*Desalination and Water Treatment*
www.deswater.com

doi: 10.5004/dwt.2020.25440

186 (2020) 144–154
May

# Impact of the selected indicators of the wastewater quality and operating parameters of the biological reactor on the simulation of sludge sedimentation: probabilistic approach

Bartosz Szeląg[a,*], Krzysztof Barbusiński[b]

[a]*Kielce University of Technology, Tysiąclecia Państwa Polskiego 7 Av., 25-314 Kielce, Poland, Tel. +41-342-47-35; email: bszelag@tu.kielce.pl (B. Szelag)*
[b]*Silesian University of Technology, Konarskiego 18 Street, 44-100 Gliwice, Poland, Tel. +32-237-11-94; email: krzysztof.barbusinski@polsl.pl (K. Barbusinski)*

Received 9 October 2019; Accepted 3 January 2020

---

## ABSTRACT

A probabilistic model was used to simulate the short- and long-term reliability of the treatment plant in relation to the sedimentation of activated sludge. In this model, it was assumed that the quantity and quality of wastewater flowing into the treatment plant and weather conditions are stochastic. While, the control variables are: the activated sludge concentration, the oxygen concentration in the nitrification chamber and the coagulant dose. The model includes the possibility of failure of wastewater quality analyzers, which are also of random nature. The probabilistic model presented in the paper consists of three components. The first of them is a classification model to identify the sedimentation capacity of activated sludge. The second component consists of generators of quantity, quality of wastewater, as well as weather conditions based on the Monte Carlo method. The third component is a failure generator for wastewater quality analyzers at the inlet to a wastewater treatment plant. Using the developed probabilistic model at work, a number of bioreactor optimization strategies were analyzed in a long-term (1 y) perspective. The analyzes carried out confirmed that the proposed probabilistic model is a valuable tool for optimizing the operation of wastewater treatment plants and allows the assessment of the impact of dynamic changes in the reactor control variables on the reliability of the facility's operation in long-term and short-term.

*Keywords:* Wastewater treatment plant; Sludge volume index; Probabilistic model; Reliability

---

## 1. Introduction

Control of operation efficiency of a wastewater treatment plant is a complex task. In order to implement, it is necessary to properly select the settings in the biological reactor in the case of dynamically changing inlet conditions and meteorological factors. The control and correction of the settings are aimed at obtaining the wastewater quality indexes at the outlet from the wastewater treatment plant that does not exceed the admissible values [1–3]. It is also important to maintain the selected technological facilities within the determined technological parameters range so that they do not deteriorate the quality of wastewater at the outlet and on the other hand do not increase the energy consumption and costs of the wastewater treatment plant. The sludge volume index (SVI) is one of the important parameters affecting the wastewater treatment plant operation. The performed laboratory examinations and computer simulations [4,5] showed that the increase in the SVI may lead to increased values of the selected wastewater quality indexes at the outlet from the treatment plant and generate problems related to excessive sludge dehydration. Therefore, it is necessary to continuously control the settings in the biological reactor in order

---

* Corresponding author.

*Presented at the 14th Conference on Microcontaminants in Human Environment, 4–6 September 2019, Czestochowa, Poland*

to maintain the SVI within the optimum range (prevent the sludge bulking; for systems with integrated removal of biogenic compounds, the value of the SVI should not exceed 150 $cm^3/kg$) and to ensure the required quality of wastewater. The reason why, on the one hand, to be able to diagnose the wastewater treatment plant on a continuous basis, to control efficiency and to optimize its operation, a mathematical model can be applied (statistical and physical).

The models mentioned above can be divided into deterministic and probabilistic models. In the first kind of model, the input variables are measurable values that can be measured online, for example, with measuring probes, which are usually used for short-term simulations. Probabilistic models are used many times for long-term simulations. Moreover, the definition of input variables in these models is more complex, because by examining the variability of these variables in time it is possible to determine their statistical distributions, but it is not possible to predict their consecutive values in time series. Physical models activated sludge model (ASM) are based on systems of differential equations describing biochemical processes taking place in particular treatment plant objects (settling tanks, activated sludge chambers, separated fermentation chambers, etc.) and allowing for simulation of wastewater quality at the plant outlet, greenhouse gas emission, biogas [5], etc. However, due to problems with the calibration of ASM models, statistical models are increasingly used. The basis for their creation is data collected at wastewater treatment plants. In implementation and calibration, these models are usually simpler than physical models, as described in the literature [1–3]. Their implementation at the stage of wastewater treatment plant operation is more difficult than physical models because it is necessary to verify the obtained models by assessing the impact of individual independent variables on the simulation results.

The implementation of mathematical models allows us to assess the facility's operation (forecast the wastewater quality at the outlet and sedimentation of the activated sediment in the secondary clarifier) under the calculation conditions, that is, assuming the proper quantity and quality of wastewater. The results of these analysis allow creating algorithms stabilizing the wastewater treatment plant operation, which has a crucial meaning during practical application [6–8]. However, the adoption of a single calculation scenario (covering the period of a week, month, year, etc.), wherein the presence of random disorders (e.g. heavy rainfall) in the physical and statistical model is assumed in advance, maybe doubtful. This results from the fact that the variability of quantity, quality of wastewater and meteorological conditions within an appropriate time frame affect the effectiveness of the wastewater treatment plant and the selected settings in real-time. Therefore, it seems difficult to assess the impact of variants related to the stabilization of the wastewater treatment plant operation on the effectiveness of its operation, energy consumption, and short- and long-term costs. In order to determine the impact of diversified conditions at the inlet on the reactor operation effectiveness, their stochastic character is modeled by using a random numbers generator (Monte Carlo method); the approach is presented in papers [9–12]. The results thus obtained allowed to determine the impact of varying quantity and quality of wastewater on the effectiveness of the adopted reactor operation strategy. On the other hand, the facility inertia, its susceptibility to varying conditions at the inlet and reliability of functioning can be determined based on such analysis. However, in the papers within this field, concerning the wastewater treatment plant, the authors [13,14] focus on the analysis of the operational effectiveness in the case of single episodes for the adopted fixed settings and the time of the wastewater treatment plant operation is not considered. As a result, this generates difficulties in the interpretation of the obtained analysis results. This problem was effectively resolved in hydrological models [15] where suitable calculation procedures were developed at the stage of Monte Carlo simulation. Despite the fact that performed simulations show the exceeding of the quality indexes (usually total nitrogen), the aspects related to the reliability of the facility operation are neglected, which form a treatment plant point of view can be a significant hint for a process engineer. An important drawback of the approach presented so far in mathematical models is the assumption of measurement continuity of wastewater quality indexes, being independent variables. Therefore, such models do not consider failures of the devices for wastewater quality measurements, which are permanent and inseparable elements of an everyday wastewater treatment plant operation.

Therefore, the paper presents an innovative probabilistic model for the analysis of the wastewater treatment plant operation with reference to the sedimentation of activated sludge. The model assumes that the independent variables that include the wastewater quantity and quality as well as meteorological conditions are of random nature and are being modeled using generators based on the Monte Carlo method. The model allows short-term (1 d) and long-term (month, year) simulations as well as considers failures of the wastewater quality analyzers at the inlet, which fact was not included in the probabilistic models of wastewater treatment plants.

## 2. Object of study

The Sitkówka–Nowiny wastewater treatment plant located in the suburbs of Kielce (Central Poland) was the basis of the studies. The treatment plant's rated capacity is 72,000 $m^3/d$, corresponding to 275,000 PE (population equivalent). This is a mechanical, biological and chemical treatment plant. Large impurities are removed on the grates, then the mineral suspension is removed in the grit, and the organic matter in the preliminary sedimentation tank. From the mechanical part, the wastewater flows to the biological reactor, which operates in a modified BARDENPHO system with a pre-denitrification chamber. In case of problems with the biological removal of the phosphorous compounds, it is possible to precipitate them using a coagulant (PIX). The treated wastewater flow to four secondary clarifiers, from where are discharged to the Bobrza River. A detailed characterization of the object is presented in the papers by Szeląg et al. [16,17].

## 3. Methodology

The paper presents a probabilistic model concerning the analysis of the wastewater treatment plant operation in relation to the sedimentation of activated sludge in the secondary sedimentation tank. The model assumes that quantity and quality of wastewater and weather conditions are of random character and are described using theoretical distributions, however, the control variables (the reactor settings) are set by the process engineer of the facility. The presented mathematical model allows us to estimate the treatment plant operation reliability within short-term and long-term and allows us to determine the impact of the wastewater quality analyzers measurement system failure on the facility operation effectiveness. The computational diagram of the probabilistic model algorithm is given in Fig. 1. The proposed model can be divided into three components. The first component represents the classification model based on the logistic regression method, used to identify the sedimentation ability of the activated sludge based on the quantity and quality of wastewater, weather conditions and control variables. The second component of the model is represented by independent variables generators included in the logit model. The paper assumes that the studied variables are independent, as shown in the paper by Szeląg et al. [18], and are modeled using the Monte Carlo generators. The generators include also a simulator for episodes during which failures occurred of analyzers of the wastewater quality selected indexes at the inlet to the wastewater treatment plant. The third component is a computational block simulating the activated sludge sedimentation in the secondary clarifier for a specified time series, based on which the reliability of the wastewater treatment plant operation is determined. The next steps of the calculation algorithm include:

- Determination of independent variables and classification model for the description of activated sludge sedimentation,
- Determination of empirical distributions and selection of theoretical ones for the independent variables including the wastewater quantity and quality, and weather conditions,
- $K$ – fold simulation using the Monte Carlo method of the annual time series concerning the quantity and quality of wastewater and weather conditions,
- Selection of the control (optimization) strategy of the treatment plant depending on the conditions at the inlet and weather conditions,
- $T$ – fold simulation using the Monte Carlo method of $Z$ – failures of analyzers within the $K$ – time series,
- Simulation of activated sludge sedimentation based on the model $p = f(x_1, x_2, x_3, \ldots, x_j)$ within the determined $T \cdot K$ time series and determination of the treatment plant operation reliability $R(t)$,
- Determination of the probabilistic solution, that is, curve describing the probability of exceeding the value $R(t)$ for the adopted control/optimization strategy.

The paper presents an example of the construction of the aforementioned probabilistic model for the Sitkówka–Nowiny wastewater treatment plant.

### 3.1. Logistic regression

The logistic regression, also referred to as a binomial logit model, is used during the analysis of binary data and hence can be applied during the probability forecasting. The model is commonly applied in medicine, social sciences, economy [19]. Because of its advantages, it is more often used during simulation of complex phenomena taking place within the technological facilities of wastewater treatment plants [20,21]. The logistic regression represents a case of a generalized linear model in the following Eq. (1):

$$p = \frac{\exp(X)}{1+\exp(X)} = \frac{\exp\left(\sum_{i=1}^{j} \alpha_i \cdot x_i + \alpha_0\right)}{1+\exp\left(X\sum_{i=1}^{j} \alpha_i \cdot x_i + \alpha_0\right)} \tag{1}$$

where $p$ – probability of exceeding the limit value of the considered dependent variable $\alpha_i$ – coefficients estimated using the highest reliability method, $x_i$ – independent variables, $X$ – linear combination of independent variables expressed as $\sum_{i=1}^{j} \alpha_i \times x_i + \alpha_0$.

The analysis performed by Szeląg et al. [17] for the Sitkówka–Nowiny wastewater treatment plant showed that to identify the sedimentation capability (activated sludge bulking), it is possible to use a relationship representing a combination of random and control variables (2)–(4):

$$X = \beta_l + \beta_c \tag{2}$$

$$\beta_l = 0.02 \cdot \frac{\mathrm{BOD}_5}{\mathrm{TN}} + 0.32 \cdot \frac{\mathrm{BOD}_5}{\mathrm{TP}} + 0.0012 \cdot L_{N-\mathrm{NH}_4} - 0.37 \cdot T_K + 14.38 \tag{3}$$

$$\beta_c = -1.36 \cdot X_{\mathrm{OC}} - 1.76 \cdot m_{\mathrm{PIX}} - 1.18 \cdot \mathrm{DO} \tag{4}$$

where $\beta_l$ – variables of stochastic character, $\beta_c$ – control variables, DO – oxygen concentration within nitrification chamber (mg/L), $X_{\mathrm{OC}}$ – concentration of activated sludge (kg/m$^3$), $m_{\mathrm{PIX}}$ – daily dosage PIX (m$^3$/d), $T_K$ – temperature in the activated sludge chambers (°C), $L_{\mathrm{N-NH}_4}$ – load of the ammonia nitrogen (kg/d), TN – content of total nitrogen (mg/L), TP – content of total phosphorous (mg/L), BOD$_5$ – biochemical oxygen demand (mg/L).

The above relationship was achieved with the assumption that sludge bulking occurs when SVI = 150 cm$^3$/g and more, which corresponds to $p = 0.50$. The above-mentioned equation and model for the description of sludge sedimentation requires comment because the variables included in the result from the data collected at the Sitkówka–Nowiny treatment plant. The independent variables are given in Eq. (4) are predictive and statistically significant at the assumed confidence level. Nevertheless, it does not mean that other factors and exploitation operations have neglected influence on the phenomenon of sedimentation. Certainly, their influence is much smaller than the variables included in the above equations. Moreover, this equation has a local character. The analysis performed by Szeląg et al. [17] showed that the sedimentation course is also influenced by the way the plant is operated. Therefore, the above equation may take different forms and may include any variables provided that the necessary data are collected at the treatment plant to allow taking them into account in the model.

### 3.2. Reliability of the wastewater treatment plant operation

The reliability is one of the most important parameters allowing estimating the facility operation effectiveness. The reliability of the facility can be expressed as follows [22]:

$$R(t) = p(t > \tau) \tag{5}$$

where $R(t)$ – reliability, $t$ – facility operation time without failures, $\tau$ – required operation time without failures.

Applying the above relationship to the wastewater treatment plant operation in any time interval ($T$) covering $N$ calculation events, one obtains the following relationship:

$$R(t) = \frac{\sum r(x_1, x_2, \ldots, x_j)}{T} \tag{6}$$

where $T$ – time period covering $N$ measurement events for which the considered index value is determined; in the paper, it was assumed that $N = 365$ d, corresponding to the annual time series; $r(x_1, x_2, \ldots, x_j)$ – is a function of statistical character or covering a physical model and its values are represented by binary variables adopting values 0 or 1. Value 0 corresponds to an event, wherein the adopted value of index $x_{gr}$ representing the basis for the treatment plant operation evaluation, indicates a drop of the facility operation effectiveness (exceeding the admissible indexes at the outlet, sludge bulking, etc.); therefore the adjustment of the reactor settings is necessary. Value 1 corresponds to an event, for which there were no interruptions in the operation of individual treatment plant facilities.

Within the analysis performed within the study, an assumption was made that the value of function $r(x_1, x_2, \ldots, x_j)$ is 1 when $p(\text{SVI} = 150\ \text{cm}^3/\text{g}) \leq 0.50$, otherwise the value of function $r$ is 0. Moreover, within the performed calculations, the coefficient of reliability (COR) related to the sedimentation and described using the following equation was used to evaluate the treatment plant operation (for individual episodes) (7):

$$\text{COR}_{\text{SVI}} = \frac{p(x_1, x_2, \ldots, x_j)}{p_{\text{SVI,min}}} \tag{7}$$

where $p$ – probability of exceeding the limit value of SVI = 150 cm$^3$/g for the independent variables combination $(x_1, x_2, x_3, \ldots, x_j)$, $p_{\text{SVI,min}}$ – minimum limit value of the probability, the exceeding of which conditions the sludge bulking.

By substituting Eqs. (7)–(6) and making the appropriate adaptations, it can be demonstrated that the reliability of wastewater treatment plant performance with respect to activated sludge sedimentation over a selected time period $T$ can be described by following stochastic integrals:

$$R(T) = \int p(x_1, x_2, \ldots, x_j)\, dT \tag{8}$$

$$\text{COR}_{\text{SVI}}(T) = \int \frac{p(x_1, x_2, \ldots, x_j)}{p_{\text{SVI,min}}}\, dT \tag{9}$$

where $x_{1,2,j}$ – independent variables describing the quantity and quality of wastewater and bioreactor temperature, described by means of empirical probability density functions $f(x_i)$ and distribution function $\int f(x_i)$.

In Eq. (8), a number of variables related to the operational variables included in Eq. (4). They are not stochastic, but their variation patterns depend on the amount and quality of wastewater at the inflow, the temperature in the biological reactor, and are described by the conditions $\varnothing_{z(S)}$ specified in Eqs. (10)–(12). The following values of the reactor operating parameters can be determined by means of appropriate conversions of Eqs. (8) and (9), for which the selected values of the settings ensure $\text{COR}_{\text{SVI}} > 1$ and $p < 0.5$:

$$x(\beta_c)_j = F^{-1}\left\{\int R(p(x_1, x_2, \ldots, x_j) > p_{\text{SVI,min}})\, dT\right\} \tag{10}$$

Due to the complexity of the above Eqs. (8) and (9) and the limited possibilities of their solution by analytical methods, the Monte Carlo method was used to solve them. The advantage of the adopted calculation methodology is the fact that it is possible to analyze the influence of variables for any form of $\varnothing_{z(S)}$, which is usually not taken into account in the probabilistic solutions of wastewater treatment plants.

### 3.3. Generators of wastewater quantity and quality, and of weather conditions

The Monte Carlo method is one of the commonly applied mathematical tools that allow simulating time series of an investigated phenomenon of random character (e.g. rainfall, air temperature, quantity and quality of wastewater at the inlet of the treatment plant). It has been confirmed by many papers in the field of hydrology, ecology and wastewater treatment [10,15,18,23]. In this method, the data gathered with the appropriate resolution are the basis for the simulation of the considered variable or variables. Based on the data, their variability is determined, expressed by means of the empirical distribution. In the next stage, the distribution is approximated by means of a theoretical distribution, used for simulation calculations. To assess the compatibility of theoretical and empirical distributions, many statistical tests are used, such as Kolmogorov–Smirnov (K–S) and Chi–square. At the stage of simulating a few variables ($x_i$) are important whether they are dependent or independent variables; this is determined based on the value of correlation coefficient $R$. If the variables are independent, $i$-independent generators are created by means of which the data are modeled [18]. However, when the variables are dependent, a modification is performed using the Monte Carlo method, introducing proper calculation algorithms (e.g. Iman Conover method, copula functions, etc.).

In this paper, based on independent random variables ($\text{BOD}_5/\text{TN}$, $\text{BOD}_5/\text{TP}$, $T_K$), in the logit model, their empirical distributions have been determined and then compared to theoretical distributions. In order to match the theoretical data to empirical one as good as possible, the following statistical distributions were considered [10,24]: Weibull, Chi–square, exponential, GEV, Gumbel, gamma, Johnson, normal, log-normal, Pareto and beta. To assess the empirical and theoretical distribution, the Kolmogorov–Smirnov (K–S) and Chi–square (Ch) tests were used. For simulation calculation purposes, it was assumed that the number of failures in the annual cycle is fixed, the days of faulty operations within a year are being modeled, and the results obtained this way cannot be used for the optimization of the treatment plant operation.

### *3.4. Selection of the biological reactor settings*

Proper selection of the settings is a key for the wastewater treatment plant operation effectiveness within the annual cycle. Their values are determined by the quality of discharge at the outlet of the wastewater treatment plant and the magnitude of technological parameters (SVI) maintained at the stage of the plant operation. The analysis performed by Szeląg et al. [17] show that in order to maintain the SVI within the optimum range, it is necessary to control simultaneously the activated sludge concentration ($X_{\mathrm{OC}}$), the concentration of oxygen in the nitrification chamber (DO) and the quantity of dosed coagulant ($m_{\mathrm{PIX}}$). The values of these parameters vary depending on the quantity and quality of wastewater at the inlet and the temperature of activated sludge in the chambers [25,26] and also affect the quality of discharge at the outlet from the wastewater treatment plant [4,5]. By optimal selection of the reactor settings at the stage of operation, it is possible to limit the sludge bulking problems, which requires a set of control rules (S1) (11):

$$\varnothing_{z(S1)} = \begin{cases} 4{,}50 \geq X_{\mathrm{OC,bz}} \geq 2{,}50 \\ 2{,}50 \geq \mathrm{DO}_{\mathrm{bz}} \geq 2{,}00 \\ 5{,}00 \geq X_{\mathrm{OC,op}} \geq 4{,}50 \\ 2{,}50 \geq \mathrm{DO}_{\mathrm{op}} \geq 2{,}00 \\ 5{,}00 \geq X_{\mathrm{OC,T}} \geq 4{,}50 \\ 2{,}50 \geq \mathrm{DO}_{T} \geq 2{,}25 \\ m_{\mathrm{PIX}} = f\left(\Phi\left(X_{\mathrm{OC}}, \mathrm{DO}\right)\right) \end{cases} \tag{11}$$

where $X_{\mathrm{OC}}$,bz, $\mathrm{DO}_{\mathrm{bz}}$ – range of variation of $X_{\mathrm{OC}}$, DO during rainy weather in spring-summer period, $X_{\mathrm{OC,op}}$, $\mathrm{DO}_{\mathrm{op}}$ – range of variation of $X_{\mathrm{OC,DO}}$ value during rainy weather and at lowered air temperature, $X_{\mathrm{OC},T}$ – $X_{\mathrm{OC}}$ value determined depending on air temperature and season [17].

An implementation example of the above-mentioned set of controlling rules to control the secondary clarifier operation is presented based on the modeled time series of wastewater quantity and quality for winter, spring, and summer [17]. Nevertheless, the set of these control rules was not tested within long-term, when the quantity and quality of wastewater were stochastic. In practical considerations, the implementation of a function describing the coagulant dose ($m_{\mathrm{PIX}}$) may be complex due to limited access to the data concerning the quality of wastewater with appropriate advance. Therefore, there may be cases, where the volume of dosed coagulant and the remaining settings ($X_{\mathrm{OC}}$, DO) are specified considering first and foremost the changes of temperature in the activated sludge chambers [17,20,27,28]. At the stage of operation, there are theoretically possible variants where the quantity of supplied coagulant is fixed, nevertheless, such an approach is not practically justified from the economic and technological standpoint.

Considering the above remarks, the following analysis present a few different sludge sedimentation control concepts in the secondary clarifier (Sitkówka–Nowiny wastewater treatment plant) by maintaining the reactor settings within an appropriate range:

- (S2 – var.) control of DO, $X_{\mathrm{OC}}$ and $m_{\mathrm{PIX}}$ values depending on $T_K$ (12):

$$\varnothing_{z(S2)} = \begin{cases} T_K \leq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}5\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 5{,}00\,\mathrm{kg/m^3} \\ m_{\mathrm{PIX}} = 1{,}14\,\mathrm{m^3/d} \\ T_K \geq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}25\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 4{,}50\,\mathrm{kg/m^3} \\ m_{\mathrm{PIX}} = 0{,}51\,\mathrm{m^3/d} \end{cases} \tag{12}$$

- (S3 – min., S4 – max.) control of DO and $X_{\mathrm{OC}}$ values depending on $T_K$ and $m_{\mathrm{PIX}}$ = const. (13) and (14):

$$\varnothing_{z(S3)} = \begin{cases} T_K \leq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}5\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 5{,}00\,\mathrm{kg/m^3} \\ T_K \geq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}25\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 4{,}50\,\mathrm{kg/m^3} \\ m_{\mathrm{PIX}} = 0{,}51\,\mathrm{m^3/d} \end{cases} \tag{13}$$

$$\varnothing_{z(S4)} = \begin{cases} T_K \leq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}5\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 5{,}00\,\mathrm{kg/m^3} \\ T_K \geq 13{,}7^{\circ}\mathrm{C} \\ \mathrm{DO} = 2{,}25\,\mathrm{mg/l} \\ X_{\mathrm{OC}} = 4{,}50\,\mathrm{kg/m^3} \\ m_{\mathrm{PIX}} = 1{,}14\,\mathrm{m^3/d} \end{cases} \tag{14}$$

In the case of S1 strategy, for random values, $\mathrm{BOD_5/TN}$, $\mathrm{BOD_5/TP}$, $L_{\mathrm{N-NH_4}}$, and $T_K$, the values of $m_{\mathrm{PIX}}$, $X_{\mathrm{OC}}$, and DO were calculated so that $p = 0.50$, which means lack of sludge bulking within the analyzed period, therefore $R(t = 365\ \mathrm{d}) = 1.0$. Therefore it is necessary to analyze the variability of operating parameters simulation of the activated sludge chambers, for the values of wastewater quality indexes and temperature in the bioreactor obtained by means of Monte Carlo simulators. Trends adopted in the paper in temperature changes of activated sludge in municipal wastewater treatment plants and the necessity to modify the setpoint values in a biological reactor are confirmed by the operation of existing facilities and simulation experiments [20,25]. For the remaining control strategies (S2–S4), owing to the fact that the values of $X_{OC}$, DO, and $m_{PIX}$ is set, values $R(t)$ may change within a proper range of variability. The S2–S4 control strategies focused on the analysis of the reliability of the facility operation $R(t)$ and technological aspects related to the dosage of coagulant PIX.

## 4. Results

### *4.1. Descriptive characteristics of data on wastewater treatment plants.*

Based on the results of measurements performed in the Sitkówka–Nowiny treatment plant in the period of 2012–2016 it was determined that the values of independent variables included in the logistic regression equation (2–4) indicate a significant variability depending on the season, which highly affects the value of SVI (Table 1). Variability of the conditions at the inlet and weather conditions significantly affects the reactor operating parameters as confirmed by the values of appropriate parameters (DO, $X_{OC}$, $m_{PIX}$). The data in Table 1 also show that the seasons and therefore reactor temperatures have a significant impact on the sedimentation of activated sludge [20,21]. At the same time, the variability of the SVI values given in Table 1 in the light of Bezak et al. [29] confirms the presence of filamentous bacteria such as Microthrix, Beggiatoa, Spherotilus natans in the sediment, which usually occurred during the transitional period (autumn-winter-spring) and which was observed in many wastewater treatment plants in Poland and abroad [28].

Moreover, the data in Table 1, in the light of the empirical formula developed, indicate that the exploitation problems at the wastewater treatment plant related to the deterioration of sedimentation capacity of the sludge are associated with a periodical decrease in the content of nutrients in the inflowing wastewater [4].

Table 1
Variability of wastewater quantity and quality, and of operating parameters of the biological reactor in the Sitkówka–Nowiny wastewater treatment plant

| Indicators | Units | Winter | | | | Spring, summer, autumn | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Min. | Mean | Max. | Standard deviation | Min. | Mean | Max. | Standard deviation |
| $Q$ | m$^3$/d | 29.952 | 39.364 | 88.986 | 6.563 | 3.0125 | 41.842 | 94.772 | 8.559 |
| $BOD_5$ | mg$O_2$/L | 151 | 290 | 489 | 81.83 | 132 | 340 | 557 | 81.2 |
| $N–NH_4$ | mg/L | 28 | 48.9 | 62 | 5.68 | 22 | 54.52 | 66.9 | 7.13 |
| TN | mg/L | 56.2 | 82.01 | 95.16 | 8.42 | 39.9 | 95.15 | 124.1 | 11.58 |
| $TN_{eff}$ | mg/L | 3.6 | 7.02 | 17.8 | 2.38 | 6.26 | 8.89 | 13.92 | 1.43 |
| $TP_{eff}$ | mg/L | 3.1 | 7.22 | 12.1 | 1.44 | 3.5 | 7.83 | 12.6 | 1.65 |
| $T_K$ | °C | 10 | 11.9 | 13.5 | 0.8 | 11.3 | 17.8 | 23 | 3.1 |
| DO | mg/L | 1.8 | 2.85 | 3.25 | 0.8 | 1.51 | 2.2 | 3.25 | 0.65 |
| $X_{OC}$ | mg/L | 2.85 | 4.95 | 6.54 | 0.84 | 2.15 | 4.11 | 5.28 | 0.95 |
| Waste activated sludge | kgMLSS/d | 12.69 | 15.35 | 18.35 | 3.51 | 10.02 | 12.35 | 17.25 | 3.77 |
| SVI | cm$^3$/g | 154 | 198 | 291 | 35 | 90 | 138 | 200 | 37 |
| $m_{PIX}$ | m$^3$/d | 0 | 0.81 | 1.75 | 0.27 | 0 | 0.84 | 1.82 | 0.28 |

Table 2
List of calculation results concerning the theoretical distributions matching with empirical data based on the Kolmogorov–Smirnov test for the Sitkówka–Nowiny wastewater treatment plant

| Indicators | Winter | | Autumn, spring, summer | |
|---|---|---|---|---|
| | Distribution | K–S ($p$) | Distribution | K–S ($p$) |
| $Q$ | N | 0.342 | N | 0.245 |
| $BOD_5$ | N | 0.782 | N | 0.569 |
| $N–NH_4$ | N | 0.881 | N | 0.752 |
| TN | N | 0.732 | N | 0.700 |
| TP | N | 0.575 | N | 0.521 |
| $T_K$ | N | 0.054 | N | 0.068 |
| DO | N | 0.156 | N | 0.231 |
| $X_{OC}$ | N | 0.135 | N | 0.156 |
| SVI | LN | 0.608 | LN | 0.52 |
| $m_{PIX}$ | N | 0.254 | N | 0.315 |

### 4.2. Adjustment of theoretical distributions to empirical data

Based on the measurement data and the independent variables determined in the logit model, the conformity of the selected theoretical distribution with the empirical one was determined using the calculated value of the test probability (Kolmogorov–Smirnov test); calculation results are given in Table 2.

On the basis of the data in Table 2 it can be concluded that the distributions of the variability of individual independent variables, concerning both the quality and quantity of wastewater and the operational parameters of the bioreactor, are described in most cases by normal distributions of Gauss. Only for the volumetric index of the sediment, it was found that it has a log-normal distribution. These results are confirmed by the results of the Kolmogorov–Smirnov test, that is, the test probability. The relationships obtained above (Table 2) are confirmed in the works of other authors [9–11] dealing with the analysis of variability of the quality of wastewater flowing into wastewater treatment plants.

### 4.3. Influence of the control strategy on the operational parameters of wastewater treatment plant

Following the developed methodology (Fig. 1), the first step was to simulate the operation of the wastewater treatment plant in terms of sedimentation of sludge in the secondary settling tank, when the control strategy S1 and $R(t) = 1$ was implemented. Based on the simulation calculations, the probability distributions of not exceeding the value of selected technological parameters (DO, $X_{OC}$, $m_{PIX}$) were determined for two variants - including and excluding failures of wastewater quality analyzers; the results of the simulation are presented in Fig. 2–4.

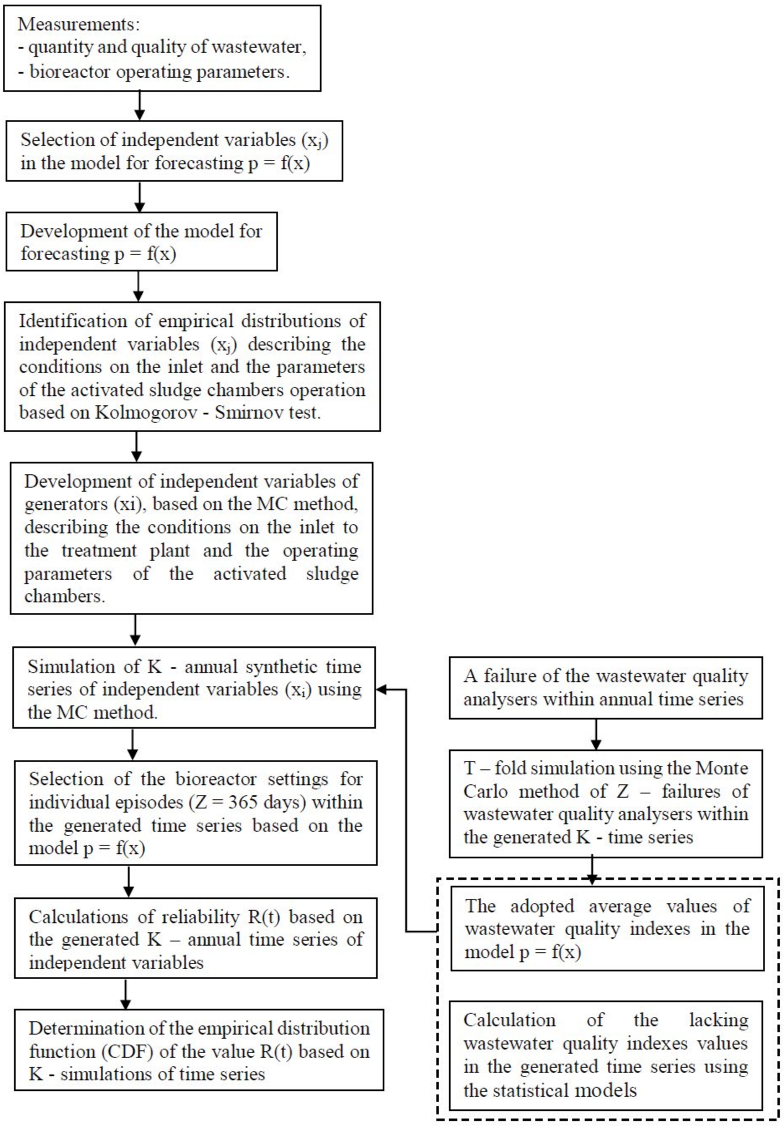

Fig. 1. Calculation method of the probabilistic model for forecasting the reliability of treatment plant operation with reference to sludge sedimentation proposed in the paper.

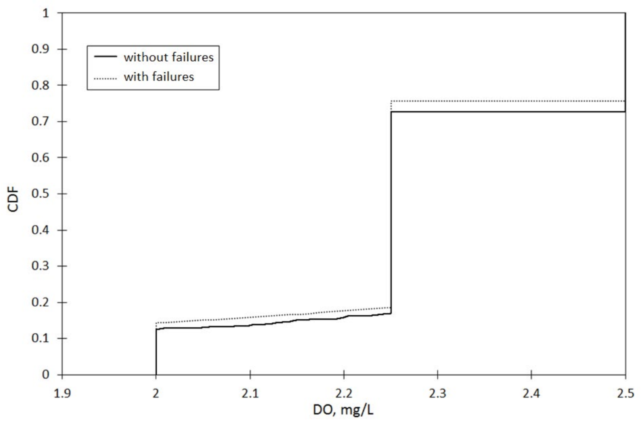

Fig. 2. Empirical distribution of the value DO within a year for the control variant S1 without and with failures in the Sitkówka–Nowiny wastewater treatment plant.

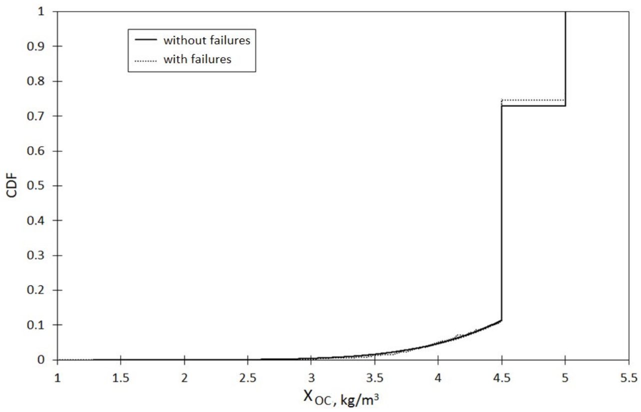

Fig. 3. Empirical distribution of the value $X_{OC}$ within a year for the control variant S1 without and with failures in the Sitkówka–Nowiny wastewater treatment plant.

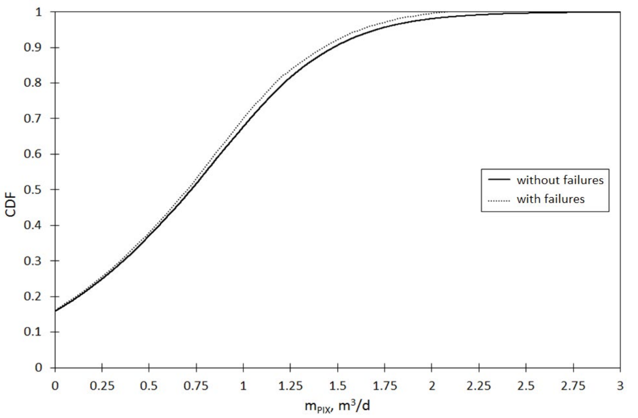

Fig. 4. Empirical distribution of the value $m_{PIX}$ within a year for the control variant S1 without and with failures in the Sitkówka–Nowiny wastewater treatment plant.

Based on the calculation results for the control variant S1, the CDF curves were determined, showing the variability of bioreactor operating parameters ($X_{OC}$, DO and $m_{PIX}$) within a year. Based on them, it was found that within an annual cycle, the calculated concentration of dissolved oxygen and the concentration of activated sludge in the reactor changed within 2.0–2.5 mg/L and 2.5–5.0 kg/m$^3$, respectively.

Obtained results of the analysis are confirmed by the results of continuous simulations conducted by Comas et al. [4], who based their analysis on the fuzzy set theory developed a model for simulation of activated sludge sedimentation and tested its usefulness on the example of a wastewater treatment plant in Catalonia. Based on continuous time series of input variables such as incoming wastewater quantity ($Q$), wastewater quality indicators (BOD, TN, TP) and reactor operating parameters (DO, MLSS), Comas et al. [4] performed sedimentation calculations of the sludge sediment, confirming its seasonal variability SVI, as also proved by Bayo et al. [20]. Based on the obtained simulation results DO ≈ 2.5 mg/L and $X_{OC} \approx 5.0$ kg/m$^3$ (this corresponds to a percentile value of $p = 0.95$) one may assume that for approx. 0.25 of a year, the temperature in the bioreactor decreased below 13.8°C. In turn, for 270 d of a year, the value of DO was below 2.25 mg/L and the value of $X_{OC} \leq 5.0$ kg/m$^3$, which can be identified with the bioreactor operation under dry weather, of which within approx. 2 months the values of DO and $X_{OC}$ changed within the range of 2.00–2.25 mg/L and 2.5–4.5 kg/m$^3$, respectively. The above variability ranges of DO and $X_{OC}$ may indicate that during the period of continuous simulation there were periods when the temperature in the activated sludge chambers was significantly increased, which at some simplifications can be identified with the summer and spring period; then, it is also possible to reduce the value of oxygen concentration in nitrification chambers and probably also the values of $X_{OC}$ (Fig. 3) and $m_{PIX}$ (Fig. 4), eliminating the problems with activated sludge bulking.

The results of the simulation obtained above on the example of the wastewater treatment plant in Sitkówka are confirmed by computer simulations performed by Flores-Alsina et al. [5], who in parallel modeled the quality of wastewater effluent the wastewater treatment plant using the calibrated ASM model, as well as modeled activated sludge sedimentation with the empirical model developed on the basis of data collected at a wastewater treatment plant in southern Spain. The simulations showed that in winter, due to temperature decrease, there is a need to increase the oxygen concentration in the activated sludge chambers in order to limit the adverse impact of activated sludge bulking on the quality of effluent from the wastewater treatment plant. The obtained dependencies were confirmed by the results of tests carried out at a wastewater treatment plant in Catalonia, on the basis of which Comas et al. [4] determined classification models for simulation of activated sludge bulking.

The performed analysis indicate also that besides the values of DO and $X_{OC}$, the volume of dosed coagulants also has a significant impact on the reactor operation stabilization. The average daily volume of dosed PIX is 0.75 m$^3$/d. However, such a significant range of daily coagulant dose variability indicates that in the episodes at the stage of simulations, it was not possible to maintain the value of SVI within a proper range by adequate selection of the values of DO and $X_{OC}$. Therefore, there was a need for coagulant dosing within these episodes.

The performed analysis shows that failures of the wastewater quality analysis are an important aspect that has been neglected in the past in the probabilistic models concerning the wastewater treatment plant operation. The performed simulations of the secondary clarifier operation (Fig. 2–4) indicated that failures of the analyzers lead to an insignificant increase in the control variables ($X_{OC}$, DO and $m_{PIX}$) within an annual cycle. An increase in these values results from the necessity to ensure high reliability of the treatment plant operation $R(t = 365\text{ d}) = 1$ with reference to the activated sludge sedimentation.

### 4.4. *Influence of the control strategy on the reliability of the wastewater treatment plant operation in the long term*

Based on the calculation algorithm (Fig. 1), the first step was to simulate the annual time series of quantities and wastewater quality indicators (BOD, TN, TP, N–$NH_4$) using the Monte Carlo method. Simultaneously, episodes involving vectors (Q, BOD, TN, TP, N–$NH_4$, T] were calculated, for which the occurrence of analyzer failures was modeled. Based on the proposed control strategies (S2–S5), calculations were made on the basis of Eqs. (8)–(10), of curves describing the variability of reliability of wastewater treatment plant operation ($R$) and of its operating parameters in the considered period of time (1 y). The performed long-term simulations of the secondary clarifier operation show that, in the case of the next operation control strategies, its operation reliability $R(t)$ within an annual cycle decreases compared to strategy S1, as shown in Fig. 5. For the considered control strategies (S2–S4), the greatest reliability of the secondary clarifier operation is achieved for the strategy S4 (constant daily dose PIX equal to 1.14 m$^3$). Assuming a stochastic nature of the wastewater quantity and quality at the inlet to the wastewater treatment plant ($BOD_5$/TN, $BOD_5$/TP, $L_{N-NH_4}$) and the temperature in the sludge chambers ($T_K$), it was found that the reliability of the wastewater treatment plant for S4 changes from 0.92 ($p = 0.05$) to 0.96 ($p = 0.95$), and the average is $R(t) = 0.94$ ($p = 0.50$). Lower reliability of the wastewater treatment plant operation is achieved for the control strategy with a variable PIX dose, depending on $T_K$, because of values of $R(t)$ change within the range of 0.76–0.84. The forecast for the Sitkówka–Nowiny wastewater treatment plant operation indicated that when applying the S4 control strategy, in 95% of events during a year, values of $COR_{SVI}$ do not exceed 1.15, which may indicate that the daily dose of PIX is not optimum and there is a need for its proper modification (Fig. 6).

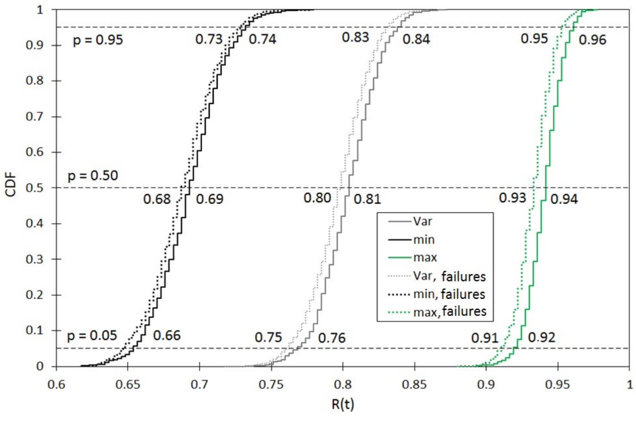

Fig. 5. Empirical distributions of reliability $R(t)$ for the assumed control strategies for a model with and without failures at the Sitkówka–Nowiny wastewater treatment plant.

These results are confirmed by the results of simulations performed with reference to the share (η) of days in a year when the volume of dosed PIX is insufficient (Fig. 7) and when the volume of dosed PIX is excessive (Figs. 6 and 9), which results in a lower value of $COR_{SVI}(t = 1–365\text{ d}) < 1.0$. Based on the determined curves in Fig. 8, one may assess the secondary clarifier optimization strategies, considering the volume of dosed PIX, the insufficiency of which leads to a lower value of $COR_{SVI}(t = 1–365\text{ d}) < 1.0$. For the wastewater treatment plant operation control option S4, conditioning the largest share of days within a year, when $COR_{SVI}(t) < 1.0$, it was found that the annual volume of excessively dosed coagulant among the considered strategies is the greatest and ranges from 95.2 m$^3$/y ($p = 0.05$) to 114.1 m$^3$/y ($p = 0.95$).

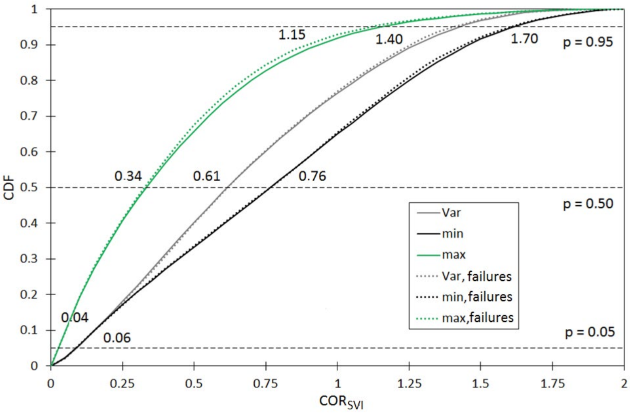

Fig. 6. Empirical distributions of reliability $COR_{SVI}$ for the assumed control strategies for a model with and without failures at the Sitkówka–Nowiny wastewater treatment plant.

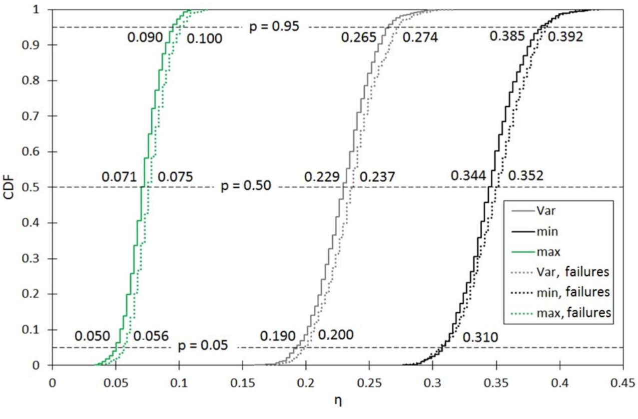

Fig. 7. Empirical distributions of values η for the adopted control strategies for a model with failures and without failures (where η – share of days within a year, when the PIX dose is too low and the sludge bulking occurs) for the Sitkówka–Nowiny wastewater treatment plant.

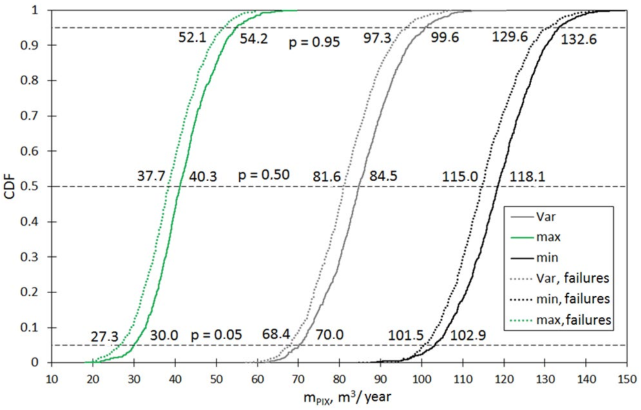

Fig. 8. Empirical distributions of annual PIX dose ($m_{\mathrm{PIX}}$) insufficiency that would eliminate the sludge bulking for the adopted control strategies for the model with and without failures at the Sitkówka–Nowiny wastewater treatment plant.

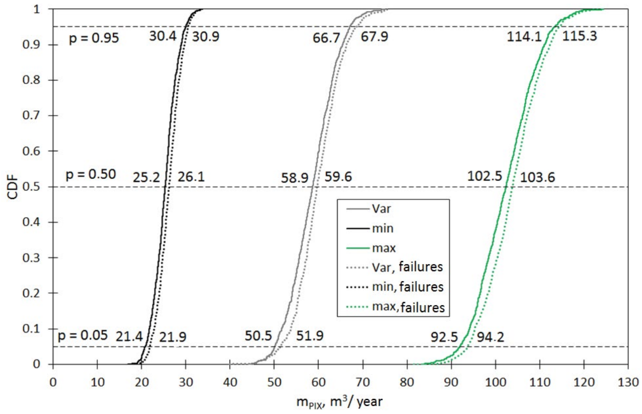

Fig. 9. Empirical distributions of excessive of the yearly PIX dose ($m_{\mathrm{PIX}}$) for the assumed control strategies for the model with and without failures at the Sitkówka–Nowiny wastewater treatment plant.

For the S2 control strategy, the volume of PIX is smaller by more than 40%, as confirmed by the results of $\mathrm{COR}_{\mathrm{SVI}}(t)$ (Fig. 6) and in the empirical distributions of the annual dosage of PIX ($m_{\mathrm{PIX}}$) that would eliminate the sludge bulking.

The performed numerical experiments, based on the Sitkówka–Nowiny wastewater treatment plant, the results of which are presented in Figs. 5–9 confirm the impact of the wastewater quality analyzers failures on the treatment plant operation reliability. The performed analysis confirmed that the failures of wastewater quality analyzers, operating online at the inlet to the wastewater treatment plant, lead to a decrease of the wastewater treatment plant operation reliability. Although the decrease of the secondary clarifier operation reliability $R(t)$ is not significant ($R(t)$ value for individual percentiles decrease by $\Delta R$~0.01), but it may indicate many operational problems, which elimination is not an easy task under a continuous operation system.

The results of calculations obtained above and curves determined on their basis confirm that the more complex the control algorithm is (it takes into account more independent variables), the higher the reliability of wastewater treatment plant operation [1,2]. By introducing further simplifications, and thus reducing the number of variables measured on-line, the reliability of a plant does not necessarily deteriorate, but it may significantly increase its operating costs [4,5,12]. It is not possible to assess the impact of a failure of analyzers on the operation of the plant in relation to the works of other authors, as this aspect has been included for the first time in this paper.

## 5. Conclusions

The paper presents an innovative probabilistic model for the simulation of a wastewater treatment plant operation (with reference to activated sludge sedimentation in the secondary clarifier) that allows assessing the impact of the reactor control strategy on the reliability of its operation. The model assumes that changes in the bioreactor settings can be performed on a continuous basis, allowing assessing their impact on the short-term and long-term reactor operation. This is a significant advantage of the model as compared to papers of other scientists because it allows assessing the impact on the reactor operation within a planned period of time, which so far was neglected in the reliability analysis models.

The innovative advantage of the probabilistic model, which has not been included in the calculation models in the past, is the fact that at the stage of simulation, it includes the option to define wastewater quality analyzer failures at the inlet to the facility. This is especially valuable at the stage of performing calculation experiments from the standpoint of their impact on the facility's short- and long-term operation reliability. Based on this approach, it is also possible to identify the conditions wherein a few analyzers fail at the same time and to develop the variants of the bioreactor settings control in the case of such circumstances.

## References

[1] A.C. Avella, T. Görner, J. Yvon, P. Chappe, P. Guinot-Thomas, P. Donato, A combined approach for a better understanding of wastewater treatment plants operation: statistical analysis of monitoring database and sludge physico-chemical characterization, Water Res., 45 (2011) 981–992.
[2] L. Hongbin, H. Mingzhi, Y. ChangKyoo, A fuzzy neural network-based soft sensor for modeling nutrient removal mechanism in a full-scale wastewater treatment system, Desal. Water Treat., 51 (2014) 6184–6193.
[3] J.F. Canete, P.D. Saz-Orozco, R. Baratti, M. Mulas, A. Ruano, A. Garcia-Cerezo, Soft-sensing estimation of plant effluent concentrations in a biological wastewater treatment plant using an optimal neural network, Expert Syst. Appl., 63 (2016) 8–19.
[4] J. Comas, I.R. Roda, K.V. Gernaey, C. Rosen, U. Jeppsson, M. Poch, Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems, Environ. Modell. Software, 23 (2008) 1250–1261.
[5] X. Flores-Alsina, J. Comas, I.R. Roda, M. Poch, K.V. Gernaey, U. Jeppsson, Evaluation of plant-wide WWTP control strategies including the effects of filamentous bulking sludge, Water Sci. Technol., 60 (2009) 2093–2103.
[6] U. Cortés, M. Martínez, J. Comas, M. Sànchez-Marrè, I. Rodríguez-Roda, A conceptual model to facilitate knowledge sharing for bulking solving in wastewater treatment plant, AI Commun., 16 (2006) 279–289.
[7] B. Béraud, J.P. Steyer, C. Lemoine, E. Latrille, G. Manic, C. Printemps-Vacquier, Towards a global multi objective optimization of wastewater treatment plant based on modeling and genetic algorithms, Water Sci. Technol., 56 (2007) 109–116.

[8] P. Kundu, A. Debsarkar, S. Mukherjee, S. Kumar, Artificial neural network modelling in biological removal of organic carbon and nitrogen for the treatment of slaughterhouse wastewater in a batch reactor, Environ. Technol., 35 (2014) 1296–1306.

[9] J.F. McCormick, B. Johnson, A. Turner, Analyzing Risk in Wastewater Process Design: Using Monte Carlo Simulation to Move Beyond Conventional Design Methods, Proc. WEFTEC, San Diego, 2007.

[10] D. Bixio, R. Carrette, I. Boonen, P. van Hauwermeiren, C. Thoeye, P. Ockier, Safeguard Your Investments for Complying with Stricter Limits - an Effective Tailor-made Plan, In: Proc. 1st IWA World Congress, Paris (France), 4–7 Jul 2000.

[11] D. Messaoud, A. Bachir, M. Maurice, Determination and analysis of daily reliability level of municipal wastewater treatment plant, Courrier du Savoir, 17 (2013) 39–46.

[12] M. Taheriyoun, S. Moradinejad, Reliability analysis of a wastewater treatment plant using fault tree analysis and Monte Carlo simulation, Environ. Monit. Assess., 187 (2015) 1–13.

[13] A. Asadi, A. Verma, K. Yang, Wastewater treatment aeration process optimization: a data mining approach, J. Environ. Manage., 203 (2016) 1–10.

[14] M. Ebrahimi, E.L. Gerber, T.D. Rockaway, Temporal performance assessment of wastewater treatment plants by using multivariate statistical analysis, J. Environ. Manage., 193 (2017) 234–246.

[15] G. Fu, D. Butler, S.T. Khu, S. Sun, Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory, Water Resour. Res., 47 (2011) 1–13.

[16] B. Szeląg, K. Barbusiński, J. Studziński, Activated sludge process modeling using selected machine learning techniques, Desal. Water Treat., 117 (2018) 78–87.

[17] B. Szeląg, K. Barbusiński, J. Studziński, Application of the model of sludge volume index forecasting to assess reliability and improvement of wastewater treatment plant operating conditions, Desal. Water Treat., 140 (2019) 132–143.

[18] B. Szeląg, Ł. Bąk, R. Suligowski, J. Górski, Statistical models to predict discharge overflow, Water Sci. Technol., 78 (2018) 1208–1218.

[19] S.C. Bagley, H. White, B.A. Golomb, Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain, J. Clin. Epidemiol., 54 (2001) 979–985.

[20] J. Bayo, J.M. Angosto, J. Serrano-Aniorte, Evaluation of physicochemical parameters influencing bulking episodes in a municipal wastewater treatment plant, Water Pollut. VIII: Model. Monit. Manage., 95 (2006) 531–542.

[21] B. Szeląg, P. Siwicki, in: B. Kaźmierczak, M. Kutyłowska, K. Piekarska, A. Trusz – Zdybek, Application of the Selected Classification Models to the Analysis of the Settling Capacity of the Activated Sludge – Case Study, E3S Web of Conferences 17, Boguszów-Gorce, 2017, pp. 1–7.

[22] S.C. Oliveira, M. Sperling, Reliability analysis of wastewater treatment plants, Water Res., 42 (2008) 1182–1194.

[23] F.C. Wu, Y.P. Tsang, Second-order Monte Carlo uncertainty/variability analysis using correlated model parameters: application to salmonid embryo survival risk assessment, Ecol. Modell., 177 (2004) 393–414.

[24] B. Bacchi, M. Balistrocchi, G. Grossi, Proposal of a semiprobabilistic approach for storage facility design, Urban Water J., 5 (2008) 195–208

[25] I. Lou, Y. Zhao, Sludge bulking prediction using principle component regression and artificial neural network, Math. Probl. Eng., 2012 (2012) 1–17.

[26] K. Barbusiński, H. Kościelniak, Influence of substrate loading intensity on floc size in activated sludge process, Water Res., 29 (1995) 1703–1710.

[27] A.M.P. Martins, J.J. Heijnen, M.C.M. van Loosdrecht, Bulking sludge in biological nutrient removal systems, Biotechnol. Bioeng., 86 (2004) 125–135.

[28] E. Kowalska, E. Paturej, M. Zielińska, Use of Lecane inermis for control of sludge bulking caused by the Haliscomenobacter genus, Desal. Water Treat., 57 (2016) 10916–10923.

[29] E. Bezak – Mazur, R. Stoinska, B. Szeląg, Evaluation of the impact of operational parameters and particular filamentous bacteria on activated sludge volume index - a case study, Annu. Set Environ. Prot., 18 (2016) 480–491.