A swarm intelligence-based method for hydrological model calibration through a simulated solution space

Juan F. Farfán[1], Luis Cea[2]

1 Environmental and Water Engineering Group, Department of Civil Engineering, University of A Coruña, A Coruña, Spain (j.farfan@udc.es),

2 Environmental and Water Engineering Group, Department of Civil Engineering, University of A Coruña, A Coruña, Spain (luis.cea@udc.es)

UNIVERSIDADE DA CORUÑA

Water and Environmental Engineering Group

Problem

Hydrological models are widely used for flood forecasting, continuous streamflow simulation and water resources management. The success of a hydrological model depends on different factors such as its formulation, data availability and parameter optimization. There are many approaches to identify the optimal parameter sets, which can be categorized in 1) Local search methods and 2) Global search methods. In the group of global search methods, swarm intelligence could provide an alternative to improve the application of surrogate models and to provide robust calibration.

Basic idea

Quasi-random sampling of parameters with the aim of mapping the feasible solution space. Montecarlo simulation for SN parameter sets, and goodness-of-fit coefficients calculation: Nash-Sutcliffe Efficiency (NSE), adapted for peaks Nash-Sutcliffe Efficiency (ANSE), Kling Gupta Efficiency (KGE), and adapted for peaks Kling Gupta Efficiency (AKGE). Configuration of the SN parameter sets and its goodness-of-fit coefficients as training set of a **Surrogate Model based on Artificial Neural Networks (ANN-SM)** [3] in order to generate a **simulated solution space**. Adaptation of a **swarm intelligence-based approach** in order to search in the simulated space. For this study, **Artificial Bee Colony algorithm (ABC)** [2] is adapted. The applied hydrological model is the **Modelo Idrológico Lumped in Continuo (MILC)** [1].

MILC model

Model parameters: Initial water content of the soil layer (W_0), maximum water capacity of the soil layer (W_{max}), exponent of drainage (Computed from the pore size distribution index ($m2 = 3 + 2/\lambda$)) ($m2$), saturated hydraulic conductivity (k_s), baseflow to drainage ratio (α), lag-area parameter (ν), Correction factor for evapotranspiration (b), initial abstraction coefficient ($\lambda 1$), coefficient to compute soil max retention from soil water content (Sr).

References

[1] L Brocca, S Liersch, F Melone, T Moramarco, and M Volk. Application of a model-based rainfall-runoff database as efficient tool for flood risk management. *Hydrology and Earth System Sciences*, 17(8):3159, 2013.

[2] Dervis Karaboga and Bahriye Basturk. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. *Journal of global optimization*, 39(3):459–471, 2007.

[3] Saman Razavi, Bryan A Tolson, and Donald H Burn. Review of surrogate modeling in water resources. *Water Resources Research*, 48(7), 2012.

Acknowledgements

This study is financed by the Galician government (*Xunta de Galicia*) as part of its pre-doctoral aid programme (*Axudas de apoio á etapa predoutoral 2019*) Register Nº ED481A-2019/014. The meteorological data was obtained from the agency MeteoGalicia. The flow data has been provided by the regional water administration *Augas de Galicia*.

ABC algorithm and adaptation

Pseudocode of Artificial Bee Colony Algorithm

Initialize the set of food sources:
x_i for $i = 1, 2, ...SN$ by
$x_i = low_d + rand[0,1](up_d - low_d)$
Calculate goodness-of-fit for:
$fit(x_i)$, for $i = 1, 2, ...SN$
while Stop criterion is not reached **do**
  for $i = 1, 2, ..., SN$ **do**
    $v_i = x_{ij} + rand[-1,1](x_{ij} - x_{kj})$,
    $k \neq j$
    if $fit(v_i) >= fit(x_i)$ **then**
      $x_i = v_i$
    end if
  end for
  for $i = 1, 2, ..., SN$ **do**
    Select an employed bee:
    $p_i = \frac{fit(x_i)}{\sum_{j=1}^{SN} fit(x_i)}$
    Repeat evaluation
    Generate a new random food source (if required).
  end for
end while

$fit()$ denotes goodness-of-fit, low is the lower limit of a parameter, up is the upper limit of a parameter, d denotes the $d-th$ parameter of the model of D parameters.

Pseudocode of Artificial Bee Colony Algorithm for simulated solution space

Initialize the set of food sources
x_i for $(i = 1, 2, ...SN)$
Use Quasi-random Sampling
Calculate goodness-of-fit for each
$f(x_i), i = 1, 2, ...SN$
Train the ANN-SM model
while Stop criterion is not reached **do**
  for $i = 1, 2, ..., SN$ **do**
    $v_i = x_{ij} + rand[-1,1](x_{ij} - x_{kj})$,
    $k \neq j$
  end for
  for $i = 1, 2, ..., SN$ **do**
    Estimate $f(v_i)$, for $(i = 1, 2, ...SN)$ **through ANN-SM**
  end for
  for $i = 1, 2, ..., SN$ **do**
    if $f(v_i) >= f(x_i)$ **then**
      $x_i = v_i$
    end if
  end for
  Generate a new random food source (if required)
end while
Selection of possible food sources by means of a **threshold criterion**
Evaluation **through the hydrological model**

Preliminary results

Figure 1 shows the parameter distribution sampled and the one obtained after 3 search cycles for those that provide goodness-of-fit values greater than 0.65 for NSE and ANSE and 0.70 for KGE and AKGE. The ANN-SM captures information about the sensitivity of the parameters and the ABC algorithm is able to use it in order to converge towards the position of good food sources. The computational burden is minimized, reducing search procedures that could take 3-10 days to 60-200 seconds. For this case study, the ANN-SM provided a 96-99 % success rate in identifying the position of food sources over the threshold criterion

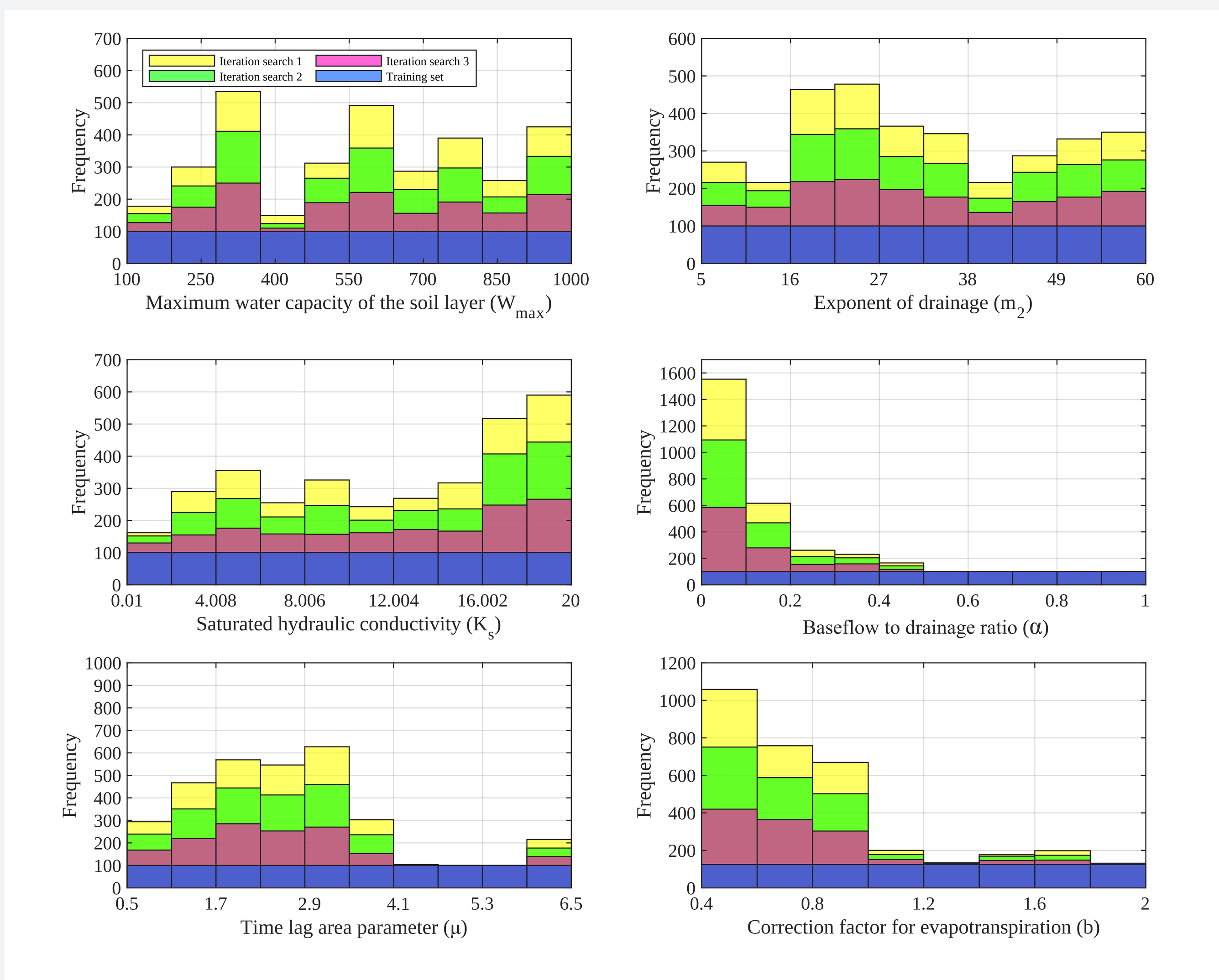

Fig 1: Frequency distribution of the parameter sets over thethreshold

The method has provided a large number of suitable parameter sets. The validation and prediction result is shown for one of these sets. The validation step shows NSE of 0.81, ANSE of 0.843, KGE of 0.77 and AKGE of 0.74. The parameter set overperformed local search method (in this case), especially in validation and prediction stages. Specifically, in the prediction stage, NSE of 0.77 and ANSE of 0.83 were obtained against NSE of 0.45 and ANSE of 0.57 for the local search parameter set.

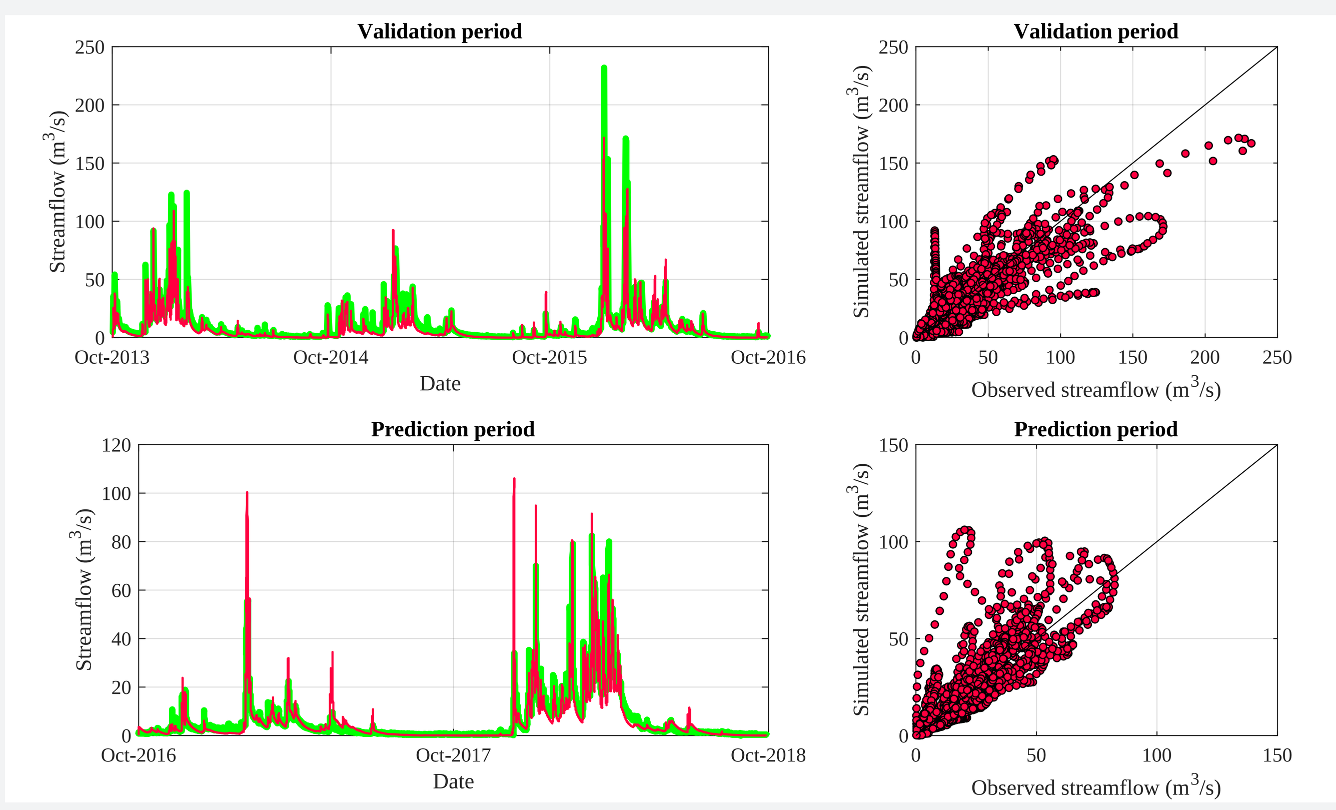

Fig 2: Validation and prediction results for a ANN-SM parameter set

Conclusion: The adaptation approach has been able to provide good preliminary results; however, this methodology should be evaluated in more computationally intensive models, where the advantages and limitations of combining ANN-SM and ABC algorithm (among others) can be explored.