Aggiornamento delle procedure di valutazione delle piene in Piemonte, con particolare riferimento ai bacini sottesi da invasi artificiali. Vol I: Dati idrologici e caratteristiche dei

Enel Produzione SpA

Politecnico di Torino
Dipartimento di Idraulica,
Trasporti e Infrastrutture Civili

Aggiornamento delle procedure di valutazione delle piene in Piemonte, con particolare riferimento ai bacini sottesi da invasi artificiali

VOLUME I

Costruzione e applicazione delle procedure di stima delle portate al colmo di piena

Pierluigi Claps, Francesco Laio

Settembre 2008

Enel Produzione SpA

Politecnico di Torino
Dipartimento di Idraulica,
Trasporti e Infrastrutture Civili

Aggiornamento delle procedure di valutazione delle piene in Piemonte, con particolare riferimento ai bacini sottesi da invasi artificiali

VOLUME I

Costruzione e applicazione delle procedure di stima delle portate al colmo di piena

Pierluigi Claps, Francesco Laio

Settembre 2008

CONTRATTO DI RICERCA N. 3000056255 - anno 2005 tra Enel Produzione S.p.a. e Politecnico di Torino, Dipartimento di Idraulica, Trasporti ed infrastrutture Civili (DITIC)

Ringraziamenti

Un lavoro durato due anni porta la firma di chi ne assume la paternità scientifica ma ne porta anche molte altre. La prima che sentiamo la necessità di aggiungere è quella di Giorgio Galeati, referente della committenza ENEL S.p.A. ma soprattutto sponda ideale di una effettiva collaborazione scientifica. La seconda, ma non in ordine di importanza, è quella di Marta Zanetta, che ha svolto, con precisione e grinta, quello che nel ciclismo è il ruolo del passista. Ad essi vanno aggiunti Elisa Bartolini, che si è occupata degli aspetti relativi alla valutazione degli estremi pluviometrici, e Daniele Ganora, che ha collaborato in alcune fasi delle analisi statistiche. Per motivazioni diverse, ma tutte buone, vogliamo ricordare qui i cortesi contributi di Paolo Villani, Secondo Barbero e, non da ultimo, le ispirazioni avute dal compianto Piero Telesca.
Vorremmo infine aggiungere, in modo forse inusuale, i nostri ringraziamenti alla dirigenza di ENEL Produzione S.p.A., per aver reso possibile una collaborazione realmente indirizzata a fini di ricerca scientifica.

Indice

Premessa

Il volume è diviso in tre parti. Dopo un capitolo introduttivo, che descrive la struttura generale del metodo indice, ha inizio la Parte I, dedicata alla stima della portata indice. In particolare, nel Capitolo 2 vengono descritti i metodi impiegati per la stima di questa grandezza e per la valutazione dell'incertezza di stima, e nel successivo Capitolo 3 tali metodi sono applicati alla regione di interesse. Nel Capitolo 4 vengono poi scelte le relazioni più efficaci per la stima della piena indice, che sono applicate nel Capitolo 5 a tutti i bacini di interesse.

La seconda parte del volume è dedicata alla stima della curva regionale di frequenza adimensionale $K_j(T)$, introdotta nel capitolo 1, che rappresenta il fattore di crescita del valore indice $Q_{ind,j}$ e misura la variabilità della portata di progetto $Q(T)$ con il tempo di ritorno T.

Nell'ultima parte si procede alla stima della piena di progetto tramite la relazione $Q_j(T) = K_j(T) \cdot Q_{ind,j}$, ovvero combinando le stime della piena indice (Parte I) e della curva di crescita (Parte II) ottenute in precedenza. Nel Capitolo 10 vengono pertanto descritti ed applicati i metodi di stima della piena di progetto, e si riportano in grafico le curve ottenute con i metodi di stima regionale, insieme ai dati osservati ed alle curve di confidenza. In chiusura di volume (capitolo 11) sono riassunti tutti i passi necessari per l'applicazione della procedura suggerita per la stima della piena di progetto.

1 Introduzione

1.1 L'analisi regionale di frequenza delle piene

Per la valutazione del rischio di piena sul territorio è importante poter far affidamento su informazioni che siano nello stesso tempo accurate e diffuse su tale territorio. Alla scala regionale le reti di monitoraggio forniscono misure puntuali delle grandezze idrologiche e climatiche, mentre si ha sovente l'esigenza di conoscere tali grandezze in un punto qualsiasi del territorio. Inoltre le serie storiche misurate sono spesso brevi al punto da rendere i campioni disponibili inadeguati ai fini dell'inferenza statistica, soprattutto nei casi in cui l'interesse è volto a tempi di ritorno *T* elevati, molto maggiori rispetto alla dimensione campionaria della serie.

Uno strumento con il quale si può ridurre l'incertezza di stima delle portate di piena *Q(T)* associate ad elevati periodi di ritorno *T* è l'analisi regionale o regionalizzazione delle portate di piena, basata sull'utilizzo dei dati disponibili attraverso le reti di monitoraggio ai fini della caratterizzazione di siti di particolare interesse ma non strumentati.

In generale i metodi di analisi regionale sono costituiti da diversi passaggi successivi, che in prima battuta prevedono l'analisi dei dati disponibili. Questo significa che per ogni *j*-esima stazione delle *s* strumentate prese in considerazione venga effettuato un attento esame delle relative osservazioni a disposizione, eliminando eventuali errori grossolani e accertandosi che la serie storica sia omogenea nel tempo. E' possibile inoltre procedere con l'integrazione della serie stessa qualora siano note ulteriori informazioni relative ad eventi eccezionali.

Per poter procedere nell'analisi regionale occorre quindi rendere confrontabili i valori misurati nelle diverse stazioni. A tal fine si procede dividendo, in ogni sito *j*, i valori di portata per un prefissato valore indice, $Q_{ind,j}$:

$$K_{i,j} = \frac{Q_{i,j}}{Q_{ind,j}}, \text{ con } i = 1,\ldots, n_j \text{ e } j = 1,\ldots, s \qquad (1.1)$$

dove n_j rappresenta il numero di osservazioni disponibili per la *j*-esima stazione in esame, ed *s* il numero totale delle stazioni. Tramite la (1.1) si ottengono *s* serie di portate ridotte, $K_{i,j}$.

Il modello più utilizzato nelle applicazioni dell'analisi di frequenza regionale è quello proposto da *Dalrymple* (1960), noto come "metodo della piena indice". In esso la portata al colmo di progetto in ciascuna sezione di misura $Q_j(T)$ risulta pari al prodotto tra la cosiddetta curva di crescita $K_j(T)$ e la piena indice $Q_{ind,j}$:

$$Q_j(T) = K_j(T) \cdot Q_{ind,j} \quad \text{con} \quad j = 1,\ldots,s \,. \tag{1.2}$$

Il fattore di scala $Q_{ind,j}$, anche detto "portata indice", è una grandezza locale, diversa per ciascun sito *j*; il fattore di crescita $K_j(T)$ misura invece la variabilità degli eventi estremi per i diversi tempi di ritorno, e risulta uguale per quei siti la cui distribuzione di frequenza è la stessa.

Nel presente volume verranno affrontate, in prima istanza, le modalità di stima della piena indice $Q_{ind,j}$ in bacini strumentati e non. Si confronteranno i risultati derivanti da una stima diretta effettuata sul campione di osservazioni disponibili, particolarmente indicata per le sezioni strumentate, e i risultati ottenuti tramite approcci basati su informazioni ausiliarie di tipo idrologico e fisico.

In secondo luogo verranno analizzate le modalità di stima della curva di crescita $K_j(T)$, la quale viene in genere espressa in forma analitica.

Infine si provvederà a fornire indicazioni utili alla scelta dell'approccio più adeguato alla stima della piena indice e della curva di crescita in relazione alla tipologia e alla quantità di informazioni idrometriche disponibili per il sito di interesse al fine di stabilire criteri oggettivi per la valutazione del rischio di piena in Piemonte e Valle d'Aosta.

1.2 La piena indice

L'ipotesi fondamentale del metodo di *Dalrymple* (1960) è che la distribuzione di probabilità della portata di progetto *Q(T)* in diversi siti appartenenti a una regione omogenea sia la stessa, a meno del parametro di scala Q_{ind}. Quest'ultimo viene detto "piena indice" e varia nella regione secondo le caratteristiche geomorfologiche e climatiche dei siti che la compongono.

Nei siti strumentati il valore della piena indice viene spesso calcolato utilizzando la media campionaria relativa al campione di n_j osservazioni disponibili. Nei casi di bacini non strumentati, o qualora non sia disponibile una serie consistente di misurazioni, è necessario ricorrere ad approcci che impieghino informazioni ausiliarie di tipo idrologico e fisico.

Questi metodi, che possono essere di tipo multiregressivo (v.es. *De Michele e Rosso*, 2001; *Bocchiola et al.*, 2003) o basati sull'impiego della formula razionale (v.es. *Furcolo et al.*, 1998; *Gioia et al.*, 2004), forniscono stime della piena indice espresse in funzione delle caratteristiche geomorfologiche e climatiche del bacino esaminato.

Considerando che i fattori fisici di interesse possono variare a seconda della zona esaminata, esistono numerose relazioni di letteratura volte alla stima del valore indice, l'applicazione delle quali conduce a risultati anche significativamente differenti. Un caso simile è stato riscontrato per i bacini dell'Italia Nord Occidentale, per i quali diversi metodi di analisi regionale, come il metodo VAPI (*Villani*, 2002) e il metodo GEV (*De Michele e Rosso*, 2001) conducono a stime della piena indice non sempre congruenti (*Claps et al.*, 2007).

La prima parte del presente volume descrive i risultati derivanti da una approfondita analisi condotta sui bacini dell'Italia Nord Occidentale con lo scopo di determinare le relazioni e gli approcci più adeguati per la valutazione della piena indice nei bacini non strumentati ricadenti nell'area di interesse. In particolare si volgerà l'attenzione ai bacini di media e alta quota, confrontando le stime ottenute nei diversi siti seguendo approcci differenti e si cercherà di stabilire criteri oggettivi per la scelta della relazione più appropriata da applicare nei diversi casi analizzati.

1.3 La curva di crescita

La curva regionale di frequenza adimensionale, Kj(T), rappresenta la curva di crescita del valore indice con il periodo di ritorno, e, quindi, misura la variabilità degli eventi estremi per i diversi T. Nelle procedure di analisi di frequenza regionale la determinazione delle curve di crescita viene effettuata adattando ai dati opportune distribuzioni di probabilità la cui forma si ritiene nota a meno di un numero finito p di parametri incogniti, $\theta_1,\ldots,\theta_p$. Ne consegue che K(T) può essere intesa come K(T; $\theta_1,\ldots,\theta_p$). I momenti campionari, quali media, varianza, asimmetria e kurtosis, vengono spesso utilizzati per la stima dei parametri delle distribuzioni di probabilità. Hosking e Wallis (1997) suggeriscono invece di utilizzare, al posto dei momenti ordinari, gli L-momenti o gli L-coefficienti perché più robusti nella stima da campioni poco consistenti di dati e perché meno soggetti a distorsioni nella stima (Appendice C). Nella seconda parte del presente volume vengono analizzate diverse modalità di stima degli L-momenti, distinguendo tra le stime effettuate direttamente a partire dal campione di osservazioni disponibili e le stime generate tramite metodi indiretti, come ad esempio l'approccio multiregressivo. Una volta presentati e applicati diversi metodi di stima, si definiranno criteri oggettivi per la definizione dell'approccio più adeguato in relazione alla quantità e al tipo di informazioni disponibili nel sito di interesse.

1.4 Costruzione del quadro conoscitivo (dati idro-geo-morfologici)

La necessità di costituire un'ampia base conoscitiva, fisica e idrologica, ai fini dello sviluppo di nuovi metodi di valutazione della piena indice e della curva di crescita, ha portato a costituire un database di tutti i bacini idrografici ricadenti nelle Regioni di Piemonte, Val d'Aosta, Lombardia occidentale ed Emilia appenninica occidentale per i quali siano disponibili osservazioni di portata massima annua al colmo di piena e giornaliera. Si è ritenuto utile includere nelle analisi anche i bacini svizzeri tributari del Ticino oltre che quelli del Rodano, i quali, trovandosi a ridosso delle Alpi, possono presentare un comportamento idrologico simile a quello dei bacini alpini piemontesi e valdostani.
Sono stati, dunque, selezionati 157 bacini idrografici, la cui gestione può competere, a seconda dei casi, a Enti diversi. In particolare si sono considerate:

- 68 stazioni che, in passato, sono state di competenza del Servizio Idrografico e Mareografico Nazionale (S.I.M.N.);
- 13 sezioni idrometriche gestite in passato da Enel;
- 4 che fanno riferimento ad A.E.M.;
- 44 che corrispondono a invasi gestiti da Enel;
- 28 che sono di competenza dell'Ufficio Federale per l'Ambiente Svizzero (U.F.A.M.).

Per approfondimenti sullo screening e sulla definizione delle serie storiche di riferimento si rimanda al Volume II relativo alla costituzione del database di dati idrologici, morfometrici e climatici per i bacini dell'Italia Nord Occidentale.
In aggiunta a quanto premesso riguardo ai dati idrologici, la definizione di parametri geomorfoclimatici caratteristici dei bacini risulta di fondamentale importanza per la stima della piena indice e della curva di crescita in siti non strumentati.
Nel presente lavoro sono state considerate tutte le caratteristiche dei bacini utili alla descrizione delle loro proprietà fisiche e idrologiche; in particolare sono stati considerati tutti quei parametri di bacino che ne descrivono:
la forma (fattore di forma, rapporto di circolarità, e di allungamento, etc.);
l'altimetria (quota media, curva ipsografica, etc.);

la lunghezza (lunghezza asta principale e del vettore orientamento, etc.);
la sollecitazione pluviometrica (curve di possibilità pluviometrica, etc.);
caratteristiche associabili a fattori di assorbimento dei suoli.

Per una disamina più completa dei metodi di valutazione dei parametri geomorfologici e climatici utilizzati si rimanda al volume II, mentre i valori calcolati di ciascun parametro sono riportati nel relativo Allegato I.

Parte I:

Stima della piena indice

2 Metodi di stima della piena indice

Nel presente capitolo si analizzano i metodi impiegati per la valutazione della piena indice nei bacini dell'Italia Nord Occidentale. In particolare si considera la possibilità di stimare tale grandezza a partire:

- dal valore empirico della media campionaria, eventualmente corretto per tenere conto di valori storici occasionali;
- da modelli di regressione;
- dall'applicazione della formula razionale;
- dalla stima derivata dagli estremi giornalieri.

Una volta stimato il valore della piena indice è importante indicare una misura della sua incertezza, in modo da conoscere la precisione con cui è stata effettuata la stima. Una statistica adeguata a tale scopo è la deviazione standard della stima, che verrà, quindi, opportunamente stimata per ciascun metodo.

2.1 Stima empirica basata sulle osservazioni storiche

2.1.1 Criteri di stima in presenza di valori storici occasionali

Considerati *s* siti strumentati, per i quali si dispone di una serie storica di portate al colmo più o meno consistente e relativa a un periodo sistematico di misurazioni, la piena indice in ciascuna sezione *j* viene stimata a partire da una statistica di scala, in genere rappresentata dalla media campionaria:

$$Q_{ind,j} = \frac{1}{n_j}\sum_{i=1}^{n_j} Q_i \qquad (2.1)$$

con $j = 1,\ldots,s$ ed n_j pari al numero di dati della stazione *j*-esima.

Nei casi in cui le serie storiche siano state integrate con dati relativi ad eventi alluvionali di particolare rilevanza occorsi in anni distanti dal periodo sistematico di misurazione, ovvero si decida di utilizzare valori occasionali significativi di portata (vedi Volume II, Allegato 3), lo stimatore della piena indice non coincide più con la media campionaria calcolata secondo la (2.1), in

quanto il campione statistico non è più costituito da osservazioni effettuate con sistematicità. E' questa una situazione abbastanza frequente nella quale una serie storica registrata inizialmente con continuità dal Servizio Idrografico viene successivamente integrata con sporadici valori di portata relativi ad eventi alluvionali più recenti od antecedenti l'inizio delle osservazioni sistematiche. I primi sono ad esempio riportati nei Rapporti di evento redatti dalla Direzione dei Servizi Tecnici di Prevenzione della Regione Piemonte (oggi ARPA Piemonte) a partire dal 1993. Una condizione ulteriore può essere riscontrata anche qualora eventi eccezionali riportati nelle Sezioni F degli Annali Idrologici non corrispondano al periodo di rilevazione continua.
Valutando la piena indice tramite la (2.1) su una stazione la cui serie sia stata integrata con dati occasionali si attribuirebbe lo stesso peso alle misure riferite al regime idrometrico medio del deflusso ed a quelle riferite agli eventi "eccezionali". Per tener conto con diverso peso della diversa natura degli eventi sporadici di intensità eccezionale si individua per essi un valore soglia Q_{soglia}, scelto pari al più piccolo dei valori relativi agli eventi occasionali considerati (Fig.2.1).

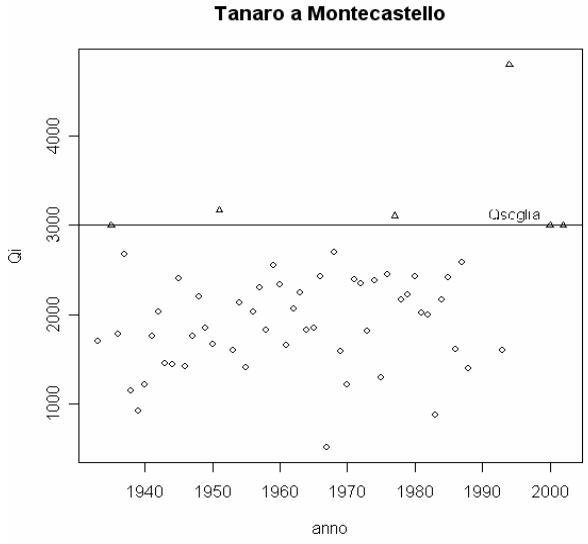


Figura 2.1. Esempio di serie storica integrata con valori occasionali significativi (triangoli). Il valore della portata soglia risulta pari a 3000 m^3/s.

In tali casi la serie storica è composta da n_{sotto_soglia} dati al di sotto della soglia prefissata e da n_{sopra_soglia} valori al di sopra di essa. Il calcolo della piena indice dovrà essere effettuato attribuendo un peso maggiore ai dati che rappresentano il regime idrometrico medio del deflusso e un peso inferiore

agli eventi occasionali significativi. Infatti nel caso in cui questi ultimi venissero trattati alla stessa stregua delle portate relative al regime ordinario il valore della piena indice potrebbe risultare distorto in modo rilevante.
La stima della piena indice viene dunque effettuata a partire dalla somma di due contributi, come definito da *Wang* (1990):

$$Q_{ind} = \frac{\sum_{i=1}^{n_{sotto\_soglia}} Q_i}{n} + \frac{\sum_{i=1}^{n_{sopra\_soglia}} Q_i}{n_{eq}} \tag{2.2}$$

in cui n rappresenta la numerosità della serie storica riferita al periodo sistematico di misurazione, mentre n_{eq} indica il periodo equivalente di osservazione, ossia la lunghezza complessiva del lasso temporale coperto dalla serie storica completa. Questo significa che il contributo dei valori occasionali significativi viene pesato non in base alla loro numerosità, bensì sul periodo totale di osservazione calcolato a partire dal primo anno di misure fino ad arrivare all'anno relativo all'ultimo evento registrato. Ad esempio nel caso del Tanaro a Montecastello, riportato in figura 2.1, il periodo equivalente di osservazione inizia nel 1926 e si conclude nel 2002, per un totale di 77 anni.

2.1.2 Valutazione dell'incertezza di stima

Una volta calcolato il valore della piena indice risulta importante valutare l'incertezza ad essa associata. A tal fine si ricorre alla determinazione della deviazione standard della portata indice, che, essendo la portata indice la media delle portate misurate, risulta esprimibile come:

$$\sigma_{Q_{ind,j}} = \frac{\sigma_{Q_j}}{\sqrt{n_j}}, \tag{2.3}$$

dove σ_{Q_j} rappresenta la deviazione standard della *j*-esima serie storica. Nel caso di campioni sistematici di misurazioni questa viene calcolata come:

$$\sigma_{Q_j} = \sqrt{\frac{1}{n_j} \cdot \sum_{i=1}^{n_j} \left[Q_i - \left(\frac{1}{n_j} \sum_{i=1}^{n_j} Q_i \right) \right]^2} \quad . \tag{2.4}$$

Qualora si stia considerando una serie integrata con informazioni di tipo storico, il calcolo della $\sigma_{Q_{ind,j}}$ non può essere effettuato tramite la (2.3), per quanto detto nel paragrafo precedente. In questo caso si fa ricorso a una formulazione analoga a quella indicata nel *Hydrology Subcommittee of the Advisory Committee on water data* (1982):

$$\sigma_{Q_{ind,j}} = \sqrt{\frac{\sum_{i=1}^{n_{sotto\_soglia,j}} \left(Q_i - Q_{ind,j} \right)^2}{n_j} + \frac{\sum_{i=1}^{n_{sopra\_soglia,j}} \left(Q_i - Q_{ind,j} \right)^2}{n_{eq,j}}} \quad , \tag{2.5}$$

dove il valore di $Q_{ind,j}$ viene calcolato tramite la (2.2).

E' importante sottolineare che la deviazione standard (2.4) viene determinata pesando la somma degli scarti $\left[Q_i - \left(\frac{1}{n_j} \sum_{i=1}^{n_j} Q_i \right) \right]^2$ su n_j e non su n_j-*1*, in modo da ottenere tramite la (2.3) valori di $\sigma_{Q_{ind,j}}$ congruenti con quelli definiti tramite la (2.5) per le serie storiche integrate con eventi occasionali significativi.

2.2 Stima statistica attraverso modelli di regressione multipla

2.2.1 Definizione dei modelli multiregressivi

I metodi multiregressivi (Appendice A) sono i più comunemente utilizzati per la stima della portata indice in siti sprovvisti di osservazioni. Nel caso in questione l'approccio multiregressivo viene utilizzato per legare la variabile piena indice alle caratteristiche del bacino, quali i parametri morfometrici e climatici descritti nel paragrafo 1.3.

Nel presente lavoro l'applicazione è avvenuta stimando $Q_{ind,j}$ mediante regressioni multiple con un numero variabile di regressori, sia nel campo lineare che in quello logaritmico.
In particolare è stato utilizzato un modello del tipo:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2.6}$$

dove:

- $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix}$

  è il vettore *(s x 1)* delle variabili dipendenti y_j di ciascuna stazione;

- $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$

  indica il vettore dei *(p+1 x 1)* coefficienti di regressione, stimati mediante la tecnica dei minimi quadrati ordinari descritta in Appendice A;

- $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1,p} \\ 1 & x_{21} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{s1} & \cdots & x_{s,p} \end{bmatrix}$

  rappresenta la matrice *(s x p+1)* delle variabili esplicative da cui viene fatta dipendere la piena indice;

- $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_s \end{bmatrix}$

  è il vettore *(s x 1)* dei residui della regressione.

La prima colonna di ***X*** è sempre costituita da un vettore unitario in modo che il primo dei coefficienti nel vettore ***β*** sia il termine noto della regressione.

Lo stimatore dei coefficienti di regressione secondo il metodo dei minimi quadrati ordinari è:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}\,, \tag{2.7}$$

mentre il corrispondente stimatore della variabile dipendente y_j in una stazione j è:

$$\hat{y}_j = \mathbf{a}_j \cdot \hat{\boldsymbol{\beta}} \tag{2.8}$$

- dove $\mathbf{a}_j = \left[x_1, x_2, \ldots, x_p\right]_j$ indica il vettore delle variabili esplicative della sezione in esame, e quindi rappresenta la j-esima riga di $\boldsymbol{X}$.

L'analisi multiregressiva è stata implementata secondo quattro diverse configurazioni:

- ponendo $\boldsymbol{y} = \boldsymbol{Q_{ind}}$ in relazione ad $\boldsymbol{X}$, matrice dei parametri geomorfologici dei bacini;
- ponendo $\boldsymbol{y} = \boldsymbol{Q_{ind}/A}$ (con A pari all'area del bacino) in relazione ad $\boldsymbol{X}$;
- ponendo $\boldsymbol{y} = ln(\boldsymbol{Q_{ind}})$ in relazione a $ln(\boldsymbol{X})$, in cui la prima colonna di $\boldsymbol{X}$ è sempre costituita da un vettore unitario;
- ponendo $\boldsymbol{y} = ln(\boldsymbol{Q_{ind}/A})$ in relazione a $ln(\boldsymbol{X})$, in cui la prima colonna di $\boldsymbol{X}$ è sempre costituita da un vettore unitario.

2.2.2 Criteri di scelta dei modelli di regressione

Ciascun modello multiregressivo consente di legare la variabile dipendente Q_{ind} a diverse caratteristiche morfologiche e climatiche dei bacini. Come accennato nel paragrafo 1.3, la disponibilità di informazioni sui descrittori di bacino è consistente. Per questo motivo è indispensabile individuare criteri che consentano di determinare quante e quali variabili esplicative debbano essere utilizzate per la stima della grandezza indice.

Per ciascuna delle quattro tipologie di regressione considerate si ottiene un elevato numero di combinazioni degli m parametri geomorfologici candidati alla formazione dei modelli. Occorre infatti considerare tutte le possibili regressioni con un numero di parametri variabile da $p = 1$ (regressione con un solo regressore) a $p = m$. Dal momento che si hanno disposizione $m = 40$ possibili variabili esplicative si ottiene un numero di possibili combinazioni dell'ordine di 10^{12} per ciascuna delle tipologie. Considerando che un buon

modello multiregressivo deve essere anche parsimonioso, si è ritenuto ragionevole estrarre esclusivamente le relazioni con non più di 4 variabili esplicative, riducendo le possibili combinazioni a 408200. Per ciascuna di queste si procede con la valutazione della sua capacità descrittiva, cercando di stabilire in quale misura ogni specifico modello multiregressivo riesce a spiegare la variabilità della piena media.

A tale scopo si calcola il coefficiente di determinazione corretto R^2_{adj} (Appendice A):

$$R^2_{adj} = 1 - \frac{\mathbf{y}^T\mathbf{y} - \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y} - \frac{1}{s}\left(\sum_{j=1}^{s} y_j\right)^2} \cdot \frac{(s-1)}{(s-p)} . \tag{2.9}$$

che consente di misurare, a parità di variabili esplicative coinvolte, la varianza spiegata dal modello.
La procedura descritta è stata ripetuta per ciascuna delle 408200 possibili combinazioni per il modello (2.6), in relazione a ciascuna delle 4 tipologie di regressione implementate.
Una volta valutata la capacità descrittiva per ciascuna delle 408200 regressioni si è ritenuto opportuno ridurre, per le successive analisi, il numero di relazioni da esaminare. In particolare si è deciso di proseguire le analisi considerando, per ciascun modello regressivo esaminato ed in relazione a ciascuna variabile dipendente esplorata, le migliori 5 relazioni in termini di R^2_{adj}. Di conseguenza, per ognuna delle 4 categorie di modelli considerati, si sono ottenute 20 configurazioni; il numero di modelli di regressione esaminati si riduce quindi da 408200 a 80.

2.2.3 Verifiche di adeguatezza del modello

Per ciascuno dei 20 modelli multiregressivi analizzati, relativi a ciascuna tipologia, si è verificato che tutte le variabili esplicative coinvolte fossero significative. A tale scopo si è utilizzato il test della *t* di Student, descritto in Appendice A, che sottopone a verifica separatamente la significatività di ciascun parametro coinvolto nella regressione attraverso la statistica (A.16). In particolare si è stabilito un livello di significatività α = *1%*, piuttosto basso

in ragione del fatto che si è ritenuto importante riscontrare la rilevanza, almeno tendenziale, di variabili di solito non utilizzate in questo tipo di analisi.
Successivamente, per i modelli che superano il test della *t* di Student, si è controllato che non ci fosse correlazione lineare tra i regressori (multicollinearità). La verifica dell'assenza di multicollinearità, spiegata in dettaglio nell'Appendice A, è volta a individuare i modelli stabili, ossia quelle regressioni ragionevolmente valide anche nel caso in cui cambiasse il campione di osservazioni a partire dal quale sono state tarate. La verifica viene effettuata calcolando il *variance inflation factor* (VIF), definito dalla (A.20); in particolare si assume che per valori di VIF > 5 l'ipotesi di assenza di multicollinearità non sia rispettata, nel qual caso il modello viene scartato.
A valle della verifica di adeguatezza il numero di modelli selezionati si può quindi ridurre rispetto agli 80 teoricamente considerati.

2.2.4 Valutazione dell'errore (capacità descrittiva del modello)

La valutazione della capacità descrittiva delle singole relazioni deve essere effettuata, in prima istanza, analizzando lo scostamento tra i valori stimati $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ e quelli osservati ***y***, utilizzati per tarare i modelli. In tal senso un primo indice di valutazione è rappresentato dal coefficiente di determinazione corretto R^2_{adj}, calcolato secondo l'Equazione (2.9). Ulteriori valutazioni possono essere effettuate a partire dall'analisi dei residui $\boldsymbol{\varepsilon} = \mathbf{y} - \hat{\mathbf{y}}$, calcolando tre differenti indici:

- errore quadratico medio:

$$RMSE = \sqrt{\frac{1}{s}\sum_{j=1}^{s}\left(y_j - \hat{y}_j\right)^2}\,, \tag{2.10}$$

  che, misurando lo scostamento quadratico tra i valori stimati e quelli osservati, attribuisce un peso maggiore agli errori più elevati;

- errore medio assoluto:

$$MAE = \frac{1}{s}\sum_{j=1}^{s}\left|y_j - \hat{y}_j\right|, \tag{2.11}$$

che esprime lo scostamento medio tra i valori stimati e quelli osservati senza considerare la loro direzione e attribuendo il medesimo peso a tutti gli scarti;

- errore medio assoluto percentuale:

$$MAPE = \frac{1}{s}\sum_{j=1}^{s}\left|\frac{y_j - \hat{y}_j}{y_j}\right|, \qquad (2.12)$$

che indica lo scostamento percentuale di stima riferito ai valori osservati.

E' importante sottolineare che il valore di RMSE è sempre maggiore, o al più uguale, rispetto all'errore medio assoluto. In particolare la differenza tra i due è tanto maggiore quanto più si riscontrano valori di residui elevati per singoli punti del campione. Al contrario, la condizione di coincidenza RMSE = MAE si verifica nel caso in cui tutti gli scostamenti assumono lo stesso valore.
I valori di RMSE, MAE e MAPE, essendo calcolati su tutti i valori stimati a partire dal campione di osservazioni, descrivono le proprietà del modello multiregressivo quando viene applicato nelle medesime stazioni usate in fase di taratura.
Per esplorare quale potrebbe essere il comportamento dei diversi modelli nel caso dei siti non strumentati è utile analizzare il RMSE, il MAE e il MAPE calcolati dopo una procedura di *cross validazione* (anche detta "jack-knife").
Tale procedura prevede che gli indici vengano calcolati a partire da un vettore stimato $\hat{\mathbf{y}}'$, il quale viene ottenuto da un modello regressivo che coinvolge le stesse variabili esplicative del modello in esame, ma con coefficienti di regressione $\boldsymbol{\beta}$ differenti da una sezione all'altra. In particolare i coefficienti vengono ristimati, ai fini della ricostruzione del valore in ciascuna sezione *j*, a partire da un vettore di osservazioni ***y'*** di dimensione *s-1*, privato della osservazione relativa alla stazione *j*. In questo modo la stima in ciascuna sezione, $\hat{y}_j'$ viene effettuata come se ci si riferisse ad un sito non strumentato, andando ad utilizzare esclusivamente le misure disponibili in altri siti. Gli indici di errore vengono dunque calcolati come:

$$RMSE_{CV} = \sqrt{\frac{1}{s}\sum_{j=1}^{s}\left(y_j - \hat{y}_j^{'}\right)^2}\ ; \tag{2.13}$$

$$MAE_{CV} = \frac{1}{s}\sum_{j=1}^{s}\left|y_j - \hat{y}_j^{'}\right|; \tag{2.14}$$

$$MAPE_{CV} = \frac{1}{s}\sum_{j=1}^{s}\left|\frac{y_j - \hat{y}_j^{'}}{y_j}\right|. \tag{2.15}$$

2.2.5 Diagrammi diagnostici

Per verificare se le assunzioni implicite nell'analisi multiregressiva, alla base del modello (2.6), vengono rispettate, è utile ricorrere ad alcuni diagrammi diagnostici. Questi consentono, ad esempio, di verificare che i residui ε abbiano media nulla, varianza costante (omoschedasticità) e che siano distribuiti normalmente.

Il grafico più immediato è quello che confronta gli andamenti della variabile dipendente osservata *y* e di quella stimata $\hat{y}$ (Fig.2.1.a), tramite uno scatter plot che consente di verificare visivamente quanto il modello regressivo riesca a rappresentare la realtà.

Tramite la rappresentazione di figura 2.1.a è inoltre possibile identificare le anomalie, rappresentate da punti che si discostano notevolmente dalla bisettrice, e che corrispondono a casi in cui il modello restituisce una stima notevolmente differente dalla misura.

Per valutare la normalità dei residui è invece utile ricorrere alla loro rappresentazione in carta probabilistica normale, la quale è definita in modo che la funzione di probabilità di Gauss sia rappresentabile con una retta su tale diagramma. Se $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$ sono i residui ordinati in senso crescente, la loro rappresentazione nei confronti della probabilità cumulata empirica $F_j = (j-1/2)/s$, $j = 1, 2, \ldots, s$, in carta probabilistica normale dovrebbe giacere approssimativamente su una linea retta quando i residui hanno una distribuzione normale (Fig.2.1.b). Deviazioni rilevanti dalla normalità dei residui implicano la mancata validità del test della *t* di Student, che è appunto basato su un'ipotesi di normalità dei residui.

La verifica grafica dell'omoschedasticità dei residui può essere effettuata mediante la costruzione di due grafici, in cui si riportano i residui in ordinata mentre in ascissa si indicano, rispettivamente, i valori osservati della variabile

dipendente e la radice della numerosità campionaria *n* (Fig.2.1.c, 2.1.d). L'ipotesi di omoschedasticità risulta verificata nel caso in cui i residui non tendano a disporsi a ventaglio al variare dell'ascissa. In caso contrario la varianza dei residui tende a crescere (o decrescere) con il valore di *y* o con $n^{0.5}$. L'eventuale violazione dell'ipotesi di omoschedasticità implica la non ottimalità degli stimatori ai minimi quadrati ordinari.

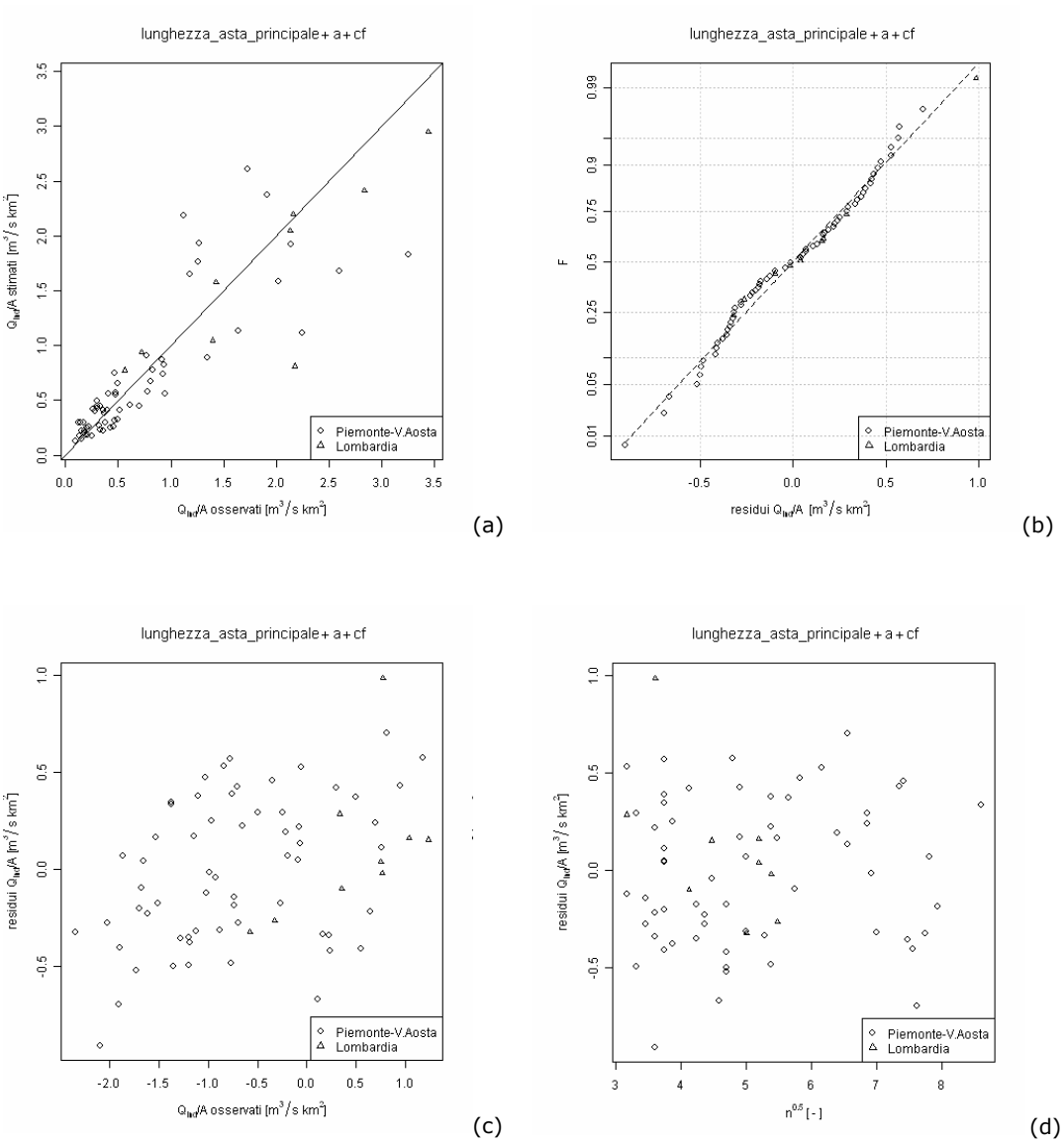


Figura 2.1. Esempio di diagrammi diagnostici utili a valutare l'affidabilità del modello regressivo. Nel caso specifico la variabile dipendente *y* è stata fissata pari a *Q/A*, mentre i descrittori impiegati nella regressione sono indicati nel titolo di ciascun grafico.

Dal confronto tra la variabile dipendente osservata e quella stimata, riportato in (a), emerge l'inadeguatezza del modello per i bacini caratterizzati da portate specifiche alte. Il diagramma (b) evidenzia la normalità dei residui della regressione, mentre (c) e (d) portano a riscontrare una lieve eteroschedasticità dei residui, caratterizzata da un andamento a ventaglio dei residui stessi, più evidente nel diagramma (c).

2.2.6 Modelli di regressione pesati

Tutte le considerazioni fin qui effettuate sono basate sull'ipotesi che la varianza dei residui $\hat{\mathbf{y}} - \mathbf{y}$ risulti circa costante per tutti i punti del campione. Come detto, l'ipotesi di omoschedasticità consente di stimare i coefficienti $\boldsymbol{\beta}$ di regressione tramite il metodo dei minimi quadrati ordinari (Ordinary Least Squares, O.L.S.) descritto in Appendice A. Tuttavia si deve sottolineare che, nel caso delle portate di piena, i valori di Q_{ind} sono a loro volta delle stime, la cui varianza cambia da un sito all'altro in ragione del numero di osservazioni n_j disponibili. In particolare, dove le serie storiche sono più consistenti la varianza è minore. Per questo motivo, una volta individuate le variabili esplicative per un certo modello multiregressivo, la valutazione dei coefficienti $\boldsymbol{\beta}$ dovrebbe essere effettuata dando un peso maggiore ai siti con più informazioni e attribuendo meno importanza alle sezioni con minore disponibilità di dati. Questo comporta che la stima dei coefficienti venga fatta ricorrendo al metodo dei minimi quadrati pesati (Weigthed Least Squares, W.L.S.), descritto in Appendice A.

Per ragioni analoghe anche nell'analisi dei residui si dovrà tenere in considerazione il peso $\sqrt{n_j}$ utilizzato in fase di stima; di conseguenza, i valori di RMSE, MAE e MAPE risultano:

$$RMSE = \sqrt{\frac{1}{s}\sum_{j=1}^{s}\left[\sqrt{n_j}\left(y_j - \hat{y}_j\right)\right]^2}\,, \tag{2.16}$$

$$MAE = \frac{1}{s}\sum_{j=1}^{s}\left|\sqrt{n_j}\left(y_j - \hat{y}_j\right)\right|, \tag{2.17}$$

$$MAPE = \frac{1}{s}\sum_{j=1}^{s}\left|\frac{\left(y_j - \hat{y}_j\right)}{y_j}\right|. \tag{2.18}$$

Nel caso in cui si voglia testare l'affidabilità dei modelli in seguito a una procedura di cross validazione, e quindi calcolare il valore di $RMSE_{CV}$, MAE_{CV}, $MAPE_{CV}$, si dovrà stimare un vettore $\hat{\mathbf{y}}'$, da sostituire a $\hat{\mathbf{y}}$ nelle (2.16), (2.17) e (2.18) applicando il metodo descritto nel paragrafo precedente.

2.2.7 Valutazione dell'incertezza di stima

Quando viene effettuata la stima di una variabile è bene associare ad essa una indicazione sull'incertezza di stima, in genere rappresentata dalla varianza. Nel caso dei modelli regressivi considerati nel presente paragrafo, l'obiettivo è quello di determinare la varianza di stima $\sigma^2_{\hat{y}_j}$ associata a ciascuna relazione e calcolata per ogni stimatore $\hat{y}_j$. Questo significa che, una volta tarata una regressione a partire da *s* misurazioni della variabile esaminata (una per ciascuna stazione), si potrà calcolare una incertezza di stima associata ad ogni punto in cui viene effettuato la stima.
La varianza di stima viene calcolata come:

$$\sigma^2_{\hat{y}_j} = \sigma^2_j \left[1 + \mathbf{a}_j \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{a}_j^T \right], \tag{2.19}$$

dove $\sigma^2_j = \dfrac{\sum_{j=1}^{s} \left(y_j - \hat{y}_j \right)^2}{s - p}$ indica la varianza dei residui, determinata tenendo in considerazione il numero *p* di parametri stimati.
Nel caso in cui il modello sia invece stato implementato nel campo logaritmico, ovvero ***y*** = *ln(**Q**_{**ind**}) o **y** = ln(**Q**_{**ind**}/**A**)* nella (2.6), la valutazione dell'incertezza di stima è più laboriosa. In particolare è necessario prima calcolare la varianza di *ln(**Q**_{**ind**}) o* di *ln(**Q**_{**ind**}/**A**)* tramite la (2.19), con $\boldsymbol{a_j}$ e ***X*** definiti in modo congruente con quelli definiti nel paragrafo 2.2.1, e considerando:

$$\sigma^2_j = \frac{\sum_{j=1}^{s} \left(\ln y_j - \ln \hat{y}_j \right)^2}{s - p}. \tag{2.20}$$

Per passare alla determinazione di $\sigma^2_{Q_{ind,j}}$ è necessario applicare le relazioni per il calcolo dei momenti relative alla distribuzione lognormale, ovvero:

$$\sigma^2_{Q_{ind,j}} = \mu^2_{Q_{ind,j}} \cdot \left[\exp\left(\sigma^2_{\ln(Q_{ind,j})} - 1 \right) \right], \tag{2.21}$$

dove la media $\mu_{Q_{ind,j}}$ della variabile stimata viene calcolata come

$$\mu_{Q_{ind,j}} = \exp\left[\mu_{\ln(Q_{ind,j})} + \frac{1}{2}\sigma^2_{\ln(Q_{ind,j})}\right] \tag{2.22}$$

in cui $\mu_{\ln(Q_{ind,j})} = \ln\left(Q_{ind,j}\right)$ e $\sigma^2_{\ln(Q_{ind,j})}$ si ottiene dalla (2.19).

2.3 Stima basata sull'impiego della formula razionale

2.3.1 Introduzione

La formula razionale è fondata sull'ipotesi che la portata di picco derivi da una pioggia netta uniforme sul bacino e caduta in una durata t_c, pari al tempo critico del bacino esaminato:

$$Q_{ind} = \frac{c_q \cdot A \cdot i(t_c)}{3.6} \tag{2.23}$$

dove:

- c_q rappresenta il coefficiente di afflusso di piena per la stazione in esame;
- A indica l'area del bacino;
- $i(t_c)$ è l'intensità di precipitazione, esprimibile come una funzione del valor medio a della pioggia estrema di 1 ora e del tempo critico: $i(t_c) = a \cdot t_c^{n-1}$, dove n è l'esponente della curva di possibilità pluviometrica.

A differenza dell'approccio regressivo, in questo caso le variabili esplicative sono fissate a priori. Tra queste, la durata critica non compare direttamente, e, come peraltro il c_q, per essa non è possibile fornire un valore oggettivo ed univoco. Si cercherà pertanto nel seguito di isolare nella formula razionale la componente di assorbimento rispetto agli effetti che le componenti fisiografiche del bacino (area, lunghezza asta principale, etc.) hanno sulla durata critica. A tal fine è utile riscrivere la (2.23) come:

$$Q_{ind} = K \cdot A \cdot a \cdot t_c^{n-1}, \tag{2.24}$$

dove $K = c_q / 3.6$.

2.3.2 Modalità di applicazione della formula razionale

La stima della piena indice a partire dalla formula razionale presuppone che sia noto il valore della durata t_c dell'evento meteorico che interessa il bacino. Tale valore non è conosciuto a priori, ma in letteratura esistono numerose relazioni per stimare t_c. In generale tale durata è assunta pari al tempo di corrivazione del bacino, stimato ad esempio tramite la formula di *Giandotti* (1934):

$$t_{c\_giand} = \frac{4\sqrt{A} + 1.5L}{0.8\sqrt{H_{media} - H_{\min}}}, \tag{2.25}$$

in cui L indica la lunghezza dell'asta principale, H_{media} e H_{min} indicano rispettivamente la quota media e minima del bacino.

Si è però visto (*Alfieri et al.*, 2008) che tale assunzione determina una sistematica sottostima del valore Q_{ind}, a differenza di quanto accade considerando t_c pari al tempo di ritardo del bacino (*Fiorentino et al.*, 1987b).
Nel caso specifico dell'uso che qui si vuole fare della relazione (2.23), e vista la particolare forma matematica della (2.24), si è ritenuto conveniente evitare di stimare t_c, per definire invece una relazione di proporzionalità tra il tempo critico t_c e le grandezze fisiografiche del bacino. In sostanza si può scrivere:

$$t_c = k_c \cdot f\left(A, L, \ldots\right), \tag{2.26}$$

dove k_c è una costante e $f\left(A, L, \ldots\right)$ una generica funzione delle grandezze fisiografiche.
Quindi, senza preoccuparsi di stimare k_c, dal momento che quest'ultima può essere assorbita nella costante K della (2.24), si possono esaminare diverse possibili espressioni del tempo critico (si veda l'Appendice B per la definizione dei simboli):

$$t_c \propto \sqrt{A}, \tag{2.27}$$

$$t_c \propto L, \tag{2.28}$$

$$t_c \propto LLDP, \tag{2.29}$$

$$t_c \propto \frac{LLDP}{\sqrt{s_{LLDP}}}. \tag{2.30}$$

Utilizzando le formule (2.27), (2.28), (2.29) e (2.30) in diversi casi è stato riscontrato che, in base ai valori delle statistiche di errore, la relazione più efficiente è la (2.27), in cui il tempo critico risulta proporzionale alla radice quadrata dell'area del bacino.
Sotto tale ipotesi il modello per la stima della piena indice tramite la formula razionale (2.24) diventa:

$$Q_{ind,j} = K_1 \cdot a_j \cdot A_j^{0.5+0.5n_j}. \tag{2.31}$$

2.3.3 Fattore di riduzione areale delle precipitazioni

Il calcolo della piena indice nel paragrafo precedente è stato effettuato ipotizzando che l'intensità di precipitazione sia uniformemente distribuita sull'intero bacino esaminato. Per tener conto della probabilità di simultaneità degli eventi si introduce nella (2.24) un fattore di riduzione areale (Areal Reduction Factor, A.R.F.), che, per definizione, assume valori inferiori a 1.
Nel presente lavoro il calcolo di tale fattore è stato effettuato secondo tre formulazioni differenti, per tentare di individuarne una rappresentativa:

1. $ARF_1 = 1 - \exp(-2.472 \cdot A^{-0.242} t_c^{0.6-\exp(-0.643 A^{0.235})})$, *Moisello e Papiri* (1986);
2. $ARF_2 = 1 - \mathrm{e}^{-1.1 t_c^{0.25}} + e^{\left(-1.1 t_c^{0.25} - 0.001 A\right)}$, *Eagleson* (1972);
3. $ARF_3 = \left[1 + 0.008 \cdot \left(\frac{A^{2.9}}{t_c}\right)^{0.211}\right]^{-2.7}$, *Barbero et al.* (2006).

Di queste, la terza è derivata dalla formulazione di *De Michele et al.* (2001), in seguito ad una specifica taratura sul territorio della Regione Piemonte. Nelle tre relazioni il tempo critico t_c è stato assunto pari al tempo di corrivazione calcolato tramite la formula di Giandotti.
Implementando le tre relazioni sopra è stato riscontrato che, in base alle statistiche di errore, la formula proposta da *Moisello e Papiri* (1986) restituisce i risultati più affidabili per i bacini considerati.

Di conseguenza la stima della piena indice secondo la formula razionale diventerebbe:

$$Q_{ind,j} = K_2 \cdot ARF_{1,j} \cdot a_j \cdot A_j^{0.5+0.5n_j} \,, \tag{2.32}$$

dove si usa il simbolo K_2 per tener conto dell'introduzione di un ulteriore fattore al secondo menbro dell'equazione.

2.3.4 Influenza della quota media

Il calcolo della piena indice è stato fin qui concepito ipotizzando che tutti i bacini rispondano in ugual modo ad una sollecitazione pluviometrica. Volendo invece distinguere il comportamento dei bacini che ricevono la precipitazione in forma liquida da quello dei bacini interessati da precipitazioni nevose, è necessario introdurre nella (2.24) un fattore correttivo che tenga conto della quota media del bacino H_{media}.
Tale fattore correttivo è assunto secondo una relazione logaritmica in base a considerazioni sull'effetto della quota media del bacino sulla riduzione dell'area contribuente del bacino stesso, come in *Allamano et al.* (2008):

$$C_{H_{med},j} = \ln\left(e - (e-1) \cdot \frac{H_{med,j}}{3500} \right), \tag{2.33}$$

il quale assume valori unitari quando $H_{media} = 0$ e decresce all'aumentare della quota media.
La stima della piena indice viene in questo caso effettuata come:

$$Q_{ind,j} = K_3 \cdot C_{H_{med},j} \cdot a_j \cdot A_j^{0.5+0.5n_j} \,. \tag{2.34}$$

2.3.5 Influenza dell'afflusso medio annuo

Un ulteriore elemento di approfondimento riguarda la possibilità di inserire nella (2.24), oltre al fattore $C_{H_{med},j}$, un coefficiente che consenta di tener conto, seppur in forma semplificata, dello stato di saturazione medio del suolo. E' infatti noto che il coefficiente di deflusso di piena tende ad aumentare con lo stato di saturazione del suolo (v.es. *Merz et al.*, 2006).

Quale variabile rappresentativa dello stato di imbibizione medio del suolo si è scelto di utilizzare l'afflusso medio annuo:

$$C_{aff,j} = aff_j \,. \tag{2.35}$$

Inserendo questo ulteriore elemento, peraltro non compreso tra 0 e 1 come i precedenti, la relazione di calcolo della portata indice risulta:

$$Q_{ind,j} = K_4 \cdot C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j} \,. \tag{2.36}$$

2.3.6 Relazione finale basata sulla formula razionale

Nei paragrafi precedenti si è visto come la stima della piena indice tramite la formula razionale possa essere effettuata inserendo in modo sequenziale nella (2.24) informazioni note a priori.
In particolare, per ciascuna sezione *j*, è possibile scrivere 4 diverse formule per il calcolo del coefficiente di afflusso:

$$K_{1,j} = \frac{Q_{ind,j}}{a_j \cdot A_j^{0.5+0.5n_j}} \,, \tag{2.37}$$

$$K_{2,j} = \frac{Q_{ind,j}}{ARF_{1,j} \cdot a_j \cdot A_j^{0.5+0.5n_j}} \,, \tag{2.38}$$

$$K_{3,j} = \frac{Q_{ind,j}}{C_{H_{med},j} \cdot a_j \cdot A_j^{0.5+0.5n_j}} \,, \tag{2.39}$$

$$K_{4,j} = \frac{Q_{ind,j}}{C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j}} \,. \tag{2.40}$$

In generale la stima della piena indice tramite la formula razionale può essere scritta come:

$$Q_{ind,j} = \hat{K}_z \cdot f_{z,j} \tag{2.41}$$

in cui $\hat{K}_z$ viene calcolato distintamente per ciascuna delle configurazioni (2.37), (2.38), (2.39), (2.40) (dunque risulta che *z = 1,…,4*).

Il valore di $\hat{K}_z$ per ogni configurazione può essere stimato come media aritmetica dei $K_{z,j}$ ottenuti per tutte le stazioni,

$$\hat{K}_{z,a} = \frac{\sum_{j=1}^{s} \hat{K}_{z,j}}{s} . \tag{2.42}$$

In alternativa si può usare una media geometrica degli stessi,

$$\hat{K}_{z,g} = \exp\left(\frac{\sum_{j=1}^{s} \ln\left(\hat{K}_{z,j}\right)}{s} \right) . \tag{2.43}$$

La variabile $f_{z,j}$ nella (2.41) è definita, a seconda dei casi, come:

$$f_{1,j} = a_j \cdot A_j^{0.5+0.5n_j} , \tag{2.44}$$

$$f_{2,j} = ARF_{1,j} \cdot a_j \cdot A_j^{0.5+0.5n_j} , \tag{2.45}$$

$$f_{3,j} = C_{H_{med},j} \cdot a_j \cdot A_j^{0.5+0.5n_j} , \tag{2.46}$$

$$f_{4,j} = C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j} . \tag{2.47}$$

2.3.7 Valutazione dell'incertezza di stima

La valutazione dell'incertezza di stima da associare alla formula razionale viene effettuata a partire dal prodotto del fattore $f_{z,j}^2$ e delle varianze di stima del coefficiente *K*:

$$\sigma_{Q_{ind},j}^2 = f_{z,j}^2 \cdot \sigma_{\hat{K}_z}^2 . \tag{2.48}$$

Nel caso in cui il fattore $\hat{K}_z$ sia stato stimato secondo la (2.42) la sua varianza di stima viene determinata a partire dalla varianza di stima del campione dei coefficienti di afflusso calcolati per le *s* stazioni:

$$\sigma^2_{\hat{K}_{z,a}} = \sigma^2_{\hat{K}_{zs}} \cdot \left(1+\frac{1}{s}\right), \tag{2.49}$$

in cui $\sigma_{\hat{K}_{zs}} = \sqrt{\frac{1}{\left(n_j - 1\right)} \cdot \sum_{i=1}^{n_j} \left[\hat{K}_{z,i} - \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \hat{K}_{z,i} \right) \right]^2}$.

Qualora $\hat{K}_z$ derivi da una media di tipo geometrico, come quella indicata nella (2.43), il procedimento per la valutazione dell'incertezza di stima porta a definire:

$$\sigma^2_{\hat{K}_z} = \mu^2_{\hat{K}_z} \cdot \exp\left(\sigma^2_{\ln\left(\hat{K}_z\right)} - 1\right) = \left[\exp\left(\mu_{\ln\left(\hat{K}_z\right)}\right)\right]^2 \cdot \exp\left(\sigma^2_{\ln\left(\hat{K}_z\right)} - 1\right), \tag{2.50}$$

dove:

- $\sigma^2_{\ln\left(\hat{K}_{z,g}\right)} = \sigma^2_{\ln\left(\hat{K}_{zs}\right)} \cdot \left(1+\frac{1}{s}\right)$ in cui per $\sigma^2_{\ln\left(\hat{K}_{zs}\right)}$ si intende la deviazione standard calcolata sul campione dei $\ln\left(K_{z,j}\right)$ definiti per ciascuna sezione;
- $\mu_{\ln\left(\hat{K}_{z,g}\right)} = \frac{1}{s}\sum_{j=1}^{s} \ln\left(\hat{K}_{z,j}\right)$.

2.4 Stima basata sull'impiego di dati di portata massima giornaliera

2.4.1 Introduzione metodologica

La valutazione della portata indice nei bacini strumentati, come si è visto nel paragrafo 2.1, viene effettuata a partire dalla serie storica delle portate al colmo disponibile per ciascuna sezione di interesse. Per alcuni bacini i dati di portata al colmo sono poco numerosi o assenti, mentre è noto un buon numero di osservazioni relative al massimo annuo della portata media giornaliera. Tenuto conto di tale disponibilità di informazioni, ci si pone l'obiettivo di ricavare una relazione che consenta di passare dalla media degli estremi giornalieri annui μ_g a quella delle portate al colmo Q_{ind}. La relazione può essere del tipo:

$$Q_{ind,j} = c_{P,j} \cdot \mu_{g,j}\,, \tag{2.51}$$

in cui c_P prende il nome di "coefficiente di punta".

A questo scopo, note le serie di $Q_{ind,j}$ e di $\mu_{g,j}$ relative a periodi di misurazione contemporanei per ciascuna *j*-esima stazione delle *s* a disposizione, si cerca di esprimere il coefficiente di punta in funzione delle caratteristiche climatiche e morfologiche dei diversi bacini.

A tal fine si ricorre ad una analisi multiregressiva analoga a quella presentata nel paragrafo 2.2.1, in cui la variabile dipendente ***y*** si identifica con c_P e le variabili indipendenti ***X*** sono rappresentate dall'insieme dei descrittori geomorfologici. In particolare vengono esplorate relazioni sia nel campo lineare che in quello logaritmico.

Il criterio di scelta dei migliori modelli multiregressivi individua le migliori 5 relazioni con al massimo 4 variabili esplicative, ordinate in termini di R^2_{adj}.

Una volta controllata l'adeguatezza dei diversi modelli tramite il test della *t* di Student e verificata l'ipotesi di non multicollinearità, si procede con la valutazione degli errori di stima mediante il calcolo di RMSE, MAE e MAPE sia prima che dopo una procedura di cross validazione, secondo quanto definito nel paragrafo 2.2.5.

2.4.2 Valutazione dell'incertezza di stima

Una volta definite diverse relazioni per la stima del coefficiente di punta c_P è possibile calcolare il valore della piena indice al colmo per qualsiasi stazione strumentata in cui sia nota una serie di portate giornaliere massime annue. L'incertezza di stima associata a ciascun valore calcolato di $Q_{ind,j}$ è funzione della media e della varianza calcolate sugli $n_{g,j}$ valori di portata giornaliera osservati nella stazione di interesse,

$$\mu_{g,j} = \frac{1}{n_{g,j}} \sum_{i}^{n_{g,j}} Q_{g,i} \,, \tag{2.52}$$

$$\sigma^2_{\mu_{g,j}} = \frac{\sum_{i=1}^{n_{g,j}} \left(Q_{g,i} - \mu_{g,j} \right)^2}{n_{g,j}}, \tag{2.53}$$

e della varianza associata alla stima del coefficiente di punta $\sigma^2_{c_P}$, calcolata tramite la (2.19).

L'incertezza di stima associata a ciascun valore calcolato di $Q_{ind,j}$ viene dunque calcolata come:

$$\sigma^2_{Q_{ind,j}} = \sigma^2_{c_P} \cdot \sigma^2_{\mu_{g,j}} + \sigma^2_{c_P} \cdot \mu^2_{\mu_{g,j}} + c^2_{P,j} \cdot \sigma^2_{\mu_{g,j}} \,. \tag{2.54}$$

3 Applicazione dei metodi nella regione di interesse

3.1 Stima empirica

La stima empirica della piena indice viene effettuata a partire dalle serie storiche di portata al colmo disponibili per le diverse stazioni, facendo riferimento alla media campionaria (2.1) per i casi in cui non sono stati impiegati dati dedotti da Rapporti di Evento o dalla Sezione F degli Annali Idrologici e ricorrendo al metodo di Wang espresso dalla (2.2) in caso contrario.

La valutazione dell'incertezza di stima da associare alla piena indice è stata effettuata determinando il valore della varianza $\sigma^2_{Q_{ind}}$, secondo il procedimento descritto nel paragrafo 2.2.

Le stime sono state effettuate su tutte le stazioni per le quali si hanno almeno 10 dati di portata al colmo, per un totale di 70 sezioni idrometriche, in modo da poter ritenere consistente la valutazione effettuata.

3.2 Stima con modelli di regressione multipla

Si considera un gruppo di 70 stazioni ricadenti nell'Italia Nord Occidentale aventi almeno 10 dati di portata al colmo sul quale tarare alcune formule multiregressive per la stima della piena indice. In particolare tali relazioni impiegheranno come possibili variabili esplicative 40 parametri geomorfologici, calcolati in precedenza per i bacini esaminati.

Si considera il modello di regressione multipla (2.6), in cui la variabile dipendente *y* viene fissata alternativamente pari a Q_{ind}/A, Q_{ind}, $ln(Q_{ind}/A)$ oppure $ln(Q_{ind})$.

Dall'insieme delle possibili combinazioni lineari tra i parametri considerati si estraggono, per ciascuna variabile dipendente esaminata, le 5 migliori soluzioni con non più di 4 variabili esplicative ordinate secondo il valore di R^2_{adj}, per un totale di 80 relazioni multiregressive. Di queste vengono considerate esclusivamente le regressioni i cui parametri risultano significativi per tutte le variabili indipendenti, ossia superano il test della *t* di Student con

livello di significatività α=1% e soddisfano l'ipotesi di non multicollinearità (Tab.3.1, Tab.3.2, Tab.3.3, Tab.3.4).

Per poter confrontare i risultati derivanti dall'impiego dei modelli multiregressivi indicati nelle tabelle 3.1, 3.2, 3.3, 3.4 si potrebbero analizzare i valori di R^2_{adj}; tuttavia tale approccio è sconsigliabile, trattandosi del confronto tra modelli sviluppati a partire da variabili dipendenti *y* diverse.

In questo caso è utile ricorrere all'analisi dei residui, e in particolare al calcolo delle statistiche di errore, quali il RMSE, MAE e MAPE. Affinché tali statistiche siano tra lori confrontabili è necessario che vengano calcolate in riferimento a una medesima grandezza; per questo motivo i valori di RMSE, MAE e MAPE indicati nelle tabelle sono stati calcolati tramite le relazioni (2.10), (2.11), (2.12), prima della cross validazione, e tramite le (2.13), (2.14), (2.15), dopo il jack-knife, ponendo sempre $y = Q_{ind}/A$ e $\hat{y} = \hat{Q}_{ind}/A$.

Variabile dipendente y = Q_{ind}/A		*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	R^2_{adj}	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
LLDP, sl_med1, c_f, aff	0.78	0.36	0.27	0.56	0.41	0.30	0.60
lun_asta_princ, sl_med1, c_f, aff	0.78	0.36	0.27	0.56	0.41	0.30	0.60
lun_vett_orient, sl_med1, c_f, aff	0.78	0.37	0.27	0.56	0.41	0.30	0.61
sl_med1, media_fa, c_f, aff	0.78	0.37	0.27	0.57	0.41	0.30	0.62
H_{media}, lun_vett_orient, aff	0.75	0.39	0.28	0.59	0.43	0.30	0.63
ΔH_2, cf, aff	0.74	0.40	0.30	0.63	0.43	0.33	0.67
lun_vett_orient, sl_med1, aff	0.74	0.40	0.28	0.54	0.43	0.30	0.58
H_{media}, LLDP, aff	0.74	0.40	0.29	0.61	0.43	0.31	0.64
H_{media}, lun_asta_princ, aff	0.74	0.40	0.29	0.61	0.43	0.31	0.64
lun_vett_orient, a	0.74	0.42	0.30	0.54	0.45	0.32	0.58
ΔH_2, aff	0.72	0.42	0.32	0.67	0.45	0.34	0.70
LLDP, a	0.71	0.43	0.31	0.55	0.45	0.32	0.57
lun_asta_princ, a	0.71	0.43	0.31	0.55	0.45	0.32	0.57
sl_med1, aff	0.71	0.43	0.29	0.49	0.46	0.31	0.51
a	0.65	0.47	0.32	0.52	0.52	0.37	0.67
aff 0.62 0.49 0.32 0.55 0.53 0.35 0.66							
X_{bar}	0.46	0.59	0.40	0.66	0.60	0.41	0.67
H_{media}	0.46	0.59	0.43	0.91	0.62	0.42	0.84
X_{sc}	0.45	0.59	0.40	0.69	0.61	0.42	0.73

Tabella 3.1. Valori di R^2_{adj} e delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente pari a Q_{ind}/A.

Variabile dipendente y = ln(Q_{ind}/A)		*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	R^2_{adj}	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
ΔH_2, CN, c_f, aff	0.85	0.36	0.23	0.29	0.40	0.25	0.31
lun_vett_orient, a, c_f, aff	0.84	0.37	0.23	0.30	0.41	0.25	0.33
pend_media_LDP, sl_med1, c_f, aff.	0.84	0.39	0.25	0.31	0.43	0.27	0.34
lun_asta_princ, sl_med1, c_f, aff	0.84	0.37	0.24	0.30	0.42	0.26	0.33
A, a, c_f, aff	0.84	0.37	0.24	0.31	0.40	0.26	0.33
lun_asta_princ, a, c_f	0.83	0.40	0.26	0.33	0.43	0.27	0.35
lun_vett_orient, a, c_f	0.83	0.41	0.26	0.33	0.44	0.27	0.35
ΔH_2, c_f, aff	0.82	0.39	0.25	0.32	0.43	0.27	0.34
A, a, c_f	0.82	0.40	0.25	0.33	0.43	0.27	0.35
LLDP, a, c_f	0.82	0.40	0.26	0.33	0.43	0.27	0.35
lun_vett_orient, a	0.78	0.41	0.26	0.37	0.42	0.27	0.38
lun_asta_princ, a	0.78	0.40	0.26	0.37	0.42	0.27	0.39
LLDP, a	0.78	0.40	0.26	0.37	0.42	0.27	0.39
A, a	0.78	0.40	0.26	0.37	0.41	0.27	0.39
media_fa, a	0.78	0.41	0.26	0.37	0.42	0.27	0.39
a	0.71	0.48	0.30	0.42	0.49	0.31	0.43
aff	0.67	0.52	0.33	0.45	0.53	0.34	0.46
X_{bar}	0.57	0.66	0.40	0.53	0.69	0.41	0.56
X_{sc}	0.56	0.65	0.40	0.54	0.68	0.41	0.57
LC_4	0.49	0.55	0.36	0.56	0.56	0.37	0.58

Tabella 3.2. Valori di R^2_{adj} e delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente pari a ln(Q_{ind}/A).

Variabile dipendente y = Q_{ind}		*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	R^2_{adj}	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
ΔH_1, LLDP, a	0.79	1.72	0.96	2.35	1.82	1.01	2.47
ΔH_1, lun_asta_princ, a	0.79	1.73	0.96	2.35	1.83	1.02	2.47
ΔH_1, media_fa, a	0.78	1.72	0.96	2.35	1.82	1.02	2.47
Y_bar, LLDP, a	0.78	1.06	0.69	1.52	1.12	0.73	1.60
LLDP, aff	0.75	1.07	0.70	1.54	1.13	0.73	1.60
lun_asta_princ, aff	0.75	1.08	0.71	1.55	1.15	0.74	1.61
media_fa, aff	0.75	1.10	0.72	1.61	1.17	0.75	1.67
media_fa, a	0.74	1.27	0.83	1.79	1.33	0.87	1.86
A, a	0.73	1.32	0.75	1.24	1.39	0.79	1.29
lun_asta_princ	0.62	0.69	0.55	1.33	1.60	0.99	2.51
media_fa	0.62	0.68	0.53	1.27	1.59	0.96	2.44
LLDP	0.62	0.71	0.56	1.37	1.54	0.94	2.36
lun_vett_orient	0.61	0.79	0.61	1.45	1.60	1.01	2.49
A	0.52	1.27	0.94	2.49	1.81	1.11	2.59

Tabella 3.3. Valori del coefficiente di determinazione corretto e delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente pari a Q_{ind}.

Variabile dipendente y = ln(Q_{ind})		*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	R^2_{adj}	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
magnitudine, c_f, aff, Hmed_A	0.94	0.38	0.24	0.30	0.42	0.26	0.33
A, a, c_f, aff	0.94	0.37	0.24	0.31	0.40	0.26	0.33
A, sl_med1, c_f, aff	0.94	0.37	0.24	0.30	0.41	0.26	0.33
magnitudine, a, c_f, aff	0.93	0.40	0.25	0.30	0.44	0.27	0.32
A, c_f, aff, LC_4	0.93	0.38	0.24	0.30	0.42	0.26	0.32
magnitudine, a, c_f	0.93	0.41	0.25	0.31	0.44	0.26	0.33
A, a, c_f	0.93	0.40	0.25	0.33	0.43	0.27	0.35
Y_lat, A, a	0.92	0.39	0.25	0.34	0.41	0.27	0.37
Y_bar, A, a	0.92	0.39	0.25	0.35	0.41	0.27	0.37
ΔH_1, aff, Hmed_A	0.92	0.38	0.26	0.36	0.41	0.28	0.38
magnitudine, a	0.91	0.42	0.26	0.37	0.43	0.27	0.38
A, a	0.91	0.40	0.26	0.37	0.41	0.27	0.39
aff, Hmed_A	0.89	0.47	0.30	0.40	0.50	0.31	0.41
media_fa, a	0.88	0.42	0.27	0.41	0.43	0.28	0.43
diam_topol, a	0.87	0.45	0.28	0.43	0.47	0.29	0.45
Hmed_A	0.82	0.67	0.40	0.54	0.70	0.42	0.56
sl_med2	0.67	1.35	0.63	0.75	1.44	0.66	0.78
media_fa	0.66	0.78	0.53	0.84	0.79	0.55	0.88
lun_asta_princ	0.65	0.78	0.54	0.87	0.79	0.56	0.91
LLDP	0.64	0.79	0.56	0.88	0.80	0.57	0.92

Tabella 3.4. Valori del coefficiente di determinazione corretto e delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente pari a ln(Q_{ind}).

Dall'analisi delle statistiche di errore emerge chiaramente che i modelli sviluppati ponendo $y = ln(Q_{ind}/A)$ oppure $y = ln(Q_{ind})$ restituiscono valori di RMSE, MAE e MAPE più bassi rispetto a quelli che si ottengono considerando una relazione lineare tra la variabile dipendente e la piena indice. In particolare si evidenzia la totale inadeguatezza dei modelli multiregressivi tarati a partire da $y = Q_{ind}$, che portano ad errori di stima estremamente elevati (Tab. 3.3). Ponendo $y = Q_{ind}/A$ si riscontra una riduzione delle statistiche di errore, che, tuttavia, evidenziano errori di stima quasi doppi rispetto a quelli che si riscontrano per $y = ln(Q_{ind}/A)$ oppure $y = ln(Q_{ind})$.
Concentrando l'attenzione sulle tabelle 3.2 e 3.4 si evidenzia come, al diminuire del numero di variabili esplicative coinvolte, il valore delle statistiche sui residui vada ad aumentare anche dell'ordine del 10%. Questo significa che la variabilità delle portate indice viene spiegata meglio usando un buon numero di descrittori. Alcuni tentativi con più di 4 regressori non hanno invece portato a sostanziali miglioramenti.

In figura 3.1 e 3.2 vengono riportati i diagrammi diagnostici relativi ad alcuni modelli multiregressivi tarati ponendo $y = ln(Q_{ind})$ o $y = (Q_{ind}/A)$.

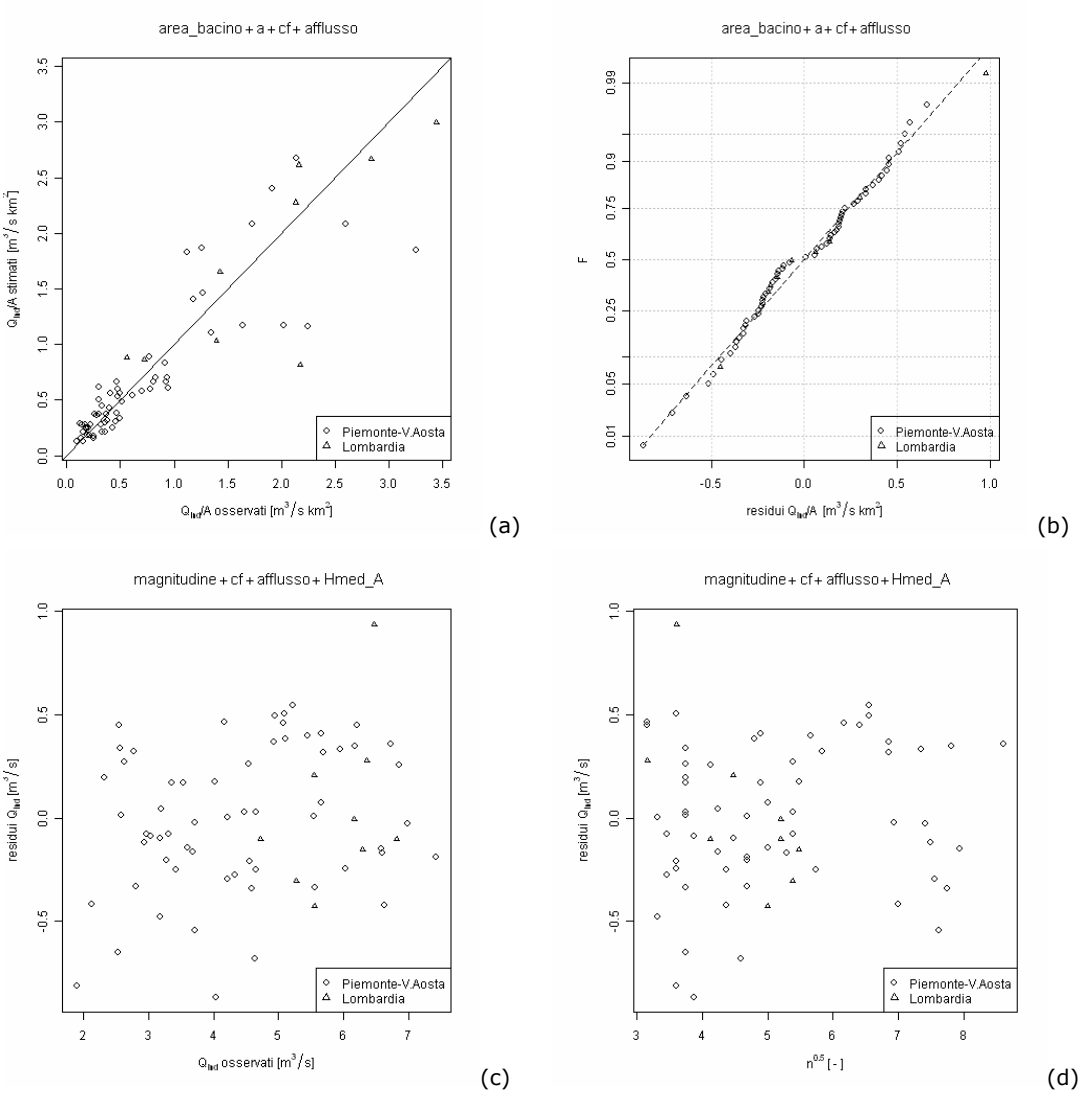


Figura 3.1. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata ponendo $y=ln(Q_{ind})$. In (a) si osserva il buon adattamento dei valori di Q_{ind} stimati a quelli osservati, mentre in (b) si verifica la normalità dei residui. Analizzando (c) e (d) si nota che i residui tendono a disporsi a ventaglio, evidenziando una lieve eteroschedasticità.

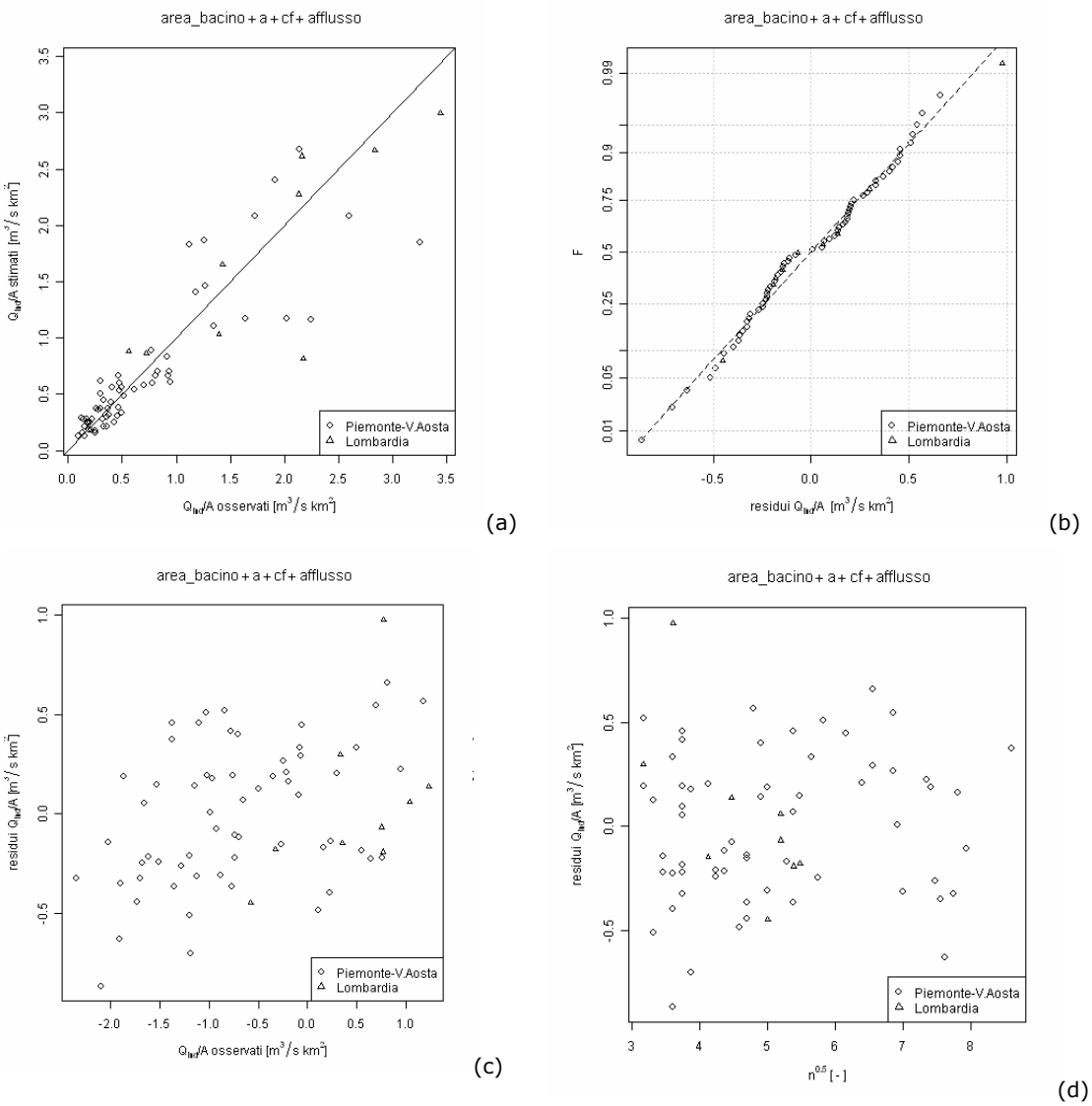


Figura 3.2. Diagrammi diagnostici per la quinta migliore relazione a 4 parametri tarata ponendo *y=ln(Q_{ind}/A)*. In (a) si osserva il buon adattamento dei valori di Q_{ind}/A stimati a quelli osservati per portate specifiche inferiori a 1.5; per valori maggiori si nota una tendenza alla sottostima dei valori osservati. In (b) si evidenzia la normalità dei residui, mentre in (c) e (d) si nota che gli stessi residui tendono a disporsi a ventaglio, facendo emergere un problema di lieve eteroschedasticità.

Dall'analisi degli scatter plot emerge il buon adattamento dei valori stimati a quelli osservati; in particolare la relazione a 4 variabili tarata per $y = ln(Q_{ind})$ porta a stimare valori di portata indice più vicini a quelli osservati (Fig.3.1.a e 3.2.a). Per entrambe le relazioni è garantita la normalità dei residui, come è possibile notare in figura 3.2.b e 3.3.b. L'apertura a ventaglio dei punti riportati in figura 3.1.c, 3.1.d, 3.2.c, 3.2.d porta a riscontrare un problema di

eteroschedasticità dei residui; tale andamento può essere giustificato dall'adozione della trasformazione logaritmica a causa della quale la reale rappresentazione risulta deformata. In questo caso infatti sono i logaritmi dei residui a dover avere varianza costante.
Per correggere l'eteroschedasticità si può attribuire un peso differente alle stazioni con più osservazioni, andando a stimare i coefficienti della regressione tramite i minimi quadrati pesati. Questo tentativo viene effettuato nel successivo paragrafo.

3.2.1 Stime multiregressive pesate

Per verificare la possibilità di correggere l'eteroschedasticità e, conseguentemente, migliorare le stime di Q_{ind}/A, si effettua il confronto tra i risultati ottenibili stimando i coefficienti di regressione con il metodo dei minimi quadrati ordinari e quelli relativi ai minimi quadrati pesati. In particolare si considerano le migliori regressioni che superano il test *t* di Student relative al caso $ln(Q_{ind}/A)$ riportate al paragrafo precedente (Tab.3.2). Per i minimi quadrati pesati si utilizza come peso la radice della numerosità campionaria n_j, secondo la procedura descritta nel paragrafo 2.2.5. Per effettuare tale confronto è necessario basarsi su statistiche di tipo oggettivo. A tal fine si calcolano per entrambi i casi i valori di RMSE e di MAE pesati (moltiplicando i residui per $n_j^{0.5}$), e il valore di MAPE, che essendo espresso in percentuale, non risente del peso come indicato nella relazione 2.18 (Tab.3.5, Tab.3.6).
Dal confronto tra le statistiche calcolate emerge che i due metodi portano a risultati sostanzialmente analoghi; inoltre il valore dei coefficienti di regressione risulta del tutto simile per i due metodi implementati. Per questi motivi si ritiene preferibile adottare il metodo dei minimi quadrati ordinari, complessivamente più robusto dell'altro.

Caso non pesato con y = ln(Q_{ind}/A)	*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	*RMSE*	*MAE*	*MAPE*	*$RMSE_{CV}$*	*MAE_{CV}*	*$MAPE_{CV}$*
ΔH_2, CN, c_f, aff	0.35	0.23	0.29	0.39	0.25	0.31
lung_vett_orient, a, c_f, aff	0.38	0.23	0.30	0.41	0.25	0.33
pen_media_LDP, sl_med1, c_f, aff	0.40	0.25	0.31	0.44	0.27	0.34
lun_asta_princ, sl_med1, c_f, aff	0.38	0.24	0.30	0.43	0.26	0.33
A, a, cf, aff	0.37	0.23	0.31	0.40	0.25	0.33
lun_asta_princ, a, c_f	0.40	0.25	0.33	0.43	0.27	0.35
lun_vett_orient, a, c_f	0.41	0.25	0.33	0.43	0.27	0.35
ΔH_2, c_f, aff	0.40	0.25	0.32	0.45	0.28	0.34
A, a, c_f	0.40	0.25	0.33	0.42	0.26	0.35
LLDP, a, c_f	0.40	0.25	0.33	0.43	0.27	0.35
lun_vett_orient, a	0.45	0.27	0.37	0.46	0.28	0.38
lun_asta_princ, a	0.45	0.27	0.37	0.46	0.28	0.39
LLDP, a	0.45	0.27	0.37	0.46	0.28	0.39
A, a	0.44	0.27	0.37	0.46	0.28	0.39
media_fa, a	0.45	0.27	0.37	0.46	0.28	0.39
a	0.53	0.31	0.42	0.54	0.32	0.43
aff	0.53	0.33	0.45	0.54	0.34	0.46
X_{bar}	0.70	0.41	0.53	0.73	0.42	0.56
X_{sc}	0.69	0.41	0.54	0.72	0.42	0.57
LC_4	0.62	0.37	0.56	0.63	0.38	0.58

Tabella 3.5. Valori delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità, tarati fissando la variabile dipendente pari a $ln(Q_{ind}/A)$ e con i coefficienti di regressione β stimati con i minimi quadrati ordinari. Il valore di RMSE, MAE e MAPE sono stati calcolati pesando i residui sulla radice della numerosità campionaria.

Caso pesato con y = ln(Q_{ind}/A)	*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	*RMSE*	*MAE*	*MAPE*	*$RMSE_{CV}$*	*MAE_{CV}*	*$MAPE_{CV}$*
ΔH_2, CN, c_f, aff	0.34	0.22	0.29	0.39	0.25	0.31
lung_vett_orient, a, c_f, aff	0.37	0.23	0.30	0.41	0.25	0.33
pen_media_LDP, sl_med1, c_f, aff	0.39	0.25	0.32	0.44	0.27	0.34
lun_asta_princ, sl_med1, c_f, aff	0.36	0.23	0.31	0.43	0.26	0.33
A, a, cf, aff	0.35	0.23	0.31	0.40	0.25	0.33
lun_asta_princ, a, c_f	0.39	0.25	0.33	0.43	0.27	0.35
lun_vett_orient, a, c_f	0.41	0.25	0.33	0.43	0.27	0.35
ΔH_2, c_f, aff	0.39	0.25	0.32	0.45	0.28	0.34
A, a, c_f	0.40	0.25	0.33	0.42	0.26	0.35
LLDP, a, c_f	0.39	0.25	0.33	0.43	0.27	0.35
lun_vett_orient, a	0.45	0.28	0.38	0.46	0.28	0.38
lun_asta_princ, a	0.44	0.27	0.38	0.46	0.28	0.39
LLDP, a	0.44	0.27	0.38	0.46	0.28	0.39
A, a	0.44	0.27	0.38	0.46	0.28	0.39
media_fa, a	0.45	0.28	0.38	0.46	0.28	0.39
a	0.52	0.31	0.41	0.54	0.32	0.43
aff	0.54	0.33	0.45	0.54	0.34	0.46
X_{bar}	0.72	0.41	0.54	0.73	0.42	0.56
X_{sc}	0.71	0.41	0.54	0.72	0.42	0.57
LC_4	0.62	0.37	0.55	0.63	0.38	0.58

Tabella 3.6. Valori delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente pari a $ln(Q_{ind}/A)$ e con i coefficienti di regressione β stimati con i minimi quadrati pesati sulla numerosità campionaria. Il valore di RMSE, MAE e MAPE sono stati calcolati pesando i residui sulla radice della numerosità campionaria.

3.3 Stima basata sulla formula razionale

Si considera lo stesso gruppo di 70 stazioni ricadenti nell'Italia Nord Occidentale esaminato nel paragrafo precedente e si stima la piena indice a partire dalla (2.41), considerando un coefficiente di afflusso $\hat{K}_z$ stimato a partire da ragionamenti progressivi fatti sulla formula razionale, come indicato nel paragrafo 2.3.6.

Per confrontare i risultati ottenuti tramite le diverse formulazioni implementate, corrispondenti all'indice *z* variabile da 1 a 4, si è ritenuto utile confrontare i valori di RMSE, MAE e MAPE calcolati fissando $y=Q_{ind}/A$. In particolare si analizzano le statistiche di errore calcolate sia prima che dopo l'applicazione di una procedura di cross validazione (Tab.3.7 e Tab.3.8).

Descrizione	*z*	Media aritmetica				Media geometrica			
		$\hat{K}_{z,a}$	*RMSE*	*MAE*	*MAPE*	$\hat{K}_{z,g}$	*RMSE*	*MAE*	*MAPE*
F.razionale	1	0.1355	0.59	0.43	0.80	0.1105	0.66	0.44	0.65
Corr. ARF_1	2	0.1478	0.59	0.42	0.78	0.1211	0.66	0.44	0.64
Corr. H_{media}	3	0.2151	0.44	0.30	0.51	0.1928	0.49	0.32	0.46
Corr. H_{media}, aff	4	$1.65\ 10^{-4}$	0.35	0.23	0.39	$1.52\ 10^{-4}$	0.35	0.23	0.37

Tabella 3.7. Confronto tra i valori di RMSE, MAE e MAPE calcolati prima della cross validazione tramite le diverse relazioni riconducibili alla formula razionale. Le statistiche di errore vengono calcolate determinando i coefficienti di afflusso sia a partire da una media aritmetica ($\hat{K}_{z,a}$) che a partire da una media geometrica ($\hat{K}_{z,g}$).

Descrizione	*z*	$\hat{K}_{z,a}$			$\hat{K}_{z,g}$		
		$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
F.razionale	1	0.59	0.43	0.81	0.66	0.45	0.66
Corr. ARF_1	2	0.60	0.42	0.79	0.67	0.45	0.65
Corr. H_{media}	3	0.44	0.30	0.51	0.49	0.32	0.47
Corr. H_{media}, aff	4	0.35	0.23	0.40	0.35	0.23	0.38

Tabella 3.8. Confronto tra i valori di RMSE, MAE e MAPE calcolati dopo la cross validazione tramite le diverse relazioni riconducibili alla formula razionale. Le statistiche di errore vengono calcolate determinando i coefficienti di afflusso sia a partire da una media aritmetica ($\hat{K}_{z,a}$) che a partire da una media geometrica ($\hat{K}_{z,g}$).

Analizzando i risultati sopra riportati emerge che le stime effettuate usando la media geometrica di $\hat{K}_z$, conducono a valori di MAPE decisamente inferiori rispetto al caso in cui $\hat{K}_z$ venga determinato a partire dalla media aritmetica espressa dalla relazione (2.44). Per quanto riguarda il RMSE e il MAE si riscontra invece la tendenza opposta; tuttavia si deve sottolineare che la media aritmetica conduce a valori di RMSE e MAE di poco inferiori rispetto a quella geometrica e, dunque, si ritiene la stima tramite media geometrica più efficiente di quella tramite media aritmetica.

E' inoltre interessante notare come le correzioni introdotte nella formula razionale conducano a un effettivo miglioramento della stima rispetto al caso della (2.31). Si noti, tuttavia, come l'introduzione del fattore di riduzione areale secondo la (2.32) non apporti un significativo contributo nello spiegare la variabilità della piena indice. Al contrario, la correzione della formula

razionale mediante la funzione della quota media, definita dalla (2.34), porta a ridurre decisamente gli errori di stima, passando da una MAPE di 0.66 a 0.47 dopo la cross validazione. L'introduzione dell'afflusso medio contribuisce a spiegare ulteriormente la variabilità della piena indice, e conduce, dopo la cross validazione, a un errore medio assoluto di 0.38, quasi dimezzando il MAPE che si aveva inizialmente applicando la (2.31).

3.4 Stima basata sui dati di portata massima giornaliera

Si è considerato un gruppo di 78 stazioni ricadenti nell'Italia Nord Occidentale e nel Sud della Svizzera aventi almeno 10 dati contemporanei di portata al colmo e di estremi giornalieri annui. Per ciascuna sezione *j* si è proceduto al calcolo del coefficiente di punta, come:

$$c_{P,j} = \frac{Q_{ind,j}}{\mu_{g,j}}. \qquad (3.1)$$

Noti i valori di c_P per le 78 stazioni esaminate si è effettuata una analisi multiregressiva, sia nel campo lineare che in quello logaritmico, con lo scopo di determinare quali parametri geomorfologici e climatici siano efficaci nel ricostruire tale coefficiente.
E' importante sottolineare che oltre ai parametri geomorfologici considerati nel paragrafo 3.2 sono stati qui introdotti altri 4 descrittori: la media μ_{Qg} e la deviazione standard σ_{Qg} degli estremi giornalieri annui, oltre che i corrispondenti contributi specifici areali $\mu_{qg}=\mu_{Qg}/A$ e $\sigma_{qg}=\sigma_{Qg}/A$.

Dall'insieme delle possibili combinazioni lineari tra i descrittori considerati si sono estratte le 5 migliori soluzioni con non più di 4 variabili esplicative, ordinate secondo il valore di R^2_{adj}. Di queste sono state considerate esclusivamente le regressioni i cui parametri spiegano significativamente la variabilità della variabile dipendente c_P, ossia superano il test *t* di Student con livello di significatività α=1% e soddisfano l'ipotesi di non multicollinearità. In tabella 3.9 vengono indicati i modelli multiregressivi che rispondono a tali requisiti, accompagnati dai valori di RMSE, MAE e MAPE calcolati dopo una procedura di cross validazione. In tabella 3.10 vengono invece riportati i valori

delle statistiche di errore ottenuti dopo l'applicazione della procedura di jack-knife.

Analizzando i risultati ottenuti è possibile notare come tutte le relazioni tarate nel campo logaritmico presentino valori di RMSE, MAE e MAPE inferiori, garantendo quindi maggiore affidabilità nella stima del coefficiente di punta.
Si sottolinea inoltre come tra i descrittori significativi ci siano quasi sempre la deviazione standard delle portate giornaliere rapportata all'area σ_{qg} e un parametro di quota, quale ΔH_2 o ΔH_1.

Variabile dipendente y = c_P	*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
μ_{Qg}, σ_{qg}, X_{sc}, sl_med1	0.26	0.20	0.11	0.28	0.22	0.12
μ_{Qg}, σ_{qg}, $X_{bar,}$,sl_med1	0.26	0.20	0.11	0.28	0.22	0.12
σ_{Qg}, σ_{qg} ,X_{lon}, sl_med1	0.26	0.20	0.11	0.28	0.22	0.12
σ_{Qg}, σ_{qg} ,X_{bar}, sl_med1	0.26	0.20	0.11	0.29	0.22	0.12
μ_{Qg}, σ_{qg}, X_{lon}, H_{media}	0.26	0.20	0.11	0.29	0.22	0.12
σ_{Qg}, σ_{qg} , H_{media}	0.28	0.22	0.12	0.30	0.23	0.12
σ_{qg}, X_{bar}, sl_med1	0.28	0.21	0.11	0.31	0.23	0.13
σ_{qg},H_{media}, lun_asta_princ	0.28	0.22	0.12	0.31	0.23	0.12
σ_{qg}, H_{media}, LLDP	0.28	0.22	0.12	0.31	0.23	0.13
σ_{qg}, H_{media}, media_fa	0.28	0.21	0.12	0.31	0.23	0.13
σ_{qg}, sl_med1	0.30	0.23	0.13	0.23	0.24	0.13
σ_{qg}, H_{media}	0.31	0.24	0.13	0.33	0.25	0.13
σ_{qg}, a	0.31	0.23	0.12	0.34	0.24	0.14
σ_{qg}, X_{bar}	0.32	0.23	0.12	0.34	0.24	0.13
σ_{qg}, ΔH_2	0.32	0.25	0.14	0.34	0.26	0.13
σ_{qg}	0.36	0.26	0.14	0.38	0.28	0.14
H_{media}	0.37	0.28	0.15	0.38	0.28	0.15
a	0.38	0.27	0.14	0.40	0.28	0.15
μ_{qg}	0.39	0.27	0.14	0.39	0.28	0.15
X_{bar}	0.42	0.30	0.16	0.43	0.31	0.16

Tabella 3.9. Valore delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente *y* pari a c_P.

Variabile dipendente y = ln(c_P)	*Prima del jack-knife*			*Dopo il jack-knife*		
Descrittori	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
σ_{qg}, ΔH_2, lun_med_vers, a	0.26	0.18	0.10	0.28	0.20	0.10
σ_{qg}, a, ΔH_2	0.27	0.20	0.10	0.29	0.21	0.11
σ_{qg}, X_{sc}, ΔH_2	0.27	0.20	0.10	0.29	0.21	0.11
σ_{qg}, X_{bar}, ΔH_2	0.27	0.20	0.10	0.29	0.21	0.11
σ_{qg}, X_{sc}, ΔH_1	0.27	0.20	0.11	0.29	0.21	0.11
σ_{qg}, X_{bar}, ΔH_1	0.27	0.21	0.11	0.29	0.21	0.11
σ_{qg} , ΔH_2	0.29	0.22	0.11	0.30	0.22	0.12
σ_{qg} , H_{media}	0.30	0.22	0.12	0.32	0.23	0.12
σ_{qg} , sl_med1	0.30	0.22	0.12	0.31	0.23	0.12
σ_{qg}, ΔH_1	0.30	0.22	0.12	0.31	0.23	0.12
μ_{qg}, H_{media}	0.31	0.25	0.12	0.34	0.23	0.12
σ_{qg}	0.36	0.26	0.13	0.37	0.26	0.14
μ_{qg}	0.38	0.26	0.14	0.39	0.27	0.14
H_{media}	0.37	0.27	0.14	0.39	0.28	0.15
a	0.39	0.27	0.14	0.40	0.28	0.14
ΔH_2	0.38	0.30	0.16	0.39	0.31	0.16

Tabella 3.10. Valore delle statistiche di errore calcolate sia prima che dopo una procedura di cross validazione per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati fissando la variabile dipendente *y* pari a *ln(c_P).*

4 Esame comparativo dei risultati dei modelli di stima

4.1 Scelta del metodo per la stima della piena indice

In questo capitolo vengono riassunti i risultati ottenuti con i tre approcci proposti per la stima della piena indice. Vengono indicate le configurazioni più efficienti di ciascun modello, e, successivamente, vengono confrontate con i metodi di stima esistenti in letteratura.

4.1.1 Modello di regressione multipla

Alla luce delle statistiche di errore calcolate per i modelli multiregressivi tarati fissando la variabile dipendente *y* pari a Q_{ind}, Q_{ind}/A, $ln(Q_{ind})$, $ln(Q_{ind}/A)$, riportate nel paragrafo 3.2, la relazione più appropriata per la stima della piena indice risulterebbe:

$$\frac{Q_{ind}}{A} = 0.5926 \cdot \Delta H_2^{-0.71} \cdot CN^{-1.92} \cdot c_f^{0.67} \cdot aff^{1.90} \qquad (4.1)$$

corrispondente alla migliore relazione a 4 parametri ottenuta per $y=ln(Q_{ind}/A)$. Tuttavia tale relazione non risulta fisicamente congruente, poiché presenta un valore negativo per l'esponente di CN. Quest'ultimo rappresenta un indice di permeabilità del suolo, che assume valori vicini a 100 quando il terreno è completamente impermeabile e tendenti a zero per terreni molto permeabili. Poiché il deflusso aumenta al crescere dell'impermeabilità del suolo, il valore di Q_{ind} può essere legato a CN esclusivamente tramite un esponente positivo; di conseguenza la (4.1) non è stata considerata utilizzabile.
La seconda migliore relazione ottenuta per la stima della piena indice è allora:

$$Q_{ind} = 1.5256 \cdot 10^{-4} \cdot A^{0.81} \cdot a^{1.10} \cdot c_f^{0.78} \cdot aff^{0.95} \qquad (4.2)$$

per la quale valgono i diagrammi diagnostici di figura 4.1.

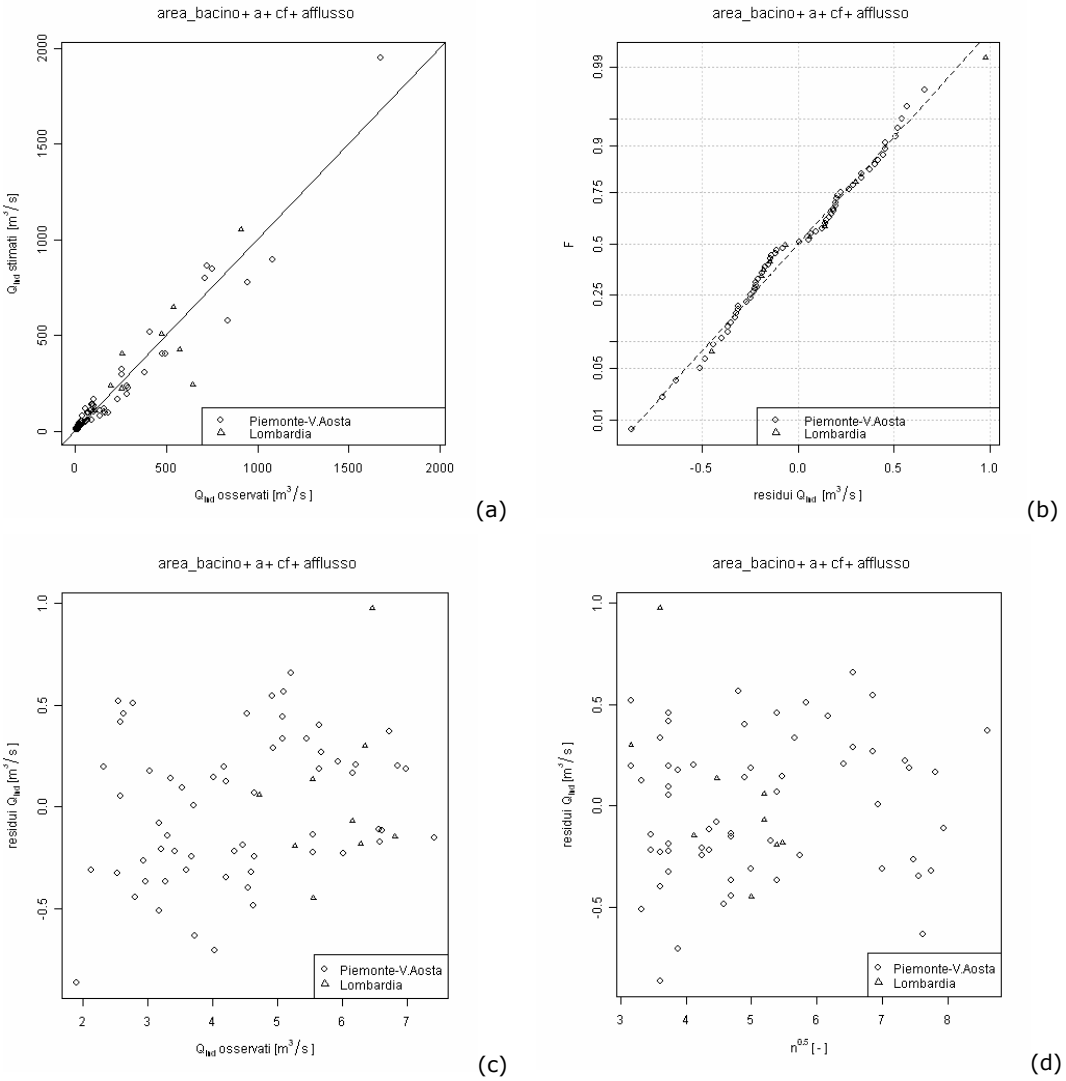


Figura 4.1. Diagrammi diagnostici per la relazione multiregressiva (4.2). In (a) si osserva il buon adattamento dei valori di Q_{ind} stimati a quelli osservati. L'andamento dei residui (b) soddisfa il requisito di normalità. In (c) e (d) viene riscontrata una apertura a ventaglio dei residui, che indica una lieve eteroschedasticità.

In figura 4.2 viene rappresentata la distribuzione sul territorio esaminato dei residui della regressione, definiti come la differenza tra il valore osservato e stimato della piena indice rapportata all'area e collocati sulla mappa in relazione al baricentro del bacino. Si nota che i residui risultano distribuiti in modo piuttosto uniforme sulla regione interessata dall'analisi; inoltre non si notano zone in cui si verifica una sottostima o una sovrastima sistematica delle portate indice osservate.

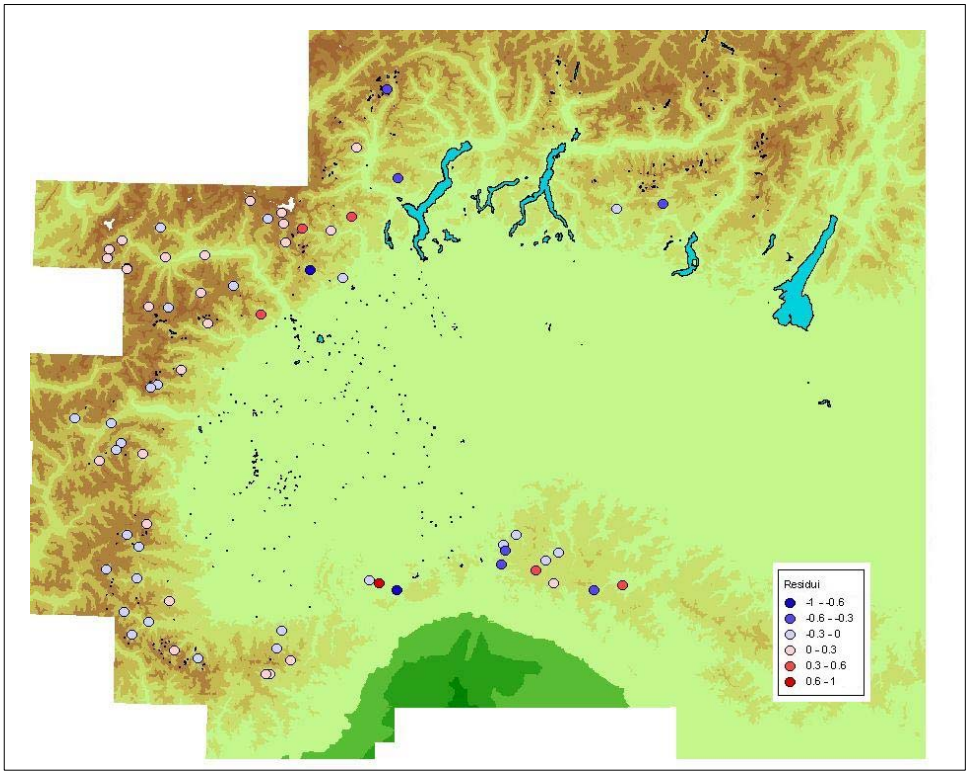


Figura 4.2. Mappa con la distribuzione dei residui calcolati su Q_{ind}/A nei siti dell'Italia Nord Occidentale relativi all'applicazione della formula multiregressiva (4.2).

4.1.2 Metodo basato sulla formula razionale

Alla luce dei risultati ottenuti tramite l'applicazione della formula razionale corretta mediante indici morfologici, climatici e di uso dei suoli, riportati nel paragrafo 3.3, viene scelta la relazione più appropriata per la stima della piena indice. In particolare la relazione più adeguata risulta essere quella che presenta il valore minore di MAPE, ovvero la (2.36), che riscritta con i coefficienti risulta:

$$Q_{ind,j} = \hat{K}_{4,g} \cdot f_{4,j} = 1.5243 \cdot 10^{-4} \cdot C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j}, \qquad (4.3)$$

alla quale è associata una certa variabilità residua, che potrebbe essere spiegata tramite l'inserimento nella (4.3) di ulteriori parametri geomorfologici. In Appendice C si verifica che la variabilità residua associata alla (4.3) non riesce ad essere spiegata in modo significativo da nessuna delle variabili morfologiche e climatiche disponibili.

In figura 4.3 viene indicata la distribuzione sul territorio esaminato dei residui calcolati tramite la formula razionale, definiti come la differenza tra il valore osservato e stimato della piena indice rapportata all'area e collocati sulla mappa in relazione al baricentro del bacino. Si nota che i residui risultano distribuiti in modo piuttosto uniforme sulla regione interessata dall'analisi; infatti non si notano zone in cui si verifica una sottostima o una sovrastima sistematica delle portate indice osservate.

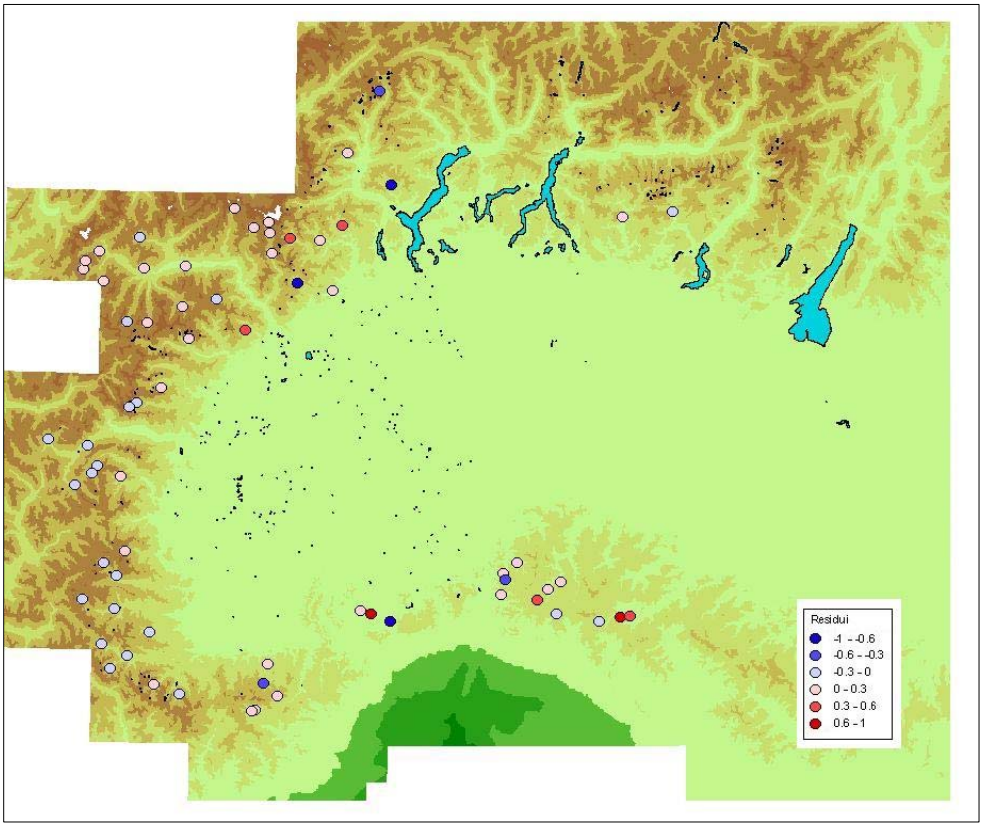


Figura 4.3. Mappa con la distribuzione dei residui calcolati su Q_{ind}/A nei siti dell'Italia Nord Occidentale relativi all'applicazione della formula razionale (4.3).

4.1.3 Metodo basato sulla trasformazione estremi giornalieri - colmi

Dovendo scegliere tra le diverse relazioni tarate nel campo logaritmico quella da impiegare per la valutazione del coefficiente di punta c_P, tramite il quale è possibile calcolare la piena indice a partire dalla media delle portate giornaliere massime annue, è necessario individuare una formula parsimoniosa, espressa in funzione di parametri facilmente determinabili.

In particolare la scelta è ricaduta sulla migliore regressione a 3 parametri tarata fissando la variabile dipendente *y* pari a $ln(c_P)$:

$$c_P = 4.0694 \cdot \sigma_{qg}^{0.12} \cdot \Delta H_2^{-0.17} \cdot a^{0.19}, \qquad (4.4)$$

per la quale si ha un errore medio percentuale del 10% prima della cross validazione e dell'11% dopo il jack-knife.

In figura 4.4 viene indicata la distribuzione sul territorio esaminato dei residui calcolati tramite equazione (4.4), definiti come la differenza tra il valore osservato e stimato della piena indice rapportata all'area e collocati sulla mappa in relazione al baricentro del bacino. Anche in questo caso i residui risultano distribuiti in modo piuttosto uniforme sulla regione interessata dall'analisi.

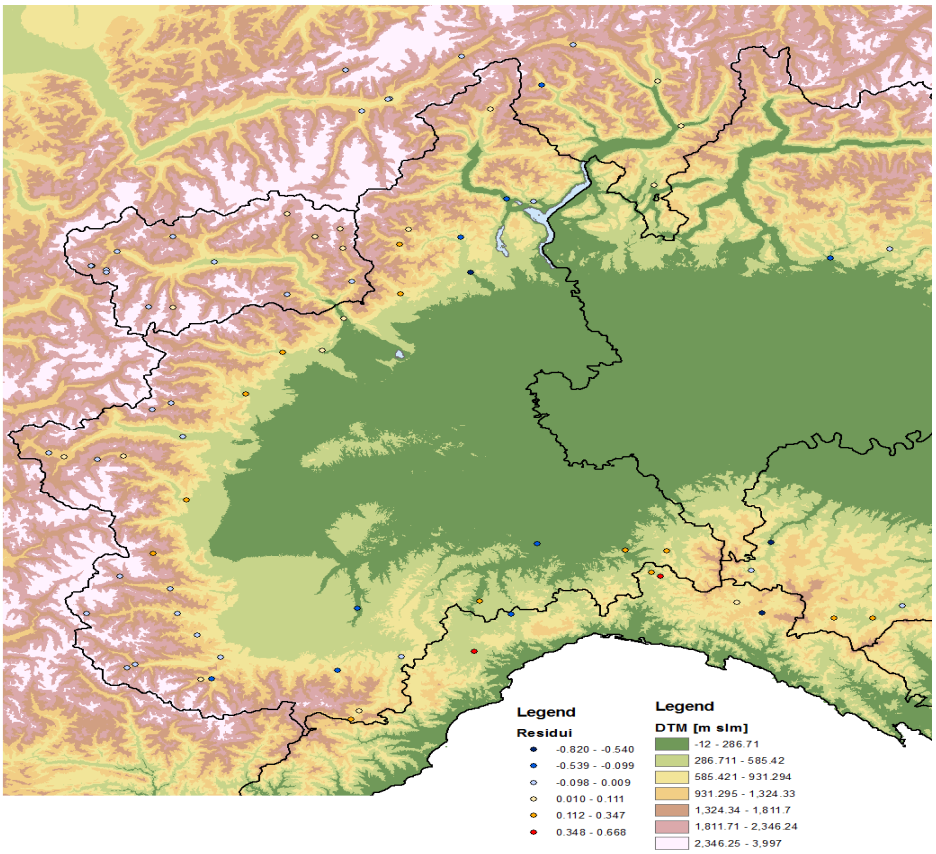


Figura 4.4. Mappa con la distribuzione dei residui calcolati su Q_{ind}/A nei siti dell'Italia Nord Occidentale relativi all'applicazione della formula razionale (4.4).

4.2 Confronto tra i metodi proposti su gruppi omogenei di bacini

Al fine di confrontare le prestazioni dei modelli prescelti nel paragrafo 4.1 si ritiene utile valutare il loro comportamento non solo globalmente su tutte le sezioni esaminate, ma anche su gruppi ristretti di bacini. Questi ultimi, in particolare, sono stati formati raggruppando in base ai valori di area e di quota media i 70 bacini ricadenti in Piemonte, Val d'Aosta, Lombardia con almeno 10 dati di portata al colmo e di estremi giornalieri. In entrambi i casi sono stati determinati tre differenti gruppi, per ciascuno dei quali è stato calcolato il valore di MAPE, così da valutare la qualità dei modelli prescelti nel paragrafo precedente.

La suddivisione operata in base al valore di area porta a selezionare 28 bacini con superficie inferiore ai 100 km^2, 35 con area compresa tra 100 e 1000 km^2, e 7 di dimensioni maggiori.

Esaminando i valori di MAPE riportati in tabella 4.1 emerge chiaramente come il metodo di trasformazione estremi giornalieri – colmi risulti particolarmente adeguato per la stima della piena indice nei bacini medio – piccoli, che, peraltro, rappresentano la quasi totalità dei bacini oggetto del presente lavoro. Per quanto riguarda i bacini di pianura, caratterizzati da elevati valori di area, si riscontra la sostanziale equivalenza tra gli errori derivanti dall'utilizzo della relazione multiregressiva (4.2) e dal metodo di trasformazione estremi giornalieri – colmi. La formula razionale (4.3), pur senza presentare valori di MAPE particolarmente alti, conduce a errori maggiori rispetto agli altri due metodi.

La suddivisione dei bacini in funzione della quota media ha portato a considerare 32 bacini di alta quota (H_{media} > 2000 m), 21 con altitudine media compresa tra 1000 e 2000 m, e 17 aventi quota inferiore a 1000 m

Metodo	A<100 km^2	100 ≤A≤1000 km^2	A>1000 km^2
Relaz. multiregressiva (4.2)	0.36	0.29	0.21
Formula razionale (4.3)	0.43	0.34	0.29
Trasf. giornalieri-colmi con c_P (4.8)	0.11	0.15	0.20

Tabella 4.1. Valori di MAPE calcolati per i modelli prescelti nel paragrafo 4.1 valutati su gruppi omogenei di stazioni definiti in base al valore di area.

Analizzando i valori di MAPE indicati in tabella 4.2 si nota che il metodo di trasformazione estremi giornalieri – colmi conduce anche in questo caso a errori di stima inferiori rispetto agli altri modelli analizzati; tuttavia si evidenzia come per i bacini con quota media inferiore a 1000 m esso produce risultati analoghi alla formula razionale (4.3). Per le sezioni a quote medio – alte la relazione multiregressiva è invece da preferire alla formula razionale, fermo restando il miglior comportamento del metodo di trasformazione giornalieri – colmi.

Metodo	**H_{media}<1000 m**	**1000 ≤H_{media}≤2000 m**	**H_{media}>2000 m**
Relaz. multiregressiva (4.2)	0.25	0.28	0.35
Formula razionale (4.3)	0.19	0.43	0.43
Tr. giornalieri-colmi con c_P (4.8)	0.19	0.14	0.10

Tabella 4.2. Valori di MAPE calcolati per i modelli prescelti nel paragrafo 4.1 valutati su gruppi omogenei di stazioni definiti in base al valore di quota media.

4.3 Condizioni di applicabilità delle relazioni proposte

Una volta stabilita, per ciascuna metodologia di stima analizzata, la relazione più appropriata per la valutazione della piena indice, è necessario determinare quali siano le condizioni nelle quali è opportuno utilizzare ciascuna formula. In particolare questo si traduce nella:

- determinazione della numerosità n_{min} di portate al colmo al di sotto della quale è preferibile ricorrere a un metodo di stima regionale, come l'approccio multiregressivo o la formula razionale, piuttosto che ad una stima *at site*, ovvero a una stima empirica di Q_{ind} effettuata a partire dal campione di osservazioni disponibile per la stazione in esame;
- determinazione della numerosità $n_{g,min}$ di estremi giornalieri annui al di sopra della quale è opportuno ricorrere a un metodo di trasformazione giornalieri – colmi piuttosto che a un metodo di stima regionale, quale l'approccio multiregressivo o la formula razionale;
- individuazione della relazione tra n_{min} ed n_g, rispettivamente la numerosità della serie delle portate al colmo e la numerosità di estremi

giornalieri, che costituisce la soglia al di sopra della quale è preferibile ricorrere ad una stima empirica a partire dalle osservazioni di Q_{ind} piuttosto che ad un metodo di trasformazione giornalieri – colmi.

La valutazione delle condizioni di applicabilità associate a ciascuna relazione è legata al valore del coefficiente di variazione di Q_{ind}:

$$CV_{Q_{ind}} = \frac{\sigma_{Q_{ind}}}{Q_{ind}}. \tag{4.5}$$

Nel caso della stima empirica, sostituendo nella (4.5) la definizione (2.3) di $\sigma_{Q_{ind}}$, si ottiene:

$$CV_{Q_{ind,emp}} = \frac{CV_Q}{\sqrt{n}}, \tag{4.6}$$

dove CV_Q è il coefficiente di variazione della serie di portata al colmo considerata.

Per quanto riguarda la relazione multiregressiva prescelta, la stima della piena indice viene effettuata tramite la (4.2), a cui è associata una varianza calcolata secondo la (2.21). Nel caso specifico si riscontra che il coefficiente di variazione risulta circa costante per le *j* stazioni in esame (si veda la tabella 5.1):

$$CV_{Q_{ind,regr}} = 0.423\,; \tag{4.7}$$

infatti il termine $\mathbf{a}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{a}^T$ nell'equazione (2.19) risulta trascurabile rispetto ad 1.

Anche nel caso della formula razionale il coefficiente di variazione assume il medesimo valore in ciascuna delle *j* stazioni, e in particolare risulta pari a:

$$CV_{Q_{ind,raz}} = 0.437\,. \tag{4.8}$$

Per quanto riguarda la trasformazione estremi giornalieri – colmi, stimando la piena indice secondo la relazione (2.51), in cui il coefficiente di punta c_P è calcolato con la relazione a 3 parametri (4.4), e la deviazione standard a partire dalla (2.54), è possibile scrivere:

$$CV_{Q_{ind,g}} = \sqrt{CV_{c_P}^2 \cdot CV_{\mu_g}^2 + CV_{c_P}^2 + CV_{\mu_g}^2} \,, \tag{4.9}$$

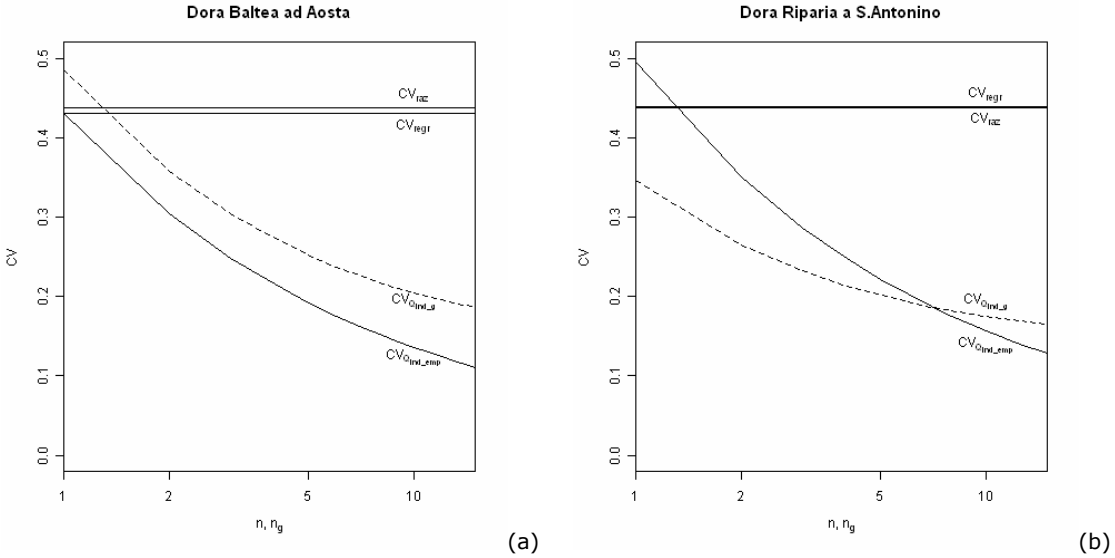
in cui:

- CV_{c_P} rappresenta il coefficiente di variazione di c_P, che risulta circa costante (per ragioni analoghe a $CV_{Q_{ind,regr}}$) e pari a 0.141;
- $CV_{\mu_g} = \dfrac{CV_{Q_g}}{\sqrt{n_g}}$ è il coefficiente di variazione degli estremi giornalieri.

Qualora si voglia stimare la piena indice Q_{ind} in un certo sito, strumentato o non, è necessario individuare quali siano le condizioni nelle quali è preferibile applicare ciascun metodo di stima. A tale scopo è inoltre utile costruire, per ciascuna sezione esaminata, un grafico nel quale riportare l'andamento dei coefficienti di variazione stimati in funzione del numero di osservazioni a disposizione. In figura 4.5 viene riportato il caso della Dora Baltea ad Aosta e della Dora Riparia a S.Antonino. In entrambi i casi il valore del *CV* associato alla relazione multiregressiva (4.2) e alla formula razionale (4.3) resta costante al variare di n e di n_g; al contrario i coefficienti di variazione $CV_{Q_{ind,emp}}$, associato alle portate al colmo osservate, e $CV_{Q_{ind,g}}$, relativo al campione di estremi giornalieri, decrescono al crescere della dimensione campionaria, secondo le relazioni (4.6) e (4.9). Analizzando il caso della Dora Baltea ad Aosta (Fig. 4.5.a) emerge che con 1 solo dato di portata al colmo si riesce ad ottenere una stima migliore di quella ottenibile tramite i metodi regionali; per quanto riguarda gli estremi giornalieri si devono invece avere almeno 2 dati per ottenere valori di CV inferiori rispetto a quelli associati alla formula razionale e alla relazione multiregressiva. Si nota inoltre che la curva relativa a $CV_{Q_{ind,emp}}$ sta sempre al di sotto di quella associata a $CV_{Q_{ind,g}}$; questo significa

che con 1 solo valore di portata al colmo si ottiene una stima di Q_{ind} più affidabile rispetto a quella ottenibile tramite la (4.4).

Per la Dora Riparia a S.Antonino (Fig.4.5.b) si riscontra che fino a 9 dati di portata al colmo è preferibile ricorrere al metodo di trasformazione estremi giornalieri – colmi, mentre al di sopra è più indicata una stima *at site* a partire dalle osservazioni al colmo. I metodi regionali, quali formula razionale o relazione multiregressiva, sono consigliabili soltanto qualora si abbiano meno di 2 dati di portata al colmo.



(a) (b)

Figura 4.5. Andamento dei coefficienti di variazione in funzione della dimensione campionaria.

L'analisi effettuata sui due bacini in figura 4.5 può essere generalizzata calcolando la numerosità minima di portate al colmo al di sopra della quale è preferibile ricorrere a una stima *at site* rispetto a un metodo di stima regionale; a tal fine si eguaglia la (4.6) al valore del coefficiente di variazione calcolato per il metodo che si vuole analizzare. Ad esempio per la relazione multiregressiva vale che:

$$\frac{CV_Q}{\sqrt{n_{\min}}} = CV_{Q_{ind,regr}}\,, \tag{4.10}$$

da cui segue:

$$n_{\min} = \frac{CV_Q^2}{CV_{Q_{ind,regr}}^2}\,. \tag{4.11}$$

Analogamente per la formula razionale:

$$n_{\min} = \frac{CV_Q^2}{CV_{Q_{ind,raz}}^2}, \tag{4.12}$$

che, per la Dora Riparia a S.Antonino (4.5b), risulta pari a 1.2.
Nel caso in cui si voglia verificare la numerosità minima per la quale è conveniente utilizzare il metodo di trasformazione estremi giornalieri – colmi si ricorre a una relazione ottenuta sostituendo l'espressione di CV_{μ_g} nella (4.9). In particolare per la formula multiregressiva si ottiene:

$$n_{g,\min} = \frac{CV_{c_P}^2 \cdot CV_{Q_g}^2 + CV_{Q_g}^2}{CV_{Q_{ind,regr}}^2 - CV_{c_P}^2}, \tag{4.13}$$

mentre per la formula razionale:

$$n_{g,\min} = \frac{CV_{c_P}^2 \cdot CV_{Q_g}^2 + CV_{Q_g}^2}{CV_{Q_{ind,raz}}^2 - CV_{c_P}^2}. \tag{4.14}$$

L'individuazione della relazione tra n ed n_g al di sopra del quale conviene ricorrere alla stima empirica effettuata a partire dal campione di portate al colmo disponibile piuttosto che ad un metodo di trasformazione giornalieri - colmi, viene determinata risolvendo l'equazione:

$$CV_{c_P}^2 \cdot \frac{CV_{Q_g}^2}{n_g} + CV_{c_P}^2 + \frac{CV_{Q_g}^2}{n_g} = \frac{CV_Q^2}{n_{\min}}. \tag{4.15}$$

Per avere delle indicazioni di massima riguardo la numerosità equivalente, valide per tutte le sezioni esaminate, si sono calcolati i valori di numerosità equivalente per tutte le stazioni, i quali sono stati poi mediati ottenendo i risultati riportati in tabella 4.3.
Analizzando i risultati indicati in tabella emerge chiaramente l'importanza di avere dati di portata osservati nelle stazioni di misura. Infatti una serie storica

anche poco consistente in una certa sezione, consente di effettuare una stima della piena indice migliore rispetto a quella ottenibile con i metodi di stima regionale tarati nel presente lavoro o sviluppati in precedenza da altri Autori per il territorio in esame (*De Michele e Rosso*, 2001; *Villani*, 2002; *Ciaponi e Moisello*, 1988).

Numerosità equivalente	***Descrizione***
$n_{min} > 2.4$	Conviene utilizzare la stima empirica sulla serie di Q_{ind} piuttosto che la formula multiregressiva (4.2)
$n_{min} > 2.2$	Conviene utilizzare la stima empirica sulla serie di Q_{ind} piuttosto che la formula razionale (4.3)
$n_{g,min} > 1.6$	Conviene utilizzare il metodo di trasformazione estremi giornalieri – colmi con il c_P stimato secondo la (4.8) piuttosto che la formula multiregressiva (4.2)
$n_{g,min} > 1.8$	Conviene utilizzare il metodo di trasformazione estremi giornalieri – colmi con il c_P stimato secondo la (4.8) piuttosto che la formula razionale (4.3)

Tabella 4.3. Valori della numerosità equivalente utile a definire le condizioni di applicabilità di ciascun metodo in ragione del numero di osservazioni disponibili.

Mediamente la numerosità minima di portate al colmo n_{min} al di sopra della quale conviene ricorrere ad una stima *at site* piuttosto che al metodo di trasformazione a partire dagli estremi giornalieri risulta pari a 1.6. Tale numerosità risulta estremamente variabile da un caso all'altro: a partire infatti da valori di 1 (ottenuto per l'Elvo a Sordevolo) si arriva ad una numerosità equivalente di 50 (per il Varaita a Rore).

Dalla tabella 4.3 emerge l'importanza di avere misure di portata nella sezione di interesse, sia in termini di estremi giornalieri che di estremi al colmo. In particolare è interessante notare come l'utilizzo della serie storica delle portate giornaliere massime annue conduca, qualora siano note poche misurazioni, a risultati comparabili a quelli ottenibili tramite la serie delle portate al colmo. Ciò è dovuto al fatto che le serie di estremi giornalieri utilizzate nel presente studio presentano minore variabilità delle serie corrispondenti di portata al colmo. Questo comporta una sostanziale equivalenza tra l'errore di stima introdotto tramite una trasformazione della media degli estremi giornalieri in Q_{ind} e l'errore associato alla stima empirica della piena indice.

4.4 Intervalli di confidenza

Utilizzando le stime della varianza (o del coefficiente di variazione) riportate nel paragrafo 4.3 si possono determinare gli intervalli di confidenza della stima della piena indice ottenuta con uno qualunque dei quattro metodi considerati nel presente volume.

La determinazione degli intervalli di confidenza richiede la conoscenza dell'intera distribuzione di probabilità della piena indice: i limiti inferiore e superiore, $\{l_{\inf}(\alpha), l_{\sup}(\alpha)\}$, dell'intervallo di confidenza simmetrico, con coefficiente di confidenza (ossia, probabilità di ricadere all'interno dell'intervallo di confidenza) α, possono infatti essere stimati tramite:

$$l_{\inf}(\alpha) = \xi_{Q_{ind}}\left(0.5 - \frac{\alpha}{2}\right) \quad , \quad l_{\sup}(\alpha) = \xi_{Q_{ind}}\left(0.5 + \frac{\alpha}{2}\right),$$

dove $\xi_{Q_{ind}}(p)$ rappresenta il quantile della distribuzione della piena indice associato ad una probabilità *p*. Per determinare i limiti degli intervalli di confidenza è quindi sufficiente trovare la generica espressione per stimare il quantile della piena indice, e poi sostituire $p = 0.5 - \frac{\alpha}{2}$ e $p = 0.5 + \frac{\alpha}{2}$.

4.4.1 Stima empirica

Nel caso della stima empirica della piena indice si ha che la piena indice viene stimata come il valor medio dei valori di portata al colmo di piena disponibili nella sezione di interesse. La distribuzione della piena indice tenderà dunque, per valori di *n* elevati, ad una distribuzione normale con media pari a $Q_{ind,emp}$ e coefficiente di variazione ottenuto dall'equazione (4.6).

La variabile casuale $\frac{Q_{ind} - Q_{ind,emp}}{CV_{Q_{ind},emp} Q_{ind,emp}}$ ha pertanto una distribuzione normale standardizzata (ossia, con media zero e varianza unitaria), ed il quantile della piena indice con probabilità *p* può essere stimato come

$$\xi_{Q_{ind}}(p) = Q_{ind,emp}\left(1 + CV_{Q_{ind,emp}} \cdot z_p\right) = Q_{ind,emp}\left(1 + \frac{CV_Q}{\sqrt{n}} \cdot z_p\right), \qquad (4.16)$$

dove z_p è il quantile con probabilità *p* della variabile normale standardizzata. E' chiaro che l'ampiezza dell'intervallo di confidenza ottenuto dall'applicazione dell'equazione (4.16) dipende dalla dimensione campionaria *n.* L'approssimazione data dall'impiego della distribuzione normale per rappresentare la distribuzione della media campionaria è buona per *n* > 10, ed è comunque accettabile per *n* > 5. Quando si considerano dimensioni campionarie più basse bisognerebbe tener conto dell'effettiva distribuzione generatrice dei dati, che è tuttavia incognita.
Per quantificare, seppur in maniera approssimata, l'incertezza associata alle stime empiriche della piena indice, si è scelto di inserire nella (4.16) il valor medio su tutte le serie dei coefficienti di variazione delle portate, $CV_Q = 0.650$, e di costruire due tabelle degli intervalli di confidenza al variare di *n*, rispettivamente con coefficiente di confidenza $\alpha = 0.3$ (*Tab. 4.4*) e $\alpha = 0.7$ (*Tab. 4.5*). Dato che le tabelle così costruite, per la loro generalità, hanno comunque significato solo orientativo, si è scelto di estendere l'utilizzo dell'equazione (4.16) anche per valori di *n* inferiori a 5.

4.4.2 Modello di regressione multipla

Nel caso dell'impiego della formula multiregressiva (4.2) per stimare la piena indice si ha che Q_{ind} ha la stessa distribuzione dei residui della regressione; dato che la regressione è stata calibrata nel campo logaritmico, si ha dunque che la variabile casuale $\frac{\ln[Q_{ind}] - \vartheta_1}{\vartheta_2}$, dove ϑ_1 e ϑ_2 sono la media e la deviazione standard di $\ln[Q_{ind}]$, ha una distribuzione normale standardizzata. Il quantile si calcola dunque come:

$$\xi_{Q_{ind}}(p) = \exp\left[\vartheta_1 + \vartheta_2 z_p\right], \tag{4.17}$$

con $\vartheta_1 = \ln\left[Q_{ind,regr}\right] - \frac{1}{2}\ln\left[1 + CV^2_{Q_{ind,regr}}\right]$ e $\vartheta_2 = \sqrt{\ln\left[1 + CV^2_{Q_{ind,regr}}\right]}$, ottenute con le classiche formule per determinare i parametri di una distribuzione lognormale. Sostituendo nella (4.17) si ottiene:

$$\xi_{Q_{ind}}(p) = Q_{ind,regr} \frac{\exp\left[z_p \sqrt{\ln\left(1+CV^2_{Q_{ind,regr}}\right)} \right]}{\sqrt{1+CV^2_{Q_{ind,regr}}}}. \qquad (4.18)$$

L'equazione (4.18) è chiaramente indipendente dalla dimensione campionaria, derivando da una regressione lineare. I risultati dell'applicazione di tale equazione, utilizzando un valore medio su tutte le serie di $CV_{Q_{ind,regr}} = 0.423$, sono riportati in Tabella 4.4 e Tabella 4.5 per coefficienti di confidenza 0.3 e 0.7. Si noti che gli intervalli di confidenza in tal caso non risultano simmetrici intorno al valore stimato. Tuttavia l'asimmetria non è, come ci si potrebbe attendere, tale da portare a valori di $l_{\inf}(\alpha)$ più vicini al valore stimato rispetto a $l_{\sup}(\alpha)$, ma accade il contrario. La ragione di questo risultato sta nella combinazione di due effetti: i ridotti valori dei coefficienti di confidenza utilizzati, ed il fatto che il valore stimato non è il valore mediano della distribuzione lognormale, ma invece il valore medio, cosa che sposta il valore stimato verso valori più elevati, ossia proprio verso $l_{\sup}(\alpha)$. Qualora si fossero scelti coefficienti di confidenza più elevati, per esempio $\alpha = 0.9$, tale effetto non si sarebbe riscontrato, perché l'asimmetria positiva della distribuzione lognormale sarebbe risultata dominante rispetto allo spostamento tra media e mediana, producendo valori di $l_{\sup}(\alpha)$ più lontani dal valore stimato rispetto a $l_{\inf}(\alpha)$.

4.4.3 Metodo basato sulla formula razionale

Il caso della formula razionale è del tutto analogo a quello della relazione multiregressiva, dal momento che anche per la formula razionale si ha che i residui (e quindi anche Q_{ind}) tendono ad avere una distribuzione lognormale. Il parametro *K* della formula razionale è stato infatti stimato tramite media geometrica (si veda paragrafo 3.3). Di conseguenza i quantili della distribuzione di Q_{ind} possono ancora essere determinati impiegando una relazione analoga alla (4.18), ossia

$$\xi_{Q_{ind}}(p) = Q_{ind,raz} \frac{\exp\left[z_p \sqrt{\ln\left(1+CV^2_{Q_{ind,raz}}\right)} \right]}{\sqrt{1+CV^2_{Q_{ind,raz}}}}. \qquad (4.19)$$

Anche in questo caso si ha che l'equazione (4.19) è indipendente dalla dimensione campionaria. I risultati dell'applicazione di tale equazione, utilizzando un valor medio su tutte le serie di $CV_{Q_{ind},raz} = 0.437$, sono riportati in Tabella 4.4 e Tabella 4.5 per coefficienti di confidenza 0.3 e 0.7.

4.4.4 Metodo basato sulla trasformazione estremi giornalieri-colmi

Nel caso della stima della piena indice tramite trasformazione estremi giornalieri-colmi si ha che la distribuzione della piena indice risulterebbe pari al prodotto di una distribuzione lognormale (distribuzione del coefficiente di punta c_p, stimato tramite regressione tarata nel campo logaritmico) e di una distribuzione normale (distribuzione della media campionaria dei valori giornalieri). Tale distribuzione non ha forma analitica nota, cosa che richiederebbe di determinare i valori dei quantili per via numerica. Si è tuttavia verificato che per dimensioni campionarie maggiori di 10 tale distribuzione è approssimabile in maniera ottima utilizzando una distribuzione lognormale con media $Q_{ind,g}$ e coefficiente di variazione ottenuto dall'equazione (4.9). L'approssimazione rimane accettabile per n_g > 5, mentre perde di validità per dimensioni campionarie ancora minori, dimensioni per le quali viene comunque a decadere anche l'approssimazione della media campionaria dei valori giornalieri tramite una distribuzione normale. Si è pertanto scelto di impiegare l'approssimazione lognormale, ossia di stimare i quantili, utili per determinare i limiti degli intervalli di confidenza, tramite la relazione:

$$\xi_{Q_{ind}}(p) = Q_{ind,g} \frac{\exp\left[z_p \sqrt{\ln\left(1+CV^2_{Q_{ind,g}}\right)} \right]}{\sqrt{1+CV^2_{Q_{ind,g}}}} = Q_{ind,g} \frac{\exp\left[z_p \sqrt{\ln\left(1+\frac{CV^2_{c_p}\cdot CV^2_{Q_g}}{n_g}+CV^2_{c_p}+\frac{CV^2_{Q_g}}{n_g}\right)} \right]}{\sqrt{1+\frac{CV^2_{c_p}\cdot CV^2_{Q_g}}{n_g}+CV^2_{c_p}+\frac{CV^2_{Q_g}}{n_g}}}. \qquad (4.20)$$

L'ampiezza dell'intervallo di confidenza ottenuto dall'applicazione dell'equazione (4.20) dipende dalla dimensione n_g del campione di dati giornalieri. Anche in questo caso si è scelto di utilizzare in Tabella 4.4 ed in Tabella 4.5 un valore medio su tutte le serie di $CV_{c_p} = 0.141$ e di $CV_{Q_g} = 0.503$. Di nuovo, dato il carattere solo orientativo delle tabelle, si è scelto di estendere l'utilizzo dell'equazione (4.26) anche per valori di n_g inferiori a 5.

n, n_g	**Stima empirica**		**Formula regressiva**	
	Limite inferiore	Limite superiore	Limite inferiore	Limite superiore
1	0.75 $Q_{ind,emp}$	1.25 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
2	0.82 $Q_{ind,emp}$	1.18 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
3	0.86 $Q_{ind,emp}$	1.14 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
5	0.89 $Q_{ind,emp}$	1.11 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
10	0.92 $Q_{ind,emp}$	1.08 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
15	0.94 $Q_{ind,emp}$	1.06 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
20	0.94 $Q_{ind,emp}$	1.06 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
30	0.95 $Q_{ind,emp}$	1.05 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
50	0.96 $Q_{ind,emp}$	1.04 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
70	0.97 $Q_{ind,emp}$	1.03 $Q_{ind,emp}$	0.79 $Q_{ind,regr}$	1.08 $Q_{ind,regr}$
n, n_g	**Formula razionale**		**Trasformazione giornalieri - colmi**	
	Limite inferiore	Limite superiore	Limite inferiore	Limite superiore
1	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.73 $Q_{ind,g}$	1.07 $Q_{ind,g}$
2	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.81 $Q_{ind,g}$	1.08 $Q_{ind,g}$
3	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.84 $Q_{ind,g}$	1.07 $Q_{ind,g}$
5	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.87 $Q_{ind,g}$	1.07 $Q_{ind,g}$
10	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.90 $Q_{ind,g}$	1.06 $Q_{ind,g}$
15	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.91 $Q_{ind,g}$	1.06 $Q_{ind,g}$
20	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.92 $Q_{ind,g}$	1.05 $Q_{ind,g}$
30	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.92 $Q_{ind,g}$	1.05 $Q_{ind,g}$
50	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.93 $Q_{ind,g}$	1.05 $Q_{ind,g}$
70	0.78 $Q_{ind,raz}$	1.08 $Q_{ind,raz}$	0.93 $Q_{ind,g}$	1.05 $Q_{ind,g}$

Tabella 4.4. Intervalli di confidenza associati ai diversi metodi di stima per α = 0.3.

Dall'analisi delle Tabelle 4.4 e 4.5 si riscontra come l'ampiezza degli intervalli di confidenza tenda ad essere molto più elevata quando si applicano metodi regionali di stima, quali le formule regressive o la formula razionale, cosa peraltro del tutto intuibile. E' infatti sufficiente avere una dimensione

campionaria pari a 2-3 elementi per avere degli intervalli di confidenza relativi alla stima diretta (colonna 1) più vicini al valore stimato di quelli delle colonne 2 e 3 . Il valore particolarmente basso della numerosità equivalente, per quanto al di sotto di quanto ci si aspettava, è del tutto consistente con quanto riscontrato in un analogo studio regionale delle portate di piena nel Regno Unito (*Robson e Reed*, 1999), da cui risulta addirittura che è sufficiente un solo anno di misura (Tab. 2.2, pag. 10, Vol. 3) per avere una stima empirica altrettanto efficiente di quella regionale. Tale risultato non è quindi da ascrivere alla scarsa qualità delle relazioni regionali per stimare la piena indice ottenute nel presente volume, ma è piuttosto un chiaro sintomo dell'elevatissimo valore dei dati per ottenere stime migliori delle variabili di progetto.

n, n_g	**Stima empirica**		**Formula regressiva**	
	Limite inferiore	Limite superiore	Limite inferiore	Limite superiore
1	0.33 $Q_{ind,emp}$	1.67 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
2	0.52 $Q_{ind,emp}$	1.48 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
3	0.61 $Q_{ind,emp}$	1.39 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
5	0.70 $Q_{ind,emp}$	1.30 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
10	0.79 $Q_{ind,emp}$	1.21 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
15	0.83 $Q_{ind,emp}$	1.17 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
20	0.85 $Q_{ind,emp}$	1.15 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
30	0.88 $Q_{ind,emp}$	1.12 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
50	0.90 $Q_{ind,emp}$	1.10 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
70	0.92 $Q_{ind,emp}$	1.08 $Q_{ind,emp}$	0.60 $Q_{ind,regr}$	1.40 $Q_{ind,regr}$
n, n_g	**Formula razionale**		**Trasformazione giornalieri - colmi**	
	Limite inferiore	Limite superiore	Limite inferiore	Limite superiore
1	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.53 $Q_{ind,g}$	1.48 $Q_{ind,g}$
2	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.63 $Q_{ind,g}$	1.37 $Q_{ind,g}$
3	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.68 $Q_{ind,g}$	1.32 $Q_{ind,g}$
5	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.74 $Q_{ind,g}$	1.27 $Q_{ind,g}$
10	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.79 $Q_{ind,g}$	1.22 $Q_{ind,g}$
15	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.81 $Q_{ind,g}$	1.20 $Q_{ind,g}$
20	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.82 $Q_{ind,g}$	1.19 $Q_{ind,g}$
30	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.83 $Q_{ind,g}$	1.17 $Q_{ind,g}$
50	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.84 $Q_{ind,g}$	1.16 $Q_{ind,g}$
70	0.59 $Q_{ind,raz}$	1.41 $Q_{ind,raz}$	0.84 $Q_{ind,g}$	1.16 $Q_{ind,g}$

Tabella 4.5. Intervalli di confidenza associati ai diversi metodi di stima per α = 0.7.

5 Applicazione dei metodi di stima della piena indice

Nei precedenti capitoli sono stati analizzati diversi approcci utili alla valutazione della piena indice. In particolare sono state sviluppate un'analisi multiregressiva, un approccio basato sulla formula razionale e, infine, è stato predisposto un metodo di trasformazione della media degli estremi giornalieri in portata indice. Inoltre,ne è stata effettuata una stima empirica a partire dal campione di osservazioni disponibile per le sezioni strumentate.

Una volta stabilita per ciascuna metodologia la relazione più adeguata per la stima della piena indice (paragrafi 4.1 e 4.2), è possibile applicarla in tutti i bacini analizzati nel presente lavoro. Nella tabella che segue vengono riportate le stime di Q_{ind} secondo i diversi metodi, accompagnate dalla corrispondente incertezza di stima, valutata in termini di coefficiente di variazione della piena indice. In particolare per i bacini strumentati sarà presente la stima empirica di Q_{ind}, mentre per quelli non strumentati sarà possibile effettuare la stima solo con i metodi regionali, quali l'analisi multiregressiva e la formula razionale. Si è ritenuto utile indicare anche la numerosità delle misurazioni di portata giornaliera e al colmo disponibili, in modo da poter associare a ciascuna stazione il metodo più appropriato per la stima della piena indice, secondo i criteri riassunti in Tabella 4.3.

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Adda a Fuentes	42	48	608.3	0.062	708.8	0.433	481.6	0.437	698.6	0.158
Adda a Tirano	11	11	191.6	0.204	243.7	0.427	169.3	0.437	111.2	0.243
Artavaz a St. Oyen	14	17	12.6	0.128	17.7	0.417	19.5	0.437	13.8	0.172
Aveto a Cabanne	27	24	112.5	0.139	105.8	0.419	121.2	0.437	141.7	0.163
Ayasse a Champorcher	29	24	19.4	0.096	27.9	0.411	22.7	0.437	20.7	0.156
Ayasse a Miserin	-	-	-	-	3.0	0.406	2.2	0.437	-	-
Borbera a Baracche	22	21	256.8	0.163	293.5	0.418	228.2	0.437	214.4	0.220
Bormida a Cassine	19	22	754.0	0.059	843.0	0.426	678.4	0.437	1159.4	0.183
Bormida ad Alessandria	4	9	1205.3	0.200	61.8	0.429	1087.7	0.437	2081.3	0.218
Bormida di Mallare a Ferrania	23	21	163.6	0.185	93.4	0.412	94.1	0.437	126.4	0.280
Bormida Spigno a Valla	47	47	138.1	0.132	80.4	0.415	88.4	0.437	121.7	0.164
Bousset a Tetti Porcera	6	5	15.9	0.237	28.5	0.425	41.1	0.437	15.9	0.278
Brembo a P.te Briolo	30	20	540.5	0.092	641.1	0.433	465.8	0.437	610.4	0.168
Breuil ad Alpette	14	14	13.1	0.071	8.5	0.419	9.6	0.437	14.2	0.161

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Bucera a Chiotas	-	-	-	-	11.0	0.414	10.6	0.437	-	-
Bucera a P.te Rovine	6	5	17.1	0.293	20.4	0.421	26.4	0.437	15.3	0.362
Buthier a Place Moulin	-	37	-	-	15.8	0.415	11.4	0.437	45.3	0.150
Cervo a Passobreve	13	15	94.7	0.167	140.0	0.418	147.5	0.437	82.7	0.191
Chiavanne ad Alpette	14	14	10.2	0.147	8.4	0.415	8.1	0.437	10.4	0.173
Chisone a Fenestrelle	19	23	30.6	0.215	38.1	0.418	39.8	0.437	23.6	0.182
Chisone a S.Basses	22	11	16.5	0.153	25.6	0.417	26.5	0.437	16.9	0.186
Chisone a S.Martino	24	37	286.5	0.144	192.7	0.425	207.4	0.437	180.1	0.192
Chiusella a Gurzia	32	31	233.1	0.129	166.2	0.419	164.6	0.437	222.3	0.208
Corsaglia a C.le Molline	25	29	36.8	0.125	50.4	0.432	86.9	0.437	44.8	0.173
Dora Baltea a Beauregard	-	40	-	-	22.0	0.416	16.5	0.437	33.7	0.149
Dora Baltea a Cignana	-	68	-	-	4.4	0.405	3.1	0.437	4.9	0.146
Dora Baltea a P.te di Mombar.	14	15	93.9	0.038	60.0	0.418	49.5	0.437	92.4	0.149
Dora Baltea a Tavagnasco	74	64	834.7	0.070	579.8	0.429	488.2	0.437	617.3	0.151

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Dora Baltea ad Aosta	25	17	285.5	0.086	239.1	0.431	223.6	0.437	264.1	0.163
Dora di Bardonecchia a Beaul.	12	17	27.4	0.107	31.5	0.431	43.1	0.437	25.7	0.170
Dora di Courm. a Prè St.Didier	7	16	60.7	0.119	42.7	0.425	39.8	0.437	52.5	0.154
Dora di Rhemes a Notre Dame	14	14	13.2	0.052	12.6	0.429	15.0	0.437	13.0	0.157
Dora di Rhemes a Pelaud	5	17	16.5	0.098	8.8	0.433	11.9	0.437	12.7	0.149
Dora di Rhemes a St.Georges	6	6	16.6	0.118	24.0	0.422	23.9	0.437	16.3	0.172
Dora Riparia a Chiomonte	-	40	-	-	87.8	0.433	113.5	0.437	64.5	0.153
Dora Riparia a Oulx	30	33	55.8	0.178	48.2	0.426	58.4	0.437	40.1	0.166
Dora Riparia a S.Antonino	60	34	99.0	0.064	135.2	0.439	203.2	0.437	97.9	0.163
Dora Riparia a Salbertrand	-	53	-	-	79.9	0.432	101.5	0.437	59.8	0.152
Elvo a Sordevolo	2	7	115.0	0.043	80.1	0.414	81.5	0.437	49.7	0.253
Entracque a Piastra	-	-	-	-	57.0	0.425	70.2	0.437	-	-
Erro a Sassello	21	16	103.4	0.072	167.0	0.415	162.6	0.437	120.5	0.177
Evancon a Brusson	-	14	-	-	45.5	0.411	31.8	0.437	15.0	0.163

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Evancon a Champoluc	22	29	26.4	0.122	38.2	0.410	23.6	0.437	23.5	0.169
Gesso Barra a S.Giacomo	6	5	14.4	0.208	20.5	0.414	22.2	0.437	12.9	0.279
Gesso Colombo a S.Giacomo	5	5	18.2	0.234	21.2	0.421	26.0	0.437	22.7	0.258
Gesso della Valletta a S.Loren.	11	10	67.6	0.245	59.3	0.421	63.1	0.437	60.3	0.168
Gesso di Entracque a Entracq.	12	12	76.4	0.255	94.2	0.428	120.5	0.437	93.8	0.278
Grana a Monterosso	48	48	40.9	0.162	40.9	0.427	63.4	0.437	44.6	0.192
Grand'Eyvia a Cretaz	10	22	65.0	0.123	53.0	0.423	44.9	0.437	48.8	0.151
Kant a Fedio	-	-	-	-	10.4	0.422	17.8	0.437	-	-
Lys a D'Ejola	10	13	12.7	0.098	7.4	0.427	5.3	0.437	11.6	0.156
Lys a Gressoney	24	23	28.8	0.103	24.8	0.426	27.4	0.437	25.9	0.152
Lys a Guillemore	29	29	105.2	0.178	98.1	0.416	84.2	0.437	114.4	0.219
Lys a Lago Gabiet	-	-	-	-	1.6	0.417	2.1	0.437	-	-
Maira a Combamala	-	-	-	-	6.0	0.419	11.1	0.437	-	-
Maira a S.Damiano Macra	57	57	67.9	0.115	96.6	0.434	130.6	0.437	68.0	0.181

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Maira a Saretto	13	13	6.7	0.069	16.2	0.423	18.7	0.437	6.9	0.155
Marmore a Balanselmo	-	-	-	-	1.9	0.396	0.8	0.437	-	-
Marmore a Balanselmo	-	-	-	-	1.9	0.396	0.8	0.437	-	-
Lys a Lago Gabiet	-	-	-	-	1.6	0.417	2.1	0.437	-	-
Marmore a Lago Goillet	-	-	-	-	1.8	0.406	1.4	0.437	-	-
Marmore a Lago Grande 1	-	-	-	-	0.5	0.394	0.2	0.437	-	-
Marmore a Lago Grande 2	-	-	-	-	0.4	0.393	0.2	0.437	-	-
Marmore a Perreres	15	14	20.9	0.201	17.6	0.411	12.0	0.437	18.1	0.253
Marmore a Ussin	-	-	-	-	32.6	0.411	24.1	0.437	-	-
Mastallone a P.te Folle	54	32	380.0	0.094	303.9	0.417	325.8	0.437	396.1	0.183
Melezet	6	-	5.3	0.081	3.3	0.458	12.1	0.437	-	-
Meris a S.Anna Valdieri	5	5	9.0	0.218	11.9	0.417	15.5	0.437	11.2	0.260
Nontey a Valnontey	5	5	13.8	0.162	14.6	0.427	14.2	0.437	11.3	0.170
Orco a Pont Canavese	41	47	498.4	0.111	403.1	0.422	329.1	0.437	331.4	0.160
Po a Crissolo	14	39	34.3	0.324	31.3	0.410	24.4	0.437	24.0	0.251
Presa Clarea	-	40	-	-	5.8	0.431	11.6	0.437	4.1	0.226

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Presa Galambra	-	40	-	-	3.1	0.426	5.7	0.437	2.4	0.157
Rio Bagni a Vinadio	20	20	24.2	0.300	26.0	0.422	32.4	0.437	23.9	0.237
Rio del Piz a Pietraporzio	-	23	-	-	9.8	0.418	12.7	0.437	5.2	0.158
Rio Freddo	7	7	32.8	0.122	16.8	0.422	22.5	0.437	42.2	0.223
Ripa a Bousson	5	5	12.5	0.140	34.9	0.424	38.7	0.437	14.7	0.207
Roccasparvera	-	-	-	-	141.2	0.435	197.8	0.437	-	-
Rochemolles	-	29	-	-	7.2	0.418	6.8	0.437	10.0	0.155
Rutor a La Joux	29	29	13.8	0.068	8.8	0.417	8.2	0.437	13.5	0.154
Rutor a Promise	34	32	16.1	0.093	9.7	0.417	9.1	0.437	16.5	0.150
Sabbione	-	47	-	-	2.4	0.469	10.5	0.437	18.0	0.153
Sampeyre	-	-	-	-	66.8	0.430	80.6	0.437	-	-
San Bernardino a Santino	14	17	259.5	0.084	324.5	0.417	349.0	0.437	255.2	0.162
San Nicolao	-	-	-	-	0.3	0.418	0.9	0.437	-	-
Savara a Eau Rousse	18	22	24.6	0.083	30.1	0.417	21.3	0.437	22.6	0.159

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Savara a Fenille	6	6	31.5	0.195	37.8	0.417	28.5	0.437	24.4	0.215
Scrivia a Isola del Cantone	13	10	412.8	0.195	515.6	0.419	396.2	0.437	342.7	0.173
Scrivia a Serravalle	28	29	725.5	0.128	860.1	0.422	601.7	0.437	600.2	0.167
Serio a P.te Cene	25	20	258.4	0.076	401.2	0.430	328.8	0.437	287.0	0.165
Sesia a Campertogno	38	28	160.8	0.151	102.6	0.419	106.9	0.437	127.5	0.215
Sesia a P.te Aranco	17	24	947.7	0.176	774.2	0.424	781.3	0.437	1287.3	0.194
Sesia a Palestro	-	10	-	-	1954.3	0.432	1461.9	0.437	2651.8	0.237
Sesia a Rimasco	43	43	184.4	0.091	95.8	0.413	97.2	0.437	170.1	0.172
Sesia a Vercelli	22	6	1673.2	0.110	1941.1	0.432	1465.6	0.437	1728.4	0.169
Stura di Demonte a Fossano	-	5	-	-	405.4	0.437	482.2	0.437	440.6	0.360
Stura di Demonte a Gaiola	33	30	104.9	0.149	135.8	0.435	187.1	0.437	101.3	0.181
Stura di Demonte a Pianche	18	23	39.5	0.190	50.4	0.427	61.9	0.437	44.4	0.214
Stura di Lanzo a Lanzo	61	53	478.1	0.079	406.7	0.419	315.9	0.437	386.0	0.164
Stura di Viu a Lago Torre	-	-	-	-	1.5	0.407	0.9	0.437	-	-

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Stura di Viu a Malciaussia	49	49	8.4	0.078	11.5	0.416	10.5	0.437	8.3	0.168
Stura di Viu a Rossa	-	53	-	-	3.0	0.410	1.9	0.437	3.2	0.157
Stura di Viu a Usseglio	11	11	24.1	0.207	40.6	0.416	34.5	0.437	27.7	0.253
Tanaro a Farigliano	63	57	714.3	0.077	790.6	0.429	666.2	0.437	843.5	0.160
Tanaro a Nucetto	47	31	292.9	0.120	225.2	0.427	262.6	0.437	305.7	0.173
Tanaro a Ormea	13	12	162.2	0.140	115.1	0.426	146.4	0.437	150.9	0.181
Tanaro a P.te Nava	43	37	139.3	0.151	105.2	0.424	127.7	0.437	103.9	0.167
Taro a Carniglia	29	25	194.9	0.090	235.7	0.422	214.3	0.437	177.7	0.157
Taro a Ostia	10	11	574.6	0.185	426.3	0.434	433.0	0.437	595.3	0.191
Taro a Pradella	13	13	646.5	0.146	243.6	0.446	381.5	0.437	529.2	0.230
Toce a Cadarese	15	20	56.9	0.107	115.2	0.421	124.4	0.437	43.1	0.158
Toce a Camplicciolì	-	-	-	-	39.3	0.406	31.9	0.437	-	-
Toce a Camposecco	-	-	-	-	5.7	0.398	3.2	0.437	-	-
Toce a Candoglia	55	37	1081.9	0.074	893.2	0.426	1027.5	0.437	1222.3	0.172

Denominazione	**n**	**n_g**	**$Q_{ind,oss}$**	**CV_{oss}**	**$Q_{ind,regr}$**	**$CV_{Qind,regr}$**	**$Q_{ind,raz}$**	**$CV_{Qind,raz}$**	**$Q_{ind,g}$**	**$CV_{Qind,g}$**
Toce a Cingino	-	-	-	-	4.6	0.398	2.6	0.437	-	-
Toce a Codelago	-	64	-	-	25.9	0.410	23.9	0.437	29.1	0.144
Toce a Lago Busin	-	60	-	-	3.5	0.401	2.6	0.437	3.6	0.146
Toce a Morasco	-	-	-	-	13.5	0.439	26.7	0.437	-	-
Toce a Obersee	-	-	-	-	3.2	0.405	2.7	0.437	-	-
Toce a Vannino	-	49	-	-	13.8	0.407	10.6	0.437	10.7	0.146
Toce ad Agaro	-	63	-	-	12.7	0.406	12.5	0.437	12.8	0.144
Toce ad Alpe Cavalli	-	-	-	-	22.6	0.407	20.7	0.437	-	-
Toce ad Avino	-	35	-	-	8.1	0.400	5.3	0.437	5.5	0.184
Trebbia a Due Ponti	20	20	256.2	0.149	222.7	0.421	222.0	0.437	248.0	0.180
Trebbia a S.Salvatore	17	21	910.5	0.120	1048.0	0.430	872.0	0.437	1261.0	0.179
Trebbia a Valsigiara	27	27	474.4	0.114	506.7	0.426	467.5	0.437	462.2	0.198
Valtoggia	-	-	-	-	7.8	0.416	9.5	0.437	-	-
Varaita a Castello	56	56	18.8	0.057	24.4	0.426	28.9	0.437	19.0	0.156

Denominazione	n	n_g	$Q_{ind,oss}$	CV_{oss}	$Q_{ind,regr}$	$CV_{Qind,regr}$	$Q_{ind,raz}$	$CV_{Qind,raz}$	$Q_{ind,g}$	$CV_{Qind,g}$
Varaita a Rore	58	62	41.3	0.153	77.6	0.430	93.3	0.437	38.2	0.166
Vermenagna a Limone	-	16	-	-	54.0	0.421	65.6	0.437	41.8	0.188
Vobbia a Vobbietta	14	13	87.6	0.298	105.5	0.416	107.8	0.437	55.3	0.293

Parte II:

Stima della curva di crescita

6 Metodi di stima della curva di crescita: stima degli L-coefficienti

Nelle classiche procedure di analisi di frequenza regionale la determinazione delle curve di crescita viene effettuata suddividendo i dati in regioni statisticamente omogenee ed attribuendo ad ogni regione omogenea una diversa curva di crescita, ottenuta adattando ai dati opportune distribuzioni la cui forma si ritiene nota a meno di un numero finito p di parametri incogniti $\theta_1,\ldots,\theta_p$. Ne consegue che $K_j(T)$ può essere intesa come $K_j(T;\ \theta_1,\ldots,\theta_p)$, con i parametri $\theta_1,\ldots,\theta_p$ che assumono valori diversi in ognuna delle sottoregioni considerate.

Hosking e Wallis (1997) suggeriscono di stimare i *p* parametri incogniti della distribuzione tramite il metodo degli *L*-momenti (Appendice C), che, al contrario dell'usuale metodo dei momenti, risulta particolarmente indicato per la stima da campioni poco numerosi, ed è meno soggetto dell'altro alla distorsione di stima.

Il metodo degli *L*-momenti consiste nell'esprimere i parametri della distribuzione di probabilità (ossia della curva di crescita) in funzione degli *L*-momenti della distribuzione, $L_1,\ldots,L_p$, ottenendo $K_j(T; L_1,\ldots,L_p)$, e successivamente nell'eguagliare i *p* *L*-momenti campionari ai corrispondenti *L*-momenti della distribuzione. Se si considera ancora che la curva di crescita è adimensionale, per essa si ha che L_1=1 (L_1 rappresenta la media della curva di crescita). Tenendo conto che i modelli probabilistici comunemente utilizzati nelle analisi regionali hanno di solito 3 (*p* = 3) oppure 4 (*p* = 4) parametri risulta conveniente esprimere la curva di crescita tramite descrittori statistici adimensionali, definiti come (si veda l'Appendice C),

$$L_{cv} = \frac{L_2}{L_1}, \qquad L_{ca} = \frac{L_3}{L_2}, \qquad L_{kur} = \frac{L_4}{L_2}$$

adatti a rappresentare rispettivamente la variabilità, l'asimmetria e l'appiattimento della curva di crescita. Si perviene pertanto ad una rappresentazione della curva di crescita del tipo

$$K_j(T; L_{cv}, L_{ca})$$

se p =3, oppure

$$K_j(T; L_{cv}, L_{ca}, L_{kur})$$

quando p =4.

Uno svantaggio dell'analisi regionale classica appena descritta è la necessità di ricorrere ad una suddivisione in regioni al cui interno si suppone che possa applicarsi un'unica curva di crescita. Sia quando le sottoregioni sono costituite su base geografica, sia quando si aggregano i bacini in base ai valori di opportuni descrittori geomorfoclimatici (cluster analysis) si incorre nel problema che la curva di crescita presenta brusche discontinuità nel passaggio da una regione ad un'altra. Inoltre non sono sempre chiare le modalità di definizione del numero di sottoregioni da considerare nell'analisi. Infine, le procedure di verifica dell'omogeneità statistica dei campioni raggruppati nelle sottoregioni sono spesso inefficienti (*Viglione et al.,* 2007) e, qualora venissero impiegate le procedure più selettive, si verrebbero ad identificare regioni troppo piccole, ossia con troppi pochi dati, perché si possa garantire una buona accuratezza delle stime con tempi di ritorno elevati.
I metodi di analisi regionale basati sulla teoria dell'Area di Influenza o "Region of Influence" tentano di superare questi svantaggi: quando si utilizzano questi metodi il numero di sottoregioni non viene infatti fissato a priori, ma ogni sito considerato viene attribuito ad una diversa sottoregione, definita associando al bacino considerato i bacini strumentati ad esso più simili (la similitudine viene di solito valutata in base ai valori dei descrittori geomorfoclimatici). Anche l'utilizzo di questi metodi non consente comunque di risolvere completamente i problemi legati alla presenza di brusche discontinuità nella curva di crescita: infatti, supponendo di muoversi lungo un'asta fluviale da monte verso valle, ci si troverà sicuramente in una situazione in cui incrementando di poco l'area del bacino si otterranno descrittori che verranno attribuiti ad una sottoregione diversa da quella attribuita al bacino più a monte. Tale diversità, anche solo minima, della sottoregione, modificherà i valori dei parametri $\theta_1,\ldots,\theta_p$ inducendo la discontinuità di cui sopra. E' chiaro che la discontinuità sarà meno brusca che nel caso di regioni fisse, dal

momento che solo una parte del campione si modifica nel passaggio da una sottoregione ad un'altra, ma comunque rimane una irregolarità spaziale nella curva di crescita, non facilmente spiegabile dal punto di vista fisico. Inoltre il metodo dell'Area di Influenza richiede di rideterminare la regione di appartenenza del sito di interesse per ogni bacino non strumentato, cosa che complica in maniera rilevante l'applicazione operativa per un utente esterno quando non si voglia ricorrere alla predisposizione di procedure automatiche (software) per la definizione della curva di crescita.

Per superare le difficoltà di cui sopra nel presente lavoro si procederà alla stima della curva di crescita con un metodo che si basa sui seguenti presupposti:

- variabilità continua della curva di crescita nello spazio. Si ipotizza che la curva di crescita sia diversa da sezione a sezione, ossia non si ricorre alla classica suddivisione della regione di interesse in sottoregioni supposte omogenee.
- Rappresentazione della curva di crescita in funzione di due *L*-coefficienti, $K_j(T; L_{cv}, L_{ca})$; si suppone dunque di utilizzare modelli probabilistici con 3 parametri. Quando si dovessero utilizzare modelli a 4 parametri, si fissa un unico valore di L_{kur} su tutta la macroregione di interesse.
- Stima degli *L*- coefficienti L_{cv} e L_{ca} tramite regressioni multiple, utilizzando come variabili dipendenti i descrittori geomorfoclimatici già utilizzati per la stima della piena indice. Verranno considerate anche procedure alternative basate sulla stima diretta (da dati di portata al colmo) o indiretta (da dati di portata giornaliera) di L_{cv} e L_{ca}.
- Scelta del modello probabilistico di rappresentazione della curva di crescita effettuata solo a valle della stima degli *L*-coefficienti. Il modello viene selezionato tra le seguenti 7 distribuzioni di probabilità: la distribuzione di Pareto generalizzata, la distribuzione Generalizzata del Valore Estremo (GEV), la distribuzione Logistica generalizzata, la distribuzione Lognormale a 3 parametri, la distribuzione Gamma a 3 parametri, la distribuzione Kappa, la

distribuzione del valore estremo a doppia componente TCEV (vedere Appendice D).

Nel presente capitolo si analizzano in dettaglio le modalità di stima degli *L*-coefficienti (6.1) per ciascuna delle sezioni considerate, considerando anche i metodi di stima per le stazioni prive di misure dirette. Le procedure vengono applicate alla regione di interesse nel successivo Capitolo 7, per poi passare nei successivi capitoli alla scelta della distribuzione di probabilità (Capitolo 8) ed alla applicazione delle procedure ai bacini ricadenti nella regione di interesse (Capitolo 9).

In particolare si considera la possibilità di stimare gli *L*-coefficienti L_{cv} e L_{ca} a partire:

- dal campione di osservazioni al colmo a disposizione, tenendo opportunamente conto di valori storici occasionali;
- da modelli di regressione;
- dalla stima derivata dai corrispondenti *L*-coefficienti calcolati sugli estremi giornalieri.

Come accennato, il valore di L-kurtosis viene invece assunto costante e pari alla media degli L-kurtosis empirici, calcolati sui dati osservati.

A ciascun metodo di stima degli L-coefficienti viene affiancato un criterio utile alla valutazione dell'incertezza con cui si effettua la stima stessa. La statistica utilizzata a tale scopo è la deviazione standard $\sigma_{L_{coef}}$, che viene opportunamente definita per ciascun metodo sviluppato.

6.1 Stima empirica basata sulle osservazioni storiche

6.1.1 Metodo di stima in presenza di valori storici occasionali

Considerato un sito strumentato, per il quale si dispone di una serie storica di portate al colmo più o meno consistente, gli L-coefficienti possono essere calcolati in funzione dei momenti pesati in probabilità b_0, b_1, b_2 e b_3, come spiegato in Appendice C. In particolare risulta:

$$L_{cv} = \frac{2 \cdot b_1}{b_0} - 1 , \qquad (6.1)$$

$$L_{ca} = \frac{2 \cdot (3b_2 - b_0)}{2b_1 - b_0} - 3 , \tag{6.2}$$

$$L_{kur} = \frac{5 \cdot [2 \cdot (2b_3 - 3b_2) + b_0]}{2b_1 - b_0} + 6 , \tag{6.3}$$

dove L_{cv}, L_{ca} ed L_{kur} rappresentano, rispettivamente, il coefficiente di L-variazione, di L-asimmetria e di L-kurtosis.

Nel caso in cui il campione di n_j osservazioni a disposizione per la j-esima sezione sia relativo a un periodo sistematico di misurazioni, il valore dei momenti pesati in probabilità (Probability Weigthed Moments, PWM) viene calcolato come:

$$b_r = \frac{1}{n} \sum_{i=1}^{n} Q_{(i)} \qquad \text{con } r=0$$

$$b_r = \frac{1}{n} \sum_{i=1}^{n} \frac{(i-1) \cdot (i-2) \cdot \ldots \cdot (i-r)}{(n-1) \cdot (n-2) \cdot \ldots \cdot (n-r)} \cdot Q_{(i)} \qquad \text{con } r=1,\ldots,3 \tag{6.4}$$

dove $Q_{(i)}$ rappresenta l'i-esimo valore di portata al colmo del campione ordinato in senso crescente.

Qualora le serie storiche siano state integrate con informazioni storiche occasionali (come descritto nel paragrafo 2.1), relative ad eventi alluvionali di particolare rilevanza occorsi in anni recenti o durante interruzioni del periodo sistematico di misurazione, il valore dei momenti pesati in probabilità non viene più calcolato tramite la relazione (6.4), in quanto il campione statistico non è più costituito da osservazioni effettuate con sistematicità.

Per tener conto con diverso peso della differente natura degli eventi sporadici di intensità eccezionale, si individua per essi un valore soglia Q_{soglia}, scelto pari al più piccolo dei valori relativi agli eventi storici occasionali considerati. E' importante sottolineare come tale valore soglia corrisponda a quello definito per la stima empirica della piena indice nel paragrafo 2.1 e a quello considerato per l'attribuzione delle plotting positions ai dati osservati seguendo il metodo di Hirsch (Appendice B, Volume II).

Una volta fissato il valore della soglia è possibile individuare quali sono gli n_{sotto_soglia} dati al di sotto di Q_{soglia} e gli n_{sopra_soglia} valori al di sopra di essa. Il calcolo dei momenti pesati in probabilità viene effettuato attribuendo un peso

maggiore ai dati sistematici ed un peso inferiore agli eventi occasionali significativi.
La stima dei PWM viene dunque effettuata a partire dalla somma di due contributi, come definito da *Wang* (1990):

$$b_r = b'_r + b''_r \,, \tag{6.5}$$

in cui b'_r è il contributo corrispondente ai valori sottosoglia del campione sistematico di misurazioni, mentre b''_r riguarda l'apporto dovuto a tutti i valori di portata che superano la Q_{soglia} prefissata.
In particolare, b'_r si calcola tramite una relazione analoga alla (6.4):

$$b'_r = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{(i-1)\cdot(i-2)\cdot\ldots\cdot(i-r)}{(n-1)\cdot(n-2)\cdot\ldots\cdot(n-r)} \cdot Q^*_{(i)} \qquad \text{con } r=0,\ldots,3 \tag{6.6}$$

dove

$$Q^*_{(i)} = \begin{cases} Q_{(i)} & Q_{(i)} < Q_{soglia} \,, \\ 0 & Q_{(i)} \geq Q_{soglia} \,. \end{cases}$$

Il contributo derivante dagli n_{sopra_soglia} eventi viene invece pesato sulla numerosità equivalente n_{eq}, ovvero sulla lunghezza complessiva del lasso temporale coperto dalla serie storica completa:

$$b''_r = \frac{1}{n_{eq}} \cdot \sum_{i=1}^{n} \frac{(i-1)\cdot(i-2)\cdot\ldots\cdot(i-r)}{(n_{eq}-1)\cdot(n_{eq}-2)\cdot\ldots\cdot(n_{eq}-r)} \cdot Q^{**}_{(i)} \tag{6.7}$$

in cui $r=0,\ldots,3$ e con

$$Q^{**}_{(i)} = \begin{cases} 0 & Q_{(i)} < Q_{soglia} \,, \\ Q_{(i)} & Q_{(i)} \geq Q_{soglia} \,. \end{cases}$$

6.1.2 Valutazione dell'incertezza di stima

Una volta stimati i coefficienti di L-variazione e di L-asimmetria risulta importante valutare l'incertezza associata alla stima. Una statistica adeguata a

descrivere l'incertezza di stima è la varianza associata alla stima stessa. La valutazione della varianza associata agli L-coefficienti è stata effettuata facendo riferimento ai risultati ottenuti da *Viglione* (2007); in particolare per L_{cv} vale che

$$\sigma^2_{L_{cv}} = \frac{\left(0.9 \cdot L_{cv}\right)^2}{n}, \tag{6.8}$$

mentre per L_{ca}:

$$\sigma^2_{L_{ca}} = \frac{\left(0.45 + 0.6 \cdot \left|L_{ca}\right|\right)^2}{n}, \tag{6.9}$$

dove n rappresenta la dimensione campionaria associata alla j-esima sezione esaminata.
Le stime campionarie di L-CV ed L-CA sono inoltre tra loro correlate. Facendo ancora riferimento a *Viglione* (2007) si ha che il coefficiente di correlazione tra i due stimatori può essere approssimato tramite la relazione

$$\rho_{L_{CV} L_{CA}} = \frac{1 - e^{-5 L_{CA}}}{1 + e^{-5 L_{CA}}}. \tag{6.10}$$

E' importante sottolineare che per le sezioni le cui serie siano state integrate con eventi occasionali significativi, la valutazione dell'incertezza di stima associata a L_{cv} o a L_{ca} viene effettuata applicando le relazioni (6.8) e (6.9) esclusivamente al campione sistematico di misurazioni. Questo comporta che il valore di L_{cv}, L_{ca} ed n da inserire nelle relazioni venga calcolato sul campione depurato degli eventi occasionali significativi, che, altrimenti, potrebbero distorcere il risultato.

6.2 Stima statistica attraverso modelli di regressione multipla

6.2.1 Definizione dei modelli multiregressivi

I metodi multiregressivi risultano utili qualora si voglia stimare una variabile idrologica in siti sprovvisti di osservazioni. Nel presente lavoro l'approccio multiregressivo è già stato impiegato per la stima della piena indice (paragrafi 2.2 e 3.2); di seguito si intende ricorrere all'analisi multiregressiva per la stima dei coefficienti di *L*-variazione e di *L*-asimmetria.

L'approccio multiregressivo viene utilizzato per legare i primi due *L*-coefficienti alle caratteristiche morfometriche e climatiche del bacino, descritte nel Volume II. A tal fine sono state implementate regressioni multiple con un numero variabile di variabili indipendenti. Il modello implementato è corrispondente a quello introdotto per la piena indice nel paragrafo 2.2, ossia:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{6.11}$$

dove:

- $\boldsymbol{y}$ è il vettore *(s x 1)* delle variabili dipendenti y_j di ciascuna stazione;
- β indica il vettore dei *(p+1 x 1)* coefficienti di regressione, stimati mediante la tecnica dei minimi quadrati ordinari;
- $\boldsymbol{X}$ rappresenta la matrice *(s x p+1)* delle variabili esplicative da cui vengono fatti dipendere gli *L*-coefficienti. La prima colonna di $\boldsymbol{X}$ è costituita da un vettore unitario in modo che il primo dei coefficienti nel vettore β sia il termine noto della regressione;
- ε è il vettore *(s x 1)* dei residui della regressione.

L'analisi regressiva è stata effettuata distintamente per i primi due *L*-coefficienti esaminati. Per entrambi è stata effettuata una stima nel campo lineare, basata sul metodo dei minimi quadrati ordinari (Ordinary Least Squares, O.L.S.) descritto in Appendice A; in particolare:

- per il coefficiente di L-variazione si è posto $\boldsymbol{y} = \boldsymbol{L_{CV}}$ in relazione alla matrice $\boldsymbol{X}$ dei parametri geomorfologici;

- per il coefficiente di L-asimmetria si è posto $\boldsymbol{y} = \boldsymbol{L_{CA}}$ in relazione alla matrice $\boldsymbol{X}$ dei descrittori geomorfologici.

6.2.2 Scelta dei modelli di regressione e verifiche di adeguatezza

Ciascun modello multiregressivo consente di legare la variabile dipendente considerata a diverse caratteristiche morfologiche e climatiche dei bacini. Poiché, come si è visto nel Volume II, ciascun bacino è caratterizzato da 33 descrittori, è indispensabile applicare dei criteri che consentano di determinare quante e quali variabili esplicative debbano essere utilizzate per la stima di L_{cv} ed L_{ca}. Come nel caso della stima della piena indice, per ciascun modello multiregressivo si ottengono circa 10^{11} possibili combinazioni con un numero di parametri variabile e compreso tra $p = 1$ (un solo regressore) e $p = 33$. Tenuto conto della parsimonia che deve caratterizzare un buon modello multiregressivo si ritiene ragionevole limitare le possibili combinazioni a relazioni con non più di 4 variabili esplicative, riducendo così il numero delle combinazioni a 163680. Da questo insieme vengono successivamente estratte le migliori 5 in termini di R^2_{adj} ottenute con p compreso tra 1 e 4. Questo significa che per ciascun modello vengono proposte 20 relazioni, ottenute in base ad un numero di variabili esplicative compreso tra 1 e 4, caratterizzate dalla migliore capacità descrittiva. Per ognuna delle 20 relazioni esaminate è stata verificata la significatività di ciascuna variabile esplicative coinvolta. A tal fine, analogamente a quanto fatto per la stima della piena indice, è stato applicato il test della t di Student (vedere Appendice A) per il quale è stato fissato un livello di significatività $\alpha = 5\%$. Contestualmente è stata effettuata la verifica dell'assenza di multicollineartità, in modo da eliminare i modelli multiregressivi meno stabili. Ne consegue, dalla verifica di adeguatezza, che il numero di modelli selezionati per una certa variabile dipendente si può ridurre rispetto ai 20 considerati inizialmente.
Per quanto riguarda l'analisi dei diagrammi diagnostici utili all'analisi dei risultati ottenuti si rimanda a quanto esposto per la piena indice nel paragrafo 2.2.5.

6.2.3 Analisi multiregressiva pesata

Tutte le considerazioni fin qui effettuate sono basate sull'ipotesi che la varianza dei residui $\hat{\mathbf{y}} - \mathbf{y}$ risulti circa costante per tutte le osservazioni del

campione (ipotesi di omoschedasticità). Ciò consente di stimare i coefficienti $\boldsymbol{\beta}$ di regressione tramite il metodo dei minimi quadrati ordinari (Ordinary Least Squares, O.L.S.) descritto in Appendice A. Tuttavia si deve sottolineare che, nel caso degli *L*-coefficienti, i valori di L_{cv} e di L_{ca} utilizzati per tarare i coefficienti di regressione sono a loro volta delle stime, la cui varianza osservata cambia, secondo le relazioni (6.8) e (6.9), in ragione del numero di osservazioni n_j disponibili per il sito *j*. In particolare, dove le serie storiche sono più consistenti la varianza è minore. Per questo motivo, la determinazione delle variabili esplicative da utilizzare per ciascun modello dovrebbe essere effettuata dando un peso maggiore ai siti con più informazioni e attribuendo meno importanza alle sezioni con minore disponibilità di dati. Questo comporta che la selezione delle variabili esplicative e la stima dei coefficienti di regressione venga fatta ricorrendo al metodo dei minimi quadrati pesati (Weigthed Least Squares, W.L.S.).
L'applicazione del metodo dei minimi quadrati pesati viene effettuata seguendo i criteri esposti in Appendice A, cioè considerando un peso w_j inversamente proporzionale alla varianza associata agli *L*-coefficienti osservati. In particolare:

- $w_j = \dfrac{n_j}{\left(0.9 \cdot L_{cv,j}\right)^2}$ quando la variabile dipendente $\boldsymbol{y}$ è rappresentata dal coefficiente di *L*-variazione;

- $w_j = \dfrac{n_j}{\left(0.45 + 0.6 \cdot \left|L_{ca,j}\right|\right)^2}$ quando la variabile dipendente $\boldsymbol{y}$ è rappresentata dal coefficiente di *L*-asimmetria.

6.2.4 Valutazione dell'incertezza di stima

Obiettivo del paragrafo è la descrizione dei metodi utilizzati per determinare la varianza di stima $\sigma^2_{\hat{y}_j}$ associata a ciascun modello di regressione costruito per il generico stimatore $\hat{y}_j$. Una volta tarata una regressione a partire da *s* osservazioni della variabile esaminata (una per

ciascuna stazione), tramite la $\sigma^2_{\hat{y}_j}$ si potrà calcolare una incertezza di stima associata ad ogni punto in cui viene effettuata la stima.

Poiché la stima $\hat{y}$ delle variabili viene effettuata tramite un modello multiregressivo tarato a partire dalle variabili osservate *y*, che sono a loro volta variabili casuali, l'incertezza ad essa associata risulta composta da due fattori (*Stedinger and Tasker,* 1985; 1986): l'incertezza campionaria, relativa alle variabili osservate, e l'incertezza del modello, relativa alla regressione utilizzata per la stima:

$$\sigma^2_{\text{regr},j} = \sigma^2_{\text{camp},j} + \sigma^2_{\text{model}} . \qquad (6.12)$$

I valori campionari della variabile *y* (ossia degli L-coefficienti) in corrispondenza del sito *j*-esimo, non sono altro che delle stime empiriche, effettuate a partire dai dati osservati. La componente di varianza campionaria nella (6.12), $\sigma^2_{\text{camp},j}$, può essere stimata facendo ricorso alle equazioni (6.8) o (6.9), a seconda che si stia considerando la stima dell'L-CV o dell'L-CA. La componente di errore di modello, σ^2_{model}, è invece indipendente dallo specifico sito nel quale si effettua la stima, ossia è indipendente dal pedice *j*, e rappresenta il valore incognito che occorre stimare per valutare l'effettiva incertezza da associare alla stima multiregressiva, ossia si ha che $\sigma^2_{\hat{y}_j} = \sigma^2_{\text{model}}$.

La stima di σ^2_{model} può essere effettuata con metodi diversi (*Stedinger and Tasker,* 1986). Nel presente lavoro si è adottata una tecnica basata sulla funzione di verosimiglianza: si suppone che i residui della regressione, $\varepsilon = \hat{y} - y$, abbiano distribuzione normale con media nulla e varianza (che cambia da sito a sito) data dall'equazione (6.12). Il logaritmo naturale della funzione di verosimiglianza risulta essere

$$\log L = C + \sum_{j=1}^{s} 0.5 \log\left(\sigma^2_{\text{regr},j}\right) + 0.5 \sum_{j=1}^{s} \left(\frac{\varepsilon_j^2}{\sigma^2_{\text{regr},j}} \right) \qquad (6.13)$$

dove:

- *C* è una costante;

- *s* è il numero di sezioni utilizzate per tarare il modello multiregressivo;
- $\varepsilon_j = \hat{y}_j - y_j$ rappresenta il *j*-esimo residuo della regressione.

Massimizzando numericamente la funzione di log-verosimiglianza nell'equazione (6.13) si ottiene una stima di $\sigma^2_{\mathrm{model}}$ che può essere utilizzata per stimare $\sigma^2_{\hat{y}_j}$, ossia per quantificare l'incertezza associata alla stima multiregressiva di L-CV o L-CA.
Anche nel caso di utilizzo del metodo dei minimi quadrati pesati, per tarare i coefficienti della regressione multipla la procedura di stima dell'errore del modello può avvenire in maniera analoga, ossia utilizzando l'equazione (6.10) con $\varepsilon_j = \hat{y}_j - y_j$ e $\hat{y}_j$ corrispondente al residuo del modello multiregressivo pesato. La stima di $\sigma^2_{\mathrm{model}}$ che si ottiene dalla (6.13) può dunque essere usata per confrontare modelli multiregressivi che utilizzano descrittori diversi per confrontare regressioni pesate e non; il modello che fornisce la stima più bassa di $\sigma^2_{\mathrm{model}}$ è quello cui si associa una minore incertezza di stima ed è quindi quello che verrà scelto nei successivi capitoli.
Prima di procedere con la descrizione delle modalità di stima di L-CV ed L-CA sono necessarie alcune precisazioni:

- la procedura qui descritta per stimare $\sigma^2_{\mathrm{model}}$ è una procedura approssimata, in quanto richiede di stimare prima i parametri della regressione (per ottenere $\hat{y}_j$), e solo successivamente $\sigma^2_{\mathrm{model}}$. In realtà $\sigma^2_{\mathrm{model}}$ andrebbe stimato contemporaneamente agli altri parametri, ad esempio massimizzando la funzione di verosimiglianza rispetto a tutti i parametri contemporaneamente (*Stedinger and Tasker,* 1985; 1986). Questo comporterebbe però una notevole complicazione delle procedure di stima ed un incremento dei tempi computazionali, ed impedirebbe l'utilizzo di procedure per la scelta automatica dei descrittori le quali, come detto, richiedono di considerare 163680 modelli regressivi.
- A rigore, si sarebbe dovuto utilizzare la procedura qui descritta anche per la stima della piena media: infatti, anche in tal caso le regressioni sono state tarate utilizzando variabili esplicative casuali, le medie

campionarie. Tuttavia nel caso delle medie si ha che i valori di $\sigma^2_{\text{camp},j}$ nell'equazione (6.12) risultano trascurabili rispetto a $\sigma^2_{\text{mod el}}$ quando si utilizzano serie con almeno 10 dati, cosa che consente di determinare l'incertezza di stima delle regressioni semplicemente considerando i residui delle stesse, evitando le complicazioni qui descritte.

- La procedura di stima viene applicata prima a L-CV e successivamente a L-CA. Tra i descrittori impiegati per stimare L-CA è quindi possibile considerare anche lo stimatore di L-CV ottenuto tramite analisi multiregressiva.

Nel caso della stima tramite analisi multiregressiva gli stimatori di LCV ed LCA possono ritenersi in prima approssimazione scorrelati ($\rho_{L_{CV}L_{CA}} = 0$), dal momento che le rispettive stime sono basate su regressori diversi.

6.3 Stima basata sull'impiego di dati di portata massima giornaliera

6.3.1 Introduzione metodologica

La valutazione degli *L*-coefficienti nei bacini strumentati, come si è visto nel paragrafo 6.1, viene effettuata a partire dalla serie storica delle portate al colmo disponibile per ciascuna sezione di interesse. Per alcuni bacini i dati di portata al colmo sono poco numerosi o assenti, mentre è noto un buon numero di osservazioni relative al massimo annuo della portata media giornaliera. Tenuto conto di tale disponibilità di informazioni, ci si pone l'obiettivo di ricavare delle relazioni che consentano di stimare i primi due *L*-coefficienti delle portate al colmo tramite i corrispondenti *L*-coefficienti delle portate estreme giornaliere. Le relazioni regressive possono includere anche altri descrittori geomorfologici, ossia:

$$L_{cv} = f\left(L_{cv,g}, A, H_{media}, \ldots\right), \tag{6.14}$$

$$L_{ca} = f\left(L_{ca,g}, A, H_{media}, \ldots\right), \tag{6.15}$$

L'individuazione di relazioni analoghe alla (6.14) e (6.15) è stata effettuata ricorrendo all'analisi multiregressiva, implementando il modello (6.11), basato sul metodo dei minimi quadrati ordinari, e considerando le stesse variabili dipendenti ***y*** utilizzate per la stima nei bacini non strumentati (paragrafo 6.2.1).

La differenza rispetto all'analisi multiregressiva sviluppata nel paragrafo precedente sta nella diversa definizione della matrice ***X***. Infatti nel caso dei bacini non strumentati la stima degli L-coefficienti può essere effettuata esclusivamente in funzione dei descrittori geomorfologici di bacino, che vengono raccolti nella matrice ***X***. Nel caso di bacini per cui siano note le serie di portata estrema giornaliera, la stima degli *L*-coefficienti coinvolge anche i rapporti tra gli *L*-momenti delle portate estreme giornaliere; in questo caso la matrice ***X*** comprende sia i parametri geomorfologici di bacino che gli *L*-coefficienti giornalieri.

I criteri applicati per la scelta e la definizione dei modelli multiregressivi sono gli stessi richiamati nel paragrafo 6.2.2 e già applicati per la stima della piena indice. Le verifiche di adeguatezza dei modelli, quali il test della *t* di Student e l'analisi di multicollinarità, sono effettuate secondo i medesimi principi utilizzati per la media. Lo stesso vale per la valutazione della capacità descrittiva dei modelli e per l'analisi dei diagrammi diagnostici.

6.3.2 Analisi multiregressiva pesata

Tutte le considerazioni fin qui effettuate sono basate sull'ipotesi di omoschedasticità, ovvero sull'ipotesi che la varianza dei residui $\hat{\mathbf{y}} - \mathbf{y}$ risulti circa costante per tutte le osservazioni del campione. Tuttavia è importante ricordare che gli *L*-coefficienti utilizzati per individuare i modelli multiregressivi presentano una varianza che cambia in funzione del numero di osservazioni n_j disponibili in ciascuna sezione di interesse. In particolare, dove le serie storiche sono più consistenti la varianza è minore. Per questo motivo si ritiene che l'individuazione delle variabili esplicative di ciascun modello debba essere effettuata dando un peso maggiore ai siti con più informazioni e attribuendo meno importanza alle sezioni con minore disponibilità di dati. Questo comporta che la stima delle variabili esplicative e dei coefficienti venga fatta ricorrendo al metodo dei minimi quadrati pesati (Weigthed Least Squares, W.L.S.), descritto in Appendice A.

L'applicazione del metodo dei minimi quadrati pesati viene effettuata seguendo i medesimi criteri e adottando gli stessi pesi w_j considerati per la stima degli *L*-coefficienti nei siti non strumentati, come descritto nel paragrafo 6.2.3.

6.3.3 Valutazione dell'incertezza di stima

Poiché la stima $\hat{y}$ delle variabili viene effettuata tramite un modello multiregressivo tarato a partire dalle variabili osservate y, coinvolgendo tra le variabili esplicative anche gli L-coefficienti calcolati sulle portate estreme giornaliere, la valutazione dell'incertezza ad essa associata risulta particolarmente complicata. Infatti, nella valutazione dell'incertezza bisogna tenere debitamente conto sia del fatto che la regressione è tarata su variabili dipendenti che sono variabili casuali (come prima), sia del fatto che anche uno dei descrittori è una variabile casuale (la stima campionaria di LCV o LCA da dati giornalieri).

Definito β_g il coefficiente di regressione relativo al descrittore $L_{CV,g}$ (o $L_{CA,g}$), si ha che la varianza associata alla regressione risulta essere composta da 4 componenti:

$$\sigma^2_{\text{regr},j} = \sigma^2_{\text{camp\_c},j} + \sigma^2_{\text{mod el}} + \beta_g^2 \cdot \sigma^2_{\text{camp\_g},j} - 2\beta_g \sigma_{\text{camp\_c},j} \sigma_{\text{camp\_g},j} \rho_{y,j} . \qquad (6.16)$$

dove

- $\sigma^2_{\text{camp\_c},j}$ è la varianza campionaria del *j*-esimo campione di portate al colmo di piena, determinata utilizzando le relazioni (6.8) e (6.9);
- $\sigma^2_{\text{mod el}}$ è l'errore di modello, ossia la variabile da determinare per definire l'incertezza di stima associata al presente metodo di stima di LCV o LCA;
- $\sigma^2_{\text{camp\_g},j}$ è la varianza campionaria del *j*-esimo campione di portate massime giornaliere, determinata utilizzando ancora le relazioni (6.8) e (6.9) ma con i valori di LCV, LCA, e numerosità campionaria relativi alla serie delle portate giornaliere;
- $\rho_{y,j}$ è il coefficiente di correlazione tra LCV (o LCA) calcolato dai dati giornalieri ed al colmo di piena. Si ha (*Kjedsen and Jones,* 2006) $\rho_{LCV,j} = \rho^2_{gc,j}$ e $\rho_{LCA,j} = \rho^3_{gc,j}$, dove $\rho_{gc,j}$ è il coefficiente di correlazione tra

le serie storiche dei massimi annui di portata giornaliera ed al colmo di piena. Dal momento che i campioni di dati giornalieri e al colmo si sovrappongono solo parzialmente (ovvero ci sono anni in cui è disponibile solo il dato giornaliero, o viceversa), si ha ancora che i coefficienti di correlazione di cui sopra vanno moltiplicati per il termine $\frac{n_{contemp,j}^2}{n_j n_{g,j}}$, dove $n_{contemp,j}$ è il numero di anni in cui sono disponibili sia il dato giornaliero che quello al colmo, n_j rappresenta la numerosità campionaria dei colmi di piena, e $n_{g,j}$ la numerosità della serie giornaliera.

La stima di $\sigma_{\mathrm{mod\,el}}^2$ si effettua, come prima, massimizzando la funzione di verosimiglianza (6.16), in cui si pone $\sigma_{\hat{y}_j}^2$ come dalla (6.19), ossia per LCV

$$\sigma_{regr,LCV_j}^2 = \frac{\left(0.9L_{CV,j}\right)^2}{n_j} + \sigma_{\mathrm{mod\,el}}^2 + \beta_g^2 \cdot \frac{\left(0.9L_{CV\_g,j}\right)^2}{n_{g,j}} + \\ - 2\beta_g \frac{0.81L_{CV,j}L_{CV\_g,j}}{\sqrt{n_j n_{g,j}}} \frac{n_{contemp,j}^2}{n_j n_{g,j}} \rho_{gc,j}^2 \tag{6.17}$$

e per LCA

$$\sigma_{regr,LCV_j}^2 = \frac{\left(0.45+0.6\left|L_{CA,j}\right|\right)^2}{n_j} + \sigma_{\mathrm{mod\,el}}^2 + \beta_g^2 \cdot \frac{\left(0.45+0.6\left|L_{CA\_g,j}\right|\right)^2}{n_{g,j}} + \\ - 2\beta_g \frac{\left(0.45+0.6\left|L_{CA,j}\right|\right)\left(0.45+0.6\left|L_{CA\_g,j}\right|\right)}{\sqrt{n_j n_{g,j}}} \frac{n_{contemp,j}^2}{n_j n_{g,j}} \rho_{gc,j}^3 \tag{6.18}$$

Come in precedenza, anche nel caso di utilizzo del metodo dei minimi quadrati pesati per tarare i coefficienti della regressione multipla la procedura di stima dell'errore di modello può avvenire utilizzando le equazioni (6.16) e (6.19). La varianza complessiva di stima di LCV o LCA dovrà infine tenere conto, anche quando le relazioni regressive sono applicate in previsione, del fatto che uno dei descrittori è una variabile casuale. Si ha allora in questo caso

$$\sigma_{\hat{y}}^2 = \sigma_{\mathrm{model}}^2 + \beta_g^2 \cdot \sigma_{\mathrm{camp\_g},j}^2 \,. \qquad (6.19)$$

Il modello che fornisce la stima più bassa di $\sigma_{\hat{y}}^2$ è quello cui si associa una minore incertezza di stima, ed è quindi quello che verrà scelto nei successivi capitoli.

Anche nel caso della stima tramite estremi giornalieri gli stimatori di LCV ed LCA possono ritenersi in prima approssimazione scorrelati ($\rho_{L_{CV}L_{CA}} = 0$), dal momento che le rispettive stime sono basate su descrittori diversi.

7 Applicazione dei metodi nella regione di interesse

7.1 Stima empirica

Le stime degli L-coefficienti sono state effettuate su tutte le stazioni ricadenti nell'Italia Nord Occidentale e nel Sud della Svizzera per le quali si hanno almeno 4 dati di portata al colmo, per un totale di 113 sezioni idrometriche, in modo da ritenere consistente la valutazione effettuata.

La stima empirica di L_{cv} e L_{ca} viene effettuata a partire dalle serie storiche di portata al colmo disponibili per le diverse stazioni, facendo riferimento alla stima campionaria espressa dalle relazioni (6.1) e (6.2).

Per le stazioni in cui non sono disponibili dati dedotti da Rapporti di Evento o dalle Sezioni F degli Annali Idrologici il valore dei momenti pesati in probabilità, necessario per il calcolo degli L-coefficienti, è stato determinato tramite la (6.4); in caso contrario si è fatto riferimento al metodo di Wang secondo la relazione (6.5).

La valutazione dell'incertezza di stima da associare a L_{cv} ed a L_{ca} è stata effettuata determinando il valore della varianza di stima, secondo il procedimento descritto nel paragrafo 6.1.2.

7.2 Stima regionale (modelli di regressione multipla)

Questo metodo è stato applicato ad un gruppo di 97 stazioni ricadenti nell'Italia Nord Occidentale e nel Sud della Svizzera aventi almeno 10 dati di portata al colmo sul quale tarare le formule multiregressive per la stima di L_{cv} e di L_{ca}, basate su 33 parametri geomorfologici, disponibili per i bacini esaminati.

Per quanto riguarda il coefficiente di L-kurtosis si è assunto, per i motivi spiegati all'inizio della parte II, un valore costante di 0.2274 pari alla media degli L_{kur} osservati.

Il modello di regressione multipla di riferimento è espresso dalla relazione (6.10), nella quale la variabile dipendente $\boldsymbol{y}$ viene fissata pari a:

- $\boldsymbol{L_{cv}}$ per la stima del coefficiente di L-variazione;

- $\boldsymbol{L_{ca}}$ per la stima nel campo lineare del coefficiente di L-asimmetria.

Una volta determinate per ciascuna variabile dipendente ***y*** le possibili combinazioni lineari tra i parametri considerati, sia con il metodo dei minimi quadrati ordinari che dei minimi quadrati pesati si estraggono le 5 migliori soluzioni con non più di 4 variabili esplicative, ordinate secondo il valore di R^2_{adj}, per un totale di 20 relazioni multiregressive. Tra queste vengono selezionate esclusivamente le regressioni i cui parametri risultano significativi per tutte le variabili indipendenti, ossia quelle che superano il test della *t* di Student con livello di significatività α=*5%* e soddisfano l'ipotesi di non multicollinearità.

7.2.1 Stima del coefficiente di L-variazione

Come spiegato nel paragrafo 7.2 la stima del coefficiente di *L*-variazione è stata effettuata nel campo lineare, fissando $\boldsymbol{y} = \boldsymbol{L_{cv}}$. In prima istanza i modelli multiregressivi sono stati determinati applicando il metodo dei minimi quadrati ordinari (O.L.S.), basato sull'ipotesi di omoschedasticità dei residui. I risultati ottenuti con tale metodo sono stati confrontati con quelli derivanti dall'applicazione del metodo dei minimi quadrati pesati (W.L.S.), ottenuti fissando un peso w_j pari a $n_j / \left(0.9 \cdot L_{cv}\right)^2$.

Per confrontare i risultati ottenuti con i due approcci si è ricorso all'analisi delle varianze di stima, calcolate secondo i criteri esposti al paragrafo 6.2.4. Le stime migliori saranno quelle caratterizzate da incertezza minore, ossia con varianza più bassa. I modelli multiregressivi ottenuti per L_{cv} nelle due configurazioni sono riportati in Tabella 7.1 e 7.2.

Variabile dipendente y = L_{cv} (O.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
Y_{sc}, ΔH_1, H_{media}, media_fa	0.93	0.00487
Y_{sc}, ΔH_1, H_{media}, LLDP	0.93	0.0048
Y_{sc}, ΔH_1, H_{media}, lun_asta_princ	0.93	0.0048
Y_{bar}, ΔH_1, H_{media}, media_fa	0.93	0.00501
Y_{bar}, ΔH_1, H_{media}, LLDP	0.93	0.00492
Y_{sc}, H_{media}, sl_med1	0.92	0.00493
Y_{bar}, H_{media}, sl_med1	0.92	0.00507
Y_{sc}	0.91	0.00637
Y_{bar}	0.90	0.0066
lungh_media_vers	0.90	0.00712
H_{media}	0.90	0.0069
R_s	0.90	0.00724

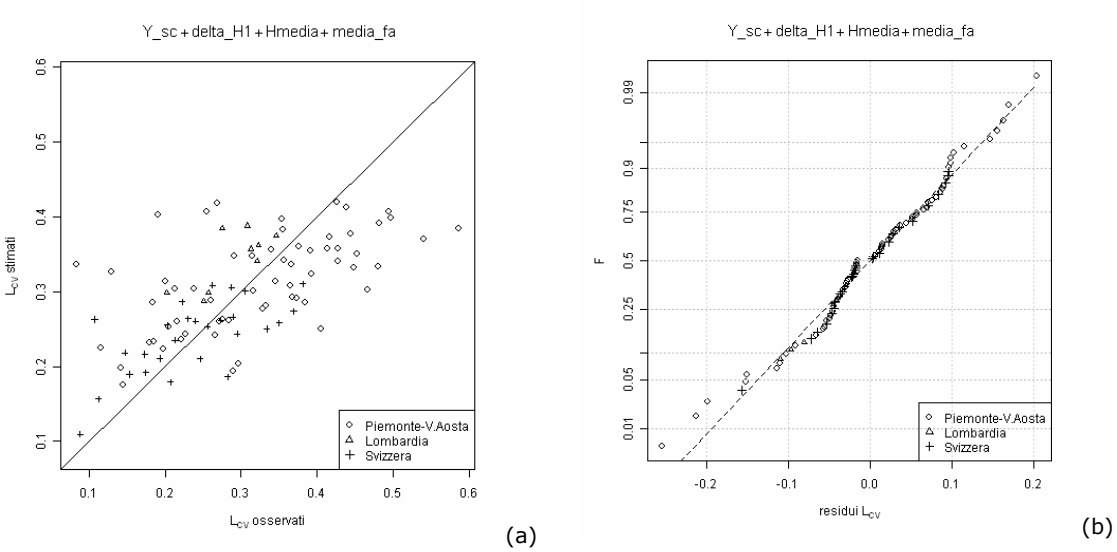
Tabella 7.1. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati ordinari fissando la variabile dipendente *y* pari a L_{cv}.

Dall'analisi dei valori di varianza associati alle diverse relazioni multiregressive emerge chiaramente come il metodo dei minimi quadrati pesati (W.L.S.) conduca a stime affette da incertezza minore rispetto a quelle determinate con il metodo dei minimi quadrati ordinari. Le differenze nei valori di $\sigma^2_{\hat{y}}$ sono significative; in particolare, tra le migliori stime a 4 parametri associate ai due metodi si riscontra una differenza in termini di varianza del 25%.

E' da notare, invece, come il valore del coefficiente di determinazione risulti confrontabile nei due casi. In tal senso si riscontra che il metodo dei minimi quadrati ordinari presenta valori di R^2_{adj} leggermente più alti rispetto a quelli indicati in Tabella 7.2; tuttavia, come detto in precedenza, il vantaggio dei minimi quadrati pesati in termini di incertezza di stima è tutt'altro che trascurabile. Si evidenzia, inoltre, come tra le migliori regressioni a quattro parametri individuate con i minimi quadrati ordinari ci sia sempre Y_{sc} (latitudine della sezione di chiusura) tra le variabili esplicative; al contrario le regressioni a quattro parametri individuate con i W.L.S. presentano sempre, tra le variabili esplicative, dei parametri di elongazione del bacino.

Variabile dipendente y = L_{cv} (W.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
H_{media}, LLDP, lungh_vett_orient, n	0.92	0.00365
H_{media}, lun_asta_princ, lungh_vett_orient, n	0.92	0.00365
H_{media}, lungh_vett_orient, media_fa, n	0.92	0.00396
H_{media}, R_al, F_f, n	0.91	0.00421
lun_asta_princ, lungh_vett_orient, a, n	0.91	0.00458
Y_{sc}, a, n	0.90	0.00444
Y_{bar}, a, n	0.90	0.00453
H_{media}, lun_asta_princ, n	0.90	0.00457
H_{media}, LLDP, n	0.90	0.00458
a, n	0.90	0.00539
H_{media}, n	0.90	0.00488
H_{media}, pend_med_LDP	0.90	0.00644
H_{media}, LLDP	0.90	0.00643
H_{media}, lun_asta_princ	0.90	0.00644
H_{media}	0.87	0.00757
lungh_media_vers	0.87	0.00789
N	0.86	0.00845
Y_{sc}	0.86	0.00722
ΔH_2	0.86	0.00893

Tabella 7.2. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati pesati fissando la variabile dipendente *y* pari a L_{cv}.



(a) (b)

Figura 7.1. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata con il metodo degli O.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia una discreta distribuzione normale dei residui.

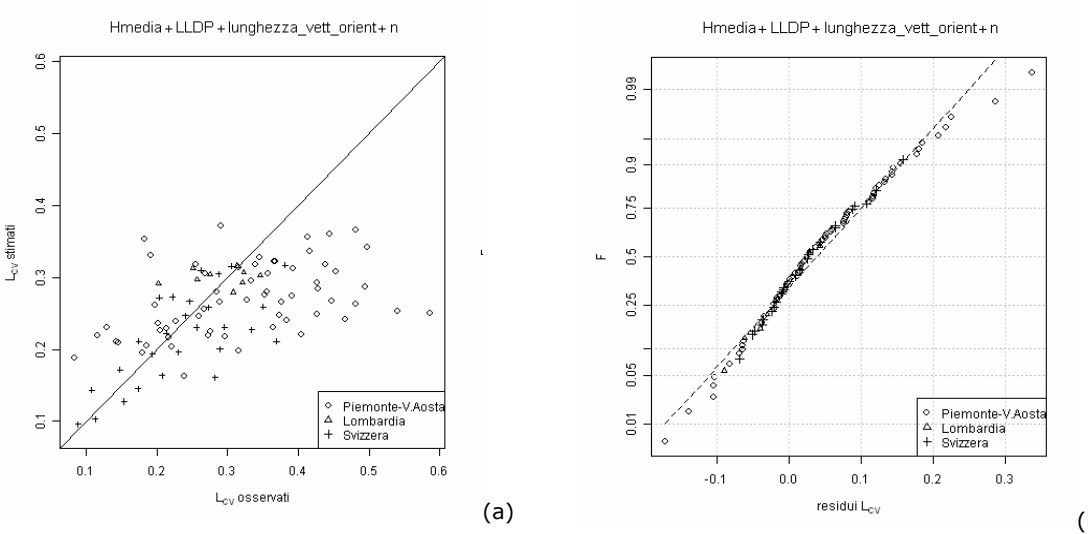


Figura 7.2. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata con il metodo dei W.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

In figura 7.1 e 7.2 vengono riportati i diagrammi diagnostici relativi alla migliore regressione a 4 parametri tarata con il metodo degli O.L.S. e dei W.L.S. Dall'analisi dei grafici a dispersione emerge, in entrambi i casi, un discreto adattamento dei valori stimati a quelli osservati (Fig.7.1.a e 7.2.a). Tuttavia si evidenzia come, per entrambi i casi, ci sia una tendenza alla sottostima per valori di L_{cv} osservati maggiori di 0.4. Inoltre per la regressione tarata con i W.L.S. si riscontra come la distribuzione dei residui sia più regolare rispetto a quella relativa alla regressione tarata con gli O.L.S. (Fig.7.1.b e 7.2.b). Dal confronto delle Figure 7.1a e 7.2a potrebbe inoltre sembrare che i risultati della regressione tarata con il metodo dei minimi quadrati ordinari siano migliori di quelli ottenuti con il metodo dei minimi quadrati pesati. Bisogna tuttavia considerare che l'utilizzo dei pesi richiede che il confronto tra LCV osservati e stimati sia effettuato utilizzando non le variabili originali (LCV), ma le variabili moltiplicate per la radice quadrata dei pesi ($\sqrt{w_i} L_{CV,i}$). Solo utilizzando le variabili opportune si ottengono degli scatter plot (Figura 7.3) congruenti con l'utilizzo dei pesi: ad esempio, le discrepanze relative a sezioni con molti dati disponibili (ad esempio le sezioni svizzere) risultano esaltate in tali diagrammi.

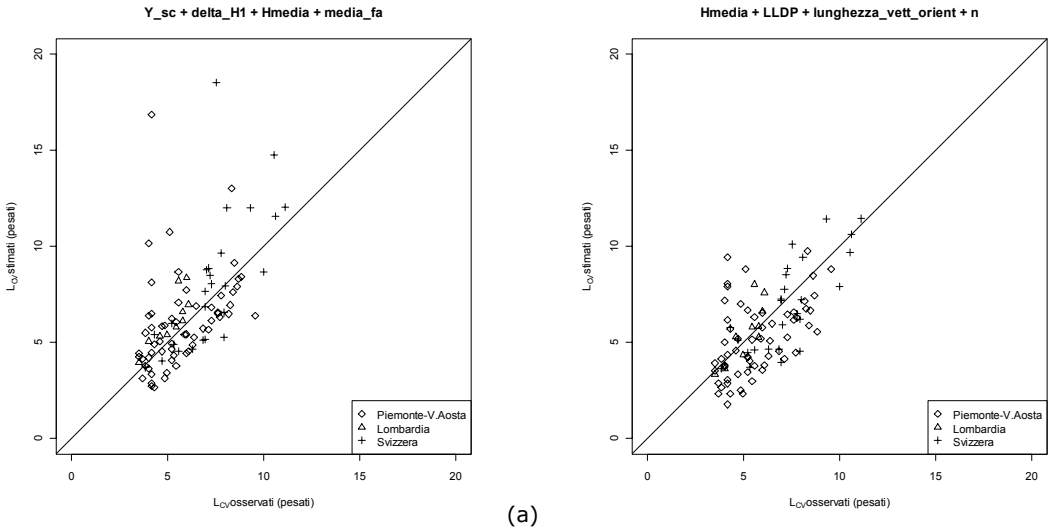


Figura 7.3. Scatter plot per LCV stimato tramite la migliore relazione a 4 parametri tarata con il metodo degli OLS (a) e dei WLS (b). Le variabili in ascissa ed in ordinata corrispondono a $\sqrt{w_j}\,L_{CV,j}$, con $w_j = n_j / \left(0.9 \cdot L_{cv,j}\right)^2$.

Dall'analisi della Figura 7.3 risulta evidente il migliore comportamento della regressione tarata con il metodo dei minimi quadrati pesati, che produce delle stime più raccolte intorno alla bisettrice dello scatter plot.

7.2.2 Stima del coefficiente di L-asimmetria

La stima del coefficiente di *L*-asimmetria è stata effettuata nel campo lineare, fissando $\boldsymbol{y} = \boldsymbol{L_{ca}}$. La determinazione dei migliori modelli regressivi è stata effettuata esaminando il valore di R^2_{adj} associato a ciascuna regressione. Analogamente a quanto fatto per il coefficiente di *L*-variazione, i modelli multiregressivi sono stati determinati applicando prima il metodo dei minimi quadrati ordinari (O.L.S.), basato sull'ipotesi di omoschedasticità dei residui, e successivamente il metodo dei minimi quadrati pesati (W.L.S.), ottenuti fissando un peso w_j pari a $n_j / \left(0.45 + 0.6 \cdot \left|L_{ca}\right|\right)^2$.

Per confrontare i risultati ottenuti con i due metodi si ricorre all'analisi delle varianze di stima, calcolate secondo i criteri esposti al paragrafo 6.2.4. Le stime migliori saranno quelle con incertezza minore, ossia con varianza più bassa. I modelli multiregressivi ottenuti per L_{cv} nelle due configurazioni sono riportati in Tabella 7.3 e 7.4.

Variabile dipendente y = L_{ca} (O.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
Y_{sc}, ΔH_1, lun_asta_princ	0.80	0.00873
X_{sc}, Y_{sc}	0.80	0.00857
X_{sc}, Y_{bar}	0.79	0.00894
Y_{sc}	0.78	0.01089
Y_{bar}	0.78	0.01115
aff	0.76	0.01153
X_{bar}	0.74	0.01466
X_{sc}	0.74	0.01488

Tabella 7.3. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati ordinari fissando la variabile dipendente *y* pari a L_{ca}.

Variabile dipendente y = L_{ca} (W.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
X_{sc},Y_{sc},$L_{cv,stim}$	0.77	0.00638
X_{sc}, Y_{bar}, $L_{cv,stim}$	0.77	0.00645
X_{sc}, Y_{sc}	0.76	0.00764
X_{sc}, $L_{cv,stim}$	0.76	0.00752
Y_{sc}, X_{bar}	0.76	0.00778
X_{sc}, Y_{bar}	0.76	0.00799
X_{bar}, Y_{bar}	0.75	0.00807
Ysc	0.73	0.01015
Y_{bar}	0.73	0.01039
aff	0.72	0.01035
X_{bar}	0.69	0.01276
X_{sc}	0.69	0.01309

Tabella 7.4. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati pesati fissando la variabile dipendente *y* pari a L_{ca}.

Dall'analisi dei valori di varianza associati alle diverse relazioni emerge come il metodo dei minimi quadrati pesati (W.L.S.) conduca a stime affette da incertezza minore rispetto a quelle determinate con il metodo dei minimi quadrati ordinari. In particolare le differenze tra i valori di $\sigma^2_{\hat{y}}$ associati alla migliore relazione a 4 parametri tarata con i due metodi risulta pari al 27%.

Il valore del coefficiente di determinazione R^2_{adj}, invece, risulta leggermente migliore nel caso delle regressioni tarate con gli O.L.S. Tuttavia, come nel caso del coefficiente di *L*-variazione, il vantaggio che emerge in termini di incertezza di stima porta a individuare il metodo ai minimi quadrati pesati come l'approccio migliore per la determinazione dei modelli multiregressivi.
Si nota inoltre come le regressioni tarate con il metodo dei W.L.S. presentino tra le variabili esplicative dei parametri geografici, quali le coordinate della sezione di chiusura; al contrario i modelli di Tabella 7.3 sono espressi quasi sempre in funzione della longitudine o della latitudine separatamente.
In figura 7.4 e 7.5 vengono riportati i diagrammi diagnostici relativi alla migliore regressione a 4 parametri tarata con il metodo degli O.L.S. e dei W.L.S.

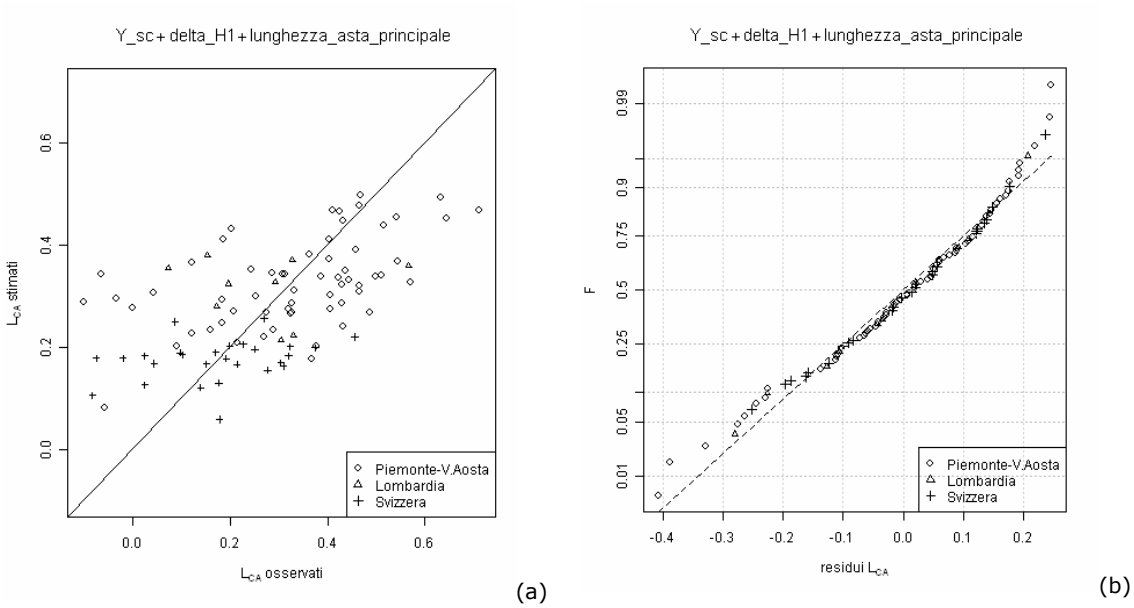


Figura 7.4. Diagrammi diagnostici per la migliore relazione a 3 parametri tarata con il metodo degli O.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

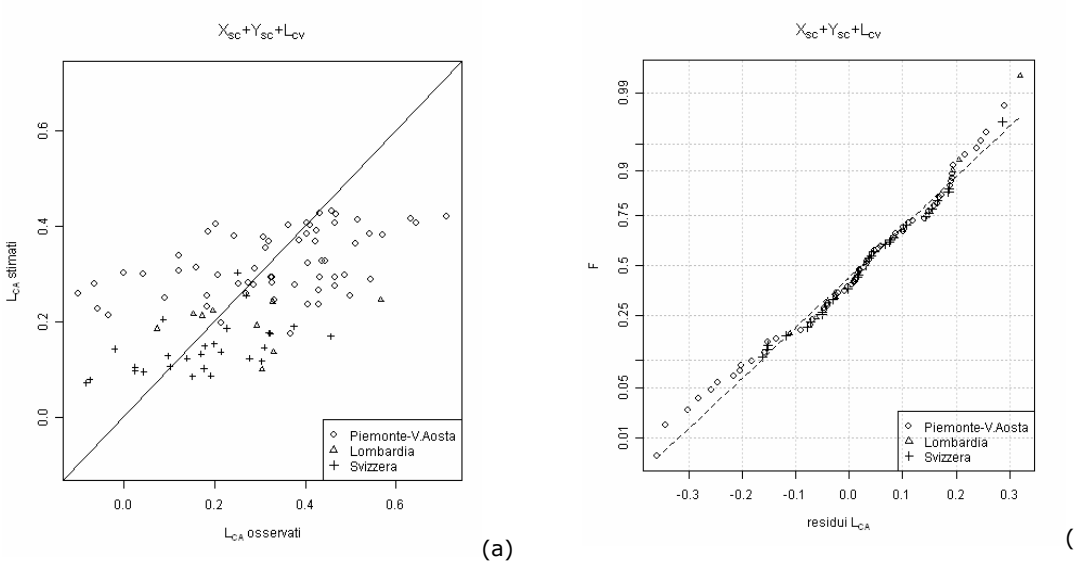


Figura 7.5. Diagrammi diagnostici per la migliore relazione a 3 parametri tarata con il metodo dei W.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia una buona distribuzione normale dei residui.

Dall'analisi dei grafici a dispersione emerge, in entrambi i casi, il discreto adattamento dei valori stimati a quelli osservati (Fig.7.4.a e 7.5.a). Tuttavia si evidenzia come, per entrambi i casi, ci sia una tendenza alla sottostima per valori di L_{ca} osservati maggiori di 0.3 e una tendenza alla sovrastima per valori di L_{ca} osservati inferiori a 0.3. Inoltre per la regressione tarata con i W.L.S. si riscontra come la distribuzione dei residui sia leggermente migliore rispetto a quella relativa alla regressione tarata con gli O.L.S. (Fig.7.4.b e 7.5.b). In questo caso i due metodi di taratura dei parametri (OLS e WLS) sembrano fornire risultati comparabili in termini di qualità dei diagrammi di dispersione prodotti. Come in precedenza, conviene anche costruire i diagrammi di dispersione utilizzando le variabili moltiplicate per la radice quadrata dei pesi ($\sqrt{w_i}L_{CA,i}$). Solo utilizzando le variabili opportune si ottengono degli scatter plot (Figura 7.6) congruenti con l'utilizzo dei pesi.

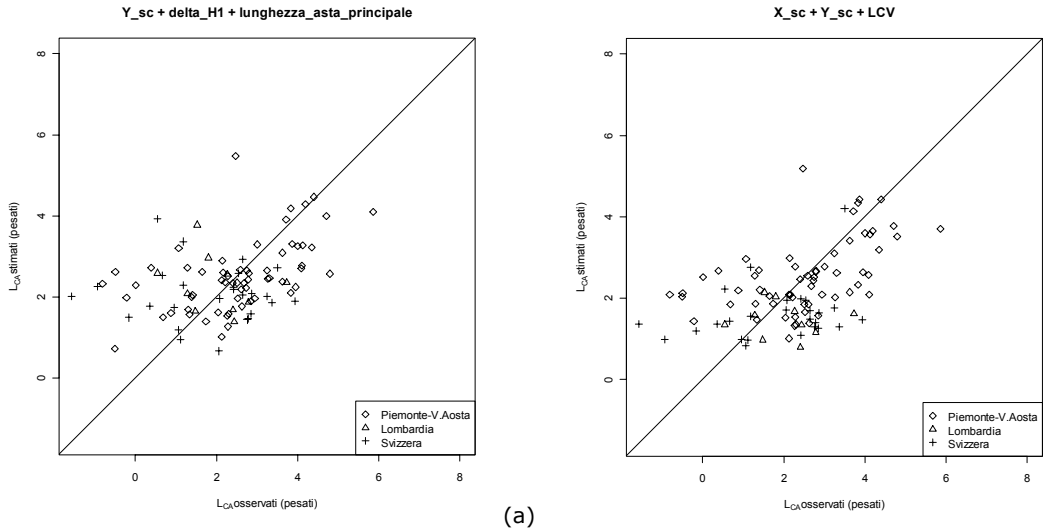


Figura 7.6. Scatter plot per LCA stimato tramite la migliore relazione a 3 parametri tarata con il metodo degli OLS (a) e dei WLS (b). Le variabili in ascissa ed in ordinata corrispondono a $\sqrt{w_j}L_{CA,j}$, con $w_j = n_j \Big/ \left(0.45 + 0.6 \cdot \left|L_{ca,j}\right|\right)^2$.

Nel caso di Figura 7.6 non si riconoscono grandi miglioramenti derivanti dall'uso del metodo dei minimi quadrati pesati. Le analisi grafiche delle Figura 7.4-7.6 evidenziano quindi una sostanziale parità nelle performance delle tecniche OLS e WLS, quando esse sono impiegate per la stima di LCA. Si è comunque scelto di utilizzare le stime di LCA ottenute tramite il metodo WLS per congruenza con quanto fatto con LCV ed in base ai migliori risultati in termini di varianza di stima, Tabella 7.3 e 7.4.

7.3 Stima basata sugli estremi giornalieri

Per questo metodo la base di dati è composta da un gruppo di 78 stazioni ricadenti nell'Italia Nord Occidentale e nel Sud della Svizzera aventi almeno 10 dati contemporanei di portata al colmo e di portata estrema giornaliera, in modo da tarare opportune formule multiregressive per la stima di L_{cv} e di L_{ca}. Tali relazioni impiegheranno come possibili variabili esplicative i 33 citati parametri geomorfologici e gli *L*-coefficienti relativi alle serie delle portate estreme giornaliere.

Il modello di regressione multipla di riferimento è espresso dalla relazione (6.10), nella quale la variabile dipendente ***y*** viene fissata pari a:

- $\boldsymbol{L_{cv}}$ per la stima del coefficiente di *L*-variazione;
- $\boldsymbol{L_{ca}}$ per la stima nel campo lineare del coefficiente di *L*-asimmetria.

I criteri utilizzati per la selezione dei modelli sono gli stessi discussi nel paragrafo 7.2.

7.3.1 Stima del coefficiente di L-variazione

Come accennato precedentemente, la stima del coefficiente di *L*-variazione è stata effettuata nel campo lineare, fissando $\boldsymbol{y} = \boldsymbol{L_{cv}}$. La determinazione dei migliori modelli regressivi è stata effettuata esaminando il valore di R^2_{adj} associato a ciascuna regressione, utilizzando sia il metodo dei minimi quadrati ordinari (O.L.S.) che quello dei minimi quadrati pesati (W.L.S.), ottenuti fissando un peso w_j pari a $n_j / (0.9 \cdot L_{cv})^2$.

Variabile dipendente y = L_{cv} (O.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
Y_{bar}, lun_asta_princ, lungh_vett_orient, $L_{cv,g}$	0.97	0.00179
Y_{sc}, lun_asta_princ, lungh_vett_orient, $L_{cv,g}$	0.97	0.00180
Y_{bar}, LLDP, lungh_vett_orient, $L_{cv,g}$	0.97	0.00181
Y_{sc}, LLDP, lungh_vett_orient, $L_{cv,g}$	0.97	0.00182
Y_{bar}, ΔH_1, lun_asta_princ, $L_{cv,g}$	0.97	0.00207
Y_{sc}, densita_dren, $L_{cv,g}$	0.96	0.00218
Y_{bar}, densita_dren, $L_{cv,g}$	0.96	0.00218
Y_{sc}, sl_med1, $L_{cv,g}$	0.96	0.00214
Y_{bar}, sl_med1, $L_{cv,g}$	0.96	0.00214
lun_asta_princ, lungh_vett_orient, $L_{cv,g}$	0.96	0.00170
Y_{sc}, $L_{cv,g}$	0.96	0.00188
Y_{bar}, $L_{cv,g}$	0.96	0.00186
$L_{cv,g}$	0.96	0.00175
Y_{sc}	0.91	0.00684
Y_{bar}	0.90	0.00706
densita_dren	0.90	0.00818
H_{media}	0.90	0.00763

Tabella 7.5. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati ordinari fissando la variabile dipendente *y* pari a L_{cv}.

Per confrontare i risultati ottenuti con i due approcci si ricorre all'analisi delle varianze di stima, calcolate secondo i criteri esposti al paragrafo 6.3.3. Le stime migliori saranno quelle con incertezza minore, ossia con errore di modello più basso (per confrontare le stime ottenute con modelli diversi conviene infatti trascurare il secondo termine in equazione (6.22)). I modelli multiregressivi ottenuti per L_{cv} nelle due configurazioni sono riportati in Tabella 7.5 e 7.6.

Variabile dipendente y = L_{cv} (W.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
A, lun_asta_princ, R_al, $L_{cv,g}$	0.96	0.00119
A, lun_asta_princ, F_f, $L_{cv,g}$	0.96	0.00121
A, LLDP, R_al, $L_{cv,g}$	0.96	0.00121
A, LLDP, F_f, $L_{cv,g}$	0.96	0.00123
lun_asta_princ, lungh_vett_orient, $L_{cv,g}$	0.96	0.00127
LLDP, lungh_vett_orient, $L_{cv,g}$	0.96	0.00130
lungh_vett_orient, media_fa, $L_{cv,g}$	0.95	0.00135
R_al, F_f, $L_{cv,g}$	0.95	0.00155
A, magnitudine, $L_{cv,g}$	0.95	0.00154
$L_{cv,g}$	0.95	0.00156
H_{media}	0.88	0.00739
Y_{sc}	0.88	0.00667
Y_{bar}	0.88	0.00699
lungh_media_vers	0.87	0.00819

Tabella 7.6. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati pesati fissando la variabile dipendente *y* pari a L_{cv}.

Dall'analisi dei valori di varianza associati ai diversi modelli emerge chiaramente come il metodo dei minimi quadrati pesati (W.L.S.) conduca a stime affette da incertezza minore rispetto a quelle determinate con il metodo dei minimi quadrati ordinari. Le differenze nei valori di $\sigma^2_{\hat{y}}$ sono significative; in particolare, tra le migliori stime a 4 parametri associate ai due metodi, si riscontra una differenza in termini di varianza del 34%.

Ancora una volta, invece, il valore del coefficiente di determinazione risulta confrontabile nei due casi. Si evidenzia, inoltre, come tra le migliori regressioni a quattro parametri individuate con i minimi quadrati ordinari ci

sia sempre una misura di latitudine, una lunghezza ed $L_{cv,g}$ tra le variabili esplicative; al contrario le analoghe regressioni individuate con i W.L.S. si basano sempre su parametri di lunghezza e di forma, oltre, naturalmente, al coefficiente di *L*-variazione delle portate estreme giornaliere.

In figura 7.7 e 7.8 vengono riportati i diagrammi diagnostici relativi alla migliore regressione a 4 parametri tarata con il metodo degli O.L.S. e dei W.L.S.
Dall'analisi dei grafici emerge, in entrambi i casi, il buon adattamento dei valori stimati a quelli osservati (Fig.7.7.a e 7.8a). Anche per quanto riguarda la verifica della distribuzione dei residui non si riscontrano problemi particolari (Fig.7.7.b e 7.8.b).

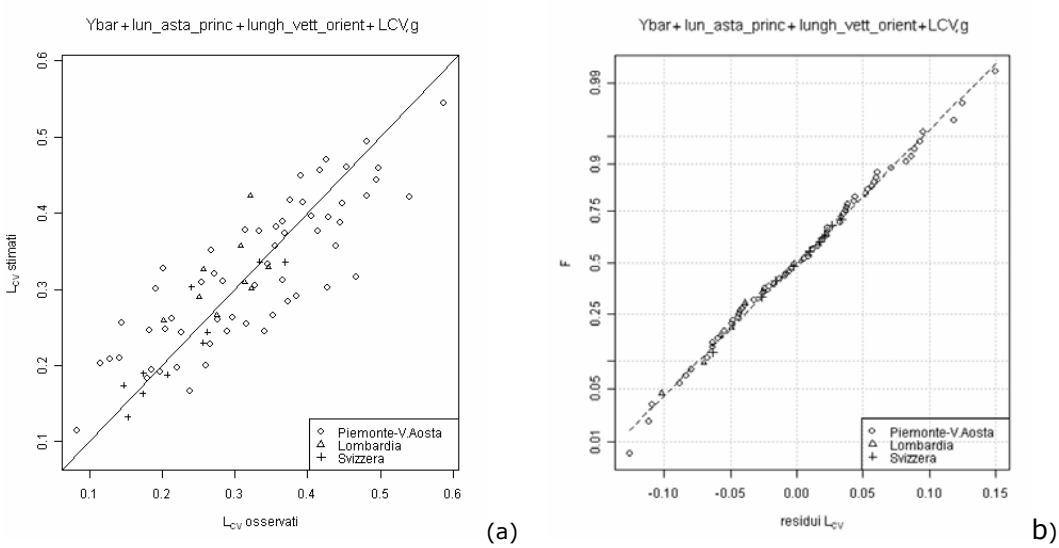


Figura 7.7. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata con il metodo degli O.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

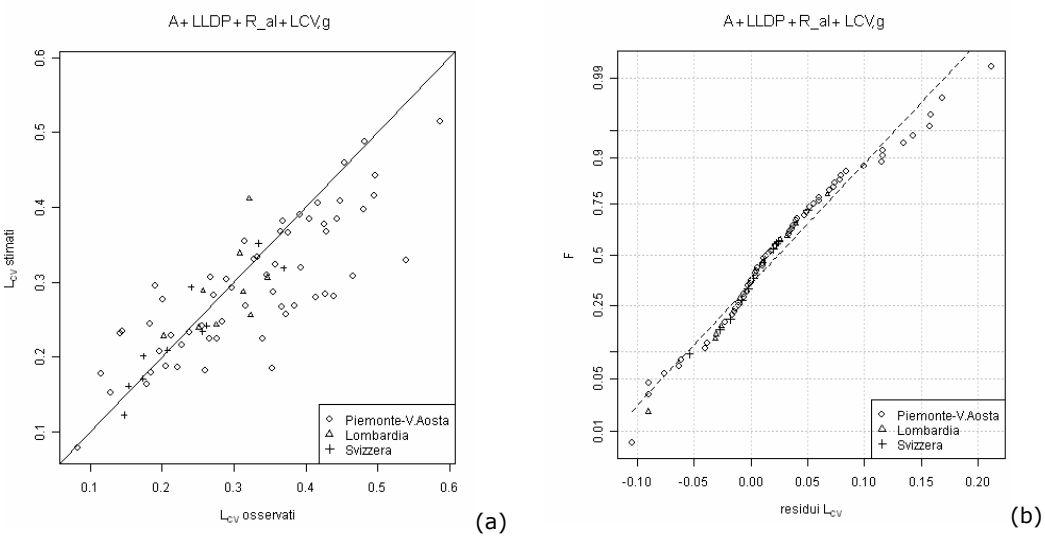


Figura 7.8. Diagrammi diagnostici per terza la migliore relazione a 4 parametri tarata con il metodo dei W.L.S. In (a) si riscontra il discreto adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

7.3.2 Stima del coefficiente di L-asimmetria

La stima del coefficiente di *L*-asimmetria è stata effettuata nel campo lineare, fissando $\boldsymbol{y} = \boldsymbol{L_{ca}}$. La determinazione dei migliori modelli regressivi è stata effettuata esaminando il valore di R^2_{adj} associato a ciascuna regressione. In prima istanza i modelli multiregressivi sono stati determinati applicando il metodo dei minimi quadrati ordinari (O.L.S.), basato sull'ipotesi di omoschedasticità dei residui. I risultati ottenuti con tale metodo sono stati confrontati con quelli derivanti dall'applicazione del metodo dei minimi quadrati pesati (W.L.S.), ottenuti fissando un peso w_j pari a $n_j / \left(0.45 + 0.6 \cdot \left|L_{ca}\right|\right)^2$.

Per confrontare i risultati ottenuti con i due approcci si ricorre all'analisi delle varianze di stima, calcolate secondo i criteri esposti al paragrafo 6.3.3. Le stime migliori saranno quelle con incertezza minore, ossia con errore di modello più basso (per confrontare le stime ottenute con modelli diversi conviene infatti trascurare il secondo termine in equazione (6.22)). Le stime dei modelli multiregressivi ottenuti per L_{cv} nelle due configurazioni sono riportati in Tabella 7.7 e 7.8.

Variabile dipendente y = L_{ca} (O.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
Y_{sc}, orientamento, sl_med1, $L_{ca,g}$	0.87	0.00410
Y_{sc}, H_{media}, orientamento, $L_{ca,g}$	0.87	0.00352
Y_{bar}, orientamento, sl_med1, $L_{ca,g}$	0.87	0.00415
Y_{sc}, orientamento, $L_{ca,g}$	0.87	0.00378
Y_{bar}, orientamento, $L_{ca,g}$	0.87	0.00390
Y_{sc}, $L_{ca,g}$	0.86	0.00451
Y_{bar}, $L_{ca,g}$	0.86	0.00468
$L_{ca,g}$	0.85	0.00623
Y_{sc}	0.79	0.01127
Y_{bar}	0.79	0.01182
aff	0.77	0.01231
X_{sc}	0.76	0.01438

Tabella 7.7. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati ordinari fissando la variabile dipendente *y* pari a L_{ca}.

Variabile dipendente y = L_{ca} (W.L.S.)	R^2_{adj}	$\sigma^2_{\hat{y}}$
Y_{sc}, X_{bar}, orientamento, $L_{ca,g}$	0.87	0.00229
X_{sc}, Y_{sc}, orientamento, $L_{ca,g}$	0.87	0.00242
R_al, F_f, aff, $L_{ca,g}$	0.87	0.00250
X_{bar}, Y_{bar}, orientamento, $L_{ca,g}$	0.87	0.00233
X_{sc}, Y_{bar}, orientamento, $L_{ca,g}$	0.87	0.00247
F_f, aff, $L_{ca,g}$	0.86	0.00229
X_{sc}, Y_{sc}, $L_{ca,g}$	0.86	0.00288
X_{sc}, F_f, Lca,g	0.86	0.00259
Y_{sc}, orientamento, $L_{ca,g}$	0.86	0.00276
aff, $L_{ca,g}$	0.86	0.00283
Y_{sc}, $L_{ca,g}$	0.85	0.00317
X_{bar}, $L_{ca,g}$	0.85	0.00333
X_{sc}, $L_{ca,g}$	0.85	0.00345
Y_{bar}, $L_{ca,g}$	0.85	0.00328
$L_{ca,g}$	0.84	0.00425
Y_{sc}	0.76	0.01050
aff	0.76	0.01078
Y_{bar}	0.76	0.01103
X_{bar}	0.74	0.01266

Tabella 7.8. Valori di R^2_{adj} e della varianza di stima calcolati per i modelli multiregressivi che superano il test della *t* di Student e soddisfano l'ipotesi di non multicollinearità tarati tramite il metodo dei minimi quadrati pesati fissando la variabile dipendente *y* pari a L_{ca}.

Esaminando valori di varianza associati alle diverse relazioni multiregressive emerge come il metodo dei minimi quadrati pesati (W.L.S.) conduca a stime affette da incertezza minore rispetto a quelle relative al metodo dei minimi quadrati ordinari. Le differenze nei valori di $\sigma^2_{\hat{y}}$ sono significative; infatti, tra le migliori stime a 4 parametri associate ai due metodi, si riscontra una differenza in termini di varianza del 44%.

E' da notare, invece, come il valore del coefficiente di determinazione risulti pressoché identico nei due casi. Si evidenzia, inoltre, come tra i migliori regressori nelle regressioni a quattro parametri individuati con il metodo O.L.S. ci sia sempre una misura di latitudine e l'orientamento del bacino ed $L_{ca,g}$; al contrario le regressioni a quattro parametri individuate con i W.L.S. presentano sempre, tra le variabili esplicative, un parametro di forma del bacino e il coefficiente di *L*-asimmetria delle portate estreme giornaliere.

In figura 7.9 e 7.10 vengono riportati i diagrammi diagnostici relativi alla migliore regressione a 4 parametri tarata con il metodo degli O.L.S. e dei W.L.S.

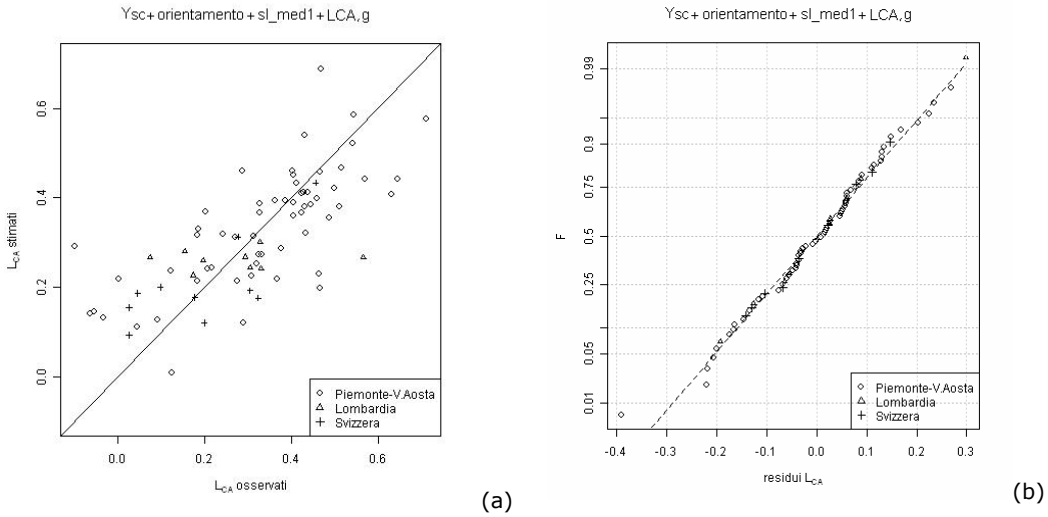


Figura 7.9. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata con il metodo degli O.L.S. In (a) si riscontra il discreto adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

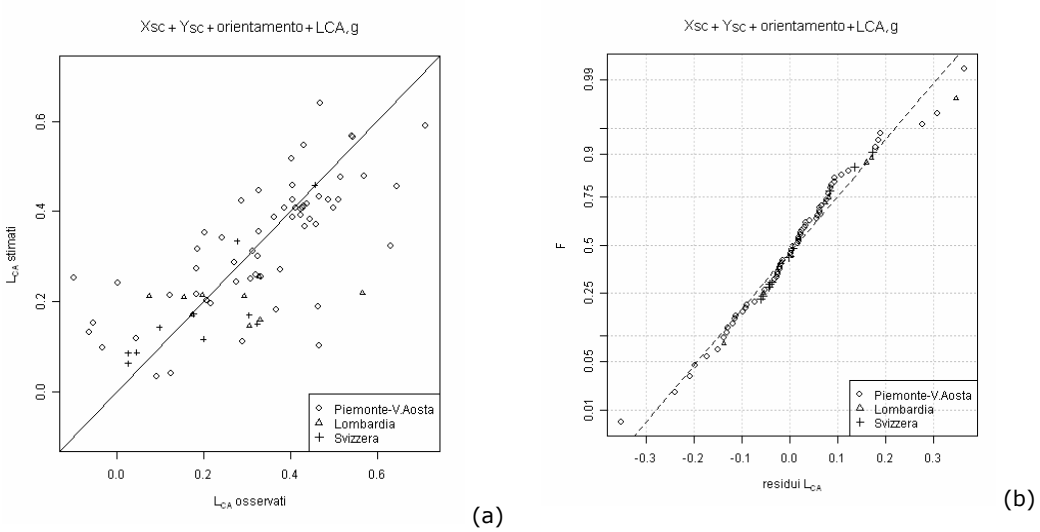


Figura 7.10. Diagrammi diagnostici per la migliore relazione a 4 parametri tarata con il metodo dei W.L.S. In (a) si riscontra il buon adattamento ai dati osservati, mentre in (b) si evidenzia la distribuzione normale dei residui.

Dall'analisi degli scatter plot emerge, in entrambi i casi, il buon adattamento dei valori stimati a quelli osservati (Fig.7.9.a e 7.10.a). Anche per quanto riguarda la verifica della distribuzione dei residui non si riscontrano problemi particolari (Fig.7.9.b e 7.10.b).

7.4 Scelta del modello per la stima degli *L*-coefficienti

In questo paragrafo vengono scelte le configurazioni più efficienti associate a ciascun metodo di stima. Lo scopo è individuare, per ogni approccio, una relazione unica per la stima di L_{cv} e di L_{ca}. A tal fine si analizzano le relazioni multiregressive tarate nel paragrafo precedente e, in base a considerazioni sull'incertezza di stima e sulle variabili esplicative coinvolte, si scelgono le regressioni più adeguate.

7.4.1 Stima regionale (regressione multipla)

7.4.1.1 Coefficiente di L-variazione

Alla luce dei valori di varianza di stima ottenuti per i modelli multiregressivi tarati con il metodo dei minimi quadrati pesati fissando la variabile dipendente y pari a L_{cv}, riportati in tabella 7.2, la relazione più appropriata per la stima del coefficiente di *L*-variazione risulta:

$$L_{cv} = 2.648 \cdot 10^{-1} - 8.392 \cdot 10^{-5} H_{media} - 4.314 \cdot 10^{-3} LLDP + \\ + 1.304 \cdot 10^{-2} lungh\_vett\_orient + 2.975 \cdot 10^{-1} n \quad , \qquad (7.1)$$

corrispondente alla migliore relazione a 4 parametri ottenuta.
Si sottolinea come quest'ultima fornisca la stessa incertezza della seconda migliore relazione a 4 parametri, la quale presenta le medesime variabili esplicative ad eccezione di LLDP, che viene sostituita dalla lunghezza dell'asta principale. Tuttavia è stata scelta la relazione (7.1) poiché il valore di LLDP è determinabile in maniera più oggettiva rispetto alla lunghezza dell'asta principale, che, invece, è dipendente dalla soglia utilizzata per la selezione delle aree non canalizzate del bacino.
Si evidenzia inoltre che i parametri *LLDP* e *lugh_vett_orient* non sono multicollineari, in quanto il primo è una lunghezza legata al reticolo, mentre il secondo rappresenta una lunghezza associata alla posizione del baricentro e della sezione di chiusura.

7.4.1.2 Coefficiente di L-asimmetria

Analizzando i valori di varianza per le diverse relazioni multiregressive tarate tramite il metodo ai minimi quadrati pesati, ottenute fissando $y = L_{ca}$, si ritiene che il modello regressivo più adatto per la stima del coefficiente di L-asimmetria sia

$$L_{ca} = 3.604 - 1.144 \cdot 10^{-6} X_{sc} - 6.052 \cdot 10^{-7} Y_{sc} + 7.190 \cdot 10^{-1} L_{cv} \, , \qquad (7.2)$$

corrispondente alla migliore relazione a 4 parametri.
Il valore di incertezza associato alla relazione (7.2) è molto simile a quello relativo alla seconda migliore relazione a 4 parametri, che differisce da essa solo per una variabile esplicativa, presentando la latitudine Y_{bar} del baricentro della bacino al posto della latitudine Y_{sc} della sezione di chiusura.

7.4.2 Stima basata sugli estremi giornalieri

7.4.2.1 Coefficiente di L-variazione

Alla luce dei valori di varianza ottenuti per i modelli multiregressivi tarati con il metodo dei minimi quadrati pesati fissando la variabile dipendente *y*

pari a L_{cv}, riportati nel paragrafo 7.3.1, la relazione più appropriata per la stima del coefficiente di *L*-variazione risulta:

$$L_{cv} = 2.485 \cdot 10^{-1} - 9.813 \cdot 10^{-5} A - 2.697 \cdot 10^{-3} LLDP + \\ - 2.402 \cdot 10^{-1} R\_al + 8.734 \cdot 10^{-1} L_{cv,g} \quad , \qquad (7.3)$$

corrispondente alla terza migliore relazione a 4 parametri ottenuta.
Si noti come il valore di incertezza ad essa relativa è circa uguale a quello ottenuto per le due migliori relazioni a 4 parametri, che, tuttavia, non sono state scelte perché presentano tra le variabili esplicative la lunghezza dell'asta principale al posto di LLDP.
Le relazioni espresse in funzione di 3 variabili esplicative non sono invece state considerate perché presentano valori di incertezza significativamente più elevati rispetto a quelli associati alla relazione (7.3).

7.4.2.2 Coefficiente di L-asimmetria

Analizzando i valori di varianza per le diverse relazioni multiregressive tarate tramite il metodo ai minimi quadrati pesati, ottenute fissando $y = L_{ca}$, si ritiene che il modello regressivo più adatto per la stima del coefficiente di *L*-asimmetria sia

$$L_{ca} = 3.043 \cdot 10^{-1} - 2.258 \cdot 10^{-1} F\_f - 8.823 \cdot 10^{-5} aff + 7.312 \cdot 10^{-1} L_{ca,g} \qquad (7.4)$$

corrispondente alla migliore relazione a 3 parametri.
Il valore di incertezza associato alla relazione (7.4) è analogo a quello relativo alla relazione migliore a 4 parametri. Si riscontra, inoltre, che le altre regressioni espresse in funzione di 4 parametri presentano valori di varianza più elevati rispetto a quello relativo alla (7.4).

7.5 Condizioni di applicabilità di ciascun approccio

Una volta stabilita, per ciascuna metodologia di stima analizzata, la relazione più appropriata per la valutazione degli *L*-coefficienti, è necessario determinare quali siano le condizioni nelle quali è opportuno utilizzare ciascuna metodologia. In particolare questo dipende dalle seguenti condizioni:

a) **alternativa stima empirica/regionale**: determinazione della numerosità n_{min} di portate al colmo al di sotto della quale è preferibile ricorrere a un metodo di stima regionale (ad es. il modello di regressione multipla), piuttosto che ad una stima *at site*, effettuata a partire dal campione di osservazioni disponibili;
b) **alternativa stima regionale/estremi giornalieri**: determinazione della numerosità $n_{g,min}$ di estremi giornalieri annui al di sopra della quale la disponibilità di dati di estremi giornalieri consente di migliorare il metodo di stima regionale;
c) **alternativa stima empirica/estremi giornalieri**: individuazione della relazione tra n ed n_g tra la numerosità della serie delle portate al colmo e la numerosità di estremi giornalieri che definisce la soglia di selezione. Al disopra di questa soglia la stima empirica basata sulle portate al colmo è preferibile alla stima basata sull'impiego delle portate giornaliere.

Con riferimento al caso a), è necessario eguagliare le incertezze di stima associata ai due criteri di stima. È utile ricordare che ciascun modello di regressione multipla presenta una varianza costante, indipendente dalla sezione in cui viene applicato; al contrario la varianza relativa alla stima empirica degli L-coefficienti è variabile da stazione a stazione, secondo la relazione (6.8) per L_{cv} e secondo la relazione (6.9) per L_{ca}. La numerosità equivalente risulta, dunque, pari a :

$$n_{\min,L_{cv}} = \frac{\left(0.9 \cdot L_{cv}\right)^2}{\sigma^2_{L_{cv},regr}}, \tag{7.5}$$

$$n_{\min,L_{ca}} = \frac{\left(0.45 + 0.6 \cdot \left|L_{ca}\right|\right)^2}{\sigma^2_{L_{ca},regr}}, \tag{7.6}$$

in cui $\sigma^2_{L_{cv},regr} = 0.00365$ (vedi tabella 7.2) e $\sigma^2_{L_{ca},regr} = 0.00638$ (vedi tabella 7.4).

Se si considera il caso b), mentre la varianza relativa a un modello di regressione multipla è costante, l'incertezza associata alla stima da estremi

giornalieri è costituita da due componenti (equazione (6.19)). Ne consegue che la numerosità equivalente risulta pari a:

$$n_{g,\min,L_{cv}} = \frac{\beta^2_{1,L_{cv,g}} \left(0.9 \cdot L_{cv,g}\right)^2}{\sigma^2_{L_{cv},regr} - \sigma^2_{L_{cv},g\_\mathrm{mod}el}}, \qquad (7.7)$$

per il coefficiente di L-variazione, e pari a

$$n_{g,\min,L_{ca}} = \frac{\beta^2_{1,L_{ca},g} \left(0.45 + 0.6 \cdot \left|L_{ca,g}\right|\right)^2}{\sigma^2_{L_{ca},regr} - \sigma^2_{L_{ca},g\_\mathrm{mod}el}}, \qquad (7.8)$$

per il coefficiente di L-asimmetria. Nelle (7.7) e (7.8) $\sigma^2_{L_{cv},g\_\mathrm{mod}el} = 0.00121$ e $\sigma^2_{L_{ca},g\_\mathrm{mod}el} = 0.00229$ rappresentano, rispettivamente, le incertezze relative ai modelli regressivi (7.3) e (7.4); $\beta_{1,L_{cv,g}} = 8.734 \cdot 10^{-1}$ e $\beta_{1,L_{ca,g}} = 7.312 \cdot 10^{-1}$ rappresentano i coefficienti moltiplicativi rispettivamente di L_{cv} e di L_{ca} nelle relazioni (7.3) e (7.4).

Per il caso c) la condizione di scelta risulta dall'imporre l'eguaglianza dell'incertezza relativa dei due metodi. Entrambe le varianze sono variabili da sito a sito; inoltre le relazioni utili al calcolo dell'incertezza sono espresse in funzione del numero di dati rispettivamente disponibili, per le portate al colmo e le portate estreme giornaliere. Ne consegue che eguagliando le due misure di incertezza si ottengono delle relazioni non esplicitabili; per L_{cv} infatti vale che

$$\frac{\left(0.9 \cdot L_{cv}\right)^2}{n} = \frac{\beta^2_{1,L_{cv,g}} \left(0.9 \cdot L_{cv,g}\right)^2}{n_g} + \sigma^2_{L_{cv},g\_\mathrm{mod}el}, \qquad (7.9)$$

mentre per L_{ca}

$$\frac{\left(0.45 + 0.6 \cdot \left|L_{ca}\right|\right)^2}{n} = \frac{\beta^2_{1,L_{ca},g} \left(0.45 + 0.6 \cdot \left|L_{ca,g}\right|\right)^2}{n_g} + \sigma^2_{L_{ca},g\_\mathrm{mod}el}. \qquad (7.10)$$

Per individuare una possibile numerosità equivalente valida per tutte le sezioni esaminate questa è stata prima calcolata per tutte le stazioni considerate. I valori ottenuti sono stati poi mediati ottenendo i risultati riportati in Tabella 7.9.

	Numerosità equivalente	***Descrizione***
L_{cv}	$n_{min} > 21.2$	Conviene utilizzare la stima empirica sulla serie di portate al colmo piuttosto che la stima regionale (formula (7.1)).
	$n_{g,min} > 18.1$	Conviene utilizzare la stima basata sulle portate estreme giornaliere tramite la relazione (7.3) piuttosto che la stima regionale (formula (7.1)).
L_{ca}	$n_{min} > 63.4$	Conviene utilizzare la stima empirica sulla serie di portate al colmo piuttosto che la formula multiregressiva (7.2).
	$n_{g,min} > 46.9$	Conviene utilizzare la stima basata sulle portate estreme giornaliere tramite la relazione (7.4) piuttosto che la formula multiregressiva (7.2).

Tabella 7.9. Valori della numerosità equivalente utile a definire le condizioni di applicabilità di ciascun metodo in ragione del numero di osservazioni disponibili.

Analizzando i risultati riportati in tabella 7.9 emerge chiaramente come la stima del coefficiente di *L*-variazione possa essere effettuata direttamente dai dati anche per campioni di portate al colmo di modeste dimensioni. Al contrario, per stimare direttamente il coefficiente di *L*-asimmetria sono necessari campioni decisamente più consistenti. Qualora essi non fossero disponibili si deve ricorrere a un modello di stima regionale.

Per quanto riguarda la stima basata sugli estremi giornalieri le numerosità equivalenti medie assumono valori superiori rispetto al caso precedente. Sono infatti sufficienti mediamente 27 dati di portata estrema giornaliera per effettuare la stima del coefficiente di *L*-variazione, mentre per il calcolo di L_{ca} ne sono mediamente necessari almeno 44.

8 Scelta della distribuzione di probabilità

8.1 Distribuzioni considerate

Nell'analisi di frequenza regionale l'obiettivo è determinare una curva di crescita $K_j(T)$, che sia rappresentativa dell'andamento degli eventi estremi in una qualsiasi sezione.

Per individuare la distribuzione che permetta la migliore stima dei quantili in una qualsiasi sezione ricadente nella regione di interesse per il presente lavoro (Italia Nord Occidentale e Sud della Svizzera), è necessario testare diverse possibili distribuzioni e verificarne la bontà di adattamento nelle sezioni strumentate appartenenti a tale territorio.

Le distribuzioni candidate alla rappresentazione della curva di crescita vengono scelte tenendo conto che la variabile idrologica di interesse è rappresentata dalle portate di piena, e, di conseguenza, si dovrà considerare con particolare attenzione l'andamento della coda superiore di ciascuna distribuzione.

Le distribuzioni considerate nel presente lavoro sono (tra parentesi si riporta il simbolo utilizzato per identificare la distribuzione nei grafici seguenti):

- la distribuzione di Gumbel (G);
- la distribuzione Pareto generalizzata (GP);
- la distribuzione Generalizzata del Valore Estremo (GEV);
- la distribuzione Logistica generalizzata (GL);
- la distribuzione Lognormale (LN);
- la distribuzione Gamma (GAM);
- la distribuzione Kappa (K);
- la distribuzione del valore estremo a doppia componente (TCEV).

Si tratta di distribuzioni a 3 parametri, ad eccezione della distribuzione di Gumbel (2 parametri) e della Kappa e della TCEV, che vengono espresse in funzione di 4 parametri.

La stima dei parametri di ogni distribuzione, in una *j*-esima sezione, viene effettuata considerando la media pari ad 1 (infatti si sta considerando la curva di crescita) ed utilizzando gli *L*-coefficienti ivi stimati. Le relazioni utilizzate per determinare i parametri delle distribuzioni noti i valori degli L-coefficienti sono riportate in Appendice D.

Una prima verifica grafica dell'efficienza delle differenti distribuzioni candidate deriva dalla rappresentazione della curva di crescita e dei dati disponibili in funzione del tempo di ritorno. Due esempi, relativi al Chisone a San Martino ed al Sesia a Rimasco sono riportati in Figura 8.1. I punti empirici sono riportati su tali diagrammi seguendo il criterio di Hirsch, adottando la plotting position di Hazen. I parametri di ciascuna distribuzione vengono stimati da L_{cv} ed L_{ca} calcolati tramite i modelli di regressione multipla (7.1) e (7.2).

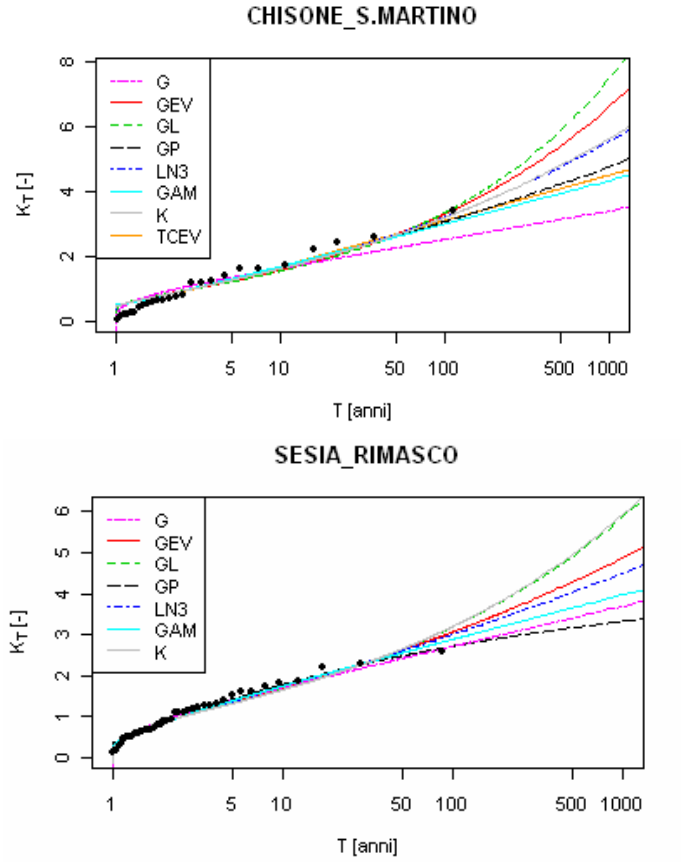


Figura 8.1. Esempi di curve di crescita empiriche (punti) e teoriche, ottenute con le 8 distribuzioni di probabilità considerate nel presente lavoro.

I diagrammi del tipo di quelli rappresentati in figura 8.1 forniscono una prima indicazione della qualità dell'approssimazione fornita dai modelli probabilistici qui considerati. Si noti che nel caso del Chisone a S. Martino sono effettivamente presenti 8 curve diverse in Figura 8.1, mentre per il Sesia a Rimasco non è presente la curva corrispondente alla distribuzione TCEV, in quanto la seconda componente della TCEV risulta irrilevante per tale bacino, cosicché la TCEV degenera in una Gumbel. In Allegato 1 si riportano le curve

di crescita ottenute per tutti i 97 bacini nei quali sono disponibili almeno 10 dati di portata al colmo di piena. Dalla figura 8.1, come anche dalle curve riportate in Allegato 1, risulta evidente come le diverse distribuzioni tendono ad avere andamenti molto simili tra loro per tempi di ritorno medio-bassi, per poi divergere nella parte destra del diagramma.
Una valutazione più sistematica ed oggettiva della qualità dell'adattamento delle diverse distribuzioni ai dati si ha applicando opportuni test di bontà di adattamento, come descritto nel paragrafo seguente.

8.2 Verifica della bontà di adattamento

8.2.1 Test di adattamento

Per ognuna delle 8 distribuzioni considerate, sono stati effettuati 4 test di adattamento, descritti nei successivi sottoparagrafi:

- Test di Cramer – von Mises;
- Test di Anderson – Darling;
- Test di uniformità;
- Test del valor massimo.

I test saranno applicati nel presente lavoro con modalità non completamente standard: infatti lo scopo dei test di adattamento è verificare se una prefissata distribuzione di probabilità possa essere considerata come la distribuzione generatrice dei dati disponibili, la cosiddetta parent distribution. Nel presente lavoro, invece, si utilizzeranno le variabili test non per scartare o accettare una data distribuzione di probabilità, ma per confrontare tra loro le 8 distribuzioni considerate. Lo scopo ultimo di questa analisi sarà dunque la selezione del modello probabilistico più appropriato a rappresentare i dati.

8.2.1.1 Test di Cramer – von Mises

Il test di Cramer – von Mises, basato sul confronto tra curva di crescita empirica e teorica, propone una misura dello scostamento quadratico medio W2 tra la distribuzione di probabilità ipotizzata e quella relativa ai dati osservati:

$$W^2 = \sum_{i=1}^{n}\left[F\left(Q_{(i)}\right) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}, \qquad (8.1)$$

in cui:

- *n* indica il numero di dati osservati disponibili per la stazione in esame;
- *i* rappresenta la posizione del valore di portata nel campione ordinato in senso crescente;
- $F\left(Q_{(i)}\right)$ è la frequenza della distribuzione di probabilità relativa al metodo considerato, determinata per il valore di portata corrispondente all'*i*-esimo dato osservato,

Nei casi in cui sono presenti dati di portata non appartenenti al campione sistematico la statistica relativa al test di Cramer – von Mises si calcola sostituendo il termine (2*i*-1)/2*n* con la frequenza empirica $P_{(i)}$ relativa all'*i*-esimo dato osservato, ottenendo la seguente relazione:

$$D^2 = \sum_{i=1}^{n}\left[F\left(Q_{(i)}\right) - P_{(i)}\right]^2 + \frac{1}{12n}, \qquad (8.2)$$

la quale rappresenta, analogamente a W^2, una misura dello scostamento medio quadratico tra la distribuzione di probabilità ipotetica e la funzione di frequenza del campione di osservazioni a disposizione. Le diverse distribuzioni saranno confrontate in base al valore assunto dalla statistica D^2, andando ovviamente a preferire distribuzioni caratterizzate da valori di D^2 inferiori. Dal momento che nell'analisi regionale sono disponibili complessivamente *s* campioni di dati, il test potrà essere ripetuto per *s* volte, ottenendo, per ogni distribuzione, $D_j^2, j = 1,..,s$; dato che è importante individuare un unico modello probabilistico che rappresenti al meglio la curva di crescita in tutte le stazioni, si possono infine rappresentare le curve di frequenza empirica relative agli 8 campioni di D_j^2 in un unico diagramma, per verificare graficamente quale distribuzione tende a produrre valori di D_j^2 più bassi (si veda il successivo paragrafo 8.2.2).

8.2.1.2 Test di Anderson - Darling

Il test di Anderson-Darling permette di verificare l'adattamento di una serie di dati ad una determinata distribuzione di probabilità, valutando lo scostamento tra curva di frequenza empirica e quella teorica e andando ad attribuire un peso maggiore agli scostamenti misurati sulle due code della distribuzione. Il test viene eseguito attraverso il calcolo della statistica A2, secondo la relazione:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\left(P_{(i)}\log\left(F\left(Q_{(i)}\right)\right) + \left(1 - P_{(i)}\right)\log\left(1 - F\left(Q_{(i)}\right)\right)\right). \qquad (8.3)$$

In questo caso le diverse distribuzioni saranno confrontate in base al valore assunto dalla statistica A^2, andando a preferire distribuzioni caratterizzate da valori di A^2 inferiori. Anche in questo caso il test potrà essere ripetuto per *s* volte, ottenendo, per ogni distribuzione, $A_j^2, j = 1,..,s$; dato che è importante individuare un unico modello probabilistico che rappresenti al meglio la curva di crescita in tutte le stazioni, si possono infine rappresentare le curve di frequenza empirica relative agli 8 campioni di A_j^2 in un unico diagramma, per verificare graficamente quale distribuzione tende a produrre valori di A_j^2 più bassi (si veda il successivo paragrafo 8.2.2).

8.2.1.3 Test di uniformità

Il terzo test considerato è basato sulla trasformazione

$$F_i = F(Q_i): \qquad (8.4)$$

se la distribuzione *F* è la distribuzione generatrice del campione di dati, allora il campione trasformato F_i ha una distribuzione uniforme. Il test è stato eseguito, per ognuna delle 8 distribuzioni considerate, (i) applicando la trasformazione (8.4) a tutti gli *s* campioni disponibili; (ii) costruendo un campione complessivo di n_{tot} elementi F_i, costituito dall'unione degli *s* sottocampioni disponibili; (iii) verificando graficamente se il campione complessivo possa ritenersi uniforme. La verifica grafica verrà effettuata

ponendo in ascissa i valori ordinati di F_i ed in ordinata le corrispondenti frequenze cumulate empiriche (i/ n_{tot}), come riportato al successivo paragrafo 8.2.2, e verificando se i punti tendono a disporsi in corrispondenza della bisettrice.

8.2.1.4 Test del valore massimo

Il test del valore massimo viene impiegato per valutare le prestazioni delle distribuzioni di frequenza ottenute con i diversi modelli per alti tempi di ritorno.

Il test in esame, infatti, anziché confrontare l'andamento della curva di frequenza calcolata con tutti i dati osservati si concentra esclusivamente sul valore più alto del campione, in modo tale da valutare se il modello è in grado di stimare adeguatamente l'estremo osservato.

Se $Q_{(n)}$ è il valor massimo del campione a disposizione, la sua distribuzione di probabilità cumulata può essere determinata a partire dalla distribuzione di probabilità teorica *F(Q)* secondo la relazione

$$\xi = F\left(Q_{(n)}\right) = \left[F(Q)\right]^n . \qquad (8.5)$$

Nota, allora, la distribuzione di probabilità dei metodi regionali impiegati è possibile determinare, per ciascuna sezione in esame, la distribuzione cumulata del valor massimo, ottenendo 8 campioni di s valori di $\xi_j , j = 1,..,s$. Se la distribuzione *F* è la distribuzione generatrice di tutti gli *s* campioni di dati, allora il campione degli $\xi_j , j = 1,..,s$ ha una distribuzione uniforme. Come prima, l'ipotesi di uniformità può essere verificata graficamente ponendo in ascissa i valori ordinati di ξ_i ed in ordinata le corrispondenti frequenze cumulate empiriche (j/ s), come riportato al successivo paragrafo 8.2.2, e verificando se i punti tendono a disporsi in corrispondenza della bisettrice.

8.2.2 Risultati dei test

8.2.2.1 Modelli di regressione multipla

I parametri delle distribuzioni considerate vengono stimati a partire dagli *L*-momenti calcolati tramite i modelli di regressione multipla tarati nel paragrafo 7.2.1 e 7.2.2. In particolare la stima di L_{cv} e di L_{ca} viene effettuata, rispettivamente, mediante le relazioni (7.1) e (7.2).

RISULTATI DEL TEST DI CRAMER VON-MISES

Come spiegato in precedenza, si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica D^2. I risultati sono riportati in Figura 8.2.

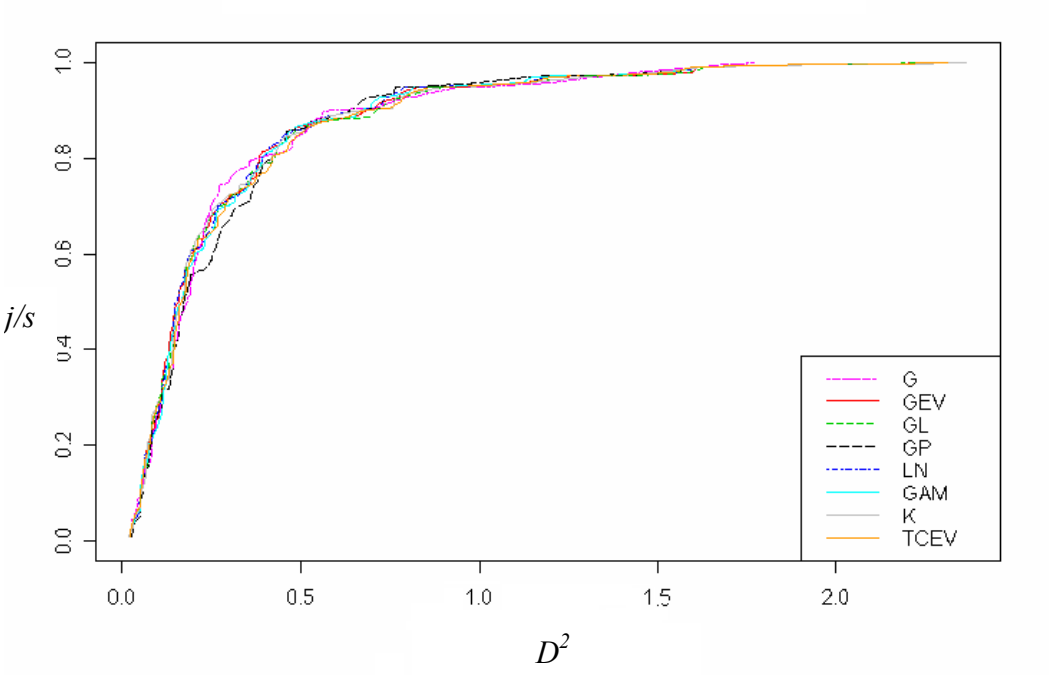


Figura 8.2. Funzione di frequenza empirica della statistica D^2 di Cramer von Mises calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Per verificare quale, tra le distribuzioni considerate, si comporta meglio secondo il test di Cramer – von Mises si verifica quale delle curve in Figura 8.2 risulti più spostata a sinistra, ossia su valori più bassi della statistica D^2. Nel caso specifico le 8 distribuzioni si comportano in maniera molto simile l'una

all'altra. Si può anche verificare qual è la distribuzione che comporta il valore di D^2 più basso per ogni stazione (Tabella 8.1).

Distribuzione	GEV	GL	GP	LN3
N_s	2	12	21	7
Distribuzione	G	GAM	K	TCEV
N_s	37	7	9	2

Tabella 8.1. Numero di stazioni in cui la distribuzione considerata restituisce il valore di D^2 più basso.

Tutte le distribuzioni considerate risultano ottime in almeno due stazioni, e le distribuzioni migliori secondo Tabella 8.1 risulterebbero la distribuzione di Gumbel e la distribuzione di Pareto generalizzata.
Un altro modo per individuare la distribuzione più adeguata a descrivere la curva di crescita delle portate di piena è quello di calcolare, per ciascuna distribuzione, il valore di D^2 in corrispondenza della mediana. La distribuzione migliore sarà caratterizzata dal D^2_{median} minore (Tabella 8.2).

Distribuzione	GEV	GL	GP	LN3
D^2_{median}	0.159	0.168	0.172	0.155
Distribuzione	G	GAM	K	TCEV
D^2_{median}	0.184	0.166	0.168	0.166

Tabella 8.2. Valori di D^2 valutati in corrispondenza della mediana.

Anche in questo caso si nota, come già in Figura 8.2, che il test di Cramer von Mises non permette di identificare grandi differenze tra le 8 distribuzioni considerate. Questo risultato è consistente con quanto rilevabile dai grafici di Figura 8.1 e riportati in Allegato I, che evidenziano come le 8 distribuzioni hanno comportamento molto simile nella parte centrale, corrispondente a tempi di ritorno medio-bassi.

RISULTATI DEL TEST DI ANDERSON - DARLING

Come spiegato in precedenza, si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica A^2. I risultati sono riportati in Figura 8.4.

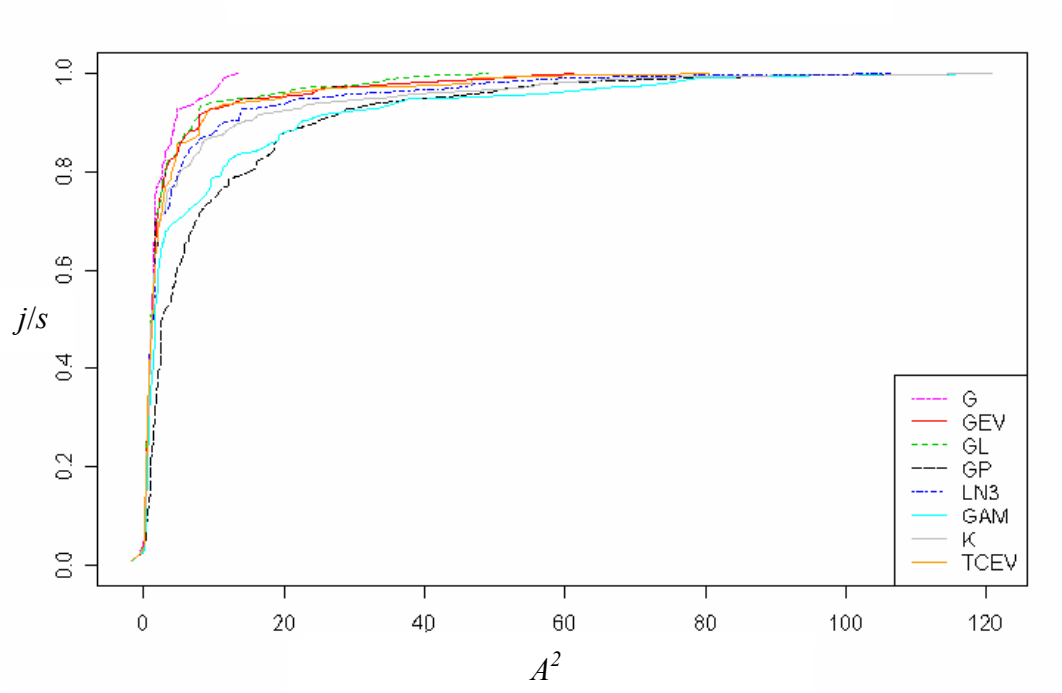


Figura 8.4. Funzione di frequenza empirica della statistica A^2 di Anderson-Darling calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Per verificare quale, tra le distribuzioni considerate, si comporta meglio secondo il test di Anderson-Darling si verifica quale delle curve in Figura 8.4 risulti più spostata a sinistra, ossia su valori più bassi della statistica A^2. Nel caso specifico le 8 distribuzioni si comportano in maniera abbastanza simile l'una all'altra, a parte le distribuzioni Gamma e di Pareto, che si comportano nettamente peggio delle altre, e la distribuzione di Gumbel, che invece si comporta mediamente meglio delle altre.

Si può anche verificare qual è la distribuzione che comporta il valore di A^2 più basso per ogni stazione (Tabella 8.3).

Distribuzione	GEV	GL	GP	LN3
N_s	6	17	1	6
Distribuzione	G	GAM	K	TCEV
N_s	46	6	11	4

Tabella 8.3. Numero di stazioni in cui la distribuzione considerata restituisce il valore di A^2 più basso.

Anche dalla Tabella 8.3 risulta che la distribuzione di Gumbel si comporta meglio delle altre secondo la statistica di Anderson-Darling.

Un altro modo per individuare la distribuzione più adeguata a descrivere la curva di crescita delle portate di piena è quello di calcolare, per ciascuna distribuzione, il valore di A^2 in corrispondenza della mediana. La distribuzione migliore sarà caratterizzata dal A^2_{median} minore (Tabella 8.4).

Distribuzione	GEV	GL	GP	LN3
A^2_{median}	1.20	1.04	2.64	1.24
Distribuzione	G	GAM	K	TCEV
A^2_{median}	1.17	1.55	1.15	1.21

Tabella 8.4. Valori di A^2 valutati in corrispondenza della mediana.

Anche la Tabella 8.4 conferma che la distribuzione di Pareto e la Gamma non sembrano appropriate a descrivere i campioni disponibili, mentre le altre si comportano circa allo stesso modo per quanto riguarda la statistica A^2_{median}.

RISULTATI DEL TEST DI UNIFORMITA'

Il diagramma in Figura 8.5 riporta in ascissa le variabili trasformate secondo l'equazione (8.4), ed in ordinata le corrispondenti frequenze empiriche. Il numero complessivo di dati considerati per costruire il diagramma è n_{tot} = 3156. Come spiegato in precedenza, le distribuzioni migliori sono quelle che si avvicinano maggiormente alla bisettrice del diagramma.

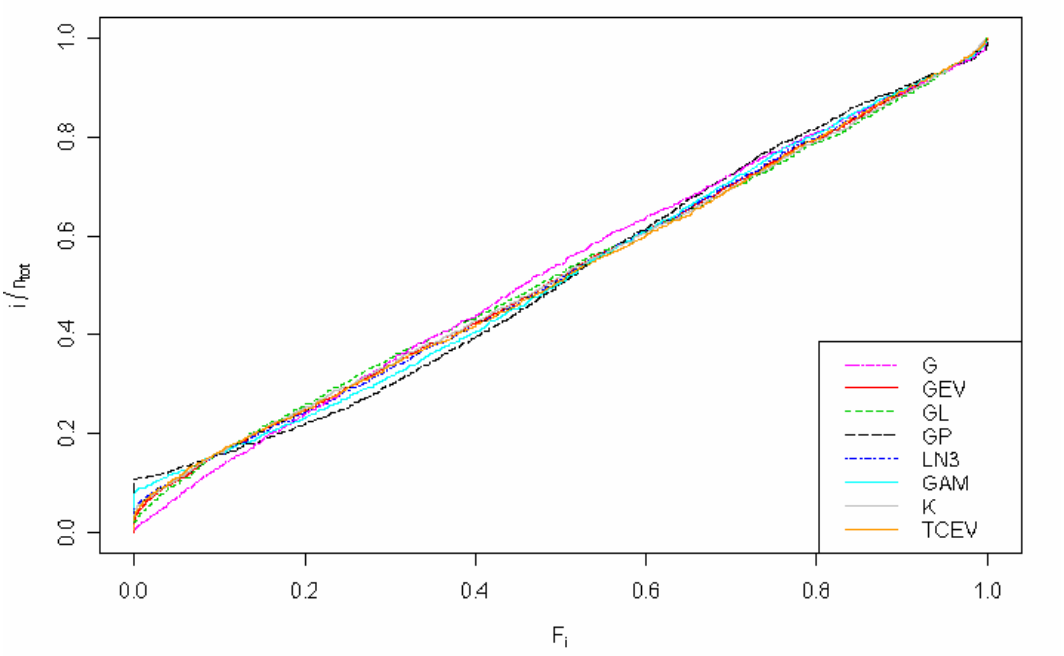


Figura 8.5. Funzione di frequenza empirica della statistica F_i di equazione (8.4) calcolata, sui 3156 dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Come nel caso del test di Cramer von Mises, anche con il test di uniformità le diverse distribuzioni producono risultai molto simili. Per provare a distinguere più accuratamente le distribuzioni, si è calcolata la massima distanza delle singole curve dalla bisettrice (Tabella 8.5).

Distribuzione	GEV	GL	GP	LN3
d_{max}	0.0614	0.0642	0.1065	0.0621
Distribuzione	G	GAM	K	TCEV
d_{max}	0.0497	0.0821	0.0626	0.0663

Tabella 8.5. Distanze massime dalla bisettrice calcolate per ciascuna distribuzione testata.

Anche la Tabella 8.5 fornisce una conferma del fatto che le diverse distribuzioni producono risultati tra loro molto simili quando si considera il test di uniformità. Anche in questo caso il risultato può essere spiegato considerando i grafici in Figura 8.1 e in Allegato 1, che evidenziano come le 8 distribuzioni hanno comportamento molto simile nella parte centrale, corrispondente a tempi di ritorno medio-bassi.

RISULTATI DEL TEST DEL VALORE MASSIMO

Come spiegato in precedenza, si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica ξ. I risultati sono riportati in Figura 8.6.

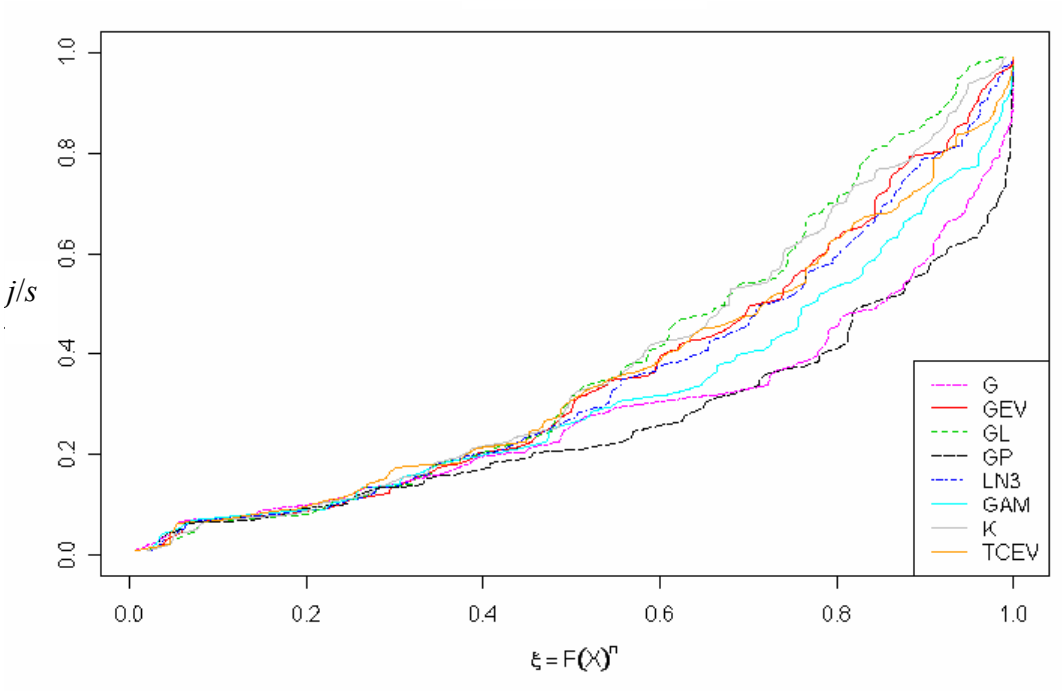


Figura 8.6. Funzione di frequenza empirica della statistica ξ del valore massimo calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Si nota chiaramente in Figura 8.6 che nessuna delle distribuzioni produce dei risultati particolarmente postivi in relazione a questo test. Infatti, tutte le curve di frequenza empirica risultano abbastanza distanti dalla bisettrice del diagramma, ed in particolare presentano con frequenza maggiore di quella attesa valori di ξ superiori a 0.5. Questo significa che le curve di crescita ottenute con L-coefficienti derivati dall'analisi multiregressiva tendono a sottostimare i valori massimi osservati. La sottostima risulta particolarmente rilevante per la distribuzione di Pareto e per la Gumbel.

A supporto dell'analisi grafica di Figura 8.6, si riporta ancora in Tabella 8.6 il numero di casi in cui, per ciascuna distribuzione si ha :

- $\xi > 0.95$, ossia si sottostima significativamente il massimo valore osservato;

- $\xi < 0.05$, ossia si sovrastima significativamente il massimo valore osservato.

Distribuzione	GEV	GL	GP	LN3
ξ < 0.05	3	3	4	4
ξ > 0.95	11	2	36	14
Distribuzione	G	GAM	K	TCEV
ξ < 0.05	4	4	3	4
ξ > 0.95	28	22	5	15

Tabella 8.6. Numero di sovrastime ($\xi < 0.05$) e di sottostime ($\xi > 0.95$) del massimo valore osservato.

Anche Tabella 8.6 fornisce una conferma del fatto che le distribuzioni, ed in particolare la distribuzione di Pareto, la Gumbel e la Gamma, tendono a sottostimare frequentemente i valori massimi. Prima di passare, nel successivo paragrafo 8.4, a considerare la scelta della distribuzione, si analizzano i risultati ottenuti del caso di stima degli L-coefficienti con dati di portata massima giornaliera.

8.2.2.2 Stima dalla serie delle portate estreme giornaliere

I parametri delle distribuzioni considerate, vengono stimati a partire dagli *L*-momenti calcolati a partire dalle serie di portata estrema giornaliera e dai parametri geomorfologici dei bacini come descritto nel paragrafo 7.3.1 e 7.3.2. In particolare la stima di L_{cv} e di L_{ca} viene effettuata, rispettivamente, mediante la relazione (7.3) e (7.4).

RISULTATI DEL TEST DI CRAMER VON-MISES

Si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica D^2. I risultati sono riportati in Figura 8.7.

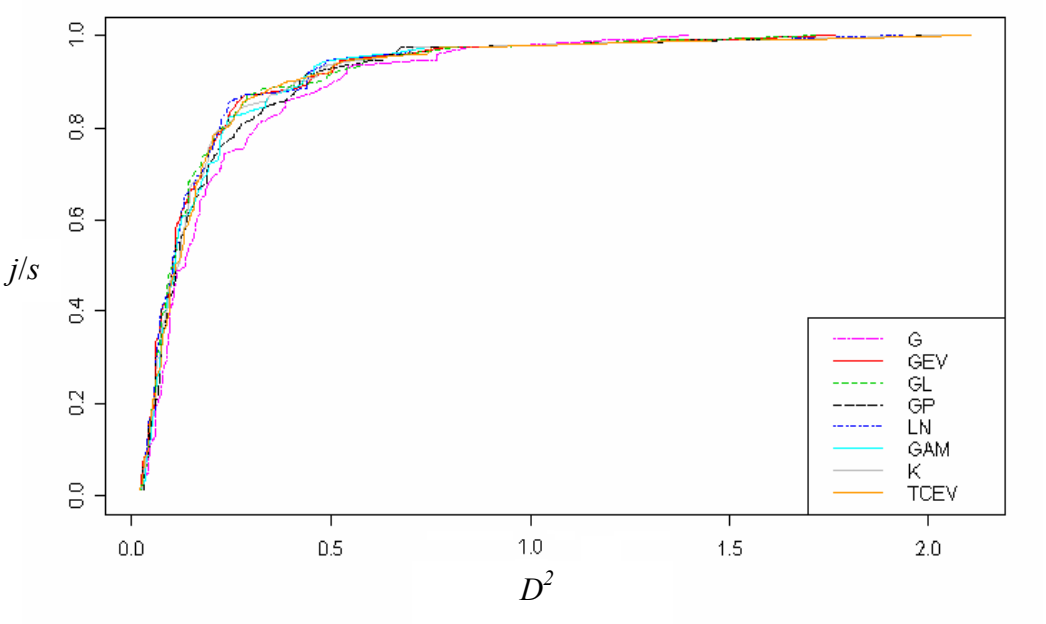


Figura 8.7. Funzione di frequenza empirica della statistica D^2 di Cramer von Mises calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Per verificare quale, tra le distribuzioni considerate, si comporta meglio secondo il test di Cramer – von Mises si verifica quale delle curve in Figura 8.7 risulti più spostata a sinistra, ossia su valori più bassi della statistica D^2. Anche in questo caso le 8 distribuzioni si comportano in maniera molto simile l'una all'altra. Si può anche verificare qual è la distribuzione che comporta il valore di D^2 più basso per ogni stazione (Tabella 8.7).

Distribuzione	GEV	GL	GP	LN3
N_s	7	18	12	5
Distribuzione	G	GAM	K	TCEV
N_s	15	12	4	5

Tabella 8.7. Numero di stazioni in cui la distribuzione considerata restituisce il valore di D^2 più basso.

Tutte le distribuzioni considerate risultano ottime in almeno 4 stazioni, e le distribuzioni migliori secondo Tabella 8.7 risulterebbero la distribuzione di Gumbel e la distribuzione Logistica, anche se le differenze sono di scarsa entità.

Come prima, si può anche calcolare, per ciascuna distribuzione, il valore di D^2 in corrispondenza della mediana. La distribuzione migliore sarà caratterizzata dal D^2_{median} minore (Tabella 8.8).

Distribuzione	GEV	GL	GP	LN3
D^2_{median}	0.103	0.107	0.117	0.103
Distribuzione	G	GAM	K	TCEV
D^2_{median}	0.135	0.107	0.118	0.115

Tabella 8.8. Valori di D^2 valutati in corrispondenza della mediana.

Anche in questo caso si nota, come già in Figura 8.6, che anche nel caso delle stime con gli estremi giornalieri il test di Cramer von Mises non permette di identificare grandi differenze tra le 8 distribuzioni considerate. Complessivamente, i valori di D^2_{median} ottenuti sono più bassi di quelli riscontrati in precedenza (Tabella 8.2), a testimonianza della maggiore qualità delle stime degli L-momenti rispetto a quelle ottenute con le tecniche multiregressive.

RISULTATI DEL TEST DI ANDERSON - DARLING

Si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica A^2. I risultati sono riportati in Figura 8.8.
Il diagramma in Figura 8.8 è abbastanza difficile da leggere, a causa di alcuni valori di A^2 molto elevati. Conviene in questo caso riferirsi ai risultati delle Tabelle 8.9 ed 8.10. In Tabella 8.9 si riporta il Numero di stazioni in cui la distribuzione considerata restituisce il valore di A^2 più basso.

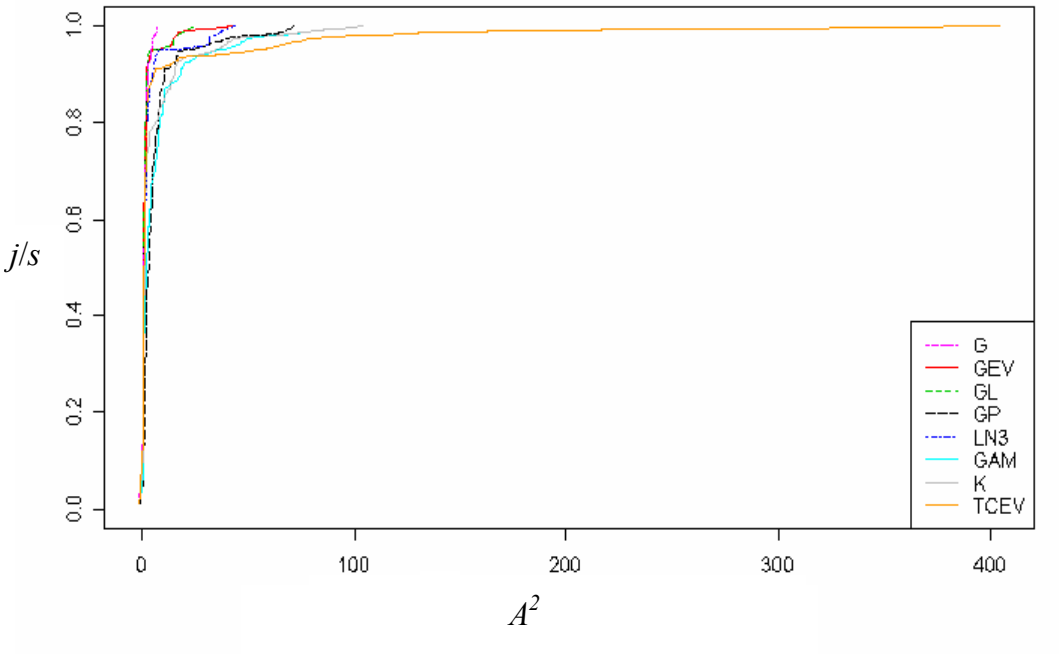


Figura 8.8. Funzione di frequenza empirica della statistica A^2 di Anderson-Darling calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Distribuzione	GEV	GL	GP	LN3
N_s	1	15	1	3
Distribuzione	G	GAM	K	TCEV
N_s	30	6	8	5

Tabella 8.9. Numero di stazioni in cui la distribuzione considerata restituisce il valore di A^2 più basso.

Dalla Tabella 8.9 risulta che la distribuzione di Gumbel si comporta meglio delle altre secondo la statistica di Anderson-Darling.

Calcolando il valore di A^2 in corrispondenza della mediana si ottengono i risultati in Tabella 8.10.

Distribuzione	GEV	GL	GP	LN3
A^2_{median}	0.580	0.511	3.298	0.844
Distribuzione	G	GAM	K	TCEV
A^2_{median}	0.658	2.243	0.882	0.786

Tabella 8.10. Valori di A^2 valutati in corrispondenza della mediana.

La Tabella 8.10 dimostra che la distribuzione di Pareto e la Gamma non sembrano appropriate a descrivere i campioni disponibili, mentre le altre si comportano circa allo stesso modo per quanto riguarda la statistica A^2_{median}. Anche per il test di Anderson-Darling i valori di A^2_{median} ottenuti sono più bassi di quelli riscontrati in precedenza (Tabella 8.5).

RISULTATI DEL TEST DI UNIFORMITA'

Come in precedenza, il diagramma in Figura 8.9 riporta in ascissa le variabili trasformate secondo l'equazione (8.4), ed in ordinata le corrispondenti frequenze empiriche.

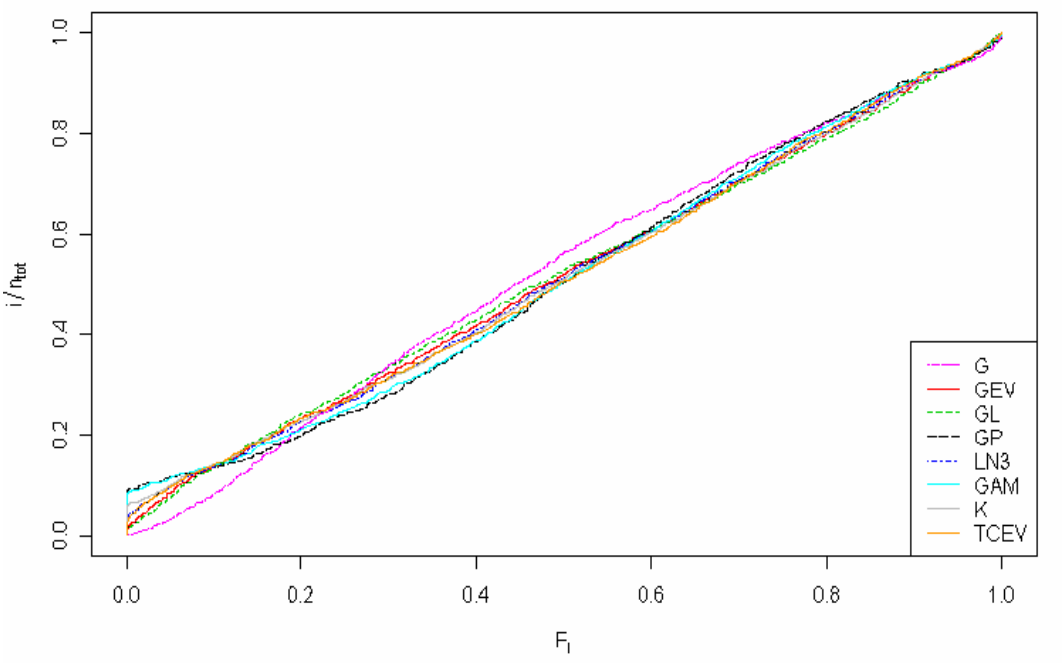


Figura 8.9. Funzione di frequenza empirica della statistica F_i di equazione (8.4) calcolata, sui 3156 dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Anche con il test di uniformità le diverse distribuzioni producono risultati molto simili. Per provare a distinguere più accuratamente le distribuzioni, si è calcolata la massima distanza delle singole curve dalla bisettrice (Tabella 8.11).

Distribuzione	GEV	GL	GP	LN3
d_{max}	0.0428	0.0459	0.0904	0.0496
Distribuzione	G	GAM	K	TCEV
d_{max}	0.0656	0.0858	0.0596	0.0501

Tabella 8.11. Distanze massime dalla bisettrice calcolate per ciascuna distribuzione testata.

Anche la Tabella 8.11 fornisce una conferma del fatto che le diverse distribuzioni producono risultati tra loro molto simili quando si considera il test di uniformità.

RISULTATI DEL TEST DEL VALORE MASSIMO

Come spiegato in precedenza, si diagramma, per ognuna delle 8 distribuzioni, la funzione di frequenza empirica della statistica ξ (Figura 8.10).

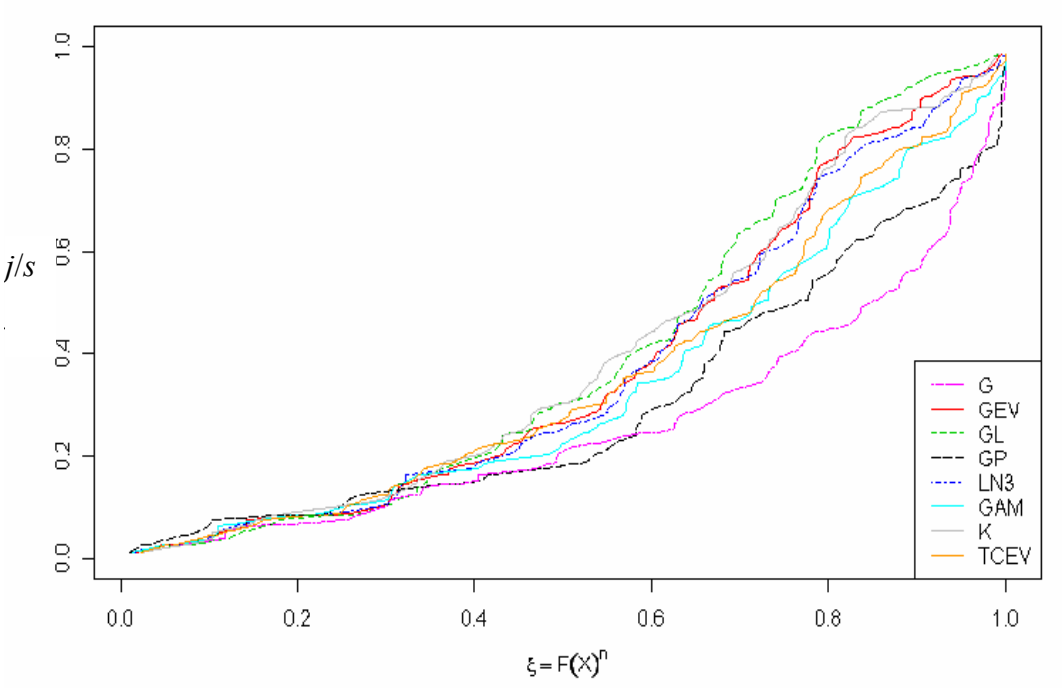


Figura 8.10. Funzione di frequenza empirica della statistica ξ del valore massimo calcolata, sulle 97 stazioni con dati disponibili, per le 8 distribuzioni considerate nel lavoro.

Come nel caso delle stime da analisi multiregressiva, nessuna delle distribuzioni produce dei risultati particolarmente positivi, in quanto tutte le curve di frequenza empirica risultano abbastanza distanti dalla bisettrice del

diagramma. In questo caso le deviazioni verso il basso (ossia la tendenza alla sottostima) sembrano essere meno marcate che in precedenza (Figura 8.6), ma la sottostima rimane rilevante in particolare per le distribuzioni di Pareto e di Gumbel.
A supporto dell'analisi grafica di Figura 8.10, si riporta ancora in Tabella 8.12 il numero di casi in cui, per ciascuna distribuzione si ha :

- $\xi > 0.95$, ossia si sottostima significativamente il massimo valore osservato;
- $\xi < 0.05$, ossia si sovrastima significativamente il massimo valore osservato.

Distribuzione	GEV	GL	GP	LN3
ξ < 0.05	2	1	2	1
ξ > 0.95	4	3	18	4
Distribuzione	G	GAM	K	TCEV
ξ < 0.05	1	1	1	1
ξ > 0.95	20	11	6	7

Tabella 8.12. Numero di sovrastime ($\xi < 0.05$) e di sottostime ($\xi > 0.95$) del massimo valore osservato.

Escludendo le distribuzioni di Gumbel, di Pareto e la distribuzione Gamma, che tendono a sottostimare frequentemente i valori massimi, le altre distribuzioni sembrano comportarsi abbastanza bene in questi casi, con un numero di sottostime che si avvicina al valore atteso (4.85). Nel successivo paragrafo 8.3 si completa il confronto tra le 8 distribuzioni, andando a selezionare quella che si ritiene più appropriata. Si noti che, per scegliere la distribuzione più adatta all'analisi regionale, non si è considerato il caso in cui gli L-momenti sono stimati dai dati empirici di portata al colmo, perché in tal caso l'analisi di frequenza non è più un'analisi regionale ma diventa un'analisi at site.

8.3 Scelta della distribuzione

Come accennato, lo scopo dei test di adattamento applicati nel precedente paragrafo 8.2 è selezionare la distribuzione più appropriata ai fini della rappresentazione della curva di crescita delle portate di piena. Tuttavia si è riscontrato che i test non risultano particolarmente efficaci a tal fine: infatti, spesso le 8 distribuzioni producono risultati molto simili tra loro, cosa che impedisce di scegliere oggettivamente quale distribuzione possa ritenersi la migliore. Il problema, come già ribadito in precedenza, sta nel fatto che le 8 distribuzioni hanno andamento molto simile per tempi di ritorno medio-bassi (cfr. Allegato 1). D'altra parte, la scelta della distribuzione influenza in maniera rilevante i risultati che si ottengono nell'analisi regionale, perché le distribuzioni hanno andamenti molto diversi per tempi di ritorno elevati (in molti dei grafici in allegato 1 il fascio di curve corrispondenti alle 8 distribuzioni si apre al di sopra di T = 100 anni). Da una parte ci si trova quindi con test di adattamento non in grado di selezionare oggettivamente una distribuzione, e dall'altra si è invece nella necessità di scegliere uno specifico modello probabilistico per le successive analisi. La situazione è analoga a quella che si ha nella modellistica idrologica, quando si trovano diversi modelli, o diversi valori dei parametri, ugualmente adatti a rappresentare i dati reali (principio dell'equifinalità). La soluzione che si può adottare in questi casi può essere il non-scegliere, ossia utilizzare tutti i modelli probabilistici per stimare K(T), e poi utilizzare come stimatore la media (o la mediana) dei valori così trovati. Questa procedura (chiamata model averaging nella letteratura anglosassone) è sicuramente adeguata dal punto di vista statistico, ma è tuttavia piuttosto complicata nelle applicazioni pratiche, dal momento che richiede di stimare 8 modelli probabilistici in ogni sezione di interesse. Si è dunque deciso di adottare una strategia simile ma meno complicata: scegliere un singolo modello probabilistico che produca delle stime di K(T) simili a quelle che si troverebbero, per tempi di ritorno elevati, adottando la procedura di model averaging.

Si è dunque deciso di procedere in questo modo:

- fissare T = 500 anni,
- calcolare il valore di K(T) per ciascuna distribuzione in esame;
- annotare, per ciascuna delle 97 sezioni in esame, qual è la distribuzione che a T = 500 restituisce il valore di K(T) più basso

(codice 1) fino ad arrivare alla distribuzione a cui corrisponde il K(T) più alto (codice 8);

- selezionare la distribuzione che si presenta più frequentemente nelle posizioni intermedie (corrispondenti al model averaging con mediana), ossia nelle caselle 4 e 5.

I risultati ottenuti con questa procedura sono riportati in tabella 8.13 per il caso in cui gli L-momenti sono stimati tramite regressioni multiple, ed in tabella 8.14 per il caso in cui la stima degli L-momenti si basa anche sugli estremi delle portate medie giornaliere.

Distribuzione	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
G	45.4	22.7	9.3	0.0	17.5	4.1	1.0	0.0
GEV	0.0	11.3	10.3	1.0	15.5	33.0	28.9	0.0
GL	0.0	0.0	0.0	0.0	4.1	3.1	44.3	48.5
GP	54.6	19.6	2.1	14.4	9.3	0.0	0.0	0.0
LN3	0.0	0.0	1.0	45.4	24.0	19.6	0.0	0.0
GAM	0.0	34	50.5	15.5	0.0	0.0	0.0	0.0
K	0.0	0.0	0.0	9.3	11.3	16.5	17.5	45.4
TCEV	0.0	12.4	26.8	14.4	8.3	23.7	8.3	6.2

Tabella 8.13. Percentuale di volte in cui ogni distribuzione restituisce un valore di K(T=500) nella posizione indicata nella prima riga (da 1 a 8) nel campione ordinato in senso crescente; L-momenti stimati tramite regressione multipla.

Distribuzione	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
G	53.9	20.5	5.1	0.0	1.3	3.9	6.4	0.0
GEV	0.0	15.4	2.6	2.6	9.0	29.5	41.0	0.0
GL	0.0	0.0	0.0	0.0	9.0	2.6	32.1	56.4
GP	44.9	15.4	3.9	14.1	15.4	6.4	0.0	0.0
LN3	0.0	0.0	2.6	33.3	38.4	56.6	0.0	0.0
GAM	0.0	24.4	50.0	25.6	0.0	0.0	0.0	0.0
K	1.3	3.9	3.9	12.8	11.5	16.7	15.4	34.6
TCEV	0.0	20.5	32.1	11.5	6.4	15.4	5.1	9.0

Tabella 8.14. Percentuale di volte in cui ogni distribuzione restituisce un valore di K(T=500) nella posizione indicata nella prima riga (da 1 a 8) nel campione ordinato in senso crescente; L-momenti stimati tramite l'uso degli estremi giornalieri.

Analizzando le Tabelle 8.13 e 8.14 si riscontra chiaramente che alcune distribuzioni hanno un comportamento molto ben definito nella coda alta della

distribuzione. Ad esempio, la distribuzione di Gumbel e la distribuzione di Pareto restituiscono valori di K(T=500) che sono quasi sempre tra i più bassi; viceversa, la distribuzione Kappa e la distribuzione Logistica tendono a produrre stime molto elevate di K(T=500); la distribuzione Gamma e la GEV sono un po' più centrali rispetto alle altre, ma denotano una significativa tendenza a porsi nella metà inferiore (Gamma) o superiore (GEV) del campione di 8. Infine, la TCEV risulta distribuita su tutte le posizioni, mentre la distribuzione Lognornale a 3 parametri risulta nettamente la migliore in termini di centralità con percentuali intorno al 70-80% nelle caselle 4 e 5 della tabella.

Un risultato analogo si ottiene considerando, per ognuna delle 97 sezioni, l'errore quadratico tra la stima di K(T=500) ottenuta con ognuna delle 8 distribuzioni, e la stima model averaging ottenuta come media degli 8 valori. Calcolando l'errore quadratico medio sulle 97 stazioni si ottiene che il valore più basso è proprio quello corrispondente alla distribuzione Lognormale a 3 parametri, ossia la distribuzione LN3 è la più simile alla stima che si otterrebbe tramite model averaging.

In base a questa analisi si è dunque deciso di scegliere la distribuzione Lognormale a 3 parametri, definita come in Appendice D (paragrafo D5), per le successive elaborazioni. Si noti che la distribuzione Lognormale è comunque tra le migliori anche in relazione ai test di adattamento presentati nel paragrafo 8.2.

8.4 Valutazione dell'incertezza di stima

Una volta scelto il modello probabilistico da impiegare per la stima della curva di crescita, si passa a considerare i metodi per la valutazione dell'incertezza di stima di K(T). Un possibile indicatore dell'incertezza di stima è dato dalla varianza di K(T), $\sigma^2_{k(T)}$, che risulta essere essa stessa funzione del tempo di ritorno considerato, oltre che della varianza di stima di L-CV ed L-CA, ottenute al capitolo 6. Tuttavia la stima della varianza di K(T) richiede l'impiego di metodi asintotici (validi per dimensioni campionarie molto elevate) e, inoltre, non risolve completamente il problema della valutazione dell'incertezza di stima di K(T); infatti, l'asimmetria delle distribuzioni considerate e la non linearità delle relazioni tra K(T) ed LCA rendono marcatamente non-normale la distribuzione di K(T), impedendo quindi la

costruzione di intervalli di confidenza anche qualora si disponesse di una buona stima di $\sigma^2_{k(T)}$.

Si è preferito dunque procedere direttamente alla definizione degli intervalli di confidenza, secondo la procedura Monte-Carlo descritta nei seguenti sotto-paragrafi.

8.4.1 Stima empirica

Si considera innanzitutto il caso in cui sono disponibili dati di portata al colmo di piena e si vuole stimare K(T) partendo dalle stime empiriche di LCV ed LCA. In tal caso la procedura Monte-Carlo per definire gli intervalli di confidenza prevede i seguenti passaggi:

1) fissata una stazione *j*, si stimano LCV ed LCA tramite i dati disponibili (stima empirica);
2) si determinano le varianze e la covarianza degli stimatori di LCV ed LCA tramite le relazioni (6.8)-(6.10);
3) si estraggono 1000 coppie di LCV ed LCA da una distribuzione normale bivariata, i cui 5 parametri sono le due medie determinate al punto 1) e le varianze e covarianza determinate al punto 2);
4) per ognuna delle *i* coppie di LCV ed LCA estratti dalla normale bivariata si calcola $K_i(T)$, i=1,…,1000, per ogni tempo di ritorno di interresse;
5) si fissa un coefficiente di confidenza α (ossia, una probabilità di ricadere all'interno dell'intervallo di confidenza) e si calcolano i limiti inferiore e superiore, $\{l_{\inf}(\alpha), l_{\sup}(\alpha)\}$, dell'intervallo di confidenza di K(T) prendendo il *p*-esimo elemento del campione ordinato in senso crescente dei $K_i(T)$, con $p = \left(0.5 - \frac{\alpha}{2}\right) \cdot 1000$ per il limite inferiore e $p = \left(0.5 + \frac{\alpha}{2}\right) \cdot 1000$ per il limite superiore;
6) si ripetono i punti da 1) a 5) per tutte le stazioni, j = 1, ..., 97.

Si noti che la procedura qui utilizzata non presuppone che la distribuzione di K(T) sia normale, ma, più ragionevolmente, che sia normale la distribuzione di LCV ed LCA.

8.4.2 Modello di regressione multipla

Nel caso di stima di LCV ed LCA tramite regressioni multiple la procedura per determinare gli intervalli di confidenza di K(T) è molto simile a quella descritta per la stima empirica. Le uniche differenze sono che al punto 1) si considerano ovviamente le stime di LCV ed LCA da regressioni invece delle stime empiriche, mentre al punto 2) si utilizza la procedura indicata al paragrafo 6.2.4 per stimare le varianze (la covarianza è in questo caso nulla).

8.4.3 Modello di stima basato sugli estremi giornalieri

Anche nel caso di stima di LCV ed LCA usando gli estremi giornalieri la procedura per determinare gli intervalli di confidenza di K(T) è simile a quella descritta per la stima empirica. Le differenze sono nuovamente al punto 1) (si considerano le stime di LCV ed LCA ottenute usando i dati giornalieri invece delle stime empiriche), oltre che al punto 2) (si utilizzano le procedure descritta al paragrafo 6.3.3 per stimare le varianze, mentre la covarianza è in anche questo caso nulla).

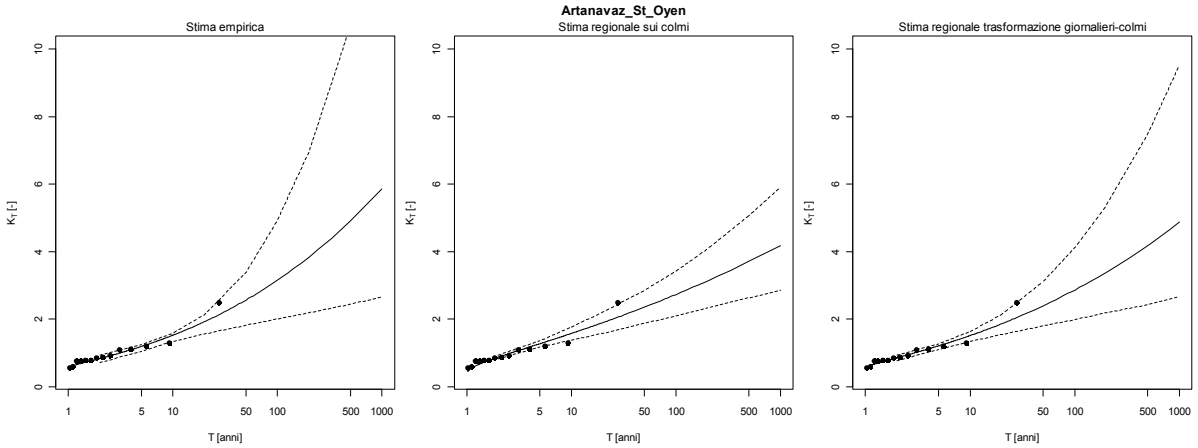
9 Applicazione dei metodi di stima della curva di crescita

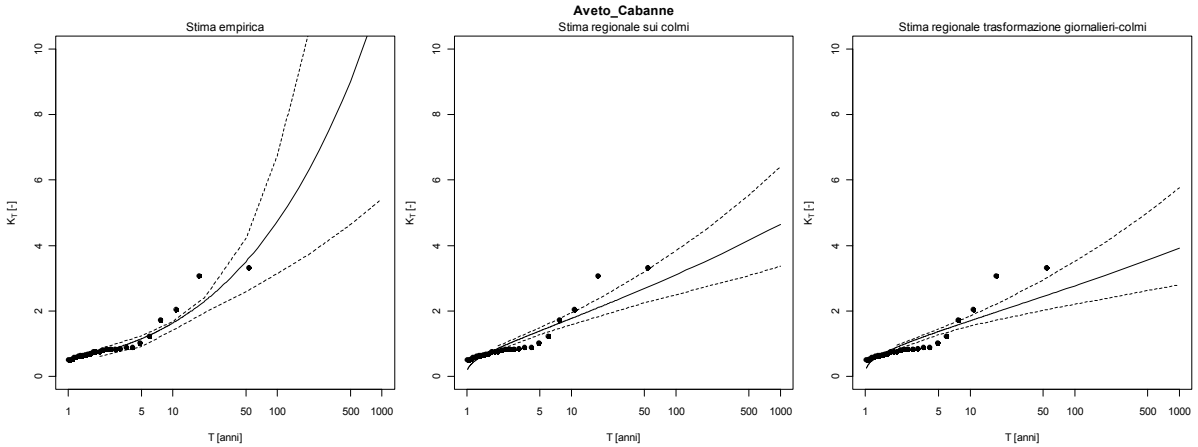
Nei precedenti capitoli sono stati presentati diversi approcci utili alla stima della curva di crescita. In particolare sono state sviluppate analisi multiregressive ed è stato predisposto un metodo per utilizzare gli estremi giornalieri nella stima della curva di crescita, oltre a considerare la stima empirica degli L-momenti a partire dal campione di osservazioni disponibile per le sezioni strumentate.

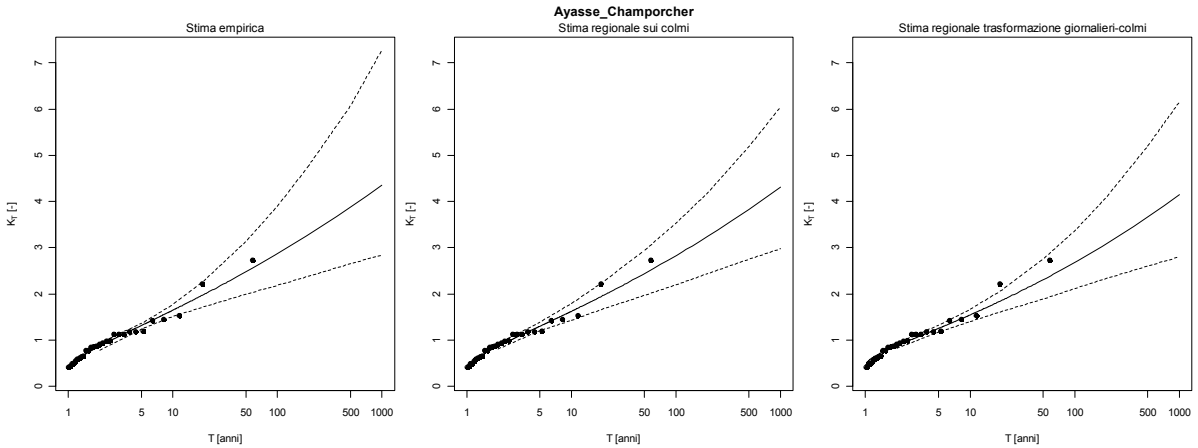
Una volta stabilita per ciascuna metodologia la relazione più adeguata per la stima della curva di crescita è possibile applicarla a tutti i bacini analizzati nel presente lavoro. Di seguito sono riportati i risultati di tale applicazione. In particolare, la prima colonna di grafici fa riferimento ai risultati che si ottengono considerando la stima degli L-momenti tramite dati empirici, la seconda colonna si riferisce alla stima regionale tramite analisi multiregressiva e la terza colonna alla stima della curva di crescita basata sugli estremi giornalieri. In ogni grafico sono riportati anche i relativi intervalli di confidenza, con livello di confidenza α=0.80 (ossia con probabilità 0.8 di stare all'interno dell'intervallo), relativi alla curva di crescita stimata con i diversi metodi. Gli intervalli di confidenza che derivano dall'applicazione della procedura descritta al paragrafo 8.5 sono spesso asimmetrici intorno al valore stimato, dal momento che non si è fatta alcuna ipotesi di normalità della distribuzione di K(T).

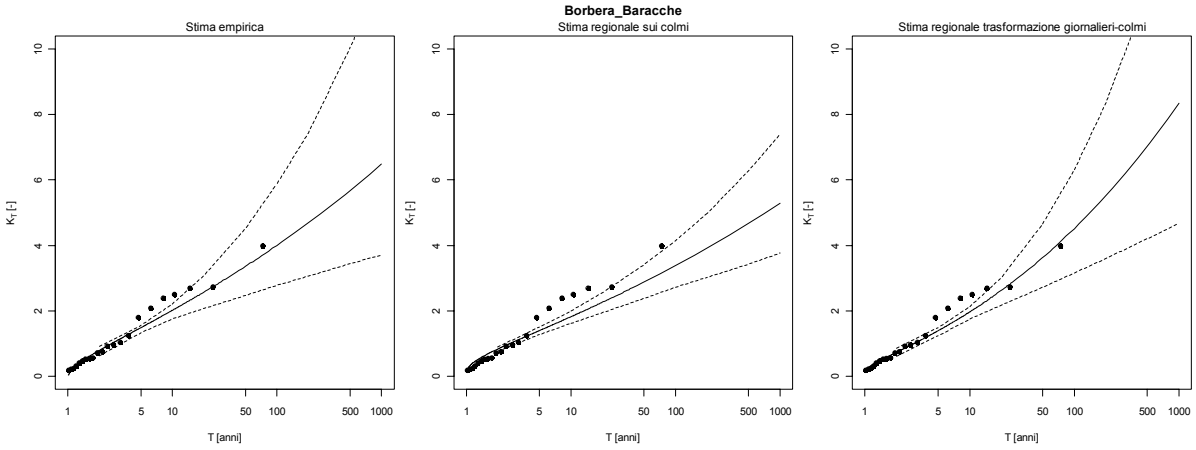
I suddetti intervalli ottenuti per la stima empirica (prima colonna di grafici) hanno ampiezza molto diversa a seconda della dimensione del campione e risultano particolarmente estesi intorno al valore centrale quando si hanno pochi dati. Nel caso di stima regionale (seconda colonna di grafici) l'ampiezza degli intervalli di confidenza è invece abbastanza costante nelle diverse sezioni, dal momento che la varianza degli stimatori regionali non dipende dalla specifica dimensione del campione. Infine l'ampiezza degli intervalli di confidenza riportati nei grafici della terza colonna varia da sezione a sezione, dal momento che la varianza degli stimatori regionali di LCV ed LCA dipende in questo caso anche dalla dimensione del campione di dati giornalieri.

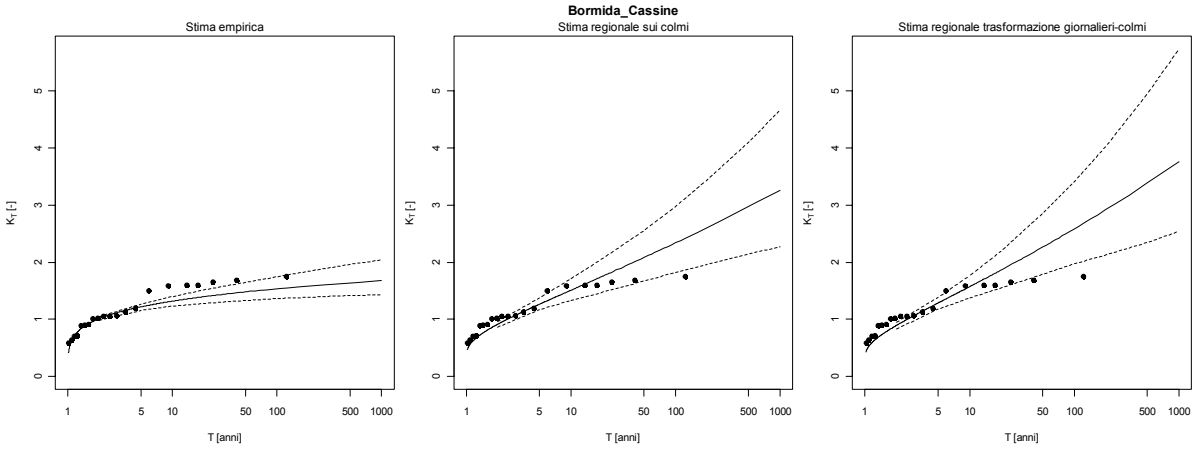
Alcuni grafici della colonna 3 sono mancanti perché in tali sezioni non sono disponibili dati di portata giornaliera.

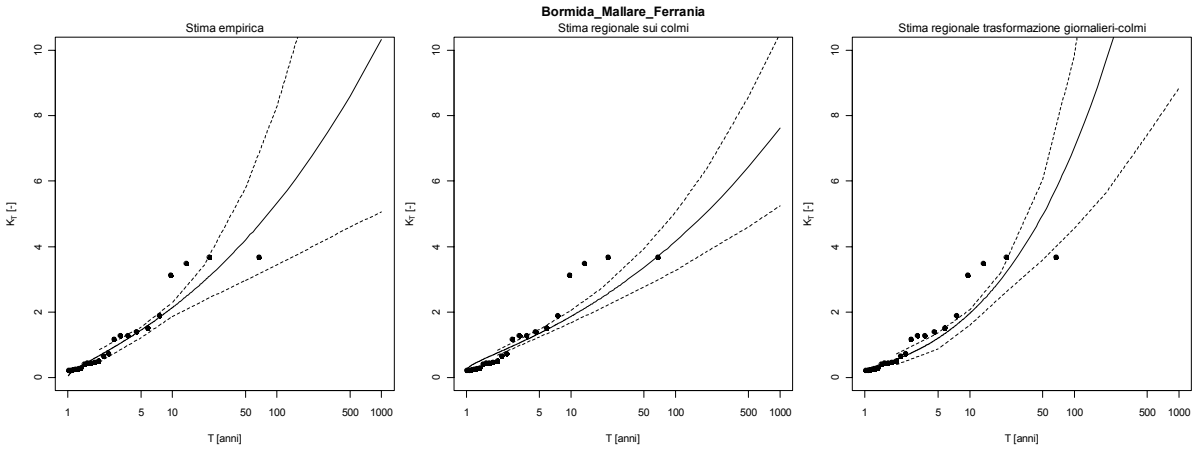

Artanavaz_St_Oyen
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000


Aveto_Cabanne
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

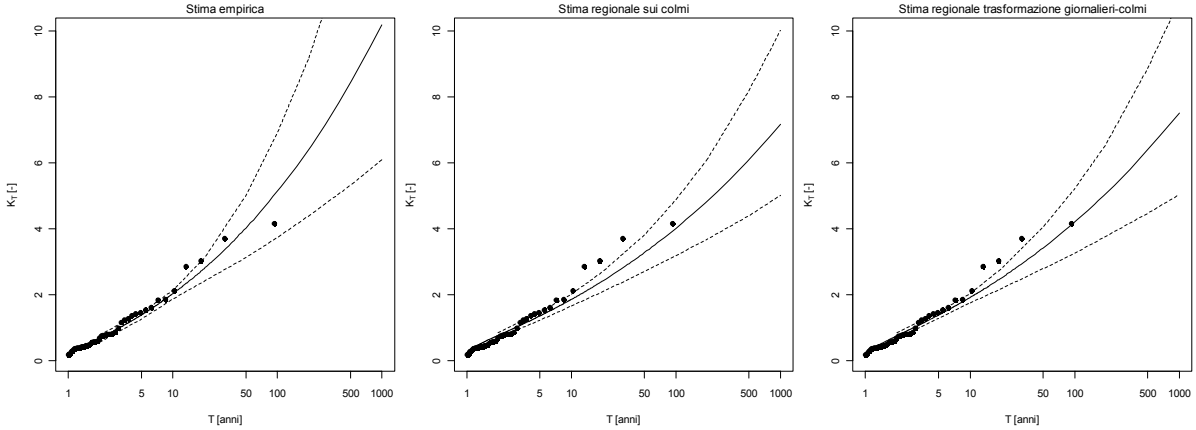

Ayasse_Champorcher
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
1
2
3
4
5
6
7
1
5
10
50
100
500
1000


Borbera_Baracche
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

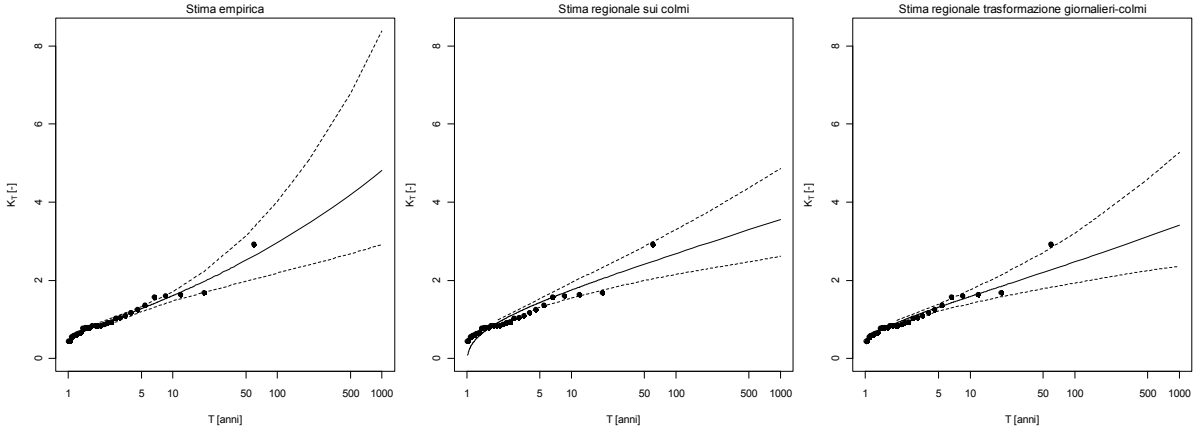

Bormida_Cassine
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
1
2
3
4
5
1
5
10
50
100
500
1000


Bormida_Mallare_Ferrania
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

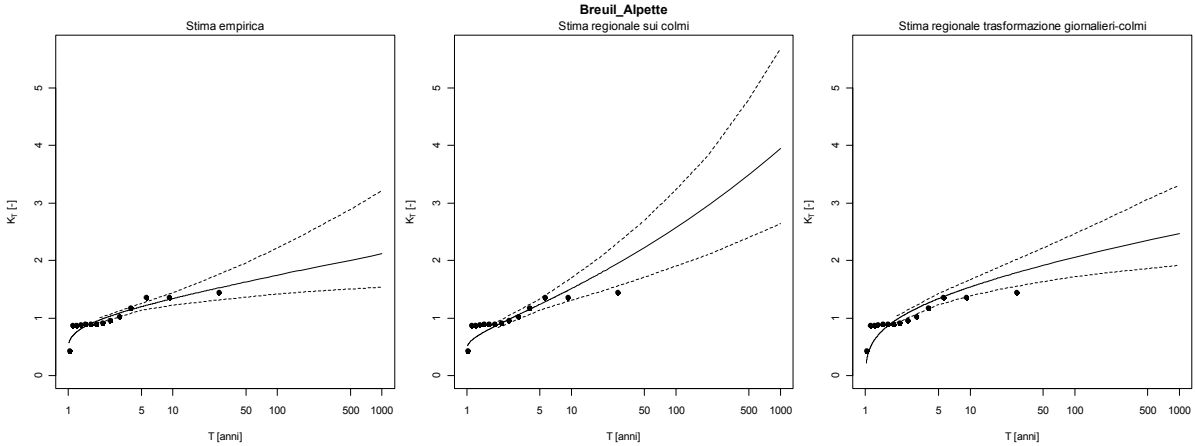
Bormida_Spigno_Valla

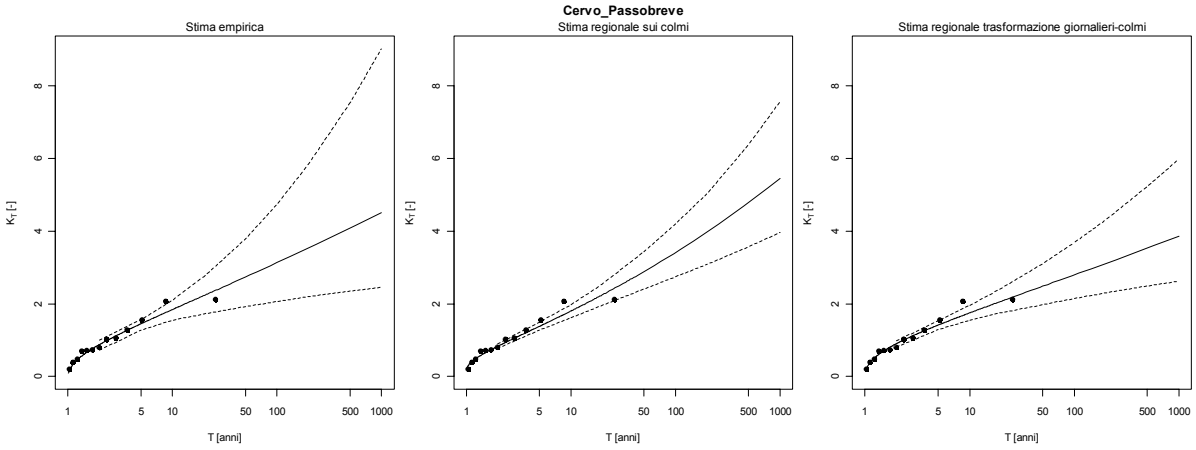


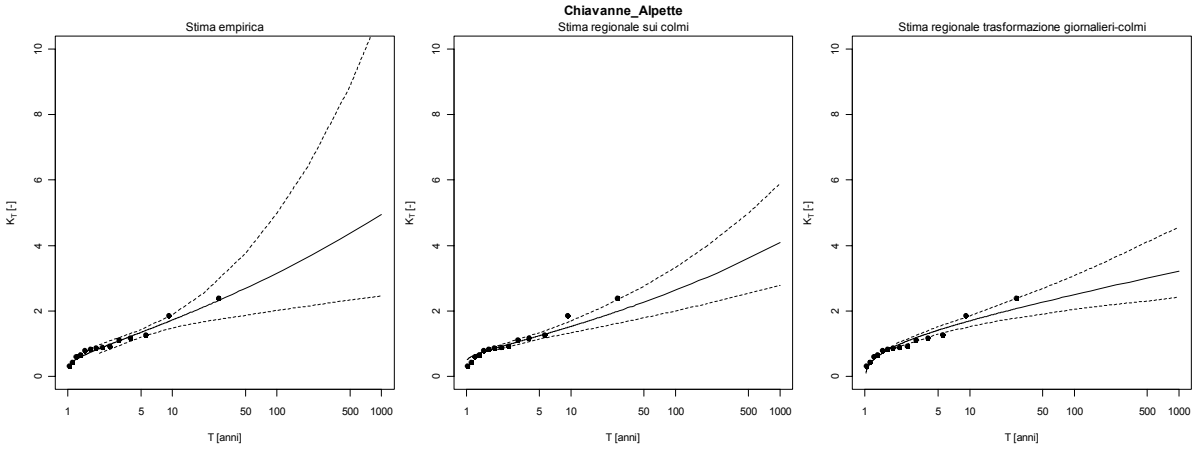
Brembo_P.te_Briolo

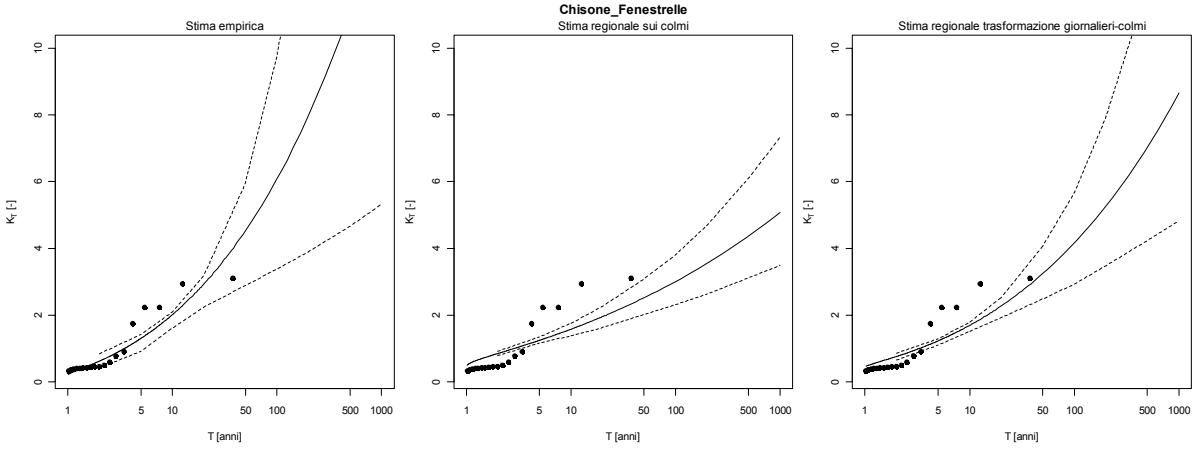


Breuil_Alpette

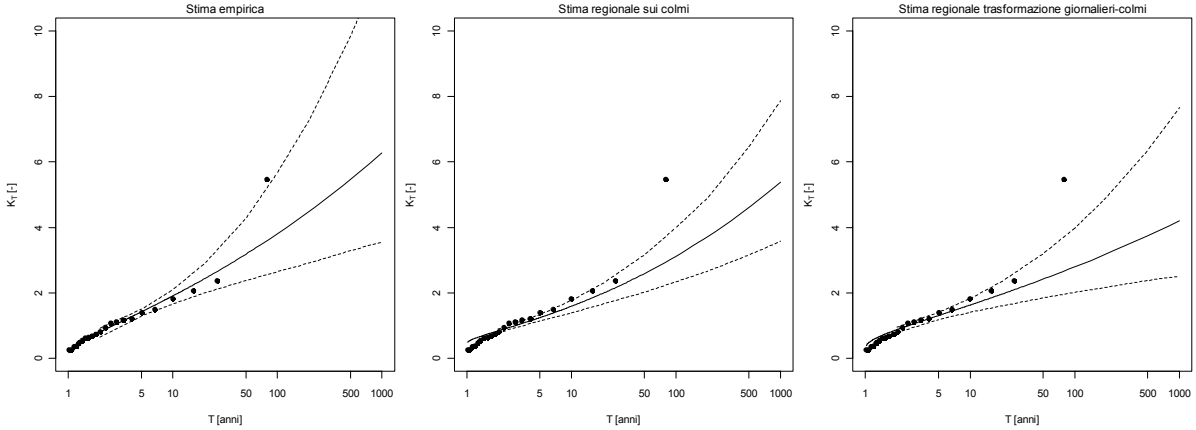



Cervo_Passobreve
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8

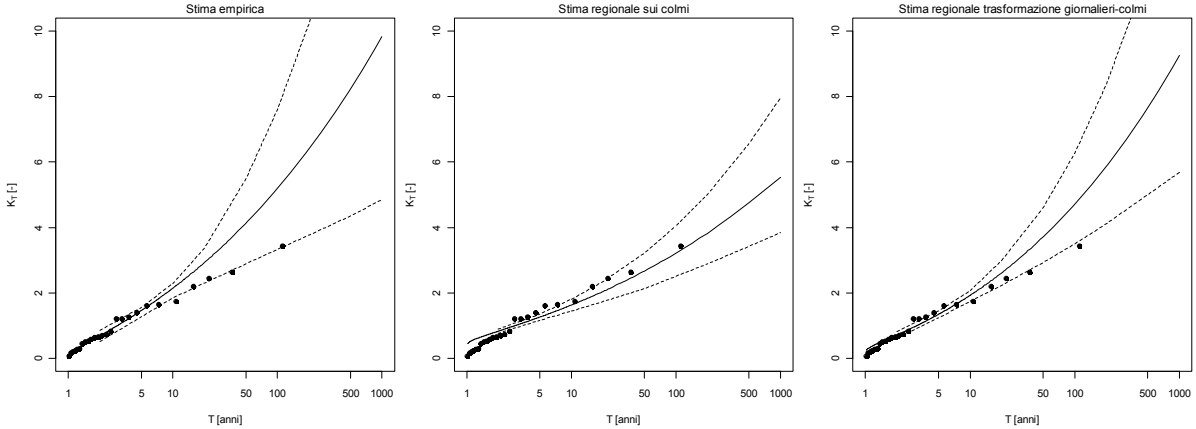

Chiavanne_Alpette
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10


Chisone_Fenestrelle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

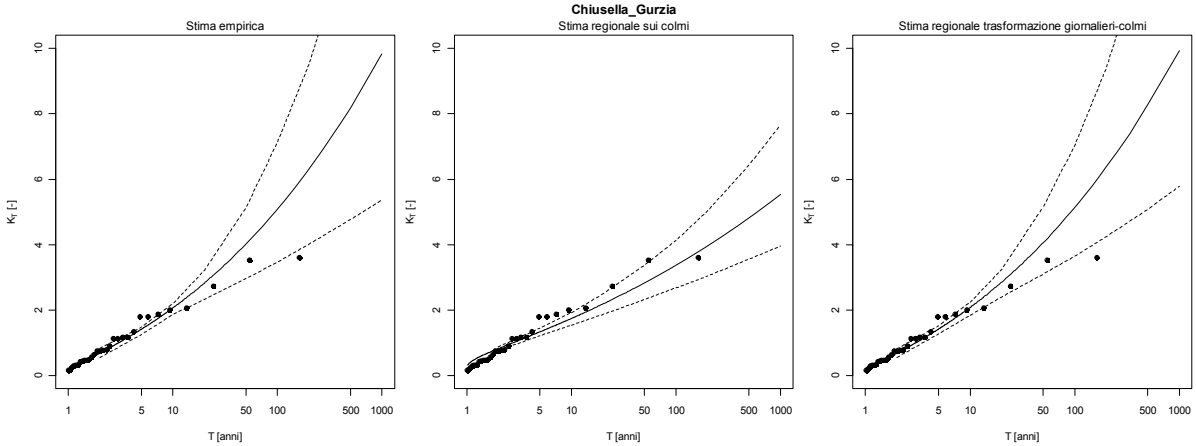
Chisone_S.Basses

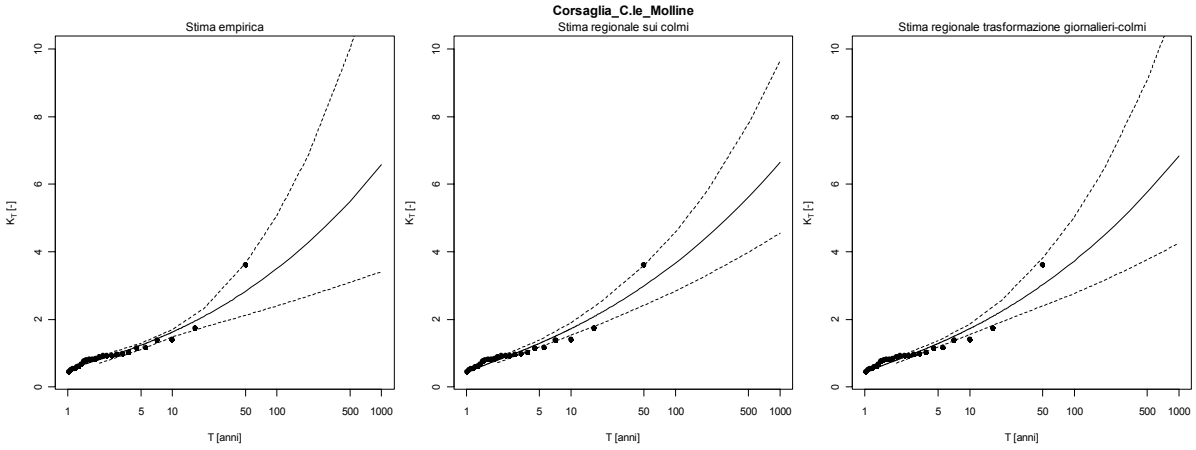


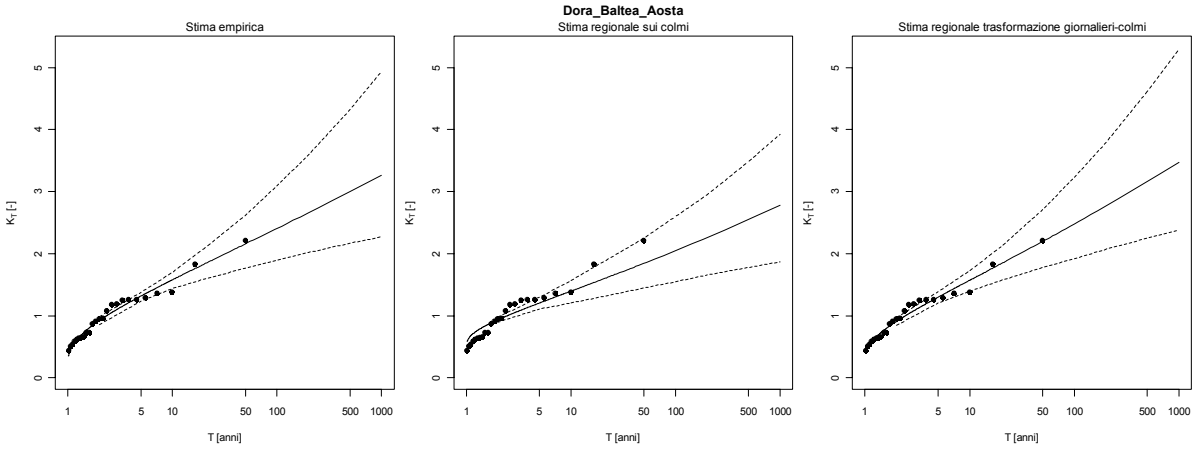
Chisone_S.Martino

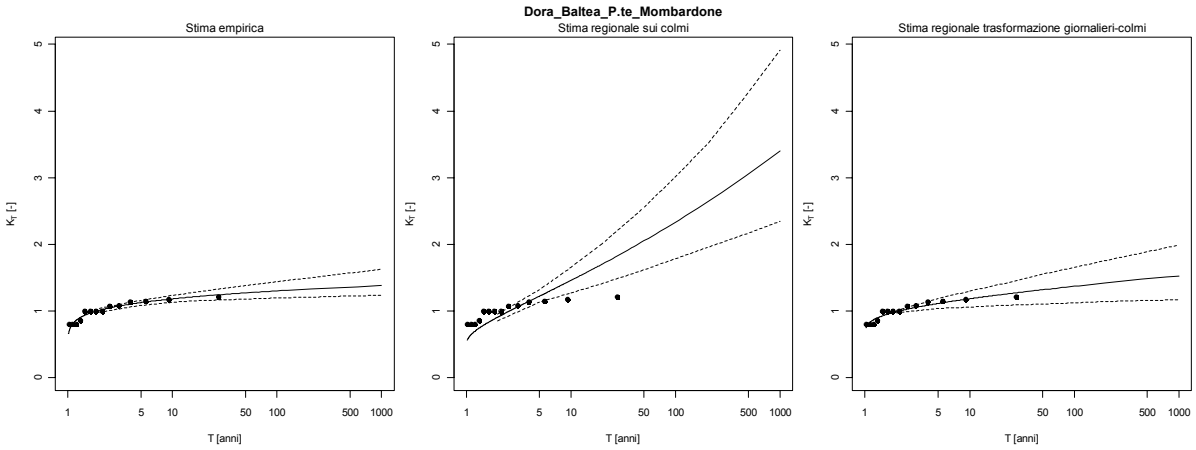


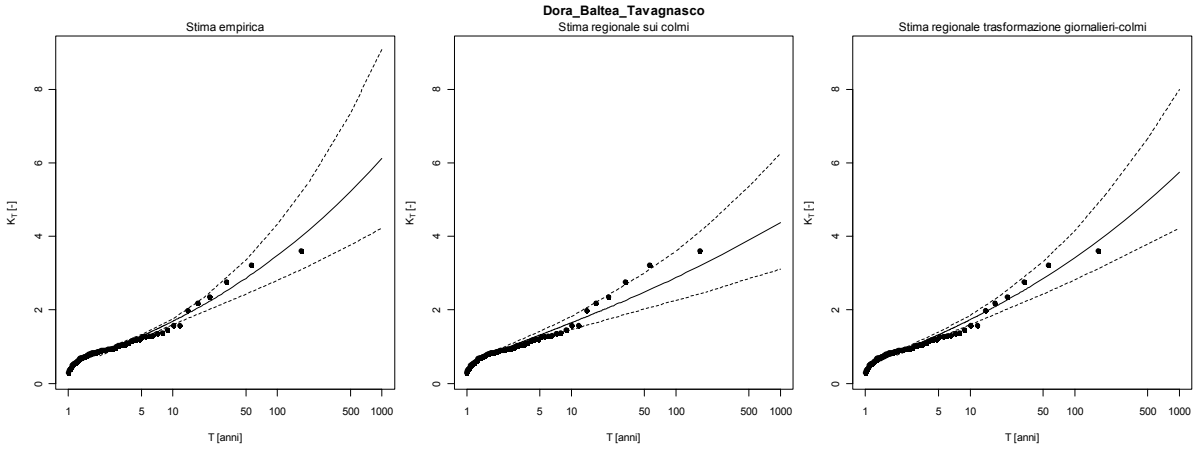
Chiusella_Gurzia

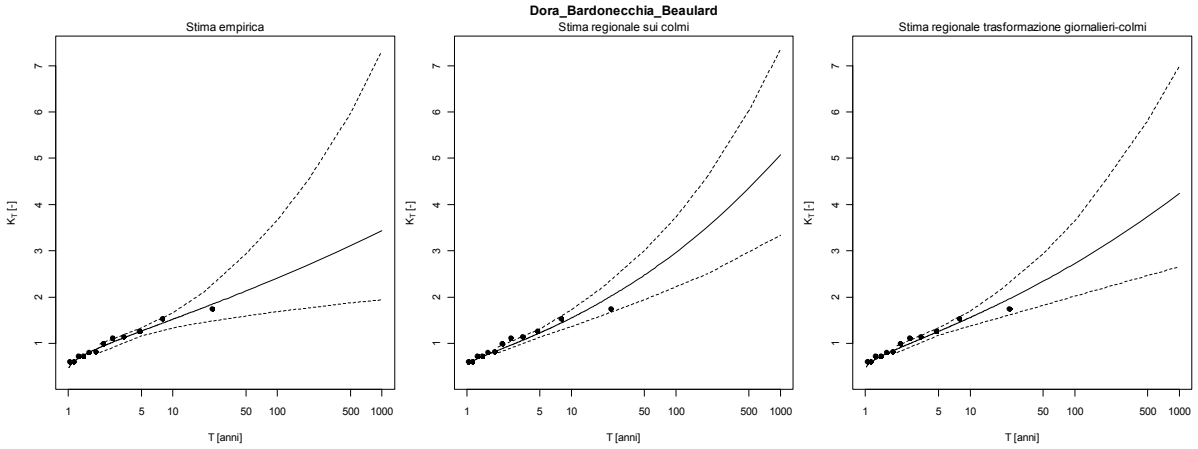


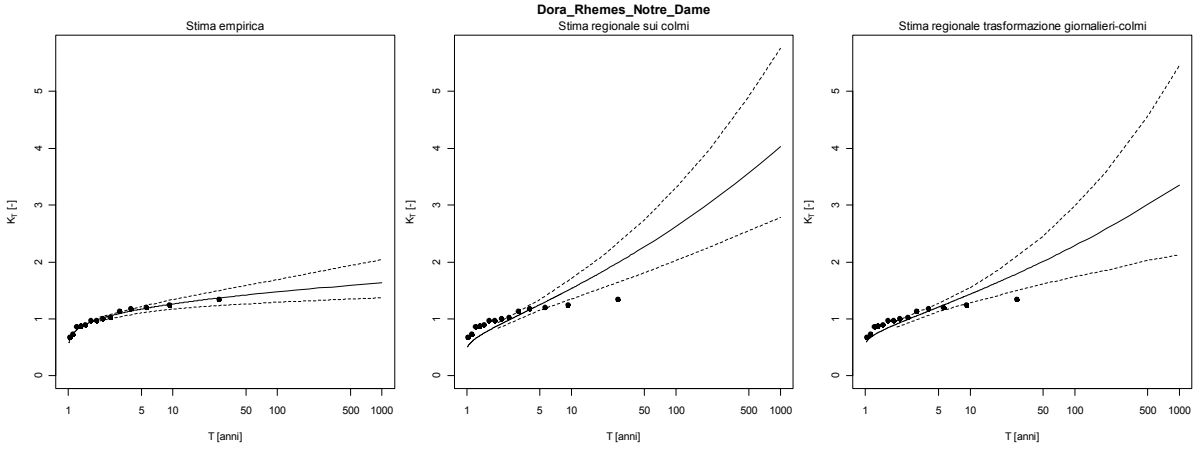

Corsaglia_C.le_Molline
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]

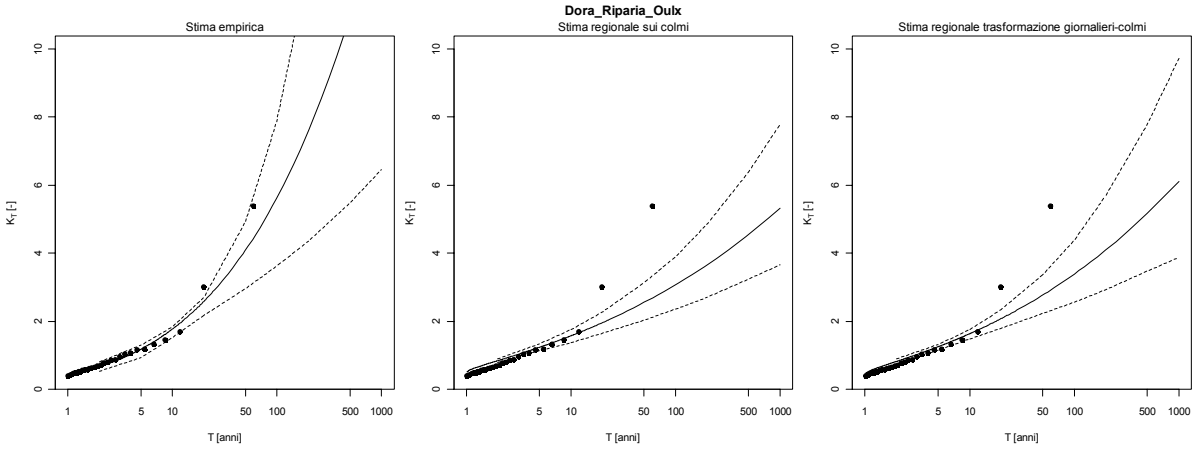

Dora_Baltea_Aosta
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]

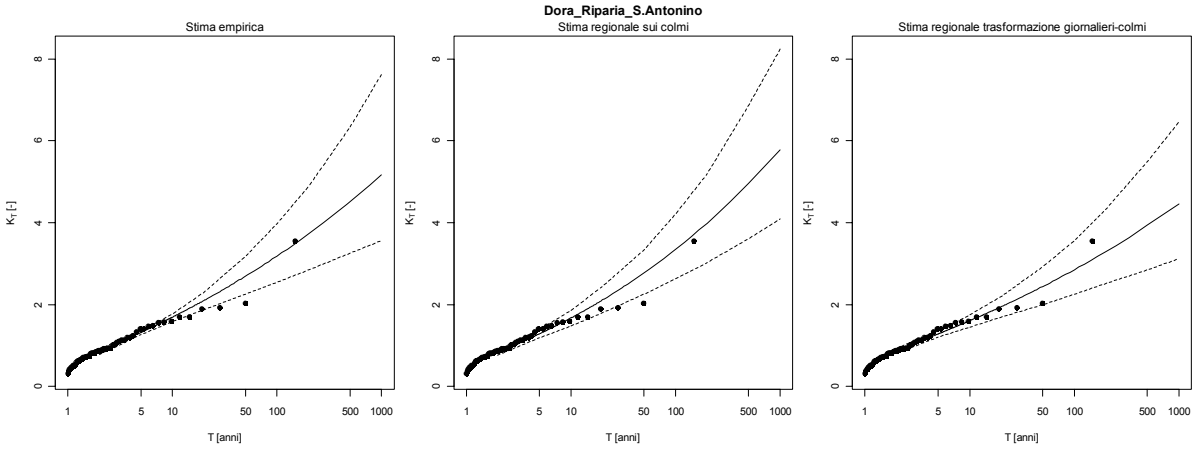

Dora_Baltea_P.te_Mombardone
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]

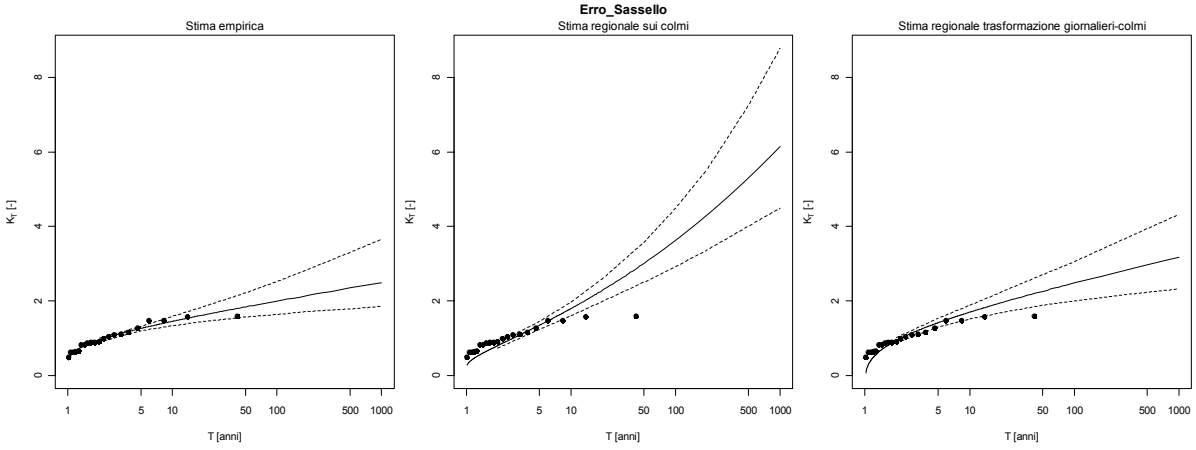

Dora_Baltea_Tavagnasco
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

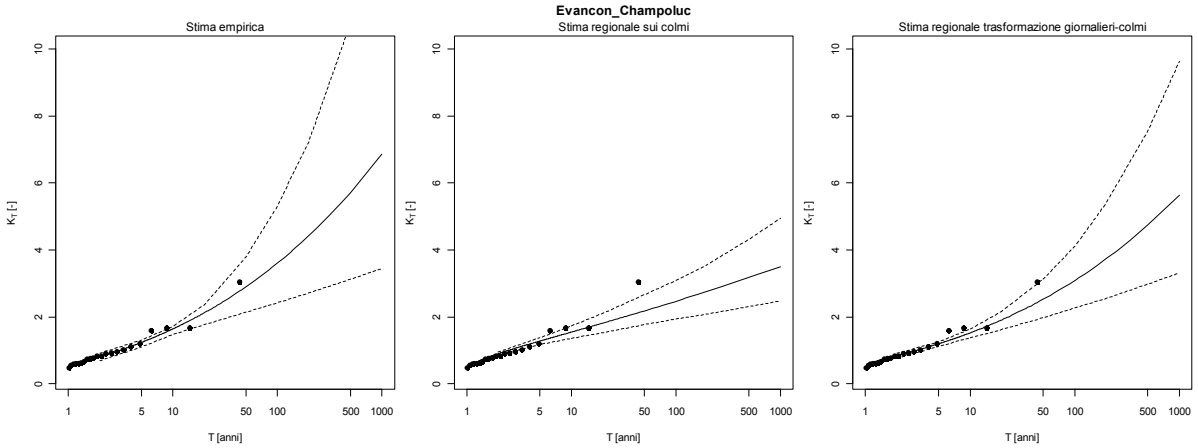

Dora_Bardonecchia_Beaulard
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

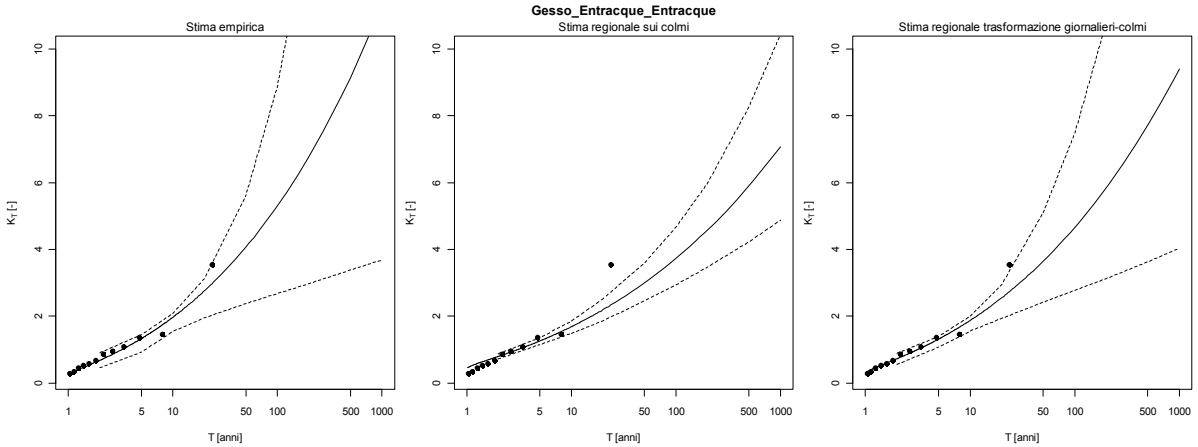

Dora_Rhemes_Notre_Dame
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

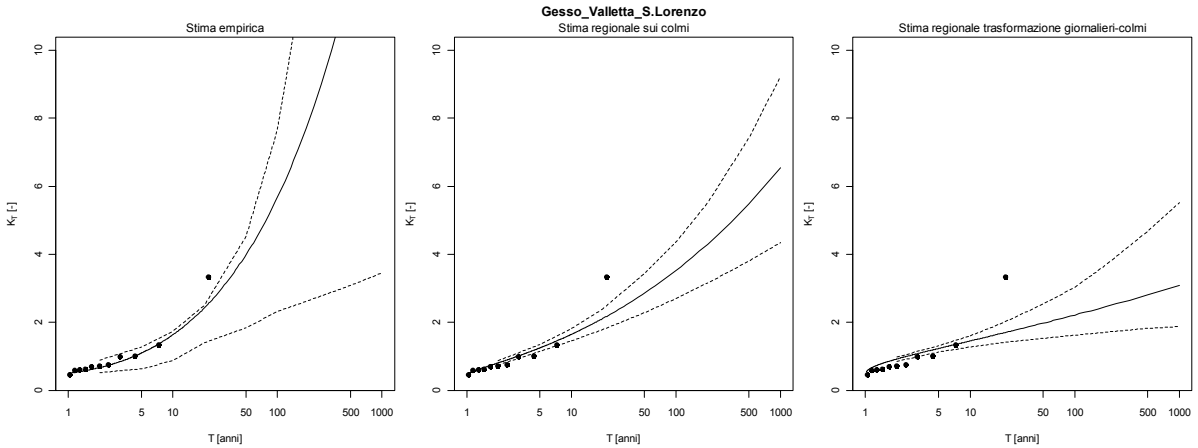

Dora_Riparia_Oulx
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10


Dora_Riparia_S.Antonino
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8

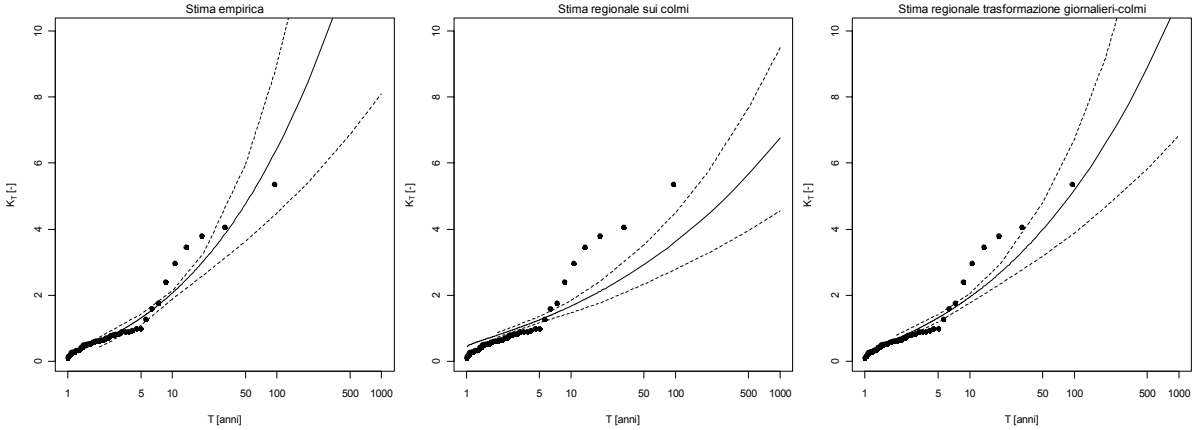

Erro_Sassello
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8


Evancon_Champoluc
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

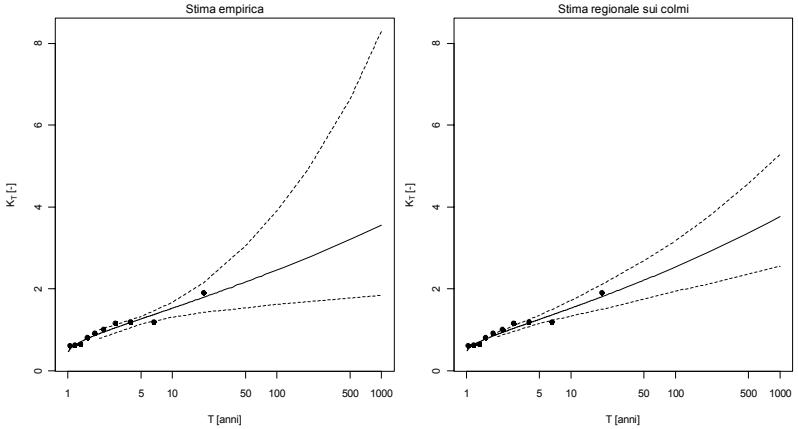

Gesso_Entracque_Entracque
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000


Gesso_Valletta_S.Lorenzo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

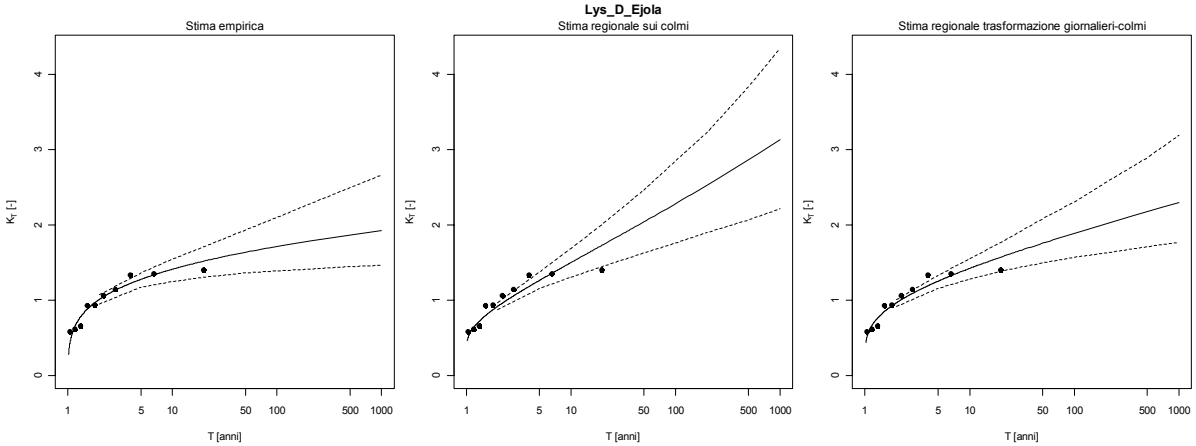
Grana_Monterosso



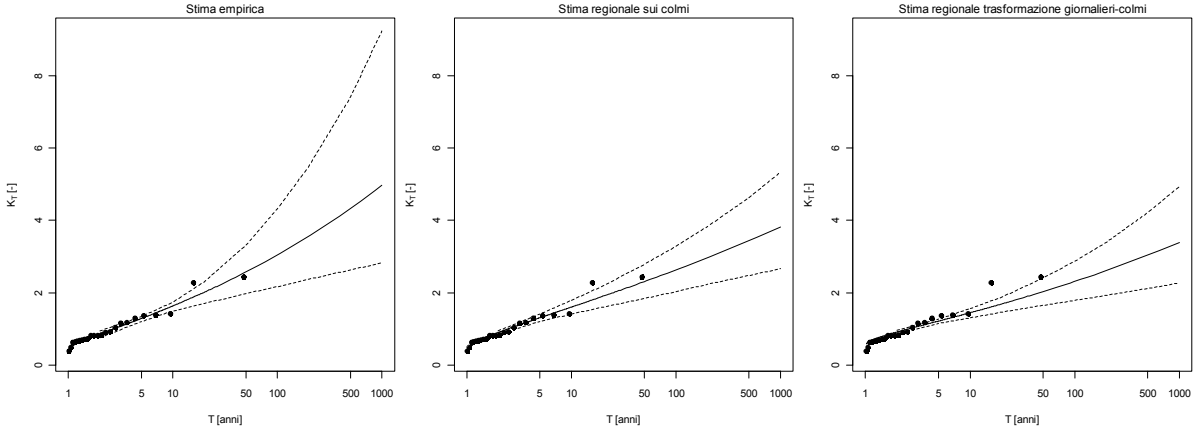
Grand_Eyvia_Cretaz



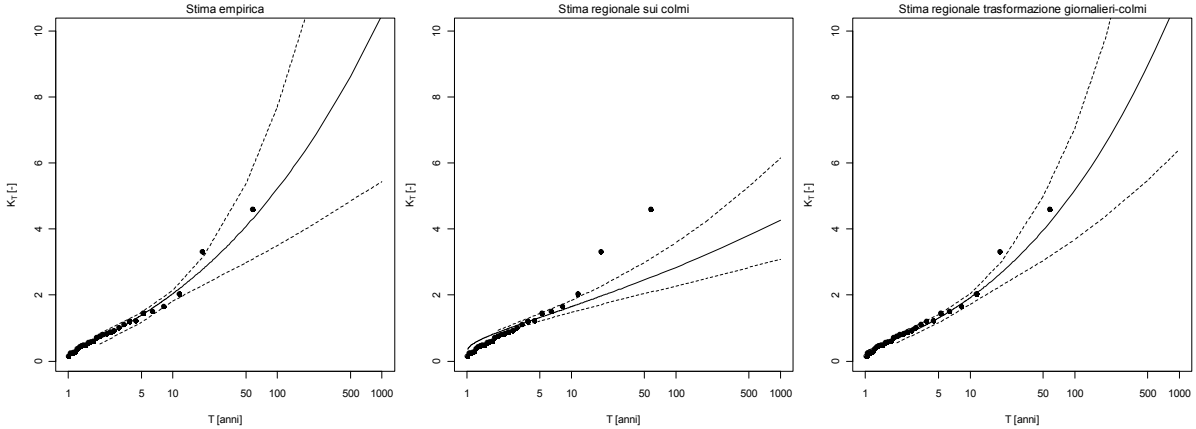
Lys_D_Ejola



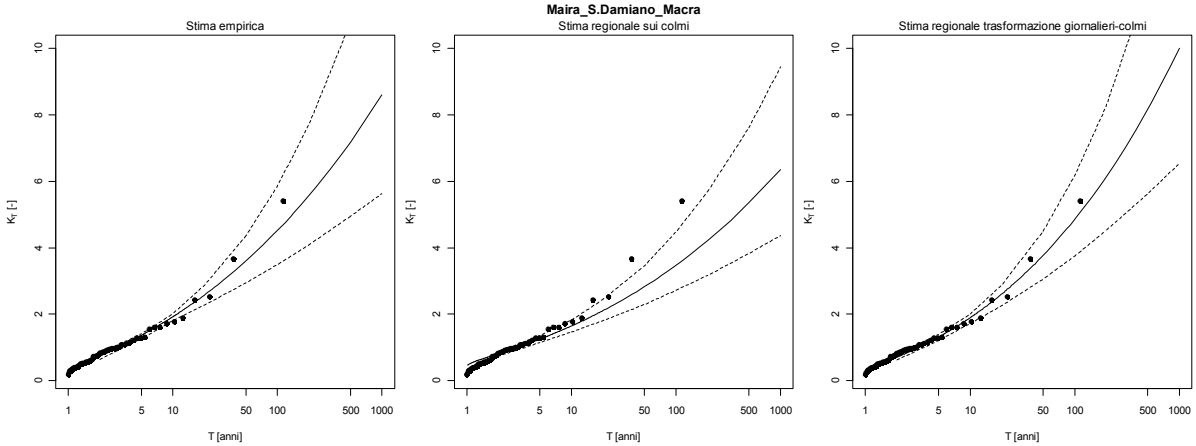
Lys_Gressoney

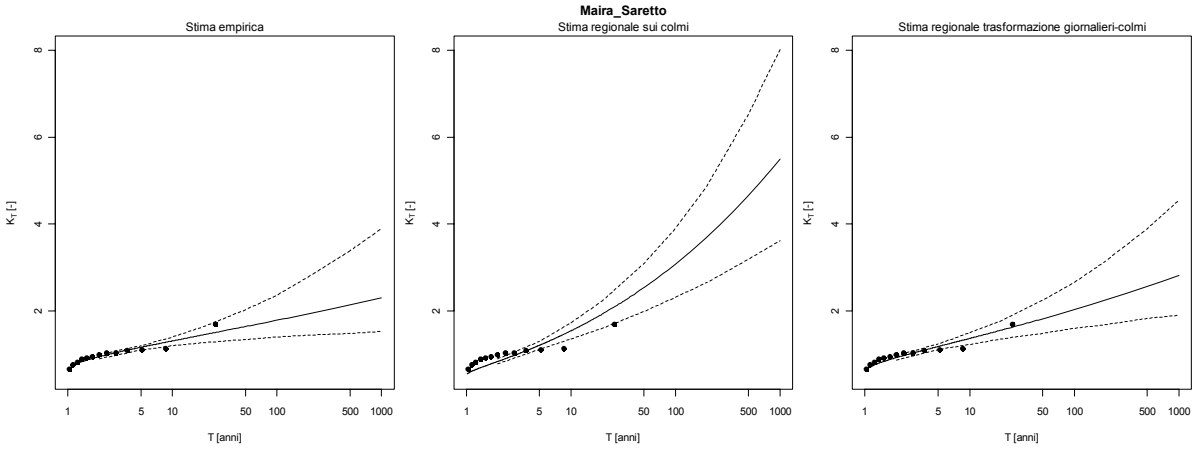


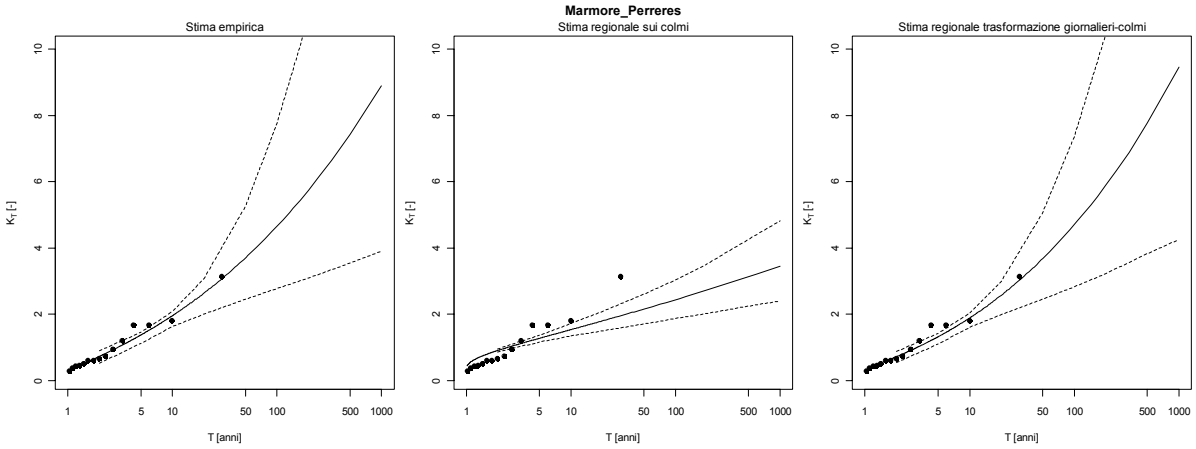
Lys_Guillemore

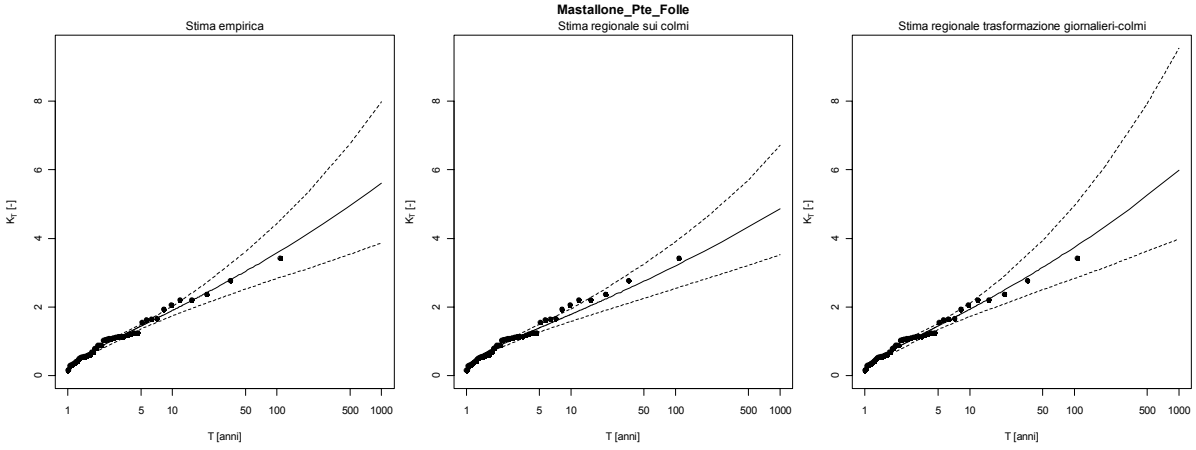


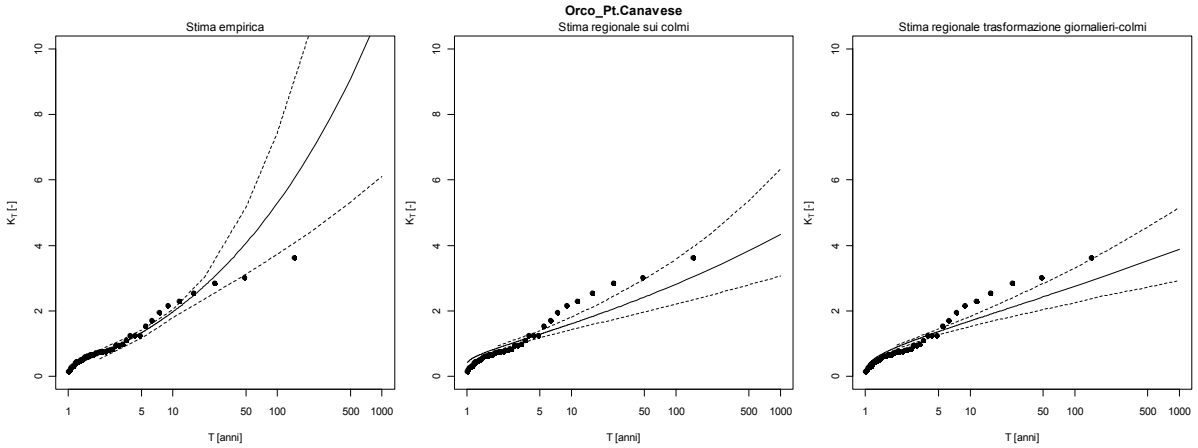
Maira_S.Damiano_Macra

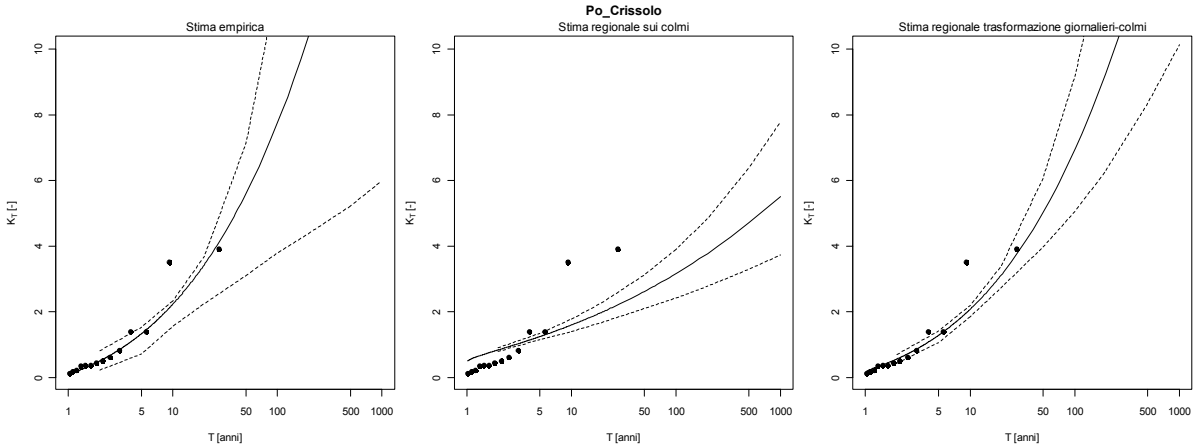


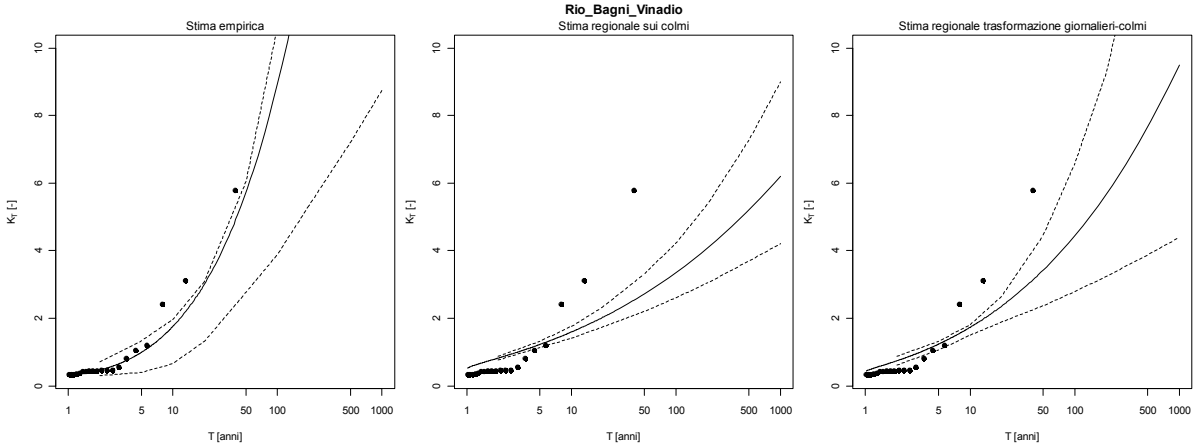

Maira_Saretto
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
2
4
6
8

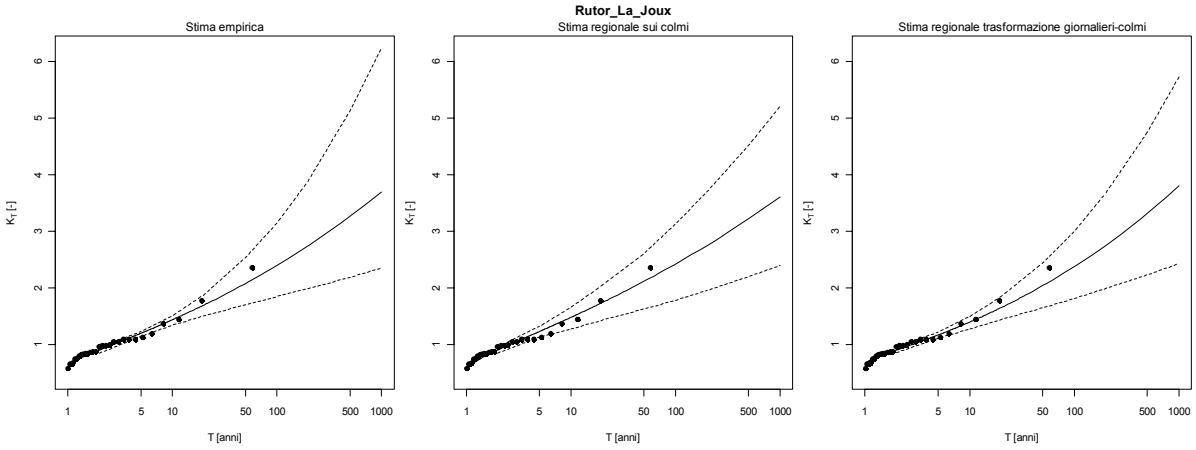

Marmore_Perreres
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

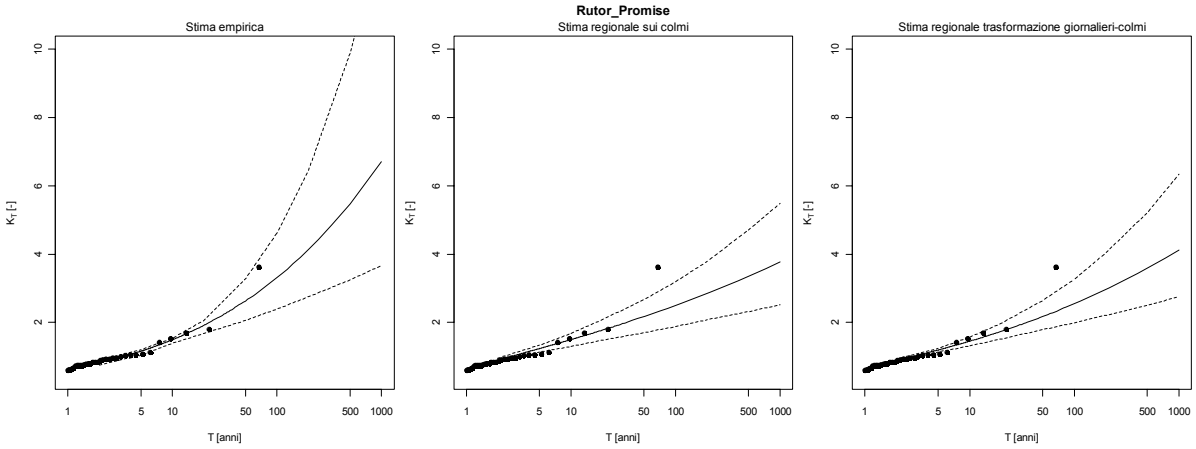

Mastallone_Pte_Folle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8

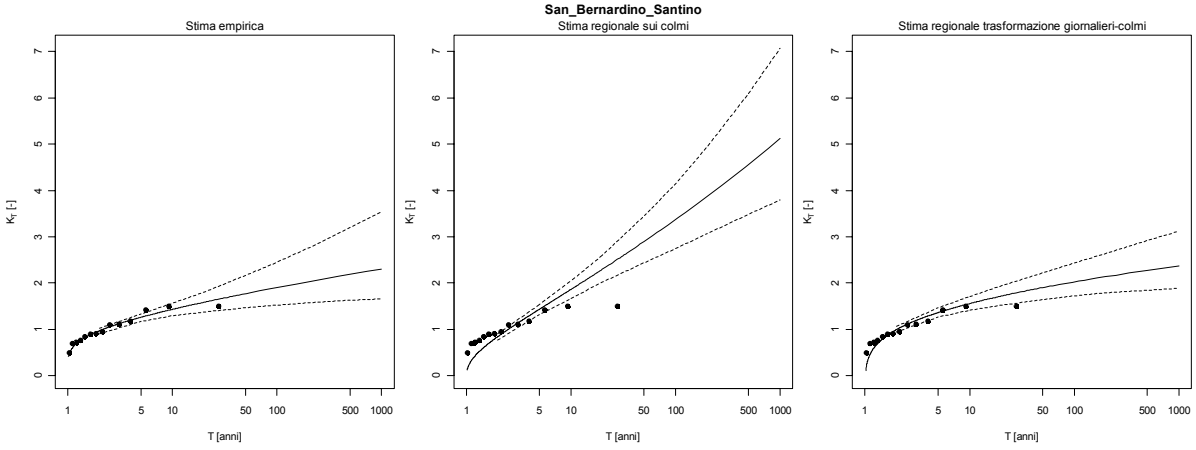

Orco_Pt.Canavese
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

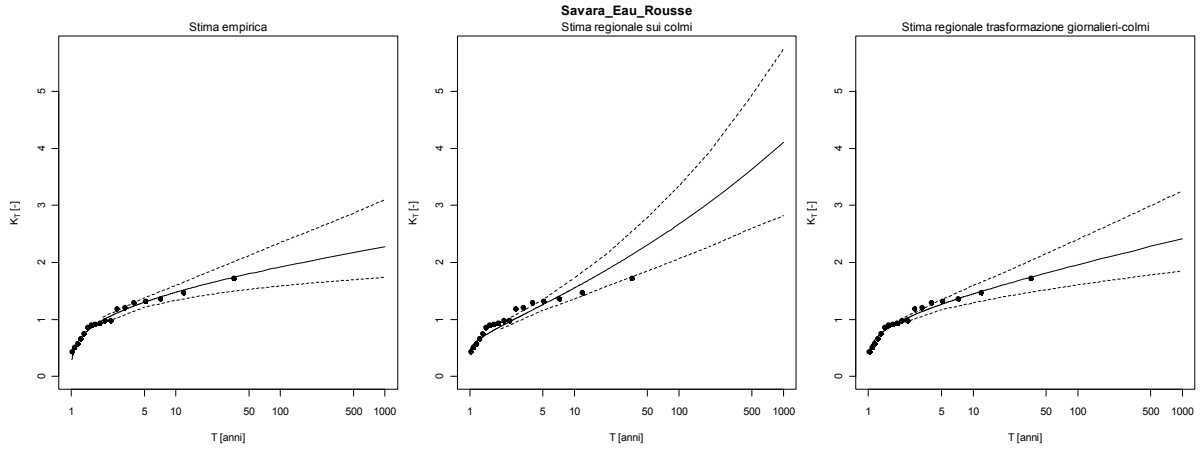

Po_Crissolo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

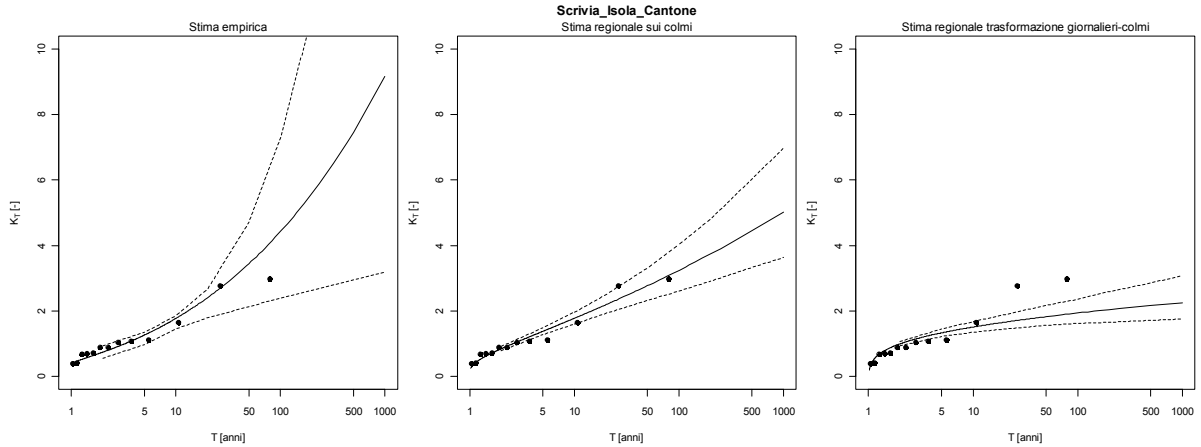

Rio_Bagni_Vinadio
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

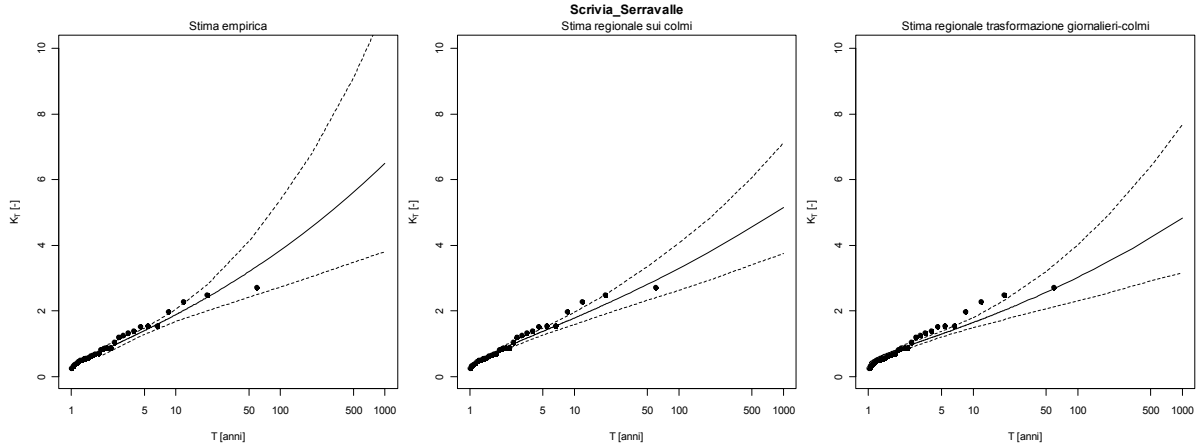

Rutor_La_Joux
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

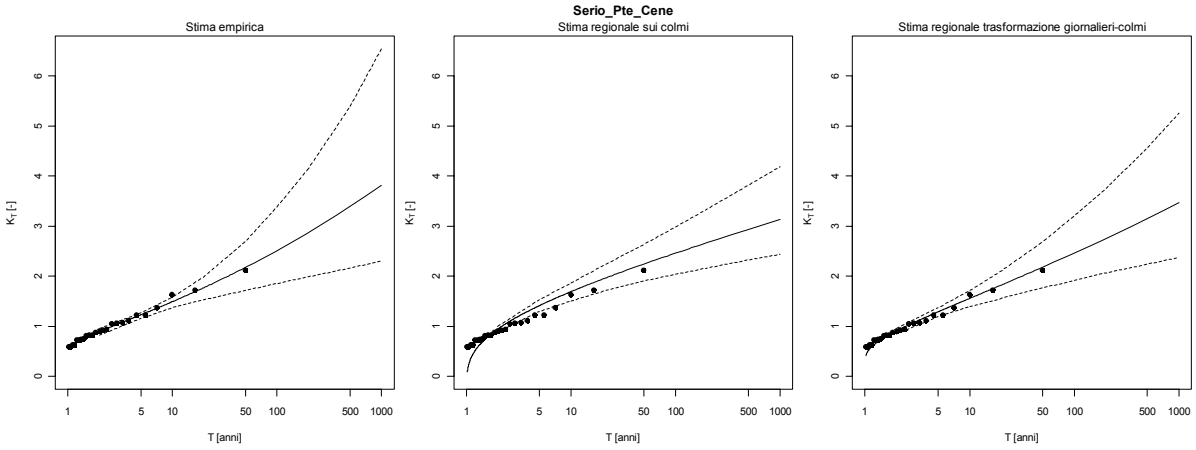

Rutor_Promise
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

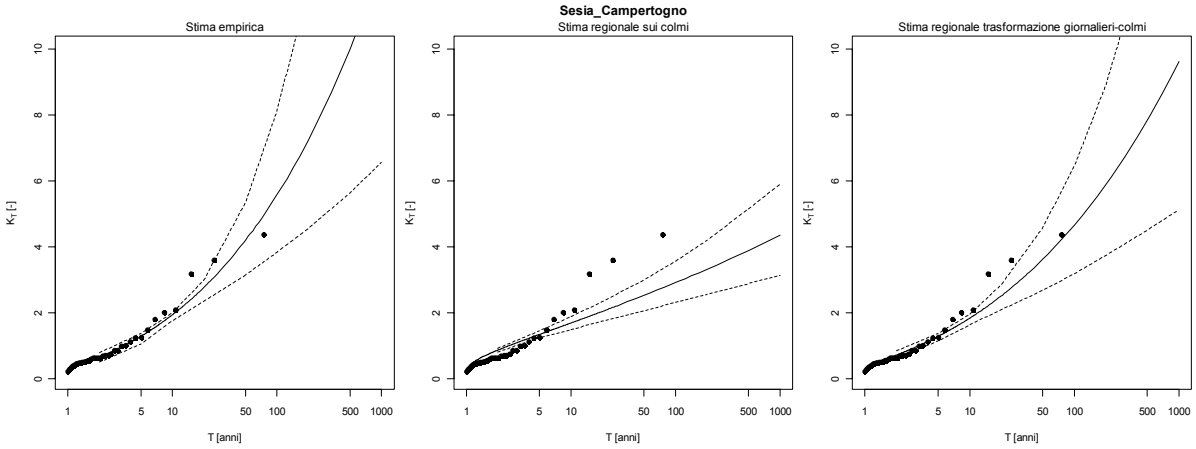

San_Bernardino_Santino
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

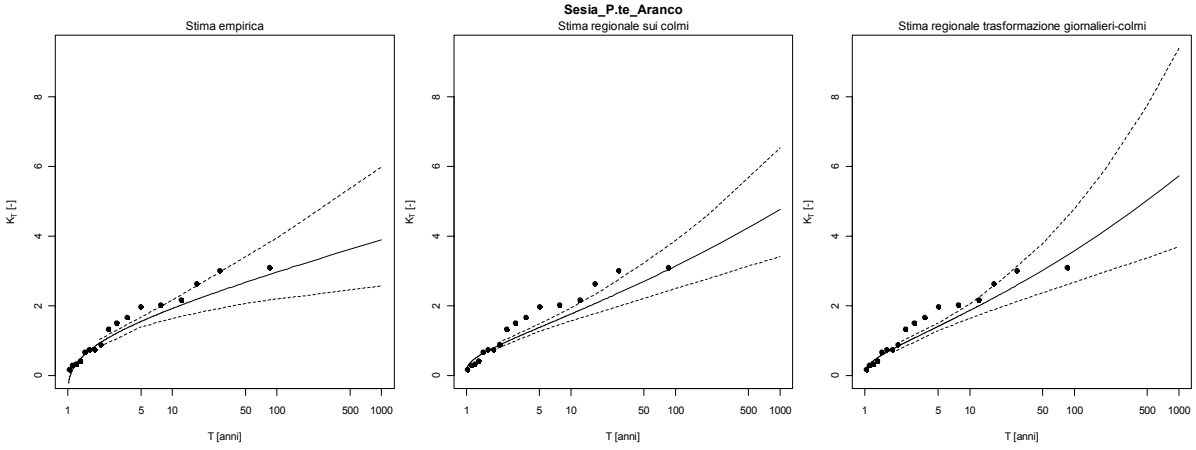

Savara_Eau_Rousse
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
1
2
3
4
5
1
5
10
50
100
500
1000

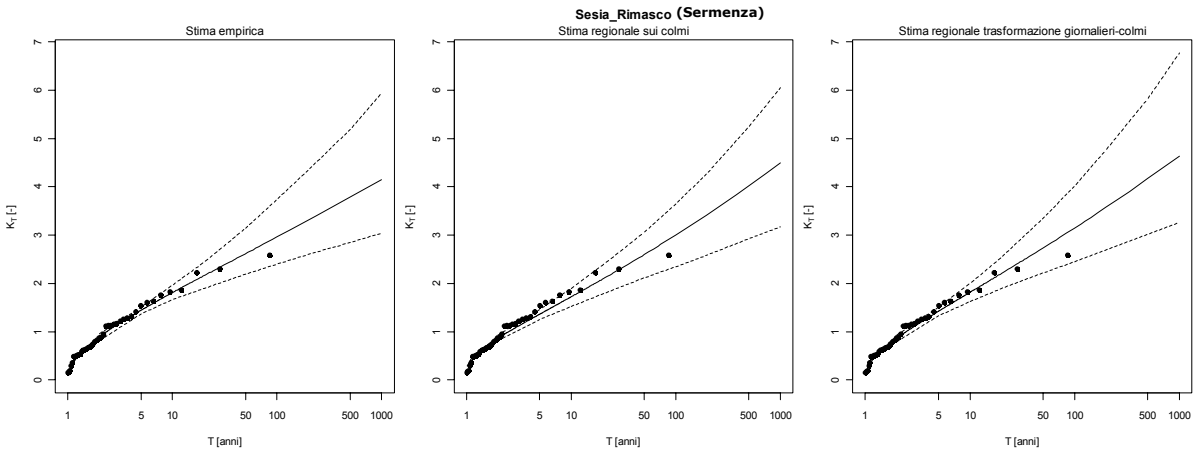

Scrivia_Isola_Cantone
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

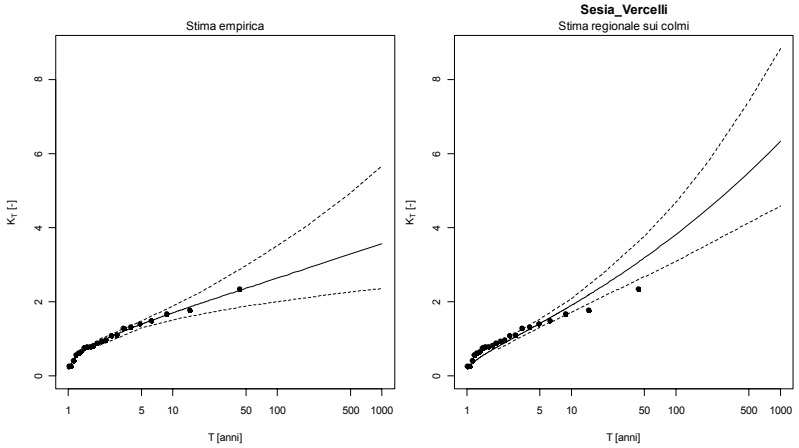

Scrivia_Serravalle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

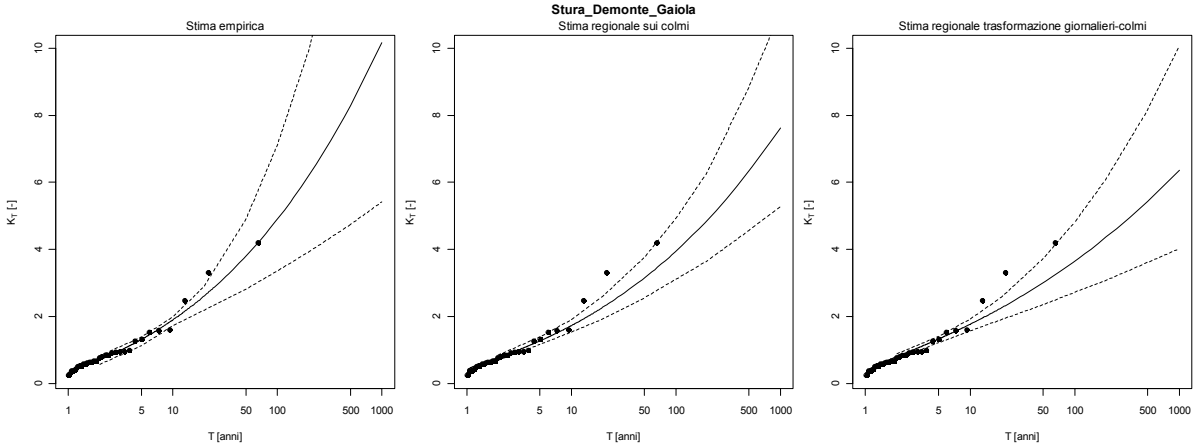

Serio_Pte_Cene
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

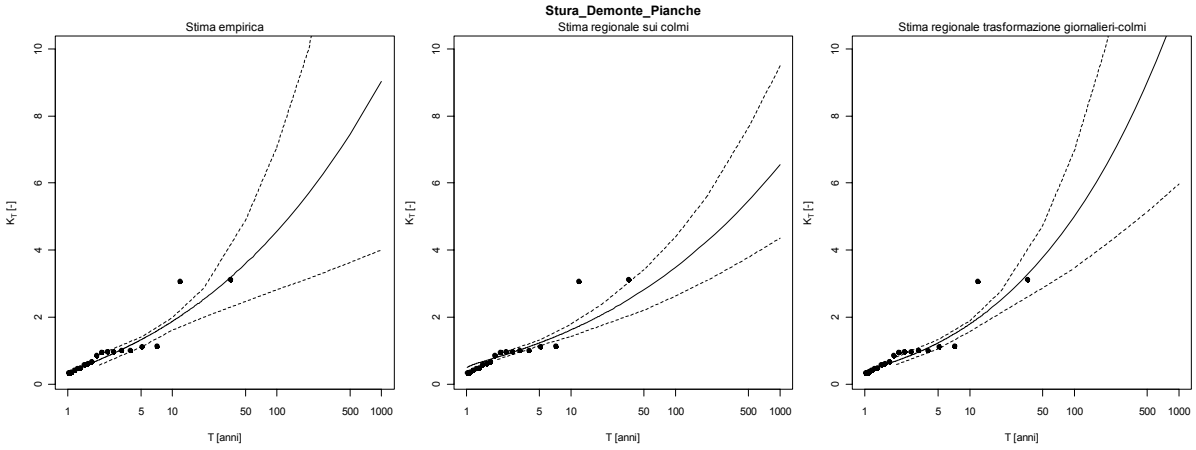

Sesia_Campertogno
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

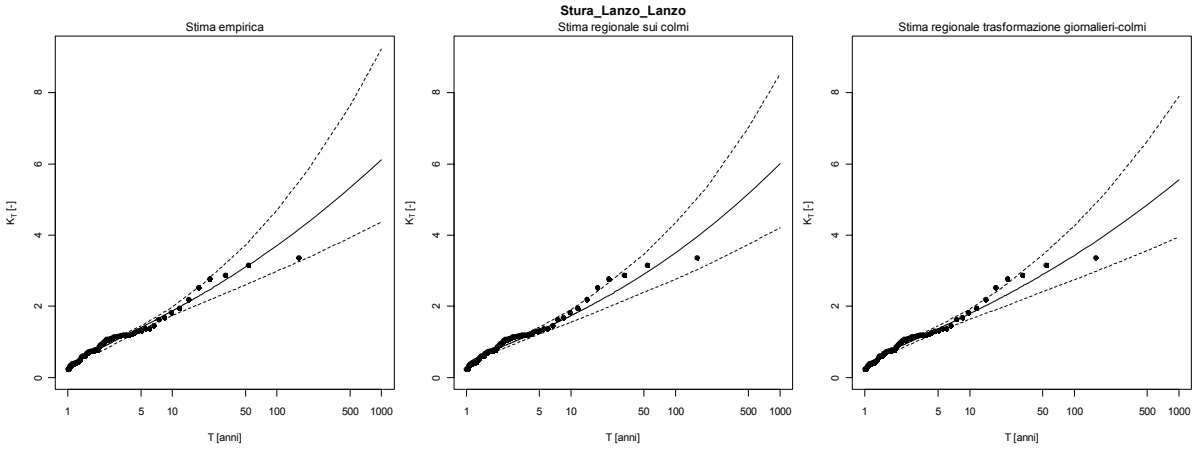

Sesia_P.te_Aranco
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

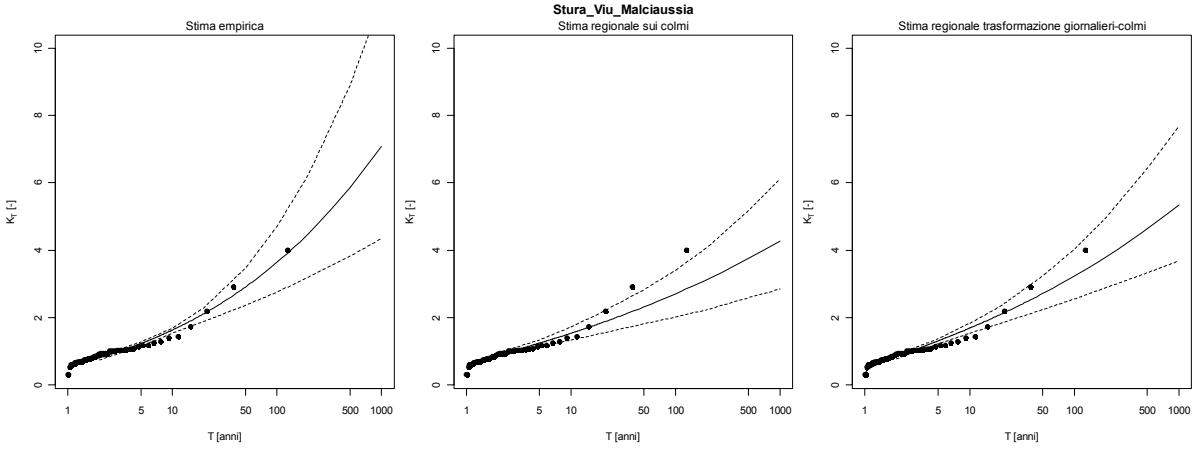

Sesia_Rimasco (Sermenza)
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

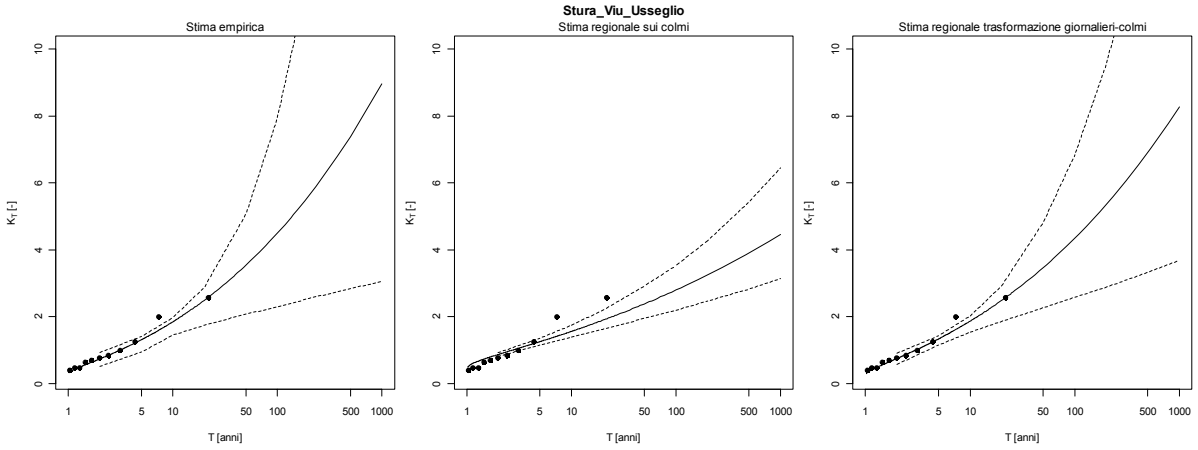

Sesia_Vercelli
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

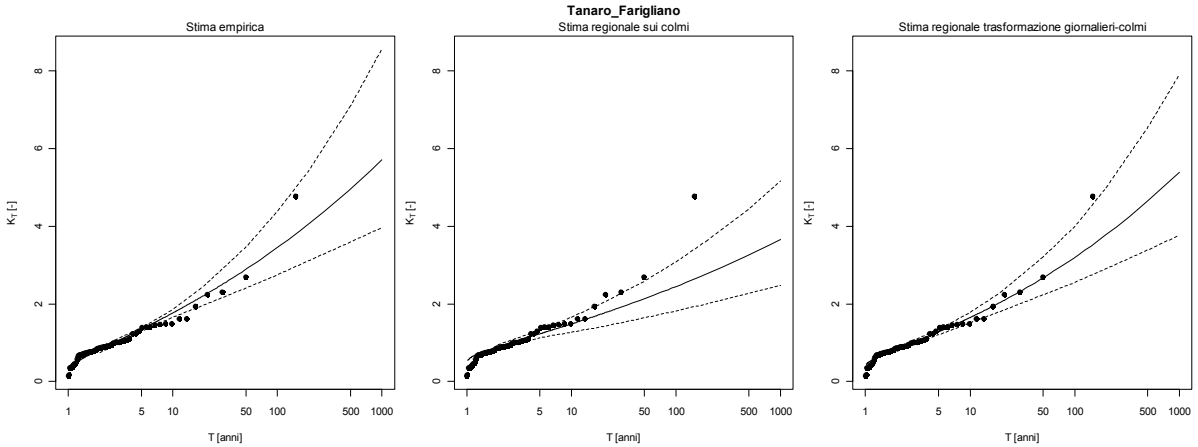

Stura_Demonte_Gaiola
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

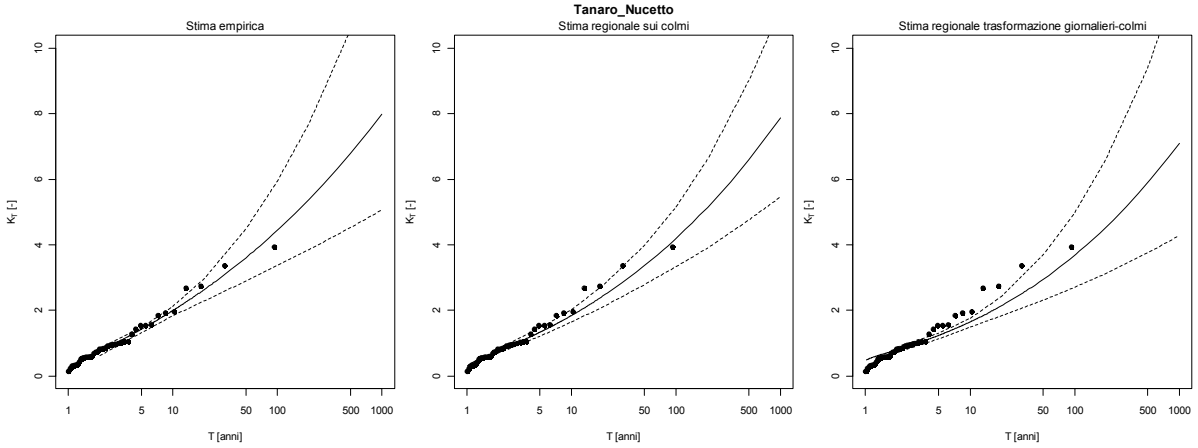

Stura_Demonte_Pianche
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

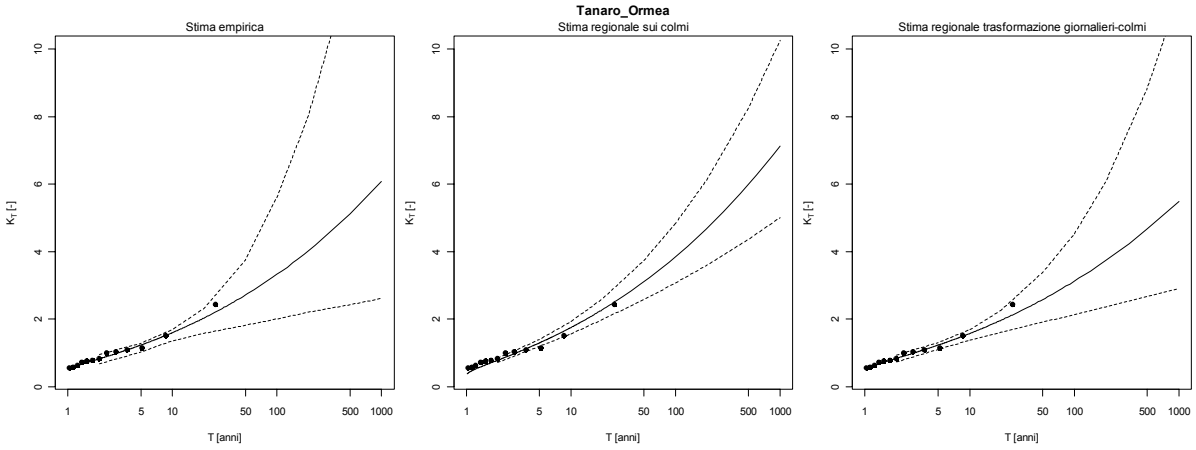

Stura_Lanzo_Lanzo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8

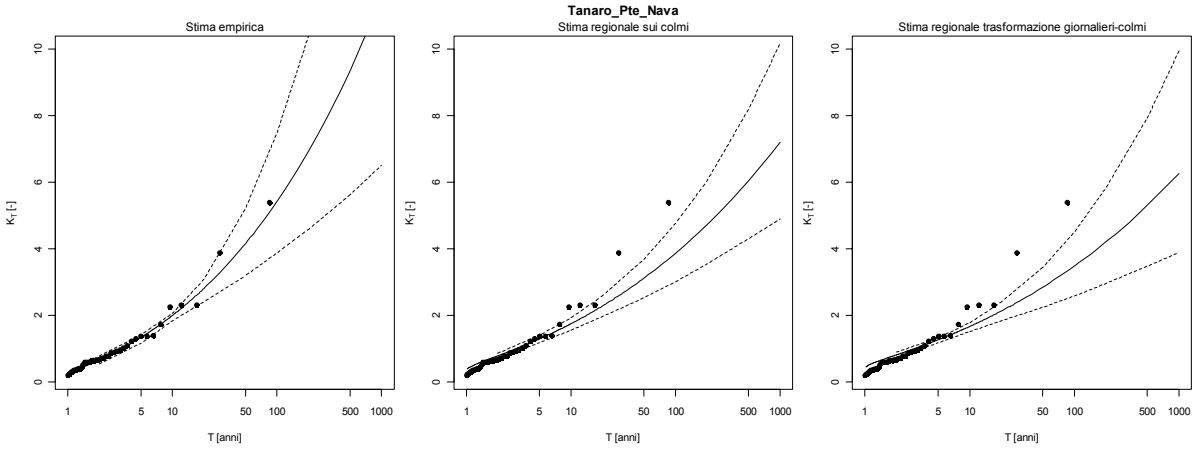

Stura_Viu_Malciaussia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

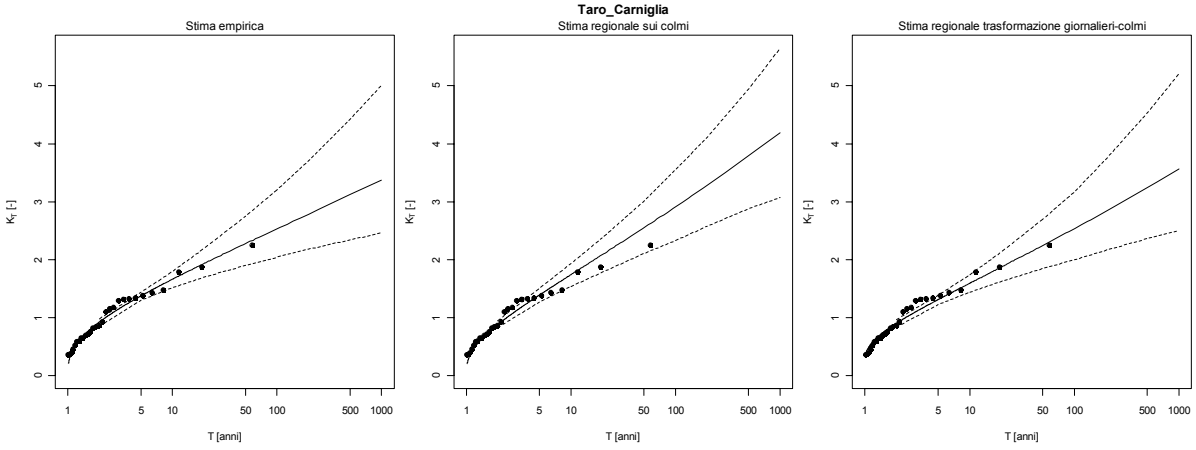

Stura_Viu_Usseglio
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

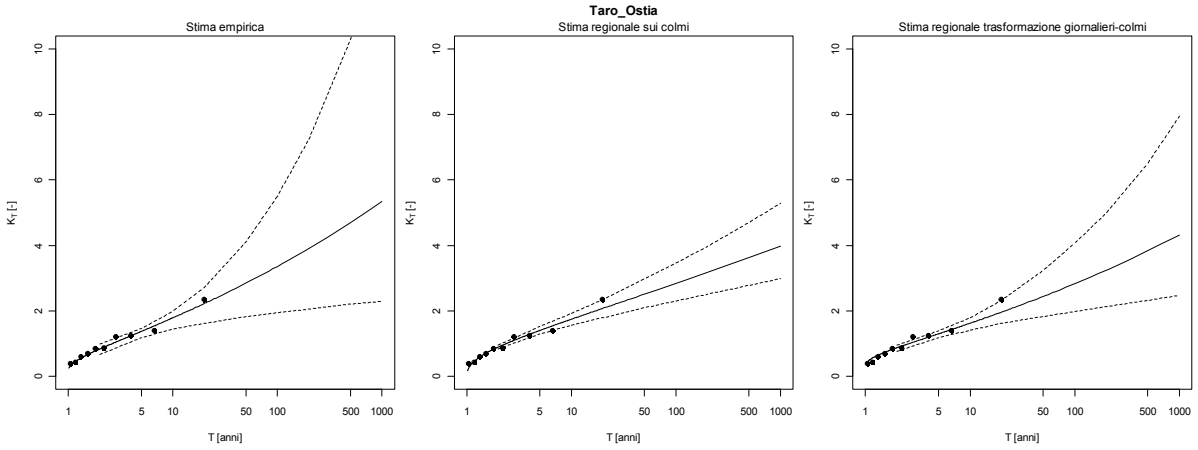

Tanaro_Farigliano
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
1
5
10
50
100
500
1000

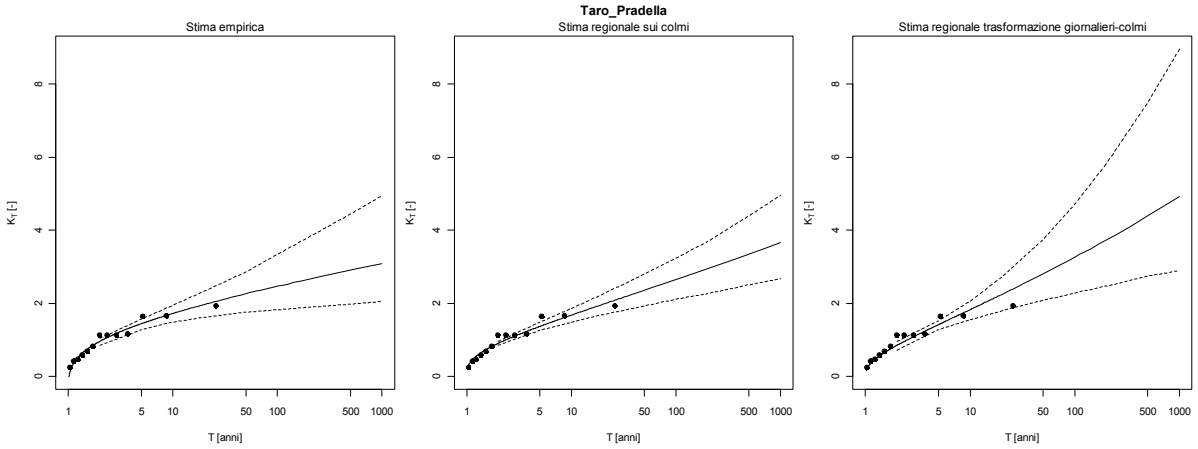

Tanaro_Nucetto
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

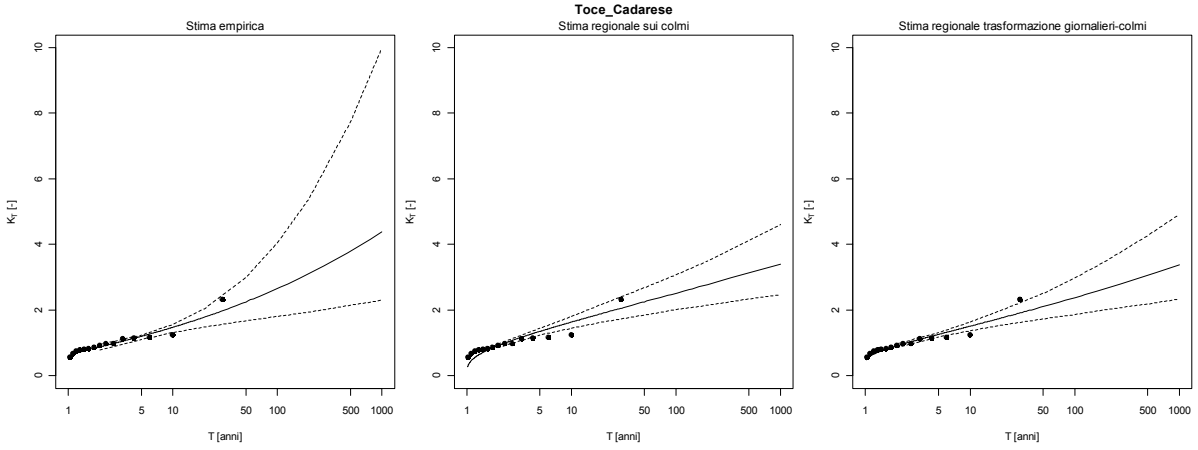

Tanaro_Ormea
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

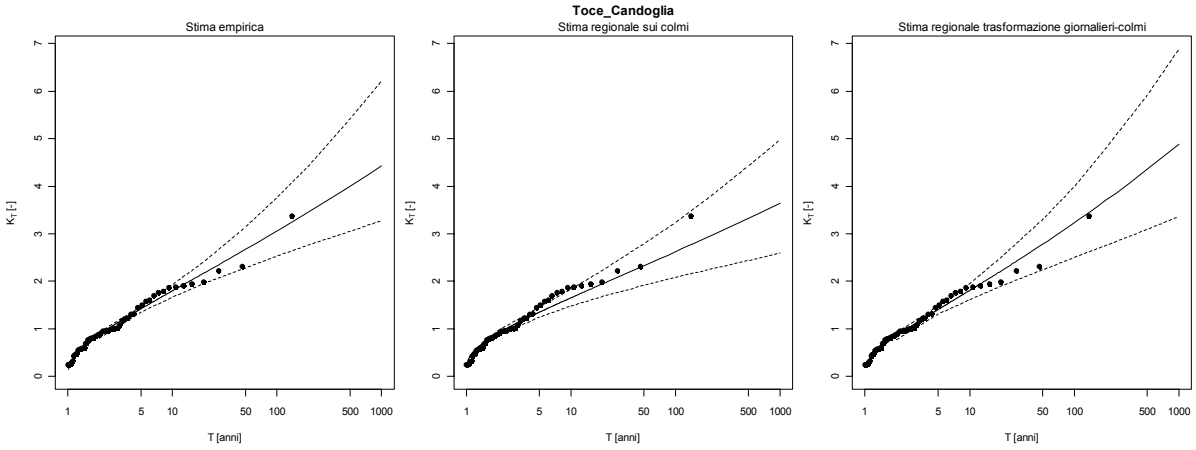

Tanaro_Pte_Nava
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

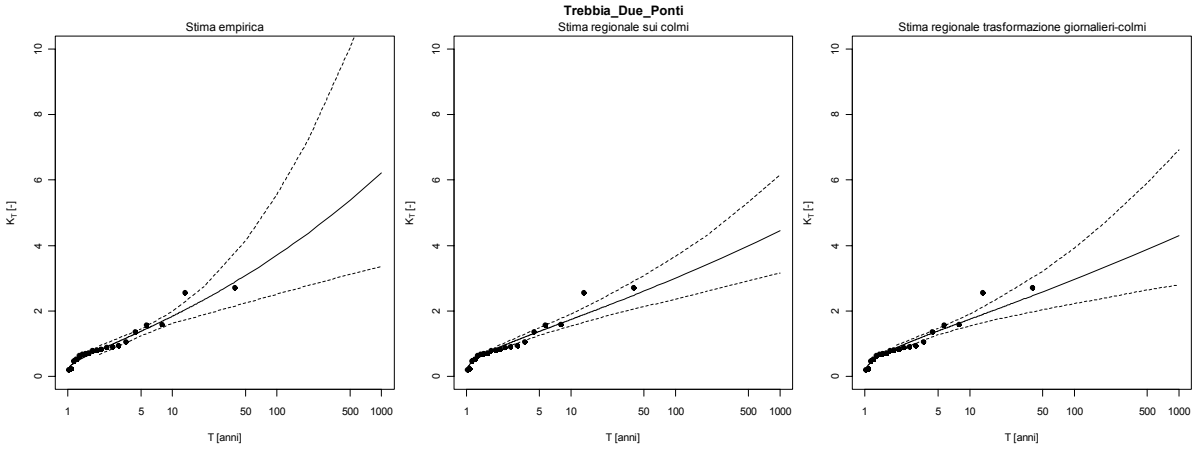

Taro_Carniglia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
1
2
3
4
5

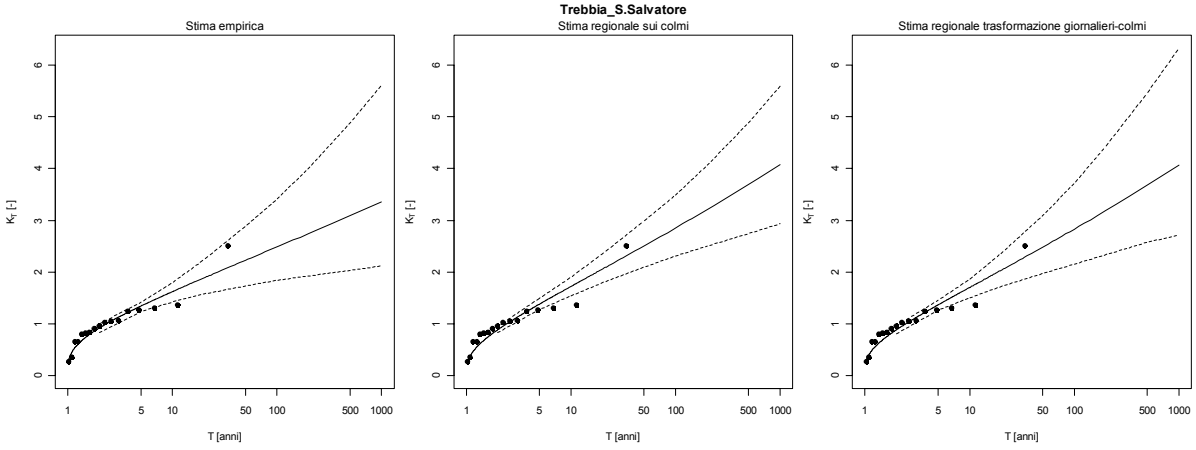

Taro_Ostia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

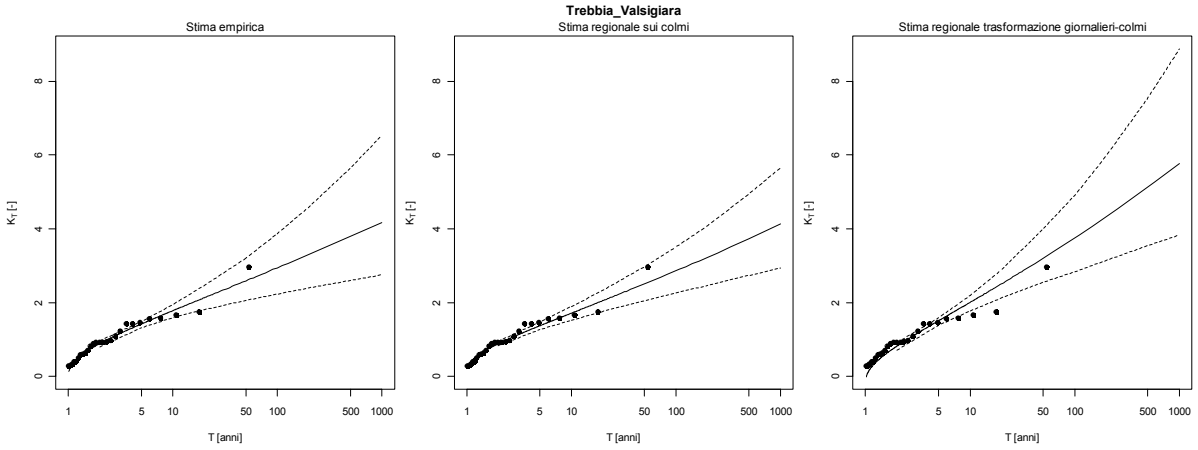

Taro_Pradella
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
1
5
10
50
100
500
1000

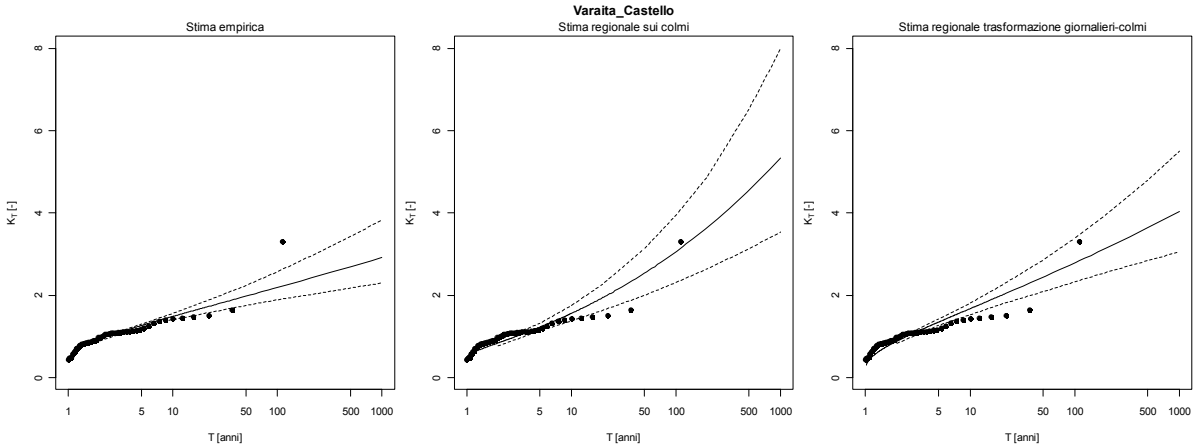

Toce_Cadarese
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

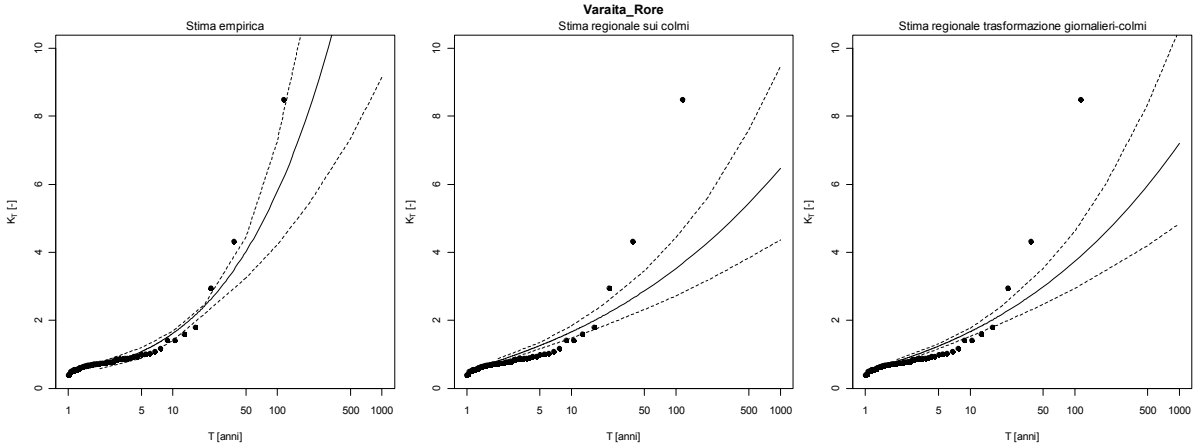

Toce_Candoglia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000


Trebbia_Due_Ponti
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

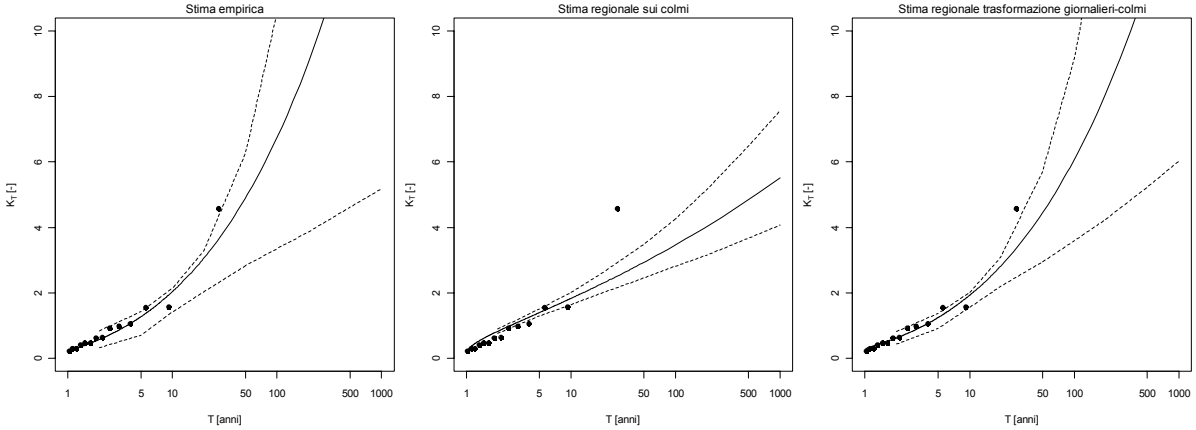

Trebbia_S.Salvatore
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000


Trebbia_Valsigiara
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
1
5
10
50
100
500
1000

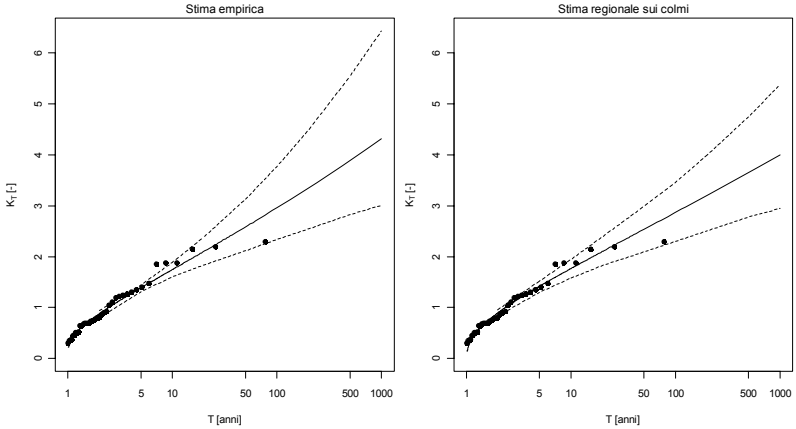

Varaita_Castello
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
1
5
10
50
100
500
1000


Varaita_Rore
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0
2
4
6
8
10
1
5
10
50
100
500
1000

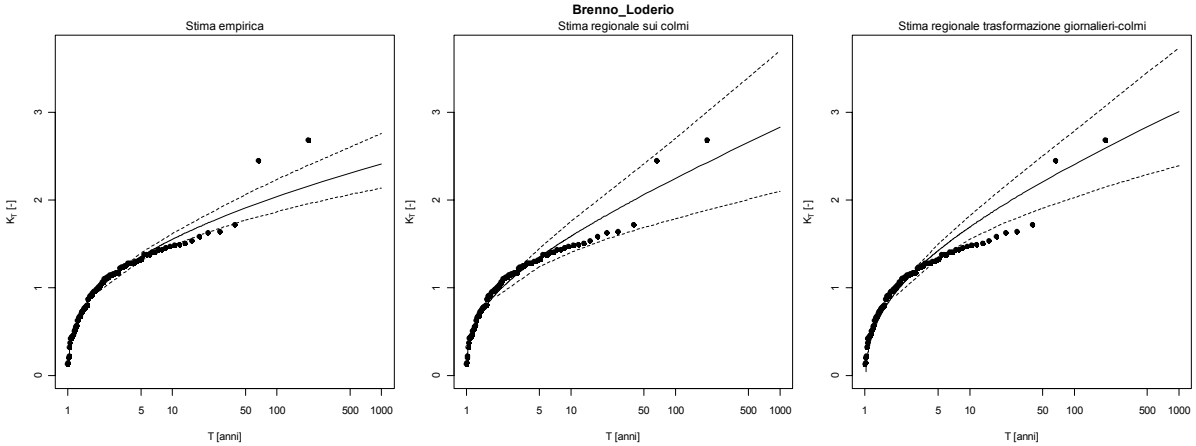
Vobbia_Vobbietta

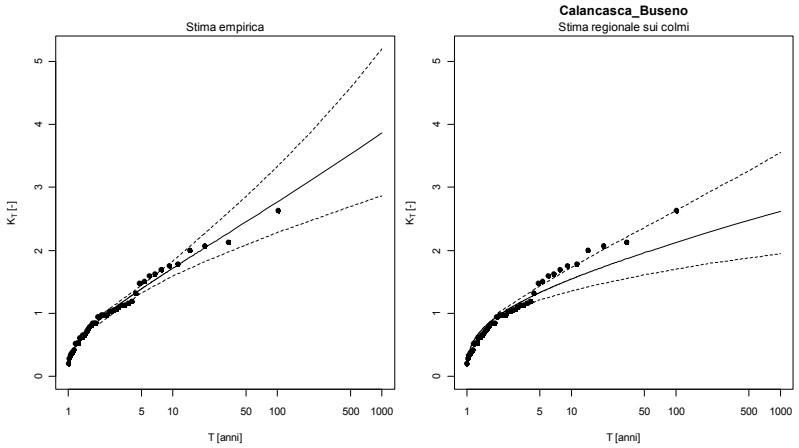


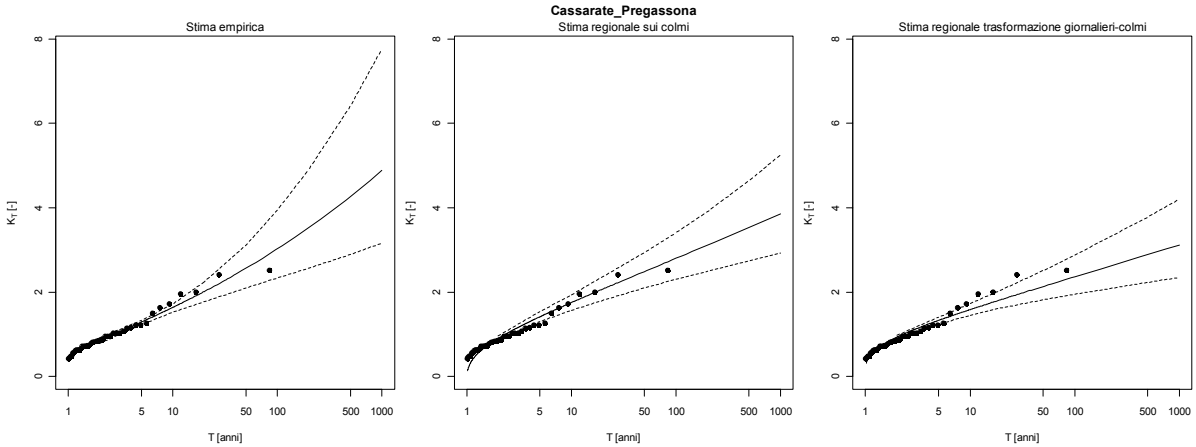
Breggia_Chiasso

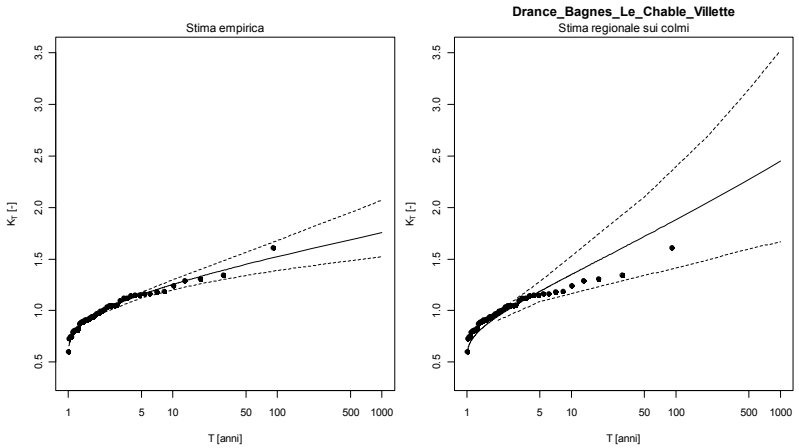


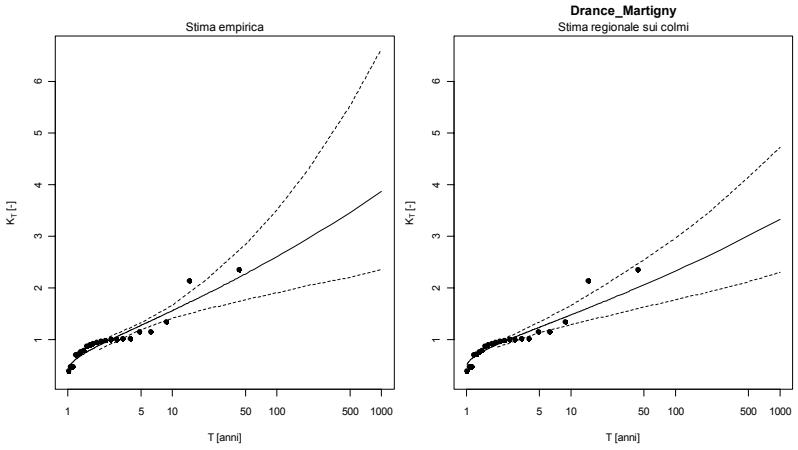
Brenno_Loderio

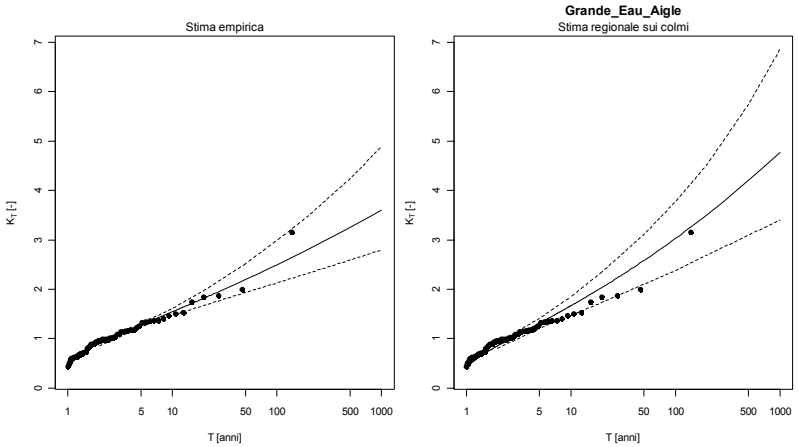


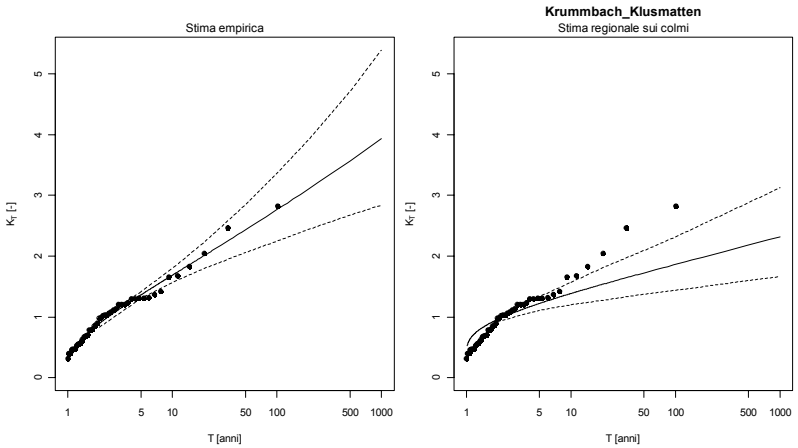

Calancasca_Buseno
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
1
2
3
4
5


Cassarate_Pregassona
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8

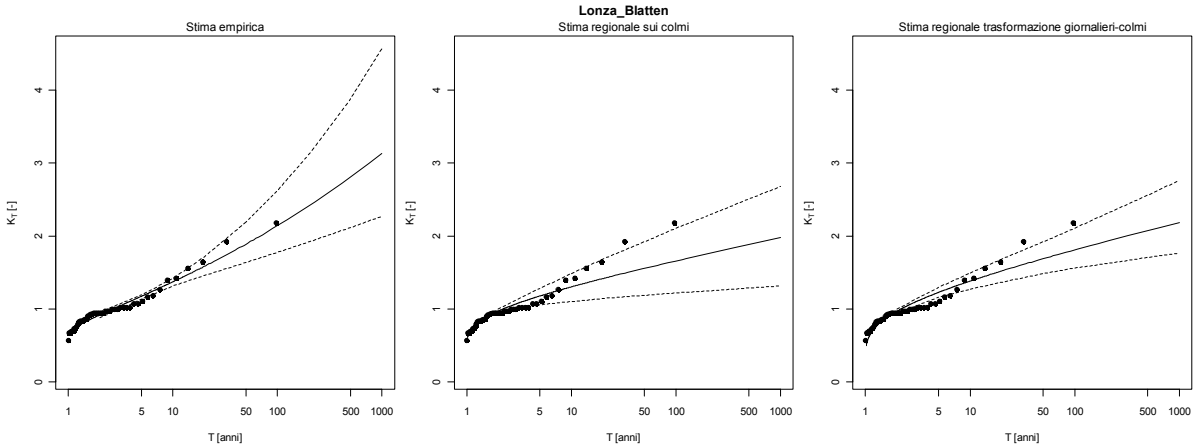

Drance_Bagnes_Le_Chable_Villette
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0.5
1.0
1.5
2.0
2.5
3.0
3.5


Drance_Martigny
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

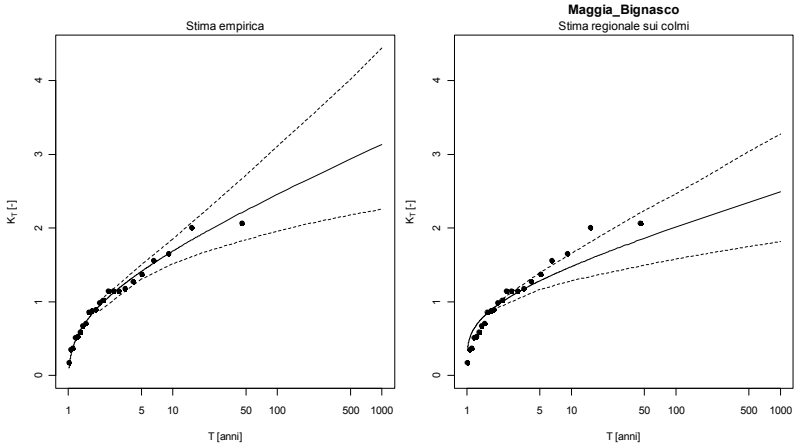

Grande_Eau_Aigle
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000


Krummbach_Klusmatten
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

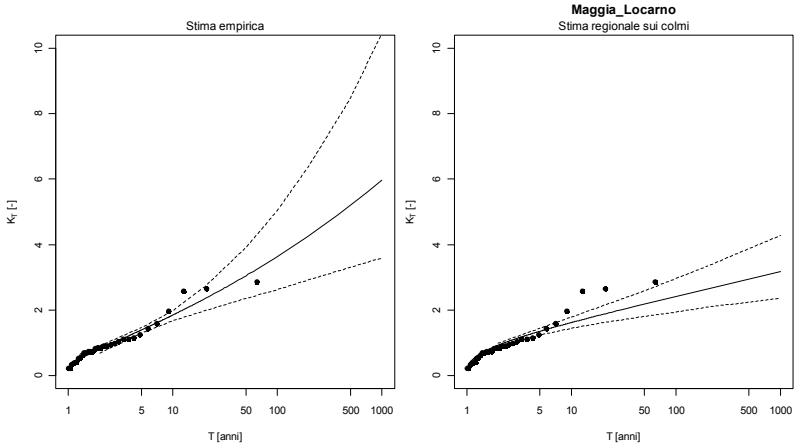
Lonza_Blatten

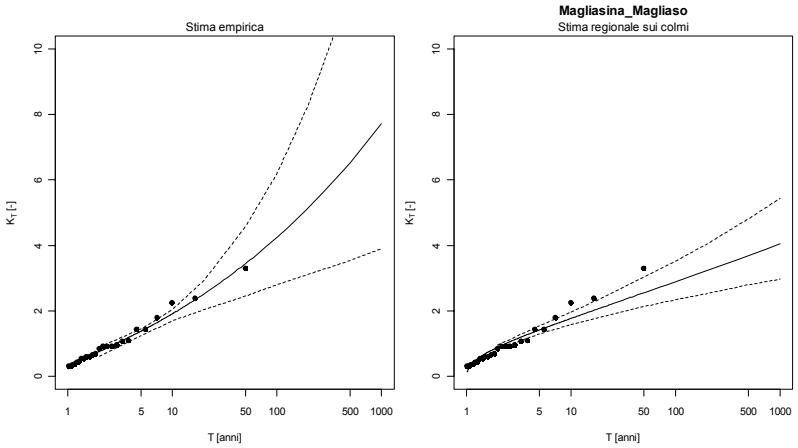


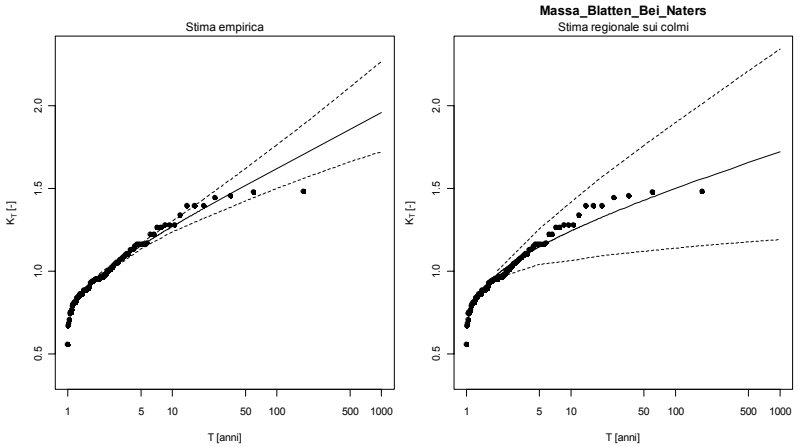
Maggia_Bignasco

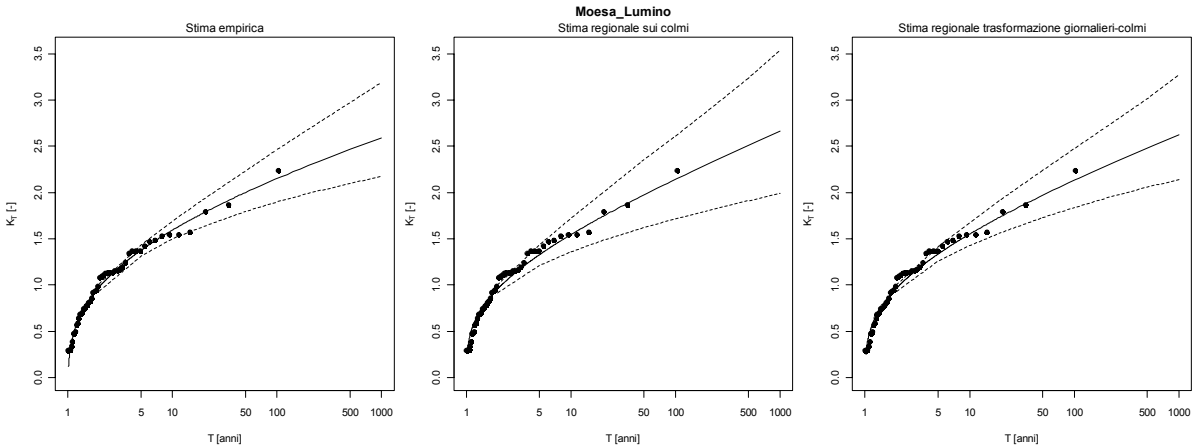


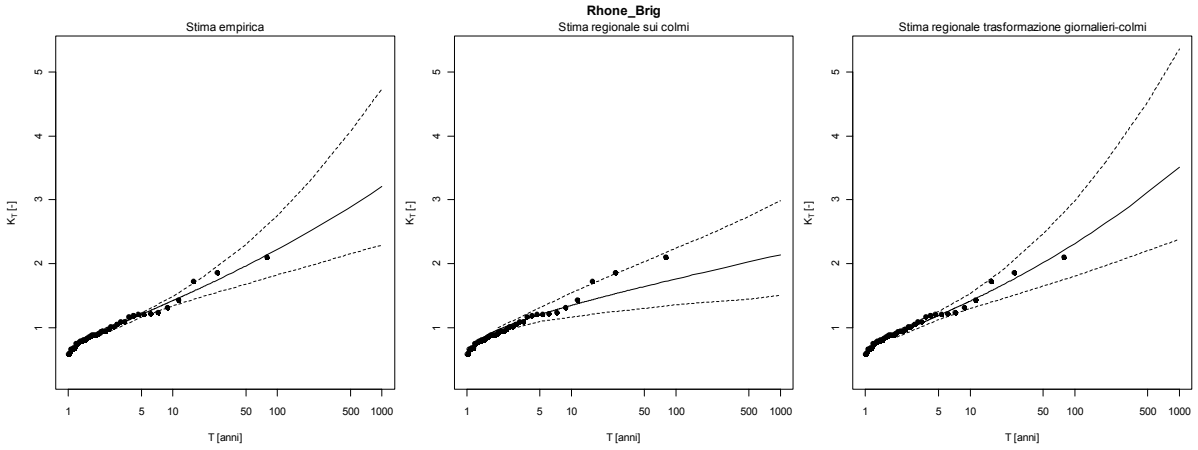
Maggia_Locarno

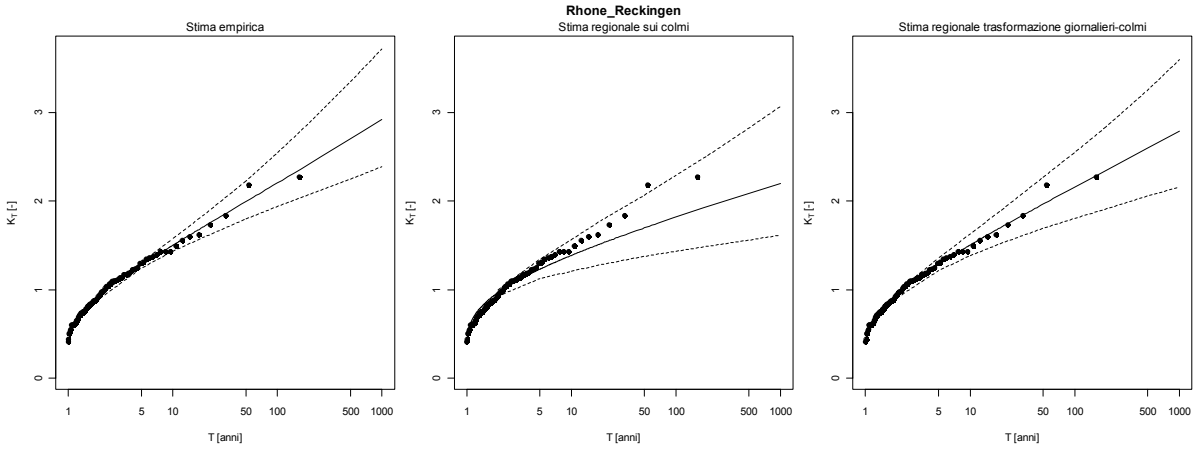


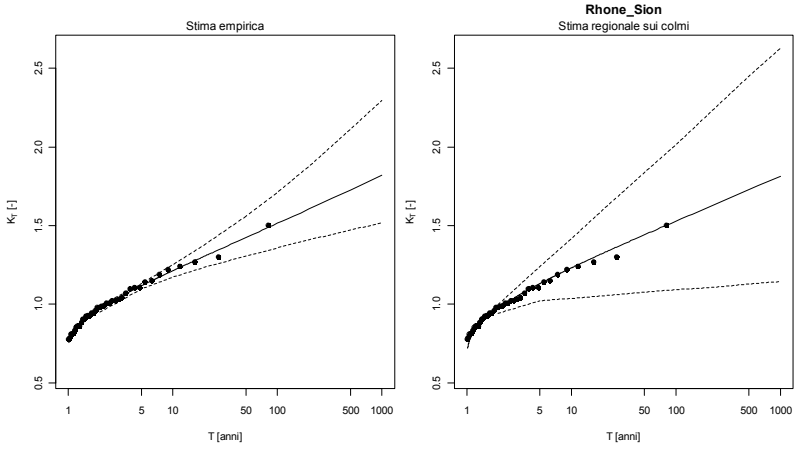

Magliasina_Magliaso
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10


Massa_Blatten_Bei_Naters
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
0.5
1.0
1.5
2.0

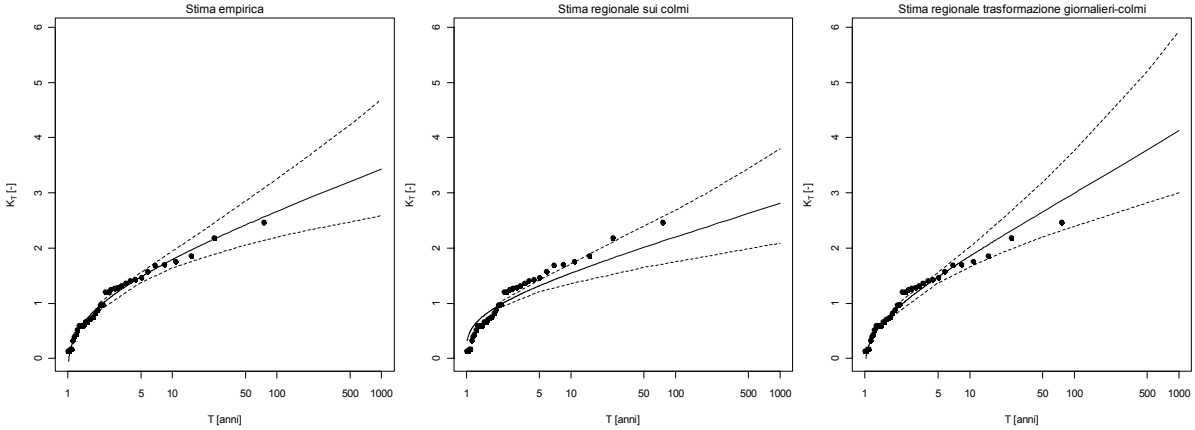

Moesa_Lumino
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5


Rhone_Brig
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

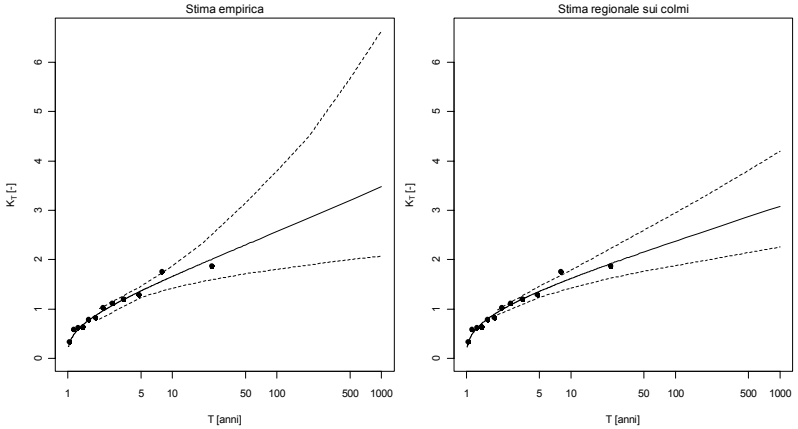

Rhone_Reckingen
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000


Rhone_Sion
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000

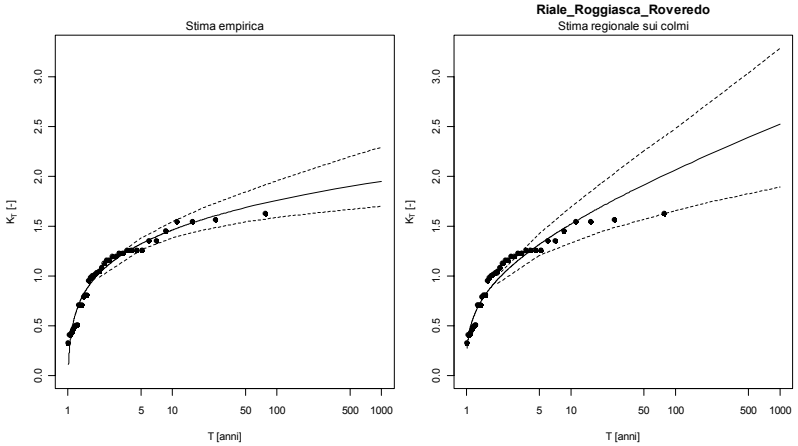
Riale_Calneggia_Cavergno

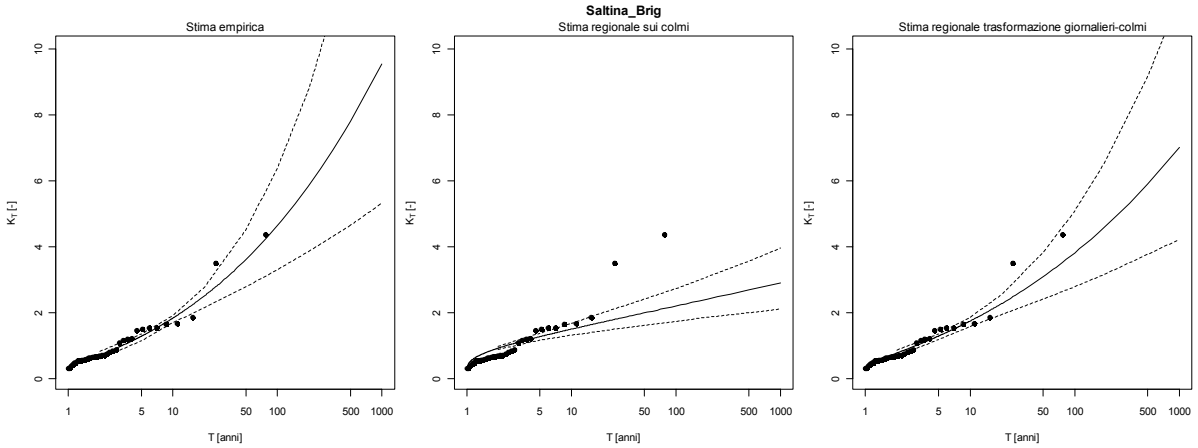


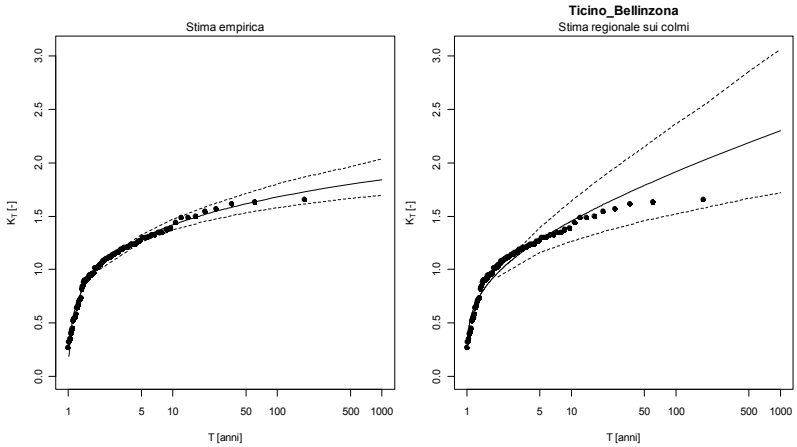
Riale_Pincascia_Lavertezzo

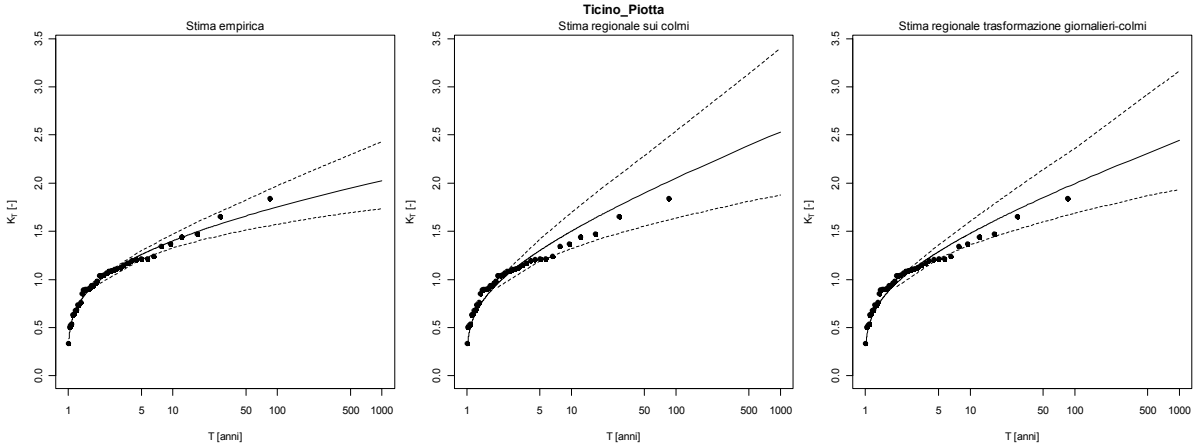


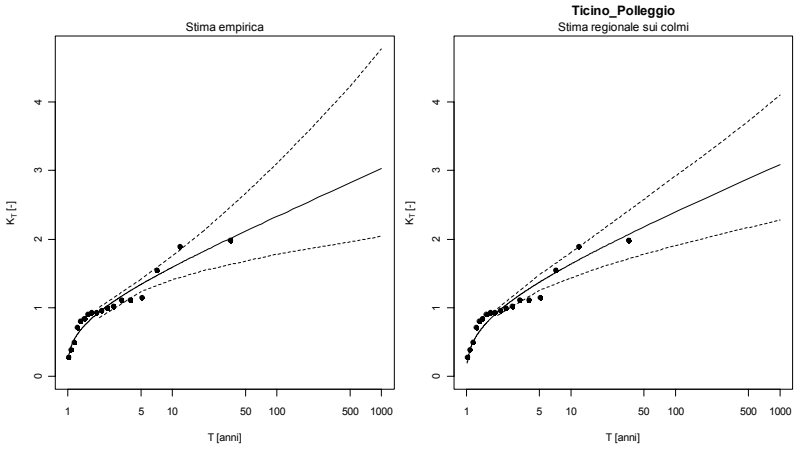
Riale_Roggiasca_Roveredo

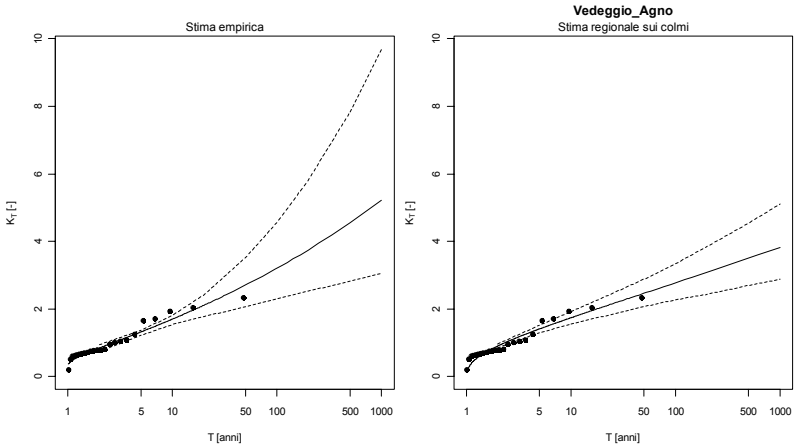


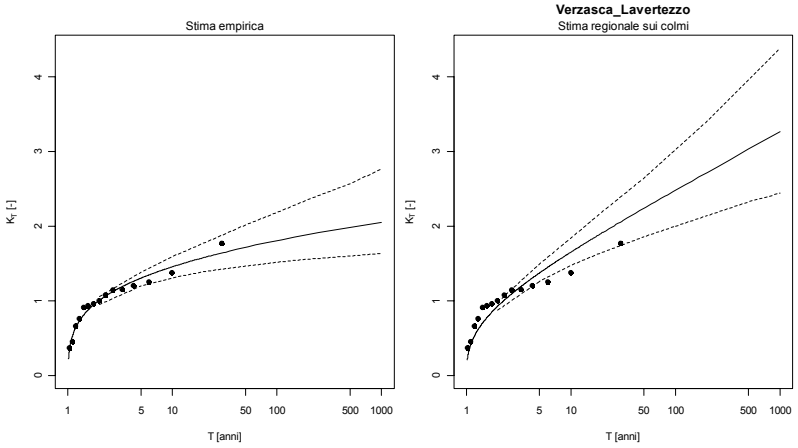

Saltina_Brig
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10

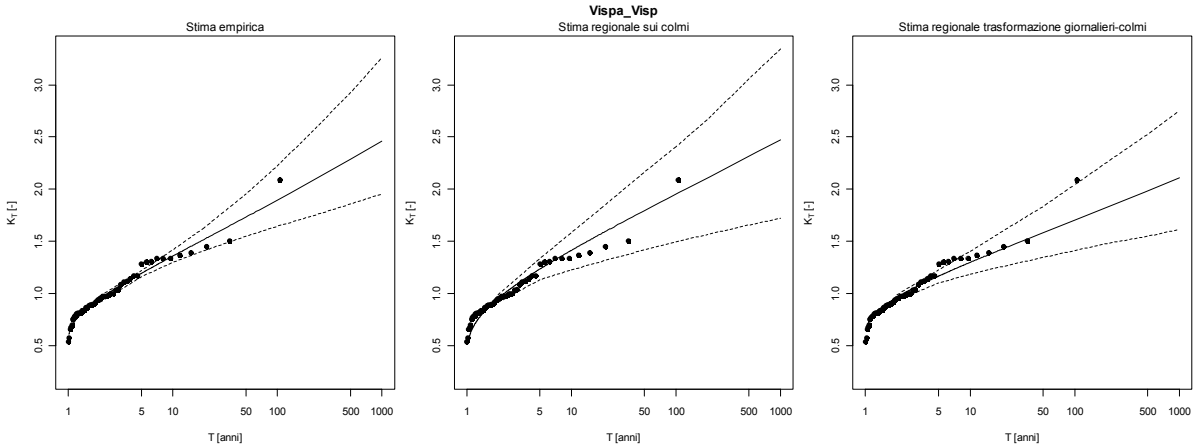

Ticino_Bellinzona
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0.0
0.5
1.0
1.5
2.0
2.5
3.0


Ticino_Piotta
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5


Ticino_Polleggio
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
1
2
3
4


Vedeggio_Agno
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
2
4
6
8
10


Verzasca_Lavertezzo
Stima empirica
Stima regionale sui colmi
K_T [-]
T [anni]
1
5
10
50
100
500
1000
0
1
2
3
4


Vispa_Visp
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
K_T [-]
T [anni]
0.5
1.0
1.5
2.0
2.5
3.0
1
5
10
50
100
500
1000

Parte III:

Stima regionale delle portate al colmo di piena

10 Stima della portata di progetto

Le procedure di stima delle portate di progetto Q_T vengono qui riferite, per consentire un riscontro con i dati, alle sezioni in cui sono presenti misure idrometriche. È però evidente che, in assenza di dati, l'unica possibile procedura applicabile sia quella interamente basata sui metodi di analisi regionale, le cui prestazioni sono discusse in questo capitolo.

In particolare, la stima delle portate di progetto con tempo di ritorno *T* anni nella generica sezione di un corso d'acqua viene effettuata combinando le stime della curva di crescita e della piena indice, tramite la relazione $Q_j(T) = K_j(T) \cdot Q_{ind,j}$. In particolare, se per la sezione sono presenti dei dati vengono esaminate 3 configurazioni:

1) Stima empirica:
   - piena indice stimata tramite la relazione (2.1) se la serie non presenta valori occasionali significativi, oppure tramite la relazione (2.2) se la serie storica è stata integrata con dati da Rapporti di evento o Sezioni F;
   - curva di crescita definita secondo la distribuzione lognormale a tre parametri, i cui parametri vengono stimati (Appendice D5) calcolando L_{cv} con la relazione (6.4) ed L_{ca} con la (6.5).

2) Stima regionale:
   - piena indice stimata con la formula multiregressiva (4.1);
   - curva di crescita lognormale con parametri stimati (Appendice D5) a partire da L_{cv} ed L_{ca} calcolati con le relazioni multiregressive (7.1) e (7.2).

3) Stima basata sugli estremi giornalieri:
   - piena indice stimata con la relazione (4.4);
   - curva di crescita definita secondo la distribuzione lognormale con parametri stimati da L_{cv} ed L_{ca} calcolati tramite le relazioni (7.3) e (7.4) espresse in funzione degli *L*-coefficienti calcolati sulle portate estreme giornaliere.

10.1 Valutazione dell'incertezza di stima

Come nel caso della stima della curva di crescita, definendo gli intervalli di confidenza della stima della piena di progetto Q(T) si ottiene una valutazione dell'incertezza della stima di Q(T). Si è adottata una procedura Monte-Carlo del tutto simile a quella impiegata per determinare gli intervalli di confidenza di K(T). La procedura prevede i seguenti passi:

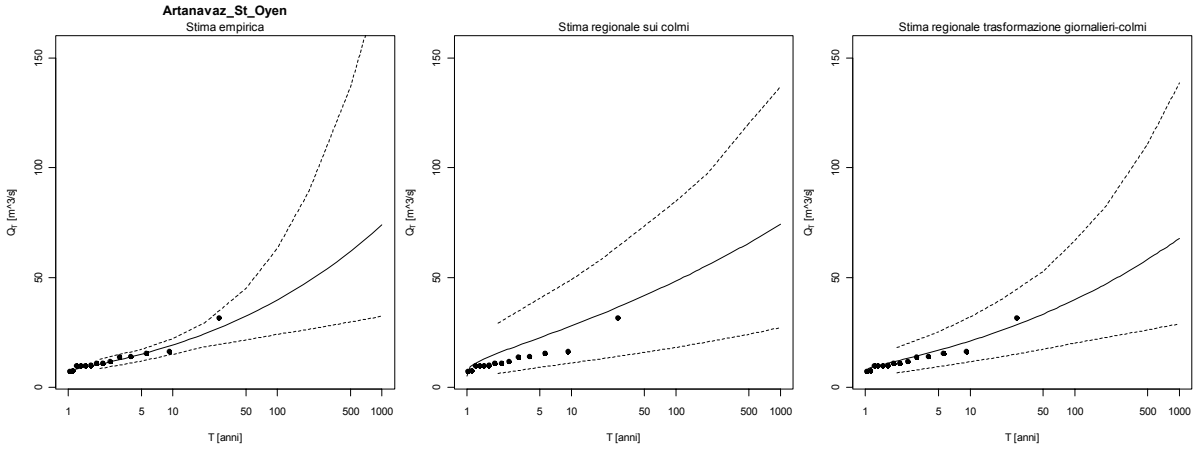
1) fissata una stazione *j*, si stimano la piena indice, LCV ed LCA tramite i dati disponibili (stima empirica) o tramite analisi multiregressiva (stima regionale) o utilizzando dati giornalieri (stima da estremi giornalieri).
2) Si determinano le varianze e la covarianza degli stimatori della piena indice, di LCV ed LCA, tramite le procedure descritte ai Capitoli 2 e 6. Si è supposto in tutti i casi che gli stimatori della piena indice non siano correlati con quelli di LCV ed LCA.
3) Si estraggono 1000 triplette di valori di piena indice, LCV ed LCA da una distribuzione normale trivariata, i cui 9 parametri sono le tre medie determinate al punto 1) e le varianze e covarianze determinate al punto 2).
4) Per ognuna delle *i* triplette di Q_{ind}, LCV ed LCA estratti dalla normale trivariata si calcola $Q_i(T)$, i=1,..,1000, per ogni tempo di ritorno di interresse.
5) Si fissa un coefficiente di confidenza α (ossia, una probabilità di ricadere all'interno dell'intervallo di confidenza) e si calcolano i limiti inferiore e superiore dell'intervallo di confidenza di Q(T) prendendo il *p*-esimo elemento del campione ordinato in senso crescente dei $Q_i(T)$, con $p = \left(0.5 - \frac{\alpha}{2}\right) \cdot 1000$ per il limite inferiore e $p = \left(0.5 + \frac{\alpha}{2}\right) \cdot 1000$ per il limite superiore.

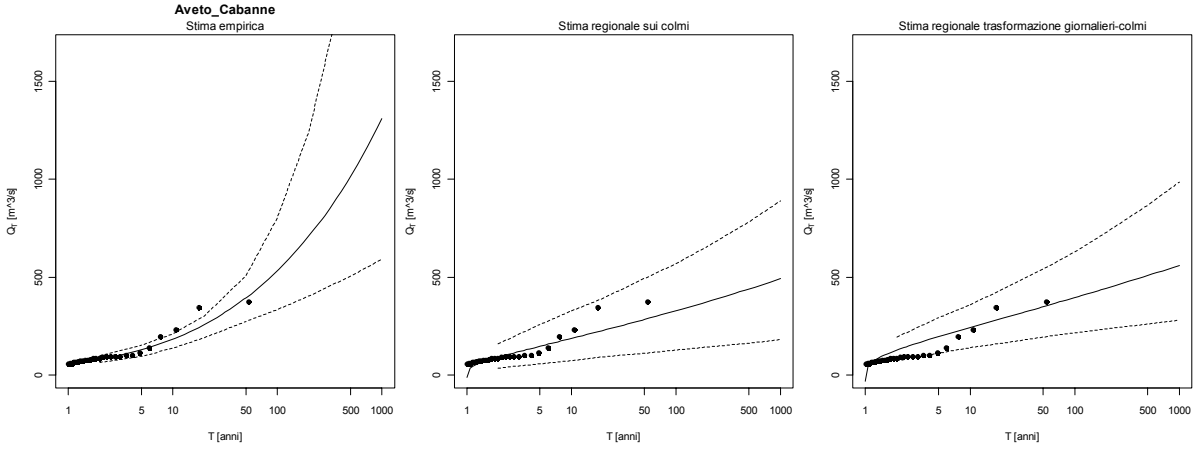
10.2 Applicazione ai bacini di interesse

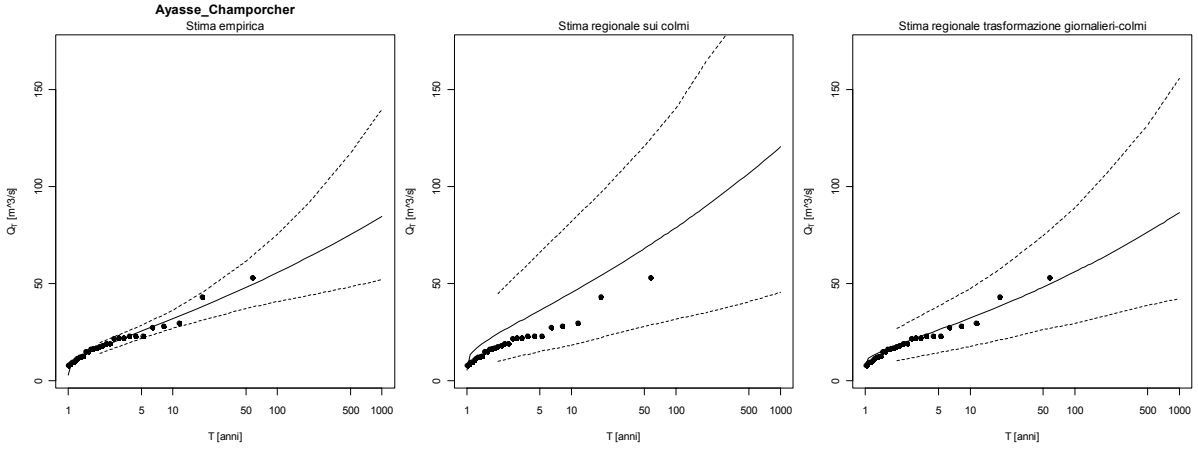
Nelle pagine seguenti sono riportati i risultati dell'applicazione al set di bacini qui considerato dei tre metodi di stima delle portate di progetto (e della relativa incertezza di stima) descritti nei paragrafi precedenti. Come per i risultati relativi alla curva di crescita, la prima colonna di grafici fa riferimento ai risultati che si ottengono considerando la stima empirica di Q(T), la seconda colonna alla stima regionale tramite analisi multiregressiva, e la terza colonna

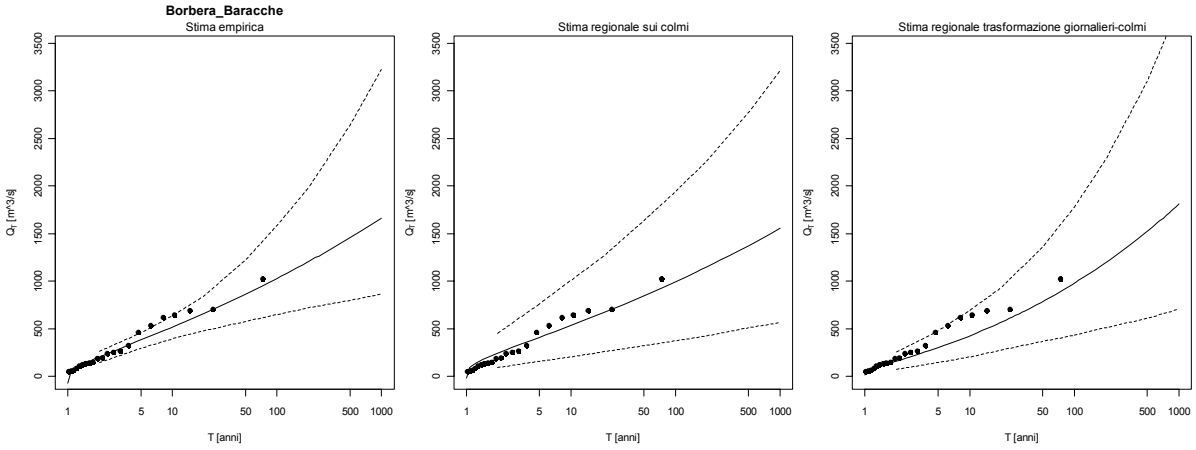
alla stima effettuata utilizzando anche gli estremi giornalieri. In ogni grafico sono riportati anche gli intervalli di confidenza, con livello di confidenza α=0.80, relativi alla piena di progetto stimata con i diversi metodi. Anche in questo caso si nota facilmente che gli intervalli di confidenza che derivano dall'applicazione della procedura descritta al paragrafo 10.1 sono spesso asimmetrici intorno al valore stimato, dal momento che non si è fatta alcuna ipotesi di normalità della distribuzione di Q(T).

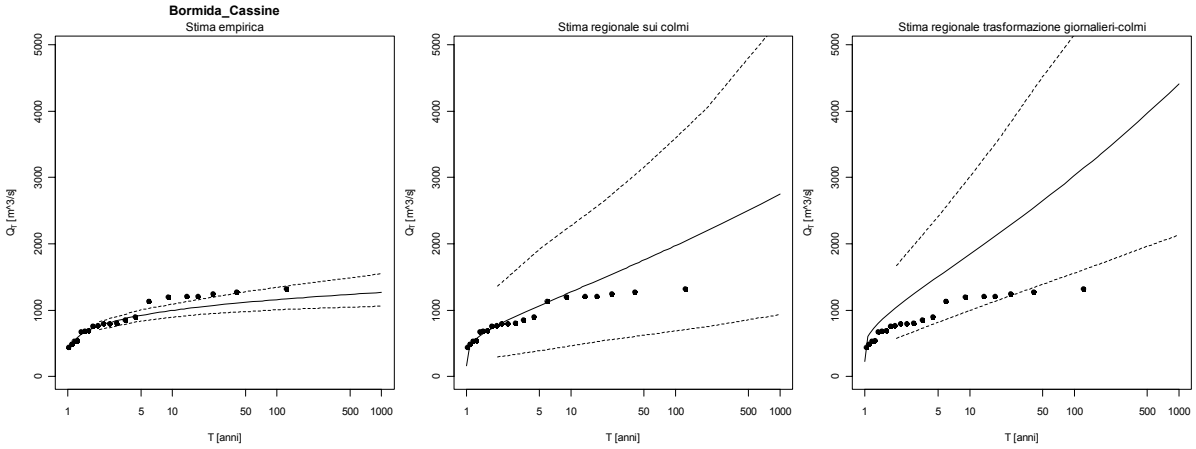
Anche ai grafici riportati nelle pagine seguenti si applicano considerazioni simili a quelle fatte per la curva di crescita: gli intervalli di confidenza ottenuti per la stima empirica (prima colonna di grafici) hanno ampiezza molto diversa a seconda della dimensione del campione di dati disponibile, e risultano particolarmente estesi intorno al valore centrale quando si hanno pochi dati disponili. Nel caso di stima regionale (seconda colonna di grafici) l'ampiezza degli intervalli di confidenza è invece abbastanza costante nelle diverse sezioni considerate, dal momento che la varianza degli stimatori regionali non dipende dalla dimensione di alcun campione. Infine l'ampiezza degli intervalli di confidenza riportati nei grafici della terza colonna, è variabile nelle diverse sezioni considerate, dal momento che la varianza degli stimatori regionali di LCV ed LCA dipende in questo caso anche dalla dimensione del campione di dati giornalieri disponibile nella specifica sezione. Alcuni grafici della seconda colonna sono mancanti perché non si può stimare la piena indice in quanto il valore di c_f è dispononibile solamente sul territorio Piemontese. Inoltre, alcuni grafici della colonna 3 sono mancanti perché in tali sezioni non sono disponibili dati di portata giornaliera.

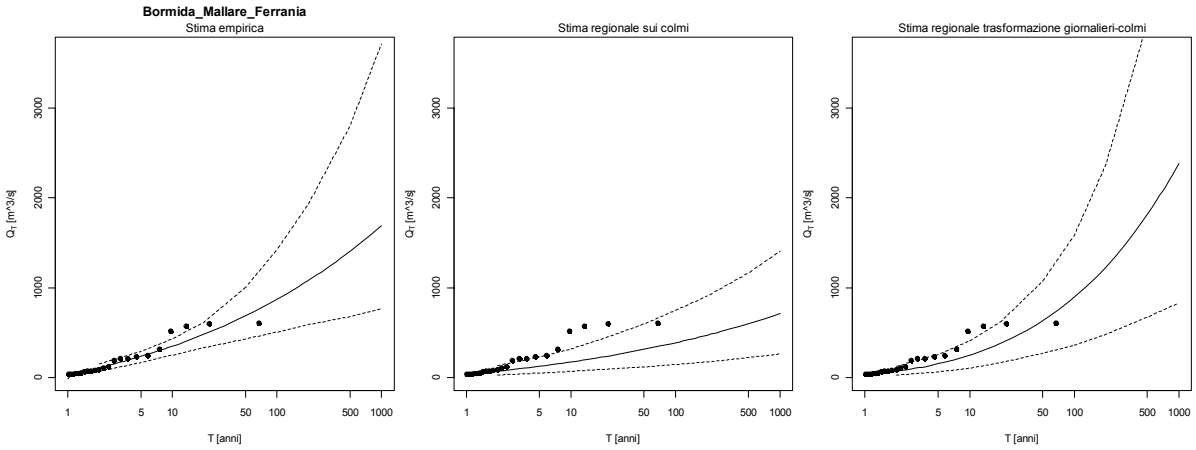

Artanavaz_St_Oyen
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
1
5
10
50
100
500
1000

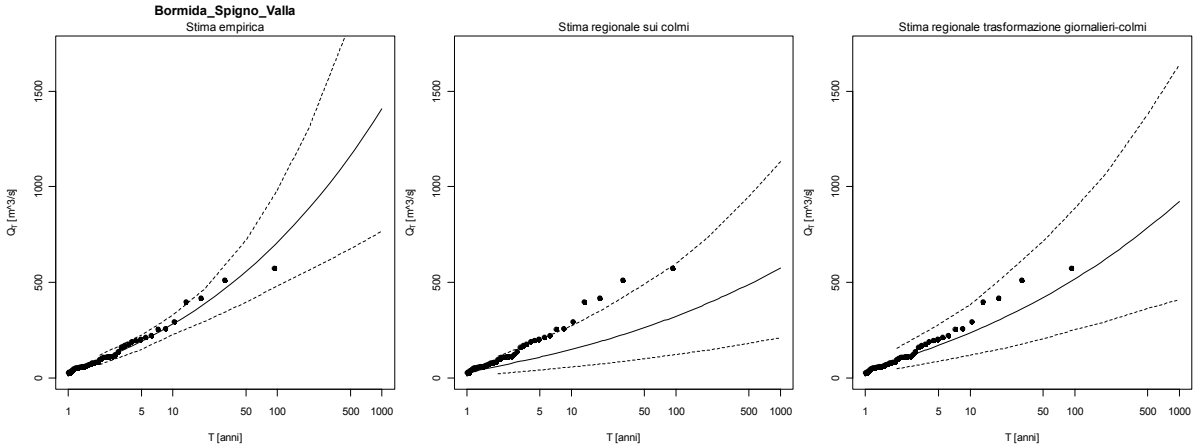

Aveto_Cabanne
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
1
5
10
50
100
500
1000

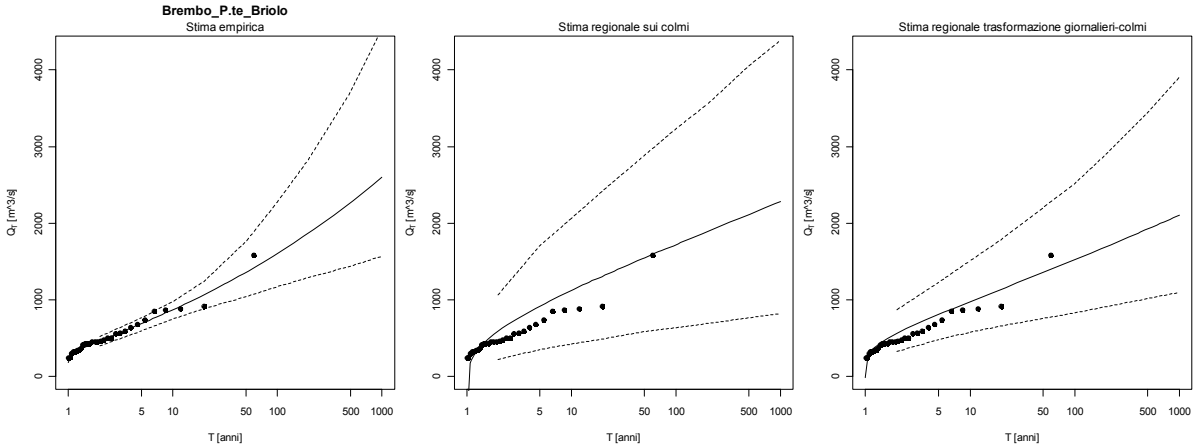

Ayasse_Champorcher
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
1
5
10
50
100
500
1000

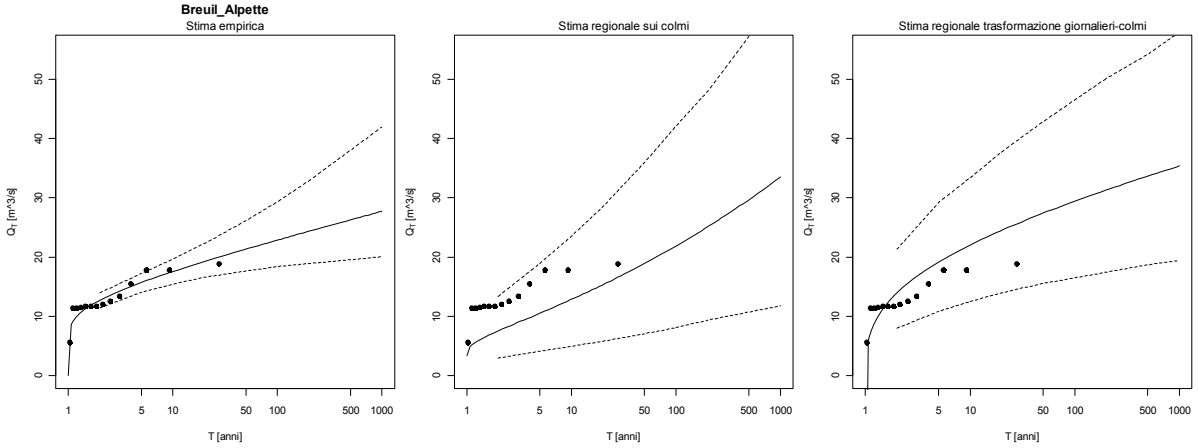

Borbera_Baracche
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

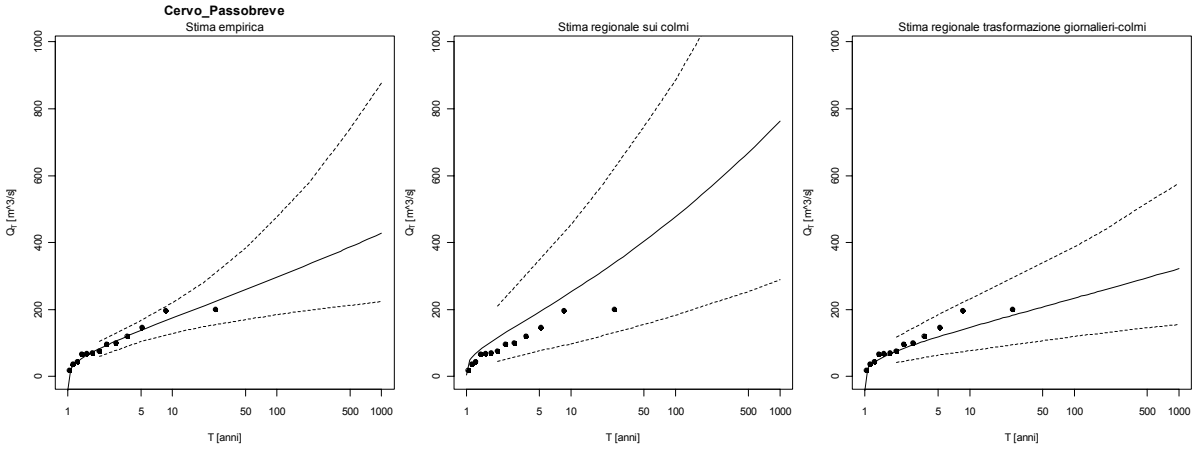

Bormida_Cassine
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

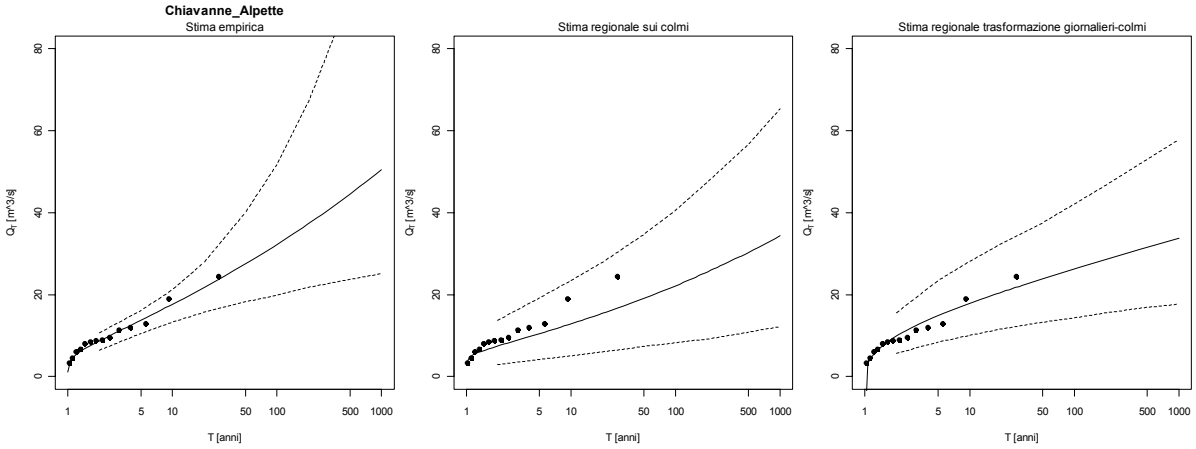

Bormida_Mallare_Ferrania
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

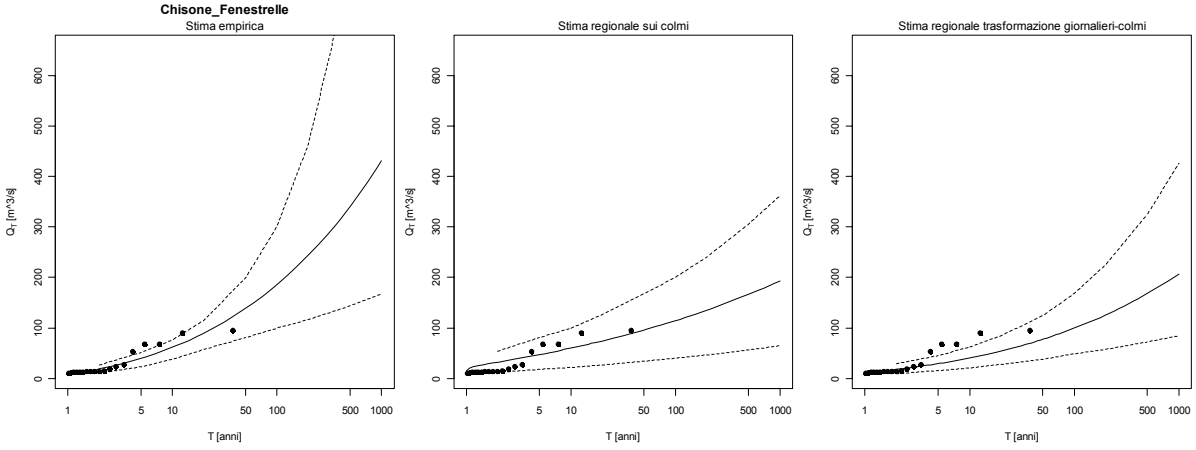

Bormida_Spigno_Valla
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
1
5
10
50
100
500
1000

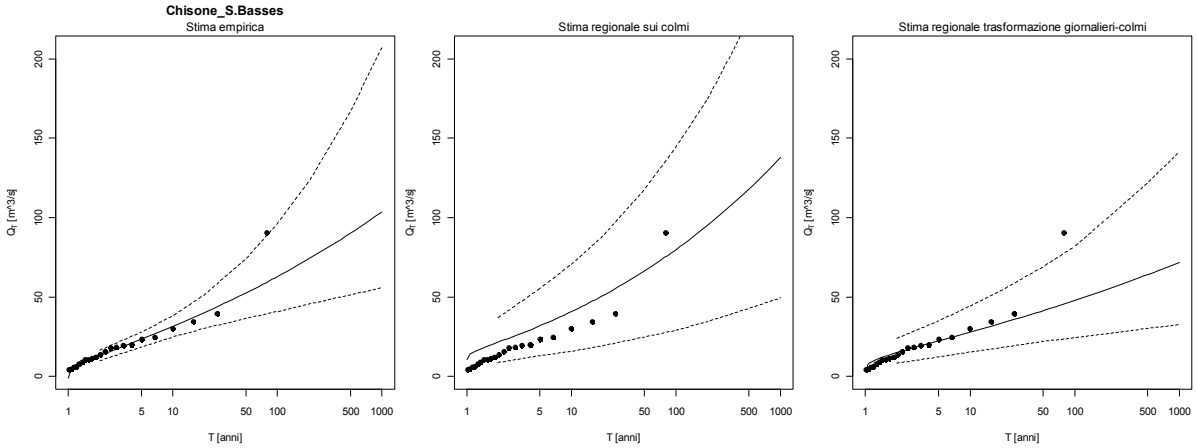

Brembo_P.te_Briolo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
1
5
10
50
100
500
1000

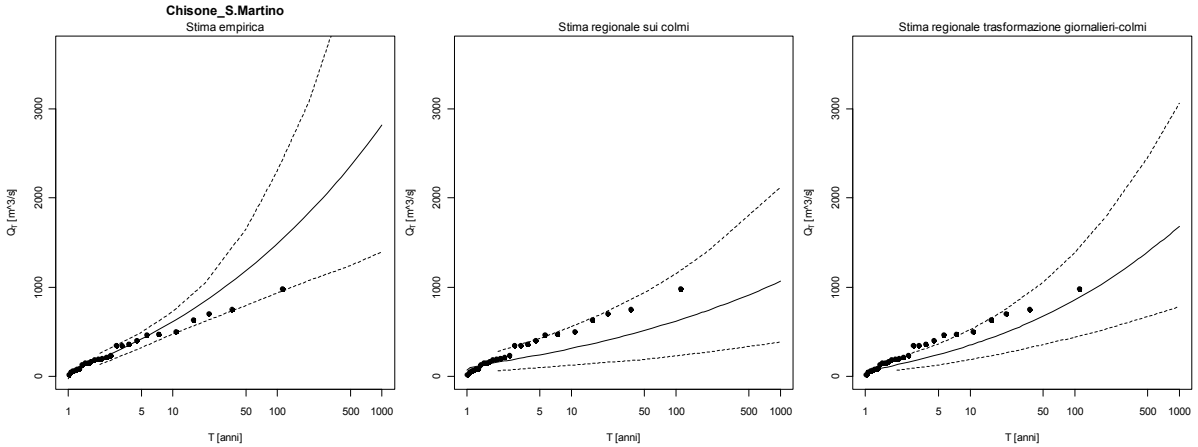

Breuil_Alpette
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
10
20
30
40
50
1
5
10
50
100
500
1000

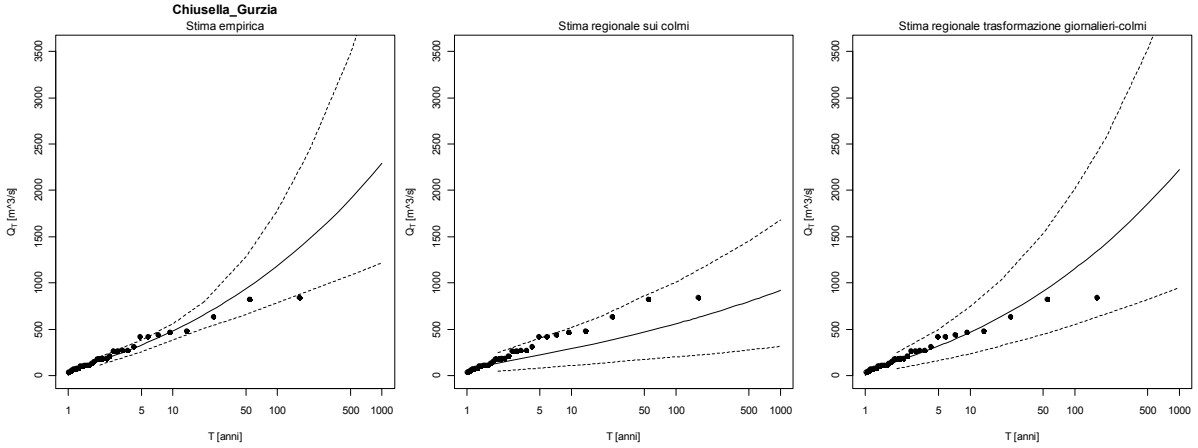

Cervo_Passobreve
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1
5
10
50
100
500
1000

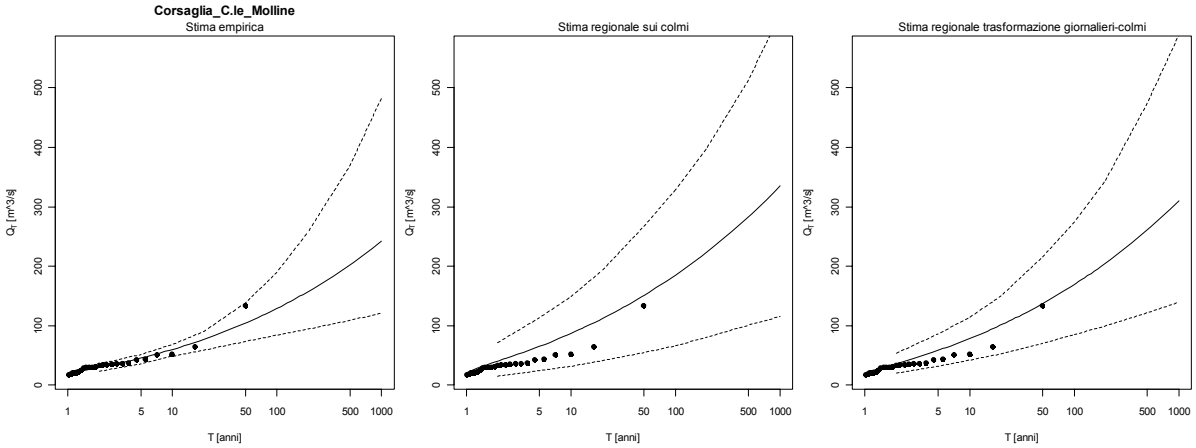

Chiavanne_Alpette
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
20
40
60
80
1
5
10
50
100
500
1000

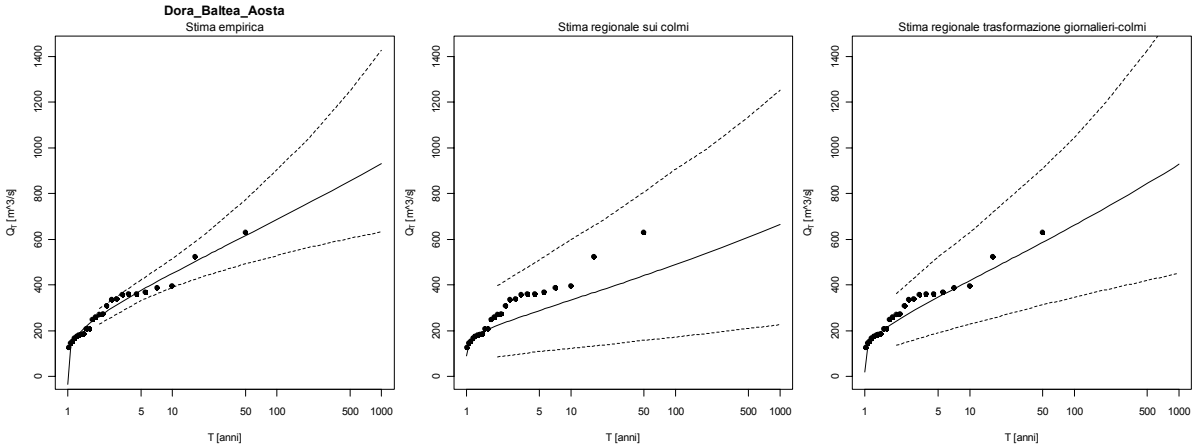

Chisone_Fenestrelle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
100
200
300
400
500
600
1
5
10
50
100
500
1000

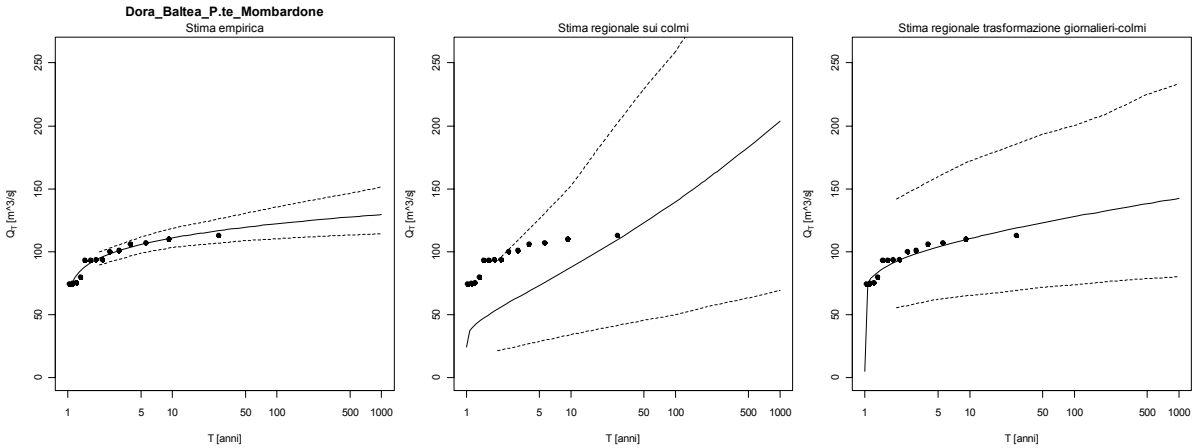

Chisone_S.Basses
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
200
1
5
10
50
100
500
1000

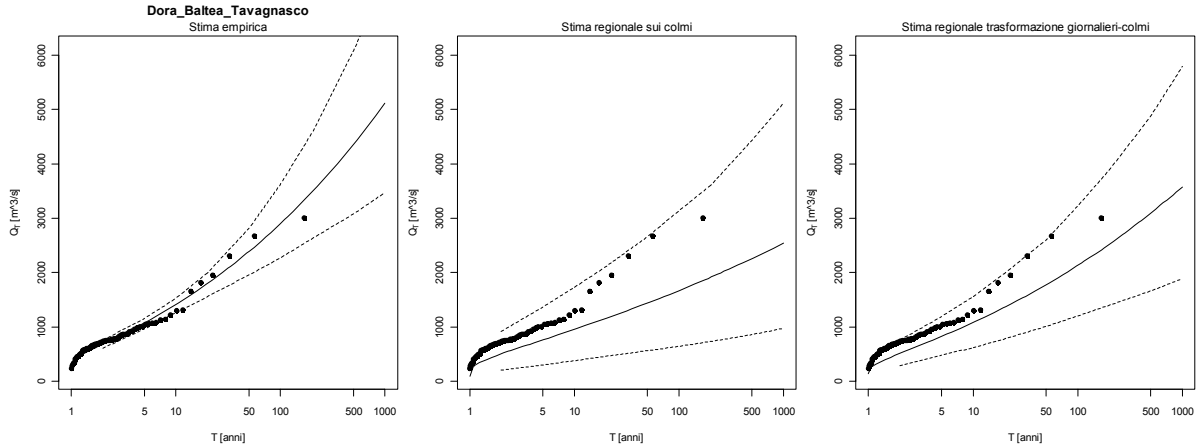

Chisone_S.Martino
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
1
5
10
50
100
500
1000

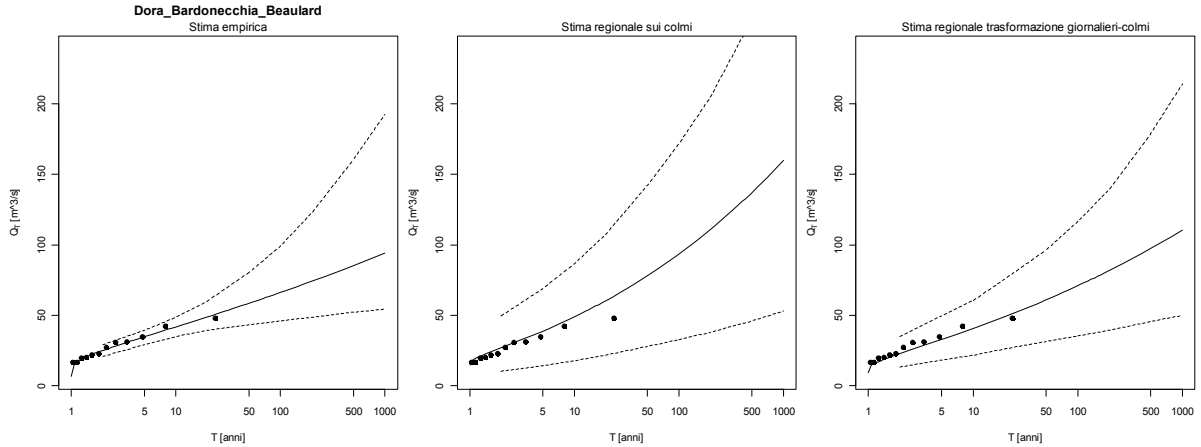

Chiusella_Gurzia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
2000
2500
3000
3500
1
5
10
50
100
500
1000

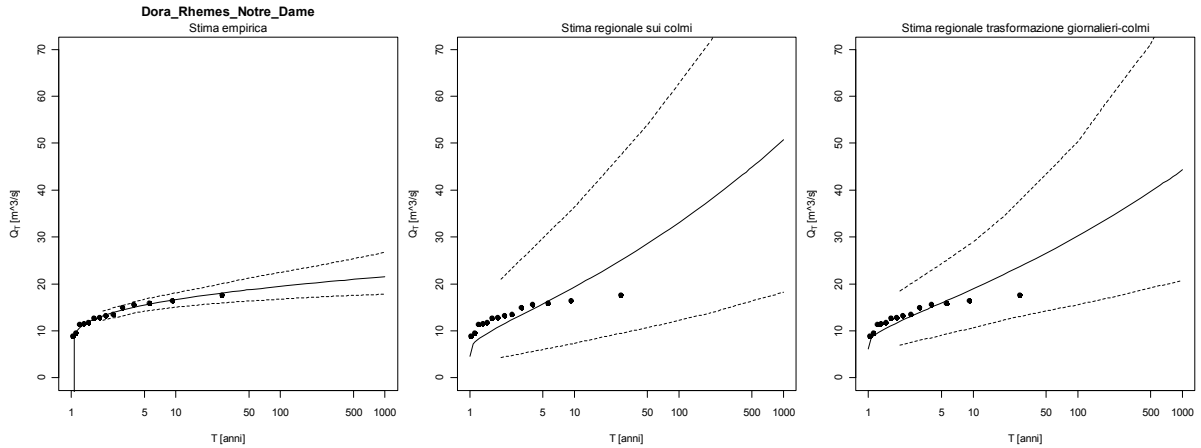

Corsaglia_C.le_Molline
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

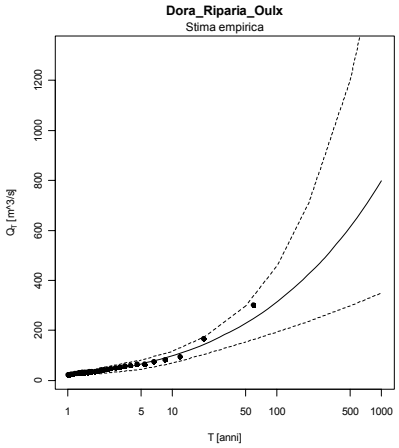

Dora_Baltea_Aosta
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

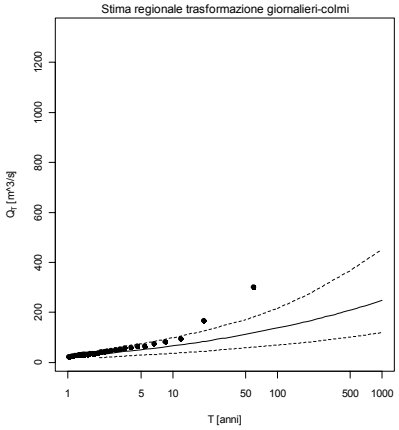

Dora_Baltea_P.te_Mombardone
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

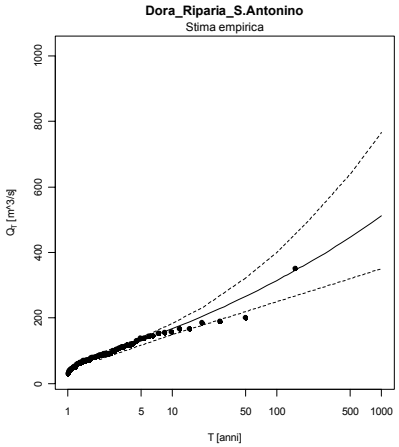

Dora_Baltea_Tavagnasco
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
5000
6000
1
5
10
50
100
500
1000

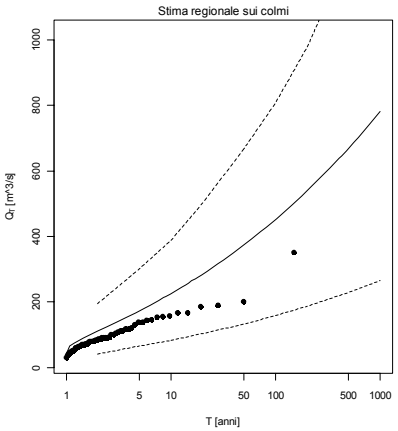

Dora_Bardonecchia_Beaulard
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
200
1
5
10
50
100
500
1000

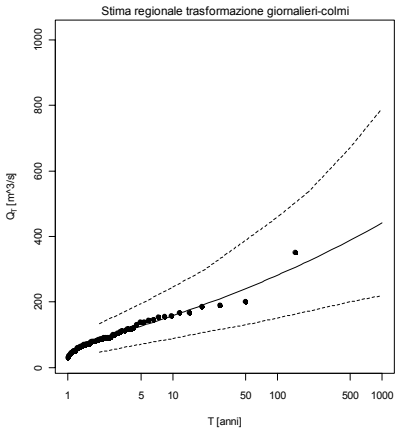

Dora_Rhemes_Notre_Dame
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
10
20
30
40
50
60
70
1
5
10
50
100
500
1000

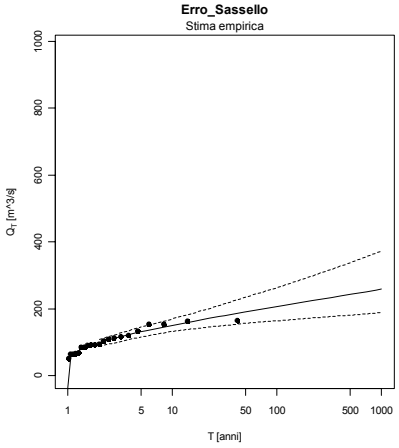

Dora_Riparia_Oulx
Stima empirica
Q_T [m^3/s]
T [anni]

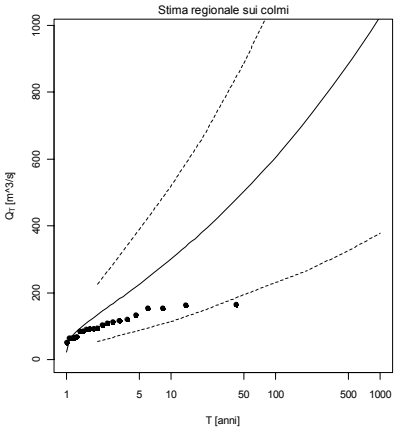

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

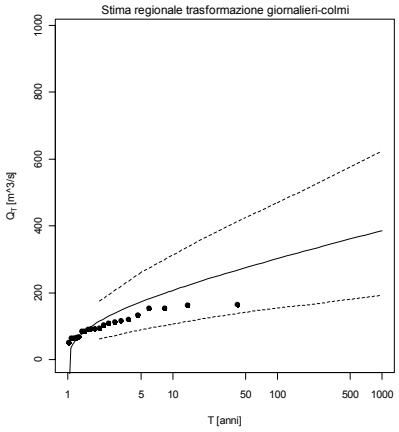

Dora_Riparia_S.Antonino
Stima empirica
Q_T [m^3/s]
T [anni]

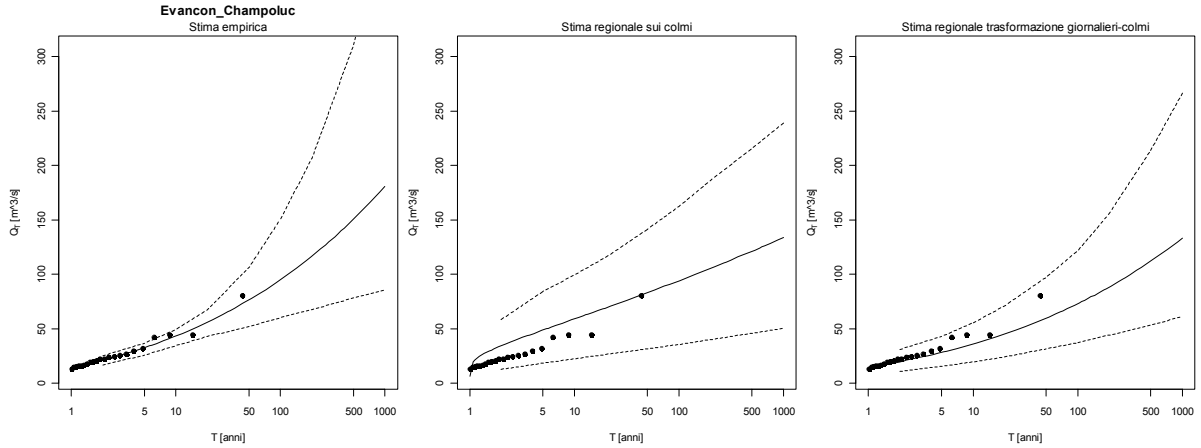

Stima regionale sui colmi
Q_T [m^3/s]
T [anni]

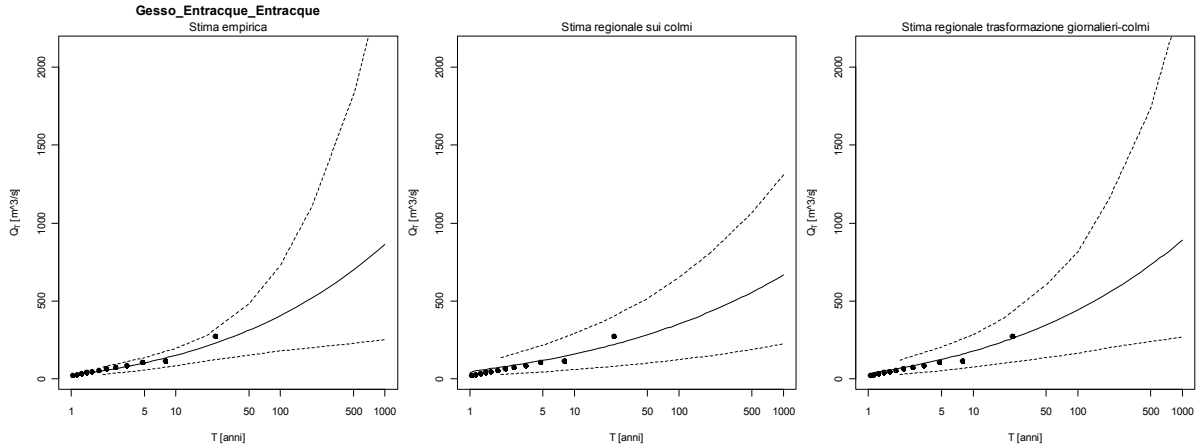

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

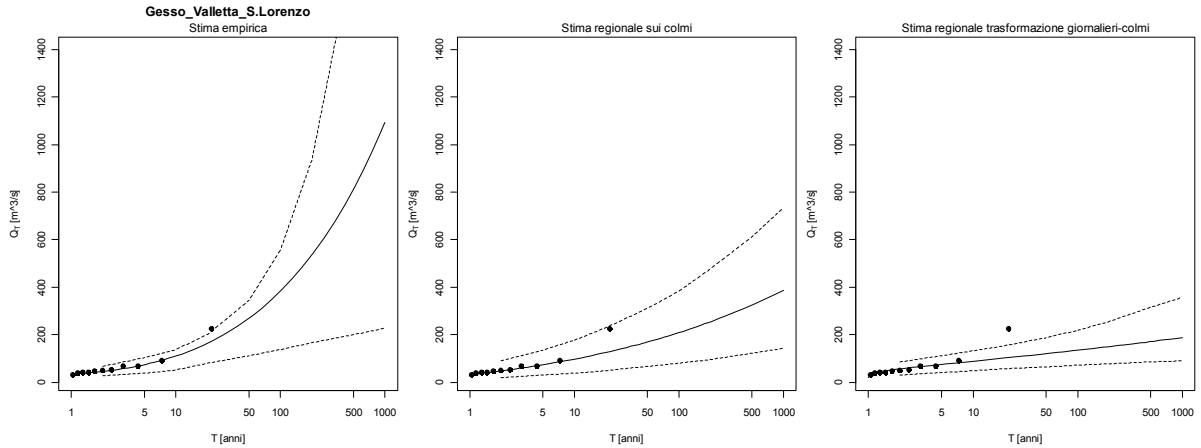

Erro_Sassello
Stima empirica
Q_T [m^3/s]
T [anni]

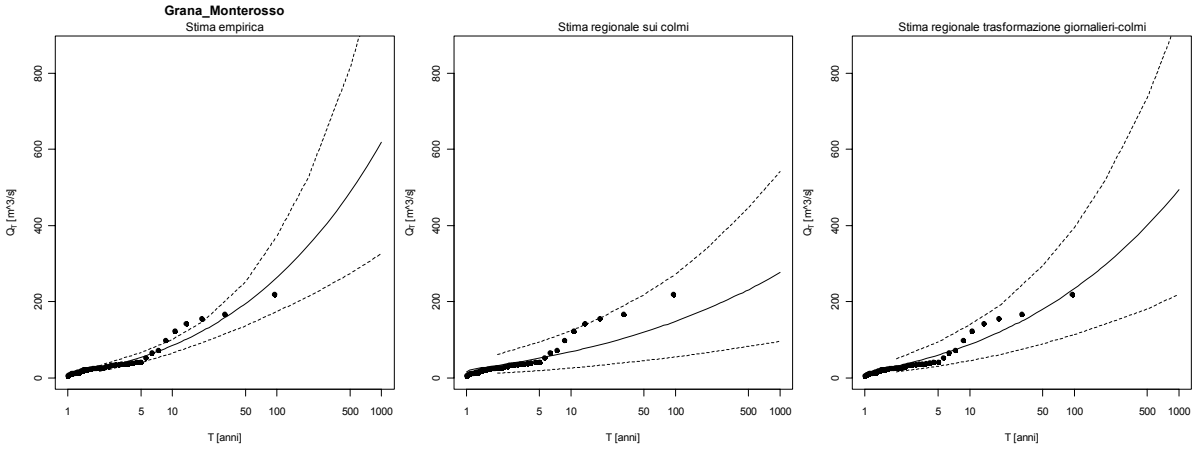

Stima regionale sui colmi
Q_T [m^3/s]
T [anni]

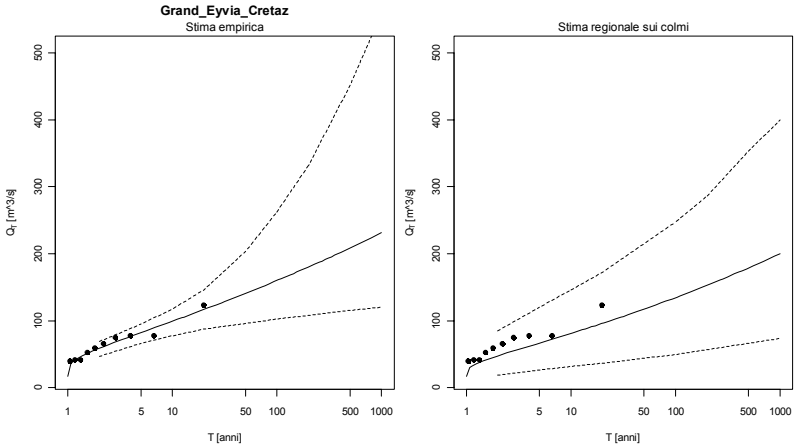

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

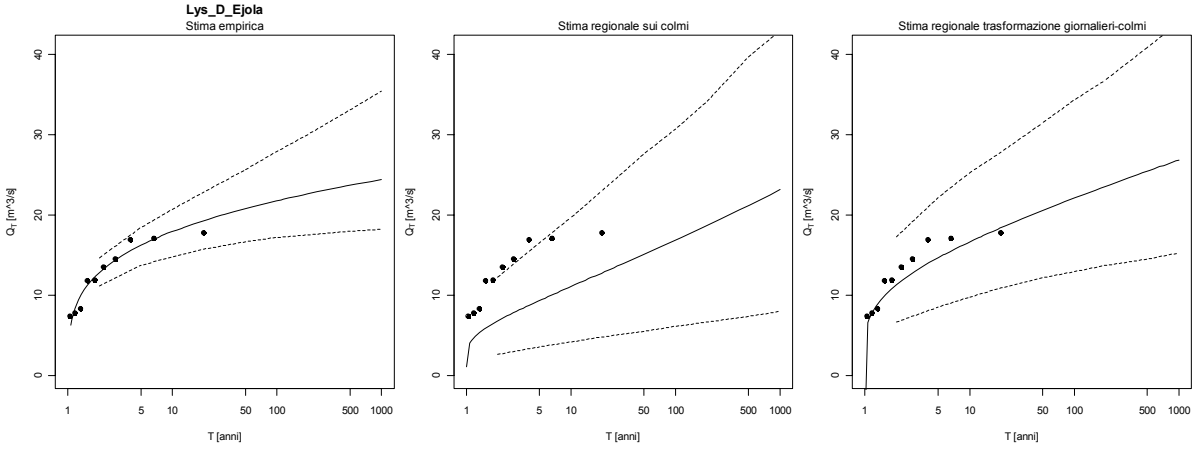

Evancon_Champoluc
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
200
250
300
1
5
10
50
100
500
1000

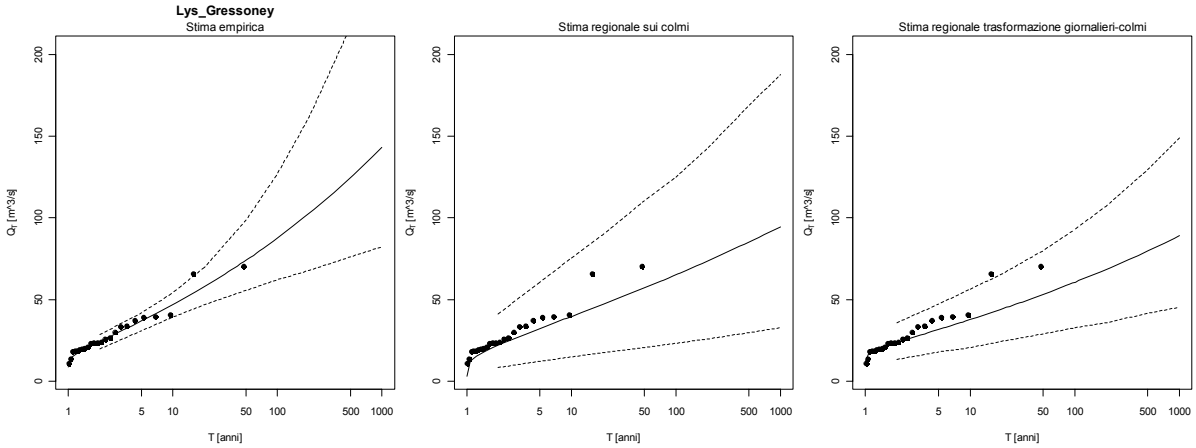

Gesso_Entracque_Entracque
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
2000
1
5
10
50
100
500
1000

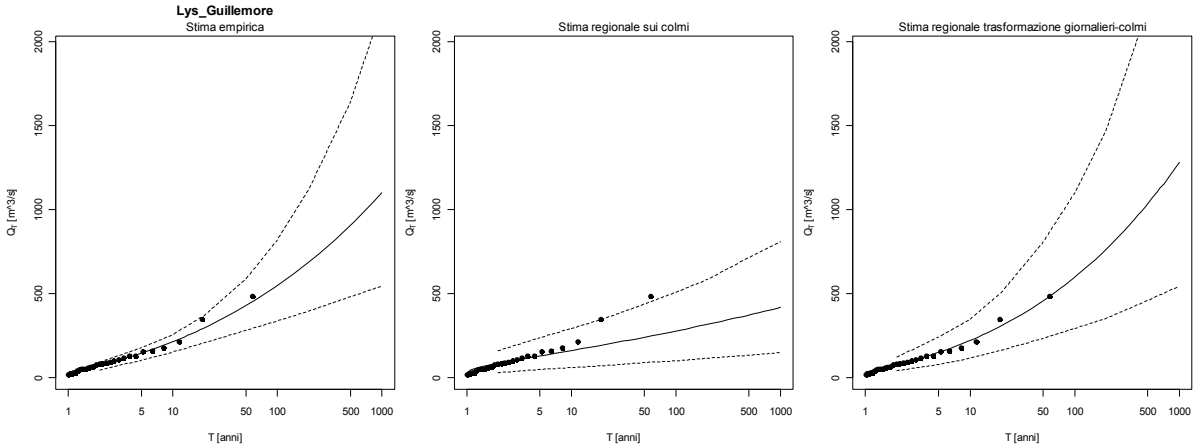

Gesso_Valletta_S.Lorenzo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000

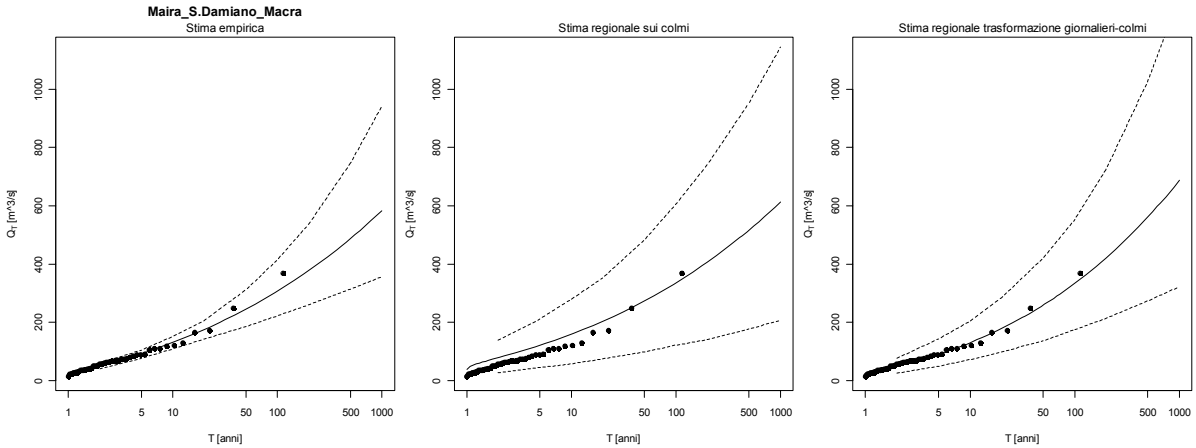

Grana_Monterosso
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1
5
10
50
100
500
1000

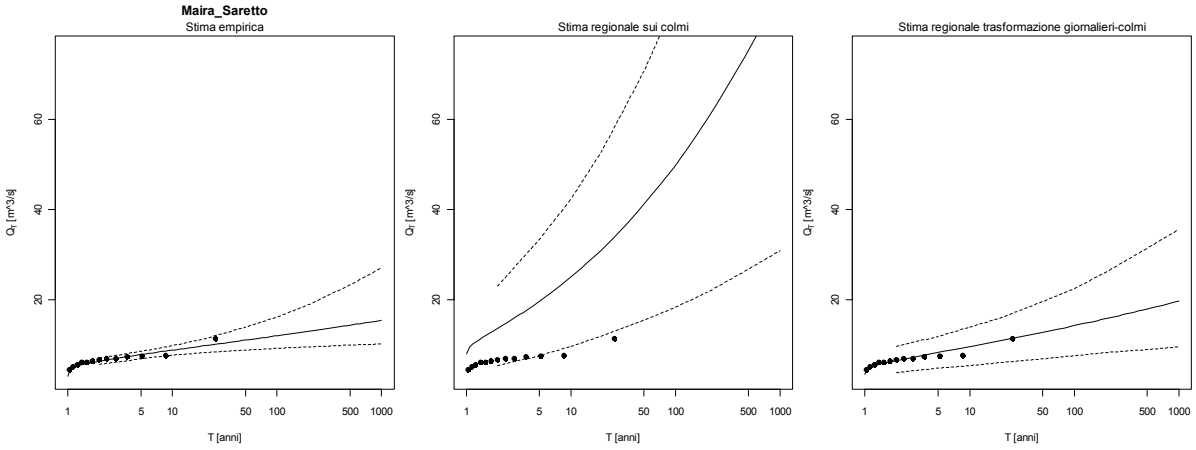

Grand_Eyvia_Cretaz
Stima empirica
Stima regionale sui colmi
Q_T [m^3/s]
T [anni]
0
100
200
300
400
500
1
5
10
50
100
500
1000

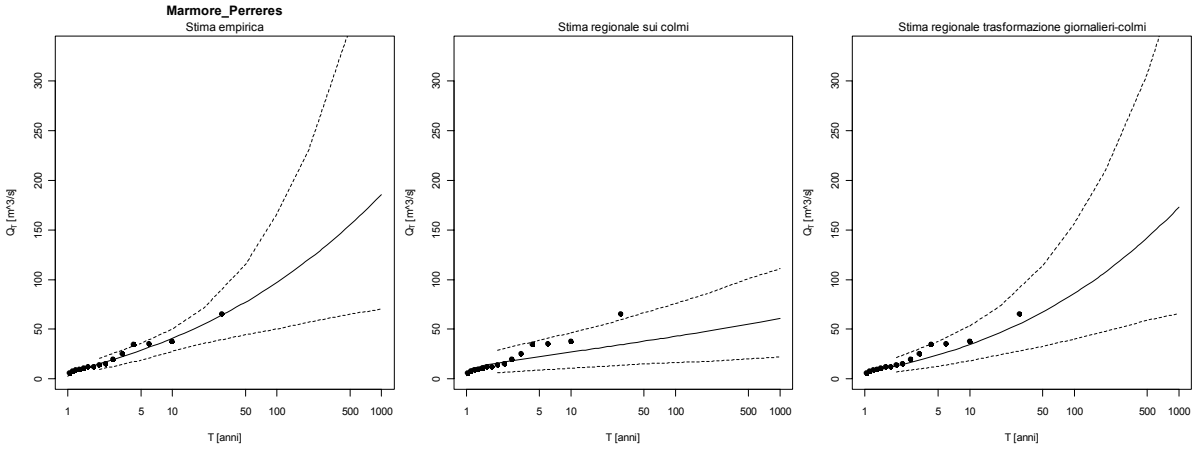

Lys_D_Ejola
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
10
20
30
40
1
5
10
50
100
500
1000

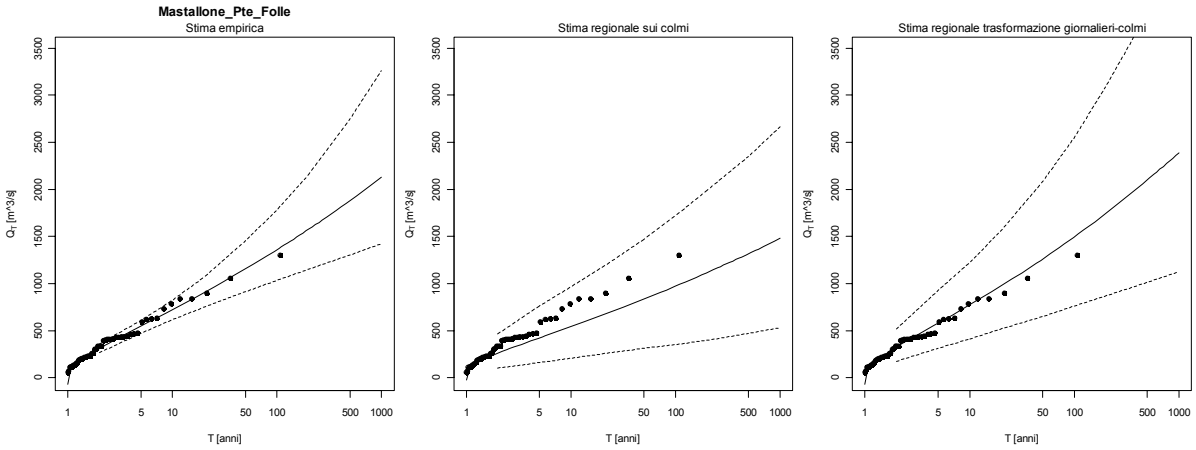

Lys_Gressoney
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
200
1
5
10
50
100
500
1000

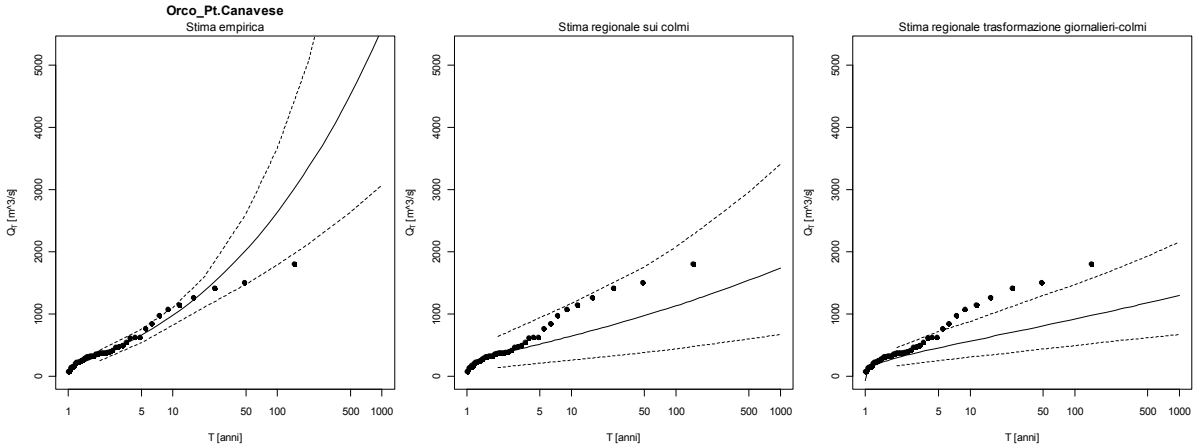

Lys_Guillemore
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
2000
1
5
10
50
100
500
1000

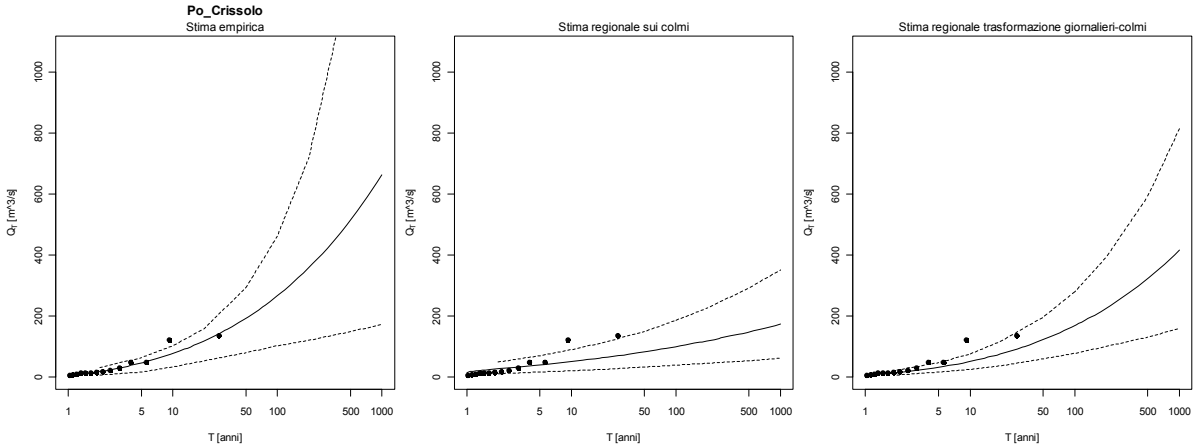

Maira_S.Damiano_Macra
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1
5
10
50
100
500
1000

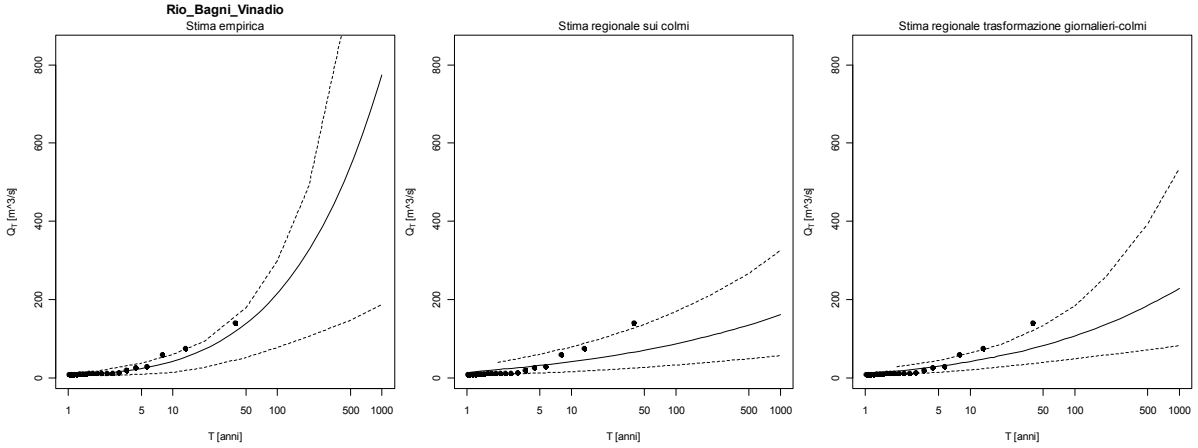

Maira_Saretto
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

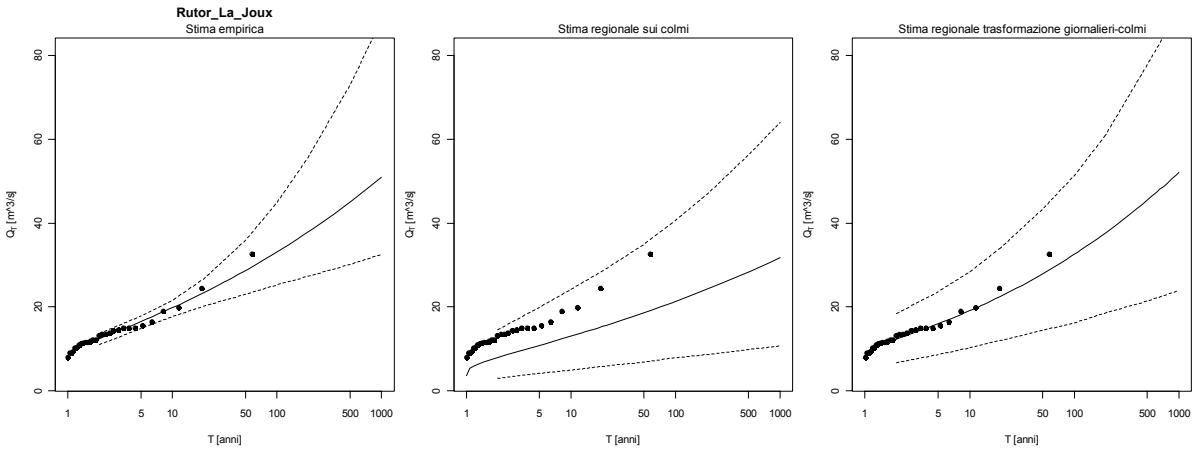

Marmore_Perreres
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

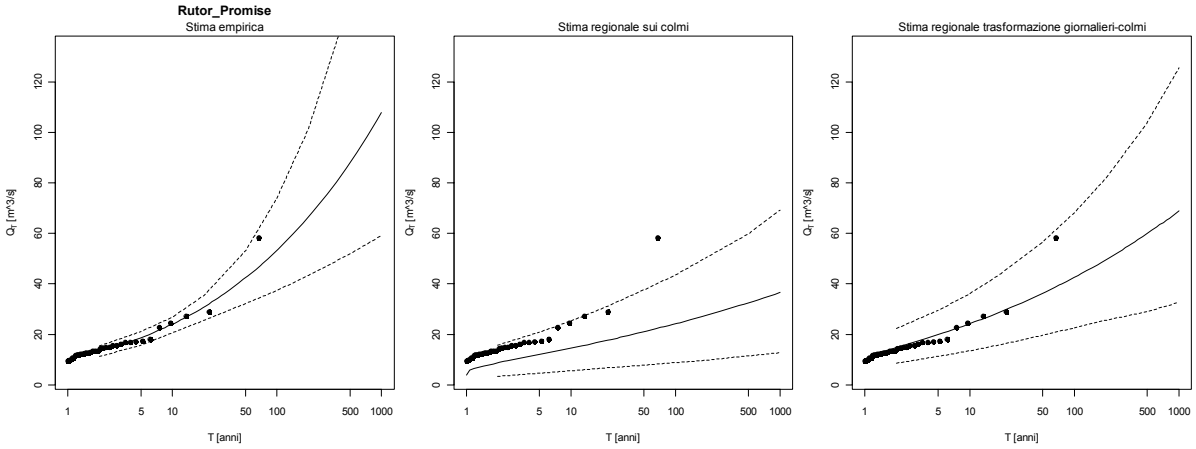

Mastallone_Pte_Folle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

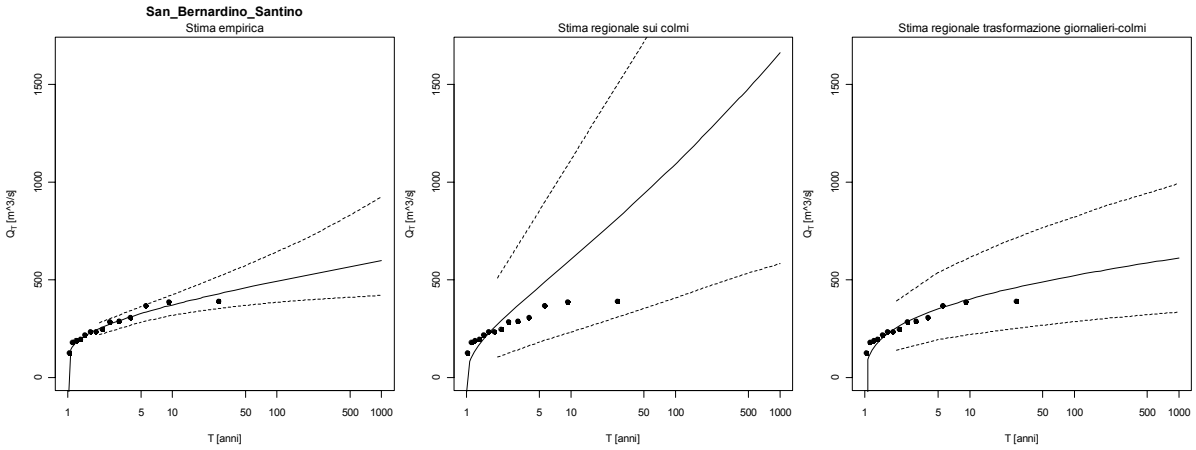

Orco_Pt.Canavese
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
5000
1
5
10
50
100
500
1000

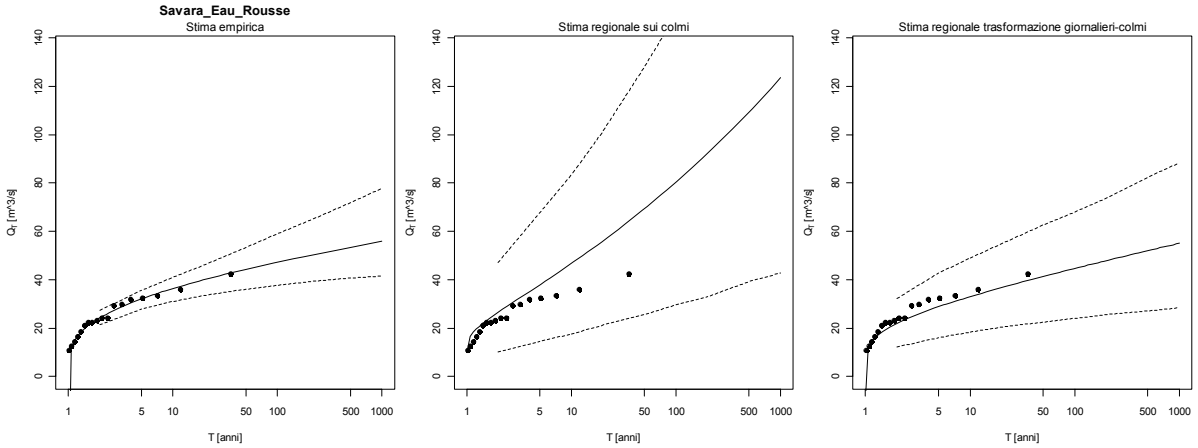

Po_Crissolo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1
5
10
50
100
500
1000

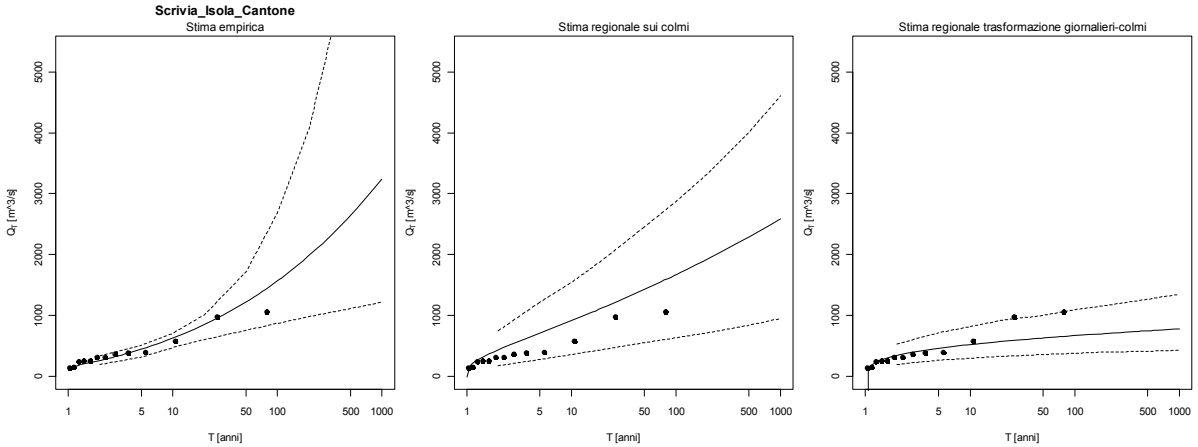

Rio_Bagni_Vinadio
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1
5
10
50
100
500
1000

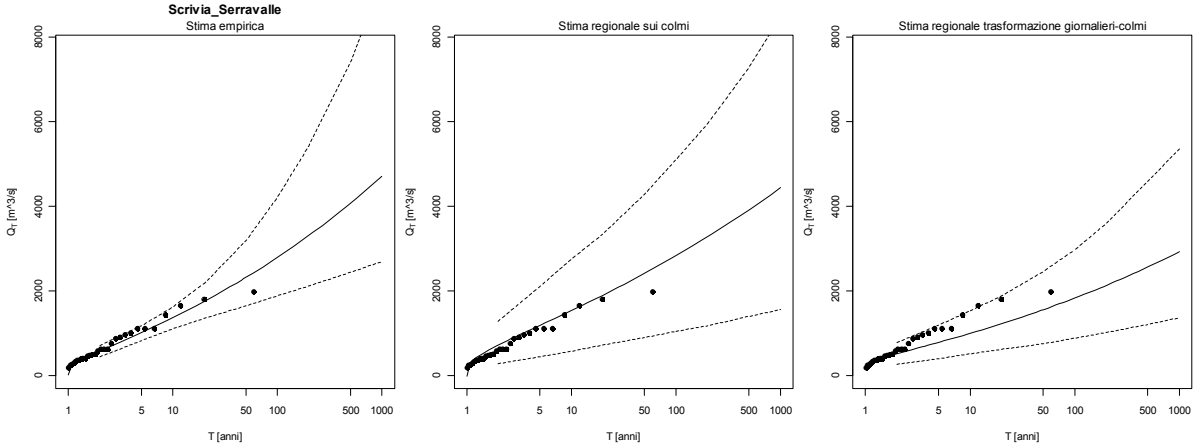

Rutor_La_Joux
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
20
40
60
80
1
5
10
50
100
500
1000

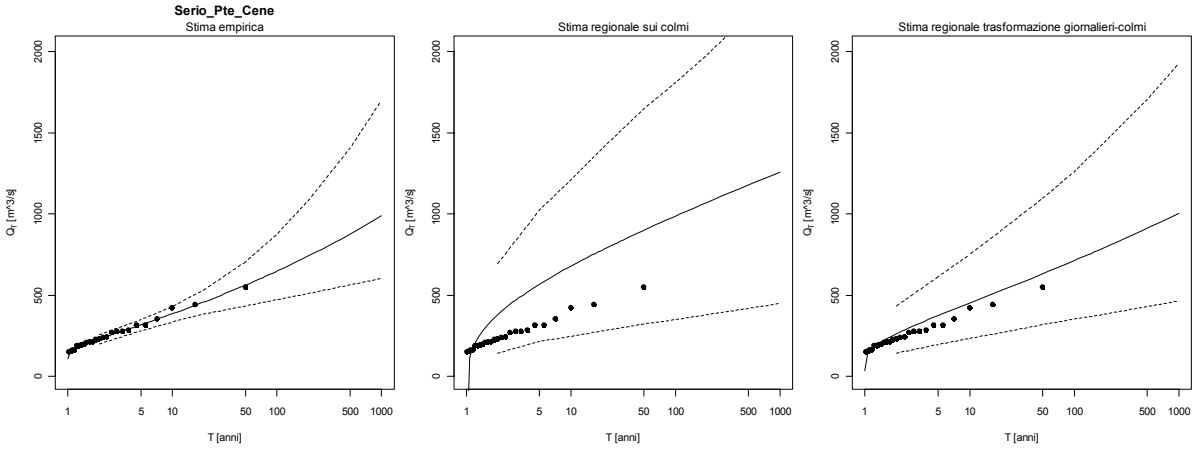

Rutor_Promise
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
20
40
60
80
100
120
1
5
10
50
100
500
1000

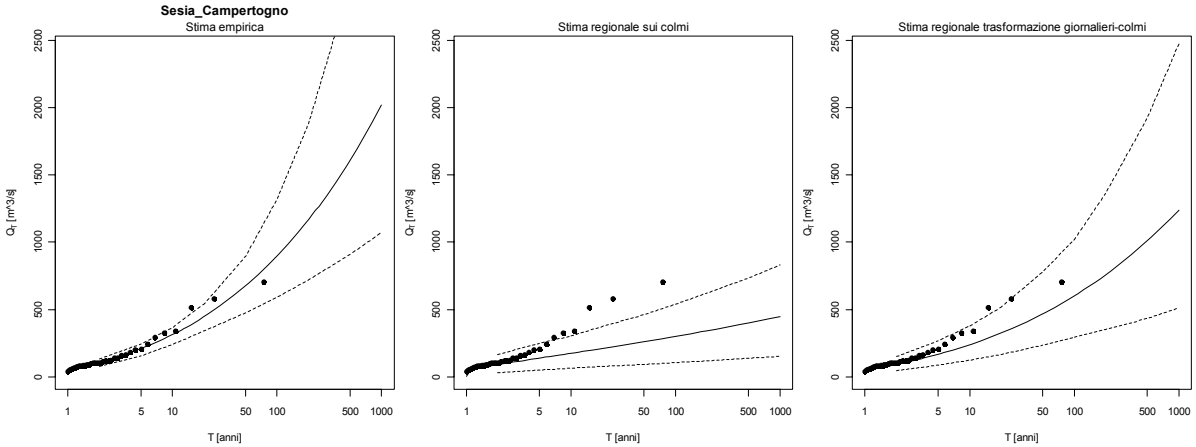

San_Bernardino_Santino
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
1
5
10
50
100
500
1000

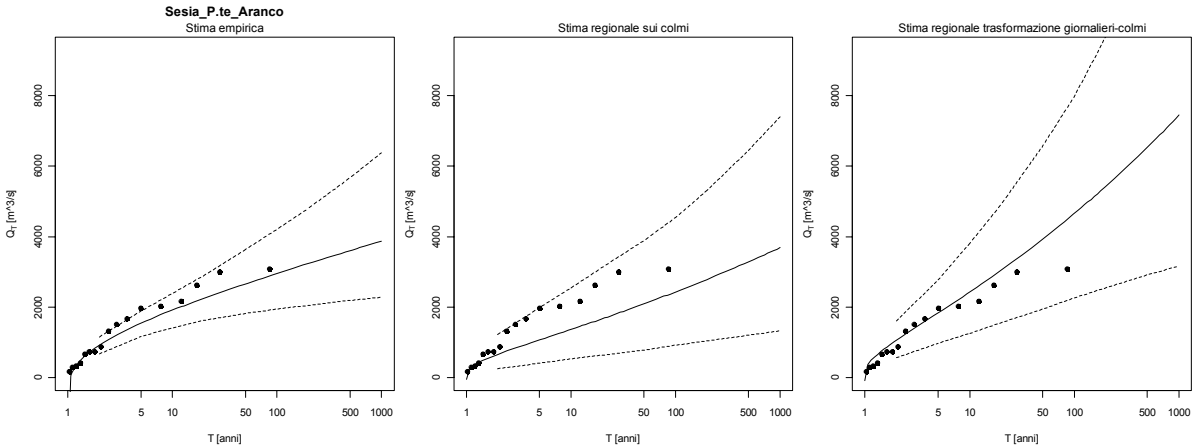

Savara_Eau_Rousse
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
20
40
60
80
100
120
140
1
5
10
50
100
500
1000

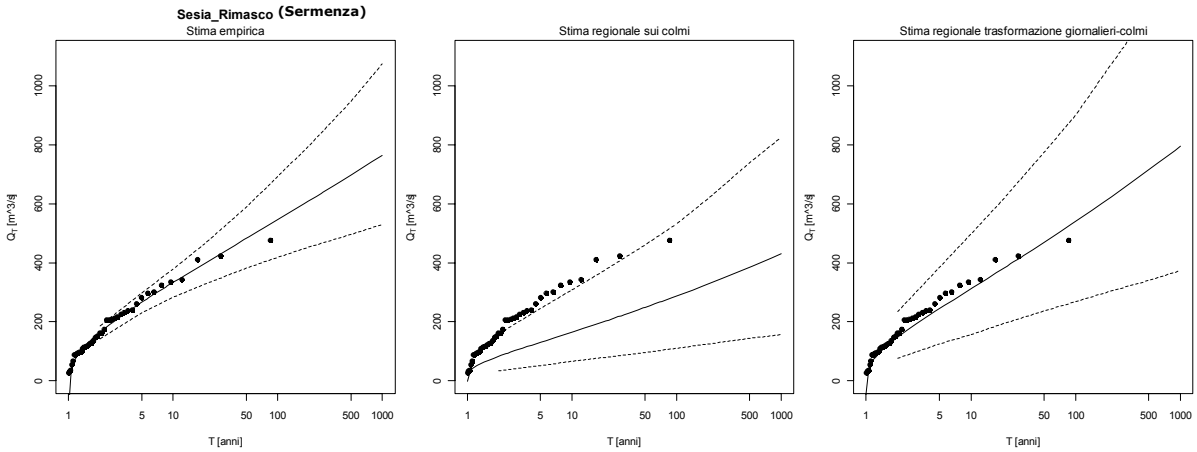

Scrivia_Isola_Cantone
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
5000
1
5
10
50
100
500
1000

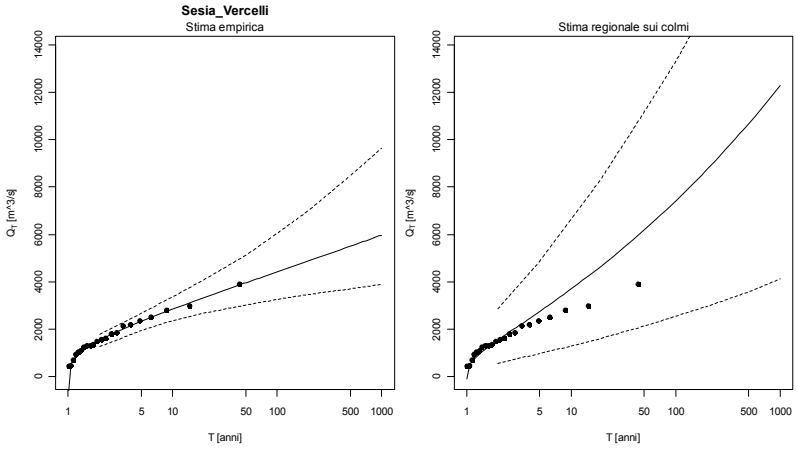

Scrivia_Serravalle
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
2000
4000
6000
8000
1
5
10
50
100
500
1000

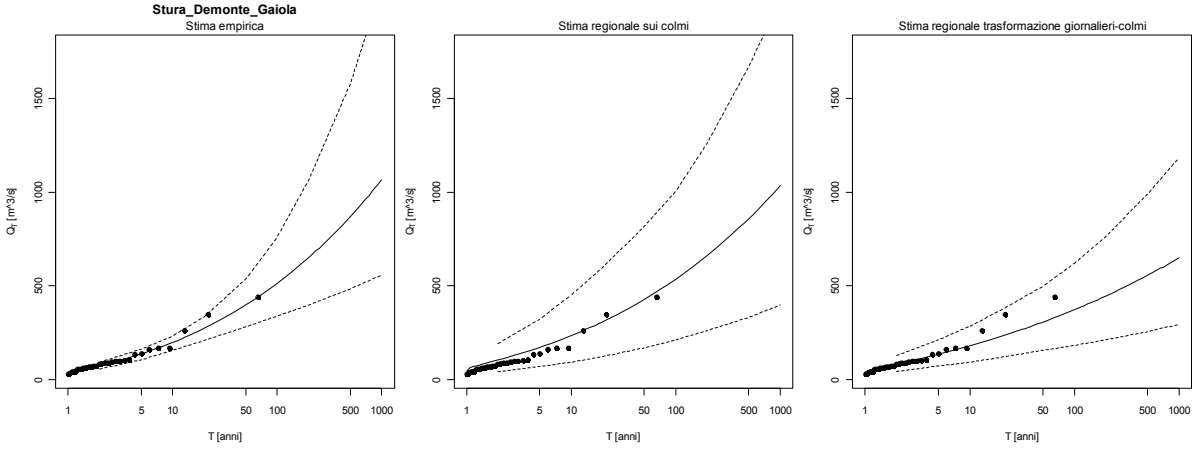

Serio_Pte_Cene
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

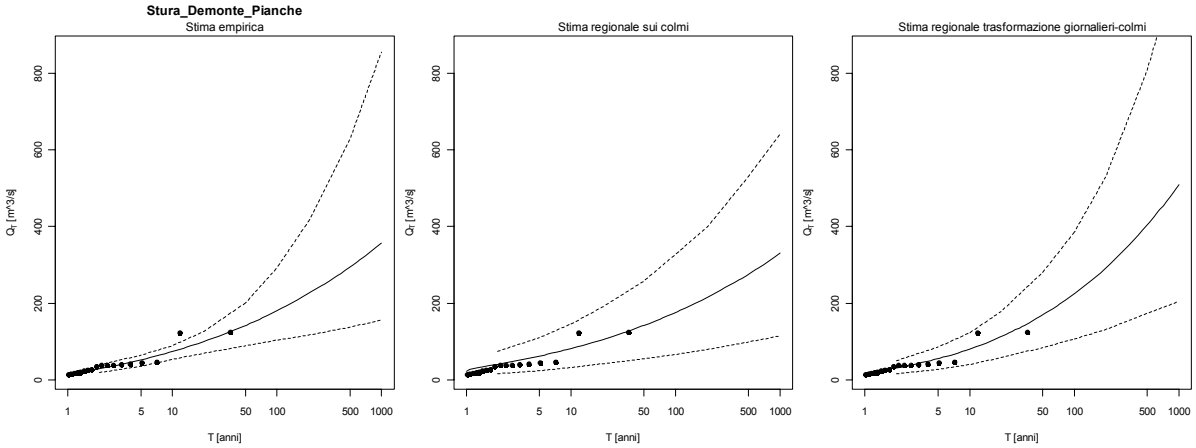

Sesia_Campertogno
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

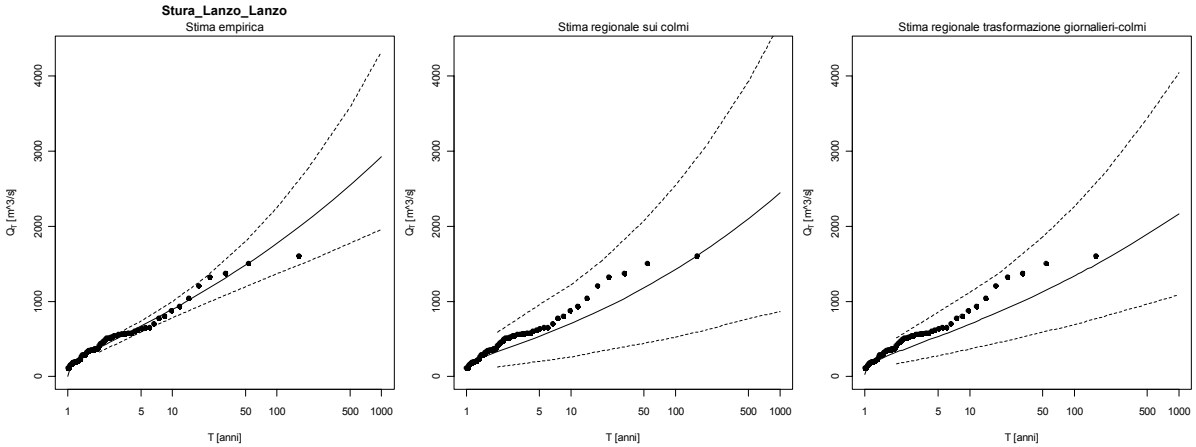

Sesia_P.te_Aranco
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

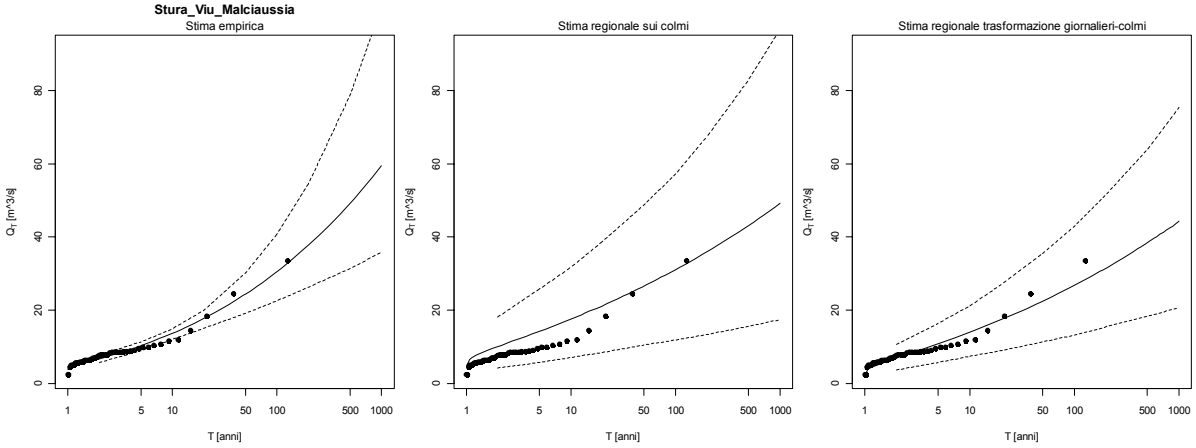

Sesia_Rimasco (Sermenza)
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1
5
10
50
100
500
1000

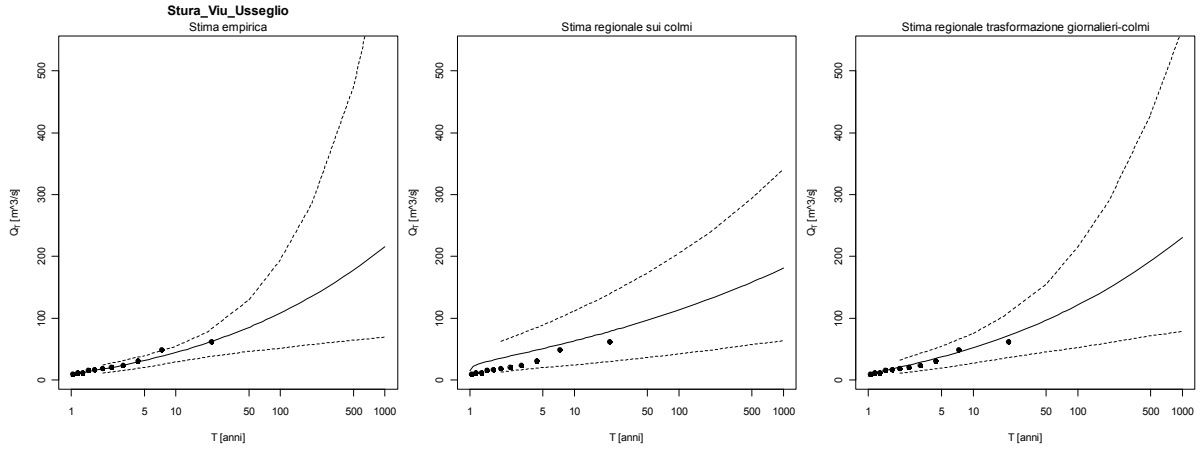

Sesia_Vercelli
Stima empirica
Stima regionale sui colmi
Q_T [m^3/s]
T [anni]
0
2000
4000
6000
8000
10000
12000
14000
1
5
10
50
100
500
1000

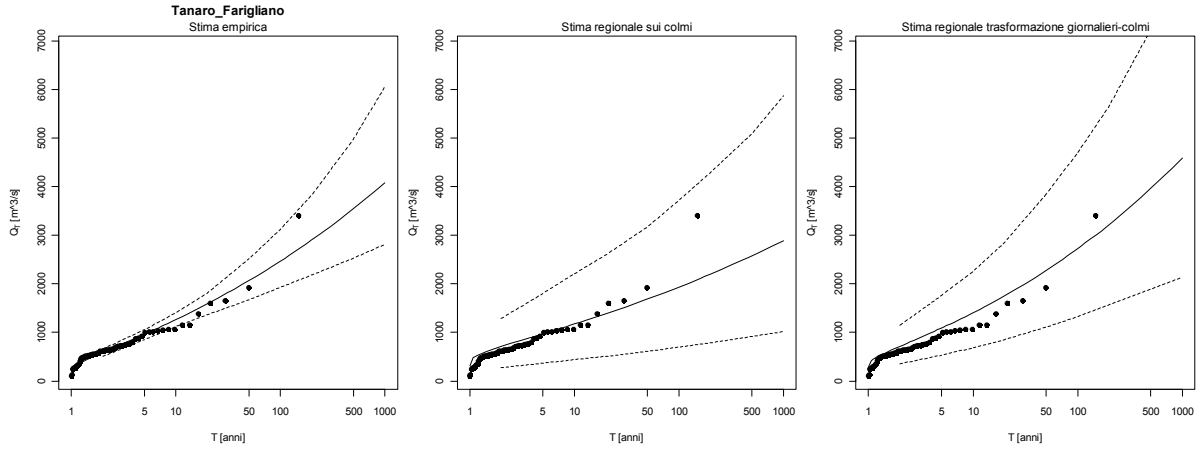

Stura_Demonte_Gaiola
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
1
5
10
50
100
500
1000

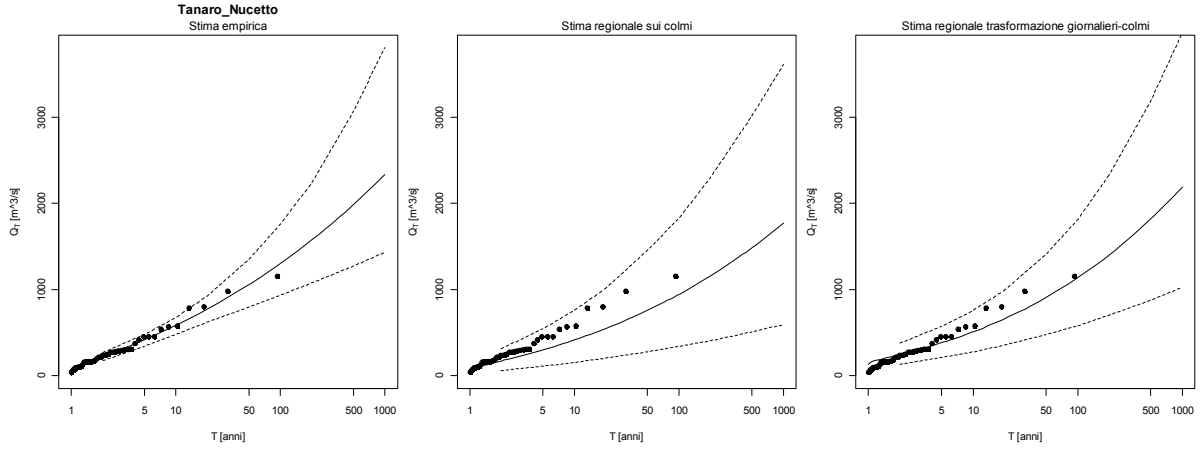

Stura_Demonte_Pianche
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1
5
10
50
100
500
1000

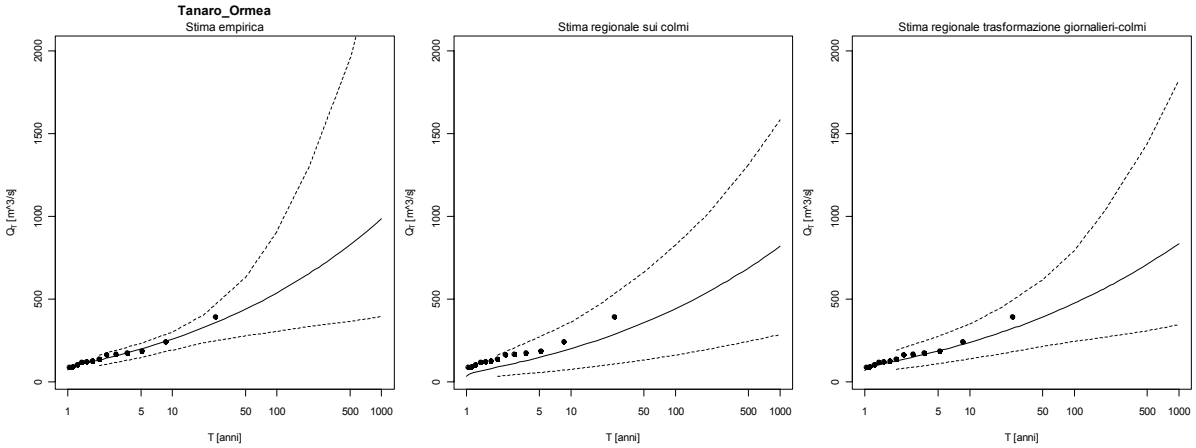

Stura_Lanzo_Lanzo
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
1
5
10
50
100
500
1000

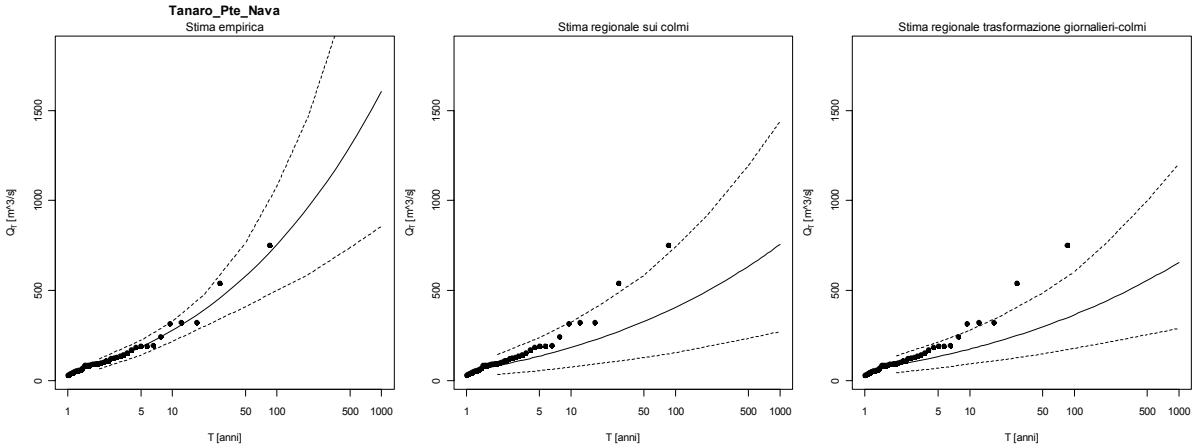

Stura_Viu_Malciaussia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
20
40
60
80
1
5
10
50
100
500
1000

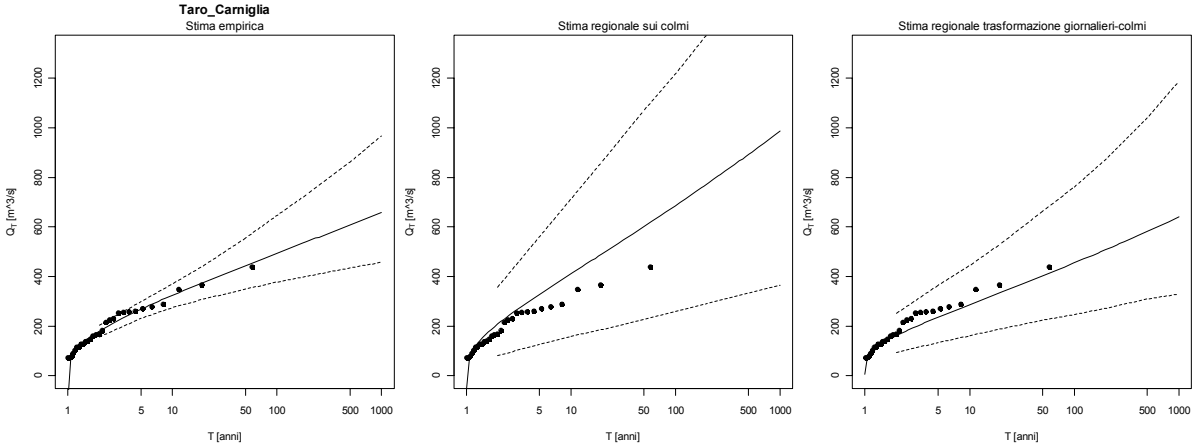

Stura_Viu_Usseglio
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
100
200
300
400
500
1
5
10
50
100
500
1000

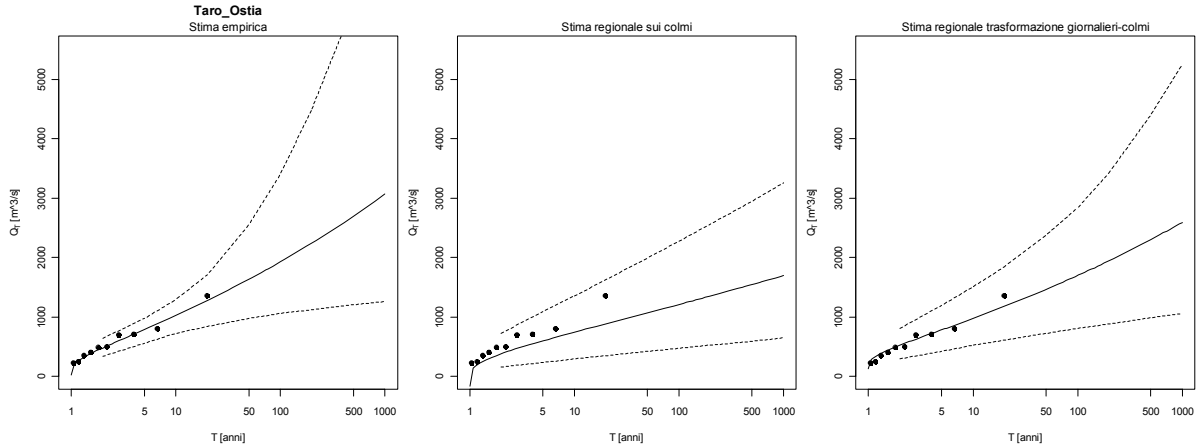

Tanaro_Farigliano
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
5000
6000
7000
1
5
10
50
100
500
1000

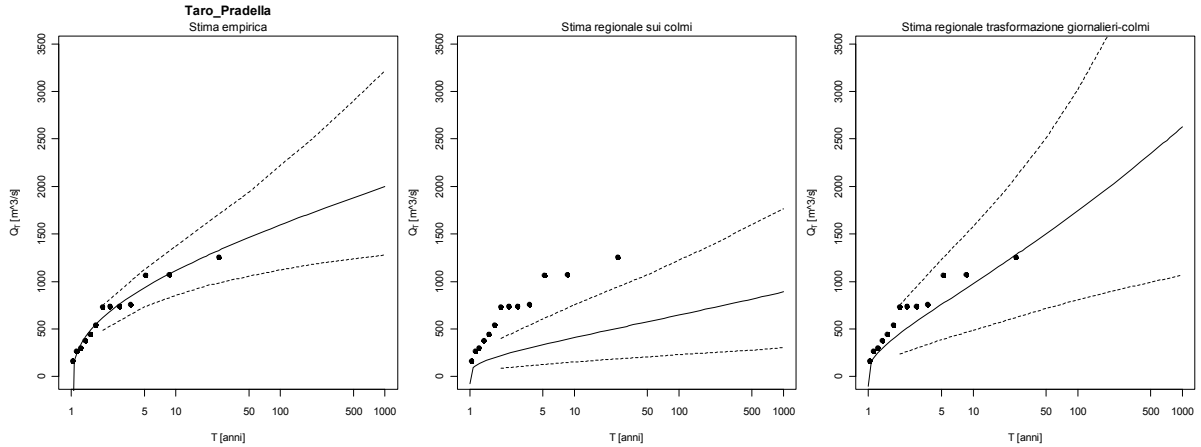

Tanaro_Nucetto
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
1
5
10
50
100
500
1000

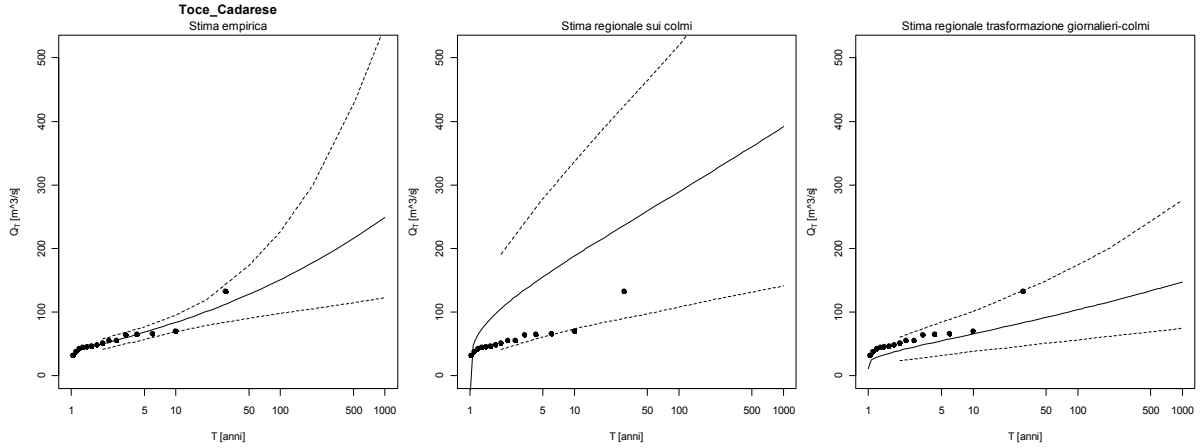

Tanaro_Ormea
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
2000
1
5
10
50
100
500
1000

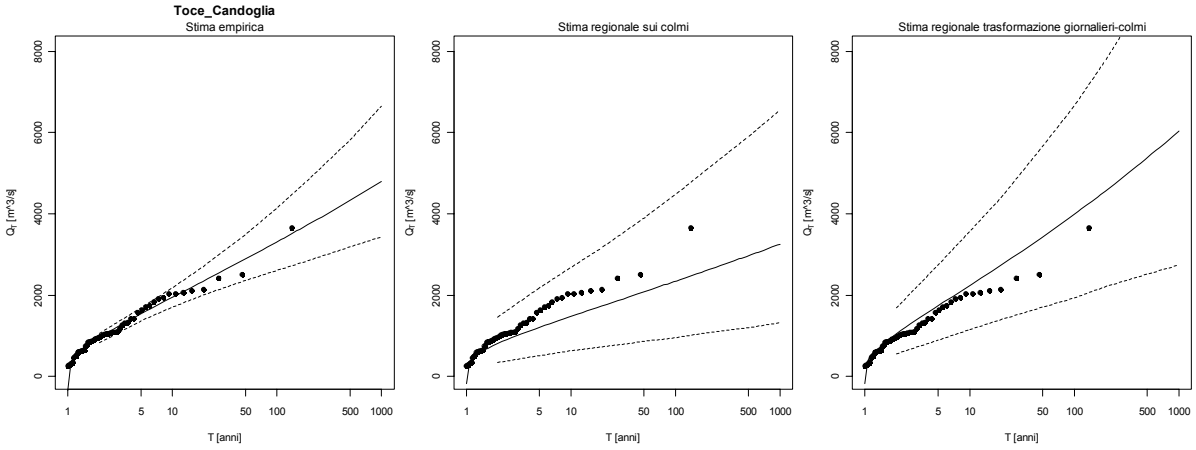

Tanaro_Pte_Nava
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
1
5
10
50
100
500
1000

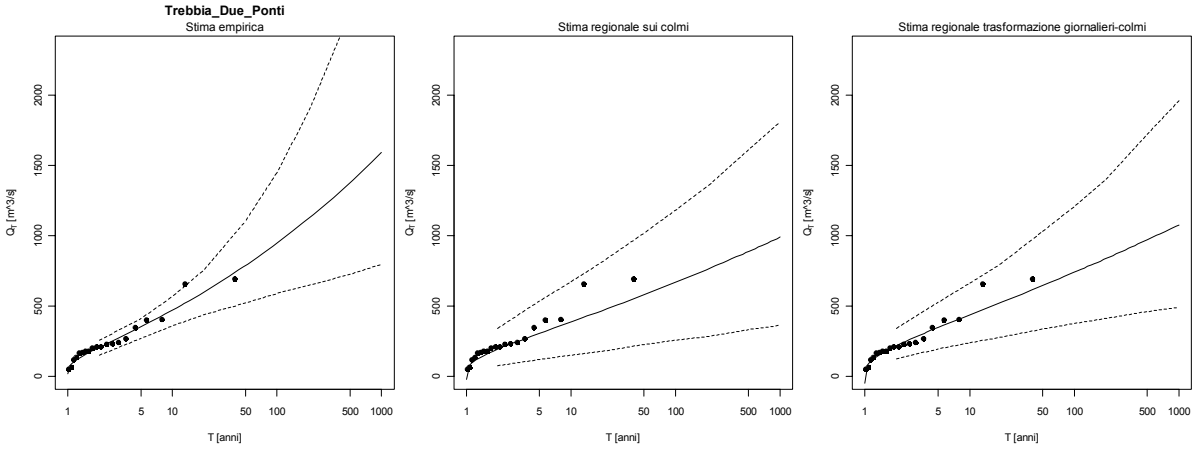

Taro_Carniglia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1000
1200
1
5
10
50
100
500
1000

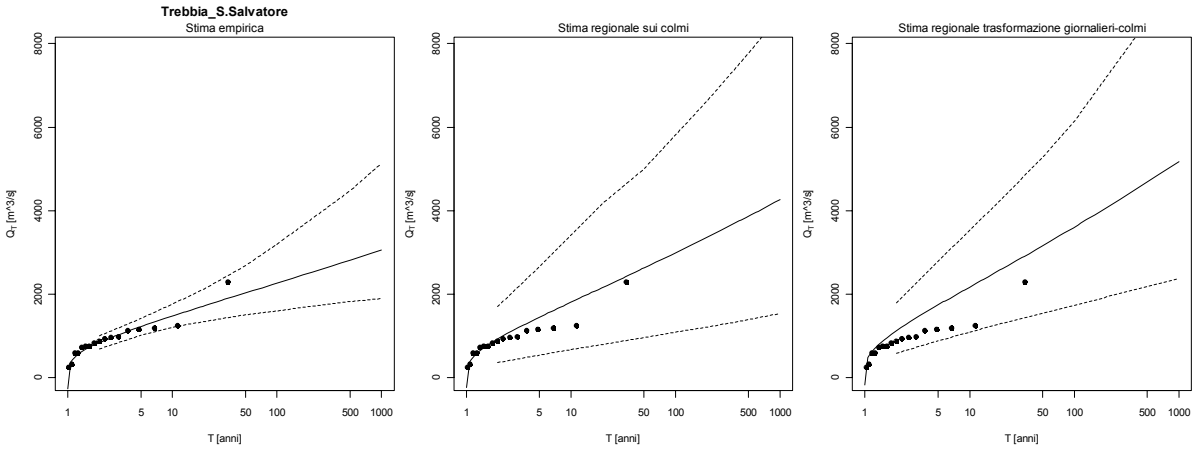

Taro_Ostia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
5000
1
5
10
50
100
500
1000

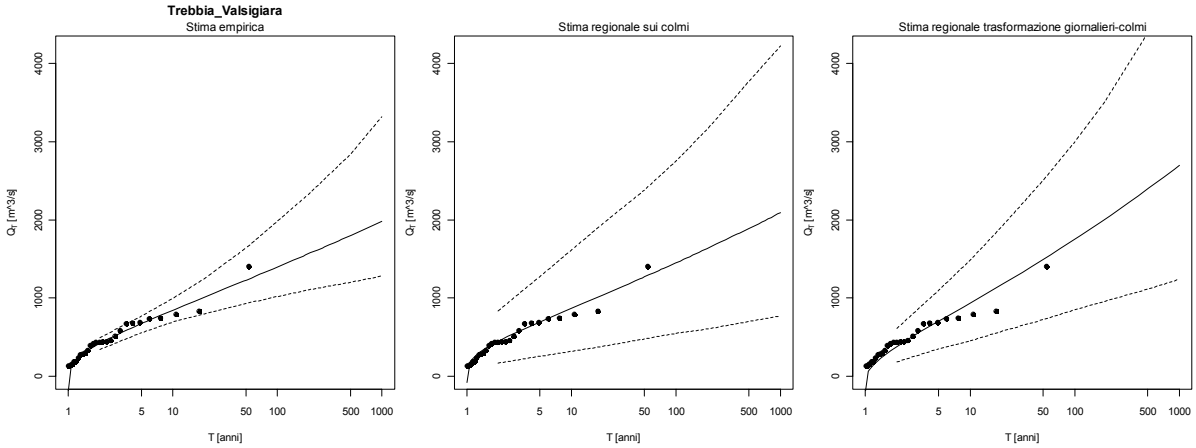

Taro_Pradella
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
500
1000
1500
2000
2500
3000
3500
1
5
10
50
100
500
1000

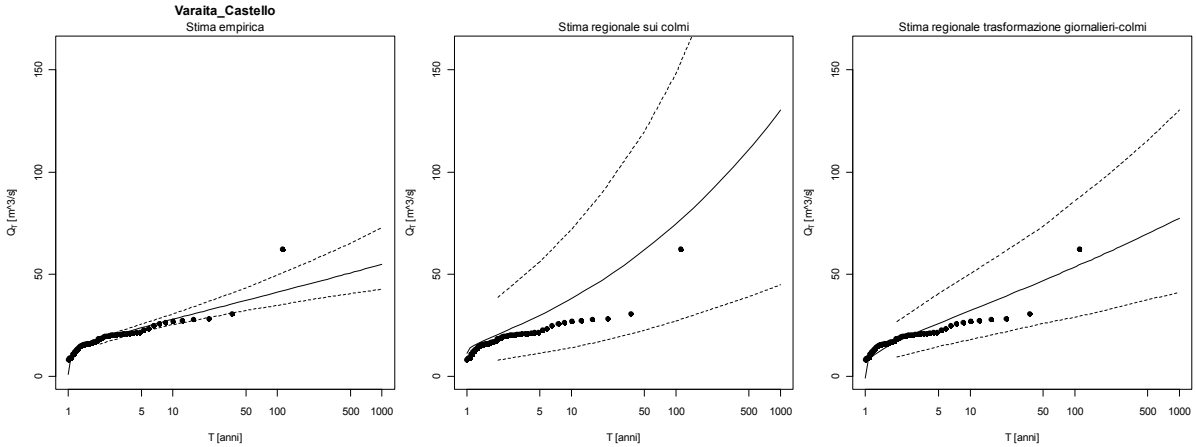

Toce_Cadarese
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
100
200
300
400
500
1
5
10
50
100
500
1000

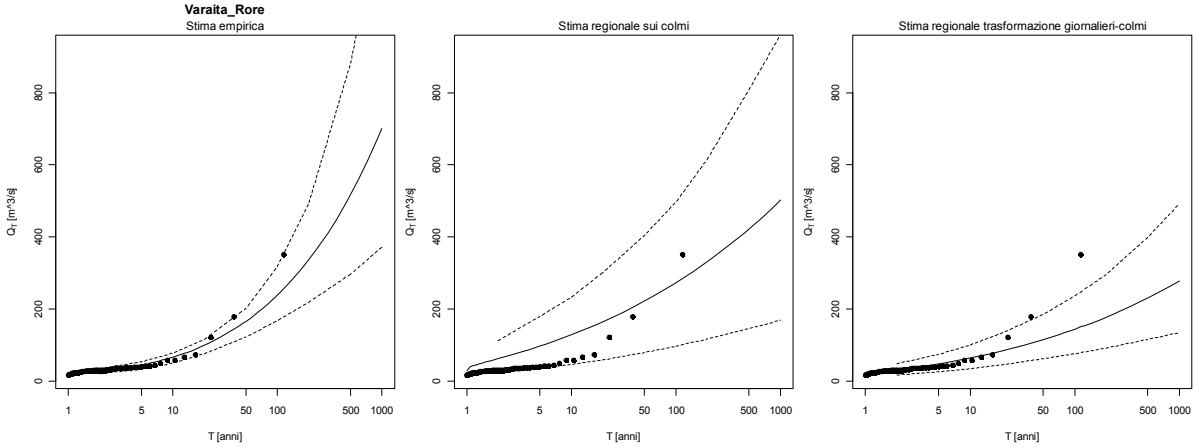

Toce_Candoglia
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

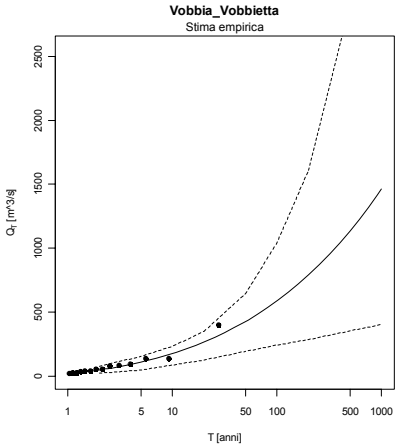

Trebbia_Due_Ponti
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

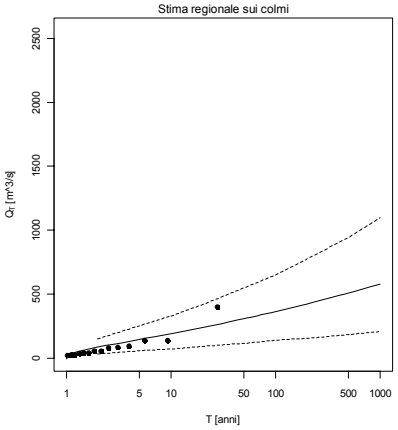

Trebbia_S.Salvatore
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

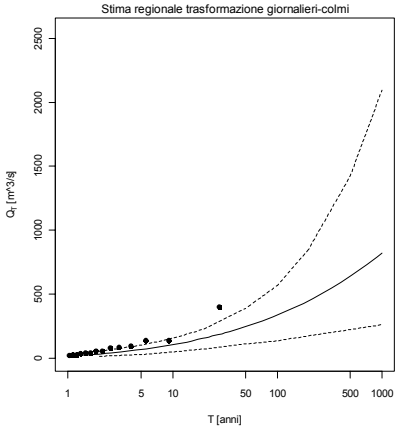

Trebbia_Valsigiara
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
1000
2000
3000
4000
1
5
10
50
100
500
1000

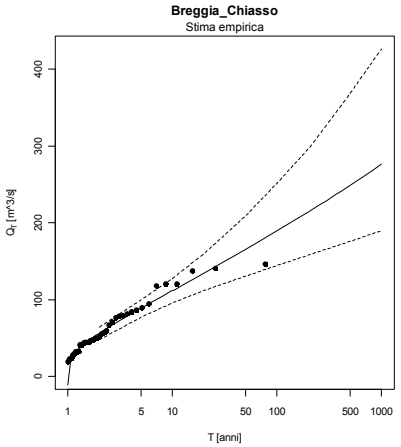

Varaita_Castello
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
50
100
150
1
5
10
50
100
500
1000

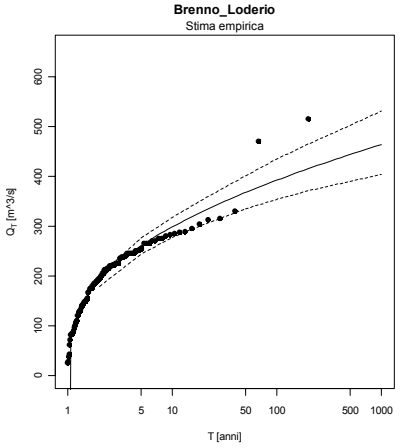

Varaita_Rore
Stima empirica
Stima regionale sui colmi
Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]
0
200
400
600
800
1
5
10
50
100
500
1000

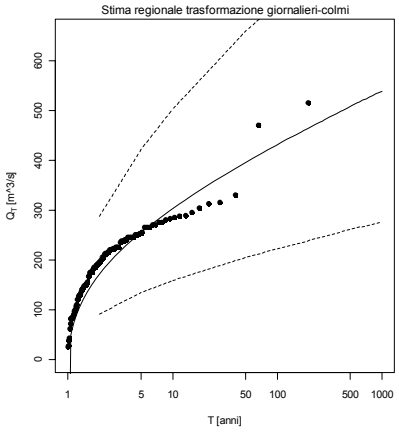

Vobbia_Vobbietta
Stima empirica
Q_T [m^3/s]
T [anni]

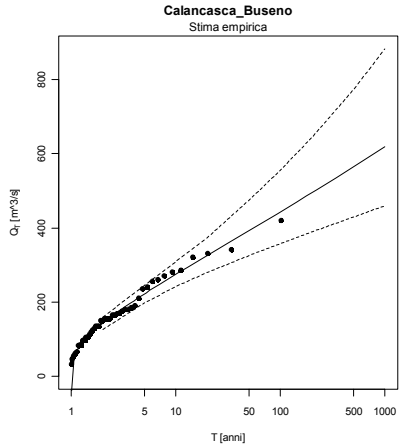

Stima regionale sui colmi
Q_T [m^3/s]
T [anni]

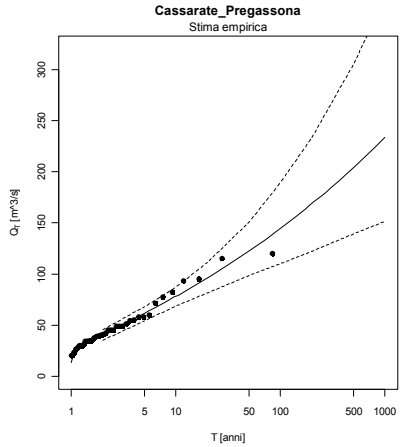

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

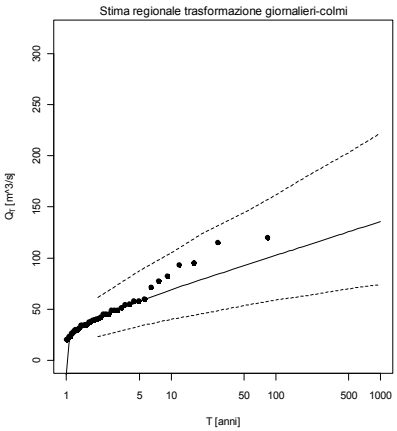

Breggia_Chiasso
Stima empirica
Q_T [m^3/s]
T [anni]

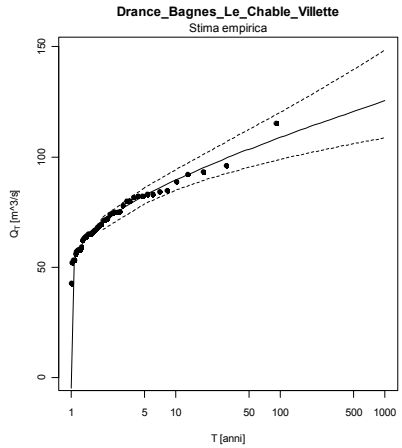

Brenno_Loderio
Stima empirica
Q_T [m^3/s]
T [anni]

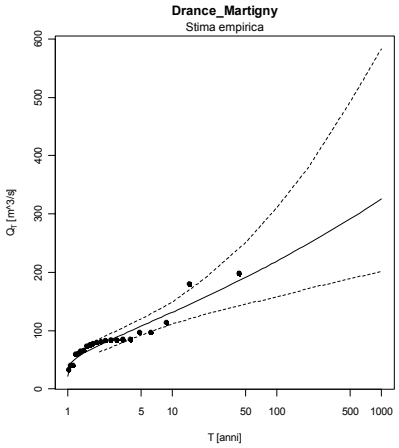

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

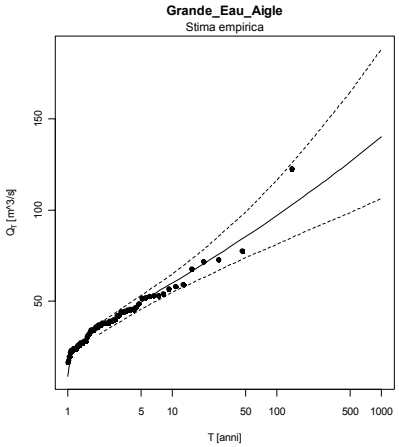

Calancasca_Buseno
Stima empirica
Q_T [m^3/s]
0
200
400
600
800
1
5
10
50
100
500
1000
T [anni]

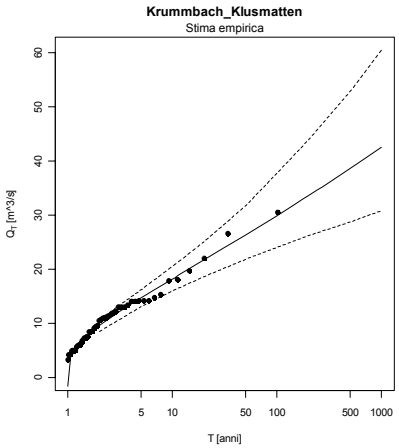

Cassarate_Pregassona
Stima empirica
Q_T [m^3/s]
0
50
100
150
200
250
300
1
5
10
50
100
500
1000
T [anni]

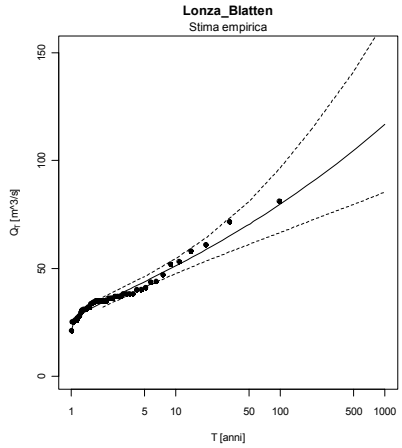

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
50
100
150
200
250
300
1
5
10
50
100
500
1000
T [anni]

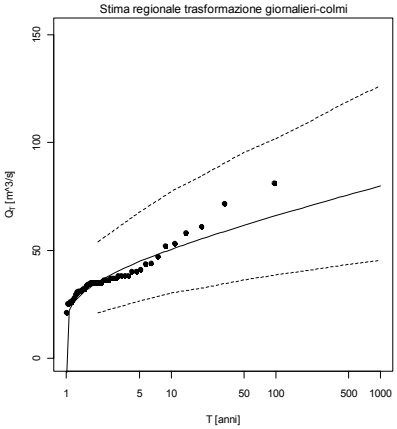

Drance_Bagnes_Le_Chable_Villette
Stima empirica
Q_T [m^3/s]
0
50
100
150
1
5
10
50
100
500
1000
T [anni]

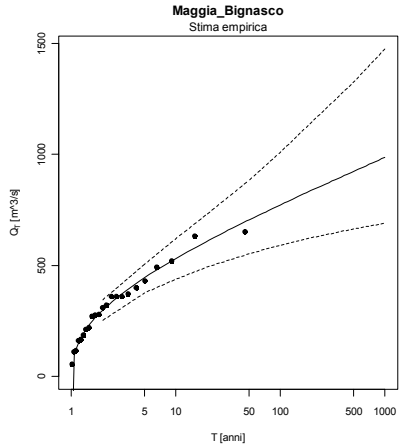

Drance_Martigny
Stima empirica
Q_T [m^3/s]
0
100
200
300
400
500
600
1
5
10
50
100
500
1000
T [anni]

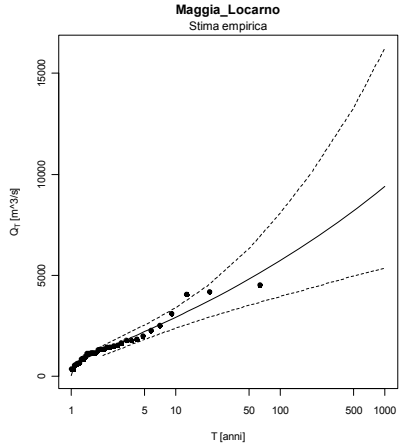

Grande_Eau_Aigle
Stima empirica
Q_T [m^3/s]
50
100
150
1
5
10
50
100
500
1000
T [anni]

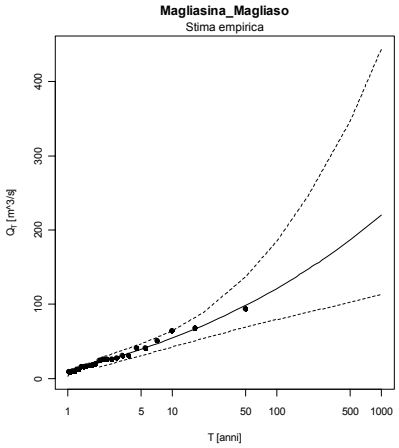

Krummbach_Klusmatten
Stima empirica
Q_T [m^3/s]
0
10
20
30
40
50
60
1
5
10
50
100
500
1000
T [anni]

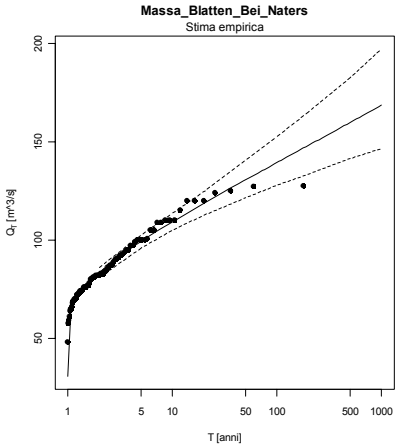

Lonza_Blatten
Stima empirica
Q_T [m^3/s]
0
50
100
150
1
5
10
50
100
500
1000
T [anni]

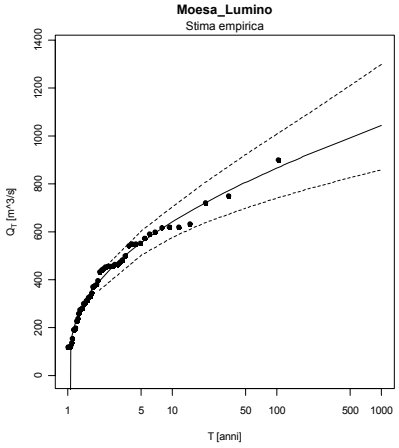

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
50
100
150
1
5
10
50
100
500
1000
T [anni]

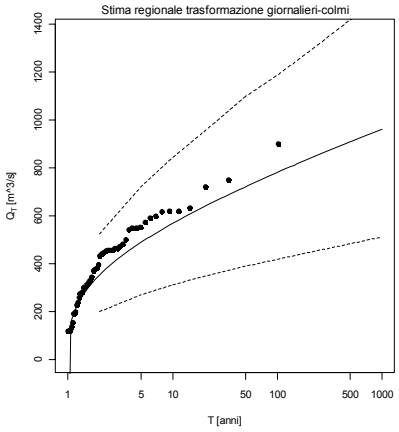

Maggia_Bignasco
Stima empirica
Q_T [m^3/s]
0
500
1000
1500
1
5
10
50
100
500
1000
T [anni]

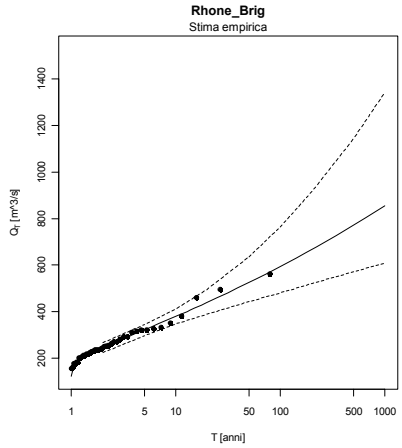

Maggia_Locarno
Stima empirica
Q_T [m^3/s]
0
5000
10000
15000
1
5
10
50
100
500
1000
T [anni]

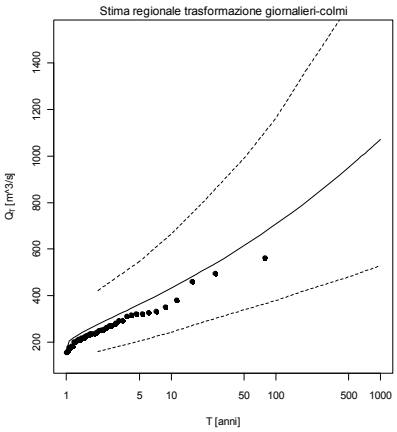

Magliasina_Magliaso
Stima empirica
Q_T [m^3/s]
0
100
200
300
400
1
5
10
50
100
500
1000
T [anni]

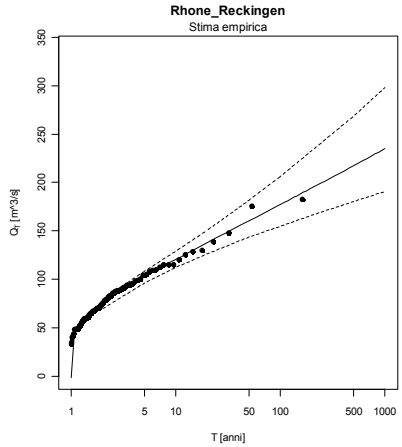

Massa_Blatten_Bei_Naters
Stima empirica
Q_T [m^3/s]
50
100
150
200
1
5
10
50
100
500
1000
T [anni]

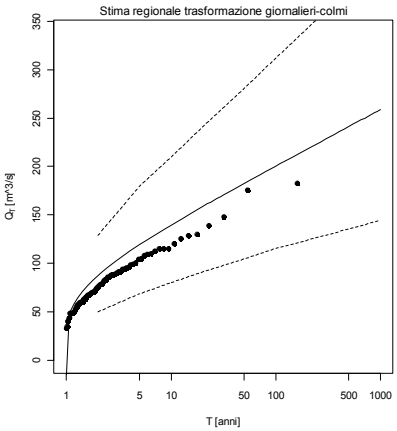

Moesa_Lumino
Stima empirica
Q_T [m^3/s]
0
200
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000
T [anni]

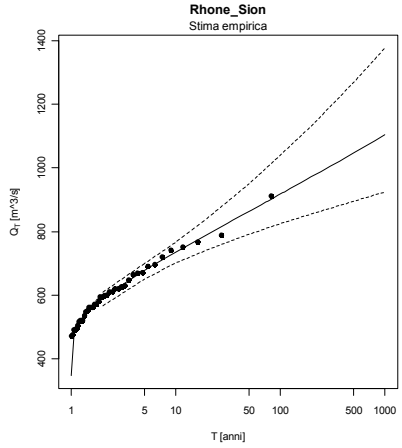

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
200
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000
T [anni]

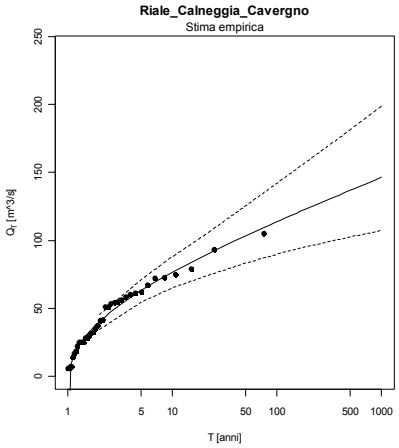

Rhone_Brig
Stima empirica
Q_T [m^3/s]
200
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000
T [anni]

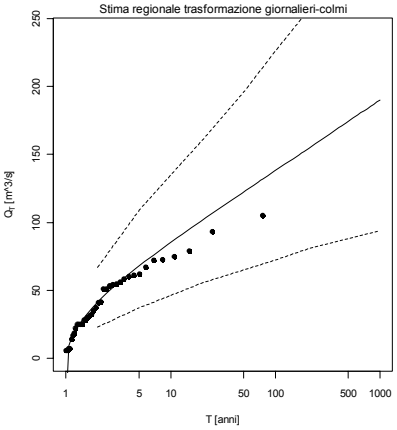

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
200
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000
T [anni]

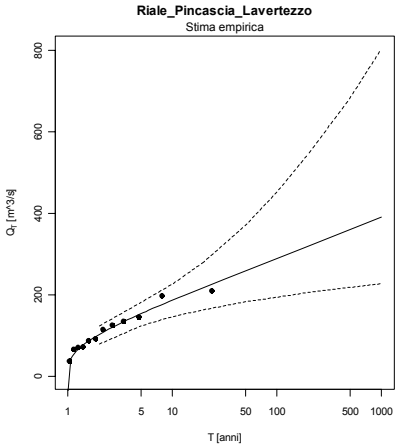

Rhone_Reckingen
Stima empirica
Q_T [m^3/s]
0
50
100
150
200
250
300
350
1
5
10
50
100
500
1000
T [anni]

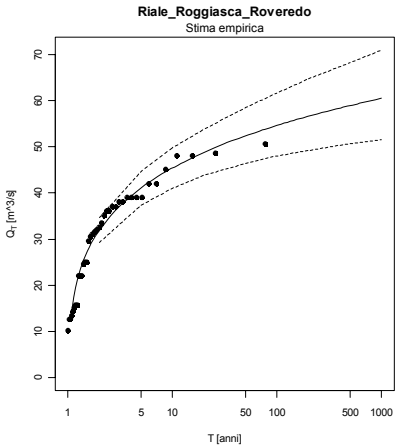

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
50
100
150
200
250
300
350
1
5
10
50
100
500
1000
T [anni]

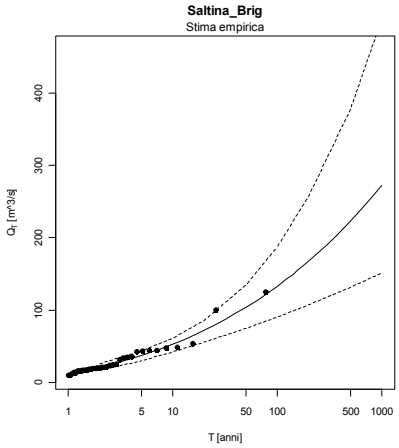

Rhone_Sion
Stima empirica
Q_T [m^3/s]
400
600
800
1000
1200
1400
1
5
10
50
100
500
1000
T [anni]

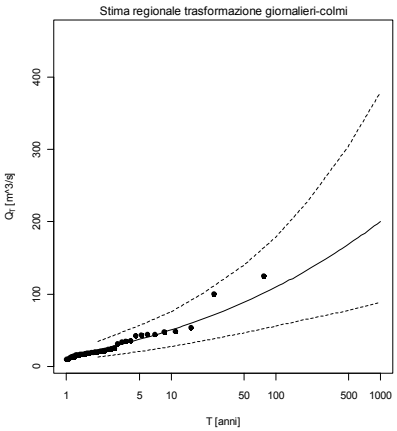

Riale_Calneggia_Cavergno
Stima empirica
Q_T [m^3/s]
T [anni]

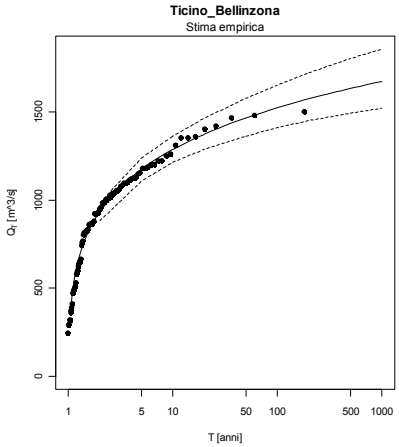

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
T [anni]

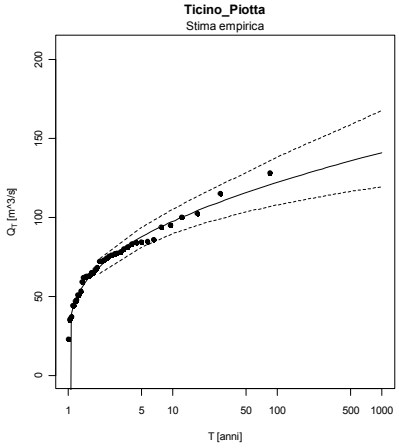

Riale_Pincascia_Lavertezzo
Stima empirica
Q_T [m^3/s]
T [anni]

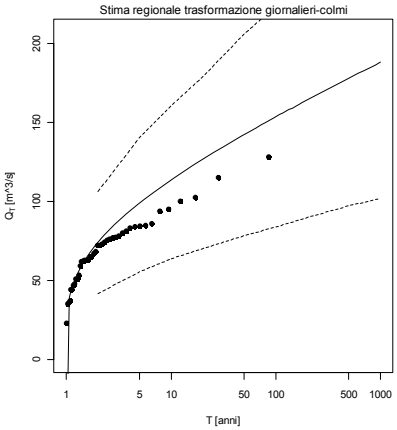

Riale_Roggiasca_Roveredo
Stima empirica
Q_T [m^3/s]
T [anni]

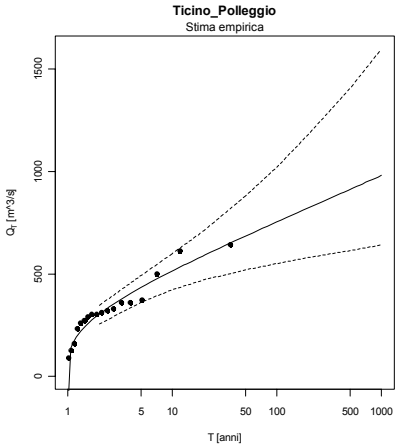

Saltina_Brig
Stima empirica
Q_T [m^3/s]
0
100
200
300
400
1
5
10
50
100
500
1000
T [anni]

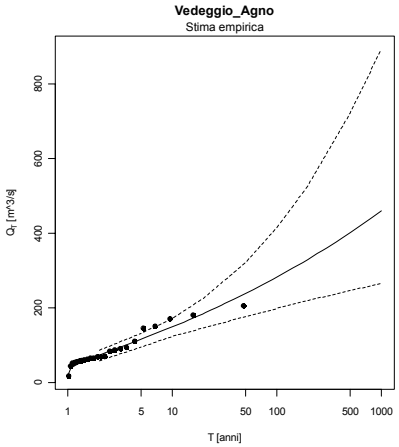

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
100
200
300
400
1
5
10
50
100
500
1000
T [anni]

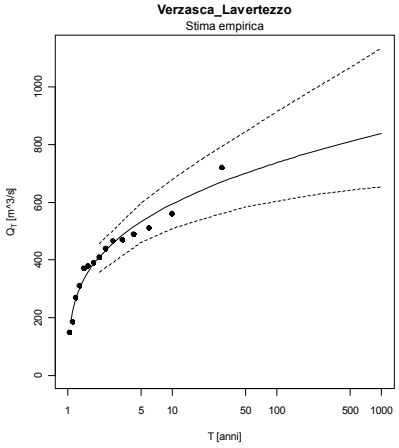

Ticino_Bellinzona
Stima empirica
Q_T [m^3/s]
0
500
1000
1500
1
5
10
50
100
500
1000
T [anni]

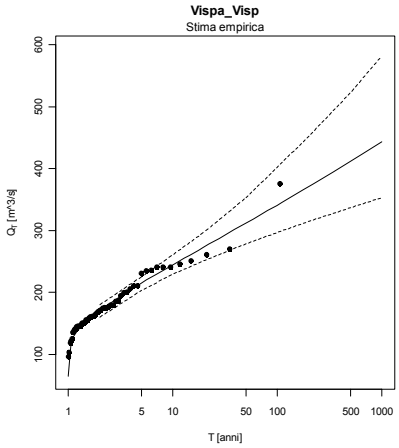

Ticino_Piotta
Stima empirica
Q_T [m^3/s]
0
50
100
150
200
1
5
10
50
100
500
1000
T [anni]

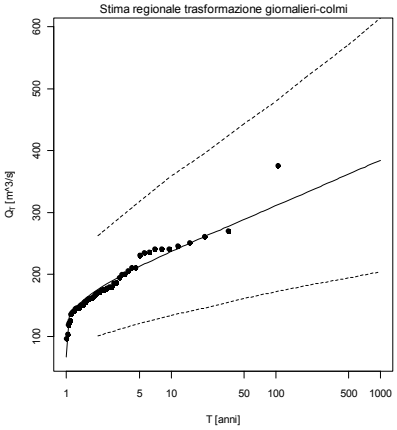

Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
0
50
100
150
200
1
5
10
50
100
500
1000
T [anni]


Ticino_Polleggio
Stima empirica
Q_T [m^3/s]
0
500
1000
1500
1
5
10
50
100
500
1000
T [anni]


Vedeggio_Agno
Stima empirica
Q_T [m^3/s]
0
200
400
600
800
1
5
10
50
100
500
1000
T [anni]


Verzasca_Lavertezzo
Stima empirica
Q_T [m^3/s]
0
200
400
600
800
1000
1
5
10
50
100
500
1000
T [anni]


Vispa_Visp
Stima empirica
Q_T [m^3/s]
100
200
300
400
500
600
1
5
10
50
100
500
1000
T [anni]


Stima regionale trasformazione giornalieri-colmi
Q_T [m^3/s]
100
200
300
400
500
600
1
5
10
50
100
500
1000
T [anni]

11 Procedura per la stima della portata di progetto

11.1 Selezione del metodo di stima

La stima della portata di progetto si effettua valutando indipendentemente la portata indice Qind e la curva di crescita K(T), che viene rappresentata in funzione dei coefficienti di L-CV ed L-CA. Determinate le due quantità, la stima della piena di progetto si ottiene tramite

$$Q_T = Q_{ind} \cdot K(T).$$

La più appropriata metodologia di stima di Q_{ind}, L-CV e L-CA nella sezione d'interesse dipende dal numero di misure disponibili e dalla loro qualità. Tra le diverse metodologie rientrano:

1. la stima empirica diretta sul campione di massimi annui al colmo di piena;
2. la stima regionale mediante modello multiregressivo;
3. la stima basata sulla trasformazione tra valori di portata massimi giornalieri a massimi al colmo di piena.

Il criterio di scelta è basato sui risultati ottenuti per tutti i bacini analizzati nello studio ed è esemplificato attraverso lo schema grafico di figura 11.1. In particolare, per ogni stazione di interesse è sufficiente entrare nel grafico con il numero n di dati di portata al colmo ed inserire in ordinata il numero di dati di massimi giornalieri n_g. La regione del piano nella quale ricade il punto così ottenuto indica il metodo di stima più adatto al set di dati a disposizione.

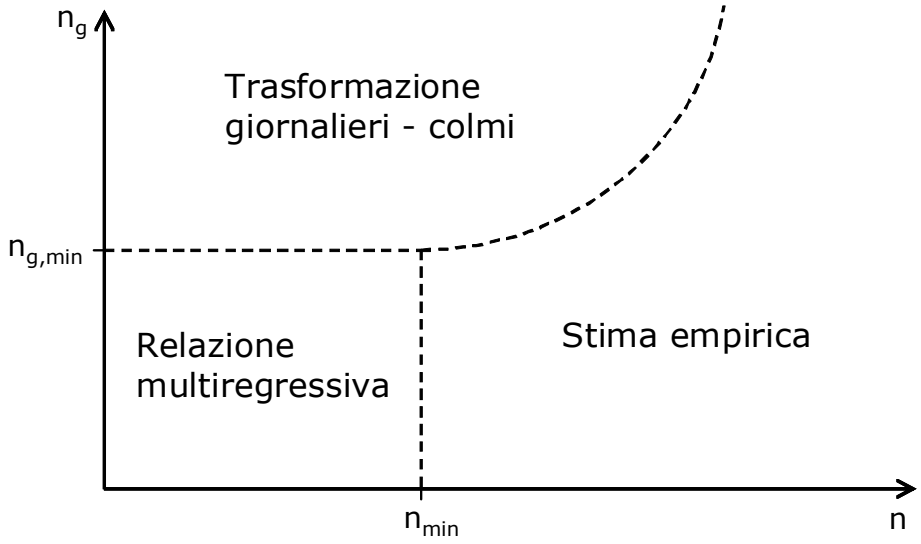


Figura 11.1. Schema per la selezione del più appropriato metodo di stima

Per la stima della portata indice lo schema di figura 11.1 viene precisato nel diagramma di figura 11.2. Per la stima di L-CV e L-CA occorre utilizzare, rispettivamente, lo schema della figura 11.3 e quello della figura 11.4. I valori (n_{min}) che definiscono il numero minimo di dati al colmo di piena per preferire una stima empirica ad una regionale ed i valori ($n_{g,min}$) che definiscono il numero minimo di dati giornalieri per cui il metodo della trasformazione giornalieri-colmi è preferibile rispetto al metodo regionale sono stati riportati nella tabella 11.1.

	n_{min}	$n_{g,min}$
Q_{ind}	2.4	1.6
L-CV	21.2	18.1
L-CA	63.4	46.9

Tabella 11.1. Soglie indicative per la selezione del più appropriato metodo di stima

I valori che determinano la scelta tra la stima empirica e la stima basata sulla trasformazione giornalieri-colmi sono invece variabili con n ed n_g, potendo essere dedotti direttamente dalla curva riportata nei grafici (vedere paragrafo 4.3 per quanto concerne la piena indice ed il paragrafo 7.5 per L-CV ed L-CA).

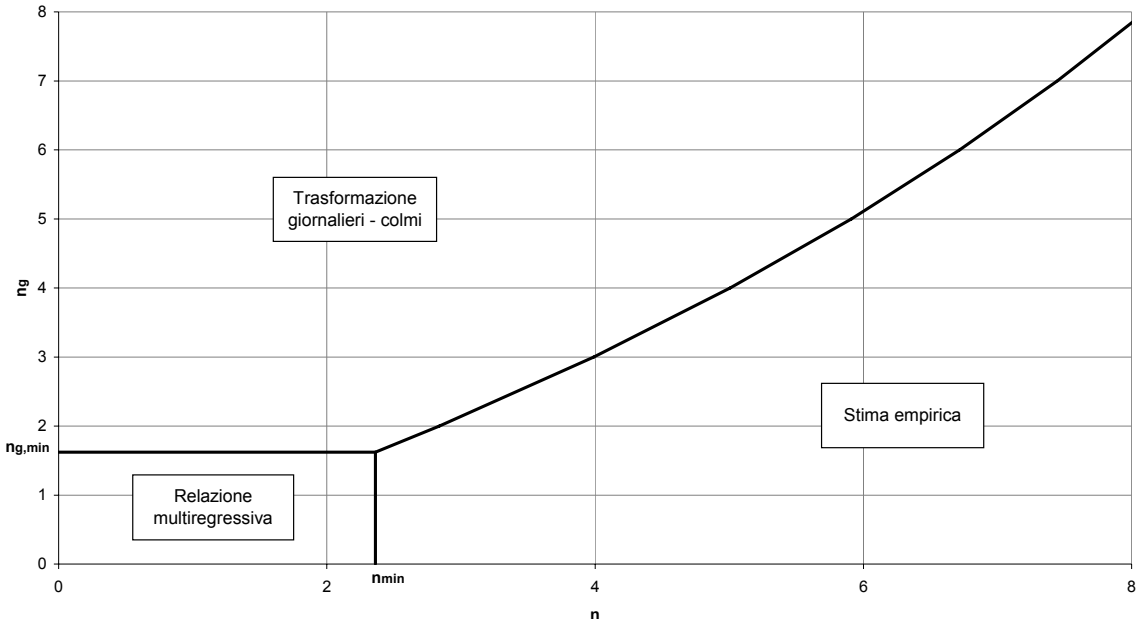


Figura 11.2. Diagramma per la scelta del metodo per la stima della portata indice

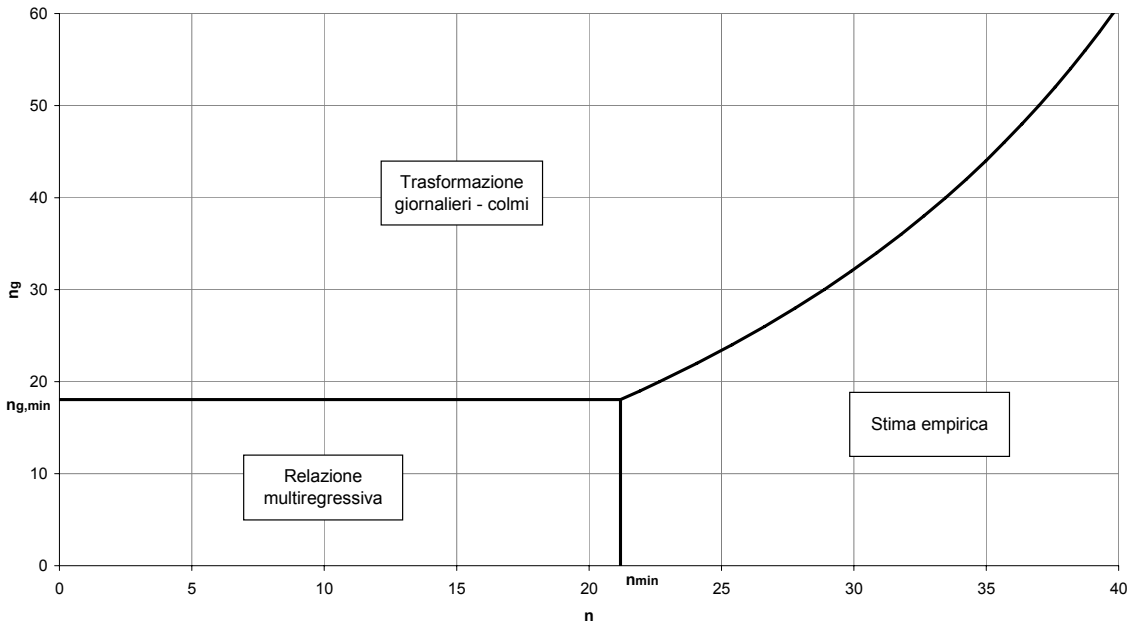


Figura 11.3. Diagramma per la scelta del metodo per la stima del L-CV

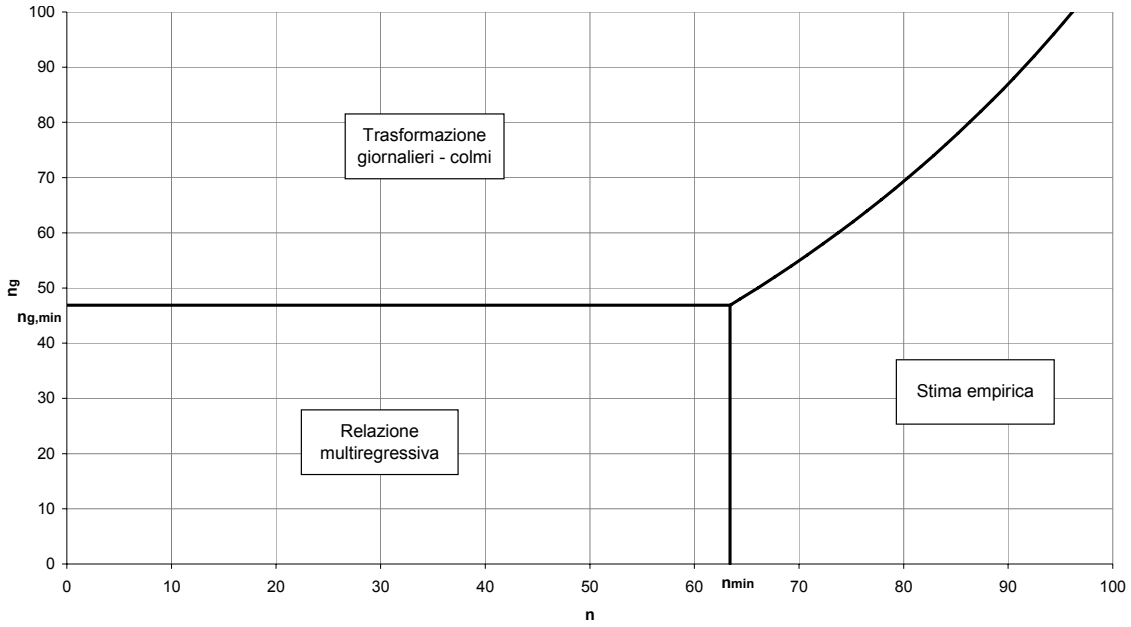


Figura 11.4. Diagramma per la scelta del metodo per la stima del L-CA

È il caso di sottolineare che gli schemi sopra riportati sono stati ricavati sulla base di valori numerici medi ottenuti su tutti i bacini utilizzati in questa analisi e sono perciò valori indicativi. Una selezione più precisa può essere fatta caso per caso, calcolando l'incertezza di stima per tutti e tre i metodi e scegliendo di conseguenza il metodo più efficiente. Tenendo conto delle scelte fatte attraverso i diagrammi delle figure 11.3 e 11.4, il metodo da utilizzare per la stima di L-CV potrebbe differire dal metodo indicato per L-CA. Di seguito vengono esposti i dettagli relativi ai diversi metodi.

11.2 Valutazione della portata indice

11.2.1 Stima empirica (v. Paragrafo 2.1.1)

La piena indice è rappresentata dalla media campionaria dei dati di portata al colmo

$$Q_{ind} = \frac{1}{n}\sum_{i=1}^{n} Q_i \ .$$

Nei casi in cui le serie storiche siano state integrate con dati relativi ad eventi alluvionali di particolare rilevanza occorsi in anni distanti dal periodo

sistematico di misurazione, ovvero si decida di utilizzare valori occasionali significativi di portata, lo stimatore della piena indice viene calcolato come:

$$Q_{ind} = \frac{\sum_{i=1}^{n_{sotto\_soglia}} Q_i}{n} + \frac{\sum_{i=1}^{n_{sopra\_soglia}} Q_i}{n_{eq}}$$

dove n rappresenta la numerosità della serie storica riferita al periodo sistematico di misurazione, mentre n_{eq} indica il periodo equivalente di osservazione, ossia la lunghezza complessiva del lasso temporale coperto dalla serie storica completa. In particolare, si fissa un valore Q_{soglia}, pari al più piccolo dei valori relativi agli eventi occasionali considerati, che permette di separare la serie storica (composta da n_{sotto_soglia} dati) dai dati occasionali (n_{sopra_soglia} valori).

11.2.2 Modello multiregressivo (v. Paragrafo 4.1.1, cap. 3 Volume II)

La stima regionale della portata indice viene effettuata con la relazione:

$$Q_{ind} = 1.526 \cdot 10^{-4} \cdot A^{0.81} \cdot a^{1.10} \cdot c_f^{0.78} \cdot aff^{0.95}$$

dove le variabili indipendenti sono:

- A [km^2], area del bacino.
- a, coefficiente (pluviale orario) della curva di possibilità pluviometrica.
- c_f, coefficiente di afflusso di piena, calcolato a partire dalla mappa dei coefficienti di afflusso di maglia 1 km x 1 km definita nel progetto VAPI Piemonte (Villani, 2003). Il valore di c_f corrisponde, quindi, alla media areale dei c_f associati alle celle ricadenti all'interno del bacino in esame, i quali sono stati definiti in relazione alla permeabilità e alla morfologia dei suoli.
- aff, afflusso totale medio annuo.

In mancanza dei dati di c_f la portata indice può essere stimata mediante la formula razionale modificata (paragrafo 4.4.2):

$$Q_{ind,j} = 1.524 \cdot 10^{-4} \cdot C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j}$$

dove $C_{H_{med},j}$ è un coefficiente dipendente dalla quota media del bacino secondo la relazione

$$C_{H_{med},j} = \ln\left(e-(e-1)\cdot\frac{H_{med,j}}{3500}\right).$$

11.2.3 Stima dagli estremi giornalieri (v. Paragrafi 3.4 e 4.1.3, cap 3 Volume II)

Quando gli estremi giornalieri sono in numero consistente (figura 11.2), note la media μ_g del campione di dati massimi giornalieri e la deviazione standard σ_{qg} dello stesso campione di dati massimi giornalieri (in questo caso divisi per l'area del bacino nella sezione considerata), si calcola il coefficiente di punta tramite l'espressione

$$c_P = 4.0694\cdot\sigma_{qg}^{0.12}\cdot\Delta H_2^{-0.17}\cdot a^{0.19}$$

che contiene due parametri morfoclimatici del bacino:

- ΔH_2 [m], differenza tra la quota media e la quota minima del bacino;
- a, coefficiente (pluviale orario) della curva di possibilità pluviometrica.

La stima della portata indice è dunque:

$$Q_{ind} = c_P\cdot\mu_g.$$

11.3 Valutazione di L-CV e L-CA (v. Cap. 6)

11.3.1 Stima empirica

Data una serie storica di valori di portata al colmo di piena (o di massimi giornalieri) è possibile calcolare gli L-momenti empirici in funzione dei momenti pesati in probabilità b_0, b_1, b_2 e b_3. In particolare risulta:

$$L_{cv} = \frac{2\cdot b_1}{b_0} - 1,$$

$$L_{ca} = \frac{2 \cdot (3b_2 - b_0)}{2b_1 - b_0} - 3 ,$$

dove L_{cv} ed L_{ca} rappresentano, rispettivamente, il coefficiente di L-variazione e di L-asimmetria.

Nel caso in cui il campione di *n* osservazioni a disposizione sia relativo a un periodo sistematico di misurazioni, il valore dei momenti pesati in probabilità (Probability Weigthed Moments, PWM) viene calcolato come:

$$b_r = \frac{1}{n} \sum_{i=1}^{n} Q_{(i)} \qquad \text{con } r=0$$

$$b_r = \frac{1}{n} \sum_{i=1}^{n} \frac{(i-1) \cdot (i-2) \cdot \ldots \cdot (i-r)}{(n-1) \cdot (n-2) \cdot \ldots \cdot (n-r)} \cdot Q_{(i)} \qquad \text{con } r=1,\ldots,3$$

dove $Q_{(i)}$ rappresenta l'*i*-esimo valore di portata del campione ordinato in senso crescente.

Qualora le serie storiche siano state integrate con informazioni storiche occasionali, si individua per essi un valore soglia Q_{soglia}, pari al più piccolo dei valori occasionali considerati e si selezionano gli n_{sotto_soglia} dati al di sotto di Q_{soglia} e gli n_{sopra_soglia} valori al di sopra di essa.

La stima dei PWM viene dunque effettuata a partire dalla somma di due contributi, come definito da *Wang* (1990):

$$b_r = b'_r + b''_r ,$$

in cui b'_r è il contributo corrispondente ai valori sotto soglia mentre b''_r è relativo ai valori di portata che superano la Q_{soglia} prefissata.

In particolare, b'_r si calcola tramite la relazione:

$$b'_r = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{(i-1) \cdot (i-2) \cdot \ldots \cdot (i-r)}{(n-1) \cdot (n-2) \cdot \ldots \cdot (n-r)} \cdot Q^*_{(i)} \qquad \text{con } r = 0,\ldots,3$$

dove

$$Q^{*}_{(i)} = \begin{cases} Q_{(i)} & Q_{(i)} < Q_{soglia} \\ 0 & Q_{(i)} \geq Q_{soglia} \end{cases} .$$

Il contributo b_r'' derivante dagli n_{sopra_soglia} eventi viene pesato sulla numerosità equivalente n_{eq}, ovvero sulla lunghezza complessiva del lasso temporale coperto dalla serie storica completa:

$$b_r'' = \frac{1}{n_{eq}} \cdot \sum_{i=1}^{n} \frac{(i-1)\cdot(i-2)\cdot\ldots\cdot(i-r)}{(n_{eq}-1)\cdot(n_{eq}-2)\cdot\ldots\cdot(n_{eq}-r)} \cdot Q^{**}_{(i)}$$

in cui $r= 0,\ldots,3$ e con

$$Q^{**}_{(i)} = \begin{cases} 0 & Q_{(i)} < Q_{soglia} \\ Q_{(i)} & Q_{(i)} \geq Q_{soglia} \end{cases} .$$

11.3.2 Stima regionale di L-CV (v. Paragrafo 7.4.1, cap. 3 Volume II)

Il coefficiente L-CV può essere stimato sulla base delle caratteristiche morfoclimatiche del bacino mediante l'espressione:

$$L_{cv} = 2.648 \cdot 10^{-1} - 8.392 \cdot 10^{-5} H_{media} - 4.314 \cdot 10^{-3} LLDP + \\ + 1.304 \cdot 10^{-2} lungh\_vett\_orient + 2.975 \cdot 10^{-1} n$$

dove:

- H_{media} [m s.l.m.], quota media del bacino.
- $LLDP$ [km], longest drainage path length. È la lunghezza del percorso tra la sezione di chiusura ed il punto più lontano da essa, posto sullo spartiacque del bacino, ottenuto seguendo le direzioni di drenaggio.
- $lungh\_vett\_orient$ [km], lunghezza del vettore orientamento, segmento che unisce il baricentro del bacino alla sezione di chiusura.
- n, esponente di invarianza di scala della curva di possibilità pluviometrica.

11.3.3 Stima regionale di L-CA (v. Paragrafo 7.4.1, cap 3 Volume II)

Il coefficiente L-CA può essere stimato sulla base delle caratteristiche morfoclimatiche del bacino mediante l'espressione:

$$L_{ca} = 3.604 - 1.144 \cdot 10^{-6} X_{sc} - 6.052 \cdot 10^{-7} Y_{sc} + 7.190 \cdot 10^{-1} L_{cv},$$

dove:

- Y_{sc} [m], latitudine della sezione di chiusura definita nel sistema di riferimento UTM ED 50.
- X_{sc} [m], latitudine del baricentro del bacino definita nel sistema di riferimento UTM ED 50.
- L_{cv}, valore del coefficiente L-CV stimato in precedenza.

11.3.4 Stima di L-CV dagli estremi giornalieri (v. Paragrafo 7.4.2, cap 3 Volume II)

Noto il valore di L-CV dal campione di dati massimi giornalieri (paragrafo 13.3.1), è possibile applicare la relazione seguente per stimare L-CV dei colmi

$$\begin{aligned} L_{cv} = {} & 2.485 \cdot 10^{-1} - 9.813 \cdot 10^{-5} A - 2.697 \cdot 10^{-3} LLDP + \\ & - 2.402 \cdot 10^{-1} R\_al + 8.734 \cdot 10^{-1} L_{cv,g} \end{aligned},$$

dove:

- A [km^2], area del bacino.
- $LLDP$ [km], longest drainage path length. È la lunghezza del percorso tra la sezione di chiusura ed il punto più lontano da essa, posto sullo spartiacque del bacino, ottenuto seguendo le direzioni di drenaggio.
- R_al [-], rapporto di allungamento, inteso come rapporto tra il diametro del cerchio di eguale area del bacino ed il valore di *LLDP*.
- $L_{cv,g}$, parametro L-CV per il campione dei massimi giornalieri.

11.3.5 Stima di L-CA dagli estremi giornalieri (v. Paragrafo 7.4.2, cap 3 Volume II)

Noto il valore di L-CA per il campione di dati massimi giornalieri (paragrafo 11.3.1) è possibile applicare la relazione seguente:

$$L_{ca} = 3.043 \cdot 10^{-1} - 2.258 \cdot 10^{-1} F\_f - 8.823 \cdot 10^{-5} aff + 7.312 \cdot 10^{-1} L_{ca,g},$$

dove:

- F_f [-], fattore di forma, corrispondente al rapporto tra l'area del bacino e il quadrato della lunghezza dell'asta principale.
- aff [mm], afflusso totale medio annuo al bacino.
- $L_{ca,g}$, coefficiente L-CA per il campione dei massimi giornalieri.

11.4 Costruzione della distribuzione di probabilità delle piene e stima della piena di progetto (v. Cap 8)

I parametri L-CV e L-CA stimati come indicato in precedenza vengono utilizzati per determinare i parametri della distribuzione di probabilità adimensionale, il cosiddetto fattore di crescita, che verrà infine moltiplicato per il valore di portata indice. La distribuzione di probabilità scelta in questo lavoro è quella lognormale, con parametri ξ (posizione), (scala) e k (forma).

Una volta fissati i parametri della distribuzione ξ, α, k è possibile calcolare il quantile con tempo di ritorno *T* della distribuzione adimensionale

$$K(T) = \begin{cases} \left. \xi + \alpha \left(1 - e^{-k\Phi^{-1}\left(1-\frac{1}{T}\right)} \right) \right/ k & k \neq 0 \\ \xi + \alpha \Phi^{-1}\left(1-\frac{1}{T}\right) & k = 0 \end{cases}$$

dove $\Phi^{-1}\left(1-\frac{1}{T}\right)$ rappresenta il quantile con tempo di ritorno *T* anni di una distribuzione normale standard. Il valore di progetto finale sarà poi pari a

$$Q_T = Q_{ind} \cdot K(T).$$

Noti i valori di L-CV e L-CA, i parametri della distribuzione lognormale possono essere calcolati mediante le seguenti espressioni:

$$k \approx -L_{ca}\frac{E_0 + E_1 L_{ca}^2 + E_2 L_{ca}^4 + E_3 L_{ca}^6}{1 + F_1\, L_{ca}^2 + F_2 L_{ca}^4 + F_3 L_{ca}^6},$$

$$\alpha = \frac{L_{cv} e^{-k^2/2} k}{1 - 2\Phi\left(-k/\sqrt{2}\right)},$$

$$\xi = 1 - \frac{\alpha}{k}\left(1 - e^{k^2/2}\right),$$

utilizzando i coefficienti riportati nella tabella 11.2 e dove Φ rappresenta la funzione cumulata della distribuzione Normale standard.

Coefficiente	Valore	Coefficiente	Valore
E_0	2.0466534	F_1	-2.0182173
E_1	-3.6544371	F_2	1.2420401
E_2	1.8396733	F_3	-0.21741801
E_3	-0.20360244		

Tabella 11.2. Coefficienti di approssimazione per le equazioni dei parametri della distribuzione lognormale.

12 Conclusioni e sviluppi futuri

La presentazione delle procedure di stima e dei relativi risultati con esse ottenuti richiede qualche commento finale, che ha lo scopo di inquadrare meglio gli aspetti innovativi del presente studio. Si elencano di seguito alcuni aspetti peculiari del lavoro svolto, per passare in seguito a segnalare alcune direzioni di eventuale ulteriore approfondimento.

1) La base di dati idrometrici utilizzati è stata sostanzialmente incrementata rispetto a studi analoghi svolti in passato. In particolare, si sono considerate complessivamente 157 stazioni idrometriche, 78 delle quali contemplano anche misure di portata al colmo di piena e sono situate in Piemonte o Valle d'Aosta (cfr. Volume II). Considerando solo queste 78 stazioni, si ha un numero complessivo di dati (massimi annui delle portate al colmo di piena) considerati pari a 1826. Per confronto, i dati riconducibili alla stessa regione (Piemonte e Valle d'Aosta) considerati da *De Michele e Rosso* (2001) sono circa 650 (il numero esatto non è riportato nei rapporti tecnici), mentre quelli utilizzati nel VAPI Piemonte (*Villani*, 2002) sono 1108 relativi a 36 stazioni idrometriche. Considerando anche le stazioni che non ricadono in Piemonte e Valle d'Aosta, il numero complessivo di massimi annui delle portate al colmo considerate in questo lavoro sale a 3250 (relativi a 114 sezioni diverse).
2) In un numero significativo di sezioni le serie storiche sono state integrate con valori estremi derivanti da misure non sistematiche (essenzialmente derivate da rapporti di evento), particolarmente importanti perché consentono di meglio ricostruire la distribuzione di probabilità delle piene nella parte corrispondente ai tempi di ritorno più elevati. L'utilizzo di questi dati non sistematici ha comportato il ricorso a specifiche metodologie statistiche (v. paragrafi 2.1.1 e 6.1.1)
3) Nelle analisi statistiche si sono considerati anche i massimi giornalieri delle portate, che sono spesso disponibili anche per sezioni dove non si hanno dati di portata al colmo di piena, o dove questi ultimi sono poco numerosi. In particolare, utilizzando tali informazioni aggiuntive si sono elaborate tecniche per stimare la piena indice ed i parametri della curva di crescita.

4) Per consentire la stima delle portate di progetto in sezioni non strumentate è essenziale disporre di una ampia base dati anche relativamente a descrittori morfologici, geo-pedologici e climatici dei bacini idrografici. Nel presente lavoro sono stati utilizzati 40 diversi descrittori (cfr. Volume II). Si è posta particolare attenzione ad utilizzare descrittori morfologici che non siano significativamente influenzati dalla risoluzione del modello digitale di terreno utilizzato per la loro determinazione.

5) Ai fini della stima della curva di crescita si è utilizzato un metodo che non richiede la suddivisione dell'area di interesse in zone (regioni) supposte omogenee. Ciò allo scopo di evitare problemi relativi a brusche discontinuità della curva di crescita nel passaggio da una zona ad un'altra e per ridurre al minimo gli elementi di soggettività dell'analisi (che sono ad esempio relativi alla definizione del numero di sottoregioni da considerare). Il metodo qui sviluppato consente la stima sia della piena indice che degli L-coefficienti L_{cv} e L_{ca} a partire dal campione di osservazioni a disposizione, oppure da modelli di regressione, oppure ancora dai corrispondenti L-coefficienti stimati sugli estremi giornalieri. Quale dei tre metodi sia conveniente utilizzare dipende dalla numerosità delle serie di valori al colmo e giornalieri disponibili nelle diverse sezioni.

Dall'analisi dei diagrammi riportati al Capitolo 11 risulta che la stima regionale della curva di crescita basata su modelli multiregressivi è conveniente quando si hanno meno di 60 anni di misure disponibili nella sezione di interesse. Tale valore è stato desunto mettendo a confronto la varianza dello stimatore regionale e di quello empirico, cioè direttamente calcolato dai dati. Il valore così ottenuto potrebbe a prima vista sembrare troppo ridotto, ossia potrebbe sembrare strano che con 65-70 anni di dati sia conveniente stimare le portate di progetto utilizzando i dati empirici, invece del metodo regionale, specie se ci si riferisce a periodi di ritorno elevati. Tuttavia una attenta considerazione del modo in cui il numero limite (60 anni) è stato ricavato consente di meglio comprenderne il senso: la varianza di stima delle portate di piena di progetto, quale che sia il metodo di stima, può essere vista come la somma di due termini, relativi all' errore di modello (errore nella scelta del modello probabilistico) ed all'errore dei parametri (errore nella stima dei parametri del modello prescelto). Le stime empiriche qui utilizzate non possono considerarsi in tutto e per tutto equivalenti a stime "at site", cioè dipendenti univocamente dai dati di misura, perché la scelta della distribuzione di probabilità (Capitolo

8) è stata fatta ricorrendo ad un'analisi "regionale". Questo ha comportato la scelta di un'unica distribuzione di probabilità per rappresentare la curva di crescita in tutti i bacini. In un'analisi "at site", infatti, anche la scelta della distribuzione sarebbe invece dipesa solo dal campione ivi disponibile. Questa scelta ha consentito di ridurre l'errore di modello nel caso della stima empirica, riducendo di conseguenza il numero minimo di dati che rende conveniente la stima empirica rispetto a quella regionale. Inoltre è importante ricordare che la regione geografica analizzata presenta caratteristiche di notevole variabilità idrologica, cosa che necessariamente rende più difficili le stime regionali e quindi, indirettamente, rende comparativamente più vantaggiose le stime empiriche.

Dalle analisi qui effettuate risulta un quadro di conoscenze e nuove metodiche di stima che produce miglioramenti nelle modalità di valutazione delle piene in Piemonte. Ovviamente, data l'intrinseca complessità dell'argomento trattato, rimangono cospicui margini di miglioramento delle procedure, ferma restando l'opportunità di riconsiderarle in caso di apprezzabile incremento della disponibilità di dati. In particolare, i seguenti punti richiedono a nostro avviso ulteriori affinamenti:

1) Dai risultati ottenuti nel presente rapporto risulta evidente il fondamentale ruolo delle misure di portata, ai fini del miglioramento delle stime delle piene di progetto. Ad esempio, appena 2-3 anni di dati sono sufficienti per avere una stima della piena media di qualità comparabile a quella derivante dall'applicazione dei metodi regionali. Al di là dell'ovvia necessità di incrementare ove possibile la base dati disponibile, si pone però anche l'eventualità che le stime empiriche basate su serie molte brevi non siano rappresentative; ad esempio, si potrebbe verificare una situazione in cui i due o tre anni disponibili siano molto più secchi (o più umidi) dell'anno medio, cosa che indurrebbe una distorsione nella stima della piena indice. Bisognerà verificare se questo effetto risulta significativo, e, in tal caso, predisporre metodi per la correzione della distorsione, metodi ad esempio basati sul confronto con serie più lunghe disponibili in stazioni limitrofe.
2) Più in generale, sembra di interesse la definizione di un nuovo impianto metodologico per la stima delle portate di piena in siti prossimi a stazioni di misura delle portate. In queste sezioni non è corretto operare stime

"locali", ma è d'altra parte ragionevole considerare in modo appropriato la somiglianza con le condizioni che si verificano nelle stazioni monitorate più vicine. L'obiettivo generale è quello di ricercare le modalità di variazione delle piene al colmo lungo le aste fluviali per poi pervenire a criteri di corretta approssimazione in sezioni prossime a siti con dati.

3) Infine, significativi miglioramenti potranno a nostro avviso derivare dalla messa a punto di metodologie per la combinazione di stime delle portate di progetto derivanti da metodi diversi, ad esempio metodi regionali e metodi locali, laddove la variabilità di stima rimanga particolarmente elevata (per esempio per T grandi).

Bibliografia

Ahmad M., Sinclair C., Werritty A., 1988. Log-logistic flood frequency analysis. Journal of Hydrology, 98, pp. 215–224.

Alfieri L., Laio F., Claps P., 2008. A simulation experiment for optimal design hyetograph selection, *Hydrological Processes* 22(6): 813-820.

Allamano P., Claps P., Laio F., 2008. An analytical model of the effects of catchment elevation on the flood frequency distribution, *Water Resources Research*, under revision.

Barbero S., Rabuffetti D, Zaccagnino M., 2006. Valutazione del fattore di riduzione areale delle precipitazioni per la Regione Piemonte, *Atti del XXX Convegno di Idraulica e Costruzioni Idrauliche*, Roma, 2006.

Beran M., Hosking J.R.M., Arnell N., 1986. Comment on "Two Component Extreme Value Distribution for Flood Frequency Analysis", *Water Resource Research*, vol.22, n.2, pp. 263 – 266.

Bocchiola D., De Michele C., Rosso R., 2003. Review of recent advances in index flood estimation. *Hydrology and Earth System Sciences*, 7(3), 283-296.

Chow V., 1954. The log-probability law and its engineering applications. In *ASCE*, 80, pp. 536–1–536–25.

Ciaponi C. e Moisello U., 1988. Un metodo per la stima delle portate al colmo con area minore di 100 km^2 attraverso le portate giornaliere, *L'energia elettrica*, n°5.

Claps P., F. Laio, M. Zanetta M., 2007. Aggiornamento delle procedure delle procedure di valutazione delle piene in Piemonte, con particolare riferimento ai bacini sottesi da invasi artificiali. Working-Paper 2007-01, Dipartimento di Idraulica, Trasporti ed Infrastrutture Civili, Politecnico di Torino.

Cox D., Hinkley D., 1974. Theoretical statistics. Chapman and Hall, London.

Dalrymple T., 1960. Flood Frequency Analysis. U.S. Geological Survey. *Water Supply Paper*, 1543-A.

De Michele C., Kottegoda N. T., and Rosso R., 2001. The derivation of areal reduction factor of storm rainfall from its scaling properties, *Water Resources Research*, Vol.37, n.12, 3247-3252.

De Michele C., Rosso R., (a cura di), 2001. Sintesi del Rapporto sulla Valutazione delle Piene Italia Nord Occidentale. Estratto dal *Rapporto Nazionale VAPI 2000 con aggiornamenti*, 4.3-4.32.

Eagleson PS., 1972. Dynamics of flood frequency. *Water Resources Research* 8: 878–898.

Fiorentino M., Gabriele S., Rossi F., Versace P., 1987a. Hierarchical approach for regional flood frequency analysis. In *Regional Flood Frequency Analysis*. A cura di Singh V., pp. 35–49. Reidel, D., Norwell, Mass..

Fiorentino M., Rossi F., Villani P., 1987b. Effect of the basin geomorphoclimatic characteristics on the mean annual flood reduction curve. Proc. Of the 18th Annual Pittsburgh Conference on Modeling and Simulation, 18(5); 1777-1784.

Furcolo P., Rossi F., Villani P., 1998. Regional geostatistical analysis of very extreme rainfall and floods, Abstract, *Proceedings XXIII EGS General Assembly, Annales Geophysicae*, suppl.II to v.16, Part II, p.C464.

Georgopoulos P. G., Seinfeld J. H., 1982. Statistical distributions of air pollutant concentrations. *Environmental Science & Technology*, 16 (7), pp. 401A–416A.

Giandotti M., 1934, Previsione delle piene e delle magre dei corsi d'acqua. *Memorie e studi idrografici*, Pubbl. 2 del Servizio Idrografico Italiano, vol. VIII, pp. 107.

Gioia A., Iacobellis V., Margotta M.R., 2004. Una Rassegna sulla piena indice su base geomorfoclimatica, *Atti del Workshop "Modelli Matematici per la simulazione di Catastrofi Idrogeologiche"*, Rende (CS).

Greenwood J., Landwehr J., Matalas N., Wallis J., 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form, *Water Resources Research*, 15, pp. 1049–1054.

Gumbel E., 1944. Ranges and midranges. *Annals of Mathematical Statistics*, 15, pp. 414–422.

Gumbel E.,1958. Statistics of Extremes. *Columbia University Press*, New Jork.

Hosking J., Wallis J., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. *Cambridge University Press*.

Hydrology Subcommittee of the Advisory Committee on Water Data, 1982. Bulletin 17B, Guidelines For Determining Flood Flow Frequency.

Jenkinson A., 1955. The frequency distribution of the annual maximum (or minimum) value of meterological elements. *Quarterly Journal of the Royal Meteorological Society*, 81, pp. 158–171.

Kjeldsen, T. R., and D. A. Jones, 2006. Prediction uncertainty in a median-based index flood method using L moments, *Water Resour. Res.*, 42, W07414, doi:10.1029/2005WR004069.

Kottegoda N.T., Rosso R., 1998. Statistics, probability and reliability for civil and environmental engineers, *Mc–Graw-Hill*, New York.

Merz R., Bolschl G., Parajka J., 2006. Spatio – temporal variability of event runoff coefficient, *Journal of Hydrology*, 331 (2006) 591 – 604.

Moisello U. e Papiri S., 1986. Relazione tra altezza di pioggia puntuale e ragguagliata, *Atti del XX Convegno di Idraulica e Costruzioni Idrauliche*, Padova, 1986.

Montgomery D.C. e Ranger C., 2003. Applied statistics and probability for engineers, *Wiley and Sons.*

Pickands J., 1975. Statistical inference using extreme order statistics. *Annuals of Statistics*, 3, pp. 491–496.

Robson A., Reed D., 1999. Flood Estimation Handbook (Procedures for flood frequency estimation), *Institute of Hydrology*, Wallingford.

Rossi F., Villani P., 1994. A project for regional analysis of flood in Italy. *Coping with floods*, (Rossi G., Harmancioglu N. and Yevjevich V. eds), 227-251.

Stedinger, J.R., and G.D. Tasker, 1985. Regional Hydrologic Analysis, 1. Ordinary, Weighted and Generalized Least Squares Compared, *Water Resour. Res.*, 21(9), 1421-1432.

Stedinger, J. R., and Tasker, G. D., 1986, Regional Hydrologic Analysis 2, Model-Error Estimators, Estimation of Sigma and Log-Pearson Type 3 Distributions. *Water Resources Research*, Vol. 22(10), pp. 1487-1499.

Viglione A., 2007. A simple method to estimate variance and covariance of sample L-CV and L-CA, *Internal report, Politecnico di Torino* (disponibile alla pagina web http://www.idrologia.polito.it/~alviglio/).

Viglione A., Laio F. and Claps P., 2007. A comparison of homogeneity tests for regional frequency analysis, *Water Resources Research*, 43, W03428, doi:10.1029/2006WR005095.

Villani P. (a cura di), 2003. Rapporto sulla Valutazione delle piene in Piemonte, in *Relazione delle attività del CUGRI fino al 2001*, 89-118, Ed. Del Paguro, Fisciano (ISBN 88-87248-35-4).

Wang Q.J., 1990. Unbiased estimation of probability weighted moments and partial probability weighted moments from systematic and historical flood information and their application to estimating the GEV distribution, *Journal of Hydrology*, 120 (1990) 115-124.

Appendice A. Modelli di regressione semplice e multipla

L'analisi regressiva è una tecnica statistica per analizzare e modellare la relazione esistente tra variabili. Nei casi in cui si voglia stabilire il legame tra due sole grandezze *x* e *y*, si parla di regressione lineare semplice; al contrario quando l'obiettivo è esprimere la variabile dipendente *y* in funzione di più variabili esplicative il modello prende il nome di regressione lineare multipla.

A.1. Regressione lineare semplice

Il modello di regressione lineare semplice è quello in cui vi è una sola variabile esplicativa ***x*** legata alla variabile dipendente ***y*** da una relazione che, geometricamente, è una linea retta (Fig.A.1). Il modello viene solitamente espresso come

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x} + \boldsymbol{\varepsilon} \tag{A.1}$$

dove l'intercetta β_0 e la pendenza β_1 sono costanti incognite (coefficienti della regressione) ed $\boldsymbol{\varepsilon}$ rappresenta il vettore dei residui con la componente di errore casuale. Su quest'ultima vengono fatte alcune ipotesi: ovvero che il suo valore atteso sia nullo ($E(\boldsymbol{\varepsilon}) = 0$) e che i valori che assume siano tra loro incorrelati.

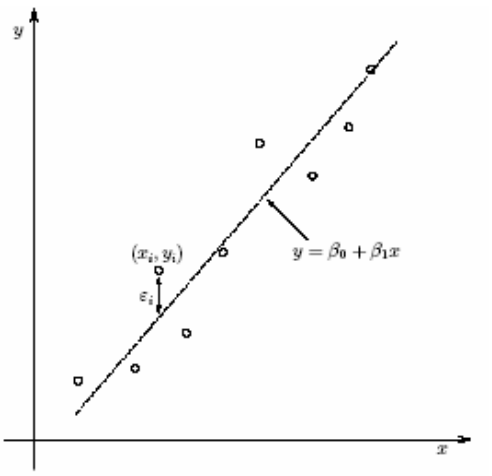


Figura A.1. Esempio di modello lineare semplice

Se si suppone di avere *s* osservazioni indipendenti della coppia (*x*, *y*), il metodo dei minimi quadrati ordinari (Ordinary Least Squares, O.L.S.) fornisce una stima dei parametri β_0 e β_1 tale che la somma dei quadrati delle differenze tra le *s* osservazioni y_j ed i corrispondenti valori stimati sia minima. Questo si traduce nella minimizzazione della funzione:

$$S\left(\beta_0,\beta_1\right)=\sum_{j=1}^{s}\left(y_j-\beta_0-\beta_1 x_j\right)^2 . \qquad \text{(A.2)}$$

Cercando gli zeri delle derivate della funzione *S* rispetto ai due coefficienti è possibile ottenere la stima di $\hat{\beta}_0$ e di $\hat{\beta}_1$:

$$\begin{cases}\hat{\beta}_0=\overline{y}-\hat{\beta}_1\overline{x}\\ \hat{\beta}_1=S_{xy}/S_{xx}\end{cases} \qquad \text{(A.3)}$$

in cui

$$\begin{aligned}&\overline{x}=1/s\sum_{j=1}^{s}x_j, \quad S_{xx}=\sum_{j=1}^{s}\left(x_j-\overline{x}\right)^2,\\ &\overline{y}=1/s\sum_{j=1}^{s}\overline{y}_j, \quad S_{xy}=\sum_{j=1}^{s}y_j\left(x_j-\overline{x}\right),\end{aligned} \qquad \text{(A.4)}$$

rappresentano, rispettivamente, le medie aritmetiche di x_j e di y_j, la somma corretta con la media dei quadrati degli x_j e la somma corretta dei prodotti incrociati tra x_j e y_j.
Le stime della variabile dipendente sono quindi:

$$\hat{y}_j=\hat{\beta}_0+\hat{\beta}_1 x_j , \qquad \text{(A.5)}$$

mentre i residui vengono definiti come:

$$\hat{\varepsilon}_j=y_j-\hat{y}_j=y_j-\left(\hat{\beta}_0-\hat{\beta}_1 x_j\right) \quad \text{con } j = 1,\ldots,s. \qquad \text{(A.6)}$$

Nel caso in cui si utilizzi il Modello (A.3) per la stima dei coefficienti della regressione, oltre alle ipotesi già citate sulla componente ε di errore casuale,

occorre che valga l'ulteriore ipotesi di omoschedasticità, ovvero che la varianza dell'errore sia costante ($var(\boldsymbol{\varepsilon}) = \boldsymbol{\sigma}^2 = cost$). Se valgono tutte queste ipotesi il metodo dei minimi quadrati ordinari fornisce la migliore stima lineare indistorta dei parametri.

A.2. Regressione lineare multipla

Qualora la variabile dipendente *y* sia messa in relazione con più di una variabile esplicativa, il modello regressivo costruito si dice "di regressione lineare multipla" ed è del tipo:

$$y_j = \beta_0 + \beta_1 x_{j,1} + \ldots + \beta_v x_{j,v} + \ldots + \beta_p x_{j,p} + \varepsilon_j \qquad \text{(A.7)}$$

dove x_j è una delle p variabili esplicative, β_v rappresentano i $p+1$ coefficienti della regressione ed ε_j è il termine di errore che è supposto essere distribuito indipendentemente ed identicamente con media 0 e varianza σ_j^2.
Nel trattare i modelli di regressione multipla è più conveniente esprimere le equazioni in notazione matriciale, per cui la (A.7) diventa

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \text{(A.8)}$$

in cui:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,p} \\ 1 & x_{2,1} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{s,1} & \cdots & x_{s,p} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_s \end{bmatrix}.$$

Avendo a disposizione *s* osservazioni (con *s* maggiore del numero di parametri *p* da stimare), ed indicando con y_j la *j*-esima osservazione della variabile dipendente e con $x_{j,v}$ la *j*-esima osservazione della *v*-esima variabile esplicativa, il metodo dei minimi quadrati ordinari consiste nel minimizzare:

$$S\left(\beta_0, \beta_1, \ldots, \beta_p\right) = \sum_{j=1}^{s} \left(y_j - \beta_0 - \sum_{v=1}^{p} \beta_j x_{j,v} \right)^2 . \qquad \text{(A.9)}$$

Lo stimatore di β col metodo dei minimi quadrati ordinari è:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} \tag{A.10}$$

se esiste la matrice inversa $\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$ ovvero se le variabili esplicative sono linearmente indipendenti tra di loro. Il vettore delle stime della regressione è:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}\,, \tag{A.11}$$

mentre quello dei residui risulta:

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\,. \tag{A.12}$$

A.2.1 Metodo dei minimi quadrati pesati

Nel caso in cui venisse meno l'ipotesi di omoschedasticità, assunta valida nei paragrafi precedenti, la stima dei parametri β_0 e β_1 dovrebbe essere effettuata utilizzando il metodo dei minimi quadrati pesati (*Weighted Least Squares*, W.L.S.). Quest'ultimo fornisce una stima di β_0, β_1, β_p tramite la minimizzazione della funzione:

$$S\left(\beta_0,\beta_1,\ldots,\beta_p\right) = \sum_{j=1}^{s}\left[y - w_j\left(\beta_0 + \sum_{v=1}^{p-1}\beta_v x_{j,v}\right)\right]^2\,, \tag{A.13}$$

dove ai valori stimati viene attribuito un peso pari a w_j, che viene fissato diversamente a seconda della variabile esplicativa ***y*** che si sta considerando.

A.3. Coefficiente di determinazione

Per valutare la capacità descrittiva di un modello regressivo è necessario stabilire in quale misura quello specifico modello riesce a spiegare la variabilità della variabile dipendente. A tal fine risulta utile calcolare il coefficiente di determinazione R^2, definito come il complemento a 1 del

rapporto tra la variabilità residua (non spiegata dalla regressione) e la variabilità delle osservazioni nelle *s* stazioni esaminate:

$$R^2 = 1 - \frac{\mathbf{y}^T\mathbf{y} - \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y} - \frac{1}{s}\left(\sum_{j=1}^{s} y_j\right)^2}. \qquad \text{(A.14)}$$

Il valore di R^2 si avvicina all'unità quanto più la variabilità di ***y*** è spiegata dal modello di regressione.
Qualora si voglia confrontare la capacità descrittiva di modelli multiregressivi con un numero diverso di regressori, è necessario apportare una correzione alla (A.14), in quanto R^2 tende ad aumentare quando si aggiunge una variabile esplicativa. Al suo posto si usa un coefficiente di determinazione corretto:

$$R^2_{adj} = 1 - \frac{\mathbf{y}^T\mathbf{y} - \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y} - \frac{1}{s}\left(\sum_{j=1}^{s} y_j\right)^2} \cdot \frac{(s-1)}{(s-p)}. \qquad \text{(A.15)}$$

Ordinando i modelli regressivi in base al valore di R^2_{adj} è possibile stabilire quali siano, a parità di numero di descrittori utilizzati, le relazioni più adeguate a spiegare la grandezza osservata. I modelli caratterizzati da valori di R^2_{adj} maggiori sono quelli aventi una migliore capacità descrittiva.

A.4. Test di Student

Il test della *t* di Student viene impiegato per individuare i modelli regressivi per i quali anche solo una delle variabili esplicative risulta non significativa per spiegare la variabilità della variabile dipendente.
Si consideri la regressione lineare semplice di equazione (A.1) con lo scopo di testare il caso in cui il coefficiente angolare della retta regressiva β_1 sia uguale ad una costante β^*. L'ipotesi nulla e l'ipotesi alternativa siano, rispettivamente, $H_0 : \beta_1 = \beta^*$ e $H_1 : \beta_1 \neq \beta^*$; gli errori siano indipendenti e distribuiti con distribuzione normale $\varepsilon \sim N(0, \sigma^2)$.

Lo stimatore del parametro β_1 dell'equazione (A.3), calcolato con il metodo dei minimi quadrati, è lineare nei confronti dei valori di y_i e, come dimostrato da *Montgomery e Ranger* (2003), distribuito come $\hat{\beta}_1 \sim N\left(\beta_1, \sigma^2 / S_{xx}\right)$ dove S_{xx} è la sommatoria dei quadrati delle differenze rispetto alla media espressa nella (A.4). Ne consegue che nel caso di validità dell'ipotesi nulla H_0:

$$Z = \frac{\hat{\beta}_1 - \beta^*}{\sigma / \sqrt{S_{xx}}} \sim N(0,1). \qquad \text{(A.16)}$$

Se conoscessimo σ^2 potremmo usare Z per testare H_0, ma, tipicamente, la varianza dell'errore non è nota. Si può però dimostrare che la statistica

$$\frac{(s-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{s-2} \qquad \text{(A.17)}$$

è distribuita secondo una distribuzione del chi-quadro con $s - 2$ gradi di libertà, e che le stime di σ^2 e β_1 sono indipendenti. In considerazione delle proprietà della distribuzione t di Student, si può dire che:

$$T = \frac{\hat{\beta}_1 - \beta^*}{\hat{\sigma} / \sqrt{S_{xx}}} \sim t_{s-2}\,, \qquad \text{(A.18)}$$

cioè la statistica T è distribuita come una distribuzione t di Student con $s - 2$ gradi di libertà.

L'Equazione (A.18) diventa quindi la statistica test da utilizzarsi nel caso in cui si voglia testare la significatività del parametro β_1: si pone β^* uguale a zero, si calcola T e si va a vedere quanto vale la corrispondente t_{limite}, relativa al livello di significatività prescelto. Se $T < t_{limite}$, il parametro non è distinguibile da zero e la variabile esplicativa non deve essere utilizzata nella regressione poiché non è significativamente legata alla variabile dipendente.

Affinché il test definito dalla statistica T dell'Equazione (A.18) possa essere usato, occorre a rigore che gli errori del modello regressivo siano distribuiti normalmente. Nella pratica si è riscontrato che per deboli "non-normalità" il test risulta essere comunque significativo (*Montgomery e Ranger*, 2003).

Nel caso della regressione lineare multipla di Equazione (A.7), il procedimento è analogo a quello presentato per la regressione lineare semplice, e viene utilizzato per valutare la significatività di ognuno dei parametri della regressione. In questo modo si possono eliminare una o più delle variabili esplicative scelte se queste non danno un contributo significativo. Nel caso di regressione lineare multipla vale che:

$$T = \frac{\hat{\beta}_j - \beta_j^*}{\sqrt{\hat{\sigma}^2 c_{vv}^{'}}} \sim t_{s-p} \tag{A.19}$$

dove $c_{vv}^{'}$ con *v = 1, …, p-1* sono gli elementi della diagonale della matrice $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$.

A.5. Multicollinearità

Per i modelli regressivi che superano il test *t* di Student è necessario verificare che non ci sia correlazione lineare tra i regressori considerati, ossia si deve accertare l'assenza di multicollinearità tra le variabili esplicative. Un esito negativo a tale verifica significherebbe la formulazione di un modello instabile, nel senso che un campione differente di osservazioni della medesima variabile potrebbe generare un modello completamente diverso. Infatti la multicollinearità influenza la procedura dei minimi quadrati determinando problemi di stima dei coefficienti β.

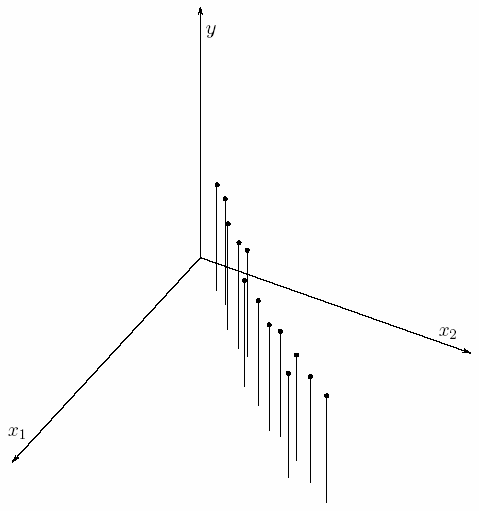


Figura A.2. Set di dati affetti da multicollinearità

Il caso rappresentato in Figura A.2 può essere utile a spiegare gli effetti dovuti alla multicollinearità. Adattare un modello regressivo ai dati (x_1, x_2, y) indicati in tale figura è analogo a far passare un piano inclinato tra i punti. Ovviamente l'inclinazione di questo piano sarà molto instabile e sensibile a piccoli cambiamenti della posizione dei punti. Inoltre, nonostante il modello possa predire abbastanza bene y nei punti (x_1, x_2) vicini a quelli dei dati, qualsiasi estrapolazione al di fuori di questi sarà verosimilmente poco affidabile.

Una statistica adeguata a misurare la presenza di multicollinearità fra le variabili esplicative di un modello multiregressivo è rappresentata dal *variance inflation factor* (*Montgomery e Ranger*, 2003):

$$VIF = \left(1 - R_j^2\right)^{-1}, \tag{A.20}$$

dove R_j^2 rappresenta il coefficiente di determinazione della regressione lineare tra la variabile indipendente x_j e i *p-1* restanti regressori.
Dall'esperienza pratica si desume che se uno dei VIF calcolati per le diverse variabili esplicative supera un valore pari a 5 la possibilità che i coefficienti della regressione siano stati stimati male a causa di multicollinearità è elevata e, di conseguenza, è consigliabile scartare il modello regressivo in questione.

Appendice B. Variabilità residua nella formula razionale

La relazione prescelta (4.3) per la stima della piena indice secondo un approccio basato sulla formula razionale corretta con la quota media e l'afflusso medio annuo conduce a un errore medio assoluto di stima pari a 0.37. Questo significa che la variabilità della piena indice viene spiegata solo in parte dai diversi fattori utilizzati per la stima. Per verificare quali potrebbero essere gli ulteriori parametri utili a spiegare la variabilità residua è utile ricorrere a una analisi regressiva. In particolare tramite tale approccio si cerca di capire se il coefficiente $\hat{K}_{4,g}$ risulta ancora legato in modo significativo a qualche parametro climatico, geomorfologico o di uso del suolo. Si effettua una analisi regressiva nel campo logaritmico secondo la (2.6), fissando la variabile dipendente y_j pari a $ln(\hat{K}_{4,g})$ e considerando la medesima matrice $\boldsymbol{X}$ di parametri climatici, geomorfologici e di uso del suolo considerati nel paragrafo 3.2:

$$K_{4,j} = X^{\beta} \cdot \varepsilon \,. \tag{B.1}$$

Nel caso specifico i descrittori che superano il test della *t* di Student, ordinati secondo il valore di R^2_{adj}, sono:

1) *CN*;
2) Y_{sc}.

Tra i parametri che aiutano a spiegare la variabilità residua della piena indice, oltre a CN, non risulta nessun altro parametro legato alla permeabilità dei suoli.

Per verificare se l'introduzione di un ulteriore parametro sia effettivamente utile a spiegare la variabilità residua della stima è necessario calcolare nuovamente le statistiche di errore calcolando la piena indice come:

$$Q_{ind,j} = \hat{K}_{4,res} \cdot a_j \cdot A_j^{0.5+0.5n_j} \tag{B.2}$$

dove $\hat{K}_{4.res} = \exp\left(\frac{\sum_{j=1}^{s} \ln\left(\hat{K}_{4,j,res}\right)}{s}\right)$ e $\hat{\mathrm{K}}_{4.\mathrm{j,res}} = \frac{\mathrm{Q}_{\mathrm{ind,j}}}{\mathrm{f}_{4,\mathrm{j}} \cdot \mathrm{CN}^{-1.38}}$ o $\hat{K}_{4.j,res} = \frac{Q_{ind,j}}{f_{4,j} \cdot Y_{sc}^{9.44}}$

a seconda del parametro geomorfologico impiegato per spiegare la variabilità residua.

Analizzando le statistiche di errore ottenute, riportate in tabella C.1, emerge chiaramente come entrambi i parametri *CN* e Y_{sc} conducano a ridurre la variabilità residua nella stessa misura. Mediante l'introduzione di tali parametri si riduce l'errore medio assoluto percentuale da 0.37 a 0.35 prima della cross validazione, mentre dopo la procedura di jack-knife si passa da 0.38 a 0.35, nel caso del *CN*, e a 0.36, nel caso di Y_{sc}. Si nota inoltre il corrispondente innalzamento dei valori di RMSE e di MAE che, dopo la cross validazione, aumentano, per *CN*, del 4% e dell'1%, mentre per Y_{sc} crescono del 5% e del 2% rispetto ai valori riportati nelle tabelle 3.7 e 3.8.
Ne consegue che, tra i due parametri, il *CN* risulta il più indicato a correggere la formula razionale espressa dalla (4.3).

Variabilità residua	*Prima della cross validazione*			*Dopo la cross validazione*		
Descrittore	*RMSE*	*MAE*	*MAPE*	$RMSE_{CV}$	MAE_{CV}	$MAPE_{CV}$
CN	0.39	0.24	0.35	0.39	0.24	0.35
Y_{sc}	0.39	0.24	0.35	0.40	0.25	0.36

Tabella C.1. Statistiche del residuo della regressione tra il K4 ed alcuni descrittori di bacino.

In tal caso risulterebbe:

$$Q_{ind,j} = 0.05 \cdot CN^{-1.38} \cdot C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j}, \qquad \text{(B.3)}$$

nella quale è da notare come il parametro *CN* abbia esponente negativo, al contrario di quello che ci si aspetterebbe. Il *CN* infatti rappresenta un indice di permeabilità dei suoli, il cui valore cresce all'aumentare del grado di impermeabilità del terreno. Al crescere di *CN* dovrebbe dunque corrispondere un aumento della piena indice, al contrario di quanto risulta dalla (B.3).

Per questi motivi si ritiene che il parametro *CN* non possa essere considerato un descrittore adeguato della piena indice; quindi per spiegare, seppur in modo marginale, la variabilità residua si dovrebbe ricorrere a Y_{sc}:

$$Q_{ind,j} = 9.47 \cdot 10^{-68} \cdot Y_{sc}^{9.44} \cdot C_{H_{med},j} \cdot aff_j \cdot a_j \cdot A_j^{0.5+0.5n_j}, \qquad \text{(B.4)}$$

la quale, non ha tuttavia un chiaro significato fisico. Per questo motivo nelle applicazioni della formula razionale si è fatto riferimento alla relazione (4.3), leggermente meno efficiente della (B.4) dal punto di vista statistico, ma decisamente più consistente dal punto di vista fisico.

Appendice C. L-momenti

Nelle procedure di analisi di frequenza regionale si adattano ai dati delle distribuzioni la cui forma si ritiene conosciuta a meno di un numero finito di parametri incogniti. I momenti campionari ordinari, in particolare media, scarto, asimmetria e kurtosis, sono spesso utilizzati per la stima dei parametri delle distribuzioni di probabilità. Hosking e Wallis (1997) suggeriscono invece di utilizzare, al posto dei momenti ordinari, gli L-momenti perché adatti a descrivere più distribuzioni, perché più robusti nella stima da campioni poco consistenti di dati in presenza di outliers e perché meno soggetti a distorsione nella stima. In questa appendice, tratta da Hosking e Wallis (1997), si definiscono gli L-momenti in maniera formale. Dopo una breve introduzione sui concetti di distribuzione di probabilità e di stimatori dei parametri, vengono definiti gli L-momenti e si discutono alcune loro proprietà.

C.1. Distribuzioni di probabilità

Si consideri una variabile casuale X, che può assumere valori appartenenti all'insieme dei numeri reali. La frequenza relativa con cui questi valori si verificano definisce la distribuzione di frequenza o distribuzione di probabilità di X, che è specificata dalla distribuzione di frequenza cumulata:

$$F(x) = \Pr[X \le x], \tag{C.1}$$

dove *Pr(A)* indica la probabilità dell'evento *A*.
F(x) è una funzione crescente di *x*, definita nell'intervallo [0, 1]. Normalmente in idrologia si ha a che fare con variabili casuali continue, per le quali *Pr[X=t]=0* per ogni *t*, ovvero a nessun valore è associata una probabilità non nulla. In questo caso *F(.)* è una funzione continua ed ha una funzione inversa corrispondente *x(.)*, detta *funzione dei quantili* di *X*. Data una qualsiasi *u*, dove *0 < u < 1*, *x(u)* è l'unico valore che soddisfa

$$F(x(u)) = u\,. \tag{C.2}$$

Per ogni probabilità *p*, *x(p)* è il *quantile di non superamento* della probabilità *p*, ovvero il valore per cui la probabilità che *X* non superi *x(p)* è pari a *p*. L'obiettivo dell'analisi di frequenza è la stima accurata dei quantili della distribuzione di una data variabile casuale. In ingegneria, e nelle applicazioni ambientali in generale, i quantili sono spesso espressi in termini di tempo di ritorno *T*; in particolare per un evento estremamente alto, nella coda superiore della distribuzione di frequenza, vale che:

$$F(x_T) = 1 - \frac{1}{T}. \tag{C.3}$$

Se *F(x)* è differenziabile, la sua derivata $f(x) = \frac{d}{dx} F(x)$ è la densità di probabilità di *X*.
Il *valore atteso* della variabile casuale *X* è definito come:

$$E(X) = \int_{-\infty}^{+\infty} x dF(x) = \int_{-\infty}^{+\infty} x f(x) dx , \tag{C.4}$$

ammesso che l'integrale esista. Se si considera la trasformazione *u = F(x)*, si può scrivere:

$$E(X) = \int_{0}^{1} x(u) du . \tag{C.5}$$

Una funzione di una variabile casuale *g(X)* è anch'essa una variabile casuale di valore atteso

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) dF(x) = \int_{-\infty}^{+\infty} g(x) f(x) dx = \int_{0}^{1} g(x(u)) du \tag{C.6}$$

La dispersione dei valori estratti dalla variabile casuale *X* può essere misurata con la *varianza* di *X*:

$$\mathrm{var}(X) = E\left[\{X - E(X)\}^2\right]. \tag{C.7}$$

In alcuni caso può essere utile misurare la tendenza di due variabili casuali *X* e *Y* ad assumere valori elevati simultaneamente. Questo può essere misurato dalla *covarianza* di *X* e *Y*:

$$\mathrm{cov}(X,Y) = E\left[\{X - E(X)\}\{Y - E(Y)\}\right]. \qquad \text{(C.8)}$$

La *correlazione* tra *X* e *Y*

$$corr(X,Y) = \mathrm{cov}(X,Y)/\{\mathrm{var}(X)\,\mathrm{var}(Y)\}^{1/2}, \qquad \text{(C.9)}$$

è il corrispettivo dimensionale della covarianza, che può assumere valori compresi tra -1 e +1.

C.2. Stimatori

Nella pratica spesso si assume che la forma di una qualche distribuzione di probabilità sia conosciuta a meno di un set di parametri incogniti θ1, ..., θp. Sia x(u; θ1, . . . , θp) la funzione dei quantili di una distribuzione con p parametri incogniti. In molte applicazioni tra i parametri incogniti si possono identificare un parametro di posizione ed un parametro di scala. Un parametro ξ di una distribuzione è un parametro di posizione se per la funzione dei quantili vale l'eguaglianza

$$x(u;\xi,\theta_2,...,\theta_p) = \xi + x(u;0,\theta_2,...,\theta_p). \qquad \text{(C.10)}$$

Si dice, invece, che α è un parametro di scala della funzione dei quantili della distribuzione se

$$x(u;\alpha,\theta_2,...,\theta_p) = \alpha \times x(u;1,\theta_2,...,\theta_p). \qquad \text{(C.11)}$$

Se per la distribuzione esistono entrambi questi parametri, allora vale l'eguaglianza

$$x(u;\xi,\alpha,\theta_3,...,\theta_p) = \xi + \alpha \times x(u;0,1,\theta_3,...,\theta_p). \qquad \text{(C.12)}$$

I parametri incogniti sono stimati a partire dai dati osservati. Dato un set di osservazioni, una funzione $\hat{\theta}$ di queste deve essere scelta come *stimatore* di *θ*. Lo stimatore $\hat{\theta}$ `e a sua volta una variabile casuale ed ha una distribuzione di probabilità. La bontà di $\hat{\theta}$ come stimatore di *θ* tipicamente dipende da quanto $\hat{\theta}$ si avvicina a *θ*. La deviazione di $\hat{\theta}$ da *θ* può essere scomposta in *distorsione* (tendenza di dare stime sistematicamente più alte, o più basse, del valore vero) e variabilità (deviazione casuale dal valore vero, che si verifica anche per gli stimatori che non presentano distorsione).

La performance di uno stimatore $\hat{\theta}$ può essere valutata con due misure, il *bias* ("distorsione") e la *radice dell'errore quadratico medio* (RMSE), definite come

$$bias\left(\hat{\theta}\right) = E\left(\hat{\theta} - \theta\right), \quad RMSE\left(\hat{\theta}\right) = \left\{E\left(\hat{\theta} - \theta\right)^2\right\}^{1/2}, \tag{C.13}$$

e caratterizzate dall'avere la stessa unità di misura del parametro *θ*. Si dice che lo stimatore $\hat{\theta}$ è *indistorto* se *bias*($\hat{\theta}$) = *0* ovvero se *E*($\hat{\theta}$) = *θ*. Diversi stimatori indistorti dello stesso parametro possono essere paragonati in termini della loro varianza: il rapporto $\mathrm{var}\left(\hat{\theta}^{(1)}\right) / \mathrm{var}\left(\hat{\theta}^{(2)}\right)$ si dice efficienza dello stimatore $\hat{\theta}^{(2)}$ rispetto allo stimatore $\hat{\theta}^{(1)}$. La radice dell'errore quadratico medio può essere anche scritta come

$$RMSE\left(\hat{\theta}\right) = \left[\left\{bias\left(\hat{\theta}\right)\right\}^2 + \mathrm{var}\left(\hat{\theta}\right)\right]^{1/2}, \tag{C.14}$$

da cui si vede come RMSE combina distorsione e variabilità di $\hat{\theta}$ e dà una misura globale dell'accuratezza della stima. Nei classici problemi di statistica in cui la stima dei parametri è basata su un campione di lunghezza *n*, sia il bias che la varianza di $\hat{\theta}$ sono asintoticamente proporzionali a n^{-1} per *n* grandi (v.es.*Cox e Hinkley*, 1974), per cui l'RMSE di $\hat{\theta}$ è proporzionale a $n^{-1/2}$.

C.3. L-momenti delle distribuzioni di probabilità

Gli *L*-momenti sono un sistema alternativo di descrivere la forma delle distribuzioni di probabilità. Storicamente essi nascono come modifica dei momenti pesati in probabilità di *Greenwood et al.* (1979). I momenti pesati in probabilità di una variabile casuale *X* con distribuzione di frequenza cumulata *F(.)* sono le quantità:

$$M_{p,r,s} = E\left[X^p \{F(X)\}^r \{1 - F(X)\}^s\right]. \qquad \text{(C.15)}$$

Particolarmente utili sono i momenti pesati in probabilità $\alpha_r = M_{1,r,0}$. Per una distribuzione caratterizzata da una funzione dei quantici *x(u)*, dalle relazioni (C.6) e (C.15) si ottiene:

$$\alpha_r = \int_0^1 x(u)(1-u)^r\, du\,, \qquad \beta_r = \int_0^1 x(u) u^r\, du\,. \qquad \text{(C.16)}$$

Mentre i momenti ordinari considerano successive elevazioni di potenza della funzione dei quantili x(u), i momenti pesati in probabilità considerano successive elevazioni di potenza di *u* oppure *1 − u* e possono essere visti come integrali di *x(u)* pesati con i polinomi u^r oppure $(1-u)^r$.

I momenti pesati in probabilità α_r e β_r sono stati usati in letteratura come base di metodi per la stima dei parametri delle distribuzioni di probabilità ma sono difficilmente interpretabili come misure di scala e forma di queste. Queste informazioni sono contenute in certe combinazioni lineari dei momenti pesati in probabilità. Ad esempio, multipli di $\alpha_0 - 2\alpha_1$ o $2\beta_1 - \beta_0$ sono stime dei parametri di scala delle distribuzioni, mentre l'asimmetria può essere misurata da $6\beta_2 - 6\beta_1 + \beta_0$. Queste combinazioni lineari derivano naturalmente dall'integrazione di *x(u)* pesata non con i polinomi u^r o $(1-u)^r$, ma con un set di polinomi ortogonali.

Si definiscano polinomi di Legendre sfasati (perchè definiti nell'intervallo [0,1] invece che nell'intervallo [−1,+1]) i polinomi $P_r^*(u)$ con *r = 0, 1, 2,…*, che godono delle seguenti proprietà:

1. $P_r^*(u)$ è un polinomio di grado *r* in *u*;
2. $P_r^*(1)=1$;
3. $\int_0^1 P_r^*(u)P_s^*(u)du=0$ se $r \neq s$ (condizione di ortogonalità).

I polinomi di Legendre sfasati hanno la forma esplicita

$$P_r^*(u)=\sum_{k=0}^{r} p^*_{r,k}\, u^k \ , \qquad \text{(C.17)}$$

dove:

$$p^*_{r,k} = (-1)^{r-k}\binom{r}{k}\binom{r+k}{k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} \qquad \text{(C.18)}$$

Gli *L*-momenti di una variabile casuale *X* con funzione dei quantili *x(u)* sono definiti come

$$\lambda_r = \int_0^1 x(u)P^*_{r-1}(u)du \ . \qquad \text{(C.19)}$$

In termini di momenti pesati in probabilità gli L-momenti sono dati da:

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p^*_{r,k}\alpha_k = \sum_{k=0}^{r} p^*_{r,k}\beta_k \ . \qquad \text{(C.20)}$$

E' conveniente definire le versioni adimensionali degli L-momenti, le quali si possono ottenere dividendo gli L-momenti di ordine superiore per la misura di scala λ_2. Si ottengono così i *rapporti degli L-momenti*, ovvero gli *L-coefficienti*

$$\tau_r = \frac{\lambda_r}{\lambda_2}, \quad r = 3, 4 \qquad \text{(C.21)}$$

che misurano la forma di una distribuzione indipendentemente dalla scala.
In particolare se *r = 3* si ottiene il coefficiente di L-asimmetria

$$L_{ca} = \frac{\lambda_3}{\lambda_2}, \tag{C.22}$$

mentre se *r = 4* si ottiene il coefficiente di L-kurtosis

$$L_{kur} = \frac{\lambda_4}{\lambda_2}. \tag{C.23}$$

Il coefficiente di L-variazione L_{cv} è invece definito come:

$$L_{cv} = \frac{\lambda_2}{\lambda_1}. \tag{C.24}$$

C.4. Proprietà degli L-momenti

Gli *L*-momenti λ_1 e λ_2, L_{cv}, L_{ca} e L_{kur} sono le quantità che Hosking e Wallis (1997) consigliano di utilizzare per descrivere le distribuzioni di probabilità.
Le loro proprietà più importanti sono:

- **Esistenza**. Se esiste la media della distribuzione, allora esistono tutti i suoi *L*-momenti e i suoi *L*-coefficienti.
- **Unicità**. Se esiste la media della distribuzione, allora gli *L*-momenti definiscono tale distribuzione in maniera univoca, ovvero non esistono due distribuzioni diverse con gli stessi *L*-momenti.
- **Terminologia**. Gli *L*-momenti e gli *L*-coefficienti che si utilizzano hanno un determinato significato, paragonabile a quello dei momenti campionari.
- **Limiti algebrici**. λ_1 pu`o assumere qualsiasi valore; $\lambda_2 \geq 0$; per una distribuzione che assume solo valori positivi $0 \leq L_{cv} < 1$; gli *L*-coefficienti soddisfano l'uguaglianza $|\tau_r| < 1$ per ogni $r \geq 3$. Limiti più precisi possono essere trovati per ogni τ_r: ad esempio, dato L_{ca}, allora

$(5L_{ca}^2 - 1)/4 \le L_{kur} < 1$ e, per distribuzioni che assumono solo valori positivi, dato L_{cv} si ha che $2L_{cv} - 1 \le L_{ca} < 1$.

- **Trasformazioni lineari**. Siano X e Y due variabili casuali con L-momenti λ_r e λ_r^* rispettivamente, e si supponga che $Y = aX+b$. Allora $\lambda_1^* = a\lambda_1 + b$; $\lambda_2^* = |a|\lambda_2$; $\tau_r^* = (sign(a))^r \tau_r$ per $r \ge 3$.
- **Simmetria**. Sia X una variabile casuale simmetrica con media μ, ossia $Pr[X \ge \mu + x] = Pr[X \le \mu - x]$ per ogni x. Allora tutti i rapporti degli L-momenti di ordine dispari valgono 0, ovvero $\tau_r = 0$ se $r = 3, 5, 7, \ldots$.

Gli L-momenti sono stati calcolati per molte distribuzioni (si veda l'Appendice D). La distribuzione che gioca un ruolo centrale nella teoria degli L-momenti, analoga alla distribuzione Normale nella teoria dei momenti ordinari, è la distribuzione uniforme. Si può dimostrare che tutti gli L-momenti λ_r e rapporti degli L-momenti τ_r di ordine superiore (con $r \ge 3$) valgono zero per la distribuzione uniforme. La distribuzione Normale, per il fatto che è simmetrica, presenta gli L-momenti di ordine dispari nulli. La distribuzione esponenziale, invece, ha dei rapporti degli L-momenti particolarmente semplici: $\tau_3 = 1/3$, $\tau_4 = 1/6$.

Un modo conveniente per rappresentare gli L-momenti di diverse distribuzioni è il diagramma dei rapporti degli L-momenti, esemplificato in Figura C.1. Questo diagramma mostra gli L-momenti in un grafico i cui assi sono il coefficiente di L-asimmetria e il coefficiente si L-kurtosis. Una distribuzione a due parametri, caratterizzata da un parametro di posizione ed uno di scala, viene rappresentata sul diagramma da un punto. Infatti se due distribuzioni differiscono solo nei parametri di posizione e di scala, allora sono distribuzioni di due variabili casuali X e $Y = aX + b$ con $a > 0$, per cui, dato la probabilità delle trasformazioni lineari degli L-momenti ($\tau_r^* = (sign(a))^r \tau_r$), hanno lo stesso L_{ca} ed L_{kur}. Una distribuzione a tre parametri, invece, dal momento che è caratterizzata dai parametri di posizione, scala e forma, viene rappresentata sul diagramma da una linea, i cui punti corrispondono a differenti valori del parametro di forma. Distribuzioni con più di un parametro di forma generalmente ricoprono aree bidimensionali sul diagramma.

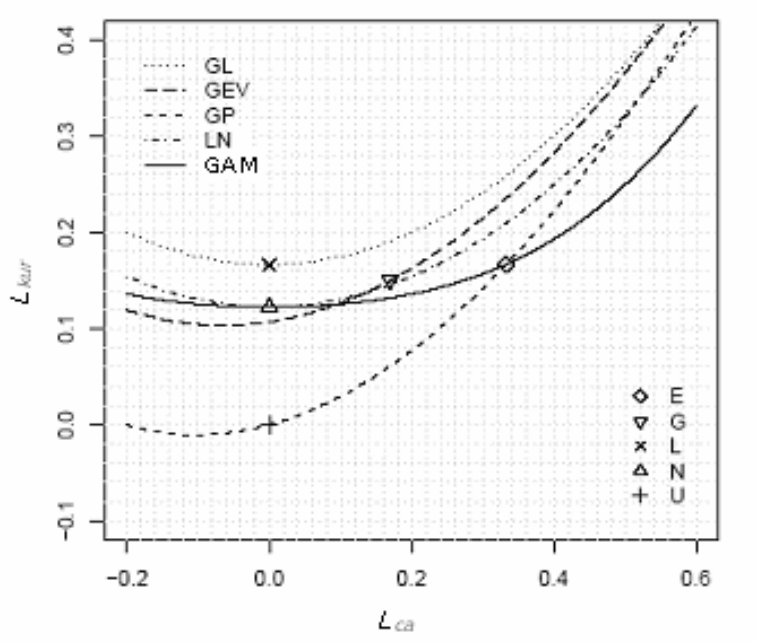


Figura C.1. Diagramma dei rapporti tra gli L-momenti. Le distribuzioni a due e a tre parametri sono riportate come punti e come linee rispettivamente. Le distribuzioni a due parametri sono: esponenziale (E), Gumbel (G), lognormale (L), normale (N), uniforme (U); quelle a tre parametri sono: logistica generalizzata (GL), generalizzata del valore estremo (GEV), Pareto generalizzata (GP), lognormale a tre parametri (LN), Pearson tipo III (GAM).

Appendice D. Distribuzioni di probabilità

D.1. Distribuzione di Gumbel

La distribuzione di Gumbel (*Gumbel*, 1958), appartenente alla famiglia esponenziale, è una delle più popolari nella modellazione di distribuzioni di frequenza di eventi naturali estremi. Da numerosi studi effettuati si è riscontrato che la Gumbel fornisce risultati molto consistenti ed è da preferire quando ci si riferisce a tempi di ritorno elevati. Tale distribuzione è caratterizzata da due parametri.

*Definizion*i

Parametri (2): ξ (posizione), α (scala).

Campo di esistenza di x: $-\infty < x < \infty$.

$$f(x) = \frac{1}{\alpha}\exp\left[-\frac{x-\xi}{\alpha}\right]\exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right], \qquad \text{(D.1)}$$

$$F(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right], \qquad \text{(D.2)}$$

$$x(F) = \xi - \alpha\log(-\log F). \qquad \text{(D.3)}$$

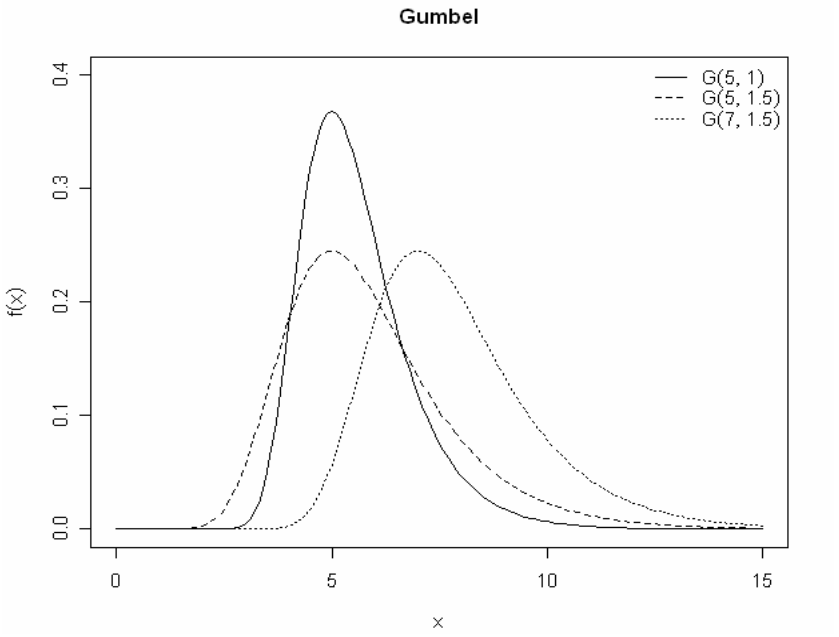


Figura D.1. Esempi di densità di probabilità di Gumbel con parametri: (1) ξ = 5, α = 1; (2) ξ = 5, α = 1.5; ξ = 7, α = 1.5.

L-coefficienti

$$L_{cv} = \frac{\alpha \log 2}{\xi + \alpha\gamma}, \tag{D.4}$$

dove γ è la costante di Eulero, 0.5772.

$$L_{ca} = \frac{\log(9/8)}{\log(2)}, \tag{D.5}$$

$$L_{kur} = \frac{[16\log(2) - 10\log(3)]}{\log(2)}. \tag{D.6}$$

Parametri

$$\alpha = \frac{L_{cv}}{\log 2}, \tag{D.7}$$

$$\xi = 1 - \gamma\alpha. \tag{D.8}$$

D.2. Distribuzione Pareto Generalizzata

La distribuzione Pareto Generalizzata è molto usata nell'analisi degli eventi estremi (*Pickands*, (1975) è stato probabilmente il primo ad utilizzarla in questo contesto), specialmente in idrologia e negli studi di affidabilità, quando occorre utilizzare alternative alla distribuzione esponenziale assumendo più spessa o più sottile la coda superiore della distribuzione.

Definizioni

Parametri (3): ξ (posizione), α (scala), k (forma).
Campo di esistenza di $\xi < x \leq \xi + \alpha/k$ se $k > 0$; $\xi \leq x < \infty$ se $k \leq 0$.

$$f(x) = \frac{1}{\alpha} e^{-(1-k)y}, \tag{D.9}$$

dove

$$y = \begin{cases} -k^{-1}\log\{1-k(x-\xi)/\alpha\}, & k \neq 0 \\ (x-\xi)/\alpha, & k = 0 \end{cases}. \qquad \text{(D.10)}$$

$$F(x) = 1 - e^{-y}, \qquad \text{(D.11)}$$

$$x(F) = \begin{cases} \xi + \alpha\left[1-(1-F)^{k}\right]/k, & k \neq 0 \\ \xi - \alpha\log(1-F), & k = 0 \end{cases}. \qquad \text{(D.12)}$$

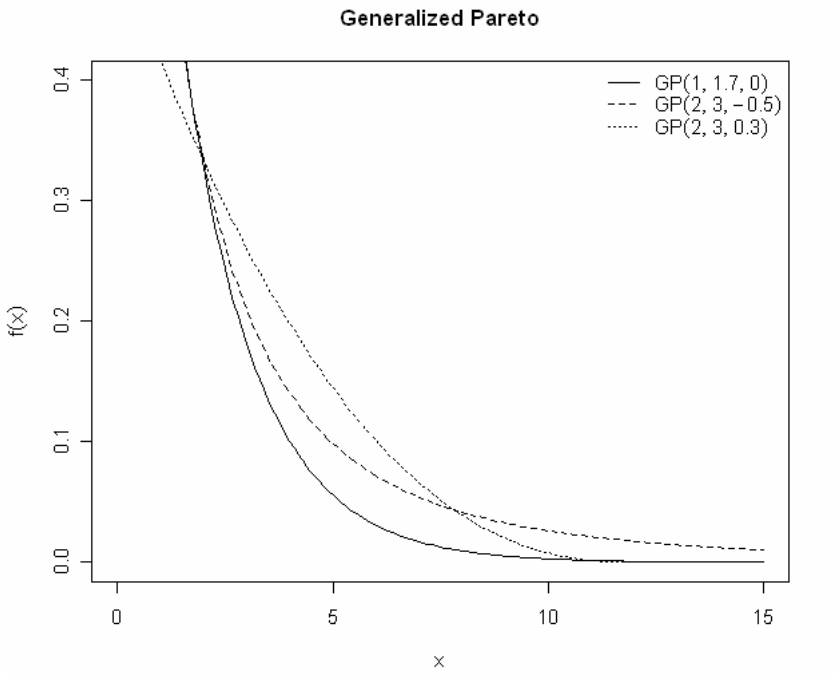


Figura D.2. Esempi di densità di probabilità della distribuzione Generalized Pareto con parametri: (1) ξ = 5, α = 1, k = 0; (2) ξ = 2, α = 3, k = -0.5; (3) ξ = 2, α = 3, k = -0.3.

L-coefficienti

Gli L-coefficienti sono definiti per k > -1.

$$L_{CV} = \frac{\alpha}{(2+k)\cdot[\alpha + \xi(1+k)]}, \qquad \text{(D.13)}$$

$$L_{ca} = \frac{1-k}{3+k}, \qquad \text{(D.14)}$$

$$L_{kur} = \frac{(1-k)(2-k)}{(3+k)(4+k)}. \tag{D.15}$$

Parametri

$$k = \frac{1-3L_{ca}}{1+L_{ca}}, \tag{D.16}$$

$$\alpha = \frac{(1+k)(2+k)L_{cv}}{L_{cv}}, \tag{D.17}$$

$$\xi = 1-(2+k)L_{cv}. \tag{D.18}$$

D.3. Distribuzione Generalizzata del Valore Estremo (GEV)

La GEV è una distribuzione a tre parametri, derivante dalla teoria dei valori estremi. Fu introdotta da *Jenkinson* (1955) per identificare la distribuzione di frequenza dei valori estremi per dati meteOrologici. La GEV è largamente utilizzata in ambito idrologico soprattutto per lo studio di piene e piogge intense.

Definizioni

Parametri (3): ξ (posizione), α (scala), k (forma).
Campo di esistenza di $-\infty < x \le \xi + \alpha/k$ se $k>0$; $-\infty < x < \infty$ se $k=0$; $\xi + \alpha/k \le x < \infty$ se $k < 0$.

$$f(x) = \frac{1}{\alpha} e^{-(1-k)y - e^{-y}}, \tag{D.19}$$

$$y = \begin{cases} -k^{-1}\log\{1-k(x-\xi)/\alpha\}, & k \neq 0 \\ (x-\xi)/\alpha, & k = 0 \end{cases}, \tag{D.20}$$

$$F(x) = e^{-e^{-y}}, \tag{D.21}$$

$$x(F)=\begin{cases}\xi+\alpha\left[1-\left(-\log F\right)^{k}\right]/k, & k\neq 0\\ \xi-\alpha\log\left(-\log\left(F\right)\right), & k=0\end{cases}. \tag{D.22}$$

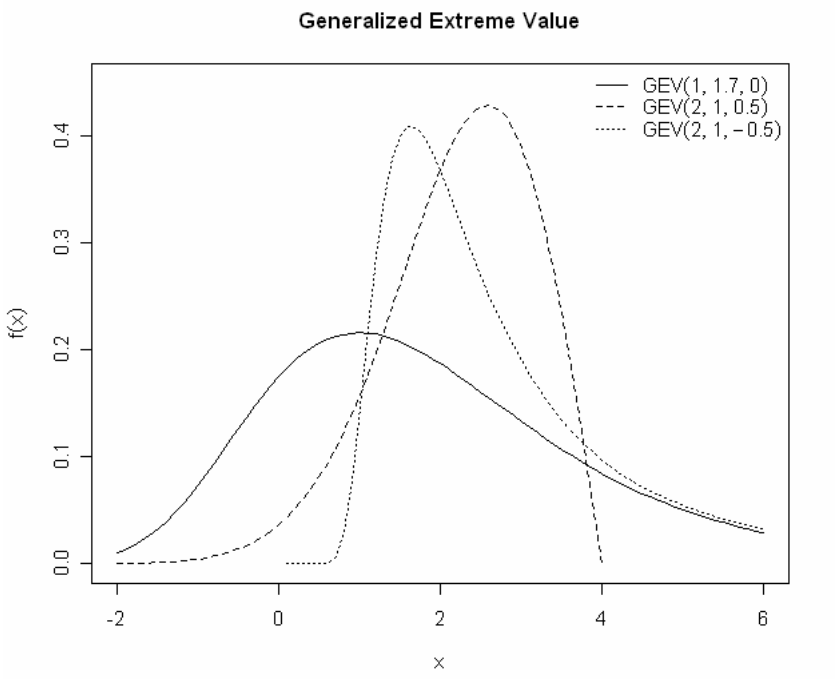


Figura D.3. Esempi di densità di probabilità della distribuzione GEV con parametri: (1) $\xi = 1$, $\alpha = 1.7$, $k = 0$; (2) $\xi = 2$, $\alpha = 1$, $k = 0.5$; (3) $\xi = 2$, $\alpha = 1$, $k = -0.5$.

L-coefficienti

Gli L-coefficienti sono definiti per k > 1.

$$L_{cv}=\frac{\alpha\left(1-2^{-k}\right)\Gamma\left(1+k\right)}{\xi+\alpha\left[1-\Gamma\left(1+k\right)\right]}, \tag{D.23}$$

$$L_{ca}=\frac{2\left(1-3^{-k}\right)}{1-2^{-k}}-3, \tag{D.24}$$

$$L_{kur}=\frac{\left[5\left(1-4^{-k}\right)-10\left(1-3^{-k}\right)+6\left(1-2^{-k}\right)\right]}{1-2^{-k}}. \tag{D.25}$$

Parametri

Per la stima del parametro *k* non esistono formule esplicite, di conseguenza si deve ricorrere ad un'approssimazione numerica

proposta da *Hosking e Wallis* (1997), che ha una precisione di 9 x 10^{-4} per $-0.5 \le L_{ca} \le 0.5$:

$$k \approx 7.8590c + 2.9554c^2 , \tag{D.26}$$

Dove

$$c = \frac{2}{3 + L_{ca}} - \frac{\log 2}{\log 3} . \tag{D.27}$$

Gli altri due parametri sono ricavabili da:

$$\alpha = \frac{L_{cv} k}{\left(1 - 2^{-k}\right)\Gamma\left(1 + k\right)} \tag{D.28}$$

$$\xi = 1 - \frac{\alpha}{k}\left[1 - \Gamma\left(1 + k\right)\right] \tag{D.29}$$

D.4. Distribuzione Logistica Generalizzata

La distribuzione logistica generalizzata nasce come distribuzione limite (n → ∞) delle medie standardizzate dei valori estremi (grandi e piccoli) di campioni casuali di lunghezza n (*Gumbel*, 1944). Esistono differenti forme di generalizzazione della distribuzione logistica. Quella qui riportata è una versione riparametrizzata della distribuzione log-logistica di *Ahmad et al.* (1988), che permette di mostrare la somiglianza della distribuzione con la Pareto generalizzata (Paragrafo D.2) e la GEV (Paragrafo D.3).

Definizioni

Parametri (3): ξ (posizione), α (scala), k (forma).
Campo di esistenza di $-\infty < x \le \xi + \alpha/k$ se k>0; $-\infty < x < \infty$ se k=0; $\xi + \alpha/k \le x < \infty$ se k < 0.

$$f(x) = \frac{\alpha^{-1} e^{-(1-k)y}}{\left(1 + e^{-y}\right)^2} , \tag{D.30}$$

dove

$$y = \begin{cases} -k^{-1}\log\{1-k(x-\xi)/\alpha\}, & k \neq 0 \\ (x-\xi)/\alpha, & k = 0 \end{cases}, \tag{D.31}$$

$$F(x) = \frac{1}{1+e^{-y}}, \tag{D.32}$$

$$x(F) = \begin{cases} \xi + \alpha\left[1-\left(\frac{1-F}{F}\right)^{k}\right]/k, & k \neq 0 \\ \xi - \alpha\log\left(\frac{1-F}{F}\right), & k = 0 \end{cases}. \tag{D.33}$$

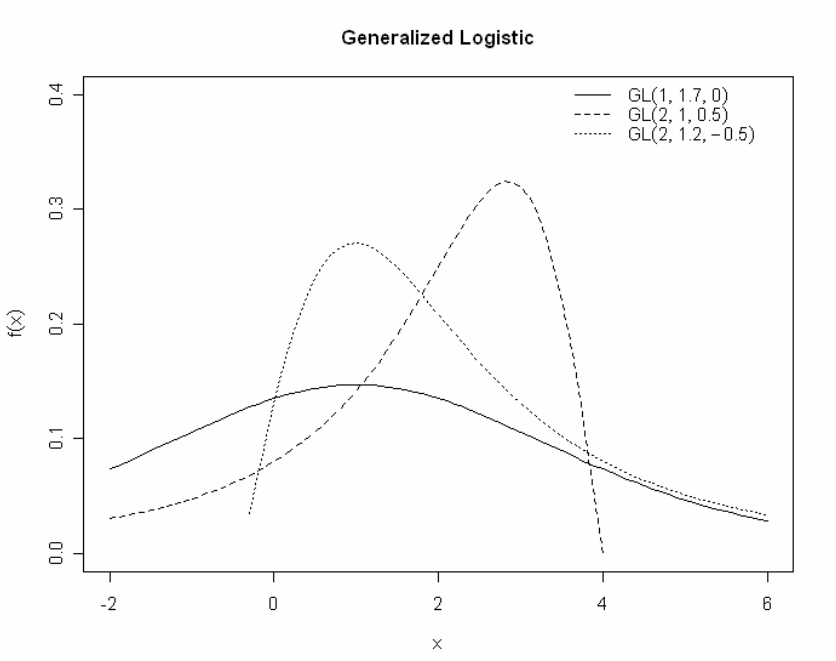


Figura D.4. Esempi di densità di probabilità della distribuzione logistica generalizzata con parametri: (1) $\xi = 1$, $\alpha = 1.7$, $k = 0$; (2) $\xi = 2$, $\alpha = 1$, $k = 0.5$; (3) $\xi = 2$, $\alpha = 1.2$, $k = -0.5$.

L-coefficienti

Gli L-momenti sono definiti per -1 < k< 1.

$$L_{cv} = \frac{\alpha k\pi}{\sin(k\pi)\cdot\{\xi + \alpha[1/k - \pi/\sin(k\pi)]\}}, \tag{D.34}$$

$$L_{ca} = -k \,, \tag{D.35}$$

$$L_{kur} = \frac{1+5k^2}{6} \,. \tag{D.36}$$

Parametri

$$k = -L_{ca} \,, \tag{D.37}$$

$$\alpha = \frac{L_{cv}\sin(k\pi)}{k\pi} \,, \tag{D.38}$$

$$\xi = 1 - \alpha\left(\frac{1}{k} - \frac{\pi}{\sin(k\pi)}\right) \,. \tag{D.39}$$

D.5. Distribuzione Lognormale

La distribuzione Lognormale è applicabile ad una gran varietà di fenomeni idrologici, specialmente quando le variabili hanno limite inferiore. Infatti la curva di densità di probabilità si presenta con andamento non simmetrico e limite inferiore. Storicamente si sono susseguiti vari studi che rivelano come la distribuzione lognormale sia adattabile all'applicazione in svariati campi, dalla modellazione di picchi di portata (*Chow*, 1954), alla descrizione delle concentrazioni di inquinanti in atmosfera (*Georgopoulos & Seinfeld*, 1982).
E' interessante notare che esiste anche una giustificazione teorica all'uso di tale distribuzione: si consideri il prodotto di una serie di variabili $X=W_1W_2...W_N$; facendo il logaritmo di ambo i membri si ricava $\ln(X)=\ln(W_1)+\ln(W_2)+. \; . \; .+\ln(W^N)$, da cui, per il teorema del limite centrale, si ottiene che X deve avere una distribuzione log-Normale, quando il numero di fattori nella moltiplicazione tende ad essere sufficientemente elevato.

Definizioni

Parametri (3): ξ (posizione), α (scala), k (forma).

Campo di esistenza di $-\infty < x \le \xi + \alpha/k$ se $k > 0$; $-\infty < x < \infty$ se $k = 0$; $\xi + \alpha/k \le x < \infty$ se $k < 0$.

$$f(x) = \frac{e^{ky - y^2/2}}{\alpha\sqrt{2\pi}}, \tag{D.40}$$

dove

$$y = \begin{cases} -k^{-1}\log\{1 - k(x-\xi)/\alpha\}, & k \neq 0 \\ (x-\xi)/\alpha, & k = 0 \end{cases}, \tag{D.41}$$

$$F(x) = \Phi(y), \tag{D.42}$$

in cui $\Phi(.)$ è la distribuzione di frequenza cumulata della distribuzione Normale standardizzata.

Per *x(F)* non esiste una forma analitica, ma si può ricorrere all'utilizzo di metodi numerici. Ad esempio nella parametrizzazione della distribuzione lognormale proposta in *Hosking e Wallis* (1997), la variabile casuale *X* è legata alla variabile casuale *Y*, che è distribuita secondo una Normale standard, in cui:

$$X = \begin{cases} \xi + \alpha\left(1 - e^{-kZ}\right)/k & k \neq 0 \\ \xi + \alpha Z & k = 0 \end{cases}. \tag{D.43}$$

La parametrizzazione standard è invece esprimibile come:

$$F(x) = \Phi\left(\frac{\log(x-\varsigma) - \mu}{\sigma}\right), \quad \varsigma \le x < \infty \tag{D.44}$$

La quale può essere ottenuta dalla parametrizzazione di Hosking e Wallis considerando che:

$$k = -\sigma, \quad \alpha = \sigma e^{\mu}, \quad \xi = \varsigma + e^{\mu}. \tag{D.45}$$

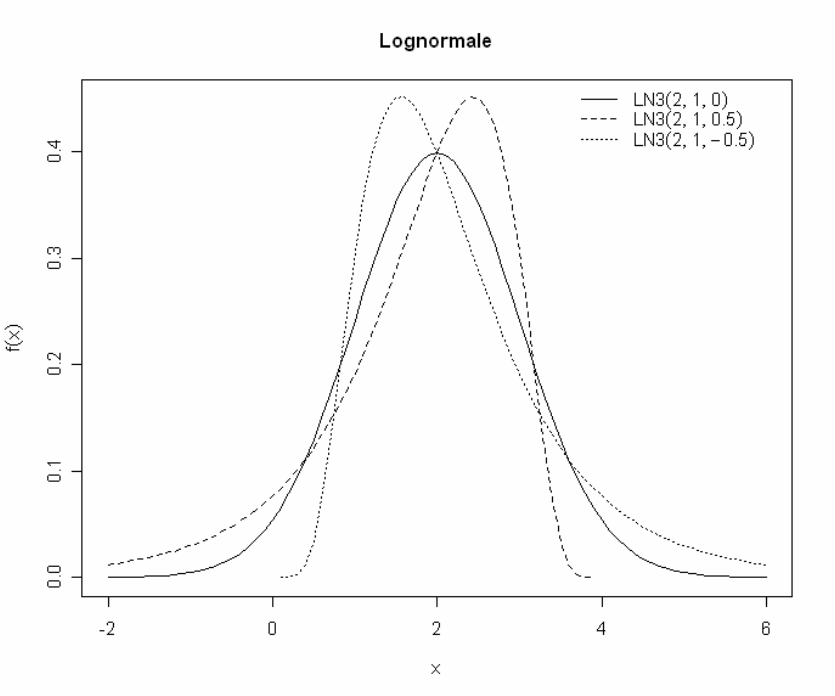


Figura D.5. Esempi di densità di probabilità della distribuzione lognormale con parametri: (1) $\xi = 2$, $\alpha = 1$, k = 0; (2) $\xi = 2$, $\alpha = 1$, k = 0.5; (3) $\xi = 2$, $\alpha = 1$, k = -0.5.

L-coefficienti

Gli L-coefficienti sono definiti per tutti i valori di *k*:

$$L_{cv} = \frac{\alpha e^{k^2/2}\left[1 - 2\Phi\left(-k/\sqrt{2}\right)\right]}{\xi k + \alpha\left(1 - e^{k^2/2}\right)}. \qquad \text{(D.46)}$$

Non esistono espressioni semplici per L_{ca} ed L_{kur}. Essi sono funzione del solo parametro k e possono essere calcolati tramite integrazione numerica. *Hosking e Wallis* (1997) propongono una approssimazione per L_{ca} ed L_{kur} che ha un'accuratezza rispettivamente di 2 x 10^{-7} e 5 x 10^{-7} per $|L_{ca}| \le 0.99$ e $|L_{kur}| \le 0.98$:

$$L_{ca} \approx -k\frac{A_0 + A_1k^2 + A_2k^4 + A_3k^6}{1 + B_1k^2 + B_2k^4 + B_3k^6}, \qquad \text{(D.47)}$$

$$L_{kur} \approx L_{kur,0} + k^2 \frac{C_0 + C_1 k^2 + C_2 k^4 + C_3 k^6}{1 + D_1 k^2 + D_2 k^4 + D_3 k^6}, \qquad \text{(D.48)}$$

dove i coefficienti sono riportati in tabella D.1.

Coefficiente	Valore	Coefficiente	Valore
$L_{kur,0}$	$1.2260172\ 10^{-1}$	C_3	$-1.8446680\ 10^{-6}$
A_0	$4.8860251\ 10^{-1}$	D_1	$8.2325617\ 10^{-2}$
A_1	$4.4493076\ 10^{-3}$	D_2	$4.2681448\ 10^{-3}$
A_2	$8.8027039\ 10^{-4}$	D_3	$1.1653690\ 10^{-4}$
A_3	$1.1507084\ 10^{-6}$	E_0	2.0466534
B_1	$6.4662924\ 10^{-2}$	E_1	-3.6544371
B_2	$3.3090406\ 10^{-3}$	E_2	1.8396733
B_3	$7.4290680\ 10^{-5}$	E_3	-0.20360244
C_0	$1.8756590\ 10^{-1}$	F_1	-2.0182173
C_1	$-2.5352147\ 10^{-3}$	F_2	1.2420401
C_2	$2.6995102\ 10^{-4}$	F_3	-0.21741801

Tabella D.1. Coefficienti di approssimazione per le equazioni (D.47), (D.48), (D.49).

Parametri

$$k \approx -L_{ca} \frac{E_0 + E_1 L_{ca}^2 + E_2 L_{ca}^4 + E_3 L_{ca}^6}{1 + F_1 L_{ca}^2 + F_2 L_{ca}^4 + F_3 L_{ca}^6}, \qquad \text{(D.49)}$$

$$\alpha = \frac{L_{cv} e^{-k^2/2} k}{1 - 2\Phi\left(-k/\sqrt{2}\right)}, \qquad \text{(D.50)}$$

$$\xi = 1 - \frac{\alpha}{k}\left(1 - e^{k^2/2}\right). \qquad \text{(D.51)}$$

D.6. Distribuzione Gamma

La distribuzione Gamma, o Pearson tipo III, è stata largamente utilizzata nel campo idrologico, ad esempio per la descrizione di grandezze quali portate massime e minime annue, volumi idrici stagionali e annuali ed anche gli eventi estremi di precipitazione. Nel caso in esame si fa riferimento alla distribuzione Gamma con tre parametri, che genera una distribuzione asimmetrica che può essere limitata superiormente o inferiormente a seconda che il valore del parametro di scala sia rispettivamente negativo o positivo.

Definizioni

Parametri (3): ξ (posizione), β (scala), α (forma).
Il legame con i momenti ordinari μ, σ e γ è dato da $\alpha = 4/\gamma^2$, $\beta = 1/2\sigma|\gamma|$ e $\xi = \mu - 2\sigma/\gamma$. Se $\beta > 0$, allora il campo di esistenza di x è $\xi \le x < \infty$, mentre se $\beta < 0$ è $-\infty < x \le \xi$.

$$f(x) = \frac{1}{|\beta|\Gamma(\alpha)}\left(\frac{x-\xi}{\beta}\right)^{\alpha-1} e^{-(x-\xi)/\beta}, \tag{D.52}$$

$$F(x)\begin{cases} G\left(\alpha, \frac{x-\xi}{\beta}\right)\Big/\Gamma(\alpha) & ,\beta > 0 \\ 1 - G\left(\alpha, \frac{x-\xi}{\beta}\right)\Big/\Gamma(\alpha), & \beta < 0 \end{cases}, \tag{D.53}$$

dove $\Gamma(.)$ è la funzione gamma:

$$\Gamma(x) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt\,, \tag{D.54}$$

mentre G(α, x) è la funzione gamma incompleta, definita come:

$$G(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t}\, dt\,. \tag{D.55}$$

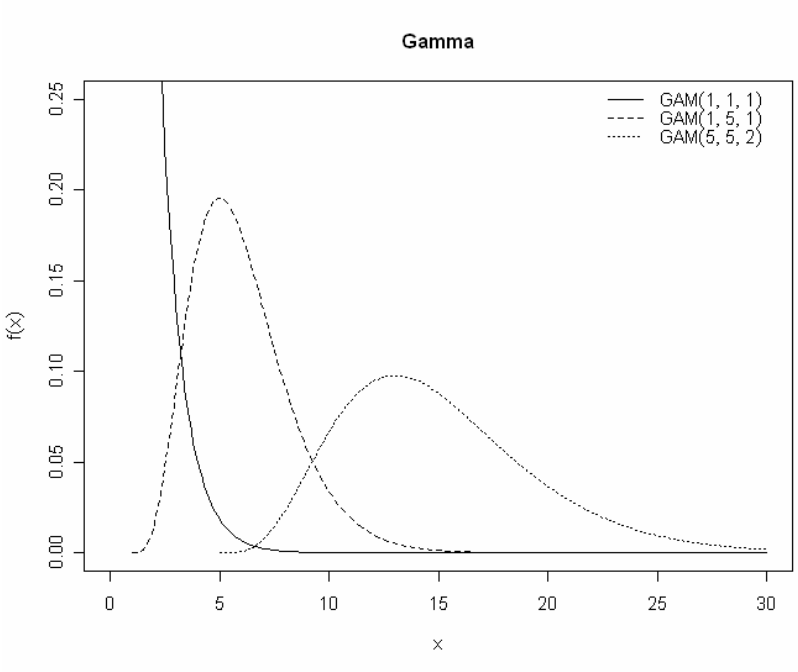


Figura D.6. Esempi di densità di probabilità della distribuzione gamma con parametri: (1) $\xi = 1$, $\alpha = 1$, $\beta = 1$; (2) $\xi = 1$, $\alpha = 2$, $\beta = 1$; (3) $\xi = 5$, $\alpha = 5$, $\beta = 2$.

L-coefficienti

Gli L-coefficienti sono definiti per $0<\alpha<\infty$.

$$L_{cv} = \frac{\beta\Gamma(\alpha+1/2)}{\sqrt{\pi}(\xi+\alpha\beta)\Gamma(\alpha)}, \tag{D.56}$$

$$L_{ca} = 6I_{1/3}(\alpha,2\alpha)-3, \tag{D.57}$$

dove $I_x(p,q)$ è la funzione beta incompleta:

$$I_x(p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\int_0^x t^{p-1}(1-t)^{q-1}\,dt\,. \tag{D.58}$$

Per il calcolo del coefficiente di L-kurtosis non esiste un'espressione semplice. Tuttavia è possibile utilizzare delle approssimazioni con accuratezza di 10^{-6}, esprimendo sia L_{ca} che L_{kur} in funzione di α:

$$L_{ca} \approx \alpha^{-0.5} \frac{A_0 + A_1\alpha^{-1} + A_2\alpha^{-2} + A_3\alpha^{-3}}{1 + B_1\alpha^{-1} + B_2\alpha^{-2}}, \quad \text{se } \alpha \geq 1 \qquad \text{(D.59)}$$

$$L_{kur} \approx \frac{C_0 + C_1\alpha^{-1} + C_2\alpha^{-2} + C_3\alpha^{-3}}{1 + D_1\alpha^{-1} + D_2\alpha^{-2}}, \qquad \text{se } \alpha \geq 1 \qquad \text{(D.60)}$$

$$L_{ca} \approx \frac{1 + E_1\alpha + E_2\alpha^2 + E_3\alpha^3}{1 + F_1\alpha + F_2\alpha^2 + F_3\alpha^3}, \qquad \text{se } \alpha < 1 \qquad \text{(D.61)}$$

$$L_{kur} \approx \frac{1 + G_1\alpha + G_2\alpha^2 + G_3\alpha^3}{1 + H_1\alpha + H_2\alpha^2 + H_3\alpha^3}, \qquad \text{se } \alpha < 1 \qquad \text{(D.62)}$$

dove i coefficienti di approssimazione assumono i valori riportati in Tabella D.2.

Coefficiente	Valore	Coefficiente	Valore
A_0	$3.2573501\ 10^{-1}$	C_0	$1.2260172\ 10^{-1}$
A_1	$1.6869150\ 10^{-1}$	C_1	$5.3730130\ 10^{-2}$
A_2	$7.8327243\ 10^{-2}$	C_2	$4.3384378\ 10^{-2}$
A_3	$-2.9120539\ 10^{-3}$	C_3	$1.1101277\ 10^{-2}$
B_1	$4.6697102\ 10^{-1}$	D_1	$1.8324466\ 10^{-1}$
B_2	$2.4255406\ 10^{-1}$	D_2	$2.0166036\ 10^{-1}$
E_1	2.3807576	G_1	2.1235833
E_2	1.5931792	G_2	4.1670213
E_3	$1.1618371\ 10^{-1}$	G_3	3.1925299
F_1	5.1533299	H_1	9.0551443
F_2	7.1425260	H_2	$2.6649995\ 10^{-1}$
F_3	1.9745056	H_3	$2.6193668\ 10^{-1}$

Tabella D.2. Coefficienti di approssimazione per le equazioni (D.59) – (D.62).

Parametri

La stima del parametro α viene effettuata mediante una relazione con precisione di 5 x 10^{-5}, assumendo che $z = 3\pi L_{ca}^2$, se $0 < |L_{ca}| < 1/3$:

$$\alpha \approx \frac{1 + 0.2906z}{z + 0.1882z^2 + 0.0442z^3}, \tag{D.63}$$

invece se $1/3 \le L_{ca} < 1$ si assume che $z = 1 - |L_{ca}|$ e si utilizza:

$$\alpha \approx \frac{0.36067z - 0.59567z^2 + 0.25361z^3}{1 - 2.78861z + 2.56096z^2 - 0.77045z^3}. \tag{D.64}$$

Noto il parametro α è possibile ricavare gli altri parametri come:

$$\gamma = 2\alpha^{-1/2} sign(L_{ca}), \tag{D.65}$$

$$\sigma = L_{cv}(\xi + \alpha\beta)\pi^{1/2}\alpha^{1/2}\,\Gamma(\alpha)/\Gamma(\alpha + 1/2), \tag{D.66}$$

$$\mu = \frac{\pi^{-1/2}\beta\,\Gamma(\alpha + 1/2)/\Gamma(\alpha)}{L_{cv}}. \tag{D.67}$$

D.7. Distribuzione Kappa

La distribuzione a quattro parametri Kappa ha avuto numerosi utilizzi in campo pratico. In particolare è utilizzata frequentemente quando le distribuzioni a tre parametri non danno un'adeguata rappresentazione dei dati, oppure in studi sulla robustezza.

Definizioni

Parametri (4): ξ (posizione), α (scala), k, h.

Campo di esistenza di x: il limite superiore è ξ + α/k se k>0, ∞ se k≤0; il limite inferiore è ξ + α (1 − h−k)/k se h>0, ξ + α/k se h≤0 e k<0, e −∞ se h≤0 e k≥0.

$$f(x)=\alpha^{-1}\left[1-k\left(x-\xi\right)/\alpha\right]^{1/k-1}\left[F(x)\right]^{1-h}, \qquad \text{(D.68)}$$

$$F(x)=\left\{1-h\left[1-k\left(x-\xi\right)/\alpha\right]^{1/k}\right\}^{1/h}, \qquad \text{(D.69)}$$

$$x(F)=\xi+\frac{\alpha}{k}\left[1-\left(\frac{1-F^{h}}{h}\right)^{k}\right]. \qquad \text{(D.70)}$$

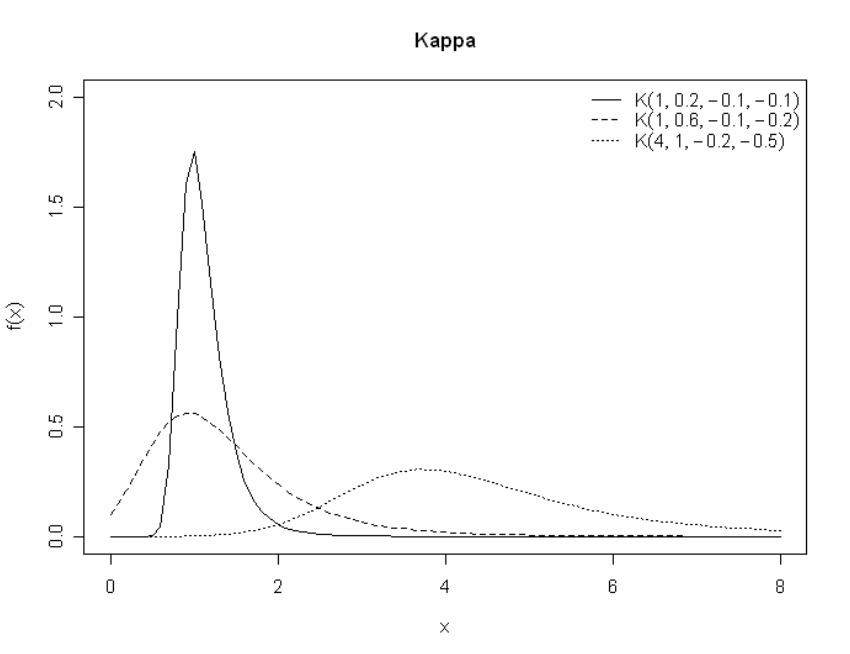


Figura D.7. Esempi di densità di probabilità della distribuzione kappa con parametri: (1) ξ = 1, α = 0.2, k = -0.1, h = -0.1; (2) ξ = 1, α = 0.6, k = -0.1, h= -0.2; (3) ξ = 4, α = 1, k = -0.2, h = -0.5.

L-coefficienti

Gli L-coefficienti sono definiti se $h\geq 0$ e $k>-1$, oppure se $h<0$ e $-1<k<-1/h$.

$$L_{cv} = \frac{\alpha(g_1 - g_2)}{k\xi + \alpha(1 - g_1)}, \tag{D.71}$$

$$L_{ca} = \frac{-g_1 + 3g_2 - 2g_3}{g_1 - g_2}, \tag{D.72}$$

$$L_{kur} = -\frac{-g_1 + 6g_2 - 10g_3 + 5g_4}{g_1 - g_2}, \tag{D.73}$$

dove

$$g_r = \begin{cases} \dfrac{r\Gamma(1+k)\Gamma(r/h)}{h^{1+k}\Gamma(1+k+r/h)}, & h > 0 \\ \dfrac{r\Gamma(1+k)\Gamma(-k-r/h)}{(-h)^{1+k}\Gamma(1-r/h)}, & h < 0 \end{cases}. \tag{D.74}$$

Parametri

Per ricavare i parametri a partire dagli L-coefficienti non ci sono semplici espressioni. Comunque è possibile ricavare *k* e *h* in funzione di L_{ca} ed L_{kur} con il metodo numerico di Newton – Raphson a partire dalle equazioni (D.72) e (D.73).

D.8. Distribuzione del valore estremo a doppia componente TCEV

Il modello TCEV (Two Component Estreme Value distribution), sviluppato da *Fiorentino et al*. (1987a), ipotizza che i massimi annuali della variabile idrologica considerata provengano da una miscela di due diverse popolazioni di eventi: una "ordinaria" e l'altra "straordinaria", rappresentata dagli outliers presenti nella serie storica. Il modello infatti assume che le portate estreme non siano originate dallo stesso gruppo di funzioni di distribuzione relativo alle portate ordinarie; sotto tale ipotesi il modello assume che la densità di probabilità e la funzione di ripartizione siano costituite da una miscela di esponenziali.

Definizioni

Parametri (4): θ_1, θ_2 (posizione), Λ_1, Λ_2 (scala).

$$f(x)=\left(\frac{\Lambda_1}{\theta_1}e^{-x/\theta_1}+\frac{\Lambda_2}{\theta_2}e^{-x/\theta_2}\right)\exp\left(-\Lambda_1 e^{-x/\theta_1}-\Lambda_2 e^{-x/\theta_2}\right), \qquad \text{(D.75)}$$

$$F(x)=\exp\left(-\Lambda_1 e^{-x/\theta_1}-\Lambda_2 e^{-x/\theta_2}\right). \qquad \text{(D.76)}$$

Il quantile *x(F)* non è esprimibile tramite espressioni analitiche dirette, poiché la relazione *F(x)* non è direttamente invertibile. Tuttavia, per *F > 0.9* è possibile ricorrere a un'espressione asintotica:

$$x(F)=\theta_2\ln\left(\frac{\Lambda_2}{\Lambda_1^{\theta_1/\theta_2}}\right)+\theta_2\ln\left(\frac{1}{1-F}\right)+\theta_1\ln(\Lambda_1). \qquad \text{(D.77)}$$

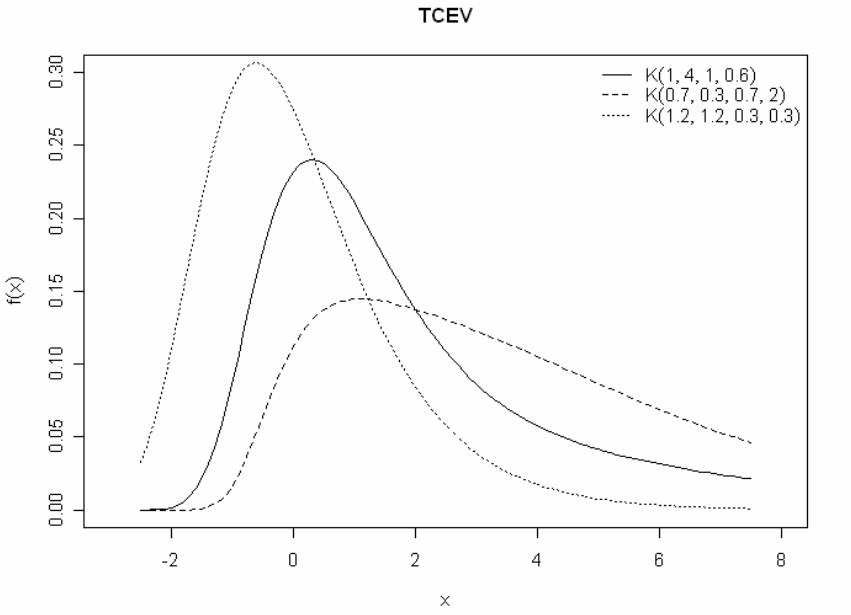


Figura D.8. Esempi di densità di probabilità della distribuzione TCEV con parametri: (1) θ_1 = 1, θ_2 = 4, Λ_1 = 1, Λ_2 = 0.6; (2) θ_1 = 0.7, θ_2 = 0.3, Λ_1 = 0.7, Λ_2 = 2; (3) θ_1 = 1.2, θ_2 = 1.2, Λ_1 = 0.3, Λ_2 = 0.3.

L-coefficienti

Non esistono relazioni semplici per il calcolo degli L-coefficienti. Questi vengono derivati dai momenti pesati in probabilità b_r (*Beran et al.*, 1986):

$$L_{cv} = \frac{2 \cdot b_1}{b_0} - 1 , \tag{D.78}$$

$$L_{ca} = \frac{2 \cdot (3b_2 - b_0)}{2b_1 - b_0} - 3 , \tag{D.79}$$

$$L_{kur} = \frac{5 \cdot [2 \cdot (2b_3 - 3b_2) + b_0]}{2b_1 - b_0} + 6 . \tag{D.80}$$

Il valore dei momenti pesati in probabilità b_r di ordine $r = 0,\ldots,3$ viene determinato come:

$$\frac{\theta_1}{r+1}[\gamma + \log \Lambda_1 + \log(r+1)], \tag{D.81}$$

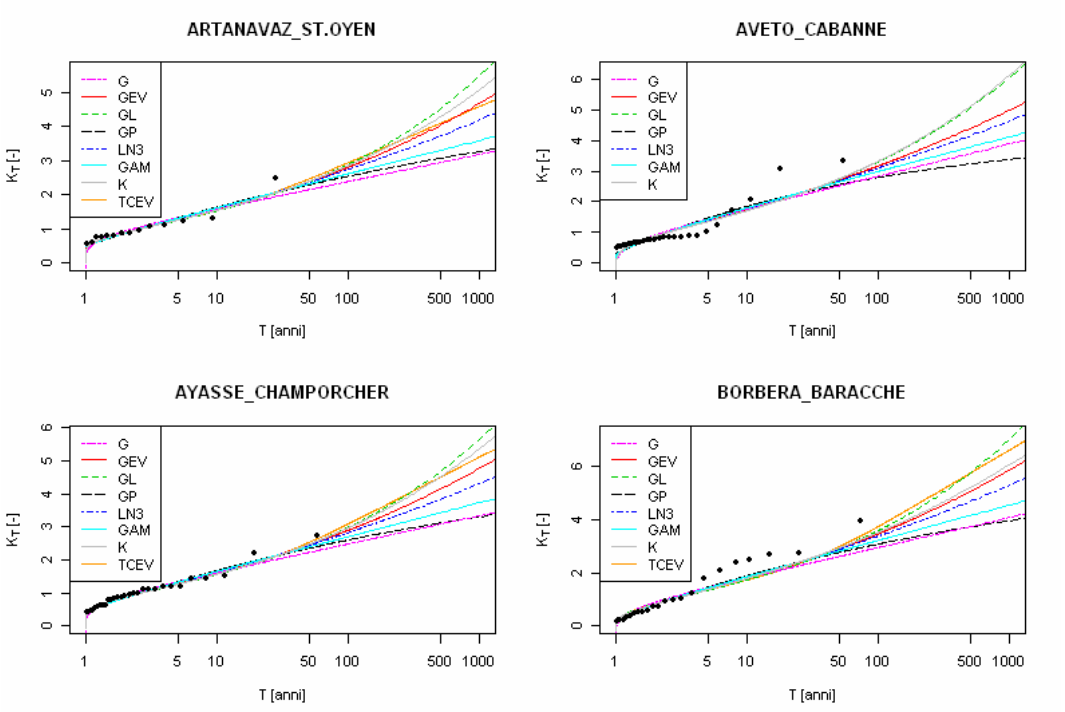
dove $\gamma = -\int_0^\infty e^z \log z \cdot dz$, assumendo $z = \Lambda_1 e^{-x/\theta_1}$.

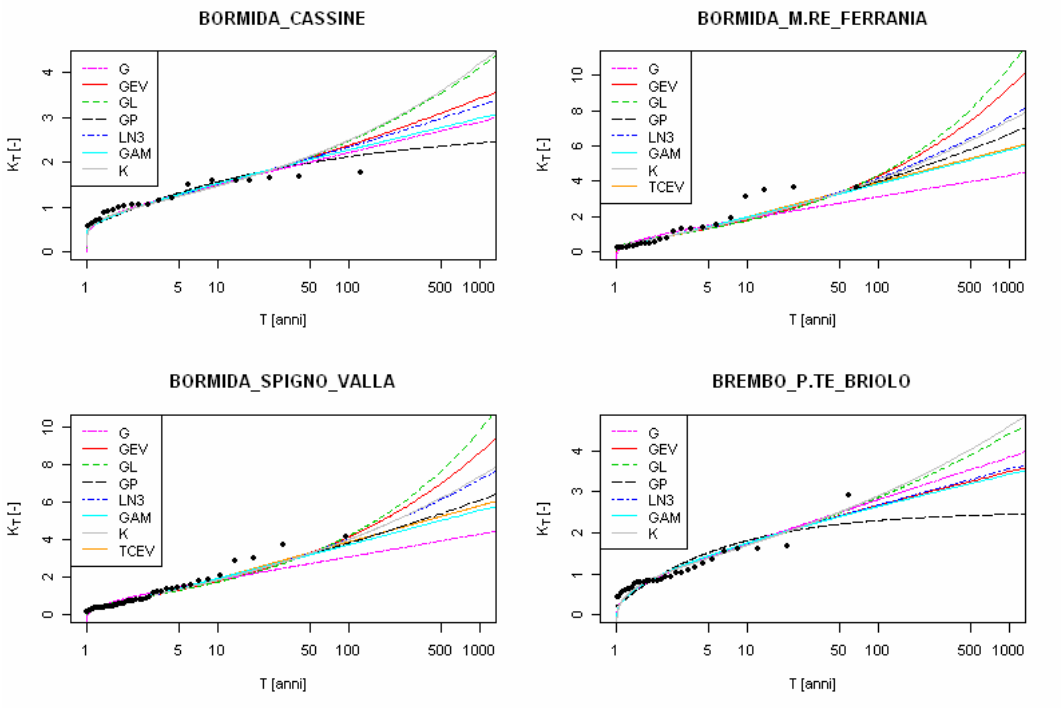
Parametri

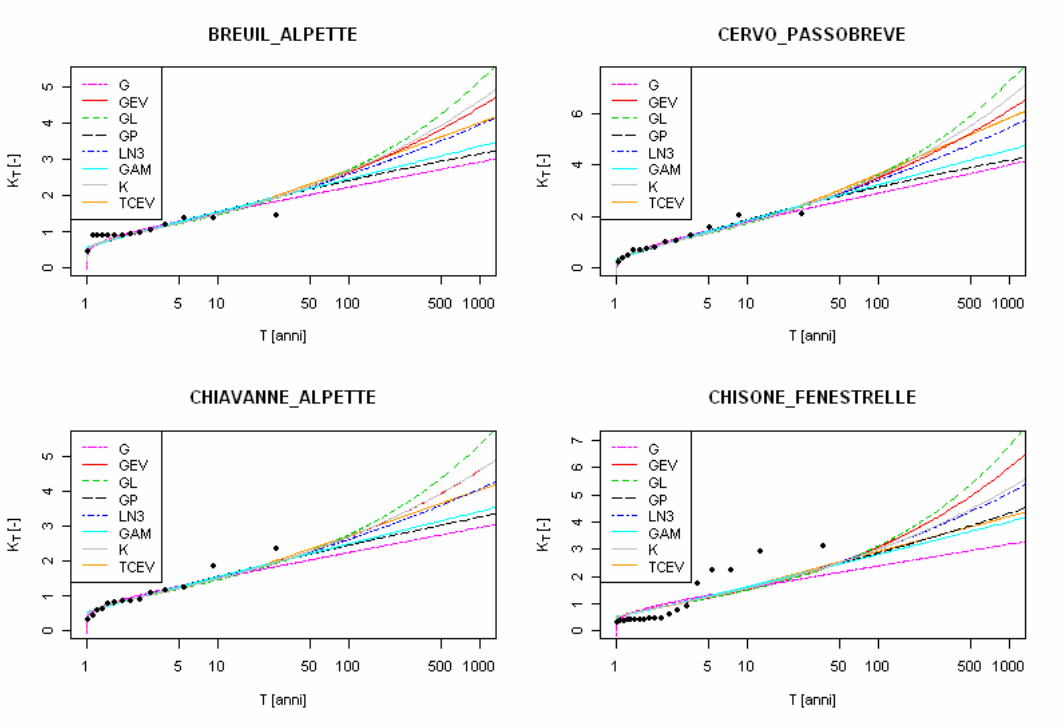
La valutazione dei parametri della TCEV deve essere effettuata ricorrendo alla procedura di stima gerarchica proposta da *Fiorentino et al.* (1987a), la quale prevede che la stima dei parametri venga fatta per fasi successive, secondo un criterio gerarchico al quale corrispondono altrettante scale spaziali. In particolare ciascun livello di regionalizzazione è basato su ipotesi differenti e sviluppa procedimenti piuttosto articolati. Per questo motivo si rimanda a *Fiorentino et al.* (1987a) per ulteriori approfondimenti.

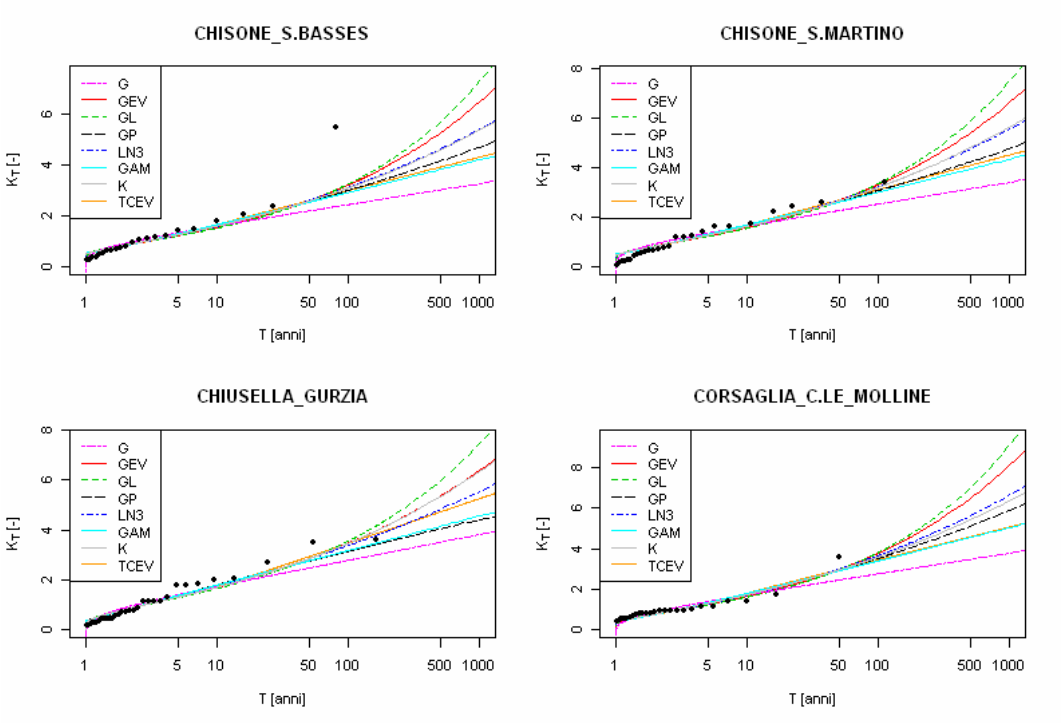
Allegato I.
Confronto tra curve di crescita regionali (8 distribuzioni)

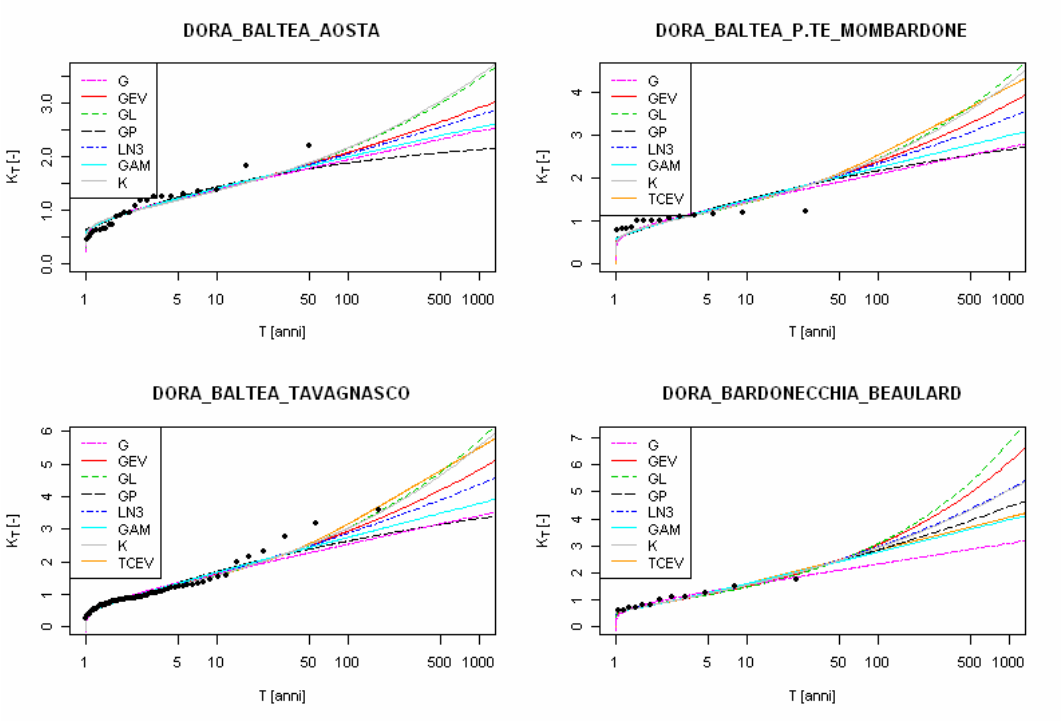
In questo allegato vengono confrontate, per ciascuna sezione considerata, le curve di crescita definite secondo le diverse distribuzioni testate. Dall'analisi dei grafici è possibile verificare l'adattamento delle diverse distribuzioni ai dati osservati, che vengono inseriti nel grafico seguendo il criterio di Hirsch, adottando la plotting position di Hazen. I parametri di ciascuna distribuzione vengono stimati da L_{cv} ed L_{ca} calcolati tramite i modelli di regressione multipla (7.1) e (7.2).

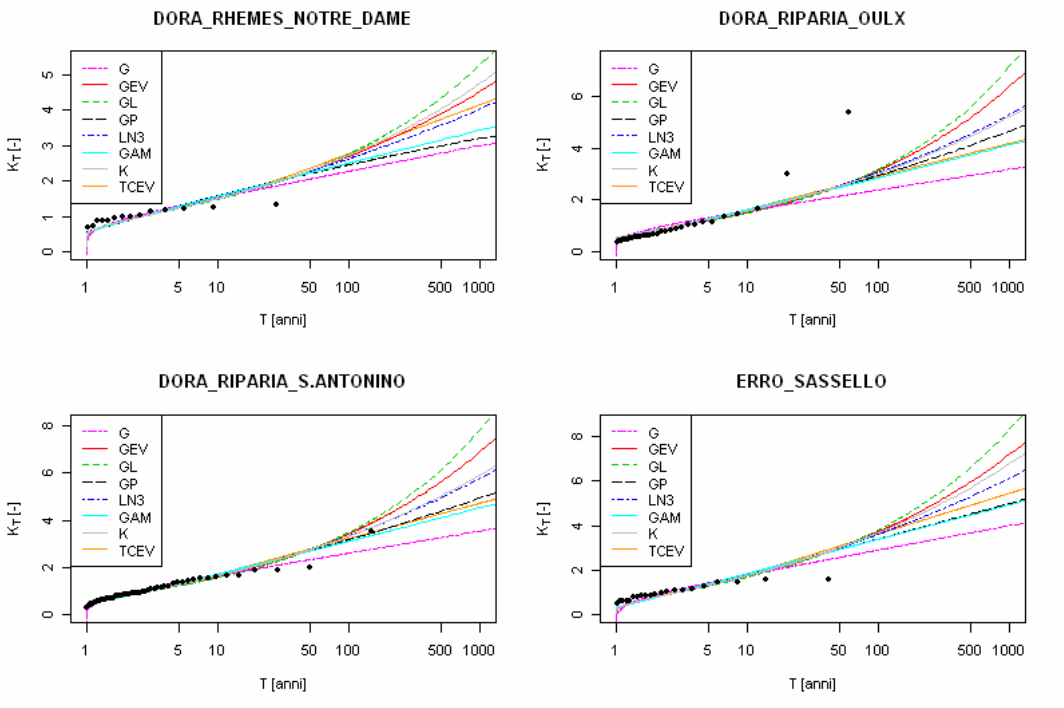


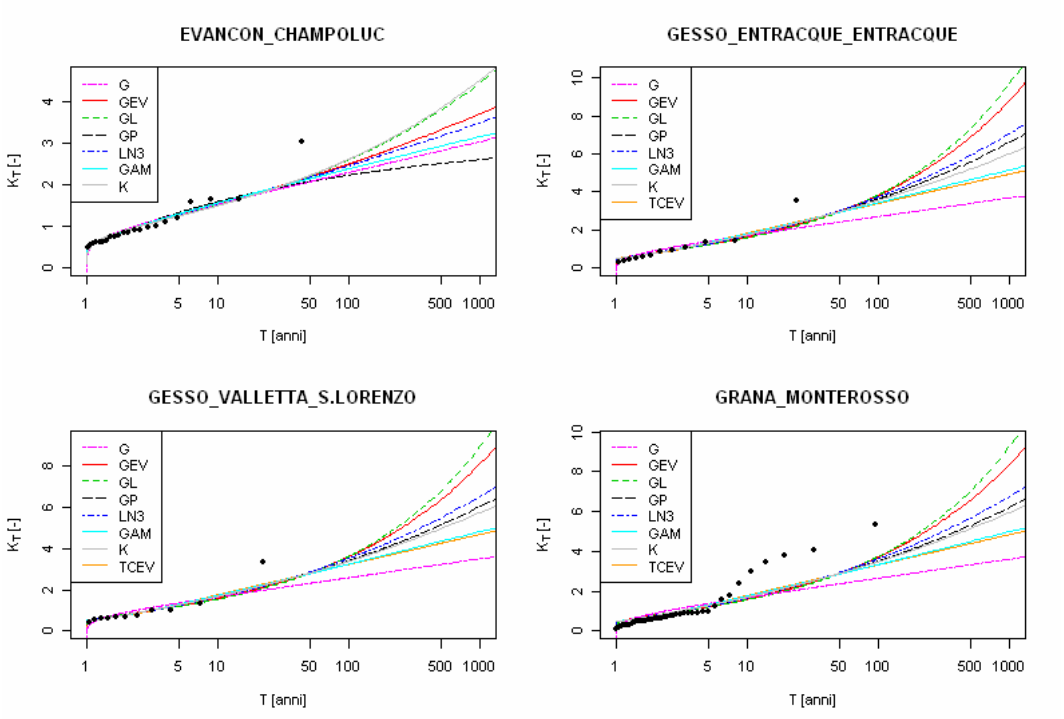

BORMIDA_CASSINE
BORMIDA_M.RE_FERRANIA
BORMIDA_SPIGNO_VALLA
BREMBO_P.TE_BRIOLO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

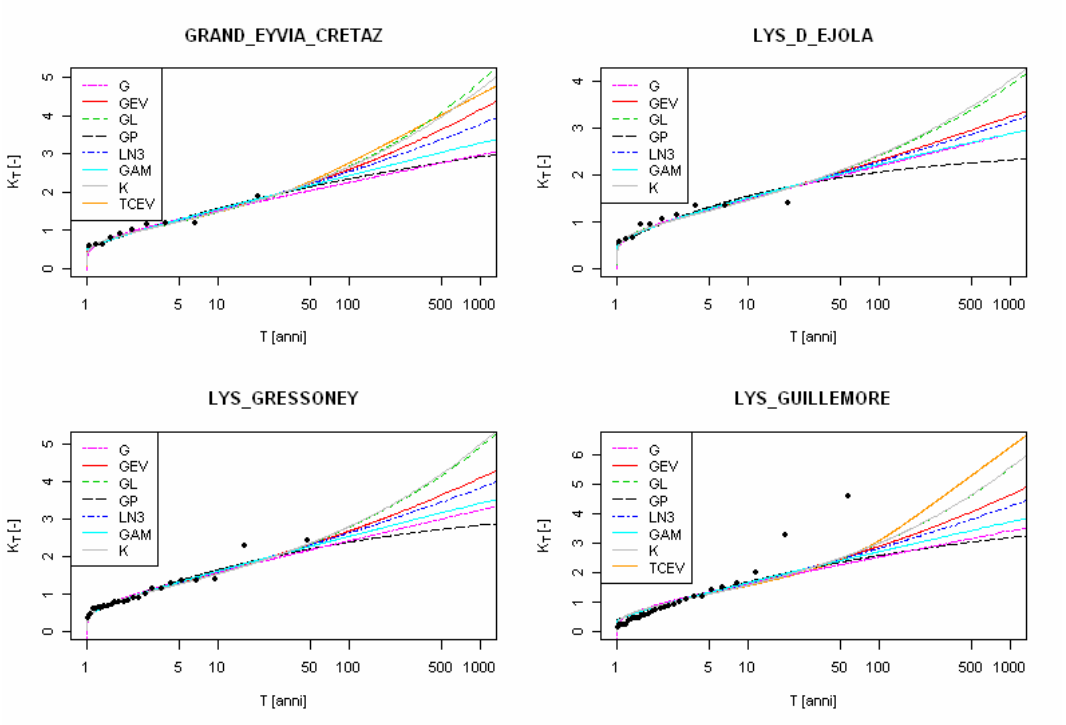

BREUIL_ALPETTE
CERVO_PASSOBREVE
CHIAVANNE_ALPETTE
CHISONE_FENESTRELLE
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

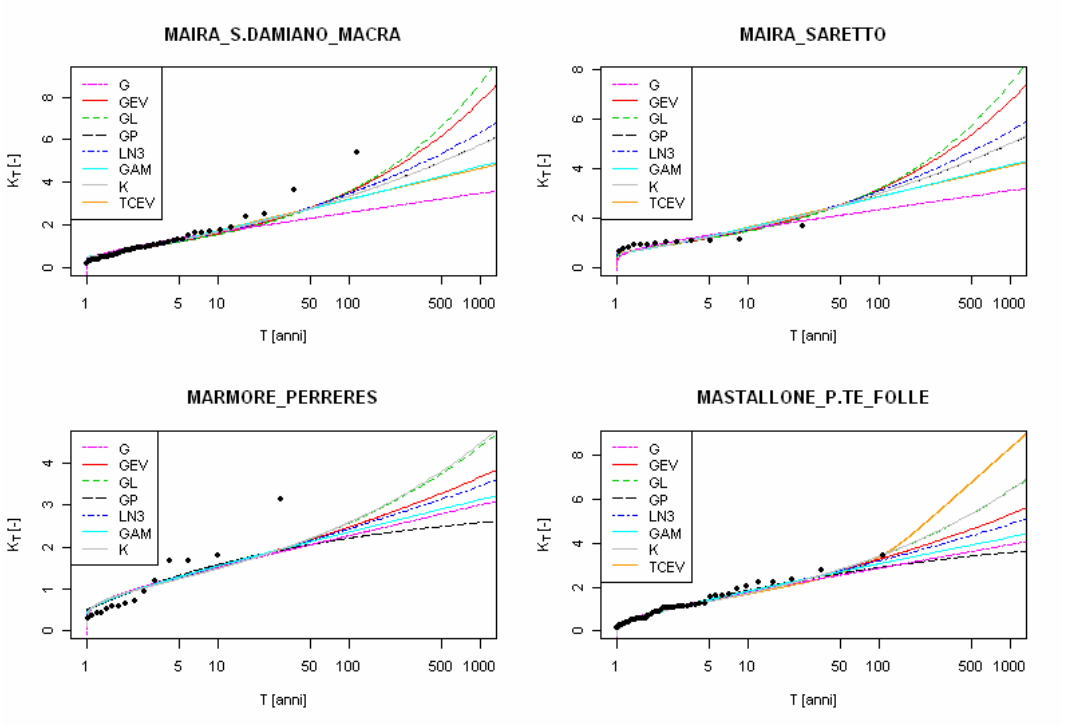

CHISONE_S.BASSES
CHISONE_S.MARTINO
CHIUSELLA_GURZIA
CORSAGLIA_C.LE_MOLLINE
G
GEV
GL
GP
LN3
GAM
K
TCEV
K_T [-]
T [anni]
1
5
10
50
100
500
1000

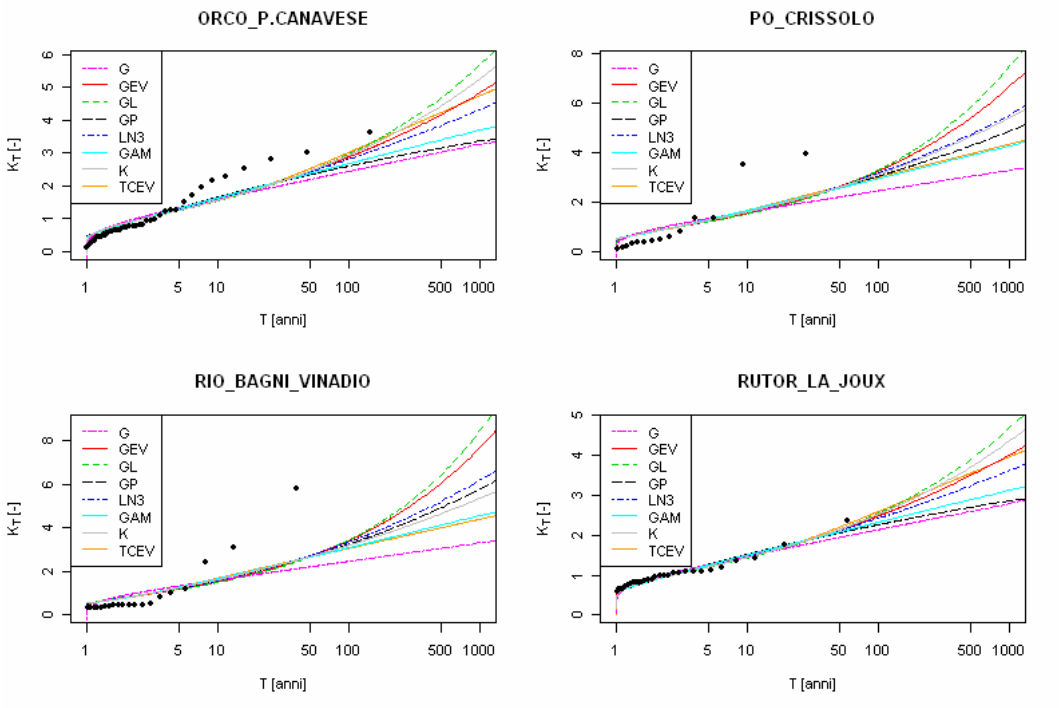

DORA_BALTEA_AOSTA
DORA_BALTEA_P.TE_MOMBARDONE
DORA_BALTEA_TAVAGNASCO
DORA_BARDONECCHIA_BEAULARD
G
GEV
GL
GP
LN3
GAM
K
TCEV
K_T [-]
T [anni]
1
5
10
50
100
500
1000

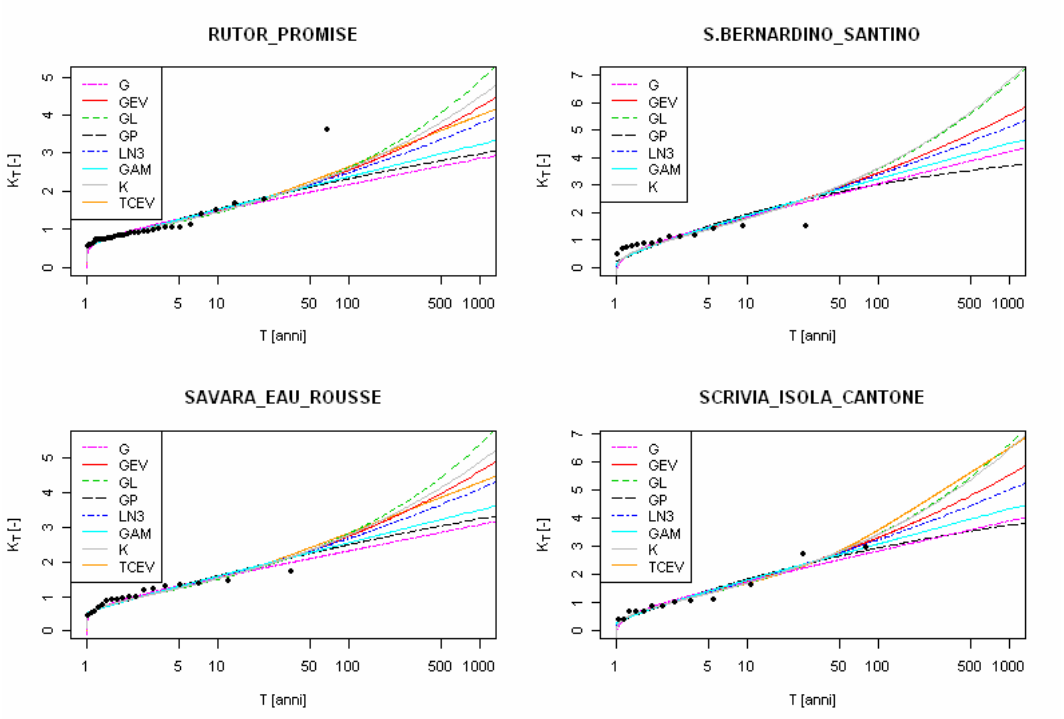

DORA_RHEMES_NOTRE_DAME
DORA_RIPARIA_OULX
DORA_RIPARIA_S.ANTONINO
ERRO_SASSELLO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

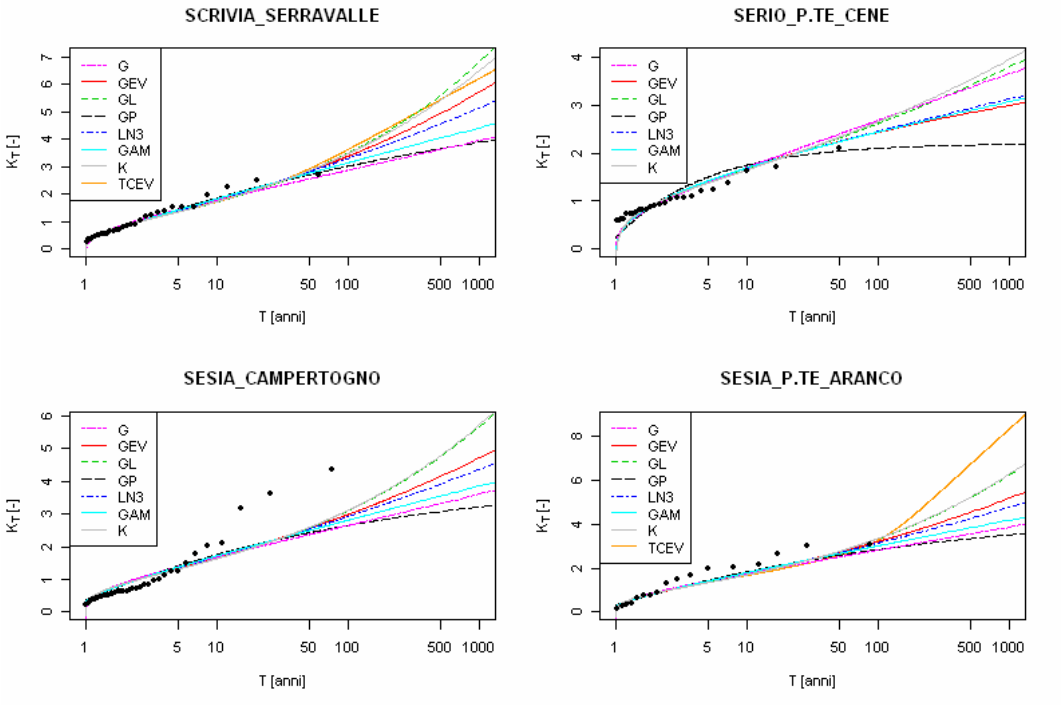

EVANCON_CHAMPOLUC
GESSO_ENTRACQUE_ENTRACQUE
GESSO_VALLETTA_S.LORENZO
GRANA_MONTEROSSO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

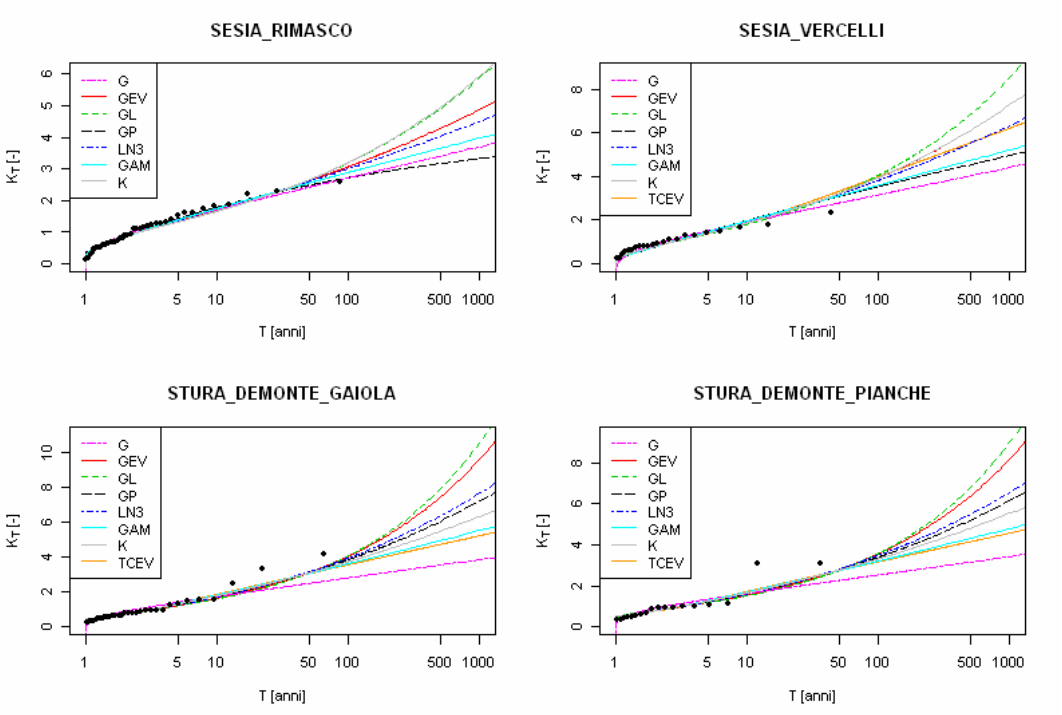

GRAND_EYVIA_CRETAZ
LYS_D_EJOLA
LYS_GRESSONEY
LYS_GUILLEMORE
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

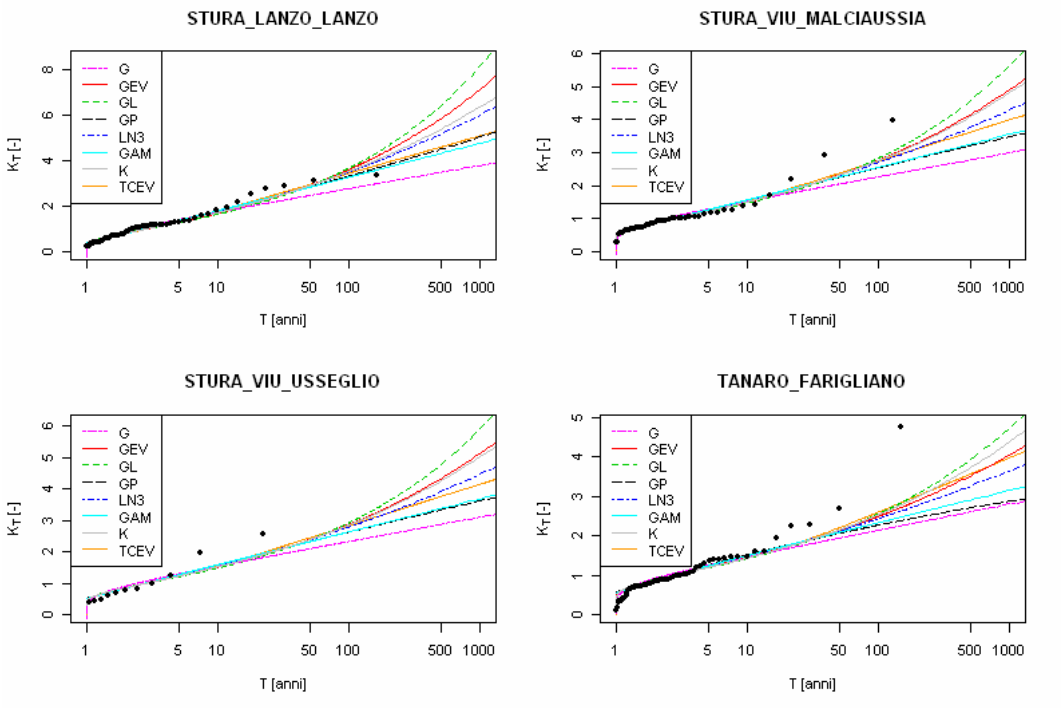

MAIRA_S.DAMIANO_MACRA
MAIRA_SARETTO
MARMORE_PERRERES
MASTALLONE_P.TE_FOLLE
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

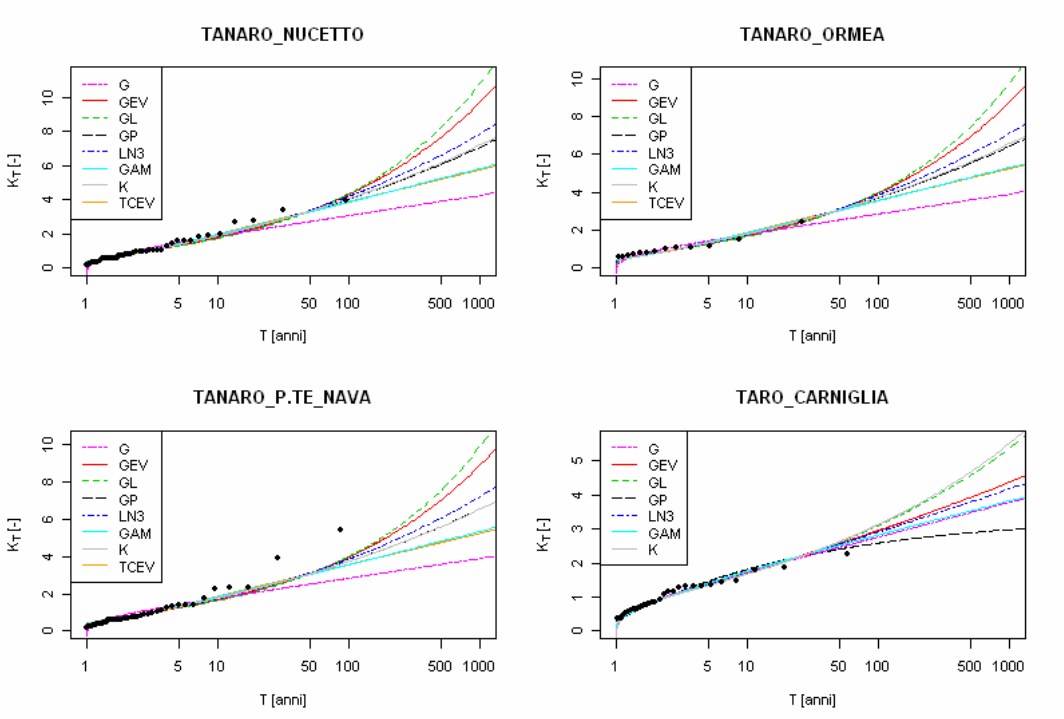

ORCO_P.CANAVESE
PO_CRISSOLO
RIO_BAGNI_VINADIO
RUTOR_LA_JOUX
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

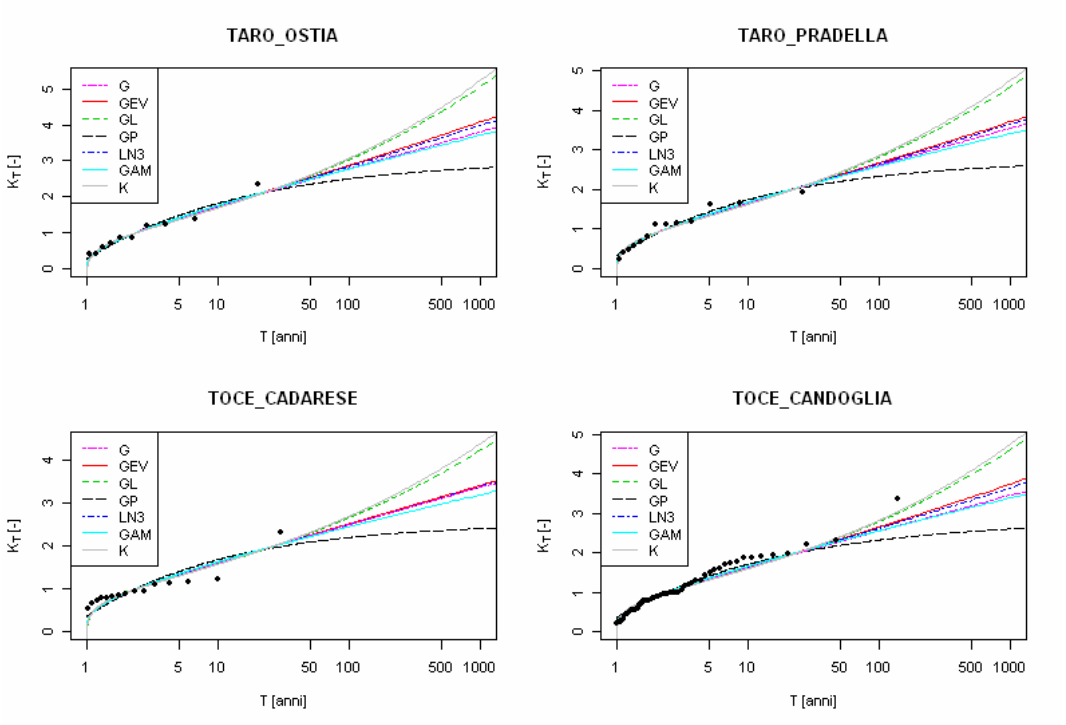

RUTOR_PROMISE
S.BERNARDINO_SANTINO
SAVARA_EAU_ROUSSE
SCRIVIA_ISOLA_CANTONE
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

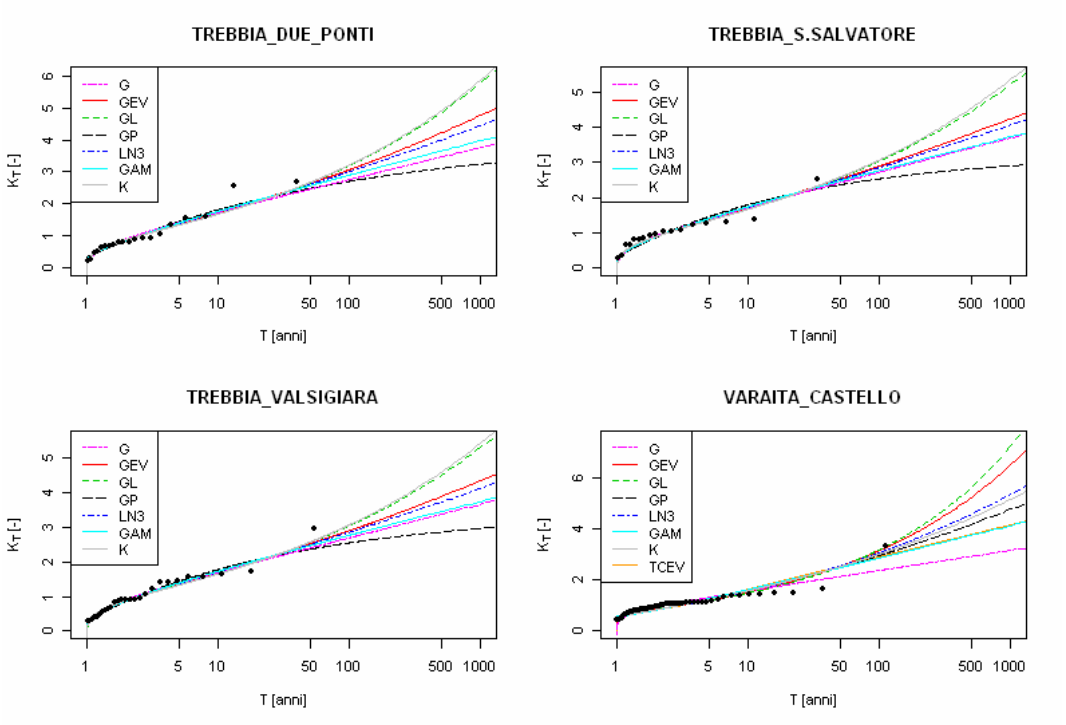

SCRIVIA_SERRAVALLE
SERIO_P.TE_CENE
SESIA_CAMPERTOGNO
SESIA_P.TE_ARANCO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

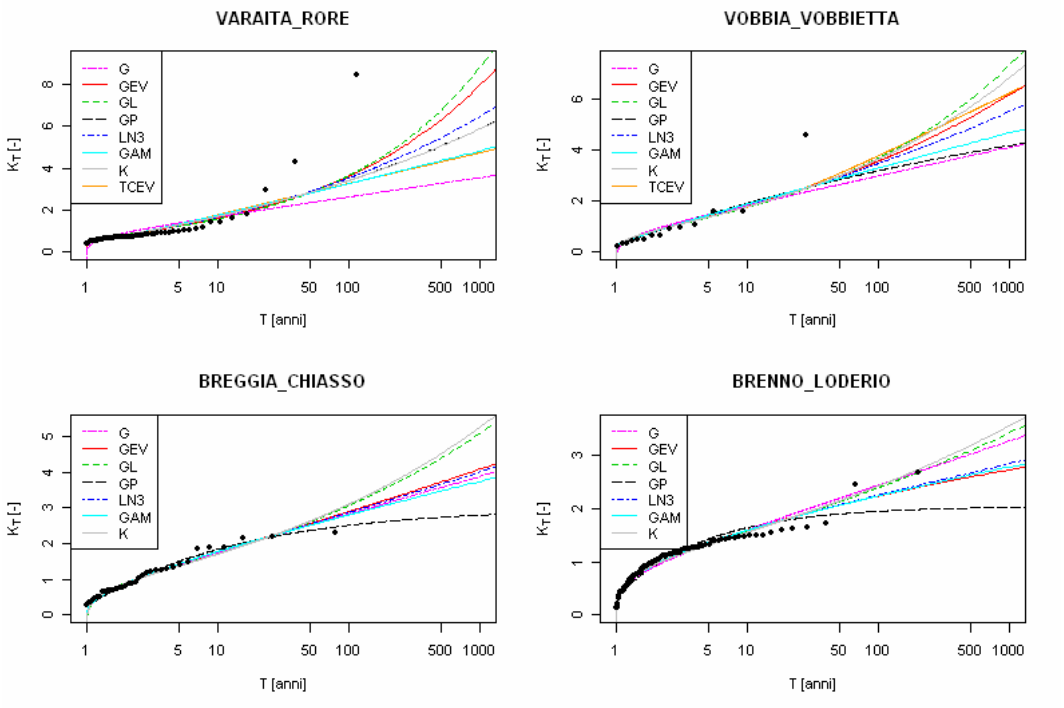

SESIA_RIMASCO
SESIA_VERCELLI
STURA_DEMONTE_GAIOLA
STURA_DEMONTE_PIANCHE
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

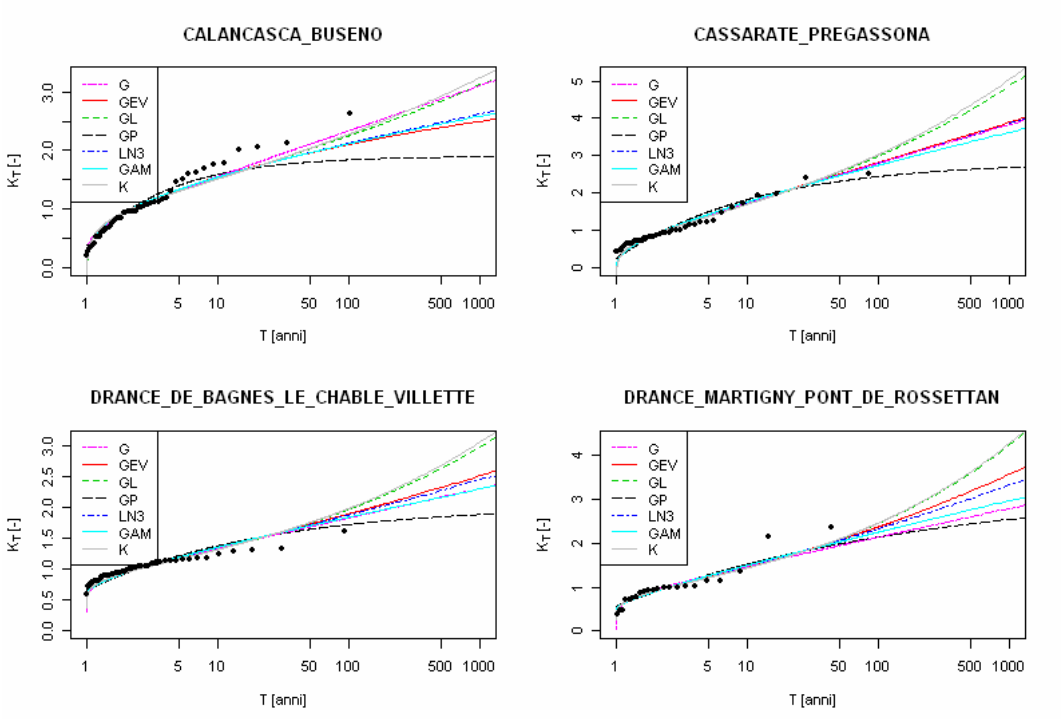

STURA_LANZO_LANZO
STURA_VIU_MALCIAUSSIA
STURA_VIU_USSEGLIO
TANARO_FARIGLIANO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

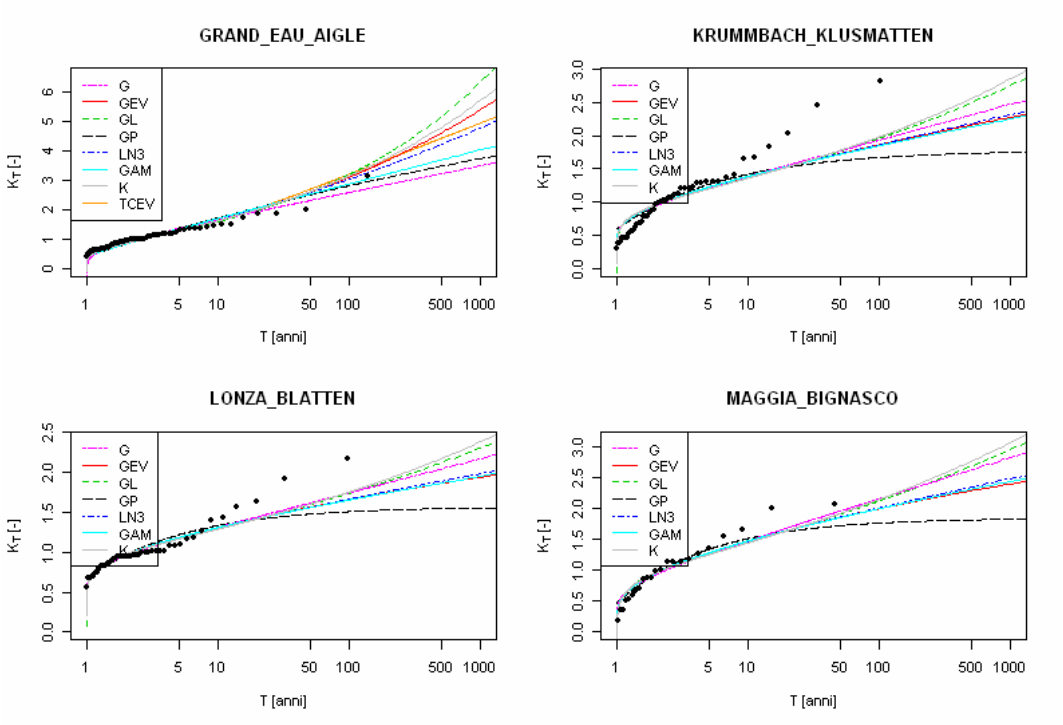

TANARO_NUCETTO
TANARO_ORMEA
TANARO_P.TE_NAVA
TARO_CARNIGLIA
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

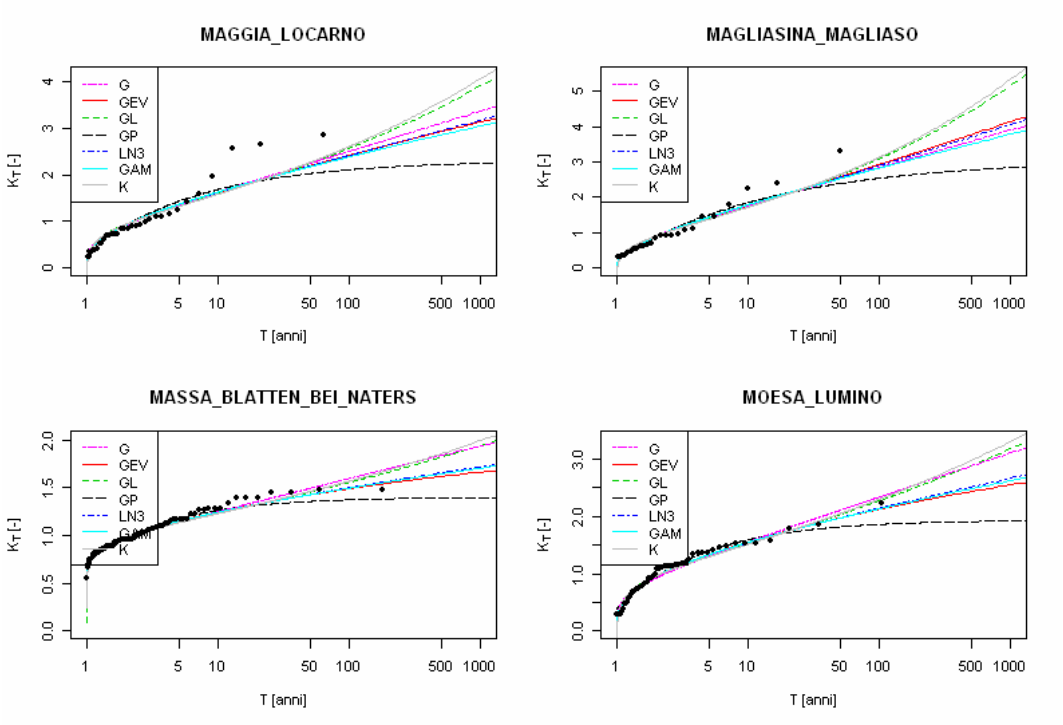

TARO_OSTIA
TARO_PRADELLA
TOCE_CADARESE
TOCE_CANDOGLIA
G
GEV
GL
GP
LN3
GAM
K
K_T [-]
T [anni]
1
5
10
50
100
500
1000

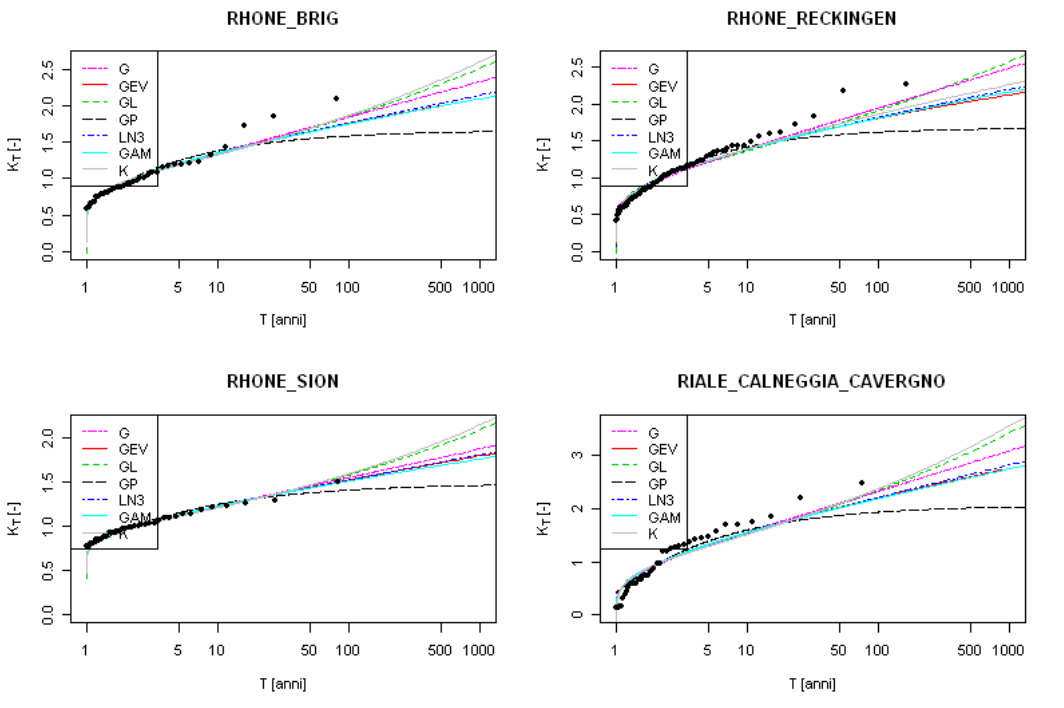

TREBBIA_DUE_PONTI
TREBBIA_S.SALVATORE
TREBBIA_VALSIGIARA
VARAITA_CASTELLO
G
GEV
GL
GP
LN3
GAM
K
TCEV
K_T [-]
T [anni]
1
5
10
50
100
500
1000

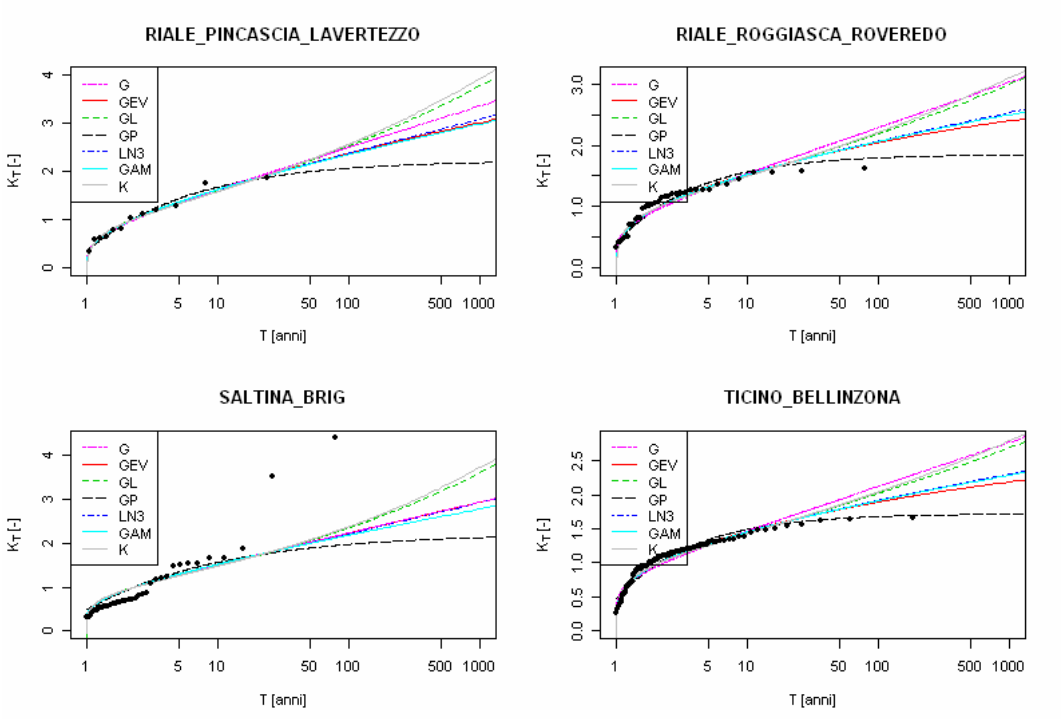

VARAITA_RORE
VOBBIA_VOBBIETTA
BREGGIA_CHIASSO
BRENNO_LODERIO
G
GEV
GL
GP
LN3
GAM
K
TCEV
KT [-]
T [anni]
1
5
10
50
100
500
1000

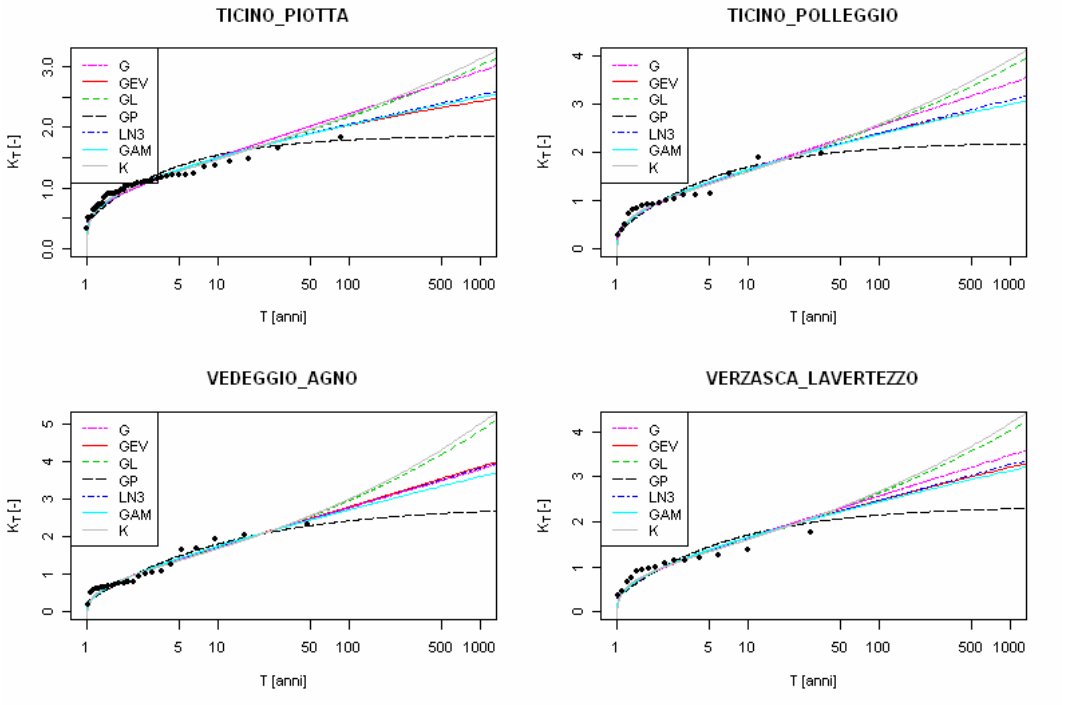

CALANCASCA_BUSENO
CASSARATE_PREGASSONA
DRANCE_DE_BAGNES_LE_CHABLE_VILLETTE
DRANCE_MARTIGNY_PONT_DE_ROSSETTAN
G
GEV
GL
GP
LN3
GAM
K
KT [-]
T [anni]
1
5
10
50
100
500
1000

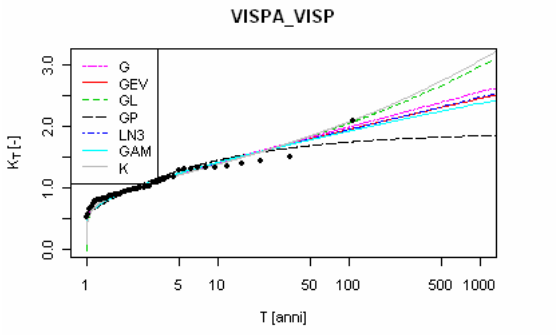

GRAND_EAU_AIGLE
KRUMMBACH_KLUSMATTEN
LONZA_BLATTEN
MAGGIA_BIGNASCO
G
GEV
GL
GP
LN3
GAM
K
TCEV
K_T [-]
T [anni]
1
5
10
50
100
500
1000


MAGGIA_LOCARNO
MAGLIASINA_MAGLIASO
MASSA_BLATTEN_BEI_NATERS
MOESA_LUMINO
G
GEV
GL
GP
LN3
GAM
K
K_T [-]
T [anni]
1
5
10
50
100
500
1000


RHONE_BRIG
RHONE_RECKINGEN
RHONE_SION
RIALE_CALNEGGIA_CAVERGNO
G
GEV
GL
GP
LN3
GAM
K
KT [-]
T [anni]
1
5
10
50
100
500
1000


RIALE_PINCASCIA_LAVERTEZZO
RIALE_ROGGIASCA_ROVEREDO
SALTINA_BRIG
TICINO_BELLINZONA
G
GEV
GL
GP
LN3
GAM
K
KT [-]
T [anni]
1
5
10
50
100
500
1000


TICINO_PIOTTA
TICINO_POLLEGGIO
VEDEGGIO_AGNO
VERZASCA_LAVERTEZZO
G
GEV
GL
GP
LN3
GAM
K
K_T [-]
T [anni]
1
5
10
50
100
500
1000


VISPA_VISP
G
GEV
GL
GP
LN3
GAM
K
K_T [-]
T [anni]
1
5
10
50
100
500
1000

Allegato II.
Descrizione dei simboli usati per i parametri geomorfologici

Num.	Param. geomorfoclimatico	Descrizione
1	A [km^2]	Area del bacino
2	a [mm/ora^n]	Media delle piogge estreme di 1 ora
3	aff [mm]	Afflusso medio annuo
4	aspect [°]	Aspect medio del bacino (angolo rispetto al Nord)
5	C_comp [-]	Coefficiente di compattezza
6	c_f [-]	Coefficiente di afflusso di piena
7	CN [-]	Curve Number
8	densita_dren [km^{-1}]	Densità di drenaggio
9	ΔH_1 [m]	Differenza tra quota massima e minima del bacino
10	ΔH_2 [m]	Differenza tra quota media e minima del bacino
11	diam_topol [-]	Diametro topologico
12	F_f [-]	Fattore di forma
13	Hmed_A [-]	Rapporto tra quota media e la radice dell'area del bacino
14	H_{media} [m s.l.m.]	Quota media del bacino
15	LC_1 [%]	Aree urbanizzate con tessuto continuo e discontinuo
16	LC_2 [%]	Boschi, vegetazione arborea e arbustiva, cespugliati
17	LC_3 [%]	Vegetazione erbacea, prato-pascolo, vigneti, seminativi
18	LC_4 [%]	Aree non vegetate
19	LC_5 [%]	Aree umide
20	LLDP [km]	Distanza tra il punto a quota massima del bacino e la sezione di chiusura
21	lun_asta_princ [km]	Lunghezza dell'asta principale
22	lungh_media_vers [km]	Lunghezza media dei versanti
23	lungh_vett_orient [km]	Lunghezza del vettore orientamento
24	magnitudine [-]	Magnitudine della rete
25	media_fa [-]	Media della funzione di ampiezza topologica del bacino
26	n [-]	Esponente di invarianza di scala
27	orientamento [°]	deviazione del vettore orientamento rispetto al Nord
28	pend_med_LDP [%]	Pendenza media del LLDP
29	R_a [-]	Rapporto tra le aree
30	R_al [-]	Rapporto di allungamento
31	R_b [-]	Rapporto di biforcazione
32	R_c [-]	Rapporto di circolarità
33	R_l [-]	Rapporto delle lunghezze
34	R_s [-]	Rapporto delle pendenze
35	sl_med1 [%]	Media delle pendenze di tutti i pixel del bacino
36	sl_med2 [%]	Pendenza media invariante del bacino
37	X_{sc} [m]	Coordinata X della sezione di chiusura (UTM)
38	X_{bar} [m]	Coordinata X del baricentro del bacino (UTM)
39	Y_{bar} [m]	Coordinata Y del baricentro del bacino (UTM)
40	Y_{sc} [m]	Coordinata Y della sezione di chiusura (UTM)