



APAT

Dipartimento difesa del suolo

Roma, 28 GIU. 2006

Prot. 18744

Dr. Gianfranco Mascazzini
Direttore Generale
Direzione Generale per la Qualità
della Vita
Ministero dell' Ambiente e
Tutela del Territorio
Via Cristoforo Colombo, n. 44
00147 Roma

Fax: 06/57225193

Oggetto: Protocollo Operativo per la determinazione dei valori di fondo di metalli/metalloidi nei suoli dei siti d'interesse nazionale.

Con la presente si trasmette il documento di cui in oggetto, che risponde ad una specifica richiesta di codesto Ministero formalizzata nel verbale della Conferenza di servizi decisoria del 13.03.2006, relativa al sito di interesse nazionale di Brindisi.

Nell'occasione, è stato affidato ad ISS e APAT l'incarico di elaborare la metodologia per affrontare il problema del "fondo naturale" di arsenico nell'area di Brindisi.

Tale documento elaborato congiuntamente con ISS, può essere considerato di valenza generale e pertanto applicabile non solo al sito di Brindisi, ma anche in tutti gli altri siti di bonifica di interesse nazionale ove sussista la problematica di definire un fondo naturale/antropizzato per presenza di metalli e/o metalloidi.

Distinti saluti.

SERVIZIO TECNOLOGIE DEL SITO
E SITI CONTAMINATI
Il Responsabile
Ing. Luciano Bonci

PO/SUO-TEC 275



**Protocollo Operativo per la determinazione dei valori di fondo
di metalli/metalloidi nei suoli dei siti d'interesse nazionale**

*Agenzia per la Protezione dell'Ambiente
e per i Servizi Tecnici*

Istituto Superiore di Sanità

Giugno 2006 (Revisione 0)

SOMMARIO

1	PREMESSA	1
2	DEFINIZIONI.....	1
3	INTRODUZIONE	2
4	CRITERI GENERALI	2
5	ASSETTO GEOLOGICO E GEOCHIMICO DELL'AREA	3
6	COSTITUZIONE DEL SET DI DATI.....	3
6.1	INDIVIDUAZIONE DELLE AREE RAPPRESENTATIVE	4
6.2	UBICAZIONE DEI PUNTI DI PRELIEVO	4
6.3	NUMERO DI CAMPIONI	5
6.4	ANALISI DI LABORATORIO	5
6.5	RACCOLTA E ANALISI DEI DATI SITO SPECIFICI	6
7	ANALISI STATISTICA DEI DATI.....	7
7.1	ANALISI PRELIMINARE DEL SET DI DATI.....	7
7.2	DEFINIZIONE DELLA DISTRIBUZIONE DEI DATI.....	7
7.3	RAPPRESENTAZIONE DELLA DISTRIBUZIONE DI DATI	8
8	DETERMINAZIONE DEI VALORI DI FONDO.....	9
	BIBLIOGRAFIA.....	11

1 PREMESSA

Il presente documento ha come obiettivi:

- definire i criteri per l'accertamento delle concentrazioni di fondo di metalli/metalloidi nei terreni in corrispondenza dei siti inquinati definiti dal Programma nazionale delle bonifiche;
- indicare le possibili opzioni per il confronto tra le concentrazioni di fondo e le concentrazioni risultanti dalla caratterizzazione di un sito specifico.

Il documento è stato redatto da un gruppo di lavoro composto dai rappresentanti dell'Agenzia Nazionale per la Protezione dell'Ambiente e per i Servizi Tecnici (APAT) e dell'Istituto Superiore di Sanità (ISS), tenendo conto delle proposte del Centro Tematico Nazionale Territorio e Suolo, cui hanno partecipato i rappresentanti di alcune Agenzie Regionali per l'Ambiente (ARPA).

Nella redazione del Protocollo Operativo ci si è basati sulle indicazioni contenute nelle principali guide tecniche prodotte a livello nazionale ed internazionale, riguardanti l'elaborazione di criteri per la determinazione delle concentrazioni di fondo nelle matrici ambientali.

2 DEFINIZIONI

Si ritiene utile riportare alcune definizioni riprese dalla documentazione esaminata perché possono contribuire a chiarire cosa si debba intendere per concentrazioni di fondo nelle matrici ambientali.

Valore di fondo (ISO 19258): concentrazione di una sostanza nel suolo derivante dai processi geologici e pedologici comprendente anche l'apporto di sorgenti diffuse.

Contenuto naturale pedo-geochimico (ISO 19258): concentrazione di sostanze nei suoli, risultante da processi naturali geologici e pedologici, senza alcuna interferenza di origine antropica.

Contenuto antropizzato (ISO 19258): concentrazione di una sostanza nei suoli derivata sia dal contenuto naturale pedo-geochimico sia della moderata immissione diffusa nel suolo.

Caratteristiche statistiche (ISO 19258): parametro statistico scelto per rappresentare la distribuzione delle concentrazioni; ad esempio il 90° percentile.

Valori di fondo naturale (ISO 19258): caratteristiche statistiche del contenuto naturale pedo-geochimico di una sostanza nei suoli.

Valori di fondo antropizzato "fondo usuale" (ISO 19258): caratteristiche statistiche del contenuto antropizzato di una sostanza nei suoli.

Immissione da sorgente diffusa (ISO 19258): l'immissione di una sostanza emessa da una sorgente mobile, da una sorgente estesa o da più sorgenti.

Inquinamento diffuso (DLgs 152/06): contaminazione o alterazioni chimiche, fisiche o biologiche delle matrici ambientali determinate da fonti diffuse e non imputabili ad una singola origine.

Inquinamento diffuso (D.M. 471/99): contaminazione o alterazioni chimiche, fisiche o biologiche del suolo o del sottosuolo o delle acque superficiali o delle acque sotterranee imputabili alla collettività indifferenziata e determinate da fonti diffuse.

Metalloidi: elementi aventi caratteristiche fisiche e chimiche intermedie tra quelle dei metalli e quelle dei non metalli; i metalloidi noti sono boro (B), silicio (Si), germanio (Ge), arsenico (As), antimonio (Sb), tellurio (Te) e polonio (Po).

Valore di fondo (US EPA, 1995): concentrazione di composti inorganici nei suoli o nei sedimenti situati in prossimità di siti inquinati ma che non sono influenzati dalle attività svolte nel sito o ad esse ricollegabili.

In questo documento, con il termine valore di fondo si fa riferimento alla distribuzione delle concentrazioni di metalli e metalloidi la cui presenza nei terreni, non è riconducibile ad alcuna sorgente puntuale e/o specifica attiva, nel presente o in passato, sull'area di interesse.

3 INTRODUZIONE

I criteri elaborati riguardano le indagini relative alla matrice suolo e sottosuolo, in cui la variabilità spaziale di alcuni elementi, data l'eterogeneità del materiale costitutivo, può essere particolarmente accentuata.

Il suolo si origina dall'alterazione, disgregazione e trasformazione della roccia madre, le cui caratteristiche ne determinano la composizione mineralogica iniziale. I caratteri fondamentali di un suolo sono pertanto determinati sia dalla composizione della roccia madre sia dai processi pedogenetici, cioè l'insieme delle interazioni tra processi chimici, fisici e biologici, che avvengono al suo interno. Un suolo poco evoluto ha caratteristiche molto simili, in termine di concentrazione e di rapporto tra gli elementi, a quelle della roccia madre. D'altra parte, invece, suoli evoluti possono presentare una composizione che differisce anche in maniera rilevante da quella del substrato originario.

Oltre agli elementi sopra riportati, la composizione chimica di un suolo può essere condizionata dalle attività antropiche, sia attraverso l'immissione diretta di composti (contaminazione da fonti puntuali e/o diffuse), sia mediante la modificazione dei parametri fisici, chimici e biologici alla base dei processi pedogenetici.

Si può quindi sostenere che, ovunque vi sia attività umana, la composizione di un suolo è data dall'insieme di una frazione pedo-geochimica naturale e di una frazione antropogenica, il cui peso nella composizione finale è difficilmente individuabile.

4 CRITERI GENERALI

Per la definizione dei criteri proposti nel presente documento si è principalmente fatto riferimento ai documenti redatti da organismi, nazionali e internazionali, di riconosciuta esperienza, di seguito riportati:

- ISO - International Organization for Standardization: Soil quality – Guidance on the determination of background values. ISO19258 - ISO TC 190/SC 7.
- FOREGS - Forum of the European Geological Surveys Directors: Geochemical mapping. Field manual
- U.S. Environmental Protection Agency (2002): Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites.
- Provincia di Milano: Linee guida per la determinazione dei valori del fondo naturale nell'ambito della bonifica dei siti contaminati.

La procedura proposta prevede i seguenti passaggi:

- raccolta e analisi dei dati esistenti;
- costituzione del set di dati;
- elaborazione statistica dei dati;
- determinazione del fondo.

Nei successivi capitoli sono descritti i principali passaggi previsti in ciascuna fase della procedura, mentre nelle appendici sono riportati alcuni approfondimenti sulle applicazioni statistiche richiamate nel testo.

5 ASSETTO GEOLOGICO E GEOCHIMICO DELL'AREA

La prima fase dello studio deve portare alla definizione dell'assetto geologico dell'area di interesse e ad una prima ricostruzione della composizione geochimica dei terreni presenti. Tale studio può essere condotto a partire dalla ricognizione delle diverse fonti di informazioni, di cui, in Tabella 1, si riporta un elenco non esaustivo:

DOCUMENTO	INFORMAZIONI
Carta geologica	Litologia
Carta geomorfologica	Aree d'accumulo e di erosione
Carta dei suoli	Caratteristiche e variabilità del suolo
Carta geochimica	Composizione geochimica dei terreni
Reti di monitoraggio	Valori di concentrazione dei principali composti nei terreni
Archivio geochimico nazionale	Principali informazioni sui suoli, sui sedimenti fluviali attivi e sulle acque del territorio italiano
FOREGS Geochemical Baseline Mapping Programme	Caratteristiche geochimiche ad ampia scala
Cartografia tecnica Regionale Piani Regolatori	Strutture presenti sul territorio Destinazione d'uso delle aree

Tabella 1: Esempio di documentazione utile per l'inquadramento generale

L'acquisizione e l'analisi dei documenti esistenti potrà fornire, anche attraverso la redazione di carte e tabelle, le indicazioni preliminari sulle concentrazioni tipiche dei parametri di interesse nell'area in studio e permetterà, inoltre, di individuare le aree aventi caratteristiche omogenee con l'area in studio, in modo da delimitare l'area rappresentativa, come meglio descritto nel paragrafo 6.1.

6 COSTITUZIONE DEL SET DI DATI

Nei successivi paragrafi sono descritti i criteri da seguire per la composizione del set di dati da utilizzare nella determinazione dei valori di fondo. I requisiti di rappresentatività, omogeneità e qualità proposti, potranno essere utilizzati sia per la validazione di dati acquisiti nel corso d'indagini già effettuate sia per la programmazione di campagne di indagine appositamente predisposte.

6.1 Individuazione delle aree rappresentative

L'area rappresentativa è intesa come una porzione di territorio, con le caratteristiche indicate nel seguito, nella quale sono raccolti i campioni le cui analisi di laboratorio forniscono il set di dati.

Per l'individuazione delle aree rappresentative si dovrà fare riferimento alla ricostruzione effettuata nella precedente fase dello studio, come descritto al capitolo 5.

I dati necessari alla determinazione del fondo, finalizzato al confronto con dati specifici di un sito potenzialmente contaminato, dovranno provenire da aree che presentino le seguenti caratteristiche:

1. siano geologicamente e geomorfologicamente confrontabili con l'area oggetto di indagine;
2. la distribuzione delle concentrazioni dei metalli/metalloidi non sia riconducibile ad alcuna sorgente puntuale e/o specifica attiva nel presente o nel passato.

Il primo criterio di selezione delle aree d'indagine prevede la compatibilità geologica e geomorfologica, per ottemperare all'esigenza di individuare un'area che presenti le medesime caratteristiche dell'area inquinata in termini di composizione chimica "naturale".

Le condizioni richieste dal secondo punto portano all'individuazione di aree di campionamento in cui la destinazione d'uso sia tale da escludere, in prima analisi, la presenza di sorgenti puntuali di contaminazione. Le aree in cui effettuare il prelievo dei campioni saranno quindi, in via prioritaria, quelle destinate a verde pubblico, i parchi, le zone protette, le aree residenziali, etc.

Appare evidente la difficoltà di disporre di aree, ubicate in prossimità di siti industriali e quindi aventi caratteristiche geologiche confrontabili, che non risentano però delle attività svolte presso il sito o di altre attività antropiche (ad esempio le aree agricole). Eventuali difformità dai requisiti sopra enunciati dovranno essere opportunamente documentate in modo da poter effettuare, in fase di elaborazione dei dati, una valutazione corretta, anche mediante l'applicazione di opportuni strumenti statistici.

6.2 Ubicazione dei punti di prelievo

Le caratteristiche geochemiche di un terreno sono funzione di numerosi fattori fra cui la natura della roccia madre, le caratteristiche climatiche, il tipo e il grado di attività antropica (es. pratiche agricole, presenza di siti industriali, traffico) che insiste o ha insistito sul sito, o nei dintorni di esso.

A differenza delle fasi di caratterizzazione di un sito potenzialmente contaminato, in cui l'obiettivo è quello di individuare sorgenti e percorsi di migrazione degli inquinanti (definizione del modello concettuale), nella progettazione di un piano di campionamento per la determinazione dei valori di fondo è utile ricorrere ad una strategia mirata all'individuazione d'aree su cui non insistono sorgenti di contaminazioni puntuali.

Le strategie di campionamento più comunemente adottate comprendono:

- campionamento sistematico o a griglia, in cui i punti ricadono sui vertici o all'interno delle celle di un reticolo immaginario a maglia quadrata o triangolare, risultando quindi equispaziati. L'intervallo tra i punti della griglia dipende dalla risoluzione desiderata e dalle dimensioni dell'area investigata;

- campionamento casuale, in cui i punti sono disposti liberamente nell'area da investigare;
- campionamento sistematico-casuale, in cui all'interno di una maglia quadrata o triangolare le singole celle contengono ciascuna un punto ma in posizione variabile da cella a cella;
- campionamento stratificato, in cui i punti vengono distribuiti casualmente o sistematicamente all'interno di sub-aree individuate col criterio della maggiore omogeneità rispetto ad un parametro prescelto (ad esempio, stessa litologia o stesso grado di umidità o stessa profondità, ecc.).

La scelta della strategia è generalmente determinata dal grado di conoscenze preesistenti sulla natura e sull'uso del complesso suolo-sottosuolo; quando si hanno scarse conoscenze l'approccio casuale e quello sistematico sono spesso più indicati. Laddove si è in possesso di informazioni specifiche circa la distribuzione spaziale di uno o più parametri "discriminanti" (es. uso del suolo, tipo di suolo) si può preferire il campionamento stratificato. È tuttavia possibile costruire una strategia ibrida mescolando alcuni aspetti dell'approccio sistematico, casuale e stratificato.

In ogni stazione di campionamento saranno prelevati 3 campioni: il top soil, un campione rappresentativo dello strato superficiale 0-1 m, e uno rappresentativo dello terreno profondo (compreso tra il primo m e il tetto del terreno saturo). L'analisi per la determinazione dei valori di fondo sarà condotta per comparti omogenei in termini di litologia e classi di profondità. Non saranno prelevati campioni costituiti da materiale di riporto.

6.3 Numero di campioni

I documenti consultati inerenti la stima dei valori del fondo, indicano che il numero minimo di campioni necessari a garantire la significatività statistica del dato è compreso fra 10 e 30.

Con più rigore, il numero di campioni su cui basare l'analisi per derivare il valore rappresentativo del fondo dipende dal tipo di distribuzione dei valori e dal livello di accettabilità dell'errore definito a priori dal decisore.

Sulla base dei criteri statistici riportati in Appendice 1, per gli scopi del presente documento, il numero minimo di campioni necessari per la determinazione della distribuzione di concentrazioni di fondo è posto pari a 30. Tale numero deve essere rispettato per ogni strato omogeneo, come definito nel paragrafo precedente.

Come ovvio un numero di campioni maggiore aumenta la significatività del dato, a patto che siano comunque rispettati i criteri descritti nel paragrafo 6.1.

6.4 Analisi di laboratorio

La normativa prevede che, nel caso di indagini di caratterizzazione di siti potenzialmente contaminati, le determinazioni analitiche siano effettuate con metodi di analisi ufficiali riconosciuti a livello nazionale e/o internazionale. Anche per i campioni prelevati per la determinazione dei valori di fondo devono pertanto essere utilizzate metodiche rispondenti a tali requisiti e in regime di buone pratiche di laboratorio e di qualità.

Nell'ambito delle indagini per l'acquisizione di dati utili ai fini della determinazione dei valori di fondo, per valutare l'effettiva biodisponibilità di inquinanti metallici, oltre alle analisi per la determinazione delle concentrazioni dei parametri di interesse è opportuno

prevedere la determinazione di ulteriori parametri caratteristici del terreno che influenzano la mobilità e le reazioni chimiche del metallo con i costituenti della matrice suolo.

Le determinazioni analitiche richieste ai fini della valutazione della biodisponibilità, sono le seguenti: *tessitura, peso specifico, pH, potenziale redox, carbonio organico, capacità scambio cationico, contenuto di carbonato, contenuto totale di ferro e alluminio*

Per ogni metallo/metalloide saranno determinati i parametri specifici riportati nella Tabella 2.

FORMA	PROCEDURA ANALITICA
Totale	Spettrofotometria XRF, dissoluzione completa in HF
Pseudo totale	Estraibili in Acqua regia
Disponibile per le piante	D.M. 13 settembre 1999 "Metodi ufficiali di analisi chimica sul suolo"
Solubilità in acqua	Test eluizione norma UNI 10802

Tabella 2: Determinazioni analitiche

6.5 Raccolta e analisi dei dati sito specifici

Per la determinazione delle concentrazioni di fondo nei siti contaminati d'interesse nazionale, si può valutare l'opportunità di utilizzare i dati acquisiti nel corso dalle attività di caratterizzazione svolte dai soggetti obbligati. In questo modo si potranno evitare inutili duplicazioni di informazioni con un conseguente contenimento degli oneri economici. L'attività comporterà:

1. verifica dei soggetti presenti all'interno dell'area perimetrata;
2. verifica dello stato di attuazione delle attività di caratterizzazione;
3. raccolta dei dati e costituzione di un archivio informatizzato.

In prima analisi si ritiene che possano essere utilizzate anche le concentrazioni di campioni prelevati all'interno dell'area perimetrata, purché esterni alle aree occupate da stabilimenti industriali, discariche o, in generale, aree in cui siano ubicate sorgenti di contaminazione attive nel presente o nel passato.

Appare evidente che l'utilizzo dei dati esistenti sarà subordinato alla verifica della completezza e dell'adeguatezza degli stessi, e alla rispondenza ai requisiti descritti nei precedenti paragrafi.

La necessità di procedere al prelievo di ulteriori campioni di terreno, rispetto a quelli già disponibili, può derivare da una o più delle seguenti motivazioni:

- il numero di campioni è insufficiente per una corretta analisi statistica;
- i campioni sono stati localizzati in aree prossime a sorgenti di contaminazione ben individuabili;
- non è nota la geologia e la geomorfologia dell'area di prelievo;
- le modalità di campionamento non sono uniformi;
- le metodiche analitiche utilizzate non permettono il confronto dei diversi set di dati;
- la qualità dei dati non è accertata;

- è necessario integrare i dati analitici con ulteriori parametri (es caratteristiche chimico/fisiche).

I dati di concentrazione dei parametri di interesse dovranno essere raggruppati per strati omogenei (top-soil, suolo superficiale, suolo profondo) distinguendo tra i vari litotipi presenti in modo da poter applicare la procedura statistica per il calcolo del valore di fondo a ciascuno strato omogeneo.

7 ANALISI STATISTICA DEI DATI

La concentrazione di un dato elemento nel suolo di un'area, può essere considerato come una popolazione statistica di dati. L'obiettivo dell'elaborazione dei dati è rappresentare e caratterizzare la suddetta popolazione utilizzando un campione statistico di n. valori individuali.

Una volta definito il set di dati corrispondente al campione rappresentativo, si procederà alla determinazione del fondo mediante l'applicazione di test statistici secondo i passaggi indicati nel seguito:

- analisi preliminare del set di dati (identificazione e trattamento di outliers e non-detect);
- definizione della distribuzione dei dati (test statistici);
- rappresentazione della distribuzione dei dati (descrittori numerici e grafici);
- definizione dei valori di concentrazione rappresentativi del fondo.

7.1 Analisi preliminare del set di dati

La valutazione dei dati, finalizzata a stabilire l'applicabilità di criteri statistici sui valori di concentrazione analiticamente determinati, prevede i seguenti passaggi:

- nel caso di utilizzo di dati acquisiti nel corso di attività di caratterizzazione pregresse, verificare che il campionamento sia uniformemente distribuito su tutta l'area di riferimento. Un campionamento più concentrato in alcune porzioni del sito può comportare una stima falsata delle concentrazioni di fondo;
- identificare gli outlier e distinguere i "veri outlier" dai "falsi outlier". I "veri outlier" possono derivare da errori di trascrizione, di codifica dei dati o da una qualsiasi inefficienza degli strumenti del sistema di rilevazione dei dati. I "falsi outlier" sono valori estremi reali. E' dunque necessario identificare e differenziare i tipi di outlier, in modo da rimuovere i primi e gestire i secondi;
- identificare i non-detect. Si ritiene opportuno porre, in ogni caso e quindi in corrispondenza a qualsiasi distribuzione dell'insieme dei dati, i non-detect pari a metà del corrispondente detection limit (n.d. = d.l.).

7.2 Definizione della distribuzione dei dati

Lo scopo è quello di individuare la distribuzione di probabilità che approssimi meglio l'insieme dei dati disponibili. Quando si ha a che fare con dati ambientali (in particolare, concentrazioni di specie chimiche nei comparti ambientali suolo, acqua, aria), le distribuzioni di probabilità che più comunemente le rappresentano sono:

- distribuzione gaussiana o normale;
- distribuzione lognormale;

- distribuzione gamma;
- distribuzione non parametrica.

L'individuazione del tipo di distribuzione che meglio approssima il campione di dati serve a definire i descrittori statistici più appropriati per stimare il valore del fondo. Dal tipo di distribuzione dipendono inoltre i test statistici da applicare per il confronto tra due set di dati (ad esempio il set relativo ai valori di fondo e quello relativo alle concentrazioni riscontrate in uno specifico sito).

Le caratteristiche delle distribuzioni suddette e i test da applicare per la selezione delle stesse sono descritti nel dettaglio in Appendice 1.

7.3 *Rappresentazione della distribuzione di dati*

Il tipo di distribuzione del set di dati può essere rappresentato con descrittori numerici o con metodi grafici, differenti in funzione del tipo di distribuzione. I principali descrittori numerici di una distribuzione di dati sono riportati nella Tabella 3.

DESCRITTORE STATISTICO	DEFINIZIONE
Massimo e Minimo	Rappresentano il valore massimo e quello minimo nell'insieme dei dati.
Media aritmetica	È data dalla somma di tutti i valori divisa per il numero dei casi.
Mediana	Rappresenta il valore centrale di una distribuzione ordinata in senso crescente.
Percentile	Sono quei valori che dividono la distribuzione in cento parti, in modo che, ad esempio, il 25° percentile (o primo quartile) sia quel valore che supera il 25% della distribuzione ed è superato dal 75%, il 50° percentile (o secondo quartile) sia il valore che divide la distribuzione in due parti uguali (e quindi il secondo quartile coincide con la mediana), il 75° percentile (terzo quartile) sia quel valore superato dal 25% della distribuzione.
Range	Rappresenta la differenza fra il valore massimo e il minimo.
Range interquartile	Rappresenta la differenza tra il 75° e il 25° percentile.
Varianza	Rappresenta la distanza di un valore dalla media aritmetica della distribuzione.
Deviazione Standard o scarto tipo o scarto quadratico medio	Rappresenta la misura della dispersione di n misure in un set di dati. È la radice quadrata della media degli scarti, al quadrato.
Coefficiente di skewness	Fornisce una stima della asimmetria della forma di distribuzione dei dati
Coefficiente di curtosi	Fornisce una stima della acutezza della curva di distribuzione dei dati
Coefficiente di variazione	E' un indice che permette di analizzare la dispersione dei valori attorno alla media indipendentemente dall'unità di misura, fornendo un'indicazione sulla variabilità delle osservazioni rilevate.

Tabella 3 Principali descrittori statistici

Le modalità di rappresentazione grafica del set di dati, di uso più comune, sono:

- istogrammi;
- box-plot;
- curve cumulative di frequenza.

Ciascuna di queste può fornire indicazioni sui principali parametri caratteristici di una distribuzione di valori; in particolare, le curve cumulative di frequenza possono essere di supporto per la determinazione dei valori di fondo, come descritto in maggiore dettaglio nella Appendice 1.

Lo studio per la determinazione dei valori di fondo dovrà presentare il numero massimo di descrittori, numerici e grafici, in modo da fornire un quadro completo della distribuzione dei dati che possa essere di supporto nella selezione del valore da utilizzare nel processo decisionale.

8 DETERMINAZIONE DEI VALORI DI FONDO

Per la definizione dei descrittori da utilizzare per la rappresentazione del valore di fondo, sono state prese in rassegna le metodologie proposte a livello nazionale e internazionale, di cui si riporta una sintesi nel seguito.

ISO 19258, 2005. La norma non indica un unico descrittore per determinare il valore di fondo, anche se raccomanda l'utilizzo dei percentili per rappresentare la distribuzione delle concentrazioni (10°, 25°, 50°, 75° e 90° percentile).

FOREGS-Geochemical Atlas of Europe, 1998. Anche in questo documento non viene riportato un unico valore rappresentativo del fondo geochimico. Le conclusioni dello studio riportano, per ciascun parametro analizzato, i valori Minimo, Massimo, Mediana, Media, Deviazione Standard, 90 Percentile.

USEPA, 2002. Il documento non fornisce criteri per la determinazione di valori di fondo di un'area ma rappresenta una guida per effettuare il confronto tra i valori di background e quelli dei siti prioritari nazionali, in funzione di obiettivi specifici. Con background si intende un'area che non ha subito impatti (rilasci) da un sito contaminato.

Provincia di Milano, 2003. La determinazione dei valori di fondo segue due linee, in funzione della dimensione del sito. Per siti di estensione inferiore a 1.000 m², i campioni che presentano concentrazioni superiori alla somma del valore medio + la deviazione standard, siano anch'essi prelevati in aree teoricamente rappresentative del fondo, sono considerati anomali. Per i siti medio-grandi (di estensione maggiore di 1.000 m²), sono proposti dei test statistici (parametrici e non parametrici) per confrontare la distribuzione dei dati del sito con la distribuzione dei dati del fondo naturale. Lo studio si completa con l'analisi della curva di distribuzione cumulativa di frequenza per determinare il valore di fondo naturale che si assume posto in corrispondenza di punti di discontinuità (gap, variazione di pendenza).

Come si evince dalla sintesi sopra riportata, la rassegna dei documenti elaborati a livello nazionale e internazionale, ha evidenziato l'assenza di un criterio condiviso sulle modalità di determinazione delle concentrazioni di fondo in un'area.

Con lo scopo di fornire una procedura basata su presupposti scientifici sufficientemente rigorosi, che permetta al momento stesso un'applicazione ai siti di interesse nazionale, si suggerisce il seguente approccio:

- verifica dell'adeguatezza del numero di dati utili disponibili (minimo 30);
- applicazione di un test statistico per la verifica del tipo di distribuzione;

- calcolo e presentazione dei descrittori statistici;
- costruzione della curva cumulativa di frequenza e individuazione di eventuali punti di discontinuità;
- selezione del valore di fondo corrispondente al 95° percentile.

Il processo decisionale potrà essere completato utilizzando il metodo comparativo. Tale metodo prevede il confronto tra la distribuzione dei dati rappresentativi del fondo con quella dei dati sito specifici, con lo scopo di verificare se le due popolazioni da cui sono originati i campioni siano statisticamente uguali o una mostri dei descrittori significativamente più alti dell'altra. Il criterio comparativo può basarsi su metodi statistici più o meno complessi e raffinati di cui, in Appendice 1, si riporta una descrizione dettagliata.

L'applicazione di test statistici si concretizza con l'accettazione o il rifiuto di ipotesi statistiche circa l'uguaglianza o meno dei set di campioni in esame. Con questi test si ha anche il controllo sulla probabilità di commettere errori decisionali. Il metodo comparativo, nelle diverse forme, richiede un certo grado di conoscenza del sito e di un consistente set di dati rappresentativi del fondo ovvero dell'area di riferimento (non interessata da contaminazione).

BIBLIOGRAFIA

APAT (2005): Criteri metodologici per l'applicazione dell'analisi assoluta di rischio ai siti contaminati, rev0, giugno 2005.

APAT, CTN-TES (2005): Metodologia per la determinazione del fondo naturale. TES-T-RAP-03-17. ver01

ARPAL (200?): Progetto Regionale per la determinazione del fondo naturale in quattro bacini idrografici della Liguria con particolare riferimento alle litologie ultrabasiche.

ARPAV, Comune di Venezia, Provincia di Venezia (2002): Determinazione del livello di fondo di metalli pesanti nei suoli dell'entroterra veneziano".

Battelle Memorial Institute Earth Tech, Inc NewFields, Inc. (2002): Guidance for environmental background analysis Volume I: Soil. Naval facilities engineering command Washington, DC 20374-5065.

Beretta G.P. (2001): Gestione dei dati analitici in fase di caratterizzazione, bonifica e certificazione dei siti contaminati. Atti della giornata di studio Bonifica di siti contaminati. Quaderni Direzione Centrale Ambiente Provincia di Milano, pp 38 – 77.

De Vivo B., Lima A., Albanese S., Cicchella D. 2003: Atlante geochimico-ambientale della Regione Campania – Dipartimento di Geofisica e Vulcanologia Università degli Studi Federico II, Napoli.

International Organization for Standardization (2005): Soil quality – Guidance on the determination of background values. ISO19258 - ISO TC 190/SC 7.

Provincia di Milano, Università degli Studi di Milano Dipartimento di Scienze della Terra 'A. Desio' (2003): Linee guida per la determinazione dei valori del fondo naturale nell'ambito della bonifica dei siti contaminati.

Salminen R., Tarvainen T., Demetriades A., Duris M., Fordyce F. M., Gregorauskiene V., Kahelin H., Kivisilla J., Klaver G., Klein H., Larson J.O., Lis J., Locutura J., Marsina K., Mjartanova H., Mouvet C., O Connor P., Odor L., Ottonello G., Paukola T., Plant J. A., Reimann C., Schermann O., Siewers U., Steenfelt A., Van der Sluys J., de Vivo B., Williams L. (1998) - FOREGS geochemical mapping. Field manual. Geologian tutkimuskeskus, Opas - Geological Survey of Finland, Guide 47. 36 pages, 15 figures, and 1 appendix.

U.S. Environmental Protection Agency (1996) – “The Data Quality Evaluation Statistical Toolbox (Data QUEST) Software”, EPA QA/G-9D. Office of Research and Development, Washington D.C.

U.S. Environmental Protection Agency (2000a): Data Quality Objectives Process for Hazardous Waste Site Investigations. EPAQA/G-4HW, Final. Office of Environmental Information, U.S. Environmental Protection Agency, Washington D.C.

U.S. Environmental Protection Agency (2000b): Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA QA/G-9, QA00 Update. Office of Environmental Information, U.S. Environmental Protection Agency, Washington D.C.

U.S. Environmental Protection Agency (2001) “RAGS: volume 3 PART A-Process for Conducting Probabilistic Risk Assessment Appendix A”.

U.S. Environmental Protection Agency (2002): Calculating Upper Confidence Limits for exposures point concentrations at hazardous waste sites. Oswer 9285.6-10. Office of Emergency and Remedial Response U.S.Environmental Protection Agency, Washington D.C.

U.S. Environmental Protection Agency (2002): Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites. Office of Emergency and Remedial Response U.S. Environmental Protection Agency Washington, DC 20460.

U.S. Environmental Protection Agency (2006): Guidance on Systematic Planning Using the Data Quality Objectives Process EPA QA/G-4 Final. Office of Environmental Information, U.S. Environmental Protection Agency, Washington D.C.

APPENDICE 1

ANALISI STATISTICA

SOMMARIO

1	INTRODUZIONE	1
2	TEST STATISTICI	1
3	NUMEROSITÀ CAMPIONARIA	2
4	GLI “OUTLIER”	3
5	I “NON-DETECT”	5
6	DISTRIBUZIONE DEI DATI	7
7	DESCRITTORI STATISTICI	9
8	DEFINIZIONE DEL TIPO DI DISTRIBUZIONE	13
9	CONFRONTO FRA I VALORI DI FONDO E I VALORI SITO SPECIFICI	15

Indice delle Tabelle

Tabella 1: Confidenza e Potenza del Test	2
Tabella 2: Criteri di selezione del test per la identificazione degli outlier.....	4
Tabella 3: Test per la selezione del tipo di distribuzione.....	15
Tabella 4: Tipologia dei Test Statistici	15
Tabella 5: Applicabilità dei Test Statistici	16
Tabella 6 Esempio di calcolo del Rango.....	20
Tabella 7 Valori critici per la distribuzione t di Student.....	21
Tabella 8 Valori critici per il Wilcoxon Rank Sum Test.....	22

Indice delle Figure

Figura 1 Esempio di distribuzione normale	8
Figura 2 Esempio di distribuzione lognormale.	9
Figura 3: Parametri statistici rappresentati nel box plot.....	12
Figura 4 Curva cumulativa di frequenza.....	13

1 INTRODUZIONE

La presente Appendice contiene indicazioni e riferimenti bibliografici per l'applicazione dei criteri finalizzati alla determinazione dei valori di fondo di metalli e metalloidi nei terreni dei siti d'interesse nazionale.

Nel seguito sono descritti i principali elementi che è necessario prendere in considerazione per stabilire l'applicabilità di criteri statistici atti ad individuare la distribuzione dei valori di fondo dall'insieme di dati a disposizione.

Si presuppone che i dati analitici a disposizione siano stati già validati, ossia sia stata verificata la loro attendibilità.

Nella redazione della presente appendice si è fatto riferimento ai contenuti dell'APPENDICE H al documento "Criteri metodologici per l'applicazione dell'analisi assoluta di rischio ai siti contaminati" rev0, disponibile sul sito dell'APAT, www.apat.it (APAT, giugno 2005).

2 TEST STATISTICI

Il test statistico può essere visto come un mezzo per verificare in maniera quantitativa la validità di un'ipotesi. In statistica, l'ipotesi da verificare si chiama ipotesi nulla e si indica con H_0 , mentre con H_a si indica l'ipotesi alternativa.

Come esempio, nel caso in cui fosse necessario verificare se un set di dati ha una distribuzione normale, è possibile prevedere le seguenti ipotesi:

- H_0 : il set di dati presenta una distribuzione normale
- H_a : il set di dati non presenta una distribuzione normale

Se, applicando il test, risultasse scartata l'ipotesi H_0 in favore di H_a , si potrebbe concludere che il set di dati, non supportando l'ipotesi nulla, deve derivare da un tipo di distribuzione non normale.

Occorre però tener presente che l'applicazione di un test statistico comporta sempre un rischio di errore. Nella pratica statistica si individuano due tipi di errori

- errore di primo tipo; è quello che porta a rifiutare H_0 quando è vera;
- errore di secondo tipo; è quello che porta ad accettare H_0 quando è falsa.

Con la lettera α si indica la probabilità di commettere un errore di primo tipo. Con $100(1-\alpha)\%$ si indica il livello di confidenza del test. Se il test non scarta H_0 , ovvero conferma l'ipotesi nulla, può significare che le informazioni del set di dati non sono sufficienti per scartare H_0 con quel livello di confidenza. Ad esempio, fissato $\alpha = 0,05(5\%)$, con i dati campionari si esegue il test e si valuta se il suo valore cade nella regione di rifiuto e nella regione di accettazione. Se, ad esempio, cade nella regione di rifiuto si dice che il test è significativo al 5%.

Con la lettera β si indica la probabilità di commettere un errore del secondo tipo. Con $100(1-\beta)\%$ si indica la potenza del test, cioè la probabilità di scartare correttamente l'ipotesi nulla.

DECISIONE BASATA SU UN CAMPIONE DI DATI	CONDIZIONE DEL SITO ATTUALE	
	H_0 è vero	H_0 non è vero
H_0 non è rigettata	Decisione corretta: $(1-\alpha)$	Errore di tipo II: Falso negativo (β)
H_0 è rigettata	Errore di tipo I: Falso positivo (α)	Decisione corretta: $(1-\beta)$

Tabella 1: Confidenza e Potenza del Test

Nell'ambito di indagini ambientali che comportino l'applicazione di test statistici, i limiti di tolleranza della probabilità di commettere errori (del primo o del secondo tipo) dovrebbero essere specificati in fase di progettazione.

Sono stati elaborati differenti metodi statistici finalizzati alla verifica di ipotesi, la scelta del più appropriato è funzione di una serie di fattori tra i quali il tipo di distribuzione dei dati (come definito al capitolo 6), ha un peso determinante. In funzione del tipo di distribuzione dei dati è possibile distinguere i Metodi Parametrici e i Metodi Non-parametrici:

- Parametrici: si tratta di metodi statistici che si basano su distribuzioni probabilistiche quale, ad esempio, la distribuzione normale. I test statistici parametrici sono utilizzati per la valutazione di ipotesi che riguardano i parametri della distribuzione;
- Non-parametrici: metodi la cui applicazione prescinde dalla conoscenza del tipo di distribuzione della popolazione. In generale i test non parametrici dovrebbero essere preferiti quando i dati non si distribuiscono secondo una normale, o comunque non si è in grado di dimostrarlo, ad esempio per numerosità ridotta.

Per riepilogare i concetti sopra riportati, nel seguito sono descritti i passaggi da seguire nella applicazione di test statistici:

- definizione dell'ipotesi nulla e dell'ipotesi alternativa;
- scelta del test da adottare;
- decisione del livello di significatività;
- esecuzione dei calcoli previsti nel test;
- decisione se accettare o meno la validità dell'ipotesi nulla, in genere confrontando il valore ottenuto nel test con un valore tabulato.

3 NUMEROSITÀ CAMPIONARIA

Per ogni data-set (suolo superficiale e suolo profondo), il numero di dati a disposizione non può essere inferiore ad un valore minimo. L'ampiezza del data set è di particolare importanza soprattutto nei casi in cui si abbia una grande variabilità della distribuzione dei dati.

Affinché l'analisi statistica sia significativa, ovvero il campione di dati sia rappresentativo della popolazione, si fa generalmente riferimento ad un numero minimo di dati, che i diversi testi consultati, riportano variabile tra 10 e 30.

La scelta del numero di campioni rappresentativi è funzione, in primo luogo, dello scopo dell'indagine: posto l'obiettivo dell'indagine di caratterizzazione, viene formulata una ipotesi da verificare mediante l'applicazione di un determinato test statistico; lo stesso test serve ad effettuare la stima del numero di campioni necessari.

Ad esempio, si vuole verificare (ipotesi nulla) se la media delle concentrazioni di un composto in un'area è superiore ad un valore soglia (ad esempio la concentrazione massima ammissibile CMA).

Una delle possibili equazioni utilizzate per determinare il numero minimo di campioni necessari per la verifica della media di una distribuzione nei confronti di un valore soglia di intervento (US EPA, 2006), mediante l'applicazione del t-test, è:

$$n = \frac{s^2 (Z_{1-\alpha} + Z_{1-\beta})^2}{\Delta^2} + 0,5Z_{1-\alpha}^2$$

in cui

n è il numero minimo di campioni

s^2 è la stima della varianza totale vera (σ^2); nel caso in cui (molto frequente) non fosse nota la σ della distribuzione, dovrà essere inserita nella formula una stima di detto parametro (ad esempio il valore della deviazione standard derivata da un campione noto della popolazione di riferimento desunto da dati bibliografici o da precedenti studi).

α è la probabilità accettabile che il test, applicato sul numero n di dati, indichi in maniera errata che la media delle concentrazioni non supera la CMA (in poche parole che un sito contaminato venga definito "pulito")

β è la probabilità accettabile che il test, applicato sul numero n di dati, indichi in maniera errata che la media delle concentrazioni supera la CMA (in poche parole che un sito "pulito" venga definito contaminato)

Δ definito come la minima differenza rilevabile, ovvero, se l'obiettivo dello studio è quello di confrontare la media di concentrazioni di un'area con le CMA, Δ rappresenta la massima differenza tra la media delle concentrazioni e le CMA, che è importante rilevare con una probabilità pari a $1 - \beta$.

Z è il valore, per una distribuzione di dati normale, per il quale la proporzione della distribuzione a sinistra di $Z_{1-\alpha}$ è pari a $1 - \alpha$. I valori di $Z_{1-\alpha}$ sono riportati in numerosi testi di statistica (es. Gilbert 1987, Table A1, pag. 254). Ad esempio se l'ipotesi nulla è quella che le concentrazioni misurate superano le CMA, $Z_{1-\alpha}$ rappresenta una quantificazione della vostra volontà di evitare di considerare contaminato un sito che in realtà è pulito.

Le assunzioni alla base della equazione sopra riportata sono che la distribuzione dei dati sia di tipo normale, i dati siano rappresentativi della popolazione, che i dati non siano correlati nel tempo e nello spazio.

Il livello accettabile di errore viene definito dal decisore e viene espresso tramite il livello di confidenza ($1-\alpha$) e potenza ($1-\beta$).

Per gli scopi del presente documento, il numero minimo di campioni necessari per la determinazione della distribuzione di concentrazioni di fondo è posto pari a 30.

4 GLI "OUTLIER"

Gli outlier sono quei valori di un data set che non sono rappresentativi dell'insieme di dati nel suo complesso. Non sono rappresentativi perché, in genere, sono quantitativamente in numero estremamente ridotto e qualitativamente assumono dei valori molto grandi o molto piccoli

rispetto al resto del data set. In campo ambientale di inquinamento dei suoli, valori di concentrazione molto alti in genere corrispondono ai picchi (hot spot) locali di concentrazione.

Comunque, in generale, tali valori estremi possono costituire dei “veri outlier” o dei “falsi outlier”. I primi possono derivare da errori di trascrizione, di codifica dei dati o da una qualsiasi inefficienza degli strumenti del sistema di rilevazione dei dati. I secondi sono quei valori estremi reali, spesso presenti in questo tipo di indagini soprattutto, come già detto, in campo ambientale. La rimozione dei secondi e/o la mancata rimozione dei primi può condurre ad una visione errata del data set (EPA 2000b, QA/G-9). Infatti è di fondamentale importanza tener conto e quindi non rimuovere i “falsi outlier” dal data set (OSWER 9285.6-10, EPA 2002).

Se il data-set a disposizione è stato già validato si esclude automaticamente la presenza di veri outlier.

L'identificazione degli outlier può essere condotta attraverso le seguenti fasi (EPA 2000b, QA/G-9).

1. Identificazione dei valore estremi che potranno essere potenziali outlier. Questo può essere fatto mediante rappresentazione grafica dell'insieme dei valori rilevati: è possibile così individuare velocemente quei punti che corrispondono a valori più elevati o più ridotti rispetto agli altri. Una volta identificati i potenziali outlier, è necessario procedere a ulteriori indagini, applicando uno dei test statistici disponibili.
2. Applicazione di un opportuno test statistico. Esistono molti test statistici atti a verificare se un outlier statistico, cioè un potenziale vero outlier, sia tale o meno. I principali test statistici utili a tale scopo sono quattro:
 - Extreme value test (Dixon's Test)
 - Discordance Test
 - Rosner's test
 - Walsh's test

Questi, descritti dettagliatamente nel seguito, si differenziano per le dimensioni del data set da considerare, il numero di potenziali outlier da analizzare e la necessità o meno di una distribuzione di tipo normale dei dati raccolti.

In particolare la guida raccomanda l'uso del Rosner's test quando il data set contiene un numero di elementi maggiore di 25; in caso contrario suggerisce quello dell'Extreme Value test. Se si ha un solo valore sospetto outlier il Discordance test può essere sostituito a uno di questi test. Se però i dati non seguono una distribuzione normale si deve considerare un test non parametrico, come il Walsh's test (Tabella 2). Per la descrizione di dettaglio dei test si rimanda al documento (EPA 2000b, QA/G-9)

DIMENSIONE DEL DATA SET	TEST	DISTRIBUZIONE NORMALE
$25 \leq n$	Extreme Value Test	si
$50 \leq n$	Discordance Test	si
$25 \geq n$	Rosner's Test	si
$50 \geq n$	Walsh's Test	no

Tabella 2: Criteri di selezione del test per la identificazione degli outlier

3. Studio scientifico degli outlier identificati per la scelta di trattamento del dato. I test statistici (fase 2) da soli non permettono di stabilire se comprendere o escludere il dato dall'insieme considerato. Le scelte possibili sono:
 - Effettuare ulteriori approfondimenti e indagini al fine di correggere il valore di outlier.
 - Utilizzare il data set comprensivo dei valori di outlier.
 - Escludere l'inserimento di tali valori dal data set. Tale scelta può avvenire solo se è possibile accompagnare i risultati dei test statistici (fase 2) con valide giustificazioni scientifiche.
4. Nel caso di esclusione degli outlier dal data set, conduzione della successiva analisi statistica dei dati sia sull'insieme dei dati comprensivo di outlier, sia su quello rivisto con l'eventuale soppressione degli outlier.
5. Documentazione dell'intero procedimento, con la descrizione di tutti i passaggi e le scelte effettuate.

5 I "NON-DETECT"

Tutte le tecniche analitiche di laboratorio hanno un "Detection Limit"(DL) (limite di rilevazione): i valori cosiddetti "non-detect" (ND) sono quelle concentrazioni realmente o virtualmente pari a zero, o comunque maggiori di zero, ma al di sotto delle possibilità di misurazione del laboratorio. Il DL dipende dalla sensibilità della metodica di estrazioni ed analisi.

Un data set contenente non-detect viene definito in letteratura "censored" a indicare la sua incompletezza, che può essere più o meno influente a seconda del DL del laboratorio che ha condotto il campionamento: per questo motivo è opportuno che il laboratorio alleghi, alla documentazione dello studio, le informazioni sul "Quantitation Limit" (limite di misura) che dipenderà dalla strumentazione di cui si è servito. Il Quantitation Limit può essere definito come il livello più basso al quale una sostanza chimica può essere misurata con precisione, generalmente pari al DL dello strumento moltiplicato per un fattore compreso fra tre e cinque, ma comunque variabile a seconda della sostanza considerata e del tipo di campione (RAGS Part A, EPA 1989).

La presenza di ND in un insieme di dati può influire pesantemente sul calcolo della media, della varianza, sull'orientamento dei dati e su vari altri parametri, pregiudicando quindi il procedimento statistico nel caso in cui questo risulti applicabile nonostante la loro presenza.

I laboratori di analisi riportano questi valori come "non-detect" (ND), oppure li pongono pari a zero o come dati "less-than"(LT) cioè "minori di" una certa quantità, in genere pari proprio al DL, o ancora capita di trovarli indicati come valori pari ad una frazione del DL (in genere a $\frac{1}{2}$ DL). E' comunque preferibile, qualora le tecniche di misurazione lo consentano, che siano riportate le loro misure esatte, benché minime, per non perdere informazioni utili all'analisi dei dati.

Nel seguito è riportato quanto proposto dai testi bibliografici presi quali riferimento.

Il documento (OSWER 9285.6-10, EPA 2002) descrive quattro possibili approcci per la trattazione dei non-detect, finalizzati all'applicazione di analisi statistiche dell'insieme dei dati e alla conseguente individuazione di un valore rappresentativo.

1. Riesame del modello concettuale del sito: da questo riesame potrebbe risultare una distribuzione dei valori di concentrazione tali da permettere l'individuazione di aree a maggior grado di contaminazione e aree a minor grado di contaminazione. In tal caso, il sito oggetto di indagine potrebbe essere suddiviso in sotto-aree, alcune delle quali presenteranno una maggiore e altre una minore concentrazione di non-detect. In tale caso potrebbe risultare necessario raccogliere un maggior numero di campioni per permettere una migliore caratterizzazione del sito.
2. Metodo della sostituzione semplice ("Simple Substitution Methods"): questo metodo prevede l'assegnazione di un valore costante ai dati non-detect. Tale valore potrà essere:
 - pari a zero;
 - pari al Detection Limit;
 - pari alla metà del DL.

L'incertezza associata a questo metodo aumenta all'aumentare del valore del DL e all'aumentare del numero di non-detect. Quindi si consiglia di scegliere, senza un preciso criterio, il valore costante da attribuire tra i tre proposti solo nel caso in cui il numero dei non-detect costituisce al massimo il 15% di tutto il data set (EPA 2000b, QA/G-9).

3. Metodo degli estremi ("Bounding Methods"): Tale metodo propone di calcolare il valore di concentrazione rappresentativo alla sorgente attribuendo, di volta in volta, uno dei valori costanti elencati sopra. Questi metodi forniscono una stima del limite superiore e di quello inferiore, calcolati sulla base dell'intero range di valori dei non-detects possibili (da 0 fino al DL).
4. Metodi della distribuzione ("Distributional Methods"): Si basano sull'ipotesi che la forma della distribuzione dei dati non-detects sia simile a quella delle concentrazioni misurate che superano il DL. Tra questi metodi il più utilizzato è il Metodo di Cohen ("Cohen's Method").

Metodo di Cohen ("Cohen's Method") (EPA 2000b, QA/G-9): è applicabile se i non-detect costituiscono il 15-50% del data set disponibile, se la forma della distribuzione dei dati senza i valori non-detect sia di tipo normale e che il DL sia sempre lo stesso. Questo metodo adatta la media e la deviazione standard per valori al di sotto del DL, basandosi sulla tecnica statistica della stima più probabile della media e della varianza, in modo che sarà possibile applicare i vari test statistici al data set. Nella applicazione di questo metodo i non-detect non si assumono mai pari a zero. Le stime derivanti dai campionamenti sono $X_1, X_2, X_3, \dots, X_n$ di cui i primi m valori rappresentano i dati sopra il DL. Quelli sotto il DL saranno dunque $n-m$.

La scelta del metodo più appropriato dipende dal grado di incompletezza del data set, dalle sue dimensioni e dalla distribuzione più idonea a rappresentare i campioni. Inoltre, sempre il documento (OSWER 9285.6-10, EPA 2002) fornisce cinque raccomandazioni su come trattare un insieme di dati in cui siano presenti dei non-detects:

- I Detection Limits devono sempre essere specificati e i non-detects riportati con il valore osservato se possibile.
- I non-detects non devono mai essere riportati come valori zero senza specifiche giustificazioni.

- Se un'analisi condotta con un Bounding Method rivela che gli effetti quantitativi della presenza di non-detects nel data set è trascurabile non sono necessari ulteriori esami.
- Se si vuole procedere ad ulteriori analisi è consigliabile usare un metodo per una specifica distribuzione.
- Se la quantità dei non-detects nel data set è alta (>75%) oppure se il numero di campioni è basso ($n < 5$) nessun metodo funzionerà bene. In tal caso si può riportare la percentuale di valori al di sotto del DL, ricorrere ancora ad un Bounding Method nel quale i non-detects saranno sostituiti dal DL nel calcolo del fondo, che sarà riportato come un numero probabile considerevolmente maggiore della media reale.

Il documento (RAGS/HHEM, EPA 1989, Volume 1) prevede la possibilità di rianalizzare i campioni cercando di riportare i dati con il loro valore esatto, di usare concentrazioni approssimate (pari al DL o alla metà) o di eliminare alcuni dei non-detect nel caso in cui si abbiano delle informazioni che facciano pensare all'assenza di queste sostanze dal sito. Quest'ultima possibilità deve essere valutata con particolare attenzione, in quanto il Quantitation Limit potrebbe essere maggiore della concentrazione di riferimento di alcuni contaminanti (con la quale deve essere confrontata la concentrazione rappresentativa alla sorgente) e perciò l'eliminazione di alcuni dati può comportare una lacuna nell'analisi di rischio globale del sito. Se la concentrazione di un certo elemento chimico non è stata rilevata in nessun campionamento nel mezzo indagato questa sostanza generalmente viene esclusa dal data set, in modo da avere alla fine dell'analisi dei campioni raccolti un data set comprendente solo quelle sostanze di cui si possiede un valore di concentrazione in almeno un campione per ogni mezzo (aria, acqua, suolo) dell'area di interesse.

Per l'applicazione delle presenti linee guida, seguendo il principio di cautela, si ritiene opportuno porre, in ogni caso e quindi in corrispondenza a qualsiasi distribuzione dell'insieme dei dati, i non-detect pari a metà del corrispondente detection limit (n.d.=d.l.).

6 DISTRIBUZIONE DEI DATI

Quando si ha a che fare con dati ambientali (in particolare, concentrazioni di specie chimiche nei comparti ambientali: suolo, acqua, aria), le distribuzioni di probabilità più comunemente utilizzate per la loro rappresentazione sono:

- distribuzione gaussiana o normale
- distribuzione lognormale
- distribuzione gamma
- distribuzione non parametrica.

Nel seguito sono descritte sinteticamente le caratteristiche delle distribuzioni e i test utili per identificare quale di queste distribuzioni rappresenti al meglio l'insieme di dati in esame.

Distribuzione Gaussiana o normale – La distribuzione Gaussiana, o normale, è una distribuzione di tipo simmetrico la cui tendenza centrale è data dal calcolo della media aritmetica dei valori $X_1, X_2, X_3, \dots, X_n$ delle grandezze considerate.

La forma della distribuzione normale è descritta dalla funzione Densità di Probabilità, definita da due parametri: la media aritmetica e la varianza del campione, che è indice della dispersione dei dati rispetto al valor medio.

Funzione
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\left(x - \bar{x}\right)^2\right]$$

Media
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Varianza
$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \bar{x}\right)^2$$

dove n è il numero di valori considerati.

In Figura 1 è riportato un esempio di distribuzione normale.

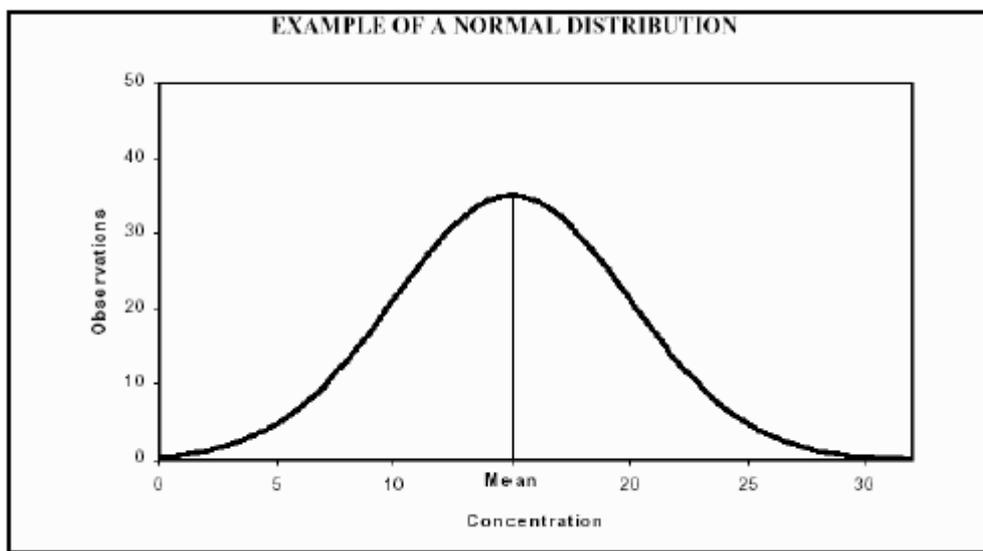


Figura 1 Esempio di distribuzione normale

Distribuzione lognormale – La distribuzione lognormale è un tipo di distribuzione asimmetrica, derivante dal calcolo della media geometrica dei valori. La sua forma è più pendente di quella di una distribuzione normale ed è delimitata a sinistra dallo zero, mentre la parte finale all'altra estremità risulta avere una specie di coda più lunga di quella normale. Quindi, la distribuzione lognormale è caratterizzata da una asimmetria positiva (coda a destra) dovuta al fatto che ad un'elevata frequenza di valori bassi si associa una coda di valori molto meno frequenti ma, allo stesso tempo, molto elevati.

La distribuzione lognormale è generalmente definita da due parametri \bar{y} e σ_y^2 (media e varianza della variabile trasformata $y = \ln x$).

Funzione
$$f(x) = \frac{1}{x\sigma_y\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_y^2}\left(\ln x - \bar{y}\right)^2\right] \quad x > 0, \quad -\infty < \bar{y} < \infty,$$

$$\sigma_y > 0$$

Media
$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n \ln x_i =$$

Varianza
$$\sigma_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

dove n è il numero di valori considerati. In Figura 2 è riportato un esempio di distribuzione lognormale.

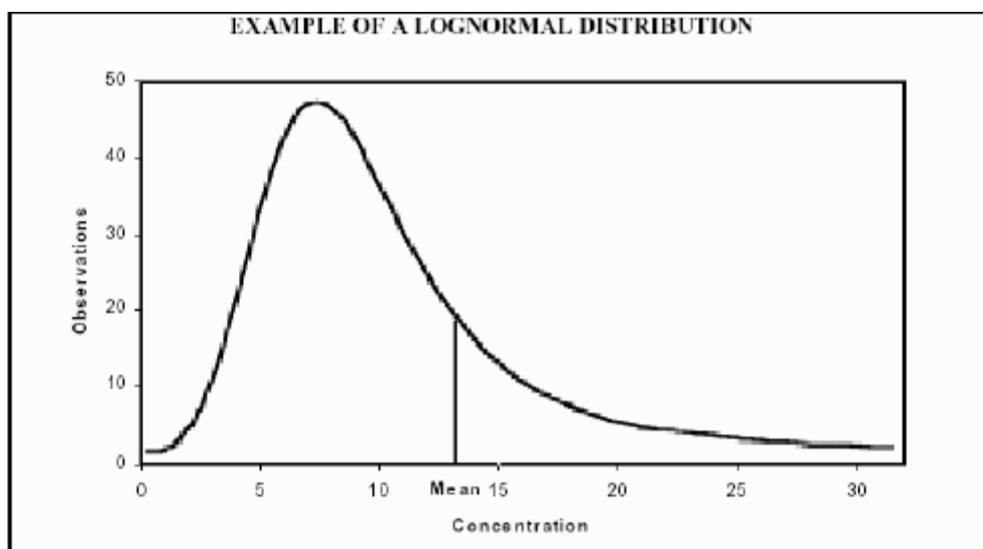


Figura 2 Esempio di distribuzione lognormale.

Distribuzione Gamma - Molti data set che presentano asimmetrie possono essere rappresentati sia mediante una distribuzione lognormale che da una distribuzione di tipo gamma, specialmente nei casi in cui il numero di campioni n è inferiore a 70-100.

La distribuzione gamma è generalmente definita da due parametri: k (parametro di forma) e θ (parametro di scala); il loro prodotto è pari alla media aritmetica

Funzione
$$f(x, k, \theta) = \frac{1}{\theta^k \Gamma(k)} x^{(k-1)} e^{(-\frac{x}{\theta})} \quad x > 0, k > 0, \theta > 0$$

Distribuzione non parametrica – Nel caso in cui non sia possibile dimostrare che i valori di un data set seguano una tra le suddette distribuzioni (ad esempio a causa dello scarso numero di campioni) o qualora risulti, dalla applicazione dei test statistici, che nessuna distribuzione approssimi bene l'insieme dei dati, allora si parla di data set non parametrici.

In tal caso esistono delle procedure specifiche, per l'individuazione del valore rappresentativo dell'insieme dei dati, indipendenti dai parametri statistici e dal tipo di distribuzione dei dati.

7 DESCRITTORI STATISTICI

Nel seguito sono descritte altre grandezze statistiche utili per lo studio del tipo di distribuzione dei dati.

Mediana - La mediana di una distribuzione è quel valore al di sopra del quale e al di sotto del quale si trova metà dell'insieme dei dati. La mediana si individua facilmente una volta ordinati in senso crescente gli n valori del data set:

se n è pari la mediana sarà il valore $x[(n+1)/2]$

se n è dispari la mediana sarà il valore $\frac{1}{2}(x_{n/2} + x_{[(n+2)/2]})$

Se la distribuzione è simmetrica, allora la mediana coincide con la media. Se la distribuzione dei dati è lognormale pendente verso destra la mediana sarà minore della media, e viceversa.

Coefficiente di skewness - Il valore di questo coefficiente fornisce una stima della asimmetria della forma di distribuzione dei dati. Si calcola secondo la seguente espressione:

$$asimmetria(skewness) = \frac{1}{\sigma^3} \sum_i \frac{(x_i - \bar{x})^3}{n}$$

Tale coefficiente può risultare:

- maggiore di zero: in tal caso la distribuzione avrà una coda verso destra;
- pari a zero: in tal caso la distribuzione sarà di tipo simmetrico, tipicamente gaussiana;
- minore di zero: in tal caso la distribuzione avrà una coda verso sinistra.

Il coefficiente di skewness non varia per traslazioni e cambiamenti di scala.

Coefficiente di curtosi - Il valore di questo coefficiente fornisce una stima della acutezza della curva di distribuzione dei dati. Si calcola secondo la seguente espressione:

$$curtosis = \frac{1}{\sigma^4} \sum_i \frac{(x_i - \bar{x})^4}{n}$$

Tale coefficiente può risultare:

- maggiore di 3: in tal caso la curva avrà un picco che determinerà una forma aguzza;
- pari a 3: in tal caso la distribuzione sarà di tipo simmetrico, con la forma a campana tipicamente gaussiana;
- minore di 3: in tal caso la forma della curva sarà appiattita.

Il coefficiente di curtosi non varia per traslazioni e cambiamenti di scala.

Coefficiente di variazione - E' un indice di dispersione che permette di analizzare la dispersione dei valori attorno alla media indipendentemente dall'unità di misura, fornendo un'indicazione sulla variabilità delle osservazioni rilevate. E' definito come il rapporto tra la deviazione standard dell'insieme dei dati ed il valore assoluto della loro media aritmetica:

$$CV = \frac{\sigma}{|\bar{x}|}$$

In particolare:

- se $CV=1$ vuol dire che $\sigma = \bar{x}$ e la media \bar{x} non è un indice corretto per la rappresentazione dei dati;

- se $CV=0$ vuol dire che $0 = \sigma$ e la media x è un indice appropriato per la rappresentazione dei dati;
- se $CV > 0,5$ la media x non è un indice corretto;
- se $CV \leq 0,5$ la media x è un indice corretto.

Per la rappresentazione della distribuzione dei dati sono utilizzate anche rappresentazioni grafiche, la cui scelta dipende dal tipo di dati da rappresentare; tra quelle di uso più comune sono:

A – istogrammi

B – box-plot,

C – curve cumulative di frequenza

Gli istogrammi

Sono grafici a barre verticali, nei quali le misure della variabile casuale sono riportate lungo l'asse orizzontale, mentre l'asse verticale rappresenta il numero assoluto, oppure la frequenza relativa o quella percentuale, con cui compaiono i valori di ogni classe.

Box plot

I box plot (Figura 3) sono dei diagrammi che riassumono gli aspetti principali di una distribuzione di valori; la base inferiore e superiore del rettangolo rappresentano rispettivamente il 25 e il 75 percentile. La linea all'interno del rettangolo rappresenta la mediana (ovvero il 50 percentile). Accanto a questi parametri statistici fondamentale, il box plot deriva altri valori importanti per l'identificazione dei valori anomali; con il termine gradino (step) si indica 1,5 volte la differenza fra il valore corrispondente al 75° percentile e quello al 25° percentile. I valori posti in corrispondenza di un gradino sopra la base superiore del rettangolo e un gradino sotto la base inferiore definiscono rispettivamente un limite superiore ed un limite inferiore (upper e lower fence). I limiti non sono solitamente visualizzati sul grafico, mentre sono riportati i valori adiacenti (cioè rispettivamente il primo valore inferiore al limite superiore, e il primo valore superiore al limite inferiore). I valori esterni a questi limiti sono usualmente considerati come outliers. Nel caso non vi siano outliers (verso i valori massimi e/o verso i valori minimi) i valori adiacenti superiore ed inferiore coincideranno rispettivamente con i valori massimo e minimo delle osservazioni.

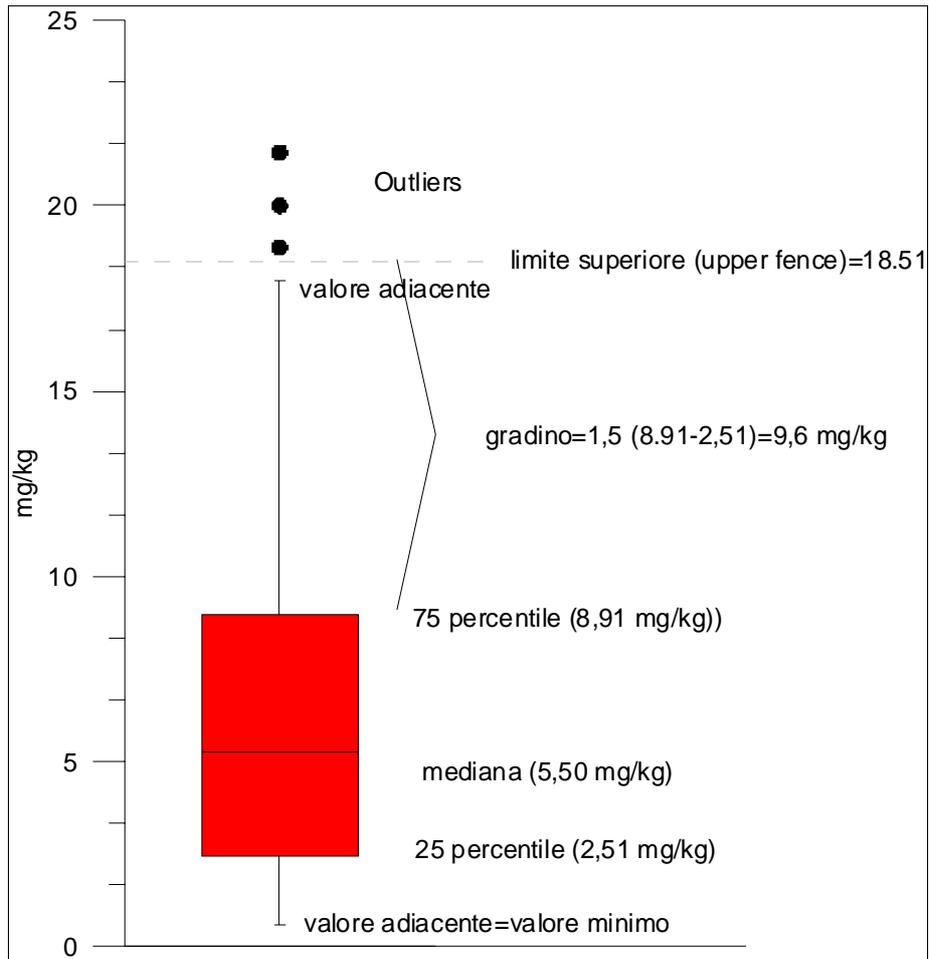


Figura 3: Parametri statistici rappresentati nel box plot

Nell'esempio, poiché il valore più basso del campione è maggiore del limite inferiore (che nel caso specifico sarebbe addirittura negativo) l'adiacente inferiore coincide con il valore minimo delle osservazioni.

Curve cumulative di frequenza

Per la costruzione della distribuzione cumulativa di frequenza, si ordinano le osservazioni in modo crescente: $x_1 < x_2 < x_3 < \dots < x_i < \dots < x_n$. Ad ogni valore delle osservazioni così ordinate si assegna il valore della frequenza assoluta AF_i (cioè in numero di volte che quel valore è stato osservato); si calcola quindi la frequenza cumulativa attraverso la relazione:

$$CF_i = \sum_{j=1}^i AF_j$$

dove il contatore j si riferisce al numero delle classi di frequenza (che possono essere uguali o minori rispetto al numero dei campioni)

CF_i rappresenta il numero di osservazioni che sono minori o uguali al valore a $x_{(i)}$,

Le percentuali cumulative per ogni valore di i si ottengono dalla normalizzazione di CF_i

$$Y(i) = 100 \frac{CF_i}{(n+1)}$$

Quando il grafico delle percentuali cumulative è costruito utilizzando, per ogni i , il valore $x(i)$ e $Y(i)$, utilizzando per l'asse $y(i)$ la scala probabilità, la curva viene definita, con la terminologia anglosassone, probability plot.

Dall'andamento della curva ottenuta (Figura 4) si possono ottenere delle informazioni circa la distribuzione del campione. Un andamento lineare è indice di un campione normalmente distribuito; in alcuni casi andamenti curvilinei possono essere resi lineari utilizzando la scala logaritmica per i valori di $x(i)$; in questo caso la distribuzione sarà log-normale.

A fronte di gap o "salti" ovvero a fronte di variazioni di pendenza della curva ottenuta potranno essere considerati dei valori soglia tali da individuare due o più popolazioni (es. il tratto rappresentativo del fondo e un segmento rappresentativo una popolazioni i cui valori sono determinati ad. es. da contaminazione).

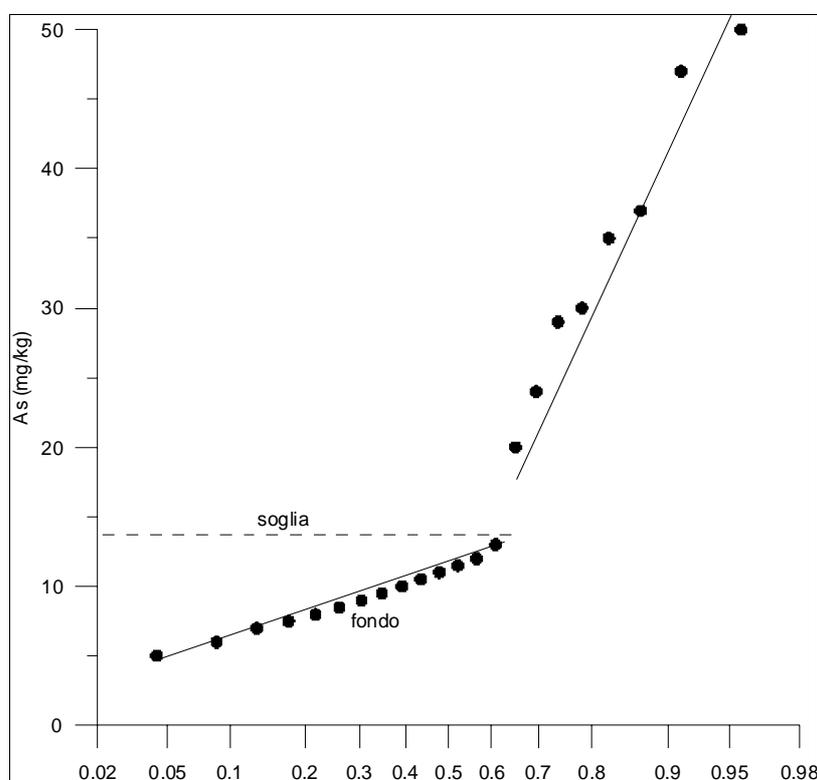


Figura 4 Curva cumulativa di frequenza

8 DEFINIZIONE DEL TIPO DI DISTRIBUZIONE

Nel seguito sono sinteticamente riportati i principali test statistici, per una trattazione di maggiore dettaglio si rimanda al riferimento bibliografico corrispondente. Nella Tabella 3 è riportata una sintesi dei test più comunemente utilizzati per lo studio del tipo di distribuzione di una serie di dati.

“Shapiro e Wilk test” (“W test”)– Con questo test si può valutare se sussistono o meno le ipotesi di distribuzione normale o lognormale nei casi in cui il numero dei dati a disposizione sia inferiore a 50 ($n < 50$).

“D’Agostino Test” – Con questo test si può valutare se sussistono o meno le ipotesi di distribuzione normale o lognormale nei casi in cui il numero dei dati a disposizione sia uguale o superiore a 50 ($n \geq 50$).

“Normal Quantile-Quantile (Q-Q) Plot” – E’ un test grafico la cui attendibilità, se non viene accompagnato da altri test più completi (come il “W test” o il “Lilliefors Test”), è piuttosto scarsa. E’ tuttavia utile per avere una prima approssimativa idea sulla distribuzione che assumono i dati in caso di ipotesi di distribuzione normale o lognormale.

“Lilliefors Test” – Viene utilizzato, nel caso di ampi data set ($n > 1000$), per verificare la normalità o la lognormalità di una distribuzione di dati.

“Quantile-Quantile (Q-Q) Plot per distribuzioni gamma” – E’ un test grafico la cui attendibilità, se non viene accompagnato da altri test più completi (come l’Anderson Darling test” o il “Kolmogorov-Smirnov test”), è piuttosto scarsa. E’ tuttavia utile per avere una prima approssimativa idea sulla distribuzione che assumono i dati in caso di ipotesi di distribuzione gamma.

“Kolmogorov-Smirnov test” - Per l’applicazione di questo test non devono essere fatte assunzioni sul tipo di distribuzione dei dati. Lo stesso viene utilizzato per dimostrare che un certo data set segue la distribuzione ipotizzata, mediante il confronto tra un determinato parametro calcolato e il corrispondente valore critico tabellato.

“Anderson Darling test” - Questo test è simile al Kolmogorov –Smirnov test, ma più preciso, in quanto fa uso di una distribuzione specifica per il calcolo dei valori critici (diversi dunque per ogni tipo di distribuzione), con i quali verrà confrontato il parametro calcolato.

TIPO DI TEST	TIPO DI DISTRIBUZIONE				RIF. BIBLIOGRAFICO
	NORMALE	LOG NORMALE	GAMMA	NON PARAMETRICA	
“Shapiro e Wilk test” (n < 50)	×	×	—	—	(Gilbert, 1987), (Software ProUCL)
”D'Agostino test“ (n = 50)	×	×	—	—	(Gilbert, 1987)
“Normal Quantile-Quantile (Q-Q) Plot”	×	×	—	—	(Software ProUCL)
“Lilliefors Test”	×	×	—	—	
“Gamma Quantile-Quantile (Q-Q) Plot”	—	—	×	—	
“Kolmogorov-Smirnov test”	—	—	×	—	
“Anderson Darling test”	—	—	×	—	

Tabella 3: Test per la selezione del tipo di distribuzione

9 CONFRONTO FRA I VALORI DI FONDO E I VALORI SITO SPECIFICI

L'obiettivo finale di una indagine statistica può essere ricondotto alla accettazione/negazione di un ipotesi previa definizione anche della probabilità di sbagliare la decisione. Nel caso specifico del confronto tra la distribuzione dei dati del fondo e del sito, il parametro di riferimento Δ sarà dato dalla differenza fra la concentrazione *rappresentativa* di un analita X in aree potenzialmente contaminate e la concentrazione *rappresentativa* del fondo.

Gli strumenti che consentono di pervenire ad un a decisione circa il rigetto o la accettazione dell'ipotesi nulla sono dei test statistici che sono adottati in funzione delle caratteristiche delle 2 popolazioni (sito e fondo), il risultato di questi test dipende anche dall'errore α connesso alla decisione.

In funzione del tipo e delle caratteristiche di distribuzione dei dati disponibili, è possibile selezionare il tipo di test più idoneo da utilizzare; nella Tabella 4 e nella Tabella 5 sono indicati alcuni dei test più comunemente usati per il confronto di popolazioni.

CARATTERISTICHE DELLE POPOLAZIONI DEI CAMPIONI DEL STO E DI FONDO	TEST DA ADOTTARE
Numero di campioni $N > 25$, distribuzione di frequenza normale o log normale, varianza simile, pochi valori inferiore al d.l.	Student t test
Distribuzione di frequenza normale o log normale, varianze delle due popolazioni diverse	Test t di Satterthwaite
Nessun limite circa la distribuzione di frequenza	Wilcoxon rank sum test

Tabella 4: Tipologia dei Test Statistici

TEST	OBIETTIVI/ASSUNZIONI	VANTAGGI	SVANTAGGI
Slippage Test	L'obiettivo è valutare la differenza nella coda di destra di una distribuzione di concentrazioni (valori più alti) di due popolazioni (sito e fondo); Si può applicare anche in presenza di un numero elevato di n.d.; È stato determinato almeno un valore del fondo diverso da n.d.; Non sono richieste assunzioni riguardanti la forma della distribuzione di concentrazioni del sito e di fondo.	Semplice da applicare; Non sono necessarie assunzioni circa il tipo di distribuzione; Si può applicare anche in presenza di numerosi n.d.; Può essere applicato parallelamente all'applicazione di test che mirano al confronto tra medie (o mediane).	Può richiedere un gran numero di dati affinché si abbia una potenza sufficiente per rilevare la differenza tra le concentrazioni di un sito e quelle di fondo
Quantile Test	L'obiettivo è valutare la differenza nella coda di destra di una distribuzione di concentrazioni (valori più alti) di due popolazioni (sito e fondo); I valori n.d. non devono essere tra i valori r più elevati nel set di dati del sito e di fondo; Non sono richieste assunzioni riguardanti la forma della distribuzione di concentrazioni del sito e di fondo.	Abbastanza semplice da applicare Non sono necessarie assunzioni circa il tipo di distribuzione; Può avere maggiore potenza per rilevare la differenza tra distribuzioni del sito e quelle di fondo rispetto ad altri test; Si può applicare anche in presenza di numerosi n.d.	Può richiedere un gran numero di dati affinché si abbia una potenza sufficiente per rilevare la differenza tra le concentrazioni di un sito e quelle di fondo Potrebbe risultare inefficace nel caso in cui fossero presenti n.d. tra i valori r più elevati.
Wilcoxon Rank Sum Test (WRS test)	L'obiettivo è valutare la differenza tra le mediane di due popolazioni (sito e fondo); Un solo detection limit (tutti i n.d. devono avere lo stesso valore) che deve essere minore del più piccolo valore di concentrazione rilevato; Non sono richieste assunzioni riguardanti la forma della distribuzione di concentrazioni del sito e di fondo.	Non sono necessarie assunzioni circa il tipo di distribuzione; Di solito, il test ha più potenza per determinare uno scostamento della mediana, rispetto a altri test, quando le distribuzioni dei valori del sito e quelle del fondo sono asimmetriche; Può essere applicato parallelamente all'applicazione di test che mirano a valutare la differenza tra la coda destra di due distribuzioni (Slippage test e Quantile test).	Relativamente più complicato da applicare; la presenza di numerosi n.d. pregiudica l'applicabilità del test.
Gehan Test	L'obiettivo è valutare la differenza tra le mediane di due popolazioni (sito e fondo); Possono essere presenti differenti valori del detection limit; Non sono richieste assunzioni riguardanti la forma della distribuzione di concentrazioni del sito e di fondo.	Può essere utilizzato in caso di presenza di differenti valori del detection limit; gli stessi vantaggi del WRS test	Il calcolo manuale può risultare relativamente complesso; Le performance del test non sono note quanto quelle del WRS test.
Test t-student	L'obiettivo è valutare la differenza tra le medie di due popolazioni (sito e fondo); Entrambe le distribuzioni devono presentare una distribuzione normale; I valori n.d. non devono avere un impatto significativo sul calcolo della media (meno del 15% dei dati sono n.d.); La distribuzione di concentrazioni del sito e di fondo deve essere la stessa (varianza).	È il test che possiede la maggiore potenza nella verifica dello scostamento dei valori medi di due popolazioni che presentano una distribuzione normale.	Il test richiede una valutazione statistica della assunzione di uguaglianza tra la varianza della distribuzione del sito e quella del fondo; in genere la potenza è inferiore al WRS test, nel caso in cui le popolazioni non presentassero una distribuzione normale; L'assunzione di "normalità" viene spesso trascurata; il risultato del test può essere influenzato dalla presenza di outliers; non si adatta a set di dati che presentano numerosi n.d.
Test t-Satterthwaite	L'obiettivo è valutare la differenza tra le medie di due popolazioni (sito e fondo); Entrambe le distribuzioni devono presentare una distribuzione normale; non devono essere presenti valori n.d.; Si presume che la distribuzione di concentrazioni del sito e di fondo non presentino la stessa forma (varianza).	Il test può essere applicato quando la distribuzione dei valori del sito e quella dei valori del fondo hanno varianze differenti	Il calcolo manuale può risultare relativamente complesso; presenta gli stessi svantaggi del t-Test

Tabella 5: Applicabilità dei Test Statistici

Test t-student

In questo test si mettono a confronto le medie di due popolazioni rappresentate rispettivamente da due set di campioni casuali:

- un set di m dati: x_1, x_2, \dots, x_m per la prima popolazione (che rappresenta ad esempio la distribuzione dei valori di background);
- un set di m dati: y_1, y_2, \dots, y_n per la seconda (che rappresenta ad esempio la distribuzione dei valori del sito).

Le condizioni necessarie per la corretta applicazione di questo test sono:

- la variabilità delle due popolazioni espressa dalle rispettive varianze sia approssimativamente uguale;
- i due campioni siano indipendenti (condizione di indipendenza);
- le due popolazioni devono avere distribuzione approssimativamente normale.

Questo test è robusto in riferimento alle condizioni di normalità e di eguaglianza delle varianze, mentre non lo è nel caso di presenza di outlier.

L'ipotesi nulla è che la differenza delle medie delle due popolazioni sia nulla:

$$H_0: \mu_1 - \mu_2 = 0$$

Ovvero che le medie delle due popolazioni non siano significativamente diverse.

Dopo avere calcolato per ogni campione le medie \bar{x} e \bar{y} e le varianze S_x^2 e S_y^2 si calcola la deviazione standard congiunta S_E data da:

$$S_E = \sqrt{\frac{(m-1)S_x^2 + (n-1)S_y^2}{(m-1) + (n-1)}}$$

Quindi si calcola il parametro t dato da:

$$t = \frac{\bar{x} - \bar{y}}{S_E \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

dalla Tabella 7 si ricava il valore critico di $t_{(1-\alpha)}$ tale che il $100(1-\alpha)\%$ della distribuzione t di Student, con $(m+n-2)$ gradi di libertà, sia inferiore a $t_{(1-\alpha)}$.

Se $t > t_{1-\alpha}$ l'ipotesi di nullità può essere rifiutata, cioè il sito è contaminato; se $t \leq t_{(1-\alpha)}$ non sussiste l'evidenza per rifiutare l'ipotesi nulla; è necessario calcolare la dimensione del campione (m^* e n^*) necessaria a ridurre le probabilità α e β di commettere errori del primo o del secondo tipo per una differenza fissata tra le medie delle due aree (nel nostro caso pari a 0).

Una volta specificata la probabilità β (es. 20%), è possibile verificare se il campione utilizzato ha una dimensione appropriata:

$$m^*=n^* = \frac{2S_E^2 (z_{(1-\alpha)} + z_{(1-\beta)})^2}{\delta_m} + 0.25z_{(1-\alpha)}^2$$

dove m^* e n^* sono il numero dei dati del campione che rendono, per determinati valori di a e b , sufficientemente attendibili i risultato del test;

$Z_{(1-\alpha)}$ e $Z_{(1-\beta)}$ rappresentano i valori del percentile della distribuzione normale standard;

δ_m la differenza dei valori medie dei due campioni $\bar{x} - \bar{y}$.

Si confrontano i valori di m e n con m^* e n^* : se $m^* \leq m$ e $n^* \leq n$, la probabilità di commettere un errore del primo tipo è accettabile. I risultati del test possono essere:

- a) è stata respinta l'ipotesi nulla e quindi sembra che $\mu_1 - \mu_2 > 0$, quindi il sito è contaminato;
- b) non è stata respinta l'ipotesi nulla ed è stata accettata la probabilità di commettere un errore del primo tipo: probabilmente è vero che $\mu_1 - \mu_2 \leq 0$, quindi il sito può essere considerato pulito;
- c) l'ipotesi nulla non è stata respinta e non è stata accettata la probabilità di commettere un errore del primo tipo: la differenza delle medie è probabilmente minore di 0, quindi il sito è probabilmente pulito, ma questa conclusione rimane incerta a causa delle dimensioni troppo piccole dei campioni.

Test t di Satterthwaite – varianze diverse

Questo test parametrico, viene usato per comparare le medie di due popolazioni quando le loro varianze sono disuguali. Esso richiede i seguenti assunti:

- i due campioni siano indipendenti (condizione di indipendenza);
- le due popolazioni devono avere distribuzione approssimativamente normale

Siano x_1, x_2, \dots, x_m e y_1, y_2, \dots, y_n i due campioni costituiti rispettivamente da m ed n misure e rappresentanti due popolazioni e caratterizzati da varianze S_x^2 e S_y^2 differenti di cui si vogliono comparare le medie \bar{x} e \bar{y} .

I passi per l'applicazione del test t di Satterthwaite prevedono:

Il calcolo della deviazione standard congiunta S_{NE} data da:

$$S_{NE} = \sqrt{\left(\frac{S_x^2}{m} + \frac{S_y^2}{n} \right)}$$

il calcolo del parametro t di Satterthwaite:

$$t = \frac{\bar{x} - \bar{y}}{S_{NE}}$$

dalla Tabella 7 si ricava il valore critico di $t_{(1-\alpha)}$ tale che il $100(1-\alpha)\%$ della distribuzione t di Student, con $(m+n-2)$ gradi di libertà, sia inferiore a $t_{(1-\alpha)}$.

Se $t > t_{1-\alpha}$ l'ipotesi di nullità può essere rifiutata, e quindi sembra che $\mu_1 - \mu_2 > 0$ cioè il sito è contaminato;

se $t \leq t_{(1-\alpha)}$ non sussiste l'evidenza per rifiutare l'ipotesi nulla; si può accettare la probabilità di commettere un errore del primo tipo: probabilmente è vero che $\mu_1 - \mu_2 \leq 0$, cioè che il sito è pulito. In alternativa si può non accettare la probabilità di commettere un errore del primo tipo: la differenza delle medie è probabilmente minore di 0, ma questa conclusione rimane incerta a causa delle dimensioni troppo piccole dei campioni.

Non esistono tuttavia formule semplici per la stima di m^* e n^* come per il test t di Student ed è quindi necessario il ricorso ad un esperto di statistica.

Wilcoxon rank sum test

Laddove le assunzioni sulle caratteristiche delle distribuzioni sono difficili da verificare o da soddisfare per entrambe le popolazioni. In questo caso è possibile utilizzare test che, mettendo a confronto la forma e la posizione di due distribuzioni anziché i relativi parametri statistici (media, mediana, ecc.), risultano svincolati dai tipi di distribuzione.

Questi test, detti non parametrici, verificano un'ipotesi nulla del tipo "H₀: la distribuzione delle popolazioni 1 e 2 sono identiche", contro l'ipotesi alternativa "H_a: parte della distribuzione della popolazione 1 è posta a destra/sinistra della distribuzione della popolazione 2".

Ad esempio si possono applicare nel caso si voglia verificare se un'area d'interesse è più contaminata di un'area di riferimento: in questo caso l'ipotesi nulla da verificare sarebbe l'uguaglianza tra le distribuzioni delle concentrazioni nei due siti.

Il Wilcoxon rank sum test (conosciuto anche come "Mann-Whitney Test") si applica nel caso si abbiano numerosi dati ($n \geq 20$ e $m \geq 20$) che descrivono le caratteristiche del sito e del fondo.

α è la probabilità che il test di Wilcoxon dichiarati in modo scorretto che le concentrazioni del sito sono superiori a quelle di fondo e cioè che vi sia un problema di contaminazione del sito da affrontare quando però tale situazione non è vera.

Dato un set di m dati: x_1, x_2, \dots, x_m per la prima popolazione (che rappresenta ad esempio la distribuzione dei valori di background) ed un set di n dati: y_1, y_2, \dots, y_n per la seconda (che rappresenta ad esempio la distribuzione dei valori del sito), tutti i dati vengono uniti ed ordinati, a prescindere dalla popolazione di partenza e viene loro attribuito il rango agli $n+m$ valori del sito e di fondo, incominciando da un rango 1 per il valore più piccolo e così via. Se si hanno valori uguali nella stessa posizione per un numero inferiore al 40% del totale si effettua una mediazione del rango. Viene calcolata la somma dei ranghi R del sito come nell'esempio seguente, posto a puro titolo esplicativo, in quanto il numero di campioni utilizzato sarebbe nel caso in esame insufficiente:

Indicativo del dato	<u>y1</u>	x4	x3	<u>y2</u>	<u>y4</u>	x1	<u>y3</u>	x2	...
Valore	2,2	2,3	2,8	2,8	3,2	3,3	3,6	3,6
posizione	1	2	3	4	5	6	7	8	n+m
Rango	<u>1</u>	2	3,5	<u>3,5</u>	<u>5</u>	6	<u>7,5</u>	7,5	

Tabella 6 Esempio di calcolo del Rango

$$Ry=1+3,5+7,5+....$$

si determina quindi il valore del parametro $W_{(1-\alpha)}$:

$$W_{(1-\alpha)} = \frac{n(n+1)}{4 + z_{(1-\alpha)} \sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

dove il valore $z_{(1-\alpha)}$ è il 100(1- α) percentile della distribuzione normale standard della Tabella 8.

Se $Ry > W_{(1-\alpha)}$ le concentrazioni del sito sono significativamente superiori a quelle del fondo, ovvero il sito risulta contaminato.

Gradi di libertà	1- α								
	,70	,75	,80	,85	,90	,95	,975	,99	,995
1	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,65
2	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,536	0,691	0,866	1,074	1,34	1,753	2,131	2,602	2,947
16	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898
18	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,533	0,6880	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,533	,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797
25	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787
26	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779
27	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771
28	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763
29	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756
30	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
40	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704
60	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660
120	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617
	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576

Tabella 7 Valori critici per la distribuzione t di Student

I valori dell'ultima riga corrispondono a valori critici per la distribuzione normale standard

n	α	m																		
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	0,05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	0,10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
3	0,05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	0,10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
4	0,05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	0,10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23
5	0,05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	0,10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
6	0,05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	0,10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
7	0,05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	0,10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
8	0,05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	0,10	3	6	8	11	4	17	20	23	25	28	31	34	37	40	43	46	49	52	55
9	0,05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	0,10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
10	0,05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	0,10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
11	0,05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	0,10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
12	0,05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	0,10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87
13	0,05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	0,10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	0,05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	0,10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	0,05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	0,10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	0,05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	0,10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120
17	0,05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	0,10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
18	0,05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	0,10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
19	0,05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	0,10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
20	0,05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	0,10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

Tabella 8 Valori critici per il Wilcoxon Rank Sum Test

(n=numero di misure nel sito, m=numero di misure del fondo)