

Ottimizzazione della modellistica di simulazione del rischio idraulico tramite algoritmi di *Optimal Sensor Location*

Grazia Fattoruso (*), Annalisa Agresta (**), Guido Guarnieri (*),
Antonio Buonanno (*), Bruno Lanza (*), Mario Molinara (***), Claudio Marrocco (***),
Saverio De Vito (*), Francesco Tortorella (***), Girolamo Di Francia (*)

(*) UTTP – ENEA – Ente per le Nuove Tecnologie, l'Energia e lo Sviluppo Economico Sostenibile, P.le E. Fermi, 1 –
80055 Portici (NA), tel. +39 0817723266, fax +39 0817723344

(**) Dipartimento di Ingegneria civile, edile e ambientale - via Claudio - Edificio C8 annalisa.agresta@unina.it

(***) Dipartimento di Ingegneria Elettrica e dell'Informazione "M. Scarano" Università degli Studi di Cassino e del
Lazio Meridionale Via G. di Biasio 43 03043 Cassino (FR)

Riassunto

L'efficacia dei modelli idrologici/idraulici che vengono utilizzati per gli scopi di previsione e allerta del rischio inondazione è funzione del grado di conoscenza del territorio di interesse e di parametri idro-climatici quali precipitazioni e portate/livelli idrometrici dei corsi d'acqua in diversi punti dei bacini idrografici. Generalmente nell'installazione di sensori idro-climatici, il numero e la posizione è fatta dipendere quasi esclusivamente da condizioni pratiche ed operative quali il costo piuttosto che l'accessibilità di un'area o la ricezione del segnale per la trasmissione del dato in quell'area. Tuttavia è possibile definire schemi di posizionamento ottimale dei sensori in funzione dello scenario applicativo di interesse e dei costi di installazione e manutenzione, tramite l'utilizzo di innovative piuttosto che consolidate metodologie di optimal sampling design ed algoritmi di intelligenza computazionale. In questo contesto si colloca il presente lavoro andando a individuare la sensoristica da installare ad integrazione di quella esistente per gli scopi di modellazione degli eventi di inondazione e quindi il numero e la posizione ottimale dove collocarli sulle aree pilota affinché la modellistica di simulazione sia ottimizzata restituendo più affidabili scenari previsionali e di allerta.

Abstract

Numerical models are instrumental to more effective flood forecasting and management services though they suffer from numerous uncertainty sources. An effective model calibration is hence essential. In this research work, a methodology of optimal sampling design has been investigated and developed for water drainage networks. Optimal hydrometer sensors locations along the Amato River (South Italy) have been defined by optimizing a two-objective function that maximizes the calibrated Model accuracy and minimizes the total metering cost. This problem has been solved by using an enumerative search solution, run on the ENEA/CRESCO HPC infrastructure, evaluating the exact Pareto-front by efficient computational time.

Introduzione

I modelli idrologici ed idraulici unitamente ai sistemi di monitoraggio in continuo e distribuiti rappresentano strumenti fondamentali per la gestione e la previsione del rischio inondazione. L'applicazione della modellistica idrologica/idraulica sui sistemi di drenaggio urbano ed extraurbano permette la simulazione dei processi di formazione e propagazione del deflusso idrico in corrispondenza di eventi meteorici, in particolare per le precipitazioni di maggiore intensità, che

danno origine ad eventi di piena. Tuttavia, al fine di ottenere scenari di rischio che rispecchino il reale comportamento dell'area oggetto di studio è necessario misurare parametri idro/climatici in diversi punti del bacino idrografico. A tal fine, è necessario realizzare delle reti di monitoraggio che rispondano a specifici criteri e permettano una calibrazione efficace ed efficiente dei modelli per l'elaborazione di scenari di rischio. La precisione della calibrazione dei modelli dipende dalla qualità, dalla quantità e dal tipo di dati misurati. Tali dati non possono essere monitorati in ogni punto del sistema, sia per l'impossibilità di fare ciò, sia perché un aumento di misure non sempre porta ad un miglioramento delle prestazioni. Non potendo disporre di tutti i sensori necessari è importante prevedere un'adeguata pianificazione delle misure, generalmente indicata in letteratura come *sampling design* (Bush and Uber, 1998). In particolare, problema della localizzazione dei sensori può essere di due tipi. In un caso può non essere presente nessun sensore sul sistema oggetto di studio e quindi il numero di sensori da aggiungere è dettata da esigenze esclusivamente economiche. Nell'altro, i sensori si aggiungono ad una rete di monitoraggio preesistente e pertanto bisogna valutare il beneficio ottenuto dall'incremento dei sensori sull'area. Il presente lavoro ricade nell'ambito dell'incremento di una rete esistente, infatti, attraverso l'utilizzo di algoritmi specifici, è stata integrata una rete di monitoraggio preesistente al fine di migliorare la calibrazione dei modelli numerici di simulazione idraulica e idrologica, andando a definire nuovi punti di misura da aggiungere alla rete. L'area oggetto di studio è il bacino Amato situato in Calabria (Italia).

La metodologia

I modelli idrologici e idraulici impiegati per la definizione di scenari di rischio inondazione presentano numerosi elementi di incertezza ovvero parametri geomorfologici, fisici difficili da definire con precisione. Pertanto per ottenere un modello affidabile del funzionamento del sistema oggetto di studio è necessario seguire una procedura di calibrazione del modello il più accurata possibile. Il processo di calibrazione modifica i parametri incerti del modello in modo da rendere minimo lo scarto tra le grandezze simulate e quelle misurate. In tale contesto si inserisce il presente lavoro con la messa a punto di una metodologia per la definizione di schemi ottimali di posizionamento degli idrometri sul bacino pilota dell'Amato, basata sulla massimizzazione dell'accuratezza della calibrazione del modello idrologico impiegato per la simulazione del flusso idrico superficiale e nel contempo sulla minimizzazione dei costi di installazione e manutenzione associati ai sensori.

I modelli idrologici sono solitamente calibrati in funzione di parametri incerti quali la percentuale di permeabilità del sistema (i.e. Curve Number-CN parameter), la scabrezza dei canali ect. Si osserva che la configurazione spaziale dei sensori che determina un errore poco significativo tra i valori misurati dai sensori e quelli simulati dal modello, andrebbe identificata sulla base di una stima accurata dei parametri di calibrazione. Tuttavia l'accuratezza della stima dei parametri di calibrazione dipende dal posizionamento ottimale dei sensori (Meier and Barkdoll, 2000). Tale problema ciclico può essere superato con la definizione di un problema stocastico, formulato come problema di ottimizzazione dell'accuratezza della calibrazione dove i parametri incerti del modello sono modellati come variabili aleatorie caratterizzate da specifiche funzioni di densità di probabilità (pdf).

Approcci di questo tipo sono riferiti in letteratura come basati sul criterio di sensitività dei punti di misura rispetto ai parametri di calibrazione. In particolare, matrici di covarianza e Jacobiani sono utilizzati per trovare schemi ottimali di monitoraggio dei parametri che entrano in gioco nei modelli di simulazione utilizzati (Bush and Uber, 1998; Lansey et al., 2001).

Sulla base di tali premesse, la metodologia proposta definisce il problema stocastico di determinazione del numero e del posizionamento ottimale degli idrometri sul bacino pilota come soluzione di un problema di ottimizzazione multi-obiettivo ovvero, (1) massimizzare l'accuratezza del modello idrologico calibrato; (2) minimizzare il costo dei sensori per l'acquisizione delle misurazioni.

Per quantificare l'accuratezza della predizione del modello calibrato, viene utilizzato il metodo FOSM (First Order Second Moment) che approssima le matrici di covarianza dei parametri rilevanti del modello e delle predizioni. In particolare, si assume di avere N_s sensori uguali, con deviazione standard σ , installati in N_s locazioni distinte e che per ognuno di questi siano disponibili N_t misure effettuate in istanti successivi.

Siano y_i l'insieme delle misure disponibili con $i=1, \dots, N_0$ e N_0 il numero di misure effettuate sia nello spazio N_s che nel tempo N_t :

$$N_0 = N_s \cdot N_t \quad (1)$$

Siano indicati con a_k i parametri di calibrazione con $k=1, \dots, N_a$ dove N_a il numero di parametri considerati. Se le misure sono tra loro indipendenti (ipotesi realistica), la matrice di covarianza associata all'insieme dei parametri si può approssimare come:

$$\text{Cov}_a = \sigma^2 \cdot (J^T \cdot J)^{-1} \quad (2)$$

dove J è la matrice Jacobiana delle derivate:

$$J_{i,k} = \delta y_i / \delta a_k \quad (3)$$

Si noti che questa equazione viene utilizzata per propagare l'incertezza delle misure effettuate su tutti i parametri di calibrazione, in quanto cattura la sensitività relativa ai parametri di calibrazione stimati rispetto tutte le misurazioni utilizzate. La sensitività dipende da numerosi fattori tra cui la vicinanza dei punti di misura, la possibile correlazione tra i parametri di calibrazione ecc.; di conseguenza, è molto difficile stabilire in anticipo quali parametri saranno più sensibili alle misure raccolte.

Definita l'incertezza dei parametri, è possibile definire l'incertezza relativa al modello calibrato come gli elementi della diagonale della matrice di covarianza delle previsioni:

$$\text{Cov}_z = J_z \cdot \text{Cov}_a \cdot J_z^T \quad (4)$$

dove in questo caso J_z è la matrice jacobiana delle derivate delle grandezze predette:

$$J_{z,i,k} = \delta z_i / \delta a_k \quad i=1, \dots, N_z \quad (5)$$

Dove N_z è il numero di grandezze predette dal modello.

Esistono diversi metodi per il calcolo delle matrici jacobiane J e J_z . Nel nostro caso per valutare gli elementi di queste due matrici si è scelto di utilizzare il metodo alle differenze finite.

Si noti che il modello FOSM presuppone che i parametri della calibrazione e le previsioni del modello siano variabili casuali che seguono una distribuzione normale. Inoltre, poiché si basa su uno sviluppo in serie di Taylor del primo ordine, esso è teoricamente valido solo per modelli lineari. Ciononostante anche per modelli non lineari, l'incertezza stimata dal FOSM, purché sufficientemente piccola, è considerata una buona approssimazione del valore reale.

Per inglobare le incertezze relative alle grandezze predette dal modello, l'incertezza ϕ del modello idrologico è stimata a partire dagli elementi sulla diagonale della matrice di covarianza Cov_z :

$$\phi = 1/N_z \cdot \sum_{i=1}^N \sqrt{\text{Cov}_{z,i,i}} \quad (6)$$

Poiché per valutare ϕ occorre un insieme di valori per i parametri di calibrazione prima che i valori della matrice jacobiana siano calcolati, assumendo che essi seguono una data pdf, si definiscono i ϕ_j come una particolare realizzazione dell'incertezza ϕ ottenuta generando un insieme di valori casuali per i parametri di calibrazione sulla base delle distribuzioni associate a ciascun parametro. Una stima più robusta dell'incertezza del modello che tiene conto anche dell'incertezza dei parametri di calibrazione si può quindi ottenere come:

$$\phi = 1/N_r \cdot \sum_{i=1}^N \phi_i \quad (7)$$

dove N_r è il numero di insiemi di valori casualmente generati per i parametri.

Per generare i valori degli N_r insiemi di parametri dalle funzioni di densità di probabilità associate è stato utilizzato la tecnica di campionamento dell'ipercubo latino (McKay et al., 1979; Helton e Davis, 2003; Post et al., 2008). Questo tipo di campionamento è un metodo statistico per la generazione di un campione di valori plausibili dei parametri da una distribuzione multidimensionale.

Il costo del posizionamento ottimale è surrogato dal numero di sensori usati, divenendo il secondo obiettivo di ottimizzazione. Generalmente, per costo si considerano sia i costi di impianto che di

esercizio. Siccome risulta difficile stimare queste quantità, di solito si procede ottimizzando il numero di sensori N_s da utilizzare. Nel caso in esame, il problema di ottimizzazione si formula come:

$$\text{Min } N_s \quad (8)$$

con $N_{\min} \leq N_s \leq N_{\max}$ numero minimo e massimo di sensori dipendente dal dimensionamento della rete di monitoraggio da realizzare.

In definitiva, il numero e il posizionamento ottimale degli idrometri si ottengono risolvendo il problema di ottimizzazione multi-obiettivo che minimizza contemporaneamente ϕ ed N_s .

Dunque la soluzione al problema posto è data da un insieme di terne:

$$(n, \phi_n, L_n) \quad (9)$$

Dove $n=1, \dots, n_{\max}$ (n_{\max} è il numero massimo di sensori da posizionare) indica il numero di sensori da localizzare, ϕ_n indica l'incertezza minima ottenibile con n sensori ed L_n il posizionamento degli n sensori corrispondente all'incertezza minima. L'equazione (9) definisce il problema di ottimizzazione multi-obiettivo relativo all'incertezza.

Si osservi che la risoluzione dell'Eq. (9) è computazionalmente onerosa per il calcolo iterativo della matrice Jacobiana. In letteratura la risoluzione di siffatti problemi di ottimizzazione è eseguita generalmente attraverso l'utilizzo di algoritmi MOGA (Multi-Objective Genetic Algorithm) unitamente anche a reti neurali artificiali (ANN) (Behzadian et al., 2009). Nell'ambito di questo lavoro, abbiamo voluto sperimentare un approccio alternativo quale quello esaustivo in virtù della disponibilità della potenza di calcolo computazionale del Data Center ENEA-CRESCO. Il problema sopra descritto si presenta come problemi di ottimizzazione in cui l'obiettivo non coincide con una sola funzione da massimizzare o minimizzare, ma è formato da due grandezze incommensurabili che non possono essere combinate in un'unica funzione. In tali problemi tipicamente non esiste una soluzione che ottimizza entrambi gli obiettivi per cui il procedimento di ottimizzazione è mirato ad ottenere un risultato la cui struttura è più complessa: l'insieme di Pareto (Van Veldhuizen et al. 1998). Questo è l'insieme delle soluzioni Pareto-ottime, dove una soluzione è Pareto-ottima se non esiste un'altra soluzione che sia migliore della prima rispettando entrambi gli obiettivi.

Il caso studio

Il problema del posizionamento ottimale di sensori idrometrici è stato applicato al bacino Amato situato in Calabria. La sua estensione planimetrica complessiva di 443,83 km², con sezione di chiusura coincidente con la foce del Mar Tirreno. Dall'analisi della rete esistente e funzionante sulle aree pilota emerge l'opportunità di integrare la rete idrometrica al fine di ottimizzare l'implementazione dei modelli e quindi le azioni di previsione e allerta. In Figura 1 è riportata la modellazione del bacino dell'Amato realizzata tramite il software SWMM5.

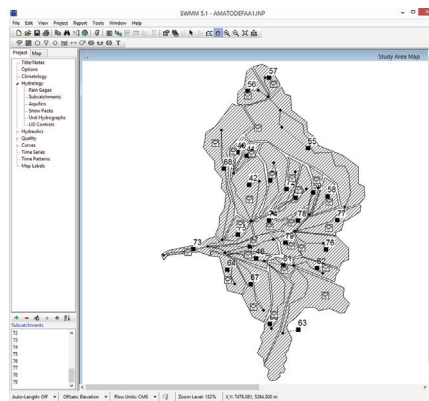


Figura 1 - Modellazione del bacino Amato, Calabria (Italia).

La procedura implementata consiste in (1) individuazione dei parametri incerti (e.g. scabrezza dei canali, percentuale di permeabilità etc.); (2) definizione del range di variazione di tali parametri; (3) individuazione della grandezza da monitorare tramite sensori installati sulla rete di drenaggio (e.g. altezza idrometrica, portata etc.); (4) ed infine, valutazione, attraverso una procedura iterativa di ottimizzazione, il valore dei parametri che riduce lo scarto tra valori simulati dal modello e valore misurato dai sensori.

Nel caso in esame, si è partiti col definire come parametro simulato da ottimizzare la portata misurata in 14 punti coincidenti con i punti di chiusura dei sottobacini in cui è stato suddiviso il bacino dell'Amato, mentre come parametro incerto si è assunto la capacità di assorbimento stimata sul bacino stesso (Curve Number). Questo parametro riassume l'attitudine propria e specifica del bacino a produrre deflusso e consente di determinare il deflusso diretto o pioggia efficace cioè la frazione della pioggia totale che direttamente e in maniera preponderante contribuisce alla formazione dell'evento di piena.

Con riferimento alla metodologia FOSM vanno specificati i seguenti insiemi di grandezze:

- ✓ Misure effettuate: y_t con $t = 1, \dots, N_s$ dove $N_s = N_z \cdot N_t$ rappresentano valori misurati su N_z locazioni in N_t intervalli di tempo;
- ✓ Misure predette: \hat{x}_t con $t = 1, \dots, N_z$ formate da valori predetti su N_z locazioni;
- ✓ Parametri di calibrazione: a_k con $k = 1, \dots, N_a$ che rappresentano i parametri del modello caratterizzati da valori incerti.

Chiaramente i parametri di calibrazione a_k saranno costituiti dai CN, per cui il numero N_a di tali parametri sarà pari al numero di sottobacini. Per quanto riguarda le misure effettuate e le misure predette, queste saranno costituite dalle portate calcolate da SWMM5, per cui avremo $N_s + N_z = 14$. Descritta la procedura per calcolare l'accuratezza di calibrazione ottenuta dall'installazione di un particolare gruppo di idrometri sul bacino, vediamo ora illustrato come tale procedura possa essere usata per valutare il fronte di Pareto del problema in esame.

L'algoritmo per valutare il fronte di Pareto può essere descritto tramite i seguenti passi:

- ✓ Per ogni valore $N_s = 1, \dots, 14$ si considerino tutti i gruppi di N_s locazioni selezionabili tra le 14 possibili. Si tenga presente che, per un fissato valore di N_s , i gruppi possibili da considerare sono $\binom{14}{N_s} = \frac{14!}{(14-N_s)!N_s!}$
- ✓ Fissato N_s , per ognuno dei $\binom{14}{N_s}$ gruppi di N_s locazioni, si calcoli la corrispondente accuratezza di calibrazione.
- ✓ Si determini il valore massimo $\Phi_{N_s}^*$ tra gli $\binom{14}{N_s}$ calcolati ed il gruppo $L_{N_s}^*$ di N_s locazioni che lo ha prodotto.

Al termine, si ottiene un insieme di terne $(N_s, \Phi_{N_s}^*, L_{N_s}^*)$ che fornisce, per ogni $N_s = 1, \dots, 14$, l'accuratezza massima $\Phi_{N_s}^*$ ottenibile con N_s sensori e la particolare localizzazione $L_{N_s}^*$ degli N_s sensori corrispondente all'incertezza minima.

Risultati

In Figura 2 è mostrato il fronte di Pareto elaborato per il problema in esame: sulle ascisse è presente il valore degli idrometri installati, mentre sulle ordinate è il valore dell'accuratezza massima normalizzata.

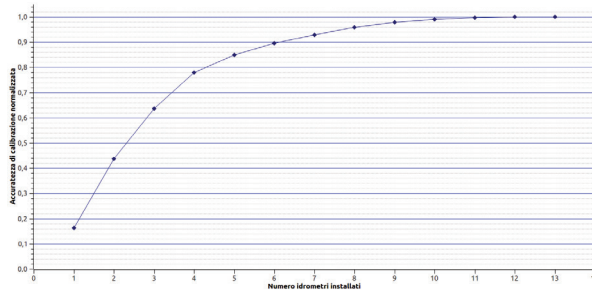


Figura 2 - Fronte di Pareto: relazione tra numeri di idrometri e accuratezza di stima.

Analizzando la Figura 2 si osserva come ci sia un miglioramento dell'accuratezza estremamente rapido nella prima parte della curva ovvero fino ad un numero massimo di 4 idrometri installati. Infatti, per valori maggiori di 4 idrometri installati, il tasso di miglioramento dell'accuratezza della calibrazione del modello in esame comincia a diminuire. Si evince pertanto che un incremento del numero di idrometri (> 4 idrometri) installati comporta un miglioramento meno significativo della calibrazione rispetto ad un numero di idrometri compreso tra 1 e 4.

Nelle figure seguenti, gli schemi di posizionamento ottimale elaborati sono visualizzati spazialmente tramite apposite mappe. Sono riportate le sole mappe relative agli schemi di posizionamento ottimale da 1 e 3 idrometri, in considerazione anche dei risultati ottenuti relativamente al miglioramento dell'accuratezza della calibrazione che cresce più lentamente per un numero maggiore di 4 idrometri installati.

Dunque, in Figura 3 si riporta la soluzione ottimale nel caso di posizionamento di un solo idrometro. L'idrometro è posizionato in prossimità della sezione di chiusura dell'intero bacino.



Figura 3 - Schema ottimale per il posizionamento di un idrometro.

La soluzione ottimale per il posizionamento di tre idrometri (Figura 4) è anch'essa conservativa. Secondo tale soluzione, il terzo idrometro va inserito nell'area nord-ovest del bacino.

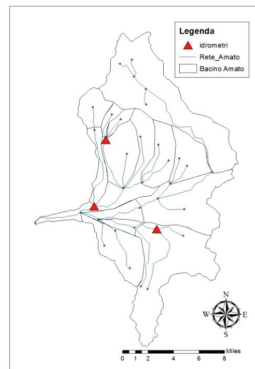


Figura 4 - Schema ottimale per il posizionamento di tre idrometri.

Le soluzioni ottenute hanno permesso di definire le coordinate spaziali dei sensori da integrare alle reti esistenti in modo da ottimizzare i modelli per la previsione del rischio idraulico ed idrologico.

Riferimenti bibliografici

- Behzadian, K., Kapelan, Z., Savic, D., & Ardeshir, A. (2009). Stochastic sampling design using a multi-objective genetic algorithm and adaptive neural networks. *Environmental Modelling & Software*, 24(4), 530-541.
- Bush, Cheryl A., and James G. Uber. "Sampling design methods for water distribution model calibration." *Journal of Water Resources Planning and Management* 124.6 (1998): 334-344.
- Helton, Jon C., and Freddie Joe Davis. "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems." *Reliability Engineering & System Safety* 81.1 (2003): 23-69.
- Lansey, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J., & Haan, C. T. (2001). Calibration assessment and data collection for water distribution networks. *Journal of Hydraulic Engineering*, 127(4), 270-279.
- McKay, Michael D., Richard J. Beckman, and William J. Conover. "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics* 21.2 (1979): 239-245.
- Meier, Roger W., and Brian D. Barkdoll. "Sampling design for network model calibration using genetic algorithms." *Journal of Water Resources Planning and Management* 126.4 (2000): 245-250.
- Post, Eric, and Christian Pedersen. "Opposing plant community responses to warming with and without herbivores." *Proceedings of the National Academy of Sciences* 105.34 (2008): 12353-12358.
- Van Veldhuizen, David A., and Gary B. Lamont. "Evolutionary computation and convergence to a pareto front." *Late breaking papers at the genetic programming 1998 conference*. Stanford University Bookstore, University of Wisconsin, Madison, Wisconsin, USA, 1998.