European Geosciences Union, General Assembly
2012 Vienna, Austria, 22 – 27 April 2012
HS2.4 - Hydrology and remote sensing for monitoring, modelling and improve the management of urban and rural water systems

OPTIMIZATION OF FLOODPLAIN MONITORING SENSORS THROUGH AN ENTROPY APPROACH

E. Ridolfi[1], K. Yan[2], L. Alfonso[2], G. Di Baldassarre[2], F. Napolitano[1], F. Russo[1], and P. D. Bates[3]

[1]Dipartimento di Idraulica, Trasporti e Strade – Università "Sapienza" , Rome, Italy, elena.ridolfi@uniroma1.it
[2]Department of Integrated Water Systems and Governance - Hydroinformatics, UNESCO-IHE Institute for Water Education, Delft, The Netherlands
[3] School of Geographical Sciences, University of Bristol, Bristol, UK

DIPARTIMENTO DI INGEGNERIA CIVILE EDILE E AMBIENTALE
SAPIENZA UNIVERSITÀ DI ROMA
UNESCO-IHE Institute for Water Education
KULTURisk

Motivation:

- To support the decision making processes of flood risk management and long term floodplain planning, a significant issue is the availability of data to build appropriate and reliable models.
- However, there remains the question of what is the real potential of those global remote sensing data, characterized by different accuracies, for global inundation monitoring and how to integrate them with inundation models.
- In order to monitor the water stage of a reach of the River Dee (UK), a network of cheap wireless sensors (GridStix) was deployed both in the channel and in the floodplain. For updating their layout, the initial number of six sensors is increased up to create a redundant network over the area.

Objective:

- The purpose is to define their location so that to provide as much information as possible and at the same time as low redundancy as possible.
- The problem is posed as a MultiObjective Optimization Problem (MOOP) in which 3 objectives are considered, namely minimizing total correlation (as a measure of redundancy (Alfonso 2010)), maximizing joint entropy (as a measure of information content) and minimizing the entropy of marginal errors between the model and TLEMAC results (as measure of precision).
- The MOOP is solved using the non-sorted genetic algorithm NSGA-II and a Pareto front of optimal sets of sensor locations is obtained. The Pareto front is then analyzed and the most appropriate solution is chosen.
- The 2D raster-based inundation model (LISFLOOD-FP) is used to generate a synthetic GridStix data set, then it is calibrated considering water depth data from the optimal sets.

1. Case study

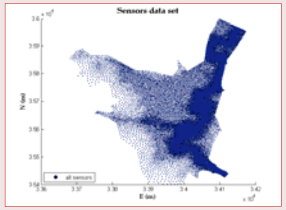
- The case study is a 10-km reach of the River Dee, between the two EA maintained gauging stations at Farndon and Iron Bridge.
- The 2D model LISFLOOD-FP (Bates & De Roo, 2000) is used to generate a synthetic GridStix data set of water stages.
- The Digital Elevation Model (DEM) used for hydraulic model building is the globally and freely available SRTM DEM.
- The boundary condition is the low-magnitude (return period of around 2 years) flooding of December 2006, (Di Baldassarre et al., 2010).

2. Restraining the uncertainty

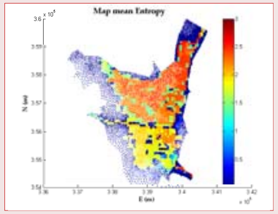
From the original number of 16163 sensors, the half is disregarded in this way:

- The entropy of the errors between each model and TELEMAC results is evaluated for each time step.
- The mean entropy over all time steps is evaluated and the half amount of sensors with the lowest mean entropy is disregarded.

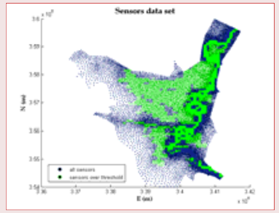
Sensors are considered to be at each computational point.



Entropy values are estimated at each sensor.



Final sensor map:
non-informative sensors are not considered, using a threshold of H=1.86 bit, so that only half sensors with highest entropy are used for analysis.



3. The Optimization problem

- A multiobject function has been considered (MOOP). The set of the optimal sensors have to fulfill:

$$Min(TotC) = Min\{C(X_{sim_1}, X_{sim_2}, ..., X_{sim_N})\}$$
$$Max(JH) = Max\{H(X_{sim_1}, X_{sim_2}, ..., X_{sim_N})\}$$
$$Min(H_{mean}) = Min\left(\frac{1}{N}\sum_{i=1}^{N} H(X_{TELEMAC_i} - X_{sim_i})\right)$$

Where TotC is the total correlation

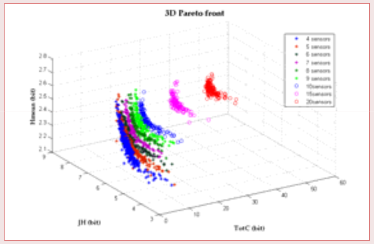
$$C(X_{sim_1}, X_{sim_2}, ...X_{sim_N}) = \sum_{i=1}^{N} H(X_{sim_i}) - H(X_{sim_1}, X_{sim_2}, ...X_{sim_N})$$

JH is the joint entropy:

$$H(X_{sim_1}, ..., X_{sim_N}) = -\sum_{i=}^{N} ... \sum_{j=1}^{M} p_{i,...,j} \log_2(p_{i,...,j})$$

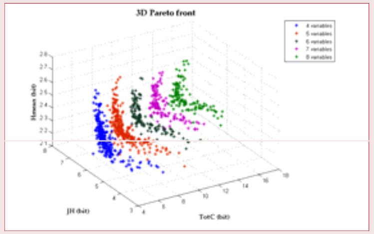
H_{mean} is the mean entropy value of errors between TELEMAC ($X_{TELEMAC}$) results (considered as "truth", in the absence of a suitable observed data set) and simulations (X_{sim})

4. The 3D Pareto front



- Errors entropy minimization increases model precision: form the Pareto front it is possible to observe that in increasing number of sensors the error decreases.
- In increasing the number of sensors, JH increases lightly, while TotC increases highly. This means that once that the optimal number of sensors is deployed, in adding more sensors, the information gain is very small, compared to the redundancy increment that is very high.
- From the Pareto front: it's not worthy considering a number of sensors higher than 10. Over this number the redundancy is much higher without any appreciable increase of information or precision.
- The optimal number of sensors is 6.
  It represents a good trade-off in fulfilling the three conditions.
- The JH is higher than for less sensors, and although its TotC is higher than for 5 sensors, H is smaller. It assures more precision.

5. Model calibration



- In varying the Manning's coefficient n, the model was calibrated against TELEMAC results (considered as "truth"). From literature (Di Baldassarre et al., 2010), the Manning's coefficient of the floodplain is that only parameter that affects results.
- A constant n value was assigned to the floodplain, it varies between 0.02 and 0.10 1/mcs.
- From the original sets given by the NSGA-II (Deb et al., 2001) that solved the MOOP, where chosen only 1/8 sets with lowest MAE.
- Through a frequency analysis, the most chosen Manning's coefficient value is equal to 0.055 1/mcs.
- This value matches with literature results (Di Baldassarre et al., 2010).

6. Conclusions

- . Using an entropy based approach, it is possible to evaluate the best location and the optimum number of sensors on a flood prone area.
- Restraining the number of sensors from the original number allows to eliminate locations that are not significant for describing flood extent.
- The multiobjective function enables to take into account the redundancy, the information content and the precision of the chosen sensors set.
- Results obtained using the entropy method match with those provided by literature guideline.
- In using the entropy method, for the same number of sensors, many configurations are possible.
  The modeler can choose the one with the lowest MAE
- Using the low resolution SRTM DEM, results are reliable, highlighting the great potential of this freely and globally available DEM.

References:

Alfonso, L., A. Lobbrecht, and R. Price (2010). Optimization of water level monitoring network in polder systems using information theory, Water Resour. Res., 46.

Bates, P.D., De Roo, A.P.J., (2000). A simple-based raster model forr flood inundation simulation. J. of Hydrology, 236 (1-2), 54-77.

Deb, K., et al. (2002), A fast and elitist multiobjective genetic algorithm:NSGA-II, IEEE Trans. Evol. Comput., 6, 182-197.

Di Baldassarre, G., Schumann, G., Bates, P.D., Freer, J.E., Beven, K.J., 2010. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches, Hydrological Sciences Journal, 55(3).