedura proposta da Pierce, è abbastanza generale e permette di derivare, anche se con una certa mole di calcoli, una soluzione al problema proposto. Resta il fatto che, come si evidenzierà nel paragrafo 4, anche se Y_t contiene componenti deterministiche e stocastiche, il modello ARIMA che lo rappresenta è, spesso, invertibile e quindi, nelle pratiche applicazioni, è quasi sempre possibile utilizzare la procedura descritta in (a).

3. L'approccio di Havenner-Swamy

A differenza di Pierce, questi Autori fanno osservare per prima cosa che nelle situazioni pratiche la decomposizione di Y_t in P_t , S_t ed e_t non è univoca a meno che quelle componenti non siano influenzate in modo differente da altre serie osservabili o da differenti politiche economiche. Per tener conto di tale non univocità nella decomposizione i due Autori citati permettono che P_t , S_t ed e_t siano fra di loro correlate e contengano, eventualmente, sia componenti deterministiche che stocastiche da derivare, a differenza di come proposto in Pierce, in modo simultaneo.

Per tener conto di queste esigenze e supponendo che la serie osservata sia mensile (ma la procedura può essere adattata a serie con cadenza diversa) si considera il seguente modello:

$$(3.1) \quad Y_t = X_t' \beta_t$$

ove X_t e β_t sono due vettori di dimensione (1,13) dati rispettivamente da

$$X_t' = \left(1, \cos\left(\pi \frac{t}{6}\right), \cos\left(2\pi \frac{t}{6}\right), \dots, \cos\left(5\pi \frac{t}{6}\right), \sin\left(\pi \frac{t}{6}\right), \sin\left(2\pi \frac{t}{6}\right), \dots, \sin\left(5\pi \frac{t}{6}\right) (-1)^t \right)$$

$$\beta_t' = (\beta_{0t}, \beta_{1t} \dots \beta_{12t})$$

in cui si assume:

$$(a) E(\beta_t) = \bar{\beta}_t = (\bar{\beta}_{0t}, \bar{\beta}_{1t} \dots \bar{\beta}_{12})$$

$$(b) (\beta_t - \bar{\beta}_t) = \phi(\beta_{t-1} - \bar{\beta}_t) + a_t$$

con ϕ matrice (13,13) generalmente non diagonale

(c) tutte le radici caratteristiche di ϕ sono in modulo minore di uno

(d) a_t è un *white noise* multivariato e precisamente $a_t \sim WN(O; \Sigma)$ con Σ matrice (13,13) definita positiva e generalmente non diagonale.

Il modello (3.1) è un modello di regressione i cui parametri sono stocastici e funzione di t . Il vantaggio di questo modello consiste nel fatto che permette di individuare in modo piuttosto semplice le seguenti componenti:

- (i) β_0 = livello medio del processo;
- (ii) $t \beta_{12}$ = trend deterministico del processo;
- (iii) $t(\beta_{12} - \bar{\beta}_{12})$ = trend stocastico del processo;
- (iv) posto

$$X_t^* = \left(\cos\left(\pi \frac{t}{6}\right), \cos\left(2\pi \frac{t}{6}\right), \dots, \cos\left(5\pi \frac{t}{6}\right), \sin\left(\pi \frac{t}{6}\right), \sin\left(2\pi \frac{t}{6}\right), \dots, \sin\left(5\pi \frac{t}{6}\right) (-1)^t t \right)$$

$$\beta_t^* = (\beta_{1t}, \beta_{2t} \dots \beta_{11t}); \quad \bar{\beta}_t^* = (\bar{\beta}_{1t}, \bar{\beta}_{2t} \dots \bar{\beta}_{11t})$$

si ha che

$$X_t^* \bar{\beta}_t^* = \text{stagionalità deterministica del processo};$$

$$X_t^* (\beta_t^* - \bar{\beta}_t^*) = \text{stagionalità stocastica del processo};$$

$$(\beta_{0t} - \bar{\beta}_0) = \text{residuo casuale del processo.}$$

La stima dei parametri del modello e quindi delle componenti viene fatta con un metodo iterativo e ricorrendo ad alcune approssimazioni. I risultati vengono applicati alla serie dei Depositi Usa dal 1969 al 1976 e confrontati con quelli ottenuti da Pierce ed applicando il metodo X-11. Un limite di questo metodo è che la sua applicazione richiede la stima di un numero elevatissimo di parametri. In caso di serie mensile il numero dei parametri raggiunge la ben ragguardevole cifra di 273. Una conseguenza di questo stato di cose è che le stime che si ottengono sono poco efficienti. Inoltre, per la stima di Σ si introducono delle approssimazioni i cui effetti sono del tutto sconosciuti.

E' piuttosto semplice verificare che il modello (3.1) è riconducibile ad un particolare modello della classe ARIMA i cui parametri MA sono dipendenti dal tempo. Infatti, da

$$(\beta_t - \bar{\beta}_t) = \phi(\beta_{t-1} - \bar{\beta}_t) + a_t$$

si ricava

$$\beta_t = \bar{\beta} + [I - \phi B]^{-1} a_t$$

che sostituita nella (3.1) implica

$$(3.2) \quad Y_t = X_t' \beta + X_t' [I - \phi B]^{-1} a_t$$

Se indichiamo con $A(B)$ l'aggiunta della matrice $(I - \phi B)$ e con $|I - \phi B|$ il suo determinante, la (3.2) si può scrivere

$$(3.3) \quad |I - \phi B| Y_t = |I - \phi B| X_t' \beta + X_t' A(B) a_t$$

Dato che è $X_t^{*'} = B^{12} X_t^{*'} = X_{t-12}^{*'}$ posto $Z_t' = (1, X_t^{*'})$ si avrà $Z_t' = B^{12} Z_t^{*' } = Z_{t-12}'$ e quindi risulta $X_t' = (Z_t', t)$; $X_{t-12}' = (Z_t', t - 12)$ per cui applicando $\nabla_{12} = 1 - B^{12}$ ad ambo i membri della (3.3) si avrà:

$$\nabla_{12} |I - \phi B| Y_t = 12 \beta_{12} |I - \phi| + \nabla_{12} X_t' A(B) a_t$$

e quindi

$$\nabla_{12} |I - \phi B| Y_t = 12 \bar{\beta}_{12} |I - \phi| + \delta_t(B) a_t$$

ove si è posto:

$$\delta_t(B) = [X_t' A(B) \nabla_{12} + 12 a_{13}'(B) B^{12}]$$

in cui $a_{13}(B)$ indica il vettore costituito dall'ultima riga di $A(B)$. Dato che $\delta_t(B) a_t$ è una combinazione lineare di *white noise* è possibile individuare un polinomio $\theta_t(B)$ ed un *white noise* ε_t per cui risulta

$$\theta_t(B) \varepsilon_t = \delta_t(B) a_t$$

In definitiva, il modello proposto da Havenner-Swamy può essere ricondotto ad uno del tipo

$$(3.4) \quad \nabla_{12} |I - \phi B| Y_t = \bar{\beta}_{12} |I - \phi| + \theta_t(B) \varepsilon_t$$

che è un modello stagionale della classe ARIMA con i parametri media mobile variabili nel tempo. In particolare, se è

$$(3.5) \quad a_t = \mathbf{K}\varepsilon_t$$

in cui si è posto $\mathbf{K}' = (K_0, K_1, \dots, K_{12})'$ allora il polinomio $\theta_t(B)$ figurante nella (3.4) è dato semplicemente da

$$\theta_t(B) = \delta_t'(B)\mathbf{K}$$

Il modello ARIMA (3.4) è molto complesso, infatti ha 13 parametri autoregressivi e 24 media mobile e questi ultimi sono variabili nel tempo. Ciò, se rende tale modello molto flessibile, lo rende anche troppo complicato per la maggior parte delle pratiche applicazioni in cui le serie possono essere rappresentate con modelli molto più semplici. La complessità di questo modello non ne giustifica, quindi, il suo uso.

Ricordando che la stagionalità del modello (3.1) è data da $S_t = X_t^{*'}\beta_t^*$ indicando con $A_{1,13}(B)$ la matrice (11,13) ottenuta da $A(B)$ dopo averne eliminato la prima e l'ultima riga, non è difficile verificare che è:

$$(3.6) \quad \nabla_{12}|I - \phi\mathbf{B}|S_t = X_t^{*'} A_{1,13}(B)\nabla_{12}a_t$$

che ha la stessa struttura AR del modello per Y_t , ed i parametri MA sono funzioni di t in modo recursivo dato che è $X_t^* = X_{t-h}^*$ $h = \pm 1, \neq 2, \dots$. Nel caso particolare in cui è vera la (3.5), la (3.6) diviene:

$$\nabla_{12}|I - \phi\mathbf{B}|S_t = (X_t^{*'} A_{1,13}(B)\nabla_{12}\mathbf{K})a_t$$

Si desume così che la stagionalità che implica il modello (3.1) è molto più complicata di quella che può attendersi nelle pratiche applicazioni o da quella presa in considerazione da altri metodi come l'X-11 (vedi Cleveland-Tiao (1976)) che pure danno risultati molto buoni.

Anche la componente ciclo-trend definita da $P_t = t\beta_{12t}$ è rappresentabile con un modello della classe ARIMA, infatti:

$$|I - \phi\mathbf{B}|P_t = t\{\bar{\beta}_{12} + a_{13}(B)a_t\}$$

e dividendo ambo i membri per t si ricava

$$|I - \phi B| \frac{P_t}{t} = \bar{\beta}_{12} + a_{13}(B)a_t$$

che è un ARIMA con i parametri sia AR che MA costanti nel tempo. In particolare, se la (3.5) è vera si ha:

$$|I - \phi B| \frac{P_t}{t} = \bar{\beta}_{12} + [a_{13}(B)K]\varepsilon_t$$

4. Il modello ARIMA e le componenti deterministiche

In questo paragrafo mostreremo, seguendo quanto dimostrato in Vitale (1982a), come la classe dei modelli ARIMA invertibile possa rappresentare, in molti casi pratici, processi che contengono sia componenti stocastiche che deterministiche.

A tal fine supponiamo che il processo Y sia dato da

$$(4.1) \quad Y_t = f(t, \Lambda) + Z_t$$

ove Z_t è un processo rappresentabile con un modello della classe ARIMA invertibile e precisamente

$$\psi_Z(B)\phi_f(B)Z_t = \theta_Z(B)a_t$$

mentre $f(t, \Lambda)$ soddisfa la seguente equazione lineare alle differenze finite omogenea:

$$\psi_f(B)\phi_f(B)f(t, \Lambda) = \theta_f$$

in cui:

- (i) $\psi_Z(B)$ e $\psi_f(B)$ sono polinomi di grado, rispettivamente, d ed r con tutte le radici sul cerchio unitario;
- (ii) $\phi_Z(B)$, $\theta_Z(B)$, $\theta_f(B)$ sono polinomi di grado, rispettivamente, p , q , c con tutte le radici fuori del cerchio unitario.

Condizione necessaria e sufficiente (vedi Vitale, 1982a) perchè Y_t sia rappresentabile con un ARIMA invertibile è che sia

$$(4.2) \quad \psi_Z(B) = \psi_f(B)\psi_{Z^*}(B)$$

ciò vuol dire che la parte stocastica non stazionaria di Y_t deve "dominare" la parte deterministica non stazionaria. D'altra parte, è molto difficile che, nelle pratiche applicazioni, in un processo osservabile vi sia una componente non stazionaria perfettamente deterministica non accompagnata da una analoga componente stocastica con la stessa struttura AR. In altri termini, è difficile immaginare in un processo l'esistenza di un trend, per esempio, perfettamente lineare che non sia accompagnato da un trend stocastico che sia lineare solo in media. Ciò è confermato anche dalle innumerevoli applicazioni che sono state fatte con i modelli ARIMA in cui i casi di ARIMA non invertibili dovuti alla presenza di componenti deterministiche sono piuttosto rari. In conclusione, si può affermare che nelle pratiche applicazioni, anche se un processo contiene componenti deterministiche può sempre essere rappresentato, magari con qualche approssimazione, con un modello della classe ARIMA invertibile. Ciò vuol dire che la ricerca di componenti come le classiche ciclo-trend, stagionalità e residuo può essere effettuata, quasi sempre, con procedure che utilizzano le medie mobili basate su modelli ARIMA invertibili. In tal modo risultano poco giustificate le procedure proposte da Pierce e da Havenner-Swamy.

5. La decomposizione di un $ARIMA(0, 1, r+1)(0, 1, 0)_r$

In questo paragrafo mostreremo come sia possibile decomporre in componenti correlate una serie a cui sia stato adattato un modello della classe $ARIMA(0, 1, r+1)(0, 1, 0)_r$ invertibile. Ovviamente, la metodologia qui delineata può essere estesa, con gli aggiustamenti del caso, per decomporre serie storiche a cui si sia adattato un modello qualsiasi della classe ARIMA.

Sia, a tale proposito Y_t il processo su cui sia stata osservata la serie storica Y_1, Y_2, \dots, Y_N e supponiamo che Y_t possa essere rappresentato dal modello

$$(5.1) \quad (1 - B)(1 - B^r)Y_t = \left(1 - \sum_{j=1}^{r+1} \theta_j B^j \right) a_t, \quad a_t \sim WN(0, \sigma_a^2)$$

I parametri del modello (5.1) possono essere stimati utilizzando le N informazioni della serie storica con gli usuali metodi di stima. Nel seguito supporremo, per rendere la trattazione più snella, che quei parametri sono noti o stimati in modo esatto. Il modello (5.1) è abbastanza generale ed è in grado di rappresentare numerose serie specialmente di tipo economico con stagionalità di periodo r . In particolare, il modello Holt-Winters (vedi Vitale, 1983) è un caso particolare della (5.1), mentre quello X-11 è molto bene approssimato da tale modello (vedi Cleveland-Tiao, 1976).

Il problema che ci si pone è di decomporre Y_t di modo che sia

$$(5.2) \quad Y_t = S_t + D_t, \quad t = 1, 2, \dots,$$

con S_t e D_t interpretabili rispettivamente come la componente stagionale e quella destagionalizzata entrambe rappresentabili con modelli della classe ARIMA. Si tratta perciò di un tipico problema di disaggregazione per cui, a meno di non imporre ulteriori restrizioni, ammette più di una (generalmente infinite) soluzioni. E' da sottolineare, per il momento, che richiedere che S e D siano rappresentabili con ARIMA è di per sè una restrizione che però non è sufficiente per individuare una soluzione unica. Nel seguito per ottenere una sola soluzione utilizzeremo sia risultati teorici, sia considerazioni di tipo empirico o di opportunità connesse al particolare problema che si sta esaminando. A tale proposito, notiamo che:

(a) dato che S_t è la stagionalità di Y_t , il suo operatore AR può essere dato proprio da $(1-B)^r$, di conseguenza l'operatore AR di D_t deve essere $(1-B)^2$;

(b) la componente D_t può contenere, a sua volta, una componente residua et supposta un *white noise* e quindi può essere $D_t = C_t + e_t$ ma in tal caso il polinomio MA di D_t deve essere di ordine $q \geq 2$ (vedi Vitale (1983)) per cui, tenuto conto del principio di parsimonia dei parametri, si può porre $q = 2$ e considerare per D_t il modello

$$(1-B)^2 D_t = (1 - \delta_1 B - \delta_2 B^2) d_t, \quad d_t \sim WN(0, \sigma_d^2)$$

(c) dato che S_t è una "parte" di Y_t è ragionevole imporre che le componenti AR ed MA di S_t siano più "semplici" (cioè di grado inferiore) delle rispettive componenti AR ed MA di Y_t , ciò vuol dire che il polinomio MA di S_t deve essere al massimo di grado r .

Tenuto conto di queste considerazioni, si può porre:

$$(5.3) \quad \begin{aligned} (1-B)^r S_t &= \theta_s(B) s_t, \quad s_t \sim WN(0, \sigma_s^2) \\ (1-B)^2 D_t &= (1 - \delta_1 B - \delta_2 B^2) d_t, \quad d_t \sim WN(0, \sigma_d^2) \end{aligned}$$

Le informazioni che ci forniscono congiuntamente le (5.1) -(5.3) non sono sufficienti per arrivare ad una decomposizione univoca. E' necessario, a tale fine, ipotizzare i legami intercorrenti fra S_t e D_t o equivalentemente fra S_t e d_t . Supponiamo, perciò, che sia

$$(5.4) \quad \begin{cases} s_t = \alpha a_t \\ d_t = \beta a_t \end{cases}$$

e quindi che S_t e D_t siano fra di loro correlate. In tal modo le (5.3) diventano

$$(5.5) \quad \begin{cases} (1-B^r)S_t = \theta_s(B)\alpha a_t \\ (1-B)^2 D_t = (1-\delta_1 B - \delta_2 B^2)\beta a_t \end{cases}$$

Da una prima analisi delle (5.5) emerge che:

- (i) se è $\alpha=0$ allora S_t è deterministica,
- (ii) se è $\beta=0$ allora D_t è deterministica ,
- (iii) se è $\theta_s(B)=0$ ha almeno una radice coincidente con almeno una di $(1-B^r)=0$ allora S_t contiene necessariamente una componente deterministica (vedi Vitale, 1982b),
- (iv) se $(1-\delta_1 B - \delta_2 B^2) = 0$ ha radici pari ad uno allora D_t contiene necessariamente una componente deterministica (vedi Vitale, 1982b).

Se ad ambo i membri della (5.2) applichiamo l'operatore $(1-B)(1-B^r)$, teniamo conto delle (5.5) e che è $(1-B^r) = (1-B)(1+B+B^2+\dots+B^r)$ avremo:

$$(5.6) \quad \begin{aligned} (1-B)\theta_s(B)\alpha + (1+B+B^2+\dots+B^r)(1-\delta_1 B - \delta_2 B^2)\beta = \\ = (1-\theta_1 B - \theta_2 B^2 - \dots - \theta_{r+1} B^{r+1}) \end{aligned}$$

Da questa uguaglianza si ricava che $\theta_s(B)$ è al massimo di grado r , se si impone che il grado $\theta_s(B)$ sia esattamente pari a $r-1$. Uguagliando i coefficienti di ambo i membri della (5.6), si ottiene il seguente sistema:

$$(5.7) \quad \begin{cases} \alpha + \beta = 1 \\ \theta_{1,s}\alpha + \alpha - \beta + \delta_1\beta = \theta_1 \\ \theta_{j,s}\alpha + \theta_{j-1,s}\alpha + \alpha - \beta + \delta_1\beta + \delta_2\beta = \theta_j \\ \theta_{r-1,s}\alpha + \alpha - \beta + \delta_2\beta = \theta_r \\ \delta_2\beta = \theta_{r+1} \end{cases}$$

Se si suppone α o β noto, posto $\theta(1) = (1-\theta_1-\theta_2-\dots-\theta_{r+1})$, si ottiene la soluzione seguente:

$$(5.8) \quad \begin{cases} \alpha = 1 - \beta \\ \delta_1\beta = \beta - \frac{\theta(1)}{r} - \theta_{r-1} \\ \delta_2\beta = \theta_{r+1} \\ \theta_{j,s}\alpha = \frac{j\theta(1)}{r} + \sum_{i=1}^j \theta_i + \theta_{r+1} - \alpha, \quad j = 1, \dots, r-1 \end{cases}$$

Dato che deve essere

$$-1 \leq \delta_2 = \frac{\theta_{r+1}}{\beta} \leq 1; \quad \delta_2 + \delta_1 \leq 1; \quad \delta_2 - \delta_1 \leq 1;$$

$$-1 \leq \theta_{r-1,s} = \frac{\theta(1)}{r(1-\beta)} + \frac{\theta_r}{(1-\beta)} - \frac{\beta}{(1-\beta)} \leq 1$$

risulta

$$(i) \quad \frac{\theta(1)}{r} \geq 0; \quad \theta_{r+1} + \frac{\theta(1)}{r} \leq 1$$

$$(ii) \quad \beta \leq |\theta_{r+1}|; \quad \beta \geq \theta_{r+1} + \frac{\theta(1)}{r}; \quad \beta \leq \frac{\left(1 + \theta_r + \frac{\theta(1)}{r}\right)}{2}$$

ove la (i) definisce quando la decomposizione è possibile con la metodologia qui proposta, mentre la (ii) definisce il campo di variazione di β (e quindi di α). Una volta accertato che, in una data applicazione, la (i) è vera, bisogna scegliere un β in:

$$(5.9) \quad \text{Max} \left\{ |\theta_{r+1}|, \theta_{r+1} + \frac{\theta(1)}{r} \leq \beta \leq \frac{\left(1 + \theta_r + \frac{\theta(1)}{r}\right)}{2} \right\}$$

In altri termini, bisogna scegliere quanto deve essere perequata la componente destagionalizzata o, equivalentemente, quanto elevato deve essere il legame fra D_t ed S_t . Si noti che per $\beta = 0$, D_t è deterministica e quindi le due componenti sono fra di loro incorrelate. Per determinare β all'interno del campo di variazione individuato si può scegliere quel valore per cui la relativa componente destagionalizzata presenti elementi stagionali il meno possibile. Questo può essere verificato considerando le autocorrelazioni di $U_t = (1-B)^2 D_t$ ai lag stagionali $r, 2r, 3r$, ecc.

Una volta effettuata la decomposizione, è necessario elaborare il filtro ottimale che applicato ad Y_t generi una stima delle componenti. A tale proposito, dalla (5.1) e dalle (5.5) si ricava immediatamente

$$\hat{S}_t = \alpha \frac{(1-B)\theta_s(B)}{\theta(B)} Y_t = W_s(B) Y_t; \quad \hat{D}_t = \beta \frac{S(B)\delta(B)}{\theta(B)} Y_t = W_D(B) Y_t$$

ove si è posto $S(B) = 1 + B + B^2 + \dots + B^{r-1}$ e $\delta(B) = 1 - \delta_1 B - \delta_2 B^2$, ed è facile verificare che è: $W_s(B) + W_D(B) = 1$. Notiamo ora che essendo il modello ARIMA per Y_t invertibile per ipotesi, è possibile scrivere:

$$W_s(B) = \alpha \frac{(1-B)\theta_s(B)}{\theta(B)} = \sum_{j=0}^{\infty} w_{sj} B^j, \quad \text{con} \quad \sum_{j=0}^{\infty} w_{sj} < \infty$$

$$W_D(B) = \beta \frac{S(B)\delta(B)}{\theta(B)} = \sum_{j=0}^{\infty} w_{Dj} B^j, \quad \text{con} \quad \sum_{j=0}^{\infty} w_{Dj} < \infty$$

che per $B = I$ forniscono la somma dei pesi dei filtri che in questo caso è facile verificare essere pari a:

$$(5.10) \quad \begin{cases} W_s(1) = \sum_{j=0}^{\infty} w_{sj} = 0 \\ W_D(1) = \sum_{j=0}^{\infty} w_{Dj} = 1 \end{cases}$$

come è da attendersi da filtri che individuano rispettivamente la componente stagionale e quella destagionalizzata. Da queste considerazioni segue che i pesi W_{sj} e W_{Dj} tendono ad essere trascurabili per valori elevati di j . Se è M quel valore di j per cui al posto dei filtri $W_s(B)$ e $W_D(B)$ possiamo considerare quelli *troncati* in M otteniamo rispettivamente

$$W_s(B; M) = \sum_{j=0}^{M-1} w_{sj} B^j; \quad W_D(B; M) = \sum_{j=0}^{M-1} w_{Dj} B^j;$$

ovviamente questi filtri non soddisfano le (5.10). Per ottenere un filtro *troncato e corretto*, consideriamo il seguente:

$$\hat{W}_D(B; M) = \frac{W_D(B; M)}{W_D(1; M)} = \frac{\sum_{j=0}^{M-1} w_{Dj} B^j}{W_D(1; M)}$$

Applicando questo filtro troncato e corretto alla serie storica osservata su Y_t si ottiene:

$$\begin{cases} \hat{D}_t = \hat{W}_D(B; M) Y_t = \sum_{j=0}^{M-1} w_{Dj} Y_{t-j} & t = M, M+1, \dots, N \\ \hat{S}_t = Y_t - \hat{D}_t \end{cases}$$

Se si vuole una stima anche delle prime $(M-1)$ osservazioni delle componenti si possono calcolare le previsioni $Y_t(-j)$ di Y_{t-j} , $j=1, 2, \dots, M-1$, considerare la serie estesa:

$$Y_1(-M+1), Y_1(-M+2), Y_1(-1), Y_1, Y_2, \dots, Y_N$$

ed applicare a questa i filtri troncati e corretti. Tuttavia, se queste correzioni garantiscono che la somma dei pesi per D_t sono pari ad uno, non garantiscono che sia anche

$$\sum_{t=1}^{N^*} \hat{S}_t = 0$$

così come si richiede alla componente stagionale. Per far in modo che ciò accada si effettuano i seguenti aggiustamenti:

(a) si calcola $\sum_{t=1}^{N^*} \hat{S}_t = A$ ove è $N^* = kr$ con $k = [N/r]$ = massimo intero di N/r .

(b) si calcola $\hat{S}_t^* = \hat{S}_t - \frac{A}{N}$, $t = 1, 2, \dots, N$

(c) si ottiene infine: $\hat{D}_t^* = Y_t - \hat{S}_t$

con questi accorgimenti il problema della decomposizione di una serie rappresentabile con un $ARIMA(0,1, r+1)(0,1,0)_r$ risulta soddisfacentemente risolto. Dato che i filtri asimmetrici inducono degli sfasamenti nelle componenti è necessario ricorrere a filtri simmetrici che evitano un tale inconveniente anche se, come accade in questo caso di correlazione fra le componenti, tali filtri sono subottimali.

Se lavoriamo *come se* le componenti fossero incorrelate, i filtri simmetrici sono dati semplicemente da:

$$(5.11) \quad \begin{cases} W_D^*(B; F) = W_D(B)W_D(F) \\ W_S^*(B; F) = W_S(B)W_S(F) \end{cases}$$

ove si è posto $F = B^{-1}$. Ovviamente anche in questo caso risulta: $W_D^*(1;1) = 1$ $W_S^*(1;1) = 0$ e quindi i pesi dei filtri tendono ad essere trascurabili dopo un certo $j \geq M$. Similmente a quanto visto per i filtri asimmetrici, si possono calcolare i pesi simmetrici, corretti e troncati ad M tramite:

$$W_D^*(B, F; M) = \frac{W_D^*(B, F; M)}{W_D^*(B, F; 1)}$$

Applicando questo filtro ad Y_t si perdono le prime e le ultime $M-1$ osservazioni. Per avere le osservazioni delle componenti anche in tali punti estremi si può estendere la serie effettuando $M-1$ previsioni in avanti a partire da N ed $M-1$ previsioni all'indietro a partire da 1 ed applicare su tale serie estesa il filtro simmetrico, corretto e troncato ad M . E' stato dimostrato che questo modo di procedere fornisce risultati ottimali nel senso che minimizza le revisioni (vedi Pierce, 1980, Wallis, 1982).

6. Una applicazione

La metodologia di decomposizione illustrata nel paragrafo precedente viene applicata alla serie Depositi Bancari negli USA (nel seguito DD_t) osservata mensilmente dal gennaio 1969 al dicembre 1976 e riportata in appendice al lavoro di Havenner-Swamy (1981).

Dato che sia Pierce (1978) che Havenner-Swamy nell'analisi di tale serie considerano il suo logaritmo, nel seguito noi faremo altrettanto e la serie così trasformata la indicheremo con LD_t . Nella Fig. 1 è riportato il grafico di tale serie trasformata da cui emerge l'esistenza di un forte trend ed una forte e regolare stagionalità. Utilizzando gli usuali metodi di identificazione e stima dei modelli ARIMA, alla serie LD_t è stato adattato il seguente modello

$$(6.1) \quad (1 - B)(1 - B^{12})D_t = (1 - \theta_{12}B^{12} - \theta_{13}B^{13})a_t$$

ottenendo le stime dei parametri qui di seguito riportate (fra parentesi è indicato il corrispondente scarto quadratico medio stimato) :

$$\begin{array}{ccc} \theta_{12} & \theta_{13} & \sigma_a^2 \\ 0.690 & 0.190 & \\ (0.081) & (0.086) & 0.29124 \times 10^{-4} \end{array}$$

che sono da considerarsi significativamente diversi da zero. Nelle Fig. 2 sono riportate le autocorrelazioni di (LD_t) , $\nabla(LD_t)$, $\nabla_{12}(LD_t)$ e quella dei residui del modello (6.1) con i relativi intervalli di confidenza approssimati al 5% che mostrano come siano compatibili con una realizzazione da *white noise*, il che è indice di adeguatezza del modello considerato. I test globali di Box-Pierce e Ljung-Box danno un valore rispettivamente di $Q = 11.9$ e $Q' = 14.5$ ed indicano un ottimo adattamento del modello alla serie dato che il relativo chi-quadrato al 5% con 18 gradi di libertà dà un valore pari a 28.9.

Nella Fig. 3 sono riportati i residui del modello stimato e questi non evidenziano né valori eccessivamente anomali, né un qualche sottostante andamento regolare. Tutto questo sta a significare che il modello (6.1) adattato a LD_t rappresenta abbastanza bene la serie tenuto anche conto che spiega più del 99%, della sua variabilità.

Nella Tab. 1 è riportata la serie DD_t dal 1969 al 1976, le previsioni per il biennio 1966-68 ed il biennio 1977-78 (i valori sono espressi in milioni di dollari). Dai parametri stimati per il modello, che nel seguito li considereremo come se fossero quelli veri, risulta:

$$\frac{\theta(1)}{r} + \theta_r = 0.7; \quad 0.2 \leq \beta \leq 0.85$$

ciò vuol dire che la decomposizione è possibile per tutti i β dell'intervallo $[0.2, 0.85]$. In

particolare, per i seguenti valori di b : 0.2; .525; .85 si ottengono i seguenti modelli ARIMA per la componente destagionalizzata:

Per $\beta= 0.2$ si ha

$$(1 - B^2)D_t = (1 - 0.95B^2)0.2a_t \Rightarrow (1 - B^2)D_t = (0 - 0.9747B)0.2a_t$$

da cui segue che entrambe le radici della parte MA di D_t sono vicine al cerchio unitario, ciò vuol dire che in D_t vi sono due componenti "quasi" deterministiche.

Per $\beta= 0.525$ si ha

$$(1 - B^2)D_t = (1 - 0.619B^2 - 0.362B^2)0.525a_t \Rightarrow (1 - B^2)D_t = (0 - 0.985B)0.525a_t$$

ed in questo caso D_t possiede almeno una componente "quasi" deterministica;

Per $\beta= 0.85$ si ha

$$(1 - B^2)D_t = (1 - 0.7647B^2 - 0.2257B^2)0.85a_t \Rightarrow \\ (1 - B^2)D_t = (0 - 0.9904B)(1 + 0.2257B)0.85a_t$$

e questa situazione è abbastanza simile alla precedente.

In conclusione, si può affermare che quando b assume il valore minimo nel suo intervallo di definizione, in D_t vi sono due componenti che in pratica possono considerarsi deterministiche, mentre negli altri casi è certo che di tali componenti ne esiste una sola. D'altra parte, nella stima del modello per la serie abbiamo visto che l'incidenza del fattore casuale sulla evoluzione di (LD_t) , è inferiore all'1%, ciò vuol dire che nella serie analizzata sono presenti essenzialmente componenti deterministiche. Questa considerazione ci porta a scegliere un $\beta=0.2$. Che questa sia la scelta più plausibile è confermato anche dal fatto che le componenti destagionalizzate utilizzando il filtro asimmetrico ed un β maggiore di .2 danno residui di stagionalità superiore a quella del caso $\beta=0.2$. I valori destagionalizzati, ottenuti utilizzando il filtro asimmetrico e $b=0.2$, sono riportati, nella scala originaria e quindi dopo averne fatto l'antilogaritmo, nella Fig.4 e nella Tab.2 dal cui esame emerge come in D_t sia ancora presente un residuo di stagionalità. Questo fenomeno è dovuto al fatto che il modello di decomposizione utilizzato prevede che D_t ed S_t siano correlate. Dato che, in media, S_t è nulla, mentre D_t è crescente, l'unico modo perchè queste due componenti siano correlate è che D_t continui a possedere un residuo stagionale così come si riscontra nell'applicazione. Questo residuo di stagionalità può essere drasticamente attenuato, ma non completamente eliminato, come vedremo fra breve, se al posto del filtro asimmetrico se ne utilizza uno simmetrico operando come se le componenti fossero incorrelate. A tal fine dalle (5.11) si ha:

$$w_{Dj}^* = \sum_{i=0}^{\infty} w_{Dj} w_{D,i+j}$$

e quindi i pesi simmetrici possono essere facilmente ottenuti da quelli asimmetrici. Utilizzando quest' ultima relazione, dopo aver troncato la sommatoria a $M=60$, abbiamo calcolato i pesi w_{Dj}^* $j=0, 1, \dots, 24$ e li abbiamo corretti di modo che la loro somma risultasse pari ad uno. I pesi così ottenuti sono stati utilizzati per una nuova stima di D_t i cui valori sono riportati nella Tab. 2 e rappresentati nella Fig. 5 e dalla loro analisi emerge come il residuo di stagionalità prima riscontrato sia praticamente, ma non completamente, scomparso.

7. Alcuni confronti

I risultati ottenuti su (LD_t) applicando la procedura di destagionalizzazione qui proposta li confrontiamo, nel seguito, con quelli ottenuti applicando l' X-11 e le procedure di Havenner e Swamy e di Pierce.

Confrontiamo, in primo, luogo, i risultati da noi ottenuti con quelli derivanti dalla applicazione della procedura X-11. A tale proposito, dal lavoro di Cleveland e Tiao (1976) emerge che la componente ciclo-trend generata da X-11 è ottimale per il modello ARIMA seguente:

$$(1 - B)^2 T_t = (1 + 0.49B - 0.49B^2) \varepsilon_t$$

Ciò vuol dire che la componente destagionalizzata è data da $D_t = T_t + e_t$; $e_t \sim WN(0, \sigma_e^2)$ che implica per D_t il seguente modello ARIMA:

$$(1 - B)^2 D_t = (1 - \delta_1 B - \delta_2 B^2) u_t$$

con

$$(1 - \delta_1 B - \delta_2 B^2) u_t = (1 + 0.49B - 0.49B^2) \varepsilon_t + (1 - 2B + B^2) e_t$$

Tenuto conto che in X-11 le componenti si suppongono incorrelate e che è $\sigma_e^2 = 14.4\sigma_u^2$ (Cleveland e Tiao, 1976) pag.583), in termini di autocovarianze si ha:

$$(1 + 6\delta_1 + 6\delta_2) \sigma_u^2 = [1 + 2(0.49)] \sigma_e^2$$

da cui, dopo alcuni passaggi, posto $\sigma_e^2 = 14.4\sigma_u^2$, si ricava

$$\begin{cases} (1 + \delta_1^2 + \delta_2^2) \sigma_u^2 = [1 + 2(0.49)^2] \sigma_e^2 + 6\sigma_e^2 \\ (-\delta_1 + \delta_2 \delta_1) \sigma_u^2 = [0.49 - (0.49)^2] \sigma_e^2 - 4\sigma_e^2 \\ -\delta_2 \sigma_u^2 = -0.49\sigma_e^2 + \sigma_e^2 \end{cases}$$

da cui, dopo alcuni passaggi, posto $\sigma_u^2 = k\sigma_\varepsilon^2$, si ricava

$$\begin{cases} \delta_1 = 57.3501 / (k + 13.91) \\ \delta_2 = -13.91 / k \\ k^4 - 60.062k^3 + 1037.6949k^2 - (13.92)^2(60.062)k + (13.92)^4 = 0 \end{cases}$$

da quest' ultimo polinomio si ricava che l'unica radice accettabile si ha per $k=41.17915$ per cui il modello ARIMA per D_t risulta

$$(1 - B)^2 D_t = (1 - 1.04119B + 0.3378B^2)u_t; u_t \sim WN(0, 41.1769\sigma_\varepsilon^2)$$

Le radici del polinomio MA di D_t generato da X-11 sono complesse e lontane dal cerchio unitario. Ciò vuol dire che in D_t esiste una periodicità stocastica, ed in particolare una stagionalità stocastica di periodo pari a circa 13 mesi, anche se molto tenue. In altri termini, l'esistenza di un residuo stagionale è sempre presente, anche se molto debole, nella componente D generata da X-11 casi come è presente, per altra via, quando si postula che le componenti sono fra di loro correlate.

La componente D_t generata da X-11 per la nostra serie è riportata nella Tab. 2 e rappresentata nella Fig. 6. Se si confrontano le Fig. 5 e 6 si nota come queste siano molto simili. Ciò vuol dire che, almeno in questo caso, i risultati ottenibili con X-11 possono essere ottenuti con poca fatica tramite una diretta decomposizione degli ARIMA in ARIMA. Nella Fig. 7 e nella Tab. 2 sono riportati i dati destagionalizzati utilizzando la procedura di Havenner e Swamy in cui è facile individuarvi un residuo di stagionalità data anche la estrema somiglianza fra la Fig. 7 e la Fig.4. Ciò è dovuto, più che all'esistenza di correlazione fra le componenti, al fatto che con tale procedura i filtri, così come avviene in tutti i metodi regressivi, sono asimmetrici. E' da sottolineare, comunque, come a fronte della complessità del modello utilizzato per ottenere i dati destagionalizzati, i risultati ottenuti siano piuttosto modesti: gli stessi si possono ottenere utilizzando il ben più semplice modello (6.1) ed il relativo filtro asimmetrico. Ovviamente, se il filtro è simmetrico i risultati che si ottengono (confronta le Fig. 5 e 7) sono decisamente migliori.

Infine, nella Fig. 8 e nella Tab. 2 sono riportati i dati destagionalizzati con la procedura proposta da Pierce. Questirisultati non sono molto simili da quelli forniti da X-11 e a quelli ottenuti con il metodo da noi proposto utilizzando un filtro simmetrico.

8. Conclusioni e prospettive

Nel paragrafo precedente abbiamo visto come la procedura di decomposizione da noi proposta fornisca, almeno nel particolare caso da noi analizzato, risultati certamente non peggiori di quelli ottenuti tramite l' X-11 e la procedura di Pierce, mentre sono certamente migliori di quelli ottenuti con la procedura proposta da Havenner e Swamy. Ovviamente, è ancora prematuro trarre conclusioni generali dato che è necessaria una sperimentazione più estesa del metodo proposto; è comunque possibile affermare che:

1. l'importanza dei filtri simmetrici è cruciale sia per decomporre in componenti correlate che incorrelate. Lo dimostra il fatto che procedure complesse quale quella di Havenner e Swamy forniscono risultati non certo molto buoni e simili a quelli derivanti da procedure estremamente più semplici come è la decomposizione degli ARIMA in componenti correlate e l' uso di filtri asimmetrici;
2. procedure piuttosto semplici e poco "costose" come quella da noi proposta possono fornire risultati abbastanza buoni e comparabili con quelli derivanti dal ben noto e collaudato X-11 solo che si utilizzino filtri simmetrici;
3. la procedura di decomposizione qui proposta può essere, con la stessa tecnica, estesa per decomporre serie che hanno una struttura molto differente da un $ARIMA(0,1,r+1)(0,1,0)_r$;
4. è bene, comunque, tener presente che il modello ARIMA qui decomposto è di per sé abbastanza generale. Infatti, ne sono suoi casi particolari il ben noto modello $ARIMA(0,1,r+1)(0,1,0)$ e quello generato da Holt-Winters (vedi per esempio Vitale (1983)). Inoltre, quel modello può approssimare altri abbastanza bene come i seguenti $ARIMA(1,1,1)(0,1,1)_r$; $ARIMA(2,1,1)(0,1,1)_r$; se i parametri AR sono "lontani" dalla regione di non stazionarietà. Così se è

$$(1-B)(1-B^r)(1-\phi B)X_t = (1-\theta B)(1-\theta B^r)a_t \sim ARIMA(1,1,1)(0,1,1)_r$$

si può scrivere

$$(1-B)(1-B^r)X_t = (1-\phi B)^{-1}(1-\theta B)(1-\theta B^r)a_t$$

Tenuto conto che per $|\phi| < 1$ è

$$(1-\phi B)^{-1} = (1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \dots +)$$

è facile verificare che se risulta, come succede in molti casi pratici:

$$\phi(\theta - \phi)(\theta - \phi^{r+1}) \cong 0 \text{ allora si avrà:}$$

$$ARIMA(1, 1, 1)(0, 1, 1)_r \geq ARIMA(0, 1, r - 1)(0, 1, 0)$$

5. la procedura proposta può essere completamente, ed abbastanza facilmente, automatizzata in modo da destagionalizzare tutte le serie rappresentabili con un $ARIMA(0, 1, r+1)(0, 1, 0)_r$. Per far questo è necessario procedere come segue:

(i) trasformare preliminarmente la serie osservata in modo da riportarla al modello additivo $Y_t = T_t + S_t + e_t$. A tale proposito si veda quanto proposto in D'Esposito (1984);

(ii) ricercare ed aggiustare i valori "anomali" ed "eccezionali" presenti nella serie; più in generale è necessario rendere la serie "omogenea" con le solite tecniche predisposte a questo scopo (vedi D'Esposito (1984));

(iii) stimare gli $r+1$ parametri MA presenti nel modello ed estrapolare la serie in entrambi i lati per due o tre anni;

(iv) utilizzare i parametri stimati per ricavare l'intervallo di definizione di β ed i valori assunti da d_1 e da d_2 nel modello per D_t ;

(v) fissare β in base ad informazioni a priori relative alla maggiore o minore presenza di accidentalità nella serie osservata o per tentativi esplorando in tre o quattro punti il campo di variazione di β ;

(vi) calcolare i pesi asimmetrici e simmetrici troncati e corretti per la componente D_t ;

(vii) calcolare la serie destagionalizzata e quella stagionale utilizzando i filtri così determinati.

Noi crediamo che una tale procedura, anche se ancora da sperimentare a fondo, sia una valida alternativa a molte di quelle attualmente in uso.

Sommario

Si affronta il problema della destagionalizzazione di serie storiche quando queste contengono sia componenti deterministiche che stocastiche. Si analizzano le procedure di Pierce ed Havenner-Swamy e se ne commentano i rispettivi limiti. In particolare, si mostra come, in contrasto a quanto affermato da Pierce, la decomposizione può essere ottenuta in pratica sempre utilizzando la modellistica ARIMA anche in presenza di componenti deterministiche nella serie da decomporre e che il modello relativo alla procedura di Havenner-Swamy è un particolare, anche se complesso, modello ARIMA con i parametri MA variabili nel tempo. Si propone un metodo alternativo per destagionalizzare serie rappresentabili con $ARIMA(0, 1, r+1)(0, 1, 0)^r$ in componenti ARIMA correlate e si applica alla serie depositi USA. I risultati, confrontati con quelli ottenuti applicando le procedure X-11, Pierce ed Havenner-Swamy, dimostrano come la procedura proposta sia una valida alternativa ad alcune attualmente in uso.

Summary

We consider the seasonal adjustment problem of a time series containing stochastic and deterministic components. We analyse the procedure proposed by Pierce and Havenner-Swamy explaining their deficiencies. Particularly, we show, in contrast to what affirmed by Pierce, that the seasonal adjustment can always be done utilising the ARIMA models, even if there are deterministic components in the observed time series, and the Havenner-Swamy model is a particular, but complex, ARIMA model with the MA parameters varying in the time. We propose an alternative procedure to seasonal adjustment of time series following $ARIMA(0, 1, r+1)(0, 1, 0)^r$ in correlated ARIMA components and we apply to the demand deposit series for the USA. The results, compared with the X-11, Pierce and Havenner-Swamy procedures, show that our approach is a valid alternative to some currently utilized procedures.

Tab.1. Depositi USA (milioni di dollari): 1966-1968, 1977-1978,previsti 1969-1976:osservati
Mesi

Anni	G	F	M	A	M	G	L	A	S	O	N	D
1966	143.415	137.723	138.483	142.116	138.774	141.308	142.116	140.702	142.586	143.961	145.510	148.851
1967	150.437	144.452	145.248	149.074	145.568	148.212	149.059	147.576	149.552	151.010	152.634	156.123
1968	157.787	151.509	152.345	156.357	152.680	155.453	156.342	154.786	156.858	158.388	160.091	163.751
1969	165.539	158.863	159.068	162.736	158.182	160.603	161.183	159.014	161.066	162.545	163.925	167.822
1970	170.235	161.463	162.652	166.990	162.317	164.491	165.158	164.879	167.843	169.148	170.907	175.785
1971	177.040	170.876	172.399	177.010	173.881	177.362	178.670	177.270	179.068	180.159	181.840	186.905
1972	188.389	182.072	184.301	189.280	184.203	187.509	190.269	189.457	192.493	194.814	197.123	204.783
1973	206.369	198.010	197.740	202.270	198.754	203.695	205.196	202.566	203.385	205.213	208.912	215.615
1974	215.332	207.447	208.810	213.523	207.446	212.226	213.734	210.694	212.189	213.696	216.285	222.160
1975	219.823	210.595	212.527	217.368	212.857	219.131	220.271	217.826	219.861	219.874	223.476	228.095
1976	227.266	218.897	220.214	227.231	221.468	224.838	226.721	224.712	226.038	229.887	231.771	239.543
1977	239.522	230.983	232.606	238.637	232.722	237.685	239.354	236.688	238.685	241.301	244.116	251.324
1978	250.922	242.002	243.677	250.021	243.799	249.023	250.722	247.979	250.046	252.812	255.736	263.313

Tab.2. Componenti destagionalizzate con i diversi metodi

Mesi												
Anni	G	F	M	A	M	G	L	A	S	O	N	D
METODO PROPOSTO: FILTRI ASIMM												
196:	161.31	161.69	160.52	160.53	161.01	160.76	161.28	161.55	161.61	162.38	163.28	164.51
197:	166.07	165.93	164.44	164.97	165.66	165.28	165.68	166.64	167.61	168.57	169.56	171.20
197:	174.15	174.83	174.43	174.98	176.48	177.26	178.39	179.13	179.27	179.92	180.92	182.50
197:	185.71	186.65	186.29	187.21	188.15	188.23	189.87	191.47	192.40	194.06	195.84	198.51
197:	201.44	201.93	200.13	200.13	201.79	203.07	204.66	205.26	205.04	205.79	207.84	210.42
197:	210.84	211.08	210.26	210.84	211.42	211.54	212.97	213.62	213.63	214.36	215.55	217.27
197:	216.76	215.88	214.61	215.36	216.60	217.51	219.09	220.00	220.41	220.69	221.55	223.00
197:	224.60	224.66	223.11	224.11	225.63	224.93	225.72	227.02	227.23	228.70	230.25	232.31
METODO PROPOSTO: FILTRI SIMM												
196:	159.77	160.88	160.60	160.12	160.75	161.00	161.08	161.88	162.35	163.16	164.44	164.40
197:	164.74	165.87	165.62	165.16	166.00	166.47	166.78	167.95	168.88	170.20	171.68	171.96
197:	172.78	174.52	174.78	175.07	176.66	177.66	178.18	179.30	180.11	181.32	182.70	182.91
197:	183.84	185.78	186.15	186.49	188.06	189.03	189.74	191.19	193.19	195.19	197.17	197.73
197:	198.65	200.31	200.05	200.02	201.55	202.37	202.63	203.82	204.75	206.17	207.63	207.89
197:	208.94	210.76	210.41	210.20	211.91	211.68	211.75	212.89	213.57	214.52	215.32	214.89
197:	215.25	216.48	215.93	215.94	217.48	216.33	218.84	220.22	221.00	221.86	222.78	222.52
197:	223.11	224.57	224.20	224.10	225.44	226.06	226.24	227.79	228.95	230.55	232.16	232.69
METODO X:												
196:	159.78	160.46	160.83	161.28	161.57	161.81	162.07	161.92	162.20	162.70	162.94	162.93
197:	164.31	163.34	164.46	165.50	165.79	165.65	165.98	167.73	169.02	169.48	169.88	170.66
197:	171.38	172.96	174.40	175.43	177.61	178.43	179.20	180.15	180.33	180.52	180.75	181.46
197:	182.81	184.47	186.44	187.59	188.15	188.54	190.45	192.32	193.85	195.20	195.94	198.43
197:	200.55	200.61	200.24	200.46	203.01	204.71	205.19	205.44	204.81	205.62	207.66	208.96
197:	209.67	210.49	211.56	211.61	211.89	213.18	213.52	213.46	213.68	214.12	214.99	215.06
197:	214.25	214.12	215.54	215.42	217.42	220.12	219.83	220.47	221.41	220.31	222.14	220.80
197:	221.50	222.90	223.57	224.98	225.98	225.85	226.04	227.21	227.63	230.34	230.38	231.89
METODO DI HAVENNER - S1												
196:	159.63	160.08	160.04	160.86	161.23	161.64	162.01	162.11	162.50	163.03	163.40	163.45
197:	163.98	162.69	164.04	165.03	165.44	165.63	166.05	168.11	169.30	169.63	170.29	171.17
197:	170.73	172.20	173.93	175.00	177.21	178.43	179.45	180.78	180.69	180.56	180.88	181.81
197:	182.44	183.54	185.99	187.16	187.75	188.45	191.00	193.17	194.28	195.17	195.85	199.08
197:	200.44	199.66	199.60	200.02	202.59	204.57	205.87	206.46	205.33	205.55	207.47	209.54
197:	209.37	209.18	210.77	211.15	211.46	213.12	214.45	214.77	214.22	214.06	214.85	215.89
197:	213.58	212.34	214.49	214.92	216.97	220.14	221.07	222.05	221.93	220.29	222.11	221.73
197:	220.55	220.69	222.24	224.67	225.75	225.92	227.59	229.07	228.15	230.35	230.40	232.91
METODO DI PIEF												
196:	159.90	160.31	160.81	161.51	161.42	161.94	162.33	161.77	162.16	162.67	162.63	162.78
197:	164.76	163.10	164.40	165.59	165.43	165.65	166.08	167.53	168.91	169.36	169.81	170.82
197:	171.41	172.64	174.15	175.44	177.26	178.75	179.64	180.02	180.19	180.42	180.80	181.53
197:	182.50	183.81	186.31	187.83	187.77	188.67	190.87	192.32	194.02	195.34	196.03	198.81
197:	189.85	199.84	199.88	200.70	202.84	205.07	205.83	205.73	205.16	205.74	207.55	209.13
197:	208.89	209.87	211.56	212.24	211.76	213.22	214.00	213.70	213.81	214.13	214.51	215.47
197:	213.43	213.15	215.21	215.94	217.33	220.34	220.76	220.98	221.25	220.07	221.56	221.30

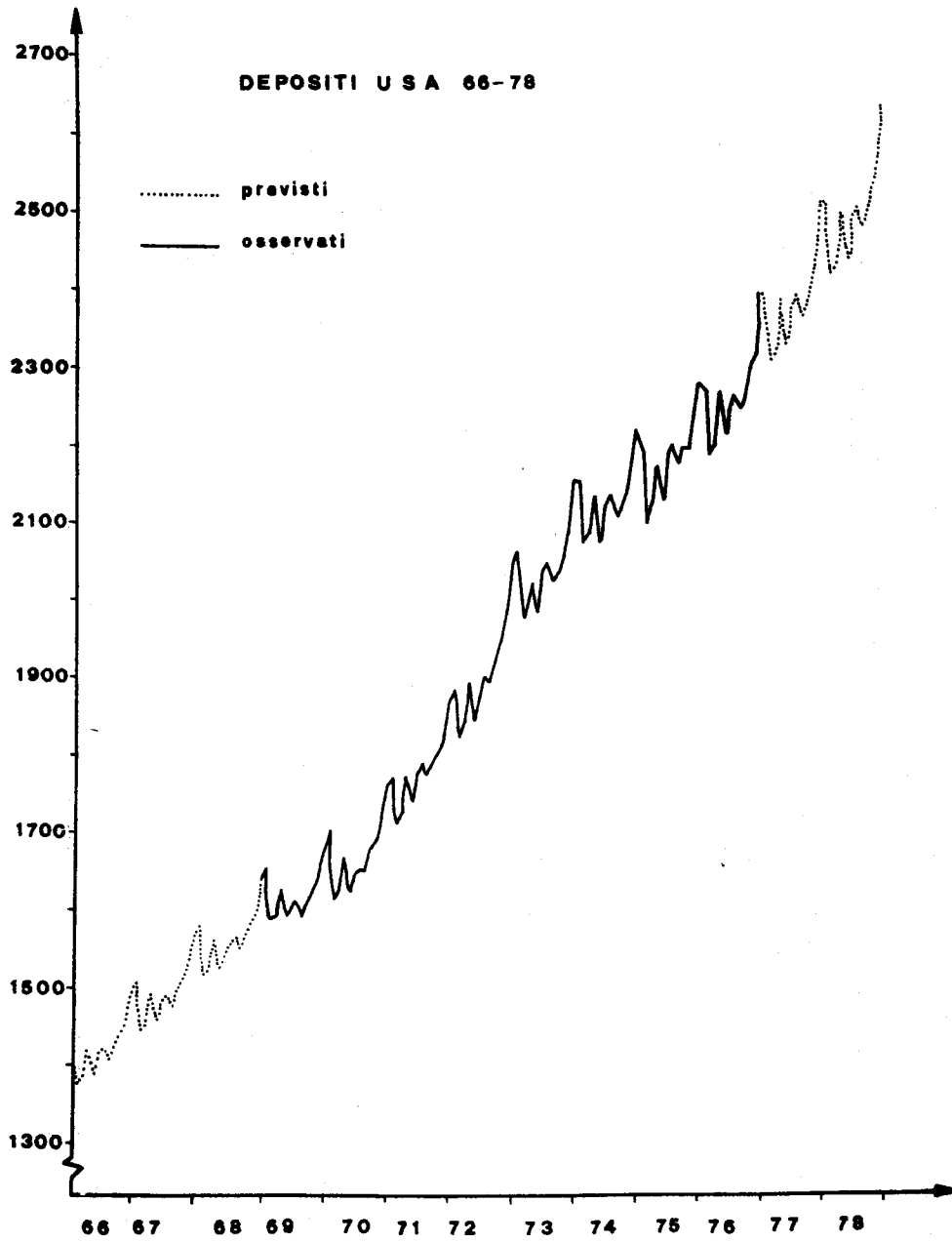


Fig.2(a) AUTOCORRELAZIONI DI

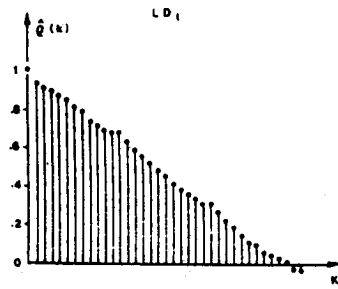


Fig.2(b) AUTOCORRELAZIONI DI

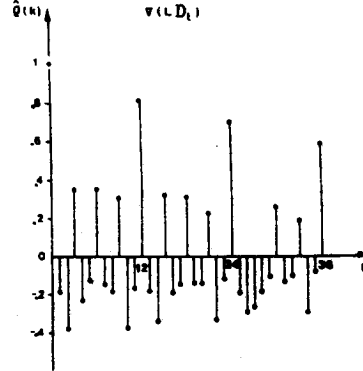


Fig.2(c) AUTOCORRELAZIONI DI

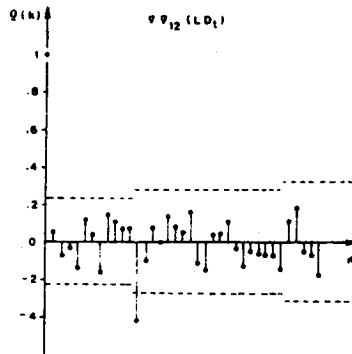
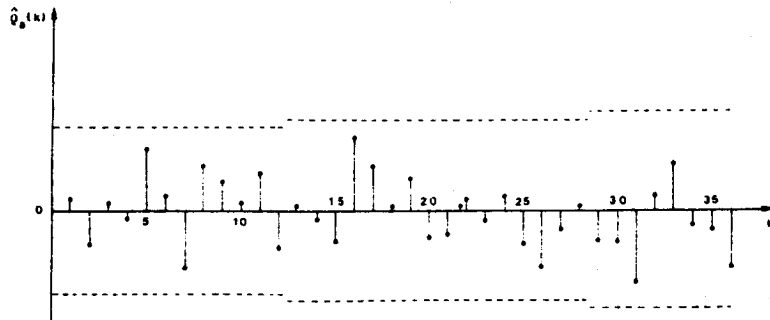
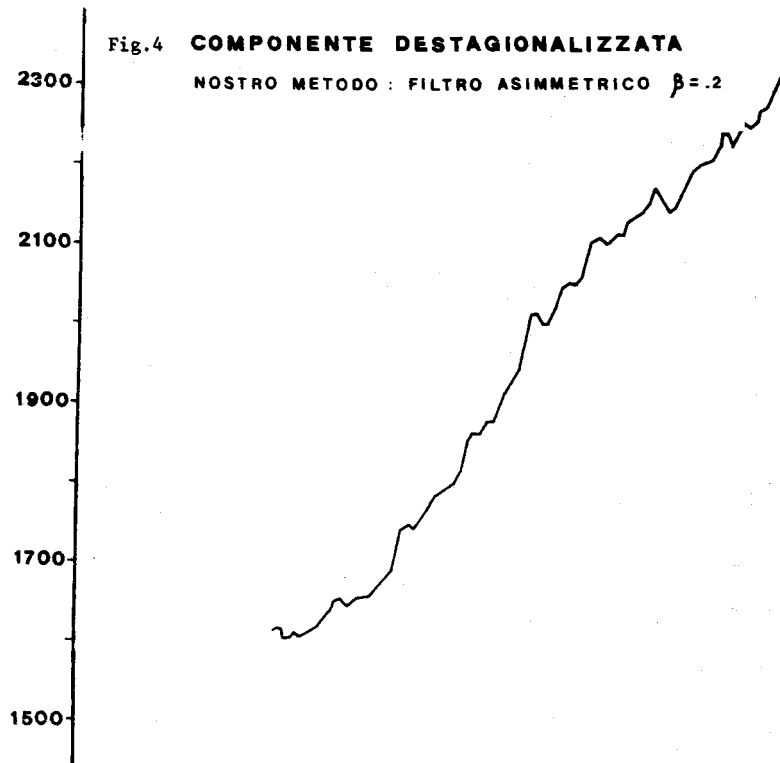
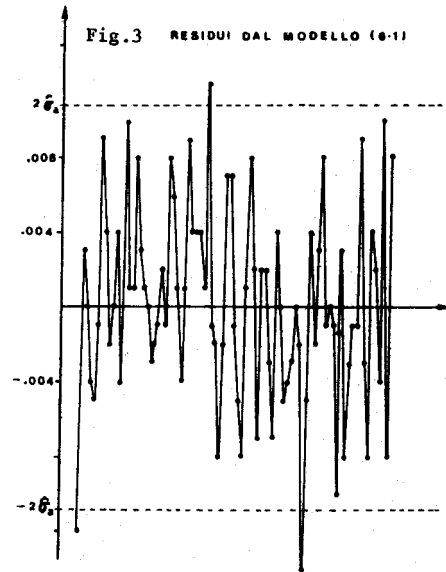
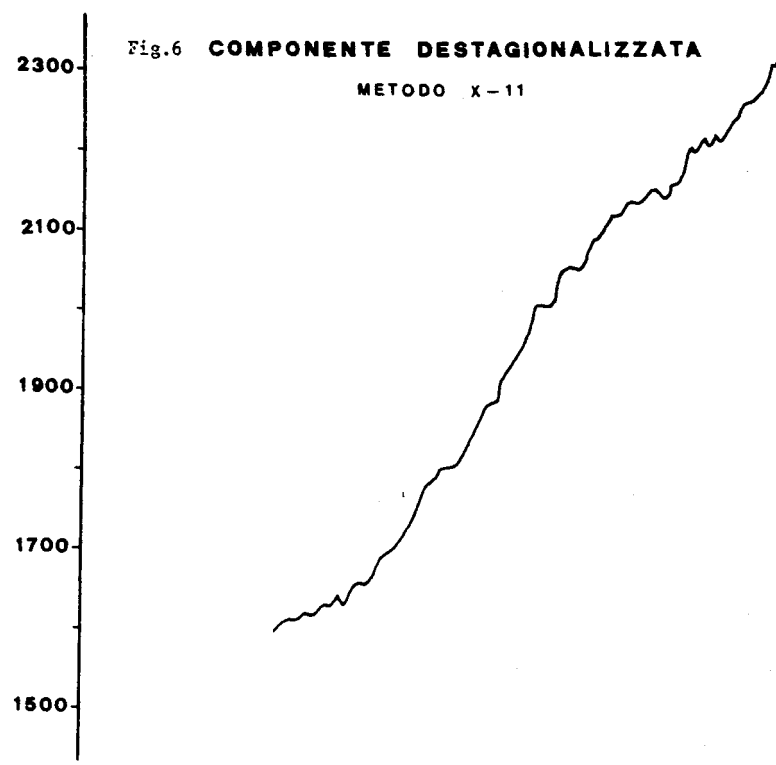
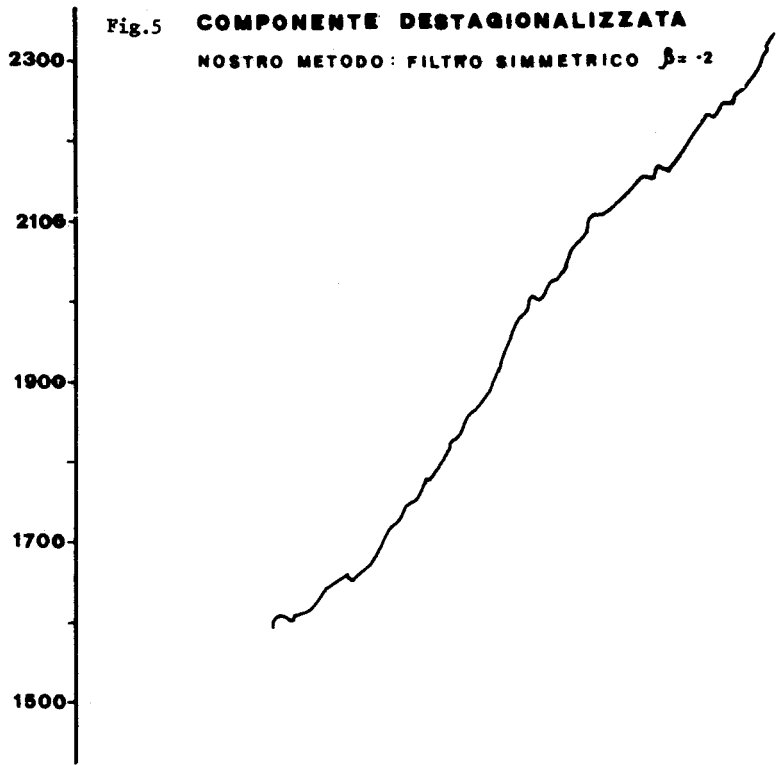


Fig.2(d) AUTOCORRELAZIONI DEI RESIDUI DAL MODELLO (8-1)



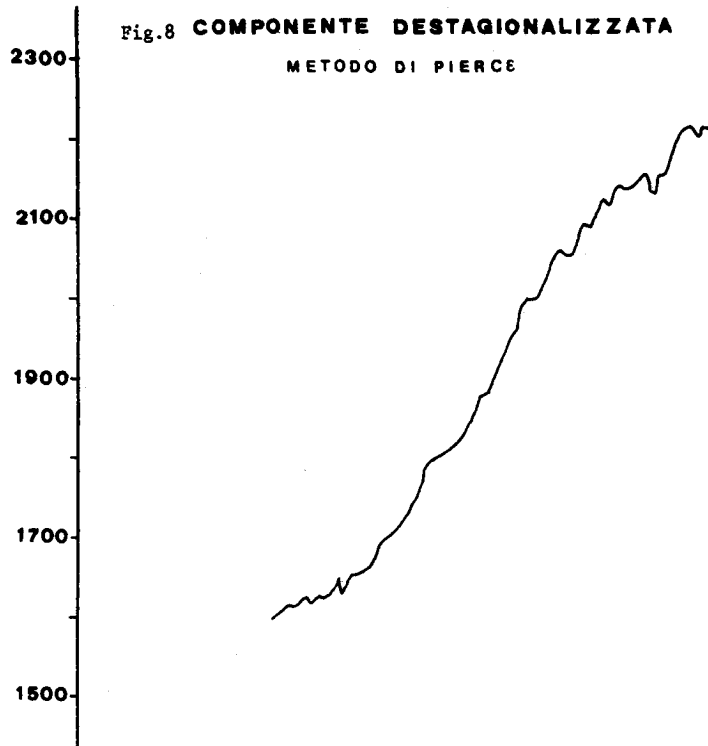
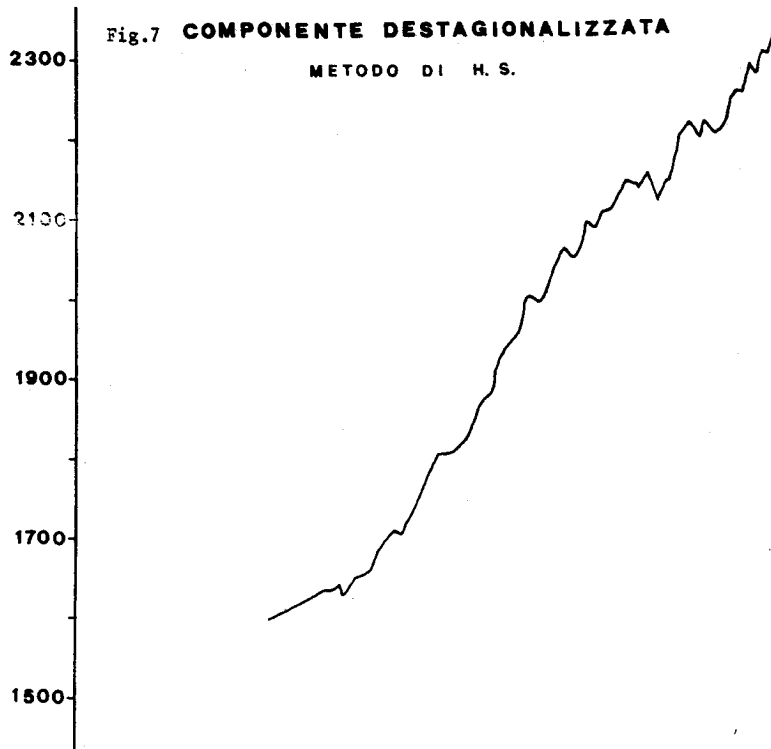




BIBLIOGRAFIA

1. ABRAHAM
models'
2. BOX, G.E
seasona
on Seas
in A. Ze
3. CLEVELA
mode fc
4. D'ESPOSIT
stagiona
5. HAVENNE
adjustm

dependent
 modelling of
 Conference
 , pubblicato
 78).
 ne series: a
 one per dati
 to seasonal
 09.



6. PIERCE, D.A., (1978): "Seasonal adjustment when both deterministic and stochastic seasonality are present" Special Studies Paper - Division of Research and Statistics, Federal Reserve Board, Washington, D.C., N. 107, in A. Zellner (edit.) "Seasonal Analysis of Economic Time Series" (1978).
7. PIERCE, D.A., (1979): "Signal extraction error in nonstationary time series" Ann. Statistics, vol. 7 pp. 1303-1320.
8. PIERCE, D.A., (1980), "Data revision with moving average seasonal adjustment process" J. Econometrics, vol. 14 pp. 95-114.
9. TIAO, G.C., HILLMER S.C., (1978): "Some consideration of decomposition of a time series" Biometrika, vol. 65 pp. 497-502.
10. VITALE, C., (1980): "Considerazioni critiche sulla scomposizione dei modelli ARIMA" Rivista di Stat. Applicata, vol. 13 pp. 91-104.
11. VITALE, C., (1982a): "La non invertibilità nei modelli ARIMA" Rivista di Statistica Applicata, vol. 15 pp. 253-270.
12. VITALE, C., (1982b): "La funzione di autocovarianza inversa nella decomposizione di modelli ARIMA" Statistica, vol. XLII pp.223-241.
13. VITALE, C., (1983): "Regioni di decomposizione e rappresentazioni del modello Holt-Winters" Note Economiche, N. 5-6 pp. 128-148.
14. WALLIS, K.F., (1982): "Seasonal adjustment and revision of current data: linear filters for the X-11 method" JRSS, A, vol. 145 pp. 76-85. Employment, Income Distribution and Development Strategy, Macmillan, London, 26-42.