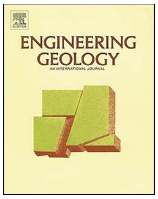
Engineering Geology 234 (2018) 132–137

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo



Analysis of the lag effect of embankment dam seepage based on delayed mutual information

Yuqun Shi[a,b], Cheng Zhao[a,c], Zhengquan Peng[a], Haiyun Yang[d], Jinping He[a,b,*]

[a] *School of Water Resources and Hydropower, Wuhan University, Wuhan, Hubei 430072, China*
[b] *State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, Hubei 430072, China*
[c] *Large Dam Safety Supervision Center, National Energy Administration, Hangzhou, Zhejiang 311122, China*
[d] *Electric Power Research Institute of State Grid Fujian Electric Power Co. Ltd., Fuzhou, Fujian 350007, China*

ARTICLE INFO

Keywords:
Embankment dam
Seepage
Lag effect
Delayed mutual information
Directed information transfer index

ABSTRACT

Environmental variables, such as the upstream level, significantly affect the embankment dam seepage lag time. The lag effect should, therefore, be given adequate consideration when determining the saturation line and establishing a mathematical model for seepage. At present, however, the lag time is mainly estimated qualitatively. A method for analyzing the seepage lag effect is presented herein based on the mutual information principle of information science theory, wherein the upstream water level is considered an information source and water level of the seepage pressure is considered an information function point. The method explores application of the directed information transfer index (*DITI*) of mutual information in order to establish an information transfer model between seepage levels of measurement points and upstream levels of the embankment dam. Furthermore, an extreme value of the *DITI* function is adopted as the judging criterion to determine the lag time of the influence of the upstream level on corresponding seepage levels. Results obtained through analysis of an engineering case study, demonstrate the effectiveness and feasibility of the proposed method.

1. Introduction

Following the impoundment of an embankment dam, the seepage field of the dam body is seen to gradually stabilize. At this point, seepage within the dam body is primarily influenced by environmental variables, such as the upstream level, which demonstrate a certain lag effect. As one of the main problems encountered in data analysis of embankment dam monitoring, lag-effect analysis of dam-body seepage plays an important and central role in health diagnosis and safety monitoring of embankment dams. For example, the lag time recorded at each measurement point must be known when using monitoring data to determine the saturation line, thereby determining the relationship between seepage pressure and upstream level. Lag-effect analysis is also required in selecting preset factors when establishing seepage models.

The causes and mechanism of the seepage lag effect of embankment dams are complicated. Gu et al. (2005) summarized the mechanism as follows. Variations in the upstream level alter the dam seepage field from one stable condition to another, and the transition process occurs over a certain time period. This transition is the main cause of the seepage lag effect. Delayed transmission of water pressure in the dam body is another factor contributing to the lag effect. Further, response times of seepage monitoring sensors, such as piezometric tubes and osmometers, with regards to variations occurring in the seepage state could also contribute to the lag effect.

At present, only a few specialized studies are being conducted with focus on the seepage lag effect in embankment dams. The basic method used in these studies involves comparison of the process line of the measured water pressure with those of the corresponding upstream water level. This technique serves to qualitatively determine the lag time of the seepage in terms of the appearance time of “peaks” and “valleys” in the process line. Gu et al. (2005) utilized the mean value of the upstream water level over several time periods as the precession factor for the hydraulic-pressure component in their seepage monitoring statistic model, thereby considering the hysteresis effect of seepage. In addition, several scholars have investigated seepage characteristics of embankment dams while also inadvertently studying aging characteristics of seepage. Ozer and Bromwell (2012) used the finite element method and measured seepage data to simulate seepage and aging characteristics of embankment dams. Simeoni (2012) investigated the relationship between the lag time and pore-water pressure using the piezometer experimental method. Wang et al. (2013) studied the influence of the crack self-healing phenomenon on the

---

[*] Corresponding author at: School of Water Resources and Hydropower, Wuhan University, Wuhan, Hubei 430072, China.
E-mail address: whuhjp@whu.edu.cn (J. He).

https://doi.org/10.1016/j.enggeo.2018.01.009
Received 23 January 2017; Received in revised form 30 December 2017; Accepted 9 January 2018

permeability of dams via 12 simulation tests. Lee et al. (2007) explored the tracer experiment method combined with specific engineering to study the concentrated seepage path of embankment dams. However, a common drawback of the studies mentioned above concerning determination of the seepage lag time is the exclusive utilization of qualitative estimations without quantitative measurements. It is, therefore, necessary to apply new ideas and methods to accurately quantify the seepage lag effect.

Seepage monitoring data reflect changes in the dam seepage field. In information science theory, any substance can be considered as an information source. This information source constantly transfers typical data to the surrounding medium through the information field, thereby altering the characteristics of space and time to generate an information effect. Based on this theory, the upstream water level can be considered as an information source while variations in the upstream water level could be considered typical information conveyed by the information source. Seepage information is transmitted to each monitoring point (information function point) through the seepage field in the embankment dam body, which in turn, alters the water level of the seepage pressure. The concept of mutual information in the information entropy theory can be used to quantitatively describe shared data that the source transfers to the function point (Zhang and Liu, 2000). The paper introduces this concept of mutual information into the study of the seepage lag effect, and proposes a new method for determining the lag time between the seepage pressure in embankment dams and upstream water level.

2. Delayed mutual information

2.1. Mutual information

The domain of the continuous random variables X and Y was set as S; their marginal distributions were represented by $f_X(x)$ and $f_Y(y)$, respectively, and their joint distribution were given by $f_{X,\ Y}(x,y)$. The information entropy $H(X)$ and $H(Y)$ of the continuous random variables X and Y could, thus, be defined as (Resconi et al., 2013) follows.

$$H(X) = -\int_S f_X(x) log f_X(x),$$
$$H(Y) = -\int_S f_Y(y) log f_Y(y) \tag{1}$$

Further, the mutual information of the two-dimensional joint distribution of X and Y could be defined as (Khademi et al., 2017) follows.

$$I(X;Y) = \iint_S f_{X,Y}(x,y)\, log \frac{f_{X,Y}(x,y)}{f_X(x) f_Y(y)} dxdy \tag{2}$$

Mutual information represents the amount of related information that is jointly owned by two or more attributes. A considerably large amount of mutual information indicates a close relationship between the concerned attributes or variables.

2.2. Directed information transfer index (DITI)

Movement is an important feature of information. Once a certain information is produced by the source, a flow is set up through its transmission, and any point in time and space of the structure (information function point) is affected, thereby resulting in different degrees of information benefits. *DITI* combines the information transfer theory with information entropy. Information entropy represents the quantity of directional transmission of specific information from the source in order to determine the degree of mutual influence between the source and the function point of information.

Assuming that X and Y refer to the information source and information function point in the information field, respectively, *DITI* $(X;Y)$ of the information source X to information function point Y can be represented as follows (Zhang and Liu, 2000).

$$DITI(X;Y) = \frac{I(X;Y)}{H(Y)} \tag{3}$$

DITI exhibits two important characteristics—(1) it measures the information transfer capability between the two information (X and Y); (2) it describes the degree of coupling between X and Y.

2.3. Delayed mutual information

The original data sequence was set to $X = \{x_1, x_2, \cdots, x_n\}$ and viewed as a source of information. The length of this data sequence was n. With an X delay considered as t, the reconstructed data sequence Y could be obtained. If Y be set as the information function point, then the length of the data sequence becomes $n - t$. Thus, $Y = \{y_1, y_2, \cdots, y_{n-t}\} = \{x_{1+t}, x_{2+t}, \cdots, x_n\}$. The mutual information between X and Y is, therefore, defined as delayed mutual information (Albers and Hripcsak, 2012).

The delay time t is theoretically desirable for either value, and is usually selected based on the characteristics of research questions. Since t is an unknown quantity, a series of reconstructed data sequences $\{Y_i, i = 1 \sim m\}$ could be obtained by setting t different delay times $\text{t} = \{t_1, \cdots t_i, \cdots, t_m\}$. That is,

$$Y = \{y_{i1}, y_{i2}, \cdots, y_{i(n-t_i)}\} = \{x_{1+t_i}, x_{2+t_i}, \cdots, x_n\} \tag{4}$$

where Y_i is the reconstructed data sequence when the delay time $t = t_i$.

Takens' theorem (Takens, 1981) indicates that when the delay time t covers the actual delay between X and Y, a reconstructed data sequence $Y_i(t_i = \tau)$ equivalent to the original data sequence X in the topological sense always exists. The mutual information between the original sequence X and reconstructed sequences $Y_i(\text{i} = 1, 2, \cdots, \text{m})$ is then calculated and analyzed. Subsequently, one reconstructed sequence $Y_i(t_i = \tau)$ that is equivalent to the original data sequence X in the topological sense could be obtained using Eq. (4) in order to determine the actual delay time, τ.

3. Determining seepage flow lag time based on delayed mutual information

3.1. Basic principles

Let us assume the data sequence of dam environment variables (such as upstream water level) to be $X = \{x_1, x_2, \cdots, x_n\}$, and the data sequence of dam effect variables (such as the water level of the seepage pressure at the measurement point) are correspondingly represented by $Y = \{y_1, y_2, \cdots, y_n\}$. Given that the environmental variables influence the monitoring effect variables of the dam, the environment variable sequence X is considered as the information source while the effect variable sequence Y is regarded as the information function point.

For seepage monitoring in embankment dams, several monitoring cross sections are usually selected, wherein several measurement points are arranged. A piezometric tube or an osmometer is buried at each measurement point to facilitate regular monitoring of the water level of seepage pressure, and the saturation line of each monitoring cross section is obtained using the monitored data. The water level of the seepage pressure at measurement points in the dam body is primarily influenced by the upstream water level, demonstrating varying degrees of hysteresis. The concept of delayed mutual information in Eq. (4) is extended to two variables with correlations X and Y. The dam monitoring effect variable $Y = \{y_1, y_2, \cdots, y_n\}$ is considered as the corresponding data sequence of the environmental variable sequence $X = \{x_1, x_2, \cdots, x_n\}$ after delay time t. The actual delay time τ of each measurement point is unknown; thus, a suitable method must utilize time to determine the actual delay time τ from amongst a set of assumed delay times $t_i(i = 1, 2, \cdots, m)$. Consequently, the approach proposed in this paper for determining the embankment dam seepage lag time is based on information entropy and delayed mutual information. Further, the