

APAT
Agenzia per la protezione dell'ambiente e per i servizi tecnici

RASSEGNA BIBLIOGRAFICA
ANCHE PER VIA TELEMATICA SUI METODI STATISTICI DI
AGGREGAZIONE DEGLI INDICATORI AMBIENTALI

Studio realizzato dal Dott. Roberto Paolini
Servizio Interdipartimentale Informativo Ambientale
Statistica Ambientale

Tutor: Dott.ssa Mariaconcetta Giunta

Marzo 2004

ABSTRACT

L'obiettivo del presente lavoro è quello di effettuare una rassegna bibliografica sui metodi statistici di aggregazione degli indicatori ambientali, reperiti nell'ambito della letteratura specifica sia nazionale che internazionale. Per la predisposizione di questa tesi si è proceduto alla raccolta, alla catalogazione e all'analisi di documenti contenenti metodologie statistiche di aggregazione dei dati, ponendo l'attenzione in particolare sull'aggregazione degli indici di qualità dell'aria.

La tesi è articolata in quattro capitoli, una bibliografia e un capitolo contenente gli allegati. Il primo capitolo l'introduzione, oltre a presentare le modalità e le finalità dello studio, fornisce i chiarimenti necessari sul ruolo e la funzione degli indicatori ambientali, sulle problematiche dell'aggregazione e sul ruolo della geostatistica. Nel secondo capitolo viene illustrato l'approccio metodologico adottato al fine di perseguire le finalità proposte, con la creazione *ad hoc* di una scheda di rilevazione e catalogazione degli articoli ed i possibili campi di applicazione. Nel terzo capitolo oltre a presentare i risultati della ricerca, fra tutti gli articoli trovati si è deciso di porre l'attenzione soprattutto sull'aggregazione degli indici di qualità dell'aria, in quanto il problema dell'inquinamento dell'aria è ormai frequente emergenza e la sua soluzione o, comunque, solo la sua gestione ha delle conseguenze economiche e sociali rilevanti.

Il quarto capitolo presenta le conclusioni, facendo notare che tale ricerca non può essere esaustiva sia per l'importanza che ricopre sia perché le tecniche di aggregazione statistica descritte risultano a tutt'oggi in fase sperimentale e di studio (in quanto analizzano modelli non ancora applicati) e riguardano soprattutto i lavori svolti dal CNR (Centro Nazionale delle Ricerche), che risulta all'avanguardia in questo tipo di studi.

ABSTRACT

The objective of this work is to report the results of a bibliographical research on aggregation statistical methods related to environmental indicators; in particular this thesis focuses on quality air indexes.

The thesis contains four chapters, a bibliography and some annexes.

The first chapter, the introduction, describes the framework and the objective of this study, and explains the problems related to geostatistic and data aggregation.

The second chapter illustrates the methodology used in this work: a survey card to catalogue articles and possible fields of application.

The third chapter introduces the results of the research activity and focuses on aggregation methods of air quality indexes; in fact nowadays the problem of the air pollution is now frequent emergency with remarkable economic and social consequences.

The last chapter concludes noticing that such a research activity cannot be exhaustive and for the importance that covers and because the techniques of data statistical aggregation, here described, are still in experimental phase.

PREFAZIONE

Il Settore delle statistiche ambientali cura la qualificazione, la validazione, l'elaborazione, le operazioni di standardizzazione o di normalizzazione, l'aggregazione e la caratterizzazione attraverso schede descrittive degli indicatori. Il progetto è proprio quello di produrre strumenti operativi per una corretta gestione dei dati e degli indicatori.

Il trattamento corretto di più fenomeni simultaneamente è uno dei nodi cruciali dell'analisi ambientale.

L'importanza di tutto questo risiede nel fatto di rendere possibile una sintesi tra molti aspetti, consentendo una valutazione globale di maggiore efficacia e significatività.

Il percorso metodologico che conduce all'aggregazione di indicatori, dunque agli indici, passa attraverso alcune fasi di calcolo, senza le quali l'esito potrebbe essere invalidato.

Tale processo è così articolato:

1. identificazione dei dati utili allo scopo;
2. normalizzazione;
3. standardizzazione;
4. aggregazione.

L'oggetto della presente tesi è quello di fornire un quadro delle metodologie statistiche di aggregazione degli indicatori ambientali tramite una rassegna bibliografica, una catalogazione e successiva analisi di alcuni documenti ritenuti più significativi.

Dott.^{ssa} Mariaconcetta Giunta

Si ringraziano per la cortese collaborazione la Dott.^{ssa} Cristina Frizza, la Dott.^{ssa} Alessandra Galosi e la Dott.^{ssa} Federica Macrì (APAT-IAM)

INDICE

ABSTRACT.....	II
ABSTRACT.....	III
PREFAZIONE	IV
INDICE	V
1. INTRODUZIONE.....	1
1.1 Il ruolo e la funzione degli indicatori ambientali.....	2
1.2 Le problematiche dell'aggregazione.....	6
1.3 Il ruolo della geostatistica	9
2. METODOLOGIA PER LA RACCOLTA E LA CATALOGAZIONE DELLA DOCUMENTAZIONE	12
3. RISULTATI DELLA RICERCA SUI METODI DI AGGREGAZIONE STATISTICA DEGLI INDICATORI AMBIENTALI DI ORGANISMI NAZIONALI E INTERNAZIONALI.....	16
4. AGGREGAZIONE INDICI QUALITA' DELL'ARIA	20
4.1 Progetto sulla costruzione di indici per la qualità dell'aria (Bruno F. – Cocchi D.)	21
4.2 Modelli statistici per dati mancanti nel monitoraggio e previsione dell'inquinamento atmosferico (Progetto di Ricerca CNR 1999).....	25
4.3 Metodi statistici per l'analisi dell'ambiente e delle interazioni ambiente-salute (Colombi R. – Lovison G.)	26
4.4 Metodi statistici per lo studio e il monitoraggio dell'ambiente (Fassò A. – Progetto CNR 2003)	32
5. CONCLUSIONI.....	36
BIBLIOGRAFIA	37

1. INTRODUZIONE

L'obiettivo della tesi è quello di fornire un quadro completo delle metodologie statistiche di aggregazione dei dati, reperite nell'ambito della letteratura specifica sia nazionale che internazionale, applicabili agli indicatori ambientali, tramite una rassegna, una catalogazione e successiva analisi dei documenti presi in esame. Con il passare degli anni il numero di indicatori ambientali è aumentato, e continua a crescere, in maniera esponenziale; ciò se da un lato fornisce un quadro più completo della situazione ambientale, dall'altro induce a definire criteri generali per la scelta degli indicatori, al fine di giungere ad una aggregazione degli stessi, per aumentare la visibilità dell'ambiente nella serie di indicatori popolati usati dal pubblico. Il sistema delle Agenzie ambientali dispone di una serie di strutture disseminate sul territorio, in grado di raccogliere e fornire dati sullo stato dell'ambiente con caratteristiche di affidabilità, qualità e comparabilità. Tali dati costituiscono una solida base informativa per l'APAT (Agenzia per la Protezione dell'Ambiente e per i servizi Tecnici), che con la pubblicazione dell' "*Annuario dei dati ambientali 2003*", presenta uno strumento efficace ed efficiente per adempiere agli obblighi di *reporting* ambientale, che rappresenta il vertice di quella che è la "*piramide della conoscenza*¹".

La piramide evidenzia il processo che, partendo dalle principali fonti di produzione ordinaria di dati ambientali, il monitoraggio e il controllo, attraverso le fasi di gestione e valutazione dell'informazione, consente di elaborare prodotti di *reporting* per i differenti potenziali utenti finali, tra i quali i decisori politici e il grande pubblico. L'Annuario, nato dall'esigenza di presentare in modo organico i dati dei monitoraggi in possesso dell'Agenzia e di integrarli con quelli di altri enti ed istituti, ha lo scopo di porre le basi per fornire annualmente al

¹ Fonte: APAT – Annuario dei dati ambientali 2002

pubblico e agli amministratori un'informazione puntuale e attendibile sullo stato dell'ambiente italiano.

Nel volume sono stati popolati e rappresentati, in molti casi sia con articolazione territoriale sia con serie storiche, circa 180 indicatori contro i circa 160 dell'edizione precedente, è dunque il punto di partenza per una riflessione più articolata e sempre più attenta al contributo che l'Agenzia può offrire in materia di protezione ambientale.

Prima di proseguire nel nostro studio è necessario soffermarsi sugli indicatori ambientali ed il loro ruolo.

1.1 Il ruolo e la funzione degli indicatori ambientali

Gli indicatori sono ormai strumenti sempre più utilizzati nelle relazioni ambientali a carattere internazionale e nazionale. *“L'indicatore è un parametro, o un valore derivato da parametri, che fornisce informazioni sullo stato di un fenomeno/ambito/area con un significato che va oltre ciò che è direttamente associato al valore del parametro”* (def. OCSE), la sua funzione principale, è la rappresentazione sintetica dei problemi indagati in modo però da conservare il contenuto informativo dell'analisi.

La scelta e l'uso di un particolare indicatore sono strettamente collegati allo scopo che si vuole raggiungere. L'OCSE (Organizzazione per la cooperazione e lo sviluppo economico) individua due principali finalità:

- ridurre il numero di misurazioni e di parametri che normalmente sono richiesti per fornire un quadro "esatto" della situazione indagata;
- semplificare il processo di comunicazione attraverso cui, i risultati delle indagini, vengono forniti all'utilizzatore e divulgati.

La tendenza generale, a livello internazionale, è quella di classificare gli indicatori in base allo schema *Pressione Stato Risposta (PSR)*, che si basa sul concetto di causa/effetto e prevede lo sviluppo di una serie di indicatori ambientali suddivisi in:

- indicatori di pressione ambientale (diverse attività antropiche che costituiscono fonti di pressione sui vari comparti ambientali);
- indicatori di stato (la qualità dell'ambiente attuale e le sue alterazioni);
- indicatori di risposta (si riferiscono alle misure prese dalla società per migliorare lo stato dell'ambiente).

L'Agenzia Europea per l'Ambiente (AEA) ha introdotto una variazione al modello *PSR* (1995), mettendo in evidenza due fasi e lasciando invariate le altre, giungendo così al modello *DPSIR* (*Driving force-Pressure-State-Impact-Response*).

Il modello *DPSIR* (fig. 1.1) ha ampliato lo schema degli indicatori aggiungendo:

- indicatori di cause primarie “*driving force*” (i settori economici e le attività umane che inducono le pressioni ambientali);
- indicatori di impatto “*impact*”, che descrivono gli effetti sull'ecosistema e sulla salute umana derivanti dai fattori di pressione ambientale.

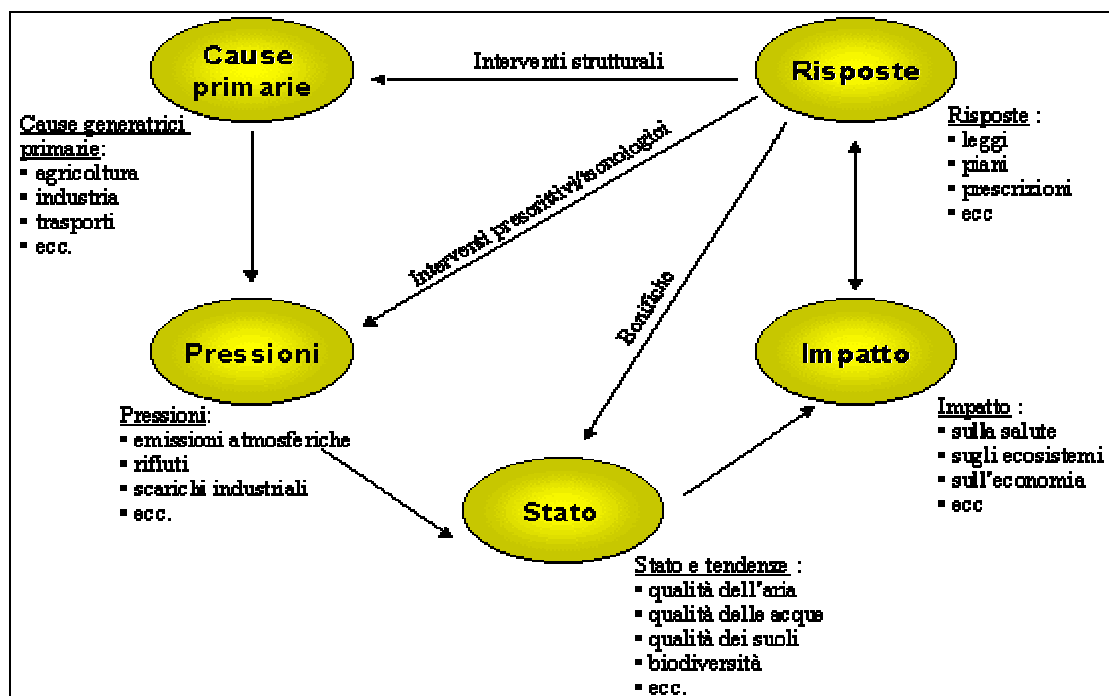


Fig. 1.1: Il modello DPSIR

L'OCSE negli anni ha individuato numerosi indicatori ambientali (circa 550), molti di questi presentavano a livello di dati rilevati, una disomogeneità delle relative serie storiche e una non buona copertura territoriale, con conseguente perdita di livello di specificità; inoltre, considerare più fenomeni contemporaneamente, implica il trattare diverse variabili, da qui la necessità, per la stessa Organizzazione, di trovare dei metodi di aggregazione che permettessero un'analisi più snella, ma puntuale dello stato dell'ambiente (al fine di ottenere un'unica misura di elevata espressività). Nacque quindi nel 1995 *ESEPI (European System for Environmental Pressure Indicators)* il progetto "Indici di Pressione", coordinato da EUROSTAT con la partecipazione di una dozzina di istituti di ricerca e nove istituti di statistica degli Stati membri, che si poneva come obiettivo principale la creazione di un'infrastruttura metodologica

per una descrizione completa e precisa delle pressioni sull'ambiente risultanti dalle attività umane. Tale progetto ha permesso di individuare **10 indici** di pressione ambientale (relativi ai 10 “campi della politica” basati sul quinto programma d'Azione Ambientale dell'Unione Europea) e **10 indicatori** ambientali per ogni indice di cui i primi **6** sono i più importanti (per un totale di 60 indicatori).²

I criteri di scelta degli indicatori hanno riguardato le seguenti tematiche:

- 1) rilevanza politica;
- 2) collegare ambiente ed attività economica;
- 3) misurabilità;
- 4) facile interpretazione (analisi, trend,...);
- 5) livelli di aggregazione.

Per giungere a questi indicatori si è concettualmente studiata ed analizzata quella che è definita la *piramide dell'informazione*³ (fig. 1.2), per poi arrivare a degli indici (cioè misure di controllo derivate da algoritmi statistici che divengono essi stessi misure di controllo) sopra citati, capaci di rappresentare una tendenza o un andamento dei fenomeni sia lungo l'asse dei tempi, sia lungo l'asse dello spazio.



Fig. 1.2: Piramide dell'informazione

² Vale la pena sottolineare che non esiste l'indicatore ideale, quanto piuttosto, al limite, è più corretto e saggio fornire dei criteri 'ideali' in grado di garantire tecnicamente la selezione e lo sviluppo di un buon indicatore.

³ Fonte: APAT – Annuario dei dati ambientali 2002.

1.2 Le problematiche dell'aggregazione

Uno dei nodi cruciali dell'analisi ambientale è il trattamento corretto di più fenomeni simultaneamente. L'importanza di ciò risiede proprio nel rendere possibile un'unica sintesi tra molti aspetti, consentendo una valutazione globale di maggior efficacia e significatività. Considerare più fenomeni contemporaneamente, implica il trattare diverse variabili per ottenere un'unica misura di elevata espressività. Il problema si pone se si desidera, appunto, trattare più variabili con la finalità di valutare le prestazioni ambientali, riducendole ad un indice unico.

Infatti ogni variabile ha una propria unità di misura e non può essere né sommata né mediata ad un'altra variabile.

L'analisi ambientale e l'analisi territoriale non possono esaurirsi certo alla valutazione della distribuzione di ogni singolo vettore di dati o di indicatori. Il trattamento di più variabili simultaneamente, impone all'analista di definire preliminarmente quale metodologia seguire per condurre correttamente il suo studio e per ottenere i risultati attesi. Infatti, nell'analisi di più variabili si affacciano nuovi problemi che ne condizionano l'elaborazione stessa: le diverse unità di misura utilizzate, i diversi fondi scala, la diversa importanza di ogni variabile rispetto alle altre, il diverso comportamento di ogni vettore di dati (stabile o disperso). Una delle tecniche per considerare più variabili consiste nel ridurle ad una o più, partendo da un set ampio, così da ottenere un indice sintetico altamente espressivo e comprensivo di più proprietà e/o fenomeni. Insomma l'analista può, in questa fase, trovare un metodo per *aggregare* i diversi dati e indicatori in modo da poterli trattare simultaneamente e da generare una soluzione soddisfacente e unica. Generalmente a questa fase si può

rispondere con due strade differenti tra loro:

1. utilizzare metodi di elaborazione dei dati e degli indicatori riconducibili all'analisi multivariata e a tutti i modelli di calcolo ad essa afferenti, consultabili ampiamente in letteratura. Si tratta di modelli matematici e statistici che sono in grado, grazie anche al calcolo automatico attraverso il calcolatore elettronico, di processare più variabili contemporaneamente, lasciando all'analista la responsabilità di fissare alcuni criteri di trattamento. L'analisi multivariata risale agli anni '70; in questo trentennio si sono sviluppate nuove teorie (analisi cluster, fuzzy,...) che hanno dato risultati decisamente interessanti in molte applicazioni territoriali;
2. utilizzare semplici metodi di trattamento dei dati e degli indicatori attraverso la preventiva standardizzazione dei valori. Con standardizzazione si intende quella procedura attraverso la quale ogni valore della serie viene trasformato in un numero indice attraverso un'operazione di divisione per un valore rappresentativo della serie. In tal modo si ottengono tanti vettori numerici quante sono le variabili considerate. Ogni vettore è composto da numeri indice confrontabili direttamente; in tal modo viene eliminato il problema delle unità di misura e dei fondi scala differenti, ma non quello delle importanze relative che dovrà essere risolto per altra via. La standardizzazione è un metodo suggerito e utilizzato più recentemente e ha il vantaggio di essere molto semplice ed efficace, oltre che adatto per condurre molte analisi di supporto alle decisioni, come, appunto quelle riconducibili alla disciplina ambientale e territoriale in genere.

Il percorso metodologico che conduce all'aggregazione di più variabili (che siano esse dati o indicatori) passa attraverso alcune fasi di calcolo molto delicate,

senza le quali l'esito dei risultati risulta essere invalidato, o non di meno, scorretto, specialmente nella fase finale dell'aggregazione.

Il processo di trattamento dei dati (indicatori) deve avvenire secondo il seguente percorso:

1. identificazione dei dati utili allo scopo;
2. normalizzazione;
3. standardizzazione;
4. aggregazione.

La fase iniziale di normalizzazione è importante per il trattamento dei dati elementari, in quanto consente di riportare il dato alla dimensione territoriale, poichè spesso si presentano non legati a quest'ultima.

La normalizzazione consiste quindi nella semplice operazione di dividere la misura originale per il numero di abitanti relativo all'area ove il dato è stato raccolto, o per l'estensione territoriale sempre relativa all'area interessata. Generalmente nell'analisi territoriale non vi sono regole valide a priori per condurre la normalizzazione dei dati. Una stessa unità di misura potrebbe essere pesata sia sugli abitanti, ottenendo una dotazione *pro capite*, sia sulla superficie dell'area amministrativa che la comprende tutta o in parte, ottenendo un grado di incidenza relativo.

Dopo aver trasformato i vettori dei dati elementari in indicatori normalizzati, la successiva fase di standardizzazione, consente di poter costruire le basi per trattare simultaneamente più variabili. Il passo è quello della trasformazione degli indicatori in numeri indici, indipendenti dalle rispettive unità di misura, che può avvenire generalmente attraverso la divisione di ogni valore del vettore degli indicatori per una delle seguenti opzioni:

1. il valore migliore della serie ("*the best positioned one*");

2. il valor medio della serie;
3. il valore riferito all'area di riferimento per lo studio.

In tutti e tre i casi si ottengono dei numeri indice che esprimono implicitamente una distanza del singolo valore rispettivamente dalla migliore “*performance*”, dal valore medio della popolazione di casi indagata e dal valore riferito all'area totale di studio.

La fase finale dell'aggregazione consente di ottenere un unico indice in grado di condensare le diverse variabili rappresentate da altrettanti indicatori trasformati a loro volta in numeri indice. Per ottenere ciò si può ricorrere almeno a tre tecniche distinte:

1. la media aritmetica;
2. la media geometrica;
3. la media aritmetica ponderata.

La grandezza finale, genericamente, consiste in un unico vettore-indice di valori che ha caratteristiche di sinteticità ed espressività ottime, ma non altrettanto di analiticità.

1.3 Il ruolo della geostatistica

Come si era accennato nei precedenti paragrafi una problematica per la rilevazione dei dati e la relativa costruzione degli indici è la non buona copertura territoriale (e temporale), tale inconveniente è stato affrontato dalla geostatistica. A partire dalla seconda metà degli anni ottanta i metodi geostatistici hanno cominciato a diffondersi in settori diversi, da quelli più tradizionali dell'ingegneria mineraria e del petrolio, nel cui ambito erano stati messi a punto, a quelli delle scienze della terra. Le discipline coinvolte sono tante, e tra queste

si possono ricordare: la scienza dei suoli, l'idrologia, l'idrogeologia, la geochimica, la meteorologia, l'oceanografia, l'igiene ambientale, l'agronomia, l'analisi delle immagini, etc. Tutte discipline che si collocano in un contesto spaziale o spaziotemporale e che sono alla base dello studio dei fenomeni ambientali che più frequentemente si manifestano sul territorio. Tra le ragioni della diffusione della disciplina se ne possono citare due: l'aspetto pratico e la sua efficacia. Il primo è dovuto al fatto che i metodi geostatistici sono stati messi a punto su problemi specifici, la seconda è testimoniata dal successo che tali metodi hanno avuto nei settori di origine, dove ormai vengono comunemente impiegati nella gestione economica della produzione. Sicché, metodi che hanno sostenuto un siffatto collaudo, sono oggi diventati disponibili per lo studio delle tematiche ambientali. L'aspetto pratico dei metodi non discende però da un approccio empirico. Data la varietà dei temi che ricorrono nei settori di recente applicazione, i metodi geostatistici sono in evoluzione e si arricchiscono sempre di più.

La caratterizzazione di un fenomeno spaziale o temporale o spaziotemporale costituisce il primo e il più importante passo di uno studio geostatistico. Essa consiste essenzialmente nell'evidenziare, in forma qualitativa e quantitativa, la variabilità del fenomeno in esame, specificandone la tipologia, in relazione alla presenza di eventuali anisotropie e all'esistenza di diverse scale di variabilità, spaziali o temporali. Per esempio nel campo temporale si possono mettere in evidenza scale di variabilità stagionali di fenomeni di inquinamento o scale di variabilità stagionali e climatiche di fenomeni idrogeologici e meteorologici, descritti da variabili quali piezometrie, precipitazioni, temperature, etc. Il comportamento mutuo delle variabili può essere diverso per le diverse scale temporali, così come diverso può apparire, rispetto alla variabilità temporale, il

comportamento dei siti delle stazioni di misura, in ragione dei loro aspetti fisici, quali quota, morfologia, esposizione, distanza dal crinale, distanza dal mare nel caso dei fenomeni meteorologici, e porosità e trasmissività del mezzo nel caso dei fenomeni idrogeologici. In campo spaziale si possono mettere in evidenza componenti locali e regionali della variabilità, per esempio del contenuto in metalli pesanti dei suoli o di altri inquinanti, mostrando come e quanto, in ogni componente, la struttura spaziale della variabilità sia legata alle fonti di inquinamento e alle caratteristiche ambientali, quali la morfologia, la litologia, la meteorologia, etc. Analogamente si può fare per le caratteristiche fisicochimiche o idrauliche dei suoli, o per le caratteristiche geochimiche delle acque e delle sostanze inquinanti in esse disciolte.

2. METODOLOGIA PER LA RACCOLTA E LA CATALOGAZIONE DELLA DOCUMENTAZIONE

Come primo passo per la raccolta e la catalogazione della documentazione, è stata effettuata un'attenta analisi sulla problematica in questione e su tutti gli aspetti che la caratterizzano. Secondo quanto descritto nei paragrafi precedenti, si è deciso di recensire tali metodologie e creare una scheda di sintesi (fig.2.1). Tale scheda è stata progettata tenendo conto delle seguenti esigenze:

- dover catalogare le metodologie organizzandole in modo da poterle poi rendere disponibili per successive consultazioni in modo semplice ed efficace;
- prevedere la possibilità di modificare ed aggiornare i contenuti della scheda;
- stabilire criteri per ricercare ed estrarre metodologie di interesse (tipo di metodologia analitica, fonte della metodologia, ecc...);
- tenere conto del fatto che, non sempre è stato possibile reperire il testo delle metodologie, in alcuni casi risulta infatti disponibile solo il riferimento bibliografico;
- i testi reperiti delle metodologie sono in formati diversi tra loro (cartaceo, file, ecc..);
- alcuni testi di metodologie erano disponibili solo su supporto cartaceo.

Successivamente si è provveduto a ricercare le metodologie di aggregazione

degli indicatori ambientali, disponibili presso fonti nazionali ed internazionali. Sono state globalmente raccolti 94 articoli e/o pubblicazioni, provenienti rispettivamente dalle seguenti fonti:

AEA Agenzia Europea dell'Ambiente (v. anche EEA)
AIEA Agenzia Internazionale per l'Energia Atomica
AM Aeronautica Militare
ANPA Agenzia Nazionale per la Protezione dell'Ambiente (ora APAT)
APPA Agenzia Provinciale per la Protezione dell'Ambiente
ARPA Agenzia Regionale per la Protezione dell'Ambiente
CE (EC) Commissione Europea (European Commission)
CEE Comunità Economica Europea
CNR Consiglio Nazionale delle Ricerche
CTN Centro Tematico Nazionale
EEA European Environment Agency
EIONET (rete informativa ambientale dell'AEA)
ENEA Ente per le Nuove tecnologie, l'Energia e l'Ambiente
ETC European Topic Centre (Centro Tematico Europeo)
EUROSTAT (Ufficio di Statistica della Commissione Europea)
FAO Food and Agriculture Organization of the United Nations
ICDM Servizio Difesa del Mare del Ministero dell'ambiente
ICRAM Istituto Centrale per la Ricerca sull'Ambiente Marino
IRSA Istituto Ricerca sulle Acque
ISS Istituto Superiore di Sanità
ISTAT Istituto Nazionale di Statistica
OCSE Organizzazione per la cooperazione e lo sviluppo economico
ONU Organizzazione delle Nazioni Unite
SINAnet Rete del Sistema Nazionale conoscitivo e dei controlli in campo ambientale
UE (EU) Unione Europea (European Union)
UNCSD United Nations Commission on Sustainable Development
US-EPA Environmental Protection Agency (Agenzia per l'Ambiente statunitense)

Nella maggior parte dei casi è stato possibile reperire il testo integrale delle metodologie, in alcuni casi direttamente su supporto informatico, in altri casi su supporto cartaceo.

Si è quindi proceduto alla catalogazione delle metodologie reperite in funzione dei seguenti criteri, presenti nella scheda:

- 1) la data della ricerca effettuata;
- 2) il tipo di ricerca effettuata (bibliografica, telematica o altro), l'argomento, le parole chiavi utilizzate e le possibili correlazioni;
- 3) l'autore ed il riferimento bibliografico;
- 4) disponibilità e reperibilità del documento (biblioteche, supporto informatico, ecc..), eventuale sito web rilevante;
- 5) campo di applicazione, comparto sul quale la metodologia è applicabile (acque, suolo, ecc...).

Gli articoli relativi all'aggregazione degli indicatori ambientali, problematica e argomento nuovissimo, ancora in fase sperimentale, non sono molti, quindi si è posta l'attenzione su quelli inerenti a metodologie che si sono ritenute più interessanti (analizzate nei successivi paragrafi).

A.P.A.T. – Agenzia per la Protezione dell’Ambiente e per i servizi Tecnici			
RICERCA			
Data della Ricerca effettuata: _ _ / _ _ / _ _ _ _			
Tipo di Ricerca: Bibliografica ? Telematica ? Altro ?			
Argomento della Ricerca:			
Parole Chiavi Utilizzate:			
Correlazioni Possibili:			
RISULTATI			
Autore:			
Titolo:			
Anno:		Pagine:	
Tipologia:			
ESITO			
Posseduto ? Non Posseduto ? Non Disponibile ? Individuato in altre Biblioteche ?			
Indirizzo Web rilevante:			
Campo di Applicazione:			
Atmosfera	?	Biosfera	?
Idrosfera	?	Geosfera	?
Rifiuti	?	Radiazioni Ionizzanti	?
Radiazioni non Ionizzanti	?	Rumore	?
Rischio Antropogenico	?	Rischio Naturale	?
Agricoltura	?	Energia	?
Trasporti	?	Turismo	?
Qualità Ambientale	?	Controlli	?

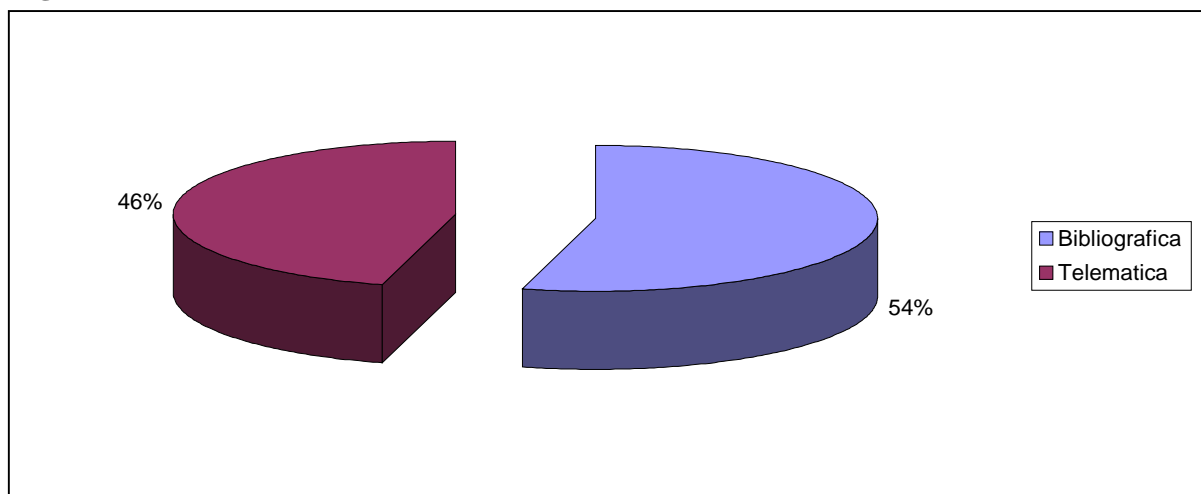
Fig. 2.1: Esempio di scheda di rilevazione e catalogazione

3. RISULTATI DELLA RICERCA SUI METODI DI AGGREGAZIONE STATISTICA DEGLI INDICATORI AMBIENTALI DI ORGANISMI NAZIONALI E INTERNAZIONALI

Tabella 3.1: Tabella di presentazione dei dati per categoria di soggetti

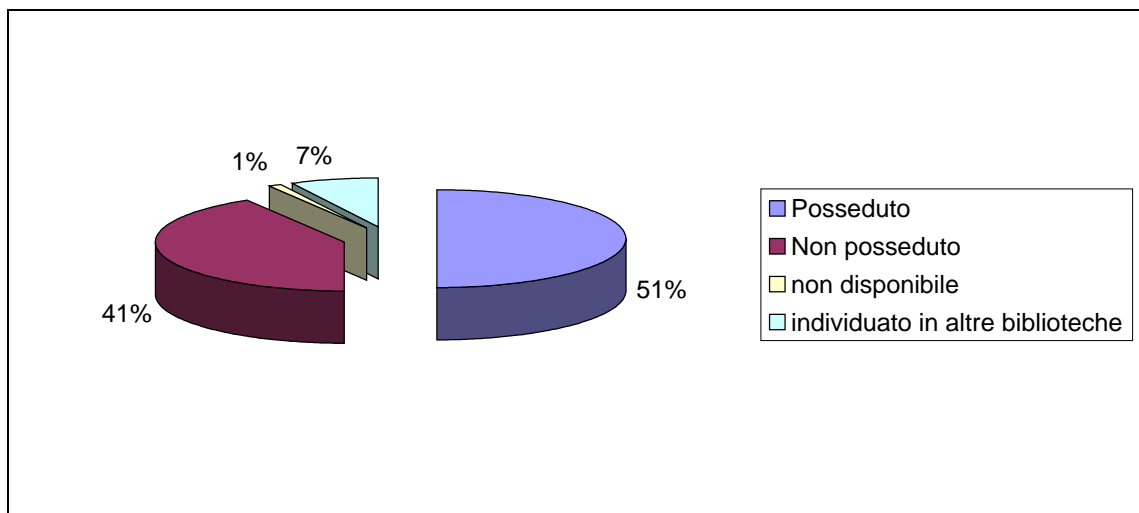
Soggetto	Documento html/sito	Documento wp/sito	Documento pdf/sito	Pubblicazioni	Libri	Manuali	Atti/ Working Paper	Totali
EPA	7	1	1	0	0	0	0	0
EEA	2	0	17	0	0	0	0	19
Studiosi/Ricercatori	0	0	0	37	1	1	6	45
Fao	1	0	0	0	0	0	0	1
UNDP	1	0	0	0	0	0	0	1
UNEP	1	0	0	0	0	0	0	1
IISD	0	0	0	2	0	0	0	2
Com Europ Amb	0	0	2	1	0	0	0	3
UN	1	0	0	0	0	0	0	1
Enea	0	0	0	1	0	0	0	1
Eurostat	1	0	0	1	0	0	2	4
OECD	2	0	0	0	0	0	0	2
CNR	0	0	0	5	0	0	0	5
Totali	16	1	20	47	1	1	8	94

Fig. 3.1: Distribuzione percentuale dei documenti trovati per tipo di ricerca effettuata



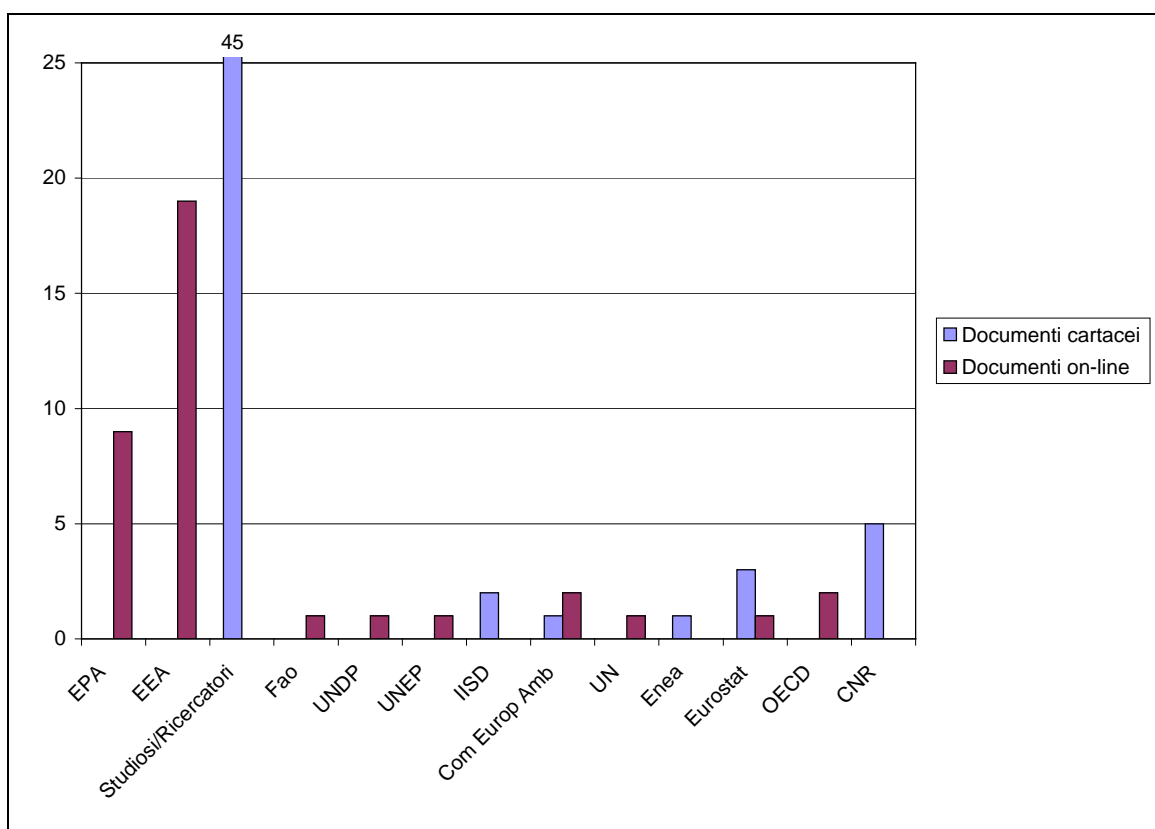
Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003

Fig. 3.2: Distribuzione percentuale delle pubblicazioni per reperibilità



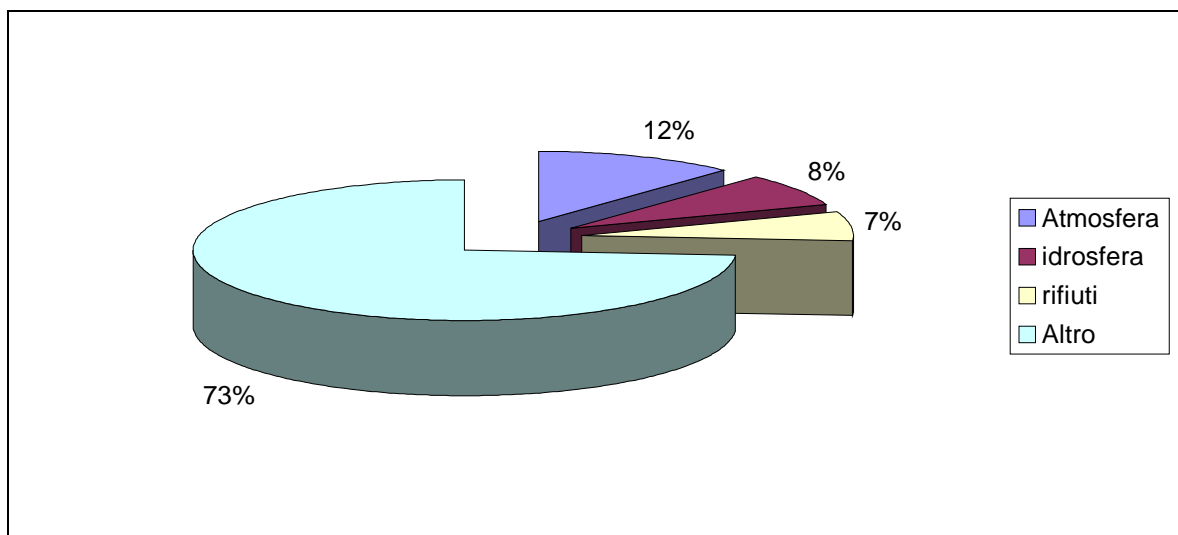
Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003

Fig. 3.3: Numero di lavori pubblicati dagli organismi nazionali e internazionali



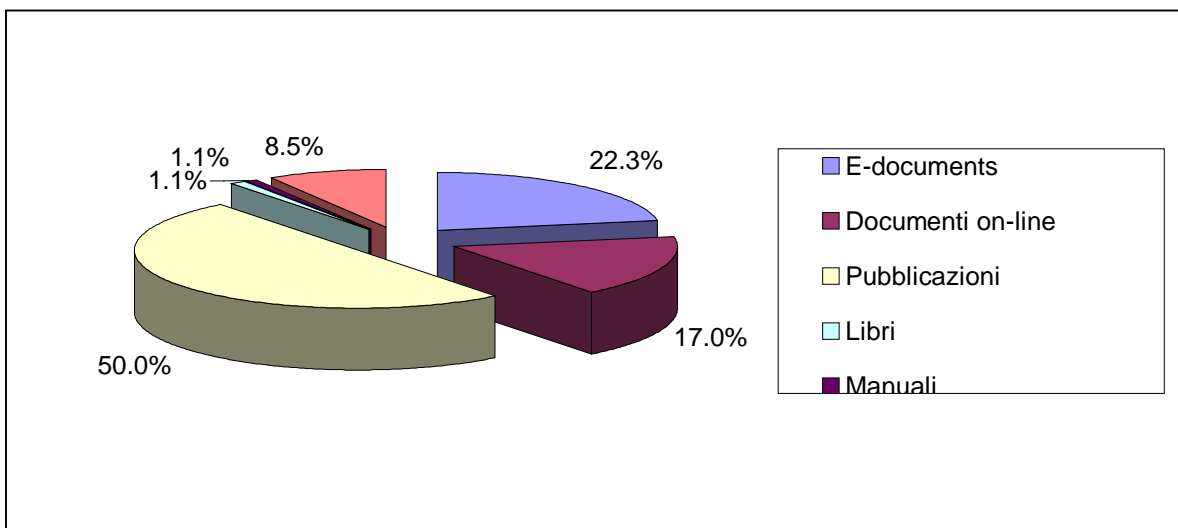
Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003

Fig. 3.4: Distribuzione percentuale delle pubblicazioni sui metodi statistici di aggregazione degli indicatori ambientali, reperiti da organismi nazionali e internazionali, per aree tematiche



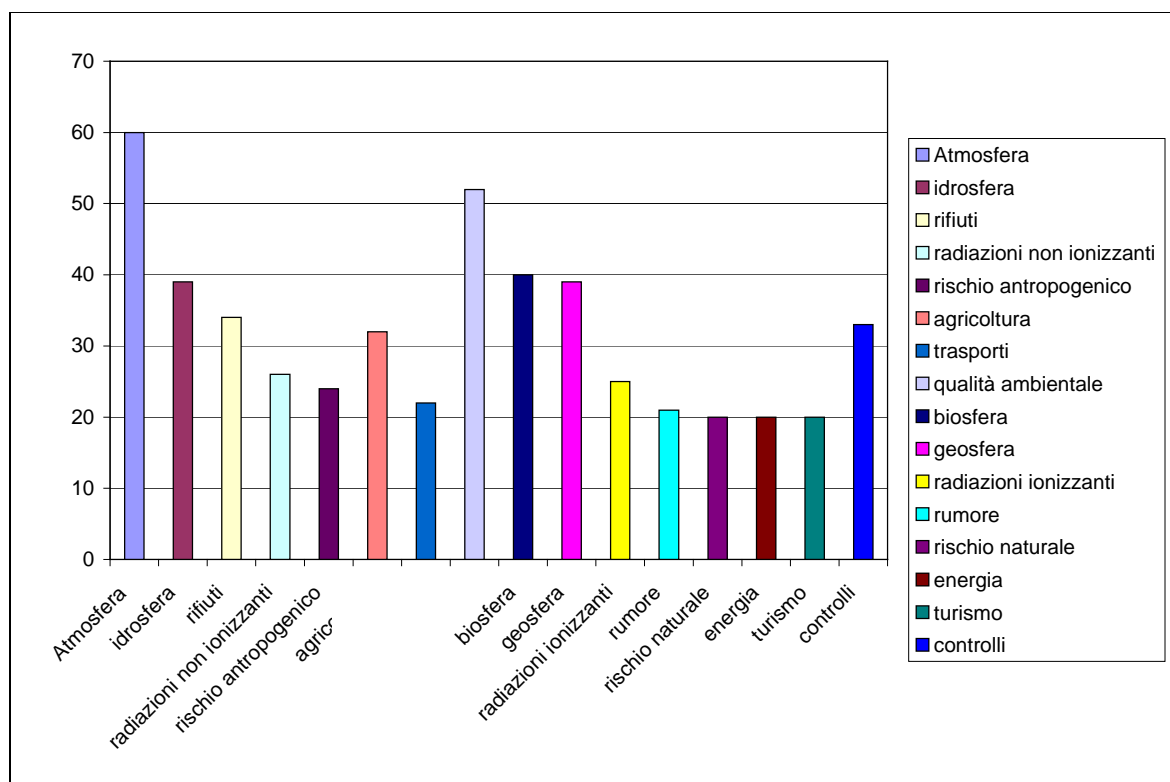
Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003

Fig. 3.5: Distribuzione percentuale delle pubblicazioni sui metodi statistici di aggregazione degli indicatori ambientali, reperiti da organismi nazionali e internazionali, per tipologia di supporto



Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003.

Fig. 3.6: Numero di articoli reperiti per aree tematiche.



Fonte: Elaborazione dei dati reperiti dai siti web - dicembre 2003.

4. AGGREGAZIONE INDICI QUALITA' DELL'ARIA

L'interesse nei confronti della costruzione di indici per la qualità dell'aria è testimoniato da molti anni di dibattito, a partire dai lavori fondamentali di Ott e Hunt (1976) e Ott (1978), fino ai contributi più recenti di Hamikoski (1998), Swamee e Tyagi (1999), Kassomenos *et al.* (1999).

La costruzione di un indice sintetico si articola secondo i seguenti criteri (fig.4.1): si deve analizzare attentamente il fenomeno e scomporre in singoli aspetti; devono essere definiti indici elementari di partenza e da questi estrarre i sub-indici; si deve proporre l'ordine con cui aggregarli (si devono scegliere sia una funzione di aggregazione sia i procedimenti per togliere l'unità di misura).

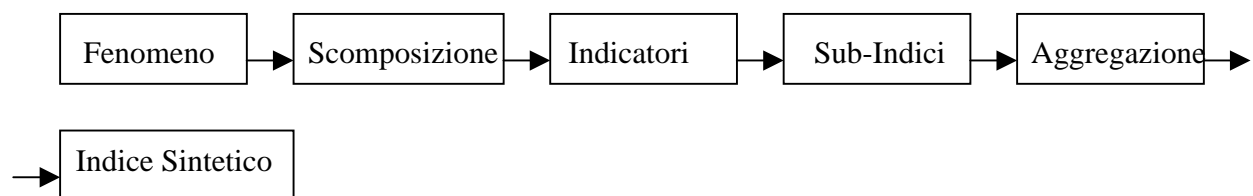


Fig. 4.1: Diagramma della costruzione di un indice sintetico (Fraire, 1989; Agostinelli et al., 1996)

Le istituzioni che hanno l'obbligo di raccogliere e fornire ai cittadini informazioni sulla qualità dell'aria usano metodi e strategie differenti, seppure tutte con l'obiettivo comune di offrire valutazioni sulla qualità dell'aria legate agli effetti sulla salute (Colombo, 1996).

Le differenze negli indici che vengono proposti sono dovute al fatto che, per la loro costruzione, è stata utilizzata una diversa metodologia sia per la standardizzazione che per l'aggregazione. La standardizzazione è resa necessaria dalla diversa unità di misura o dal diverso ordine di grandezza degli inquinanti.

L'aggregazione è dovuta alla molteplicità degli inquinanti, al numero delle

postazioni in cui sono rilevati e al ripetersi delle rilevazioni nel tempo.

4.1 Progetto sulla costruzione di indici per la qualità dell'aria (Bruno F. – Cocchi D.)

In letteratura, gli orientamenti per la definizione degli indici sono quelli delle medie analitiche, ponderate ed eventualmente potenziate (Swaame e Tyagi, 1999) e quelli delle medie di posizione (Thomas, 1972; Attanasio e Capursi, 1997). Nella pubblicazione di Bruno-Cocchi (2000) si privilegiano le medie di posizione. Le ragioni di tale scelta risiedono nelle proprietà di tali medie, più robuste di quelle analitiche, in caso di dati con qualità non sempre sotto controllo, e nella difficoltà di proporre elevamenti a potenza e ponderazione con rilevanza non solo matematica, ma anche sostanziale e interpretativa.

In questo lavoro viene proposta una descrizione del processo di aggregazione dei dati per costruire gli indici di qualità dell'aria. Il processo, che si svolge attraverso fasi successive di aggregazione, può richiedere l'eliminazione dell'unità di misura del fenomeno e permette di mettere in rilievo quando siano possibili scelte alternative, di verificare in quali casi l'ordine di aggregazione influenzi il risultato finale e di valutare a che punto sia necessaria una standardizzazione.

Per la standardizzazione proposta viene utilizzata, per ogni inquinante considerato, una classificazione in base a soglie di rischio per la salute. La classificazione è, in genere, dovuta al contributo di esperti, ed è esterna e precedente alla costruzione dell'indice. La scelta di una particolare classificazione delle concentrazioni degli inquinanti in base a soglie "critiche" di inquinamento influenza il comportamento dell'indice e l'interpretazione dei suoi risultati. Se si usano in modo flessibile le informazioni raccolte calcolando

diverse sintesi, si traggono comunque indicazioni utili anche subordinatamente alla scelta effettuata.

Poiché il procedimento di aggregazione non è unico, può essere utile presentare l'indice finale affiancandolo con l'indicazione della postazione o dell'inquinante che l'hanno determinato.

Ai vari passi del processo di aggregazione può essere proposta più di una sintesi. Sfruttando questa opportunità, si possono calcolare simultaneamente indici che controllano il comportamento dei valori più alti e dei valori medio bassi. Il comportamento dell'indice (ed eventualmente di più indici congiuntamente) può essere studiato, in particolare, in corrispondenza di valori medio-bassi, per verificare la presenza di innalzamenti dei valori di fondo degli inquinanti.

L'utilizzazione contemporanea di più indici permette inoltre di dare risalto ad aspetti complementari dell'inquinamento. In particolare un certo tipo di sintesi mette in evidenza la presenza di episodi acuti di inquinamento isolati nel territorio, mentre la sintesi complementare dà rilievo all'eventualità della diffusione spaziale di valori elevati di inquinamento.

A fianco della proposta di impiegare più di una sintesi, suggeriamo di utilizzare le informazioni in modo da tenere conto della dispersione dei valori rilevati in fasi particolari del processo di aggregazione, in quanto, nella presentazione degli indici sintetici sulla qualità dell'aria, solitamente non si considera la variabilità dei valori che vengono sintetizzati.

Proponendo l'uso simultaneo di più indicatori, sembra di avere trovato una via che permetta di seguire in modo continuo l'evoluzione dell'inquinamento, anche in presenza di spostamenti delle postazioni.

Il primo passo è quello di ridurre la dimensionalità tramite un processo di aggregazione temporale che permetta di ricavare per ciascun inquinante un

valore giornaliero a partire dai valori orari rilevati. Successivamente l'aggregazione può essere svolta per:

- 1) postazione-inquinante (utilizzando una funzione lineare a tratti) che considera per ultima l'aggregazione degli inquinanti e permette di ritardare la standardizzazione;
- 2) inquinante-postazione (tramite funzione lineare) che considera per ultima l'aggregazione spaziale, dove è necessario calcolare prima un vettore standardizzato per ciascuna postazione, ottenendo un indice sintetico di postazione su tutti gli inquinanti.

Come visto in fig. 4.1, la letteratura non sembra tenere conto esplicitamente delle tre fasi e parla, senza distinzioni, di calcolo di sub-indici, dando rilievo quasi esclusivamente all'aggregazione dei dati relativi ad una singola postazione. Invece le sintesi intermedie hanno significati molto diversi e il valore finale, proposto varia al variare dell'ordine di aggregazione, così come non sono uguali le sintesi finali ottenute tramite il calcolo di mediane. I valori standardizzati degli indici di qualità dell'aria possono essere risultato sia di aggregazioni intermedie sia di sintesi finali. La qualità dell'aria può essere infatti valutata nell'ambito di un'aggregazione parziale o al momento della sintesi in un unico valore.

I valori soglia dell'indice standardizzato che definiscono le classi di qualità dell'aria costituiscono i punti di riferimento più importanti. Nelle applicazioni proposte, il valore 100 discrimina tra qualità buona e accettabile, il valore 200 discrimina tra qualità accettabile e scadente e il valore 300 tra scadente e pessima.

Quando il giudizio sulla qualità dell'aria è basato sull'impiego contemporaneo di

più indici, non consiste soltanto nell'attribuzione del livello di qualità ad una classe, ma anche nella valutazione della posizione relativa dei diversi indici.

Proponendo l'uso simultaneo di più indicatori, si è trovata una via per seguire in modo continuo l'evoluzione dell'inquinamento anche in presenza di spostamenti delle postazioni di rilevazione.

La scelta tra i diversi indicatori è governata dall'obiettivo finale, che va dalla più tradizionale segnalazione di episodi acuti all'individuazione di tendenze di fondo. Il comportamento congiunto degli indici proposti può evidenziare se si stiano verificando innalzamenti dei valori di fondo degli inquinanti. La scelta di una particolare classificazione in base a soglie "critiche" di inquinamento influenza le valutazioni complessive. Tuttavia, se si usano le informazioni raccolte in modo flessibile calcolando sintesi diverse, si traggono comunque informazioni utili anche subordinatamente alla scelta effettuata.

L'utilizzazione contemporanea di indici che seguono i due percorsi di aggregazione permette di evidenziare aspetti complementari dell'inquinamento. In particolare la sintesi postazioni-inquinanti dà risalto alla presenza di episodi acuti di inquinamento isolati nel territorio, mentre la sintesi inquinanti-postazioni dà rilievo all'eventualità della diffusione spaziale di valori elevati di inquinamento.

L'esame congiunto degli indici di qualità dell'aria ottenuti secondo i diversi procedimenti di aggregazione e delle corrispondenti misure di dispersione fornisce un quadro completo che può essere usato per confronti nel tempo e nello spazio.

4.2 Modelli statistici per dati mancanti nel monitoraggio e previsione dell'inquinamento atmosferico (Progetto di Ricerca CNR 1999)

(Statistical Models for Missing Data: Monitoring and Forecasting Air pollution)

Dall'articolo preso in esame si evince che la ricerca effettuata prende in considerazione i dati ad alta frequenza rilevati dalle centraline della Regione Lombardia, con particolare riferimento all'ozono. Ciò perché, mentre alcuni inquinanti come il biossido di zolfo ed il monossido di carbonio sono diminuiti per l'evoluzione di combustibili industriali e da riscaldamento (Ballarin Denti, 1998 e Cantagallo, 1997), altrettanto non accade per gli ossidi di azoto e per l'ozono. In particolare quest'ultimo costituisce un'emergenza sia in numerose città che in zone rurali e montane ed ha registrato numerosi episodi acuti di allarme nella regione Lombardia (cfr Bolasso et al. 1997).

Negli ultimi anni si è avuto un grande sviluppo di modelli statistico-matematici sul problema del monitoraggio e della previsione della qualità dell'aria. In particolare le classiche tecniche di previsione delle serie storiche si sono evolute dai più semplici modelli lineari tipo ARMA a modelli più sofisticati come i modelli a scatola grigia (cfr. Finzi e Brusasca, 1991), le approssimazioni locali (cfr. Lapenna et al. 1966) i modelli con nonlinearità dipendenti dallo stato (Young, 1995) i modelli eteroschedastici (cfr. Graf e Jaunin, 1998,) ed i modelli di reti neurali artificiali (cfr. Bordignon et al. 1997).

Promettenti anche i modelli Markoviani latenti per la previsione dei cambiamenti di regime (I. L. MacDonald, \ W. Zucchini, 1997).

Quando si considerano i dati ad alta frequenza emerge, fra gli altri, il problema dei dati mancanti in quanto spesso gli strumenti delle centraline hanno delle fermate per taratura e/o per malfunzionamento. In questo studio si utilizzano le

tecniche per i dati mancanti nell'identificazione e stima del modello statistico. Da un lato confrontando quelle basate su interpolazione (Karanta, 1993), quelle basate su modelli previsivi (Nijman e Palm, 1990, Tiao e Pena, 1991 e Beveridge, 1992) e sul filtro di Kalman (Corradi e Guagnano, 1993) sia in ambito lineare che nonlineare (Bovas e Thavaneswaran, 1991) e, in ambito nonparametrico, su stimatori kernell (Cheng e Philip 1990) nel caso concreto in studio (Lee et al.). Dall'altro studiando l'influenza dei dati mancanti sul modello risultante con particolare riferimento alla previsione degli eventi di interesse.

Inoltre si studia da un punto di vista metodologico l'influenza dei dati mancanti sui test di validazione (Stoffer e Tolo, 1992) e di linearità nonché di cross validazione (Fassò, 1997d). Prendendo in considerazione sia i classici test tipo Ljung-Box che test non parametrici tipo Hallin *et al.* (1988) e Fassò (1997-a e - b). La metodologia utilizzata fa riferimento ai dati osservati e valuta, quindi, i diversi metodi nello specifico. Inoltre si considerano delle simulazioni stocastiche dei modelli considerati per evidenziare come le diverse metodologie si comportano in determinate situazioni controllate.

4.3 Metodi statistici per l'analisi dell'ambiente e delle interazioni ambiente-salute (Colombi R. – Lovison G.)

In tale ricerca si sono studiati, messi a punto ed applicati a casi concreti vari metodi e modelli statistici per lo studio dell'inquinamento. Particolare attenzione si è posta all'inquinamento dell'aria (ozono) nella zona di Bergamo.

Tutti i punti centrali del progetto sono stati variamente trattati. Questi sono:

I. previsione in tempo reale;

II. monitoraggio statistico;

III. diffusione dell'inquinamento;

IV. software.

La modellistica sviluppata per la descrizione e la previsione dell'inquinamento atmosferico si articola in tre tipologie:

M1: modelli non lineari di tipo autoregressivo;

M2: modelli markoviani latenti;

M3: modelli logit con cambiamenti di regime markoviani.

Durante la parte di validazione del modello è spesso necessario utilizzare un set di dati diverso da quello usato per la stima e valutare le proprietà di adattamento messe in luce dai residui (indipendenza, normalità etc.) su questo secondo data set.

Per i modelli M1 l'obiettivo è avere un modello globale, valido per tutta la serie storica considerata, con un comportamento particolarmente buono anche e soprattutto quando il problema ambientale è più grave, cioè nella coda destra della distribuzione, altamente asimmetrica, dell'inquinante.

A tal fine si sono utilizzati modelli non lineari del tipo autoregressivo stagionale a soglia endogena, con memoria lunga, variabili esogene, date da grandezze meteorologiche, altri inquinanti collegati (precursori) e alta eteroschedasticità condizionale. La sigla che definisce questo tipo di modelli è SFI-SETARX-ARCH che sta per *Seasonally Fractionally Integrated - Self Exciting Threshold AutoRegressive with eXogenous - AutoRegressive Conditionally Heteroscedastic*. Nell'applicazione ai dati orari dell'ozono a Bergamo, negli anni dal 1993 al 1997, si sono usati modelli con due soglie, cioè tre diverse

approssimazioni lineari rispettivamente per il basso, medio e alto inquinamento. Dei tre anni (93-95) usati per il *fitting*, in particolare il 1995 è stato un anno di alti livelli dell'ozono estivo. Ciò ha consentito una buona stima della dinamica nella coda destra che normalmente è composta di eventi rari. La stagionalità è da intendersi in senso lato di periodicità giornaliera, settimanali, mensili e, propriamente, annuali. Emergono dei modelli che hanno buone proprietà di adattamento e buone capacità di previsione.

In particolare si è sviluppata la parte metodologica relativa alla modellistica e agli algoritmi di stima dei parametri nonché la parte applicativa di identificazione e stima nel caso concreto di cui sopra. Si sono potuti valutare quindi i vantaggi del modello M1 adottato rispetto ai più tradizionali modelli lineari per serie storiche.

Inoltre si sono sviluppati gli algoritmi per il calcolo Monte Carlo del revisore ottimale a più passi che, a causa della forte non linearità, non può essere approssimato dal previsore *naïve* ottenuto dall'iterazione di più previsori ad un passo.

L'applicazione ai dati mostra l'utilità dei modelli di tipo M1 per la previsione a più passi rispetto ai più semplici modelli lineari. In entrambi i lavori, di cui sopra, emerge, in particolare, che l'approccio utilizzato dà vantaggi tanto più rilevanti in termini di adattamento e di previsione quanto più alti sono i livelli di inquinamento. Inoltre, per i primi tre quartili di basso e medio inquinamento dà dei risultati simili, e non inferiori, ai modelli lineari classici.

Proseguendo nella logica della modellazione diretta, si sono considerati dei modelli tempo varianti eteroschedastici, basati sull'applicazione dei minimi quadrati ricorsivi sia all'equazione del previsore che della sua varianza. I risultati mostrano che questo approccio cattura la forte non linearità stagionale del

fenomeno considerato anche dopo aver scontato l'eteroschedasticità.

I modelli markoviani latenti gaussiani (MMLG: M2) descrivono l'evoluzione nel tempo di una o più variabili osservate, la cui dinamica si pensa sia influenzata dalla dinamica di altre variabili che non sono osservabili. Il modello statistico tiene quindi conto sia delle prime che delle seconde; esso viene definito attraverso una coppia di processi stocastici a tempo discreto, $(Y_t; X_t)$, dove (X_t) è una catena di Markov non osservata con un numero finito di stati e (Y_t) , data (X_t) , è una successione di variabili casuali osservate condizionatamente indipendenti, con la distribuzione di ogni variabile Y_t che dipende dalla successione (X_t) solo attraverso la singola variabile X_t , cioè solo dallo stato in cui si trova la variabile non osservata, nell'istante considerato. Come probabilità condizionata dallo stato si è ipotizzata la gaussiana.

L'attenzione è stata rivolta alla stima di massima verosimiglianza dei parametri del modello, rappresentati dalla matrice delle probabilità di transizione e dal vettore delle probabilità iniziali della catena di Markov e dai parametri della particolare distribuzione ipotizzata, per ogni stato della catena di Markov. Data la complessità analitica della funzione di verosimiglianza e la presenza nel modello di una variabile non osservata, non è possibile derivare direttamente, per via analitica, gli stimatori di massima verosimiglianza. Il metodo numerico comunemente utilizzato in questi casi è l'algoritmo EM (Expectation-Maximization) che viene qui applicato al modello in studio. In tale caso si dimostra che l'applicazione iterativa dei due passi dell'algoritmo EM porta a definire le espressioni esplicite degli stimatori di massima verosimiglianza dei parametri incogniti.

Tali modelli sono stati utilizzati per analizzare serie temporali riguardanti la

qualità dell'aria a Bergamo.

Nel lavoro si fa riferimento a nuove versioni delle formule esplicite degli stimatori dei parametri dei MMLG; questi nuovi stimatori consentono l'analisi anche di serie temporali con osservazioni mancanti.

Dal punto di vista delle decisioni che devono essere prese dagli amministratori in relazione ad episodi di inquinamento da ozono spesso più che l'esatto valore della concentrazione di O_3 è interessante sapere se saranno superati i livelli di attenzione o di allarme.

Per questo motivo e, perché si ritiene che dal punto di vista dell'amministratore è utile saper prevedere non tanto l'esatto valore della concentrazione di O_3 ma l'intervallo di valori più probabile, si è studiata una classe di modelli Logit Dinamici il cui scopo è quello di assegnare, alla luce delle informazioni inerenti, i livelli di concentrazione precedenti e di alcune variabili meteo, una probabilità ai livelli di attenzione e allarme della concentrazione di O_3 per l'istante temporale successivo.

La classe di modelli è stata specificata assumendo che condizionatamente ad una catena di Markov a due stati non osservabile i livelli di O_3 siano a loro volta una catena di Markov di ordine maggiore o uguale ad uno.

La catena di Markov non osservabile è stata introdotta per modellare il fatto che tipicamente le serie storiche della concentrazione di O_3 al suolo mostrano due differenti dinamiche: una, che potremmo definire di rischio, tipica dei periodi estivi e caratterizzata da alti livelli di concentrazione e una, tipicamente invernale, con bassi livelli di concentrazione. Modellare la transizione tra queste due dinamiche è appunto la funzione della catena di Markov latente.

I differenti modelli Logit che sono stati specificati differiscono, oltre che per il tipo di Logits utilizzato (locale, di continuazione, globale), per l'insieme di

parametri che dipendono dagli stati della catena di Markov non osservabile.

L'approccio utilizzato generalizza sia l'usuale approccio di modellare serie storiche categoriche con catene di Markov sia il più recente approccio basato sui modelli noti come Hidden Markov Models. Nell'applicazione fatta si sono usati dati inerenti massimi giornalieri di concentrazione O_3 , relativi al periodo 93-97, e la capacità di prevedere i livelli di attenzione e allarme che è stata verificata per svariate specificazioni del modello logit dinamico.

In particolare si è evidenziato come i modelli proposti hanno una capacità di previsione decisamente superiore a quella dei tradizionali modelli basati su catene di Markov e a quelli noti come Hidden Markov Models.

Siccome la performance di previsione dei modelli utilizzati dipende fortemente dal tipo di parametrizzazione usata si è anche dedicata notevole attenzione alle proprietà delle diverse parametrizzazioni Logit (locale, di continuazione, globale) rilevanti nelle applicazioni a dati ambientali.

La modellazione diretta di tipo M1 consente poi varie tecniche di monitoraggio statistico. Fra queste ricordiamo quelle orientate al riconoscimento di errori strumentali e quelle orientate al riconoscimento di trend e dinamiche non gradite (per esempio l'aumento dei livelli medi, l'aumento dei massimi, etc.). Infatti, caratteristica dell'inquinamento ambientale, in generale, e dell'aria, in particolare, è che questa si deteriora man mano che le concentrazioni di una o più sostanze aumentano. Nonostante ciò la legislazione fa in genere riferimento alle singole sostanze più che al loro mix. Alcuni tentativi di trovare degli indici sintetici di qualità dell'aria sono stati fatti dalle unità di Bologna e Palermo.

Nello studio dell'inquinamento ambientale sono di grande importanza complessi modelli di simulazione basati sull'uso intensivo del calcolatore che, benchè siano

basati su modelli (fisico/chimico-) matematici, sono di difficile trattazione analitica.

Si pensi ad esempio:

1. ai modelli atmosferici per la previsione delle grandezze meteorologiche e, conseguentemente, la previsione dell'inquinamento atmosferico;
2. ai modelli di diffusione dell'inquinamento;
3. ai modelli idrologici di deflusso ed al loro uso per la valutazione dell'inquinamento dei corsi d'acqua.

Lo studio dell'influenza dei "parametri" di input sul comportamento di questo tipo di modelli prende il nome di analisi di sensitività (*sensitivity analysis of computer models*). In talune situazioni questa si basa sulla sostituzione del modello originario usato per il calcolo con una semplificazione detta modello "surrogato", data per esempio da un modello di regressione costruito su base statistica.

4.4 Metodi statistici per lo studio e il monitoraggio dell'ambiente (Fassò A. – Progetto CNR 2003)

Fra le applicazioni più importanti della statistica allo studio dell'ambiente si trovano il campionamento di dati spesso raccolti “manualmente” e il monitoraggio di dati rilevati in automatico da apposita strumentazione. Le due problematiche, fortemente connesse, anche se spesso trattate separatamente, sono affrontate entrambe in questo progetto di ricerca.

Per quanto riguarda i problemi di campionamento, Skalski (1994) si è occupato di disegni campionari stratificati complessi che combinano, per esempio, sia il campionamento areale che tecniche di campionamento probabilistico. L'UR di Siena si muove in questa direzione occupandosi delle stime derivate dagli

stimatori per rapporti o da regressioni localmente smorzate.

Alla seconda problematica, collegata con il monitoraggio e l'analisi delle serie storiche e spazio-temporali in generale si riferiscono le altre UR.

I modelli per le serie storiche ambientali considerati utilizzano dati ad alta frequenza: giornalieri, orari etc.. In questo ambito Graf-Jaccottet (1997) ha discusso l'importanza della modellistica eteroschedastica e Fassò - Negri (2000a e b) e Bauer et al. (2001) i limiti dei tradizionali modelli lineari. L'UR di Bergamo prosegue la ricerca sui modelli eteroschedastici non lineari.

Inoltre l'UR di Bergamo si occupa di inferenza su ordinamenti stocastici per catene di Markov. Le ipotesi analoghe a Colombi e Forcina (2000a,b) sono rilevanti nel monitoraggio della qualità dell'aria. Punti di partenza sono Kaijima (1997) che ha discusso il ruolo degli ordinamenti stocastici per le catene di Markov e Fahrmeir e Kaufman (1987) che hanno adattato i risultati per l'analisi delle tabelle di contingenza alle catene di Markov. Partendo da McCulloch & Searle (2000) l'UR di Bologna considera l'applicazione ambientale dei GLM⁴ a dati non gaussiani come valori estremi o conteggi e dei GLM misti che, tramite la tecnica degli effetti casuali, permettono di modellare congiuntamente la variazione spaziale e temporale di fenomeni ambientali.

I modelli spazio-temporali autoregressivi e geostatistici dinamici (Christensen, 1997, Christakos e Hristopulos, 1998, Wackernagel et al. 2001) hanno ricevuto una notevole attenzione di recente. L'UR di Roma considera problemi di scelta del modello e identificazione automatica, tecniche di riduzione della dimensionalità e studio dell'associazione in processi spazio-temporali multivariati. In termini generali gli obiettivi principali del progetto sono:

⁴ Modelli Lineari Generalizzati

1. aggiornare ed estendere gli strumenti attualmente disponibili in campo ambientale per l'analisi ed il monitoraggio delle serie storiche ad alta frequenza e lo studio della dinamica spazio-temporale delle grandezze ambientali;
- 2 sviluppare delle significative applicazioni che illustrino l'uso e i limiti dei metodi nello studio di casi concreti anche con riferimento a problemi di campionamento;
- 3 sviluppare il software opportuno, almeno a livello di prototipo, per risolvere i problemi di calcolo connessi ai metodi proposti.

Maggior dettaglio viene di seguito illustrato usando le consuete classificazioni:

1- modelli spazio-temporali. Obiettivo di scelta del modello e problemi di stima; confronto tra modelli di tipo autoregressivo (modelli STAR, STARMA) e modelli di tipo geostatistico dinamico con particolare riguardo ai criteri di scelta automatica; tecniche esplorative per serie spazio-temporali multivariate. Estensione di tecniche esplorative già sviluppate nel contesto puramente spaziale (Jona-Lasinio 1999) al caso spazio-temporale;

2- modelli per serie storiche. Applicazione a metodi di previsione e monitoraggio di modelli nonlineari ed eteroschedastici per i dati sull'ozono e le polveri fini sia in Italia che in altri paesi con condizioni climatologiche diverse. Con particolare riferimento ai modelli SFI-SETARX-ARCH e localmente lineari eteroschedastici. Verifica di ipotesi su vincoli di disuguaglianza. Sviluppare la modellistica e la metodologia per la verifica di ipotesi di ordinamento stocastico

su catene di Markov e dei processi Hidden Markov applicati a problemi di monitoraggio ambientale e discutere l'applicazione;

3- modelli per valori estremi e conteggi. Andamento di dati di conteggio che presentano sovradisersione, le manifestazioni estreme dei livelli di inquinamento e dei fenomeni atmosferici; in particolare, per quanto riguarda l'inquinamento, un'importante finalità dello studio della distribuzione dei massimi nel tempo riguarda l'individuazione di trend;

4- campionamento e stima. La metodologia usata per lo studio dei metodi proposti si basa prevalentemente sul comportamento asintotico per grandi campioni, sulle simulazioni di tipo Monte Carlo e Bootstrap e sullo studio di casi concreti.

Per quanto riguarda il punto 1 è da tenere presente che le serie storiche ambientali ad alta frequenza sono spesso costituite da un grande numero di osservazioni e l'approccio asintotico è perciò appropriato.

5. CONCLUSIONI

L'esperienza effettuata è risultata molto interessante oltre che dal punto di vista formativo ed informativo, anche perché ha costituito l'occasione per dare avvio ad una attività di rassegna e catalogazione dei metodi di aggregazione di diversi indicatori ambientali.

Il lavoro effettuato può sicuramente essere migliorato ed arricchito, anche grazie a eventuali suggerimenti legati a specifiche esigenze dell'APAT; a puro titolo di esempio, si ritiene che sarebbe utile, nel futuro, provvedere ad apportare le seguenti modifiche ed integrazioni:

- Inserire tali schede, in sola lettura e consultazione, sul sito o la banca dati dell'APAT;
- Ricercare altri metodi di aggregazione diversi da quelli studiati in questo lavoro;
- Creare un gruppo di ricerca che permetti di approfondire l'argomento e che possa esso stesso sviluppare dei metodi di aggregazione statistica che siano differenti ed innovativi, per contribuire così a migliorare la precisione degli indicatori presenti in letteratura e poterli proporre, dopo averne accertato la bontà a tutti quegli organi nazionali ed internazionali che si occupano di ricerca, studio e tutela dell'ambiente.

BIBLIOGRAFIA

- Aa.Vv. (1997), *Land Quality Indicators and Their Use in Sustainable Agriculture and Rural Development*, FAO, UNDP, UNEP and World Bank, Roma (<http://www.fao.org/docrep/W4745E/w4745e00.htm#Contents>).
- AaVv. (2001), *Sustainability Indicators*, (Cap. 7) Penrith Rural Lands Study, Sidney, Australia.
- Adriaanse A. (1993), *Environmental Policy Performance Indicators: a study of indicators for environmental policy in the Netherlands*, SDU Uitgeverij Koninginnegracht, The Hague.
- Agostinelli C., Capiluppi C., Fabbris L. (1996), *Indici sintetici di qualità di una risorsa ambientale*, Annali di Statistica, 10, pp.239-256.
- Attanasio M., Capursi V. (1997), *Graduatorie sulla qualità della vita: prime analisi di sensibilità delle tecniche adottate*, Atti della XXXV Riunione Scientifica della SIEDS Politiche e tecniche di valutazione dell'attività della Pubblica Amministrazione e degli interventi sociali, Alghero.
- Atti del Convegno, *Workshop on indicators as a tool for managing and monitoring a sustainable local and regional planning process* (1999), Bruxelles, November 18. www.europa.eu.int.

- Becciu G., Bianchi A., Fassò A., Fassò C. A., Larcán E.(2000), *Quick calculation of minimum instream flow in drainage basins of Central Alps*, in Proceeding of New Trends in Water and Environmental Engineering for Safety and Life, Maione, Majone Lehto & Monti (eds), Balkema, Rotterdam.

- Bosch P. (2001), *The European Environment Agency focuses on EU-Policy in its Approach to sustainable development indicators*. Paper presentato da EEA al Joint ECE/Eurostat Work Session on Methodological Issues of Environment Statistics, Ottawa, Canada.

- Colombo A. (1996), *Indicatori per la valutazione di impatto ambientale*, Annali di Statistica, 10, pp.147-154.

- (1999), *Indicators for sustainable development: Theory, method, applications*, IISD- International Institute for Sustainable Development, Winnipeg, Manitoba, Canada.

- (2001), *Assessing Viability and Sustainability: a Systembased Approach for Deriving Comprehensive indicator Sets*, Conservation Ecology 5 (2): 12, (online) URL: <http://www.consecol.org/vol5/iss2/art12>.

- (1999a), *Documento di lavoro della Commissione. Relazione sugli indicatori ambientali e sugli indicatori di integrazione per il vertice di Helsinki*. Bruxelles. SEC (1999) 1942 def. (<http://europa.eu.int/comm/environment/newprg/sec99> 1942 it.pdf).

- (2000b), *Comunicazione della commissione – Gli indicatori strutturali*, Bruxelles, COM (2000) 594 def. (http://europa.eu.int/eur-lex/it/com/pdf/2000/com2000_0594it01.pdf).

- (2001c), *Decisione del Parlamento Europeo e del Consiglio relativa al programma statistico comunitario 2003-2007*. Bruxelles, COM(2001) 683 def.

- (2001f), *Gli indicatori strutturali Bruxelles*, COM(2001) 619 def.

- (2001h), *Statistical Information needed for Indicators to monitor the Integration of Environmental concerns into the Common Agricultural Policy*, Bruxelles, COM (2001) 144 final.

- Di Sano J.A. (2001), *Indicators of sustainable development: Guidelines and Methodologies*, UN Division for Sustainable Development, www.uncsd.org; <http://www.un.org/esa/sustdev/isd.htm>.

- D.M. Ambiente 15/4/1994, *Norme tecniche in materia di livelli e di stati di attenzione e di allarme per gli inquinanti atmosferici nelle aree urbane ai sensi degli articoli 3 e 4 del D.P.R. 24 Maggio 1988 n. 203, e dell'art. 9 del D.M. 20 Maggio 1991*, Gazzetta Ufficiale della Repubblica Italiana n. 107 del 10/05/1994.

- D.M. Ambiente 16/5/1996, *Attivazione di un sistema di sorveglianza di inquinamento da ozono*, Gazzetta Ufficiale della Repubblica Italiana n. 163 del 13/07/1996.

- Enea (1998) *Metodologie per lo Sviluppo Sostenibile. Indicatori e Schemi Pressione Stato Risposta*. Enea (www.enea.it).

- Eurostat (2001), *Key indicators for sustainable development*, Paper presentato da EEA al Joint ECE/Eurostat Work Session on Methodological Issues of Environment Statistics, Ottawa, Canada, 1-4 ottobre 2001.

- Fassò A. (2000), *Recursive least squares with ARCH errors*, GRASPA working paper n.6 (sirio.unipd.it/graspa/graspa.htm).

- Fassò A. (2001), *Time Varying Least Squares with ARCH Errors for Environmental Time Series Modelling*, 53rd ISI Conference, Seoul. Bulletin of the International Statistical Institute, Tome LIX, Volume 3, 139-140.

- Fassò, S. Locatelli (2002), *Semiparametric Modeling of High Frequency Ground Ozone*, in corso di stampa su Statistical Monitoring for Environmental Engineering: Models and Applications to Bergamo County.

- Fassò A., Negri I. (2002a), *Nonlinear statistical modelling of high frequency ground ozone data. Environmetrics*, Vol. 13, 3, 225-241.

- Fassò A., Negri I. (2002b), *Multi step forecasting for nonlinear models of high frequency ground ozone data: a Monte Carlo approach. Environmetrics*, Vol. 13, 4, 365-378.

- Fassò A., Perri P.F. (2001), *Sensitivity Analysis for Environmental Models*, in Abdel H. El-Shaarawi and Walter W. Piegorsch (eds) *Encyclopedia of Environmetrics*, Volume 4, pp 1968—1982, Wiley.

- Federico A. (2001), *Indicatori per uno sviluppo sostenibile in Italia*, Seminario ISSI sugli indicatori, ENEA, Roma (<http://prog2000.casaccia.enea.it>).

- Fraire M. (1989), *Problemi e metodologie statistiche di misurazione di fenomeni complessi tramite indicatori e indici sintetici*, *Statistica*, anno XLIX, n. 2, 1989, pp. 245-263.

- Global Urban Observatory (1997), *Monitoring Human Settlements with urban indicators* (draft), United Nation centre for Human Settlements (Habitat), Nairobi, Kenya.

- Greeuw S., Kok K., Rothman D. (2001), *Factors, actors, sectors and indicators. The Concepts and Application in MedAction*, ICIS, Maastricht University, Maastricht, Olanda (<http://www.icis.unimaas.nl>).

- Griguolo S., Palermo P.C. (1980), «Significati e metodologie dell'analisi territoriale multivariata», in *Archivio di studi urbani e regionali*, nn. 8-9, F. Angeli, Milano.

- Guinomet I. (1999), *Eurostat's tools for the development of indicators of sustainable development & The relationships between indicators of sustainable*

development. Paper presentato alla Euroconference "Quality of Life - Sustainability - Environmental Changes", Burgenland, Austria.

- Hamekoski K. (1998), *The use of a simple air quality index in the Helsinki Area, Finland*, Environmental Management, vol. 22 n.4, pp.517-520.

- Hartigan, J.A. (1975), *Clustering Algorithms*, John Wiley & Sons, New York.

- Hertin J., Berkout F., Moll S., Schepelmann P. (2001), *Indicators for monitoring integration of environment and sustainable development in enterprise policy*, SPRU - Science and Technology Policy Research, University of Sussex, Brighton, UK.

- Hoon P., Singh N., Wanmali S.S. (1997), *Sustainable Livelihoods: Concepts, Principles and Approaches to Indicator development*. (Draft Discussion Paper), Poverty and Sustainable Livelihoods, Social Development and Poverty Eradication Division, Bureau for Development Policy, United Nations Development Programme (UNDP), New York.

- Kassomenos P., Skouloudis A.N., Lykoudis S., Flocas H.A. (1999), *'Air quality indicators' for uniform indexing of atmospheric pollution over large metropolitan areas*, Atmospheric Environment, 33, pp.1861-1879.

- Mazziotta C. (1998), *La definizione degli indicatori*, in Istituto Guglielmo Tagliacarne e Unioncamere, Statistica e territorio, E Angeli, Milano.

- Newton P. et al. (1998), *Environmental indicators for national state of the environment reporting – human settlements*. Department of the Environment, Canberra Australia.

- Nicolis O., A. Fassò A.(2002), *La Qualità dell'Aria nella Provincia di Bergamo: un'Analisi Descrittiva dell'Ultimo Decennio*, in corso di stampa su Statistical Monitoring for Environmental Engineering: Models and Applications to Bergamo County. R. Colombi and A. Fassò Ed.'s, Bergamo University Press, Ed. Sestante, Bergamo.

- OECD (1991), *Environmental Indicators: a preliminary set*, Parigi (1993) *Core set of indicators for environmental performance reviews*, Environment Monographs n°83, Parigi (OECD/GD(93)179) (<http://www.oecd.org/env/indicators/publications .htm>).

- (2001a), *Key environmental indicators*, (<http://www.oecdwash.org/DATA/online.htm>).

- Ohman I., (1999), *Environmental Indicators - Towards a European System of Environmental Pressure Indicators and Indices*, Envstats Issue 6, United Nations Statistics Division.

- Ott. W.R. (1978), *Environmental indices: theory and practice*, Ann Arbor Science Publishers, Ann Arbor.

- Ott W.R., Hunt W.F. (1976), *A quantitative evaluation of the pollutant*

standards index, Journal of the Air Pollution Control Association, 26, pp. 1051-1054.

- Pender A., Dunne L., Convery EJ. (2000), *Environmental indicators for the urban environment: a literature review*. Department of Environmental Studies, University College Dublin, Richview, Dublin, Ireland (www.environmentaleconomics.net).

- Pilot Action Archi-Med (2001), *Risultanze del team degli esperti italiani sugli indicatori della base metodologica comune Italia-Grecia*, Sub action n. 2.2 "Enhancement of the Environment as a Factor of Development in the mediterranean Space", ERDF.

- Sahara and Sahel Observatory (1997, in co-operation with CILSS, IGAD, UMA, and Club du Sahel, and in concertation with Mali, Tunisia, Senegal, UNDP/UNSO, UNESCO), *Implementation and impact indicators*, Interim Secretariat of the Convention to Combat Desertification, Paris.

- Savoia M. (1993), *La documentazione statistica. Panoramica su alcune principali fonti internazionali*. In Cronache 1/93 IVAY, Venezia (www-ivar.it/doest).

- Shah R. (2000), *International Frameworks of Environmental Statistics and Indicators*, contributo presentato a Inception Workshop on the Institutional Strengthening and Collection of Environment Statistics, 25-28 April 2000, Samarkand, Uzbekistan.

- Smeets E., Weterings R. (1999), *Environmental indicators: typology and overview*, Technical report No 25, EEA, Copenhagen, Denmark.
- Spangenberg J.H., (2000), *Sustainable Development. Concepts and Indicators*. SERI, Sustainable Europe Research Institute, Cologne, FR Germany.
- Swamee P.K, Tyagi A. (1999), *Formation of an air pollution index*, Journal of the Air & Waste Management Association, 49, pp. 88-91.
- Thomas W.A. (Ed.) (1972), *Indicators of environmental quality*, Plenum Press, New York