

# Reverse flood routing in an open channel using genetic algorithm

Ali. Azizipour<sup>1</sup>, Seyed Mahmood Kashefipour<sup>2</sup>, Ali Haghighi<sup>3</sup>

1- M.SC Graduate Student of Civil Engineering, Shahid Chamran University of Ahvaz, Iran (aliazizi583@yahoo.com)

2- Professor of Water and Environmental Engineering Faculty, Shahid Chamran University of Ahvaz, Iran (kashefipour@scu.ac.ir)

3- Professor of Engineering Faculty, Shahid Chamran University of Ahvaz, Iran (a.haghighi@scu.ac.ir)

## ABSTRACT

Flood routing in flood forecasting issue, calculation the height of flood bands, determining the river boundaries, and estimation of protective facilities for flood –exposed building is applicable. In many cases, due to the lack of measuring stations, the status of the upstream flood generating hydrograph is not known. The purpose of this study is to present an integrated method comprising of an optimization model and a hydrodynamic numerical model for flood modeling to determine the upstream hydrograph using the provided hydrograph at the downstream measuring station of a river. The routing procedure consists of three steps: (1) generating a hypothetical upstream hydrograph using genetic algorithm method; (2) hydrodynamic modeling using a numerical simulation model for flood routing according to the hypothetical hydrograph which is generated in the first step; (3) compare the calculated and observed hydrograph in downstream by using a fitness function. This recommended procedure was named as Reverse Flood Routing Method (RFRM) and was then applied to Karun River, the largest river in Iran. Comparing the generated upstream hydrograph by the RFRM model with the corresponding measured hydrograph at Ahvaz hydrometric station, as an ungauged river location, shows the high accuracy of the recommended model in this study.

**Keywords:** generating missed hydrograph, genetic algorithm, reverse flood routing, Karun river, numerical FASTER model

## 1. INTRODUCTION

Since the earliest recorded civilization, humans has tended to settle near rivers because of proximity to water supplies, natural needs for urban and rural populations, and advantages for agriculture activities, resulting for living a significant percentage of people around the world near rivers (Baldassarre et al., 2013). But living along rivers has dangers, like a flood, which is a natural phenomenon and its occurrence is inevitable in many places and times and sometimes no humanmade structures can stop it. However, there are some methods for mitigating the risk of its happening and consequences. Many researchers has studied the causes and presented strategies for reducing the destructive effects of the flood. Floods control and management and also the design of the suitable flood-resistance structures and specifying flood catchment areas on both sides of a river is dependent on the flood hydrograph, its peak discharge, and base flow time. Flow discharge is obtained from velocity, depth, and width of flow in the river, so river, and some cross-sections

may be prepared for flow depth measurement (Zucco et al., 2015). These equipped cross-sections are generally called hydrometric stations. In these types of stations, water elevations are continuously recorded and are converted to discharges using a pre-prepared accurate rating curve or stage-discharge relationship. Therefore, the flood hydrographs are provided at these stations (Perumal et al., 2007; D'Oria and Tanda, 2012). Construction, equipment, and maintenance of hydrometric stations require high costs for governments or water agencies, especially for natural waterways with complex and challenging conditions. For this reason, the number of such stations along a river or canal may be limited and is mostly dependent on the importance of the river area and also on the current and future development projects. Therefore, providing hydrograph at any point of a waterway or river needs to set up a hydrodynamic numerical model and run it for the desired reach of channel and time period. These types of models need all information like stage-discharge at upstream and downstream boundaries of the chosen distance during the selected period.

To mitigate the flood risk, the knowledge of discharge at gauging stations and upstream or downstream of the section is essential (Zucco et al., 2015). Accurate information about flood characteristics is one of the critical issues for managing its disaster (Soleimani-Alyar et al., 2016). There are several hydraulic and hydrologic mathematical methods used for flood modeling in rivers and waterways (Pashazadeh and Javan, 2020). A hydraulic model consists of the conservation and energy or momentum equations that use geometric and hydraulic data for flood routing. But a hydrologic model relies on limited parameters coupled with a linear or nonlinear equations based on the continuity equation for flood routing, which is simple in comparison to a hydraulic model (Bozorg Hadad et al., 2018). Flow hydrographs are one of the fundamental elements for flood modeling which is recorded as time series at hydrometric or gauged stations also, the information of discharge hydrograph is essential for flood-risk assessment and management of water resource systems (Todaro et al., 2019). However, due to the lack of hydrometric and, or gauged stations, these hydrographs may not be available everywhere (Kaya et al., 2017). The Muskingum method is one of the hydrological approaches that been used for many years due to its simplicity and acceptable accuracy (Norouzi and Bazargan, 2020). The usual downstream routing of flood flow, i.e., using the existing upstream hydrograph, is a stable process for the Froude numbers less than one and, the results are reliable. In contrast, the reverse flood routing determines the upstream hydrograph based on the river geometry, flow hydraulic characteristics, and downstream hydrograph (Bruen and Dooge, 2007). A part of performing reverse routing is to solve the 1-D Saint-Venant equations, solving these equations requires initial and boundary conditions, and discharge hydrograph at the end of the river reach (Spada et al., 2017).

In this study, a novel procedure is deployed to develop an integrated model comprising a hydrodynamic mathematical model, which solves the 1-D Saint-Venant equations, and an optimization model accompanied by a genetic algorithm method. This combined program can generate the upstream missed hydrograph using the existing downstream hydrograph.

## THEORY AND PREVIOUS STUDIES

### Governing 1D Hydrodynamic Equations

The basic formulation of unsteady one-dimensional flow in open channels defined as St.venant equations. These equations consist of continuity and momentum equation, which written in the following form, respectively (Cunge et al., 1980; Mohamadi and Kashefipour, 2014, 2020):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q_L \quad (1)$$

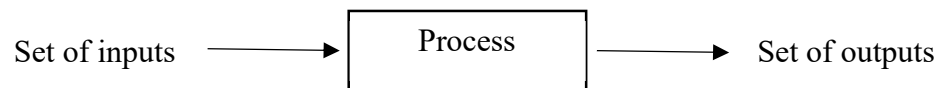
$$\frac{\partial Q}{\partial t} + \frac{2\beta Q q_L}{A} - \frac{2\beta Q T_W}{A} - \beta \frac{Q^2}{A^2} \frac{\partial A}{\partial x} = -gA \frac{\partial Z}{\partial x} - \frac{n^2 Q |Q|}{AR^{\frac{4}{3}}} \quad (2)$$

Where  $Q$  = discharge,  $A$  = cross section area,  $x$  = distance in flow direction,  $g$  = gravitational acceleration,  $Z$  = water surface elevation,  $R$  = hydraulic radius,  $n$  = Manning roughness coefficient,  $q_L$  = lateral inflow or out flow (negative for out flow),  $T_W$  = top width of channel,  $t$  = time and  $\beta$  = momentum correction coefficient.

These partial differential equations are solved using different types of numerical methods. For example, in FASTER (Flow And Solute Transport in Estuaries and Rivers) model, these equations are numerically solved using an implicit staggered central finite difference scheme with variable grid size, which is accurate and unconditionally stable. More details regarding this type of numerical solution of the Saint-Venant equations can be found in Kashefipour (2002). In the MIKE11 commercial model, which was developed and released by DHI (Danish Hydraulic Institute), the Saint-Venant equations are solved based on an implicit finite difference scheme developed by Abbott and Ionescu (DHI Reference Manual, 2017).

### Introduction to Optimization

To minimize an error, effort or energy, or to maximize profit or outcomes, the optimization process can be used in all sciences such as math, computer, engineering, and so on (Kramer, 2017). In other words, optimization is a technical and iterative procedure for making something better. In any process, there is a set of inputs and a set of outputs, as shown in Figure 1.



**Fig.1.** Optimization process

Optimization refers to finding the values of inputs in such a way to get the “best” output values. The definition of “best” varies from a problem to another, but in mathematical terms, it refers to maximizing or minimizing one or more objective functions, by changing the input parameters. The set of all possible solutions or values, which the inputs can take, make up the search space. In this search space, lies a facts or a set of points that gives the optimal solution. The optimizations aim is to find that point or set of points in the search space.

## ***Genetic Algorithm***

Genetic Algorithms (GA) guide search strategies inspired by a biological process, derived from Darwin's principal of natural selection and survival of fittest (Sivapragasam et al., 2008). This method is frequently used to solve optimization problems, research and in machine learning. The general framework of the most evolutionary algorithm are similar, that means they start with a population of random solution which is evaluated by a fitness function. The most popular components of this algorithm are crossover, mutation, and selection (Mirjalili et al., 2020). Crossover is an operator that allows exchanging genes between parents to produce a new solution. There are three main ways of crossover methods in the literature, single-point crossover, double point crossover and multipoint crossover. For example, in single-point crossover, the chromosome of the two-parent solution change before and after single-point. Crossover is the primary operator of the exploitation of GA. The mutation causes random changes in genes and is the primary mechanism of exploration in GA. There are many mutation techniques such as uniform, non-uniform etc. that can be used in a genetic algorithms, by using a selection operator, the best offspring solutions selected to be parents in the new parental population to allow convergence towards optimal solutions. Selection process is based on the fitness value in a population. When we face a minimize situation, the low value is acceptable but in a maximize situation, the high value is preferred. There are many methods for simulating the selection, Roulette wheel selection is one way. More details about GA can be found in Haupt and Haupt (2004).

## **2. MATERIALS AND METHODS**

Describe the procedure of developing the integrated model in this research is followed using an application for Karun River, which flows in the southwest of Iran.

### ***Case Study Description***

Karun River is the largest and the only navigation river in the southwest of Iran. The basin area of this river is about  $45220 \text{ km}^2$ , with the 50 years average discharge being reported about  $600 \text{ m}^3/\text{s}$ . The selected reach length was about  $50 \text{ km}$  with, 49 cross-sections (Figure 2). This figure shows study area domain; Ahvaz (cross-section 49 or C.S49) and Farsiat (cross-section 1 or C.S1) are the upstream and downstream boundaries, respectively. The measured hydrographs ( $Q-t$  chart) and time-series of water elevations ( $Z-t$  chart) were available at both boundaries for several periods, i.e., the continuous measurements at both hydrometric stations.

In this study, the numerical FASTER model, which was initially developed by Kashefipour (2002), and genetic algorithm was used for simulation the flow in Karun River and optimization, respectively. A piecewise equation was added to the FASTER model enabling it to calculate the Manning roughness coefficient at any spatial point and any time during simulation using flow discharge and depth. This dynamic equation makes the model very accurate for determining time-dependent discharge and depth at any point of the domain. This part of the FASTER model is fully described by Mohamadi and Kashefipour (2014). The FASTER model was set up for the aforementioned domain. All data needed for implementing the model for the years 2016-2017 and also the bed elevations for all cross-sections were provided by KWPA (Khuzestan Water and Power Authorities) company. Ahvaz and Farsiat hydrometric stations were defined as the upstream

and downstream boundaries, respectively (see Figure 2). The  $Q-t$  data is usually defined for the upstream boundary and the  $Z-t$  data for the downstream boundary.

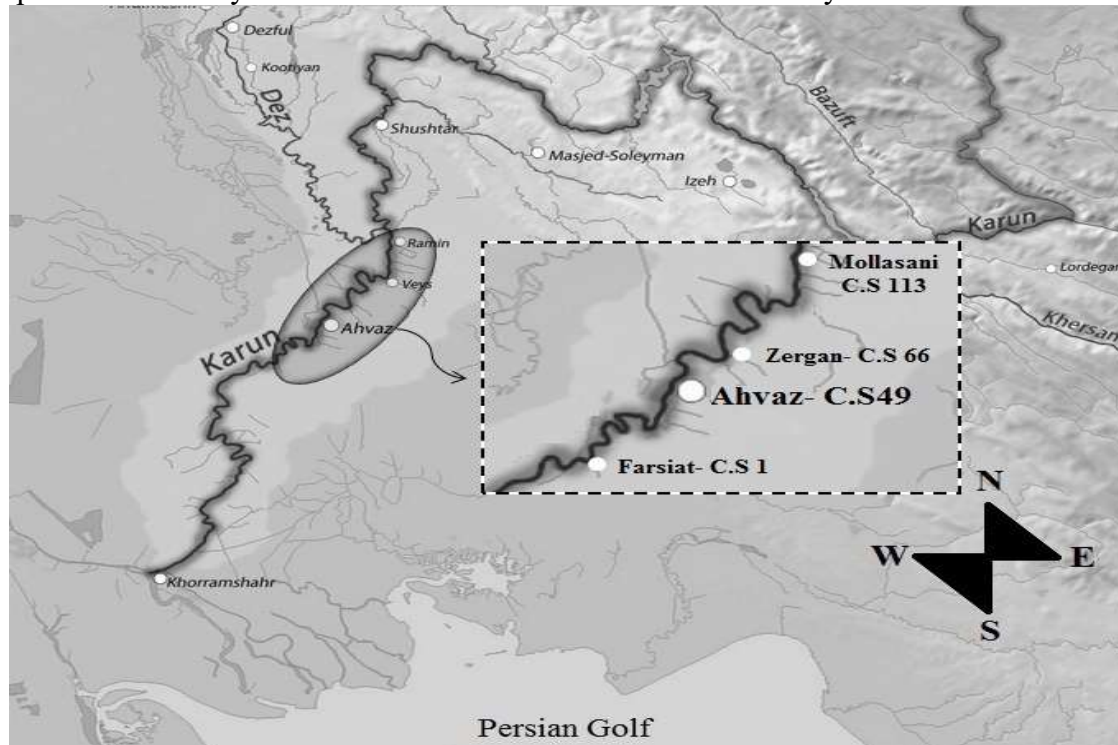


Fig. 2. – The selected reach and hydrometric stations along this reach in Karun River

#### Calibration of FASTER Model

Since the main part of the developed integrated model in this study is the flow simulation model, it was essential to ensure from the FASTER model accuracy in the prediction of discharges and water surface elevations. For this reason, the FASTER model was first calibrated using the existing data. It was found that the dynamic piecewise equations introduced for the Manning roughness coefficient by Mohamadi and Kashefipour (2014) in the FASTER model need to be slightly revised for this domain of Karun River. The original equations were replaced by the revised ones in the FASTER model. So we were assured of the accuracy of the model predictions for the discharges and water surface elevations. These two equations are written as Equations 3 and 4. Comparison of these equations with the original ones in the FASTER model shows that for the discharges more than  $1100 \text{ m}^3/\text{s}$ , the roughness coefficients increase about 9%, and this value was about 27% for the discharges less than  $1100 \text{ m}^3/\text{s}$ .

$$\text{If } Q > 1100 \text{ m}^3/\text{s} \quad n = \frac{a * 0.037}{y_{ave}^{0.124} * u_{ave}^{0.53}} \quad (3)$$

$$\text{If } Q \leq 1100 \text{ m}^3/\text{s} \quad n = \frac{b * 0.037}{y_{ave}^{0.023} * u_{ave}^{0.46}} \quad (4)$$

Where  $y_{ave}$  and  $u_{ave}$  are the cross-sectional average of flow depth and velocity, respectively;  $a=1.09$  and  $b=1.27$ .



### Optimization Method

The optimization model is written based on the genetic algorithm (GA) using the well-known and free access MATLAB software. To write a genetic algorithm first, a function as a fitness function is defined. The fitness function in this study is defined as Equation 5.

$$f = \sum_{i=1}^N (Q_{obsi} - Q_{cali})^2 \quad (5)$$

Where  $Q_{obs}$  and  $Q_{cal}$  are measured and calculated discharge, respectively. Given that we face a minimization situation, the goal is to bring this function closer to zero as close as possible. By changing some of the parameters in the genetic algorithm, at last, seven different optimization models were written. These parameters were the number of members of the population ( $N_{pop}$ ), percentage of crossover ( $P_c$ ), number of mutations ( $N_{mut}$ ) and, number of iteration. These models were produced based on trial and error, previous experiences, and using two types of numbers, Integer (Models 1 to 3) and Real (Models 4 to 7) numbers, Table 1. The task of this part of the main program is to select the new set of discharges for upstream boundary based on the defined models and the previous set of data and check for minimizing Equation 5 by comparison of the new set of calculated discharges for the downstream boundary with the corresponding measured values. The models defined in Table 1, were finally evaluated by the statistical parameter of Root Mean Square Error (RMSE), which is defined as:

$$RMSE\% = 100 \times f / \sum_{i=1}^N Q_{obsi}^2 \quad (6)$$

**Table 1.** Genetic algorithm parameters for Integer and Real models

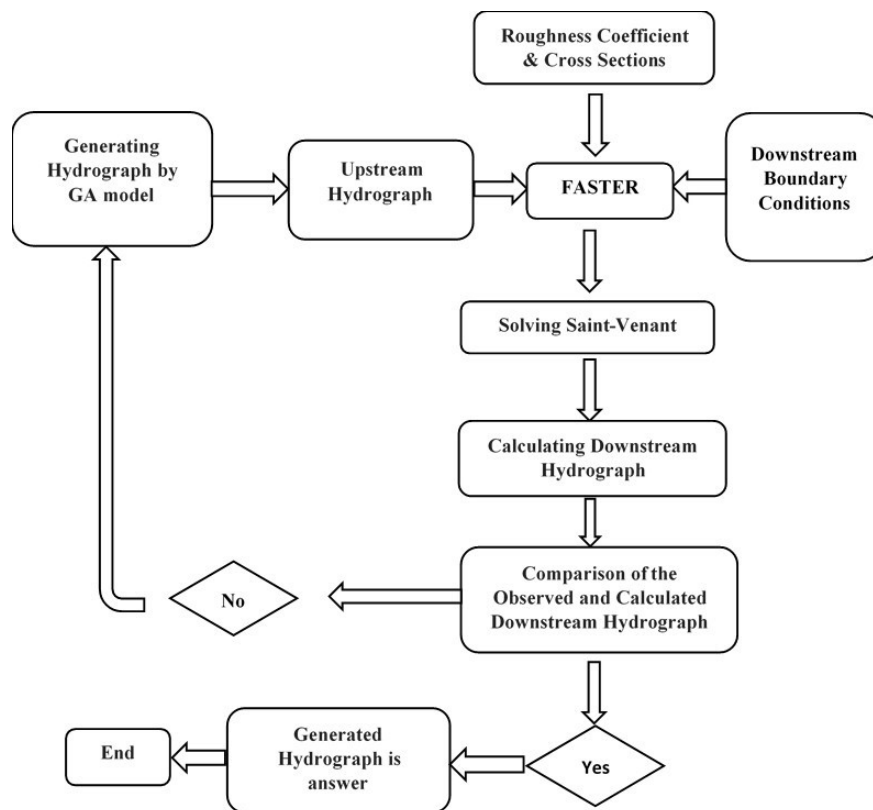
Model	Type	GA Parameters		
		$N_{pop}$	$N_{mut}$	$P_c$
1	Integer	6	2	0.9
2	Integer	10	4	0.9
3	Integer	3	1	0.9
4	Real	6	2	0.9
5	Real	10	4	0.9
6	Real	10	1	0.9
7	Real	3	1	0.9

### Model Development

The main model consists of three parts including: 1-a numerical hydrodynamic model (here FASTER model), which solves the Saint Venant equations and determines water elevations and discharges at any point and time of simulation. 2-a genetic algorithm (GA) model, which produces the new set of discharges (hydrograph) for upstream boundary using the previous set and checking the fitness function (Equation 5) for minimization, and 3-an interface for linking the above models and manage the number of iterations towards the best-optimized results.

Ahvaz and Farsiat hydrometric stations were set as the upstream and downstream boundaries, respectively. The  $Q$ - $t$  data (hydrograph) and  $Z$ - $t$  data for ten days period in 2016-2017 were defined for these boundaries, respectively. It is assumed that the upstream hydrograph is not existed, and the hydrodynamic model cannot be implemented without this hydrograph. Therefore first, a hypothetical hydrograph is produced for upstream boundary using the GA part of the model based

on the downstream hydrograph. For producing the first hydrograph, a range of discharge was defined for the GA model. The hydrograph was randomly chosen between twice of the maximum discharge and 0.8 of the minimum discharge the downstream hydrograph. This range was set for converging the solution faster and receiving the final upstream hydrograph quicker with the minimum necessary iteration. The flow routing is then started along with this domain and continued up to the end of simulation time (i.e., ten days or 240 hours). At the end of the simulation time of the FASTER model, the downstream hydrograph is provided. By the defined fitness function, the calculated and observed hydrographs in the downstream boundary are compared. At last, for the reasonable error, the generated hypothetical hydrograph is the final upstream hydrograph. Otherwise, the GA model produces another hypothetical hydrograph, and this cycle continues until the calculated and observed hydrographs at the downstream boundary had an acceptable error. The interface part of the model transfers the information between the FASTER and GA models, i.e., the upstream hydrograph from GA to FASTER and the downstream hydrograph from FASTER to GA. In this part of the model, a conditional statement is placed to control the number of iterations based on the calculated error. According to the number of iterations, i.e., 10000 and running FASTER and GA models, it is necessary to use a supercomputer to reduce the program run time. A flowchart of the model is shown in Figure 3.



**Figure 3: Flowchart of the model**

### 3. RESULTS AND DISCUSSION

Optimization problems are generally solved using a series of repetitive mathematical calculations and, by defining one or more conditional statements, tries to get the best answers. For example, in this study, in each iteration, a GA model, a numerical hydrodynamic model, and an interface model are implemented, and running these models together with a usual computer takes a very long time, especially when the number of iterations is high. Also, it should be mentioned that according to the simulation time and the amount of the time step, the hydrodynamic model may take a considerable time to be completed in each iteration. Therefore, in this type of computer program, the time consuming of running program is a key parameter. For minimizing the time-consuming running a program, we used a supercomputer with a 64-core processor. However, selecting a suitable combination of the parameters and their values for the GA model (i.e., models 1-7, see Table 1) may significantly reduce time of implementing the main model. In addition to program execution time, the performance and accuracy of the GA model, which produces the upstream hydrograph, is another key parameter.

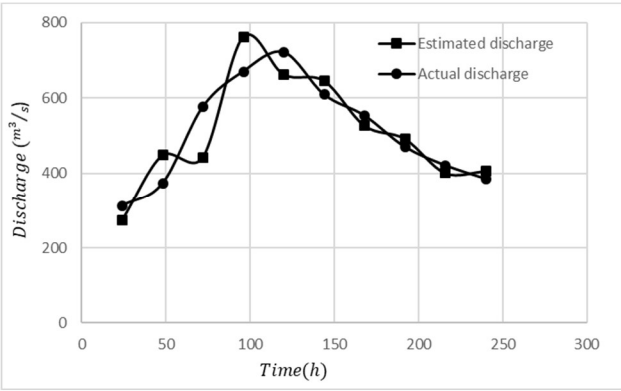
In this study, the number of iteration for all seven considered models in the GA model was the same and equal to 10000. The most critical decision for implementing a genetic algorithm is to choose a good representation, otherwise leads to poor performance and answer. The results showed that the best combination of the GA parameters is model No.6 (see Table 1). Running the program revealed that after 3030 iterations, the RMSE% of comparing the produced and real hydrographs for the downstream boundary was calculated less than 0.005, which indicated the high accuracy of the GA model. Furthermore, the implementation time of the program to reach the best answer is much shorter.

The produced hydrographs for the upstream boundary by the different GA models are illustrated in Table 2, and the real hydrograph is also added to this table. The RMSE% based on Equation 6 is also calculated and added to this table. As it is shown in Table 2, model No.6 performed better than the other considered models. Figures 4a-4g compares the produced or estimated hydrographs by seven GA models with the measured or actual hydrograph at the upper boundary, i.e., Ahvaz hydrometric station in this study.

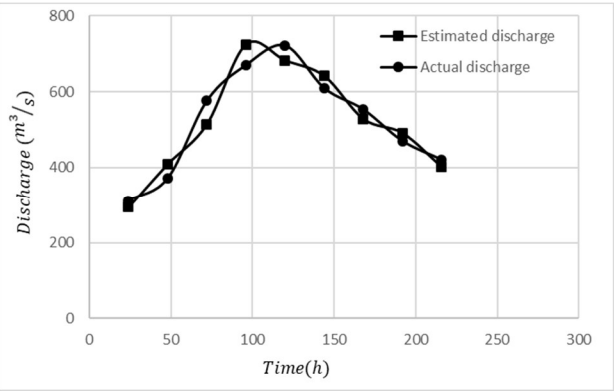
**Table 2.** Calculated and Observed discharges for the upstream boundary (Ahvaz station)

Time (h)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Observed
24	273	294	297	273.66	295.97	304.62	298.83	312
48	448	409	403	448.14	402.64	386.37	397.24	372
72	443	514	526	442.70	525.34	554.89	534.51	576
96	762	724	717	761.78	714.33	687.87	707.26	670
120	663	682	683	665.23	688.85	710.33	691.85	721
144	646	641	640	642.00	630.13	615.23	632.13	609
168	527	528	530	532.25	539.74	550.81	531.51	553
192	492	491	490	483.83	476.43	470.26	489.17	470
216	401	402	402	413.08	419.48	424.33	403.84	421
240	406	402	404	392.11	385.49	381.26	402.54	385
RMSE%	1.45	0.46	0.36	1.39	0.27	0.04	0.24	-

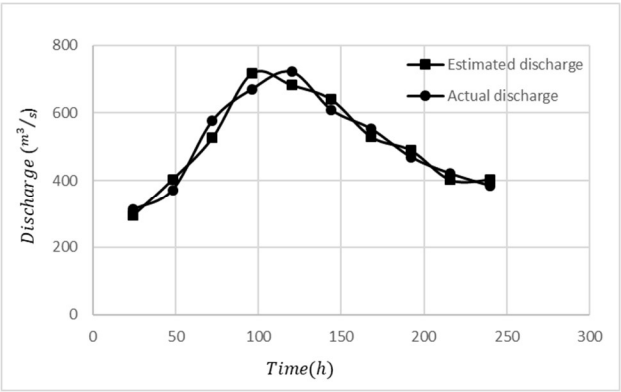




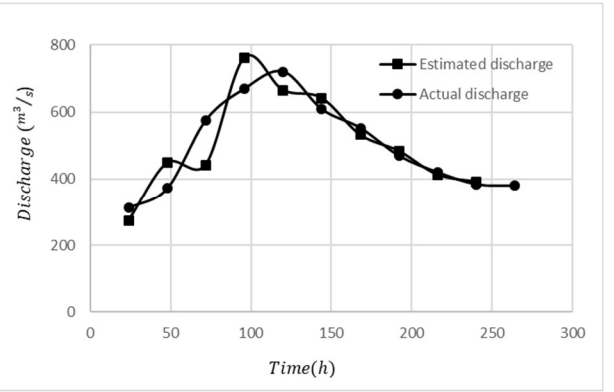
(a) Model 1



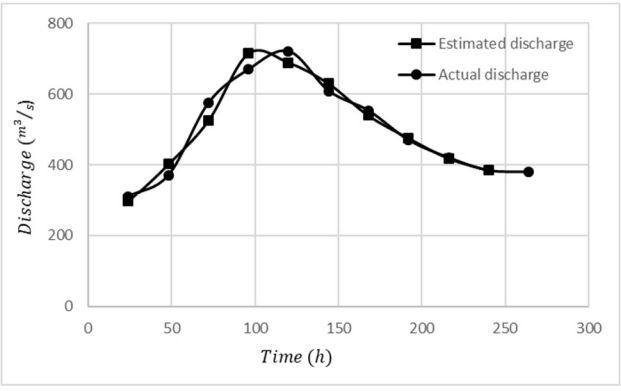
(b) Model 2



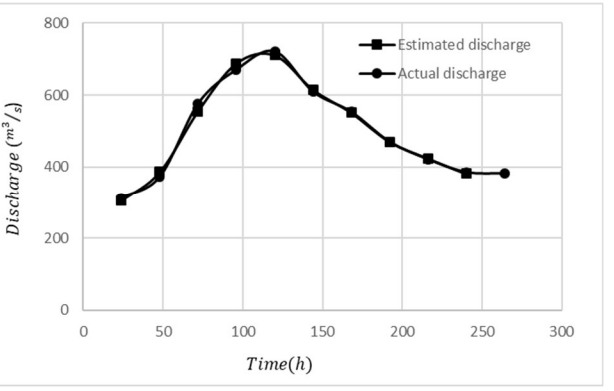
(c) Model 3



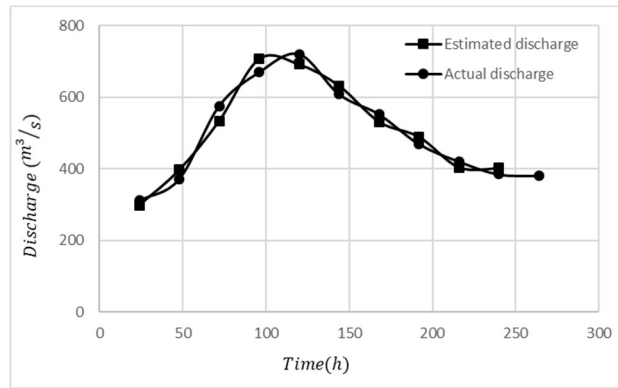
(d) Model 4



(e) Model 5



(f) Model 6



(g) Model 7

**Figure 4: Comparison of the estimated and actual hydrographs for upstream boundary**

Knowing the river conditions in terms of different discharges is essential for many river engineering projects, such as numerical flow and water quality modeling in riverine basins, watershed management, designing hydraulic and water supply structures, and flood control and distribution plans. However, continuous measurement of discharges along a river network may need to establish several hydrometric stations, and this may be impossible due to budget shortage or severe conditions, or high instability of river sections. This research teaches us how to produce the missed or non-measured hydrographs for upstream from downstream or vice-versa.

#### 4. Conclusion

In this study, an integrated model was developed comprising a GA model, a hydrodynamic numerical 1-D model, and an interface program. This mathematical model was written so it can be used for the reverse flow routing to produce an upstream boundary hydrograph from the known downstream hydrograph. It was found that the GA parameters play an important role in not only producing precise hydrographs and best answers, but also reducing consuming computer running time. Here, we could generate the upstream hydrograph with a minimum RMSE error 0.04% for the GA parameters of  $N_{pop}=10$  (number of members of population),  $N_{mut}=1$  (number of mutation), and  $PC=0.9$  (percentage of crossover) with minimum iteration.

## Reference

1. Baldassarre, G.D., Viglione, A., Carr, G., Kuil, L., Salinas, J.L, and Blöschl, G., 2013. Socio-hydrology: conceptualizing human-flood interactions. *Hydrology and Earth System Sciences*, 17(8), pp.3295-3303.
2. Bozorg-Haddad, O., Hamed, F., Fallah-Mehdipour, E. and Loáiciga, H.A., 2018. Upstream flood pattern recognition based on downstream events. *Environmental monitoring and assessment*, 190, pp.1-8.
3. Bruen, M. and Dooge, J.C.I., 2007. Harmonic analysis of the stability of reverse routing in channels. *Hydrology and Earth System Sciences Discussions*, 11(1), pp.559-568.
4. Cunge, J.A., Holly, F.M., and Verwey, A. (1980) *Practical aspects of computational river hydraulics*. Pitman Publishing Limited, London, 420pp.
5. DHI MIKE 11, 2017. *A modeling system for rivers and channels, Reference Manual, Powered by DHI*, 500pp.
6. D'Oria, M. and Tanda, M.G., 2012. Reverse flow routing in open channels: a Bayesian geostatistical approach. *Journal of hydrology*, 460, pp.130-135.
7. Haupt R.L., and Haupt S.E., 2004. *Practical Genetic Algorithms*. John Wiley & Sons, Inc., Publications, 274 pp.
8. Kashefipour, S.M., 2002. *Modeling flow, water quality and sediment transport processes in riverine basins* (Doctoral dissertation, Cardiff University).
9. Kaya, C.M., Tayfur, G. and Gungor, O., 2017. Predicting flood plain inundation for natural channels having no upstream gauged stations. *Journal of Water and Climate Change*.
10. Kramer, O., 2017. *Genetic algorithm essentials* (Vol. 679). Springer.
11. Mirjalili, S., Dong, J.S., Sadiq, A.S. and Faris, H., 2020. Genetic algorithm: Theory, literature review, and application in image reconstruction. In *Nature-Inspired Optimizers* (pp. 69-85). Springer, Cham.
12. Mohammadi, S., Kashefipour, S.M. (2014) Numerical Modeling of Flow in Riverine Basins Using an Improved Dynamic Roughness Coefficient. *Water Resources*, Springer, 41(4): 464-472.
13. Mohamadi, S., and Kashefipour, S.M. 2020, Numerical modeling of FC bacteria using a dynamic and variable mortality rate, *Journal of Water, Sanitation & Hygiene for Development*, 10(3): 569-579.
14. Norouzi, H. and Bazargan, J., 2020. Flood routing by linear Muskingum method using two basic floods data using particle swarm optimization (PSO) algorithm. *Water Supply*.
15. Pashazadeh, A. and Javan, M., 2020. Comparison of the gene expression programming, artificial neural network (ANN), and equivalent Muskingum inflow models in the flood routing of multiple branched rivers. *Theoretical and Applied Climatology*, 139(3-4), pp.1349-1362.

16. Perumal, M., Moramarco, T., Sahoo, B. and Barbetta, S., 2007. A methodology for discharge estimation and rating curve development at ungauged river sites. *Water Resources Research*, 43(2).
17. Sivapragasam, C., Maheswaran, R. and Venkatesh, V., 2008. Genetic programming approach for flood routing in natural channels. *Hydrological Processes: An International Journal*, 22(5), pp.623-628.
18. Soleimani-Alyar, M., Ghaffari-Hadigheh, A. and Sadeghi, F., 2016. Controlling floods by optimization methods. *Water Resources Management*, 30(12), pp.4053-4062
19. Spada, E., Sinagra, M., Tucciarelli, T., Barbetta, S., Moramarco, T. and Corato, G., 2017. Assessment of river flow with significant lateral inflow through reverse routing modeling. *Hydrological Processes*, 31(7), pp.1539-1557.
20. Todaro, V., D'Oria, M., Tanda, M.G. and Gómez-Hernández, J.J., 2019. Ensemble smoother with multiple data assimilation for reverse flow routing. *Computers & Geosciences*, 131, pp.32-40.
21. Zucco, G., Tayfur, G. and Moramarco, T., 2015. Reverse flood routing in natural channels using genetic algorithm. *Water resources management*, 29(12), pp.4241-4267.