

SCOMPOSIZIONE DI SERIE STORICHE ECONOMICHE
MEDIANTE UNA METODOLOGIA BAYESIANA

Claudia Scarani

Marzo 1988

N. 45

SCOMPOSIZIONE DI SERIE STORICHE ECONOMICHE MEDIANTE UNA
METODOLOGIA BAYESIANA

Claudia Scarani

Riassunto

In questo lavoro si presentano alcuni risultati ottenuti mediante un programma di destagionalizzazione che utilizza un modello lineare bayesiano.

Lo schema teorico, proposto da Akaike (1980), consiste nella risoluzione di un problema di minimi quadrati e nella minimizzazione di una funzione di costo ABIC. L'implementazione dell'algoritmo nel programma SEA ha tenuto conto, nella risoluzione del problema di minimi quadrati, delle strutture del modello, rendendo così possibile una drastica riduzione del numero di calcoli e di occupazione di memoria. L'algoritmo utilizzato e il confronto con altri algoritmi basati sulla stessa metodologia si trovano nel lavoro di Scarani (1986).

Di seguito vengono esposti alcuni risultati relativi alla convergenza del modello.

Il programma è stato utilizzato per la destagionalizzazione di sei serie monetarie dell'economia italiana ed è stato effettuato un confronto con i risultati ottenuti dall'applicazione del programma X11ARIMA, Dagum (1980).

Posizione del problema.

Si suppone che una serie temporale $\{y_i\}$, $i=1,2,\dots,N$, che presenta stagionalità di periodo p (cioè $p=4$ se la serie è trimestrale, $p=12$ se è mensile) possa essere rappresentata come $y_i = t_i + s_i + i_i$ dove t_i è la componente di lungo periodo, s_i la componente stagionale e i_i la componente irregolare che si suppone $NID(0, \sigma^2)$. Quindi si ipotizza un modello lineare per il graduale cambiamento della componente stagionale e di quella a lungo periodo. Inoltre si impone il vincolo che la somma dei coefficienti stagionali di un anno sia nulla. Queste informazioni sono immagazzinate nella matrice D che risulta rigidamente strutturata in blocchi relativi ai diversi vincoli. Vedi Corradi e Scarani (1984) per i dettagli su D .

Si risolve poi il seguente problema di minimi quadrati:

$$\min f(a) = \min \| D(a-\hat{a}) \|^2 + q^2 \| Xa - y \|^2 \quad (1)$$

dove a è il vettore incognito delle componenti da stimare $a = (s_N, t_N, \dots, s_1, t_1)'$, \hat{a} è una stima iniziale per a , D è la matrice del modello di dimensione $3N \times 2N$, X è una matrice $N \times 2N$ che presenta $x_{i,j}=1$ per $j=2*i-1$ o $j=2*i$, $=0$ altrove.

Ulteriori parametri, talvolta detti superparametri, sono q , rigid, zersum. Gli ultimi due pesano i blocchi relativi alla componente stagionale della matrice D . Il parametro q permette di bilanciare l'adattamento al modello e ai dati. Esso viene

Le matrici L_i , $i=1,2,\dots,n$ e W_i , $i=1,2,\dots,n-1$, sono matrici triangolari inferiori di dimensione $l \times l$. Le matrici L_i sono invertibili.

Il sistema $Pa = b$ viene risolto in due passi:

$$Lz = b \quad \text{dove} \quad z' = [z_1', z_2', \dots, z_n'] \quad \text{e poi} \quad L'a = z.$$

Il calcolo di ABIC è immediato osservando che il determinante di P si ottiene dal quadrato del prodotto degli elementi diagonali di L perciò

$$\ln |P| = \ln |D'D + q^2 X'X| = 2 \sum_{i=1}^n \sum_{j=1}^{l-1} \ln [L_i(j,j)]$$

e il logaritmo del quadrato della norma dei residui risulta

$$\ln \|f(a_*)\|^2 = \ln [(D\hat{a})'(D\hat{a}) + q^2 y'y - z'z]$$

(vedi Lawson and Hanson, 1974, pag. 122). La fattorizzazione di P e il calcolo di z non sono eseguiti in un unico passo, ma attraverso n fattorizzazioni di sottomatrici di dimensione $l+1$ e la risoluzione di n sistemi lineari. La matrice L_{i+1} si ottiene dalla fattorizzazione di

$$C - W_i'W_i = C - B'(L_i L_i')^{-1}B$$

e dipende da L_i . L' algoritmo fornisce la soluzione di (1) a un costo di $N(l^2+5l+8)$ operazioni per q , rigid e zersum fissati. Per i dettagli dell' algoritmo si veda Scarani (1986).

Convergenza dell'algoritmo.

Si può notare che la matrice P , di dimensione $2N \times 2N$, è indipendente dai valori assunti dai dati, ma non dalla loro numerosità. La fattorizzazione di P fornisce le matrici L_i , W_i che per $i > \tau$, dove τ è un indice che varia con i valori assunti dai superparametri, convergono alle matrici $L.$, $W.$. Questo fatto è stato verificato empiricamente, calcolando per quale indice $\tau = \tau(q, \text{rigid}, \text{zersum})$ le matrici L_τ e $L_{\tau-1}$ hanno tutti gli elementi che differiscono in modulo meno di $\epsilon = 10^{-6}$. Nella figura 1 si può vedere come varia l'indice di convergenza τ per q fissato e relativo ad un modello di destagionalizzazione trimestrale con $k=2$ e $m=1$. Quando la convergenza è raggiunta non è più necessario calcolare L_i e W_i , ma è sufficiente ricavare z_i da

$$L \cdot z_i = b_i - W \cdot z_{i-1} \quad \text{per } i = \tau, \tau+1, \dots, n. \quad \text{Per serie}$$

che presentano un numero di osservazioni elevato, raggiunta la convergenza, si ha una riduzione ulteriore delle operazioni e dell'occupazione di memoria poichè le matrici L_i e W_i sono calcolate e memorizzate solo per $i=1, 2, \dots, \tau$. Empiricamente si è trovato che la convergenza è raggiunta spesso in meno di 15 iterazioni.

Applicazioni a serie economiche.

Il programma di calcolo SEA è stato utilizzato per destagionalizzare alcune serie monetarie che presentano un

andamento stagionale alquanto variabile nel tempo. Inoltre è stato effettuato un confronto con la destagionalizzazione ottenuta utilizzando il metodo X11ARIMA, vedi Dagum (1978) per i dettagli sul metodo. Il confronto con il metodo X11ARIMA viene proposto poichè esso è particolarmente diffuso. Le serie monetarie trimestrali prese in esame sono:

CIRP, circolante del pubblico, dati 116 rilevati dal 1958.1 al 1986.4;

BMT, base monetaria, dati 100 rilevati dal 1962.1 al 1986.4;

ACD, depositi bancari totali, dati 116 rilevati dal 1958.1 al 1986.4;

M2, moneta dell'economia, dati 68 rilevati dal 1970.1 al 1986.4;

CTINS, credito totale del settore non statale, dati 52 rilevati dal 1974.1 al 1986.4;

ACIMPI, impieghi bancari, dati 116 rilevati dal 1958.1 al 1986.4.

La fonte comune è il Bollettino Statistico della Banca d'Italia e si tratta di consistenze in miliardi di lire.

Le serie possono essere divise in due gruppi in base al loro comportamento. Il circolante CIRP, la base monetaria BMT, la moneta dell'economia M2 e i depositi bancari ACD presentano il quarto trimestre fortemente stagionale a causa del maggior fabbisogno di liquidità per circolante e base monetaria e per l'accREDITAMENTO degli interessi di fine d'anno per i depositi bancari; la moneta dell'economia ha un andamento analogo poichè contiene il circolante e i depositi bancari. Il secondo gruppo è formato dal credito CTINS e dagli impieghi

bancari ACIMPI; queste serie presentano una forte accidentalità dovuta a provvedimenti amministrativi avvenuti dal 1974 in poi; inoltre, anche in questo caso, la stagionalità del quarto trimestre è elevata per le numerose operazioni bancarie di fine d'anno.

Per la destagionalizzazione con il metodo di Akaike si è ipotizzato un modello in cui la componente di lungo periodo è soggetta al vincolo $k=2$ e quella stagionale al vincolo $m=1$.

Il confronto fra le componenti ottenute con i due metodi permette di affermare che le componenti di lungo periodo sono molto simili fra loro, mentre la componente stagionale stimata con il metodo X11ARIMA è più rigida di quella ottenuta con il metodo di Akaike; la componente irregolare ottenuta con X11ARIMA ha varianza superiore all'altra.

Nelle figure 2,3,4,5,6,7 sono riportati i grafici delle serie originali e delle componenti stagionali ottenute con i due metodi citati.

Nella tabella 1 si trovano media (μ), deviazione standard (σ) e primo coefficiente di autocorrelazione (ρ) delle componenti irregolari.

Conclusioni.

L'analisi dei risultati, simile per le sei serie, porta a valutare positivamente la destagionalizzazione ottenuta con il metodo di Akaike poichè la componente residua ha media più vicina al valore atteso zero, varianza inferiore e coefficiente di autocorrelazione più vicino al valore nullo

della componente ottenuta con il metodo X11ARIMA. Il test di Ljung-Box, calcolato sulle prime diciotto autocorrelazioni, fornisce valori simili per entrambi i metodi.

Un giudizio globale sulla destagionalizzazione di una serie temporale dovrebbe tener conto non solo del comportamento casuale della componente irregolare, ma anche di altri criteri quali ad esempio la stabilità dei coefficienti, vedi Bacchilega e Gambetta (1984) e Gambetta (1985). Tuttavia è opportuno notare che il metodo di Akaike fornisce risultati migliori del metodo X11ARIMA per quanto riguarda il comportamento della componente irregolare, criterio per cui questo metodo viene privilegiato su altri.

Si ringrazia la Associazione Prometeia e in particolare la dott.ssa Chiara Fornasari per aver fornito le serie economiche prese in esame e le componenti di lungo periodo e stagionale ottenute con il metodo X11ARIMA.

Bibliografia

- Akaike, H. (1980), Seasonal Adjustment by a Bayesian Modelling, J. Time Ser. Anal. 1,1-13.
- Bacchilega, G. e G. Gambetta (1984), Appunti per un confronto empirico fra diversi metodi di depurazione stagionale. Progetto DESEC, promosso dal Centro di Specializzazione e Ricerche Economiche Agrarie per il Mezzogiorno, Portici, Univ. di Napoli R. S. 16/1984.
- Dagum, E. B. (1978), Modelling, Forecasting and Seasonally Adjusting Economic Time Series with the X-11 ARIMA Method, The Statistician, 27, 203-216.
- Gambetta, G. (1985), Relazione finale della fase sperimentale del progetto DESEC. Progetto DESEC, promosso dal Centro di Specializzazione e Ricerche Economiche Agrarie per il Mezzogiorno, Portici, Univ. di Napoli, R. S. 38/1985.
- Corradi, C. e C. Scarani (1984), A Note on the Computation of the Bayesian Decomposition of a Time Series, J. Time Ser. Anal. 5, 205-212.
- George, A. (1981), Direct Solution of Sparse Positive Definite System; Some Basic Ideas and Open Problems, in "Sparse Matrices and their Uses", Duff ed., Academic Press, London, 283-306.
- Golub, G.H. and C. F. Van Loan (1983), Matrix Computation, the Johns Hopkins Univ. Press, Baltimore, Maryland.

Lawson, C. L. e R. J. Hanson (1974), Solving Least Squares Problems, Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Scarani, C. (1986), An Efficient Algorithm for Time Series Decomposition. COMPSTAT 1986, Physica Verlag, Heidelberg for IASC, 399-404.

Tabella 1

	ACD	ACIMPI	BMT	CIRP	CTINS	M2
μ	-.47	-.53	-2.45	-.28	-9.17	-4.32
σ	196.65	778.73	281.35	35.92	286.96	287.53
ρ	-.02	-.32	-.22	-.29	-.41	-.17
μ'	47.04	24.44	19.76	4.28	4.37	85.25
σ'	1056.62	973.72	390.63	130.11	1122.14	1639.47
ρ'	-.22	-.52	-.33	-.52	-.59	-.37

Tab. 1. $\mu, \sigma, \rho, (\mu', \sigma', \rho')$ sono media, deviazione standard, coefficiente di autocorrelazione del primo ordine della componente irregolare $i = y - t - s$ ottenuta col programma SEA e con X11ARIMA (entro parentesi) .

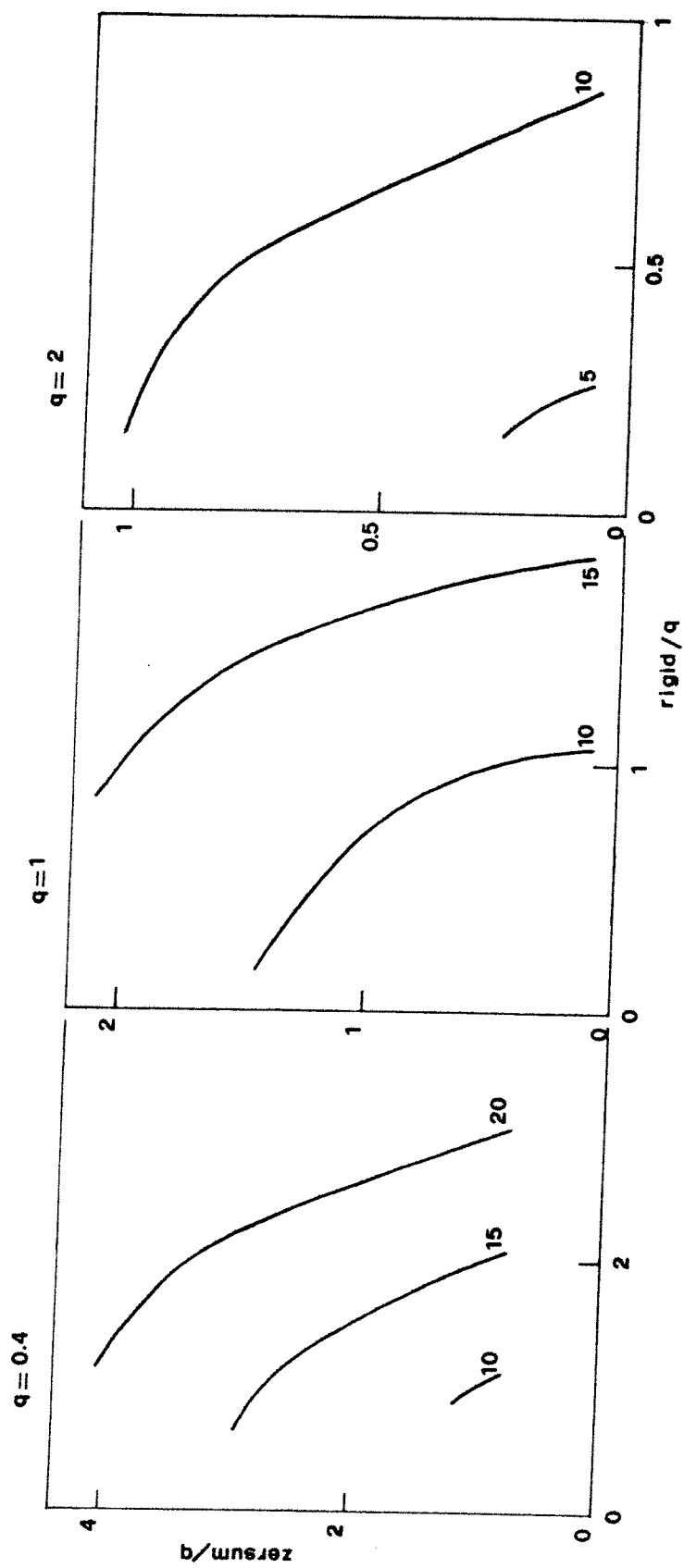


Figura 1. Indice di convergenza r .

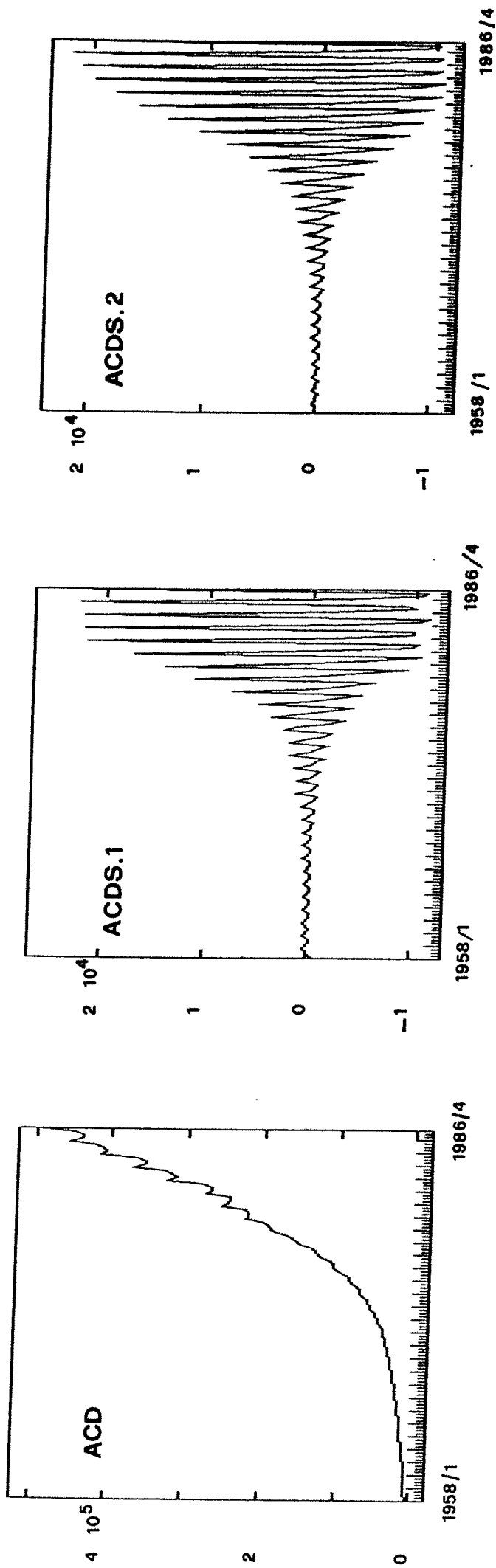


Figura 2. ACD serie grezza.
 ACDS.1 comp. stag. ottenuta con SEA.
 ACDS.2 comp. stag. ottenuta con X1LARIMA.

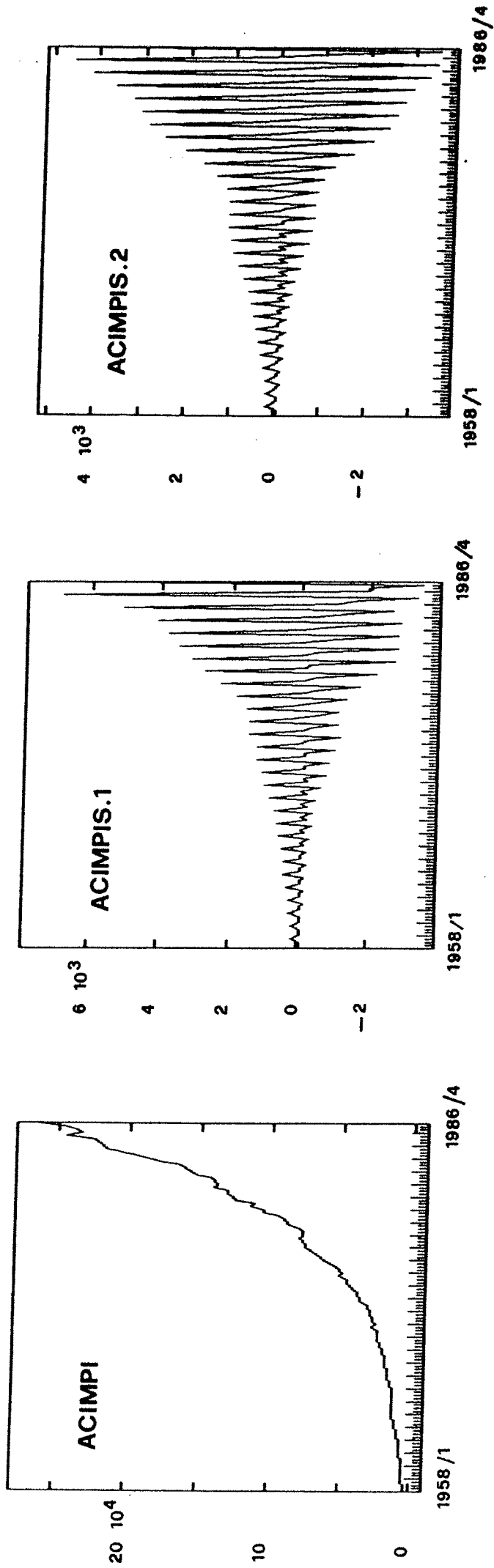


Figura 3. ACIMPI serie grezza.
 ACIMPIS.1 comp. stag. ottenuta con SEA.
 ACIMPIS.2 comp. stag. ottenuta con X11ARIMA.

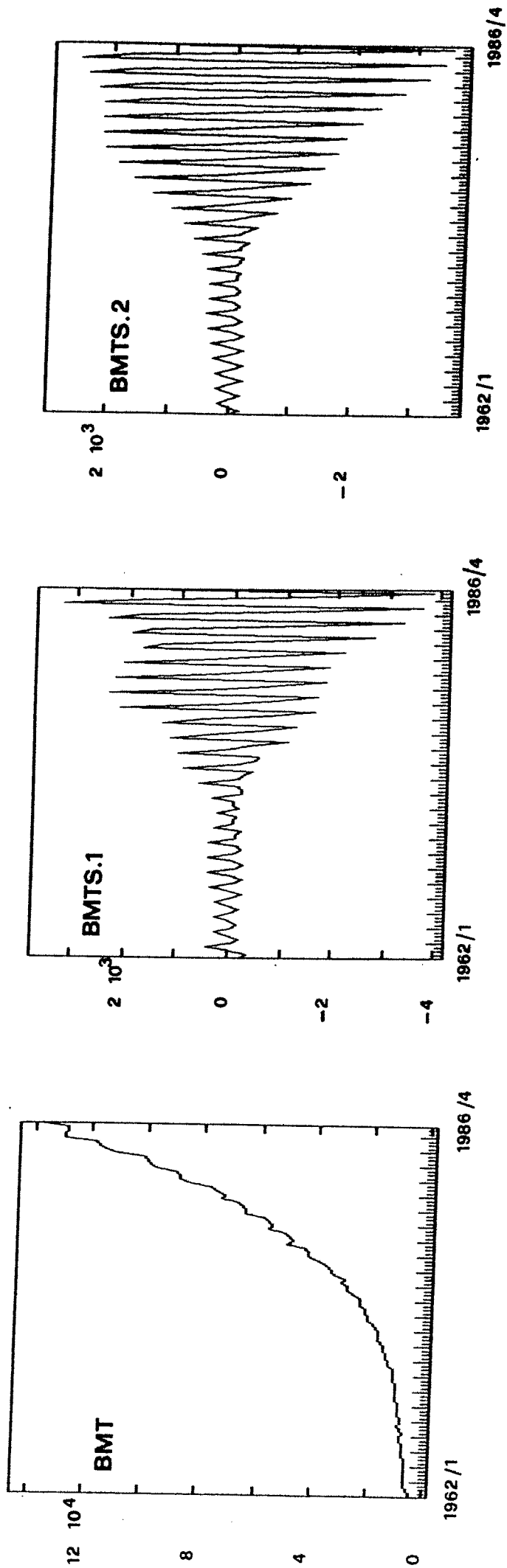


Figura 4. BMT serie grezza.
 BMTS.1 comp. stag. ottenuta con SEA.
 BMTS.2 comp. stag. ottenuta con X1LARIMA.

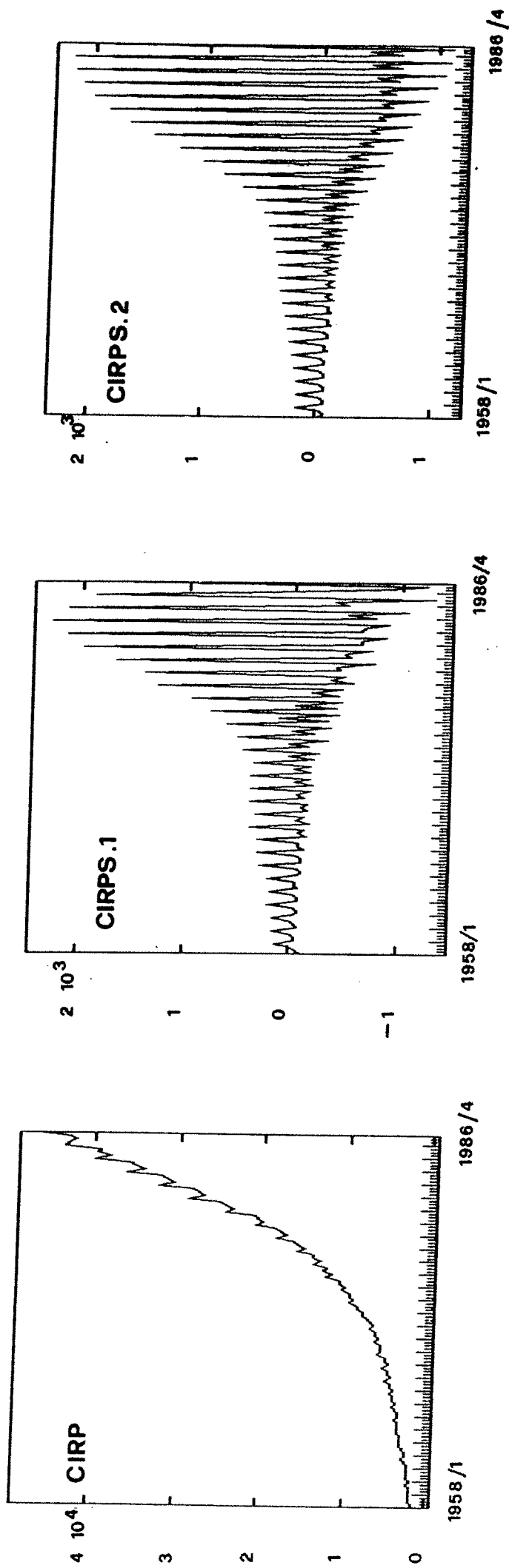


Figura 5. CIRP serie grezza.
 CIRPS.1 comp. stag. ottenuta con SEA.
 CIRPS.2 comp. stag. ottenuta con X11ARIMA.

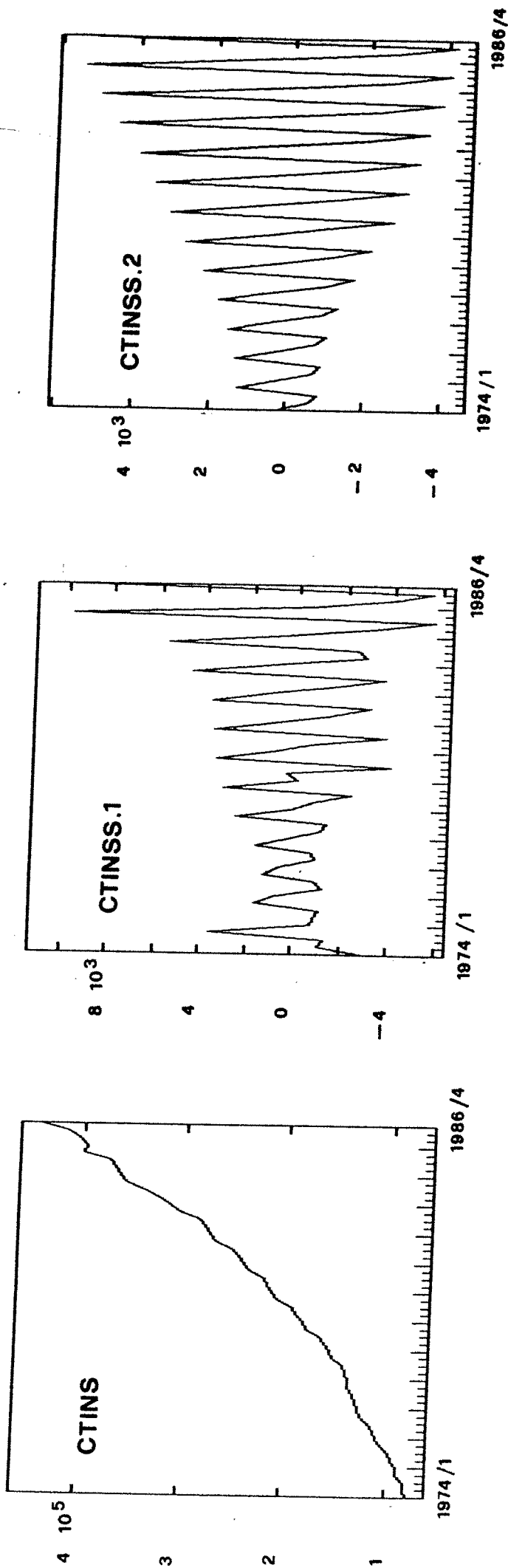


Figura 6. CTINS serie grezza.
 CTINSS.1 comp. stag. ottenuta con SEA.
 CTINSS.2 comp. stag. ottenuta con XILARIMA.

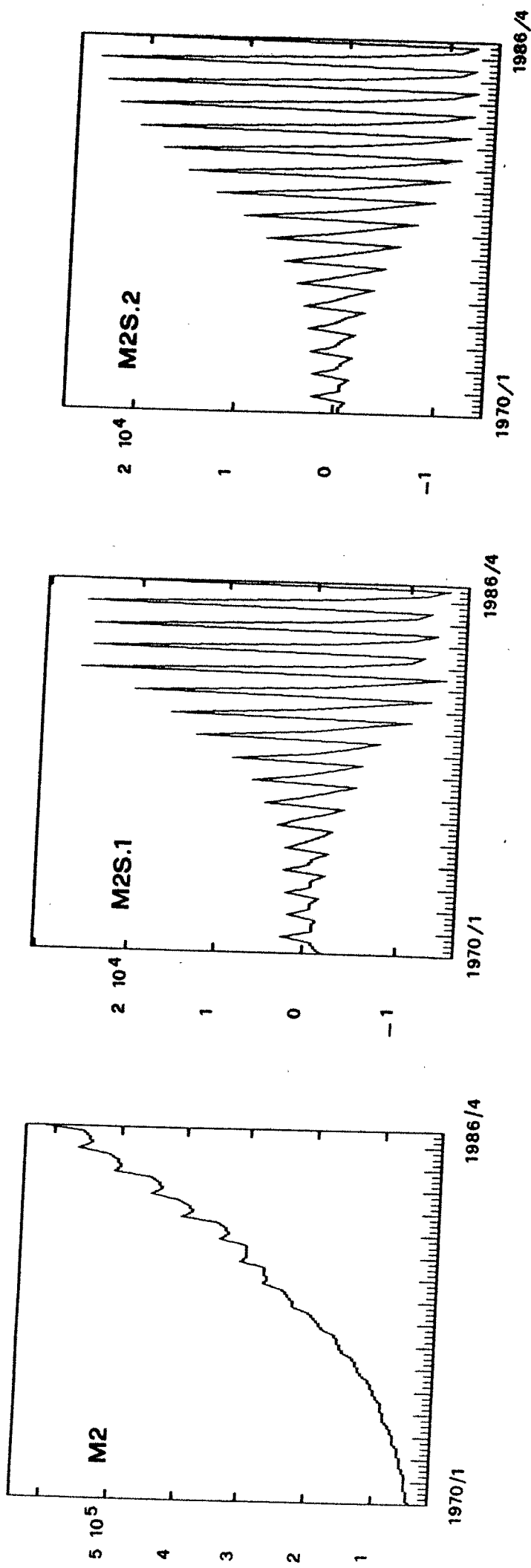


Figura 7. M2 serie grezza.
 M2S.1 comp. stag. ottenuta con SEA.
 M2S.2 comp. stag. ottenuta con X11ARIMA.