

Gruppo di lavoro per la definizione di standard per l'Istat

Istituito con delibera n. 5/DGEN del 10 gennaio 2014 e successive n. 36/DGEN del 31 marzo 2014, n. 98/DGEN del 29 ottobre 2014 e 27/DGEN del 25 febbraio 2015.

Destagionalizzazione di serie storiche con metodologia Arima model based (AMB) implementata nel software JDemetra+

P. Anzini, M. Ascione, G. Bruno, M.S. Causo, F. Ceccato, A. Ciammola, C. Cicconi, R. Corradini, D. Fantozzi, L. Fenga, C. Graziani, B. Iaconelli, R. Iannaccone, S. Loriga, C. Pappalardo, A. Spizzichino, D. Tuzi, F. Tuzi, D. Zurlo.

12 ottobre 2015

Attività: Destagionalizzazione di serie storiche

1. Sommario

Nell'ambito dell'analisi della congiuntura economica, svolta correntemente sulla base di numerosi indicatori, è ampiamente diffuso l'impiego di dati presentati in forma destagionalizzata, cioè al netto delle fluttuazioni di carattere stagionale, al fine di poter cogliere, in maniera più chiara, l'evoluzione di breve termine dei fenomeni considerati.

La stagionalità, nella dinamica di una serie storica, è quella componente che si ripete ad intervalli regolari ogni anno, con variazioni di intensità più o meno analoga nello stesso periodo (mese, trimestre, semestre, etc.) di anni successivi e di intensità diversa nel corso di uno stesso anno. Ne rappresentano degli esempi tipici il calo della produzione industriale nel mese di agosto, in seguito alla chiusura per ferie di molte aziende, e l'aumento delle vendite al dettaglio nel mese di dicembre, per effetto delle festività natalizie.

Le fluttuazioni stagionali, mascherando altri movimenti di interesse, tipicamente le fluttuazioni cicliche, sono spesso considerate di disturbo nell'analisi della congiuntura economica. La presenza di stagionalità rende, ad esempio, problematiche l'analisi e l'interpretazione delle variazioni osservate su una serie storica tra due periodi (mesi, trimestri) consecutivi dell'anno (cd. variazione congiunturale), essendo queste spesso influenzate in misura prevalente dalle oscillazioni stagionali piuttosto che da movimenti dovuti ad altre cause (ad es. al ciclo economico). Questi ultimi possono essere, invece, correttamente evidenziati calcolando le variazioni congiunturali sui dati destagionalizzati. Inoltre, essendo ciascuna serie storica caratterizzata da uno specifico profilo stagionale, l'impiego di dati destagionalizzati permette di comparare l'evoluzione di diverse serie storiche e trova ampia applicazione nell'utilizzo congiunto delle statistiche prodotte da diversi Paesi.

Un'altra pratica connessa alla precedente è quella di correggere i dati per la cosiddetta componente di calendario, determinata dalla diversa composizione del calendario nei vari mesi o trimestri, poiché

anche tale componente contribuisce ad offuscare il segnale congiunturale di interesse. Il diverso numero di giorni lavorativi o di giorni specifici della settimana in essi contenuti, come anche il modo in cui si collocano, nei periodi messi a confronto, le festività nazionali civili e religiose, fisse e mobili, e gli anni bisestili, possono costituire una fonte di variazione di breve periodo per molte serie storiche. Tali effetti, non necessariamente analoghi tra paesi o settori, inficiano la comparabilità nel tempo dei fenomeni economici e per queste ragioni sono spesso rimossi unitamente alla componente stagionale. In molti casi, accanto ai dati destagionalizzati e corretti, vengono prodotte anche serie storiche al netto dei soli effetti di calendario.

2. Descrizione generale

Nell'ultimo ventennio la richiesta di dati più agevolmente interpretabili e comparabili, in particolare di indicatori congiunturali presentati in forma destagionalizzata, è andata crescendo, in seguito agli sviluppi compiuti in campo metodologico, all'acquisizione di strumenti informatici sempre più innovativi, nonché all'internazionalizzazione dell'economia e alla realizzazione dell'Unione monetaria europea. In questo quadro, numerosi sono stati gli sforzi compiuti a livello internazionale ed europeo per giungere ad una maggiore armonizzazione delle metodologie e delle pratiche impiegate per la destagionalizzazione. In particolare, Eurostat e la Banca centrale europea hanno contribuito a produrre le linee guida sulla destagionalizzazione (Eurostat, 2015), allo scopo di promuovere, nell'ambito del Sistema statistico europeo (SSE), le migliori pratiche sulla destagionalizzazione e sulla correzione per gli effetti di calendario.

2.1 Principali approcci alla destagionalizzazione

Generalmente, l'ipotesi sottostante alla costruzione di una procedura di destagionalizzazione è che ogni serie storica Y_t , osservata a cadenza infra-annuale (ove $t = 1, 2, \dots, T$ è un indice temporale), sia esprimibile come una combinazione delle seguenti componenti non osservabili:

1. una componente di trend T_t , che rappresenta la tendenza di medio-lungo periodo, talvolta denominata anche ciclo-trend;
2. una componente stagionale S_t , costituita da oscillazioni di periodo annuale;
3. una componente irregolare I_t , dovuta a movimenti erratici, cioè a fluttuazioni di breve periodo non sistematiche e non prevedibili.

Nell'ambito della produzione statistica ufficiale, e in linea con quanto riportato nelle linee guida europee, gli approcci metodologici più diffusi alla destagionalizzazione sono essenzialmente due:

1. Metodi di tipo *Arima model based* (AMB), sviluppati tra gli altri da Burman (1980), Box, Hillmer e Tiao (1978) e Hillmer e Tiao (1982), basati sull'ipotesi che esista un particolare modello statistico parametrico (Arima) in grado di descrivere adeguatamente la struttura probabilistica del processo stocastico generatore della serie storica osservata, essendo quest'ultima concepita come la parte finita di una particolare realizzazione di un processo stocastico. I filtri lineari utilizzati in questo approccio dipendono, conseguentemente, dalle caratteristiche della serie storica considerata. Questo tipo di approccio metodologico è adottato dalla procedura TRAMO-SEATS, sviluppata da Gómez e Maravall (1996).

2. Metodi filter based (FLB), di tipo non parametrico o semiparametrico, in cui, al contrario, la stima delle componenti avviene senza ipotizzare l'esistenza di un modello statistico rappresentante la serie analizzata ma mediante l'applicazione iterativa di una serie di filtri lineari costituiti da medie mobili centrate di diversa lunghezza. Tali procedure sono dette ad hoc, poiché i filtri adottati derivano da regole meramente empiriche piuttosto che dalla struttura probabilistica del processo stocastico che ha generato la serie. Appartengono a questo gruppo i classici metodi della famiglia X-11 (X11): dai primi X11 e X-11-ARIMA (X-11A), ai più attuali X-12-ARIMA (X-12A) (Findley et al., 1998) e X-13-ARIMA-SEATS (X-13AS) (Findley, 2005), che incorporano al loro interno numerosi miglioramenti rispetto alle precedenti versioni. Tra questi, il ricorso a modelli reg-Arima finalizzato al trattamento preliminare dei dati e a una migliore previsione della serie, che si traduce in un miglioramento dei filtri simmetrici a media mobile impiegati, e cioè, generalmente, in una maggiore stabilità dei fattori stagionali stimati. Inoltre, la procedura X-13AS incorpora, oltre alla classica metodologia FLB, anche quella AMB.

In entrambi le metodologie, la fase vera e propria di destagionalizzazione, cioè di stima e di conseguente rimozione della componente stagionale, è preceduta da una fase di pretrattamento dei dati, sostanzialmente comune ai due approcci, in cui avviene la scelta dello schema di scomposizione che lega le diverse componenti della serie storica (additiva, moltiplicativa, log-additiva, ecc.) e sono identificati ed eliminati una serie di effetti, quali i valori anomali (outlier) e quelli legati agli effetti di calendario. È su questa serie corretta preliminarmente che viene condotta, sia nell'approccio AMB che in quello FLB, la fase successiva di destagionalizzazione. Eliminando dalla serie storica originaria la componente stagionale stimata, è possibile ottenere la serie destagionalizzata (SA), che presenta normalmente un andamento meno oscillante di quella "grezza", pur mantenendo al suo interno la componente irregolare. A questa fase segue il reinserimento, nella serie SA, di alcuni elementi identificati nella fase di pretrattamento, attribuiti o al trend (ad es. i cambiamenti di livello - *level shifts* LS) o alla componente irregolare (valori anomali additivi - *additive outliers* AO); vengono invece esclusi dalla serie SA gli effetti di calendario e gli outliers stagionali.

2.2 *La situazione italiana con riferimento alla metodologia impiegata*

All'interno dell'Istat, la situazione attuale si presenta omogenea per quanto riguarda la metodologia utilizzata (AMB) per la destagionalizzazione, mentre permangono alcune differenziazioni, tra i vari domini di indagine, per quanto concerne gli applicativi impiegati per la sua implementazione. A partire dal 1999, viene infatti adottata la procedura TRAMO-SEATS, basata sulla metodologia AMB, la quale ha sostituito quella precedentemente impiegata dalla metà degli anni ottanta, X-11A. Tale scelta, coerente con quella effettuata alla fine degli anni novanta da altri Istituti nazionali di statistica e da Eurostat, ha fatto seguito alle conclusioni della commissione scientifica *Seasonal Adjustment Research Appraisal* (SARA) (Giovannini e Piccolo, 2000) che, confrontando le caratteristiche dei principali approcci alla destagionalizzazione disponibili, ha espresso una preferenza a favore della metodologia AMB.

3. **Attività collegate**

Elaborazioni successive, ad esempio estrazione del trend.

4. Fasi e sottoprocessi GSBPM (version 5.0) in cui l'attività è riconducibile

Tradizionalmente, gli Istituti nazionali di statistica provvedono a calcolare e pubblicare i dati in forma SA per una selezione di indicatori macroeconomici. Dal punto di vista operativo, la procedura di destagionalizzazione si colloca a valle del ciclo produttivo del dato. In particolare, secondo la suddivisione in fasi adottata dal *Generic Statistical Business Process Model* (GSBPM), all'interno del primo dei cinque sottoprocessi (*Prepare draft outputs*) in cui si articola la fase di analisi statistica dell'output, che segue quella di produzione dei macrodati e precede quella di diffusione.

Tale sottoprocesso prevede l'applicazione di metodologie statistiche finalizzate alla produzione di misure aggiuntive – come il calcolo di indici, di dati di ciclo-trend, di serie destagionalizzate e/o corrette per gli effetti di calendario, nonché di varie misure di qualità – che consentono di arrivare al prodotto finale, ossia l'informazione statistica da diffondere.

5. Riferimenti bibliografici

Box, G., S. Hillmer e G. Tiao (1978), «Analysis and Modeling of Seasonal Time Series». In: *Seasonal Analysis of Economic Time Series*. A cura di A. Zellner. U.S. Dept. of Commerce – Bureau of the Census, pp. 309–334.

Burman, J. P. (1980), «Seasonal Adjustment by Signal Extraction». English. In: *Journal of the Royal Statistical Society. Series A (General)* 143.3, pp. 321–337. ISSN: 00359238. URL: <http://www.jstor.org/stable/2982132>.

Eurostat (2015), *Ess Guidelines on Seasonal Adjustment*. Manuals and Guidelines. ISBN: 978-92-79-45176-8. DOI: 10.2785/317290. URL: <http://dx.doi.org/10.2785/317290>.

Findley, D. F. (giu. 2005), «Some Recent Developments and Directions in Seasonal Adjustment». In: *Journal of Official Statistics* 21.2.

Findley, D. F. et al. (1998), «New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program». In: *Journal of Business & Economic Statistics* 16.2, pp. 127–152. DOI: 10.1080/07350015.1998.10524743.

Giovannini, E. e D. Piccolo, cur. (2000), *Seasonal adjustment procedures - experiences and perspectives*. Istat. ISBN: 8845802167.

Gómez, V. e A. Maravall (1996), *Programs TRAMO and SEATS, Instruction for User (Beta Version: September 1996)*. Banco de España Working Papers 9628. Banco de España. URL: <http://ideas.repec.org/p/bde/wpaper/9628.html>.

Hillmer, S. C. e G. C. Tiao (1982), «An ARIMA-Model-Based Approach to Seasonal Adjustment». English. In: *Journal of the American Statistical Association* 77.377, pp. 63–70. ISSN: 01621459. URL: <http://www.jstor.org/stable/2287770>.

Metodo: TRAMO-SEATS (Gómez e Maravall, 1996)

Data di pubblicazione: 1996

6. Sommario

L'attività del team, finalizzata a proporre uno standard metodologico per l'Istat in relazione alla destagionalizzazione di serie storiche, ha comportato, in prima istanza, la valutazione di approcci alternativi alla destagionalizzazione e, in seconda battuta, di alcuni dei principali software che implementano tali metodologie. Per quanto riguarda il primo aspetto, l'attenzione si è focalizzata sui due principali approcci metodologici, di tipo AMB e FLB, il cui impiego viene anche incoraggiato dalle linee guida Eurostat; diverse considerazioni hanno portato a confermare la scelta metodologica, già effettuata nel 1998 dalla commissione SARA, di adottare un approccio AMB, in particolare quello incorporato nella procedura TRAMO-SEATS (*Time series Regression with Arima noise, Missing observations and Outliers e Signal Extraction in Arima Time Series - TS*), sviluppato da Gómez e Maravall (1996); inoltre, viene ammessa la possibilità di ricorrere ai metodi FLB, limitatamente a quei domini per i quali questi possano portare a risultati più soddisfacenti.

La procedura TS, appartenendo alla metodologia AMB, si basa 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.

La procedura TS si compone di due parti integrate. Con la prima, TRAMO, si sceglie il tipo di scomposizione, si interpolano eventuali dati mancanti e si eliminano i cosiddetti effetti deterministici (effetti di calendario e valori anomali); inoltre si procede all'identificazione e alla stima di un idoneo modello ARIMA per la serie depurata dagli effetti deterministici, cioè per la parte stocastica della serie originaria (nota anche come serie "linearizzata"). La seconda parte della procedura, SEATS, effettua, sulla base di opportune ipotesi, l'individuazione della componente stagionale e la sua rimozione.

7. Descrizione

Nel seguito vengono descritte in maggiore dettaglio le funzioni principali dei due moduli TRAMO e SEATS di cui si compone la procedura TS.

Dal punto di vista operativo, le fasi essenziali di TRAMO possono essere così riassunte:

1. Fase preliminare (pretrattamento della serie storica osservata) costituita, nell'ordine, da due passi:
 - a) **SCOMPOSIZIONE**. Viene valutato il ricorso ad opportune trasformazioni dei valori originari della serie, cioè individuato lo schema di aggregazione delle componenti non osservabili. La scelta del tipo di scomposizione avviene tra due tipologie:
 - additiva, ove si assume che la serie storica osservata sia ottenibile come somma delle componenti di trend, stagionale e irregolare; la serie SA sarà pertanto ottenuta sottraendo dalla serie originaria la componente relativa alla stagionalità. L'ipotesi alla

base di una relazione di tipo additivo è che le diverse componenti della serie siano tra loro indipendenti;

- log-additiva. Tale relazione consente di considerare una relazione moltiplicativa tra le componenti. Essa ipotizza una scomposizione additiva del logaritmo della serie storica originaria e pertanto la serie SA sarà ottenuta dividendo la serie originaria per la trasformata esponenziale della componente stagionale.

b) EFFETTI DETERMINISTICI. Vengono individuati e rimossi eventuali effetti deterministici la cui presenza può influire negativamente sulla corretta identificazione della componente stagionale nella parte successiva. Tra questi, i principali sono rappresentati da EFFETTI DI CALENDARIO, OUTLIER e ALTRI EFFETTI DETERMINISTICI.

Tra gli EFFETTI DI CALENDARIO vengono solitamente considerati: l'effetto giorni lavorativi, che cattura le fluttuazioni dovute alla diversa distribuzione dei giorni della settimana nei vari mesi, in cui è possibile assumere che ciascun giorno della settimana influenzi in modo diverso il fenomeno rappresentato dalla serie storica (effetto *trading-day*), o che gli effetti dei giorni dal lunedì al venerdì siano tra loro uguali e si confrontino con un effetto comune dei sabati e delle domeniche (effetto *working-day*); l'effetto festività mobili, che, come la Pasqua, cambiano data di anno in anno; l'effetto *leap-year*, che cattura gli eventuali cicli di quattro anni dovuti alla presenza di un giorno in più nel mese di febbraio degli anni bisestili; l'effetto dovuto alle festività nazionali pubbliche e religiose, la cui influenza può essere analizzata separatamente mediante l'introduzione di un regressore esterno, ovvero considerando le modifiche apportate all'effetto giorni lavorativi, cioè assimilando il giorno della settimana in cui cade una specifica festività (es. Ferragosto) ad una domenica.

La presenza di OUTLIER, ossia di osservazioni anomale nella serie storica causate da eventi straordinari, può determinare significative distorsioni nell'identificazione dell'ordine del modello reg-Arima e/o nella stima dei suoi coefficienti. TS consente di ricercare automaticamente tre tipi di valori anomali: additivi (puntuali, di durata limitata a un periodo), temporanei (che rappresentano un *temporary change*, cioè un mutamento transitorio di durata superiore ad un periodo) o permanenti (che determinano un cambio di livello, *level shift*).

Oltre ai suddetti effetti, la serie storica potrebbe essere influenzata da ALTRE COMPONENTI DETERMINISTICHE, la cui rimozione può favorire una più corretta identificazione e stima del modello Arima. Tra i principali di questi effetti, che è possibile cogliere mediante l'introduzione di regressori *ad hoc*, vanno annoverati: l'effetto rampa, che rappresenta i cambiamenti di livello permanenti che si realizzano in modo graduale nel tempo; gli outlier stagionali per modellare i cambi di livello nei fattori stagionali di una serie storica.

La rimozione delle suddette componenti, stimate utilizzando un modello di regressione (modello reg-Arima) dove tali effetti sono catturati dalla combinazione lineare di k variabili deterministiche e dai rispettivi coefficienti, darà come risultato la componente stocastica Z_t della serie storica, rappresentata da un modello $Arima(p,d,q)(P,D,Q)s$.

2. IDENTIFICAZIONE del modello $Arima(p,d,q)(P,D,Q)_s$, per la componente stocastica Z_t , che consente di specificarne gli ordini dei polinomi AR e MA per la parte non stagionale (p e q) e per la parte stagionale (P e Q) e l'eventuale relativo grado di integrazione (d e D) del modello ARIMA.

Nell'ambito di un modello reg-ARIMA, le fasi 1b) di stima degli effetti deterministici e 2) di identificazione del modello Arima, sono tra loro interdipendenti e rappresentano un aspetto molto delicato del processo di destagionalizzazione. Nei metodi AMB, come TS, è fondamentale che il modello $Arima(p,d,q)(P,D,Q)_s$ per serie Z_t , sulla quale verrà successivamente stimata la componente stagionale, sia correttamente specificato, in quanto questo concorre direttamente alla stima della componente stagionale della serie. Le principali procedure disponibili per la destagionalizzazione sono dotate di routine per l'identificazione automatica sia dei principali effetti deterministici che del modello Arima. Tale funzionalità è molto utile soprattutto nei casi in cui si debba trattare un elevato numero di serie storiche, ma è consigliabile analizzarne attentamente i risultati e intervenire se necessario per modificarli.

TS consente di seguire, infatti, relativamente all'individuazione degli ordini dei polinomi AR ed MA stagionali e non stagionali, due strade: quella automatica, basata sul confronto di vari modelli utilizzando criteri standard sviluppati in letteratura (in particolare, valutando le statistiche espresse dagli indici BIC e AIC); quella manuale, che richiede una conoscenza più approfondita della modellistica ARIMA, in cui è l'utente a specificare gli ordini del modello.

3. STIMA del modello ARIMA identificato, utilizzando il metodo della massima verosimiglianza;
4. VALIDAZIONE DEL MODELLO, basata sull'elaborazione di test statistici per verificare se il modello individuato sia una buona rappresentazione del processo stocastico che ha generato la serie osservata.

Le tre fasi (identificazione, stima e verifica delle ipotesi) di cui consta la costruzione di un modello ARIMA (cfr. Box e Jenkins, 1970), possono essere iterate finché non si perviene al modello ritenuto migliore. Le serie così trattate vengono, quindi, sottoposte a SEATS in cui, in estrema sintesi, vengono compiute le seguenti operazioni:

1. VERIFICA del modello individuato da TRAMO ed eventuale nuova identificazione e stima del modello ARIMA;
2. IDENTIFICAZIONE dei modelli ARIMA per le componenti della serie linearizzata Z_t , coerenti con quello seguito dalla serie Z_t . La scomposizione avviene attraverso la fattorizzazione dello spettro della serie storica Z_t , imponendo come condizione identificativa la cosiddetta scomposizione *canonica* (Hillmer e Tiao, 1982) che, tra le scomposizioni ammissibili, seleziona quella che assegna la maggior quota di varianza alla componente irregolare. In questo modo si perviene alla componente stagionale il più possibile stabile, date le caratteristiche stocastiche della serie Z_t .
3. STIMA delle componenti e ANALISI DEGLI ERRORI. Identificati i modelli ARIMA per ciascuna componente, se ne derivano i relativi stimatori applicando un opportuno filtro lineare alla serie Z_t . SEATS utilizza, in particolare, il filtro di Wiener-Kolmogorov, che è infinito, simmetrico e convergente. Nella pratica, ovviamente, si utilizzano filtri "troncati", in quanto si dispone di una realizzazione finita per la serie Z_t ; pertanto, se T è l'istante temporale corrispondente

all'osservazione più recente, a seconda dell'istante temporale t in base al quale si calcola una stima della generica componente, lo stesso stimatore può essere definito: *preliminare* se $t < T$ e vicino a T , *concurrent* se $t = T$ e *finale* se t cade nella parte centrale di una serie storica abbastanza lunga (almeno 10 anni). Per ciascuno stimatore, di conseguenza, sarà possibile valutare rispettivamente l'errore di stima preliminare, l'errore di revisione e l'errore di stima finale.

Sottolineiamo ancora come, nell'approccio TS, la scelta del filtro ottimale si riconduca all'identificazione ottimale del modello ARIMA, la cui corretta specificazione, operata nella fase di pretrattamento, è un'operazione fondamentale, in quanto condiziona in maniera determinante le fasi successive.

4. STIMA FINALE che prevede l'allocatione, tra le varie componenti, delle parti deterministiche individuate da TRAMO. Vengono, ad es., attribuiti al trend i *level shift* e le rampe, alla componente irregolare gli outlier additivi e temporanei. Verranno invece esclusi dalla serie SA gli effetti di calendario e gli outlier stagionali.

SEATS offre l'opportunità di valutare i risultati raggiunti sia attraverso una diagnostica molto ampia sia tramite un'analisi grafica delle componenti non osservabili.

8. Ambito di applicazione

La procedura TS viene impiegata per rimuovere la componente stagionale di una serie storica, ottenendo la relativa serie destagionalizzata. Inoltre, il metodo è impiegato anche per ottenere serie corrette solo per gli effetti di calendario.

8.1 Obiettivo del metodo

L'obiettivo che si intende perseguire con l'applicazione del metodo TS di destagionalizzazione è quello di ottenere indicatori che consentano di cogliere in maniera più chiara l'evoluzione di breve termine dei fenomeni considerati.

8.2 Indicazioni relative all'applicabilità del metodo

Per una trattazione dettagliata delle raccomandazioni relative all'uso ottimale del metodo, si rimanda alle sezioni 2 e 6 delle linee guida Eurostat (2015). Le principali indicazioni qui segnalate attengono ad alcuni casi problematici che possono sorgere nella pratica corrente della destagionalizzazione, effettuata con il programma TS, e che possono richiedere una valutazione più attenta da parte dell'utente. Tra questi citiamo:

1. La lunghezza delle serie storiche da destagionalizzare (serie troppo brevi/lunghe)
 - a) Sebbene il programma TS sia capace di trattare serie storiche brevi, in cui si dispone anche di sole 12 o 36 osservazioni (a seconda che la frequenza sia trimestrale o mensile), la lunghezza delle serie da destagionalizzare dovrebbe essere tale da consentire una stima ragionevolmente stabile dei coefficienti. Pertanto, nella pratica corrente, è preferibile che la serie storica abbia una lunghezza minima di 6-7 anni ed è opportuno prestare molta

cautela nell'analizzare i risultati della destagionalizzazione nei casi in cui si dispone di un numero inferiore di osservazioni (si veda la sezione 6.1 delle succitate linee guida).

- b) Se da un lato esiste la necessità di utilizzare serie che contengano un adeguato numero di osservazioni, dall'altro il trattamento di serie storiche troppo lunghe può far sorgere il problema della scelta dell'intervallo temporale su cui basare l'analisi. Serie troppo lunghe sono infatti spesso caratterizzate da una forte eterogeneità, dovuta ad eventuali cambiamenti di base e concatenamenti (nel caso di numeri indici), a cambiamenti nella rilevazione, a ricostruzioni dei dati, etc., rendendo difficile l'individuazione di un modello ARIMA soddisfacente. Difficoltà di questo tipo possono rendere preferibile non condurre la destagionalizzazione sull'intero periodo (si veda il paragrafo 6.2 delle linee guida Eurostat).
2. Alcuni accorgimenti da adottare si riferiscono a situazioni che possono condizionare in maniera determinante il pretrattamento delle serie storiche, un'operazione che, come già evidenziato, riveste un ruolo fondamentale nella procedura completa di destagionalizzazione. Grande cautela bisogna prestare, ad es., ai seguenti casi:
- a) Identificazione automatica di valori anomali alle estremità della serie, in particolare in corrispondenza degli ultimi anni della serie, da cui potrebbe scaturirne un'informazione congiunturale fortemente distorta (si veda la sezione 2.8 delle linee guida Eurostat 2015);
 - b) Eventuale identificazione automatica di un numero eccessivo di outlier, che potrebbe indicare l'opportunità di procedere ad una diversa specificazione del modello ARIMA. Alcuni accorgimenti adottati per ridurre il numero consistono nell'abbreviare l'intervallo temporale di ricerca degli stessi e/o nell'aumentare il valore soglia dei test statistico impiegato nella procedura di identificazione adottata (Chen e Liu, 1993).

Infine, una pratica fortemente raccomandata, basilare sia per evidenziare eventuali aspetti problematici che potrebbero sorgere nell'applicazione della procedura TS che per individuare le caratteristiche fondamentali dei dati, è quella di esaminare graficamente la serie grezza nella fase preliminare e prima di effettuare qualunque altra analisi. L'alternativa raccomandata (si veda la sezione 2.2 delle succitate linee guida) è di utilizzare sia strumenti grafici di base, sia strumenti più sofisticati (come lo spettro e il correlogramma), almeno una volta l'anno, per i principali indicatori economici, documentandone i risultati. Si consiglia inoltre l'esame degli stessi grafici applicato ai residui del modello reg-Arima.

8.3 *Limitazioni e inconvenienti del metodo*

Tra i possibili effetti indesiderabili risultanti dall'applicazione del metodo, va annoverata la sua applicazione su serie storiche non stagionali. È pertanto necessario verificare la significatività della componente stagionale prima di procedere alla sua eliminazione. A tale scopo possono essere impiegati vari test statistici presenti nei principali pacchetti software.

9. Modalità di applicazione

9.1 Dati di input

I dati richiesti in input dalla procedura TS sono serie storiche di dati quantitativi aventi frequenza infra-annuale ed equi-spaziati. È possibile gestire un numero, limitato, di valori mancanti.

Oltre ai nomi delle serie da trattare e alle relative osservazioni in sequenza, le modalità in cui organizzare le informazioni da fornire in input per l'applicazione della procedura TS (sintassi dei comandi, parametri da inserire, eventuali regressori specificati dall'utente e forniti in data-set separati, etc.) variano a seconda dello specifico strumento IT impiegato per l'implementazione del metodo TS. Per tali dettagli applicativi, pertanto, occorre far riferimento ai manuali relativi al particolare software scelto a tale scopo (si veda Grudkowska, 2015b).

9.2 Requisiti relativi ai dati di input

Le uniche condizioni poste dal metodo sui dati di input sono:

1. La frequenza dei dati che si richiede sia infra-annuale (tipicamente, mensile o trimestrale) ed uniforme sull'intera serie;
2. La lunghezza minima delle serie storiche che deve essere tale da coprire almeno 3 anni (12 o 36 osservazioni a seconda che si tratti di dati a cadenza trimestrale o mensile).

9.3 Parametri da definire

Per una trattazione dettagliata delle numerose opzioni occorre fare riferimento alla manualistica del software.

9.4 Dati di output

La procedura TS fornisce, separatamente per i due programmi integrati di cui il metodo si compone, un output molto ampio, prodotto sia in forma dettagliata per ciascuna serie storica trattata, sia in forma sintetica per l'insieme di serie storiche trattate contemporaneamente. Tale output include serie storiche (es.: serie originaria, linearizzata, componenti stocastiche, componenti finali, tassi di crescita, residui, effetti deterministici, etc.) e vari strumenti diagnostici riferiti alle diverse fasi in cui si articola la procedura (sia test statistici che strumenti grafici, di base e complessi).

Più nel dettaglio, l'attuale output di TRAMO contiene:

- informazioni relative al modello (numero di osservazioni, livelli o eventuale trasformazione logaritmica della serie, ordine del modello ARIMA);
- diagnostica sui residui (*standard error* dei residui, BIC, test Q di Ljung-Box-Pierce per l'autocorrelazione dei residui, test di Bowman-Shenton per la normalità dei residui, t -value del coefficiente di asimmetria e della curtosi dei residui, test QS di Pierce per l'autocorrelazione dei residui in corrispondenza dei ritardi stagionali, t -test per i *runs* dei segni dei residui);
- informazioni sulla capacità previsiva del modello (viene effettuato un test F sulle previsioni *out-of-sample* confrontando la varianza degli errori di previsione *out-of-sample* a un passo con la varianza dei residui *in sample*. La procedura prevede la stima del modello sulla serie storica

senza gli ultimi due anni di osservazioni, il calcolo degli errori di previsione a un passo, tenendo fissi i parametri del modello, e il calcolo della statistica test).

L'attuale output di SEATS contiene:

- informazioni generali sulla scomposizione (approssimazione del modello qualora non esista una scomposizione ammissibile per il modello originario, "fattorizzazione spettrale" che restituisce la scomposizione del modello, check su ACF basato sul confronto delle varianze tra componente teorica, stimatore teorico e stima empirica, assegnazione delle radici autoregressive alla componente transitoria, di trend, o alla componente stagionale, assegnazione delle componenti deterministiche stimate in Tramo alle componenti stocastiche di Seats);
- analisi degli *standard errors* (*standard deviation* delle innovazioni delle componenti, *standard error* dello stimatore *concurrent* per CT e SA, *standard error* delle revisioni dello stimatore *concurrent* per CT e SA, *standard error* dei tassi di crescita delle componenti stimate CT e SA espressi in punti percentuali);
- informazioni relative alle proprietà della scomposizione (convergenza dopo 1 anno e dopo 5 anni per CT e SA espressa come riduzione percentuale della varianza dell'errore di revisione dello stimatore *concurrent* dopo 1 e 5 anni di ulteriori osservazioni, significatività della stagionalità su tre diversi periodi, media delle differenze assolute tra le medie annuali delle serie originaria, SA e del CT);
- numerosi grafici distinti in: serie storiche, pseudo-spettri (componenti teoriche e stimatori teorici), autocorrelogrammi (serie differenziata, residui, componenti teoriche, stimatori teorici e componenti stimate), filtri (dominio temporale e delle frequenze), previsioni con intervalli di confidenza (serie originaria, componenti deterministiche, componenti stocastiche e componenti finali), effetti deterministici (di calendario, outlier e altri effetti specificati dall'utente).

9.5 *Proprietà dei dati di output*

Le serie destagionalizzate fornite in output devono avere la stessa lunghezza delle serie storiche fornite in input, essere prive di dati mancanti e caratterizzate da un profilo più regolare rispetto alle serie originarie.

9.6 *Indicatori di qualità dei dati di output*

Per valutare la qualità della destagionalizzazione assumono rilievo le diagnostiche relative all'adeguatezza del modello ARIMA utilizzato (i più importanti test impiegati a tal fine sono quelli relativi all'assenza di autocorrelazione dei residui, alla normalità della distribuzione dei residui e all'assenza di autocorrelazione stagionale dei residui) e quelle relative alla qualità della scomposizione (per ciascuna componente stimata: errore di stima, errori di previsione, varianza delle revisioni, etc.). Inoltre è essenziale verificare l'assenza di stagionalità nella serie di output. Per una trattazione più completa si vedano la sezione 5 delle linee guida Eurostat (2015) e Eurostat (2014).

9.7 Interazione con l'utente

Per l'applicazione del metodo è richiesto che l'utente fornisca in input un file le cui specifiche possono variare secondo lo specifico strumento IT impiegato. Indipendentemente da quest'ultimo, tuttavia, il dataset di input dovrà contenere, per ciascuna serie storica, il nome della serie, i suoi valori con i relativi riferimenti temporali e i parametri scelti per effettuare la destagionalizzazione. Ove necessario, l'utente dovrà fornire informazioni aggiuntive, quali regressori esterni.

Dopo l'esecuzione del programma di destagionalizzazione l'utente avrà a disposizione un ampio output per ciascuna serie trattata, costituito, in particolare, dalla serie destagionalizzata e dalle altre componenti, nonché un insieme di diagnostiche. L'analisi grafica e statistica dell'output consentirà quindi di validare i risultati per la pubblicazione.

10. Riferimenti bibliografici

- Anitori, P. et al. (2000), *Guida all'utilizzo di TRAMO-SEATS per la destagionalizzazione delle serie storiche*. Documenti 4. Istat. URL: http://www3.istat.it/dati/pubbsci/documenti/Documenti/doc_2000/manuale.zip.
- Box G. E. P., Jenkins G. M. (1970), *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco.
- Chen C., Liu L. (1993), *Joint Estimation of Model Parameters and Outlier Effects in Time Series* in *Journal of the American Statistical Association*, 88, 284:297.
- Eurostat (2015), *Ess Guidelines on Seasonal Adjustment*. Manuals and Guidelines. ISBN: 978-92-79-45176-8. DOI: 10.2785/317290. URL: <http://dx.doi.org/10.2785/317290>.
- Eurostat (2014), *ESS Handbook for Quality Reports*. ISSN 2315-0815. URL: <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf/18dd4bf0-8de6-4f3f-9adb-fab92db1a568>
- Gómez, V. e A. Maravall (1996), *Programs TRAMO and SEATS, Instruction for User (Beta Version: September 1996)*. Banco de España Working Papers 9628. Banco de España. URL: <http://ideas.repec.org/p/bde/wpaper/9628.html>.
- Gomez V. e Maravall A. (1998). *Guide for Using the Programs TRAMO and SEATS (Beta Version)*.
- Grudkowska, S. (2015b). *JDemetra+ Reference Manual, v. 0.1*. URL: <http://www.crosportal.eu/sites/default/files/JDemetra%2BReference%20Manual.%20Version%200.1.pdf>.
- Hillmer, S. C. e G. C. Tiao (1982), «An ARIMA-Model-Based Approach to Seasonal Adjustment». English. In: *Journal of the American Statistical Association* 77.377, pp. 63–70. ISSN: 01621459. URL: <http://www.jstor.org/stable/2287770>.
- Kaiser, R. e A. Maravall (2000a), *Notes on the time series analysis, ARIMA models and signal extraction*. Working Paper 12. Bank of Spain. URL: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosTrabajo/00/Fic/dt0012e.pdf>.

Maravall, A. et al. (2014), *New Features and Modifications in Tramo-Seats*. URL: http://www.bde.es/f/webbde/SES/servicio/Programas_estadisticos_y_econometricos/Programas/ficheros/TSWnewFeatures20141222.pdf.

Software: JDemetra+

Versione 2.0.0

Pubblicato a febbraio 2015

11. Descrizione

JDemetra+ (JD+) è un programma *open source* sviluppato dalla Banca nazionale del Belgio, con il sostegno finanziario di Eurostat, che costituisce una reingegnerizzazione in Java delle versioni più recenti che implementano le metodologie di tipo AMB e FLB, cioè della procedura TRAMO-SEATS e della procedura X-13AS. Il software, recentemente diffuso da Eurostat, è nato con l'obiettivo di sostenere gli Istituti di statistica nell'applicazione della destagionalizzazione, tenendo conto delle indicazioni presenti in Eurostat (2015).

Il programma JD+ comprende un'avanzata interfaccia grafica per l'uso delle due procedure di destagionalizzazione ed include un insieme aggiuntivo di librerie che possono essere utilizzate per effettuare specifiche analisi (come la destagionalizzazione di dataset di grandi dimensioni, la disaggregazione temporale di serie storiche, il *benchmarking* e l'analisi del ciclo economico). Le sue funzionalità, inoltre, possono essere estese mediante lo sviluppo di *plug-in* aggiuntivi, richiamabili dalle procedure di destagionalizzazione. JD+ fornisce, inoltre, strumenti di analisi analoghi e output standardizzati per i diversi approcci alla destagionalizzazione, consentendo un confronto più agevole dei risultati ad essi associati.

La diffusione di tale strumento nei processi correnti di produzione da parte degli Istituti nazionali di statistica e delle Banche centrali, viene incoraggiata da Eurostat con la costituzione del *Seasonal Adjustment Centre of Excellence* (SACE), al quale partecipano diversi Istituti nazionali di statistica, tra cui l'Istat. La versione recentemente diffusa del programma ha già subito una serie di verifiche da parte del SACE e del *Seasonal Adjustment Users Group* (SAUG) e rappresenta il software proposto da Eurostat come standard per la destagionalizzazione presso gli Istituti di statistica dei paesi membri. Per quanto riguarda le azioni intraprese a garanzia di un futuro supporto operativo e metodologico del software, sono da segnalare: la creazione di un servizio di *help-desk* gestito da Eurostat tramite il SACE; per i prossimi anni il mantenimento e lo sviluppo del software saranno gestiti dalla Banca nazionale del Belgio, per la parte AMB, e della Banca centrale tedesca, per la parte FLB; a regime dovrebbe crearsi una comunità autonoma di sviluppatori in grado mantenere, aggiornare e sviluppare i *core engines* e i *plug-in*, sotto la supervisione del gruppo permanente di esperti sulla destagionalizzazione denominato SAEG (*Seasonal Adjustment Expert Group*).

12. Requisiti per l'applicazione

Un'importante caratteristica applicativa del software JD+ è quella di essere multiplatforma; essendo il programma implementato in Java, esso risulta eseguibile sui principali sistemi operativi (Linux, Windows e Mac).

13. Riferimenti bibliografici

Findley, D. F. (giu. 2005). «Some Recent Developments and Directions in Seasonal Adjustment». In: *Journal of Official Statistics* 21.2.

Grudkowska, S. (2015a). *JDemetra+ Quick Start*. URL: http://www.cros-portal.eu/sites/default/files/JDemetra%2B%20Quick%20Start_0.pdf.

Grudkowska, S. (2015b). *JDemetra+ Reference Manual, v. 0.1*. URL: <http://www.cros-portal.eu/sites/default/files/JDemetra%2BReference%20Manual.%20Version%200.1.pdf>.

Grudkowska, S. (2015c). *JDemetra+ User Guide*. URL: <http://www.cros-portal.eu/sites/default/files/JDemetra%2B%20User%20Guide.pdf>.

Maravall, A. et al. (2014), *New Features and Modifications in Tramo-Seats*. URL: http://www.bde.es/f/webbde/SES/servicio/Programas_estadisticos_y_econometricos/Programas/ficheros/TSWnewFeatures20141222.pdf.