

Research Article

Robust Anomaly Recognition in Hydraulic Structural Safety Monitoring: A Methodology Based on Deconfounding Boosted Regression Trees

Hairui Li ¹, Xuemei Liu ¹, Xiaolu Chen,² and Xianfeng Huai²

¹North China University of Water Resources and Electric Power, Zhengzhou, China

²China South-to-North Water Diversion Group Middle Route Co., Ltd., Beijing, China

Correspondence should be addressed to Xuemei Liu; lizgrf@163.com

Received 23 February 2023; Revised 19 July 2023; Accepted 25 July 2023; Published 16 August 2023

Academic Editor: Xiang Peng

Copyright © 2023 Hairui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Analyzing monitoring data to recognize structural anomalies is a typical intelligent application of structural safety monitoring, which is of great significance to hydraulic engineering operational management. Many regression modeling methods have been developed to describe the complex statistical relationships between engineering safety monitoring points, which in turn can be used to recognize abnormal data. However, existing studies are devoted to introducing the correlation between adjacent response points to improve prediction accuracy, ignoring the detrimental effects on anomaly recognition, especially the pseudo-regression problem. In this paper, an anomaly recognition method is proposed from the perspective of causal inference to realize the best exploitation of various types of monitoring information in model construction, including four steps of constructing causal graph, regression modeling, model interpretation, and anomaly recognition. In regression modeling stage, two deconfounding machine learning models, two-stage boosted regression trees and copula debiased boosted regression trees, are proposed for recovering the causal effects of correlated response points. The validation was carried out with Shanmen River culvert monitoring data, and experiment results showed that the proposed method in this paper has better anomaly recognition compared to existing regression modeling methods, as shown by lower false alarm rates and lower averaged missing alarm rates under different structural anomaly scenarios.

1. Introduction

A large number of hydraulic engineering projects in the world are in operation and maintenance. These projects may encounter engineering defects or structural failures caused by design defects, flooding, geological movement, material aging, etc., during operation, which may result in serious accident hazards. Understanding whether there are safety hazards in the structure through activities such as engineering monitoring and inspection is an important task in the process of engineering operation and maintenance. For this reason, more and more technologies of sensors [1], data acquisition [2], and data analysis [3] are developed and applied to engineering structural safety monitoring [4] and engineering safety state evaluation [5], so as to improve the efficiency of safety monitoring and reduce the human cost of engineering operation and maintenance.

Data-based structural anomaly recognition is a monitoring data analysis technique that aims to establish recognition criteria for the presence of structural anomalies by mining and analyzing historical data [6]. In the process, common data mining methods include statistical feature extraction [7], cluster analysis [8, 9] and regression modeling [10, 11], which has been widely discussed in the engineering safety monitoring literature. Statistical feature extraction assumes that the extreme values, rates of change, and other features of structural response quantity should be consistent with engineering experience, and this method relies too much on the design of domain experts [12]. Cluster analysis assumes that monitoring data samples are distributed in clusters of different sizes, and “outliers” that are too far from the cluster centers or too small in density are recognized as anomalous data. This method takes into account the correlation between the environment quantity and the effect quantity but is prone

to misreporting sparse samples in extreme environments [13]. Regression modeling assumes that the environment quantity x directly affects the change of the effect quantity y in the form of load $f(x)$, and by regression modeling of historical data, a structural response model $y = f(x) + u$ can be constructed, and samples that do not satisfy the conditional distribution $P(y|x)$ are recognized as abnormal data.

It is not difficult to see that anomaly recognition methods based on regression modeling have more restrictive assumptions and thus are more reliable in recognizing structural anomalies and are widely studied [14]. However, the complex response mechanism between environment and response quantities, the nonlinear feature interaction and nonstationary properties of engineering safety monitoring data make it difficult to construct a reasonable regression model using simple statistical modeling techniques [15]. In recent years, the introduction of machine learning techniques has improved this situation [16]. Machine learning models not only provide sufficient parameters for data fitting [17] but also provide efficient optimization methods to learn the complex interactions between features and thus have higher prediction accuracy [12]. In addition, machine learning models can effectively identify long-term stable signals and short-term fluctuation signals in monitoring data to adapt to the changing environment–response relationship of engineering structure in different operation periods [18].

Usually, there is more than one measurement point for response variables of the same type. This is because regression analysis of single response point tends to give false alarms under limited monitoring of environmental variables, while correlated data consisting of multiple response points can assist in inferring the presence of abnormal loads on the structure. Engineering examples exhibit the consistent behavior of hydraulic structure under normal conditions [19]. However, existing studies tend to focus only on the prediction accuracy of regression models, especially assuming that the variance of response residuals u are uncorrelated [20], ignore that the inference of correlated data often requires special methods, otherwise the results are likely to be biased or yield paradoxes [21]. This brings significant risks to the task of structural anomaly recognition. Thus, there is still a lack of debiased anomaly recognition model construction methods within consideration of correlated response points in hydraulic structural safety monitoring.

This paper aims at providing a machine learning regression modeling idea from the perspective of causal inference, so that the introduced correlated response variables not only assist the inference of abnormal data but also can improve the variance of model prediction, thus reducing the chance of false alarm and missed alarm. To achieve this purpose, this paper constructs a causal graph between the physical quantities of structural monitoring, proposes two deconfounding regression modeling methods based on boosted regression trees, and validates these methods on the Shanmen River culvert. The structure of this paper is as follows: Section 2 describes the correlation and its mining value of adjacent response points in engineering safety monitoring, and reviews the literature taking account of adjacent response points in

regression modeling. Then, this paper proposes a structural anomaly recognition modeling process in Section 3, in which two deconfounding regression modeling approaches based on causal inference are proposed to improve the problem of robustness of anomaly recognition brought by the introduction of adjacent response points. The method proposed in this paper is applied and validated using Shanmen River culvert as an example in Section 4. Section 5 provides an overview of the advantages and limitations of the suggested models, and prospects for future work.

2. Literature Review

Existing studies have shown that many machine learning methods such as neural networks [22], support vector machines [23], random forests [24], and extreme learning machines [25], and among others can construct well-performing models for structural response trends prediction and structural anomaly recognition by their strong fitting ability. In machine learning research, it is a common idea to mine and exploit the correlation in the data as much as possible to improve the model prediction accuracy. As mentioned above, the correlation in structural safety monitoring data is not only reflected in the causation between environment and response points but also in the correlation between adjacent response points. Although the existing literature is optimistic about the prospect of utilizing correlations between response points, it remains cautious in research and application in structural anomaly detection [26].

There are two main problems with introducing correlations between adjacent response points when constructing structural response models. One problem is that along with the change of structural properties, the correlation between adjacent response points may change, i.e., there is a phenomenon of covariate drift with uncertainty [19]. Under normal operating conditions, adjacent response points will exhibit a high degree of correlation, but when the structure is subjected to abnormal loading, the correlation may either remain or change significantly, depending on whether the structure is experiencing an overall or local anomaly. Thus, if the constructed regression model relies too much on the correlation between adjacent response points, the anomaly recognition model will exhibit a lack of generalization ability.

Further, another problem is the interpretation of machine learning models [27, 28]. Machine learning models generally have a more complex structure, which gives them the ability to better explore the complex relationships between variables while also posing difficulties in model interpretation. Due to the lack of a unified theory and method to explain the constructed regression models, most of the existing studies only focus on the model prediction accuracy performance and lack the comparison and analysis of the significance of input features. This leads to the problem of model overfitting not being easily revealed after introducing the correlation of adjacent response points.

In contrast to without consideration of using adjacent response points in regression modeling (named “without consideration” in Table 1), there are three major approaches introducing this correlation in literatures. The most direct

TABLE 1: Regression modeling approaches consideration on adjacent response points and their limitations.

Modeling approach	Limitations
Without consideration	Insufficient monitoring of environmental factors, the regression model is prone to missing alarms
Direct modeling	The regression model suffers from severe overfitting and is prone to false alarms
Two-stage regression	The powerful nonlinear fitting capability of machine learning hinders overfitting correction
Separate regression	The integrated regression model ignores interactions between causal factors and correlation factors

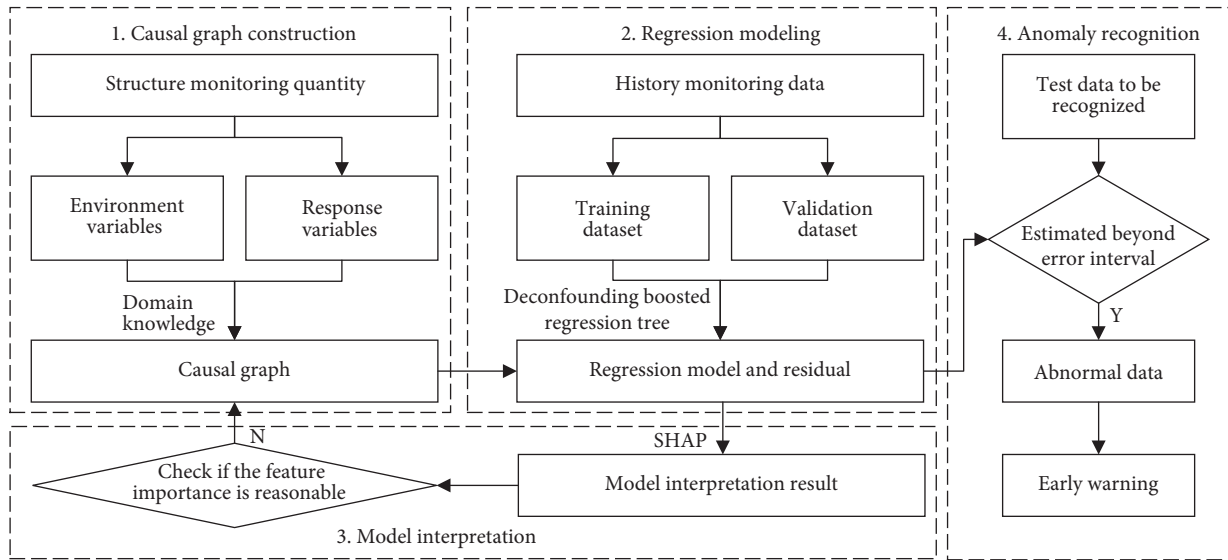


FIGURE 1: The overall workflow of anomaly recognition model construction and application.

way is to use the adjacent response points as the input features of regression model [20, 29], and then select a machine learning model with good generalization ability such as boosted regression trees (BRT) for training (named “direct modeling” in Table 1). Although the simulation experimental results show that model’s prediction performance get better, the problem of overfitting exists obviously. Another alternative approach is to construct the regressor of the adjacent response points using the environment quantity in first-stage regression, and then the prediction results of adjacent response points of first-stage regressor are used to construct the second-stage regression model with environment quantity as model inputs (named “two-stage regression” in Table 1) [30]. However, some studies have shown that introducing a nonlinear regressor constructed by machine learning in the first-stage regression can cause significant bias in the second-stage regression modeling [31]. In addition, there is another way of regression modeling using environment and adjacent response points separately and then weighting multiple regressors together to obtain an integrated regression model (named “separate regression” in Table 1) [19]. This approach distinguishes effect between causal factors and correlation factors but ignores the interaction between the two. The above-related studies are summarized in Table 1.

In summary, although existing studies consider the introduction of adjacent response points in the regression modeling phase, they lack the exploration of the anomaly recognition performance and require a suitable theoretical guide in regression correction strategy, so that there are many limitations in practice.

3. Methodology

As a common data analysis method, regression modeling has been used in many areas of business, with prediction tasks being the most popular. Prediction aims to make full use of the correlations in the data to estimate the changes of system under different scenarios; however, prediction models cannot always be used to make counterfactual inferences due to the possible confounding variables used in modeling the system, which makes the constructed statistical models suffer from pseudo-regression problems. Consequently, the introduction of causal analysis methods to remove confounding effects is usually considered in decision-making tasks. Deconfounding regression modeling, i.e., combining causal analysis methods to correct the process of regression modeling or the estimated results, often entails drawing a graph of what may be causing what, identifying confounders, and stratifying those to find the effect of a treatment on an outcome. Doing this properly helps decision makers stay clear of absurd claims.

This paper proposes deconfounding regression modeling methods to implement a structural anomaly recognition technique based on the analysis of causal relationships among measurement points, including four steps of constructing causal graph, regression modeling, model interpretation, and anomaly recognition, as shown in Figure 1. First, according to the layout of measurement points in the monitoring section, the measurement points are classified into exogenous variables (adjacent response points) and endogenous

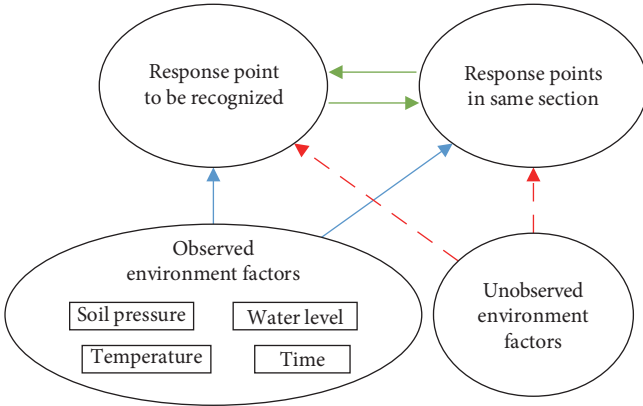


FIGURE 2: The causal graph of measurement points in the monitoring section.

variables (environment points) using domain knowledge, and a cause graph is constructed to guide the next regression modeling process. Then, in the regression modeling stage, inspired by econometrics, two improved machine learning modeling methods are proposed to minimize the effects arising from confounding bias. After the regression modeling is completed, a suitable model interpretation method is selected to understand whether the endogeneity problem of the regression model is mitigated by comparing the importance of different features at sample estimation. Finally, in the anomaly recognition stage, the anomaly discriminant interval is considered, which is used to establish the anomaly recognition rules. The details are shown below.

3.1. Causal Graph Construction. A reasonable hypothesis on the causal relationship between variables is a prerequisite for causal inference. During the design and construction of engineering safety monitoring, various types of sensors are generally installed in a vertical monitoring section. These sensors include several response points for monitoring the structural response of interest, such as displacement, seepage, stress, strain, etc., and also some environment points for monitoring the environment factors that produce loads on the structural response, such as water level, temperature, etc. Although the historical monitoring data collected by different sensors are correlated with each other, the reasons to produce the correlation are not the same. First, changes in environment quantities are the direct cause of changes in response quantities, and thus there is causality between the environment and response quantities. In contrast, the environment factors and the principle of producing load on response measures in the same section are similar, and thus these response points show a high correlation, but this correlation does not mean that there is a strong direct interaction between them. Based on the above analysis, a causal graph is constructed, as shown in Figure 2.

There are three paths in Figure 2 to reveal the source of correlation between the adjacent response points, which are indicated by green, blue, and red arrows. The green path indicates that the engineering structure as a whole, local load changes will have some direct effects on the loads at

other locations. The blue path and the red path indicate that the environment quantities as confounding variables result in correlations between the same section response points, where the blue path is the observed environment quantity, and this confounding effect can be eliminated by the condition control of the environment quantity, while the red path represents the confounding effect of the unobserved environment quantity, and this confounding effect is not easily eliminated. It is worth noting that there is another reason for the correlation with response points, which is the most of the monitoring data are collected in good conditions, which makes the monitoring data used for regression modeling subject to sample selection bias. In econometrics, sample selection bias can also be considered as a confounding effect of unobserved variables.

In the causal inference perspective, in order to restore the causal effects between adjacent response points, it is key to condition the environment quantities, otherwise the constructed regression model will suffer from endogeneity problems. Endogeneity refers to the presence of confounding factors both affecting x and y , leading to inequality of statistical association distribution and interventional distribution, i.e., $P(y|x) \neq P(y|do(x))$, this is also referred to as confounding bias. In econometrics, endogeneity is described as the presence of correlation between x and u in the ideal regression model $y = f(x) + u$. As an example, in a linear model $y = \beta x + u$, if x is an endogenous variable, it is correlated with the residual u , that is the covariance of two variables $cov(x, u) \neq 0$. When doing ordinary linear regression, the model parameter β would be biased estimated for:

$$\hat{\beta} = \frac{cov(y, x) - cov(x, u)}{cov(x, x)} \neq \frac{cov(y, x)}{cov(x, x)}. \quad (1)$$

3.2. Regression Modeling. Machine learning-based regression modeling methods can be very good at learning correlations between complex factors, and in practice, often the more correlated features have a greater impact on the estimation results. However, correlation is not equal to causality, and the purpose of causal inference is to eliminate confounding bias as much as possible and identify causality from correlation. In this paper, inspired by the methods of dealing with endogeneity problems in econometrics, we propose two deconfounding regression modeling methods based on boosted regression trees models, including two-stage BRT (TSBRT) and copula debiased BRT (CDBRT). Assume that the ideal regression model has the form $y = f(x, z) + u$, where z is the exogenous variable, i.e., the environment quantity, and x is the endogenous variable, i.e., the adjacent response quantity. The principle of the two methods proposed in this paper is to reduce the correlation between x and u as much as possible in the regression modeling process, and the difference between the two methods is whether the confounding effect from unobserved environment factors is considered in the correction process, as described below.

3.2.1. Boosted Regression Trees. BRT is an ensemble learning model based on classification and regression tree (CART), where the model constructs multiple decision trees with training data and sums the estimates of each decision tree as the final estimation at the prediction stage. The method fully combines the flexibility of CART with the effectiveness of the boosting learning mechanism. CART serves as a sub-model to compose BRT, and during the regression model training, the algorithm is used by recursively dividing the dataset into smaller subsets until certain stopping conditions are satisfied, such as minimum number of samples or maximum tree height.

Decision tree splits during model learning with dataset $R = \{(x, y)\}$, the dataset is divided as $R_1(j, s) = \{x|x_j \leq s\}$ and $R_2(j, s) = \{x|x_j > s\}$, recursively. Everytime subset is divided, the CART algorithm selects a split sample with feature j and feature value s that satisfies the minimum mean square error (MSE) condition in a greedy strategy:

$$\operatorname{argmin}_{j,s} \sum_{x_i \in R_1(j,s)} (y_i - \hat{c}_1)^2 + \sum_{x_i \in R_2(j,s)} (y_i - \hat{c}_2)^2. \quad (2)$$

In Equation (2), the estimate value \hat{c} is the mean of the labels y of the samples contained in each child node. After the decision tree h is built, a common regularization technique to avoid overfitting is to reduce the complexity of the decision tree by pruning. The regularized loss function $L(h)$ is:

$$L(h) = \text{MSE}(h) + \alpha|h|, \quad (3)$$

where $\text{MSE}(h)$ is the mean square error of samples estimated by h , $|h|$ is the number of leaf nodes, and α is a predefined regularization weight to trade-off different losses. The pruning process is traversed upward from each leaf node, and the parent node is pruned if the regularized loss function will become smaller after the parent node is deleted.

The CART algorithm can handle both continuous and discrete variables, so it is well-suited to handle various types of datasets. However, the fitting ability of a decision tree as a weak classifier is very limited. BRT performs multiple rounds of training based on the boosting mechanism for each build a group of decision trees. After a decision tree $h_i(x)$ is trained, the regression function of the whole model $f_i(x) = f_{i-1}(x) + \lambda h_i(x)$, where λ is the learning rate, and the fitted residual of the sample $u_i = y - f_i(x)$ is just the gradient of trained regression model $f_i(x)$ and will be used as the new label y_{i+1} for the next round of decision tree training.

The hyperparameters in model training process include max iteration number M , minimal leaf sample, max tree depth, regularization weight α , and learning rate λ . These parameter values need to be set empirically during model training, and the values taken in this study are shown in Section 4.2. The decision tree construction process uses different features for division, thus can better mine the interaction effects between variables, and decision tree is easier to explain. However, when the decision trees are complex or too

many, it still requires sophisticated interpretation tools, which will be discussed in Section 3.3.

3.2.2. Two-Stage Boosted Regression Trees. When causal inference models are constructed, the input features are divided into two categories, one for exogenous variables z and the other one for endogenous variables x . The difference between them is that endogenous variables may lead to endogeneity problems when they are used directly in regression modeling. Two-stage regression is a common idea in causal inference, where in the first stage, a regression model is constructed for each endogenous variable x , and then in the second stage, exogenous variables z and the estimated endogenous \hat{x} are used to construct a regression model for the variables y to be estimated.

Based on the above principles, this paper proposes a two-stage regression model based on BRT named TSBRT, in which the environment factors are considered as exogenous variables z and the other response points in same section are considered as endogenous variables x . In the first stage regression modeling progress, a part of correlation between different response points produced by unobserved environment factors (red path in Figure 2) is filtered out, and another part produced by common observed environment factors (blue path in Figure 2), thus the second-stage regression model predict account for adjacent response points with remaining correlation (green path in Figure 2). The algorithm details are shown in Algorithm 1.

It is worth noting that original BRT is an ensemble learning method, which means that each decision tree generated in the modeling is a weak regressor. The proposed TSBRT constructs the regression model $f_x(z)$ for the endogenous variable x and the regression model $f(z, \hat{x})$ for the variable y in a stepwise “tree by tree” manner, instead of constructing the whole regression model one by one. The reason for this is that it has been found in practice that machine learning methods have a strong nonlinear fitting ability, and thus the instrumental variables generated tend to introduce an overfitting risk to the second-stage regression training, and empirically the complexity of the first-stage trained regression model should not exceed that of the second-stage model.

3.2.3. Copula Debiased Boosted Regression Trees. Filtering out correlations between response points produced by unobserved environment factors is a conservative modeling strategy. Retaining a portion of the correlation between response points under reasonable assumptions can further improve the prediction accuracy of the model while ensuring model robustness. As there are no suitable instrumental variables to estimate the unobserved environment factors, some instrument-free modeling approaches in econometrics can be introduced.

In causal inference, there is endogeneity in the constructed regression models, implying that there is a correlation between the endogenous variables x and the model residual u . The problem of endogeneity can be greatly alleviated if the joint distribution between the endogenous variables and the model residuals can be modeled. Copula is a

```

Input: Train data collection  $C = \{(z, x, y)\}$ , loss function  $L$ , max iteration number  $M$ , and learning rate  $\lambda$ 
Output: Second-stage regression function  $\hat{y} = f_M(z, \hat{x})$  and each first stage regression function  $\hat{x}^i = f_M^i(z), i = 1, 2, \dots, |x|$ 
1  $f_0^i(z) = \text{mean}(x^i), i = 1, 2, \dots, |x|$ //Initialize each first stage regression function
2  $f_0(z, \hat{x}) = \text{mean}(y)$ //Initialize second-stage regression function
3 while  $m = 1 \rightarrow M$  do
4   while  $i = 1 \rightarrow |x|$  do
5      $u_m^i = -\frac{\partial L(x^i, f_{m-1}^i(z))}{\partial f_{m-1}^i(z)} = x^i - f_{m-1}^i(z)$  //Get residual's gradient of  $f_{m-1}^i(z)$ 
6      $h_m^i(z) \leftarrow \text{CART}(C_m^i = \{(z, u_m^i)\})$ 
7      $f_m^i(z) = f_{m-1}^i(z) + \lambda h_m^i(z)$ 
8     //Train a new CART tree to update regression function of  $\hat{x}^i = f_m^i(z)$ 
9   end
10   $v_m = -\frac{\partial L(y, f_{m-1}(z, \hat{x}_m))}{\partial f_{m-1}(z, \hat{x}_m)} = y - f_{m-1}(z, \hat{x}_m)$  //Get residual's gradient of  $f_{m-1}(z, \hat{x}_m)$ 
11   $g_m(z, \hat{x}_m) \leftarrow \text{CART}(C_m = \{(z, \hat{x}_m, v_m)\})$ 
12   $f_m(z, \hat{x}_m) = f_{m-1}(z, \hat{x}_m) + \lambda g_m(z, \hat{x}_m)$ 
13  //Train a new CART tree to update regression function of  $\hat{y} = f_m(z, \hat{x}_m)$ 
14 end
15 Return  $f_M(z, \hat{x})$  and  $f_M^i(z), i = 1, 2, \dots, |x|$ 

```

ALGORITHM 1: Training of a two-stage boosted regression trees model.

tool for constructing complex joint distributions between variables, and this section proposes a copula-based modeling improvement strategy.

Using the copula method, the correlation between the two can be modeled, that is, the joint distribution of x and u , and then more consistent model parameter estimates can be obtained. According to Sklar theorem, for two variables with marginal distributions H and G , there exists a copula function c such that the joint distribution:

$$F(x, u) = c(H(x), G(u)) = c(U_x, U_u). \quad (4)$$

In Equation (4), $U_x = H(x)$ and $U_u = G(u)$ are variables generated by probability integral transformation that satisfies the uniform distribution in $[0, 1]$, which means their probability density functions satisfy $f(x, u) = c(U_x, U_u)h(x)g(u)$. Similar to many existing model assumptions, this paper uses a Gaussian copula function to establish the joint density function of endogenous features x and residual u :

$$f(x, u) = \frac{1}{(1 - \rho^2)^{1/2}} \exp \left[-\frac{\rho^2 (\Phi^{-1}(U_x))^2 + \Phi^{-1}(U_u)^2}{2(1 - \rho^2)} + \frac{\rho \Phi^{-1}(U_x) \Phi^{-1}(U_u)}{(1 - \rho^2)} \right] h(x)g(u). \quad (5)$$

In Equation (5), ρ is the correlation coefficient of endogenous features x and residual u , U_x can be obtained through nonparametric density estimation method. Assuming u is sampled from a normal density function with mean 0 and standard error σ_u^2 . This linear Gaussian correlation assumption implies that the correlated residual u can be rewritten as:

$$\tilde{u} = \sigma_u \cdot \rho \cdot \tilde{x} + \sigma_u \cdot \sqrt{1 - \rho^2} \cdot \tilde{w}, \quad (6)$$

where \tilde{x} is got from $\Phi^{-1}(U_x)$ and \tilde{w} is a variable sampled from standardized normal distribution [32]. The implementation of the proposed method introduces a generated feature \tilde{u} in the process of BRT training as Algorithm 2 described.

3.3. Model Interpretation. In a linear regression-based causal inference task, although it is impossible to confirm whether the parameters learned by the model are in accordance with causality, it is possible to judge whether the improved model is more consistent with engineering experience and

perception based on the changes in model parameters before and after the implementation of the deconfounding method, such as a significant increase in the weights of the causal variables. However, the model parameters involved in machine learning methods are often very numerous and complex in their relationships, which poses a great challenge to model parameter identification.

In recent years, the development of explainable artificial intelligence techniques has provided many tools to help model users understand the behavior of the trained model. Researchers tend to focus on the degree to which each feature of the model input affects the model output, often referred to as feature importance [33], and in linear regression models, feature importance is measured in the form of variable coefficients. Although decision tree-based feature importance quantification methods have been proposed for a long time, the traditional interpretation methods based on the importance of decision tree features are often misleading [34]. To address this problem, Lundberg and Lee [35] proposed SHapley

Input: Train data collection $C = \{(z, x, y)\}$, correlation coefficient ρ , loss function L , max iteration number M , and learning rate λ

Output: Regression function $\hat{y} = f_M(z, x)$

```

1  $f_0(z, x) = \text{mean}(y)$  //Initialize regression function
2 while  $m = 1 \rightarrow M$  do
3    $v_m = -\frac{\partial L(y, f_{m-1}(z, x))}{\partial f_{m-1}(z, x)} = y - f_{m-1}(z, x)$  //Get residual's gradient of  $f_{m-1}(z, x)$ 
4    $\tilde{C}_m(z, x) \leftarrow \text{CART}(\tilde{C}_m = \{(z, \tilde{x}, \tilde{v}_m)\})$ 
5   //Train CART tree with bootstrap method and sample residual's standard error  $\sigma_{u, m}$ 
6    $\tilde{u}_m = \sigma_{u, m} \cdot \rho \cdot \tilde{x} + \sigma_{u, m} \cdot \sqrt{1 - \rho^2} \cdot \tilde{w}$ 
7    $h_m(z, x) \leftarrow \text{CART}(C_m = \{(z, x, v_m - \tilde{u}_m)\})$ 
8    $f_m(z, x) = f_{m-1}(z, x) + \lambda h_m(z, x)$ 
9   //Train a new CART tree to update regression function of  $\hat{y} = f_m(z, x)$ 
10 end
11 Return  $f_M(z, x)$ 

```

ALGORITHM 2: Training of a copula debiased boosted regression trees model.

Additive exPlanations (SHAP), a computational framework dedicated to unifying the field of interpretable machine learning based on the concept of Shapley value.

Assuming that players cooperate in a union and receive a certain amount of income from this cooperation, for a joint game, each player contributes differently to the total income, and thus a method is needed to allocate income to players based on their contribution to the total spent. Similarly, regression is the process of estimating the label jointly based on multiple input features, and SHAP uses Shapley values to explain the importance of each feature in the model, gradually becoming a general method for interpretability of complex models.

As stated in Section 3.2.1, the decision tree model estimates the sample by traversing different sets of features s and eventually obtaining different estimates $g(s)$, depending on the splitting point. SHAP takes each input feature of the machine learning model as player i , then for the arrangement of all features $z' \in \{0, 1\}^M$, there is a linear equation for the constructed explanation model:

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i. \quad (7)$$

In Equation (7), $z'_i = 1$ means that feature i participate in s , otherwise it is 0, and ϕ_i represents the contribution of feature i . The conditional expectation $h(x) = E[f(x)|x_s]$ of different features sets s are obtained by estimating a large number of samples of the model $f(x)$. And the final explanation of the feature importance is the marginal contribution of a player i in various alliance s , thus $h(s \cup i) - h(s)$.

3.4. Anomaly Detection. After a regression equation $f(x)$ has been trained, it can be used to perform anomaly recognition tasks. This is done by first estimating the label y of data $\{(x)\}$ to be tested with $\hat{y} = f(x)$, and discriminate data (x_i, y_i) is abnormal, if Equation (8) is satisfied:

$$|y_i - \hat{y}_i| > 3\sigma. \quad (8)$$

In Equation (8), σ is the sample standard error of the residuals u of the regression model $f(x)$ in the validation dataset. The basic principle of this discriminant rule is that if $f(x)$ learns the mapping $x \rightarrow y$ well, the sample residuals u of the normal data should conform to the normal distribution with means 0 and standard error σ^2 . The statistical meaning of 3σ is that, y is in about 95% confidential interval of $f(x)$, which is a common setting in literatures. This threshold is given as 2σ in some papers as well. From an engineering point of view, when the estimated response value exceed the anomaly discrimination interval, it indicates that the structure is subjected to a load that cannot be explained, i.e., the identification of a possible anomaly in the current structure.

According to Equation (8), it can be found that an anomaly recognition model with good performance tends to have two properties. On the one hand, $f(x)$ has good predictive performance in the validation dataset, which means that the standard error of the model residuals is smaller, and thus a narrower anomaly discrimination interval allows the model to identify minor abnormal loads at the early stage when structural anomalies occur. On the other hand, $f(x)$ needs to have a good generalization ability, i.e., the model reasonably extracts the feature information as the main basis for estimation. The engineering operation process is complex and the distribution of environmental variables may change. Better generalization ability can better adapt to covariate shifts and reduce the possibility of false alarms and missed alarms.

However, the two goals often conflict in practice, especially when model training pursues only in-bag accuracy, it may mislead machine learning with the problem of overconfidence and poor generalization performance. Thus, the trade-off problem of variance-bias needs to be considered. The main focus of this paper is on how to effectively use the correlation between adjacent response points to improve the

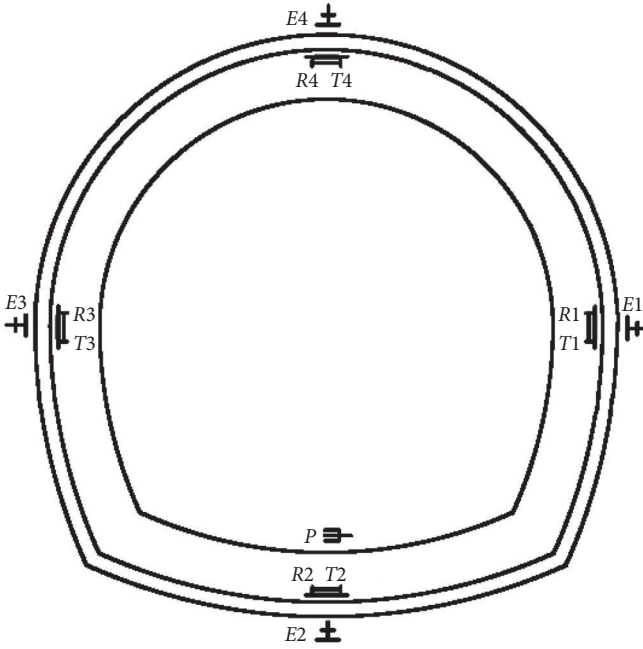


FIGURE 3: Monitoring equipment layout of Shanmen River culvert.

anomaly recognition performance. This paper argues that because the correlation between adjacent response points is unstable, directly using them as input features for model training may easily lead to insufficient generalization ability of the regression model. Considering that the forms of structural anomalies are diverse, such as single point anomalies, partial structural anomalies, and overall structural anomalies, this paper aims to present the proposed deconfounding BRT methods have stable anomaly recognition ability, as tested in Section 4.

4. Case Study

The case study concerns the Shanmen River culvert in Jiaozuo with coordinates (113.19094° E, 35.162414° N) according to the World Geodetic System (WGS84). Culvert is a common water delivery building and is widely constructed in South-to-North Water Diversion Project. Shanmen River culvert was completed in June 2012 and has been in use since December 2014. The culvert is a pressureless tunnel, 550 m long. The maximum excavation height of the cave is 11.75 m, the span is 11.75 m, and the distance between the left and right holes is 24 m. The cover layer between the bottom of Shanmen River and the culvert is thin, and the minimum burial depth is about 18 m. The surrounding rocks are loose laminated bulk structure of the fourth series, mainly composed of pebbles and heavy powdery loam, in which the pebble cementation is mostly uneven, staggered layers or lenticular distribution, and the distribution is more irregular. The instruments layout of SHM culvert monitoring section is shown in Figure 3.

4.1. Regression Model Construction. Under the influence of material aging, geological movement, or other factors during operation, the surrounding rock structure may be cracked

and damaged, which may result in structural instability. By monitoring the changes of the steel stress in the surrounding body, it is possible to recognize structural anomaly and take disposal measures in time. As illustrated in Figure 3, there are four steel stress gauges (R1–R4) that were installed on the palm cavity at 3, 6, 9, and 12 o'clock, respectively. Some environment sensors were installed to perceive the potential load from environment changing near them, the environment sensors including four temperature gauges (T1–T4) to monitor the steel temperature, four soil pressure gauges (E1–E4) to monitor load from the river above and surrounding soil, and a water level gauge (P) to monitor water level in the cavity. In addition, taking into account the effect of concrete creep is related to time, the input features often include time t in regression process.

To achieve regression-based structural stress anomaly recognition, the key step is to construct a regression model for each steel stress measurement point. The main question discussed in this paper is how to effectively use the correlation between adjacent response points to compensate for the incomplete monitoring of environment factors. To demonstrate the effectiveness of the proposed method in this paper, two existed model construction ideas are implemented for comparison. One is a causal model, in which the model selects some of the environment variables that directly generate loads on the structural stress measurement points as input features. In addition, a noncausal model is constructed by introducing other steel stress measurement points as input features, considering the existence of unobserved environment factors on the structure. The input features of the two regression models are shown in Table 2. Taking the regression model for steel stress gauge R3 as an example, the input features of the noncausal model include not only the environment variables E3, T3, P, and t but also the steel stresses R1, R2, and R4 at other positions.

As described in Section 3.3, the input features of the two improved models proposed in this paper are identical to those of the noncausal model. The difference is that the regression model proposed in this paper divides the input features into environment variables z and adjacent steel stress points x . The noncausal model directly constructs the mapping of input features to the estimated steel stress \hat{y} as $(x, z) \rightarrow \hat{y}$, while TSBRT estimate $z \rightarrow \hat{x}$ before constructing a mapping of $(\hat{x}, z) \rightarrow y$, and CDBRT constructs a mapping $(x, y) \rightarrow \tilde{u}$, then $(x, z) \rightarrow \hat{y} - \tilde{u}$ is constructed.

4.2. Model Training and Interpretation. In this paper, monitoring data of Shanmen River culvert from 2,018.6 to 2,022.6 were used for model training and validation. Regression modeling assumes that the data used for model training and accuracy evaluation are all data collected under normal working conditions, while the actual engineering monitoring process may generate error data due to problems such as unstable instrument operation or manual mistakes. Before the model training, the monitoring data were processed for the coarse error exclusion, which was based on the Grubbs criterion. For each measurement point, the historical data were sorted according to the measured value from smallest

TABLE 2: Regression models and related input features.

Input feature	Causal model	Noncausal model
Soil pressure	Adjacent soil pressure gauge T	Adjacent soil pressure gauge T
Temperature	Adjacent temperature gauge E	Adjacent temperature gauge E
Water level	P	P
Time	t	t
Noncausal features	—	Other steel stress gauges R

TABLE 3: Mean absolute error (KPa) of different BRT models.

Model	$R1$	$R2$	$R3$	$R4$	Average
Causal model	0.387	0.163	0.234	0.643	0.357
Noncausal model	0.079	0.085	0.068	0.163	0.099
Two-stage BRT	0.116	0.124	0.128	0.252	0.155
Copula debiased BRT	0.08	0.091	0.076	0.182	0.107

to largest. The data series of each variable x is obtained, and then the following equation is calculated:

$$g_i = \frac{\text{abs}(x_i - \bar{x})}{\sigma}. \quad (9)$$

In Equation (9), x_i is the i th sample in dataset, \bar{x} is the sample mean of $\{x\}$, σ is the sample standard error, and abs means the absolute value of a scalar value. If $g_i \geq g_0(i, 0.05)$, then the sample data x_i is judged to be coarse error and is excluded, and $g_0(i, 0.05)$ is the threshold value of the judgment criterion, which is obtained by checking the Grubbs threshold table.

After coarse error exclusion, 70% of the historical data is used for model training and the remaining 30% is used to evaluate the model performance. The same hyperparameters were used for all model training processes, including a training loss function of MSE, a max iteration number of 50, a minimal leaf sample of 5, a max tree depth of 3, a regularization weight α of 1.0, and a learning rate λ of 0.1. Comparing the model residuals using different models and constructing different measurement points, the model prediction accuracy of test data was measured using the mean absolute error, as shown in Table 3.

From the experimental results, it can be seen that the causal model has the lowest model prediction accuracy because it only considers the observed environment factors. In contrast, the other models have higher model prediction accuracy because they consider the effect of unobserved environment factors effect on other response points. The noncausal model still has the highest prediction accuracy, but as the results of the analysis in Section 3.4 show that the model uses the correlation between response points as the main basis for prediction and thus has serious robustness problems in the anomaly recognition stage.

As stated in the content of Section 3.3, a good anomaly recognition model should not only have more accurate prediction but also a reasonable distribution of feature importance. Still taking the steel stress measurement point $R3$ as an

example, the feature importance of each model is expressed in SHAP summary plot, as shown in Figure 4. The summary plot ranks each input feature from top to bottom according to feature importance. Each row depicts the relationship between the feature value and its SHAP value, with the magnitude of the feature value indicated by color and the SHAP value indicated by the horizontal coordinate of the sample point.

The summary plot of the causal model shows that the model considers the environment factors to be influencing in the order of soil pressure, water level, time, and temperature on the steel stress. The summary plot of the noncausal model shows that although the ranking of the importance of environmental factors by the noncausal model does not change, the model bases its predictions more on the stress measurement points at other locations, where the importance of the $R1$ feature in particular almost dominating. Then, it can be seen that the distribution of feature importance of two improved models is more reasonable. First, there is no change in the importance ranking of the observed environment factors by two improved models compared to causal model. Second, compared to noncausal model, two improved models consider the importance of the observed environmental factors to be higher, but the correlation between adjacent response points has a nonnegligible influence, as shown by the fact that the $R2$ measurement point with the strongest data correlation is no longer used as the dominant basis in the prediction, and its importance is lower than that of the soil pressure $E3$.

4.3. False Alarm Performance. Regression-based anomaly recognition method uses a regression model to accomplish a classification task. Although regression models are trained with similarly distributed data, some training data cannot be fitted well when the input features considered for constructing regression models are not sufficient, such that some normal data in the historical data can be misclassified as abnormal data. In the anomaly recognition task, this phenomenon is called false alarm, and the generated alarms do not imply that there is a potential risk to structural safety and thus do not directly have serious consequences. However, a high rate of false alarms can lead to frequent use of human

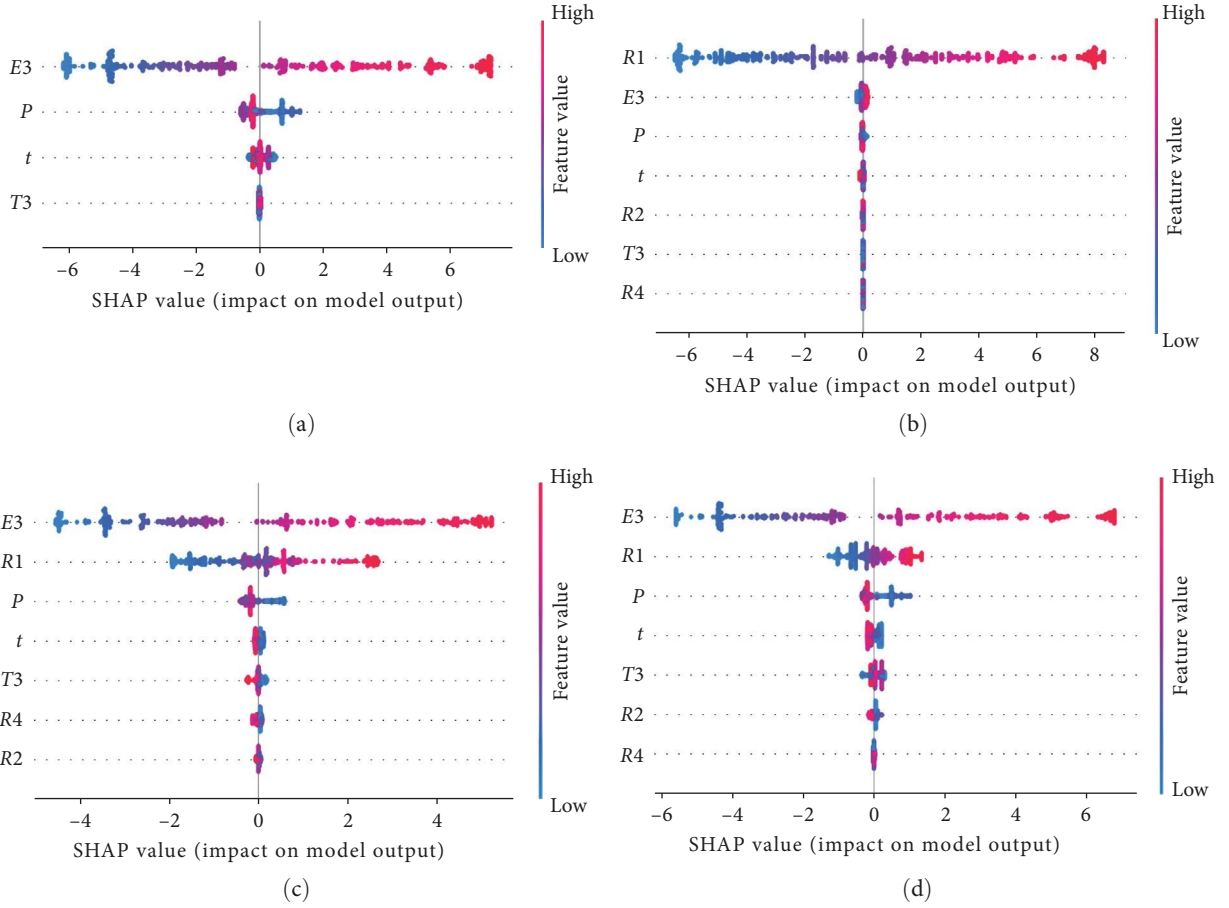


FIGURE 4: SHAP summary plot of each models: causal model summary plot (a); noncausal model summary plot (b); TSBRT summary plot (c); and CDBRT summary plot (d).

TABLE 4: False alarm rate (%) of different BRT models.

Model	R1	R2	R3	R4	Average
Causal model	7.91	9.28	8.47	7.53	8.30
Noncausal model	1.61	4.88	2.27	1.98	2.69
Two-stage BRT	3.24	7.06	5.82	5.29	5.35
Copula debiased BRT	1.63	5.14	2.71	2.15	2.91

resources for verification, which in the long run can lead to mistrust of the model by the operation managers, so that when a real structural anomaly is identified by the model, it cannot be effectively brought to the attention of the operation managers. Thus, false alarms need to be attended to.

In order to compare the false alarm performance of different models, this paper performs anomaly recognition on the validation dataset, and the false alarm performance of the model can be evaluated based on the ratio of the number of anomalies recognized by the model. The false alarm rates of different models and different structural stress measurement points are shown in Table 4. Comparing with Table 3, we can find that the model false alarm rate is inversely proportional to the prediction accuracy of the model in general. The noncausal model has the lowest false alarm rate, and the two

models proposed in this paper, TSBRT and CDBRT, perform close to it, while the noncausal model has the highest false alarm rate. Anomaly recognition on selected data of stress measurement R3 is plotted and compared, as shown in Figures 5 and 6.

By comparing the anomaly recognition plots, the differences between different models can be identified. As mentioned in Section 3.2, regression model tends to be less effective in good operating conditions of the structure if it does not take into account the effects of unobserved environmental factors. Figure 5 depicts the differences between the causal model and the noncausal model for anomaly recognition on the validation dataset. The causal model has a lower prediction accuracy and a wider anomaly discrimination interval. At the same time, when the unobserved environmental

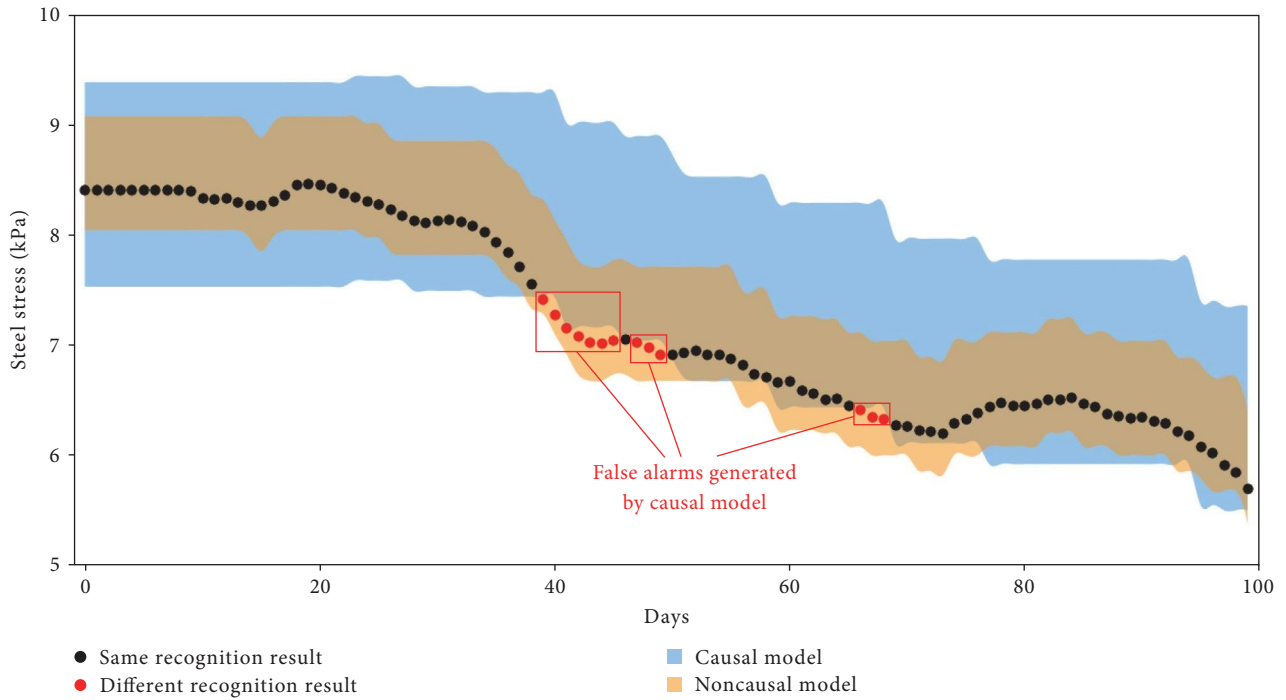


FIGURE 5: Anomaly detection result of causal model and noncausal model.

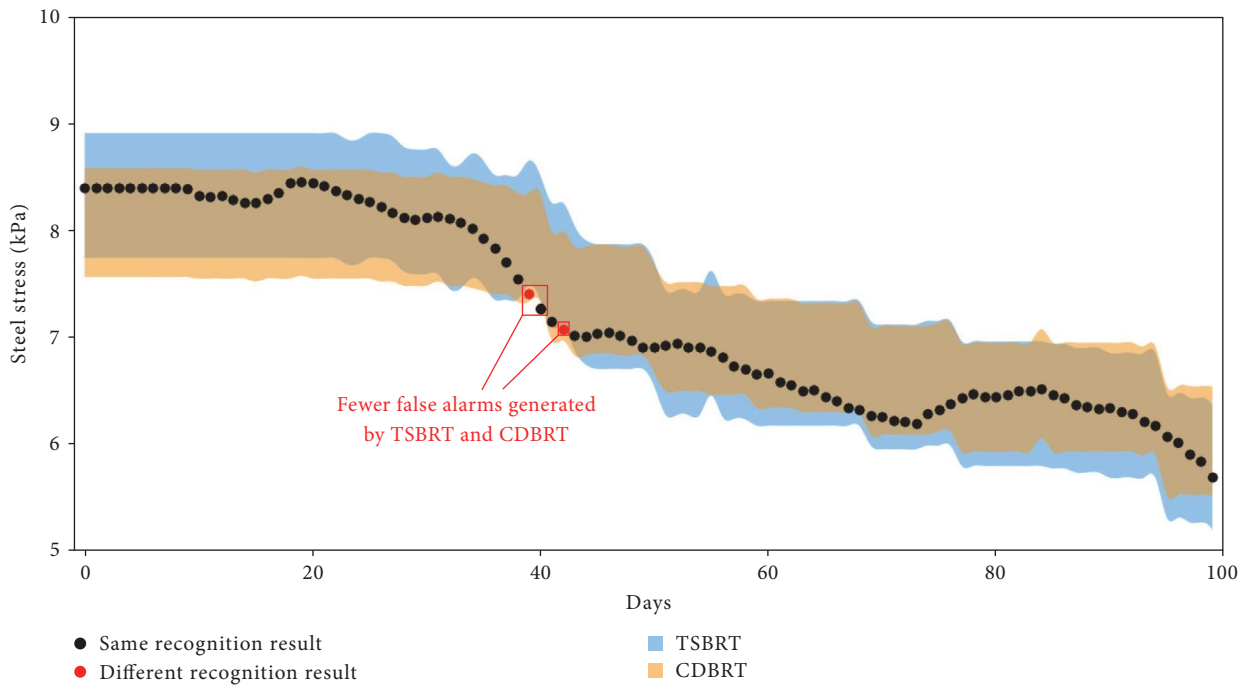


FIGURE 6: Anomaly detection result of TSBRt and CDBRt.

variables changed, the causal model was unable to identify them and could easily misclassify the structure as anomalous. The improved TSBRt and CDBRt take into account the influence of unobserved environmental variables in the modeling process from adjacent response points, and therefore rarely misjudge, as shown in Figure 6.

4.4. Anomaly Recognition Performance. The key purpose of anomaly recognition models is to accurately identify abnormal data when structural anomalies occur, and this section uses simulated data to evaluate the anomaly recognition performance of different models. In order to verify the models' robustness of anomaly detection, this paper processes the

TABLE 5: Anomaly detection accuracy of models in different scenarios.

Model	Scenario 1	Scenario 2	Scenario 3	Average
Causal model	0.862	0.862	0.862	0.862
Noncausal model	0.936	0.824	0.745	0.835
TSBRT	0.917	0.912	0.908	0.912
CDBRT	0.922	0.919	0.904	0.915

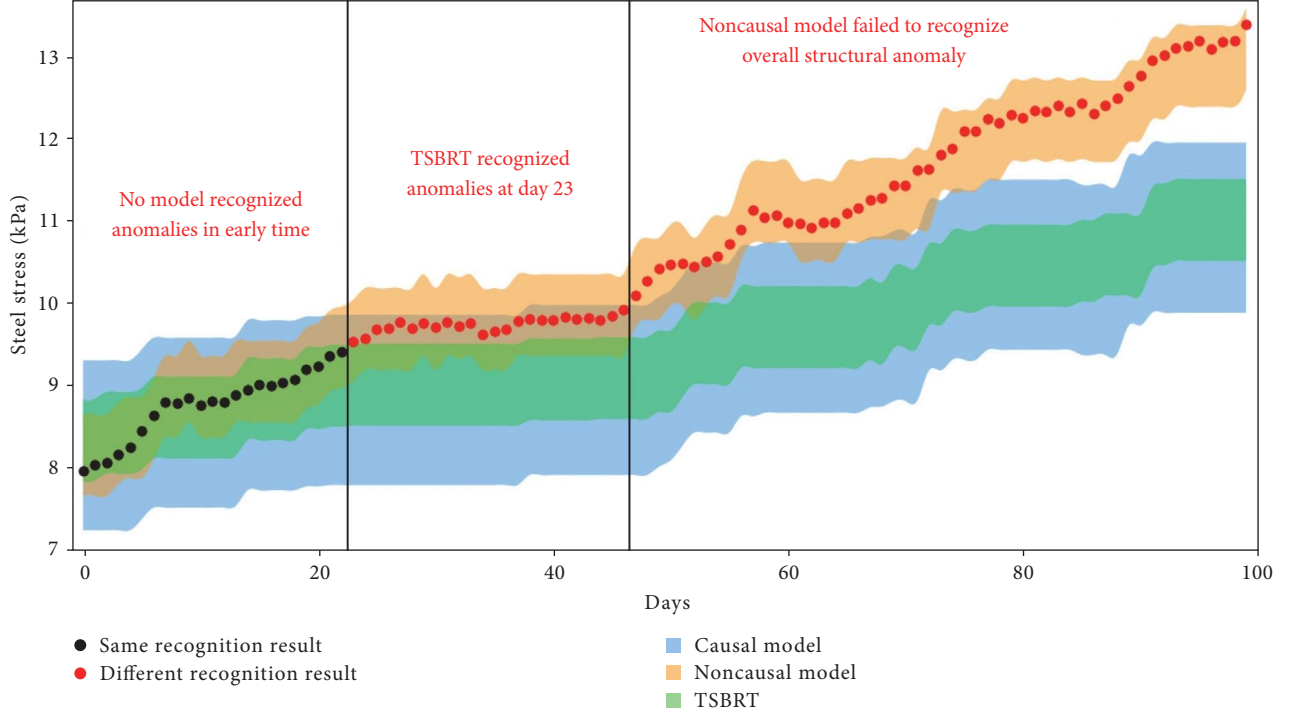


FIGURE 7: Anomaly detection result of causal model, noncausal model, and TSBRT in scenario 3.

raw monitoring data to simulate the different forms of anomalies that occur on the culvert structure. This is done by randomly selecting 30% of the segmented data from the original dataset, and then applying different forms of deviations to these segmented data to simulate three structural anomaly scenarios, including single measurement point local anomaly (scenario 1), nonuniform overall anomaly (scenario 2), and uniform overall anomaly (scenario 3). They are generated by adding time-linearly correlated deviations to 10% of original data for each exception scenario. Scenario 1 only adds deviations to the steel stress measurement points to be tested; scenario 3 adds equivalent deviations to all steel stress measurement points; and scenario 2 also adds deviations to all steel stress measurement points, but with different deviations for different points.

A robust anomaly detection model should accurately identify normal data (true positive, TP) and abnormal data (true negative, TN), thus have a lower frequency of missing alarm (false positive, FP) and false alarms (false negative, FN). This paper evaluates the anomaly recognition performance of each model by accuracy, measured as below:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (10)$$

TP, TN, FP, and FN in Equation (10) mean the count of anomaly recognition result for each test monitoring data. In this paper, three structural anomaly scenarios are used to test four models, including causal model, noncausal model, TSBRT, CDBRT, and the experiment results are shown in Table 5.

As can be seen in Table 5, the improved TSBRT and CDBRT models have a more robust anomaly recognition performance. This is demonstrated by the fact that the improved models not only has the better anomaly recognition ability than causal model in scenario 1 but also has the best anomaly recognition in scenarios 2 and 3 where overall structural anomalies are detected. Although the noncausal model has the best anomaly recognition accuracy in scenario 1, it is unable to effectively identify deviations of overall structural stress distribution, which is demonstrated in Figures 7 and 8.

From the anomaly discrimination intervals of the different models in Figures 7 and 8, it can be seen that the

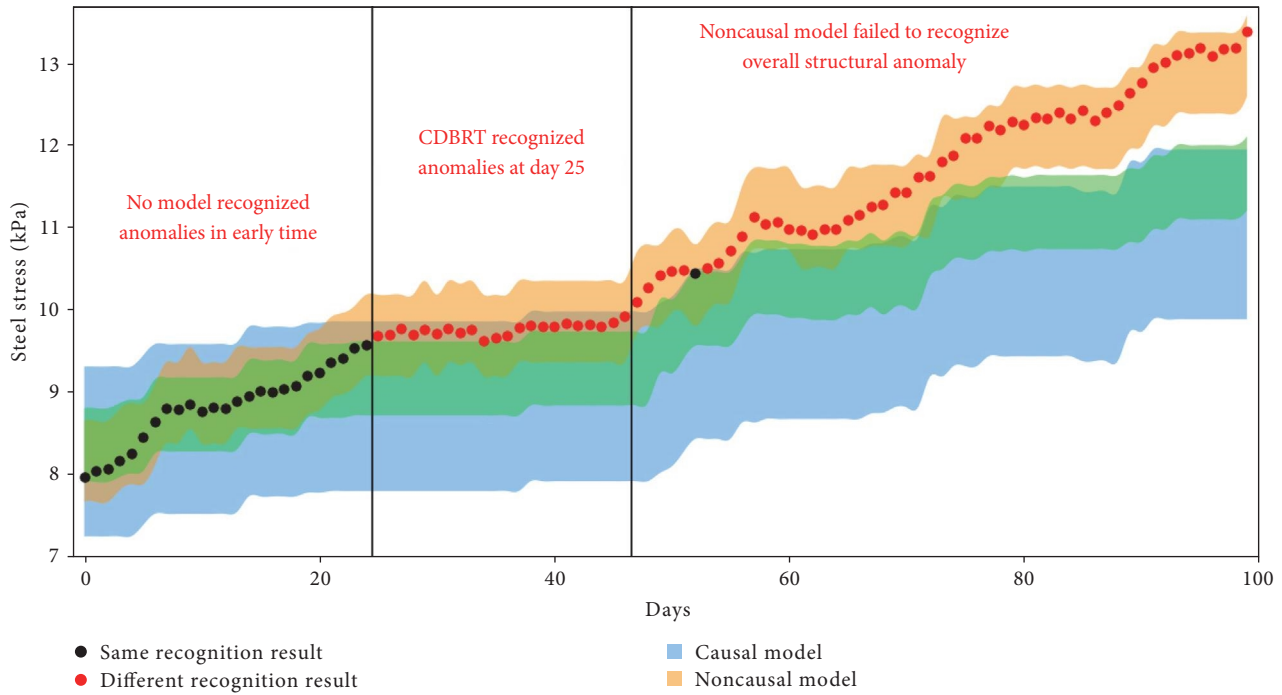


FIGURE 8: Anomaly detection result of causal model, noncausal model, and CDBRT in scenario 3.

noncausal model is unable to identify overall structural anomalies. It is because its estimation dependent on other steel stress measurement points too much. The improved TSBRT and CDBRT, however, are able to accurately identify causation from the correlation of adjacent effect measure points, thus have more reasonable anomaly recognition results. In addition, the improved model is able to recognize anomalies and alerts earlier relative to the causal model. It is because the improved model takes into account the effect of unobserved environmental quantities in the estimation and has a higher prediction accuracy, i.e., a narrower discriminatory interval for anomalies.

5. Conclusions

The use of monitoring data to recognize structural anomalies is a typical intelligent application of structural safety monitoring, which is of great significance to hydraulic engineering operational management. A large number of studies have paid attention to constructing regression models with advanced method to obtain a better in-bag prediction accuracy, like machine learning, deep learning, etc. However, few researchers focus on how to improve the anomaly recognition performance which is the ultimate goal of regression modeling. This paper proposed a novel anomaly detection method and revealed that integration of causal inference method could reduce the risks of false alarms and missing alarms significantly, especially when correlated response points are taken into consideration. In the process, two deconfounding machine learning models inspired by methods of handling endogeneity problems in economics, TSBRT and CDBRT, are proposed to restore the causal effect between adjacent response points. The validation was carried

out with the Shanmen River culvert monitoring data, and the experimental results showed that anomaly recognition methods proposed in this paper has higher average recognition accuracy in different structural anomaly scenarios than existing regression models, which has good application prospects.

The solutions and methods proposed in this paper still hold incomplete issues that should be further investigated. These issues to study in future research include: first, the causal graph of culvert monitoring variables proposed in this paper is derived from expert experience, and how to improve the causal graph to better guide the regression modeling process needs further exploration. Second, the anomaly data used in the case study are generated based on hypothetical structural anomaly scenarios, and research on combining with physical simulation methods such as finite element analysis should be conducted to obtain more realistic anomaly data to evaluate the model anomaly recognition performance more accurately. In addition, the causal inference model proposed in this paper is based on boosted regression trees, and future research will attempt to adapt more machine learning models and validate them on more types of hydraulic structures.

Nomenclature

- x, y, z, u, v : Variables (set) used in regression modeling
- $f(\cdot), g(\cdot), h(\cdot)$: Regression functions with a set of input variables
- $P(y|x)$: The conditional probability distribution of y on x , while $P(y|do(x))$ means interference distribution on x
- $\alpha, \beta, \lambda, \phi, \rho$: Scalar value of specific parameters

$\{(\cdot)\}$:	Dataset with a set of features
$ \cdot $:	Cardinality of a set
$F(\cdot), G(\cdot), H(\cdot)$:	Probability integral transformation function of variables
σ :	Standard error of a variable
$\Phi^{-1}(\cdot)$:	Inverse normal distribution of a variable
$\frac{\partial L(x)}{\partial x}$:	Derivative of loss function $L(x)$ with respect to a variable x
$\hat{x}, \hat{y}, \hat{w}, \hat{c}, \hat{\beta}$:	Estimated value of variables or parameters
$\tilde{x}, \tilde{y}, \tilde{w}$:	Variables generated by sampling from dataset or distributions.

Data Availability

The data that support the findings of this study are available from the corresponding author Xuemei Liu upon reasonable request. Requests may be sent at lizgrf@163.com.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Projects of Open Cooperation of Henan Academy of Sciences under grant number 220901008, Major Science and Technology Projects of the Ministry of Water Resources under grant number SKS-2022029, and Henan Water Conservancy Science and Technology Research Project Plan under grant number GG202259.

References

- [1] H.-N. Li, D.-S. Li, L. Ren, T.-H. Yi, Z.-G. Jia, and K.-P. Li, "Structural health monitoring of innovative civil engineering structures in Mainland China," *Structural Monitoring and Maintenance*, vol. 3, no. 1, pp. 1–32, 2016.
- [2] A. B. Noel, A. Abdaoui, T. Elfouly, M. H. Ahmed, A. Badawy, and M. S. Shehata, "Structural health monitoring using wireless sensor networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1403–1423, 2017.
- [3] A. Pyayt, I. Mokhov, A. Kozionov et al., "Artificial intelligence and finite element modelling for monitoring flood defence structures," in *2011 IEEE Workshop on Environmental Energy and Structural Monitoring Systems*, IEEE, 2011.
- [4] P. Bukenya, P. Moyo, H. Beushausen, and C. Oosthuizen, "Health monitoring of concrete dams: a literature review," *Journal of Civil Structural Health Monitoring*, vol. 4, no. 4, pp. 235–244, 2014.
- [5] W. Belcher, T. Camp, and V. V. Krzhizhanovskaya, "Detecting erosion events in earth dam and levee passive seismic data with clustering," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2015.
- [6] A. Sivasuriyan, D. S. Vijayan, R. Munusami, and P. Devarajan, "Health assessment of dams under various environmental conditions using structural health monitoring techniques: a state-of-art review," *Environmental Science and Pollution Research*, vol. 29, no. 57, pp. 86180–86191, 2022.
- [7] H. Yu, Z. Wu, T. Bao, and L. Zhang, "Multivariate analysis in dam monitoring data with PCA," *Science China Technological Sciences*, vol. 53, no. 4, pp. 1088–1097, 2010.
- [8] B. Chen, T. Hu, Z. Huang, and C. Fang, "A spatio-temporal clustering and diagnosis method for concrete arch dams using deformation monitoring data," *Structural Health Monitoring*, vol. 18, no. 5–6, pp. 1355–1371, 2019.
- [9] C. Ma, T. Zhao, G. Li, A. Zhang, and L. Cheng, "Intelligent anomaly identification of uplift pressure monitoring data and structural diagnosis of concrete dam," *Applied Sciences*, vol. 12, no. 2, Article ID 612, 2022.
- [10] W. D. Fisher, T. K. Camp, and V. V. Krzhizhanovskaya, "Anomaly detection in earth dam and levee passive seismic data using support vector machines and automatic feature selection," *Journal of Computational Science*, vol. 20, pp. 143–153, 2017.
- [11] X. Li, Y. Li, X. Lu, Y. Wang, H. Zhang, and P. Zhang, "An online anomaly recognition and early warning model for dam safety monitoring data," *Structural Health Monitoring*, vol. 19, no. 3, pp. 796–809, 2020.
- [12] F. Salazar, A. Conde, J. Irazábal, and D. J. Vicente, "Anomaly detection in dam behaviour with machine learning classification models," *Water*, vol. 13, no. 17, Article ID 2387, 2021.
- [13] J. Hu, F. Ma, and S. Wu, "Anomaly identification of foundation uplift pressures of gravity dams based on DTW and LOF," *Structural Control and Health Monitoring*, vol. 25, no. 5, Article ID e2153, 2018.
- [14] R. Hellgren, R. Malm, and A. Ansell, "Performance of data-based models for early detection of damage in concrete dams," *Structure and Infrastructure Engineering*, vol. 17, no. 2, pp. 275–289, 2021.
- [15] F. Salazar, R. Morán, M.Á. Toledo, and E. Oñate, "Data-based models for the prediction of dam behaviour: a review and some methodological considerations," *Archives of Computational Methods in Engineering*, vol. 24, no. 1, pp. 1–21, 2017.
- [16] M. Flah, I. Nunez, W. Ben Chaabene, and M. L. Nehdi, "Machine learning algorithms in civil structural health monitoring: a systematic review," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2621–2643, 2021.
- [17] G. Prakash, R. Dugalam, M. Barbosh, and A. Sadhu, "Recent advancement of concrete dam health monitoring technology: a systematic literature review," *Structures*, vol. 44, pp. 766–784, 2022.
- [18] G. Prakash, A. Sadhu, S. Narasimhan, and J.-M. Brehe, "Initial service life data towards structural health monitoring of a concrete arch dam," *Structural Control and Health Monitoring*, vol. 25, no. 1, Article ID e2036, 2018.
- [19] S. Wang, Y. Xu, C. Gu, Q. Xia, and K. Hu, "Two spatial association—considered mathematical models for diagnosing the long-term balanced relationship and short-term fluctuation of the deformation behaviour of high concrete arch dams," *Structural Health Monitoring*, vol. 19, no. 5, pp. 1421–1439, 2020.
- [20] F. Salazar, M.Á. Toledo, J. M. González, and E. Oñate, "Early detection of anomalies in dam performance: a methodology based on boosted regression trees," *Structural Control and Health Monitoring*, vol. 24, no. 11, Article ID e2012, 2017.
- [21] K. Sainani, "The importance of accounting for correlated observations," *PM&R*, vol. 2, no. 9, pp. 858–861, 2010.
- [22] J. Mata, "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models," *Engineering Structures*, vol. 33, no. 3, pp. 903–910, 2011.

- [23] H. Su, Z. Chen, and Z. Wen, "Performance improvement method of support vector machine-based model monitoring dam safety," *Structural Control and Health Monitoring*, vol. 23, no. 2, pp. 252–266, 2016.
- [24] B. Dai, C. Gu, E. Zhao, and X. Qin, "Statistical model optimized random forest regression model for concrete dam deformation monitoring," *Structural Control and Health Monitoring*, vol. 25, no. 6, Article ID e2170, 2018.
- [25] F. Kang, J. Liu, J. Li, and S. Li, "Concrete dam deformation prediction model for health monitoring based on extreme learning machine," *Structural Control and Health Monitoring*, vol. 24, no. 10, Article ID e2153, 2017.
- [26] J. Mata, F. Salazar, J. Barateiro, and A. Antunes, "Validation of machine learning models for structural dam behaviour interpretation and prediction," *Water*, vol. 13, no. 19, Article ID 2717, 2021.
- [27] F. Salazar, M.Á. Toledo, E. Oñate, and B. Suárez, "Interpretation of dam deformation and leakage with boosted regression trees," *Engineering Structures*, vol. 119, pp. 230–251, 2016.
- [28] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box: automated decisions and the GDPR," *Harvard Journal of Law & Technology*, vol. 31, Article ID 841, 2017.
- [29] A. El Bilali, M. Moukhliiss, A. Taleb et al., "Predicting daily pore water pressure in embankment dam: empowering machine learning-based modeling," *Environmental Science and Pollution Research*, vol. 29, no. 31, pp. 47382–47398, 2022.
- [30] S. Chen, C. Gu, C. Lin, and M. A. Hariri-Ardebili, "Prediction of arch dam deformation via correlated multi-target stacking," *Applied Mathematical Modelling*, vol. 91, pp. 1175–1193, 2021.
- [31] J. Hartford, G. Lewis, K. Leyton-Brown, and M. Taddy, "Deep IV: a flexible approach for counterfactual prediction," in *International Conference on Machine Learning*, PMLR, 2017.
- [32] S. Park and S. Gupta, "Handling endogenous regressors by joint estimation using copulas," *Marketing Science*, vol. 31, no. 4, pp. 567–586, 2012.
- [33] E. Štrumbelj and I. Kononenko, "Explaining prediction models and individual predictions with feature contributions," *Knowledge and Information Systems*, vol. 41, no. 3, pp. 647–665, 2014.
- [34] S. M. Lundberg, G. Erion, H. Chen et al., "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56–67, 2020.
- [35] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, 2017.