

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA

in

Ingegneria Idraulica: Scienze e Tecnologie dell'Acqua

Ciclo XXII

Settore scientifico-disciplinare di afferenza: ICAR/02

**CALIBRAZIONE DI MODELLI IDROLOGICI
CON ALGORITMI MULTIOBIETTIVO**

Laura Lombardi

Coordinatore Dottorato:

Prof. Ing. Alberto Lamberti

Relatore:

Prof. Ing. Alberto Montanari

Correlatore:

Ing. Elena Toth

Esame finale anno 2011

Ringraziamenti

Vorrei rivolgere un sentito ringraziamento a tutti coloro che hanno collaborato alla realizzazione del presente lavoro. Tra questi desidero ringraziare in modo particolare il Prof. Ing. Alberto Montanari e l'Ing. Elena Toth per tutto il tempo che mi hanno dedicato oltre che per le conoscenze che mi hanno trasmesso.

Un grazie speciale all'Ing. Attilio Castellarin e all'Ing. Simone Castiglioni per avermi fornito le informazioni necessarie alle applicazioni realizzate nel presente lavoro oltre che per l'amicizia dimostratami in questi anni.

Infine desidero ringraziare il Prof. Jasper Vrugt per avermi fornito il software per l'utilizzo dell'algoritmo AMALGAM.

Sommaro

INTRODUZIONE	9
1 LA CALIBRAZIONE DEI MODELLI AFFLUSSI-DEFLUSSI	11
1.1 La calibrazione automatica	12
1.1.1 La funzione obiettivo	12
1.1.2 L' Algoritmo di ottimizzazione	13
1.1.2.1 Metodi di ricerca locale	14
1.1.2.2 Metodi di ricerca globale	15
1.1.3 Criterio di arresto	17
1.1.4 Periodo di calibrazione	18
1.2 La calibrazione mono-obiettivo	19
1.3 La calibrazione multi-obiettivo	22
2 GLI ALGORITMI EVOLUTIVI	27
2.1 Gli Algoritmi Euristici (Heuristic Optimisation)	27
2.2 Gli algoritmi Evolutivi (Evolutionary Algorithms)	29
2.3 Gli Algoritmi Genetici (Genetic Algorithms)	30
2.3.1 Modalità di rappresentazione	31
2.3.2 Modalità di selezione	32
2.3.3 Modalità di cambiamento	34
2.3.4 Vantaggi degli algoritmi genetici	35
2.3.5 Limiti degli algoritmi genetici	38
2.4 Cenni su altre tecniche di problem-solving	40
2.4.1 Le reti neurali	40
2.4.2 L'Hill climbing	41
2.4.3 IL Simulated annealing	42
3 L'ALGORITMO DI OTTIMIZZAZIONE AMALGAM	45
3.1 Descrizione della procedura implementata in AMALGAM	46
3.2 Elitist Non-Dominated Sorting GA-II (NSGA-II)	54
3.2.1 Inizializzazione della popolazione	56
3.2.2 Ordinamento delle soluzioni non dominate	56
3.2.3 Calcolo della Crowding Distance	57

3.2.4	Modalità di selezione	58
3.2.5	Operatori genetici	58
3.2.5.1	Simulated Binary Crossover	58
3.2.5.2	Polinomial mutation	59
3.2.6	Ricombinazione e selezione	60
3.3	Particle Swarm Optimization (PSO)	60
3.3.1	Struttura dell'algoritmo	61
3.3.2	Swarm analyses	63
3.3.3	Swarm stagnation	64
3.4	Adaptive Metropolis Search (AMS)	65
3.4.1	Descrizione dell'algoritmo	67
3.5	Differential Evolution (DE)	70
3.5.1	Descrizione dell'algoritmo	71
3.5.2	Mutazione	71
3.5.3	Crossover	72
3.5.4	Selezione	73
4	AREA DI STUDIO E MODELLO AFFLUSSI-DEFLUSSI	75
4.1	Inquadramento geografico	75
4.2	Informazioni Idrometriche	76
4.3	Informazioni Geomorfologiche	78
4.4	Informazioni Climatiche	79
4.5	Il Modello Afflussi-Deflussi Hymod	83
5	LA CALIBRAZIONE DEI MODELLI AFFLUSSI-DEFLUSSI IN BACINI STRUMENTATI	87
5.1	Area di indagine	87
5.2	Funzioni obiettivo	89
5.3	Risultati e confronti	91
5.3.1	Modalità mono-obiettivo	91
5.3.2	Modalità multi-obiettivo (due funzioni)	101
5.3.3	Modalità multi-obiettivo (tre funzioni)	114
5.4	Conclusioni	127
6	LA CALIBRAZIONE DEI MODELLI AFFLUSSI-DEFLUSSI IN BACINI NON STRUMENTATI	131
6.1	L'approssimazione di Whittle della funzione di massima verosimiglianza Gaussiana	134
6.1.1	Regionalizzazione delle grandezze statistiche	137
6.1.2	Indici di Magra	140
6.1.2.1	Procedura di ricampionamento "jack - knife"	141

6.2	Scelta delle Funzioni Obiettivo per la calibrazione in bacini non strumentati	142
6.3	Risultati e confronti	145
6.3.1	Analisi degli indici di qualità nel passaggio da bacino Strumentato a bacino non Strumentato per la calibrazione multi-obiettivo	146
6.3.2	Analisi dei risultati ottenuti in modalità mono-obiettivo	151
6.3.3	Analisi dei risultati ottenuti in modalità multi-obiettivo (tre e quattro funzioni)	165
6.3.4	Confronto modalità mono e multi-obiettivo: incidenza delle statistiche	178
6.3.5	Analisi dei risultati relativi alla riproduzione delle portate	182
6.4	Conclusioni	187
BIBLIOGRAFIA		193
INDICE DELLE FIGURE		199
INDICE DELLE TABELLE		205

Introduzione

I modelli idrologici prima di poter essere utilizzati per la soluzione di problemi pratici devono essere “calibrati”. Con il termine calibrazione si intende la stima dei valori dei “parametri” del modello al fine di ottenere che la simulazione del modello sia il più possibile simile a quello del sistema reale che si sta studiando.

In alcuni casi i valori più appropriati per i parametri del modello possono essere determinati tramite misure realizzate direttamente sul sistema in studio ma, in molte altre situazioni, tali parametri sono rappresentazioni concettuali di caratteristiche astratte del bacino e devono essere determinate attraverso un processo *trial-and-error*, ossia modificando di volta in volta il valore dei parametri fino ad ottenere il risultato desiderato. Questa procedura veniva fatta in passato tramite tarature manuali dei parametri e la bontà del risultato veniva giudicata prevalentemente tramite un esame visivo dell'idrogramma simulato rispetto a quello osservato. Tale prassi, che eseguita da un idrologo esperto può portare anche ad ottimi risultati, non ha però nessun valore oggettivo e può risultare molto dispendiosa in termini di tempo (*Duan et al., 1994 - Yapo et al., 1998 - Gupta, 1998*).

Nella procedura automatica (*computer-based*) i valori dei parametri vengono invece aggiustati automaticamente in accordo con uno schema di ricerca predefinito e in base ad una valutazione oggettiva della capacità del modello di riprodurre il fenomeno in studio. Conseguentemente, rispetto alla prassi manuale, i metodi automatici sono basati su meccanismi oggettivi ed espliciti.

Per questi motivi negli ultimi decenni molti studiosi si sono concentrati sulla possibilità di sviluppare metodi automatici per la stima dei parametri cercando di affrontare i molti aspetti che compongono questo ambito di studio, quali lo sviluppo di tecniche per la gestione degli errori presenti nei valori misurati, la ricerca di procedure di ottimizzazione che possano fornire in modo affidabile la stima dei parametri del modello, la possibilità di ricavare dai dati informazioni quantitativamente e qualitativamente affidabili basandosi anche sulla rappresentazione dell'incertezza che si ripercuote sulla risposta del modello.

Lo sviluppo di procedure di calibrazione automatiche si è concentrato in passato principalmente sull'utilizzo di un'unica funzione obiettivo (ad esempio la radice dell'errore quadratico medio tra osservato e simulato) per il raggiungimento di una simulazione accettabile (*Duan et al., 1992, Duan et al., 1993*). Tuttavia studi più recenti hanno messo in luce l'impossibilità di basarsi su un unico obiettivo per ottenere la simulazione contemporanea dei molteplici aspetti che sono in realtà racchiusi nei dati osservati (*Gupta et al., 1998 - Yapo et al., 1998 - Boyle et al., 2000*).

Questo approccio infatti deriva da un' impostazione legata a tecniche statistiche di tipo classico finalizzate più che altro alla perfetta riproduzione dell'idrogramma misurato piuttosto che alla comprensione dei fenomeni in gioco (*Madsen, 2000*). La ricerca sui metodi di ottimizzazione ha portato all'uso di strategie di tipo evolutivo raggiungendo risultati che si sono dimostrati efficienti ed efficaci nell'obiettivo di identificare i parametri del modello idrologico in uso.

L'obiettivo di questo lavoro è effettuare l'analisi di un algoritmo evolutivo multi obiettivo in applicazione ad un modello afflussi-deflussi di tipo concentrato con particolare attenzione alle strategie di calibrazione utilizzabili e all'applicazione al caso di bacini non strumentati. I risultati presentati sono stati ottenuti utilizzando l'algoritmo di ottimizzazione *multi-obiettivo* AMALGAM proposto da Vrugt e Robinson (2007) in applicazione al modello afflussi-deflussi Hymod proposto da Boyle (2000).

1 La calibrazione dei modelli afflussi-deflussi

La calibrazione di un modello afflussi-deflussi di tipo concettuale comporta che il bacino in esame venga considerato come un sistema soggetto ad un ingresso, costituito dal volume di pioggia e dalla temperatura necessaria per valutare l'evapotraspirazione oppure dall'evapotraspirazione stessa, e da un'uscita, rappresentata dalla portata defluente attraverso la sezione di chiusura.

Il modello deve quindi reagire all'immissione delle variabili di input generando una risposta che si accordi correttamente con il sistema fisico che vuole rappresentare.

Perché questo avvenga un ruolo fondamentale viene svolto dai parametri del modello, i cui valori numerici devono venire calibrati in modo che i dati simulati dallo stesso siano il più vicino possibile ai dati osservati (generalmente il confronto viene fatto tra gli idrogrammi di portata in una determinata sezione del bacino d'interesse).

Tutti i parametri presenti in un dato modello (il cui numero varia a seconda della complessità dello stesso) devono essere stimati perché il modello si adatti bene al caso di interesse: alcuni possono essere determinati tramite lo studio di specifiche caratteristiche del bacino, ma per altri, legati a processi interni del bacino, è necessario procedere ad una stima indiretta o ad una calibrazione, basandosi sulle informazioni disponibili quali serie storiche di precipitazioni e deflussi osservati simultaneamente.

L'importanza dei parametri (e quindi della loro calibrazione) è fondamentale perché da essi dipende l'utilità del modello stesso: se i parametri sono mal specificati anche un modello estremamente sofisticato produrrà portate sempre relativamente lontane da quelle osservate. Ne consegue che la calibrazione deve essere condotta in modo ottimale per massimizzare l'affidabilità del modello.

Fino agli anni '70 la calibrazione manuale costituiva l'approccio più utilizzato. L'idrologo, sfruttando l'esperienza e le informazioni sul bacino, utilizzava un processo iterativo semi-intuitivo per aggiustare i parametri. In pratica dopo ogni aggiustamento procedeva al confronto dei risultati simulati con quelli osservati (tipicamente gli idrogrammi).

Questo metodo però risultava del tutto soggettivo e strettamente legato alla conoscenza personale e all'esperienza dell'idrologo in merito a quel determinato sistema con l'evidente difficoltà di trasferire tale esperienza e conoscenza ad altri idrologi o ad altri bacini.

Accadeva inoltre che, nel caso in cui il numero dei parametri da stimare fosse elevato, la procedura di aggiustamento diventasse troppo laboriosa, dispendiosa in termini di tempo e con effetti non prevedibili sull'output del modello. Lo stesso modello infatti si presenta nella grande maggioranza dei casi con struttura di tipo non-lineare, rendendo difficile prevedere come le varie parti di un idrogramma possano essere sensibili ad un tale processo di aggiustamento dei parametri.

Per questi motivi si è ritenuto opportuno avviare numerose ricerche per realizzare metodi automatici di calibrazione che, grazie all'utilizzo dei calcolatori, hanno consentito di raggiungere alcuni obiettivi fondamentali quali:

- la necessità di velocizzare il processo di calibrazione,
- la necessità di poter prescindere dalla presenza di idrologi esperti su ogni bacino,
- la necessità di eliminare la soggettività legata ad un processo manuale.

1.1 La calibrazione automatica

Una tipica procedura di calibrazione automatica dei parametri richiede la precedente determinazione delle seguenti caratteristiche:

- definizione della funzione obiettivo,
- determinazione dell'algoritmo di ottimizzazione,
- determinazione del criterio di arresto,
- determinazione del periodo di calibrazione.

1.1.1 La funzione obiettivo

Una funzione obiettivo è un'equazione che quantifica numericamente la differenza tra l'output simulato dal modello e quello osservato del bacino; l'obiettivo della calibrazione automatica del modello è quindi quello di trovare quei valori dei parametri che minimizzano/massimizzano a seconda dei casi i valori numerici della funzione obiettivo.

Un esempio di funzione obiettivo molto utilizzata può essere la sommatoria degli scarti quadratici pesati (*Weighted Least Squares*):

$$F(\theta) = \sum_{i=1}^n W_t [q_t^{obs} - q_t(\theta)]^2 \quad (1.1)$$

dove:

$$\begin{aligned} \theta &= \text{vettore dei parametri,} \\ n &= \text{numero degli istanti che devono essere confrontati,} \\ W_t &= \text{peso all'istante } t, \\ q_t^{obs} &= \text{portata osservata all'istante } t, \\ q_t(\theta) &= \text{portata simulata all'istante } t. \end{aligned}$$

In questo caso il peso quantifica il buon adattamento di un particolare valore dell'idrogramma simulato al rispettivo valore osservato. Se i pesi sono tutti posti uguali a 1, la funzione *WLS* si riduce alla sommatoria degli scarti quadratici medi (*Simple Least Squares*).

Il minimo assoluto di questa funzione obiettivo è zero che si raggiungerebbe se il modello riproducesse perfettamente l'idrogramma osservato. Generalmente questo non accade e pertanto l'obiettivo diventa trovare il valore di θ che rende minima la funzione stessa.

Altre funzioni obiettivo di comune utilizzo sono riportate nella Tabella 1.1 dove i simboli assumono il seguente significato:

$$\begin{aligned} O_t &= \text{dato osservato all'istante } t, \\ S_t(\theta) &= \text{dato simulato all'istante } t, \\ n &= \text{numero dei parametri,} \\ \overline{O}_t &= \text{media dei dati osservati.} \end{aligned}$$

1.1.2 L' Algoritmo di ottimizzazione

Ciò che viene descritto dalla funzione obiettivo viene comunemente detto "superficie" della risposta anche se si tratta effettivamente di una superficie solo nel caso a due dimensioni, cioè nel caso di due parametri da calibrare.

L'algoritmo di ottimizzazione è proprio la procedura usata per la ricerca di tale superficie che, a partire dagli intervalli dei valori ammessi per i parametri, determina esattamente quelli che ottimizzano il valore numerico della funzione obiettivo (nella

maggior parte dei casi si tratta di minimizzare tale valore, in quanto la funzione, quasi sempre, è una misura dello scostamento tra valori osservati e simulati).

Esistono diversi metodi di ottimizzazione: in generale si tratta di procedure matematiche che lavorano in modo iterativo al fine di perseguire il miglioramento del risultato.

Abbreviazione	Descrizione	Espressione Matematica
RMSE	Radice dello scarto quadratico medio	$\sqrt{\frac{1}{n} \sum_{t=1}^n (O_t - S_t(\theta))^2}$
ABSERR	Valore assoluto dello scarto medio	$\frac{1}{n} \sum_{t=1}^n O_t - S_t(\theta) $
ABSMAX	Valore assoluto dello scarto medio	$\max_{1 \leq t < n} O_t - S_t(\theta) $
BIAS	Scarto medio	$\frac{1}{n} \sum_{t=1}^n (O_t - S_t(\theta))$
PDIFF	Errore di picco	$\max_{1 \leq t < n} \{O_t\} - \max_{1 \leq t < n} \{S_t(\theta)\}$
NS	Efficienza di Nash	$1 - \frac{\sum_{t=1}^n (O_t - S_t(\theta))^2}{\sum_{t=1}^n (O_t - \bar{O}_t)^2}$
NSC	Numero di cambi di segno	Calcola il numero di cambi di segno della sequenza dei residui

Tabella 1.1. Esempi di funzioni obiettivo di comune utilizzo.

1.1.2.1 Metodi di ricerca locale

Questi metodi sono adatti a cercare in maniera efficiente il minimo di funzioni monotone, cioè funzioni per le quali ogni strategia, che si muova continuamente in "discesa" (nella direzione di miglioramento del valore della funzione), deve arrivare alla posizione del minimo indipendentemente dal punto dello spazio dei parametri da cui è partita la ricerca.

In questo caso la procedura richiede 3 decisioni distinte:

1. in quale direzione muovere,
2. di quanto spostarsi in tale direzione,
3. quando non è più possibile un miglioramento ulteriore.

Le diverse strategie di ricerca locale differiscono in questi punti e per questo motivo si distingue tra metodi di ricerca *diretta* e a *gradiente*.

Nei metodi di ricerca *diretta* si parte da un punto iniziale dello spazio dei parametri, la strategia sceglie alcune direzioni e la dimensione dei passi e, di volta in volta, valuta la funzione nel nuovo punto. Successivamente, basandosi sulle differenze nei valori della funzione tra punto iniziale e punto nuovo, viene fatta una previsione sulla direzione migliore in cui spostarsi e sulla grandezza dello spostamento da compiere per ottimizzare la funzione.

In genere il funzionamento prevede che venga fatto un passo in una direzione di tentativo e, se il nuovo punto presenta un valore più basso della funzione rispetto al precedente, lo sostituisce e la procedura prosegue nella stessa direzione, altrimenti si riduce la lunghezza dello spostamento e si effettua un altro tentativo. La ricerca si ferma quando la strategia non è più in grado di trovare una direzione in cui sia possibile il miglioramento.

Tra questi metodi il più usato è il *Simplex* (Nelder e Mead, 1965) che prevede di individuare un gruppo di $n+1$ punti nello spazio dei parametri possibili (dove n è la dimensione del problema, cioè il numero di parametri da calibrare) e successivamente di operare in modo autonomo più ricerche su quel singolo gruppo senza alcuna condivisione delle informazioni.

Nei metodi a *gradiente* si utilizzano informazioni sui valori della funzione e del rispettivo gradiente per prendere le decisioni citate in precedenza, mentre il funzionamento e il criterio di arresto sono analoghi.

Queste tipologie di metodi non danno però sempre risultati soddisfacenti poiché molti problemi di calibrazione di modelli idrologici, a causa della forte non linearità, presentano punti dello spazio dei parametri in cui il valore della funzione è un minimo locale e non il minimo assoluto. Questo comporta che il punto in cui l'algoritmo termina la ricerca dipende dal punto di partenza e diventa quindi impossibile sapere con certezza se la procedura di ottimizzazione ha trovato il minimo globale della funzione.

Per questo motivo si è sviluppato lo studio di metodi di ricerca *globale*.

1.1.2.2 Metodi di ricerca globale

Le strategie di ricerca globale vengono divise in *deterministiche* o *stocastiche*. Per la calibrazione dei modelli idrologici si predilige l'utilizzo di tecniche stocastiche o al più di una combinazione delle due tipologie di tecniche appena citate.

Una strategia possibile può basarsi su metodi di ricerca *casuale*: essi utilizzano dei generatori di numeri casuali per campionare casualmente lo spazio dei parametri in cerca di punti che migliorino il valore della funzione obiettivo. I campioni sono generati secondo distribuzioni di probabilità applicate allo spazio dei parametri possibili. Nella ricerca *casuale pura* il campionamento è fatto utilizzando una distribuzione uniforme, cioè senza l'utilizzo di nessuna informazione a priori su dove si trovi il set ottimale nello spazio dei parametri. Il fatto che non si utilizzino le informazioni sui valori della funzione ottenute dalle iterazioni precedenti rende questa tecnica di ricerca poco efficiente e per questo motivo, soprattutto nei problemi con molti parametri, vengono utilizzati metodi di ricerca casuale che effettuano il campionamento dello spazio dei parametri basandosi sulle informazioni legate ai valori della funzione ottenuti durante la ricerca medesima. L'obiettivo è quello di aumentare la probabilità di cercare nelle aree dove sono stati trovati valori più bassi della funzione nelle iterazioni precedenti.

Di fatto ciò che noi ricerchiamo nei risultati di una procedura stocastica è che la probabilità di errore sia la minore possibile: per questo vengono utilizzati gli algoritmi *multi-start* che agiscono sviluppando diversi tentativi di ricerca locale da più punti di partenza distinti scelti casualmente nello spazio dei parametri possibili, ottenendo percentuali di successi molto più alte rispetto ai metodi visti finora.

Sebbene le procedure *multi-start* funzionino bene su modelli di bacino semplici, nei casi più complessi ottengono risultati scarsi e per questo è stata sviluppata una sofisticata combinazione del metodo *simplex* non lineare con il metodo *multi-start*, detta algoritmo *SCE-UA* (Duan et al., 1992).

L'algoritmo *SCE-UA* è un metodo di ottimizzazione *globale* che prevede la condivisione di informazioni (al contrario dei metodi puramente *simplex*) e si è dimostrato molto efficace ed efficiente nella calibrazione di modelli di diversa natura.

L'algoritmo parte con la selezione di un insieme di punti, campionati casualmente nello spazio dei parametri, diviso in una serie di *complessi*, ciascuno dei quali contiene $2n+1$ punti (n è sempre la dimensione del problema). Ogni complesso evolve indipendentemente dagli altri, basandosi su un processo di riproduzione statistica che utilizza il metodo *simplex* (insieme di procedure geometriche per dirigere la ricerca in una direzione di miglioramento). Periodicamente però l'intero insieme viene "rimescolato" e vengono formati nuovi complessi in modo che le informazioni ottenute dai complessi precedenti vengano condivise. Se la dimensione iniziale dell'insieme è sufficientemente grande, mentre la ricerca procede, l'insieme tende a convergere nelle

vicinanze dell'ottimo globale. I passi di evoluzione e rimescolamento vengono ripetuti fino a che il criterio di convergenza prestabilito non è soddisfatto.

Il metodo *SCE-UA* contiene molte componenti probabilistiche e deterministiche che vengono controllate da coefficienti dell'algoritmo ed è quindi necessario che tali parametri vengano scelti correttamente al fine di garantire un corretto funzionamento della procedura.

1.1.3 Criterio di arresto

Tutte quante le strategie di ottimizzazione, per la loro natura iterativa, necessitano di un criterio che fermi la ricerca. Ovviamente, la soluzione ottimale sarebbe fermare la ricerca quando è stato individuato il valore della funzione che rappresenta il minimo assoluto. Essendo però difficile sapere quando si è raggiunto questo punto si usano comunemente altri criteri.

Uno di questi è quello di fermare la ricerca quando l'algoritmo non è in grado di migliorare in modo apprezzabile il valore della funzione obiettivo: questo può indicare che è stata raggiunta la posizione di un ottimo oppure una zona molto "piatta" della superficie della risposta. La convergenza della funzione è un criterio utile di stop e un esempio dell'espressione matematica utilizzabile può essere il seguente:

$$\frac{(f_{i-1} - f_i)}{f_i} \leq \varepsilon f \quad (1.2)$$

dove:

$$\begin{aligned} f_{i-1} &= \text{valore della funzione al passo } i-1, \\ f_i &= \text{valore della funzione al passo } i, \\ \varepsilon f &= \text{valore prestabilito.} \end{aligned}$$

Un altro metodo utilizzato consiste nel fermare la ricerca quando l'algoritmo non è in grado di cambiare in modo apprezzabile i valori dei parametri e contemporaneamente migliorare il valore della funzione obiettivo: questo può indicare la localizzazione di un ottimo oppure che è stata raggiunta una zona della superficie della risposta di alta interazione tra i parametri.

Un esempio dell'espressione matematica utilizzabile è il seguente:

$$\frac{\theta(j)_{i-1} - \theta(j)_i}{\theta(j)_{\max} - \theta(j)_{\min}} \leq \varepsilon_{\theta} \text{ per ogni } \theta_j \quad (1.3)$$

dove:

$$\begin{aligned} \theta(j)_{i-1} &= \text{valore del parametro al passo } i-1, \\ \theta(j)_i &= \text{valore del parametro al passo } i, \\ \varepsilon_{\theta} &= \text{valore predefinito.} \end{aligned}$$

Molto spesso invece si termina la ricerca quando un numero massimo predefinito di iterazioni è stato raggiunto (se i due criteri precedenti non sono già stati soddisfatti) anche se può non essere semplice definire tale numero in modo opportuno perché il valore può essere legato sia alla situazione oggetto di studio che all'algoritmo in uso. Per questi motivi dunque si effettuano di solito alcuni tentativi variando il valore del numero di iterazioni e si stabilisce solo successivamente, in base ai risultati ottenuti, quale valore assegnare al numero stesso.

In generale nessuno di questi criteri è totalmente affidabile nell'indicare il raggiungimento dell'ottimo globale tranne nei casi di funzioni convesse e ben poste, ma l'esperienza indica che il miglior procedimento nel caso della calibrazione è il metodo che controlla la convergenza dei parametri.

1.1.4 Periodo di calibrazione

In genere per la calibrazione si usa una parte dei dati disponibili, destinando la parte rimanente per la verifica dei risultati. Tuttavia molti studi hanno dimostrato che, superata una certa lunghezza del set di dati, l'utilizzo di set di lunghezza maggiore migliora in modo solo marginale la stima.

Si può dimostrare infatti che l'errore standard nella stima del parametro j diminuisce con la dimensione n del campionamento, secondo l'espressione:

$$\sigma(j) \propto \frac{1}{\sqrt{n}} \quad (1.4)$$

La qualità dei dati utilizzati per la calibrazione invece è legata alle informazioni contenute nei dati stessi e agli errori dai quali possono essere affetti.

Un set di dati viene considerato *informativo* quando rappresenta sufficientemente bene i diversi processi idrologici e la variabilità nel funzionamento del bacino: questi

aspetti infatti aumentano la possibilità che i dati contengano informazioni sufficienti per determinare valori corretti dei parametri. Se ad esempio i dati sono relativi ad un anno molto secco, alcuni processi del deflusso non vengono attivati e il modello risulta *insensibile* nei confronti di alcuni parametri rendendone difficile la calibrazione. La soluzione migliore sembra essere quella di poter usufruire di dati che contengano la maggior variabilità idrologica possibile per l'area in studio.

Bisogna inoltre considerare che la presenza degli errori di misura riduce la qualità dei dati e conseguentemente l'affidabilità che possiamo fare sulla stima dei parametri. Le sorgenti di errore sono molteplici:

- malfunzionamento degli strumenti,
- problemi legati ai metodi di trasferimento dei dati sui computer
- approssimazione delle curve e delle formule per convertire le osservazioni sul campo in volumi stimati etc.

Spesso viene suggerito che gli effetti combinati di questi errori tendano ad una distribuzione normale, ma questo non è dimostrato e probabilmente questi errori contengono delle distorsioni. Per questo motivo, nella selezione di dati per la calibrazione bisogna esaminare bene la presenza di eventuali errori al fine di non compromettere la stima dei parametri.

Una volta terminata la calibrazione è opportuno fare un test di verifica (validazione): di solito si valutano i risultati del modello calibrato sull'insieme dei dati non utilizzati per la calibrazione sfruttando il fatto che i dati stessi, essendo sempre legati all'area di studio, dovrebbero contenere gli elementi idrologici che il modello è chiamato a stimare o prevedere. L'obiettivo di questa analisi è trovare delle distorsioni nella stima dei parametri in quanto spesso succede che il modello approssimi bene il set di dati in calibrazione ma non quello in validazione (fenomeno della *divergenza del modello*). Se questo avviene la calibrazione deve essere rivista partendo dall'analisi degli elementi che possono avere compromesso il risultato.

1.2 La calibrazione mono-obiettivo

In generale si possono distinguere due tipi di metodi di calibrazione automatica dei parametri. Il primo è costituito dalla calibrazione *mono-obiettivo*: questa tecnica prevede di definire uno spazio possibile dei parametri (generalmente attraverso la definizione dei limiti superiori ed inferiori per ognuno dei parametri) e, successivamente, di cercare in questo spazio un singolo punto (cioè un set di

parametri) che ottimizzi il valore della funzione obiettivo, vale a dire il criterio matematico scelto per misurare le differenze tra simulato e osservato.

L'iter dell'approccio *mono-obiettivo* è il seguente:

- si considera un sistema S per cui un modello idrologico H deve essere calibrato
- si assume che la struttura matematica del modello sia predeterminata e fissata
- si assume che il modello debba simulare un solo aspetto del sistema (tipicamente la portata)
- si assume che i limiti inferiore e superiore (fisicamente realistici) per ogni parametro siano specificati a priori.

Definendo quindi:

$$Q^{obs}(t) = (Q^{obs}(t_1), Q^{obs}(t_2), \dots, Q^{obs}(t_n)) \quad (1.5)$$

il vettore delle portate misurate ai passi $1, \dots, n$ e

$$Q_{\theta}^{sim}(t) = (Q_{\theta}^{sim}(t_1), Q_{\theta}^{sim}(t_2), \dots, Q_{\theta}^{sim}(t_n)) \quad (1.6)$$

il vettore delle corrispondenti portate simulate dal modello usando i parametri θ_i , le differenze tra osservato e simulato sono date dal vettore dei residui:

$$E(\theta) = G[Q^{obs}] - G[Q_{\theta}^{sim}] = [e_1(\theta), e_2(\theta), \dots, e_n(\theta)] \quad (1.7)$$

dove G rappresenta una funzione generica che permette trasformazioni lineari e non lineari.

Nell'approccio classico, l'obiettivo è trovare i valori ottimali dei parametri tali che E sia il più possibile vicino a zero e il procedimento standard è definire una misura L del vettore E (la funzione obiettivo), cercando i valori di θ che minimizzano L .

Si ottiene quindi un solo punto nello spazio dei parametri (Figura 1.1) nel caso a due dimensioni) che rappresenta il set *ottimale* e di conseguenza un solo idrogramma (Figura 1.2).

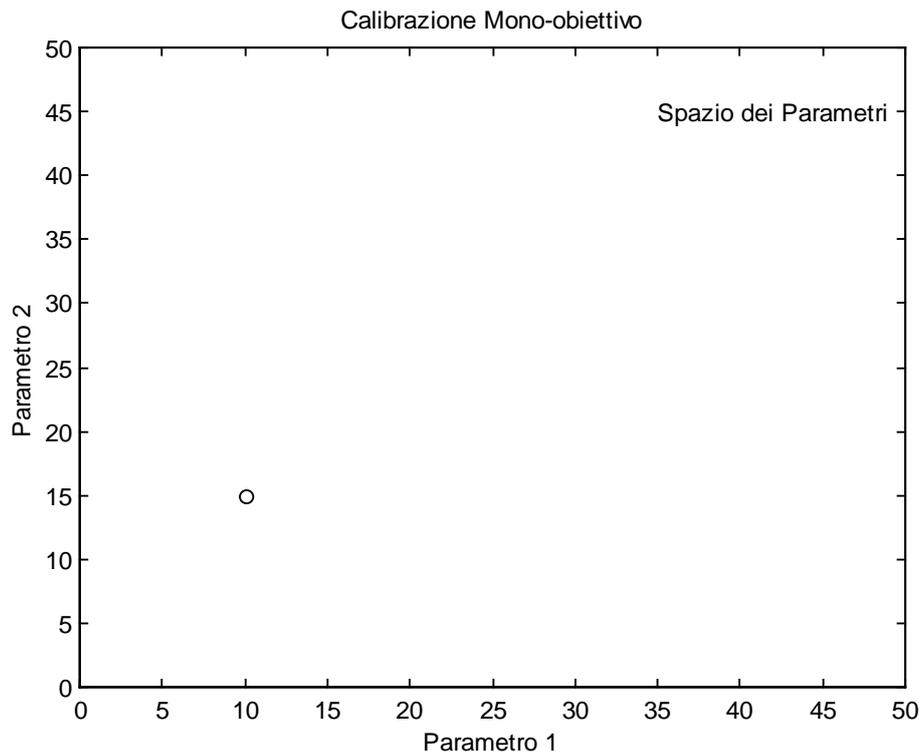


Figura 1.1. Esempio di risultato di una calibrazione mono-obiettivo per un problema a due parametri.

Si è visto però dall'applicazione di questo approccio che il criterio *mono-obiettivo* può fornire un set di parametri che soddisfa il criterio matematico ma che si può dimostrare non realistico dal punto di vista idrologico. Inoltre, pur presentando un punto di forza nella "robustezza" della procedura di ottimizzazione in presenza di forte interazione tra i parametri, questa tecnica ha un punto debole nella sua completa dipendenza da una singola funzione matematica. Ciò comporta che mentre un esperto, durante la calibrazione manuale, può tentare di tenere nella dovuta considerazione tutti gli aspetti del fenomeno in gioco, la scelta di un singolo criterio matematico in un processo di calibrazione automatica, può portare la procedura verso una soluzione inappropriata (ad esempio si è visto che la funzione RMSE approssima bene i picchi e meno bene le recessioni, cioè i periodi successivi ai picchi). Il risultato che emerge è che spesso risulta impossibile trovare un unico set ottimale di parametri.

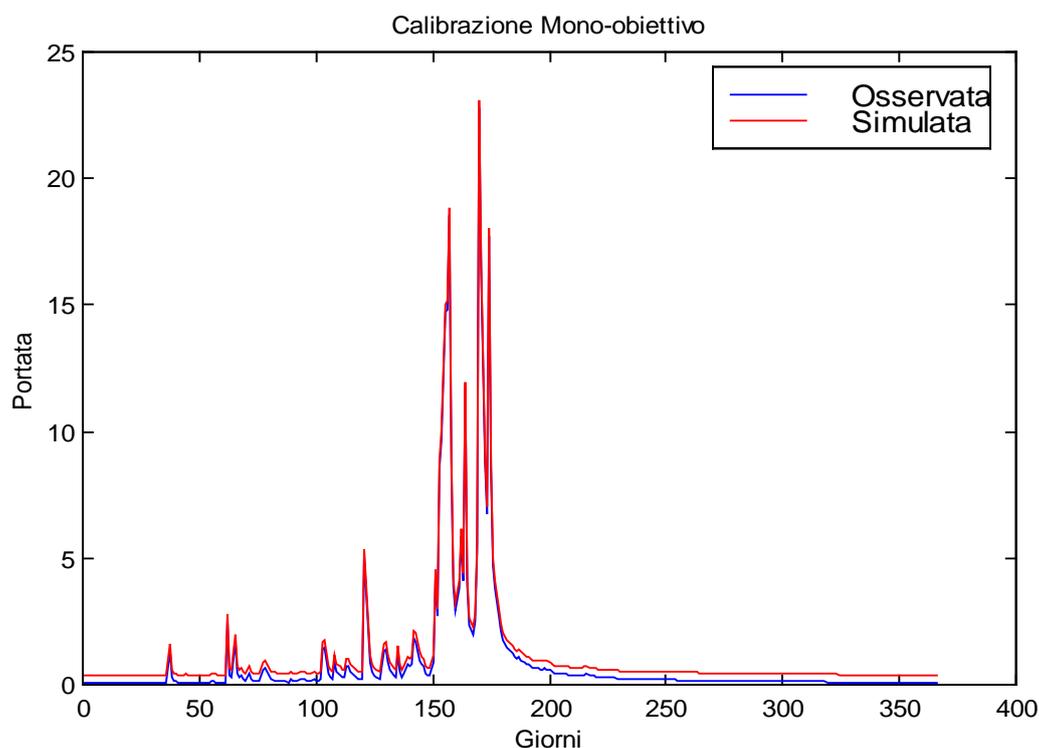


Figura 1.2. Esempio di confronto tra idrogramma osservato e simulato per una calibrazione mono-obiettivo.

1.3 La calibrazione multi-obiettivo

Per risolvere questo problema molti studiosi hanno proposto algoritmi per quella che viene definita *calibrazione multi obiettivo* dimostrando, in questo modo, di poter ottenere un approccio più versatile al problema della definizione del set di parametri del modello (Lindstrom, 1997; Liong et al., 1996 - Liong et al., 1998; Gupta et al., 1998; Yapo et al., 1998).

Tale strategia richiede di analizzare diversi aspetti dell'idrogramma osservato che corrispondono a diverse caratteristiche del funzionamento del bacino e di definire, allo stesso modo, più criteri matematici ognuno dei quali dovrà valutare la somiglianza tra simulato e osservato per un singolo aspetto individuato dell'idrogramma.

Al fine di ottenere una buona riproduzione è necessario infatti considerare più obiettivi contemporaneamente quali ad esempio:

- una buona riproduzione del volume medio dell'idrogramma osservato (corrispondente ad un buon bilancio idrico),
- una buona riproduzione dell'andamento dell'idrogramma osservato,
- una buona riproduzione dei picchi di piena in riferimento al tempo, alla frequenza e al volume idrico coinvolto,

- una buona riproduzione delle magre (*Madsen, 2000*).

Con queste premesse il problema appare implicitamente di natura *multi-obiettivo* e tale approccio si dimostra opportuno per ottenere la simultanea minimizzazione di tutti i criteri matematici utilizzati per quantificare lo scarto tra simulato e osservato.

Nella calibrazione *multi-obiettivo* si fissano dunque più funzioni, l'algoritmo le minimizza/massimizza tutte simultaneamente e si ottengono più set di parametri (Figura 1.3) e più idrogrammi (Figura 1.4), tutti equivalenti seppur diversi tra loro.

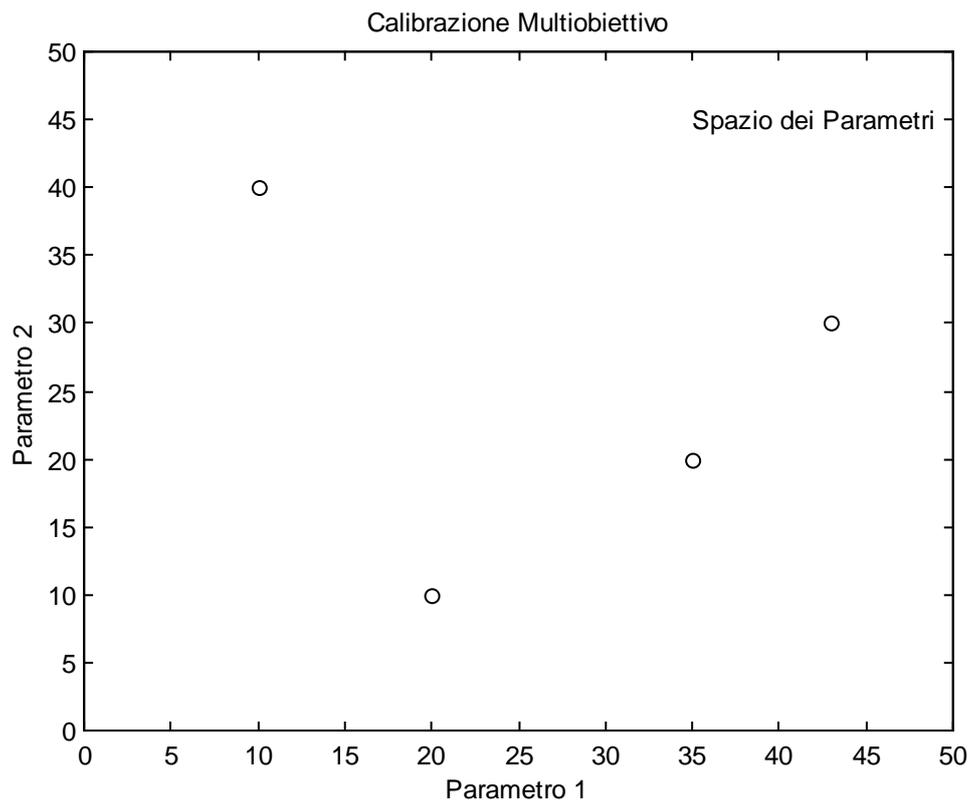


Figura 1.3. Esempio di risultato di una calibrazione multi-obiettivo per un problema a due parametri.

L'equivalenza *multi-obiettivo* fra i set di parametri non è un concetto statistico, ma si basa sulla metodologia con cui si può definire il miglior avvicinamento del modello ai dati osservati e spesso viene chiamata *PARETO OTTIMALE* (cioè di pari efficienza), proprio per indicare che tutti i set sono ugualmente validi. Tale definizione prende origine dall'Italiano Vilfredo Pareto, economista vissuto tra la fine del'800 e i primi anni del '900, che per primo definì le soluzioni di pari efficienza come quelle oltre le quali è impossibile il miglioramento di una condizione senza implicitamente provocare il peggioramento di un'altra.

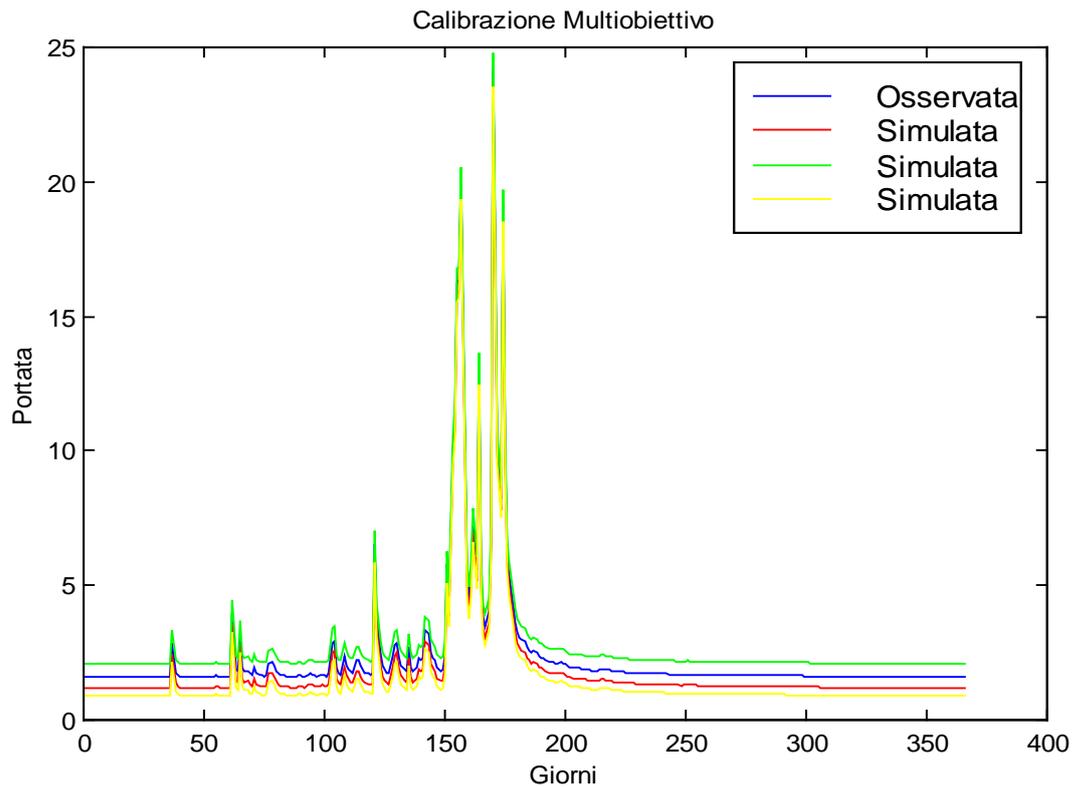


Figura 1.4. Esempio di confronto tra idrogramma osservato e idrogrammi simulati per una calibrazione multi-obiettivo.

Per meglio comprendere il concetto di *PARETO* si può fare il seguente esempio: prendiamo un modello con due parametri (θ_1 e θ_2) e due funzioni obiettivo F_1 e F_2 come indicato in Figura 1.5. I punti *A* e *B* sono quelli che ottimizzano separatamente le due funzioni F_1 e F_2 mentre la linea che li unisce è il *set PARETO* di soluzioni. Tutti i punti che costituiscono questa retta rappresentano diversi compromessi tra le funzioni obiettivo utilizzate (in questo caso due) e, in assenza di informazioni aggiuntive, non è possibile stabilire se una soluzione del set è migliore di un'altra, perché ciascuna di esse può rappresentare meglio una caratteristica dei dati osservati meglio di altre. Tutti i set così identificati formano la *Frontiera Pareto* che ha tipicamente la forma riportata Figura 1.5-b (nel caso di due funzioni obiettivo da minimizzare).

Da un punto di vista matematico la definizione dei *set di PARETO* è tale che ogni membro θ_i del set ha le seguenti proprietà:

- per ciascun escluso θ_j esiste almeno un membro θ_i tale che $F(\theta_i) < F(\theta_j)$, cioè

$$F_k(\theta_i) < F_k(\theta_j) \text{ per } k=1, \dots, m$$

- non è possibile trovare un membro θ_j nei set di PARETO tale che $F(\theta_j) < F(\theta_i)$,
cioè $F_k(\theta_j) < F_k(\theta_i)$ per $k=1, \dots, m$.

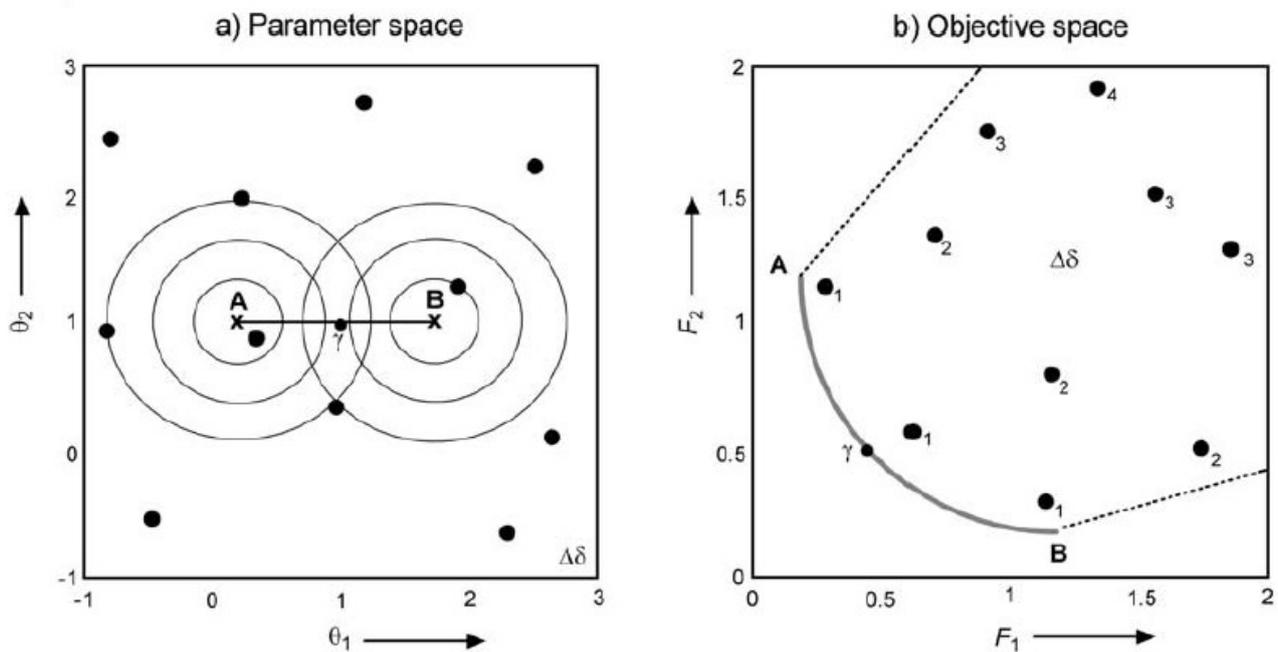


Figura 1.5. Esempio di concetto di Frontiera Pareto nel caso a due parametri e due criteri F_1 e F_2 . a) indica lo spazio dei parametri e b) lo spazio delle funzioni obiettivo. I punti A e B indicano le soluzioni che riducono al minimo ciascuno dei singoli criteri.

I vantaggi della formulazione *multi-obiettivo* sono evidenti sul piano dell'incertezza. Noi abbiamo inizialmente un'incertezza del modello $U^0(\Theta)$, rappresentata dalla dimensione dello spazio possibile dei parametri, a cui corrisponde lo spazio $U^0(Q)$ dei possibili idrogrammi. L'obiettivo della funzione *mono-obiettivo* è di trovare una soluzione unica θ^* che riduca la dimensione finale dell'incertezza sui parametri $U^f(\Theta)$ e, di conseguenza su $U^f(Q)$, a zero. Tuttavia a causa degli errori dei dati e di quelli del modello $U^f(Q)$ rimane diversa da zero. L'approccio *multi-obiettivo* invece raggiunge lo scopo perché il valore minimo di $U^f(\Theta)$ è lo spazio dei parametri PARETO $P(\Theta)$ e il valore minimo di $U^f(Q)$ è lo spazio associato degli idrogrammi $P(Q)$.

Inoltre le dimensioni e le caratteristiche di $P(\Theta)$ e $P(Q)$ ci danno informazioni sui limiti del modello: se ad esempio il range PARETO per alcuni parametri è molto elevato, questo può suggerire la presenza di problemi nella struttura del modello legati a quei parametri.

2 Gli algoritmi evolutivi

L'algoritmo *AMALGAM*, utilizzato per il raggiungimento dei risultati esposti nella presente trattazione, è un *Algoritmo Genetico*. Questi algoritmi costituiscono un sottoinsieme degli *Algoritmi Evolutivi* che a loro volta fanno parte della categoria degli *Algoritmi Euristici*. In questo capitolo quindi si è cercato di descrivere sommariamente le caratteristiche di queste tipologie di algoritmi al fine di comprendere al meglio il funzionamento dell'algoritmo utilizzato per il presente lavoro.

2.1 Gli Algoritmi Euristici (Heuristic Optimisation)

La difficoltà nella risoluzione di problemi che coinvolgono funzioni obiettivo per le quali spesso è difficile conoscere la trattazione di caratteristiche come continuità e derivabilità, ha portato allo sviluppo di un campo di ricerca molto promettente, l'*Euristica*. Si tratta di un campo al limite tra la *Ricerca Operativa* e l'*Intelligenza Artificiale* che utilizza alcune analogie con sistemi sociali e naturali per derivare metodi non deterministici che possano fornire buoni risultati nelle applicazioni ai problemi di ottimizzazione. I sistemi simulati da questi nuovi metodi appartengono a scienze quali medicina, biologia e sociologia (*Colomi et al., 1996*).

Gli *Algoritmi Euristici* sono processi iterativi non-deterministici che simulano un fenomeno esistente in natura. Spesso presentano implicitamente una struttura parallela e hanno la capacità di gestire informazioni retroattive per modificare i parametri e la loro struttura interna. Per questo vengono detti *adattativi*.

Questi algoritmi sono spesso capaci di trovare soluzioni a problemi di ottimizzazione combinatoriale e poiché molti problemi ingegneristici sono di questo tipo, è evidente l'importanza di questo campo di ricerca.

Il vantaggio più importante degli *Algoritmi Euristici* è che sfruttano regole di transizione probabilistiche per guidare la loro ricerca verso quelle regioni dello spazio con probabilità di miglioramento delle soluzioni (*Arnò & Mazzola, 2000*).

D'altro canto invece lo svantaggio fondamentale sta proprio nella loro forma non deterministica ossia non precisa. Utilizzando infatti regole probabilistiche non si hanno criteri di *eccellenza* per valutare la qualità dei *sub-ottimi* ottenuti né per capire il grado di esplorazione della regione ammissibile compiuto dall'algoritmo euristico, se non per il fatto che i valori iniziali della funzione obiettivo siano stati migliorati notevolmente. Un aspetto negativo di molti *Algoritmi Euristici* è infatti la tendenza ad una ricerca di tipo *locale* (paragrafo 1.1.2.1), cioè a rimanere intrappolati in un minimo locale senza riuscire ad esplorare (quindi a vedere in modo più *globale*) tutta la regione ammissibile.

Nella costruzione di un *Algoritmo Euristico* bisogna bilanciare in maniera corretta due caratteristiche fondamentali:

1. il *grado di ricerca*, che misura lo sforzo diretto verso la ricerca locale in una regione dello spazio di studio (se la regione è promettente la ricerca deve essere più approfondita),
2. il *grado di esplorazione*, che misura lo sforzo dell'algoritmo speso nella ricerca in regioni distanti dello spazio (deve essere possibile scegliere una soluzione in una regione distante e/o accettare una cattiva soluzione per avere la possibilità di scoprire nuove buone soluzioni) (Colorni et al., 1996).

Queste due strategie di ricerca sono in conflitto tra loro: l'*Hill-Climbing* (si veda il paragrafo 2.4.2) è un esempio di strategia con un alto grado di ricerca della migliore soluzione ma per contro trascura l'esplorazione dello spazio di ricerca.

La ricerca *casuale* (paragrafo 1.1.2.2) è invece un tipico esempio di strategia che esplora in maniera spinta lo spazio, ignorando però qualsiasi tipo di ricerca di regioni promettenti.

Un *Algoritmo Euristico* è un metodo che cerca un equilibrio tra questi due estremi e per questo in esso è fondamentale un attento bilanciamento tra ricerca ed esplorazione (Michalewicz, 1996).

Nel prossimo paragrafo verranno illustrate le caratteristiche di una tipologia di *Algoritmi Euristici* tra le più utilizzate per l'ottimizzazione dei modelli afflussi-deflussi: gli *Algoritmi Evolutivi*.

2.2 Gli algoritmi Evolutivi (Evolutionary Algorithms)

La ricerca nei metodi di ottimizzazione e la possibilità di utilizzare una programmazione automatica (*Friedberg, 1958*) ha portato allo sviluppo di strategie basate sui principi dell'evoluzione delle popolazioni (e.g. *Brazil and Krajewski, 1987; Brazil, 1988; Wang, 1991; Duan et al., 1992-1993; Sorooshian et al., 1993*) e conseguentemente all'utilizzo di algoritmi definiti appunto evolutivi (*Evolutionary Algorithms - EAs* in seguito).

Gli *EAs* permettono di ottenere una soluzione stocastica, basata sulla filogenesi (cioè sull'evoluzione) della popolazione raggiunta sfruttando le conoscenze delle generazioni precedenti e la competizione degli individui per la sopravvivenza nell'ambiente.

Nell'intelligenza artificiale, gli *EAs*, rappresentano un sottoinsieme del calcolo evolutivo che si basa sul concetto di ottimizzazione della popolazione iniziale. Un *EA* utilizza alcuni meccanismi ispirati all'evoluzione biologica, alla riproduzione, alla mutazione, alla ricombinazione e alla selezione delle caratteristiche delle specie viventi.

Le soluzioni possibili al problema di ottimizzazione sono paragonabili quindi agli individui di una popolazione e la funzione obiettivo utilizzata determina l'ambiente all'interno del quale le soluzioni/individui vivono. L'evoluzione quindi avviene proprio per la successiva applicazione delle condizioni imposte dall'ambiente (funzione obiettivo) e solo le soluzioni/individui che sopravvivono risultano essere adatti a soddisfare le condizioni imposte.

Gli *EAs* hanno portato ad importanti risultati nella risoluzione di problemi di ottimizzazione che coinvolgono più funzioni obiettivo in conflitto l'una con l'altra. Grazie alla loro abilità di esplorare un'ampia gamma di soluzioni, questi algoritmi permettono di ottenere diversi set che vengono iterativamente migliorati grazie alla ricombinazione delle soluzioni stesse. Questa caratteristica permette di raggiungere la convergenza delle soluzioni a quelle di *Frontiera* in un singolo processo di ricerca (*Zitzler et al., 1999*).

Esistono vari esempi di algoritmi evolutivi tra i quali:

- *Genetic programming (GP)* - *Koza, 1992*,
- *Evolutionary programming (EP)* - *Fogel et al., 1966*,
- *Evolution Strategy (ESs)* - *Rechenberg, 1965*,
- *Differential Evolution (DE)* - *Storn & Price, 1995*,
- *Genetic Algorithms (GAs)* - *Bremermann, 1958, Holland, 1962-1975, Goldberg, 1989*.

Questi ultimi sono i più utilizzati e sono caratterizzati dalla ricerca di soluzioni del problema in forma di stringhe di numeri (tradizionalmente binari, anche se le migliori sono solitamente quelle che riflettono alcune caratteristiche del problema che deve essere risolto), ottenute tramite l'applicazione di operatori come la ricombinazione o la mutazione (singolarmente o insieme). Gli *Algoritmi Genetici* sono particolarmente adatti ai problemi di ottimizzazione e per questo trovano ampia applicazione nella calibrazione dei modelli afflussi-deflussi.

2.3 Gli Algoritmi Genetici (Genetic Algorithms)

Come detto in precedenza gli *Algoritmi Genetici* (*GAs* in seguito) sono algoritmi di analisi dei dati di tipo *euristico* e sono ispirati al principio della selezione naturale di Charles Darwin che regola l'evoluzione biologica. Questo tipo di algoritmo è detto *genetico* per similitudine con la genetica, branca della biologia.

Gli *GAs* sono applicabili alla risoluzione di un'ampia varietà di problemi d'ottimizzazione non indicati per gli algoritmi classici, compresi quelli in cui la funzione obiettivo è discontinua, non derivabile, stocastica o fortemente non lineare.

In pratica un *GA* è una tecnica di programmazione che simula l'evoluzione biologica come una strategia di *problem-solving*. Dato un problema specifico da risolvere, l'input per l'*GA* sono una serie di possibili soluzioni a questo problema, codificato in qualche modo, e una o più funzioni obiettivo, altrimenti dette funzioni di *fitness*, che vengono valutate quantitativamente per ogni soluzione possibile. Queste soluzioni possono essere già note, e in questo caso l'obiettivo del *GA* è quello di migliorarne la qualità, ma più spesso vengono generate a caso.

Generati quindi una serie di candidati in modo casuale, l'*GA* valuta per ciascun candidato il valore della/delle funzioni di *fitness*. Naturalmente la maggior parte delle soluzioni risulterà inadatto e per questo verrà scartato ma alcune di esse, seppur in modo marginale, presenterà un certo grado di soddisfazione dei requisiti. Tali candidati risultano quindi promettenti e verranno mantenuti per generare nuove soluzioni. Le nuove generazioni non saranno ancora perfette ma grazie alle variazioni casuali introdotte durante il processo di riproduzione e ripetendo più volte la sequenza di eliminazione e ricombinazione degli individui è possibile avvicinarsi sempre più ad una generazione di ottimi.

Il processo viene ripetuto centinaia o migliaia di volte proprio per consentire il graduale miglioramento del valore di *fitness* medio della popolazione generata ad ogni step.

L'algoritmo evolve quindi attraverso i seguenti punti:

- generazione, in maniera casuale, di una popolazione iniziale;
- creazione di una sequenza di nuove popolazioni, o generazioni. In ciascuna iterazione, gli individui della popolazione corrente sono usati per creare la generazione successiva, e a questo scopo si compiono degli ulteriori passi:
 - ciascun membro della popolazione corrente è valutato calcolandone il rispettivo valore di *fitness* (o *idoneità*);
 - si determina un opportuno ordinamento di tali individui sulla base dei valori di *fitness*;
 - gli individui più promettenti sono selezionati come genitori;
 - a partire da tali individui si genera un pari numero di individui della generazione successiva, e ciò può avvenire secondo due modalità distinte, vale a dire effettuando cambiamenti casuali su un singolo genitore (*mutazione*) oppure combinando opportunamente le caratteristiche di una coppia di genitori (*incrocio*);
 - gli individui così generati vanno a sostituire i genitori consentendo la formazione della generazione successiva;
- infine, l'algoritmo s'interrompe quando uno dei criteri d'arresto (si veda il paragrafo 1.1.3) è soddisfatto.

Per quanto sorprendenti e contro-intuitivi come possono sembrare ad alcuni, gli GAs hanno dimostrato di essere una strategia estremamente potente e di essere in grado di risolvere problemi di svariata natura.

2.3.1 Modalità di rappresentazione

Prima che un GA possa essere utilizzato è necessario codificare un metodo per rappresentare le potenziali soluzioni al problema in studio. Un approccio comune è quello di codificare le soluzioni sotto forma di stringhe binarie ossia sequenze di 1 e 0, dove la cifra in ciascuna posizione rappresenta il valore di qualche aspetto della soluzione. Un altro approccio simile è quello di codificare soluzioni come *array* di interi o numeri decimali dove ogni posizione ancora una volta rappresenta un particolare

aspetto della soluzione. Questo approccio consente una maggiore precisione e complessità del metodo relativamente ristretto dei numeri binari poiché intuitivamente più vicino al problema di ispezione dello spazio possibile. Un terzo approccio è quello di rappresentare gli individui in generale, come stringhe di lettere, dove ogni lettera rappresenta ancora una volta un aspetto specifico della soluzione.

Tutti e tre questi metodi di rappresentazione hanno la virtù di rendere facile il modo in cui si definiscono gli operatori che causano i cambiamenti casuali nei candidati prescelti:

- *shift* da 0 a 1 o viceversa,
- somma o sottrazione di un determinato importo scelto a caso dal valore del numero ottenuto,
- sostituzione di una lettera.

Nel *genetic programming* ad esempio i programmi sono rappresentati come strutture di dati ramificate chiamate per questo *alberi*. In questo approccio, i cambiamenti casuali possono essere realizzati modificando l'operatore, modificando il valore in un dato nodo nell'albero oppure sostituendo una ramificazione dell'albero con un'altra (Figura 2.1).

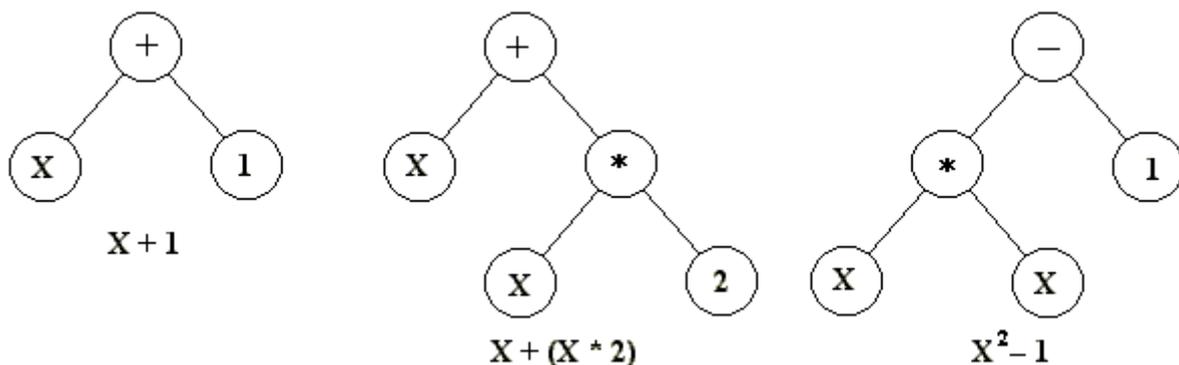


Figura 2.1. Esempi di tre strutture ad albero normalmente utilizzate nella programmazione genetica. L'espressione matematica che ognuno rappresenta è riportata sotto alla corrispondente rappresentazione.

2.3.2 Modalità di selezione

Ci sono molte tecniche diverse attraverso le quali un *GA* può *selezionare* gli individui da utilizzare per la creazione della generazione successiva. Di seguito vengono elencati e brevemente spiegati i più comuni. Alcuni di essi sono mutualmente esclusivi mentre altri possono e sono spesso usati in combinazione.

- *Elitist selection*: i membri più adatti di ogni generazione sono sicuramente selezionati (in realtà la maggior parte degli GAs non applicano l'elitarismo puro ma ne utilizzano una forma modificata: in pratica il miglior individuo o un piccolo gruppo di migliori di una generazione vengono portati nella generazione successiva solo nel caso in cui non ci siano individui che forniscano migliori risultati).
- *Fitness-proportionate selection*: i membri più adatti hanno più probabilità ma non sono certi di essere selezionati per il passaggio alla generazione successiva.
- *Roulette-wheel selection*: è una forma di *selezione* proporzionata al grado di attitudine. In pratica la probabilità di un individuo di essere selezionato è proporzionale al fatto che si dimostri più in forma dei suoi *competitors*. In modo figurato può essere immaginato come una *roulette* dove ogni individuo ha una fetta più o meno grande in funzione della sua attitudine a rispecchiare le caratteristiche richieste e per questo avrà maggiori probabilità di essere selezionato.
- *Scaling selection*: la funzione di *fitness* diventa più discriminante man mano che aumenta il *fitness* medio della popolazione e la pressione della selezione si appesantisce. Questo metodo può essere molto utile quando la procedura è arrivata già ad un certo livello di selezione e la popolazione ha mediamente un buon livello di *fitness* con solo piccole differenze tra i membri.
- *Tournament selection*: dalla popolazione iniziale si individuano dei sottogruppi e gli individui di ogni sottogruppo competono solo con gli appartenenti allo stesso gruppo. Solo i vincitori di ogni sottogruppo saranno selezionati per la riproduzione.
- *Rank selection*: ad ogni elemento della popolazione è assegnato un grado numerico (*numerical rank*) in base al suo valore di *fitness* e la selezione si basa su questo grado piuttosto che sul valore di *fitness* in senso assoluto. Il vantaggio di questo approccio è che previene il fatto che individui con valore di *fitness* molto elevato predominino in modo schiacciante quelli con valore più basso mettendo a rischio il mantenimento della varietà dei membri della popolazione e rischiando di portare ad una soluzione solo apparentemente migliore.
- *Generational selection*: gli individui generati costituiscono le nuove generazioni mentre nessun individuo della generazione precedente viene mantenuto.
- *Steady-state selection*: gli individui di nuova generazione ritornano indietro a far parte della generazione precedente e vanno a sostituire quelli che si erano dimostrati meno adatti. Alcuni individui vengono mantenuti attraverso le generazioni.

- *Hierarchical selection*: gli individui passano attraverso vari cicli di selezione per ogni generazione. Le prime valutazioni sono più veloci e meno discriminanti, mentre quelli che sopravvivono a più alti livelli vengono valutati con maggior rigore. Il vantaggio di questo metodo è che riduce il tempo complessivo di calcolo utilizzando appunto criteri di valutazione più veloci e meno impegnativi per estirpare la maggior parte degli individui che mostrano poca o nessuna attitudine e sottoponendo solo quelli che sopravvivono a questa prova iniziale a criteri di selezione più rigorosi e computazionalmente più impegnativi.

2.3.3 Modalità di cambiamento

Una volta che tramite la modalità di *selezione* si sono individuati gli individui adatti, questi devono essere alterati in modo casuale nella speranza di migliorare la loro idoneità per la prossima generazione. Ci sono due strategie di base per raggiungere questo scopo. La prima e più semplice si chiama *mutazione*. Proprio come la *mutazione* negli esseri viventi cambia un gene con un altro, così la *mutazione* in un *GA* provoca alterazioni di singoli punti nel codice di un individuo. La *mutazione* quindi consiste nella modifica casuale di alcune parti dei geni con valore di *fitness* più basso, in base a coefficienti definiti inizialmente. Queste modifiche puntano a migliorare il valore della funzione per il gene in questione.

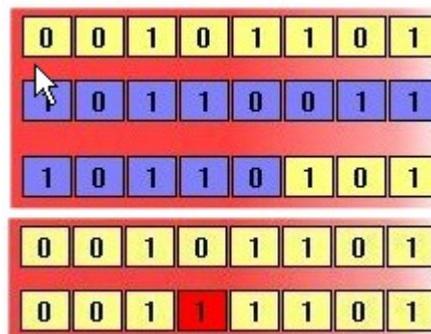


Figura 2.2. Crossover e mutazione.

In realtà non è corretto pensare di *mutare* solo i cromosomi con *fitness* più basso poiché al fine di garantire una maggiore capacità esplorativa dell'algoritmo (e non finire in punti di ottimo locale) sono ritenute utili anche le *mutazioni* di cromosomi con valore di *fitness* alto. In pratica quindi è corretto sostenere che le *mutazioni* servono soprattutto a esplorare lo spazio di ricerca e non puramente a migliorare il risultato.

Il secondo metodo è chiamato *crossover* e comporta lo scambio di due segmenti del codice di due individui, portando così ad un nuovo individuo che sarà una

combinazione delle caratteristiche dei due *genitori*. Questo processo ha lo scopo di simulare la ricombinazione che si verifica ai cromosomi durante la riproduzione sessuale.

Forme comuni di *crossover* includono il *crossover a punto singolo*, in cui il punto di scambio è fissato in una posizione casuale del genoma dei due individui originari e il primo individuo contribuisce con tutto il suo codice prima di tale punto e l'altro contribuisce con tutto il suo codice dal punto di scambio in poi, e il *crossover uniforme*, in cui il valore in qualsiasi sito nel genoma della prole è il valore del genoma di un genitore in quella posizione o il valore del genoma dell'altro genitore in quel punto, scelto con probabilità del 50%. La Figura 2.2 illustra questi operatori genetici sugli individui in una popolazione rappresentata con stringhe di 8 bit. Il codice superiore mostra rappresenta due individui sottoposti a *single-point crossover*: il punto di scambio è fissato tra la quinta e la sesta posizione del genoma producendo un nuovo individuo che risulta un ibrido tra i genitori. Il codice inferiore rappresenta un individuo sottoposto a *mutazione* nella quarta posizione sostituendo lo 0 del genoma originario con un 1.

2.3.4 Vantaggi degli algoritmi genetici

Il primo e più importante vantaggio degli *GAs* è che sono intrinsecamente paralleli. La maggior parte degli altri algoritmi funziona in modo seriale e può esplorare lo spazio delle soluzioni di un problema in una sola direzione alla volta. Questo comporta che se la soluzione scoperta risulta essere *sub-ottimale*, tutto il lavoro svolto fino a quel momento deve essere abbandonato ed è necessario avviare una nuova procedura. Gli *GAs* invece, generando più individui contemporaneamente, permettono di esplorare lo spazio delle soluzioni in molteplici direzioni nello stesso momento. Se un percorso risulta essere un vicolo cieco, si può facilmente eliminare e proseguire il lavoro sulle strade più promettenti, dando loro una maggiore possibilità di trovare ogni volta la soluzione ottimale.

Il vantaggio del parallelismo va oltre quanto appena detto. Si consideri il seguente esempio: tutte le stringhe di 8 cifre binarie (stringhe di 0 e di 1) formano uno spazio di ricerca, che può essere rappresentato come $*****$ (dove * sta per " 0 o 1 "). La stringa 01101010 è un membro di questo spazio. Tuttavia, è anche un membro dello spazio $0*****$, dello spazio $01*****$, dello spazio $*****00$, dello spazio $0*1*1*1*$, dello spazio di $01*01**0$ e così via.

Valutando l'idoneità di questa stringa in particolare, un *GA* campionerebbe i molti spazi a cui la stringa appartiene. Per tutte queste valutazioni calcolerebbe un valore sempre più accurato per l'idoneità media in ciascuno di questi spazi che a sua volta comprende molteplici individui. Pertanto, un *GA* che valuta in modo esplicito un piccolo numero di individui è implicitamente concentrato su un gruppo molto più ampio di persone così come un sondaggista che pone delle domande ad un membro di una certa etnia o gruppo religioso o sociale spera di imparare qualcosa circa le opinioni di tutti i membri di tale gruppo, e quindi tenta di prevedere in modo attendibile il parere di molti campionando solo una piccola percentuale della popolazione. Allo stesso modo quindi un *GA* può dall'interno dello spazio degli individui studiare quelli con valore di *fitness* più elevato e che rappresentano al meglio le caratteristiche del gruppo. Nel contesto degli *Algoritmi Evolutivi* questo è noto come *Schema Theorem* e rappresenta il vantaggio principale degli *GA* rispetto ad altri metodi di soluzione dei problemi (*Holland, 1992; Mitchell 1996; Goldberg 1989*).

Grazie al parallelismo, che permette di valutare implicitamente ogni schema in una volta sola, gli *GAs* sono particolarmente adatti a risolvere i problemi in cui lo spazio di tutte le soluzioni possibili è particolarmente vasto, tanto cioè da rendere impossibile la ricerca esaustiva di un'entità alla volta in un tempo ragionevole. La maggior parte dei problemi che rientrano in questa categoria sono di tipo *non lineare*. In un problema *lineare*, l'idoneità di ciascun componente è indipendente, quindi qualsiasi miglioramento a qualunque parte si tradurrà in un miglioramento del sistema nel suo complesso. Inutile dire che pochi problemi del mondo reale sono di questo tipo. Di fatto la *non linearità* comporta che un cambiamento su un componente abbia effetti a catena sull'intero sistema e, ancor più difficile da prevedere, può accadere che più modifiche, che individualmente sarebbero dannose, possono portare a miglioramenti decisamente maggiori dei singoli interventi.

La *non-linearità* comporta un incremento esponenziale di casistiche: se il problema è *lineare* uno spazio di stringhe binarie di 1.000 cifre può essere studiato esaustivamente valutando solo 2.000 possibilità, mentre se il problema è *non lineare*, per avere un quadro esaustivo è necessario valutare 21000 possibilità.

Fortunatamente, il *parallelismo* implicito di un *GA* consente di superare anche questo numero enorme di casistiche, riuscendo a trovare ottimi risultati in un breve periodo di tempo dopo aver campionato direttamente solo piccole regioni dello spazio possibile (*Forrest, 1993*).

Altro punto di forza notevole degli *GAs* è la loro particolare attitudine a risolvere problemi che presentano impostazioni complesse come ad esempio funzioni di tipo

discontinuo, variabili nel tempo o con molti ottimi locali. Quest'ultimo punto è di particolare delicatezza poiché la maggior parte dei problemi pratici ha un vasto spazio di soluzioni che difficilmente può essere ispezionato in modo esaustivo e conseguentemente la difficoltà sta proprio nel riuscire ad evitare un ottimo locale. Il problema diventa quindi legato a come individuare quelle soluzioni che appaiono meglio di tutte quelle limitrofe ma che in realtà non sono accettabili rispetto ad altre meno vicine.

In effetti non esiste alcun modo per capire se una soluzione è effettivamente l'ottimo ma grazie alle quattro componenti che caratterizzano un *GA* - *parallelismo*, *selezione*, *mutazione* e *crossover* – si è dimostrata la loro attitudine a riconoscere gli ottimi locali e a fornire una soluzione quanto meno molto buona.

Il *crossover* è l'elemento chiave che distingue gli *GAs* da altri metodi come *hill-climber* e il *simulated annealing*. Senza *crossover*, ogni soluzione agisce in modo autonomo ed esplora lo spazio di ricerca nelle sue immediate vicinanze, senza riferimento a ciò che altri soggetti possono avere scoperto. Invece con il *crossover* vi è un trasferimento di informazioni tra i candidati selezionati, gli individui possono trarre vantaggio da ciò che gli altri hanno imparato e gli schemi possono essere mescolati e combinati, con il potenziale di produrre una progenie che ha i punti di forza di entrambi i suoi genitori e le debolezze di nessuno dei due.

Un altro pregio fondamentale degli *GAs* è la loro capacità di manipolare molti parametri contemporaneamente (*Forrest, 1993*). Molti problemi del mondo reale non si possono concretizzare in termini di un singolo valore da minimizzare o massimizzare ma devono essere affrontati in termini di obiettivi multipli con i conseguenti compromessi. Gli *GAs*, sempre grazie al *parallelismo*, permettono di produrre più soluzioni di pari efficienza anche se ovviamente ci saranno soluzioni che tenderanno a massimizzare un aspetto rispetto ad altre che tenderanno a massimizzarne un altro. Come accennato in precedenza (paragrafo 1.3) se una particolare soluzione ad un problema *multi-obiettivo* ottimizza un parametro ad un livello tale che tale parametro non può essere ulteriormente migliorato senza causare una corrispondente diminuzione della qualità di qualche altro parametro, tale soluzione prende il nome di *Pareto ottimale* o *non dominato* (*Coello, 2000*).

La virtù di queste tecniche infine sta anche nel fatto che, basandosi su decisioni casuali, non sono vincolate da preconcetti o da conoscenze che a priori potrebbero imporre delle esclusioni o accettazioni. Infatti tutti i percorsi di ricerca possibili sono teoricamente aperti mentre qualsiasi strategia di soluzione dei problemi che si basa su una conoscenza preliminare deve necessariamente cominciare con l'escludere alcuni

percorsi. Infine qualsiasi tecnica che si basa su conoscenze pregresse è vincolata alla presenza delle stesse e quindi si blocca quando tali conoscenze non sono disponibili.

2.3.5 Limiti degli algoritmi genetici

Sebbene gli *GAs* abbiano dimostrato di essere una strategia efficace e potente per la soluzione di molti tipi di problemi, non sono privi di alcune limitazioni che possono comunque essere superate grazie all'indiscutibile validità della teoria sulla quale si basano: l'evoluzione biologica.

La prima considerazione da fare è che nell'utilizzo di un *GA* non è sempre semplice definire una rappresentazione del problema che sia robusta ossia in grado di tollerare variazioni casuali che non alterino la qualità della soluzione finale. Per raggiungere tale obiettivo ci sono due modi. Il primo e più utilizzato è quello di definire individui come elenchi di numeri (a valori binari, valori interi o valori reali) dove ogni numero rappresenti un aspetto di una soluzione candidata. Se gli individui sono stringhe binarie, 0 o 1 rappresentano l'assenza o la presenza di una determinata caratteristica. Se essi sono liste di numeri, questi numeri potrebbero rappresentare molte cose diverse: il peso dei *link* in una rete neurale, l'ordine delle città visitate in un tour, la collocazione spaziale di componenti elettronici, i valori inseriti in un controllore, gli angoli di torsione di legami in una proteina, e così via.

La *mutazione* quindi comporta la modifica di questi numeri, spostando *bit*, aggiungendo o sottraendo valori casuali. In questo metodo, il codice effettivo del programma non cambia: il codice è quello che gestisce la simulazione e registra i singoli individui, la valutazione della loro idoneità e tenta di assicurare che solo i valori realistici e possibili per il risultato del problema vengano tenuti.

Il secondo metodo prevede invece che il codice vero e proprio del programma cambi. Come descritto in precedenza la programmazione genetica prevede che gli individui possano essere rappresentati con codice strutturato ad albero e quindi è possibile prevedere la sostituzione di una ramificazione, cioè di una parte di codice, con un'altra.

Entrambi questi metodi permettono di realizzare rappresentazioni robuste rispetto alla *mutazione* e possono essere applicati a diverse tipologie di casi di studio. Ovviamente il problema appena esplicitato non si presenta in natura poiché il codice genetico è intrinsecamente robusto e conseguentemente qualsiasi *mutazione* ha una buona probabilità di produrre un miglioramento.

Il problema di come scrivere la funzione di *fitness* dunque deve essere valutato attentamente poiché il raggiungimento di una maggior adeguatezza corrisponde alla possibilità di avere una soluzione migliore al problema dato. Viceversa se la funzione di *fitness* è mal posta, l'algoritmo può non essere in grado di trovare una soluzione o può raggiungere una soluzione ad un problema in realtà diverso. In questo tipo di problema la natura non ci soccorre in quanto nell'evoluzione biologica c'è solo una funzione di *fitness* che è la stessa per tutti gli esseri viventi e cioè la spinta a sopravvivere e a riprodursi. Ne risulta che quegli organismi che si riproducono in abbondanza rispetto ai loro concorrenti sono più adatti alla sopravvivenza in quell'ambiente mentre quelli che non riescono a riprodursi sono destinati all'estinzione.

Oltre alla scelta della funzione di *fitness*, è necessario porre particolare attenzione anche agli altri parametri coinvolti in un *GA* come la dimensione della popolazione, la consistenza della *mutazione* e del *crossover*, la tipologia e la metodologia di *selezione*. Se ad esempio la popolazione iniziale è troppo piccola, l'*Algoritmo Genetico* non può esplorare lo spazio sufficiente per trovare una soluzione coerente e soddisfacente. Se invece la consistenza della variazione genetica è troppo elevata o il sistema di selezione è stato scelto male, lo schema elaborato per il miglioramento potrebbe essere disturbato e portare al contrario ad una catastrofe legata all'eccessiva velocità con la quale viene messa in atto la selezione. Anche in questo caso possiamo fare una analogia con gli esseri viventi poiché si è visto che se la dimensione di una popolazione diventa troppo bassa, le percentuali di *mutazione* diventano molto alte o la pressione della selezione troppo forte (una tale situazione potrebbe essere causata da drastici cambiamenti ambientali), con possibile estinzione della specie.

Un tipo di problema di difficile risoluzione negli *GAs* è quello detto con funzioni di *fitness ingannevoli* (Mitchell, 1996) ossia dove i punti che migliorano la soluzione danno informazioni fuorvianti sulla posizione dell'*optimum*. Come esempio immaginiamo un problema in cui lo spazio di ricerca sia costituito da tutte le stringhe binarie di otto caratteri e l'idoneità di un individuo sia direttamente proporzionale al numero di 1 in essa - vale a dire 00000001 sarebbe meno adatta a 00000011, che a sua volta sarebbe meno adatta di 00000111 e così via. Tuttavia può accadere che la stringa 11111111 risulti avere una funzione di *fitness* molto bassa e la stringa 00000000 una molto alta. In questo tipo di situazione un *GA* non ha più probabilità di trovare la soluzione migliore di una semplice ricerca casuale. Nel mondo reale pochi sono i problemi completamente ingannevoli come l'esempio artificioso fatto poco

sopra e solitamente la posizione dell'ottimo locale offre almeno alcune informazioni sulla posizione dell'ottimo globale.

Un problema ben noto che può verificarsi con un GA è dato dalla possibile convergenza prematura anche se tale fenomeno si presenti anche in natura e sia pertanto ritenuto del tutto normale. Può accadere infatti che se un individuo risulta più adatto di altri in quell'ambiente, esso emerge molto velocemente e può riprodursi così in fretta da limitare la diversità della popolazione troppo presto. In questo caso l'algoritmo converge all'ottimo locale rappresentato da quell'individuo piuttosto che esplorare completamente tutto il panorama possibile (Forrest, 1993; Mitchell, 1996). Questo è un problema comune soprattutto se la popolazione iniziale è piccola poiché le variazioni anche limitate legate alla riproduzione possono fare in modo che un individuo diventi dominante sugli altri con una certa facilità. Il metodo più comune per correggere questo problema è esercitare un controllo sulla *selezione* che può essere messo in pratica utilizzando le tecniche di *rank*, *scaling* e *tournament selection* descritte in precedenza.

Infine alcuni ricercatori (Holland, 1992; Forrest, 1993) consigliano di non utilizzare gli GAs con problemi risolvibili analiticamente poiché i tradizionali metodi analitici necessitano di molto meno tempo e sforzo computazionale rispetto ad essi e, a differenza di essi, sono di solito matematicamente garantiti per fornire proprio la soluzione esatta.

2.4 Cenni su altre tecniche di problem-solving

Con l'avvento dell'*intelligenza artificiale* e lo sviluppo di *metodi euristici*, sono emerse altre tecniche computerizzate di *problem-solving* simili agli GAs. Questa sezione ha lo scopo di accennarne alcune al solo fine di evidenziarne determinati aspetti che possono renderli simili o diversi agli GAs.

2.4.1 Le reti neurali

Una *rete neurale* è un metodo di risoluzione dei problemi che si basa su come i neuroni sono collegati al cervello ed è modellato da strati (*layers*) di unità di trasformazione chiamati *nodi* collegati da *link direzionali*. In particolare sono presenti uno strato di *input* e uno strato di *output* direttamente collegati o separati da alcuni

strati nascosti. L'*input* viene conferito al primo *layer* della rete neurale e i *nod*i vengono poi stimolati a trasmettere un segnale ai *nod*i del *layer* successivo a cui sono collegati. Se la somma di tutti i segnali entranti in uno di questi neuroni virtuali è superiore a quella che viene definita soglia di attivazione, il neurone stesso si attiva e trasmette il proprio segnale ai neuroni del *layer* successivo. Il segnale di attivazione si diffonde quindi in avanti fino a raggiungere il livello di uscita e viene riutilizzato successivamente come nuovo *input*. Proprio come nel sistema nervoso degli organismi biologici, le *reti neurali* apprendono e perfezionano i risultati nel tempo attraverso cicli ripetuti fino a quando l'uscita corrisponde a quella desiderata. Questo processo può essere controllato direttamente o può essere eseguito automaticamente utilizzando un algoritmo di apprendimento (Mitchell, 1996). Gli *GAs* sono stati utilizzati sia per costruire che per implementare le reti neurali.

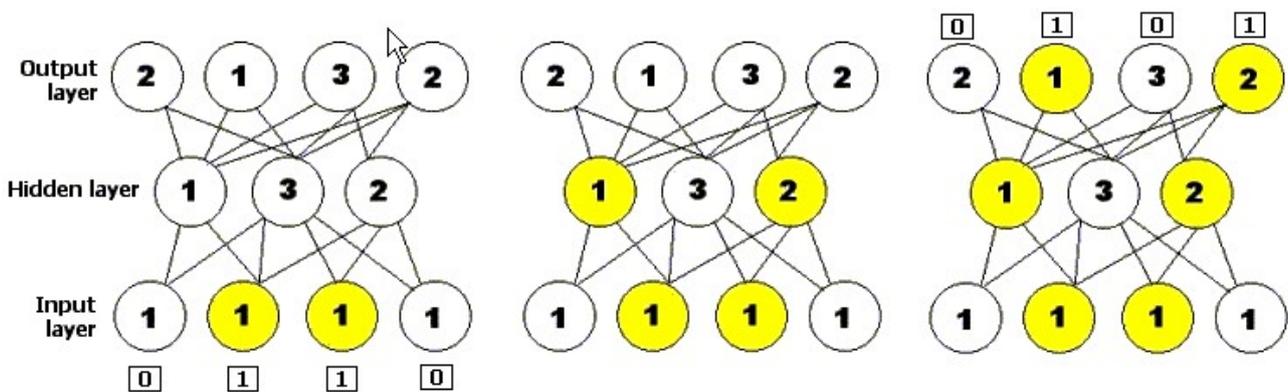


Figura 2.3. Esempio di semplice rete neurale feedforward con un livello di input che consiste in quattro neuroni, un layer nascosto da tre neuroni e un layer di output composto da quattro neuroni. Il numero su ogni neurone rappresenta la sua soglia di attivazione.

2.4.2 L'Hill climbing

L'*hill climbing* è una metodologia simile a quella vista per gli *GAs* anche se più sistematica e meno casuale. Un algoritmo di questo tipo procede identificando una prima soluzione al problema in oggetto, di solito scelta a caso. La soluzione viene successivamente modificata e se essa produce risultati migliori rispetto a quella precedente viene mantenuta mentre in caso contrario viene rifiutata. La procedura viene poi ripetuta fino a quando nessuna modifica può apportare miglioramenti alla soluzione ottenuta che si accetta pertanto come ottima (Koza et al., 2003).

Il nome di questa tecnica diventa comprensibile se si immagina che lo spazio di tutte le possibili soluzioni ad un dato problema sia rappresentato come un paesaggio in tre dimensioni (Figura 2.4). Un dato insieme di coordinate in quel paesaggio rappresenta

una soluzione particolare: le soluzioni migliori sono più in alto e formano colline o vette mentre le soluzioni peggiori si trovano a quota più bassa. L'algoritmo che inizia in un dato punto sul paesaggio si muove inesorabilmente in salita e fa sempre la migliore scelta disponibile ad ogni passo, nella speranza di raggiungere il risultato migliore/più in alto in assoluto.

Un problema di questa procedura è che si rischia di interrompere il processo una volta raggiunto un ottimo locale (Figura 2.4) anziché un ottimo globale oppure che si arrivi ad una zona paragonabile ad un *altopiano* dove le soluzioni appaiono tutte simili mandando in crisi la ricerca.

Altri algoritmi di ricerca locale, come ad esempio lo *stochastic hill-climbing*, il *random walks* e il *simulated annealing*, tentano di ovviare a questo problema

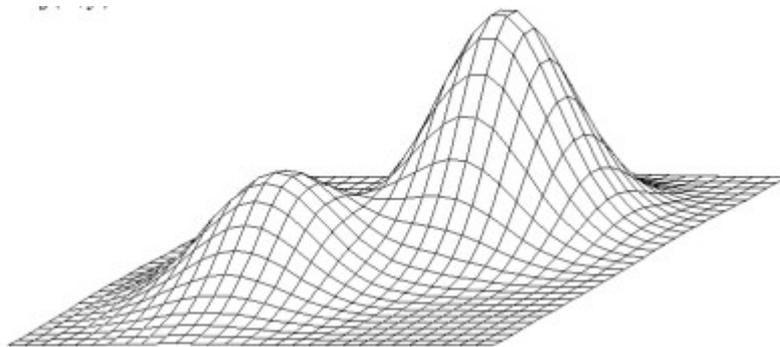


Figura 2.4. Esempio di spazio delle soluzioni in una procedura di ottimizzazione tramite hill-climbing: l'algoritmo inizia in un determinato punto dello spazio e poi si muove in salita alla ricerca del punto più alto che dovrebbe corrispondere all'ottimo globale.

2.4.3 IL Simulated annealing

Un'altra tecnica di ottimizzazione simile agli *GAs* è nota come *simulated annealing*. L'idea prende in prestito il nome dalla lavorazione industriale di *ricottura*: tale procedimento prevede che un materiale venga riscaldato ad una temperatura sopra di un punto critico per essere ammorbidito e poi gradualmente raffreddato per cancellare i difetti nella sua struttura cristallina, producendo un sistema a reticolo più stabile e regolare di atomi. Anche in tale tecnica esiste una funzione di *fitness* che stabilisce l'adattamento all'ambiente, ma anziché una popolazione di candidati, come negli *GAs*, si parte da una sola soluzione candidata. Il *simulated annealing* introduce anche il concetto di *temperatura* come quantità numerica che diminuisce gradualmente nel tempo. Ad ogni passo dell'algoritmo, la soluzione si modifica cioè si sposta in un punto limitrofo nel paesaggio di *fitness* e il confronto tra l'idoneità della nuova soluzione con quella vecchia determina quale delle due viene mantenuta. In particolare però se

l'idoneità della nuova soluzione è inferiore a quella della vecchia, la decisione di disfarsene dipende anche dalla *temperatura*. Se è alta come all'inizio significa che anche se le modifiche apportate hanno causato una diminuzione del *fitness*, possono essere conservate e utilizzate come base per il prossimo ciclo dell'algoritmo, ma se è diminuita l'algoritmo diventa incline a mantenere solo le soluzioni che hanno portato ad un aumento della funzione di *fitness*. Infine se la *temperatura* arriva allo zero la procedura si blocca e qualunque soluzione si sia raggiunta rappresenta la soluzione finale.

3 L'algoritmo di ottimizzazione AMALGAM

Il problema dell'ottimizzazione *multi-obiiettivo* è oggetto di studio ormai da molti anni e pertanto è possibile sostenere che il tema sia stato sufficientemente approfondito. Gli *Algoritmi Evolutivi* comunemente utilizzati implementano una singola procedura per la modellazione dell'evoluzione della popolazione. Tuttavia tale modellazione del processo di selezione naturale presuppone che esista un singolo modo attraverso il quale possiamo raggiungere l'ottimo nello spazio dei parametri possibili, mentre studi e teorie recenti hanno dimostrato che è impossibile sviluppare un singolo algoritmo per l'evoluzione della popolazione che risulti sempre efficiente per tutti i tipi di ottimizzazione. Negli ultimi anni sono stati proposti gli *hybrid genetic algorithms* o *memetic algorithms* con lo scopo di aumentare l'efficienza di questo tipo di implementazioni (Hart W.E. et al., 2005) Questi modelli sono ispirati ai processi di adattamento osservabili in natura e utilizzano un *GA* per l'esplorazione globale dello spazio dei parametri combinato con un *Modello Euristico* di sfruttamento delle condizioni locali. I *memetic algorithms* hanno dimostrato di poter portare ad una accelerazione notevole nella ricerca della soluzione ottimale tramite il percorso dell'evoluzione per una grande varietà di problemi di ottimizzazione. Per questo si è arrivati a ipotizzare che una procedura di ricerca che *adattivamente* cambi il modo di generare i nuovi individui basandosi sulla condivisione delle caratteristiche e sulle peculiarità locali dell'ambiente di *fitness* potesse aumentare definitivamente l'efficienza e l'efficacia dei metodi di ricerca evolutivi. Questo approccio è apparso immediatamente funzionale poiché la natura dell'ambiente di *fitness* (dato dalle funzioni obiettivo che tracciano la superficie della risposta) spesso varia considerevolmente da un problema di ottimizzazione all'altro ma anche in modo dinamico durante la ricerca delle soluzioni ottime globali. Con queste premesse è nato l'algoritmo utilizzato nel presente lavoro che è stato implementato nel 2007 da Jasper A. Vrugt e Bruce A. Robinson presso il Laboratorio Nazionale di Los Alamos in Nuovo Messico. Il metodo combina il concetto di simultanea ricerca *multi-metodo* e di creazione *self-adaptive* di individui allo scopo di ottenere una procedura affidabile,

veloce ed efficiente dal punto di vista computazionale per i problemi di ottimizzazione *multi-obiettivo*.

Questo approccio è stato chiamato *A MultiAlgorithm Genetically Adaptive Multiobjective* o AMALGAM in modo da evocare l'immagine di una procedura che unisce la forza di diversi algoritmi di ottimizzazione (*Vrugt et al., 2007*).

Per l'implementazione di AMALGAM gli autori hanno cercato di approfondire i seguenti tre argomenti:

- ricercare un metodo per far comunicare più algoritmi al fine di condividere in modo efficace le loro informazioni,
- scegliere in modo opportuno i metodi più efficaci per la creazione di individui che si adattino all'ambiente in studio,
- scegliere gli algoritmi che meglio rispondono alle necessità precedentemente individuate.

3.1 Descrizione della procedura implementata in AMALGAM

AMALGAM implementa una procedura di ricerca basata sull'*elitarismo* al fine di identificare una corretta distribuzione delle soluzioni di *Pareto* in un'unica attivazione. Il codice in linguaggio MATLAB[®] viene di seguito descritto.

Nella fase di inizializzazione è necessario fornire alcuni dati relativi al problema in studio: in particolare l'utente deve inserire la *dimensione della popolazione* che si intende utilizzare (100 individui nel presente studio), *il numero massimo di funzioni* che si intende valutare (10000 nel presente studio) oltre al *numero di parametri* da stimare (5 in relazione al modello afflussi-deflussi utilizzato nel presente studio) e gli *intervalli di valori* che tali parametri potranno assumere (si veda il paragrafo 4.5). Inoltre deve essere specificato il *numero di funzioni obiettivo* che si intende utilizzare poiché anche questo è un parametro che viene deciso dall'utilizzatore.

L'algoritmo si avvia generando una popolazione iniziale in modo casuale, che chiameremo P_0 di dimensione N , attraverso il *Latin Hypercube Sampling* o *LHS* (*McKay et al., 1979*). Si tratta di una procedura statistica che genera una distribuzione di combinazioni di possibili valori dei parametri partendo da una distribuzione multidimensionale ed è spesso utilizzata nell'analisi dell'incertezza. In seguito viene effettuata una prima simulazione con tutti i valori casuali che costituiscono la

popolazione iniziale. Nel presente studio le simulazioni sono state effettuate tramite il modello afflussi-deflussi Hymod descritto in seguito (paragrafo 4.5).

Grazie ai valori delle funzioni obiettivo così ottenuti è possibile analizzare la qualità di questi individui/valori: tale valutazione viene effettuata tramite l'algoritmo *Fast Nondominated Sorting (FNS)* (Deb et al., 2002). In pratica gli individui vengono ordinati in base ad un criterio di *non dominanza* (Figura 3.1) che significa che sul primo fronte vengono posizionati i set *non dominati* ai quali viene assegnato un valore di *rank* (=rango) pari a 1, sul secondo i *set dominati* da quelli sul primo fronte ai quali viene assegnato un valore di *rank* pari a 2, sul terzo i *set dominati* da quelli che stanno sul primo e sul secondo fronte ai quali viene assegnato un valore di *rank* pari a 3 e così via.

Poiché non tutte le soluzioni potranno entrare a far parte della nuova popolazione si eliminano quelle appartenenti ai fronti con *rank* superiore (Figura 3.1). Nel caso il fronte non dominato, ossia quello da inserire nella nuova popolazione, non possa entrare completamente in esso, è necessario scegliere quali soluzioni scartare e quali mantenere. A tal fine, oltre al valore di *rank* già attribuito ad ogni individuo viene assegnato un secondo valore chiamato *crowding distance* che indica quanto le soluzioni appartenenti ad uno stesso fronte siano vicine tra loro. La scelta verrà fatta in modo da preferire le soluzioni più distanti, cioè quelle che si trovano in regioni meno affollate dello spazio di ricerca al fine di garantire una buona diversità della popolazione. La Figura 3.2 chiarisce con un esempio il meccanismo del *crowding distance*.

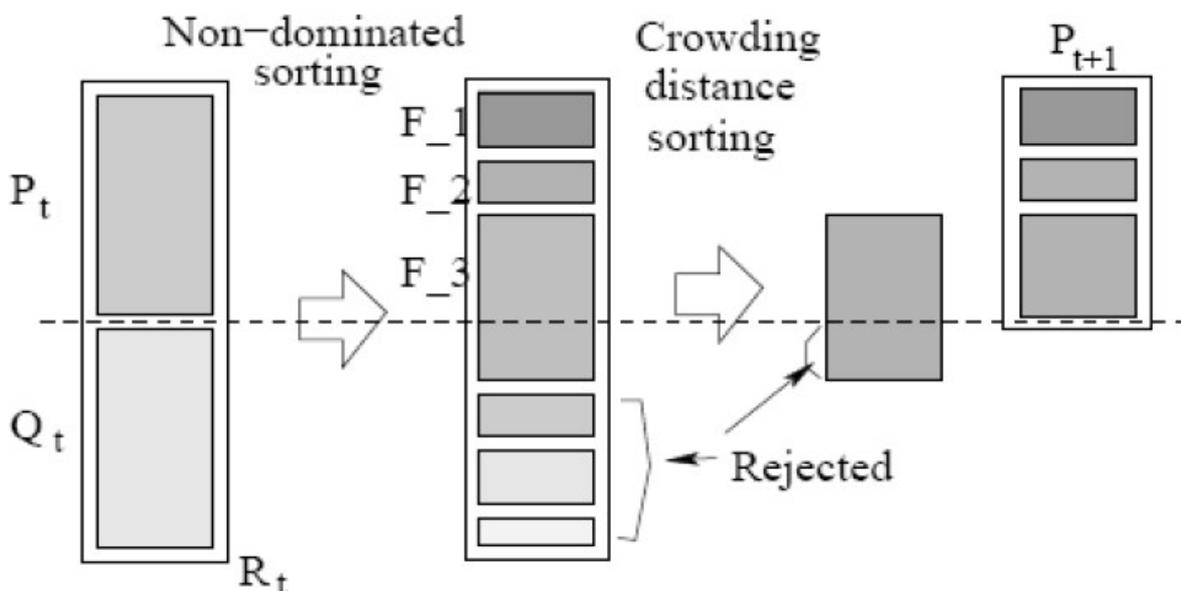


Figura 3.1. Esempio di assegnazione del rank e concetto di crowding distance.

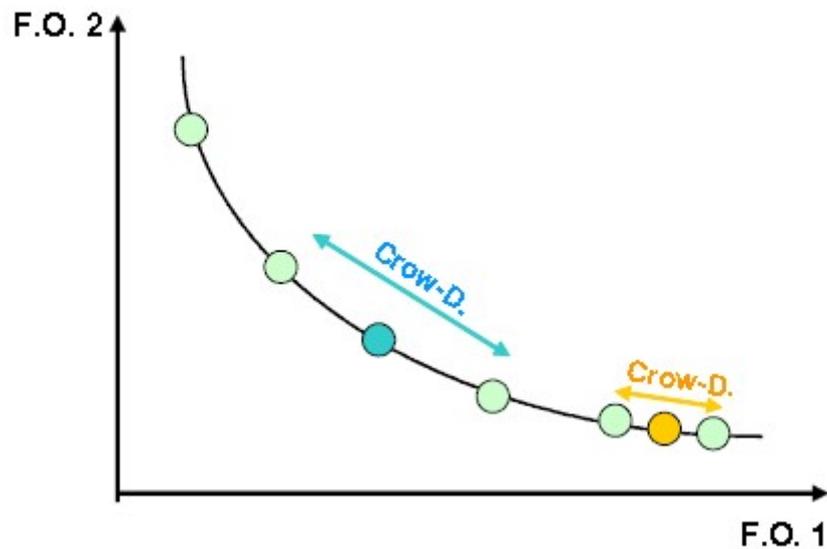


Figura 3.2: Meccanismo del crowding distance per mantenere la diversità nel fronte.

Il calcolo della *crowding distance* comporta l'assegnazione di un valore grande, ∞ (infinito), alle soluzioni che hanno i valori estremi delle funzioni obiettivo, mentre per tutte le altre soluzioni il valore è definito dalla sommatoria delle distanze sugli obiettivi:

$$CrowDi = \sum_{m=1}^M d_{i,m} \quad (3.1)$$

$$d_{i,m} = \frac{f_{i+1}^m - f_{i-1}^m}{f_{\max}^m - f_{\min}^m} \quad (3.2)$$

dove

- $CrowDi$ è il valore di *crowding distance* della soluzione i ,
- M è il numero di obiettivi,
- $d_{i,m}$ rappresenta il valore di distanza per l'obiettivo m relativo alla soluzione i ,
- f_{i+1}^m e f_{i-1}^m rappresentano i valori della m -esima funzione obiettivo delle soluzioni adiacenti ad i e
- f_{\max}^m e f_{\min}^m rappresentano i valori estremi della m -esima funzione obiettivo (Figura 3.3).

La diversità tra le soluzioni non dominate viene mantenuta usando il criterio di *selezione a torneo* (*tournament selection* - paragrafo 2.3.2) confrontando due

individui alla volta. Questo significa che un individuo sarà selezionato se appartiene ad un *rank* di ordine maggiore (ossia valore minore) oppure se, all'interno dello stesso *rank*, è caratterizzato da un valore di *crowding distance* maggiore rispetto ad un altro individuo. In questo modo si ha la certezza che nessuna soluzione ottima verrà sostituita con una soluzione peggiore ma, d'altra parte, può accadere che l'algoritmo perda le sue proprietà di convergenza se accade che tutte le soluzioni appartengano al fronte non dominato.

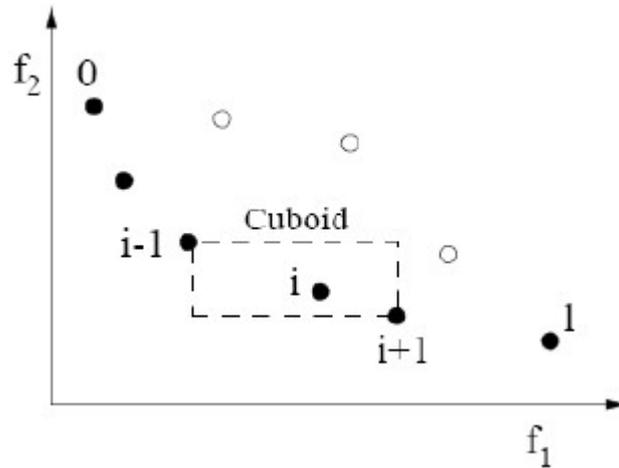


Figura 3.3: Significato del parametro *crowding distance*

A questo punto, tramite l'utilizzo della ricerca *multi-metodo* alla base di AMALGAM, viene creata una nuova generazione di individui Q_0 di dimensione N . In dettaglio anziché implementare un unico operatore per la *riproduzione*, si utilizzano k singoli algoritmi che danno vita ad una nuova generazione del tipo:

$$Q_0 = \{Q_0^1, Q_0^2, \dots, Q_0^k\}. \quad (3.3)$$

Ciascuno di questi algoritmi crea un numero predefinito di nuovi individui che contribuiscono a formare l'insieme N_i

$$N_i = \{N_i^1, N_i^2, \dots, N_i^k\} \quad (3.4)$$

partendo da P_0 e utilizzando diverse tecniche adattive. Successivamente alla generazione dei nuovi individui, viene creata una popolazione R_0 di dimensione $2N$ data dalla combinazione delle due popolazioni in essere

$$R_0 = P_0 \cup Q_0 \quad (3.5)$$

che a sua volta viene sottoposta alla procedura di *rank selection* con l'*FNS*.

Confrontando la nuova generazione con la precedente la procedura di *elitarismo* è assicurata perché tutti i membri, compresi quelli *non-dominati*, della vecchia generazione saranno sempre presenti in R_0 . Infine vengono selezionati i membri che andranno a formare la nuova generazione P_1 scegliendo tra quelli *non-dominati* di R_0 secondo quanto emerso dai loro valori di *rank* e *crowding distance*. La popolazione P_1 viene poi utilizzata per creare i nuovi individui usando di nuovo la procedura appena spiegata finché non si raggiunge la convergenza del risultato.

La novità introdotta in AMALGAM è che il metodo per la creazione della nuova generazione è realizzato in modo da favorire i singoli algoritmi che hanno mostrato un maggior successo nel contribuire nello step precedente. In pratica questo si traduce pesando il loro contributo alla formazione della popolazione successiva e quindi aggiornando l'insieme

$$N_t = \{N_t^1, N_t^2, \dots, N_t^k\} \quad (3.6)$$

con la seguente formulazione:

$$N_t^i = N \frac{\frac{P_t^i}{N_{t-1}^i}}{\sum_{i=1}^N \left(\frac{P_t^i}{N_{t-1}^i} \right)}. \quad (3.7)$$

Il termine al numeratore è il rapporto tra il numero di nuovi punti che genera l'algoritmo per la popolazione successiva P_t^i e il corrispondente numero di individui con i quali ha partecipato alla popolazione precedente N_{t-1}^i . Il resto dell'espressione adegua il peso di ogni singolo algoritmo rispetto al successo combinato di tutti gli altri. Nell'applicazione è stato imposto un valore minimo di individui pari a 5 con i quali ciascun algoritmo deve comunque contribuire al fine di evitare che venga escluso dalla formazione delle generazioni successive.

Gli algoritmi che sono stati implementati dagli autori e utilizzati nel presente lavoro sono:

- *Elitist Non-Dominated Sorting GA (NSGA II)* - (Deb et al., 2002),
- *Particle Swarm Optimization (PSO)* - (Kennedy et al., 1995),
- *Adaptive Metropolis Search (AMS)* - (Haario et al., 2001),
- *Differential Evolution (DE)* - (Storn et al., 1997).

Nei paragrafi 3.2 - 3.3 - 3.4 - 3.5 è riportata una descrizione delle caratteristiche più importanti di questi algoritmi.

Tale scelta è stata giustificata dai numerosi risultati numerici ottenuti dagli autori che dimostrano che questi algoritmi risultano mutualmente consistenti e complementari. Infatti i risultati riportati in Figura 3.4, ottenuti direttamente dagli autori, dimostrano i vantaggi di un'ottimizzazione *multi-metodo* evidenziando l'evoluzione delle *Frontiere di Pareto* generate con i singoli algoritmi *NSGA-II* (quadrantini rossi), *PSO* (cerchietti azzurri), *AMS* (+) e *DE* (diamanti verdi) rispetto ad *AMALGAM* (x blu) per un tipo di test normalmente utilizzato (*ZDT4*). Questo problema presenta 219 differenti soluzioni nello spazio di ricerca dei quali solo uno corrisponde alla *Frontiera di Pareto Globale*. Combinando i singoli algoritmi per una ricerca *multi-metodo* si ottengono risultati più affidabili in modo più veloce. Gli autori sottolineano anche che i risultati mostrano i vantaggi di un metodo di ricerca *evolutivo* e *adattivo*: infatti dopo solo 7500 valutazioni, *AMALGAM* ha generato la *Frontiera Pareto* (indicata in nero) e ha trovato soluzioni distribuite sulla stessa in modo molto più uniforme di tutti gli altri algoritmi presi singolarmente.

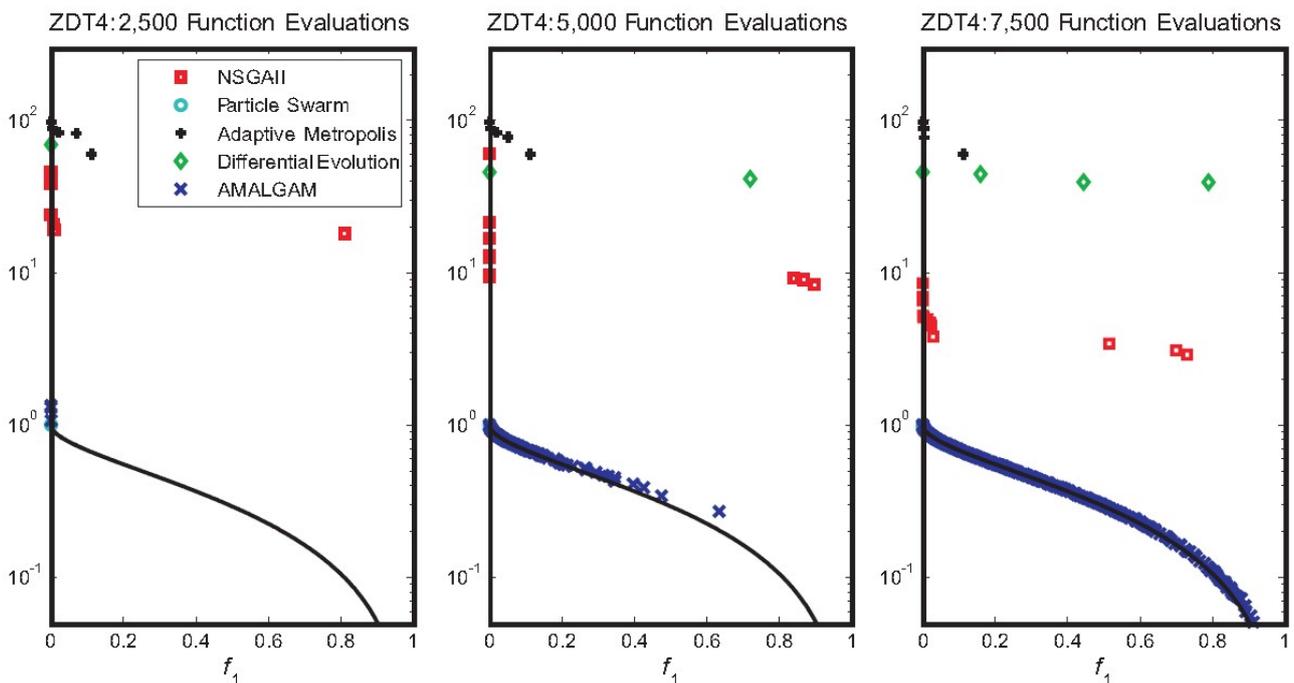


Figura 3.4. Posizione delle soluzioni rispetto alla posizione della frontiera Pareto dopo 25, 50 e 75 generazioni con *NSGA-II* (quadrantini) *PSO* (cerchietti), *AMS* (rombi neri), *DE* (diamanti) e *AMALGAM* (x) per il test *ZDT4*.

Questo miglioramento della *performance* è quantificato nella Tabella 3.1 dove vengono confrontate le convergenze delle statistiche (è riportato un valore medio su 30 simulazioni) utilizzando l'*NSGA-II* e *AMALGAM* per alcuni problemi test presi in esame dagli autori (*Vrugt et al., 2007*).

I valori indicano l'affidabilità della convergenza per un set di soluzioni *Pareto* noto (Y) e la capacità o meno delle soluzioni di distribuirsi in modo uniforme sulla frontiera stessa (Δ). Valori minori per entrambi gli indicatori indicano performance migliori.

Function	Number of function evaluations*		After 150 generations			
	NSGA-II	AMALGAM	Convergence metric: Y		Diversity metric: Δ	
			NSGA-II	AMALGAM	NSGA-II	AMALGAM
ZDT1	5,920	840	0.0053	0.0011	0.34	0.33
ZDT2	7,680	1,240	0.0068	0.0009	0.36	0.35
ZDT3	5,100	1,090	0.0027	0.0010	0.56	0.55
ZDT4	13,050 (24) [†]	6,320	0.0523	0.0022	0.73	0.32
ZDT6	8,940	950	0.0504	0.0011	0.53	0.40
SCH	6,350 (25)	3,610	0.0036	0.0032	0.49	0.37
SCH2	12,000 (22)	3,320	0.0052	0.0041	0.69	0.52
FON	4,930	1,590	0.0026	0.0017	0.38	0.33
KUR	6,544	6,122	0.0108	0.0099	0.48	0.47
ROT	N/C [‡]	9,440	1.50	0.35	1.14	0.55

Results for the NSGA-II and AMALGAM methods are presented. Statistics represent averages over 30 optimization runs. A detailed description of the various test problems and performance metrics appears in *SI Text*.

*Average number of function evaluations needed to result in a relative hypervolume smaller than 0.005.

[†]Number of runs that have failed to converge are shown in parentheses.

[‡]None of the 30 optimization runs have converged after 150 generations.

Tabella 3.1. Numero di valutazioni necessarie per raggiungere la convergenza, valori di convergenza Y e indicatori di diversità 'delta' dopo 150 generazioni per i 10 test utilizzati nello studio preliminare.

I risultati dimostrano che AMALGAM è significativamente più efficiente nella localizzazione delle soluzioni di *Pareto* rispetto all'NSAGA-II anche per i problemi più complessi o di dimensione superiore (ZDT1-ZDT6 e ROT). Questo metodo inoltre raggiunge la convergenza anche nel caso rotazionale entro le 150 generazioni dimostrando la capacità di trattare con variabili di decisione correlate che classicamente mandano in crisi la *mutazione* genetica e la *selezione* degli operatori.

La Figura 3.5 mostra l'evoluzione del numero di nuovi punti generati tramite i singoli algoritmi per il problema test ZDT4. Da questa immagine emerge la motivazione che rende un'ottimizzazione *multi-metodo* migliore: infatti durante la prima parte dell'ottimizzazione l'NSGA-II (quadrati verdi) presenta il massimo successo riproduttivo grazie agli operatori in esso implementati ossia il *crossover* e la *mutazione*. Tuttavia, dopo 20 generazioni, la capacità dell'NSGA-II diminuisce bruscamente in favore rispettivamente e nell'ordine degli algoritmi *DE* (diamanti verdi), *AMS* (+) e *PSO* (cerchietti azzurri). La combinazione di tutti e quattro i metodi quindi risulta essere estremamente efficiente e capace di poter aumentare la diversità delle soluzioni lungo la *Frontiera Pareto* anche dopo che l'NSGA-II perde la sua funzionalità iniziale confermando che una strategia in grado di passare da un metodo all'altro garantisce sempre un forte supporto per una ricerca di tipo *evolutivo*.

Gli autori hanno voluto infine testare le capacità di AMALGAM anche per problemi di dimensioni maggiori e a tale scopo hanno preso in esame il problema test DTLZ6 con tre funzioni obiettivo e 12 parametri da calibrare. Tale scelta è stata giustificata proprio dal fatto che gli algoritmi normalmente utilizzati per la risoluzione di tale problema spesso dimostrano notevoli difficoltà nell'individuare le soluzioni di *Frontiera*.

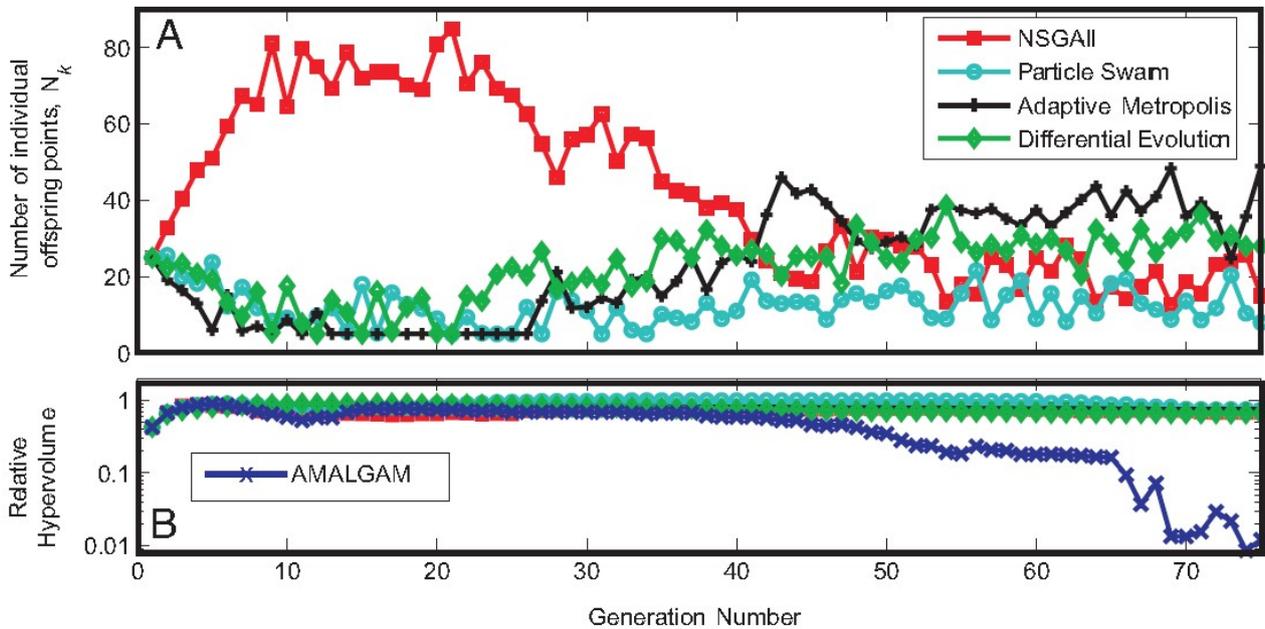


Figura 3.5. Visualizzazione del concetto di creazione delle nuove generazioni in modo auto-adattivo. (A) Evoluzione del numero di nuovi punti generati con gli algoritmi NSGA-II (quadrantini verdi), PSO (cerchietti azzurri), AMS (+), PSO DE (diamanti verdi) rispetto all'uso di AMALGAM come metodo di ricerca multi metodo per il problema test ZDT4. (B) Converggenza metrica di AMALGAM e degli altri quattro algoritmi presi singolarmente. L'immagine mostra la capacità di ogni singolo algoritmo di avvicinarsi maggiormente alla soluzione durante le varie fasi dell'ottimizzazione e produce un esempio numerico del teorema "No Free Lunch Theorems for Optimization" di Wolpert and Macready evidenziando che è impossibile sviluppare un singolo algoritmo di ricerca che sia sempre superiore ad ogni altro per tutto il periodo necessario per la convergenza.

La Figura 3.6 (A) riporta la curva di *Frontiera Pareto* teorica e i risultati ottenuti dopo un'ottimizzazione con 50 e 1000 generazioni rispettivamente per AMALGAM e NSGA-II. AMALGAM individua le soluzioni di frontiera in 5000 valutazioni di funzione mentre l'NSGA-II non è in grado di raggiungere lo stesso risultato prima di 100000 iterazioni. La Figura 3.6 (B) illustra infine ancora una volta come i singoli algoritmi contribuiscano all'evoluzione in funzione della loro capacità di adattarsi al problema in esame.

I risultati presentati dimostrano che un *metodo evolutivo* di ottimizzazione di tipo combinato che utilizza i vari approcci in modo adattivo per la creazione di nuove generazioni è un potente mezzo per la risoluzione di problemi di ottimizzazione complessi. Questa metodologia porta a vantaggi di notevole interesse che combinati

con l'aumento del potere di calcolo potrebbe fornire nuove opportunità per la soluzione di problemi di ottimizzazione considerati fino ad ora di difficile risoluzione.

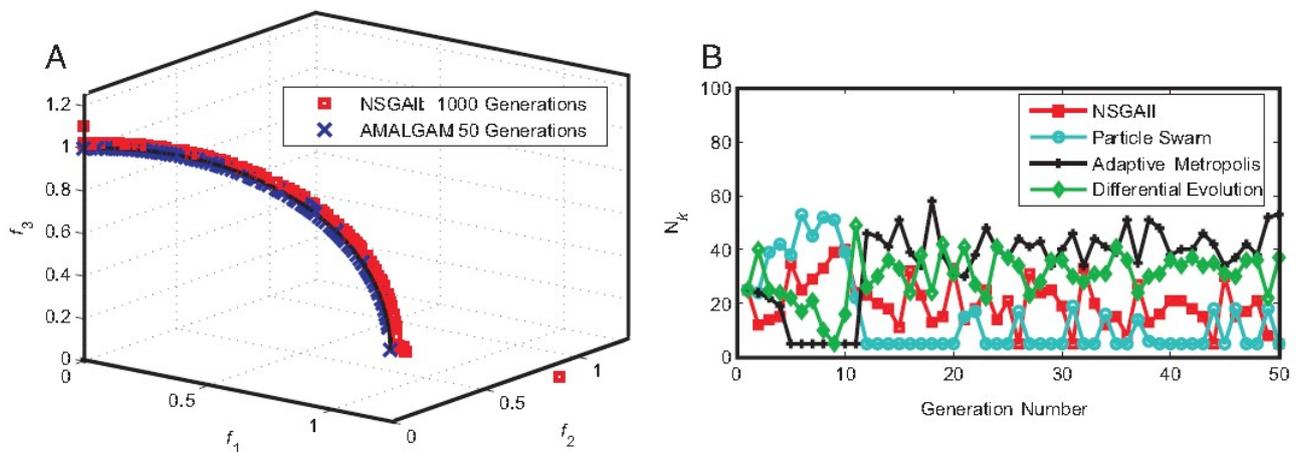


Figura 3.6. Soluzioni non dominate individuate per il problema test DTLZ6 con l'NSGA-II e AMALGAM dopo rispettivamente 100,000 and 5,000 valutazioni di funzione. (A) La linea nera rappresenta la frontiera ed è possibile notare che AMALGAM vi converge perfettamente dopo un numero esiguo di valutazioni. (B) Creazione delle nuove generazioni in maniera auto-adattiva ottenuta variando la relativa importanza degli algoritmi durante l'ottimizzazione in modo da rendere il processo più efficiente e veloce.

Il presente lavoro ha voluto quindi fornire un contributo nella sperimentazione e nell'applicazione di questo metodo utilizzandolo nello scenario di problemi reali. Al fine di comprendere e approfondire in modo completo la conoscenza dell'algoritmo AMALGAM, sembra opportuno riprendere alcuni concetti alla base dei singoli algoritmi che in esso si avvicendano rimandando alle pubblicazioni specifiche per ulteriori approfondimenti.

3.2 Elitist Non-Dominated Sorting GA-II (NSGA-II)

Non-Dominated Sorting in Genetic Algorithms (Srinivas, 1994) è un *Algoritmo Genetico* che si basa sul concetto della selezione degli individui dominanti, molto utilizzato per l'ottimizzazione *multi-obiettivo*. E' un algoritmo decisamente efficace ma che è stato molto criticato per la sua complessità dal punto di vista computazionale, per la mancanza del concetto di *elitarismo* oltre che per la metodologia con la quale sceglie il valore ottimo per i parametri senza condividere le informazioni disponibili. Per questo è stata proposta una versione modificata di questo algoritmo che prende il nome di *NSGA-II* (Deb et al., 2002). Tale versione ovvia ad alcune limitazioni presenti nella versione originale.

Questa versione dell'algoritmo è stata presentata da *Deb et al., 2002* e usa un meccanismo esplicito per preservare la diversità delle soluzioni. La popolazione viene inizializzata come descritto nella sessione 3.2.1 e il procedimento consiste nella creazione delle soluzioni figlie a partire dal numero di soluzioni genitrici presenti nella popolazione di partenza. In seguito le soluzioni figlie vengono unite alla popolazione corrente e l'insieme così costituito, contenente il doppio del numero di soluzioni della popolazione iniziale, viene ordinato attraverso il criterio di non dominanza esattamente come descritto in precedenza per l'algoritmo AMALGAM.



Figura 3.7. Schema di funzionamento dell'algoritmo NSGA-II.

Il diagramma di flusso relativo all'algoritmo *NSGA-II* è presentato nella Figura 3.7. L'algoritmo, come tutti gli algoritmi *euristici*, è iterativo: dopo una prima fase di inizializzazione, una serie di operazioni vengono ripetute fino a quando un certo criterio di termine non è soddisfatto. Questo criterio può consistere nel raggiungimento di dati valori delle funzioni obiettivo, in un numero di iterazioni in cui le soluzioni non sono state migliorate o nel raggiungimento di un numero fissato di iterazioni. Il criterio di termine qui scelto è l'ultimo, cioè si fissa un numero di iterazioni raggiunto il quale l'algoritmo si ferma.

3.2.1 Inizializzazione della popolazione

Nel caso specifico relativo al presente studio, la fase di inizializzazione viene svolta prima di attivare ogni singolo algoritmo perché viene affrontata nella parte di codice comune a tutti gli algoritmi che lavorano contemporaneamente in AMALGAM. Arrivati quindi alla parte di codice relativo strettamente all'*NSGA-II* avremo già a disposizione una matrice in ambiente MATLAB® con un numero di righe pari alla dimensione della popolazione scelta e un numero di colonne pari al numero di variabili del problema. Inoltre, ad ogni set di parametri ipotizzati, saranno associati i valori della/delle funzioni obiettivo ottenute da AMALGAM nella prima simulazione fatta con l'utilizzo del modello afflussi-deflussi scelto.

3.2.2 Ordinamento delle soluzioni non dominate

La popolazione iniziale è ordinata secondo il principio della *non-dominanza*. Un individuo domina un altro individuo se il suo valore della funzione obiettivo è migliore (maggiore o minore a seconda del caso) dell'altro (almeno un valore nel caso di più funzioni obiettivo).

L'*FNS* (Deb et al., 2002) procede come di seguito descritto:

- per ogni individuo p della popolazione principale P si esegue la seguente procedura:
 - Si inizializza $S_p=0$. Questo set dovrebbe contenere tutti gli individui che saranno dominati da p .
 - Si inizializza $n_p=0$. Questo sarà il numero di individui che dominerà p .
 - Per ogni individuo q in P :
 - se p ha dominato q allora si aggiunge q al set S_p ossia $S_p=S_p \cup \{q\}$;
 - altrimenti se q domina p allora si incrementa il contatore dei dominatori di p ossia $n_p=n_p+1$.
 - Se $n_p=0$ ossia non ci sono individui che dominano p allora p appartiene al primo fronte (*rank*) con $prank=1$. Di seguito si aggiorna il numero di individui che appartengono a quel *rank* e l'insieme vero e proprio degli individui che appartengono a quel *rank* vale a dire $F_1= F_1 \cup \{p\}$.
- Questa procedura è ripetuta per ogni individuo della popolazione iniziale.
- Viene inizializzato il contatore dei fronti $i=1$.

- Successivamente, finché l' i -esimo fronte è non vuoto ossia $F_i \neq \emptyset$,
 - $Q = \emptyset$ e rappresenta il set degli individui dell' $(i+1)$ -esimo fronte.
 - Per ogni individuo p nel fronte F_i :
 - Per ogni individuo q in S_p (S_p è il set di individui dominati da p) $nq = nq - 1$ ossia di decrementa il contatore per l'individuo q . Se $nq = 0$ allora nessuno degli individui nei fronti successivi dominerebbe q . Quindi si imposta $q_{rank} = i + 1$ e si aggiorna il set Q con l'individuo q ossia $Q = Q \cup q$.
 - Si incrementa il contatore del fronte di 1.
 - Ora il set Q è il fronte successivo e quindi $F_{i+1} = Q$.

Questo algoritmo è migliore rispetto a quello originale poiché utilizza l'informazione data dal set che l'individuo riesce a dominare (S_p) e il numero di individui che ne fanno parte (n_p).

3.2.3 Calcolo della Crowding Distance

Una volta che l'ordinamento delle soluzioni in base al principio della *dominanza* è completato, viene assegnato il valore di *crowding distance*. Dal momento che tutti gli individui sono selezionati in base al valore di *rank* e di *crowding distance*, a tutti deve essere assegnata questa coppia di valori. Come detto il valore di *crowding distance* è assegnato in base al fronte poiché comparare due individui che appartengono a fronti diversi è meno significativo. Il valore di *crowding distance* invece è calcolato come di seguito descritto.

- Per ogni fronte F_i , n è il numero degli individui, quindi:
 - si inizializza la distanza per tutti gli individui uguale a zero ossia $F_i(d_j) = 0$ dove j corrisponde al j -esimo individuo sul fronte F_i .
- Per ogni funzione obiettivo m :
 - si ordinano gli individui sul fronte F_i basandosi sul valore della funzione obiettivo m ossia $I = \text{sort}(F_i, m)$.
 - si assegna una distanza infinita ai valori limite per ogni individuo in F_i ossia $I(d_1) = \infty$ e $I(d_n) = \infty$.
 - da $k=2$ fino a $k=(n-1)$

$$I(d_k) = I(d_{k-1}) + \frac{I(d_{k+1}) - I(d_{k-1})}{f_m^{\max} - f_m^{\min}} \quad (3.8)$$

dove $I(k)m$ è il valore dell' m -esima funzione obiettivo relativa al k -esimo individuo in I .

L'idea alla base del concetto di *crowding distance* è la ricerca della distanza euclidea di ogni individuo dello stesso fronte basandosi sui loro m obiettivi in un *iper-spazio* m -dimensionale. Gli individui sul contorno sono sempre selezionati poiché hanno un valore di *crowding distance* particolarmente grande.

3.2.4 Modalità di selezione

Una volta che gli individui sono stati ordinati in base al principio di *non-dominanza* e che gli è stato assegnato un valore di *crowding distance*, la selezione prosegue utilizzando un *crode comparison operator* (\prec_n). Il confronto è effettuato come descritto in seguito basandosi su

- il *rank* di non dominanza p_{rank} , ossia gli individui nel fronte F_i avranno valore di *rank* dato da $p_{rank} = i$;
- il valore di *crowding distance* $F_i = d_j$.

Quindi si pone $p \prec_n q$ se

- $p_{rank} < q_{rank}$ oppure
- se p e q appartengono allo stesso fronte F_i allora $F_i(d_p) > F_i(d_q)$ ossia il valore di *crowding distance* deve essere maggiore.

Gli individui sono selezionati usando una metodologia a *torneo* tenendo conto del confronto dell'operatore di *crowding*.

3.2.5 Operatori genetici

Il codice dell'*GA* usa l'operatore definito *Simulated Binary Crossover (SBX)* (Deb, 1995 – Beyer, 2001) per il *crossover* e l'operatore *polynomial mutation* per la mutazione.

3.2.5.1 Simulated Binary Crossover

Questo operatore simula il *crossover binario* osservabile in natura procedendo nel seguente modo:

$$C_{1,K} = \frac{1}{2}[(1 - \beta_K)p_{1,K} + (1 + \beta_K)p_{2,K}]$$

$$C_{2,K} = \frac{1}{2}[(1 + \beta_K)p_{1,K} + (1 - \beta_K)p_{2,K}]$$
(3.9)

dove $C_{i,K}$ è l' i -esimo *figlio* con k componenti mentre $p_{i,K}$ è il *genitore* selezionato. $\beta_K (\geq 0)$ è un campione selezionato tra i numeri generati in modo casuale di densità pari a

$$p(\beta) = \frac{1}{2}(\eta_c + 1)\beta^{\eta_c}$$
(3.10)

se $0 \leq \beta \leq 1$,

$$p(\beta) = \frac{1}{2}(\eta_c + 1)\frac{1}{\beta^{\eta_c+2}}$$
(3.11)

se $\beta > 1$.

Questa distribuzione può essere ottenuta da un numero casuale u uniformemente campionato tra 0 e 1 mentre η_c è l'indice di distribuzione per il *crossover* (questo valore in pratica determina quanto saranno sparsi i *figli* rispetto ai *genitori*).

Otteniamo:

$$\beta(u) = 2u^{\eta_c+1}$$
(3.12)

$$\beta(u) = \frac{1}{[2(1-u)]^{\frac{1}{\eta_c+1}}}$$
(3.13)

3.2.5.2 Polinomial mutation

Si utilizza la seguente esprime:

$$c_k = p_k + (p_k^u - p_k^l)\delta_k$$
(3.14)

dove c_k è il *figlio* e p_k è il *genitore*, p_k^u è il limite superiore dello spazio dei genitori e p_k^l rappresenta il limite inferiore. δ_k è una piccola variazione che viene calcolata da una distribuzione polinomiale usando:

$$\delta_k = (2r_k)^{\frac{1}{\eta_m+1}} - 1 \quad (3.15)$$

se $r_k < 0.5$,

$$\delta_k = 1 - [2(1-r_k)]^{\frac{1}{\eta_m+1}} \quad (3.16)$$

se $r_k \geq 0.5$,

dove r_k è un numero campionato uniformemente tra 0 e 1 e η_m è l'indice di distribuzione della mutazione.

3.2.6 Ricombinazione e selezione

La popolazione dei *figli* è combinata con la popolazione dei *genitori* e viene avviata la procedura di *selezione* per individuare i membri della popolazione successiva. L'*elitarismo* è assicurato dal fatto che sia la generazione precedente che quella successiva contribuiscono alla selezione dei *migliori* per la creazione della nuova generazione. La popolazione che si ricava è dunque ordinata in base al principio di *non dominanza*. La nuova generazione viene *riempita* procedendo per fronti successivi fino al raggiungimento del numero di membri fissato. Se il numero di individui selezionabili sullo stesso fronte supera quello consentito, si procede selezionando i membri in base al valore di *crowding distance*.

Al termine di questa operazione tutta la procedura viene ripetuta per la creazione della generazione successiva.

3.3 Particle Swarm Optimization (PSO)

Il *Natural Computing* è un campo di ricerca che si basa sulla possibilità di prendere spunto dalle strutture e dai meccanismi naturali per la risoluzione di problematiche di natura scientifica che si mostrano resistenti ai metodi tradizionali. Dal punto di vista computazionale sono state implementate strutture molto semplici ispirate ai processi o ai movimenti naturali che permettono di risolvere problemi molto articolati. Il *Particle Swarm Optimization (PSO)* è uno di questi metodi introdotto da *Kennedy & Eberhart, (1995)* caratterizzato da elementi ispirati al comportamento sociale di sciame esistenti

in natura e dalle connessioni con l'*evolutionary computation* con la quale condivide gran parte dei campi di applicazione.

Molti algoritmi che appartengono a questa categoria condividono proprietà importanti come l'abilità di agire anche in caso di dati incompleti o con errori, l'abilità di risolvere problemi combinatoriali, di simulare comportamenti complessi tramite l'interazione di più azioni semplici nonché l'uso di un mix di interazione e cooperazione tra gli elementi. Questi algoritmi, nonostante la complessità dei processi che simulano come ad esempio l'eventuale percorso di apprendimento degli individui, risultano spesso efficienti dal punto di vista computazionale permettendo anche un'implementazione di tipo distribuito (*Banks et al., 2007*).

Come già sottolineato nel Capitolo 2, gli *Algoritmi Evolutivi* sono principalmente basati sulla competizione tra gli individui. Il *PSO* invece adotta una strategia di tipo cooperativo potendolo così definire ontogenetico cioè basato sulla capacità degli individui di adattarsi all'ambiente circostante per massimizzare le loro qualità nel corso della vita. Nel *PSO* gli individui comunicano la loro migliore soluzione al gruppo e ciascun membro segue una direzione data dalla combinazione delle soluzioni migliori precedenti del gruppo e la propria con un elemento stocastico ulteriore che garantisce l'esplorazione dello spazio delle soluzioni. Le ricerche più recenti mostrano che il *PSO* trova le soluzioni vicino all'ottimo in modo più veloce rispetto agli *GAs* anche se il successivo sfruttamento di queste soluzioni è più debole. Rispetto agli *EAs* in genere il *PSO* risulta più efficiente ma anche maggiormente incline ad una convergenza prematura verso un ottimo locale.

Il *PSO* procede generando stocasticamente una serie di punti/particelle/valori in movimento che vengono inizialmente collocati in una zona definita dello spazio possibile. Ad ogni particella è assegnato un vettore velocità iniziale che successivamente, in modo iterativo, viene modificato/aggiustato in modo *random*. Ogni particella viene quindi spostata, in base alla sua velocità, dalla sua posizione iniziale verso una nuova posizione che viene scelta in modo che tale movimento sembri realistico.

3.3.1 Struttura dell'algoritmo

La differenza fondamentale tra gli *GAs* e il *PSO* può sembrare semplicemente legata al modo in cui si riproduce l'evoluzione della popolazione. Negli *GAs* i nuovi individui sono generati tramite la ricombinazione di alcuni *genitori* selezionati secondo alcuni

criteri che vengono utilizzati per sostituire alcuni membri della popolazione precedente. Nel *PSO* invece i nuovi individui sono generati dalla *perturbazione* delle soluzioni esistenti. Quest'ultimo approccio però può portare a problemi di stabilità e al fine di ridurre tale rischio *Eberhart et al., (1996)* hanno proposto una versione rivista dell'algoritmo originale che concepisce di poter limitare la velocità di ciascuna particella all'interno di un *range* predefinito e successivamente, al fine di mediare la tendenza delle soluzioni a convergere velocemente (versione originale) o a espandersi eccessivamente nello spazio (versione modificata), si è pensato di introdurre un peso nell'equazione che si utilizza per il calcolo della nuova posizione della particella nello spazio. In AMALGAM è stata implementata questa versione modificata che si traduce nell'utilizzo della seguente equazione:

$$x_{t+1} = x_t + \chi V_{t+1} \quad (3.17)$$

dove x_t rappresenta la posizione delle particelle allo step precedente (ossia i valori che i parametri assumono allo step precedente) mentre

$$V_{t+1} = \omega V_t + \varphi_1 \beta_1 (p_i - x_i) + \varphi_2 \beta_2 (p_g - x_i) \quad (3.18)$$

rappresenta la velocità di movimento della particella. Il significato dei simboli è il seguente:

- ω = fattore di peso, pari in AMALGAM a $0.5 + \sqrt{2}$,
- φ_1 = costante che considera la conoscenza dell'individuo/particella,
- φ_2 = costante che considera la conoscenza del gruppo,
- p_i = precedente posizione migliore dell'individuo,
- p_g = precedente posizione migliore del gruppo,
- x_i = posizione corrente nella dimensione considerata,
- β_1 e β_2 = numeri generati in modo random con distribuzione uniforme e valore comunque inferiore ad un limite superiore β_{\max} che è un parametro dell'algoritmo.

I termini φ_1 e φ_2 in AMALGAM sono posti uguali a 2 mentre il termine χ è posto pari a 1. Quest'ultimo ha il significato di fattore restrittivo della velocità ed è stato introdotto in una versione modificata dell'algoritmo proposta da *Clerc & Kennedy, 2002*. rendendo non più necessaria l'introduzione di un limite superiore della velocità da parte dell'utilizzatore.

Una volta aggiornata la posizione delle particelle in base alla velocità si effettua un cambiamento *random* della posizione di alcune particelle e si aggiorna la distribuzione della posizione di tutte le particelle. A questo punto la *nuova generazione* è pronta per generare una simulazione e controllare se questa distribuzione di particelle ha prodotto un risultato migliore o peggiore della precedente. La procedura viene ripetuta in modo iterativo fino alla convergenza verso una distribuzione ottimale.

3.3.2 Swarm analyses

Una delle prime e particolarmente dettagliate analisi del *PSO* è stata condotta da *Kennedy (1997)* con lo scopo di stabilire il beneficio dell'interazione sociale tra gli individui nel raggiungimento dell'obiettivo comune. Questa analisi è stata condotta confrontando l'algoritmo originale con tre nuovi modelli: il primo prendeva in considerazione solo l'aspetto cognitivo, escludendo l'influenza dagli altri individui del gruppo sull'individuo in esame; il secondo prendeva in esame solo l'aspetto sociale considerando quindi solo l'individuo migliore di tutto il gruppo; il terzo infine escludeva l'individuo dalla competizione. I primi due modelli, di natura sociale, permettevano il raggiungimento di risultati migliori ma l'autore non consigliava la sostituzione della versione originale perché la trattazione era stata condotta con una funzione obiettivo troppo semplice e poteva accadere che, in condizioni più complesse, si sarebbe giunti a *ottimi-locali*.

Un approccio più teorico è stata proposto da *Ozcan & Mohan (1998 e 1999)*. I due autori osservano che le particelle invece di *volare* nello spazio ammissibile più correttamente *navigano* su onde sinusoidali e forniscono quindi un'idea del ruolo di φ_1 e φ_2 nell'influenza dell'ampiezza e della frequenza della sinusoide stessa. I due autori inoltre mostrano che un valore *adattivo* di inerzia potrebbe muovere una particella bloccata in una regione non corretta verso una regione migliore. Questo lavoro ha prodotto alcune semplificazioni all'algoritmo la più significativa delle quali è la rimozione dell'elemento stocastico. *Clerc e Kennedy (2002)* in uno studio approfondito delle dinamiche sottostanti, hanno reintrodotta questo elemento sostenendo che senza restrizioni di natura stocastica, le particelle vanno inesorabilmente verso l'infinito (esplosione). I due autori non negano l'idea del movimento sinusoidale delle particelle ma lo limitano a quelle particelle il cui movimento viene contenuto e a situazioni nelle quali il tempo viene trattato in modo discreto. Tuttavia, qualora il tempo venisse trattato in modo continuo la particella si

muoverebbe a spirale verso il suo obiettivo. Da questa analisi si è ritenuto di dover introdurre il fattore di contenimento della velocità già introdotto in precedenza proprio con lo scopo di evitare l'esplosione dello sciame di particelle grazie anche all'introduzione degli elementi stocastici β_1 e β_2 .

Non è solo la traiettoria delle particelle che è importante per il comportamento a sciame; il metodo con cui p_g è calcolato influisce sul modo in cui è permesso alle particelle di esplorare le zone al di fuori dall'area attualmente considerata come migliore. *Kennedy (1997)* ha studiato alcune tipologie di reti sociali testandone i tipi di interazioni: esse sono di tipo circolare, rotazionale, a stella e con bordi casuali. La ricerca ha concluso che la tipologia influisce sulle prestazioni dello sciame, ma che la configurazione ottimale è influenzata dalla funzione obiettivo. I collegamenti possono subire effetti molto imprevedibili. Un dato desunto dal lavoro è che le reti nelle quali il livello di comunicazione rallenta presentano una minor possibilità di convergere in modo prematuro.

Il *PSO* ha in comune con molti algoritmi che si basano sull'evoluzione della popolazione iniziale la tendenza non intenzionale a produrre risultati migliori in problemi che hanno soluzioni ottimali verso determinate aree rispetto alla posizione iniziale dello sciame come ad esempio il centro dello spazio delle soluzioni o l'origine assoluta. *Monson & Seppi (2005)* hanno analizzato questo comportamento e hanno proposto una tecnica denominata *Center Offset*. Per illustrare questa tecnica hanno utilizzato uno strumento ausiliario al *PSO* denominato *Clerc's TRIBES (2003)* che consente un'analisi più coerente del moto delle particelle. Il *Center Offset* ha dimostrato particolare utilità nell'affrontare i problemi che dimostrano convergenza verso l'origine assoluta ma minore utilità per quelli che convergono verso il centro dello spazio per i quali è più utile un approccio di tipo regionale. Gli autori hanno concluso che nessuna tecnica è sufficiente di per sé e per affrontare l'analisi in modo completo è necessario utilizzarle entrambe.

3.3.3 Swarm stagnation

La rapida convergenza del *PSO* è uno dei suoi vantaggi principali, ma questo può anche essere problematico perché se la soluzione risulta *sub-ottimale*, lo sciame può facilmente ristagnare senza alcuna pressione per continuare l'esplorazione. Per evitare questo, *van den Bergh & Engelbrecht (2002)* hanno introdotto la versione *Guaranteed Convergence PSO (GCPSO)*, che utilizza una diversa equazione per aggiornare la

velocità delle particelle e la posizione p_g . Questa equazione comporta che la particella esegua una ricerca casuale nell'intorno di p_g con un raggio definito da un fattore di scala che si aggiorna in modo che la posizione della particella non sia in stallo in p_g . Lo svantaggio principale di questa versione è che può essere necessaria una preventiva conoscenza per ottimizzare l'adattamento del fattore di scala. Il punto di forza di questo metodo invece è la capacità di operare con sciame di piccole dimensioni.

Un sistema di prevenzione delle collisioni è stato introdotto da *Krink et al. (2002)*, in modo che le particelle possano evitare di *rimbalzare* a causa del raggruppamento attorno ad una soluzione ottimale avvicinandosi magari ad una soluzione *sub-ottimale*. Sono state testate tre differenti strategie per prevenire l'effetto *rimbalzo*: un cambiatore di direzione casuale, un sistema di previsione del rimbalzo e un cambiatore di velocità casuale. La prima porta a piccoli vantaggi probabilmente perché fa in modo che la ricerca diventi più o meno casuale. Viceversa le altre due strategie portano a discreti miglioramenti legati al fatto che non consentono una stagnazione o un raggruppamento delle particelle.

I bassi livelli di diversità delle particelle possono essere indicativi di un effetto di stagnazione dello sciame. *Riget e Vesterstrøm (2002)* hanno implementato una misura di diversità per controllare l'alternanza di attrazione e repulsione delle particelle da e verso la posizione migliore nello sciame. Inizialmente le particelle sono attratte l'una dall'altra finché non venga raggiunto un livello di diversità minimo. Successivamente lo sciame inizia a respingersi fino a raggiungere un livello di diversità massima. La fase di attrazione è implementata attraverso la normale applicazione del canonico aggiornamento dell'equazione della velocità del *PSO* mentre la fase di repulsione utilizza un'inversione dell'aggiornamento della velocità. L'approccio ha dimostrato risultati migliori rispetto a quelli raggiungibili con il *PSO* originale e con gli *GAs* in genere.

3.4 Adaptive Metropolis Search (AMS)

Un processo stocastico *markoviano* o *processo di Markov* è un processo stocastico nel quale la probabilità di transizione che determina il passaggio ad uno stato di sistema dipende unicamente dallo stato di sistema immediatamente precedente (proprietà di Markov) e non dal come si è giunti a tale stato (in quest'ultima ipotesi si parla di

processo non markoviano). Tale processo prende il nome dal matematico russo Andrej Andreevič Markov che per primo ne sviluppò la teoria.

Markov chain Monte Carlo (MCMC) è una tecnica per stimare tramite la simulazione il valore atteso di una statistica in un modello complesso. Successive selezioni casuali formano una catena di Markov, la distribuzione stazionaria della quale è la distribuzione di destinazione. Questa tecnica è particolarmente utile per la valutazione delle distribuzioni in modelli *bayesiani* complessi. Nell' algoritmo *Metropolis-Hastings* (*Hastings, 1970*), gli elementi vengono selezionati da una distribuzione arbitraria "proposta" e sono conservati o no secondo una regola di accettazione.

L'*Adaptive Metropolis Serach (AMS)* proposto da *Haario et al., (2001)* è caratterizzato dal fatto che la distribuzione gaussiana proposta viene aggiornata lungo il processo utilizzando le informazioni accumulate. A causa della natura adattiva del processo, l'algoritmo *AMS* è *non markoviano*, ma ha comunque la proprietà di *ergodicità* (si dice di un sistema o processo in cui la media temporale delle grandezze che lo descrivono coincide con un'opportuna media presa su un insieme di stati possibili del sistema stesso).

Gli autori sono partiti dalla considerazione che non è affatto semplice scegliere una distribuzione efficace nel caso di utilizzo di algoritmi come ad esempio il *random walk Metropolis* (*Metropolis et al., 1953*), ma che allo stesso tempo tale scelta risulta di fatto cruciale per ottenere risultati ragionevoli in un lasso di tempo limitato. Questa scelta difficoltosa riguarda sia le dimensioni che l'orientamento spaziale della distribuzione proposta e le difficoltà sono legate al fatto che la densità di destinazione è sconosciuta (*Gelman et al., 1996 - Gilks et al., 1995 e 1998; Haario et al., 1999; Roberts et al., 1997*).

L'adeguamento proposto nell'*AMS* riguarda sia la dimensione che l'orientamento spaziale della distribuzione proposta e il risultato è un algoritmo semplice da implementare e utilizzare. L'implementazione dell'algoritmo *AMS* si basa su quella classica del *random walk Metropolis*, (*Metropolis et al., 1953*) e su quelle delle successive modifiche proposte nell'algoritmo *Adaptive Proposal (AP)* introdotto da *Haario et al. (1999)*.

Nel algoritmo *AP* la distribuzione proposta è una distribuzione gaussiana centrata sulla situazione attuale, e la covarianza è calcolata da un numero fisso finito di stati precedenti. Nell' algoritmo *AM* invece la covarianza della distribuzione proposta è calcolata utilizzando tutti gli stati precedenti.

Un vantaggio importante dell'algoritmo *AM* è che inizia sfruttando l'informazione cumulata. L'avvio rapido dell'adeguamento assicura che la ricerca diventi più efficace

subito nella fase iniziale della simulazione richiedendo così un numero inferiore di valutazione delle funzioni necessarie. In particolare si suppone che al tempo t gli stati della catena di AM già sottoposti a campionamento siano X_0, X_1, \dots, X_t , alcuni dei quali possono essere multipli. La nuova distribuzione proposta per il prossimo punto candidato è quindi una distribuzione gaussiana con media al punto corrente X_t e covarianza data da $s_d R$ dove R è la matrice di covarianza determinata dalla distribuzione spaziale degli stati $X_0, X_1, \dots, X_t \in R^d$. Il parametro di scala s_d dipende solo dalla dimensione d dei vettori. Questa strategia di adeguamento spinge la distribuzione proposta ad avvicinarsi ad una distribuzione gaussiana approssimativamente scalata della distribuzione di destinazione aumentando l'efficienza della simulazione.

Una delle difficoltà nella costruzione di algoritmi $MCMC$ di tipo adattativo è garantire che il l'algoritmo mantenga la propria *ergodicità* (tale proprietà non è mantenuta ad esempio nell'algoritmo AP). I teoremi introdotti dagli autori fanno in modo invece che l' AMS mantenga tale proprietà grazie all'assunzione che la densità di destinazione sia delimitata da condizioni al contorno. Come già accennato la catena AM non è *markoviana*, ma si dimostra che la dipendenza asintotica tra gli elementi della catena è tanto debole da consentire l'applicazione dei più noti teoremi dei grandi numeri (*McLeish, 1975*). Risultati analoghi possono essere raggiunti anche per varianti dell'algoritmo dove la covarianza è calcolata tenendo conto di un segmento sempre maggiore di informazioni subito precedenti a quella presa in esame.

3.4.1 Descrizione dell'algoritmo

Supponendo che la distribuzione obiettivo sia supportata dal *sub-set* $S \in R^d$ e che abbia la densità (non scalata) $\pi(x)$ rispetto alla misura di Lebesgue su S . Con un leggero abuso di notazione, è possibile indicare la distribuzione di destinazione con π . Supponiamo che al tempo $t-1$ si sia giunti alla risoluzione delle condizioni X_0, X_1, \dots, X_{t-1} dove X_0 rappresenta lo stato iniziale. Il nuovo punto Y è trattato a partire dalla distribuzione proposta (asintoticamente simmetrica) che dipende dall'intera situazione descritta da X_0, X_1, \dots, X_{t-1} .

Il nuovo punto Y sarà accettato con probabilità

$$\alpha(X_{t-1}, Y) = \min\left(1, \frac{\pi(Y)}{\pi(X_{t-1})}\right) \quad (3.19)$$

nel caso in cui si ponga $X_t=Y$ e diversamente $X_t = X_{t-1}$.

Si può osservare che la probabilità scelta per l'accettazione ricorda la probabilità di accettazione dell'algoritmo *Metropolis*. Tuttavia, qui la scelta per la probabilità di accettazione non è basata sulla *simmetria (reversibilità)* delle condizioni perché non può essere soddisfatta.

Come precedentemente accennato la distribuzione proposta e impiegata nel AMS è una distribuzione Gaussiana del tipo

$$q_t(\cdot | X_0, \dots, X_{t-1}) \quad (3.20)$$

con media al punto corrente X_{t-1} e covarianza data da

$$C_t = C_t(X_0, \dots, X_{t-1}). \quad (3.21)$$

Il punto cruciale in merito all'adattamento è legato a come la covarianza della distribuzione proposta dipende dalla storia della catena. Nell'AMS tale problema è risolto ponendo dopo un periodo iniziale

$$C_t = s_d \text{cov}(X_0, \dots, X_{t-1}) + s_d \varepsilon I_d \quad (3.22)$$

dove s_d è un parametro che dipende solo dalla dimensione d e $\varepsilon > 0$ è una costante con valore molto piccolo rispetto alla dimensione di S . I_d si riferisce alla matrice identità di dimensione d .

Al fine di avviare la procedura si sceglie una covarianza iniziale C_0 arbitraria ma strettamente positiva e legata alle conoscenze iniziali possedute. Si seleziona un indice $t_0 > 0$ per la lunghezza di un periodo iniziale e si definisce

$$C_t = C_0 \quad (3.23)$$

per $t \leq t_0$

$$C_t = s_d \text{cov}(X_0, \dots, X_{t-1}) + s_d \varepsilon I_d \quad (3.24)$$

per $t > t_0$.

La covarianza C_t può essere vista come una funzione di $t \in \mathbb{R}^d$ che può assumere valori in matrici definite uniformemente positive.

La definizione della matrice empirica di covarianza determinata dai punti $X_0, X_1, \dots, X_t \in R^d$ è la seguente:

$$\text{cov}(x_0, \dots, x_k) = \frac{1}{k} \left(\sum_{i=0}^k x_i x_i^T - (k+1) \bar{x}_k \bar{x}_k^T \right) \quad (3.25)$$

dove

$$\bar{x}_k = \frac{1}{k+1} \left(\sum_{i=0}^k x_i \right) \quad (3.26)$$

e gli elementi sono $X_i \in R^d$ sono considerati come vettori colonna.

Con l'equazione 3.24 per $t \geq t_0 + 1$ si ottiene che la covarianza C_t soddisfa la seguente equazione:

$$C_{t+1} = \frac{t-1}{t} C_t + \frac{s_d}{t} \left(t \bar{X}_{t-1} \bar{X}_{t-1}^T - (t+1) \bar{X}_t \bar{X}_t^T + X_t X_t^T + \varepsilon I_d \right). \quad (3.27)$$

In questo modo, finché la media soddisfa una formula analoga, è possibile calcolare C_t senza uno sforzo computazionale esagerato.

La scelta della lunghezza del segmento iniziale $t_0 > 0$ è libera ma più viene scelto grande più lento sarà l'effetto dell'adattamento. La dimensione di t_0 è legata in un certo senso alla correttezza della covarianza iniziale C_0 .

Il ruolo del parametro ε è di assicurare che C_t non diventi singolare mentre per il parametro s_d si adotta la seguente espressione (*Gelman et al., 1996*):

$$s_d = \frac{(2.4)^2}{d} \quad (3.28)$$

poiché è stato dimostrato che in un certo senso questa scelta consente di ottimizzare le proprietà di ricerca dell'algoritmo *Metropolis*.

3.5 Differential Evolution (DE)

Il *Differential Evolution (DE)* è un algoritmo *euristico* proposto da *Storn e Price (1997)* come alternativa ai modelli esistenti per minimizzare equazioni continue non lineari e non differenziabili. In particolare gli autori hanno voluto realizzare un algoritmo che rispondesse ai seguenti requisiti:

1. capacità di gestire funzioni obiettivo non differenziabili, non lineare e multimodali,
2. capacità di lavorare in parallelo per gestire funzioni obiettivo particolarmente pesanti dal punto di vista computazionale,
3. facilità d'uso, vale a dire presenza di poche variabili di controllo per gestire la minimizzazione. Queste variabili devono essere comunque robuste e di facile scelta,
4. buone proprietà di convergenza, cioè la convergenza al minimo globale deve essere confermata da prove consecutive ed indipendenti.

Per far fronte al primo punto, il *DE* è stato concepito come metodo di ricerca stocastico e diretto in modo da sfruttare il vantaggio di essere facilmente applicato nella sperimentazione dove il valore della funzione obiettivo può essere derivato da una trattazione fisica piuttosto che un calcolo computazionale.

Il punto 2 è importante in caso di ottimizzazioni impegnative dal punto di vista computazionale poiché la valutazione di una funzione obiettivo può richiedere da minuti a ore e quindi al fine di ottenere risultati utilizzabili in un ragionevole lasso di tempo, l'unico approccio praticabile può essere il ricorrere a un computer in parallelo o ad una rete di computer. Il *DE* soddisfa il requisito 2 utilizzando un vettore di popolazione in cui la perturbazione stocastica dei vettori può essere fatta in maniera indipendente.

Al fine di soddisfare il terzo requisito è utile che il metodo sia *auto-organizzato* in modo che l'utilizzatore debba intervenire solo in piccola parte. Il sistema di auto-organizzazione del *DE* prende il vettore differenza di due vettori popolazione scelti in modo casuale e lo utilizza per perturbare un vettore esistente. La perturbazione viene fatta per ogni vettore di popolazione. Questa idea è fondamentalmente in contrasto con il metodo utilizzato dagli *ESs* tradizionali nei quali sono funzioni di distribuzione di probabilità predeterminate a determinare le perturbazioni vettoriali.

Infine il quarto punto sottolinea la necessità che un buon algoritmo di minimizzazione goda della proprietà di *convergenza* e gli autori sottolineano che test approfonditi in varie condizioni hanno dimostrato l'eccellente comportamento del *DE*.

3.5.1 Descrizione dell'algoritmo

Il *Differential Evolution* è un metodo di ricerca diretta parallela che utilizza vettori dei parametri di dimensione NP . L'espressione seguente si riferisce alla popolazione per ciascuna generazione G :

$$x_{i,G}, i = 1, 2, \dots, NP. \quad (3.29)$$

NP non cambia durante la procedura di minimizzazione. Il vettore iniziale della popolazione è scelto in modo casuale e dovrebbe coprire l'intero spazio dei parametri. Come regola generale, il *DE* utilizza una distribuzione di probabilità uniforme per tutte le decisioni casuali. Nel caso in cui una soluzione preliminare sia disponibile, la popolazione iniziale, potrebbe essere generata con l'aggiunta di deviazioni casuali normalmente distribuite alla soluzione nominale $x_{nom,0}$. *DE* genera i nuovi vettori dei parametri aggiungendo la differenza pesata tra due vettori di popolazione ad un terzo vettore. Questa è l'operazione che assolve il ruolo della *mutazione*. I parametri dei vettori che sono mutati sono successivamente mischiati con i parametri di un altro vettore predeterminato (il vettore di *destinazione*) e formano il cosiddetto vettore di *processo*. Il mix di parametri è comunemente detto *cross-over* ed è meglio spiegato nel successivo paragrafo 3.5.3. Se il vettore di processo porta ad un valore minore della funzione obiettivo allora questo vettore sostituisce il vettore di destinazione nella generazione successiva. Quest'ultima operazione è quella definita di *selezione*. Poiché ogni vettore di popolazione così come ogni vettore di destinazione deve essere utilizzato una volta, in una generazione vengono effettuati NP confronti.

3.5.2 Mutazione

Per ogni vettore di destinazione

$$x_{i,G}, i = 1, 2, \dots, NP \quad (3.30)$$

viene generato un vettore di mutazione del tipo

$$v_{i,G+1} = x_{r1,G} + F(x_{r2,G} - x_{r3,G}) \quad (3.31)$$

con indici casuali

$$r_1, r_2, r_3 \in \{1, 2, \dots, NP\} \quad (3.32)$$

interi, mutualmente differenti e con $F > 0$.

La scelta casuale degli indici interi r è inoltre effettuata in modo che siano differenti anche dall'indice i , conseguentemente NP deve essere maggiore o uguale a 4.

F è una costante reale $\in [0, 2]$ che controlla l'amplificazione della variazione differenziale $(x_{r2,G} - x_{r3,G})$. Nella versione utilizzata in AMALGAM il valore di F è scelto in modo casuale nell'intervallo $[0.6, 1]$.

3.5.3 Crossover

Il *crossover* ha il compito di aumentare la diversità dei vettori dei parametri perturbati. A tal fine si assume il vettore di processo pari a

$$u_{i,G+1} = u_{1i,G+1} + u_{2i,G+1} + \dots + u_{Di,G+1} \quad (3.33)$$

dove

$$u_{ji,G+1} = v_{ji,G+1} \quad (3.34)$$

se

$$randb(j) \leq CR \quad (3.35)$$

o

$$j = rnbr(i) \quad (3.36)$$

altrimenti

$$u_{ji,G+1} = x_{ji,G} \quad (3.37)$$

se

$$randb(j) > CR \quad (3.38)$$

e

$$j \neq rnbr(i) \tag{3.39}$$

con $j = 1, 2, \dots, D$.

Nelle equazioni precedenti $randb(j)$ rappresenta la j -esima valutazione di un generatore di numeri casuali uniformemente distribuiti nello spazio $\in [0,1]$. CR è la costante di crossover $\in [0,1]$ che deve essere determinata dall'utilizzatore mentre $rnbr(i)$ è un indice scelto in modo casuale $\in [1, 2, \dots, D]$ che assicura che $u_{i,G+1}$ prenda al minimo un parametro da $v_{i,G+1}$.

3.5.4 Selezione

Per decidere se il vettore di processo $u_{i,G+1}$ deve diventare un membro della generazione $G+1$, esso viene paragonato al vettore di destinazione $x_{i,G}$. In pratica si comparano i valori delle rispettive funzioni obiettivo e in base ad esse si decide per la sostituzione o la conservazione del valore precedente.

4 Area di studio e Modello Afflussi-Deflussi

4.1 Inquadramento geografico

I risultati riportati nella presente trattazione sono legati a due tipologie di casi di studio: applicazione al caso di bacini strumentati e applicazione al caso di bacini non strumentati. In particolare in questo secondo caso è stato possibile procedere con la trattazione grazie alla procedura di regionalizzazione precedentemente realizzata da *Castiglioni (2010)* durante il Dottorato di ricerca presso il DICAM di Bologna.

Le analisi si riferiscono in entrambi i casi ad alcuni bacini ricadenti in un'area dell'Italia centrale identificabile con le Regioni Marche ed Abruzzo (Figura 4.1). Tale area si espande per una superficie complessiva di circa 18257 km² ed è delimitata a nord dal bacino del Fiume Foglia, a sud dal bacino del Fiume Sangro, ad ovest dalla dorsale appenninica e ad est dal Mar Adriatico.

La regione rappresenta un complesso insieme di bacini assai disomogeneo rispetto alle caratteristiche climatiche e geomorfologiche. E' caratterizzata perlopiù da colline e montagne con brevi tratti pianeggianti in prossimità della costa. I corsi d'acqua hanno limitati bacini imbriferi, corso breve e rapido, andamento parallelo con direzione principale da sud-ovest verso nord-est. I bacini presentano un regime idrologico irregolare, a carattere prevalentemente torrentizio e contraddistinto da forti magre durante il periodo estivo.

L'area in oggetto coincide con i compartimenti che erano di competenza delle Sezioni di Bologna e Pescara del Servizio Idrografico e Mareografico Nazionale (S.I.M.N.). Per i bacini di quest'area sono disponibili osservazioni idrometriche per almeno 4 anni di dati, oltre a numerosi dati relativi ai parametri di sintesi delle caratteristiche geomorfologiche e climatiche (Pubblicazione n° 17 del S.I.M.N.). La regione presenta sul territorio un numero considerevole di stazioni di misura delle precipitazioni e temperatura le quali hanno reso possibile, per ben 13 bacini (evidenziati in azzurro in Figura 4.1), l'individuazione e la determinazione di serie di portate, precipitazioni ed evapotraspirazione a scala giornaliera riferite al medesimo intervallo temporale.

Nei seguenti paragrafi si riporta una descrizione delle informazioni disponibili in formato elettronico per i bacini appartenenti a quest'area e che sono state utilizzate da Castiglioni (2010) per la procedura di regionalizzazione richiamata in seguito (si veda il Paragrafo 6.1.1).

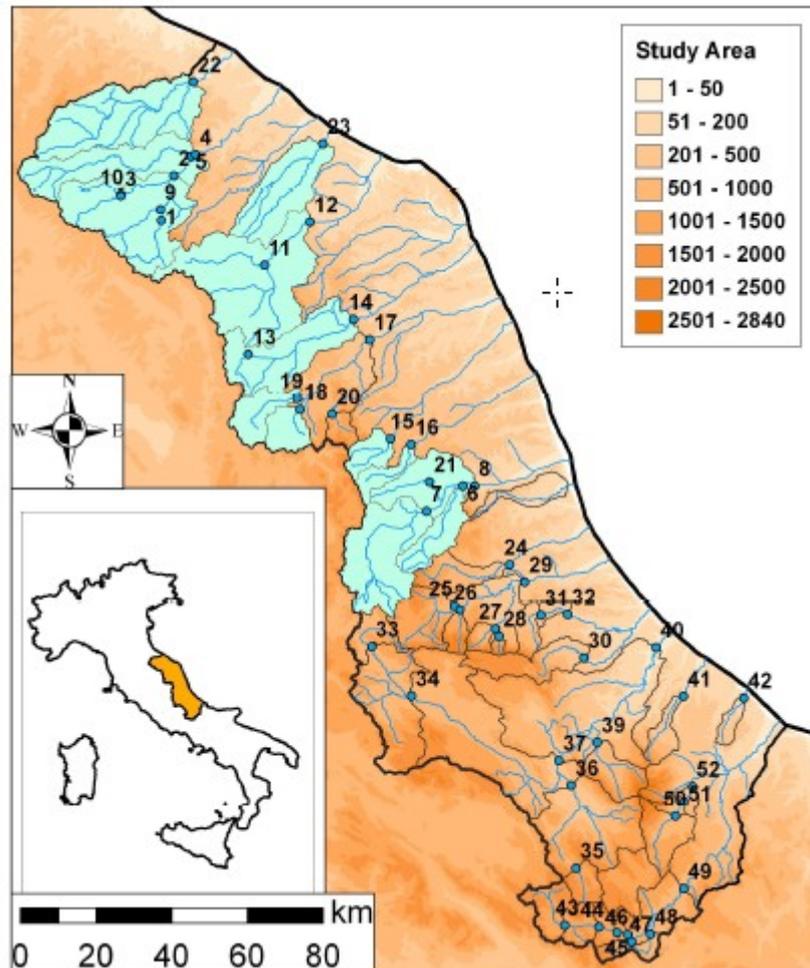


Figura 4.1: Area di studio: bacini idrografici e relative stazioni idrometriche. In azzurro sono evidenziati i bacini per i quali sono disponibili serie di portate, piogge, ed evapotraspirazione riferite al medesimo intervallo temporale.

4.2 Informazioni Idrometriche

Per ciascun bacino appartenente all'area appena descritta sono disponibili su supporto informatico i dati di osservazioni di portata media giornaliera (Annali Idrologici degli Uffici di Bologna e Pescara del S.I.M.N.). Le serie delle portate si riferiscono ad un intervallo temporale di durata minima di 4 anni fino ad un massimo di 67 anni di rilevazioni. Per ogni sezione di misura idrometrica è stata esclusa la presenza di condizioni particolari che potrebbero influenzare in maniera significativa il regime dei

deflussi, come ad esempio la presenza di invasi artificiali o derivazioni a monte del punto di misura.

<i>N.</i>	<i>Cod.</i>	<i>Corso d'acqua</i>	<i>Località</i>	<i>N. Oss.</i>	Q_{355} [L/s]	$Q_{95\%}$ [L/s]	$7Q10$ [L/s]	μ [m ³ /s]	σ^2	ρ_1
1	801	Burano	Foci	7	350	360	272	2.76	12.55	0.72
2	901	Candigliano	Acqualagna	22	940	910	517	13.88	477.39	0.71
3	902	Candigliano	Piobbico	8	80	70	44	4.45	56.03	0.64
4	1002	Metauro	Barco di Bellag.	26	980	1000	659	21.29	1082.1	0.68
5	1004	Metauro	Calmazzo	10	110	140	56	6.91	122.32	0.65
6	1101	Castellano	Ascoli Piceno	9	1080	780	527	3.49	21.09	0.69
7	1202	Tronto	Ponte d'Arli	26	2310	2080	1444	8.74	77.62	0.72
8	1204	Tronto	Tolignano di M.	40	3690	3270	2517	16.77	375.14	0.69
9	1701	Bosso	Cagli	10	330	330	254	3.20	26.06	0.67
10	2101	Biscuvio	Piobbico	5	40	40	35	2.54	22.32	0.64
11	2201	Sentino	S. Vittore	12	440	400	269	6.86	105.56	0.73
12	2502	Esino	Moie	11	3230	2610	1372	14.39	294.49	0.72
13	2601	Potenza	Spindoli	10	620	600	473	2.60	5.11	0.80
14	2602	Potenza	Cannucciaro	40	2550	2350	1859	7.75	46.30	0.80
15	2801	Tenna	Amandola	39	1000	930	722	2.72	4.66	0.77
16	2901	Aso	Comunanza	21	1260	1190	956	2.86	4.58	0.65
17	3002	Chienti	Tolentino	5	3060	2830	2210	13.06	126.23	0.80
18	3003	Chienti	Pieve Torina	29	580	480	383	2.10	3.27	0.85
19	3006	Chienti	Ponte Giove	26	380	230	172	1.49	1.80	0.82
20	3101	Fiastrone	Fiume	6	540	540	470	1.64	1.40	0.80
21	3301	Fluvione	Ponte Pugliese	5	220	170	106	1.67	4.56	0.71
22	3701	Foglia	Montecchio	38	60	30	11	6.65	223.34	0.65
23	4701	Misa	Vallone	9	10	0	0	2.21	33.79	0.74
24	5010	Vibrata	Alba Adriatica	12	0	0	0	0.57	1.71	0.74
25	5210	Tordino	Teramo	44	260	240	143	1.80	12.30	0.62
26	5330	Vomano	Fano Adriano	21	660	620	383	4.79	42.92	0.67
27	5360	Rio Arno	Ponte Rio Arno	67	170	140	74	1.00	3.70	0.75
28	5380	Mavone	Isola del G. Sasso	18	480	490	352	1.81	3.63	0.71
29	5390	Ruzzo	Pretara	16	660	660	503	1.23	0.50	0.74
30	5410	Vomano	Ponte Vomano	9	3090	2960	1903	16.11	377.64	0.68
31	5640	Tavo	S. Pellegrino	36	550	480	342	1.95	6.74	0.70
32	5660	Fino	Bisenti	11	40	40	18	1.00	3.56	0.71
33	5670	Fino	Castiglione M. R.	6	60	30	8	1.32	5.99	0.60
34	5710	Aterno	Treporti	36	80	70	35	1.05	2.87	0.69
35	5730	Aterno	L'Aquila	24	960	820	567	3.86	51.48	0.69
36	5800	Tasso	Scanno	18	220	150	87	0.69	0.35	0.79
37	5870	Sagittario	Capo Canale	53	2630	2220	1467	6.85	9.84	0.76
38	5920	Pescara	Maraone	59	17700	16000	13908	26.62	75.98	0.79
39	5990	Orte	Bolognano	14	320	320	188	3.53	24.80	0.67
40	6120	Pescara	S. Teresa	59	32840	29100	25579	49.21	385.50	0.73
41	6320	Foro	Ponte di Vacri	14	140	80	28	0.57	1.48	0.72
42	6560	Feltrino	S. Vito	37	40	30	11	0.35	1.49	0.63
43	6610	Sangro	Opi	27	20	0	2	0.70	2.67	0.68
44	6620	Sangro	Villetta Barrea	29	1120	1060	763	3.21	10.00	0.79
45	6630	Sangro	Barrea	16	2090	2080	1666	5.47	22.31	0.77
46	6650	Sangro	Alfedena	7	1890	1940	1656	5.12	19.24	0.78
47	6660	Rio Torto	Alfedena	9	130	120	71	0.84	0.83	0.76
48	6670	Zittola	Montenero V. S.	62	100	90	60	0.90	1.93	0.82
49	6690	Sangro	Ateleta	62	2250	2040	1468	8.88	89.86	0.68
50	6750	Aventino	Lama dei Peligni	5	2290	2170	1659	5.06	12.44	0.80
51	6760	Aventino	Vicenne	24	1710	1670	1162	4.34	22.68	0.75
52	6770	Aventino	Casoli	13	2150	1910	1254	5.61	36.62	0.72

Tabella 4.1: Stazioni Idrometriche interessate dallo studio e indici caratteristici delle portate di magra (Q_{355} , $Q_{95\%}$ e $7Q10$), media di lungo periodo (μ), varianza (σ^2) e coefficiente di correlazione lineare a ritardo un giorno (ρ_1) relative alle serie di portata giornaliera disponibili.

Al fine della presente trattazione è importante richiamare anche il fatto che nello studio di Castiglioni (*Castiglioni et al., 2010a*), a partire dalle serie idrometriche, è stato possibile determinare tre diversi indici di magra oltre ad alcune caratteristiche statistiche delle serie dei deflussi fluviali per ciascun bacino. La Tabella 4.1 ripropone tali risultati.

Nel presente lavoro si è fatto riferimento in particolare alla portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (in breve Q_{355}) e ad alcune statistiche dei deflussi quali

- la media (μ),
- la varianza (σ^2) e
- il coefficiente di autocorrelazione a ritardo unitario (ρ_1).

4.3 Informazioni Geomorfologiche

Per la caratterizzazione geomorfologica dei bacini si sono presi in considerazione 8 parametri di sintesi delle caratteristiche geomorfologiche dei bacini in esame (Tabella 4.2).

In particolare la superficie A [km²] e la lunghezza L [km] dell'asta principale sono state valutate a partire dalla cartografia IGM con scala 1:100.000. Le caratteristiche relative all'altezza massima H_{max} , media H_{med} e minima H_{min} [m s.l.m.] del bacino idrografico e la permeabilità del suolo P [%] sono state tratte dalla Pubblicazione n°17 del S.I.M.N. (Min. LLPP., annate varie). Da tali informazioni sono state determinate la quota media del bacino rispetto allo zero idrometrico W_H [m], definito come $W_H = H_{med} - H_{min}$, e il tempo di corrivazione del bacino τ_c [ore], calcolato mediante la formula di Giandotti secondo l'espressione

$$t_c = \frac{4\sqrt{A} + 1,5L}{0,8\sqrt{\Delta H}} \quad (4.1)$$

con significato dei simboli già esplicitato.

A completamento delle informazioni geomorfologiche del bacino in esame si è tenuto conto dell'aspetto litologico dell'area grazie ad una carta, in formato digitale, contenente le principali unità litologiche delle Regioni Marche ed Abruzzo che riporta la suddivisione dell'area in classi a diverse permeabilità e porosità (Figura 4.2).

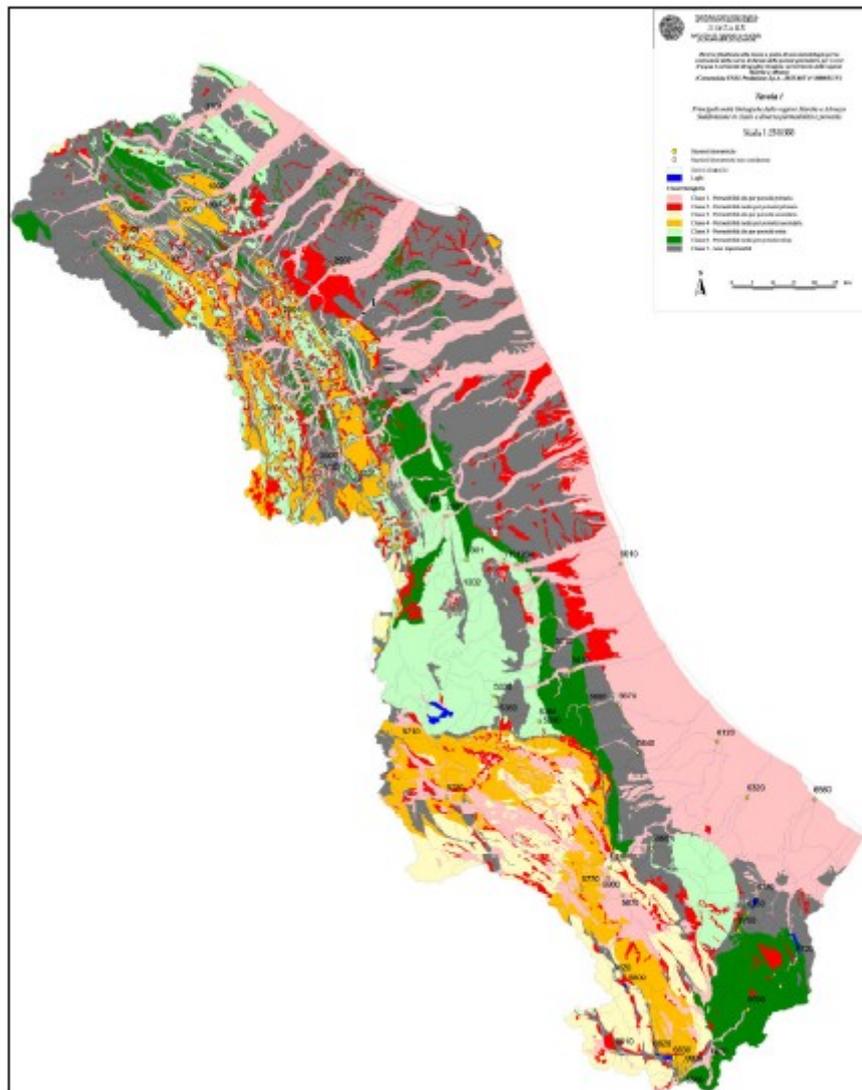


Figura 4.2: Principali unità litologiche delle regioni Marche ed Abruzzo. Suddivisione in classi a diverse permeabilità e porosità (fonte DICAM, Facoltà di Ingegneria di Bologna).

4.4 Informazioni Climatiche

Le informazioni climatiche sono state desunte dai dati pluviometrici e termometrici riferiti alle stazioni di misura distribuite sul territorio (Figura 4.2 e Figura 4.3). In particolare sono disponibili su supporto informatico le precipitazioni totali mensili [mm] per tutte le stazioni pluviometriche-pluviometrografiche ubicate nelle Regioni Marche ed Abruzzo, nel territorio compreso tra i bacini del Fiume Foglia a nord e del Fiume Sangro a sud (Annali Idrologici degli Uffici di Bologna e Pescara del S.I.M.N.). L'insieme di tutte le stazioni pluviometriche disponibili è composto da 337 stazioni pluviometriche, 187 ubicate nella regione Marche e le restanti 150 nella regione Abruzzo e sono caratterizzate da almeno 10 anni di osservazione (Figura 4.3). Per 44 stazioni collocate nella regione Marche è stato possibile reperire dal Dipartimento per

le Politiche Integrate di Sicurezza e per la Protezione Civile di Ancona, le serie di osservazioni giornaliere dal 1951 al 2006.

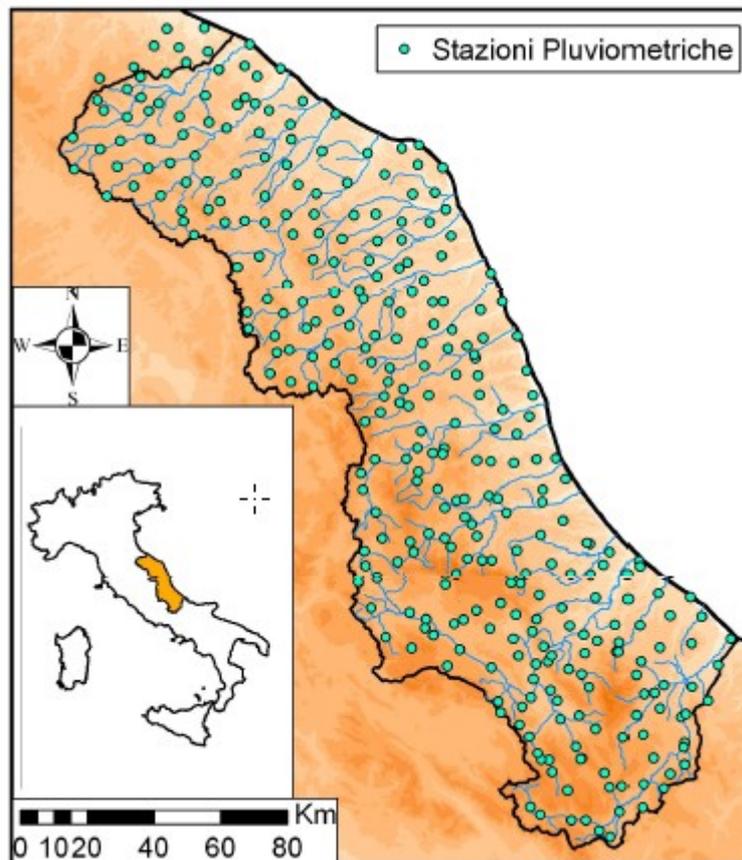


Figura 4.3: Sensori pluviometrici disponibili nell'area di studio.

Analogamente per le osservazioni di temperatura si è potuto disporre di dati riferiti alle temperature medie mensili [C°] per tutte le stazioni termometriche-termometrografiche ubicate nelle regioni Marche ed Abruzzo (Annali Idrologici degli Uffici di Bologna e Pescara del S.I.M.N.). Tali grandezze sono determinate per la generica stazione, secondo la definizione del S.I.M.N., come la semisomma della media delle temperature massime giornaliere e della media delle temperature minime giornaliere osservate alla stazione stessa nel mese considerato. Le stazioni termometriche disponibili sono in tutto 88 delle quali, 28 ubicate nella regione Marche e 60 nella regione Abruzzo. Per 16 di esse, collocate tutte nella regione Marche, è stato possibile reperire dal Dipartimento per le Politiche Integrate di Sicurezza e per la Protezione Civile di Ancona, le serie di temperatura medie giornaliere dal 1951 al 1979.

A partire dalle informazioni puntuali di precipitazione e temperatura, sono stati ottenuti i valori a scala di bacino della precipitazione *MAP* [mm/anno] e della temperatura *TAM* [C°], entrambi medi annuali, ricorrendo al metodo della poligonazione di Thiessen (Tabella 4.3).

<i>N.</i>	<i>Cod.</i>	<i>Corso d'acqua</i>	<i>Località</i>	<i>A</i> [km ²]	<i>P</i> [%]	<i>H_{max}</i> [m slm]	<i>H_{med}</i> [m slm]	<i>H_{min}</i> [m slm]	ΔH [m]	<i>L</i> [km]	τ_c [ore]
1	801	Burano	Foci	124.1	35.0	1702	655	285	370	16.9	4.5
2	901	Candigliano	Acqualagna	613.8	26.0	1702	600	183	417	56.3	11.2
3	902	Candigliano	Piobbico	186.7	10.0	1526	624	329	295	33.6	7.6
4	1002	Metauro	Barco di Bellag.	1043.6	20.0	1702	560	110	450	75.7	14.3
5	1004	Metauro	Calmazzo	375.9	6.0	1384	507	115	392	74.3	11.9
6	1101	Castellano	Ascoli Piceno	168.4	19.0	2455	1035	131	904	37.4	4.5
7	1202	Tronto	Ponte d'Arli	470.2	21.0	2478	1135	268	867	55.8	7.2
8	1204	Tronto	Tolignano di M.	900.5	16.0	2478	959	90	869	77.2	10.0
9	1701	Bosso	Cagli	126.1	36.0	1526	677	250	427	26.6	5.1
10	2101	Biscuvio	Piobbico	95.2	15.0	1526	633	334	299	19.0	4.9
11	2201	Sentino	S. Vittore	263.6	57.0	1702	571	191	380	48.8	8.9
12	2502	Esino	Moie	798.2	47.5	1702	529	96	433	69.9	13.1
13	2601	Potenza	Spindoli	88.2	82.0	1570	812	466	346	15.4	4.1
14	2602	Potenza	Cannucciaro	430.7	57.0	1570	616	168	448	58.4	10.1
15	2801	Tenna	Amandola	99.7	71.0	2334	1170	425	745	19.4	3.2
16	2901	Aso	Comunanza	83.5	43.0	2478	1100	437	663	26.0	3.7
17	3002	Chienti	Tolentino	691.0	66.0	2103	780	186	594	59.4	10.0
18	3003	Chienti	Pieve Torina	115.7	97.0	1575	924	445	479	17.1	3.9
19	3006	Chienti	Ponte Giove	141.4	99.0	1434	887	429	458	26.7	5.1
20	3101	Fiastrone	Fiume	60.9	90.0	2103	1268	621	647	12.9	2.5
21	3301	Fluvione	Ponte Pugliese	98.4	4.0	2478	778	269	509	19.0	3.8
22	3701	Foglia	Montecchio	605.6	0.1	1415	384	39	345	80.0	14.7
23	4701	Misa	Vallone	360.4	8.5	825	220	12	208	39.4	11.7
24	5010	Vibrata	Alba Adriatica	114.5	46.0	1815	214	4	210	36.0	8.4
25	5210	Tordino	Teramo	147.7	5.0	2435	930	222	708	35.1	4.8
26	5330	Vomano	Fano Adriano	216.4	15.0	2914	1530	505	1025	24.3	3.7
27	5360	Rio Arno	Ponte Rio Arno	57.5	52.0	2914	1950	407	1543	11.9	1.5
28	5380	Mavone	Isola del G. Sasso	44.1	77.0	2914	1200	400	800	10.0	1.8
29	5390	Ruzzo	Pretara	14.4	90.0	2566	1500	476	1024	5.3	0.9
30	5410	Vomano	Ponte Vomano	570.1	25.0	2914	1065	136	929	51.8	7.1
31	5640	Tavo	S. Pellegrino	210.8	66.0	2570	1200	179	1021	23.4	3.6
32	5660	Fino	Bisenti	74.3	17.0	2570	631	249	382	13.3	3.5
33	5670	Fino	Castiglione M. R.	108.1	22.0	2570	550	149	401	24.1	4.9
34	5710	Aterno	Treporti	112.6	34.0	1616	1026	760	266	20.8	5.6
35	5730	Aterno	L'Aquila	489.3	46.0	2206	1105	640	465	41.2	9.0
36	5800	Tasso	Scanno	80.3	97.0	2186	1550	926	624	10.0	2.5
37	5870	Sagittario	Capo Canale	603.7	93.0	2795	1105	269	836	26.1	5.9
38	5920	Pescara	Maraone	1961.3	70.0	2795	1080	240	840	103.5	14.4
39	5990	Orte	Bolognano	156.7	43.0	2795	1077	198	879	18.8	3.3
40	6120	Pescara	S. Teresa	3082.0	58.0	2795	940	5	935	159.9	18.9
41	6320	Foro	Ponte di Vacri	88.5	50.0	1955	620	95	525	23.7	4.0
42	6560	Feltrino	S. Vito	49.4	5.0	460	178	3	175	15.6	4.9
43	6610	Sangro	Opi	132.5	91.0	2208	1405	1103	302	15.8	5.0
44	6620	Sangro	Villetta Barrea	219.2	90.0	2249	1515	990	525	25.5	5.3
45	6630	Sangro	Barrea	271.7	90.0	2283	1515	944	571	30.7	5.9
46	6650	Sangro	Alfedena	289.2	89.0	2008	1500	876	624	33.9	6.0
47	6660	Rio Torto	Alfedena	31.7	83.0	2241	1040	882	158	9.6	3.7
48	6670	Zittola	Montenero V. S.	31.6	81.0	1260	1080	816	264	10.6	3.0
49	6690	Sangro	Ateleta	561.9	65.0	2795	1320	720	600	61.2	9.5
50	6750	Aventino	Lama dei Peligni	167.4	44.0	280	1350	350	1000	19.8	3.2
51	6760	Aventino	Vicenne	198.7	40.0	2795	1191	271	920	24.6	3.8
52	6770	Aventino	Casoli	229.3	41.0	2795	1090	225	865	30.1	4.5

Tabella 4.2. Stazioni idrometriche interessate dallo studio e relativi dati geomorfologici caratteristici dei bacini idrografici da esse sottese.

A completamento delle informazioni climatiche erano infine disponibili i valori medi mensili di evapotraspirazione potenziale calcolati attraverso la relazione empirica di Thornthwaite (v. ad es. *Moisello, 1999*) a scala di bacino per tutti i bacini appartenenti

all'area presa in esame da Castiglioni (2010). Da questi dati, attraverso opportune trasformazioni, è stato possibile determinare i valori medi giornalieri.

<i>N.</i>	<i>Cod.</i>	<i>Corso d'acqua</i>	<i>Località</i>	<i>MAP</i> [mm]	<i>TAM</i> [C°]
1	801	Burano	Foci	1223	12.0
2	901	Candigliano	Acqualagna	1118	12.1
3	902	Candigliano	Piobbico	1174	12.6
4	1002	Metauro	Barco di Bellag.	1115	12.5
5	1004	Metauro	Calmazzo	1159	12.8
6	1101	Castellano	Ascoli Piceno	932	13.6
7	1202	Tronto	Ponte d'Arli	1006	10.9
8	1204	Tronto	Tolignano di M.	988	11.8
9	1701	Bosso	Cagli	1251	12.2
10	2101	Biscuvio	Piobbico	1129	13.0
11	2201	Sentino	S. Vittore	1185	12.5
12	2502	Esino	Moie	1228	12.4
13	2601	Potenza	Spindoli	1038	12.1
14	2602	Potenza	Cannuciaro	1097	12.2
15	2801	Tenna	Amandola	900	11.1
16	2901	Aso	Comunanza	1164	11.5
17	3002	Chienti	Tolentino	1219	11.5
18	3003	Chienti	Pieve Torina	1223	13.0
19	3006	Chienti	Ponte Giove	958	13.1
20	3101	Fiastrone	Fiume	1144	11.4
21	3301	Fluvione	Ponte Pugliese	1189	12.0
22	3701	Foglia	Montecchio	949	12.3
23	4701	Misa	Vallone	886	12.6
24	5010	Vibrata	Alba Adriatica	824	15.3
25	5210	Tordino	Teramo	987	11.8
26	5330	Vomano	Fano Adriano	1139	10.1
27	5360	Rio Arno	Ponte Rio Arno	1047	9.7
28	5380	Mavone	Isola del G. Sasso	1363	11.8
29	5390	Ruzzo	Pretara	1336	12.9
30	5410	Vomano	Ponte Vomano	1169	9.8
31	5640	Tavo	S. Pellegrino	986	9.5
32	5660	Fino	Bisenti	1097	9.9
33	5670	Fino	Castiglione M. R.	1098	12.0
34	5710	Aterno	Treporti	997	11.0
35	5730	Aterno	L'Aquila	1011	10.0
36	5800	Tasso	Scanno	1046	9.7
37	5870	Sagittario	Capo Canale	888	11.8
38	5920	Pescara	Maraone	874	11.2
39	5990	Orte	Bolognano	1419	13.1
40	6120	Pescara	S. Teresa	870	12.1
41	6320	Foro	Ponte di Vaci	974	12.0
42	6560	Feltrino	S. Vito	820	14.5
43	6610	Sangro	Opi	1530	8.3
44	6620	Sangro	Villetta Barrea	1273	8.6
45	6630	Sangro	Barrea	1516	10.0
46	6650	Sangro	Alfedena	1336	12.4
47	6660	Rio Torto	Alfedena	1219	14.7
48	6670	Zittola	Montenero V. S.	1231	10.6
49	6690	Sangro	Ateleta	1243	10.2
50	6750	Aventino	Lama dei Peligni	822	10.2
51	6760	Aventino	Vicenne	867	11.2
52	6770	Aventino	Casoli	893	11.4

Tabella 4.3. Stazioni idrometriche interessate dallo studio e relativi dati climatici caratteristici dei bacini idrografici da esse sottese.

4.5 Il Modello Afflussi-Deflussi Hymod

L'algoritmo *multi-obiettivo* AMALGAM può essere utilizzato in linea di principio in applicazione a qualsiasi modello afflussi-deflussi. Tuttavia, dovendo effettuare un'analisi anche nell'ipotesi di bacini non strumentati, si è preferito diminuire l'incertezza associata ai parametri utilizzando un modello parsimonioso. Si è mantenuta quindi la scelta effettuata dall'autore dell'algoritmo in fase di verifica del comportamento dello stesso ad un caso test, confermando l'utilizzo di HYMOD, un modello concettuale a 5 parametri concentrati proposto da Boyle (2000) e recentemente usato anche da molti altri autori (si veda esempio Wagener, 2001 - Montanari e Toth, 2007).

Il concetto generale alla base del modello HYMOD si basa su quanto proposto da R.J. Moore (1985) ed è fondato sull'analisi della distribuzione di probabilità del processo di produzione di deflusso in un punto all'interno del bacino. Al fine di ricavare delle espressioni algebriche che integrino la risposta del bacino, vengono utilizzate delle distribuzioni di probabilità che permettono di descrivere la variazione spaziale dei parametri di processo sul bacino medesimo.

In breve possiamo riassumere le caratteristiche di tale modello (Figura 4.4) spiegando che si basa sulla propagazione del deflusso tramite due serie di serbatoi lineari. La simulazione del deflusso veloce viene effettuata tramite tre serbatoi identici posti in serie, mentre la propagazione del deflusso lento viene realizzata mediante un solo serbatoio di caratteristiche ovviamente differenti dai primi.

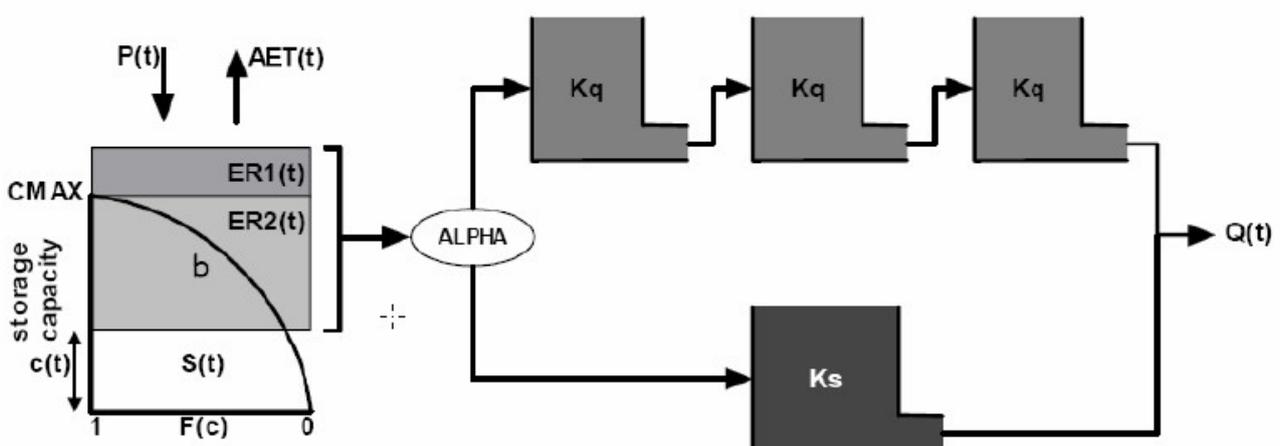


Figura 4.4: Schema di funzionamento del modello Hymod.

I parametri del modello da calibrare sono 5 ossia:

- la capacità massima di immagazzinamento del terreno, C_{MAX} [mm];

- il grado di variabilità spaziale della capacità di immagazzinamento sul bacino, b [-];
- il fattore di ripartizione del deflusso tra le due serie di serbatoi, $ALPHA$ [-];
- costante di proporzionalità fra portata uscente dallo scarico di fondo e volume invasato (rispettivamente nelle due tipologie di serbatoi) K_q e K_s [-].

Per arrivare a questa schematizzazione, l'autore è partito dal considerare il bacino come un insieme di porzioni di area indipendenti tra loro. Ciascuna porzione sarà caratterizzata, oltre che da un certo grado di precipitazione e di evapotraspirazione legati all'intervallo di tempo, da una capacità propria di *stoccaggio* dell'acqua (denominata appunto C) che varierà nello spazio in funzione della tessitura e della struttura del terreno. La funzione di distribuzione di tale capacità di stoccaggio dell'acqua è definita come:

$$F(C) = 1 - \left(1 - \frac{C}{C_{\max}}\right)^b \quad (4.2)$$

con

$$0 \leq C \leq C_{\max} \quad (4.3)$$

dove i simboli assumono il significato già esplicitato in precedenza e con F che rappresenta la probabilità cumulata della capacità di stoccaggio dell'acqua.

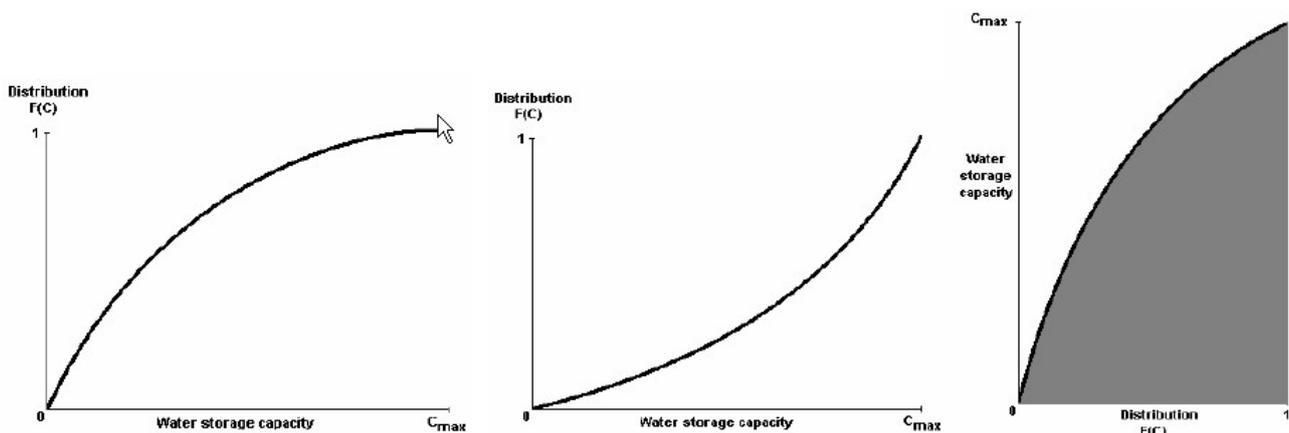


Figura 4.5. Andamento della funzione di distribuzione della capacità di immagazzinamento del bacino in funzione del valore di b .

In funzione del valore di b , l'andamento della funzione di distribuzione può apparire come nella Figura 4.5: in particolare l'immagine di sinistra si riferisce al valore di $b=2$

mentre quella centrale al valore di $b=0.5$, mentre valori intermedi di b produrranno andamenti intermedi della curva.

L'immagine di destra evidenzia in grigio la capacità di immagazzinamento complessiva del bacino (si osservi che gli assi sono stati invertiti rispetto alle due immagini precedenti).

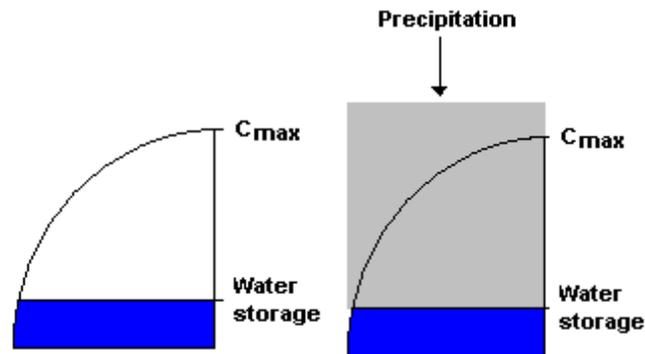


Figura 4.6. Situazione iniziale del bacino e raggiungimento della massima capacità di immagazzinamento.

Nella situazione iniziale il bacino contiene una quantità generica di acqua (Figura 4.6 - immagine di sinistra); in presenza di precipitazione tale quantità aumenta (Figura 4.6 - immagine di destra) fino a diventare superiore al valore massimo immagazzinabile. Il volume d'acqua superiore a C_{max} (indicato in rosso nella Figura 4.7) viene direttamente destinato al deflusso superficiale poiché il terreno sottostante è saturo e non può ricevere ulteriori quantitativi che saranno quindi smaltiti in base alle caratteristiche dei tre serbatoi di superficie.

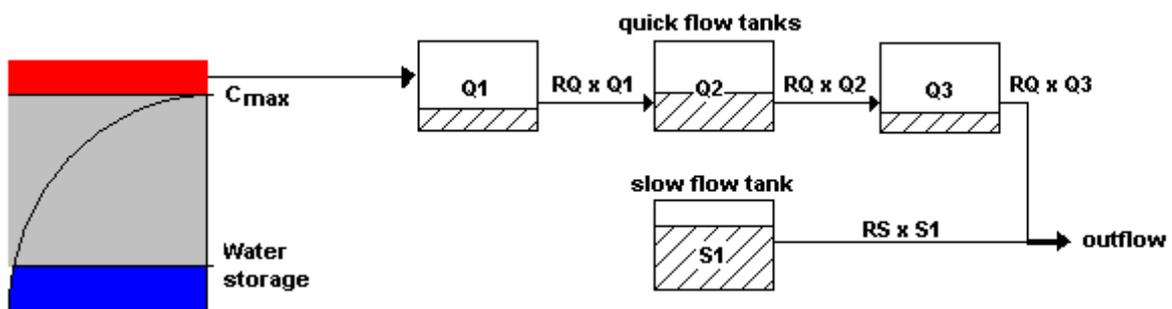


Figura 4.7. Porzione di precipitazione che viene smaltita esclusivamente come deflusso superficiale.

La parte di precipitazione rimanente che comunque eccede la capacità di immagazzinamento del bacino (indicata in rosso nella Figura 4.8) viene divisa tra il deflusso lento e quello veloce in funzione del valore del parametro $ALPHA$. La quantità ancora presente, a meno di una parte destinata all'evapotraspirazione, rappresenta il volume di partenza per ripetere il procedimento allo step temporale successivo.

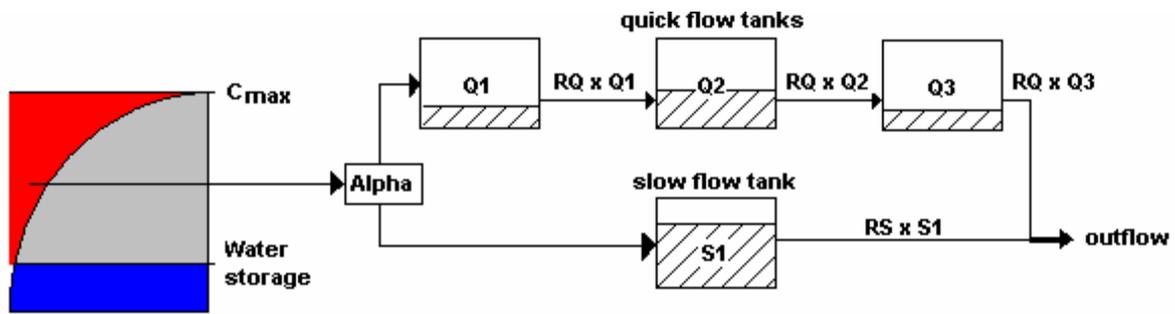


Figura 4.8. Porzione di precipitazione suddivisa tra deflusso profondo e deflusso superficiale.

Il deflusso complessivo sarà dato dalla quantità di acqua che fuoriesce dalla serie dei tre serbatoi di superficie sommata a quella proveniente dal serbatoio relativo al deflusso profondo.

Come si evince da quanto spiegato fino ad ora i dati di input necessari per utilizzare il modello sono le serie corrispondenti di precipitazioni ed evapotraspirazione uniti alla dimensione dell'area del bacino idrografico.

5 La calibrazione dei modelli afflussi-deflussi in bacini strumentati

La prima parte della ricerca è stata rivolta all'applicazione dell'algoritmo di ottimizzazione AMALGAM ai bacini dell'area di studio considerati come strumentati.

Questa applicazione ha avuto i seguenti obiettivi:

1. ottenere i primi risultati applicativi sul funzionamento dell'algoritmo di ottimizzazione AMALGAM in riferimento a casi di studio reali,
2. effettuare un confronto tra le modalità di calibrazione *mono* e *multi-obiettivo*,
3. ricavare alcuni indici di qualità relativi al caso di calibrazione in bacini strumentati in modo da poter poi effettuare una stima del deterioramento dei risultati nel passaggio al caso di bacini non strumentati.

Nei seguenti paragrafi sono riportati i dettagli relativi a questo tipo di approccio.

5.1 Area di indagine

L'area di studio individuata per questa applicazione è costituita da 4 bacini ricadenti nella parte nord dell'area ben più vasta descritta in precedenza. Tale area è interamente compresa nella Regione Marche ed è costituita dal bacino del Metauro con sezione di chiusura a Barco di Bellaguardia, dal Fiume Candigliano ad Acqualagna, dal fiume Esino a Moie e dal Fiume Potenza a Cannuciaro (Figura 5.1).

In Tabella 5.1 sono riportati i periodi di tempo relativi ai dati disponibili per questi 4 bacini che consistono in particolare nelle portate fluviali medie giornaliere, nelle precipitazioni medie areali calcolate con il metodo dei Poligoni di Thiessen e nelle serie di evapotraspirazione giornaliere ricavate tramite l'eq. di Thornthwaite.

Nei giorni presi in esame per la calibrazione si è tenuto conto di un periodo di *warm-up* variabile da 65 a 283 giorni in funzione della disponibilità dei dati, necessario per

mandare a regime il funzionamento dei serbatoi lineari del modello afflussi-deflussi Hymod.

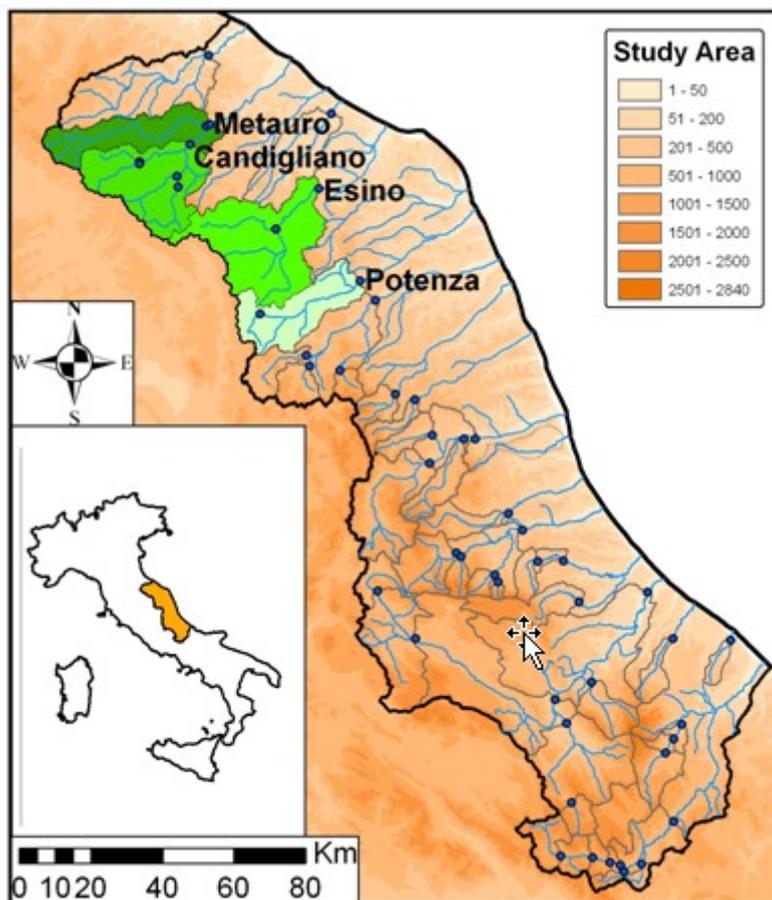


Figura 5.1. Area di studio per la trattazione del caso di bacini strumentati.

Fiume (codice sezione)	Dati disponibili (anni-giorni)	Calibrazione (giorni)	Validazione (giorni)
Candigliano (901)	dal 1970 al 1975 (2191)	1460	731
Metauro (1002)	dal 1951 al 1956 (2192)	1460	732
Esino (2502)	dal 1955 al 1958 (1461)	730	731
Potenza (2602)	dal 1967 al 1972 (2192)	1460	732

Tabella 5.1. Dati disponibili per l'area di studio presa in esame nella trattazione di bacini considerati strumentati.

5.2 Funzioni obiettivo

Per effettuare le indagini necessarie al raggiungimento degli obiettivi preposti sono state selezionate alcune funzioni obiettivo di comune utilizzo e si è operato considerandole singolarmente oppure in alcune combinazioni a coppie e a tre a tre.

Simbolo	Denominazione	Espressione matematica
BIAS	Scarto medio	$\frac{1}{n} \sum_{i=1}^n (Q_i^{sim} - Q_i^{oss})$
EFF	Efficienza di Nash	$1 - \frac{\sum_{i=1}^n (Q_i^{sim} - Q_i^{oss})^2}{\sum_{i=1}^n (Q_i^{oss} - \bar{Q})^2}$
MAE	Errore assoluto medio (Mean Absolute Error)	$\frac{1}{n} \left \sum_{i=1}^n Q_i^{sim} - Q_i^{oss} \right $
MRE	Errore medio relativo (Mean Relative Error)	$\frac{1}{n} \sum_{i=1}^n \left(\frac{ Q_i^{sim} - Q_i^{oss} }{Q_i^{oss}} \right)$
MSE	Errore quadratico medio (Mean Square Error)	$\frac{1}{n} \sum_{i=1}^n (Q_i^{sim} - Q_i^{oss})^2$
RMSE	Radice dell'errore quadratico medio (Root Mean Square Error)	$\sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^{sim} - Q_i^{oss})^2}$
MSLE	Errore quadratico medio logaritmico (Mean Square Logaritmico Error)	$\frac{1}{n} \sum_{i=1}^n (\log(Q_i^{sim}) - \log(Q_i^{oss}))^2$

Tabella 5.2. Funzioni obiettivo prese in esame nel caso di bacini considerati come strumentati e loro espressione matematica.

La scelta di tali funzioni è legata al fatto che ciascuna di esse permette di riprodurre in modo accurato un aspetto dell'idrogramma piuttosto di altri come già accennato nel paragrafo 1.3. Ad esempio, una buona riproduzione del volume medio

dell'idrogramma osservato, corrispondente ad un buon bilancio idrico, si ottiene utilizzando lo Scarto Medio (BIAS) mentre una buona riproduzione dell'andamento dell'idrogramma osservato è meglio ottenibile tramite l'utilizzo del Mean Absolute Error (MAE). In dettaglio inoltre, una buona riproduzione dei picchi di piena in riferimento al tempo, alla frequenza e al volume idrico coinvolto è meglio ottenibile tramite l'uso di indici quadratici come il Mean Square Error (MSE), il Root Mean Square Error (RMSE) o l'indice di Nash-Sutcliffe (EFF) mentre una buona riproduzione delle magre può essere ottenuta tramite l'uso del Mean Relative Error (MRE) o di indici logaritmici come il Mean Square Logarithmic Error (MSLE) (Lindstrom, 1997; Liong et al., 1996; Liong et al., 1998; Gupta et al., 1998; Yapo et al., 1998; Madsen, 2000).

Simbolo	Modalità
BIAS EFF MAE MRE RMSE	Mono obiettivo
RMSE (driven-no driven) MAE-MRE MAE-MSE MAE-MSLE MRE-MSE RMSE-BIAS	Multi obiettivo (2 funzioni)
MAE-MSE-MRE MAE-MSLE-MRE	Multi obiettivo (3 funzioni)

Tabella 5.3. Combinazioni di Funzioni Obiettivo prese in esame nel caso di bacini considerati come strumentati.

Nella presente trattazione la funzione RMSE è stata considerata in modalità *mono-obiettivo* e in modalità *multi-obiettivo* calcolandola separatamente per i giorni con eventi di pioggia (no driven) o per i giorni privi di tali eventi (driven).

In Tabella 5.2 sono riportate le espressioni matematiche delle funzioni obiettivo utilizzate mentre in Tabella 5.3 sono riportate le combinazioni di tali funzioni prese in esame.

L'analisi dei risultati è stata condotta attraverso il controllo della qualità visiva degli idrogrammi simulati, degli scatterplot di confronto tra portata simulata e osservata nonché dei valori dell'Efficienza di Nash-Sutcliffe, del Mean Absolute Error e del Mean Relative Error (secondo le formulazioni già riportate in Tabella 5.2) considerati questa volta come indici di qualità delle prestazioni del modello.

In particolare nei casi di calibrazione *multi-obiettivo*, per poter effettuare un confronto visivo diretto e un confronto degli indici di prestazione, si è presa in considerazione la portata *media* tra tutte quelle ottenute tramite i set della *Frontiera Pareto* (secondo quanto descritto nel paragrafo 1.3). La scelta della portata *media* è legata alla necessità di avere una valutazione intermedia tra tutte quelle ugualmente possibili e identificate dall'algoritmo di ottimizzazione in modalità *multi-obiettivo*.

5.3 Risultati e confronti

Le calibrazioni con l'algoritmo di ottimizzazione AMALGAM hanno permesso il raggiungimento dei risultati di seguito riportati. Nei seguenti paragrafi quindi si riportano gli indici di qualità ottenuti per tutti i bacini e con tutte le modalità di ottimizzazione prese in esame.

Al fine di rendere più semplice un eventuale confronto tra modalità *mono* e *multi-obiettivo* si è deciso di riportare i grafici (per ciascun bacino) relativi solo a quella modalità che si è dimostrata migliore per ciascuna tipologia.

5.3.1 Modalità mono-obiettivo

Nella Tabella 5.4 sono riportati i valori degli indici di qualità (indicati in lilla) ottenuti per il Fiume Candigliano in calibrazione con l'utilizzo delle funzioni obiettivo (indicate in grigio) ciascuna considerata in modalità *mono-obiettivo*.

I valori evidenziati (con sfondo azzurro) sono rispettivamente il valore massimo di EFF e minimi di MAE ed MRE che si sono ottenuti tra tutte le modalità *mono-obiettivo*. E' abbastanza evidente che ciascun indice di qualità è massimo o minimo quando esso stesso viene utilizzato come funzione obiettivo, per cui si preferisce riportare di

seguito i grafici ottenuti per la modalità *mono-obiettivo* che utilizza la funzione RMSE dimostratasi in realtà decisamente affidabile per la calibrazione di tutti i bacini.

CANDIGLIANO (calibrazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,24	0,74	0,68	0,68	0,74
MAE	6,11	3,33	3,01	3,04	3,33
MRE	1,92	0,61	0,42	0,41	0,61

Tabella 5.4. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

In Tabella 5.5 sono riportati gli stessi indici ottenuti però relativamente al periodo di validazione.

Nel dettaglio, la simulazione delle portate per il Fiume Candigliano è soddisfacente sia in calibrazione che in validazione (Figura 5.2 e Figura 5.3) a meno di pochissimi eventi di piena che vengono sottostimati dal modello. Sono presenti al contrario alcuni eventi con picchi poco significativi che vengono enfatizzati.

CANDIGLIANO (validazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,09	0,65	0,57	0,57	0,65
MAE	4,92	2,73	2,31	2,21	2,73
MRE	1,53	0,62	0,43	0,4	0,62

Tabella 5.5. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Gli scatterplot appaiono nel complesso simmetrici e poco dispersi (Figura 5.4).

La modalità peggiore per il Fiume Candigliano risulta quella ottenuta tramite l'utilizzo del BIAS per il quale l'EFF in calibrazione raggiunge un valore di 0.24.

Decisamente meno soddisfacenti, per quanto accettabili, le simulazioni ottenute per il Fiume Esino (Tabella 5.6 e Tabella 5.7). Anche in questo caso si è deciso di prendere come riferimento la simulazione ottenuta tramite l'utilizzo del RMSE, nonostante i

valori migliori degli indici - considerati singolarmente - si ottengano tramite la corrispondente funzione obiettivo, a causa delle motivazioni già citate in precedenza.

Fiume Candigliano - Calibrazione (RMSE)

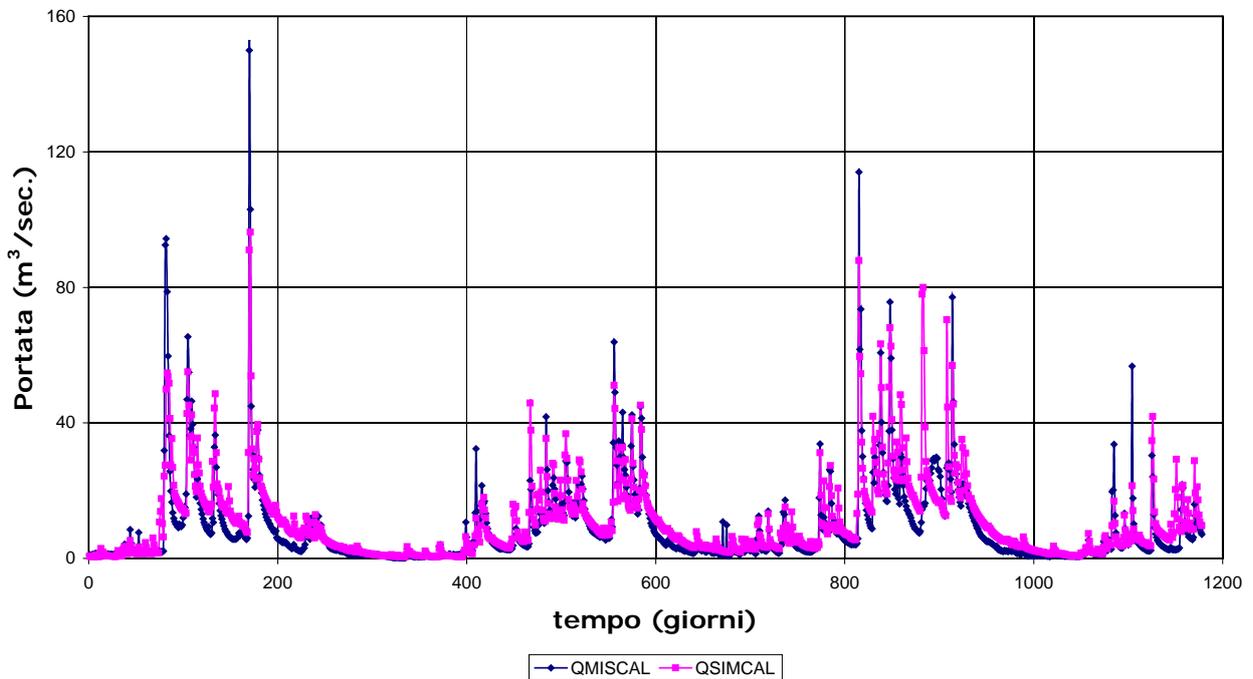


Figura 5.2. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (RMSE).

Fiume Candigliano - Validazione (RMSE)

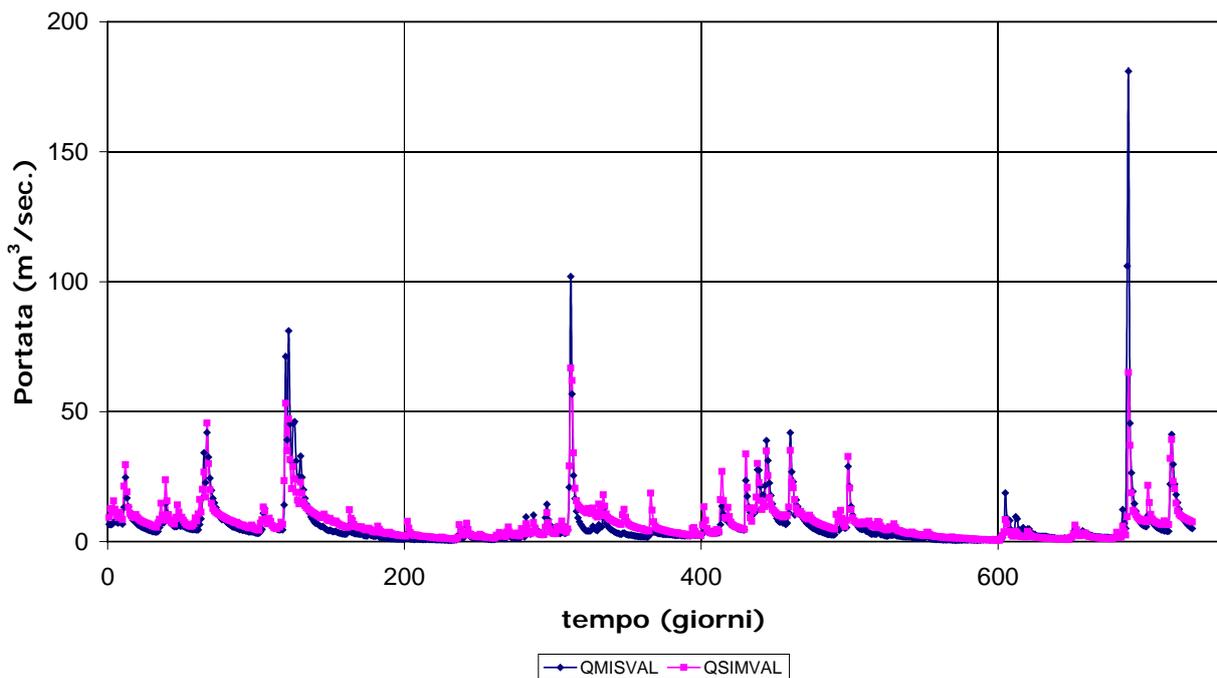


Figura 5.3. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (RMSE).

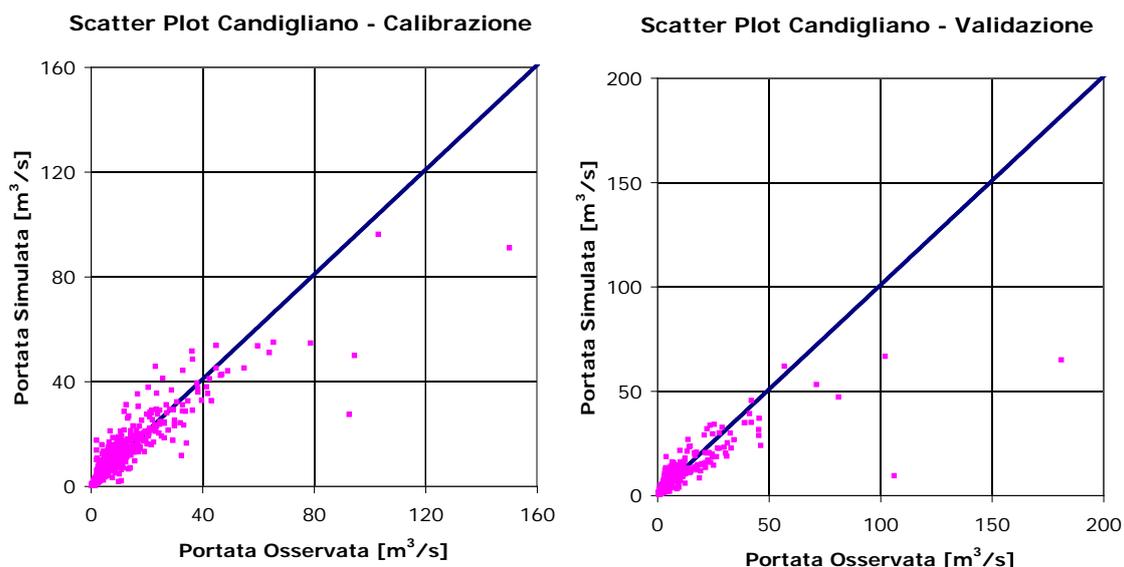


Figura 5.4. scatterplot in calibrazione e in validazione delle portate relative al Fiume Candigliano a Acqualagna (RMSE).

In particolare le simulazioni per tale bacino generano risultati migliori in validazione (Figura 5.6) piuttosto che in calibrazione (Figura 5.5) a causa della difficoltà a riprodurre gli eventi di piena del primo periodo. In validazione, sebbene non venga riprodotto il maggior picco, l'andamento delle portate osservate è più regolare e quindi meglio riproducibile dal modello. Conseguentemente lo scatterplot in validazione è meno disperso e più simmetrico di quello in calibrazione (Figura 5.7).

ESINO (calibrazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,48	0,58	0,56	0,08	0,48
MAE	7,78	6,98	6,54	8,28	6,98
MRE	0,59	0,51	0,44	0,35	0,51

Tabella 5.6. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Per il Fiume Esino la modalità di calibrazione peggiore sembra quella ottenuta tramite l'utilizzo del MRE per il quale l'EFF in calibrazione ha valore di 0.08.

Le simulazioni per il Fiume Metauro presentano valori dell'indice MAE molto elevati (Tabella 5.8 e Tabella 5.9). Osservando gli idrogrammi ottenuti con l'utilizzo del RMSE come riferimento (Figura 5.8 e Figura 5.9), il motivo di tali valori sembra apparentemente legato ad una inadeguata riproduzione dei fenomeni di piena. Questo

particolare è confermato da quanto si osserva anche nello scatterplot in calibrazione (Figura 5.10) dove risulta evidente una migliore riproduzione delle portate basse piuttosto che delle portate elevate. In validazione invece lo scatterplot è maggiormente disperso anche in riferimento ai valori di portata meno elevati (Figura 5.10).

ESINO (VALIDAZIONE)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,73	0,71	0,69	0,11	0,71
MAE	5,5	5,43	5,28	7,7	5,43
MRE	0,51	0,41	0,41	0,37	0,41

Tabella 5.7. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Per il Fiume Metauro la simulazione peggiore si ottiene con l'utilizzo del BIAS seppur anche in questo caso i valori degli indici non siano particolarmente scarsi.

Fiume Esino - Calibrazione (RMSE)

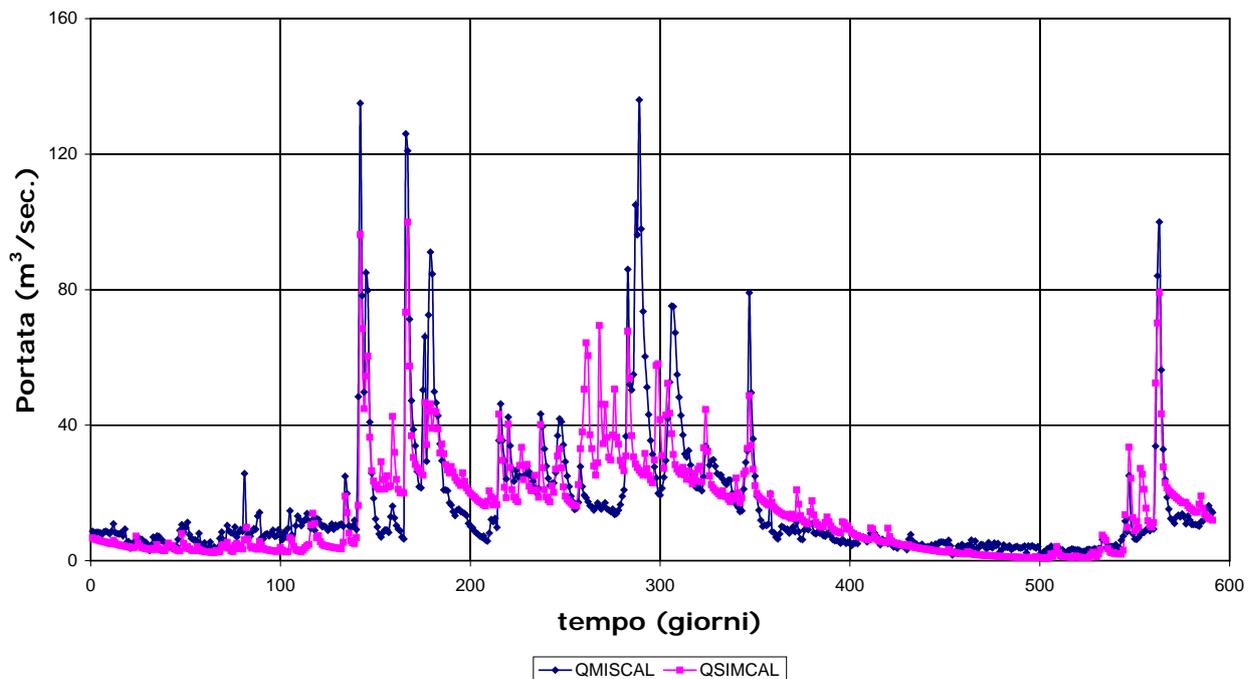


Figura 5.5. Idrogramma in calibrazione del Fiume Esino a Cannucciaro (RMSE).

Fiume Esino - Validazione (RMSE)

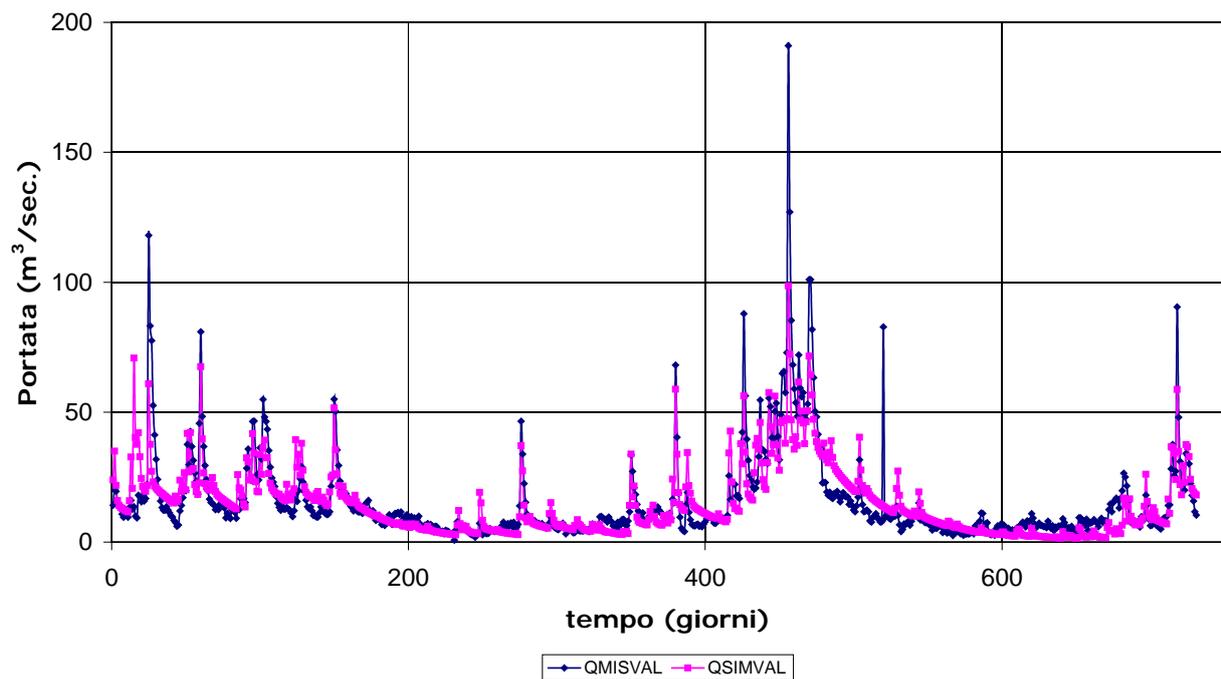


Figura 5.6. Idrogramma in validazione del Fiume Esino a Moie (RMSE).

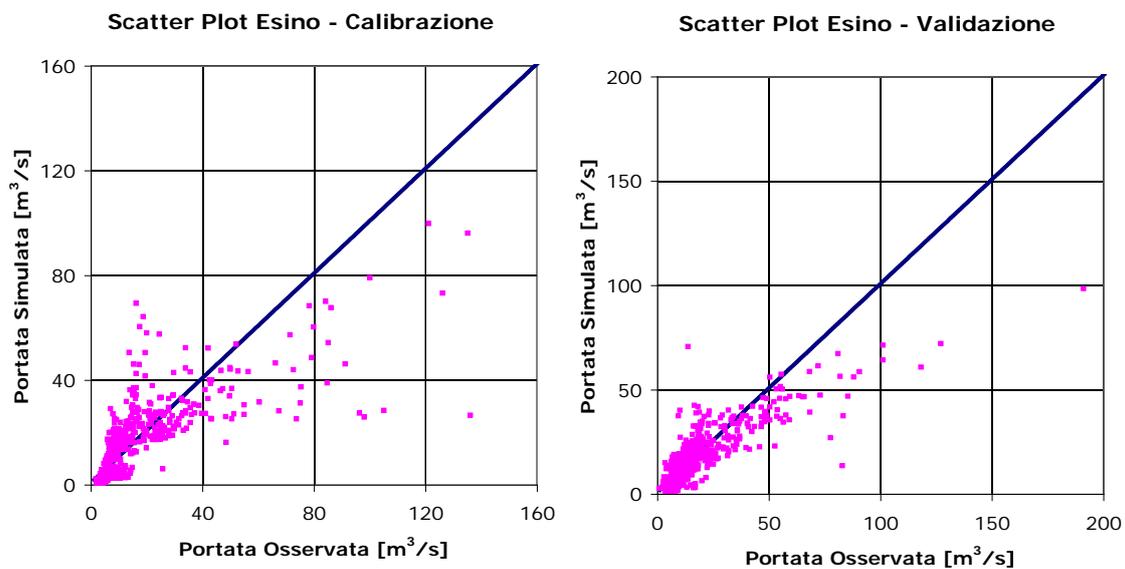


Figura 5.7. scatterplot in calibrazione e in validazione delle portate relative al Fiume Esino a Moie (RMSE).

La simulazione delle portate per il Fiume Potenza è decisamente soddisfacente con l'utilizzo dell'EFF e del RMSE per i quali i valori di tutti gli indici sono molto buoni (Tabella 5.10).

METAURO (calibrazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,65	0,72	0,71	0,71	0,72
MAE	7,61	6,95	6,64	6,7	6,95
MRE	0,73	0,69	0,44	0,44	0,69

Tabella 5.8. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

METAURO (validazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,59	0,6	0,59	0,57	0,6
MAE	8,81	8,8	8,32	8,51	8,8
MRE	0,74	0,73	0,54	0,55	0,73

Tabella 5.9. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Fiume Metauro - Calibrazione (RMSE)

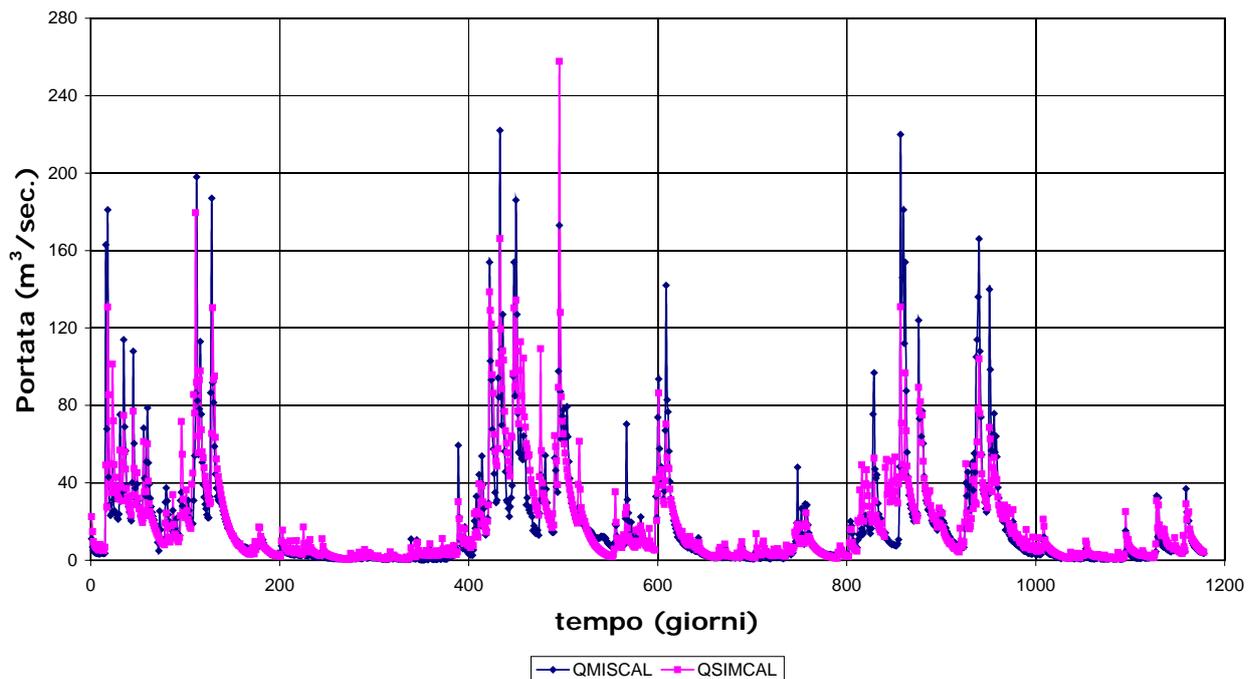


Figura 5.8. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (RMSE).

Fiume Metauro - Validazione (RMSE)

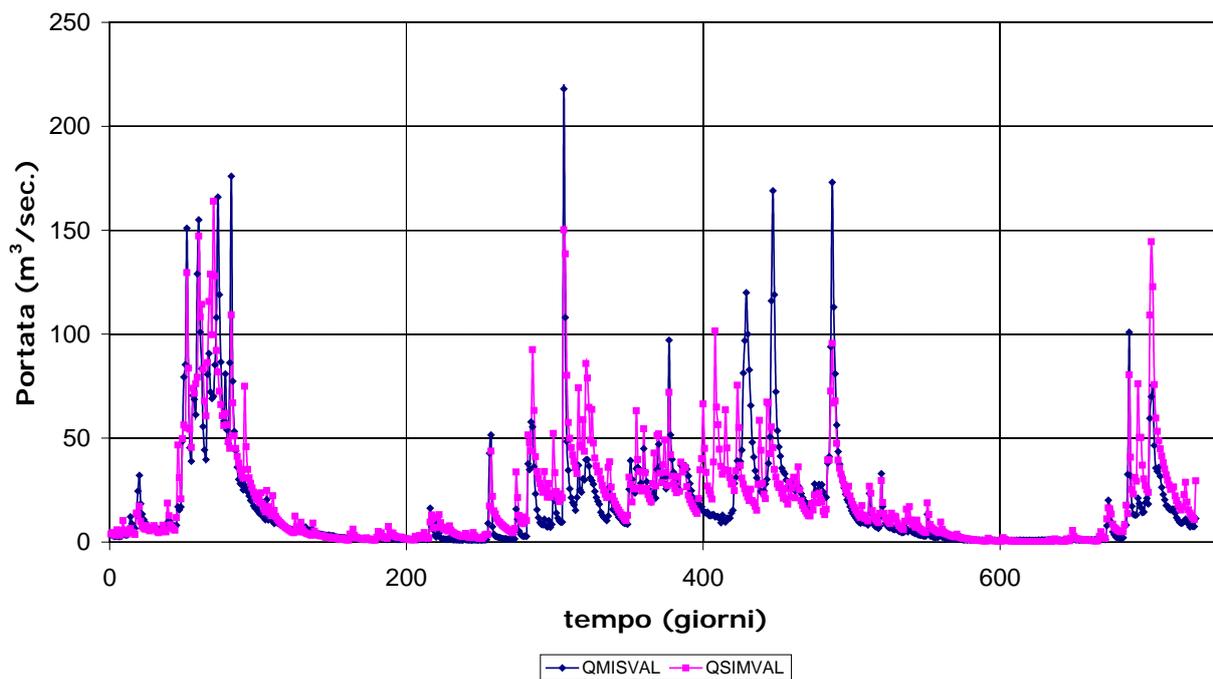


Figura 5.9. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (RMSE).

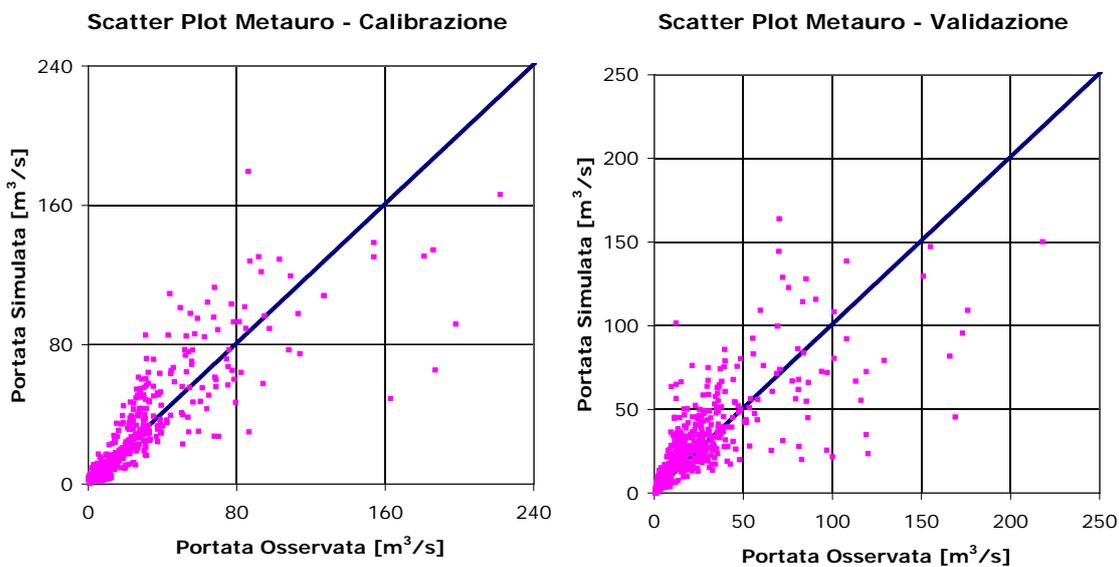


Figura 5.10. scatterplot delle portate in calibrazione e validazione relative al Fiume Metauro a Barco di Bellaguardia (RMSE).

POTENZA (calibrazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	0,58	0,78	0,73	0,5	0,78
MAE	2,77	1,77	1,66	1,86	1,77
MRE	0,47	0,26	0,22	0,2	0,26

Tabella 5.10. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

POTENZA (validazione)					
Funzione Obiettivo	BIAS	EFF	MAE	MRE	RMSE
EFF	-0,28	0,52	0,57	0,31	-0,28
MAE	3,31	2,12	1,95	1,86	2,12
MRE	0,66	0,42	0,37	0,28	0,42

Tabella 5.11. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Tuttavia si nota come la modalità che utilizza il RMSE produca un valore di EFF in validazione decisamente poco soddisfacente (Tabella 5.11). Nonostante questo dettaglio si è comunque preso in esame tale esempio come riferimento dato che un'ipotetica scelta delle modalità basata solo sulle conoscenze dei dati relativi al periodo di calibrazione, porterebbe a procedere in questo modo. Nel caso del RMSE quindi le portate sono ben riprodotte in calibrazione a meno di pochi eventi di piena (Figura 5.11) mentre in validazione si evidenzia una leggera sovrastima dei valori di portata intermedi (Figura 5.12). Conseguentemente lo scatterplot in calibrazione risulta abbastanza simmetrico e poco disperso (soprattutto nei valori di portata bassi e medi) mentre quello in validazione evidenzia la sovrastima dei volumi già percepita dall'idrogramma (Figura 5.13).

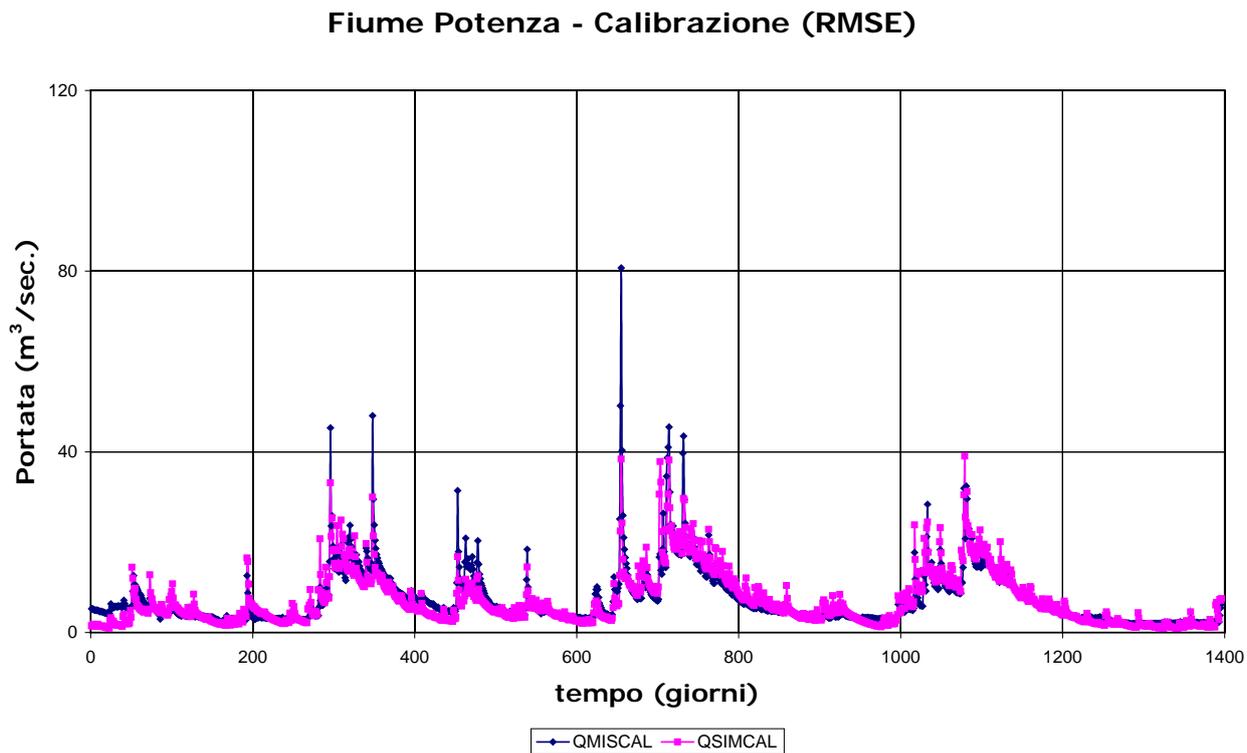


Figura 5.11. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro (RMSE).

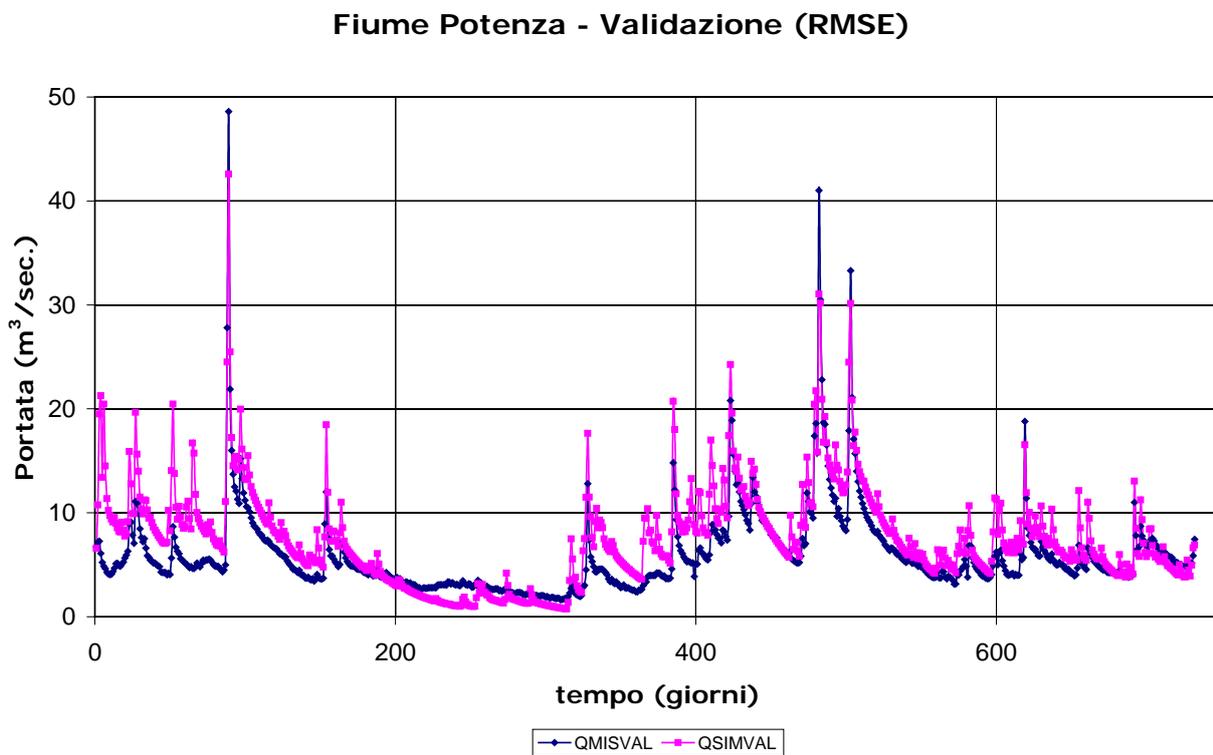


Figura 5.12. Idrogramma in validazione del Fiume Potenza a Cannuciaro (RMSE).

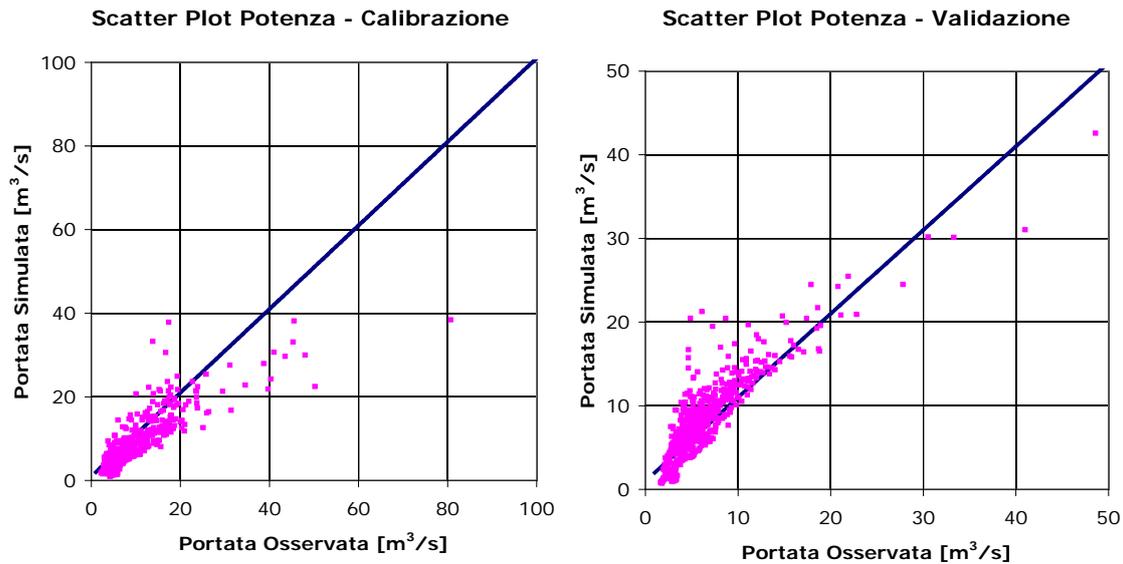


Figura 5.13. scatterplot delle portate in calibrazione e validazione relative al Fiume Potenza a Cannuciaro (RMSE).

5.3.2 Modalità multi-obiettivo (due funzioni)

Come già accennato in precedenza, l'ottimizzazione *multi-obiettivo* genera un insieme di set dei parametri ugualmente possibili e che per questo appartengono alla *Frontiera Pareto* come precedentemente definita (secondo quanto descritto nel paragrafo 1.3). Per poter effettuare un confronto visivo tra idrogramma osservato e simulato nonché per poter comparare gli indici di prestazione ottenuti con tale modalità rispetto a quelli ottenuti in modalità *mono-obiettivo*, si è presa in considerazione la portata *media* tra tutte quelle ottenute tramite i set della *Frontiera Pareto*.

Per il Fiume Candigliano i valori migliori (evidenziati in azzurro) si ottengono con la combinazione MAE-MSE e MRE-MSE per quanto riguarda l'indice EFF, mentre con le combinazioni MAE-MRE per quanto riguarda gli indici MAE ed MRE (Tabella 5.12). Ipotizziamo allora di prendere come riferimento proprio quest'ultima combinazione poiché il valore dell'EFF è comunque accettabile e non tanto minore del massimo ottenuto negli altri casi. Inoltre i valori degli indici relativi a questa modalità si mantengono buoni anche in validazione (Tabella 5.13).

Da una prima analisi visiva si nota che l'idrogramma in calibrazione relativo alla portata *media* è migliore di quello ottenuto con la modalità *mono-obiettivo* (RMSE) mostrata in precedenza poiché meglio sembrano riprodotte le portate di magra. Analoghe considerazioni possono essere fatte per l'idrogramma in validazione anche se permane la sottostima di alcuni eventi di piena significativi.

CANDIGLIANO (calibrazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,7	0,68	0,72	0,7	0,72	0,7
MAE	3,27	3,03	3,09	3,07	3,14	3,84
MRE	0,55	0,41	0,45	0,42	0,43	0,89

Tabella 5.12. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

CANDIGLIANO (validazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,63	0,57	0,63	0,58	0,62	0,52
MAE	2,62	2,26	2,39	2,33	2,33	3,13
MRE	0,56	0,41	0,47	0,44	0,43	0,81

Tabella 5.13. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

Come già sottolineato i vantaggi del *multi-obiettivo* sono legati al fatto di ottenere più set di uguale probabilità: per questo sembra interessante, oltre ad un'analisi degli indici legati alla portata media, prendere in considerazione anche gli involucri massimo e minimo di tutte le portate ottenute dai set appartenenti alla *Frontiera Pareto*. Si riportano di seguito tali involucri dove, al fine di rendere maggiormente leggibili i risultati, si sono presi in considerazione solo 150 giorni attorno all'evento più importante in termini di portata sia in calibrazione che in validazione.

In validazione (Figura 5.17) il picco è ancora sottostimato anche dall'involuppo massimo mentre le portate di magra, caratterizzate da valori molto bassi, sono riprodotte in modo soddisfacente.

Per il Fiume Esino i valori migliori (evidenziati in azzurro) si ottengono con la combinazione RMSE-BIAS per quanto riguarda l'indice EFF, con la combinazione MAE-MSLE per quanto riguarda l'indice MAE, mentre con la combinazione MAE-MRE per quanto riguarda l'indice MRE (Tabella 5.14).

Fiume Candigliano - Calibrazione (MAE-MRE)

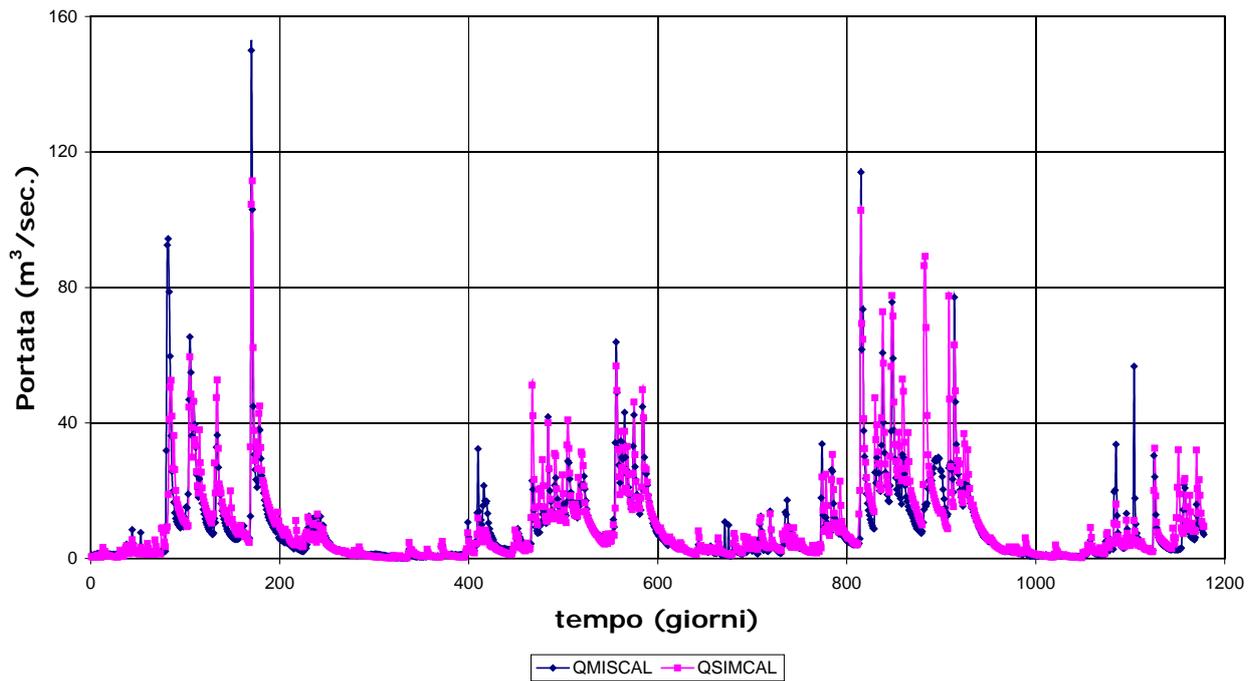


Figura 5.14. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (MAE-MRE).

Fiume Candigliano - Validazione (MAE-MRE)

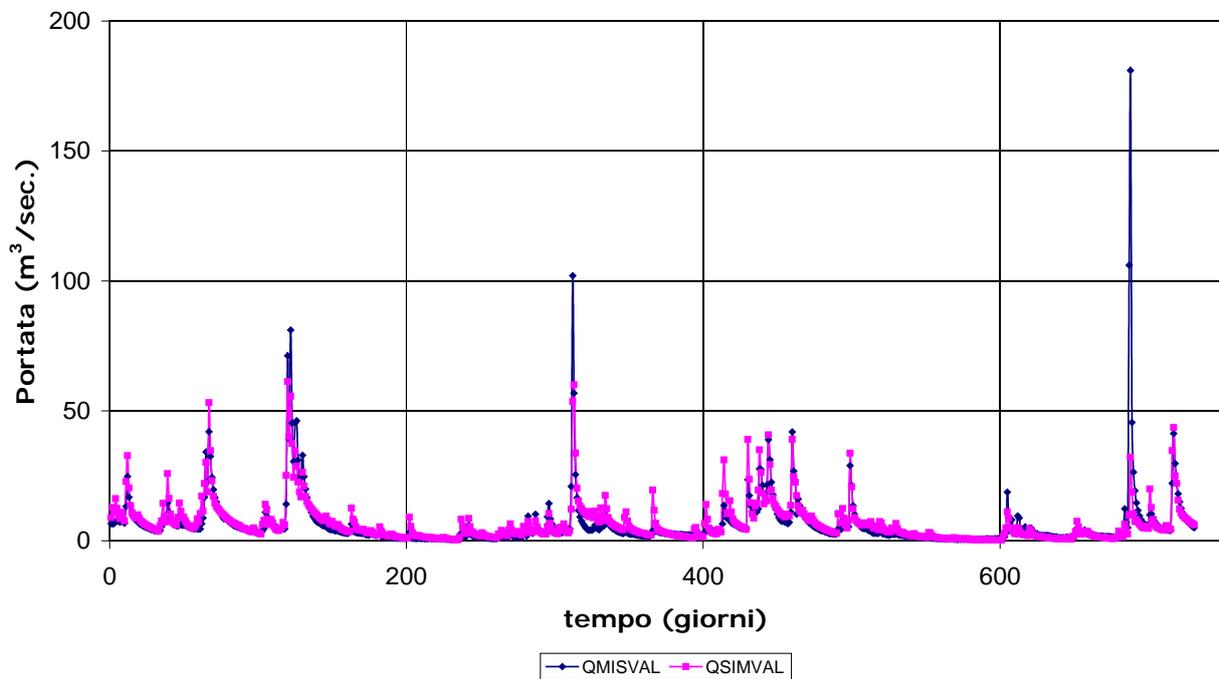


Figura 5.15. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (MAE-MRE).

Fiume Candigliano - Calibrazione (MAE-MRE)

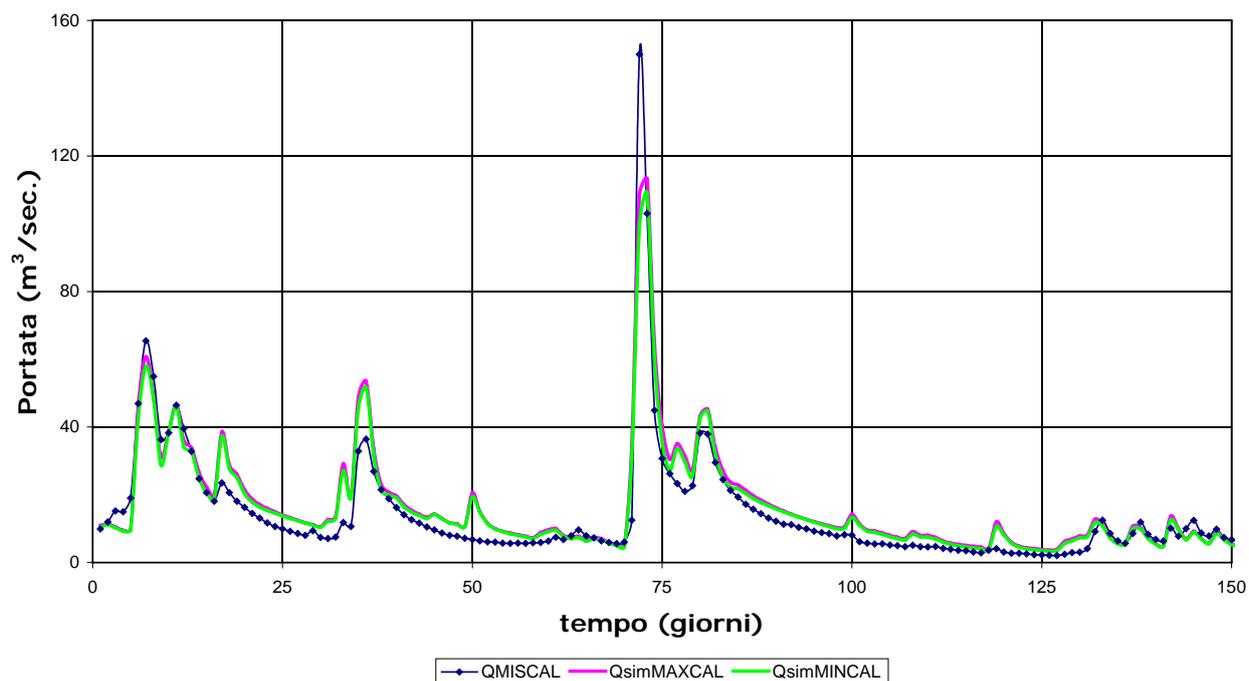


Figura 5.16. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).

Fiume Candigliano - Validazione (MAE-MRE)

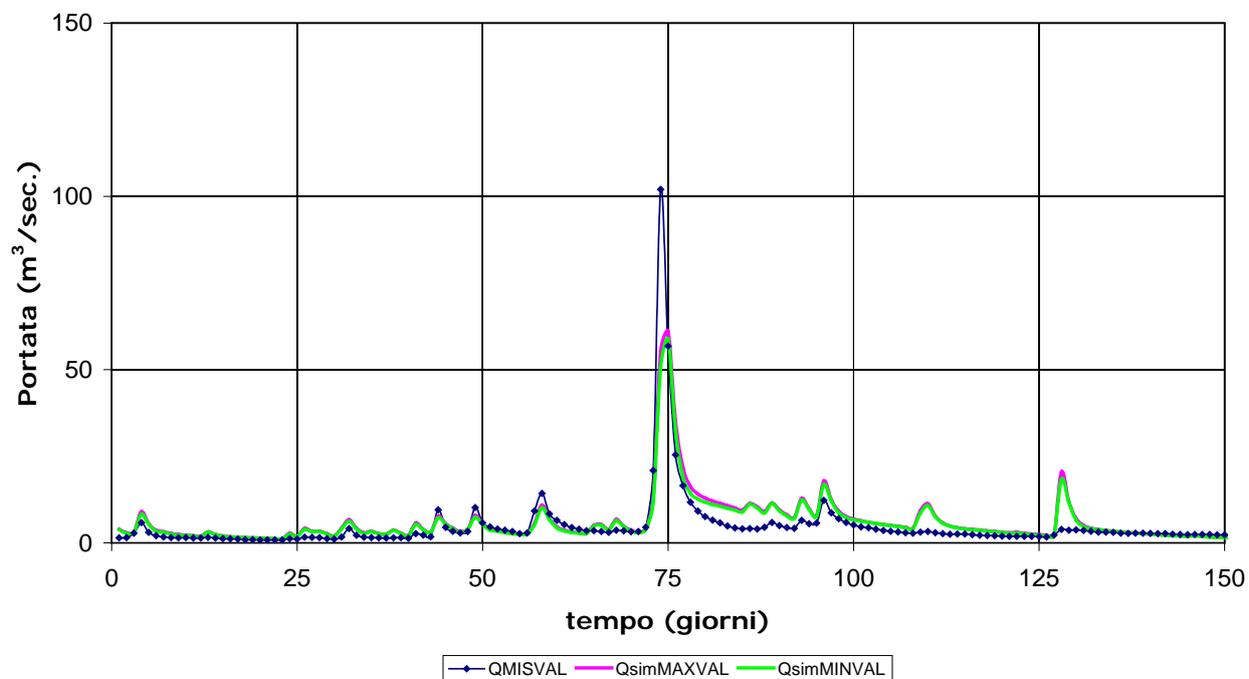


Figura 5.17. Involuppi massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto e portata osservata in validazione per il Fiume Candigliano ad Acqualagna (periodo di 150 giorni attorno al massimo evento registrato).

ESINO (calibrazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,49	0,36	0,57	0,55	0,41	0,58
MAE	7,58	7,63	6,67	6,64	7,67	6,97
MRE	0,52	0,38	0,47	0,46	0,39	0,5

Tabella 5.14. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.

Di queste combinazioni la migliore in validazione è la RMSE-BIAS (Tabella 5.15) pertanto prendiamo proprio questa come combinazione di riferimento per effettuare un confronto con la modalità *mono-obiettivo* per il Fiume Esino.

L'idrogramma in calibrazione (Figura 5.18) e in validazione (Figura 5.19) ottenuti in questo modo sono pressoché identici a quelli ottenuti con la modalità *mono-obiettivo* RMSE mostrata in precedenza (Figura 5.2 e Figura 5.3).

Analizzando un intervallo temporale ridotto nell'intorno del picco di piena osserviamo che in calibrazione ci sono evidenti difficoltà sia nel riprodurre il picco medesimo che nel seguire l'andamento nei periodi di magra (Figura 5.20); in validazione invece la riproduzione delle magre migliora in modo significativo ma rimane insoddisfacente la riproduzione delle portate di piena che vengono costantemente sottostimate (Figura 5.21).

ESINO (validazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,61	0,43	0,71	0,67	0,52	0,7
MAE	6,39	6,83	5,33	5,62	6,69	5,61
MRE	0,47	0,39	0,41	0,46	0,4	0,44

Tabella 5.15. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.

Fiume Esino - Calibrazione (RMSE-BIAS)

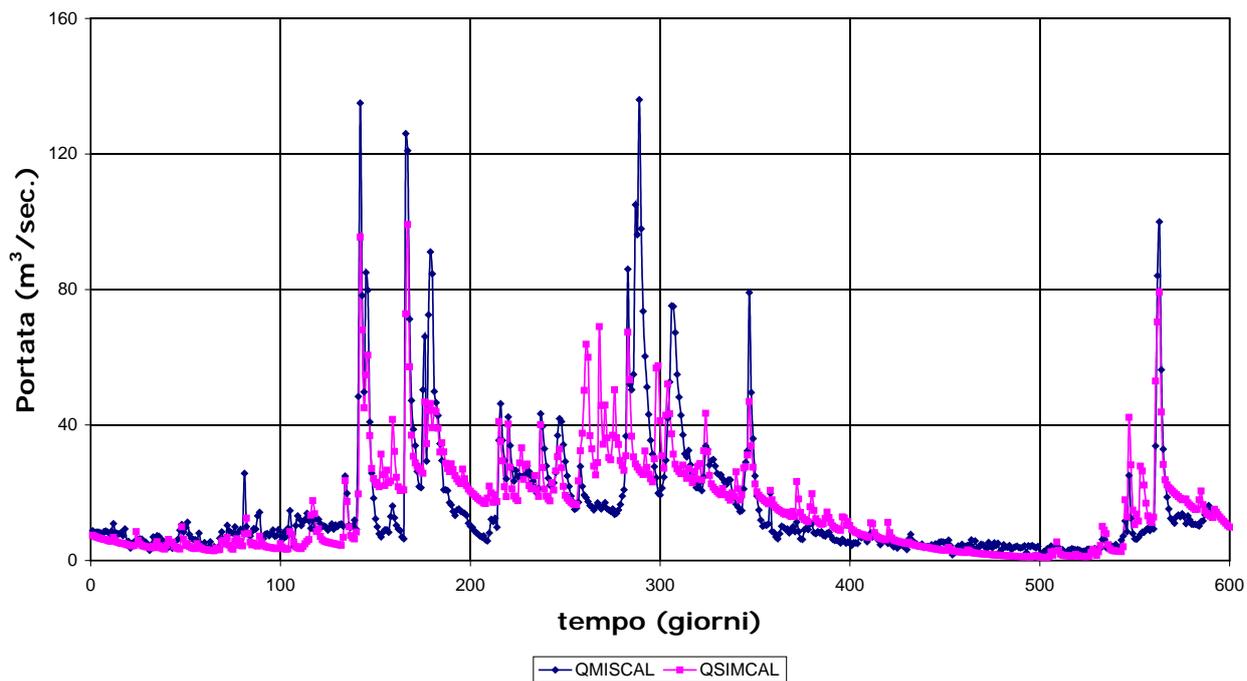


Figura 5.18. Idrogramma in calibrazione del Fiume Esino a Moie (RMSE-BIAS).

Fiume Esino - Validazione (RMSE-BIAS)

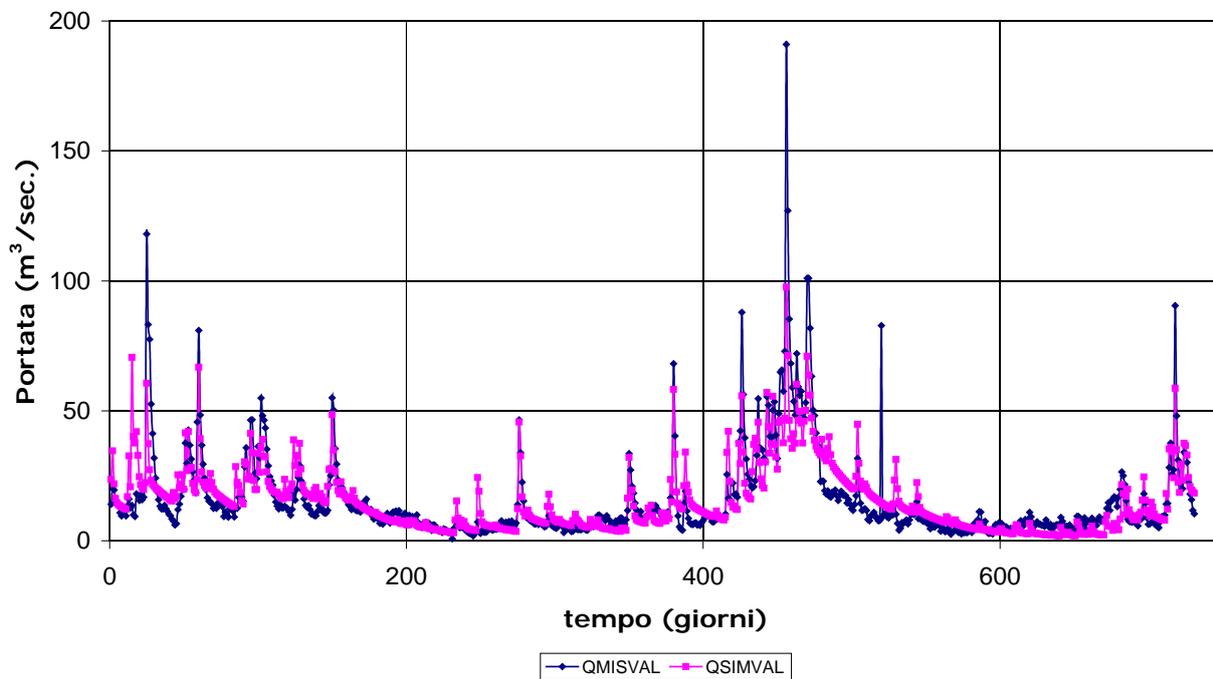


Figura 5.19. Idrogramma in validazione del Fiume Esino a Moie (RMSE-BIAS)

Fiume Esino - Calibrazione (RMSE-BIAS)

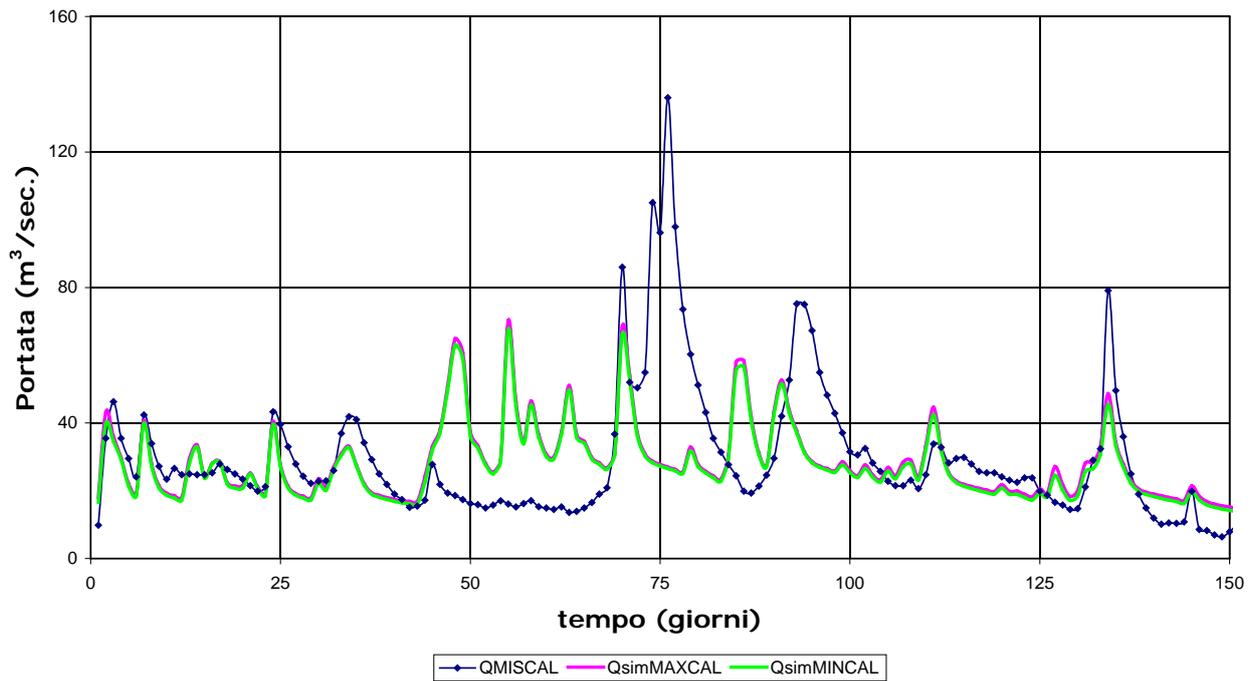


Figura 5.20. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (RMSE-BIAS).

Fiume Esino - Validazione (RMSE-BIAS)

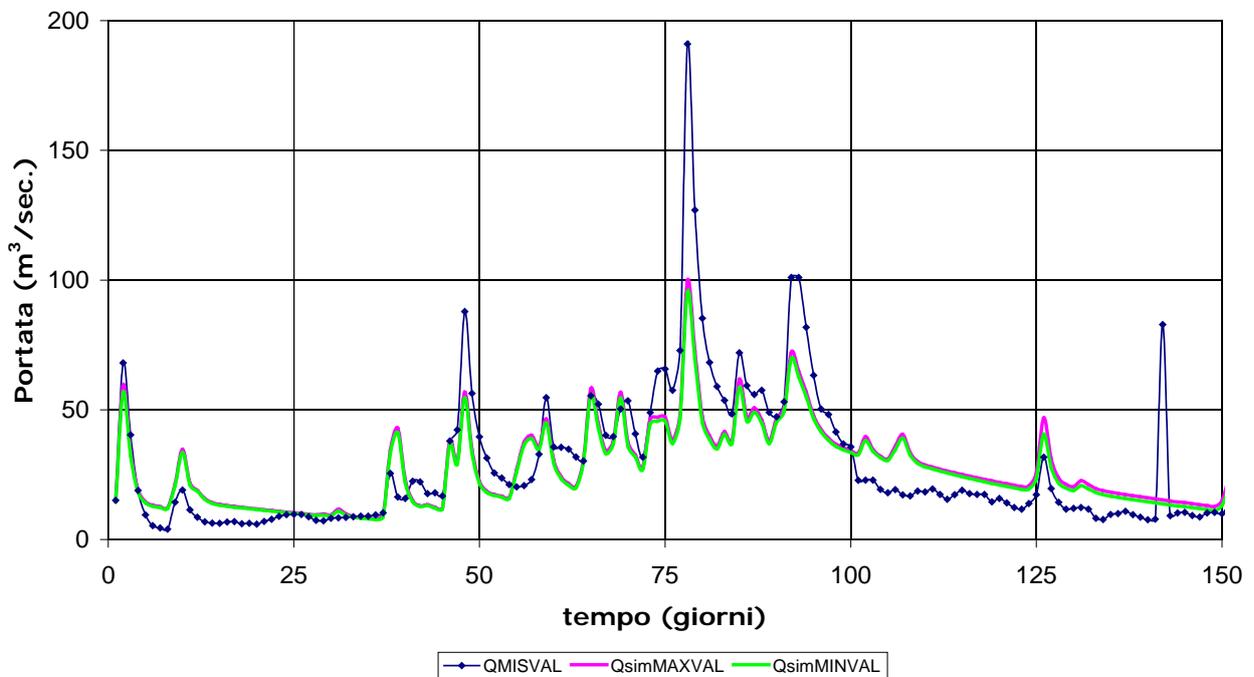


Figura 5.21. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (RMSE-BIAS).

METAURO (calibrazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,72	0,71	0,72	0,72	0,72	0,72
MAE	6,82	6,66	6,72	6,71	6,74	6,84
MRE	0,59	0,44	0,51	0,44	0,46	0,62

Tabella 5.16. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

Per il Fiume Metauro i valori migliori (evidenziati in azzurro) degli indici di prestazione si ottengono con le combinazioni RMSE-D/ND, MAE-MSE, MAE-MSLE, MRE-MSE e BIAS-RMSE per quanto riguarda l'indice EFF, con la combinazione MAE-MRE per quanto riguarda l'indice MAE, mentre con le combinazioni MAE-MRE e MAE-MSLE per quanto riguarda l'indice MRE (Tabella 5.16).

Poiché in validazione (Tabella 5.17) la coppia di funzioni obiettivo MAE-MRE mantiene prestazioni molto buone, si prende questa modalità come riferimento per un confronto con le simulazioni *mono-obiettivo* già esaminate.

METAURO (validazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,6	0,58	0,6	0,59	0,59	0,6
MAE	8,58	8,37	8,44	8,53	8,57	8,62
MRE	0,66	0,54	0,6	0,56	0,58	0,68

Tabella 5.17. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

La simulazione del modello così ottenuta (Figura 5.22 e Figura 5.23) è in realtà molto simile a quella ottenuta tramite l'utilizzo del RMSE (Figura 5.8). Osservando gli inviluppi (Figura 5.24 e Figura 5.25) si nota che essi sono molto vicini e pertanto la portata osservata è quasi sempre fuori dagli inviluppi medesimi. In calibrazione poi (Figura 5.24) si verifica una sovrastima del picco di piena mentre viene enfatizzato l'evento di piena in prossimità della fine del periodo graficato, a fronte invece di una portata osservata non particolarmente significativa.

Fiume Metauro - Calibrazione (MAE-MRE)

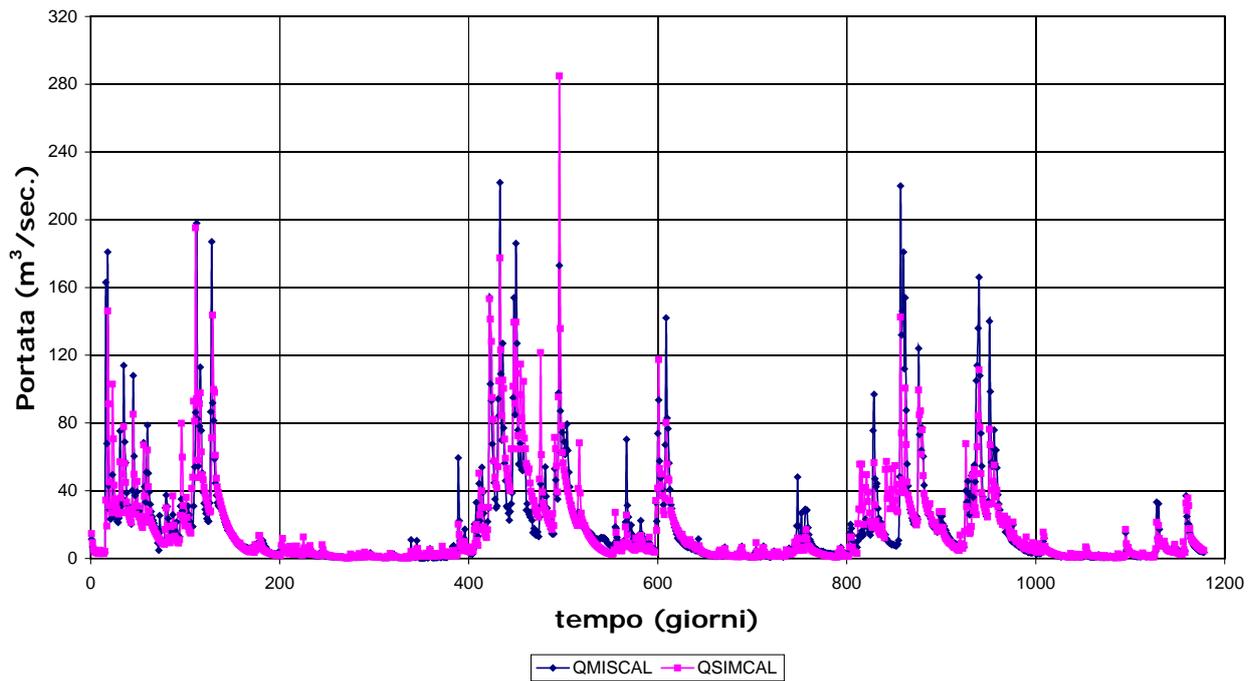


Figura 5.22. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (MAE-MRE).

Fiume Metauro - Validazione (MAE-MRE)

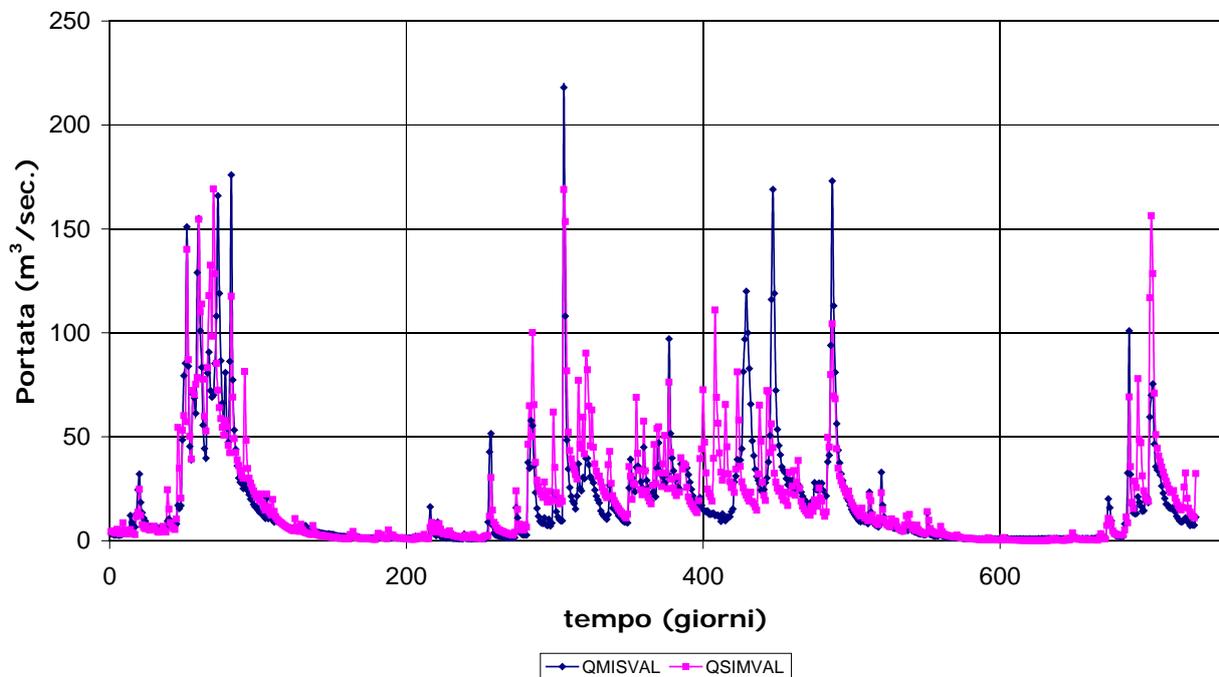


Figura 5.23. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (MAE-MRE).

Una possibile spiegazione può essere legata al fatto che Hymod non tiene conto, nella modellazione afflussi-deflussi, del diverso contributo apportato dagli eventi piovosi

piuttosto che dagli eventi nivali. I bacini che compongono l'area di studio invece sono caratterizzati da quote elevate e molti eventi modellati come piovosi sono in realtà precipitazioni nevose che danno vita ad idrogrammi con andamenti meno ripidi e più distesi nel tempo. In questi casi è quindi evidente che il modello risponde con una crescita improvvisa ed elevata dell'idrogramma, come se i volumi contribuissero tutti molto velocemente, quando in realtà l'onda reale cresce in modo più lento e moderato consentendo uno smaltimento dei volumi che non permette di raggiungere valori di portata tanto elevati.

Fiume Metauro - Calibrazione (MAE-MRE)

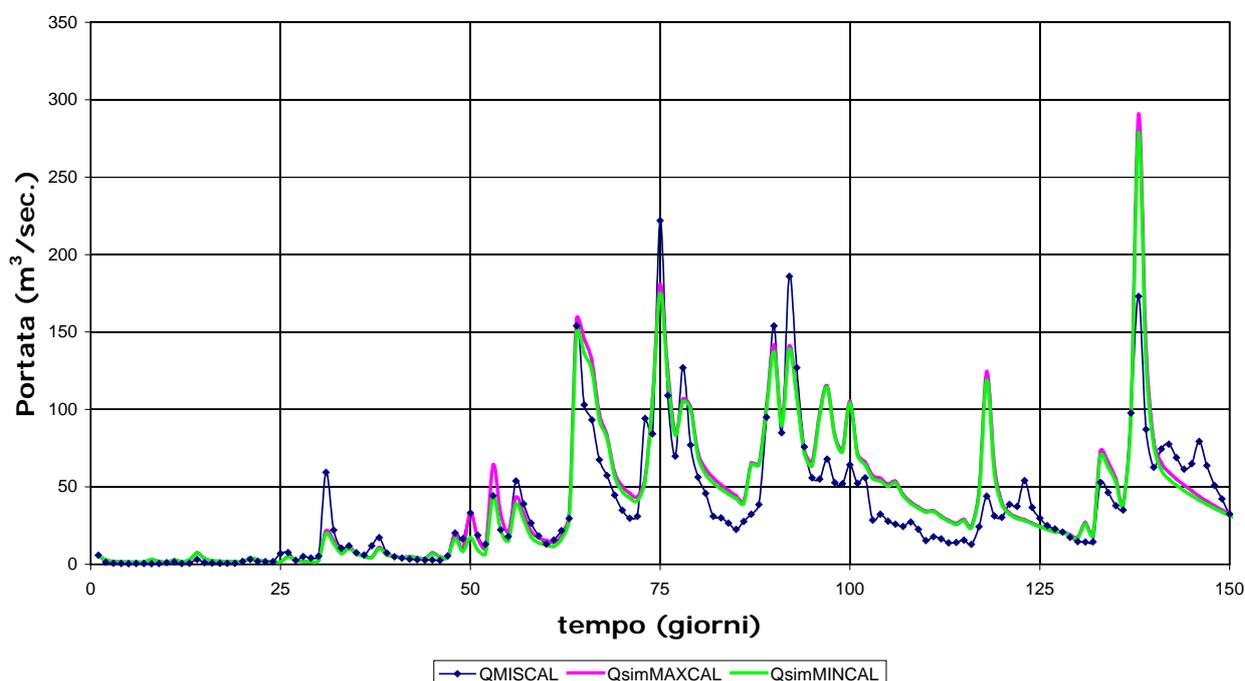


Figura 5.24. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).

I valori osservati sono fuori dall'involuppo delle portate simulate anche in riferimento al periodo di validazione (Figura 5.25): i valori di MAE in tal caso sono decisamente importanti e questo suggerisce che, nonostante gli involuppi riproducano indicativamente l'andamento delle portate, gli scostamenti nel complesso sono significativi sia in periodi di scarsi deflussi che in periodi di picchi di portata.

La calibrazione per il Fiume Potenza ha prodotto valori degli indici di qualità molto buoni (Tabella 5.18). Per quanto riguarda l'EFF i valori migliori si raggiungono con l'utilizzo delle combinazioni RMSE-D/ND e RMSE-BIAS mentre per quanto riguarda il MAE e il MRE le combinazioni migliori sono rispettivamente il MAE-MSLE e il MAE-MRE (indicate in azzurro). Nel complesso però la combinazione MAE-MSLE è quella che

mantiene risultati più soddisfacenti anche in validazione (Tabella 5.19) e per questo viene presa come riferimento per il confronto con la modalità *mono-obiettivo*.

L'andamento della media delle portate simulate in calibrazione (Figura 5.26) è molto vicino alla portata osservata soprattutto in riferimento agli eventi di magra, mentre anche in questo caso i picchi di portata più significativi non sono ben riprodotti.

In validazione invece (Figura 5.27), per quanto l'andamento generale dell'idrogramma medio sia molto simile a quello dell'osservato, si evidenzia una tendenza alla sovrastima delle portate basse oltre alla permanenza di una sottostima dei picchi.

Fiume Metauro - Validazione (MAE-MRE)

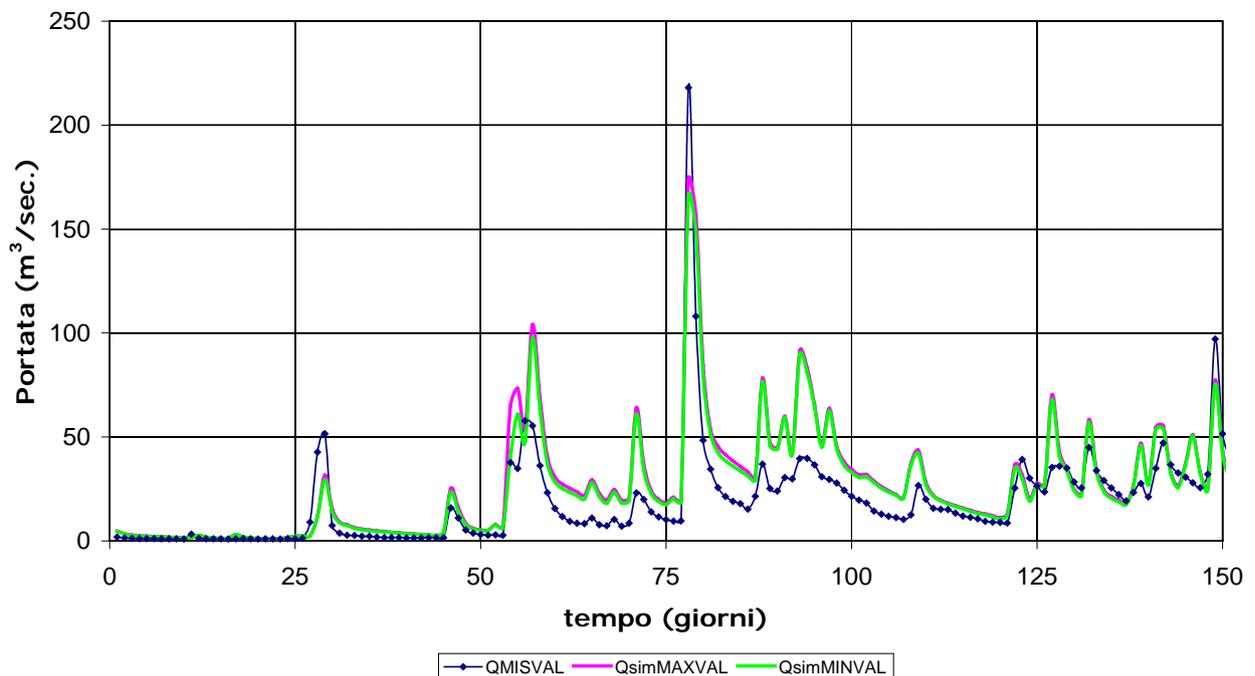


Figura 5.25. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).

POTENZA (calibrazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,78	0,67	0,77	0,72	0,75	0,78
MAE	1,77	1,74	1,69	1,68	1,69	1,77
MRE	0,26	0,21	0,23	0,22	0,23	0,26

Tabella 5.18. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

POTENZA (validazione)						
Funzione Obiettivo	RMSE D/ND	MAE-MRE	MAE-MSE	MAE MSLE	MRE-MSE	RMSE BIAS
EFF	0,58	0,59	0,63	0,62	0,66	0,57
MAE	2,2	1,92	2,01	1,95	1,92	2,21
MRE	0,43	0,34	0,39	0,37	0,38	0,43

Tabella 5.19. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).

Fiume Potenza - Calibrazione (MAE-MSLE)

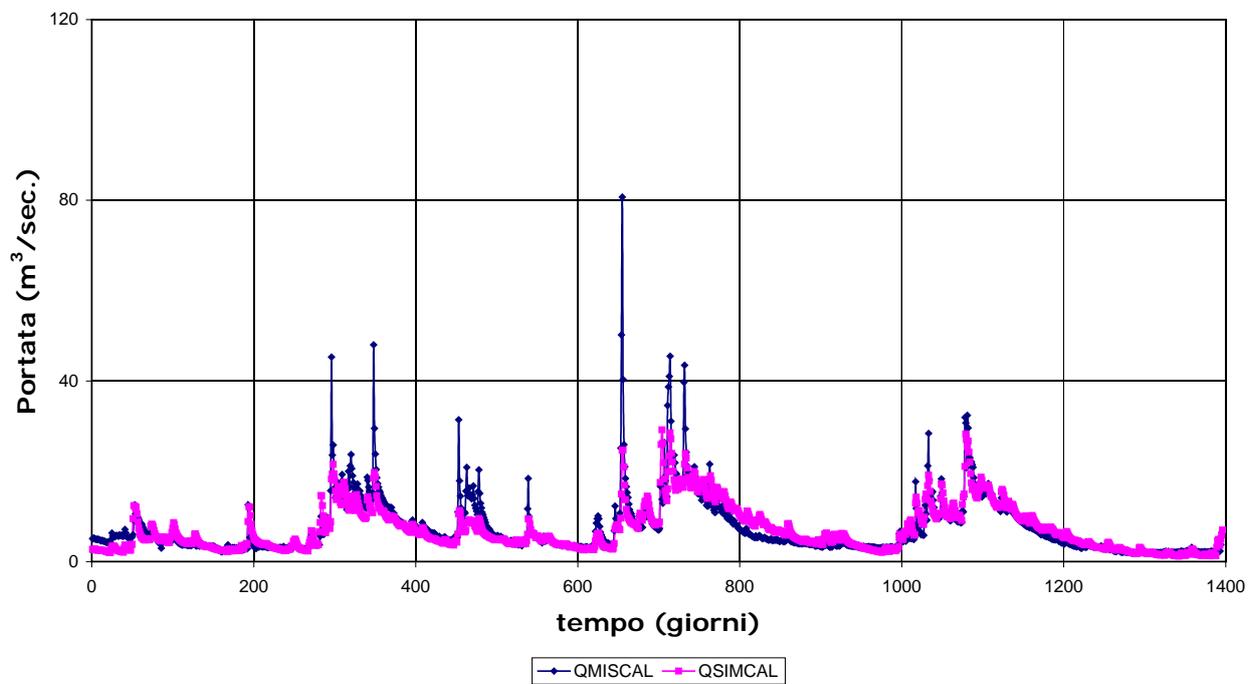


Figura 5.26. Idrogramma in calibrazione del Fiume Potenza a Cannuccio (MAE-MSLE).

Fiume Potenza - Validazione (MAE-MSLE)

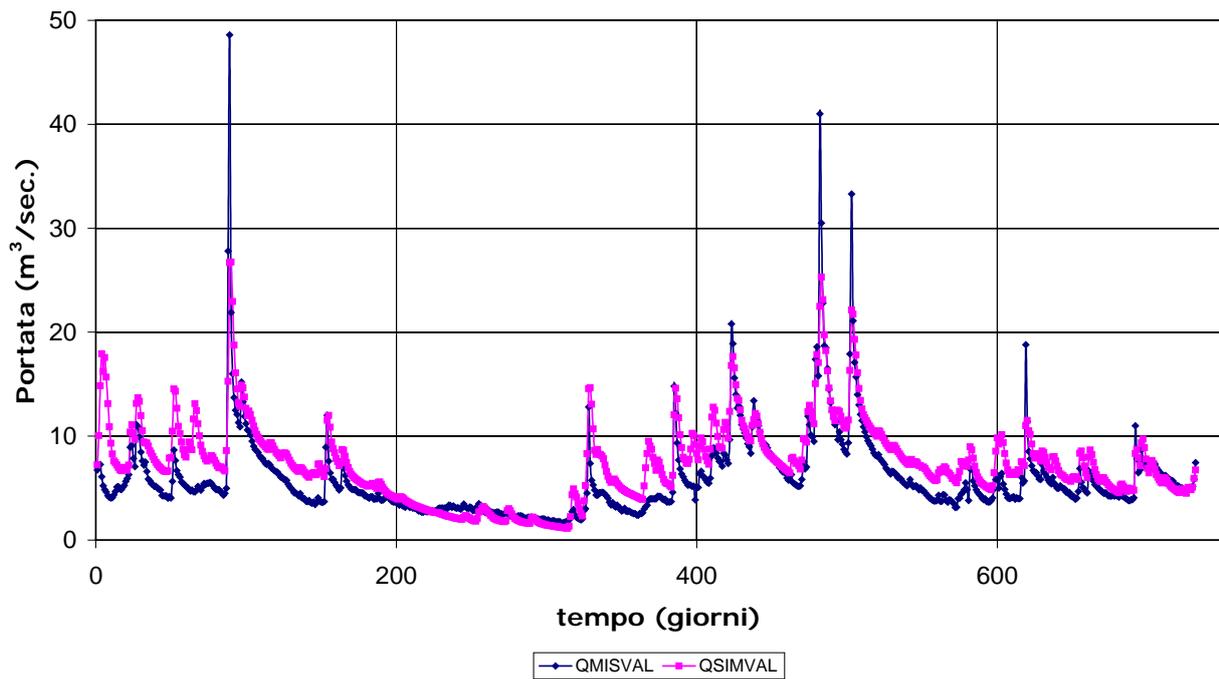


Figura 5.27. Idrogramma in validazione del Fiume Potenza a Cannucciaro (MAE-MSLE).

Fiume Potenza - Calibrazione (MAE-MSLE)

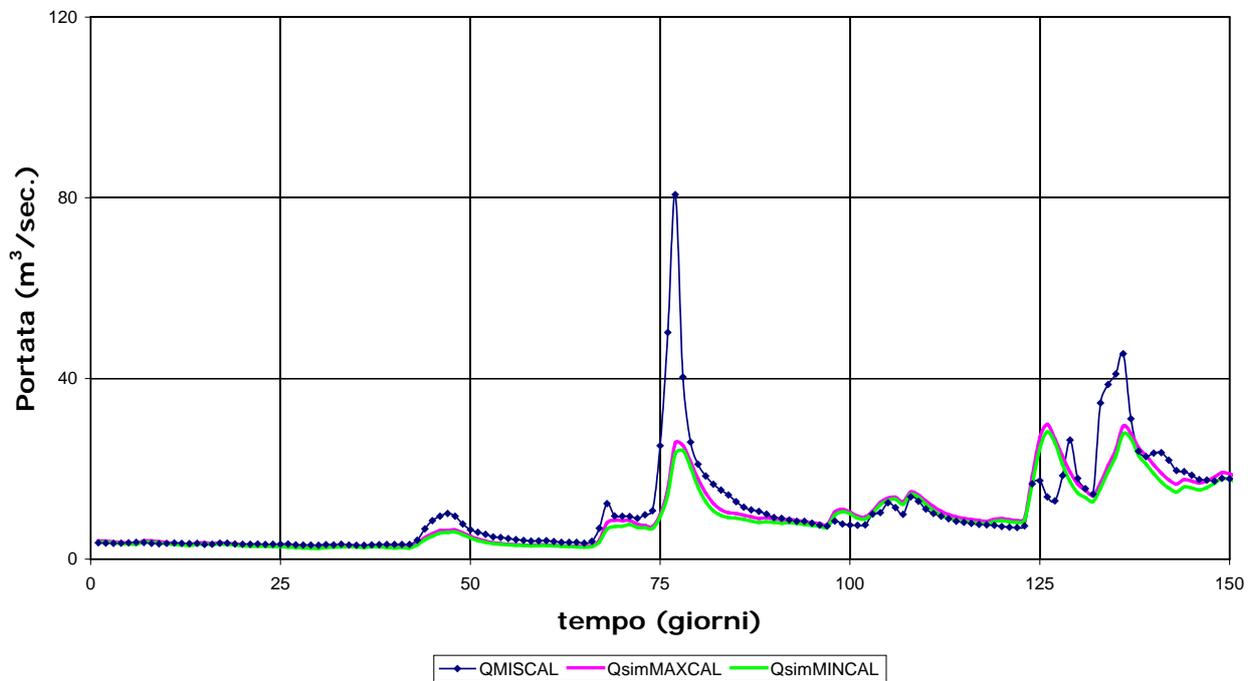


Figura 5.28. Involupi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSLE).

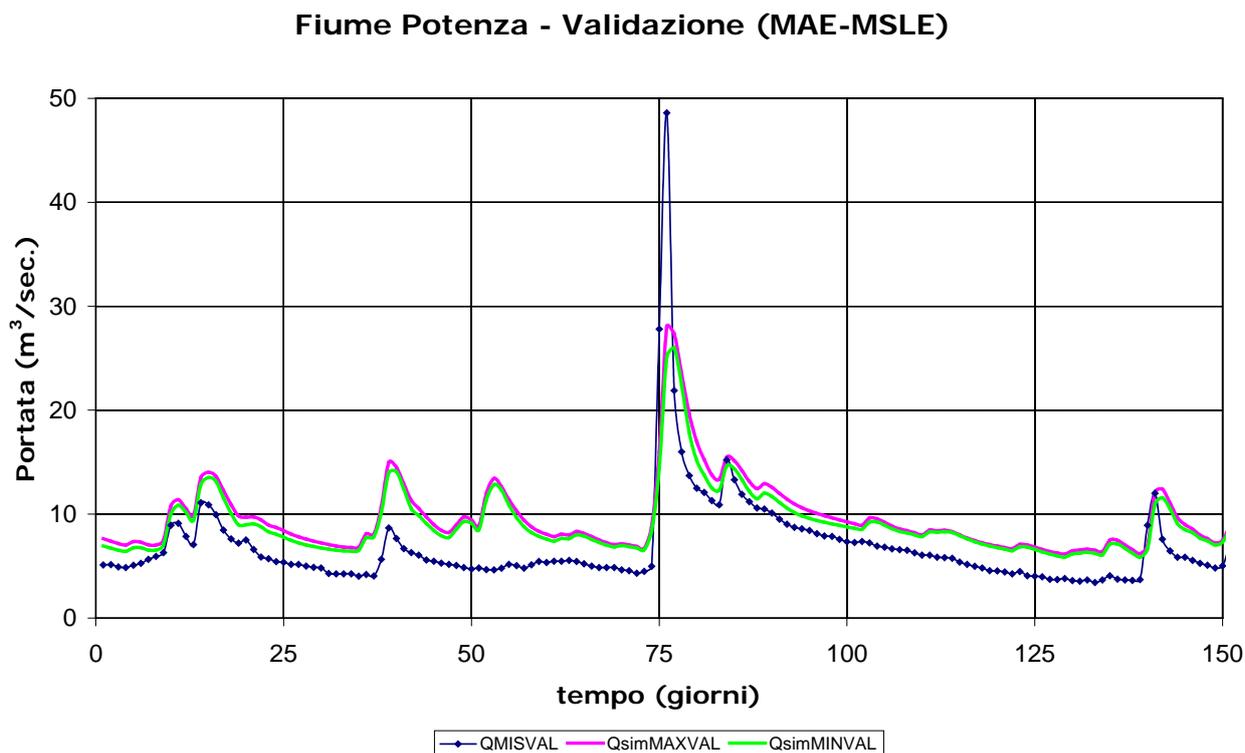


Figura 5.29. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSLE).

Osservando gli inviluppi massimo e minimo (Figura 5.28 e Figura 5.29) si nota subito che la loro vicinanza impedisce alla curva delle portate osservate di essere inglobata al loro interno. Tuttavia gli andamenti sia in calibrazione che in validazione sono molto vicini a quelli osservati e per questo la simulazione è da ritenersi soddisfacente.

E' importante sottolineare che per quest'ultimo bacino tutte le combinazioni di funzioni obiettivo hanno portato a risultati soddisfacenti a dimostrazione che il modello riproduce bene il comportamento di quest'area.

5.3.3 Modalità multi-obiettivo (tre funzioni)

Al fine di completare l'analisi del comportamento dell'algoritmo AMALGAM in modalità *multi-obiettivo* e di analizzare le combinazioni di funzioni obiettivo che producono i risultati più soddisfacenti per l'area in esame, sono state effettuate due calibrazioni a tre funzioni-obiettivo ottenute combinando MAE-MSE-MRE e MAE-MSLE-MRE.

Anche in questo caso, al fine di effettuare un confronto visivo tra idrogramma osservato e simulato nonché per poter comparare gli indici di prestazione ottenuti con tale modalità rispetto a quelli ottenuti in modalità *mono-obiettivo* e *multi-obiettivo* a

tre funzioni, si è presa in considerazione la portata *media* tra tutte quelle ottenute tramite i set della frontiera Pareto.

Per il Fiume Candigliano le due combinazioni danno risultati molto simili (Tabella 5.20 e Tabella 5.21) e l'idrogramma medio è pressoché identico. Quello mostrato in Figura 5.30 è riferito al caso MAE-MSE-MRE.

Da un'attenta osservazione dell'idrogramma medio in calibrazione (Figura 5.30) ottenuto con questa combinazione rispetto a quella che ha prodotto i valori degli indici migliori in modalità *mono-obiettivo* (RMSE - Figura 5.2), è evidente come i picchi di piena siano ora meglio avvicinati e come le portate di magra siano riprodotte in modo più dettagliato sia in termini di entità che in termini di sincronismo temporale.

CANDIGLIANO (calibrazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,71	0,7
MAE	3,1	3,07
MRE	0,43	0,42

Tabella 5.20. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

CANDIGLIANO (validazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,62	0,57
MAE	2,33	2,3
MRE	0,44	0,43

Tabella 5.21. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

La stessa considerazione può essere fatta in merito al periodo di validazione (Figura 5.31) anche se il picco di portata massimo non viene riprodotto neppure in questo caso.

L'entità di tale picco, pari a quasi il doppio del secondo picco di portata registrato in tutto il periodo di calibrazione, sembra tuttavia indicare un possibile errore nei dati.

Fiume Candigliano - Calibrazione (MAE-MSE-MRE)

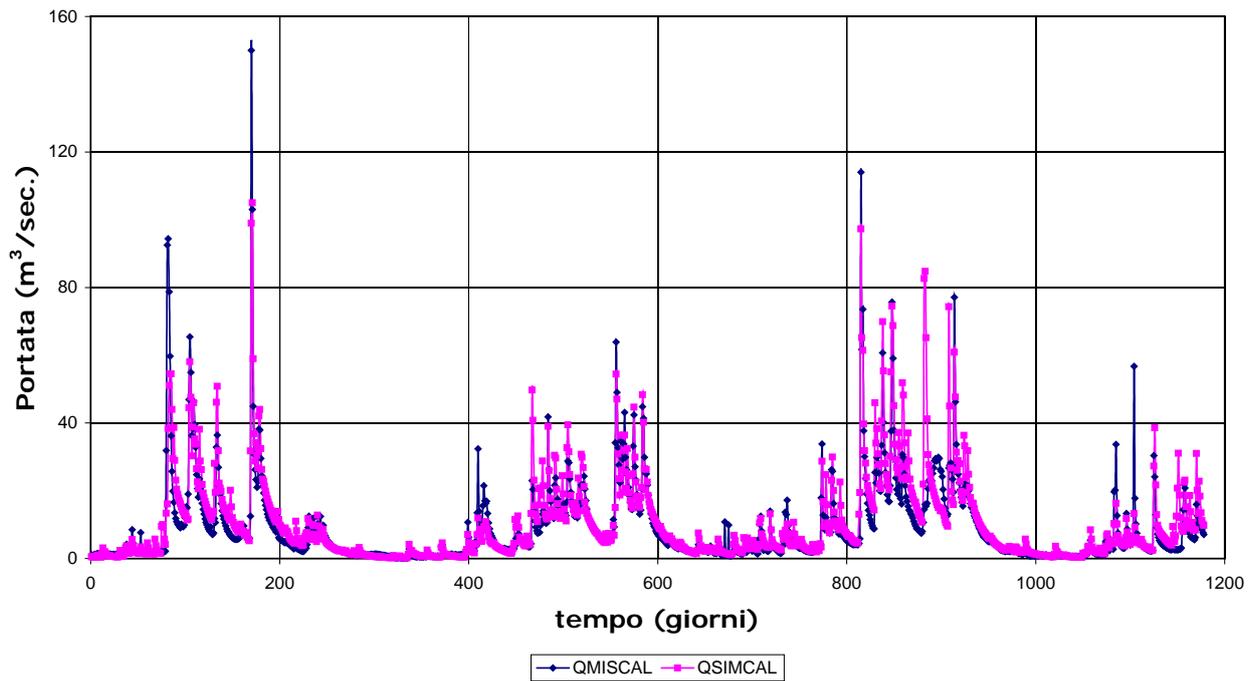


Figura 5.30. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (MAE-MSE-MRE).

Fiume Candigliano - Validazione (MAE-MSE-MRE)

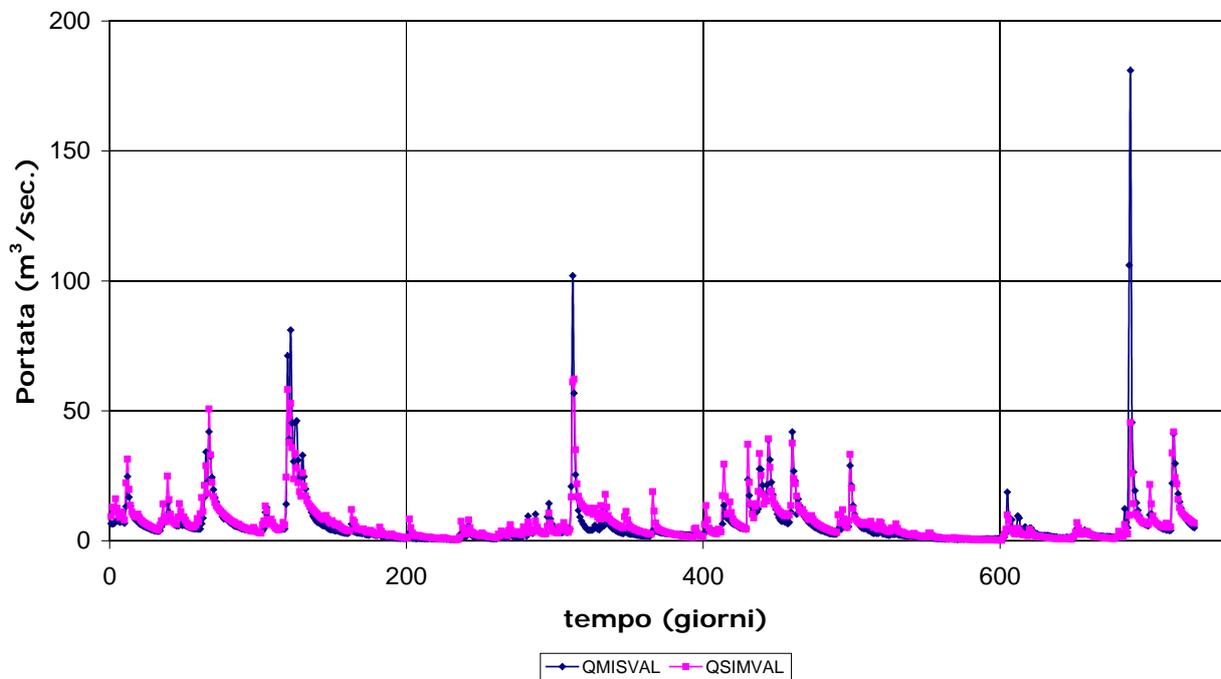


Figura 5.31. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (MAE-MSE-MRE).

Fiume Candigliano - Calibrazione (MAE-MSE-MRE)

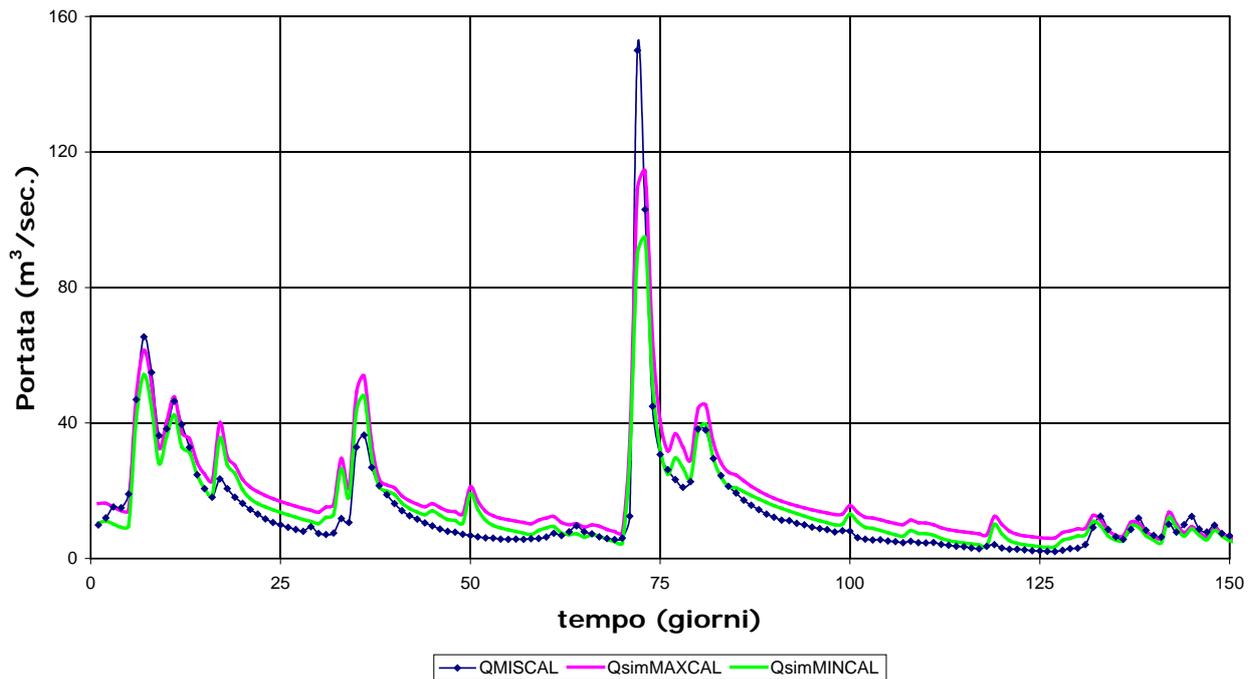


Figura 5.32. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

L'inviluppo massimo e minimo delle portate in calibrazione (Figura 5.32) è decisamente soddisfacente poiché riproduce in modo estremamente preciso l'andamento delle portate osservate e si avvicina molto di più rispetto alle altre simulazioni al picco di portata. Per quanto riguarda il periodo di validazione si è deciso di graficare il secondo picco proprio per quanto sottolineato precedentemente (Figura 5.33): anche in questo caso il risultato è decisamente buono poiché in prossimità di portate molto basse, il modello riesce a riprodurre bene le piccole oscillazioni registrate, mentre il valore di picco è sottostimato in maniera contenuta dall'inviluppo massimo.

Per quanto riguarda il Fiume Esino, le simulazioni migliori in base agli indici di prestazione, si ottengono con entrambe le combinazioni a seconda che si consideri l'EFF, il MAE o l'MRE (Tabella 5.22). Osservando però quanto accade in validazione, si preferisce analizzare i grafici relativi alla combinazione *multi-obiettivo* MAE-MSE-MRE (Tabella 5.23).

Gli indici in realtà non sono in nessun caso particolarmente buoni come risulta evidente dall'idrogramma delle portate medie che in calibrazione è molto lontano dall'osservato (Figura 5.34).

Fiume Candigliano - Validazione (MAE-MSE-MRE)

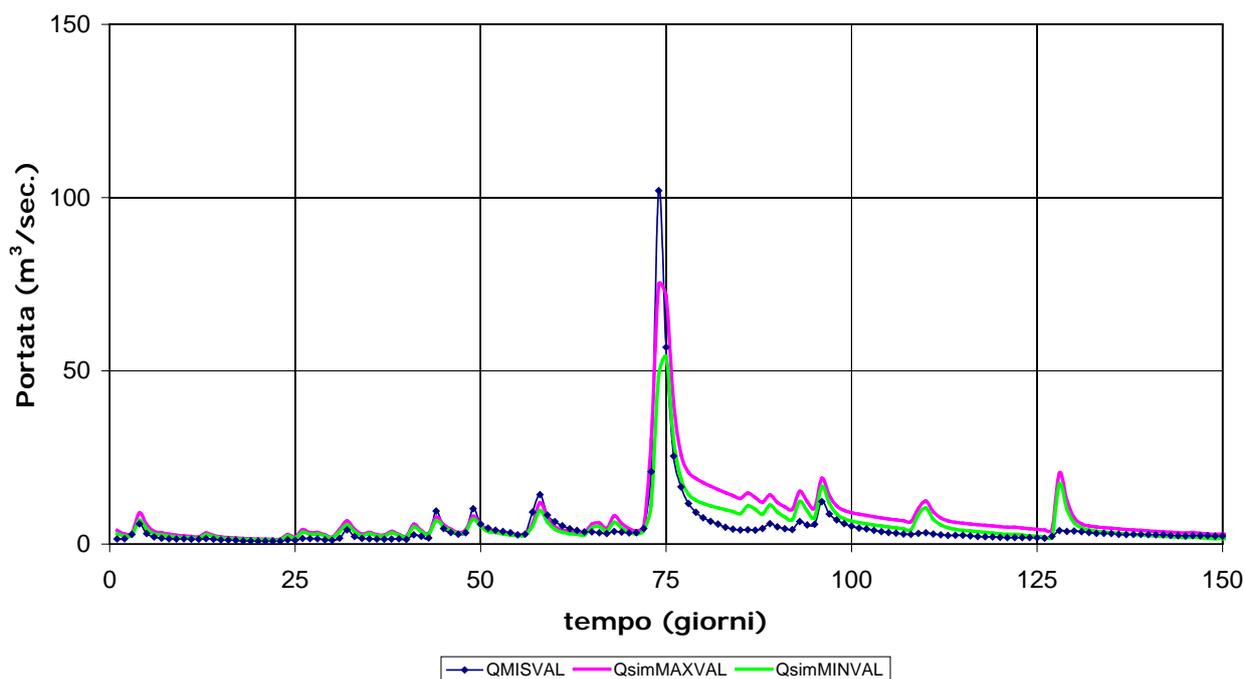


Figura 5.33. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

ESINO (calibrazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,42	0,38
MAE	7,6	7,52
MRE	0,39	0,39

Tabella 5.22. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

Le portate di magra sono riprodotte in modo discreto ma nessun picco è colto dalla simulazione del modello che non supera mai valori di portata pari in realtà ad un terzo di quella realmente registrata.

In validazione gli indici di qualità sono leggermente migliori ma l'osservazione degli idrogrammi ci porta sostanzialmente alle stesse considerazioni (Figura 5.35).

Il modello non riesce a riprodurre bene il comportamento di questo bacino: osservando gli inviluppi infatti (Figura 5.36 e Figura 5.37) sembrerebbe che la portata osservata si trovi spesso inglobata in tali inviluppi. Tuttavia la loro distanza e il loro

andamento ben poco simile a quello dell'idrogramma osservato indica un' incapacità di simulare i processi di deflusso sia superficiale che profondo.

L'andamento in validazione è leggermente migliore (Figura 5.37): ad eccezione dei valori di picco, la portata è sempre inclusa negli involuipi che questa volta (soprattutto l'involuppo massimo) hanno un andamento leggermente più simile a quello della portata osservata.

ESINO (validazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,53	0,48
MAE	6,65	6,67
MRE	0,4	0,4

Tabella 5.23. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

Fiume Esino - Calibrazione (MAE-MSE-MRE)

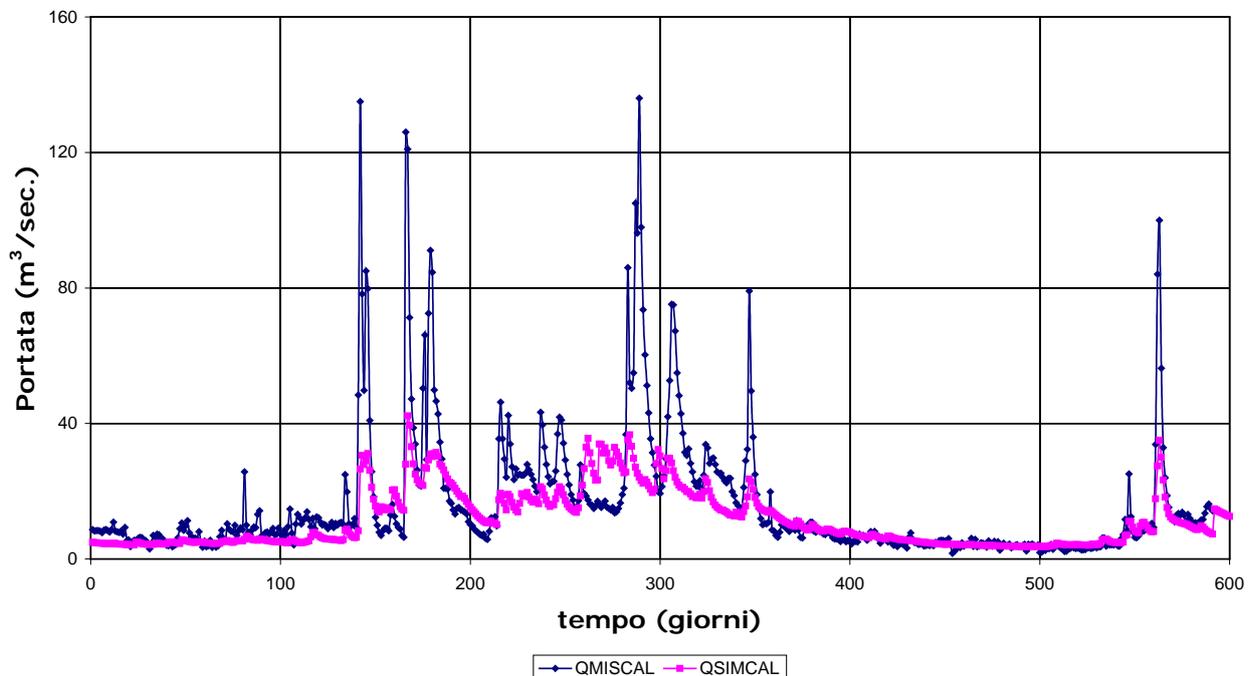


Figura 5.34. Idrogramma in calibrazione del Fiume Esino a Moie (MAE-MSE-MRE).

Fiume Esino - Validazione (MAE-MSE-MRE)

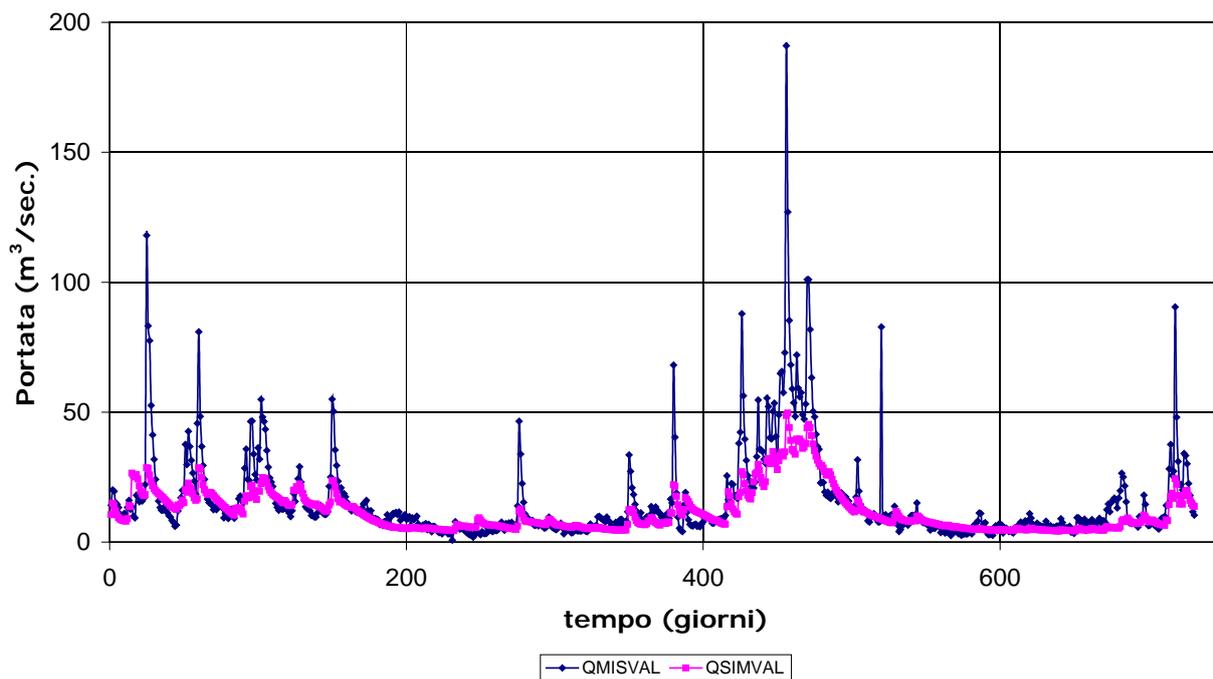


Figura 5.35. Idrogramma in validazione del Fiume Esino a Moie (MAE-MSE-MRE).

Fiume Esino - Calibrazione (MAE-MSE-MRE)

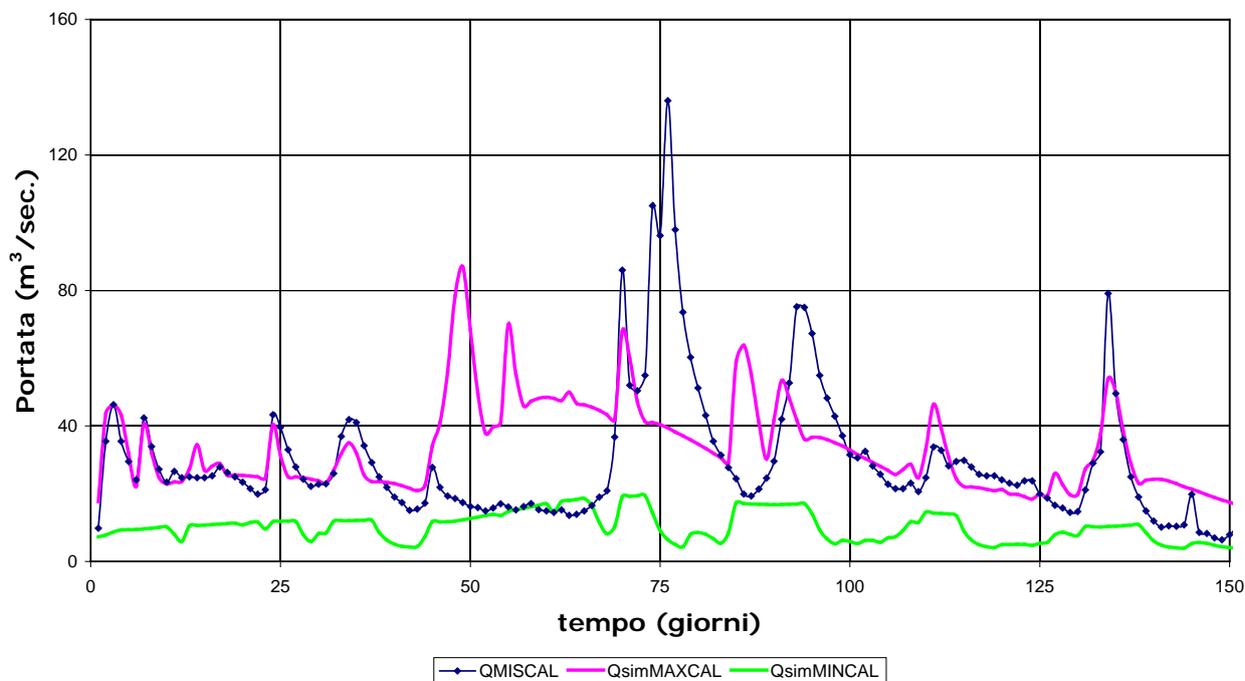


Figura 5.36. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

Fiume Esino - Validazione (MAE-MSE-MRE)

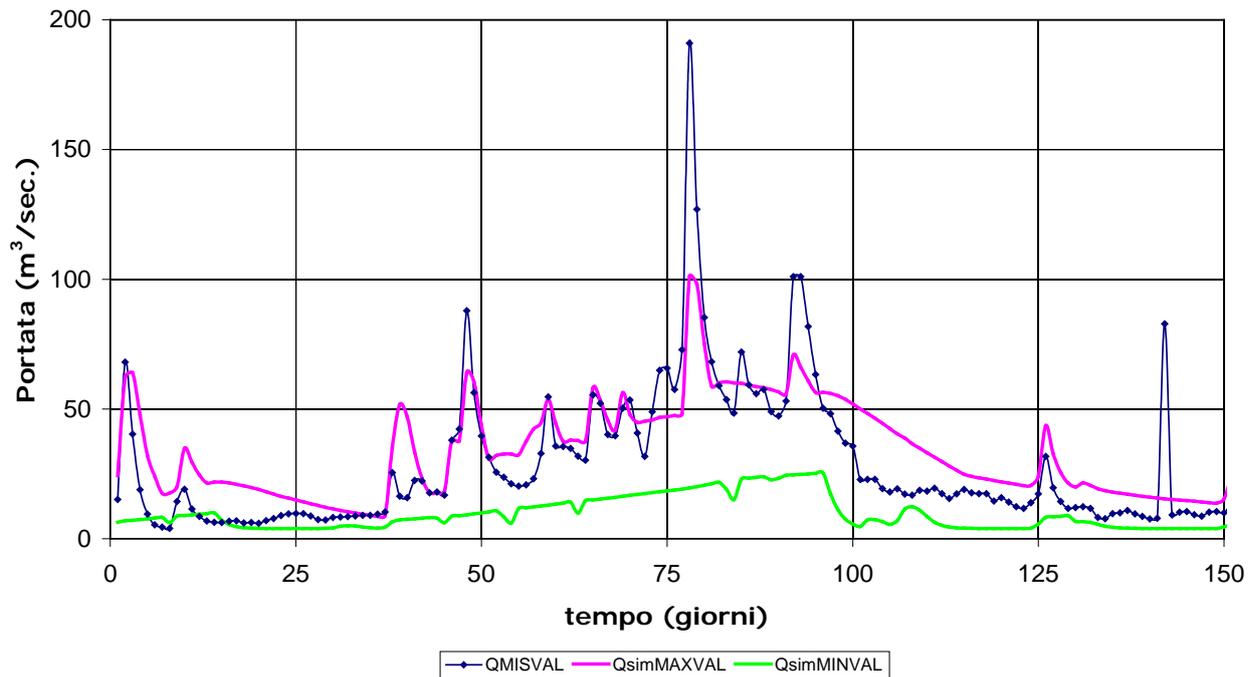


Figura 5.37. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

Per quanto riguarda il Fiume Metauro, le considerazioni che possono essere fatte in merito agli indici di qualità ottenuti con le due combinazioni *multi-obiettivo* (Tabella 5.24 e Tabella 5.25), sono molto simili a quelle fatte per il Fiume Esino.

METAURO (calibrazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,72	0,71
MAE	6,7	6,71
MRE	0,46	0,44

Tabella 5.24. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

I valori degli indici di qualità sono cioè sostanzialmente paragonabili e solo in considerazione dei valori ottenuti in validazione si preferisce la combinazione MAE-MSE-MRE.

In dettaglio i valori di EFF in calibrazione sono molto buoni mentre i valori di MAE in validazione rimangono anche in questa modalità molto elevati. L'idrogramma simulato medio in calibrazione (Figura 5.38) è decisamente molto simile a quello osservato e

rispetto alla modalità *mono-obiettivo* porta ad una migliore riproduzione dei valori di portata estrema. In validazione può essere fatta la stessa considerazione anche se è innegabile la presenza di alcuni eventi estremi mal riprodotti (Figura 5.39).

METAURO (validazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,6	0,58
MAE	8,43	8,53
MRE	0,57	0,56

Tabella 5.25. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

L'osservazione degli inviluppi conferma il fatto che il modello sia adatto alla simulazione dei processi in questo bacino (Figura 5.40 e Figura 5.41): essi infatti si presentano molto vicini e con andamento sostanzialmente uguale a quello dell'idrogramma osservato. Esiste un unico evento anomalo verso la fine del periodo di calibrazione che viene, anche in questa modalità, decisamente sovrastimato dal modello. In validazione infine il modello sembra dare ottimi risultati anche se permane una sovrastima delle portate che, essendo sostanzialmente diffusa a tutto il periodo, in generale produce i valori di MAE elevati già sottolineati in precedenza (Figura 5.41). Per quanto riguarda il Fiume Potenza gli indici relativi alla portata media in calibrazione sembrano indicare che la combinazione migliore di funzioni obiettivo è quella data da MAE-MSE-MRE (migliore per EFF e MAE) anche se le due modalità si dimostrano tutto sommato equivalenti (Tabella 5.26).

L'analisi degli indici in validazione conferma quanto appena sostenuto poiché le differenze nei valori degli indici sono pressoché trascurabili (Tabella 5.27).

L'idrogramma simulato medio in calibrazione è decisamente soddisfacente nella riproduzione delle portate di magra e meno soddisfacente nella riproduzione dei picchi (Figura 5.42).

In validazione l'idrogramma simulato medio ha un andamento molto simile a quello osservato anche se è evidente una sovrastima dei processi di magra e una sottostima dei picchi di piena (Figura 5.43).

Fiume Metauro - Calibrazione (MAE-MSE-MRE)

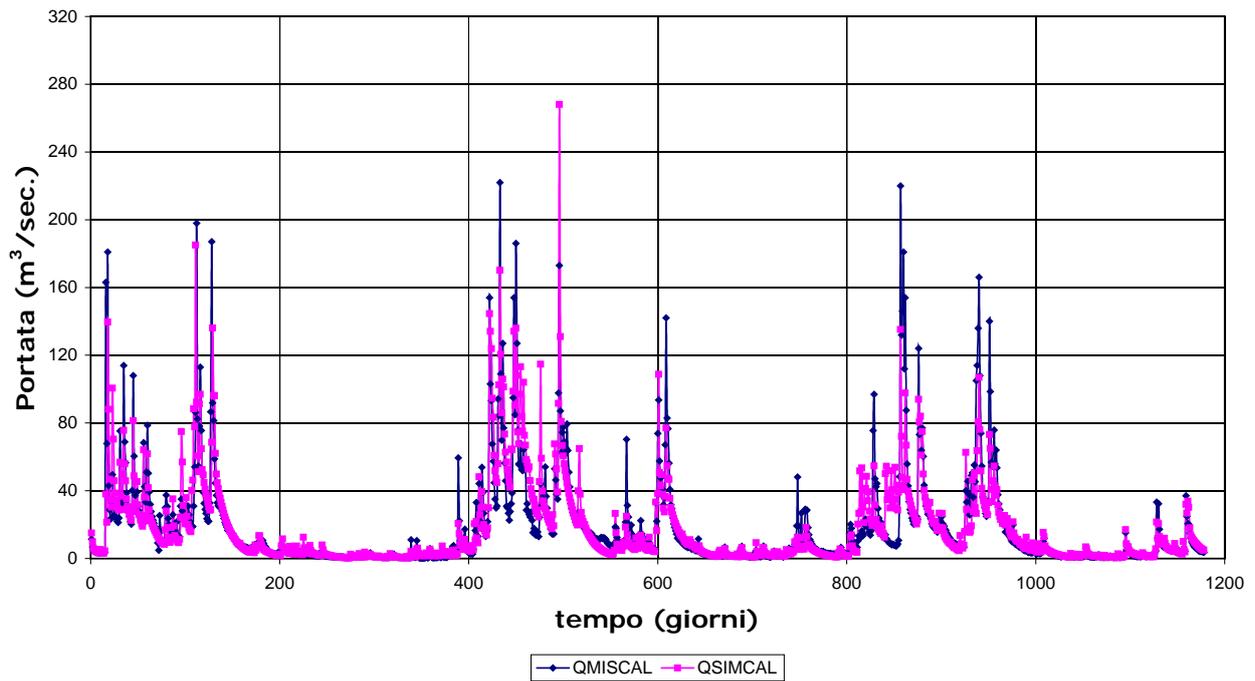


Figura 5.38. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (MAE-MSE-MRE).

Fiume Metauro - Validazione (MAE-MSE-MRE)

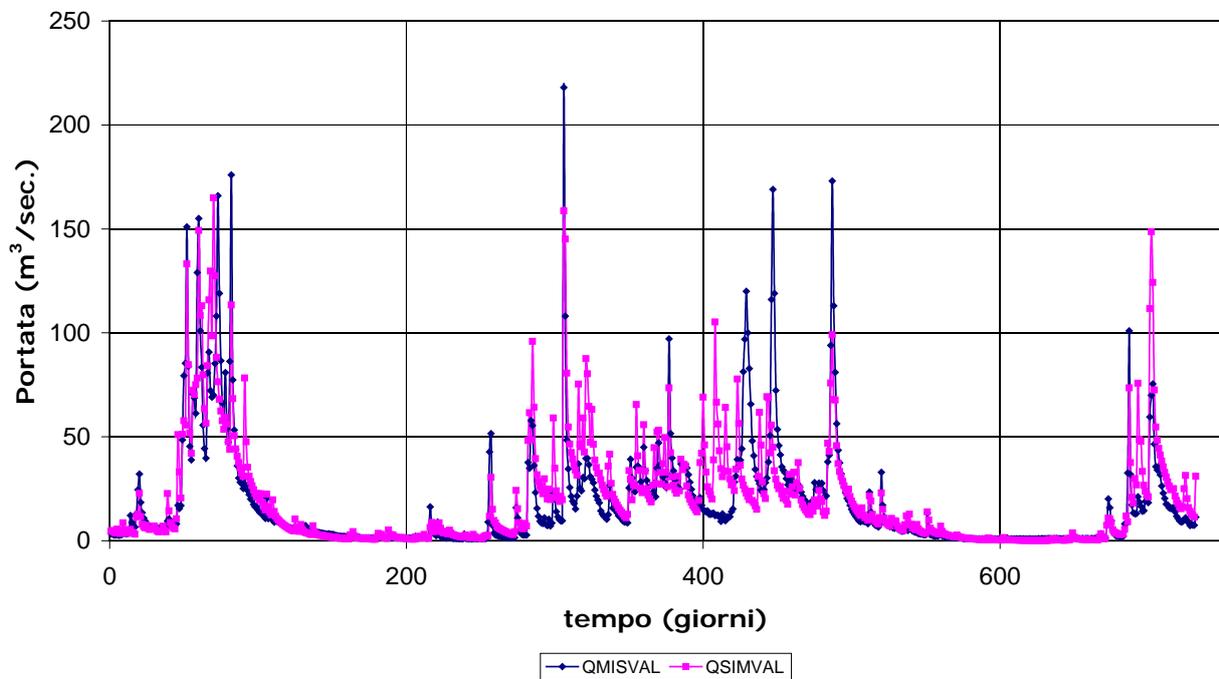


Figura 5.39. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (MAE-MSE-MRE).

Fiume Metauro - Calibrazione (MAE-MSE-MRE)

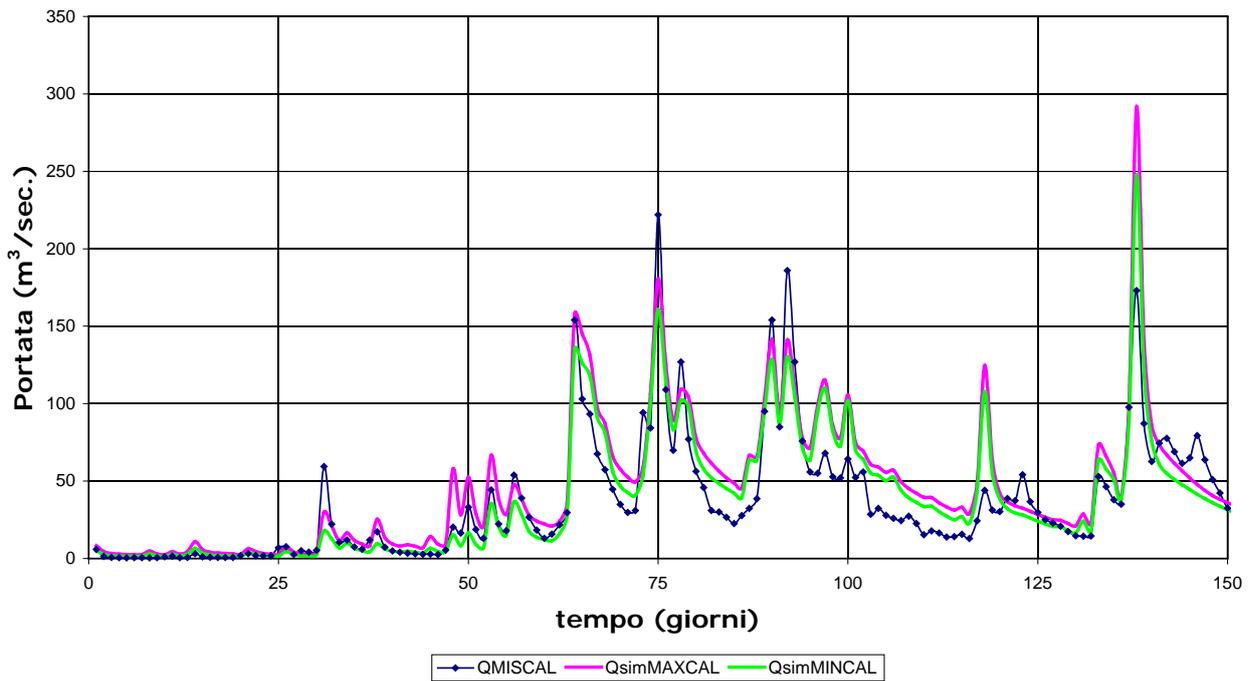


Figura 5.40. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

Fiume Metauro - Validazione (MAE-MSE-MRE)

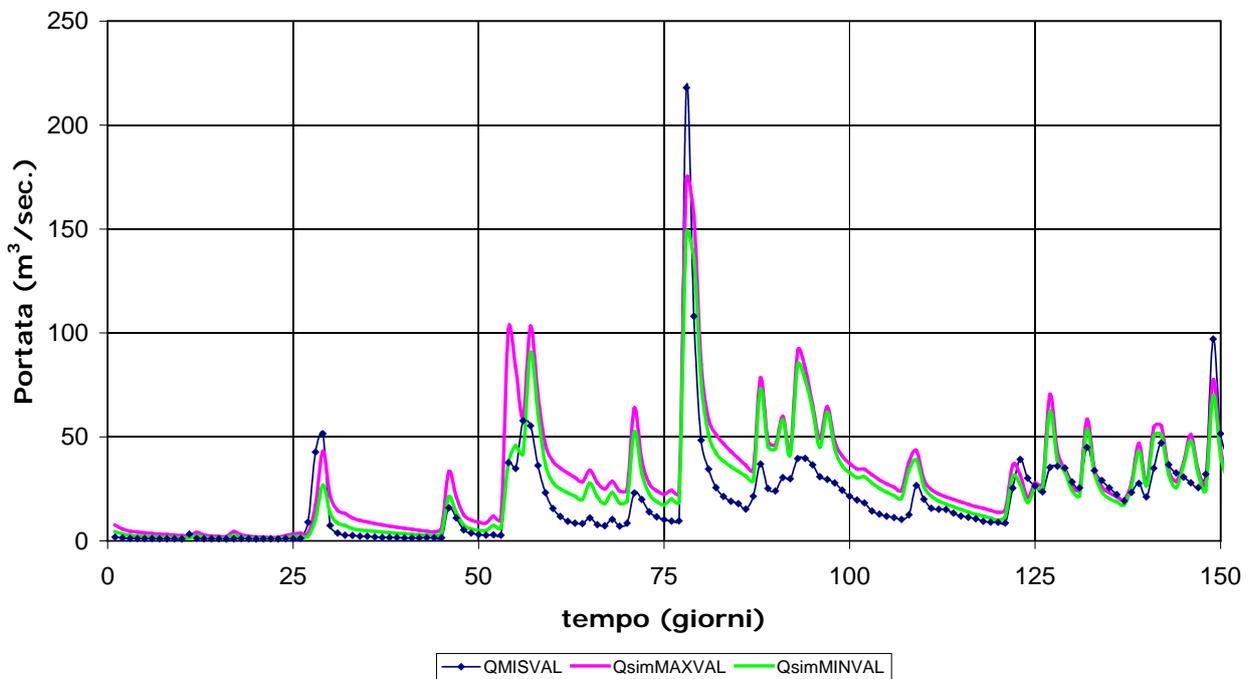


Figura 5.41. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

POTENZA (calibrazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,75	0,7
MAE	1,68	1,72
MRE	0,23	0,22

Tabella 5.26. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

POTENZA (validazione)		
Funzione Obiettivo	MAE-MSE-MRE	MAE-MSLE-MRE
EFF	0,65	0,64
MAE	1,94	1,91
MRE	0,38	0,35

Tabella 5.27. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).

Fiume Potenza - Calibrazione (MAE-MSE-MRE)

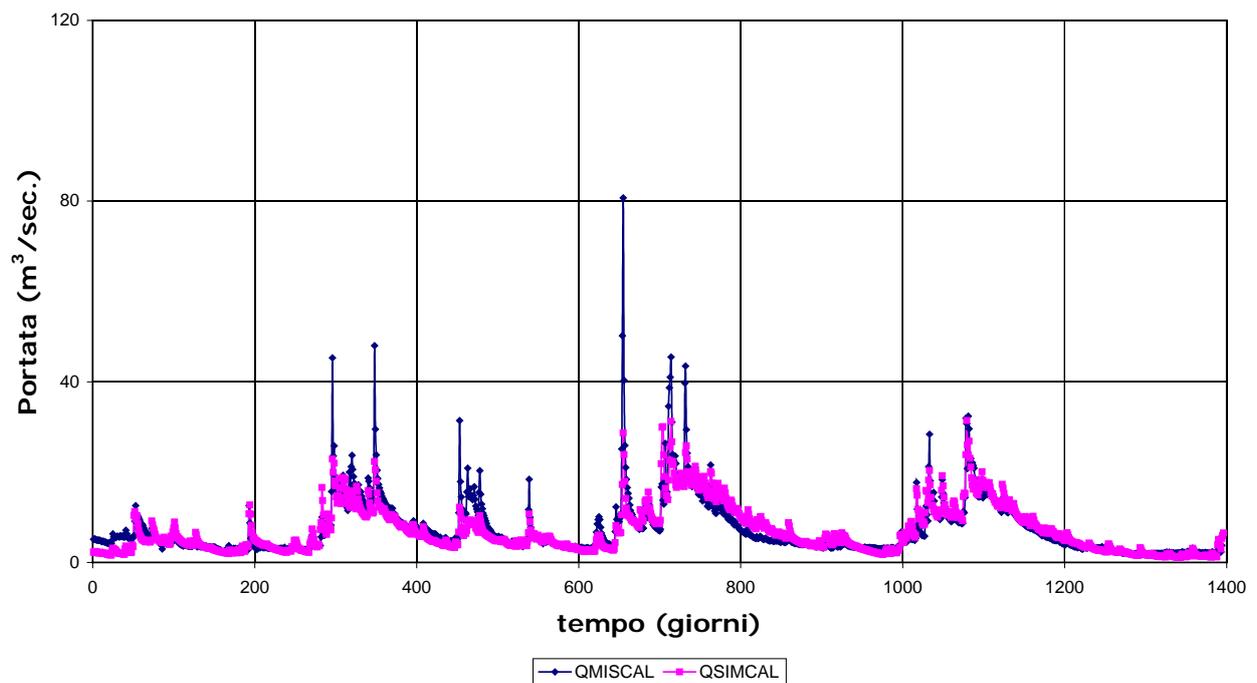


Figura 5.42. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro (MAE-MSE-MRE).

Fiume Potenza - Validazione (MAE-MSE-MRE)

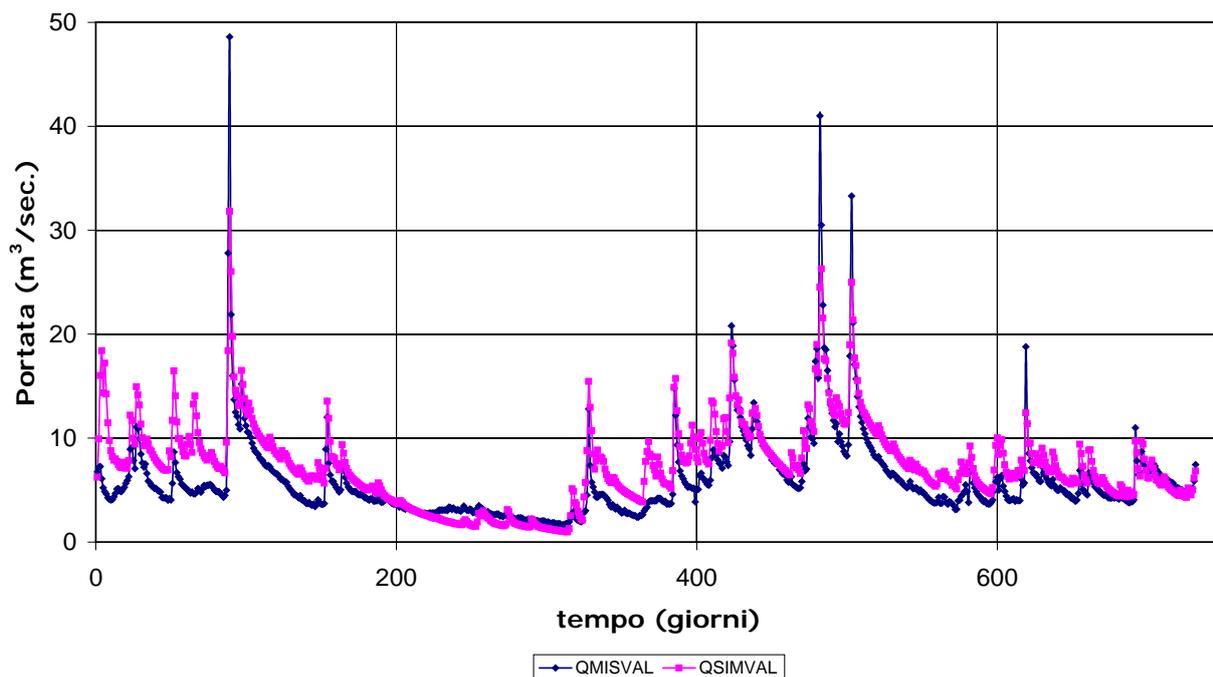


Figura 5.43. Idrogramma in validazione del Fiume Potenza a Cannucciaro (MAE-MSE-MRE).

Fiume Potenza - Calibrazione (MAE-MSE-MRE)

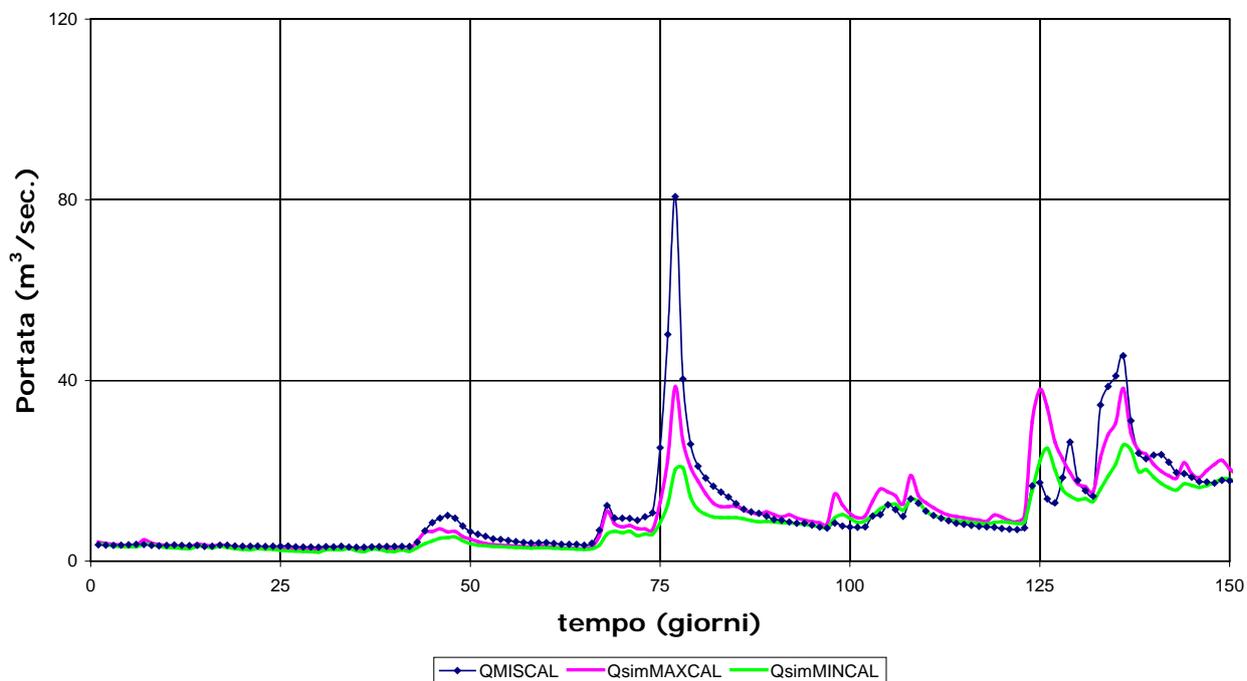


Figura 5.44. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

Fiume Potenza - Validazione (MAE-MSE-MRE)

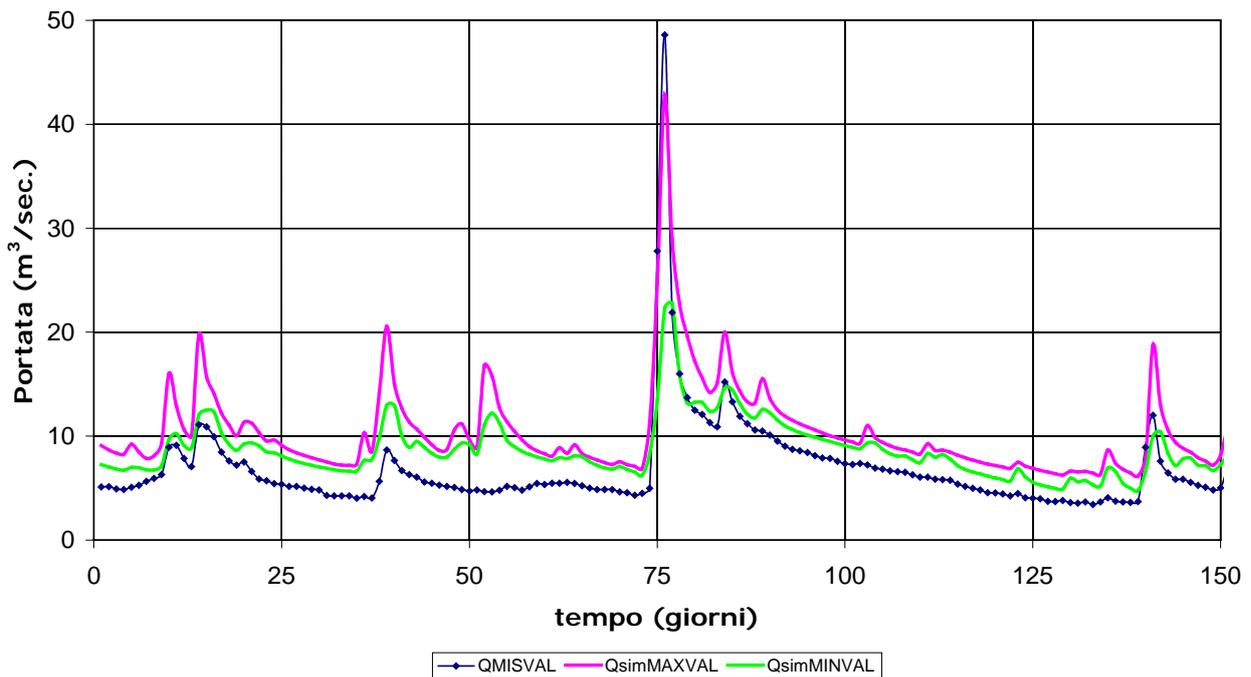


Figura 5.45. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).

Gli inviluppi massimo e minimo in calibrazione dimostrano un buon andamento anche se il picco è stimato in misura circa pari alla metà rispetto alla sua vera entità (Figura 5.44).

In validazione invece c'è una evidente sovrastima delle portate di magra, ma l'inviluppo massimo si avvicina in modo soddisfacente alla riproduzione del picco di piena sia in termini di entità che di sincronia (Figura 5.45).

5.4 Conclusioni

Le simulazioni realizzate con le modalità *mono* e *multi-obiettivo* hanno permesso di analizzare il comportamento del modello Hymod e dell'algoritmo di ottimizzazione AMALGAM in relazione all'area di studio presa in considerazione.

A titolo di confronto si riportano i valori degli indici di prestazione già analizzati in precedenza ma organizzati in modo da poter evidenziare in modo diretto le differenze tra le varie modalità di simulazione indagate.

In Tabella 5.28 sono riportati tutti i valori di EFF ottenuti per i 4 bacini con le varie modalità di simulazione in calibrazione mentre in Tabella 5.29 sono riportati gli stessi

valori in validazione. Successivamente la Tabella 5.30, la Tabella 5.31, la Tabella 5.32 e la Tabella 5.33 riportano con lo stesso criterio i valori di MAE e MRE.

I risultati ottenuti in modalità *mono-obiettivo* sono soddisfacenti in termini di indici di qualità per tutti bacini. Fa eccezione la modalità BIAS che, analizzando i valori di EFF prodotti sia in calibrazione che in validazione, dimostra di non essere adatta alla simulazione dei processi per il Fiume Candigliano e Potenza. E' da sottolineare che se si fossero analizzati solo i valori di MAE ed MRE come indici di qualità, questa considerazione non sarebbe stata altrettanto evidente e di conseguenza possiamo concludere che anche in termini di valutazione di una procedura di calibrazione non è corretto considerare un solo aspetto dell'idrogramma ottenuto piuttosto che approfondire una valutazione globale del processo.

Le calibrazioni *mono-obiettivo* si dimostrano, come ci si aspetterebbe, migliori o peggiori in funzione dell'indice preso in esame per analizzarle: se si considera l'EFF è ovviamente la modalità di calibrazione con l'EFF quella con prestazioni migliori, se invece si considera il MAE, è la modalità che utilizza questo indice come funzione *mono-obiettivo* a produrre le prestazioni migliori; analoghe considerazioni possono essere fatte per il MRE.

Lo studio dei risultati ottenuti in modalità *multi-obiettivo* consente invece un'analisi globale del processo afflussi-deflussi: le funzioni obiettivo agiscono in questo caso contemporaneamente nella ricerca dei valori dei parametri e conseguentemente il loro contributo è tenuto in considerazione dal processo stesso. La combinazione migliore in termini di EFF sembra essere RMSE-BIAS mentre se si considerano gli indici MAE e MRE la scelta si sposta sulla combinazione MAE-MRE.

EFF (calibrazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	0,24	0,74	0,68	0,68	0,74	0,7	0,68	0,72	0,7	0,72	0,7	0,71	0,7
Esino	0,48	0,58	0,56	0,08	0,58	0,49	0,36	0,57	0,55	0,41	0,58	0,42	0,38
Metauro	0,65	0,72	0,71	0,71	0,72	0,72	0,71	0,72	0,72	0,72	0,72	0,72	0,71
Potenza	0,58	0,78	0,73	0,5	0,78	0,78	0,67	0,77	0,72	0,75	0,78	0,75	0,7

Tabella 5.28. Valori di EFF in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

EFF (calibrazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	0,09	0,65	0,57	0,57	0,65	0,63	0,57	0,63	0,58	0,62	0,52	0,62	0,57
Esino	0,73	0,71	0,69	0,11	0,71	0,61	0,43	0,71	0,67	0,52	0,7	0,53	0,48
Metauro	0,59	0,6	0,59	0,57	0,6	0,6	0,58	0,6	0,59	0,59	0,6	0,6	0,58
Potenza	-0,28	0,52	0,57	0,31	0,52	0,58	0,59	0,63	0,62	0,66	0,57	0,65	0,64

Tabella 5.29. Valori di EFF in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

MAE (calibrazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	6,11	3,33	3,01	3,04	3,33	3,27	3,03	3,09	3,07	3,14	3,84	3,1	3,07
Esino	7,78	6,98	6,54	8,28	6,98	7,58	7,63	6,67	6,64	7,67	6,97	7,6	7,52
Metauro	7,61	6,95	6,64	6,7	6,95	6,82	6,66	6,72	6,71	6,74	6,84	6,7	6,71
Potenza	2,77	1,77	1,66	1,86	1,77	1,77	1,74	1,69	1,68	1,69	1,77	1,68	1,72

Tabella 5.30. Valori di MAE in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

MAE (validazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	4,92	2,73	2,31	2,21	2,73	2,62	2,26	2,39	2,33	2,33	3,13	2,33	2,3
Esino	5,5	5,43	5,28	7,7	5,43	6,39	6,83	5,33	5,62	6,69	5,61	6,65	6,67
Metauro	8,81	8,8	8,32	8,51	8,8	8,58	8,37	8,44	8,53	8,57	8,62	8,43	8,53
Potenza	3,31	2,12	1,95	1,86	2,12	2,2	1,92	2,01	1,95	1,92	2,21	1,94	1,91

Tabella 5.31. Valori di MAE in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

MRE (calibrazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	1,92	0,61	0,42	0,41	0,61	0,55	0,41	0,45	0,42	0,43	0,89	0,43	0,42
Esino	0,59	0,51	0,44	0,35	0,51	0,52	0,38	0,47	0,46	0,39	0,5	0,39	0,39
Metauro	0,73	0,69	0,44	0,44	0,69	0,59	0,44	0,51	0,44	0,46	0,62	0,46	0,44
Potenza	0,47	0,26	0,22	0,2	0,26	0,26	0,21	0,23	0,22	0,23	0,26	0,23	0,22

Tabella 5.32. Valori di MRE in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

MRE (validazione)	FUNZIONI OBIETTIVO												
Fiume	BIAS	EFF	MAE	MRE	RMSE	RMSE D_ND	MAE MRE	MAE MSE	MAE MSLE	MRE MSE	RMSE BIAS	MAE MSE MRE	MAE MSLE MRE
Candigliano	1,53	0,62	0,43	0,4	0,62	0,56	0,41	0,47	0,44	0,43	0,81	0,44	0,43
Esino	0,51	0,41	0,41	0,37	0,41	0,47	0,39	0,41	0,46	0,4	0,44	0,4	0,4
Metauro	0,74	0,73	0,54	0,55	0,73	0,66	0,54	0,6	0,56	0,58	0,68	0,57	0,56
Potenza	0,66	0,42	0,37	0,28	0,42	0,43	0,34	0,39	0,37	0,38	0,43	0,38	0,35

Tabella 5.33. Valori di MRE in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.

Confrontando direttamente i grafici della modalità *mono-obiettivo* con quelli della modalità *multi-obiettivo*, è possibile confermare che anche se i singoli indici possono peggiorare leggermente tenendo conto di entrambe le funzioni obiettivo, la simulazione del modello migliora (seppur in modo limitato). Questo è legato al fatto che diminuisce l'errore che nel complesso si compie nella simulazione dei processi di magra e di piena nonché nella ricerca di un maggior sincronismo degli eventi. Per il Fiume Potenza la simulazione *multi-obiettivo* è migliore anche in termini di singoli indici di quelle ottenute con la modalità *mono-obiettivo*.

Per il Fiume Esino, nonostante i valori degli indici siano discreti, l'analisi degli idrogrammi dimostra che il modello non sembra ideale per la riproduzione del processo afflussi-deflussi per lo meno in fase di calibrazione. Utilizzando tre funzioni obiettivo la riproduzione grafica dell'idrogramma migliora ulteriormente mentre gli indici di qualità si confermano pressoché invariati e comunque decisamente soddisfacenti. Questo ci porta a concludere che :

- la scelta di una calibrazione *multi-obiettivo* è nel complesso più oggettiva di una modalità *mono-obiettivo* perché consente di non dover privilegiare un aspetto del processo che si vuole simulare piuttosto che un altro;
- la calibrazione *multi-obiettivo* è più semplice da valutare poiché tenendo sin dall'inizio conto di più aspetti del fenomeno in gioco, in fase di valutazione della qualità della simulazione non si rischia di tener conto nuovamente solo dello stesso aspetto e di trascurarne altri;
- la calibrazione *multi-obiettivo* si è dimostrata infine altrettanto semplice da implementare grazie all'utilizzo di AMALGAM, che consente di ottenere i valori dei parametri senza incremento eccessivo dei tempi di analisi e delle risorse necessarie.

6 La calibrazione dei modelli afflussi-deflussi in bacini non strumentati

E' stato già sottolineato in precedenza che la calibrazione di un modello idrologico viene realizzata attraverso il confronto fra i valori di portata simulati ed i corrispondenti valori osservati, ricercando i parametri che rendono tale confronto ottimale. Tuttavia, qualora si debba far fronte all'incompletezza e alla carenza dei dati idrometrici, può essere necessario ricorrere a metodi indiretti per il raggiungimento di tale stima.

Negli ultimi anni numerosi contributi in ambito scientifico hanno portato all'approfondimento di tecniche di regionalizzazione dei parametri dei modelli afflussi-deflussi (*Seibert, 1999; Merz & Blöschl, 2004; Wagener & Wheater, 2006; Bardossy, 2007; Hundecha et al., 2008*). In particolare, la procedura comune a tali metodi si articola in due fasi: la prima prevede una calibrazione che presuppone di poter considerare i bacini come strumentati e porta alla determinazione per ciascun bacino di un set di parametri ottimale. A partire poi da tali parametri, vengono messi a punto dei modelli *multiregressivi* per la stima degli stessi in funzione, questa volta, delle caratteristiche geomorfologiche e climatiche dei bacini. Nella seconda fase invece si riconsiderano i bacini come non strumentati e si sfruttano le relazioni precedentemente ottenute per stimare i parametri del modello afflussi-deflussi.

I metodi di regionalizzazione dei parametri presentano tuttavia alcune problematiche che non possono essere omesse: innanzi tutto, come già ampiamente dibattuto in precedenza, non sembra opportuno parlare di set ottimale dei parametri ma di insieme di set ugualmente accettabili e possibili in funzione della complessità del fenomeno che si vuole riprodurre. Inoltre è senza dubbio vero che può non essere semplice trovare la relazione fra i parametri del modello e le caratteristiche fisiche di ciascun bacino. Infine, la diversa disponibilità di dati e le differenti caratteristiche idrologiche dei bacini, suggeriscono in alcuni casi l'utilizzo di modelli afflussi-deflussi di natura differente a seconda dei casi (*Wagener & Wheater, 2006*). In sintesi si può

affermare che l'applicazione di tali procedure è condizionata alla difficoltà di giungere ad una stima attendibile dei parametri dei modelli che a loro volta possono rendere poco affidabili i risultati delle simulazioni eseguite mediante il modello afflussi-deflussi. I valori utilizzati nel presente lavoro sono stati ottenuti dall'Ing. Castiglioni durante il Dottorato di Ricerca svolto presso il DICAM di Bologna (*Castiglioni, 2010*). Il concetto alla base della metodologia di regionalizzazione dei parametri utilizzata nel suo lavoro e che ha portato alla determinazione dei valori utilizzati per il raggiungimento dei risultati presentati in questa trattazione, è quello di consentire la modellazione afflussi-deflussi in bacini non strumentati sfruttando le informazioni disponibili a scala regionale attraverso una procedura parsimoniosa e flessibile, la cui complessità varia a seconda della disponibilità di dati. In tale contesto le informazioni regionali rappresentano una valida opportunità e la flessibilità della struttura del modello, un elemento essenziale. L'idea iniziale era quella di ottenere delle relazioni regionali che fornissero informazioni quantitative circa il regime dei deflussi fluviali (come ad esempio le statistiche delle serie delle portate) e di effettuare la calibrazione del modello afflussi-deflussi determinando quei parametri che consentono di riprodurre al meglio le grandezze caratteristiche dei deflussi.

La letteratura propone diversi approcci per la regionalizzazione delle statistiche delle serie fluviali (*Stedinger et al., 1993; Hosking & Wallis, 1997; Pandey & Nguyen, 1999; FEH, 1999*) e presenta applicazioni a problemi e contesti idrologici differenti come la stima delle portate di piena (v. ad es. *Dalrymple, 1960; Burn, 1990; Castellarin et al., 2001; Merz & Blöschl, 2005*) e l'analisi della frequenza delle portate di magra (v. ad es. *Vogel & Kroll, 1992; Furey & Gupta, 2000; Smakthin, 2001; Castellarin et al., 2004b; Laaha & Blöschl, 2006*).

La scelta delle statistiche da utilizzare come indicatori di riferimento relativi ai deflussi fluviali è stata basata su quanto emerso dallo studio condotto da *Montanari & Toth, 2007* dal quale si evince come la calibrazione spettrale di modelli afflussi-deflussi costituisca un'interessante opportunità per la definizione dei parametri dei modelli in bacini non strumentati. In questo studio gli autori hanno fatto riferimento ad una tecnica di calibrazione basata sull'approssimazione della funzione di verosimiglianza Gaussiana introdotta da *Whittle, 1953*. In dettaglio viene proposta una calibrazione effettuata minimizzando una funzione obiettivo che si basa sul confronto fra la funzione di densità spettrale dei dati simulati dal modello ed il *periodogramma* della serie temporale dei corrispondenti dati osservati. Il *periodogramma* però deve essere a sua volta stimato ed è stato dimostrato che può essere completamente definito in funzione della deviazione standard e del correlogramma delle portate fluviali

(*Brockwell & Davis, 1987*). Di recente inoltre *Gupta et al., 2009* hanno mostrato come l'errore quadratico medio, che viene generalmente utilizzato nella fase di calibrazione dei modelli afflussi-deflussi per il confronto delle serie simulate con quelle osservate, possa essere rappresentato mediante un'opportuna combinazione dei momenti del primo e del secondo ordine delle serie dei deflussi simulati ed osservati (media e deviazione standard), e del coefficiente di cross-correlazione calcolato fra valori corrispondenti di portate simulate e osservate. Quest'ultimo consente di avere indicazioni relative, per esempio, alla sincronia fra i picchi di portata simulati ed osservati, agli andamenti delle portate propri delle fasi di formazione ed esaurimento delle piene, etc. etc..

Questi elementi, insieme ad altri aspetti dell'idrogramma che sono caratteristici delle serie di portate fluviali, possono essere appunto espresse mediante la struttura di correlazione e pertanto un'ulteriore semplificazione può essere quella di considerare l'aspetto dell'autocorrelazione delle portate (*Gupta et al., 2009*).

Alla luce di queste considerazioni era stata proposta una procedura di calibrazione di modelli afflussi-deflussi in bacini non strumentati basata sull'utilizzo delle seguenti caratteristiche statistiche delle serie di portata:

- il loro valore medio μ ,
- la loro varianza σ^2 ,
- il coefficiente di autocorrelazione a ritardo unitario, ρ_1 .

Tali grandezze consentono di fornire importanti informazioni circa l'entità e la variabilità dei deflussi fluviali e circa la struttura di correlazione delle serie, presentando il vantaggio di poter essere efficacemente stimate mediante modelli *multiregressivi* in funzione delle caratteristiche geomorfologiche e climatiche di ciascun bacino. Pertanto selezionando un'opportuna funzione obiettivo dipendente dalle caratteristiche statistiche considerate, è possibile calibrare il modello afflussi-deflussi anche in bacini privi di informazioni idrometriche attraverso il confronto realizzato fra le statistiche delle serie simulate con quelle opportunamente regionalizzate mediante modelli *multiregressivi*.

In seguito vengono riportati due paragrafi nei quali sono richiamati i principi su cui si basa l'approssimazione della funzione di massima verosimiglianza proposta da Whittle. Successivamente è presentato un paragrafo relativo agli indici di magra poiché, sempre a partire dallo studio dell'Ing. Castiglioni, è stata approfondita l'ipotesi di introdurre anche un indice di riproduzione dei bassi deflussi tra le statistiche utilizzate per la calibrazione del modello Hymod nel caso di bacini non strumentati. Tali

approfondimenti hanno il compito di mostrare le motivazioni che hanno portato alla scelta delle citate statistiche.

6.1 L'approssimazione di Whittle della funzione di massima verosimiglianza Gaussiana

L'approssimazione proposta da *Whittle (1953)* della funzione di verosimiglianza Gaussiana si basa sul calcolo di una misura di verosimiglianza, di seguito indicata come $W(\theta)$ con θ vettore dei parametri del modello idrologico, intesa fra una serie temporale stazionaria osservata e la corrispondente simulata dal modello ed e calcolata nel dominio spettrale (*Montanari & Toth, 2006 - Lombardi et al., 2010a*). Riferendosi al caso dei modelli afflussi-deflussi, al fine di introdurre $W(\theta)$, l'approssimazione operata dal modello idrologico della serie temporale di portata fluviale osservata $Q_{obs}(t)$ si può descrivere matematicamente nella forma:

$$Q_{obs}(t) = M[\theta, I(t)] + e(t) = M[\theta, I(t)] + \Phi^{-1}(B)\varepsilon(t), \quad (6.1)$$

dove $M[\theta, I(t)]$ è la trasformazione operata dal modello stesso; $I(t)$ è il vettore delle variabili in ingresso (ad esempio precipitazione e temperature, oppure evapotraspirazione all'istante t); $e(t)$ è la serie temporale dei residui del modello (il cui valor medio è ipotizzato nullo); $\Phi(B)$ è un operatore auto-regressivo (*Brockwell et al. 1987*) introdotto al fine di tenere conto dell'autocorrelazione di $e(t)$ e la cui struttura è condizionata dal modello idrologico scelto; infine $\varepsilon(t)$ rappresenta una variabile decorrelata a media nulla (rumore bianco).

La misura di verosimiglianza di Whittle in riferimento al modello descritto nell'equazione precedente può essere calcolata con la seguente espressione:

$$L(\theta) = \exp \left[- \sum_{j=1}^{\frac{N}{2}} \left\{ \log f(\lambda_j, \theta) + \frac{J(\lambda_j)}{f_M(\lambda_j, \theta) + f_e(\lambda_j, \Phi)} \right\} \right] \quad (6.2)$$

con $-\pi \leq \lambda_j \leq \pi$ e dove

- λ_j indica la frequenza di Fourier;

- J indica il *periodogramma* della serie temporale osservata;
- f_M è la funzione di densità spettrale del modello idrologico che dipende dal vettore dei parametri θ ;
- f_e è la funzione di densità spettrale dell'operatore auto regressivo Φ che può essere così determinata:

$$f_e(\lambda_j, \Phi) = \frac{\sigma_t^2}{2\pi |\Phi(e^{-i\lambda_j})|^2} \quad (6.3)$$

dove σ_t^2 è la deviazione standard di $\varepsilon(t)$.

La calibrazione del modello si effettua quindi minimizzando la quantità:

$$W(\theta) = \sum_{j=1}^{\frac{N}{2}} \left\{ \log f(\lambda_j, \theta) + \frac{J(\lambda_j)}{f_M(\lambda_j, \theta) + f_e(\lambda_j, \Phi)} \right\}. \quad (6.4)$$

Qualora siano disponibili serie contemporanee di dati in ingresso e di misure di portata per un periodo di tempo $[0, T]$ di estensione sufficiente, lo stimatore dato dalla precedente eq. può essere utilizzato per operare una calibrazione del modello di tipo classico, ottimizzando quindi la riproduzione delle variabili osservate. Questa procedura può essere indicata con il termine "calibrazione diretta". In tal caso, la funzione di densità spettrale J del campione osservato può essere stimato calcolando il *periodogramma* delle osservazioni idrometriche disponibili per il periodo $[0, T]$. Inoltre, una volta individuato un valore di tentativo per il vettore dei parametri θ , la funzione di densità spettrale del modello, $f_M(\lambda_j, \theta)$, può essere stimata calcolando il *periodogramma* della simulazione effettuata sul medesimo periodo $[0, T]$ fornendo in ingresso al modello le variabili meteorologiche osservate. Infine, i residui del modello nel periodo $[0, T]$ possono essere calcolati, così da poter individuare l'operatore auto regressivo $\Phi(B)$ e stimare i parametri. Durante l'operazione di calibrazione devono essere esclusi a priori valori di tentativo di θ che comportino differenze significative $\Delta\mu_{obs, sim}$ tra le medie delle portate osservate e simulate, poiché in tal caso è violata l'ipotesi di media nulla per $\varepsilon(t)$. A rigore, occorre escludere a priori i valori di θ che comportino differenze significative.

Qualora non siano disponibili serie storiche contemporanee di dati in ingresso e di portata fluviale, ossia nel caso di bacini non strumentati o scarsamente strumentati, la densità spettrale così come il *periodogramma* degli errori delle simulazioni non è noto e pertanto lo si può assumere identicamente nullo. In tali condizioni, assicurandosi che la media delle portate simulate sia riprodotta in maniera soddisfacente, la funzione obiettivo che approssima la funzione di massima verosimiglianza di Whittle risulta essere (Montanari & Toth, 2007):

$$W(\theta) = \sum_{j=1}^N \left\{ \log[J_M(\lambda_j, \theta)] + \frac{J(\lambda_j)}{J_M(\lambda_j, \theta)} \right\} \quad (6.5)$$

dove $J(\lambda_j)$ è il *periodogramma* delle serie di portate osservate che è necessario stimare attraverso metodi indiretti.

In questo modo la calibrazione di un modello afflussi – deflussi è possibile anche in assenza di osservazioni idrometriche, purché siano note le variabili in ingresso al modello stesso. Tale assunzione risulta solitamente accettabile in quanto i dati meteorologici e le informazioni a scala di bacino sono solitamente disponibili.

Alla luce delle considerazioni esposte, la calibrazione del modello afflussi – deflussi attraverso la minimizzazione della funzione 6.5 implica che la calibrazione sia effettuata essenzialmente attraverso un confronto fra la media ed il *periodogramma* delle serie delle portate simulate e quelle osservate. Assunto tale presupposto, ciascuna serie può essere scomposta in una somma di componenti periodiche attraverso l'analisi armonica ed il *periodogramma*.

Il passo successivo consiste nel notare che il *periodogramma* può essere derivato dalla funzione di *autocovarianza* dei dati, portandoci ad affermare che la funzione di verosimiglianza di Whittle consente di calibrare un modello afflussi – deflussi sostanzialmente attraverso la determinazione di quei parametri che consentono di ottenere delle serie fluviali simulate aventi *media* e funzione di *autocovarianza* il più possibile simile a quello dei corrispondenti valori osservati.

Riassumendo possiamo dire che per la calibrazione di un modello afflussi – deflussi in bacini non strumentati attraverso l'utilizzo della funzione di Whittle, è necessario ricorrere alle stime di *media* e del *periodogramma* in assenza di osservazioni ma, qualora si volesse ricorrere all'applicazione di metodi regionali, il *periodogramma* non potrebbe essere stimato riferendosi direttamente al dominio spettrale, poiché sarebbe impossibile dedurre la variabilità delle portate fluviali spiegate da ciascuna frequenza di Fourier. Una stima regionale del *periodogramma* è tuttavia possibile nel dominio

temporale stimando, attraverso metodi regionali, la *media*, la *varianza* e la *struttura di correlazione* delle portate fluviali medesime.

Se da un lato esiste una vasta letteratura dedicata alla stima regionale di media e varianza, la regionalizzazione della struttura di correlazione non è stata ancora trattata in modo diffuso. Per quest'ultima dunque, una possibile soluzione è quella di fittare la funzione di *autocorrelazione* attraverso una funzione parametrica ed in seguito regionalizzare i relativi parametri. A tal fine nello studio dell'Ing. Castiglioni al quale si è fatto riferimento, si è assunto che i deflussi fluviali possano essere rappresentati attraverso un processo *autoregressivo* del primo ordine *AR1* (*Brockwell e Davis, 1987*). È evidente la limitazione di quest'ultima semplificazione che è stata introdotta per consentire una stima più robusta delle proprietà di *autocorrelazione* dei deflussi fluviali e che ad ogni modo conduce ad una buona riproduzione dei fenomeni di breve memoria. Tuttavia tale semplificazione sembra accettabile in riferimento al fatto che la stima è riferita al caso di bacini non strumentati.

La regionalizzazione più estesa della struttura di correlazione potrebbe essere oggetto di un possibile studio futuro.

In definitiva alla luce delle assunzioni fatte, la calibrazione di modelli afflussi – deflussi in bacini non strumentati può essere eseguita ricorrendo al confronto della media μ , della varianza σ^2 e del coefficiente di autocorrelazione a ritardo unitario ρ_1 delle serie di deflussi fluviali simulate e i corrispondenti valori stimati mediante opportuni modelli regionali in funzione delle caratteristiche geomorfologiche e climatiche di ciascun bacino (*Castiglioni et al., 2010b*).

6.1.1 Regionalizzazione delle grandezze statistiche

Per approfondimenti in merito ai modelli *multiregressivi* utilizzati al fine dell'individuazione delle eq. matematiche che hanno consentito la stima delle tre statistiche appena definite, si rimanda alla Tesi di Dottorato dell'Ing. S. Castiglioni dal titolo "Modelli per la stima delle risorse idriche superficiali in bacini idrografici non strumentati". Tuttavia si ritiene opportuno richiamare alcuni passaggi dello studio citato al fine di aiutare la comprensione del percorso che ha portato al raggiungimento dei risultati mostrati nel presente lavoro.

I modelli multi regressivi considerati sono stati i tre seguenti:

$$\gamma(Q) = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n \quad (6.6)$$

$$\gamma(Q) = a_0 + a_1 \log(x_1) + a_2 \log(x_2) + \dots + a_n \log(x_n) \quad (6.7)$$

$$\gamma(Q) = a_0 x_1^{a_1} x_2^{a_2} x_3^{a_3} \dots x_n^{a_n} \quad (6.8)$$

dove

- $\gamma(Q)$ rappresenta la media di lungo periodo, la varianza, o il coefficiente di autocorrelazione a ritardo di un giorno (ovvero μ , σ^2 e ρ_1),
- x_i , con $i = 1, 2, \dots, N$, indica l' i -esima variabile esplicativa del modello *multiregressivo* (ovvero gli indici geomorfologici e climatici) e
- a_i , con $i = 0, 1, 2, \dots, k$ indica i coefficienti della regressione.

Il numero ottimale delle variabili esplicative e i relativi coefficienti sono stati stimati attraverso una procedura *stepwise*. La procedura è stata applicata per ciascun modello *multiregressivo* andando ad aggiungere ad ogni step una variabile esplicativa, minimizzando di volta in volta l'indice di Mallow (*Weisberg, 1985*) fino a quando la variabile aggiunta determinava una variazione significativa dell'indice considerato.

Identificato il modello ottimale, la procedura verifica se una tra le variabili esplicative può essere rimossa senza determinare variazioni significative dell'indice preso come riferimento. In questo modo, per ciascuna statistica, è stato selezionato il modello che consentiva di minimizzare l'indice di Mallow e contemporaneamente massimizzare l'efficienza di Nash (secondo la formulazione già citata nella Tabella 1.1), calcolato confrontando i valori empirici delle grandezze con quelli stimati dal modello *multiregressivo*.

L'analisi ha consentito di individuare per ciascuna statistica le combinazioni ottimali dei parametri geomorfologici e climatici per la stima di μ , σ^2 e ρ_1 che vengono di seguito riportate:

$$\mu = -3.907 + 0.016A - 0.003H_{\min} - 0.118L - 0.422t_c + 0.0055MAP \quad (6.9)$$

$$\sigma^2 = 0.28 \cdot 10^{-3} \cdot P^{-0.265} \cdot \Delta H^{0.577} \cdot L^{2.400} \cdot MAP^{0.046} \quad (6.10)$$

$$\rho_1 = 0.761 + 10^{-3} \cdot (1.6P - 0.03H_{\max} + 0.04\Delta H - 0.069MAP) \quad (6.11)$$

I modelli *multiregressivi* rivelatisi più efficienti nella fase di calibrazione sono stati applicati in seguito mediante la procedura di ricampionamento *jack-knife* (un richiamo in merito a tale procedura è riportato nel paragrafo 6.1.2.1). In conclusione

l'applicazione ha consentito di stimare in maniera indiretta le grandezze statistiche per ciascun bacino simulando per essi condizioni non strumentate.

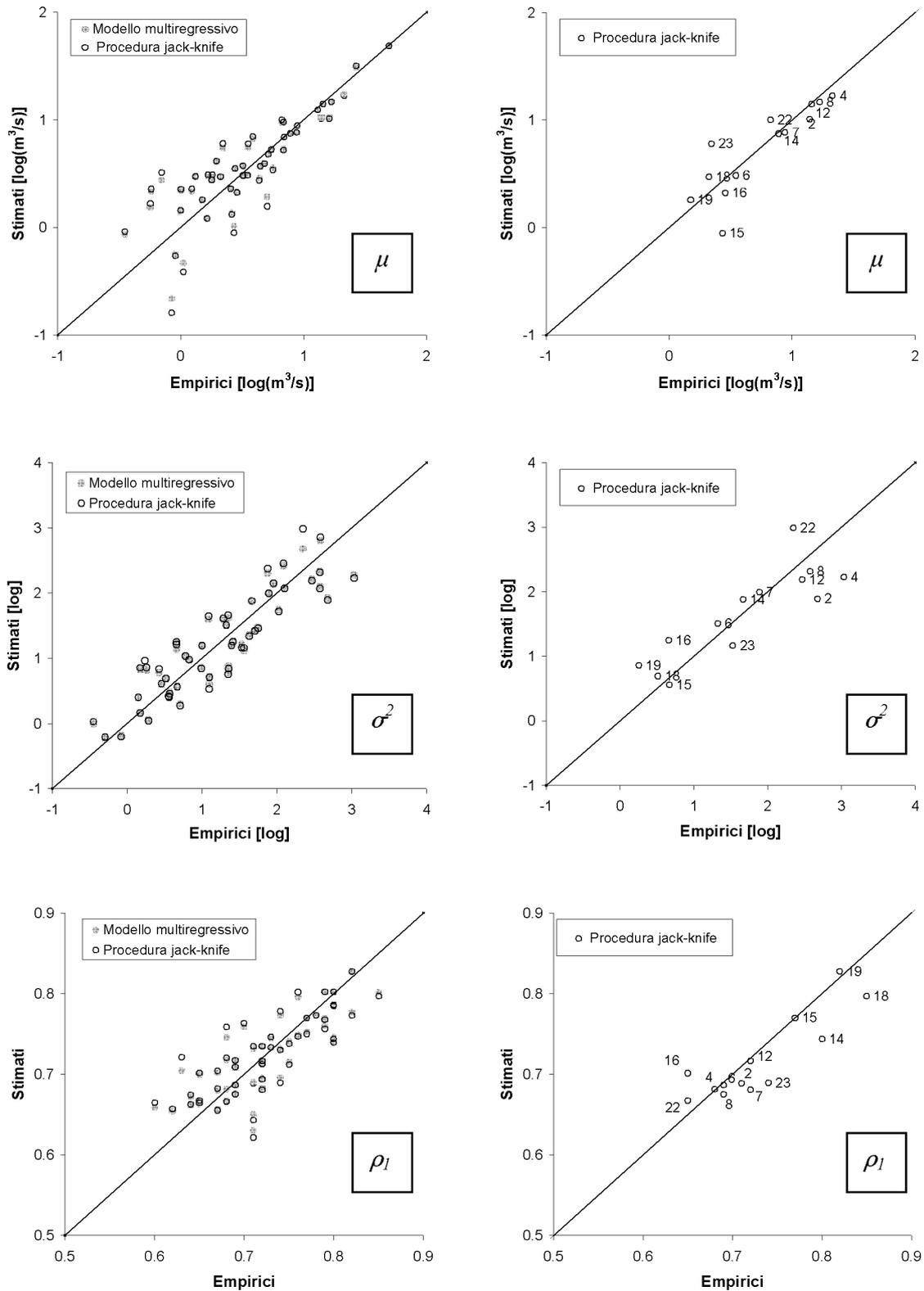


Figura 6.1. Scatter-plots che riportano un confronto diretto fra le grandezze statistiche empiriche e quelle stimate mediante modelli multiregressivi applicati con procedura di ricampionamento jack-knife. In particolare nei grafici di sinistra vengono riportati i valori relativi ai 52 bacini, mentre a destra un particolare di quelli ottenuti per i 13 bacini per i quali è stata applicata la procedura di calibrazione.

La procedura di regionalizzazione è stata applicata a tutta l'area descritta in precedenza (Capitolo 4) comprendente 52 bacini mentre la procedura di calibrazione è successivamente stata applicata ad una porzione di tale area comprendente solo 13 bacini. Al fine di considerare un confronto diretto con la trattazione nel caso di bacini strumentati già esposta nel Capitolo 5, si riportano di seguito i risultati ottenuti per i medesimi 4 bacini trattati questa volta come non strumentati.

I risultati dell'analisi sono riportati in Figura 6.1 dove si mostra un confronto diretto fra i valori empirici e quelli stimati in validazione *jack-knife*. In particolare, a fianco di ciascun grafico relativo all'applicazione in validazione per i 52 bacini, vengono riportati i risultati delle stime relative ai 13 bacini per i quali è stato possibile svolgere la successiva analisi della procedura di calibrazione.

6.1.2 Indici di Magra

In riferimento al lavoro svolto dall'Ing. Castiglioni precedentemente citato, si è preso in considerazione anche il valore regionalizzato della portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (Q_{355}). Tale valore assume il significato di indice di magra ed è utile al fine di tentare di riprodurre al meglio i bassi deflussi ed avere quindi una stima attendibile della disponibilità di risorsa idrica nei bacini oggetto di indagine (Castiglioni, 2010a - Lombardi et al., 2010b).

Il valore della Q_{355} è stato regionalizzato tramite una tecnica di interpolazione spaziale operando all'interno di un dominio di calcolo bidimensionale definito in base a due coordinate cartesiane ortogonali determinate in funzione dei descrittori dei bacini. Pertanto in primis è stato necessario ridurre la dimensione dello spazio delle caratteristiche geomorfoclimatiche attraverso l'impiego di una procedura statistica di analisi multivariata che prende il nome di Analisi delle Componenti Principali (*Principal Component Analysis*, PCA, Basilevsky, 1994). Tale metodologia permette di ottenere, a partire da un insieme di n variabili osservate (in questo caso dieci: A , L , P , H_{max} , H_{med} , H_{min} , W_{Hr} , τ_c , MAP , TAM), uno spazio definito da n nuove variabili, determinate come combinazione lineare delle precedenti. In particolare è possibile concentrare la maggior parte delle informazioni possedute dall'insieme di variabili originarie nelle prime due componenti principali che, fra loro non correlate e ortogonali, possono essere utilizzate come base spaziale di un sistema bidimensionale cartesiano sul quale applicare la tecnica di interpolazione scelta. La tecnica PCA è stata applicata all'area di studio prendendo inizialmente in considerazione le dieci caratteristiche

geomorfoclimatiche precedentemente elencate. Le prime applicazioni della procedura hanno mostrato lo scarso potere descrittivo della *TAM*, i cui valori empirici risultano assai poco diversificati sull'area di studio e praticamente indipendenti dai tre indici di deflusso di magra precedentemente menzionati. La *TAM* è stata quindi rimossa dall'insieme di caratteristiche considerate nelle analisi successive.

Le prime due componenti principali valutate per l'insieme di nove descrittori rimanenti ha permesso di definire una base spaziale bidimensionale spiegando circa il 70% della varianza delle caratteristiche geomorfoclimatiche originarie.

6.1.2.1 Procedura di ricampionamento "jack - knife"

L'accuratezza di stima della metodologia è stata valutata mediante l'utilizzo della procedura di ricampionamento *jack-knife* (si veda *Efron, 1982; Shao & Tu, 1995; Brath et al., 2003; Castellarin et al., 2004b; Zhang & Kroll, 2007*). Essa rappresenta uno strumento estremamente versatile, di semplice applicazione, in grado di fornire un adeguato controllo sulle prestazioni degli interpolatori utilizzati, simulando le condizioni non strumentate in ciascuno dei bacini dell'area di studio. L'applicazione della procedura di cross-validazione *jack-knife* può essere riassunta come segue:

1. dato un insieme di N bacini, se ne elimina uno, di coordinate (x_i, y_i) dall'insieme delle osservazioni disponibili;
2. viene definito uno spazio bidimensionale utilizzando le prime due componenti principali determinate a partire dai descrittori geomorfoclimatici degli $N-1$ bacini rimanenti;
3. le tecniche di interpolazione scelte vengono applicate allo spazio bidimensionale così costruito interpolando i valori empirici di Q_{355} di ciascun bacino e ottenendo quindi delle superfici che rappresentano la variazione dell'indice di magra nello spazio costituito;
4. la stima della Q_{355} del bacino i -esimo è determinabile dalle superfici generate a partire dalle sue coordinate (x_i, y_i) ;
5. gli step 1 - 4 vengono ripetuti considerando a turno uno dei restanti $N-1$ bacini;
6. le stime delle Q_{355} ottenute in cross-validazione vengono successivamente confrontate con i corrispondenti valori empirici.

Per la quantificazione dell'accuratezza di stima caratteristica della tecnica utilizzata sono stati utilizzati tre diversi indici di prestazione: l'efficienza di Nash-Sutcliffe EFF (*Nash & Sutcliffe, 1970*), l'errore medio relativo in modulo MRE, e l'errore quadratico

medio relativo, RRMSE secondo le espressioni già riportate in precedenza (Catiglioni, 2010).

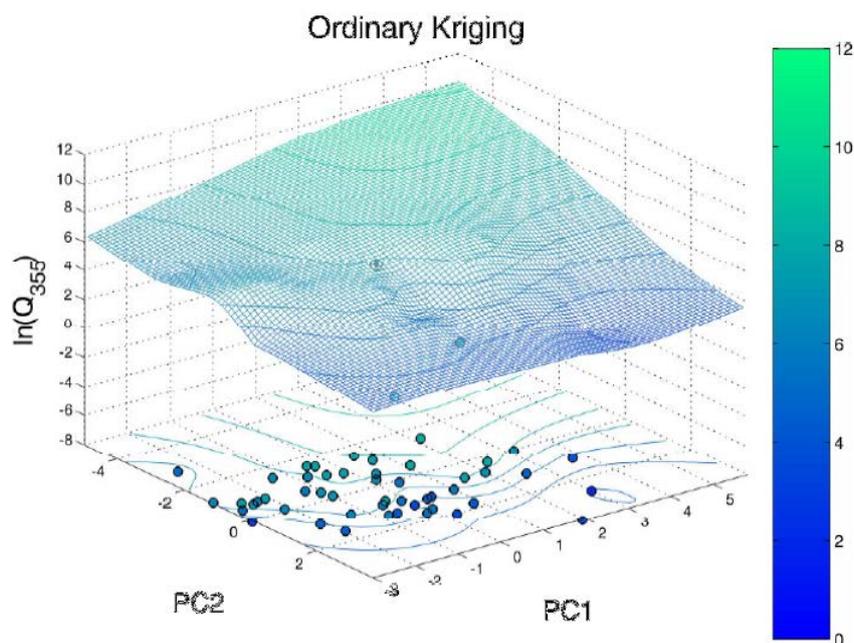


Figura 6.2. Esempio di applicazione della tecnica di interpolazione spaziale Ordinary Kriging in cross-validazione per il bacino del Fiume Aterno facente parte dell'area di studio di 52 bacini utilizzati per la regionalizzazione delle statistiche.

In Figura 6.2 viene riportato un esempio di applicazione delle tecniche di interpolazione in cross-validazione per uno dei bacini dell'area di studio relativo al fiume Aterno nei pressi della sezione di chiusura de L'Aquila ($A=489.3 \text{ km}^2$). La figura mostra la superficie tridimensionale generata con l'Ordinary Kriging e riporta i differenti valori della Q_{355} mediante scala cromatica. Tale superficie è rappresentata sul piano x-y attraverso delle isolinee (piano delle componenti principali, o spazio dei descrittori geomorfoclimatici) insieme ai punti che rappresentano ciascuno dei 52 bacini considerati. Il colore dei punti rappresenta il valore empirico della Q_{355} adottando la stessa scala cromatica della superficie. La Figura 6.2 consente quindi di confrontare tridimensionalmente i valori empirici della Q_{355} (puntini neri) con i corrispondenti valori stimati in validazione (cerchietti neri).

6.2 Scelta delle Funzioni Obiettivo per la calibrazione in bacini non strumentati

Determinate le caratteristiche statistiche dei deflussi fluviali, è possibile utilizzare queste grandezze per calibrare il generico modello afflussi – deflussi (ad es. Hymod, Boyle, 2000), il quale viene applicato ponendo in ingresso dati meteorologici, quali per

esempio serie di precipitazioni, di temperatura e di evapotraspirazione. Queste grandezze possono essere considerate ragionevolmente note in quanto gli strumenti di misura delle precipitazioni e della temperatura sono generalmente più diffusi nel territorio rispetto agli idrometri, perciò le serie di precipitazione e temperatura possono essere ricostruite a scala di bacino attraverso le comuni tecniche di interpolazione spaziale, quali per esempio la poligonazione di Thiessen o la distanza inversa (Goovaerts, 2000).

Nel caso di calibrazione *mono-obiettivo* è stato necessario combinare le grandezze statistiche in un'unica equazione. Poiché a priori non è possibile sapere se una statistica può avere un peso maggiore di un'altra nel processo di calibrazione, si è deciso di utilizzare un peso uguale per tutte.

Riassumendo per le modalità *mono-obiettivo* sono state utilizzate le seguenti funzioni obiettivo:

- media, varianza e coefficiente di correlazione lineare a ritardo un giorno (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}\rho_1$),

$$\phi(x) = \frac{1}{3} \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 + \frac{1}{3} \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 + \frac{1}{3} \left(\frac{\rho_1 - \rho_1^{2*}}{\rho_1^{2*}} \right)^2 \quad (6.12)$$

- media, varianza e portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}Q_{355}$),

$$\phi(x) = \frac{1}{3} \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 + \frac{1}{3} \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 + \frac{1}{3} \left(\frac{Q_{355} - Q_{355}^*}{Q_{355}^*} \right)^2 \quad (6.13)$$

- media, varianza, coefficiente di correlazione lineare a ritardo un giorno e portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$),

$$\phi(x) = \frac{1}{4} \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 + \frac{1}{4} \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 + \frac{1}{4} \left(\frac{\rho_1 - \rho_1^{2*}}{\rho_1^{2*}} \right)^2 + \frac{1}{4} \left(\frac{Q_{355} - Q_{355}^*}{Q_{355}^*} \right)^2 \quad (6.14)$$

dove le statistiche con pedice *sim* si riferiscono ai valori simulati dal modello, mentre i valori delle statistiche indicate con apice * sono quelli ottenuti dalla procedura di regionalizzazione descritta in precedenza.

Per quanto riguarda le simulazioni con modalità *multi-obiettivo* sono state utilizzate le seguenti funzioni obiettivo:

- CASO A TRE FUNZIONI: media, varianza e coefficiente di correlazione lineare a ritardo un giorno (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$),

$$\phi_1(x) = \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 \quad (6.15)$$

$$\phi_2(x) = \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 \quad (6.16)$$

$$\phi_3(x) = \left(\frac{\rho_1 - \rho_1^{2*}}{\rho_1^{2*}} \right)^2 \quad (6.17)$$

- CASO A TRE FUNZIONI: media, varianza e portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$),

$$\phi_1(x) = \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 \quad (6.18)$$

$$\phi_2(x) = \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 \quad (6.19)$$

$$\phi_3(x) = \left(\frac{Q_{355} - Q_{355}^*}{Q_{355}^*} \right)^2 \quad (6.20)$$

- CASO A QUATTRO FUNZIONI: media, varianza, coefficiente di correlazione lineare a ritardo un giorno e portata media giornaliera con frequenza di superamento di 355 giorni su base annuale (abbreviato in seguito $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$),

$$\phi_1(x) = \left(\frac{\mu_{sim} - \mu^*}{\mu^*} \right)^2 \quad (6.21)$$

$$\phi_2(x) = \left(\frac{\sigma_{sim}^2 - \sigma^{2*}}{\sigma^{2*}} \right)^2 \quad (6.22)$$

$$\phi_3(x) = \left(\frac{\rho_1 - \rho_1^{2*}}{\rho_1^{2*}} \right)^2 \quad (6.23)$$

$$\phi_4(x) = \left(\frac{Q_{355} - Q_{355}^*}{Q_{355}^*} \right)^2 \quad (6.24)$$

In parallelo con quanto svolto nel caso di bacini strumentati si è deciso di procedere ad un confronto tra la modalità *mono-obiettivo* e la modalità *multi-obiettivo*. A partire dal confronto degli indici di qualità si è deciso di verificare quanti set ottenuti con la modalità *multi-obiettivo* fossero migliori rispetto all'unico risultato ottenuto con la modalità *mono-obiettivo*. La possibilità di controllare la loro posizione sulla *Frontiera Pareto* ha permesso inoltre di arrivare ad alcune considerazioni in merito all'incidenza delle statistiche selezionate sull'identificazione dei valori dei parametri.

Infine si è effettuato un controllo sugli involucri massimo e minimo delle portate simulate in modalità *multi-obiettivo* al fine di verificare la percentuale di portate inglobate in tali involucri. Lo scopo specifico era verificare se le difficoltà maggiori di riproduzione fossero riferite a periodi di scarsi o elevati deflussi. Nei paragrafi successivi sono riportati i risultati di queste indagini.

6.3 Risultati e confronti

In prima analisi si è deciso di confrontare i valori degli indici di qualità ottenuti con la calibrazione *multi-obiettivo* realizzata utilizzando i valori delle statistiche direttamente ottenute dai valori osservati (bacini Strumentati) rispetto al caso di valori delle statistiche ottenute dalla procedura di regionalizzazione (bacini non Strumentati).

Questo confronto ha lo scopo di verificare l'entità del deterioramento dei risultati che, seppur atteso, deve essere contenuto al fine di poter procedere con le simulazioni in bacini non strumentati. Inoltre, tramite questi valori, è possibile avere una stima della qualità della procedura di regionalizzazione. Questi risultati sono riportati nel paragrafo 6.3.1.

Successivamente, dimostrata l'attendibilità della procedura di regionalizzazione, ci si è addentrati nell'esame specifico del caso di bacini non Strumentati e si è operato, analogamente a quanto visto nei paragrafi precedenti, facendo un confronto diretto tra modalità di calibrazione *mono-obiettivo* e *multi-obiettivo*. Lo scopo è stato indagare se quanto emerso per il caso di bacini strumentati potesse essere

confermato anche al caso di bacini non strumentati. In merito a questo confronto si è voluto anche identificare la posizione dei set della *Frontiera Pareto* che portassero a risultati migliori rispetto alla calibrazione *mono-obiettivo* al fine di evidenziare l'eventuale incidenza delle statistiche sui risultati ottenuti. Questi risultati sono riportati nei paragrafi 6.3.2 - 6.3.3 - 6.3.4.

In fine, nello specifico caso di calibrazione *multi-obiettivo*, si è analizzata la percentuale di portata inglobata negli inviluppi massimo e minimo con lo scopo di verificare se le difficoltà di riproduzione fossero maggiormente legate alla simulazione delle portate di magra o delle portate di picco. Questi risultati sono riportati nel paragrafo 6.3.5.

Al fine di rendere più semplice un eventuale confronto tra modalità *mono* e *multi obiettivo* si è deciso di riportare i grafici (per ciascun bacino) relativi solo a quella modalità che si è dimostrata migliore per ciascuna categoria così come già effettuato per il caso di bacini strumentati.

6.3.1 Analisi degli indici di qualità nel passaggio da bacino Strumentato a bacino non Strumentato per la calibrazione multi-obiettivo

Gli idrogrammi ottenuti dalla calibrazione realizzata con i valori delle statistiche ottenuti dai dati osservati sono molto simili a quelli ottenuti con l'utilizzo dei valori regionalizzati, tanto da rendere quasi inutile un confronto visivo. Si riportano pertanto solo le tabelle con gli indici di qualità in calibrazione e validazione in modo da poter fare delle considerazioni direttamente sui valori ottenuti (*Lombardi et al., 2010*).

Come già accennato in precedenza, l'ottimizzazione *multi-obiettivo* genera un insieme di set dei parametri ugualmente possibili e che per questo appartengono alla *Frontiera Pareto* come precedentemente definita (secondo quanto descritto nel paragrafo 1.3). In questo caso, oltre all'analisi del comportamento della portata media così come definita e ottenuta per le modalità *multi-obiettivo* applicate ai bacini considerati come strumentati, si è deciso di utilizzare la portata ottenuta con il set di parametri più vicino all'origine degli assi sulla *Frontiera Pareto* (Figura 6.3) indicata di seguito come Q_{co} (*Closest to the Origin of the axis*). Il motivo di tale scelta è legato al fatto che per applicare la modalità *mono-obiettivo* è stato necessario attribuire dei pesi equi alle varie statistiche (equazioni 6.12 - 6.13 - 6.14). In teoria quindi, se la frontiera avesse la forma teorica indicata in Figura 6.3 nel caso a due funzioni (o una forma analoga nel caso a tre o quattro funzioni) il punto che pesa in modo equo le varie statistiche

sarebbe proprio il più vicino possibile all'origine. Questa considerazione ci permette di avere, oltre ad un confronto con la modalità *mono-obiettivo*, anche un'idea della forma della *Frontiera Pareto*.

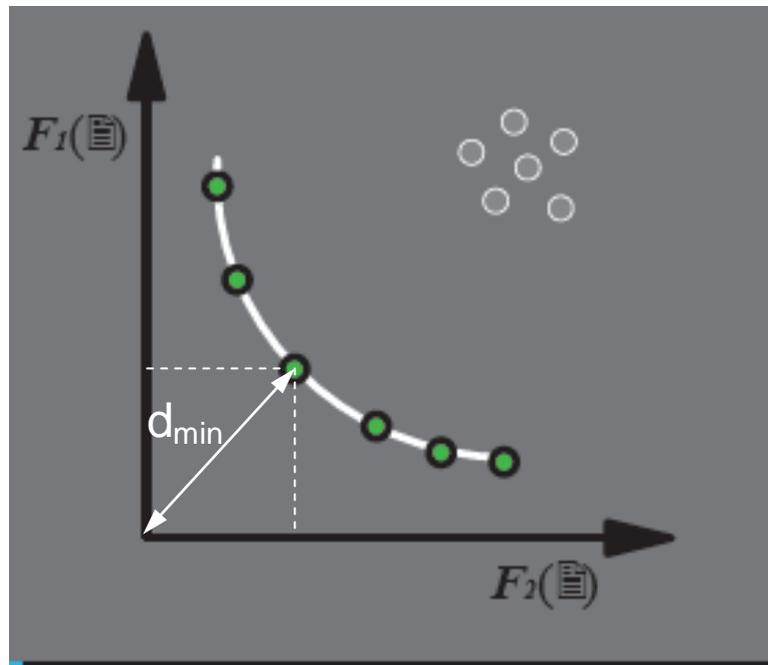


Figura 6.3. Frontiera Pareto: set della frontiera a minima distanza dall'origine degli assi.

Le tabelle seguenti riportano gli indici di qualità ottenuti per tutti i bacini in calibrazione e in validazione ottenute considerando il set più vicino all'origine degli assi della *Frontiera Pareto*. Le funzioni obiettivo sono indicate in grigio e con colori uguali sono evidenziati i valori da confrontare poiché riferiti al caso di bacini strumentati (colonna a sinistra delle due di medesimo colore) e di bacini non strumentati (colonna a destra delle due di medesimo colore) ottenuti con le stesse funzioni obiettivo utilizzate in modalità *multi-obiettivo*.

Il Fiume Candigliano ha un comportamento regolare poiché, come atteso, gli indici peggiorano passando dal caso di bacino strumentato a quello di bacino non strumentato sia in calibrazione che in validazione (Tabella 6.1 e Tabella 6.2). Il peggioramento è tuttavia contenuto e pertanto la regionalizzazione delle statistiche sembra poter essere accettabile. Fa eccezione il caso $\mu\text{-}\sigma^2\text{-}Q_{35}$ per il quale il passaggio al caso di bacino non strumentato provoca un miglioramento degli indici: questo può essere spiegato tenendo conto che a volte può accadere che aggiungendo fattori di incertezza, si ottenga la compensazione con errori legati ad altri fattori (incertezza legata ai dati, alle approssimazioni computazionali, etc...), generando un complessivo miglioramento dei risultati.

Questo potrebbe spiegare il comportamento del Fiume Esino per il quale si è ottenuto un miglioramento di tutte le performances passando alla simulazione tramite le statistiche generate con la regionalizzazione (Tabella 6.3). Questo comportamento si nota anche in validazione (Tabella 6.4) dove accade addirittura che nel complesso i valori degli indici siano maggiori rispetto al periodo di calibrazione. Dall'analisi degli idrogrammi mostrati in seguito si evince infatti che il periodo di calibrazione non viene riprodotto in modo soddisfacente a causa di alcuni fenomeni di carattere nivale che non sono simulati dal modello afflussi-deflussi.

Il Fiume Metauro presenta indici adeguati alle attese ad eccezione della modalità $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ per la quale gli indici migliorano leggermente sia in calibrazione che in validazione passando dall'utilizzo dei valori osservati a quelli regionalizzati (Tabella 6.7 e Tabella 6.8).

CANDIGLIANO (calibrazione-Qco)						
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$ (oss.)	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$ (oss.)	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ (oss.)	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0.64	0.62	0.37	0,7	0.69	0.66
MAE	3.76	4.09	4.52	3.34	3.20	3.68
MRE	0.9	0.9	0.62	0.61	0.55	0.61

Tabella 6.1. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Candigliano.

CANDIGLIANO (validazione-Qco)						
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$ (oss.)	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$ (oss.)	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ (oss.)	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0.54	0.48	0.33	0.55	0.54	0.49
MAE	2.98	3.46	3.08	2.66	2.67	3.13
MRE	0.76	0.84	0.54	0.64	0.57	0.65

Tabella 6.2. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Candigliano.

ESINO (calibrazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.39	0.51	0.24	0.29	0.26	0,53
MAE	8.36	7.47	9.24	8.87	8.91	7,27
MRE	0.65	0.50	0.69	0.6	0.62	0,52

Tabella 6.3. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Esino.

ESINO (validazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.65	0.60	0.42	0.55	0.51	0.67
MAE	6.38	6.36	7.77	6.63	7.31	5.87
MRE	0.57	0.44	0.62	0.49	0.59	0.42

Tabella 6.4. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Esino.

METAURO (calibrazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.64	0.54	0.56	0.24	0.64	0,64
MAE	7.75	10.71	8.65	12.98	7.84	7,82
MRE	0.59	1.64	0.61	0.96	0.71	0,63

Tabella 6.5. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Metauro.

Infine il Fiume Potenza presenta indici che, come atteso, peggiorano nel passaggio alla calibrazione in bacini non strumentati: il peggioramento è contenuto ad indicare una buona qualità della procedura di regionalizzazione, ma i valori di EFF, già bassi in

validazione per il caso di bacini strumentati, arrivano ad essere negativi nel caso di bacini non Strumentati (Tabella 6.7 e Tabella 6.8).

METAURO (validazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.46	0.53	0.42	0.22	0.52	0.63
MAE	9.37	10,69	9.95	12.88	8.90	8.05
MRE	0.64	1.70	0.62	1.02	0.68	0.61

Tabella 6.6. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Metauro.

POTENZA (calibrazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.62	0.33	0.53	0.52	0.56	0,44
MAE	2.35	3.07	2.42	2.96	2.58	3,08
MRE	0.34	0.46	0.32	0.49	0.38	0,53

Tabella 6.7. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Potenza.

POTENZA (validazione-Qco)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ (oss.)	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$ (oss.)	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$ (oss.)	$\mu-\sigma^2-\rho_1-Q_{355}$
EFF	0.20	-0.82	-0.28	-0.72	0.13	-0.87
MAE	2.28	3.14	2.96	3.84	2.35	3.84
MRE	0.41	0.55	0.49	0.71	0.41	0.71

Tabella 6.8. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Potenza.

6.3.2 Analisi dei risultati ottenuti in modalità mono-obiettivo

Nella Tabella 6.9 sono riportati i valori degli indici di qualità (indicati in lilla) ottenuti per il Fiume Candigliano in calibrazione con l'utilizzo delle funzioni obiettivo (indicate in grigio) ciascuna considerata in modalità *mono-obiettivo* secondo le espressioni matematiche riportate in precedenza (equazioni 6.12 - 6.13 - 6.14).

CANDIGLIANO (calibrazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,62	0,64	0,63
MAE	4,25	3,8	3,84
MRE	1,18	0,74	0,73

Tabella 6.9. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

I valori evidenziati (con sfondo azzurro) sono rispettivamente il valore massimo di EFF e minimi di MAE ed MRE che si sono ottenuti tra tutte le modalità *mono-obiettivo*. In questo caso la simulazione ottenuta tramite la combinazione *mono-obiettivo* $\mu\text{-}\sigma^2\text{-}Q_{355}$ sembra aver prodotto i migliori risultati. In Tabella 6.10 sono riportati gli stessi indici ottenuti però relativamente al periodo di validazione.

CANDIGLIANO (validazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,5	0,49	0,49
MAE	3,41	3,24	3,29
MRE	0,96	0,73	0,73

Tabella 6.10. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

Il confronto degli indici in merito al periodo di validazione conferma quanto sostenuto in precedenza e per questo in seguito sono riportati i grafici in calibrazione e in validazione per il Fiume Candigliano relativi alla combinazione mono-obiettivo $\mu\text{-}\sigma^2\text{-}Q_{355}$.

La Figura 6.4 riporta l'idrogramma in calibrazione mentre la Figura 6.5 si riferisce al periodo di validazione. L'idrogramma in calibrazione appare soddisfacente nel suo andamento e riproduce in modo adeguato anche quasi tutti i fenomeni di piena. Il processo di esaurimento di tali piene sembra essere simulato con eccessiva lentezza ma nel complesso l'idrogramma simulato appare conforme all'osservato.

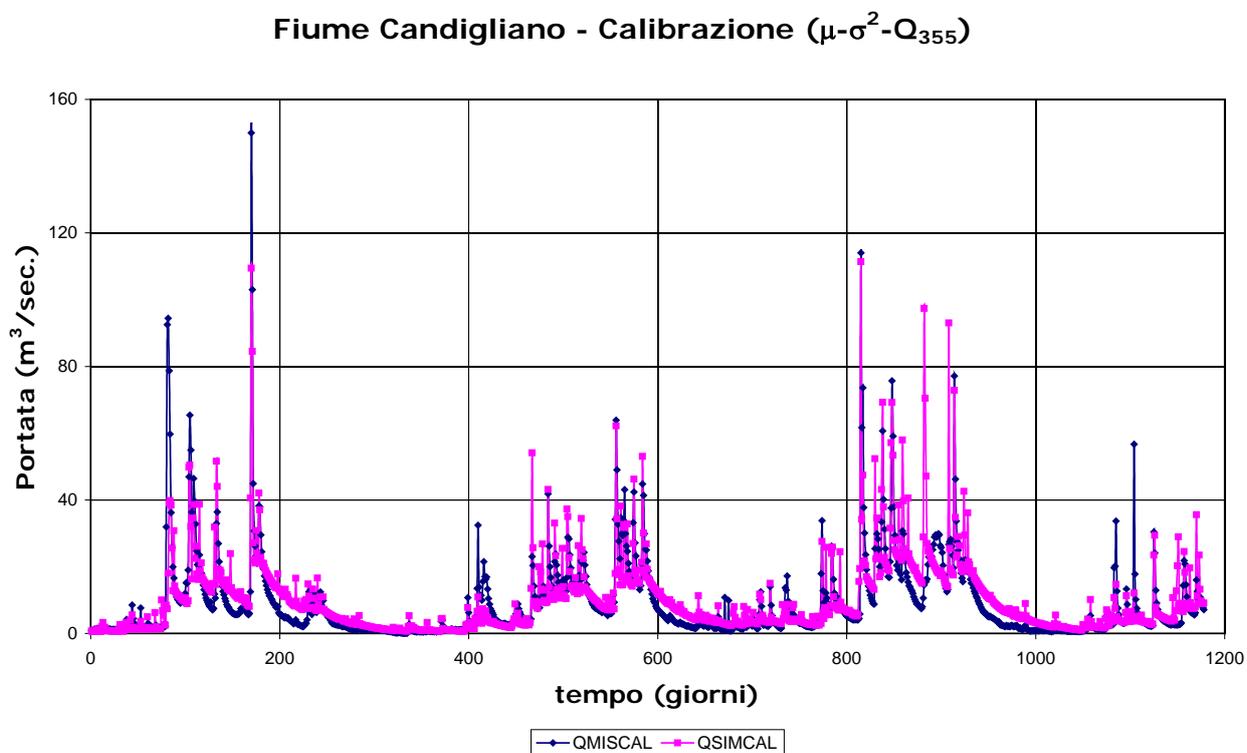
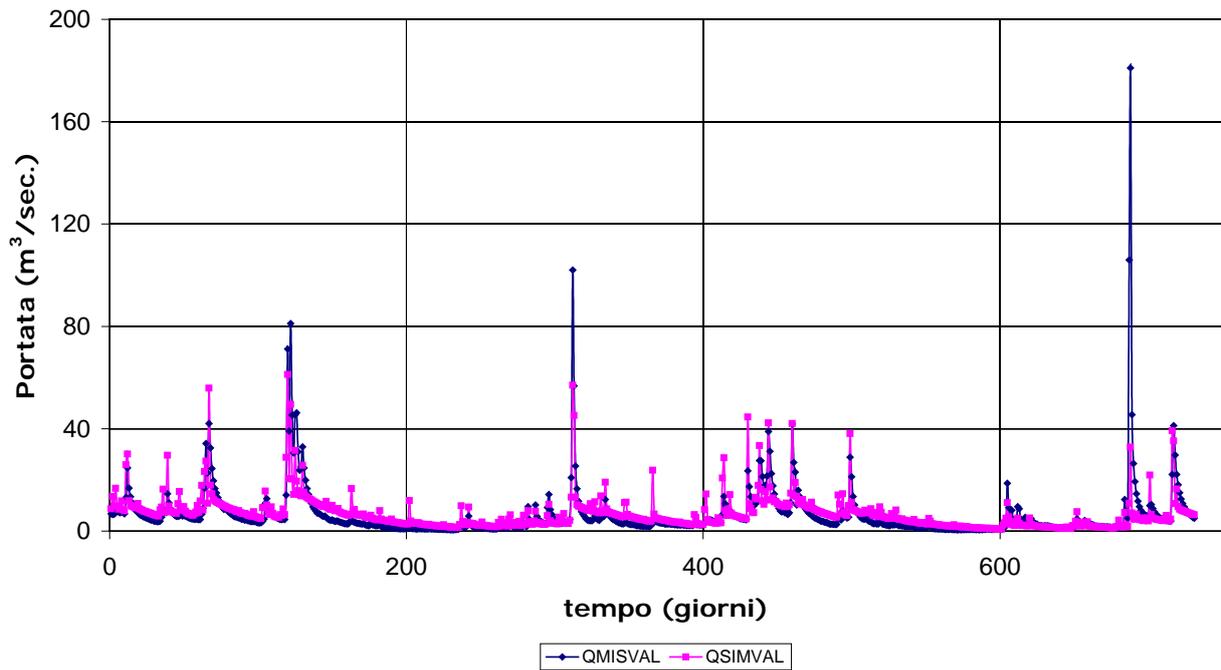
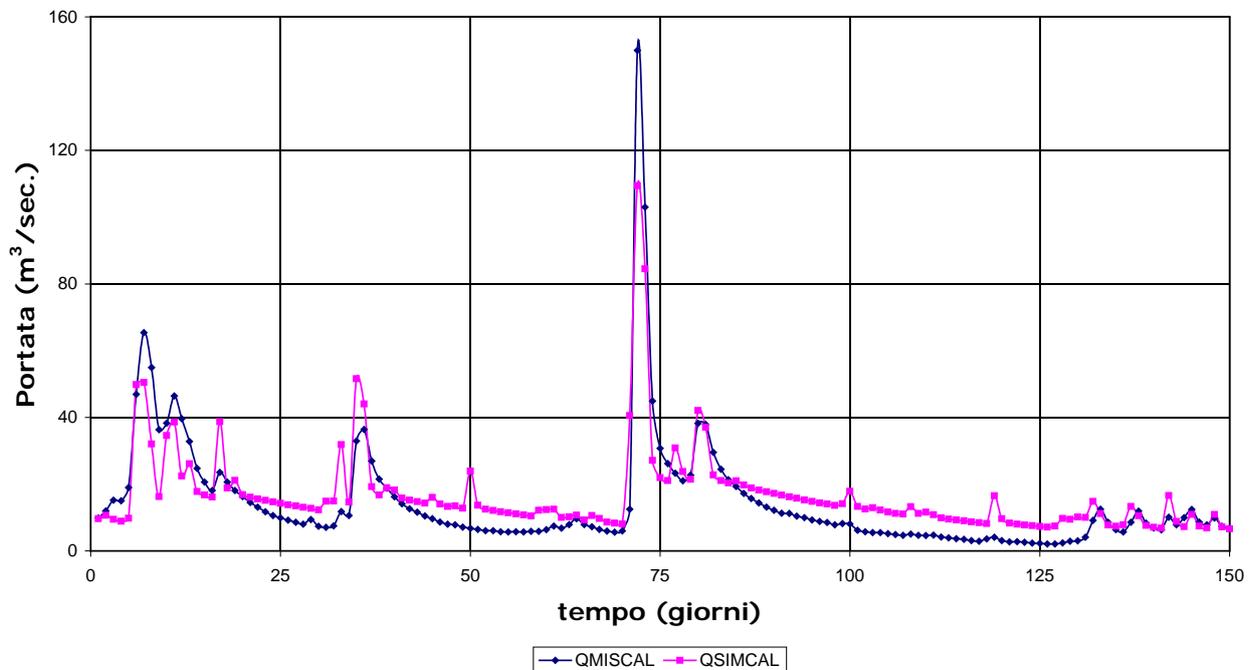


Figura 6.4. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

Fiume Candigliano - Validazione ($\mu\text{-}\sigma^2\text{-}Q_{355}$)Figura 6.5. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu\text{-}\sigma^2\text{-}Q_{355}$).Fiume Candigliano - Calibrazione ($\mu\text{-}\sigma^2\text{-}Q_{355}$)Figura 6.6. Idrogramma in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

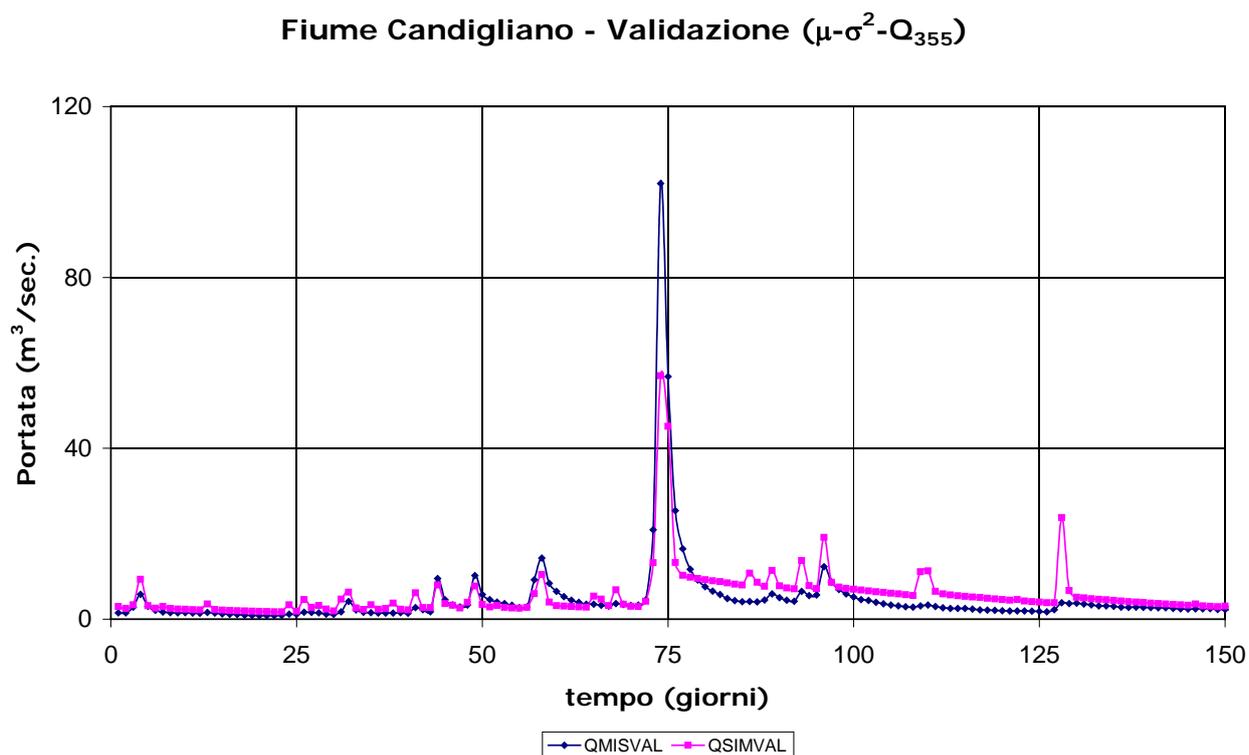


Figura 6.7. Idrogramma in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

Analoghe considerazioni possono essere fatte per il periodo di validazione.

Per avere un confronto diretto con quanto ottenuto nel caso di calibrazione *multi-obiettivo* (paragrafo 6.3.3), si riporta anche il grafico relativo alla finestra temporale di 150 giorni intorno al periodo di massimo picco sia in calibrazione che in validazione.

Questa analisi più dettagliata consente di verificare che la riproduzione degli idrogrammi non è affatto lontana da quella ottenuta con le funzioni obiettivo di utilizzo più classico; l'idrogramma in calibrazione infatti ha un andamento molto simile a quello osservato e si avvicina molto a quello riportato in Figura 5.2 relativo alla calibrazione con l'indice RMSE.

L'idrogramma in validazione si dimostra altrettanto soddisfacente ad indicare che la procedura di regionalizzazione è stata realizzata in modo adeguato e funzionale all'obiettivo preposto.

Nella Tabella 6.11 sono riportati i valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con l'utilizzo delle funzioni obiettivo ciascuna considerata in modalità *mono-obiettivo*.

ESINO (calibrazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,44	0,41	0,44
MAE	7,79	7,58	7,8
MRE	0,52	0,52	0,52

Tabella 6.11. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

In questo caso le simulazioni sono pressoché paragonabili poiché ciascuna ha due indici su tre migliori o uguali alle altre. Controllando anche i valori in validazione (Tabella 6.12) sembra opportuno prediligere anche per questo bacino la simulazione ottenuta tramite la combinazione *mono-obiettivo* $\mu\text{-}\sigma^2\text{-}Q_{355}$.

La Figura 6.8 riporta l'idrogramma in calibrazione mentre la Figura 6.9 si riferisce al periodo di validazione. L'idrogramma in calibrazione appare soddisfacente nella riproduzione delle magra e meno realistico nella riproduzione delle piene. Tali eventi infatti non sono adeguatamente riprodotti né in termini di entità né in termini di sincronismo. Se confrontiamo questo idrogramma con quello ottenuto in modalità *mono-obiettivo* nel caso di bacino considerato come strumentato (Figura 5.5), notiamo un miglioramento nella riproduzione del fenomeno. Le magre sembrano riprodotte in modo più fedele mentre la riproduzione dei picchi è confrontabile nelle due simulazioni.

Gli idrogrammi di Figura 6.10 e Figura 6.11 confermano la considerazione appena fatta in merito alla discreta riproduzione delle portate di magra e alla difficoltà relativa alla riproduzione delle piene. Come già emerso dagli indici e dai risultati ottenuti nel caso di bacino strumentato, la simulazione in validazione si avvicina maggiormente all'osservata piuttosto che quella ottenuta in calibrazione.

ESINO (validazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,63	0,7	0,64
MAE	6,11	5,37	6
MRE	0,42	0,44	0,42

Tabella 6.12. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

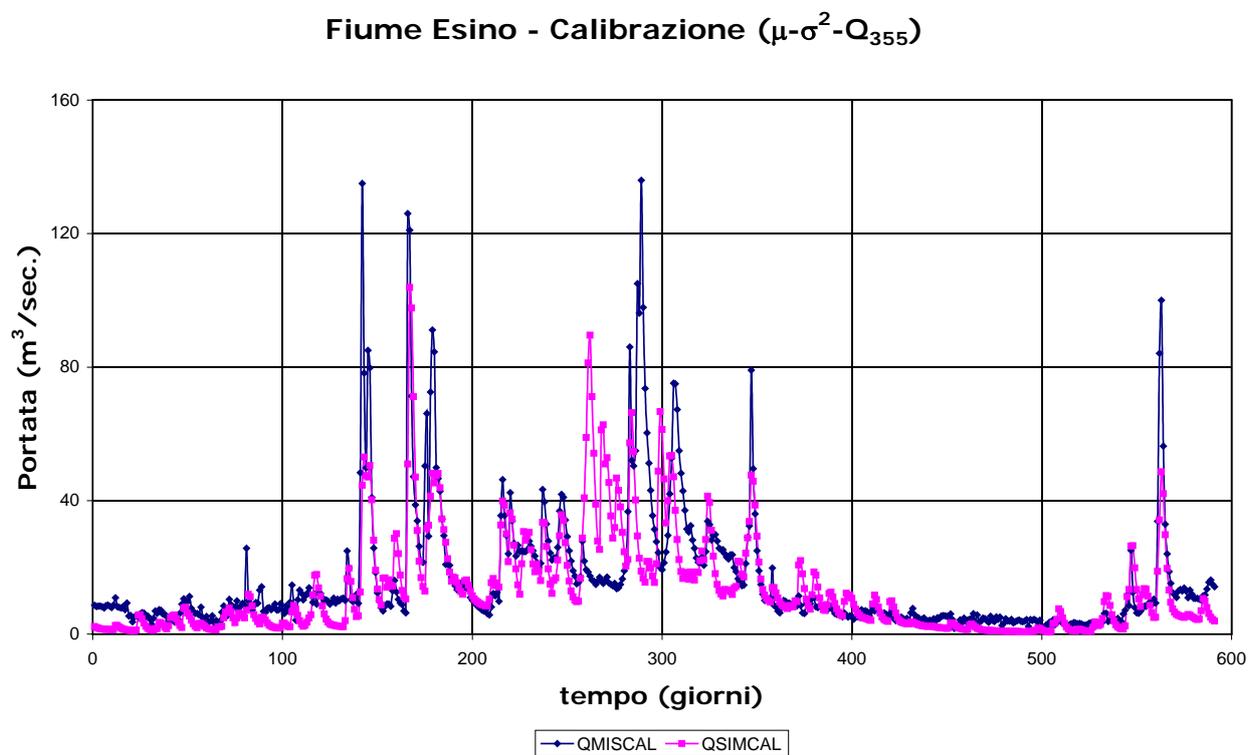


Figura 6.8. Idrogramma in calibrazione del Fiume Esino a Moie ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

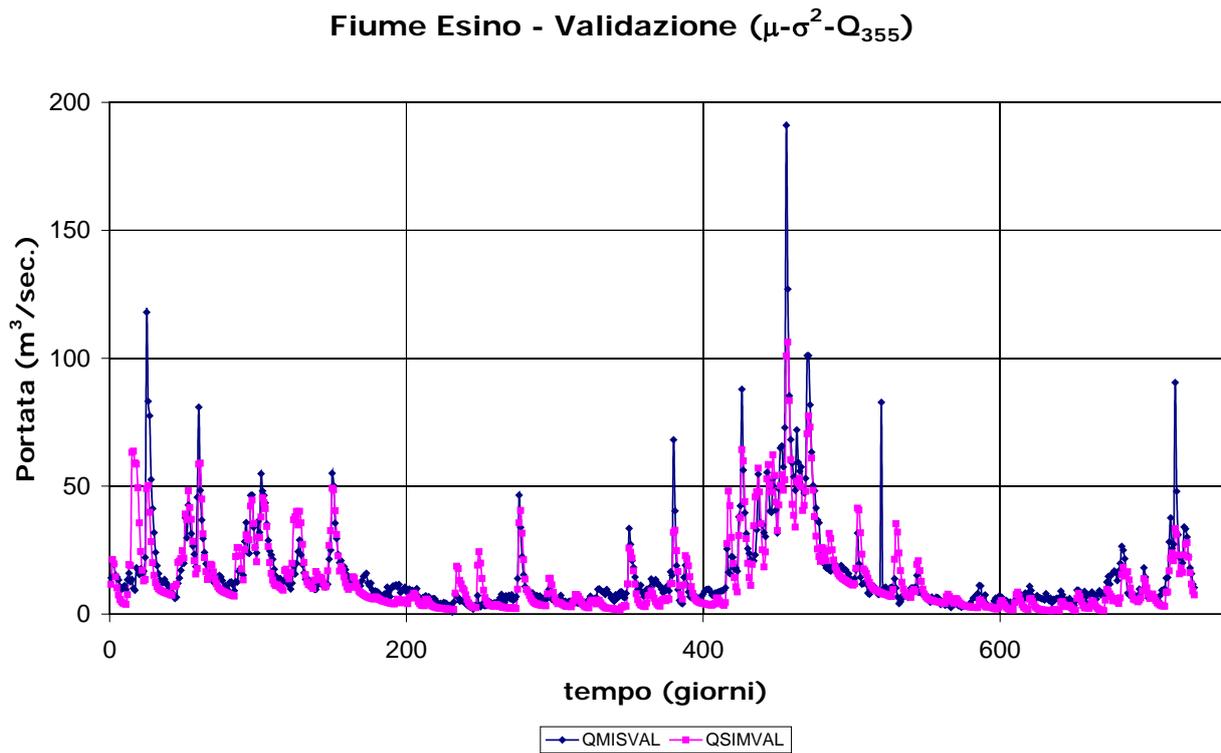


Figura 6.9. Idrogramma in validazione del Fiume Esino a Moie ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

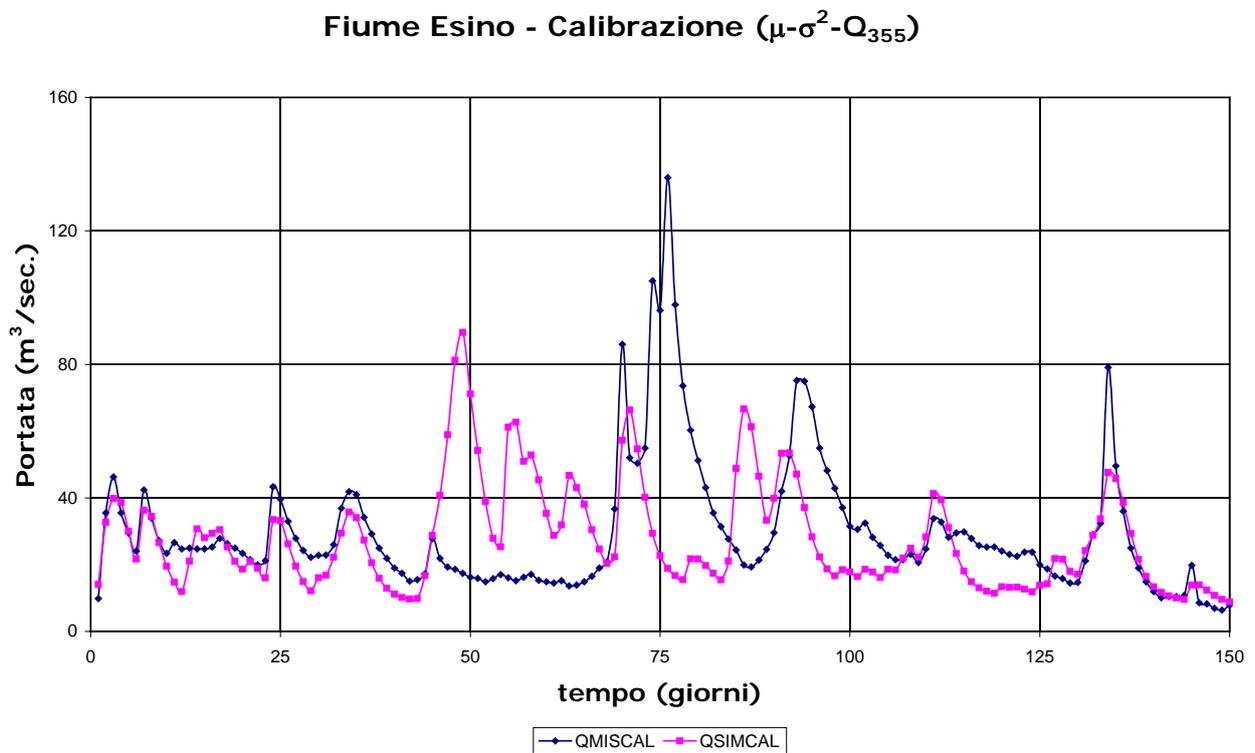


Figura 6.10. Idrogramma in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

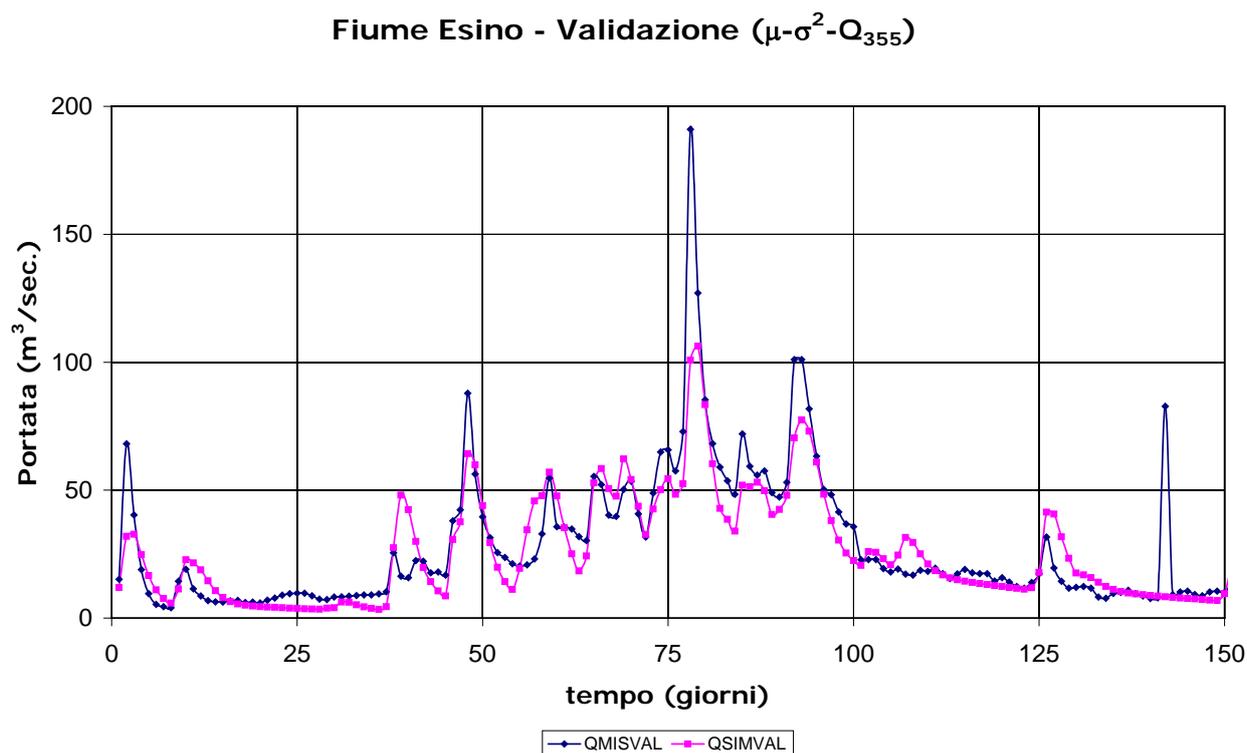


Figura 6.11. Idrogramma in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}$).

Questa considerazione ci permette comunque di confermare la qualità della procedura di regionalizzazione che non sembra aver introdotto fattori di incertezza sui risultati ottenuti.

Nella Tabella 6.13 sono riportati i valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con l'utilizzo delle funzioni obiettivo sempre in modalità *mono-obiettivo* mentre la Tabella 6.14 riporta gli stessi valori per il periodo di validazione.

METAURO (calibrazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,55	0,36	0,36
MAE	10,38	11,54	11,54
MRE	1,49	0,85	0,85

Tabella 6.13. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

In questo caso le simulazioni sono coerenti in calibrazione e in validazione. In entrambi i casi infatti gli indici EFF e MAE sono migliori per la simulazione $\mu\text{-}\sigma^2\text{-}\rho_1$,

mentre il MRE è migliore con le modalità $\mu\text{-}\sigma^2\text{-}Q_{355}$ e $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ pressoché paragonabili poiché ciascuna ha due indici su tre migliori o uguali alle altre. Si predilige allora la simulazione $\mu\text{-}\sigma^2\text{-}\rho_1$ tenendo conto che il deterioramento degli indici appare più significativo se si considera l' EFF e il MAE piuttosto che il MRE.

METAURO (validazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,54	0,34	0,34
MAE	10,34	11,8	11,8
MRE	1,56	0,92	0,92

Tabella 6.14. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

La Figura 6.12 riporta l'idrogramma in calibrazione mentre la Figura 6.13 si riferisce al periodo di validazione. Entrambi gli idrogrammi appaiono anomali nella riproduzione delle magre poiché i rami di esaurimento risultano molto più lenti rispetto a quanto misurato e conseguentemente i valori di portata più bassi vengono riprodotti solo per brevi periodi corrispondenti alla fase subito precedente alla piena successiva piuttosto che per tutto il periodo durante il quale essi si manifestano.

Il fenomeno è meno evidente osservando gli idrogrammi relativi ai 150 giorni attorno all'evento di picco (Figura 6.14 e Figura 6.15): in calibrazione infatti l'andamento dell'idrogramma simulato è molto vicino a quello osservato anche se risulta evidente una sottostima dei volumi complessivi.

In validazione invece le due curve sono molto simili a meno dell'evento di piena che viene sottostimato dal modello in termini di entità ma perfettamente riprodotto in termini temporali.

Rispetto a quanto ottenuto per il caso di bacino strumentato con la modalità *mono-obiettivo* RMSE (Figura 5.8 e Figura 5.9), la simulazione con l'utilizzo delle statistiche porta ad una riproduzione meno fedele delle portate con qualche difficoltà specifica legata alla riproduzione delle magre. La simulazione con l'utilizzo dell'indice di magra Q_{355} probabilmente permette di compensare questo peggioramento, ma provoca un incremento dell'errore sulle portate di picco.

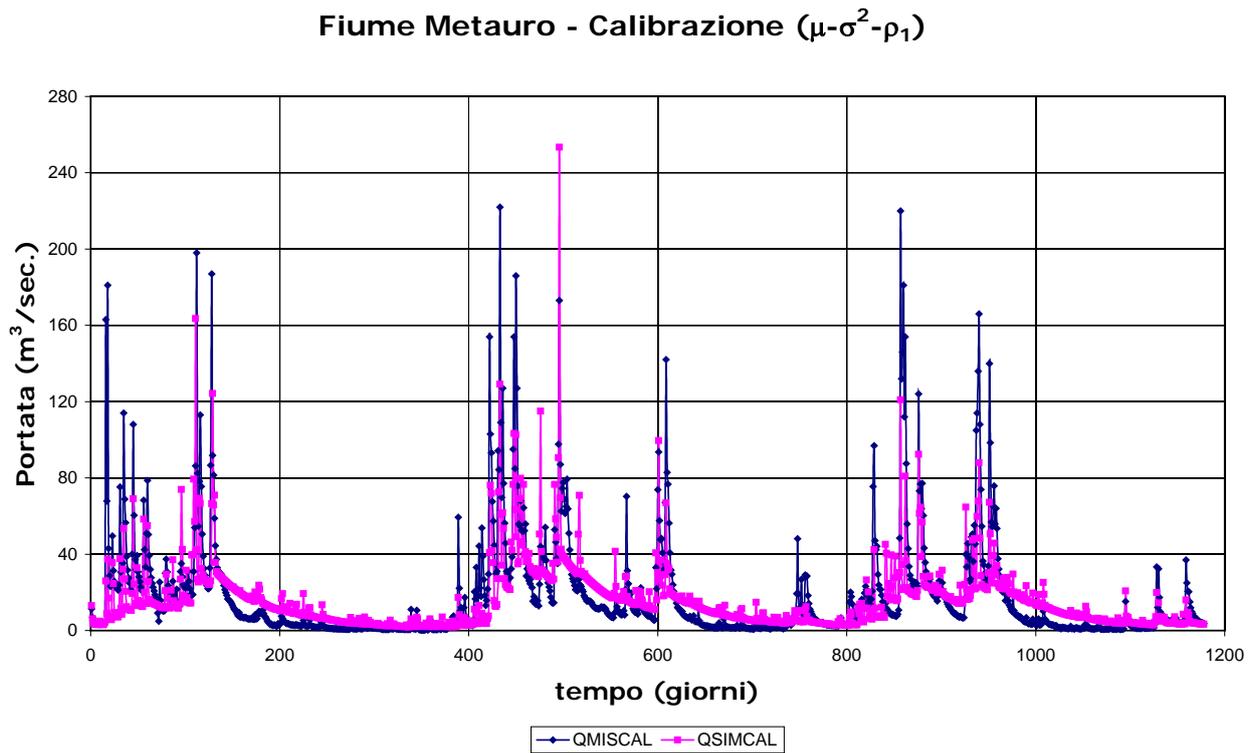


Figura 6.12. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1$).

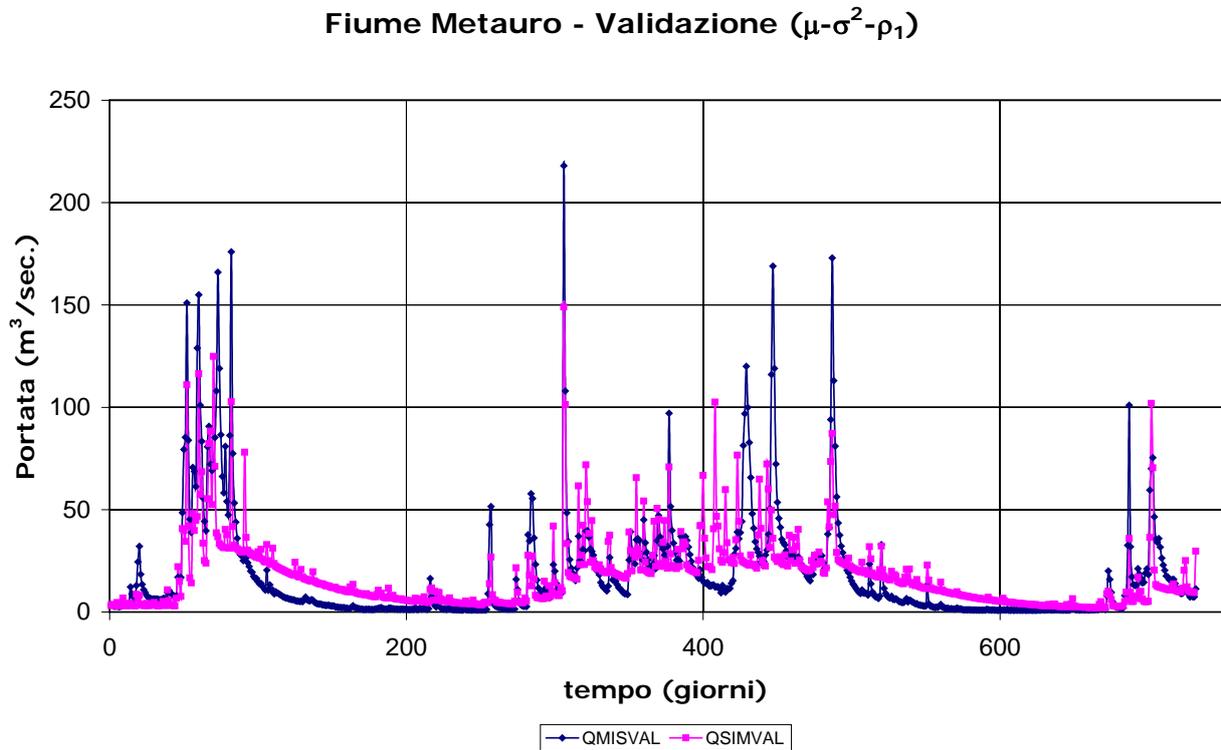


Figura 6.13. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1$).

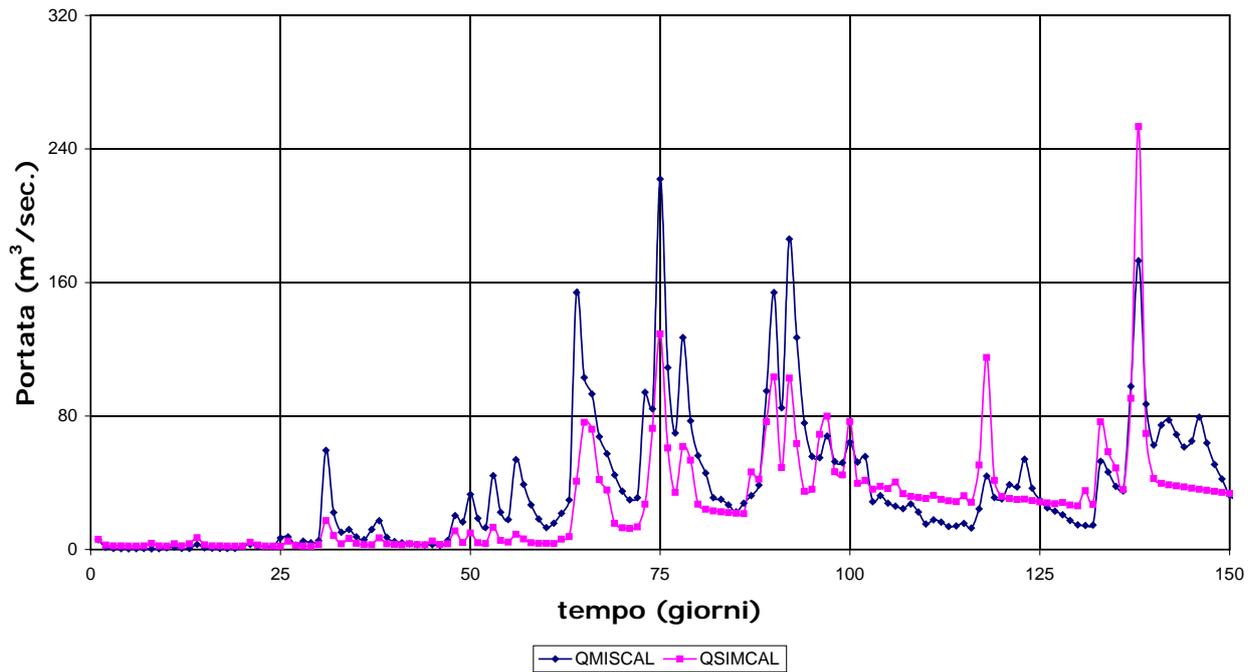
Fiume Metauro - Calibrazione ($\mu-\sigma^2-\rho_1$)

Figura 6.14. Idrogramma in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).

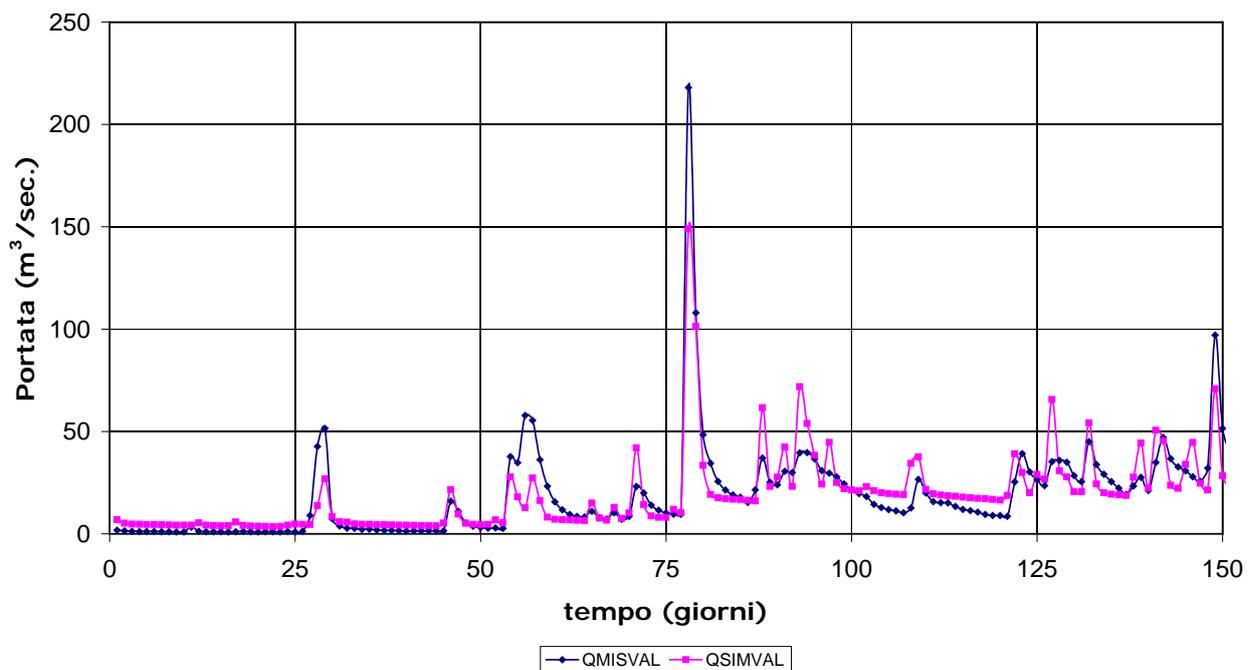
Fiume Metauro - Validazione ($\mu-\sigma^2-\rho_1$)

Figura 6.15. Idrogramma in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).

Nella Tabella 6.15 sono riportati i valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con l'utilizzo delle funzioni obiettivo sempre in modalità *mono-obiettivo* mentre la Tabella 6.16 riporta gli stessi valori per il periodo di validazione.

POTENZA (calibrazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	0,4	0,46	0,45
MAE	2,92	3,04	2,99
MRE	0,46	0,49	0,5

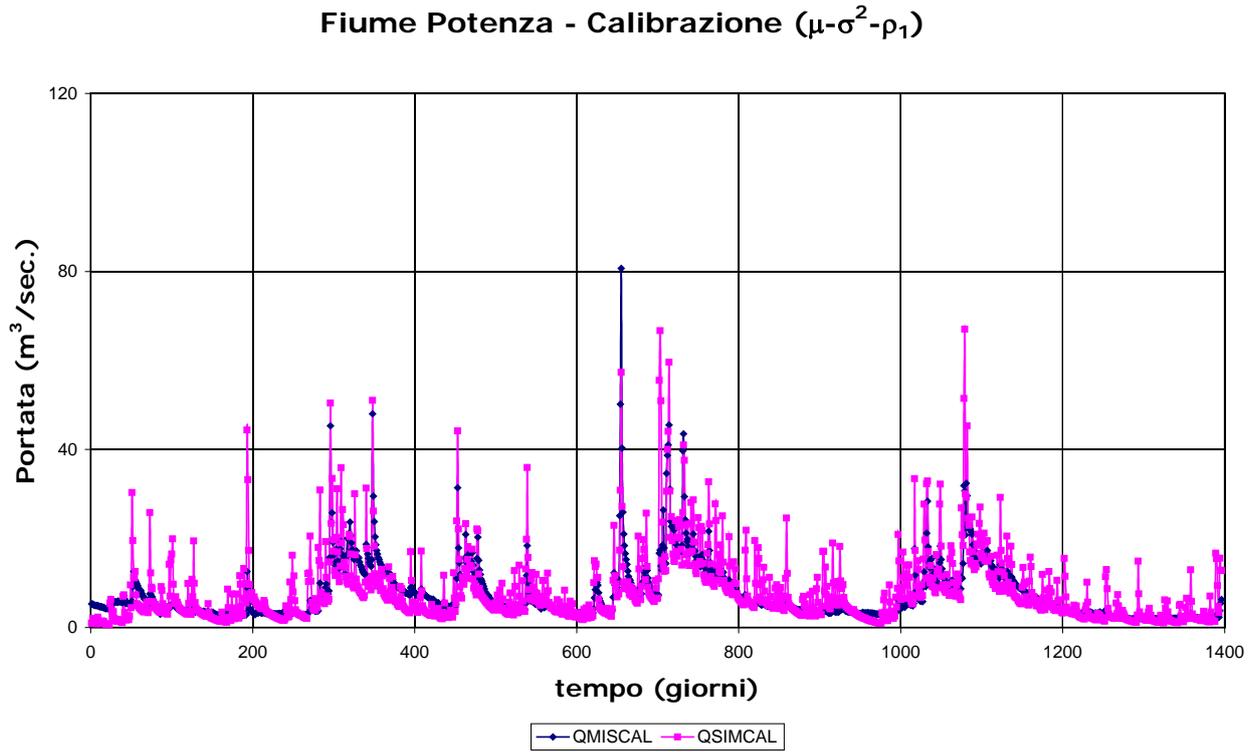
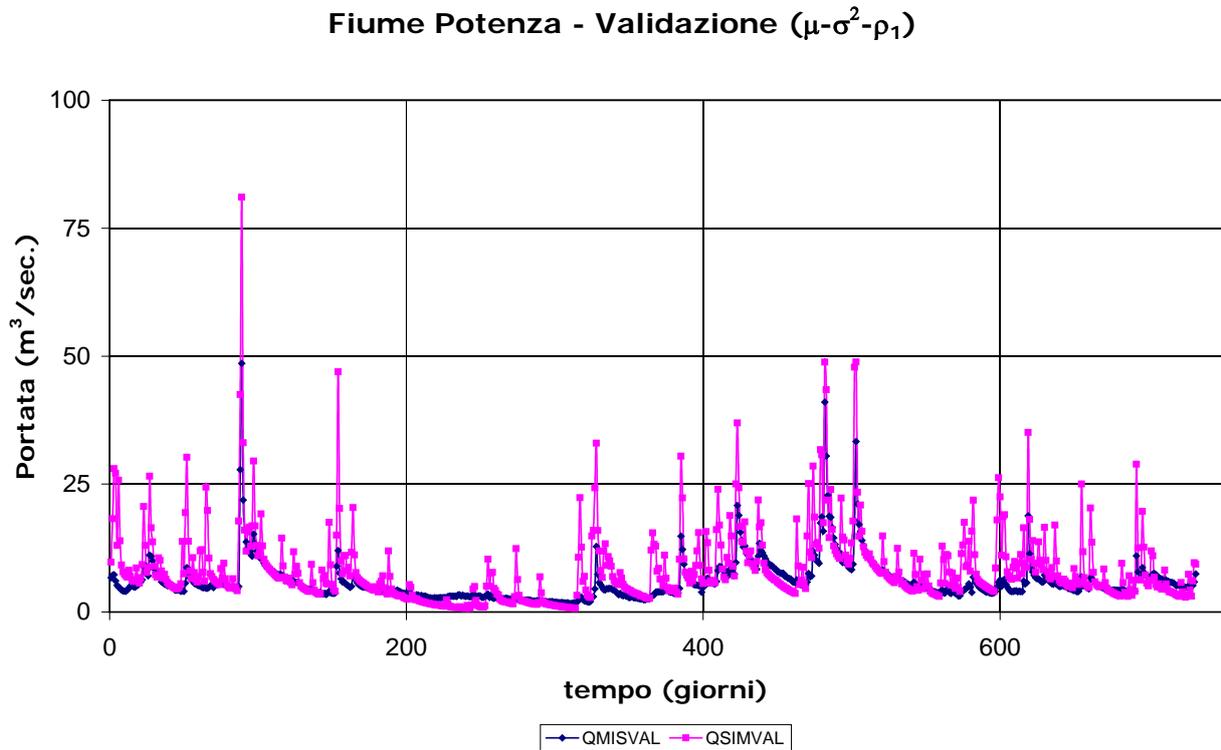
Tabella 6.15. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.

In questo caso i valori dell'EFF, che in calibrazione risultano accettabili, diventano decisamente poco soddisfacenti in fase di validazione presentando addirittura valori negativi. I valori del MAE risultano decisamente più accettabili seppur indicando una certa difficoltà nella riproduzione delle portate elevate. I valori del MRE invece non sono particolarmente negativi. La combinazione migliore sembra quella ottenuta con l'uso della media, della varianza e del coefficiente di correlazione lineare a ritardo un giorno, $\mu\text{-}\sigma^2\text{-}\rho_1$, dunque in seguito sono riportati gli idrogrammi ottenuti con questa tipologia di funzione obiettivo.

POTENZA (validazione)			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1$	$\mu\text{-}\sigma^2\text{-}Q_{355}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$
EFF	-0,71	-0,45	-0,68
MAE	3,09	3,99	3,43
MRE	0,56	0,73	0,63

Tabella 6.16. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.

La Figura 6.16 riporta l'idrogramma in calibrazione mentre la Figura 6.17 si riferisce al periodo di validazione. Entrambi gli idrogrammi seguono nel complesso l'andamento delle portate osservate ma è evidente che la portata simulata ha oscillazioni molto più frequenti, anche se non particolarmente significative, rispetto alla portata osservata.

Figura 6.16. Idrogramma in calibrazione del Fiume Potenza a Cannucciaro ($\mu-\sigma^2-\rho_1$).Figura 6.17. Idrogramma in validazione del Fiume Potenza a Cannucciaro ($\mu-\sigma^2-\rho_1$).

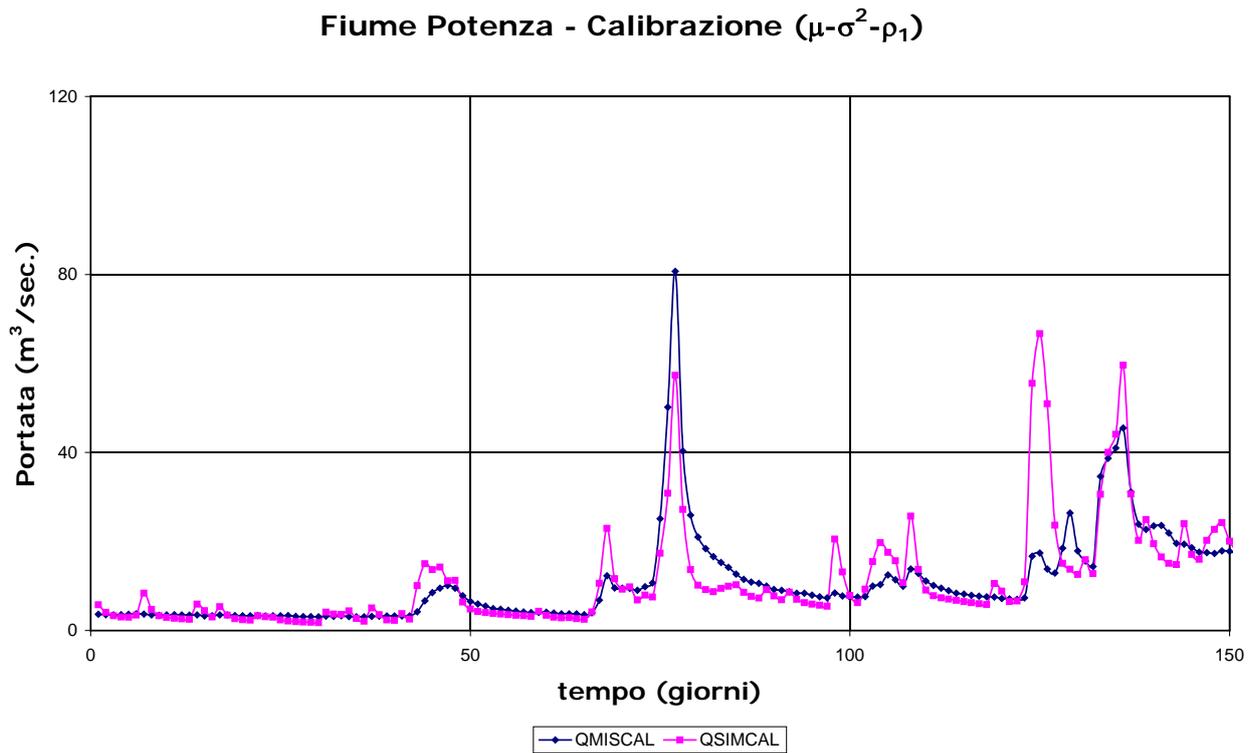


Figura 6.18. Idrogramma in calibrazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1).

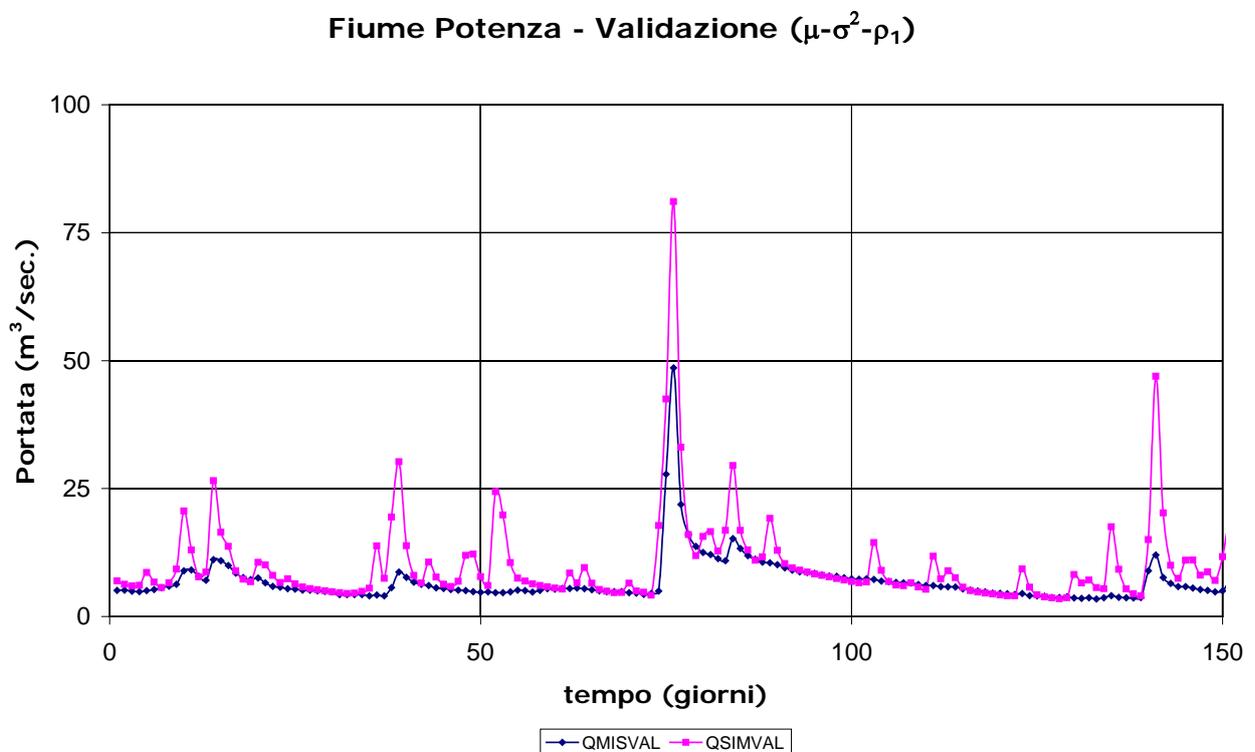


Figura 6.19. Idrogramma in validazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1).

L'esame degli idrogrammi focalizzati attorno al periodo di maggior picco (Figura 6.18 e Figura 6.19) confermano questa osservazione. Il modello risponde evidentemente in modo troppo rapido alle sollecitazioni piovose generando picchi che, nonostante la modesta entità, portano a valori degli indici di qualità non proprio soddisfacenti. Il valore del MRE si mantiene accettabile poiché essendo un indice che pesa maggiormente le magre rispetto alle piene, non risente in modo significativo di questo comportamento del modello sul bacino.

Il fenomeno è ancora più evidente in validazione dove l'idrogramma osservato, a meno del maggior picco, si presenta praticamente piatto rispetto all'idrogramma simulato che presenta numerosi picchi di portata.

6.3.3 Analisi dei risultati ottenuti in modalità multi-obiettivo (tre e quattro funzioni)

Come già visto nel Paragrafo 6.3.1 gli indici di qualità ottenuti nel caso di bacini considerati come non strumentati sono abbastanza soddisfacenti. Di seguito si riportano nuovamente i valori di tali indici per i casi di calibrazione *multi-obiettivo* e, relativamente alla modalità che ha prodotto i risultati migliori, si riportano gli idrogrammi sia in calibrazione che in validazione. Come già sottolineato i valori si riferiscono alla portata ottenuta con il set più vicino all'origine della *Frontiera Pareto* (indicata con Q_{CO}).

Nella Tabella 6.17 sono riportati i valori degli indici di qualità (indicati in lilla) ottenuti per il Fiume Candigliano in calibrazione con l'utilizzo delle funzioni obiettivo (indicate in grigio) ciascuna considerata in modalità *multi-obiettivo* secondo le espressioni matematiche riportate in precedenza. Il valore si riferisce appunto al set più vicino all'origine degli assi sulla *Frontiera Pareto*.

CANDIGLIANO (calibrazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$ -MO	$\mu-\sigma^2-Q_{355}$ -MO	$\mu-\sigma^2-\rho_1-Q_{355}$ -MO
EFF	0,62	0,7	0,66
MAE	4,09	3,34	3,68
MRE	0,9	0,61	0,61

Tabella 6.17. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.

I valori evidenziati (con sfondo azzurro) sono rispettivamente il valore massimo di EFF e minimi di MAE ed MRE che si sono ottenuti tra tutte le modalità *multi-obiettivo*. In questo caso la simulazione ottenuta tramite la combinazione *multi-obiettivo* $\mu\text{-}\sigma^2\text{-}Q_{355}$ sembra aver prodotto i migliori risultati. In Tabella 6.18 sono riportati gli stessi indici ottenuti però relativamente al periodo di validazione.

CANDIGLIANO (validazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}MO$	$\mu\text{-}\sigma^2\text{-}Q_{355}\text{-}MO$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-}MO$
EFF	0,48	0,55	0,49
MAE	3,46	2,66	3,13
MRE	0,84	0,64	0,65

Tabella 6.18. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.

Il confronto degli indici in merito al periodo di validazione conferma quanto sostenuto in precedenza e per questo in seguito sono riportati i grafici in calibrazione e in validazione per il Fiume Candigliano relativi alla combinazione *multi-obiettivo* $\mu\text{-}\sigma^2\text{-}Q_{355}$.

La Figura 6.20 riporta l'idrogramma in calibrazione mentre la Figura 6.21 si riferisce al periodo di validazione. L'idrogramma in calibrazione appare soddisfacente nel suo andamento e riproduce in modo adeguato tutti i fenomeni di magra al contrario di quanto accadeva con la simulazione *mono-obiettivo* equivalente (Figura 6.4). In questo caso infatti il processo di esaurimento delle piene viene simulato con una certa precisione anche se permangono le difficoltà nella riproduzione delle portate di picco. Analoghe considerazioni possono essere fatte per il periodo di validazione (Figura 6.21). Per avere un confronto diretto con quanto ottenuto nel caso di calibrazione *mono-obiettivo* (paragrafo 6.3.2), si riporta anche il grafico relativo agli involuipi massimo e minimo per una finestra temporale di 150 giorni intorno al periodo di massimo picco sia in calibrazione che in validazione ottenuti ancora una volta tramite l'utilizzo del set più vicino all'origine degli assi sulla *Frontiera Pareto*. Questa analisi più dettagliata consente di confermare che la riproduzione degli idrogrammi è soddisfacente; l'idrogramma in calibrazione ha un andamento meno oscillante rispetto a quello ottenuto in modalità *mono-obiettivo* con le medesime statistiche (Figura 6.6) e nel complesso genera una sottostima paragonabile della portata di picco. In validazione (Figura 6.23) l'idrogramma si presenta soddisfacente nella riproduzione

delle magra mentre anche in questo caso il picco è sottostimato in modo significativo seppur paragonabile alla modalità *mono-obiettivo*.

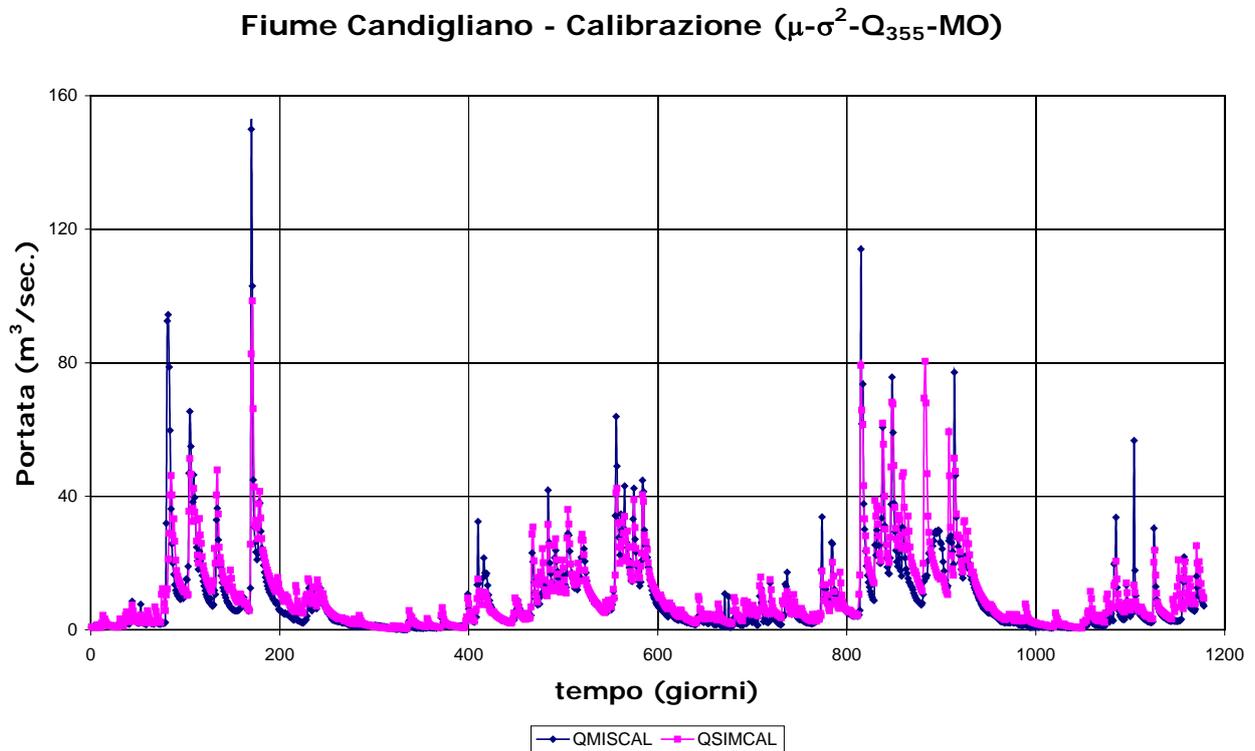


Figura 6.20. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu\text{-}\sigma^2\text{-}p_1\text{-MO}$).

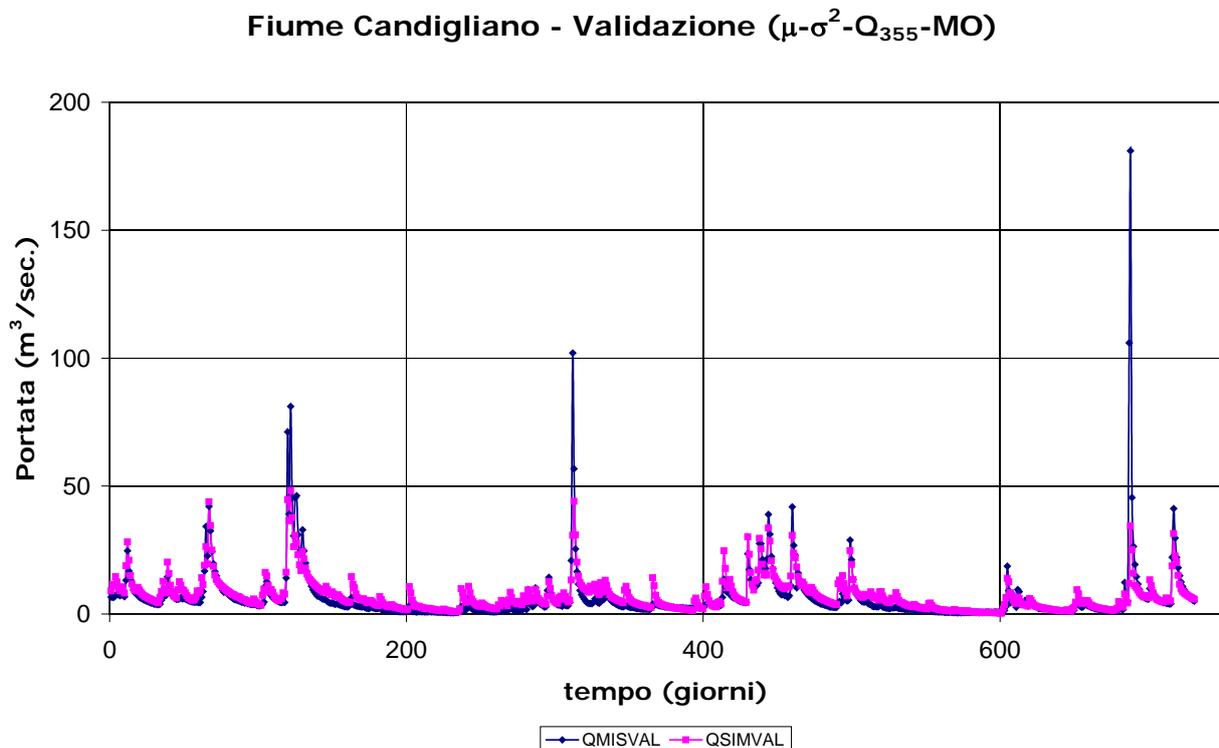


Figura 6.21. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu\text{-}\sigma^2\text{-}p_1\text{-MO}$).

Fiume Candigliano - Calibrazione ($\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$)

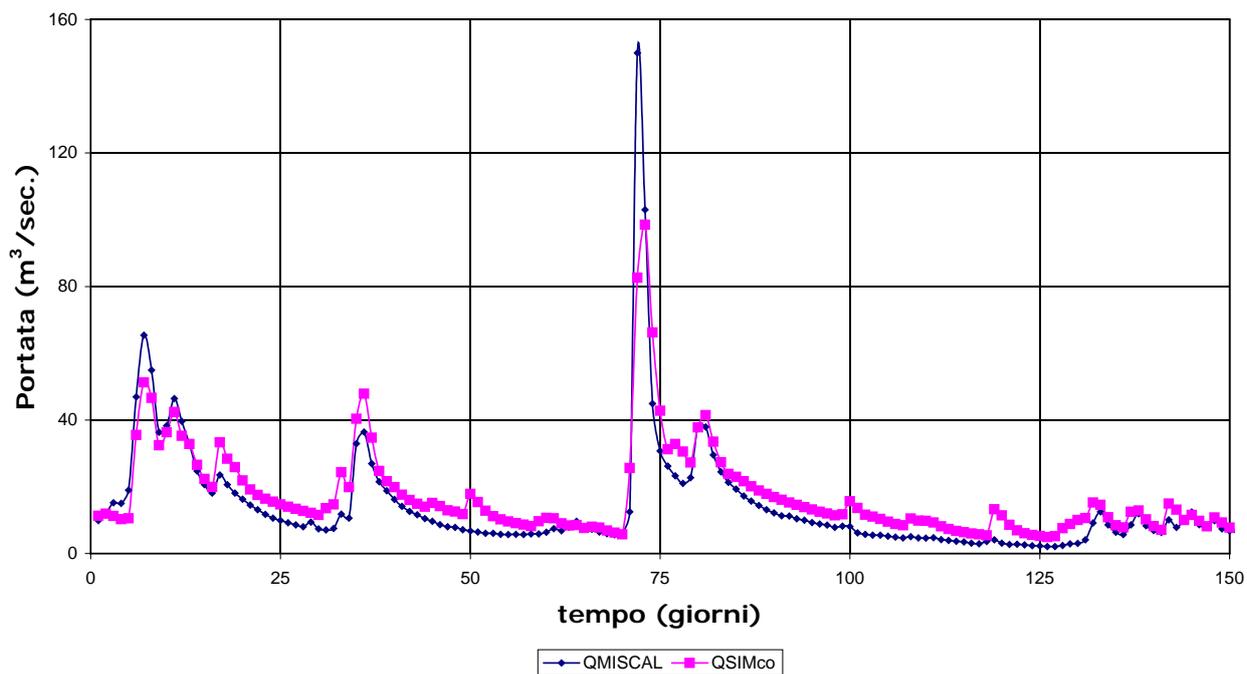


Figura 6.22. Idrogramma in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$).

Fiume Candigliano - Validazione ($\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$)

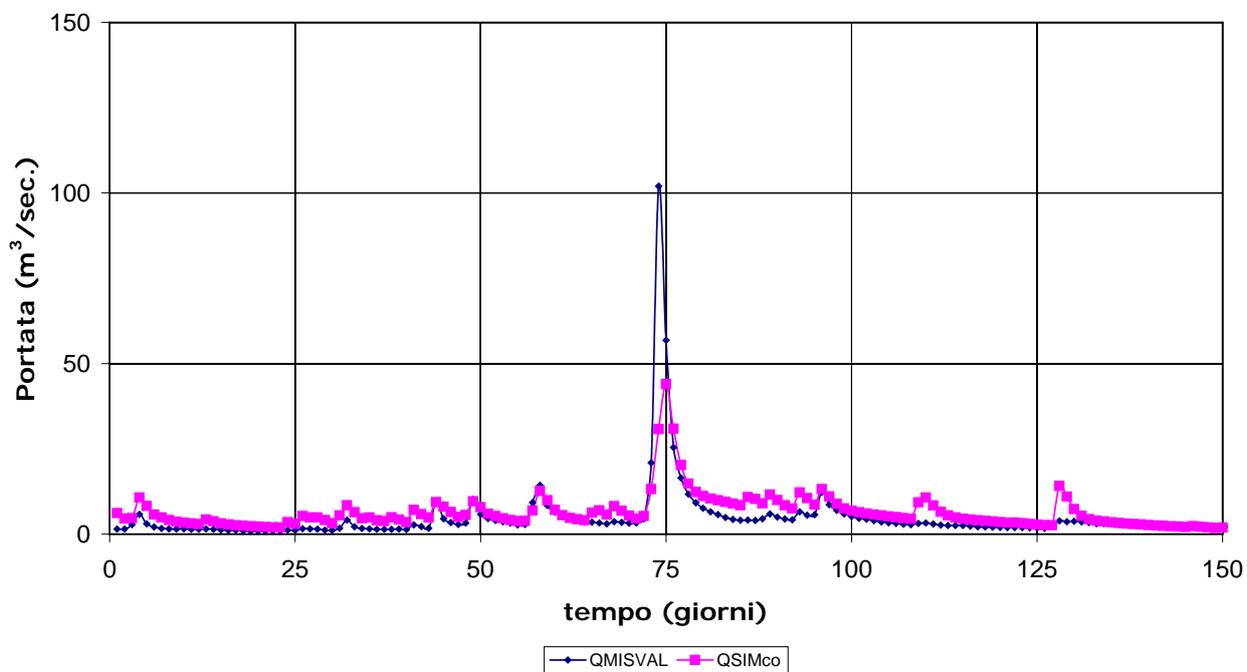


Figura 6.23. Idrogramma in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$).

Nella Tabella 6.19 sono riportati i valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con l'utilizzo delle funzioni obiettivo ciascuna considerata in modalità *multi-obiettivo*.

ESINO (calibrazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu-\sigma^2-\rho_1-MO$	$\mu-\sigma^2-Q_{355}-MO$	$\mu-\sigma^2-\rho_1-Q_{355}-MO$
EFF	0,51	0,29	0,53
MAE	7,47	8,87	7,27
MRE	0,5	0,6	0,52

Tabella 6.19. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.

In questo caso la simulazione migliore risulta quella ottenuta con l'utilizzo della combinazione di tutte e quattro le funzioni obiettivo $\mu-\sigma^2-\rho_1-Q_{355}$. A tale conclusione si arriva considerando anche i valori in validazione riportati in Tabella 6.20 poiché in calibrazione il valore del MRE risulta minore con la combinazione $\mu-\sigma^2-\rho_1$. In questo caso quindi si ottiene un risultato differente rispetto a quanto ottenuto con la modalità *mono-obiettivo* per la quale si erano ottenuti indici migliori con la modalità $\mu-\sigma^2-Q_{355}$ (Tabella 6.11 e Tabella 6.12). Gli idrogrammi tuttavia non risultavano particolarmente soddisfacenti pertanto sembra interessante esaminare quanto ottenuto in questo caso specifico.

ESINO (validazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu-\sigma^2-\rho_1-MO$	$\mu-\sigma^2-Q_{355}-MO$	$\mu-\sigma^2-\rho_1-Q_{355}-MO$
EFF	0,6	0,55	0,67
MAE	6,36	6,63	5,87
MRE	0,44	0,49	0,42

Tabella 6.20. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.

La Figura 6.24 riporta l'idrogramma in calibrazione ottenuto con le quattro funzioni obiettivo: da questa immagine risulta evidente che il modello riproduce bene, soprattutto in termini temporali, il primo picco di piena importante, mentre anticipa decisamente il secondo picco, riproducendolo con una serie di eventi veloci di entità

minore. Queste considerazioni sono confermate dall'analisi della Figura 6.26 che focalizza l'attenzione solamente sul secondo picco di portata. Gli altri eventi sembrano ben riprodotti anche se persistono alcune imprecisioni nei rami di esaurimento.

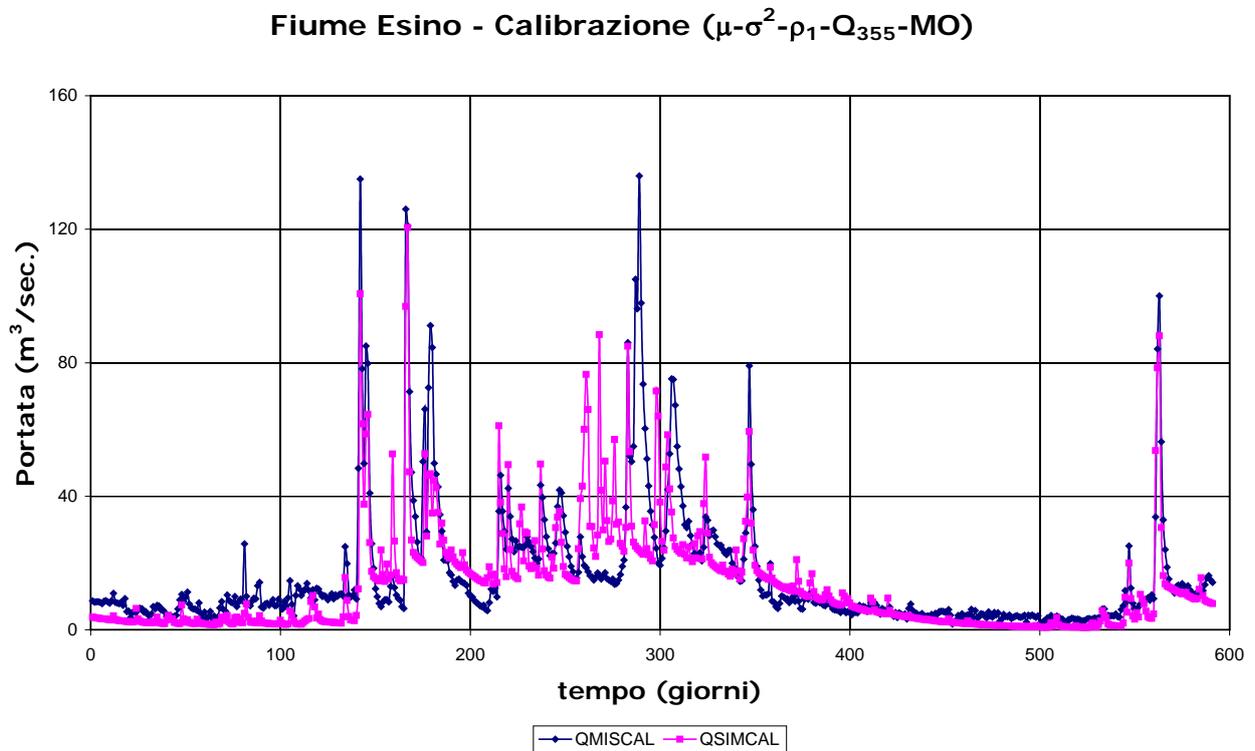


Figura 6.24. Idrogramma in calibrazione del Fiume Esino a Moie ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

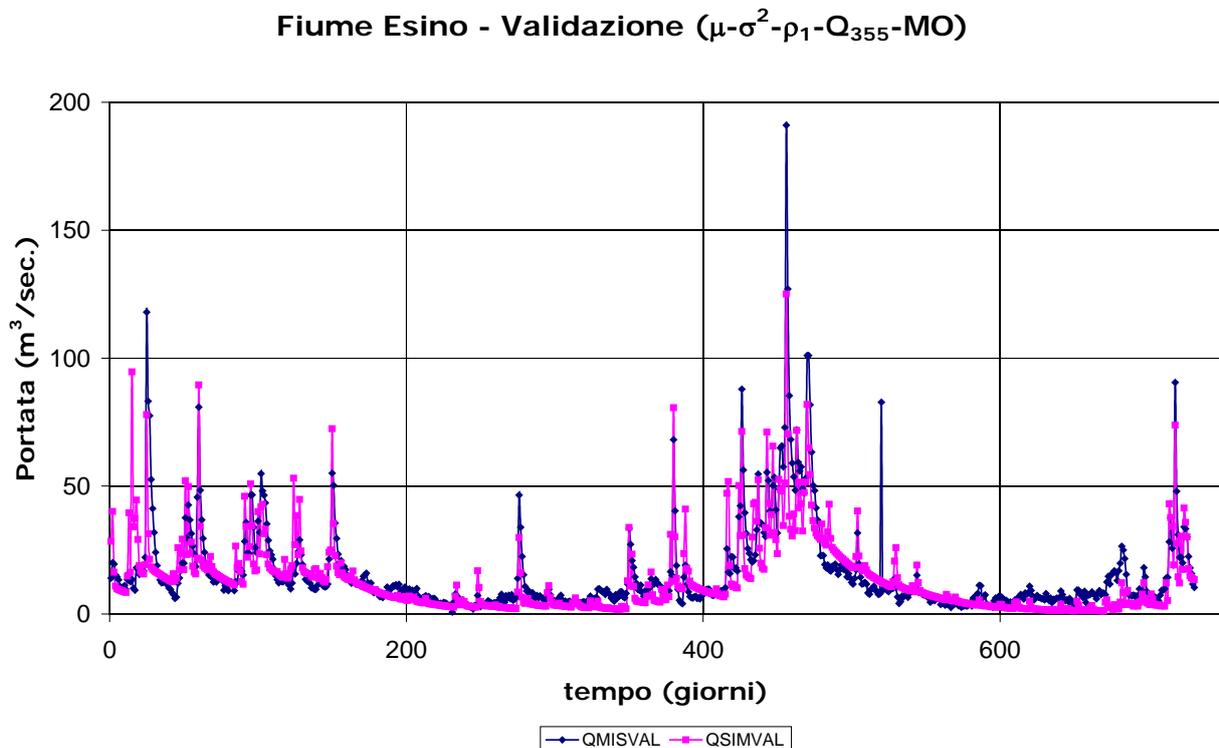


Figura 6.25. Idrogramma in validazione del Fiume Esino a Moie ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

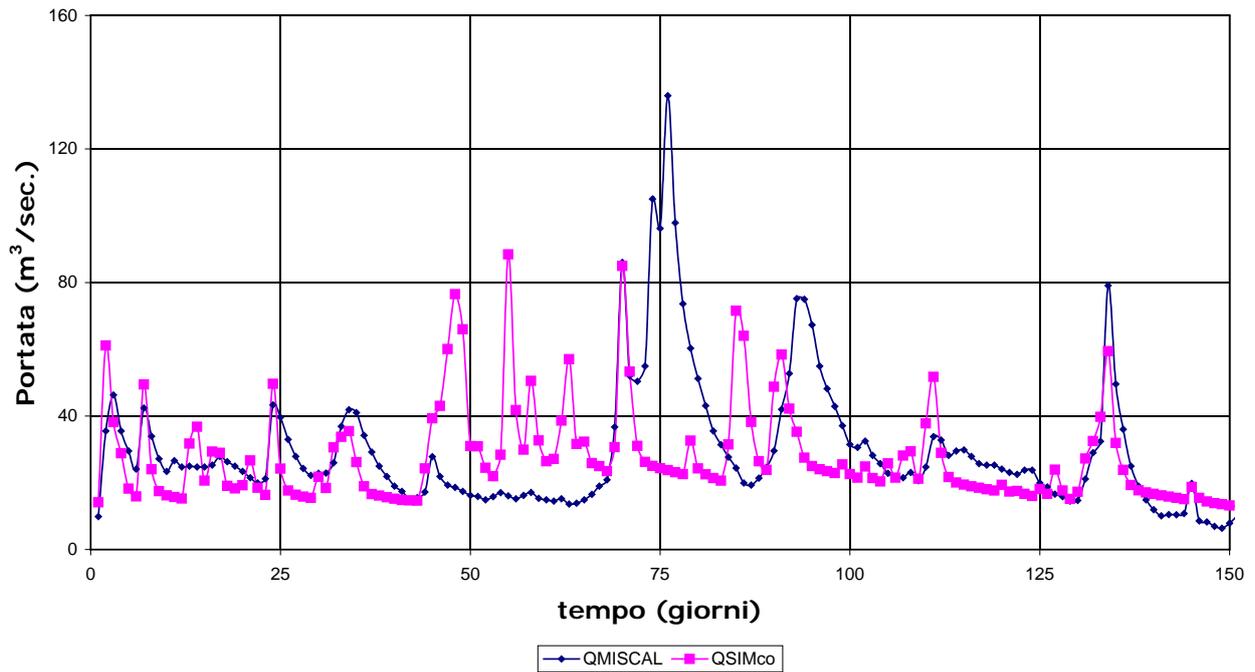
Fiume Esino - Calibrazione (μ - σ^2 - ρ_1 - Q_{355} -MO)

Figura 6.26. Idrogramma in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1 - Q_{355} -MO).

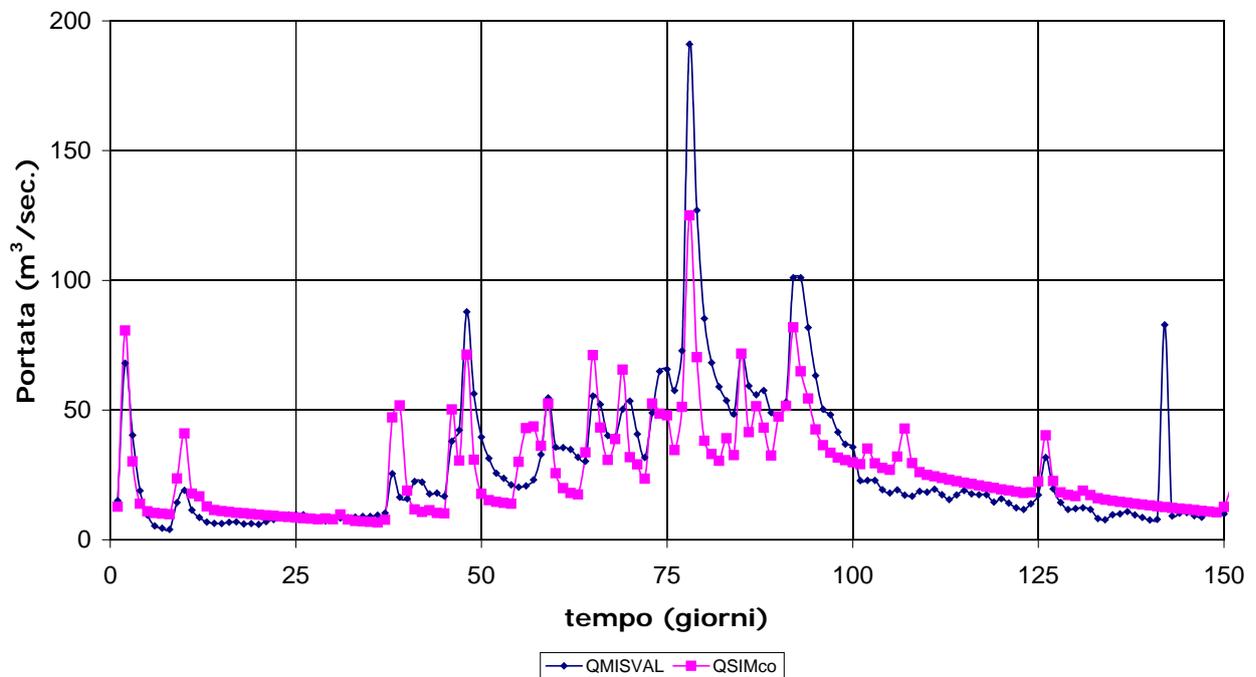
Fiume Esino - Validazione (μ - σ^2 - ρ_1 - Q_{355} -MO)

Figura 6.27. Idrogramma in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1 - Q_{355} -MO).

L'idrogramma in validazione (Figura 6.25) è più preciso in termini temporali e la finestra di 150 giorni attorno al picco conferma una discreta sensibilità del modello che

riproduce tutte le oscillazioni ad eccezione del picco verso la fine della finestra temporale. Permane una sottostima dei volumi già presente anche nella modalità *mono-obiettivo* (Figura 6.11).

Nella Tabella 6.21 sono riportati i valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con l'utilizzo delle funzioni obiettivo ciascuna considerata in modalità *multi-obiettivo*.

METAURO (calibrazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$	$\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$
EFF	0,54	0,24	0,64
MAE	10,71	12,98	7,82
MRE	1,64	0,96	0,63

Tabella 6.21. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.

In questo caso i valori degli indici di qualità sono nettamente a favore della modalità $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ sia in calibrazione che in validazione (Tabella 6.22) e pertanto la scelta ricade evidentemente su questa tipologia di combinazione di funzioni obiettivo. Anche per questo bacino quindi si è operata una scelta diversa rispetto a come proceduto nel caso di modalità *mono-obiettivo* secondo la quale la combinazione migliore era risultata quella che utilizza solamente $\mu\text{-}\sigma^2\text{-}\rho_1$ a meno dell'indice di magra.

METAURO (validazione-valori riferiti a Q_{CO})			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$	$\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$
EFF	0,53	0,22	0,63
MAE	10,69	12,88	8,05
MRE	1,70	1,02	0,61

Tabella 6.22. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.

In merito ai valori degli indici è da sottolineare che anche in questo caso i valori di MAE risultano particolarmente elevati.

La Figura 6.28 riporta l'idrogramma in calibrazione mentre la Figura 6.29 si riferisce al periodo di validazione. Entrambi gli idrogrammi appaiono decisamente migliori di quelli ottenuti nel caso di modalità *mono-obiettivo* (Figura 6.12 e Figura 6.13) poiché i

rami di esaurimento sono ora riprodotti in modo decisamente più adeguato e la portata di magra viene riprodotta dal modello in modo fedele.

Fiume Metauro - Calibrazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$)

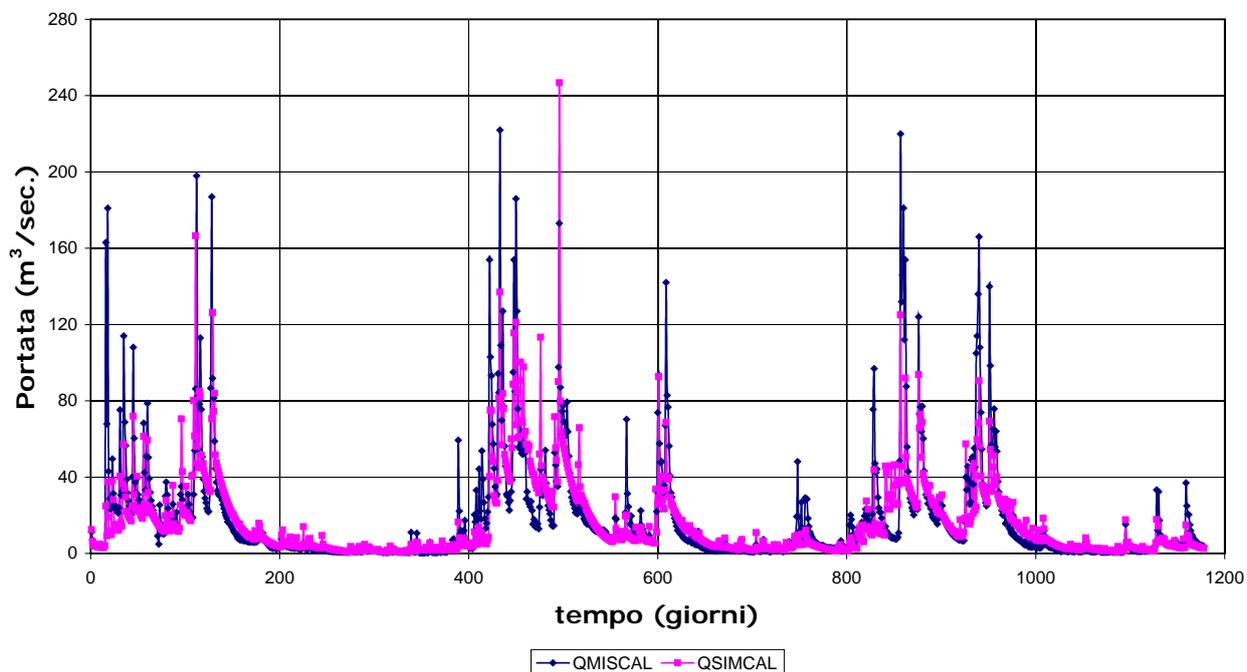


Figura 6.28. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

Fiume Metauro - Validazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$)

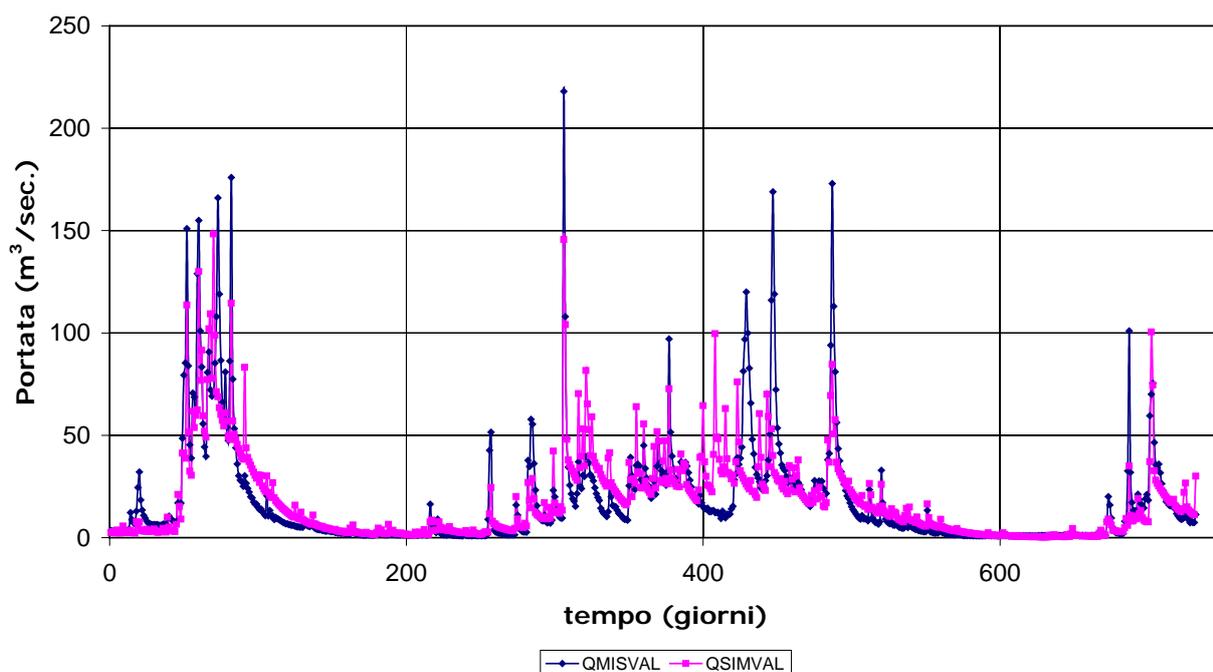


Figura 6.29. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

Fiume Metauro - Calibrazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$)

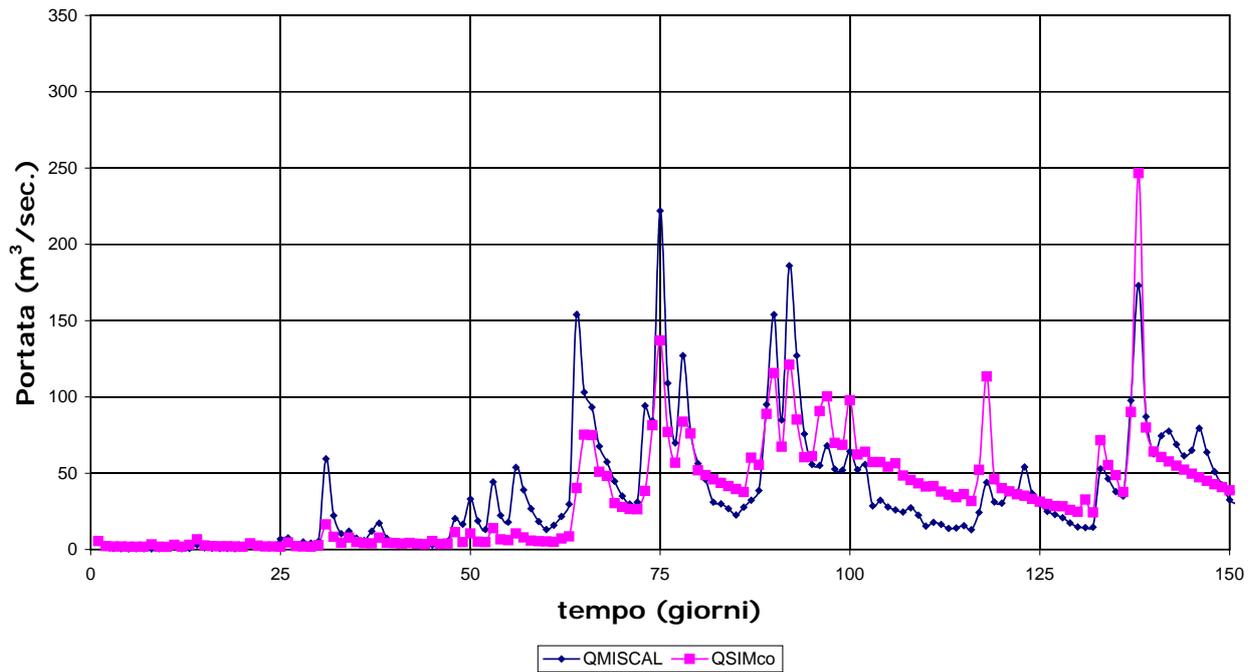


Figura 6.30. Idrogramma in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

Fiume Metauro - Validazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$)

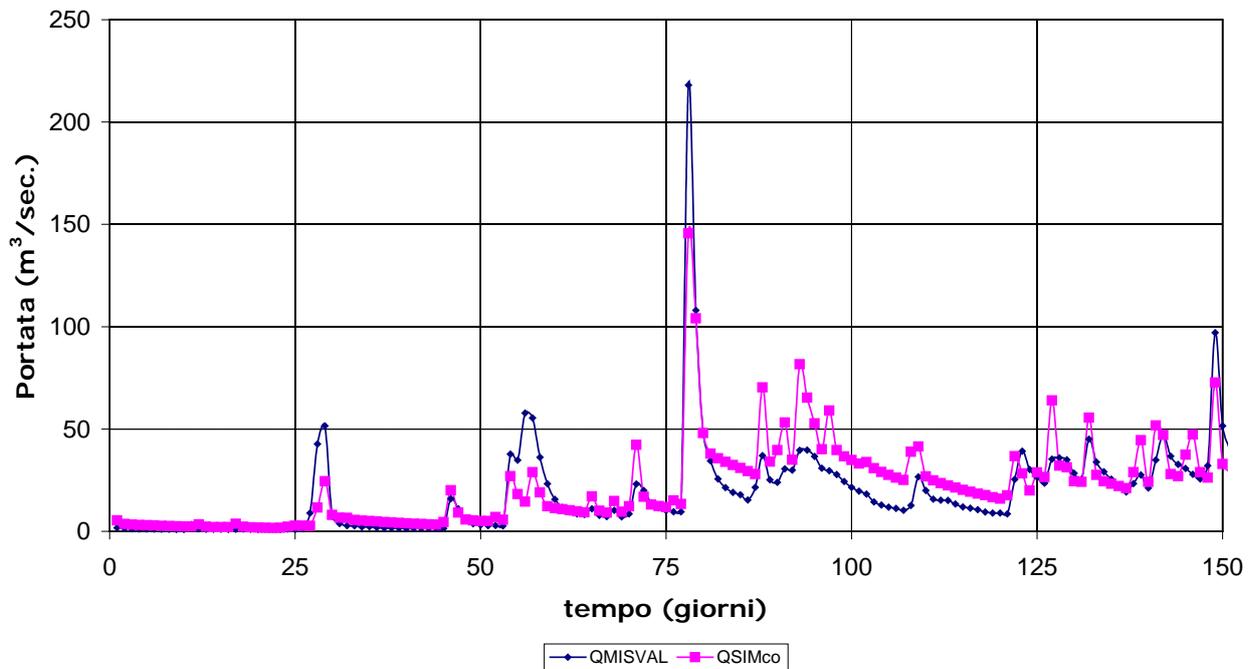


Figura 6.31. Idrogramma in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$).

Come sottolineato anche per il Fiume Esino, il poter considerare tutte e quattro le grandezze statistiche in modalità *multi-obiettivo*, sembra permettere una simulazione migliore delle portate di magra e dei processi di esaurimento delle piene.

Permangono le difficoltà nella riproduzione delle portate di picco che sfuggono praticamente sempre alla simulazione del modello là dove il valore di circa 100 m³/sec viene superato. Le immagini riportate in Figura 6.30 e in Figura 6.31 confermano quanto appena sottolineato.

Nella Tabella 6.23 sono riportati i valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con l'utilizzo delle funzioni obiettivo ciascuna considerata in modalità *multi-obiettivo*.

POTENZA (calibrazione-valori riferiti a Q _{CO})			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$	$\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$
EFF	0,33	0,52	0,44
MAE	3,07	2,96	3,08
MRE	0,46	0,49	0,53

Tabella 6.23. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.

POTENZA (validazione-valori riferiti a Q _{CO})			
Funzione Obiettivo	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$	$\mu\text{-}\sigma^2\text{-}Q_{355}\text{-MO}$	$\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}\text{-MO}$
EFF	-0,82	-0,72	-0,87
MAE	3,14	3,84	3,84
MRE	0,55	0,71	0,71

Tabella 6.24. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.

In questo caso i valori degli indici di qualità non permettono una veloce identificazione della combinazione migliore. La modalità $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$ a 4 funzioni sembra da escludere in partenza mentre la modalità $\mu\text{-}\sigma^2\text{-}Q_{355}$ sembrerebbe accettabile in considerazione dei valori di EFF e MAE. In realtà i valori in validazione (Tabella 6.24) ribaltano la situazione in favore della combinazione $\mu\text{-}\sigma^2\text{-}\rho_1$ che si dimostra migliore in termini di MAE ed MRE. I valori di EFF in validazione sono tutti pessimi raggiungendo addirittura valori negativi. Questo accadeva però anche con la modalità *mono-*

obiettivo (Tabella 6.16) pertanto, al fine di poter effettuare un confronto diretto, si sceglie di riportare la simulazione ottenuta con la modalità $\mu\text{-}\sigma^2\text{-}\rho_1$.

Fiume Potenza - Calibrazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$)

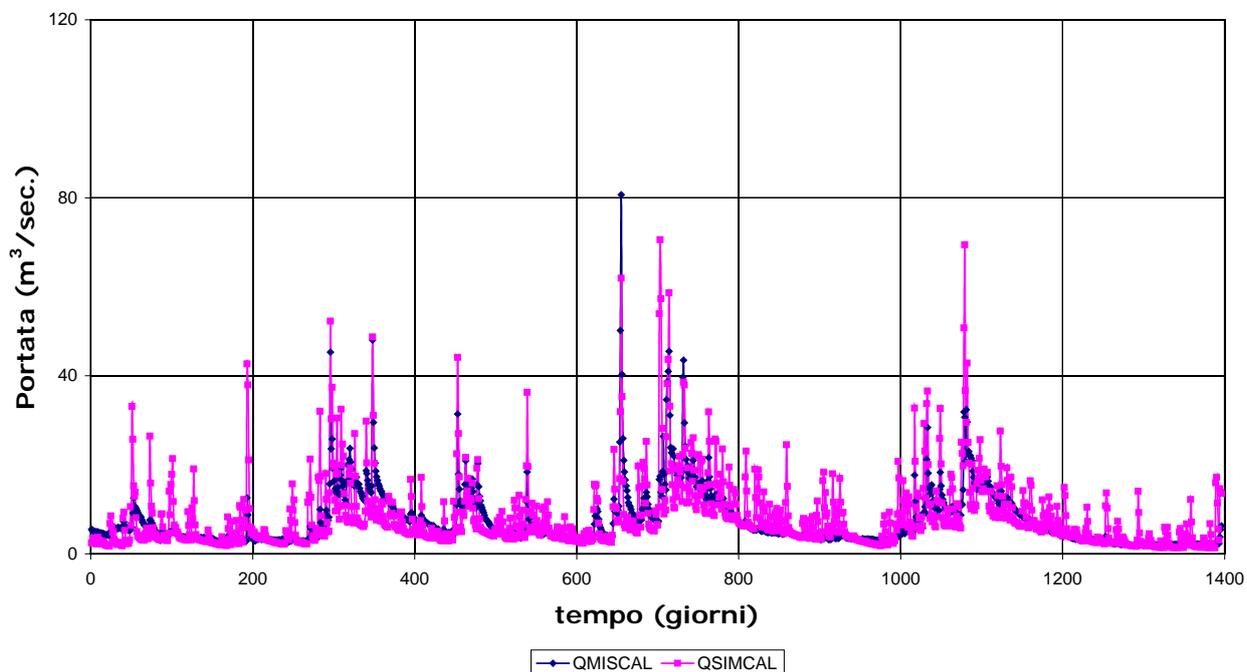


Figura 6.32. Idrogramma in calibrazione del Fiume Potenza a Cannucciaro ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$).

Fiume Potenza - Validazione ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$)

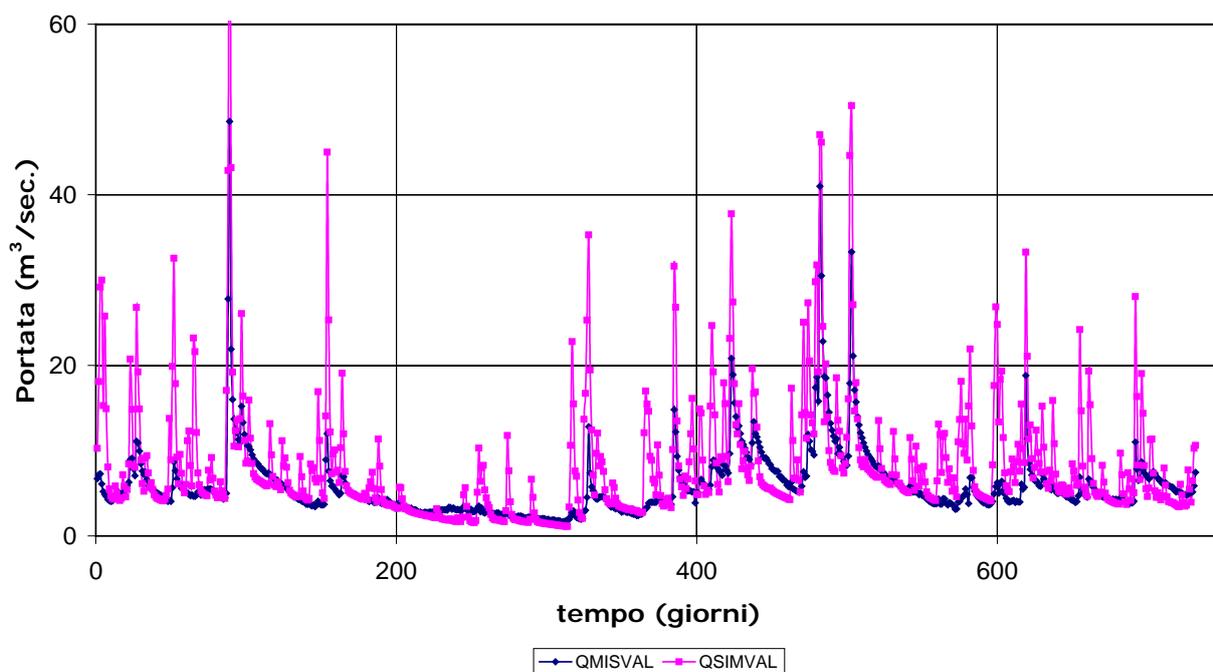


Figura 6.33. Idrogramma in validazione del Fiume Potenza a Cannucciaro ($\mu\text{-}\sigma^2\text{-}\rho_1\text{-MO}$).

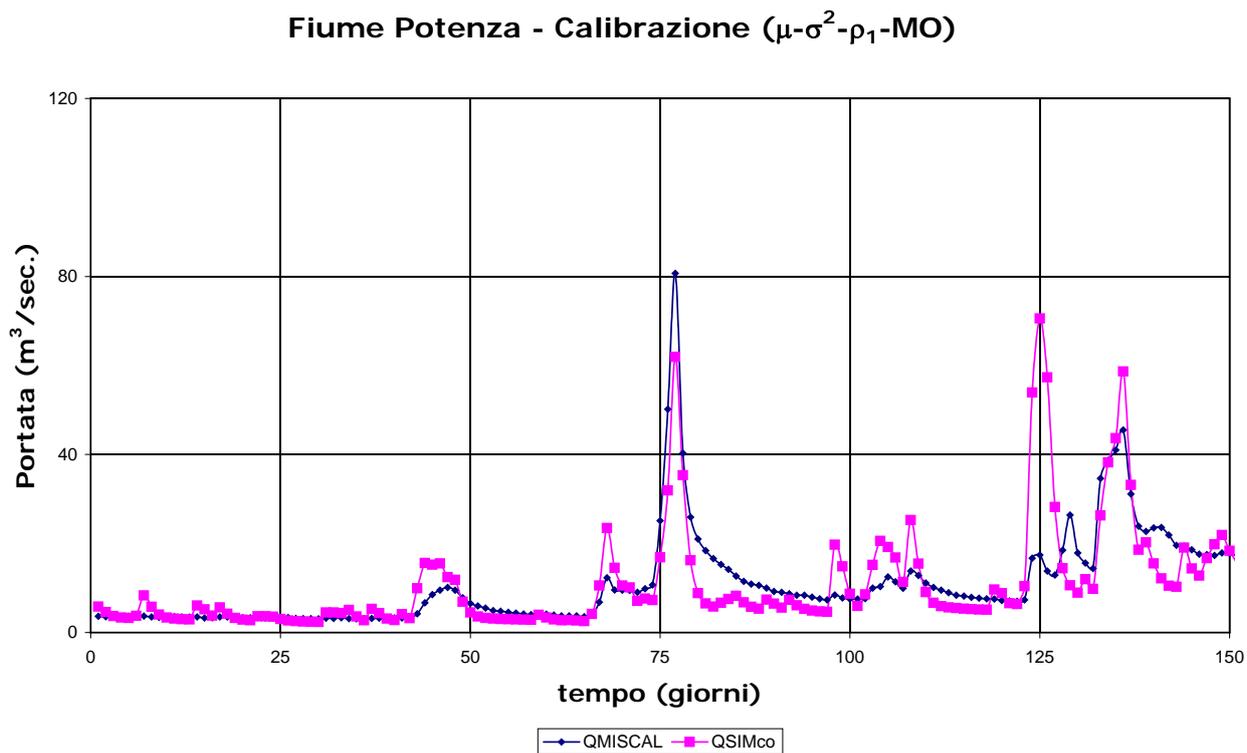


Figura 6.34. Idrogramma in calibrazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1 -MO).

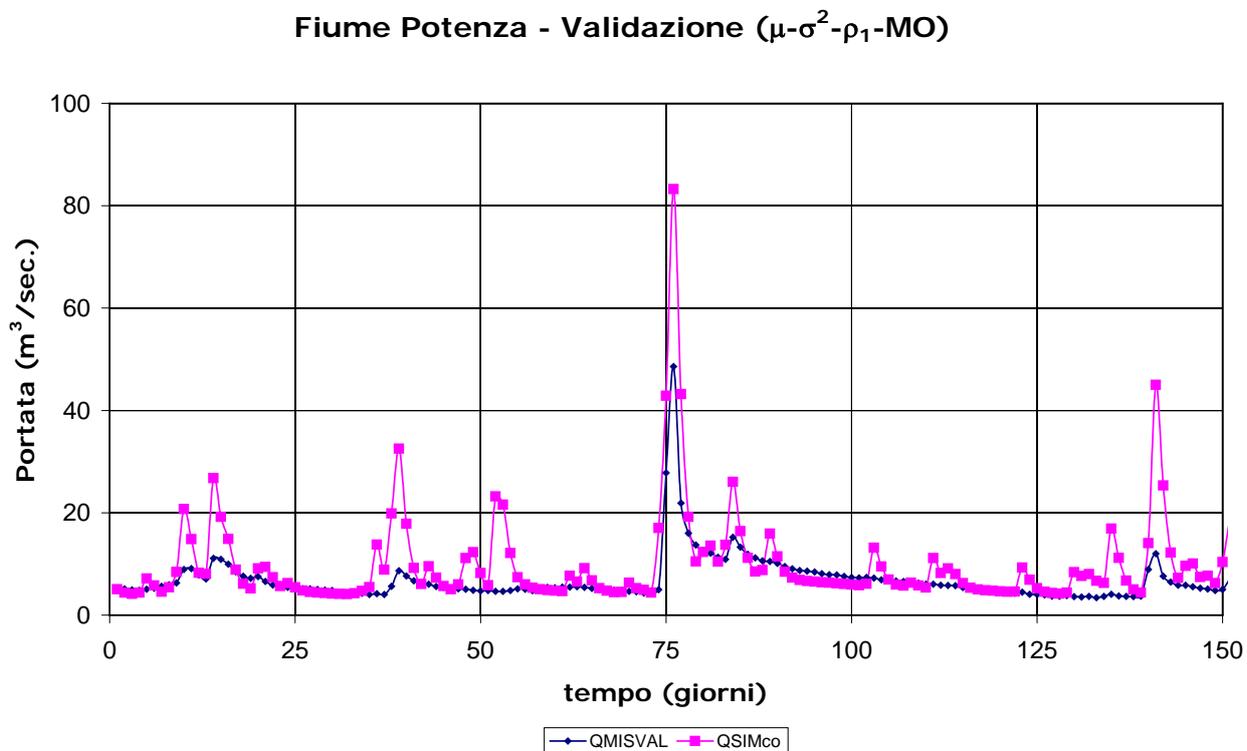


Figura 6.35. Idrogramma in validazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato (μ - σ^2 - ρ_1 -MO).

La Figura 6.32 riporta l'idrogramma in calibrazione mentre la Figura 6.33 si riferisce al periodo di validazione. Le conclusioni che si possono elaborare sono molto simili a

quelle tratte per il caso *mono-obiettivo*. Entrambi gli idrogrammi seguono nel complesso l'andamento delle portate osservate ma è evidente anche in questo caso che la portata simulata ha oscillazioni molto più frequenti rispetto alla portata osservata. L'esame degli idrogrammi focalizzati attorno al periodo di maggior picco (Figura 6.34 e Figura 6.35) confermano questa osservazione. In aggiunta a quanto già sottolineato per la modalità *mono-obiettivo* in merito alla capacità del modello di riprodurre i deflussi, possiamo solamente sottolineare come per questo bacino, al contrario di ciò che avviene per gli altri, i picchi di piena siano addirittura sovrastimati (soprattutto in validazione - Figura 6.33).

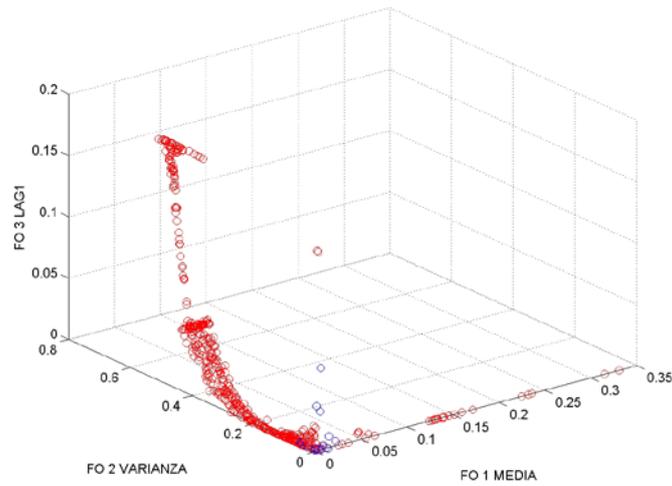
6.3.4 Confronto modalità mono e multi-obiettivo: incidenza delle statistiche

La modalità di calibrazione *multi-obiettivo* fornisce un numero di set dei parametri ugualmente possibili che formano la *Frontiera Pareto*. Non tutti i set in realtà forniscono risultati che si dimostrano eccellenti in validazione poiché ciascuno di essi è un compromesso tra le varie funzioni obiettivo che a propri sono tutte ugualmente importanti. Al fine di approfondire l'analisi ed il confronto delle modalità *mono* e *multi-obiettivo* anche nel caso di bacini non strumentati e di capire se una statistica può avere in realtà un peso più o meno significativo rispetto alle altre, si è ritenuto interessante indagare la posizione dei set ottenuti dalla calibrazione *multi-obiettivo* che forniscono un risultato migliore rispetto all'unico set ottenuto dalla modalità *mono-obiettivo*. Questa analisi è stata condotta per tutti i bacini ma in questa sede, a titolo di esempio, si riportano i risultati relativi al solo Fiume Candigliano che, avendo un comportamento regolare, può essere utilizzato per l'analisi dell'influenza delle statistiche.

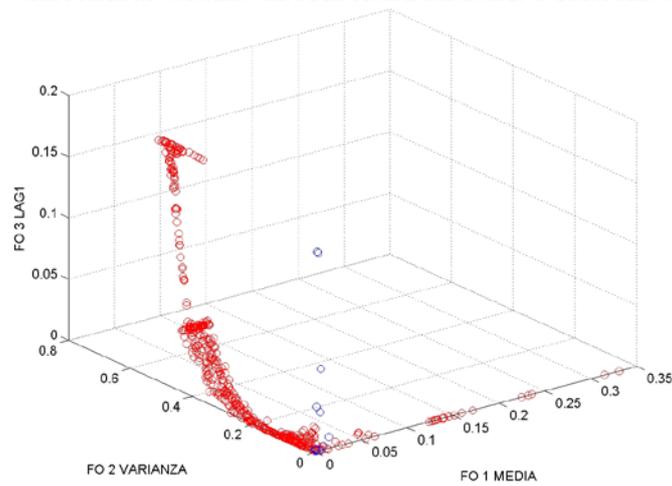
Analizzando le immagini a partire dalla Figura 6.36 fino alla Figura 6.38 è possibile notare che, in tutte le combinazioni di funzioni obiettivo, i set con indici di qualità migliori rispetto al caso *mono-obiettivo* che utilizza le medesime statistiche, sono sommariamente in prossimità dell'origine degli assi per quanto riguarda tutte le statistiche ad eccezione del coefficiente di correlazione lineare a ritardo unitario.

Questo significa che la minimizzazione di questa statistica non è strettamente necessaria per ottenere una simulazione soddisfacente dal punto di vista degli indici di qualità presi in esame.

FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON EFF MAGGIORE RISPETTO AL MONO-OBIETTIVO



FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON MAE MINORE RISPETTO AL MONO-OBIETTIVO



FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON MRE MINORE RISPETTO AL MONO-OBIETTIVO

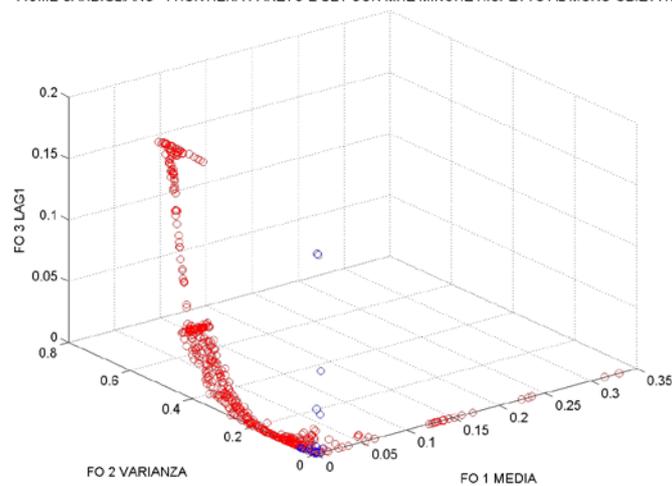
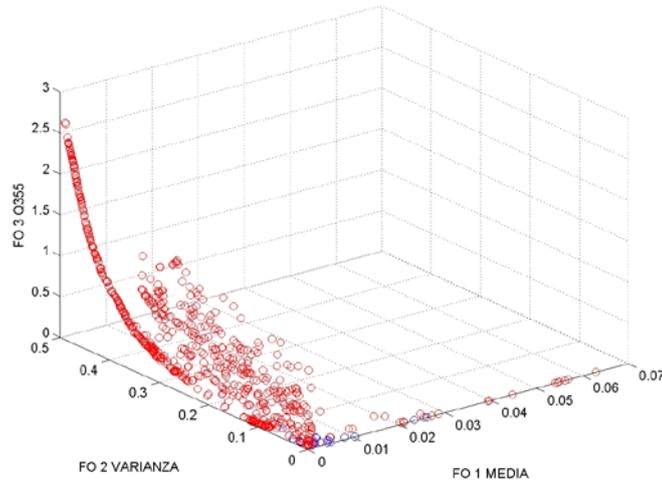
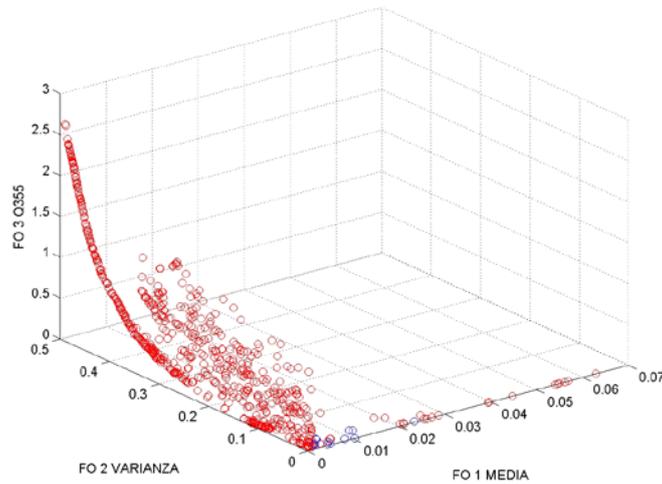


Figura 6.36. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo $\mu\text{-}\sigma^2\text{-}\rho_1$. I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.

FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON EFF MAGGIORE RISPETTO AL MONO-OBIETTIVO



FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON MAE MINORE RISPETTO AL MONO-OBIETTIVO



FIUME CANDIGLIANO - FRONTIERA PARETO E SET CON MRE MINORE RISPETTO AL MONO-OBIETTIVO

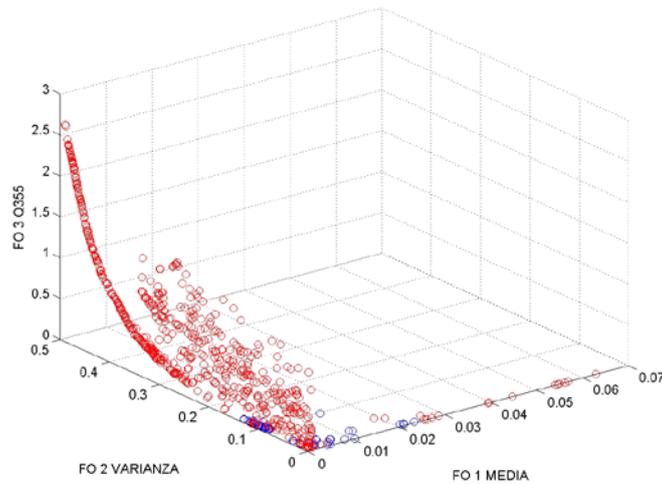


Figura 6.37. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo μ - σ^2 - Q_{355} . I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.

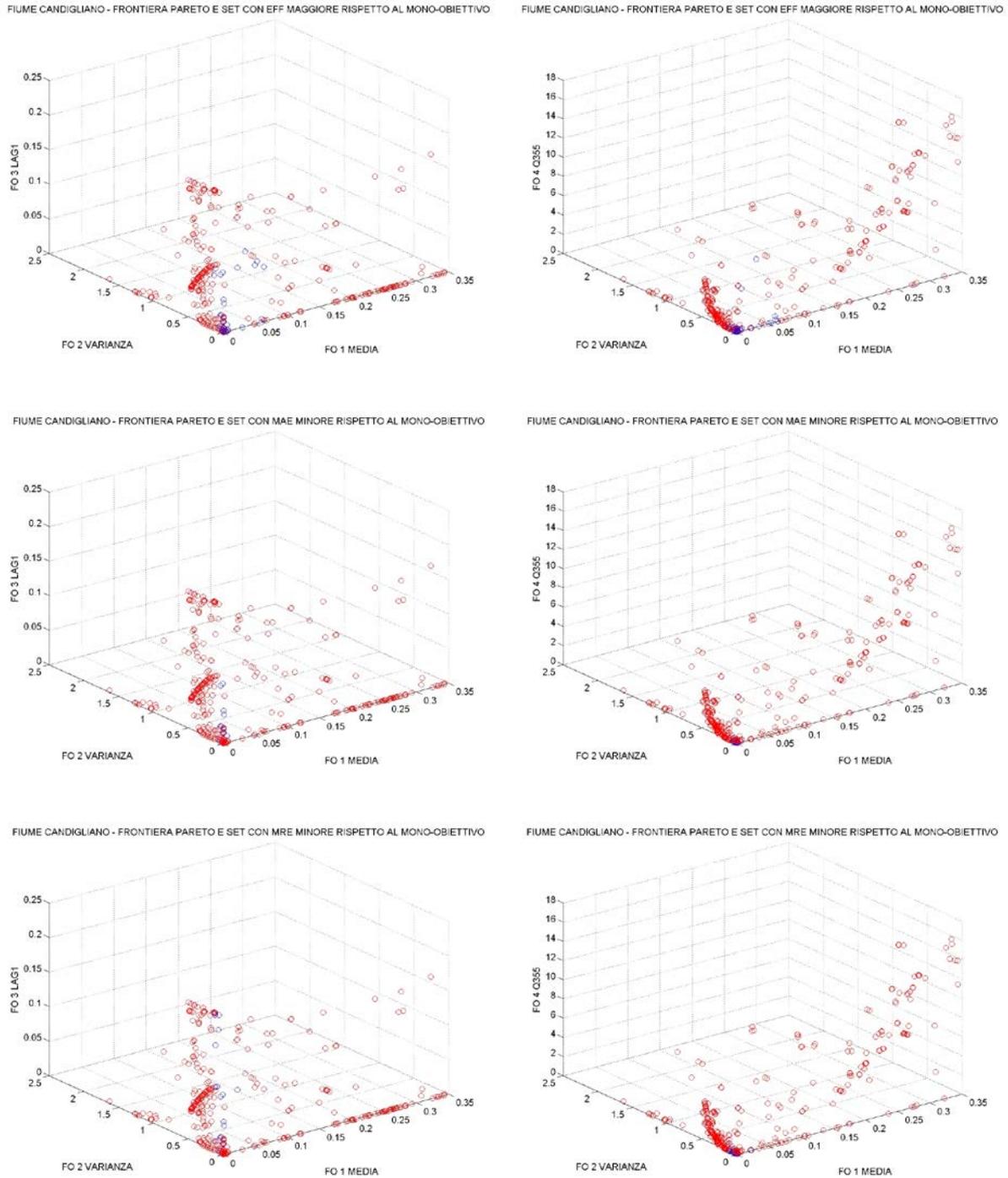


Figura 6.38. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo $\mu\text{-}\sigma^2\text{-}\rho_1\text{-}Q_{355}$. I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.

Da queste immagini possiamo anche fare alcune considerazioni sulla forma della *Frontiera Pareto*: nel caso di due (e analogamente tre o quattro) funzioni da minimizzare, la forma della Frontiera dovrebbe essere simile a quella mostrata in Figura 6.3. Solo le simulazioni con la combinazione $\mu\text{-}\sigma^2\text{-}\rho_1$ hanno un andamento che in parte può somigliare a quello teorico, mentre le altre combinazioni si presentano prevalentemente come una nuvola di punti pressoché dispersa. Fa in parte eccezione

la *Frontiera* ottenuta con quattro funzioni obiettivo che, nel grafico dove non viene riportato il coefficiente di correlazione lineare, accenna in parte un andamento simile a quello teorico (Figura 6.38 - immagini di destra).

Sempre in queste immagini però si nota come i valori della funzione obiettivo relativa alla Q_{355} raggiungano valori pari anche a 13, 14. Ricordando l'espressione matematica di tale funzione obiettivo (equazione 6.24), tale valore indica che la differenza adimensionalizzata tra il valore della statistica simulata dal modello e quello regionalizzato è molto elevata. Questa osservazione ci dice in pratica che l'algoritmo ha incluso nella *Frontiera* anche punti che in realtà non minimizzano questa funzione obiettivo probabilmente perché, contemporaneamente, i valori delle altre tre funzioni obiettivo risultavano particolarmente bassi indicando di fatto un conflitto tra le statistiche. Questo fenomeno si presenta in modo molto ridotto nelle combinazioni a tre funzioni obiettivo dimostrando quindi che in tal caso l'algoritmo riesce a garantire la minimizzazione di tutte le funzioni obiettivo.

6.3.5 Analisi dei risultati relativi alla riproduzione delle portate

Tramite l'utilizzo dell'algoritmo di ottimizzazione AMALGAM abbiamo ottenuto più set appartenenti alla *Frontiera Pareto* ipotizzando di considerare i bacini dell'area di studio come non Strumentati. Per effettuare un confronto diretto con la modalità *mono-obiettivo* (realizzata con l'utilizzo di una funzione data dalla combinazione pesata, con pesi equi, delle statistiche regionalizzate) abbiamo selezionato il set sulla *Frontiera Pareto* più vicino all'origine degli assi. Tuttavia, tenendo conto del fatto che tutti i set ottenuti sono ugualmente probabili e possibili, può risultare utile analizzare gli involuppi massimi e minimi delle portate ottenibili da tali set. Questa analisi è stata condotta per tutti i bacini facendo particolare attenzione al fatto che le difficoltà riproduttive di tale portata si presentassero principalmente in periodi di scarsi deflussi o di deflussi abbondanti. Al fine della presente trattazione si definiscono deflussi di magra quelli con valore inferiore alla portata media su tutto il periodo di dati disponibili e viceversa si definiscono deflussi di piena i valori di portata al di sopra di tale valore medio.

La Tabella 6.25 e la Tabella 6.26 riportano le percentuali di portata esclusa dagli involuppi massimo e minimo, ottenuti in calibrazione e validazione rispettivamente, tramite tutti i set della *Frontiera Pareto* generati da AMALGAM. In particolare il valore

fuori dalla parentesi indica la percentuale totale di portate escluse mentre il valore dentro alla parentesi si riferisce alla percentuale esclusa in periodi considerati di scarsi deflussi.

Percentuale di portata esclusa dagli inviluppi (calibrazione)			
Funzioni Obiettivo	$\mu-\sigma^2-\rho_1$ -MO	$\mu-\sigma^2-Q_{355}$ -MO	$\mu-\sigma^2-\rho_1-Q_{355}$ -MO
Candigliano	32 % (27 %)	27 % (21 %)	19 % (17 %)
Esino	32 % (19 %)	35 % (24 %)	18 % (9 %)
Metauro	48 % (40 %)	25 % (19 %)	8 % (5 %)
Potenza	24 % (11 %)	42 % (30 %)	7 % (1 %)

Tabella 6.25. Percentuale esclusa dagli inviluppi di portata massimo e minimo ottenuti in calibrazione da tutti i set appartenenti alla Frontiera Pareto. Il valore tra parentesi indica la percentuale di portata di magra esclusa dagli inviluppi medesimi.

Sono state evidenziate in azzurro quelle caselle nelle quali il valore escluso in periodi di magra è superiore alla metà dei deflussi esclusi in totale al fine di sottolineare che le maggiori difficoltà si manifestano nella riproduzione delle magre piuttosto che delle piene.

Percentuale di portata esclusa dagli inviluppi (validazione)			
Funzioni Obiettivo	$\mu-\sigma^2-\rho_1$ -MO	$\mu-\sigma^2-Q_{355}$ -MO	$\mu-\sigma^2-\rho_1-Q_{355}$ -MO
Candigliano	42 % (37 %)	39 % (36 %)	33 % (31 %)
Esino	17 % (10 %)	23 % (18 %)	6 % (3 %)
Metauro	45 % (38 %)	53 % (34 %)	10 % (7 %)
Potenza	30 % (16 %)	24 % (11 %)	6 % (2 %)

Tabella 6.26. Percentuale esclusa dagli inviluppi di portata massimo e minimo ottenuti in validazione da tutti i set appartenenti alla Frontiera Pareto. Il valore tra parentesi indica la percentuale di portata di magra esclusa dagli inviluppi medesimi.

Particolarmente interessanti appaiono i risultati ottenuti per il Fiume Metauro e per il Fiume Potenza per i quali si arriva in calibrazione ad ottenere che la percentuale di portata esclusa dagli inviluppi è rispettivamente ridotta al 8% e 7%.

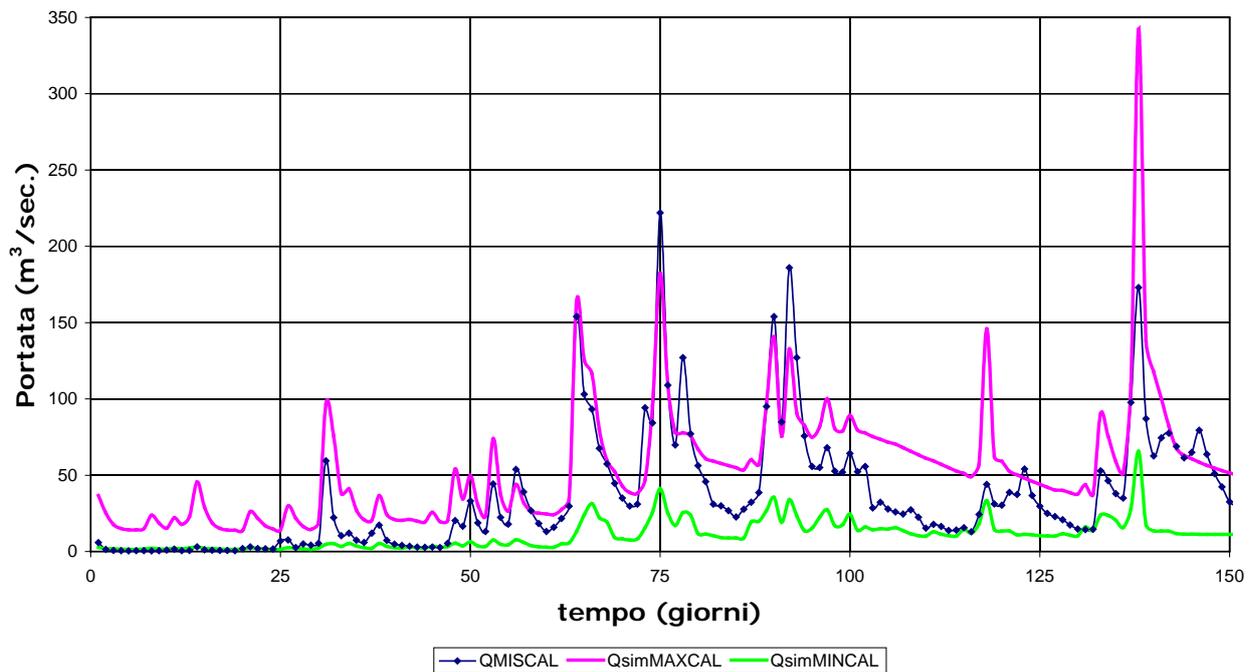
Fiume Metauro - Calibrazione (μ - σ^2 - ρ_1 -MO)

Figura 6.39. Fiume Metauro: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 . (periodo di 150 giorni attorno al massimo picco di portata).

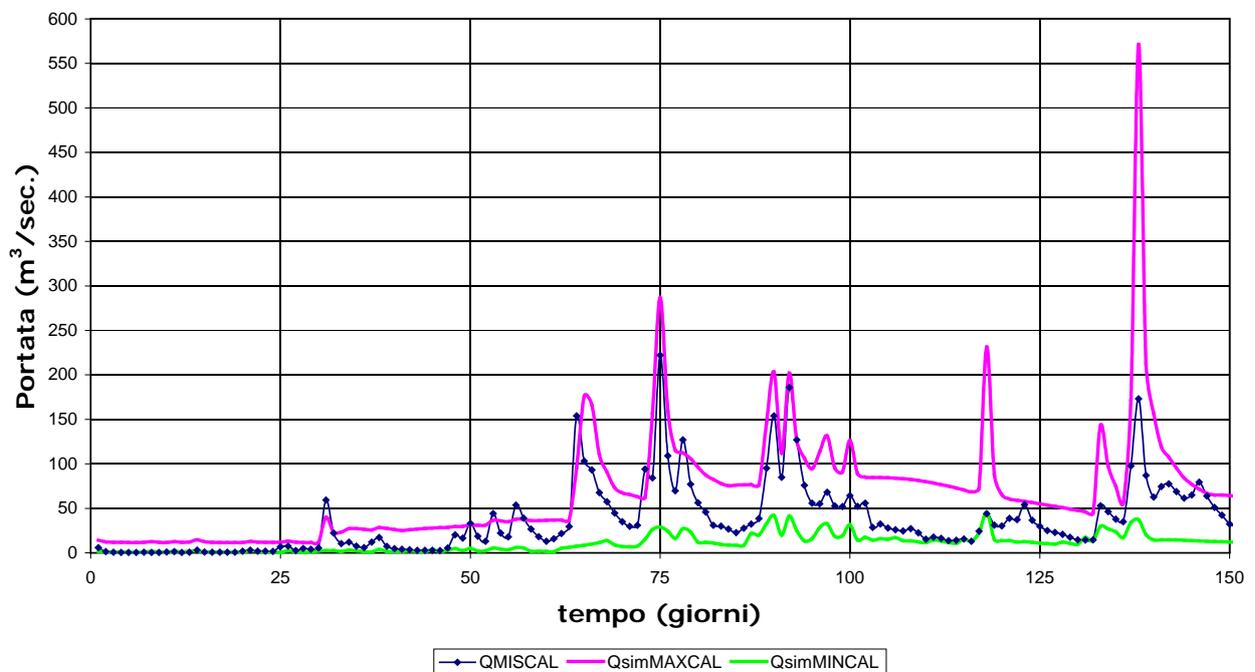
Fiume Metauro - Calibrazione (μ - σ^2 - Q_{355} -MO)

Figura 6.40. Fiume Metauro: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata).

Fiume Metauro - Calibrazione (μ - σ^2 - ρ_1 - Q_{355} -MO)

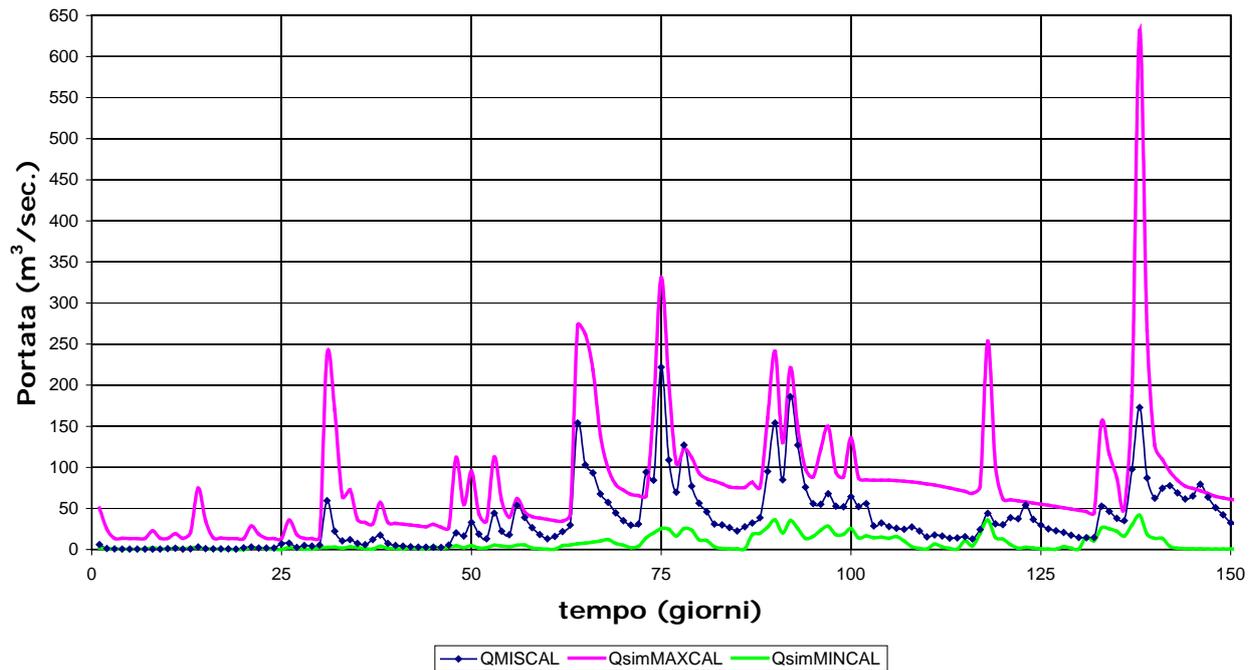


Figura 6.41. Fiume Metauro: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata).

Fiume Potenza - Calibrazione (μ - σ^2 - ρ_1 -MO)

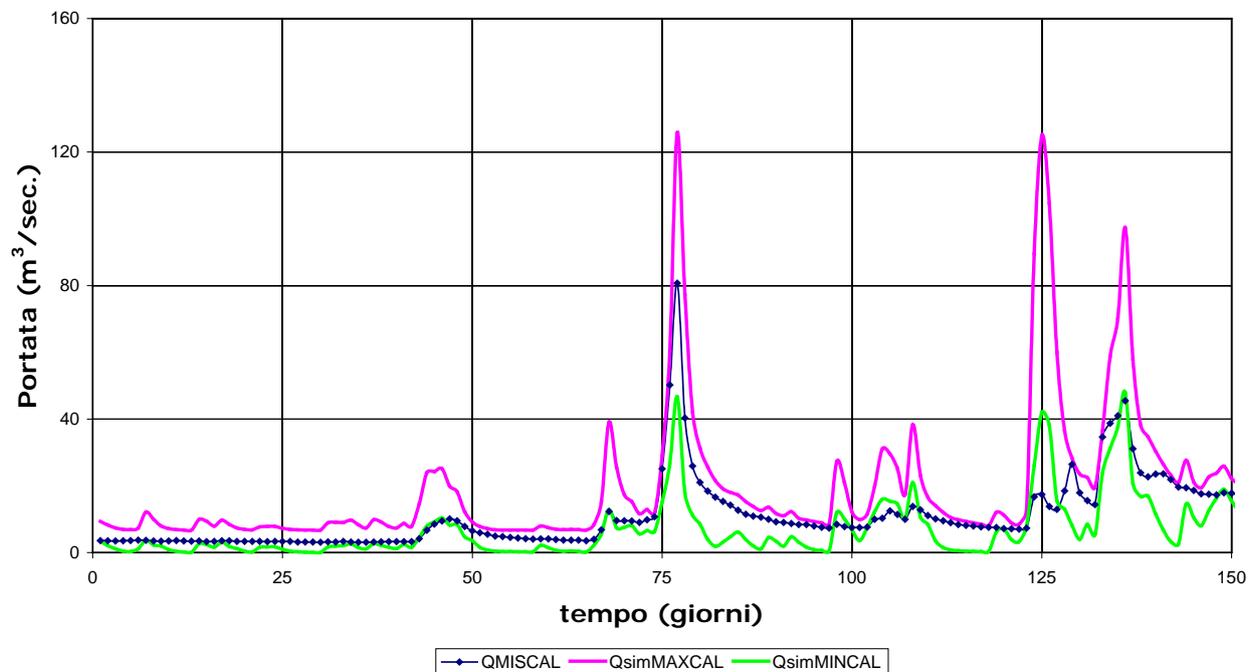


Figura 6.42. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 . (periodo di 150 giorni attorno al massimo picco di portata).

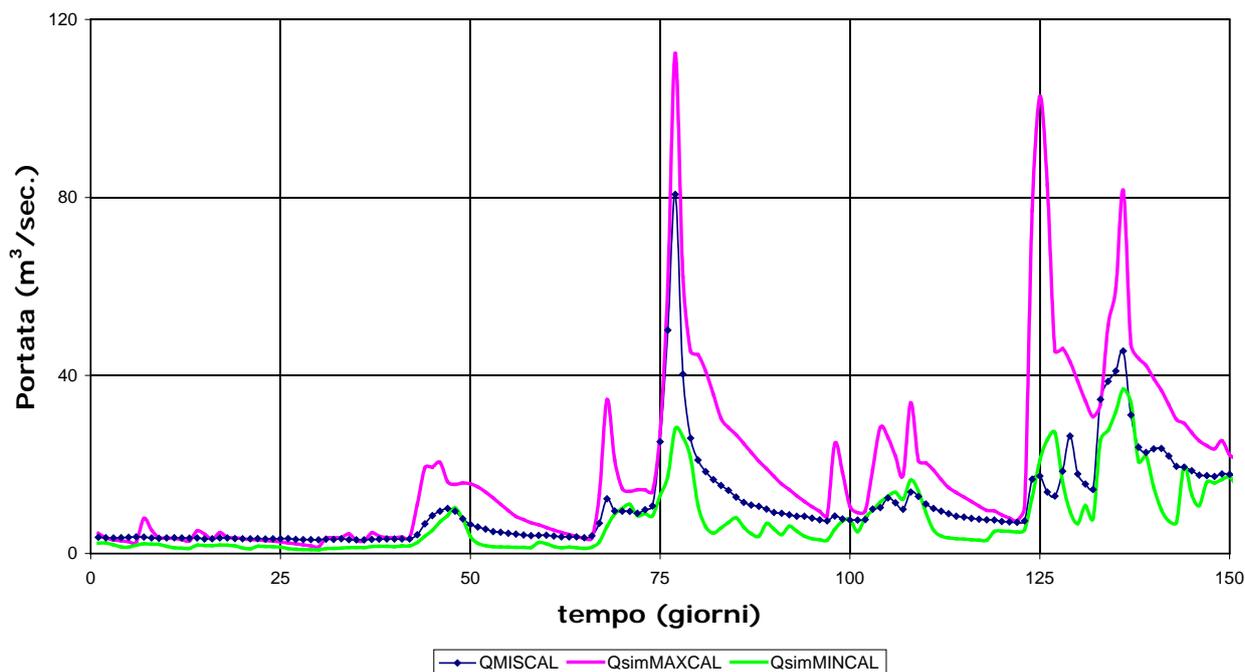
Fiume Potenza - Calibrazione (μ - σ^2 - Q_{355} -MO)

Figura 6.43. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata).

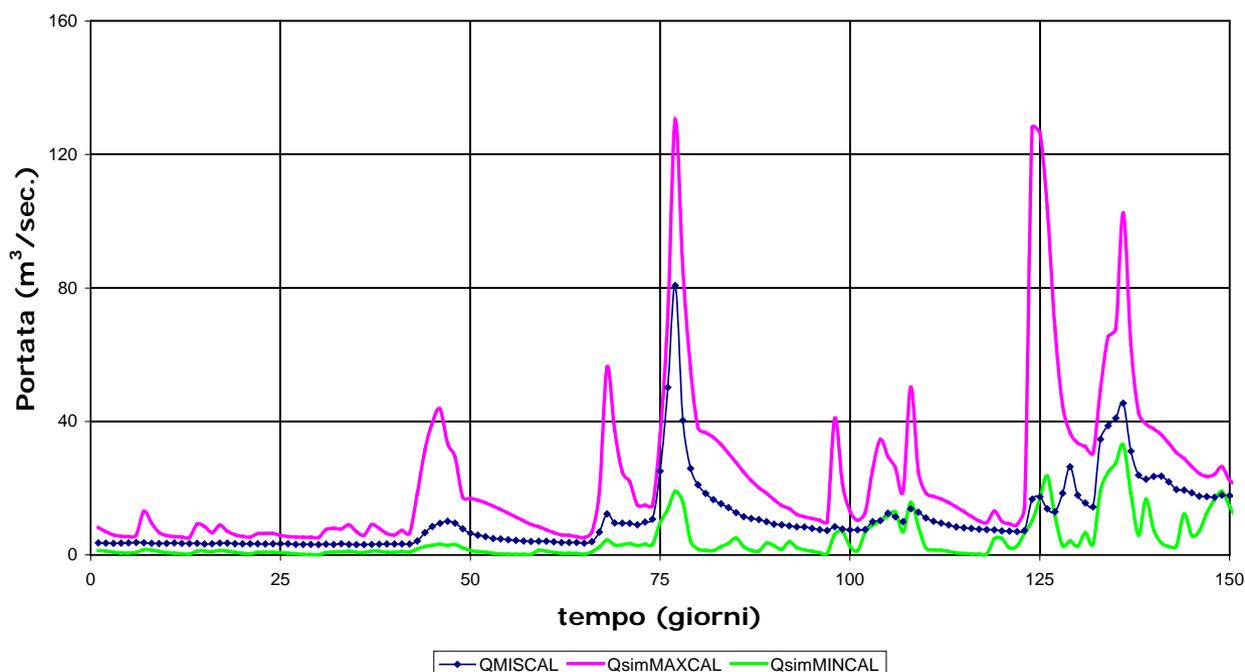
Fiume Potenza - Calibrazione (μ - σ^2 - ρ_1 - Q_{355} -MO)

Figura 6.44. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata).

Da un'analisi più dettagliata degli idrogrammi però osserviamo che ad una diminuzione della percentuale di portate escluse dagli inviluppi non corrisponde un affinamento della simulazione, bensì un aumento dell'ampiezza dell'intervallo dei valori di portata che si possono verificare.

In dettaglio nella Figura 6.39 è riportato l'idrogramma osservato (in blu) e gli inviluppi massimo e minimo (rispettivamente in fucsia e verde) ottenuti con la calibrazione *multi-obiettivo* attraverso l'uso delle statistiche $\mu-\sigma^2-\rho_1$: l'analisi di tale idrogramma conferma i risultati di Tabella 6.25 ossia l'esclusione dagli inviluppi di alcune porzioni di portata osservata. Nella Figura 6.40 e nella Figura 6.41 sono riportati gli stessi idrogrammi per il caso di calibrazione *multi-obiettivo* attraverso l'utilizzo delle statistiche $\mu-\sigma^2-Q_{355}$ e $\mu-\sigma^2-\rho_1-Q_{355}$ rispettivamente. Si nota immediatamente che, mentre l'inviluppo minimo si mantiene pressoché invariato, l'inviluppo massimo tende ad enfatizzare sempre di più i valori di portata fino a raggiungere picchi almeno 3 o 4 volte superiori al valore osservato.

Per il Fiume Potenza (Figura 6.42 - Figura 6.43 - Figura 6.44) il fenomeno descritto per Fiume Metauro è leggermente più contenuto tanto che da una prima analisi puramente visiva sembrerebbe che la simulazione ottenuta tramite le $\mu-\sigma^2-Q_{355}$ sia migliore dato che i picchi raggiunti dall'inviluppo massimo sono più contenuti. In realtà non è così poiché da un'analisi più approfondita si nota che i rami di risalita e di esaurimento dell'inviluppo massimo sono più ampi e più lontani conseguentemente dalla portata osservata. Gli indici di Tabella 6.25 inoltre indicano che in questa combinazione si raggiunge la maggior percentuale di portata esclusa dagli inviluppi per il Fiume Potenza in calibrazione.

6.4 Conclusioni

L'idea alla base di questa seconda parte della presente trattazione era testare la procedura di calibrazione del modello afflussi-deflussi già visto nella prima parte del lavoro considerando però i bacini dell'area di studio come non strumentati. Questo è stato possibile partendo dall'approssimazione della funzione di massima verosimiglianza proposta da Whittle e dalla procedura di regionalizzazione delle statistiche approfondita dall'Ing. Castiglioni (*Castiglioni, (2010)*).

A partire da questi valori è stato possibile identificare le funzioni obiettivo utilizzate sia in modalità *mono-obiettivo* che in modalità *multi-obiettivo*. Le prime sono state ricavate sommando il valore adimensionalizzato della differenza di ciascuna statistica

simulata dal modello rispetto al valore regionalizzato e attribuendo poi a ciascun valore il medesimo peso in modo da non conferire ad una statistica un peso maggiore rispetto ad un'altra. In modalità *multi-obiettivo* invece ogni statistica è stata considerata indipendentemente dalle altre portando così al caso a tre o a quattro funzioni distinte.

I risultati hanno mostrato innanzi tutto che la procedura di regionalizzazione è utilizzabile per consentire l'applicazione di procedure di calibrazioni in siti non strumentati e comunque può essere presa in considerazione per consentire una riduzione iniziale dello spazio dei parametri possibili. E' altresì da sottolineare che l'approccio suggerito nella presente trattazione ha solamente lo scopo di fornire una stima indicativa degli involuppi delle portate ottenuti tenendo conto di tutti i set di pari efficienza calcolati dall'algoritmo. Tale stima tuttavia può consentire di ottenere un'idea dell'incertezza della simulazione dato che, nel caso di bacini non strumentati, non si può quantificare lo scostamento tra simulato e osservato in modo classico. E' comunque necessario riconoscere che le difficoltà legate alla mancanza di dati rendono difficile ipotizzare l'applicabilità di questa procedura nelle problematiche reali. Ad ogni modo la massima verosimiglianza regionale può essere un aiuto importante per realizzare la calibrazione dei modelli afflussi-deflussi in funzione del fatto che, in siti non strumentati, l'integrazione di informazioni differenti, è l'unico modo per ridurre lo spazio possibile dei parametri e ottenere alcune indicazioni sull'area oggetto di studio.

I risultati ottenuti direttamente nel corso del presente studio permettono di concludere che:

- la calibrazione con le statistiche regionalizzate rispetto al caso di statistiche ottenute dai dati osservati comporta un peggioramento contenuto dei valori degli indici di prestazione (da Tabella 6.1 fino a Tabella 6.8);
- tra le statistiche regionalizzate il coefficiente di autocorrelazione lineare a ritardo unitario sembra essere quello meno incidente sulla determinazione dei valori dei parametri anche se nella combinazione a quattro funzioni i valori delle funzioni obiettivo assumono valori molto differenti tanto da rendere difficile la minimizzazione delle stesse nel medesimo momento (paragrafo 6.3.4);
- l'analisi degli involuppi delle portate sembra indicare che la calibrazione migliore avvenga con l'utilizzo in modalità *multi-obiettivo* di quattro funzioni. Un'attenta osservazione degli idrogrammi però permette di concludere che l'apparente miglior riproduzione delle portate avviene in realtà a scapito dell'accuratezza della stessa;

- l'approccio *multi-obiettivo* è confrontabile con l'approccio *mono-obiettivo* e consente di considerare più aspetti dell'idrogramma anche nel caso di bacini non strumentati. Le tabelle di seguito riportate mettono a confronto gli indici di qualità ottenuti con le varie modalità *mono* e *multi-obiettivo* analizzate nei paragrafi 6.3.2 e 6.3.3. In particolare i valori indicati in azzurro si riferiscono al miglior risultato ottenuto. Per il Fiume Candigliano (Tabella 6.27) la modalità di calibrazione che ha portato i risultati migliori è stata quella *multi-obiettivo* con l'utilizzo delle funzioni obiettivo legate a media, varianza e coefficiente di correlazione lineare a ritardo unitario. Per il Fiume Esino (Tabella 6.28) i risultati sono meno netti poiché secondo gli indici EFF e MAE la combinazione che dà i risultati migliori è quella che utilizza le funzioni obiettivo date dalla media, dalla varianza, dal coefficiente di correlazione lineare a ritardo unitario e dalla portata presente in alveo almeno 355 all'anno, mentre secondo il MRE sembrerebbe migliore l'utilizzo di media, varianza e coefficiente di correlazione lineare. Ad ogni modo in tutti i casi si tratta di calibrazione con modalità *multi-obiettivo*.

CANDIGLIANO (calibrazione)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$	$\mu-\sigma^2-\rho_1-MO(Q_{CO})$	$\mu-\sigma^2-Q_{355}-MO(Q_{CO})$	$\mu-\sigma^2-\rho_1-Q_{355}-MO(Q_{CO})$
EFF	0,62	0,64	0,63	0,62	0,7	0,66
MAE	4,25	3,8	3,84	4,09	3,34	3,68
MRE	1,18	0,74	0,73	0,9	0,61	0,61

Tabella 6.27. Fiume Candigliano: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato.

ESINO (calibrazione)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$	$\mu-\sigma^2-\rho_1-MO(Q_{CO})$	$\mu-\sigma^2-Q_{355}-MO(Q_{CO})$	$\mu-\sigma^2-\rho_1-Q_{355}-MO(Q_{CO})$
EFF	0,44	0,41	0,44	0,51	0,26	0,53
MAE	7,79	7,58	7,8	7,47	8,87	7,27
MRE	0,52	0,52	0,52	0,5	0,6	0,52

Tabella 6.28. Fiume Esino: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato.

METAURO (calibrazione)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$	$\mu-\sigma^2-\rho_1-MO(Q_{CO})$	$\mu-\sigma^2-Q_{355}-MO(Q_{CO})$	$\mu-\sigma^2-\rho_1-Q_{355}-MO(Q_{CO})$
EFF	0,55	0,36	0,36	0,56	0,38	0,64
MAE	10,38	11,54	11,54	9,98	11,35	7,82
MRE	1,49	0,85	0,85	1,24	0,83	0,63

Tabella 6.29. Fiume Metauro: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato.

Per il Fiume Metauro (Tabella 6.29) il risultato è netto e a favore della calibrazione *multi-obiettivo* con l'utilizzo di tutte e quattro le funzioni ricavate dalle statistiche.

POTENZA (calibrazione)						
Funzione Obiettivo	$\mu-\sigma^2-\rho_1$	$\mu-\sigma^2-Q_{355}$	$\mu-\sigma^2-\rho_1-Q_{355}$	$\mu-\sigma^2-\rho_1-MO(Q_{CO})$	$\mu-\sigma^2-Q_{355}-MO(Q_{CO})$	$\mu-\sigma^2-\rho_1-Q_{355}-MO(Q_{CO})$
EFF	0,4	0,46	0,45	0,33	0,52	0,44
MAE	2,92	3,04	2,99	3,07	2,96	3,08
MRE	0,46	0,49	0,5	0,46	0,49	0,53

Tabella 6.30. Fiume Potenza: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato.

Infine per il Fiume Potenza il risultato è intermedio (Tabella 6.30): l'EFF si presenta con valori migliori nella modalità *multi-obiettivo* che utilizza le funzioni obiettivo date dalla media, dalla varianza e dalla portata presente in alveo almeno 355 giorni all'anno mentre il MAE è migliore in modalità *mono-obiettivo* con l'equazione ottenuta da media, varianza e coefficiente di correlazione lineare a ritardo unitario. Il valore del MRE infine presenta valori paragonabili in entrambe le modalità appena citate.

Quanto emerso nel presente studio evidenzia le potenzialità dell'algoritmo di ottimizzazione AMALGAM che si è dimostrato versatile, di semplice interpretazione e poco dispendioso in termini di tempo consentendo la realizzazione di molte tipologie di simulazioni e la conseguente raccolta di un notevole quantitativo di dati qui riportati solo in parte.

Ulteriori studi potrebbero consentire l'applicazione a tutta l'area utilizzata per la procedura di regionalizzazione sia in casi di bacini strumentati che non strumentati portando eventualmente anche ad un confronto legato alla dimensione dei bacini stessi nonché alle diverse caratteristiche geomorfologiche e climatiche che li caratterizzano (alcuni primi risultati sono riportati in *Lombardi et al., 2011*). Potrebbe essere inoltre interessante utilizzare l'algoritmo di ottimizzazione AMALGAM in applicazione ad un modello afflussi deflussi di tipo distribuito.

Bibliografia

- Arno' D. & Mazzola M.R. (2000). "Ottimizzazione del progetto di reti idriche mediante algoritmi genetici", IDRA 2000 – XXVII Convegno di Idraulica e Costruzioni Idrauliche, Genova..... 27
- Banks A. & Vincent J. & Anyakoha C. (2007). "A review of particle swarm optimization. Part I: background and development", Springer Science + Business Media B.V. 61
- Bardossy A. (2007). "Calibration of hydrological model parameters for ungauged catchments", Hydrol. Earth Syst. Sci., 11, 703–710. 131
- Basilevsky A. (1994). "Statistical Factor Analysis and Related Methods: Theory and 140
- Beyer H. G. & Deb K. (2001). "On Self-Adaptive Features in Real-Parameter Evolutionary Algorithm", IEEE Transactions on Evolutionary Computation 5. 58
- Boyle D.P. (2000). "Multicriteria calibration of hydrological models", Ph.D. dissertation, Dep. Of Hydrol. and Water Resour., Univ. of Ariz., Tempe. 10; 83; 142
- Brath A. & Castellarin A. & Montanari A. (2003). "Assessing the reliability of regional depthduration-frequency equations for gaged and ungauged sites", Water Resour. Res., 39 141
- Brazil L.E. & Krajewski W.F. (1987). "Optimization of complex hydrologic models using random search methods", paper presented at Conference on Engineering Hydrology, Hydraulics Division, American Society of Civil Engineering, Williams-burg, VA. 29
- Bremermann H.J. (1958). "The evolution of intelligence. The Nervous System as a Model of its Environment", Technical report, no 1, contract no 477(17), Dept Mathematics, Univ. Washington, Seattle. 29
- Brockwell P.J. & Davis R.A. (1987). "Time Series: Theory and Methods", Springer, New York. 133; 137
- Burn D.H. (1990). "Evaluation of regional flood frequency analysis with a region of influence approach", Water Resour. Res., 26 (10), 2257–2265. 132
- Castellarin A. & Burn D.H. & Brath A. (2001). "Assessing the effectiveness of hydrological similarity measures for flood frequency analysis", J. Hydrol., 241, 270-285. 132
- Castellarin A. & Galeati G. & Brandimarte L. & Montanari A. & Brath A. (2004b). "Regional flowduration curves: reliability for ungauged basins", Adv. Water Res., 27 (10), 953-965. 132; 141
- Castiglioni S. & Castellarin A. & Montanari A. & Skoien J. O. & Blöschl G. (2010a). "Regionalizzazione geostatistica dei deflussi di magra: PSBI e Top-kriging a confronto", Atti del 32° Convegno di Idraulica e Costruzioni Idrauliche, Palermo 14-17 Settembre 2010. ... 78; 140

- Castiglioni S. & Lombardi L. & Castellarin A. & Toth E. & Montanari A. (2010b). "Calibration of rainfall-runoff models in ungauged basins: a regional maximum likelihood approach", *Advances in Water Resources*, Vol. 33..... 137
- Clerc M. & Kennedy J. (2002). "The particle swarm: explosion, stability and convergence in a multi-dimensional complex space", *IEEE Trans Evol Comput* 6:58–73. 62; 63
- Clerc M. (2003). TRIBES – "Un Exemple D'Optimisation par Essaim Particulaire Sans Parametres de Contrôle", In: *Optimisation par Essaim Particulaire (OEP 2003)*, Paris, France. 64
- Coello C. (2000). "An updated survey of GA-based multiobjective optimization techniques", *ACM Computing Surveys*, vol.32..... 37
- Coloni A. & Dorigo M. & Maoli F. & Maniezzo V. & Righini G. & Trubian M. & (1996). " Heuristics From Nature For Hard Combinatorial Optimization Problems ", *International Transactions in Operational Research*, Vol. 3, N. 1, pp. 1-21. 27; 28
- Dalrymple T. (1960). "Flood Frequency Analyses", Reston (VA, USA): USGS Wat. Supply Paper 1543-A. 132
- Deb K. & Agarwal R. B. (1995). "Simulated Binary Crossover for Continuous Search Space", *Complex Systems* 9. 58
- Deb K. & Pratap A. & Agarwal S. & Meyarivan T. (2002). "A fast and elitist multiobjective genetic algorithm: Nsgall", *Transactions on Evolutionary Computation*. 47; 50; 54; 56
- Duan Q. & Sorooshian S. & Gupta V.K. (1992). "Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models", *Water Resources Research*, 28(4), 1015-1031. .. 10; 16; 29
- Duan Q. & Sorooshian S. & Gupta V.K. (1993). "A Shuffled Complex Evolution Approach for Effective and Efficient Global Optimization", *Journal of Optimization Theory and Its Applications*, 76(3), 501-521. 10; 29
- Duan Q. & Sorooshian S. & Gupta V.K. (1994). "Optimal Use of the SCE-UA Global Optimization Method for Calibrating Watershed Models", *Journal of Hydrology*, 158, 265-284. 9
- Eberhart R.C. & Simpson P. & Dobbins R. (1996). "Computational intelligence PC tools", *AP Professional*, San Diego, CA, Chapter 6, pp 212–226. 62
- Efron B. (1982). "The Jackknife, the Bootstrap and Other Resampling Plans", *Society for* 141
- FEH (1999). "Flood Estimation Handbook" (FEH), Institute of Hydrology, Wallingford (UK). . 132
- Fogel L.J. & Owens A.J. & Walsh M.J. (1966). "Artificial intelligence thorough simulated evolution", John Wiley & Sons, Ltd, Chichester, UK. 29
- Forrest S. (1993). "Genetic algorithms: principles of natural selection applied to computation", *Science*, vol.261..... 37; 40
- Friedberg R.M. (1958). "A learning machine: Part I", *IBM J* 2–13..... 29
- Furey P.R. & Gupta V.A. (2000). "Space-time variability of low streamflows in river networks", *Water Resour. Res.*, 36 (9), 2679-2690..... 132
- Gelman A.G. & Roberts G.O. & Gilks W.R. (1996). "Efficient Metropolis jumping rules", In J.M. Bernardo & J.O. Berger & A.F. David & A.F.M. Smith (eds), *Bayesian Statistics V*, pp. 599-608, Oxford: Oxford University Press. 66

- Gilks W.R. & Roberts G.O. & Sahu S.K. (1998). "Adaptive Markov chain Monte Carlo", J Amer. 66
- Gilks W.R. & Roberts G.O. (1995). "Strategies for improving MCMC", In W.R. Gilks & S. Richardson & D.J. Spiegelhalter (eds), *Markov Chain Monte Carlo in Practice*, pp. 75-88. London: Chapman & Hall. 66
- Goldberg D. (1989). "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley. 29; 36
- Goovaerts P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall", *Journal of Hydrology*, 228, 113-129. 143
- Gupta H. V. & Kling H. & Yilmaz K. K. & Martinez G. F. (2009). "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling", *J. Hydrol.*, 377, 80-91. 133
- Gupta H.V. & Sorooshian S. & Yapo P.O. (1998). "A Toward Improved Calibration of Hydrologic Models: Multiple and Non-Commensurable Measures of Information", *Water Resources Research*, Vol. 34, No. 4, pp. 751-763. 9; 10
- Haario H. & Saksman E. & Tamminen J. (1999). "Adaptive proposal distribution for random walk Metropolis algorithm", *Comput. Statist.*, 14, 375-395. 66
- Haario H. & Saksman E. & Tamminen J. (2001). "An adaptive Metropolis algorithm", *Bernoulli*, vol. 7. 50; 66
- Hart W.E. & Krasnogor N. & Smith J.E. (2005). "Recent Advances in Memetic Algorithms", Springer, Berlin. 45
- Hastings W.K. (1970). "Monte Carlo Sampling Methods Using Markov Chains and Their Applications", *Biometrika* 57 (1): 97-109. 66
- Holland J. (1992). "Genetic algorithms", *Scientific American*. 36; 40
- Holland J.H. (1962). "Outline for a logical theory of adaptive systems", *J Assoc Comput Machinery* 3: 297-314. 29
- Holland J.H. (1975). "Adaptation in natural and artificial systems", University of Michigan Press, Ann Arbor. 29
- Hosking J.R.M. & Wallis J.R. (1997). "Regional Frequency Analysis: An Approach Based on LMoments", Cambridge (UK): Cambridge University Press. 132
- Hundecha Y. & Ouarda T. B. M. J. & Bardossy A. (2008). "Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the "spatial" structures of the parameters within a canonical physiographic-climatic space", *Water Resour. Res.*, 44, W01427, doi: 10.1029/2006WR005439. 131
- Kennedy J & Eberhart R.C. (1995). "Particle swarm optimization", *Proceedings of IEEE International Conference on Neural Networks*, Piscataway, NJ, pp 1942-1948. 50; 60
- Kennedy J. (1997). "The particle swarm: social adaptation of knowledge", In: *Proceedings of the international conference on evolutionary computation*, IEEE, Piscataway, NJ, pp 303-308. 63; 64
- Koza J. (1992). "Genetic Programming: on the programming of computers by means of natural selection", MIT Press, Cambridge, MA. 29

- Koza J. et al. (2003). "Genetic Programming IV: Routine Human-Competitive Machine Intelligence", Kluwer Academic Publishers. 41
- Krink T. & Vestertroem J.S. & Riget J. (2002). "Particle swarm optimization with spatial particle extension", Proceedings of the IEEE congress on evolutionary computation (CEC 2002), Honolulu, Hawaii. 65
- Laaha G. & Blöschl G. (2006). "Seasonality indices for regionalizing low flows", *Hydrol. Process.*, 20, 3851–3878. 132
- Lindström G. (1997). "A simple automatic calibration routine for the HBV model", *Nordic Hydrol.* 28 (3), 153–168. 22; 90
- Liong S.Y. & Khu S.T. & Chan W.T. (1996). "Construction of multiobjective function response surface with genetic algorithm and neural network", Proceedings of the International Conference on Water Resources and Environmental Research, 29–31 October, Kyoto, Japan, vol. II, pp. 31–38. 22; 90
- Liong S.Y. & Khu S.T. & Chan W.T. (1998). "Derivation of Pareto front with accelerated convergence genetic algorithm", ACGA, in: Babovic V. & Larsen L.C. (Eds.). *Hydroinformatics'98*, Balkema, Rotterdam, The Netherlands, pp. 889–896. 22; 90
- Lombardi L. & Castiglioni S. & Toth E. & Castellarin A. & Montanari A. (2010a). "A regional maximum likelihood approach for calibrating rainfall-runoff models in ungauged basins", *Atti del XXXII Convegno Nazionale di Idraulica e Costruzioni Idrauliche – Palermo*. 134
- Lombardi L. & Toth E. & Castellarin A. & Montanari A. & Brath A. (2011). "Calibration of a rainfall-runoff model at regional scale by optimising river discharge statistics: performance analysis for different river flow regimes", *Physics and Chemistry of the Earth B* (manoscritto proposto per la pubblicazione), 2011. 191
- Lombardi L. & Toth E. & Castellarin A. & Montanari A. (2010b). Calibration of hydrological models in ungauged basins: the value of multiobjective optimisation, EGU Leonardo Topical Conference, Lussemburgo. 140
- Lombardi L. & Toth E. & Montanari A. (2010). "Simulation of average and low flows through the regional calibration of a rainfall-runoff model II part: model calibration and simulation results", EGU Vienna. 146
- Madsen H. (2000). "Automatic calibration of a conceptual rainfall-runoff model using multiple objectives", *Journal of Hydrology* 235 276–288. 10; 23; 90
- McKay M.D. & Beckman R.J. & Conover W.J. (1979). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code", *Technometrics American Statistical Association and American Society for Quality*. 46
- McLeish D.L. (1975). "A maximal inequality and dependent strong laws", *Ann. Probab.*, 3, 829–839. 67
- Merz R. & Blöschl G. (2004). "Regionalisation of catchment model parameters", *J. Hydrol.*, 287, 95–123. 131
- Merz R. & Blöschl G. (2005). "Flood frequency regionalisation-spatial proximity vs. catchment attributes", *J. Hydrol.*, 302 (1-4), 283–306. 132
- Metropolis N. & Rosenbluth A.W. & Rosenbluth M.N. & Teller A.H. & Teller E. (1953). "Equations of State Calculations by Fast Computing Machines", *Journal of Chemical Physics* 21 (6): 1087–1092. 66

- Michalewicz Z. (1996). "Genetic Algorithms + Data Structures: Evolution Programs", Springer Editor. 28
- Mitchell M. (1996). "An Introduction to Genetic Algorithms", MIT Press. 36; 39; 40; 41
- Moisello U. (1999). "Idrologia tecnica", La Goliardica Pavese, PAVIA. 81
- Monson C.K. & Seppi K.D. (2005). "Improving on the Kalman swarm extracting its essential characteristics", Proceedings of genetic and evolutionary computation conference (GECCO), ACM, Washington, DC. 64
- Montanari A. & Toth E. (2006). "Calibrazione di modelli afflussi – deflussi nel dominio spettrale: una possibilità per bacini idrografici non strumentati?", Atti del 30° Convegno di Idraulica e Costruzioni Idrauliche, Roma..... 134
- Montanari A. & Toth E. (2007). "Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?", Water Resour. Res., 43, W05434, doi: 10.1029/2006WR005184. 83; 132; 136
- Moore R.J. (1985). "The probability-distributed principle and runoff production at point and basin scale", Hydrological Sciences Journal 30(2), 273-297. 83
- Nash, J. & Sutcliffe J. (1970). "River flow forecasting through conceptual models, Part 1: a discussion of principles", J. Hydrol., 10, 282-290..... 141
- Nelder J. A. & Mead R. (1965). "A simplex method for function minimization", Computer Journal 7, 308-313. 15
- Ozcan E. & Mohan C.K. (1998). "Analysis of a Simple PSO System", Intelligent Engineering Systems Through Artificial Neural Networks, Vol. 8, pp. 253-258. 63
- Ozcan E. (1999). "Analysis of Particle Velocities in a Particle Swarm Optimization System", 8th Turkish Symposium on Artificial Intelligence and Neural Networks, pp. 257-266..... 63
- Pandey G.R. & Nguyen V.T.V. (1999). "A comparative study of regression based methods in regional flood frequency analysis", J. Hydrol., 225(1-2), 92–101. 132
- Rechenberg I. (1965). "Cybernetic solution path of an experimental problem", Royal Aircraft Establishment, Library Translation No 1122. 29
- Riget J. & Vesterstrøm J.S. (2002). "A diversity-guided particle swarm optimizer – the ARPSO", EVALife Technical Report no 2002..... 65
- Roberts G.O. & Gelman A. & Gilks W.R. (1997). "Weak convergence and optimal scaling of random walk Metropolis algorithms", Ann. Appl. Probab., 7, 110-120. 66
- Seibert J. (1999). "Regionalisation of parameters for a conceptual rainfall-runoff model", Agric. For. Meteorol., 98– 99, 279– 293. 131
- Shao J. & Tu D. (1995). "The Jackknife and Bootstrap", Springer-Verlag, New York. 141
- Smakhtin V.U. (2001). "Low flow in hydrology: a review", J. Hydrol., 240, 147-186. 132
- Sorooshian S. & Duan Q. & Gupta V.K. (1993). "Calibration of rainfall-runoff models: Application of global optimization to the Sacramento soil moisture accounting model", Water Resour. Res., 29, 1185–1194..... 29
- Srinivas N. & Deb K. (1994). "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms", Evolutionary Computation 2. 54

- Stedinger J.R. & Vogel R.M. & Foufoula-Georgiou E. (1993). "Frequency analysis of extreme events", Maidment DR (ed.) Handbook of Hydrology, New York (USA): McGraw-Hill-Inc... 132
- Storn R. & Price K. (1997). "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", Journal of Global Optimization. 50; 70
- Storn R. & Price K. (1995). "Differential Evolution – A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", Technical Report TR-95-012, ICSI. 29
- van den Bergh F. & Engelbrecht A.P. (2002). "A new locally convergent particle swarm optimizer", Proceedings of IEEE conference on systems, man and cybernetics, Hammamet, Tunisia. 64
- Vogel R.M. & Kroll C.N. (1992). "Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics", Water Resour. Res., 28 (9), 2451–2458. 132
- Vrugt A. J. & Robinson A. B. (2007). "Improved evolutionary optimization from genetically adaptive multimethod search", PNAS, vol. 104. 46; 51
- Wagner T. & Boyle D.P. & Lees M.P. & Weather H.S. & Gupta H.V. & Sorooshian S. (2001), "A framework for development and application of hydrological models", Hydrology and Earth System Sciences 5(1), 13-26. 83
- Wagner T. & Wheater H.S. (2006). "Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty", J. Hydrol., 320, 132-154. 131
- Wang Q.J. (1991). "The genetic algorithm and its application to calibrating conceptual rainfall-runoff models", Water Resour. Res., 27, 2467–2471. 29
- Weisberg S. (1985). "Applied linear regression", 2nd ed., John Wiley, New York. 138
- Whittle P. (1953). "Estimation and information in stationary time series". Ark. Mat., 2, 423-434. 132; 134
- Wolpert D.H. & Macready W.G. (1997). "No Free Lunch Theorems for Optimization", IEEE Transactions on Evolutionary Computation. 53
- Yapo P.O. & Gupta H.V. & Sorooshian S. (1998). "Multi objective global optimization for hydrologic models", Journal of Hydrology, 204, 83-97. 9; 10; 22; 90
- Zhang Z. & Kroll C.N. (2007). "The baseflow correlation method with multiple gauged sites", J. Hydrol., 347(3-4), 371-380, doi: 10.1016/j.jhydrol.2007.09.025. 141
- Zitzler E. & Thiele L. (1999). IEEE Trans. Evol. Comp. 3:257–271. 29

Indice delle Figure

Figura 1.1. Esempio di risultato di una calibrazione mono-obiettivo per un problema a due parametri.	21
Figura 1.2. Esempio di confronto tra idrogramma osservato e simulato per una calibrazione mono-obiettivo.	22
Figura 1.3. Esempio di risultato di una calibrazione multi-obiettivo per un problema a due parametri.	23
Figura 1.4. Esempio di confronto tra idrogramma osservato e idrogrammi simulati per una calibrazione multi-obiettivo.	24
Figura 1.5. Esempio di concetto di Frontiera Pareto nel caso a due parametri e due criteri F_1 e F_2 . a) indica lo spazio dei parametri e b) lo spazio delle funzioni obiettivo. I punti A e B indicano le soluzioni che riducono al minimo ciascuno dei singoli criteri.	25
Figura 2.1. Esempi di tre strutture ad albero normalmente utilizzate nella programmazione genetica. L'espressione matematica che ognuno rappresenta è riportata sotto alla corrispondente rappresentazione.	32
Figura 2.2. Crossover e mutazione.	34
Figura 2.3. Esempio di semplice rete neurale feedforward con un livello di input che consiste in quattro neuroni, un layer nascosto da tre neuroni e un layer di output composto da quattro neuroni. Il numero su ogni neurone rappresenta la sua soglia di attivazione.	41
Figura 2.4. Esempio di spazio delle soluzioni in una procedura di ottimizzazione tramite hill-climbing: l'algoritmo inizia in un determinato punto dello spazio e poi si muove in salita alla ricerca del punto più alto che dovrebbe corrispondere all'ottimo globale.	42
Figura 3.1. Esempio di assegnazione del rank e concetto di crowding distance.	47
Figura 3.2: Meccanismo del crowding distance per mantenere la diversità nel fronte.	48
Figura 3.3: Significato del parametro crowding distance	49
Figura 3.4. Posizione delle soluzioni rispetto alla posizione della frontiera Pareto dopo 25, 50 e 75 generazioni con NSGA-II (quadrantini) PSO (cerchietti), AMS (rombi neri), DE (diamanti) e AMALGAM (x) per il test ZDT4.	51
Figura 3.5. Visualizzazione del concetto di creazione delle nuove generazioni in modo auto-adattivo. (A) Evoluzione del numero di nuovi punti generati con gli algoritmi NSGA-II (quadrantini verdi), PSO (cerchietti azzurri), AMS (+), PSO DE (diamanti verdi) rispetto all'uso di AMALGAM come metodo di ricerca multi metodo per il problema test ZDT4. (B) Convergenza metrica di AMALGAM e degli altri quattro algoritmi presi singolarmente. L'immagine mostra la capacità di ogni singolo algoritmo di avvicinarsi maggiormente alla soluzione durante le varie fasi dell'ottimizzazione e produce un esempio numerico del teorema "No Free Lunch Theorems for Optimization" di Wolpert and Macready evidenziando che è impossibile sviluppare un singolo algoritmo di ricerca che sia sempre superiore ad ogni altro per tutto il periodo necessario per la convergenza.	53
Figura 3.6. Soluzioni non dominate individuate per il problema test DTLZ6 con l'NSGA-II e AMALGAM dopo rispettivamente 100,000 and 5,000 valutazioni di funzione. (A) La linea nera rappresenta la frontiera ed è possibile notare che AMALGAM vi converge perfettamente dopo un numero esiguo di valutazioni. (B) Creazione delle nuove generazioni in maniera auto-	

adattiva ottenuta variando la relativa importanza degli algoritmi durante l'ottimizzazione in modo da rendere il processo più efficiente e veloce.....	54
Figura 3.7. Schema di funzionamento dell'algoritmo NSGA-II.....	55
Figura 4.1: Area di studio: bacini idrografici e relative stazioni idrometriche. In azzurro sono evidenziati i bacini per i quali sono disponibili serie di portate, piogge, ed evapotraspirazione riferite al medesimo intervallo temporale.....	76
Figura 4.2: Principali unità litologiche delle regioni Marche ed Abruzzo. Suddivisione in classi a diverse permeabilità e porosità (fonte DICAM, Facoltà di Ingegneria di Bologna).	79
Figura 4.3: Sensori pluviometrici disponibili nell'area di studio.	80
Figura 4.4: Schema di funzionamento del modello Hymod.	83
Figura 4.5. Andamento della funzione di distribuzione della capacità di immagazzinamento del bacino in funzione del valore di b.	84
Figura 4.6. Situazione iniziale del bacino e raggiungimento della massima capacità di immagazzinamento.	85
Figura 4.7. Porzione di precipitazione che viene smaltita esclusivamente come deflusso superficiale.	85
Figura 4.8. Porzione di precipitazione suddivisa tra deflusso profondo e deflusso superficiale.	86
Figura 5.1. Area di studio per la trattazione del caso di bacini strumentati.	88
Figura 5.2. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (RMSE).	93
Figura 5.3. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (RMSE).	93
Figura 5.4. scatterplot in calibrazione e in validazione delle portate relative al Fiume Candigliano a Acqualagna (RMSE).	94
Figura 5.5. Idrogramma in calibrazione del Fiume Esino a Cannuciaro (RMSE).	95
Figura 5.6. Idrogramma in validazione del Fiume Esino a Moie (RMSE).....	96
Figura 5.7. scatterplot in calibrazione e in validazione delle portate relative al Fiume Esino a Moie (RMSE).	96
Figura 5.8. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (RMSE)...	97
Figura 5.9. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (RMSE)....	98
Figura 5.10. scatterplot delle portate in calibrazione e validazione relative al Fiume Metauro a Barco di Bellaguardia (RMSE).	98
Figura 5.11. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro (RMSE).	100
Figura 5.12. Idrogramma in validazione del Fiume Potenza a Cannuciaro (RMSE).....	100
Figura 5.13. scatterplot delle portate in calibrazione e validazione relative al Fiume Potenza a Cannuciaro (RMSE).	101
Figura 5.14. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (MAE-MRE).	103
Figura 5.15. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (MAE-MRE). .	103
Figura 5.16. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).	104
Figura 5.17. Involuppi massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto e portata osservata in validazione per il Fiume Candigliano ad Acqualagna (periodo di 150 giorni attorno al massimo evento registrato).	104
Figura 5.18. Idrogramma in calibrazione del Fiume Esino a Moie (RMSE-BIAS).....	106
Figura 5.19. Idrogramma in validazione del Fiume Esino a Moie (RMSE-BIAS).....	106

Figura 5.20. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (RMSE-BIAS).	107
Figura 5.21. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (RMSE-BIAS).	107
Figura 5.22. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (MAE-MRE).	109
Figura 5.23. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (MAE-MRE).	109
Figura 5.24. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).	110
Figura 5.25. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MRE).	111
Figura 5.26. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro (MAE-MSLE).	112
Figura 5.27. Idrogramma in validazione del Fiume Potenza a Cannuciaro (MAE-MSLE).	113
Figura 5.28. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSLE).	113
Figura 5.29. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSLE).	114
Figura 5.30. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna (MAE-MSE-MRE).	116
Figura 5.31. Idrogramma in validazione del Fiume Candigliano ad Acqualagna (MAE-MSE-MRE).	116
Figura 5.32. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	117
Figura 5.33. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	118
Figura 5.34. Idrogramma in calibrazione del Fiume Esino a Moie (MAE-MSE-MRE).	119
Figura 5.35. Idrogramma in validazione del Fiume Esino a Moie (MAE-MSE-MRE).	120
Figura 5.36. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	120
Figura 5.37. Inviluppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	121
Figura 5.38. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia (MAE-MSE-MRE).	123
Figura 5.39. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia (MAE-MSE-MRE).	123
Figura 5.40. Inviluppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	124

Figura 5.41. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).....	124
Figura 5.42. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro (MAE-MSE-MRE).	125
Figura 5.43. Idrogramma in validazione del Fiume Potenza a Cannuciaro (MAE-MSE-MRE).	126
Figura 5.44. Involuppi massimo e minimo delle portate e portata osservata in calibrazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	126
Figura 5.45. Involuppi massimo e minimo delle portate e portata osservata in validazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato (MAE-MSE-MRE).	127
Figura 6.1. Scatter-plots che riportano un confronto diretto fra le grandezze statistiche empiriche e quelle stimate mediante modelli multiregressivi applicati con procedura di ricampionamento jack-knife. In particolare nei grafici di sinistra vengono riportati i valori relativi ai 52 bacini, mentre a destra un particolare di quelli ottenuti per i 13 bacini per i quali è stata applicata la procedura di calibrazione.....	139
Figura 6.2. Esempio di applicazione della tecnica di interpolazione spaziale Ordinary Kriging in cross-validazione per il bacino del Fiume Aterno facente parte dell'area di studio di 52 bacini utilizzati per la regionalizzazione delle statistiche.....	142
Figura 6.3. Frontiera Pareto: set della frontiera a minima distanza dall'origine degli assi.	147
Figura 6.4. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu-\sigma^2-Q_{355}$)....	152
Figura 6.5. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu-\sigma^2-Q_{355}$)....	153
Figura 6.6. Idrogramma in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-Q_{355}$).	153
Figura 6.7. Idrogramma in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-Q_{355}$).	154
Figura 6.8. Idrogramma in calibrazione del Fiume Esino a Moie ($\mu-\sigma^2-Q_{355}$).....	156
Figura 6.9. Idrogramma in validazione del Fiume Esino a Moie ($\mu-\sigma^2-Q_{355}$).	157
Figura 6.10. Idrogramma in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-Q_{355}$).	157
Figura 6.11. Idrogramma in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-Q_{355}$).	158
Figura 6.12. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1$).	160
Figura 6.13. Idrogramma in validazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1$).	160
Figura 6.14. Idrogramma in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).	161
Figura 6.15. Idrogramma in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).	161
Figura 6.16. Idrogramma in calibrazione del Fiume Potenza a Cannuciaro ($\mu-\sigma^2-\rho_1$).	163
Figura 6.17. Idrogramma in validazione del Fiume Potenza a Cannuciaro ($\mu-\sigma^2-\rho_1$).	163
Figura 6.18. Idrogramma in calibrazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).	164
Figura 6.19. Idrogramma in validazione per il Fiume Potenza a Cannuciaro per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$).	164

Figura 6.20. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu-\sigma^2-\rho_1$ -MO).	167
Figura 6.21. Idrogramma in calibrazione del Fiume Candigliano ad Acqualagna ($\mu-\sigma^2-\rho_1$ -MO).	167
Figura 6.22. Idrogramma in calibrazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($(\mu-\sigma^2-Q_{355}$ -MO).	168
Figura 6.23. Idrogramma in validazione per il Fiume Candigliano ad Acqualagna per un periodo di 150 giorni attorno al massimo evento registrato ($(\mu-\sigma^2-Q_{355}$ -MO).	168
Figura 6.24. Idrogramma in calibrazione del Fiume Esino a Moie ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	170
Figura 6.25. Idrogramma in validazione del Fiume Esino a Moie ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	170
Figura 6.26. Idrogramma in calibrazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	171
Figura 6.27. Idrogramma in validazione per il Fiume Esino a Moie per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	171
Figura 6.28. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	173
Figura 6.29. Idrogramma in calibrazione del Fiume Metauro a Barco di Bellaguardia ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	173
Figura 6.30. Idrogramma in calibrazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	174
Figura 6.31. Idrogramma in validazione per il Fiume Metauro a Barco di Bellaguardia per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1-Q_{355}$ -MO).	174
Figura 6.32. Idrogramma in calibrazione del Fiume Potenza a Cannucciaro ($\mu-\sigma^2-\rho_1$ -MO).	176
Figura 6.33. Idrogramma in validazione del Fiume Potenza a Cannucciaro ($\mu-\sigma^2-\rho_1$ -MO).	176
Figura 6.34. Idrogramma in calibrazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$ -MO).	177
Figura 6.35. Idrogramma in validazione per il Fiume Potenza a Cannucciaro per un periodo di 150 giorni attorno al massimo evento registrato ($\mu-\sigma^2-\rho_1$ -MO).	177
Figura 6.36. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo $\mu-\sigma^2-\rho_1$. I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.	179
Figura 6.37. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo $\mu-\sigma^2-Q_{355}$. I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.	180
Figura 6.38. Fiume Candigliano: Frontiera Pareto ottenuta con l'utilizzo della combinazione multi-obiettivo $\mu-\sigma^2-\rho_1-Q_{355}$. I cerchietti blu indicano i set che hanno indici di qualità migliori rispetto alla simulazione mono-obiettivo che utilizza le stesse statistiche.	181
Figura 6.39. Fiume Metauro: involuppi massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche $\mu-\sigma^2-\rho_1$. (periodo di 150 giorni attorno al massimo picco di portata).	184
Figura 6.40. Fiume Metauro: involuppi massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche $\mu-\sigma^2-Q_{355}$ (periodo di 150 giorni attorno al massimo picco di portata).	184
Figura 6.41. Fiume Metauro: involuppi massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche $\mu-\sigma^2-\rho_1-Q_{355}$ (periodo di 150 giorni attorno al massimo picco di portata).	185

Figura 6.42. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 .(periodo di 150 giorni attorno al massimo picco di portata)..... 185

Figura 6.43. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata)..... 186

Figura 6.44. Fiume Potenza: involucri massimo e minimo delle portate ottenute da tutti i set della Frontiera Pareto attraverso la calibrazione multi-obiettivo con le statistiche μ - σ^2 - ρ_1 - Q_{355} (periodo di 150 giorni attorno al massimo picco di portata)..... 186

Indice delle Tabelle

Tabella 1.1. Esempi di funzioni obiettivo di comune utilizzo.....	14
Tabella 3.1. Numero di valutazioni necessarie per raggiungere la convergenza, valori di convergenza Y e indicatori di diversità 'delta' dopo 150 generazioni per i 10 test utilizzati nello studio preliminare.....	52
Tabella 4.1: Stazioni Idrometriche interessate dallo studio e indici caratteristici delle portate di magra (Q_{355} , $Q_{95\%}$ e $7Q_{10}$), media di lungo periodo (μ), varianza (σ^2) e coefficiente di correlazione lineare a ritardo un giorno (ρ_1) relative alle serie di portata giornaliera disponibili.....	77
Tabella 4.2. Stazioni idrometriche interessate dallo studio e relativi dati geomorfologici caratteristici dei bacini idrografici da esse sottese.....	81
Tabella 4.3. Stazioni idrometriche interessate dallo studio e relativi dati climatici caratteristici dei bacini idrografici da esse sottese.....	82
Tabella 5.1. Dati disponibili per l'area di studio presa in esame nella trattazione di bacini considerati strumentati.....	88
Tabella 5.2. Funzioni obiettivo prese in esame nel caso di bacini considerati come strumentati e loro espressione matematica.....	89
Tabella 5.3. Combinazioni di Funzioni Obiettivo prese in esame nel caso di bacini considerati come strumentati.....	90
Tabella 5.4. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	92
Tabella 5.5. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	92
Tabella 5.6. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	94
Tabella 5.7. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	95
Tabella 5.8. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	97
Tabella 5.9. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	97
Tabella 5.10. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	99
Tabella 5.11. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.....	99
Tabella 5.12. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).....	102
Tabella 5.13. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).....	102

Tabella 5.14. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.	105
Tabella 5.15. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.	105
Tabella 5.16. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).	108
Tabella 5.17. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).	108
Tabella 5.18. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).	111
Tabella 5.19. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (due funzioni).	112
Tabella 5.20. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	115
Tabella 5.21. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	115
Tabella 5.22. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	118
Tabella 5.23. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	119
Tabella 5.24. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	121
Tabella 5.25. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	122
Tabella 5.26. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	125
Tabella 5.27. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo (tre funzioni).	125
Tabella 5.28. Valori di EFF in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	128
Tabella 5.29. Valori di EFF in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	128
Tabella 5.30. Valori di MAE in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	129
Tabella 5.31. Valori di MAE in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	129
Tabella 5.32. Valori di MRE in calibrazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	129
Tabella 5.33. Valori di MRE in validazione ottenuti tramite tutte le combinazioni di funzioni obiettivo prese in considerazione per i bacini esaminati.	129
Tabella 6.1. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Candigliano.	148
Tabella 6.2. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Candigliano.	148
Tabella 6.3. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Esino.	149
Tabella 6.4. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Esino.	149

Tabella 6.5. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Metauro.	149
Tabella 6.6. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Metauro.	150
Tabella 6.7. Confronto in calibrazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Potenza.	150
Tabella 6.8. Confronto in validazione degli indici di qualità ottenuti con le modalità multi-obiettivo (strumentato vs. non strumentato) per il Fiume Potenza.	150
Tabella 6.9. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.	151
Tabella 6.10. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.	151
Tabella 6.11. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.	155
Tabella 6.12. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.	156
Tabella 6.13. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.	158
Tabella 6.14. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.	159
Tabella 6.15. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità mono-obiettivo.	162
Tabella 6.16. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità mono-obiettivo.	162
Tabella 6.17. Valori degli indici di qualità ottenuti per il Fiume Candigliano in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.	165
Tabella 6.18. Valori degli indici di qualità ottenuti per il Fiume Candigliano in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.	166
Tabella 6.19. Valori degli indici di qualità ottenuti per il Fiume Esino in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.	169
Tabella 6.20. Valori degli indici di qualità ottenuti per il Fiume Esino in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.	169
Tabella 6.21. Valori degli indici di qualità ottenuti per il Fiume Metauro in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.	172
Tabella 6.22. Valori degli indici di qualità ottenuti per il Fiume Metauro in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.	172
Tabella 6.23. Valori degli indici di qualità ottenuti per il Fiume Potenza in calibrazione con le varie funzioni utilizzate in modalità multi-obiettivo.	175
Tabella 6.24. Valori degli indici di qualità ottenuti per il Fiume Potenza in validazione con le varie funzioni utilizzate in modalità multi-obiettivo.	175
Tabella 6.25. Percentuale esclusa dagli inviluppi di portata massimo e minimo ottenuti in calibrazione da tutti i set appartenenti alla Frontiera Pareto. Il valore tra parentesi indica la percentuale di portata di magra esclusa dagli inviluppi medesimi.	183
Tabella 6.26. Percentuale esclusa dagli inviluppi di portata massimo e minimo ottenuti in validazione da tutti i set appartenenti alla Frontiera Pareto. Il valore tra parentesi indica la percentuale di portata di magra esclusa dagli inviluppi medesimi.	183
Tabella 6.27. Fiume Candigliano: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato.	189

Tabella 6.28. Fiume Esino: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato. 190

Tabella 6.29. Fiume Metauro: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato. 190

Tabella 6.30. Fiume Potenza: confronto degli indici di qualità ottenuti in calibrazione in modalità mono e multi-obiettivo nel caso di bacino non Strumentato. 190