

La metrica autoregressiva per la valutazione e ottimizzazione di sistemi di monitoraggio ambientale¹

G. Damiana Costanzo e Emma Sarno

Dipartimento di Economia Politica, Università della Calabria
dm.costanzo@unical.it; e.sarno@unical.it

Summary: This paper deals with the evaluation and the optimization of a monitoring network. We propose an innovative use of the AR metric (Piccolo, 1984; 1990) which is a measure of structural discrepancy between ARIMA models. The analysis of the mutual dissimilarity among models corresponding to each station in the network and the selection of subsets of sites which maximizes it lead to a new optimal design by reducing redundant information. Hourly data of atmospheric carbon monoxide in the urban area of Rome are considered to illustrate the new procedure.

Key words: AR Metric, ARIMA Models, Monitoring Networks

1. Introduzione

I problemi ambientali ed in particolar modo l'inquinamento sono argomenti di grande interesse ed attualità, di forte impatto nella vita sociale. Studi di monitoraggio e ricerche vengono condotti per quantificare l'ammontare di inquinanti immessi nell'ambiente, per individuare trend

¹Questo lavoro costituisce un parziale adattamento di una relazione presentata al Convegno: "Conoscere per decidere: le ricerche del Dipartimento di Economia Politica dell'UNICAL a sostegno delle scelte per lo sviluppo socio-economico della Calabria", tenutosi a Rende (CS) nell'ottobre 2000.

di lungo periodo, per valutare la loro conformità agli standard ambientali o a quelli relativi alle emissioni; altri studi si occupano invece di determinare come gli inquinanti si distribuiscono nell'aria, nell'acqua, sul suolo e nel sottosuolo, i loro tempi di persistenza e gli effetti sull'uomo e l'ambiente.

Lo stadio fondamentale per la conoscenza dei fenomeni d'inquinamento è costituito dalla rilevazione delle sostanze inquinanti. Le concentrazioni vengono misurate dalle reti di monitoraggio, ossia da un insieme di siti di misurazione coordinati e gestiti da un centro operativo in base a criteri omogenei, per i quali la messa in rete agevola la determinazione della variabilità spaziale delle concentrazioni ma anche, come nel caso dell'inquinamento atmosferico, di altri parametri meteorologici e climatici quali direzione del vento, quantità di pioggia, umidità.

La definizione dei siti rappresenta una delle fasi più qualificanti della costruzione di una rete di rilevamento. Il disegno della rete deve essere infatti tale da fornire un'adeguata copertura spaziale ed una ragionevole risoluzione temporale del fenomeno ambientale allo studio; mentre, il numero dei siti sui quali articolare una rete può risultare limitato dall'elevato costo di investimento e gestione delle apparecchiature utilizzate.

Esistono in letteratura ormai numerosi studi sulla dislocazione dei siti di monitoraggio. In alcuni il problema dell'allocazione ottimale dei siti è formulato come un tipico problema di ricerca operativa, ad esempio negli algoritmi di Noll-Mitsutomi (1983) o di Langstaff (Langstaff *et al.*, 1987). In approcci di natura più eminentemente statistica, il problema del disegno ottimale di una rete di monitoraggio è affrontato o secondo l'ottica della teoria del disegno degli esperimenti (Fedorov e Mueller 1989) oppure utilizzando tecniche di kriging tipiche della geostatistica (Cressie *et al.*, 1990; Warrick e Myers, 1989). Altri metodi statistici si basano, invece, sulla determinazione della similarità dei siti. Spesso infatti, le misurazioni dei fenomeni ambientali sono altamente correlate nel tempo e/o nello spazio. In tali casi, i dati ottenuti da una o più centraline possono essere adeguatamente previsti a partire dalle osservazioni disponibili in corrispondenza di altre postazioni. L'analisi della ridondanza informativa eventualmente presente nella rete può essere sfruttata per apportare modifiche al suo assetto, migliorandone le prestazioni in

termini di rapporto tra informazione acquisibile e numero di siti. Ciò può condurre alla rimozione dei siti che sono meglio predetti dal resto della rete (Caselton e Husain, 1980; Caselton e Zidek, 1984), alla loro eventuale sostituzione, oppure all'aggiunta di nuovi siti di modo che nel complesso migliori l'informazione attesa nel senso di Shannon e sia massimizzata l'entropia della rete (Guttorp *et al.*, 1992; Sampson e Guttorp, 1994).

L'obiettivo di questo lavoro è valutare e ottimizzare l'efficienza informativa di una rete di monitoraggio ambientale analizzando l'eventuale presenza di ridondanza di informazione fornita dalle postazioni di misurazione. In particolare, viene proposto un nuovo metodo per scegliere una configurazione ottimale di dimensione k da un insieme di n siti monitorati. Tale metodo si differenzia da quelli precedentemente menzionati in quanto prende esplicitamente in considerazione la dimensione temporale del fenomeno e non solo quella spaziale. A questo fine viene introdotta come misura di dissimilarità tra i siti la distanza tra i modelli che descrivono le serie storiche corrispondenti. L'utilizzo di una distanza per valutare la ridondanza informativa di una rete è stato già considerato in Costanzo & D'Urso (2000) con riferimento ad una formulazione di tipo stato-spazio del modello strutturale per serie storiche, utilizzando algoritmi di stima ricorsiva basati sul filtro di Kalman (Costanzo, 1999). La procedura che presentiamo in questo lavoro si avvale invece dei modelli ARIMA (Box e Jenkins, 1976) e della metrica autoregressiva introdotta da Piccolo (1984, 1990) per confrontare e classificare dati dinamici attraverso i modelli ARIMA che li rappresentano.

In particolare, nel paragrafo 2 introduciamo in termini generali la metrica autoregressiva. Nel paragrafo 3, illustriamo la procedura di valutazione e ottimizzazione di una rete di monitoraggio ambientale, facendo riferimento ad un caso reale relativo alle misurazioni del monossido di carbonio rilevate nella città di Roma nell'anno 1992-1993, cui erano state applicate metodologie classiche di ottimizzazione in un precedente rapporto tecnico (La Sala, 1995). Più esattamente, discutiamo il trattamento preliminare dei dati (§3.1) e quindi i risultati dell'applicazione (§3.2). Infine nel paragrafo 4, concludiamo con alcune considerazioni generali e le indicazioni sui possibili indirizzi di ricerca futuri.

2. La metrica autoregressiva

In letteratura, la metrica autoregressiva ha trovato applicazione in molti campi. Ad esempio, è stata impiegata nella cluster analysis (Piccolo, 1990), per valutare mutamenti di fenomeni demografici (Corduas e Piccolo, 1995), per individuare dati anomali (Corduas, 1990), per classificare serie stazionarie (Maharaj, 1996), per stabilire l'adeguatezza di filtri stagionali (Corduas e Piccolo, 1999). In questo lavoro, proponiamo di utilizzarla nell'ambito del monitoraggio ambientale come strumento per la scelta di combinazioni ottimali di centraline di rilevamento.

Più precisamente, le serie storiche $W_{i,t}$, $i = 1, 2, \dots, n$, $t = 1, 2, \dots, T_i$, ottenute in corrispondenza di n siti di misurazione presenti nella rete e rappresentate mediante modelli ARIMA, vengono confrontate a due a due producendo, per ogni possibile coppia, una misura della loro diversità strutturale attraverso la metrica autoregressiva. La procedura di valutazione e ottimizzazione della rete che suggeriamo consiste nell'individuare per ogni sottoinsieme k , $k < n$, la dislocazione di centraline che rende massima la loro mutua dissimilarità. Ciò consente di ridurre la ridondanza delle informazioni raccolte. Infatti, se la localizzazione dei siti di misurazione è tale che in due punti diversi e non necessariamente vicini dello spazio le serie storiche corrispondenti sono ascrivibili a modelli simili, allora l'informazione di una delle due centraline può ritenersi ridondante in quanto prevedibile a partire dalle altre. È caratteristica specifica della metrica autoregressiva la proprietà per cui essa tende a zero se (per valori iniziali delle serie storiche e ordini dei modelli noti) le funzioni di previsione corrispondenti alle due serie confrontate tendono a coincidere (Piccolo, 1989, p.232). Ne segue che, all'interno di un sistema di monitoraggio, la scelta delle postazioni che forniscono serie storiche più dissimili corrisponde ad un'allocazione efficiente delle centraline.

In questo lavoro, facendo riferimento alla notazione classica di Box e Jenkins (1976), con $Z_t \sim ARIMA(p, d, q) \times (P, D, Q)_s$ indichiamo il processo stocastico $\phi(B)\Phi(B^s)\nabla^d\nabla_s^D Z_t = \theta(B)\Theta(B^s)a_t$, in cui a_t è un processo *white noise* gaussiano. Quando $\theta(B)\Theta(B^s) = 0 \rightarrow |B| > 1$, il processo Z_t è invertibile e ammette la rappresentazione $AR(\infty)$ data da $\pi(B)Z_t = a_t$, con $\pi(B) = \phi(B)\Phi(B^s)\nabla^d\nabla_s^D\theta^{-1}(B)\Theta^{-1}(B^s) = 1 -$

$$\pi_1 B - \pi_2 B^2 - \dots$$

La metrica introdotta da Piccolo (1984, 1990) per valutare la differenza strutturale tra due processi X_t e Y_t invertibili confronta le sequenze $\{\pi_{X,j}, j = 1, 2, \dots\}$ e $\{\pi_{Y,j}, j = 1, 2, \dots\}$ dei pesi autoregressivi mediante la distanza euclidea $d(X_t, Y_t) = \sqrt{\sum_{j=1}^{\infty} (\pi_{X,j} - \pi_{Y,j})^2}$.

Scegliendo un'adeguata approssimazione autoregressiva finita $AR(L)$ del processo invertibile ed opportune stime dei pesi $\hat{\pi}$, possiamo utilizzare lo stimatore della distanza $\hat{d}_L(X_t, Y_t) = \sqrt{\sum_{j=1}^L (\hat{\pi}_{X,j} - \hat{\pi}_{Y,j})^2}$. La sua distribuzione campionaria asintotica è nota nel caso di confronti tra modelli AR (Piccolo, 1989) e ARMA (Corduas, 1996) per stime dei pesi autoregressivi ottenute con il metodo della massima verosimiglianza, ma è stata recentemente studiata anche nel caso di confronti tra modelli MA (Sarno, 2000) sotto l'ipotesi di stime dei minimi quadrati.

3. La rete di monitoraggio sulla qualità dell'aria nella città di Roma

L'esempio preso in considerazione in questo lavoro riguarda le concentrazioni medie orarie del monossido di carbonio rilevate nella città di Roma nel periodo che va dal 1 aprile 1992 al 31 marzo 1993 (Fonte: Regione Lazio). La rete è costituita complessivamente da 9 stazioni di diversa tipologia² che riportiamo nella Tabella 1 e il cui nome corrisponde alla rispettiva ubicazione nell'area cittadina.

²Il D.M.A. 20.05.1991 pianifica la struttura di una rete urbana per il monitoraggio ambientale su 4 tipologie di stazioni: di tipo A, poste in zone non direttamente interessate dalle sorgenti di emissione urbana (parchi, isole pedonali); di tipo B, in zone ad elevata densità abitativa; di tipo C, in zone ad elevato traffico e bassa ventilazione; di tipo D, in periferia o in aree suburbane (finalizzate alla misura di inquinanti fotochimici, quali l'ozono). La rete di Roma esaminata, in quanto preposta alla rilevazione di sostanze inquinanti di diretto impatto sulla salute dell'uomo, è costituita esclusivamente da stazioni di tipo B e C.

Tabella 1. Le centraline di rilevamento.

L.go Arenula (<i>AREN</i>) dall'1.04.92 al 25.06.92	B
P.za Fermi (<i>FERM</i>) dall'1.04.92 al 14.07.92	B
C.so Francia (<i>FRAN</i>) dall'1.04.92 al 28.08.92	C
P.zza Gondar (<i>GOND</i>) dall'1.04.92 al 7.07.92	B
L.go Gregorio XIII (<i>GREG</i>) dall'1.04.92 al 28.08.92	C
L.go Magnagrecia (<i>MGRE</i>) dall'1.04.92 al 29.07.92	B
L.go Montezemolo (<i>MONT</i>) dall'1.04.92 al 28.08.92	C
L.go Preneste (<i>PREN</i>) dall'1.04.92 al 11.08.92	C
Via Tiburtina (<i>TIBU</i>) dal 29.04.92 al 21.06.92	C

3.1 Descrizione e analisi preliminari dei dati

Le stazioni di monitoraggio sono completamente automatizzate e ogni giorno forniscono 23 dati orari poichè la prevista calibrazione degli strumenti causa la mancata disponibilità di un'osservazione sempre alla stessa ora. Nel periodo in esame tutte le apparecchiature hanno avuto bisogno di manutenzione straordinaria, compromettendo in parte la disponibilità e la qualità dei dati; la temporanea inattività delle centraline in periodi diversi, ma a volte relativamente lunghi dell'anno, ha determinato un elevato numero di dati mancanti anche consecutivi e quindi impossibili da ricostruire. Di conseguenza, si è deciso di effettuare le analisi su un arco temporale più ristretto, durante il quale fosse però disponibile, per tutti i siti, la maggior parte dei dati contemporaneamente. Tale periodo (vedi Tabella 1) è coinciso approssimativamente con i primi quattro mesi dell'anno in esame.

Anche nell'ambito dei sottoperiodi menzionati, il primo problema che si è posto è stato quello dei dati mancanti: di quelli giornalieri corrispondenti alla calibrazione di routine quotidiana e di altri determinati da cause non note ma presumibilmente dovuti all'occasionale cattivo funzionamento delle apparecchiature, cui seguiva una calibrazione straordinaria. Per ricostruire le osservazioni mancanti si è utilizzata una media mobile a due termini; successivamente, quel numero di termini è stato avvalorato

dall'ordine autoregressivo identificato per i modelli ARIMA pari proprio a due.

Per rendere possibili le successive analisi, a causa della elevata asimmetria positiva di questo tipo di dati ambientali, si è reso necessario trasformare le serie storiche di modo che l'ipotesi di normalità fosse legittimata; a tal fine ai dati originari è stata applicata la consueta trasformazione logaritmica: $Z_{i,t} = \log(1 + W_{i,t})$.

Infine, anche la presenza di pochissimi valori estremi (ossia esterni all'intervallo compreso tra la media più o meno cinque volte lo scarto quadratico medio) ha richiesto uno specifico trattamento, che è consistito nella sostituzione di quei dati con le stesse medie mobili di cui sopra.

3.2 Procedura di valutazione e risultati

La procedura di valutazione e ottimizzazione della rete di monitoraggio sulla qualità dell'aria della città di Roma si può riassumere nei seguenti punti:

1. Per ogni centralina, seguendo lo schema di Box e Jenkins, sono stati identificati i modelli ARIMA sulla base dell'analisi delle funzioni di autocorrelazione globale (Figura 1) e parziale, nonché degli spettri delle serie storiche. Osserviamo che, già a questo stadio, il grafico delle funzioni di autocorrelazione stimate ci consente di cogliere strutture di correlazione più o meno simili tra i dati in questione, ma che l'effettiva omogeneità strutturale dei modelli atti a rappresentarli sarà misurata proprio grazie alla metrica autoregressiva. A titolo puramente esemplificativo, rileviamo da un lato la diversità tra le stazioni *FERM*, *MGRE* e *MONT*, dall'altra la similitudine di *FRAN* e *TIBU*.

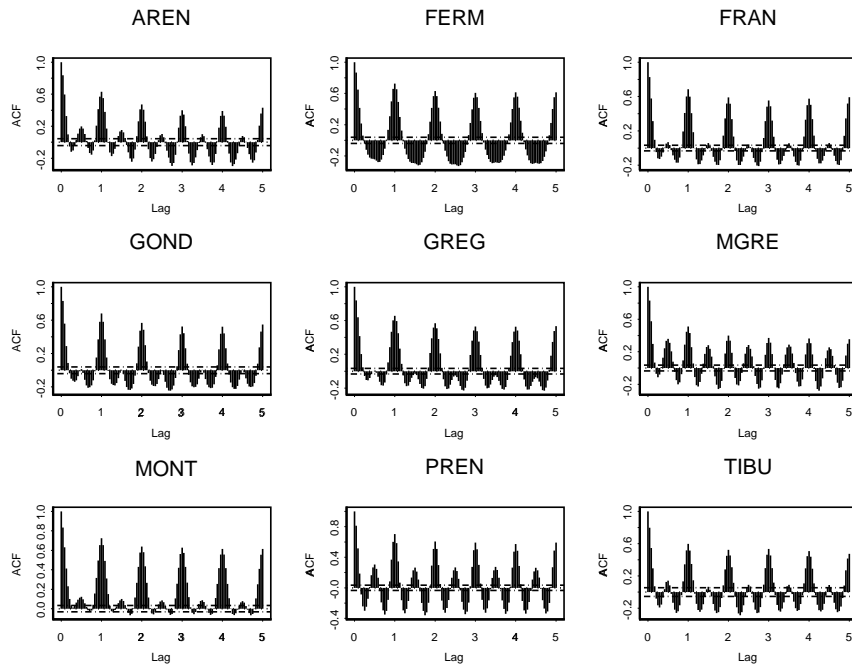


Figura 1. Funzioni di autocorrelazione globale delle serie storiche per centraline di rilevamento (max lag=24*5).

2. I modelli, una volta identificati, sono stati stimati con il metodo della massima verosimiglianza utilizzando la funzione `arima.mle` del programma S-Plus, e poi validati mediante l'analisi dei residui e test specifici. Seguono i risultati con le stime dei parametri:

$$\cdot AREN \sim ARIMA(2, 0, 0) \times (1, 1, 1)_{24}$$

$$(1 - 0.8234B + 0.012B^2)(1 - 0.0395B^{24})(1 - B^{24})Z_t = (1 - 0.9757B^{24})a_t$$

$$\cdot FERM \sim ARIMA(2, 0, 0) \times (1, 1, 1)_{24}$$

$$(1 - 0.8513B + 0.0683B^2)(1 - 0.0604B^{24})(1 - B^{24})Z_t = (1 - 0.9935B^{24})a_t$$

$$\cdot FRAN \sim ARIMA(2, 0, 0) \times (1, 1, 1)_{24}$$

$$(1 - 0.7155B - 0.0471B^2)(1 - 0.0895B^{24})(1 - B^{24})Z_t = (1 - 0.9591B^{24})a_t$$

$$\cdot GOND \sim ARIMA(2, 0, 0) \times (1, 1, 1)_{24}$$

$$(1 - 0.8556B + 0.0624B^2)(1 - 0.0954B^{24})(1 - B^{24})Z_t = (1 - 0.9776B^{24})a_t$$

$$\begin{aligned}
& \cdot \textit{GREG} \sim \textit{ARIMA}(2, 0, 0) \times (1, 1, 1)_{24} \\
& (1 - 0.7065B - 0.069B^2)(1 - 0.0515B^{24})(1 - B^{24})Z_t = (1 - 0.9601B^{24})a_t \\
& \cdot \textit{MGRE} \sim \textit{ARIMA}(2, 0, 0) \times (1, 1, 1)_{24} \\
& (1 - 0.8992B + 0.0805B^2)(1 - 0.0486B^{24})(1 - B^{24})Z_t = (1 - 0.9592B^{24})a_t \\
& \cdot \textit{MONT} \sim \textit{ARIMA}(2, 0, 0) \times (1, 1, 1)_{24} \\
& (1 - 0.6526B - 0.1024B^2)(1 - 0.0736B^{24})(1 - B^{24})Z_t = (1 - 0.934B^{24})a_t \\
& \cdot \textit{PREN} \sim \textit{ARIMA}(1, 0, 0) \times (1, 1, 1)_{24} \\
& (1 - 0.7657B)(1 - 0.1011B^{24})(1 - B^{24})Z_t = (1 - 0.9592B^{24})a_t \\
& \cdot \textit{TIBU} \sim \textit{ARIMA}(2, 0, 0) \times (1, 1, 1)_{24} \\
& (1 - 0.6927B - 0.0326B^2)(1 - 0.051B^{24})(1 - B^{24})Z_t = (1 - 0.9643B^{24})a_t
\end{aligned}$$

Constatiamo che l'output di tutte le centraline è riconducibile ad un unico meccanismo generatore dei dati, ossia il modello $\textit{ARIMA}(2, 0, 0) \times (1, 1, 1)_{24}$, con un'unica parziale eccezione costituita dalla stazione *PREN*, il cui ordine autoregressivo non stagionale è 1 anziché 2. Inoltre, precisiamo che la periodicità giornaliera ha richiesto l'operatore differenza stagionale di ordine 24. Questo perchè l'alternativa di inserire nel modello una componente periodica deterministica (in taluni casi fortemente suggerita dalla quasi-cancellazione degli operatori per la presenza di un coefficiente prossimo a 1 nell'operatore MA stagionale) è stata scartata a causa dei risultati di gran lunga peggiori che si ottenevano dall'analisi dei residui.

3. Per consentire i successivi confronti, i modelli sono stati riscritti nella versione *AR*, scegliendo come punto di troncamento $L = 200$. Tale scelta³ ha consentito alla rappresentazione puramente autoregressiva di includere da un lato i coefficienti più "rilevanti", dall'altro di coprire almeno un arco temporale settimanale. Tuttavia, è importante evidenziare che le analisi successive si sono dimostrate robuste rispetto alla scelta di L ; infatti risultati pressoché

³Per processi stazionari esistono condizioni sul punto di troncamento (tra esse $L \simeq n^{1/3}$) a garanzia delle proprietà asintotiche degli stimatori dei pesi autoregressivi (Berks, 1974; Bhansali, 1978), ma le stesse non sono generalizzabili al nostro caso.

simili si ottenevano anche per punti di troncamento molto più bassi, ad esempio per $L = 100$. I pesi $\hat{\pi}$ sono rappresentati graficamente nella Figura 2. In essa riemergono, da altra prospettiva, le similitudini tra centraline già evidenziate a partire dalle funzioni di autocorrelazione.

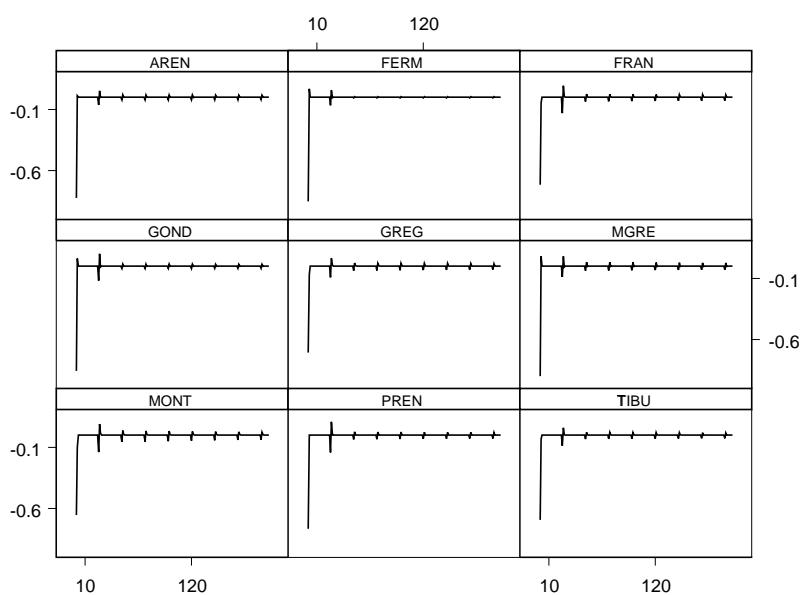


Figura 2. I pesi della rappresentazione autoregressiva per centraline di rilevamento.

4. Utilizzando i pesi $\hat{\pi}$, sono state calcolate le distanze autoregressive per ogni coppia di modelli. A scopo illustrativo, sulla matrice di distanze è stato condotto uno scaling multidimensionale classico (infatti $d(X_t, Y_t)$ soddisfa tutte le proprietà di una metrica) mediante la funzione `cmdscale` di S-Plus. Dal relativo grafico ottenuto in due dimensioni (Figura 3) è immediato cogliere la similarità (punti vicini) e la dissimilarità strutturale (punti lontani) tra i vari siti che, in generale, non trovano necessariamente corrispondenza con l'effettiva distanza geografica.

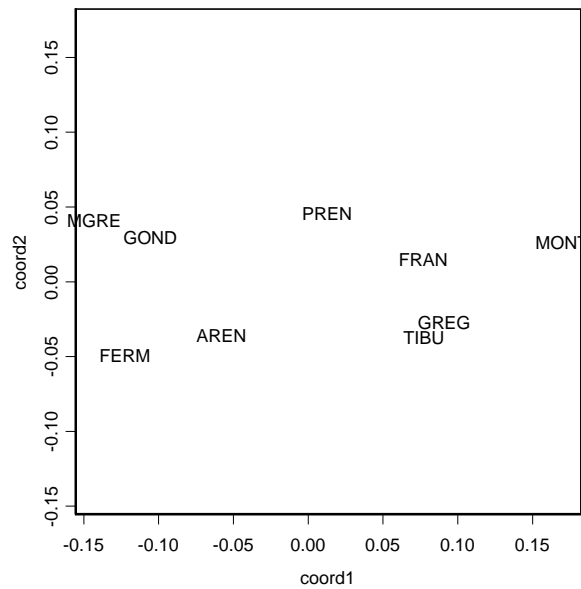


Figura 3. Rappresentazione MDS delle centraline di rilevamento basata sulla metrica autoregressiva.

5. Infine, sono state ordinate tutte le possibili combinazioni di n centraline in sottoinsiemi di k elementi, per $k = 2, 3, \dots, 9$, in base al criterio di dissimilarità decrescente. Più precisamente, per ogni sottoinsieme k , le possibili combinazioni sono state classificate in ordine decrescente in base alla somma delle rispettive distanze autoregressive. La prima combinazione di ciascun sottoinsieme è perciò quella che assicura la minore ridondanza di informazione, ossia la combinazione "ottimale". Per motivi di spazio, nella Tabella 2 vengono mostrate solo queste ultime. È interessante osservare come, per $k = 3$, la procedura abbia individuato proprio la combinazione {FERM, MGRE, MONT} in cui ritroviamo le stazioni con strutture di autocorrelazione più dissimili, così come era emerso al

precedente punto 1. La classifica completa (non riportata) consente inoltre di valutare le potenziali configurazioni di sottoinsiemi alternativi a quelli ottimali, alle quali gli operatori ambientali potrebbero essere comunque interessati per ragioni di natura non strettamente statistica.

Tabella 2. Le combinazioni ottimali di centraline

k	Siti	$\sum_{i>j} d(Z_{i,t}, Z_{j,t})$
2	<i>MGRE, MONT</i>	0.315
3	<i>FERM, MGRE, MONT</i>	0.732
4	<i>FERM, GREG, MGRE, MONT</i>	1.292
5	<i>FERM, GOND, GREG, MGRE, MONT</i>	1.938
6	<i>FERM, GOND, GREG, MGRE, MONT, TIBU</i>	2.735
7	<i>FERM, GOND, GREG, MGRE, MONT, PREN, TIBU</i>	3.572
8	<i>AREN, FERM, FRAN, GOND, GREG, MGRE, MONT, TIBU</i>	4.543
9	<i>AREN, FERM, FRAN, GOND, GREG, MGRE, MONT, PREN, TIBU</i>	5.568

4. Considerazioni finali

Ai fini della interpretazione dei risultati ottenuti è utile considerare la tipologia delle stazioni che costituiscono la rete di Roma, in quanto essa consente di cogliere caratteristiche spaziali differenti, legate all'intensità con cui si manifesta il fenomeno dell'inquinamento. Ricordiamo che le centraline qui analizzate sono localizzate o in aree ad elevata densità abitativa (quelle di tipo B) oppure in aree ad elevata intensità di traffico e bassa ventilazione (quelle di tipo C). Emerge che tutte le configurazioni ottimali qui proposte sono equilibrate quanto alla presenza dei due tipi di stazioni: ad esempio, si evince dalla Tabella 2 che per $k = 2$ la combi-

nazione ottimale prevede una centralina di tipo B e una di tipo C, mentre per $k = 4$ se ne hanno due di tipo B e due di tipo C.

Inoltre, l'ordinamento completo dei sottoinsiemi di centraline ci consente di affermare che, se ci si limitasse a considerare stazioni di un solo tipo, si finirebbe con lo scegliere le configurazioni peggiori secondo il nostro criterio: infatti, nel caso $k = 4$, l'insieme costituito dalle centraline di tipo B ossia {AREN, FERM, GOND, MGRE} occupa il terzultimo posto tra le possibili 126 combinazioni mentre, per $k = 5$, quello delle centraline di tipo C ossia {FRAN, GREG, MONT, PREN, TIBU} è l'ultimo di altrettante combinazioni.

In questo lavoro abbiamo esplorato le potenzialità della metrica autoregressiva come strumento per l'ottimizzazione delle reti essenzialmente nei suoi aspetti descrittivi. I futuri sviluppi di questa ricerca dovranno tener conto anche di valutazioni di tipo probabilistico. Solo in questo modo sarà possibile determinare la dimensione ottimale k^* del sottoinsieme ottimale oppure effettuare previsioni su punti della rete non monitorati. Si potrà ottenere così, nella fase di realizzazione di un sistema di controllo ambientale, un'allocazione efficiente di nuovi siti di misurazione.

Ringraziamenti: Desideriamo ringraziare il Prof. D. Piccolo per i suoi preziosi commenti. Resta comunque responsabilità delle sole autrici il contenuto di questo lavoro, di cui i paragrafi 1 e 3.1 vanno attribuiti a G.D. Costanzo e i paragrafi 2 e 3.2 a E. Sarno. La ricerca è stata finanziata con fondi di Ateneo e PRIN 1998-2000: "Modelli statistici per l'analisi delle serie temporali".

Riferimenti bibliografici

Berks, K.N. (1974) Consistent autoregressive spectral estimates, *The Annals of Statistics*, 2, 3, 489-502

- Bhansali, R.J. (1978) Linear prediction by autoregressive model fitting in the time domain, *The Annals of Statistics*, 6, 1, 224-231
- Box, G.E.P. e Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- Caselton, W.F. e Husain, T. (1980) Hydrologic networks: information transmission, *Journal of Water Resources Planning and Management Division*, ASCE, 106(WR”), 503-20
- Caselton, W.F. e Zideck, J.V.(1984) Optimal monitoring network designs, *Statistics and Probability Letters*, 2, 223-7
- Corduas, M. (1990) Detecting influential observations in ARIMA models, *Centro di Specializzazione e Ricerche*, Portici (NA), RS 2/90
- Corduas, M. (1996) Uno studio sulla distribuzione asintotica della metrica autoregressiva, *Statistica*, LVI, 3, 321-332
- Corduas, M. e Piccolo, D. (1995) Mutamenti strutturali della natalità e metodi statistici, *Atti del Convegno SIS: Continuità e discontinuità dei processi demografici*, Arcavacata di Rende, 315-322
- Corduas, M. e Piccolo, D. (1999) On the use of the AR metric for seasonal adjustments, *Atti del CLADAG*, Roma, 1-4
- Costanzo, G.D. (1999) Il filtro di Kalman nell’analisi di serie storiche di dati ambientali: la serie storica delle polveri nella città di Genova, in: *Modelli Statistici per l’Ambiente, Quaderni del Dipartimento di Statistica, Probabilità e Statistiche Applicate*, Roma, 8, 13-37
- Costanzo, G.D. e D’Urso, P. (2000) A distance measure between structural models, *Proceedings of COMPSTAT 2000*, Utrecht, 25-27
- Cressie, N., Gotway, C.A. e Grondona, M.O. (1990) Spatial prediction from networks, *Chemical Intelligence Lab. Systems*, 7, 251-71
- Fedorov, V. e Mueller, W. (1989), Comparison of two approaches in the optimal design of an observation network, *Statistics*, 20, 339-51
- Guttorp, P., Sampson P.D. e Newman K. (1992) Nonparametric estimation of spatial covariance with application to monitoring network evaluation, in *Statistics in the Environmental & Earth Sciences*, Walden & Guttorp (eds), Edward Arnold, London, 39-51
- Langstaff, J., Seigneur C., Behar J. e Mcelroy J.L. (1987) Design of an optimal air monitoring network for exposure assessments, *Atmospheric Environment*, 21

La Sala, E. (1995) Valutazione ed ottimizzazione di una rete di rilevamento del monossido di carbonio in area urbana: un'applicazione al caso di Roma, *ENEA RT/AMB/95*

Maharaj, E. A. (1996) A significance test for classifying ARMA models, *Journal of Computation and Simulation*, 54, 305-331

Noll, K.E. e Mitsutomi S. (1983) Design methodology for optimal dosage air monitoring site selection, *Atmospheric Environment*, 17

Piccolo, D. (1984) Una topologia per i modelli ARIMA, *Statistica*, XLIV, 47-59

Piccolo, D. (1989) On a measure of dissimilarity between ARIMA models, *Proceedings of the Business and Economic Statistics Section of the ASA Meeting*, Washington, 231-236

Piccolo, D. (1990) A distance measure for classifying ARIMA models, *Journal of Time Series Analysis*, 11, 1153-1164

Sampson, P.D. e Guttorp P. (1994) Nonparametric estimation of non-stationary spatial covariance structure, *Journal of the American Statistical Association*, 87, 417, 108-119

Sarno, E. (2000) The behaviour of the AR metric for MA model comparisons, *Atti della XL Riunione Scientifica SIS*, Firenze, 149-152

Warrick, A.W. e Myers, D.E. (1987) Optimization of sampling locations for variogram calculations, *Water Resources*, 23, 496-500