

ANALISI E SINTESI STATISTICA DI VARIABILI IDROLOGICHE ATTRAVERSO FUNZIONI COPULA

B. Bacchi, M. Balistrocchi

DICATAM – Università di Brescia

SOMMARIO

Numerosi problemi affrontati dall'idrologia applicata richiedono la messa in conto di più variabili. In questi casi, l'approccio multivariato è, teoricamente, il più opportuno. L'impiego di distribuzioni congiunte basate su modelli tradizionali non si è però mai dimostrato del tutto soddisfacente. L'introduzione delle funzioni copula anche in questa disciplina ha però condotto ad un sensibile miglioramento nella capacità di adattare le funzioni teoriche alla naturale variabilità delle grandezze idrologiche. In questa memoria si illustra l'impiego di tali funzioni nell'ambito dell'analisi bivariata delle altezze e delle durate di precipitazione. Una volta derivata la distribuzione congiunta, si propone un'applicazione inerente la stima delle portate al colmo per un ipotetico bacino di piccole dimensioni.

1. Introduzione

Nell'ambito dell'idrologia tecnica è molto spesso opportuno, se non indispensabile, affrontare alcuni problemi di carattere applicativo tenendo in considerazione più di una variabile naturale. Si pensi, ad esempio, alla perimetrazione delle aree allagabili, la progettazione degli invasi destinati alla laminazione delle piene o la valutazione del rischio di innesco di un movimento franoso. In tutti questi casi, la sollecitazione naturale dovrebbe essere rappresentata da almeno due variabili: il colmo ed il volume di piena, nei primi due casi, l'intensità e la durata di precipitazione nel terzo.

L'impiego dell'approccio multivariato, in linea di principio quello più appropriato, si è però scontrato all'atto pratico con la difficoltà di adattare le funzioni di probabilità disponibili, principalmente quelle basate sui modelli di tipo normale (Sackl & Bergmann, 1987, Chang et al., 1994) ed esponenziale (Correia, 1987; Bacchi et al., 1994), alle distribuzioni empiriche osservate. I limiti più restrittivi derivavano sia dal numero relativamente modesto di funzioni di probabilità disponibili, sia dal fatto che esse erano diretta estensione di funzioni di probabilità univariate.

Per tale motivo le distribuzioni marginali dovevano appartenere necessariamente allo stesso modello probabilistico, rendendo quasi sempre necessario trasformare le osservazioni campionarie allo scopo di modificarne opportunamente la distribuzione (Goel et al., 1998; Asklar, 2007). La stima della distribuzione congiunta diveniva perciò più articolata e maggiormente affetta da incertezza, dovendo individuare le funzioni della procedura di trasformazione. In presenza di dati non ellittici, sussisteva inoltre il rischio di alterare la struttura di dipendenza originale.

Infine, la stima dei parametri di dipendenza, essendo, di fatto, effettuata sulla base del coefficiente di correlazione lineare di Pearson, era inevitabilmente influenzata dalle stesse distribuzioni marginali, non essendo questo indice invariante rispetto a trasformazioni monotone crescenti delle variabili. Per tali motivi, i risultati conseguiti mediante questo approccio convenzionale non sono mai stati del tutto soddisfacenti.

In anni più recenti, l'applicazione delle funzioni copula (Joe, 1997; Nelsen, 2006) è stata proposta anche nel campo della ricerca idrologica (Dupuis, 2007; Salvadori et al., 2007), in quanto esse permettono di superare in modo agevole praticamente tutte le difficoltà anzidette. Secondo questo approccio, infatti, struttura di dipendenza, la cui rappresentazione è demandata esclusivamente alla funzione copula, e distribuzioni marginali possono essere analizzate separatamente.

Ciò amplia significativamente i margini di inferenza della distribuzione di probabilità congiunta, nella quale è possibile implementare funzioni marginali diverse, anche complesse ed appartenenti a qualsivoglia modello probabilistico. La funzione copula è invece scelta esclusivamente in relazione al modo di addensarsi dei dati nello spazio delle probabilità, per la cui rappresentazione è già disponibile una grande varietà di modelli.

In questa memoria, dopo una sintetica illustrazione della metodologia di analisi statistica bivariata mediante copule, si propone un'applicazione riferita alle precipitazioni osservate presso il pluviometro di Brescia: l'esempio mostra come sia possibile, tramite elaborazioni relativamente semplici, costruire la distribuzione di probabilità congiunta per l'altezza e la durata degli eventi meteorici e come essa sia utilizzabile nella pratica per la derivazione di distribuzioni delle portate al colmo prodotte da bacini di piccole dimensioni.

2. Analisi statistica bivariata tramite funzioni copula

Lo strumento fondamentale sul quale si basa l'analisi multivariata tramite funzioni copula è il teorema di Sklar (1959). Come mostrato dall'equazione (1) per il caso bivariato, secondo questo teorema data una distribuzione congiunta P_{XY} avente marginali P_X e P_Y esiste una funzione bivariata C_{XY} tale da soddisfare l'uguaglianza (1).

$$P_{XY}(x, y) = C_{XY}[P_X(x), P_Y(y)] \quad (1)$$

La funzione bivariata C_{XY} è comunemente detta copula e, se le marginali P_X e P_Y sono continue, è unica. Intuitivamente essa rappresenta la funzione di distribuzione di variabili casuali distribuite uniformemente nell'intervallo unitario $\mathbb{I} = [0,1]$ e ricavate da quelle originali attraverso la cosiddetta "Probability Integral Transform" (PIT), come mostrato nelle equazioni (2).

$$u = P_X(x); \quad v = P_Y(y) \quad \text{con} \quad u, v \in \mathbb{I} \quad (2)$$

La funzione $C_{XY}(u, v): \mathbb{I}^2 \rightarrow \mathbb{I}$ deve perciò soddisfare le condizioni di uniformità delle marginali (3) "border constrains" e di non decrescente monotonicità bivariata (4) "measure constrain".

$$C_{XY}(u, 0) = 0, C_{XY}(u, 1) = u, C_{XY}(0, v) = 0, C_{XY}(0, 1) = v \quad (3)$$

$$C_{XY}(u_2, v_2) - C_{XY}(u_2, v_1) - C_{XY}(u_1, v_2) + C_{XY}(u_1, v_1) \geq 0 \quad (4)$$

con $u_1 \leq u_2, v_1 \leq v_2$

Tra le infinite funzioni copula che soddisfano i requisiti (3) e (4), tre assumono un significato particolare. Le copule bivariate definite dalle equazioni (5) e (6) esprimono una perfetta dipendenza tra le variabili marginali, co-monotona nel primo caso e contro-monotona nel secondo. La copula M_2 è denominata limite superiore di Fréchet-Hoeffding, mentre la copula W_2 limite inferiore di Fréchet-Hoeffding, dato che per ogni coppia $(u, v) \in \mathbb{I}^2$ una generica funzione copula bivariata C_{XY} deve soddisfare il sistema di diseuguaglianze (7).

$$M_2(u, v) = \min \{ u, v \} \quad (5)$$

$$W_2(u, v) = \max \{ u + v - 1, 0 \} \quad (6)$$

$$W_2(u, v) \leq C_{XY}(u, v) \leq M_2(u, v) \quad (7)$$

La terza copula fondamentale Π_2 è quella che esprime la perfetta indipendenza statistica e che è espressa dal semplice prodotto delle variabili marginali, come mostrato dall'equazione (8) sempre per il caso bivariato.

$$\Pi_2(u, v) = uv \quad (8)$$

La figura 1 illustra queste tre copule bivariate fondamentali attraverso sia la loro rappresentazione tridimensionale, sia le corrispondenti linee di livello.

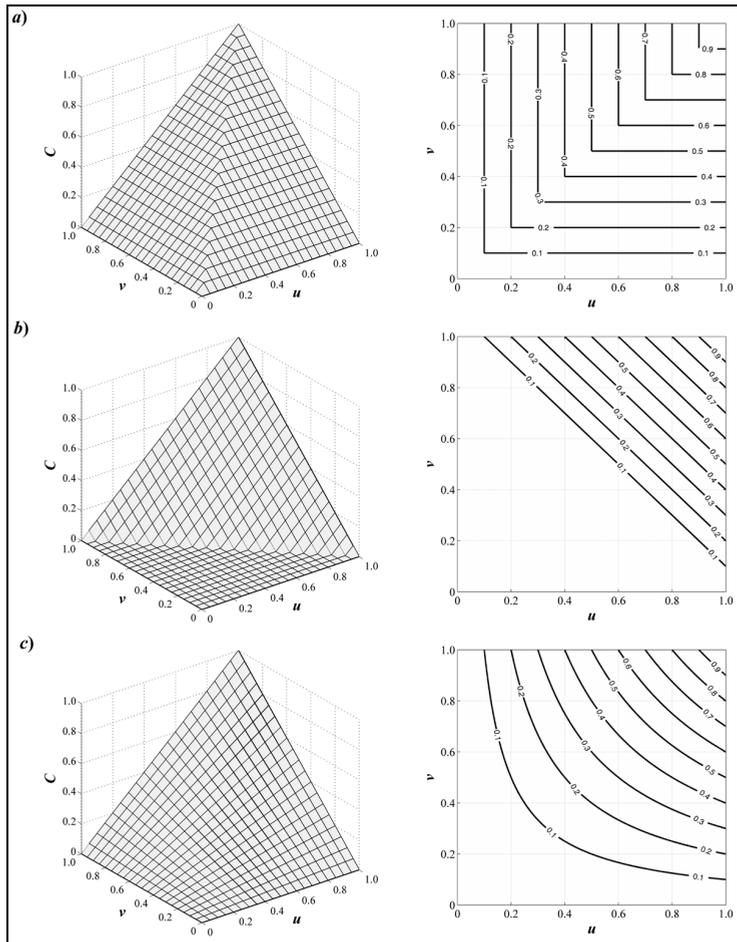


Fig. 1 - Funzioni copula bivariate fondamentali a) M_2 , b) W_2 e c) Π_2

Una macro-famiglia di funzioni copula, alla quale si ricorrerà nell'esempio applicativo proposto di seguito, è quella archimedeica. Essa presenta svariati vantaggi rispetto ad altre soluzioni, tanto da risultare tra le più popolari. Si tratta infatti di funzioni esplicite e mono-parametriche, nelle quali il parametro che rap-

presenta la dipendenza tra le variabili marginali è esprimibile in ragione di misure di associazione tramite semplici espressioni algebriche. Ciò fa sì che la calibrazione della funzione tramite il cosiddetto metodo dei momenti, illustrato nella sezione 2.3, sia immediata. Ciò riduce grandemente gli oneri computazionali in fase di stima della distribuzione e spesso anche in quella successiva di verifica. Dal punto di vista prettamente matematico le copule archimedee ammettono l'esistenza di una funzione generatrice $\varphi: \mathbb{I} \rightarrow [0, \infty]$ continua, strettamente decrescente, convessa con $\varphi(1)=0$, tale che sussista la condizione (9).

$$C_{XY}(u, v) = \varphi^{[-1]}[\varphi(u) + \varphi(v)] \quad (9)$$

Tra le proprietà che conseguono dalla relazione (9), quella più rilevante a fini pratici è la simmetria della copula C_{XY} . Questo se da un lato può rappresentare una limitazione, tanto che sono stati sviluppati appositi test statistici allo scopo di verificarne l'attendibilità (Genest et al., 2012), dall'altro implica l'interscambiabilità delle variabili u e v . Questo invero non pone particolari problemi nel momento della costruzione della distribuzione congiunta la quale, dovendo in generale implementare marginali differenti, non sarà simmetrica.

2.1 Stima del tempo di ritorno

Il concetto di tempo di ritorno $T(x)$, generalmente espresso in anni, è ben definito nel caso univariato in ragione della probabilità di non superamento $P(x)$ di un evento x e del numero medio di eventi annui ω (anni^{-1}), secondo quanto indicato dall'equazione (10). Tale formulazione risulta purtroppo di difficile estensione al caso multivariato. Il motivo risiede essenzialmente nel fatto che, a differenza del primo caso, nel secondo lo spazio delle variabili non risulta essere totalmente ordinato.

$$T(x) = \frac{1}{\omega(1 - P(x))} \quad (10)$$

Nel caso bivariato esistono infinite possibilità di ripetire lo spazio delle variabili \mathbb{I}^2 in relazione ad un evento, al fine di delimitare una regione sub-critica nonché quella complementare super-critica. Convenzionalmente si è sempre fatto riferimento alle due soluzioni limite in figura 2, dove la regione sub-critica è evidenziata in grigio.

Nel riquadro *a*) è rappresentata la regione rispetto alla quale stimare la probabilità di non superamento di un generico evento (U, V) quando si intende come super-critico un evento nel quale almeno una delle due variabili sia superata; nel riquadro *b*) invece la stessa regione è rappresentata quando si intende come super-critico un evento nel quale entrambe le variabili siano superate. Nella prima situazione il tempo di ritorno è detto OR, mentre nella seconda è detto AND e possono essere stimati rispettivamente come indicato nelle espressioni (11) e (12).

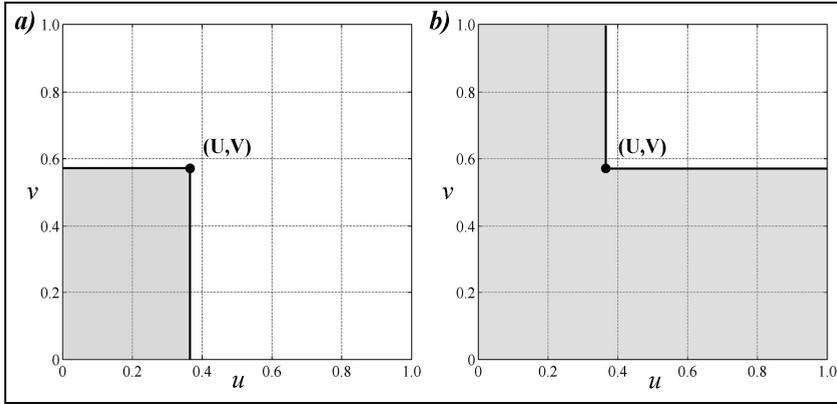


Fig. 2 - Regione sub-critica per un generico evento (U,V) quando si vogliono calcolare T^{OR} a) e T^{AND} b)

$$T^{\text{OR}}(u,v) = \frac{1}{\omega(1 - C_{XY}(u,v))} \quad (11)$$

$$T^{\text{AND}}(u,v) = \frac{1}{\omega(1 - u - v + C_{XY}(u,v))} \quad (12)$$

Il tempo di ritorno effettivo di un evento dovrebbe quindi essere incluso tra questi due valori estremi. La differenza tra queste stime tuttavia è generalmente molto ampia, in virtù dell'estrema sensibilità di questa grandezza nei confronti della probabilità di non superamento. Mediante l'uso delle funzioni copula sono state proposte diverse soluzioni, allo scopo di ovviare a questo problema concettuale di importanza non indifferente. In generale esse si sono concentrate nel tentativo di ricondurre la stima ad uno schema analogo a quello del caso univariato: in alcuni casi sfruttando analisi di regressione, distribuzioni condizionali o funzioni di misura dello spazio delle probabilità (Gräler et al., 2013).

In questo senso, Salvadori et al. (2010) propongono di ripartire lo spazio delle variabili uniformi \mathbb{I}^2 in relazione alle proiezioni \mathcal{L}_t sul piano (u,v) delle curve di livello della funzione copula, così come definito dall'equazione (13), dato che, per un qualsiasi evento (U,V) , esse sono univocamente determinate e ripartiscono lo spazio in due regioni dicotomiche.

$$\mathcal{L}_t = \{(u,v) \in \mathbb{I}^2 : C_{XY}(u,v) = t\} \quad \text{con} \quad 0 \leq t \leq 1 \quad (13)$$

In riferimento a queste regioni è possibile calcolare la probabilità che si verifichi un evento caratterizzato da un valore della copula C_{XY} inferiore al livello t , la quale sarà comune a tutti gli eventi appartenenti alla stessa curva \mathcal{L}_t . Questo permette di definire la funzione di Kendall K_C , formalmente definita in (14) per il caso bivariato.

$$K_C(t) = \text{Prob}\{C_{XY}(u, v) \leq t\} \quad \text{con } 0 \leq t \leq 1 \quad (14)$$

Dato che questa funzione presenta tutte le caratteristiche di una distribuzione univariata, il suo valore può essere formalmente sostituito alla generica probabilità di non superamento $P(x)$ nella definizione di tempo di ritorno (10), pervenendo quindi alla (15).

$$T^K(t) = \frac{1}{\omega(1-K_C(t))} \quad (15)$$

In generale i valori assunti da K_C possono essere calcolati integrando la densità della copula sul dominio delimitato dalla curva \mathcal{L} . Per le copule archimedee, essa assume una semplice formulazione analitica, che può essere determinata attraverso la funzione generatrice φ , come indicato dall'equazione (16).

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t)} \quad (16)$$

2.3 Stima dei parametri

Nell'ambito di questa metodologia, la rappresentazione dei dati campionari (X_i, Y_i) avviene attraverso la definizione delle pseudo-osservazioni (U_i, V_i) ricavate tramite le espressioni (17), controparti empiriche delle equazioni (2), nelle quali R rappresenta la funzione rango ed N la numerosità campionaria.

$$U_i = \frac{R(X_i)}{N+1}; V_i = \frac{R(Y_i)}{N+1} \quad \text{con } i = 1, \dots, N \quad (17)$$

Le pseudo-osservazioni consentono di definire la copula empirica C_N , la cui funzione è riportata dall'espressione (18), dove $\mathbf{1}$ è la cosiddetta “*indicator function*”, che vale 1, quando la condizione nel suo argomento è soddisfatta, e 0 viceversa. Rispetto a questa funzione empirica avviene la scelta della funzione copula teorica e la sua capacità di adattamento alle pseudo-osservazioni. Essa infatti sintetizza il comportamento congiunto dei dati campionari, essendo una stima non parametrica della funzione teorica C_{XY} (Deheuvels, 1979).

$$C_N(u, v) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(U_i \leq u, V_i \leq v) \quad (18)$$

Per poter eseguire il confronto tra funzioni teoriche ed empiriche, occorre che prime siano calibrate. I metodi che si possono seguire sono essenzialmente due: quello di massima verosimiglianza e quello dei momenti. Il primo può essere implementato secondo un approccio semi-parametrico impiegando lo stimatore seguente (Genest et al., 1995).

$$L(\boldsymbol{\theta}) = \sum_{i=1}^N \log[c_{XY}(U_i, V_i)] \quad (19)$$

Nell'espressione (19) c_{XY} rappresenta la densità della copula C_{XY} e può essere facilmente derivata come mostrato all'equazione (20).

$$c_{XY}(u, v) = \frac{\partial^2 C_{XY}}{\partial u \partial v} \quad (20)$$

Il vettore θ contiene invece i parametri di dipendenza, ovvero i parametri della copula; esso è ottenuto, come di consueto nel metodo di massima verosimiglianza, massimizzando lo stimatore (19). Quest'ultimo è detto di pseudo-verosimiglianza in quanto basato solo sull'impiego della funzione C_{XY} . Metodi alternativi che impiegano invece la densità di probabilità della funzione (1), definiscono stimatori completi che portano simultaneamente alla stima dei parametri della copula e delle marginali. Lo stimatore $L(\theta)$, sebbene non efficiente (Genest & Werker, 2002), è tuttavia da preferire sia dal punto di vista dell'onere computazionale sia nell'intento di mantenere completamente distinta la stima delle diverse componenti della distribuzione (1).

Qualora la funzione copula prescelta sia mono-parametrica, nel caso bivariato è possibile ricorrere ad un metodo analogo a quello dei momenti, di uso comune nella statistica univariata. In questo caso, si ricorre a misure di associazione le quali, in modi diversi, sono costruite in riferimento alle probabilità di concordanza e di discordanza delle variabili.

Esse presentano due vantaggi fondamentali rispetto al tradizionale coefficiente di correlazione lineare di Pearson: in quanto indici di correlazione di rango esistono sempre, sono scala invarianti e, perciò, non dipendono dalle distribuzioni marginali. Le misure di associazione sono quindi legate esclusivamente dalla copula e, in molte di esse, come quelle archimedee, possono essere messe in relazione direttamente al parametro di dipendenza θ .

Una delle misure di associazione più comunemente usate è il coefficiente di Kendall (1938), la cui versione teorica è definita in termini di differenza tra la probabilità di concordanza e la probabilità di discordanza (21).

$$\tau_k = \mathbb{P}\{(x_i - x_j)(y_i - y_j) > 0\} - \mathbb{P}\{(x_i - x_j)(y_i - y_j) < 0\} \quad (21)$$

La versione campionaria utilizzata per stimarne il valore è invece data dall'espressione (22), nella quale n_c ed n_d indicano il numero di coppie concordanti ed il numero di coppie discordanti.

$$\tau_k = \frac{n_c - n_d}{n_c + n_d} \quad (22)$$

2.4 Test di adattamento

La valutazione dell'adattamento delle funzioni copula alle pseudo-osservazioni può essere condotta tramite idonei test statistici. Tra le innumerevoli metodologie proposte, quelle classificabili come “*blanket tests*” hanno evi-

denziato significativi vantaggi (Genest et al., 2006), dato che possono essere adottati per ogni famiglia di copule e non richiedono alcuna scelta soggettiva per poter essere implementati. Tra i test ritenuti di maggiore efficacia da Genest et al. (2009) vi è quello basato sul criterio di adattamento di Cramer-Von Mises (23), che misura la sommatoria gli scarti quadratici tra copula teorica (1) e copula empirica (11) in corrispondenza delle pseudo-osservazioni.

$$S_N = \sum_{i=1}^N [C_N(U_i, V_i) - C_{XY}(U_i, V_i)]^2 \quad (23)$$

Al fine di testare un'ipotesi nulla, secondo la quale la funzione copula parentale dalla quale sono estratte le pseudo-osservazioni corrisponde a quella prescelta, è possibile ricorrere ad una procedura i cui passaggi sono i seguenti:

- 1) Si valuta la statistica S_N confrontando la copula empirica C_N e la copula teorica C_{XY} definite rispetto alle pseudo-osservazioni;
- 2) Sulla base dell'ipotesi nulla si genera un nuovo campione di pseudo-osservazioni della stessa numerosità di quello originale;
- 3) Si calibra nuovamente una copula teorica $C_{XY,k}$ rispetto alle pseudo-osservazioni generate;
- 4) Si costruisce una nuova copula empirica $C_{N,k}$ rispetto alle pseudo-osservazioni generate;
- 5) Si valuta la statistica $S_{N,k}$ confrontando la copula empirica $C_{N,k}$ e la copula teorica $C_{XY,k}$ definite rispetto alle pseudo-osservazioni generate;
- 6) Si ripetono i punti da 2) a 5) per un numero m sufficientemente grande (almeno cento volte la numerosità campionaria N) in modo da poter valutare la sommatoria (24).

L'indice p ottenuto dalla (24), comunemente detto “ p -value”, rappresenta una stima empirica della significatività del test, ovvero della probabilità di rigettare l'ipotesi nulla quando questa è vera.

$$p = \frac{1}{m} \sum_{k=1}^m \mathbf{1}(S_{N,k} > S_N) \quad (24)$$

Per poter generare un campione come richiesto al punto 2), infine, si deve ricorrere ad una procedura di simulazione della copula testata (Salvadori et al., 2007). Quest'ultima può essere articolata nel caso bivariato nei passaggi seguenti:

- 1) Si generano due numeri casuali β_1 e β_2 da una distribuzione uniforme in \mathbb{I} ;
- 2) Si costruisce la distribuzione della variabile v condizionata alla variabile $u=\beta_1$ sfruttando le proprietà delle funzioni copula, come mostrato dall'equazione (25);
- 3) Una coppia di pseudo-osservazioni (U_i, V_i) estratta dalla copula C_{XY} è quindi data dalle uguaglianze (26).

$$c_{v|\beta_1}(v) = \text{Prob}\{v' \leq v \mid u = \beta_1\} = \frac{\partial}{\partial u} C_{XY}(\beta_1, v) \quad (25)$$

$$\begin{cases} U_i = \beta_1 \\ V_i = c_{v|\beta_1}^{-1}(\beta_2) \end{cases} \quad (26)$$

3. Analisi di serie di precipitazioni

L'analisi statistica delle precipitazioni più comunemente impiegata nella pratica idrologica è condotta secondo un approccio univariato in riferimento alle sole altezze, o volumi, di precipitazione massime annuali di assegnata durata. Essa porta alla definizione delle cosiddette linee segnalatrici di possibilità pluviometrica (LSPP), (Moisello, 1999) mediante le quali è possibile ricavare, per un fissato tempo di ritorno, ossia per una certa frequenza di accadimento, l'altezza di precipitazione corrispondente ad una generica durata. Attraverso le LSPP è quindi possibile costruire eventi pluviometrici sintetici di progetto, o ietogrammi di progetto.

Problema essenziale da risolvere in questa operazione è la scelta della distribuzione spazio-temporale delle altezze di precipitazione, la cui caratteristica più rilevante è appunto la durata. Sebbene nella pratica progettuale si adottino spesso distribuzioni uniformi sia nello spazio che nel tempo, questi due aspetti sono però legati a considerazioni di tipo diverso. Nel primo caso risultano rilevanti l'estensione del bacino di interesse e le sue caratteristiche morfometriche. Nel secondo è invece necessario tenere presente che, al crescere della durata dell'evento pluviometrico, aumentano le altezze di precipitazione ma si riducono le intensità medie.

Ne consegue che gli effetti di una precipitazione breve ed intensa, per fenomeni caratterizzati da tempi di risposta non troppo piccoli, possono essere meno importanti di quelli prodotti da precipitazioni meno intense ma di durata maggiore. Nella scelta della durata il tempo di risposta idrologica del bacino di interesse assume pertanto un'importanza fondamentale. Al fine di ovviare a questa difficoltà si ricorre, spesso, al criterio dell'evento critico, ovvero si ricerca quella particolare durata che, sotto alcune ipotesi semplificative, massimizzi la sollecitazione idrologica del manufatto per il tempo di ritorno prescelto.

I principali inconvenienti di questo approccio univariato risiedono nell'assenza di ogni riferimento della durata dell'evento pluviometrico alla naturale variabilità del processo di precipitazione e nell'associare un tempo di ritorno all'intero evento il quale è, in realtà, espressione di una sola delle variabili che lo caratterizzano. Un certo grado di avanzamento può quindi essere conseguito affrontando il problema della costruzione di eventi di progetto attraverso uno schema bivariato, che esprima la naturale variabilità dell'altezza e della durata di precipitazione.

L'applicazione della metodologia di analisi statistica tramite funzioni copula qui illustrata esamina una lunga serie di precipitazioni registrate presso il pluviometro ITAS Pastori di Brescia. La serie è costituita da 45 anni di osservazioni, tra il 1949 ed il 1993, a cadenza semioraria. Il clima è caratterizzato da un regime pluviometrico di tipo sub-litoraneo alpino, con una precipitazione media annua di poco inferiore ai 1000 mm.

3.1 Analisi dei dati

Allo scopo di analizzare la variabilità congiunta di altezze e durate di precipitazione, si rende necessario separare le serie continue in eventi isolati. Per far questo è opportuno definire due parametri: un tempo minimo di inter-evento ed una soglia minima per l'altezza di precipitazione. Il primo ha lo scopo di definire il tempo minimo asciutto che deve trascorrere affinché due scrosci successivi possano essere considerati indipendenti, il secondo quello di eliminare dalla serie le precipitazioni non significative.

I valori da attribuire a questi parametri esercitano una considerevole influenza sulle proprietà statistiche dei campioni ottenuti, sicché la loro scelta si deve effettuare in maniera opportuna. Un criterio che si è dimostrato efficace in alcuni problemi pratici (Balistrocchi et al., 2013) è quello di legarli a parametri idrologici caratterizzanti la specifica applicazione: talvolta, nei problemi di idrologia urbana, il tempo di inter-evento minimo IETD può quindi essere indentificato con il tempo di corrivazione, mentre la soglia sull'altezza IA può essere assimilata alla perdita idrologica iniziale, che deve essere superata per poter avere deflusso superficiale.

La procedura di identificazione degli eventi di precipitazione indipendenti prevede pertanto che gli scrosci individuali, separati da tempi asciutti inferiori a IETD, siano aggregati in un unico evento. L'altezza e la durata della precipitazione sono così valutate in riferimento all'istante iniziale del primo scroscio e quello finale dell'ultimo; nei casi in cui l'altezza di precipitazione risulti inferiore ad IA, l'evento viene soppresso e la corrispondente durata è assimilata ad un periodo di tempo asciutto.

3.2 Analisi delle distribuzioni marginali

Un primo aspetto da tenere in considerazione è certamente la sensibilità mostrata dal numero medio di eventi annui ω nei confronti dei parametri della procedura di identificazione degli eventi indipendenti. Come si può osservare dalle curve in figura 3, si evidenzia una generale tendenza alla diminuzione di ω al crescere sia del tempo minimo di inter-evento IETD sia della soglia sull'altezza di precipitazione IA. Nel primo caso, ciò è dovuto alla progressiva aggregazione degli scrosci in eventi maggiori, mentre nel secondo alla soppressione degli eventi minori.

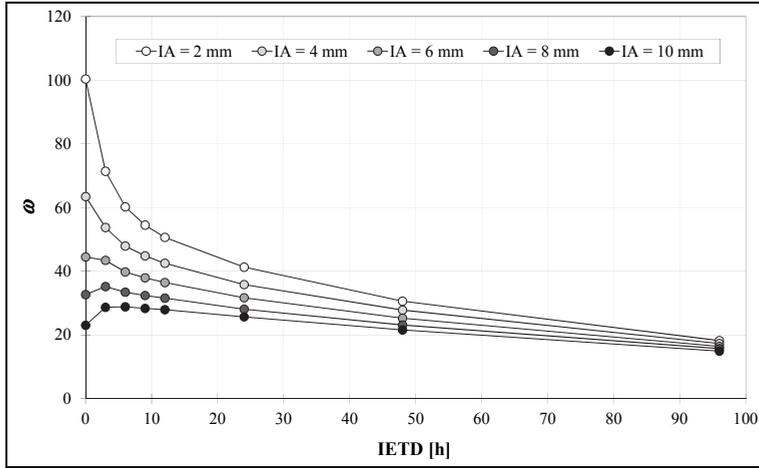


Fig. 3 - Variabilità del numero medio di eventi annui ω in relazione ai parametri della procedura di individuazione degli eventi indipendenti IA e IETD.

Come conseguenza la sensitività di ω rispetto ad IA risulta meno rilevante quando IETD è maggiore. Solo per soglie IA molto alte e tempi di inter-evento IETD piccoli si evidenzia, tuttavia, un incremento di ω al crescere del secondo parametro. Ciò è giustificato dal fatto che, in questa condizione, risulta predominante il fenomeno dell'incremento di volume di pioggia degli eventi, conseguente l'aggregazione degli scrosci, che consente ad un numero maggiore di eventi di superare il criterio della soglia sull'altezza di totale di precipitazione; tale fenomeno, nelle altre condizioni, appare invece irrilevante.

Per quanto concerne invece la scelta della distribuzione teorica per la rappresentazione delle marginali, un modello impiegato molto frequentemente quando si trattano eventi non necessariamente estremi, è la distribuzione esponenziale. Il suo non soddisfacente adattamento ai campioni estratti dalla serie per svariate combinazioni dei parametri IETD ed IA, tuttavia, ha suggerito il ricorso a modelli più complessi. Una soluzione idonea è stata individuata nella distribuzione di Weibull (Balistocchi & Bacchi; 2011), che può essere considerata una generalizzazione dello stesso modello esponenziale.

Le funzioni marginali definite dalle equazioni (27) e (28) per l'altezza h (mm) e la durata d (h) di precipitazione, si contraddistinguono oltre che per i parametri di scala ζ (mm) e λ (h), anche per la presenza di parametri di forma β e γ , adimensionali. Nella funzione marginale (27) si introduce inoltre un limite inferiore pari ad IA, in modo da renderla coerente con la procedura di individuazione degli eventi indipendenti.

$$P_H(h) = 1 - \exp \left[- \left(\frac{h - IA}{\zeta} \right)^\beta \right] \quad (27)$$

$$P_D(d) = 1 - \exp \left[- \left(\frac{d}{\lambda} \right)^\gamma \right] \quad (28)$$

La stima dei parametri delle marginali (27) e (28), effettuata secondo il criterio della massima verosimiglianza al variare di IETD ed IA, è sintetizzata in tabella 1: mentre il parametro di forma β assume sistematicamente valori inferiori all'unità, γ appare nella maggioranza delle situazioni superiore a tale limite, anche se si osservano valori pari od inferiori.

Le conseguenze sulla forma della marginale sono rilevanti, dato che per valori del parametro di forma inferiori all'unità, la densità della distribuzione è contraddistinta da un asintoto verticale in corrispondenza del limite inferiore, per valori unitari la distribuzione diviene semplicemente esponenziale, la moda assume un valore finito, pari al reciproco del parametro di scala, ma si verifica comunque in corrispondenza del limite inferiore, infine al crescere di tale parametro oltre l'unità la moda tende ad allontanarsi sempre più dal limite inferiore.

In generale si può quindi osservare come nel caso delle altezze di precipitazione al crescere di entrambi i parametri della procedura di individuazione degli eventi indipendenti, la densità di probabilità rimanga comunque concentrata in corrispondenza del limite inferiore IA, sebbene di evidenze una leggera tendenza all'aumento della varianza. Nel caso invece della durata di precipitazione, al crescere della soglia IA essa tende a discostarsi sempre più dall'origine degli assi mentre l'aumento del tempo minimo IETD determina l'effetto opposto. Il comportamento dei parametri di scala è, come era peraltro facilmente prevedibile, più regolare dato che, per entrambe le variabili, essi crescono in modo sostanzialmente lineare sia con IETD che con IA.

Più interessante è notare invece come, quando si confrontano i parametri delle marginali (27) e (28) stimati in vari regimi pluviometrici italiani, quelli di forma non manifestino significative differenze rispetto a quanto descritto in questo caso. I parametri di scala, invece, pur mantenendo andamenti qualitativamente analoghi rispetto a IETD ed IA, presentano valori maggiori, o minori, a seconda che il clima sia più piovoso o più secco (Balitrocchi & Bacchi, 2011).

Un esempio della qualità dell'adattamento conseguibile attraverso il modello di Weibull è offerto nelle figure 4 e 5, nelle quali si mostrano i test della fascia fiduciaria per un livello di significatività del 5%, effettuati fissando IETD a 12 h ed IA pari a 4 mm. In questi grafici ν e δ rappresentano le variabili ridotte corrispondenti ad h e d , date dal rapporto tra la variabile naturale, a meno di IA nel caso dell'altezza di precipitazione, ed il parametro di scala.

IETD	3			6			12			24			48		
IA	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10
β	0.78	0.88	0.88	0.79	0.88	0.88	0.81	0.90	0.90	0.84	0.91	0.92	0.87	0.89	0.93
ζ	8.0	10.9	12.2	10.3	13.3	14.4	13.2	16.1	17.1	17.5	20.3	21.3	25.5	27.3	29.1
γ	1.06	1.18	1.32	1.03	1.22	1.35	1.00	1.19	1.30	0.97	1.16	1.30	0.90	1.02	1.15
λ	6.6	9.5	12.1	10.4	14.7	18.0	16.5	22.4	26.9	29.3	38.2	45.5	60.8	74.5	88.7

Tab. 1 - Stima dei parametri per le distribuzioni marginali al variare di IETD ed IA, β ($^{\circ}$), γ ($^{\circ}$), ζ (mm), λ (h)

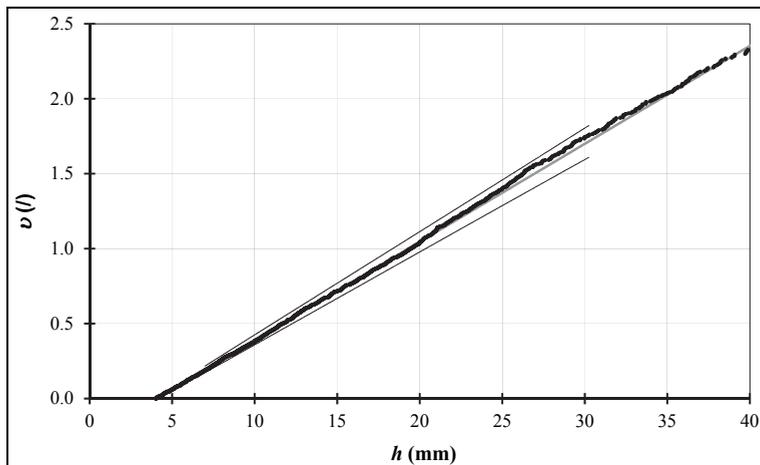


Fig. 4 - Rappresentazione della funzione $v = (h-IA)/\zeta$ e della fascia fiduciaria con significatività 5% quando IETD = 12 h e IA = 4 mm (funzione teorica linea grigia marcata, valori empirici punti neri, fasce di confidenza linee nere sottili)

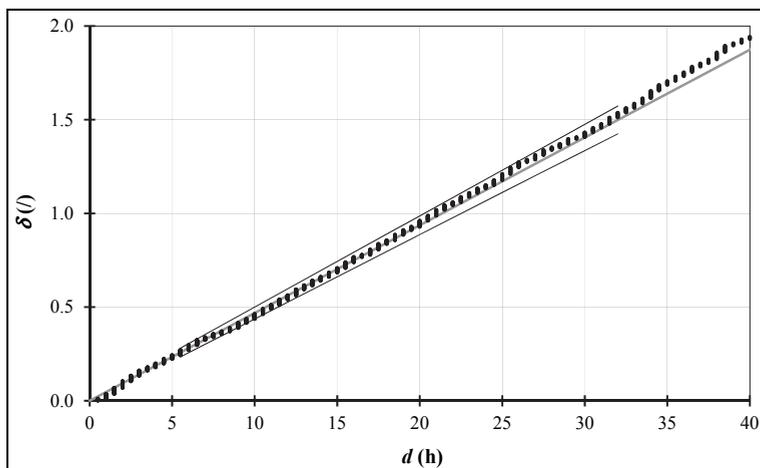


Fig. 5 - Rappresentazione della funzione $\delta = d/\lambda$ e della fascia fiduciaria con significatività 5% quando IETD = 12 h e IA = 4 mm (funzione teorica linea grigia marcata, valori empirici punti neri, fasce di confidenza linee nere sottili)

3.3 Analisi della struttura di dipendenza

Il primo passo nell'analisi della struttura di dipendenza che intercorre tra due variabili casuali è quello di valutare il grado di associazione empirico, allo scopo di stabilire innanzi tutto la forza del legame e, nel caso in cui sia significativo, se esso è concordante o discordante. Questo permette di escludere immedia-

tamente quelle famiglie di copule che non sono in grado di rappresentare quel particolare tipo di legame. La figura 6 propone quindi questo tipo di valutazione, facendo riferimento alla misura di Kendall τ_k (22) valutata per diverse coppie di IETD ed IA. Per tutte le combinazioni esaminate, τ_k si dimostra significativamente diverso da zero, e positivo, evidenziando un'associazione sempre concordante tra le due variabili. Il grado dell'associazione tende inoltre a crescere con IETD ma a diminuire con IA.

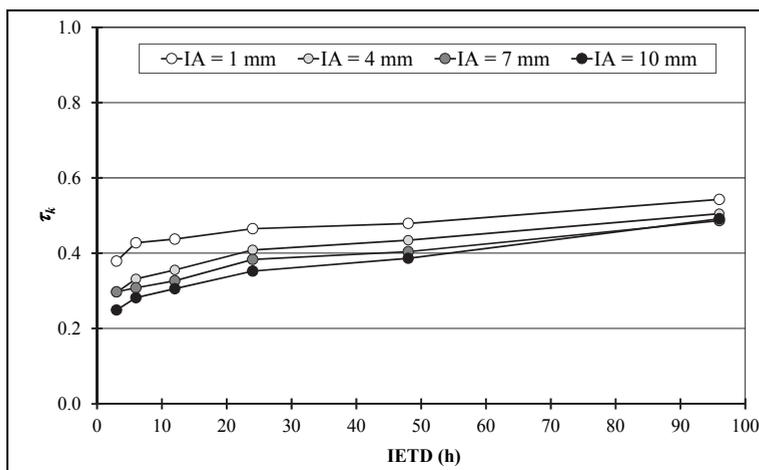


Fig. 6 - Andamento del coefficiente di Kendall τ_k per vari IETD ed IA

La prima tendenza può essere spiegata considerando che l'aggregazione degli scrosci in eventi di maggiore entità tende a ridurre il numero e, di conseguenza, la variabilità. Il secondo è legato alla soppressione degli scrosci di minor volume, cui sono tendenzialmente associate brevi durate; permangono tuttavia eventi di breve durata e forte intensità e, quindi, di maggiore volume. Essendo i primi tendenzialmente concordanti ed i secondi discordanti, l'effetto complessivo sul campione è quello di diminuire il valore del coefficiente τ_k .

Una volta individuate le famiglie di copule in grado di rappresentare questo tipo di associazione, si procede a stimarne i parametri mediante il metodo di massima verosimiglianza o quello dei momenti ed a confrontarle graficamente con la copula empirica (18). Dopo aver circoscritto la scelta della copula teorica a quelle che palesemente mostrano di spiegare meglio la variabilità naturale delle osservazioni, è consigliabile condurre dei test di statistici al fine di stimare in modo non soggettivo la qualità dell'adattamento. In riferimento al test qui illustrato, si adotterà, a meno che non sussistano motivazioni particolari, quel modello che conduce al minor S_N ed al p -value maggiore.

Tra le numerose famiglie prese in esame, la copula di Gumbel-Hougaard (29) ha dimostrato la migliore capacità di adattare le pseudo-osservazioni (17) al modello teorico, per ogni scelta dei parametri di separazione degli eventi indipendenti. Questa è una famiglia di copule strettamente archimedee avente

come funzione di generazione la funzione (30), i cui membri sono tutti assolutamente continui. La funzione copula è definita da un unico parametro di dipendenza θ , legato al coefficiente di Kendall tramite la semplice relazione algebrica (31). Date le limitazioni indicate sempre in quest'ultima espressione, la copula di Gumbel-Hougaard è in grado di rappresentare solamente variabili associate positivamente, o indipendenti.

$$C_{\text{HD}}(u, v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\} \quad (29)$$

$$\varphi(t) = (-\ln t)^\theta \quad (30)$$

$$\theta = \frac{1}{1 - \tau_K} \quad \text{con } \theta \geq 1 \quad (31)$$

La famiglia Gumbel-Hougaard include infatti la copula indipendenza Π_2 (8), quando θ vale 1. Quando, viceversa, θ tende ad infinito, la copula approssima il limite superiore M_2 (6). La funzione di Kendall (14) per questa famiglia si scrive, sostituendo funzione di generazione (31) nella relazione (16), come indicato nella (32).

$$K_C(t) = t \left(1 - \frac{\ln(t)}{\theta}\right) \quad (32)$$

Una proprietà di un certo interesse applicativo di questa funzione riguarda i coefficienti di dipendenza di coda. Questi parametri statistici, nella loro versione teorica, sono definibili in termini di copula, come indicato nelle equazioni (33), e misurano la propensione delle variabili a formare coppie di valori fortemente associati nella coda inferiore (33a) e superiore (33b). Rispetto alla misura τ_K , che sintetizza questo tipo di propensione globalmente, questi due ulteriori parametri offrono pertanto la possibilità di indagare il comportamento congiunto delle variabili focalizzandosi su alcune porzioni dello spazio \mathbb{I}^2 .

In questa famiglia, mentre il coefficiente λ_L è identicamente nullo, quello della coda superiore λ_U , quando θ è diverso dall'unità, è maggiore di zero e soddisfa la relazione (34). Il buon adattamento tra la copula di Gumbel-Hougaard e le pseudo-osservazioni, dimostra come le altezze e le durate di precipitazione si contraddistinguano per una maggiore indipendenza negli eventi minori ma, al contrario, mostrino una spiccata tendenza a formare coppie concordanti negli eventi maggiori.

$$\begin{aligned} \lambda_L &= \lim_{t \rightarrow 0^+} \left[\frac{C_{XY}(t, t)}{t} \right] & a) \\ \lambda_U &= \lim_{t \rightarrow 1^-} \left[\frac{1 - 2t + C_{XY}(t, t)}{1 - t} \right] & b) \end{aligned} \quad (33)$$

Questo si accorda molto bene con l'evidenza sperimentale, dato che gli eventi di minor durata possono essere associati anche a forte intensità e, quindi, a significative altezze di precipitazione. Viceversa, questo non avviene per gli eventi di lunga durata, che sono quasi sempre associati a grandi volumi di precipitazione.

$$\lambda_U = 2 - 2^{1/\theta} \quad (34)$$

La capacità di adattamento di questa famiglia è mostrata in figura 7. Nel riquadro *a*) si riporta il confronto tra le linee di livello della copula teorica, calibrata secondo il metodo dei momenti, e della copula empirica, le cui pseudo-osservazioni sono ricavate assumendo IETD pari a 3 h ed IA 5 pari a mm. Nel riquadro *b*) è rappresentata invece un'immagine tridimensionale della funzione teorica e della copula empirica stimata nelle pseudo-osservazioni.

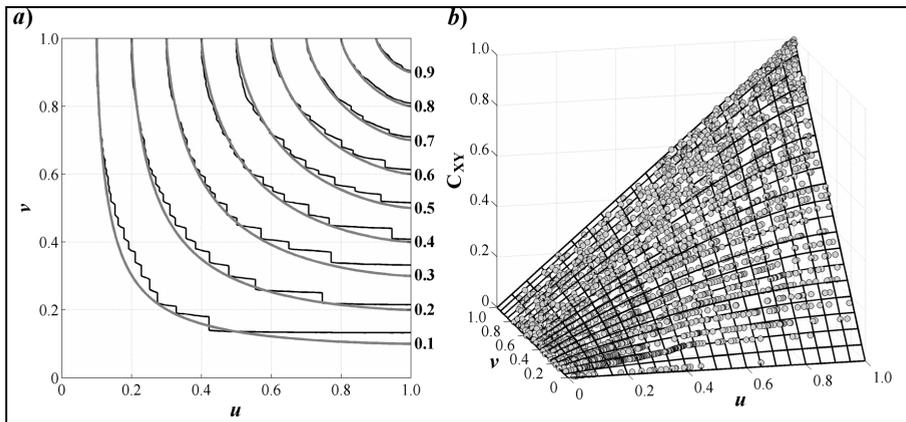


Fig. 7 - Adattamento della copula di Gumbel-Hougaard alla copula empirica nel caso in cui IETD 3 h ed IA 5 mm *a*) linee di livello (linea grigia continua copula teorica, linea nera sottile copula empirica), *b*) superficie tridimensionale in \mathbb{I}^2 (superficie continua, copula teorica; punti isolati copula, empirica)

La tabella 2 riporta infine i risultati del test di adattamento basato sul criterio di Cramer-Von Mises (23), per scelte diverse dei parametri di separazione degli eventi indipendenti IETD e IA. Si può osservare come l'ipotesi nulla, secondo la quale la copula teorica appartiene alla famiglia (29), non sia rigettabile in tutti i casi per significatività molto elevate. In particolare l'adattamento si dimostra migliore per valori elevati sia di IETD che di IA. Come ulteriore conferma della significatività dell'associazione è stato verificato anche l'adattamento della copula indipendenza Π_2 (8). Sotto questa ipotesi nulla, il test conduce per ogni combinazione di IETD ed IA a *p-value* praticamente nulli, evidenziando come questa copula sia largamente rigettabile.

IETD	3			6			12			24			48		
IA	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10
$\theta(l)$	1.61	1.41	1.33	1.75	1.47	1.39	1.78	1.52	1.44	1.87	1.66	1.54	1.92	1.75	1.63
$S_N(l)$	0.14	0.07	0.02	0.07	0.04	0.03	0.05	0.04	0.02	0.03	0.02	0.02	0.04	0.02	0.02
p (%)	40.4	79.2	99.9	78.7	97.6	99.8	93.4	95.9	99.9	99.7	99.9	99.9	98.8	99.9	99.9

Tab. 2 - Statistica S_N e stima dei p -value per il test di adattamento della copula di Gumbel-Hougaard al variare di IETD ed IA

3.4 Costruzione della distribuzione congiunta

L'applicazione del teorema di Sklar (1) conduce infine all'inferenza della distribuzione congiunta delle altezze e delle durate di precipitazione P_{HD} . In figura 8 si propone un confronto tra questa distribuzione teorica e quella empirica, per una particolare combinazione di IETD ed IA. In considerazione della complessità del problema, il grado di adattamento del modello probabilistico appare buono in ogni parte del dominio esaminato.

Un'applicazione immediata della distribuzione così ottenuta è legata alla possibilità di implementare procedure numeriche, atte a generare campioni di dati delle variabili indagate di qualsivoglia dimensione. I campioni di pseudo-osservazioni ricavati attraverso la procedura di simulazione, qui impiegata nell'ambito dei test statistici, sono immediatamente trasformabili tramite l'inversione delle rispettive distribuzioni marginali (27) e (28), in campioni di variabili naturali. Un esempio in questo senso è riportato in figura 9 per la distribuzione bivariata rappresentata in figura 8.

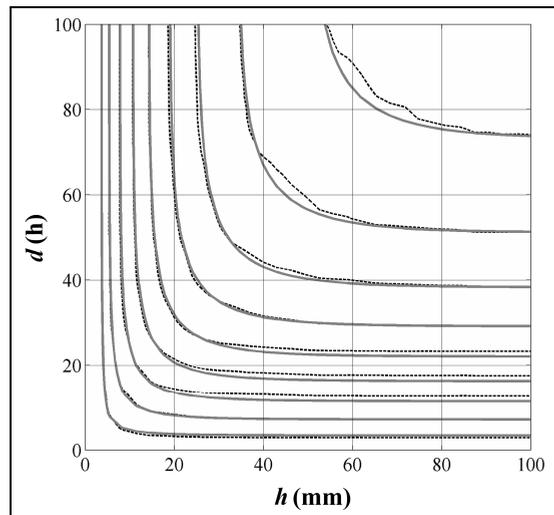


Fig. 8 - Adattamento della distribuzione congiunta P_{HD} alla distribuzione empirica nel caso in cui IETD 12 h ed IA 2 mm, confronto tra linee di livello (linea grigia continua, distribuzione teorica; linea nera tratteggiata, distribuzione empirica)

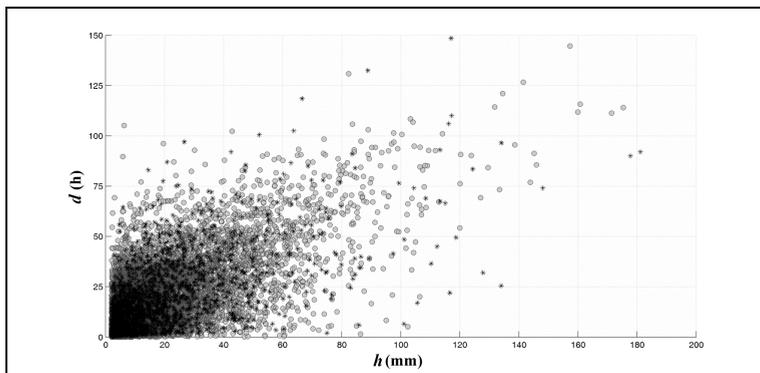


Fig. 9 - Confronto tra eventi di precipitazione campionari e generati attraverso la distribuzione P_{HD} nel caso in cui $IETD = 12$ h ed $IA = 2$ mm, eventi osservati 2281 (asterischi neri), eventi generati 10000 (punti grigi)

4. Derivazione della distribuzione delle portate al colmo

Una possibile applicazione pratica della distribuzione di probabilità congiunta P_{HD} consiste nella derivazione di distribuzioni di frequenza delle portate al colmo prodotte da un bacino di ridotte dimensioni. Le altezze e le durate di precipitazione generate mediante la funzione P_{HD} possono infatti essere trasformate, senza necessità di ragguaglio all'area, in idrogrammi di piena alla sezione di chiusura, una volta che sia stato scelto un opportuno modello di trasformazione afflussi-deflussi.

In questa memoria si propone un primo semplice esempio sviluppato mediante modelli semplificati sia per la depurazione della precipitazione lorda dalle perdite idrologiche, sia per la successiva convoluzione degli afflussi netti in portate di deflusso superficiale. Nel primo caso si è adottato un criterio basato su due parametri, come indicato nell'equazione (35), secondo il quale per il calcolo dell'altezza di precipitazione netta h_n si è tenuto conto della "initial abstraction" IA e di una perdita percentuale, per la parte eccedente, introducendo un coefficiente di afflusso Φ .

$$h_n = \Phi (h - IA) \quad (35)$$

Nel secondo caso, avendo presente lo schema suggerito da Wycoff & Singh (1976), si è assunto che l'altezza di precipitazione netta h_n coincida con il volume di un'onda di piena triangolare avente base data dalla somma della durata di precipitazione efficace d_n e del tempo di corrivazione del bacino t_c ed altezza pari alla portata al colmo q_p . La stima della durata di pioggia efficace d_n può essere dedotta immediatamente (36), se si ipotizza che l'intensità di pioggia lorda i sia costante; la forma triangolare dell'idrogramma consente poi di calcolare la portata al colmo q_p tramite la semplice espressione algebrica (37).

$$d_n = d - \frac{IA}{i} = d \frac{h - IA}{h} \quad (36)$$

$$q_p = \frac{2 h_n}{d_n + t_c} = \frac{2 \Phi (h - IA)}{d_n + t_c} \quad (37)$$

La procedura di generazione delle variabili permette quindi, attraverso questo semplice schema algebrico, di ricavare un campione di portate al colmo di piena, le quali possono essere associate al tempo di ritorno T grazie alla consueta espressione (10). In questa fase, si è tenuto conto del fatto che, per ogni anno di simulazione, sarà necessario generare un numero di eventi coerente con il numero medio annuo ω e la sua variabilità.

Il caso preso in esame consiste in un ipotetico bacino urbano avente superficie di area pari ad 100 ha e caratterizzato da una perdita iniziale IA di 3 mm e da un coefficiente di afflusso Φ di 0,45. Per un bacino di tali dimensioni si può assumere un tempo di corrivazione di t_c di 20 minuti, mentre il tempo di interevento IETD, da impiegare nella procedura di separazione degli eventi indipendenti assieme ad IA , può essere fissato in 3 h, valore minimo considerato ragionevole per applicazioni di drenaggio urbano (Adams & Papa, 2000).

Allo scopo di valutare l'affidabilità del modello proposto, sono state effettuate anche delle simulazioni continue impiegando la serie di precipitazioni osservate. Queste simulazioni sono state condotte adottando lo stesso modello di depurazione delle piogge, ma eseguendo la convoluzione degli afflussi netti tramite un idrogramma unitario istantaneo triangolare di base pari al tempo di corrivazione t_c . La serie di portate così ottenuta è stata poi separata in eventi indipendenti utilizzando ancora uno IETD di 3 h. I colmi dei singoli eventi di piena sono stati infine associati ai tempi di ritorno T ancora secondo l'espressione (10).

La figura 10 propone quindi il confronto tra le due metodologie, mettendo in evidenza una soddisfacente corrispondenza tra le portate ricavate a partire dalla distribuzione P_{HD} e trasformate attraverso il modello semplificato e quelle ricavate dalla convoluzione delle precipitazioni osservate. In particolare si può notare come la crescita meno che proporzionale di q_p rispetto al tempo di ritorno T sia concettualmente corretta e come i tempi di ritorno stimati siano sostanzialmente accettabili. Gli scostamenti maggiori sono peraltro ascrivibili ai tempi di ritorno maggiori, ovvero quelli prossimi all'estensione temporale della serie, per i quali è ragionevole attendersi un maggiore grado di incertezza.

Un'ultima considerazione merita sicuramente la stima dei tempi di ritorno delle portate q_p secondo le tre diverse formulazioni illustrate T^{OR} (11), T^{AND} (12) e T^K (15). Un primo problema di ordine concettuale risiede nel fatto che, secondo ciascuna di esse, una qualsiasi portata al colmo, potendo essere teoricamente prodotta da infinite combinazioni di altezze e durate di precipitazione, non è associabile ad un unico valore bensì ad un intervallo di tempi di ritorno. Dal punto di vista pratico operativo, ciò potrebbe essere un problema superabile se tale intervallo fosse piccolo ed i valori ragionevoli. Purtroppo in nessuna delle tre formulazioni illustrate queste condizioni sono soddisfatte.

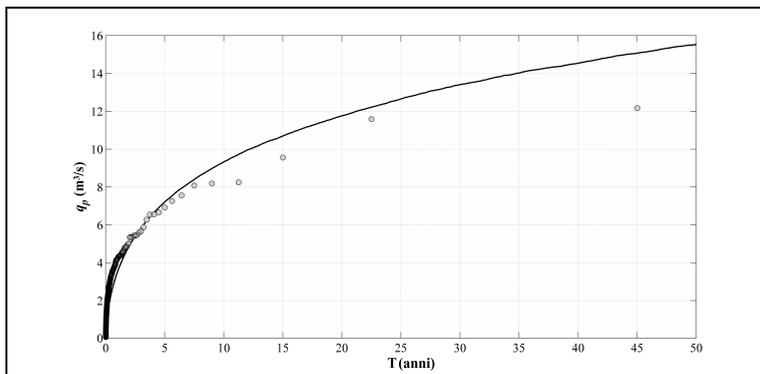


Fig. 10 – Confronto tra portate al colmo di piena generate a partire dalla distribuzione P_{HD} (linea continua) e quelle stimate tramite simulazione delle precipitazioni osservate (punti grigi)

Prendendo in esame, ad esempio, la portata al colmo di $9,33 \text{ m}^3/\text{s}$, associabile al tempo di ritorno di 10 anni secondo il grafico in figura 10, al variare di altezza e durata di precipitazione si ottengono tempi di ritorno T^{OR} e T^K abbastanza simili, inclusi in un intervallo piuttosto limitato, ma molto inferiori all'anno e, quindi, complessivamente inaccettabili. Nel caso della stima secondo la formulazione T^{AND} , il tempo di ritorno, anche se prevalentemente inferiori ai 2 anni, risulta estremamente variabile arrivando anche a valori superiori ai 100 anni. Nessuna delle tre formulazioni si dimostra pertanto idonea a stimare il tempo di ritorno di 10 anni ricavato dalle simulazioni.

5. Conclusioni

Nella presente memoria è stata illustrata una procedura di analisi statistica delle precipitazioni, di tipo multivariato, condotta tramite l'impiego di funzioni copula. Questo approccio ha consentito la costruzione di una distribuzione di probabilità congiunta delle altezze e delle durate di precipitazione basata su un modello probabilistico, ottenuto combinando la copula di Gumbel-Hougaard, per la struttura di dipendenza, e distribuzioni di Weibull, per le marginali. L'adattamento ai campioni estratti da serie continue di osservazioni, per svariate combinazioni dei parametri di separazione degli eventi indipendenti, è infatti risultato sempre molto soddisfacente.

La procedura adottata si basa, come suggerito da vari autori (Salvadori et al., 2007), sulla separazione tra le analisi univariate delle marginali, le quali sono condotte secondo consolidate metodologie tradizionali, e quella della struttura di dipendenza. Ciò non appare unicamente conveniente dal punto di vista delle procedure di stima delle funzioni e dei loro parametri, le quali risultano molto più rigorose e meno soggette a scelte arbitrarie, ma anche da quello dell'interpretazione delle proprietà statistiche delle variabili idrologiche.

La teoria delle funzioni copula rende inoltre piuttosto agevole l'implementazione delle distribuzioni multivariate nell'ambito di metodi di tipo Montecarlo. Grazie ad esse è infatti possibile delineare una procedura, relativamente semplice, in grado di generare lunghe serie di variabili idrologiche da impiegare, ad esempio, nella verifica o nel progetto di opere idrauliche. Purtroppo il problema dell'estensione del concetto di tempo di ritorno ai casi multivariati, nonostante alcune proposte recentemente avanzate, resta ancora irrisolto ed oggetto di ricerca. Il ricorso all'approccio univariato tramite derivazione, in via analitica o per simulazione numerica, di una variabile dipendente sembra per ora l'approccio più promettente.

Bibliografia

1. Adams B. J., Papa F. (2000) – *Urban stormwater management planning with analytical probabilistic models* – John Wiley, New York.
2. Ashkar F. (2007) – *Bivariate generalized-Pareto-based models for hydrological frequency analysis* – Proc. 18th Canadian Hydrotechnical Conf. and Symp. of the Canadian Society for Civil Engineering, Winnipeg Manitoba, Canada.
3. Bacchi B., Becciu G., Kottegoda N. T. (1994) – *Bivariate exponential model applied to intensities and durations of extreme rainfall* – J. Hydrol. 155 (1-2), pagg. 225-36.
4. Balistrocchi M., Bacchi B. (2011) – *Modelling the statistical dependence of rainfall event variables through copula functions* – Hydrol Earth Sys Science, 15(6), pagg. 1959-77.
5. Balistrocchi M., Grossi G., Bacchi B. (2013) – *Deriving a practical analytical-probabilistic method to size flood routing reservoirs* – Adv. Water Resour., 62, pagg. 37-46.
6. Chang C. H., Tung Y. K., Yang J. C. (1994) – *Monte Carlo simulations for correlated variables with marginal distributions* – J. Hydraul. Eng.-ASCE, 120(3), pagg. 313-31.
7. Correia F. N. (1987) – *Multivariate partial duration series in flood risk analysis* – In Hydrologic Frequency Modelling, edited by V. P. Singh, , Reidel, Dordrecht, The Netherlands, pagg. 541-54.
8. Deheuvels P. (1979) – *Empirical dependence function and properties: nonparametric test of independence* – Bulletin de la classe des sciences Academie Royale de Belgique, 65(6), pagg. 274-92.
9. Dupuis D. J. (2007) – *Using copulas in hydrology: benefits, cautions, and issues* – J. Hydrol. Eng., 12(4), pagg. 381-93.
10. Genest C., Ghoudi K., Rivest L. P. (1995) – *A semiparametric estimation procedure of dependence parameters in multivariate families of distributions* – Biometrika 82 (3), pagg. 543-52.
11. Genest C., Werker B. J. M. (2002) – *Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models* – In: Distribution with Given Marginals and Statistical Modelling. Kluwer, Dordrecht, The Netherlands, pagg. 103-12.
12. Genest C., Quessy J. F., Rémillard B. (2006) – *Goodness-of-fit procedures for copula models based on the probability integral transformation* – Scand J Stat, 32, 2, pagg. 337-66.
13. Genest C., Rémillard B., Beaudoin D. (2009) – *Goodness-of-fit tests for copulas: a review and a power study* – Insur Math Econ, 44, 2, pagg. 199-213.
14. Genest C., Nešlehová J., Quessy J. F. (2012) – *Test of symmetry for bivariate copulas* – Ann Inst Stat Math, 64, pagg. 811-34.

15. Goel N. K., Seth S. M., Chandra S. (1998) – *Multivariate modeling of flood flows* – J. Hydraul. Eng.-ASCE, 124(2), pagg. 146-55.
16. Gräler B., van den Berg M. J., Vandenberghe S., Petroselli A., Grimaldi S., De Baets B., Verhoest N. E. C. (2013) – *Multivariate return periods in hydrology: a critical review focusing on synthetic design hydrograph estimation* – Hydrol Earth Sys Science, 17(4), pagg. 1281–96.
17. Joe H. (1997) – *Multivariate models and dependence concepts* – Chapman and Hall, London.
18. Kendall M. G. (1938) – *A new measure of the rank correlation* – Biometrika, 30, pagg. 81-93.
19. Moisélo U. (1999) – *Idrologia tecnica* – La Goliardica Pavese, Pavia.
20. Nelsen R. B. (2006) – *An introduction to copulas, second ed.* – Springer, New York.
21. Sackl B., Bergmann H. (1987) – *A bivariate flood model and its applications* – In: Hydrologic Frequency Modelling, edited by Singh V. P., , Reidel, Dordrecht, The Netherlands, pagg. 571-82.
22. Salvadori G., De Michele C., Kottegoda N. T., Rosso R. (2007) – *Extremes in nature: an approach using copulas* – Springer, Dordrecht, The Netherlands.
23. Salvadori G., De Michele C., Kottegoda N. T., Rosso R. (2010) – *Multivariate multiparameter extreme value models and return periods: a copula approach* – Water Resour. Res. 46(10), W10501.
24. Sklar A. (1959) – *Fonctions de répartition à n dimensions et leurs marges* – Publ. Inst. Statist. Univ. Paris, 8, pagg. 229-31.
25. Wycoff R. L., Singh U. P. (1976) – *Preliminary hydrologic design of small flood detention reservoirs* – Water Resour. Bull., 12, pagg. 337-349.