

**F. ANGELINI, S. DI FRISCHIA  
F. COLAO, A. PALUCCI**

Divisione Tecnologie Fisiche per la Sicurezza e la Salute  
Laboratorio Diagnostiche e Metrologia  
Centro Ricerche Frascati, Roma

# IMPLEMENTAZIONE DI UN MODELLO INVERSO DI SOURCE APPORTIONMENT PER LA STIMA DI EMISSIONI INDUSTRIALI

RT/2017/21/ENEA



AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE,  
L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

F. ANGELINI, S. DI FRISCHIA  
F. COLAO, A. PALUCCI

Divisione Tecnologie Fisiche per la Sicurezza e la Salute  
Laboratorio Diagnostiche e Metrologia  
Centro Ricerche Frascati, Roma

# IMPLEMENTAZIONE DI UN MODELLO INVERSO DI SOURCE APPORTIONMENT PER LA STIMA DI EMISSIONI INDUSTRIALI

RT/2017/21/ENEA



AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE,  
L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

I rapporti tecnici sono scaricabili in formato pdf dal sito web ENEA alla pagina <http://www.enea.it/it/produzione-scientifica/rapporti-tecnici>

I contenuti tecnico-scientifici dei rapporti tecnici dell'ENEA rispecchiano l'opinione degli autori e non necessariamente quella dell'Agenzia

The technical and scientific contents of these reports express the opinion of the authors but not necessarily the opinion of ENEA.

# IMPLEMENTAZIONE DI UN MODELLO INVERSO DI SOURCE APPORTIONMENT PER LA STIMA DI EMISSIONI INDUSTRIALI

Federico Angelini, Stefano di Frischia, Francesco Colao, Antonio Palucci

## Riassunto

Questo documento riporta la descrizione e il commento del processo di implementazione di una catena modellistica di source apportionment per la quantificazione delle emissioni diffuse da sorgenti industriali. Lo schema seguito è quello delineato dalla norma EN UNI 15445 e documentato in letteratura, che prevede una simulazione atmosferica con un modello di dispersione, seguita da una regressione multilineare per risolvere il problema inverso.

Un modello di dispersione (AERMOD), è un modello stazionario a plume che calcola la dispersione atmosferica a partire dalla struttura della turbolenza nello strato limite planetario e concetti di scaling, e tratta sorgenti sia al suolo che elevate, così come orografia semplice o complessa). Il modello è quindi stato implementato ed è utilizzato per calcolare l'impatto di ogni sorgente sulla rete di recettori, previa inizializzazione con caratteristiche del suolo e dati meteo misurati. Viene qui allora descritta la procedura di acquisizione automatica dei dati e pre-processamento necessario all'utilizzo di AERMOD. Infine, un modello di regressione multilineare è utilizzato per stimare i valori delle emissioni diffuse dalle due acciaierie, sia su base oraria che su base giornaliera. E' stimata anche la sensibilità del modello a vari parametri in input, e la possibile interferenza con altre sorgenti prossime.

Vengono infine discussi i punti critici dell'intero procedimento, sia sulla parte diretta che inversa, suggerendo possibili strategie per il loro superamento.

**Parole chiave:** modellistica di dispersione atmosferica, emissioni diffuse EN UNI 15445, AERMOD, regressione lineare multipla.

## Abstract

*This document describes the implementation process of a source apportionment modeling chain for quantifying emissions coming from industrial sources. The scheme is outlined in EN UNI 15445 and documented in literature, which provides an atmospheric simulation with a dispersion model, followed by a multilinear regression to solve the reverse problem.*

*A dispersion model (AERMOD) is a stationary plume model that calculates atmospheric dispersion from the structure of turbulence in the planetary boundary layer and scaling concepts, and involves both ground and elevated springs, as well as simple or complex orography). The model has been implemented and used to calculate the impact of each source on the receptor network, after initialization with soil characteristics and measured weather data. Here comes the procedure for automatic data acquisition and pre-processing required for the use of AERMOD.*

*Finally, a multilinear regression model is used to estimate the emission values of diffused by the two steel mills, both on hourly and daily basis. The model's sensitivity to various input parameters, as well as possible interference with other near sources, has been also estimated.*

*Finally, the critical points of the entire process are discussed, in both direct and the inverse side, suggesting possible strategies for overcoming them are presented.*

**Keywords:** multiple linear regression, atmospheric dispersion model, diffuse emissions, source apportionment, EN UNI 15445, AERMOD.



# INDICE

1	Introduzione	7
2	Il problema del source apportionment	7
3	Modello dispersione (ADM)	9
3.1	Tassonomia dei modelli di dispersione atmosferica	9
3.2	Modelli stazionari e scelta dell'ADM	10
3.3	AERMOD	11
3.3.1	AERMET: preprocessore meteorologico di AERMOD	13
3.3.2	AERMOD: input e output	16
3.3.3	Studio di sensibilità per i parametri CC, $\alpha$ , $z_0$ e B.	18
3.3.4	Conclusioni sull'utilizzo di AERMOD	19
4	Modello di regressione multilineare (MLR)	21
4.1	Procedura Multilinear Regression Model	22
4.1.1	Correlazione tra sorgenti	23
4.1.2	Coefficienti negativi e regressione vincolata	24
4.1.3	Normalità dei residui	24
4.1.4	Varianza spiegata	25
4.1.5	Test F	26
4.2	Restituzione dei valori	26
4.2.1	Dati in uscita	26
4.2.2	Raggruppamento delle sorgenti	28
5	Conclusioni generali	29
6	Bibliografia essenziale	31
6.1	Documenti su AERMOD:	32



## 1 Introduzione

Il problema della determinazione delle emissioni diffuse da impianti industriali è molto vasto e riveste notevole importanza nell'ambito della tutela della salute, come testimoniato dalla lunga storia di tentativi di soluzione, e dalla varietà di approcci adottati: per una panoramica si possono consultare i lavori di Viana et al. 2008, Fragkou et al., 2012. La norma EN UNI 15445 [Emissioni da fughe e diffuse relative ai settori industriali - Determinazione di sorgenti di polveri da fughe mediante metodo RDM (modello di dispersione inversa)] fa da riferimento per il source apportionment da emissioni fuggitive in ambito industriale, accoppiando un modello di dispersione ad una regressione lineare multipla. Pertanto, tale norma è stata adottata come modello per il presente lavoro, che implementa quindi le direttive ivi descritte; tale metodologia è riportata in letteratura già dai primi anni 2000 (Fushimi et al., 2005). Questa ha il vantaggio di non necessitare di conoscenze a priori sulla composizione o sull'andamento temporale delle sorgenti; prevede invece l'utilizzo di un modello di dispersione atmosferica per simulare l'accoppiamento tra una serie di sorgenti e un insieme di recettori (problema diretto). Dopo di ciò, occorre risolvere un problema inverso che consiste nello stimare le emissioni delle sorgenti che meglio approssimano il campo osservato. In generale questo tipo di problemi è molto complesso dal punto di vista matematico, in quanto non è in genere garantita l'esistenza né l'unicità della soluzione, che inoltre mostra una notevole sensibilità dalle condizioni al contorno. Essendo in questo caso particolare il problema lineare, è possibile risolverlo con una regressione lineare multipla, come descritto con qualche dettaglio nei prossimi paragrafi. Ciò semplifica molto le cose da un punto di vista analitico.

E' da notare, infine, che un approccio di questo genere è applicabile a qualsiasi inquinante, purché il modello di dispersione sia costruito per calcolarne il trasporto (es. tenga conto di deposizione, reazioni chimiche, decadimento, ecc.).

## 2 Il problema del source apportionment

Nell'approccio proposto da Fushimi, il problema si può enunciare in maniera molto semplice, notando che per ogni recettore  $j$  la concentrazione misurata  $C_j$  dell'inquinante considerato può essere scomposta nella somma delle concentrazioni provenienti da ognuna delle  $m$  possibili sorgenti  $E_i$ , pesate per il corrispettivo coefficiente di accoppiamento  $\beta_{ij}$  tra la sorgente  $i$  ed il recettore  $j$ :

$$C_j = \sum_{i=1}^m \beta_{ij} E_i + E_0$$

Ove  $E_0$  rappresenta un termine di fondo, che deve ovviamente essere lo stesso per ogni recettore. Generalizzando a  $n$  ricettori forma matriciale il problema si può riscrivere:

$$\begin{bmatrix} C_1 \\ \dots \\ C_n \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \dots & \beta_{1m} \\ \vdots & \ddots & \vdots \\ \mathbf{1} & \dots & \beta_{nm} \end{bmatrix} \begin{bmatrix} E_0 \\ \dots \\ E_m \end{bmatrix}$$

ove  $C$  rappresenta il vettore delle concentrazioni  $C_j$  ( $j=1,n$ ) misurate alle  $n$  centraline di monitoraggio,  $E_0$  rappresenta il termine di fondo,  $\beta$  la matrice degli accoppiamenti ed  $E_i$  ( $i=0,m$ ) le sorgenti. Nella matrice  $\beta$ , il vettore  $E_0$  è stato accorpato immettendo una colonna costante e unitaria nella matrice degli accoppiamenti. Il vettore delle emissioni contiene così un termine aggiuntivo corrispondente a tale colonna, che rappresenta quindi il termine di fondo. Qualora fossero noti i coefficienti della matrice  $\beta$ , il problema da risolvere è un cosiddetto problema inverso, che può essere affrontato in vari modi.

Poiché il problema in esame è intrinsecamente lineare nei coefficienti  $E_0, \dots, E_m$ , il metodo più semplice è utilizzare una regressione multilineare (MLR), basata sull'approccio statistico della massima verosimiglianza, e simile concettualmente al fit di una regressione lineare nel piano, ma con più variabili concorrenti.

A seconda del numero di sorgenti e di recettori, il sistema può essere sovra- o sotto- determinato. Oltretutto, il sistema può passare da sovradeterminato a sottodeterminato in base alle caratteristiche della matrice  $\beta$ . Tali casi saranno analizzati nel seguito (rango, correlazione tra colonne). Se il sistema è ridondante (cioè sovra determinato,  $n>m$ ), si può cercare un algoritmo che calcoli la combinazione dei coefficienti del vettore delle emissioni tale da minimizzare gli scarti della soluzione del sistema. Per fare ciò, ovviamente, esso deve essere sovradeterminato (come nel calcolo di un fit a due dimensioni, del resto), cioè presentare ridondanza. Questo porta in maniera naturale al calcolo di un termine di residui  $\varepsilon$ , che tiene conto delle differenze tra quanto previsto dalla regressione e quanto osservato nel vettore  $C$ . Il problema allora può essere scritto nella forma:

$$\begin{bmatrix} C_1 \\ \dots \\ C_n \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \dots & \beta_{1m} \\ \vdots & \ddots & \vdots \\ \mathbf{1} & \dots & \beta_{nm} \end{bmatrix} \begin{bmatrix} E_0 \\ \dots \\ E_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \dots \\ \varepsilon_n \end{bmatrix}. \quad (1)$$

Oltre a fornire i valori di  $\varepsilon$  ed  $E$ , l'approccio MLR permette di calcolare vari indici statistici per valutare la significatività del risultato: in particolare si utilizzano il coefficiente di correlazione multiplo  $R^2$  (totale e parziale) ed il test F, che saranno spiegati in seguito. L'approccio MLR è piuttosto potente, a patto che il problema sia sufficientemente condizionato, sebbene possieda varie criticità che saranno discusse in seguito. Nella Figura 1 è mostrato lo schema logico di funzionamento della catena di processamento del modello di source apportionment.

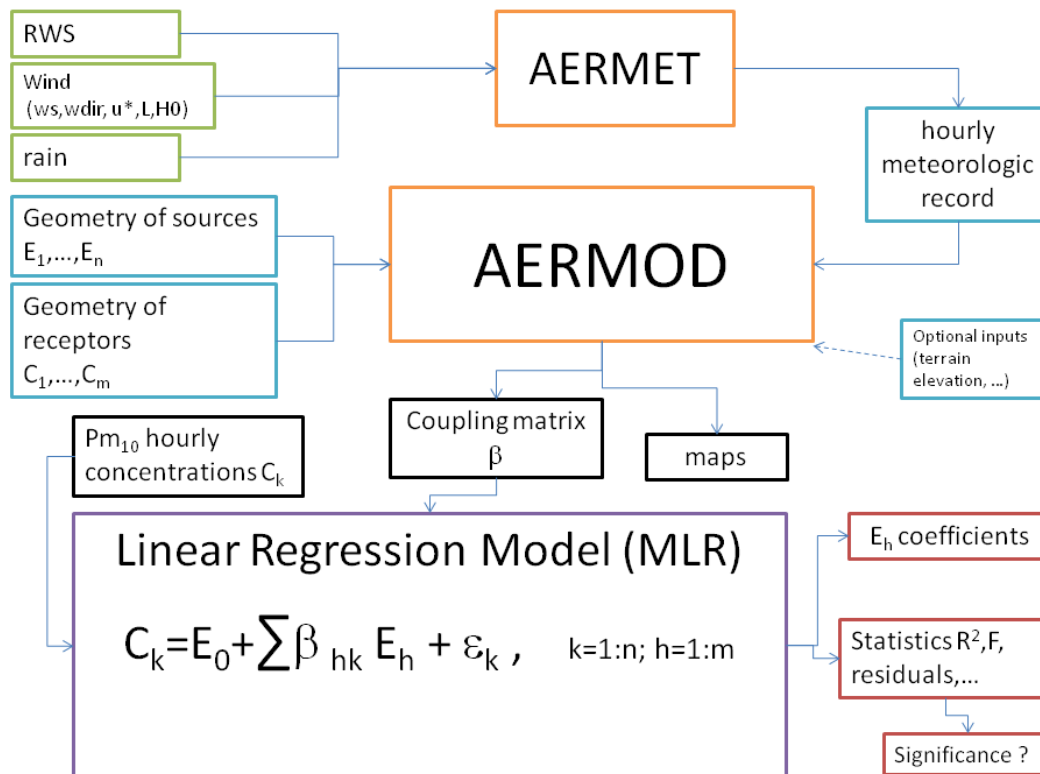


Figura 1 - Schema logico di funzionamento della catena di processamento del source apportionment implementato.

### 3 Modello dispersione (ADM)

Resta ora il problema di come determinare i coefficienti della matrice  $\beta$ . Secondo la citata norma EN UNI, e secondo la prassi corrente nella risoluzione di questo tipo di problemi (Fushimi et al., 2005; Laupsa et al., 2009; Denby et al., 2010), occorre stimarli mediante un modello di dispersione. Poche indicazioni sono fornite sulla scelta del modello, tranne che si richiede una risoluzione oraria per la determinazione dei parametri. Sta di fatto che tutta la letteratura sull'argomento converge sull'utilizzo dei cosiddetti modelli *gaussiani a plume*, e ciò per una serie di ragioni che saranno ora discusse. L'offerta infatti è molto vasta e le tipologie di modelli innumerevole, per non parlare delle singole implementazioni e varianti. Qui di seguito è riportata una breve descrizione delle possibilità allo stato dell'arte attuale.

#### 3.1 Tassonomia dei modelli di dispersione atmosferica

I modelli statistici sono basati sulla sola analisi statistica delle emissioni, senza alcun apporto di informazioni meteo-dispersive. Rappresentano un approccio potente se si hanno a disposizione molti dati (serie temporali molto lunghe, molti recettori e tendenzialmente una serie di fingerprint delle emissioni di ogni sorgente. I metodi utilizzati all'interno di questi modelli variano dalla Principal Component Analysis (PCA) alla Positive Matrix Factorization (PMF), a modelli ibridi e al Chemical Mass Balance (CMB). Non sono,

viceversa, validi per trattare con una sola specie o qualora non si abbiano a disposizione molte informazioni sulla composizione dei contributi delle varie sorgenti concorrenti.

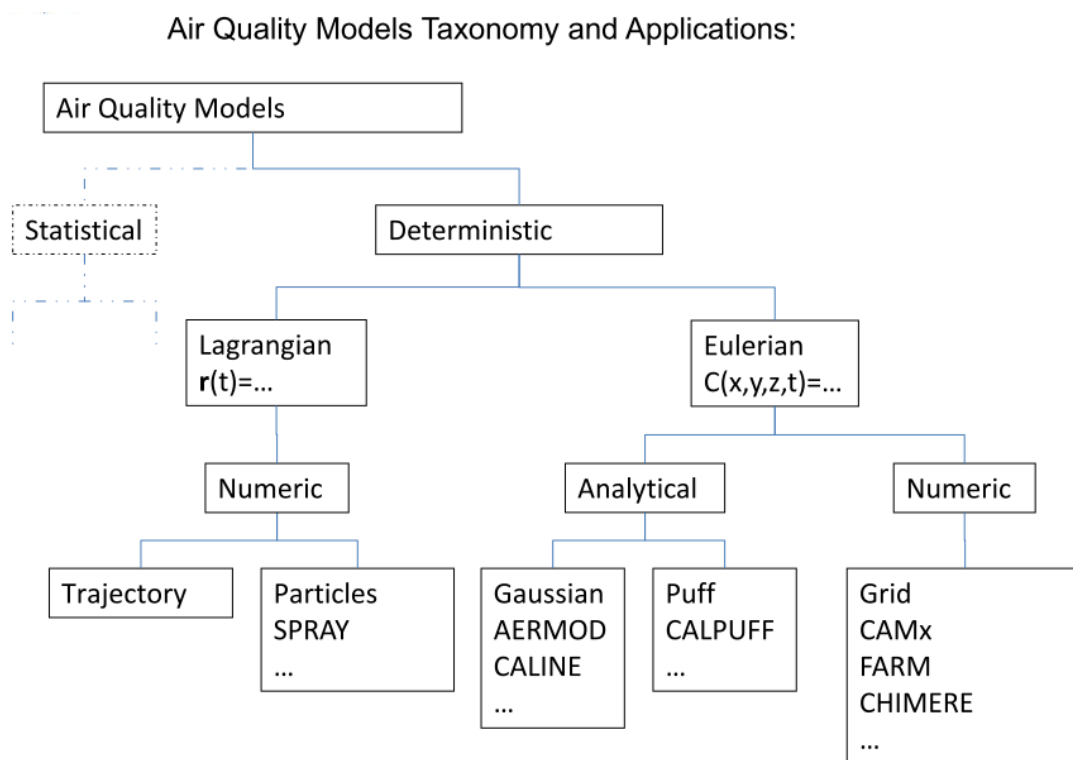


Figura 2 – Classificazione dei modelli di dispersione atmosferica.

I modelli denominati ‘deterministici’, invece, simulano i processi fisici e chimici per calcolare il trasporto, la dispersione, la deposizione e la trasformazione dei vari inquinanti emessi, con uno spettro molto variabile di gradi di complessità e precisione. Una tipica classificazione è basata sull’impiego di coordinate euleriane o lagrangiane; poiché un approfondimento di queste caratteristiche è oltre gli scopi del presente lavoro, si rimanda a testi specializzati (es.: Sozzi, 2003, Stull, 2012) per eventuali approfondimenti.

### 3.2 Modelli stazionari e scelta dell’ADM

A prescindere dal punto di vista adottato nella formulazione, una importante famiglia di modelli di dispersione comprende i cosiddetti modelli stazionari: questi sono sviluppati per processi poco variabili, sulle scale dei tempi e degli spazi, all’interno del dominio di calcolo. Viceversa, i modelli non stazionari tengono conto della variazione delle emissioni e del campo meteorologico ( e quindi della termodinamica e turbolenza atmosferiche) nel periodo di simulazione. Nella pratica corrente i modelli di tipo stazionario sono frequentemente impiegati per la stima del trasporto e della dispersione degli inquinanti in atmosfera, limitatamente alla determinazione dell’impatto che varie sorgenti inquinanti presenti hanno nei confronti di un territorio di estensione limitata e di caratteristiche estremamente regolari. Le ipotesi su cui si basano tutti questi tipi di modelli sono:

- le situazioni meteorologiche devono essere lentamente variabili nel tempo in modo tale che si possa immaginare di trattare una sequenza di situazioni quasi-stazionarie;
- il campo di vento e delle altre variabili meteorologiche che determinano il trasporto e la dispersione degli inquinanti in aria è orizzontalmente omogeneo;
- per tutte le variabili meteorologiche si ipotizza un profilo verticale noto;
- la turbolenza atmosferica deve essere descritta nella maniera più semplice e più compatta possibile.

All'interno dei modelli stazionari, un'importante categoria è quella dei modelli gaussiani a *plume*, i primi ad essere sviluppati storicamente per simulare la dispersione di pennacchi provenienti da sorgenti puntuali. I modelli a *plume* sono tutt'ora i modelli più semplici ed utilizzati nello studio della dispersione atmosferica su brevi distanze, nonostante le limitazioni operative dettate dalla elevata schematizzazione dei processi. Nel tempo, tuttavia, molti modelli gaussiani a *plume* sono stati migliorati per tenere conto degli effetti del galleggiamento, del *downwash* a valle di edifici, la possibilità di introdurre sorgenti non puntuali e la descrizione non gaussiana delle *plume* distorte da fenomeni di *updraft* e *downdraft* (modelli ibridi). La limitata richiesta di parametri in input e la velocità di calcolo di questo tipo di modelli ne rende ancora molto competitivo l'utilizzo in molte occasioni. Infatti, nella scelta del modello operativo per il caso in esame occorre tenere conto della necessità di contenere i tempi di calcolo entro limiti ragionevoli, con un dataset in entrata piuttosto semplice (sono disponibili solo dati di vento a livello del suolo) e risoluzione dell'ordine delle centinaia di metri, distanza tipica tra le centraline operanti. E' pertanto necessario un modello operante tra la microscala (100 m, 1 km) e la scala locale (1 km, 100 km). Da un'attenta analisi di tutti questi fattori e della letteratura recente è stato scelto l'impiego di un modello gaussiano a *plume*, precisamente il modello AERMOD sviluppato e rilasciato dall'Environmental Protection Agency degli Stati Uniti (EPA), e da questa elencato tra i modelli consigliati, assieme a CALPUFF, per lo studio della dispersione degli inquinanti.

### 3.3 AERMOD

AERMOD è un modello stazionario, nel senso che si assume che il processo di diffusione della *plume* sia costante all'interno del tempo di propagazione e di rivelazione. Con emissioni lentamente variabili nel tempo rispetto alla risoluzione temporale delle misure, e su scale spaziali piccole (fino alle decine di km), si può ragionevolmente assumere la stazionarietà del processo. Il modello assume, come tutti i modelli gaussiani, un allargamento gaussiano della *plume* ed un trasporto dati dalle caratteristiche del vento medio e della stabilità atmosferica. Mentre la diffusione è guidata dalle caratteristiche turbolente dell'atmosfera, il trasporto e il galleggiamento della *plume* sono dettate dal vento medio e dalla struttura termodinamica dello strato limite planetario. Per approfondire il discorso si può consultare il testo di micrometeorologia e dispersione degli inquinanti disponibile gratuitamente al sito: <http://www.arpat.toscana.it/temi-ambientali/aria/modellistica-per-la-qualita-dellaria/linee-guida/apat-micrometeorologia.pdf>.

E' da notare che questo tipo di modelli non è capace di calcolare la diffusione in caso di calma di vento. Questo è un problema importante quando si ha a che fare con dati di vento provenienti da anemometri

tradizionali (a banderuola e coppette), che presentano una inerzia e un attrito considerevoli, tanto che velocità del vento minori di qualche decimetro al secondo vengono registrate come nulle. Gli anemometri sonici sono, invece, capaci di misurare velocità fino a frazioni del cm/s, e pertanto, per l'implementazione di un modello di dispersione, è altamente consigliabile l'impiego di tali strumenti..

Altro punto da sottolineare è che si tratta di un modello analitico e non numerico, per cui non opera su una griglia da definire, ma utilizza parametrizzazioni opportune per determinare i coefficienti dell'allargamento e del trasporto del pennacchio. Per questo motivo richiede pochi calcoli ed è pertanto molto veloce. L'output consiste nel calcolo delle concentrazioni esclusivamente nei punti prescelti, e non richiede il calcolo continuo su tutti i punti griglia. Essendo inoltre stazionario, non è considerata alcuna evoluzione temporale, e ciò semplifica ulteriormente i calcoli. Ovviamente questo ha per contro ovvie limitazioni sulle scale temporali e spaziali maggiori, e per situazioni orografiche complesse (interfacce terra-mare, terreni montagnosi o urbani), ma il prezzo da pagare per l'utilizzo di un modello più sofisticato, quale ad esempio Calpuff, è una complessità molto maggiore dei dati richiesti in entrata (campo meteorologico tridimensionale, mappe di uso del suolo, DEM,...), e un'accurata sintonizzazione di tutti i parametri con conseguente validazione, prima di poter ottenere risultati più attendibili rispetto ad un più semplice modello gaussiano stazionario. In ogni caso, la citata norma EN 15445 fa riferimento al modello Fugitive Dust Model (FDM), sempre sviluppato da EPA, ma risalente agli anni '80 e ormai obsoleto, tanto che non rientra più tra i 'recommended models'. Pertanto, si è scelto di utilizzare la sua evoluzione allo stato dell'arte, cioè appunto AERMOD.

Senza entrare troppo nel dettaglio della fisica implementata nel modello AERMOD (per il quale si rimanda al manuale disponibile al sito:

[https://www3.epa.gov/ttn/scram/7thconf/aermod/aermod\\_mfd.pdf](https://www3.epa.gov/ttn/scram/7thconf/aermod/aermod_mfd.pdf)), si riepiloga qui lo schema operativo del modello.

Previo inserimento dei dati di geometria e portata della sorgente, AERMOD calcola l'innalzamento del pennacchio e la spinta di galleggiamento. Esso è in grado di simulare i cosiddetti "plume lofting" cioè le porzioni di massa degli inquinanti che in situazioni convettive prima di diffondersi nello strato limite, tendono ad innalzarsi e a rimanere in prossimità della fascia più elevata dello strato limite. Inoltre, riesce a tenere conto della penetrazione del pennacchio in presenza di inversioni in quota; può trattare lo strato limite in situazioni urbane, ed infine predispone un'analisi dei parametri meteorologici allo scopo di definire la struttura verticale dello strato limite. Maggiori dettagli sul funzionamento di AERMOD sono reperibili nei due lavori di Cimorelli et al. (2004) e Perry et al. (2004).

Per quanto riguarda l'output del modello, si possono specificare recettori in tutti i tipi di terreno, a quote e posizioni arbitrarie. Come sorgenti può prendere in esame i rilasci da sorgenti singole o multiple, sia puntuali che areali o volumetriche, a quote arbitrariamente specificate. AERMOD fornisce allora le concentrazioni medie, sulla base del tempo di media specificato in input, ai vari recettori, in  $\mu\text{g}/\text{m}^3$ .

Con i preprocessori AERMET e AERMAP si possono prendere in considerazione diversi tipi di dati meteorologici in input, e si può introdurre una topografia complessa per simulare meglio l'effetto di turbolenza indotta da edifici o la compressione delle linee di flusso determinate dall'orografia. Moltissima

letteratura è disponibile in rete riguardo alla validazione di AERMOD e al confronto con altri modelli di dispersione. Consideriamo quindi il modello allo stato dell'arte per questo tipo di problematiche. AERMAP richiede in ingresso un modello DEM della topografia ad alta risoluzione (~30 m), per cui non è stato implementato vista la geometria piuttosto semplice del terreno nei dintorni delle centraline considerate.

### 3.3.1 AERMET: preprocessore meteorologico di AERMOD

AERMET è un preprocessore meteorologico che accetta una varietà di dati meteorologici in input, e li elabora fornendo un record che può essere accettato da AERMOD per i calcoli di dispersione veri e propri. In pratica, per soddisfare le necessità meteorologiche e micrometeorologiche di un modello stazionario è sufficiente considerare un solo punto del dominio di calcolo da essi considerato, che per quel che si è detto, risulta orizzontalmente omogeneo: le informazioni micrometeorologiche e meteorologiche relative a tale punto saranno dunque rappresentative dell'intero dominio di calcolo. Pertanto si può dire che tale classe di modelli richiedono un modello monodimensionale di PBL (Planet Boundary Layer) che inoltre è stazionario, dato che questi modelli di dispersione ipotizzano che in ogni intervallo temporale considerato il PBL resti stazionario e quindi simulano una successione di stati quasi stazionari. Per quanto riguarda i modelli ibridi di nuova generazione, come per esempio AERMOD, la struttura della turbolenza del PBL risulta completamente descritta una volta che risultino noti nel punto di riferimento del dominio di calcolo:

- la lunghezza di rugosità  $z_0$ , caratteristica dell'intero dominio di calcolo stesso;
- la turbolenza di tipo meccanico, rappresentata dalla velocità di frizione  $u^*$  ;
- la turbolenza di tipo convettivo, caratterizzata dal flusso turbolento di calore sensibile  $H_0$  ;
- l'estensione verticale del PBL,  $Z_i$  (normalmente indicata col termine di altezza di rimescolamento), sia nelle ore convettive che nelle ore stabili. Nelle ore convettive  $Z_i$  è determinata essenzialmente da  $H_0$ , mentre nelle ore stabili da  $u^*$ .

Il processing dei dati meteo da parte di AERMET si articola in tre fasi principali:

- il primo passo è l'estrazione/recupero dei dati meteo *in situ*, e la valutazione della qualità dei dati stessi
- il secondo passo consiste nell'unire (*merging*) i file provenienti dal passo 1 in un unico file ASCII. Questo file è organizzato in modo che ogni blocco di dati contiene le osservazioni di un periodo di 24 ore
- il terzo e ultimo passo consiste nel leggere il file del passo precedente e produrre due file input per AERMOD. Il primo conterrà parametri legati allo strato limite e ai venti d'altitudine (**Surface File**), il secondo conterrà uno o più profili di vento, temperatura e altri parametri (**Profile File**).

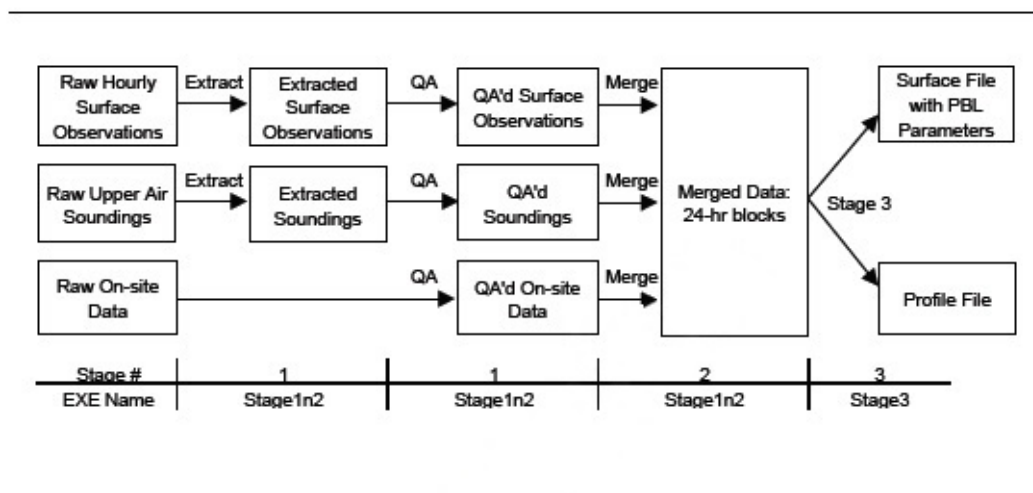


Figura 3: Gli step di procesamiento di AERMET

I dati raw iniziali che AERMET accetta come input sono di tre tipi:

- «**Raw hourly surface observations**» - dati orari di superficie (direzione e velocità vento, temperatura, ...) solitamente registrati da stazioni di monitoraggio pubbliche (negli Stati Uniti dal National Weather Service)
- «**Raw upper air soundings**» - dati atmosferici di altitudine raccolti da radiosonde lanciate da stazioni fisse (sul web sono resi disponibili dall'Università del Wyoming – <http://weather.uwyo.edu/upperair/sounding.html>)
- «**Raw on-site data**» - dati specifici raccolti nel sito d'interesse che dovranno essere formattati e processati in modo da poter essere elaborati da AERMET

### 3.3.1.1 Dati al suolo

I dati orari di superficie contengono informazioni circa i venti, la temperatura e la copertura delle nuvole che saranno usati per i calcoli del modello di dispersione. Sono tipicamente raccolti da stazioni del NWS e vengono resi disponibili in alcuni formati standard.

I formati che AERMET accetta per questi dati sono molteplici (**CD-144**, un vecchio formato rimpiazzato dal TD-3280, **SAMSON** e **SCRAM** (MET144), una versione ridotta del vecchio formato CD-144).

Poiché il formato dei dati meteorologici in situ dipende dalla strumentazione disponibile e dal sistema informatico di acquisizione e archiviazione, è da valutare caso per caso quali e quante variabili passare al modello. Per tale motivo, è stato scelto di introdurre una procedura che traduca i dati esistenti nel formato SCRAM, il più recente ed il più essenziale fra quelli proposti, in modo da mantenere inalterato il sistema di input e trasferire la flessibilità della catena di procesamiento nella modifica del solo script di traduzione.

Un file meteo nel formato SCRAM è un file di testo composto da un record per ora. Ogni riga contenente un record orario ha una lunghezza fissa di 28 caratteri, la cui posizione definisce il tipo di parametro analizzato.

Un esempio di file SCRAM con 5 record orari è il seguente:

	<u>Columns</u>	<u>Element</u>
	1- 5	Surface Station Number
	6- 7	Year
	8- 9	Month
	10-11	Day
	12-13	Hour
	14-16	Ceiling Height (Hundreds of Feet)
	17-18	Wind Direction (Tens of Degrees)
	19-21	Wind Speed (Knots)
	22-24	Dry Bulb Temperature (Degrees Fahrenheit)
	25-26	Total Cloud Cover
	27-28	Opaque Cloud Cover

```
0000115110101---080010610000
0000115110102---050010600000
0000115110103---050000590000
0000115110104---020010580000
0000115110105---020020580000
```

**Wind Direction** - The direction from which the wind is blowing, based on the 36 point compass, e.g. 09=East, 18=South, 27=West, 36=North, 00=Calm.

**Cloud Cover** - There are two cloud cover parameters, *opaque cloud cover* and *total cloud cover* in the SCRAM meteorological data files. Both parameters identify the amount of cloud cover measured in tens of percent, 0 = clear or less than 10%, 4 = 40-49%, '-' = overcast or 100%. Note: PCRAMMET reads the field for opaque cloud cover.

$$\text{*knots} = \text{m/s} * 1.94$$

$$\text{*Fahrenheit} = \text{°C} * 1.8 + 32$$

Figura 4: Definizione formato SCRAM

### 3.3.1.2 Dati in quota

I dati atmosferici in quota forniscono informazioni sulla struttura verticale dell'atmosfera, e includono, tra gli altri, valori di altezza, pressione, temperatura, umidità e vento. Sono raccolti per mezzo di radio-sonde lanciate in atmosfera da stazioni di monitoraggio prestabilite, e rilasciati in formati standard sul web. I radio-sondaggi vengono eseguiti tipicamente due volte al giorno: alle 0:00 e alle 12:00

Il formato accettato da AERMET è il **TD-6201**. Nel caso più generale, i dati di radiosondaggio non sono di pubblico dominio, ma vanno acquistati dai rispettivi gestori. Tuttavia, nel caso non siano disponibili dati in alta risoluzione specifici, il software è in grado di scaricare automaticamente dal sito web dell'Università del Wyoming i dati del radiosondaggio di una stazione vicina, scelta dall'utente:

(<http://weather.uwyo.edu/upperair/sounding.html>).

Un file meteo nel formato TD-6201 è un file di testo composto da un record per ogni radio-sondaggio (ad es. 2 record al giorno). Ogni riga contenente un record è composta da un blocco fisso iniziale di 32 caratteri (**ID Portion**), e da un numero variabile di blocchi da 36 caratteri (**Data Portion**) in base a quante misure in altezza sono state effettuate dalla sonda.

Un esempio di record TD-6201 con 1 ID-Portion iniziale (blocco blu) e 2 Data-Portion (blocchi rosso e verde) è il seguente:

```
000163204039N01757E20151101001270999910230+00010+13408416500200000090999910200+00034+1
680681610030000009
```

Per una spiegazione completa del significato dei parametri contenuti nel record, rimandiamo al link <http://www.webmet.com/MetGuide/TD6200.html>

In caso di errori nell'acquisizione dei dati di radiosondaggio, è possibile utilizzare un profilo medio climatologico per la stazione prescelta. Un messaggio di warning viene in questo caso comunque riportato per avvertire dell'utilizzo dei dati climatologici.

#### 3.3.1.3 *Dati al suolo ON-SITE*

I dati on-site sono misure specifiche raccolte da un sito di osservazione meteorologica che l'utente vuole includere nel modello di AERMET. Generalmente, questi dati possiedono una loro struttura specifica che non coincide con quella richiesta da AERMET. L'utente deve perciò indicare al processore come sono strutturati questi dati e come gestire una serie di opzioni (ad esempio, i valori mancanti).

Tuttavia AERMET permette all'utente di specificare un insieme di parametri che è in grado di processare (per un elenco completo di questi parametri, si rimanda all'Appendice B della guida utente di AERMET: <https://www3.epa.gov/scram001/7thconf/aermod/aermetugb.pdf>).

esempio Nel caso in cui questi siano disponibili, vengono utilizzati parametri aggiuntivi quali la Componente Verticale del Vento (VVW), la Surface Friction Velocity ( $u^*$ ), il Surface Heat Flux (H0) e la distanza di Monin-Obukhov (L).

#### 3.3.1.4 *Risultati dell'elaborazione*

Dopo le diverse fasi dell'elaborazione di AERMET, illustrate sopra, il risultato del modello sarà costituito da due file dati ASCII e dai file di reportistica generati analogamente anche negli stage intermedi.

I due file dati sono:

- **AERMET.PFL** – un file di testo contenente uno o più livelli di vento, temperatura o deviazione standard del vento calcolati da un'osservazione specifica on-site o dai dati NWS
- **AERMET.SFC** – un file di testo contenente l'insieme dei parametri micro-meteo di superficie calcolati dal modello

### 3.3.2 **AERMOD: input e output**

Una volta terminato il *processing* dei dati meteo da parte di AERMET, AERMOD può essere avviato, avendo come input i due file meteo e un ulteriore file di input (che chiameremo AERMET.INP) dove vengono elencati i parametri per il calcolo del modello di dispersione.

In particolare, il file di input AERMOD sarà composto dalle seguenti sezioni:

- Una sezione **CONTROL**, per specificare le principali opzioni di avvio e controllo
  - ad esempio il tempo di media, il tipo di inquinante e la configurazione del terreno
- Una sezione **SOURCE**, per definire le sorgenti di emissione
  - ad esempio la posizione delle sorgenti, il loro rateo di emissione, o la loro tipologia (puntuali, areali o volumetriche)
- Una sezione **RECEPTOR**, per definire i recettori
  - ad esempio la loro posizione e la loro tipologia (griglia di recettori, recettori in posizioni discrete, ecc...)
- Una sezione **METEO**, per specificare le opzioni degli input meteo
  - ad esempio il nome dei file meteo di input e delle stazioni meteo utilizzate in AERMET

- Una sezione **EVENT**, per includere eventuali eventi particolari nel modello
  - ad esempio per specificare eventi che vanno a modificare la normale concentrazione modellata da AERMOD
- Una sezione **OUTPUT**, per definire le opzioni di output
  - ad esempio il nome e il formato dei file dati generati da AERMOD

Dal punto di vista del sistema informativo, l'elaborazione iniziata con l'esecuzione del pre-processore AERMET continuerà senza interruzioni con l'esecuzione di AERMOD (ovvero dell'eseguibile AERMOD.EXE), mostrando all'utente i risultati finali del modello di dispersione, come se si trattasse di un unico processo.

Un'analisi più approfondita riguarda una serie di parametri meteo aggiuntivi che AERMOD necessita per la sua esecuzione. Questi parametri di input, sono forniti da AERMET, ma devono essere definiti manualmente all'inizio del pre-processamento meteo. In particolare, si tratta di:

- Copertura nuvolosa (CC, totale e opaca). La copertura nuvolosa serve per stimare la radiazione incidente sulla superficie, che serve a sua volta per stimare il flusso di calore latente. Questo è utilizzato per calcolare la condizione di stabilità atmosferica, e tramite un modello unidimensionale termodinamico l'altezza dello strato limite planetario convettivo.
- Albedo ( $\alpha$ ). L'albedo rappresenta la frazione di energia elettromagnetica riflessa dal suolo terrestre, e che quindi non risulta disponibile per essere accumulata nel suolo ed innescare la convezione. Tanto maggiore è l'albedo, tanto minore risulterà il calore accumulato dal suolo. Il tool AERSURFACE, un programma nato per fornire questi dati ad AERMOD sulla base di catalogazione del land use, fornisce per aree industriali e urbane un valore medio costante annuale pari a 0.18, che è stato assunto come valido.
- Rapporto di Bowen (B). Una volta accumulato il calore nel suolo, esso è rilasciato in atmosfera all'aria a contatto con esso. A seconda della disponibilità di acqua, però, il calore rilasciato può trasferirsi sotto forma sensibile (aria a temperatura superiore), ovvero sotto forma di calore latente, cioè di vapore che ha assorbito calore ed è pronto a rilasciarlo in caso di condensazione. Il rapporto tra calore sensibile e calore latente, detto rapporto di Bowen, è utilizzato come parametrizzazione in molti modelli. Esso è variabile tra 10 nel caso di terreno estremamente secco e 0.1 in caso di oceano. Tipicamente in situazioni di *land use* industriale è considerato attorno a 1.5, variabile tra 0.5 e 3 a seconda della condizione del suolo (bagnato-secco), e non ci si aspettano grandi variazioni stagionali, per cui il valore è stato assunto costante e pari a 1.5.
- Rugosità superficiale ( $Z_0$ ). Essa è definita nell'ambito della teoria della similarità di Monin-Obukov, e rappresenta la quota a cui il profilo logaritmico della velocità del vento si annulla. Anche questo valore è piuttosto stabile in un ambiente industriale, e sulla base delle indicazioni di AERSURFACE è assunto pari a 0.7 m.

### 3.3.3 Studio di sensibilità per i parametri $CC$ , $\alpha$ , $z_0$ e $B$ .

E' evidente che parametrizzazioni come queste influenzano in cascata vari calcoli del modello: è pertanto buona norma stimarne la sensibilità al cambiamento dei parametri in input per evitare comportamenti instabili in caso di piccole variazioni. In letteratura sono reperibili vari studi di sensibilità per il modello AERMOD (tra cui: Faulkner et al., 2008; Karvounis et al., 2007; Zou et al., 2010), ma a causa della non linearità tra le posizioni relative di sorgenti e recettori e delle condizioni meteorologiche, i vari studi sono difficilmente comparabili, sia tra loro che con il caso qui in esame.

Pertanto, uno studio di sensibilità è stato condotto per il giorno 10 novembre 2015, ed i risultati sono mostrati in Tabella 1. Le concentrazioni fornite dal modello sono state confrontate, per le cinque stazioni utilizzate, con quelle calcolate utilizzando i valori standard di  $\alpha$ ,  $z_0$  e  $B$ .

Parametro	Valore	Variazione media giornaliera delle concentrazioni stimate rispetto ai valori standard dei parametri(%)				
		Recettore	Recettore	Recettore	Recettore	Recettore
		A	B	C	D	E
$\alpha$	0.1	9.43	-0.55	-0.38	0	-0.55
$\alpha$	0.5	-4.59	-15.3	-8.1	0	-15.3
$z_0$	0.5 m	9.76	41.39	0.42	0	41.39
$z_0$	0.9 m	-1.56	-23.37	-0.56	0	-23.37
$B$	0.5	-3.98	-4.12	-7.49	0	-4.12
$B$	3	10.36	-1.3	-0.85	0	-1.3

*Tabella 1 - Variazioni della concentrazione predetta, al variare dei parametri  $\alpha$ ,  $z_0$  e  $B$ . I valori standard di riferimento sono, rispettivamente, 0,18, 0,7 e 1,5. Per il giorno prescelto i valori predetti al recettore D sono sempre nulli, a causa del vento prevalente da N-NO.*

Si nota come, anche per variazioni piuttosto grandi dei parametri, l'errore è generalmente contenuto. Inoltre è da considerare che il parametro  $z_0$  è importante soprattutto di notte, quando l'unica turbolenza è legata allo shear del vento, mentre i parametri  $B$  ed  $\alpha$  contano nelle ore diurne, essendo connesse ai regimi convettivi. C'è da osservare che il parametro  $z_0$  risulta piuttosto critico per le concentrazioni stimate in regimi stabili: nelle stazioni B ed E le variazioni di concentrazione connesse a variazioni di  $z_0$  tra 0,5 e 0,9 m risultano superiori al 50% del valor medio. In letteratura, inoltre, sono riportati valori tipici di  $z_0$  per ambienti industriali e urbani variabili tra 0,8 e 3 m. Sarebbe quindi importante stimare con maggiore precisione tale parametro, direttamente dai dati di anemometria sonica (Sozzi et al., 1998, Martano, 1999) oppure dai dati geometrici degli edifici e suolo (Macdonald et al., 1997).

L'importanza di questo studio è evidente se si nota che ai vari recettori si possono avere errori di segno opposto, il che di fatto cambia il pattern di distribuzione delle concentrazioni con cui la regressione lineare dovrà poi giustificare i dati osservati. Negli studi citati in letteratura si riscontrano anche calcoli di variabilità a seconda delle condizioni meteo, qualora le parametrizzazioni della micrometeorologia non siano efficaci. Infatti, sebbene allo stato della release attuale (15181 del 30/06/2015), alcune di queste quantità non vengono utilizzate da AERMOD - ma servono solo come step intermedi in AERMET per la stima di  $Z_i$  -, è possibile bypassare il preprocessore AERMET e fornire direttamente i dati di  $H_0$ ,  $L$ ,  $u^*$  misurati dagli anemometri sonici e, avendone la possibilità, anche dell'altezza  $Z_i$  dello strato limite. Infine, a tale proposito, si rammenta che la stima della copertura nuvolosa e dell'altezza dello strato limite planetario sarebbero in linea di principio possibile dall'analisi supervisionata dei dati della rete lidar installata nell'area. Mentre la stima di  $Z_i$  richiede l'intervento dell'operatore, in quanto allo stato attuale non esistono algoritmi universalmente riconosciuti per la stima attendibile dell'altezza del PBL dai dati lidar elastici, per la copertura nuvolosa sono invece disponibili diversi algoritmi automatici che potrebbero essere implementati per cercare di migliorare le stime dal modello; occorre tuttavia valutarne l'opportunità alla luce dei costi e benefici nell'economia generale della catena di processamento.

### 3.3.4 Conclusioni sull'utilizzo di AERMOD

Il modello AERMOD, validato in moltissime campagne come riportato in un gran numero di articoli, produce dati puntuali o mappe di concentrazione a partire da uno o più dati meteo e dal valore dei ratei emissivi delle singole sorgenti.

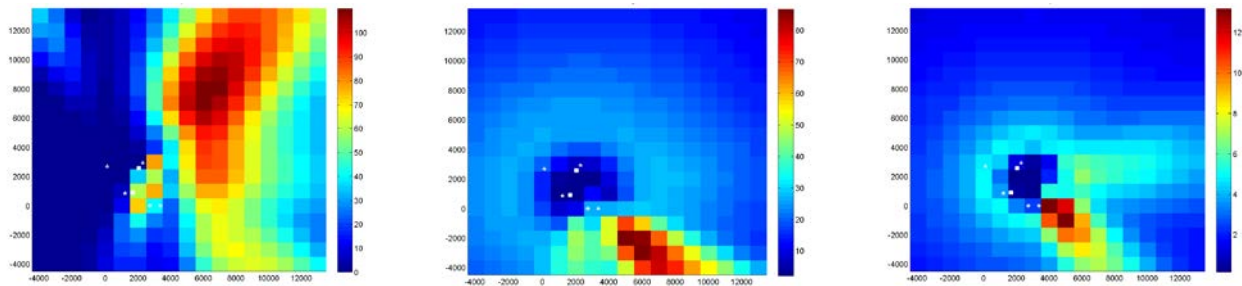


Figura 5. Esempio di mappe di Concentrazione relativa, predetta dal modello AERMOD per il giorno 29/12/2015. A sinistra è mostrato il pattern dovuto alle acciaierie, al centro all'altoforno e a destra all'agglomerato. I valori sono riferiti ad una emissione standard pari a 1000 g/s per ogni sorgente, per cui non forniscono valori assoluti, ma vanno considerate in relazione tra loro. I quadrati bianchi rappresentano le due acciaierie, mentre le stelline i recettori. La scala è in km. Per raffronto, in basso è mostrata un'immagine satellitare più o meno sovrapponibile alle mappe sovrastanti.

L'uso del modello AERMOD permette, ovviamente, lo studio dell'importanza relativa di eventuali sorgenti concorrenti: per conoscere i contributi effettivi i valori andrebbero moltiplicati per le rispettive emissioni. Per

le sorgenti areali si suppone che il rilascio sia affidato al vento, mentre per sorgenti puntuali occorre specificare la velocità di salita e la temperatura dei fumi, quindi la simulazione è soggetta ad ulteriori errori. Il valore dei campi di concentrazione può essere valutato su una serie di punti a scelta oppure su un grigliato regolare che dà luogo alle mappe riportate in figura 3. Dallo studio delle mappe si può notare che spesso, a causa probabilmente del galleggiamento e dello scarso rimescolamento, i valori di concentrazione in prossimità delle sorgenti sono nulli al suolo, salvo poi presentare massimi a qualche km di distanza.

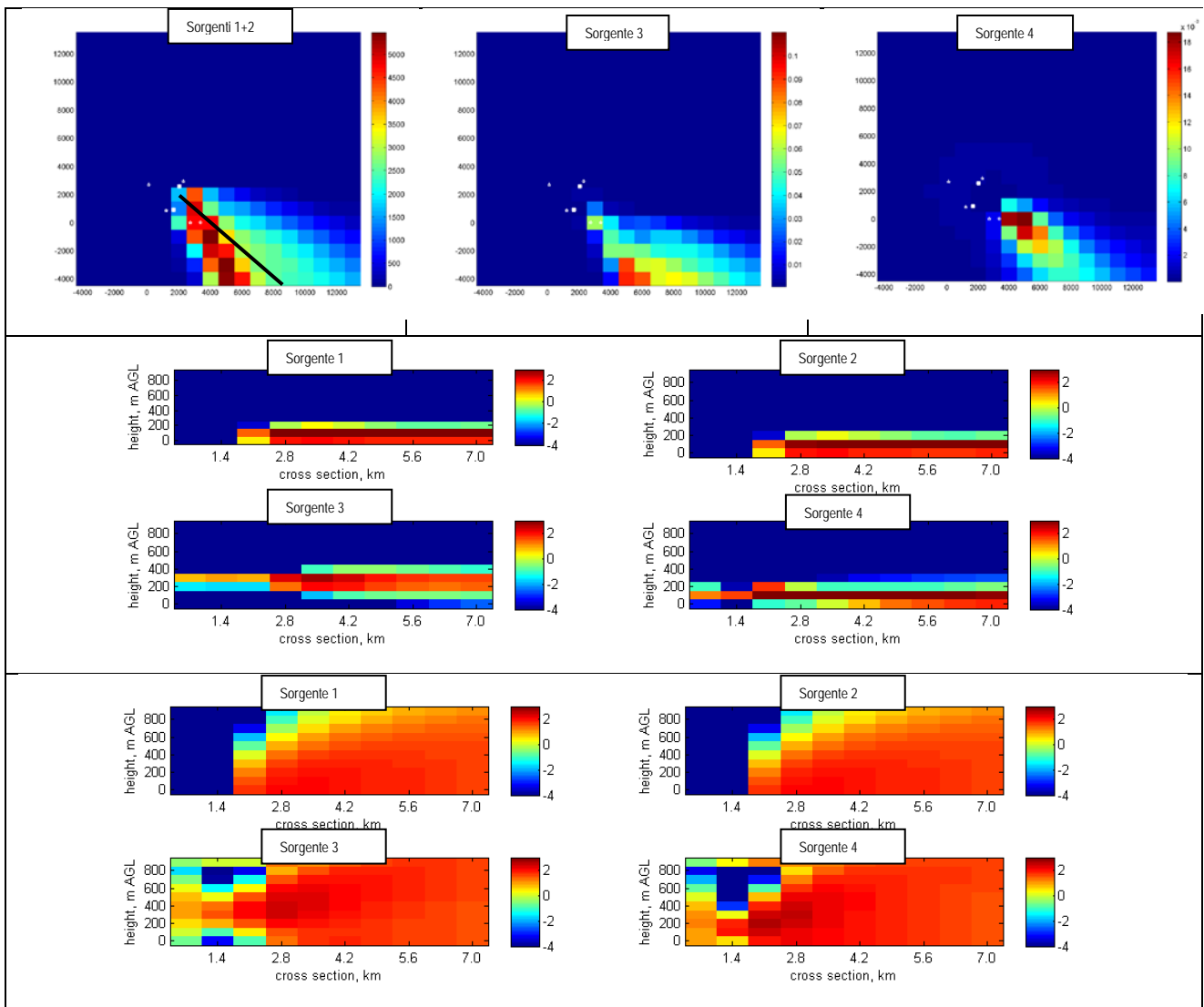


Figura 6. Esempio di mappe e sezioni della dispersione prevista da AERMOD. Nelle tre figure in alto sono mostrate le mappe al suolo (in unità arbitrarie) previste per ipotetiche emissioni da sorgenti diffuse (raggruppate, immagine a sinistra), da sorgenti puntuali basse (al centro) e da un camino elevato (a destra). Tutte le emissioni sono state poste uguali, per cui i grafici hanno il solo scopo di comparare i pattern relativi. I quadrati bianchi rappresentano le sorgenti, mentre le stelline i recettori. La linea nera rappresenta la sezione mostrata nelle immagini sotto, che mostrano le concentrazioni stimate lungo la sezione indicata, in direzione NW-SE, suddivise per ore notturne (grafici al centro) e giorno (grafici in basso). Le scale di colore indicano il logaritmo decimale della concentrazione, in unità arbitrarie.

Ciò si verifica principalmente durante le ore notturne, quando il flusso di calore negativo è indice di assenza di convezione e le plume restano per lo più confinate attorno alla quota di rilascio e si allargano lentamente a causa della turbolenza meccanica, schematizzata dalla variabile  $u^*$ . In tali occasioni l'effetto del down wash da parte degli edifici e dell'orografia può diventare importante. AERMOD può simulare effetti del genere, ma occorre immettere in input una mappa DEM ad alta risoluzione (<50m) per ottenere risultati attendibili. Per il momento tale opzione non è stata implementata.

In Figura 6 sono mostrate le mappe e le sezioni di concentrazione di PM calcolate per il giorno 12/12/2016. Il giorno è stato scelto in base al fatto che la direzione del vento è rimasta più o meno costante nell'arco delle 24 ore, e pertanto una rappresentazione integrata sulla giornata risulta rappresentativa di una situazione valida ora per ora lungo la stessa sezione. Tutte le concentrazioni sono riferite ad una emissione unitaria (1kg/s per ogni sorgente). Dalle mappe mostrate si nota che, specialmente per le sorgenti puntiformi (calde ed elevate), nei pressi della sorgente (~ 1km) i campi di concentrazione al suolo sono bassi, soprattutto nelle ore notturne, per i motivi già discussi.

Da un lato, ciò dovrebbe garantire lo scarso peso sulle centraline di queste sorgenti rispetto alle sorgenti diffuse, dall'altro mostra come molte emissioni potrebbero divenire importanti a distanze maggiori di quelle della rete attuale. Vista l'utilità della visualizzazione delle sezioni verticali, nel sistema informativo rilasciato è stata implementata l'opzione per calcolare le mappe su un dominio tridimensionale di 20x20 km, e dello spessore di 500 m in sei livelli. Altrimenti, l'opzione standard prevede il salvataggio delle sole griglie al suolo con dominio di 20 x 20 km.

#### **4 Modello di regressione multilineare (MLR)**

Una volta determinati i coefficienti di accoppiamento della matrice  $\beta$ , occorre risolvere il problema inverso di calcolare i valori delle emissioni e del fondo che giustifichino al meglio il vettore delle osservazioni.

La linearità del problema descritto dalla equazione (1) rispetto ai coefficienti di  $\beta$  permette l'inversione mediante la tecnica standard di regressione lineare multipla (MLR). La MLR è basata sullo stesso approccio statistico alla base delle regressioni lineari tra due n-uple di valori, con la differenza che la combinazione lineare di più variabili concorre a descrivere il vettore dei dati osservati. Affinché tale operazione sia attendibile, però, occorre che siano verificate un certo numero di assunzioni, come descritto anche nella norma EN UNI 15445.

1. Le colonne della matrice  $b$  rappresentano i vettori di accoppiamento di ogni sorgente con tutti i recettori. E' evidente che se i pattern sono molto simili tra loro, non è facile risolvere il problema in maniera attendibile, in quanto le sorgenti verranno 'viste' dall'array dei recettori più o meno nella stessa maniera. Da un punto di vista matematico, ciò è espresso dal fatto che la matrice  $b$  è quasi singolare (o al limite singolare, se due colonne sono uguali). Per evitare che ciò accada occorre controllare che i coefficienti di correlazione lineare tra tutte le coppie di sorgenti sia bassa (è

suggerita una soglia pari a 0.5). In caso ciò non avvenga, le due sorgenti interessate vanno accorpate e considerate come un'unica sorgente.

2. La risoluzione raccomandata dalla norma è di un'ora. Le matrici di accoppiamento per le sorgenti diffuse mostrano pattern molto direzionali, a causa probabilmente del rilascio a velocità verticale nulla. Ciò comporta che molti coefficienti della matrice  $b$  siano nulli, rendendo il problema sempre meno determinato. In tal caso, i valori misurati in queste centraline devono fornire il fondo, per lo meno rispetto alle sorgenti considerate.
3. Si può verificare l'eventualità che i dati di PM in qualche centralina o i fattori di accoppiamento di qualche sorgente siano mancanti. In tal caso la matrice scende di dimensionalità e, pur non segnalando errore da un punto di vista matematico, l'attendibilità dell'inversione diviene fortemente dubbia (è come far passare una retta per due soli punti, da un punto di vista formale non ci sono problemi, ma non c'è alcuna ridondanza per stimare l'errore).
4. Mediando la matrice degli accoppiamenti sulle 24 ore si ottengono matrici con meno coefficienti nulli, il che dovrebbe migliorare la stabilità del calcolo della regressione. Questa ulteriore possibilità è stata implementata, e il sistema fornisce anche i risultati su base giornaliera.
5. In caso di pioggia il modello AERMOD non è in grado di simulare la dispersione, e l'analisi non può essere portata a termine.

Ulteriori dettagli sulle regressioni lineari multiple sono riportati in molti testi di statistica descrittiva, ad esempio: Radhakrishna Rao, *Linear models and generalizations* (Springer 2008).

Spendiamo a questo punto una parola sulla linearità del problema. Per come è stato formulato, il problema è, istante per istante, lineare nei coefficienti  $E_0 \dots E_m$ , come già osservato. Tuttavia, trattandosi di sorgenti fuggitive, per definizione queste dipendono dal vento (si pensi ad un cumulo di terra che viene eroso dal vento, o alla polvere estratta dall'interno di un edificio a causa del vento). Pertanto, sebbene il problema sia istante per istante lineare, ci aspettiamo che i coefficienti di emissione possano, in generale, risultare dipendenti da velocità e direzione del vento. Questo ovviamente potrà essere studiato su base statistica solo dopo un'ampia casistica di *retrieval* dei fattori di emissione sotto diverse condizioni meteorologiche.

#### **4.1 Procedura Multilinear Regression Model**

La procedura di regressione lineare fornisce, oltre ai parametri  $E_1, \dots, E_m$ , al valore dell'intercetta  $E_0$  e ai residui  $\varepsilon_1, \dots, \varepsilon_n$ , alcuni importanti parametri statistici per valutare la bontà dei dati calcolati. Il valore  $E_0$  indica la concentrazione di fondo media osservata, in  $\text{mg}/\text{m}^3$ . I valori  $E_1, E_m$  rappresentano i coefficienti da utilizzare per normalizzare le emissioni imposte in input al modello di dispersione. Infatti, affinché le colonne di  $\beta$  siano confrontabili, i valori di emissione impostati nel modello devono essere riferiti ad una emissione unitaria. Nel nostro caso, per evitare arrotondamenti a zero o overflow nell'output è stato impostato un valore di 1000 g/s.

##### **4.1.1 Correlazione tra sorgenti**

La norma EN UNI 15445 raccomanda come primo step il controllo della correlazione tra le colonne della matrice  $b$ . Come accennato in precedenza, se le due colonne sono troppo simili ( $R > 0.5$ ), da un punto di vista matematico la matrice risulta quasi singolare, e diventa di difficile trattamento. Da un punto di vista fisico,

invece, non ci si può aspettare di riuscire a separare i contributi di due sorgenti che creano dei campi molto simili. Questo fenomeno deve essere limitato il più possibile progettando la rete di recettori in maniera opportuna. In particolare, è opportuno che alcuni recettori siano compresi tra le varie sorgenti, in modo che l'accoppiamento cambi in maniera sostanziale al variare della direzione del vento. Come regola generale, la norma UNI EN 15445 richiede che il numero dei recettori sia più che doppio rispetto al numero delle sorgenti, così da avere un problema sufficientemente sovradimensionato da essere statisticamente robusto.

In caso di correlazione alta occorre accorpare le due sorgenti interessate e considerarle come un'unica sorgente. Se le sorgenti sono solo due, accorpendole (sommando cioè i vettori colonna della matrice  $\beta$ ) il problema diventa una semplice regressione lineare: in pratica un fit lineare tra le concentrazioni osservate e il campo predetto dal modello. Avendo più di due recettori a disposizione il problema è correttamente sovra-determinato, ma se consideriamo che spesso (soprattutto trattando dati a risoluzione oraria) i coefficienti di  $\beta$  possono essere nulli su due o più stazioni, il problema può diventare esattamente determinato. Infatti, quando le sorgenti sono accoppiate, in genere si assiste anche alla presenza di varie righe nulle nella matrice in quanto alcune centraline risultano 'sopravento' e quindi non accoppiate alle sorgenti. Questo effetto, invece che migliorare le condizioni di inversione le peggiora, poiché la procedura di regressione viene a condursi su due soli parametri da stimare (fondo + emissione) e due o tre dati noti (le centraline non nulle, in quanto le altre rappresentano per forza il fondo).

In tal caso si può ancora far passare la retta di regressione (il cui coefficiente angolare rappresenta l'emissione cumulata delle due sorgenti), ma non vi è alcuna ridondanza per calcolare la significatività né la possibilità che altre sorgenti influiscano, pertanto l'intera procedura perde di senso.

Per questo motivo è stata implementata anche la possibilità di condurre un'analisi su base giornaliera, in quanto in questo modo la matrice  $\beta$  risulta molto meno vuota, permettendo una inversione meno critica. Purtroppo, però, all'aumentare del periodo di media il pericolo di includere altre sorgenti di limitata durata aumenta (es.: traffico veicolare, movimentazione mezzi e terra), e l'analisi può quindi essere inficiata da motivi differenti.

Osserviamo, infine, che i singoli coefficienti di correlazione tra i vettori della matrice  $\beta$  e il vettore delle osservazioni rendono conto di quanto ogni sorgente sia correlata con il pattern osservato. E' possibile infatti che, nonostante nessuna delle sorgenti sia linearmente correlata all'osservazione, una loro combinazione lineare lo sia (e questo è il ragionamento alla base del metodo MLR), se i vettori risultano scorrelati tra loro. Nel sistema in questione, se i vettori di accoppiamento risultano invece correlati significativamente, ed occorre quindi accorpare le sorgenti, ci si aspetta allora che il coefficiente di correlazione tra il vettore delle osservazioni e l'unico vettore degli accoppiamenti superstiti sia alto. In tal caso, questa correlazione rappresenta un'indicazione utile a capire se vi può essere una buona spiegazione in termini di una sorgente sola.

### 4.1.2 Coefficienti negativi e regressione vincolata

Per quanto riguarda i coefficienti  $E_1, E_2$ , occorre notare che spesso si ricavano dei coefficienti negativi. Questo accade quando il vettore delle concentrazioni è anticorrelato ai vettori degli accoppiamenti, per cui per spiegare la varianza osservata occorre ‘togliere’ da un punto di vista formale una parte di una sorgente. Ovviamente questa operazione non ha senso fisico, in quanto ciò succede presumibilmente a causa di sorgenti non considerate, che alterano il pattern di concentrazione rispetto alle informazioni inserite nel modello. Purtroppo in questo caso non c’è niente da fare, se non introdurre altri dati di sorgenti e di osservazioni e aumentare la dimensionalità del problema. Per risolvere questo problema, per lo meno da un punto di vista formale, è stata implementata una procedura di regressione vincolata, in cui si è imposto che i valori dei coefficienti siano non negativi.

In dettaglio, il problema da risolvere è trovare i valori di  $E$  tali che:

$$\min_E \frac{1}{2} \|\beta E - c\|, \text{ con il vincolo su } E \text{ tale che } E > 0.$$

Tale modello è ben più complesso di un semplice modello di regressione multipla, e viene risolto con metodi di programmazione lineare. La descrizione del metodo si può trovare in Gill et al. (1981).

Il risultato di tale modello è dato dalla coppia di valori  $E_1, E_2$ , dal valore dell’intercetta  $E_0$  e dai residui  $\varepsilon_1, \dots, \varepsilon_n$ . Per definizione, i valori  $E_1, E_2$  sono non negativi, mentre  $E_0$  ed i residui sono simili a quelli calcolati con il metodo MLR. Proprio per l’approccio non statistico, ma piuttosto variazionale adottato, non sono disponibili test statistici di validità dell’inversione. Cionondimeno, l’analisi dei residui è sempre possibile.

### 4.1.3 Normalità dei residui

Infatti, una delle ipotesi alla base delle procedure di MLR è che i residui siano distribuiti normalmente attorno allo zero. Ciò si può verificare sia mediante i cosiddetti normal plot, sia utilizzando test statistici appropriati, come i test di Kolmogorov-Smirnov (K-S) e di Shapiro-Wilk (S-W). Infine, lo scarto quadratico medio dei residui dà un’indicazione dell’errore commesso nell’approssimazione lineare.

Il sistema implementato restituisce il risultato del test K-S per i residui, sia su base oraria che su base giornaliera. Il test K-S è un test non parametrico che confronta due distribuzioni arbitrarie, in particolare la distribuzione dei residui osservata con la distribuzione normale standard. L’ipotesi nulla  $H_0$  è che le due distribuzioni siano uguali, l’ipotesi alternativa è che esse siano diverse. Un risultato pari a 0 indica quindi che, con un livello di confidenza del 95%, i residui sono distribuiti normalmente, mentre un risultato di 1 indica una distribuzione non normale. Ciò può essere attribuito ad errori sistematici nel modello, come ad esempio il contributo di sorgenti spurie: in tal caso non ci si aspetta infatti che i residui siano distribuiti normalmente, ma si evidenzia un pattern sistematico che altera la distribuzione degli errori (i residui). E’ da osservare, in conclusione, che la condizione di normalità dei residui è condizione necessaria ma non sufficiente per la bontà dell’inversione del modello.

Lo stesso risultato è prodotto anche mediante il test di S-W, che è fornito in aggiunta al test K-S. Un risultato di 1 in uno dei due indica quindi che l'inversione può essere cattiva, seppur determinata da un punto di vista matematico.

Infine, il normal-plot mostra lo scatterplot dei residui con quello teorico associato ad un profilo gaussiano. Tanto più i punti corrispondenti ai residui sono posti sulla bisettrice, tanto migliore è la distribuzione dei residui. Un normal-plot dei residui è mostrato in Figura 7. Il test di S-W è più conservativo di quello di K-S, per cui è più probabile che rigetti l'ipotesi nulla.

Il normal-plot viene prodotto su indicazione della norma EN UNI 15445, sebbene il valore del test S-W racchiuda in sé il risultato del test di normalità e di fatto risulta più utile e conciso del normal-plot.

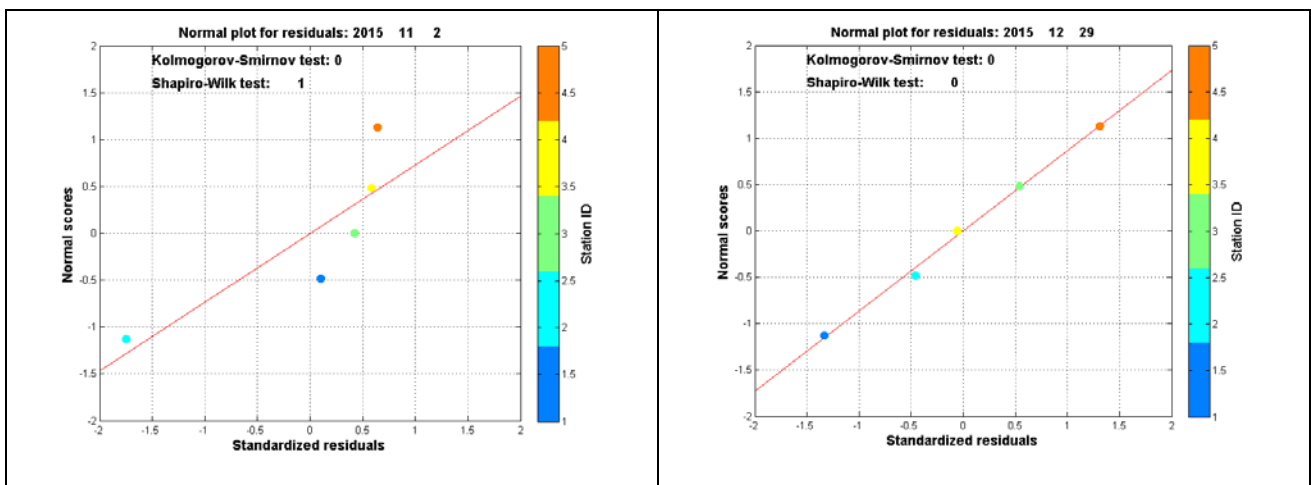


Figura 7. Normal plot per i residui del MLR. A sinistra i dati relativi al 2 novembre, a destra i dati relativi al 29 dicembre. E' mostrato anche il risultato dei test K-S e S-W. Il test fallito indica che i residui non sono distribuiti normalmente.

#### 4.1.4 Varianza spiegata

Oltre al test di normalità per i residui, la procedura MLR produce un valore di correlazione  $R^2$  che indica la percentuale di varianza del vettore osservazione spiegata dal modello. Ovviamente, tanto maggiore è  $R^2$ , tanto migliore è l'interpolazione trovata dalla procedura di inversione.

#### 4.1.5 Test F

Oltre al valore di  $R^2$ , si può effettuare il test F, che testa invece l'ipotesi nulla che tutti i coefficienti della regressione siano nulli:

$$H_0: E_1=E_2=0$$

contro l'ipotesi alternativa

$$H_1: \exists n | E_n \neq 0.$$

Ciò descrive in pratica se il modello spiega i valori osservati meglio di una costante. E' da notare che il superamento del test non indica quale sia il coefficiente più importante, dice solo che almeno uno è diverso da zero. Il superamento del test F è dato dal confronto del suo p-valore con la soglia di riferimento (es. 0.05, pari al 95% di confidenza). Il p-valore è mostrato in uscita (piuttosto che la variabile logica 1/0 di accettazione/rigetto) per dar modo di accettare o rigettare l'ipotesi anche con valori diversi di confidenza. Anche il superamento del test F è condizione necessaria ma non sufficiente per la bontà dell'inversione dei dati.

## 4.2 Restituzione dei valori

Verrà qui discusso l'insieme dei dati che il modello restituisce alla fine di ogni simulazione. Il file di uscita contiene infatti i valori della regressione, sia vincolata che non, con gli intervalli di confidenza associati (al 95%), assieme ai valori dei residui, e i valori dei test statistici richiesti.

### 4.2.1 Dati in uscita

Alla fine del processo, sia per le analisi orarie che per le analisi giornaliere, il sistema restituisce un file con le informazioni contenute in Figura 8. E' possibile creare sia un documento ascii che racchiuda tutti i dati come in figura (in formato csv), che un documento pdf che oltre alla tabella racchiude un header di descrizione e classificazione dei risultati.

Nel file di testo sono contenuti tutti i parametri utili a valutare la riuscita o meno del test, ed il grado di attendibilità, oltre ai valori dei ratei di emissioni dalle sorgenti calcolate dal metodo di source apportionment. Per semplicità di interpretazione, nel documento pdf è stata apposta una intestazione colorata e con un'etichetta che descrive la riuscita del modello. Ciò è stato fatto in base ad una scelta arbitraria degli autori, e non trova supporto nella norma EN UNI 15445. *Le informazioni fornite dall'intestazione sono pertanto da considerarsi puramente indicative*; infatti sono accompagnate sempre dalla lista completa di variabili e risultati dei test calcolati dal modello.

```

===== STATISTICS =====
Mean observed PM10 ug/m3, 44.718
Correlation among sources (grouping if >0.5), 0.872
Corr. PM-ACC1, 0.827
Corr. PM-ACC2, NaN
===== MultiLinear Regression =====
MLR total R^2 (% variance explained), 68.339
MLR F-test (accept under 0.05), 0.084
Kolmogorov-Smirnov normality test on residuals (accept if 0), 0.000
Shapiro-Wilk normality test on residuals (accept if 0), 0.000
PM10 background (with 95% confidence bounds) ug/m3, 26.586 , -2.061 , 55.232
MLR Acc1 coefficient g/s (with 95% confidence bounds), 2.34e+000 , -5.86e-001 , 5.26e+000
MLR Acc2 coefficient g/s (with 95% confidence bounds), NaN , NaN , NaN
Residuals ug/m3, -0.213 , 11.117 , 5.247 , 1.406 , -17.557
===== Constrained least squares regression =====
CLRR Intercept ug/m3, 26.586
CLSR Acc1 coefficient g/s, 2.34e+000
CLSR Acc2 coefficient g/s, NaN
Residuals ug/m3, 0.213 , -11.117 , -5.247 , -1.406 , 17.557
Kolmogorov-Smirnov normality test on residuals (accept if 0), 0
Shapiro-Wilk normality test on residuals (accept if 0), 0

```

-----  
 Figura 8. File di uscita della catena modellistica di source apportionment per i dati medi giornalieri (dati del 29 dicembre). Il formato è un file ascii in formato csv utilizzabile con qualsiasi foglio di calcolo.

La logica del metodo di classificazione del risultato è mostrato in Figura 9. Il parametro che domina la classificazione è il test S-W, che definisce se i residui sono distribuiti normalmente. Infatti, come discusso in precedenza, il fallimento del test indica che probabilmente ci sono sorgenti di cui non si è tenuto conto, invalidando quindi l'intera procedura.

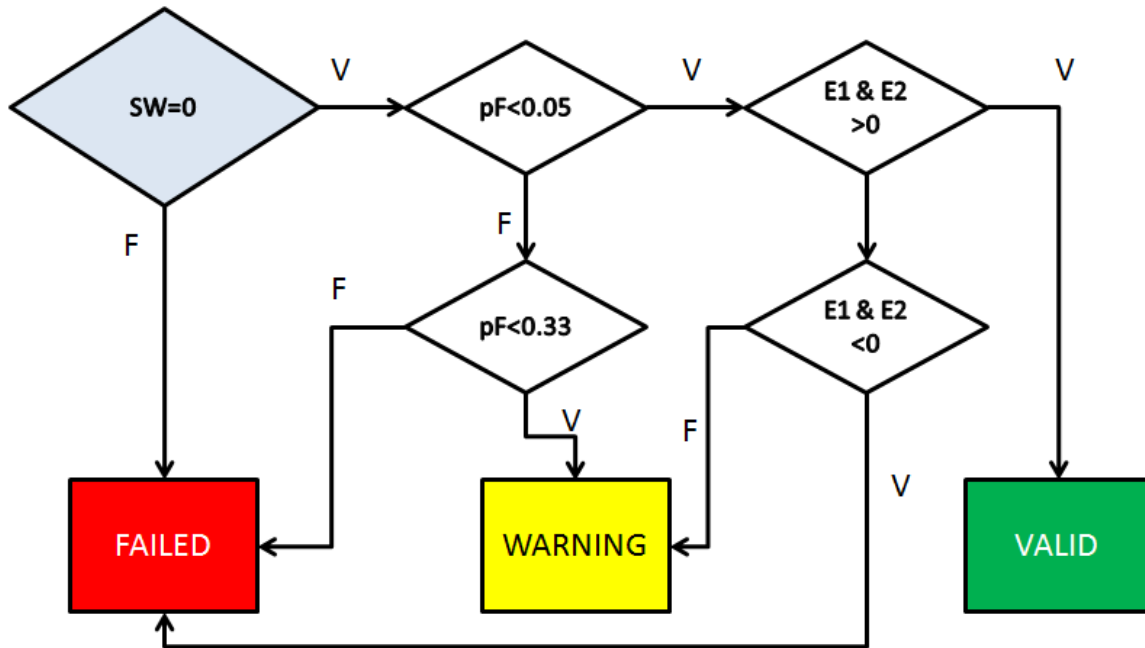


Figura 9. Logica di classificazione dei risultati del modello.

Se il test di S-W è superato, si osserva il p-valore del test F: come detto, anche questo è importante in quanto stabilisce se il modello di regressione lineare spiega la distribuzione dei dati meglio di un modello costante. Si è scelto di utilizzare due livelli di confidenza: il 95% per stabilire se il modello ha funzionato bene, ed il 67% se ha funzionato in modo non del tutto soddisfacente. Valori di confidenza inferiori al 67% sono considerati insoddisfacenti, e la prova viene considerata fallita.

Se il p-valore è buono (<0.05, 95% di confidenza) si controllano i valori delle emissioni: se sono entrambi positivi il test è superato. Se uno dei due è negativo, la procedura restituisce un warning poiché, essendo i valori negativi inammissibili, occorre valutare attentamente tutti i parametri. In questo caso si possono ad esempio controllare i dati della regressione vincolata. Se invece i coefficienti di emissione sono entrambi negativi, il test si considera fallito. E' da notare che il valore di 0.33 corrisponde ad una scelta arbitraria.

Se infine il p-value è maggiore di 0.05, si controlla se esso è inferiore a 0.33 (67% di confidenza). Se il test è superato, la procedura restituisce warning, altrimenti il test è fallito (troppa poca confidenza: utilizzare un

modello costante spiega altrettanto bene i dati in input). I casi di warning ovviamente sono i più delicati, in cui occorre valutare attentamente tutti i parametri per trarre conclusioni.

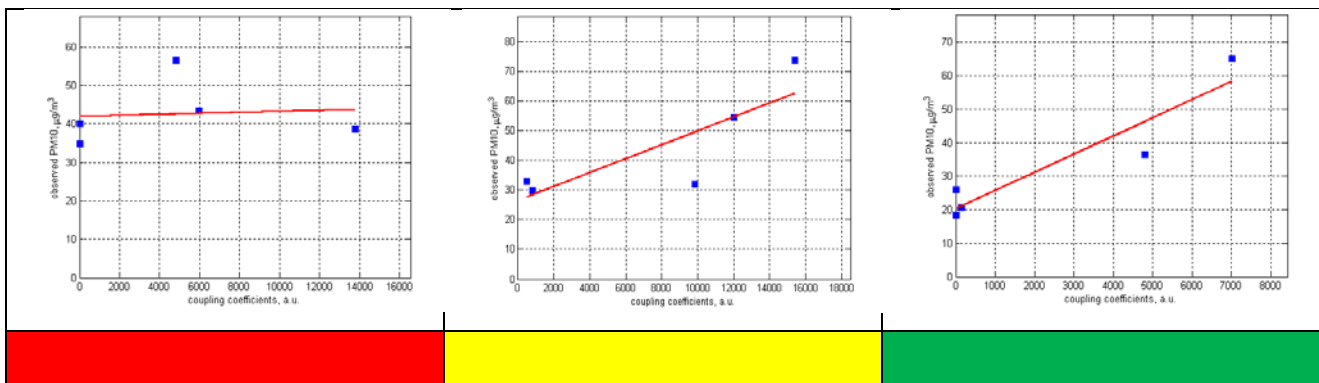
#### 4.2.2 Raggruppamento delle sorgenti

Come già discusso, all'inizio della procedura MLR si effettua un test di correlazione tra i vettori di accoppiamento sorgenti-recettori per valutare se le sorgenti possono essere considerate indipendenti o meno. Nel caso in cui il coefficiente di correlazione lineare tra due vettori risultino correlati oltre 0.5, occorre raggruppare le sorgenti corrispondenti e considerarle una sola. Nel caso in cui le sorgenti siano solo due, ciò equivale a considerare una sola sorgente, per cui il vettore degli accoppiamenti *deve* risultare linearmente correlato al vettore delle osservazioni (le concentrazioni medie osservate nelle corrispondenti centraline). Questo vuol dire che la regressione collassa su una regressione lineare ad una dimensione, e la procedura equivale a calcolare il best fit ad una sola dimensione.

Se succede ciò, il sistema passa automaticamente ad escludere la seconda sorgente (nan è posto come valore di emissione per la seconda sorgente), ma è possibile visualizzare il grafico che unisce le osservazioni al vettore di accoppiamento medio delle due sorgenti.

Un esempio che illustra bene il significato dei test statistici che corredano le analisi è mostrato in Figura 10 (dati giornalieri):

- Nel grafico di sinistra si ha un test fallito, a causa del fatto che il test F ha un p-valore molto alto: 0.9 e il modello spiega di conseguenza meno dell'1% della distribuzione delle osservazioni. Ciò è ragionevole in quanto non è evidente una correlazione tra i dati osservati e quelli predetti dal modello.
- Al centro (grafico relativo ai dati di Figura 8) diviene evidente la correlazione tra le due misure: in tal caso la classificazione è di warning solo perché il p-valore del test F è superiore a 0.05, seppur di poco ( $\approx 0.09$ ). Sebbene non rispetti tutti i criteri richiesti dalla norma, si può dire che il risultato è attendibile con il 91% di confidenza circa. La moderata dispersione dei dati attorno alla retta è condensata nel coefficiente  $R^2 = 68.339\%$ , che significa che il modello spiega quasi il 70% della varianza osservata delle osservazioni.
- Nel grafico di destra si ha infine un test valido. In questo caso la correlazione tra le sorgenti è pari al 93%, che quindi sono state raggruppate, ed il rateo totale di emissione risulta essere 5.43 g/s. Il test F è passato con 0.019 e la varianza spiegata è molto alta, pari all'87%.



*Figura 10. Grafico della regressione nel caso di accorpamento delle sorgenti.*

*A sinistra caso 'failed', al centro caso di 'warning', a destra caso 'valid'.*

## 5 Conclusioni generali

Per valutare le prestazioni del modello inverso c'è da tenere conto della serie di test statistici che la procedura MLR deve soddisfare. D'altra parte c'è però da considerare la significatività dei risultati del modello diretto. Data infatti la grande varietà di fattori che possono inficiare il meccanismo di *apportionment*, affinché la regressione possa essere considerata valida occorre che tutti i test siano superati con successo.

Più difficile è stimare le prestazioni del modello, il quale è dimostrato funzionare bene in generale, tanto da essere considerato tra i *preferred models* dall'EPA. Come è noto, però, specie nei periodi di transizione tra condizioni stabili e convettive (e viceversa) le stime possono essere falsate da una dinamica e una parametrizzazione troppo semplificata (ma questo è vero anche per modelli più complessi!). Allungare il tempo di media sicuramente diminuisce l'incidenza di questi errori.

Saranno qui discusse eventuali migliorie che possono essere introdotte per migliorare le prestazioni della procedura. Per il modello diretto:

- una fonte di possibile errore è data dall'effetto di downwash da parte degli edifici. Questo effetto può essere considerato in AERMOD previo inserimento della topografia ad alta risoluzione, che al momento non è stato possibile implementare, in quanto richiede un notevole lavoro;
- Una migliore stima dei parametri del suolo (rapporto di Bowen, albedo, rugosità superficiale) può sicuramente migliorare le simulazioni di AERMOD, eventualmente introducendo la differenziazione per settori di provenienza del vento (per  $z_0$ );
- Inoltre, AERMOD ma soprattutto il suo preprocessore AERMET potrebbe essere modificato per calcolare il flusso di calore sensibile direttamente dai dati micrometeorologici, saltando le parametrizzazioni, così come l'altezza dello strato mescolato convettivo e stabile, e la copertura nuvolosa. Questi valori potrebbero essere estratti dai dati della rete lidar, ma occorre valutare accuratamente il rapporto costi/benefici in quanto alcune tipologie di dato non sono al momento unanimemente considerate valide dalla comunità scientifica (es.: metodi di calcolo automatico dell'altezza dello strato limite convettivo, e soprattutto di quello stabile).

Per quanto riguarda il miglioramento della procedura di regressione, si è già notato che il problema principale di questo passo è dato dal fatto che spesso molti coefficienti di accoppiamento sorgenti-recettori vengono nulli, specialmente su tempi di media brevi (le plume sono molto direzionali e pertanto investono ora per ora solo una o due centraline). Ciò fa sì che la matrice da invertire risulti piena di zeri, e a limite quasi singolare, e costringa quindi a scendere di dimensione. Ciò richiede l'accorpamento delle sorgenti, obbligando ad una regressione lineare singola. Pertanto:

- l'unico margine per migliorare la parte di inversione risiede nel posizionamento e nel numero delle centraline, in modo da evitare il più possibile altre sorgenti interferenti ed avere in media un accoppiamento buono (non nullo e scorrelato) tra le varie sorgenti e i recettori, ma ciò ovviamente

richiede una pianificazione preventiva, possibile dopo l'analisi di una lunga serie temporale di misure meteorologiche e stime di dispersione.

D'altro canto, come notato da Mensink et al., 2011, la regressione risulta molto sensibile alle sorgenti non considerate, per cui semplificare troppo il problema può portare ad avere soluzioni prive di senso fisico (ad esempio quando alcune sorgenti sono sistematicamente associate ad emissioni negative).

Ovviamente, l'utilizzo del modello diretto può risultare comunque utile a prescindere dalla procedura di inversione, per valutare gli impatti delle sorgenti considerate ma anche di altre eventualmente implementabili (camini, strade, parchi) sia su base puntuale che da un punto di vista statistico. Questo studio, inoltre, permetterebbe di quantificare, per lo meno in prima approssimazione, l'impatto relativo delle diverse sorgenti alle varie centraline, per capire quanto sia ragionevole l'ipotesi di trascurare queste sorgenti nel processo di source apportionment, che come detto, è particolarmente sensibile alle sorgenti mancanti.

Concluderemo ora con la discussione di una possibilità che è emersa alla luce dei risultati dei *run* del modello di source apportionment su vari giorni. Questa trae spunto sia dalle caratteristiche della matrice di accoppiamento che si ottiene normalmente, che dall'utilizzo delle concentrazioni di fondo, usate in alcuni lavori per identificare e quantificare vari tipi di emissioni (Cosemans et al., 2008; Cosemans and Mensink, 2008; Cosemans and Mensink, 2011; Mensink et al., 2011).

In questi lavori gli autori hanno sottratto, dalle concentrazioni osservate, i valori misurati in apposite stazioni di fondo, per calcolare i contributi delle sorgenti locali. Questa è una difficoltà che non può essere superata in maniera diretta per tutte le aree in cui non siano disponibili dati orari di fondo. L'utilizzo di AERMOD ci può però venire in aiuto: si potrebbero infatti utilizzare, per stimare il fondo ora per ora, i dati delle centraline per cui gli accoppiamenti con le sorgenti sono nulli. Visto che, specie su base oraria, almeno due o tre centraline risultano non esposte alle sorgenti in base ai calcoli di AERMOD, la media di tali valori permette un semplice calcolo del fondo (per lo meno rispetto alle sorgenti considerate). Questo, come evidenziato dalle analisi dei dati delle centraline di monitoraggio della qualità dell'aria, è variabile nell'arco della giornata, per cui non è pensabile di confrontare direttamente tra loro le concentrazioni misurate ad ore diverse. Se invece la sorgente si può considerare uniforme, i contributi calcolati possono essere confrontati tra le varie ore, previa normalizzazione con i coefficienti di accoppiamento. Infatti, a questo punto si può costruire una matrice  $b$  che contiene tutti e soli i vettori di accoppiamento non nulli per l'intera giornata, ed effettuare la regressione confrontandola con il vettore dei contributi. Ci si aspetta allora di dover invertire una matrice ampiamente sovradimensionata (la matrice  $b$  conterrà così fino ad un massimo di 120 righe per tre colonne), che quindi porta ad una regressione ragionevolmente più significativa di quella ottenibile considerando solo i cinque punti delle stazioni, anche se mediati su tutta la giornata. Altro utile output di questo approccio è rappresentato dal vettore dei contributi: nei casi in cui questi risultano negativi vuol dire che, per lo meno nell'ora corrispondente, il fondo è maggiore dell'incremento dovuto alle sorgenti considerate. Quindi il problema è fisicamente mal condizionato: il fatto che il fondo risulti maggiore delle

concentrazioni sotto vento significa che ci sono in gioco delle altre sorgenti che impattano nelle zone in cui dovrei osservare valori minimi rispetto alle sorgenti considerate.

Unico requisito di tutto ciò è che la sorgente sia abbastanza costante da non immettere troppa dispersione nella serie temporale dei contributi.

Questo approccio non è ancora stato proposto da nessuno, per lo meno a conoscenza degli autori, pertanto non è descritto né in letteratura né tantomeno in linee guida, sembra però in teoria promettente.

## 6 Bibliografia essenziale

- Cimorelli et al., *AERMOD: a dispersion model for industrial source applications. Part I: general model formulation and boundary layer characterization*, J. Appl. Met., 44, 682-693, 2005
- Cosemans, G., Kretzschmar, J., & Mensink, C. (2008). Pollutant roses for daily averaged ambient air pollutant concentrations. *Atmospheric environment*, 42(29), 6982-6991.
- Cosemans and Mensink, *Reverse modelling for the determination of fugitive sources of PM10*, proc. 11<sup>th</sup> HARMO conference, Cambridge, UK, July 2-5, 2007
- Cosemans and Mensink, *Reverse modelling for the determination of fugitive sources of PM10*, Int. J. Env. Poll., 44, 1-4, 2011
- Denby et al., *Estimating domestic wood burning emissions of particulate matter in two nordic cities by combining ambient air observations with receptor and dispersion models*, CI&CEQ 16 (3) 237-241 (2010)
- Duan et al., *Evaluation and source apportionment of heavy metals (HMs) in sewage sludge of municipal wastewater treatment plants (WWTPs) in Shanxi, China*, Int. J. Of environmental research and public health, 2015, 12, 15807-15818
- Faulkner et al., *Sensitivity of Two Dispersion Models (AERMOD and ISCST3) to Input Parameters for a Rural Ground-Level Area Source*, J. Air & Waste Managemen. Assoc. 58: 1288-1296, 2008
- Frangkou et al., *Current trends in the use of source apportionment of air pollutant in Europe*, Int. J. Environment and Pollution, Vol. 50, Nos. 1/2/3/4, 2012
- Fushimi, *Source apportionment based on an atmospheric dispersion model and multiple linear regression*, Atmos. Env. 39 (2005) 1323-1334
- Gill, P.E., W. Murray, and M.H. Wright, *Practical Optimization*, Academic Press, London, UK, 1981
- Karvounis et al., *Sensitivity of Two Dispersion Models (AERMOD and ISCST3) to Input Parameters for a Rural Ground-Level Area Source*, Proc. 11th Int. Conf. on harmonisation within atmospheric dispersion modelling for regulatory purposes, Cambridge, UK, July 2-5, 2007
- Laupsa et al., *Source apportionment of particulate matter (PM2.5) in an urban area using dispersion, receptor and inverse modelling*, Atmos. Env. 43 (2009) 4733-4744
- Macdonald, R. W.; Griffiths, R. F.; Hall, D. J. *An improved method for the estimation of surface roughness of obstacle arrays*. Atmospheric environment, 1998, 32.11: 1857-1864.
- Martano, *Estimation of Surface Roughness Length and Displacement Height from Single-Level Sonic Anemometer Data*, J appl. Met 1999
- Mensink, C., et al. *Quantification of diffuse and fugitive PM 10 sources by integrated "hot-spot" method*. Atmospheric Environment, 2011, 45.13: 2233-2241.
- Norma UNI EN 15445: *Emissioni da fughe diffuse relative ai settori industriali – Determinazione dei sorgenti di polveri da fughe mediante metodo RDM (modello di dispersione inversa) – luglio 2008*

- PERRY, Steven G., et al. *AERMOD: A dispersion model for industrial source applications. Part II: Model performance against 17 field study databases*. Journal of Applied Meteorology, 2005, 44.5: 694-708.
- Radhakrishna Rao, Shalabh, Toutenburg, Heumann, *Linear Models and Generalizations - Least Squares and Alternatives*, Springer Series in Statistics (2008), ISBN: 978-3-540-74226-5
- Sozzi R., *La micrometeorologia e la dispersione degli inquinanti in aria*, APAT CTN\_ACE, 2003, disponibile all'indirizzo: <http://www.arpad.toscana.it/temi-ambientali/aria/modellistica-per-la-qualita-dellaria/linee-guida/apat-micrometeorologia.pdf>
- Sozzi, R.; Favaron, M.; Georgiadis, T., *Method for estimation of surface roughness and similarity function of wind speed vertical profile*. Journal of Applied Meteorology, 1998, 37.5: 461-469.
- STULL, *An introduction to boundary layer meteorology*. Springer Science & Business Media, 2012.
- Viana et al., *Source apportionment of particulate matter in Europe: A review of methods and results*, Aerosol Science 39, 827–849, 2008

## 6.1 Documenti su AERMOD:

Tutti i documenti elencati sono disponibili alla pagina

[https://www3.epa.gov/scram001/dispersion\\_prefrec.htm](https://www3.epa.gov/scram001/dispersion_prefrec.htm):

AERMET User's guide

AERSURFACE User's guide

AERMOD User's guide

AERMOD User's guide addendum

AERMOD: description of model formulation

AERMOD: Latest Features and Evaluation Results.

ENEA  
Servizio Promozione e Comunicazione  
[www.enea.it](http://www.enea.it)

Stampa: Laboratorio Tecnografico ENEA - C.R. Frascati  
giugno 2017