

Presentazione del simulatore

Università di Pisa

Emanuele Urselli, Mercoledì 20 Gennaio 2021

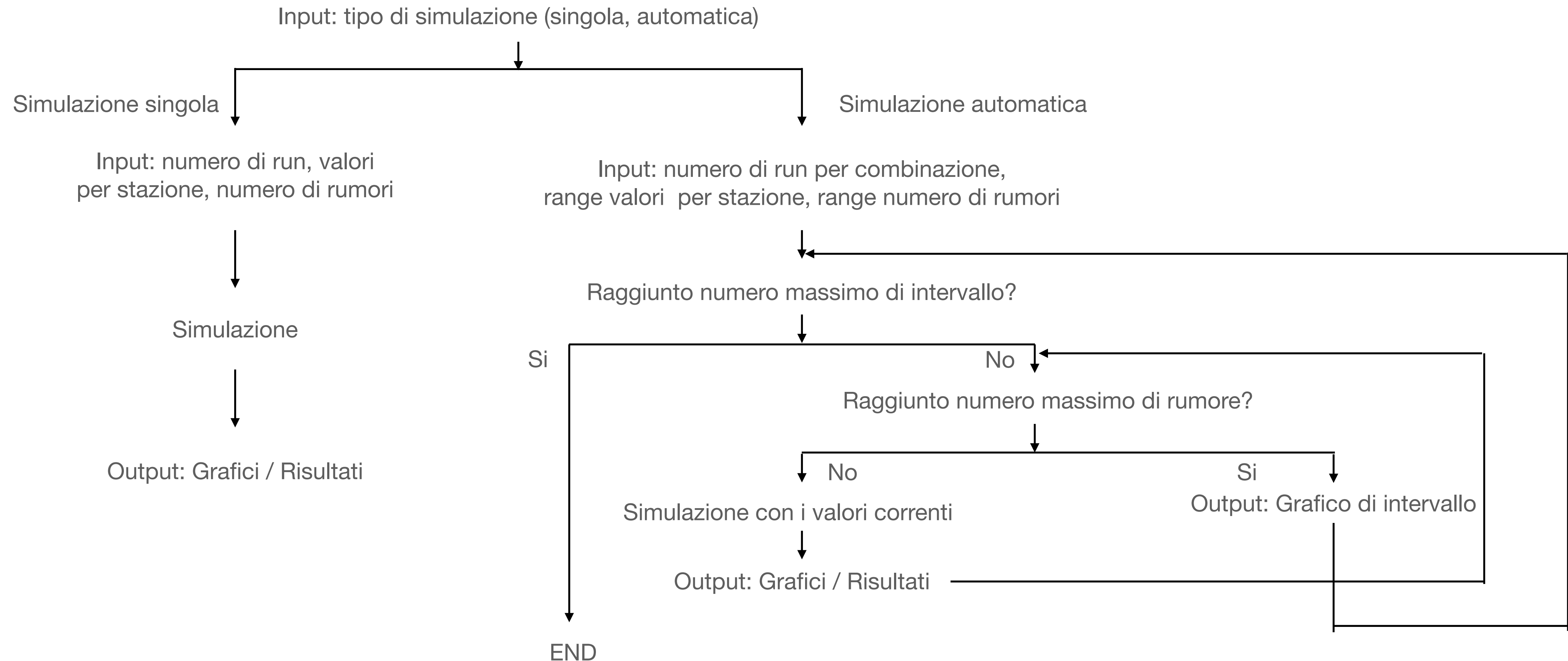
Presentazione del simulatore

Breve introduzione

- Il simulatore prendendo dei valori di input, esegue il processo di simulazione e proietta i risultati attraverso l'utilizzo di grafici. (Approfondito in seguito)
- Il simulatore permette di eseguire due tipi di simulazioni:
 - **Simulazione singola:**
Esegue la simulazione con un numero fisso di valori per stazione e un numero fisso di valori sostituiti da 'rumore'.
 - **Simulazione automatica:**
I valori relativi al numero di valori per stazione e al numero di valori sostituiti da 'rumore' spaziano in un range che va da 1 all'input inserito.
Per ogni combinazione, sono generati grafici diversi. (Approfondito in seguito)

Presentazione del simulatore

Diagramma di funzionamento (Una iterazione del software)



Presentazione del simulatore

Input

- Come accennato in precedenza, il software esegue due tipi di simulazione:
 - **Singola:** In questa simulazione, i valori di input sono i seguenti:
Numero di run da eseguire, numero di valori per stazione, numero di valori sostituiti dal 'rumore'.
 - **Automatica:** In questa simulazione, i valori di input sono identici ai precedenti, ma sono utilizzati in modo diverso.
In breve, si esegue una simulazione singola per ogni combinazione tra ***numero di valori per stazione*** e ***numero di valori sostituiti dal rumore***, fino al raggiungimento dei valori di input inseriti dall'utente.
Per ognuna delle combinazioni, viene eseguito un ***numero di run*** pari al valore rispettivo ricevuto in input dall'utente.
- I dati utilizzati sono all'interno della directory StazioniBolognaWU.
12 valori, 10 stazioni.
Ogni valore fa riferimento ad un intervallo di 10 minuti.
I valori sono tutti compressi in un intervallo di 2 ore.

Presentazione del simulatore

Output

- I risultati delle simulazioni vengono memorizzati nella directory *Results*.
- Simulazione:
 - **Singola:** Al termine della simulazione, i risultati vengono memorizzati nella directory definita dal timestamp a tempo della simulazione, sotto la directory 'Single Simulations'.
I file generati sono:
 - **Result_runs.csv:** Risultati delle singole run eseguite
 - **Noisy_matrix.pdf:** Grafico andamento delle medie finali delle singole run
 - **Final_Results.csv:** Risultati complessivi della simulazione eseguita
 - **Bar_chart_avg.pdf:** Grafico a barre relativo alle medie delle tre matrici (reale, rumorosa, e con valori sostituiti dal 'rumore')
 - **Bar_chart_diff_avg.pdf:** Grafico a barre relativo alle differenze tra le medie delle matrici con i dati reali, i dati rumorosi e i dati con valori mancanti e sostituiti dal rumore.
 - Automatica: Per ogni combinazione (numero di valori per stazione, numero di valori sostituiti da 'rumore') genera i risultati appena mostrati, nella directory relativa all'intervallo corrente. All'interno di quest'ultima directory, i risultati con un numero diverso di rumore, sono inseriti nelle rispettive directory.
Prima di incrementare il numero di valori per stazione, viene generato un grafico di intervallo: mostra l'andamento della media con tutte le combinazioni relative al rumore.

Presentazione del simulatore

Metriche utilizzate e considerazioni

- Nel paper originale, per:
 - Laplace-distribution:
Parametro di posizione = 0
Viene considerato solo il valore di scale.
 - Gamma-distribution:
Parametro shape = #partecipanti (10 in questo caso)
- Ho considerato l'utilizzo di diversi valori di scale per i due tipi di rumore.
Le variabili con gamma-distribution sono utilizzate per aggiungere rumore a tutti i dati iniziali, per questo, si è utilizzato un valore di scale basso (0.01) per ottenere un rumore tollerabile.
Per le variabili con laplace-distribution, si è utilizzato un valore contenuto nell'intervallo $[1/3 - 1/4]$ della media dei valori reali, in modo tale da generare un valore quasi vicino ai valori reali, ma comunque compreso di rumore.
- Sistema poco tollerabile al rumore. Col 10% si ha una perdita di circa 1,3 gradi. (Varia - ovviamente - in base ai settings delle distribuzioni)

Presentazione del simulatore

Risultati: Result_runs.csv

- Questo file contiene i risultati delle singole runs.
#Run: Numero della run
N_of_values: Numero dei valori considerati (totali)
N_of_noises: Numero di valori sostituiti da 'rumore'
Real_Avg: Media dei valori reali (senza rumore)
Avg_modified_mat: Media dei valori (col rumore)
Noisy_mat_avg: Media dei valori (con n_of_noises valori sostituiti da 'rumore')
Real-Modified_avg: Differenza tra i primi due valori
Real-Noisy_avg: Differenza tra il primo e il terzo valore
Modified-Noisy_avg: Differenza tra il secondo e il terzo valore

Presentazione del simulatore

Risultati: Final_Results.csv

- Questo file contiene i risultati complessivi della simulazione.
Total_avg_modified_mat: Media complessiva ottenuta considerando i valori delle matrici rumorose
Total_noisy_mat_avg: Media complessiva ottenuta considerando i valori delle matrici con un determinato numero di elementi sostituiti da 'rumore'
Total_real-modified_avg: Media complessiva della differenza tra la media reale e la media rumorosa
Total_real-noisy_avg: Media complessiva della differenza tra la media reale e la media con un numero di valori sostituiti da 'rumore'
Total_modified-noisy_avg: Media complessiva della differenza tra la media rumorosa e la media con un numero di valori sostituiti dal 'rumore'

Presentazione del simulatore

Risultati: Grafico di intervallo (Per simulazioni automatiche)

- Nelle simulazioni automatiche, per ogni intervallo (numero di valori per stazione) diverso, viene generato un grafico che mostra l'andamento della media dei valori con un numero di valori sostituiti da 'rumore'. che varia da 1 all'input dell'utente.

Esempio:

1 valore per stazione, max 9 rumori.
(Numero di run molto basso)

