

Prediction of side thermal buoyant discharge in the cross flow using multi-objective evolutionary polynomial regression (EPR-MOGA)

Milad Khosravi and Mitra Javan

ABSTRACT

The capability to predict the distribution of pollutants in water bodies is one of the most important issues in the design of jet outfalls. Three-dimensional computational fluid dynamics (CFD) model and multi-objective evolutionary polynomial regression (EPR-MOGA) are used and compared in modeling the temperature field in the side thermal buoyant discharge in the cross flow. The input variables used for training the EPR-MOGA models are spatial coordinates (x , y , z), jet to cross flow velocity ratio (R), depth of the channel (d), and the temperature excess (T_0). A previous experimental study is used to verify and compare the performance of the EPR-MOGA and CFD models. The results show that the EPR-MOGA model predicts the thermal cross section of the flow and the spread of pollutants at the surface with a better accuracy than the CFD model. However, the CFD method performs significantly better than EPR-MOGA in predicting temperature profiles. The uncertainty analysis indicated that the EPR-MOGA model had lower mean prediction error and smaller uncertainty band than the CFD model. The relationships achieved by the EPR-MOGA model are very useful to predict temperature profiles, temperature half-thickness, and temperature spread on surface in practice.

Key words | buoyant discharge, computational fluid dynamics, evolutionary polynomial regression, temperature field

Milad Khosravi
Mitra Javan (corresponding author)
Department of Civil Engineering,
Razi University,
Kermanshah, 6714414971,
Iran
E-mail: javanmi@gmail.com

INTRODUCTION

The rapid pace of industrialization in countries has led to an increase in demand for electricity. Thermal power plants are one of the most popular sources of electrical energy in the world. These power plants produce large amounts of thermal waste, which mostly discharge into water bodies such as lakes, rivers, and seas (Kuhlman & Prahl 1976). Temperature is a key environmental variable, which has a great impact on many chemical and physical properties of water, including the ability to dissolve oxygen and other gases, the rate of chemical reaction, and microbial activity including the presence or absence of pathogens (Dallas 2009). Hence, regulations have been set in some countries to explain the environmental standards for the discharge of pollutants. The United States Environmental Protection

Agency (EPA) has determined a set of standards for industrial factories in relation to the maximum permissible temperature rise and maximum allowable surface isotherm areas (Kuhlman & Prahl 1976). The rules determined make industrial plants and factories utilize discharge systems with the least damage to the environment.

Surface discharge through side channels with free surface is a common method in power plants to discharge large volumes of thermal contaminants. The discharge system of open channels is much more cost-effective than submerged single or multiport diffuser systems, although the initial mixing is greater in diffusers (Fossati *et al.* 2011). Discharging hot water into water bodies through a side channel forms a complex phenomenon, which covers a

spatial range from the near field with initial mixing of the jet to the far field with mixing due to buoyancy and diffusion. A comprehensive classification for different flow regimes from a side channel has been conducted based on the geometry and momentum and buoyancy flow by Jones *et al.* (2007).

Numerous experimental studies have been carried out on the discharge of thermal contaminants from a side channel (Motz & Benedict 1970; Stolzenbach & Harleman 1971; Carter *et al.* 1973; Carter & Regier 1974). Hydraulics Research Wallingford (1985) studied the mixing behavior in the near field of buoyant surface jet discharge through a rectangular channel in a strong cross flow. Abessi *et al.* (2012) investigated the discharge of negatively buoyant effluent from a side channel shallower than the cross flow into a stationary homogeneous receiver fluid. They analyzed the flow mixing behavior using digital video recordings and determined the plume trajectories by the data obtained from the probe. Stefan *et al.* (1975) conducted many experiments to study the impact of the aspect ratio of channel (width-to-depth ratio), volumetric discharge rate, and various discharge temperatures on thermal distribution and provided mathematical models to predict them. Teng *et al.* (2018) studied the impacts of lateral jet on the turbulent characteristics in an open channel with rigid vegetation. They compared their results with flow current without vegetation. There are also field studies regarding the pollutant discharge of power plants into rivers and seas (Frigo & Frye 1972; Vaillancourt & Couture 1978). However, the extraction of experimental results is generally difficult for cases with different flow variables, hydraulics (e.g., Reynolds number, Froude number, and roughness), and geometries. Experimental data are usually used as benchmark data for the validation of models.

Many numerical studies have been performed to identify the properties of discharged buoyant jet (McGuirk & Rodi 1978; Kalita *et al.* 2002; Bodart *et al.* 2013; Zhang & Yang 2017). One of the advantages of numerical modeling is the availability of continuous profiles for the temperature distribution rather than discrete experimental data. The computational fluid dynamics (CFD) models also provide the possibility of visualizing three-dimensional structures of flow at less cost than the experimental models. McGuirk & Rodi (1979) developed a three-dimensional numerical model by the Reynolds-averaged Navier–Stokes (RANS)

equations in order to simulate the surface thermal jet discharge in a side channel shallower than the cross flow into a stationary receiver fluid. Addad *et al.* (2004) studied the negatively buoyant wall jets in a cross flow by large eddy simulations (LES). Kim & Cho (2006) investigated the thermal flow discharge in a cross flow by the commercial FLOW-3D model from a side channel. The authors analyzed the temperature distribution in a cross flow under the condition of an adjacent channel with a free surface and submerged states. They used the renormalization group (RNG) k - ϵ model to consider the turbulence effects. Yan & Mohammadian (2017) examined the mixing characteristics in the near field for different confinement indexes and densimetric Froude numbers in a vertical buoyant jet. They used the turbulent buoyancy-modified k - ϵ model in the simulation process. CFD models have provided suitable and satisfactory results, but CFD solutions are heavily influenced by selected modeling parameters such as spatial and temporal discretization schemes, boundary conditions, the velocity-pressure coupling algorithm, turbulence model, under-relaxation factors, etc. (Sun *et al.* 2014). The computation time of CFD models is usually expensive, especially in three-dimensional simulations. Also, a valid numerical simulation requires a great deal of experience and expertise.

Soft computing methods have been considered by researchers due to their high capability in analyzing sophisticated issues. These methods can establish a logical connection between the input and output using the observed data. Artificial neural network (ANN) is one of the most common methods of soft computing. This method has many applications, due to the excellent ability to solve nonlinear and complex phenomena, in various water engineering issues such as local scour depth prediction (Choi & Cheong 2006; Bashiri *et al.* 2018), suspended and bedload transport (Nagy *et al.* 2002; Melesse *et al.* 2011), prediction of the flow mean velocity in an open channel (Sun *et al.* 2014; Gholami *et al.* 2015), discharge capacity of a side weir (Eghbalzadeh *et al.* 2016; Bilhan *et al.* 2018), flow characteristics in open channels (Donmez 2011; Baghalian *et al.* 2012; Huai *et al.* 2013), flood management (Wu & Chau 2011; Fotovatikhah *et al.* 2018), river flow forecasting (Taormina *et al.* 2015; Yaseen *et al.* 2019), and evaporation forecast (Ali Ghorbani *et al.* 2018; Moazenzadeh *et al.* 2018). Despite all the advantages of ANN-based soft

computing models, there are also some shortcomings, among which, being a black box model is the most important, which prevents them from being utilized in practical applications. These methods present the behavior of the system in the form of weighted matrix terms along with bias, which is not easy to understand. Setting up the parameters and over-fitting are among other problems in utilizing these methods. Genetic programming (GP) is another modeling technique used to model various engineering problems. GP as an evolutionary computing method produces transparent mathematical relationships to express the phenomenon being studied. However, GP also has some limitations. It is not very potent to find coefficients that leads to generate functions which grow in length over time.

Recently, a modern data mining method has been developed under the title of evolutionary polynomial regression – multi-objective genetic algorithm to overcome the above-mentioned problem and derive transparent mathematical relations to be used in practical works. This method was first proposed by Giustolisi & Savic (2006) for solving environmental issues, then it was widely used in solving various engineering problems (Javadi & Rezanian 2009; Bonakdari *et al.* 2017; Morano *et al.* 2017; Najafzadeh *et al.* 2017). The key idea of this method is to use the capabilities of genetic algorithms in finding an appropriate symbolic structure for the model in combination with the method of least squares to find the parameters of the symbolic model. The main advantage of the EPR-MOGA technique is the possibility to select models from a set of optimal solutions on the basis of simplicity and its performances on a test set of data. The prediction quality can be improved by retraining the EPR-MOGA model using the new data.

A literature review based on data mining methods showed that no study has been performed in the field of predicting the distribution of pollutants in the junction of channels with cross flow, despite numerous studies in the field of the hydraulics of flow and sediment transport in open channels. The prediction of thermal contaminants' distribution at the junction of channels with cross flow is studied using the EPR-MOGA model. Three distinct mathematical relations are provided by the EPR-MOGA model: one for the prediction of temperature distribution in the main channel (first case); one for determining the temperature half-

thickness profiles as the cross-sectional flow of the pollutants (second case); and another model just to predict the amount of thermal jet spread in the cross flow (third case). In this paper, CFD simulations are performed to assess the capability of the proposed EPR-MOGA models. Previous experimental data have been used to verify the results of the CFD model and training the EPR-MOGA. Temperature distribution profiles, temperature half-thickness profiles, spread of thermal contamination on the surface, and the maximum temperature in different cross sections were obtained separately for the CFD and EPR-MOGA models, and were then compared with each other. The results of this study are very effectual in predicting the distribution of thermal contaminants at the junction of channels with cross flow and designing and constructing the jet outlets from the side channel.

EXPERIMENTAL MODEL

The experimental data measured by Abdelwahed (1981) were selected as the training and testing data in the EPR-MOGA model and verification of CFD models. Abdelwahed (1981) conducted many experiments regarding jet discharge from a side channel under different geometries and temperatures, and measured various experimental data. The data related to the flow discharge from a side channel are selected at the same depth with cross flow in this article. The experiments were conducted in an open wide channel with the length of 900 cm, width of 61 cm, and depth of 15 cm of hot water discharge in cross flow under eight different modes. The channel length in the numerical modeling is selected 1 m at the upstream of the discharge point in order to develop the flow and reduce the computational costs. The channel length at the downstream of the discharge point is selected 1.5 m in order to prevent outflow impacts on the results (Figure 1).

Other test conditions include water velocity at the entry of the main channel, U ; main and side channel depths, d ; the volumetric flow rate at the discharge point of the side channel, Q_0 ; the temperature at the entry of the main channel, T_a ; amount of excess temperature for the outflow jet, T_0 ; and the ratio of outflow jet velocity to the flow velocity at the main channel, R ; these are presented in Table 1 for the tests.

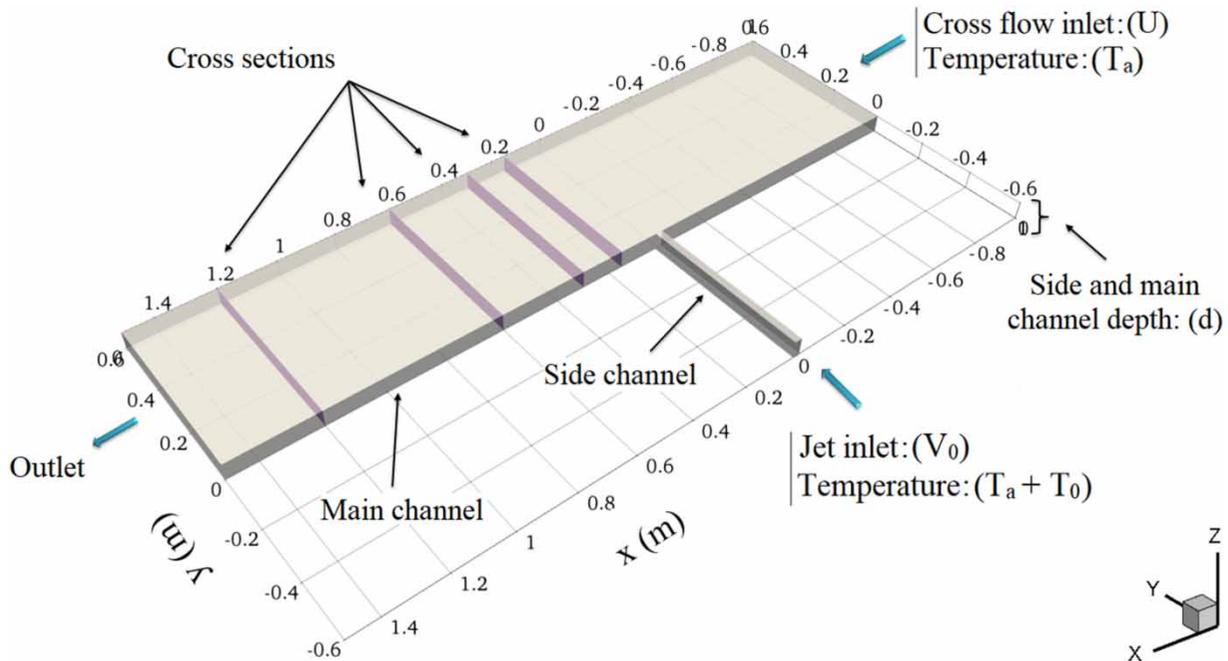


Figure 1 | Situation of cross sections studied in the models.

Table 1 | The experimental conditions in the experiments performed by Abdelwahed (1981)

Test no.	U (cm/sec)	d (cm)	Q_0 (cm ³ /sec)	T_a (°C)	T_0 (°C)	R (-)
1E	6.12	6.00	157.73	21.2	7.80	1.69
2E	6.06	9.00	457.73	20.20	7.80	1.14
3E	8.10	9.08	234.89	20.10	5.00	1.26
4E	6.01	12.5	181.851	19.40	6.50	0.95
5E	5.99	12.5	272.78	18.90	4.40	1.44
6E	6.18	12.5	363.70	18.70	3.35	1.86
7E	6.06	12.5	151.54	18.40	7.70	0.79
8E	6.05	12.5	454.63	18.20	2.75	2.37

The experimental temperature data for each test is measured at four cross sections located at the downstream of the side channel (Figure 1). The location of probes at each cross section was placed in a grid with eight side positions and six different levels for each test. Figure 2 shows a sample of probe locations for the 1E test at a distance of 15 cm from the discharge point. The grids at each test cover all the turbulent cross section area. The level of the first probe at each cross section is set at 0.5 cm from the free surface.

Abdelwahed (1981) calculated the excess temperature at each time step by considering the inflow temperature of the

main channel (T_a) as the reference temperature and subtracting it from the instant temperature of each probe. All the temperature analyses were conducted by the average time of excess temperature during the analysis, which is calculated as follows:

$$T = \frac{1}{N} \sum_{i=1}^N T_i \quad (1)$$

where N is the number of data recorded during the time of analysis. Since the duration of analysis is 55 seconds, and the data were collected at the rate of 25 records in a second, the amount of N is equal to 1,375 for all the analyses. Abdelwahed (1981) used a parameter named temperature half-thickness (δ) in his experiments to visualize the cross section of flow. The temperature half-thickness is equal to a position in which it is half on the free surface at the same position (Figure 3(a)). The spread of thermal contamination in the surface, (r_s), is equal to the maximum distance between the connection point of the temperature half-thickness with a free surface and the internal wall. The maximum time-averaged temperature, T_m , which is equal to the maximum T at the last second ($t = 55$ s), is presented as the test result in addition to temperature profiles and the temperature half-thickness, as shown in Figure 3(b).

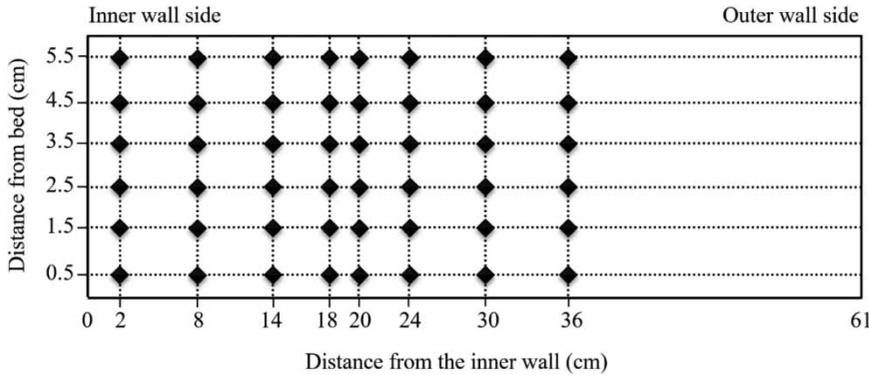


Figure 2 | Points where the velocities were measured in the experimental model cross section.

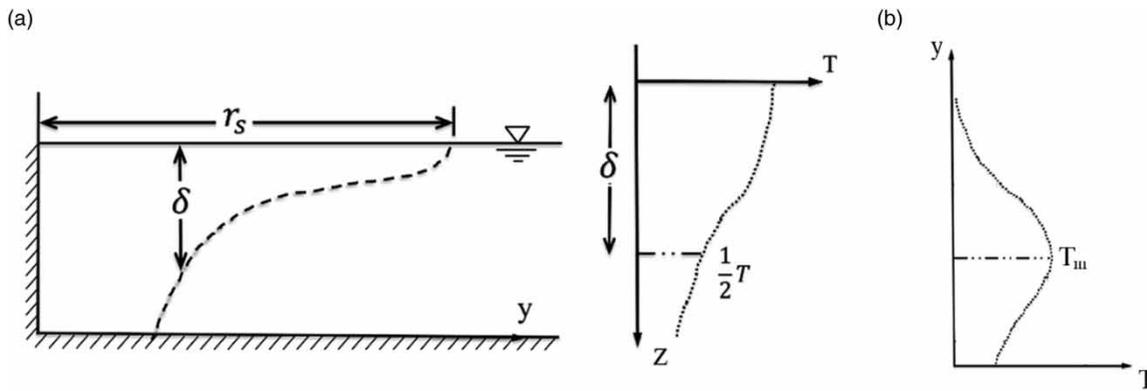


Figure 3 | (a) Temperature half-thickness and (b) position of the maximum time-averaged temperature.

COMPUTATIONAL FLUID DYNAMICS MODEL

Governing equations

The governing equations for incompressible and turbulent fluids in three-dimensional thermal jets are the RANS equations with the Boussinesq approximation for buoyancy effects, written as follows (Ferziger & Peric 2012):

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0 \tag{2}$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_j \bar{u}_i) - \frac{\partial}{\partial x_j} \left\{ \nu_0 \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \overline{u'_i u'_j} \right\} = - \frac{\partial}{\partial x_i} \left(\frac{\bar{p}}{\rho_0} \right) + g_i \rho_k \tag{3}$$

where \bar{u}_i and \bar{u}_j are the average velocity components in the i and j direction, g_i is the acceleration due to gravity in the

i -direction, x_i and x_j are the Cartesian coordinates in the i and j direction, t is time, δ_{ij} is the Kronecker delta function, $\nu_0 = \mu_0 / \rho_0$ is the kinematic viscosity, ρ_0 is the density of the fluid, μ_0 is the constant molecular viscosity, and \bar{p} is the average pressure. ρ_k is the effective kinematic density, which can be expressed as (Ferziger & Peric 2012):

$$\rho_k = 1 - \beta(\bar{T} - T_k) \tag{4}$$

where T_k and \bar{T} are the mean temperature and reference temperature in kelvin, and β is the expansion coefficient with the fluid temperature in kelvin⁻¹. The Reynolds stress term is calculated using Boussinesq hypothesis (Stoesser et al. 2010):

$$-\overline{u'_i u'_j} = \nu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij} \tag{5}$$

where $\nu_t = C_\mu k^2 / \epsilon$ is the turbulent viscosity, k is the turbulent kinetic energy, and ϵ is the rate of dissipation of

turbulence energy. The turbulent viscosity is modeled by the realizable $k-\varepsilon$ turbulence model, first introduced by Shih *et al.* (1995). Huai *et al.* (2010), Kheirkhah Gildeh *et al.* (2014), and Yan & Mohammadian (2017) simulated turbulent buoyant jets by the realizable $k-\varepsilon$ model. In the case of the realizable $k-\varepsilon$ model, the turbulent kinetic energy and its dissipation rate are obtained with the following equations (Shih *et al.* 1995):

$$\frac{\partial k}{\partial t} + \frac{\partial}{\partial x_j} (u_j k) = \frac{\partial}{\partial x_j} \left[\left(\nu + \frac{\nu_t}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right] + \frac{1}{\rho} G_k - \varepsilon \quad (6)$$

$$\frac{\partial \varepsilon}{\partial t} + \frac{\partial}{\partial x_j} (u_j \varepsilon) = \frac{\partial}{\partial x_j} \left[\left(\nu + \frac{\nu_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right] + C_1 S \varepsilon - C_2 \frac{\varepsilon^2}{k + \sqrt{\nu \varepsilon}} \quad (7)$$

where $G_k = \nu_t S^2$ is generation of turbulence kinetic energy due to the mean velocity gradients, $S = \sqrt{2S_{ij}S_{ij}}$ is the modulus of the mean rate-of-strain tensor. In turbulence viscosity, C_μ is the function of the mean strain and rotation rates and given by:

$$C_\mu = \frac{1}{A_0 + A_s U^* \frac{k}{\varepsilon}} \quad (8)$$

In the present study, the parameters A_0 , σ_k , and σ_ε are constant and proposed by Shih *et al.* (1995) as the values of 4.0, 1.0, 1.2, respectively. Relationships were presented to calculate C_1 , A_s , and U^* by Shih *et al.* (1995). The temperature distribution is determined by heat advection diffusion (Currie 2002; White & Corfield 2006; Ferziger & Peric 2012):

$$\frac{\partial \bar{T}}{\partial t} + \frac{\partial}{\partial x_j} (\bar{T} \bar{u}_j) - \frac{\partial}{\partial x_k} \left(k_{eff} \frac{\partial \bar{T}}{\partial x_k} \right) = 0 \quad (9)$$

where \bar{T} is the mean temperature, $k_{eff} = (\nu_t / Pr_t) + (\nu_0 / Pr)$ is the effective thermal diffusivity, $Pr_t = 0.85$ and $Pr = 7$ are the Prandtl number and turbulent Prandtl number, respectively.

CFD solving methods, gridding, and boundary conditions

The Open Field Operation and Manipulation (OpenFOAM) package version 5.0.X is used for simulation. Three-dimensional governing equations are discretized by the

finite volume method. The pressure-velocity couple has been applied by PISO (pressure implicit with splitting of operators) algorithm. The required number of modification steps is low in this scheme to achieve the desired accuracy (Jasak 1996; Oliveira & Issa 2001; Ferziger & Peric 2012). The flow field is solved for all the variables to achieve the absolute error of 1×10^{-8} . Discretization of existing terms in the equations is performed with different methods and orders of accuracy, which is provided in Table 2.

A meshing with a total of 800,000–1,000,000 nodes for all the geometry was selected after checking different meshes and performing sensitivity analysis on the size of grid. Elements with small size were used in areas near to the bed and wall, considering the extreme gradients near these places. The size of elements is selected along the flow in the main channel in a way that elements get smaller by getting close to the side channel, and they increase a little towards downstream by getting away from the side channel to make the calculations more effective and finally reach a constant element size. The size of elements also have a constant value along the depth of channels except for near the bed. The plan and cross-sectional views are shown in Figure 4 with magnification of details.

The boundary conditions used for the inlets include uniform velocity and constant temperature. The pressure boundary condition at all boundaries is zero gradient, except for the outlet of the main channel, in which a constant static pressure has been defined. No-slip condition for velocity and zero gradient condition for temperature have been defined at the solid walls. Air impact has been disregarded and symmetry boundary condition is used for the free surface in this research.

Table 2 | Boundary conditions applied in the CFD model

Term	Type of scheme	Notes
Time derivatives	Backward scheme	Implicit second order accurate
Gradient terms	Gauss linear	Cell limited
Convective terms	Filtered linear 2 V	A combined scheme
Laplacian term	Gauss linear limited	

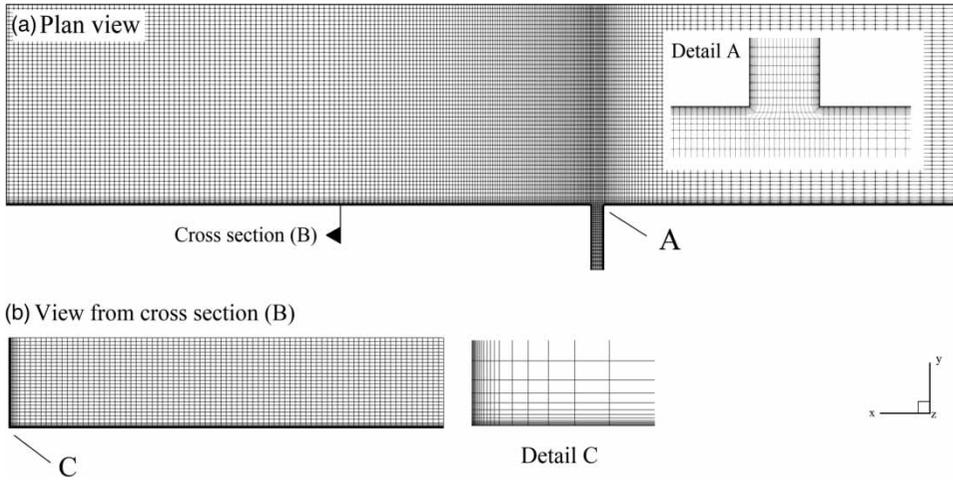


Figure 4 | Gridding in CFD model for (a) plan and (b) cross section.

EVOLUTIONARY POLYNOMIAL REGRESSION (EPR) MODELS

EPR structure

The EPR technique is a hybrid data mining method, which was first developed by Giustolisi & Savic (2006) in order to provide mathematical expressions to predict engineering problems. The symbolic models are formed in this technique by exploiting the capabilities of genetic algorithm and numerical regression. In fact, the structure of symbolic models is determined by the genetic algorithm in the first step, and then a linear least squares problem is solved in the second step to determine the parameters of the symbolic model. The proposed general symbolic expression in this method is as follows:

$$y = \sum_{j=1}^m F(X, f(x), a_j) + a_0 \tag{10}$$

in which y is the estimated output during the analysis process by the EPR method, m refers to the maximum number of final polynomial terms by ignoring the bias term a_0 , F is a function determined during the search process of the symbolic structure for each term, X is the matrix of independent variables used for the construction of the model, f is a function that should be selected by the user beforehand, and a_j is the constant factor for the j -th term.

Equation (10) must be changed into the following vector form at the first step in order to use it in identifying the structure of the model:

$$\begin{aligned} \mathbf{Y}_{N \times 1}(\boldsymbol{\theta}, \mathbf{Z}) &= [\mathbf{I}_{N \times 1} \quad \mathbf{Z}_{N \times m}^j] \times [a_0 \quad a_1 \quad \dots \quad a_m]^T \\ &= \mathbf{Z}_{N \times d} \times \boldsymbol{\theta}_{d \times 1}^T \end{aligned} \tag{11}$$

in which, $\mathbf{Y}_{N \times 1}(\boldsymbol{\theta}, \mathbf{Z})$ is the vector of calculated values for N objective value, $\boldsymbol{\theta}_{d \times 1}^T$ is the transpose vector, $d = m + 1$ is the constant factor of a_j and a_0 , as $j = 1:m$, $\mathbf{Z}_{N \times d}$ is the matrix consisting of the unit matrix \mathbf{I} and variables of \mathbf{Z}^j , the variables of \mathbf{Z}^j consist of independent input parameters of $X = \langle X_1 X_2 X_3 \dots X_k \rangle$.

We start the EPR from Equation (12) of the process in order to find the best combination of vector input variables $X_{s=1:k}$.

The input variable matrix is defined as follows:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Nk} \end{bmatrix} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_k] \tag{12}$$

The candidate input variables in the j th term of Equation (12) in the k th column of X are shown. Therefore, the j th term of Equation (12) can be written as follows:

$$\mathbf{Z}_{N \times 1}^j = [(\mathbf{X}_1)^{\text{ES}(j,1)} \cdot (\mathbf{X}_2)^{\text{ES}(j,2)} \cdot \dots \cdot (\mathbf{X}_k)^{\text{ES}(j,k)}] \tag{13}$$

in which Z^j is the j th column vector, the members of which are composed from the vector of independent candidate inputs and ES is the matrix of exponent. Thus, the main issue is to search for the matrix of exponent $ES_{k \times m}$. For example, if the user has chosen the vector of exponents of independent variables, X , as $EX = [0, 1, 2, -1, -2]$, and the number of symbolic terms without considering the bias is 4, and the number of inputs, k , is 3, the main purpose in the following would be to find the matrix of exponents $ES_{4 \times 3}$. For example, the matrix $ES_{k \times m}$ can be written as follows:

$$ES_{4 \times 3} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 0 \\ -2 & 1 & 0 \end{bmatrix} \tag{14}$$

The elements of matrix ES are the values of proposed exponents of the user in the EX vector. Also, each row of the above matrix shows the exponent of independent input variables for the j th term from Equation (12). The variables of Z^j are obtained by substituting the above matrix into Equation (14) as follows:

$$\begin{aligned} Z_1 &= (X_1)^{-1} \cdot (X_2)^0 \cdot (X_3)^1 = X_1^{-1} \cdot X_3 \\ Z_2 &= (X_1)^0 \cdot (X_2)^2 \cdot (X_3)^{-1} = X_2^2 \cdot X_3^{-1} \\ Z_3 &= (X_1)^1 \cdot (X_2)^0 \cdot (X_3)^0 = X_2 \\ Z_4 &= (X_1)^{-2} \cdot (X_2)^1 \cdot (X_3)^0 = X_1^{-2} \cdot X_2 \end{aligned} \tag{15}$$

The following equation is obtained after replacing the resulted parameters of matrix 15 into Equation (12):

$$\begin{aligned} Y &= a_0 + a_1 Z_1 + a_2 Z_2 + a_3 Z_3 + a_4 Z_4 \\ &= a_0 + a_1 \frac{X_3}{X_1} + a_2 \frac{X_2^2}{X_3} + a_3 X_2 + a_4 \frac{X_2}{X_1^2} \end{aligned} \tag{16}$$

The adjustable parameters of the mentioned symbolic term, a_j , are now determined by solving a least squares problem.

The accuracy level of the model is determined by the coefficient of determination (COD) during the search process, which is derived from the following relation:

$$COD = \left(\frac{\sum_{i=1}^n (Fr_{observed_i} - \overline{Fr_{observed_i}})}{(Fr_{predicted_i} - \overline{Fr_{predicted_i}})} \right)^2 \tag{17}$$

$$= \frac{\left(\frac{\sum_{i=1}^n (Fr_{observed_i} - \overline{Fr_{observed_i}})}{\sqrt{\sum_{i=1}^n (Fr_{observed_i} - \overline{Fr_{observed_i}})^2}} \right)^2}{\left(\frac{\sum_{i=1}^n (Fr_{predicted_i} - \overline{Fr_{predicted_i}})}{\sqrt{\sum_{i=1}^n (Fr_{predicted_i} - \overline{Fr_{predicted_i}})^2}} \right)^2}$$

in which, Y_a is the real value measured in the laboratory, Y_p is the value of calculated output variable by EPR method, and N is the number of data used.

The above-mentioned steps form a complete cycle of a search process to find the best model by EPR method. This process ends when the termination criterion is fulfilled. This criterion depends on the number of input variables, the number of maximum defined terms for the model, and the length of chromosome. The process diagram for the EPR method is shown in Figure 5.

Multi-objective EPR

Despite all the capabilities of the original EPR version, the method has some shortcomings, including: the efficiency of the model is reduced by the increase of the problem search space (due to the increase in the number of inputs and terms of the model). It has also been proven that the simpler model is a more appropriate model, but appropriate approaches have not been considered in the main method of EPR to reduce the complexity of model, and models are built only based on the accuracy increase of a specific statistical indicator. The existence of these shortcomings caused Giustolisi & Savic (2009) to use a multi-objective genetic algorithm (MOGA) instead of a single-objective genetic algorithm (SOGA) for the search process. This change created a new version of the method called the evolutionary polynomial regression – multi-objective genetic algorithm (EPR-MOGA). At least two functions are used in the EPR-MOGA method to search in the problem space, so that a target function maximizes the accuracy of the model and at least one target function minimizes the complexity of the model. Applying these changes in the improved model decreased the input variables, increased the accuracy of the model, and decreased the number of terms in comparison to the EPR model.

Setting of EPR-MOGA initial parameters and data preparation

It is necessary to adjust some initial parameters to conduct an analysis with the EPR-MOGA model, including: specifying the structure of the model (with or without bias), number of polynomial terms, range of suggested exponents for the

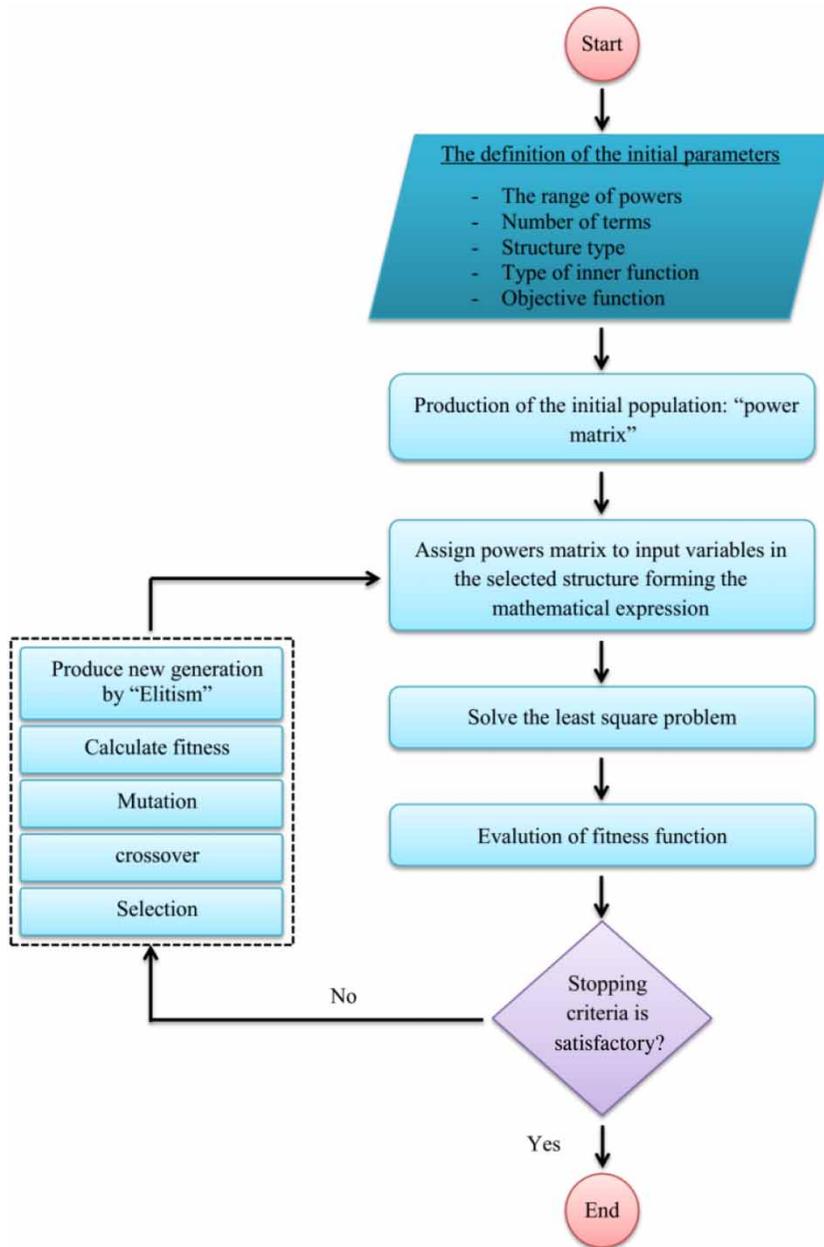


Figure 5 | Algorithm of evolutionary polynomial regression (EPR) method.

variables of the model, determining the target functions for the searching process, and finally, setting some parameters related to the evolutionary algorithm such as initial population and number of iterations. The number of terms and different ranges of exponents were investigated with or without bias condition in order to produce the best model, and the best one of them is chosen according to the statistical indicators obtained from the produced models. The bias

term is considered in the produced models. The range of exponents is between -2 and 2 with steps of 0.5 . The structure of the model is considered in its simplest form. The maximum number of terms for modeling the first case (average temperature distribution) was selected to be 100 , and for the second (temperature half-thickness) and third (jet spread across the main channel) cases was 30 . The maximum number of terms is selected in a way that a model be

produced with maximum capability without considering the complexity of the model. The conducted settings are summarized in Table 3.

Overall, 1,634 experimental data for the first case, 1,632 data for the second case, and 34 data for the third case were used in this study. The data were equally divided into two groups for training and testing the model. The temperature at each point depends on fluid characteristics (such as jet (V) and cross flow velocity (U)), flow characteristics (such as temperatures of cross flow (T_a) and side channel (T_j)), channel dimensions (such as depth of main and side channel (d)) and the Cartesian coordinates of the studied point (x, y, z) as follows:

$$F = f(x, y, z, d, T_a, T_j, U, V) \tag{18}$$

Abdelwahed (1981) suggested using the temperature excess ($T_o = T_j - T_a$) parameter in the definition of specifications of the fluid temperature and jet exit to cross flow velocity ratio ($R = V/U$) in the flow field definition. Also, in this paper, the variables of $x, y,$ and $z,$ which are being used to describe the spatial coordinates of the studying point, became dimensionless in order to derive a mathematical relation which is general and independent of dimension, as follows:

$$F = f\left(X = \frac{x}{d}, Y = \frac{y}{d}, Z = \frac{z}{d}, T_o, R\right) \tag{19}$$

Initial settings of the EPR-MOGA analysis and the data overview are presented in Table 3.

According to Table 3, three relations were obtained for predicting the temperature distribution (Equation (20)), temperature half-thickness (Equation (21)), and the spread of thermal contamination at the free surface (Equation (22)):

$$\begin{aligned} T_{EPR-MOGA} = & 0.27764 \frac{Z^{1.5} Y^2 T_o}{X^{1.5} R} - 1.4172 \frac{Z^2}{X^{1.5} R^{0.5}} \\ & - 1.988 \frac{Z^{1.5} Y^{1.5} T_o^{0.5}}{XR} + 0.025364 \frac{T_o^{1.5}}{X^{0.5} Y} + 14.4746 \frac{Z^2}{X^{0.5} R T_o^{0.5}} \\ & + 2.0106 \frac{1}{Y^2 T_o^{1.5}} - 3.8549 \frac{1}{Y^2} + 26.9065 \frac{1}{Y^{1.5}} - 77.7664 \frac{1}{Y} \\ & + 26.9543 \frac{1}{Y^{0.5} R^{0.5}} + 89.9413 \frac{1}{Y^{0.5}} - 1.0029 \frac{R^2}{Y^{0.5}} \\ & - 4.9872 \frac{T_o^{1.5}}{R^{0.5}} + 750.2023 \frac{1}{T_o} + 1452.2152 T_o^{0.5} \\ & - 521.4064 T_o + 71.0843 T_o^{1.5} + 29.2388 \frac{Y^{0.5}}{R^{1.5}} \\ & + 17.9947 \frac{Y^{0.5}}{T_o^{1.5}} + 13.3904 Y^{0.5} R^{0.5} - 7.4645 Y^{0.5} R \\ & - 5.2406 \frac{Y}{R^2} - 0.63774 \frac{Z^2 T_o^{0.5}}{Y^{0.5}} + 0.092633 Z^2 T_o^2 \\ & - 0.0048306 X^{0.5} R^{0.5} T_o^2 - 0.0028508 \frac{X Z^2 T_o^2}{Y^{0.5} R} - 1616.995 \end{aligned} \tag{20}$$

Table 3 | Initial settings of the EPR-MOGA analysis and the data overview

Parameters	EPR-MOGA setting		
	Case 1: Mean temperature	Case 2: Temperature half-thickness	Case 3: Spread of jet across the channel
Regression type	Linear regression		
Polynomial structure	$Y = \sum (a \times f(X1 \times X2)) + a_0$		
Inner function type	No function		
Constant estimation method	Least square		
Range of exponents	[-2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2]		
Maximum number of terms	[1:100]	[1:30]	[1:30]
Number of data	1,634	1,632	34
Number of training data	817	816	17
Number of testing data	817	816	17
Input variables	(x, y, z, R, d, T_o)	(x, R, d, T_o)	(x, y, R, d, T_o)
Output variables	T	H	S

$$\begin{aligned}
 H_{EPR-MOGA} = & -1.8808 \frac{Y^{0.5}}{X^2 R^{1.5} T_0^{0.5}} + 0.0037944 \frac{Y^2 R^{1.5} T_0^2}{X^2} \\
 & + 0.24141 \frac{R^{0.5} T_0^{1.5}}{X^{1.5}} - 0.46883 \frac{R^{1.5} T_0^{1.5}}{X^{1.5}} \\
 & - 0.20337 \frac{Y^{1.5} R^{1.5} T_0^{0.5}}{X^{1.5}} + 23.9225 \frac{1}{X} - 59.1894 \frac{1}{X^{0.5} R^{0.5}} \\
 & + 27.8289 \frac{Y^{0.5}}{X^{0.5}} - 3.4181 \frac{Y T_0^{0.5}}{X^{0.5}} + 1.2837e - 09 \frac{1}{Y^{0.5}} \\
 & - 114.5158 \frac{T_0^{0.5}}{R^2} - 532.3097 \frac{1}{R T_0^{0.5}} + 1148.9758 T_0^{0.5} \\
 & + 53.9165 T_0^{1.5} - 1014.0181 R^{0.5} - 493.5939 R^{0.5} T_0 \\
 & + 439.4998 R^{1.5} T_0^{0.5} - 171.3011 R^2 - 11.4575 Y^{0.5} R^2 \\
 & + 12.1222 Y R^{1.5} - 5.2212 Y^{1.5} R^{0.5} + 0.89121 Y^2 T_0 \\
 & - 0.21721 Y^2 T_0^{1.5} - 1.0892 Y^2 R^{0.5} - 5.8832e - 09 \frac{X^{0.5}}{Y^{0.5} T_0} \\
 & - 0.34471 X^{0.5} T_0^2 - 38.245 \frac{X^{0.5} Y^{0.5}}{R^{1.5} T_0} + 11.4294 \frac{X^{0.5} Y^{0.5} R^{0.5}}{T_0^{0.5}} \\
 & + 1.3826 \frac{X^{0.5} Y}{R^2} - 0.055091 X^{0.5} Y T_0^{1.5} - 0.50113 X^{0.5} Y R^2 \\
 & + 6.2513e - 09 \frac{X}{Y^{0.5} R^{0.5} T_0^2} - 2.3628 X T_0^{0.5} + 1.5791 X T_0 \\
 & + 0.01933 X Y^{1.5} R^{1.5} - 0.071444 X^{1.5} T_0 + 0
 \end{aligned} \tag{21}$$

$$\begin{aligned}
 S_{EPR-MOGA} = & -37.6974 \frac{1}{X T_0} + 498482.3257 \frac{1}{R^{0.5} T_0^{0.5}} \\
 & - 144902.7363 \frac{1}{R^{0.5}} - 190602.2304 \frac{T_0^{0.5}}{R^{0.5}} + 63939.45 \frac{T_0}{R^{0.5}} \\
 & + 155514.4722 T_0^{0.5} - 3123.3424 T_0^2 + 2.3303 \frac{X^{1.5}}{R^{1.5} T_0^{0.5}} \\
 & - 244262.1956
 \end{aligned} \tag{22}$$

where X , Y , and Z are dimensionless spatial coordinates relative to the depth of channel, d ; R is the ratio of velocity in the side channel to the cross flow; and T_0 indicates the temperature difference between the inflow of the side channel and inflow of the cross flow, which is a dimensionless value.

MODEL PERFORMANCE

There are several methods to quantify the difference between predicted and observed data. In this research, the statistical indicators of COD, mean squared error (MSE),

mean absolute error (MAE), root mean square error (RMSE), and relative error (RE) are used with Equations (23)–(27) in order to evaluate the accuracy of the results for both the CFD and EPR-MOGA models. One of the most common of statistical indicators is MSE defined as average of squares of the errors. If all cases within a cluster are identical the MSE would then be equal to 0. This paper uses RMSE to evaluate the closeness of the results of numerical models, results compared to those of experimental data. RMSE measures how much error there is between two data sets. In other words, it compares a predicted value and an observed or known value. RMSE is always non-negative, and a value of 0 (almost never achieved in practice) would indicate a perfect fit to the data. In general, a lower RMSM is better than a higher one. MAE is the average of absolute residuals and represents the closeness between forecasted and actual values. The advantage of RMSE and MAE models in evaluating the model error is that both use the same unit of variables. COD provides a measure of how well the observed outcomes are replicated by the model. The closer COD gets to 1, the better the results. Also, COD shows how well the numerical methods replicated the observed results. Thus, to achieve a fair judgment about the performance of a numerical method, it is good for all of the statistics to be investigated together.

$$COD = \left(\frac{\sum_{i=1}^n (Fr_{observed_i} - \overline{Fr_{observed_i}})}{(Fr_{predicted_i} - \overline{Fr_{predicted_i}})} \right)^2 \tag{23}$$

$$MSE = \frac{\sum_{i=1}^n (Fr_{observed_i} - Fr_{predicted_i})^2}{n} \tag{24}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Fr_{observed_i} - Fr_{predicted_i})^2} \tag{25}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Fr_{observed_i} - Fr_{predicted_i}| \tag{26}$$

$$RE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Fr_{observed_i} - Fr_{predicted_i}}{Fr_{observed_i}} \right| \tag{27}$$

In the above relations, n is the number of data, $Fr_{observed_i}$ and $Fr_{predicted_i}$ are the measured and predicted outputs,

respectively, $\overline{Fr}_{observed,i}$ and $\overline{Fr}_{predicted,i}$ are the average of the measured and predicted outputs, respectively.

RESULTS AND DISCUSSION

The hot water flow forms a recirculation region after being discharged from the side channel. The simulation result of the recirculation region length is investigated to verify the CFD model in predicting the flow field. Figure 6 shows the time-averaged field flow streamlines for the 1E test. In this figure, the horizontal axis is in the direction of the cross section towards the downstream of the side channel and the vertical axis is in the transverse direction of the main channel perpendicular to the cross flow. The numerical simulation and the experimental data of the flow separation point are shown in Figure 6. Abdelwahed (1981) reported separation point equal to 0.5785. The separation point in the numerical model occurs approximately at the distance between 0.575 and 0.58 m. As can be observed, the CFD model has a good agreement with the experimental results.

Figure 7 shows the time-averaged temperature distribution profile of experimental, numerical simulation and calculated by the EPR-MOGA model at distances of $x = 15, 30, 60,$ and 120 cm and depths of $z = 0.5, 1.5, 2.5, 3.5, 4.5,$ and 5.5 cm for the 1E test. The temperature difference in this test between the ambient fluid and the jet fluid is 7.8 K, which has a maximum temperature difference similar to the 2E test among all other conducted tests. The value of R is equal to 1.69 in this test. The horizontal axis represents the flow in the transverse direction and the vertical axis of the diagram represents the amount of time-averaged

temperature. The solid line and the symbols in this figure imply the results of the CFD model and experimental model, respectively. Examining the profiles obtained from the CFD model in Figure 7(a) indicates that temperature values near to the inner wall and also the peak value of temperature in the depths near to the bed ($z = 0.5$ and $z = 1.5$) are appropriate, and the predicted value decreases in comparison with the experimental results when getting closer to the free surface. The temperature decay rate is almost good after the peak point and spreading toward the outer wall. Figure 7(b) shows that the error rate in the peak value and the approximate position increases by getting away from the discharge point. The model behavior is similar to Figure 7(a) near the inner wall and after the peak point. Moving towards downstream of the discharge point (Figure 7(c) and 7(d)), a large error is observed in the amount of peak point and its position and the rate of temperature progression towards the outer wall for depths close to the free surface. The quantity and position of peak point, temperature decay after the peak point, and temperature near to the inner wall were predicted with an appropriate accuracy for other depths. The dotted lines in this figure indicate the results of EPR-MOGA model (relation (18)). Using the proposed relation for drawing the temperature distribution profiles by the EPR-MOGA model, negative values are calculated for the time-averaged temperature at some distances near to the bed, which is very close to zero. The reason for this is the wrong prediction of the model for the results with zero values. However, the lack of data points for training the EPR-MOGA model is one of the most important reasons to predict these negative values. Hence, the negative values calculated by

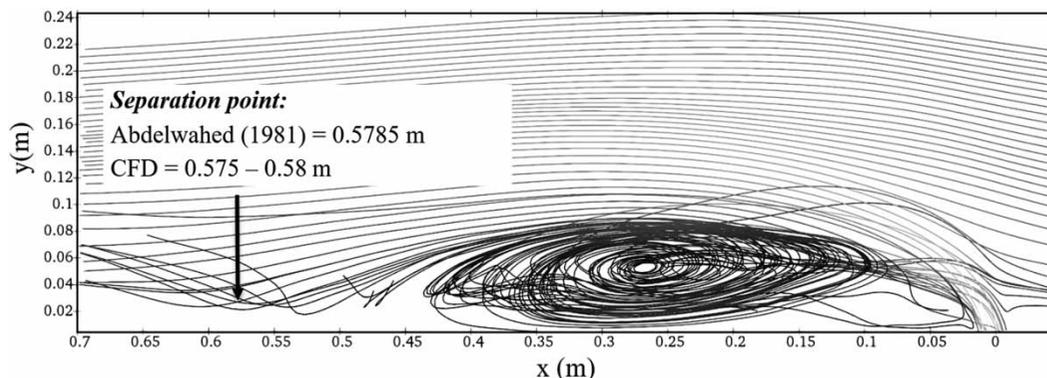


Figure 6 | The free surface streamlines illustrated by the time-averaged flow field of the 1E test.

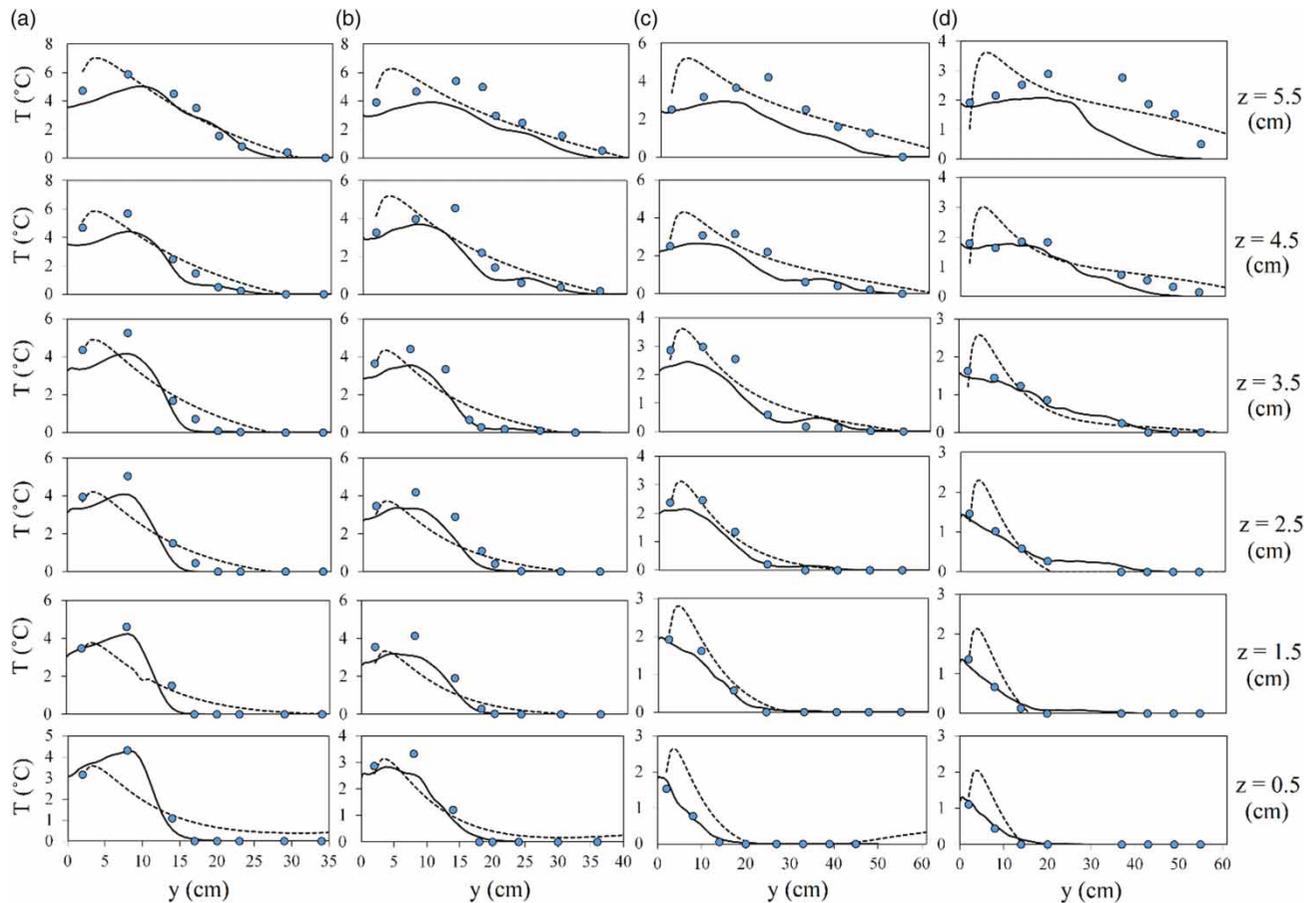


Figure 7 | Experimental (●), numerical (solid line), and EPR-MOGA results (dashed line) of the time-averaged temperature for the 1E test: (a) $x = 15$ cm; (b) $x = 30$ cm; (c) $x = 60$; (d) $x = 120$ cm.

the model for T are limited to zero in order to improve the efficiency of the proposed relation. As observed, the EPR-MOGA model has been able to get the temperature value in the closest position to the inner wall in all the sections in good agreement with the experimental results. The pollutant spread rate towards the outer wall and the time-averaged temperature value are also in good agreement with the experimental results. A peak position is formed in all the cross sections at a distance of 2.5–7.5 cm from the inner wall, which indicated a high error of the EPR-MOGA model in predicting the position of the peak point formation, which increases by getting away from the discharge point. However, as can be observed from the figure, the peak value has no large error with regard to the experimental value.

The values for MAE, RMSE, and COD obtained from the time-averaged temperature predictions for all the 1–8E tests are presented in Table 4. Reviewing the results of this

table based on the error indicators (RMSE and MAE) for the CFD model indicates that it does not have good accuracy in predicting the temperature distribution for tests with the values of $R < 1$ and has large errors. The best results are for the 8E test, which has the maximum R and minimum T_0 . The higher the T_0 value, the greater the error rate of the numerical results is. For the tests 1E and 2E with the same value of T_0 , the 1E test with the higher amount of R has a lower error rate. According to the comparison made on the basis of R and T_0 , it can be concluded that the CFD model error becomes less as the flow behavior becomes closer to the pure jet (i.e., the high value of R and the lower density difference between the jet fluid and the ambient fluid). Investigating the results for the EPR-MOGA model based on the error indicators shows that the accuracy of results gets better by T_0 getting lower and R getting greater values, and vice versa. As can be concluded from Table 4,

Table 4 | Statistical indicators for the results of the temperature distribution predicted by the CFD and EPR-MOGA models for 1–8E tests

Run	CFD			EPR-MOGA		
	COD	RMSE (m)	MAE (m)	COD	RMSE (m)	MAE (m)
E1	92.52	0.512	0.376	84.21	0.630	0.461
E2	71.91	0.562	0.382	84.74	0.613	0.458
E3	94.96	0.322	0.231	86.43	0.458	0.369
E4	58.43	0.573	0.415	83.39	0.525	0.399
E5	75.17	0.295	0.208	81.04	0.398	0.328
E6	94.68	0.166	0.117	87.54	0.259	0.217
E7	86.72	0.688	0.451	85.08	0.626	0.486
E8	98.47	0.114	0.087	81.89	0.265	0.206
Averaged values	84.11	0.404	0.283	84.29	0.472	0.365

the best accuracy is for the 5E, 6E, and 8E tests, which have the highest R and lowest T_0 , and the most inappropriate results are for the 1E, 2E, and 7E tests, which have the lowest R and highest T_0 . Comparing the average values of MAE, RMSE, and COD (0.365, 0.472, and 84.29, respectively) in Table 4 indicates an acceptable ability of the EPR-MOGA model in comparison with the CFD model (MAE = 0.283, RMSE = 0.404, and COD = 84.11) in the prediction of temperature distribution in the channel. Also, the behavior of the model relative to the parameters R and T_0 in the EPR-MOGA model is similar to the CFD model.

The buoyant jet gradually detaches from the bed and spreads on the surface after the recirculation region due to lower density than the receiver fluid. Figure 8 illustrates

the cross section of flow for the 1E test at various distances from the discharge point. As described in the previous section, the parameter of temperature half-thickness (δ) is used to calculate the cross-sectional flow. The horizontal axis is along the width of the main channel and the vertical axis is along the channel depth. The solid line and the symbols in this figure imply the results of the CFD model and experimental model, respectively. The analysis of the profiles obtained from the CFD model indicated that the density separation point from the bed is predicted less than the experimental results from the distance of $x = 15$ cm to the distance of $x = 60$ cm. The connection point to the free surface is correctly predicted for the distance of $x = 15$ cm, but this value is greater than the experimental model for the distance of $x = 30$ cm onwards. However, the model has not been able to define any connection point to the free surface at $x = 60$ cm. The dotted lines in this figure indicate the results of EPR-MOGA model (relation (2)). To increase the efficiency of the relation used by the EPR-MOGA model, the negative results calculated by Equation (2) are considered to be zero. It can be observed in Figure 8 for the EPR-MOGA model that the separation point from the bed and spread length on the surface is predicted less than the experimental model at the distance of $x = 15$ cm. The separation point from the bed is predicted as lower than the experimental model at the distance of $x = 30$ cm, while the propagation rate on the surface is approximately consistent with the experimental results. The separation point from the bed is approximately calculated 2 cm less than the experimental model at a distance of $x = 60$ cm. The experimental results

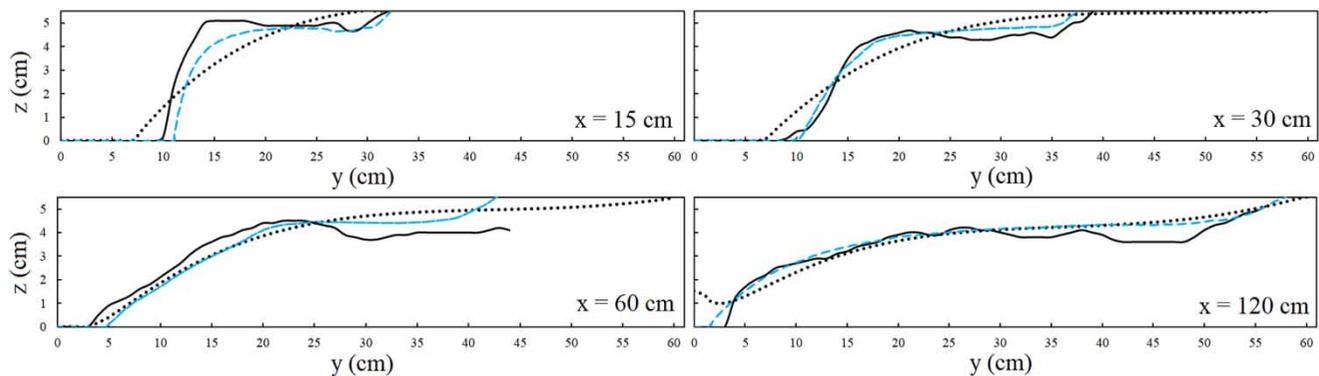
**Figure 8** | The thermal cross section (temperature half-thickness) of the full-depth plume flow in the 1E test; numerical (solid line), experimental (dashed line), and EPR-MOGA results (dots).

Table 5 | Statistical indicators for the results of the *temperature half-thickness* predicted by the CFD and EPR-MOGA model for 1–8E tests

Run	CFD			EPR-MOGA		
	COD	RMSE (m)	MAE (m)	COD	RMSE (m)	MAE (m)
E1	89.85	0.563	0.446	89.94	0.419	0.304
E2	92.08	0.774	0.618	91.47	0.425	0.329
E3	83.60	1.267	0.990	83.01	1.076	0.814
E4	72.32	0.933	0.791	76.60	0.962	0.715
E5	84.99	1.178	0.970	75.94	1.329	1.033
E6	85.43	1.331	1.087	83.99	1.419	1.147
E7	91.27	1.030	0.889	85.66	0.836	0.591
E8	88.00	2.555	1.956	76.12	1.567	1.183
Averaged values	85.94	1.204	0.968	82.84	1.004	0.764

have expanded to a distance of $y = 42$ cm from the inner wall for the distance of $x = 60$ cm, while this expansion extends almost to the outer wall in the EPR-MOGA model. However, the thickness of the occupied space of free surface decreases much more moving towards the outer wall. This model is in very good agreement with the experimental data at a distance of $x = 120$ cm, except the identification of the separation point from the bed. Comparing the results of the CFD model with EPR-MOGA model indicates the capability of both models in

predicting the cross section of flow in terms of temperature half-thickness. However, the results presented in Table 5 for the models of CFD and EPR-MOGA indicated a higher ability of the EPR-MOGA model compared to the CFD model, respectively.

The values for MAE, RMSE, and COD obtained from the predicted cross-sectional flow by the temperature half-thickness term for all the tests (1–8E), are presented in Table 5.

The correlation analysis between the experimental data and the predicted results of the CFD and EPR-MOGA models (relation (3)) for the amount of thermal contamination propagation on the surface for all the tests (1–8E) are presented in Figure 9. The COD index for the CFD model is equal to 0.8752 (Figure 9(a)). The scatter diagram for the test data and training are presented in Figure 9(b) for the EPR-MOGA model separately to ensure the universality of the model. The COD indexes for the test and training data are 0.89 and 0.9887, respectively, which are close to 1 and indicate the proper accuracy of the predictions. The comparison of COD values indicates the lower accuracy of the CFD model than the results of the EPR-MOGA model. The fitting line is also drawn on the figures by the relation of $y = C_1x + C_2$. The more the C_1 and C_2 factors get close to values of 1 and 0, respectively, the more accurate the predictions will be. The exact line is also plotted in the figure. Any predictions placed on the exact line are in perfect agreement with the experimental results. The predicted results are

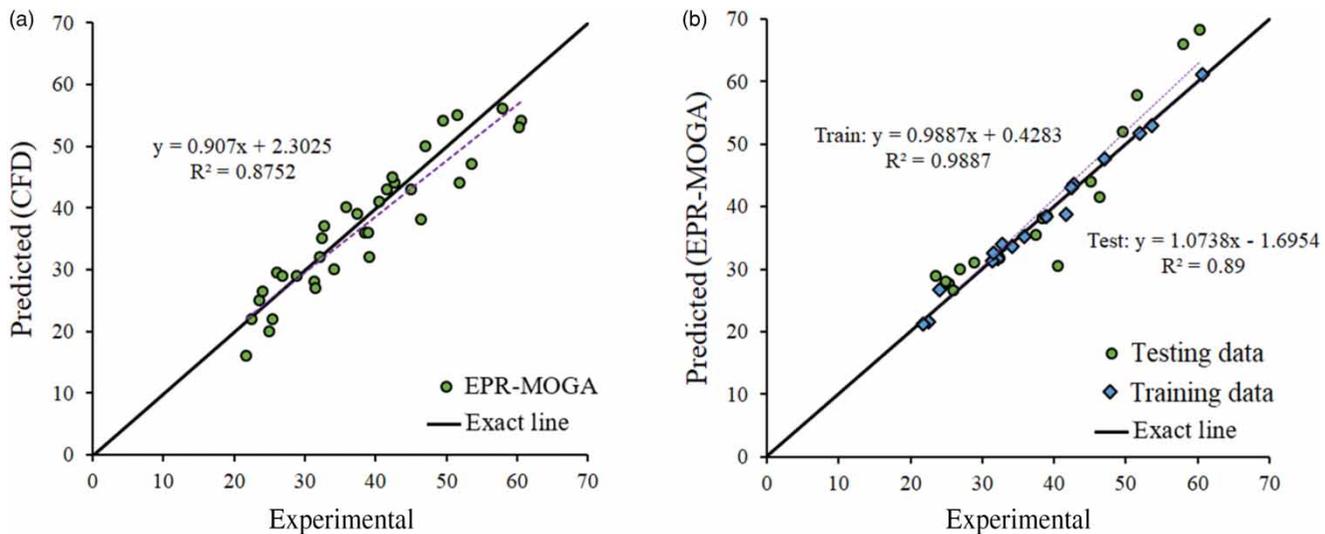


Figure 9 | Scatter plots of experimental data for (a) EPR-MOGA and (b) CFD model to predict the spread of thermal pollution at the surface.

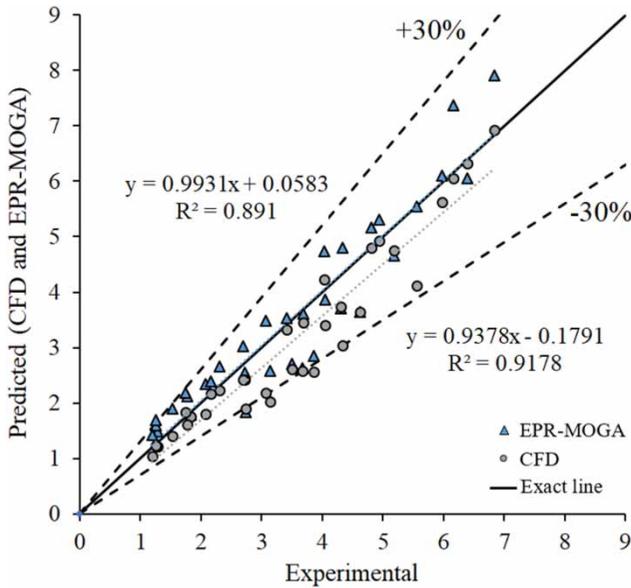


Figure 10 | Scatter plots of experimental data for EPR-MOGA (▲) and CFD (●) model to predict the maximum time-averaged temperature.

overestimations in comparison to the experimental results if the data are placed above this line, and they are underestimations if the data are placed below this line. The CFD model is neither overestimated nor underestimated, according to Figure 9(a). Investigating the data by the exact line for the EPR-MOGA model in Figure 9(b) also indicates that the results of analysis for both categories are neither overestimated nor underestimated.

The correct prediction of the maximum temperature is one of the design requirements for the thermal dischargers. The correlation analysis in comparison with the experimental data for both the models is presented in Figure 10, in order to demonstrate the capabilities of the EPR-MOGA model in predicting the peak value in different sections and compare it with the results of the CFD model. The CFD model, according to the figure, is obviously underestimated in the predictions, while no underestimation or overestimation could be defined for the EPR-MOGA model. The CFD model with an index value of 0.9178 for COD presents better accuracy in comparison with the EPR-MOGA model with an index value of 0.891 for COD. The calculation of relative error values for the EPR-MOGA model represents RE equal to 15%, while it is 11.92% for the CFD model, which indicates a better accuracy of the CFD model compared to

the EPR-MOGA model in predicting the maximum temperature of thermal contaminant.

In this section, an uncertainty analysis is applied to assess the proposed EPR-MOGA models versus the CFD model for predicting the parameters of mean temperature, temperature half-thickness, and spread of thermal effluent across the channel. The uncertainty analysis is used for all the data sets of this study. A reasonable comparison of the uncertainty in forecasting various variables of models is presented using these data sets. The calculated prediction error ($E_i = P_i - T_i$) for all the data sets are employed to compute the mean (\bar{E}) and standard deviation (S_E) of the prediction errors as follows:

$$\bar{E} = \frac{1}{n} \sum_{i=1}^n E_i \quad (28)$$

$$S_E = \sqrt{\frac{\sum_{i=1}^n (E_i - \bar{E})^2}{n - 1}} \quad (29)$$

where E_i is the i th error between the predicted value (P_i) and measured data (T_i), and n is the number of the data. A negative mean value shows that the model underestimates the actual values, and a positive value indicates that the model overestimates the actual values. The confidence band was defined by using \bar{E} and S_E values, and the Wilson score method without continuity correction. The 95% confidence band was achieved at $\pm 1.95996S_E$.

Table 6 shows the mean estimation errors of the EPR-MOGA and CFD models, the uncertainty band width, and 95% prediction interval error. The EPR-MOGA model predicts better than the CFD model with lower uncertainty. The mean prediction errors of the EPR-MOGA model are -0.006 , -0.021 , and 0.574 in comparison to -1.656 , 0.512 , and -1.257 for the CFD model. Both EPR-MOGA and CFD models underestimate the mean temperature. The EPR-MOGA model overestimates the spread of thermal effluent across the channel and underestimates the temperature half-thickness, while the opposite sign is for the CFD model. The mean error values of the EPR-MOGA model relatively is less than the CFD model. The uncertainty band of the EPR-MOGA model is smaller than the CFD model in the mean temperature, temperature half-thickness, and spread of thermal effluent across the channel. Likewise,

Table 6 | Uncertainty analysis of the EPR-MOGA and CFD models

Parameters	Mean prediction error		Width of uncertainty band		95% prediction error interval	
	EPR-MOGA	CFD	EPR-MOGA	CFD	EPR-MOGA	CFD
Mean temperature	-0.006	-1.656	±0.024	±0.073	-0.030 to 0.019	-1.749 to 1.603
Temperature Half-thickness	-0.021	0.512	±0.038	±0.072	-0.058 to 0.017	0.440 to 0.584
Spread of jet	0.574	-0.1257	±1.130	±1.399	-0.556 to 1.704	-2.655 to 0.142

the lowest 95% confidence prediction error interval is observed for the EPR-MOGA model.

CONCLUSION

In this study, the three-dimensional CFD model and evolutionary polynomial regression – multi-objective genetic algorithm are used to simulate the thermal contaminant distribution discharged from a side channel into a cross flow. The data from previous experimental tests are used to validate the CFD model and train the EPR-MOGA model. The Boussinesq hypothesis is used to consider the buoyancy effects and the k- ϵ turbulence model is used to consider the turbulent effect in the CFD model. Three distinct mathematical relations are proposed by the EPR-MOGA model to predict the temperature distribution, predict the temperature half-thickness as the thermal cross section, and the pollutant spread on the surface. The input variables used in the EPR-MOGA model include the spatial coordinates of the studying point, temperature difference between the fluid in the side channel and the receiver fluid, the ratio of discharged fluid velocity to cross flow velocity, and the depth of channel. The results of CFD and EPR-MOGA models in terms of temperature profiles, temperature half-thickness, temperature spread on the surface, and the maximum temperature in different sections are compared with each other. The results show that both the models have reasonable results with acceptable error. However, the CFD models have less error in predicting the temperature distribution profiles (RMSE = 0.404 and MAE = 0.283 compared with RMSE = 0.472 and MAE = 0.365), and the maximum averaged temperature and its position for different sections are predicted better than the EPR-MOGA model. The EPR-MOGA model, in turn, acts better than the CFD model in predicting

the advancement of pollutant flow on the surface. Both models can predict the occupied surface by the thermal pollutant in the prediction of flow cross section in terms of temperature half-thickness. However, it can be concluded, by comparing the results of Table 5 for the CFD and EPR-MOGA models, that the EPR-MOGA model has a better accuracy than the CFD model in predicting the thermal cross section. The CFD and EPR-MOGA models are able to predict the general pattern of temperature distribution, and the results of both models are consistent with each other. Therefore, both models could be used in predicting the thermal contaminant distribution discharged from a side channel. The uncertainty analysis indicated that the EPR-MOGA model had lower mean prediction error and smaller uncertainty band than the CFD model in all predictions. The relationships achieved by the EPR-MOGA model are very useful to predict temperature profiles, temperature half-thickness, and temperature spread on the surface in practical. Each model has its own advantages and disadvantages. The CFD model has a very large computational cost in comparison to the EPR-MOGA model. The relation presented by the EPR-MOGA model is a simple formula, which could provide a great deal of information about the discharged flow by conducting very few calculations. Furthermore, the proposed EPR-MOGA model is a general one that can predict the basic flow parameters including temperature values and pollutant propagation on the flow surface without the need for the CFD model or existing experimental data. It is easy to access a wide range of parameters in a CFD model, while the only parameters that can be accessed through an EPR-MOGA model are the ones that have been considered for the output of the model. Another very important point to mention in this research, is to prove the high capability of the new data mining EPR-MOGA model in predicting the sophisticated

transferring and mixing phenomena of discharged pollutant from a side channel in the cross flow.

Since, in this paper, lateral channels with the same depth with cross flow are investigated, prediction of side thermal buoyant discharge in deep cross flow can be suggested as a direction for further studies.

REFERENCES

- Abdelwahed, M. S. T. 1981 *Surface Jets and Surface Plumes in Cross-Flows*. PhD Thesis, Department of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada.
- Abessi, O., Saeedi, M., Bleninger, T. & Davidson, M. 2012 [Surface discharge of negatively buoyant effluent in unstratified stagnant water](#). *Journal of Hydro-Environment Research* **6** (3), 181–193.
- Addad, Y., Benhamadouche, S. & Laurence, D. 2004 [The negatively buoyant wall-jet: LES results](#). *International Journal of Heat and Fluid Flow* **25** (5), 795–808.
- Ali Ghorbani, M., Kazempour, R., Chau, K. W., Shamshirband, S. & Taherei Ghazvinei, P. 2018 [Forecasting pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: a case study in Talesh, Northern Iran](#). *Engineering Applications of Computational Fluid Mechanics* **12** (1), 724–737.
- Baghalian, S., Bonakdari, H., Nazari, F. & Fazli, M. 2012 [Closed-form solution for flow field in curved channels in comparison with experimental and numerical analyses and artificial neural network](#). *Engineering Applications of Computational Fluid Mechanics* **6** (4), 514–526.
- Bashiri, H., Sharifi, E. & Singh, V. P. 2018 [Prediction of local scour depth downstream of sluice gates using harmony search algorithm and artificial neural networks](#). *Journal of Irrigation and Drainage Engineering* **144** (5), 06018002.
- Bilhan, O., Emiroglu, M. E., Miller, C. J. & Ulas, M. 2018 [The evaluation of the effect of nappe breakers on the discharge capacity of trapezoidal labyrinth weirs by ELM and SVR approaches](#). *Flow Measurement and Instrumentation* **64**, 71–82.
- Bodart, J., Coletti, F., Bermejo-Moreno, I. & Eaton, J. K. 2013 [High-fidelity simulation of a turbulent inclined jet in a crossflow](#). *Center for Turbulence Research Annual Research Briefs* **19**, 263–275.
- Bonakdari, H., Ebtehaj, I. & Akhbari, A. 2017 [Multi-objective evolutionary polynomial regression-based prediction of energy consumption probing](#). *Water Science and Technology* **75** (12), 2791–2799.
- Carter, H. H. & Regier, R. 1974 *Three Dimensional Heated Surface Jet in a Cross Flow*. Technical report 88. No. COO-3062-19. Chesapeake Bay Institute, Johns Hopkins University, Baltimore, MD, USA.
- Carter, H. H., Schiemer, E. W. & Regier, R. 1973 *Buoyant Surface Jet Discharging Normal to an Ambient Flow of Various Depths*. Technical report 81. No. COO-3062-13. Chesapeake Bay Institute, Johns Hopkins University, Baltimore, MD, USA.
- Choi, S. U. & Cheong, S. 2006 [Prediction of local scour around bridge piers using artificial neural networks 1](#). *JAWRA Journal of the American Water Resources Association* **42** (2), 487–494.
- Currie, I. G. 2002 *Fundamental Mechanics of Fluid*, 3rd edn. Marcel Dekker, New York, USA.
- Dallas, H. 2009 *The Effect of Water Temperature on Aquatic Organisms: A Review of Knowledge and Methods for Assessing Biotic Responses to Temperature*. WRC report no. KV 213/09. Water Research Commission, Gezina, South Africa.
- Donmez, S. 2011 [Using artificial neural networks for prediction of alternate depth shaped on rectangular channel in open channel flow](#). *Energy Education Science and Technology Part A-Energy Science and Research* **28** (1), 339–348.
- Eghbalzadeh, A., Javan, M., Hayati, M. & Amini, A. 2016 [Discharge prediction of circular and rectangular side orifices using artificial neural networks](#). *KSCE Journal of Civil Engineering* **20** (2), 990–996.
- Ferziger, J. H. & Peric, M. 2012 *Computational Methods for Fluid Dynamics*. Springer Science & Business Media, Berlin, Germany.
- Fossati, M., Santoro, P., Urrestarazu, S. & Piedra-Cueva, I. 2011 [Numerical study of the effect of a power plant cooling water discharge in the Montevideo Bay](#). *Journal of Applied Mathematics* **2011**, 1–23.
- Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K. W., Faizollahzadeh Ardabili, S. & Piran, M. J. 2018 [Survey of computational intelligence as basis to big flood management: challenges, research directions and future work](#). *Engineering Applications of Computational Fluid Mechanics* **12** (1), 411–437.
- Frigo, A. A. & Frye, D. E. 1972 *Physical Measurements of Thermal Discharges Into Lake Michigan, 1971*. No. ANL/ES-16. Argonne National Laboratory, Lemont, IL, USA.
- Gholami, A., Bonakdari, H., Zaji, A. H. & Akhtari, A. A. 2015 [Simulation of open channel bend characteristics using computational fluid dynamics and artificial neural networks](#). *Engineering Applications of Computational Fluid Mechanics* **9** (1), 355–369.
- Giustolisi, O. & Savic, D. A. 2006 [A symbolic data-driven technique based on evolutionary polynomial regression](#). *Journal of Hydroinformatics* **8** (3), 207–222.
- Giustolisi, O. & Savic, D. A. 2009 [Advances in data-driven analyses and modelling using EPR-MOGA](#). *Journal of Hydroinformatics* **11** (3–4), 225–236.
- Huai, W.-X., Li, Z.-W., Qian, Z.-D., Zeng, Y.-H., Han, J. & Peng, W.-Q. 2010 [Numerical simulation of horizontal buoyant wall jet](#). *Journal of Hydrodynamics* **1**, 58–65.
- Huai, W., Chen, G. & Zeng, Y. 2013 [Predicting apparent shear stress in prismatic compound open channels using artificial neural networks](#). *Journal of Hydroinformatics* **15** (1), 138–146.

- Hydraulics Research Wallingford 1985 *A Buoyant Surface jet in A Strong Crossflow*. Report No. SR 68. Hydraulics Research, Wallingford, UK.
- Jasak, H. 1996 *Error Analysis and Estimation for the Finite-Volume Method with Applications to Fluid Flows*. PhD Thesis, University of London, UK.
- Javadi, A. A. & Rezania, M. 2009 Applications of artificial intelligence and data mining techniques in soil modeling. *Geomechanics and Engineering* **1** (1), 53–74.
- Jones, G. R., Nash, J. D., Doneker, R. L. & Jirka, G. H. 2007 Buoyant surface discharges into water bodies. I: flow classification and prediction methodology. *Journal of Hydraulic Engineering* **133**, 1010–1020.
- Kalita, K., Dewan, A. & Dass, A. K. 2002 Prediction of turbulent plane jet in crossflow. *Numerical Heat Transfer: Part A: Applications* **41** (1), 101–111.
- Kheirkhah Gildeh, H., Mohammadian, A., Nistor, I. & Qiblawey, H. 2014 Numerical modeling of turbulent buoyant wall jets in stationary ambient water. *Journal of Hydraulic Engineering* **6**, 04014012.
- Kim, D. G. & Cho, H. Y. 2006 Modeling the buoyant flow of heated water discharged from surface and submerged side outfalls in shallow and deep water with a cross flow. *Environmental Fluid Mechanics* **6**, 501–518.
- Kuhlman, J. M. & Prahl, J. M. 1976 Buoyant rectangular surface thermal plumes. *Journal of Great Lakes Research* **2** (2), 340–356.
- McGuirk, J. J. & Rodi, W. 1978 A depth-averaged mathematical model for the near field of side discharges into open-channel flow. *Journal of Fluid Mechanics* **86**, 761–781.
- McGuirk, J. J. & Rodi, W. 1979 Mathematical modelling of three-dimensional heated surface jets. *Journal of Fluid Mechanics* **95**, 609–633.
- Melesse, A. M., Ahmad, S., McClain, M. E., Wang, X. & Lim, Y. H. 2011 Suspended sediment load prediction of river systems: an artificial neural network approach. *Agricultural Water Management* **98** (5), 855–866.
- Moazenzadeh, R., Mohammadi, B., Shamshirband, S. & Chau, K. W. 2018 Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran. *Engineering Applications of Computational Fluid Mechanics* **12** (1), 584–597.
- Morano, P., Tajani, F. & Locurcio, M. 2017 GIS application and econometric analysis for the verification of the financial feasibility of roof-top wind turbines in the city of Bari (Italy). *Renewable and Sustainable Energy Reviews* **70**, 999–1010.
- Motz, L. H. & Benedict, B. A. 1970 *Heated Surface jet Discharged Into A Flowing Ambient Stream*. Report 4. Department of Environmental and Water Resources Engineering, Vanderbilt University, Nashville, TN, USA.
- Nagy, H. M., Watanabe, K. A. N. D. & Hirano, M. 2002 Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering* **128** (6), 588–595.
- Najafzadeh, M., Laucelli, D. B. & Zahiri, A. 2017 Application of model tree and evolutionary polynomial regression for evaluation of sediment transport in pipes. *KSCE Journal of Civil Engineering* **21** (5), 1956–1963.
- Oliveira, P. J. & Issa, R. I. 2001 An improved PISO algorithm for the computation of buoyancy-driven flows. *Numerical Heat Transfer: Part B: Fundamentals* **40**, 473–493.
- Shih, T. H., Liou, W. W., Shabbir, A., Yang, Z. & Zhu, J. 1995 A new $k-\epsilon$ eddy viscosity model for high reynolds number turbulent flows. *Computers & Fluids* **24** (3), 227–238.
- Stefan, H. G., Bergstedt, L. M. & Mrosla, E. 1975 *Flow Establishment and Initial Entrainment of Heated Water Surface Jets*. National Environmental Research Center, Minnesota University, Minneapolis, MN, USA.
- Stoesser, T., Ruether, N. & Olsen, N. R. B. 2010 Calculation of primary and secondary flow and boundary shear stresses in a meandering channel. *Advances in Water Resources* **33** (2), 158–170.
- Stolzenbach, K. D. & Harleman, D. R. F. 1971 *An Analytical and Experimental Investigation of Surface Discharges of Heated Water*. Report 135. Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Massachusetts Institute of Technology, Cambridge, MA, USA.
- Sun, S., Yan, H. & Kouyi, G. L. 2014 Artificial neural network modelling in simulation of complex flow at open channel junctions based on large data sets. *Environmental Modelling & Software* **62**, 178–187.
- Taormina, R., Chau, K. W. & Sivakumar, B. 2015 Neural network river forecasting through baseflow separation and binary-coded swarm optimization. *Journal of Hydrology* **529**, 1788–1797.
- Teng, S., Feng, M., Chen, K., Wang, W. & Zheng, B. 2018 Effect of a lateral jet on the turbulent flow characteristics of an open channel flow with rigid vegetation. *Water* **10** (9), 1204.
- Vaillancourt, G. & Couture, R. 1978 Influence of heat from the Gentilly nuclear power station on water temperature and Gastropoda. *Canadian Water Resources Journal* **3**, 121–133.
- White, F. M. & Corfield, I. 2006 *Viscous Fluid Flow*, Vol. 3. McGraw-Hill, New York, USA.
- Wu, C. L. & Chau, K. W. 2011 Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology* **399** (3–4), 394–409.
- Yan, X. & Mohammadian, A. 2017 Numerical modeling of vertical buoyant jets subjected to lateral confinement. *Journal of Hydraulic Engineering* **143**, 04017016.
- Yaseen, Z. M., Sulaiman, S. O., Deo, R. C. & Chau, K. W. 2019 An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology* **569**, 387–408.
- Zhang, L. & Yang, V. 2017 Flow dynamics and mixing of a transverse jet in crossflow – part I: steady crossflow. *Journal of Engineering for Gas Turbines and Power* **139**, 082601.

First received 9 January 2019; accepted in revised form 22 July 2019. Available online 19 September 2019