Water Supply

IWA PUBLISHING



Water Supply Vol 22 No 5, 5221 doi: 10.2166/ws.2022.172

# Water surface profile in converging compound channel using gene expression programming

Bandita Naik [a,*], Vijay Kaushik [b] and Munendra Kumar [b]

[a] Department of Civil Engineering, Methodist College of Engineering, Hyderabad, India
[b] Department of Civil Engineering, Delhi Technological University, Delhi, India
*Corresponding author. E-mail: banditanaik@methodist.edu.in

BN, 0000-0002-9488-2184

## ABSTRACT

Assessment of water surface profile in compound channels is essential for flood defence systems. Agriculture and development activities in floodplains affect the floodplain shape over the length, leading in a converging compound channel. Few laboratory investigations proved overbank flow in converging floodplains. Therefore, innovative and precise approaches are still in great demand. In this paper, new approach has been proposed to forecast the water surface profile of various compound channels with converging floodplains using gene expression programming (GEP). The models are constructed utilising pertinent experimental data from past studies. A new equation is devised to compute water surface profile in such channels using non-dimensional geometric and flow parameters such as converging angle, width ratio, relative distance, relative depth, aspect ratio and bed slope. The findings demonstrate that the GEP-derived water surface profile is in good correlation with the experimental data and data from other studies ($R^2 = 0.99$ and RMSE $= 0.028$ for the training data and $R^2 = 0.99$ and RMSE $= 0.027$ for the testing data). According to the results of statistically based investigations, the GEP model created for the study of compound channel flow is reliable and can be used in this domain.

**Key words:** converging compound channel, error analysis, flow parameters, gene expression programming, water surface profile

## HIGHLIGHTS

- In this paper, new approach has been proposed to forecast the water surface profile of various compound channels with converging floodplains using gene expression programming (GEP).
- The computed water surface profile using gene expression programming is found to be more accurate when compared to previous methods.

## NOTATION

The following symbols are used in this paper:

| | |
|---|---|
| B | total width of compound channel |
| b | total width of the main channel |
| H | bank full depth |
| h | total height of the main channel |
| L | converging length |
| S | longitudinal bed slope |
| $\Psi$ | non-dimensional water surface profile (H/h) |
| $X_r$ | relative distance(x/L) |
| x | distance between two consecutive sections |
| $\alpha$ | width ratio (B/b) |
| $\beta$ | relative flow depth [(H–h)/H] |
| $\delta$ | aspect ratio (b/h) |
| $\theta$ | converging angle |

## INTRODUCTION

Rivers are vital to human civilizations since people rely on them for daily needs. Larger communities have formed on the waterway floodplains as a result of the continued usage of rivers, causing the floodplain to converge. As a result of the incorrect calculation of the water surface profile during the flood, there were more life loss and economic damages. Natural compound rivers are classified as non-prismatic compound rivers since their floodplains vary. The flow of natural rivers is continually changing from uniform to non-uniform since they are non-prismatic in nature. It's difficult to simulate flow in a non-prismatic compound channel. The flow models of straight and meandering prismatic compound channels have been extensively studied by a number of researchers, but very little is known about non-prismatic compound channels (Sellin 1964; Myers & Elsawy 1975; Knight *et al.* 2010; Khatua *et al.* 2012). The flow on spreading floodplains accelerates, whereas the flow on converging floodplains diminishes due to channel shape convergence (James & Brown 1977). The studies were carried out by Bousmar & Zech (2002), Bousmar *et al.* (2004), Rezaei (2006) and Rezaei & Knight (2009) on compound channels with symmetrically narrowing floodplains. In these experiments, the geometrical momentum transfer and the associated additional head loss were emphasised. Proust *et al.* (2006) investigated asymmetric geometry with a greater convergence rate. More mass transfer and total head loss observed with the higher convergence angle (22°). Chlebek *et al.* (2010) investigated and analysed the flow behaviour of various compound channels, such as skewed channels, symmetrically converging channels, and diverging channels. Rezaei & Knight (2011) and Hojjat *et al.* (2013) recently carried out a unique experiment on the converging compound channels that yielded substantially more accurate data sets than previously accessible. All of the studies cited above, have focused on the effect of changes in floodplain sections to evaluate the discharge. In non-prismatic compound channels, the impact of geometry and flow conditions on water surface profile has not been thoroughly explored. As a consequence, precise water surface profile modeling is required to identify flooded areas, which will improve flood mitigation and risk management studies. The impact of floodplain width contraction on water depth prediction in a compound channel is currently being investigated in order to determine the relationship between many geometrical and hydraulic factors in predicting the water surface profile of a converging compound channel. In the present work, an effort was made to create a trustworthy mathematical model to anticipate the water surface profile of a compound channel with converging floodplains for different converging angles based on the experimental data of Naik & Khatua (2016a, 2016b) and Rezaei (2006).

Developing a water surface profile model using mathematical, analytic, or numerical approaches is quite challenging to analyze the connections between dependent and independent factors. In addition, these models become significantly clumsy and time-consuming. Therefore, a simple technique like Gene Expression Programming (GEP) may estimate water surface profile. As a consequence, not only is experimental time reduced by half, but laborious calculations are also reduced. Due to the increasing reliance on machine learning algorithms to estimate flow in compound channels, these channels are increasingly calculated using support vector machines (SVM), gene expression programming (GEP), artificial neural networks, fuzzy neural networks, and the M5 tree decision models (Seckin 2004; Unal *et al.* 2010; Sahu *et al.* 2011; Zahiri & Azamathulla 2014; Najafzadeh & Zahiri 2015; Parsaie *et al.* 2017). The advantage of GEP over other soft computing techniques is that it generates simpler equations without assuming a previous form of the existing relation. GEP is a multigene, one-of-a-kind coding language that enables the modification of more complicated equations that are divided into many sub-equations. It also uses gene generations, fitness-based gene selection, and the introduction of genetic diversity through the use of one or more genetic operators. The GEP's capacity to generate mathematical correlations distinguishes it from other soft computing approaches such as ANN and SVM (Cousin & Savic 1997; Drecourt 1999; Savic *et al.* 1999; Whigham & Crapper 1999, 2001; Babovic & Keijzer 2002; Karimi *et al.* 2015). However, river engineering using the GEP method has received far less attention (Harris *et al.* 2003; Giustolisi 2004; Guven & Gunal 2008; Guven & Aytek 2009; Azamathulla *et al.* 2013; Pradhan & Khatua 2017b). For the estimation of discharge in diverging and converging compound channels, an equation has been created by Das *et al.* (2019). This equation encourages the usage of GEP. Mohanta & Patra (2021) have developed a model equation for calculating discharge in meandering compound channels, validating the use of GEP over the classic channel division technique. An evolution process known as the GEP technique was used to create mathematical expressions, decision trees, polynomial structures, and logical expressions for the current investigation. The authors attempt to provide a novel equation for water surface profile in compound channels with convergent floodplains. The method considers variables such as the width ratio ($\alpha$), relative depth ($\beta$), converging angle ($\theta$), aspect ratio($\delta$), relative distance ($X_r$) and bed slope (S). The effectiveness of the proposed equation and its performance are assessed using previous models.

## EXPERIMENTAL SETUP

A series of experiments were conducted in a concrete flume containing three converging compound channels. A Perspex sheet of 15 m long, 0.90 m wide, and 0.5 m deep was used to make non-prismatic part of compound channel. The channel's width ratio was 1.8, and its aspect ratio was 5. The channel's converging angles were 12.38°, 9°, and 5°, respectively, keeping the geometry constant. The non-prismatic compound channel has converging lengths of 0.84, 1.26, and 2.28 m, respectively. Longitudinal bed slope of the channel was 0.0011; it satisfied subcritical flow conditions at all sections of the non-prismatic compound channels. The roughness of the floodplain and main channel was maintained identical, and the Manning's n was determined as 0.011 from the experimental runs in the channel. The flowing portion of the experiment was turbulent. This system recirculates the water supply by pumping it from an underground sump to a reservoir located in the experimental channel. The rectangular notch has been surrounded by adjustable vertical gates and flows strengthens. The flume's removable tailgates help maintain a consistent flow over the test reach. A volumetric tank collects the water that flows from the channel. It then goes back to the underground sump. The geometric characteristics and a cross-section of a compound channel are shown in Figure 1.

Figure 2 illustrates the experimental setup from the top. The plan view of several tests with cross-sectional dimensions of Naik & Khatua (2016a, 2016b) and Rezaei (2006) channels is shown in Figure 3. As part of the compound channel design, each point on the channel's plan could be accessed for measuring. A moveable bridge could be used to collect the measurements. The width ratio and aspect ratio of the channel are the essential parameters in the study. A micro-pitot tube with a diameter of 4.77 mm was used to measure the flow grid's velocity. The order of maximum velocity for a given flow path was determined using a flow detector with a minimum least count of 0.1°. Use of circular scale and pointer configuration on flow direction sensor to measure pitot tube leg angle in relation to channel longitudinal direction. When combining the longitudinal velocity plot with the volumetric tank collection, the total discharge computed was within ±3% of the actual data. This study used velocity data and a semi-log plot to predict channel bed and wall shear stresses. The boundary shear stresses were calculated using Patel's (1965) relationship and manometer measurements of Preston tube head differences. Shear values were corrected by comparing them to the equivalent values calculated using the energy gradient

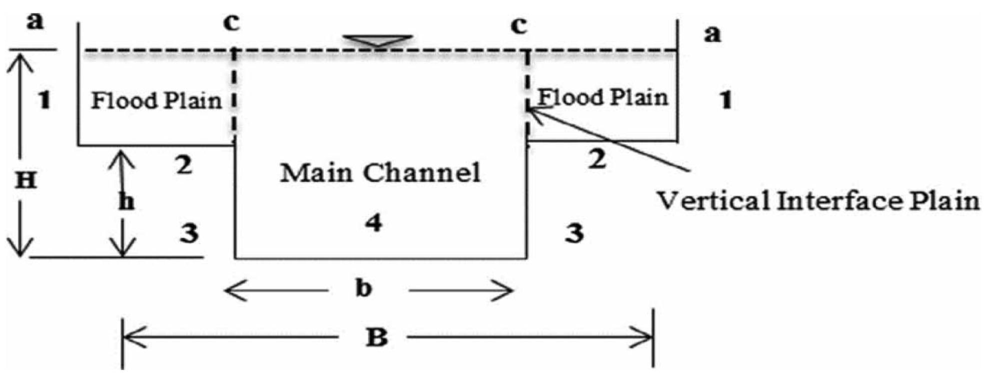


**Figure 1** | Cross-section of a compound channel.

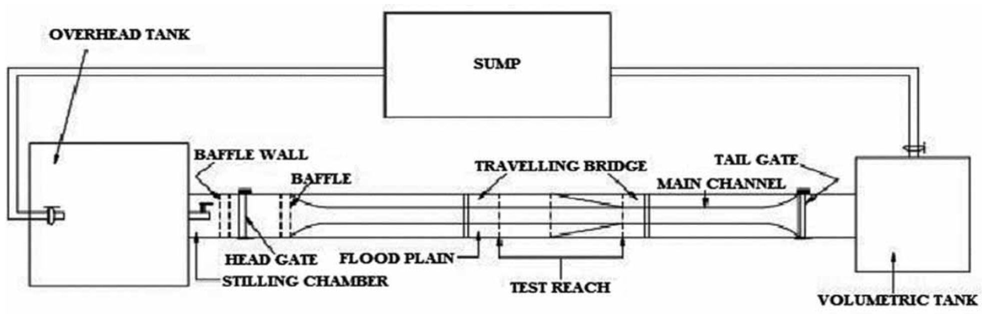


**Figure 2** | Experimental flume setup.

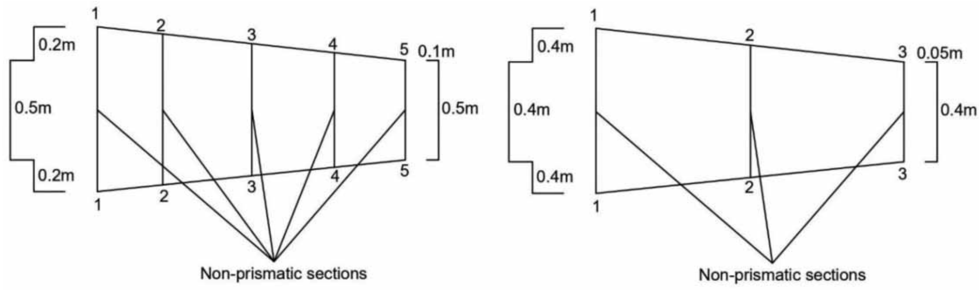


**Figure 3** | Cross-sectional dimensions of both Naik & Khatua (2016a, 2016b) & Rezaei (2006) channels.

technique. As a result, the findings were constantly within $\pm 3\%$ of the value. According to lab data analysis, pitot tube shear stress computation is superior to ADV in terms of accuracy. For one thing, measuring velocity at the boundary with ADV is never trustworthy. In addition, ADV has certain limits when it comes to measuring the velocity near the bed. It can penetrate up to 5 cm below the top edge. Consequently, the micro-ADV down probe could not reach a distance of 5 cm from the free surface. In order to measure the transient decrease, a pitot tube was used. Verification of the validity of this approach was carried out using the energy gradient methodology (i.e., weight component of the flow) (Naik & Khatua 2016a, 2016b).

## THEORETICAL BACKGROUND

Accurate prediction of water surface profile in overbank flows is important for river engineers in the design of drainage channels, flood defence schemes, river training works, floodplain management, etc. In natural rivers, providing sufficiently precise and comprehensive field data under unstable flood flow conditions is difficult. To increase understanding of the water surface profile in compound channels with prismatic and non-prismatic floodplains, laboratory experiments are required. Naik & Khatua (2016a, 2016b) proposed the following equations for the water surface profile in a compound channel with converging floodplains:

$$\Psi = 1.06\alpha^{0.22} \quad \text{for lower aspect ratio channel 1} \tag{1}$$

$$\Psi = 1.16\alpha^{0.22} \quad \text{for lower aspect ratio channel 2} \tag{2}$$

$$\Psi = 1.21\alpha^{0.29} \quad \text{for lower aspect ratio channel 3} \tag{3}$$

$$\Psi = 0.07\alpha + 1.78 \quad \text{for higher aspect ratio channel 1} \tag{4}$$

$$\Psi = 0.05\alpha + 1.28 \quad \text{for higher aspect ratio channel 2} \tag{5}$$

$$\Psi = 0.13\alpha + 1.25 \quad \text{for higher aspect ratio channel 3} \tag{6}$$

$$\Psi = -0.14X_r + 1.22 \quad \text{for lower aspect ratio channel 1} \tag{7}$$

$$\Psi = -0.15X_r + 1.32 \quad \text{for lower aspect ratio channel 2} \tag{8}$$

$$\Psi = -0.22X_r + 1.37 \quad \text{for lower aspect ratio channel 3} \tag{9}$$

$$\Psi = -0.15X_r + 2.01 \quad \text{for higher aspect ratio channel 1} \tag{10}$$

$$\Psi = -0.16X_r + 1.45 \quad \text{for higher aspect ratio channel 2} \tag{11}$$

$$\Psi = -0.21X_r + 1.67 \quad \text{for higher aspect ratio channel 3} \tag{12}$$

$$\Psi = -1.22 + 2.27\alpha^{0.22} + 0.18X_r \quad \text{for lower aspect ratio channel 1} \tag{13}$$

$$\Psi = -1.21 + 2.28\alpha^{0.22} + 0.19X_r \quad \text{for lower aspect ratio channel 2} \tag{14}$$

$$\Psi = -0.58 + 1.63\alpha^{0.29} + 0.18X_r \quad \text{for lower aspect ratio channel 3} \tag{15}$$

$$\Psi = -0.66 + 0.29\alpha + 0.12X_r \quad \text{for higher aspect ratio channel 1} \tag{16}$$

$$\Psi = 0.86 + 0.29\alpha + 0.11X_r \quad \text{for higher aspect ratio channel 2} \tag{17}$$

$$\Psi = 0.86 + 0.29\alpha + 0.12X_r \quad \text{for higher aspect ratio channel 3} \tag{18}$$

$$\Psi^*(\theta) = \frac{Actual\ \Psi}{Eq.\ (12)} \tag{19}$$

$$\frac{Actual\ \Psi}{Eq.\ (12)} = e^{0.0017\theta} \tag{20}$$

$$\Psi = e^{0.0017\theta}(-1.21 + 2.25\alpha^{0.22} + 0.18X_r) \tag{21}$$

Using gene expression programming, an effort was made to predict the water surface profile for the compound channel with distinct converging floodplains. The flow may be expected to be uniform until it reaches the prismatic region, but it is discovered to be non-uniform in the non-prismatic part. Three distinct kinds of converging compound channels of Naik & Khatua (2016a, 2016b), as well as three sets of Rezaei (2006) data, were used to create a non-dimensional water surface profile (details of the data sets are given in Table 1). The main channel and floodplain surfaces of all of these channels have been given a uniform roughness. For all of these smooth surfaces, Manning's n values are determined to be 0.01. To estimate the water surface profile of converging compound channels, most influencing factors such as width ratio ($\alpha$), relative depth ratio ($\beta$), aspect ratio ($\delta$), converging angle ($\theta$), relative distance ($X_r$) and longitudinal slope (S) are taken into account. The author uses a GEP approach with the help of GeneXproTools 5.0 (2014) to develop an equation for the non-dimensional water surface profile parameter ($\Psi$) in a converging compound channel, taking into account all aspects. The generation of models is chosen based on the fitness of the training and testing datasets. Using one or more genetic operators, such as mutation or recombination, chosen models are reproduced in GEP. Brief theoretical overviews of GEP are described in past studies (Mallick *et al.* 2020). A variety of parameters are selected to allow the resulting model equations to be applied to various compound channels. To simplify the equation for a straight compound channel, assume the converging angle as zero and simplify the equation for the straight compound channel.

**Table 1** | Details of parameters used in present analysis

| Verified test channel | Type of channel | Angle of convergent (θ) | Longitudinal slope (S) | Cross-sectional geometry | Total channel width (B) (m) | Main channel width (b) (m) | Main channel depth (h) (m) | Converging length ($X_r$) (m) | Aspect ratio (δ) |
|---|---|---|---|---|---|---|---|---|---|
| Channel 1 Rezaei (2006) | Converging | 11.31° | 0.002 | Rectangular | 1.2 | 0.398 | 0.05 | 2 | 7.96 |
| Channel 2 Rezaei (2006) | Converging | 3.81° | 0.002 | Rectangular | 1.2 | 0.398 | 0.05 | 6 | 7.96 |
| Channel 3 Rezaei (2006) | Converging | 1.91° | 0.002 | Rectangular | 1.2 | 0.398 | 0.05 | 6 | 7.96 |
| Channel 1 Naik & Khatua (2016a, 2016b) | Converging | 5° | 0.0011 | Rectangular | 0.9 | 0.5 | 0.1 | 2.28 | 5 |
| Channel 2 Naik & Khatua (2016a, 2016b) | Converging | 9° | 0.0011 | Rectangular | 0.9 | 0.5 | 0.1 | 1.26 | 5 |
| Channel 3 Naik & Khatua (2016a, 2016b) | Converging | 12.38° | 0.0011 | Rectangular | 0.9 | 0.5 | 0.1 | 0.84 | 5 |

The required dimensionless equation can be written as

$$\Psi = F(\alpha, \beta, \delta, \theta, X_r, S) \quad (22)$$

## Gene expression programming model for water surface profile

The relationship (Equation (22)) indicates the water surface profile of converging compound channels as a function of geometric and hydraulic variables. In this study, the modelling procedure uses $\Psi$ as the target value and the six independent factors ($\alpha$, $\beta$, $\delta$, $\theta$, $X_r$, S) as input variables discussed in Equation (22). The model is constructed using four fundamental arithmetic operators (+, −, ×, /). There are 318 data sets used and randomly distributed for the two different phases of the modeling process. For the current study, 50% of the data is used for training, while the other 50% is used for testing. In this study, RMSE was the fitness function (Ei), and the fitness (fi) was computed by Equation (23) that yields the target value's total sum of errors. Starting with just one gene and two head sizes, the first model was constructed. The genes and heads were then increased one by one throughout each run, and the outcomes of the training and testing datasets were recorded. For head lengths of more than eight and more than three genes, the performance of the training and testing data phase did not improve significantly.

Consequently, eight were chosen as head length for inclusion in the GEP model, with three genes per chromosome. In order to connect three genes, the addition was used as linking function. After 5,000 generations, the fitness function value and coefficient of determination of training and testing data had not changed, suggesting that generations may have come to an end. Table 2 summarises the major parameters that determine GEP modeling effectiveness and are used to build a model for forecasting water surface profile in compound channels with converging floodplains. The final model of the GEP was discovered to be an algebraic equation between output and input variables after a lot of trial and error. In this study, GeneXproTools (2014), a robust soft computing software package was used in the modeling process.

**Table 2** | Functional set and operational parameters used in GEP Model

| Description of parameter | Parameter setting |
|---|---|
| Function set | +, −, ×, / |
| Number of chromosomes | 30 |
| Head size | 8 |
| Number of genes | 3 |
| Gene size | 26 |
| Linking function | Addition |
| Fitness function | RMSE |
| Program size | 47 |
| Literals | 18 |
| Number of generations | 5,000 |
| Constants per gene | 10 |
| Data type | Floating-point |
| Mutation | 0.00138 |
| Inversion | 0.00546 |
| Gene recombination rate | 0.00277 |
| One-point recombination rate | 0.00277 |
| Two-point recombination rate | 0.00277 |
| Gene transposition rate | 0.00277 |
| Insertion sequence (IS) transposition rate | 0.00546 |
| Root insertion sequence (RIS) transposition rate | 0.00546 |

GEP is modeled in the following way in terms of analytical form as

$$\Psi = \left[\frac{3.72 + \ ((\delta + S) \times \beta)}{\left(\frac{\alpha}{-8.31}\right) + (3.72 - S)}\right] + \left[\frac{\left(\left(\frac{9.63}{\delta}\right) - \alpha\right)/(5.73 + \theta)}{(\delta \times \beta) + (X_r + X_r)}\right] + \left[\left(\left(\frac{\frac{4.72}{\beta}}{-0.32}\right) - \left(\frac{\alpha}{\beta}\right)\right) \times (\beta \times (S \times \delta))\right] \tag{23}$$

Further, Equation (23) is simplified as

$$\Psi = \left[\frac{3.72 + \delta\beta + S\beta}{3.72 - S - 0.12\alpha} + \frac{9.63 - \alpha\delta}{(5.73 + \theta) \times (\delta^2\beta + 2\delta X_r)} - 14.75\delta S - \alpha\delta S\right] \tag{24}$$

In Figure 4, the GEP model for water surface profile is represented by an expression tree (ET) representation as in Equation (24). Thus, in Figure 4, d0 denotes the $\alpha$, d1 denotes the $\beta$, d2 denotes the $\delta$, d3 denotes the $\theta$, d4 denotes the $X_r$, and d5 denotes S. G1c5 and G1c7 represent the numerical constants employed in the model's first gene. Similarly, G2c6, G2c9, and G3c1, G3c2 are the constants used in the second and third genes of the model.

## RESULTS AND DISCUSSION

Figure 5 depicts the water surface profile at various relative depths over a 9° converging floodplain angle. The depth of flow diminishes as the longitudinal distance rises, as seen in the figure. According to Naik & Khatua (2016a, 2016b), this drop is

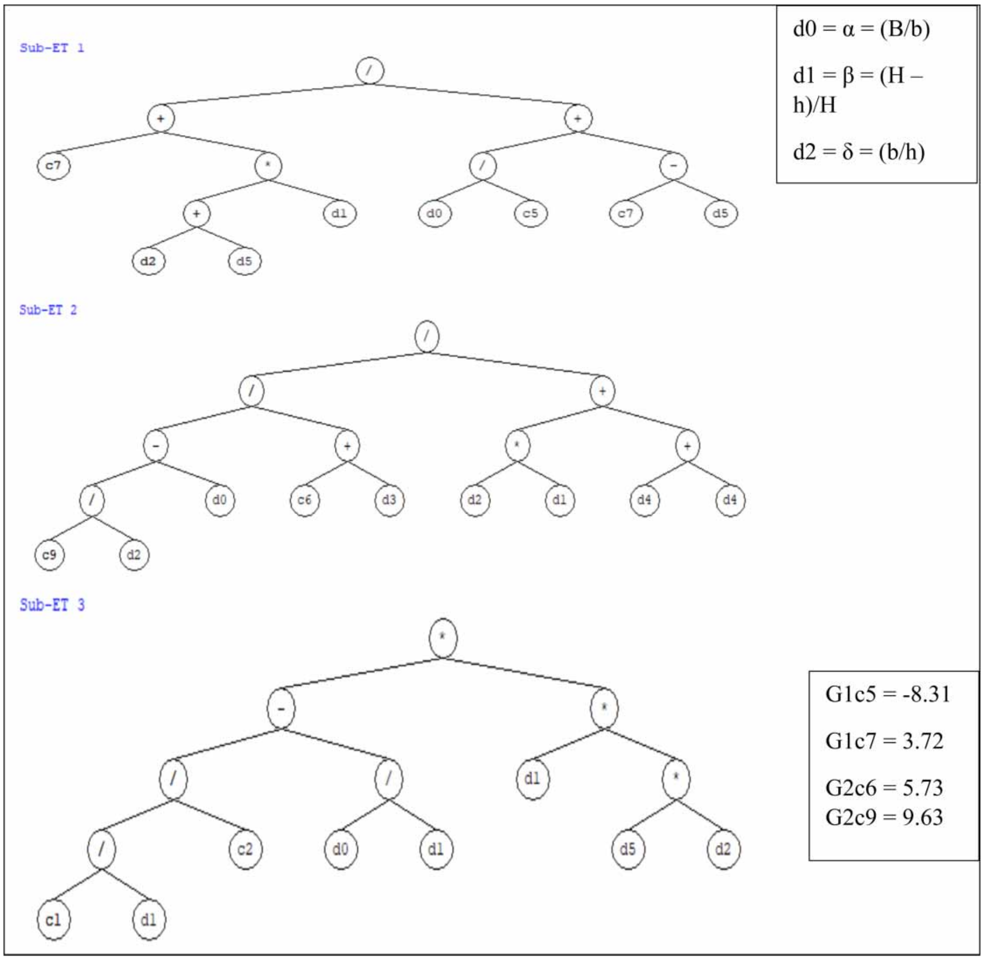


**Figure 4** | Expression tree (ET) for GEP formulation.

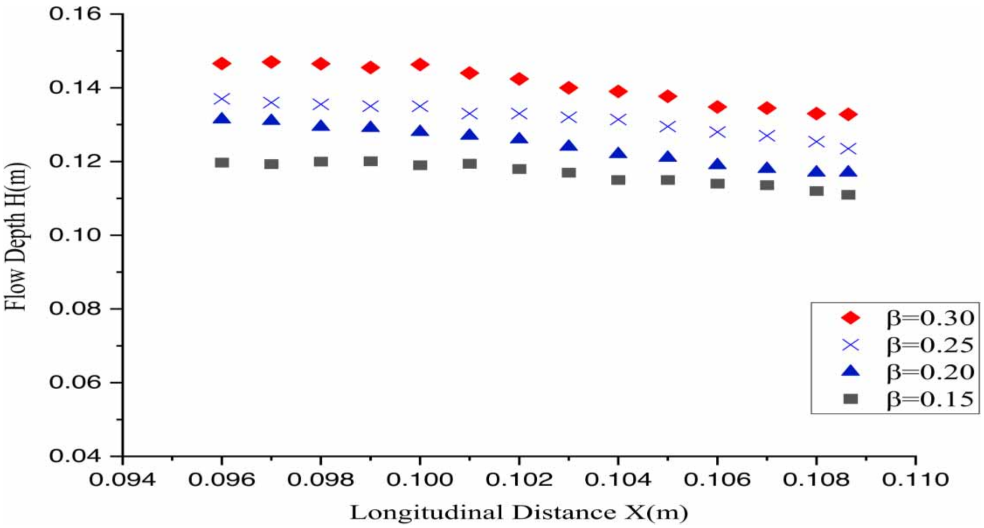


**Figure 5** | Water surface profile for compound channel at 9° floodplain converging angle for different relative flow depths.

more pronounced for lower converging angles than for greater converging angles. This is owing to the fact that the main channel's bottom corner and floodplain have the lowest velocity whereas maximum velocity occurs on central region of main channel. When different relative flow depths are used in conjunction with different converging angles, the fluctuation of the non-dimensional water surface profile with width ratio is shown in Figure 6. The water surface profile grows as the width ratio increases at all rising flow depths. The water surface profile follows the same pattern of fluctuation with width ratio at all converging angles. For varied converging angles, Figure 7 shows that the water surface profile rises non-linearly with relative flow depths. Figure 8 investigates the influence of relative distance on non-dimensional water surface profiles. The water surface profile is seen to decline as the relative distance $X_r$ rises, as shown in the figure. It demonstrates that a converging transition quickly raises the velocity head, lowers the potential head. The velocity contours presented by Naik & Khatua (2016a, 2016b) help to clarify this. For higher converging floodplain angles, the fall is exceedingly steep. For all converging compound channels, the fall trends are determined to be linear. Rezaei (2006) channels have a lower width ratio, resulting in a narrower floodplain than Naik & Khatua (2016a, 2016b) channels, resulting in smoother water surface profile fluctuations.

Figure 9 shows the comparison for the training and testing data phases of the GEP method for the actual and predicted $\Psi$. The authors used Equation (24) to compare the predicted value to the actual value for each data set used in the training and testing data phases independently, as shown in Figure 9. For converging compound channels, the GEP method reveals a highly non-linear relationship between $\Psi$ and the input parameters ($\alpha$, $\beta$, $\delta$, $\theta$, $X_r$, S) with high precision and relatively very low residuals, as shown in Figure 10. Figure 11 shows the performance matrices for the training and testing data phases of the GEP method for the actual and predicted $\Psi$ and it is observed that both the data are in high correlation with each other. For the training data ($R^2 = 0.99$ and RMSE = 0.028) and the testing data ($R^2 = 0.99$ and RMSE = 0.027), the GEP model produced the lowest errors. The GEP model can reflect the underlying phenomena of water surface profile in converging compound channel, as seen in Figure 9. It claims that although choosing an input vector based on linear measurements between the variables of interest is common, it may include certain misleading characteristics that add to the model's difficulty.

For further testing of the accuracy of the developed model by GEP approach, various types of error analysis, such as the coefficient of determination ($R^2$), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE), are analysed using the following equations (Mohanta *et al.* 2018)

$$R^2 = \frac{\sum_{i=1}^{N} (a_i - \bar{a})^2 \; (p_i - \bar{p})^2}{\sum_{i=1}^{N} (a_i - \bar{a})^2 \; \sum_{i=1}^{N} (p_i - \bar{p})^2} \tag{25}$$

$$\mathrm{MAE} = \frac{1}{N} \sum_{i=1}^{N} |p_i - a_i| \tag{26}$$

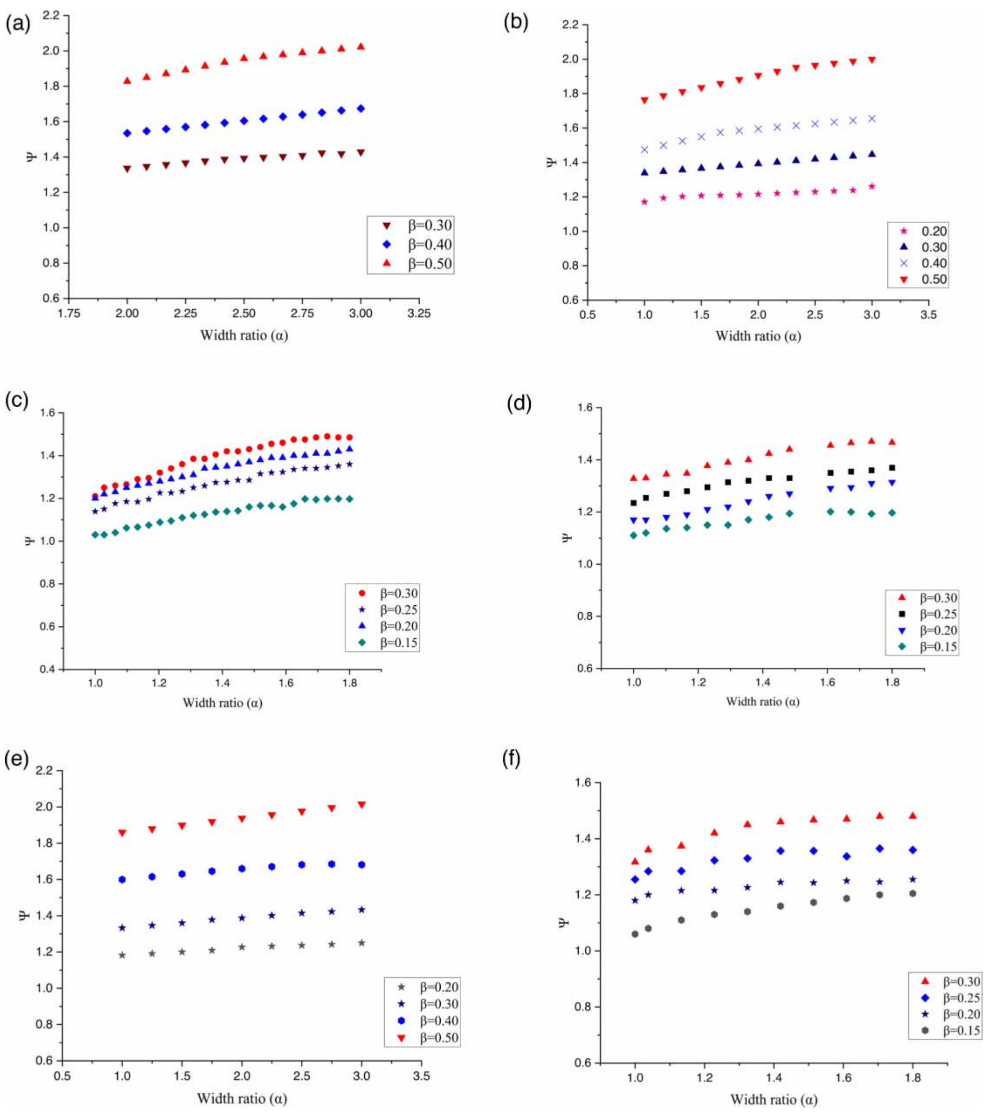


**Figure 6** | Variation of non-dimensional water surface profile with width ratio for different relative flow depths at various converging angles (a) 1.91° (b) 3.81° (c) 5° (d) 9° (e) 11.31° (f) 12.38°.

$$\mathrm{MAPE}(\%) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|p_i - a_i|}{a_i} \times 100\right) \tag{27}$$

$$\mathrm{RMSE} = \sqrt{1/N \ \textstyle\sum_{i=1}^{N} (p_i - a_i)^2} \tag{28}$$

where $a$ and $p$ are the actual and predicted values, respectively, $\bar{a}$ and $\bar{p}$ are the mean of actual and predicted values, respectively, and $N$ is the number of datasets.

For both the training and testing datasets, performance assessments are shown in Table 3 in terms of $R^2$ (coefficient of determination), correlation coefficient, mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE ) for the created GEP model were calculated. The model's adequacy is demonstrated by the $R^2$ value for the two datasets (training = 0.99 and testing = 0.99) in the current method. The RMSE values for the training and testing data sets were 0.028 and 0.027, respectively, according to the GEP model. MAE and MAPE values for training data were 0.022 and 1.543%, respectively, while MAE and MAPE values for testing data were 0.022 and 1.546%, indicating the predicted model's

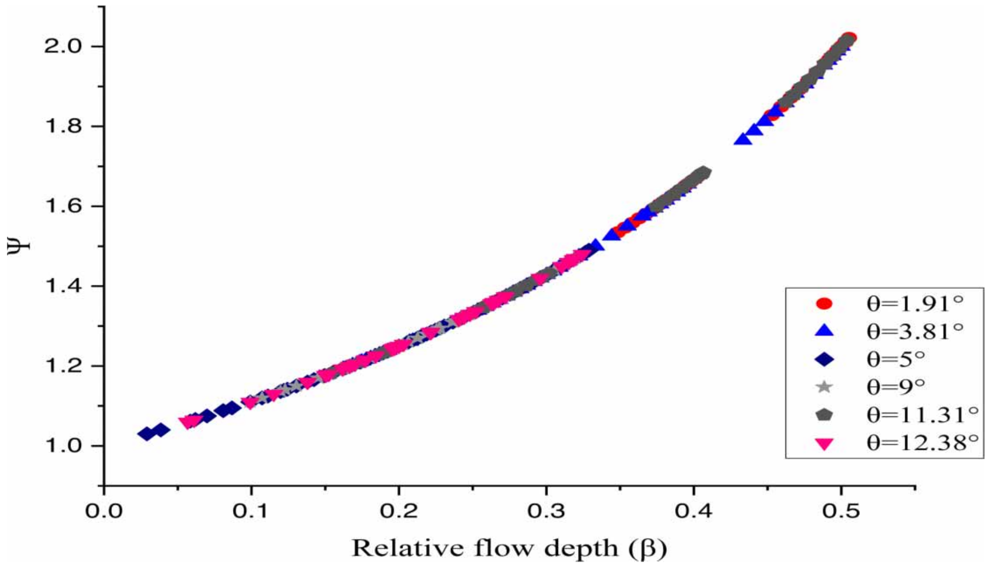


**Figure 7** | Variation of non-dimensional water surface profile with different relative flow depths for various converging angles.

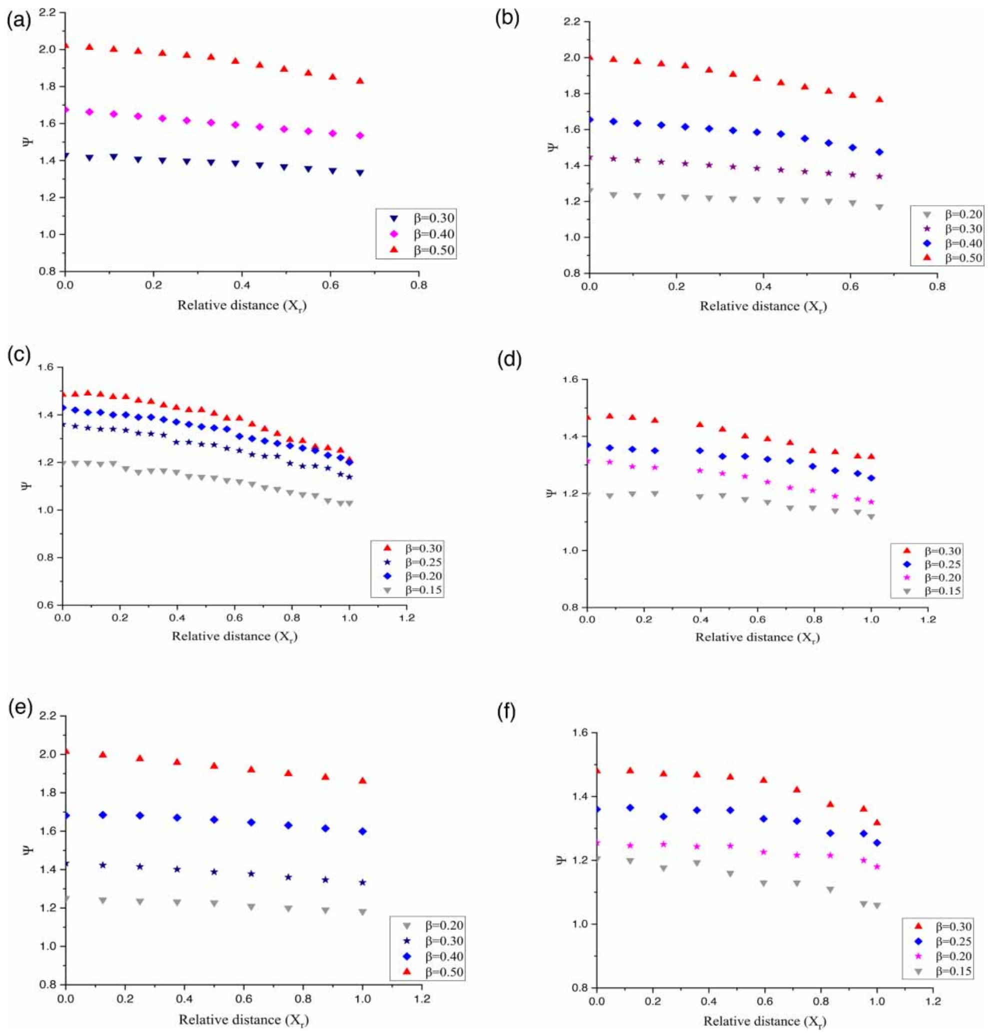


**Figure 8** | Variation of non-dimensional water surface profile with relative distance for different relative flow depths at various converging angles (a) 1.91° (b) 3.81° (c) 5° (d) 9° (e) 11.31° (f) 12.38°.

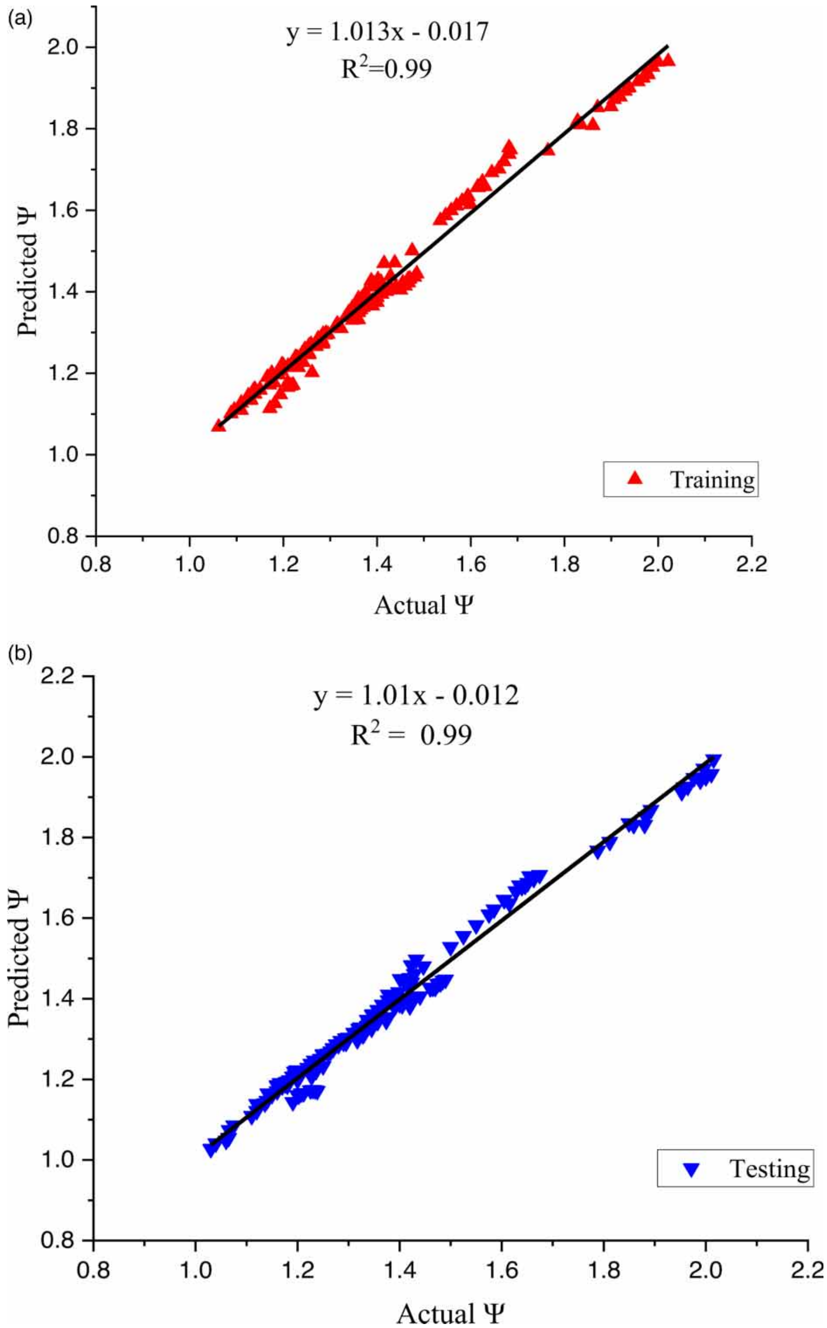


**Figure 9** | Comparison between actual and predicted $\Psi$ for the different data phase: (a) Training phase (b) Testing phase.

performance and accuracy. As a result, the proposed GEP is a reliable approach for predicting water surface profile in compound channels with a high generalization capacity and does not exhibit overtraining. The disparity between the anticipated non-dimensional water surface profile values for compound channels with converging floodplains and the actual values for all six kinds of channels is depicted in Figure 12. When comparing the current model to the prior model for Naik & Khatua (2016a, 2016b), the percentage error in the estimation of $\Psi$ is lower using the GEP approach shown in Figure 12. Table 4 compares statistical error analysis of present study and Naik & Khatua (2016a, 2016b) approaches, with the new GEP method appearing to be the best. The suggested GEP model has a substantial benefit over traditional regression-based models (standard equations). It can project data into a high-dimensional feature space, where many approaches may be utilized to identify relationships in the data. However, because the mapping is so broad, the connections are very general. To

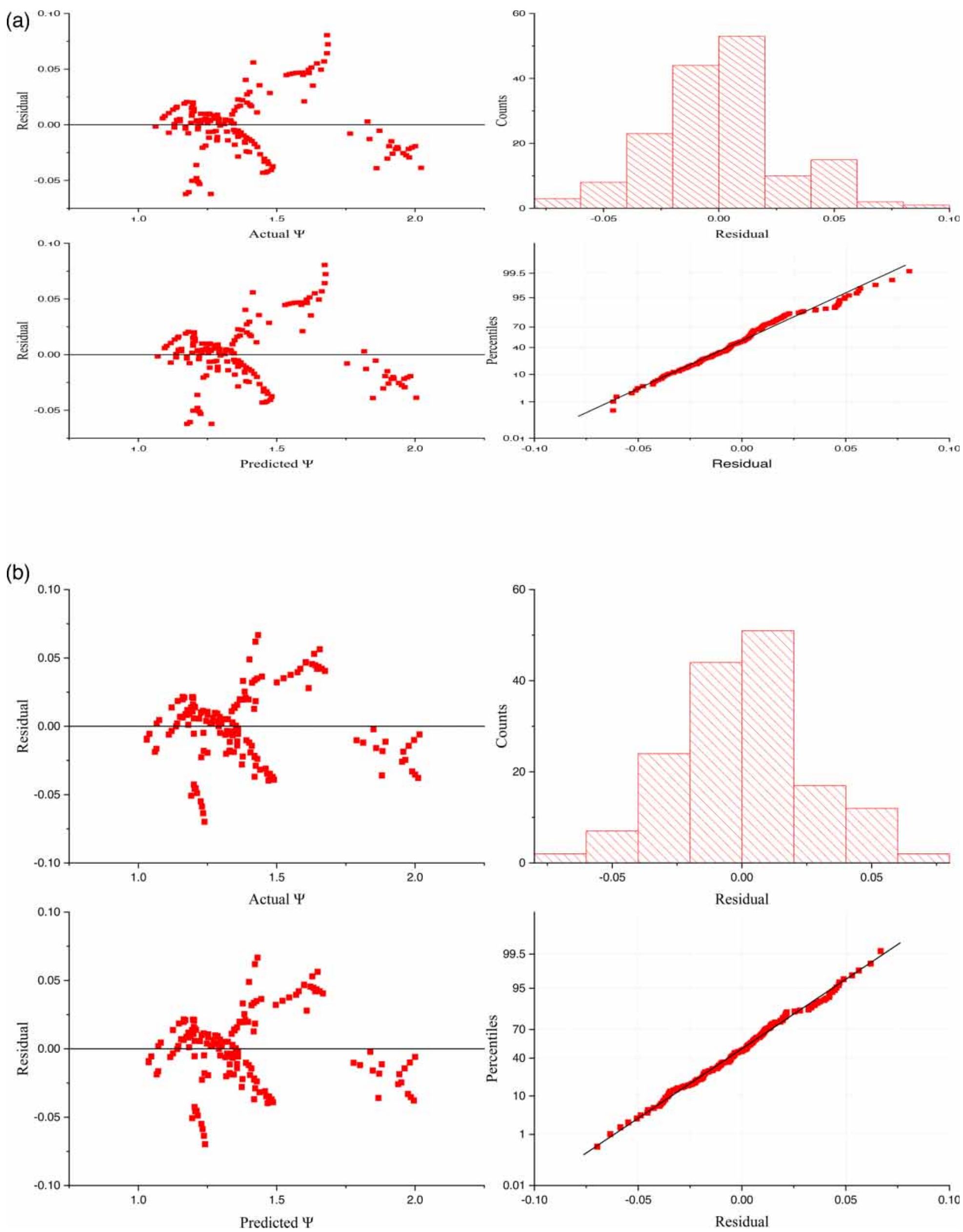

**Figure 10** | Residual plot for $\Psi$ for the different data phase: (a) Training phase (b) Testing phase.

decrease related uncertainties, physical-based, process-oriented riverine hydraulics research is required to develop techniques for predicting water surface profile, boundary shear stresses, in-channel and overbank roughness coefficients, and other variables impacting energy losses, velocity, and discharge. These issues have considerable influence on water resources and related studies, such as flood inundation studies, flood hazard warnings, hydraulic structure design and safety in channels and floodplains, sediment transport, geomorphic, and other studies that require reliable riverine hydraulic data.

## CONCLUSIONS

GEP is a computer programming language that leverages a fixed-length gene expression representation to encapsulate computer programs and rapidly discover succinct and understandable solutions. In this study, the GEP is utilized to estimate the water surface profile of converging compound channels. The GEP model can accurately forecast water surface profile with a

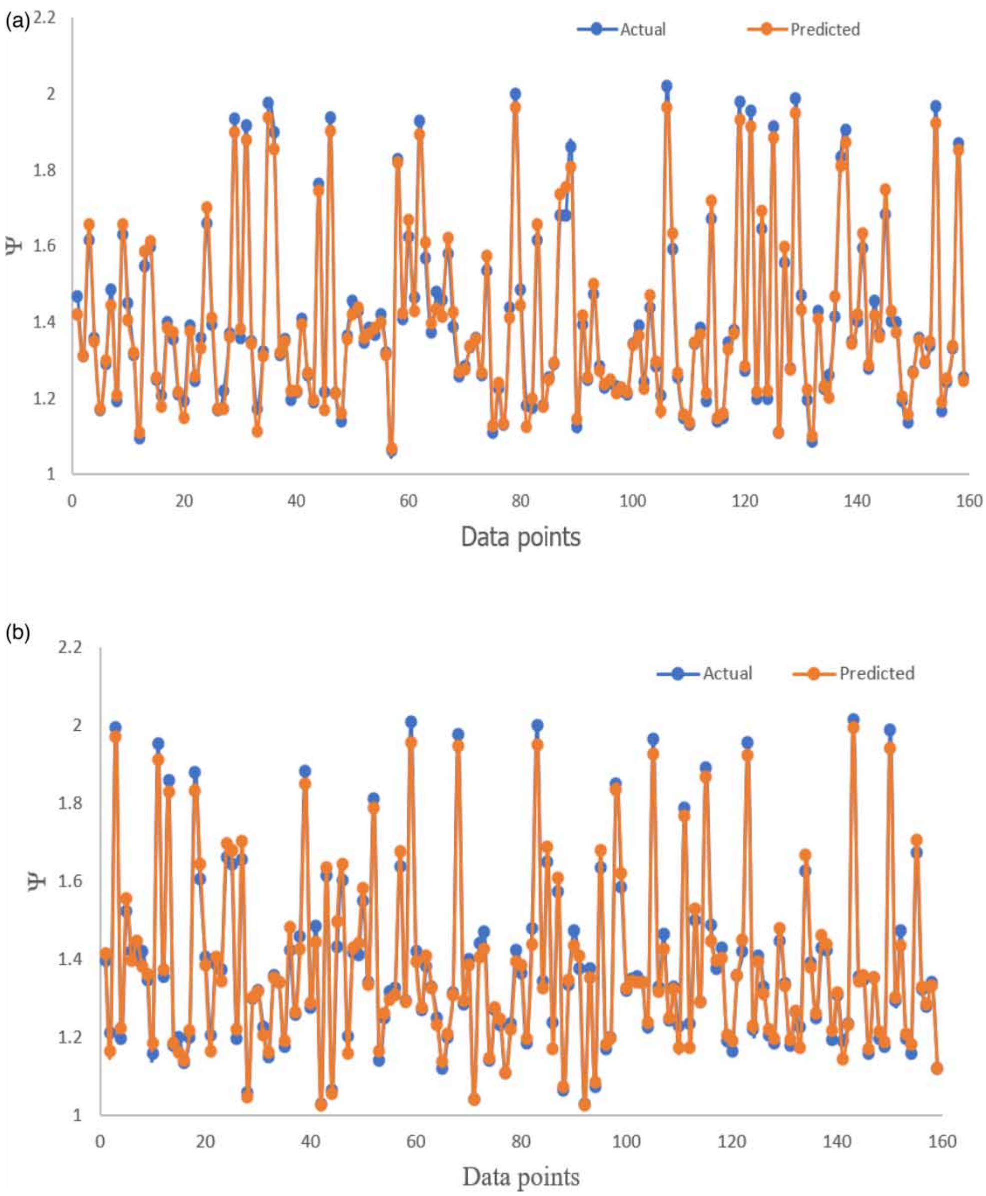


**Figure 11** | Performance matrices for the prediction of $\Psi$ for the different data phase: (a) Training phase (b) Testing phase.

**Table 3** | Performance evaluation of predicted $\Psi$ by GEP model for training and testing dataset

| Dataset | $R^2$ | Correlation coefficient | MAE | RMSE | MAPE(%) |
|---|---|---|---|---|---|
| Training | 0.99 | 0.99 | 0.022 | 0.028 | 1.543 |
| Testing | 0.99 | 0.99 | 0.022 | 0.027 | 1.546 |

short run time, which is dependent on the number of generations. The author conducts experimental investigations using a novel set of converging compound channels, including a converging compound channel with varied angles of converging floodplains ranging from 1.91° to 12.38° and varying width ratios ranging from 1.0 to 3.0. Using 318 high-quality data from experimental compound channels with converging floodplains, the GEP model was constructed to determine the exact values of the water surface profile. The following are the conclusions derived as a result of this study:

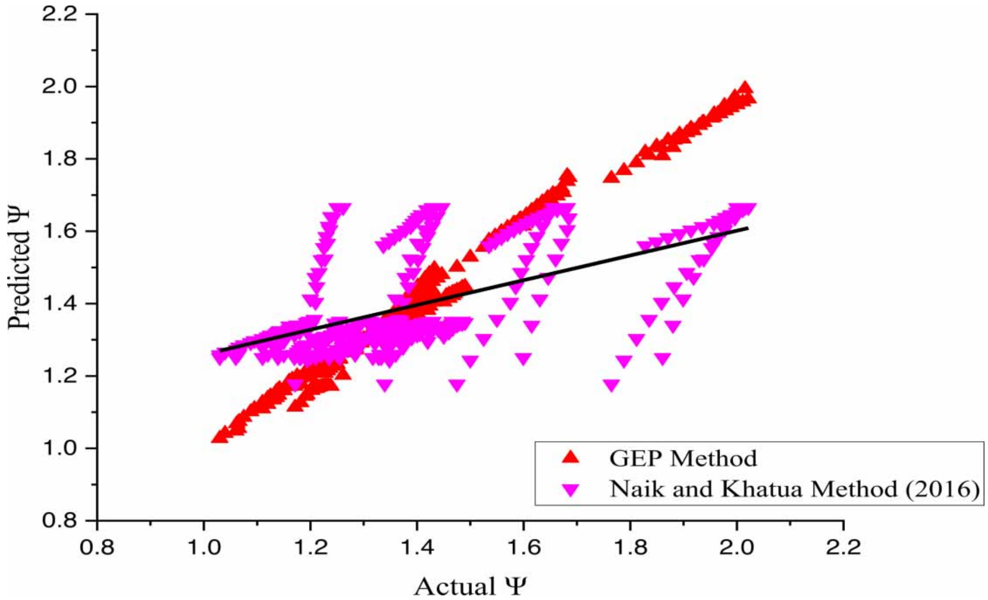


**Figure 12** | Comparison of predicted value of $\Psi$ for different models.

**Table 4** | Statistical error analysis of different methods

| Statistical parameters | New GEP Method Training | New GEP Method Testing | Naik and Khatua Method |
|---|---|---|---|
| $R^2$ | 0.99 | 0.99 | 0.896 |
| MSE | 0.0008 | 0.0007 | 0.0019 |
| RMSE | 0.028 | 0.027 | 0.043 |
| MAE | 0.022 | 0.022 | 0.0015 |
| MAPE (%) | 1.543 | 1.546 | 2.429 |

The proposed model appears to be influenced by parameters such as width ratio, relative flow depth, converging angle, relative distance, aspect ratio and longitudinal slope. With increasing width ratio and relative distance of converging compound channels, the non-dimensional water surface profile is seen to grow. Furthermore, with increasing overbank flow depth, the nondimensional water surface profile is seen to expand exponentially. For various converging angles, the non-dimensional water surface profile is seen to rise as the relative depth increases. The relationship between the most affecting non-dimensional geometric and hydraulic factors of a converging compound channel and the nondimensional water surface profile is investigated. It is discovered to have a nonlinear connection for all parameters. The GEP approach's suggested model is highly suitable for all of these kinds of channel systems, covering various laboratory models. In comparison to other flow parameters such as converging angle and relative distance, the relative flow depth and width ratio was found to be more appropriate in computing water surface profile. In contrast to previous methods such as Naik & Khatua (2016a, 2016b), the created GEP model shows better results in terms of $R^2$, MAE, RMSE, and MAPE for various datasets. The GEP approach's compatibility is determined by the model's mean percentage of error and the approaches' suitability. Within a close range of non-dimensional parameters examined in this study, the findings demonstrate the effectiveness of the GEP model and its potential utility for real applications. The results of this investigation show that the GEP model is more beneficial in any circumstance.

The limitation of the model is that it can be utilized to predict the water surface profile of compound channel with converging flood plain for homogeneous roughness only. The model can be improved with more data sets from wider flood plains and for differential roughness in the main channel and flood plains.

## ACKNOWLEDGEMENTS

The authors would like to convey their heartfelt gratitude to anonymous reviewers for their valuable time on this article.

## DISCLOSURE STATEMENT

The authors reported no potential conflicts of interest.

## FUNDING

Not applicable.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

Azamathulla, H. M., Ahmad, Z. & Ghani, A. A. 2013 An expert system for predicting Manning's roughness coefficient in open channels by using gene expression programming. *Neural Comput. Appl.* **23** (5), 1343–1349.

Babovic, V. & Keijzer, M. 2002 Rainfall runoff modelling based on genetic programming. *Hydrol. Res.* **33** (5), 331–346.

Bousmar, D. & Zech, Y. 2002 Periodical turbulent structures in compound channels. In *River Flow International Conference on Fluvial Hydraulics*, Louvain-la-Neuve, Belgium, pp. 177–185.

Bousmar, D., Wilkin, N., Jacquemart, J. H. & Zech, Y. 2004 Overbank flow in symmetrically narrowing floodplains. *J. Hydraul. Eng. ASCE* **130** (4), 305–312.

Chlebek, J., Bousmar, D., Knight, D. W. & Sterling, M. A. 2010 Comparison of overbank flow conditions in skewed and converging/diverging channels. In: *River Flows International Conference*, pp. 503–511.

Cousin, N. & Savic, D. A. 1997 *A Rainfall-Runoff Model Using Genetic Programming*. Centre for Systems and Control Engineering, Rep. No. 97, 3.

Das, B. S., Devi, K. & Khatua, K. K. 2019 Prediction of discharge in converging and diverging compound channel by gene expression programming. *J. Hydraul. Eng.* doi:10.1080/09715010.2018.1558116.

Drecourt, J. P. 1999 Application of neural networks and genetic programming to rainfall runoff modeling. *Water Resour. Manage.* **13** (3), 219–231.

Gepsoft, G. 2014 *Version 5.0*.

Giustolisi, O. 2004 Using genetic programming to determine Chezy resistance coefficient in corrugated channels. *J. Hydroinf.* **6** (3), 157.

Guven, A. & Aytek, A. 2009 New approach for stage-discharge relationship: gene-expression programming. *J. Hydrol. Eng.* **14** (8), 812–820.

Guven, A. & Gunal, M. 2008 Genetic programming approach for prediction of local scour downstream of hydraulic structures. *J. Irrig. Drain. Eng.* **134** (2), 241–249.

Harris, E. L., Babovic, V. & Falconer, R. A. 2003 Velocity predictions in compound channels with vegetated floodplains using genetic programming. *Int. J. River Basin Manage.* **1** (2), 117–123.

Hojjat, A. Y., Mohammad, H. O. & Seyed, A. A. 2013 The hydraulics of flow in non-prismatic compound channels. *J. Civ. Eng. Urbanism* **3** (6), 342–356.

James, M. & Brown, R. J. 1977 Geometric parameters that influence floodplain flow. In *U.S. Army Engineer Waterways Experimental Station*, June, Vicksburg Miss. Research report H-77.

Karimi, S., Shiri, J., Kisi, O. & Shiri, A. A. 2015 Short-term and long-term streamflow prediction by using 'wavelet-gene expression' programming approach. *ISH J. Hydraul. Eng.* **22** (2), 148–162.

Khatua, K. K., Patra, K. C. & Mohanty, P. K. 2012 Stage-discharge prediction for straight and smooth compound channels with wide floodplains. *J. Hydraul. Eng. ASCE* **138** (1), 93–99.

Knight, D. W., Tang, X., Sterling, M., Shiono, K. & McGahey, C. 2010 Solving open channel flow problems with a simple lateral distribution model. *River Flow* **1**, 41–48.

Mallick, M., Mohanta, A., Kumar, A. & Patra, K. C. 2020 Gene-expression programming for the assessment of surface mean pressure coefficient on building surfaces. *Build. Simul.* **13**, 401–418.

Mohanta, A. & Patra, K. C. 2021 Gene-expression programming for calculating discharge in meandering compound channels. *Sustainable Water Resour. Manage.* **7**, 33. https://doi.org/10.1007/s40899-021-00504-0.

Mohanta, A., Patra, K. C. & Sahoo, B. B. 2018 Anticipate Manning's coefficient in meandering compound channels. *Hydrology* **5** (3), 47.

Myers, W. R. C. & Elsawy, E. M. 1975 Boundary shears in channel with flood plain. *J. Hydraul. Div. ASCE* **101** (7), 933–946.

Naik, B. & Khatua, K. K. 2016a Water surface profile computation for compound channels with narrow flood plains. *Arabian J. Sci. Eng.* **42** (3), 941–955. doi:10.1007/s13369-016-2236-x.

Naik, B. & Khatua, K. K. 2016b Boundary shear stress distribution for a converging compound channel. *ISH J. Hydraul. Eng.* doi:10.1080/09715010.2016.1165633.

Najafzadeh, M. & Zahiri, A. 2015 Neuro-fuzzy GMDH-based evolutionary algorithms to predict flow discharge in straight compound channels. *J. Hydrol. Eng.* **20** (12), 04015035.
Parsaie, A., Yonesi, H. & Najafian, S. 2017 Prediction of flow discharge in compound open channels using adaptive neuro fuzzy inference system method. *Flow Meas. Instrum.* **54**, 288–297.
Patel, V. C. 1965 Calibration of the Preston tube and limitations on its use in pressure gradients. *J. Fluid Mech.* **231**, 85–208.
Pradhan, A. & Khatua, K. K. 2017b Gene expression programming to predict Manning's n in meandering flows. *Can. J. Civ. Eng.* **45** (4), 304–313.
Proust, S., Rivière, N., Bousmar, D., Paquier, A. & Zech, Y. 2006 Flow in the compound channel with abrupt floodplain contraction. *J. Hydraul. Eng.* **132** (9), 958–970.
Rezaei, B. 2006 *Overbank Flow in Compound Channels with Prismatic and non-Prismatic Floodplains*. Ph.D. Thesis, University of Birmingham, Birmingham, UK.
Rezaei, B. & Knight, D. W. 2009 Application of the Shiono and Knight Method in the compound channel with non-prismatic floodplains. *J. Hydraul. Res.* **47** (6), 716–726.
Rezaei, B. & Knight, D. W. 2011 Overbank flow in compound channels with non-prismatic floodplains. *J. Hydraul.* **137**, 815–824.
Sahu, M., Khatua, K. K. & Mahapatra, S. S. 2011 A neural network approach for prediction of discharge in straight compound open channel flow. *Flow Meas. Instrum.* **22** (5), 438–446.
Savic, D. A., Walters, G. A. & Davidson, J. W. 1999 A genetic programming approach to rainfall-runoff modelling. *Water Resour. Manage.* **13** (3), 219–231.
Seckin, G. 2004 A comparison of one-dimensional methods for estimating discharge capacity of straight compound channels. *Can. J. Civ. Eng.* **31** (4), 619–631.
Sellin, R. H. J. 1964 A laboratory investigation into the interaction between flow in the channel of a river and that of its flood plain. *LaHouille Blanche* **7**, 793–801.
Unal, B., Mamak, M., Seckin, G. & Cobaner, M. 2010 Comparison of an ANN approach with 1-D and 2-D methods for estimating discharge capacity of straight compound channels. *Adv. Eng. Software* **41** (2), 120–129.
Whigham, P. A. & Crapper, P. F. 1999 Time series modelling using genetic programming: an application to rainfall-runoff models. *Adv. Genet. Program.* **3**, 89–104.
Whigham, P. A. & Crapper, P. F. 2001 Modelling rainfall-runoff using genetic programming. *Math. Comput. Modell.* **33** (6–7), 707–721.
Zahiri, A. & Azamathulla, H. M. 2014 Comparison between linear genetic programming and M5 tree models to predict flow discharge in compound channels. *Neural Comput. Appl.* **24** (2), 413–420.

First received 9 December 2021; accepted in revised form 1 April 2022. Available online 18 April 2022