1. Introduction

Temperature is an important factor that affects massive concrete structures’ health [1]. When the volume of a concrete structure is so large that it generates excessive heat and associated expansions [2], which may lead to a risk of temperature cracks, we can name it “massive concrete” [3, 4]. Postcooling measures, such as embedding cool-pipes, are commonly included in the design to mitigate such effects.

The first significant application case of cool-pipes was in a small experimental section of the Owyhee Dam in Oregon in 1931, which was the tallest dam in the world at the time of its completion [5]. The Bureau of Reclamation believed the Owyhee experiment was successful and extended its application in the construction of the Hoover Dam in 1936. In every slab coil of 1-inch-thick-walled steel pipe running water was embedded. More than 582-miles- (937 km) long cool-pipes were placed within the concrete [5–7]. The success of bedding steel cool-pipes in the Hoover Dam made the method as famous as the dam itself. The method has thus been utilized in a number of large-scale massive concrete projects [8–10] over the last 70 years.

Several researchers have carried out studies on the thermal effect of cool-pipes [7, 8, 11–16]. The Bureau of Reclamation first studied a theory solution of long iron pipes [7]. Zhu and Cai have made a great effort to conduct calculation for cool-pipes by using finite element methods [12, 14]. Much research has focused on the calculation model and method of cool-pipes based on assumptions (i.e., the pipe is straight; heat transfer between pipe and concrete is constant) that make the problem simple and solvable. However, during the construction of a massive concrete structure, the heat transfer process is very sophisticated. The cool-pipes are embedded in multiple U-shapes and the heat transfer between fluid (water) and solid (pipe) is very complicated. A recent study pointed out that the convection coefficient cannot be considered constant [13]. So study of in situ thermal performance of cool-pipes is needed, which can help us acknowledge the thermal mechanism of cool-pipes well.

To the knowledge of the authors, few studies have reported the statistics of the thermal performance of cool-pipes in actual concrete engineering studies, based on real measured data, especially for large-scale projects. An important reason for this scarcity is the lack of adequate monitoring
measurements and detailed field data analysis. Fortunately, a 285 m high arch dam project under construction gives a perfect opportunity to conduct the work. During the construction period of the dam, the temperature and the flow of every cool-pipe in every slab are monitored. A large amount of data is produced; we implement data mining of thermal performance of cool-pipes in this paper.

Many mathematical methods can be employed in data mining [17–27]. Support vector machine (SVM) is a novel machine learning method developed on the basis of statistical learning theory [28, 29]. It is one of the most popular methods for pattern recognition, regression estimation, and time series prediction problems [30]. A typical process of SVM training can be summarized into four steps as follows [31]: selecting kernel function, selecting the smoothing parameters of the kernel function, choosing penalty factor, and solving the quadratic problem. It has been shown that compared with other machine learning methods, such as artificial neural network, SVM not only is easier to use [32] but also leads to higher accuracy and robustness. For specific cases, it converges 10 to 100 times faster in training [33, 34].

Our paper is organized as follows: Section 2 describes the delicate thermal monitoring program, including the temperature measurement of concrete, inlet-outlet temperature, and flow of cool-pipes. Section 3 shows the establishment of the thermal model to guide the data mining. Section 4 describes the SVM method and its relevant concepts. Section 5 presents the details of the monitoring database and a developed automatic data mining program flow. Section 6 provides the data mining results, including the calibrated parameters, numerical model to predict outlet cool-pipe temperature, the evaluation of the thermal performance, and pipe material classification. And Section 7 provides our conclusions.

2. In Situ Monitoring

2.1. Engineering Description. Xiluodu Dam is located in the middle of the Jinsha River, Yunnan Province in Southwest China. It is designed as an arch dam with a maximum height of 285.5 m and a crest length at the top of 680 m [8]. Based on the project plan, the amount of concrete utilized was estimated to be 1315 × 10^3 m^3 during the construction period from 2009 to 2014. The rear view of the dam being constructed is shown in Figure 1.

Across the river flow direction, the dam is divided by 31 monoliths [35]. Vertically, each monolith consists of a number of slabs poured at different times. Figure 2 shows the slab zoning map. Each slab is individually poured in a formed hexahedron block with a vertical thickness of 3.0 m or 1.5 m. The first slab was poured on March 27, 2009, and 2158 slabs altogether had been placed by December 2013. These slabs are the main study objects of the research presented in this paper.

Cool-pipes were horizontally embedded in every slab to control the concrete temperature properly. Circulating water was supplied via two large cooling water units, which could stabilize the supply temperature at a constant level on demand. Considering that water temperature increases along the flow direction, the flow direction was changed twice a day to make the concrete temperature field more uniform. Iron pipes and high-density polyethylene (HDPE) pipes are employed in this project.

2.2. In Situ Thermal Monitoring Setup

2.2.1. Concrete Temperatures. Since the start of the dam construction, digital temperature sensors were installed during the pouring of every concrete slab. Staggered cool-pipes were vertically embedded every 1.5 m (i.e., two layers in a 3.0 m thick slab). The sensors were vertically positioned in the middle of two cool-pipe layers, as shown in Figure 3. Concrete temperature was measured and recorded for four or five times per day.

2.2.2. Cooling Water Temperatures and Flows. Figure 4 shows the schematic layout of the cool-pipes in one slab. Faucets were connected to both ends of each cool-pipe. By letting water run from the tap, the inlet and outlet temperatures of the cool-pipes were measured using a mercury thermometer and recorded from four or five times per day. At one end of every pipe, one water meter was installed (Figure 4). The cooling water flow rate was measured and recorded.

3. Thermal Model

A simplified thermal model of cool-pipe and concrete is sketched in Figure 5. Water flows through a pipe whose wall clings to the concrete. Heat transfer per unit time is discussed below.

The heat energy loss from water per unit length, which is denoted as \( Q_w \), can be expressed as follows:

\[
Q_w = \rho_w C_{p,w} dT_w \cdot u A_p,
\]

where \( \rho_w \) is the density of the water, \( C_{p,w} \) is the specific heat capacity of water, \( T_w \) is the water temperature, \( dT_w \) is the temperature difference along the pipe (per unit length), \( u \) is the water velocity, and \( A_p \) is cross section area of the pipe.

The heat energy income for the concrete per unit length, which is denoted as \( Q_c \), can be calculated by Newton's law of cooling, which is also called the Robin boundary condition

\[
Q_c = k (T_c - T_w) dA_c,
\]

where \( k \) is the heat transfer coefficient between water and concrete, \( T_c \) is the concrete temperature, and \( A_c \) is the contact area of pipe wall per unit length.
$A_p$ and $A_c$ can be further expressed by the geometry property of the pipe

$$A_p = \frac{1}{4} \pi D^2, \quad dA_c = \pi D dL,$$

where $D$ is the diameter of the pipe and $L$ is the length of the pipe.

Propose that $T_c$ is constant along the pipe and let $\Delta T$ represent $(T_c - T_w)$. The differential of $T_w$ can be regarded as the differential of $\Delta T$:

$$d\Delta T = d(T_c - T_w) = -dT_w.$$  (4)

Based on the conservation of energy, $Q_w = Q_c$. We can obtain that

$$\frac{1}{4} \pi D^2 C_{p,w} \rho_w u (-d\Delta T) = k \Delta T \pi D dL.$$  (5)

Integrating (5) along the whole pipe gives

$$\ln \frac{\Delta T_{\text{out}}}{\Delta T_{\text{in}}} = -\frac{4k}{C_{p,w} \rho_w D u} L,$$  (6)

where subscripts in and out represent the inlet and outlet of the pipe, respectively. If the inlet status of the water pipe is known, the outlet water temperature can be determined by the following equation:

$$(T_c - T_{w\text{-out}}) = (T_c - T_{w\text{-in}}) \exp \left( -\frac{4k}{C_{p,w} \rho_w D u} L \right).$$  (7)

Temperature rise along the pipe $T_{w\text{-out}} - T_{w\text{-in}}$ is an important indicator which directly gives the total energy loss of water. It can be obtained by (7):

$$T_{w\text{-out}} - T_{w\text{-in}} = (T_c - T_{w\text{-in}}) + (T_c - T_{w\text{-out}})$$
$$= (T_c - T_{w\text{-in}}) \left[ \exp \left( -\frac{4k}{C_{p,w} \rho_w D u} L \right) - 1 \right].$$  (8)

The physical properties of the water, $C_{p,w}$ and $\rho_w$, are constant in usual environment conditions. $k$ is determined by the pipe material. In engineering cases, the diameter of a pipe material is always fixed. In China’s dam projects $D$ for HDPE pipe and iron pipe are 32 mm. Let $T_{w*}$ denote
4 Mathematical Problems in Engineering

Thermometer
Concrete slab
Water
Cool-pipes

Figure 4: Schematic diagram of the cool-pipes layout.

\[ T_{w-out} - T_{w-in} \] and \( T^*_w \) can be seen as the function of some parameter as follows:

\[ T^*_w = f(\Delta T_{in}, u, k, L). \] (9)

With \( T^*_w \) known, the absorbed heat power by cool-pipe, which is denoted as \( P \), can be calculated by the following equation:

\[ P = \rho_w C_{pw} T^*_w q_w, \] (10)

where \( q_w = u A_p \) is the flow rate of the cooling water.

Actually, the thermal process between the concrete and the cool-pipe is very complex, which has some differences with the theory model proposed above:

1. Different pipe materials have different \( k \) values. The value of \( k \) is hard to determine and may change with flow states.
2. Concrete temperature is not constant along the pipe.
3. Pipe is embedded in multiple U-shapes like a long crawling snake, not in a straight line.

But (9) can give a good guide for us to conduct data mining, which tells us that \( T^*_w \) is relevant with the inlet temperature difference \( (\Delta T_{in}) \), cooling water flow speed \( (u) \), pipe length \( (L) \), and the pipe material \((k)\).

4. Data Mining Model

4.1 SVM Classification. For a classification problem of two classes, a set of sample data \( S \) is given as

\[ S = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n)\}, \] (11)

where \( x_i \in \mathbb{R}^d \) \( (i = 1, 2, \ldots, n) \) is a \( d \)-dimensional input vector and \( y_i \in \{-1, 1\} \) \( (i = 1, 2, \ldots, n) \) is the classification label. The object is to find a function \( f(x) \in \mathbb{R}^d \) that can separate the two classes with a margin as large as possible. Linear estimation function to achieve the goals can be expressed as follows:

\[ f(x) = w \cdot x + b \quad \text{with} \quad w \in \mathbb{R}^d, \ b \in \mathbb{R}, \] (12)

where \( w \) and \( b \) are weight factors to be determined. Many possible linear classifiers exist that can separate the data. As shown in Figure 6, \( H^* \) is the classifier plane, and \( H1 \) and \( H2 \) are the planes that cross the nearest data point and parallel to the classifier plane (the sample points on \( H1 \) and \( H2 \) can be called support vector). The basic idea of SVM is to find an optimal plane that maximizes the classification margin
between $H_1$ and $H_2$. One way to ensure this is to minimize the norm of $w$ [36], that is, $\|w\|^2$:

Minimize $\frac{1}{2}\|w\|^2$

subject to:

$$f(x_i) > 0, \quad \text{when } y_i = 1, \ i = 1, 2, \ldots, n$$

$$f(x_i) < 0, \quad \text{when } y_i = -1, \ i = 1, 2, \ldots, n.$$  \hspace{1cm} (13)

The constraint condition of (13) can be expressed as

$$y_i f(x_i) \geq 1, \quad i = 1, 2, \ldots, n.$$  \hspace{1cm} (14)

In particular cases, data of different labels cannot be clearly separated. A set of slack variable $\{\xi_i\}$ is introduced that measure the amount by which the constraints are violated, as shown in Figure 6. Equation (13) is then transformed into the following form:

Minimize $\frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i$

subject to:

$$y_i f(x_i) \geq 1 - \xi_i, \quad i = 1, 2, \ldots, n$$

$$\xi_i \geq 0, \quad i = 1, 2, \ldots, n.$$  \hspace{1cm} (15)

where $C$ is a penalty constant chosen a priori, which determines the cost of constraint violation. As is well known, (15) is a quadratic programming optimization problem. The method of Lagrange multipliers is a good strategy to find the solution of an optimization problem considering the constraint conditions. Solving (15) can be converted into finding the saddle point of the following Lagrange function:

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i$$

$$- \sum_{i=1}^{n} \alpha_i \left[ y_i f(x_i) - 1 + \xi_i \right] - \sum_{i=1}^{n} \eta_i \xi_i,$$  \hspace{1cm} (16)

where $\alpha_i \geq 0, \eta_i \geq 0$ is the Lagrange factor. The saddle points can be obtained by

$$\frac{\partial L}{\partial w} = 0 \implies w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = 0 \implies \sum_{i=1}^{n} \alpha_i y_i = 0$$  \hspace{1cm} (17)

$$\frac{\partial L}{\partial \xi_i} = 0 \implies C - \alpha_i - \eta_i = 0$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \implies \left[ y_i f(x_i) \right] - 1 + \xi_i = 0.$$

Substituting (17) into (16) yields the dual optimization problem:

$$\max_{\alpha_i} Q(\alpha_i) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i y_i x_i)^T \cdot (\alpha_j y_j x_j)$$

subject to: $0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, n$

$$\sum_{i=1}^{n} \alpha_i y_i = 0.$$  \hspace{1cm} (18)

The solution $\alpha_i^*$ for (18) can determine the optimal parameter $w^*$:

$$w^* = \sum_{i=1}^{n} \alpha_i^* y_i x_i.$$  \hspace{1cm} (19)

With $w^*$, the optimal parameter $b^*$ can be obtained from the boundary condition in (17), which can be written as follows:

$$\left[ y_i (w^* x_i + b^*) \right] - 1 + \xi_i = 0.$$  \hspace{1cm} (20)

The optimal hyperplane decision function can be finally written as

$$f(x) = \text{sgn}(w^* \cdot x + b^*).$$  \hspace{1cm} (21)

4.2. SVM Regression. Regression can also be performed by SVM. The goal is to find a function $f(x)$ that can achieve the targets for all data points and is as flat as possible. By flat, we mean a small $w$, which can generate a problem as follows:

Minimize $\frac{1}{2}\|w\|^2$

subject to: $|y_i - f(x_i)| \leq \varepsilon, \quad i = 1, 2, \ldots, n,$

where $\varepsilon$ is a loss precision, which is employed to create an intensive band near the regression line, as shown in Figure 7. A double-set of slack variable $\{\xi_i^{up}\}, \{\xi_i^{down}\}$ is introduced to
make the constraint more feasible. Equation (13) is then transformed into the following form:

\[
\begin{align*}
\text{Minimize} & \quad \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} (\varepsilon_{i}^{up} + \varepsilon_{i}^{down}) \\
\text{subject to:} & \quad y_{i} - f(x) \leq \varepsilon + \varepsilon_{i}^{up}, \quad i = 1, 2, \ldots, n \\
& \quad f(x) - y_{i} \leq \varepsilon + \varepsilon_{i}^{down}, \quad i = 1, 2, \ldots, n \\
& \quad \varepsilon_{i}^{up}, \varepsilon_{i}^{down} > 0.
\end{align*}
\]  
(23)

Similar to classification, (23) can be converted into a Lagrange function

\[
L(w, b, \alpha) = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} (\varepsilon_{i}^{up} + \varepsilon_{i}^{down}) \\
- \sum_{i=1}^{n} \alpha_{i}^{up} \varepsilon_{i}^{up} + \varepsilon - [y_{i} - f(x)] \\
- \sum_{i=1}^{n} \alpha_{i}^{down} \varepsilon_{i}^{down} + \varepsilon - [f(x) - y_{i}] \\
- \sum_{i=1}^{n} (\eta_{i}^{up} \varepsilon_{i}^{up} + \eta_{i}^{down} \varepsilon_{i}^{down}),
\]  
(24)

where \(\alpha_{i}^{up} \geq 0, \alpha_{i}^{down} \geq 0, \eta_{i}^{up} \geq 0, \eta_{i}^{down} \geq 0\) is the Lagrange factor. The saddle point of (24) is the solution of (23), which can be obtained by

\[
\begin{align*}
\frac{\partial L}{\partial w} = 0 \implies w = \sum_{i=1}^{n} (\alpha_{i}^{down} - \alpha_{i}^{up}) x_{i} \\
\frac{\partial L}{\partial b} = 0 \implies \sum_{i=1}^{n} (\alpha_{i}^{up} - \alpha_{i}^{down}) = 0
\end{align*}
\]

The solution \(\alpha_{i}^{up,*}, \alpha_{i}^{down,*}\) for (25) can determine the optimal parameter \(w^*\),

\[
w^* = \sum_{i=1}^{n} (\alpha_{i}^{up,*} - \alpha_{i}^{down,*}) x_{i}.
\]  
(27)

With \(w^*\), the optimal parameter \(b^*\) can be calculated from the boundary condition in (25). Then the optimal regression function can be finally obtained.

### 4.3. Kernel Function.

Actually the algorithms in Sections 4.1 and 4.2 are specified in linear space. Either in classification or regression problem, the input data \(x\) is projected into \(y\) by function \(f\), which is built during training. Furthermore, \(f(x) = w \cdot x + b\) can be expanded into \(f(x) = w \cdot \varphi(x) + b\) to solve more complicated problems, such as nonlinear.
problems. Thus, formulas (18) and (26) can be rewritten as follows:

\[
\max_{\alpha_i} \quad Q(\alpha_i) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi(x_i)^T \cdot \varphi(x_j) \\
\text{subject to:} \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, n \quad (28)
\]

where \( V \) is a positive integer. Every subset becomes the test set once, so that the training results get closer to the distribution of original data without much increase in computing cost.

Algorithm 1: Grid-search \( V \)-fold CV method.

5. Data Processor and Programming

5.1. Database. We established a database to store quantities of monitoring data. Microsoft SQL Server (in Version 10.50.1600) is applied as the database engine. All the monitored data introduced in Section 2 were stored in the database.

5.2. Data Cleaning. Data cleaning is carried out as a preprocessing of data mining; some work is done as follows.

(1) Monitored data that has a character of instrument failure is filtered out.

(2) The flow direction is changed every day, and each pair of \( T_{w-out} \) and \( T_{w-in} \) are automatically compared to guarantee that \( T_{w-out} > T_{w-in} \).

(3) Pipe length is not studied as a variable in the data mining process, for most pipes are 220 m long in the engineering. So for all the pipes, the length is considered as a constant and not selected during the data mining procedure.

(4) HDPE pipe’s diameter is identical as, in the iron pipe in the engineering, flow rate \( q_w \) is adopted in data mining instead of the water speed \( u \) for convenience.

5.3. Data Extractor. A set of programs is developed in Python language to seek the database and export data samples. Given the massive stored data, the program is developed in parallel to accelerate the processing speed.

Taking data mining on \( T_{w-out} \) as an example, the relevant data include pipe material, inlet and outlet water temperature, water flow, and the concrete temperature. They are stored in three different data tables, the structures of which are shown in Figure 8. Data extraction is carried out based on Algorithm 2.

5.4. Data Mining Implementation. The SVM method is developed in Python programs to model the thermal performance. Based on the integration of database, data extractor, and SVM model engine, the whole program flow of data mining is shown in Figure 9.
(1) Select * from TbPipeInfo where Material = "HDPE", record them as Pipes.
(2) for each Pipe in Pipes do:
(3) Select * from TbSensorInfo where SlabId = Pipe.SlabId, record them as Sensors.
(4) Select * from TbCoolRecord where Id = pipe.PipeId, record them as Rows.
(5) for each Row in Rows do:
(6) for each Sensor in Sensors do:
(7) Select Temperature from TbTemRecord
where TemTime nearest * to Row.CoolTime
and SensorId = Sensor.SensorId
(8) end loop Sensors
(9) \( T_c = \text{average}(\text{sum}(\text{Temperature})) \)
(10) output \( T_c, T_{w-in}, T_{w-out}, \text{Flow} \)
(11) end loop Rows
(12) end loop Pipes

Note *: all the temperature sensors’ reading in the slab are collected. If the nearest time difference is larger than a threshold, the data of the time will not be used.

Algorithm 2: Data extraction flow (HDPE pipes as examples).

Figure 8: Data structures.

6. Results and Discussions

6.1. Parametric Analysis. The SVM coefficients including \( \varepsilon \), \( C \), and \( y \) need to be determined before utilizing the method. To the knowledge of authors, no determinate function exists that directly obtains the optimal coefficients for a model. So, parametric analysis is carried out to calibrate the model.

Taking prediction of outlet temperature as a study case, \( T_{w}^* \) is employed as the target variable whereas \( \Delta T_{in} \) and \( q_{w} \) are adopted as the observed variables. The data of cool-pipes embedded in monolith 15# is used as the training dataset, and monolith 16# is used as the testing dataset. The grid search method integrated with \( V \)-fold CV is used here to find optimal parameters, as mentioned in Section 4.4. The mean squared error is applied to evaluate the model performance, which is expressed as

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} [f(x_i) - y_i]^2, \tag{31}
\]
where \( x \) stands for the combination of \( \Delta T_{\text{in}} \) and \( q_w \) and \( y \) stands for \( T^*_w \). A small MSE indicates a good fitness of the numerical model. The ranges of \( C \) and \( \gamma \) are both set to be \((2^{-5} \sim 2^5)\). The results of MSE of different \( \epsilon \) are drawn in logarithmic plots with \( \log_2(C) \) and \( \log_2(\gamma) \) as the \( x \)- and \( y \)-axis, respectively.

As shown in Figure 10, the model has a very well fit effect when \( \gamma \) and \( C \) are large for training data set. But it performs relatively poor for the testing data set. This situation is called the overtraining phenomenon. Optimal SVM coefficients are chosen as the one that fits the testing data set best, in order to avoid the phenomenon. The optimal results are listed in Table 1.

### Table 1: Optimal SVM coefficients.

<table>
<thead>
<tr>
<th>Material</th>
<th>( \epsilon )</th>
<th>Training data set</th>
<th>Testing data set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( C ) ( \gamma ) MSE</td>
<td>( C ) ( \gamma ) MSE</td>
<td></td>
</tr>
<tr>
<td>HDPE</td>
<td>0.1 32 32 0.00915854</td>
<td>0.25 8 0.006821</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.01 32 32 0.00916964</td>
<td>2 8 0.006485</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001 16 32 0.00917025</td>
<td>0.5 4 0.003299</td>
<td></td>
</tr>
<tr>
<td>Iron</td>
<td>0.1 32 32 0.00616273</td>
<td>8 0.5 0.003472</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.01 32 32 0.00585564</td>
<td>0.25 4 0.003036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001 16 32 0.00586528</td>
<td>0.5 4 0.003299</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001 16 32 0.00587013</td>
<td>0.0625 8 0.003301</td>
<td></td>
</tr>
</tbody>
</table>

Optimal results:

- HDPE: \( \epsilon = 0.01 \) \( C = 2 \) \( \gamma = 8 \)
- Iron: \( \epsilon = 0.001 \) \( C = 0.5 \) \( \gamma = 4 \)

6.2. Prediction of Water Temperature Rise. As written in Section 3, the water temperature rise of cool-pipe, \( T^*_w \), is relevant to the inlet temperature difference and cooling flow rate. We employ the data of monolith 15# as the training data set to model the cooling performance. All the other slabs are used as testing cases. The predicted outlet temperature agrees well with the monitoring result. Four slabs are taken as examples in Figure 11.

6.3. Thermal Performance Comparison. The pattern of thermal performance of cool-pipes can be found in the numerical model that has mined the data mentioned above. The numerical results of \( T^*_w \) at different combination of \( \Delta T_{\text{in}} \) and \( q_w \) are listed in Table 2. Three-dimensional visualization results of HDPE and iron pipes are also compared in Figure 12. From the results, conclusions can be given as follows.

1. Outlet water temperature of HDPE pipes decreases with the water flow rate in a lower flow range, which is less than 30 L/min. But in a higher flow range, say more than 40 L/min, the outlet temperature is intensive with the flow rate.

2. Outlet water temperatures for iron pipes have similar phenomenon with HDPE pipes in a lower flow rate range. And an obvious turning point can be found in the result. \( T^*_w \) increases with \( q_w \) when \( q_w \) is
Figure 10: Mean squared errors of different parameter combinations.
Figure 11: Prediction outlet temperature and monitoring temperature.

Figure 12: 3D results of the numerical model.
Table 2: Result of $T_w^*$ (°C) output by the numerical model.

<table>
<thead>
<tr>
<th>$q_w$ (L/min)</th>
<th>2.0</th>
<th>4.0</th>
<th>6.0</th>
<th>8.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.69 (2.28)*</td>
<td>3.36 (4.43)</td>
<td>5.11 (6.65)</td>
<td>6.76 (8.41)</td>
</tr>
<tr>
<td>10</td>
<td>1.58 (1.93)</td>
<td>3.24 (3.82)</td>
<td>4.41 (5.74)</td>
<td>5.65 (7.25)</td>
</tr>
<tr>
<td>20</td>
<td>1.17 (1.47)</td>
<td>2.75 (2.93)</td>
<td>3.36 (4.29)</td>
<td>4.17 (5.30)</td>
</tr>
<tr>
<td>30</td>
<td>0.86 (1.27)</td>
<td>2.23 (2.44)</td>
<td>2.90 (3.38)</td>
<td>3.73 (3.99)</td>
</tr>
<tr>
<td>40</td>
<td>0.79 (1.18)</td>
<td>1.93 (2.15)</td>
<td>2.83 (2.86)</td>
<td>3.75 (3.23)</td>
</tr>
<tr>
<td>50</td>
<td>0.80 (1.05)</td>
<td>1.83 (1.87)</td>
<td>2.81 (2.46)</td>
<td>3.61 (2.77)</td>
</tr>
<tr>
<td>60</td>
<td>0.79 (0.87)</td>
<td>1.80 (1.50)</td>
<td>2.64 (2.02)</td>
<td>3.18 (2.38)</td>
</tr>
<tr>
<td>70</td>
<td>0.87 (0.70)</td>
<td>1.72 (1.12)</td>
<td>2.29 (1.58)</td>
<td>2.63 (2.03)</td>
</tr>
<tr>
<td>80</td>
<td>1.02 (0.72)</td>
<td>1.59 (0.93)</td>
<td>1.91 (1.31)</td>
<td>2.23 (1.81)</td>
</tr>
<tr>
<td>90</td>
<td>1.01 (1.03)</td>
<td>1.45 (1.08)</td>
<td>1.78 (1.37)</td>
<td>2.15 (1.84)</td>
</tr>
<tr>
<td>100</td>
<td>0.76 (1.62)</td>
<td>1.45 (1.58)</td>
<td>2.03 (1.78)</td>
<td>2.31 (2.16)</td>
</tr>
</tbody>
</table>

The number enclosed in brackets is the iron result, and the number outside of the brackets is the HDPE result.

6.4. Classification of Different Pipe Materials. The pipe material determines the value of $k$ in Section 3. In order to correctly classify the pipe materials, $\Delta T_{in}$, $T_{w}^*$, and $q_w$ are adopted as the variables, and iron/HDPE is used as the classification label. The data of monolith 15# is used as the training data set and 108 slabs are used as the testing data set. A large number of data exist for a slab as the product of the monitoring program. The labels of HDPE pipes are defined as 0 and iron ones are defined as 1. Every day’s data of the slab is tested and labeled by the SVM classifier, and the following equation is used to estimate the pipe material of the slab:

$$Material = \begin{cases} 
    \text{HDPE} & \text{when } \text{avg} < 0.4 \\
    \text{iron} & \text{when } \text{avg} > 0.6 \\
    \text{alternative} & \text{otherwise}
\end{cases}$$

where avg is the average value of the label values. The classification result is shown in Figure 15. The detailed classification results of two slabs are plotted as examples, as shown in Figure 16.

From the classification result, we can see that the SVM model can classify the different pipe materials well, with accuracy at 83%. The HDPE pipes have higher accuracy than iron pipes. The result shows that the classification by SVM method is effective.
7. Conclusions

The heat transfer process inside the cool-pipe embedded concrete is very complex. Most current researches focus on the theory calculations. Statistical study on the actual measurement is limited. Based on a delicate monitoring program during the construction period of a super high arch dam, data mining on the thermal performance of the cool-pipes is carried out. SVM is applied as the data mining method. With the direction of thermal model built in this paper, the relative factors of outlet temperature of cool-pipes are determined and the relationship is numerically mined. The prediction result has a good agreement with the monitoring data, which verifies the validation of the approach proposed in this paper. The thermal performances between iron pipes and HDPE pipes are also analyzed in detail in this paper. Iron pipes have better performance when flow rate is low, and HDPE have better performance when flow rate is large. Iron pipes have a turning flow rate of 80 L/min. The classification of pipe material is also conducted in this paper, which has correctly automatically distinguished HDPE pipes and iron pipes from monitoring data.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
Acknowledgments

This work is supported by the National “973” Researching Project of China (no. 2013CB035902), Research Project of State Key Laboratory of Hydroscience and Engineering of Tsinghua University (no. 2012-KY-4), and National Nature Science Foundation of China (nos. 51279087 and 51339003).

References


[34] F. Mosteller and J. W. Tukey, Data Analysis, including Statistics, 1968.


Submit your manuscripts at http://www.hindawi.com