USING GENETIC ALGORITHMS TO CALCULATE FLOODWAY STATIONS
WITH HEC-RAS

Thesis
Submitted to
The School of Engineering of the
UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree of
Master of Science in Civil Engineering

By
Lu Yu

UNIVERSITY OF DAYTON
Dayton, Ohio
December, 2015
USING GENETIC ALGORITHMS TO CALCULATE FLOODWAY STATIONS
WITH HEC-RAS

Name: Yu, Lu

APPROVED BY:

Donald V. Chase, Ph.D., P.E.
Advisory Committee Chairman
Department Chairman
Department of Civil & Environmental Engineering & Engineering Mechanics

Denise Taylor, Ph.D., P.E.
Committee Member
Associated Professor
Department of Civil & Environmental Engineering & Engineering Mechanics

Dan Zalewski, Ph.D.
Committee Member
Assistant Professor
Department of Engineering Management & Systems

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering
ABSTRACT

USING GENETIC ALGORITHMS TO CALCULATE FLOODWAY STATIONS WITH HEC-RAS

Name: Yu, Lu
University of Dayton

Advisor: Dr. Donald V. Chase

The paper describes the current problem associated with HEC-RAS that could not provide an accurate 1.00 surcharge for every cross-section of a studied river in encroachment analysis. And also developed a computer model that uses an optimization technique called genetic algorithms to call HEC-RAS to do the encroachment analysis. Genetic algorithm has been applied widely in water resources field since several decades ago, such as in water distribution system design, modeling pipes, pumps, valves and tanks associated with minimum cost issue. This GA program is developed upon the classical process consisting of building the initial population, fitness test, selection, crossover and mutation. In this case, the decision variable of the problem is the percent of the overbank width and could express the encroachment station within the overbanks. And the objective function is to minimizing the difference between the actual surcharge and the target surcharge for every cross-section. An artificial channel with 12
cross-sections is tested for accuracy and responding time, and results have been compared with HEC-RAS method 4 results and manual solution results. Currently, the GA program still could not provide a result better than HEC-RAS provided, however, it is in good progress and is doing well on cutting modeling time and effort. To improve the program, some inspirations about build population, cross over, mutation, recording system, checking system, and bisection method obtained from the manual solution to pursue a more reasonable and closer result.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................................................... iii

LIST OF FIGURES ................................................................................................................................... vii

LIST OF TABLES ........................................................................................................................................ ix

LIST OF ABBREVIATIONS AND NOTATIONS ................................................................................. xi

CHAPTER 1  INTRODUCTION ...........................................................................................................1
  Causes of Flood .......................................................................................................................................................... 2
  Flood Damage .............................................................................................................................................................. 3
  Flood Control ................................................................................................................................................................ 4
  Federal Emergency Management Agency (FEMA) ................................................................................................. 4
  Floodplain and Floodway ........................................................................................................................................... 5
  Floodplain Modeling .................................................................................................................................................. 6
  HEC-RAS and HEC-RAS Modeling Floodplain ................................................................................................. 8

CHAPTER 2  LITERATURE REVIEW ..............................................................................................17
  GA Introduction ........................................................................................................................................................ 17
  History of FEMA Floodplain Management ........................................................................................................ 30
  Brief Review of Genetic Algorithms .................................................................................................................. 32
  Review of GA in Water Resources Field ........................................................................................................... 34
  History of FEMA Floodway .................................................................................................................................. 35
CHAPTER 3  DEVELOPMENT OF THE OPTIMIZATION MODEL FOR FLOODWAY STATION’S CALCULATION

Define Terminologies
Develop GA Processes
Discuss GA

CHAPTER 4  RESULTS AND DISCUSSION OF SIMULATIONS

Parameters Tested
Methods Comparison

CHAPTER 5  OBSERVATIONS

Further Program Improvement

CHAPTER 6  SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

REFERENCES

APPENDIX-1

APPENDIX-2
**LIST OF FIGURES**

4. Bisection method illustration ............................................................................................... 15
5. One-point crossover .............................................................................................................. 24
6. Two-point crossover .............................................................................................................. 25
7. Uniform crossover .................................................................................................................. 25
8. Single point mutation at 11\textsuperscript{th} gene from 0 to 1 ........................................... 26
9. Multi point mutation at 3\textsuperscript{rd} gene from 1 to 0, and 11\textsuperscript{th} gene from 0 to 1 .... 27
10. GA processes ....................................................................................................................... 28
11. Define percent overbank width as decision variable ......................................................... 46
12. One-point crossover example ............................................................................................. 51
13. Single point mutation example ............................................................................................ 52
14. Sample channel illustration ................................................................................................ 60
15. Encroachment result comparison .......................................................................................... 62
16. Example of the bisectional method applied in a cross-section ......................................... 73
17. Example of bisection method solution .................................................................................. 73
18. Example of conveyance calculation ......................................................... 81
19. Encroachment station calculation example .............................................. 86
20. Example of the station points selection of the quadratic curve fit .......... 89
21. Left over bank curve fitting .................................................................... 91
22. Right over bank curve fitting ................................................................. 94
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Selection example</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of using absolute value and square value in fitness function (A)</td>
<td>44</td>
</tr>
<tr>
<td>3</td>
<td>Comparison of using absolute value and square value in fitness function (B)</td>
<td>45</td>
</tr>
<tr>
<td>4</td>
<td>Population size tests</td>
<td>56</td>
</tr>
<tr>
<td>5</td>
<td>Mutation percentage tests</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>Parents’ rate tests</td>
<td>59</td>
</tr>
<tr>
<td>7</td>
<td>Encroachment results comparison</td>
<td>61</td>
</tr>
<tr>
<td>8</td>
<td>Best fitness comparison</td>
<td>63</td>
</tr>
<tr>
<td>9</td>
<td>Directed mutation</td>
<td>70</td>
</tr>
<tr>
<td>10</td>
<td>Geometry data of the cross-section</td>
<td>82</td>
</tr>
<tr>
<td>11</td>
<td>Area calculation</td>
<td>83</td>
</tr>
<tr>
<td>12</td>
<td>Wetted perimeter calculation</td>
<td>84</td>
</tr>
<tr>
<td>13</td>
<td>Hydraulic radius and conveyance calculation</td>
<td>85</td>
</tr>
<tr>
<td>14</td>
<td>Area calculation</td>
<td>87</td>
</tr>
<tr>
<td>15</td>
<td>Wetted perimeter calculation</td>
<td>88</td>
</tr>
<tr>
<td>16</td>
<td>Hydraulic radius and conveyance calculation</td>
<td>88</td>
</tr>
<tr>
<td>17</td>
<td>Left over bank quadratic curve fitting data</td>
<td>90</td>
</tr>
</tbody>
</table>
18. Right over bank quadratic curve fitting data......................................................... 93
LIST OF ABBREVIATIONS AND NOTATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEFs</td>
<td>Base Flood Elevations</td>
</tr>
<tr>
<td>BL</td>
<td>Binary length</td>
</tr>
<tr>
<td>CLOMR</td>
<td>Conditional Letter of Map Revision</td>
</tr>
<tr>
<td>DWSL</td>
<td>Delta water surface elevation</td>
</tr>
<tr>
<td>EGL</td>
<td>Energy grade line</td>
</tr>
<tr>
<td>ENCL</td>
<td>Left encroachment station</td>
</tr>
<tr>
<td>ENCR</td>
<td>Right encroachment station</td>
</tr>
<tr>
<td>FEMA</td>
<td>Federal Emergency Management Agency</td>
</tr>
<tr>
<td>FIA</td>
<td>Federal Insurance Administration</td>
</tr>
<tr>
<td>FIRM</td>
<td>Flood Insurance Rate Map</td>
</tr>
<tr>
<td>FORM</td>
<td>First-order reliability method</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithms</td>
</tr>
<tr>
<td>GP</td>
<td>Genetic Programming</td>
</tr>
<tr>
<td>HEC-2</td>
<td>Previous version of HEC-RAS</td>
</tr>
<tr>
<td>HEC-RAS</td>
<td>Hydrologic Engineering Centers River Analysis System</td>
</tr>
<tr>
<td>HGL</td>
<td>Hydraulic Grade Line</td>
</tr>
<tr>
<td>IFitness</td>
<td>Individual fitness</td>
</tr>
<tr>
<td>LOB</td>
<td>Left overbank</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>LOBB</td>
<td>Left overbank boundary</td>
</tr>
<tr>
<td>ROBB</td>
<td>Right overbank boundary</td>
</tr>
<tr>
<td>LOMR</td>
<td>Letter of Map Revision</td>
</tr>
<tr>
<td>MCH</td>
<td>Main channel</td>
</tr>
<tr>
<td>MT-2</td>
<td>Riverine Hydrology &amp; Hydraulics form</td>
</tr>
<tr>
<td>ND</td>
<td>Numbers to the right of decimal point</td>
</tr>
<tr>
<td>NFIP</td>
<td>National Flood Insurance Program</td>
</tr>
<tr>
<td>PP</td>
<td>Potential parent</td>
</tr>
<tr>
<td>ROB</td>
<td>Right overbank</td>
</tr>
<tr>
<td>SFHA</td>
<td>Special Flood Hazard Area</td>
</tr>
<tr>
<td>USACE</td>
<td>U.S Army Corps of Engineers</td>
</tr>
<tr>
<td>WES</td>
<td>Waterways Experiment Station</td>
</tr>
<tr>
<td>WSE</td>
<td>Water surface elevation</td>
</tr>
</tbody>
</table>
A flood is an overflow of water from rivers, creeks, and other waterways that submerge dry land. Floods are one of the most frequent and costly natural disasters and cause severe hardship among people and severe economic losses\[1\]. For example, in August 1931, a huge flood hit the Yangtze River. Because of heavy and long-lasting rain in July of that year, water filled up the rivers and lakes within the Yangtze River basin. In August, floods within the Yangtze River basin inundated, as many as eight provinces. According to some estimates, more than 8 million acres of farmland were flooded and as many as 145,000 people drowned\[2\]. For three months thereafter, HanJiang Plain was submerged in water, streets turned to rivers, boats became the prime transportation, and starvation and disease spread everywhere. The Central China Flood is generally regarded as the world’s deadliest flood and was the catalyst for the Three Gorges Dam\[3\].

Among all the natural disasters worldwide (excluding droughts), damage caused by floods and flood-caused debris flows account for over 90 percent of fatalities.
As an extreme natural disaster, flood could damage farmland and homes, kill crops, people, pets and livestock, and also disrupt agriculture and business [4].

**Causes of Flood**

Figure 1 shows the hydrologic cycle, that is, the circular movement of water on the earth’s surface. Water evaporated from the oceans will condense and form rainfall-bearing clouds. The resulting precipitation strikes the earth’s surface and is either infiltrated just below the earth’s surface or is intercepted by vegetation or caught in small depressions. That fraction of precipitation that is not lost is called runoff and eventually finds its way into rivers, creeks, and streams. Eventually runoff will work its way back to the oceans where the process is continually repeated.

![Figure 1 Hydraulic cycle][5]
During periods of long duration and intense precipitation, the ground may become saturated thus significantly reducing the infiltration capacity of the soil. Thus the only outlet for excess water is from surface runoff flowing to low-lying topography and eventually flowing into the nearest watercourse. Over all, flood could occur in known floodplains when long duration and intense precipitation falls, such as after a heavy rain, snowmelt, and structural damage like dam and levee failures. Plus, if it is saturated soil or hard ground cover, the excess water will be hard to absorb; and if it is a low-lying topography, pools will be easily formed.

**Flood Damage**

One reason that floods are very dangerous is due to the force of flowing water. Six inches of water could knock people off their feet, and two feet of water can lift automobiles. And it is also because that the differences in water surface levels and water volumes cause the fast water movement with great power. Floods could often result in disease and mortality, harm lives and damage property. And even, flood could cause mud rock flow to happen. Floods can be deadly to human, animals and plant species. The overall economic impact of floods including the cost of repairing flood damage and lost productivity is estimated at $2 billion annually in the US. [6].
**Flood Control**

While humans generally cannot prevent floods, they can help to mitigate flood damage. Engineers design and build flood control structures such as dikes, spurs, levees and seawalls [7]. When developing an area adjacent to a watercourse, engineers must insure that any homes or buildings be located outside of the limits of the flood.

Regardless if the engineered system is intended for flood control, for allowing traffic to cross a waterway, or as part of a residential or commercial development; it is important that these systems be designed so that they function properly and at low cost. In order to effectively and efficiently design such system, engineers must have a method available to them to study the dynamics of floodplains.

**Federal Emergency Management Agency (FEMA)**

The United States Federal Emergency Management Agency (FEMA) has been tasked with providing disaster relief and assistance to states and US Territories. Chief among the many diverse tasks that FEMA performs is managing the National Flood Insurance Program (NFIP). The NFIP is a program designed to lessen the impact of flooding on public and private structures. This is done in a variety of ways including providing mechanisms for property owners likely to be affected by floods to secure flood insurance. More importantly, however, is that FEMA encourages communities to adopt effective floodplain management strategies and enforce floodplain regulations. Communities that participate in the
NFIP typically have maps of stream systems within their boundaries that show the limits of flooding.

**Floodplain and Floodway**

A floodplain is the area adjacent to a river or stream that stretches from the banks of its channel to the base of the enclosing valley walls and experiences flooding during periods of high discharge [8]. Frequently stream systems will be evaluated under a discharge event called the 100-year flood – especially if the analysis involves a stream system that has been mapped by FEMA. A 100-year flood is a discharge of such a magnitude that it has a 1 percent chance of occurring of being exceeded during any given year. If the waterway has been mapped by FEMA, then the floodplain will include a corridor called the Regulatory Floodway.

The Regulatory Floodway, as defined by the Federal Emergency Management Agency is the channel of a river or other watercourse and the adjacent land areas that must be reserved in order to pass the base flood without cumulatively increasing the water surface elevation more than a designated height [9]. The maximum surcharge is typically 1.00 Ft. Communities are free to establish the maximum surcharge as long as it does not exceed one foot.
By establishing the floodway boundaries that produce a surcharge within 1.00 Ft, the rest of the floodplain, called the floodway fringe, can be developed for other purposes and the flood hazards will not be increased significantly. Consequently, a clear delineation of floodplains could give instruction to the land planning departments to avoid or reduce flood damage caused by floodplain development.

All development must meet the minimum floodplain management standards required for participation in the NFIP.

**Floodplain Modeling**

In some cases, a community may request to modify a regulatory floodway for better usage of land sources by shifting the floodway or changing its
configuration. Before any encroachment, the community must obtain the Conditional Letter of Map Revision (CLOMR) from FEMA.

CLOMR is FEMA’s comment on a proposed project that would, upon construction, affect the hydrologic or hydraulic characteristics of a flooding source; and thus result in the modification of the existing regulatory floodway, the effective Base Flood Elevations (BFEs), or the Special Flood Hazard Area (SFHA). After the project has been completed and built as proposed, the community needs to request a revision to the Flood Insurance Rate Map (FIRM) by requesting the Letter of Map Revision (LOMR) from FEMA within 6 months of construction completion [11].

LOMR is FEMA’s modification to the currently effective FEMA maps used to show changes in BFEs, floodplains and floodways [12]. LOMR requires MT-2 [13], annotated FIRM, data certified by a registered professional engineer and revision request acknowledged by all impacted communities [14]. For floodway revisions, it requires floodway analysis including hydraulic modeling and analysis for the city to review and see if the BEFs will be impacted by the proposed development. In the floodplain modeling report, the requestors need to clearly document all the proposed floodplain development activities comply with local, state, and federal (FEMA) floodplain regulations [13].
HEC-RAS and HEC-RAS Modeling Floodplain

Typically, the floodway boundaries are calculated using computer programs like the Hydrologic Engineering Centers River Analysis System (HEC-RAS). HEC-RAS uses an approach called the equal loss of conveyance method to establish floodway boundaries. Conveyance is a measure of the carrying capacity of a channel. The fundamental idea underlying the equal loss of conveyance is that the conveyance within the encroached cross-section after the flood fringe area are 100% blocked with a higher water surface elevation is equal to the conveyance of the natural cross-section with the original water surface elevation.
How Encroachments (Floodways) are Computed

HEC-RAS allows encroachment stations to be computed using one of five different methods:

- Method 1 – directly specify encroachment stations
- Method 2 – specify a fixed top width
- Method 3 – specify a percent reduction in conveyance
- Method 4 – Equal Loss of Conveyance based on target HGL rise
- Method 5 – Equal Loss of Conveyance based on target EGL rise

Typically Method 4 – Equal Loss of Conveyance based on a target rise in the water level is used when identifying encroachment stations as part of a CLOMR or LOMR.

Conveyance is an important term in open channel flow modeling, and it is a measure of the carrying capacity of a channel. As shown in Equation 1, conveyance considers Manning’s roughness, cross-sectional area, and hydraulic radius.

Equation 1

\[ K = \frac{1.486}{N} AR^{2/3} \]

K----Conveyance (Cfs)

N----Manning’s roughness coefficient
A----Cross-sectional area (Ft\(^2\))

R----Hydraulic radius (Ft)

\textbf{Equation 2}

\[ R = \frac{A}{P} \]

P----Wetted Perimeter (Ft)

Frequently Manning’s roughness values in the overbanks are higher than those in the main channel. As a result, the conveyance should be calculated separately for three sub-channels: the left overbank, the main channel, and the right overbank. An example of the conveyance computation for an open channel section where the roughness coefficients vary among the three sections could be found in the Appendix-1.

Appendix-1 provides the details of calculating the encroachment stations using Method 4 – Equal Loss of Conveyance based on a target rise in HGL.

1. Calculate the conveyance of the left overbank \(K_{LOB}\), the main channel \(K_{MCH}\) and the right overbank \(K_{ROB}\) for the flood event being examined. Add the conveyances together to obtain a total conveyance \(K_{TOTAL1}\).
2. For a specified rise in water level for different regulations, also called a surcharge, calculate the new conveyance of the three sub-channels for the flood event, and calculate the new total conveyance \(K_{TOTAL2}\).
3. Find the difference between $K_{TOTAL2}$ and $K_{TOTAL1}$. Compute the percent reduction of conveyance by dividing the difference between $K_{TOTAL2}$ and $K_{TOTAL1}$ by $K_{TOTAL1}$.

4. The Equal Loss of Conveyance method means that the reduction of conveyance is split equally among the left and right overbanks. That means 50% of the conveyance reduction is from the left overbank, and the other 50% is from the right overbank. If one overbank has less conveyance than the proposed reduction amount for the overbank, then HEC-RAS deducts all the conveyance available in this overbank and removes the rest from the other overbank. For example, if the conveyance reduction is 100,000 Cfs then 50,000 Cfs of conveyance must be reduced from the left overbank and 50,000 Cfs must be removed from the right overbank. Suppose the left overbank only has 30,000 Cfs of conveyance. In this case, HEC-RAS will remove 30,000 Cfs of conveyance from the left overbank and 70,000 Cfs from the right overbank – assuming that the right overbank has 70,000 Cfs of conveyance within it; if not, HEC-RAS will stop at the bank station instead of place the floodway boundary into the main channel.

5. HEC-2, the previous version of HEC-RAS was using quadratic curve fitting method to define a function that relates the amount of conveyance available in the overbank to the encroachment station. The example in Appendix-2 provides details on how this can be accomplished. HEC-RAS
has definitely improved the solve routine by using some advanced regression technologies.

6. The following example is using HEC-2’s quadratic curve fitting method. Compute the left and right encroachment station using the quadratic relationship along with the amount of conveyance that should remain in the overbank after the reduction has been removed.

7. The final step is to compute water surface elevations using the geometry associated with the newly found encroachment stations.

The procedure outlined above will not guarantee that the increase in water surface elevation associated with the encroachments produces the specified target rise. The reasons that the actual rise in water level may be different than the specified surcharge include:

1. HEC-RAS does not permit encroachment to enter main channel;

2. Curve fitting itself is the source of error, and it could only provide an approximate result;

3. Water surface elevation is obtained by using standard step method. It is a computational technique utilized to estimate one-dimensional surface water profiles in open channels with gradually varied flow under steady state conditions. As shown in the following equation, the water surface elevation is estimated by assuming a value, then replace the assumed water surface elevation with the calculated water surface elevation to do
the calculation again until an acceptable result achieved. And this is an error-associated process.

Equation 3

\[ WSE_2 = WSE_1 + \frac{\alpha V_2^2}{2g} - \frac{\alpha V_1^2}{2g} + S_f L_Q + |\frac{\alpha V_2^2}{2g} - \frac{\alpha V_1^2}{2g}| \]

Hand solving floodway stations for a cross-section

Appendix-2 is the calculation of one cross-section from a channel, while for long and complex watercourses, there could be hundreds of cross-sections and tens of branches. To develop a set of accurate encroachment stations with specified surcharge for each cross-section is much more complicated by hand solving. Thus, HEC-RAS is widely used for floodway stations’ calculation.

Once the floodway boundaries (also called encroachment stations) are found for each cross-section, then a water surface profile analysis is conducted using the cross-sectional geometry associated with the encroachment stations. There is no guarantee that the water surface elevation associated with the encroached geometry will produce a water level increase of exactly the desired amount. There are several reasons why.

1. The actual conveyance found from the curve fit may not be equal to the desired conveyance due to inaccuracies in using a 2nd order curve fit. This is well explained in the example explained in Appendix-2 that the left overbank conveyance is $339731.64\text{cfs}$ while the proposed left overbank
conveyance is 403065.117 cfs. The trend line added to express the relation between conveyance and station is only true for the picked dataset. For those are not picked and calculated, the trend line is no more than a guessed expression. Even though HEC-RAS has improved this method and could eliminate most errors occurred in the quadratic curve fitting method, it could still not provide the exactly the desired surcharge.

2. The computed water surface elevations are found using the energy equation and the geometry associated with the encroached section. Since the energy equation is used, water surface elevations are not only a function of the channel geometry but also of the previous cross-section’s water surface elevation.

3. HAC-RAS will not enter into the main channel to set the floodway boundaries.

But in order to be more confident with the floodway decision that a design could come to, the error should be minimized, in other words an exact 1.00 Ft surcharge should be achieved by the floodway modification. This thesis is using GA (Genetic Algorism) to calculate the floodway boundaries.

**Manual solution using HEC-RAS**

Steps:

1. Open the project and edit the steady flow data.
2. Edit the encroachment query and do the steady flow analysis using Method 4.

3. Check the result using a spreadsheet to see the surcharge of each cross-section.

4. Import the encroachment data to Method 1 and do the adjustment using bisectional method from downstream end to upstream end.

*Bisection method for encroachment adjustment*

Bisection method is a root solving method that repeatedly bisects an interval, and then selects a subinterval in which a root must lie for further processes. It is very simple and robust.

For a cross-section, if there is a specific encroachment station on the left or the right-overbank that could lead to a perfect solution within the geometry
boundaries, the midpoint of the boundaries, and use it as one boundary. The new boundary and one of the original boundaries consists the new boundary that contains the solution. In this case, the new boundaries are A and M-1. Within point A and M-1, find the midpoint, and name it M-2. And then find the new midpoint of A and M-2, and name it M-3. Now, M-3 replaced point A, and the new boundaries are point M-3 and M-2, so on so forth. With a few test, the encroachment that provide an accurate surcharge could be find out.

Not all the channels could result in an accurate surcharge in the encroachment finding task. There are cases that no solution would be found, in other word, no root. Most of these cases are because that the 100-year flood event flow is not large enough to provide a certain surcharge. In this case, the bisection method could find out a optimum solution with the goal to make F(X) closest to 0.

Modifying the cross-sections will not affect the downstream cross-sections, but have effect on upstream cross-sections. As a result, a small improvement in a downstream cross-section could cause tremendous change in upstream cross-sections, and sometimes the change is not in a good way.
CHAPTER 2

LITERATURE REVIEW

*GA Introduction*

**Optimization**

Optimization is the selection of the best decision variables (with regard to some criteria) that produces the best result of a given function (Objective function).

Simply speaking, optimization is the mathematical process of systematically examining different solutions to a problem, and selecting the best solution. Often times in engineering problems the best solution is the one that minimizes costs. Optimization is commonly used to solve complex problems in fields such as economics, operations research, and especially in engineering.

*Genetic algorithm (GA)*

Genetic Algorithm (GA) is an optimization technique, developed by John Holland in 1975, that mimics the process of natural selection. The underlying premise of GA is the biological processes of survival and adaptation that the strong tend to adapt and survive while the weak tend to die out [16]. Translated to optimization,
good solutions are kept and improved upon and poor solutions are discarded.

The paragraphs that followed describe the overall process associated with Genetic Algorithm optimization. Firstly though, some terms commonly used in GA optimization will be defined.

- Fitness function – a function to be optimized.
- Individual – a potential solution to the optimization problem.
- Population – a collection of individuals.
- Fitness – a number that describes how well the individual performs
- Parents – a selection of individuals in the current population to create the next generation.
- Children – the individuals created by parents from the previous generation [17].
- Crossover – an operator used to create new individuals.
- Mutation – an operator used to maintain genetic diversity.

Optimization is based on evolution, and the "survival of the fittest" concept. Generally, GA is based upon randomization. A set of randomly generated individual buildup an initial population. Every individual is evaluated by computing its fitness. Both those individuals and their fitness values are ranked and probabilities are assigned to those individuals by their rank. Then select parents randomly that a higher probability has a better chance to be chosen, and then crossover and mutation takes place to create children for next generation of population. After successive iterations, the best fitness value could be obtained.
GA processes

The process of finding an optimal solution to a problem using genetic algorithms involves several steps that are described in detail below. These individual steps include:

- Build the initial population
- Find the fitness for each individual
- Rank order the initial population
- Select the parents
- Crossover to create children
- Mutation
- Check for termination criteria

*Build initial population*

In GA coding, decision variable sets can be described as binary strings, that is, a sequence of 1’s and 0’s. This is similar to a gene chain containing individual chromosomes. Using binary strings to represent decision variables was common in early GA applications. Later in 1997, Oliveira and Loucks found real-value coding to be more effective for problems [18]. Individuals are potential solutions to the optimization problem and depending upon the nature of the optimization problem, an individual could consist of many decision variables.

The initial population is usually constructed by randomly generating a set of decision variables for each individual making up the population. The GA code
used in this research effort uses binary strings to represent the real value of the
decision variables. Either a 0 or 1 is selected randomly for each gene of the
chromosome, and this process repeats for each individual in the initial population.
The value of each decision variable is the value of the binary string.

*Find fitness for each individual*

The objective function is a mathematic function used to quantify the set of
decision variables. The objective function may be a simple formula or a complex
set of operations that require a computer-based solution. In genetic algorithms a
term called fitness is typically used to describe how well an individual performs.
Conceptually there is no difference between the objective function and the fitness
as both quantify how well a set of decision variables (an individual) satisfies the
optimization problem. The fitness value is the value of the objective function that
is computed for each individual to differentiate between individuals. Additionally,
the fitness value is the key indicator for the likelihood of an individual to survive.
A better fitness indicates the individual has a better chance of survival and thus
has stronger influences on future generations.

*Rank order from lowest to largest*

Once the fitness value for each individual is found, the GA program sorts the
individuals of the first generations based on their fitness value. For a
minimization problem, individuals are ranked in ascending order – the best-fit
individual has the smallest fitness value. For maximization problem the population is ranked in descending order.

**Selection**

The next step in GA is to create a mating pool that is a collection of individuals that may (or may not) produce offspring. The size of the mating pool is typically determined by the size of the population. For example a mating pool size of 40% for a population size of 100 individuals has 40 potential parents. In this case, the 40 individuals with the best fitness are placed in the mating pool and are also included in the next generation. The remainder of the next generation is populated with children.

Not all individuals in the mating pool will become parents. Several approaches are available to identify which individuals in the mating pool will become parents. One popular approach is fitness proportionate selection introduced by Goldberg in 1989 [19].

**Equation 4**

\[
P_i = \frac{f_i}{\sum_{j=1}^{n} f_j}
\]

The probability of a potential parent \(i\) to be chosen as a mate \(P_i\) among \(n\) individuals is its fitness value over the sum of all potential parents' fitness values in the mating pool. Equation 4 is the general expression for fitness proportionate
selection, when minimizing the fitness value, the expression modified to Equation 5:

\[ P_i = \frac{f_{n+1-i}}{\sum_{j=1}^{n} f_j} \]

There are many other approaches such as equal probability selection and tournament selection [20].

With the fitness proportionate approach, those members of the mating pool that have a better fitness value have a greater probability of becoming a parent. One way that this can be accomplished is to develop a probability band that is correlated with the fitness value.

<table>
<thead>
<tr>
<th>PP Index</th>
<th>Fitness</th>
<th>PP Index</th>
<th>Rnd Fitness</th>
<th>Probability</th>
<th>Select Zone</th>
<th>Rnd pick</th>
<th>Picked Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.10</td>
<td>6</td>
<td>1.09</td>
<td>0.16</td>
<td>0.16</td>
<td>0.37</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>3.61</td>
<td>5</td>
<td>1.79</td>
<td>0.14</td>
<td>0.29</td>
<td>0.09</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>2.86</td>
<td>10</td>
<td>2.04</td>
<td>0.13</td>
<td>0.42</td>
<td>0.92</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>3.37</td>
<td>9</td>
<td>2.33</td>
<td>0.11</td>
<td>0.53</td>
<td>0.16</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>1.79</td>
<td>7</td>
<td>2.39</td>
<td>0.11</td>
<td>0.63</td>
<td>0.98</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1.09</td>
<td>8</td>
<td>2.78</td>
<td>0.09</td>
<td>0.72</td>
<td>0.54</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>2.39</td>
<td>3</td>
<td>2.86</td>
<td>0.09</td>
<td>0.81</td>
<td>0.69</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>2.78</td>
<td>4</td>
<td>3.37</td>
<td>0.08</td>
<td>0.89</td>
<td>0.44</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>2.33</td>
<td>2</td>
<td>3.61</td>
<td>0.07</td>
<td>0.96</td>
<td>0.08</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>2.04</td>
<td>1</td>
<td>4.10</td>
<td>0.04</td>
<td>1.00</td>
<td>0.67</td>
<td>8</td>
</tr>
<tr>
<td>SUM</td>
<td>26.37</td>
<td>SUM</td>
<td>26.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Selection example
For example, as shown in Table 1 above, the mating pool is 10, which means 10 mates will be selected from the 10 potential parents. Each of them has its fitness. The fitness values are ranked from low to high in this minimization problem as shown in Column 4.

The probability for potential parent 6 in row 1 is \( P_1 = \frac{f_{10} + 1 - 1}{\sum_{j=1}^{10} f_j} = \frac{4.10}{26.37} = 0.16 \). The selection zone for potential parent 6 in row 1 is \([0, 0.16)\), and the selection zone for potential parent 5 in row 2 is \([0.16, 0.16 + P_2)\). A random number between 0 and 1 is selected and the individual associated with the value of the random number is selected from the mating pool to become a parent. This is illustrated in Table 1 above through Columns 7 and 8. For example, a random number having a value of 0.37 is generated. The value of 0.37 falls within the probability band \([0.29, 0.42)\) which is associated with individual #10. Thus individual #10 is the first of 10 parents.

As the probability band gets smaller and smaller, the chance for a random number between 0 and 1 falling in to the band also gets smaller. This approach, therefore, generally results in those individuals who have a better fitness also having a greater chance of being selected as a parent.

Whichever scheme is used, the total mates’ quantity has to equal the amount of parents in order to maintain the population. For example, the quantity of potential parents in the mating pool is 10, and then, the quantity of parents selected using the proportionate selection method is 10 too. As the example above illustrates, one potential parent can produce more than one mate, for example, both
potential parent 6 and 8 have been chosen twice and potential parent 3 and 4 has never been chosen.

**Crossover to create children**

Crossover is used to generate new individuals called children. Children and the individuals from the mating pool form the next generation or population. Crossover methods include one-point crossover, two-points crossover and uniform crossover.

![One-point crossover](image)

**One-point crossover**

In one-point crossover, a random position within the parent's binary string is selected and the gene or genes on either side of that position are switched as shown in Figure 5.
In two-point crossover, two random positions within the parent’s binary string are selected and they divide the string into three parts of genes. The second part of genes is switched with the first and the third part of genes remains in the same position in the children created as shown in Figure 6.
Unlike one or two-point crossover, uniform crossover is using a fixed mixing ratio between two parents and enables the parent chromosomes to contribute by ratio [21]. For example, as shown in Figure 7, the mixing ratio is 0.5, that means 50% of genes from the first parent is transferred to the first child and the other 50% is transferred to the second child with the crossover positions randomly selected.

**Mutation**

After a few generations, all the individuals in a generation might converge to the same solution. Indeed, the genetic algorithm optimization process is completed when all individuals converge to the same solution for several consecutive generations. Mutation is a way to introduce variability in the population when there is none. When representing decision variables using binary strings, basic mutation can be achieved by the random flipping of bit values from 0 to 1 (or from 1 to 0). In the mutation process, a random individual is selected, and then a gene in a random position is selected to mutate as illustrated in Figure 8.

![Figure 8 Single point mutation at 11th gene from 0 to 1](image-url)
Mutation could alter one or more bit values in a chromosome from its initial state; thus, there is single point mutation and multi point mutation. Even though mutation may not occur frequently, it may change the entire solution for future generations.

![Multi point mutation at 3rd gene from 1 to 0, and 11th gene from 0 to 1](image)

**Check for termination criteria**

If the best fitness is good enough, then the GA evolution procedure is completed and a solution is found. The indicator for “good enough” is related to the nature of the problem and how much error could be accepted. The GA process described above is repeated with a new generation being generated for each iteration. Figure 10 describes the overall genetic algorithm process.
Advantages and disadvantages

GA has so many great advantages that it has received wide acceptance as an optimization technique particularly for engineering problems. Some of the advantages of GA include:

1. GA is easily understood and does not require the knowledge of complex
mathematics since the calculations are rather simple and straightforward.

2. GA is capable of solving very complex problems such as multi-dimensional problems and problems with multiple solutions.

3. GA is easy to transfer to existing simulations and models, so that it could use existing models as fitness functions to solve even more complex problems such as being combined with HEC-RAS to calculation encroachment stations.

4. GA deals directly with a bunch of solutions within a range every time. As a result that getting an optimum solution could be faster than using traditional methods for complicated problems.

5. Within GA processes, different solutions that are close to the optimum solution could be obtained as well to provide comparison of different designs.

While GA has distinct advantages over other optimization techniques, there are some drawbacks:

1. Constrained optimization problems are difficult to solve without using penalty-based methods.

2. There is no absolute assurance that a genetic algorithm will find a global optimum solution. [22]

3. Genetic Algorithm cannot assure constant optimization response times.
4. There are many parameters needed in GA such as parent rate, mutation rate, how the mating pool is selected, how the cross-over is performed, etc. These parameters can greatly affect the final result. Consequently much testing should be done with GA parameters to determine how the parameters affect the problem being solved.

History of FEMA Floodplain Management

Historically people have been attracted to rivers and streams for living, industry, commerce, and recreation since these bodies of water could provide water supply, power, and transportation. As people are using up more and more spaces near the water, people want more from filling the floodplain. This action could obstruct flood flows, back up floodwaters onto upstream and adjacent properties, and most importantly, it reduce the floodplain’s ability to store excess water, send more excess water downstream and cause floods to rise to higher levels with floodwater velocity increase.

Floodplain management is the operation of a community program of corrective and preventative measures for reducing flood damage. Even before this term was created, our ancestors started to manage the floodplain thousands of years ago. Around 4000 B.C the first dam was built by Egyptians to protect their homeland from flooding, deliver and store water.
Since the early 20th century, the federal government took a major role in flood control and got significant progress in floodplain hydraulic analysis. In 1927, the U.S Army Corps of Engineers (USACE) founded the Waterways Experiment Station (WES) and lead floodplain hydraulic analysis for many years. Building massive flood control structures was the primary way to reduce flood losses at that time, but flood losses were still increasing even though more and more dams, levees, and floodwalls were built. The main reason for that is people continued to build in floodplains. Since then, floodplain management started to use a combination of many tools.

In 1968, the National Flood Insurance Program (NFIP) was created to set a national standard for regulating new development in floodplains as well as to provide financial aid after floods under the Federal Insurance Administration (FIA)’s full responsibility.

In 1979, FIA and the NFIP were transferred to the newly created the Federal Emergency Management Agency (FEMA).

FEMA provides federally guaranteed flood insurance to properties in those communities that agree to regulate development in their mapped floodplains. NFIP has three basic parts:

1. Mapping: FEMA funds floodplain and flood hazard maps to be primarily used as a basis for regulating new flood prone construction, and rating flood insurance policies.
2. Insurance: FEMA certifies communities’ eligibility of joining NFIP and provides relief for all floods, including those not large enough or severe enough to warrant federal disaster aid to every building located in a participating community.

3. Regulation: FEMA designs the floodplain regulations, helps local flood prone communities to adopt the regulations, and evaluates its enforcement to ensure that new buildings will be protected from the flood levels shown on the FIRM and that development will not make the flood hazard worse [23].

Now, FEMA concluded the definition of floodplain management as a decision-making process that aims to achieve the wise use of the nation’s floodplains. Its purpose is to reduce flood losses and protect natural resources and function of floodplains. The basic principle of floodplain management is that the flood hazard on other properties must not be increased by the development. FEMA’s goal is to convert all communities to the regular phase of the NFIP as quickly as possible.

**Brief Review of Genetic Algorithms**

GA is a search algorithm based on the mechanics of the natural selection process (biological evolution) that was developed by John Holland in 1975 [24].

In 1962, John Holland at University of Michigan began to publish and teach adaptive systems, and in 1965 first described the importance of crossover [25].
In 1965 Ingo Rechenberg devised evolution strategy without population or crossover, it was one parent mutated to produce one offspring with the better individual being kept [26].


In 1973, Rechenberg developed the evolutionary strategy in Berlin, Germany [28].

In 1975, John Holland published *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, which is considered by most to be the seminal work in the field. The population-based method was a huge innovation [25].

In the 1970s to 1980s, GA research increased rapidly due to advances in computer technology in both theory and application. It has been applied successfully in many fields in engineering, mathematics, water resources and science, etc.

In 1992, Koza developed Genetic Programming (GP) that is a new methodology to provide a way to run a search. This fulfills his suggestion that a desired program should evolve itself to search for the most fit computer program to solve a particular problem [29].

Genetic Algorithms being used today is the result of decades of hard work from
former scientists and engineers in many fields by trying to solve different types of problems testing reasonable innovative ideas.

**Review of GA in Water Resources Field**

In water resources field, many works have been done using GA programming. In the past several decades, the most application of GA on water resources is water distribution system. By modeling pipes, hydraulic control elements such as pumps, valves, and tanks to insure the serving area could get sufficient supply with appropriate pressure and minimum cost.

In 1977, Cembrowicz and Krauter used Rechenberg’s biological evolution strategy concepts to optimize pipe networks [30]. In 1987, David and Chie applied GA to the steady state optimization of a serial liquid pipeline. They found it effective and should permit to more complex, highly dimensional problems [31].

In 1991, Wang applied a GA to the calibration of a conceptual rainfall-runoff model [32]. Similar, in 1996, Franchini, who used a GA in combination with sequential quadratic programming to calibrate a conceptual rainfall-runoff model [33].

In 1994, Angus and Graeme applied GA to the problem of designing pipe networks and compared its performance with the techniques of complete enumeration and nonlinear programming. The pipe sizes as decision variables were coded as binary strings, and they used reproduction, crossover, and
mutation in the process. They found GA effective in finding global optimal or near-optimal solutions for the pipe network optimization [34].

In 1997, Dragan and Godfrey applied genetic algorithms to the problem of least-cost design of water distribution networks [35]. In 1998, Ann and Linfield used GA to estimate water quality model parameters in a calibration scenario. They used synthetic data and found it capable of accurately finding optimum parameter sets [36].

In 2004, Bryan and Holger linked GA with the first-order reliability method (FORM) for reliability-based optimization of water distribution networks effectively, this paper demonstrates the novel combination of a GA with FORM [37].

It has also been used in storm sewer system management and water and wastewater treatment plants as well as water quality management problems, even though it is still facing some problems.

History of FEMA Floodway

Floodway definition

FEMA defined floodway as the channel of a river or other watercourse and the adjacent land areas that must be reserved in order to discharge the base flood without cumulatively increasing the water surface elevation by more than a designated height. And FEMA criteria state that the maximum increase or surcharge is 1.0 Ft at any location. Floodway boundaries are virtual boundaries
that are found by computer modeling for land use planning and flood insurance study. Frankly, the hypothetical boundary is a “safety range” for floodplain development.

The importance of floodways

Since floodways provide information and regulation of land use planning and flood prone development, they are crucial tools for floodplain management. Once the floodway is blocked, flood characteristics could be greatly impacted, such as increasing the peak discharge, velocity that could lead to a much more severe flood casualty and damage. Additionally, the floodway fringe, means the rest of the floodplain, could be used for construction that removes conveyance from the floodplain, and water surface elevation would increase within a safe range.

Floodway management method

Floodways must remain open and prohibit obstruction to carry the deeper, faster moving water during a flood. FEMA allows floodplain encroachment within a 1.00 Ft surcharge for 100-year floods. Some states have a more stringent surcharge limit compared to FEMA’s allowance, such as Illinois’s limit is 0.1 Ft and Wisconsin is 0.0 Ft. The smaller the allowance is, the less encroachment could take place. There is no development permitted within the regulatory floodway to keep it free of obstructions to pass large flood discharges, on the other hand, the flood fringe area could be filled completely for blocking. But there are still possibilities for a project within the floodway to take place—conducting an
engineering analysis to get a no-rise certification supported by technical data that is signed by a registered professional engineer. “No rise” means the proposed encroachment causes no rise in 100-year flood surface elevation, even 0.001 Ft rise is unacceptable.

**Fair and Equitable Delineation of Floodway**

In determination of floodways, not only the surcharge amount is considered. Fair and equitable delineation also plays an important role. Properties on both sides of a river or stream should be treated equitably. Even if there is only one side of land planned to project the encroachment, development could only fill up half of the hydraulic capacity and leave the other half for the future projects on the other side of the river. For example, if an owner wants his river plain to be filled, he could only fill half of the allowed reduction of conveyance. Otherwise, excess reduction is considered wrongfully detained.

**Equal Loss of Conveyance**

Equal flood capacity reduction on opposite sides of stream is used due to fair and equitable delineation of floodways. The loss of capacity for flow is computed in terms of conveyance. In floodplain delineation calculation, after a total conveyance reduction is obtained, divide this value into two equal terms, and then use the new conveyance to calculate the location of floodway for each side. However, there are circumstances that use equal loss of conveyance could lead to one boundary located within the floodway, and this is inhibited. For example,
one side of the cross-section has less conveyance than the amount intended to be subtracted from, HEC-RAS will reduce the rest from the opposite side of the floodplain.

**Ways to Determine Floodway Stations**

The HEC-RAS floodplain encroachment procedure is based on calculating the water surface elevation in natural conditions and encroached conditions. Then target water surface difference is used to find deductible conveyance and this conveyance is used to calculate floodway stations with equal loss of conveyance as a choice. There are five methods being used by HEC-RAS:

1. The left and right encroachment stations are provided as input. Method 1 specifies the exact location of encroachment stations on each side of the channel. This method is used for modelers to make small adjustments to individual cross-sections to get a better floodway.

2. The desired floodway width is provided as input. Method 2 sets a specific floodway top width. This method is used for studies where the communities want to set equal widths from the stream centerline to establish a floodway.

3. The percent conveyance reduction is specified as input. Method 3 specifies a percent reduction in conveyance, $x\%$ of which applied to the left side and $1-x\%$ applied to the right side.

4. Target surcharge with equal conveyance reduction is specified as input.
Method 4 is similar to Method 3 except the target WSE above the 100-year unencroached water surface elevation is specified and half of the conveyance applied to the left side floodplain, and the other half applied to the right side floodplain.

5. Two targets are optimized.

Methods 5 builds on Method 4 by adding allowable increase in the energy grade line elevation. This method is an optimization technique to achieve the target water surface elevation for all cross-sections by up to 20 iterations.

The modeler develops a floodway using HEC-RAS following these steps:

1. Develop base flood elevation when no floodway is present.
3. Modify the floodway target increases for portions of reach and finalize floodways through Method 1 [21].

Modifying the floodway target manually by adding or subtracting a number from the floodway stations is an unpredictable process. The water surface elevation is varying with the floodway encroachment station. The water surface elevation difference for encroached and unencroached situations are not directly linked. For the same cross-section, a wider floodway does not necessarily provide a smaller water surface elevation difference. This is because different parameters could be used to calculate the difference of water surface elevations. So they are neither positive correlational nor negative correlational. To get good results that
make the surcharge close to the target value, this process is time and labor consuming.

**Optimal Locations of Encroachment Stations**

Method 4 is the basic method to find floodway boundaries on both sides of the channel as it considers fair and equitable delineation of conveyance. Generally speaking, Method 4 is the preferred method for identifying the location of encroachment stations, i.e. delineating the floodway. Method 4 – the Equal Loss of Conveyance method usually results in a fair and equitable floodway delineation as the approach attempts to achieve the same conveyance reduction from both overbanks. Recall that “Ways to determine floodway stations” section on page 63 provides details on the Equal Loss of Conveyance method.

Method 4 provides encroachment stations that frequently produce surcharges at or near the target water level rise, but sometimes not exactly equal to the target rise. As discussed above, in such cases modelers have to manually manipulate the encroachment stations (Method 1) until the specified target rise is obtained exactly. Since the use of Method 1 can often times be laborious and time consuming, an optimization technique was developed in an effort to automate the process.

**Unique Contribution of This Effort**

This article innovatively uses Genetic Algorithms combined with HEC-RAS in calculating floodway boundaries. A review of the literature indicates that very little
research has been done on this topic. Thus the optimization approach
documented below could provide modelers with a bfes given the fact that using
Method 1 to find encroachment stations that exactly meet a specified surcharge
can be very time consuming.

By using the GA program, the average time required to find floodway stations
could be decreased multifold, but it simplifies modelers' work and make getting a
result faster.
CHAPTER 3

DEVELOPMENT OF THE OPTIMIZATION MODEL FOR FLOODWAY STATION’S CALCULATION

The goal of the methodology proposed here is to find encroachment stations via Genetic Algorithm optimization that exactly produce the specified surcharge or water level increase. Mathematically this can be expressed as minimizing the sum of the differences between the computed water surface elevation associated with the encroached cross-section and the target water surface elevation. Note that the target water surface elevation is the water level associated with the un-encroached cross-section plus the specified surcharge.

Equation 5 below represents the objective function or fitness function for this problem.

Equation 6

\[
Obj = \min \sum_{i=1}^{n} |\text{Computed WSEL} - (\text{Original WSEL} + \text{Surcharge})|
\]

Where:
• Computed \( WSEL \): computed water surface elevation associated with the encroached geometry

• Original \( WSEL \): original water surface elevation associated with the original geometry

• Surcharge: the target water surface elevation rise, normally is 1.00 Ft or follow state or local surcharge standard.

**Define Terminologies**

**Fitness function**

For this problem, the fitness function is the sum of the absolute value of the difference between the water surface elevation associated with the encroached geometry and water surface elevation associated with the undisturbed or un-encroached geometry plus the surcharge.

Using the absolute value of the difference could eliminate the situation of compensating errors. Compensating errors may occur when summing positive and negative numbers differences result in a value of zero. Using the square of the differences instead of absolute value could also avoid this problem, but the fitness value is also act as a parameter to estimate how close the optimized result is to the desired result, and it is more convenient to represent the fitness value in Ft rater than Ft\(^2\). For example, as shown in Table 2, if we use the absolute value of the differences of the original water surface elevation and the surcharged new water surface elevation, a 1.89 fitness value is obtained. This
means that the average error on each cross section is 0.189 ft. But using square instead of the absolute value provides a 0.50 fitness value. There is no direct estimation from this 0.50 fitness value to any average error of each cross section. This might mislead the modeler to overestimate the current solution, or underestimate the current solution as shown in Table 3. It is clear that using the absolute value in calculation of the fitness function is more convenient.

Table 2 Comparison of using absolute value and square value in fitness function (A)

<table>
<thead>
<tr>
<th>Cross-Section</th>
<th>DWSL</th>
<th>ABSOLUTE</th>
<th>SQUARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.26</td>
<td>0.26</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>1.13</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>0.96</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>1.21</td>
<td>0.21</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>1.38</td>
<td>0.38</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>1.31</td>
<td>0.31</td>
<td>0.10</td>
</tr>
<tr>
<td>7</td>
<td>0.89</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>8</td>
<td>1.03</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>9</td>
<td>1.32</td>
<td>0.32</td>
<td>0.11</td>
</tr>
<tr>
<td>10</td>
<td>0.92</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>SUM</td>
<td>11.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitness Val</td>
<td>1.89</td>
<td></td>
<td>0.50</td>
</tr>
</tbody>
</table>
Table 3 Comparison of using absolute value and square value in fitness function (B)

<table>
<thead>
<tr>
<th>Cross-Section</th>
<th>DWSL</th>
<th>Absolute</th>
<th>Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.26</td>
<td>1.26</td>
<td>1.60</td>
</tr>
<tr>
<td>2</td>
<td>2.13</td>
<td>1.13</td>
<td>1.28</td>
</tr>
<tr>
<td>3</td>
<td>0.96</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>3.21</td>
<td>2.21</td>
<td>4.87</td>
</tr>
<tr>
<td>5</td>
<td>1.38</td>
<td>0.38</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>2.31</td>
<td>1.31</td>
<td>1.73</td>
</tr>
<tr>
<td>7</td>
<td>0.89</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>8</td>
<td>2.33</td>
<td>1.33</td>
<td>1.77</td>
</tr>
<tr>
<td>9</td>
<td>1.32</td>
<td>0.32</td>
<td>0.11</td>
</tr>
<tr>
<td>10</td>
<td>0.92</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>SUM</td>
<td>17.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitness Val</td>
<td>8.18</td>
<td>11.52</td>
<td></td>
</tr>
</tbody>
</table>

**Decision variables**

In the research effort documented in this report, two different decision variables were examined and compared. The first decision variable that was examined was the distance from the bank stations as shown in Equation 6 and Equation 7.

**Equation 7**

\[ X(L) = \text{Left bank station} - \text{Left encroachment station} \]

**Equation 8**

\[ X(L + 1) = \text{Right encroachment station} - \text{Right bank station} \]

One of the disadvantages associated with this definition of the decision variable is that unreasonable encroachment stations could be produced. For example, if
the left bank station was 100 Ft, and X(L) was 373 Ft then the left encroachment station is -273 Ft, and that is normally outside of the mapping zone and could cause a lot unnecessary engineering work.

In an effort to address this disadvantage, another decision variables examined in the research effort was the percent of the overbank width. Consider the cross-section shown in the Figure 11 below. The location of the left encroachment station is within the left overbank. And the distance between the left encroachment station and the left overbank is the left over bank width (W_{LOB}) multiply by the percent of overbank length which is the decision variable as shown in Equation 8.

\[
ENCL = Left \text{ bank station} - X(L) \times L_{LOB}
\]

Figure 11 Define percent overbank width as decision variable

Equation 9
Equation 10

\[
ENCR = X(L + 1) \ast L_{ROB} - \text{Right bank station}
\]

In this case a decision variable is the percent value. The left encroachment station (ENCL) is equal to the left bank station less the value of the decision variable times the left overbank distance. The right encroachment station is computed in a similar manner as shown in the second equation above.

In the approach documented here, a percentage value was obtained by setting the maximum allowable value to approximately 100%. Compared to binaries, 10 digits of binaries values 1024. After a value adjustment by dividing 1000, the decision variables are limited between 0 and 1.024, thus, the maximum value is very similar to 1.

**Develop GA Processes**

In this paper, a customized GA process is built with some modification of the traditional process. These modifications are to make the program more suitable for finding floodway stations with HEC-RAS.

**Build population**

This research of GA uses binary strings to describe decision variables. Initial population is built by randomly selecting each digit in every decision variable of each individual. If an individual has n cross-sections, and the total decision variable quantity is 2*n; this is because there are two encroachment stations in each cross-section. For each decision variable's binary string, either 0 or 1 is
selected randomly for each gene. After all the individuals are created, they form the first generation.

Find fitness for each individual

Fitness for each individual is calculated from a successive process involving the following steps:

1. Find the numerical value of each decision variable by converting the bit-string to a real number.
2. Replace the original encroachment stations in HEC-RAS' plan file with the current set of encroachment stations, and conduct HEC-RAS simulation to get the original water surface elevation and encroached water surface elevation.
3. Finally, the fitness value is the result of objective function with the two water surface elevations as substitution.

The computer time required for each individual depends on the data set quantity; for instance, a study area has many cross-sections, and complicated geometry data will cost more time to run the simulation than a case has a few cross-sections with simple geometry data. The process time is about 0.5s for a channel with 12 cross-sections and 12 ground-points in each cross-section.
**Ranking**

The goal of finding the floodway stations is to find a set of locations that lead the surcharge to the target surcharge as accurate as possible. This is a minimization problem; thus, the individuals are ranked by their fitness in an increasing order. Record the best fitness and its correlative individual’s information.

**Selection**

In this paper a reversed fitness proportionate selection is used. It is based on the popular fitness proportionate selection.

The procedure of fitness proportionate selection is:

1. Get the fitness value for each individual, and normalize by dividing the fitness value of each individual by the sum of all fitness values.

\[ P_i = \frac{f_i}{\sum_{i=1}^{n} f_i} \]  

2. Sort them in a descending order.

3. Calculate each cumulative fitness probability by the following equation:

\[ P_{c,i} = P_i + \sum_{i=1}^{i} P_{i-1} \]

4. Randomly select a number \( R \) between 0 and 1.
5. Select the first individual whose cumulative fitness \( p_{c,i} \) is greater than \( R \).

The modification is from Step 2. This project is a minimization problem, thus the individual with the smallest fitness should have the greatest chance to be chosen. The individuals are ranked in an ascending order of \( p_i \), and then the largest \( p_i \) is assigned to the first individual, the second largest \( p_i \) to the second individual, etc. The smallest \( p_i \) is assigned to the last individual. The rest of the steps remain the same.

Selection is conducted from the potential parents in the mating pool that are related to parents’ rate. The rest of individuals besides the potential parents are “killed” because they couldn’t provide a good fitness result. Those individuals in the mating pool that haven’t been chosen are “killed” as well. Selection will be terminated when the newly selected parents are filling up the mating pool. In this way, the parents’ quantity will remain the same for every generation.

Additionally, a direct selection of the top 4 individuals as the parents is also an option in this program. This option might be useful when dealing with small projects.

**Crossover**

Crossover is to produce new children from parents to replace the killed individuals outside of the mating pool. Determine the crossover position via a
random number generator that a random integer within the range of genome length is generated to indicate the position to flip the two strings. For example, as shown below, the random number generated is 9, which means the first 9 genes of the first child come from the first parent, and the rest of the genes of the first child come from the second parent. And the other two halves consist the second child.

When the newly produced children’s quantity is equal to the individuals’ quantity outside of the mating pool, the crossover is finished for this generation.

**Mutation**

As discussed in the “Mutation” section on page 50, after a few generations of evolution, the whole generation is much more likely to converge to a few similar
individuals which could not provide enough variability of gene combination. For example, if the population is 100, and the parents' rate is 20, this means 20 best individuals are selected to become potential parents. And the mates are selected from the 20 individuals based on reversed fitness proportionate selection. There is a very high chance to choose one individual multiple times in the same generation. Then after crossover, there is still a big chance to have several individuals that are the same in the first 20 individuals of the ranked second generation. As the process moves on, the diversity decreases dramatically. As the standard divisions of the individuals start to get below 1, mutation is activated.

Conventional single point mutation is randomly selecting the gene from a randomly selected individual to mutate by switching the binary value of the gene for this individual. And in this program, half of the individuals are mutated instead of one. As shown below, the random number generated is 11, and the 11th gene in the string is converted from 0 to 1 (white to black).

Figure 13 Single point mutation example
Additionally, a uniform mutation method is an option. Uniform mutation in this program is adding or subtracting a real number between 0 and 50 to each decision variable, and then converting the real number to a binary string. This option has not been achieved yet and needs to be tested afterward.

**Check for termination criteria**

There are three ways to end the GA process including solution is found, max generation is reached, and fatal error in HEC-RAS computation occurred.

1. **Solution is found.**

   When individual fitness is less or equal to a small number such as 0.5 or 0.01*CrossSectionNumber, the solution is found.

2. **Max generation is reached.**

   If the preset parameter of max generation has been reached, the GA process has to end too.

3. **Fatal error in HEC-RAS computation occurred**

   Whenever HEC-RAS encounters a computation problem, the GA process is terminated.
Discuss GA

The GA process is working great on tons of problems, and being applied on finding floodway stations with HEC-RAS is very convenient. GA cultivates a good result by giving it pressure to survive. Since the GA process kills all the individuals that are outside of the parents’ range, it lost tons of good genes that might lead to a success later on. But there are always new gene combinations that could come out during the mutation process.

Using HEC-RAS alone to find the floodway station of a river is very complicated as discussed in "Ways to determine floodway stations" section on page 64. In consideration of equal loss of conveyance and the quadratic curve fitting, the result from Method 4 must be calibrated by hand in Method 1, again and again. The calibration is totally based on experience and guessing. It would take days to get a relatively usable floodway station result. By using the GA program with HEC-RAS, all the manual work could be simplified. The program guesses and tries and keeps the best genes to generate better results. This saves much energy and time for modelers to use this program. But as an optimization process, it does not provide an absolute assurance to get a perfect solution nor a constant response time, neither does HEC-RAS. In the spirit of Genetic Algorithm, the process is to try to provide a better solution automatically.
CHAPTER 4

RESULTS AND DISCUSSION OF SIMULATIONS

Parameters Tested

The parameters tested are population size, percent to mutate, and parents’ rate.

The solving methods include HEC-RAS calculation, HEC-RAS manual modification, and GA simulation.

Population size

This section is to test the best population size from 30 to 200 for the case study. Other uncertain parameters assume 20% parents’ rate and mutate 80% of the population. Varying the population size from 30 to 200 is because a small population size could not provide enough individual diversity; and even though a large population size could provide more chances to have a better guess in the first place, it also requires time to process such a large quantity of calculations.

The population size test results are shown in Table 4. The best fitness value obtained is from the generation with 50 individuals. However, when the
population size varies from 50 to 200, the best fitness value does not change much. Considering the time issue, the population size of 50 is used for later tests.

Table 4 Population size tests

<table>
<thead>
<tr>
<th>Tests</th>
<th>Population size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td>1</td>
<td>4.51</td>
</tr>
<tr>
<td>2</td>
<td>3.40</td>
</tr>
<tr>
<td>3</td>
<td>3.29</td>
</tr>
<tr>
<td>4</td>
<td>3.48</td>
</tr>
<tr>
<td>5</td>
<td>3.77</td>
</tr>
<tr>
<td>6</td>
<td>3.88</td>
</tr>
<tr>
<td>7</td>
<td>3.19</td>
</tr>
<tr>
<td>8</td>
<td>4.21</td>
</tr>
<tr>
<td>9</td>
<td>3.37</td>
</tr>
<tr>
<td>10</td>
<td>3.69</td>
</tr>
<tr>
<td>Average</td>
<td>3.68</td>
</tr>
<tr>
<td>Minimum</td>
<td>3.19</td>
</tr>
</tbody>
</table>

Mutation percentage

The mutation percentage is the control of the percentage of population that will mutate during the mutation process. As the standard division gets smaller and smaller, technically, for a smaller percent of mutation, the result mostly depends on the initial guess, and the individuals do not vary a lot after mutation.

This section is to varying the mutation percentage from 20% to 80% to test its influence on the best fitness value. The tested percentage interval is evenly distributed to get a general idea of its influence.
Tests result in Table 5 shows that the mutation percent does not cause any obvious impact on the result. In this test, all of the best fitness values are obtained within the 10th generation and no further improvement in the mutation after that. This might be because that the value of mutation position is either too large that it changed a good decision variable to a bad one or too small that it couldn't reflect change in the decision variable. Thus, a modified mutation process should be developed that will be discussed in the improvement section.

Considering the potential shortcoming, a moderate 50% mutation rate is chosen for later tests. Once the new mutation method is finished, the mutation percentage parameter should be tested again to reveal the real relationship between the mutation percentage and the best fitness value.

<table>
<thead>
<tr>
<th>Tests</th>
<th>Mutation Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>3.67</td>
</tr>
<tr>
<td>2</td>
<td>3.79</td>
</tr>
<tr>
<td>3</td>
<td>3.08</td>
</tr>
<tr>
<td>4</td>
<td>2.58</td>
</tr>
<tr>
<td>5</td>
<td>3.85</td>
</tr>
<tr>
<td>6</td>
<td>3.99</td>
</tr>
<tr>
<td>7</td>
<td>4.20</td>
</tr>
<tr>
<td>8</td>
<td>3.72</td>
</tr>
<tr>
<td>9</td>
<td>2.31</td>
</tr>
<tr>
<td>10</td>
<td>3.13</td>
</tr>
<tr>
<td>Average</td>
<td>3.43</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.31</td>
</tr>
</tbody>
</table>
Parents’ rate

The parents’ rate decides how many parents should be kept to the next generation. A smaller parents’ rate could lead to a smaller variety of individual genes in the next generation, and a larger parents’ rate could lead to a longer duration of evolution to reach the goal fitness. This is the famous population diversity and the selective pressure relationship. The two factors are inversely related. As selective pressure increases, the parents’ rate is smaller, which indicates the search is focusing on the top few individuals in the population, and because of this, the available genetic diversity decreases. Reducing the selective pressure (parents rate is larger) could increase the genetic diversity because more parents are involved in the search [38].

This section is to varying the parents’ rate from 20% to 60% to test its influence on the best fitness value. The tested percentage interval is evenly distributed and the max tested value is 60% is because if the selective pressure is too small it will cause the decrease of quality of the selected individuals.

From the results shown in Table 6, there is no obvious relationship between parents’ rate and generation time. This could be because GA is an optimization process and there is no guarantee for a certain result to be obtained in certain conditions, the selection of those parameters will mostly depend on trails for different situations. A organized and large amount of test should be done to reveal the real relationship between the parents’ rate and the best fitness value.
Table 6 Parents’ rate tests

<table>
<thead>
<tr>
<th>Tests</th>
<th>Parents’ Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>2.69</td>
</tr>
<tr>
<td>2</td>
<td>2.97</td>
</tr>
<tr>
<td>3</td>
<td>4.34</td>
</tr>
<tr>
<td>4</td>
<td>3.16</td>
</tr>
<tr>
<td>5</td>
<td>3.33</td>
</tr>
<tr>
<td>6</td>
<td>2.62</td>
</tr>
<tr>
<td>7</td>
<td>3.81</td>
</tr>
<tr>
<td>8</td>
<td>3.49</td>
</tr>
<tr>
<td>9</td>
<td>2.89</td>
</tr>
<tr>
<td>10</td>
<td>3.49</td>
</tr>
<tr>
<td>Average</td>
<td>3.28</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.62</td>
</tr>
</tbody>
</table>
Methods Comparison

In the GA process experiment, an artificial channel showed in Figure 14 is used. There are 12 cross-sections on this sample channel and each of them has two decision variables, in total, there are 24 decision variables for each individual. In this section, results from three methods are compared. Note that WS PF is the water surface profile.

![Figure 14 Sample channel illustration](image)

Encroachment result comparison

Table 7 is the result of encroachment stations’ calculation from manual solution, HEC-RAS solution and the GA program. And Figure 15 is the chart reveals the actual location of the encroachment stations along the channel.
Table 7 Encroachment results comparison

<table>
<thead>
<tr>
<th>Section</th>
<th>Manual</th>
<th>HECRAS</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enc L</td>
<td>Enc R</td>
<td>Enc L</td>
</tr>
<tr>
<td>1</td>
<td>1.938</td>
<td>1611.5</td>
<td>1781.6</td>
</tr>
<tr>
<td>2</td>
<td>1.762</td>
<td>364.83</td>
<td>1045.48</td>
</tr>
<tr>
<td>3</td>
<td>1.45</td>
<td>1177</td>
<td>1280</td>
</tr>
<tr>
<td>4</td>
<td>1.364</td>
<td>964.2</td>
<td>1560.3</td>
</tr>
<tr>
<td>5</td>
<td>1.197</td>
<td>2074.4</td>
<td>2192.3</td>
</tr>
<tr>
<td>6</td>
<td>1.172</td>
<td>905.78</td>
<td>1648.24</td>
</tr>
<tr>
<td>7</td>
<td>1.16</td>
<td>1107.77</td>
<td>1518.64</td>
</tr>
<tr>
<td>8</td>
<td>0.939</td>
<td>1927.47</td>
<td>2211.06</td>
</tr>
<tr>
<td>9</td>
<td>0.928</td>
<td>882.8</td>
<td>1088.82</td>
</tr>
<tr>
<td>10</td>
<td>0.912</td>
<td>1765.89</td>
<td>2076.64</td>
</tr>
<tr>
<td>11</td>
<td>0.656</td>
<td>2120.63</td>
<td>2510.09</td>
</tr>
<tr>
<td>12</td>
<td>0.273</td>
<td>1378</td>
<td>2234.89</td>
</tr>
</tbody>
</table>
Figure 15 Encroachment result comparison
Figure 15 shows that the manual solution is very similar to HEC-RAS solution, and the differences get more obvious towards the upstream end. This is because the manual solution is adopting HEC-RAS Method 4 result and doing the modification upon it from downstream toward upstream. Apparently, GA encroachment station results have some differences from the other two, but it is still in a reasonable range and follows along the channel quite well.

**Best fitness achieved comparison**

<table>
<thead>
<tr>
<th>CS No.</th>
<th>HECRAS</th>
<th>Manual</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prof Delta (Ft)</td>
<td>Delta WSE (Ft)</td>
<td>Prof Delta (Ft)</td>
</tr>
<tr>
<td>1</td>
<td>0.94</td>
<td>0.06</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-0.22</td>
<td>1.22</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.01</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1.01</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1.03</td>
<td>0.03</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1.01</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1.02</td>
<td>0.02</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.99</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.99</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.99</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fitness</td>
<td>1.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1s</td>
<td></td>
<td>5hr</td>
</tr>
</tbody>
</table>

Table 8 is the fitness of the result from HEC-RAS calculation, HEC-RAS manual modification, and GA simulation. Among all these three methods, the HEC-RAS’ initial calculation is the fastest, and it only requires a second to do the calculation.
The result is reasonable, but it is not as accurate as desired. Manual modification of the HEC-RAS result improved the best result by achieving a specific 1.00 Ft surcharge for every cross-section, and this method took 5 hours of a professional labor. But the GA program took 2 minutes to achieve a result that is 3 times worse than HEC-RAS’s result.

These comparisons reveal that the HEC-RAS method has the advantage to get at least a closest result in a shortest time. The manual solution is more precise but takes time. And the GA method couldn’t provide a good solution compare to HEC-RAS but the time it took is satisfying. However, the GA method has the greatest potential to be modified and improved to achieve a better goal. Thus, the GA process method still has chances to catch up and surpass HEC-RAS method.
CHAPTER 5

OBSERVATIONS

Further Program Improvement

The results of the previous tests are not as good as the author desired. HEC-RAS could provide a result that is better than this GA program provided, and the manual solution’s technique could provide a good solution. Thus, some improvements of the program should be tested with a two cross-section model in the future study. These inspirations about build population, cross over, mutation, checking system and bisection method are obtained from the manual solution to pursue a more reasonable and closer result.

Build population

Instead of building the population by randomly selecting each gene in every decision variable of each individual, the initial population could be built based upon the initial HEC-RAS computation. After assigning the default parameters and importing the project file to the GA code and HEC-RAS, an initial computation could be conducted by HEC-RAS with Method 4 with a target water
surface difference of 1.0 to find the initial floodway stations. Then duplicate this individual to the quantity of individuals in the same generation.

This modification has several advantages over the conventional randomly built population process.

- First, it gives the program a better start level that equals to the best result that could be achieved in the first attempt that has already considered equal loss of conveyance.
- Second, in defining floodway stations, there could be many solutions rather than one, so having a good start could lead to a much more reasonable result and could save time from the first a few generations which might not convert to a solution as good as HEC-RAS computation.
- Third, this is much more alike with the manual solution to find the floodway stations, which uses Method 4 to get the initial floodway station, and changes the floodway station piece by piece using Method 1 to find a better solution. Even though there is not much gene variation at first, this situation could be solved during the mutation process to generate a variety of gene strings of the individuals that indicate a much better control of the evolution.

**Crossover**

To improve the crossover power and generate more variety of new binary strings, the two-point crossover could be used. This modification could give the medium
variety population a better chance to get new genes in lack of mutation. And in combination with a modified, non-uniform mutation discussed below, it could provide more combinations within a smaller range of mutation.

**Mutation**

The conventional single-point mutation is to randomly select the gene from randomly selected individuals to mutate by switching the binary value of the gene for this individual. The result of this mutation method has limited efficiency.

Considering the bisectional theory in manual method, the mutation process could be modified toward the following ways:

1. In the manual method, the cross-sections are modified from the downstream to upstream, however, in the GA method all the decision variables are changed by the same random amount for the selected individuals. In reality, the most cases could be only a few of the cross-sections need to be modified, but not all. So the change could be in two ways:
   a. Modify one cross-section after another following bisection method discussed in “*Bisection method for encroachment adjustment*” section on page 39. To fulfill this modification, the basic GA process could be changed, and this will be illustrated in detail in the “*Bisection method*” section on page 90.
b. Modify selected cross-sections at once by mutate a portion of the
decision variables. For example, within each individual, assign a
random number between 0 and 1 to every cross-section. If the
number is greater than 0.5, it will be mutated; and if it is smaller
than 0.5, it will not be mutated.

2. In the current GA process, the mutation is randomly adding or subtracting
from the left overbank or the right overbank. There is a big chance for the
encroachments of a specific cross-section to move towards the same
direction. With this mutation method, even the mutations take place, there
might be very little impact on the actual surcharge. For example, for a
cross section, if the current left encroachment station is 100 Ft, and right
encroachment station is 200 Ft, and both of them are mutating by adding
15 Ft. The new encroachment stations are changed to 115 Ft and 215 Ft.
The surcharge could be remaining no improvement or less improvement
towards the target surcharge than with an encroachment station of 85 Ft
and 215 Ft, or 115 Ft and 185 Ft. Since every left encroachment station
has an odd decision variable index number, thus, this modification could
be done by assign different signs to decision variable n and n+1.

3. The two-point mutation and the real-number mutation are more powerful
than the single-point mutation.

4. Mutation positions could be specified within a range for more effective
simulations. For example, a 10 digits binary string has the maximum value
of 1024, if the overbank width is 100 Ft, a desired change is 1.00 Ft, then
any change occurred within the last three digits of the binary is useless because \( 2^3 = 8 \leq 1024/100 = 10.24 \). Additionally, from the experience of manually solving the problem using HEC-RAS, if a change of the decision variable is less than 1, there could be no observed change in fitness. Following this clue, more research should be done to reveal the effective zones for different mutation situations.

5. After several generations, a smaller range of mutation targeted, a larger chance to improve the result could be gotten. For instance, let the unknown target be 2.3, a best estimation is 2.5, and a mutation with a value change range from 0 to 1 is much better than 0 to 100. Besides, to better estimate the mutation value, an equal sub range could be defined to equally distribute the estimations. This is the idea of modified, non-uniform mutation. A modified, uniform mutation adds or subtracts a randomly selected number in a certain range. Additionally, both the selected range and the not-ranked individuals are divided into some number of sections.

\[ \text{IndividualSectionNumber} = \text{RangeSection Number} + 1. \]

For example, a population of 100 individuals could be divided into 5 groups, and a change range of 50 could be divided into 4 sections. The first 20 individuals could randomly add or subtract a number within 0 to 12.5. The second 20 individuals could mutate by adding or subtracting a number within 12.5 to 25, etc. The last group of individuals remains unchanged. This is a real number modification to the decision variables.
that have been converted to real numbers. Thus, after the mutation, the new decision variables should be converted to binary again.

Table 9 Directed mutation

<table>
<thead>
<tr>
<th>Individual Group</th>
<th>Individuals</th>
<th>Mutation Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1~20</td>
<td>[0, 12.5)</td>
</tr>
<tr>
<td>2</td>
<td>21~40</td>
<td>[12.5, 25)</td>
</tr>
<tr>
<td>3</td>
<td>41~60</td>
<td>[25, 37.5)</td>
</tr>
<tr>
<td>4</td>
<td>61~80</td>
<td>[37.5, 50]</td>
</tr>
<tr>
<td>5</td>
<td>81~100</td>
<td>No Mutation</td>
</tr>
</tbody>
</table>

This modification could result in a more targeted mutation that distributes evenly within a specified range. A randomly generated number could be far away from the preferred result.

For further generations, the mutation range could downsize as the result is smaller and the decision variable variance should be downsized to a smaller random number. For example, the first 20% of generations could have a mutation range of 50, and 20%~40% of generations could have a total range of 25, and 40%~60% of generations could have a total range of 12.5, and 60%~80% of generations have a total range of 6.25, the 80%~100% of generations could have a total range of 3. This range and match should be found out depending on laboratory tests.

**Checking system**

The checking system could be a choice for targeted mutation to check if any specific cross-section contributes to the fitness more than any other cross-section, and then do the mutation on this cross-section.
Recording System

The recording system is to record each GA project operation results and keep it as an individual of the next project operation, so that the highest quality individual could be kept alive and improved upon for several executions.

Bisection method

In the manual solution, it gets the best result by using bisection method for each cross-section from downstream to upstream. This is because bisection method could provide accurate roots for any resoluble problem; and even if the problem is irresoluble, it could provide an accurate extreme value.

Processes:

1. Get HEC-RAS initial encroachment stations result.
2. From downstream cross-section, define boundary arrays. They are left over bank boundary (LOBB) and right over bank boundary (ROBB). Each boundary has a left boundary, midpoint, and right boundary. The boundaries are expressed by their stations.
3. Define individual fitness (IFitness)= Surcharge-Target Surcharge.
4. Conduct HEC-RAS simulation, and find left boundary’s IFitness, midpoint’s IFitness, and right boundary’s IFitness for each over bank. For the first generation, the midpoints are adopted from HEC-RAS result. The two over banks in the cross-section could be computed one by one or at the same time, as long as the termination criteria are achieved.
5. Compare the signs of each overbank’s IFitnesses. Set the new boundaries by replacing the boundary that has the same sign with the midpoint by the midpoint, and calculate a new midpoint.

6. If the midpoint’s IFitness is larger than an acceptable value of error, repeat process 1 to 5 with the new boundaries obtained in process 5 until the midpoint’s IFitness is smaller than the acceptable error.

7. When midpoint’s IFitness is less than an acceptable value, the encroachment stations for this cross-section are found, and move to the up stream cross-section and repeat process 1 to 7.

In this process, the decision variables are the stations of the midpoints of each boundary for each over bank.

For the example cross-section shown in Figure 16. Assume that the HEC-RAS initial encroachment stations are 360 Ft and 740 Ft. And assume the IFitness for this encroachment station is -0.25, the IFitness for encroachment stations shown in black is -0.30, and the IFitness for encroachment stations shown in green is 0.30. The midpoint of red station and green station for left over bank is 260 Ft, and the midpoint of red station and green station for right over bank is 840 Ft. Assume the calculated IFitness for this encroachment stations shown in red is 0.25, the new boundary is consist by the red stations and blue stations. The midpoints are 310 Ft and 790 Ft shown in purple. After calculation, it’s IFitness is 0, and the encroachment station for this cross-section is 310 Ft and 790 Ft.
Figure 16 Example of the bisectional method applied in a cross-section

Figure 17 Example of bisection method solution
CHAPTER 6

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This paper has introduced a useful technique for defining the encroachment stations for a regular channel, Genetic Algorithms program. A simple modified GA process is developed which is consisting of randomly building the initial population, fitness test, reversed fitness proportionate selection, one-point crossover and single-point mutation. And in combination with HEC-RAS, the GA program is developed to optimize the best result of the encroachment.

Experiments using synthetic data were performed to compare HEC-RAS calculation, HEC-RAS manual modification, and the GA optimization. Currently, this GA program could not provide a good solution compare HEC-RAS provided, however, it is automating the process and in good progress. To improve the program, some further modification should be investigated in the suggested order:

1. Trying to use the bisection method process described on page 90;
2. Directing the mutation process by applying it to selected cross-sections more effectively on a targeted position zone;
3. Using HEC-RAS Method 4 result to generate initial population;
4. Using more powerful crossover and mutation methods;
5. Adding a recording system to keep the best individual of the current execution to the next one;

6. Adding a checking system to identify the largest fitness contributor.

After encoding these modifications into the program in above order, compare results to identify if it has any improvement in further study.
REFERENCES

   NOAA.GOV. 26 Jun 2015


6. USGS. Floods And Flood Plains. 11 Jan 2013. 28 6 2015


11.—. Conditional letter of map revision. 26 04 2015. 01 05 2015

12.—. Letter of map revision. 26 04 2015. 01 05 2015


15.—. Flood map service center. 6 Jan 2005. 20 Jun 2015


17. MathWorks. Genetic algorithm terminology. 4 7 2015


22. Advantages and disadvantages of genetic algorithms. 2015 July 7

23. FEMA. UNIT2: The National flood insurance program.


APPENDIX-1

Conveyance Calculation Example:

Known: Left overbank roughness, $N_{LOB} = 0.040$

Main channel roughness, $N_{MC} = 0.030$

Right overbank roughness, $N_{ROB} = 0.040$
Water surface elevation = 709.9 Ft

<table>
<thead>
<tr>
<th>Ground Points</th>
<th>Station (Ft)</th>
<th>Elevation (Ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150</td>
<td>730</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>695</td>
</tr>
<tr>
<td>3</td>
<td>400</td>
<td>695</td>
</tr>
<tr>
<td>4</td>
<td>500</td>
<td>670</td>
</tr>
<tr>
<td>5</td>
<td>600</td>
<td>670</td>
</tr>
<tr>
<td>6</td>
<td>700</td>
<td>695</td>
</tr>
<tr>
<td>7</td>
<td>850</td>
<td>695</td>
</tr>
<tr>
<td>8</td>
<td>950</td>
<td>730</td>
</tr>
</tbody>
</table>

Solve:

1. Calculate the water surface intersection with ground geometry using linear interpolation. Note that $X_1$ refers to the left intersection of ground geometry and water surface, $X_2$ refers to the right intersection of ground geometry and the water surface.

\[
\frac{730 - 695}{150 - 250} = \frac{709.9 - 695}{X_1 - 250} \]

$X_1 = 207.43\, ft$

\[
\frac{730 - 695}{950 - 850} = \frac{709.9 - 695}{X_2 - 850} \]

$X_2 = 892.57\, ft$
2. Calculate the conveyances of the left overbank, the main channel and the right overbank associated with a water surface elevation of 709.9 Ft. Note that areas are computed using an equation for a trapezoid, which works for triangular and rectangular shapes as well.

Area of left overbank, $A_{LOB} = A_1 + A_2$

$$A_1 = \frac{1}{2} \left| [(WSE - y_1) + (WSE - y_2)](x_2 - x_1) \right|$$

$$= \frac{1}{2} \left| [(709.9 - 709.9) + (709.9 - 695)](250 - 207.4) \right| = 317.146 \text{ ft}^2$$

Area of left overbank, $A_{LOB} = A_1 + A_2 = 2,552.147 \text{ ft}^2$

Area of main channel, $A_{MCH} = A_3 + A_4 + A_5 = 9,470 \text{ ft}^2$

Area of right overbank, $A_{ROB} = A_6 + A_7 = 2,552.147 \text{ ft}^2$

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>A (Ft$^2$)</th>
<th>$A_{LOB}$ (Ft$^2$)</th>
<th>$A_{MCH}$ (Ft$^2$)</th>
<th>$A_{ROB}$ (Ft$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>207.43</td>
<td>709.9</td>
<td>317.146</td>
<td>317.146</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>695</td>
<td>2,235.000</td>
<td>2,235.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>695</td>
<td>2,740.000</td>
<td>2,740.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>670</td>
<td>3,990.000</td>
<td>3,990.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>670</td>
<td>2,740.000</td>
<td>2,740.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>695</td>
<td>2,235.000</td>
<td>2,235.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>850</td>
<td>695</td>
<td>317.147</td>
<td>317.147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>892.57</td>
<td>709.9</td>
<td>Sum</td>
<td>2,552.147</td>
<td>9,470.000</td>
<td>2,552.147</td>
</tr>
</tbody>
</table>

Wetted perimeter of Left overbank, $P_{LOB} = p_1 + p_2 + p_3 + \ldots + p_n$. Wetted perimeter calculations make use of Pythagorean’s Theorem.
\[ P_1 = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2} = [(250 - 207.43)^2 + (695 - 709.9)^2]^{1/2} = 45.102 \text{ft} \]

Wetted perimeter of left overbank, \( P_{LOB} = P_1 + P_2 = 195.102 \text{ft} \)

Wetted perimeter of main channel, \( P_{MCH} = P_3 + P_4 + P_5 = 306.155 \text{ft} \)

Wetted perimeter of right overbank, \( P_{ROB} = P_6 + P_7 = 195.102 \text{ft} \)

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>P (Ft)</th>
<th>( P_{LOB} ) (Ft)</th>
<th>( P_{MCH} ) (Ft)</th>
<th>( P_{ROB} ) (Ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>207.43</td>
<td>709.9</td>
<td>45.102</td>
<td>45.102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>695</td>
<td>150.000</td>
<td>150.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>695</td>
<td>103.078</td>
<td>103.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>670</td>
<td>100.000</td>
<td>100.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>670</td>
<td>103.078</td>
<td>103.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>695</td>
<td>150.000</td>
<td></td>
<td>150.000</td>
<td></td>
</tr>
<tr>
<td>850</td>
<td>695</td>
<td>45.102</td>
<td></td>
<td></td>
<td>45.102</td>
</tr>
<tr>
<td>892.57</td>
<td>709.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>195.102</td>
<td>306.155</td>
<td>195.102</td>
<td></td>
</tr>
</tbody>
</table>

Hydraulic Radius of Left overbank, \( R_{LOB} = \frac{A_{LOB}}{P_{LOB}} = \frac{2,552.147}{195.102} = 13.081 \text{ft} \)

Hydraulic Radius of Main channel, \( R_{MCH} = \frac{A_{MCH}}{P_{MCH}} = 30.932 \text{ft} \)

Hydraulic Radius of Right overbank, \( R_{ROB} = \frac{A_{ROB}}{P_{ROB}} = 13.081 \text{ft} \)

Conveyance of Left overbank, \( K_{LOB} = \frac{1.486}{N_{LOB}} A_{LOB} R_{LOB}^{2/3} = 526,372.593 \text{cfs} \)

Conveyance of Main channel, \( K_{MC} = \frac{1.486}{N_{MC}} A_{MC} R_{MC}^{2/3} = 4,622,245.582 \text{cfs} \)
Conveyance of Right overbank, \( K_{ROB} = \frac{1.486}{N_{ROB}} A_{ROB} R_{ROB}^{2/3} = 526,372.593 \, \text{cfs} \)

3. Obtain total conveyance

Total conveyance of a cross-section,

\[ K_{TOT} = K_{LOB} + K_{ROB} + K_{MCH} = 5,674,990.767 \, \text{cfs} \]

<table>
<thead>
<tr>
<th>Table 13 Hydraulic radius and conveyance calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>LOB</td>
</tr>
<tr>
<td>MC</td>
</tr>
<tr>
<td>ROB</td>
</tr>
<tr>
<td>Sum</td>
</tr>
</tbody>
</table>

The hypothetical floodway boundaries are assumed vertical and frictionless.
Example of manually calculating floodway stations for a cross-section:

**Known:** Adopt the same example data and result from previous conveyance calculation example, and continue with the calculation from Step 4.

4. For a specified surcharge, calculate the new conveyance of the three sub-channels at the same discharge.
Let surcharge=1.00 Ft

New water surface elevation, $WSE_{new} = WSE_0 + 1.00 = 710.9\text{ft}$

5. Calculate the water surface intersection with the ground points using linear interpolation.

\[
\frac{730 - 695}{150 - 250} = \frac{710.9 - 695}{X_1 - 250}
\]

\[X_1 = 204.57\text{ft}\]

\[
\frac{730 - 695}{950 - 850} = \frac{709.9 - 695}{X_2 - 850}
\]

\[X_2 = 895.43\text{ft}\]

6. Using the same sectional geometry, calculate the area, wetted perimeter, hydraulic radius and conveyance at the higher water surface elevation.

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$A$ (Ft$^2$)</th>
<th>$A_{LOB}$ (Ft$^2$)</th>
<th>$A_{MCH}$ (Ft$^2$)</th>
<th>$A_{ROB}$ (Ft$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>204.57</td>
<td>710.9</td>
<td>361.169</td>
<td>361.169</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>695</td>
<td>2,385.000</td>
<td>2,385.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>695</td>
<td>2,840.000</td>
<td></td>
<td>2,840.000</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>670</td>
<td>4,090.000</td>
<td></td>
<td>4,090.000</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>670</td>
<td>2,840.000</td>
<td></td>
<td>2,840.000</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>695</td>
<td>2,385.000</td>
<td></td>
<td></td>
<td>2,385.000</td>
</tr>
<tr>
<td>850</td>
<td>695</td>
<td>361.168</td>
<td></td>
<td>361.168</td>
<td></td>
</tr>
<tr>
<td>895.43</td>
<td>710.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td>2,746.169</td>
<td>9,770.000</td>
<td>2,746.169</td>
</tr>
</tbody>
</table>
Table 15 Wetted perimeter calculation

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>P (Ft)</th>
<th>P_{LOB} (Ft)</th>
<th>P_{MCH} (Ft)</th>
<th>P_{ROB} (Ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>204.57</td>
<td>710.9</td>
<td>48.132</td>
<td>48.132</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>695</td>
<td>150.000</td>
<td>150.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>695</td>
<td>103.078</td>
<td>103.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>670</td>
<td>100.000</td>
<td>100.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>670</td>
<td>103.078</td>
<td>103.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>695</td>
<td>150.000</td>
<td>150.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>850</td>
<td>695</td>
<td>48.132</td>
<td>48.132</td>
<td></td>
<td></td>
</tr>
<tr>
<td>895.43</td>
<td>709.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

`Sum` | 198.13 | 306.155 | 198.13 |

Table 16 Hydraulic radius and conveyance calculation

<table>
<thead>
<tr>
<th>N</th>
<th>A (Ft$^2$)</th>
<th>P (Ft)</th>
<th>R (Ft)</th>
<th>K (Cfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOB</td>
<td>0.04</td>
<td>2,746.169</td>
<td>198.132</td>
<td>13.860</td>
</tr>
<tr>
<td>MC</td>
<td>0.03</td>
<td>9,770.000</td>
<td>306.155</td>
<td>31.912</td>
</tr>
<tr>
<td>ROB</td>
<td>0.04</td>
<td>2,746.169</td>
<td>198.132</td>
<td>13.860</td>
</tr>
</tbody>
</table>

`Sum` | 6,046,188.805 |

Calculate the difference between the previous and the new total conveyance.

$\Delta K = K_{TOT,NEW} - K_{TOT,0} = 6046188.805\text{cfs} - 5674990.767\text{cfs} = 371198.038\text{cfs}$

7. Find the percent reduction of conveyance by dividing the conveyance difference by $K_{TOT,0}$ and insure that both overbanks see an equal loss of conveyance.

$\text{%Reduction} = \frac{371,198.038}{6,046,188.805} = 0.0614 = 6.14%$

Conveyance reduction of left overbank,
$\Delta K_{LOB} = 0.5 \times \Delta K = 185,599.019 \text{cfs} < 588,664.196 \text{cfs}$

Conveyance reduction of right overbank,

$\Delta K_{ROB} = 0.5 \times \Delta K = 185,599.019 \text{cfs} < 588,664.196 \text{cfs}$

8. Find the left encroachment using a second order curve fit. To do so, select three station points ($X$) and calculate the remaining conveyance in the overbank. For example, select $X_1, X_2, X_3$

Let the left encroachment station $X_1 = 204.57 \text{ft}$, all left over bank conveyance is the remaining conveyance, $K_{LOB,R1}=K_{LOB} = 588,664.196 \text{cfs}$

If the left encroachment station $X_2 = 250 \text{ft}$, the first ground point is $(250 \text{ft}, 695 \text{ft})$
\[ A_{LOB,NEW} = \frac{1}{2} \left| ((WSE - y_1) + (WSE - y_2))(X_2 - X_1) \right| \]
\[ = \frac{1}{2} \left| ((710.9 - 695) + (710.9 - 695))(400 - 250) \right| = 2385 ft^2 \]

\[ P_{LOB,NEW} = (250 - 400)^2 + (695 - 695)^2)^{0.5} + (710.9 - 695) = 165.90 ft \]

\[ R_{LOB,NEW} = 14.38 ft \]

\[ K_{LOB,NEW} = 523851.91 cfs \]

The remaining conveyance, \[ K_{LOB,R2} = 523,851.91 cfs \]

\[ X_3 = X_{LOB} = 400, \text{ the remaining conveyance, } K_{LOB,R3} = 0 \]

**Table 17 Left over bank quadratic curve fitting data**

<table>
<thead>
<tr>
<th>Station, X (Ft)</th>
<th>Remaining Conveyance, K (Cfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>204.57</td>
<td>588,664.2</td>
</tr>
<tr>
<td>250</td>
<td>523,851.91</td>
</tr>
<tr>
<td>400</td>
<td>0</td>
</tr>
</tbody>
</table>
9. Curve fit the data set into a trend line to represent the relationship.

\[ y = -10.57x^2 + 3378.2x - 339933 \]

Where:  
- \( X \) – Encroachment station 
- \( K \) – Available conveyance in the overbank at encroachment station \( X \)

10. Find the left encroachment station

\[ \Delta K_{LOB} = 185,599.019 \text{ cfs} \]

The remaining conveyance,
\[ K_L = K_{LOB,NEW} - \Delta K_L = 588,664.196 \text{cfs} - 185,599.019 \text{cfs} = 403,065.177 \text{cfs}, \]

\[ 403,065.177 = -10.57005X^2 + 3,378.18786X + 339,933.20707 \]

\[ X_L = 19.931\text{ft} \text{ (exceed the boundary), or } X_L = 299.668\text{ft} \text{ (acceptable)}, \]

Check when \( X_L = 299.668\text{ft}, Y_2 = 695\text{ft} \)

\[ A_{LOB,\text{CHECK}} = 1,595.28\text{ft}^2 \]

\[ P_{LOB,NEW} = 100.33 + (710.9 - 695) = 116.23\text{ft} \]

\[ R_{LOB,\text{CHECK}} = 13.72 \]

\[ K_{LOB,\text{CHECK}} = 339,731.64\text{cfs} \]

\[ K_{LOB,\text{GOAL}} = 588,664.196\text{cfs} - 185,599.019 \text{cfs} = 403,065.117 \text{cfs} \]

By checking, it is clear that the \( K_{LOB,\text{CHECK}} = 339,731.64\text{cfs} \) is not equal to \( K_{LOB,\text{GOAL}} = 403,065.117\text{cfs} \). This is because the curve fitting method is imprecise.

The real curve formed by real conveyance-station relationship could be very different to the 3-point quadratic curve fitting.

\[ \Delta K_{ROB} = 0.5 \times \Delta K_{TOT} + (K_{LOB,GOAL} - K_{LOB,CHECK}) \]

\[ = 185,599.019\text{cfs} + (403,065.117\text{cfs} - 339,731.64\text{cfs}) \]

\[ = 248,932.496\text{cfs} \]
11. Find the right encroachment using a second order curve fit, and select three station points (X) to calculate remaining conveyance. For example, select $X_1, X_2, X_3$

$$X_1 = X_{ROB} = 700,$$ and the remaining conveyance, $K_{ROB,R1} = 0$

$$X_2 = 850 \text{ ft}, \ Y_2 = 695 \text{ ft}$$

$$A_{ROB,NEW} = 2385 \text{ ft}^2$$

$$P_{ROB,NEW} = 150 \text{ ft} + (710.9 - 695) \text{ ft} = 165.9 \text{ ft}$$

$$R_{ROB,NEW} = 14.38 \text{ ft}$$

$$K_{ROB,NEW} = 523,851.91 \text{ cfs}$$

The remaining conveyance, $K_{ROB,R2} = 523,851.91 \text{ cfs}$

$$X_1 = 892.57,$$ the remaining conveyance, $K_{LOB,R1} = K_{LOB} = 556,111.37 \text{ cfs}$

The remaining conveyance $K_{ROB} = 556,111.37 - 248,932.496 = 307,178.874 \text{ cfs}$

<table>
<thead>
<tr>
<th>Station (Ft)</th>
<th>Remaining Conveyance (Cfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>700</td>
<td>0.00</td>
</tr>
<tr>
<td>850</td>
<td>523,851.91</td>
</tr>
<tr>
<td>892.57</td>
<td>556,111.37</td>
</tr>
</tbody>
</table>

12. Curve fit the data set into a trend line to represent the relationship.
Figure 22 Right over bank curve fitting

\[ K = -14.20028X^2 + 25,502.78141X - 10,893,809.36349 \]

Use the remaining conveyance of the right overbank in a quadratic curve fit to calculate the right encroachment station, \( K_{ROB} = 307,178.874\, cfs \)

\[ 307,178.874 = -14.20028X^2 + 25,502.78141X - 10,893,809.36349 \]

\( X_R = 1,030.479\, ft \) (exceed the boundary), or \( X_R = 765.456\, ft \) (acceptable).

Check when \( X_R = 765.456\, ft, Y_2 = 695\, ft \)

\[ A_{ROB, CHECK} = 1,040.75\, ft^2 \]
\[ P_{\text{ROB,NEW}} = 65.46 + (710.9 - 695) = 81.356 ft \]

\[ R_{\text{ROB,CHECK}} = 12.79 \]

\[ K_{\text{ROB,CHECK}} = 211,483.62 \text{ cfs} \]

Again, by checking, it is obvious that the \( K_{\text{ROB,CHECK}} = 211,483.62 \text{ cfs} \) is not equal to \( K_{\text{ROB,GOAL}} = 307,178.874 \text{ cfs} \). Thus, this widely used technique could not provide an accurate result and requires a huge amount of calculation workload.

From this calculation, the results are:

1. Left over bank encroachment=299.67 Ft
2. Right over bank encroachment=765.46 Ft

If Excel Solver is used instead of quadratic curve fitting, by setting the expected left over bank conveyance to be \( K_{\text{LOB,GOAL}} \) and changing the starting ground point station along the geometric cross-section, the left encroachment station is 282.55 Ft. And follow the same routine by setting the expected right over bank conveyance to be \( K_{\text{ROB,GOAL}} \) and changing the ending ground point station along the geometric cross-section, the left encroachment station is 817.55 Ft. This is the same result obtained using HEC-RAS.