

Stanisław Lach¹

Detection of Outlier Observations in Piezometric Measurements: A Case Study in the Southern Region of Poland

Abstract: One of the main modes of monitoring the geotechnical conditions of earth dams is piezometric measurement, which measures water levels in an open piezometer or water pressure in a closed piezometer. During piezometric measurements, various types of factors can cause disturbances in these measurements that take the form of systematic, accidental, or obvious mistakes. Before measurements from open or closed piezometers are analyzed, outliers due to coarse errors should be detected and rejected. Such observations may significantly influence the result of the analysis and cause erroneous assessment and interpretation of the phenomenon studied. To do this, statistical tests must be applied so that the doubtful measurement can be accepted or rejected at the assumed significance level. This paper uses five statistical tests for identifying and rejecting outliers: the Q-Dixon test, the Grubbs test, as well as the Hampel test, the Iglewicz and Hoaglin test, and the Rosner test. The aim of this article is to try to identify the most suitable test for periodic piezometric measurements. The scope of the study includes the analysis of piezometric measurements for the Czaniec Dam for the multi-year period 2017–2020.

Keywords: dam, outlier observations, safety, seepage, piezometer

Received: 13 July 2021; accepted: 29 October 2021

© 2022 Author. This is an open access publication, which can be used, distributed and reproduced in any medium according to the Creative Commons CC-BY 4.0 License.

¹ AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Department of Environmental Management and Protection, Krakow, Poland, email: slach@agh.edu.pl,  <https://orcid.org/0000-0002-6378-7838>

1. Introduction

The construction of dams and the existence of the reservoirs they create are an essential part of society. Reservoirs make it possible to store water when it is in excess and to use it when it is scarce. Dams, on the other hand, make it possible to prevent floods by modifying the course of a flood wave and, to a large extent, lowering its peak. Hydroelectric dams operate under varying meteorological and hydrological conditions. Such facilities are constantly exposed to intense precipitation, floods, as well as landslides, lightning, and ice phenomena. The vulnerability of hydroelectric dams to damage or disaster increases with the time of operation. This concerns about 30% of the Polish damming facilities that have been operating for more than 50 years. Such a long period of operation, according to the assessment of the International Commission on Large Dams (ICOLD), results in a higher number of damage events and an increased probability of failure [1].

Embankment dams are the most common type of dam built today. Embankment dams are constructed from natural earthen materials, usually local soil, and rock [2]. In embankment dams, water seeps through the soil layers of the dam, and any change in this behavior may be an indication of emerging problems. Special attention should be paid to the safety of the dams since the number of disasters and major dam failures is increasing successively [3]. In the case of dams on earth and embankments, the most common cause of disasters was overtopping (31% of cases as the main cause and 18% as an additional cause), followed by internal erosion of the dam body (15% of cases as the main cause and 13% as an additional cause), as well as foundation defects, including settlement and slope instability (12% of cases as the main cause and 5% as an additional cause) [4–9]. According to ICOLD, in 15% of all cases, suffusion is the main cause of earth dam failure, while in 13% of cases this process is an additional cause [1].

The primary form of dam monitoring is by means of piezometric measurements [10]. These measurements make it possible to measure the level of the water table in an open piezometer or to measure the water pressure in a closed piezometer [11]. These measurements allow monitoring of the intensity of the seepage phenomenon through the dam [12, 13]. An increasing (or decreasing) trend in piezometers can indicate the movement of fine particles in the body or subsoil of the structure, which over time can lead to a local exceedance of the permissible seepage gradients, leading to a situation that poses a threat to the safety of dam operation [14–16]. With systematic measurements, a possible catastrophe can be effectively prevented by activating warning or alarm systems, as well as planning the upgrade of the facility far in advance [17].

Although the phenomenon of moderate intensity seepage is not in itself undesirable, the particularly intensive seepage process can cause changes in soil structure and be the reason for the loss of object stability [18–23]. The phenomenon of seepage can be recognized using classical measuring equipment (mainly piezometers and drainage flow measurements) and a number of research procedures ranging from

indicator methods through acoustic methods, resistivity-based methods, numerical modeling (Finite Element Methods) to methods using relations between seepage and heat flow in soil medium [24–39].

During the performance of measurements in piezometers, some factors may appear that cause disturbances in the obtained results. These can be systematic, accidental, or an obvious mistake. Before analyzing such data, coarse errors that may significantly affect the result and cause a false assessment or interpretation of the phenomenon studied should be detected and removed [40]. It is also worth mentioning the problem of the difference between an outlier and the detection of unusual geotechnical behavior. Geotechnical measurements are inevitably subject to various uncertainties [41–45]. The probability distributions of specific geotechnical parameters depend significantly on the quality of the measurements obtained, which are affected by measurement errors, changing measurement conditions (e.g. severe weather, icing, etc.) or other unknown environmental disturbances [46–49]. Estimated statistics of geotechnical parameters may be subject to high statistical uncertainty, and therefore it would be advisable not only to detect and remove outliers, but also to try to detect components that influence the formation of unusual geotechnical behavior, for which additional in situ and/or laboratory tests may be necessary [50, 51]. For this purpose, for example, a probabilistic outlier detection method can be applied to sparse multivariate data obtained during geotechnical investigations [52].

Diagnostics of the structure and its monitoring allows to know the technical condition of the hydrotechnical object, it is especially useful in the assessment of water dam structures [53]. It not only means the adoption of a safety coefficient that guarantees the integrity and stability of the structure, but it also becomes an essential component of the risk of catastrophe caused by dam failure, where the risk is understood here as the product of the probability of dam failure and human and material loss caused downstream of the structure due to its sudden failure [54]. Hydrotechnical structures have large volumes and are exposed to continuous contact with water, usually surface water. The function of the water dam is of particular importance, since seepage has a strong influence on the object and the ground [55]. In addition to seepage, the condition of a dam is affected by contact with flowing water, which can result in erosion or siltation [56].

The principles on which the monitoring of hydraulic structures is based are contained in the following documents [57–59]: the Construction Law (Act of 7 July 1994, Journal of Laws 2006, No. 156, item 1118, as amended) and the Water Law (Act of 20 July 2017, Journal of Laws 2017, item 1566, as amended) and regulations, including technical and construction regulations (e.g. Regulation of the Minister of Environment of 20 April 2007 on technical conditions to be met by hydrotechnical structures and their location (Journal of Laws 2007, no. 86, item 579), as well as various industry guidelines and instructions. For each hydrotechnical object, the organizational and legal supervision is selected individually based on the above-mentioned documents.

This paper uses five statistical tests for identifying and rejecting outliers: the Q-Dixon test, the Grubbs test, as well as the Hampel test, the Iglewicz and Hoaglin test, and the Rosner test. The aim of this paper is to propose the most suitable test for determining outlier observations for periodic piezometric measurements. The scope of the study includes the analysis of piezometric measurements for Czaniec Dam for the multiyear period 2017–2020.

Section 2 shows the scope of the work, which includes the analysis of piezometric measurements for the Czaniec Dam over the multi-year period 2017–2020, and describes all statistical tests used to detect outlier observations. Section 3 briefly describes the object studied, Czaniec Dam. The area of the reservoir and its basic functions are defined. For the dam studied, its location and essential elements such as height of the dam, length, and width of the crest as well as downstream and upstream slopes and their insurance were presented. Section 4 tabulates the number of outlier results for all the statistical tests performed and discusses them. Section 5 describes the main conclusions of the study.

2. Materials and Methods

For the Czaniec Dam located in the Silesian Province, an analysis of changes in the measurements of water table, changes in open piezometers within the front dam and side dams (64 piezometers on total), in open piezometers at the dike (8 piezometers) and in wells (4 piezometers) covering the period from 17.01.2017 to 23.12.2020. Measurements were almost always made twice a month, which resulted in 95 measurement results for a single piezometer during the analysis period. Piezometric data were provided by the Regional Water Management Board in Krakow.

In this study, five statistical tests were used to identify and reject outliers: the Q-Dixon test, the Grubbs test, as well as the Hampel test, the Iglewicz and Hoaglin test, and the Rosner test.

The Q-Dixon test is used to check whether a particular data set has a result that is subject to coarse error. The prerequisite for the application of the test is its numerical size. Two variants of the test were used in this study: for the single outliers – the N9 test and for pairs of outliers – the N13 test. The Q-Dixon test can be used to reject only a single outlier or a pair of outliers from a data set [60]. Note that before performing the test described above, the set of piezometric measurements should be arranged in a nondecreasing sequence. Table 1 shows the pairs of hypotheses tested in each variant of the test.

To reject the hypothesis of the absence of an outlier (variant N9) or a pair of outliers (variant N13), the value of the Q statistic is compared with the value read from the table of critical values Q_n of the Q-Dixon test in the variant N9 or N13 at the significance level α [61].

Table 1. Hypotheses tested in variants N9 and N13 using the Q-Dixon test

Test N9	
Testing the upper outlier	Testing the lower outlier
Null hypothesis (H_0): $x_{(n)}$ is not an outlier. Alternative hypothesis (H_1): $x_{(n)}$ is an outlier. Test statistic: $Q_n = \frac{x_n - x_{n-1}}{x_n - x_2}$	Null hypothesis (H_0): $x_{(1)}$ is not an outlier. Alternative hypothesis (H_1): $x_{(1)}$ is an outlier. Test statistic: $Q_1 = \frac{x_2 - x_1}{x_{n-1} - x_1}$
Test N13	
Testing the pair of largest outliers	Testing the pair of smallest outliers
Null hypothesis (H_0): pair ($x_{(n)}, x_{(n-1)}$) is not an outlier pair. Alternative hypothesis (H_1): pair ($x_{(n)}, x_{(n-1)}$) is a pair of outliers. Test statistic: $Q_n = \frac{x_n - x_{n-2}}{x_n - x_3}$	Null hypothesis (H_0): pair ($x_{(1)}, x_{(2)}$) is not an outlier pair. Alternative hypothesis (H_1): pair ($x_{(1)}, x_{(2)}$) is a pair of outliers. Test statistic: $Q_1 = \frac{x_3 - x_1}{x_{n-2} - x_1}$

Source: [62]

Before performing the Grubbs test, the set of experimental results, as in the case of the Q-Dixon test, should be ranked in a non-decreasing sequence. It is clear that the coarse error may be the largest (x_{max}) or smallest (x_{min}) result value in the sample under analysis. This test at one time, like the Q-Dixon test, only gives the possibility of detecting one outlier, so it should be repeated until no further outliers are observed in the data set [63, 64]. The value of the G_p test statistic for the Grubbs test can be calculated using the formula:

$$G_p = \frac{\max_{i=1, \dots, n} |x_i - \bar{x}|}{s} \tag{1}$$

where:

- \bar{x} – mean value of the series of measurements tested,
- s – standard deviation.

The critical value of the Grubbs test statistic for the critical value of the Grubbs test statistic for the assumed significance level α can be calculated from the following formula [61]:

$$G_{kr} = \frac{(n-1)}{\sqrt{n}} \cdot \sqrt{\frac{t_{(\alpha/(2n), n-2)}^2}{n-2 + t_{(\alpha/(2n), n-2)}^2}} \quad (2)$$

Thus, it can be seen that the value of G_{kr} is calculated from the critical value of the Student's t distribution for a significance level of $\alpha/(2n)$ and a number of degrees of freedom equal to $n - 2$, where n is the number of piezometric measurements in the series.

Another test used is the Hampel test. An advantage of this test is its simplicity, as there is no limitation on the size of the set tested. The Hampel test is used to detect in the analyzed data set results that significantly deviate from the average values. This test also tends to generate a significant number of errors [60]. Inference about the nature of the observation under study is based on the evaluation of the obtained results of the analysis based on specific formulas. By performing the Hampel test, it is necessary to calculate the median value M_e and then the deviations r_i from the median and the absolute values $|r_i|$ and median deviations $Me_{|r_i|}$ in the analyzed set. In the case of the Hampel test, unusual observations are those that satisfy the condition $|r_i| \geq 4.5Me_{|r_i|}$ [64]:

For another statistical test of Iglewicz and Hoaglin, calculate the M_i value according to the formula:

$$M_i = \frac{0.6745(x_i - \tilde{x})}{MAD} \quad (3)$$

where:

\tilde{x} – median of the piezometric data set,

MAD – median absolute deviation, calculated as:

$$MAD = \text{median}(|x_i - \tilde{x}|) \quad (4)$$

where $|x|$ – absolute value of x .

The authors of the test recommend that M_i with a nonsingular value greater than 3.5 be considered an outlier.

The last static test used is the Rosner test [65]. The methodology to calculate outliers in this test is to run a series of test statistics and remove the measurement that is farthest from the mean and recalculate the test statistic according to the following equation [61]:

$$R_{i+1} = \frac{|x^{(i)} - \bar{x}^{(i)}|}{s^{(i)}} \quad (5)$$

where:

$$\bar{x}^{(i)} = \frac{1}{(n-i)} \sum_{j=1}^{n-i} x_j \quad (6)$$

$$s^{(i)} = \sqrt{\frac{1}{(n-i)} \sum_{j=1}^{n-i} (x_j - \bar{x}^{(i)})^2} \quad (7)$$

and $x^{(i)}$ – the most distant value from the measurement from $\bar{x}^{(i)}$.

3. Case Study

Czaniecki Reservoir is a reservoir located in the municipality of Porąbka, Bielsko-Biała county, Silesian voivodeship (Fig. 1). Figure 2 presents a situation plan for the Czaniec Dam. The water reservoir is located on 28.8 km of the Soła River. Its area is 43 ha. The main task of the Czaniecki reservoir is to equalize the daily flow of the Soła River, as well as to enable water intake for users downstream of the reservoir.



Fig. 1. Czaniecki Reservoir and Czaniec Dam

Source: <https://haskoningdhv.pl/wp-content/uploads/2020/09/Czaniec-2-1.jpg> [access: 22.10.2021]

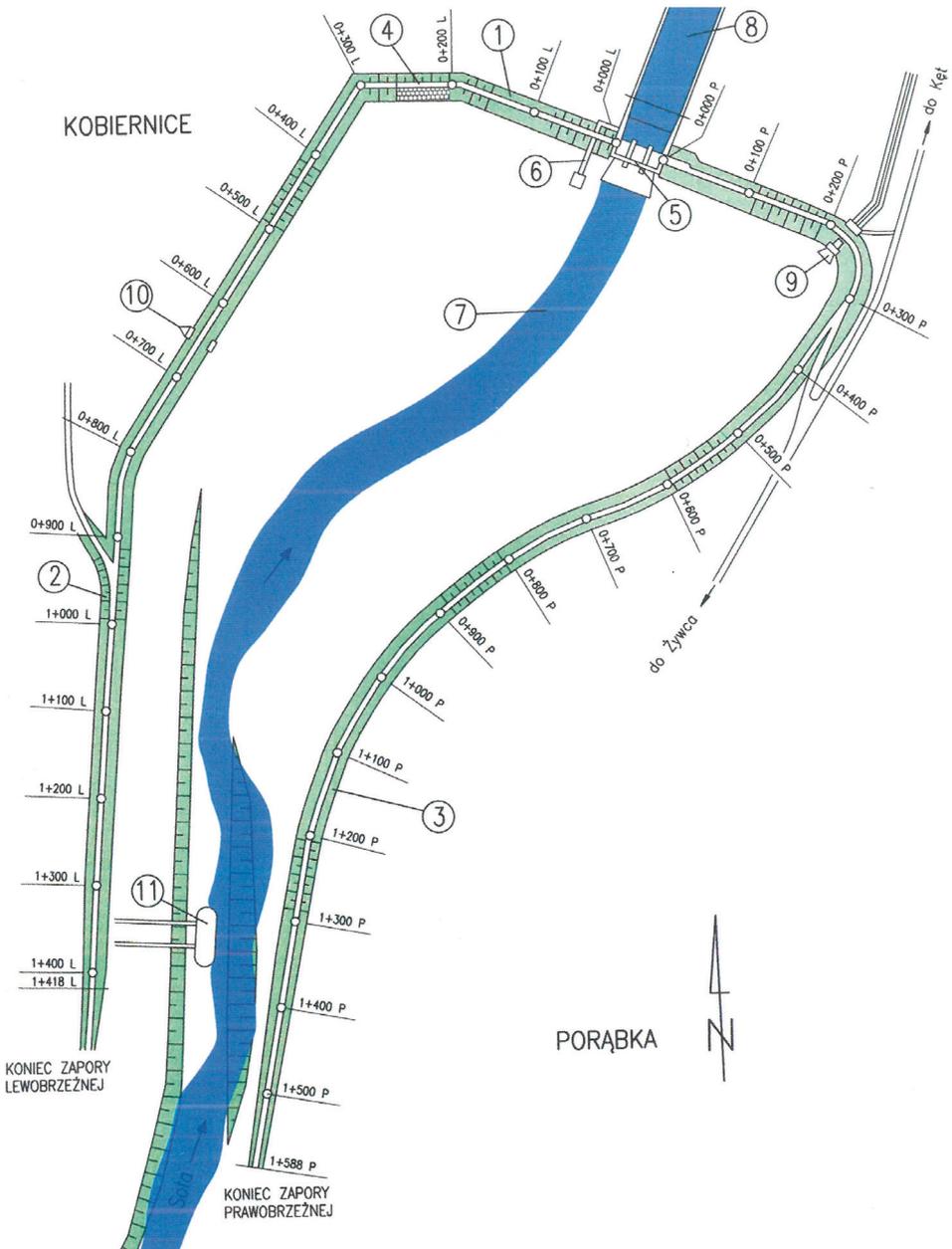


Fig. 2. Plan of situation of the Czaniec Dam:

1 – main dam, 2 – left-bank dam, 3 – right-bank dam, 4 – emergency spillway, 5 – weir, 6 – fish ladder, 7 – the Soła River, 8 – lower weir station, 9 – water intake for the Kęcka Water Company, 10 – water intake for fishponds in Kobiernice, 11 – Upper Silesian Industrial District water intake

Source: [66]

The Czaniec compensating reservoir is a dual-purpose reservoir because it lines up with the retention reservoirs (Tresna and Porąbka) and peak power plants that are located next to them. Optimal efficiency of power plants during the highest daily energy demand requires full power operation. It is ensured by the Czaniec compensating reservoir, whose usable volume is equal to the volume of water used during peak demand. The second task that the reservoir must perform is to evenly discharge the accumulated water into the river, so that because of the maximum operation of the power plants in the watercourse, there are no outflow waves that have the character of flood waves. The water levels in the equalization reservoir are subject to change, depending on the operation of the power plant, and the highest levels are recorded at the end of operation during peak periods.

The main dam is divided by a dike into two parts and its extensions are side dams. The right part of the main dam is 300 m long, while the left part, due to the emergency passage, consists of two sections with a total length of 248 m. The dam crest is 7 m wide and is located at the ordinate of 299.50 m above sea level. The slope of the downstream slope is 1:2.5, while that of the upstream slope is 1:2. In the right section, 15 m long from the abutment of the width of the dam structure, the crest width was extended from 7 m to 12 m, which allows easy maneuvering of vehicles that deliver equipment to the weir. Both the dam crest and the downstream slope are exposed to destructive weathering. In turn, the upstream slope and sometimes the lower area of the downstream slope are exposed to wave action. These surfaces require appropriate reinforcement [8]. The upstream slope is protected with a screen made of reinforced concrete slabs along its entire length. The bottom edge of the reinforced concrete slabs on the upstream slope is fixed in the apron made of clay and located at the ordinate of 294 m above sea level. Both the clay apron and the gravel cover are 0.5 m thick. The width of the apron is constant and equals 20 m except for the section of the Soła old riverbed where it increases to 25 m.

The drainage of the dam is made in the lower part of the vent layer in the form of a trapezoidal prism of stone and gravel. Drainage is an element that enables the intake and discharge of filtration and groundwater from the protection zone, reducing the zone's waterlogging and load caused by water pressure or filtration. The drain is equipped with a reverse filter, which prevents the leaking water from washing out the soil particles. The material used for drainage is characterized by high strength and resistance to frost. The dike was made in the dike, the crown of which has a width of 1 m and is located at the ordinate of 299.00 m above sea level. The dike formed in this way will easily be washed out if the water level exceeds 299.00 m above sea level in the reservoir. To prevent this, the bench and spillway scarp were reinforced with 15 cm thick cobbles. The overflow drainage slope is covered with cobbles, on which a clay-covered screen was laid. The layers made in this way seal the whole body against the overflow. The downstream slope is made in the same way as the upstream slope of the main dam and is grassed over with a layer of humus. The overtopping body has a fixed sill to reduce the level of scour and is made

of a steel wall with a reinforced concrete ring, a wooden palisade, and stone insulation. The river abutments are angular reinforced concrete retaining walls and 6.5 m high. The 15 m wide overbank floodplain towards the reservoir has been insured with concrete slabs [66].

The Czaniec Dam, as well as other facilities connected with it, is equipped with control, and measuring devices used to check the compliance of construction works with the design documentation. The equipment is distributed in a network and the type of devices, number, and place of mounting of measurement points must be selected in such a way that it is possible to trace the intensity of phenomena, probable safety hazards, and durability of the structure.

4. Results and Discussion

The results obtained are summarized in Table 2. Furthermore, Figure 3 shows an example of changes in the graph of water level changes in the P10B piezometer (with a clearly visible outlier at the Czaniec Dam before identification and removal of outliers, while Figure 4 shows the graph of changes in water level in the same piezometer at the Czaniec Dam after identification and removal of outliers.

Table 2. Number of outliers for the statistical tests for Czaniec Dam in the period from 17.01.2017 to 23.12.2020

Section no.	Designation	Number of outlier observations detected for statistical tests:					
		Q-Dixon		Grubbs	Hampel	Rosner	Iglewicz and Hoaglin
		N9	N13				
Water reservoir							
I	P1	0	2	0	1	0	1
	P2	0	2	2	14	2	10
	P3	1	0	0	2	0	2
Ia	P1A	0	0	0	2	0	2
	P3A	1	2	1	1	1	1
Ib	P1B	1	2	1	3	1	3
	S1	0	0	0	1	0	0
II	P4	0	0	0	0	0	0
	P5	0	2	0	1	0	1
	P6	0	0	0	2	0	1
III	P7	0	0	3	3	3	3
	P8	1	2	1	1	1	1
	P9	0	0	0	6	0	2

Table 2. cont.

IV	P10A	0	0	0	20	0	20
	P11A	1	2	1	3	1	1
	P12A	0	2	0	2	0	1
	P10B	1	2	1	1	1	1
V	P13B	0	0	0	0	0	0
	P14	1	0	0	0	0	0
	P15	0	0	0	1	0	0
Va	P51	0	0	0	1	0	0
	PB	0	0	0	7	0	5
VI	P16A	2	4	1	2	1	1
	P17	0	0	0	3	0	2
	P18	0	0	0	2	0	1
VII	P19A	0	0	0	1	0	1
	P20	0	0	0	1	0	0
	P21	0	0	0	0	0	0
VIII	P22B	0	0	0	0	0	0
	P22A	1	4	1	1	1	1
	P23	0	0	0	0	0	0
IX	P52B	0	0	0	0	0	0
	P53	2	2	2	2	2	2
	P54	0	4	3	3	3	2
X	P24	0	0	0	1	0	1
	P25A	0	2	2	2	2	2
	P26	0	0	1	12	1	11
XI	P27B	1	2	1	2	1	1
	P27C	1	2	1	2	1	1
	P28B	1	2	1	2	1	2
	P29	1	2	1	4	1	1
XII	P55B	2	2	2	2	2	1
	P56	0	0	0	1	0	0
	P60	0	0	2	8	2	4
XIII	P30B	1	2	3	3	3	3
	P31A	0	0	0	0	0	0
	P61	0	0	0	5	0	2

Table 2. cont.

Section no.	Designation	Number of outlier observations detected for statistical tests:					
		Q-Dixon		Grubbs	Hampel	Rosner	Iglewicz and Hoaglin
		N9	N13				
XIV	P33B	0	0	2	3	2	2
	P57	0	0	2	2	2	2
XVa	P35C	0	0	0	4	0	2
	P36A	1	2	1	6	1	3
XVb	P35D	1	2	1	1	0	1
	P36B	0	0	2	4	2	4
XV	P35B	0	0	4	6	4	4
	P36	1	2	2	3	2	3
	P62	0	0	2	2	2	2
XVI	P58	1	2	1	2	1	1
	P59	0	2	2	2	2	2
XVII	P38A	2	2	2	2	2	2
	P39A	0	0	1	1	1	1
XVIII	P41A	0	0	0	0	0	0
	P42A	0	0	0	0	0	0
XIX	P44	0	0	0	0	0	0
	P45	1	2	1	3	1	1
Dam							
I	P1	1	2	1	8	1	6
	P2	1	2	3	5	3	5
II	P3	1	2	1	4	3	3
	P4	1	2	10	6	4	8
III	P5	2	4	3	3	3	3
	P6	0	0	4	7	4	5
IV	P7	0	2	2	13	2	9
	P8	2	2	3	11	3	8
Wells							
	S2	0	0	0	4	0	2
	S5	0	0	0	2	0	0
	S6	0	0	0	0	0	0
	S7	2	2	2	2	2	2

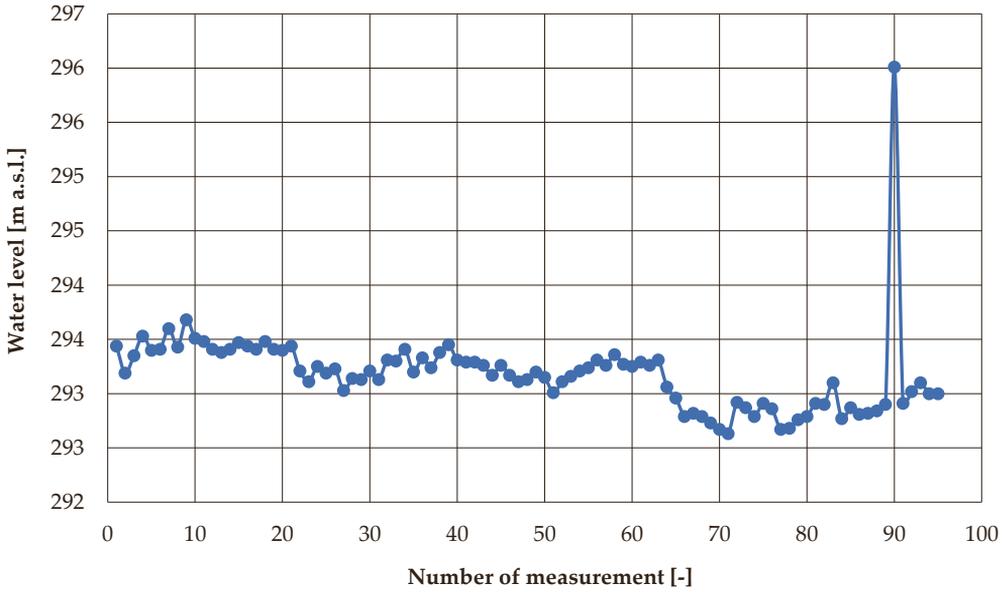


Fig. 3. Diagram of on the P10B water level changes in piezometer P10B at the Czaniec Dam before identification and removal of outliers

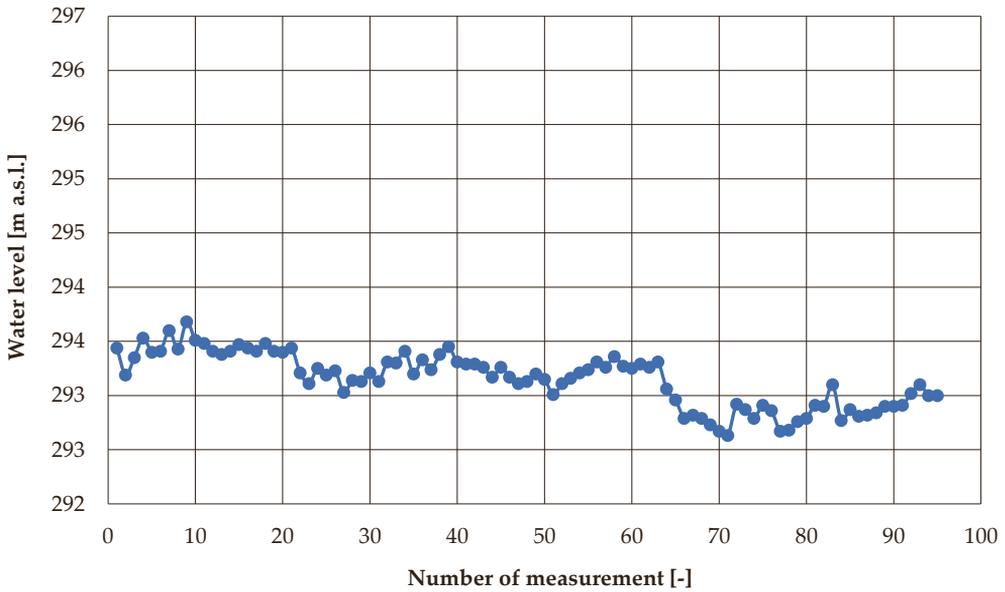


Fig. 4. Diagram in changes of water levels in piezometer P10B at Czaniec Dam after identification and removal of an outlier

For Czaniec Dam in the period from 17.01.2017 to 23.12.2020 for the analyzed open piezometers, the Q-Dixon test (variant N9) showed the existence of 36 outlier observations (0.50% of all results). On the other hand, 39 pairs of outlier observations were determined using variant N13 (1.08% of all results). Grubbs test identified 83 outlier observations (1.15% of all results), while Rosner test showed the existence of 78 outlier observations (1.08% of all results). The Iglewicz and Hoaglin test identified 175 outlier observations (2.42% of all measurements). The highest number of outlier observations was determined by the Hampel test with 237 outlier observations (3.28% of all results).

Figure 5 shows a graph of changes in water level in piezometer P10A, for which the Hampel test revealed 20 outlier observations (Tab. 2). Analyzing this graph, it can be seen that this piezometer reacts to changes of water level in the reservoir (seasonality), and it can be certainly stated that in this data set there are not so many outliers.

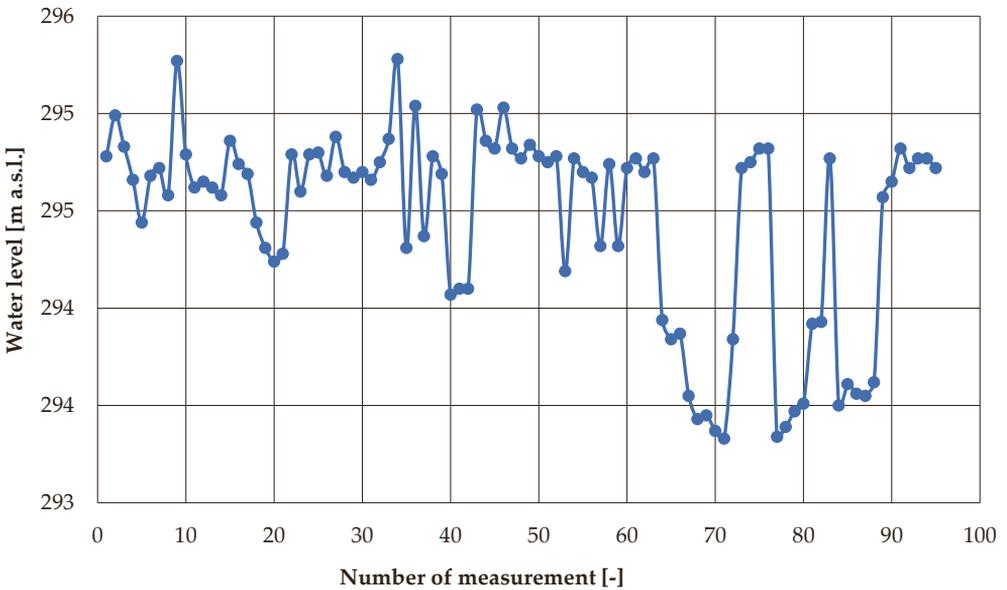


Fig. 5. Diagram of changes in the P10A water level in piezometer P10A at the Czaniec Dam before identification and removal of outliers

The large number of outliers detected for the Hampel statistical test (for example, for piezometer P10A or P2) is related to the design of this test. When there are many observations with the same or similar values in the data set, when calculating the value of deviations r_i from the median value we get a result equal to 0. Then the median value of the set of deviations is also obtained equal to 0, so if the modulus of deviations r_i takes a value greater than 0 then the observation is treated as an outlier.

The Iglewicz and Hoaglin test behaves similarly and shows a similar number of outliers detected; therefore, it can be concluded that both the Hampel test and the Iglewicz and Hoaglin test are not suitable for detecting outliers in piezometric measurements as they are too rigid.

In Table 2, for the P10B piezometer, it is easy to see that the Q-Dixon N13 test detected two outliers, while all others detected only one. These results illustrate the important differences between the Q-Dixon test for pairs of observations (e.g., variant N13) and the test for a single observation (e.g., variant N9), which, due to the design of the test statistic, is unable to detect the pair of largest or smallest outliers. On the other hand, the test for a pair of observations may not be an effective tool when there is only one outlier in the data set. In summary, these results therefore confirm that in practical applications it is worth bearing in mind the relative advantages and disadvantages of both types of Q-Dixon test: the test for a single outlier and the test for a pair of outliers. An important drawback of the Q-Dixon test is that only a single outlier or a pair of outliers can be rejected each time from the analyzed data set. Therefore, the procedure of identification and elimination of coarse errors for a measurement data set devoid of a previously rejected observation must be performed until all outliers are eliminated, which is certainly a cumbersome and time-consuming task.

The Grubbs and Rosner tests showed the existence of a very similar number of outlier observations. Only for three piezometers were differences observed: for the piezometer labelled P35D, P3, and P4. The Grubbs test, as with the Q-Dixon test, can only detect one outlier at a time and should therefore be repeated until there are no further outliers in the data set. The Rosner test, better known as the Extreme Studentized Deviate test (ESD), is a modification of the Grubbs test. The Rosner test can be performed iteratively by analyzing the most deviant values in turn, whereas in the Grubbs test the number of questionable values must be determined a priori. Before performing the Rosner test, the measurement set should be arranged in a nondecreasing sequence, and it should be checked whether the analyzed data set has a normal distribution. The Doornik–Hansen or Shapiro–Wilk test can be used for this purpose. The application of the test requires the maximum number of outliers r in the test sample to be given. This test is applied to sample sizes of $n \geq 25$ observations in which up to 10 outliers are recorded. The ability to detect up to 10 outlier observations at a time is undoubtedly a very important advantage of this test due to the timing of the test. From the point of view of periodic piezometric measurements. The Rosner test therefore appears to be the most suitable test, as it is neither too rigid nor too flexible.

It is an undeniable fact that outliers that are clearly the result of undesirable performance should be removed. So, the question arises: what to do after removing an outlier from the data set? It is possible to replace missing data by the arithmetic mean of neighboring data in the corresponding cell [67]. It is worth remembering, however, that this procedure will reduce the spread of the population and make the

observed distribution more leptokurtic and may increase the probability of making a Type I error (the error of rejecting a null hypothesis that is not in fact false). A more complicated multiple imputation method can also be used, which involves replacing outliers (or missing data) with possible values [68].

Error assessment is a set of issues at the intersection of mathematics, metrology, and statistics that deals with the evaluation and analysis of measurement uncertainty. It encompasses the principles of elaboration and presentation of experimental results. The results of any measurement without error analysis, more specifically without the determination of measurement error (defined as the deviation between the measurement result and its true value), are, in fact, treated as indications only. Measurement error is an intrinsic factor of the measurement process that does not arise only from a mistake. In fact, each measure and is influenced by a large number of factors and variables, resulting in numerous sources of error, including imperfect sense, inaccuracy of measuring instruments and methods, and uncontrolled variability of environmental conditions. Some factors and variables cannot be fully controlled, and anyone can make some increase or decrease in the result. Recognizing the sources of error can help reduce them, but it is important to realize that we can never eliminate all of them. The error in a single measurement cannot be calculated as the difference between the result of the measurement and the true value of the measurement since this value is not known. It can only be estimated, or its components calculated. However, the procedure depends on the recognition of the interactions that affect the result of the measurement. Considering the types of interaction (accidental or systematic), the measurement errors are divided into accidental, systematic, and coarse. Most often the errors are of accidental nature, so the tools of mathematical statistics can be used to process the obtained results.

A single measurement result with an outlier error is usually an extreme value (minimum or maximum) of an ordered set of results. Coarse errors that can occur during piezometric measurements are caused by a number of factors, among which are: damage to the measuring equipment, change in measurement conditions (e.g., icing), incorrect numbering of points or accidental change in the order of two adjacent numbers, storage or preparation for analysis, improper use of the measuring equipment, mechanical damage to the measurement points, mistakes in reading or recording the readings of the measuring instrument, improper method of measurement (data collection) or improper entry of measurement data into the database.

Measurements that are considered questionable are among the most troublesome problems when performing any data analysis. Doubtful results are the result of a one-time influence of a disturbing cause that does not operate continuously and only affects certain measurements.

For measurement series that include the results of measurements made under repeatability conditions, such errors can be easily detected and identified. In case of periodic piezometric measurements, observations are recorded only once for each of the piezometers under study. Therefore, only independent measurements are

available. Possible disturbances and changes during water level changes in piezometers can be noticed by comparing them with the image from previous measurement periods. As a rule, the person who performs piezometric calculations and analyses only receives the results of the measurements themselves, without providing additional information about their course. In such a situation, it is necessary to apply statistical tests by which it will be possible to accept or reject a doubtful measurement at the assumed significance level α .

The concept of error, as applied to scientific measurement, is closely related to the completely unavoidable uncertainty that is intrinsically linked to the essence of making a measurement using a given method. In this sense, errors do not characterize mistakes that can be avoided if greater care is taken in performing measurements. Therefore, one should strive to minimize the size of errors and find a way to estimate their magnitude.

5. Conclusions

All the statistical tests used in this research are used to identify outlier observations in measurement data sets. The purpose of this paper is to identify the test that would be most appropriate for periodically performed piezometric measurements. A very important part of a thorough analysis of measurement data is developing techniques to look for outliers and understand their impact on the analysis performed. Statistical tests based on sample mean and variance can be biased when outliers are present in the data set. Before analyzing piezometric data, it is important to reject coarse errors that, even for a single outlier observation, can significantly affect the result and cause a false assessment or interpretation of the phenomenon under study.

Considering all the static tests used in this paper to identify outlier observations, from the point of view of periodic piezometric measurements, it seems that the Rosner test is the most appropriate test. The test handles piezometric measurements well and is neither too rigid nor too flexible. A great and unquestionable advantage of the Rosner test is its short execution time, as it can detect up to 10 outliers at a time (other tests can only reject a single outlier (Q-Dixon test in the N9 variant or Grubbs test) or a pair of outliers (Q-Dixon test in the N13 variant). Moreover, Rosner test, unlike Hampel's test and Iglewicz and Hoaglin test, also performs well when there are a large number of observations with the same or similar values in the measurement data set and does not indicate them as outliers.

Working effectively with outliers in datasets is certainly a difficult and tedious task. If an observation is found to be an outlier by the chosen statistical test, then each time the analyst should attempt to explain this phenomenon before excluding it from further analysis and decide whether to remove the observation. If no explanation can be found, then such an observation should be treated as extreme but valid and included in further analysis.

References

- [1] ICOLD (International Commission on Large Dams): *Dam Failures Statistical Analysis*. Bulletin, 99, Commission Internationale des Grands Barrages, Paris 1995.
- [2] *Living With Dams: Know Your Risks*. Association of State Dam Safety, Federal Emergency Management Agency, 2013, https://www.fema.gov/sites/default/files/2020-08/fema_living-with-dams_p-956.pdf [access: 21.10.2021].
- [3] McCann M.: *Learning from the US experience*. International Water Power and Dam Construction, no. 12, 1998, pp. 30–32.
- [4] Fiedler K., Hrabowski W.: *Bezpieczeństwo budowy wodnych*. Wydawnictwo Prasa ZSL, Warszawa 1980.
- [5] Fell R., Foster M., Spannagle M.: *The statistics of embankment dam failures and accidents*. Canadian Geotechnical Journal, vol. 37, no. 5, 2000, pp. 1000–1024. <https://doi.org/10.1139/t00-030>.
- [6] Fiedler K.: *Kilka uwag o awariach i katastrofach nasypowych budowy wodnych*. [in:] *Problemy budownictwa wodnego i gospodarki wodno-ściekowej w Polsce ze szczególnym uwzględnieniem regionu Podkarpacia*, Wydawnictwo Politechniki Rzeszowskiej, Rzeszów 2001, pp. 182–185 [conference materials].
- [7] Fiedler K., Gamdzyk J., Jankowski W., Opyrczał L., Selerski S., Wita A., Wróblewski M.: *Awarye i katastrofy zapór – zagrożenia, ich przyczyny i skutki oraz działania zapobiegawcze*, Monografie Instytutu Meteorologii i Gospodarki Wodnej, Instytut Meteorologii i Gospodarki Wodnej, Warszawa 2007.
- [8] Ribas J.R., Severo J.C.R., Guimaraes L.F., Perpetuo K.P.C.: *A fuzzy FMEA assessment of hydroelectric earth dam failure modes: A case study in Central Brazil*. Energy Reports, vol. 7, 2021, pp. 4412–4424. <https://doi.org/10.1016/j.egy.2021.07.012>.
- [9] Fry J.J.: *Lessons on internal erosion in embankment dams from failures and physical models*. [in:] Harris J., Whitehouse H., Moxon S. (eds.), *Scour and Erosion*, CRC Press, 2016, pp. 41–58.
- [10] Crum D.: *Rogue Piezometers*. [in:] *21st Century Dam Design – Advances and Adaptations*. 31st Annual USSD Conference. San Diego, California, April 11–15, 2011, U.S. Society on Dams, Denver 2011, pp. 1591–1602.
- [11] Pelton F.: *Guidelines for Instrumentation and Measurements for Monitoring Dam Performance*. American Society of Civil Engineers, Reston 2000.
- [12] *The Aging of Embankment Dams*. United States Society on Dams, Committee on Materials for Embankment Dams, Denver 2010.
- [13] Leng Y.B., Zhu W.Z., He J.: *Current Situation and Prospects of Dike Anomaly and Infiltration Detecting Technology in China*. Advances in Science and Technology of Water Resources, no. 22(2), 2002, pp. 59–62.
- [14] Molski T.: *Ziemne budowle hydrotechniczne i ich podłoże w warunkach filtracji naporowej*. Infrastruktura i Ekologia Terenów Wiejskich – Infrastructure and Ecology of Rural Areas, nr 3/III, 2012, pp. 221–233.

- [15] Foster M., Fell R.: *Use of Event Trees to Estimate the Probability of Failure of Embankment Dams by Internal Erosion and Piping*. [in:] *Twentieth International Congress on Large Dams: 19–22 September 2000, Beijing, China; [transactions]*, Vol. 1, Question 76, International Commission on Large Dams (ICOLD), Paris 2000, pp. 237–260.
- [16] Torres R.L.: *Considerations for detection of internal erosion in embankment dams*. [in:] *GEO-Velopment: The Role of Geological and Geotechnical Engineering in New and Redevelopment Projects*, American Society of Civil Engineers, 2009, pp. 82–98. [https://doi.org/10.1061/41006\(332\)7](https://doi.org/10.1061/41006(332)7).
- [17] Wang Y., Guo N., Wang Sh., Gu Y.: *Detection of Internal Erosion and Piping in Embankment Dams*. [in:] *Proceedings of the 2016 International Forum on Energy, Environment and Sustainable Development*, Atlantis Press, 2016, pp. 114–122. <https://doi.org/10.2991/ifeesd-16.2016.21>.
- [18] Cwetschek-Wisniewska M.: *Filtracja wody w zaporze Smukała i wybór metod zabezpieczenia przed niekorzystnymi jej skutkami*. *Gospodarka Wodna*, nr 4, 1998, pp. 147–148.
- [19] Dłużewski J.M., Popielski P., Ciuhak K., Hrabowski W.: *Stateczność skarp i osuwisk w ujęciu metody elementów skończonych*. *Inżynieria Morska i Geotechnika*, nr 3, 1999, pp. 118–123.
- [20] Bolt A., Szudek W., Duszyński R., Sukowski T.: *Zabezpieczenie kanału odpływowego stopnia wodnego w Starszynie przed zjawiskami sufozyjnymi*. *Inżynieria Morska i Geotechnika*, nr 3, 2005, pp. 224–239.
- [21] Richards K.S., Reddy K.R.: *Critical appraisal of piping phenomena in earth dams*. *Bulletin of Engineering Geology and the Environment*, vol. 66, 2007, pp. 381–402. <https://doi.org/10.1007/s10064-007-0095-0>.
- [22] Sterpejkowicz-Wersocki W., Szudek W.: *Zabezpieczenie przed sufozją zapory w Smukale*. [in:] *Nachlik E. (red.), HYDROTECHNIKA X '2008: sympozjum ogólnokrajowe: materiały, Ustroń, 13–15 Maja 2008, Śląska Rada Naczelnej Organizacji Technicznej FSNT, Katowice 2008*, pp. 167–174.
- [23] Popielski P., Stasiński J., Wrzosek K.: *Wykorzystanie trójwymiarowej analizy stateczności jako narzędzia wspomaganie projektowania zapory w zmiennych warunkach geologicznych i morfologicznych*. *Czasopismo Techniczne. Środowisko*, R. 108, z. 21, 2011, pp. 137–155.
- [24] Finn W.D.L.: *Finite-Element Analysis of seepage through dams*. *Journal of the Soil Mechanics and Foundations Division*, vol. 93, no. 6, 1967, pp. 41–48. <https://doi.org/10.1061/JSFEAQ.0001074>.
- [25] Johansson S., Dahlin T.: *Seepage monitoring in an earth embankment dam by repeated resistivity measurements*. *European Journal of Engineering and Environmental Geophysics*, no. 1(3), 1996, pp. 229–247.
- [26] Panthulu T.V., Krishnaiah C., Shirke J.M.: *Detection of seepage paths in earth dams using self-potential and electrical resistivity methods*. *Engineering Geology*, vol. 59, no. 3–4, 2001, pp. 281–295. [https://doi.org/10.1016/S0013-7952\(00\)00082-X](https://doi.org/10.1016/S0013-7952(00)00082-X).

- [27] Opyrczał L., Popielski P., Tomaszewicz A., Jankowski W.: *Wpływ zmienności parametrów gruntu na współczynnik bezpieczeństwa skarpy oraz kształt krzywej poślizgu*. Prace Naukowe Instytutu Geotechniki i Hydrotechniki Politechniki Wrocławskiej, vol. 75, nr 41, 2005, pp. 415–420.
- [28] Radzicki K.: *Wykorzystanie metod termodekacji do badań filtracji w ziemnych budowlach piętujących*. Gospodarka Wodna, nr 9, 2005, pp. 372–376.
- [29] Bolt A., Sukowski T., Szudek W.: *Wykorzystanie modelowania numerycznego procesów sufozji w ocenie stanu i rewitalizacji małych elektrowni wodnych*. Zeszyty Naukowe Politechniki Gdańskiej. Budownictwo Lądowe, nr 57, 2006, pp. 335–342.
- [30] Popielski P.: *Wykorzystanie „analizy wstecz” i nowoczesnych badań gruntu do weryfikacji modeli numerycznych w geotechnice i hydrotechnice*. Prace Naukowe Politechniki Warszawskiej. Inżynieria Środowiska, z. 54, 2007, pp. 259–280.
- [31] Popielski P., Stasiński J.: *Możliwości zastosowania MES do oceny stateczności zboczy ziemnych*. Prace Naukowe Politechniki Warszawskiej. Inżynieria Środowiska, z. 54, 2007, pp. 281–291.
- [32] Popielski P., Stasiński J.: *Symulacja procesu deformacji wielkokubaturowych budowli ziemnych*. Czasopismo Techniczne. Środowisko, R. 104, z. 1, 2007, pp. 117–127.
- [33] Popielski P., Zaczek-Peplinska J.: *Wykorzystanie modeli numerycznych w eksploatacji budowli piętujących*. Gospodarka Wodna, nr 2, 2008, pp. 73–80.
- [34] Hung M.-H., Lauchle G.C., Wang M.C.: *Seepage-Induced acoustic emission in granular soils*. Journal of Geotechnical and Geoenvironmental Engineering, vol. 135, no. 4, 2009, pp. 566–572. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2009\)135:4\(566\)](https://doi.org/10.1061/(ASCE)1090-0241(2009)135:4(566)).
- [35] Sjö Dahl P., Dahlin T., Johansson S.: *Embankment dam seepage evaluation from resistivity monitoring data*. Near Surface Geophysics, vol. 7(5), 2009, pp. 463–474. <https://doi.org/10.3997/2214-4609.20146268>.
- [36] Radzicki K., Bonelli S.: *Lokalizacja procesów filtracyjnych oraz określenie stopnia ich nasilenia za pomocą analizy modelem IRFTA światłowodowych pomiarów temperatury*. Czasopismo Techniczne. Środowisko, R. 107, z. 1, 2010, pp. 55–62.
- [37] Freeze A.: *Influence of the unsaturated flow domain on seepage through earth dams*. Water Resources Research, vol. 7, no. 4, 1971, pp. 929–941. <https://doi.org/10.1029/wr007i004p00929>.
- [38] Gao Y., Jin F., Wang X., Wang J.: *Finite Element Analysis of dam-reservoir interaction using high-order doubly asymptotic open boundary*. [in:] *Seismic Safety Evaluation of Concrete Dams: A Nonlinear Behavioral Approach*, Butterworth-Heinemann, 2013, pp. 173–198.
- [39] Elliotis M.C.: *A mathematical model for a steady-state seepage flow of groundwater under a reinforced concrete dam*. Applied Computing and Geosciences, vol. 1, 2019, pp. 1–9. <https://doi.org/10.1016/j.acags.2019.100003>.

- [40] Chandola V., Banerjee A., Kuma V.: *Anomaly Detection: A Survey*. ACM Computing Surveys, vol. 4, no. 3, 2009, pp. 1–58. <https://doi.org/10.1145/1541880.1541882>.
- [41] Phoon K.-K., Kulhawy F.H.: *Evaluation of geotechnical property variability*. Canadian Geotechnical Journal, vol. 36, no. 4, 1999, pp. 625–639. <https://doi.org/10.1139/t99-039>.
- [42] Baecher G.B., Christian J.T.: *Reliability and Statistics in Geotechnical Engineering*. John Wiley & Sons, New York 2005.
- [43] Zhang W.G., Goh A.T.C., Zhang Y.M., Chen Y.M., Xiao Y.: *Assessment of soil liquefaction based on capacity energy concept and multivariate adaptive regression splines*. Engineering Geology, vol. 188, no. 7, 2015, pp. 29–37. <https://doi.org/10.1016/j.enggeo.2015.01.009>.
- [44] Zhang W.G., Goh A.T.C.: *Multivariate adaptive regression splines and neural network models for prediction of pile drivability*. Geoscience Frontiers, vol. 7, no. 1, 2016, pp. 45–52. <https://doi.org/10.1016/j.gsf.2014.10.003>.
- [45] Yong R., Ye J., Li B., Du S.: *Determining the maximum sampling interval in rock joint roughness measurements using Fourier series*. International Journal of Rock Mechanics and Mining Sciences, vol. 101, no. 1, 2018, pp. 78–88. <https://doi.org/10.1016/j.IJRMMS.2017.11.008>.
- [46] Hawkins D.: *Identification of Outliers*. Chapman and Hall, London 1980.
- [47] Yuen K.V., Mu H.Q.: *A novel probabilistic method for robust parametric identification and outlier detection*. Probabilistic Engineering Mechanics, vol. 30, 2012, pp. 48–59. <https://doi.org/10.1016/j.probenmech.2012.06.002>.
- [48] Rousseeuw P.J.: *Tutorial to robust statistics*. Journal of Chemometrics, vol. 5, no. 1, 1991, pp. 1–20. <https://doi.org/10.1002/cem.1180050103>.
- [49] Han J., Kamber M.: *Data Mining*. Morgan Kaufmann Publishers, New York 2001.
- [50] Wang Y., Cao Z.J., Li D.Q.: *Bayesian perspective on geotechnical variability and site characterization*. Engineering Geology, vol. 203, 2016, pp. 117–125. <https://doi.org/10.1016/j.enggeo.2015.08.017>.
- [51] Ching J., Phoon K.K.: *Constructing site-specific multivariate probability distribution model using Bayesian machine learning*. Journal of Engineering Mechanics, vol. 145, no. 1, 2019, pp. 1–46. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0001537](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001537).
- [52] Zheng S., Zhu Y.-X., Li D.-Q., Cao Z.-J., Deng Q.-X., Phoon K.-K.: *Probabilistic outlier detection for sparse multivariate geotechnical site investigation data using Bayesian learning*. Geoscience Frontiers, vol. 12, 2021, pp. 425–439. <https://doi.org/10.1016/j.gsf.2020.03.017>.
- [53] Lach S., Opyrchal L.: *Using the modified scalar product approach for testing the direction of seepage through the earth-fill dam in Piecryska*. Journal of Water and Land Development, no. 33, 2017, pp. 89–98. <https://doi.org/10.1515/jwld-2017-0023>.