ON THE CALIBRATION OF MULTIGENE GENETIC PROGRAMMING TO SIMULATE LOW FLOWS IN THE MOSELLE RIVER

Ali DANANDEH MEHR *
Mehmet C. DEMIREL **

Abstract: The aim of this paper is to calibrate a data-driven model to simulate Moselle River flows and compare the performance with three different hydrologic models from a previous study. For consistency a similar set up and error metric are used to evaluate the model results. Precipitation, potential evapotranspiration and streamflow from previous day have been used as inputs. Based on the calibration and validation results, the proposed multigene genetic programming model is the best performing model among four models. The timing and the magnitude of extreme low flow events could be captured even when we use root mean squared error as the objective function for model calibration. Although the model is developed and calibrated for Moselle River flows, the multigene genetic algorithm offers a great opportunity for hydrologic prediction and forecast problems in the river basins with scarce data issues.

Keywords: Low flows, calibration, genetic programming, ANN, HBV and GR4J

Moselle Nehri’ndeki Düşük Debilerin Benzetimi için Çoklu Genetik Programlama Modelinin Kalibre Edilmesi


Anahtar Kelimeler: Düşük debi, model ayarlanma, çoklu genetik programlama, ANN, HBV ve GR4J

* Marine Engineering Department, Istanbul Technical University–North Cyprus Education and Research Campuses, Famagusta, Turkish Republic of North Cyprus
** Geological Survey of Denmark and Greenland, Øster Voldgade 10, 1350 Copenhagen, Denmark
Correspondence Author: Ali Danandeh Mehr (danandehmehr@yahoo.com)
1. INTRODUCTION

High and low flows are two important phases of streamflow. Since high flows, e.g. floods are direct threat to human life, the topic has been extensively studied (Vormoor et al., 2015; Zhang et al., 2015). Low flows and droughts are also important in water management especially after obviously intensified climate change impacts in different part of the world (Demirel et al., 2013a; Griffin and Anchukaitis, 2014; Hesami et al., 2016; Pal et al., 2015). Different definitions of low flows exist based on the occurrence period such as winter low flows and summer low flows (Smakhtin, 2001). To anticipate these events, first, the basin behavior has to be reasonably modelled using hydrologic or data-driven models. Then, forecasted precipitation (P), potential evapotranspiration (PET) as well as other inputs can be provided to the calibrated model for issuing streamflow forecasts. Similarly for assessing climate change impacts on streamflow, a calibrated model is necessary to execute long-term simulations. In this study, we are interested in calibrating a data-driven low flow model based on multigene genetic programming and comparing its performance with existing three models in Demirel et al. (Demirel et al., 2015).

Hydrologic models are usually classified into three categories: physically distributed conceptual and data-driven models. Distributed models are complex models with numerous parameters and physically based formulations of hydrologic process valid for every grid cell or hydrological unit in a river basin. They are often used for soil and vegetation parameterization, and assessing land use and land cover change on river flows (Livneh et al., 2015; Samaniego et al., 2010). Conceptual models are lumped hydrologic models assuming model parameters and inputs are uniformly distributed over the basin. This includes precipitation and potential evapotranspiration which are the only inputs required for the water balance of a basin. They are the most commonly used hydrologic models as they are computationally efficient for different assessments e.g. Monte Carlo simulations and climate change impacts (Demirel et al., 2013b; Tian et al., 2014). Data-driven and statistical methods such as artificial neural network (ANN), autoregressive moving average models (ARMA), fuzzy logic, and genetic programing (GP) have been extensively used as alternative to physical and conceptual hydrologic models (Demirel et al., 2015, 2009; Nourani et al., 2009).

All of these models are calibrated with different local and global methods. A well-known example for local methods is Gauss–Marquard–Levenberg gradient search method (Arsenault et al., 2014; Madsen, 2000), whereas there are different global search methods like Shuffle Complex Evolution Algorithm (Duan et al., 1993), Covariance Matrix Adaptation Evolution Strategy (Hansen and Ostermeier, 2001) and Ant algorithm (Madadgar and Afshar, 2009). Both local and global methods require objective functions that are selected based on the modeler’s objective and target problem. For example bias, correlation coefficient ($R^2$), Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe (1970)) and Kling-Gupta Efficiency (KGE, Gupta et al., (2009)) all focus on water balance and prone to the peak flows. Logarithmic transform, inverse values of the streamflow observations can be used to reduce the effect of peak flows on the model performance and highlight the model performance on low and mean flows. Pushpalatha et al (2012) reported suitable objective functions and performance metrics to evaluate low flows. Recently, a neural network model and two well-known conceptual models (GR4J and HBV) are used to simulate low flows in the Moselle River (Demirel et al., 2015). GR4J model has been used in other low flow studies too (Nicolle et al., 2014; Pushpalatha et al., 2011). However, to our knowledge, the new data-driven method, i.e. multigene genetic programming, has not been calibrated and used for low flows.

In this study, the main objective is to benchmark the performance of multigene genetic programming with three existing models (ANN, GR4J and HBV) for low flows in the Moselle
River. We use a hybrid metric comprised of mean absolute error calculated based on low flows and inverse values of hydrograph for the calibration period (Demirel et al., 2015).

The outline of the paper is as follows. The Moselle River basin and model input data are presented in Section 2. Section 3 describes methods used in this study. The results are presented and discussed in Section 4 and the conclusions are drawn in Section 0.

2. STUDY AREA AND DATA

2.1 Study Area

The models are tested in the Moselle River basin which has an area of 27262 km² and river length of 545 km (Figure 1). This basin is the largest part of the River Rhine basin with varying landforms from 59 to 1326 m (Demirel et al., 2015). The minimum streamflow observed in the basin outlet (Cochem) is 14 m³/s in dry summers and maximum streamflow is ~4000 m³/s in winter period. The river flow is regulated by different dams and weirs to allow river navigation along the year.

![Figure 1: The River Rhine and Moselle River basin.](image)

2.2 Model Input

Daily streamflow (Q) data observed at Cochem Station from 1951 to 2006 are provided by the Global Runoff Data Centre (GRDC), Koblenz. The daily observed P and PET, available for the same period, are provided by the German Federal Institute of Hydrology (BfG) in Koblenz, Germany. The statistical parameters of observational Q, P, and PET time series for the period 1951-2006 used in this study is presented in Table 1.
Table 1. Descriptive statistics of the modelling variables

<table>
<thead>
<tr>
<th>Parameters</th>
<th>P (mm)</th>
<th>PET (mm)</th>
<th>Discharge (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>20454</td>
<td>20454</td>
<td>20454</td>
</tr>
<tr>
<td>Maximum</td>
<td>39.6</td>
<td>5.41</td>
<td>4020</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0.14</td>
<td>10</td>
</tr>
<tr>
<td>Average</td>
<td>2.45</td>
<td>1.57</td>
<td>323.29</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.91</td>
<td>1.18</td>
<td>354.17</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.75</td>
<td>0.89</td>
<td>3.11</td>
</tr>
</tbody>
</table>

3. METHODS

In this study four hydrologic models are compared based on their performance during calibration and evaluation periods. Demirel et al. (2015) presented a case study adapting model parameters and all required calibration runs by the three models namely ANN, HBV and GR4J for the same study area (Table 2). Here, we present only the additional model i.e. multigene genetic programming. For calibration and validation of this model, we use the same periods (i.e., 1971-2001 for calibration and 1951-1970 for validation) as three previous models to fairly compare four models. It must be noted that the entire data are normalized by the well-known Min-Max method before using them as modelling variables. Low flows occur mostly after a long dry period. If the dry period is in summer then it is called summer low flows and if low flow occurs because precipitation forms snow then it is called winter low flows (Smakhtin, 2001). Demirel et al. (2015) selected exceedance probability of 75% (Q75) as a threshold for the low flows. In a similar way, if exceedance probability of daily flow at Moselle River is more than 75%, it is considered as low flow in this study.

Table 2. Model details from Demirel et al. (2015)

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of buckets</th>
<th>Number of parameters</th>
<th>Inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR4J</td>
<td>2</td>
<td>4</td>
<td>P and PET</td>
</tr>
<tr>
<td>HBV</td>
<td>3</td>
<td>8</td>
<td>P and PET</td>
</tr>
<tr>
<td>ANN</td>
<td>1</td>
<td>6</td>
<td>P, PET and Q(t-1)</td>
</tr>
</tbody>
</table>

3.1 Multigene Genetic Programming

The state of the art Genetic Programming (Koza, 1992) is another data-driven method that evolves computer programs to automatically solve problems using Darwinian natural selection. In hydrological applications, it is commonly used to infer the underlying structure of either natural (Danandeh Mehr et al., 2013; Ghorbani et al., 2010; Guven, 2009; Sattar and Gharabaghi, 2015) for experimental processes (Khan et al., 2012; Roushangar et al., 2014; Uyumaz et al., 2014). In such applications, GP generates some possible solutions to identify the process numerically. Multigene
genetic programming (Searson, 2009) is one of the most recent advancement of GP that linearly combines low depth GP trees in order to improve fitness of traditional GP approach. Owing to the using of smaller trees, the multigene genetic programming (MGGP) is expected to provide simpler models than those of traditional monolithic GP. In MGGP, predictand variables are computed by the weighted output of each gene (i.e. trees) in the multigene program plus a bias term. The MGGP employed in this study is based on evolving low-depth-low-number GP trees (i.e., Max tree depth=4 and Max genes=4) without any implicit or explicit reference to any of the other genes in the same chromosome. The standard GP sub-tree crossover, direct reproduction and mutation transformation (Poli, R., Langdon, W.B., McPhee, 2008) as well as common mathematical functions (i.e., addition, subtraction, multiplication, square, sin, cos, and exp) are used to create the best model. In addition to the standard GP crossover operator, we also used a tree crossover operator called “two point high-level crossovers” to exchange of genes between chromosomes. The bias and weights (i.e. regression coefficients) are determined by a least squares procedure for each multigene chromosome. More details on the MGGP can be obtained from Searson et al. (2010) and Gandomi and Alavi (2012).

3.2 Performance evaluation criteria

In this study we used the hybrid metric comprised of mean absolute error calculated based on low flows and inverse values of hydrograph for the calibration and validation periods, shown at Eq. (1), as the results should be comparable to Demirel et al. (Demirel et al., 2015). In addition root mean squared error (RMSE) is used as the objective function to calibrate the model.

\[
MAE_{hybrid} = \frac{1}{m} \sum_{i=1}^{m} |Q_{sim}(i) - Q_{obs}(i)| + \frac{1}{n} \sum_{j=1}^{n} \frac{1}{Q_{sim}(j)} + \frac{1}{Q_{obs}(j) + \epsilon}
\]

(1)

Where \(Q_{obs}\) and \(Q_{sim}\) are the observed and simulated streamflows for the \(i^{th}\) observed low flow days (i.e. \(Q_{obs}<Q_{75}\)) and \(m\) is the number of low flow days in total. In the second part of the equation \(j\) is the day index; \(n\) is the total number of days and \(\epsilon\) is the 1% of the mean observed streamflow to avoid infinity during zero values.

4. RESULTS AND DISCUSSION

In order to develop a MGGP-based model for low flow prediction in the Moselle River, an open-source software platform for symbolic data mining in MATLAB®, namely GPTIPS 2 (Searson, 2015) is used in this study. Figure 2 shows the summary of runs configured to minimize the error metric (RMSE) over the calibration data. The upper part of the figure shows the log10 value of the best RMSE achieved in the population over the generations of a run and the lower part depicts the mean RMSE achieved in the population.
Figure 2: The best and mean RMSE achieved in each generation.

Figure 2 illustrates low quality of population at the beginning. Good quality results begin to dominate in the population quickly, along with the increasing number of generations. The method converges to the good quality of results very fast. Only seven generations are enough to obtain good mean RMSE. Further generation improves the result very slight, but the best RMSE occurs after 110 generation.

The multigene expression of the best MGGP model developed for one-day in advance discharge prediction in Moselle River is shown in Figure 3. The parameters $x_1$, $x_2$, and $x_3$ given in the terminal nodes of the genes represent the P and PET at time t as well as Q at time t-1, respectively. The simplified mathematical expression of the model for normalized values of data including its constituent bias and the genes weights is also presented in Eq. (2).

$$
y = 0.8857 \sin(x_3) + 0.4039 x_3 \sin(x_1)^{1/2} - 0.2389 x_2 x_3 (x_1 + x_2)^2
- 74.67 x_1^4 x_2 x_3^7 \sin(x_1)^2 + 0.004167
$$

(2)
Figure 3:
Multigene expression of the best MGGP model developed for one-day lead time discharge prediction in Moslile River

Figure 4 shows the Q-Q plots for different periods. While the MGGP models captures low flows and high flows well, some of the mean flows (between 1500 and 2500 cms) are underestimated. The picture is slightly improved for validation for flows below 1000 cms. However, the model usually overestimates the very high flows. We evaluated the model also for a test period between 2002 and 2006. The performance of the model outside calibration and validation period is promising. For a forecast model, this can be the perfect forecast using observations. In addition, European Centre for Medium-Range Weather Forecasts (ECMWF) such as P and PET can be included to assess the forecast performance of the model. However this is beyond the scope of this current study.

Figure 4:
Scatter plot of simulated and observed streamflow using the MGGP model.
Figure 5 shows only the low flow hydrographs for the selected years in both calibration and validation periods. The MGGP model can simulate extreme low flows and miss very few events (empty blue circles) close to mean values showing the skill of model for predicting extreme events.

![Hydrograph of simulated and observed low flows using the MGGP model.](image)

Efficiency values of the best evolved MGGP model are summarized in Table 3 and compared with those of ANN, HBV and GR4J models. Based on the results, it is obvious that MGGP model performs better than all three models previously used to simulate low flows in the Moselle River basin. The significant difference between MGGP and ANN models can be explained by the number of hidden neurons in ANN as both models incorporate the streamflow values from previous day to mimic storages/states and memory in the typical conceptual models e.g. GR4J and HBV. When the streamflow from previous day is excluded from the inputs, the model performance is significantly hampered.

<table>
<thead>
<tr>
<th>Model</th>
<th>Calibration Period (mm)</th>
<th>Validation Period (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGGP</td>
<td>0.049</td>
<td>0.045</td>
</tr>
<tr>
<td>GR4J</td>
<td>0.550</td>
<td>0.720</td>
</tr>
<tr>
<td>HBV</td>
<td>0.550</td>
<td>0.580</td>
</tr>
<tr>
<td>ANN</td>
<td>0.850</td>
<td>0.980</td>
</tr>
</tbody>
</table>

The performance of the model would obviously be better when low flow oriented objective function (Eq. (1)) is used. However, such function may lead the MGGP to over-train for low flow prediction so that the best solution would not be applicable for entire time series. In addition, our
study showed that the MGGP model could perform better than other three models even it considers the performance over the entire hydrograph. Also the number of peaks and high flows apparently didn’t dominate the RMSE performance.

5. CONCLUSIONS

The performance of a new generation data-driven model on simulating low flows is compared with three different hydrologic models. Similar set up for all models are used to do a fair comparison. Precipitation, potential evapotranspiration and streamflow from previous day have been used as input to both data-driven models i.e. ANN and MGGP, whereas only basin averaged precipitation and potential evapotranspiration are required for GR4J and HBV models. Based on the calibration and validation results significant improvement in low flow simulation performance is achieved using the MGGP. Obviously this model is the best performing model from the four models. The timing and the magnitude of most of the low flow events could be predicted. The proposed model is successful especially for the extreme low flow events. Although the model is developed and calibrated for Moselle River flows, the multigene genetic algorithm offers a great opportunity for hydrologic prediction and forecast problems in the river basins with scarce data issues.

ACKNOWLEDGEMENTS

We acknowledge the financial support for the SPACE project (http://www.space.geus.dk/) by the Villum Foundation (http://villumfonden.dk/) through their Young Investigator Programme and the Scientific (grant VKR023443) and and the Scientific the financial support of the Ir. Cornelis Lely Stichting (CLS), Project No. 20957310.

REFERENCES


36. Vormoor, K., Lawrence, D., Heistermann, M., Bronstert, A., 2015. Climate change impacts on the seasonality and generation processes of floods – projections and uncertainties for
Demirel C.M., Mehr D.A.: On the Calibration of Multigene Genetic Prog. to Sim. Low Flows in the Moselle River
