

Previsione e selezione automatica dei modelli per serie storiche regionali: metodo bi-fase a conciliazione esterna

La comprensione dei fattori socio-economici è elemento chiave per lo sviluppo dei sistemi più evoluti. L'analisi dei fenomeni in contesti non più solo nazionali, ma anche regionali, sta diventando un fattore di grande rilevanza in un'Europa che unisce componenti nazionali e attitudini/esigenze regionali. Tuttavia, il ritardo con cui i dati regionali sono resi disponibili non consente alle amministrazioni locali di prendere decisioni basate su informazioni recenti.

Obiettivo

Obiettivo del lavoro è la sperimentazione di una metodologia innovativa ESeC per le analisi territoriali, finalizzata a stime correnti ed affidabili della situazione economico-sociale delle Regioni d'Europa. In tal senso, oltre alla tipica strategia Bottom-Up per il Forecasting regionale, è stata implementata la più robusta Middle-Out esterna (a livello NUTS0). Si tratta di una metodologia bi-fase basata sulla stima del dato regionale, riconciliato ad un dato nazionale ottenuto esternamente al processo tramite l'applicazione di una modellistica stocastica e l'aggiunzione delle stime nazionali di periodicità maggiore.

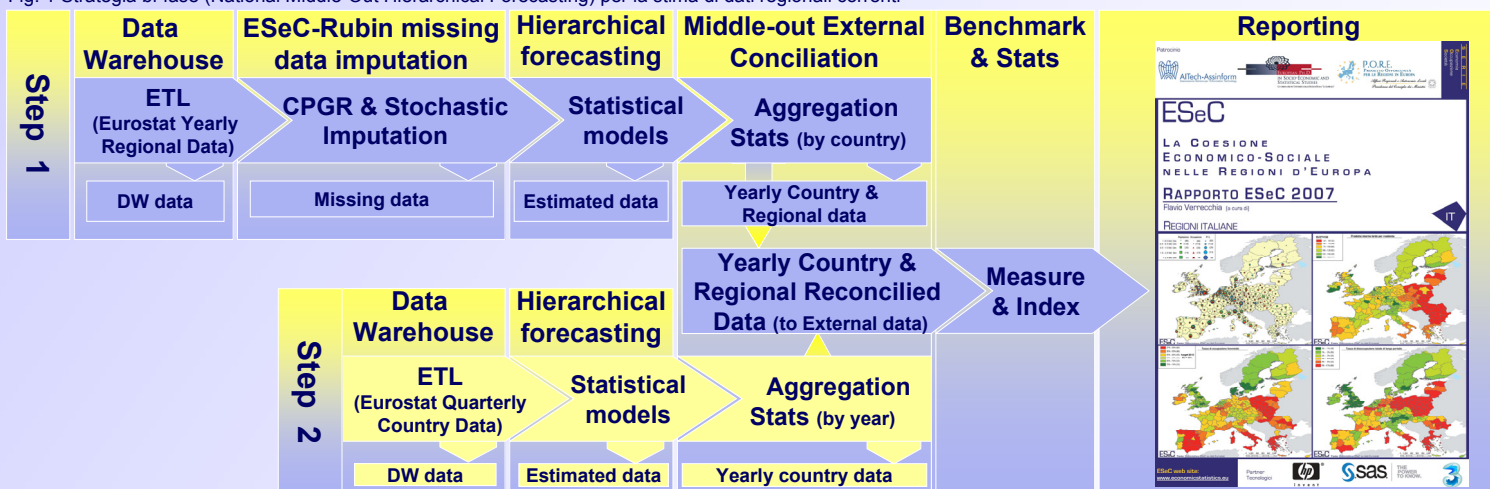
Soluzione

SAS Forecast Server, utilizzato per la gestione delle informazioni gerarchiche, consente la selezione automatica dei modelli sulla base di statistiche di Fit (e.g. MAPE). Si fa qui uso della modellistica univariata tipica dell'analisi delle serie storiche e delle tecniche di conciliazione. I dati Eurostat, disponibili per il monitoraggio dei progressi economico/sociali a livello regionale, hanno due particolarità:

- storicità limitata;
- periodicità annuale del dato.

La **strategia Bottom-up** è, per esempio, applicabile con successo alla demografia regionale dove si hanno a disposizione serie storiche più estese, più complete e di scarsa variabilità. Se però la qualità delle serie storiche a disposizione non è del tutto soddisfacente, e/o si ha la disponibilità di serie storiche macro-regionali con maggior frequenza di rilevazione, si può optare per una diversa strategia (**External Middle-Out Hierarchical Forecasting**) basata su Forecast separati per la base dati regionale e per quella nazionale e su di una conciliazione (esterna) alla base dati nazionale di tutti i livelli della gerarchia della base dati regionale (Fig. 1).

Fig. 1 Strategia bi-fase (National Middle-Out Hierarchical Forecasting) per la stima di dati regionali correnti



Benefici

La strategia Middle-Out esterna mira al superamento del gap informativo consentendo una prima stima del dato regionale corrente a cui si applica un aggiustamento tramite la conciliazione al dato nazionale esterno (i.e. non ottenuto dall'aggregazione del dato regionale), stimato a partire da una base dati diversa e migliore rispetto a quella di livello NUTS2. **Con la Middle-Out esterna si beneficia di stime correnti robuste, migliorate appunto dalla conciliazione al dato nazionale.**

Ad esempio, l'occupazione italiana derivata dall'aggregazione delle stime regionali, risulta sovrastimata (+0,3%) rispetto alla previsione ottenuta come aggregazione trimestrale delle stime nazionali esterne (Tab. 1, Fig. 2), e a livello delle regioni d'Europa, la previsione con conciliazione esterna risulterà corretta dai coefficienti di conciliazione nazionali come avviene per la regione Emilia-Romagna (Tab. 2, Fig. 3).

Tab. 1 Occupazione (in migliaia), classe età 15-64, aggregazione e conciliazione esterna, Italia, 1999-2008

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Aggregazione regionale (in migliaia)	20357	20752	21169	21478	21710	22060	22214	22618	22941	23275
Aggregazione trimestrale esterna (in migliaia)	22875	23193
Coefficiente di conciliazione	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9971	0,9965

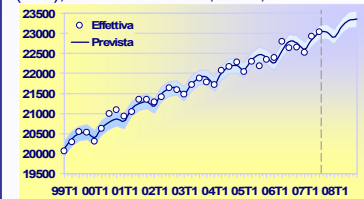
Fonte: Stime ESeC su dati Eurostat (Labour Force Survey). Note: Dati per la specificazione dei modelli: I. annuali dal 1999 al 2006 (previsioni 2007-2008); II. trimestrali dal 1999T1 al 2007T3 (previsioni 2007T4-2008T4).

Tab. 2 Occupazione (in migliaia), età 15-64, stima da modello regionale e con conciliazione esterna, Emilia-Romagna, 1999-2008

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Stima da modello regionale	1704	1737	1757	1786	1810	1813	1837	1876	1891	1913
Stima da modello reg. e conciliazione esterna	1885	1907

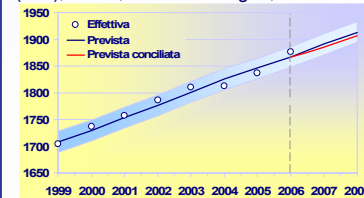
Fonte: Stime ESeC su dati Eurostat (Labour Force Survey) - dati disponibili al 12 marzo 2008. Note: Dati per la specificazione dei modelli annuali dal 1999 al 2006 (previsioni 2007-2008).

Fig. 2 Modello e previsione occupazione (000), classe età 15-64, Italia, 1999-2008



Fonte: stime ESeC su dati Eurostat (Labour Force Survey). Note: Modello: Additive (Winters) Seasonal Smoothing (MAPE = 0,37).

Fig. 3 Modello e previsione occupazione (000), 15-64, Emilia-Romagna, 1999-2008



Fonte: stime ESeC su dati Eurostat (Labour Force Survey). Note: Modello: Linear (Holt) Exponential Smoothing (MAPE = 0,45).