Intelligent Approaches for Predicting Failure of Watermains

Zainab Almheiri[1], Mohamed Meguid[2] and Tarek Zayed[3]

ABSTRACT

Watermains are indispensable infrastructures in many countries around the world. Several factors may be responsible for the failure of these essential pipelines and negatively impact their integrity and service life. The purpose of this study is to propose models that can predict the average time to failure of watermains by using intelligent approaches, including artificial neural network (ANN), ridge regression (l_2), and ensemble decision tree (EDT) models. The developed models were trained by using collected data from Quebec City watermains, including records of the possible factors, such as the materials, length, and diameter of pipes, that contributed to the failure. The ensemble learning model was applied by using a boosting technique to improve the performance of the decision tree model. All models, however, were able to predict reasonably the failure of watermains. A global sensitivity analysis (GSA) was then conducted to test the robustness of the model and to show clearly the relationship between the input and output of the model. The GSA results show that gray cast iron (CI), hyprescon/concrete (Hy) and ductile iron with lining (DIL) are the most vulnerable materials for the model output. The results also indicate that the failure of watermains mostly depends on pipe material and its length. It is hoped that this study will help decision makers to avoid unexpected watermain failure.

Keywords: Watermains; Intelligent approaches; Pipe failure modelling; Global sensitivity analysis.

[1] PhD Candidate, Dept. of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada. Email: zainab.almheiri@mail.mcgill.ca

[2] Professor and Chair, Dept. of Civil Engineering and Applied Mechanics, McGill University, Montreal, Quebec, Canada. Email: mohamed.meguid@mcgill.ca.

[3] Professor, Building and Real Estate Dept (BRE), The Hong Kong Polytechnic University, Hong Kong. Email: tarek.zayed@polyu.edu.hk.

INTRODUCTION

Watermains are crucial infrastructures to cities around the world, and their failure can lead to significant economic, environmental, and social losses (Ana and Bauwens 2010). Watermains are expected to have a service life of between 50 and 100 years (Li et al. 2013; Ormsby 2009). More than 300 breaks of large water utilities could occur in a year and thus require urgent repair or replacement. The failure of both small and large diameters of watermains could cost millions of dollars for replacement and damage mitigation resulting from the flooding and repair of averaging roads (Makar et al. 2001).

The deterioration of watermains can be categorized into two types: structural and inner surface deterioration (Kleiner and Rajani 2001). Structural deterioration leads to the diminishing of the ability and resilience of watermains to carry loads (Kleiner and Rajani 2001). The deterioration of the inner surfaces of the pipes may result in the degradation of the water quality, in the diminishing of the hydraulic capacity, and in the reduction in the ability of the pipes to withstand internal corrosion. Both types of deteriorations can contribute to the failure of water distribution networks (Kleiner and Rajani 2001).

Ductile iron, cast iron, asbestos cement, and polyvinyl chloride are among the popular materials used in pipelines. Gray cast iron (CI) is commonly used as pipe material for watermains in North America (Makar et al. 2001; Sipos 2010), covering about 50% of the total installed watermains (Kleiner and Rajani 2001). Whereas 11% of the watermains are made of polyvinyl chloride (PVC) in North America (Rajani and Kleiner 2001). Kettler and Goulter (1985) found that the failure rate of watermains made of asbestos-cement and cast-iron pipes increases with time. A significant relation was also found between pipe diameter and the number of breaks (Christodoulou 2011; Kettler and Goulter 1985; Yamijala et al. 2009).

51　According to a water infrastructure report card published by the American Society of Civil

52　Engineering (ASCE) in 2017, the water systems in the US have attained an overall grade of "D".

53　This means that the water network is operating below standards and is furthermore in poor to

54　fair condition. Over two trillion gallons of treated drinking water is wasted owing to an estimated

55　240,000 annual watermain breaks (ASCE 2017). As reported by the American Water Works

56　Association (AWWA) in 2016, $1 trillion is needed for the maintenance of water infrastructure

57　to meet the future demand for water (ASCE 2017). Similarly, the Canadian Infrastructure Report

58　Card revealed that 23% of the water infrastructure is in poor to fair condition (CIRC 2016). As

59　reported by Folkman (2018), the failure rate of watermains is increasing exponentially over time

60　with the majority of failures occurring in gray cast iron (CI) pipes, whereas the lowest failure

61　rate occurs in polyvinyl chloride (PVC) pipes.

62　In Canada, the investment in pipe maintenance between 1985 and 2006 has increased from $57

63　billion to $125 billion, respectively (Mirza 2007) . Rehabilitation and replacement (R&R) of

64　watermains are ongoing activities owing to aging caused by operational and environmental

65　factors, such as climate change and temperature (Rogers and Grigg 2009).

66　**Purpose of the Study**

67　The physical mechanisms that lead to the failure of watermains are complex and not fully

68　understood. Therefore, the main objectives of the present study are: i) to develop prediction

69　models for determining the average time to failure; and ii) to assess the sensitivity of the

70　prediction model to different factors using GSA.

71

LITERATURE REVIEW

There are several factors that can negatively impact the condition of buried pipes. In structural and geotechnical studies, Kamel and Meguid (2012) found that loss of contact between pipelines and the surrounding soil can increase the contact pressure and consequently stresses in the pipe material. Kaddoura and Zayed (2018) examined the effect of erosion voids around sewer pipelines as such voids may reduce the service lives of the pipelines. They identified five factors that contribute the most to the occurrence of voids around sewer pipelines—soil type, bedding type, pipeline depth, pipeline age, and water table.

Kishawy and Gabbar (2010) concluded that the integrity of pipelines has to meet the increasing pressure of demand on the pipelines. They summarized the factors that could threaten the integrity of subsurface pipelines as (1) Incorrect operation, (2) Pipe material, (3) Corrosion and cracking mechanics, (4) Earth forces, such as earthquakes and landslides, (5) Weather-related factors, such as rough seas, high winds and temperature. Thus, the design of pipelines should not only rely on pressure and stress criteria but also on other indispensable factors responsible for pipeline integrity.

Evaluating the factors that affect the condition of watermains is essential for owners to meet the following criteria: (1) Developing strategies that mitigate the likelihood of pipe failure; (2) the cost that result from pipe failure, such as flooding and traffic disruption; (3) Avoiding the early replacement of a pipe, that is, before the end of its economic life. Deterioration of watermains can result from static and dynamic factors that may cause long and/or mid-term failure of water distribution systems. Fig. 1 illustrates static, dynamic and operational factors that affect the failure of watermains in both short and long terms. Static factors, such as pipe material, diameter, and age, are fixed over time, whereas dynamic factors including environmental and operational

95   factors, such as soil moisture, temperature, landslide, external stresses, and water pressure,

96   potentially chang over time (Farmani et al. 2017; Kleiner and Rajani 2001; Wang et al. 2009).

97   In addition, there are physical mechanisms that can lead to pipe breakage and failure, such as

98   (1) internal loads resulting from operational pressure and external loads due to traffic, soil

99   overburden, and frost loads; and (2) structural properties of the pipe, pipe-soil interaction, and

100   quality of installation (Rajani and Kleiner 2001). Manufacturing flaws, excessive forces, and

101   human errors have been found to have significant impact on the performance and service life of

102   cast-iron watermains (Makar et al. 2001).

103   <Fig. 1>

104

105   **Failure Prediction Models of Watermains**

106   *Physical, Statistical, and Machine Learning Models*

107   Previous studies extensively implemented physical and statistical models to examine the

108   condition of utility pipes, such as watermains. In physical models, pipe failure was determined

109   by measuring the response under a variety of environmental or other loading conditions that

110   would negatively affect the integrity of the pipe structure. Physical models do not require a huge

111   amount of historical data to develop. However, they require material properties, which are

112   generally not readily available (Wilson et al. 2017).

113   Statistical models, on the other hand, embody a set of assumptions based on observed data that

114   are generally obtained from many cases that describe the pipe behavior. They can be

115   implemented in a case study using various levels of input data for distribution watermains

116   (Kleiner and Rajani 2001). As reported by Mavin (1996), the main fragility of statistical models

117   is that "while failures generally increase with pipe age, there is a wide variation of individual

118   assets" where "links appear with soil type and weather fluctuations".

119   In recent years, machine learning has become more favorable in the prediction of pipe failure.

120   Algorithms can be developed to help decision makers to identify pipes that require urgent

121   attention or those in need of immediate replacement (Liu and Kleiner 2014; Zhou and Chen

122   2018). Machine learning could be applied to predict the failure of watermains with the growth

123   of population and aging of pipelines (Kettler and Goulter 1985; Zhou and Chen 2018). Data

124   mining is defined as the process of extracting new and useful information from big databases.

125   Machine learning is the technical basis of data mining that is used for finding, studying and

126   describing constitutional patterns in data (Witten et al. 2016). It is also a subset of artificial

127   intelligence (AI) that emerged from computer science focusing on the study of algorithms

128   (Bishop 2006). Several articles have been published on the use of statistical and machine

129   learning models to predict the failure of pipelines, such as watermains. Table 1 presents a

130   summary of some of the most recent failure prediction models of watermains and variables were

131   considered in the model. Yamijala et al. (2009) developed statistical models that are multilinear

132   regression, multivariate exponential regression, and logistic generalized linear model (GLM) to

133   estimate the likelihood of pipe breaks on the basis of several factors—pipe diameter, material,

134   length, land use, temperature, soil moisture, and soil type. Results showed that GLM performed

135   better than other models. In another work, Christodoulou et al. (2011) developed possession and

136   Cox statistical regression models based on various factors, such as pipe age, material, incident

137   type, diameter, and number of breaks, in order to predict the failure rate. Francis et al. (2013)

138   built a model that contained several factors (pipe material, diameter, age, demographic variables,

139   and temperature) to predict pipe break by using Bayesian Belief Networks (BBNs) of

140   watermains on the basis of historical data for a large city in the Mid-Atlantic United States.

141   Shirzad et al. (2014) suggested that support vector machine (SVM) over artificial neural network

142   (ANN) could be used to predict the failure rate of watermains. Both models were developed on

143   the basis of several factors: (1) hydraulic pressure, (2) pipe diameter, (3) length, (4) age, (5) and

144   depth. Kabir et al. (2015) conducted a study to predict the failure rate of watermains using

145   Bayesian and regression models based on number of breaks, pipe age, diameter, length, soil

146   resistivity, and soil corrosivity. Results proved that the Bayesian model performed better than

147   the regression models when limited data were available. Demissie et al. (2017) developed

148   dynamic Bayesian network (DBN) to predict pipe break on the basis of pipe length, diameter,

149   number of previous failures, type of service connection, freezing index, thawing index, rainfall

150   deficit, and soil corrosion. Farmani et al. (2017) studied static and dynamic factors that impact

151   the condition of watermains. The dataset was divided into homogeneous groups according to

152   the similarity in watermain features and following K-mean clustering approach. The

153   Evolutionary Polynomial Regression (EPR) model was then developed to predict the number of

154   failures based on soil type, diameter, and age. Kaushik et al. (2017) developed a model using

155   logistic regression model to predict pipe breaks using static features, such as pipe diameter,

156   length, and pipe material. The study was applied to a real-world dataset obtained from a water

157   utility in Europe. Sattar et al. (2019) suggested the use of the extreme learning machine (ELM)

158   to predict the failure of watermains based on historical data from the Greater Toronto Area,

159   Canada. The model was trained by using 9500 instances of failure records based on pipe length,

160   material, pipe protection method, and diameter. Using a decision tree to predict the time of the

161   next watermain failure, Sinder and McBean (2018) developed a gradient-boosting algorithm.

162   The researchers trained the model on a dataset derived from water utility for ductile iron pipes

163 in North America. The results suggested the use of gradient boosting algorithm when dealing

164 with a large dataset of pipe failure to predict the time of the next watermain failure. Winkler et

165 al. (2018) presented the EDT-based model using historical data from Austria to predict

166 watermain failure. The boosting technique was used to improve the model performance, and

167 50% of the available data were utilized to train the model based on various factors that are

168 physically, geographically, and historically derived. Some researchers have claimed that there

169 is a significant relationship between pipe diameter and number of breaks (Christodoulou 2011;

170 Kettler and Goulter 1985; Yamijala et al. 2009).

171 <Table 1>

172

173 **The Motivation to the Study and Research Gaps**

174 The failure of watermains is considered as a serious problem in North America and has been

175 studied for the last 40 years through the application of a variety of methods. Through the use of

176 physical and statistical models, extensive research has been conducted to study the conditions

177 and failure of watermains, respectively. However, in recent years, machine learning has become

178 more favorable in the prediction of watermain failure and can be implemented to assess the

179 future state of watermains whenever proactive maintenance is needed. A prediction model is

180 one of the fundamental components that helps decision makers to determine the best

181 maintenance plans and to prioritize the rehabilitation actions. Computational intelligence, on the

182 other hand, promotes high-level awareness of system conditions using historical records and

183 sensory data. Using statistical models, previous studies have focused on the failure prediction

184 of watermains with small or large diameters. But, as regards the techniques of AI, such as

185 ensemble modeling, in the prediction of watermain failure, research efforts remain limited,

186   hence the need for this study. Therefore, three models are developed; i) to predict the failure of

187   various types of materials and different sizes of watermains; and ii) to choose the best model in

188   terms of performance. These models are: ensemble decision tree, artificial neural network, and

189   ridg regression. In general, ensemble model performs better than a single model (Drucker 1997).

190   In previous studies, ensemble models were applied in classification-based models to classify

191   whether there is failure or not , according to Table 1, refer to Winkler et al. (2018). However,

192   in this study the ensemble decision tree is of a regression-based model that aims to predict the

193   average time of watermain failure which is a continuous variable. Whereas, ridge regression is

194   developed as it performs better than linear regression model in case of multicollinearity between

195   predictors (Khalaf and Shukur 2005). In other words, ridge regression is a form of linear

196   regression with a penalty term that is imposed to address multicolinearity as well as to improve

197   model performance. The ANN, on the other hand can capture nonlinear correlation between

198   predictors and the model output. This study is an attempt to use machine learning to predict

199   watermain failure.

200

201   **Modeling Techniques**

202   The study mainly utilizes three main intelligent approach schemes. These schemes are as

203   follows: Artificial Neural Network (ANN), Ridge Regression and Ensemble Decision Tree

204   (EDT).

205   *Artificial Neural Network (ANN)*

206   ANNs are tools that have been widely utilized in solving many complex real-world problems.

207   Feedforward network is extensively applied to recognize patterns and learn from their

208   interactions with the environment. The architecture of multilayer perceptron (MLP) of

209   feedforward neural network includes a set of artificial neurons within the input and output layers.

210   Intermediate layers of neurons, on the other hand, are placed between the input and output

211   layers. They are called hidden layers as they contain hidden neurons and do not interact with the

212   external environment (Basheer and Hajmeer 2000; Jafar et al. 2010). These neurons calculate a

213   weighted sum of the input signals (I) and then pass it (the sum) into an activation function,

214   according to Eq. (1), where (x_i) for $i = 1, 2, ..., I$, is the input signal, w_i is the estimated

215   parameter; (f) is the activation function; (α_0) is the bias/intercept (Haykin 1994).

216
$$y = f\left(\sum_{i=1}^{I} w_i x_i + \alpha_0\right) \qquad (1)$$

Ridge Regression Models

218   The main purpose of the simple linear regression analysis is to predict the relationship between

219   one or more continuous target variables, given a set of input (x_i). A linear combination of these

220   variables is built by fitting a linear relationship between target/dependent variable (y) and

221   predictor/independent variables (x_i), as per Eq. (2) (Tibshirani 1996). The linear combination

222   is called linear regression model, which is a linear function of the estimated parameters(w_i)

223   and input variables (x_i). Althought, some predictors are linearly correlated while others are

224   nonlinearly correlated with the model output; linear regression remains an efficient model that

225   has been widely applied to predict the failure of watermains. The least-square method can be

226   applied to calculate the estimated weights, as expressed in Eq. (3). Ridge regression (l_2) is a

227   linear regression, where constraints (λ) are set to control the size of the weights in order to

228   reduce the variance between actual and predicted values, Eq. (4). Solution are indexed by tuning

229   the parameter (λ), where (λ) is chosen for whichever coefficients are not rapidly changing and

230   have meaningful signs (Tibshirani 1996; Zou and Hastie 2005).

231
$$y\,(x,w) = \alpha_0 + \sum_{i=1:n} w_i x_i \qquad (2)$$

232
$$w = (x^T x)^{-1}\, x^T y \qquad (3)$$

233
$$w_\lambda^{ridge} = (x^T x + \lambda\, I_P)^{-1}\, x^T y \qquad (4)$$

234

Ensemble Decision Tree-based Model

236 Decision tree models are employed to predict the deterioration of pipelines because of the

237 simplicity of the method in creating trees that help with decision making (Winkler et al. 2018).

238 The ensemble model is applied to improve the accuracy of the prediction model and reduce the

239 prediction error by using multiple predictors. The two well-known techniques of ensemble

240 learning are called bagging and boosting. Both can be used to build a committee of regressors

241 that may be superior to a single regressor. Each regressor in bagging and boosting is trained on

242 different subsets of the training dataset, and these are randomly selected. In bagging, each

243 regressor is trained independently on samples that are randomly picked with replacement from

244 the original samples of the training dataset. Different models are constructed by applying the

245 same base learning algorithm and then averaging the prediction of all the regressors (Bühlmann

246 and Hothorn 2007; Shrestha and Solomatine 2006). Whereas in boosting, the first regressor is

247 trained on samples that are randomly picked from the original training dataset. Then all the

248 training patterns pass through the first regressor to adjust the prediction error (Drucker 1997;

249 Elith et al. 2008).

250 Each function prediction is associated with a weight (α_m) that is updated during the iterative

251 training algorithm. Eq. (5) shows the combination of all weak predictors to obtain a strong

252 predictor. In boosting, that is by the AdaBoost-based technique, the weights ($w_1, w_2, \ldots, w_N$) of

253 the individual observations (x_i, y_i), $i = 1, \ldots, N$ are initialized to $w_i = 1/N$.

$$254 \qquad G(x) = sign(\sum_{m=1}^{M} \alpha_m\, G_m(x)) \qquad (5)$$

255   These weights are reweighted after each iteration, so the probability patterns that are most in

256   error are more likely to be fixed in the training datasets. The patterns that are most in error are

257   those in which the difference between their predicted and actual values is significant (Bühlmann

258   and Hothorn 2007; Winkler et al. 2018).

259

260   *Global Sensitivity Analysis (GSA)*

261   GSA was carried out to test the robustness of the model. This was achieved by searching for

262   errors in the model and observing unexpected relationships between the inputs and outputs. In

263   GSA, all inputs are varied simultaneously in comparison with local sensitivity analysis. GSA

264   can be defined as the study of the uncertainty in the output of a mathematical model to identify

265   the key variables whose uncertainty affects most the output over their entire range of interest

266   (Homma and Saltelli 1996; Ramakrishnan and Bailey-Kellogg 2008; Tarantola et al. 2006).

267   A variance-based method is the most common one to compute sensitivity indices for GSA

268   (Homma and Saltelli 1996; Pianosi et al. 2015; Song et al. 2013; Tarantola et al. 2006).

269   Suppose the model:

$$270 \qquad \boldsymbol{Y = f(X_1, \dots, X_p),} \qquad (6)$$

271   where (Y) is the output, and the inputs $X_1, \dots, X_p$ are independent random variables defined by

272   probability distributions that define the uncertain knowledge in the system. The purpose of this

273   method is to study the contribution of each input variable to the output variance. In other words,

274   the importance of an input variable (X_i) on the variance of (Y) (Homma and Saltelli 1996;

275 Pianosi et al. 2015; Sobol 1993; Sobol 2001). The first order indices can be calculated using the

276 conditional variance, where $S_i < 1$, as shown in Eq. (7) (Sobol 1993; Zhan et al. 2013).

277
$$S_i = \frac{V_i}{V(Y)} = \frac{V(E(Y|X))}{V(Y)} \qquad (7)$$

278

279 **Research Methodology and Model Development**

280 Fig. 2 illustrates the framework of the proposed model that can predict average time to failure

281 in watermains in order to assess the condition of subsurface infrastructure and help decision

282 makers either in main rehabilitation or/and replacement services. The study has developed three

283 models including artificial neural network (ANN), ridge regression (l_2) and ensemble decision

284 tree (EDT) in order to select the one with satisfactory performance for failure prediction of

285 watermains. The developed models are evaluated employing root mean square error (RMSE),

286 mean absolute error (MAE), and coefficient correlation (R). In addition, the developed models

287 are tested on data collected from Quebec City watermains.

288 <Fig. 2>

289

290 **Data Collection Method**

291 The historical failure database used in this study is obtained from the municipality of Sainte-

292 Foy, Quebec, Canada. The database consists of recorded information, such as pipe material,

293 diameter, and length, year of installation, and breaks. It spans a 15-year period (1987–2001) for

294 each individual pipe within 432km of a water distribution system. The database contains five

295 types of pipe materials, namely, gray CI, ductile iron with lining (DIL), ductile iron without

296 lining (DIN), PVC, and Hy pipes. The majority of breaks occurred in CI, DIN, DIL, PVC, and

297 Hy at a rate of 47.05%, 19.41%, 18.34%, 11.18%, and 4.02%, respectively, as summarized in

298   Table 2. In a statistically based model, dummy variables can be used to represent qualitative

299   data that take the value of either 0 or 1 (Garavaglia and Sharma 1998). Thus, each material was

300   defined as a dummy input and fed into the model as a vector of two possible values, 0 or 1.

301   <Table 2>

302

303   Data Preprocessing and Cross-Validation

304   A descriptive analysis was utilized to check the missing values of the dataset. The database was

305   divided up so that 90% of the total database was dedicated to training and validation (Cross-

306   validation), and 10% to testing the final model. Cross-validation (k-fold) was applied to validate

307   and examine the performance of each selected model. It was also applied to evaluate the

308   parameter selection and learning process of all the developed models. The database (N) was

309   randomly split into five subsets (the folds), $N_1, N_2, N_3, \ldots, N_5$, of approximately equal sizes. The

310   model was tested and validated k times, hence $t \in \{1, 2, \ldots, k\}$. The model was trained on (N/N_t)

311   and validated on (N_t). The testing dataset, on the other hand, was then used to test the

312   performance of the final developed models. Some outliers were removed to improve the

313   performance of the model.

314

315   Model Development

316   Models were developed and coded to predict the time to watermain failure on the basis of pipe

317   material, diameter, and length by using MATLAB (R2018a) software. The time to watermain

318   failure is defined as the time from installation date to the 1st break.

319 *Artificial Neural Network Model*

320 ANN-MLP model was developed by using ANN Toolbox in MATLAB (R2018a) software. As

321 previously explained, MLP is a class of feedforward neural network, and used to refer to ANN

322 that consists of multiple layers of perceptrons. It was applied using backpropagation algorithm

323 with adaptive learning rate when searching for the best parameters for building a final model for

324 predicting the failure of watermains. The parameter-tuning approach was conducted for the

325 ANN-MLP model to reach the optimal solution. Fig. 3 depicts the representation of ANN-MLP

326 model used to search for the optimal solution. ANN-MLP consisted of one input and one output

327 layer with 6 neurons, and one neuron, respectively. The performance of ANN-MLP was also

328 examined at different numbers of hidden neurons in the range of [2, 20].

329 <Fig. 3>

330

331 *Ridge Regression Model*

332 Ridge regression has one parameter called lambda (λ) as explained previously in Eq. (4). The

333 model was coded to calculate the weights of each quantitative and qualitative variables in the

334 model. The quantitative variables are the pipe length (L) and diameter (D), whereas the

335 qualitative variables are the pipe materials that were defined as dummy variables in the model,

336 namely, PVC, DIL, DIN, CI, and Hy. The ridge regression model, on the other hand, was tested

337 on different values of (λ) in the range of [0, 2] to find the optimal solution when validating the

338 performance of the model. Eq (8) depicts the developed equations of ridge regression, where (y)

339 is the predicted time to pipe failure, (α_o) is the intercept/bias, and (β_k) is the estimated

340 weights/coefficients.

341 $$y = \alpha_o + \beta_1 L + \beta_2 D + \beta_3\,PVC + \beta_3\,DIN + \beta_4\,DIL + \beta_5\,CI \quad (8)$$

342

Ensemble Decision Tree (EDT) Model

344   Fig. 4 depicts the typical structure of bagging (a) and boosting (b) ensemble learning models

345   which were explained previously in ensemble decision tree-based model. Drucker (1997)

346   reported that if boosting is not equivalent to bagging, then in most cases, it is better than bagging.

347   Therefore, the boosting technique was applied in the present study to improve the prediction

348   accuracy. Ensemble decision tree model using the boosting technique is a combination of two

349   algorithms: (1) a regression tree that creates a relationship between predictors and response on

350   the basis of recursive binary splits, and (2) boosting, which is an ensemble technique used to

351   combine simple models to improve the accuracy of the model. Fig. 5 depicts a comprehensive

352   methodology of ensemble decision tree model using the boosting technique. Regression trees

353   were constructed and trained on randomly selected samples from the training dataset. The

354   models were then combined to predict the failure of watermains. A modification was applied to

355   the Adaboost.R algorithm as suggested by (Drucker 1997; Freund and Schapire 1996). First, a

356   single decision tree model was developed by tuning different hyperparameters such as maximum

357   split and parent size to eliminate the risk of overfitting. However, the model performance was

358   similar to the "default" setting. Afterward, the ensemble decision tree model was developed and

359   coded as follows (see Fig. 6):

360   •   Assign a weight for each sample in the training dataset $Z_i = \frac{1}{N_k}$, where (N_k) is the sample

361   number of training dataset), thus, the initial probability of each sample being in the

362   training set is $p_i = Z_i / \sum Z_i$.

363   •   Pick (N_k) with replacement to form a training dataset for each regressor (regression tree).

364 • Construct regression tree machines such that each machine makes a hypothesis (h_r: $\boldsymbol{x} \rightarrow$

365 $\boldsymbol{y}$, where (R) is the number of iterations/regressors).

366 • Train each regression tree machine (R) by using every (N_k).

367 • Calculate the loss using a nonlinear function ($Loss_i = \frac{|y_i - \hat{y}|^2}{\max(|y_i - \hat{y}|)^2}$) and the average loss

368 $\overline{(Loss} = \sum_{i=1}^{N_k} L_i\, p_i)$.

369 • Update the weight by $Z_i = Z_i \times \beta$, where ($\beta = \frac{\overline{Loss}}{1 - \overline{Loss}}$).

370 • Calculate the weighted median (W) for each regression tree; $W = \sum_{I:h_r \leq y} \log(\frac{1}{\beta_r})$.

371 • Calculate the cumulative predictions $CP(x) = sign(\sum_{r=1}^{R} W_r\, CP_r(x))$.

372 <Fig. 4>

373 <Fig. 5>

374 <Fig. 6>

375 *Global Sensitivity Analysis (GSA) of the Developed Models*

376 GSA was later carried out to test which variable was more critical to the model output. Random

377 samples of different sizes (N) were generated to obtain the desired probability density function

378 for (x_i) that covers the input space of interest. For the simplification of testing the model, a

379 uniform distribution was assumed for each variable where each outcome has the same

380 probability that it will be the outcome. The model was then evaluated at different levels of N

381 designs to test the reliability of the model. Finally, a sensitivity index (S_i) was computed for

382 each (x_i) at different (N) point designs by using Eq. (10). The first-order main effect was

383 calculated as it does not require a significant computational cost, and higher orders are often

384 neglected in sensitivity analysis (SA).

385

Performance Assessment of the Developed Models

387 A set of mathematical validation equations was used to evaluate the performance of the model.

388 A comparison between the actual (y_i) and the predicted values ($\hat{y}$) was accomplished as a main

389 evaluation procedure for assessing the performance of the model, using three techniques,

390 namely, RMSE, MAE, and correlation coefficient (R), Eqs. (9), (10), and (11), respectively

391 (Botchkarev 2018; El-Abbasy et al. 2014; Sarsam 2019). (R) is a statistical measure that

392 calculates the association between two variables. It is also well known as Pearson's correlation

393 coefficient. In this study, the strength of the relationship between two variables (A, B) was

394 calculated using (R) expressed in Eq. (11), (Sarsam 2019):

395
$$\text{RMSE} = \sum_{i=1}^{n} \sqrt{(\hat{y} - y_i)^2} \qquad (9)$$

396
$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{y} - y_i| \qquad (10)$$

397
$$R = \frac{cov\,(A, B)}{\sigma_A \sigma_B} \qquad (11)$$

398

MODEL IMPLEMENTATION AND VALIDATION

Model Setup for Artificial Neural Network (ANN-MLP)

401 The aim of the cross validation is to examine the performance of ANN-MLP, using

402 backpropagation algorithm with adaptive learning rate when searching for the best parameters

403 to build failure prediction model of watermains for decision makers. ANN-MLP trained until

404 the validation error reached the optimal at epochs 52. The output and input layers consisted of

405 one neuron with a linear activation function and 6 neurons, respectively. The weights are

406 randomly initialized in the range of [0, 1]. After tuning the parameters, the best ANN-MLP was

407 constructed with one hidden layer and 5 hidden neurons with tangent hyperbolic sigmoid as

408 shown in Fig. 7a. The learning rate was also set at 0.38. These parameters were chosen based

409 on the performance evaluation of ANN-MLP in cross-validation. Fig. 7b shows the performance

410 of ANN-MLP when the lowest MSE reached the minimum value (optimal solution) at epoch

411 37. The final model was tested on a test dataset and the result shows satisfactory performance

412 in terms of the prediction error of watermain failure prediction model, using ANN-MLP (Fig.

413 7c).

414 <center><Fig. 7></center>

415

416 ## Model Setup for Ridge Regression (l_2)

417 Cross validation showed that the validation error reached optimal solution at $\lambda = 0.2$ (Fig. 8a).

418 Therefore, after tuning the parameter and performance evaluation of ridge regression, (λ) was

419 set at 0.2. The performance of the final model of watermain failure prediction using ridge

420 regression is shown in Fig. 8b. Table 3 illustrates ridge coefficients of qualitative and

421 quantitative variables, and descriptive statistics of quantitative variables of the ridge regression

422 model. Eq (12) shows the equation of the developed model, using ridge regression.

423 $$y = 17.6394 + 1.1428\,L - 3.8146D - 12.7269\,PVC - 5.0187\,DIN + 4.9950\,DIL$$

424 $$+ 16.8929\,CI \quad (12)$$

425 <center><Table 3></center>

426

Model Setup for Ensemble Decision Tree (EDT)

The study utilized the parameter-tuning approach for ensemble decision tree model to reach the optimal solution. The model was trained on different numbers of iterations/regressors (R) and the initial weight was assigned for each sample in the training ($Z_i = 1/N_K$). EDT model, on the other hand, was trained and validated on different scenarios. Scenario (1): Some outliers were removed. Scenario (2): all outliers were removed. Scenario (3): Assumed minimum time to watermain failure = 5 years. Scenario (4): Assumed minimum time to watermain failure = 10 years. Scenarios (3) and (4) showed the highest performance. However, scenario (1) was selected to train the model as, logically speaking, pipe failure could happen at an early age because of manufacturing flaws, human errors, or any other odd/unknown reasons. Fig. 8c depicts the performance of the evaluation of the model at different numbers of iterations/regressors (TR). The optimal solution was achieved when TR = 60. Thus, after parameter tuning and model validation, (TR) was set at 60. Fig. 8d illustrates the final performance of the ensemble decision tree model for the failure prediction of watermains. Additionally, different decision tree-based models were developed and the results show that EDT proposed in this study performed better than other investigated methods (Table 4).

<Table 4>

<Fig. 8>

FINDINGS AND DISCUSSION

The Comparison of the Three Intelligent Models

The study employed a set of mathematical validation equations to test the performance of the each model . The evaluation matrices showed that l_2 has RMSE = 5.42, MAE = 4.21, and R =

450   0.90. However, ANN has RMSE = 6.47, MAE = 5.25, and R = 0.84. EDT, on the other hand,

451   has RMSE = 6.34, MAE = 5.01, and R = 0.88 (see Table 5). Results revealed that all models

452   were able to predict failure of watermains. Fig. 9 illustrates the predicted values versus actual

453   values of time-to-pipe failure and shows high correlation coefficient (R) was achieved by all the

454   models developed. Results also prove that there is no high variation between predicted values

455   and actual values, and there are no outliers (Fig. 10).

456   <Fig. 9>

457   <Table 5>

458

459   **Prediction Error of Different Pipe Categories**

460   The RMSE of different pipe categories was calculated on the basis of the pipe material. Results

461   indicated that the prediction error of watermain pipes made of PVC was lower than that of other

462   materials and the total average error. However, the prediction error resulting from CI material

463   was higher than the total average error (see Fig. 10a). Thus, the results suggest that the model

464   may perform better when predicting the failure of each material individually. For materials such

465   as PVC, Hy, DIN, and DIL, the prediction error is lower than the total average error.

466

467   **Global Sensitivity Analysis for EDT**

468   Ensemble decision tree model was again selected for analyzing and testing the reliability of the

469   model because of its simplicity and computational efficiency in terms of creating trees that assist

470   decision making. This is also the case because of the flexibility of decision tree based-model in

471   coping with both discrete and continuous variables. Overall, however, all models can help

472   decision makers to avoid the unexpected failure of watermains in the future.

473 The results of the GSA showed that CI, Hy, and DIL are the most critical variables to the output

474 of the model (average age to failure of individual pipe) followed by DIN and PVC (Fig. 10b).

475 Fig. 10 shows the sensitivity indices and standard error for each material at N = 50000, N =

476 100000, and N = 1000000. The standard error for each material decreases as the size of

477 generated samples increases.

478 <Fig. 10>

479

480

481 Different pipe geometry and material types are considered in this study. The relationship

482 between the pipe materials with different pipe lengths (regardless of its diameter) and the

483 average time-to-pipe failure is depicted in Fig. 11a. Whereas, the correlation of pipe materials

484 with different pipe materials (regardless of its length) versus the average time-to-pipe failure is

485 shown in Fig. 11b. Results showed that a relatively strong correlation exists between the pipe

486 length and the time-to-pipe failure (in years) for CI, DIN, and Hy. However, the time-to-pipe

487 failure (in years) is relatively constant with the length of PVC and DIL (Fig. 11a). On the other

488 hand, time-to-pipe failure (in years) has slightly changed with the diameter of DIL, DIN, and

489 Hy material at a certain point and then remained constant (Fig. 11b). This is attributed to the

490 fact that the majority of pipe failure occurred in pipes with small diameters, that is, between 100

491 and 250mm. Furthermore, the output is almost constant with the diameter for other materials. It

492 can be concluded, therefore, that the failure of watermains mostly relies on pipe material and

493 length. Similarly, previous studies showed that pipe length affects the condition of the pipe

494 (Demissie et al. 2017; Fares 2008; Ismaeel 2016; Karimian 2015; Mohammed 2016; Sattar et

495 al. 2019; Wang 2006). Additionally, as pointed out by Zangenehmadar (2016), the potentail of

496 water main failure increases with the increase of the pipe length. The results proved that the

497   failure of water mains is correlated with the pipe length. El Chanati (2014), on the other hand,

498   reported that the pipe length contributes to the failure of water mains.

499   \<Fig. 11\>

500

CONCLUSION

501

502   The purpose of this study was to develop intelligent models that could assist decision makers to

503   avoid unexpected watermain failure. The developed models were trained on data collected from

504   the municipality of Sainte-Foy of Quebec City. The data consisted of three physical variables

505   that were expected to contribute to the failure of watermains with respect to pipe material,

506   length, and diameter. The results revealed that the boosting technique reduced the prediction

507   error of single decision tree.

508   Overall, all models showed good performance and were satisfactorily able to predict the failure

509   of watermains. However, ensemble decision tree is recommended owing to the simplicity and

510   computational efficiency of decision tree-based models in terms of creating trees for decision

511   making and coping with both discrete and continuous variables. GSA showed that CI, Hy, and

512   DIL are the most critical variables to the output of the model. Besides, the results revealed that

513   a relatively strong correlation exists between time-to-pipe failure (in years) and the length of CI,

514   DIN, and Hy. The results also indicated that the output slightly changes with the diameter for

515   Hy, DIN and DIL, whereas the model output is almost constant with the diameter for other

516   materials. Nevertheless, some pipe materials, such as CI, DIL and Hy, last longer than others. It

517   can, therefore, be concluded that the failure of watermains mostly relies on pipe material and

518   length. The provided data, on the other hand, are limited and there are other variables that are

519   crucial for identifying the cause of watermain failure. Considering especially, the fact that the

520 failure of watermains is a dynamic process and time-dependent in nature,

521 dynamic/environmental variables will be considered in future. Finally, it is worth mentioning

522 that the performance of prediction models highly relies on the quality of the data collected and

523 on the selection of appropriate algorithmic techniques. However, the database is limited and

524 adding more variables that affect the failure of watermains may improve the accuracy of the

525 model.

526

DATA AVAILABILITY STATEMENT

528 Some data, models, codes are used during the study are available from the corresponding author

529 by request. Data provided by the municipality of Sainte-Foy, Quebec cannot be shared as it is

530 confidential.

531

536

REFERENCES

538 Ana, E. V., and Bauwens, W. (2010). "Modeling the structural deterioration of urban drainage
539   pipes: the state-of-the-art in statistical methods." *Urban Water Journal*, 7(1), 47-59.
540 ASCE (2017). "2017 infrastructure report card." https://www.infrastructurereportcard.org/
541 (July 10 2019).
542 Basheer, I. A., and Hajmeer, M. (2000). "Artificial neural networks: fundamentals, computing,
543   design, and application." *Journal of microbiological methods*, 43(1), 3-31.
544 Bishop, C. M. (2006). *Pattern recognition and machine learning*, springer, New York, United
545   States.

546   Botchkarev, A. (2018). "Performance Metrics (Error Measures) in Machine Learning
547   Regression, Forecasting and Prognostics: Properties and Typology."
548   https://arxiv.org/abs/1809.03006, (June 16, 2019).
549   Bühlmann, P., and Hothorn, T. (2007). "Boosting algorithms: Regularization, prediction and
550   model fitting." *Statistical Science*, 22(4), 477-505.
551   Christodoulou, S. E. (2011). "Water network assessment and reliability analysis by use of
552   survival analysis." *Water resources management*, 25(4), 1229-1238.
553   CIRC (2016). "Canadian Infrastructure Report Card."
554   http://canadianinfrastructure.ca/en/index.html, (July 22,2019).
555   Demissie, G., Tesfamariam, S., and Sadiq, R. (2017). "Prediction of pipe failure by
556   considering time-dependent factors: Dynamic Bayesian belief network model." *ASCE-*
557   *ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil*
558   *Engineering*, 3(4), 04017017.
559   Drucker, H. (1997). "Improving regressors using boosting techniques." *Proc., In Machine*
560   *Learning: Proceeding of the Fourteenth International Conference*, 107-115.
561   El-Abbasy, M. S., Senouci, A., Zayed, T., Mirahadi, F., and Parvizsedghy, L. (2014).
562   "Artificial neural network models for predicting condition of offshore oil and gas
563   pipelines." *Automation in Construction*, 45, 50-65.
564   El Chanati, H. (2014). "Performance assessment of water network infrastructure."Doctoral
565   dissertation, Concordia University.
566   Elith, J., Leathwick, J. R., and Hastie, T. (2008). "A working guide to boosted regression
567   trees." *Journal of Animal Ecology*, 77(4), 802-813.
568   Fares, H. A. (2008). "Evaluating the risk of water main failure using a hierarchical fuzzy
569   expert system." Concordia University.
570   Farmani, R., Kakoudakis, K., Behzadian, K., and Butler, D. (2017). "Pipe Failure Prediction in
571   Water Distribution Systems Considering Static and Dynamic Factors." *Procedia*
572   *Engineering*, 186, 117-126.
573   Folkman, S. (2018). "Water main break rates in the USA and Canada: A comprehensive
574   study." Buried Structure Laboratory, Utah State University, Logan, UT.
575   Francis, R. A., Guikema, S. D., and Henneman, L. (2014). "Bayesian belief networks for
576   predicting drinking water distribution system pipe breaks." *Reliability Engineering &*
577   *System Safety*, 130, 1-11.
578   Freund, Y., and Schapire, R. E. (1996). "Experiments with a new boosting algorithm." *Proc.,*
579   *Machine Learning: Proceedings of the Thirteenth International Conference*
580   148-156.
581   Garavaglia, S., and Sharma, A. (1998). "A smart guide to dummy variables: Four applications
582   and a macro." *Proc., Proceedings of the Northeast SAS Users Group Conference*, 43.
583   Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation. 2nd edition.*
584   , Prentice Hall PTR, Ontario, Canada.
585   Homma, T., and Saltelli, A. (1996). "Importance measures in global sensitivity analysis of
586   nonlinear models." *Reliability Engineering & System Safety*, 52(1), 1-17.
587   Ismaeel, M. (2016). "Performance based budget allocation model for water networks."
588   Concordia University.
589   Jafar, R., Shahrour, I., and Juran, I. (2010). "Application of Artificial Neural Networks (ANN)
590   to model the failure of urban water mains." *Mathematical and Computer Modelling*,
591   51(9-10), 1170-1180.

Kabir, G., Tesfamariam, S., and Sadiq, R. (2015). "Bayesian model averaging for the prediction of water main failure for small to large Canadian municipalities." *Canadian Journal of Civil Engineering*, 43(3), 233-240.

Kaddoura, K., and Zayed, T. (2018). "Erosion Void Condition Prediction Models for Buried Linear Assets." *Journal of Pipeline Systems Engineering and Practice*, 10(1). 10.1061/(ASCE)ps.1949-1204.0000357.

Kamel, S., and Meguid, M. A. (2012). "Investigating the effects of local contact loss on the earth pressure distribution on rigid pipes." *Geotechnical and Geological Engineering*, 31(1), 199-212.

Karimian, S. F. (2015). "Failure Rate Prediction Models of Water Distribution Networks." Concordia University.

Kaushik, G., Manimaran, A., Vasan, A., Sarangan, V., and Sivasubramaniam, A. (2017). "Cracks under pressure? burst prediction in water networks using dynamic metrics." *Proc., Twenty-Ninth IAAI Conference*Texas, United States.

Kettler, A., and Goulter, I. (1985). "An analysis of pipe breakage in urban water distribution networks." *Canadian journal of civil engineering*, 12(2), 286-293.

Khalaf, G., and Shukur, G. (2005). "Choosing Ridge Parameter for Regression Problems." *Communications in Statistics-Theory and Methods*, 34(5), 1177-1182.

Kishawy, H. A., and Gabbar, H. A. (2010). "Review of pipeline integrity management practices." *International Journal of Pressure Vessels and Piping*, 87(7), 373-380.

Kleiner, Y., and Rajani, B. (2001). "Comprehensive review of structural deterioration of water mains: statistical models." *Urban water*, 3(3), 131-150.

Li, Z., Zhang, B., Wang, Y., Chen, F., Taib, R., Whiffin, V., and Wang, Y. (2013). "Water pipe condition assessment: a hierarchical beta process approach for sparse incident data." *Machine Learning*, 95(1), 11-26.

Liu, Z., and Kleiner, Y. (2014). "Computational intelligence for urban infrastructure condition assessment: Water transmission and distribution systems." *IEEE Sensors Journal*, 14(12), 4122-4133.

Makar, J., Desnoyers, R., and McDonald, S. (2001). "Failure modes and mechanisms in gray cast iron pipe." *Underground Infrastructure Research*, 1-10.

Mavin, K. (1996). *Predicting the failure performance of individual water mains*, Urban Water Research Association of Australia, Melbourne, Australia.

Mirza, S. (2007). "Danger ahead: The coming collapse of Canada's municipal infrastructure." Federation of Canadian Municipalities.

Mohammed, A. U. (2016). "Integrated Reliability Assessment of Water Distribution Networks." Concordia University.

Ormsby, C. M. (2009). "A framework for estimating the total cost of buried municipal infrastructure renewal projects." Doctoral dissertation, McGill University, Montreal, Canada.

Parvizsedghy, L., Gkountis, I., Senouci, A., Zayed, T., Alsharqawi, M., El Chanati, H., El-Abbasy, M., and Mosleh, F. (2017). "Deterioration Assessment Models for Water Pipelines." *World Academy of Science, Engineering and Technology, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering*, 11(7), 1013-1022.

Pianosi, F., Sarrazin, F., and Wagener, T. (2015). "A Matlab toolbox for global sensitivity analysis." *Environmental Modelling & Software*, 70, 80-85.

638 Rajani, B., and Kleiner, Y. (2001). "Comprehensive review of structural deterioration of water
639   mains: physically based models." *Urban water*, 3(3), 151-164.
640 Ramakrishnan, N., and Bailey-Kellogg, C. (2008). "Gaussian Process Models in Spatial Data
641   Mining." *Encyclopedia of GIS*, 325-329.
642 Rogers, P. D., and Grigg, N. S. (2009). "Failure assessment modeling to prioritize water pipe
643   renewal: two case studies." *Journal of Infrastructure Systems*, 15(3), 162-171.
644   10.1061/(ASCE)1076-0342(2009)15:3(162).
645 Sarsam, S. M. "Reinforcing the Decision-making Process in Chemometrics: Feature Selection
646   and Algorithm Optimization." *Proc., Proceedings of the 2019 8th International
647   Conference on Software and Computer Applications*, ACM, 11-16.
648 Sattar, A. M., Ertuğrul, Ö. F., Gharabaghi, B., McBean, E., and Cao, J. (2019). "Extreme
649   learning machine model for water network management." *Neural Computing and
650   Applications*, 31(1), 157-169.
651 Shirzad, A., Tabesh, M., and Farmani, R. (2014). "A comparison between performance of
652   support vector regression and artificial neural network in prediction of pipe burst rate
653   in water distribution networks." *KSCE Journal of Civil Engineering*, 18(4), 941-948.
654 Shrestha, D. L., and Solomatine, D. P. (2006). "Experiments with AdaBoost. RT, an improved
655   boosting scheme for regression." *Neural computation*, 18(7), 1678-1710.
656 Sipos, C. (2010). "GIS-inventory and Condition Rating of Water Supply System at McGill
657   Downtown Campus." Doctoral dissertation, McGill University Montreal, Canada.
658 Snider, B., and McBean, E. A. (2018). "Improving Time to Failure Predictions for Water
659   Distribution Systems Using Extreme Gradient Boosting Algorithm." *Proc.,
660   WDSA/CCWI Joint Conference Proceedings*.
661 Sobol, I. M. (1993). "Sensitivity estimates for nonlinear mathematical models." *Mathematical
662   modelling and computational experiments*, 1(4), 407-414.
663 Sobol, I. M. (2001). "Global sensitivity indices for nonlinear mathematical models and their
664   Monte Carlo estimates." *Mathematics and computers in simulation*, 55(1-3), 271-280.
665 Song, X.-M., Kong, F.-Z., Zhan, C.-S., Han, J.-W., and Zhang, X.-H. (2013). "Parameter
666   identification and global sensitivity analysis of Xin'anjiang model using meta-modeling
667   approach." *Water Science and Engineering*, 6(1), 1-17.
668 Syachrani, S., Jeong, H. S. D., and Chung, C. S. (2013). "Decision Tree–Based Deterioration
669   Model for Buried Wastewater Pipelines." *Journal of Performance of Constructed
670   Facilities*, 27(5), 633-645. 10.1061/(ASCE)cf.1943-5509.0000349.
671 Tarantola, S., Gatelli, D., and Mara, T. A. (2006). "Random balance designs for the estimation
672   of first order global sensitivity indices." *Reliability Engineering & System Safety*,
673   91(6), 717-727.
674 Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso." *Journal of the Royal
675   Statistical Society: Series B (Methodological)*, 58(1), 267-288.
676 Wang, Y. (2006). "Deterioration and condition rating analysis of water mains." Concordia
677   University.
678 Wang, Y., Zayed, T., and Moselhi, O. (2009). "Prediction models for annual break rates of
679   water mains." *Journal of performance of constructed facilities*, 23(1), 47-54.
680   10.1061/(ASCE)0887-3828(2009)23:1(47).
681 Wilson, D., Filion, Y., and Moore, I. (2017). "State-of-the-art review of water pipe failure
682   prediction models and applicability to large-diameter mains." *Urban Water Journal*,
683   14(2), 173-184.

684 Winkler, D., Haltmeier, M., Kleidorfer, M., Rauch, W., and Tscheikner-Gratl, F. (2018). "Pipe
685   failure modelling for water distribution networks using boosted decision trees."
686   *Structure and Infrastructure Engineering*, 14(10), 1402-1411.
687 Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical machine*
688   *learning tools and techniques*, Elsevier, United States.
689 Yamijala, S., Guikema, S. D., and Brumbelow, K. (2009). "Statistical models for the analysis
690   of water distribution system pipe break data." *Reliability Engineering & System Safety*,
691   94(2), 282-293.
692 Zangenehmadar, Z. (2016). "Asset Management Tools for Sustainable Water Distribution
693   Networks." Concordia University.
694 Zhan, C.-S., Song, X.-M., Xia, J., and Tong, C. (2013). "An efficient integrated approach for
695   global sensitivity analysis of hydrological model parameters." *Environmental*
696   *Modelling & Software*, 41, 39-52.
697 Zhou, J., and Chen, F. (2018). *Human and Machine Learning: Visible, Explainable,*
698   *Trustworthy and Transparent*, Springer, Cham, Switzerland.
699 Zou, H., and Hastie, T. (2005). "Regularization and variable selection via the elastic net."
700   *Journal of the royal statistical society: series B (statistical methodology)*, 67(2), 301-
701   320.

702

703

704

705

706

707

708

709

710

711

712

713

714 **List of Tables**

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741 **Table 1**. A Summary of Failure Prediction Models of Watermains.

Reference	Variables	Methodology	Output
Yamijala (2009)	Diameter, material, length, land use, soil type, soil moisture, temperature	MLRM MERM GLM	Likelihood of break
Christodoulou (2011)	Age, material, incident type, diameter, number of breaks	PRM	Failure rate
Francis et al. (2014)	material, diameter, age, demographic variables and temperature	BBNs	Pipe breaks
Shirzad et al. (2014)	Hydraulic pressure, diameter, length, age and depth	ANN, SVR	Failure rate
Kabir et al. (2015)	# of bursts, age, diameter, length, soil resistivity, soil corrosivity	BMA RM	Failure rate
Demissie et al. (2017)	Length, Diameter, number of previous failure, type of service connection, freezing index, thawing index, rainfall deficit and soil corrosion	DBN	Pipe breaks
Farmani et al.(2017)	Length, diameter, age, temperature, freezing index	EPR	Pipe breaks
Parvizsedghy et al. (2017)	Age, pipe size, soil type, groundwater depth, installation quality, and C-factor, and surface type	RM	Failure rate
Kaushik et al. (2017)	Length, diameter, age, # of bursts, depth, material, pressure, and flow	LRM	Pipe breaks
Sattar et al. (2019)	Length, diameter, soil type, and # of pipe failures	ELM	Failure time
Sinder and McBean (2018)	Length, diameter, soil type, material, and # of pipe failures	GBDTM	Failure time
Winkler et al. (2018)	Length, material, age, diameter, variables, and # of pipe failures	BDT	Failure or not

742 MLRM= Multilinear Regression Model, MERM= Multivariate Exponential Regression model, GLM= Generalized
743 linear model, SVM= Support Vector Machine, PRM= Possession Regression Model, BDT= Boosted Decision Tree,
744 RM= Regression Model, ANN= Artificial Neural Network, BMA= Bayesian Model Average, EPR= Evolutionary
745 Polynomial Regression, LRM=Logistic Regression Model, EML= Extreme Machine Learning, GBDTM=Gradient
746 Boosting Decision Tree Model, DTM= Decision Tree Model, SVR= Support Vector Regression, BBNs= Bayesian
747 Belief Networks, DBN= Dynamic Bayesian Network.
748

749
750
751
752
753
754
755
756
757
758
759
760
761

762

763 **Table 2**. A summary of the failure database records from municipality of Sainte-Foy, Quebec City, Canada.

	Pipe Material				
Variable	CI	DIL	DIN	PVC	Hy
Total Number of pipes	526	205	217	125	45
Total pipe length (m)	3734.04	15673.96	14657.9	9196.245	3371.854
Average Pipe length (m)	71.0	71.50	72.23	73.57	74.93
Diameter (mm)	100-600	38-600	100-450	50-750	150-400

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783
784 **Table 3.** Ridge coefficients and model equation of ridge regression model.

Variable	Ridge Coefficient	Mean	Std	Ste	Max
Intercept	17.6394	×	×	×	×
L	1.1428	0.391	0.214	0.0068	1
D	-3.8146	0.231	0.081	0.0026	1
PVC*	-12.7269	×	×	×	×
DIL*	-5.0187	×	×	×	×
DIN*	4.9950	×	×	×	×
CI*	16.8929	×	×	×	×

785 *Categorical variables; L: Length, D: Diameter; Std: Standard deviation; Ste: Standard Error.

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804 **Table 4**. Performance of different decision tree-based models based on validation error.

Models	Method	Cross-validation	# of iterations	RMSE	MAE
Single Decision Tree	"Fitrtree" *	5-fold	Non	7.58	5.98
Ensemble Model	"Fitrensemble"* Boosting Technique	5-fold	50	7.31	5.69
Ensemble Decision Tree Model	"Fitrtree"* Boosting Technique	5-fold	60	3.10	2.7138

805 *Applied function in MATLAB (2018a).

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

Table 5. An evaluation of different models using a set of mathematical validation equations.

Models	RMSE	MAE	R-Correlation
l_2	5.42	4.21	0.90
ANN	6.47	5.25	0.84
EDT	6.34	5.01	0.88

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

Fig. 1 . Factors affecting the condition of watermains in both short and long terms.

Fig. 2. Proposed framework for failure prediction of watermains.

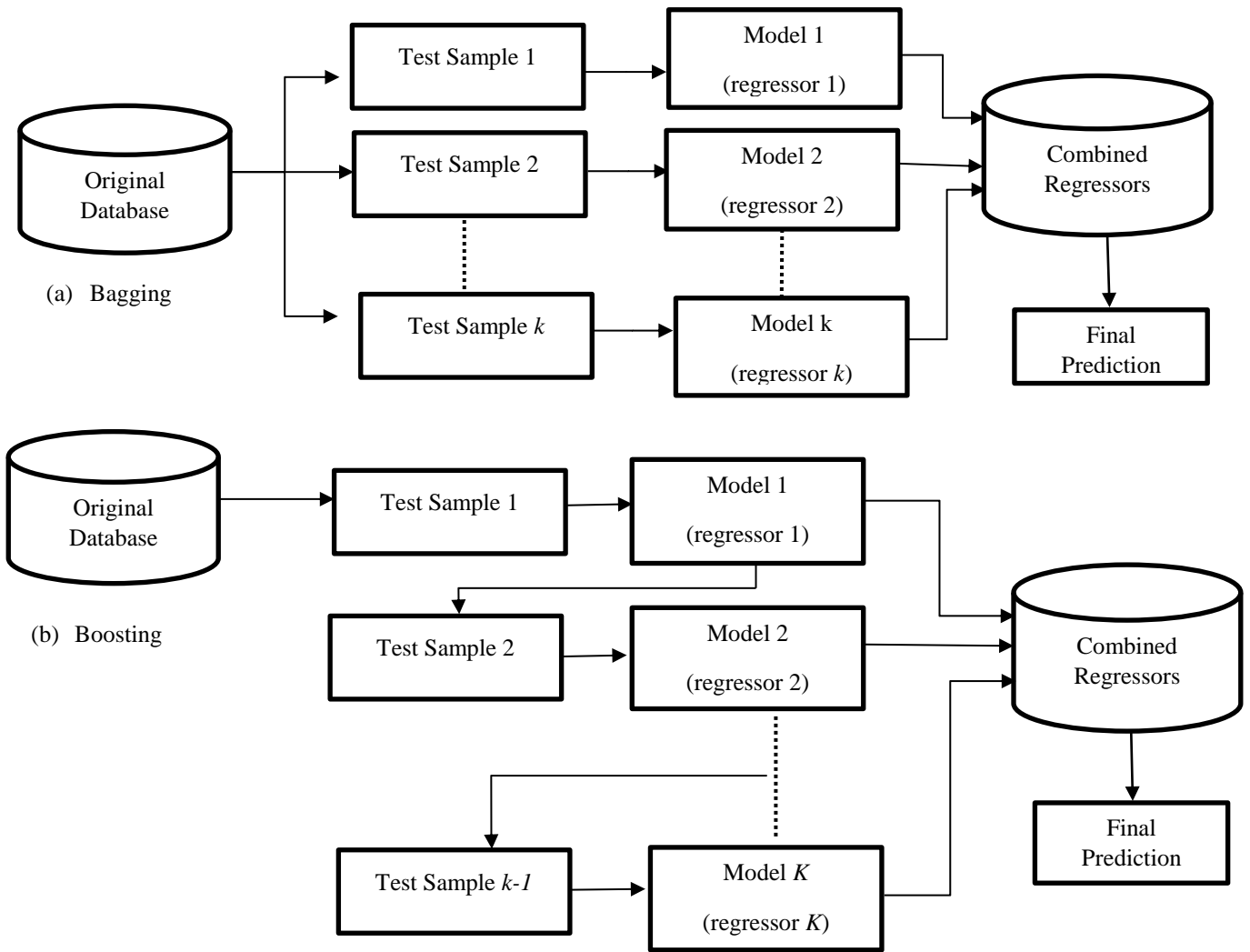**Fig. 3**. The initial representation of ANN-MLP.

Fig. 4. The typical architecture of ensemble learning models; (a) bagging based-model; (b) boosting based-model.
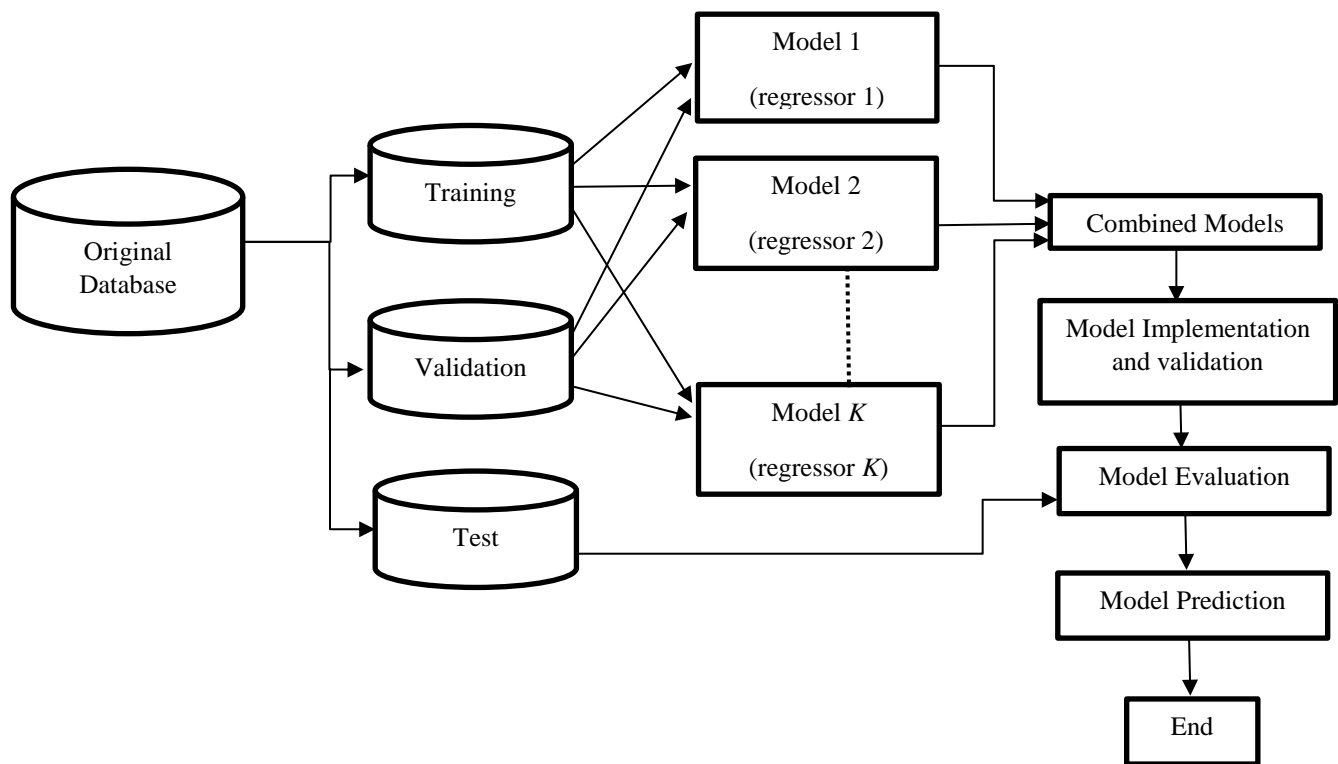
Fig.5. A comprehensive methodology of ensemble decision tree model using boosting technique.
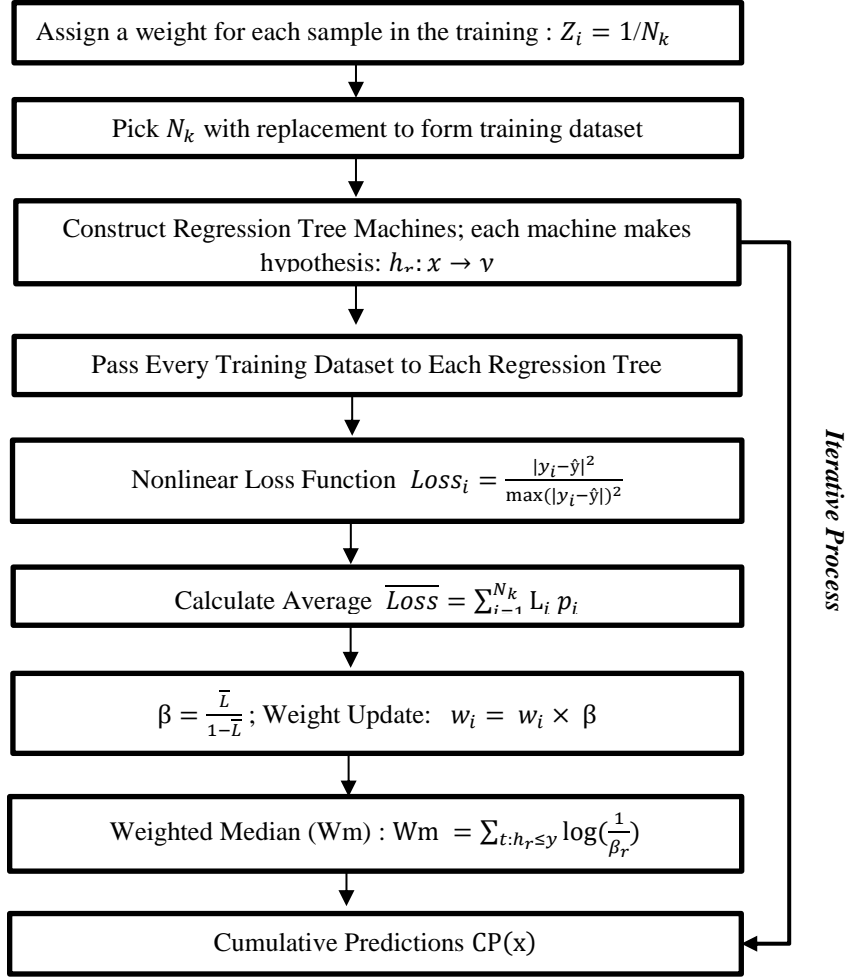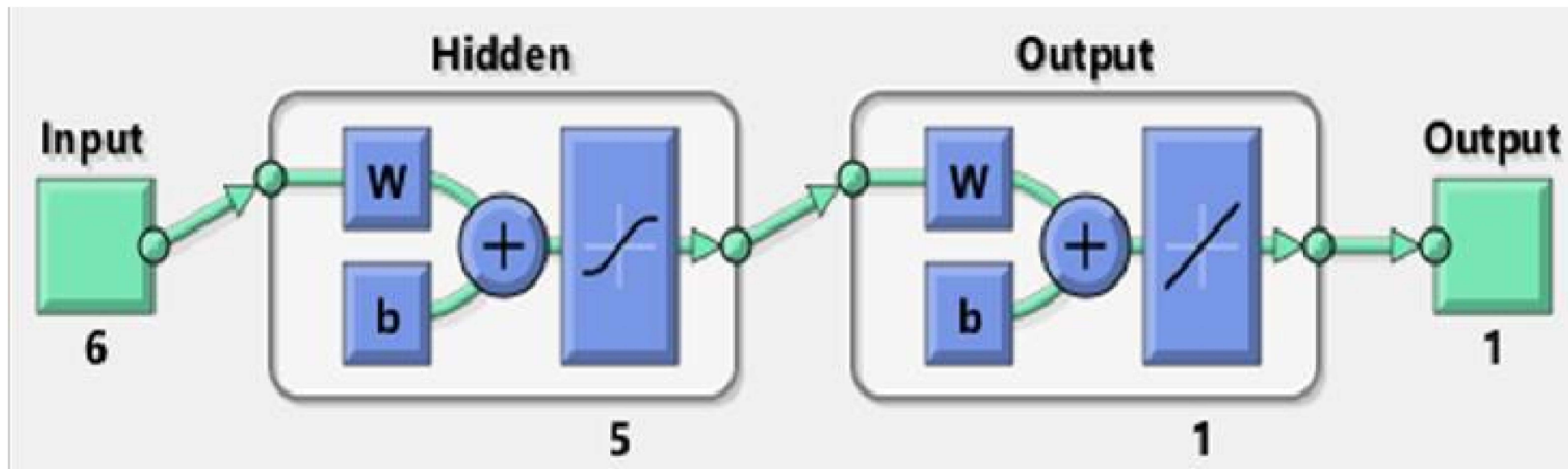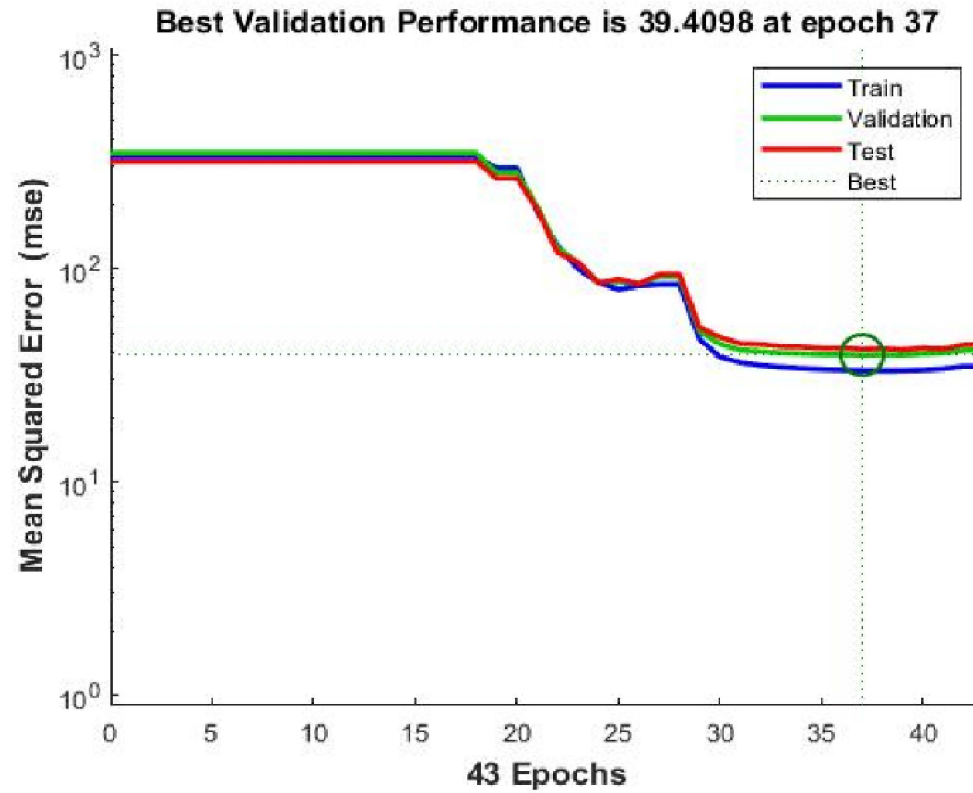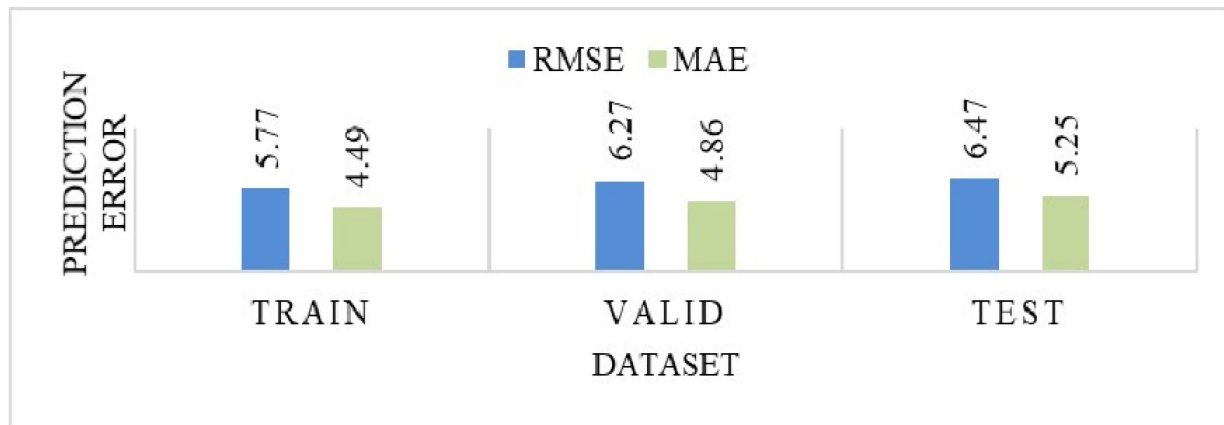
```
┌─────────────────────────────────────────────────────────────┐
│ Assign a weight for each sample in the training : $Z_i = 1/N_k$ │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│       Pick $N_k$ with replacement to form training dataset    │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│   Construct Regression Tree Machines; each machine makes      │
│              hypothesis: $h_r: x \rightarrow v$               │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│       Pass Every Training Dataset to Each Regression Tree     │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│   Nonlinear Loss Function  $Loss_i = \dfrac{|y_i - \hat{y}|^2}{\max(|y_i - \hat{y}|)^2}$ │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│   Calculate Average $\overline{Loss} = \sum_{i-1}^{N_k} L_i\, p_i$ │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│   $\beta = \dfrac{\overline{L}}{1-\overline{L}}$ ; Weight Update:  $w_i = w_i \times \beta$ │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│   Weighted Median (Wm) : Wm $= \sum_{t:h_r \le y} \log(\frac{1}{\beta_r})$ │
└─────────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────────┐
│              Cumulative Predictions $CP(x)$                    │
└─────────────────────────────────────────────────────────────┘
```

Iterative Process

Fig. 6. Model development of ensemble decision tress model.
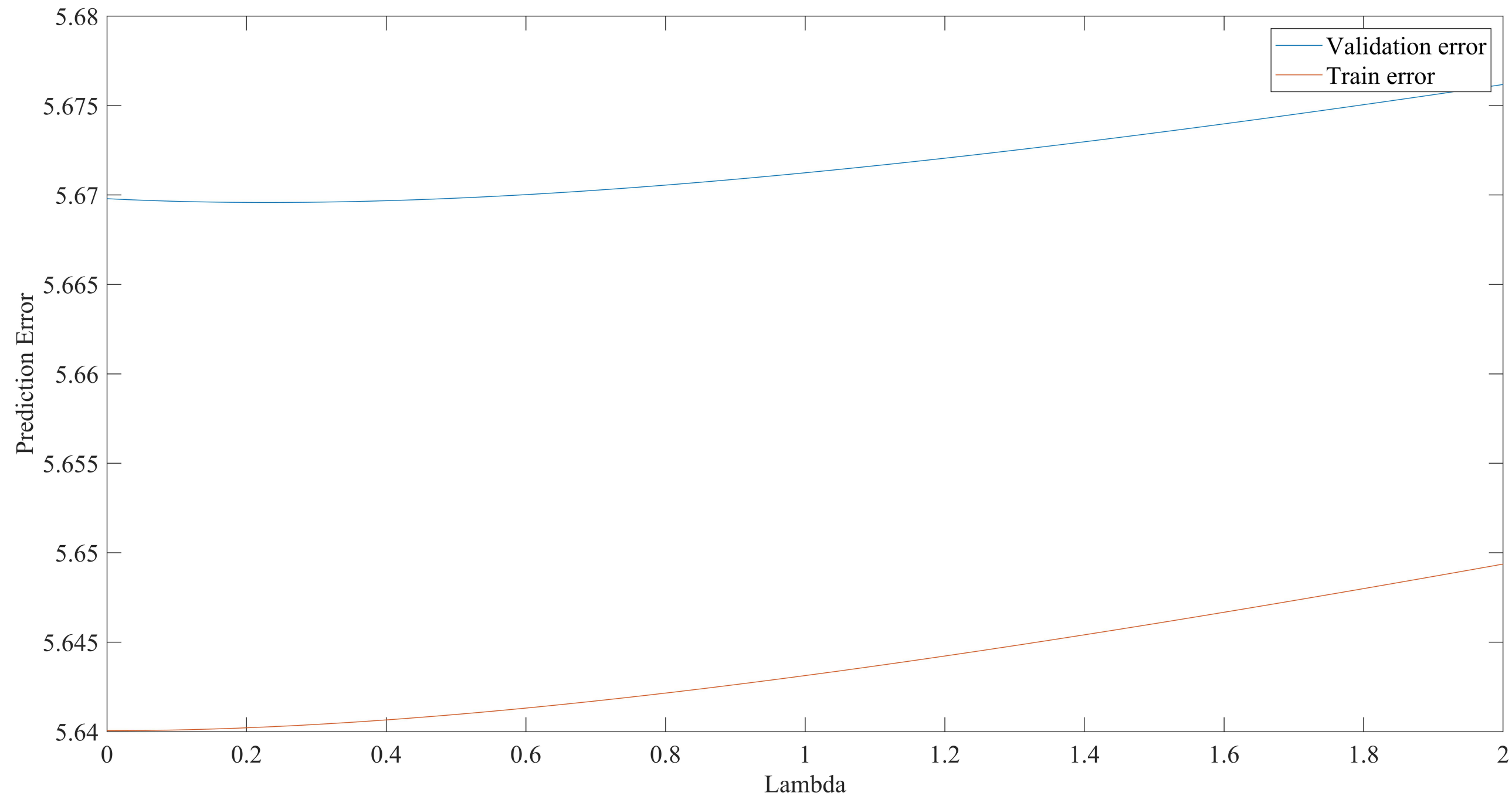
(a) ANN-MLP Model.

(b) ANN-MLP Model Training Performance.
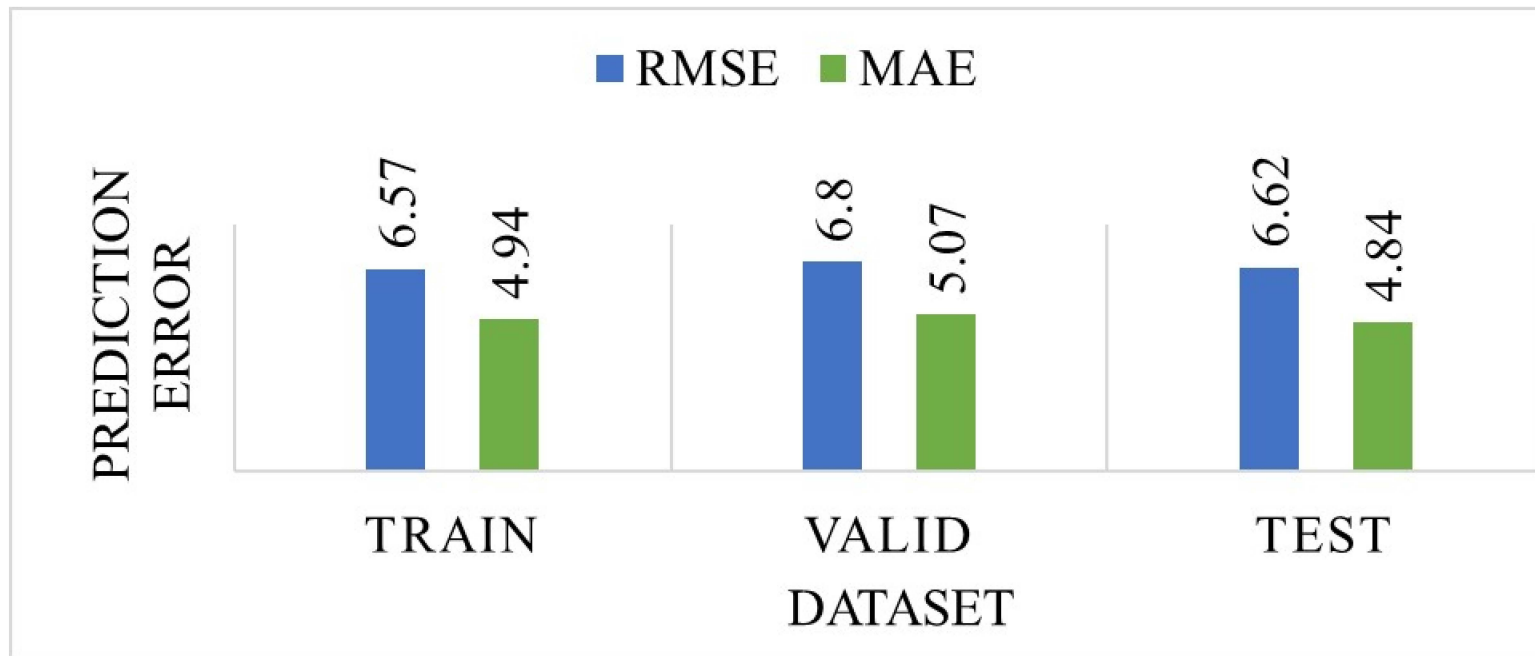
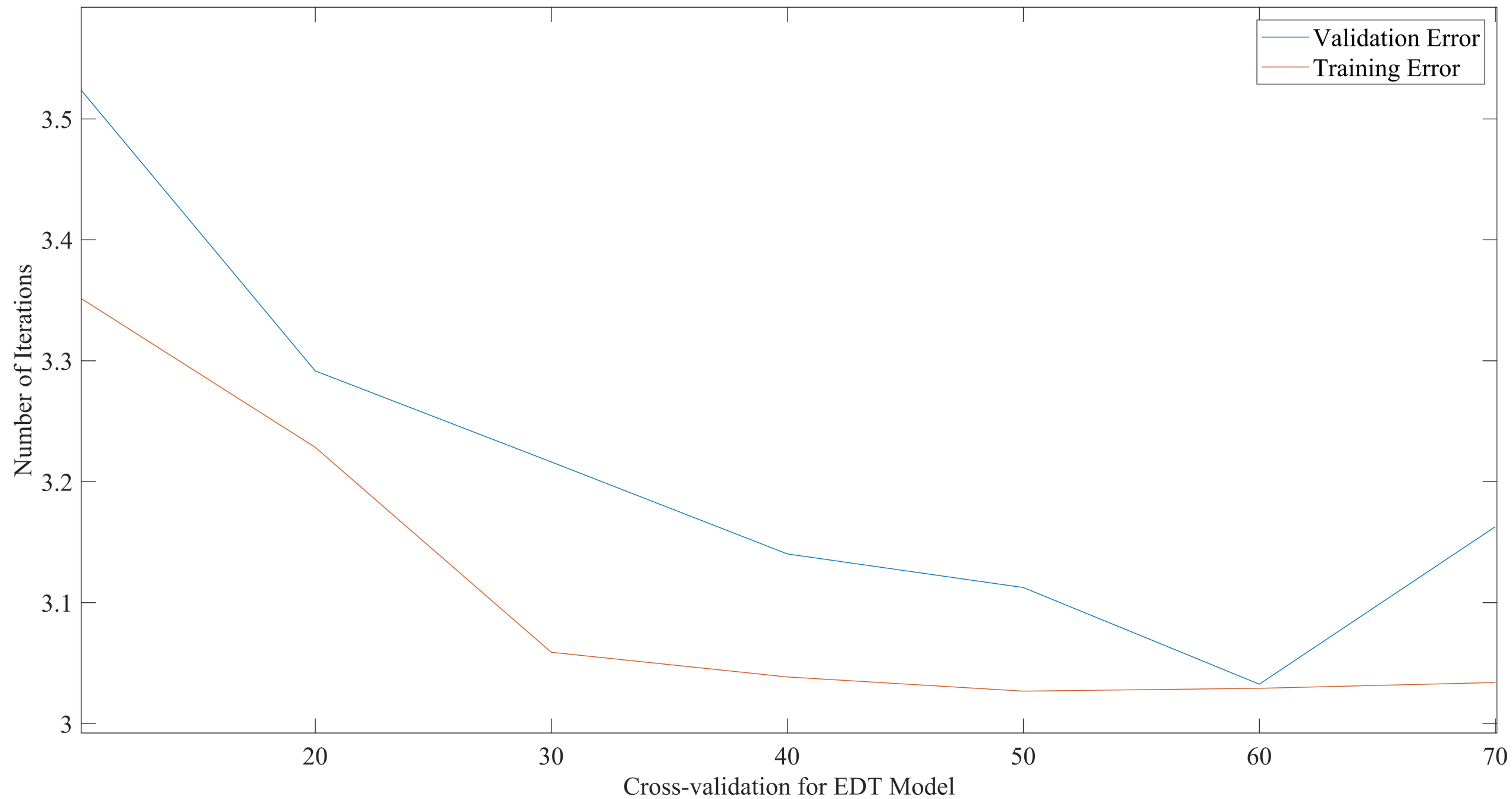**(c)** ANN-MLP Model Prediction Performance.

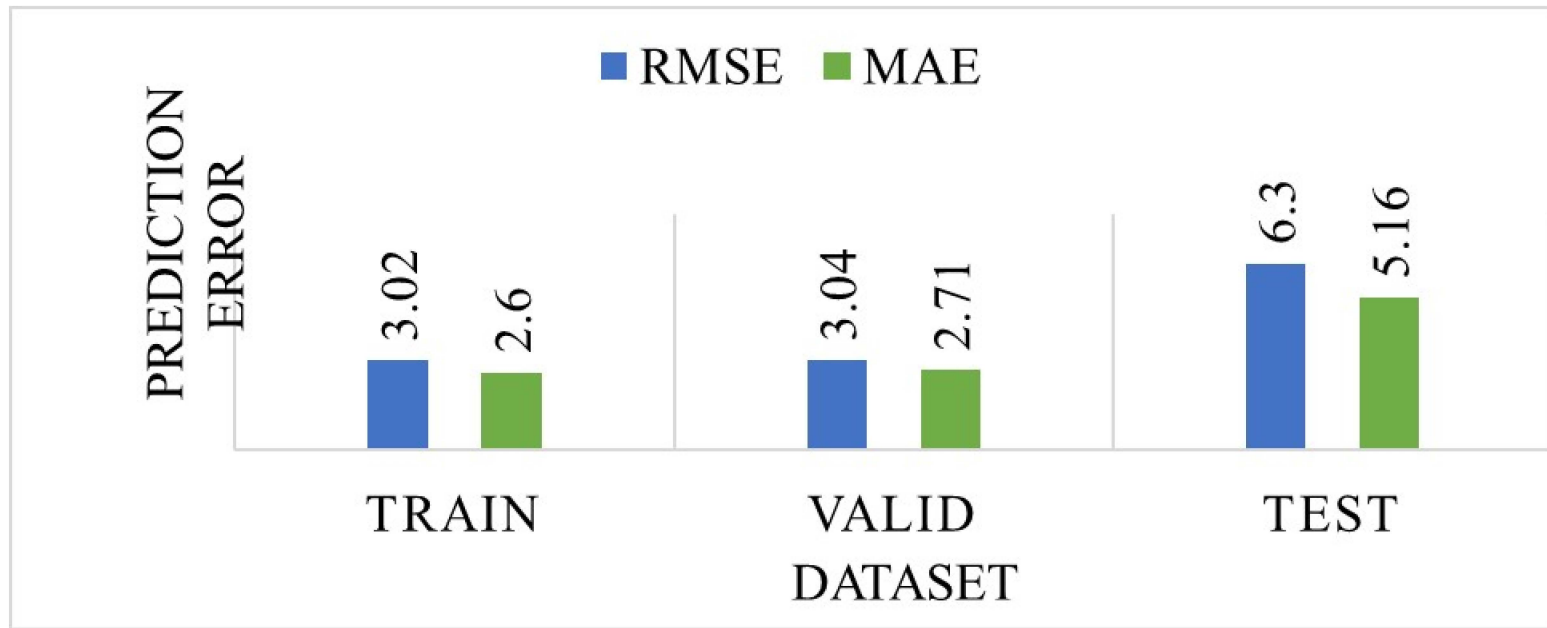**Fig. 7.** ANN-MLP and its Performance.

(a) Cross-validation for Ridge Regression Model.

(b) Prediction Performance for Ridge Regression Model.

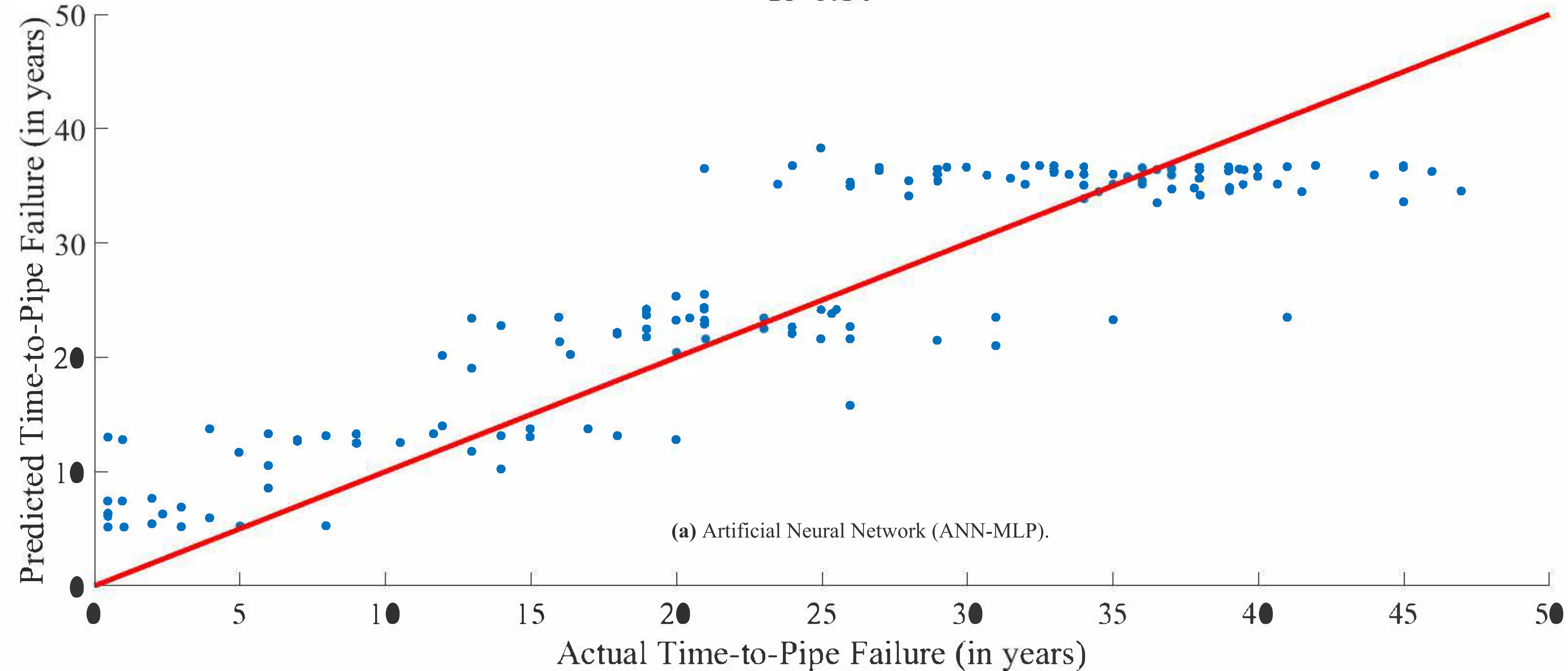**(c)** Cross-validation for Ensemble Decision Tree Model.

(d) Prediction Performance for Ensemble Decision Tree Model.

Fig. 8. Ridge Regression Model Outputs [(a) Cross-validation (b) Prediction Performance] and Ensemble Decision Tree Model Outputs [(c) Cross-validation (d) Prediction Performance].

Artificial Neural Network (ANN-MLP)
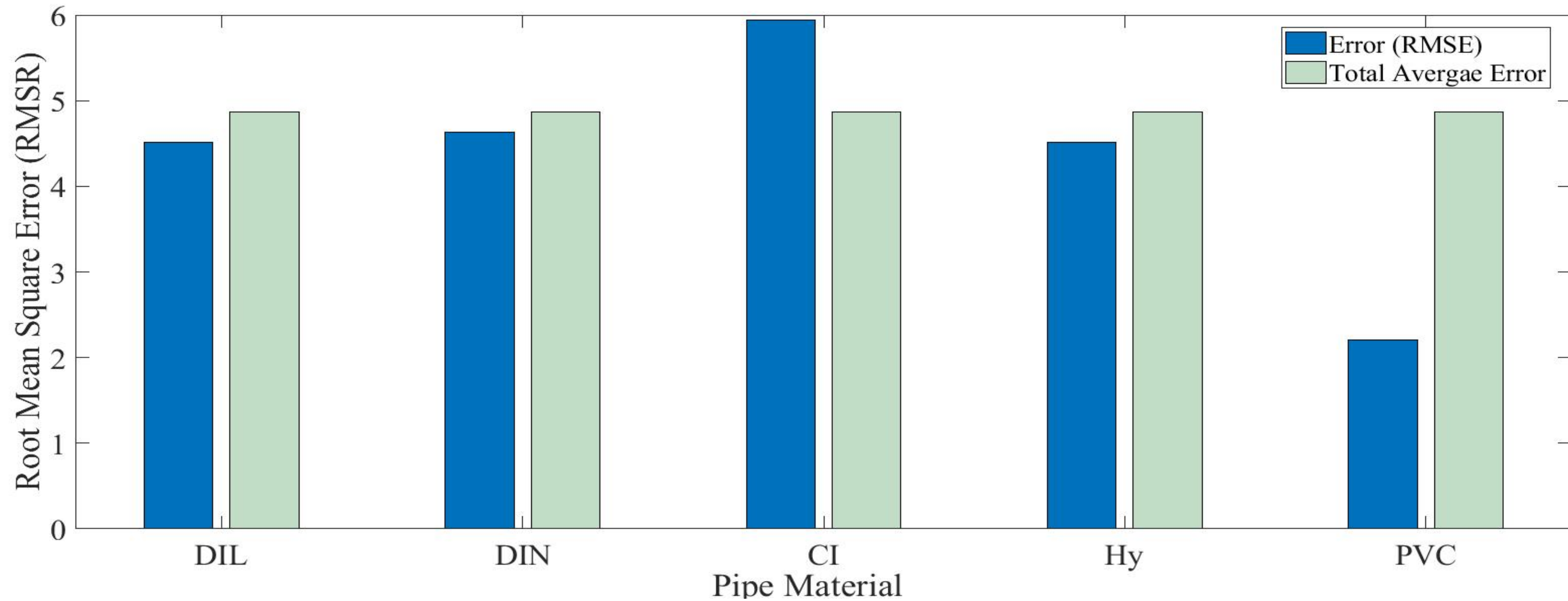**R=0.84**

(a) Artificial Neural Network (ANN-MLP).

Predicted Time-to-Pipe Failure (in years)

Actual Time-to-Pipe Failure (in years)

(b) Ridge Regression (*l2*).

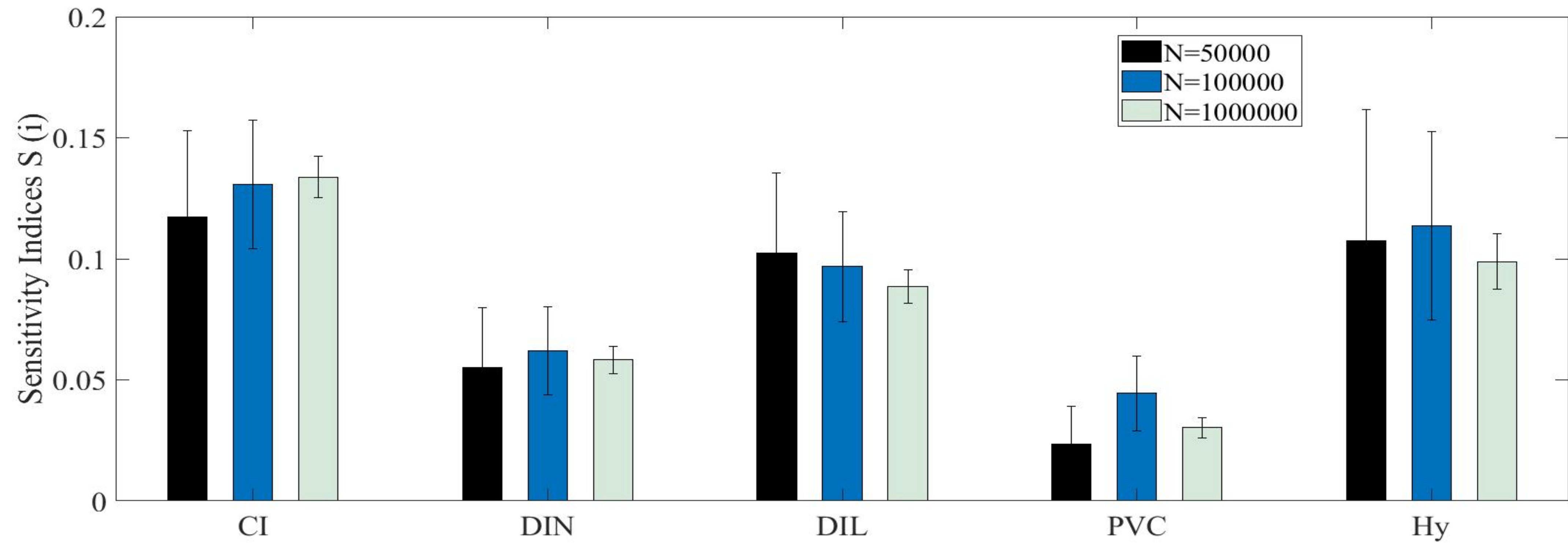**Ensemble Decision Tree (EDT)**
R=0.88

(c) Ensemble Decision Tree (EDT).

Fig. 9. Actual vs. predicted values of watermain failure prediction.

(a) The error (RMSE).

(b) Sensitivity indices of pipe materials.
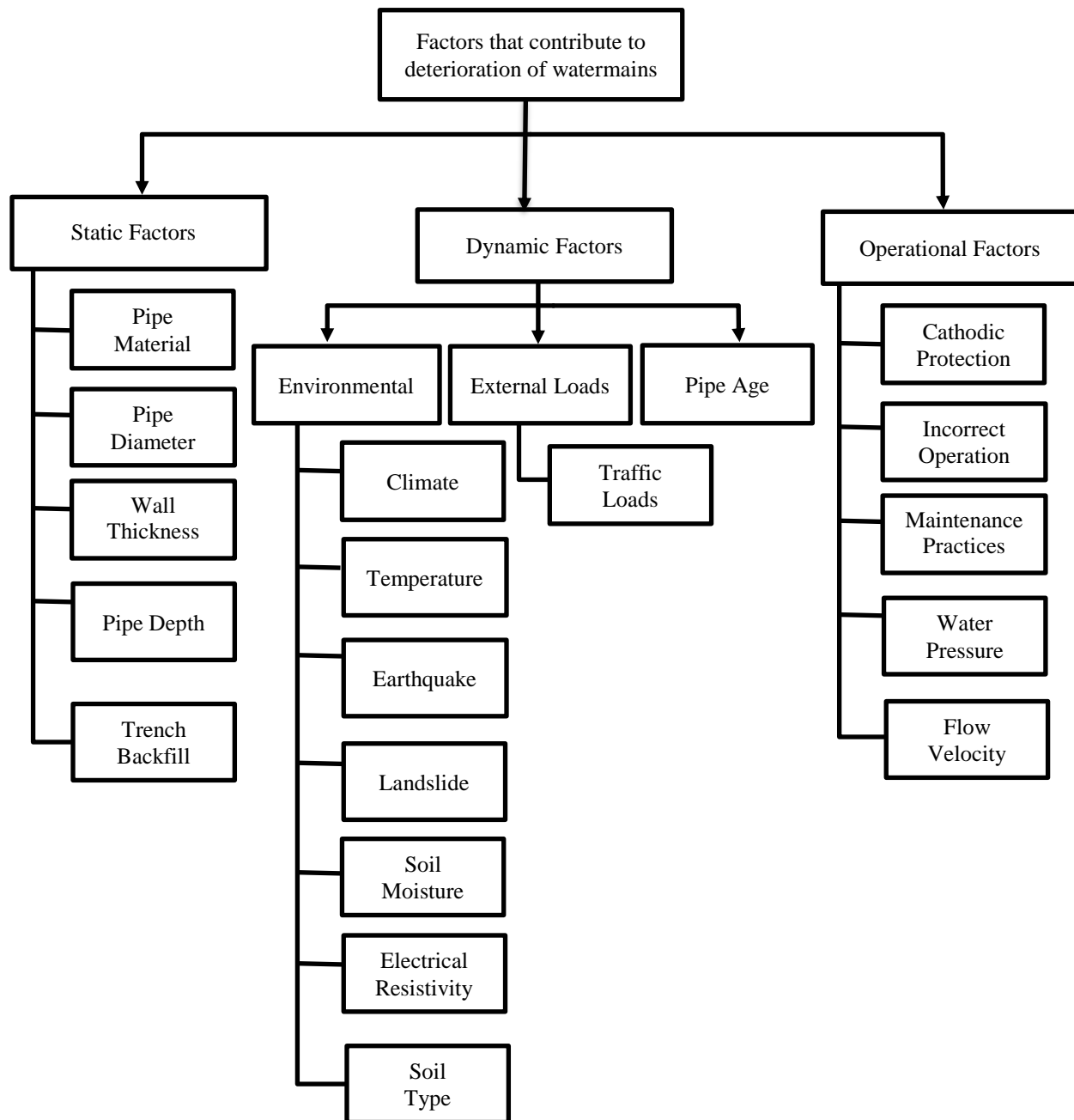**Fig. 10.** The error (RMSE) and sensitivity indices of pipe material.

Fig. 1 . Factors affecting the condition of watermains in both short and long terms.