

A Paolo

*"Essere amata è più che essere ricca,
perché vuol dire essere felice"*

INDICE

INTRODUZIONE.....	I
CAPITOLO 1 – LA GEOMETRIA FRATTALE E LE CURVE DI RITENZIONE IDRICA.....	1
1.1 Premessa.....	1
1.2 I frattali	4
1.2.1 Teoria frattale e proprietà dei frattali	4
1.2.2 Dimensione di <i>Hausdorff</i>	6
1.2.3 Dimensione di <i>box-counting</i> e algoritmi di risoluzione.....	8
1.2.4 Dimensione frattale per oggetti frattali deterministici.....	11
1.3 Il terreno come mezzo poroso	14
1.3.1 Contenuto di acqua nel suolo	15
1.3.2 Zona insatura e potenziale dell'acqua nel suolo	17
1.3.3 Curva di ritenzione idrica nel suolo	22
1.3.4 Cenni sul moto dell'acqua del suolo	26
1.4 Modelli frattali per la caratterizzazione delle curve di ritenzione idrica (<i>WRC</i>)	29
1.4.1 I modelli di <i>Tyler & Wheatcraft</i> [1990] e di <i>Rieu & Sposito</i> [1991]	30
1.4.2 Il modello <i>Pore-Solid Fractal</i> [<i>Bird et al.</i> , 2000]	33
CAPITOLO 2 – LA GEOSTATISTICA COME STUDIO DELLA CORRELAZIONE SPAZIALE	36
2.1 Premessa	36

2.2 La distribuzione dei punti campionati: analisi del vicinaggio	38
2.2.1 <i>Nearest Neighbor Analysis (NNA)</i>	38
2.3 Caratterizzazione dei processi spaziali	40
2.3.1 Variabile regionalizzata.....	40
2.3.2 Funzioni aleatorie.....	41
2.3.3 Momenti di primo e secondo ordine	43
2.3.4 Le ipotesi di stazionarietà	44
2.4 Analisi esplorativa dei dati (<i>Exploratory Data Analysis, EDA</i>).....	48
2.4.1 Analisi statistica descrittiva	48
2.4.2 Descrizione grafica del campione.....	49
2.4.3 Test statistici e verifica delle ipotesi.....	52
2.5 Analisi strutturale	60
2.5.1 Il variogramma sperimentale.....	61
2.5.2 Comportamento del variogramma in prossimità dell'origine.....	64
2.5.3 Comportamento tipico del variogramma	65
2.5.4 Lo studio delle anisotropie	66
2.5.5 Modelli di variogramma	68
CAPITOLO 3 – METODI DI STIMA PUNTUALE.....	72
3.1 Premessa.....	72
3.2 Alcune definizioni fondamentali.....	73
3.2.1 Stimatore, stima lineare e interpolatori.....	73
3.3 Metodi di interpolazione deterministici.....	76
3.3.1 Interpolazione dei vicini naturali (<i>Natural Neighbour Interpolation, NNI</i>)	76
3.3.2 Funzioni inverse della distanza (<i>Inverse Distance, ID</i>).....	76
3.3.3 <i>Splines</i>	77

3.4	Interpolazione spaziale stocastica: il <i>Kriging</i>	80
3.4.1	Premessa	80
3.4.2	Il <i>Kriging</i> semplice (<i>Simple Kriging, SK</i>).....	81
3.4.3	Il <i>Kriging</i> ordinario (<i>Ordinary Kriging, OK</i>).....	83
3.4.4	Proprietà del <i>Kriging</i> stazionario.....	86
3.4.5	Equazioni del <i>Kriging</i> non stazionario	87
3.5	La validazione incrociata (<i>Cross Validation</i>)	92
CAPITOLO 4 – CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO		
DEL TORRENTE TURBOLO		
		94
4.1	Descrizione del bacino oggetto di studio e del campo sperimentale	94
4.2	Misura del contenuto d’acqua	99
4.3	Stima della dimensione frattale	102
4.4	Analisi dei risultati	106
4.4.1	<i>Exploratory Data Analysis (EDA)</i>	106
4.4.2	Descrizione quantitativa della distribuzione dei punti e analisi del trend spaziale....	120
4.4.3	Analisi della correlazione spaziale.....	128
4.4.4	Validazione incrociata (<i>Cross Validation</i>)	142
4.4.5	Elaborazione delle mappe ottenute mediante l’interpolatore <i>Kriging</i> ordinario	146
CONCLUSIONI		168
RIFERIMENTI BIBLIOGRAFICI		171

Introduzione

L'obiettivo generale del seguente lavoro di tesi si propone di applicare particolari strumenti legati all'analisi spaziale di una specifica variabile (dimensione frattale), con lo scopo di identificare la variabilità di quest'ultima in una data area esaminata e, di conseguenza, conoscere il suo valore anche in punti in cui non sono state effettuate misure di campo.

Le caratteristiche idrauliche dei terreni dipendono fortemente dalla geometria dello spazio poroso e la quantificazione di tale geometria è stata caratterizzata negli ultimi anni attraverso la geometria frattale [Mandelbrot, 1982] poiché essa offre l'opportunità di relazionare le proprietà dell'acqua nel suolo alle proprietà strutturali dello stesso. La dimensione frattale [Falconer, 1990], quindi, rappresenta un parametro strettamente legato alle proprietà delle strutture dei pori nei suoli [Tyler & Wheatcraft, 1989; Perfect & Kay, 1991] che, quando misurate a diverse scale di risoluzione, sembrano spesso essere funzioni di una legge di potenza su un'appropriata scala di lunghezza. Tali leggi di potenza intese come i modelli frattali per la caratterizzazione delle curve di ritenzione idrica [Tyler & Wheatcraft, 1990; Rieu & Sposito, 1991; Bird et al., 2000] hanno come esponente proprio la dimensione frattale.

La distribuzione spaziale della dimensione frattale su un'area generica assegnata potrebbe risultare di grande importanza, quindi, allo scopo di rappresentare il fenomeno della ritenzione idrica nei suoli anche in punti privi di misura, o anche per evidenziare variazioni sensibili nella struttura del suolo stesso. Questo perché la curva di ritenzione idrica (*Water Retention Curve, WRC*) dei suoli influenza un vastissimo numero di processi naturali quali la germinazione dei semi e la disponibilità idrica per le piante, i flussi idrici nel suolo, l'attività microbica e quindi la degradazione della sostanza organica, l'evaporazione, i flussi di soluti ed in generale il bilancio idrico.

Avendo a disposizione osservazioni limitate e puntuali, insufficienti, quindi, a fornire una dettagliata descrizione spaziale della variabile esaminata, è stato possibile effettuare un'interpolazione spaziale allo scopo di valutare la variabilità della dimensione frattale nell'area oggetto di studio e stimare il rispettivo valore in qualsiasi punto del campo sperimentale preso in esame. Poiché la variazione delle proprietà del suolo dipende da molti processi fisici e chimici che interagiscono tra di loro in modo a volte apparentemente caotico, il risultato di questi processi è così complesso che la loro variazione sembra essere casuale. Questo significa che non siamo in grado di descrivere il determinismo della variazione delle proprietà del suolo e le funzioni matematiche non sono adeguate a questo scopo. A tal proposito, nel presente studio, la dimensione frattale è stata trattata come una funzione casuale descritta secondo un approccio di tipo stocastico utilizzando i metodi della teoria delle variabili regionalizzate proposta da *Matheron* [1971], nota come Geostatistica [*Matheron*, 1963; *Journel & Huijbregts*, 1978; *Clark*, 1984; *Hohn*, 1988; *Isaaks & Srivastava*, 1989; *Cressie*, 1991; *Goovaerts*, 1997; *Deutsch & Journel*, 1998; *Houlding*, 2000].

Il seguente studio è strutturato in modo tale da avere una visione completa e generale circa la parte teorica riguardante la geometria dei frattali, la descrizione del terreno e delle rispettive *WRC* e particolare importanza verrà data ai modelli frattali esaminati per la caratterizzazione delle stesse. Sarà effettuata una breve esposizione circa l'interpretazione dei dati spaziali che si hanno a disposizione e verrà data particolare enfasi all'interpolatore *Kriging* [*Krige*, 1951] in quanto è l'unico metodo di stima che tiene conto del modello di correlazione spaziale (variogramma).

Di conseguenza, i risultati ottenuti verranno interpretati in modo tale da pervenire allo scopo principale del lavoro. Nel presente studio, quindi, si avranno a disposizione come *output* per ciascun modello frattale la mappatura spaziale della variabile dimensione frattale nel dominio in cui è stato effettuato il campionamento, l'errore di stima commesso per mezzo dell'interpolatore *Kriging* e una misura circa la precisione della previsione fatta. Si cercherà, quindi, di individuare il valore della dimensione frattale in qualsiasi punto dell'area esaminata, verificando contemporaneamente l'attendibilità e l'applicabilità dei modelli frattali per la caratterizzazione

delle *WRC* e tentando di definire quali tra tutti i modelli è quello che meglio interpreta i dati reali (misurati).

Capitolo 1 – La geometria frattale e Le curve di ritenzione idrica

1.1 Premessa

La curva di ritenzione idrica (*Water Retention Curve, WRC*) è un'importante funzione idraulica per lo studio dei processi di flusso dell'acqua e dei soluti nella zona non satura del suolo; la stima e lo studio di tale funzione e/o delle grandezze ad essa correlate risulta essere di particolare interesse per la valutazione e, ovviamente, l'eventuale limitazione dei fenomeni di inquinamento riguardante le falde acquifere. La caratterizzazione delle curve di ritenzione idrica può essere effettuata attraverso precisi metodi (diretti), sviluppati sia in laboratorio che in pieno campo. I metodi di laboratorio, anche se presentano le difficoltà connesse al prelievo di campioni non rimaneggiati di suolo, consentono misure in ambienti controllati, queste ultime effettuate più agevolmente con l'impiego di strumenti di misura anche molto sofisticati i quali forniscono di solito rilievi più accurati. In laboratorio è possibile, inoltre, condurre prove simultanee su più campioni, anche provenienti da zone diverse, consentendo una caratterizzazione più spedita ed efficiente. L'applicabilità dei metodi di laboratorio è però vincolata dalle tecniche di prelievo, dalla dimensione dei campioni di suolo da prelevare, dall'estensione dell'area che è necessario caratterizzare e spesso anche da costi eccessivi. Se si è in presenza di un suolo con evidenti stratificazioni, oppure molto strutturato e con notevole contenuto di scheletro, è alquanto difficile prelevare campioni con minimo disturbo e che possano considerarsi sufficientemente rappresentativi delle reali situazioni di campo. Per ovviare ai problemi cui ora si è accennato si preferisce ricorrere a prove idrauliche eseguite direttamente in campo, ma anche in tal caso ci sono limitazioni essenzialmente legate all'onerosità delle misure, che peraltro sono eseguite utilizzando strumenti che talvolta danno

luogo a incertezze su alcune entità. Pertanto, data la complessità del problema, non è stato possibile mettere a punto un unico sistema di misura delle proprietà idrauliche del suolo in grado di operare bene per tutti i tipi di suolo e per le diverse condizioni di prova. Un'ampia rassegna di metodi diretti, sviluppati per prove da condurre in laboratorio o in campo, per la misura delle proprietà idrauliche del suolo è stata fornita da *Gee & Ward* [1999]. Sebbene relativamente semplici dal punto di vista concettuale, i metodi diretti presentano, quindi, limitazioni che ne riducono l'impiego pratico. Queste difficoltà, di conseguenza, hanno motivato numerosi tentativi di determinare le proprietà idrauliche del suolo da informazioni già disponibili o più facilmente acquisibili e meno costose da misurare come, ad esempio, i dati di tessitura, di contenuto di materia organica, di densità apparente e di altre proprietà fisico-chimiche del suolo [*van Genuchten*, 1980; *Wosten & van Genuchten*, 1988]. Queste tecniche semplificate sono in pratica delle relazioni empiriche o semi-empiriche che trasferiscono l'informazione dalle proprietà fisico-chimiche di un suolo alle sue caratteristiche idrauliche e sono note in letteratura anglosassone con il nome di *pedotransfer functions (PTF)* [*Pachepsky & Rawls*, 1999; *Minasny et al.*, 2008]. Il concetto di una funzione di trasferimento pedologico-idraulico (*PTF*), anche se formalmente denominato in tal modo solo nel 1989 [*Bouma*, 1989], era in realtà ben noto da tempo e si fonda essenzialmente sulla ricerca di una relazione semplificata per stimare una proprietà del suolo tramite altre variabili più agevoli da misurare. Tuttavia, poiché fra le proprietà di un suolo, quelle idrauliche sono difficili da misurare, era logico che i primi sforzi si sono orientati verso una stima semplificata proprio delle principali variabili che possono sintetizzare la risposta idrologica di un suolo. Diversi Autori hanno dimostrato la validità dell'applicazione delle *PTF*, ma ne hanno anche messo in evidenza alcuni limiti importanti di applicabilità [*Tietje & Tapkenhinrichs*, 1993; *Espino et al.*, 1995]. Altri ancora [*Vereecken et al.*, 1992] hanno valutato gli errori nelle previsioni delle *PTF* e i loro effetti sulla simulazione del movimento dell'acqua nel suolo; *Schaap & Leij* [1998] hanno considerato l'incertezza delle *PTF* in funzione del livello di accuratezza dei dati di suolo da cui le *PTF* sono state definite e calibrate; *Minasny et al.*, [1999] hanno analizzato l'incertezza di previsione della ritenzione idrica, analizzando l'incertezza associata ai parametri delle *PTF*, evidenziando come l'incertezza nei parametri è piccola rispetto all'incertezza delle variabili di ingresso.

Una delle più recenti novità nel campo dell'idrogeologia è costituita dal tentativo di applicare e utilizzare i concetti e i metodi della cosiddetta geometria frattale, sviluppata negli anni settanta dello scorso secolo dal matematico di origine polacca *B. B. Mandelbrot*. I recenti sforzi di applicare i concetti frattali ai fenomeni idraulici nei suoli, hanno permesso di identificare nuovi principi di organizzazione per comprendere la struttura dei medesimi [*van Damme*, 1995]. Specifiche applicazioni della teoria dei frattali hanno permesso di sviluppare modelli matematici in grado di definire la *WRC* correlandola contemporaneamente alla struttura intrinseca del terreno stesso, fornendo così un ulteriore strumento utile per superare il vuoto tra l'uso dei modelli semi-empirici e il significato fisico stesso dei loro parametri. In questo capitolo verrà descritta in primis la parte teorica riguardante la geometria dei frattali, verrà effettuata una descrizione del terreno come mezzo poroso e delle caratteristiche delle *WRC* e, infine, verrà esposto lo studio riguardante alcuni dei modelli frattali scelti nel presente lavoro per la caratterizzazione delle medesime.

1.2 I frattali

1.2.1 Teoria frattale e proprietà dei frattali

Il termine frattale, coniato nel 1975 da *Mandelbrot*, trae la sua etimologia dal latino *fractus*, ossia dal verbo frangere (rotto, spezzato) per indicare insiemi frastagliati o spezzettati. Non c'è una definizione universalmente riconosciuta di frattale ma, in modo informale, si può affermare che un frattale è un oggetto geometrico di forma estremamente irregolare o estremamente interrotta o frammentata che, se ingrandito, nasconde sempre maggiori dettagli. La geometria frattale può essere considerata come un campo intermedio tra l'ideale ordine geometrico euclideo e il caos geometrico dell'irregolarità e della frammentazione [*Mandelbrot*, 1989], ovvero come primo persuasivo linguaggio visuale delle forme complesse presenti in natura [*Musgrave*, 1994]. Se osserviamo la natura, infatti, vediamo che essa rifiuta la simmetria e che inoltre in essa non esistono linee, piani e solidi perfetti così come descrivono le regole canoniche euclidee. "...*le nuvole non sono sfere, le montagne non sono coni... Io credo che la natura esibisca delle forme che abbiano un grado di complessità più elevato e totalmente differente se paragonato alla descrizione euclidea della natura*", così *Mandelbrot* [1982] è solito descrivere l'inadeguatezza della geometria euclidea nella descrizione della natura. La non idoneità della geometria euclidea nel descrivere la forma degli oggetti presenti in natura può essere compresa pensando al fatto che, per quest'ultima, un oggetto diventa sempre più semplice via via che lo si guardi nella sua particolarità. Ad esempio, un solido tridimensionale diventa un piano bidimensionale, poi una linea fino ad arrivare ad un punto. Un oggetto reale dimostra invece di avere tanti più dettagli quanto più approfonditamente lo si guarda. Gli oggetti frattali, a tal proposito, godono della stessa proprietà: più da vicino sono esaminati, più dettagli possono essere osservati. Questa proprietà è detta auto-somiglianza (*self-similarity*) e, cioè, a qualsiasi scala si osservi, l'oggetto presenta sempre gli stessi caratteri globali. Una lecita e intuitiva definizione di frattale, quindi, potrebbe essere come quell'insieme in cui esiste un'auto-somiglianza tra la parte e il tutto [*Mandelbrot*, 1986]. Considerando l'invarianza rispetto alle trasformazioni auto-somiglianti e auto-affini che li hanno generati [*Mandelbrot*, 1985], i frattali possono essere suddivisi in auto-somiglianti (*self-similar fractals*) e auto-affini (*self-affine*

fractals). I primi saranno pertanto caratterizzati dalla cosiddetta auto-somiglianza di scala, ossia dall'invarianza rispetto alle trasformazioni di auto-somiglianza, mentre i frattali auto-affini dall'auto-affinità di scala, ovvero dall'invarianza rispetto alle trasformazioni affini (trasformazioni lineari associate a traslazioni). Si ricorda brevemente che un oggetto è definito auto-affine quando può essere decomposto nell'unione di n distinti sottoinsiemi non-intersecantisi ognuno dei quali in modo anisotropo congruente all'originario; cambiamento di scala anisotropo significa inoltre che il fattore usato per calcolare le coordinate dell'oggetto ingrandito o ridotto dipende dalla direzione, ovvero dai differenti rapporti di scala utilizzati [Mandelbrot, 1985; Mandelbrot, 1986]. Una ulteriore suddivisione differenzia i frattali in statisticamente auto-somiglianti, caratterizzati dalla stessa distribuzione statistica delle parti rispetto al tutto, e frattali statisticamente auto-affini, caratterizzati invece dalla medesima distribuzione delle parti in ragione dei differenti fattori di scala utilizzati [Falconer, 1990]. Molto spesso tale ulteriore suddivisione viene utilizzata per raggruppare i cosiddetti frattali casuali (*random fractals*), cioè quei frattali in cui viene introdotto un elemento casuale ad ogni passo di costruzione, producendo come risultato finale la medesima variabilità su tutte le scale [Falconer, 1990]. I concetti di auto-somiglianza e auto-affinità di scala sono ben noti in fisica come leggi di riscaldamento (*scaling laws*). Nel campo dei mezzi porosi, come si vedrà nel paragrafo (1.4), tale legge costituisce il *trait d'union* tra il contenuto d'acqua, la geometria del mezzo e il potenziale matriciale, il tutto relazionato attraverso l'invariante di scala definito dalla dimensione frattale.

La dimensione frattale contiene molte informazioni circa quelle che sono le proprietà geometriche dell'oggetto esaminato. In prima approssimazione si può affermare che essa è una misura dell'interruzione, o irregolarità, di una figura frattale quando questa viene osservata ad una determinata scala. Nel corso degli anni sono state formulate differenti definizioni della dimensione frattale le quali risultano essere tutte legate tra di loro; alcune hanno senso solo in specifiche situazioni; in certi casi le diverse definizioni coincidono, mentre in altre non tutte conducono allo stesso risultato numerico.

La dimensione frattale risulta essere di difficile determinazione nel senso rigoroso di *Hausdorff* (\dim_H , vedi par. 1.2.2). Nei casi reali essa viene invece stimata generalmente mediante la dimensione di *box-counting*, \dim_B , (*box dimension o box counting dimension*, vedi par. 1.2.3). Vediamo di seguito, quindi, cosa si intende per dimensione di *Hausdorff* e dimensione di *box-counting*, quando e per quali oggetti frattali possono essere determinate.

1.2.2 Dimensione di *Hausdorff*

In questo paragrafo verranno introdotti due concetti di fondamentale importanza per la comprensione dell'intera geometria frattale, noti rispettivamente come misura e dimensione di *Hausdorff*. Il concetto di misura è stato introdotto per la prima volta agli inizi del secolo da *Carathéodory* e ripreso in seguito dallo stesso *Hausdorff* con l'utilizzo degli insiemi di ricoprimento. La misura di un insieme S di punti nello spazio d -dimensionale euclideo E^d , con $S \subset E^d$, può essere effettuata ricoprendo l'insieme stesso con una funzione detta di prova, o *test function*:

$$h(r) = \gamma(d) \cdot r^d \quad (1.1)$$

in cui $\gamma(d)$ è un coefficiente geometrico che dipende dalla figura adottata per effettuare il ricoprimento, r rappresenta il lato della figura (cella o ipercubo) e d è la dimensione della misura. Detta Γ la funzione gamma di Eulero, il coefficiente $\gamma(d)$ può essere calcolato attraverso la relazione:

$$\gamma(d) = \frac{[\Gamma(\frac{1}{2})]^d}{\Gamma(1 + \frac{d}{2})} \quad (1.2)$$

Per ogni coefficiente $r > 0$ risulta possibile determinare la seguente sommatoria:

$$H_r^d(S) = \sum h(r) = \sum \gamma(d) \cdot r^d = N(r)\gamma(d)r^d \quad (1.3)$$

in cui $N(r)$ rappresenta il numero di figure dipendente dal passo scelto [Feder, 1988]. Sulla scorta di questa sommatoria, viene definita misura d -dimensionale di *Hausdorff* (H^d) dell'insieme S la relazione:

$$H^d(S) = \lim_{r \rightarrow 0} H_r^d(S) \quad (1.4)$$

Quest'ultima relazione gode di una proprietà importante nota come legge di scala delle dilatazioni, cioè se $S \subset \mathbb{R}^n$ e scelto un valore di $\lambda > 0$, vale la proporzione:

$$H^d(\lambda S) = \lambda^d H^d(S) \quad (1.5)$$

in cui $\lambda S = \{\lambda x : x \in S\}$ è l'insieme S riscalato di un fattore λ . E' facile riscontrare che la (1.4) varia con il lato dell'ipercubo. Quindi, per r tendente a zero, essa può assumere, in generale, valore nullo o infinito, secondo l'esponente d scelto. La dimensione di *Hausdorff* (definita anche come dimensione di *Hausdorff-Besicovitch*), indicata come $\dim_H S$, è quel valore critico di d della dimensione per cui la misura, per r tendente a zero, subisce il salto da zero e infinito:

$$H^d(S) = \lim_{r \rightarrow 0} H_r^d(S) = \lim_{r \rightarrow 0} N(r) \gamma(d) r^d = \begin{cases} 0, & \text{se } d > \dim_H S \\ \infty, & \text{se } d < \dim_H S \end{cases} \quad (1.6)$$

Con maggiore formalismo matematico si può scrivere:

$$\dim_H S = \inf\{d: H^d(S) = 0\} = \sup\{d: H^d(S) = \infty\} \quad (1.7)$$

Se $d = \dim_H S$, allora $H^d(S)$ può assumere rispettivamente i valori zero o infinito, oppure soddisfare la disequazione:

$$0 < H^d(S) < \infty \quad (1.8)$$

In questo caso, quando $\dim_H S$ assume valori non interi, la dimensione è detta frattale [Falconer, 1990]. E' possibile fornire adesso una definizione più rigorosa di frattale, data dallo stesso *Mandelbrot*, come quell'insieme la cui dimensione di *Hausdorff* \dim_H è strettamente maggiore della dimensione topologica (\dim_T) [Mandelbrot, 1987], dimensione, secondo la geometria euclidea, che risulta essere sempre un numero naturale intero non superiore a tre. La determinazione della dimensione frattale secondo *Hausdorff* (\dim_H), benchè storicamente

importante, è di scarso interesse pratico poiché nasconde in se serie difficoltà di calcolo in casi concreti (non deterministici). L'esperienza, che ha come origine i primi studi di *Pontrjagin* e *Schnirelman* [1932], ha permesso di individuare per la valutazione di \dim_H una tecnica numerica semplice basata sul concetto di ricoprimento dell'insieme, mediante celle, sfere o cubi e il successivo conteggio. Tale tecnica prende il nome di *box-counting* (conteggio di celle, appunto) e la dimensione con essa calcolata viene definita come dimensione di *box-counting*, \dim_B .

1.2.3 Dimensione di *box-counting* e algoritmi di risoluzione

Sia E un qualsiasi sottoinsieme limitato non vuoto di \mathbb{R}^d e sia $N_r(E)$ il più piccolo numero di insiemi di diametro pari al più a r con cui è possibile ricoprire E . Si definiscono, rispettivamente, le dimensioni inferiore (*lower*) e superiore (*upper*) di *box-counting* di E , ossia $\underline{\dim}_B E$, con le relazioni seguenti [Edgar, 1990; Edgar, 1998; Falconer, 1990; Falconer, 1997]:

$$\liminf_{r \rightarrow 0} \frac{\log N_r(E)}{\log \left(\frac{1}{r}\right)} \quad (1.9)$$

$$\limsup_{r \rightarrow 0} \frac{\log N_r(E)}{\log \left(\frac{1}{r}\right)} \quad (1.10)$$

Quando questi limiti sono uguali ed esistono, il loro valore rappresenta proprio la dimensione di *box-counting*; in tal caso la dimensione di *box-counting* di E , può essere indicata come:

$$\dim_B = \lim_{r \rightarrow 0} \frac{\log N_r(E)}{\log \left(\frac{1}{r}\right)} \quad (1.11)$$

Si può dimostrare [Falconer, 1990] che il limite così definito esiste e che si può prendere in modo equivalente $N_r(E)$ come numero di cubi della maglia di lato r intersecante E , essendo $N_r(E) \sim r^d$, in cui $d = \dim_B E$, in accordo alla (1.6) attraverso le relazioni [Falconer, 1990; Pascolini, 1992]:

$$N_r(E)r^d \rightarrow \begin{cases} 0, & \text{se } d > \dim_B E \\ \infty, & \text{se } d < \dim_B E \end{cases} \quad (1.12)$$

e

$$N_r(E)r^d \rightarrow \inf \left\{ \sum_{i=1}^{\infty} r^d : \{U_i\} \text{ ricoprimento di } E \text{ con } |U_i| \leq r \right\} \quad (1.13)$$

quest'ultima espressa in conformità ad una collezione di ricoprimenti aventi diametri $|U_i|$ minori o uguali a r . In definitiva, la relazione (1.11) viene definita nella maggior parte dei testi come dimensione di *box-counting*, sebbene in alcuni casi [Edgar, 1998; Pascolini, 1992] viene indicata con termini differenti e, cioè, di dimensione di capacità, dimensione metrica, densità logaritmica, dimensione di *Minkowski*.

Uno dei punti cruciali nella valutazione della dimensione di *Hausdorff*, e, quindi, della corretta stima della dimensione di un qualsiasi oggetto frattale, è costituito dalla somma dei termini $\sum_{i=1}^{\infty} |U_i|^d$ presenti nella relazione a partire dalla quale è possibile ricavare la d -misura di *Hausdorff* come processo limite per r tendente a zero [Falconer, 1990]:

$$H_r^d(S) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^d : \{U_i\} r - \text{ricoprimento di } S \right\} \quad (1.14)$$

qui di seguito riscritta in conformità all'espressione (1.3) e in accordo ai predetti diametri. Con la dimensione di *box-counting* (\dim_B) tale problema viene superato mediante la sostituzione dei termini diametrali con i termini r^d . Se consideriamo nuovamente l'esistenza del limite fornito dalla relazione (1.11), il numero $N_r(E)$ di cubi della maglia di lato r intersecante E (con $E \subset \mathbb{R}^d$) potrà essere rappresentato attraverso una delle seguenti definizioni [Falconer, 1990; Falconer, 1997]:

1. il più piccolo numero di sfere di lato chiuse di raggio r ricoprenti E ;
2. il più piccolo numero di cubi di lato r che ricoprono E ;
3. il numero di r maglie che intersecano E ;
4. il più piccolo numero di insieme con diametro pari al più a r che ricoprono E ;
5. il più grande numero di sfere disgiunte di raggio r con centri in E .

Nelle condizioni comuni si tende ad adottare la terza proprietà descritta poiché consente di superare le complessità algoritmiche intrinseche nel metodo.

A tal proposito, la dimensione di *box-counting* (1.11) può essere stimata mediante opportuni algoritmi di calcolo, basati essenzialmente sul conteggio delle celle intersecanti l'insieme da analizzare, nello spazio d-dimensionale, al variare del lato r. In generale, tali algoritmi vengono raggruppati nei cosiddetti algoritmi normalizzati o standardizzati (*Standard Algorithms*), cioè quelli la cui corrispondenza lineare tra conteggio delle celle e lato r avviene mediante la costruzione di *boxes* B del tipo [Falconer, 1990; Parker & Chua, 1989]:

$$B = \prod_{i=1}^d [l_i r, (l_i + 1)r[\quad (1.15)$$

in cui gli l_i sono costituiti da numeri interi. L'espressione (1.15) permette di costruire una sequenza lineare di maglie di lato r scalate fino a un prefissato valore soglia r_{\min} [Hunt & Sullivan, 1986]. Per la relazione (1.11) e (1.12), il legame che sussiste tra il numero N di r-maglie, intersecanti l'insieme oggetto di misura, risulta essere proporzionale ad una qualche potenza di r e, dunque, tale legame può essere espresso attraverso tale legge di potenza [Falconer, 1990; Feder, 1988]:

$$N(r) = k \cdot r^{-D} \quad (1.16)$$

in cui D e k rappresentano, rispettivamente, la dimensione frattale da stimare e una costante dipendente dalla geometria dell'insieme esaminato. Dalla (1.16) segue:

$$\log N(r) = \log k - D \log r \quad (1.17)$$

la dimensione frattale D può essere calcolata dalla pendenza della curva del diagramma doppio logaritmico di N(r) in funzione di r [Parker & Chua, 1989]. Di seguito vengono sintetizzati i passi necessari per la determinazione della dimensione frattale tramite *box-counting* in un caso piano:

1. sia assegnata una sequenza finita di N punti costituenti l'insieme oggetto di misura, di coordinate $\{(x_1, y_1), \dots, (x_N, y_N)\}$;

2. si sceglie r_{\min} ;
3. si sceglie una sequenza di maglie aventi rispettivamente celle di lato (r_1, \dots, r_N) ;
4. si ricopre di volta in volta l'insieme in esame con tale sequenza di maglia, contando il numero di celle in cui ricadono le coppie (x_i, y_i) ; si ottiene in tal modo la sequenza $N(r_1), \dots, N(r_{\min})$;
5. si rappresenta su un diagramma log-log l'insieme di punti di coordinate $\{(r_1, N(r_1)), \dots, (r_{\min}, N(r_{\min}))\}$;
6. si calcola la dimensione frattale di *box-counting* come coefficiente angolare della retta di regressione dei minimi quadrati che interpola i suddetti punti.

Altro algoritmo interessante e altamente efficiente dal punto di vista computazionale per la determinazione della dimensione di *box-counting* è quello proposto da *Hunt e Sullivan* [1986]. In generale, si può affermare che per frattali deterministici (ad esempio i frattali generati attraverso la procedura *Iterated Function System, IFS*), frattali, cioè, esatti, aventi la stessa dimensione in ogni parte, contrariamente ai frattali aleatori (*random fractal*), la cui dimensione può essere definita solo localmente e come media statistica, la dimensione di *Hausdorff* coincide con la dimensione di *box-counting* [*Falconer, 1990*], ossia:

$$dim_H = dim_B = D \quad (1.18)$$

1.2.4 Dimensione frattale per oggetti frattali deterministici

I frattali auto-somiglianti sono quelli più comuni e costruiti dalla ripetizione di un modello (generatore) su un oggetto di partenza (iniziatore). L'iniziatore determina la dimensionalità, mentre il generatore definisce la simmetria complessiva dell'oggetto e produce caratteristiche geometriche a differenti scale di lunghezza. Il modello generatore è ripetuto n volte e può dar luogo ad un aumento o ad una riduzione dell'iniziatore. L'oggetto frattale finale racchiude un intervallo di lunghezza pari a $r = 1/b^n$, in cui b è un fattore di riscaldamento mentre n varierà tra 1 ed n .

La barra di *Cantor*, il triangolo o tappeto di *Sierpiński* e la spugna di *Menger* [Mandelbrot, 1982] sono generati da questo processo (vedi fig. 1.1); i frattali sopra summenzionati appartengono alla categoria di frattali deterministici poiché la stessa operazione è ripetuta a tutte le scale. L'equazione fondamentale applicabile a tutti i frattali deterministici [Mandelbrot, 1982; Feder, 1988] è rappresentata, pertanto, dalle relazioni (1.16) e (1.18), in cui $N(r)$ è il numero di elementi di lunghezza uguale ad r , k il numero di iniziatori di lunghezza unitaria e D la dimensione frattale. Una definizione della dimensione frattale D segue dall'equazione (1.16):

$$D = \frac{(\log N(r) - \log k)}{\log\left(\frac{1}{r}\right)} \quad (1.19)$$

ovvero anche:

$$D = \frac{\log N\left(\frac{1}{b^{i+1}}\right) - \log N\left(\frac{1}{b^i}\right)}{\log(b)} \quad (1.20)$$

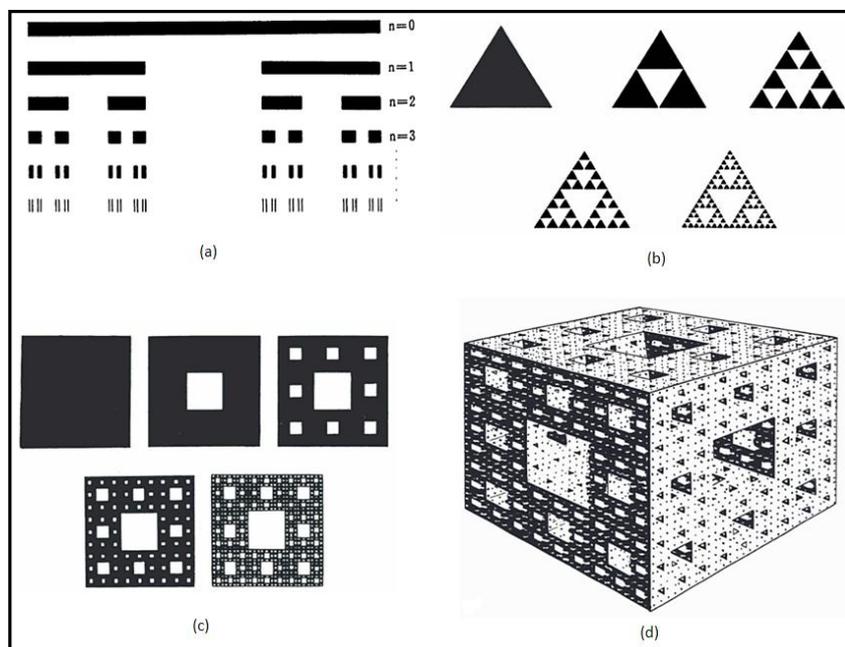


Figura 1.1 – Esempi di frattali deterministici: (a) barra di Cantor; (b) triangolo di Sierpiński; (c) tappeto di Sierpiński; (d) spugna di Menger [Xie, 1993]

Ponendo $k=1$, si può facilmente trovare la dimensione frattale degli oggetti frattali deterministici sopra indicati (tab. 1.1).

Oggetto frattale	N	r	D
Insieme di <i>Cantor</i>	2	3	0.630
Triangolo di <i>Sierpiński</i>	3	2	1.584
Tappeto di <i>Sierpiński</i>	8	3	1.892
Spugna di <i>Menger</i>	20	3	2.726

Tabella 1.1 – Proprietà degli oggetti frattali esaminati

1.3 Il terreno come mezzo poroso

Lo studio di molti problemi ambientali quali desertificazione, il ruscellamento superficiale, l'erosione, l'inquinamento delle falde acquifere, la tutela delle risorse idriche, la difesa dei versanti, sono regolati dai processi che si sviluppano nel suolo e vanno affrontati con una visione globale che tenga conto delle diverse interazioni di natura fisica, chimica e biologica. Pertanto, la comprensione delle dinamiche dei principali processi idrologici nel suolo rappresenta ancora oggi una delle più grandi sfide in quanto abbraccia diverse scale spaziali e temporali e coinvolge approcci disciplinari che comprendono l'idrologia, la geomorfologia, la geologia, l'agronomia, l'ecologia, la microbiologia del suolo [Eagleson, 1970]. Tra le emergenze ambientali più pressanti, quelle cui le ricerche e i recenti sviluppi nel campo dell'idrologia del suolo possono fornire risposte riguardano principalmente il problema della gestione sostenibile delle risorse idriche e del suolo. Le problematiche ambientali sono oggetto di studio dell'idrologia nel suolo in quanto legate a specifici processi del ciclo idrologico, ma soprattutto lo studio dei processi idrologici che evolvono nei primi strati di suolo è indispensabile per la comprensione e la previsione del degrado delle risorse naturali. Il suolo, infatti, per la sua posizione di contatto con l'atmosfera e per la presenza della vegetazione, svolge un ruolo particolarmente importante nel ciclo idrologico poiché i processi di moto dell'acqua che vi hanno sede regolano la partizione della pioggia in infiltrazione, il ruscellamento superficiale, i processi di trasporto dell'acqua e dei soluti nel mezzo poroso, l'evoluzione temporale e spaziale del contenuto idrico del suolo nello strato esplorato dalle radici [Kutilek & Nielsen, 1994]. Anche i processi di traspirazione delle piante e quelli di evaporazione dalla superficie del suolo sono strettamente legati al moto dell'acqua nel suolo. La risposta idrologica di un suolo è individuata attraverso le proprietà idrauliche del mezzo poroso, esprimibili con la funzione di ritenzione idrica e quella di conducibilità idraulica, che sintetizzano i complessi legami fra il contenuto di acqua (θ), il potenziale matriciale (h) e la conducibilità idraulica del suolo (K). La determinazione delle proprietà idrauliche costituisce una fase importante del complesso di indagini volte all'accertamento del rischio di degrado del suolo ed, eventualmente, all'individuazione di strategie di mitigazione del fenomeno.

1.3.1 Contenuto di acqua nel suolo

La misura del contenuto d'acqua nel suolo assume notevole importanza nell'idrologia, nell'ingegneria idraulica, in campo agricolo, meteorologico e nella meccanica dei suoli. L'umidità del suolo determina la sua temperatura e la partizione della radiazione netta in calore sensibile e latente, condiziona i processi di evapotraspirazione ed il moto dei soluti e degli inquinanti a seguito delle piogge e degli interventi irrigui. Numerose proprietà idrauliche (ritenzione idrica e conducibilità idraulica) e meccaniche del suolo (consistenza, plasticità, penetrabilità) dipendono dal contenuto d'acqua. Nei suoli argillosi i fenomeni di rigonfiamento e contrazione sono associati rispettivamente ad incrementi o decrementi del contenuto idrico. La conoscenza delle relazioni tra potenziale matriciale ed il contenuto idrico del suolo è necessaria in quanto interviene in numerose applicazioni, quali: stesura di bilanci idrologici, definizione delle proprietà idrauliche del suolo, valutazione della quantità di acqua disponibile per la vegetazione, valutazione del rischio di inquinamento o di asportazione dal suolo di composti chimici, ecc.

Il contenuto d'acqua del suolo in situ o in un campione di suolo da analizzare in laboratorio, può essere espresso sia in termini gravimetrici sia volumetrici. In termini gravimetrici, il contenuto d'acqua w si esprime come rapporto adimensionale tra la massa d'acqua presente nel suolo (M_w) e la massa di suolo secco (M_s) secondo la seguente espressione:

$$w = \frac{M_w}{M_s} \quad (1.21)$$

In termini volumetrici, invece, il contenuto d'acqua (θ) si esprime attraverso il rapporto adimensionale tra il volume occupato dall'acqua (V_w) nel campione di suolo ed il volume totale di suolo (V), secondo la seguente relazione:

$$\theta = \frac{V_w}{V} \quad (1.22)$$

Il contenuto d'acqua volumetrico θ ed il contenuto d'acqua gravimetrico w , possono essere correlati tra loro tramite la densità apparente del suolo (ρ_a) e la densità dell'acqua (ρ_w) mediante la seguente relazione:

$$\theta = w \frac{\rho_a}{\rho_w} \quad (1.23)$$

Il contenuto d'acqua massimo nel suolo, definito dalla condizione in cui tutti i pori sono pieni d'acqua, viene indicato come contenuto d'acqua in condizioni di saturazione (θ_s) ed in alcune applicazioni per la valutazione dell'acqua presente nel suolo, si fa riferimento al grado di saturazione (s) definito dal rapporto fra il contenuto d'acqua nel suolo e quello alla saturazione completa:

$$s = \frac{\theta}{\theta_s} \quad (1.24)$$

Il parametro dimensionale s viene di solito espresso in percentuale e varia da 0 a 100%. Tuttavia, come la condizione di saturazione completa ($s = 100\%$) è molto difficile da raggiungere, specie in campo ma pure in laboratorio, per la presenza sia di minuscole bolle d'aria sia di pori ciechi, anche l'assenza assoluta di acqua nel suolo è una situazione estrema altrettanto difficile da conseguire ($s = 0\%$). Se si indica con θ_r il contenuto d'acqua residuo nel suolo, si preferisce fare riferimento alla saturazione effettiva Θ :

$$\Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \quad (1.25)$$

che coincide con s solo se $\theta_r = 0$. Per soddisfare alle sempre più pressanti necessità di valutare il contenuto d'acqua nel suolo, sia saltuariamente che in modo continuo, sono stati proposti diversi metodi che utilizzano tecnologie sempre più sofisticate. Tali metodi possono raggrupparsi in due categorie: metodi diretti e metodi indiretti. I metodi diretti comportano il prelievo di campioni di suolo e la successiva estrazione e valutazione dell'acqua presente nel campione. Il loro impiego determina necessariamente la distruzione del campione per cui non è più possibile ripetere la misura nello stesso luogo. Il principale e più noto metodo diretto è quello termo-gravimetrico. Tuttavia, quando si devono valutare in campo le variazioni nel tempo del contenuto d'acqua, è necessario prelevare, a causa dell'eterogeneità, un numero notevole di campioni necessari per determinare ogni volta la varianza delle misure. Questa esigenza può rendere gravosa l'applicazione dei metodi diretti ed, inoltre, ripetute misure

possono determinare una inaccettabile distruzione del sistema sotto osservazione. I metodi indiretti prevedono la determinazione di alcune proprietà fisiche del suolo o chimico-fisiche che siano fortemente influenzate dal contenuto d'acqua. In genere questi metodi non prevedono la distruzione del campione ed utilizzano strumentazioni che è possibile installare anche in modo permanente nel suolo o sensori remoti localizzati su piattaforme (ad esempio satelliti). Si prestano bene per misure da ripetere nel tempo e consentono la registrazione automatica dei dati.

1.3.2 Zona insatura e potenziale dell'acqua nel suolo

Un terreno sciolto, da un punto di vista fisico e idrogeologico, è generalmente considerato come un sistema di tubi capillari costituiti da spazi non occupati dalla fase solida detti vuoti. Questi, in condizioni naturali, sono disponibili ad ospitare sia una fase liquida che una gassosa, acqua ed aria rispettivamente, rendendo il mezzo in condizioni naturali un perfetto sistema bifase. Parametri quali porosità e saturazione sono necessari per descrivere le caratteristiche di un mezzo e ottenere il profilo di distribuzione verticale dell'acqua naturalmente contenuta all'interno dello stesso. In tal senso è possibile discriminare all'interno di un suolo tre differenti settori ciascuno con un differente grado di saturazione compreso tra un valore minimo e massimo:

1. zona insatura o aerata in cui i pori tra i granuli di terreno sono parzialmente riempiti con acqua o da liquido che filtra lentamente verso il basso per gravità e il restante spazio dei pori libero dall'acqua è occupato dall'aria.
2. zona di transizione in cui la saturazione è intorno al 75%; essa è posta al centro della zona insatura e in essa le acque circolano solo verso la falda.
3. frangia capillare è la zona situata immediatamente al di sopra della zona satura (superficie piezometrica) ed ha con essa un legame stretto [Bear, 1972]. I pori sono parzialmente riempiti d'acqua ma la pressione della stessa è inferiore a quella atmosferica. La frangia capillare presenta spessori molto diversi a seconda della granulometria dei terreni:

qualche centimetro per le ghiaie fino a un paio di metri per i terreni argillosi. Il movimento dell'acqua è essenzialmente verticale con componente verso il basso dovuta a gravità e componente verso l'alto dovuta a principalmente alle forze capillari.

Alla base della frangia capillare è localizzata la zona satura o acquifero; in essa i pori tra i granuli di terreno sono tutti occupati dall'acqua di falda che si muove per moto proprio; la pressione dell'acqua aumenta con la profondità e il livello dove questa è uguale alla pressione atmosferica, rappresenta il livello della superficie freatica. La porzione di terreno così rappresentabile, ad esclusione della parte interessata dalla falda acquifera, va sotto il nome di zona insatura ed è nota come la regione compresa tra il piano campagna e la superficie piezometrica (fig. 1.2).

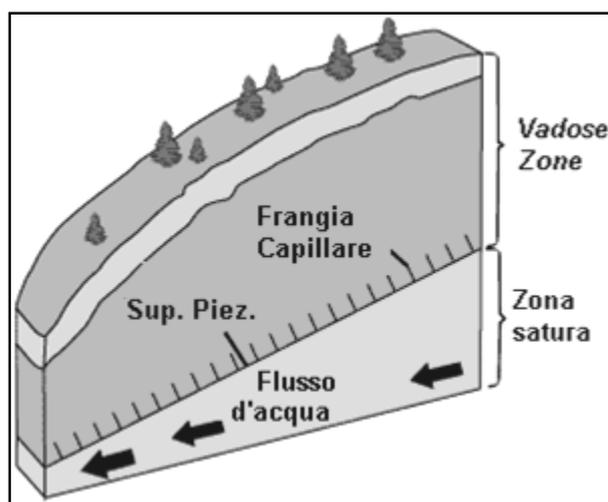


Figura 1.2 – Rappresentazione della zona insatura

L'acqua presente in un mezzo poroso non saturo, qual è il suolo, è soggetta ad una grande varietà di forze che agiscono secondo diverse direzioni. Il campo gravitazionale terrestre ed i sovraccarichi determinati dal peso degli strati sovrastanti ad ammassi porosi non completamente rigidi spingono l'acqua essenzialmente in direzione verticale. Le forze di attrazione fra le molecole dell'acqua e le superfici delle particelle solide e quelle che si determinano in corrispondenza dell'interfaccia di separazione acqua-aria agiscono invece in varie direzioni. Inoltre, gli ioni in soluzione esercitano azioni di attrazione che si oppongono al movimento dell'acqua. A causa delle difficoltà che si incontrano nel descrivere questo sistema complesso di

forze, per lo studio dei processi di moto nel terreno, si preferisce far riferimento piuttosto all'energia che si determina sull'unità di quantità d'acqua in conseguenza del campo di forze. Inoltre, per le modeste velocità in gioco che caratterizzano gli usuali fenomeni di filtrazione, si può senz'altro trascurare l'energia cinetica e considerare solo la componente potenziale dell'energia totale. Poiché ai fini del moto assume significato fisico la differenza di energia potenziale in due diversi punti del dominio del moto, i potenziali dell'acqua nel mezzo poroso non sono valutati mediante una scala assoluta dell'energia, bensì sono definiti in relazione ad una condizione di riferimento standard. Questa condizione viene usualmente fissata attraverso l'energia posseduta dalla quantità unitaria dell'acqua pura (senza soluti), libera, (contenuta in un ipotetico recipiente e soggetta alla sola forza di gravità), alla pressione atmosferica normale, alla stessa temperatura dell'acqua nel terreno (o ad un'altra temperatura da specificare) e ad una stabilita quota di riferimento. Pertanto, i movimenti dell'acqua risultano determinati da differenze di potenziale e il liquido si sposta da regioni caratterizzate da potenziali più elevati verso regioni con valori minori del potenziale; l'acqua è in condizioni di equilibrio nel mezzo poroso se il potenziale è costante in tutti i punti del sistema. Il concetto di potenziale dell'acqua in un mezzo poroso è di fondamentale e grande importanza per lo studio dei processi di trasporto e fornisce anche un modo unificato per poter valutare nel tempo lo stato dell'acqua in tutto il sistema continuo suolo-vegetazione-atmosfera. Questo concetto di potenziale ha sostituito le arbitrarie classificazioni proposte nel passato che facevano riferimento a diverse forme di acqua nel terreno quali l'acqua gravitazionale, l'acqua capillare, l'acqua igroscopica. Queste suddivisioni risultano inesatte e generano ambiguità in quanto tutta l'acqua presente nel mezzo poroso e non solo una parte, è soggetta al campo gravitazionale terrestre ed, inoltre, in condizioni di saturazione parziale, i fenomeni di capillarità sono presenti per ogni valore del contenuto d'acqua [Hillel, 1982]. Per poter considerare separatamente le diverse forze del campo è invalso l'uso di definire il potenziale con un approccio termodinamico considerandolo in termini di differenza di energia libera specifica tra l'acqua nel mezzo poroso e quella nella condizione di riferimento. Il Comitato sulla Terminologia della Società Internazionale di Scienza del Suolo (ISSS) [Bolt, 1976] ha definito il potenziale totale dell'acqua nel terreno come il lavoro, per unità di quantità di acqua pura, che bisogna compiere per trasportare in modo

reversibile ed isoterma una quantità infinitesima d'acqua da un recipiente contenente acqua pura, ad una prefissata quota ed alla pressione atmosferica, nell'acqua del terreno nel punto in esame. Questa definizione, anche se risulta puramente formale e non può essere utilizzata per una effettiva misura, consente però di considerare il potenziale totale come la somma di diverse componenti ciascuna delle quali si riferisce ad una trasformazione isoterma e reversibile che converte parzialmente le condizioni dell'acqua dalla situazione di riferimento a quella finale nel terreno. Con riferimento ai suggerimenti del Comitato sulla Terminologia della ISSS, il potenziale totale ψ_t viene suddiviso in tre componenti:

$$\psi_t = \psi_g + \psi_p + \psi_o \quad (1.26)$$

in cui i pedici g, p ed o denotano rispettivamente il potenziale gravitazionale, di pressione ed osmotico. I potenziali ψ_g e ψ_o si riferiscono agli effetti sullo stato energetico dell'acqua dovuti alle differenze di quota ed ai soluti disciolti, mentre il potenziale di pressione ψ_p comprende tutte le altre azioni che si esercitano sull'acqua nel mezzo poroso e tiene conto dei legami con la matrice solida, della curvatura dei menischi acqua aria, del peso dell'ammasso poroso sovrastante e, nel caso di saturazione, della pressione idrostatica. La definizione del potenziale di pressione ψ_p non è però di uso generale in quanto nell'Idrologia del Suolo tradizionalmente le variazioni di energia associate al trasporto dell'acqua dalla condizione di riferimento standard a quella nel mezzo poroso nel punto considerato sono state suddivise in altre componenti che considerano separatamente gli effetti della pressione della fase gassosa, del peso del terreno sovrastante, della pressione idrostatica e dei legami con la matrice solida. In particolare la componente del potenziale di pressione ψ_p che tiene conto delle forze di adsorbimento e di capillarità dovute all'affinità dell'acqua con la matrice del mezzo poroso, viene indicato come potenziale matriciale h. Le forze che attraggono e legano l'acqua nel terreno ne riducono l'energia potenziale al disotto di quella dell'acqua libera, pertanto h assume valori negativi o al più pari a zero che aumentano al ridursi dei legami con la matrice solida. Nel caso di mezzi porosi non rigonfiabili, per i quali quindi la matrice è rigida e sopporta interamente il carico dell'ammasso sovrastante, ed in presenza di una fase gassosa interconnessa a pressione atmosferica, il potenziale di pressione coincide con quello matriciale: $\psi_p = h$.

Il potenziale totale dell'acqua nel mezzo poroso e le sue componenti sono definiti con riferimento all'unità di acqua pura e le loro dimensioni variano a seconda che questa unità venga espressa in termini di una massa, un volume o un peso. Quando si fa riferimento all'unità di massa, nel S.I. il potenziale ha le dimensioni $[L^2 T^{-2}]$ e unità di misura in $J kg^{-1}$. Sebbene il riferimento alla massa appaia il più corretto, questa definizione di potenziale è usata in modo esteso solo in termodinamica in quanto per le applicazioni nel campo della Fisica del Suolo conduce spesso ad espressioni con termini più complessi. Poiché nella maggior parte delle situazioni di interesse applicativo si può assumere che l'acqua sia incompressibile e la sua densità praticamente indipendente dal potenziale totale, si può far riferimento nella definizione del potenziale all'unità di volume invece che all'unità di massa dell'acqua. Con questa definizione, il potenziale assume le dimensioni di una pressione $[M L^{-1} T^{-2}]$ e si esprime nel S.I. in $N m^{-2} = Pa$. Quando, infine, l'energia è riferita all'unità di peso dell'acqua, il potenziale ha unità di misura in $J N^{-1}$, o più semplicemente in m , e ha le dimensioni di una lunghezza $[L]$. Questo terzo modo di valutare i potenziali è considerato più conveniente e preferito per la sua semplicità e comporta le stesse approssimazioni impiegate nel riferire l'energia all'unità di volume d'acqua. Se si utilizza questa terza definizione, il potenziale totale viene più specificamente indicato con il simbolo H in analogia a quello che in Idraulica esprime l'energia specifica totale, e per la maggior parte delle situazioni in cui si possono trascurare gli effetti dei soluti viene impiegata la relazione:

$$H = z + h \quad (1.27)$$

in cui z è la quota del punto in esame rispetto al riferimento adottato (potenziale gravitazionale) ed h equivale approssimativamente al termine che in Idraulica esprime l'altezza piezometrica (potenziale matriciale). La misura del potenziale riveste un'importanza fondamentale per lo studio dei processi di trasporto dell'acqua nel suolo e per il controllo del regime idrico delle colture: a tal proposito, la letteratura fornisce un'ampia casistica di impiego delle diverse tecniche di misura a problemi applicativi [*Richards & Gardner, 1936; Talsma, 1977; Towner, 1981*].

1.3.3 Curva di ritenzione idrica nel suolo

Ogni suolo, in relazione all'organizzazione geometrica delle sue particelle, alla loro aggregazione, e ad altre caratteristiche chimiche e biologiche, è caratterizzata da una relazione funzionale tra $h(\theta)$, come mostrato in figura 1.3. La forma della curva di ritenzione dipende dalla dimensione dei pori e quindi dalla composizione granulometrica e dallo stato di addensamento del suolo. I suoli a tessitura grossolana che hanno pori interconnessi e di grandi dimensioni, sono caratterizzati da bassi valori di potenziale matriciale, e da una curva ripida nella zona di transizione. I suoli a tessitura fine, le cui particelle hanno elevata superficie specifica e quindi forti legami elettrochimici con le molecole d'acqua, sono caratterizzati da alti valore del potenziale matriciale e da una minore pendenza della curva di ritenzione nella zona di transizione. Si osserva che quando il potenziale matriciale è nullo, che presuppone la sola presenza della pressione atmosferica, il contenuto d'acqua del suolo eguaglia la porosità e rappresenta il contenuto d'acqua a saturazione. Se il contenuto d'acqua diminuisce, per esempio per effetto dell'evaporazione, dapprima la pressione diminuisce lentamente, poi a partire da un punto caratteristico chiamato *air entry point*, al di là del quale è possibile lo svuotamento dei pori più grossi.

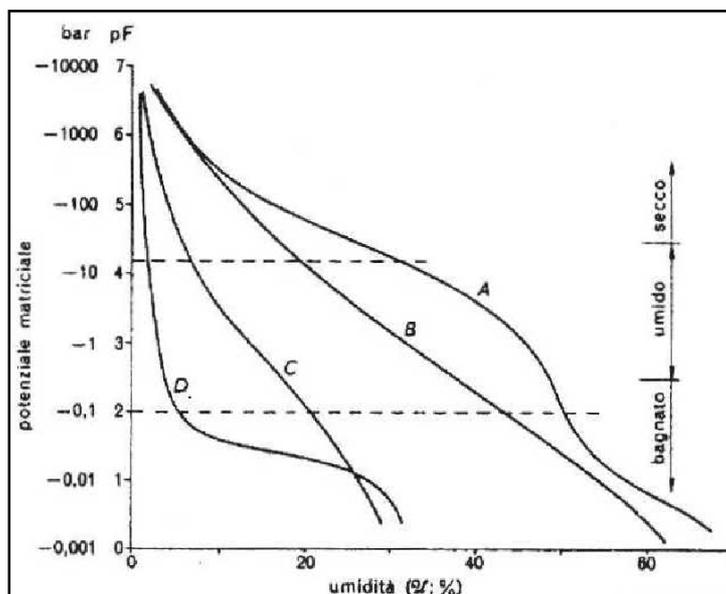


Figura 1.3 – Curve di ritenzione idrica per diverse tipologie di suoli: (A) argilla sabbiosa di pessima struttura; (B) terra rossa argillosa; (C) terreno sabbioso, ricco di parte fina; (D) sabbia grossolana [Cavazza, 1981]

Questo si nota in un grafico in scala semi-logaritmica (figura 1.4) nel quale la curva di ritenzione idrica può essere considerata come un braccio di iperbole con una inflessione in prossimità della saturazione, più marcata nel caso della curva di drenaggio. Per valori molto piccoli del contenuto d'acqua del suolo il potenziale tende ad un valore asintotico. Nella fase di infiltrazione d'acqua il processo segue un cammino diverso da quello seguito durante la fase di drenaggio (in cui si passa dallo stato saturo allo stato secco), come mostrato nel diagramma di figura 1.4. Entrambe le curve sono continue e si avvicinano in corrispondenza dei rami principali della curva caratteristica. Per un determinato valore di potenziale, quindi, corrispondono due valori di contenuto idrico, rispettivamente minore e maggiore, a seconda che si percorra la curva di capacità idrica oppure la curva di ritenzione idrica. Il caso di maggior interesse agronomico è rappresentato dalla curva di ritenzione idrica o di drenaggio.

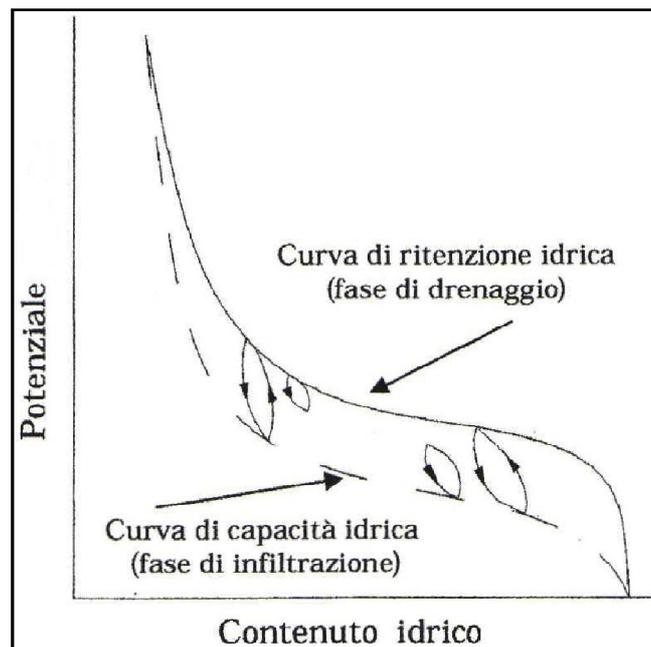


Figura 1.4 – Relazione tra potenziale e contenuto idrico. Sono indicate la curva di ritenzione idrica e la curva di capacità idrica

La forma di questa curva dipende fortemente dalla tessitura del suolo; infatti, all'aumentare del contenuto d'argilla, la curva di ritenzione trasla verso destra, mentre in suoli sabbiosi si sposta verso sinistra, come indicato in figura 1.5. In suoli sabbiosi i pori generalmente sono grandi ed

una volta svuotati, vi rimane solo una piccola quantità d'acqua; al contrario nei suoli argillosi la distribuzione volumetrica dei pori (di piccole dimensioni) è più uniforme e la quantità d'acqua che rimane è superiore. Quando un suolo umido comincia a drenare e successivamente è nuovamente irrigato, la relazione tra potenziale e contenuto idrico segue spostamenti intermedi, passando dal ramo di drenaggio a quello di irrigazione e viceversa.

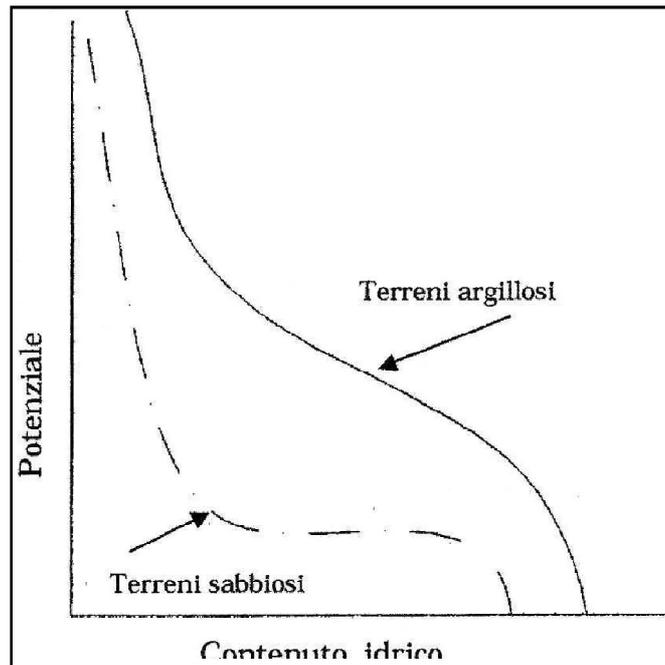


Figura 1.5 – Differenza tra la curva di ritenzione idrica per suoli sabbiosi e argillosi

La ciclicità degli eventi di precipitazione può causare dei cicli di isteresi come mostrato in figura 1.4. I cicli di isteresi sono dovuti a quattro cause concomitanti: l'effetto geometrico dei suoli chiamato anche effetto bottiglia di inchiostro; la variazione degli angoli di contatto; la presenza o meno di aria all'intero dei pori; le contrazioni volumetriche della matrice solida. L'effetto geometrico è causato dalla irregolarità dei pori nei suoli. Le forze capillari dipendono inversamente dal raggio di curvatura dei menischi che si vanno a formare al contatto tra l'acqua, l'aria ed il suolo. Questo raggio di curvatura dipende a sua volta dalla forma geometrica dei pori. All'interno del suolo sono presenti pori di dimensioni diverse, dislocati in diversi punti della matrice solida, e quindi le forze di capillarità sono di diversa intensità. In particolare, in presenza di variazioni di dimensione dei pori si ottiene l'effetto boccetta di inchiostro che è responsabile della formazione dei cicli di isteresi.

Quando il poro è riempito d'acqua il raggio di curvatura del menisco che si forma, indicato con r , è determinato dalle dimensioni del tubo di collegamento tra più pori che ha dimensioni minori del singolo poro. Quando invece il poro è vuoto, il raggio di curvatura del menisco è proporzionale al raggio del poro R , e quindi vale la relazione $R > r$. Essendo il potenziale capillare proporzionale al reciproco del raggio di curvatura, nel primo caso si avrà:

$$\psi_r = \frac{2T}{r} \quad (1.28)$$

mentre nel secondo caso si avrà:

$$\psi_R = \frac{2T}{R} \quad (1.29)$$

che rispetta la seguente disuguaglianza:

$$|\psi_r| > |\psi_R| \quad (1.30)$$

Il termine T rappresenta la tensione superficiale all'interfaccia acqua-aria ed il suo valore dipende dalla temperatura. Il segno negativo sta ad indicare la direzione della pressione generata dalla tensione tangenziale superficiale agente all'interfaccia acqua-aria che è diretta dalla parte concava del menisco, cioè verso l'esterno, come evidenziato nell'equazione di *Laplace*. Utilizzando la legge di distribuzione idrostatica della pressione riferita ad una superficie isobara ed uguagliata con l'equazione di *Laplace* si ottiene una formulazione per quantificare il valore della risalita capillare secondo la relazione:

$$h_c = \frac{\psi_{r,R} \cos \gamma}{(\gamma_{acqua} - \gamma_{aria})} \quad (1.31)$$

in cui h_c è la risalita capillare espressa in m; $\psi_{r,R}$ è il potenziale espresso in termini del raggio del poro R o del raggio del condotto r ; γ_{acqua} e γ_{aria} sono rispettivamente peso specifico dell'acqua e dell'aria; T è la tensione superficiale. Dal ciclo di isteresi si evidenzia come, a parità di contenuto d'acqua nel suolo, il lavoro necessario per la fase di essiccamento, fase in cui il poro si svuota, è maggiore rispetto al lavoro necessario per la fase di inumidimento, fase in cui il poro si riempie.

1.3.4 Cenni sul moto dell'acqua del suolo

Per la difficoltà di pervenire ad una esatta descrizione della reale geometria del mezzo poroso, la descrizione dettagliata del processo di moto in ogni punto del suolo è pressoché impossibile e anche di modesto interesse applicativo. Pertanto, si preferisce descrivere i processi di trasporto nel suolo su scala macroscopica facendo riferimento ad un mezzo poroso considerato continuo ed in cui le varie grandezze e le proprietà fisiche sono considerate funzioni della posizione e del tempo. Generalmente si fa riferimento a processi di moto isotermi, ad una fase liquida newtoniana e a una fase gassosa caratterizzata da una pressione pari a quella atmosferica. Inoltre, a causa delle basse resistenze al moto dell'aria, si trascura il movimento della fase gassosa e si fa riferimento solo al flusso dell'acqua. In questa sezione si rappresentano le principali equazioni che descrivono il flusso idrico nel suolo alla scala macroscopica. La legge di *Darcy* [Marshall et al., 1999] esprime il flusso d'acqua in un mezzo poroso; considerando un cilindro orizzontale di lunghezza L , contenente un certo volume di suolo connesso con due serbatoi mantenuti a diversi livelli costanti di acqua che determinano una differenza di carico Δh , il flusso della portata $Q = V/t$ viene semplicemente misurato dal flusso volumetrico d'acqua, V nel tempo t . La densità del flusso q è data da $q=V/(At)$ in cui A è l'area della sezione del cilindro di suolo attraversato dal flusso. Nel 1856 *Darcy* ha sperimentalmente dimostrato per colonne di suolo sabbioso una relazione lineare tra la densità di flusso idrico q ed il gradiente idraulico $\Delta h/L$:

$$q = K_s \frac{\Delta h}{L} \quad (1.32)$$

dove K_s rappresenta la conducibilità idraulica per terreni saturi. Nella tabella 1.2 viene riportata una generica classificazione dei suoli rispetto al valore della conducibilità idraulica.

La maggior parte dei processi di moto dell'acqua nel suolo non sono stazionari, poiché nel mezzo poroso il contenuto d'acqua varia nel tempo ($d\theta/dt \neq 0$). Per la corretta descrizione di tali processi l'equazione di *Darcy* deve essere accompagnata dall'equazione di continuità con l'assunzione che il mezzo poroso sia rigido. Il principio della conservazione della massa richiede che la variazione nel tempo della massa contenuta in un volume elementare di suolo debba

eguagliare la differenza tra la massa entrante e uscente attraverso la superficie di contorno del volume considerato (fig. 1.6).

Soil	k (m s ⁻¹)	Drainage
Gravel	10 ⁻³ -1	Good
Coarse sand	10 ⁻⁶ -10 ⁻²	Good
Medium sand	10 ⁻⁶ -10 ⁻³	Good
Fine sand	10 ⁻⁷ -10 ⁻⁴	Poor-good
Silty sand	10 ⁻⁷ -10 ⁻³	Poor-good
Silt, loess	10 ⁻⁹ -10 ⁻⁵	Poor-fair
Glacial till ^a	10 ⁻¹² -10 ⁻⁶	Very poor ^b -poor
Marine clay ^a	10 ⁻¹² -10 ⁻⁹	Very poor ^b -poor
Clay ^a	10 ⁻¹¹ -10 ⁻⁸	Very poor ^b -poor

Tabella 1.2 – Classificazione dei suoli rispetto alla conducibilità idraulica [Lade, 2001]

Si ottiene:

$$\frac{\partial(\rho_w \theta)}{\partial t} = -\nabla(\rho_w q) \quad (1.33)$$

in cui ρ_w è la densità dell'acqua e

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \quad (1.34)$$

è l'operatore gradiente.

Nel caso di liquido incomprimibile ($\rho = \text{costante}$) e scheletro solido indeformabile, si segue la schematizzazione di Richards introducendo la legge di *Darcy* nell'equazione di continuità e si ottiene la seguente equazione differenziale parziale che viene di solito indicata come equazione di *Richards*:

$$\frac{\partial(\rho_w \theta)}{\partial t} = -\nabla \cdot (\rho_w K \Delta \varphi) \quad (1.35)$$

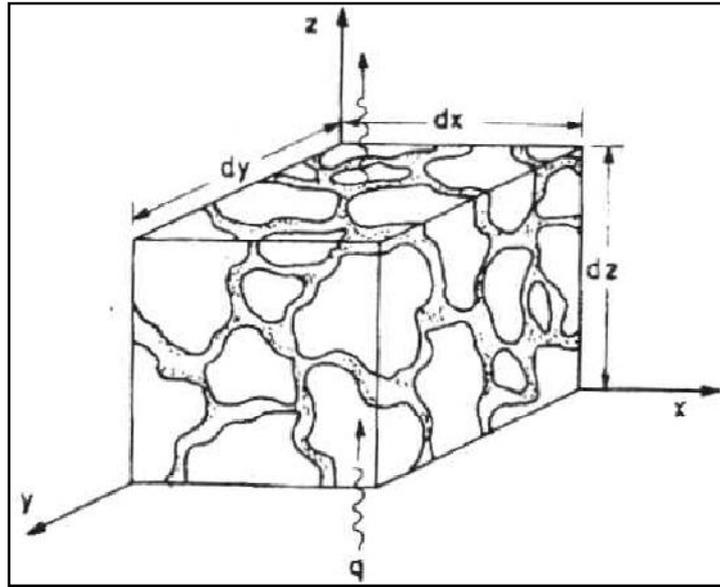


Figura 1.6 – Schema di volume elementare di matrice porosa dove q_x , q_y , q_z sono le componenti del vettore velocità di filtrazione nelle direzioni x , y , z del riferimento cartesiano

L'equazione che regge il moto dell'acqua nel suolo non saturo, in condizioni isoterme e in direzione verticale, considerando la massa volumica dell'acqua è l'equazione di *Richards*:

$$C(h) \frac{\partial h}{\partial t} = \frac{\partial \left[K(h) \left(\frac{\partial h}{\partial z} + 1 \right) \right]}{\partial z} \quad (1.36)$$

scritta nella sua versione unidimensionale, nella quale z rappresenta la quota considerata positiva verso l'alto, K è la conducibilità idraulica del suolo e $C(h)$ è la capacità idrica differenziale. I parametri presenti nell'equazione precedente sono la funzione di ritenzione idrica $h(\theta)$, e la funzione di conducibilità idraulica $K(\theta)$, che sintetizzano le complesse relazioni tra h , θ e K .

1.4 Modelli frattali per la caratterizzazione delle curve di ritenzione idrica (*WRC*)

Come precedentemente accennato, i frattali stanno diventando sempre più popolari nella ricerca fisica del terreno, come strumenti per caratterizzare le varie proprietà e le caratteristiche dei mezzi porosi, tra cui la ritenzione idrica o la variabilità della conducibilità idraulica. Un oggetto frattale, infatti, ha una struttura geometrica rappresentabile mediante una gerarchia tra le parti costituenti lo stesso, ovvero con le unità più grandi che contengono le unità più piccole, le quali includono a loro volta altre unità più piccole e così via procedendo fino all'infinito.

Facendo riferimento alle caratteristiche delle curve di ritenzione idrica del suolo, è ben noto che, sia il contenuto di acqua che il potenziale matriciale, dipendono fortemente dalla geometria del mezzo poroso. I modelli frattali studiati per la caratterizzazione delle *WRC* derivano da differenti lavori portati avanti per descrivere la struttura frattale del suolo in termini di distribuzione delle dimensioni delle particelle [Tyler & Wheatcraft, 1990], distribuzione del tipo di aggregati solidi [Perfect & Kay, 1991; Crawford et al., 1993], area dell'interfaccia poro-solido [de Gennes, 1985; Toledo et al., 1990], massa frattale della fase poro [Katz & Thompson, 1985; Bird et al., 1996] e massa frattale della fase solido [Rieu & Sposito, 1991a, b; Crawford et al., 1995]. È stato dimostrato, quindi, attraverso tali studi, che le proprietà delle strutture porose, misurate a diverse scale di risoluzione, sembrano spesso essere delle funzioni caratterizzate da una legge di potenza definita su una appropriata scala di lunghezza. L'esponente di questa legge di potenza può essere interpretato in termini di dimensione frattale non intera, la quale, a sua volta, può essere essa stessa associata alle principali caratteristiche strutturali del suolo. Essenzialmente, i vari modelli utilizzano come costruzione frattale di base il tappeto di *Sierpiński* oppure il suo duplicato tridimensionale rappresentato dalla spugna di *Menger* [Bird et al., 2000; Mandelbrot, 1982; Perrier et al., 1996; Toledo et al., 1990]. È stato dimostrato, inoltre, che essa risulta essere adatta alla descrizione del comportamento dello *scaling* frattale del suolo, sia in termini della fase solido che della fase poro di cui un mezzo poroso è costituito. Effettivamente, osservando lo stadio finale della spugna di *Menger*, per come riportato nella figura 1.7, si può osservare come essa esibisca una struttura dei vuoti e dei pieni tale da richiamare alla mente la geometria delle micro e macrostrutture riscontrabili nei mezzi porosi, quale è il terreno.

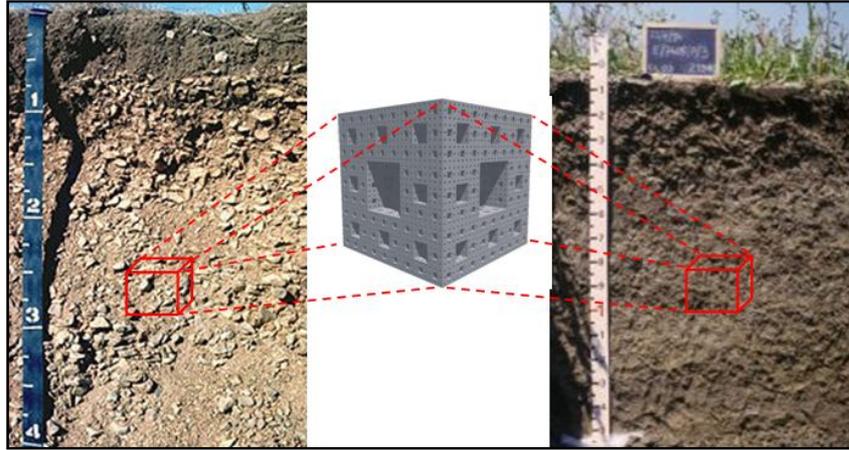


Figura 1.7 – Confronto tra spugna di Menger e geometria delle micro e macrostrutture riscontrabili nel mezzo poroso

I modelli frattali utilizzati nell'ambito di tale studio e che verranno descritti specificamente nel successivo paragrafo, sono tutti caratterizzati da un unico parametro geometrico che è, appunto, la dimensione frattale, essendo quest'ultima l'invariante di *scaling* nel processo di ritenzione idrica [Avnir *et al.*, 1998].

1.4.1 I modelli di Tyler & Wheatcraft [1990] e di Rieu & Sposito [1991]

I modelli di Tyler & Wheatcraft [1990] e di Rieu & Sposito [1991] descrivono la struttura porosa del suolo attraverso un frattale rappresentante una fase poro e una fase solido. In particolare, considerando la relazione tra le proprietà idrauliche del suolo e la geometria dei pori, nonché assumendo un modello di capillarità semplice, le WRC possono essere associate alla distribuzione della grandezza dei pori [Rieu & Perrier, 1998]. Sotto tali ipotesi, il contenuto di acqua in volume, θ , relazionato al potenziale matriciale generico, h , è assunto essere quello della porosità $\Phi \leq 1$, ossia per quei pori di grandezza più piccola di $l = \alpha/h$ (legge di Laplace), in cui α è una costante strumentale ricavabile dalle WRC.

I modelli di Tyler & Wheatcraft [1990] e di Rieu & Sposito [1991] possono essere ricavati sulla base della distribuzione dei pori di un frattale di massa, quest'ultimo inteso come aggregato solido. Tyler & Wheatcraft [1990] rappresentano la funzione idraulica delle curve di ritenzione

idrica usando l'algoritmo del tappeto di *Sierpiński*, il quale può essere esteso in un dominio di rappresentazione tridimensionale mediante la spugna di *Menger* (figura 1.8).

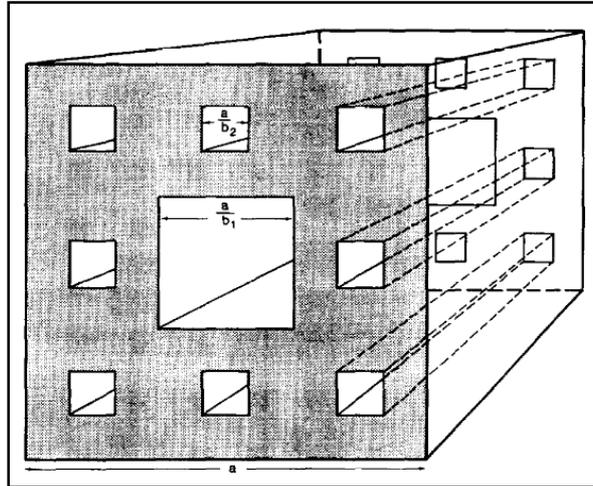


Figura 1.8 – Visione concettuale del tappeto di *Sierpiński* tridimensionale utilizzato per rappresentare un mezzo poroso [Tyler & Wheatcraft, 1990]

Tyler & Wheatcraft [1990], assumendo che non esiste alcun *cut-off* inferiore di scala, quando la grandezza del poro più piccolo l_{min} tende a 0, e, quindi, in considerazione della relazione:

$$[\Phi \leq l] = \Phi \left(\frac{l}{l_{max}} \right)^{E-D} \quad (1.37)$$

in cui l_{max} è la misura del poro più grande, E la dimensione di contenimento euclidea e D la dimensione frattale, attraverso l'uso diretto della legge di Laplace, essi pervengono alla seguente relazione:

$$\theta(h) = \left[\Phi \leq \frac{\alpha}{h} \right] = \Phi \left(\frac{h}{h_{min}} \right)^{D-E} \quad (1.38)$$

Nella (1.38) h_{min} rappresenta il valore del potenziale matriciale o carico di gorgogliamento [Danielson & Sutherland, 1986], ossia il valore di penetrazione dell'aria nel mezzo detto anche punto di bolla, inversamente proporzionale a l_{max} . In conclusione, l'espressione analitica del modello di Tyler & Wheatcraft [1990] è data dalla seguente equazione:

$$\theta(h) = \theta_s \left(\frac{h}{h_{min}} \right)^{D-E} \quad (1.39)$$

Tale relazione è equivalente a quella semplificata di *Brooks & Corey* [1964]. Nell'equazione (1.39), θ_s rappresenta il contenuto idrico a saturazione ed E risulta essere pari a 3 per un dominio euclideo tri-dimensionale.

Nel loro studio, *Rieu & Sposito* [1991] hanno sviluppato un modello di struttura del suolo sempre basato sulla spugna di *Menger*. Gli stessi Autori hanno, inoltre, introdotto il concetto di frammentazione incompleta (figura 1.9) per spiegare la coesione meccanica dei suoli.

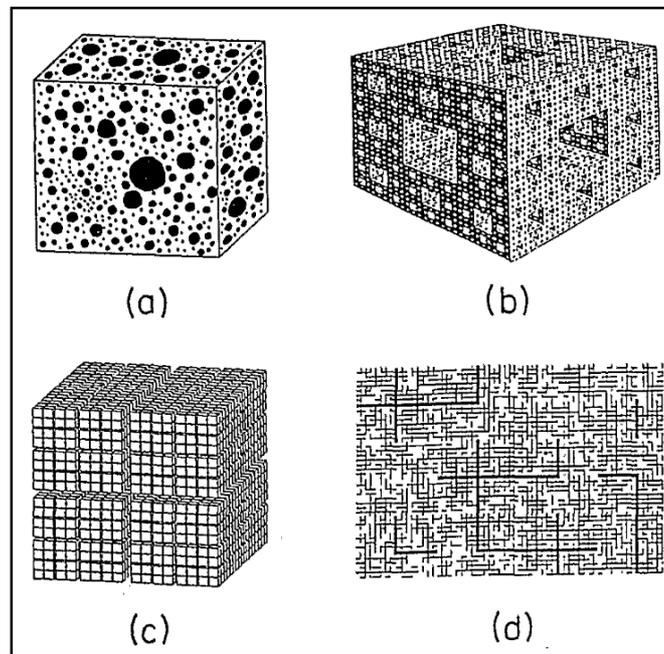


Figura 1.9 – Modelli concettuali di mezzo poroso: (a) mezzo poroso non necessariamente frattale; (b) spugna di Menger con $D=2.7268$; (c) mezzo poroso frattale completamente frammentato con $D=2.8$; (d) sezione trasversale di un mezzo poroso non completamente frammentato con $D=2.87$, che mostra un sistema di rete di fratture che non è attraversato da fratture più grandi [Rieu & Sposito, 1991a]

Pertanto, il loro modello per la caratterizzazione delle *WRC* si basa sull'assunzione che in un frattale di massa rappresentato da un modello geometrico bi-fase (poro e solido), esiste un limite di *cut-off* minimo finito della scala e questo risulta in corrispondenza di $l_{min} \neq 0$. In tal caso, il valore di $\Phi \leq l$ è equivalente alla relazione:

$$[\Phi \leq l] = \Phi_{max} - 1 + \left(\frac{l}{l_{max}}\right)^{E-D} \quad (1.40)$$

in cui Φ_{max} rappresenta la porosità a saturazione. Analogamente al caso precedente e considerando la legge di *Laplace*, si perviene alla seguente espressione:

$$\theta(h) = \left[\Phi \leq \frac{\alpha}{h}\right] = \Phi_{max} - 1 + \left(\frac{h}{h_{min}}\right)^{D-E} \quad (1.41)$$

In conclusione, *Rieu & Sposito* [1991a] hanno fornito per uno spazio di pori frattali la seguente espressione:

$$\theta(h) = \theta_s - 1 + \left(\frac{h}{h_{min}}\right)^{D-E} \quad (1.42)$$

1.4.2 Il modello *Pore-Solid Fractal* [*Bird et al.*, 2000]

Se si accetta che entrambe le fasi solido e poro presentano proprietà di auto-somiglianza, allora il terreno può essere modellato mediante la geometria frattale e descritto attraverso leggi di scala utilizzando la teoria dello *scaling* frattale [*Huang et al.*, 2006]. In tal senso *Bird et al.* [2000] hanno derivato un modello frattale, del tutto generale, atto a schematizzare sia la struttura del terreno che la simmetria fra la fase solido e la fase poro. Nello specifico, secondo procedure di carattere generale iterativo e descrittivo della spugna di *Menger*, gli stessi Autori hanno messo in relazione la distribuzione delle due fasi con il volume (o dimensione) della fase poro e della fase solido, mediante una legge di potenza. Tale legge definisce la funzione contenuto d'acqua nel terreno attraverso la relazione:

$$\theta(h) = \Phi - \frac{p}{p+s} \left[1 - \left(\frac{h}{h_{min}}\right)^{D-E} \right] \quad h_{min} \leq h \leq h_{max} \quad (1.43)$$

in cui Φ è la porosità totale, h_{min} e h_{max} rispettivamente i valori del potenziale nel momento in cui entra aria all'interno dei pori e quello necessario per drenare il più piccolo vuoto presente nel mezzo, p rappresenta la fase poro mentre s la fase solido. Ponendo E uguale a 3 per un dominio euclideo tri-dimensionale e definendo il parametro A come [*Wang et al.*, 2005]:

$$A = \frac{p}{p + s} \quad (1.44)$$

l'equazione (1.43) diventa:

$$\theta(h) = \Phi - A \left[1 - \left(\frac{h}{h_{min}} \right)^{D-3} \right] \quad (1.45)$$

Nel modello di *Bird et al.* [2000], gli Autori hanno calcolato la dimensione frattale come distribuzione cumulata della massa solida. Più nello specifico, l'espressione utilizzata è la seguente:

$$M_{s(i)} = L^3 \xi d_s \frac{s}{s + p} \left(\frac{r_{s(i)}}{r_{s(1)}} \right)^{3-D} \quad (1.46)$$

in cui $M_{s(i)}$ è la massa totale degli elementi solidi di taglia minori o uguali a $r_{s(i)}$, determinata dalla somma del volume dei solidi di taglia minori o uguale a $r_{s(i)}$ moltiplicato per la loro densità d_s ; $r_{s(1)}$ è il limite superiore della grandezza delle particelle, mentre l'indice i rappresenta il numero di iterazioni, L la lunghezza della regione studiata e ξ una costante tale che il prodotto ξL^3 uguaglia il volume di massa. Assumendo che tutti i pori del suolo possono essere completamente saturi [*Perrier et al.*, 1996], l'equazione (1.43) diviene [*Bird et al.*, 2000]:

$$\theta = (\theta_s - A) + A \left[\left(\frac{h}{h_{min}} \right)^{D-3} \right] \quad (1.47)$$

in cui θ_s è il contenuto idrico a saturazione ($\theta_s = \Phi$).

Nel modello di *Bird et al.* [2000], i valori limiti del parametro A , oggetto di studio e definizione nel campo frattale, sono rispettivamente il valore del contenuto idrico a saturazione θ_s ed il valore 1, ossia $\theta_s \leq A \leq 1$. È da osservare che i modelli di *Tyler & Wheatcraft* [1990] e di *Rieu & Sposito* [1991] costituiscono due casi limite simmetrici del più generale modello di *Bird et al.* [2000]. Infatti, assumendo il parametro $A = \theta_s$ (cioè quando la fase solido risulta preponderante), l'equazione (1.43) descrive una struttura che tende ad una massa frattale porosa. In questo caso, il modello *Pore-Solid Fractal* [2000], e, perciò, la funzione di ritenzione idrica, può essere espressa in termini del modello di *Tyler & Wheatcraft* [1990]. Analogamente, assumendo il

parametro $A = 1$ (cioè quando la fase solido tende a zero), l'equazione (1.43) si riconduce all'equazione del modello di *Rieu & Sposito* [1991]. Quindi le equazioni (1.39) e (1.42) possono essere considerati come casi specifici dell'equazione generale (1.43)

Capitolo 2 – La Geostatistica come studio della correlazione spaziale

2.1 Premessa

In ambito geologico-ambientale vi è spesso l'esigenza di conoscere come i valori di una proprietà fisico-chimica siano distribuiti entro un certo dominio spaziale. Tale compito è reso spesso difficile sia a causa della complessità ed eterogeneità dei processi fisico-chimici che si riflettono nella particolare distribuzione spaziale della proprietà studiata, sia per il carattere frammentario, eterogeneo, incerto e vago dell'informazione disponibile. Risulta necessario, quindi, introdurre metodi e modelli statistici e probabilistici per studiare parametri naturali campionati o conosciuti solo in parte.

A tal proposito, la Geostatistica rappresenta un importante strumento che fornisce le basi necessarie per capire e utilizzare varie tecniche per la stima del valore di una variabile spaziale nelle aree dove tale variabile non è stata misurata, ovvero è uno strumento per fare delle ipotesi il più possibile corrette sul continuo spaziale. La Geostatistica è nata nell'ambito delle scienze legate al suolo come la pedologia e l'agronomia, sebbene negli anni '50 del secolo scorso sia stata portata alla ribalta come metodologia applicata all'industria mineraria [Krige, 1951]. I principi teorici di questa recente disciplina sono stati sviluppati da *George Matheron* [1963; 1971] del *Centre de Morphologie Mathématique (Fontainebleau, Francia)*, che ha definito tale tecnica come teoria delle variabili regionalizzate. Essa si è evoluta e perfezionata nel settore dell'industria del petrolio già dagli anni settanta e le discipline oggi coinvolte dalle metodologie geostatistiche sono molteplici, la geologia, l'idrologia, l'oceanografia, la geochimica e tante altre, tutte discipline che si collocano in un contesto spaziale e/o spazio-temporale e che sono alla base

dello studio dei fenomeni ambientali che più frequentemente si manifestano sul territorio. Negli ultimi decenni sono stati pubblicati numerosi libri sugli aspetti teorici e sulle varie applicazioni della Geostatistica. Questi includono trattazioni classiche [Journel & Huijbregts, 1978], testi introduttivi [Clark, 1984; Hohn, 1988; Isaaks & Srivastava, 1989], testi con recenti sviluppi degli aspetti teorici e pratici dell'analisi spaziale dei dati [Cressie, 1991], della Geostatistica [Gooavaerts, 1997; Deutsch & Journel, 1998; Houlding, 2000] e della Geostatistica multivariata [Wackernagel, 1995].

La maggior parte dei metodi statistici permette soltanto di avere una visione qualitativa dei fenomeni naturali ed essi non utilizzano le informazioni spaziali che costituiscono la caratteristica fondamentale dei dati ambientali e delle scienze della terra. Il maggior contributo fornito dalla Geostatistica, rispetto alla statistica classica, è, quindi, quello di implementare tecniche opportune per studiare fenomeni naturali in cui bisogna tener conto della dipendenza spaziale tra le osservazioni [Oliver *et al.*, 1989]: questa situazione è definita autocorrelazione. Questo significa che le proprietà dell'ambiente sono in relazione fra di loro in una qualche scala, grande o piccola che sia. I valori campionati in luoghi vicini tra di loro, quindi, tendono ad avere comportamenti simili, mentre valori di una stessa variabile misurati in campioni raccolti in luoghi lontani tra di loro tendono ad avere comportamenti differenti [Tobler, 1970], o almeno tendono a differire dai valori medi che si riscontrano nei due luoghi stessi. In tal senso, la correlazione fra i valori della variabile tende a diminuire con l'aumentare della distanza. La Geostatistica, quindi, esprime quantitativamente il principio dell'autocorrelazione, utilizza tale proprietà per la predizione (*estimation*) della variabile in punti non campionati minimizzando l'errore legato proprio a queste stime. Questo capitolo contiene inizialmente una breve ma fondamentale descrizione dell'interpretazione dei dati spaziali che si hanno a disposizione, anche in senso statistico; successivamente sarà descritto in tutte le sue sfaccettature il mezzo principale utilizzato in Geostatistica per la descrizione della correlazione spaziale: il variogramma.

2.2 La distribuzione dei punti campionati: analisi del vicinaggio

Con il termine analisi del vicinaggio si intende una serie di tecniche che, partendo dalla localizzazione geografica dei punti-eventi, mirano a stabilire se la loro distribuzione spaziale è di tipo casuale, clusterizzata (processo aggregativo) o dispersa (processo repulsivo) (fig.2.1).

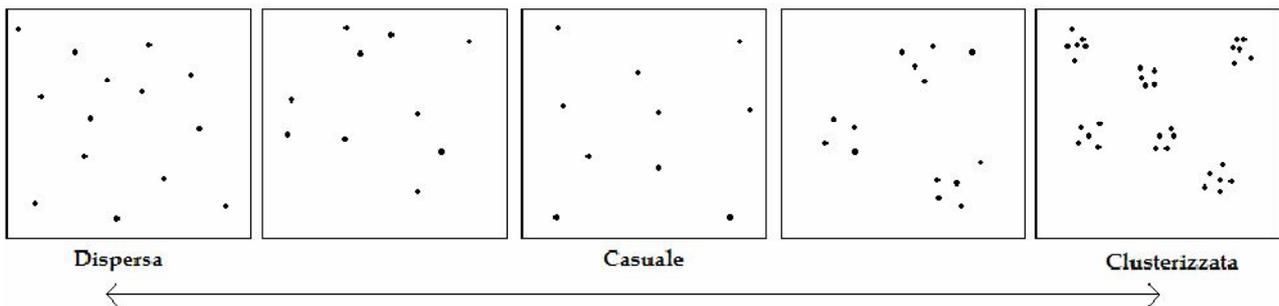


Figura 2.1 – Tipologie di configurazione spaziale dei punti

La logica di questo tipo di analisi si basa sullo studio della similarità tra distanze che possono essere calcolate tra:

1. ogni punto-evento ed il suo vicino più prossimo;
2. i diversi punti-evento e una località prefissata.

In letteratura sono note diverse procedure per l'analisi del vicinaggio, come la *Nearest Neighbor Analysis (NNA)* e alle sue generalizzazioni quali il *K-Order Nearest Neighbor Index* di ordine k (NNI_k) [Chang, 2002], oppure la *K-function*, messa a punto da Ripley [1981]. Nel presente lavoro si prenderà in considerazione la prima metodologia, di seguito illustrata.

2.2.1 *Nearest Neighbor Analysis (NNA)*

L'impianto teorico di questa tecnica è quello tipico delle procedure statistiche; essa consiste essenzialmente nel confrontare la media delle distanze tra vicini, calcolata sui dati osservati, con la media calcolata su una distribuzione casuale dei punti-evento. Se la media delle distanze tra

vicini è maggiore della distanza media teorica, si può affermare che la configurazione spaziale dei punti-evento tende ad essere dispersa, per cui il processo è di tipo repulsivo-inibitorio. Al contrario, se la media delle distanze osservate è minore di quella casuale, il processo è di tipo aggregativo, per cui i punti-evento tendono a raggrupparsi. Per stabilire, infine, se eventuali differenze tra le distanze osservate e quelle teoriche sono statisticamente significative si ricorre ad un apposito test probabilistico.

Indicando con $\min(d_{ij})$ la distanza minima tra i punti-evento, la media $d(NN)$ di queste distanze può essere così definita:

$$d(NN) = \sum_{i=1}^N \frac{\min(d_{ij})}{N} \quad (2.1)$$

In uno spazio omogeneo, e sotto la condizione d'isotropia (la densità media dei punti è uguale in ogni località dell'area presa in esame), la distanza media sotto l'ipotesi di casualità è:

$$d(att) = 0.5 \sqrt{\frac{A}{N}} \quad (2.2)$$

in cui A è l'area del dominio spaziale di interesse ed N il numero di eventi distribuzione. Dal rapporto tra le due medie si ottiene il *Nearest Neighbor Index (NNI)*:

$$NNI = \frac{d(NN)}{d(att)} \quad (2.3)$$

Il termine al numeratore rappresenta la media della distanza minima tra gli N punti-evento; il termine al denominatore è il valore atteso della medesima grandezza quando si considerino gli eventi distribuiti in maniera puramente casuale. Per cui se NNI è minore di uno la distanza media osservata è inferiore rispetto a quella attesa, per cui il processo tende all'aggregazione dei punti-evento, quindi gli eventi sono più vicini tra loro di quanto ci si aspetterebbe. Se NNI è maggiore di uno la distanza media osservata è maggiore di quella attesa per cui il processo è inibitorio e quindi i punti-evento tendono a disperdersi sul territorio.

2.3 Caratterizzazione dei processi spaziali

2.3.1 Variabile regionalizzata

Una variabile regionalizzata è la realizzazione di un processo stocastico, cioè di un insieme di variabili casuali corrispondenti ai punti di un dominio D oggetto di studio. Dunque, per ogni punto u in uno spazio d -dimensionale, il valore del parametro di interesse $z(u)$ è una delle realizzazioni del processo $Z(u)$. *Matheron* [1971] ha definito le variabili spaziali come variabili regionalizzate per descrivere la loro natura, espressa da due aspetti complementari: uno casuale, che tiene conto di irregolarità locali presentate dal fenomeno (piccola scala), e un aspetto strutturale, che riflette le caratteristiche globali dello stesso (larga scala). Assegnata una generica variabile regionalizzata Z , il suo modello probabilistico fondamentale può essere espresso nel seguente modo:

$$Z(u) = m(u) + Y(u) \quad u \in D \quad (2.4)$$

in cui D rappresenta il dominio oggetto di indagine, u è una qualsiasi localizzazione nel dominio D , il termine $m(u)$, denominato *trend*, rappresenta la componente deterministica (in tal caso lineare), $Y(u)$ è la componente casuale. In figura 2.2 è riportata una rappresentazione grafica del profilo di una variabile regionalizzata e delle sue componenti.

Analizzando l'espressione (2.4), è evidente che, se si trascura la componente casuale, il modello risulta essere deterministico, ovvero la distribuzione spaziale del fenomeno è univocamente ed esattamente descritta dalla funzione *trend*. D'altra parte, l'informazione contenuta nella componente residua $Y(u)$ del modello (2.4), racchiude caratteristiche fondamentali dei dati osservati. In conclusione, è fondamentale sottolineare che uno degli obiettivi dell'analisi Geostatistica è quello di identificare le due componenti suddette per risolvere il problema della stima di una variabile regionalizzata, a partire dall'informazione, seppur frammentaria, contenuta nel campione osservato. La componente strutturale, se esistente, viene solitamente individuata ricorrendo al metodo dei minimi quadrati [*Watson*, 1971].

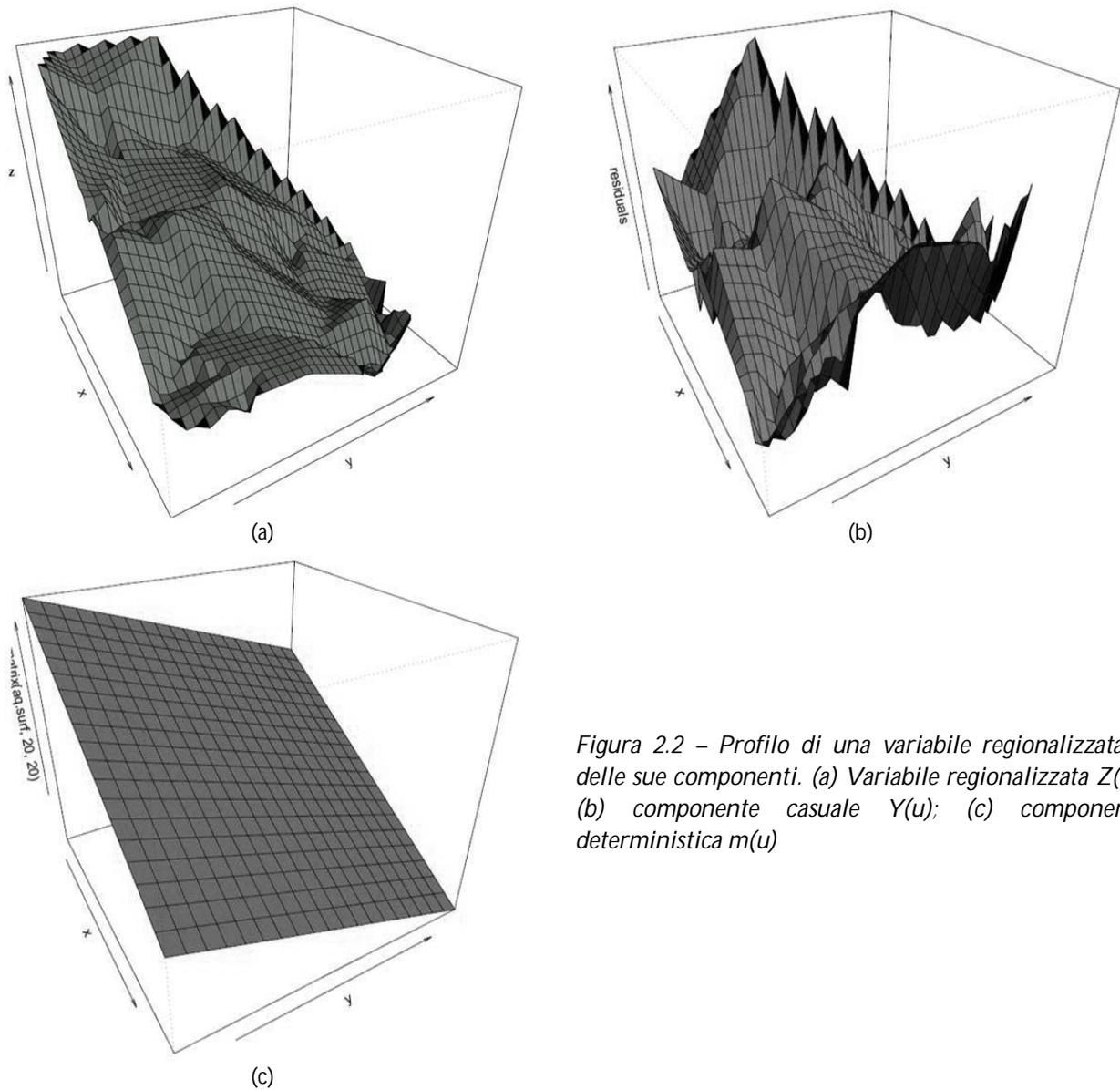


Figura 2.2 – Profilo di una variabile regionalizzata e delle sue componenti. (a) Variabile regionalizzata $Z(u)$; (b) componente casuale $Y(u)$; (c) componente deterministica $m(u)$

2.3.2 Funzioni aleatorie

La necessità di modellare la distribuzione spaziale della variabile Z su un dominio D deriva dall'impossibilità di disporre di un'informazione esaustiva sul territorio. In tale contesto, il ricorso alla teoria delle funzioni aleatorie consente di riconoscere la componente di incertezza che caratterizza la distribuzione dei fenomeni spaziali, per loro natura eterogenei e complessi. In altri termini, mediante il concetto di funzione aleatoria, è possibile associare, ad ogni punto u

del dominio D di interesse, una variabile aleatoria, indicata con $Z(u)$, in grado di descrivere le possibili manifestazioni del fenomeno in questione.

L'insieme delle variabili aleatorie, idoneo ad interpretare il comportamento di un fenomeno spaziale nell'ambito di una regione, è denominato funzione aleatoria o campo aleatorio ed è semplicemente indicato con Z oppure con:

$$\{Z(u), \quad u \in D\} \quad (2.5)$$

in cui D risulta essere un sottoinsieme dello spazio euclideo \mathbb{R}^d di dimensione d , con $d \in \mathbb{N}_+$ ($D \subseteq \mathbb{R}^d$). In letteratura, se $D \subseteq \mathbb{R}$, allora la suddetta collezione di variabili aleatorie viene denominata processo stocastico; d'altra parte, se $D \subseteq \mathbb{R}^d$ ($d \geq 2$), è preferibile la denominazione di campo aleatorio. Da un punto di vista geometrico, è ben noto che, assegnato un riferimento cartesiano in uno spazio euclideo di dimensione d , ogni punto u di tale spazio è identificato da una d -pla ordinata di numeri reali, denominati coordinate cartesiane di u . Evidentemente, se $D \subseteq \mathbb{R}^2$, ogni suo punto è individuato da una coppia ordinata di coordinate, ovvero $u_i = (x_i, y_i)$ rappresenta una localizzazione dello spazio D , e, conseguentemente, $z(u_i)$ il valore osservato in u_i . È ben noto che, una generica variabile aleatoria, è completamente caratterizzata dalla propria funzione di ripartizione, analogamente una funzione aleatoria è univocamente individuata dalla legge di probabilità congiunta delle variabili aleatorie che compongono il campo aleatorio stesso. In particolare, è interessante introdurre l'insieme delle distribuzioni finito-dimensionali della funzione aleatoria Z , ovvero l'insieme delle funzioni di ripartizione:

$$F_{u_1, u_2, \dots, u_n}(z_1, z_2, \dots, z_n) = P[Z(u_1) \leq z_1, Z(u_2) \leq z_2, \dots, Z(u_n) \leq z_n] \quad (2.6)$$

definite per ogni valore di n e per ogni possibile scelta dei punti di supporto u_1, u_2, \dots, u_n in \mathbb{R}^d . In tale contesto, le distribuzioni di probabilità congiunta forniscono un'informazione fondamentale sulla probabilità che si osservi uno specifico campione $z(u_1), z(u_2), \dots, z(u_n)$, per le variabili aleatorie, rispettivamente, $Z(u_1), Z(u_2), \dots, Z(u_n)$, dipendenti tra loro. In molte applicazioni, la conoscenza di tutte le distribuzioni fisico-dimensionali non viene richiesta, dal momento che i primi due momenti di tali distribuzioni sono sufficienti a fornire una soluzione approssimata accettabile per gran parte dei problemi applicativi [Posa & De Iaco, 2009].

2.3.3 Momenti di primo e secondo ordine

Assegnato un campo aleatorio Z , su un dominio D , è possibile definire i momenti del primo e del secondo ordine.

Il valore atteso (o momento del primo ordine) del campo aleatorio Z , se esiste finito, coincide con la componente strutturale del modello fondamentale (2.4), ovvero:

$$E[Z(u)] = m(u) \quad u \in D \quad (2.7)$$

essendo $E[Y(u)] = 0$.

Il valore atteso del campo aleatorio Z è dunque una funzione della posizione spaziale u e descrive le variazioni sistematiche di Z . Se la funzione $m(u)$ non risulta essere costante sul dominio D , è possibile affermare che il fenomeno descritto dal campo aleatorio Z presenta un *trend*.

In Geostatistica, i momenti del secondo ordine risultano essere particolarmente utili, dal momento che uno dei metodi di stima più utilizzato, ossia il metodo del *Kriging*, richiede che tali momenti esistano finiti e siano noti.

La varianza del campo aleatorio Z , se esiste finita, viene definita come segue:

$$Var[Z(u)] = E[Z(u) - m(u)]^2 \quad u \in D \quad (2.8)$$

Essa, denotata anche con $\sigma^2(u)$, dipende in generale dal punto u . Come sarà specificato in seguito, tale valore caratteristico esiste ed è costante per ogni u del dominio D nel caso di campi aleatori stazionari del secondo ordine.

Se il campo aleatorio Z ammette varianza, allora esso ammette anche covarianza. Quest'ultima, denominata covariogramma, è definita come segue:

$$Cov[Z(u), Z(u')] = E\{[Z(u) - m(u)] \cdot [Z(u') - m(u')]\} \quad u \in D \quad (2.9)$$

con u e $u' \in D$ o equivalentemente,

$$Cov[Z(u), Z(u')] = E[Z(u)Z(u') - m(u)m(u')] \quad u \in D \quad (2.10)$$

posto che $E[|Z(u)Z(u')|]$ esista finito. Il covariogramma è quindi funzione dei punti u e u' del dominio. È facile verificare che se la varianza esiste finita in ogni punto del dominio, la

covarianza di Z è sempre definita per ogni coppia di punti. Infatti, dalla disuguaglianza di *Cauchy-Schwartz* [Chow, 2000], segue immediatamente che:

$$E[|Z(u)Z(u')|] \leq E[Z^2(u)]^{1/2} E[Z^2(u')]^{1/2} \quad (2.11)$$

per cui, se la varianza è sempre ben definita, ovvero:

$$E[Z^2(u)] < +\infty \quad \forall u \in D \quad (2.12)$$

allora risulta:

$$E[|Z(u)Z(u')|] < +\infty \quad (2.13)$$

Se la varianza degli incrementi $[Z(u) - Z(u')]$ del campo aleatorio Z , con $u, u' \in D$, esiste finita, tale momento, indicato con 2γ , è denominato variogramma:

$$2\gamma[Z(u), Z(u')] = \text{Var}[Z(u) - Z(u')] \quad (2.14)$$

con $u, u' \in D$. Il variogramma, denotato anche con $2\gamma(u, u')$, è funzione dei punti u ed u' del dominio. È opportuno sottolineare che l'esistenza del variogramma non implica l'esistenza della varianza e, quindi, del covariogramma. La funzione γ è denominata semivariogramma. In seguito si utilizzerà indifferentemente il termine variogramma o semivariogramma, poiché, a meno di un fattore di proporzionalità pari a 2, essi sono equivalenti.

2.3.4 Le ipotesi di stazionarietà

I momenti di un campo aleatorio sono solitamente incogniti; pertanto, si presenta spesso il problema inferenziale riguardante la loro stima. In particolare, con riferimento ai momenti del secondo ordine, risulta evidente dalle definizioni che le funzioni variogramma e covariogramma dipendono dai punti di supporto u ed u' ; di conseguenza, sarebbero necessarie diverse osservazioni della coppia di variabili aleatorie $[Z(u), Z(u')]$ per effettuare inferenza statistica su tali momenti del secondo ordine. Tuttavia, per la caratteristica di non ripetitività dei dati a struttura spaziale, risulta impossibile disporre di più osservazioni nella stessa locazione [Posa, 1995].

Le ipotesi di stazionarietà consentono di risolvere il suddetto problema di natura inferenziale. Infatti, assumendo la cosiddetta stazionarietà del secondo ordine, è possibile supporre che le funzioni covariogramma e variogramma non dipendano dagli specifici punti di supporto u ed u' , ma solamente dalla distanza tra i punti di supporto, o meglio, dal vettore $h = (u - u')$. Pertanto, la stima di tali momenti del secondo ordine, può essere agevolmente ottenuta individuando, all'interno dello schema di campionamento spaziale, le diverse coppie di valori osservati in localizzazioni caratterizzate dallo stesso vettore di separazione h . In altri termini, se si ipotizza la stazionarietà del secondo ordine, ogni coppia di valori, corrispondenti a localizzazioni il cui vettore di separazione è pari ad h , può essere considerata una diversa realizzazione della coppia di variabili aleatorie $[Z(u), Z(u')]$.

Una funzione aleatoria Z risulta essere stazionaria del secondo ordine [Kitanidis, 1997] se il valore atteso della variabile aleatoria $Z(u)$ esiste e non dipende dal punto u :

$$E[Z(u)] = m \quad u \in D \quad (2.15)$$

e se, per ogni coppia di variabili aleatorie $[Z(u), Z(u+h)]$, il covariogramma esiste e dipende solo dal vettore di separazione h :

$$Cov(h) = E\{[Z(u+h) - m] \cdot [Z(u) - m]\} \quad u, u+h \in D \quad (2.16)$$

o equivalentemente:

$$Cov(h) = E[Z(u+h)Z(u)] - m^2 \quad (2.17)$$

Di seguito verranno elencate le conseguenze dell'ipotesi di stazionarietà del secondo ordine.

La stazionarietà del variogramma implica la stazionarietà della varianza e del variogramma, per cui è opportuno evidenziare quanto segue:

1. la varianza della variabile regionalizzata $Z(u)$ corrisponde al covariogramma valutato nell'origine, ovvero:

$$Var[Z(u)] = E[Z(u) - m]^2 = Cov(0) \quad u \in D \quad (2.18)$$

Infatti, richiamando la (2.16) e ponendo $h=0$, risulta:

$$Cov(0) = E\{[Z(u + 0) - m] \cdot [Z(u) - m]\} = E[Z(u) - m]^2 \quad (2.19)$$

Per cui, se la funzione aleatoria è stazionaria del secondo ordine, sia il valore atteso che la varianza non dipendono dalla localizzazione $u \in D$.

2. Per ogni coppia di variabili aleatorie $[Z(u), Z(u+h)]$, il variogramma esiste e dipende solo dal vettore di separazione h :

$$2\gamma(h) = Var[Z(u + h) - Z(u)] \quad (2.20)$$

o equivalentemente:

$$2\gamma(h) = E[Z(u + h) - Z(u)]^2 \quad (2.21)$$

con $u, u+h \in D$.

3. Il semivariogramma e il covariogramma sono legati dalla seguente relazione:

$$\gamma(h) = Cov(0) - Cov(h) \quad (2.22)$$

cioè covariogramma e semivariogramma sono due strutture equivalenti per caratterizzare la correlazione spaziale di un campo aleatorio stazionario del secondo ordine.

Esistono molti fenomeni fisici per i quali non esiste né la varianza, né il covarianza: questi fenomeni, tuttavia, ammettono un variogramma. Per cui, in tali situazioni empiriche, può essere opportuno assumere solamente l'esistenza e la stazionarietà del variogramma, ovvero indebolire le ipotesi che caratterizzano la stazionarietà del secondo ordine. Le ipotesi intrinseche, presentate di seguito, rispondono a tale esigenza.

Una funzione Z risulta essere intrinseca se, per ogni coppia di variabili aleatorie $[Z(u), Z(u+h)]$, il valore atteso degli incrementi $[Z(u+h)-Z(u)]$ esiste e non dipende dal punto u :

$$E[Z(u + h) - Z(u)] = 0 \quad u, u + h \in D \quad (2.23)$$

o equivalentemente:

$$E[Z(u)] = m \quad u \in D \quad (2.24)$$

e se il variogramma esiste e dipende solo dal vettore h di separazione, ossia:

$$2\gamma(h) = \text{Var}[Z(u+h) - Z(u)] \quad u, u+h \in D \quad (2.25)$$

In altri termini per ogni vettore h , gli incrementi $[Z(u+h)-Z(u)]$ ammettono una varianza finita che non dipende dal punto u . Per cui valgono le espressioni (2.20) e (2.21).

La stazionarietà del secondo ordine implica l'esistenza delle ipotesi intrinseche, ma non è vero il viceversa. In altri termini, la famiglia delle funzioni aleatorie che soddisfano le ipotesi intrinseche include quella delle funzioni aleatorie stazionarie del secondo ordine.

2.4 Analisi esplorativa dei dati (*Exploratory Data Analysis, EDA*)

Prima di intraprendere il trattamento geostatistico di un qualunque insieme di dati, è conveniente familiarizzare tanto con la natura fisica del fenomeno quanto con i dati osservati. Tale fase preliminare è particolarmente importante dal momento che un'attenta validazione dei dati e, quindi, un'accurata analisi descrittiva degli stessi, può evidenziare caratteristiche varie del campione ma, in modo particolare, la presenza di errori sistematici o materiali dei dati, che, se non opportunamente corretti o eliminati, possono riflettersi in ogni fase successiva dell'analisi, inficiando i risultati finali. Al fine di disporre di una visione del fenomeno spaziale non distorta ed eliminare ogni possibile errore, è quindi opportuno eseguire un'attenta analisi esplorativa dei dati disponibili mediante un'analisi statistica descrittiva e la rispettiva rappresentazione grafica.

2.4.1 Analisi statistica descrittiva

L'analisi statistica elementare riguarda l'insieme delle metodologie e degli strumenti che consentono di sintetizzare ed interpretare le osservazioni relative ad uno o più aspetti di un determinato fenomeno, mediante opportune rappresentazioni tabellari [*Bhattacharyya & Johnson, 1977*].

Alle misure di posizione appartengono la media aritmetica, la mediana, la moda, il valore minimo e il valore massimo. Le prime tre forniscono indicazioni sul valore rappresentativo di una distribuzione. La mediana, che è il valore che occupa la posizione centrale in un insieme ordinato di dati, è una misura robusta, in quanto poco influenzata dalla presenza di dati anomali. La moda, valore più frequente di una distribuzione, non è influenzata dalla presenza di nessun valore estremo; tuttavia viene utilizzata solamente a scopi descrittivi, perché è meno stabile e meno oggettiva delle altre misure di tendenza centrale. Essa, inoltre, consente di indagare sull'omogeneità delle osservazioni: se la distribuzione dei valori risulta plurimodale, allora il fenomeno si presenta eterogeneo.

Le misure di forma sono rappresentate dal coefficiente di asimmetria e dall'indice di curtosi: essi valutano gli aspetti riguardanti la forma di una distribuzione. In particolare, il coefficiente di asimmetria consente di quantificare l'asimmetria, appunto, di una distribuzione, mentre l'indice di curtosi fornisce una valutazione sulle frequenze corrispondenti alle modalità estreme rispetto a quelle centrali per una distribuzione unimodale o simmetrica.

Infine, varianza e deviazione standard rappresentano e danno informazioni circa la variabilità di un campione. La prima, nello specifico, è la misura della dispersione dei valori intorno alla media. Di più facile interpretazione è la deviazione standard, definita come la radice quadrata della varianza. Quest'ultima può essere intesa in termini di distanza di natura euclidea fra le singole osservazioni provenienti dalla popolazione e la loro media.

2.4.2 Descrizione grafica del campione

In genere, la rappresentazione grafica dei valori sperimentali, fornisce una valutazione preliminare della distribuzione dei dati; nonostante la relativa semplicità delle elaborazioni grafiche, la rappresentazione visiva dei dati può fornire una grande quantità di informazioni.

In primis, è interessante valutare se una la distribuzione del campione della variabile in esame si discosta o meno da una distribuzione normale: alcuni metodi di stima in Geostatistica (come il *Kriging* ordinario) forniscono risultati più attendibili se la distribuzione delle osservazioni è vicina a quella gaussiana. L'importanza di questa distribuzione nella letteratura statistica risiede nel teorema del limite centrale [Davis, 2002] il quale afferma che se si calcolano i valori medi di una serie di campioni prelevati casualmente da una popolazione, questi avranno una distribuzione normale; questa tendenza alla normalità sarà sempre più pronunciata all'aumentare del numero dei campioni.

Inoltre, attraverso i metodi grafici, è possibile identificare eventuali valori anomali (*outliers*) all'interno del campione, valori che devono essere opportunamente trattati e/o eliminati allo scopo di non falsificare l'eventuale interpolazione spaziale dei dati misurati che si hanno a disposizione.

2.4.2.1 Istogrammi di frequenza

Permette di ottenere un andamento approssimato della forma della distribuzione in funzione delle frequenze assolute ottenute dalle osservazioni; si tratta di un grafico che divide le osservazioni in classi, di solito equispaziate, e per ogni classe mostra il numero di osservazioni che vi appartengono; vi sono diversi modi per dividere l'intervallo delle osservazioni in classi, nessuno dei quali risulta ottimale.

La prima regola risale a *Sturges* [1926] e tiene conto solo della numerosità del campione: maggiore è il numero dei dati, minore è l'influenza delle fluttuazioni casuali e possiamo quindi utilizzare classi più strette. Essa può essere espressa come:

$$n_{class} = \log_2 n + 1 \quad (2.26)$$

in cui n rappresenta il numero dei dati del campione a disposizione.

Sono state introdotte altre formule che utilizzano solo la numerosità del campione, simile a quella di *Sturges*. Riportiamo ad esempio la regola di *Larson* [1975]:

$$n_{class} = 1 + 2.2 \log_{10} n \quad (2.27)$$

Infine, la formula di *Scott* [1978] tiene conto non solo del numero di dati, ma anche della deviazione standard σ :

$$n_{class} = 3.49 \sigma n^{1/3} \quad (2.28)$$

2.4.2.2 Box-Plot

Fornisce, in modo robusto, un'informazione sulla dispersione e sulla presenza di *outliers*. Utilizza la cosiddetta regola del *box plot* [Tukey, 1977], in cui si sfrutta il concetto di *range* interquartile (*IQR*) per identificare i possibili *outliers*: se un'osservazione x è compresa tra Q_3 e Q_1 (rispettivamente, primo e terzo quartile) non si tratta di un valore anomalo; diversamente si parla di *mild outlier* se $x \in (Q_3 + k \text{ IQR}, Q_3)$ o $x \in (Q_1, Q_1 - k \text{ IQR})$, o di *extreme outlier* negli altri casi. La costante k di solito viene posta pari a 1.5.

Oltre alla versione originale di *Tukey* [1977], ne esistono numerose varianti che differiscono principalmente per come calcolano i quartili [*Frigge et al.*, 1989].

2.4.2.3 Quantile-Quantile plot (Q-Q plot) e Normal Probability Plot (NPP)

I grafici Quantile-Quantile [*Wilk & Gnanadesikan*, 1968] e i grafici normali di probabilità fanno parte, appunto, della famiglia dei cosiddetti grafici di probabilità (*probability plot*). Il primo può essere utilizzato sia per test a un campione sia per quelli a due campioni; nel caso di test a un campione, il campione dei dati viene confrontato con uno generato sinteticamente dalla distribuzione di probabilità teorica per cui si sta conducendo il test (nel caso specifico si parla di distribuzione normale). Questo grafico non è altro che un diagramma a dispersione (*scatter plot*) relativo ai quantili di una popolazione e ai quantili campionari.

Il grafico normale di probabilità è un tipo di grafico cumulato che consente di verificare se una distribuzione è gaussiana. In questo tipo di grafico l'asse delle ordinate è riportato su una scala in maniera tale che il grafico cumulato delle frequenze sarà rappresentato come una linea retta nel caso di una distribuzione normale.

Infatti, in entrambe i casi, se il grafico è lineare i dati presentano una distribuzione simile a quella gaussiana; nel caso di grafici non lineari, gli scostamenti della linearità daranno informazioni importanti sul grado di asimmetria dei dati rispetto alla distribuzione normale. I grafici *Q-Q* e *NPP* sono anche utilizzati per l'identificazione di possibili *outliers*: la presenza di valori molto più grandi o molto più piccoli degli altri si manifesterà con la concentrazione di questi ultimi in zone più o meno distaccate rispetto a quella in cui risiedono gli altri valori della variabile.

2.4.3 Test statistici e verifica delle ipotesi

2.4.3.1 Verifica della presenza di outliers: the Dixon test [1953]

Questi test permettono di verificare, per un livello di significatività α fissato dallo sperimentatore, se un valore estremo può essere ritenuto anomalo, utilizzando le informazioni fornite dagli altri dati che costituiscono un campione [Castino & Roletto, 1991].

Il test di *Dixon* [1953] comprende in realtà un insieme di test, ognuno dei quali riferito ad una ben determinata situazione [Barnett & Lewis, 1978]. I dati x_1, x_2, \dots, x_n sono stati ordinati in senso crescente tale che:

$$x_1 \leq x_2 \leq \dots \leq x_n \quad (2.29)$$

Si impone la condizione che i dati provengono da una distribuzione normale $N(\mu, \sigma)$ della quale ignoriamo i parametri. Le ipotesi, inoltre, sono diverse a seconda del numero e della localizzazione dei dati anomali. Due sono le situazioni possibili:

1. la serie presenta un solo valore anomalo che sarà necessariamente o il primo (x_1 , il più piccolo) o l'ultimo (x_n , il più grande);
2. la serie presenta più valori sospetti alle estremità: si tratta, allora, o di x_1 e di x_n , oppure di due o più valori iniziali e/o di due o più valori finali.

Per entrambe le situazioni l'ipotesi nulla H_0 è la seguente: tutti i valori campionari provengono dalla stessa popolazione d'origine $N(\mu, \sigma)$. L'ipotesi alternativa H_1 afferma che uno o più valori estremi sono anomali.

La funzione discriminante è diversa a seconda che il valore anomalo sia l'estremo superiore o inferiore della serie ordinata.

1. Il valore anomalo è l'estremo inferiore (x_1) o superiore (x_n); la funzione discriminante può essere:

$$r_{10} = \frac{x_2 - x_1}{x_n - x_1} \quad \text{estremo inferiore} \quad (2.30)$$

$$r_{10} = \frac{x_n - x_{n-1}}{x_n - x_1} \quad \text{estremo superiore} \quad (2.31)$$

I valori critici sono riportati nella tabella 2.1 [Castino & Roletto, 1991] in funzione di n e di α .

2. La serie presenta due valori anomali: l'estremo inferiore e l'estremo superiore. In tal caso la funzione discriminante è:

$$r_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1} \quad \text{estremo inferiore} \quad (2.32)$$

$$r_{11} = \frac{x_n - x_{n-1}}{x_n - x_2} \quad \text{estremo superiore} \quad (2.33)$$

che permette di evitare l'effetto mascherante di un valore x_n o x_1 anomalo. I valori critici sono riportati nella Tavola XIIb [Castino & Roletto, 1991] in funzione di n e di α .

$\alpha \backslash n$	0.005	0.01	0.02	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
3	0.994	0.988	0.976	0.941	0.886	0.781	0.684	0.591	0.500	0.409	0.316	0.219	0.114	0.059
4	0.926	0.889	0.846	0.765	0.679	0.560	0.471	0.394	0.324	0.257	0.193	0.130	0.065	0.033
5	0.821	0.780	0.729	0.642	0.557	0.451	0.373	0.308	0.250	0.196	0.146	0.097	0.048	0.023
6	0.740	0.698	0.644	0.560	0.482	0.386	0.318	0.261	0.210	0.164	0.121	0.079	0.038	0.018
7	0.680	0.637	0.586	0.507	0.434	0.344	0.281	0.230	0.184	0.143	0.105	0.068	0.032	0.016
8	0.634	0.590	0.543	0.468	0.399	0.314	0.255	0.208	0.166	0.128	0.094	0.060	0.029	0.014
9	0.598	0.555	0.510	0.437	0.370	0.290	0.234	0.191	0.152	0.118	0.086	0.055	0.026	0.013
10	0.568	0.527	0.483	0.412	0.349	0.273	0.219	0.178	0.142	0.110	0.080	0.051	0.025	0.012
11	0.542	0.502	0.460	0.392	0.332	0.259	0.208	0.168	0.133	0.103	0.074	0.048	0.023	0.011
12	0.522	0.482	0.441	0.376	0.318	0.247	0.197	0.160	0.126	0.097	0.070	0.045	0.022	0.011
13	0.503	0.465	0.425	0.361	0.305	0.237	0.188	0.153	0.120	0.092	0.067	0.043	0.021	0.010
14	0.488	0.450	0.411	0.349	0.294	0.228	0.181	0.147	0.115	0.088	0.064	0.041	0.020	0.010
15	0.475	0.438	0.399	0.338	0.285	0.220	0.175	0.141	0.111	0.085	0.062	0.040	0.019	0.010
16	0.463	0.426	0.388	0.329	0.277	0.213	0.169	0.136	0.107	0.082	0.060	0.039	0.019	0.009
17	0.452	0.416	0.379	0.320	0.269	0.207	0.165	0.132	0.104	0.080	0.058	0.038	0.018	0.009
18	0.442	0.407	0.370	0.313	0.263	0.202	0.160	0.128	0.101	0.078	0.056	0.036	0.018	0.009
19	0.433	0.398	0.363	0.306	0.258	0.197	0.157	0.125	0.098	0.076	0.055	0.036	0.017	0.008
20	0.425	0.391	0.356	0.300	0.252	0.193	0.153	0.122	0.096	0.074	0.053	0.035	0.017	0.008
21	0.418	0.384	0.350	0.295	0.247	0.189	0.150	0.119	0.094	0.072	0.052	0.034	0.016	0.008
22	0.411	0.378	0.344	0.290	0.242	0.185	0.147	0.117	0.092	0.071	0.051	0.033	0.016	0.008
23	0.404	0.372	0.338	0.285	0.238	0.182	0.144	0.115	0.090	0.069	0.050	0.033	0.016	0.008
24	0.399	0.367	0.333	0.281	0.234	0.179	0.142	0.113	0.089	0.068	0.049	0.032	0.016	0.008
25	0.393	0.362	0.329	0.277	0.230	0.176	0.139	0.111	0.088	0.067	0.048	0.032	0.015	0.008
26	0.388	0.357	0.324	0.273	0.227	0.173	0.137	0.109	0.086	0.066	0.047	0.031	0.015	0.007
27	0.384	0.353	0.320	0.269	0.224	0.171	0.135	0.108	0.085	0.065	0.047	0.031	0.015	0.007
28	0.380	0.349	0.316	0.266	0.220	0.168	0.133	0.106	0.084	0.064	0.046	0.030	0.015	0.007
29	0.376	0.345	0.312	0.263	0.218	0.166	0.131	0.105	0.083	0.063	0.046	0.030	0.014	0.007
30	0.371	0.341	0.309	0.260	0.215	0.164	0.130	0.103	0.082	0.062	0.045	0.029	0.014	0.007

Tabella 2.1 – Valori critici del test di Dixon [1953] nel caso in cui il valore anomalo è rappresentato dall'estremo inferiore o superiore [Castino & Roletto, 1991]

3. La serie presenta una coppia di valori anomali. Può trattarsi della coppia superiore (x_n e x_{n-1}) oppure della coppia inferiore x_1 e x_2 . La funzione discriminante sarà:

$$r_{20} = \frac{x_3 - x_1}{x_n - x_1} \quad \text{estremo inferiore} \quad (2.34)$$

$$r_{20} = \frac{x_n - x_{n-2}}{x_n - x_1} \quad \text{estremo superiore} \quad (2.35)$$

I valori critici delle funzione r_{20} sono riportati nella tabella 2.2 [Castino & Roletto, 1991] in funzione di n e di α .

$n \backslash \alpha$	0.005	0.01	0.02	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
4	0.996	0.992	0.987	0.967	0.935	0.871	0.807	0.743	0.676	0.606	0.529	0.440	0.321	0.235
5	0.950	0.929	0.901	0.845	0.782	0.694	0.623	0.560	0.500	0.440	0.377	0.306	0.218	0.155
6	0.865	0.836	0.800	0.736	0.670	0.585	0.520	0.463	0.411	0.358	0.305	0.245	0.172	0.126
7	0.814	0.778	0.732	0.661	0.596	0.516	0.454	0.402	0.355	0.306	0.261	0.208	0.144	0.099
8	0.746	0.710	0.670	0.607	0.545	0.468	0.410	0.361	0.317	0.274	0.230	0.184	0.125	0.085
9	0.700	0.667	0.627	0.565	0.505	0.432	0.378	0.331	0.288	0.250	0.208	0.166	0.114	0.077
10	0.664	0.632	0.592	0.531	0.474	0.404	0.354	0.307	0.268	0.231	0.192	0.153	0.104	0.070
11	0.627	0.603	0.564	0.504	0.449	0.381	0.334	0.290	0.253	0.217	0.181	0.143	0.097	0.065
12	0.612	0.579	0.540	0.481	0.429	0.362	0.316	0.274	0.239	0.205	0.172	0.136	0.091	0.060
13	0.590	0.557	0.520	0.461	0.411	0.345	0.301	0.261	0.227	0.195	0.164	0.129	0.086	0.057
14	0.571	0.538	0.502	0.445	0.395	0.332	0.288	0.250	0.217	0.187	0.157	0.123	0.082	0.054
15	0.554	0.522	0.486	0.430	0.382	0.320	0.277	0.241	0.209	0.179	0.150	0.118	0.079	0.052
16	0.539	0.508	0.472	0.418	0.370	0.310	0.268	0.233	0.202	0.173	0.144	0.113	0.076	0.050
17	0.526	0.495	0.460	0.406	0.359	0.301	0.260	0.226	0.195	0.167	0.139	0.109	0.074	0.049
18	0.514	0.484	0.449	0.397	0.350	0.293	0.252	0.219	0.189	0.162	0.134	0.105	0.071	0.048
19	0.503	0.473	0.439	0.379	0.341	0.286	0.246	0.213	0.184	0.157	0.130	0.101	0.069	0.047
20	0.494	0.464	0.430	0.372	0.333	0.279	0.240	0.208	0.179	0.152	0.126	0.098	0.067	0.046
21	0.485	0.455	0.422	0.365	0.326	0.273	0.235	0.203	0.175	0.148	0.123	0.096	0.065	0.045
22	0.477	0.447	0.414	0.358	0.320	0.267	0.230	0.199	0.171	0.145	0.120	0.094	0.064	0.044
23	0.469	0.440	0.407	0.352	0.314	0.262	0.225	0.195	0.167	0.142	0.117	0.092	0.062	0.043
24	0.462	0.434	0.401	0.347	0.309	0.258	0.221	0.192	0.164	0.139	0.114	0.090	0.061	0.042
25	0.456	0.428	0.395	0.343	0.304	0.254	0.217	0.189	0.161	0.136	0.112	0.089	0.060	0.041
26	0.450	0.422	0.390	0.338	0.300	0.250	0.214	0.186	0.158	0.134	0.110	0.087	0.059	0.041
27	0.444	0.417	0.385	0.334	0.296	0.246	0.211	0.183	0.156	0.132	0.109	0.086	0.058	0.040
28	0.439	0.412	0.381	0.330	0.292	0.243	0.208	0.180	0.154	0.130	0.107	0.085	0.058	0.040
29	0.434	0.407	0.376	0.326	0.288	0.239	0.205	0.177	0.151	0.128	0.106	0.083	0.057	0.039
30	0.428	0.402	0.372	0.322	0.285	0.236	0.202	0.175	0.149	0.126	0.104	0.082	0.056	0.039

Tabella 2.2 – Valori critici del test di Dixon [1953] nel caso in cui i valori anomali sono rappresentati da una coppia di dati [Castino & Roletto, 1991]

4. La serie presenta una coppia di valori anomali, i due inferiori oppure i due superiori, ed un ulteriore valore anomalo all'altra estremità della serie (x_n se la coppia sospetta è x_1 e x_2 ; x_1 se la coppia sospetta è x_n e x_{n-1}). La funzione discriminante è:

$$r_{21} = \frac{x_3 - x_1}{x_{n-1} - x_1} \quad (2.36)$$

per il valore anomalo x_1 , in modo da evitare l'effetto mascherante di x_2 e x_n ;

$$r_{21} = \frac{x_n - x_{n-2}}{x_n - x_3} \quad (2.37)$$

per il valore anomalo x_n , in modo da evitare l'effetto mascherante di x_{n-1} e x_1 . I valori critici di r_{21} sono riportati nella Tavola XIId [Castino & Roletto, 1991]. Sono possibili, inoltre, altre situazioni per le quali esistono altrettanti opportune funzioni discriminanti [Barnett & Lewis, 1978].

2.4.3.2 Verifica dell'ipotesi di normalità

Il test χ^2 (chi-quadrato) rappresenta uno dei test statistici più utilizzati ed è usato per verificare l'ipotesi H_0 che un certo campione di dati provenga da una specifica distribuzione nota (normale). Per valutare, infatti, la normalità di una distribuzione di dati sperimentali *Snedecor & Cochran* [1974] propongono di ricorrere al test del χ^2 , chiamato appunto test per la bontà di adattamento (*goodness of fit test*) confrontando la distribuzione osservata con quella attesa, quest'ultima costruita mediante la media e la varianza del campione applicate alla normale. L'ipotesi nulla è che non esistano differenze significative tra la distribuzione dei dati raccolti e quella normale corrispondente, con stessa media e stessa varianza. L'ipotesi alternativa è che la distribuzione osservata se ne discosti in modo significativo, per un effetto combinato di asimmetria e curtosi. Partendo dai dati campionari è necessario stimare le frequenze attese e calcolare il valore del χ^2 . Si individuano i valori centrali (\bar{X}_i) di ciascuna classe; si calcola la media generale ($\bar{\bar{X}}$) della distribuzione osservata:

$$\bar{\bar{X}} = \frac{\sum_{i=1}^k (\bar{X}_i \cdot n_i)}{n} \quad (2.38)$$

in cui k è il numero di classi, n_i il numero di osservazioni della classe i ed n il numero totale delle osservazioni. Si calcola, quindi, la deviazione standard (σ) della distribuzione e si stima il valore di Z per gli estremi di ogni classe (X_i) calcolato mediante la relazione:

$$Z = \frac{X_i - \bar{\bar{X}}}{\sigma} \quad (2.39)$$

Dal valore Z di ogni estremo di classe si ricava, attraverso la tavola della distribuzione normale [Snedecor & Cochran, 1974], la frequenza relativa corrispondente; per differenza, si stima la frequenza relativa di ogni classe. Rapportando ad n tali frequenze relative, si ottiene la frequenza attesa di ciascuna classe. Il χ^2 si stima per ogni classe con la seguente relazione:

$$\chi^2 = \frac{(\text{Freq. Oss.} - \text{Freq. Att.})^2}{\text{Freq. Att.}} \quad (2.40)$$

La somma dei valori χ^2 fornisce il valore del χ^2 totale, sotto condizione di validità di questo test che ogni frequenza attesa non sia inferiore a 5. Il numero di gradi di libertà del χ^2 è pari a:

$$\nu = k - r - 1 \quad (2.41)$$

in cui r è il numero dei parametri della distribuzione indagata. La tabella 2.3 riporta il valore critico (V_{cr}) del χ^2 per un determinato numero di gradi di libertà ν e per un certo livello di significatività α . Se il χ^2 calcolato risulta essere minore rispetto al V_{cr} riportato nelle apposite tabelle, si ha l'accettazione dell'ipotesi H_0 .

ν	0.20	0.10	0.05	0.025	α 0.01	0.005	0.001	0.0005	0.0001
1	1.6424	2.7055	3.8415	5.0239	6.6349	7.8794	10.8274	12.1153	15.1343
2	3.2189	4.6052	5.9915	7.3778	9.2104	10.5965	13.8150	15.2014	18.4247
3	4.6416	6.2514	7.8147	9.3484	11.3449	12.8381	16.2660	17.7311	21.1040
4	5.9886	7.7794	9.4877	11.1433	13.2767	14.8602	18.4662	19.9977	23.5064
5	7.2893	9.2363	11.0705	12.8325	15.0863	16.7496	20.5147	22.1057	25.7507
6	8.5581	10.6446	12.5916	14.4494	16.8119	18.5475	22.4575	24.1016	27.8527
7	9.8032	12.0170	14.0671	16.0128	18.4753	20.2777	24.3213	26.0179	29.8814
8	11.0301	13.3616	15.5073	17.5345	20.0902	21.9549	26.1239	27.8674	31.8268
9	12.2421	14.6837	16.9190	19.0228	21.6660	23.5893	27.8767	29.6669	33.7247
10	13.4420	15.9872	18.3070	20.4832	23.2093	25.1881	29.5879	31.4195	35.5572
11	14.6314	17.2750	19.6752	21.9200	24.7250	26.7569	31.2635	33.1382	37.3647
12	15.8120	18.5493	21.0261	23.3367	26.2170	28.2997	32.9092	34.8211	39.1306
13	16.9848	19.8119	22.3620	24.7356	27.6882	29.8193	34.5274	36.4768	40.8735
14	18.1508	21.0641	23.6848	26.1189	29.1412	31.3194	36.1239	38.1085	42.5752
15	19.3107	22.3071	24.9958	27.4884	30.5780	32.8015	37.6978	39.7173	44.2596
16	20.4651	23.5418	26.2962	28.8453	31.9999	34.2671	39.2518	41.3077	45.9255
17	21.6146	24.7690	27.5871	30.1910	33.4087	35.7184	40.7911	42.8808	47.5591
18	22.7595	25.9894	28.8693	31.5264	34.8052	37.1564	42.3119	44.4337	49.1853
19	23.9004	27.2036	30.1435	32.8523	36.1908	38.5821	43.8194	45.9738	50.7873
20	25.0375	28.4120	31.4104	34.1696	37.5663	39.9969	45.3142	47.4977	52.3832
21	26.1711	29.6151	32.6706	35.4789	38.9322	41.4009	46.7963	49.0096	53.9599
22	27.3015	30.8133	33.9245	36.7807	40.2894	42.7957	48.2676	50.5105	55.5244
23	28.4288	32.0069	35.1725	38.0756	41.6383	44.1814	49.7276	51.9995	57.0668
24	29.5533	33.1962	36.4150	39.3641	42.9798	45.5584	51.1790	53.4776	58.6071
25	30.6752	34.3816	37.6525	40.6465	44.3140	46.9280	52.6187	54.9475	60.1360
26	31.7946	35.5632	38.8851	41.9231	45.6416	48.2898	54.0511	56.4068	61.6666
27	32.9117	36.7412	40.1133	43.1945	46.9628	49.6450	55.4751	57.8556	63.1660
28	34.0266	37.9159	41.3372	44.4608	48.2782	50.9936	56.8918	59.2990	64.6561
29	35.1394	39.0875	42.5569	45.7223	49.5878	52.3355	58.3006	60.7342	66.1524
30	36.2502	40.2560	43.7730	46.9792	50.8922	53.6719	59.7022	62.1600	67.6230
35	41.7780	46.0588	49.8018	53.2033	57.3420	60.2746	66.6192	69.1975	74.9253
40	47.2685	51.8050	55.7585	59.3417	63.6908	66.7660	73.4029	76.0963	82.0551
45	52.7288	57.5053	61.6562	65.4101	69.9569	73.1660	80.0776	82.8734	89.0704
50	58.1638	63.1671	67.5048	71.4202	76.1538	79.4898	86.6603	89.5597	95.9713
55	63.5772	68.7962	73.3115	77.3804	82.2920	85.7491	93.1671	96.1607	102.7735
60	68.9721	74.3970	79.0820	83.2977	88.3794	91.9518	99.6078	102.6971	109.4967
70	79.7147	85.5270	90.5313	95.0231	100.4251	104.2148	112.3167	115.5766	122.7443
80	90.4053	96.5782	101.8795	106.6285	112.3288	116.3209	124.8389	128.2636	135.7728
90	101.0537	107.5650	113.1452	118.1359	124.1162	128.2987	137.2082	140.7804	148.6198
100	111.6667	118.4980	124.3421	129.5613	135.8069	140.1697	149.4488	153.1638	161.3297
150	164.3492	172.5812	179.5806	185.8004	193.2075	198.3599	209.2652	213.6135	223.1209
200	216.6088	226.0210	233.9942	241.0578	249.4452	255.2638	267.5388	272.4220	283.0448

Tabella 2.3 – Valori critici per la distribuzione del test del χ^2

Il test di *Kolmogorov-Smirnov* [Kolmogorov, 1933; Smirnov, 1939] è di solito più efficace del test di bontà dell'adattamento tramite il chi-quadrato [Snedecor & Cochran, 1974] per campioni di taglia piccola e può essere utilizzato per campioni con pochi dati quando quest'ultimo non si possa impiegare. La relativa statistica fa parte delle cosiddette *Supremum statistics* e rappresenta una misura di distanza verticale massime fra due funzioni. Questo tipo di test si concentra sulla forma della densità di probabilità del campione osservato e la confronta con quella ipotizzata. In realtà, più che analizzare la densità di probabilità, il test stima la funzione di distribuzione osservata ($F_0(X)$) del campione di dati X_1, X_2, \dots, X_n . Si ricava così una curva continua connettendo tutti i punti determinati al passo precedente. Si calcolano i valori

della distribuzione ipotizzata (normale) definita come $F_n(X)$; la funzione di distribuzione osservata si mette a confronto con quella ipotizzata costruendo la seguente variabile D che verrà impiegata come statistica del test:

$$D = \max\{|F_0(X_i) - F_n(X_i)|\} \quad (2.42)$$

Essa rappresenta il punto di massima distanza tra la curva ipotizzata e quella osservata (fig. 2.3). Se D risulta minore rispetto a un valore critico V_{cr} , tabellato in funzione di n e di α (tab. 2.4), l'ipotesi fatta (campione gaussiano) può essere accettata.

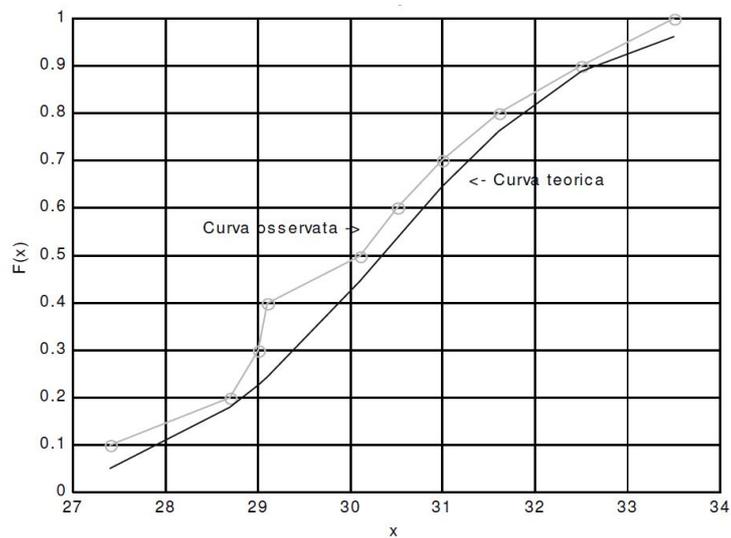


Figura 2.3 – Rappresentazione grafica della curva teorica e di quella osservata per la valutazione della statistica D del test di Kolmogorov-Smirnov

Taglia campionaria n	$D_{.10}$	$D_{.05}$	$D_{.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29

Tabella 2.3 – Valori della statistica D del test di Kolmogorov-Smirnov

Infine, il test di *Anderson-Darling* [1952] è un test numerico, derivante dal test di *Kolmogorov-Smirnov* [Kolmogorov, 1933; Smirnov, 1939], basato sempre sulla funzione di distribuzione empirica [Stephens, 1974], la cui statistica fa parte delle cosiddette *Quadratic statistics*; rappresenta una misura di distanza verticale quadratica pesata fra due funzioni di distribuzione. La versione del test a un campione misura, quindi, la distanza quadratica pesata tra la funzione di distribuzione cumulativa empirica $F_0(X)$ di un campione e la funzione di distribuzione cumulativa teorica $F_n(X)$, applicata allo stesso campione. Il valore della statistica del test (A^2) è data dalla seguente relazione:

$$A^2 = -n - \sum_{i=1}^n \frac{2i-1}{n} [\ln F_0(X_i) + \ln (1 - F_0(X_{n-i+1}))] \quad (2.43)$$

Dall'equazione (2.43) emerge che il test è non-parametrico, cioè è indipendente dalla distribuzione (*distribution free*). L'ipotesi nulla del test (campione normale) viene rifiutata quando la statistica A^2 risulta maggiore di un certo valore critico V_{cr} [Anderson & Darling, 1952].

2.5 Analisi strutturale

Le problematiche che si presentano più frequentemente nelle applicazioni e che si affrontano mediante tecniche geostatistiche riguardano, da un lato, la stima del valore di una variabile d'interesse in un punto in cui non si dispone di alcuna misura (obiettivi inferenziali), e, dall'altro, la generazione di diverse mappe spaziali alternative ed equiprobabili della variabile in esame (obiettivi di simulazione). Sia in un caso che nell'altro è comunque necessario disporre di un modello di correlazione spaziale: in Geostatistica, lo studio della correlazione spaziale, viene generalmente effettuato mediante l'analisi del variogramma. Quest'ultimo, infatti, essendo un momento del secondo ordine, consente di modellare la variabilità spaziale del campo aleatorio in esame [Matheron, 1965]. Il diffuso utilizzo del variogramma è attestato, inoltre, dai numerosi contributi scientifici dedicati a tale strumento. In particolare, si pensi ai lavori specialistici riguardanti gli stimatori robusti o la valutazione dell'incertezza associata ai parametri del modello del variogramma [Cressie, 1985; Ortiz & Deutsch, 2002]. È opportuno sottolineare che l'uso del variogramma è apparso nella letteratura scientifica nel 1941, ovvero prima che la Geostatistica divenisse una vera e propria disciplina: Kolmogorov [1941] lo introdusse per studiare la struttura locale della turbolenza di un fluido, Yaglom [1957] lo definì funzione di struttura nell'ambito della teoria sui processi stocastici, infine Gandin [1963] lo utilizzò in meteorologia.

L'analisi della correlazione spaziale, denominata anche analisi strutturale, rappresenta il passo più delicato e importante di uno studio geostatistico e richiede la stima del variogramma e il successivo adattamento tramite un modello, in grado di riassumere le caratteristiche spaziali riscontrate. Di seguito verranno descritti le fasi principali per effettuare una corretta e precisa analisi strutturale.

2.5.1 Il variogramma sperimentale

Il metodo più semplice per confrontare due valori è quello di calcolare la differenza tra due campioni. Nel caso si devono confrontare le differenze di un elevato numero di coppie di osservazioni, è possibile osservare quanto segue:

1. alcune differenze saranno positive e altre negative, pertanto elevando al quadrato le differenze si otterranno tutti valori positivi;
2. le differenze misurate tra campioni separati da distanze simili sono elevate al quadrato e mediate.

L'espressione più comunemente utilizzata per il calcolo del variogramma è quella proposta da Matheron [1970]:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(u_i) - z(u_i + h)]^2 \quad (2.44)$$

dove h appartiene a classi di distanze in cui viene diviso l'intervallo ed $N(h)$ il numero di coppie la cui distanza è compresa nell'intervallo considerato. Le distanze fra le coppie di punti campione sono rappresentate sull'asse delle ascisse ed etichettate con la lettera h ; le medie delle differenze elevate al quadrato relative a tutte le coppie ad uguale distanza sono rappresentate sull'asse delle ordinate ed etichettate con $\gamma(h)$. Il grafico così costruito prende il nome di variogramma sperimentale (fig. 2.4).

La determinazione del variogramma sperimentale richiede la definizione di diversi parametri.

In generale, la grandezza principale da definire è chiamata *lag* e rappresenta il passo (o distanza) utilizzata per confrontare le differenze tra i campioni. Se i punti campionati sono collocati su una griglia pseudo-regolare, la distanza tra i punti della griglia può costituire un buon *lag*; se la distribuzione dei punti è casuale, come spesso succede, si può utilizzare come *lag*, almeno inizialmente, la distanza minima tra tutti i campioni o la media delle minime distanze. In effetti, non esiste una regola ben precisa che permette di definire la giusta ed esatta misura del *lag*: la scelta di tale parametro dipende essenzialmente dalla tipologia di campionamento effettuato e dall'estensione dell'area esaminata.

Il numero di *lag* è il numero di intervalli di distanza che saranno presi in esame dal variogramma. La scelta del numero di *lag* è naturalmente dettata dalla grandezza dell'area investigata: in generale i variogrammi non sono validi se calcolati per distanze maggiori di un terzo o la metà della distanza esistente tra la coppia di campioni più lontani tra di loro [Ciotoli & Finoia, 2005].

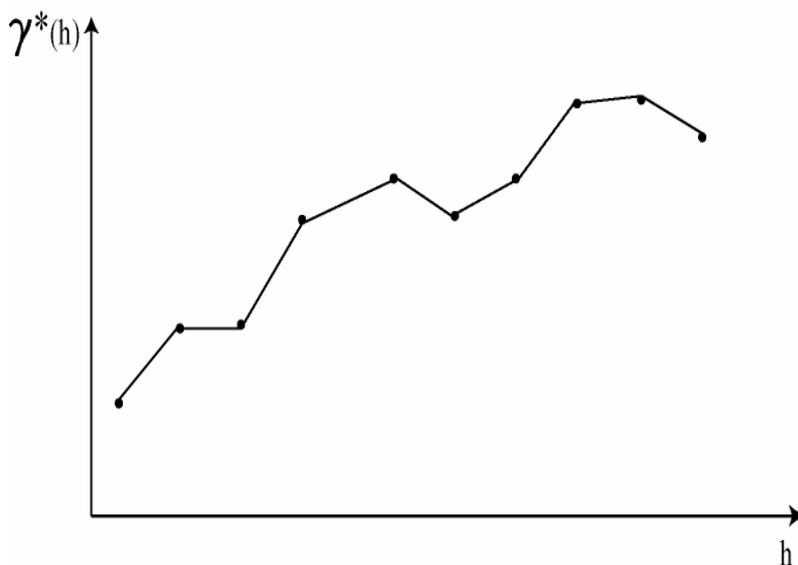


Figura 2.4 – Rappresentazione di un generico variogramma sperimentale

Nel caso in cui si ha una distribuzione dei campioni uniforme, ma non regolare, è necessario introdurre nella scelta del *lag* una tolleranza lineare dovuta al fatto che è possibile non trovare coppie per una distanza definita. Per la scelta della distanza di tolleranza comunemente si tende utilizzare la metà del *lag*, anche se spesso non è errato scegliere una tolleranza minore poiché in questa maniera verranno utilizzate meno coppie nella costruzione del variogramma che risulterà di conseguenza più chiaro. La figura 2.5 riporta uno schema generico delle caratteristiche dei parametri appena esposti: *lag*, numero di *lag* e tolleranza lineare.

Nel caso in cui si vuole indagare sulla presenza di anisotropie (par. 2.5.4) nella distribuzione spaziale dei dati, è necessario calcolare variogrammi direzionali; in tal caso bisognerà definire non solo una direzione specifica del vettore *lag* ma anche una tolleranza angolare. Teoricamente bisognerebbe utilizzare una tolleranza angolare più piccola possibile in maniera tale da limitare la confusione che potrebbe nascere combinando coppie con differenti direzioni.

Sfortunatamente utilizzando una tolleranza angolare molto piccola entreranno poche coppie nel calcolo del variogramma che si presenterà troppo irregolare per una buona descrizione. In questi casi il migliore approccio consiste nel tentare la costruzione del variogramma usando differenti tolleranze e scegliendo la più piccola che fornisce il miglior risultato.

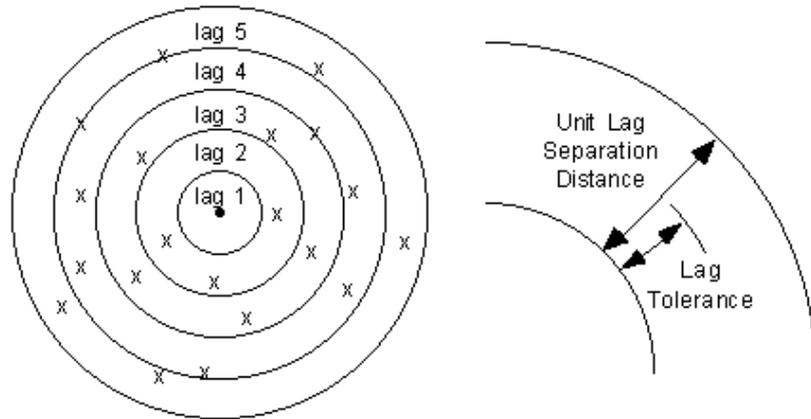


Figura 2.5 – Schematizzazione generica del lag, numero di lag e tolleranza lineare

Infine, per definire un cono di ricerca troncato, bisogna stabilire un ulteriore parametro detto *maximum bandwidth* (fig. 2.6).

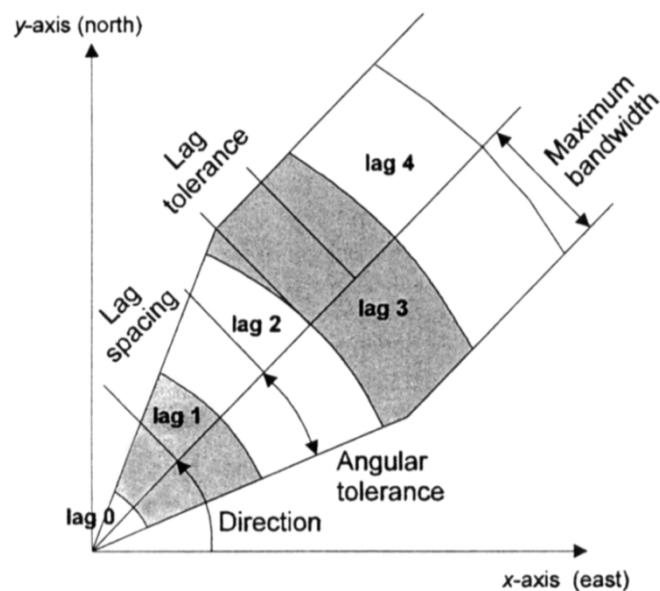


Figura 2.6 – Schema rappresentante il cono di ricerca per la definizione di un variogramma sperimentale direzionale [Pannatier, 1996]

2.5.2 Comportamento del variogramma in prossimità dell'origine

Il comportamento del variogramma in prossimità dell'origine risulta essere significativo al fine di valutare la continuità e la regolarità spaziale della variabile regionalizzata [Posa, 1991]. Di seguito vengono elencati i quattro andamenti tipici di un variogramma in prossimità dell'origine.

Il variogramma presenta un comportamento parabolico (fig. 2.7A) in prossimità dell'origine se $\gamma(h) \approx |h|^2$, per $|h| \rightarrow 0$. Tale comportamento è tipico di una variabile regionalizzata continua e differenziabile (nel senso della media quadratica).

Il variogramma è caratterizzato da un comportamento lineare (fig. 2.7B) in prossimità dell'origine se $\gamma(h) \approx |h|$, per $|h| \rightarrow 0$. In tal caso la variabile regionalizzata risulta essere continua, nel senso della media quadratica, ma non differenziabile, e quindi meno regolare rispetto a quella con comportamento parabolico.

Il variogramma presenta un effetto *nugget* (fig. 2.7C) se:

$$\lim_{|h| \rightarrow 0} \gamma(h) \neq \gamma(0) \quad (2.45)$$

ovvero, se la funzione γ è discontinua nell'origine. In altri termini in presenza di un effetto *nugget*, il variogramma tende, per piccole distanza $|h|$, ad un vettore $k \neq 0$, sebbene per definizione risulti $\gamma(0) = 0$. Questo comportamento indica che il fenomeno spaziale non è continuo, nel senso della media quadratica, e quindi si presenta molto irregolare. La denominazione effetto *nugget* deriva proprio da termine *nugget*, che nel gergo anglosassone significa pepita, ed è storicamente riconducibile all'andamento tipico del variogramma calcolato sui dati relativi ai quantitativi di metallo prezioso ritrovati in un deposito aurifero. In generale l'effetto pepita è attribuibile a:

1. microstrutture, ovvero componenti del fenomeno in cui il *range* (par. 2.5.3) è inferiore alla più piccola distanza disponibile per il calcolo del variogramma. L'eventuale presenza di un effetto *nugget* dipende dal campionamento effettuato (distanza minima tra le localizzazioni effettuate) e dalla variabilità presentata dal fenomeno;
2. errori di misura;
3. errori di localizzazione.

Infine, nel caso di comportamento a curva piatta (fig. 2.7D), il variogramma presenta solo una discontinuità nell'origine ed è costante per ogni valore di $|h|$. Ciò indica che la variabile regionalizzata non presenta alcuna struttura: $Z(u)$ e $Z(u+h)$ non sono correlate, qualunque sia il vettore di separazione h tra i punti u e $u+h$.

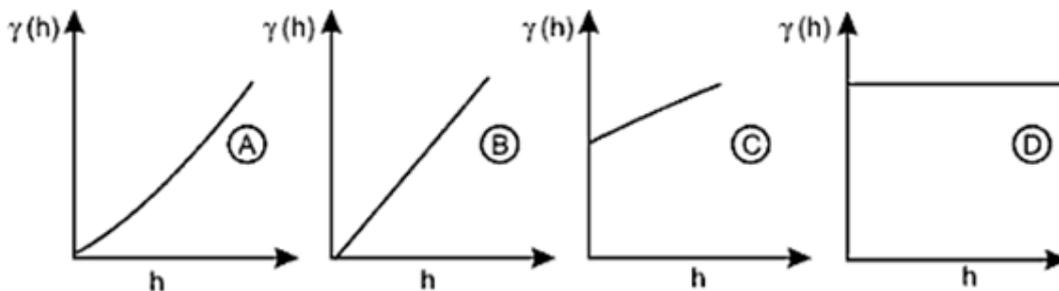


Figura 2.7 – Comportamento di un variogramma in prossimità dell'origine. A) parabolico, B) lineare, C) effetto nugget, D) curva piatta [Sarma, 2009]

2.5.3 Comportamento tipico del variogramma

Il variogramma, generalmente, presenta il seguente comportamento (fig. 2.8):

1. parte dall'origine, ovvero per $h = 0$ e risulta $\gamma(0) = 0$ oppure presenta all'origine una discontinuità, il *nugget* (C_0), come riportato nel paragrafo precedente;
2. cresce all'aumentare della distanza h ;
3. conserva un andamento crescente anche per distanze molto elevate, oppure si stabilizza intorno ad un valore limite (fig. 2.8), detto *sill* (C).

Se il variogramma raggiunge un valore limite, ossia se è possibile identificare un valore del *sill*, esiste una distanza a , denominata *range*, oltre la quale le variabili aleatorie $Z(u)$ e $Z(u+h)$ non sono più correlate. Da questo punto di vista il *range* fornisce un significato preciso al tradizionale concetto di zona di influenza di una variabile aleatoria: una variabile $Z(u)$ sarà correlata con qualsiasi altra variabile la cui localizzazione appartenga all'intorno centrato in u e di raggio a [Posa & De Iaco, 2009]. Come sarà specificato nel paragrafo 2.5.5 dedicato alle

anisotropie, il *range* e il *sill* dipendono solitamente dalla direzione lungo la quale viene definito il variogramma.

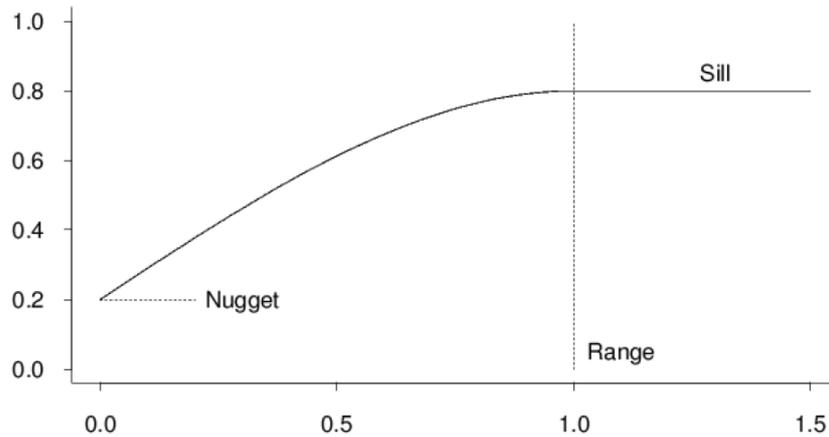


Figura 2.8 – Forma idealizzata di un variogramma con nugget (C_0), sill (C) e range (a)

2.5.4 Lo studio delle anisotropie

Gran parte dei fenomeni naturali si comportano in maniera anisotropa, cioè le variabili che li caratterizzano non si presentano uniformi in tutte le direzioni [Journel & Huijbregts, 1978]. Vi sono due tipi di anisotropia: quella più semplice si verifica quando lo stesso modello di variogramma e lo stesso *sill* sono comuni a tutte le direzioni, mentre cambia solamente il *range* (fig. 2.9a). In queste situazioni si ha, quindi, un unico *sill* che viene raggiunto dal semivariogramma a distanze diverse. In particolare si avrà una determinata direzione in cui il *range* risulta massimo (a_{max}), mentre in quella ortogonale risulta minimo (a_{min}). Questo tipo di anisotropia è detta geometrica (o anisotropia ellittica) e il rapporto fra il *range* minimo e quello massimo è noto come rapporto di anisotropia:

$$R = \frac{a_{min}}{a_{max}} \quad (2.46)$$

Il *range* a_{max} corrisponde, quindi, alla direzione dell'asse maggiore di un'ellisse lungo la quale il processo spaziale appare più continuo e quindi più fortemente correlato; il *range* a_{min} definisce la

direzione di minima continuità spaziale lungo la quale il fenomeno spaziale appare meno continuo.

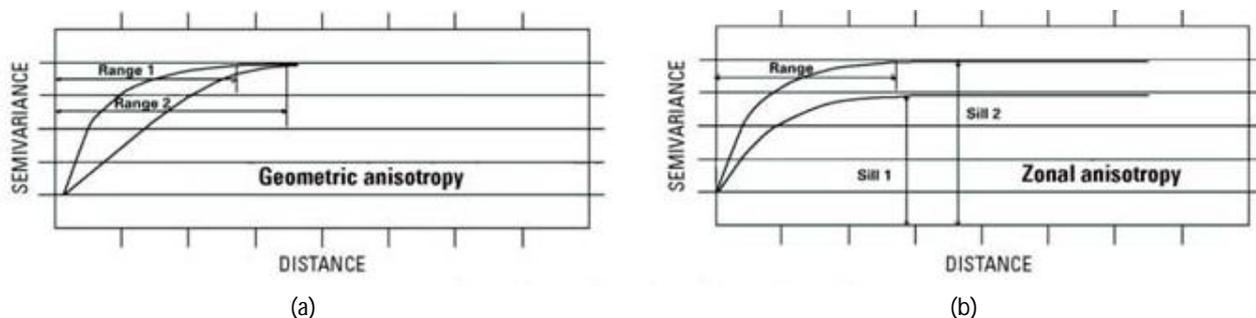


Figura 2.9 – Esempio di variogrammi con a) anisotropia geometrica e b) anisotropia zonale [Hack, 2005]

Differentemente, l'anisotropia zonale (fig. 2.9b) è caratterizzata da variogrammi con *sill* differenti, ma stesso *range*; ad esempio, l'anisotropia zonale, nel campo della geologia, è tipica di depositi stratificati laddove la variabilità di una certa proprietà può mantenersi bassa anche per grandi distanze all'interno di un orizzonte stratigrafico, ma può variare molto per brevi distanze quando si esaminano proprietà tra gli orizzonti stratigrafici [Ciotoli & Finoia, 2005]. Si può inoltre affermare che l'anisotropia geometrica è un caso particolare di anisotropia zonale [Journel & Huijbregts, 1978].

Al fine di poter individuare la presenza di anisotropie nella distribuzione spaziale di una variabile regionalizzata si possono individuare i variogrammi lungo direzioni selezionate; se questi mostrano un comportamento simile per le differenti direzioni, allora si può affermare che le relazioni distanza-differenze dipendono solo dal valore della distanza alla quale sono confrontate le coppie: in questo tipo di condizioni il fenomeno può considerarsi isotropo. Al contrario, se la costruzione dei variogrammi direzionali mette in evidenza un comportamento molto differente (cambiamenti nei valori di *sill* e *range*) a seconda della direzione secondo la quale viene calcolato, in tal caso si parla di comportamento anisotropo del fenomeno. Generalmente, il calcolo dei variogrammi si effettua lungo le direzioni Est, Nord-Est, Nord, Nord-Ovest, utilizzando una tolleranza angolare di 22.5°.

Un'altra tecnica per l'esplorazione di eventuali anisotropie del fenomeno esaminato consiste nell'elaborazione della mappa della varianza o superficie del variogramma (*variogram surface*).

Immaginiamo un grafico in cui ogni punto campionato abbia un'origine comune. Questa proprietà rende il variogramma simmetrico. Questo sistema fa sì che si possano raggruppare tutte le distanze vettoriali h tra un punto e tutti gli altri secondo il modulo vettoriale (lunghezza del vettore, e quindi distanza) e l'orientazione del vettore stesso. Tale grafico si costruisce utilizzando una griglia con celle regolari o quadrate oppure una griglia con celle a settori circolari. Le due rappresentazioni sono analoghe e dipendono solamente dalla convenzione del *software* utilizzato. La presenza di celle caratterizzate dallo stesso colore e che definiscono una certa continuità e direzione sulla mappa del variogramma denota l'esistenza di una direzione principale di anisotropia (fig. 2.10).

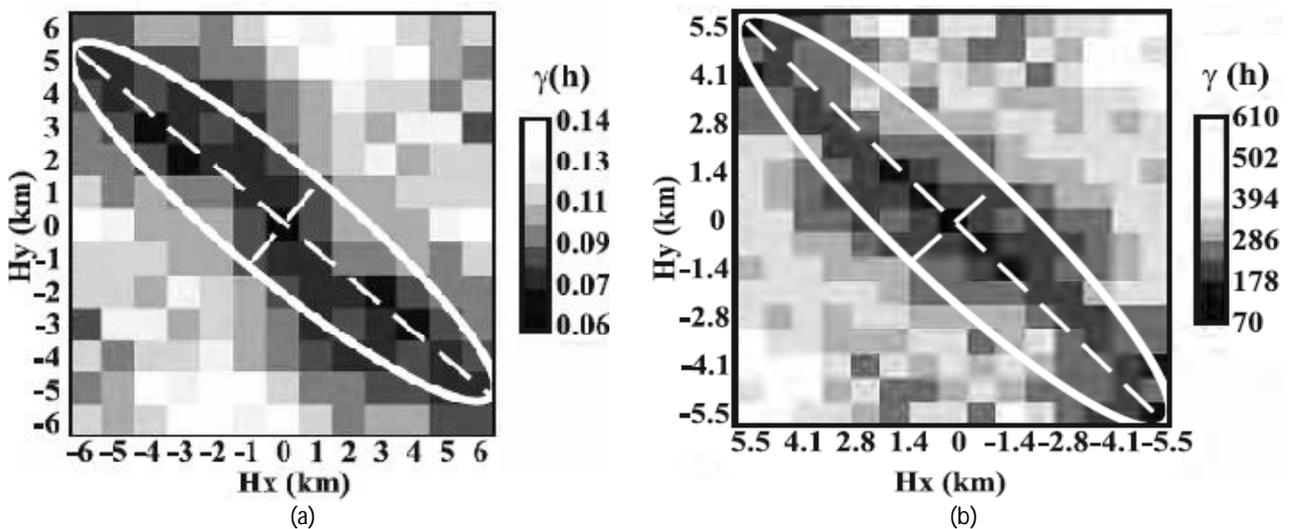


Figura 2.10 – Mappa del variogramma per due tipologie di variabili: concentrazione di a) CO_2 e di b) Rn riferita al bacino del Fucino [Ciotoli et al., 2007]

2.5.5 Modelli di variogramma

Per modello s'intende una funzione matematica continua che rappresenta adeguatamente il grafico del variogramma sperimentale, garantendo il rispetto di tutte le seguenti proprietà:

- a. è una funzione definita positiva

$$\gamma(h) \geq 0 \tag{2.47}$$

b. è una funzione pari

$$\gamma(h) = \gamma(-h) \quad (2.48)$$

c. nel caso di stazionarietà è legato alla covarianza dalla relazione

$$\gamma(h) = Cov(0) - Cov(h) \quad (2.49)$$

d. all'infinito cresce meno velocemente di h^2

$$\lim_{|h| \rightarrow \infty} \frac{\gamma(h)}{|h|^2} = 0 \quad (2.50)$$

e. deve dare origine a combinazioni lineari autorizzate, ovvero le combinazioni lineari nelle quali partecipano per l'applicazione del *Kriging* devono ammettere una varianza finita.

Nella letteratura geostatistica i modelli più comunemente utilizzati nelle applicazioni pratiche sono i seguenti [de Marsily, 1986; Kitanidis, 1997]:

1. Modello sferico

È rappresentato da una curva crescente fino al *range* a (fig. 2.11a). Dopo tale valore le differenze quadratiche medie non cambiano e la curva, raggiunta la varianza (σ^2), rimane costante. È un variogramma caratteristico di dati con aree d'influenza ben sviluppate e buona continuità. La formula che ne regola l'andamento è la seguente:

$$\begin{cases} \gamma(h) = \sigma^2 \left[1.5 \frac{h}{a} - 0.5 \left(\frac{h}{a} \right)^3 \right] \\ \sigma^2 \end{cases} \quad (2.51)$$

in cui la prima equazione è definita per $h \leq a$, mentre la seconda per $h \geq a$. Nella formula (2.51) la varianza σ^2 coincide con il *sill* (C).

2. Modello esponenziale

È descritto da una curva che cresce al crescere delle distanze senza raggiungere il valore della varianza, ovvero lo raggiunge solo asintoticamente (fig. 2.11b). È caratteristico di dati che hanno una limitata area entro cui si manifestano le relazioni d'influenza oppure che mostrano un'elevata distanza di continuità. La sua legge è la seguente:

$$\gamma(h) = \sigma^2 \left(1 - e^{-\frac{h}{a}}\right) \quad (2.52)$$

Anche in tal caso il *sill* (C) coincide con la varianza.

3. Modello gaussiano

È rappresentato da una curva che inizialmente cresce lentamente con la distanza. Da una certa distanza in poi, il tasso di crescita è accelerato, assestandosi sul valore della varianza ad un ben definito valore della scala di correlazione (fig. 2.11c). È caratteristico di dati con una elevata presenza di tendenza che si manifesta a piccola scala e, nello stesso tempo, un alto livello di continuità regionale. La formula che regola tale modello è la seguente:

$$\gamma(h) = \sigma^2 \left(1 - e^{-\frac{h^2}{a^2}}\right) \quad (2.53)$$

Come al solito il *sill* (C) coincide con la varianza; è da precisare che il modello gaussiano è l'unico modello ad avere un andamento parabolico all'origine: ciò sta ad indicare che esso rappresenta una variabile regionalizzata che è abbastanza *smooth* da essere differenziabile.

4. Modello potenza

È rappresentato da una curva che evidenzia una crescita delle differenze quadratiche medie al crescere delle distanze (fig. 2.11d). La formula è la seguente:

$$\gamma(h) = \omega h^s \quad 0 \leq s \leq 2 \quad (2.54)$$

Il comportamento di $\gamma(h)$ all'origine varia con s . Se $s \leq 1$ il variogramma ha una concavità verso il basso; se $s = 1$ il modello potenza coincide con il modello lineare ed il variogramma diventa

una retta di coefficiente angolare ω ; se $s \geq 1$ il variogramma presenta una concavità verso l'alto. Tali modelli sono caratterizzati da una dispersione spaziale non limitata. Non possono essere usati per funzioni aleatorie stazionarie del secondo ordine.

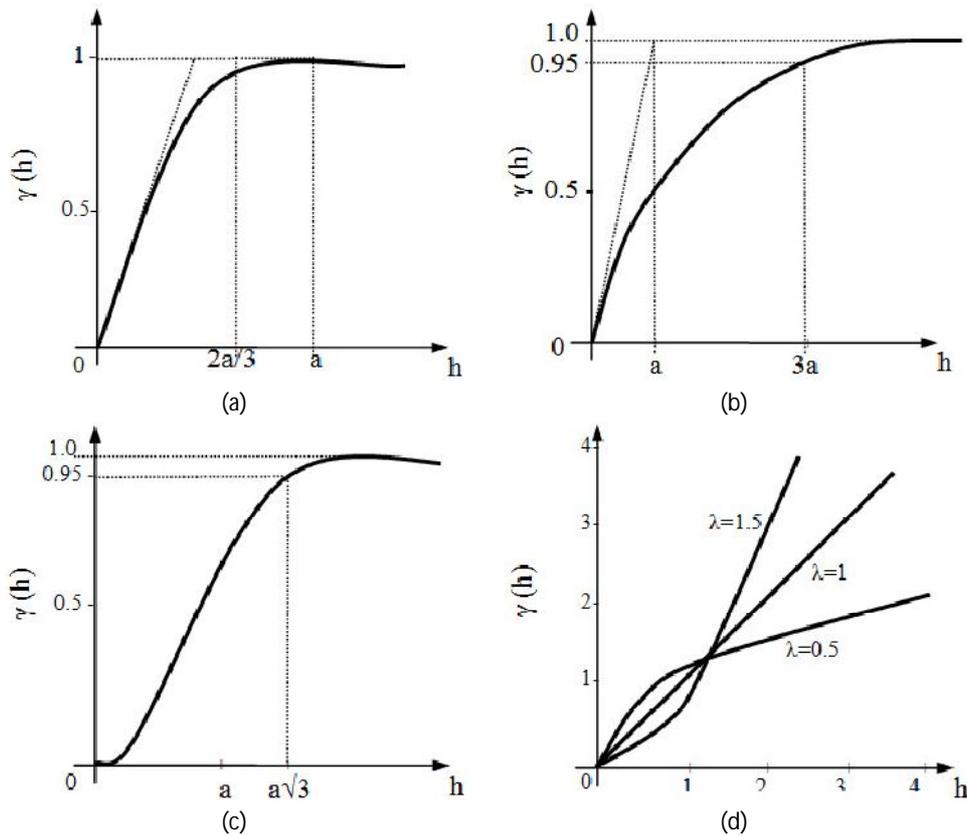


Figura 2.11 – Modelli di variogramma: a) sferico, b) esponenziale, c) gaussiano, d) potenza [de Marsily, 1986]

Capitolo 3 – Metodi di stima puntuale

3.1 Premessa

La stima della variabile oggetto d'indagine in localizzazioni non campionate rappresenta uno dei principali obiettivi in Geostatistica. A tal fine, si ricorre a tecniche di interpolazione stocastica che, utilizzando opportunamente l'informazione derivante dai valori osservati in alcuni punti del dominio, consentono di valutare le dinamiche della caratteristica di interesse in localizzazioni dello stesso dominio in cui non si dispone di alcuna rilevazione.

In Geostatistica, le procedure inferenziali si propongono di risalire alle caratteristiche di un territorio, a partire dall'osservazione di più campioni in differenti localizzazioni. In particolare, di seguito, sarà oggetto di inferenza statistica il valore atteso della funzione aleatoria Z idonea a descrivere un determinato carattere del dominio di riferimento, condizionato dai valori campionati. A differenza di altre tecniche di stima parametrica, la stima puntuale consente di utilizzare l'informazione derivante dall'osservazione della caratteristica in esame nelle localizzazioni campionate per individuare un solo valore, denominato stima, che fornisca una valutazione del parametro incognito, nel caso specifico del valore atteso condizionato.

Nel presente capitolo saranno descritti alcune metodologie di stima puntuale, quali il metodo *Natural Neighbour Interpolation* [Sibson, 1981], il metodo *Inverse Distance* [Shepard, 1968; Burrough & McDonnell, 1998], l'insieme delle tecniche *splines* [Hutchinson, 2005; Mitasova & Mitas, 1993]. Tuttavia sarà data particolare enfasi all'interpolatore *Kriging* [Krige, 1951], l'unico metodo di stima che tiene conto del modello di correlazione spaziale costruito mediante l'approccio variografico descritto nel capitolo precedente.

3.2 Alcune definizioni fondamentali

3.2.1 Stimatore, stima lineare e interpolatori

È stato più volte sottolineato che, assegnato un particolare fenomeno che si manifesti con continuità sul territorio, è possibile definire una funzione aleatoria Z , con i suoi momenti del primo e del secondo ordine, che sia idonea a descriverlo. Fissata una localizzazione u non campionata, l'obiettivo inferenziale, che si pone in questo contesto, consiste nel fornire una stima della variabile aleatoria $Z(u)$, con $u \in D \subseteq \mathbb{R}^d$, utilizzando l'informazione disponibile. Al fine di effettuare inferenza sulla variabile regionalizzata in una localizzazione non campionata, può essere ragionevolmente utilizzata l'informazione derivante dall'osservazione della caratteristica Z su un campione di localizzazioni u_1, u_2, \dots, u_n , selezionate nel dominio D di riferimento. Prima di formalizzare le definizioni di stimatore lineare e stima lineare, è opportuno premettere che le variabili aleatorie $Z(u_1), Z(u_2), \dots, Z(u_n)$, che descrivono tutti i possibili risultati di un campionamento spaziale, sono, in generale, dipendenti [Posa & De Iaco, 2009]. Pertanto, si ribadisce che, a differenza delle tecniche di inferenza classica, in Geostatistica viene abbandonata l'ipotesi di indipendenza tra le variabili aleatorie che compongono il campione. Si ricordi, inoltre, che le osservazioni campionarie, ovvero i valori rilevati in corrispondenza di un insieme finito di localizzazioni relative al campione, u_1, u_2, \dots, u_n del dominio D , sono indicate con $z(u_1), z(u_2), \dots, z(u_n)$, e vengono interpretate come realizzazioni delle variabili aleatorie $Z(u_1), Z(u_2), \dots, Z(u_n)$, rispettivamente. In altri termini, i valori di una caratteristica, osservati nelle localizzazioni spaziali u_1, u_2, \dots, u_n del dominio D , rappresentano una realizzazione parziale di una funzione aleatoria Z definita sullo stesso dominio D .

Assegnata una funzione aleatoria Z , siano $Z(u_1), Z(u_2), \dots, Z(u_n)$ le variabili aleatorie definite in corrispondenza delle localizzazioni campionate u_1, u_2, \dots, u_n nel dominio D ($D \subseteq \mathbb{R}^d$) e sia $Z(u)$ la variabile aleatoria in una prefissata localizzazione $u \in D$ non campionata.

Si definisce stimatore lineare di $Z(u)$, una qualsiasi combinazione lineare $Z^*(u)$ delle variabili aleatorie $Z(u_i)$, con coefficienti reali $\lambda_i(u)$, $i = 1, 2, \dots, n$, come di seguito riportato:

$$Z^*(u) = \sum_{i=1}^n \lambda_i(u) Z(u_i) \quad (3.1)$$

In particolare, i coefficienti $\lambda_i(u)$, $i = 1, 2, \dots, n$, della combinazione lineare (3.1), rappresentano i pesi (o coefficienti ponderatori) da attribuire ai dati campionari, allo scopo di fornire una valutazione appropriata della variabile $Z(u)$ in corrispondenza della localizzazione spaziale u . In questo senso, si giustifica la loro dipendenza dal punto u .

Il valore $z^*(u)$ che la combinazione lineare $Z^*(u)$ assume in corrispondenza dei valori $z(u_1), z(u_2), \dots, z(u_n)$ osservati nelle localizzazioni spaziali campionate u_1, u_2, \dots, u_n , viene denominato stima lineare di $Z(u)$. In altri termini, la stima

$$z^*(u) = \sum_{i=1}^n \lambda_i(u) z(u_i) \quad (3.2)$$

fornisce una valutazione puntuale della variabile aleatoria $Z(u)$ nel punto $u \in D$.

In generale, i metodi di stima per l'interpolazione spaziale dei dati [Lam, 1983; Myers, 1994] si suddividono in deterministici e stocastici. Nei modelli deterministici le variabili di *input* assumono valori fissi. È vero che i risultati (*output*) generati da questi modelli possono tener conto, entro certi limiti, della variabilità e dell'effetto del caso; tuttavia, si tratta sempre di elaborazioni di tipo deterministico, in quanto non si tiene in considerazione l'incertezza associata alle variabili di *input*. Al contrario, i modelli stocastici tengono in considerazione le variazioni (casuali e non) di tali variabili, e, quindi forniscono risultati in termini di probabilità. I principali elementi che caratterizzano i metodi deterministici e stocastici di interpolazione spaziale dei dati sperimentali sono riportati nella tabella 5.1. Uno schema generale delle caratteristiche principali tra metodi deterministici e stocastici è riportato in tabella 5.2, in cui vengono riassunti la maggior parte dei metodi di interpolazione conosciuti [Beretta, 2004].

Modelli deterministici	Modelli stocastici
Individuazione dell'andamento regionale della variabile	Individuazione dell'andamento regionale della variabile
Non è possibile controllare l'attendibilità delle correlazioni al di fuori dei punti di misura	È possibile controllare l'attendibilità delle correlazioni al di fuori dei punti di misura
Non è possibile quantificare l'errore di stima	È possibile quantificare l'errore di stima

Tabella 5.1 – Differenze principali tra metodi deterministici e metodi stocastici

<i>Tipo di metodo</i>	<i>Metodo</i>	<i>Descrizione</i>	<i>Numero di dati disponibili</i>	<i>Vantaggi</i>	<i>Svantaggi</i>
Deterministico	Triangolazione con interpolazione lineare	Metodo basato sull'interpolazione lineare tra tre punti di misura.	Produce buone rappresentazioni con un numero medio (200-1000) ed elevato di dati (> 1000). L'elaborazione dei dati è veloce.	I valori sperimentali misurati sono conservati (interpolatore esatto)	Se i dati sono pochi (< 20) e/o sono disomogeneamente distribuiti si creano andamenti influenzati dalla maglia triangolare e buchi nella cartografia. È necessario correggere manualmente gli andamenti e i buchi. Non sono estrapolati valori al di là dell'intervallo dei dati.
	Regressione polinomiale	I dati sono elaborati con funzioni polinomiali per mostrare l'andamento generale della variabile.	Rappresentazione della tendenza della distribuzione anche con numero molto ridotto di valori (< 10).	Metodo rapido di interpolazione dei dati. Utilizzato per l'analisi della tendenza della superficie interpolata.	Nella ricostruzione sono persi alcuni andamenti locali. Non sono estrapolati valori al di là dell'intervallo dei dati.
	Inverso delle distanze	Il metodo si basa sull'interpolazione di dati in modo che l'influenza di un punto su un altro decresce con la potenza dell'inverso della distanza.	Elaborazioni veloci con numero medio (200-1000) di dati.	Possibile elaborazione rapida di un gran numero di valori. Interpolatore esatto se non viene specificato il fattore di smoothing.	Non produce ricostruzioni affidabili con un numero ridotto di valori sperimentali. Si producono contorni circolari non reali intorno ai punti di misura che possono essere corretti solo parzialmente. Non sono estrapolati valori al di là dell'intervallo dei dati.
	Minima curvatura	Sono prodotti andamenti curvilinei di interpolazione dei dati.	Produce buone rappresentazioni con un numero elevato di dati (> 1000). L'elaborazione dei dati è veloce.	I valori sperimentali sono conservati.	Non sono correttamente rappresentati gli andamenti in presenza di forti variazioni del parametro considerato. Si creano andamenti molto artefatti in area con assenza di dati.
	Natural neighbour	Basandosi sui poligoni di Thiessen vengono effettuate medie pesate dei valori, i cui pesi sono in funzione dell'estensione di ciascun poligono.	Elaborazioni veloci con numero medio (200-1000) di dati.	Produzione di buone ricostruzioni in aree con dati sia scarsi sia sufficienti. Interpolatore esatto.	Non sono estrapolati valori al di là dell'intervallo dei dati. Non sono prodotte ricostruzioni in settori privi di dati.
	Funzioni radiali di base (splines)	Vengono impiegate funzioni analoghe al variogramma nel kriging.	Può produrre buone rappresentazioni anche con un numero limitato di valori (< 250). Con molti dati non si incrementa significativamente il tempo di elaborazione.	L'interpolatore è esatto (se non viene specificato il fattore di smoothing) e i valori originali sono preservati. L'andamento generale della variabile viene rappresentato.	Talora si devono inserire punti ausiliari per poter correggere gli andamenti in vicinanza dei limiti dell'area.
Stocastico	Kriging	Individuazione della struttura di correlazione spaziale della variabile e ricostruzione della distribuzione dei valori per mezzo di stima di pesi nell'interpolazione	Anche con un numero medio (200-1000) di dati si hanno buone rappresentazioni. Con molti dati non si incrementa significativamente il tempo di elaborazione.	Metodo raccomandato di interpolazione. Può estrapolare valori al di fuori dell'intervallo di dati. Produce una stima dell'incertezza della ricostruzione al di fuori dei punti di misura. Interpolatore esatto se non si deve specificare l'effetto pepita.	Elaborazione lenta con molti dati.

Tabella 5.2 – Caratteristiche dei metodi deterministici e stocastici di interpolazione dei dati [Beretta, 2004]

Verranno di seguito trattati alcuni metodi di interpolazione spaziale; per uno studio completo di tutti i metodi di interpolazione, anche dei più semplici, come i poligoni di *Thiessen* e il metodo delle triangolazioni, si può far riferimento a testi specifici riguardanti l'analisi spaziale dei dati [Webster & Oliver, 2007; Lloyd, 2010; Chandra & Ghosh, 2006]

3.3 Metodi di interpolazione deterministici

3.3.1 Interpolazione dei vicini naturali (*Natural Neighbour Interpolation, NNI*)

Questo metodo [Sibson, 1981] parte dalla costruzione di una triangolazione di *Delauney* [Fisher, 2005], in cui i vertici dei triangoli sono i punti campionati in adiacenti poligoni di *Thiessen*. Per determinare il valore in ogni punto u si selezionano i nodi più vicini che vanno a formare un guscio convesso attorno al punto da interpolare u ed i pesi sono proporzionati alle aree dei poligoni vicini. Per tale motivo, questo metodo è stato definito interpolazione dei vicini naturali (*Natural Neighbour Interpolation*). Per ogni vicino viene calcolata l'area A della porzione del poligono di *Thiessen* che è incorporata nel poligono del nuovo punto da calcolare. Queste aree, una volta scalate da 0 a 1, se sommate, danno una somma uguale ad 1. Lo stimatore *NNI* appartiene alla classe (3.1) degli stimatori lineari ed è tale che ognuna di tali aree è un peso uguale a:

$$\lambda_i = \frac{A_i}{\sum_{k=1}^n A_k} \quad i = 1, 2, \dots, n \quad (3.3)$$

Questo significa che, se un punto u_i non è un punto vicino, allora esso non ha un'area in comune con il punto u (*target point*) per il quale vogliamo stimare il valore della variabile Z , ed il suo peso $\lambda_i(u)$ è nullo. Questo interpolatore è continuo e liscio, cioè non ha dei cambiamenti bruschi come la triangolazione o i poligoni di *Thiessen*. Questo interpolatore, inoltre, dà i valori reali nei punti campionati, ovvero è un interpolatore esatto (par. 3.4.4). Esso è semplice da utilizzare, non richiede di specificare nessun parametro, come ad esempio il raggio di ricerca dei vicini, o il numero dei vicini, ed è piuttosto appropriato quando i campioni sono distribuiti in maniera irregolare. Il problema sostanziale di questo metodo è che esso produce dei risultati inaccettabili dove i dati sono molto addensati.

3.3.2 Funzioni inverse della distanza (*Inverse Distance, ID*)

Il metodo *ID* [Shepard, 1968; Burrough & McDonnell, 1998] si basa su una particolare classe di stimatori lineari $Z^*(u)$, dove i pesi attribuiti alle variabili $Z(u_i)$ nelle localizzazioni u_i

campionate, risultano essere inversamente proporzionali a una potenza p ($p \in \mathbb{R}$) [Isaak & Srivastava, 1989] della distanza tra le localizzazioni u_i campionate e il punto u non campionato. Quindi, assegnato uno schema di campionamento spaziale, in cui u_1, u_2, \dots, u_n sono le localizzazioni campionate nel dominio D , lo stimatore ID del campo aleatorio Z in una localizzazione $u \in D$ non campionata, con $u \neq u_i, i = 1, 2, \dots, n$, è costruito come segue:

$$Z^*(u) = \frac{\sum_{i=1}^n \frac{1}{d_i^p} Z(u_i)}{\sum_{i=1}^n \frac{1}{d_i^p}} \quad (3.4)$$

essendo $d_i = d(u_i, u)$ le distanze di ciascuna delle n localizzazioni u_i campionate dal punto u in cui si intende stimare Z . Solitamente, la distanza utilizzata è quella euclidea. Tale stimatore, che fa parte della famiglia (3.1) degli stimatori, è tale che la somma dei pesi

$$\lambda_i = \frac{\left(\frac{1}{d_i^p}\right)}{\sum_{i=1}^n \frac{1}{d_i^p}} \quad i = 1, 2, \dots, n \quad (3.5)$$

sia uguale ad uno; per cui è uno stimatore corretto (par. 3.4.1). Tale metodo, a differenza di molti altri, come il metodo delle triangolazioni, consente di ottenere una stima $Z(u)$ nella localizzazione u non campionata, utilizzando tutti i valori osservati. Inoltre, a differenza del metodo NNI , l' ID non può considerarsi uno stimatore esatto; a tal proposito, si osservi che l'interpolatore (3.4) non è definito se la localizzazione u non campionata coincide con una delle localizzazioni campionate [Posa & De Iaco, 2009].

3.3.3 *Splines*

L'insieme delle tecniche di interpolazione *splines* stima i valori utilizzando una funzione radiale di base matematica che minimizza le curvature della superficie che si adatta sui campioni, fornendo come risultato una superficie addolcita che passa esattamente per i punti di *input*. Esso adatta una funzione matematica ad uno specificato numero di punti più vicini, mentre passa

attraverso i punti di campionamento. Maggiore sarà il numero di campioni utilizzati, maggiore risulta l'influenza dei punti lontani in ogni cella e, di conseguenza, maggiormente addolcita è la superficie interpolata. Esistono differenti tipologie di *splines* [Hutchinson, 2005; Mitasova & Mitas, 1993], basati su differenti funzioni e che producono differenti risultati (fig. 3.1). Questo metodo esatto è abbastanza buono quando si devono interpolare campioni che variano dolcemente l'un l'altro, come l'elevazione, le altezze piezometriche o la concentrazione di inquinanti. È un metodo relativamente semplice e veloce, richiede il settaggio di pochi parametri e non necessita che i dati siano correlati nello spazio. Quando i punti campionati sono numerosi e distribuiti in modo uniforme sul territorio può fornire risultati soddisfacenti, comparabili a quelli ottenuti con tecniche di tipo stocastico [Posa, 1988]. Per questo interpolatore, come per tutti i metodi deterministici, non si quantifica l'errore di stima [Hofierka et al., 2002], non sono ben rappresentati bruschi aumenti dei valori e si creano dati artefatti in aree con assenza di punti campionati, specialmente ai limiti dell'area indagata.

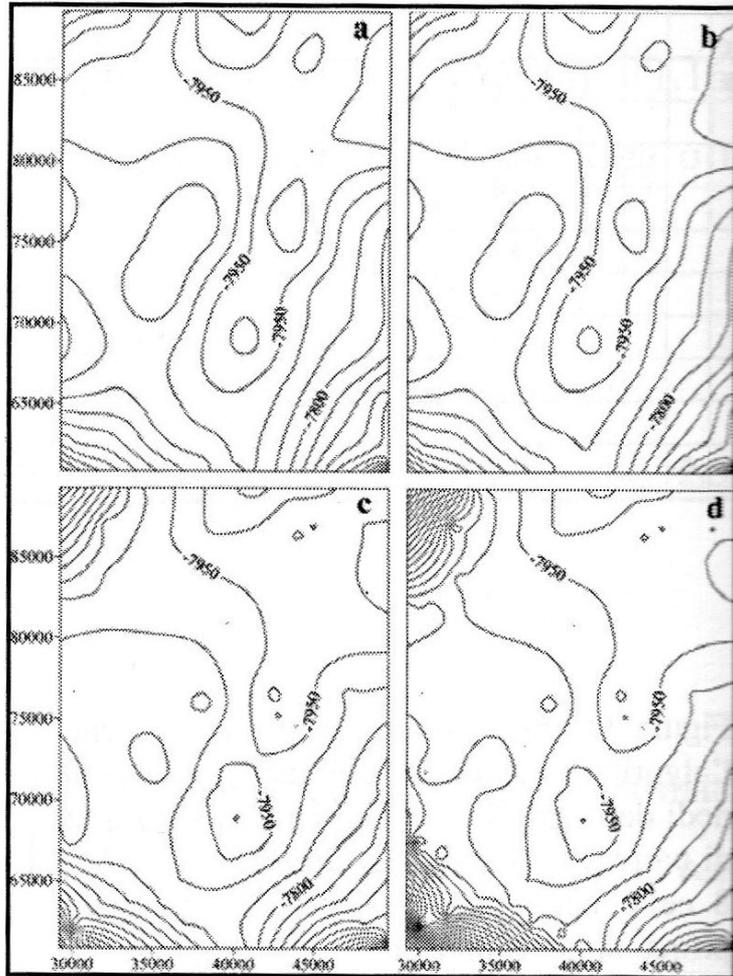


Fig. 3.1 – Mappe ad isolinee relative al tetto della formazione sabbiosa di Lamont del Mississipiano (Ohio, USA). Le mappe sono state elaborate mediante la costruzione di una griglia regolare a partire da dati ottenuti da sondaggi distribuiti irregolarmente nell'area di indagine. Per la stima è stato utilizzato il metodo di minima curvatura (spline) con diverso valore del fattore di tensione [Jones et al., 1986]

3.4 Interpolazione spaziale stocastica: il *Kriging*

3.4.1 Premessa

Nessun metodo di stima, per quanto sofisticato, è in grado di fornire il valore vero di una variabile regionalizzata in una localizzazione non campionata. Tuttavia, un buon metodo di stima dovrebbe utilizzare in modo efficiente l'informazione contenuta nei dati [Posa & De Iaco, 2009]. L'accuratezza di una stima, per una variabile regionalizzata, dipende da diversi elementi, tra i quali:

1. Il numero delle osservazioni e la qualità dei dati, poiché un numero considerevole di osservazioni, sufficientemente attendibili, ovvero non affette da errori di misura o di localizzazione, costituisce una garanzia per l'accuratezza delle stime.
2. La geometria dei punti campionati, ossia la posizione dei punti all'interno del dominio in esame; infatti localizzazioni disposte su una griglia regolare forniscono, in generale, un'informazione sul fenomeno più attendibile rispetto al caso in cui i punti siano irregolarmente distribuiti o raggruppati in una particolare area.
3. La distanza tra i punti, in quanto ci si aspetta che il livello di accuratezza della stima diminuisca all'aumentare della distanza tra il punto in cui si desidera una stima della variabile spaziale e le localizzazioni campionate.
4. La regolarità della variabile spaziale in esame, perché evidentemente la stima risulta essere più efficiente per una variabile regionalizzata caratterizzata da continuità e da un ridotto livello di dispersione.

Il *Kriging* è un metodo di stima che consente di considerare opportunamente tutti i fattori appena elencati.

Nelle applicazioni geostatistiche, quindi, subito dopo la modellazione dei variogrammi e la determinazione dei corrispondenti parametri, è possibile procedere alla stima dei valori nei punti non noti usando il metodo del *Kriging*.

Esso è spesso associato all'acronimo inglese *B.L.U.E.* che sta per *Best Linear Unbiased Estimated*, cioè stimatore ottimale, lineare e corretto [Delhomme, 1978]. È uno stimatore ottimale dato che tende a minimizzare la varianza (σ^2) degli errori di stima, lineare in quanto le sue stime sono

combinazioni lineari pesate di dati noti e, infine, corretto perché cerca di produrre stime tali che la media dell'errore sia nulla.

Il metodo *Kriging* viene denominato stazionario se il problema della stima di una variabile regionalizzata in una localizzazione non campionata viene affrontato assumendo che il campo aleatorio di interesse soddisfi le ipotesi di stazionarietà (stazionarietà del secondo ordine e ipotesi intrinseche); d'altra parte si ricorre al metodo del *Kriging* non stazionario se non si assume che il campo aleatorio d'interesse soddisfi le ipotesi di stazionarietà.

3.4.2 Il *Kriging* semplice (*Simple Kriging, SK*)

In questo caso si assume che la variabile regionalizzata abbia una stazionarietà del secondo ordine, cioè valgono le condizioni (2.15) e (2.16) [Kitanidis, 1997]. Definiamo inoltre un processo a media nulla:

$$Y(u) = Z(u) - m \quad (3.6)$$

che definisce, a sua volta, $E[Y(u)] = 0$.

Lo stimatore lineare nel punto incognito u è rappresentato dalla seguente espressione:

$$Y^*(u) = \sum_{i=1}^n \lambda_i(u) Y(u_i) \quad (3.7)$$

Per poter determinare la migliore stima della variabile regionalizzata in u si impone la condizione di ottimalità ovvero che sia minima la varianza dell'errore di stima (o residuo):

$$\sigma^2 = Var[Y(u) - Y^*(u)] = \min \quad (3.8)$$

la quale, prendendo in esame le definizioni (2.9), (2.19) e di processo a media nulla, si può scrivere come:

$$\begin{aligned}
 \sigma^2 &= E \left\{ \left[Y(u) - \sum_{i=1}^n \lambda_i(u) Y(u_i) \right]^2 \right\} \\
 &= E \left[Y^2(u) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) Y(u_i) Y(u_j) - 2 \sum_{i=1}^n \lambda_i(u) Y(u_i) Y(u) \right] \\
 &= Cov(0) \\
 &+ \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) Cov[Y(u_i), Y(u_j)] - 2 \sum_{i=1}^n \lambda_i(u) Cov[Y(u_i), Y(u)] \quad (3.9)
 \end{aligned}$$

Alla relazione della varianza estimativa in funzione dei pesi $\lambda_i(u)$, si applica la condizione di ottimalità.

Quest'ultima, matematicamente definita come processo di minimizzazione, è generalmente condotta uguagliando a zero le n-derivate parziali del primo ordine rispetto ai pesi:

$$\frac{\partial \sigma^2}{\partial \lambda_i(u)} = 0 \quad (3.10)$$

Tale processo di minimizzazione produce un sistema di n equazioni in n incognite. Ponendo la derivata della varianza estimativa pari a zero, si ottiene:

$$\frac{\partial \sigma^2}{\partial \lambda_i(u)} = 2 \sum_{j=1}^n \lambda_j(u) Cov[Y(u_i), Y(u_j)] - 2 Cov[Y(u_i), Y(u)] = 0 \quad (3.11)$$

pertanto la condizione di ottimalità fornisce un sistema di n equazioni con n pesi incogniti:

$$\sum_{j=1}^n \lambda_j(u) Cov[Y(u_i), Y(u_j)] = Cov[Y(u_i), Y(u)] \quad (3.12)$$

La derivazione di σ^2 permette di calcolare agevolmente la varianza estimativa. Infatti, sostituendola nell'equazione di σ^2 , possiamo esprimere la varianza estimativa nella seguente forma compatta:

$$\sigma^2 = Cov(0) - \sum_{i=1}^n \lambda_i(u) Cov[Y(u_i), Y(u)] \quad (3.13)$$

3.4.3 Il *Kriging* ordinario (*Ordinary Kriging, OK*)

Per questo tipo di interpolatore si assume che per la variabile regionalizzata valgano le ipotesi intrinseche, ossia le condizioni (2.23) e (2.25). Lo stimatore lineare nel punto u è rappresentato dall'espressione (3.1).

La condizione di non distorsione (*unbiased*) della stima può essere espressa come:

$$E[Z(u) - Z^*(u)] = E[Z(u)] - \sum_{i=1}^n \lambda_i(u) E[Z(u_i)] = 0 \quad (3.14)$$

che per la validità dell'ipotesi intrinseca:

$$E[Z(u)] = E[Z(u_i)] = m \quad (3.15)$$

diviene:

$$m \left[1 - \sum_{i=1}^n \lambda_i(u) \right] = 0 \quad (3.16)$$

che è equivalente alla condizione sui ponderatori:

$$\sum_{i=1}^n \lambda_i(u) = 1 \quad (3.17)$$

Per poter determinare la migliore stima della variabile regionalizzata nel punto incognito u si impone la condizione di ottimalità, ovvero che sia minima la varianza dell'errore di stima:

$$\sigma^2 = \text{Var}[Z(u) - Z^*(u)] = \min \quad (3.18)$$

la quale, attraverso le condizioni (2.17) e (2.19), si può scrivere come:

$$\begin{aligned}
\sigma^2 &= E \left\{ \left[Z(u) - \sum_{i=1}^n \lambda_i(u) Z(u_i) \right]^2 \right\} - \left\{ E \left[Z(u) - \sum_{i=1}^n \lambda_i(u) - Z(u_i) \right] \right\}^2 \\
&= E \left[Z^2(u) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) Z(u_i) Z(u_j) - 2 \sum_{i=1}^n \lambda_i(u) Z(u_i) Z(u) \right] - 2m^2 + 2m^2 \\
&= \{E[Z^2(u) - m^2]\} \\
&+ \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) \{E[Z(u_i) Z(u_j) - m^2]\} - 2 \sum_{i=1}^n \lambda_i(u) \{E[Z(u_i) Z(u)] - m^2\} \\
&= Cov(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) Cov[Z(u_i) Z(u_j)] \\
&- 2 \sum_{i=1}^n \lambda_i(u) Cov[Z(u_i), Z(u)] \tag{3.19}
\end{aligned}$$

Alla relazione finale (3.19), che esprime la varianza estimativa in funzione dei pesi $\lambda_i(u)$, si applica la condizione di ottimalità. Quest'ultima, matematicamente definita come processo di minimizzazione, è generalmente condotta uguagliando a zero le n-derivate parziali del primo ordine rispetto ai pesi, come riportato dalla (3.10). Tale processo di minimizzazione produce n+1 equazioni in n incognite. La soluzione di un tale sistema fornisce un set qualunque di n pesi che non garantiscono la condizione di non distorsione. Si deve perciò restringere il campo delle possibili soluzioni a quei set di pesi la cui somma è pari ad uno. Questo problema di ottimizzazione vincolata si risolve usando la tecnica dei moltiplicatori di *Lagrange* [Chapra & Canale, 1985], ovvero aggiungendo alla varianza estimativa il termine $2\mu\lambda$. Pertanto si ottiene:

$$\frac{\partial \sigma^2}{\partial \lambda_i(u)} = 2 \sum_{j=1}^n \lambda_j(u) Cov[Z(u_i), Z(u_j)] - 2Cov[Z(u_i), Z(u)] + 2\mu = 0 \tag{3.20}$$

la condizione di ottimalità, unitamente alla condizione di correttezza, fornisce un sistema di n+1 equazioni con incogniti gli n pesi e il moltiplicatore di *Lagrange* [Chapra & Canale, 1985] μ :

$$\left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i(u) Cov[Z(u_i), Z(u_j)] + \mu = Cov[Z(u_i), Z(u)] \\ \sum_{i=1}^n \lambda_i(u) = 1 \end{array} \right. \quad (3.21)$$

L'equazione (3.20) ci permette di calcolare agevolmente la varianza estimativa. Infatti sostituendo l'equazione (3.20) nell'equazione (3.19) possiamo esprimere la varianza estimativa nella seguente forma compatta:

$$\sigma^2 = Cov(0) - \sum_{i=1}^n \lambda_i(u) Cov[Z(u_i), Z(u)] - \mu \quad (3.22)$$

Nell'ipotesi di stazionarietà del secondo ordine è possibile riscrivere il sistema del *Kriging* e la varianza estimativa in funzione del variogramma. Infatti, per la stazionarietà si ha che:

$$\gamma[Z(u_i), Z(u_j)] = Cov(0) - Cov[Z(u_i), Z(u_j)] \quad (3.23)$$

quindi la condizione di ottimalità, in termini di variogramma, diventa:

$$\frac{\partial \sigma^2}{\partial \lambda_i(u)} = 2 \sum_{j=1}^n \lambda_j(u) [Cov(0) - \gamma[Z(u_i), Z(u_j)]] + 2[Cov(0) - \gamma[Z(u_i), Z(u)] + 2\mu = 0 \quad (3.24)$$

e ora, la condizione di ottimalità, unitamente alla condizione di correttezza, ci permette di ottenere il sistema del *Kriging* in funzione del variogramma:

$$\left\{ \begin{array}{l} \sum_{j=1}^n \lambda_j(u) \gamma[Z(u_i), Z(u_j)] - \mu = \gamma[Z(u_i), Z(u)] \\ \sum_{i=1}^n \lambda_i(u) = 1 \end{array} \right. \quad (3.25)$$

In forma matriciale il sistema *Kriging* scritto in funzione della funzione variogramma si presenta come:

$$Gw = D \Rightarrow \begin{bmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \dots & \gamma_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix} \quad (3.26)$$

in cui G è la matrice dei coefficienti, w il vettore delle incognite e D il vettore dei termini noti; γ_{ij} rappresenta, inoltre, il valore della funzione variogramma $\gamma[Z(u_i), Z(u_j)]$. Si noti che il valore di γ svanisce quando la distanza di separazione è zero. Quindi, come si può vedere nella matrice dei coefficienti, gli elementi lungo la diagonale sono nulli in quanto si riferiscono a due variabili collocate nello stesso punto, cioè a distanza nulla. Conseguentemente, utilizzando l'espressione (3.19) è possibile esprimere anche la varianza estimativa in funzione del variogramma:

$$\begin{aligned} \sigma^2 &= Cov(0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) [\gamma[Z(u_i), Z(u_j)] - Cov(0)] \\ &\quad + 2 \sum_{i=1}^n \lambda_i(u) [\gamma[Z(u_i), Z(u)] - Cov(0)] \\ &= 2 \sum_{i=1}^n \lambda_i(u) \gamma[Z(u_i), Z(u)] - \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) \gamma[Z(u_i), Z(u_j)] \end{aligned} \quad (3.27)$$

L'equazione (3.24) ci permette di calcolare agevolmente la varianza estimativa. Infatti, sostituendo alla (3.27) l'equazione (3.24), possiamo esprimere la varianza estimativa in funzione del variogramma:

$$\sigma^2 = \sum_{i=1}^n \lambda_i(u) \gamma[Z(u_i), Z(u)] - \mu \quad (3.28)$$

3.4.4 Proprietà del *Kriging* stazionario

Il sistema del *Kriging* stazionario presenta una e una sola soluzione se e solo se la matrice di covarianza è definita positiva e, quindi, il suo determinante risulta essere positivo. Se sono valide le ipotesi intrinseche (*Ordinary Kriging*), e, quindi, ha senso definire solo il variogramma,

il sistema del *Kriging* ammette una sola soluzione se la funzione γ è definita positiva in modo condizionato.

Il *Kriging*, che per costruzione è uno stimatore non distorto, è anche un interpolatore esatto. In altri termini, se la localizzazione in cui si intende stimare il valore della variabile regionalizzata coincide con una delle localizzazioni u_1, u_2, \dots, u_n campionate, il sistema *Kriging* fornisce:

1. una stima $z^*(u_i)$ identica al valore osservato $z(u_i)$ nella localizzazione u_i ;
2. una varianza dell'errore nulla.

Il sistema del *Kriging* e la varianza dell'errore dipendono solamente dal modello covarianza o variogramma e dalla geometria relativa alle localizzazioni, ma non dipendono dai valori osservati $z(u_i)$.

Il *Kriging*, rispetto ai metodi di stima precedentemente affrontati, tiene conto della correlazione spaziale presentata dai valori osservati. Inoltre, è possibile considerare eventuali anisotropie presentate dallo stesso fenomeno mediante i modelli covarianza e variogramma.

3.4.5 Equazioni del *Kriging* non stazionario

Nel metodo del *Kriging* non stazionario [Deutsch & Journel, 1998; Wackernagel, 1995; Davis, 2002] si presuppone che il campo aleatorio oggetto d'interesse non soddisfi le ipotesi di stazionarietà, ovvero che la variabile regionalizzata in esame, modellata secondo la (2.4), sia caratterizzata da (par. 2.3.1):

1. una componente deterministica $m(u)$, denominata *trend*, che descrive la variabilità presentata su larga scala della funzione aleatoria Z e che dipende dalla localizzazione u del dominio D ;
2. una componente residua $Y(u)$, che esprime le fluttuazioni casuali presentate da Z e che risulta essere stazionaria del secondo ordine.

In altri termini, si presuppone che il valore atteso del campo aleatorio Z non sia costante, ossia:

$$E[Z(u)] = m(u) \quad (3.29)$$

Tale metodo di interpolazione viene anche denominato *Kriging* con un modello *trend* o *Kriging* universale (*Universal Kriging, UK*); in particolare si assume che il modello *trend* $m(u)$ sia descritto dalla seguente forma polinomiale:

$$m(u) = \sum_{l=0}^L a_l f_l(u) \quad (3.30)$$

dove a_l , $l = 0, 1, \dots, L$, sono coefficienti incogniti ed f_l , $l = 0, 1, \dots, L$, sono monomi. Infatti, sotto opportune ipotesi di regolarità, un generico *trend* può essere sempre approssimato localmente da un'espressione (3.30). In figura 3.2, ad esempio, è rappresentata una realizzazione di un campo aleatorio non stazionario in un dominio bidimensionale ($D \subseteq \mathbb{R}^2$), il cui *trend* è lineare [Posa & De Iaco, 2009]. L'espressione esplicita di un *trend* lineare in termini di coordinate (x, y) risulta essere:

$$m(x, y) = a_0 + a_1 x + a_2 y \quad (3.31)$$

In tal caso, $f_0(u) = 1$, $f_1(u) = x$, $f_2(u) = y$.

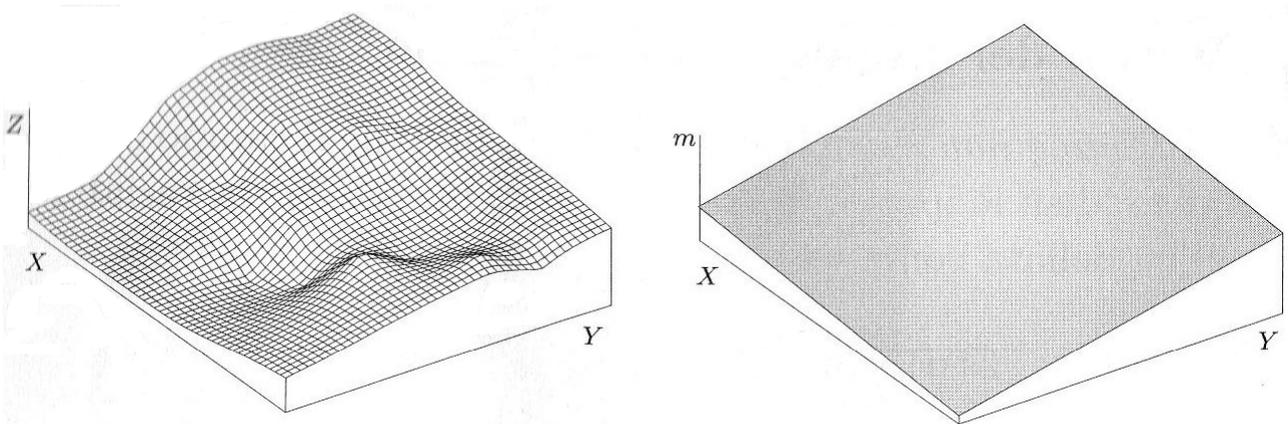


Figura 3.2 – Realizzazione di un campo aleatorio non stazionario e relativo trend [Posa & De Iaco, 2009]

Come già discusso nel caso stazionario, il problema da affrontare consiste nello stimare il campo aleatorio Z in una localizzazione u non campionata, mediante lo stimatore lineare (3.1). I

coefficienti $\lambda_i(u)$ vengono sempre stimati imponendo la condizione di non distorsione e la condizione di ottimalità.

La condizione di non distorsione dello stimatore richiede che sia nullo il valore atteso dell'errore di stima, ossia:

$$\begin{aligned} E[Z^*(u) - Z(u)] &= E[Z^*(u)] - E[Z(u)] \\ &= \sum_{i=1}^n \lambda_i(u) m(u_i) - m(u) = \sum_{i=1}^n \lambda_i(u) \left[\sum_{l=0}^L a_l f_l(u_i) \right] - \sum_{l=0}^L a_l f_l(u) \\ &= 0 \end{aligned} \quad (3.32)$$

da cui segue:

$$\sum_{l=0}^L a_l \left[\sum_{i=1}^n \lambda_i(u) f_l(u_i) - f_l(u) \right] = 0 \quad (3.33)$$

Evidentemente, la suddetta equazione è soddisfatta, per qualunque valore dei coefficienti incogniti a_l , se risulta:

$$\sum_{i=1}^n \lambda_i(u) f_l(u_i) = f_l(u) \quad (3.34)$$

Per cui, la condizione di non distorsione impone $L+1$ vincoli.

È immediato verificare che la condizione di non distorsione introdotta nel *Kriging* stazionario rappresenta un caso particolare della (3.34) ; infatti, se si pone $L=0$, $f_0(u) = 1$, risulta la (3.17). Inoltre, al fine di garantire la condizione di efficienza, occorre determinare quei pesi $\lambda_i(u)$ che, nel rispetto dei vincoli lineari precedentemente individuati, rendano minima la varianza dell'errore, definita come segue:

$$\begin{aligned} \text{Var}[Z(u) - Z^*(u)] &= \sigma^2 \\ &+ \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) \text{Cov}[Z(u_i), Z(u_j)] - 2 \sum_{i=1}^n \lambda_i(u) \text{Cov}[Z(u_i), Z(u)] \end{aligned} \quad (3.35)$$

in cui $\sigma^2 = \text{Var} [Y(u)]$. In altri termini, assegnata la funzione (3.35) che esprime la varianza dell'errore, è necessario individuarne il minimo, rispetto ai coefficienti $\lambda_i(u)$, vincolato alle condizioni di non distorsione riportate nella (3.34). Per cui, ricorrendo sempre al metodo dei moltiplicatori di *Lagrange* [Chapra & Canale, 1985], è sufficiente definire la seguente funzione lagrangiana:

$$\begin{aligned} \mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_L) &= \sigma^2 \sum_{i=1}^n \sum_{j=1}^n \lambda_i(u) \lambda_j(u) \text{Cov}[Z(u_i), Z(u_j)] - 2 \sum_{i=1}^n \lambda_i(u) \text{Cov}[Z(u_i), Z(u)] \\ &+ 2 \sum_{l=0}^L \mu_l(u) \left[\sum_{i=1}^n \lambda_i(u) f_l(u_i) - f_l(u) \right] \end{aligned} \quad (3.36)$$

e determinarne i punti stazionari rispetto ai coefficienti incogniti $\lambda_i(u)$ ed ai moltiplicatori di *Lagrange* $\mu_l(u)$. Si osservi che, la covarianza utilizzata per definire la varianza dell'errore, e quindi, la funzione lagrangiana, è quella relativa ai residui Y del modello (2.4). Inoltre, a differenza del *Kriging* stazionario, il formalismo lagrangiano viene riproposto nel *Kriging* non stazionario considerando $L+1$ equazioni vincolari che richiedono l'introduzione di $L+1$ moltiplicatori di *Lagrange*. I punti di minimo della funzione lagrangiana sono determinati annullando le derivate parziali prime della funzione stessa rispetto alle incognite $\lambda_i(u)$ e $\mu_l(u)$. Di conseguenza, si ottiene il seguente sistema lineare in $n+L+1$ equazioni nelle $n+L+1$ incognite, denominato sistema delle equazioni del *Kriging* non stazionario:

$$\begin{cases} \sum_{j=1}^n \lambda_j(u) \text{Cov}[Z(u_i), Z(u_j)] + \sum_{l=0}^L \mu_l(u) f_l(u_i) = \text{Cov}[Z(u_i), Z(u)] \\ \sum_{i=1}^n \lambda_i(u) f_l(u_i) = f_l(u) \end{cases} \quad (3.37)$$

Questo sistema può essere espresso in forma matriciale, come di seguito riportato:

$$Cw = D \Rightarrow \begin{bmatrix} Cov_{11} & \dots & Cov_{1n} & f_0(u_1) & \dots & f_L(u_1) \\ Cov_{21} & \dots & Cov_{2n} & f_0(u_2) & \dots & f_L(u_2) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Cov_{n1} & \dots & Cov_{nn} & f_0(u_n) & \dots & f_L(u_n) \\ f_0(u_1) & \dots & f_0(u_n) & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_L(u_1) & \dots & f_L(u_n) & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu_0 \\ \vdots \\ \mu_L \end{bmatrix} = \begin{bmatrix} Cov_{10} \\ Cov_{20} \\ \vdots \\ Cov_{n0} \\ f_0(u) \\ \vdots \\ f_L(u) \end{bmatrix} \quad (3.38)$$

in cui C rappresenta la matrice dei coefficienti, w il vettore delle incognite e D il vettore dei termini noti.

L'annullamento delle derivate parziali prime rappresenta, in questo caso, non soltanto la condizione necessaria per la determinazione del minimo, ma anche quella sufficiente, poiché la funzione (3.36), nel caso in cui la covarianza sia definita positiva, risulta essere strettamente convessa; per cui, l'unico punto che annulla le derivate parziali prime è quello di minimo. Pertanto, i pesi $\lambda_i(u)$, ottenuti risolvendo il precedente sistema, consentono di determinare una stima ottimale di $Z(u)$ nella localizzazione u non campionata.

Il *Kriging* universale è utilizzato principalmente in due casi. Il primo, come già accennato, è quando vogliamo predire una caratteristica di un campo aleatorio in cui è presente un *trend* nei dati. Il secondo caso invece si presenta quando i dati sono influenzati da caratteristiche del campo non modellizzabili nel nostro modello. Per esempio, se si sta indagando sulla temperatura presente in una certa regione, è ragionevole pensare che i dati a disposizione siano influenzati dal tipo di ambiente presente nei siti di osservazione (dati presi vicino a città o strade, e quindi in prossimità di asfalto, saranno influenzati in maniera completamente diversa rispetto ad altri dati presi vicino a fiumi o campi irrigati).

Il *Kriging* universale, quindi, è una tecnica molto flessibile che si può adattare alla maggior parte dei casi in cui si debba lavorare con campi aleatori che non presentano un valore atteso costante. A causa della grande varietà di casi in cui il *Kriging* universale può essere utilizzato, alcuni Autori (*Hengl et al.*, 2004; *Monestiez et al.*, 2001; *Deutsch & Journel*, 1998) hanno fatto distinzione tra i vari casi creando le cosiddette tecniche ibride, cioè una famiglia di tecniche che sfrutta l'uso del *Kriging* associato a informazioni presenti nel campo.

3.5 La validazione incrociata (*Cross Validation*)

Una qualunque procedura di stima necessita di uno strumento in grado di misurare la bontà del modello, ritenuto idoneo a descrivere il fenomeno oggetto di studio. Un metodo molto utilizzato è la *Cross Validation* [Davis, 1987; Cressie, 1993; Wackernagel, 1995; Armstrong, 1998], o validazione incrociata, basata sulla tecnica del *jackknife* [Miller, 1994a, 1994b], che usa sostanzialmente i valori misurati come informazione di riferimento rispetto ai valori stimati. Per implementare un confronto robusto si rimuovono, uno alla volta, i punti misurati effettuando la stima in quel punto mancante (*missing value*) con tutti gli altri rimanenti.

In pratica, la *Cross Validation* non prova che il modello di variogramma scelto è corretto, ma semplicemente dice che il modello utilizzato non è grossolanamente errato; inoltre, essa può essere utilizzata in tutti i casi di stazionarietà e non stazionarietà.

Iterando la procedura per tutti i valori noti, si può confrontare il risultato della stima con le informazioni misurate. In questo modo, per ciascun campione, si hanno a disposizione tre valori, ossia:

1. $z(u_i)$, il suo valore reale (misurato);
2. $z^*(u_i)$, il suo valore stimato;
3. $\sigma(u_i)$, la varianza della stima.

I valori stimati tramite l'interpolatore *Kriging* e i rispettivi valori reali vengono paragonati: la differenza tra queste due grandezze è definita residuo della *Cross Validation*.

Essenzialmente, le quantità da prendere in considerazione per la validazione del modello di variogramma scelto sono le seguenti [Davis, 1987]:

1. l'errore medio (*Mean Error, ME*)

$$ME = \frac{1}{n} \sum_{i=1}^n [z^*(u_i) - z(u_i)] \quad (3.39)$$

2. l'errore medio standardizzato (*Mean Standardized Error, MSE*)

$$MSE = \frac{1}{n} \sum_{i=1}^n \frac{[z^*(u_i) - z(u_i)]}{\sigma(u_i)} \quad (3.40)$$

3. la varianza dell'errore standardizzato (*Variance of Standardized Error, VSE*)

$$VSE = \frac{1}{n} \sum_{i=1}^n \frac{[z^*(u_i) - z(u_i)]^2}{\sigma^2(u_i)} \quad (3.41)$$

Le due medie, quella dell'errore e quella dell'errore standardizzato, misurano il grado di correttezza (*degree of unbiasedness*); affinché il modello di variogramma scelto non sia grossolanamente sbagliato, tali grandezze devono assumere un valore prossimo allo zero [Hengl, 2009], proprio per la condizione di non distorsione che deve essere imposta all'interpolatore *Kriging* (par. 3.4.1). Tuttavia, la quantità più interessante da analizzare è la varianza dell'errore standardizzato, la quale corrisponde al rapporto tra le varianze sperimentali e quelle teoriche. In tal caso, il criterio per stimare la bontà di adattamento del modello ai dati sperimentali presuppone che la varianza dell'errore standardizzato deve tendere ad uno [Armstrong, 1998], vale a dire la varianza sperimentale tende al valore di quella teorica.

Capitolo 4 – Caso di studio: analisi sperimentale condotta sul bacino del torrente Turbolo

4.1 Descrizione del bacino oggetto di studio e del campo sperimentale

L'area su cui si è svolta la campagna di misura ricade nel bacino del torrente Turbolo, sottobacino del fiume Crati (Calabria). Il Torrente Turbolo è un affluente di sinistra del medio corso del Crati e nasce dai rilievi della Catena Costiera che si sviluppa ad Ovest degli abitati di S. Giacomo, Cerzeto, Cavallerizzo e Mongrassano; la sezione di chiusura è invece sita poco a sud della stazione ferroviaria di Mongrassano-Cervicati, prima della confluenza col fiume Crati. Nella figura 4.1 viene riportata la localizzazione regionale del bacino del Crati, con in evidenza anche l'intero sottobacino del Turbolo e l'ubicazione del campo sperimentale; nella figura 4.2 si riporta un inquadramento totale del sottobacino e la rappresentazione della sua rete idrografica con le stazioni di misura allocate in esso. Dal punto di vista morfometrico, il bacino idrografico del Torrente Turbolo si sviluppa longitudinalmente da Ovest verso Est, per una lunghezza parallela all'asta principale di 11.5 km, su una superficie di circa 29 km²; inoltre, la pendenza media del bacino, di forma notevolmente allungata, è pari a circa il 28%. L'area di studio esibisce forma planimetrica rettangolare di superficie pari a 2800 m² (70 m x 40 m), con il lato maggiore parallelo al torrente Turbolo, ovvero secondo la direzione est-ovest. Essa ricade nel Comune di Mongrassano Scalo, località La Colombra, a circa 250 m dalla sponda sinistra del torrente.

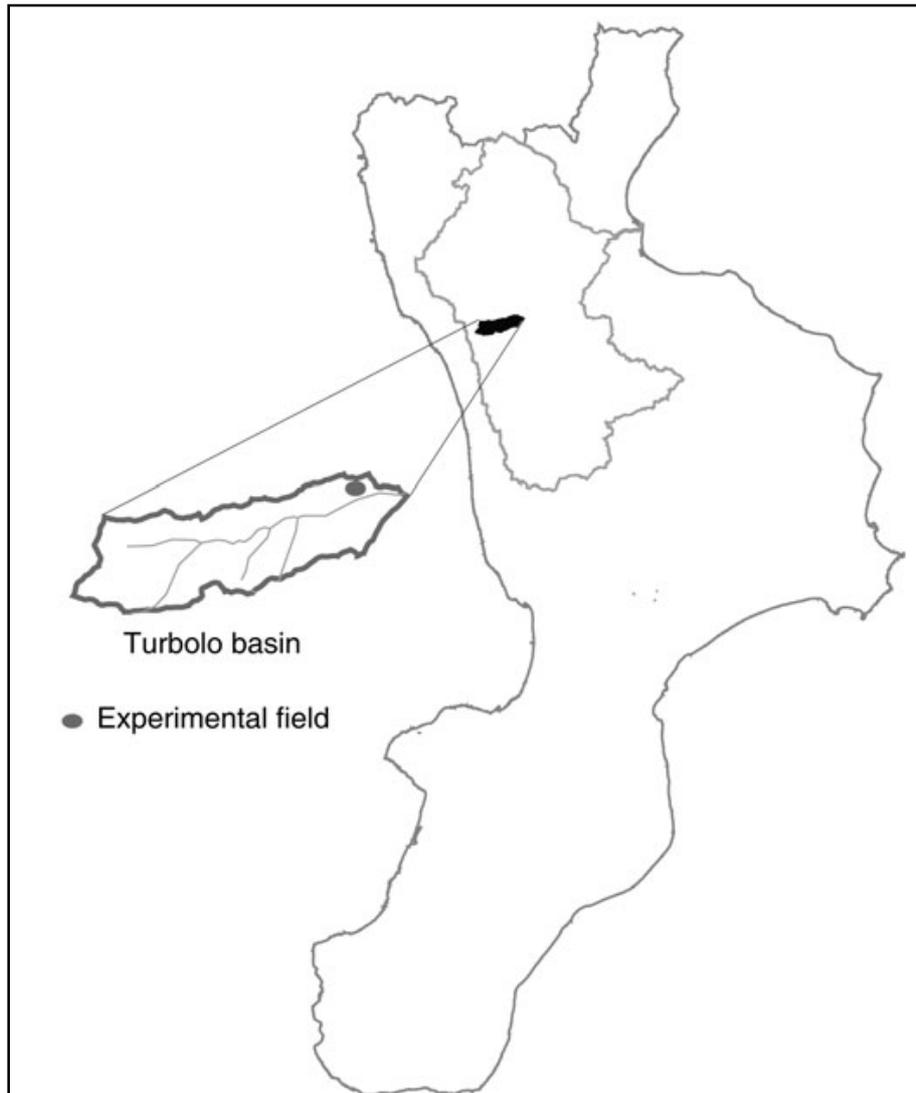


Figura 4.1 – Campo sperimentale all'interno del bacino del Turbolo, sotto-bacino del Crati, Calabria [Fallico et al., 2010]

La stessa area è da alcuni anni oggetto di ricerche particolarmente orientate verso lo studio del comportamento idraulico del terreno, dell'influenza della formazione delle piene ed della produzione dei sedimenti; essa è stata scelta per le sue caratteristiche di esposizione, per l'adeguata pendenza e per la vicinanza al torrente. A tal fine, il sito è stato opportunamente attrezzato, con la collaborazione dell'*Institute for Land and Water Management della Katholieke Universiteit (I.L.W.M.)* di Lovaino del Belgio, con una particolare strumentazione per il monitoraggio in continuo, in più punti e a diverse profondità, del contenuto d'acqua e del potenziale matriciale della parte non satura del terreno, oltre che per il monitoraggio

dell'afflusso meteorico e dello scorrimento superficiale. Sono state anche condotte indagini mirate a caratterizzare, dal punto di vista idraulico, le diverse tipologie geologiche affioranti nel bacino, a esclusione delle formazioni rocciose, con l'obiettivo principale di definire una maggiore attendibilità dei modelli di previsione degli eventi catastrofici [Fallico *et al.*, 2004].

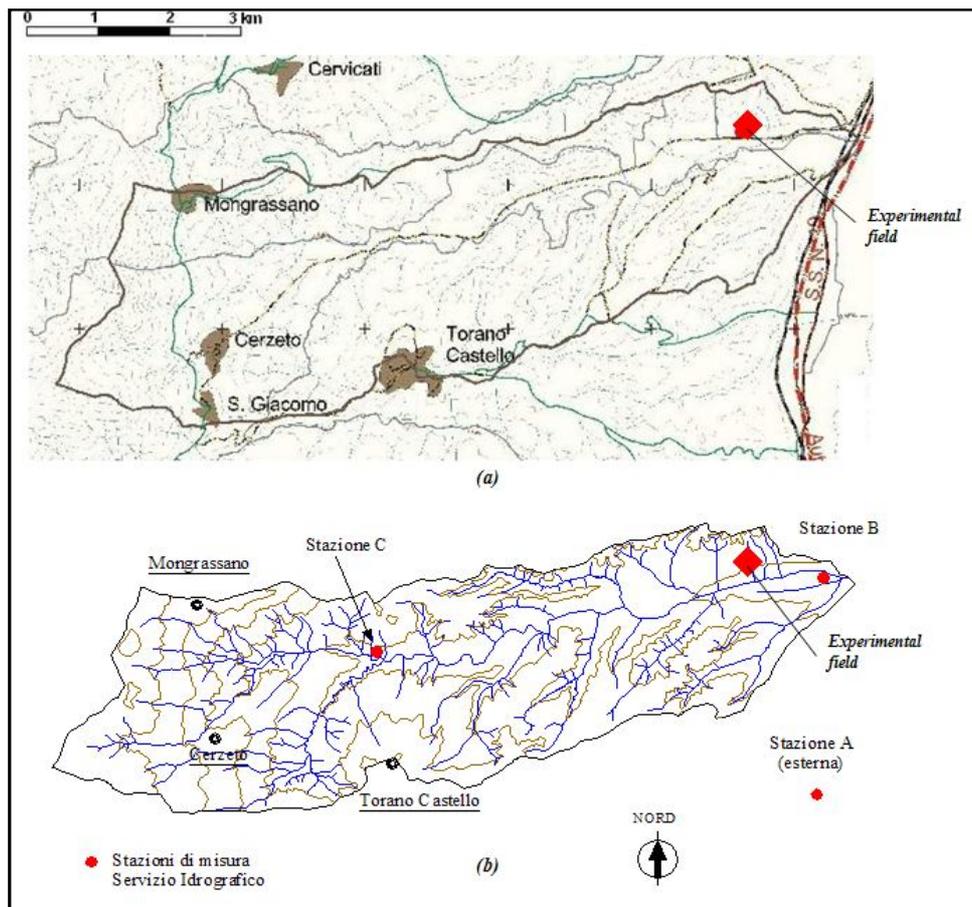


Figura 4.2 – (a) Inquadramento del bacino idrografico del torrente Turbolo e (b) rete idrografica e stazioni di misura allocate in esso

Per la caratterizzazione del non saturo è stato scelto, all'interno di tale area, un campo sperimentale avente una superficie rettangolare di circa 800 m² (40 m x 20 m), con pendenza modesta (6% ÷ 8%) verso il Torrente Turbolo. Su tale area è stata effettuata la campagna di misure atta all'acquisizione di trenta campioni cilindrici di terreno indisturbato, di dimensioni 5 cm x 5 cm (figura 4.3). Ciò ha permesso di definire, quindi, un campione di trenta dati, ciascuno dei quali relativo ad una particolare locazione all'interno dell'area oggetto di studio. Le locazioni

sono state numerate progressivamente da 1 a 30 essendo note per ciascuna di esse le relative coordinate nel piano di riferimento in esame (tabella 4.1). La figura 4.4 fornisce, inoltre, una planimetria dell'area e del campo sperimentale evidenziando le posizioni delle varie locazioni.



(a)



(b)

Figura 4.3 – (a) Posizionamento del provino cilindrico ad una generica locazione e (b) successiva estrazione

Localazione	x (m)	y (m)	Localazione	x (m)	y (m)	Localazione	x (m)	y (m)
loc.1	2.50	3.30	loc.11	13.00	13.00	loc.21	1.00	28.00
loc.2	7.50	3.30	loc.12	17.00	11.00	loc.22	8.00	27.00
loc.3	2.50	9.90	loc.13	2.00	17.00	loc.23	16.00	29.00
loc.4	7.50	9.90	loc.14	10.00	18.00	loc.24	7.00	32.00
loc.5	5.00	6.60	loc.15	18.00	19.00	loc.25	13.00	31.00
loc.6	10.00	5.00	loc.16	5.00	22.00	loc.26	9.00	35.00
loc.7	14.00	4.00	loc.17	15.00	23.00	loc.27	17.00	36.00
loc.8	15.00	7.00	loc.18	3.00	25.00	loc.28	6.00	37.00
loc.9	12.00	7.50	loc.19	12.00	21.00	loc.29	15.00	36.00
loc.10	6.00	12.00	loc.20	18.00	25.00	loc.30	19.00	36.00

Tabella 4.1 – Coordinate delle varie localazioni del campo sperimentale

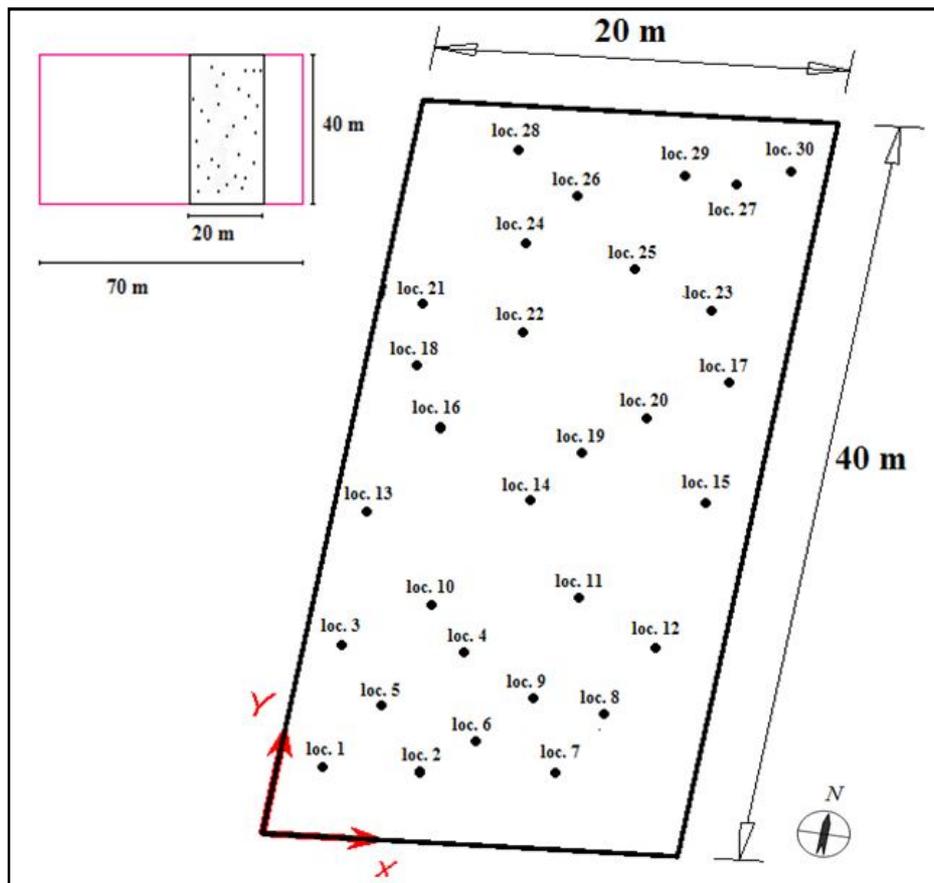


Figura 4.4 – Campo sperimentale e punti di prelievo dei campioni di terreno indisturbati

4.2 Misura del contenuto d'acqua

Le curve di ritenzione idrica sperimentali $\theta = \theta(h)$, relative a ciascuna locazione, sono state ricavate in laboratorio tramite il metodo di misura gravimetrico. Occorre precisare che si è ricorsi alla determinazione delle curve di ritenzione idrica con tale metodo indiretto poiché, a differenza dei metodi diretti applicabili in campo (ad esempio del tensiometro o della sonda T.D.R.), le misure di laboratorio sono più precise. Questo perché la saturazione ottenuta in campo è generalmente minore di quella ottenibile in laboratorio, a causa della struttura dei mezzi porosi la quale presenta i cosiddetti pori ciechi, cioè sistemi che favoriscono l'intrappolamento dell'aria. In particolare, per le misure delle coppie (θ, h) alle varie locazioni è stato utilizzato in laboratorio l'apparato a pressione di *Richards*, di cui si riporta nella figura 4.5 la rappresentazione schematica e in figura 4.6 il particolare della camera a pressione. Tale strumentazione è costituita essenzialmente da una camera a pressione, da una membrana permeabile, da un sistema di regolazione della pressione, da una sorgente di aria compressa e da un sistema di controllo del flusso idrico uscente dalla camera a pressione.

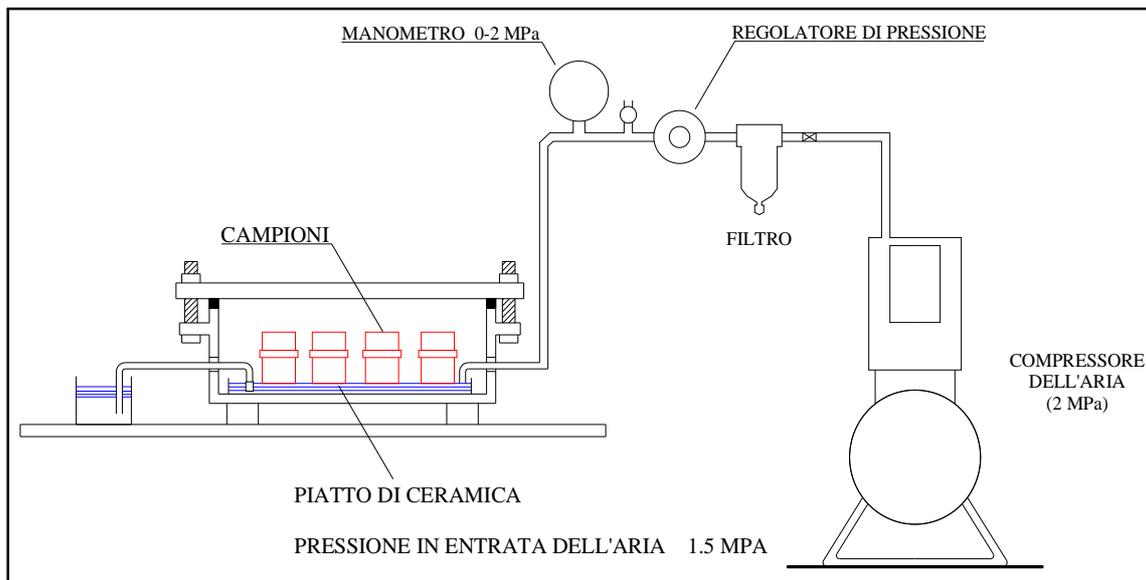


Figura 4.5 – Schema generico dell'apparato a pressione di Richards

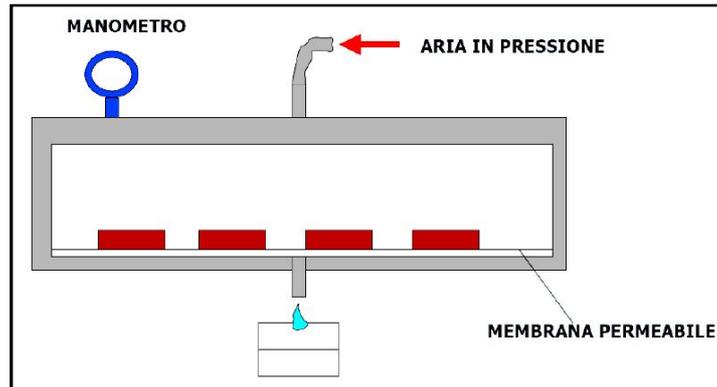


Figura 4.6 – Particolare della camera a pressione di Richards

Il sistema di regolazione della pressione deve essere tale da poter controllare almeno due camere, di cui una per basse ed una per alte pressioni; si possono così condurre più prove nello stesso tempo, riducendo i costi di analisi. Per un sistema di questo tipo occorre un regolatore di pressione generale, con adeguato manometro, agente nel campo da 0 kPa fino a 1800 kPa, corrispondenti a 18000 cm in termini di carico h . Il flusso idrico, uscente dalla camera a pressione, è misurato con una buretta calibrata con graduazioni di 1 mm^3 , collegata a sua volta con un tubicino al foro di uscita della camera a pressione. Le operazioni eseguite con la camera a pressione, su ciascun campione cilindrico, possono essere riassunte nei seguenti passi:

1. i campioni vengono rastremati con un coltellino alla base e alla sommità, in modo da definire un volume noto, pari a quello del cilindro contenitore; alla base viene posta una cuffietta di tessuto di nylon, previamente pesata, con maglie di $30 \mu\text{m}$, in modo da non perdere suolo nel movimento del campione;
2. il piatto poroso, o la membrana di cellulosa, sono portati a saturazione versando preventivamente un passante limoso a $50 \mu\text{m}$ per un'altezza di circa $1/2 \text{ mm}$; tale artificio manterrà il contatto idraulico tra il campione e il piatto (o membrana) senza aumentare l'impedenza al flusso;
3. il campione, essendo in contatto con il piatto poroso, perviene anch'esso a saturazione per imbibizione dal basso e viene pesato; per la saturazione si utilizza preferibilmente una soluzione di CaSO_4 a $0,005 \text{ M}$, poichè l'acqua potrebbe causare la dispersione delle

argille, specialmente se queste sono rigonfiabili o sodiche; la prova si ripete fino a quando il peso non rimane pressoché costante;

4. raggiunta la completa saturazione si iniziano le misure applicando pressioni via via crescenti; per ogni pressione si deve raggiungere l'equilibrio nel campione, cosa che viene rilevata dalla costanza dell'efflusso raccolto nella buretta graduata. Ad equilibrio raggiunto si scarica la pressione e si apre la cella per effettuare la pesatura del campione, non senza aver prima pulito attentamente la cuffietta posta alla sua base. In tale maniera vengono ottenuti per ogni ciclo i pesi umidi lordi P_{lu} ;
5. effettuata la pesata si ripone il campione nella cella, dopo aver bagnato con alcune gocce d'acqua lo straterello di limo in modo da ristabilire il contatto idraulico tra piatto poroso e campione;
6. chiusa la cella si impone una pressione superiore raggiunta con piccoli incrementi, dal momento che una sua rapida imposizione può arrecare danni alla membrana di cellulosa. Le pressioni applicate sono state stabilite nei seguenti passi espressi in kPa: 0.1 (1 cm), 0.3 (3 cm), 1.0 (10cm), 3.2 (32 cm), 10.0 (100 cm), 20.0 (200 cm), 63.1 (631 cm), 251.2 (2512 cm) e 1584.9 (15849 cm);
7. raggiunto l'equilibrio alla massima pressione stabilita il campione viene pesato e, quindi, posto ad asciugare in stufa a 105 °C per non meno di 24 ore;
8. trovato il peso asciutto lordo si risale a quello asciutto netto P_{na} eliminando le tare note (peso del cilindro e della cuffietta); similmente si risale ai pesi netti umidi P_{nu} dai relativi pesi lordi misurati per ogni gradino di pressione h ;
9. si calcolano quindi i vari contenuti idrici $\theta(h)$, in percentuale rispetto al peso secco, con la seguente relazione:

$$\theta(h) = \frac{P_{nu} - P_{na}}{P_{na}} \quad (4.1)$$

Il generico contenuto idrico θ presenterà quindi un potenziale matriciale pari alla pressione h imposta, cambiato di segno.

4.3 Stima della dimensione frattale

La dimensione frattale è stata stimata per ciascuna locazione attraverso le equazioni (1.39), (1.42) e (1.45), o più precisamente, analizzando lo *scaling* relativo ai grafici in scala doppio-logaritmica (θ^* , h/h_{\min}) in cui θ^* viene definito in funzione del rapporto h/h_{\min} . In particolare, il valore del $\log \theta^*$ è definito attraverso le seguenti relazioni:

$$\log \theta^* = \log \frac{\theta}{\theta_S}, \quad (4.2)$$

$$\log \theta^* = \log(\theta + 1 - \theta_S), \quad (4.3)$$

$$\log \theta^* = \log \frac{\theta + A - \theta_S}{A}, \quad (4.4)$$

e cioè, rispettivamente, la (4.2) per il modello di *Tyler & Wheatcraft* [1990], la (4.3) di *Rieu & Sposito* [1991] e la (4.4) del *Pore-Solid Fractal* [*Bird et al.*, 2000]. In ciascun modello esaminato, il valore della dimensione frattale è definito nel *range* dei valori h_{\max} e h_{\min} di *cut-off* rispetto al quale esso risulta essere l'invariante di scala. Studi sperimentali [*Toledo et al.*, 1990; *Millàn & Gonzàles-Posada*, 2005] hanno dimostrato che le curve di ritenzione idrica sono caratterizzate da due tratti di *scaling*, riferiti rispettivamente ad alti e bassi contenuti di acqua. Perciò la dimensione frattale stimata dipende molto dalla grandezza dell'intervallo in cui si definisce, essendo quest'ultimo rappresentativo del fenomeno fisico di ritenzione idrica. In analogia ad altri fenomeni fisici [*Malcai et al.*, 1997], nella ritenzione idrica del suolo un simile comportamento è riscontrabile all'interno di un intervallo di *scaling* abbastanza ridotto, che dipende dalla natura del terreno preso in esame. Nel caso di studio specifico, tale intervallo occupa una posizione intermedia tra i limiti vicini ai più grandi caratterizzanti il carico a saturazione e il contenuto di acqua più piccolo (gravimetrico) vicino allo zero. A tal proposito, il criterio per valutare il valore della dimensione frattale adottato nello studio di *Fallico et al.* [2010] ha permesso di individuare l'esponente $m = 3 - D > 0$, stimato tra i valori degli intervalli di *cut-off* in scala logaritmica espressi dalle relazioni $L_{lower} = \log\left(\frac{h_{min}}{h_{upper}}\right)$ e $L_{upper} = \log\left(\frac{h_{min}}{h_{lower}}\right)$ nel diagramma doppio-logaritmico di cui sopra. La stessa dimensione frattale non è altro che il complemento del coefficiente angolare della retta di regressione che meglio

interpola i punti nell'intervallo predetto, cioè con il più alto coefficiente di determinazione R^2 [Meakin, 1998]. Il suo valore è quindi ottenuto come differenza $D = 3 - m$. Di conseguenza, i rispettivi limiti di *cut-off* dello scaling (legge di potenza associata a ciascuna curva di ritenzione idrica), ossia i valori del carico potenziale minimo e massimo, possono essere dedotti attraverso le seguenti equazioni:

$$h_{upper} = \frac{h_{min}}{e^{L_{lower}}} \quad (4.5)$$

$$h_{lower} = \frac{h_{min}}{e^{L_{upper}}} \quad (4.6)$$

Risulta evidente dalle relazioni (4.5) e (4.6) che il *range* di *cut-off* (h_{lower} , h_{upper}) non risulta direttamente confrontabile con i valori h_{min} e h_{max} definiti nella relazione (1.43). Inoltre, è importante sottolineare, che per ciascun modello frattale esaminato, la stima della dimensione frattale non dipende dal parametro h_{min} .

Per quanto riguarda il parametro A , secondo gli studi di Wang *et al.* [2005], esso assume valori compresi tra 0.25 e 0.59 per un suolo caratterizzato da sabbia limosa; nello studio di Fallico *et al.* [2010], effettuato sul bacino sperimentale del Turbolo, il parametro A è stato posto rispettivamente pari a 0.30, 0.45 e 0.60, essendo quest'ultimi rappresentativi del tipo di terreno oggetto d'indagine. Alcuni risultati di tale studio vengono riportati nella tabella 4.2 e nella figura 4.7.

Model	$D_{f(1^\circ \text{ range})}$	R_1^2	$D_{f(2^\circ \text{ range})}$	R_2^2
TW	2.941 ± 0.003	0.999	2.779 ± 0.043	0.963
RS	2.981 ± 0.001	0.992	2.952 ± 0.013	0.930
PSF, $A = 0.30$	2.913 ± 0.004	0.988	2.466 ± 0.027	0.997
PSF, $A = 0.45$	2.951 ± 0.002	0.991	2.837 ± 0.036	0.953
PSF, $A = 0.60$	2.966 ± 0.001	0.992	2.902 ± 0.020	0.940

Tabella 4.2 – Valori della dimensione frattale per i modelli di Tyler & Wheatcraft (TW), di Rieu & Sposito (RS) e per il modello Pore-Solid Fractal (PSF) con parametro $A = 0.30, 0.45$ e 0.60 . I valori delle dimensioni frattali stimate si riferiscono agli intervalli di *cut-off* rappresentati dagli estremi $h_{lower} = 3.1$ cm e $h_{upper} = 631$ cm per la locazione 10. Sono riportati, inoltre, i coefficienti di determinazione R^2 relativi alla stima della dimensione frattale per ciascun modello esaminato [Fallico *et al.*, 2010]

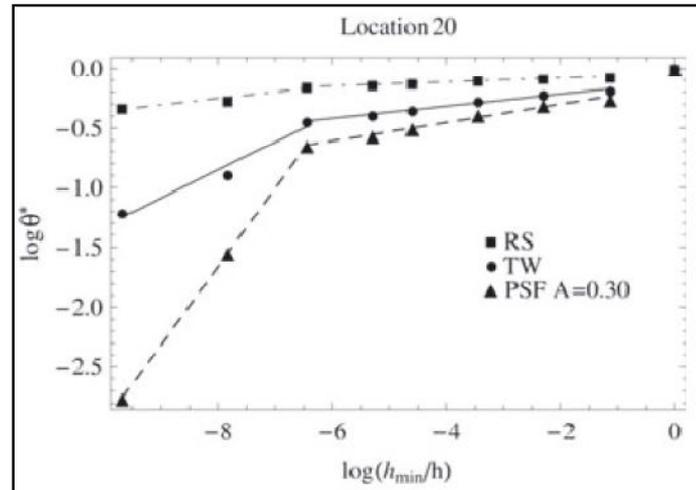


Figura 4.7 – Analisi dello scaling lineare riferito alla locazione 20. Nel grafico doppio-logaritmico θ^* è funzione di h_{min}/h per ciascun modello frattale esaminato. Le linee a tratto continuo e non evidenziano i punti selezionati per la determinazione del range di cut-off in cui è massimo il coefficiente di determinazione R^2 e la dimensione frattale rappresenta l'invariante di scala [Fallico et al., 2010].

Considerando che i modelli frattali delle curve di ritenzione idrica considerati nel presente studio sono rappresentati dai tre precedentemente esposti, *Tyler & Wheatcraft* [1990], *Rieu & Sposito* [1991], *Pore-Solid Fractal* [Bird et al., 2000], considerando ancora che per quest'ultimo il valore del parametro A è stato assunto pari a 0.30, 0.45, 0.60, i set di valori della dimensione frattale di cui si dispone sono complessivamente dieci⁽¹⁾ (tabella 4.3). Cinque sono relativi all'intervallo di *cut-off* superiore (alti contenuti di acqua) e gli altri a quello inferiore (bassi contenuti di acqua). Per ogni campione, quindi, è stata effettuata un'analisi Geostatistica con lo scopo di definire la variabilità spaziale della dimensione frattale nell'area oggetto di studio.

Loc.	X (m)	Y (m)	TW (I)	RS (I)	PSF030 (I)	PSF045 (I)	PSF060 (I)	TW (II)	RS (II)	PSF030 (II)	PSF045 (II)	PSF060 (II)
1	2.5	3.3	2.93905	2.97868	2.90699	2.94603	2.96194	2.76567	2.94308	2.32914	2.80834	2.88401
2	7.5	3.3	2.89426	2.96688	2.84677	2.914	2.94012	2.83109	2.9616	2.66387	2.8757	2.92292
3	2.5	9.9	2.95097	2.98243	2.92376	2.95465	2.96833					
4	7.5	9.9	2.94289	2.97991	2.91562	2.95002	2.96446	2.75856	2.94114	2.38224	2.80634	2.88129
5	5	6.6	2.93136	2.97595	2.89586	2.93937	2.95717	2.77475	2.94547	2.37929	2.81658	2.88889
6	10	5	2.97785	2.99311	2.97305	2.98341	2.98802	2.75685	2.94564	2.65401	2.84186	2.89627
7	14	4	2.97048	2.99142	2.96613	2.97924	2.98503	2.76151	2.95098	2.69601	2.85826	2.90668
8	15	7	2.96323	2.98921	2.95733	2.97388	2.98118	2.7471	2.94897	2.66648	2.84985	2.90206
9	12	7.5	2.94771	2.98388	2.93028	2.95932	2.97126	2.76084	2.94889	2.57095	2.83825	2.89858
10	16	12	2.94099	2.98123	2.91288	2.95117	2.96603	2.77945	2.95187	2.46624	2.83724	2.90162
11	13	13	2.95774	2.98649	2.94514	2.96689	2.97628	2.76015	2.94684	2.59581	2.83645	2.89586
12	17	11	2.94522	2.98183	2.91732	2.95311	2.96725	2.77151	2.94685	2.41118	2.82181	2.89185
13	2	17	2.94906	2.98455	2.93436	2.96133	2.97258	2.7723	2.95159	2.63036	2.85036	2.90493
14	10	18	2.96944	2.98974	2.96099	2.97563	2.98228	2.77014	2.9443	2.63871	2.83703	2.89343
15	18	19	2.95938	2.98667	2.94417	2.96686	2.97642	2.74229	2.94198	2.47543	2.81466	2.88449
16	5	22	2.97069	2.99005	2.96046	2.97586	2.98263	2.74668	2.93879	2.53443	2.81312	2.88063
17	18	25	2.96411	2.98825	2.95232	2.97121	2.97937	2.75333	2.9436	2.55972	2.82575	2.88935
18	3	25	2.96843	2.98945	2.95942	2.97481	2.98173					
19	12	21	2.96513	2.98852	2.95447	2.97218	2.97996	2.77507	2.94654	2.63034	2.84014	2.89665
20	15	23	2.94945	2.98339	2.92313	2.95682	2.96995	2.76465	2.9467	2.34466	2.8179	2.89064
21	1	28	2.94482	2.98212	2.93059	2.9586	2.97049	2.81537	2.95384	2.67019	2.85987	2.91007
22	8	27	2.95353	2.98466	2.94439	2.96642	2.97594	2.76514	2.94457	2.57256	2.82912	2.89134
23	16	29	2.96225	2.98669	2.94499	2.9671	2.97653	2.78924	2.94447	2.56878	2.82857	2.89109
24	7	32	2.94236	2.97912	2.91317	2.94826	2.96313	2.79592	2.94456	2.54898	2.82661	2.89061
25	13	31	2.95828	2.986	2.94471	2.96615	2.97559	2.80297	2.95105	2.67586	2.85518	2.90579
26	9	35	2.90098	2.96295		2.89643	2.93053	2.83621	2.95379	2.10942	2.82431	2.90078
27	17	36	2.94437	2.98078	2.918	2.95186	2.96588	2.78433	2.94806	2.48745	2.82936	2.89521
28	6	36	2.9534	2.98308	2.92832	2.95772	2.96999	2.8273	2.95244	2.61589	2.85057	2.9059
29	15	36	2.95757	2.98483	2.93233	2.96119	2.97277	2.80408	2.94771	2.49275	2.8289	2.89469
30	19	36	2.95262	2.98181	2.91098	2.95038	2.96629					

Tabella 4.3 – Set di campioni della variabile dimensione frattale, riferita ai modelli frattali esaminati (TW, RS, PSF con $A=0.30$, PSF con $A=0.45$ e PSF con $A=0.60$), definita per alti (I) e bassi (II) contenuti di acqua.

⁽¹⁾I valori mancanti sono considerati anomali (*outliers*) e quindi non corretti per intraprendere uno studio geostatistico [Fallico et al., 2010]

4.4 Analisi dei risultati

4.4.1 Exploratory Data Analysis (EDA)

Vengono di seguito descritti i risultati riferiti alla fase preliminare dell'analisi Geostatistica, cioè l'analisi esplorativa dei dati, per ciascun campione della variabile dimensione frattale, con lo scopo di approfondire la conoscenza del fenomeno indagato mediante la semplice indagine sulle distribuzioni statistiche. Tale fase comprende essenzialmente un'analisi descrittiva e un'analisi grafica del campione che permette di captare le caratteristiche salienti dello stesso, individuare particolari informazioni su di esso come, ad esempio, la presenza di *outliers* e/o di valutare di quanto la distribuzione dei dati si discosta da una distribuzione normale. La tabella 4.4 rappresenta i valori dei parametri di forma, di posizione e di variabilità dei campioni esaminati mentre le figure dalla 4.8.1 alla 4.8.10 riportano la descrizione grafica della variabile dimensione frattale per ciascun modello e per alti e bassi contenuti di acqua.

E' da precisare che il numero di classi per la definizione degli istogrammi di frequenza è stato scelto pari a cinque, definito, cioè, come media delle tre formule di *Sturges* [1926], di *Larson* [1969] e di *Scott* [1978] (tabella 4.5).

I risultati rappresentati nella tabella 4.4 dimostrano che il modello di *Rieu & Sposito* [1991], riferito ad alti e a bassi contenuti di acqua, presenta un valore medio della dimensione frattale pari rispettivamente a 2.9836 e 2.9476, e, quindi, per tale modello la variabile risulta avere i valori più alti. Le medie dei campioni degli alti contenuti di acqua sono comprese in un *range* di valori tra 2.9327 e 2.9836, mentre per bassi contenuti di acqua tale *range* è molto più ampio, limitato, cioè, dai valori 2.5322 e 2.9476; questo rappresenta una prima considerazione che dimostra una consistente eterogeneità dei valori della dimensione frattale per bassi contenuti di acqua. Per tutti i modelli frattali, ad eccezione del modello *Pore-Solid Fractal* [2000] con $A=0.30$ e per bassi contenuti d'acqua, la mediana è molto vicina al valore medio: di conseguenza è da aspettarsi che tali campioni possano essere interpretati con una legge di distribuzione normale; tale ipotesi può essere confermata (ma non verificata) attraverso l'utilizzo dei grafici *Q-Q* e *NPP* (fig. 4.8 b, c) in quanto la distribuzione dei punti della variabile nei grafici appare allineata approssimativamente lungo una retta.

Modello	Min	Max	Media	Mediana	Errore Standard	Deviazione Standard	Varianza	Curtosi	Asim.	
I	TW	2.8943	2.9779	2.9509	2.9530	$3.3 \cdot 10^{-3}$	$1.8 \cdot 10^{-2}$	$3.3 \cdot 10^{-4}$	3.5408	-1.5992
	RS	2.9630	2.9931	2.9836	2.9842	$1.18 \cdot 10^{-3}$	$6.5 \cdot 10^{-3}$	$4.22 \cdot 10^{-5}$	3.2441	-1.5044
	PSF _{0.30}	2.8468	2.9731	2.9327	2.9323	$4.81 \cdot 10^{-3}$	$2.59 \cdot 10^{-2}$	$6.72 \cdot 10^{-4}$	2.9248	-1.1723
	PSF _{0.45}	2.8964	2.9834	2.9583	2.9603	$3.32 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$	$3.31 \cdot 10^{-4}$	4.3097	-1.7475
	PSF _{0.60}	2.9305	2.9880	2.9706	2.9719	$2.23 \cdot 10^{-3}$	$1.22 \cdot 10^{-2}$	$1.49 \cdot 10^{-4}$	3.7734	-1.6387
II	TW	2.7423	2.8362	2.7782	2.7715	$5.06 \cdot 10^{-3}$	$2.63 \cdot 10^{-2}$	$6.93 \cdot 10^{-4}$	-0.1022	0.8807
	RS	2.9388	2.9616	2.9476	2.9468	$9.22 \cdot 10^{-4}$	$4.79 \cdot 10^{-3}$	$2.29 \cdot 10^{-5}$	1.4767	0.8048
	PSF _{0.30}	2.1094	2.6960	2.5322	2.5688	$2.64 \cdot 10^{-2}$	0.1373	$1.89 \cdot 10^{-2}$	1.9191	-1.2707
	PSF _{0.45}	2.8063	2.8757	2.8342	2.8294	$3.30 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	$2.94 \cdot 10^{-4}$	-0.1562	0.4601
	PSF _{0.60}	2.8806	2.9229	2.8961	2.8952	$1.83 \cdot 10^{-3}$	$9.53 \cdot 10^{-3}$	$9.08 \cdot 10^{-5}$	1.001	0.6895

Tabella 4.4 – Parametri principali della statistica descrittiva per ciascun modello frattale e per alti (I) e bassi (II) contenuti di acqua

Numero di classi	Modello frattale									
	TW (I)	RS (I)	PSF _{0.30} (I)	PSF _{0.45} (I)	PSF _{0.60} (I)	TW (II)	RS (II)	PSF _{0.30} (II)	PSF _{0.45} (II)	PSF _{0.60} (II)
Sturges [1926]	5.87	5.87	5.83	5.87	5.87	5.72	5.72	5.72	5.72	5.72
Larson [1975]	4.25	4.25	4.22	4.25	4.25	4.15	4.15	4.15	4.15	4.15
Scott [1979]	4.14	4.12	4.29	4.26	4.20	3.07	4.09	3.67	3.49	3.82

Tabella 4.5 – Calcolo del numero di classi per la definizione degli istogrammi di frequenza

Per quanto riguarda gli indici che valutano la forma della distribuzione dei campioni (Curtosi e asimmetria), tutti i set di dati riferiti ad alti contenuti di acqua presentano un valore di Curtosi positivo e, quindi, una tipologia di distribuzione più appuntita (leptocurtica) rispetto a quella normale; per bassi contenuti di acqua, invece, i campioni del modello di *Tyler & Wheatcraft* [1990] e del modello *Pore-Solid Fractal* [2000] con $A=0.45$, hanno un valore di Curtosi negativo

e, per questo, presentano una distribuzione più schiacciata verso il basso (platicurtica) rispetto alla distribuzione normale. La distribuzione della variabile dimensione frattale per tutti i campioni con alti contenuti di acqua presenta una forma con coda allungata verso i valori più bassi e per questo essi sono caratterizzati da un coefficiente di asimmetria negativo; diversamente la distribuzione dei campioni riferiti a bassi contenuti di acqua, ad eccezione del modello *Pore-Solid Fractal* con $A=0.30$, hanno una forma con coda allungata verso destra e, quindi, un coefficiente di asimmetria positivo.

Per effettuare la verifica dell'ipotesi circa la legge di distribuzione normale che descrive i set di campioni della variabile dimensione frattale, sono stati presi in considerazione i test di adattamento di *Kolmogorov-Smirnov* [Kolmogorov, 1933; Smirnov, 1939], di *Anderson-Darling* [1952] e del chi-quadrato [Snedecor & Cochran, 1974], i cui risultati sono riportati in tabella 4.6. Esaminando tali valori e considerando, inoltre, un livello di significatività $\alpha = 0.05$, i campioni della variabile dimensione frattale, per ciascun modello frattale, per alti e bassi contenuti d'acqua, possono essere interpretati con una legge gaussiana poiché il valore della statistica (D , $A^2 \cdot \chi^2$) risulta sempre essere minore rispetto al valore critico (V_{crit}), qualsiasi sia il test di adattamento preso in considerazione.

La valutazione di eventuali *outliers* è stata preliminarmente effettuata attraverso l'utilizzo della rappresentazione grafica della variabile (fig. 4.8.1 – 4.8.10); il test di *Dixon* [1953] per l'individuazione dei valori anomali, conferma le ipotesi fatte circa i valori anomali presenti in ciascun campione della dimensione frattale, in quanto il valore della funzione discriminante (di seguito chiamata Q), per un livello di significatività $\alpha = 0.1$, risulta essere maggiore rispetto al valore critico Q_{crit} (tabella 4.7). Nello specifico, per valori della dimensione frattale riferiti ad alti contenuti di acqua, gli *outliers* presenti corrispondono alle locazioni 2 e 26; diversamente, per bassi contenuti di acqua si è in presenza di una situazione differente: il modello di *Tyler & Wheatcraft* [1990] non è caratterizzato da valori anomali; gli *outliers* per i rimanenti modelli frattali coincidono con la locazione 2 ad eccezione del modello *Pore-Solid Fractal* [2000] con $A=0.30$ per il quale la locazione contraddistinta da un valore anomalo è la 26.

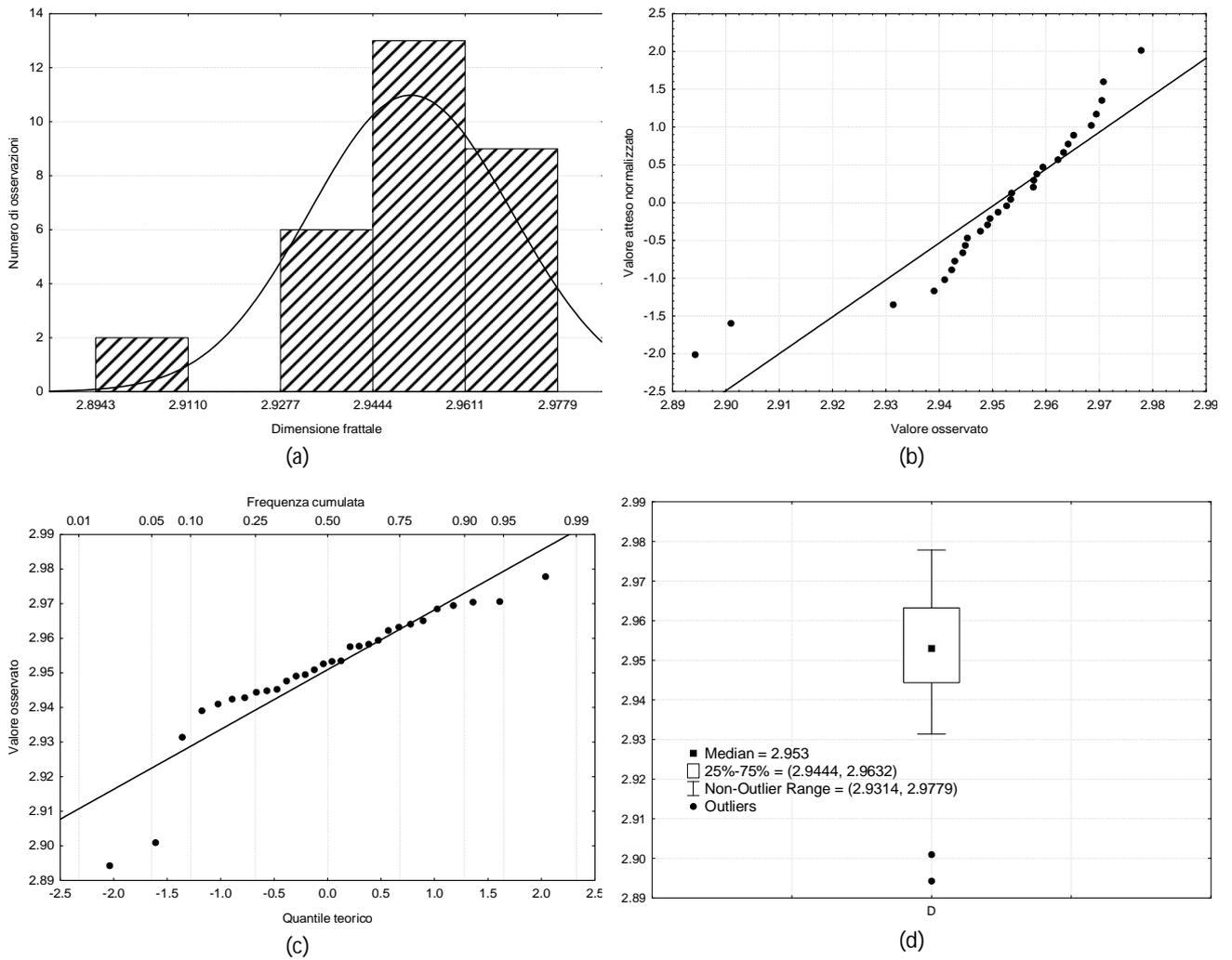


Figura 4.8.1 – Rappresentazione grafica della variabile dimensione frattale per il modello di Tyler & Wheatcraft [1990] e per alti contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

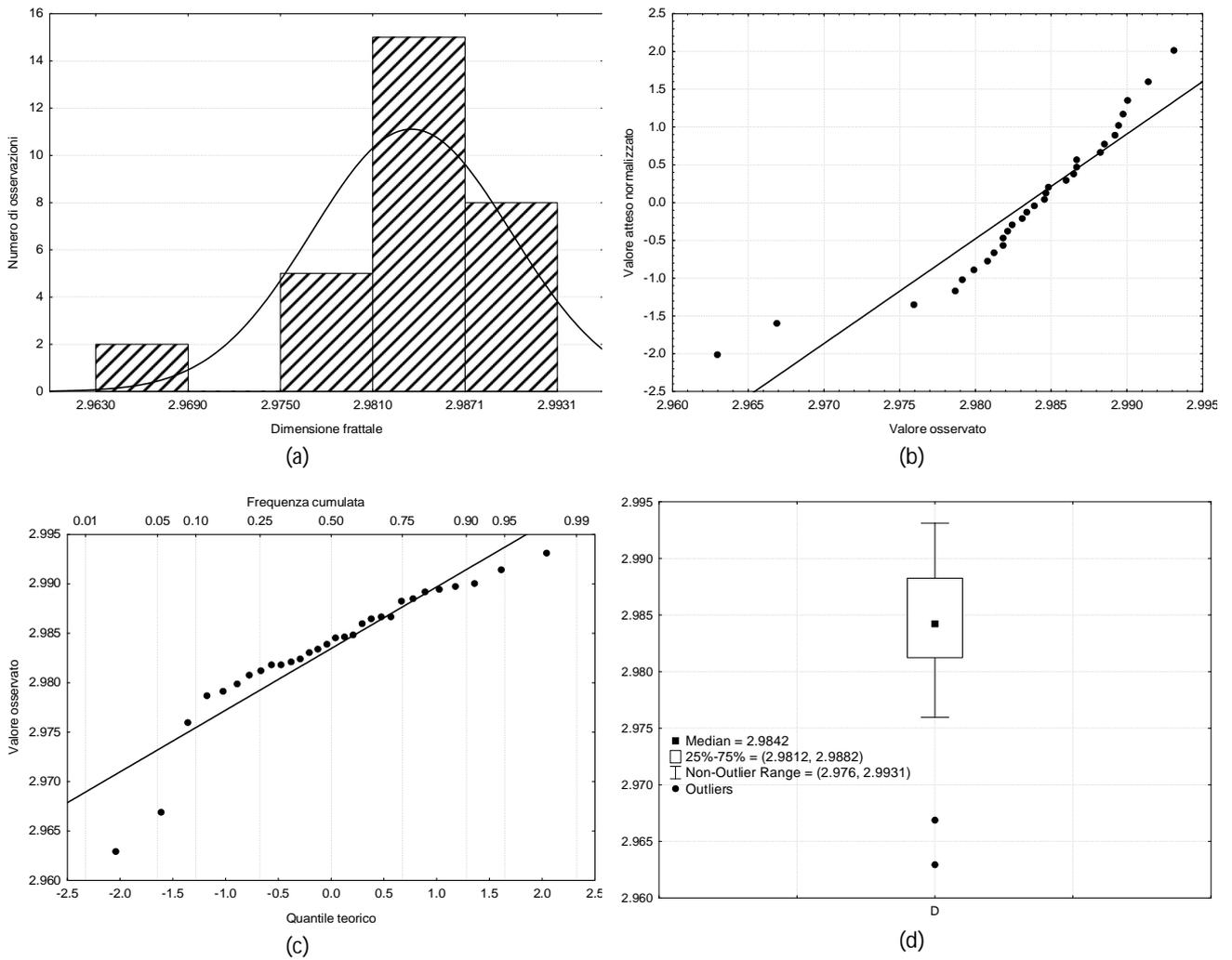


Figura 4.8.2 – Rappresentazione grafica della variabile dimensione frattale per il modello di Rieu & Sposito [1991] e per alti contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

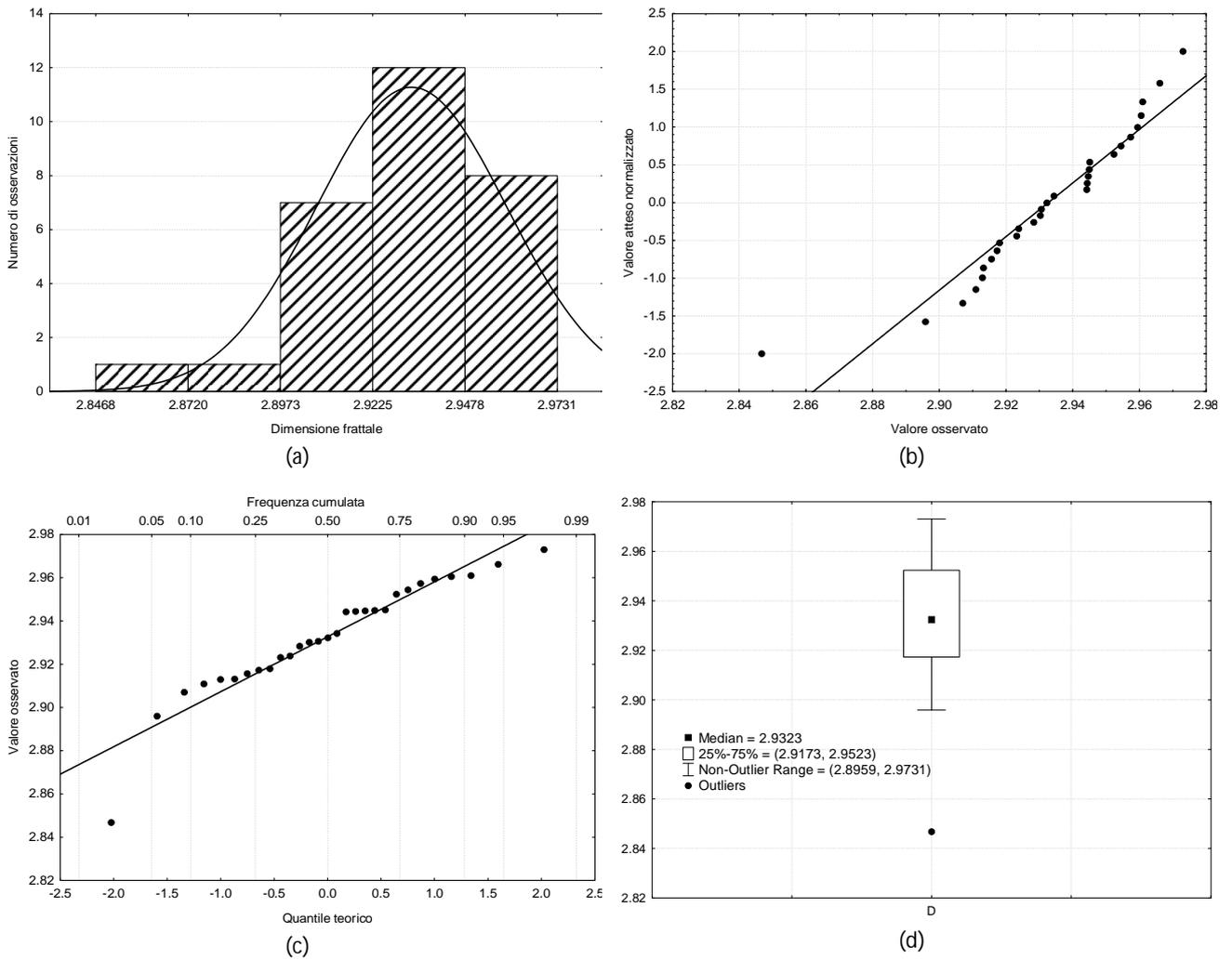


Figura 4.8.3 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per alti contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

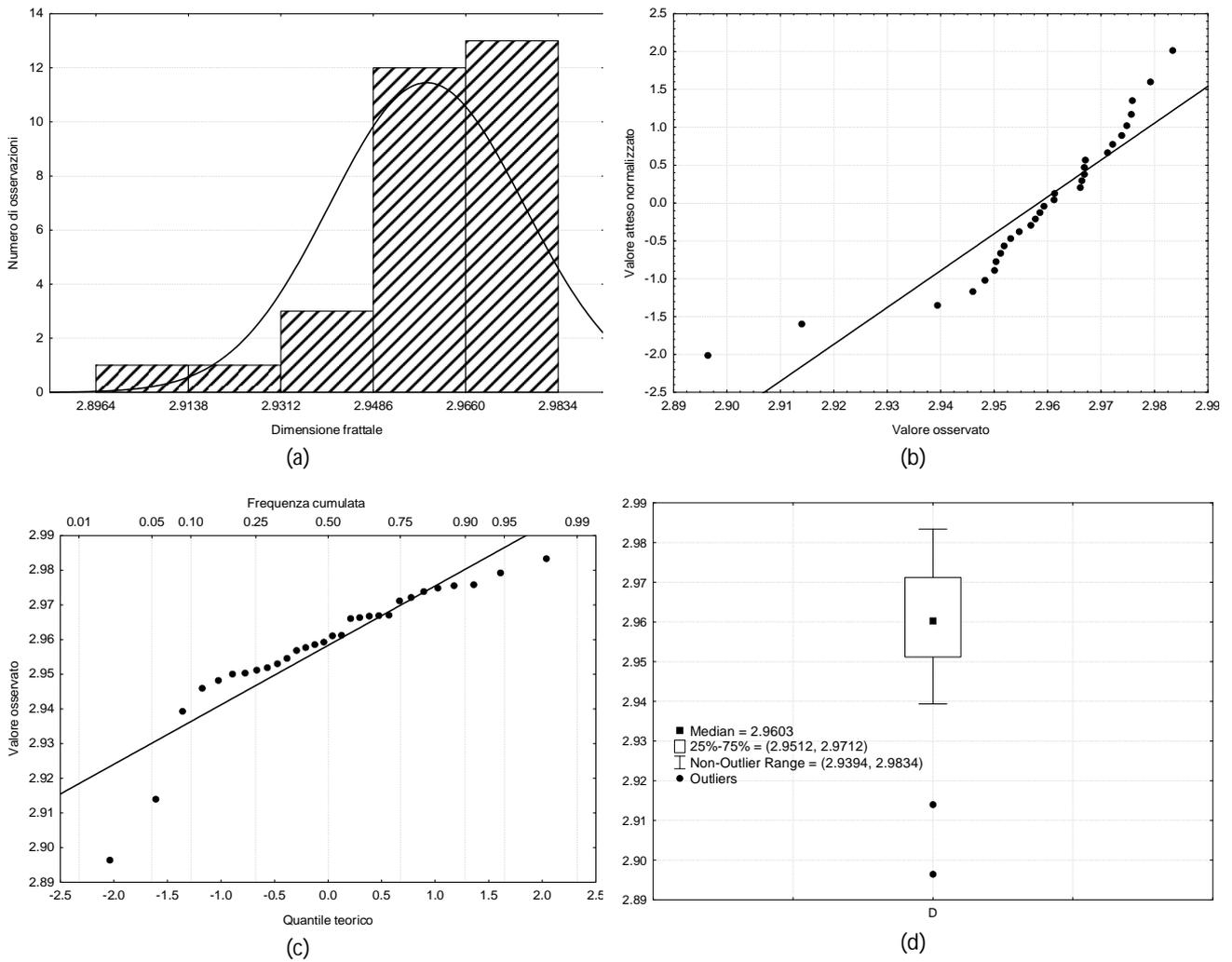


Figura 4.8.4 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per alti contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

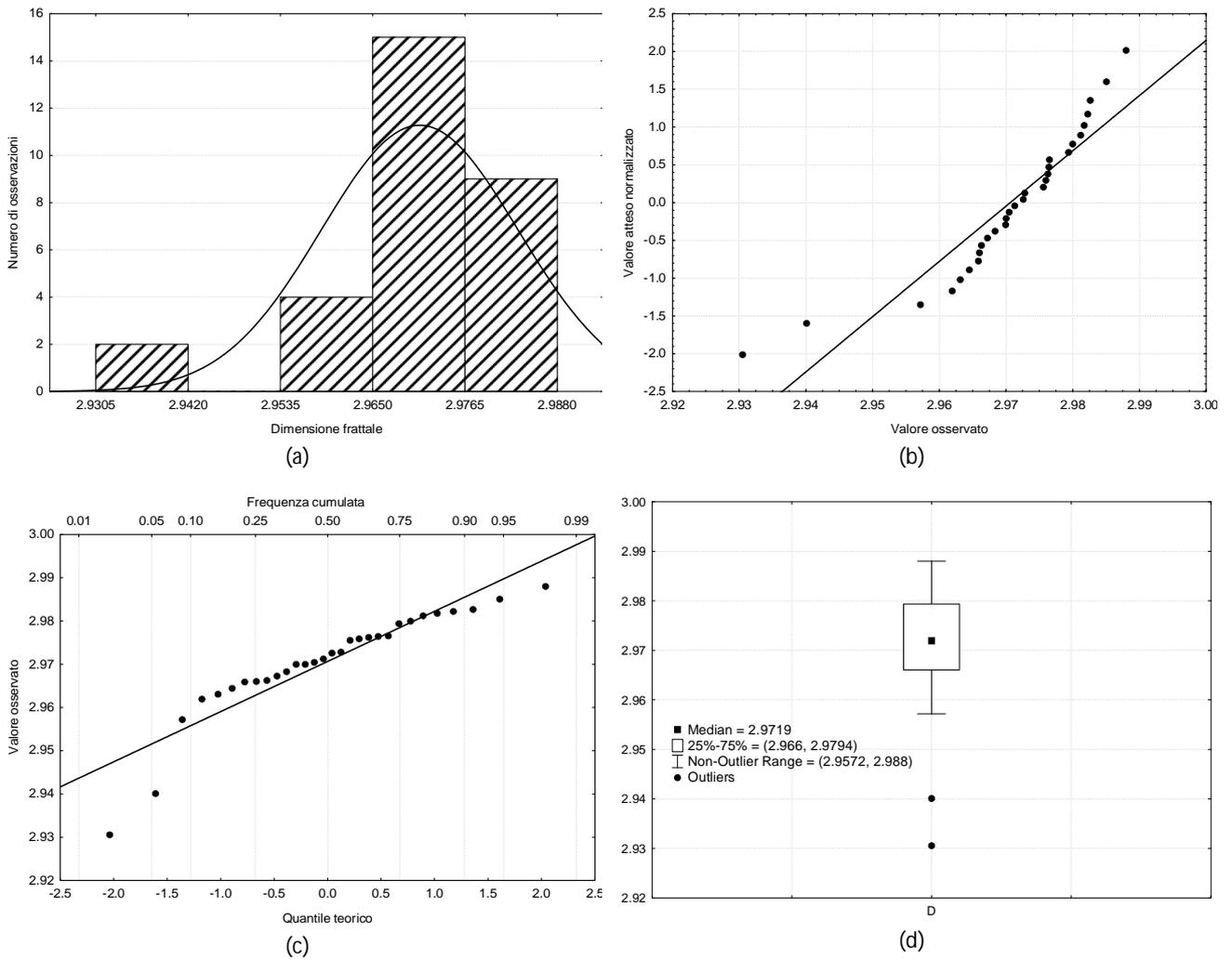


Figura 4.8.5 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per alti contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

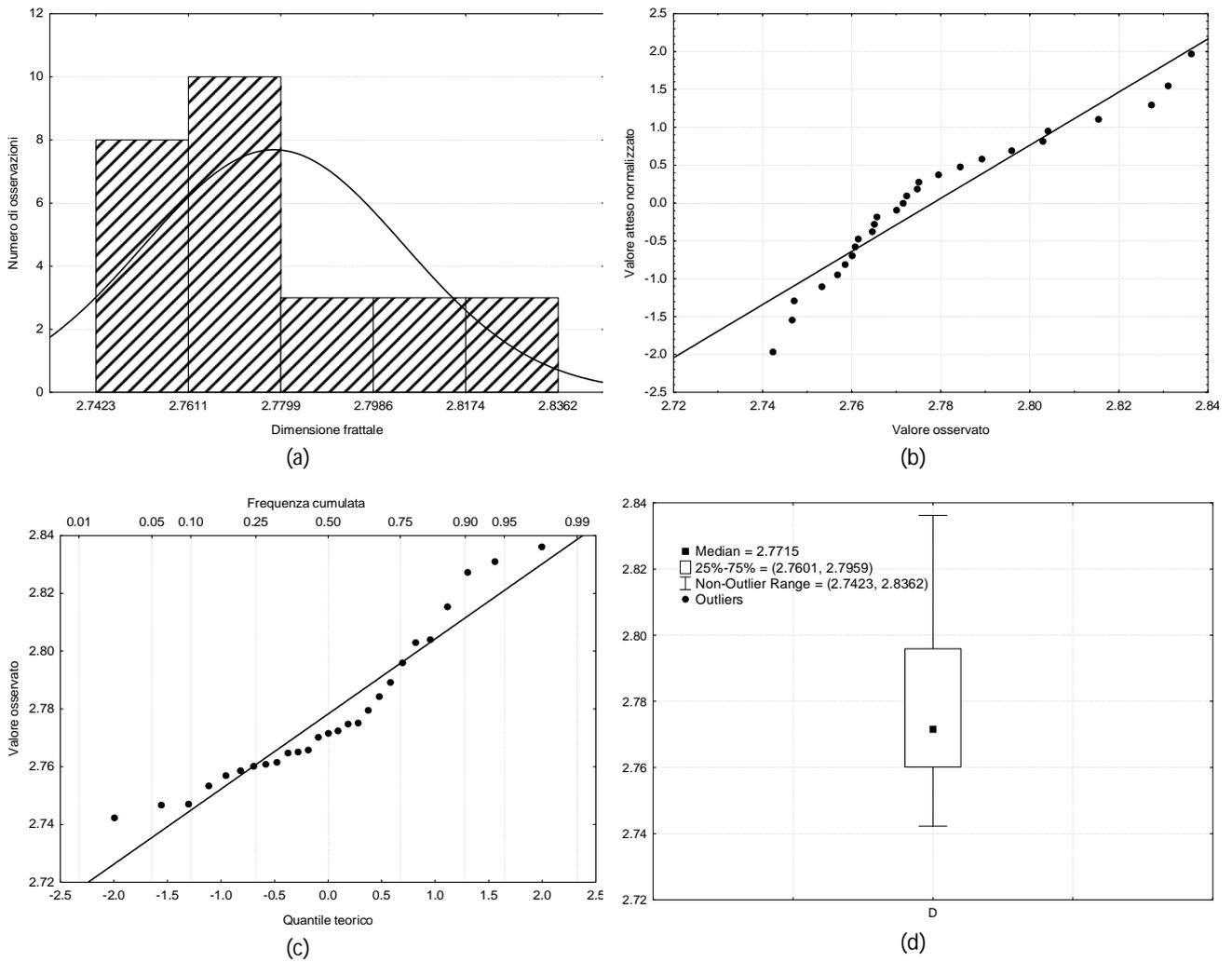


Figura 4.8.6 – Rappresentazione grafica della variabile dimensione frattale per il modello di Tyler & Wheatcraft [1990] e per bassi contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

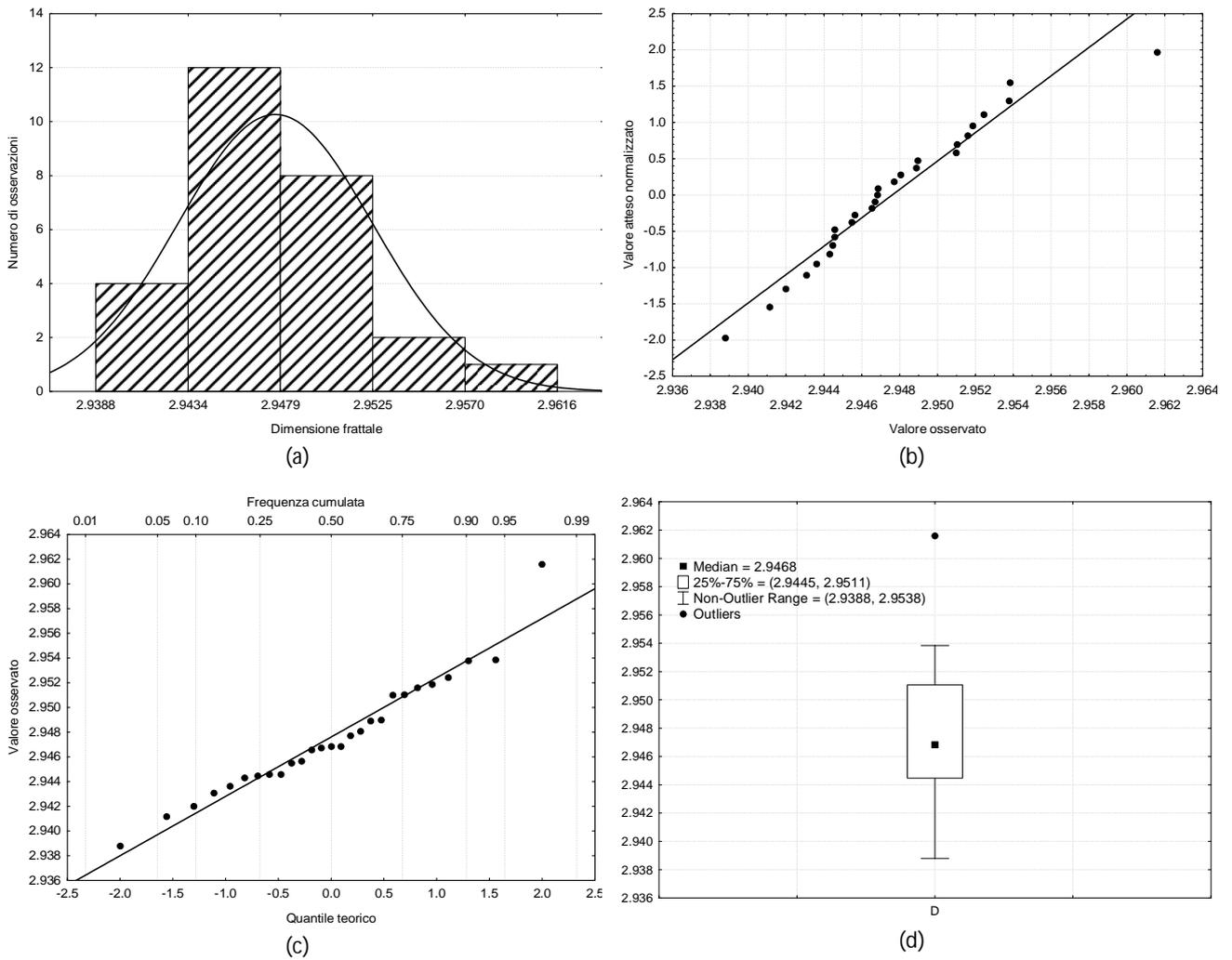


Figura 4.8.7 – Rappresentazione grafica della variabile dimensione frattale per il modello di Rieu & Sposito [1991] e per bassi contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

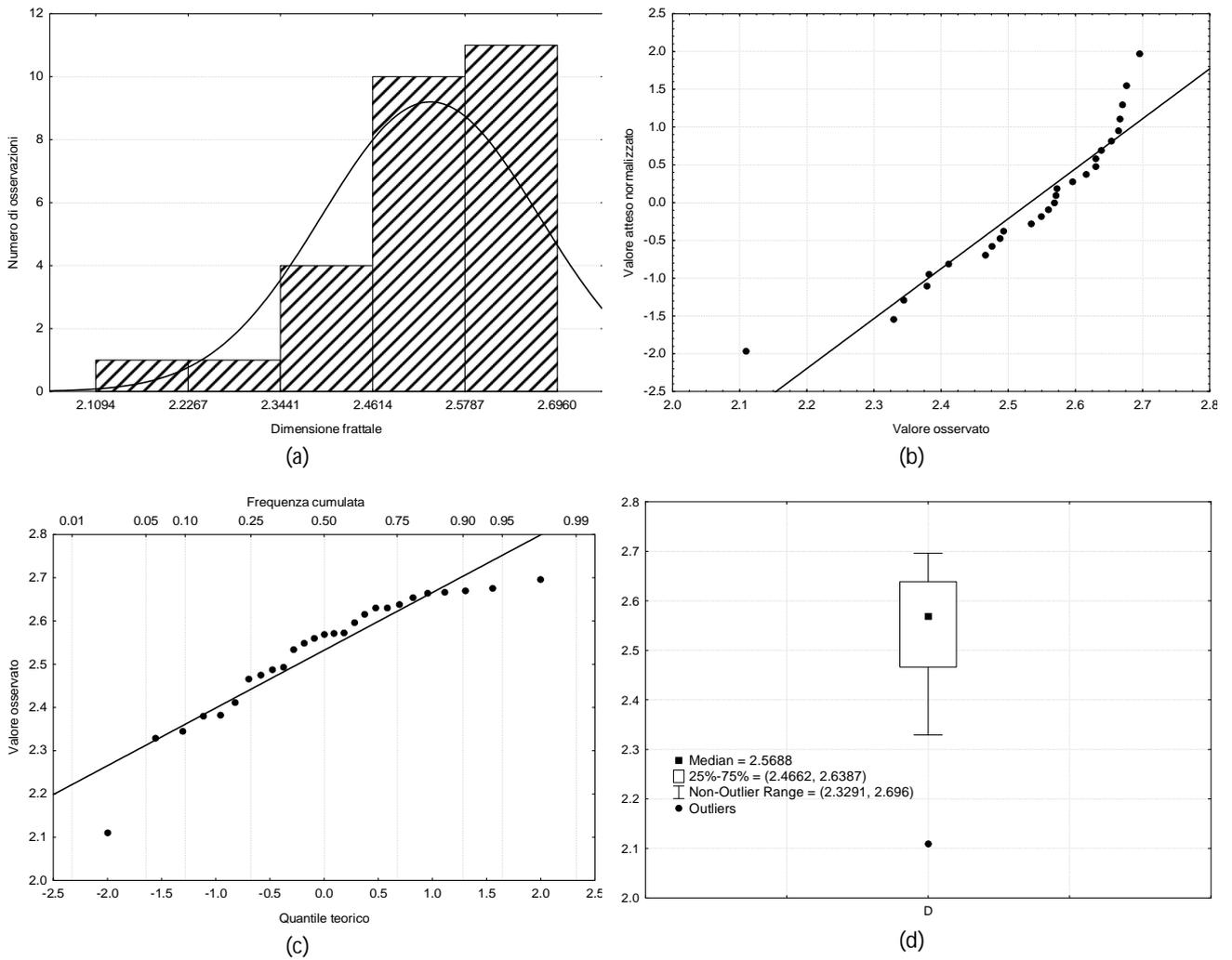


Figura 4.8.8 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con un parametro $A=0.30$ per bassi contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

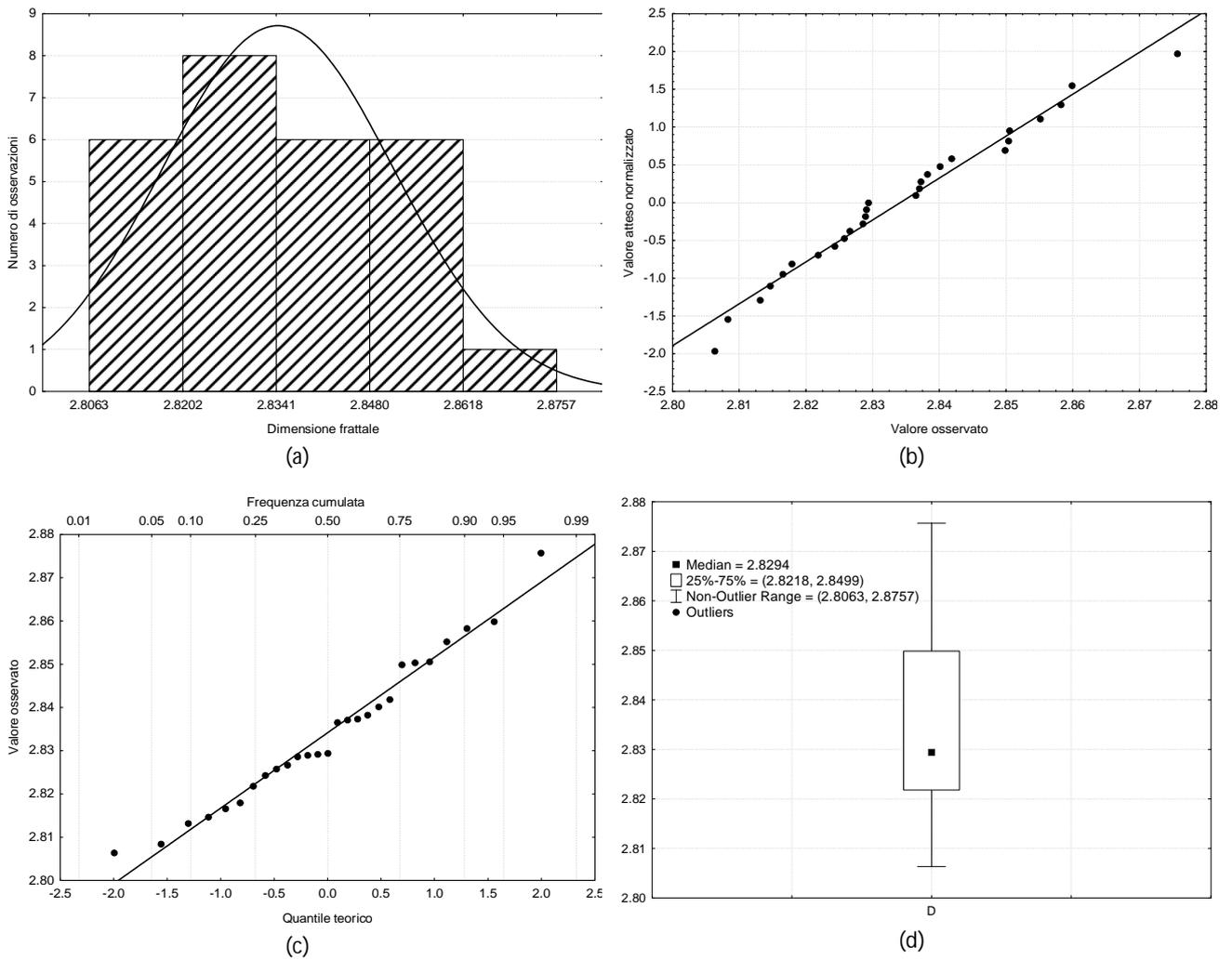


Figura 4.8.9 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con un parametro $A=0.45$ per bassi contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

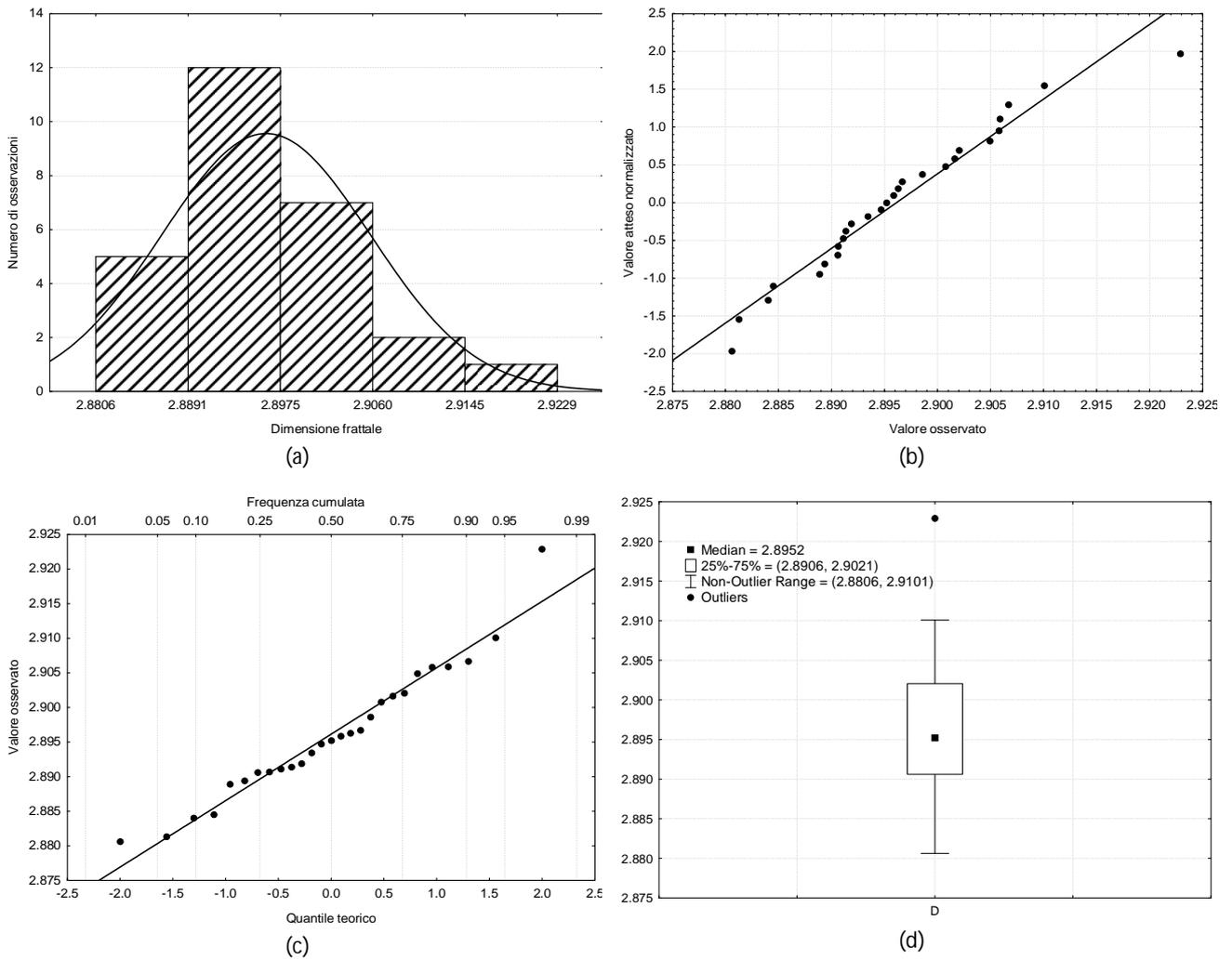


Figura 4.8.10 – Rappresentazione grafica della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] con un parametro $A=0.60$ per bassi contenuti di acqua. (a) Istogramma di frequenza; (b) NPP plot; (c) Q-Q plot; (d) Box-Plot (STATISTICA 8.0)

Modello frattale	Test statistico		
	Kolmogorov-Smirnov D	Anderson-Darling A ²	Chi-Square χ^2
TW(I)	0.160	1.120	1.553
RS(I)	0.140	0.956	1.645
PSF _{0.30} (I)	0.119	0.478	0.168
PSF _{0.45} (I)	0.157	1.142	0.822
PSF _{0.60} (I)	0.148	1.061	0.822
TW(II)	0.178	0.884	2.264
RS(II)	0.118	0.350	1.136
PSF _{0.30} (II)	0.141	0.793	2.187
PSF _{0.45} (II)	0.129	0.257	3.026
PSF _{0.60} (II)	0.108	0.292	0.512
V_{crit}	0.242	2.502	7.815

Tabella 4.6 – Risultati dei test di adattamento contr. v_{crit}

Modello frattale	Loc.	Outliers	Q	Q _{crit} ($\alpha=0.1$)
TW(I)	2	2.89426	0.4438	0.285
	26	2.90098		
RS(I)	2	2.96688	0.4310	0.285
	26	2.96295		
PSF _{0.30} (I)	2	2.84677	0.3887	0.218
PSF _{0.45} (I)	2	2.9140	0.4937	0.285
	26	2.89643		
PSF _{0.60} (I)	2	2.94012	0.4634	0.285
	26	2.93053		
RS(II)	2	2.9616	0.340	0.224
PSF _{0.30} (II)	26	2.10942	0.3746	0.224
PSF _{0.45} (II)	2	2.8757	0.2282	0.224
PSF _{0.60} (II)	2	2.92292	0.3039	0.224

Tabella 4.7 – Valori degli outliers per ciascun modello frattale e risultati del test di Dixon [1953]

4.4.2 Descrizione quantitativa della distribuzione dei punti e analisi del trend spaziale

Attraverso la tecnica *Nearest Neighbor Analysis (NNA)* è stato stabilito se la distribuzione spaziale dei punti, in cui sono stati effettuati i prelievi dei campioni per la stima delle curve di ritenzione idrica, risulti essere di tipo casuale, clusterizzata o dispersa. L'indice *Nearest Neighbor (NNI)* risulta essere sempre maggiore di uno (tabella 4.8) e, di conseguenza, ciascun processo può essere definito di tipo inibitorio; questo implica che i punti-evento tendono a disperdersi sul territorio e che, quindi, non vi è la necessità di applicare alcuna tecnica di declusterizzazione. Nella tabella 4.8 vengono inoltre riportati i valori della media della distanza minima tra gli N punti-evento, $d(NN)$, e il valore atteso della medesima grandezza quando si considerano gli eventi distribuiti in maniera puramente casuale, $d(att)$.

TW (I)		RS (I)		PSF _{0.30} (I)		PSF _{0.45} (I)		PSF _{0.60} (I)	
d(NN)	3.487	d(NN)	3.487	d(NN)	3.487	d(NN)	3.487	d(NN)	3.487
d(att)	2.283	d(att)	2.283	d(att)	2.283	d(att)	2.283	d(att)	2.283
<u>NNI</u>	<u>1.527</u>	<u>NNI</u>	<u>1.527</u>	<u>NNI</u>	<u>1.527</u>	<u>NNI</u>	<u>1.527</u>	<u>NNI</u>	<u>1.527</u>
TW (II)		RS (II)		PSF _{0.30} (II)		PSF _{0.45} (II)		PSF _{0.60} (II)	
d(NN)	3.633	d(NN)	3.664	d(NN)	3.664	d(NN)	3.664	d(NN)	3.664
d(att)	2.232	d(att)	2.264	d(att)	2.264	d(att)	2.264	d(att)	2.264
<u>NNI</u>	<u>1.627</u>	<u>NNI</u>	<u>1.618</u>	<u>NNI</u>	<u>1.618</u>	<u>NNI</u>	<u>1.618</u>	<u>NNI</u>	<u>1.618</u>

Tabella 4.8 – Risultati relativi all'analisi del vicinaggio

La verifica dell'ipotesi di stazionarietà e, quindi, l'esclusione della presenza di un trend spaziale, può essere effettuata con differenti metodologie. Nel presente lavoro un primo semplice controllo per la verifica di tale ipotesi è stato eseguito attraverso la rappresentazione delle mappe dei poligoni Voroni (fig. 4.9.1 - 4.9.5) in cui il valore assegnato alla singola cella corrisponde al valore misurato nel punto all'interno del poligono. Dalla distribuzione di tali poligoni si può notare, per ciascun modello frattale, la tendenza della variabile ad assumere valori uniformemente distribuiti nell'area di studio e, quindi, l'assenza nello spazio di un cambiamento sistematico e graduale della stessa.

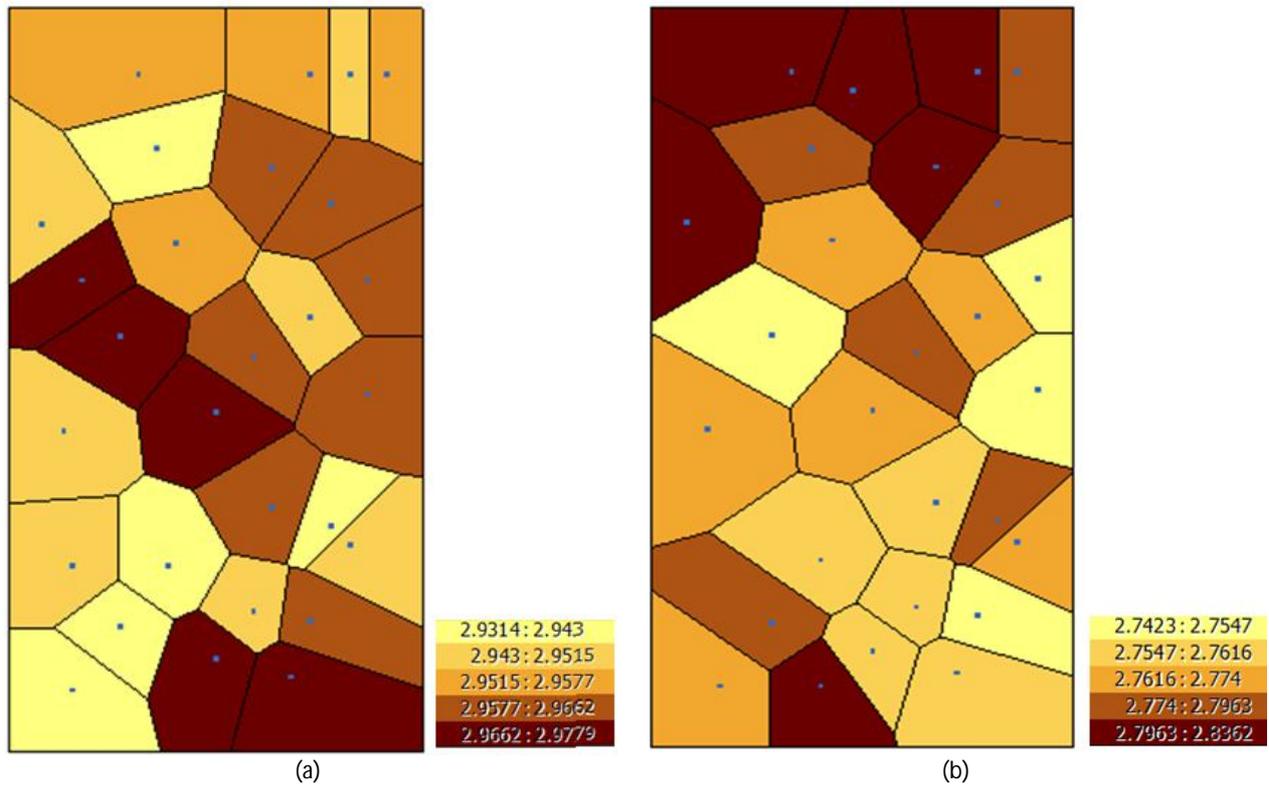


Figura 4.9.1 – Rappresentazione della mappa dei poligoni di Voronoi della variabile dimensione frattale per il modello di Tyler & Wheatcraft [1990]; (a) alti contenuti di acqua e (b) bassi contenuti di acqua (ARCGIS 9.3)

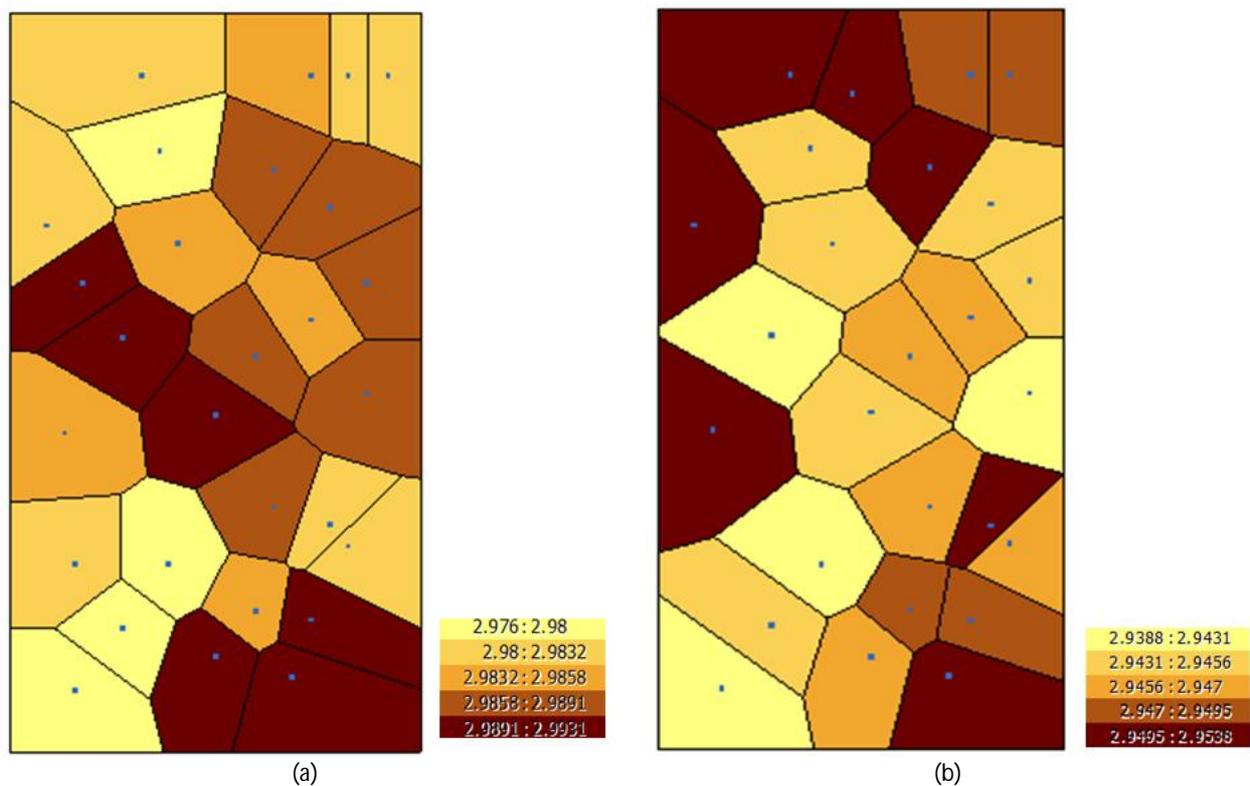


Figura 4.9.2 – Rappresentazione della mappa dei poligoni di Voronoi della variabile dimensione frattale per il modello di Rieu & Sposito [1991]; (a) alti contenuti di acqua e (b) bassi contenuti di acqua (ARCGIS 9.3)

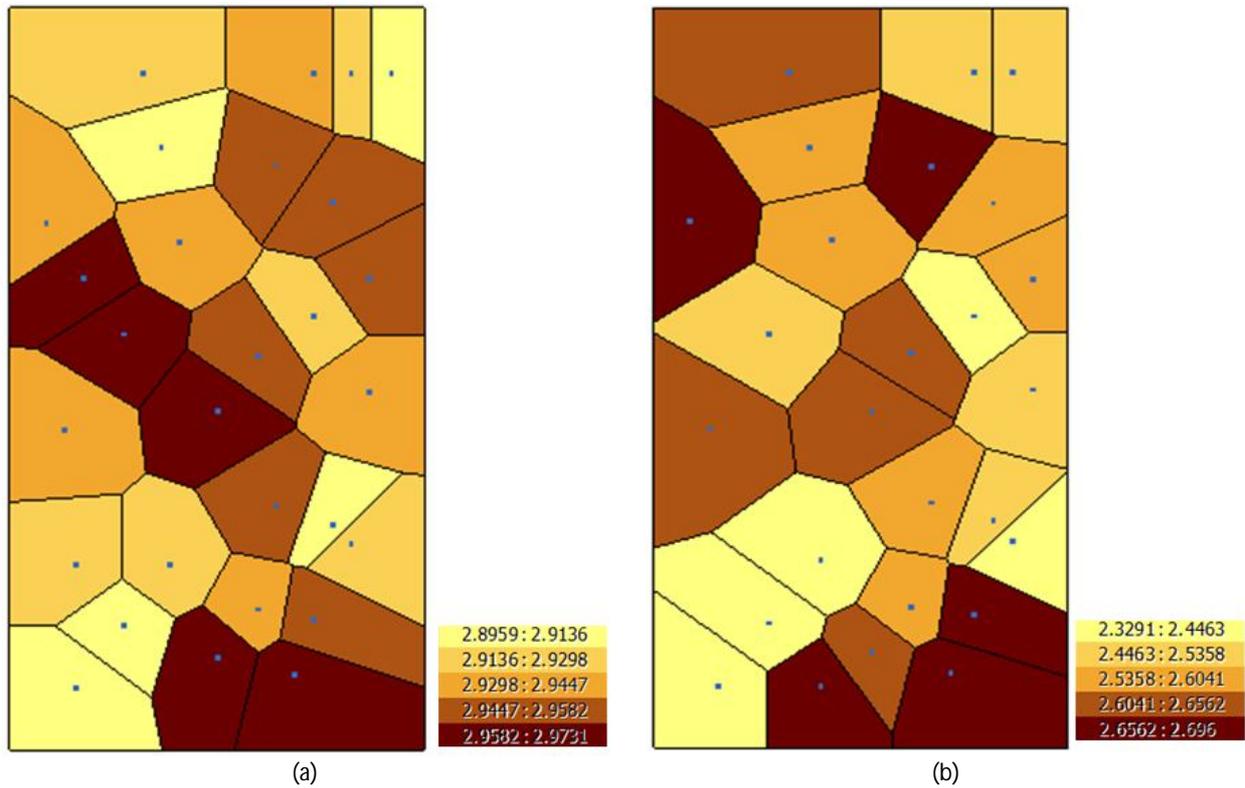


Figura 4.9.3 – Rappresentazione della mappa dei poligoni di Voronoi della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] e $A=0.30$; (a) alti contenuti di acqua e (b) bassi contenuti di acqua (ARCGIS 9.3)

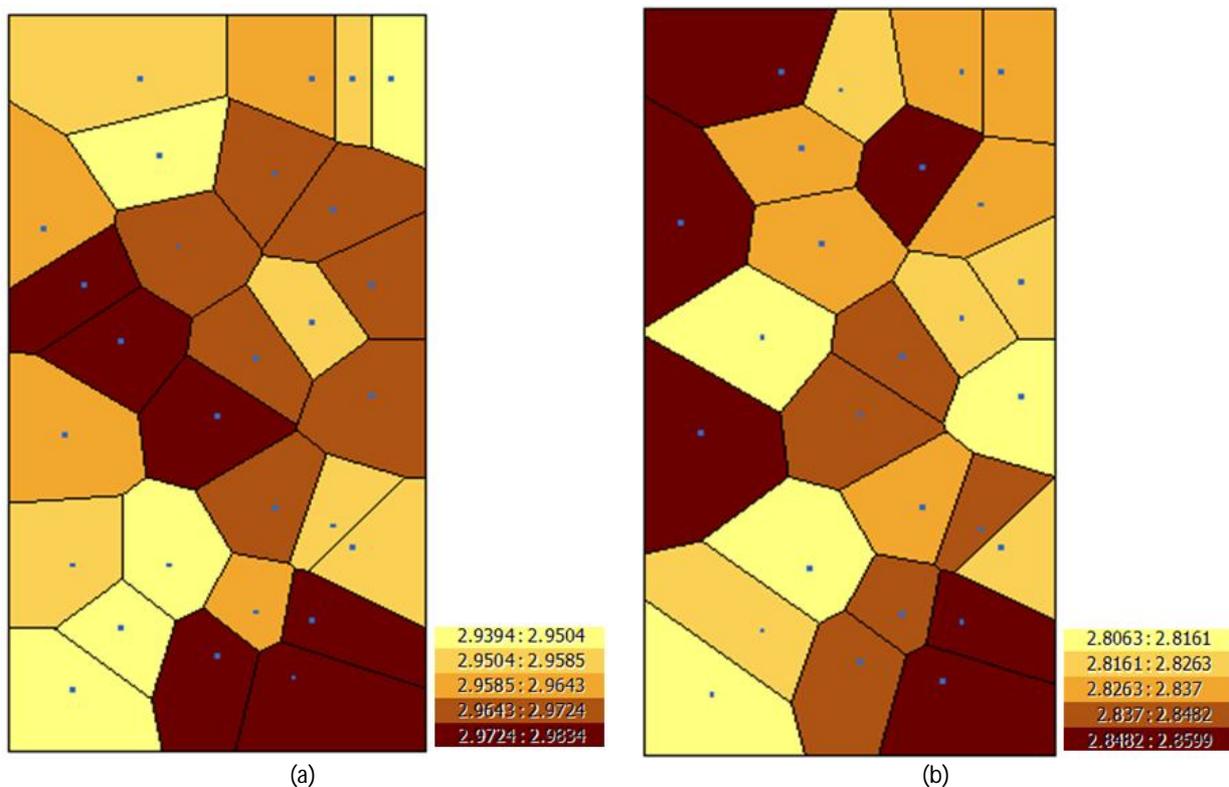


Figura 4.9.4 – Rappresentazione della mappa dei poligoni di Voronoi della variabile dimensione frattale per il modello Pore-Solid Fractal [2000] e $A=0.45$; (a) alti contenuti di acqua e (b) bassi contenuti di acqua (ARCGIS 9.3)

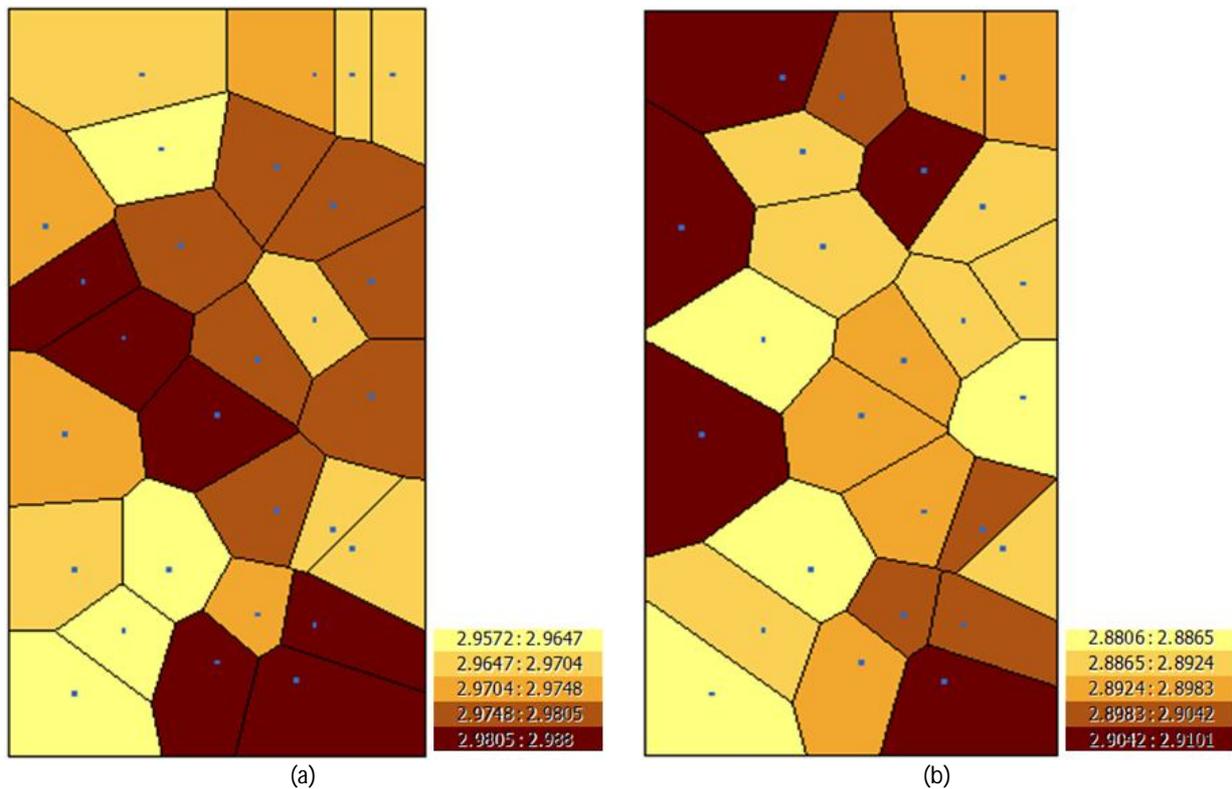


Figura 4.9.5 – Rappresentazione della mappa dei poligoni di Voronoi della variabile dimensione frattale per il modello *Pore-Solid Fractal* [2000] e $A=0.60$; (a) alti contenuti di acqua e (b) bassi contenuti di acqua (ARCGIS 9.3)

Un altro metodo per la verifica dell'ipotesi di stazionarietà consiste nell'esaminare l'andamento della media della variabile dimensione frattale nel piano, sia lungo la direzione x che lungo la direzione y ; nello specifico sono stati utilizzati dei diagrammi *box-plot* nei quali il valore centrale in ciascuna scatola non corrisponde alla mediana bensì alla media (fig. 4.10.1 – 4.10.4); tale metodologia conferma l'ipotesi precedentemente assunta: poiché il valore medio, lungo le direzioni x e y , non presenta un abbassamento e/o innalzamento graduale ed evidente, si può dedurre che i set di campioni esaminati non sono caratterizzati da un trend spaziale e, di conseguenza, può essere confermata la verifica dell'ipotesi di stazionarietà della variabile dimensione frattale nell'area oggetto di studio.

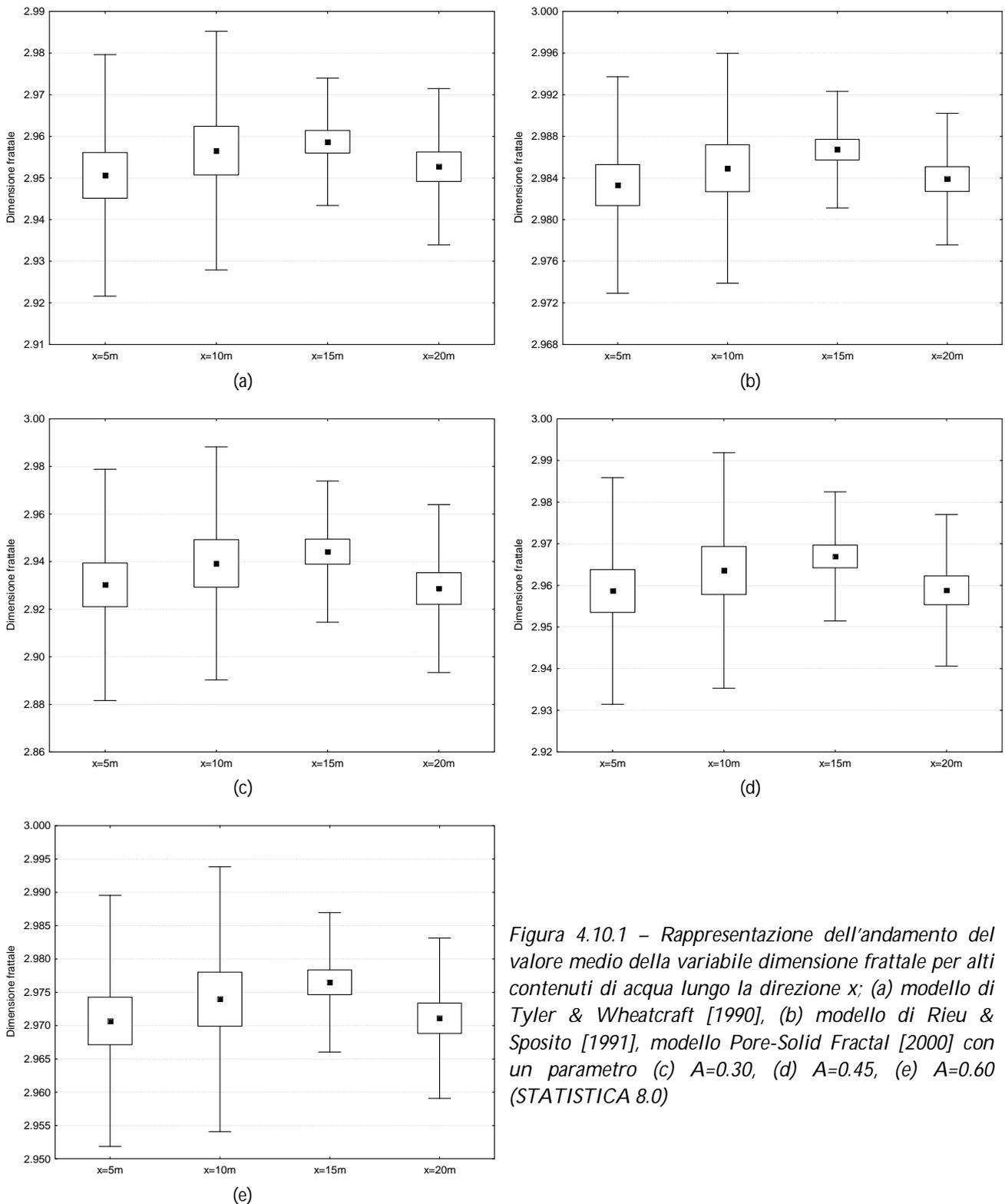


Figura 4.10.1 – Rappresentazione dell'andamento del valore medio della variabile dimensione frattale per alti contenuti di acqua lungo la direzione x ; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (STATISTICA 8.0)

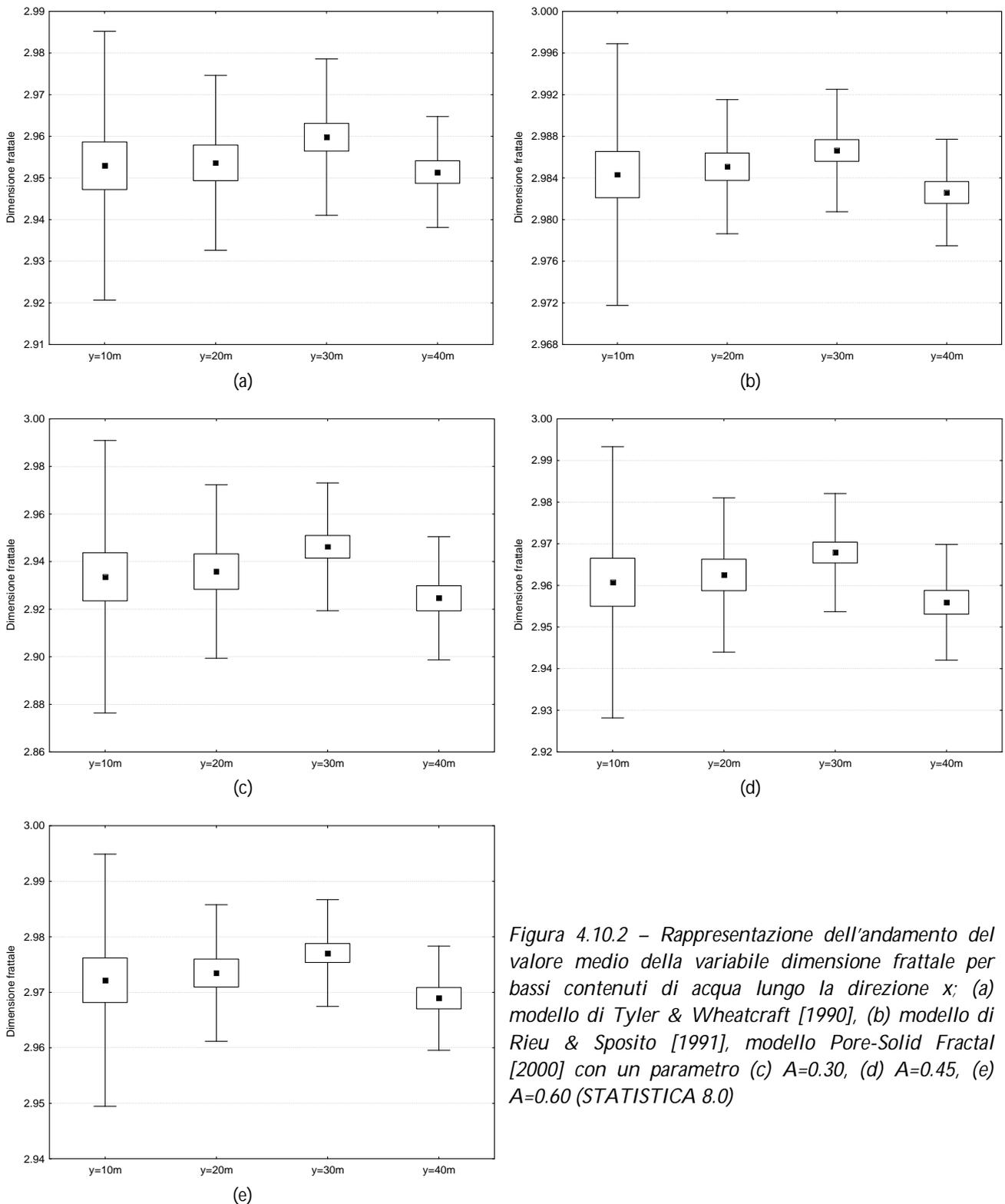


Figura 4.10.2 – Rappresentazione dell'andamento del valore medio della variabile dimensione frattale per bassi contenuti di acqua lungo la direzione x; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (STATISTICA 8.0)

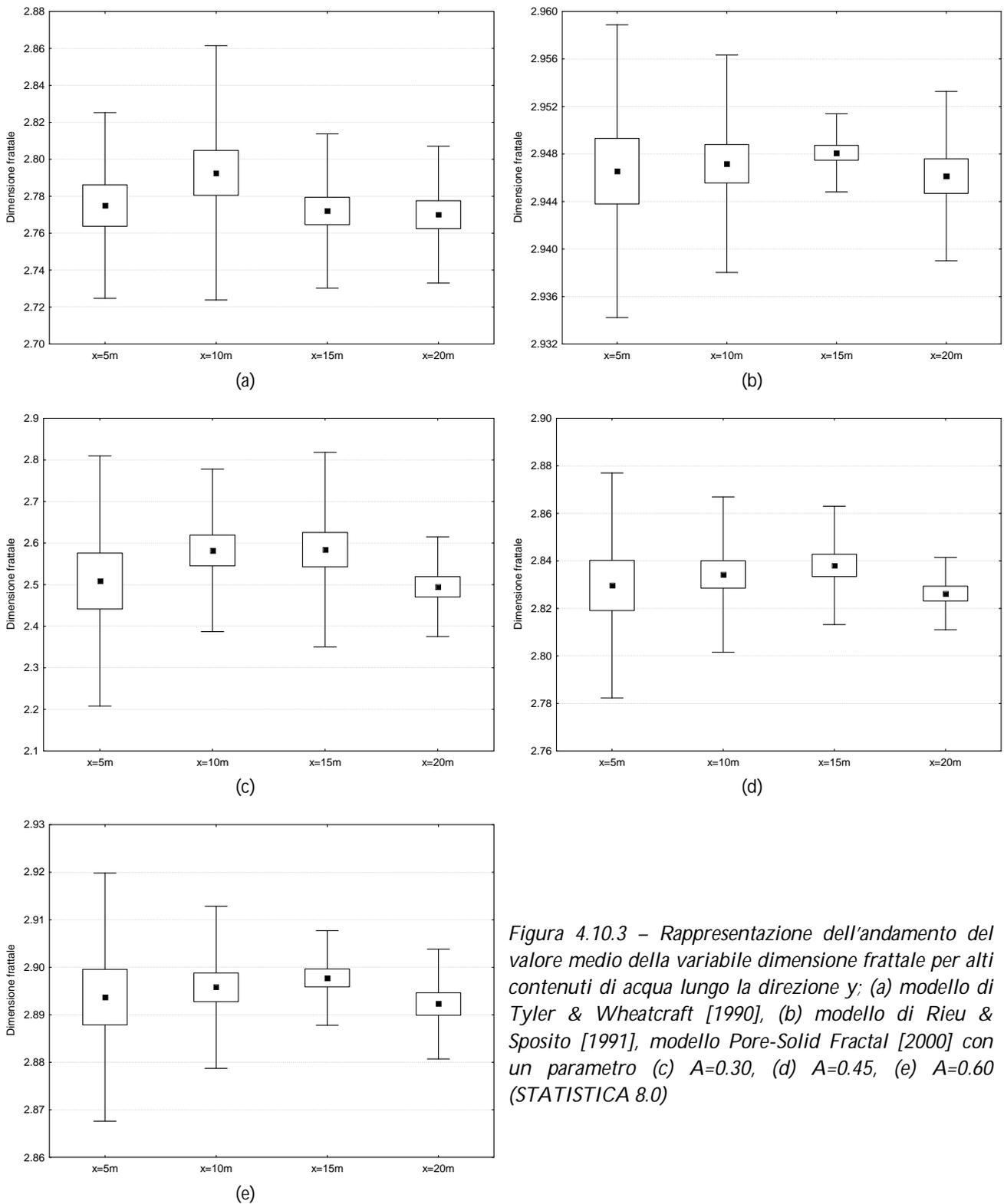


Figura 4.10.3 – Rappresentazione dell'andamento del valore medio della variabile dimensione frattale per alti contenuti di acqua lungo la direzione y; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (STATISTICA 8.0)

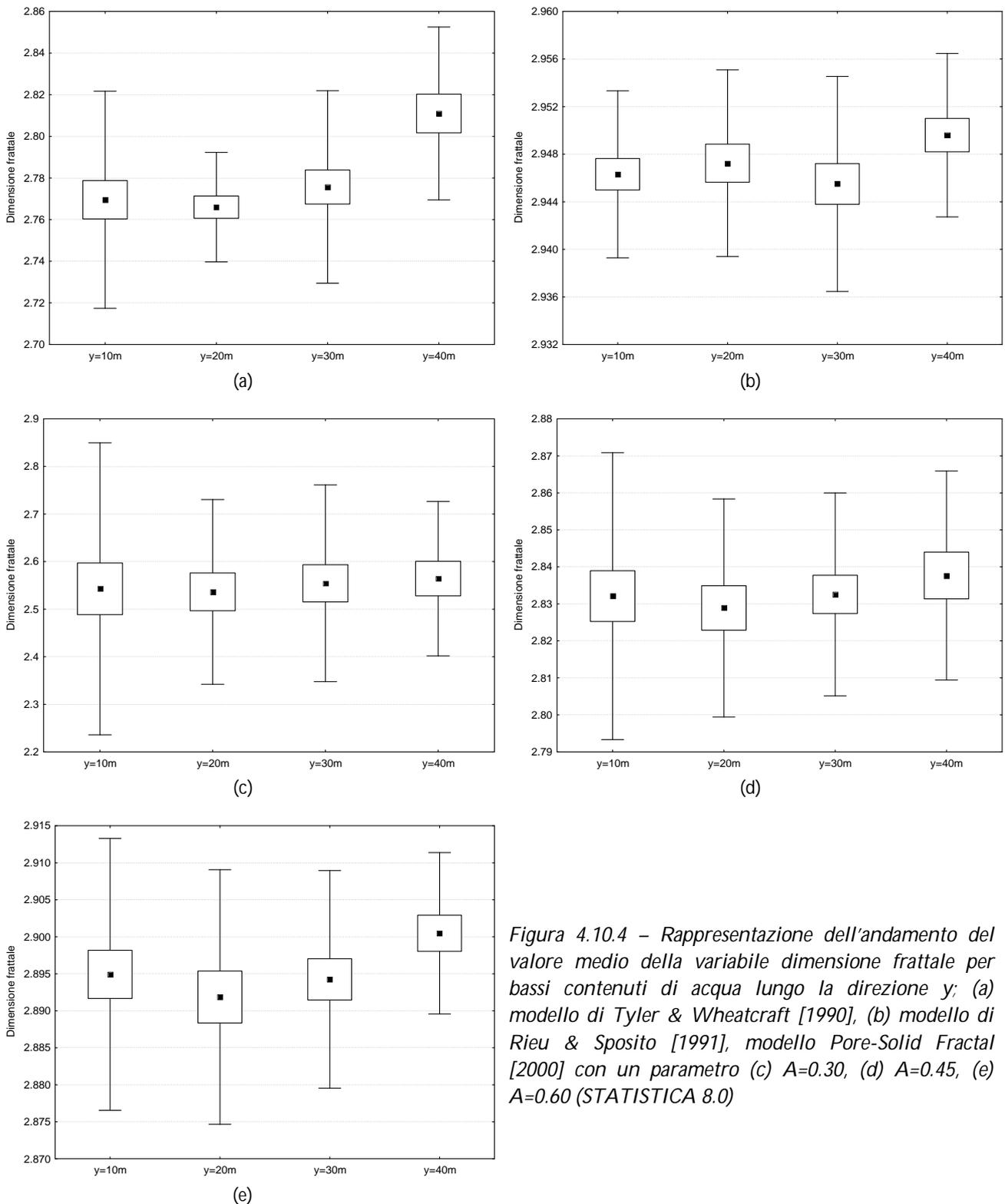


Figura 4.10.4 – Rappresentazione dell'andamento del valore medio della variabile dimensione frattale per bassi contenuti di acqua lungo la direzione y; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (STATISTICA 8.0)

4.4.3 Analisi della correlazione spaziale

4.4.3.1 Studio delle anisotropie e individuazione del variogramma sperimentale

La mappa del variogramma, che rappresenta l'andamento della varianza nello spazio, fa sì che sia subito comprensibile come vari un eventuale fenomeno anisotropo per poter decidere, eventualmente, su quali direzioni principali concentrare l'attenzione del variogramma. Ogni cella della mappa è rappresentata infatti con una simbologia di colori variabile a seconda del valore del variogramma secondo una scala cromatica propria (fig. 4.11.1 e 4.11.2). Da questa prima analisi circa le direzioni preferenziali di anisotropia si può affermare che la variabile dimensione frattale, per ciascun modello, per bassi e alti contenuti di acqua, non rappresenta un fenomeno che si sviluppa con un preciso andamento spaziale, bensì esso risulta uniformemente e ugualmente distribuito in qualsiasi direzione.

Tale ipotesi può essere confermata mediante l'utilizzo dei variogrammi direzionali (fig. 4.12.1 – 4.12.10); questi ultimi sono stati ottenuti precisamente individuando le direzioni Est, Nord-Est, Nord, e Nord-Ovest con un angolo di tolleranza pari a 22.5°. Attraverso l'analisi e lo studio dei variogrammi direzionali è possibile vedere se essi presentano o meno un comportamento simile nelle differenti direzioni e confrontarli anche con ciascun variogramma omnidirezionale (fig. 4.13). E' da precisare che, per ogni tipologia di variogramma, i numeri in corrispondenza dei punti con coordinate $(h, \gamma(h))$, rappresentano le coppie di punti campionati presenti per ciascun *lag*; differentemente, la linea tratteggiata indica il valore della varianza statistica dei campioni. L'andamento dei variogrammi direzionali per i valori della dimensione frattale riferiti ad alti e bassi contenuti di acqua, presenta un andamento molto simile in direzione Nord, Nord-Est e Nord-Ovest.

Per ciascun modello tutti i variogrammi direzionali sono caratterizzati da valori del *sill* che si mantengono più o meno intorno al valore della varianza, mentre il *range* risulta essere pari a circa 10-12m. In tal caso, considerando che i variogrammi nelle direzioni Nord, Nord-Est e Nord-Ovest hanno tra l'altro un andamento forzatamente simile a quello omnidirezionale, per quel che riguarda l'anisotropia e per alti contenuti di acqua, si può, quindi, mantenere la tesi assunta tramite l'analisi della mappa del variogramma: la variabile dimensione frattale risulta uniformemente distribuita qualsiasi sia la direzione all'interno dell'area oggetto di studio.

I variogrammi sperimentali presi in esame per effettuare la successiva interpolazione col *Kriging*, saranno, quindi, quelli omnidirezionali, per ciascun modello frattale e per alti e bassi contenuti di acqua.

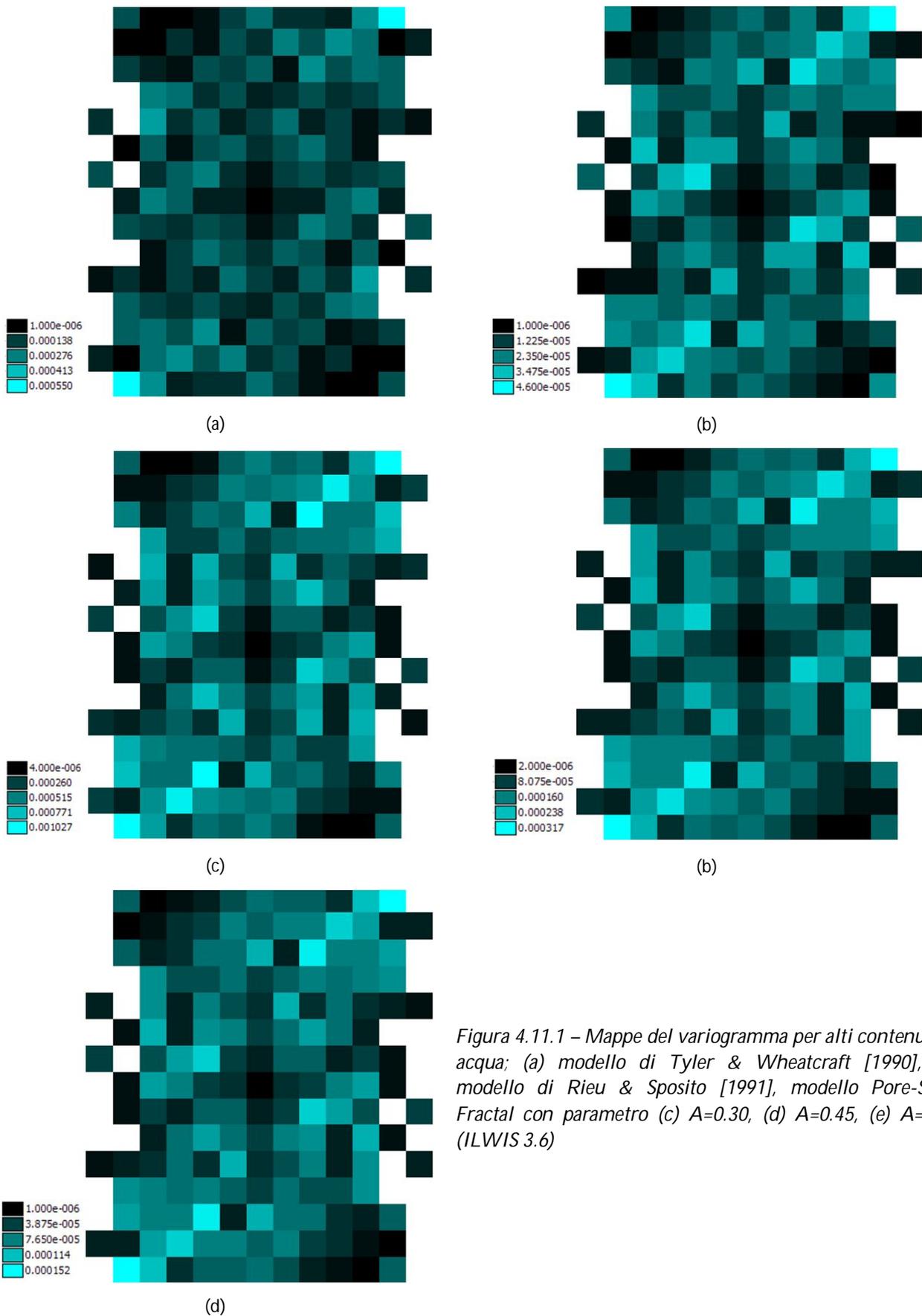


Figura 4.11.1 – Mappe del variogramma per alti contenuti di acqua; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal con parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (ILWIS 3.6)

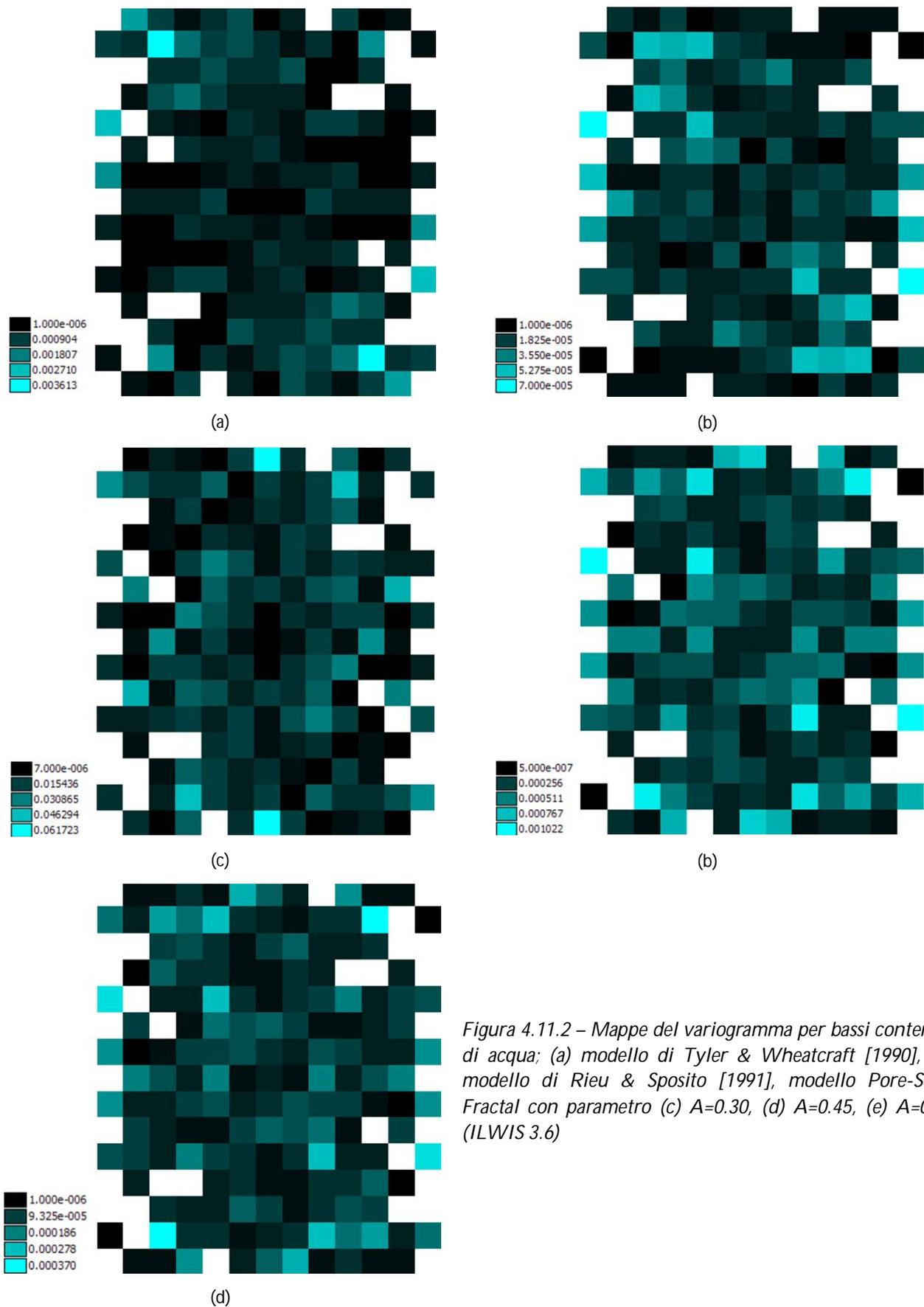


Figura 4.11.2 – Mappe del variogramma per bassi contenuti di acqua; (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal con parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ (ILWIS 3.6)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

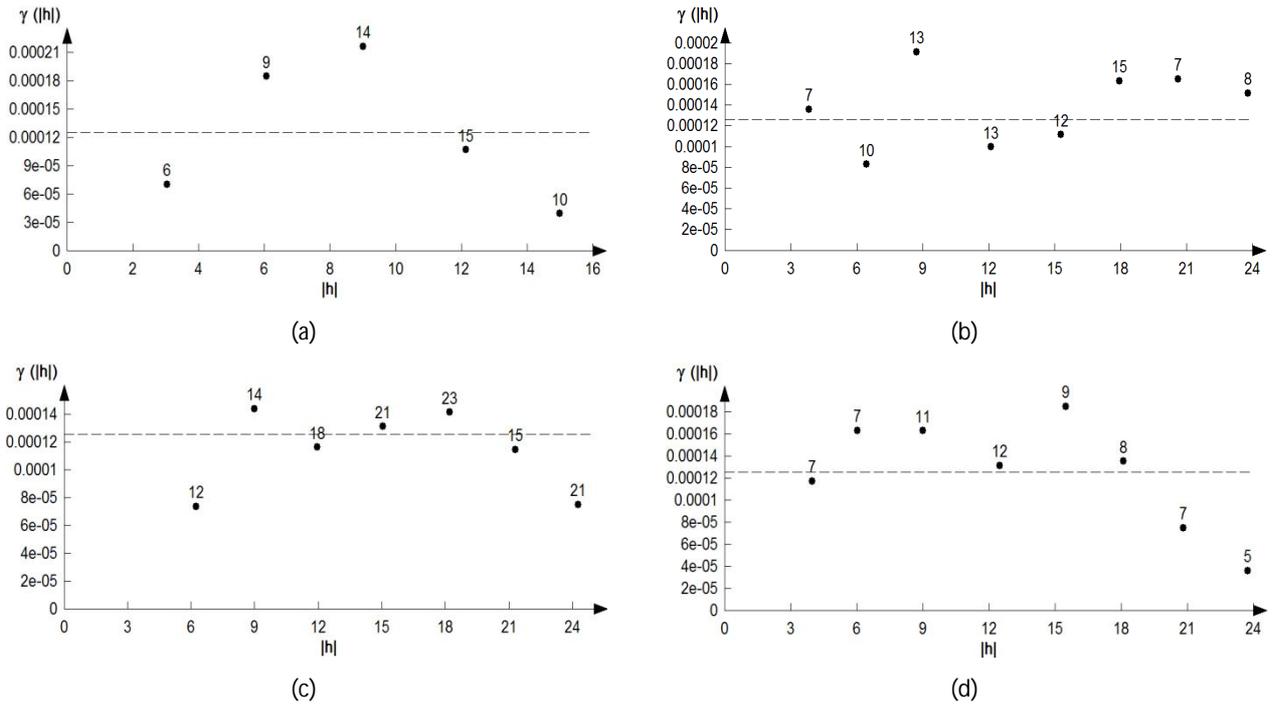


Figura 4.12.1 – Rappresentazione dei variogrammi direzionali per il modello di Tyler & Wheatcraft [1990] e per alti contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

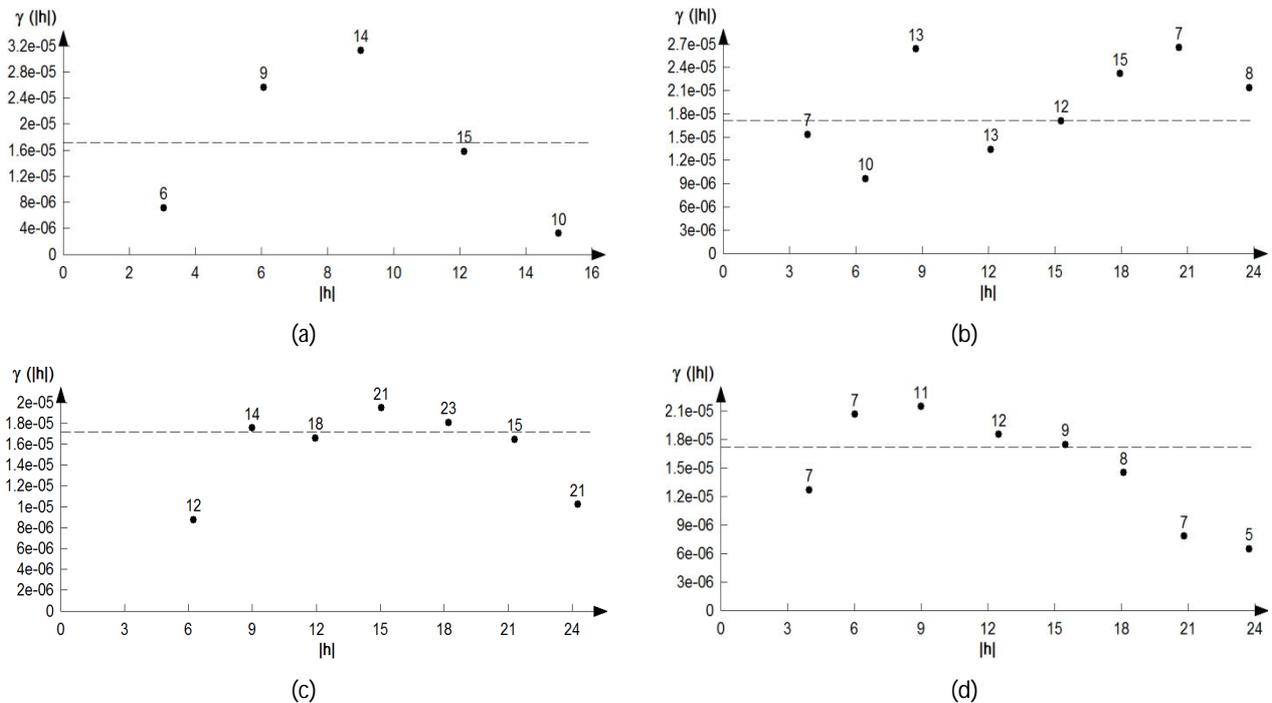


Figura 4.12.2 – Rappresentazione dei variogrammi direzionali per il modello di Rieu & Sposito [1991] e per alti contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

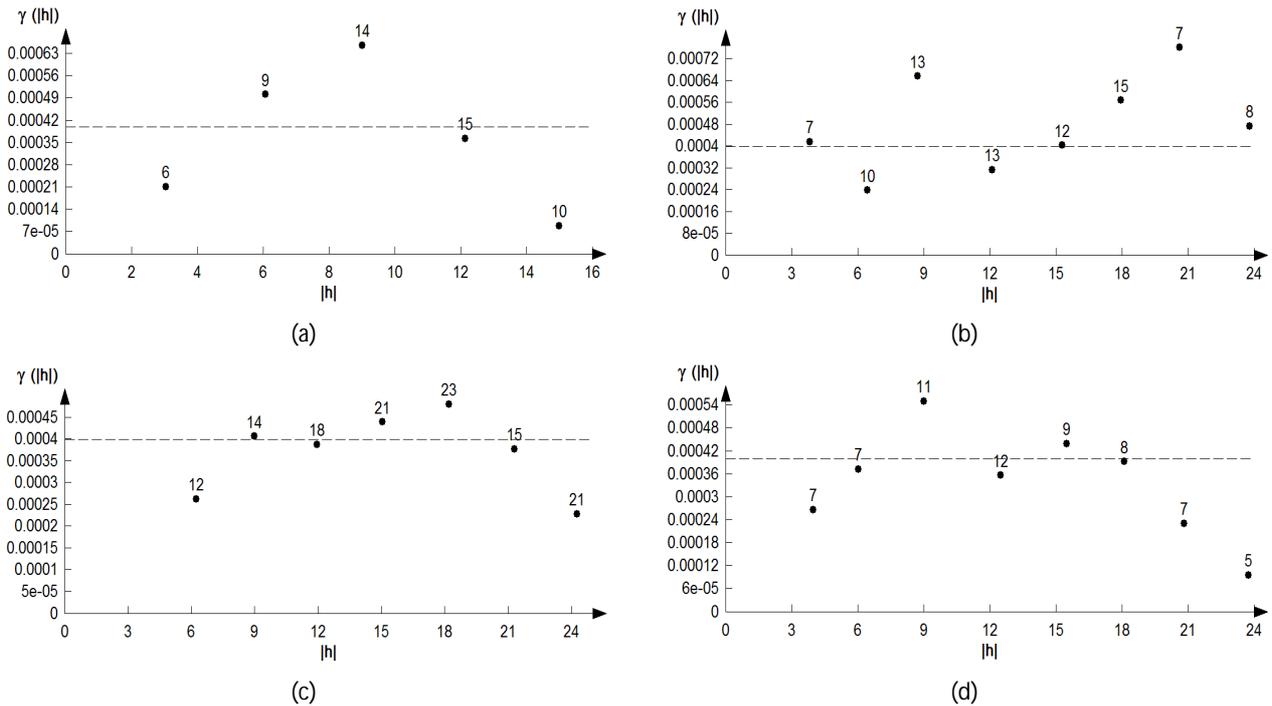


Figura 4.12.3 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per alti contenuti di acqua: (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

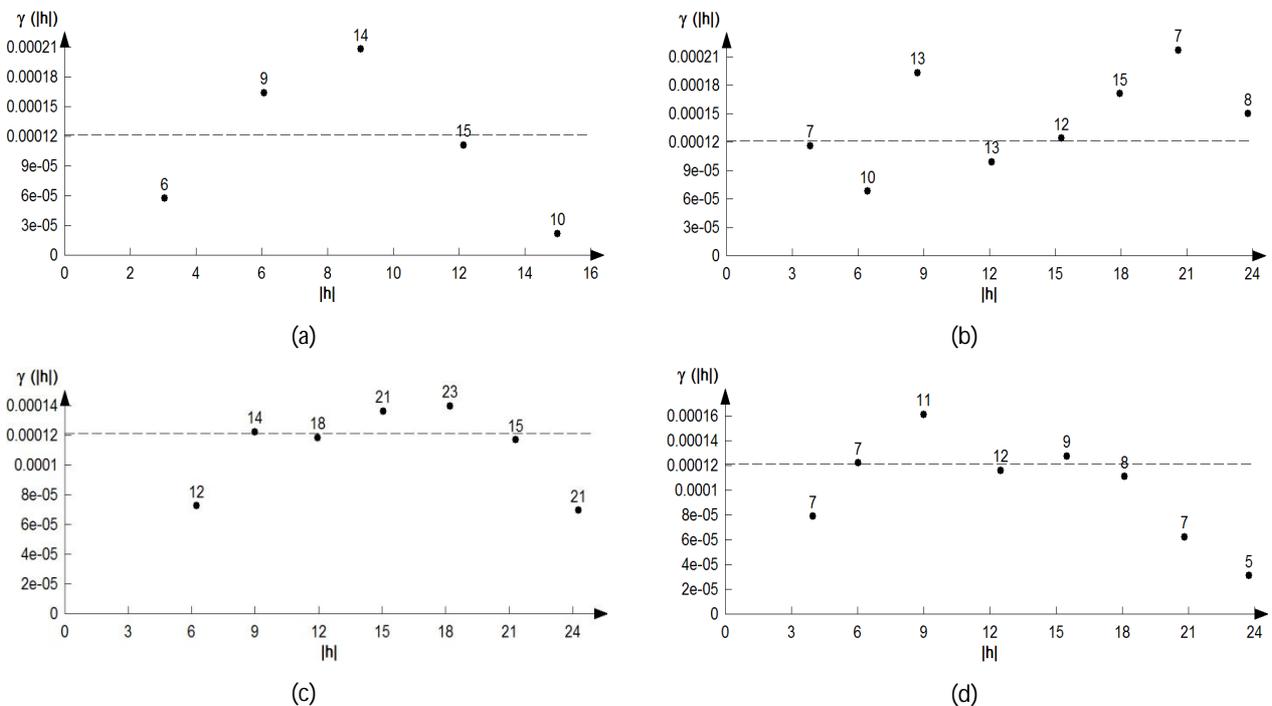


Figura 4.12.4 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per alti contenuti di acqua: (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

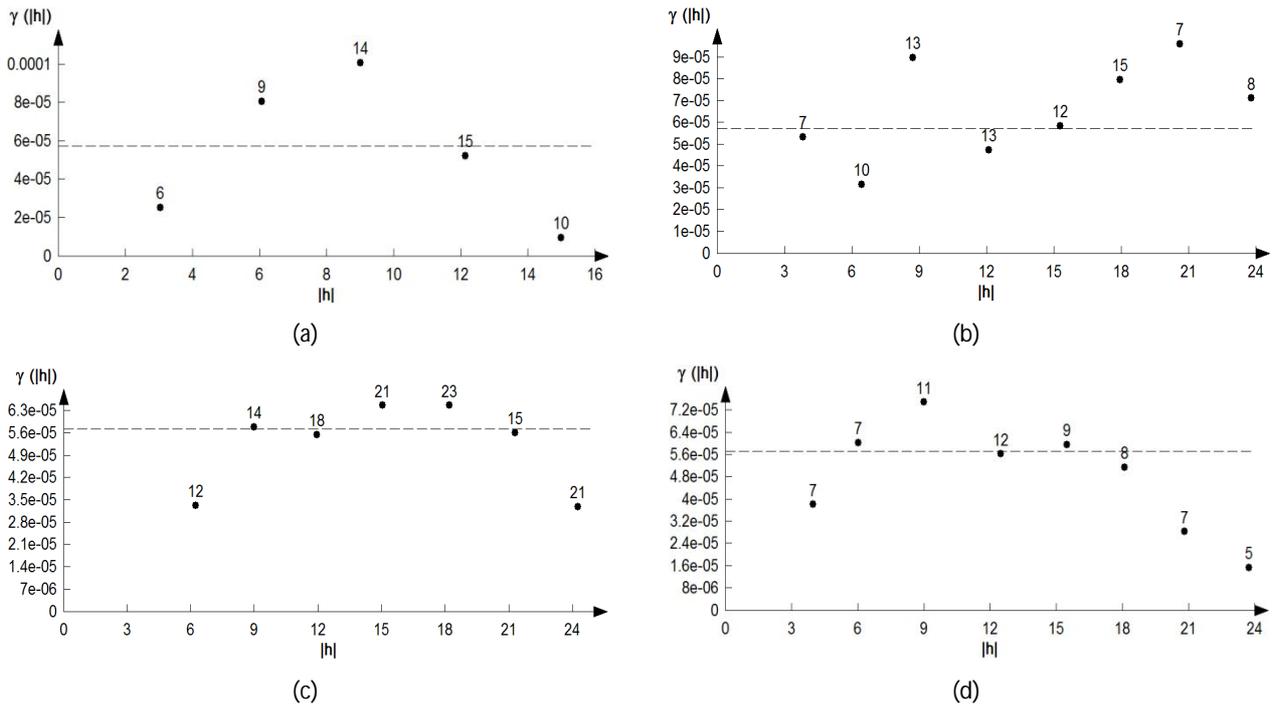


Figura 4.12.5 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per alti contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

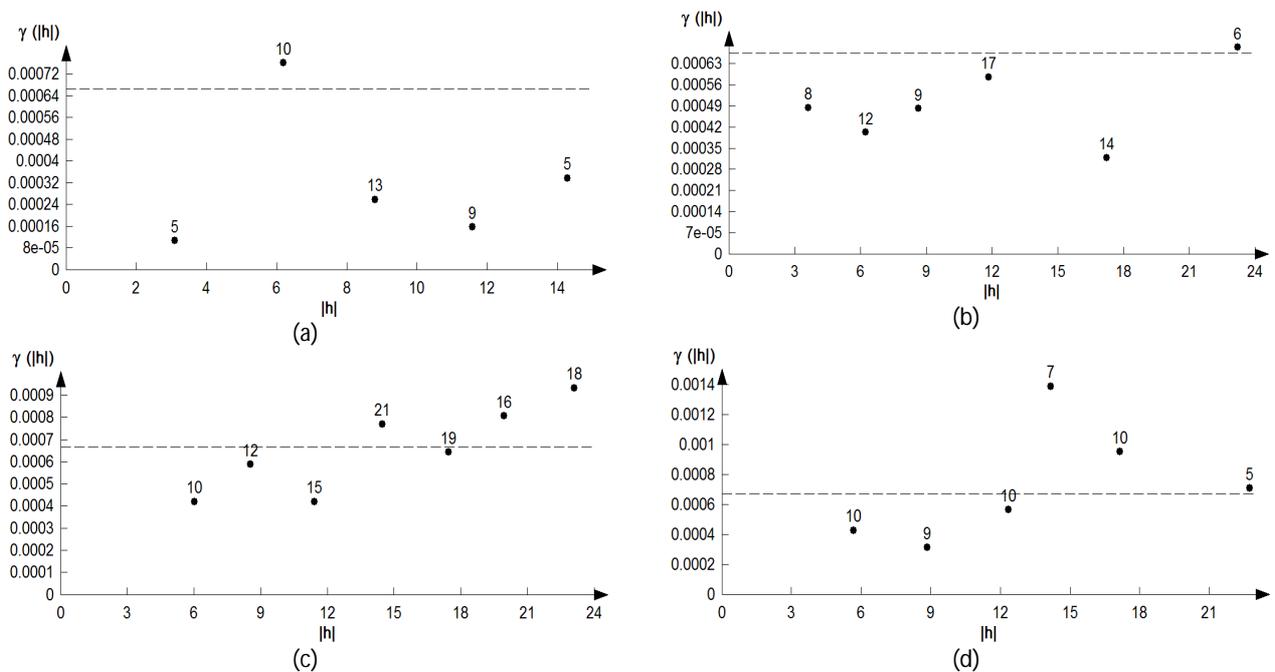


Figura 4.12.6 – Rappresentazione dei variogrammi direzionali per il modello di Tyler & Wheatcraft [1990] e per bassi contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

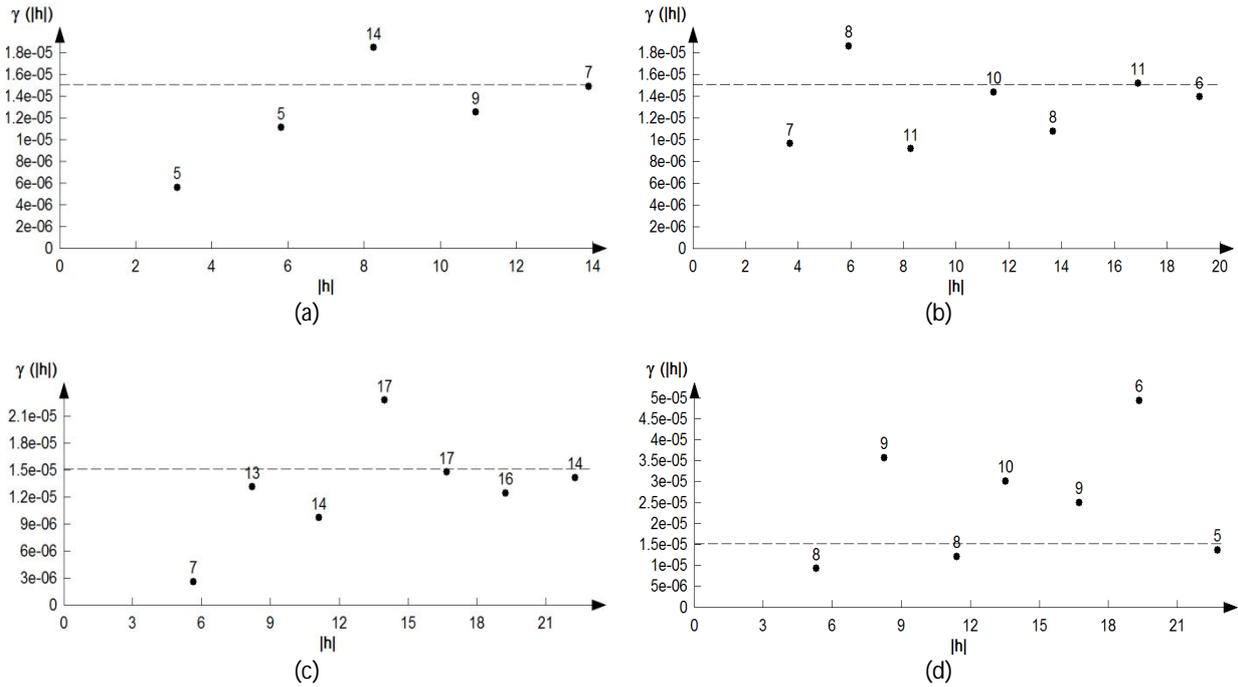


Figura 4.12.7 – Rappresentazione dei variogrammi direzionali per il modello di Rieu & Sposito [1991] e per bassi contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

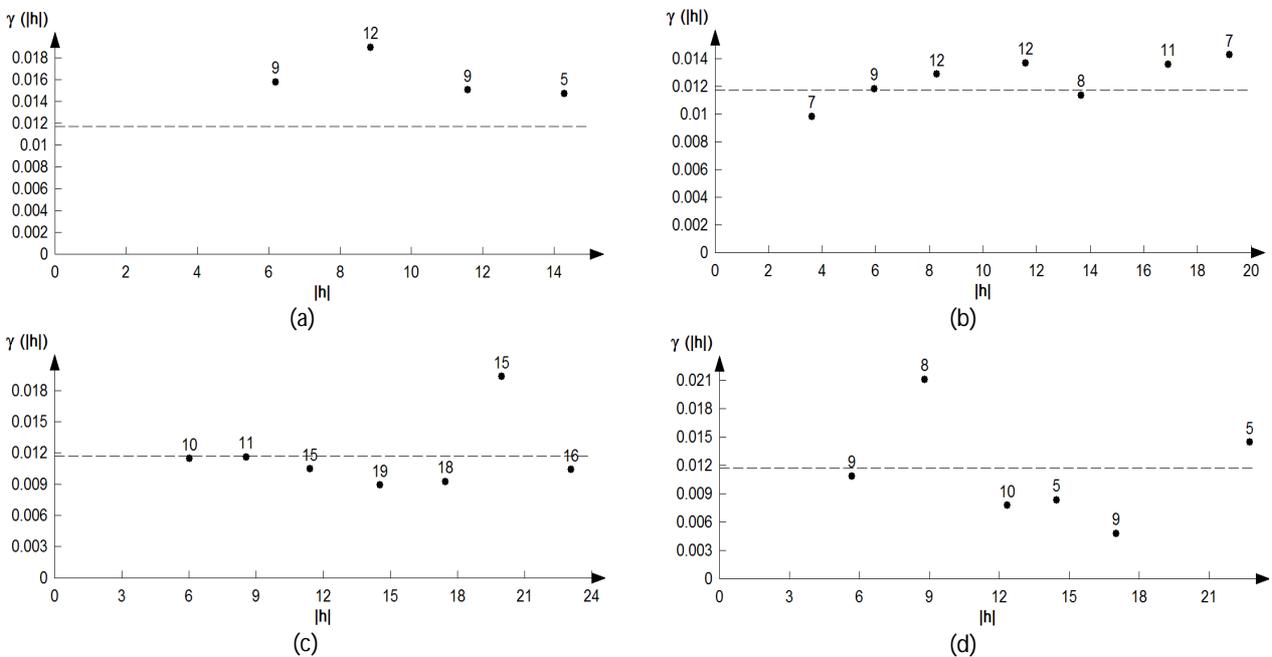


Figura 4.12.8 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per bassi contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

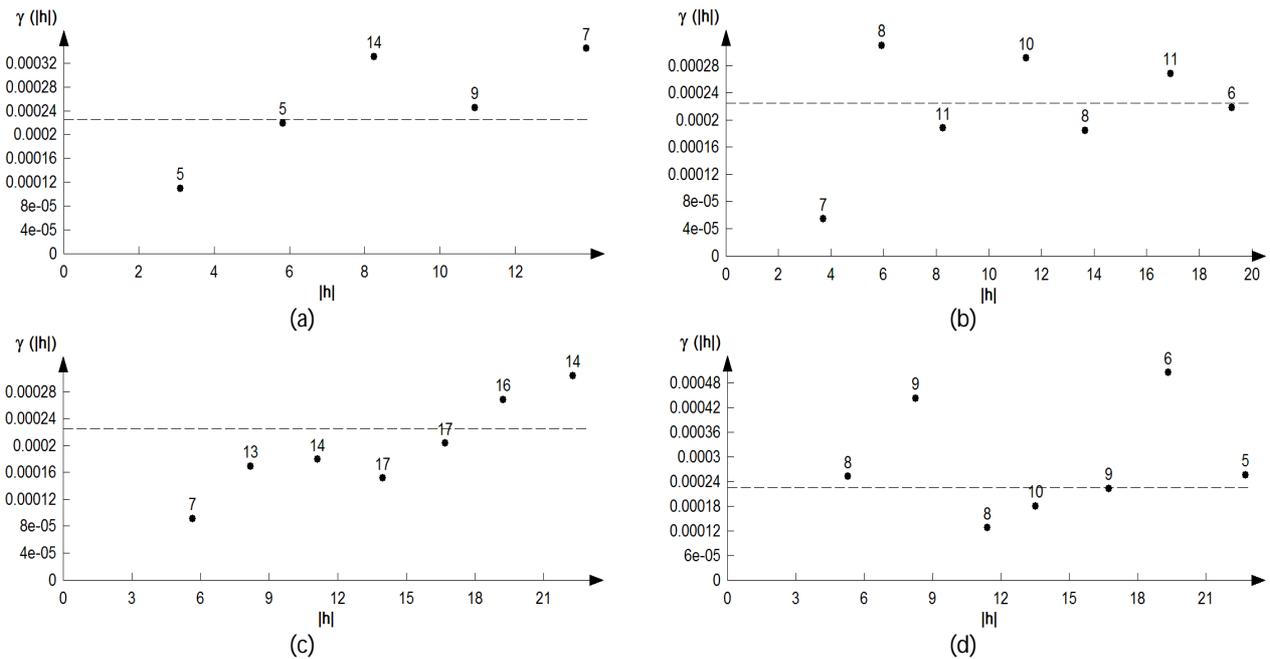


Figura 4.12.9 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per bassi contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

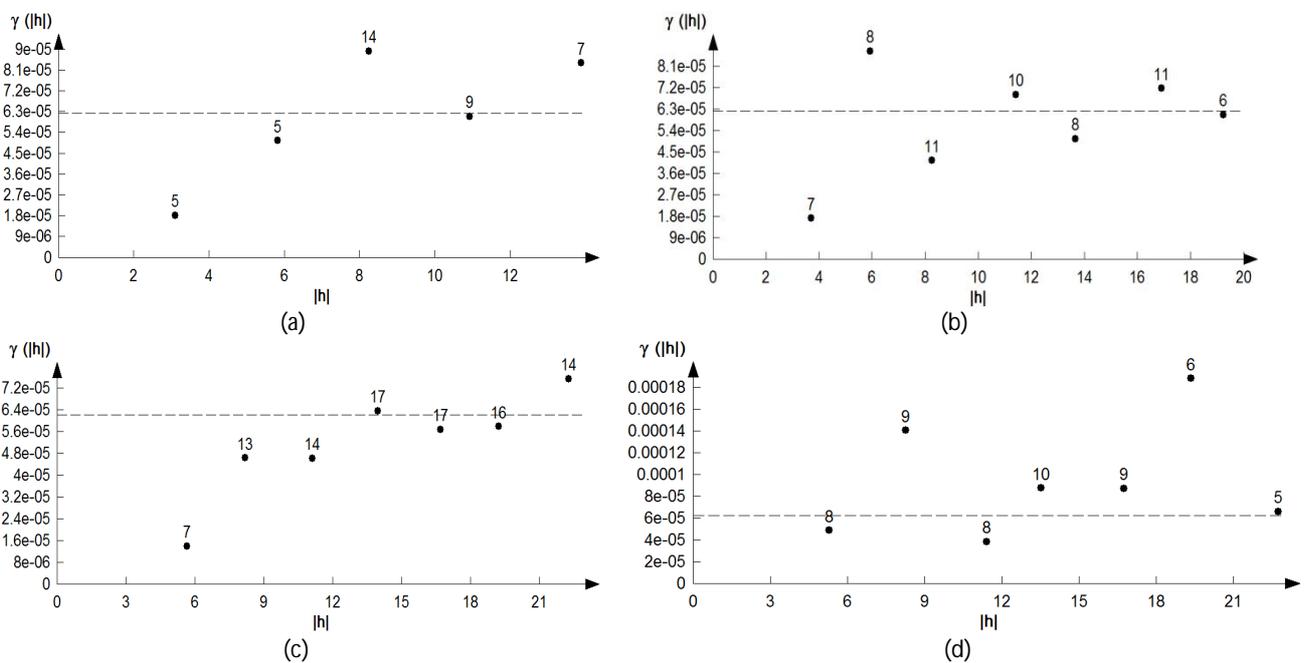


Figura 4.12.10 – Rappresentazione dei variogrammi direzionali per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per bassi contenuti di acqua; (a) Est, (b) Nord-Est, (c) Nord, (d) Nord-Ovest (VARIOWIN 2.2)

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

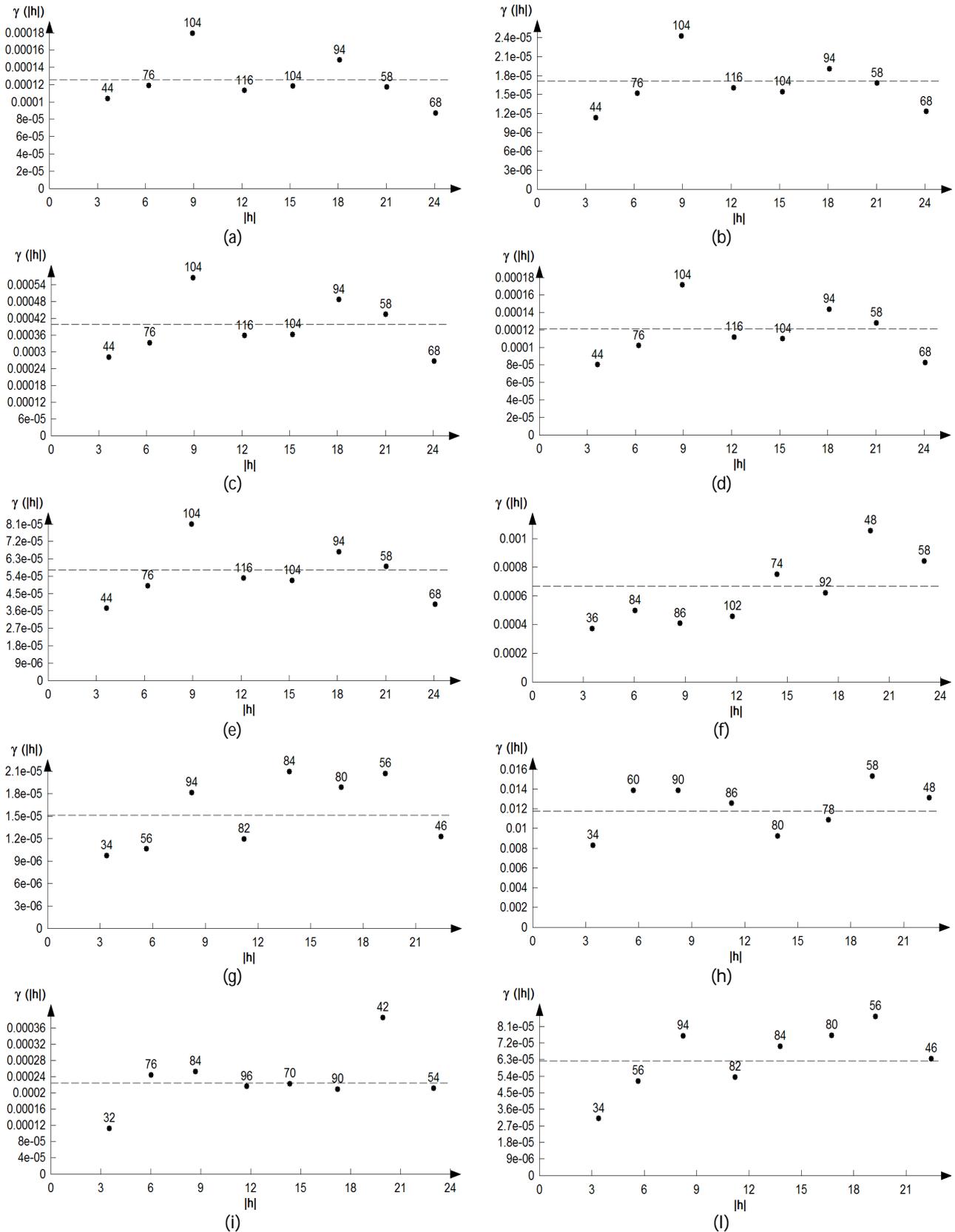


Figura 4.13 – Variogrammi omnidirezionali per alti contenuti di acqua (a) modello di Tyler & Wheatcraft [1990], (b) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (c) $A=0.30$, (d) $A=0.45$, (e) $A=0.60$ e per bassi contenuti d'acqua (f) modello di Tyler & Wheatcraft [1990], (g) modello di Rieu & Sposito [1991], modello Pore-Solid Fractal [2000] con un parametro (h) $A=0.30$, (i) $A=0.45$, (l) $A=0.60$ (VARIOWIN 2.2)

4.4.3.2 Identificazione del variogramma modello

Per descrivere e caratterizzare in maniera completa i variogrammi sperimentali prodotti dai dati a nostra disposizione è stato necessario ricondurli ad un modello matematico. In questo modo è stato creato per ogni campione di dati della variabile dimensione frattale un variogramma modello (fig. 4.14), il quale rappresenta la funzione matematica che meglio descrive la semivarianza (γ) sulla base della distribuzione dei punti nel variogramma sperimentale. Esso è stato definito per ciascun modello frattale per mezzo dell'individuazione dei tre parametri fondamentali (*nugget*, *sill*, *range*), i cui valori sono necessari per una corretta applicazione delle procedure di interpolazione. I risultati relativi alle caratteristiche dei variogrammi sperimentali (*lag*, numero di *lag*), alla scelta del modello di variogramma e dei rispettivi parametri sono riportati nella tabella 4.9. La grandezza del *lag* varia tra 3m, per modelli frattali con alti contenuti di acqua, a 2.8m, per modelli con bassi contenuti di acqua. Il numero dei *lags* è uguale a 8 in tutti i casi, di conseguenza la distanza entro la quale il variogramma sperimentale fornisce la correlazione spaziale tra i punti campionati risulta essere pari a 22.4m – 24m.

La tipologia di modello di variogramma scelto è per tutti i campioni di tipo sferico il quale presenta un comportamento lineare all'origine e la tangente all'origine raggiunge il *sill* a 2/3 del *range*. I modelli sferici scelti sono caratterizzati da un valore del *range* compreso tra 11m - 15m per i modelli con alti contenuti di acqua e 12m – 14m per i modelli frattali con bassi contenuti di acqua. Oltre tale distanza, quindi, i dati sperimentali possono considerarsi statisticamente indipendente. Il *sill* e il *nugget* variano rispettivamente tra 1.22E-04 - 1.74E-05 e 1.96E-04 - 6.0E-06 per alti contenuti di acqua e 0.0131 - 7.7E-05 e 4.8E-03 - 8.5E-06 per bassi contenuti di acqua. La tabella 4.10 mostra la percentuale del *nugget* relativa al *sill* totale in quanto per valori superiori al 75% - 80% non si ha alcun vantaggio ad utilizzare il *Kriging*, per cui si potrebbe ricorrere ad altri metodi di interpolazione (metodi deterministici). Nessun valore risulta essere maggiore del 75%: conseguentemente l'interpolazione mediante *Kriging* può fornire risultati attendibili. Inoltre, per alti contenuti di acqua, il modello frattale con il minore effetto *nugget* relativo (espresso in percentuale del *sill* totale) è il modello di *Rieu & Sposito* [1991]; differentemente, per bassi contenuti di acqua il valore più piccolo si riscontra per il modello di *Bird et al.* [2000] con parametro $A=0.30$: poiché maggiore è l'effetto *nugget* relativo, più ristretta

è la zona di correlazione spaziale fra i campioni e meno efficace risulterà la successiva interpolazione con il *Kriging*, i modelli frattali precedentemente descritti potrebbero dare i risultati più attendibili. La tabella 4.11, infine, descrive le equazioni di ciascun modello di variogramma sferico associato a tutti i set di dati a disposizione.

Mod. frat.	Lag (m)	N. lags	Mod. variog	Range (m)	Partial sill	Sill	Nugget
TW(I)	3	8	Sferico	11	4.1E-05	1.27E-04	8.6E-05
RS(I)	3	8	Sferico	13	1.14E-05	1.74E-05	6.0E-06
PSF _{0.30} (I)	3	8	Sferico	14.5	2.0E-04	3.96E-04	1.96E-04
PSF _{0.45} (I)	3	8	Sferico	15	6.8E-05	1.22E-04	5.4E-05
PSF _{0.60} (I)	3	8	Sferico	14.5	3.36E-05	5.78E-05	2.42E-05
TW(II)	2.9	8	Sferico	14	5.2E-04	6.7E-04	1.5E-04
RS(II)	2.8	8	Sferico	12	8.7E-06	1.72E-05	8.5E-06
PSF _{0.30} (II)	2.8	8	Sferico	13	8.3E-03	0.0131	4.8E-03
PSF _{0.45} (II)	2.8	8	Sferico	12	1.55E-04	2.69E-04	1.14E-04
PSF _{0.60} (II)	2.8	8	Sferico	12	4.9E-05	7.7E-05	2.8E-05

Tabella 4.9 – Risultati finali dell'analisi variografica per alti (I) e bassi (II) contenuti di acqua

Alti contenuti di acqua					Bassi contenuti di acqua				
<i>Tyler & Wheatcraft</i> [1990]	<i>Rieu & Sposito</i> [1991]	<i>Bird et al.</i> [2000] A=0.30	<i>Bird et al.</i> [2000] A=0.45	<i>Bird et al.</i> [2000] A=0.60	<i>Tyler & Wheatcraft</i> [1990]	<i>Rieu & Sposito</i> [1991]	<i>Bird et al.</i> [2000] A=0.30	<i>Bird et al.</i> [2000] A=0.45	<i>Bird et al.</i> [2000] A=0.60
67.7%	34.5%	49.5%	44.26%	41.9%	22.4%	49.4%	18.5%	42.4%	36.4%

Tabella 4.10 – Percentuale dell'effetto nugget relativo al sill totale per ciascun modello frattale, per alti e bassi contenuti di acqua

Modello frattale	Equazione variogramma modello per alti contenuti d'acqua	Equazione variogramma modello per bassi contenuti d'acqua
Tyler & Wheatcraft [1990]	$\gamma(h) = \begin{cases} 8.6 \cdot 10^{-5} + 4.1 \cdot 10^{-5} \left(1.5 \left(\frac{h}{11} \right) - 0.5 \left(\frac{h}{11} \right)^3 \right) & h \leq 11 \\ 8.6 \cdot 10^{-5} + 4.1 \cdot 10^{-5} & h > 11 \end{cases}$	$\gamma(h) = \begin{cases} 1.5 \cdot 10^{-4} + 5.2 \cdot 10^{-4} \left(1.5 \left(\frac{h}{14} \right) - 0.5 \left(\frac{h}{14} \right)^3 \right) & h \leq 14 \\ 1.5 \cdot 10^{-4} + 5.2 \cdot 10^{-4} & h > 14 \end{cases}$
Rieu & Sposito [1991]	$\gamma(h) = \begin{cases} 6 \cdot 10^{-6} + 1.14 \cdot 10^{-5} \left(1.5 \left(\frac{h}{11} \right) - 0.5 \left(\frac{h}{11} \right)^3 \right) & h \leq 13 \\ 6 \cdot 10^{-6} + 1.14 \cdot 10^{-5} & h > 13 \end{cases}$	$\gamma(h) = \begin{cases} 8.5 \cdot 10^{-6} + 8.7 \cdot 10^{-6} \left(1.5 \left(\frac{h}{12} \right) - 0.5 \left(\frac{h}{12} \right)^3 \right) & h \leq 12 \\ 8.5 \cdot 10^{-6} + 8.7 \cdot 10^{-6} & h > 12 \end{cases}$
Bird et al. [2000] A=0.30	$\gamma(h) = \begin{cases} 1.96 \cdot 10^{-4} + 2 \cdot 10^{-4} \left(1.5 \left(\frac{h}{14.5} \right) - 0.5 \left(\frac{h}{14.5} \right)^3 \right) & h \leq 14.5 \\ 1.96 \cdot 10^{-4} + 2 \cdot 10^{-4} & h > 14.5 \end{cases}$	$\gamma(h) = \begin{cases} 4.8 \cdot 10^{-3} + 8.3 \cdot 10^{-3} \left(1.5 \left(\frac{h}{11} \right) - 0.5 \left(\frac{h}{11} \right)^3 \right) & h \leq 13 \\ 4.8 \cdot 10^{-3} + 8.3 \cdot 10^{-3} & h > 13 \end{cases}$
Bird et al. [2000] A=0.45	$\gamma(h) = \begin{cases} 5.4 \cdot 10^{-5} + 6.8 \cdot 10^{-5} \left(1.5 \left(\frac{h}{15} \right) - 0.5 \left(\frac{h}{15} \right)^3 \right) & h \leq 15 \\ 5.4 \cdot 10^{-5} + 6.8 \cdot 10^{-5} & h > 15 \end{cases}$	$\gamma(h) = \begin{cases} 1.14 \cdot 10^{-4} + 1.55 \cdot 10^{-4} \left(1.5 \left(\frac{h}{12} \right) - 0.5 \left(\frac{h}{12} \right)^3 \right) & h \leq 12 \\ 1.14 \cdot 10^{-4} + 1.55 \cdot 10^{-4} & h > 12 \end{cases}$
Bird et al. [2000] A=0.60	$\gamma(h) = \begin{cases} 2.42 \cdot 10^{-5} + 3.36 \cdot 10^{-5} \left(1.5 \left(\frac{h}{14.5} \right) - 0.5 \left(\frac{h}{14.5} \right)^3 \right) & h \leq 14.5 \\ 2.42 \cdot 10^{-5} + 3.36 \cdot 10^{-5} & h > 14.5 \end{cases}$	$\gamma(h) = \begin{cases} 2.8 \cdot 10^{-5} + 4.9 \cdot 10^{-5} \left(1.5 \left(\frac{h}{12} \right) - 0.5 \left(\frac{h}{12} \right)^3 \right) & h \leq 12 \\ 2.8 \cdot 10^{-5} + 4.9 \cdot 10^{-5} & h > 12 \end{cases}$

Tabella 4.11 – Equazioni dei variogrammi modello di tipo sferico per ciascun modello frattale

CASO DI STUDIO: ANALISI SPERIMENTALE CONDOTTA SUL BACINO DEL TORRENTE TURBOLO

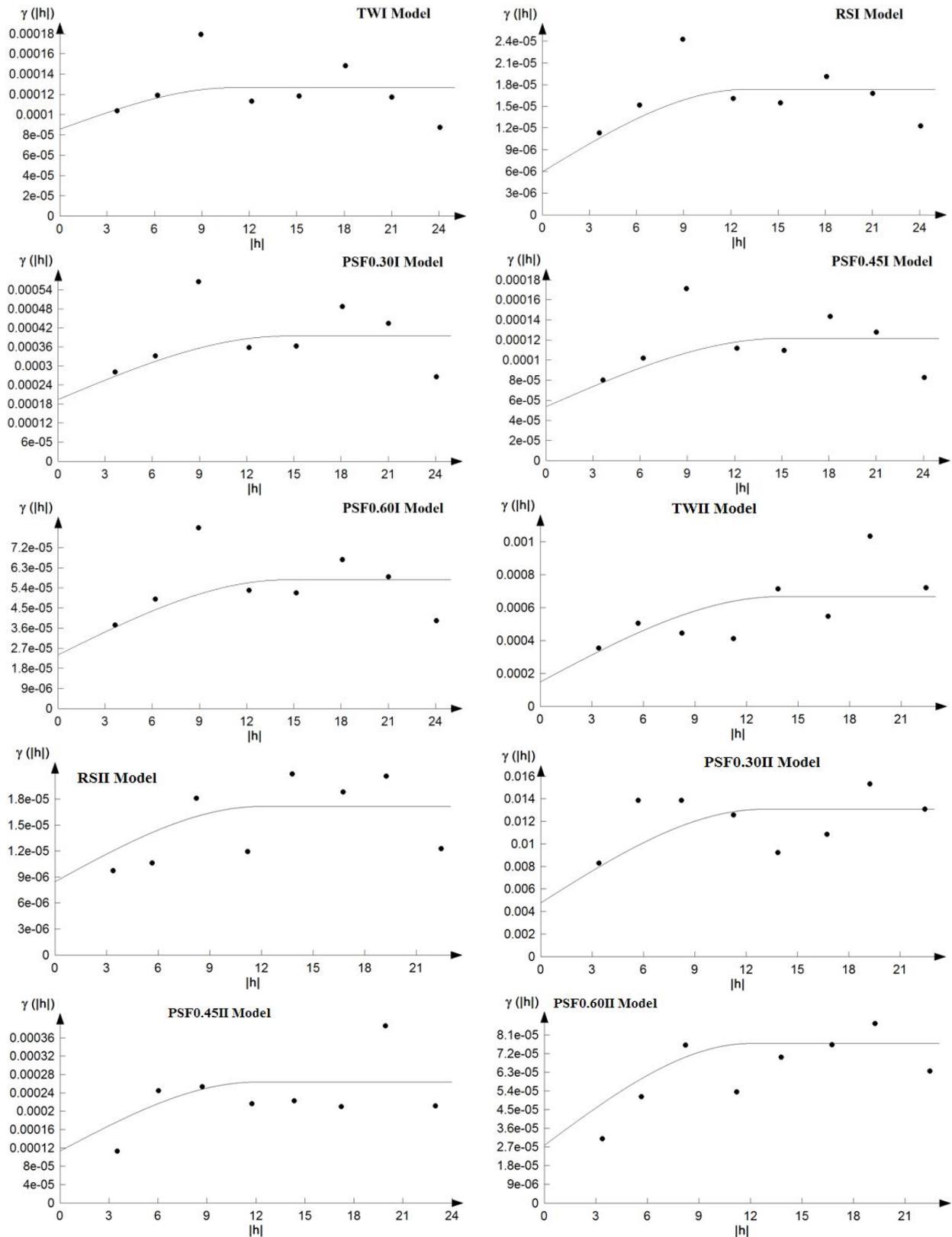


Figura 4.14 – Rappresentazione dei variogrammi sperimentali omnidirezionali e rispettivi modelli per alti contenuti di acqua (I) per bassi contenuti d'acqua (II) (VARIOWIN 2.2)

4.4.4 Validazione incrociata (*Cross Validation*)

La convalida del modello di variogramma utilizzato per ciascun campione è stata portata avanti attraverso la valutazione di tre parametri specifici: la media degli errori (*ME*), la media degli errori standardizzati (*MSE*) e la varianza degli errori standardizzati (*VSE*). Affinché il modello non risulti grossolanamente errato, i primi due parametri devono risultare il più possibile vicino a zero, mentre la varianza degli errori standardizzati deve essere prossima a uno. Le tabelle 4.11.1 – 4.11.10 riportano i risultati riferiti a ciascun modello frattale. I valori sottolineati identificano il modello di variogramma e i rispettivi parametri che meglio interpretano il singolo variogramma sperimentale esaminato per ciascun modello frattale analizzato.

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	13.5	8.0E-05	4.7E-05	1.27E-04	3.98E-04	1.0843	3.55E-02
Sferico	14	8.0E-05	4.7E-05	1.27E-04	3.55E-04	1.0924	3.19E-02
Sferico	13	8.0E-05	4.7E-05	1.27E-04	3.82E-04	1.0616	3.38E-02
Sferico	13	8.6E-05	4.1E-05	1.27E-04	4.23E-04	1.0487	3.68E-02
Sferico	12	8.6E-05	4.1E-05	1.27E-04	4.25E-04	1.0287	3.68E-02
<u>Sferico</u>	<u>11</u>	<u>8.6E-05</u>	<u>4.1E-05</u>	<u>1.27E-04</u>	<u>2.53E-04</u>	<u>1.0089</u>	<u>2.16E-02</u>

Tabella 4.11.1 – Risultati della cross-validation per il modello di Tyler & Wheatcraft [1990] e per alti contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	11	6.0E-06	1.14E-05	1.74E-05	1.0E-06	0.9125	4.92E-03
<u>Sferico</u>	<u>13</u>	<u>6.0E-06</u>	<u>1.14E-05</u>	<u>1.74E-05</u>	<u>4.7E-05</u>	<u>1.0056</u>	<u>1.6E-02</u>

Tabella 4.11.2 – Risultati della cross-validation per il modello di Rieu & Sposito [1991] e per alti contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	13	1.84E-04	2.2E-04	4.04E-04	6.13E-04	0.9479	3.43E-02
Sferico	13	1.96E-04	2.0E-04	3.96E-04	6.99E-04	0.9566	3.81E-02
Sferico	14	1.96E-04	2.0E-05	3.96E-04	7.29E-04	0.9883	3.98E-02
<u>Sferico</u>	<u>14.5</u>	<u>1.96E-04</u>	<u>2.0E-05</u>	<u>3.96E-04</u>	<u>7.22E-04</u>	<u>1.0035</u>	3.94E-02

Tabella 4.11.3 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per alti contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	14.5	6.2E-05	6.0E-05	1.22E-04	3.80E-04	0.9648	3.69E-02
Sferico	13	5.4E-05	6.8E-05	1.22E-04	2.90E-04	0.9349	3.0E-02
Sferico	14	5.4E-05	6.8E-05	1.22E-04	3.05E-04	0.9707	3.15E-02
Sferico	14.5	5.4E-05	6.8E-05	1.22E-04	3.02E-04	0.9868	3.11E-02
<u>Sferico</u>	<u>15</u>	<u>5.4E-05</u>	<u>6.8E-05</u>	<u>1.22E-04</u>	<u>2.80E-04</u>	<u>1.0042</u>	<u>2.97E-02</u>

Tabella 4.11.4 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per alti contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	14.5	2.52E-05	2.82E-05	5.34E-05	2.18E-04	1.0586	3.22E-02
Sferico	14	2.52E-05	2.82E-05	5.34E-05	2.14E-04	1.0423	3.26E-02
Sferico	14	2.42E-05	3.36E-05	5.78E-05	1.71E-04	0.9840	2.67E-02
<u>Sferico</u>	<u>14.5</u>	<u>2.42E-05</u>	<u>3.36E-05</u>	<u>5.78E-05</u>	<u>1.69E-04</u>	<u>1.0008</u>	<u>2.63E-02</u>

Tabella 4.11.5 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per alti contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	14	2.1E-04	4.6E-04	6.7E-04	-3.42E-04	0.9084	-1.85E-02
Sferico	13	2.1E-04	4.6E-04	6.7E-04	-3.36E-04	0.8701	-1.91E-02
Sferico	15	2.1E-04	4.6E-04	6.7E-04	-2.88E-04	0.9443	-1.61E-02
Sferico	15	1.5E-04	5.2E-04	6.7E-04	-1.37E-04	1.0518	-1.04E-02
<u>Sferico</u>	<u>14</u>	<u>1.5E-04</u>	<u>5.2E-04</u>	<u>6.7E-04</u>	<u>-1.99E-04</u>	<u>1.0048</u>	<u>-1.33E-02</u>

Tabella 4.11.6 – Risultati della cross-validation per il modello di Tyler & Wheatcraft [1990] per bassi contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	13	6.2 E-06	9.2E-06	1.54E-05	-1.41E-04	1.2635	-3.73E-02
Sferico	13	6.2 E-06	1.0E-05	1.62E-05	-1.38E-04	1.2264	-3.57E-02
<u>Sferico</u>	<u>12</u>	<u>8.5 E-06</u>	<u>8.7E-06</u>	<u>1.72E-05</u>	<u>-1.31E-04</u>	<u>1.0018</u>	<u>-3.31E-02</u>

Tabella 4.11.7 – Risultati della cross-validation per il modello di Rieu & Sposito [1991] per bassi contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	14	2.2E-03	9.7E-03	0.0119	-1.33E-03	1.3780	-4.15E-03
Sferico	10	4.8E-03	8.3E-03	0.0131	-7.5E-04	0.8884	-1.08E-03
Sferico	12	4.8E-03	8.3E-03	0.0131	-4.8E-04	0.9803	8.12E-03
<u>Sferico</u>	<u>13</u>	<u>4.8E-03</u>	<u>8.3E-03</u>	<u>0.0131</u>	<u>-2.1E-04</u>	<u>1.0166</u>	<u>5.02E-03</u>

Tabella 4.11.8 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per bassi contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	12	1.14E-04	1.5E-04	2.64E-04	-2.95E-04	1.0225	-1.51E-02
<u>Sferico</u>	<u>12</u>	<u>1.14E-04</u>	<u>1.55E-04</u>	<u>2.69E-04</u>	<u>-2.95E-04</u>	<u>1.0084</u>	<u>-1.48E-02</u>

Tabella 4.11.9 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per bassi contenuti di acqua

Modello	Range (m)	Nugget	Partial sill	Sill	ME	VSE	MSE
Sferico	13	9.1E-06	5.4E-05	6.31E-05	-2.05E-04	1.7077	-1.86E-02
Sferico	12	2.8E-06	6.5E-05	6.78E-05	-8.2E-05	1.7461	1.0E-03
Sferico	13	2.52E-05	4.0E-05	6.52E-05	-2.4E-04	1.2269	-2.86E-02
Sferico	12	8.4E-06	6.4E-05	7.24E-05	-1.27E-04	1.4350	-8.01E-03
Sferico	14	2.52E-05	4.0E-05	6.52E-05	-1.73E-04	1.2963	-1.93E-02
Sferico	13	2.8E-05	3.5E-05	6.3E-05	-2.45E-04	1.2058	-3.03E-02
Sferico	12	2.8E-05	3.5E-05	6.3E-05	-2.14E-04	1.1661	-2.64E-02
Sferico	11	2.8E-05	3.5E-05	6.3E-05	-4.54E-04	1.1903	-5.29E-02
Sferico	12	2.8E-05	3.9E-05	6.7E-05	-2.1E-04	1.1185	-2.49E-02
Sferico	13	2.8E-05	3.9E-05	6.7E-05	-2.43E-04	1.1602	-2.89E-02
<u>Sferico</u>	<u>12</u>	<u>2.8E-05</u>	<u>4.9E-05</u>	<u>7.7E-05</u>	<u>-2.0E-04</u>	<u>1.0171</u>	<u>-2.16E-02</u>

Tabella 4.11.10 – Risultati della cross-validation per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per bassi contenuti di acqua

4.4.5 Elaborazione delle mappe ottenute mediante l'interpolatore *Kriging* ordinario

Per poter ricostruire la superficie che rappresenta la variabilità spaziale della dimensione frattale è stato necessario interpolare i dati disponibili per stimare i valori dove non si hanno campioni. In questo senso il *Kriging* ordinario ci viene incontro, in quanto dà una soluzione al problema della stima basato su un modello continuo di variazione spaziale stocastica. Esso fa il miglior uso della conoscenza della variabile, prendendo in considerazione il modo in cui la dimensione frattale varia nello spazio attraverso il modello del variogramma che è stato scelto e validato. Una volta scelta e costruita la griglia di interpolazione (0.25m x 0.25m), si è proceduto con l'interpolazione in corrispondenza di ciascun nodo della griglia allo scopo di realizzare un *grid* regolare di punti interpolati. La mappatura della variabile dimensione frattale (bi e tri-dimensionale) per ciascun modello frattale, per alti e bassi contenuti di acqua, è riportata nelle figure 4.15.1- 4.15.10.

Per il modello di *Tyler & Wheatcraft* [1990], la dimensione frattale assume un valore compreso tra 2.9314 e 2.9779 per alti contenuti di acqua (fig. 4.15.1); essa si presenta con valori molto bassi in due ben distinte zone dell'area esaminata e, precisamente, nella parte Sud-Ovest e Nord-Ovest. I valori più alti, compresi cioè tra 2.96 e 2.978, si hanno in diversi punti del campo sperimentale, in modo particolare nella zona centrale e a Sud-Est. Sempre per il modello di *Tyler & Wheatcraft* [1990] e per bassi contenuti di acqua, la dimensione frattale varia tra 2.7423 e 2.8362 (fig. 4.15.2); in questo caso l'area oggetto di studio presenta per la maggior parte valori molto bassi compresi tra 2.7423 e 2.78 ad eccezione della porzione posta più a nord e di una piccola regione collocata a Sud-Ovest.

Esaminando il modello *Rieu & Sposito* [1991], la variabile dimensione frattale assume valori compresi tra 2.97595 e 2.99311 in riferimento ad alti contenuti di acqua (fig. 4.15.3); anche in tal caso, essa presenta i valori più bassi nella parte più a nord del dominio di studio e in una piccola porzione a Sud-Ovest. I valori maggiori della variabile (2.987 – 2.99), come per il modello *Tyler & Wheatcraft* [1990], si presentano in diverse zone dell'area esaminata, a Sud-Est e nella fascia centrale. Per bassi contenuti di acqua e per lo stesso modello, il valore della dimensione frattale è compreso tra 2.9388 e 2.9538 (fig. 4.15.4). In questo caso l'area in esame può essere suddivisa in tre sotto-aree ben distinte: nella prima la variabile presenta valori molto

bassi (2.94359 e 2.947), nella seconda valori medi (2.947 e 2.949) e nell'ultima i valori maggiori (2.949 e 2.9519).

Per il modello *Pore-Solid Fractal* [2000] e parametro $A=0.30$, la variabile dimensione frattale assume un valore compreso tra 2.8959 e 2.9731 (fig. 4.15.5) per alti contenuti di acqua; i valori minori di detta grandezza si presentano nelle parti Sud-Ovest e Nord-Est dell'area oggetto di studio mentre, i valori maggiori si hanno nella zona Sud-Est e quella centrale. Per bassi contenuti di acqua e per lo stesso modello, il valore della dimensione frattale è compreso tra 2.33291 e 2.69601 (fig. 4.15.6). Osservando la mappatura, in questo caso la variabile assume i valori più bassi (compresi tra 2.33291 e 2.45) nella parte a Sud-Ovest del dominio di studio, con coordinate x comprese tra 0 e 10 m, mentre i valori medi e maggiori si possono riscontrare nella maggior parte dell'area in esame.

Per il modello *Pore-Solid Fractal* [2000] e parametro $A=0.45$ la variabile assume valori tra 2.9394 e 2.9834 (fig. 4.15.7) e tra 2.8063 e 2.8599 (fig. 4.15.8) rispettivamente per alti e bassi contenuti di acqua. Per il primo caso si può notare dalla mappatura la presenza di valori molto bassi nella parte più a nord e in una sotto-area posta a Sud-Ovest, mentre la presenza dei valori maggiori si ha in diversi punti della stessa area, e precisamente a Sud-Est e nella fascia centrale. Per bassi contenuti di acqua la dimensione frattale assume valori molto alti nella fascia centro-settentrionale posta ad Ovest (con coordinate x comprese tra 0 e 8m) e in una porzione a Sud-Est. La rimanente area è caratterizzata dalla zona Sud-Ovest e da una piccola area centrale in cui sussistono valori molto bassi.

Esaminando, infine, il modello *Pore-Solid Fractal* [2000] e parametro $A=0.60$ si può notare che la variabile presenta valori compresi tra 2.9572 e 2.988 (fig. 4.15.9), in riferimento ad alti contenuti di acqua. In tal caso la dimensione frattale assume valori molto bassi nella parte più a nord del dominio di studio e in una piccola area a Sud-Ovest. Al contrario, i valori maggiori si presentano a Sud-Est e in una vasta area centrale. Per bassi contenuti di acqua e sempre per il modello *Pore-Solid Fractal* [2000] e parametro $A=0.60$, la variabile presenta valori tra 2.8806 e 2.91007 (fig. 4.15.10). I valori più alti sono presenti in due zone opposte, cioè a Nord-Ovest e a Sud-Est. I valori minori della dimensione frattale si riscontrano in quasi tutta la parte centrale e a Sud-Ovest.

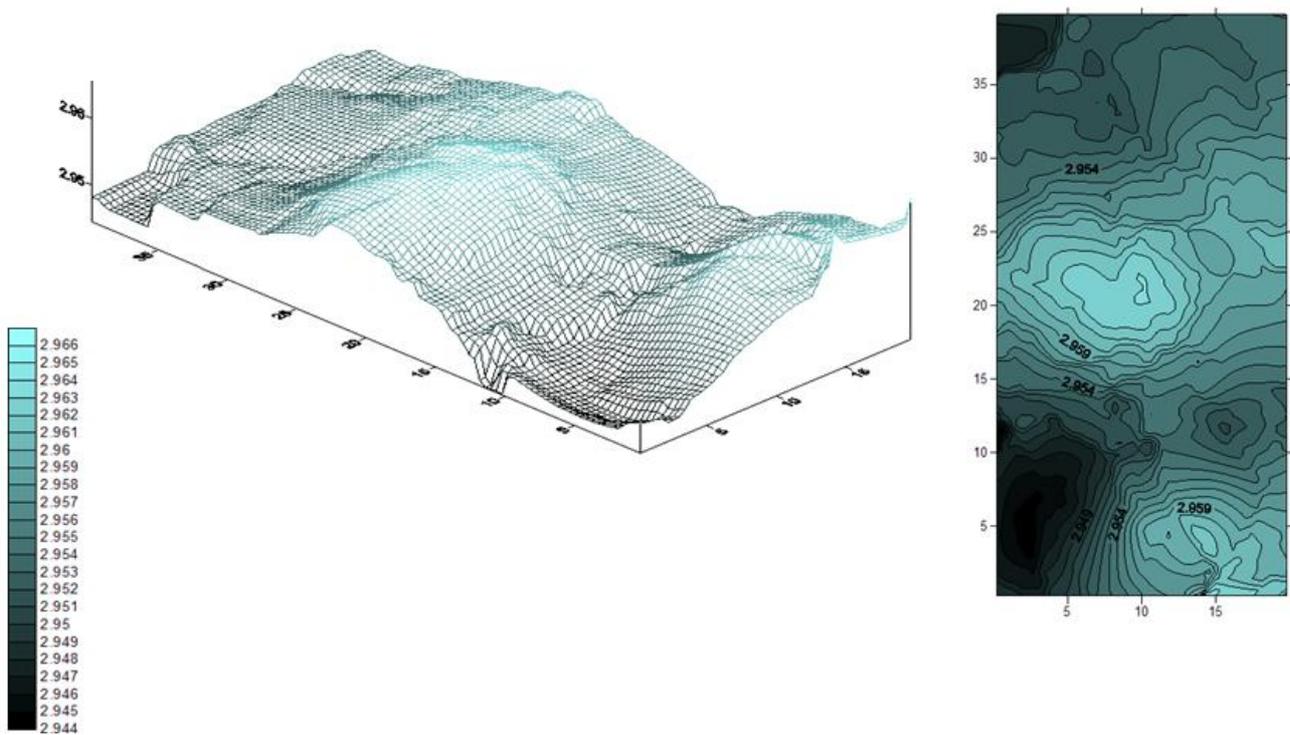


Figura 4.15.1 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello di Tyler & Wheatcraft [1990] e per alti contenuti di acqua

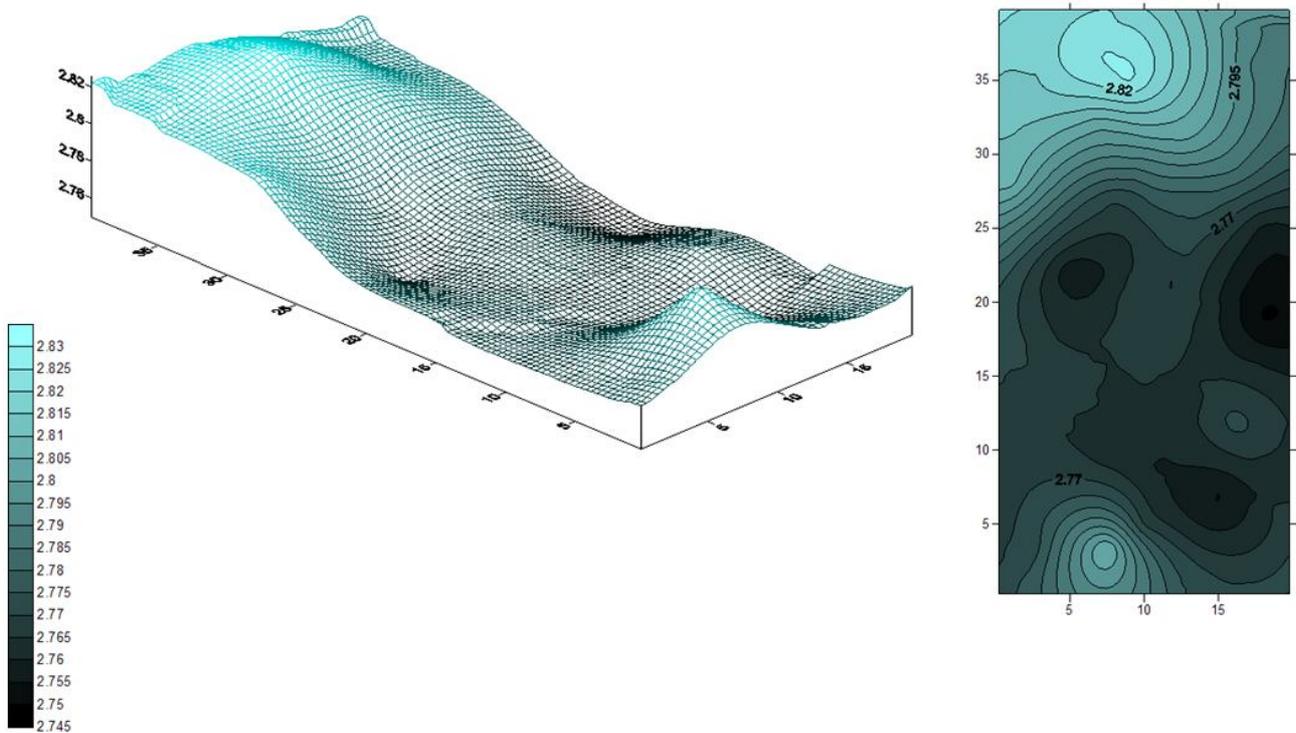


Figura 4.15.2 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello di Tyler & Wheatcraft [1990] per bassi contenuti di acqua

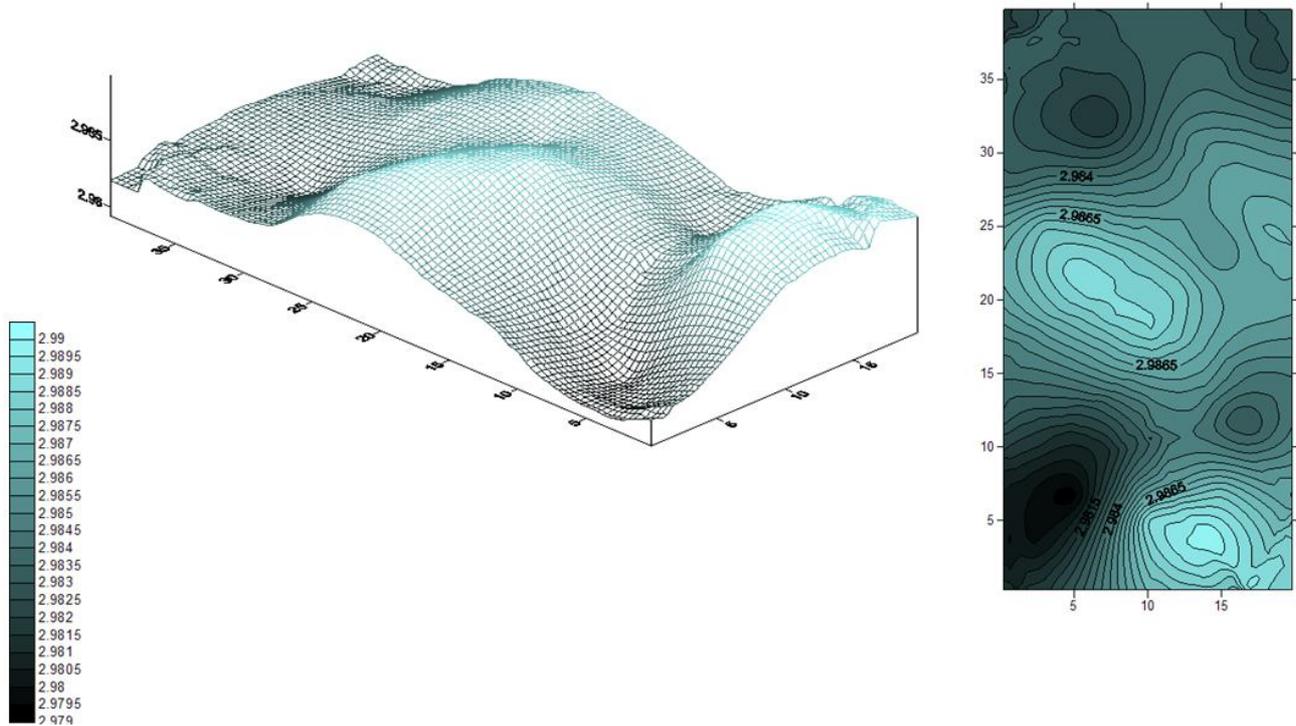


Figura 4.15.3 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello di Rieu & Sposito [1991] e per alti contenuti di acqua

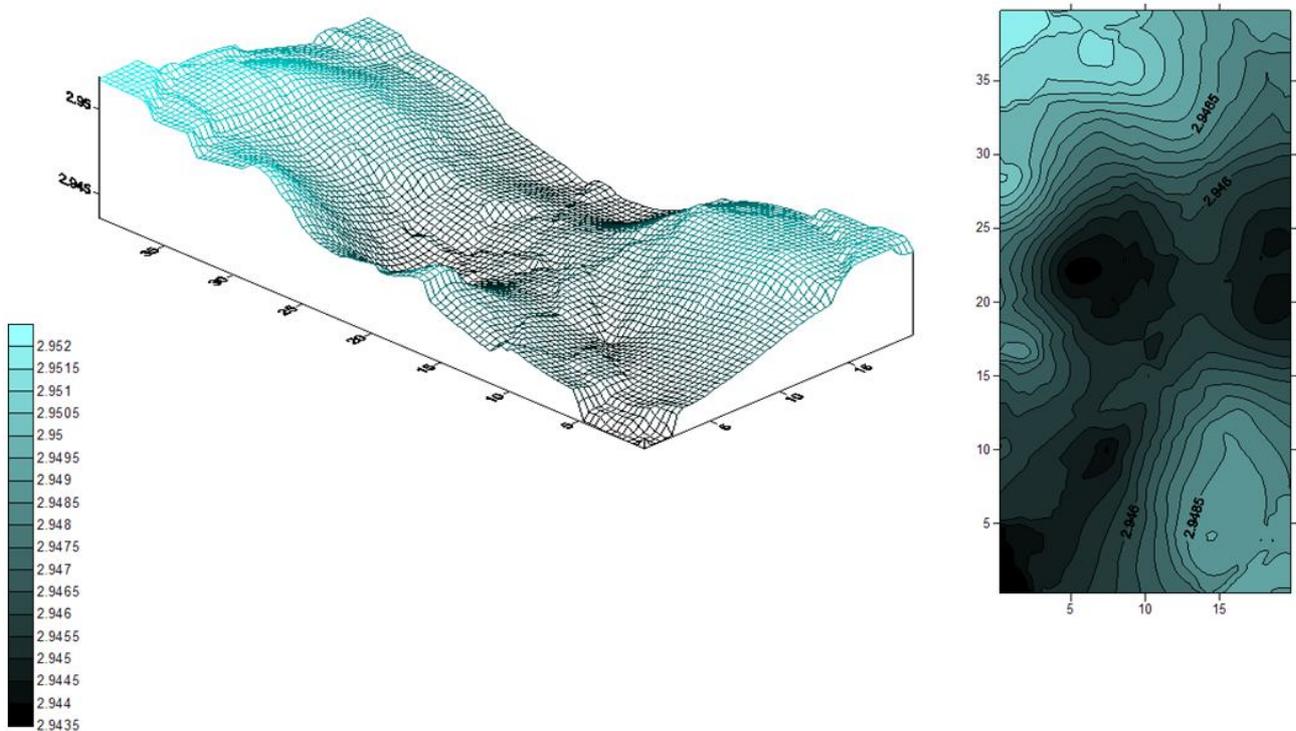


Figura 4.15.4 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello di Rieu & Sposito [1991] e per bassi contenuti di acqua

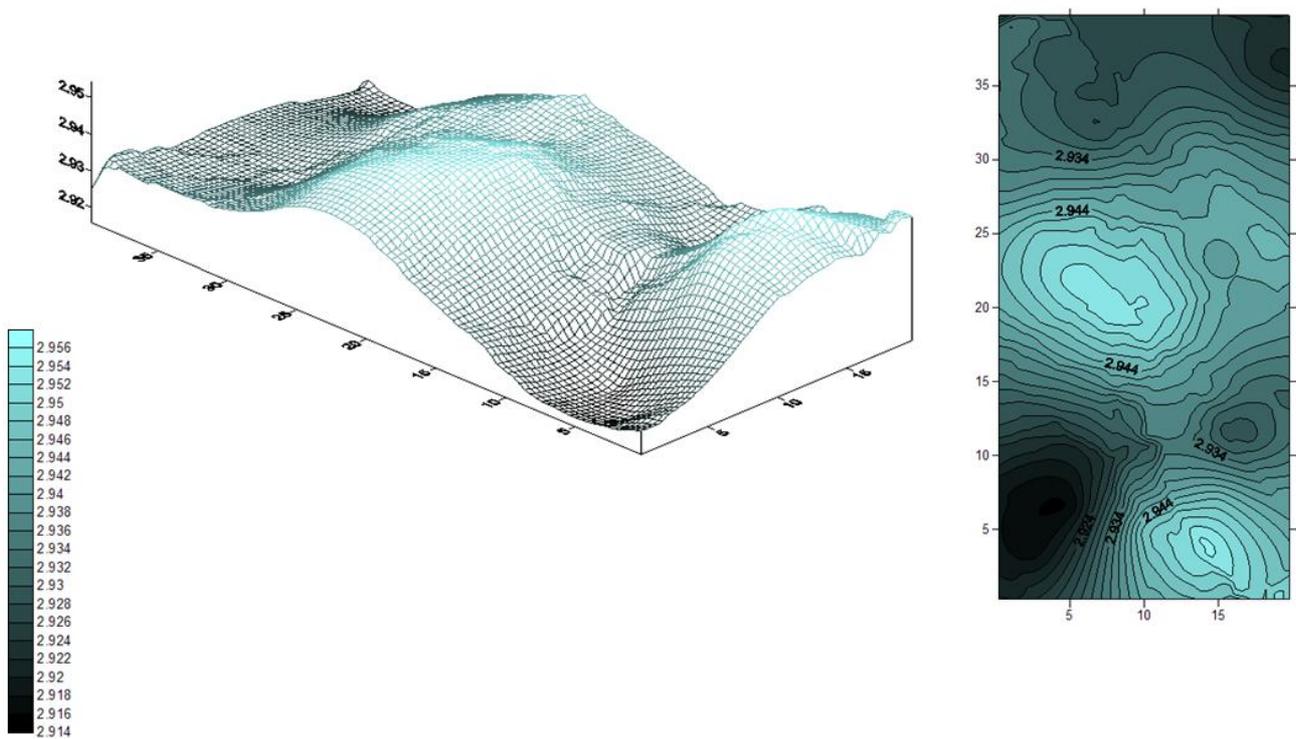


Figura 4.15.5 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per alti contenuti di acqua

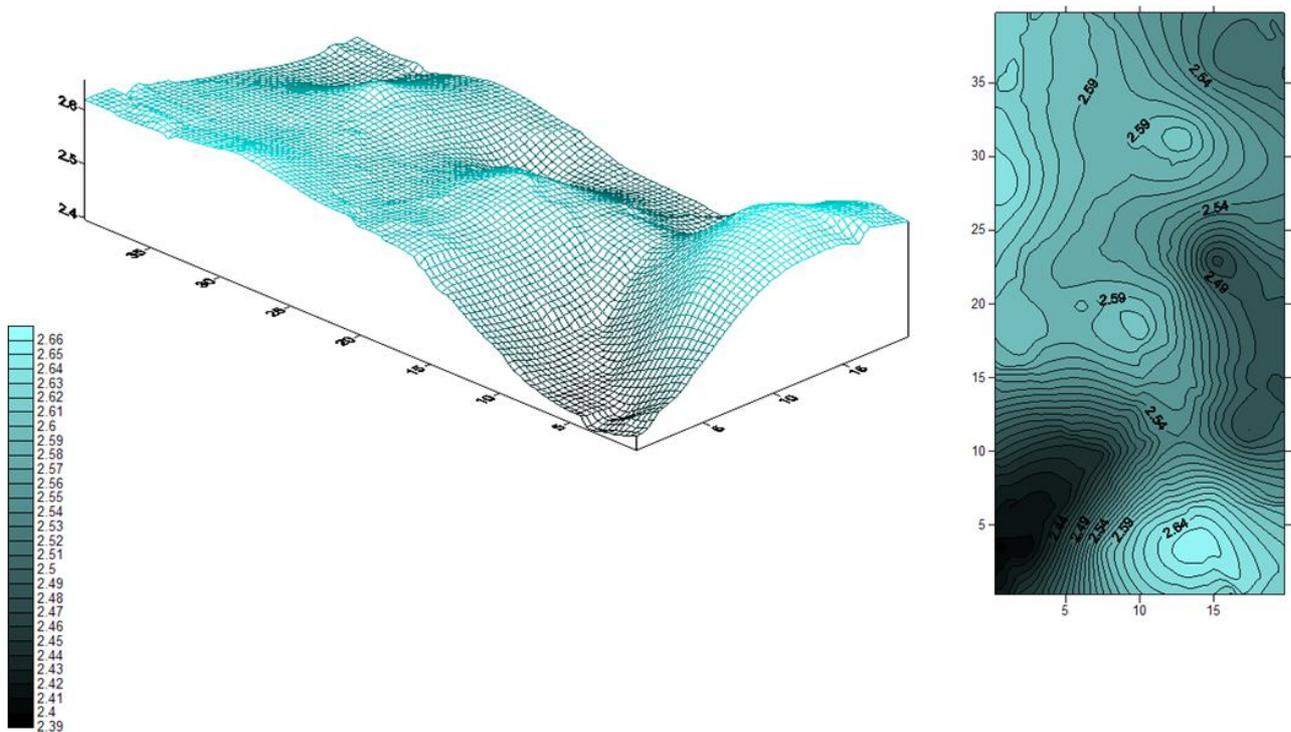


Figura 4.15.6 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per bassi contenuti di acqua

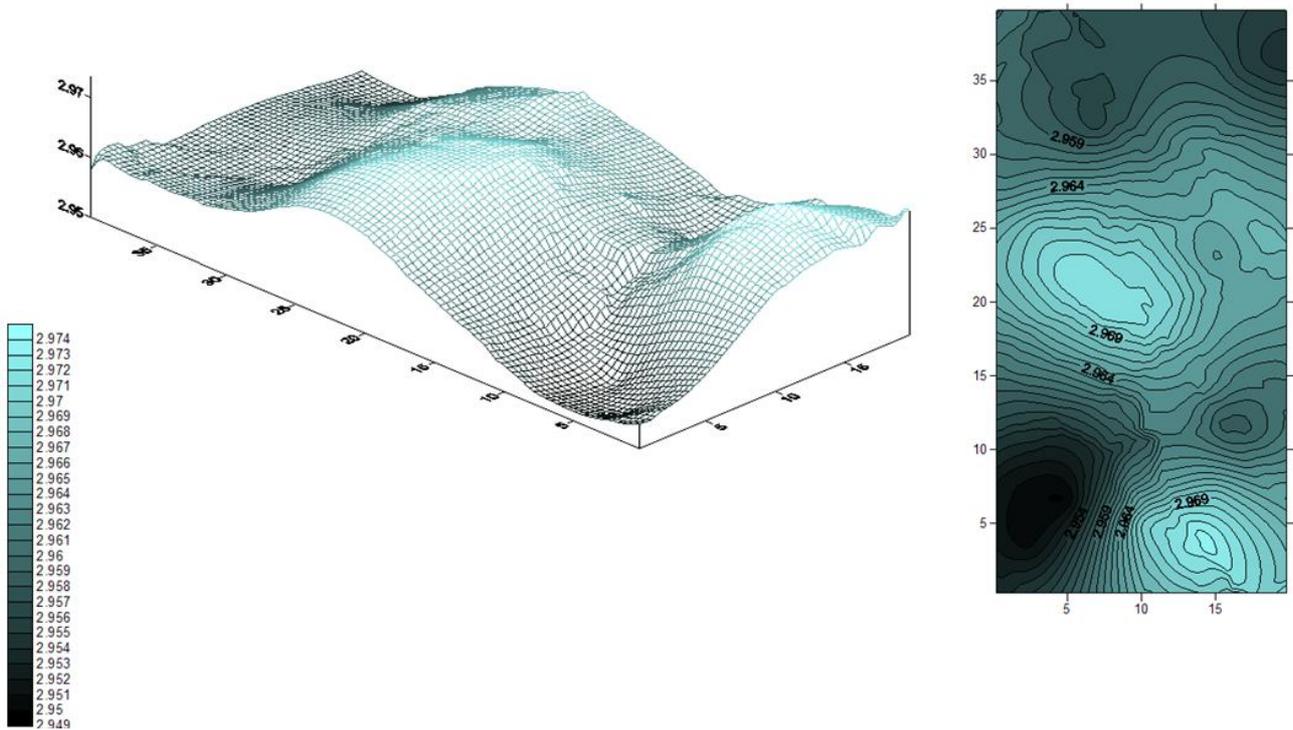


Figura 4.15.7 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per alti contenuti di acqua

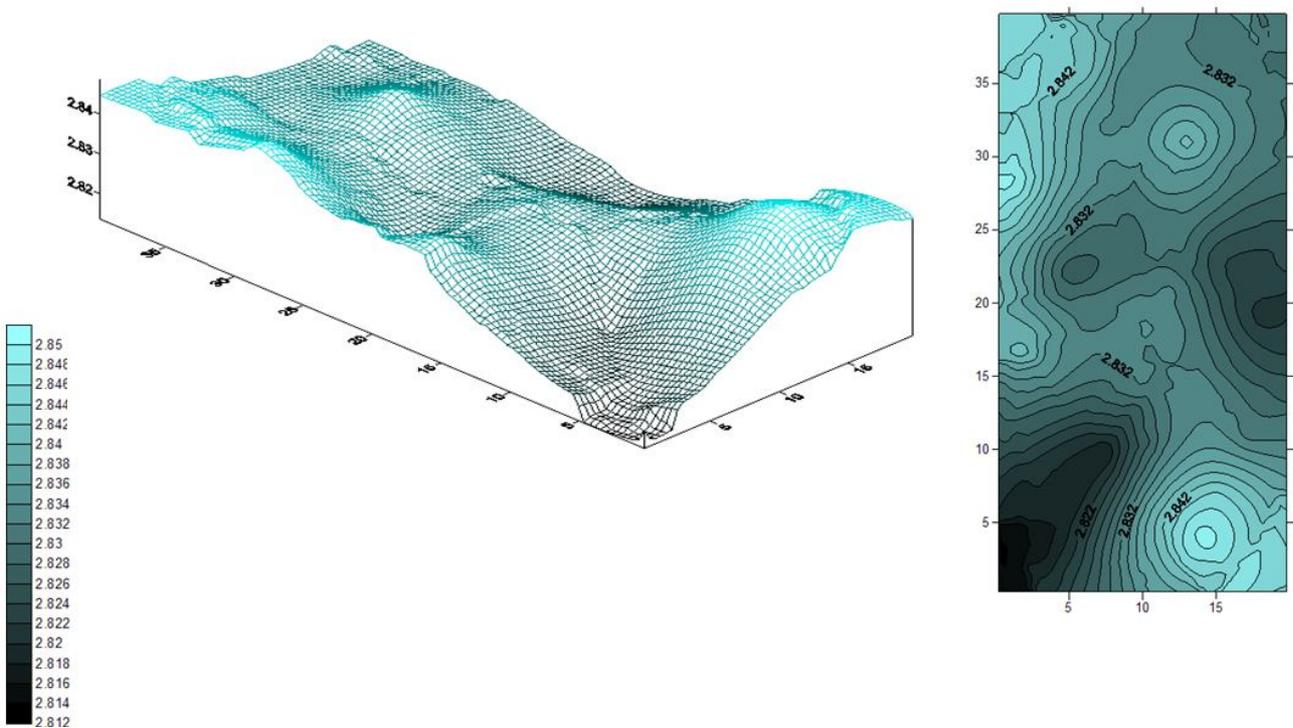


Figura 4.15.8 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per bassi contenuti di acqua

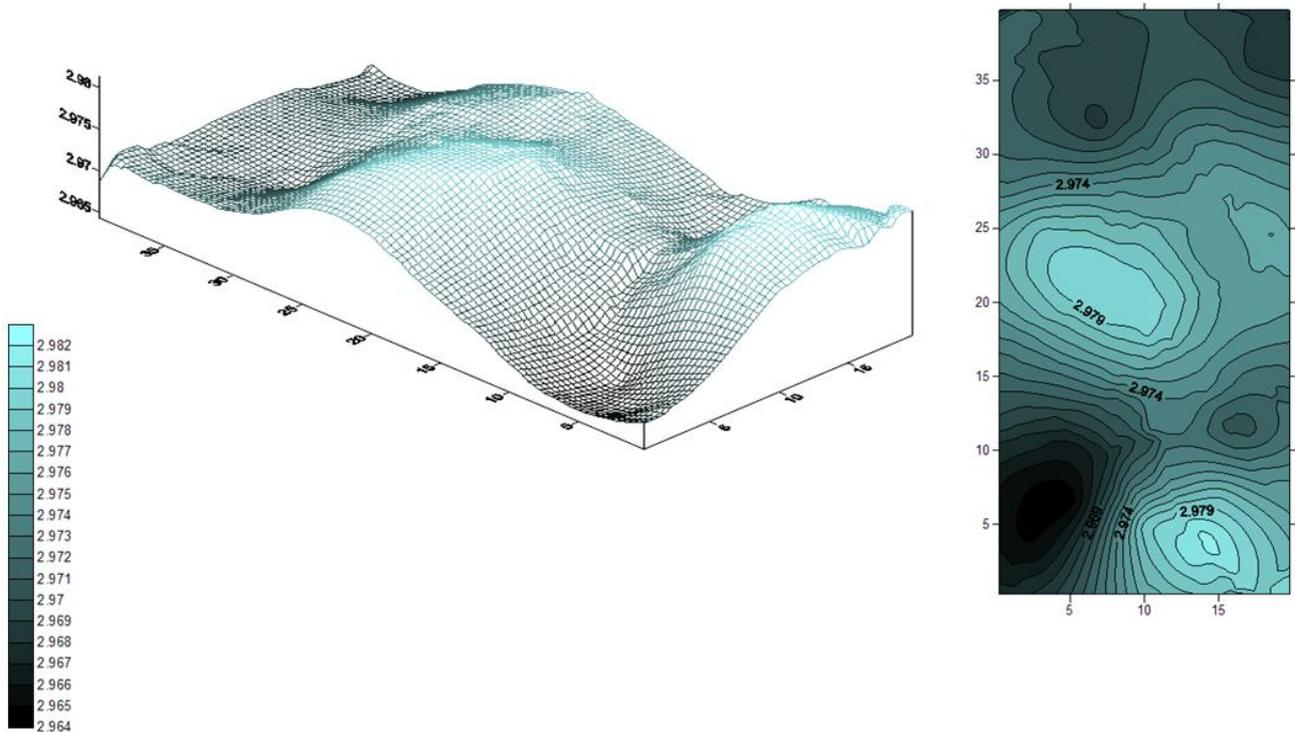


Figura 4.15.9 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per alti contenuti di acqua

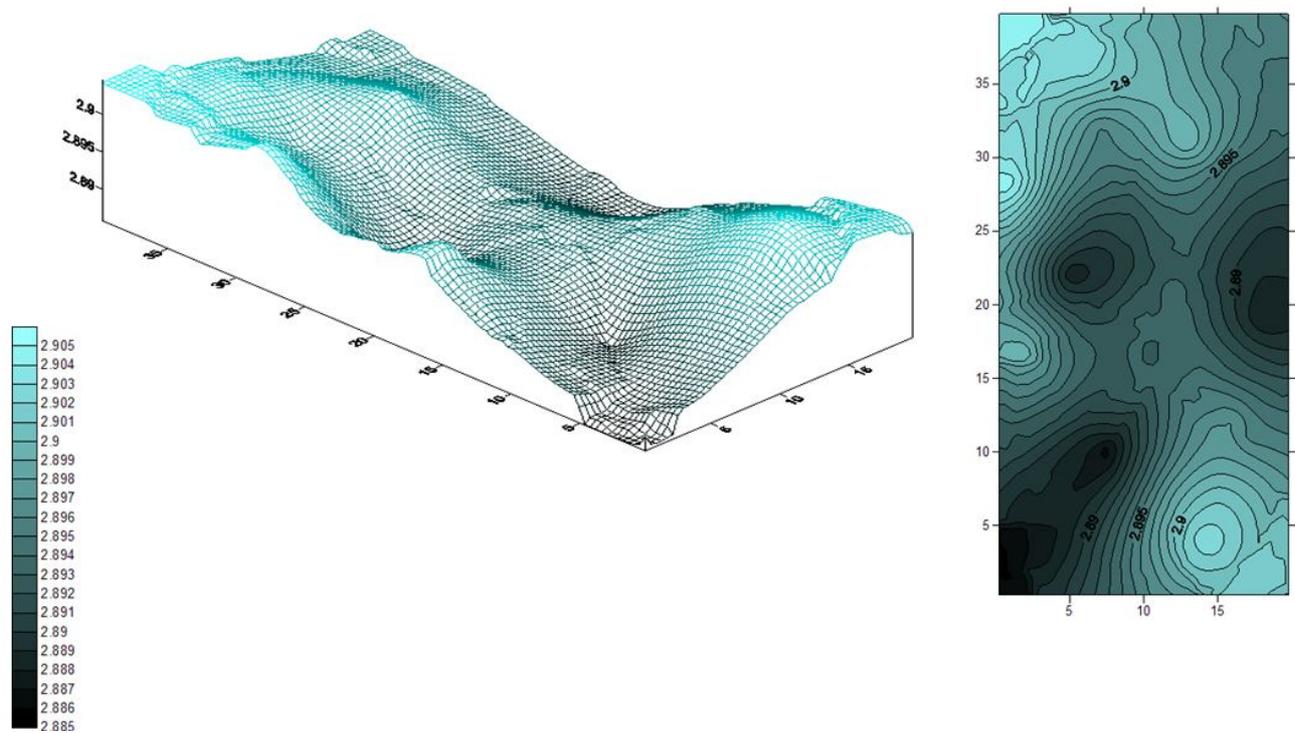


Figura 4.15.10 – Interpolazione spaziale della variabile dimensione frattale mediante Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per bassi contenuti di acqua

Nelle figure 4.16.1 – 4.16.10 è mostrata la distribuzione spaziale (bi e tri-dimensionale) della varianza estimativa, dalla quale si possono evincere le zone all'interno del dominio di studio in cui la stima effettuata è molto affidabile oppure poco attendibile. Come è facile immaginare, la maggior parte dell'area esaminata, per ciascun modello frattale e per alti e bassi contenuti d'acqua, presenta un valore della varianza molto basso in prossimità dei punti campionati, aumentando notevolmente a mano a mano che ci si avvicina ai bordi. Questo indica che le stime risultano meno accurate in corrispondenza dei limiti del dominio di studio, in quanto in tali aree l'interpolazione avviene ovviamente con minori informazioni.

Per il modello di *Tyler & Wheatcraft* [1990], la varianza assume un valore compreso tra $1.15E-04$ e $1.9E-04$ (fig. 4.16.1) per alti contenuti di acqua; per bassi contenuti di acqua (fig. 4.16.2) essa varia all'interno di un intervallo maggiore, ossia tra $2.5E-04$ e $8.0E-04$.

Analizzando la mappa della varianza relativa al modello di *Rieu & Sposito* [1991] e per alti contenuti di acqua, essa risulta compresa tra $9.0E-06$ e $2.2E-05$ (fig. 4.16.3); diversamente per bassi contenuti di acqua, essa assume valori tra $1.2E-05$ e $2.2E-05$ (fig. 4.16.4).

Per il modello di *Bird et al.* [2000] con parametro $A=0.30$, la varianza varia all'interno di un range tra $2.7E-04$ e $5.0E-04$ (fig. 4.16.5) e tra $7.5E-03$ e $1.65E-02$ (fig. 4.16.6) rispettivamente per alti e per bassi contenuti di acqua.

Il valore della varianza per il modello di *Bird et al.* [2000] con parametro $A=0.45$ è compreso tra $7.5E-05$ e $1.45E-04$ (fig. 4.16.7) per alti contenuti di acqua, tra $1.7E-04$ e $3.4E-04$ per bassi contenuti di acqua (fig. 4.16.8).

Infine, per il modello di *Bird et al.* [2000] con parametro $A=0.60$, la varianza assume un valore compreso tra $3.6E-05$ e $7.2E-05$ (fig. 4.16.9) per alti contenuti di acqua; per bassi contenuti di acqua (fig. 4.16.10) essa varia all'interno di un intervallo tra $4.5E-05$ e $9.5E-05$.

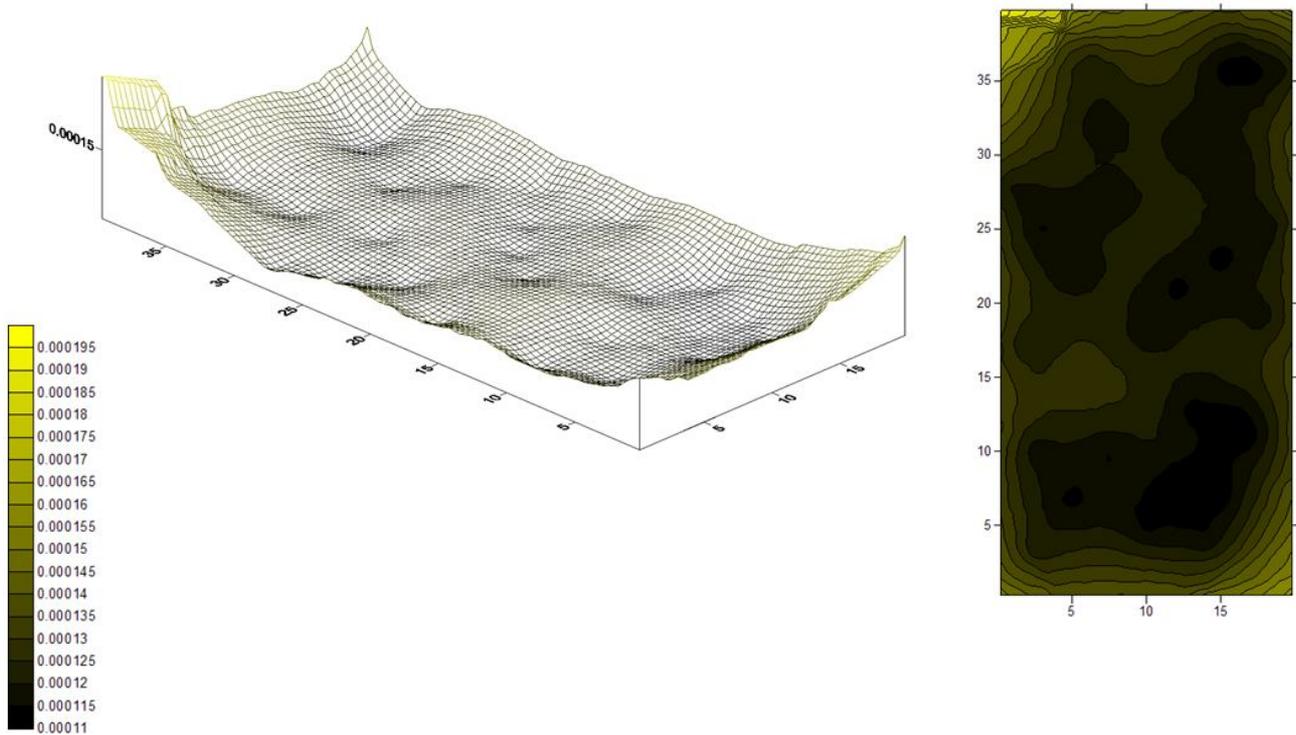


Figura 4.16.1 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello di Tyler & Wheatcraft [1990] e per alti contenuti di acqua

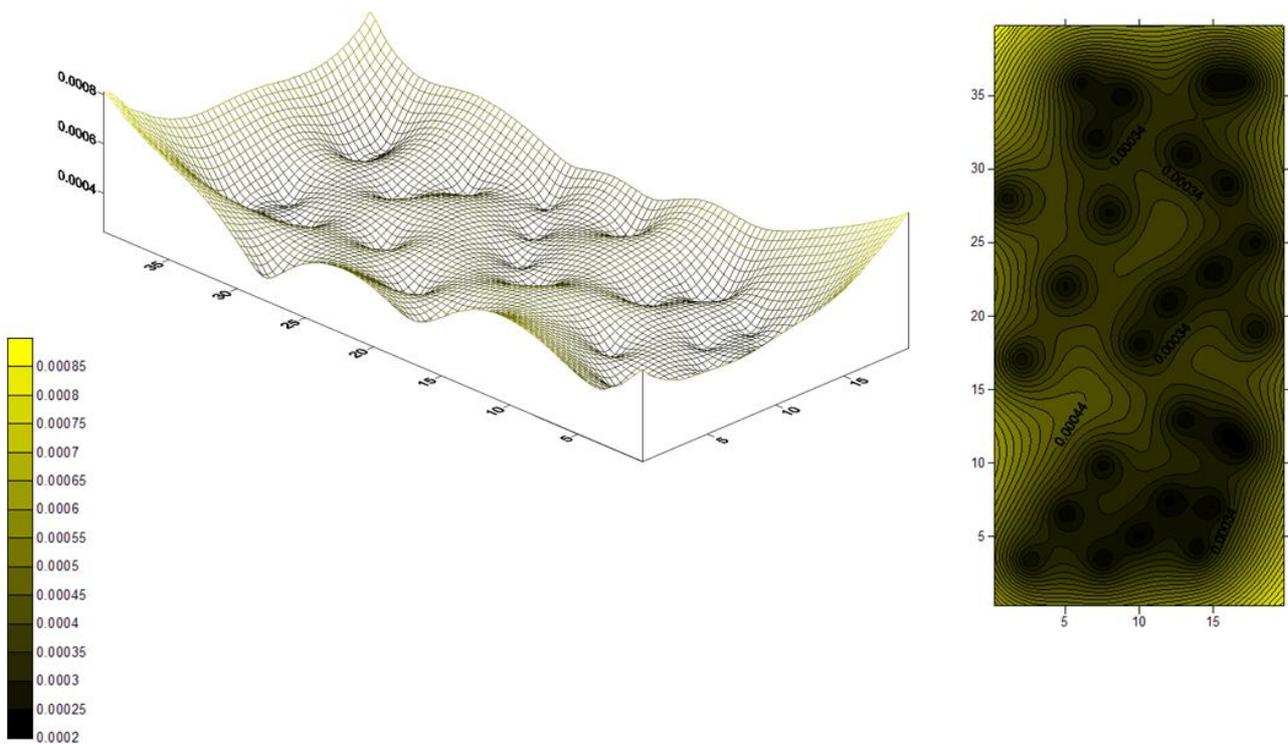


Figura 4.16.2 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello di Tyler & Wheatcraft [1990] e per bassi contenuti di acqua

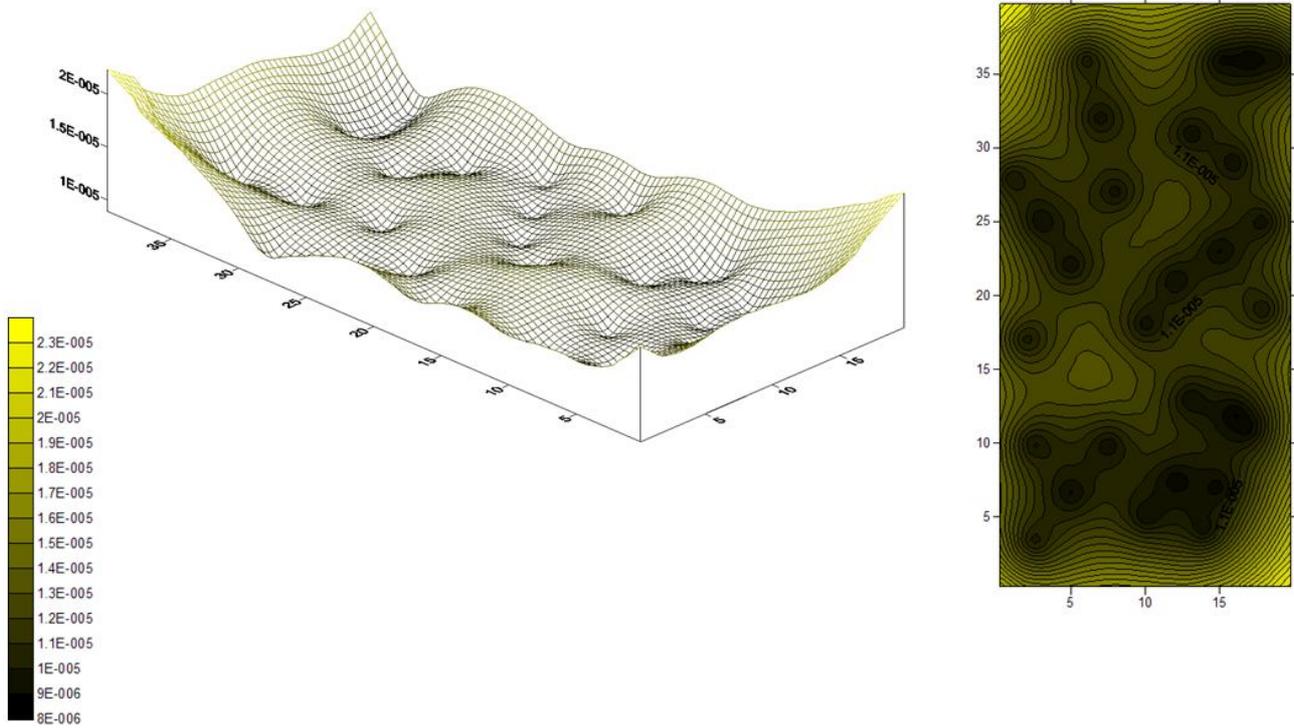


Figura 4.16.3 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello di Rieu & Sposito [1991] e per alti contenuti di acqua

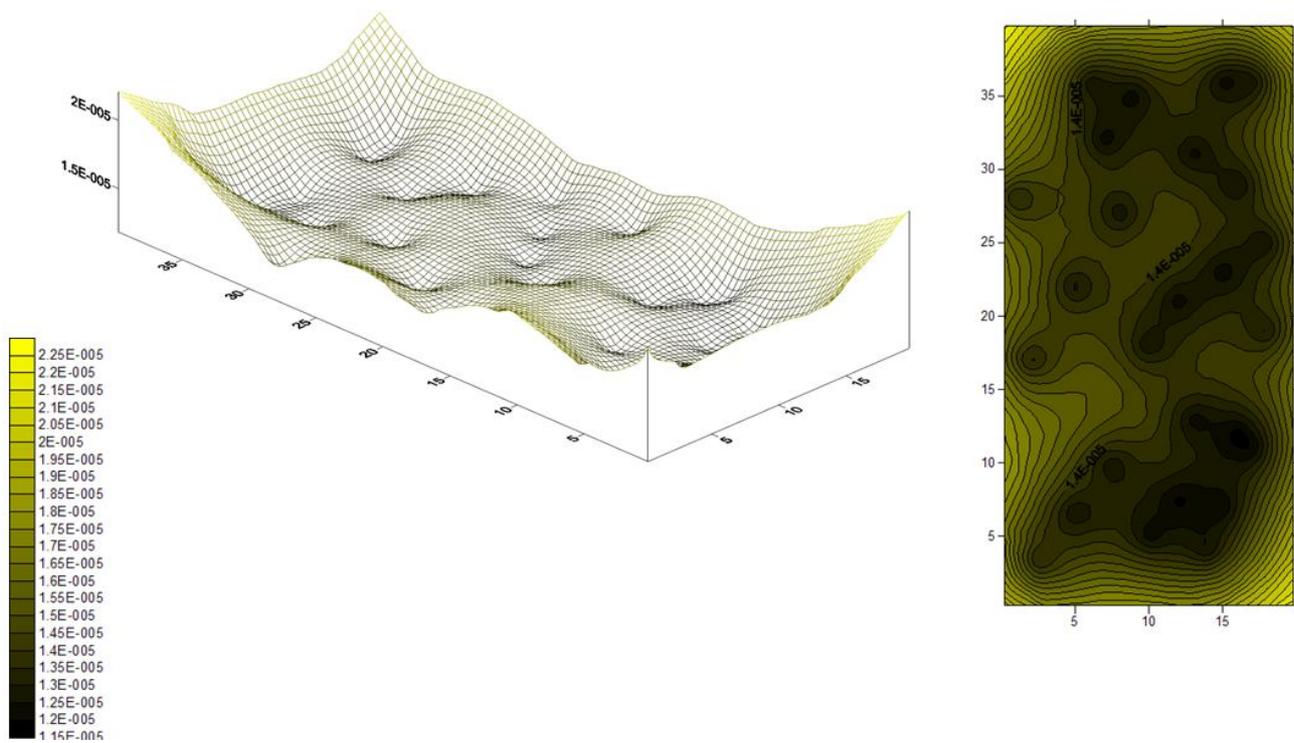


Figura 4.16.4 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello di Rieu & Sposito [1991] e per bassi contenuti di acqua

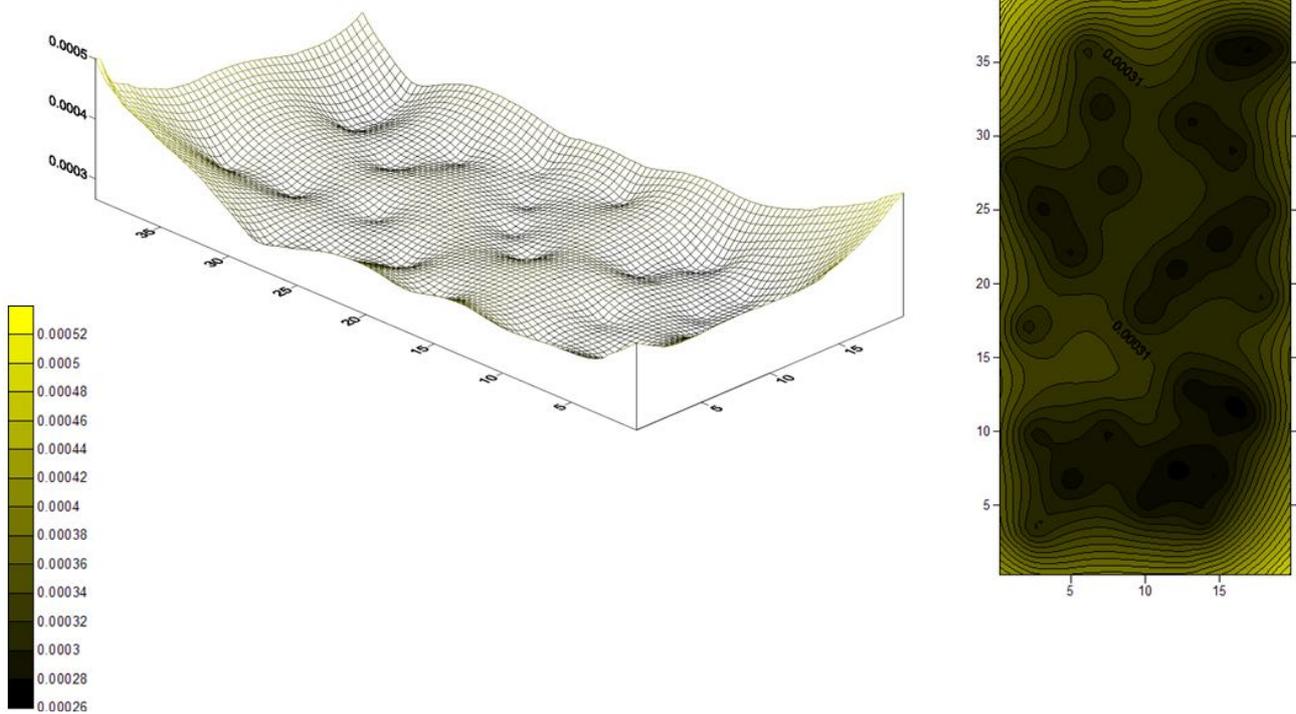


Figura 4.16.5 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per alti contenuti di acqua

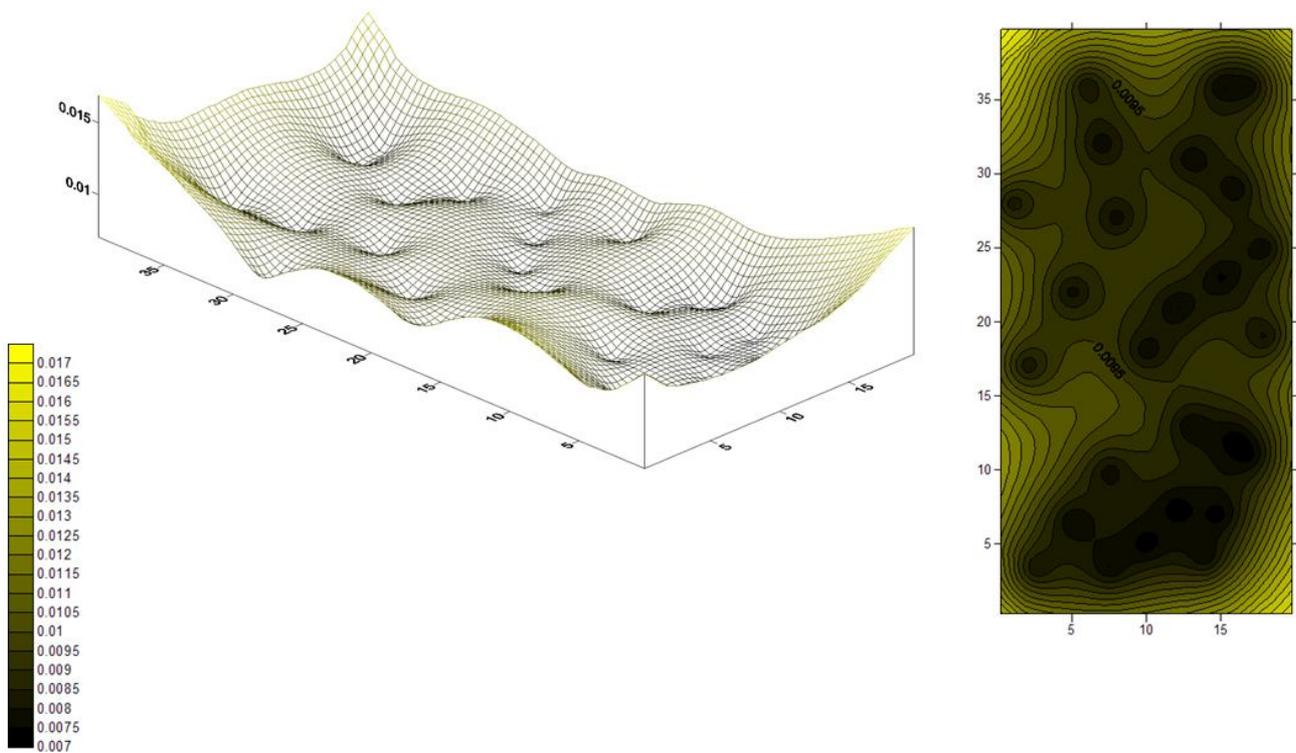


Figura 4.16.6 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.30$ e per bassi contenuti di acqua

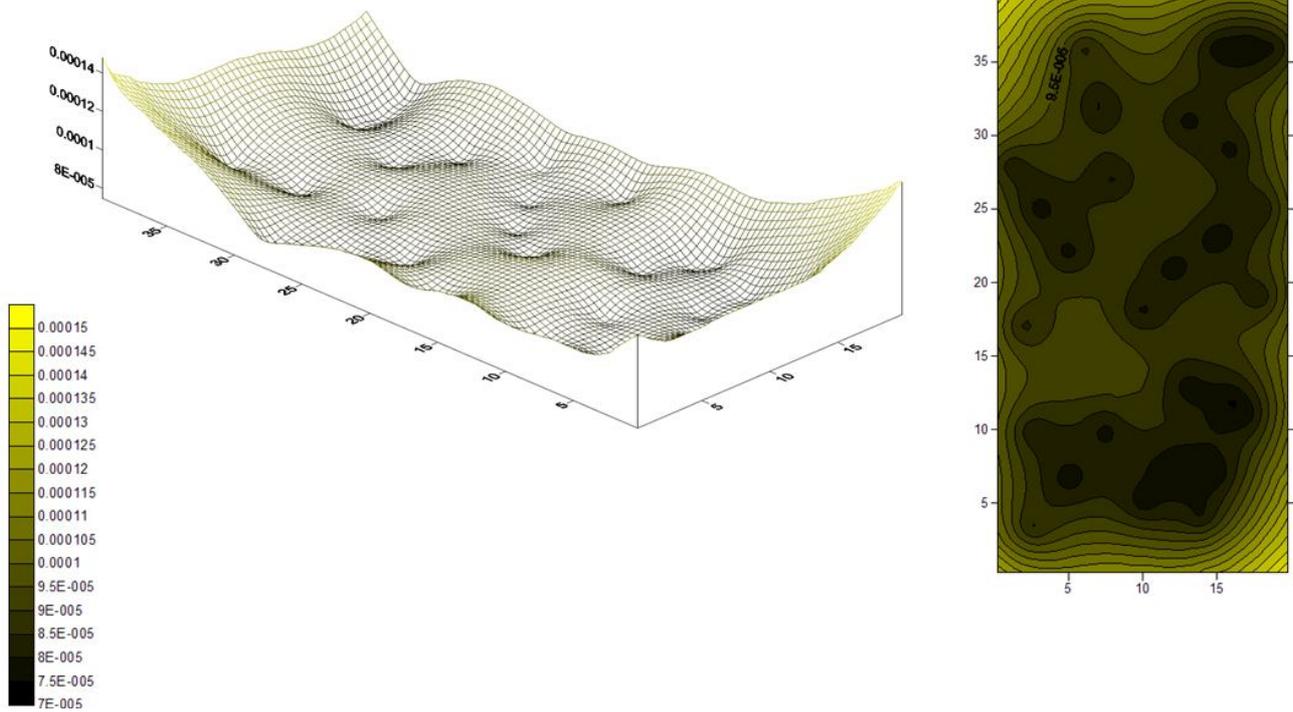


Figura 4.16.7 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per alti contenuti di acqua

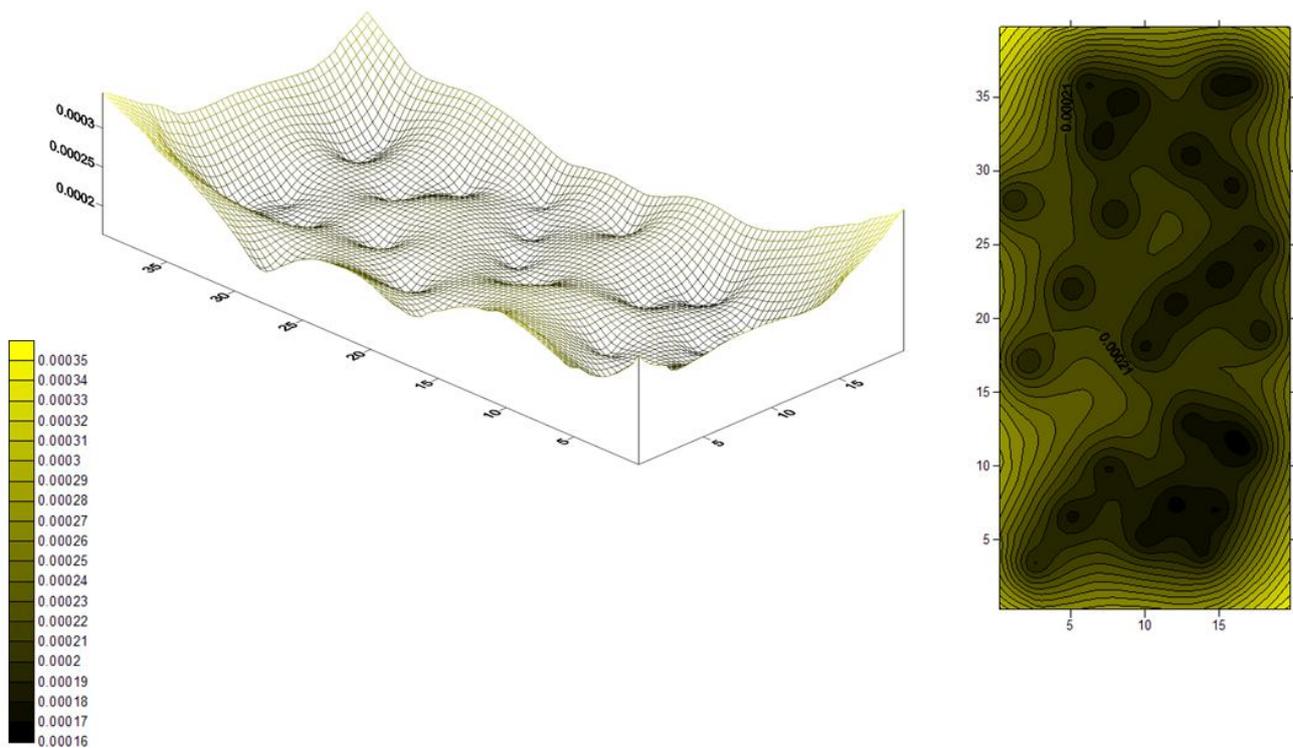


Figura 4.16.8 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.45$ e per bassi contenuti di acqua

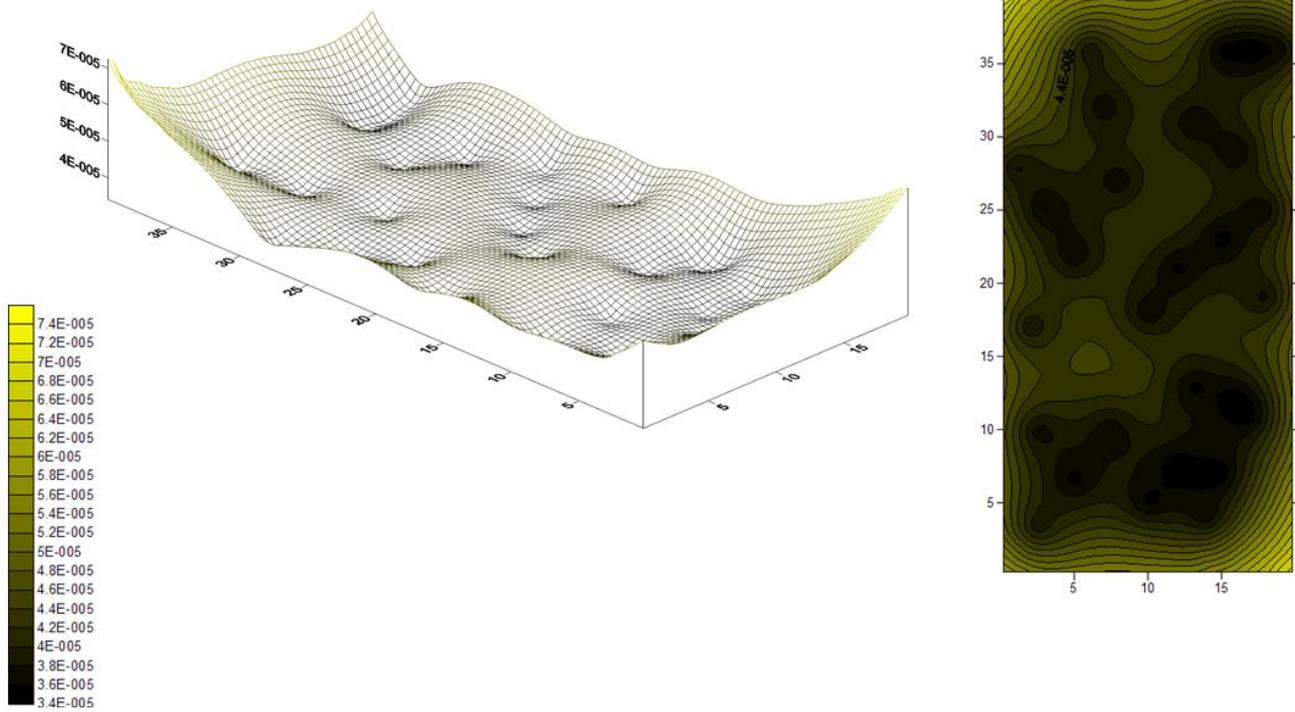


Figura 4.16.9 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per alti contenuti di acqua

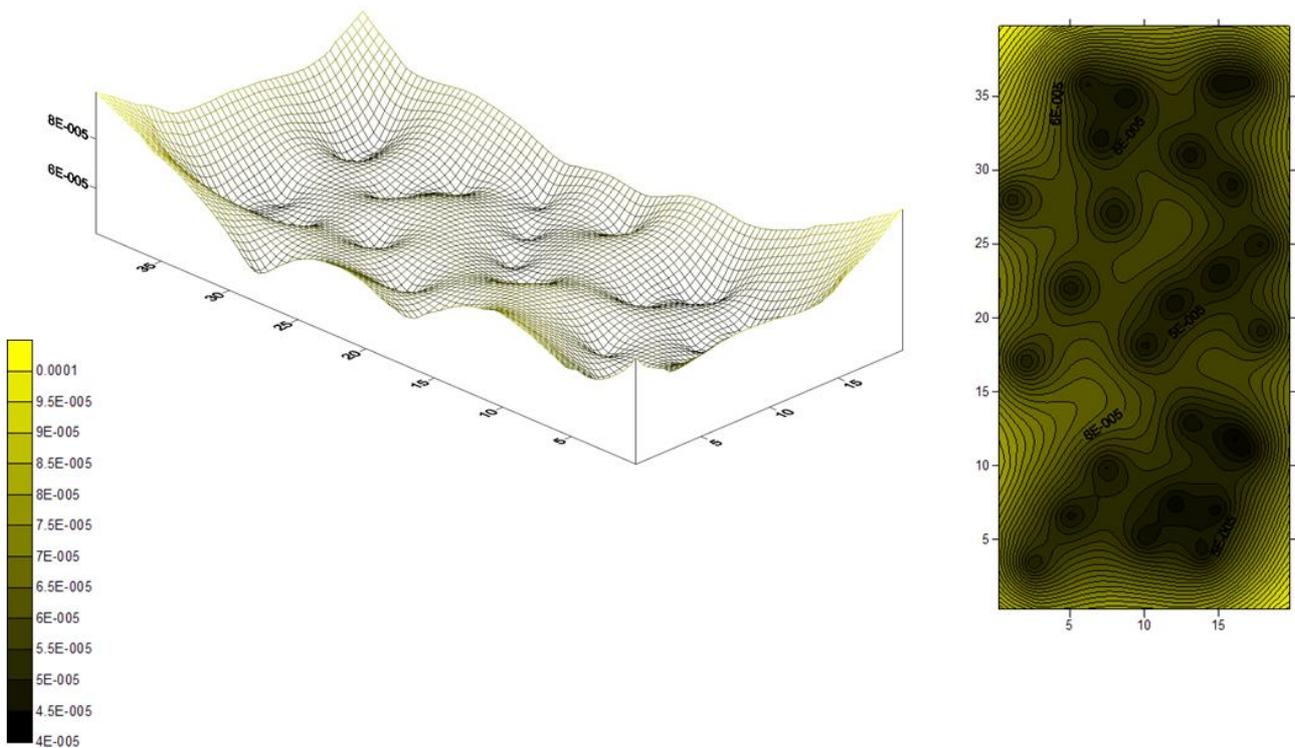


Figura 4.16.10 – Rappresentazione spaziale della varianza del Kriging ordinario per il modello Pore-Solid Fractal [2000] con $A=0.60$ e per bassi contenuti di acqua

Dopo aver effettuato un'interpolazione spaziale tramite l'interpolatore *Kriging* ordinario, è indispensabile avere una visione completa circa l'entità dell'errore e la precisione della previsione fatta (valutazione della varianza estimativa) riferite a ciascuna locazione e per ciascun modello frattale. Le figure 4.17.1 – 4.17.3 illustrano tali risultati per alti contenuti di acqua. In generale si può affermare che per quasi tutte le locazioni (ad eccezione della 28) il modello di *Rieu & Sposito* [1991] è il modello caratterizzato dal valore dell'errore minore mentre gli errori che si commettono maggiormente riguardano nella maggior parte dei casi il modello *Pore-Solid Fractal* [2000] con $A=0.30$ e il modello di *Tyler & Wheatcraft* [1990], precisamente per le locazioni 3, 13, 21, 22, 27, 28. Nella maggior parte delle locazioni, la variabile dimensione frattale è stata sovrastimata, poiché l'errore di stima è contraddistinto da un valore negativo per ciascun modello frattale. A tal proposito bisogna precisare che nella locazione 22, la dimensione frattale è stata sovrastimata per i modelli di *Tyler & Wheatcraft* [1990] e di *Rieu & Sposito* [1991], mentre la stessa viene sottostimata per il modello di *Bird et al.* [2000], qualsiasi sia il parametro A preso in considerazione. Esaminando inoltre la locazione 28, l'unico modello a cui è associato un valore della stima maggiore rispetto alla misura effettuata è il *Pore-Solid Fractal* [2000] caratterizzato da un parametro $A=0.30$. In definitiva, come già accennato in precedenza, si può affermare che per alti contenuti di acqua, il modello di *Rieu & Sposito* [1991] rappresenta il miglior modello che ben interpreta i valori misurati (o reali), proprio perché esso è il modello che dispone dell'errore di stima minore (ad eccezione della locazione 28). Inoltre, il valore della varianza dell'errore, che rappresenta la precisione della previsione effettuata, conferma tale ipotesi in quanto essa assume sempre il valore minimo proprio in riferimento al modello di *Rieu & Sposito* [1991].

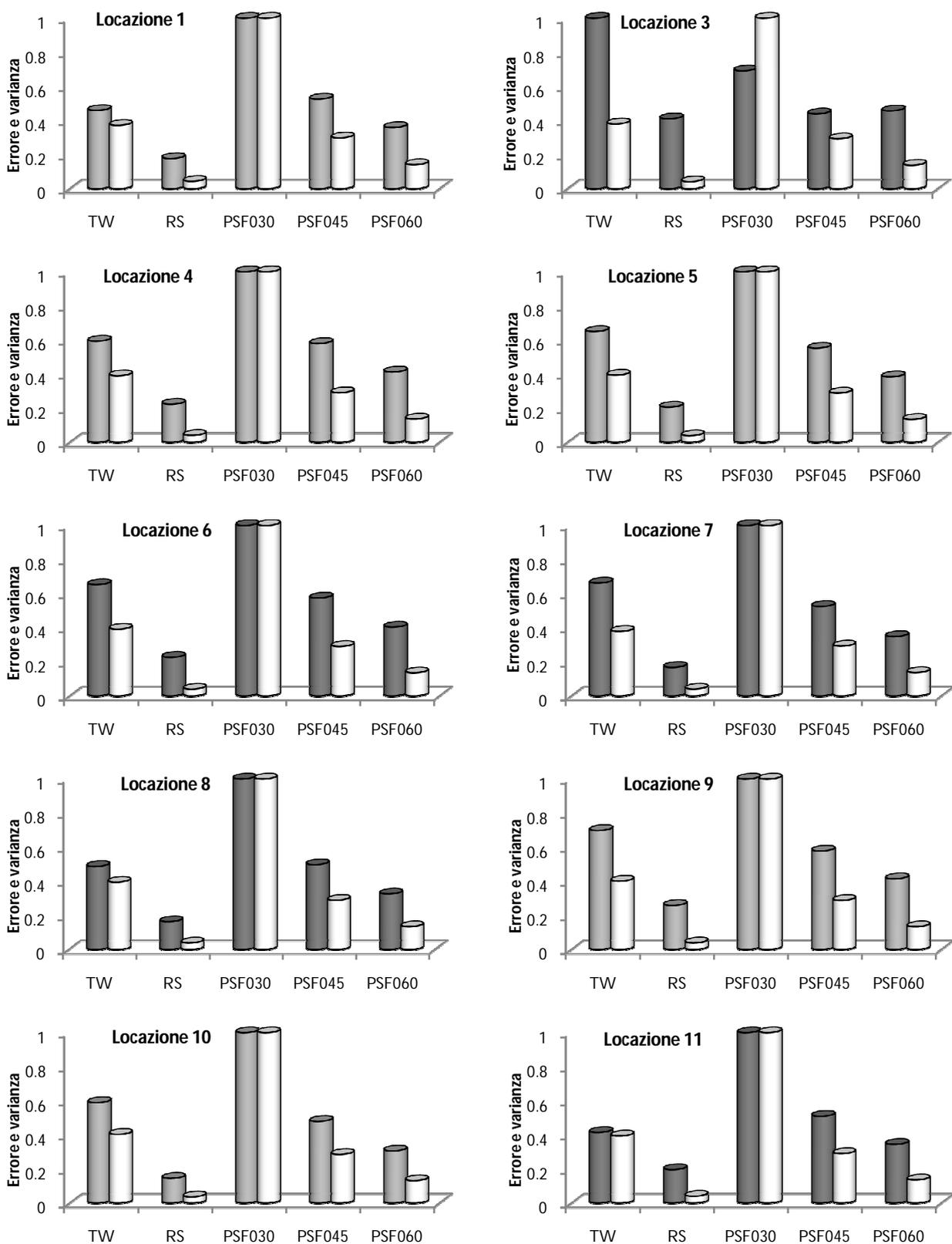


Figura 4.17.1 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e alti contenuti di acqua dalla locazione 1 alla locazione 11. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; diversamente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

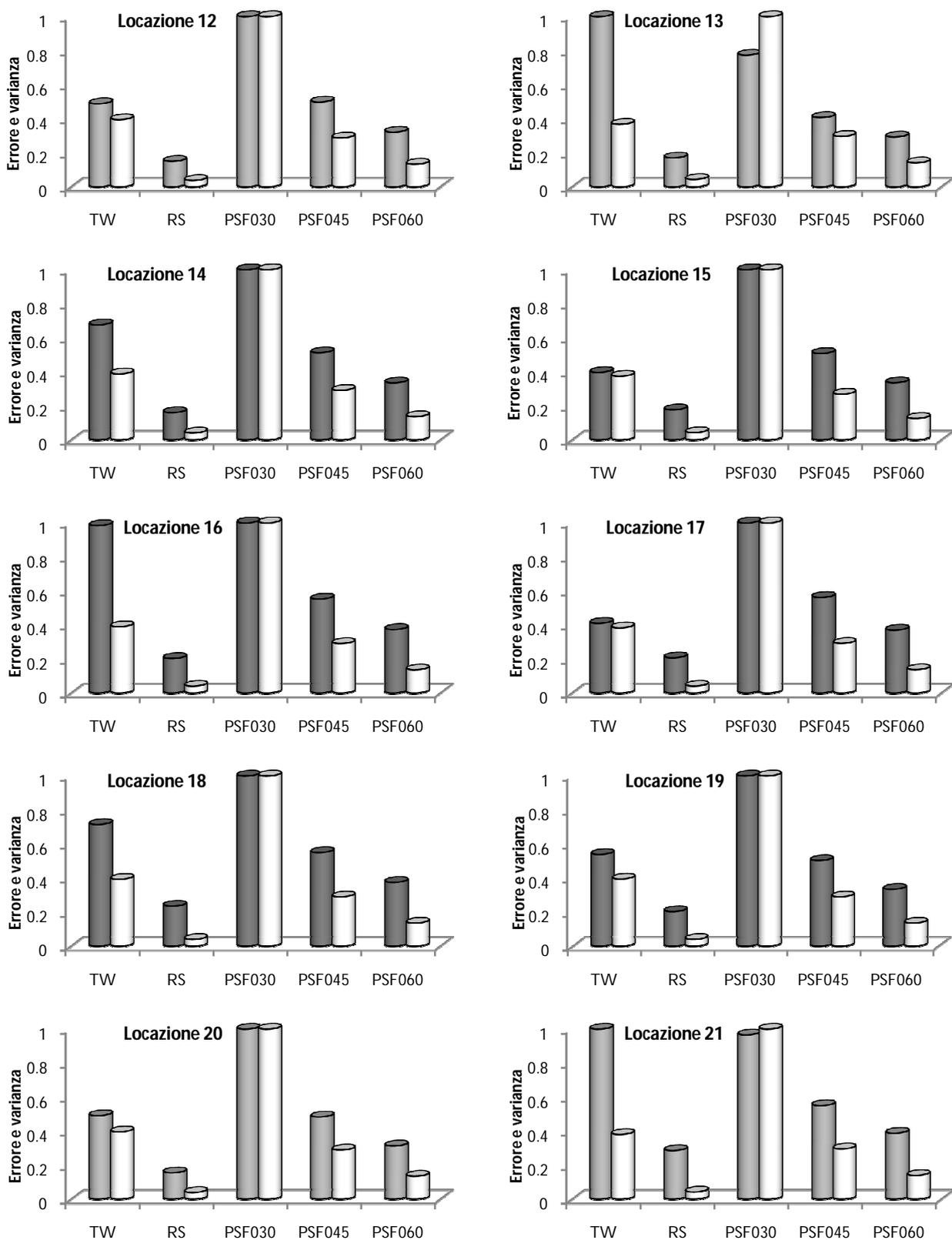


Figura 4.17.2 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e alti contenuti di acqua dalla locazione 12 alla locazione 21. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; diversamente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

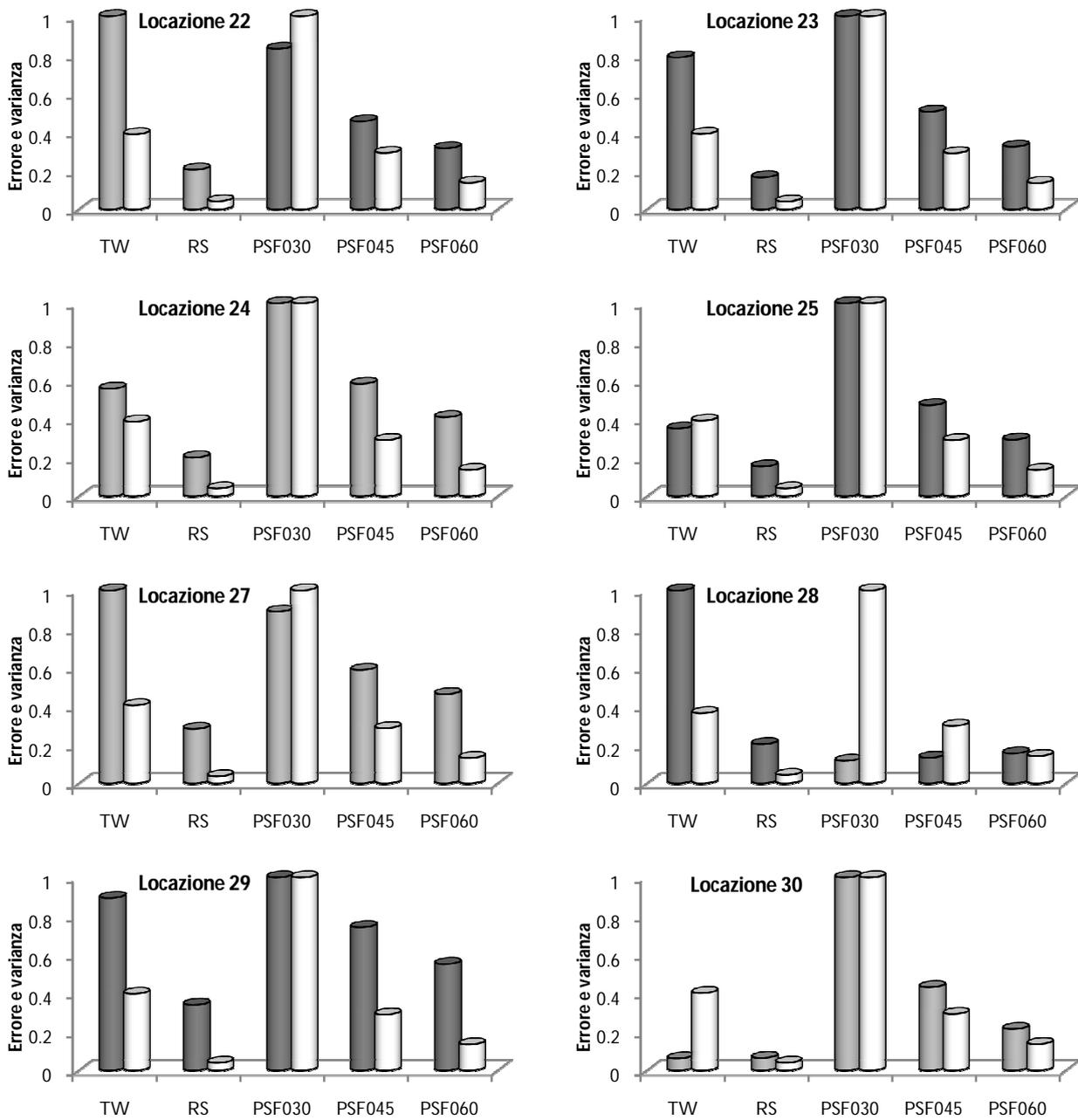


Figura 4.17.3 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e alti contenuti di acqua dalla locazione 22 alla locazione 30. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; differentemente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

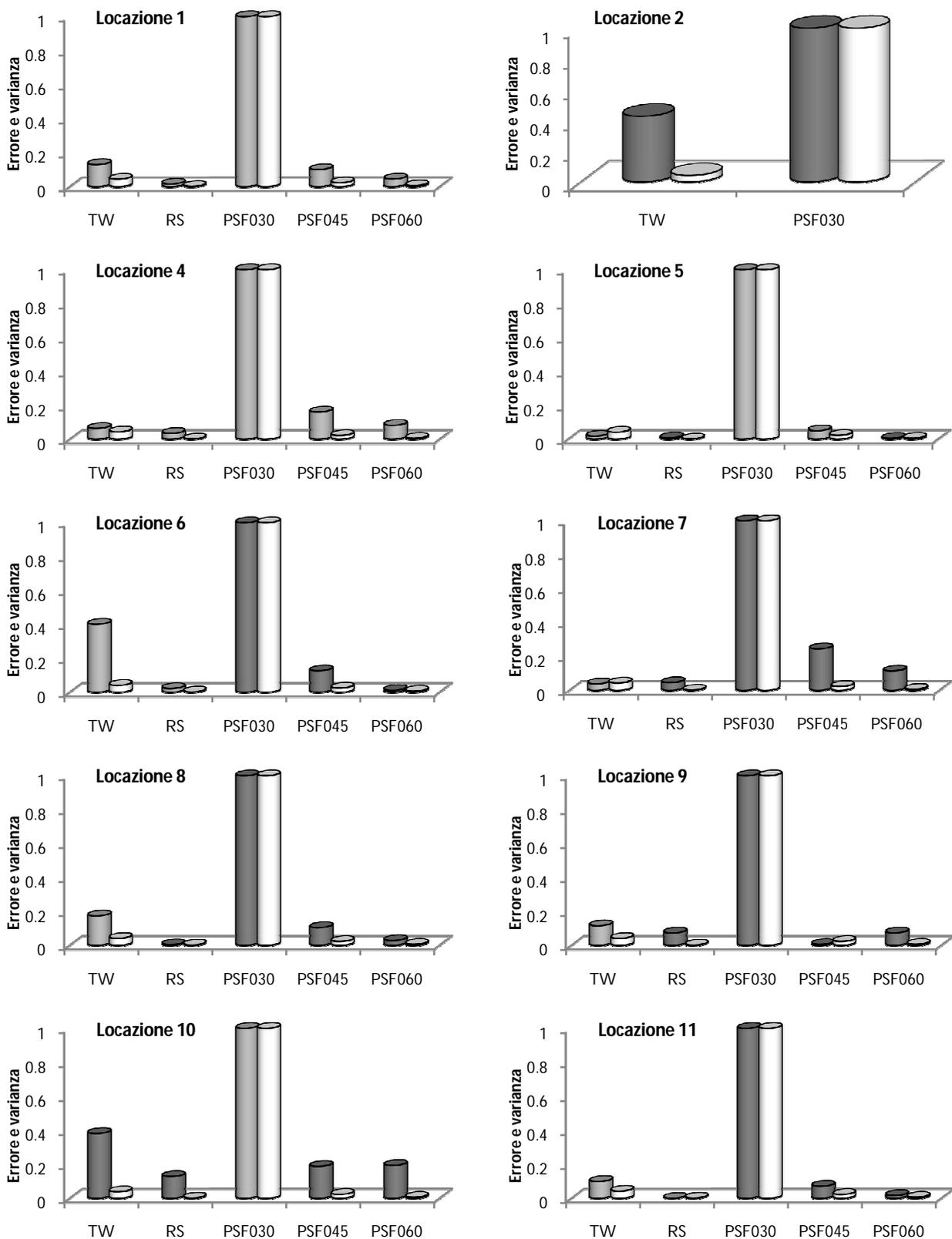


Figura 4.18.1 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e bassi contenuti di acqua dalla locazione 1 alla locazione 11. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; diversamente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

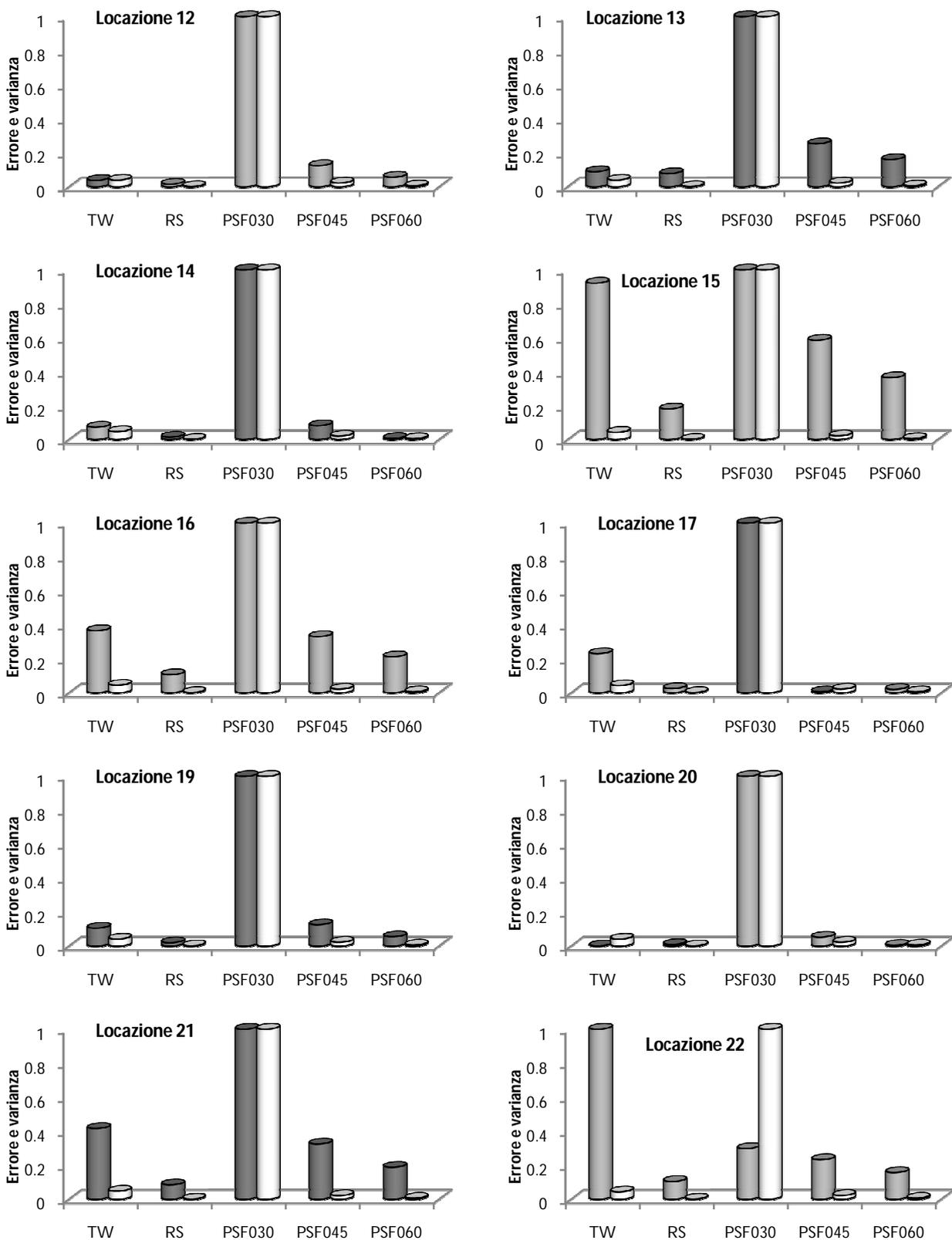


Figura 4.18.2 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e bassi contenuti di acqua dalla locazione 12 alla locazione 22. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; diversamente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

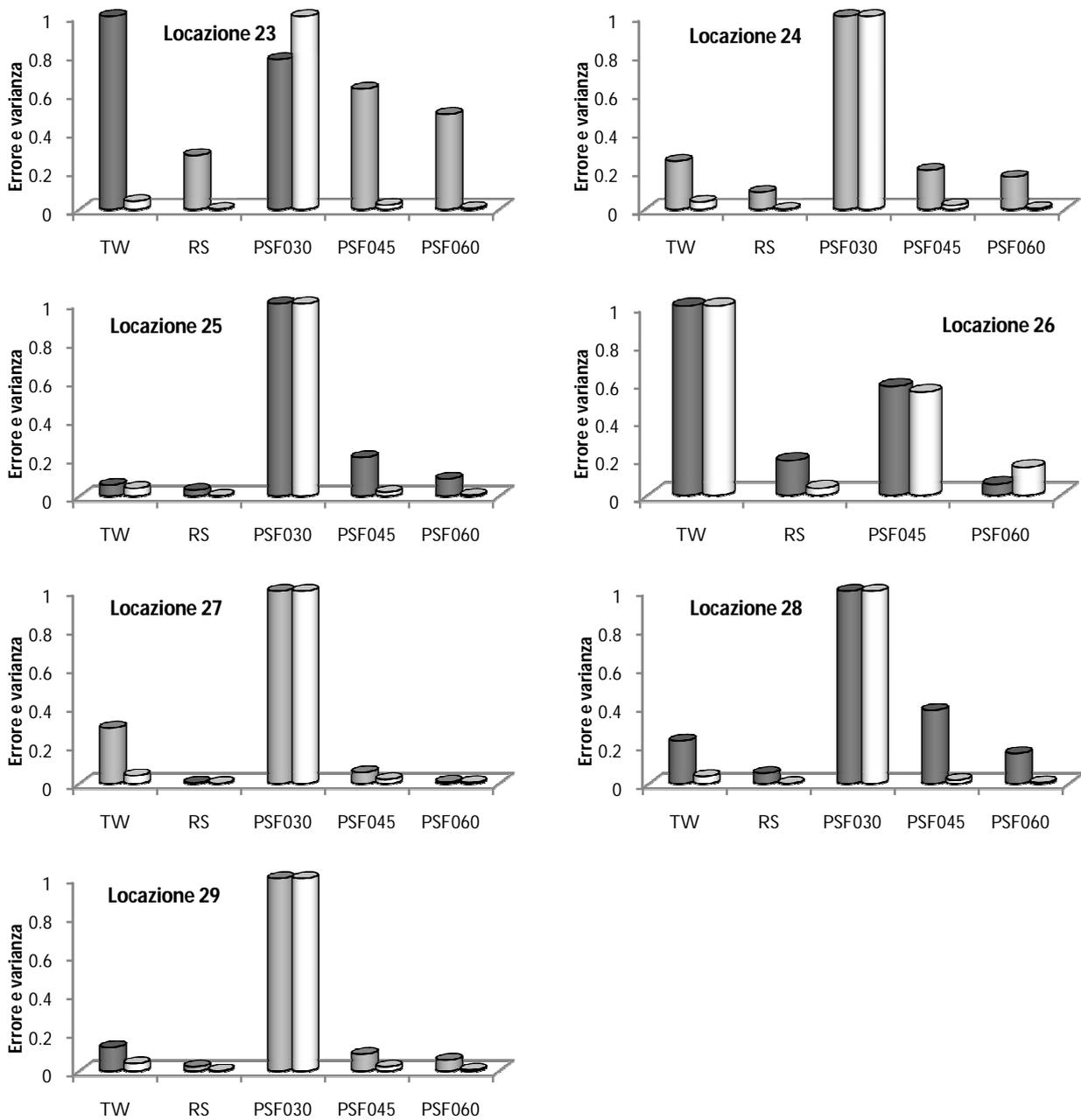


Figura 4.18.3 – Rappresentazione per mezzo di istogrammi dei valori dell'errore e della varianza normalizzati per ciascun modello frattale e bassi contenuti di acqua dalla locazione 23 alla locazione 29. Gli istogrammi di colore grigio chiaro indicano una sovrastima della dimensione frattale; differentemente, gli istogrammi di colore grigio scuro denotano una sottostima della variabile. La varianza è rappresentata dagli istogrammi bianchi

Le figure 4.18.1 – 4.18.3 illustrano i risultati relativi agli errori e ai valori della varianza estimativa, riferiti a ciascuna locazione, per ciascun modello frattale e per bassi contenuti di acqua. Indubbiamente i risultati dei valori della varianza mostrano una condizione molto esplicita e chiara: nel modello di *Rieu & Sposito* [1991] la precisione della previsione fatta risulta la migliore, a differenza del modello *Pore-Solid Fractal* [2000] con $A=0.30$ il quale appare il meno attendibile. Nella locazione 2 si tiene conto dei risultati di soli due modelli frattali, ossia i modelli di *Tyler & Wheatcraft* [1990] e di *Bird et al.*, [2000] con $A=0.30$, poiché per i rimanenti modelli, la dimensione frattale, in tale locazione, è risultata un valore anomalo (*outlier*); per la medesima motivazione, nella locazione 26 non sussistono i valori degli errori e della varianza del modello *Pore-Solid Fractal* [2000], con $A=0.30$. In generale si può affermare che, per la maggior parte delle locazioni, il modello di *Rieu & Sposito* [1991] è il modello caratterizzato dal valore dell'errore minore mentre gli errori che si commettono maggiormente riguardano, nella maggior parte dei casi, il modello *Pore-Solid Fractal* [2000] con $A=0.30$, ad eccezione delle locazioni 22, 23, 26, nelle quali è proprio il modello di *Tyler & Wheatcraft* [1990] a risultare il meno attendibile. In effetti, il numero di punti campionati nei quali il modello di *Rieu & Sposito* [1991] non dispone del minimo errore di stima (2, 5, 6, 7, 9, 14, 17, 20, 26), è sicuramente maggiore rispetto ai dati riferiti ad alti contenuti di acqua. D'altra parte è necessario precisare che in tali locazioni, il valore dell'errore di stima per il modello di *Rieu & Sposito* [1991] è notevolmente basso e, comunque, decisamente vicino all'errore minimo tra tutti i modelli frattali. Tra alti e bassi contenuti d'acqua si possono evidenziare differenze alquanto interessanti. Nel primo caso, infatti, in ciascuna locazione, ed esaminando tutti i modelli frattali, sussiste quasi sempre una simile condizione di sottostima e/o sovrastima dei nuovi valori della variabile dimensione frattale. Differentemente, per bassi contenuti di acqua, questa circostanza non si presenta, poiché esistono situazioni in cui in una specifica locazione e per un determinato modello frattale, la stima della variabile risulta maggiore del valore reale, e, contemporaneamente, per un altro modello la situazione opposta. In definitiva si può comunque affermare che, per bassi contenuti di acqua, il modello di *Rieu & Sposito* [1991] rappresenta il miglior modello che ben interpreta i valori misurati (o reali), in quanto esso è il modello

caratterizzato dall'errore di stima minore nella maggior parte delle locazioni, ipotesi che può essere comunque confermata dai valori estremamente piccoli della varianza estimativa.

Conclusioni

Nel presente lavoro, lo studio di metodologie geostatistiche applicate alla dimensione frattale ha permesso di quantificare il valore di tale variabile su tutta l'area di un sito sperimentale e, quindi, anche in zone in cui non sono state effettuate misure.

La dimensione frattale è stata stimata analizzando lo *scaling* relativo ai grafici in scala doppio-logaritmica (θ^* , h/h_{\min}) [Fallico et al., 2010] e utilizzando le equazioni che definiscono alcuni modelli frattali per la caratterizzazione delle curve di ritenzione idrica [Tyler & Wheatcraft, 1990; Rieu & Sposito, 1991; Bird et al., 2000]; in accordo a diversi studi sperimentali [Toledo et al., 1990; Millàn & Gonzàles-Posada, 2005; Fallico et al., 2010], le curve di ritenzione idrica sono caratterizzate da due tratti di *scaling* riferiti ad alti e bassi contenuti di acqua ed essa rappresenta per ciascuno di tali tratti il complemento del coefficiente angolare della retta che meglio interpola i dati sperimentali, ossia con il massimo coefficiente di correlazione R^2 [Meakin, 1998]. I risultati ottenuti nel presente studio possono essere utilizzati per varie applicazioni e discipline, come l'agronomia, l'idrogeologia, la pedologia, l'idraulica sotterranea, in quanto la curva di ritenzione idrica influenza un vastissimo numero di processi naturali.

Considerando che i modelli frattali delle curve di ritenzione idrica considerati nel presente studio sono stati quelli di Tyler & Wheatcraft [1990], di Rieu & Sposito [1991] e il *Pore-Solid Fractal* [Bird et al., 2000], considerando ancora che per quest'ultimo il valore del parametro A [Wang et al., 2005] è stato assunto pari a 0.30, 0.45, 0.60 [Fallico et al., 2010], l'insieme dei valori della dimensione frattale di cui si dispone è rappresentato complessivamente da dieci campioni; cinque sono relativi all'intervallo di *cut-off* superiore (alti contenuti di acqua) e gli altri cinque a quello inferiore (bassi contenuti di acqua).

Per ciascun set di dati è stata applicata un'analisi geostatistica con lo scopo di valutare il comportamento della variabile dimensione frattale in un sito sperimentale del bacino del

torrente Turbolo, sottobacino del fiume Crati, nel quale sono stati prelevati trenta campioni di terreno indisturbato.

Attraverso l'analisi esplorativa dei dati [Tukey, 1977], che rappresenta la fase preliminare di uno studio geostatistico, è stato dimostrato che ciascun campione può essere descritto con una legge di tipo normale o lognormale, condizione necessaria in geostatistica per la verifica delle ipotesi di stazionarietà; inoltre, in ciascun set di dati, sono stati individuati diversi *outliers*, valori che sono stati opportunamente eliminati allo scopo di non falsificare l'interpolazione spaziale dei dati misurati che si hanno a disposizione.

L'analisi spaziale della dimensione frattale è stata effettuata mediante il *Kriging* [Krige, 1951], tecnica di interpolazione spaziale che consiste nella risoluzione di un sistema di equazioni lineari e che permette di ottenere stime ottimali e non distorte (B.L.U.E.) della variabile regionalizzata in punti non campionati; esso utilizza la serie dei dati sperimentali a disposizione e considera il modello di correlazione spaziale esaminato (variogramma) nell'analisi strutturale. In particolare, nel seguente lavoro, tenendo conto della condizione di stazionarietà del fenomeno, è stato utilizzato il *Kriging* ordinario [Deutsch & Journel, 1992; Isaaks & Srivastava, 1989] che rappresenta sempre una combinazione lineare dei valori noti della variabile in prossimità dell'entità da stimare, sotto specifiche ipotesi intrinseche.

Esaminando le mappature della variabile dimensione frattale, è stato verificato che, per alti contenuti di acqua, sussiste una forte uguaglianza tra i modelli frattali e, cioè, essa presenta nel complesso valori simili nelle stesse sottoaree del sito sperimentale oggetto di studio. Differentemente, per bassi contenuti di acqua, le mappe presentano una grande disuguaglianza poiché la variabile mostra caratteristiche differenti tra un modello frattale e l'altro. I risultati relativi ai bassi contenuti di acqua risultano meno attendibili perché, molto probabilmente, il numero di punti sperimentali relativi al secondo tratto è molto limitato.

Per il primo e il secondo *range* dello *scaling*, il modello di Rieu & Sposito [1991] mostra il miglior adattamento ai dati misurati perché, spesso, esso presenta il valore dell'errore minore e, la relativa varianza del *Kriging* ordinario, assume i valori più piccoli in molte località.

Paragonato allo studio precedente di *Fallico et al.* [2010], questo lavoro mostra un'analisi geostatistica più dettagliata poiché sono stati esaminati diversi altri aspetti come, ad esempio, la valutazione di possibili anisotropie riguardanti il fenomeno indagato.

Tuttavia, i risultati del presente studio confermano le conclusioni del lavoro di *Fallico et al.* [2010], poiché in quest'ultimo, sia per alti che per bassi contenuti di acqua, il modello di *Rieu & Sposito* [1991] è risultato il miglior modello frattale che ben interpreta i dati reali.

Per quanto riguarda gli obiettivi futuri, di particolare interesse potrebbe essere il nuovo approccio geostatistico multivariato [*Wackernagel*, 1995] che incorpora alla variabile primaria (dimensione frattale), la variabile secondaria, ossia la conducibilità idraulica. Il valore della conducibilità idraulica, infatti, è disponibile nei punti campionati in cui è stato stimato il valore della dimensione frattale. Per tale elaborazione si pensa di utilizzare il paradigma *Kriging* con *external drift* [*Isaak & Srivastava*, 1989; *Wackernagel*, 1995; *Deutsch & Journel*, 1998] che presuppone anche la conoscenza della variabile ausiliaria in tutti i punti in cui deve essere stimata la variabile primaria; ovviamente quest'ultima ipotesi può essere convalidata solo andando a ricostruire l'andamento della conducibilità idraulica nel dominio oggetto di studio attraverso l'interpolatore *Kriging* ordinario. Il lavoro effettuato nel campo della geostatistica multivariata presupporrebbe come obiettivo primario, quindi, la verifica dell'attendibilità del modello di *Rieu & Sposito* [1991], come dimostrato nello studio di *Fallico et al.* [2010] e nel presente lavoro, in quanto la presenza e la conoscenza di una seconda grandezza in fase analitica comporterebbe e implicherebbe, ovviamente, una precisione e affidabilità maggiore del modello frattale.

Riferimenti bibliografici

- Anderson, T. W., & Darling, D. A. (1952). Asimptotic theory of certain goodness of fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23 (2), 193-212.
- Armstrong, M. (1998). *Basic Linear Geostatistics*. Berlin, GERMANY: Springer Verlag.
- Avnir, D., Biham, O., Lidar, D., & Malcai, O. (1998). Is the geometry of nature fractal? *Science*, 279, 38-40.
- Barnett, V., & Lewis, T. (1978). *Outliers in Statistical Data*. New York, USA: John Wiley & Sons.
- Bear, J. (1972). *Dynamics of Fluids in Porous Media*. New York, USA: Elsevier.
- Beretta, G. P. (2004). *Il trattamento e interpretazione dei dati ambientali; qualità dei suoli e delle acque sotterranee nella bonifica dei siti contaminati*. Bologna, ITALIA: Pitagora Editrice.
- Bhattacharyya, G. K., & Johnson, R. A. (1977). *Statistics-Principles and Methods*. New York, USA: John Wiley & Sons.
- Bird, N., Bartoli, F., & Dexter, A. R. (1996). Water retention models for fractal soil structure. *European Journal of Soil Science*, 47, 1-6.
- Bird, N., Perrier, E., & Rieu, M. (2000). The water retention function for a model of soil structure with pore and solid fractal distributions. *European Journal of Soil Science*, 51 (1), 55-63.
- Bolt, G. H. (1976). Soil physics terminology. *Bulletin International. Society Soil Science*, 49, 26-36.
- Bouma, J. (1989). Using soil survey data for quantitative land evaluation. *Advances in Soil Science*, 9, 177-213.
- Brooks, R. H., & Corey, A. T. (1964). *Hydraulic Properties of Porous Media*. Colorado State University: Fort Collins, CO.
- Castino, M., & Roletto, E. (1991). *Statistica applicata*. Sassari, ITALIA: Piccin.

- Cavazza, L. (1981). *Fisica del Terreno Agrario*. Torino, ITALIA: Utet.
- Chandra, A. M., & Ghosh, S. K. (2006). *Remote sensing and geographical information system*. Oxford, England: Alpha Science International Ltd.
- Chang, K. T. (2002). *Introduction to Geographic Information Systems*. New York, USA: Mc Graw-Hill.
- Chapra, S. C., & Canale, R. P. (1985). *Numerical methods for engineers with personal computer applications*. New York, USA: McGraw-Hill.
- Chow, T. L. (2000). *Mathematical Methods for Physicists*. Melbourne, AUSTRALIA: Cambridge University Press.
- Ciotoli, G., & Finoia, M. (2005). *Dalla Statistica alla Geostatistica*. Roma, ITALIA: Aracne editrice srl.
- Ciotoli, G., Lombardi, S., & Annunziatelli, A. (2007). Geostatistical analysis of soil gas data in a high seismic intermontane basin: Fucino Plain, central Italy. *Journal of Geophysical Research*, 112, 1-23.
- Clark, I. (1984). *Practical Geostatistics*. London, ENGLAND: Elsevier Applied Science Publishers.
- Crawford, J. W., Sleeman, B. D., & Young, I. M. (1993). On the relation between number-size distribution and the fractal dimension of aggregates. *Journal of Soil Science*, 44, 555-565.
- Cressie, N. (1991). *Statistics for Spatial Data*. New York, USA: John Wiley & Sons.
- Cressie, N. (1985). When are relative variograms useful in Geostatistics. *Mathematical Geology*, 17, 693-702.
- Danielson, R. E., & Sutherland, P. L. (1986). Porosity. In A. Klute, *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods* (p. 443-461). Madison: American Society of Agronomy.
- Davis, B. M. (1987). Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, 29, 241-248.
- Davis, C. J. (2002). *Statistics and Data Analysis in Geology*. New York, USA: John Wiley & Sons.
- de Gennes, P. G. (1985). Partial filling of a fractal structure by a wetting fluid. In: *Physics of Disordered Materials* (eds D. Adler, E. Frietsche and S. R. Ovshinsky), 227-241, New York, USA: Plenum Press.

- de Marsily, G. (1986). *Quantitative Hydrogeology. Groundwater hydrology for engineers*. San Diego, USA: Academic Press.
- Delhomme, J. P. (1978). Kriging in hydrosociences. *Advances in Water Resources*, 1 (5), 251-256.
- Deutsch, C. V., & Journel, A. G. (1998). *GSLIB Geostatistical Software Library and User's Guide*. New York, USA: Oxford University Press.
- Dixon, W. J. (1953). Processing Data for Outliers. *Biometrics*, 9, 74-89.
- Eagleson, P. S. (1970). *Dynamics hydrology*. New York, USA: McGraw-Hill. Book Company.
- Edgar, G. A. (1998). *Integral, probability, and fractal measures*. New York, USA: Springer Verlag.
- Edgar, G. A. (1990). *Measure, topology, and fractal geometry*. New York, USA: Springer Verlag.
- Espino, A., Mallants, D., Vanclooster, M., & Feyen, J. (1995). Cautionary notes on the use of pedo-transfer functions for estimating soil hydraulic properties. *Agricultural Water Management*, 29, 235-253.
- Falconer, K. (1990). *Fractal geometry: mathematical foundations and applications*. Chichester, ENGLAND: John Wiley & Sons
- Falconer, K. (1997). *Techniques in fractal geometry*. Chichester, ENGLAND: John Wiley & Sons
- Fallico, C., Tarquis, A. M., De Bartolo, S., & Veltri, M. (2010). Scaling analysis of water retention curves for unsaturated sandy loam soils by using fractal geometry. *European Journal of Soil Science*, 61, 425-436.
- Fallico, C., Troisi, S., Amato, G., & Mazzuca, L. (2004). Il campo prove del bacino del torrente Turbolo per il monitoraggio del sottosuolo non saturo. *XXV Corso di Aggiornamento in Tecniche per la Difesa dall'inquinamento*. Guardia Piemontese (CS), ITALIA: Bios.
- Feder, J. (1988). *Fractals*. New York, USA: Plenum.
- Fisher, P. F. (2005). *Developments in Spatial Data Handling*. Berlin, Germany: Springer Verlag.
- Frank, J., & Massey, J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46 (253), 68-78.
- Frigge, M., Hoaglin, D. C., & Iglewicz, B. (1989). Some implementations of the box-plot. *The American Statistician*, 43 (1), 50-54.

- Gandin, L. S. (1963). *Objective Analysis of Meteorological Fields*. Leningrad: Gidrometeorologicheskoe Izdatel'stvo (GIMIZ).
- Gooavaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. New York, USA: Oxford University Press.
- Hack, D. R. (2005). *Issues and challenges in the applications of Geostatistics and Spatial-Data Analysis to the characterization of sand and gravel resources*. U.S. Geological survey.
- Hengl, T., Heuvelink, G., & Stain, A. (2004). A generic framework for spatial prediction of soil variables on regression-kriging. *Geoderma*, 120, 75-93.
- Hillel, D. (1982). *Introduction to Soil Physics*. Orlando, USA: Academic Press.
- Hofierka, J., Parajka, J., Mitasova, M., & Mitas, L. (2002). Multivariate interpolation of precipitation using regularized spline with tension. *Transactions in GIS*, 6, 135-150.
- Hohn, M. E. (1988). *Geostatistics and Petroleum Geology*. New York, USA: Van Nostrand Reinhold.
- Houlding. (2000). *Practical Geostatistics*. Berlin, GERMANY: Springer Verlag.
- Huang, G. H., Zhang, R. D., & Huang, Q. Z. (2006). Modeling soil water retention curve with a fractal method. *Pedosphere*, 16, 137-146.
- Hunt, F., & Sullivan, F. (1986). *Efficient algorithms for computing fractal dimensions*. New York, USA: Springer Verlag.
- Hutchinson, M. F. (1995). Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems*, 9, 385-403.
- Isaaks, E. H., & Srivastava, M. (1989). *An Introduction to Applied Geostatistics*. New York, USA: Oxford University Press.
- Italiana, U. M. (2001). *Bollettino della Unione matematica italiana: Matematica nella società e nella cultura*. Bologna, ITALIA.
- Jones, T. A., Hamilton, D. E., & Johnson, C. R. (1986). *Countouring geologic surfaces with the computer*. New York, USA: Van Nostrand Reinhold.
- Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics*. London, ENGLAND: Academic Press.

- Katz, A. J., & Thompson, A. H. (1985). Fractal sandstone pores: implications for conductivity and pore formation. *Physical Review Letters*, 54, 1325-1327.
- Kitanidis, P. K. (1997). *Introduction to Geostatistics*. Cambridge University Press.
- Kolmogorov, A. N. (1933). *Foundations of the Theory of Probability*. New York, USA: Chelsea Publishing Company.
- Kolmogorov, A. N. (1941). The local structure of turbulence in a incompressible viscous fluid at very large Reynolds numbers. *Doklady Akademii Nauk SSSR*, 30, 301-305.
- Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Chemical Metallurgical and Mining Society of South Africa Journal*, 52, 119-139.
- Kutilek, M., & Nielsen, D. R. (1994). *Soil Hydrology*. Berlin, GERMANY: Springer Verlag.
- Lade, P. V. (2001). Engineering Properties of Soils and Typical Correlations. In *Geotechnical and Geoenvironmental Engineering Handbook* (p. 54). Norwell, Massachusetts: R. Kerry Rowe.
- Lam, N. S. (1983). Spatial interpolation methods: a review. *The American Cartographer*, 10, 129-149.
- Larson, H. J. (1969). *Introduction to Probability Theory and Statistical Inference*. New York, USA: John Wiley & Sons.
- Lloyd, C. D. (2010). *Local Models of Spatial Analysis*. Boca Raton, USA: CRC Press.
- Malcai, O., Lidar, D. A., Biham, O., & D., A. (1997). Scaling range and cutoffs in empirical fractals. *Physical Review*, 56, 2817 - 2828.
- Mandelbrot, B. B. (1989). Fractal geometry: what is, and what does it? In T. D. Fleischmann M., *Fractals in the natural sciences* (p. 3-16). New Jersey: Princeton University Press, Princeton.
- Mandelbrot, B. B. (1987). *Gli Oggetti Frattali: forma caso e dimensione*. Torino, ITALIA: Einaudi.
- Mandelbrot, B. B. (1986). Self-affine fractal sets. In L. Tosatti, & E. Pietronero, *Fractals in physics* (p. 3-28). Amsterdam, Nederland.
- Mandelbrot, B. B. (1985). Self-affine fractals and the fractal dimension. *Physica Scripta*, 32, 257-260.

- Mandelbrot, B. B. (1982). *The fractal geometry of nature*. New York, USA: W H Freeman & Co.
- Marshall, T. J., Holmes, J. W., & Rose, C. W. (1999). *Soil Physics*. Cambridge, ENGLAND: Cambridge University Press.
- Matheron, G. (1965). *Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature*. Paris, FRANCE: Masson.
- Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58, 1246-1266.
- Matheron, G. (1971). The theory of regionalized variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau*, Fascicolo 5.
- Meakin, P. (1998). *Fractals, Scaling and Growth Far From Equilibrium*. Cambridge, ENGLAND: Cambridge University Press.
- Millán, H., & Gonzàles-Posada, M. (2005). Modelling soil water retention scaling. Comparison of a classical fractal model with a piecewise approach. *Geoderma*, 125, 25-38.
- Miller, R. G. (1974a). An unbalanced jackknife. *The Annals of Statistics*, 2, 880-891.
- Miller, R. G. (1974b). The jackknife - a review. *Biometrika*, 61, 1-15.
- Minasny, B., McBratney, A. B., & Bristow, K. L. (1999). Comparison of different approaches to the development of pedotransfer functions for water retention curves. *Geoderma*, 93, 225-253.
- Minasny, B., McBratney, A. B., & Salvador-Blanes, S. (2008). Quantitative models for pedogenesis – a review. *Geoderma*, 144, 140-157.
- Mitasova, H., & Mitas, L. (1993). Interpolation by regularized spline with tension. I. Theory and implementation. *Mathematical Geology*, 25, 641-655.
- Monestiez, P., Courault, D., Allard, D., & Ruget, F. (2001). Spatial interpolation of air temperature using environmental context: Application to a crop model. *Environmental and Ecological Statistics*, 8, 297-309.
- Musgrave, K. F. (1994). A brief introduction to fractals. In E. D. S., *Texturing and modeling. A procedural approach* (p. 249-265). Massachusetts, USA: Academic Press.
- Myers, D. E. (1994). Spatial interpolation: an overview. *Geoderma*, 62, 17-28.
- Nikulín, M. S. (1973). Chi-Squared test for normality. *International Vilnius Conference on Probability Theory and Mathematical Statistics*, 2, 119-122.

- Oliver, M., Webster, R., & Gerrard, J. (1989). Geostatistics in Physical Geography. Part I: Theory. *Transactions of the Institute of British Geographers, New Series*, 14 (3), 259-269.
- Ortiz, J., & Deutsch, C. (2002). Calculation of uncertainty in the variogram. *Mathematical Geology*, 34, 169-183.
- Pachepsky, Y. A., & Rawls, W. L. (1999). Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Science Society of America Journal*, 63, 1748-1756.
- Pannatier, Y. (1996). *Variowin: Software for Spatial Data Analysis in 2D*. New York, USA: Springer Verlag.
- Parker, T. S., & Chua, L. O. (1989). *Practical numerical algorithms for chaotic systems*. New York, USA: Springer Verlag.
- Pascolini, A. (1992). *Insiemi frattali e sistemi dinamici non lineari*. Padova, ITALIA: CLEUP.
- Peitgen, H., Jurgens, H., & Saupe, D. (2004). *Chaos and Fractals: New Frontiers of Science*. New York, USA: Springer Verlag.
- Perfect, E. (1999). Estimating soil mass fractal dimension from water retention curves. *Geoderma*, 88, 221-231.
- Perfect, E., & Kay, B. D. (1991). Fractal theory applied to soil aggregation. *Soil Science Society of American Journal*, 55, 1552-1558.
- Perfect, E., Kenst, A. B., Diaz-Zorita, M., & Grove, J. H. (2004). Fractal analysis of soil water desorption data collected on disturbed samples with water activity meters. *Soil Science Society of America Journal*, 68, 1177-1184.
- Perrier, E., Bird, N., & Rieu, M. (1999). Generalizing the fractal model of soil structure: the PSF approach. *Geoderma*, 88, 137-164.
- Perrier, E., Rieu, M., Sposito, G., & Marsily, G. (1996). Models of the water retention curve for soils with a fractal pore size distribution. *Water Resources Research*, 32 (10), 3025-3031.
- Pontrjagin, L., & Schnirelman, L. (1932). Sur une propriété métrique de la dimension. *Annales of Mathematics*, 33, 156-162.
- Posa, D. (1988). Interpolation by splines and kriging. *Journal of Statistical Research*, 22, 61-71.
- Posa, D. (1995). *Introduzione alla geostatistica*. Lecce, ITALIA: Adriatica Editrice Salentina.

- Posa, D. (1991). Limiting Stochastic Operations for Stationary Spatial Processes. *Mathematical Geology*, 23, 695-701.
- Posa, D., & De Iaco, S. (2009). *Geostatistica teoria e applicazioni*. Torino, ITALIA: G. GIAPPICHELLI EDITORE.
- Richards, L. A., & Gardner, W. (1936). Tensiometers for measuring the capillarity tension of soil water. *Journal of the American Society of Agronomy*, 28, 352-358.
- Rieu, M., & Perrier, E. (1998). Models of fragmented and aggregated soils. In P. Baveye, & e. al, *Fractals in Soil Science* (p. 169-202). Boca Raton, USA: CRC.
- Rieu, M., & Sposito, G. (1991a). Fractal fragmentation, soil porosity and soil water properties: I. Theory. *Soil Science Society of American Journal*, 55, 1231-1238.
- Rieu, M., & Sposito, G. (1991b). Fractal fragmentation, soil porosity and soil water properties: II. Applications. *Soil Science Society of America Journal*, 55, 1239-1244.
- Ripley, B. D. (1981). *Spatial Statistics*. New York, USA: Wiley & Sons.
- Sarma, D. D. (2009). *Geostatistics with applications in earth sciences*. New Delhi, INDIA: Springer Velag.
- Schaap, M. G., & Leij, F. J. (1998). Database-related accuracy and uncertainty of pedotransfer functions. *Journal of Soil Science*, 163, 765-769.
- Scott, D. W. (1978). On optimal data-based histograms. *Biometrika*, 66, 605-610.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference* (p. 517-524). New York, USA: ACM Press.
- Sibson. (1981). A brief description of natural neighbour interpolation. In V. Barnett, *Interpreting Multivariate Data* (p. 21-36). John Wiley & Sons.
- Smirnov, N. V. (1939). Sur les Ecart de la Courbe de Distribution Empirique. *Matematicheskii Sbornik*, 6, 3-26.
- Snedecor, G. W., & Cochran, W. G. (1974). *Statistical Methods* (6th ed.). Ames, USA: Iowa State University Press.
- Sthephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69, 730-737.

- Sturges, H. A. (1926). The choice of class interval. *Journal of the American Statistical Association*, 21 (153), 65-66.
- Talsma, T. (1977). Measurement of the overburden component of total potential in swelling soil. *Journal of Soil Science*, 32, 499-504.
- Tietje, O., & Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *American Journal of Soil Science*, 57, 1088-1095.
- Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234-240.
- Toledo, G., Novy, R. A., Davis, H. T., & Scriven, L. E. (1990). Hydraulic conductivity of porous media at low water content. *Soil Science Society of American Journal*, 54, 673-679.
- Towner, G. D. (1981). The connection of in situ tensiometer readings for overburden pressure in swelling soil. *Journal of Soil Science*, 32, 499-504.
- Tukey, J. (1977). *Exploratory Data Analysis*. Reading, ENGLAND: Addison Wesley.
- Tyler, S. W., & Wheatcraft, S. W. (1989). Application of fractal mathematics to soil water retention estimation. *Soil Society of American Journal*, 53, 987-996.
- Tyler, S. W., & Wheatcraft, S. W. (1990). Fractal processes in soil water retention. *Water Resources Research*, 26, 1047-1054.
- van Damme, H. (1995). Scale invariance and hydric behaviour of soils and clays. *C. R. Acad. Sci. Paris*, 320, 665-681.
- van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 892-898.
- Vereecken, K., Diels, J., van Orshoven, J., Feyen, J., & Bouma, J. (1992). Functional evaluation of pedotransfer functions for the estimation of soil hydraulic properties. *American Journal of Soil Science*, 56, 1371-1378.
- Wackernagel, H. (1995). *Multivariate Geostatistics*. Berlin, GERMANY: Springer Verlag.
- Wang, K., Zhang, R., & Wang, F. (2005). Testing the pore-solid fractal model for the soil water retention function. *Soil Science Society of America Journal*, 69, 776-782.
- Watson, G. S. (1971). Trend-surface analysis. *Mathematical Geology*, 3 (3), 215-226.

- Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. Chichester, ENGLAND: John Wiley & Sons.
- Wilk, M. B., & Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika*, 55 (1), 1-17.
- Wosten, J. H., & van Genuchten, M. T. (1988). Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *American Journal of Soil Science*, 52, 1762-1770.
- Xie, H. (1993). *Fractal in Rock Mechanics*. Rotterdam, NEDERLAND: A. A. Balkena P. O.
- Yaglom, A. M. (1957). Some classes of random fields in n-dimensional space related to stationary random processes. *Theory of Probability and its Applications*, 2 (3), 273-320.
- Young, I. M., & Crawford, J. W. (1991). The fractal structure of soil aggregates: its measurement and interpretation. *Journal of Soil Science*, 42, 187-192.
- Zhang, C., & McGrath, D. (2004). Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. *Geoderma* (119), 261-265.