DIPARTIMENTO DI ELETTRONICA E INFORMAZIONE

POLITECNICO DI MILANO

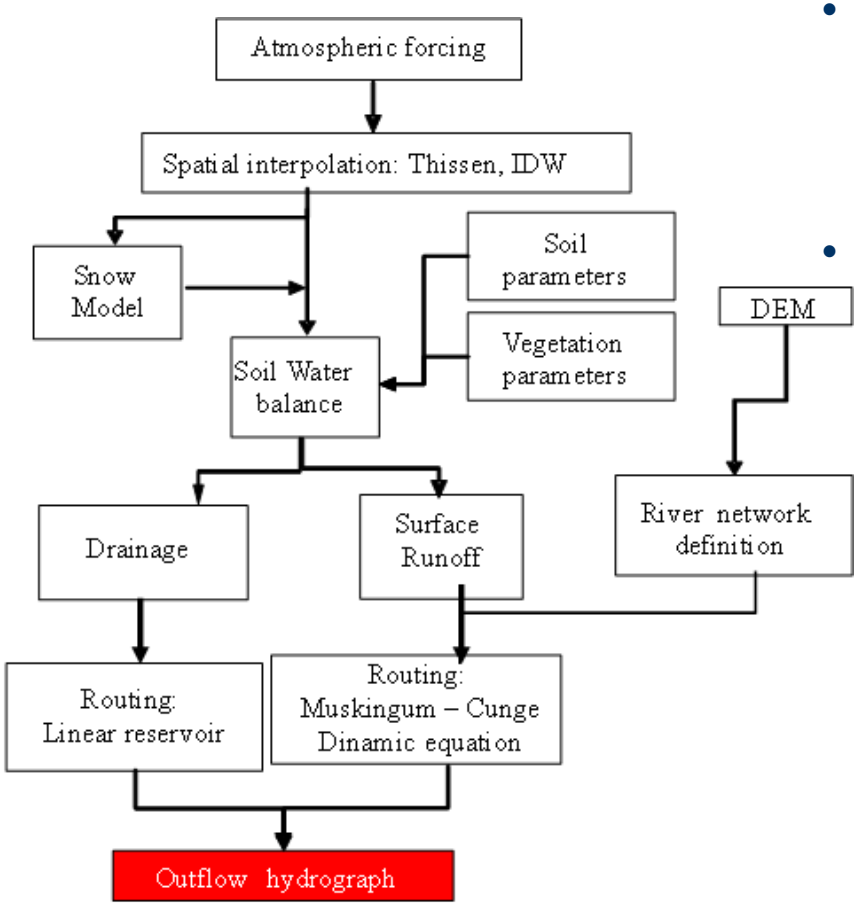
EGU 2012

Automatic identification and placement of measurement stations for hydrological discharge simulations at basin scale

P. R. Grassi, A. Ceppi, F. Cancarè, G. Ravazzani, M. Mancini, D. Sciuto

Paolo Roberto Grassi (grassi@elet.polimi.it)

Introduction



Full scheme of the rainfall-runoff distributed hydrological model FEST-WB, physically based

(Mancini, 1990 → Ravazzani et al., 2010)

- Accurate forecasts of hydro-meteorological events are important to prevent and mitigate the effect of dangerous events such as floods or drought
- Politecnico di Milano developed a distributed physically based rainfall-runoff FEST-WB model, which is daily used to generate runoff simulations and hydrological forecasts for nowcasting monitoring, and as a civil-protection tool:
  - **Input**: a set of observed meteorological data (solar radiation, precipitation, air temperature and relative humidity)
  - **Collected by**: ARPAP (Environment Protection Regional Agency of Piedmont) hydro-meteorological station network and MeteoSwiss
  - **Output**: discharge forecasts in various hydrological basins

Introduction

- Deployment and maintenance of measurements stations is **expensive**
- Density and position of stations depends on topological and hydrological characteristics of the area of interest
- Placement of measurement stations dramatically affects the prediction accuracy and thus it must be designed correctly
- Moreover, the kind of measurement of the station affects the placement:
  - Pluviometric stations placement should follows the topology of the area according to the impact of the hydro flows
  - Thermometric stations placement should provide a complete and accurate map of the temperatures in the area

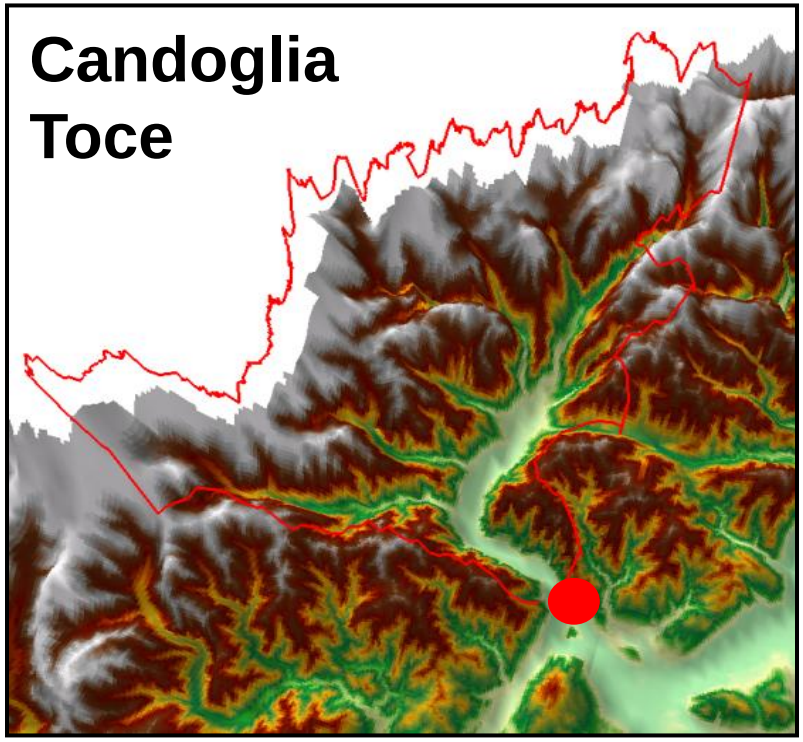
Thermo-hygrometer

Rain gauge

Weather station

Objectives

- The aim of this study is to identify a **topology-independent** and **model-independent** methodology for the analysis of input-output relationships in hydro-meteorological forecasting
- It detects the minimum number and the position of the measurements stations that provides a low prediction error
- This work has been validated on the Toce river basin, an Alpine watershed, located in North-West of Italy
- The methodology can be extended to other areas



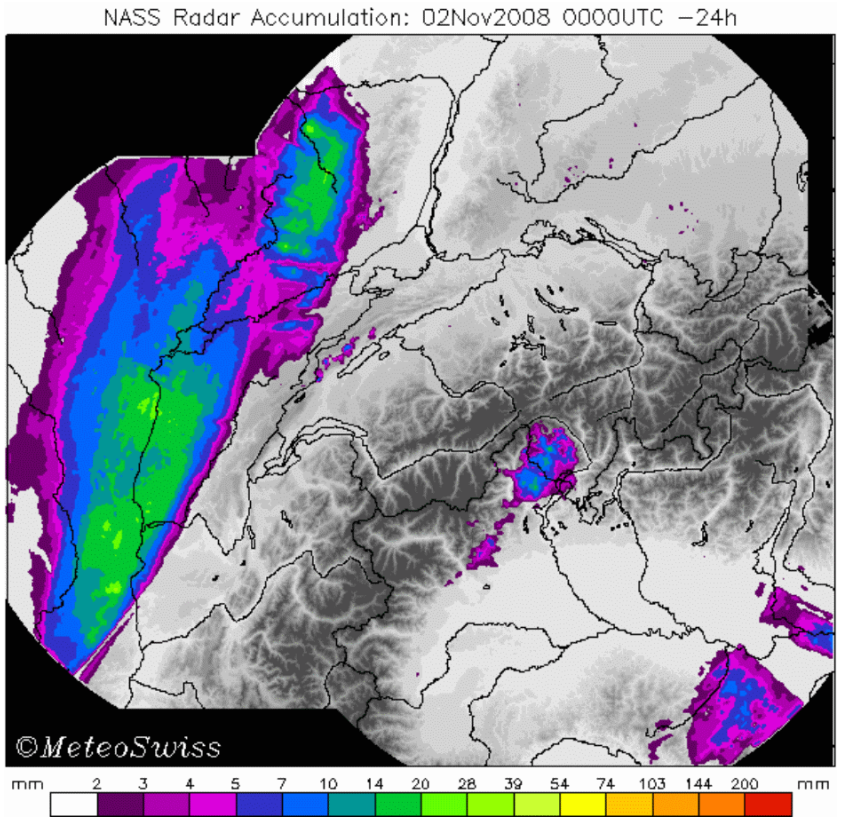
Case Study

Days	Toce
1 November 2008	6.5
2 November 2008	7.2
3 November 2008	52.2
4 November 2008	79.0
5 November 2008	95.4
Cumulated Precipitation [mm]	**240.3**

Peak discharge (Q_{max}) on 5 November 2008 @ Candoglia:

Q_{max} observed: 916 $m^3 s^{-1}$

Q_{max} Simulated with FEST-WB: 992 $m^3 s^{-1}$

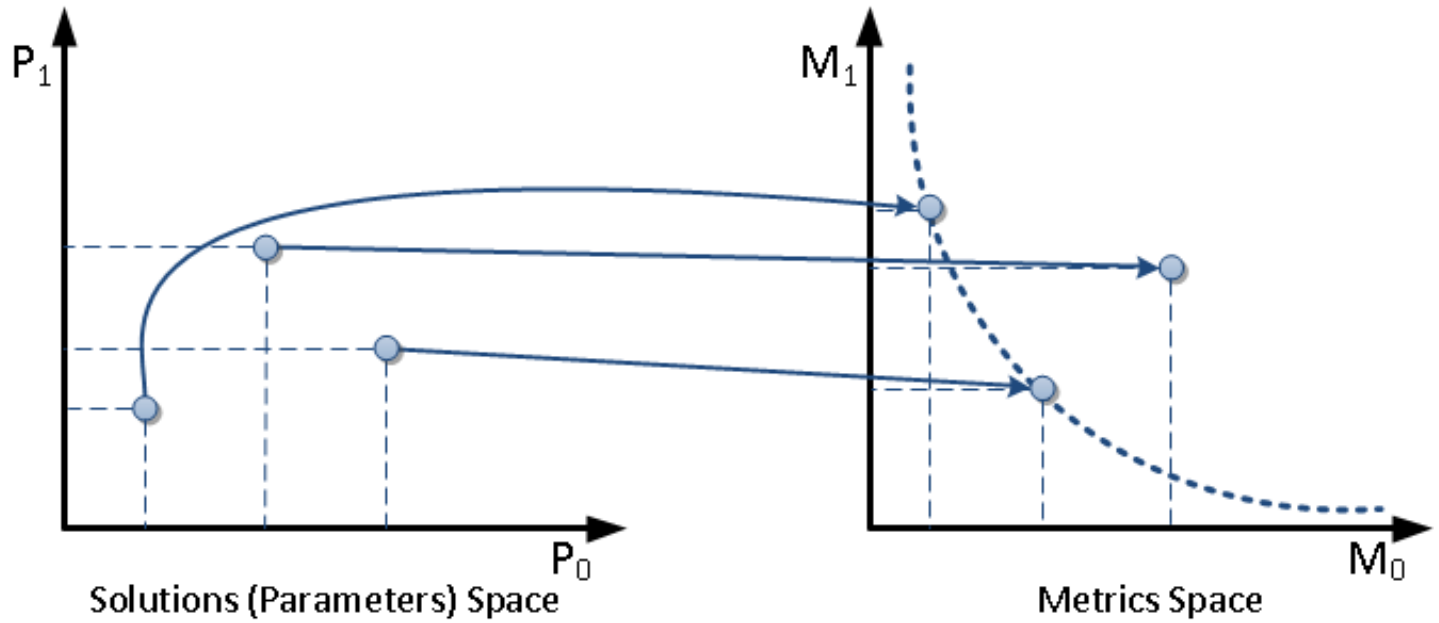


Radar accumulation (mm) over Piemonte and South Switzerland

Design Space Exploration

- The **Design** is a process that aims at identifing the optimal configurations of parameters such as given metrics are optimized

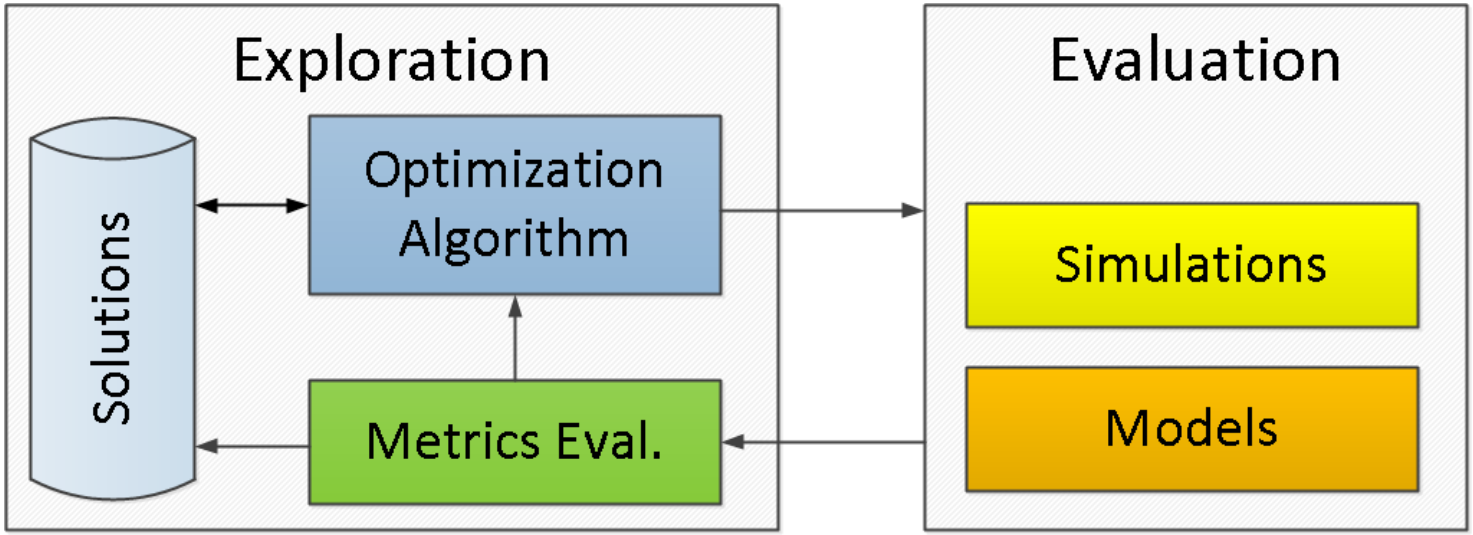
- The **Design Space Exploration** (DSE) is an iterative process aiming at automatically discover Pareto-optimal solutions to the design problem



Design Space Exploration

- Design Space Exploration is composed by two parts:
  - **Evaluation**: evaluates the solutions in order to provide a characterization of the given metrics. It is usually performed using simulations or analytical models;
  - **Exploration**: analyzes the (evaluated) solutions and determines a set of new solutions (hopefully) better than the previous

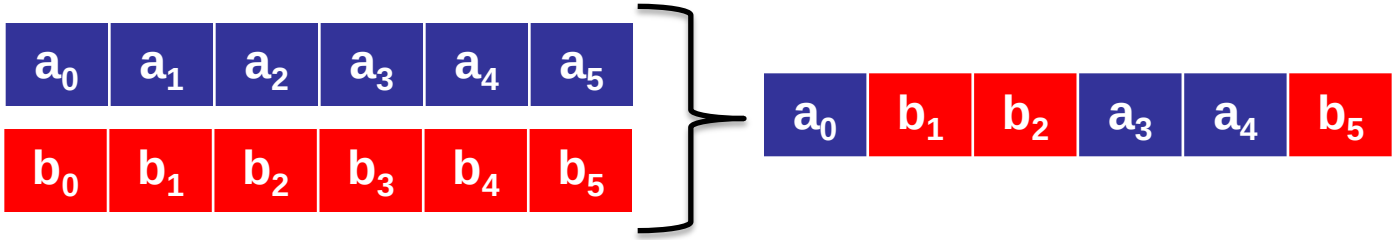
- The Design Space Exploration iterates on the solutions until a certain amount of solutions have been discovered



Genetic Algorithms

- The Genetic Algorithms (GA) were invented by John Holland in the 1970s
- The solutions in a GA are defined as numeric vectors, called **chromosomes**:

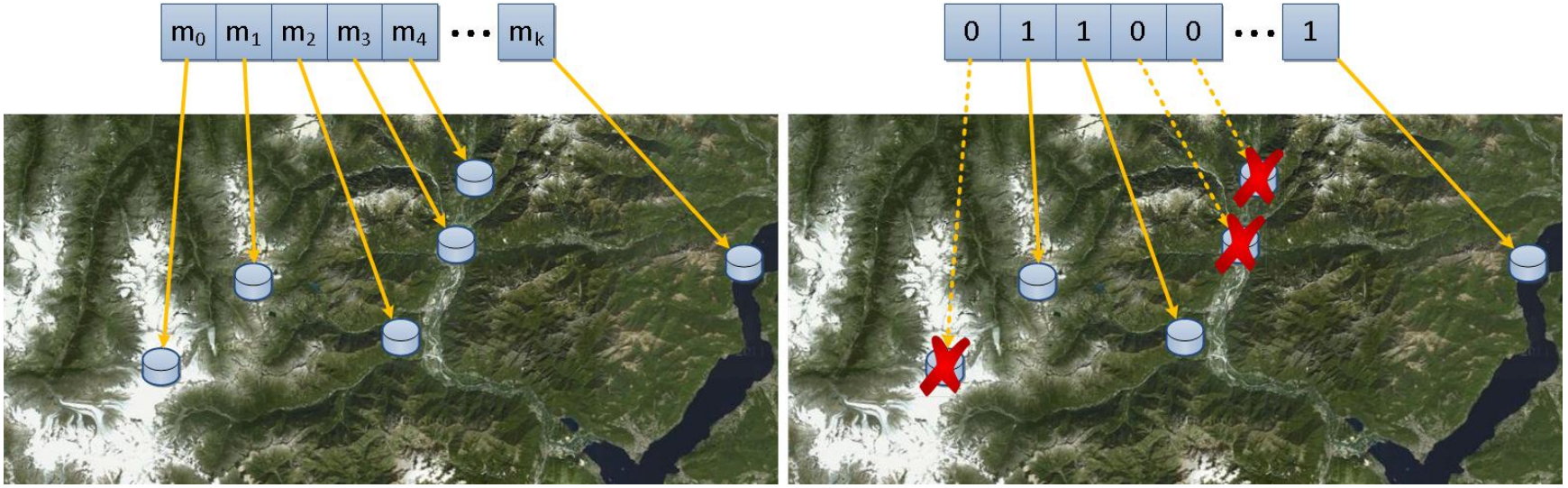
- The GA combines the solutions using two basic operations:
  - **Crossover**: the child solution is generated by randomly mixing the parameters of two parent solutions



  - **Mutation**: a random parameter is changed randomly

DSE for Hydrological Forecasting

- A solution specifies which **measurement stations are used**
- It is a **Boolean vector**, where 0 means “station is not used” and 1 “station is used”



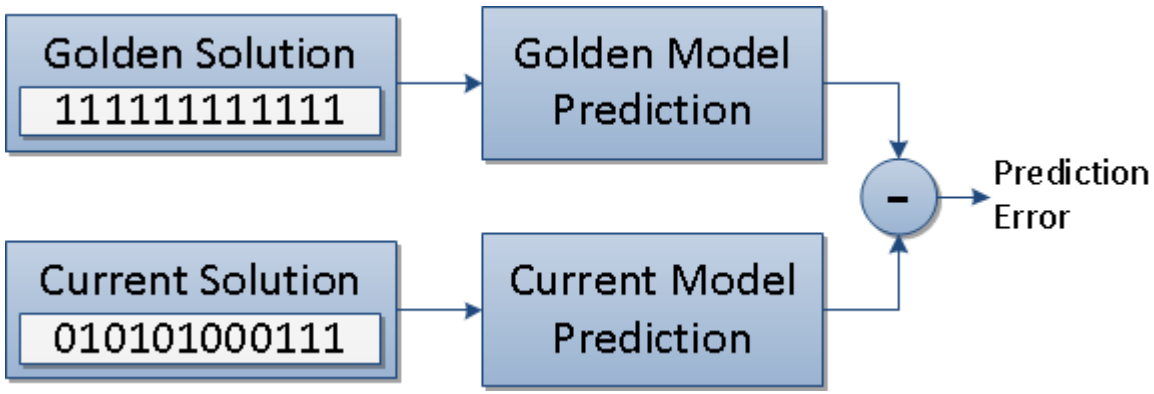
- The hydrological forecasts are performed using selected stations only, assuming that measurements from non-selected stations are not available
- In the analyzed application case study, **118 measurement stations** have been considered: **60 Thermometric** stations and **58 Hydrometric** stations

DSE for Hydrological Forecasting

- Each solution is validated using a **distributed physically based rainfall-rainoff FEST-WB model** developed by the Politecnico di Milano

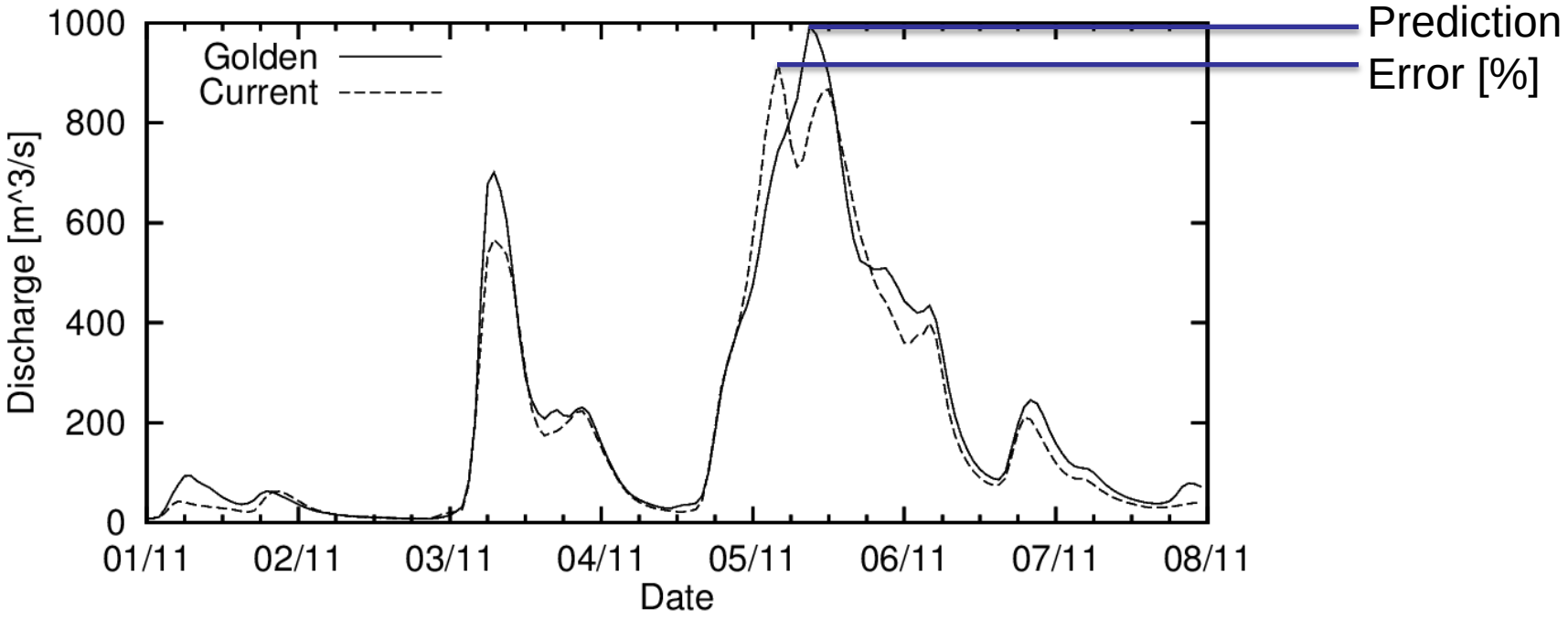
- The outcome of the prediction using the current solution is compared with the prediction using a **golden solution** (a prediction using **all** the stations)

- The objective of the exploration is to reduce the amount of measurement stations such as prediction error minimized



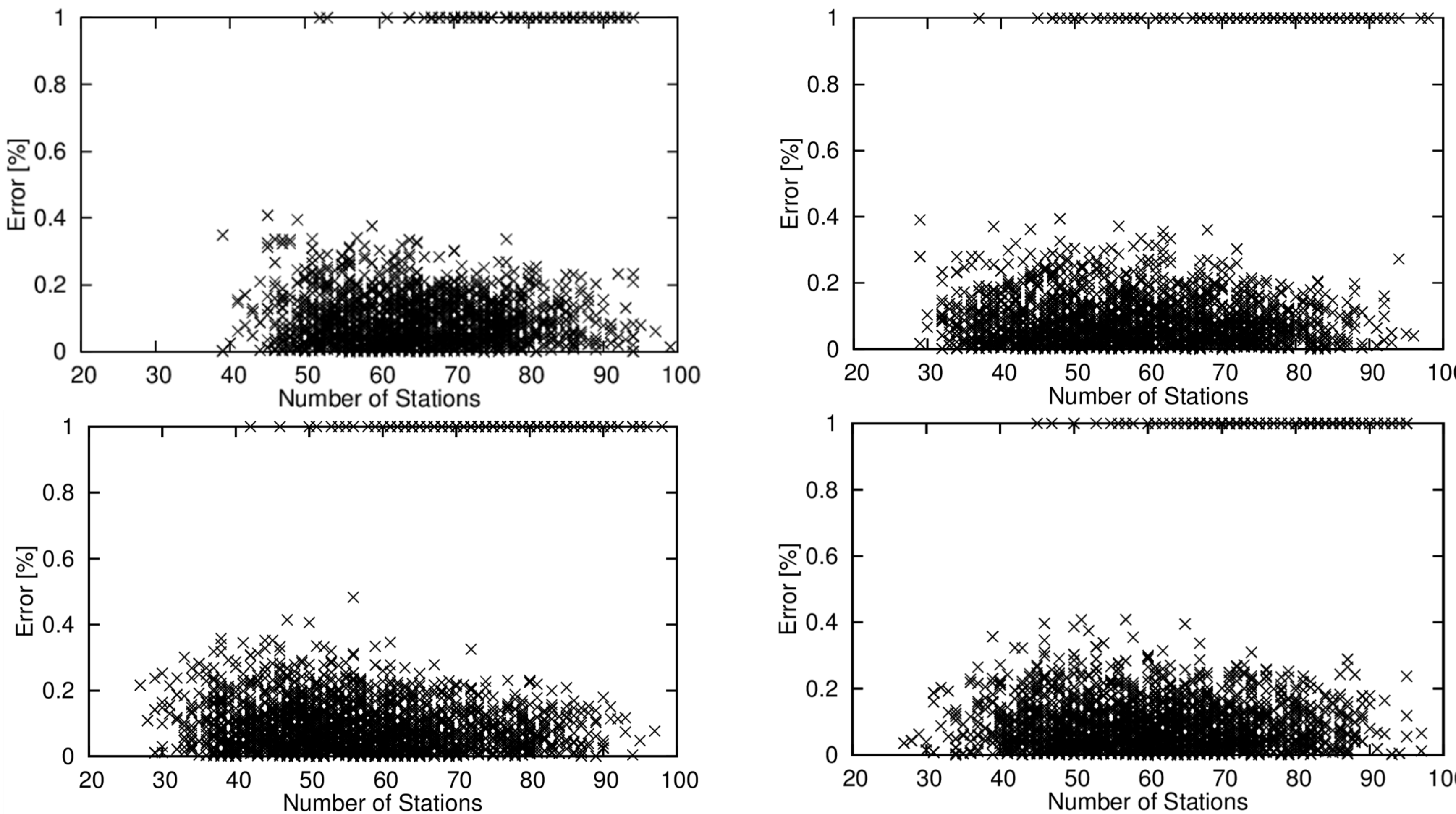
Experimental Case Study

- The proposed approach has been validated on a flooding event occurred on the **Toce river Basin at 5 of November 2008**
- Prediction error is computed as the difference between the maximum value of the Golden prediction and the current prediction



Experimental Results

- At the end of each run (after 3000 solutions have been evaluated), the metrics of the explored space can be viewed on a two-dimensional space
- Each point represents a configuration of the meteorological stations in the metric space



Experimental Results

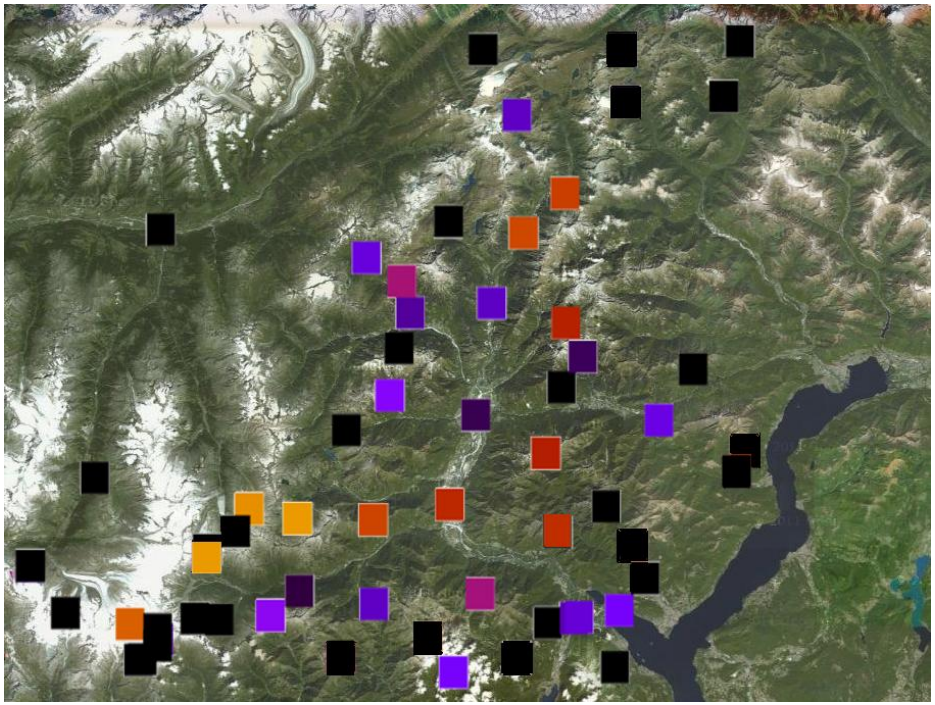
- If we are interested in detect the relevance of a metereological station we must analyze the results
- To get the relevance of the metereological stations we analyze the effect of these stations on the metrics by computing the conditional probabilities:

	Quality of the Solution	
	GOOD	**BAD**
CHOSEN (1)	G1	B1
NOT CHOSEN (0)	G0	B0

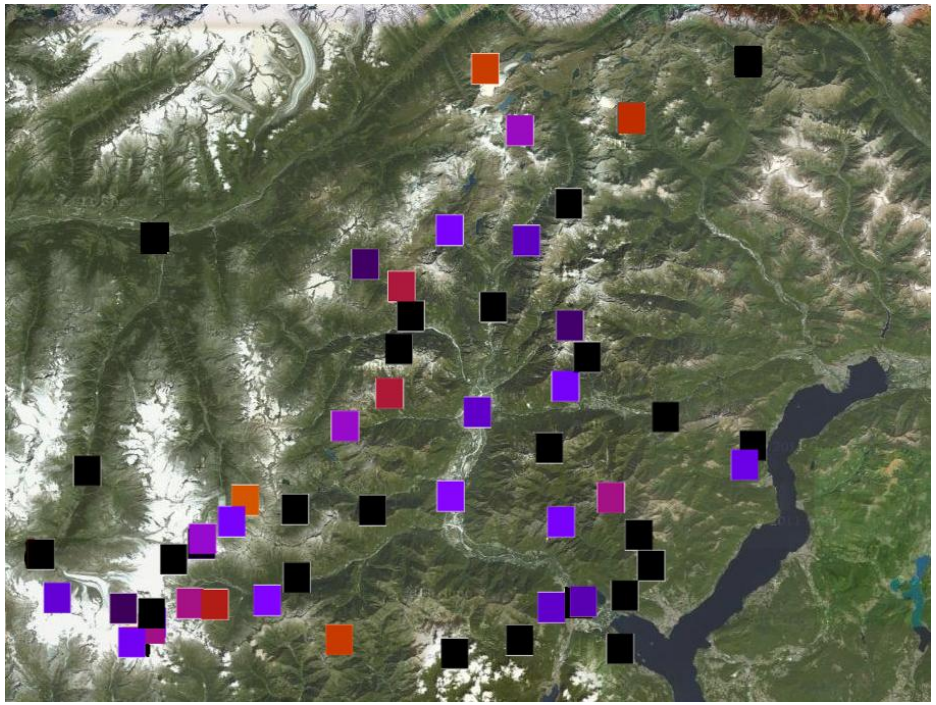
- The higher are G1 and B0 the more relevant is the station in the model:
  - G1: the station is **chosen** and the solution is **GOOD**
  - B0: the station is **not chosen** and the solution is **BAD**

Experimental Results

- Plotting the results on the map…

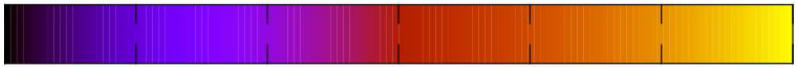


PLUVIOMETRIC



THERMOMETRIC

Station's Relevance



Conclusions

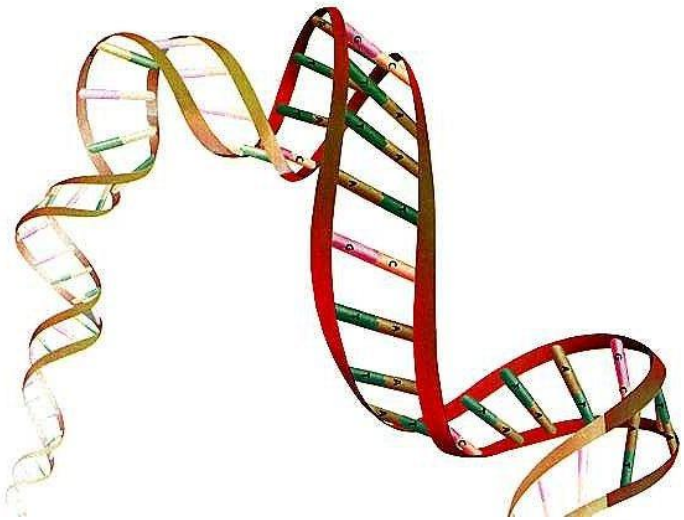
- Accurate forecasts of hydro-meteorological events are important to prevent and mitigate the effect of dangerous events such as floods or drought
- Deployment and maintenance of measurements station is expensive
- The identification of the optimal number and position of measurement stations is currently performed manually
- Thanks to Design Space Exploration technique, it is possible to estimate the number and the position of measurement stations in order to reduce the cost keeping the prediction error as low as possible
- The proposed methodology has been validated on a flooding event of the Toce river but is generic, thus it can be applied on other scenarios

Future Works

- As future work we aim at compare many meta-heuristics algorithms other than GAs in order to identify which one better fits the given problem

- Apply the proposed algorithm on other events and places

- Study the correlation among variables:
  - Geographical Position
  - Density
  - Measurement Kind
  - …



THANKS FOR YOUR ATTENTION

Meta-Heuristics for DSE

- There are several kind of meta-heuristics that can be used for the DSE:
  - **Genetic Algorithms**
  - Simulated Annealing
  - Particle Swarm Analysis
  - Pareto Memetic Algorithm
  - Strength Pareto Evolutionary Algorithm
  - Multiple Start Local Search
  - Ant Colony Optimization
  - Tabu Search

- All the approaches uses historical information about explored solutions to derive better solutions

- In this presentation we illustrates how Genetic Algorithms can be used for our purposes

Genetic Algorithms

- The pseudo-code of a generic Genetic Algorithm is:

```
popsize ← desired population size
P ← {}
for popsize times do
    P ← P U {new random individual}
Best ← {}
repeat
    for each individual P_i ∈ P do
        AssessFitness(P_i)
        if Best = {} or Fitness(P_i) > Fitness(Best) then
            Best ← P_i
    Q ← {}
    for popsize/2 times do
        Parent P_a ← SelectWithReplacement(P)
        Parent P_a ← SelectWithReplacement(P)
        Children C_a, C_b ← Crossover(P_a, P_a)
        Q ← P U {Mutate(C_a), Mutate(C_b)}
    P ← Q
```