LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
LUT School of Engineering Science
Degree Program of Chemical Engineering

Master's Thesis
2018

Rostislav Konoplin

IDENTIFICATION AND PREDICTIVE CAPABILITIES OF IBM WATSON IN INDUSTRIAL WASTEWATER TREATMENT

Examiners: Professor Tuomas Koiranen
Tech.Lic Esko Lahdenperä
Supervisor: Professor Tuomas Koiranen

ABSTRACT

Lappeenranta University of Technology

LUT School of Engineering Science

Degree Program of Chemical Engineering

Rostislav Konoplin

Identification and predictive capabilities of IBM Watson in industrial wastewater treatment

Master's Thesis

2018

67 pages, 42 figures, 13 tables and 7 appendices

Examiners: Professor Tuomas Koiranen

Tech.Lic Esko Lahdenperä

Supervisor: Professor Tuomas Koiranen

Key-words: wastewater treatment, predictive analysis, big data processing.

With current rate of development of technologies and researches in different spheres it may be challenging for scientists to be kept up-to-date on. One of the main reasons is that only small part of this information is suitable for coherent analysis, integration and classification for end users. The issue lies in deployment and extraction of data from many sources and conversion into suitable formats.

For efficient wastewater treatment it is absolutely crucial to achieve precise control and management of operating parameters. Temperature, flowrates, concentration of inlet compounds and amount of added chemicals may cause significant fluctuations of the wastewater purification efficiency. Some of the impurities promote the growth of the hazardous bacteria and pathogens, which are the "top priority" for removing during the wastewater treatment processes. In spite of the fact that all purification processes are tightly-controlled, the concentration of inlet compounds may fluctuate heavily and does not correlate with processing parameters, and as a result, it may affect the treatment efficiency.

New technologies for efficient data analysis allow to process and evaluate big datasets according to user's request. These systems are able to understand and analyze different types of data, such as laboratory results, industry and technical-specific content (chemistry,

pharmaceuticals, medicine, economy), compare the scientific results, summarize numeric values through the instrumentality of predictive modeling and statistical algorithms.

One of the potential solution is IBM Watson - a question-answering system, which uses combination different predictive algorithms. In question-answering system the question may be defined as an initial input in data processing. Program extracts key elements of the questions and provides precise answers according to the user's input. IBM Watson Analytics is a cloud-based system, which is used for data analysis and pattern recognition. This tool is suitable for patterns and dependencies searching in data results and determination of the statistical drivers of analyzed values. Application of IBM Watson Analytics for analysis of wastewater purification process related data provides an opportunity to effectively investigate and evaluate operating parameters thus enabling in-depth analysis of the process. Manual estimation of the most significant parameters is time-consuming and, in some cases, impossible due to the high amount of data and large quantity of variables. Also, non-controlled inlet parameters (such as a concentration of impurities in wastewater) may be taken into account as affecting factors for purification processes.

Table of contents

List of Figures

List of Tables

LITERATURE REVIEW

1 Wastewater treatment

Wastewater treatment process may be defined as a process for cleaning and reusing wastewater (no longer needed and suitable for use), which after is discharged back into the water storages. Wastewater streams are formed by the following ways: municipal wastewater (bathing, washing, sewage, and rainwater), agriculture and industrial wastes. Depending on the origin, wastewater streams may contain bacteria, chemicals and other toxins (Kukreja 2009). Primary purpose of treatment is to reduce all harmful contaminants to appropriate levels and corresponding environmental requirements for municipal and industrial water. Figure 1 represent sources of wastewater streams.

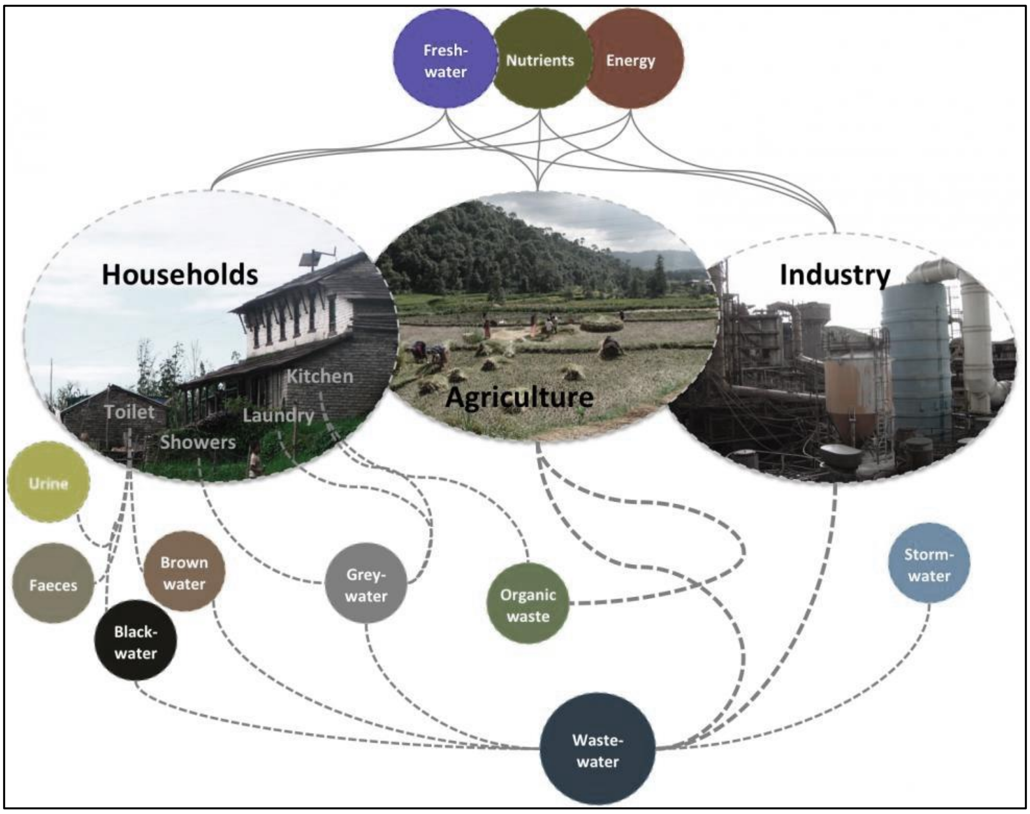


Figure 1 - Wastewater sources

The description of the wastewater is presented in Table 1 (de Graaff 2010).

Table 1- Wastewater specification

Type of water	Source	Harmful constituents
Brown and black	Domestic sewage	Pathogens and nutrients
Grey	Households (kitchen, laundry)	Soap particles, fats and oils

There are several wastewater treatment plants which are designed for cleaning relevant impurities (UNEP 2013):

- Physical and chemical treatment plant – uses physical processes together with chemical reactions. It is meant for cleaning industrial and manufacture wastewater streams due to high concentration of chemicals and toxins, which are very harmful for the environment.
- Biological wastewater treatment plant - uses biological matter and bacteria processing to demolish waste matter and detach impurities. It is meant for cleaning municipal and business premises wastewater.(Kukreja 2009)

1.2 Description of wastewater treatment process

First step of water treatment process consists of accumulation of wastewater streams in lifting stations **(1)**. The control for collection of wastewater is conducted by municipal administration, households or industries. Collected water streams are preliminary treated with chemicals. At this stage odor control also is conducted since odor threshold of cleaned water is a significant parameter. Therefore, the odor control is the initial step in water treatment. After pretreatment step separation or screening process **(2)** is conducted where solid compounds larger than 6 mm (nappies, cotton buds, plastics, diapers, rags, personal sanitary items, face wipes, pieces of glass or bottle tops) are removed (Kukreja 2009). All this removed waste is directed to further processing or buried, depending on its toxicity and water treatment plant capabilities. If these steps are not conducted, large impurities may damage the equipment in further processing. During next sedimentation step **(3)** the heavy residual compounds and macrobiotic solid matter (UNEP 2013) gravitate to the bottom of settling basins, wherefrom are removed and used as fertilizers. The remaining water after sedimentation is pumped in bioreactor **(4)**, where bacteria demolish harmful matter and purify the water. In order to achieve rapid growth of bacteria, which consumes remaining organic matter and oxygen (as a growth source), the air is supplied into a bioreactor and mixed with wastewater stream and seed sludge. (UNEP 2013). Overall time of this process is 3-6 hours.

In clarification process **(5)** bacteria continue to purify water stream until corresponding environmental requirements are achieved. After primary treatment processes cleaned water still contains diseases causing organisms. Therefore, chlorine treatment **(6)** is designed for removing bacteria from water, that is being disinfected for 20-25 minutes in tanks. The treated water passes to weir, basin or as circulating water return to industrial facility. Figure 2 represents the wastewater treatment plant operation concept.

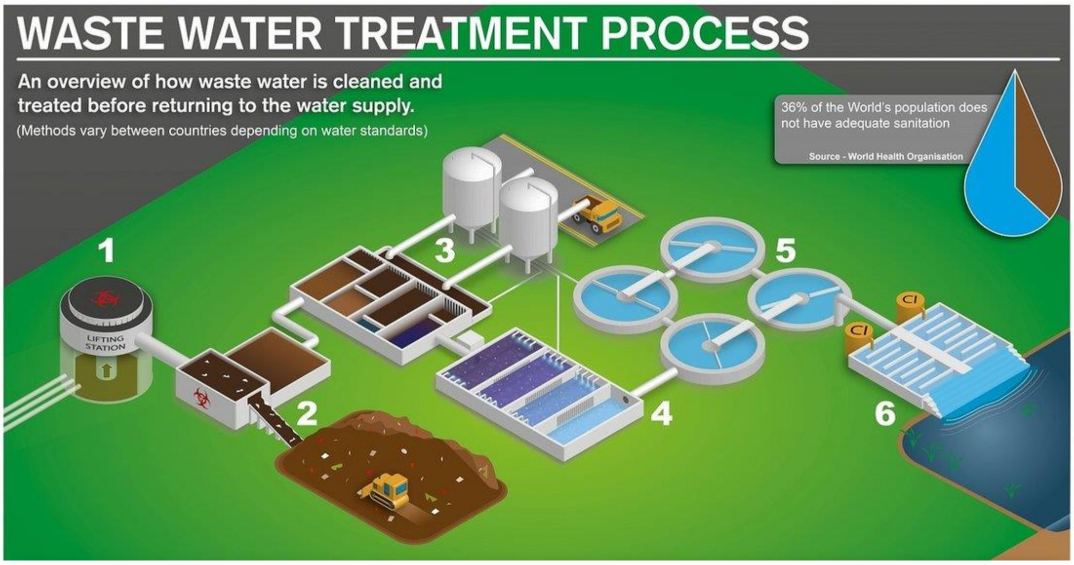


Figure 2 - Wastewater treatment process flowchart (Scutt 2012) 1 – Collection and pretreatment of wastewater; 2 - The separation process; 3 – Sedimentation; 4- Bioreactor, 5- The clarification process, 6- Final chemical treatment with chlorine.

1.3 Wastewater constituents

Current wastewater treatment processes are classified depending on constituent type. For high efficient wastewater treatment optimal process should be selected. The variety of water treatment operations is represented in Table 2 (Tchobanoglous et al. 2003).

Table 2 - Processes of constituent's removal

№	Constituent	Average concentration in domestic water, mg/L	Average concentration in industrial water, mg/L	Operation or process	Chapter
1	Suspended solids	210	200-1500 (Nasr, Doma & Abdel-Halim 2007)	•Flotation •Grit removal •Sedimentation •High-rate clarification •Screening •Chemical precipitation •Depth and surface filtration	1.3
2	Biodegradable organics	100	100-900	• Aerobic suspended growth •Aerobic attached growth •Anaerobic suspended growth •Anaerobic attached growth •Lagoon variations •Physical-Chemical systems •Chemical oxidation •Advanced oxidation •Membrane filtration	1.4

3	Nutrients			• Chemical oxidation •Suspended-growth nitrification and denitrifications •Fixed-film nitrification and denitrifications •Air stripping •Ion exchange	1.5
	Nitrogen	20-40	(organic) 19		
	Phosphorus	7	9	•Chemical treatment •Biological phosphorus removal	1.5
4	Pathogens			• Chlorine compounds • Chlorine dioxide • Ozone • Ultraviolet radiation	1.4,1.9
5	Colloidal and dissolved solids	500	500-1400	• Membranes • Chemical treatment • Carbon adsorption • Ion exchange	1.3-1.6
6	Volatile organic compounds	100-400	100-900	• Air stripping • Carbon absorption • Advances oxidation	1.4-1.6
7	Odors	-	-	•Chemical scrubbers •Carbon adsorption •Biofilters •Compost filters	1.6,1.7

1.4 Processes for removing suspended solids

Suspended solids may be defined as particles which are remaining in suspended form in water either because the density of these particles lower or the same as water. The content of suspended solids in water is one of the most important environmental requirements. There are several methods of removing of these constituents (Table 2).

1.4.1 Screening

Screening is one of the initial unit operations in wastewater treatment process. The main purpose of this step is to remove coarse particles from water streams since impurities may damage process equipment and decrease treatment efficiency. The elements of the screens consist of perforated surfaces, wedge wire elements and different wire cloth compositions (Tchobanoglous et al. 2003).

The different types of screening are represented in Table 3. Coarse screens are used for cleaning particles with diameter ranging from 6 mm to 150 mm. Fine screens were designed for removing particles of diameter less than 6 mm. Both these types of screens are used in preliminary wastewater treatment processes. Microscreens were designed for removing parts with diameter less than 5 µm during intermediate stages of wastewater treatment processes.

Table 3- Classification of screening processes (Tchobanoglous et al. 2003)

Type of screening		Advantages	Drawbacks	Size of removing particles, mm
Mechanically cleaned:	- Chain-driver	Use for heavy – duty applications	Before maintenance, submerged parts need to be dewatered	6-150
	- Reciprocating rake	No submerged moving parts; Handle large objects	Emergency water level increasing may cause motor burnout	
	- Catenary	Sprockets are not submerged Handle large objects	Heavy chain, which difficult to handle; Open design may emit odors	
	- Continuous bet	Units difficult to jam	High operation costs	0.5-30.0
Fine screening	- Static wedge wire	Suitable for small industrial plants	Must be cleaned once or twice per day (hot water)	0.2-1.2
	- Drum	Variety of implementation in processes	structural complexity	<6
	- Step	Can be also use for removal of solids from septage, primary sludge, biosolids	Maintenance of moving part and plates	3-6
	Microscreenig	Remove small sizes particles	Incomplete solids removal (average 55 %)	0.001

1.4.2 Grit removal

Grit composite consists of gravel, sands or different types of cinder. It also may include solid organic matter like shells, seeds, animal bone ships etc. (X. Wang 2015).

Removing this type of constituents may be achieved in grit chambers or by centrifugal solid separation. Grit chambers were designed for protection of equipment moving parts and minimization of composites formation in pipelines (Tchobanoglous et al. 2003). Technical design of a grit chamber depends on grit compositions and particle sizes. Grit chamber classification is provided in Table 4. Currently, four types of grit chambers exist: rectangular and square horizontal-flow grit chambers, Aerated grit chambers and Vortex-type grit chambers.

Table 4- Classification of grit removal processes

Type of grit chamber	Description	Removed particles	Removal rates, %
Rectangular Horizontal-Flow	Conveyor with scrapers, buckets and plows	0.21 mm-diameter, organic	-
Square Horizontal-Flow	Solids removed by rotating and collected in accumulation tanks	0.15 mm-diameter organic matter and concentrate grit	95
Aerated	Air creates spiral flow pattern and separates heavy grit and light organic particles	0.21 mm-diameter or larger	≈100
Vortex-Type	Vortex flow pattern removed the grid from the water, gravitas and centrifugal forces separate the grit	0.33 mm 0.24 mm 0.15 mm	92-98 80-90 60-70

1.4.3 Sedimentation

The sedimentation process was designed for removing of settleable solids and floated matter. The desired outcome of sedimentation process is removal of 50-70% total suspended solids from wastewater stream and disposal of organic matter. Selection of sedimentation equipment depends on plant scale and local wastewater composition. All current plants provide several sedimentations tanks in order to maintained overall continuous wastewater treatment processes. The classification of sedimentation processes is presented in Table 5 (Tchobanoglous et al. 2003).

Table 5 - Classification of sedimentation processes

Sedimentation equipment	Particles	Description	Advantages
Rectangular tanks (X. L. Wang et al. 2010)	Combination of two or more tanks allow to separate different sizes varieties of particles	Particles in water stream hold by baffles, then collected and removed (Figure 1.3)	Simple structure, flexible during operating conditions
Circular tanks		Water stream enter in the center of the tank	Large-scale implementation treatment plant
Combination Flocculator-Clarifier	Industrial wastewater streams with biosolids impurities	Flow of the wastewater charges in the tank center and are being mixed with flocculants by low-speed mixer.	Suitable for industrial water treatment with bioorganic solids

The operating concept of rectangular and circular tanks are represented in Figures 3 and 4.

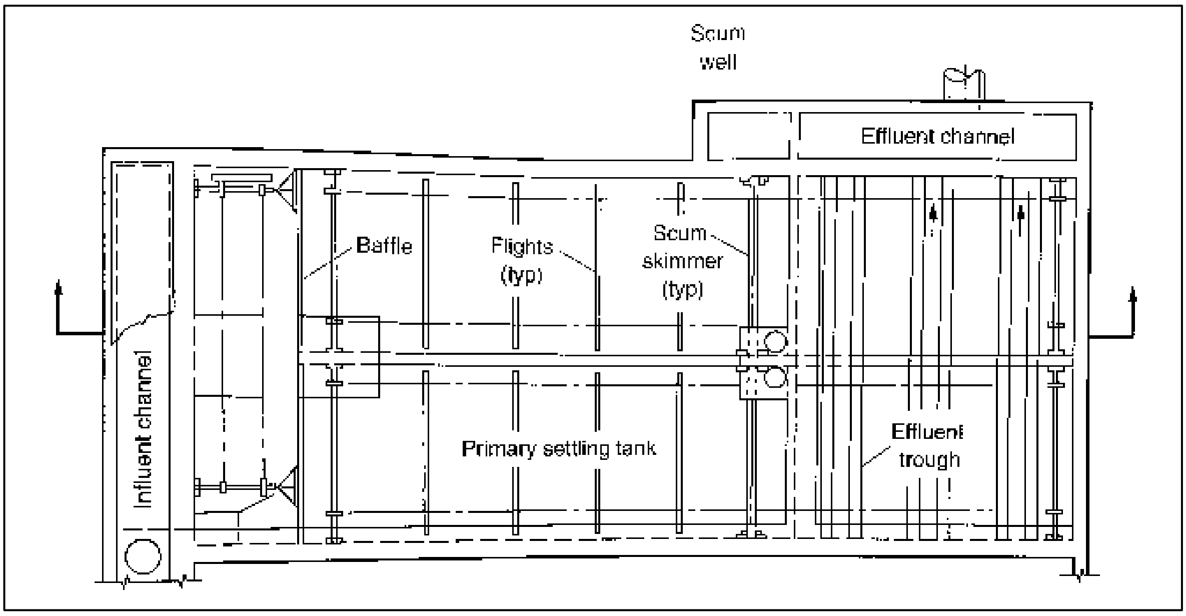


Figure 3- Rectangular sedimentation tank (Tchobanoglous et al. 2003)

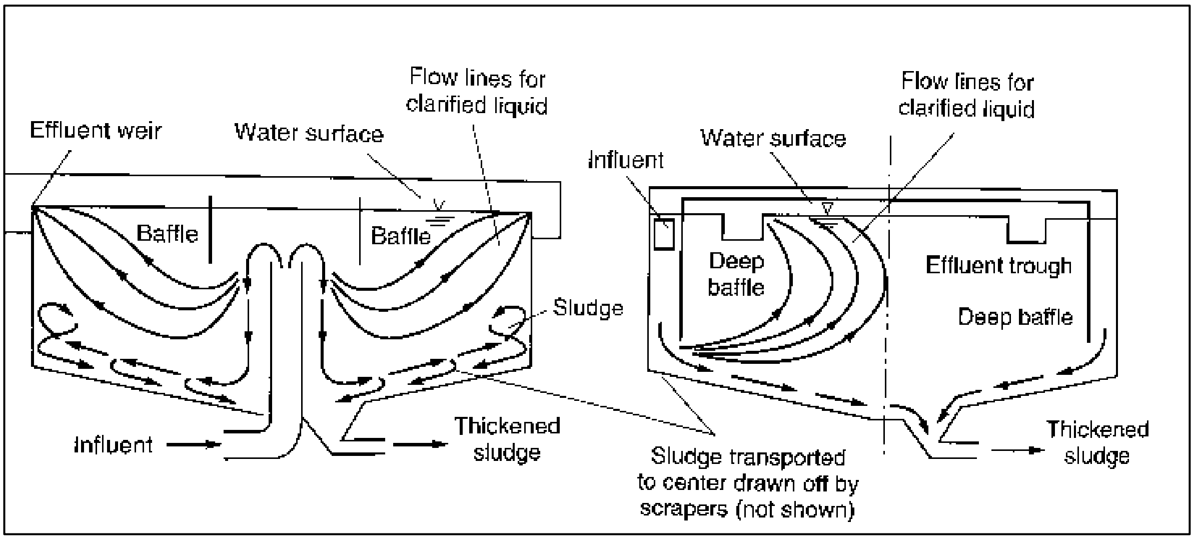


Figure 4 - Circular sedimentation tank (Tchobanoglous et al. 2003)

1.4.4 High-rate clarification

This type of process uses chemical and physical treatment and combination of flocculation and sedimentation systems to improve the settling process. Units of high-rate clarification process are compact; therefore, this process is easy to implement and maintain. The classification of clarification processes is presented in Table 6 (Tchobanoglous et al. 2003, Russell 2006).

Table 6 - Classification of clarification processes

Clarification processes	Particles	Description	Flowrates, $m^3/m^2 \cdot d$	Removing of total suspended solids, %
Ballasted flocculation	Particles sizes 80-100 µm	Chemical coagulant provides inoculating for floc formation; Floc is dense and settles rapidly; High efficiency in small tank volume	Low- 1200-2900	70-90
			Medium – 1800-3500	40-80
			High - 2300-4100	30-80
Lamella plate clarification	Particles sizes 80-100 µm	Use chemical additions, three-stage flocculation increase floc formation; High efficiency in small tank volume	Low- 880	60-70
			Medium - 1200	65-75
			High - 1800	40-50
Dense sludge	Light particles, like algae	The recycled solids are return to water streams and accelerate settling process, no additives are needed	Low- 2300	80-90
			Medium - 2900	70-80
			High - 3500	70-80

1.4.5 Flotation

Main purpose of this process is separation solids and liquid constituents from liquid phase. Fine gas bubbles flow through liquid phase, attach impurities and as a result of buoyant forces raise these particles to the surface. Current flotation units at wastewater treatment plants use air as flotation agent (Russell 2006). The main advantage of flotation process over other sedimentation processes is possibility to remove more efficient and rapidly very small and light particles.

Currently, at wastewater treatment plant two type of flotation processes are used: dissolved-air flotation and dispersed-air flotation. In the first process, air under pressure of several atmospheres dissolves in wastewater streams, and after relief of pressure to atmosphere conditions, raises impurities to the surface of a flotation tank. Figure 5 represents operation concept of dissolved-air flotation unit. Pressure-types units designed mainly for industrials wastewater treatment.

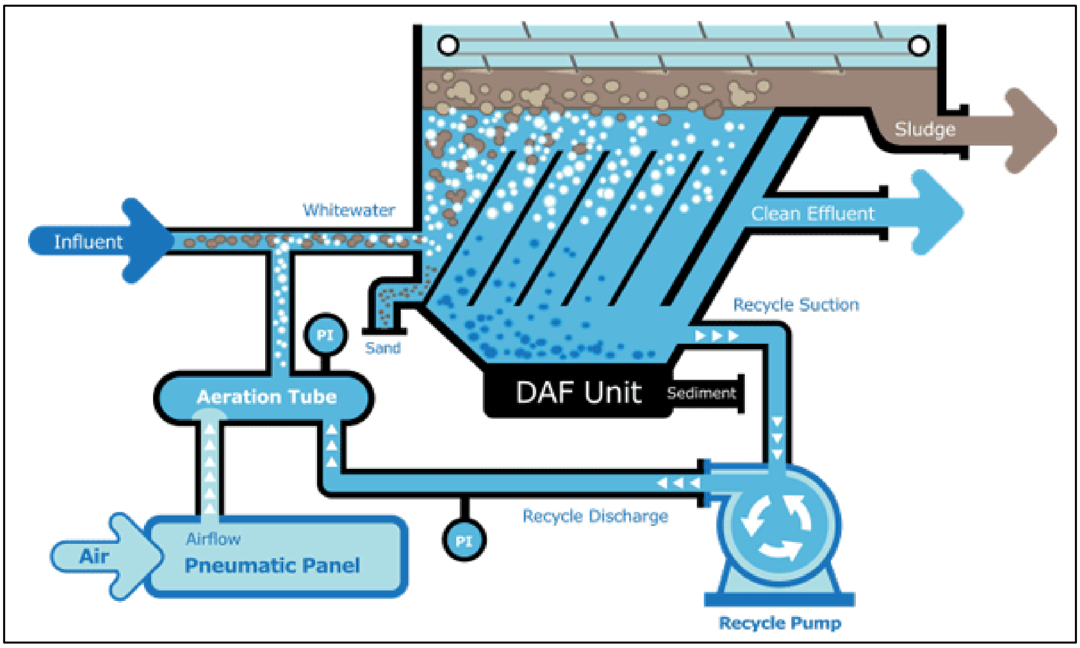


Figure 5- Dissolved-air flotation process ((FRC 2018))

Dispersed-air flotation units commonly are used in industrial water streams and seldom in municipal wastewater treatment. During this process, air bubbles are formed by direct injection in liquid phase from impeller, which is located in the middle of the tank. The Figure 6 describes the Dispersed-air flotation operating principle.

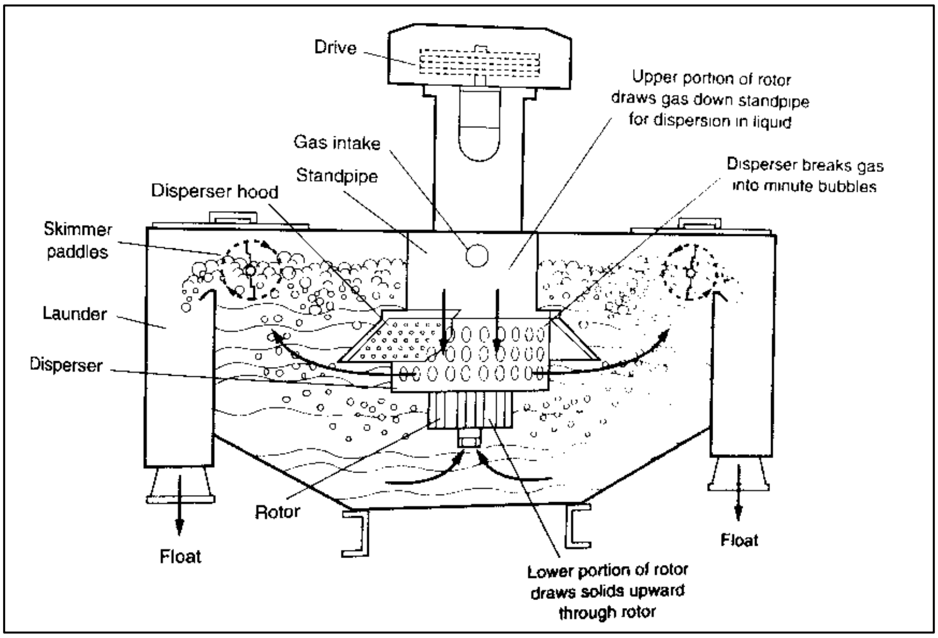


Figure 6- Dispersed-air flotation process (Tchobanoglous et al. 2003)

1.4.6 Chemical precipitation

Chemical precipitation is a method for improving water quality, which is used for increasing efficiency of primary settling facilities, or as a basic stage in physical-chemical processes and phosphorus or heavy metal removal. During this process, chemical components are added in wastewater streams and they precipitate required constituents. Table 7 represents types of impurities and chemicals, which remove them (IAEA 1992).

Table 7- Chemical precipitation

Method	Removing component	
	Calcium or Magnesium	Phosphorus
Al	$3Ca(HCO_3)_2+Al_2(SO_4)_3 \cdot 18H_2O \rightarrow$ **$2Al(OH)_3$**$\downarrow + 3CaSO_4 + CO_2\uparrow + 18H_2O$ $Al(OH)_3$ removed suspended material in water stream	$Al^{3+}+H_nPO_n^{3-n}\leftrightarrow$ **$AlPO_4$**$\downarrow +nH^+$
Lime	$Ca(HCO_3)_2 + Ca(OH)_2 \leftrightarrow$ **$3CaCO_3$**$\downarrow + 2H_2O$	-
$FeCl_3$	$2FeCl_3 + 3Ca(HCO_3)_2 \cdot \leftrightarrow$ **$2Fe(OH)_3$**$\downarrow + 3CaCl_2 + 6CO_2\uparrow$	$Fe^{3+}+H_nPO_n^{3-n}\leftrightarrow$ **$FePO_4$**$\downarrow +nH^+$
$FeSO_4$	1-s stage: $FeSO_4 \cdot 7H_2O + Ca(HCO_3)_2 \cdot \leftrightarrow 2Fe(HCO_3)_2 + CaSO_4 + 7H_2O$ 2-d stage: $Fe(HCO_3)_2 + Ca(OH)_2 \cdot \leftrightarrow$ **$Fe(OH)_2$**$\downarrow + CaCO_3 + 2H_2O$ 2-stage process, since only $Fe(OH)_2$ are precipitate	-

1.4.7 Depth filtration

Depth filtration is defined as a cluster of porous layers with different depth sizes, which are used to capture the suspended solids from the liquid phase. Depth filtration is one of the initial operations in wastewater treatment processes or it may be integrated as an intermediate stage. Due to the asymmetric shapes of the filtration medium, the particles are retained within the filter structure. The main advantage of depth filters is their ability to restrain a high quantity of particles without decreasing the separation efficiency. Depth filters are commonly characterized by the sand filter and are suitable for application with significantly higher filter rates compared with other units. (Purchas, Sutherland 2008). The scheme of depth filtration is presented in Figure 7.

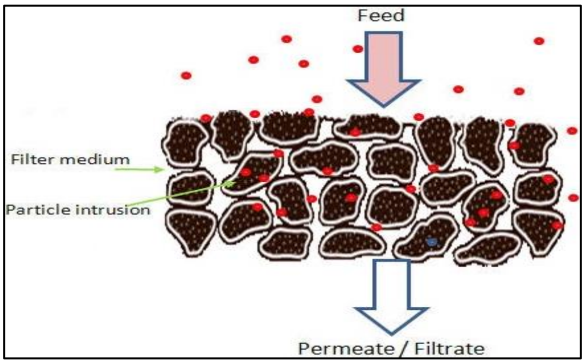


Figure 7- Process of depth filtration description (Yoon 2015)

1.4.8 Surface filtration

In surface filtration process, constituents are retained on the filter surface and do not intrude into the filter medium. The material of filter septum includes cloth or metal fabrics and synthetic materials. The cloth fabrics membrane filters have opening in sizes of pores 0.01-.0.03 mm, and membranes filters have 0.0001-1.0 μm of pore sizes (Tchobanoglous et al. 2003).

This type of filtration was designed for removing residual suspended solids from stabilized water streams and from secondary effluents. The scheme of surface filtration is presented in Figure 8.

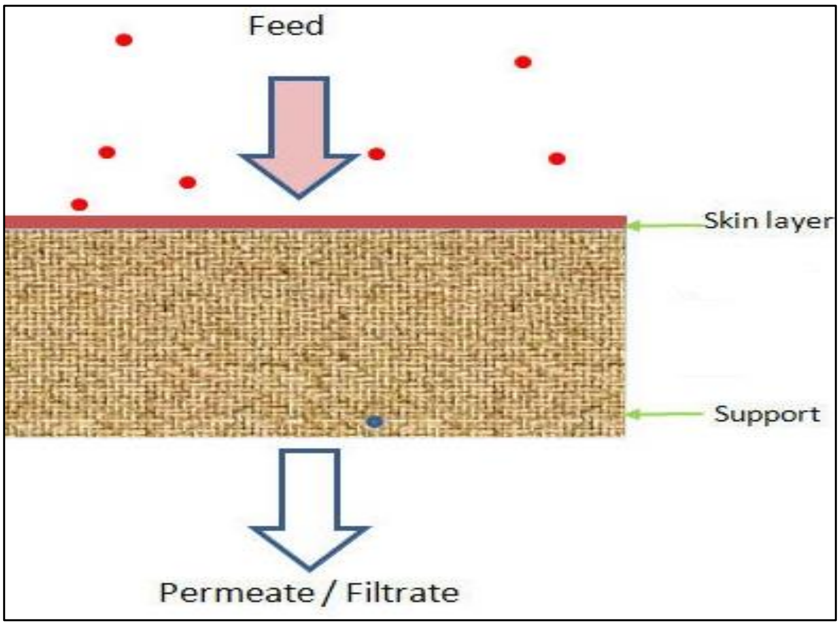


Figure 8- Process of surface filtration description (Yoon 2015)

1.5 Processes and technologies for biodegradable organics removal

Biodegradable wastes are the compounds, that mostly consist of organic matter, which may be demolished to carbon dioxide (CO_2), water (H_2O), methane (CH_4) or low molecular weight organics by means of micro-organisms using following approaches: composting, aerobic and anaerobic digestion and other similar processes. In wastewater treatment processes it may also include non-organic materials which can be decomposed by different bacteria, such as gypsum, plastic matter, sulfates.

1.5.1 Lagoon variations

The aerated lagoon is a wastewater treatment system, which is a reservoir with artificial aeration for stimulation of biological oxidation of treated streams (Ratchford 2003). Oxygen is supplied in order to reduce of pollutants and keep biosolids in suspended form. The overview diagram of lagoons processes is presented in Figure 9.

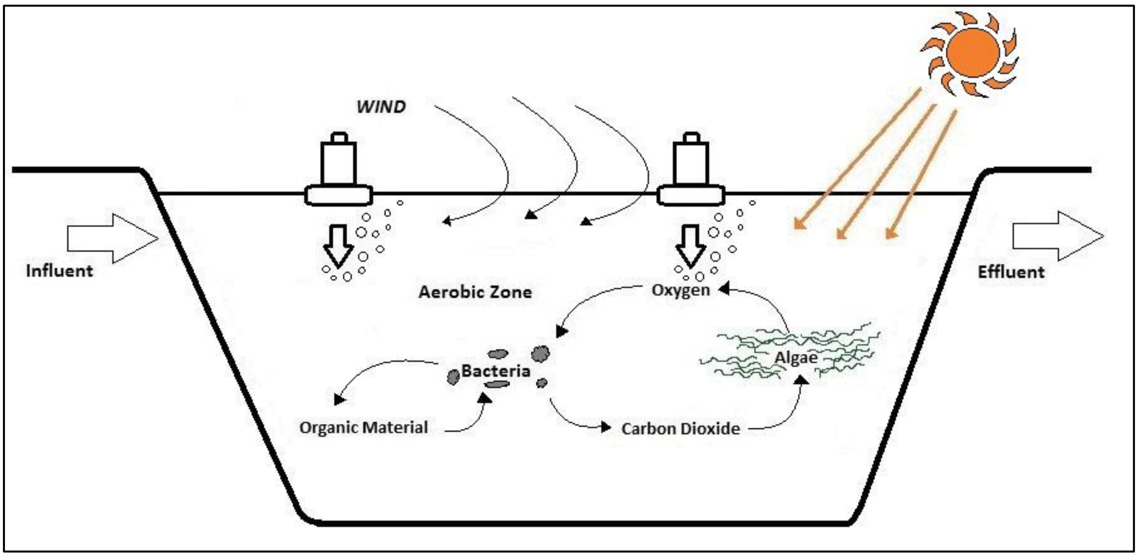


Figure 9-Overview diagram of lagoons processes (Weyenberg et al. 2013)

The classification of suspended growth lagoon processes depends on solids characteristics in the wastewater streams and presented in Table 8 (Tchobanoglous et al. 2003).

Table 8-Classification of Suspended growth lagoon processes

Type of process	Retention time, days	Particles	Environmental considerations
Facultative partially mixed	4-10	Energy input in this process sufficient for biological treatment of pathogens and ammonia, but insufficient for solid suspension,	No nutrients formation
Aerobic flow through with partial mixing	3-6	Designed for stabilizing organic matter, but insufficient for overall solid suspension	Nutrients may aggregate
Aerobic with solids recycle	0.25-2.00	Designed for maintain solids in suspension form	Nutrients formation

1.5.2 Chemical oxidation

Chemical oxidation process in water treatment is defined as a process which involves using the strong oxidizing agent, such as:

- Ozone O_3;
- Permanganate ion $(MnO_4)^-$;
- Hydrogen peroxide (H_2O_2);
- Chloride dioxide ClO^2;
- Chlorine (HClO);
- Oxygen (O_2).

Each chemical reacts with corresponding constituent in water treatment and transformed it for efficient removal. Chemical oxidation processes were designed for organic components concentration, bacteria reduction and odor controls. There are several applications of chemical oxidations processes, which are presented in Table 9 (Tchobanoglous et al. 2003).

Table 9-Applications of chemical oxidation processes in wastewater treatment

Application	Chemical	Description
Slime-growth control	Cl_2, H_2O_2	Fungi and Slime production control
Corrosion control	Cl_2, H_2O_2, O_3	Reduce oxidation by H_2S
Odor control	Cl_2, H_2O_2, O_3	Odor control at pump stations
Grease removal	Cl_2	Preaeration stage
$FeSO_4$ oxidation	Cl_2, O_3	Production of $FeCl_3$
Filter-ponding control	Cl_2	Remove residuals at filter nozzles
Digester foaming control	Cl_2	Prevent foaming formation
Ammonia oxidation	Cl_2	Convert ammonia to $N_2\uparrow$
Bacteria reduction	Cl_2, H_2O_2, O_3	Biodestruction

1.5.3 Advanced oxidation

During the last years, advanced oxidation processes have been intensively studied (Lapertot 2008):

- the biological treatment of effluents from impurities, with subsequent increase of the organics biodegradability for next step of biological treatment;
- organic constituents removal;
- to remove organics compounds as a post-treatment unit for other technologies and minimize the toxicity values.

The advanced oxidation processes may be defined as processes with the highly reactive $HO^{\bullet}$ radicals presence, which are suitable for a rapid and indiscriminate reaction with organic compounds, as a result achieving approximately 100 % mineralization. There are several current processes that can be used for hydroxyl radical generation, which are powerful oxidizing species. Among all these processes the photocatalytic oxidation by titanium dioxide (TiO_2) is a promising chemical procedure for environmental treatment. Titanium dioxide is widely used as a photocatalyst, since it is not toxic, has high photochemical stability and low cost (Wu 2009).

Current processes of wastewater advanced oxidation are described below:

Another efficient process in advanced oxidation is Ozone (O_3)/ultraviolet process. Production of $HO^{\bullet}$ radicals are initialized by reaction of Ozone and Ultraviolet radiation (UV) according to the following transformation chain:

$$O_3+UV \rightarrow O_2+O' \quad (1)$$

$$O'+H_2O \rightarrow HO^{\bullet} + HO^{\bullet} \text{(in wet air)} \quad (2)$$

$$O' + H_2O \rightarrow HO^{\bullet} + HO^{\bullet} \rightarrow H_2O_2 \text{ (in water)} \quad (3)$$

Where O' – exited oxygen atom

Since formation of hydroxyl radicals in water caused the molecule of hydrogen peroxide formation, therefore the use of ozone is not cost-effective comparing with other processes in this sphere (Tchobanoglous et al. 2003). This method is designed for treatment compounds, which are able to absorb ultraviolet radiation.

Otherwise, for compounds which are not able to absorb ultraviolet radiation Ozone/Hydrogen peroxide method is best-performing. The example of such chemicals may be trichloroethylene and perchloroethylene. The overall reaction for hydroxyl radical production can be described as follows: (Karimi 1997)

$$H_2O_2 + O_3 \rightarrow HO^{\bullet}\ HO^{\bullet} + O_2 \quad (4)$$

In terms of after treatment methods Hydrogen peroxide/UV method is frequently used. More recently, this method was applied for tracing constituents in treated water, since H_2O_2 has a small molar extinction coefficient. The overall reaction can be described as follows (Tchobanoglous et al. 2003):

$$H_2O_2 + UV \rightarrow HO^{\bullet}\ HO^{\bullet} \quad (5)$$

1.5.4 Membrane filtration

This type of process involves separation and removal of macroscopic particles and colloidal matter from waste water. The range of particles sizes varies from 0.0001 to 1.0 µm. This process includes microfiltration, ultrafiltration, nanofiltration and reverse osmosis. Table 10 presents the description of each process (Crittenden et al. 2012).

Table 10-Classification of membrane filtration processes

Process	Separation mechanism	Pore size, nm	Operating range, µm	Constituent removed	Description
Microfiltration	Sieve	>50	0.08-2.00	Total suspended solids, cysts, some viruses and bacteria	Separation proceed by straining or sieving
Ultrafiltration	Sieve	2-50	0.005-0.200	Macromolecules, most bacteria, dome viruses and proteins	Separation proceed by straining or sieving

Continuation of table 10

Nanofiltration	Sieve + diffusion + exclusion	<2	0.001-0.01	Small molecules, viruses	Particles adsorbed on the membrane surfaces
Reverse osmosis	diffusion + exclusion	<2	0.0001-0.001	Very small molecules, Ions of nitrates, sulfates and sodium	Particles adsorbed on the membrane surfaces

1.6 Processes for removing nutrients

Recently, along with removing suspended solids the esterification process has drawn extensive attention since the algal and other types of aquatic plants growth excessively promoted by nutrients. The main suppliers of the nutrients ponds and surface waters are industrial and municipal wastewater disposals. One of the examples may be total concentration of nitrogen and phosphorus compounds in domestic wastewater streams which is 35-45 (for nitrogen) and 10-15 mg/L (for phosphorus). The concentration of nutrients for initializing of eutrophication is 0.3–0.5 mg/L for nitrogen and 0.01–0.05 mg/L for phosphorus. Therefore, for solving this issue, at wastewater treatment plants high nutrient removal efficiency should be achieved. Nutrient removal is a pretreatment process of wastewater prior to discharged into groundwater or surface by means of removing nitrogen and phosphorus compounds.

The high concentration of nutrients compounds in wastewater may cause cultural eutrophication or nutrient enrichment in surface waters. One of the consequences are algal blooms, therefore the following problems with ecosystems in water pods can be encountered: low concentration of dissolved oxygen, depletion of flora and fauna, and murky water (Rahman 2016).

1.6.1 Nitrification

Biological nitrification may be defined as a process of oxidizing ammonia (NH_4^+) to remove nitrogenous impurities from wastewaters. The most important chemical reactions of the process are the following:

$$2NH_4^+ + 3O_2 \rightarrow 2NO_2^- + 4H^+ + 2H_2O \quad (6)$$

$$2NO_2^- + O_2 \rightarrow 2NO_3^- \quad (7)$$

Domestic sewage contains approximately 20 to 40 mg/L (ppm) of ammonia nitrogen (Tchobanoglous et al. 2003). The organic matter in wastewater, which contains nitrogen in form of protein and nucleic acid, is able to biodegrade thus releasing ammonia. Ammonia released into water streams has a negative toxic impact, such as aquatic organisms formation and significant oxygen depletion. Thus, many domestic and industrial wastewater treatment plants are required to remove the ammonia before discharge of the treated water (Res 2014).

1.6.2 Ion exchange

Ion exchange process was designed for displacing insoluble exchange material in the solution by different species ions. Currently, this process used for municipal water softening by removal ions of calcium and magnesium in the purified water. Ion exchange processes also use nitrogen (N_2) as a source for heavy metals and total dissolved solids removal (Tchobanoglous et al. 2003).

This process proceeds by two variations:

- Batch mode, where resin is stirred with water in the reactor for complete water cleaning, waste resin is removed and regenerated;
- Second type is continuous mode, where exchange material is placed in a bed or packed column; water stream enters to the column bottom and passes through the bed of resin. The column packing is periodically regenerated in order to maintain acceptable properties.

Zeolites, which are complex of aluminosilicates with sodium as mobile ion, resins and phenolic polymers can be used as ion-exchange materials (Tchobanoglous et al. 2003). Typical ion-exchange reaction is provided below.

For natural Zeolites (Z): (8)

$$ZNa_2 + \begin{bmatrix} Ca^{2+} \\ Mg^{2+} \\ Fe^{2+} \end{bmatrix} \rightarrow Z \begin{bmatrix} Ca^{2+} \\ Mg^{2+} \\ Fe^{2+} \end{bmatrix} + 2Na^{+}$$

For synthetic resins: (9)

$$RSO_3H + Na^+ \leftrightarrow RSO_3Na + H^+$$
$$2RSO_3Na + Ca^{2+} \leftrightarrow (RSO_3)_2Ca + 2Na^+$$

$$RCOOH + Na^+ \leftrightarrow RCOONa + H^+ \quad (10)$$
$$2RCOONa + Ca^{2+} \leftrightarrow (RCOO)_2Ca + 2Na^+$$

$$RR'_3NOH + Cl^- \leftrightarrow RR'_3NCl + OH^- \quad (11)$$

$$RNH_3OH + Cl^- \leftrightarrow RNH_3Cl + OH^- \quad (12)$$

$$2RNH_3Cl+SO_4^{2-} \leftrightarrow (RNH_3)_2SO_4+2Cl^-$$

1.6.3 Biological phosphorus removal

Unremoved and untreated nutrients (especially phosphorus and nitrogen) in water may cause significant consequences, which were described earlier. Currently, there are three different types of processes for the nutrients removal, which are chemical, physical and biological. Biological process uses microorganisms for removing nitrogen and phosphorus from wastewater under different environmental conditions.

Biological process of removing phosphorus consists of three zones located consequently: anaerobic, aerobic zone and former zones. The main advantage of this process is that organic compounds uptake and phosphorus release take place under anaerobic condition, meanwhile phosphorus uptake takes place under subsequent aerobic zone. Overall phosphorus compounds are concentrated in the sludge, which are removed by sedimentation process. (Srinivasan, Chowdhury & Viraraghavan 2009). The overall reactions are represented below (Tchobanoglous et al. 2003).

Using lime:

1. $10Ca^{2+} + 6PO_4^{2-} + 2OH^- \leftrightarrow Ca_{10}(PO_4)_6(OH)_2$ (13)
2. $Mg^{2+}+2OH^- \leftrightarrow Mg(OH)_2$ (14)
3. $Ca^{2+}+CO_3^{2-} \leftrightarrow CaCO_3$ (15)

Using Al^{3+} :

1. $Al^{3+} + PO_4^{2-} \leftrightarrow AlPO_4$ (16)
2. $Al^{3+}+3OH^- \leftrightarrow Al(OH)_3$ (17)

Using Fe^{3+}:

1. $Fe^{3+} + PO_4^{2-} \leftrightarrow FePO_4$ (18)
2. $Fe^{3+}+3OH^- \leftrightarrow Fe(OH)_3$ (19)

1.7 Processes for removal colloidal and dissolved solids

In water filtration process activated carbon is a sort of filter whereon contaminants adhere to the surface carbon granules or also may be trapped in pores as different types. The scheme is presented in Figure 10.

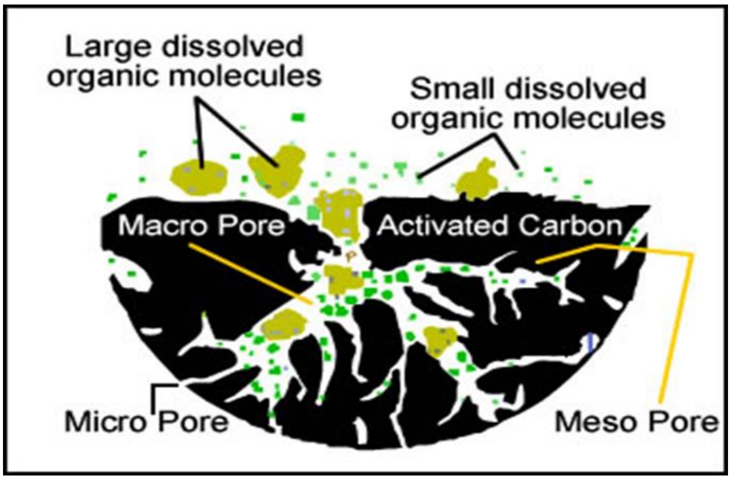


Figure 10-Structure of activated carbon (Begg Cousland 2017)

Activated carbon filters were designed for certain organics removal (unwanted taste and odors, micropollutants, chlorine, fluorine, radon, etc.) from drinking and wastewater streams. However, it is not efficient for microbial contaminants, and inorganic (metals, nitrates and other) contaminants removal. This type of filtration may be used in centralized treatment plants for effluents treatment and at household level integrated in hygiene equipment for drinking water production. It is also implemented in water treatment for the micropollutants removal in drinking water production and for the more efficient purification of treated wastewater before disposal (Tchobanoglous et al. 2003).

Currently there are two types of water filters: the first are particulate filters, which exclude particles depending on their sizes and the second are adsorptive or reactive filters, which contain a material (medium) designed for adsorption and reaction with contaminant impurities in water. The background of adsorptive activated carbon filtration is the same as to any other adsorption material. The different types of contaminants are drawn and adsorbed on the carbon particles surfaces. In order to achieve maximum efficiency of adsorption processes the different specification of the carbon material, for example pore and particle sizes, surface area, surface of different chemistry and other, should be determined.

The characteristics, specification and properties of the chemical contaminant are significant factors as well. For example, compounds that are insoluble in water are more efficiently adsorbed to a solid compound. Another characteristic is the similarity of contaminants in water streams with the carbon surface. This affinity property depends on the charge and are more prominent for molecules with less charge comparing with activated carbon filter (Ho et al. 2011).

1.8 Processes for removal volatile organic compounds

For many years one of the main steps in wastewater treatment processes were the removing of suspended solids, various bacteria and biochemical oxygen demand (BOD). The volatile organic compounds are also taken into account along these steps since they may cause nitrification. As an example, nitrogen organic matter is converted by means of hydrolysis into ammonia compounds (NH_4^+) during almost every water treatment stage. Furthermore, this ion is utilized by bacteria in cell synthesis and excess amount of ammonia enters to the final step of treatment as residual ammonia nitrogen (Sun 2003).

Active biological nitrification also may take place inside the aeration tank. The presence of organic components in wastewater sources which may be potentially hazard or toxic is detected by advanced analytical methods.

One of the most efficient processes for volatile organics removal is gas or usually air stripping. During this process liquid (water or wastewater stream) contacts with air, and as a result undesirable substances present in the liquid phase can be released and removed with the gas stream. (Lawrence K. Wang 2006). The basic description of counter-current air-stripping technology is presented in the Figure 11.

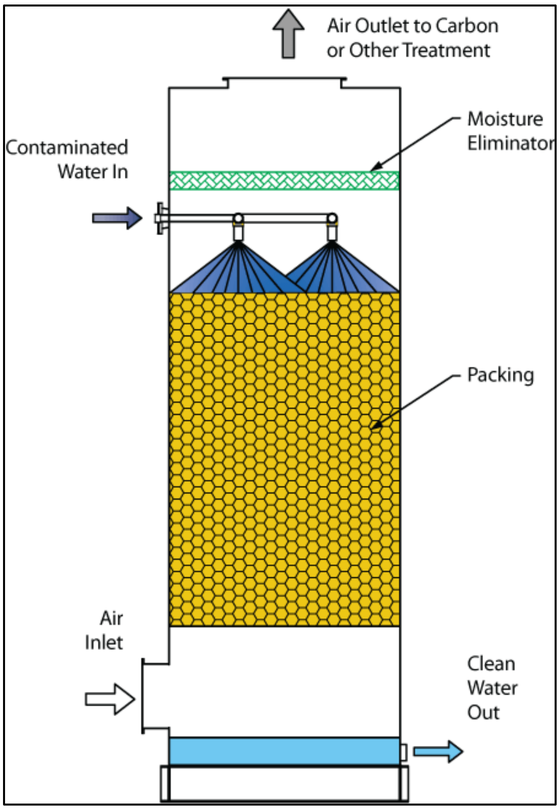


Figure 11-Counter-current Air-stripping technology (Srinivasan et al. 2009)

The counter-current flow of the air flow from the bottom part of the column goes upwards and during flow through different packing materials and faces with treated liquid which is supplied by liquid distribution system. Because of highly dispersed structure of packing materials the gas and wastewater distributed equally among the column volume. The separated air outflow from the water constituents goes to the air-treated units for further recirculation.

1.9 Odor control

The odor substance are usually inorganic compounds in gas form or highly volatile organic matter, which are: mercaptans (R-S), indoles (C_8H_7N), inorganic acids, skatoles (C_9H_9N), aldehydes, ketones and other organic compounds which contains nitrogen and sulfur atoms. All these compounds are formed by biological activity in sewage and in industrial wastes. For example, anaerobic decomposition of high molecular weight compounds (proteins) may cause formation of odor-responsible sources. The ammonia and hydrogen sulfide, among all inorganic compounds, are the most significant sources of odor from household's sewage drainage (Rajbansi 2014).

1.9.1 Chemical scrubbers

For the wastewater vapor-phase odor control wet air scrubbing is currently used since it is flexible and reliable technology. It may be implemented for water-soluble contaminant treating. This type of scrubbers also may be used for removing hydrogen sulfide, different organics odors and ammonia compounds. In wet-air scrubber units, impurities in water dissolve liquid-phase solution in vapor phase. For preventing leakages of odor from units under changing conditions of continuous processes the chemical balance between reagents and odor-sources are continuously maintained.

One of the possibilities of undesirable odor utilization is the multi-stage scrubber technology. This type of units has high efficient chemical usage and may treated wide range of contaminants and odor-sources (Tchobanoglous et al. 2003). The principle of two-stage wet scrubber unit is presented in Figure 12.

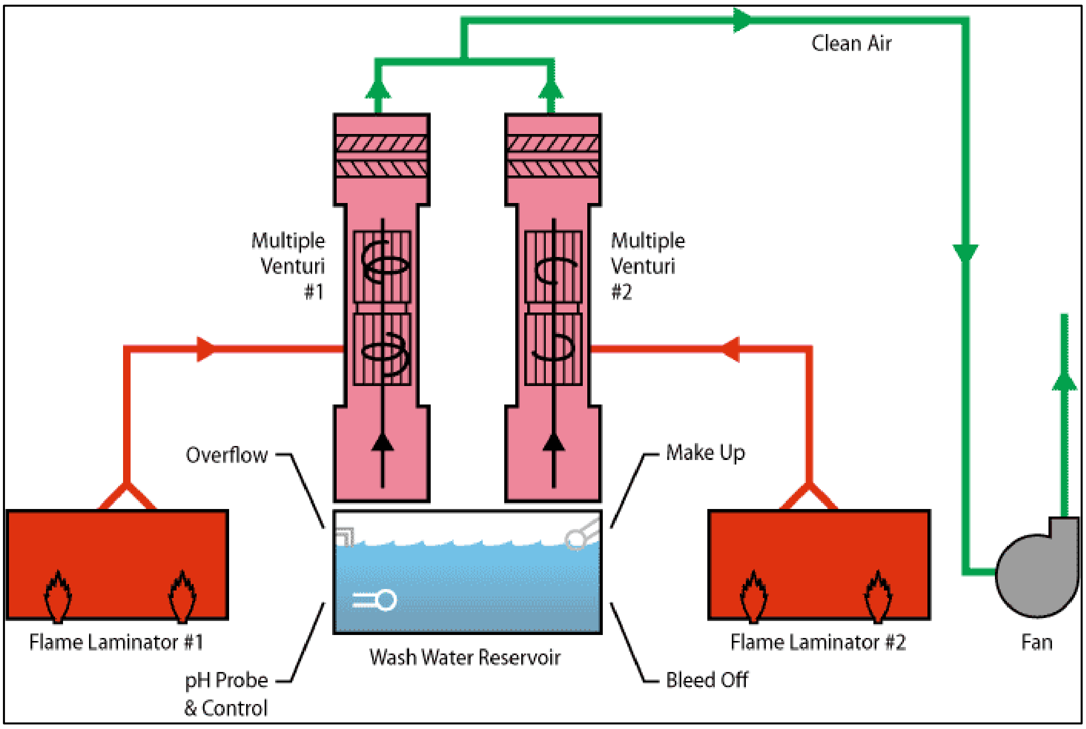


Figure 12-Operating principle of wet scrubber (Monroe Environmental 2018)

The wet-scrubbing systems are flexible and reliable due to using chemical reactions for water treatment processes. However, one of the process bottlenecks are high operating costs, which are required due to high amounts of chemicals. Therefore, specially designed units are used for reduction chemicals. The hydrogen sulfide may be solubilized with sodium hydroxide (NaOH) solution, along with organic-matter of odor sources in wastewater. Most efficient way for removing these organic derived odor-sources is the use of sodium hypochlorite (NaClO). In scrubber system these chemicals (NaClO+NaOH) are used in terms of recirculating chemical solutions. Sodium hydroxide (NaOH) and sodium hypochlorite (NaClO) are used in the single-stage scrubber systems. Due to strong oxidized properties of sodium hypochlorite, it rapidly reacts with sulfide compounds, which are solubilized by sodium hydroxide. Thus, for efficient hydrogen sulfide oxidation and treatment of other odor compounds the optimal quantities of sodium hypochlorite must be added in the two-stage system. (Harshman 2000).

1.9.2 Biofilters

The technology of biofiltration is used for treatment of different biodegradable and water-soluble contaminants. In biofilter units, the odor compounds are solubilized from the vapor phase into the liquid form on the various organic matter surfaces, such as compost, mulch or peat. Further these compounds are demolished by means of the bacteriological population on this media. The principle of biofilter is presented in Figure 13.

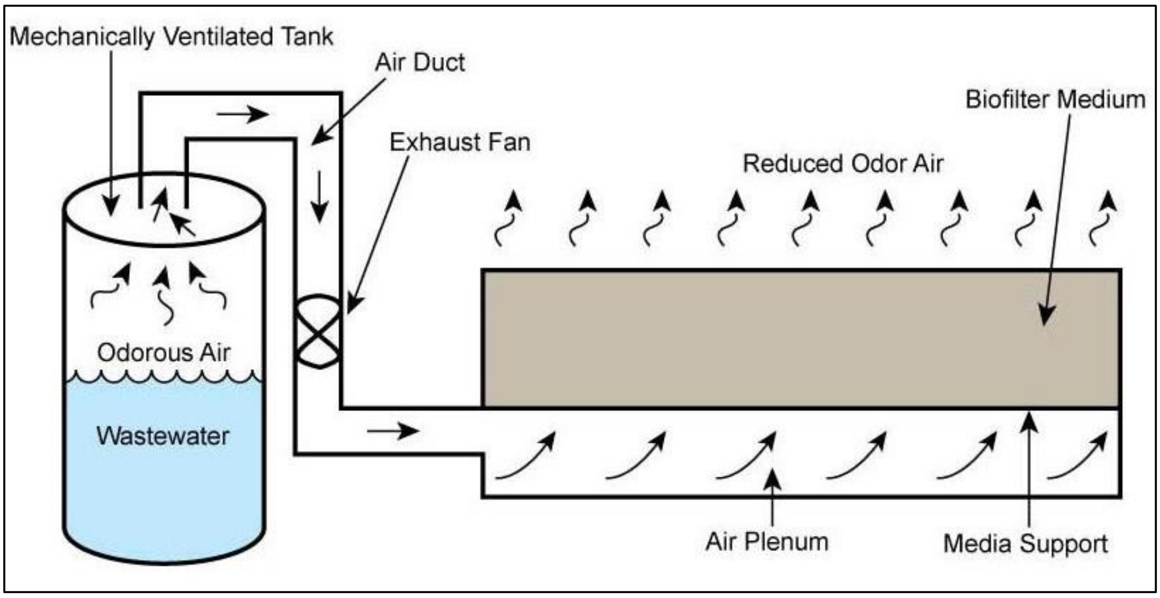


Figure 13-Open bed biofilter (Advanced Waste 2016)

These types of filters are advantageous for removing odor compounds, especially sulfur-based such as hydrogen sulfide, sulfur-containing organic matter and mercaptans. However, in case of nitrogen-containing compounds (ammonia and ammines) biofilters are not beneficial.

Biofiltration systems consist of two main parts, which are needed to be regulate:

1) Stability of the media since it may breakdown during the process. As a result, the head loss through the filter is increasing due to filter's bed sedimentation and compacts. This also may cause airflow decreasing and odor compound emissions.

2) Control of biofiltration process, since bacteria population upsets leads to break-through odors occurring.

Biofilters maintain media stability and control issues, which occur during the process. The filter media typically consist of the following materials: compost, mulch and wood chips. The materials selection depends on their local availability near the water treatment plant. The media composition consists of organic materials combination, which is set in accordance with wastewater specification for achieving repeatable performance and for ensuring consistency. The main process control is contingent on the filter's properties and design. Temperature and humidity fluctuations are the controlled parameters, which drive the biofilters properties. (Harshman 2000).

1.10 Pathogens removal

Domestic wastewater streams contain health-hazard materials or pathogens, which may be divided into 4 following groups: bacteria, viruses, protozoa, and parasites. The most typical bacteria in waste water are E. coli, Salmonella, Shigella, Helicobacter, Campylobacter and Vibrio cholerae. These types of pathogens are most common in household wastewater streams (Hu 2016).

The most hazardous viruses which contain in wastewater are hepatitis A, rotavirus and enteroviruses. The example of protozoa would be Entamoeba histolytica, Giardia lamblia, Cryptosporidium parvum. These types of pathogens are frequently detected in wastewater streams and may occur because of cysts and oocysts. As an example of parasites, which are detected in waste water, are helminths and their eggs, Ascaris (roundworm), Ancylostoma (hookworm) and Trichuris (whipworm) (Hu 2016).

Removal of pathogens is conducted by various of processes such as wastewater stabilization ponds, activated sludge (more energy intensive), chlorination, using ozone and ultraviolet radiation. However, these methods are not highly efficient for removing helminth eggs and protozoa since this type of pathogens have different behavior and resistance comparing with viruses and bacteria (Jimenez 2007).

Conventional wastewater treatment plants currently consist of four following purification steps:

1) Preliminary (pretreatment), which comprises grit removal and screening for large and coarse suspended solids removal, non-organic (sand, gravel, rocks) and organic (fats, oils and greases, plastics and wood particles) matter. During this step, the levels of pathogens and nutrients are not affected;

2) Primary treatment is usually sedimentation process in septic tanks, anaerobic ponds (including high-rate anaerobic ponds) or in upward current anaerobic sludge-blanket reactors. These processes have a few hours of hydraulic retention time, whereby almost all solids, which able to settle, sediment to the bottom of the treatment units. All settled particles are regularly removed during the continuous processes (depending on the process: once a day, once a week and every 1-3 years). The sludge, which occurs in these processes, may contain pathogens, particularly helminth eggs, thus further treatment is required before using water for agriculture need and to avoid hazardous compounds injection in subsurface soils (Jimenez 2010);

3) After the primary treatment processes follow secondary systems, which are basically the combination of biological processes with solid-liquid separation. Secondary aerobic treatment processes may be defined as a biological reactor with sedimentation tank, which combination

is used for organic components in wastewater derived biomass removing. Also, aerobic reactors may be combined with suspended growth processes, such as aerated lagoons, oxidation ditches and activated sludge, and with fixed-film processes, such as biological contactors and tricking filters;

4) Tertiary (or final) treatment processes are applied for additional purification downstream of secondary treatment. Among these processes there are additional solids removal by means of flocculation, filtration in granular medium, coagulation and sedimentation and disinfection. If some of these processes are implemented in wastewater treatment plants, the overall purification process is defined as advanced wastewater treatment (Jimenez 2007, Jimenez 2010).

2 IBM Watson Analytics

Current digital world consists of enormous amount of information, which may be fundamental for researches to direct their development. Unfortunately, only small part of this information is analyzed, integrated and catalogized for users. The main challenge lies in extracting exact points and values from this information stream and hundreds of sources. New technologies, such as natural language processing and cognitive computing provide systemization of big datasets according to user's request.

New technologies for efficient data analysis allow to process and evaluate big datasets according to user's request. Software is programed to understand and analyze technical and industry-specific information, big experimental data from chemistry, medicine and pharmaceuticals (Chen 2016). IBM Watson company provides various of different tools for data processing and analyzing. It is founded on Unstructured Information Management Architecture (UIMA) algorithms which are an open framework source. These toolkits may use different natural processing technologies such as Lucene and Solr, program applies corpus processing, semantic evaluations and parallelism for answers searching, prediction and pattern recognition (S. S. Murtaza et al. 2016).

In question-answering system the question may be defined as an initial input in data processing. Program extracts key elements of the questions and provides precise answers according to the user's input (BANSAL 2013). The IBM Watson's Architecture is presented in Figure 14:

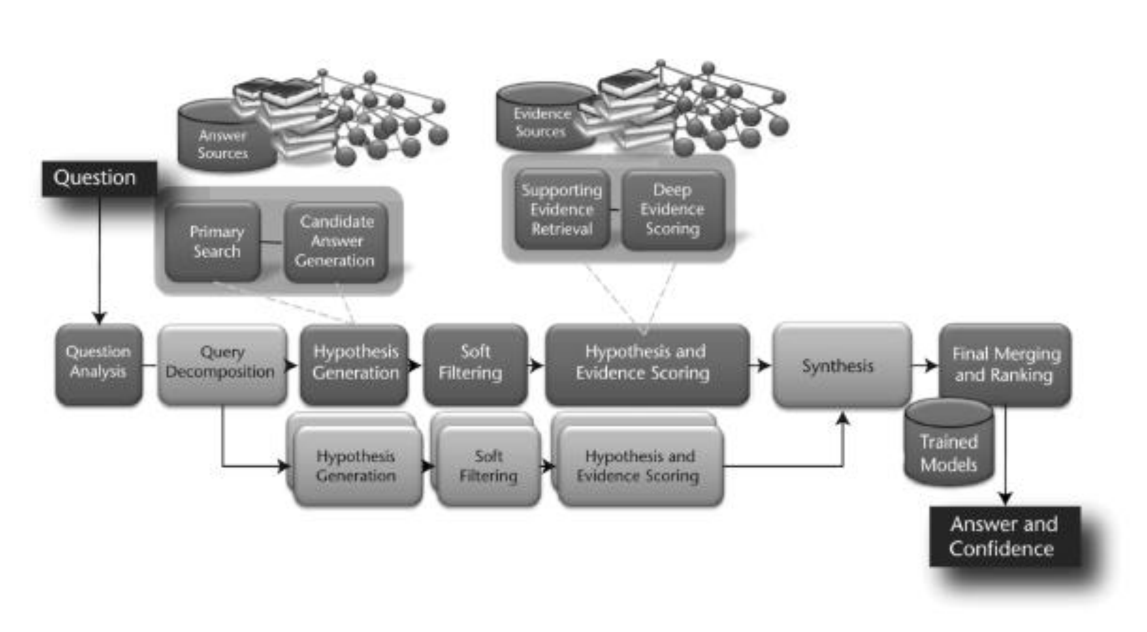


Figure 14-IBM Watson architecture (S. S. Murtaza et al. 2016)

First step in this architecture is questions analyzing, where keywords are segregated from the user's input. Also, during this step the type of the answer is defined based on predefined classification of name-entities, which basically are common nouns (person, place, values). The predicate structure detector establishes the relationship between these entities. All these objects initialized as an input for the second step, where IBM Watson uses key points for relevant information searching through all available sources. This search is based on Lucene and Solr natural language processing technologies. The results of the search are ranked according to keywords and answer type in document's titles and passages contents. During third stage ranked information is scored based on its integrity and matched to user's input by different analytics algorithms. During final fourth stage program provides final version of merged and ranked information to the user according information's score or confidence. The high confidence values provide more relevant and useful information (BANSAL 2013).
IBM Watson Analytics is a cloud-based toolkit, which allows the user to conduct complex analytics, using a web browser. The main aim is to allow users focus only on experiment results and case study (IBM 2018). Due to simple interface, users may be familiar with data analytics techniques or not.

After uploading the database on the cloud, the IBM Watson Analytics offers three categories of inputs, which are: Explore, Predict and Assemble. The "Explore" mode provides data clusterization for patterns recognition and relationships determination between data. The "Predict" mode allows to perform predictions according to the data, discovering the predictive strength of the most significant parameters, compared to a single target parameter set up by the user. “Assemble” mode is providing efficient data visualization using infographics (Guidi et al. 2016).

One of the significant possibilities of Watson Analytics, is the full automation of analysis steps, simplify the using of the program for non-expert users. The main automated step is represented below (Guidi et al. 2016):

- data preparation, including transformation and quality index determination (analysis of empty fields and identification of constant values);

- modeling, including most relevant models setting and identification of strong patterns between data values.

The “Predict” mode provides analyses by setting a prediction target. By using this mode, it is possible to look through the predictive power of other parameters.

IBM Watson Analytics provides the graphics and text results, in following ways (Guidi et al. 2016):

1) Single predictor: the predictive value of the most significant parameter;
2) Double predictor: the first two most significant parameters;
3) Combination: the combination of several parameters for a more efficient prediction.

After the switching from “single predictor” to number of predictors, it may cause the increase of the overall prediction accuracy, however, it demands more precise result analysis. In some fields this can be less acceptable compared to losing some percentage points in accuracy (IBM 2018).

EXPERIMENTAL PART

In experimental part the databases, which contain the information regarding wastewater treatment plant, were analyzed using IBM Watson Analytics and MODDE Pro software. Both programs were designed for data analyzing, pattern recognition, determination of most significant variables and process/experiment optimization, according to obtained results. However, each program has features and limitation, thus comparative analysis was obtained for efficient data processing.

3 Aims of the research

The following aims were assigned to the experimental part:

1) Study of the possibilities of implementing IBM Watson Analytics for improvement wastewater treatment process;
2) Conduct the study: how the removal efficiency for all constituents, amount of added chemicals and influence of flowrate, temperature of process and fluctuations of pH values interrelate with each other (as an example: “How does flowrate and temperature affect removal efficiency parameters”?);
3) Determine the main driving parameters for most significant and relevant values in wastewater treatment processes;
4) Conduct and compare the predictive model in IBM Watson Analytics and MODDE Pro;
5) Obtain the same results in MODDE Pro software and compare them with IBM Watson results;
6) Determine limitations and benefits of IBM Watson Analytics and MODDE Pro.

4 Description of database

4.1 Source of information

The databases were assembled according to Wastewater treatment annual report (2012-2016) provided by Southwest Finland Water and Environmental Research Ltd. The company is located in Turku, Finland and provides reports and studies regarding water and environmental protection. All water properties are measured according to ISO/IEC 17025 requirements.

4.2 Database assembling

For the experimental part, the typical results from the wastewater plant were used, according to the information from five annuals reports. After extraction, the overall results were combined in one table, the assembling approach will be onward described. The description of studied wastewater characteristics parameters is presented in Table 11.

Table 11-Wastewater characteristics

Parameter	Description
Flowrate, m^3/d	Amount of charged wastewater
Temperature (°C)	temperature in the system during the process
Conductivity, mS/m	Changing value (inlet and outlet) before and after treatment
pH	Changing value (inlet and outlet) before and after treatment
BOD, mg/L	Biochemical oxygen demand, inlet and outlet values, concentration after treatment
COD, mg/L	Chemical oxygen demand, inlet and outlet values, concentration after treatment
Inlet/outlet total P, mg/L	Concentration of total phosphorus constituent before and after treatment
Diss. P, mg/L	Concentration of dissolved phosphorus after treatment
Inlet/outlet NH_4, mg/L	Concentration of ammonia ions constituent before and after treatment
NO_2^- treated, mg/L	Concentration of NO_2^- ions in water after treatment
NO_3^- treated, mg/L	Concentration of NO_3^- ions in water after treatment
Inlet/outlet solids, mg/L	Concentration of total solids constituents before and after treatment
Fe, mg/L	Concentration of ferrous ions in water after treatment
Diss. Fe (mg/L)	Concentration of dissolved iron in water after treatment
Nitrification efficiency, %	Efficiency of nutrients removal after treatment
Added Fe sulfate, mg/L	The amount of added ferrous sulfate for water treatment
Added polymers, mg/L	The amount of added polymers for water treatment

As a quantity measure of wastewater purification efficiency, the value efficiency removal was used. For all constituents these values were calculated according to the following equation:

$$Efficiency\ removal = (1 - \frac{Outlet\ concentration}{Inlet\ concentration}) \cdot 100\% \qquad (20)$$

Overall data was imported from TSP-annual reports for years from 2012 to 2016. The example for one year is presented in Figure 15.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Date	3.1.	6.1.	7.1.	8.1.	11.1.	12.1.	15.1.	18.1.	20.1.	23.1.	26.1.	28.1.	31.1.
2	Inlet flow m³/d	68600	66200	66400	67900	66200	66400	63900	62900	63300	59200	123000	213000	113000
3	Outlet flow m³/d	68600	66200	66400	67900	66200	66400	63900	62900	63300	59200	123000	213000	113000
4	Temperature °C	13,5	13,7	13,7	13,7	13,7	13,8	14,1	13,9	14,2	14,1	13,1	9	9,4
5	Inlet Alkali (vl) mmol/l			6,4				7,1	7,2			4,3		
6	Outlet alkali mmol/l			2,1				1,7	2			2,5		
7	Inlet conductivity (vl) mS/m			100				110	100			74		
8	Outlet mS/m			73				70	73			78		
9	Inlet pH (vl)	7,7	7,8	7,6	7,6	7,5	7,6	7,5	7,6	7,6	7,7	7,4	7,4	7,6
10	Outlet pH	6,7	6,9	6,9	6,9	6,8	6,8	6,9	6,9	6,8	7,2	6,8	6,6	6,8
11	Inlet COD (vl) mg/l	570	670	750	740	790	860	770	850	880	810	700	370	400
12	Outlet COD mg/l	35	36	34	34	36	36	39	38	47	30	48	37	27
13	COD removal efficiency, %													
14	Inlet BOD (vl) mg/l	250	290	380	320	370	370	300	410	390	310	300	180	180

Figure 15-Case 1 data analyzing

Since, after importing data to IBM Watson Analytics, the quality of provided database was 45%, it was decided to transpose tables using Matlab program. After this approach, the database visual environment has become as presented in Figure 16.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Day	Inlet flow m³/d	Outlet flow m³/d	Temperature °C	Inlet Alkali (vl) mmol/l	Outlet alkali mmol/l	Inlet conductivity (vl) mS/m	Outlet conductivity mS/m	Inlet pH (vl)	Outlet pH	Inlet COD (vl) mg/l	Outlet COD mg/l	COD removal efficiency, %	Inlet BOD (vl) mg/l	Outlet BOD mg/l
2	03.01.2016	68600,00	68600,00	13,50					7,70	6,70	570,00	35,00	93,86	250,00	1,90
3	06.01.2016	66200,00	66200,00	13,70					7,80	6,90	670,00	36,00	94,63	290,00	2,30
4	07.01.2016	66400,00	66400,00	13,70	6,40	2,10	100,00	73,00	7,60	6,90	750,00	34,00	95,47	380,00	2,40
5	08.01.2016	67900,00	67900,00	13,70					7,60	6,90	740,00	34,00	95,41	320,00	2,50
6	11.01.2016	66200,00	66200,00	13,70					7,50	6,80	790,00	36,00	95,44	370,00	2,70
7	12.01.2016	66400,00	66400,00	13,80					7,60	6,80	860,00	36,00	95,81	370,00	2,20
8	15.01.2016	63900,00	63900,00	14,10	7,10	1,70	110,00	70,00	7,50	6,90	770,00	39,00	94,94	300,00	1,90
9	18.01.2016	62900,00	62900,00	13,90	7,20	2,00	100,00	73,00	7,60	6,90	850,00	38,00	95,53	410,00	2,20
10	20.01.2016	63300,00	63300,00	14,20					7,60	6,80	880,00	47,00	94,66	390,00	3,00
11	23.01.2016	59200,00	59200,00	14,10					7,70	7,20	810,00	30,00	96,30	310,00	2,50
12	26.01.2016	123000,00	123000,00	13,10	4,30	2,50	74,00	78,00	7,40	6,80	700,00	48,00	93,14	300,00	10,00
13	28.01.2016	213000,00	213000,00	9,00					7,40	6,60	370,00	37,00	90,00	180,00	15,00
14	31.01.2016	113000,00	113000,00	9,40					7,60	6,80	400,00	27,00	93,25	180,00	2,10

Figure 16-Case 2 data analyzing

The quality value of case 2 database is higher compared with case 1 database and equals 78%. The quality value of data describes the efficiency IBM Watson Analytics in data processing. It means lower the quality value is, the lower accuracy of predictions and exploration are gained.

The average values were calculated according to the approach, which is presented in Figure 18.

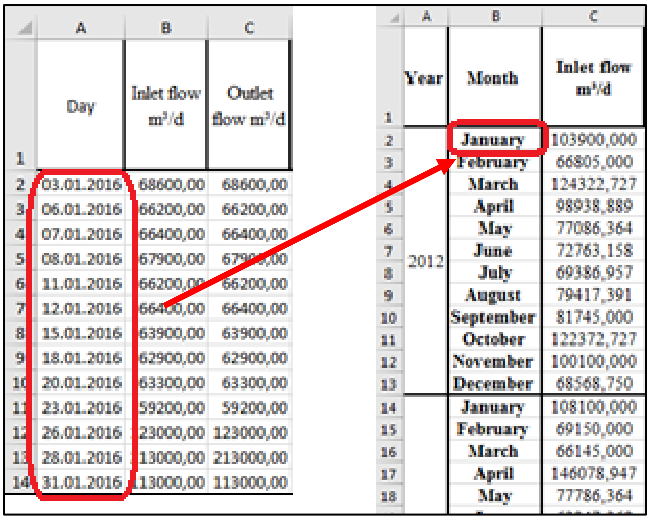


Figure 17-Average values calculation

Furthermore, monthly average values were calculated for overall wastewater parameters and added chemicals. The example of final database is presented in Appendix 1.

The quantity representation for all databases is presented in Table 12.

Table 12-Number of values in studied databases

Studied databases	Number of rows	Number of columns	Total quantity of values
Month database	60	34	2040
Total database for 5 years	1153	34	39 202

All studies and results were obtained with both databases according to the program's limitations, which will be discussed in the following chapters.

5 IBM Watson setup

For analyzing data in IBM Watson Analytics, it needs to be downloaded to cloud storage using personal account at webpage https://watson.analytics.ibmcloud.com. After pressing "New data" (Figure 18), the file browsing window opens. The data may be imported via IBM Watson applications and websites. Also, it is possible to download files from the personal device hard drive disk according to following order: Local file – browse – Import.

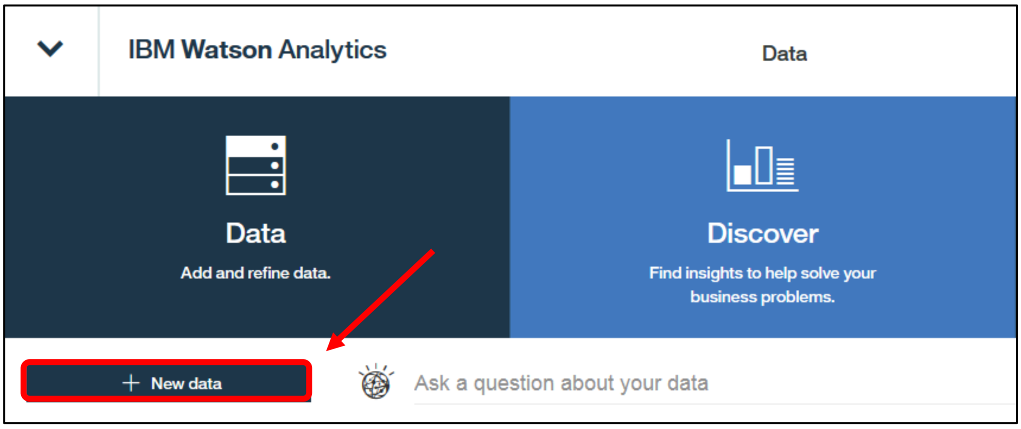


Figure 18-Data importing in IBW Watson Analytics

After importing the data, it is emerged in Cloud storage **(1)**, where further studies are conducted and saved **(2)**. In the main account window, it is possible to sort the databases according to user's purposes **(3)**. The user databases interface is presented in Figure 19.

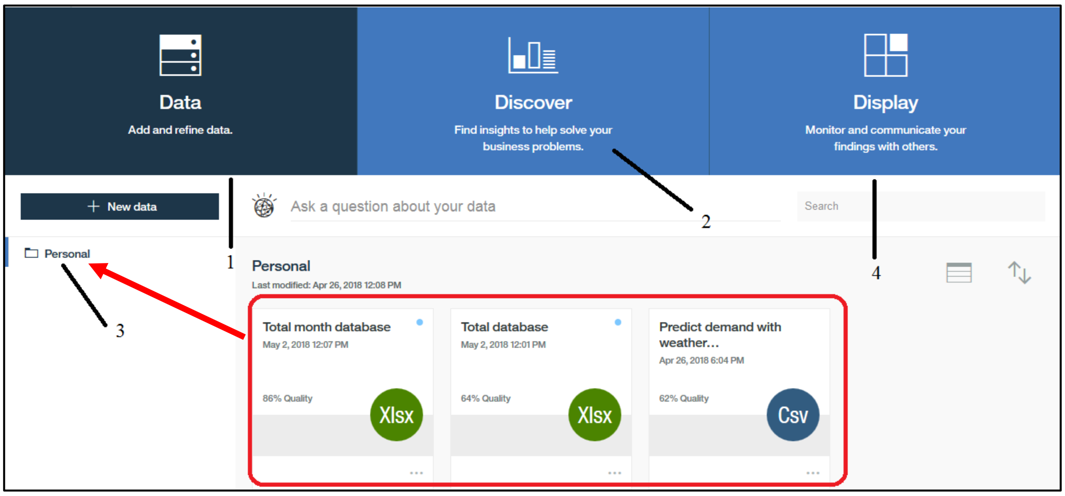


Figure 19-IBW Watson Analytics user interface
1 – Window with overall downloaded databases; 2 – Window will overall conducted studies;
3 - Folders with sorted databases; 4- Window for sharing databases with other users

For convenient group researching IBM Watson Analytics provided the databases sharing and results with other users **(4)** by Dashboard and Infographic creation.

For conducting the study, the selection of visualization is required. In other words, for efficient result providing the optimal visualization is needed to be fitted. The selection of different studies is presented in Figure 20.

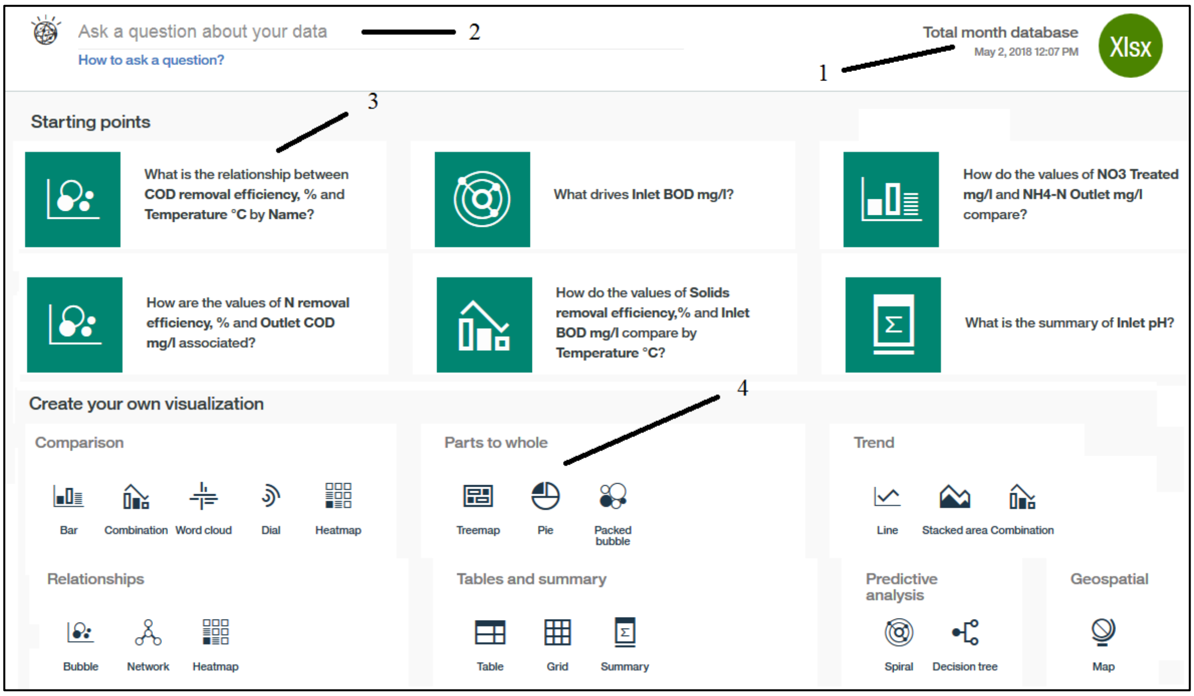


Figure 20-IBW Watson Analytics Discovery

On the result screen the name of the database **(1)**, space for the question input **(2)**, recommended by IBM Watson Analytics results, variations of possible study visualization **(4)** are presented. Starting points **(3)** are the recommendation provided by the program, according to the obtained results. In the question space user may manually input the request and list of results will be automatically proposed to the user. The examples of questions are represented below:

1) How the values of "Parameter 1" and "Parameter 2" associated?
2) What drives parameter 1"?
3) How do values of "Parameter 1" compared to "Parameter 2"?
4) What is the trend of "Parameter 1" over Parameter 2"?
5) What is predictive model for studied "Parameter"?

Also, IBW Watson Analytics provided recommended and most relevant results or "Starting points", which further may be modify. The discoveries may be set manually by using different types of visualizations as following: Comparison (for analyzing several parameters), Parts to whole (for determination of components share), Trend (for pattern recognition between values), Relationships (how several values associated with each other), Tables and summary, Predictive analysis (for prediction of possible variations depending on values) and Geospatial (sorting results according to geographic location - maps).

Key appointment of starting points is to determine correlation and patterns between initial values. The most relevant parameters of wastewater are efficiency of hazardous and undesirable compounds removing, amount of added chemicals and influence of flowrate, temperature of process, fluctuations of pH values and conductivity.

5.1 Inlet parameters

As inlet parameters the temperature, flowrate, inlet iron concentration and inlet alkali concentration were set. All these process variables may be controlled and maintained for increasing efficiency of wastewater purification processes. In this chapter the influence of these parameters on removal efficiencies were analyzed by means of IBM Watson Analytics.

5.1.1 Inlet flow rate

Inlet flow rate has significant impact on removal efficiency parameters. In all cases the following pattern is occurred: the decreasing of flowrate leads to increasing of removal efficiency. The example for Nitrification efficiency is presented in Figure 21.

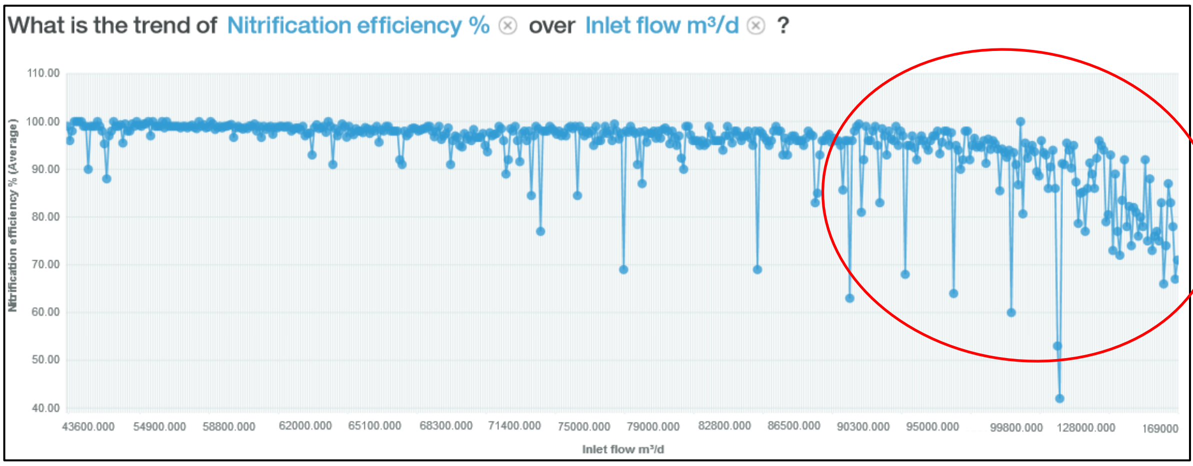


Figure 21-What is the trend of Nitrification efficiency over Inlet flowrate

According to results, which are provided in Figure 21, the Nitrification efficiency decreased under high flowrates values, and as a result the distribution is less linear. Therefore, the flowrate should maintain below 90 300 m^3/d values to achieve desired level of water treatment in terms of nutrients removal.

The results for Biochemical oxygen demand are presented in Figure 22.

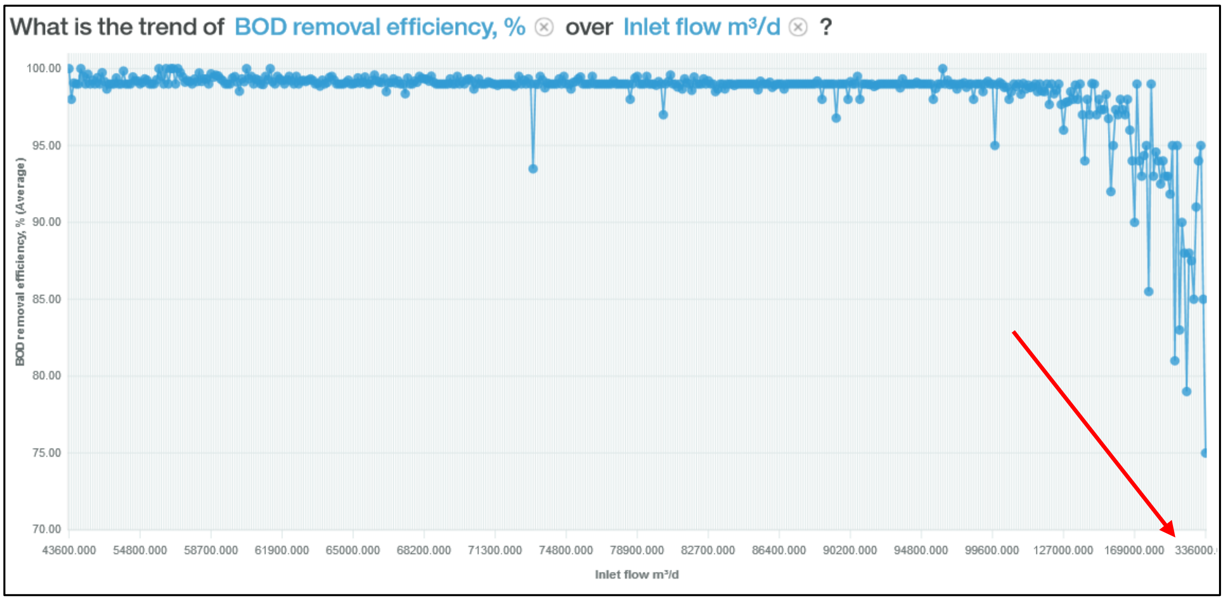


Figure 22-Trend for BOD over inlet flowrate

The flowrate values also have an impact on removal efficiency, however in case of BOD removal efficiency the significant influence observed at flowrate range higher from 127 000 m^3/d to 169 000 m^3/d.

For COD the results of flowrate influence are presented in Figure 23.

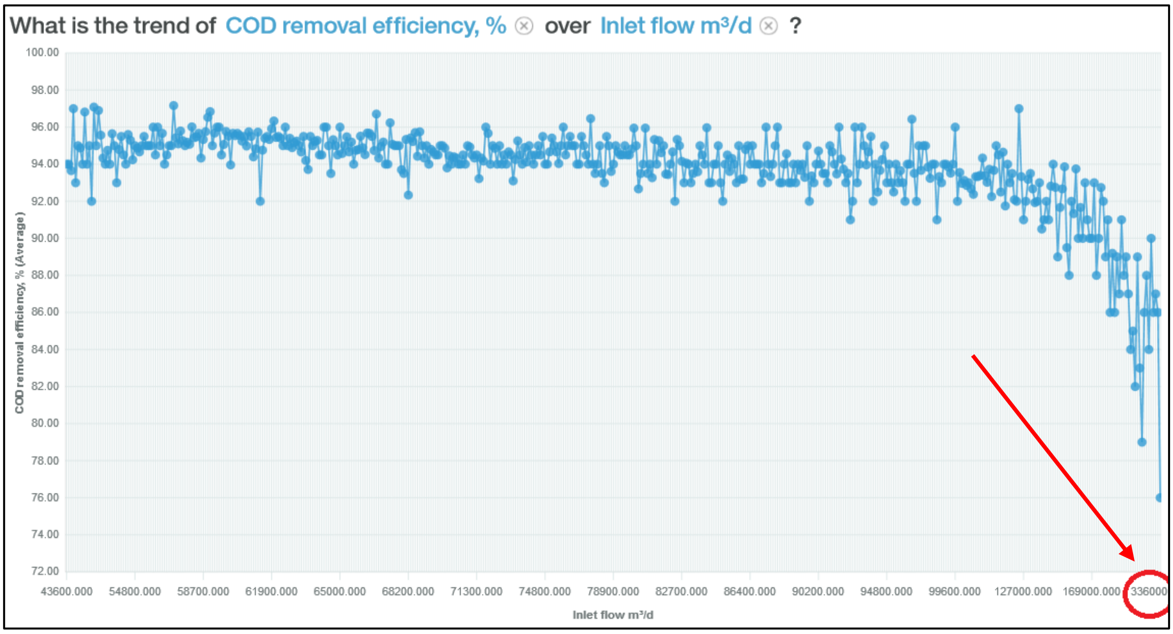


Figure 23-Trend for COD over inlet flowrate

The same pattern as in Biochemical oxygen demand case is observed. The significant impact of flowrate on the purification factors is observed in range from 96 000 m3/d to 169 000 m3/d. The flowrate, which is equal 336 000 m^3/d, was proposed by IBW Watson Analytics.

As may be noted from the results for Solids and Phosphorus removal, which are presented in Appendix 2, the flowrate does not exert a significant impact on hard particles and phosphorus treatment. The slight decrease of removal efficiency is occurred at large flowrates (more than 170 000 m^3/d) of wastewater.

In terms of Nitrogen removal, the flowrate values also have influence, however, the distribution of treatment efficiency is less linear and equal, than other studied parameters. One of the possible reasons is high quantity of Drivers for nitrogen-containing impurities removal. It means that fluctuations in other parameters values may cause nitrogen treatment variations as well. The figure is presented in Appendix 2.

For ammonia removal efficiency the flowrate increasing influence significantly in the same way as in previous studies. The critical decreases of treatment are observed with range of flowrate from 78 900 m^3/d to 127 000 m^3/d. Further amount of wastewater increasing is also has not any meaningful positive effect on ammonia removal. The figure is presented in Appendix 2.

Analysis of the results, provided by IBM Watson Analytics in terms of the influence of flowrate on wastewater treatment efficiency, revealed that water supply on wastewater treatment plant is needed to be maintained on the defined level for efficient removal of constituents. High values of flowrates may provoke declining of all purification factors due to lack of time for chemical reactions, which are used in a major of wastewater treatment processes. For all studied parameters of treatment efficiency, the optimal range of flowrate is from 79 000 m^3/d to 90 000 m^3/d. In case of Phosphorus and Solids removal with value may be increased up to 170 000 m^3/d, nevertheless it will endure declining of other parameters treatment efficiency.

5.1.2 Inlet alkali concentration

The alkali concentration in wastewater severely tracking due their impact on water pH level. In initial part of the purification processes alkali as treatment chemical added to the water for neutralization of acidic ions H^+. Therefore, the removal efficiency is contingent on alkali amount. As an example, the results for nitrification efficiency are presented in Figure 24.

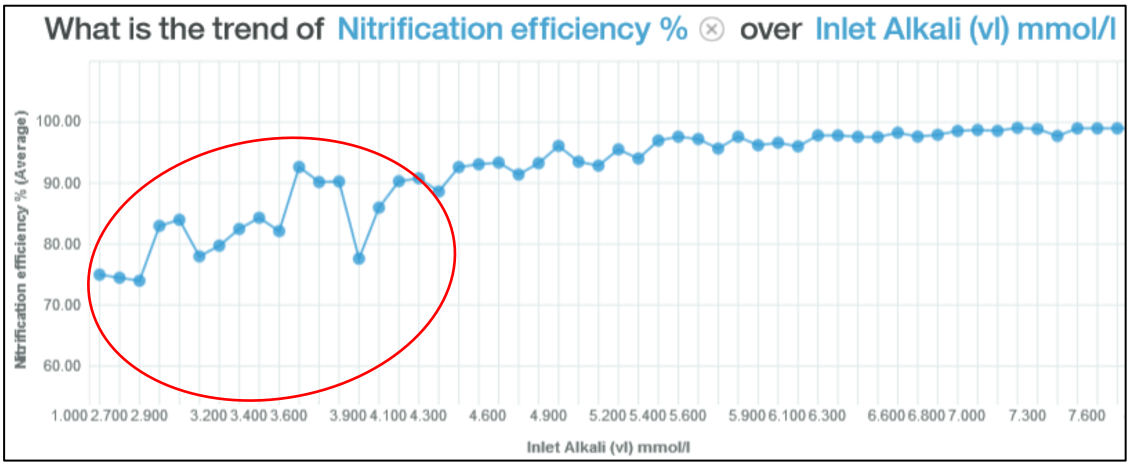


Figure 19-What is the trend of Nitrification efficiency over Inlet alkali concentration

Analyzing the results, low values of inlet alkali concentration decrease the efficiency of nutrients treatment. During nitrification the acid formation is occur, however, these acids may be neutralized by the added alkali (Marietta 2015). Thus, the optimal value of these chemicals should be maintained in terms of increasing water treatment degree of purification and compliance with pure water requirements.

In terms of Biochemical Oxygen Demand and Chemical Oxygen Demand treatment the Inlet alkali influence is similar as in nitrification case – low values of alkali decrease BOD and COD treatment in wastewater. The optimal value of alkali concentration is in range from 4.7 to 7.6 mmol/l of water. The further increasing may cause strong alkaline conditions of purified water, thus the optimal for neutralization of acids amount of alkali need to be added to the wastewater. The figure is presented in Appendix 3. The impact of Inlet alkali concentration on solids and phosphorus is insignificant. The figure is also presented in Appendix 3.

Low values of alkali compounds in wastewater induce extremely low purification measures of nitrogen and ammonia after treatment. The results are presented in Appendix 3. During the nitrogen containing compound treatment chemical reactions are meant to be used, which demand specified conditions as temperature and pH.

According to the obtained results, it may be noted that Inlet alkali concentration has significant impact on water treatment efficiency substantially on Nitrogen, Ammonia and Nutrients removal. Therefore, the similar to the flowrate case, alkali concentration need to be maintained on optimal level, which is in range from 5.2 mmol/L to 7.6 mmol/L. The higher amount of alkali, as it was described above, is not recommended due to strong alkaline conditions. The lower amount may be insufficient for ions H^+ neutralization.

5.1.3 Temperature

The average ambient temperature in Finland is 4.5 ^{o}C, in range from 7^{o}C to 17 ^{o}C. The fluctuation of this parameter is insignificant, however has influence on purification and compounds removal in wastewater treatment plants. The dependence of nutrients treatment over temperature is presented in Figure 26.

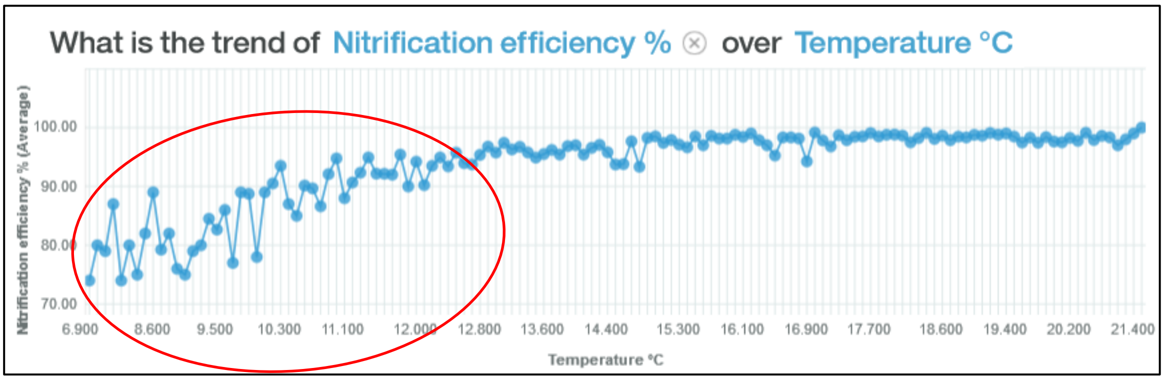


Figure 20-What is the trend of Nitrification efficiency over Temperature

At low temperature (below 12 ^{o}C) the decreasing of purification efficiency is observed. With low average ambient temperatures heat sources in water plants are demanded. Also, during the winter seasons temperature maintenance is required.

According to the Figure 26 the low temperature values significantly decrease COD purification.

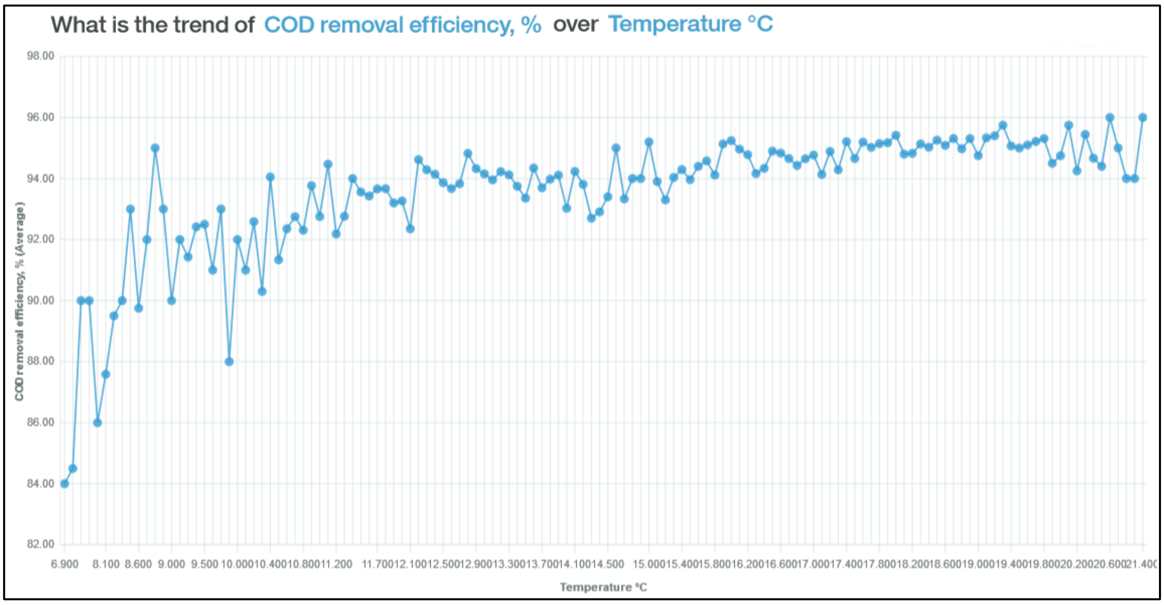


Figure 21-The trend of COD over Temperature

The same pattern may be observed in terms of BOD removal efficiency and temperature influence. The results are presented in Figure 27.

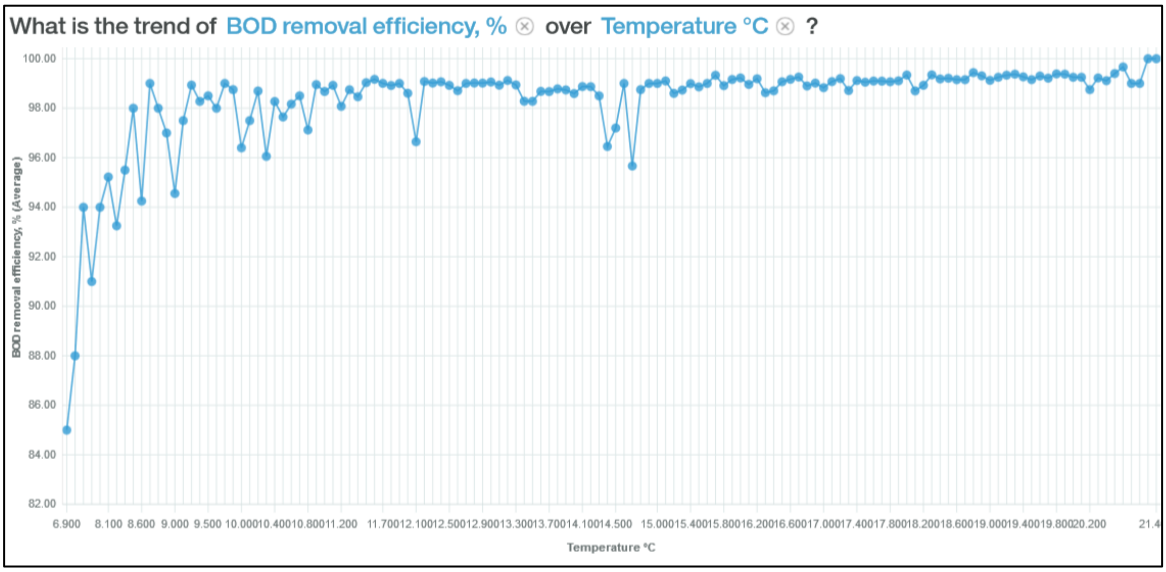


Figure 22- Trend of BOD over Temperature

Therefore, temperature maintenance for purification improvements is similar to inlet flowrate.

Nitrogen and ammonia are strongly depended on temperature fluctuations, which are presented in figures in Appendix 4. As it was described in alkali case, due to chemical reactions conducting, the optimal parameters ranges are required.

Analysis of the results of temperature influence suggests that for increasing the nitrogen compounds removal the temperature needs to be severely maintained. Also, other parameters may be fitted to the required temperature for treatment of nitrogen compounds.

5.1.4 Inlet conductivity

The conductivity in water is related to cations and anions concentration, which are obtained from dissolved salts and inorganic compounds. The dependence of nutrients treatment over inlet conductivity values is presented in Figure 28.

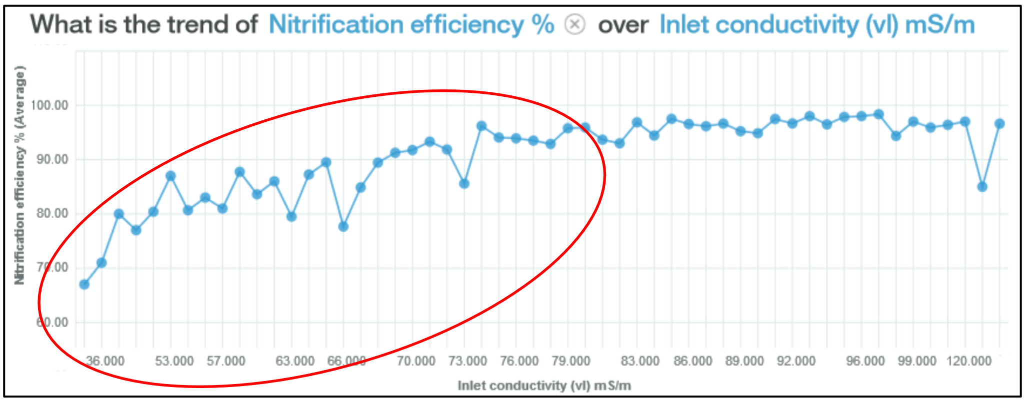


Figure 23-What is the trend of Nitrification efficiency over Inlet conductivity

According to Figure 28 the decreasing of conductivity leads to lower nutrients removal during purification processes. The same pattern is observed in nitrogen and ammonia cases, which are presented in Appendix 5.

The inlet conductivity has not any significant influence on Biochemical and Chemical Oxygen Demand, solids and phosphorus removal efficiency values, thus the same approach as in temperature case is need to be conducted.

5.1.5 Iron concentration

The presence of iron causes deterioration drinking water taste qualities, increases rate of corrosion and forms colloidal precipitates. Also, it may affect efficiency of purification processes. The dependence of nutrients treatment over Inlet iron concertation values is presented in Figure 29.

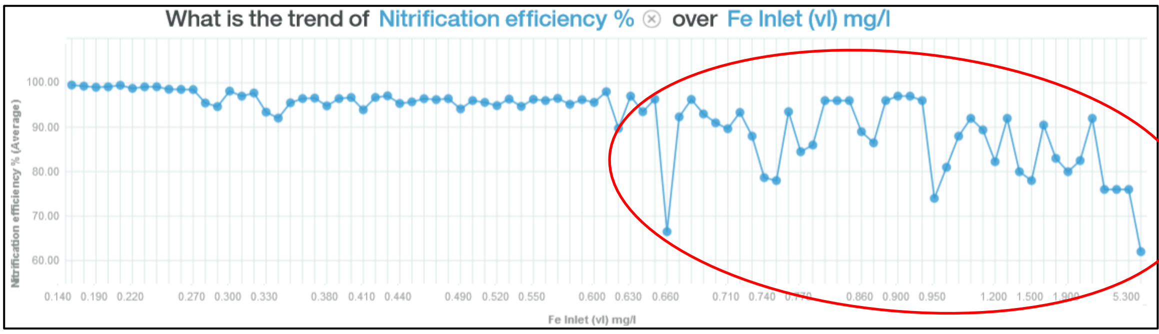


Figure 29-What is the trend of Nitrification efficiency over Iron concentration

With high iron concertation in water streams the removing of nutrients is decreasing. The same pattern is observed in Ammonia and Nitrogen cases. The results for these studies are presented in Appendix 6. One of the possible reasons is colloidal precipitance formation during water treatment processes in pipelines.

For BOD and COD removal efficiency Iron also has impact, however it is less significant in comparison with nitrogen-contained compounds. The results are presented in Appendix 6. For Solids and Phosphorus, the Iron concentration has no any effect on purification processes. The results are also presented in Appendix 6.

5.2 Outlet parameters

Due to high emphasis on removing hazardous components the following most relevant parameters were studied: the removal efficiency of Chemical Oxygen Demand, Biochemical Oxygen Demand, Nitrification, Phosphorus, Nitrogen, Ammonia and Solids. In this study the drivers of parameters were determined. Drivers describe which value has the highest influence on removal efficiency. The evaluation of influence is provided according to driver's percent. The higher the percent is, the higher impact exerts the given parameter. The IBM Watson application provided the drivers description as presented in Figure 30.

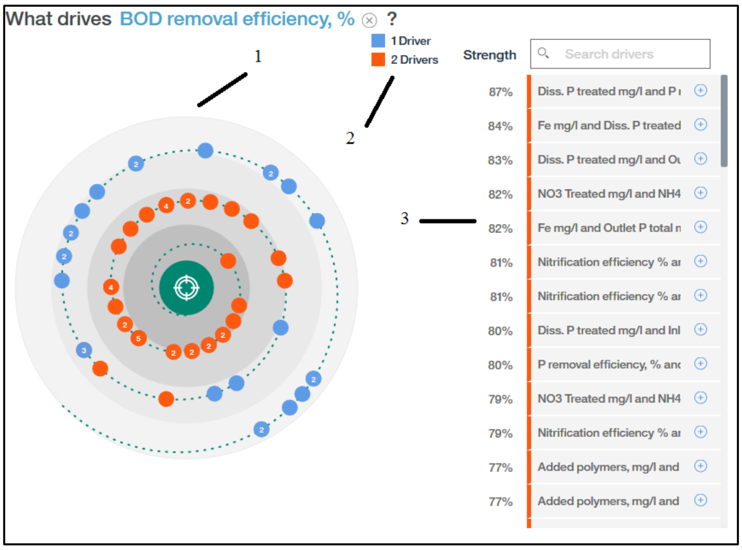


Figure 30-IBM Watson drivers
1- Driver spiral, 2- Number of drivers, 3- Strength and value

In figure 2.3 Driver Spiral **(1)** visualized which values have the highest impact on removal efficiency parameter. Number of values **(2)** represent the number of drivers on the spiral. 2 Drivers explain how the combination of two parameters effect on studied object. The strength and value **(3)** describe which exact parameter has the influence (percentage of the strength) on studied removal efficiency. The overall most relevant drivers of each removal efficiency are presented in Appendix 7.

5.3 Predictive models

One of the most relevant features of IBM Watson Analytics are predictive models. A decision tree visualization (Predictive model) is used to illustrate how studied data predicts a chosen target and highlights key insights about the decision tree. However, with small amount of data this type of visualization does not provide relevant results due to insufficient quantity of information for pattern recognition. Therefore, the analysis was not conducted for monthly based.

The number of rows in five-years database was equal 1153 rows, which provided more efficient pattern recognition. The prediction tree for nitrification efficiency provides relevant results according to parameters, which have impact on studied parameter. The Nitrification correlated to Ammonia concentration, since it converts into nitrate in water, thus Outlet Ammonia concentration is the most significant value. The prediction model is presented in Figure 31.

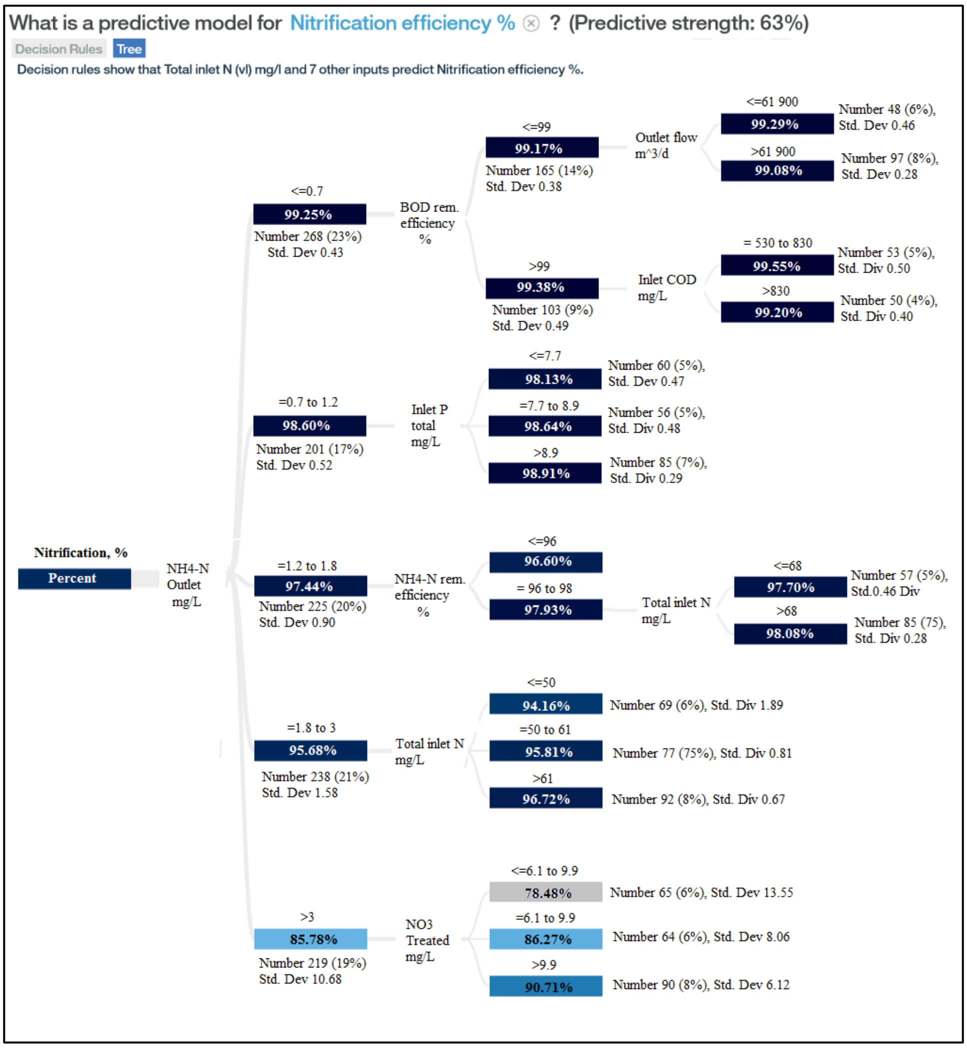


Figure 31-Prediction model for nitrification efficiency

Predictive model determines the most relevant drivers of the studied component and describes the various alternative of Nitrification removal efficiency. Each driver is affected by its own significant parameters (or his own predictive model) according to possible scenario, therefore it provides relevant results regarding controlling and maintaining of parameters.

5.4 Clarification of the results

Analyzing the results obtained by IBM Watson Analytics, it may be conducted that the most significant technological values are inlet flowrate and temperature. The purification efficiency factors may be severely declined to such an extent, that water quality does not fulfill the requirements. The highest impact flowrate and temperature have on nitrification and nitrogen removal particles. The significant impact may be explained in terms of sensitive conditions of nitrogen-containing compounds removal reactions. There is no effect on phosphorus treatment processes, however, the technical part is similar to nitrogen removal and nitrification. The range of phosphorus removal efficiency values is from 96.25% to 99.01%, thus such slight fluctuations do not provide consistent results, and the larger amount of accurate data is required.

The added concentration of alkali in waste water influences purification factors, such as nitrification and ammonia. Since during nitrification the H+ cations are obtained, the anions OH- neutralize and stabilize of water pH value. As it was described above, maintenance of this chemical concentration in pretreatment stages may reduce strong acid condition in purified water.

The increase of conductivity values has positive impact on nitrogen, ammonia and nitrification treatment. However, high conductivity is not compatible with treated water quality. Typically for pure drinking water the conductivity value is in range from 5 mS/m to 50 mS/m (Lenntech BV 2018). According to wastewater treatment report, the conductivity values do not fulfill drinking water requirements, however, for efficient water treatment these parameters are in optimal range, which is 76-96 mS/m.

Since removal of solids is mechanical only, the technological parameters do not have effect on purification quality. However, Iron Concentration may cause declining of solids removal efficiency due to iron precipitation and proceeding in solid phase. Conductivity may be defined as a measure of the ionic activity of a solution in term of its capacity to transmit current. Dissolved solids are the measure of the different ions in solution, thus in wastewater, dissolved solids and conductivity are reasonably comparable. The predictive model for solids removal efficiency is presented in Figure 32.

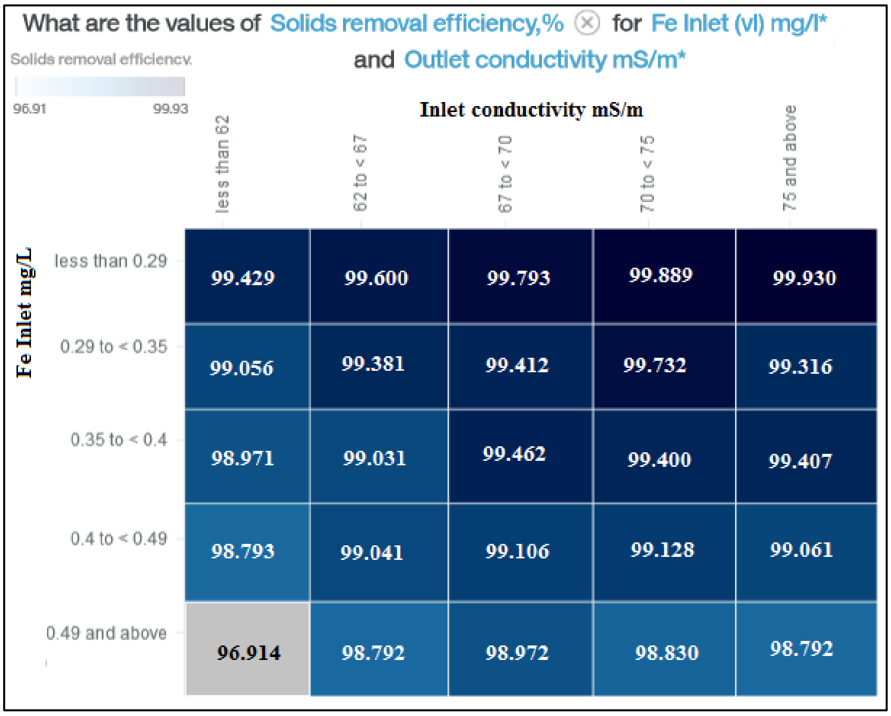


Figure 32-Prediction model for solids removal

Chemical oxygen demand is a measurement, which describes the amount of the oxygen required for oxidation of soluble and macroparticles organic matter in waste water. Therefore, the better the oxidation is, the more efficient is the purification factor. In nitrates the nitrogen atom is highest degree of oxidation, thus increasing amount of Treated NO_3^- indicates efficient COD removal efficiency. The results are presented in Figure 33.

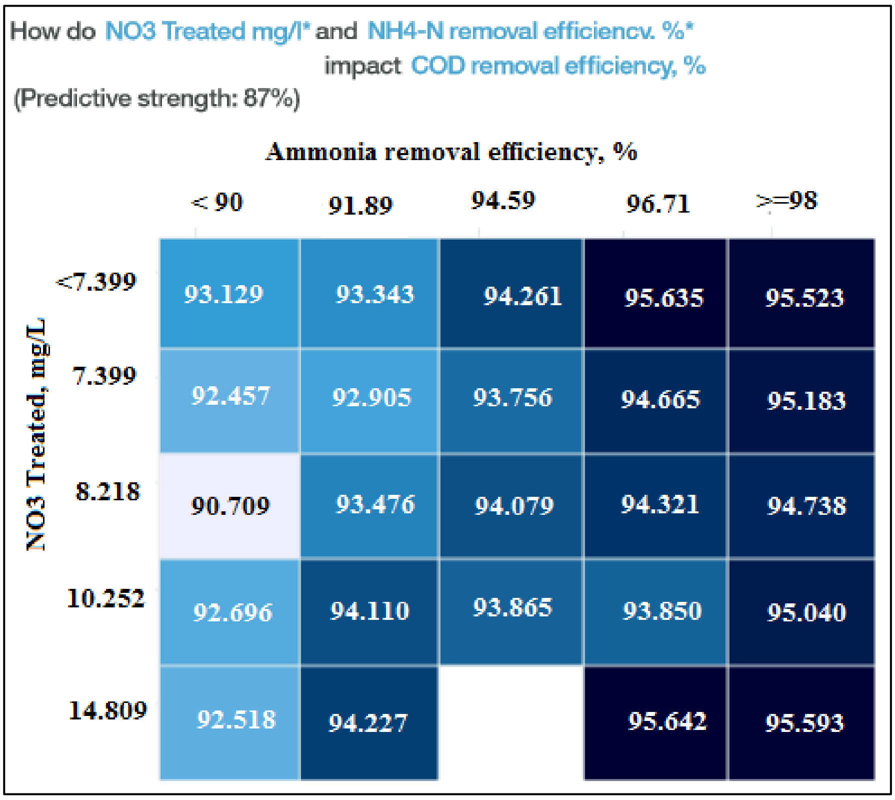


Figure 33- Prediction model for chemical oxygen demand

Ammonia is converted into the nitrates during nitrification, thus ammonia removal efficiency factor also has significant impact on COD value. Efficient nitrification removal efficiency provides less nitrates formation. Therefore, for effective chemical oxygen demand treatment, the nitrogen, ammonia and nitrates should to be controlled.

Biochemical oxygen demand is a quantifiable factor, which describes the required amount of oxygen for bioorganic matter demolition in wastewater. BOD purification factor is similar to COD significant factors. In BOD case phosphorus affects the removal of biological compounds, since phosphorus acts as nutrient for bacteria and other organisms. The predictive model for BOD removal efficiency is presented in Figure 34.

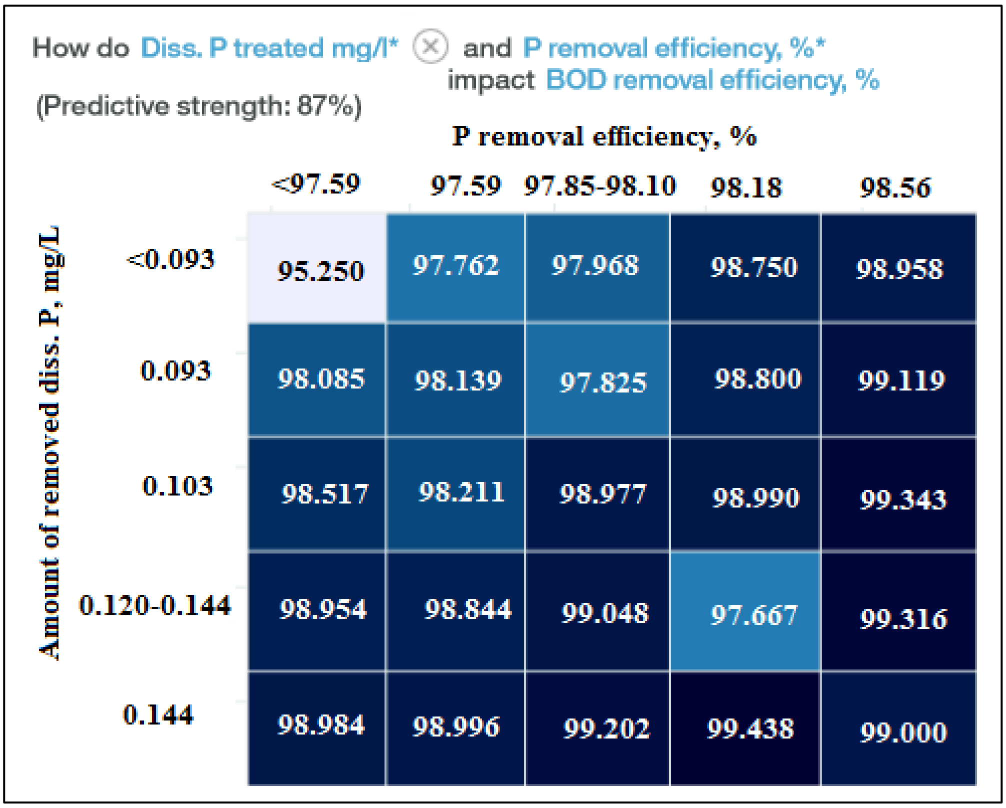


Figure 34-Prediction model for biochemical oxygen demand

The increasing amount of removed dissolved phosphorus leads to improvement of the treatment effectiveness in terms of BOD. The similar effect has the efficiency of phosphorus removal. Therefore, it may be conducted, that determination of phosphorus amount in wastewater is one of significant steps in water pretreatment processes since it may cause growth of biological organisms.

6 MODDE Pro

The MODDE Pro is a software, which is used for experimental design and data analysis, developed by Swedish company "Umetrics". MODDE Pro software uses two regression methods for fitting a model to the data: multiple linear regression (MLR) and partial least squares regression (PLS). Both these models may provide prediction of one or several dependent variables, which are also called responses (analyzed value). A set of independent variables called the factors (influencing factor).

Multiple linear regression minimizes the squared deviation between the dependent variables, and the linear prediction model. MODDE uses the singular value decomposition to obtain the regression coefficients (Sartorius Stedim Data Analytics 2017).

With poorly conditioned problem or when the variables are correlated with each other, regularization or sorting is required for the model stabilization and limit overfitting. For these cases Partial Least Squares (PLS) regression is used, which may be defined as a hidden variable method that provides regularization by representing the dependent variables by means of a small set of linear combinations (the hidden variables). And as a result, the dimension of the space spanned by the independent variables is reduced. PLS regression also provides improved interpretation of the regression model (Sartorius Stedim Data Analytics 2017).

The numerical rating of the model quality is coefficient R^2, describes a variance of data compared to model. Q^2 coefficient estimates the predictive ability of the model. R^2 coefficient specifies the fraction of the response, provided by the model and may be defined by the following equation (Rybinska 2016):

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (21)$$

Where

SS_{res}- sum of squares of the residual;

SS_{tot}- total sum of squares;

y_i- data variable;

$\bar{y}_i$- mean value of the variable;

$\hat{y}_i$- predicted by the model variable.

The Q^2 coefficient estimates the model predictive ability and may be defined according to the following equation (Rybinska 2016):

$$Q^2 = 1 - \frac{Press}{SS_{tot}} = 1 - \frac{Press}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (22)$$

Where

$Press$- prediction residual sum of squares;

SS_{tot}- total sum of squares;

y_i- data variable;

$\bar{y}_i$- mean value of the variable.

This program provides analysis of the experiment and risk estimation data, which mean the determination of the optimal parameters. It is also possible to analyze the values with have the most significant impact on the studied parameters – similar feature as in IBW Watson Analytics (Driver values). However, this application imposes the following constrains on data analysis:

- The program provides the opportunity to analyzed only 32 rows of data;
- The program requires high data quality, single missing value is unacceptable;
- Simultaneous analysis of overall factor influence is impossible;

- Each study requires initial input of “Response” (analyzed parameter), it is desirable to select one or several values for research;
- The increasing of analyzed parameters caused the enhancing relevance of the results, however, the quantity of values more than eight does not provide efficient risk estimation, therefore, the optimal combination of two or three responses is recommended;
- For adding new “Responses” to the study it is required to create new document, since integrate new analyzed parameters into existing model is impossible;

After uploading the data in the application, it is required to set the “Name”, which is the comparative quantity variable (number of rows/days), and “Responses”, which are defined as the studied parameters. For analyzing experimental parts, for pattern recognition and risk estimation all studied parameters must be set as “quantitative”.

For more efficient visualization schematic results MODDE Pro program abbreviate the parameter’s names, therefore, the value description is presented in Table 13.

Table 13-Values description

Parameter	**Unit**	**Abbreviation**	**Range**	**Average**
Inlet flowrate	m^3/d	Inl	54 860 – 146 100	100 480
Temperature	°C	Temp	9.44 – 20.38	14.91
Inlet alkali concentration	mmol/l	In2	4.132 – 7.275	5.703
Outlet alkali concentration	mmol/l	Out	1.00 – 2.77	1.886
Inlet conductivity	mS/m	In3	72.75 – 114	93.38
Inlet pH	-	In4	7.391 – 7.617	7.504
Inlet COD	mg/l	In5	439 – 1043	741
Outlet COD	mg/l	Ou2	31.89 – 48.96	40.42
Inlet BOD	mg/l	In6	200.5 – 453.3	326.9
Outlet BOD	mg/l	Ou3	1.625 – 7.037	4.331
Inlet P total concentration	mg/l	Int	5.58 – 12.58	9.08
Outlet P total concentration	mg/l	Ou4	0.1104 – 0.2119	0.1612
Dissolved P treated concentration	mg/l	Dis	0.0738 – 0.1777	0.1258
Total inlet Nitrogen concentration	mg/l	Tot	41.14 – 89.33	65.24
Total outlet Nitrogen concentration	mg/l	To2	8.77 – 23.11	15.94
NH4-N Inlet concentration	mg/l	NH4	30.05 – 68.45	49.205
NH4-N Outlet concentration	mg/l	NH2	0.56 – 16.31	8.43
NO_2^- Treated concentration	mg/l	NO2	0.031 – 5.175	2.603
NO_3^- Treated concentration	mg/l	NO3	5.15 – 763.10	384.12

Continuation of table 13

Inlet solids concentration	mg/l	In7	202.3 – 465.8	334,053
Outlet solids concentration	mg/l	Ou5	0.523 – 5.172	2.848
Iron concentration	mg/l	Fem	0.221 – 0,895	0.558
Dissolved Fe concentration	mg/l	Di2	0,117 – 0,251	0.184
Added Fe sulfate concentration	mg/l	Add	71.0 – 216.0	143.5
Added polymers concertation	mg/l	Ad2	0.6 – 3.0	1.8
COD removal efficiency	%	COD	89.74 – 96.43	93.09
BOD removal efficiency	%	BOD	95.00 – 99.68	97.34
Nitrogen removal efficiency	%	Nre	55.45 – 89.75	72.60
Phosphorus removal efficiency	%	Pre	96.52 – 98.92	97.72
Solids removal efficiency	%	Sol	97.39 – 100.00	98.70
Ammonia removal efficiency	%	NH3	56.97 – 99.08	78.03
Nitrification	%	Nit	69.80 – 99.33	84.57

The MODDE Pro program focuses on determination of significant values, which are affect purification efficiency (Nitrification, COD, BOD, Nitrogen-contained compounds, Phosphorus and solids), the similar approach as IBW Watson Analytics. After model generation, the comparison analysis may be conducted.

6.1 Inlet parameters

Similar to approach, which were applied in IBM Watson Analytics case, the flowrate, inlet iron concentration and inlet alkali concentration were set as inlet parameters. In this chapter these parameters influence on removal efficiencies were analyzed by means of MODDE Pro for comparative analysis with IBW Watson Analytics as month-based 5 years data.

One of the most relevant features of MODDE Pro program is “Contour”, which determines the worst- and best-case scenarios for the analyzed factor and compare it with the Target value. Target is the variable, which is set by the user according to quality requirements of studied values (in this case - efficiency of impurities removal). The influence of the inlet parameters, such as Temperature and Inlet flowrate, is set by the user and other variable are considered as the constants. The constants variables located in the right corner of the pattern (Figures 35 and 36). The contour plot for COD removal efficiency is presented in Figure 35.

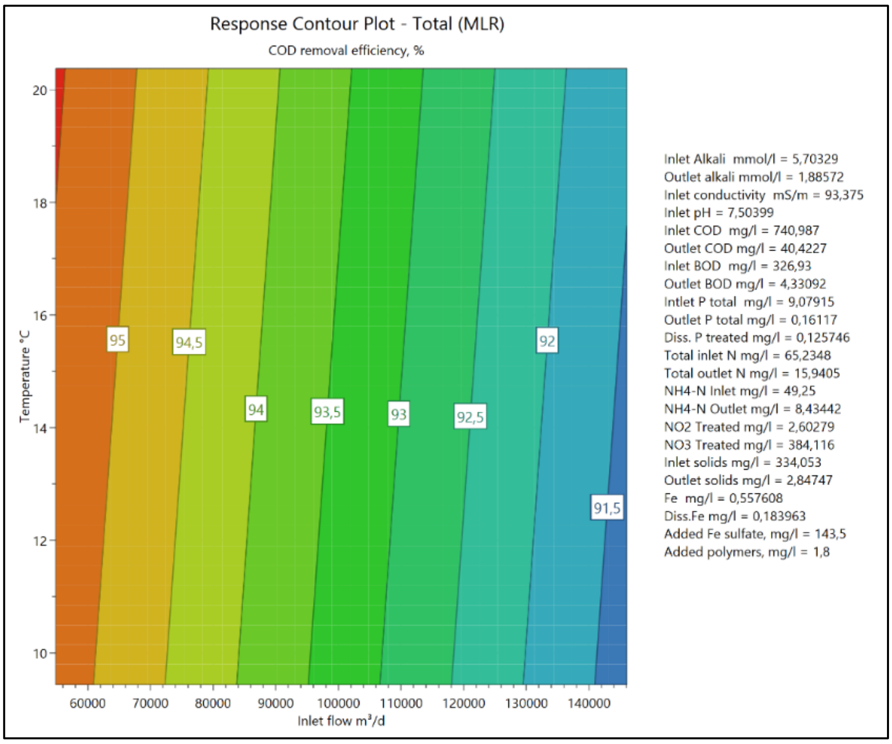


Figure 35-Design space results for COD, R^2=0.67, Q^2=0.65

According to the results, it may be conducted that the temperature and inlet flowrate have impact on the COD removal efficiency, which is similar to the IBW Watson Analytics obtained results, however, the impacts of the other parameters are not including, since they were set as constants. The same analysis for BOD is presented in Figure 36.

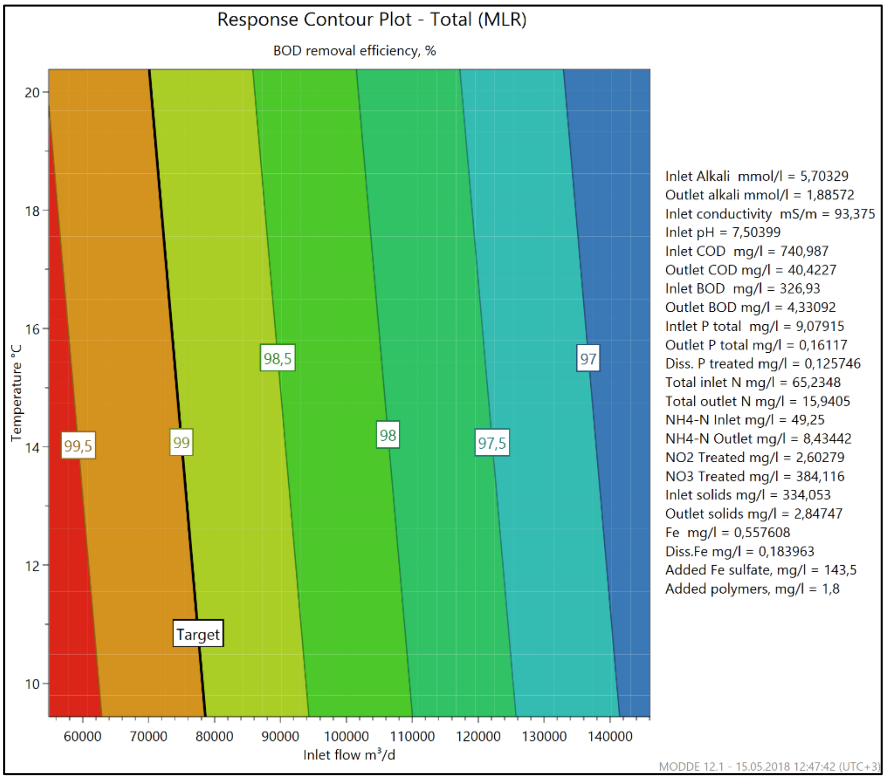


Figure 36-Design space results for BOD, R^2=0.67, Q^2=0.65

The same pattern is occurred for BOD removal efficiency, as for COD. Therefore, it may be conducted that the temperature and inlet flowrate have significant impact on purification processes.

Other relevant features of MODDE Pro software is "Risk Estimation", which provides the most efficient for parameter optimization After analyzing of two parameters (in this example Chemical and Biochemical oxygen demand were studied) the model with lower quality was obtained. It is worth noting that risk estimation for combination of two parameters with using the most significant values provides the pattern, which may be used for optimal parameters determination. The design space for COD and BOD parameters is presented in Figure 37.

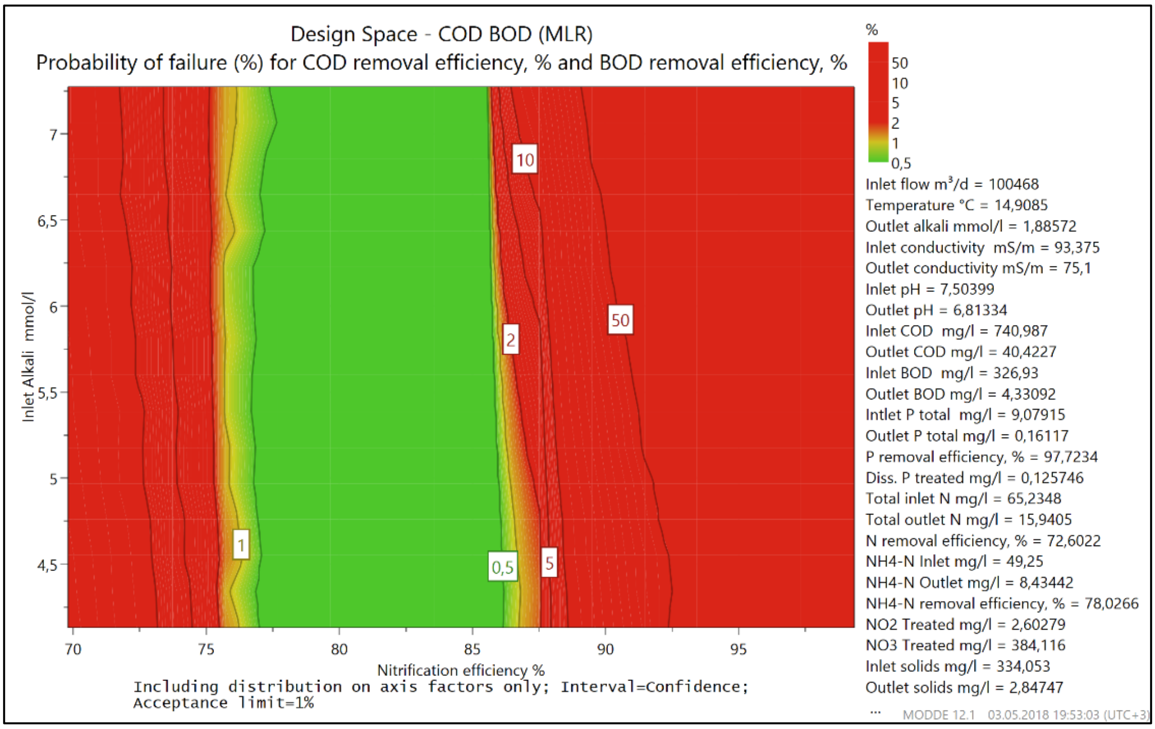


Figure 37-Design space plot for COD and BOD (Nitrification/Alkali)

Inlet alkali concentration and Nitrification were set as influenced factor; however, nitrification is outlet parameter, which is not affect process conditions. MODDE Pro conducts the risk estimation according to all variables, thus the separation for process parameters and results is requires before the initialization of the factors. The design space for COD and BOD is presented in Figure 38.

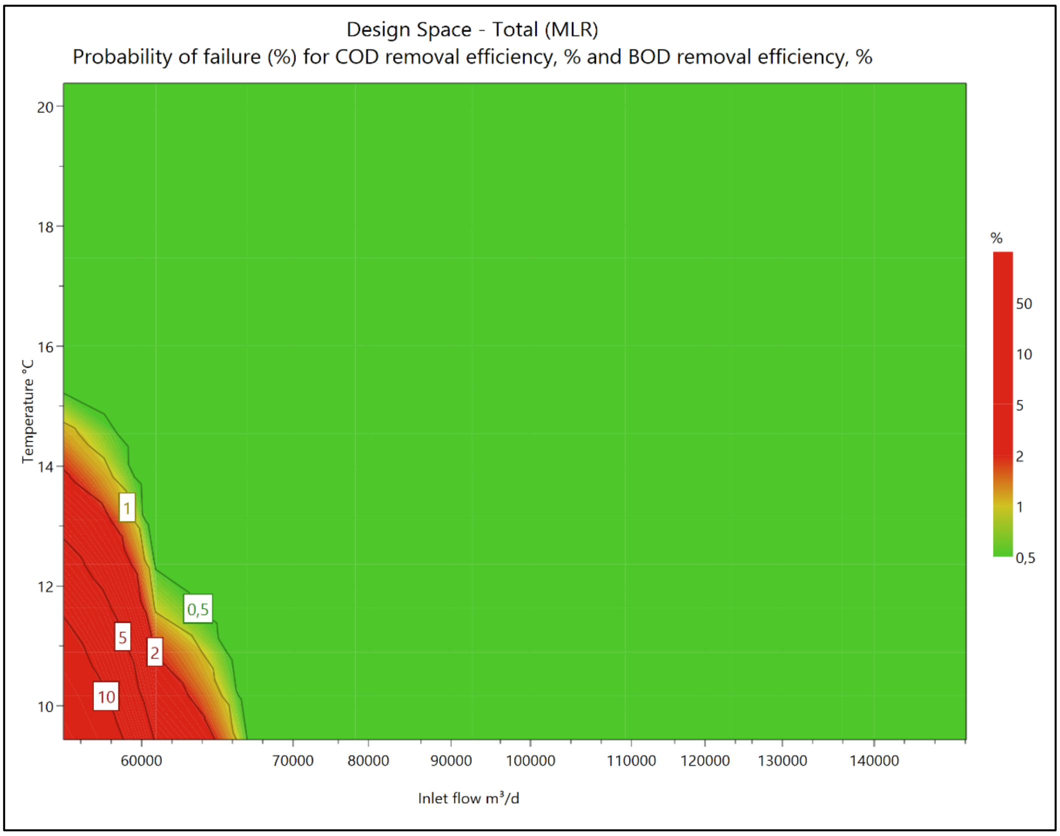


Figure 38-Design space plot for COD and BOD (Flowrate/Temperature)

The obtained results for COD BOD do not correlated with results, provided by IBW Watson Analytics in case of flowrate and temperature, since in MODDE Pro other parameters were set as constants.

6.2 Outlet parameters

The tab page “Coefficients” provides the parameters, which has significant impact on the studied parameters. As an example, the single index of water purification Nitrification efficiency was set. The Coefficients are presented in Figure 39. The combination of thin lines and bars provides the significance of the values of the model:

- High bars with parameter confidence intervals determine the significance of the studied parameter as “high”, the example is marked by red shape (Figure 39);
- Small bars with parameter confidence intervals determine significance of the studied parameter as “low”, the example is marked by blue shape (Figure 39).

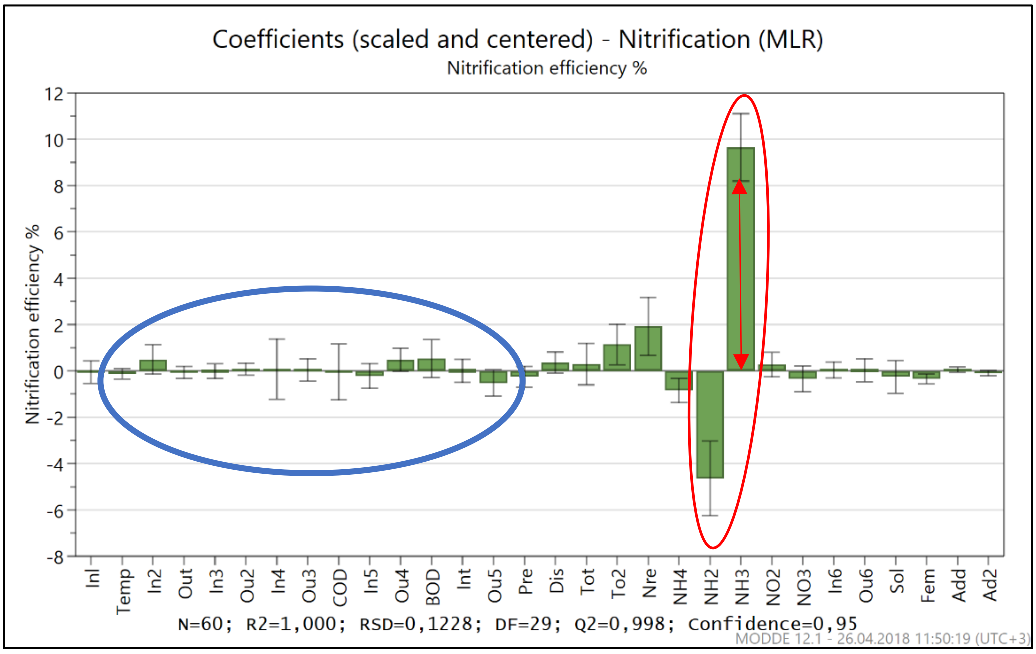


Figure 39-Coefficients for Nitrification

Using "Auto-Tune" function, which are automatically remove non-significant terms from the model, the using of the coefficients are more efficient and accessible. The updated results are presented in Figure 40.

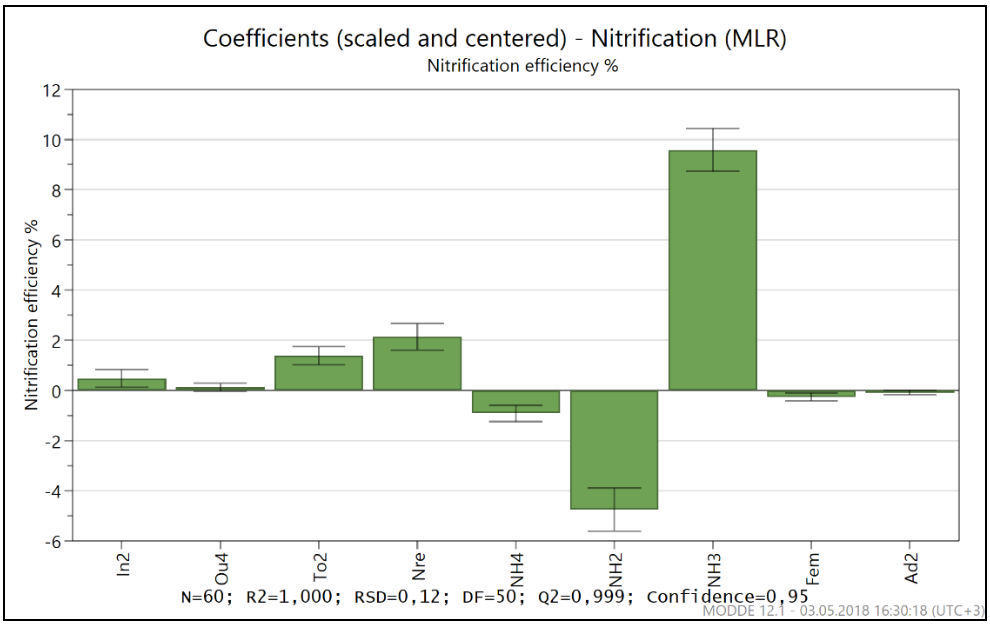


Figure 40-Relevant coefficients for Nitrification

As it was described in previous chapters, the R^2 and Q^2 coefficients define the fitting of the model to the studied values. The coefficients close to 1 indicate model fit the data with a high accuracy. According to the results, the most valuable parameters for nitrification efficiency are the following: Ammonia Inlet and Outlet concentration, the degree of Ammonia and Nitrogen purification and concentration of inlet Nitrogen. The results are similar to IBM Watson Analytics; however, the influence of other parameters combinations has not been observed.

Observed vs. Predicted plots describe the qualities of the studied models. Plots with values close to the straight line indicates efficient model, which may provide relevant results. For one parameter (Nitrification efficiency removal) the observed vs. predictive model the pattern is as presented in Figure 41.

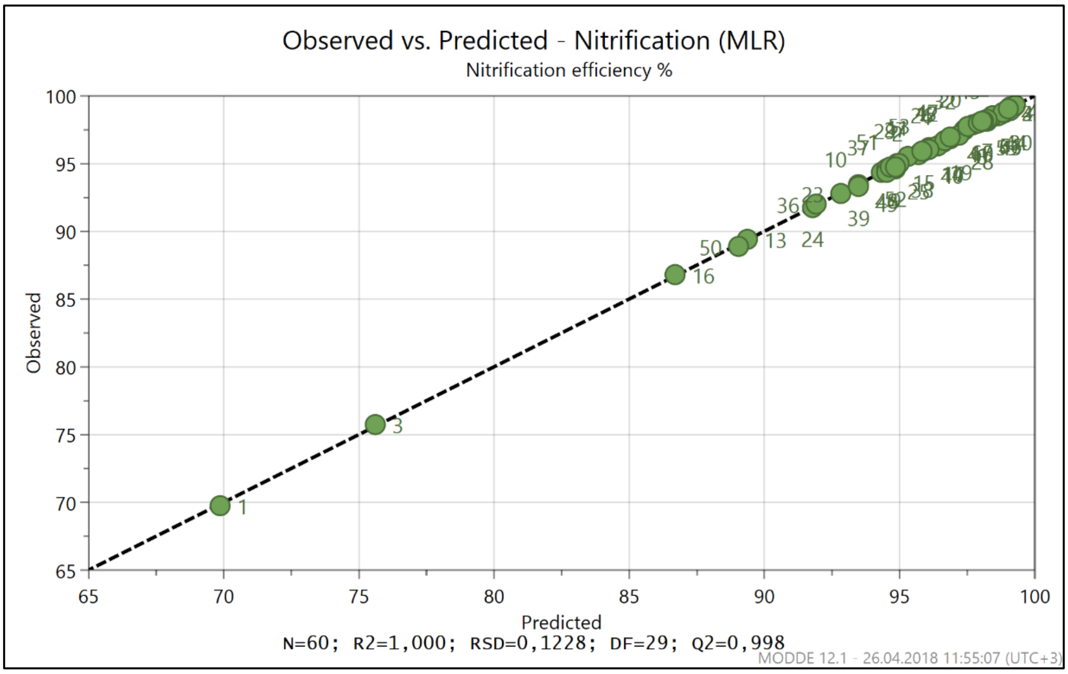


Figure 41-Observed vs Predicted results for Nitrification

Analysis of the data shows that it is the high-quality model with observed parameter range, as evidenced by near the straight-line distribution of values, parameter Q^2 equals to 0.988 and R^2 equals 1. Therefore, analyzing one parameter provides high quality model, however, it is desirable to compare simultaneous impacts of different wastewater characteristics on efficiency of impurities removal.

6.3 Clarification of the results

The risk estimation may be conducted by using the most significant parameters. However, the simultaneous analysis of all parameters of purification efficiency may provide illustrative example in combination significant and non-significant value impacts.

For comparison with IBM Watson Analytics, the Inlet flowrate and Inlet alkali concentration, which were one of the most significant drivers, were also implemented in MODDE Pro as axis in risk estimation approach. According to Figure 43 the Inlet flowrate **(Inl)** and Temperature **(Tem -** this value was not even considered as significant by MODDE Pro “Auto tune” feature) have no strong impact on studied purification parameters. the risk estimation for all seven purification factors is presented in the Figure 42.

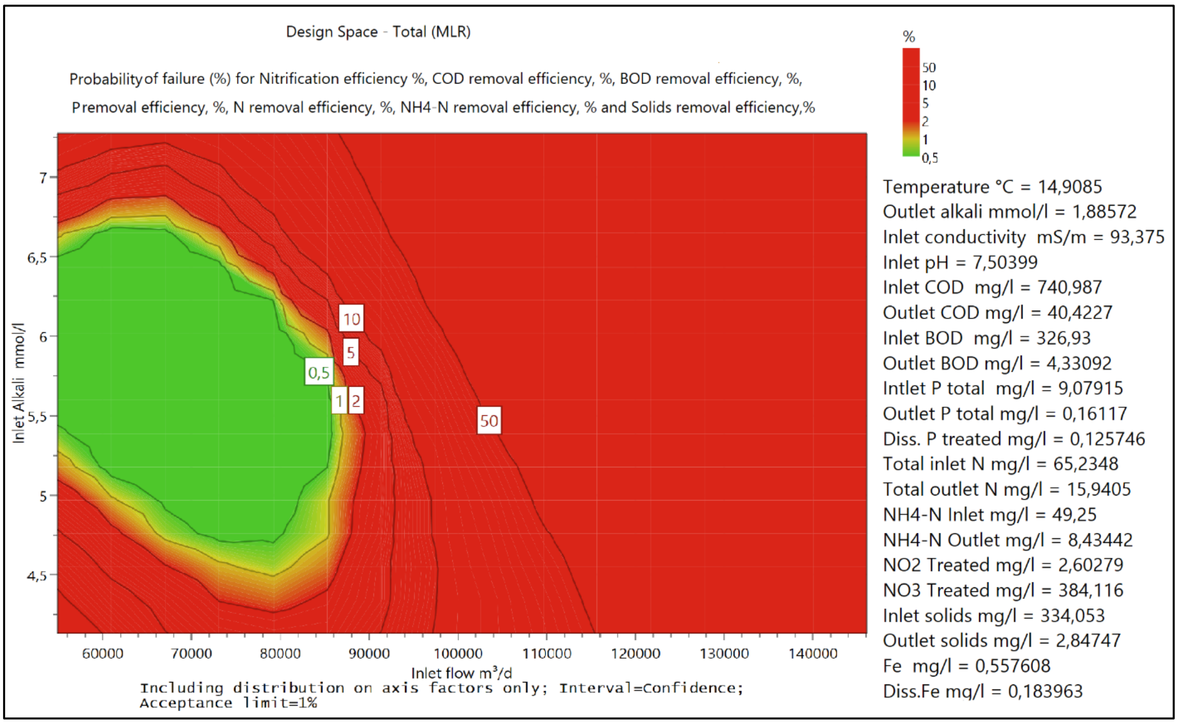


Figure 42-Design space plot for all parameters

Despite of the fact, that the Inlet alkali concentration and Inlet flowrate have not significant impact on purification factor separately, the combination of them provides relevant risk estimation. These results allow to determine the optimal maintenance parameters (concentration of added alkali and inlet flowrate), which are fitted to all removal efficiency factors.

7 Comparative analysis of IBM Watson Analytics and MODDE Pro

Comparing the results and possibilities of IBM Watson Analytics and MODDE Pro, it may be conducted that each of the program has their own limitations and possibilities. All these points may be divided according to the following subtitles:

1. Storage, software program location and serviceability

The IBM Watson Analytics is a cloud-based software, which provides personal account with virtual data storage, therefore, for the result achievements only Internet access is required. The cloud storage provides convenient management and sorting of the data in personal account, thus the access to the user's data is possible at any connected device.

MODDE Pro is a software, which require installation on the personal computer hard drive. Therefore, additional process load is occurred during the risk estimation and model fitting to the data. Unavailability of cloud storage reduces the range and demand specific platform for research.

2. Data importing

As it was mentioned above, the MODDE Pro does not allow to import data with missing values, since for the model generation each variable is used. At the same time, it is impossible to use for research databases with more than 32 number of columns (factors and studied parameters). Described limitations demand high quality of data preparation, which may cause difficulties with large amount of variables.

IBM Watson Analytics does not have high requirements to the data quality, therefore it is possible to upload or import files with missing variables. However, with high amount of missing values may deteriorate pattern recognition results.

3. Features

It is important to note the possibility of IBM Watson Analytics generation of predictive model. This feature automatically determines the significant variables and predict the conduct of studied parameter according to Driver values fluctuations.

On the most useful features of MODDE Pro is risk estimation, which provides the most efficient for parameter optimization. It is possible to study the impact of several parameters on efficiency of the wastewater treatment processes.

4. Data analyzing

The IBM Watson Analytics obtains efficient results regarding determination of the driver values, their combination and index number of significance (Strength); pattern recognition and visualization; predictive of the data variation.

MODDE Pro program is focused on optimization of the laboratory results, determination of the most significant values and their impact, estimation of possible consequences of variable fluctuation and propose of the most optimal values in the furtherance of researching goals.

Therefore, it may be conducted that both these programs provide efficient results and user-friendly visualization. In accordance with the ultimate objective and data quality, the optimal program should be selected for obtaining relevant results.

7.1 Comparison of data analysis result

Since both IBM Watson Analytics and MODDE Pro are able to provide similar results, the description of data analysis was conducted in this chapter. The temperature and flowrate influence on COD and BOD purification results was studied and compared.

In case of the significant values determination, both programs provide similar results. The strongest influence for COD (as an example) has inlet COD concentration, which is obvious and irrelevant in this case since the high inlet concentration requires the stronger purification. The same results regarding drivers were obtained in MODDE Pro. However, IBM Watson has the possibility to combine drivers influence; thus more precise dataset association may be conducted.

The flowrate is the process control parameter, which is possible to maintain according to the required processing rate. The temperature is external parameter, the opposite of the flowrate case. Since the IBM Watson Analytics uses whole dataset for analyzing and drivers determination, the correlation between temperature and flowrate has not been discovered. However, the separate analysis of these values and their influence on COD and BOD provide relevant results, which are presented in Figures 22-23 for the flowrate and in Figures 26-27 for the temperature.

The Inlet flowrate and Inlet alkali concentration are one of the most significant drivers in both programs. In MODDE Pro separately these values do not have strong impact on studied purification parameters. The combination of them provides relevant risk estimation, which is presented in Figure 42. The similar results were obtained in IBM Watson Analytics, which are presented in Appendix 3: the alkali concentration lower than 4.5 mmol/L provide insufficient degree the purification.

In MODDE Pro for data analyzing the factors, which affect the responses, are set, and other parameters set as constants, thus the simultaneous influence of all variables are not possible to study. However, two factors provide the relevant results in terms of risk estimation and contour plot. The simultaneous influence of temperature and flowrate are similar to IBM Watson Analytics results: high flowrate and low temperature decline the efficiency of the purification. The MODDE Pro results are presented in Figures 35 and 36.

The predictive model in IBM Watson Analytics (or the decision tree, Figure 32) automatically applies the most significant drivers of the studied factor, which means the whole dataset are considered. However, if the variable is not determined as “Significant”, the decision tree does not include it into the model, only the predictive model with 2 parameters (Figures 32-34) may be set. Since in MODDE Pro the responses are defined, the design space and contour conduction require factors determination. The influence on all responses may be studied simultaneously, while only two factors are set.

CONCLUSIONS

Process parameters and different chemicals may cause the fluctuations of the wastewater purification efficiency. Nitrogen and phosphorus contained compounds promote the growth of the hazardous bacteria and pathogens. The results comparison and analysis of impurities removal may provide more efficient purification processes and reduction of wastewater treatment plants operation costs due to lower energy and chemicals consumption.

Since the manual data processing may be time-consuming, and sometimes even impossible to analyze information for several years, the application of platforms, such as IBW Watson Analytics and MODDE Pro, provides effective databases analysis, determination of interdependent and optimal process parameters, and describes most significant values and their influence on wastewater purification processes.

According to the results, provided by IBM Watson Analytics the most significant technological parameters are inlet flowrate and temperature. High values of flowrates may provoke declining of all purification factors, thus for all studied parameters of treatment efficiency, the optimal range of flowrate is from 79 000 m^3/d to 90 000 m^3/d. The similar results were obtained in MODDE Pro's feature "Risk Estimation", which is presented in Figure 43. Temperature, as technological parameter, does not significantly affect water treatment processes. However, after using Temperature in combination with the flowrate as factor, MODDE Pro determines the optimal value as14-16 ^oC (Figures 36 and 37) for COD and BOD. IBM Watson results do not indicated temperature as driving factor for COD and BOD, but for nitrogen-contained compounds. The highest impact flowrate and temperature have on Nitrification and nitrogen removal particles. The significant impact may be caused sensitive conditions of nitrogen-containing compounds removal reactions.

Added chemical, such as Iron Sulfate and alkali also influence the water treatment efficiency due chemical reactions, or precipitation of constituents. The concentration of inlet compounds may be highly ranged and does not correlate with processes parameters. High iron concentration may cause declining of solids removal efficiency due to iron precipitation and proceeding in solid phase Thus the optimal amount of added chemical is required to avoid colloidal particles formation. The concentration of addition alkali in wastewater is tracking due their impact on water pH level. In initial part of the purification processes alkali as treatment chemical added to the water for neutralization of acidic ions H^+. Small amount of inlet alkali decreases the efficiency of wastewater purification. During nitrification, the acid formation occurs, however, these acids may be neutralized by the added alkali. Thus, the optimal value of these chemicals need to be maintained in terms of increasing water treatment degree of purification and compliance of pure water requirements. The similar results were obtained in IBM Watson Analytics (Figure 27, chapter 5.1.2, Appendix 3) and in MODDE Pro (Figures 38 and 43). Low values of alkali decrease BOD and COD treatment in wastewater. The optimal value of alkali concentration is in range from 4.7 to 7.6 mmol/l of water. As it was mentioned before, the further increase may cause strong alkaline conditions of purified water.

Comparing the IBM Watson Analytics MODDE Pro, the data specification should be taking into account, since both programs use different algorithms for obtaining efficient results. Watson Analytics may consider simultaneous influence of all variables in whole dataset, which provides wide database connections, however it requires significant amount of data. MODDE Pro algorithms focuses on result optimization, determination of the most significant values and their impact, estimation of possible consequences of variable fluctuation. The drawback of this method is that part of the factors considered as constants. In accordance with the ultimate objective and data quality, the optimal program should be selected for obtaining relevant results. Both IBM Watson Analytics and MOODE Pro may be applied as data analyzing program for optimization of wastewater treatment processes, and also in other manufactured sectors, where process databases are severely maintained.

References

1. BANSAL, M. 2013, "Surface Web Semantics for Structured Natural Language Processing", .

2.Chen, Y. 2016, "IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research", *Clinical therapeutics*, vol. 38, no. 4, pp. 688-701.

3. Crittenden, J.C., Trussell, R.R., Hand, D.W., Howe, K.J. & Tchobanoglous, G. 2012, "Membrane filtration" in , pp. 819-902.

4. de Graaff, M.S. 2010, "Long term partial nitritation of anaerobically treated black water and the emission of nitrous oxide", *Water research*, vol. 44, no. 7, pp. 2171-2178.

5. FRC, S. 2018, , *Trusted wastewater solutions*. Available: http://frcsystems.com/pcl-dissolved-air-flotation-systems [2018, 01.23].

6. Guidi, G., Miniati, R., Mazzola, M. & Iadanza, E. 2016, "Case study: IBM watson analytics cloud platform as analytics-as-a-service system for heart failure early detection. *Future Internet*", , no. 8(3), pp. 32.

7. Harshman, V. 2000, "Wastewater odor control: An evaluation of technologies", *Water Engineering & Management*, vol. 147, no. 5, pp. 34-46.

8. Ho, L., Lambling, P., Bustamante, H., Duker, P. & Newcombe, G. 2011, *Application of powdered activated carbon for the adsorption of cylindrospermopsin and microcystin toxins from drinking water supplies*.

9. Hu, M. 2016, "Constructing the ecological sanitation: a review on technology and methods", *Journal of Cleaner Production*, vol. 125, pp. 1-21.

10. IAEA 1992, *Chemical precipitation processes for the treatment of aqueous radioactive waste*, International Atomic Energy Agency, Vienna.

11. IBM 2018, , *System Designed for Answers: the Future of Workload Optimized Systems Design.* Available: https://www.ibm.com/us-en/ [2018, 02.26].

12. Jimenez, B. 2010, "Wastewater treatment for pathogen removal and nutrient conservation: suitable systems for use in developing countries", *IDEAS Working Paper Series from RePEc*, .

13. Jimenez, B. 2007, "Helminth ova removal from wastewater for agriculture and aquaculture reuse", *Water science and technology : a journal of the International Association on Water Pollution Research*, vol. 55, no. 1-2, pp. 485.

14. Karimi, A.A. 1997, "Evaluating an AOP for TCE and PCE removal", *Journal of the American Water Works Association*, vol. 89, no. 8, pp. 41-53.

15. Kukreja, R. 2009, , *What is waswtewater treatment*. Available: https://www.conserve-energy-future.com/process-of-wastewater-treatment.php [2018, 01.25].

16. Lapertot, M. 2008, "Evaluating Microtox © as a tool for biodegradability assessment of partially treated solutions of pesticides using Fe 3+ and TiO 2 solar photo-assisted processes", *Ecotoxicology and environmental safety,* vol. 69, no. 3, pp. 546-555.

17. Lawrence K. Wang 2006, *Advanced Physicochemical Treatment Processes,* .

18. Lenntech BV 2018, , *Water Treatment Solutions* [Homepage of Lenntech BV], [Online]. Available: https://www.lenntech.com/applications/ultrapure/conductivity/water-conductivity.htm [2018, 05.06].

19. Marietta, M. 2015, *The Role of Alkalinity in Aerobic Wastewater Treatment Plants,* Quality system edn, Magnesia Specialties, LLC, Baltimore, Maryland 21236 USA.

20. Nasr, F.A., Doma, H.S. & Abdel-Halim, H.S. 2007, "Environmentalist", vol. 27, pp. 255.

21. Purchas, D. & Sutherland, K. 2008, "Filter Media-Section 2" in .

22. Rahman, S.M. 2016, "Life-Cycle Assessment of Advanced Nutrient Removal Technologies for Wastewater Treatment", *Environmental science & technology,* vol. 50, no. 6, pp. 3020.

23. Rajbansi, B. 2014, "Hazardous odor markers from sewage wastewater: A step towards simultaneous assessment, dearomatization and removal", *Journal of the Taiwan Institute of Chemical Engineers,* vol. 45, no. 4, pp. 1549-1557.

24. Ratchford, I.A.J. 2003, "The effect of light:dark cycles of medium frequency on photosynthesis by Chlorella vulgaris and the implications for waste stabilisation pond design and performance", *Water science and technology : a journal of the International Association on Water Pollution Research,* vol. 48, no. 2, pp. 69.

25. Res, J. 2014, "Nitrification In Wastewater Treatment Plants", vol. 63, pp. 198-207.

26. Russell, D.L. 2006, *Practical wastewater treatment,* Wiley, Hoboken (NJ).

27. Rybinska, A. 2016, "Geometry optimization method versus predictive ability in QSPR modeling for ionic liquids", *Journal of computer-aided molecular design,* vol. 30, no. 2, pp. 165-176.

28. S. S. Murtaza, P. Lak, A. Bener & A. Pischdotchian 2016, "How to Effectively Train IBM Watson: Classroom Experience", *2016 49th Hawaii International Conference on System Sciences (HICSS),* pp. 1663.

29. Sartorius Stedim Data Analytics 2017, *User Guide to MODDE Pro,* 12th edn, Sartorius Stedim Data Analytics AB, Umeå, Sweden.

30. Srinivasan, A., Chowdhury, P. & Viraraghavan, T. 2009, "Air Stripping in Industrial Wastewater Treatment" in *Water and Wastewater Treatment Tecnologies,* Bioresource Technologies edn,Elsevier SCI ltd, the boulevard, Langford lane, Killington, oxford ox5 1gb, Oxon, England, pp. 439-449.

31. Sun, J. 2003, "Co-removal of hexavalent chromium through copper precipitation in synthetic wastewater", *Environmental science & technology,* vol. 37, no. 18, pp. 4281.

32. Tchobanoglous, G., Tchobanoglous, G., Burton, F.L. & Stensel, H.D. 2003, *Wastewater engineering : treatment and reuse*, 4. ed. / revised by George Tchobanoglous, Franklin L. Burton, H. David Stensel edn, McGraw-Hill, Boston.

33. UNEP 2013, , *United Nations Environment Program Division of Technology, Industry and Economics*. Available: http://www.unep.or.jp/ietc/publications/freshwater/fms1/2.asp [2018, 01.25].

34. Wang, X.L., Lang, J., Zhou, S.S., Zhang, L.L. & Chen, M.X. 2010, "Numerical analysis of solid-liquid two-phase flow on sandstone wastewater of hydropower stations in a rectangular sedimentation tank. *Industrial and Engineering Chemistry Research*", vol. 22, pp. 11714-11723.

35. Wang, X. 2015, "Probabilistic evaluation of integrating resource recovery into wastewater treatment to improve environmental sustainability", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 5, pp. 1630.

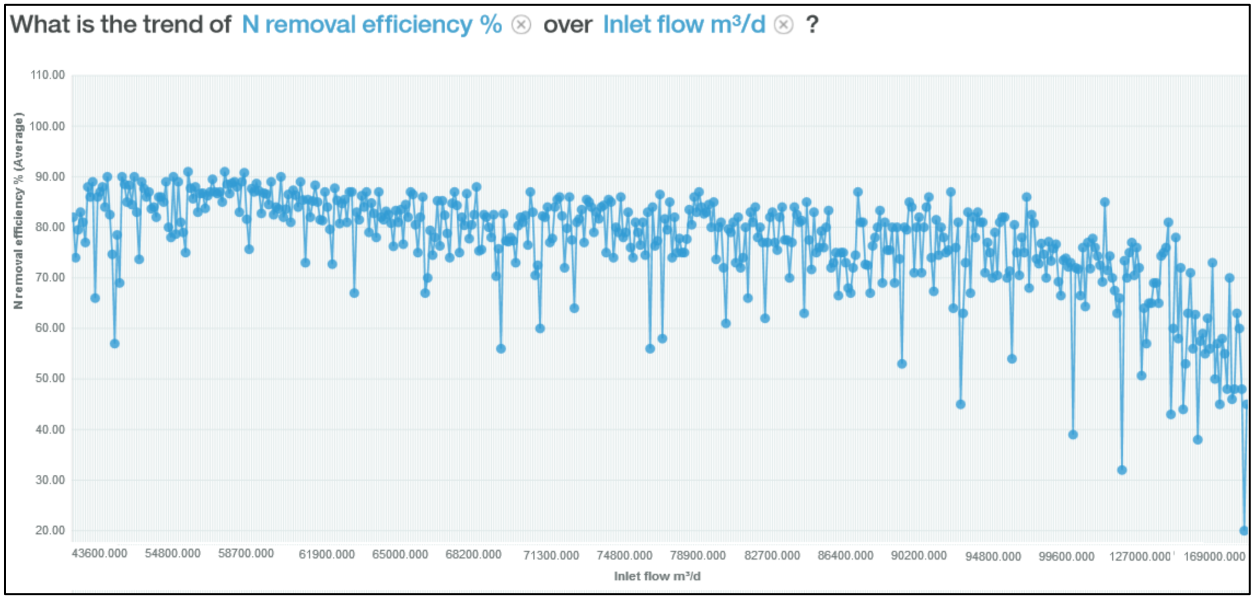
36. Wu, R. 2009, "Titanium dioxide-mediated heterogeneous photocatalytic degradation of terbufos: Parameter study and reaction pathways", *Journal of hazardous materials*, vol. 162, no. 2, pp. 945-953.

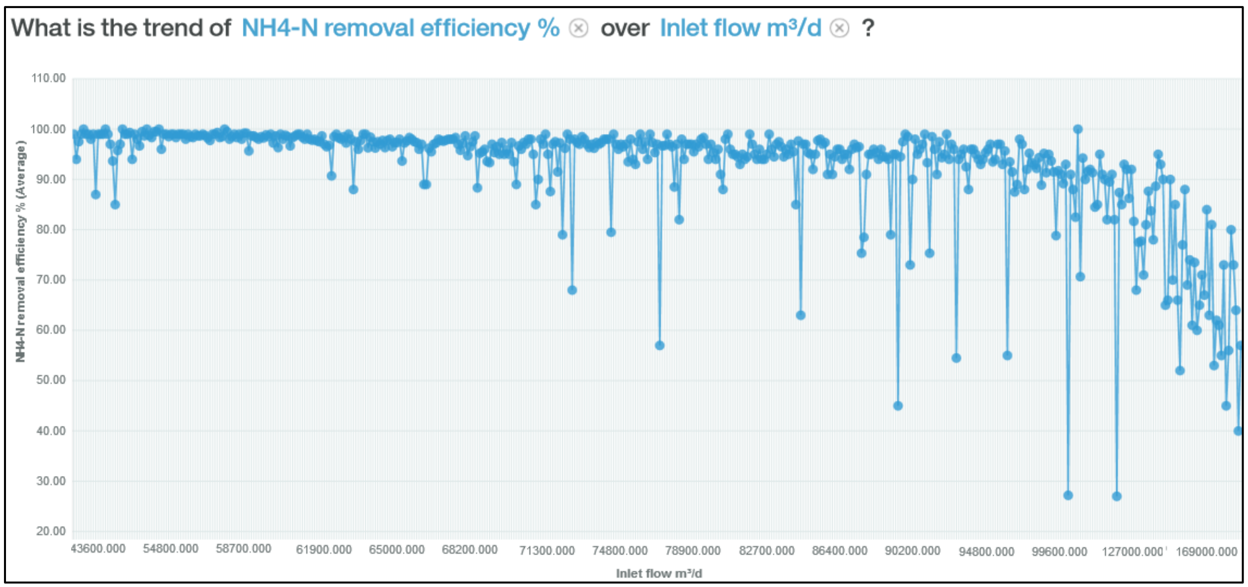
Appendix 1 – Example of overall monthly database for 2012-2013 years

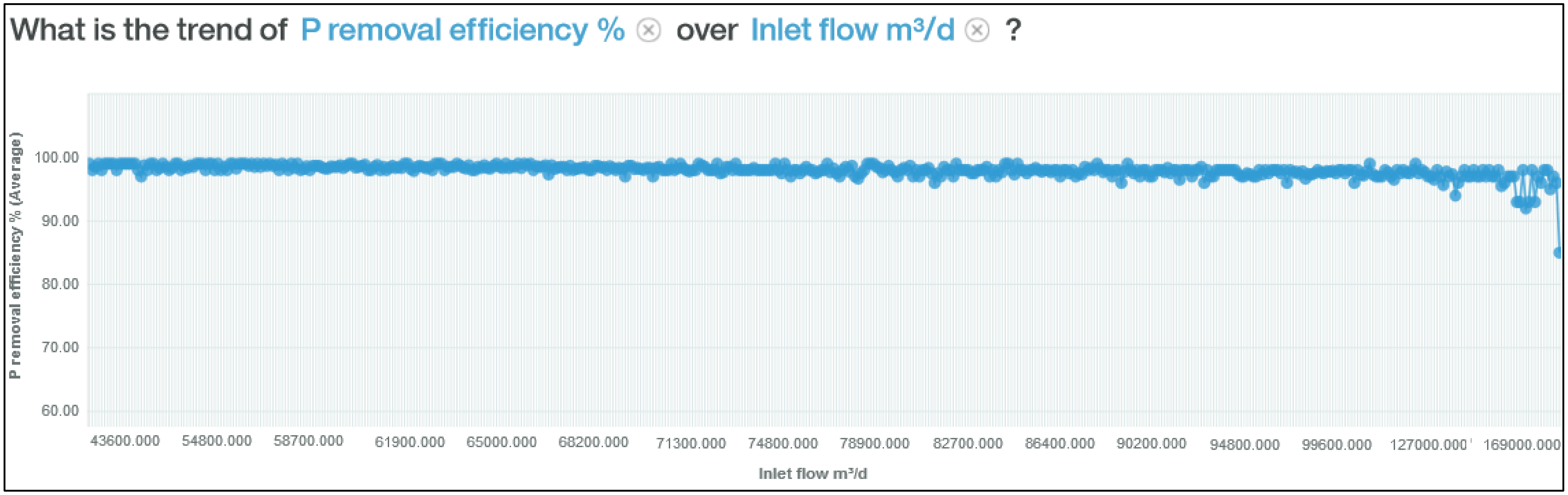
	Year	Month	Inlet flow m³/d	Outlet flow m³/d	Temperature °C	Inlet Alkali mmol/l	Outlet alkali mmol/l	Inlet conductivity mS/m	Outlet conductivity mS/m	Inlet pH	Outlet pH	Inlet COD mg/l	Outlet COD mg/l	COD removal efficiency, %	Inlet BOD mg/l	Outlet BOD mg/l	BOD removal efficiency, %
2	2012	January	103900,000	104233,333	10,962	5,070	2,771	102,900	91,950	7,530	6,385	617,500	32,738	94,650	247,000	3,014	98,500
3		February	66805,000	66805,000	12,720	6,220	1,305	113,850	81,750	7,545	6,625	798,500	35,700	95,450	331,500	2,670	99,000
4		March	124322,727	124322,727	9,925	4,445	1,938	75,636	64,450	7,532	6,838	470,455	32,636	92,636	215,682	4,673	97,409
5		April	98938,889	98938,889	10,839	4,850	1,394	84,000	61,889	7,539	6,683	546,111	31,889	94,056	248,889	2,406	99,000
6		May	77086,364	77086,364	13,918	5,609	1,491	96,636	67,364	7,427	6,768	715,909	41,818	93,909	283,636	3,209	98,955
7		June	72763,158	72763,158	16,126	5,821	1,263	97,579	69,632	7,521	6,847	693,684	41,211	94,158	273,684	2,379	99,000
8		July	69386,957	69473,913	16,852	5,696	1,309	96,957	66,478	7,430	6,961	674,348	36,000	94,696	258,261	3,078	98,783
9		August	79417,391	79417,391	18,759	5,491	1,423	91,522	63,864	7,426	6,950	684,783	38,043	94,261	273,913	2,470	98,913
10		September	81745,000	81745,000	18,330	5,425	1,600	99,000	75,650	7,405	6,960	687,500	39,550	93,950	307,500	3,110	98,800
11		October	122372,727	122372,727	15,386	4,490	1,471	78,000	61,682	7,432	6,759	475,909	37,636	91,682	207,727	3,900	98,000
12		November	100100,000	98736,364	14,032	4,845	1,532	85,727	62,045	7,423	6,859	548,182	36,455	93,227	231,364	3,436	98,500
13		December	68568,750	68568,750	13,994	6,044	1,200	102,750	71,250	7,469	6,750	710,625	44,188	93,500	313,750	3,206	99,000
14	2013	January	108100,000	108100,000	11,439	5,057	1,239	89,913	63,522	7,535	6,709	566,957	39,000	92,696	252,435	3,217	98,043
15		February	69150,000	69150,000	12,485	5,900	1,000	108,800	76,850	7,560	6,605	704,000	45,650	93,450	316,000	3,110	99,000
16		March	66145,000	66145,000	12,520	6,200	1,445	104,300	76,450	7,585	6,750	769,000	47,200	93,750	337,500	2,530	99,050
17		April	146078,947	146078,947	9,435	4,132	1,695	74,737	61,684	7,553	6,732	439,474	40,842	89,737	200,526	7,037	95,000
18		May	77786,364	77786,364	13,436	5,623	1,764	92,955	67,455	7,555	6,864	652,273	41,773	93,318	306,818	2,386	99,136
19		June	62947,368	62947,368	17,205	6,005	1,284	98,368	66,579	7,505	6,821	804,737	44,316	94,421	336,842	2,489	99,000
20		July	54856,522	54856,522	19,426	6,483	1,283	105,565	73,957	7,491	6,852	873,913	48,957	94,304	358,261	3,017	99,000
21		August	64136,364	64136,364	20,382	6,045	1,432	98,455	67,000	7,423	6,855	761,364	38,045	94,909	322,273	2,305	99,318
22		September	61038,095	61038,095	20,095	6,262	1,405	102,714	69,000	7,414	6,848	845,714	41,905	95,048	345,238	3,010	99,000
23		October	73552,174	73552,174	18,017	6,039	1,657	101,565	70,522	7,391	6,891	813,478	40,870	94,739	334,783	3,583	98,696
24		November	118280,000	118280,000	14,375	4,480	1,665	77,350	63,100	7,415	6,740	527,000	39,350	92,450	210,500	5,495	97,150
25		December	110222,222	110222,222	12,844	4,550	1,867	88,278	72,389	7,394	6,828	531,111	37,083	92,778	226,333	3,267	98,278

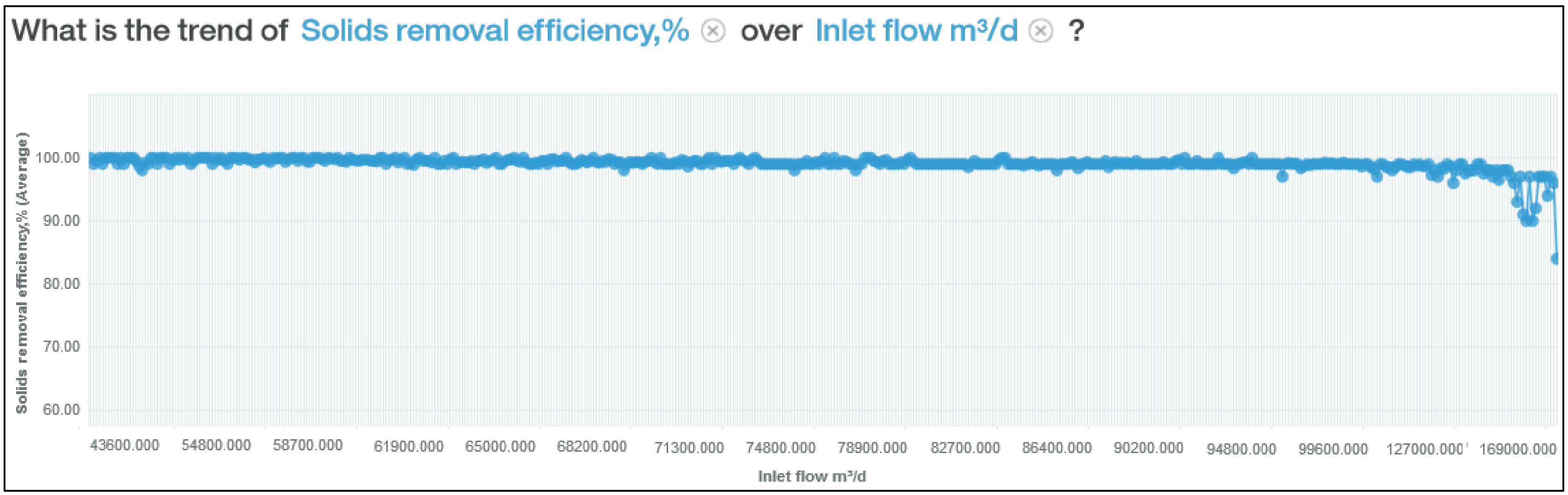
	Intlet P total mg/l	Outlet P total mg/l	P removal efficiency , %	Diss. P treated mg/l	Total inlet N mg/l	Total outlet N mg/l	N removal efficiency, %	NH4-N Inlet mg/l	NH4-N Outlet mg/l	NH4-N removal efficiency, %	NO2 Treated mg/l	NO3 Treated mg/l	Inlet solids mg/l	Outlet solids mg/l	Solids removal efficiency, %	Fe mg/l	Diss.Fe mg/l	Nitrificatio n efficiency %	Added Fe sulfate, mg/l	Added polymer s, mg/l
2	9,250	0,121	98,350	0,081	52,200	23,114	56,000	37,300	16,314	56,970	0,065	5,154	356,500	3,833	98,700	0,489	0,128	69,800	115,000	1,500
3	11,605	0,139	98,750	0,089	72,800	16,450	77,300	54,450	3,610	93,350	0,086	10,950	363,000	2,995	99,000	0,469	0,185	95,000	97,000	1,200
4	6,077	0,116	98,045	0,083	45,818	19,855	55,455	31,136	11,105	63,591	0,045	6,682	202,273	2,468	98,818	0,363	0,142	75,682	105,000	1,400
5	6,678	0,138	97,944	0,107	48,278	11,789	75,556	34,722	1,611	95,333	0,034	8,744	251,667	2,767	99,000	0,345	0,148	96,611	123,000	1,000
6	9,341	0,155	98,182	0,112	64,682	12,336	80,682	43,045	2,050	94,909	0,061	8,173	339,091	3,014	99,045	0,415	0,183	96,773	121,000	1,000
7	8,447	0,129	98,474	0,086	64,421	12,200	80,947	44,263	1,105	97,158	0,055	9,226	312,632	2,789	99,158	0,406	0,162	98,211	142,000	1,000
8	8,961	0,120	98,739	0,080	62,739	11,109	81,870	47,870	1,117	97,609	0,054	8,243	363,043	3,726	98,826	0,409	0,167	98,217	180,000	1,800
9	8,439	0,110	98,565	0,084	58,087	9,143	83,609	45,130	1,313	96,478	0,049	6,522	345,652	2,239	99,391	0,355	0,189	97,435	164,000	2,200
10	8,385	0,119	98,500	0,094	58,200	11,215	79,450	42,250	2,710	92,650	0,062	6,730	337,000	2,235	99,350	0,377	0,164	94,650	158,000	2,000
11	5,850	0,132	97,727	0,098	41,136	12,077	70,273	30,045	2,518	90,818	0,065	8,218	280,455	3,655	98,682	0,517	0,140	93,381	190,000	1,700
12	6,564	0,133	97,864	0,099	47,455	11,241	74,955	36,364	1,900	94,591	0,063	7,655	293,636	2,768	99,045	0,461	0,141	95,727	136,000	2,200
13	8,888	0,164	98,000	0,124	65,875	17,138	74,063	52,125	2,031	96,063	0,137	13,281	362,500	3,931	98,813	0,667	0,208	96,875	160,000	2,300
14	7,030	0,127	98,130	0,094	52,391	18,735	62,478	40,713	4,943	85,435	0,141	12,291	275,652	2,835	98,826	0,511	0,194	89,391	92,000	1,600
15	9,275	0,180	98,100	0,127	67,200	19,400	71,100	52,850	2,570	95,150	0,110	14,810	348,500	3,615	98,950	0,546	0,208	96,150	84,000	1,700
16	10,245	0,161	98,300	0,116	75,800	19,900	73,700	55,900	3,090	94,500	0,079	14,040	380,500	3,130	99,000	0,521	0,193	95,850	114,000	1,500
17	5,853	0,127	97,579	0,074	44,842	16,000	62,737	31,526	4,863	79,737	0,051	9,794	232,105	4,558	97,474	0,562	0,145	86,737	124,000	2,000
18	8,705	0,121	98,636	0,087	61,045	12,132	79,227	46,182	2,123	94,955	0,044	8,127	342,273	2,277	99,227	0,412	0,155	96,227	100,000	2,000
19	10,447	0,155	98,632	0,128	68,316	11,311	83,105	53,211	1,158	97,737	0,042	8,195	416,842	2,989	99,105	0,514	0,187	98,158	140,000	2,000
20	11,478	0,173	98,565	0,147	77,261	18,135	76,261	57,957	2,735	95,261	0,111	13,017	438,261	3,696	99,217	0,545	0,181	96,304	144,000	1,300
21	9,986	0,130	98,818	0,112	70,500	10,682	84,591	54,955	1,469	97,045	0,050	7,850	414,091	1,727	99,773	0,381	0,158	97,818	186,000	1,100
22	11,390	0,131	98,810	0,098	79,238	10,648	86,571	55,619	0,924	98,381	0,053	7,705	454,762	1,767	99,619	0,436	0,178	98,810	125,000	1,300
23	10,709	0,180	98,174	0,148	75,391	11,348	84,696	54,870	0,809	98,000	0,040	8,704	444,783	2,600	99,435	0,423	0,142	98,696	179,000	1,200
24	6,245	0,132	97,850	0,095	47,650	10,610	77,050	32,700	3,120	89,400	0,065	6,370	300,500	3,275	98,900	0,586	0,136	92,850	128,000	1,500
25	6,406	0,138	97,778	0,101	50,500	10,867	76,389	36,111	3,706	88,167	0,292	5,294	278,333	2,772	98,833	0,462	0,166	91,722	83,000	1,700

Appendix 2 – Inlet flowrate

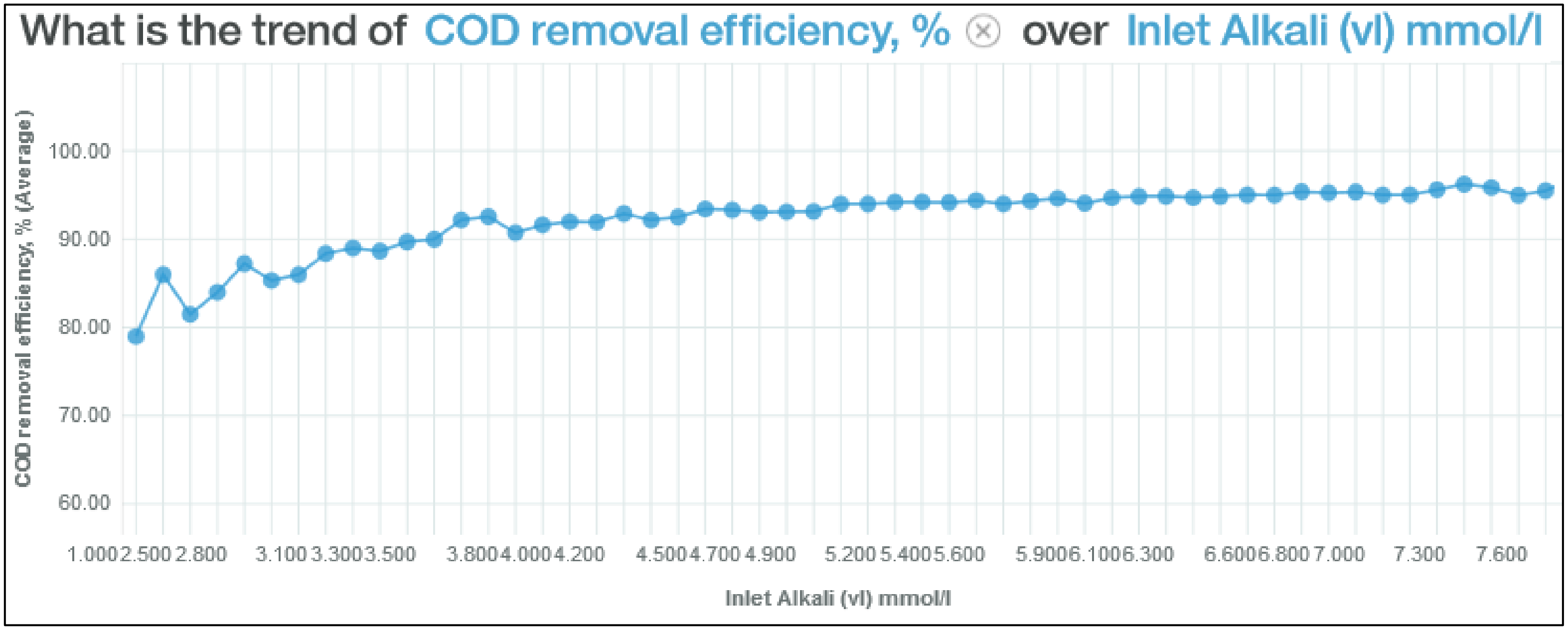


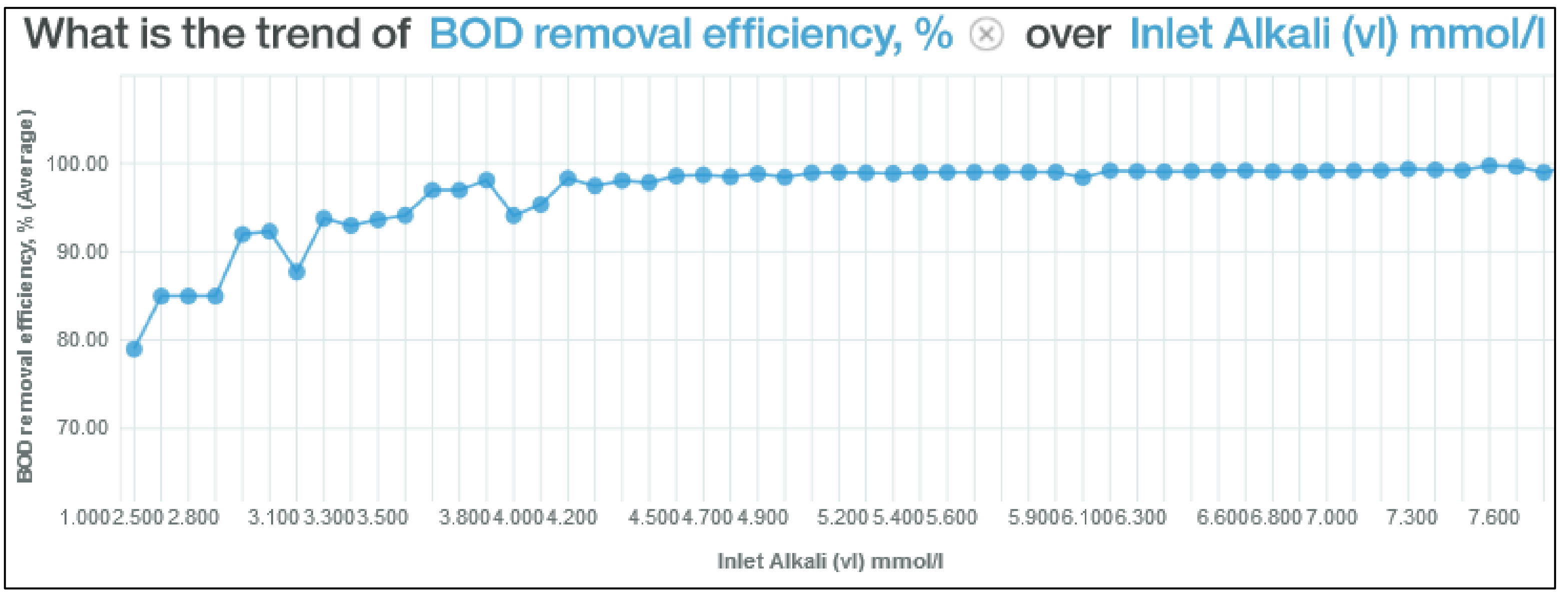


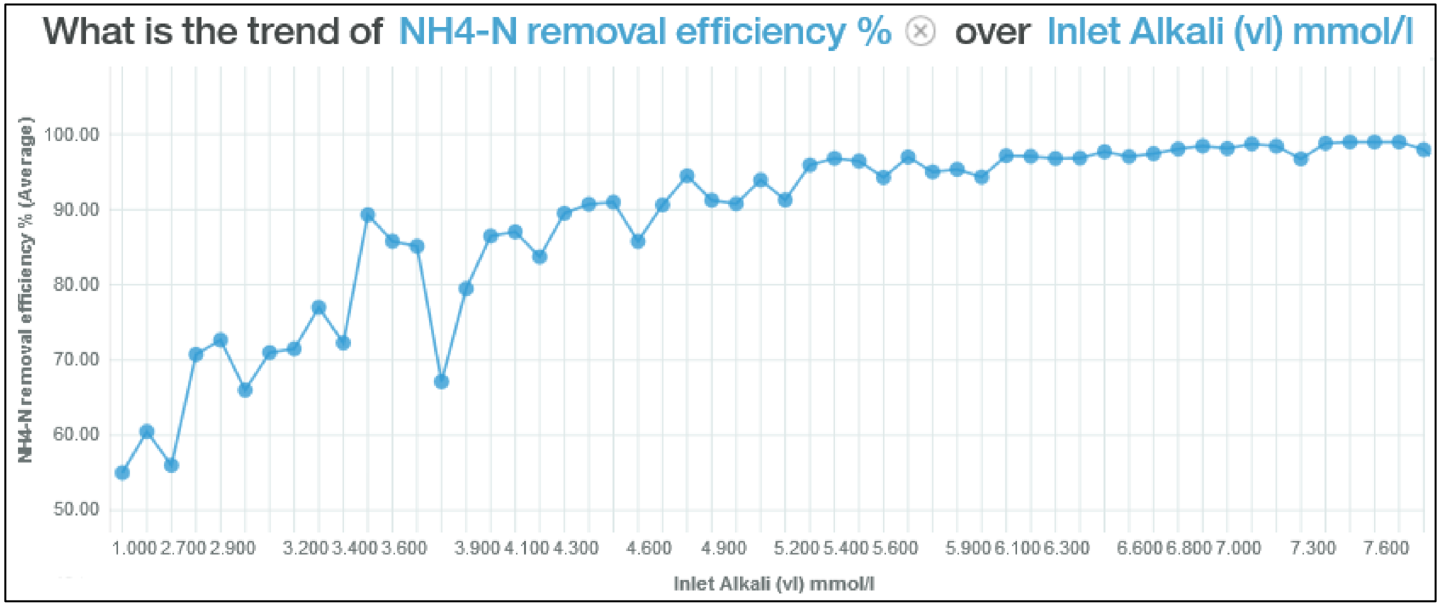

What is the trend of P removal efficiency % ⊗ over Inlet flow m³/d ⊗ ?
P removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
43600.000
54800.000
58700.000
61900.000
65000.000
68200.000
71300.000
74800.000
78900.000
82700.000
86400.000
90200.000
94800.000
99600.000
127000.000
169000.000
Inlet flow m³/d

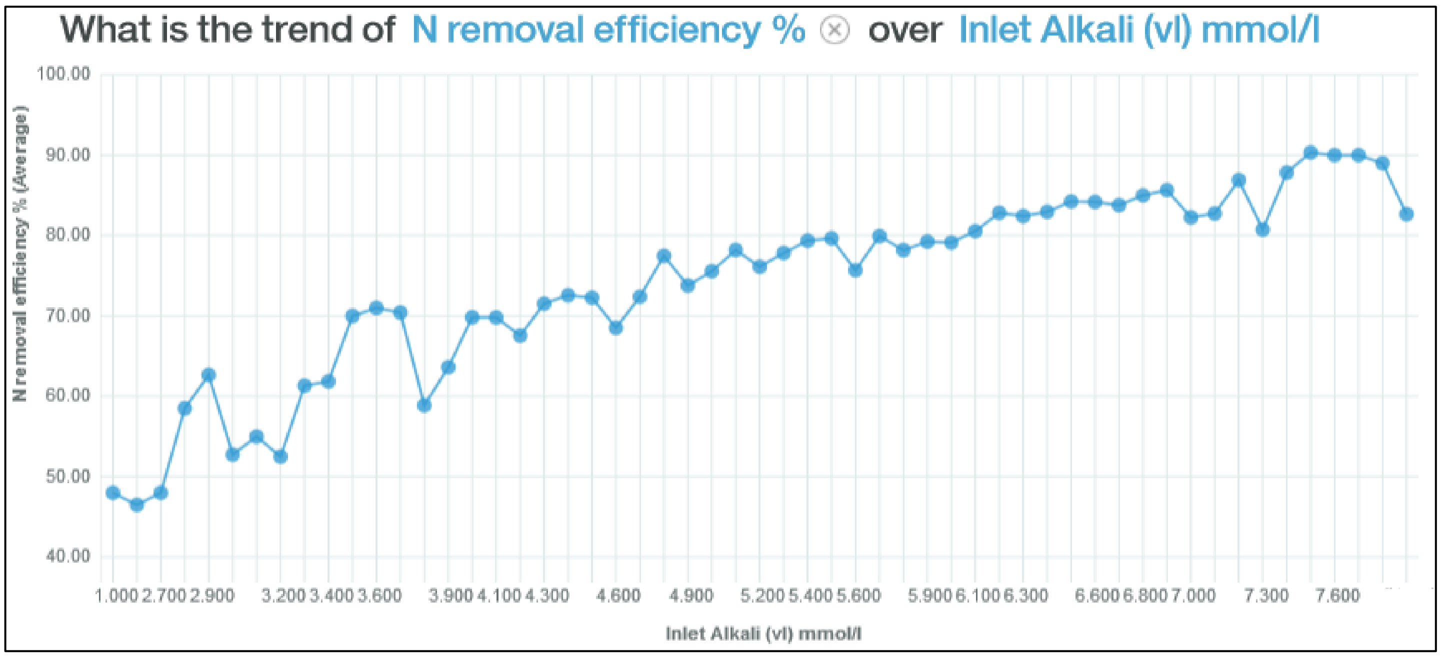

What is the trend of Solids removal efficiency,% ⊗ over Inlet flow m³/d ⊗ ?
Solids removal efficiency,% (Average)
100.00
90.00
80.00
70.00
60.00
43600.000
54800.000
58700.000
61900.000
65000.000
68200.000
71300.000
74800.000
78900.000
82700.000
86400.000
90200.000
94800.000
99600.000
127000.000
169000.000
Inlet flow m³/d

Appendix 3 – Inlet Alkali concentration

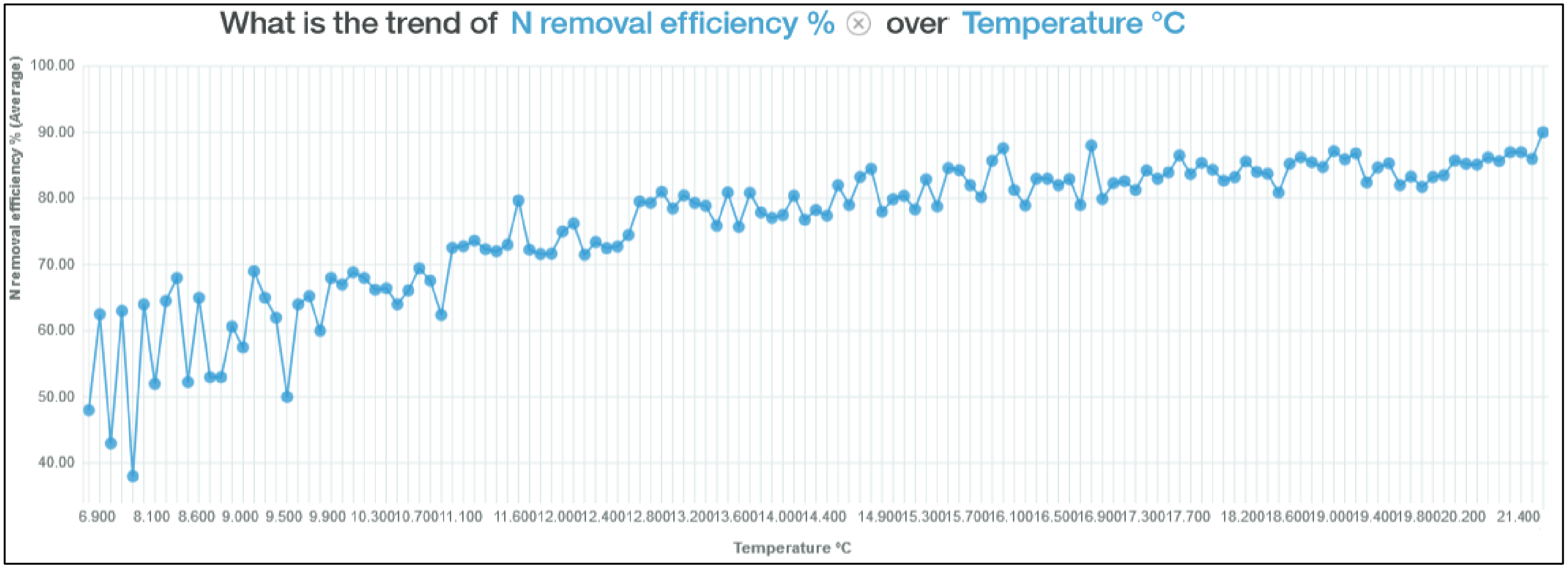


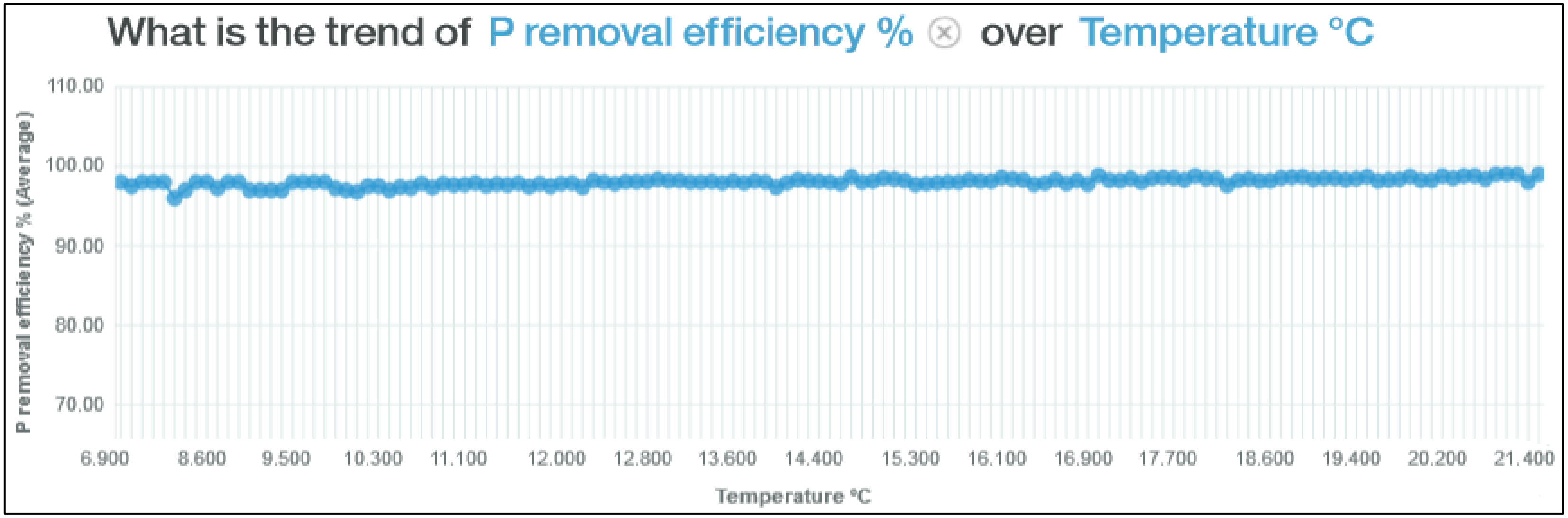


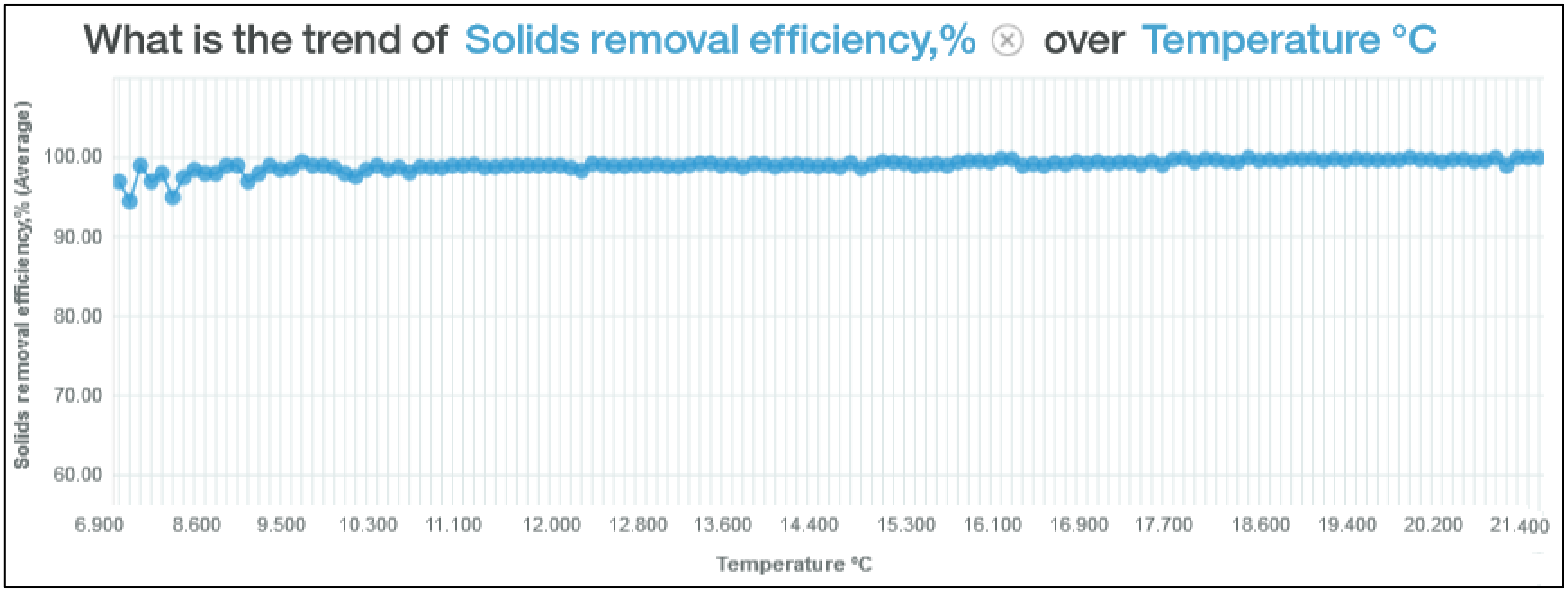

What is the trend of NH4-N removal efficiency % over Inlet Alkali (vl) mmol/l
NH4-N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
1.000 2.700 2.900 3.200 3.400 3.600 3.900 4.100 4.300 4.600 4.900 5.200 5.400 5.600 5.900 6.100 6.300 6.600 6.800 7.000 7.300 7.600
Inlet Alkali (vl) mmol/l

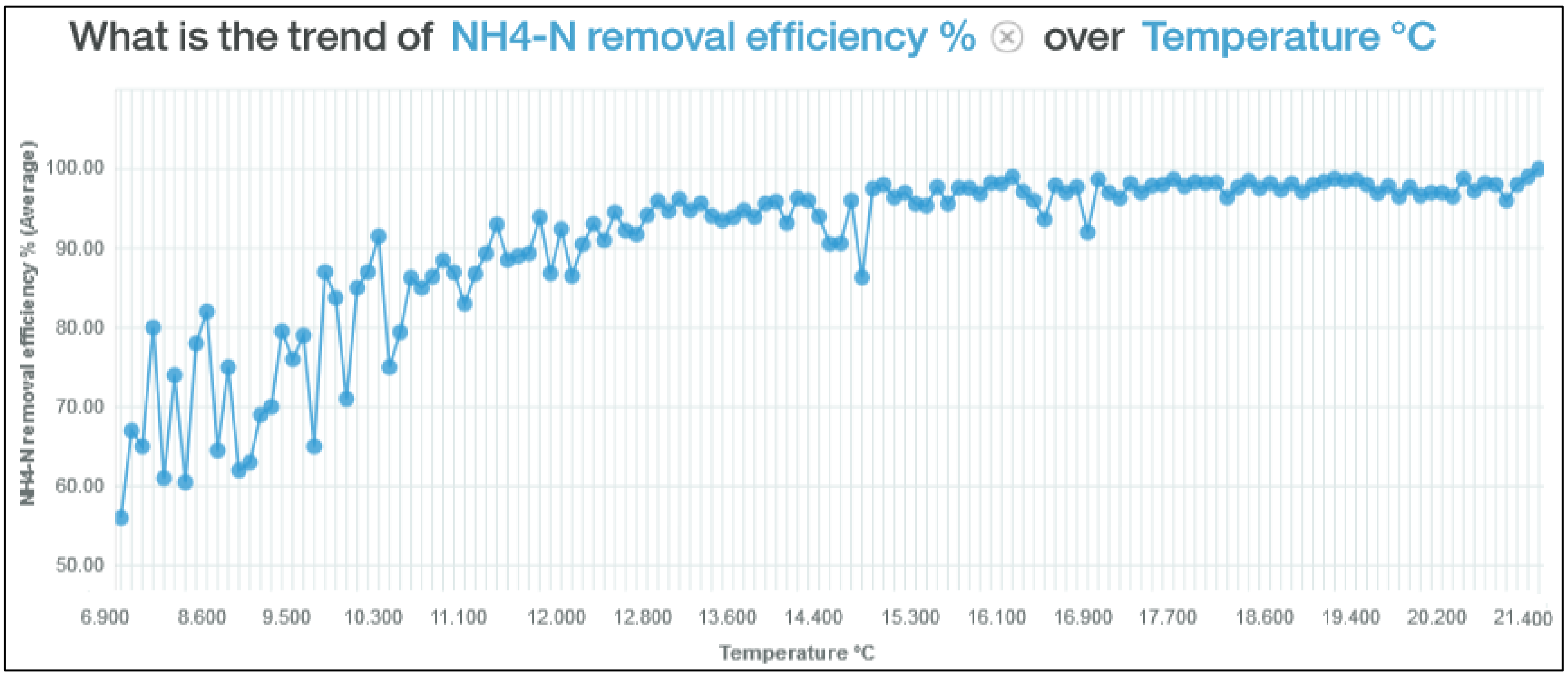

What is the trend of N removal efficiency % over Inlet Alkali (vl) mmol/l
N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
40.00
1.000 2.700 2.900 3.200 3.400 3.600 3.900 4.100 4.300 4.600 4.900 5.200 5.400 5.600 5.900 6.100 6.300 6.600 6.800 7.000 7.300 7.600
Inlet Alkali (vl) mmol/l

Appendix 4 – Temperature

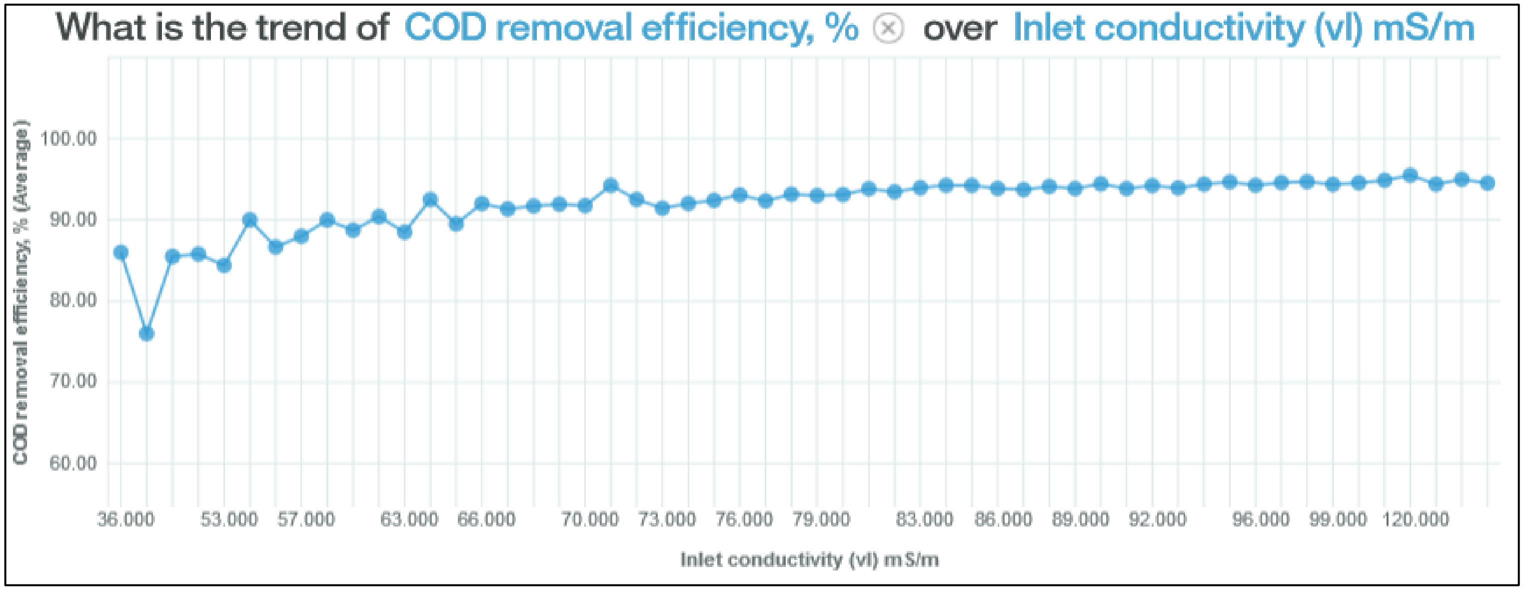


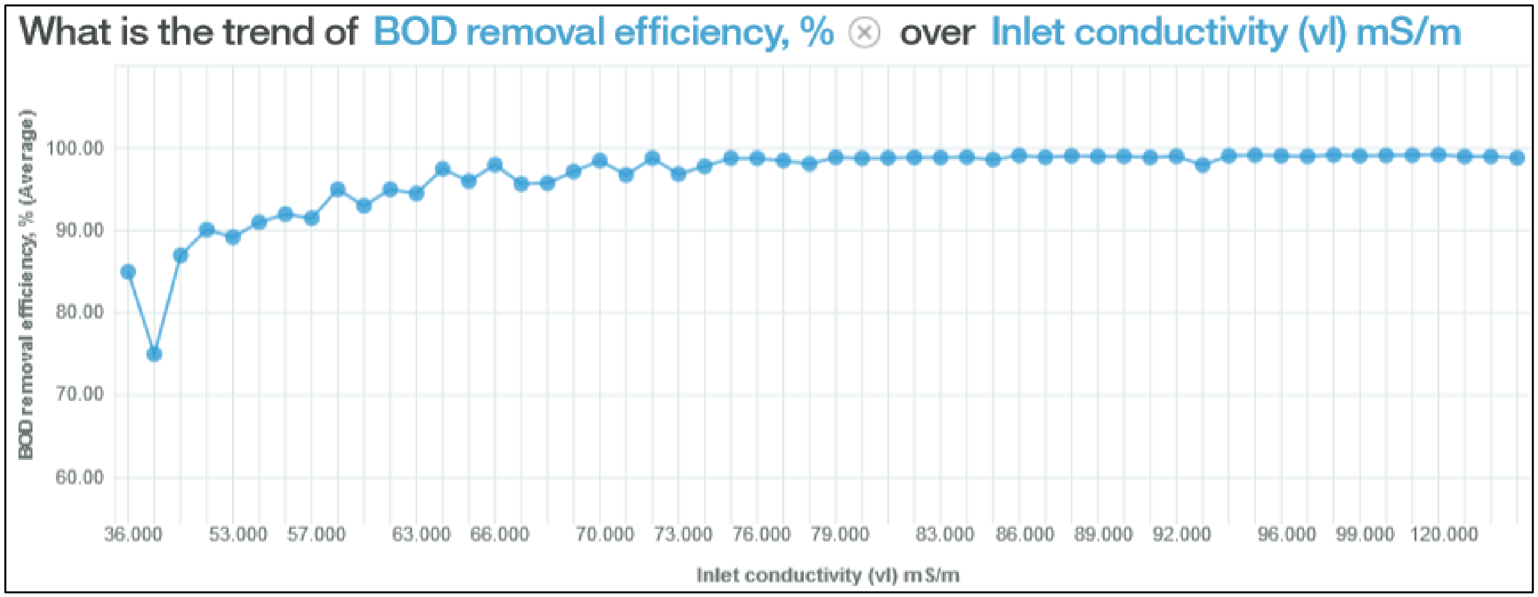


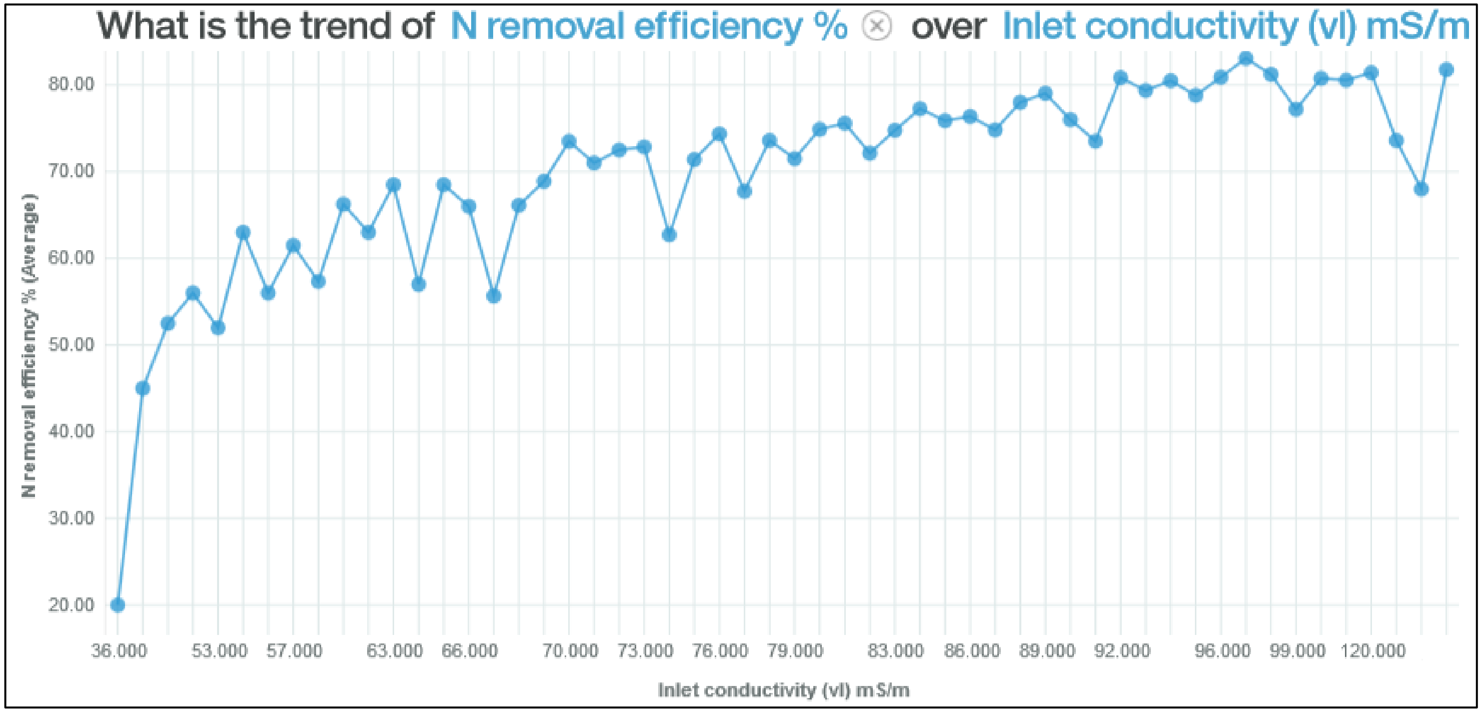

What is the trend of Solids removal efficiency,% over Temperature °C
Solids removal efficiency,% (Average)
100.00
90.00
80.00
70.00
60.00
6.900
8.600
9.500
10.300
11.100
12.000
12.800
13.600
14.400
15.300
16.100
16.900
17.700
18.600
19.400
20.200
21.400
Temperature °C

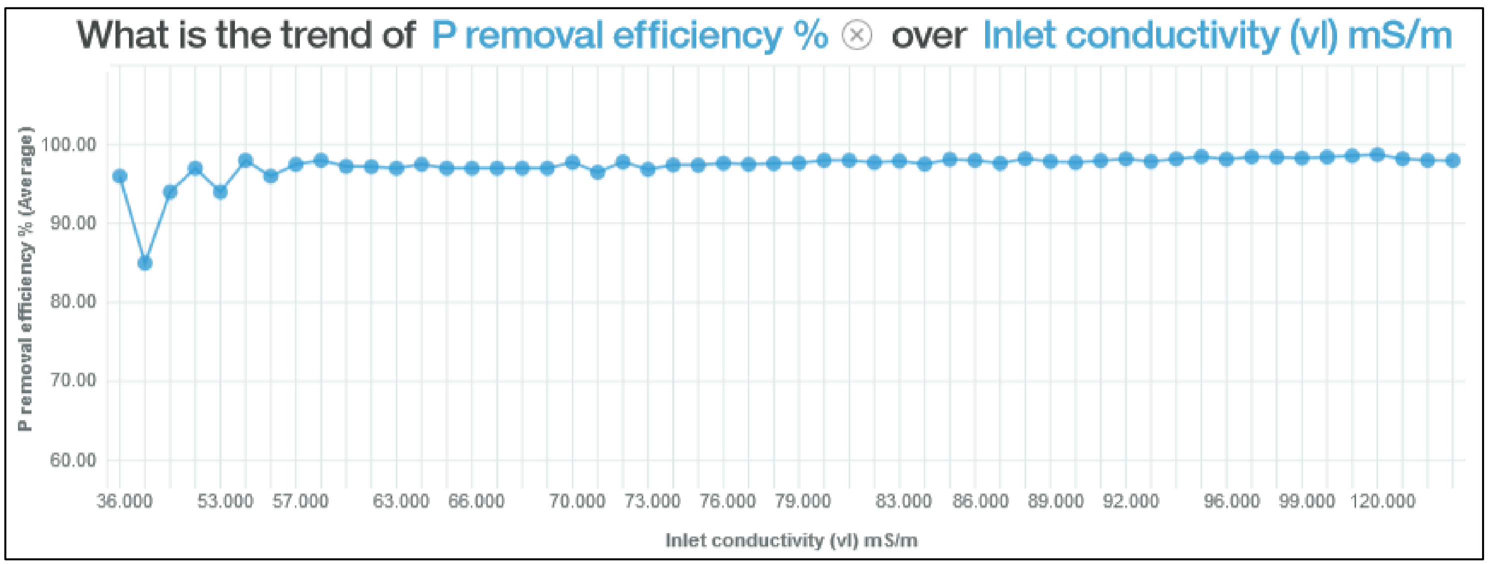

What is the trend of NH4-N removal efficiency % over Temperature °C
NH4-N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
6.900
8.600
9.500
10.300
11.100
12.000
12.800
13.600
14.400
15.300
16.100
16.900
17.700
18.600
19.400
20.200
21.400
Temperature °C

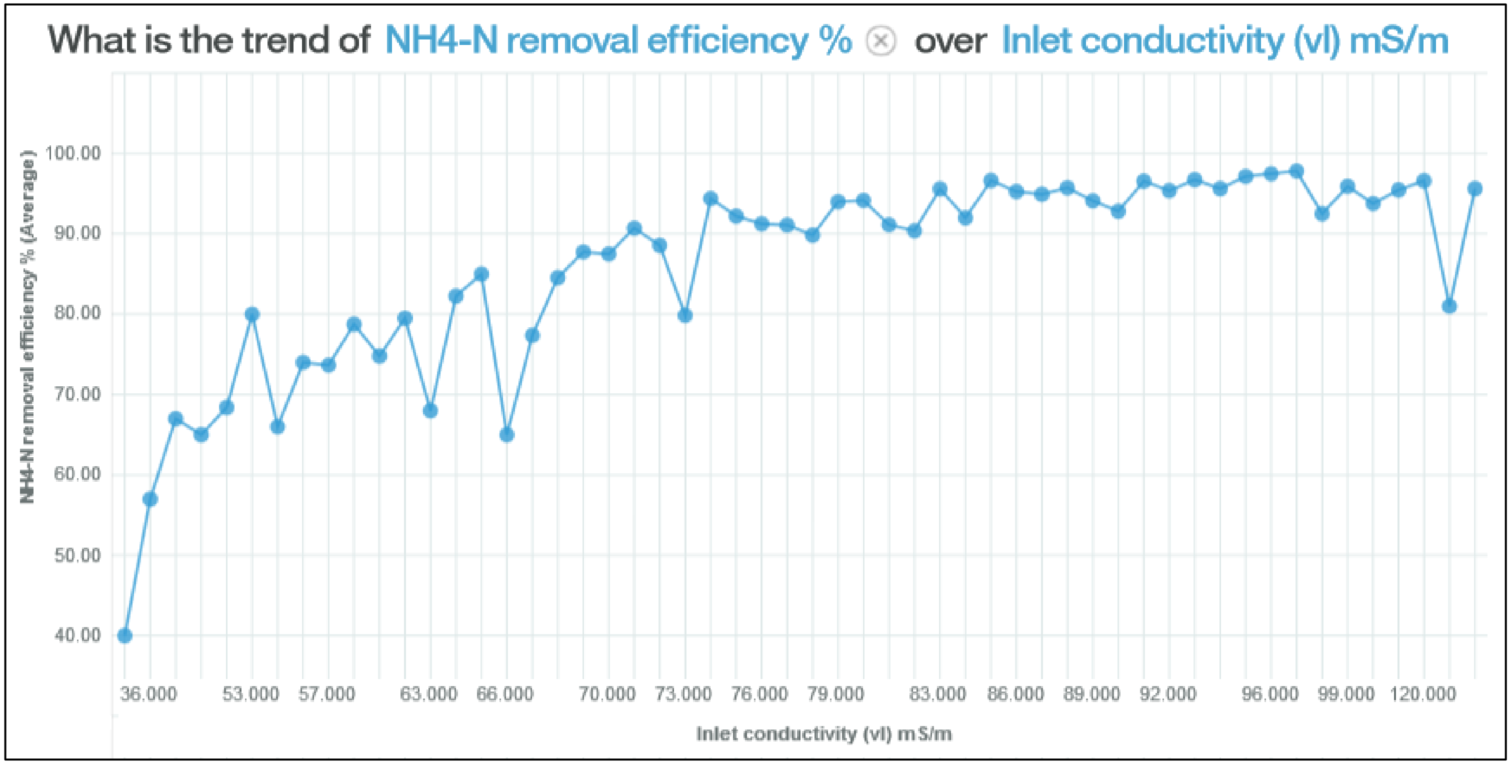
Appendix 5 – Inlet conductivity



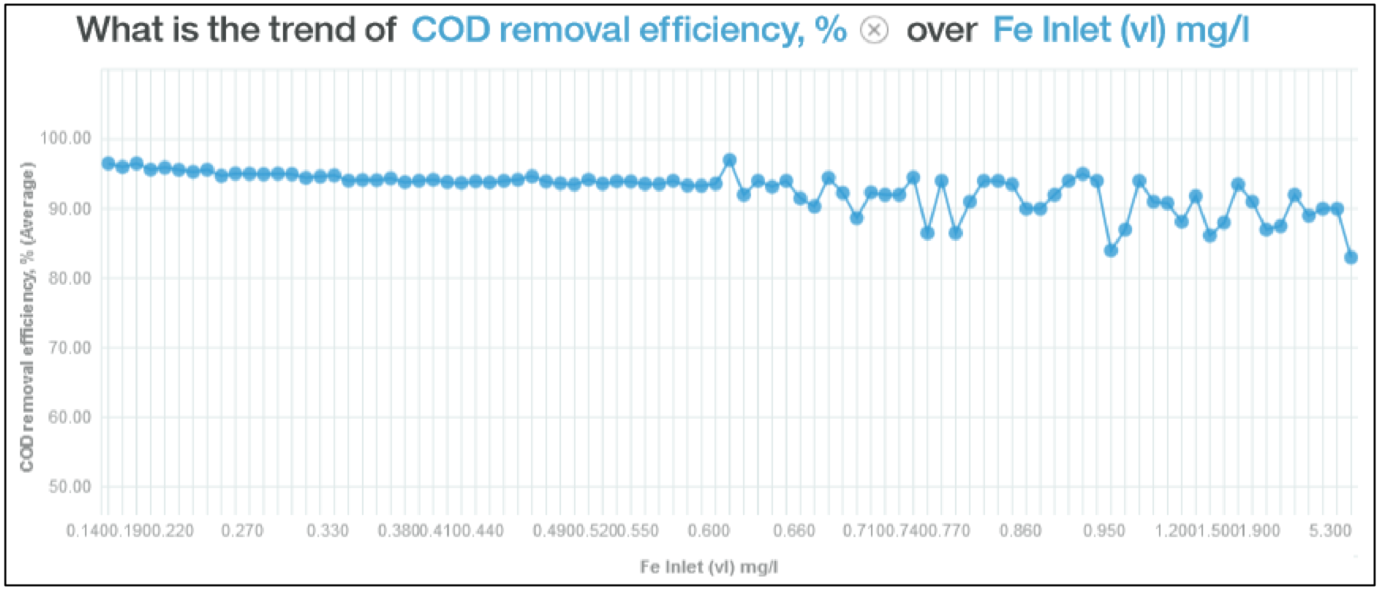


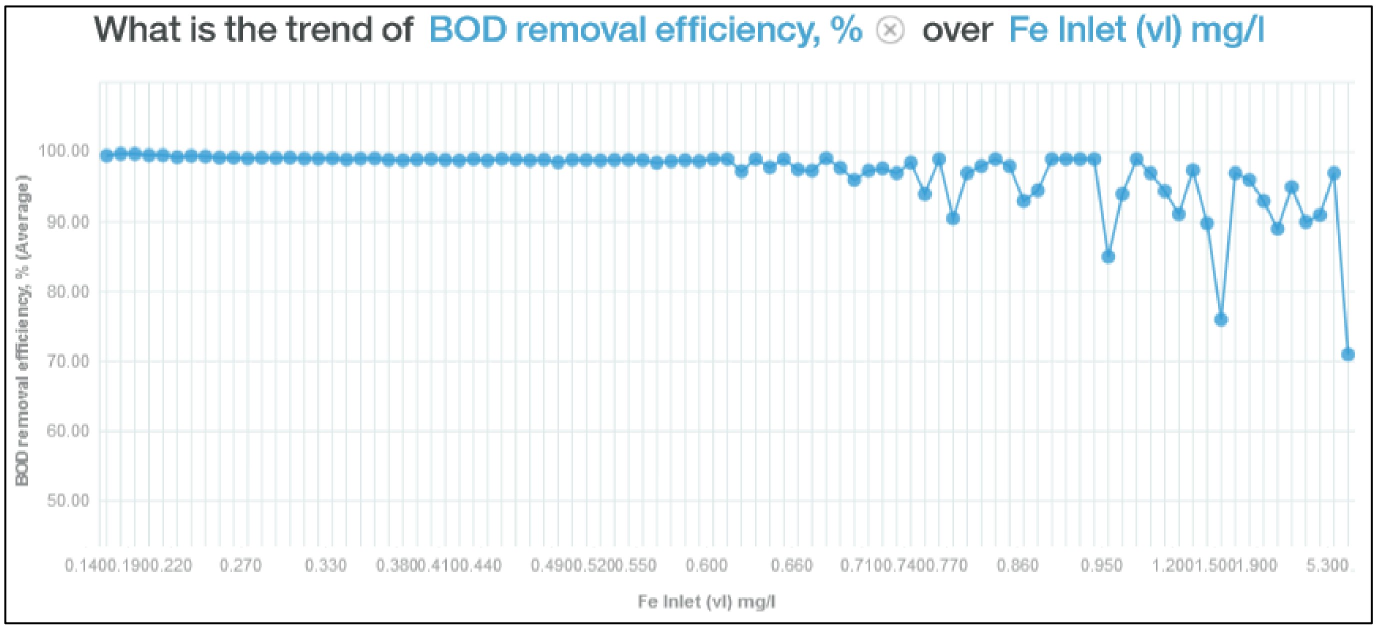

What is the trend of N removal efficiency % over Inlet conductivity (vl) mS/m
N removal efficiency % (Average)
80.00
70.00
60.00
50.00
40.00
30.00
20.00
36.000
53.000
57.000
63.000
66.000
70.000
73.000
76.000
79.000
83.000
86.000
89.000
92.000
96.000
99.000
120.000
Inlet conductivity (vl) mS/m

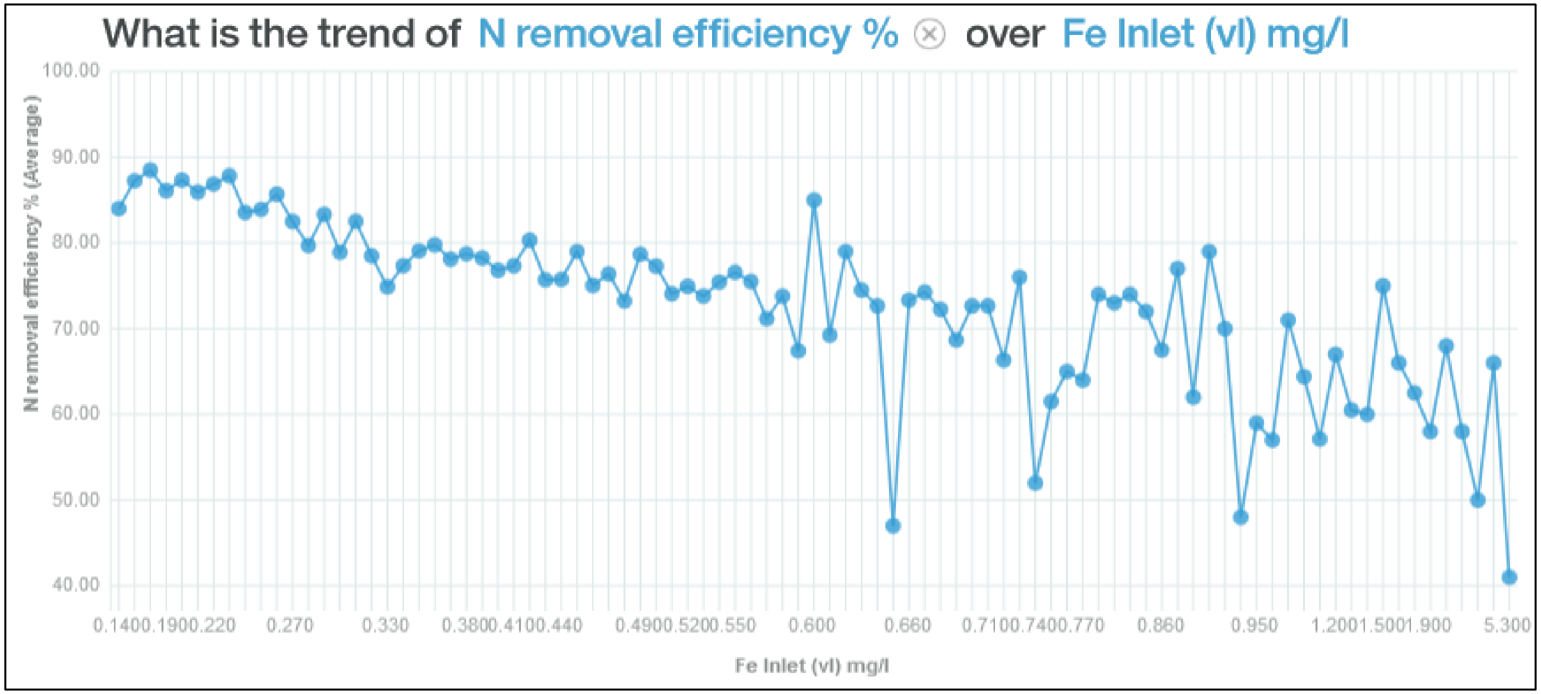

What is the trend of P removal efficiency % over Inlet conductivity (vl) mS/m
P removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
36.000
53.000
57.000
63.000
66.000
70.000
73.000
76.000
79.000
83.000
86.000
89.000
92.000
96.000
99.000
120.000
Inlet conductivity (vl) mS/m


What is the trend of NH4-N removal efficiency % over Inlet conductivity (vl) mS/m
NH4-N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
40.00
36.000
53.000
57.000
63.000
66.000
70.000
73.000
76.000
79.000
83.000
86.000
89.000
92.000
96.000
99.000
120.000
Inlet conductivity (vl) mS/m

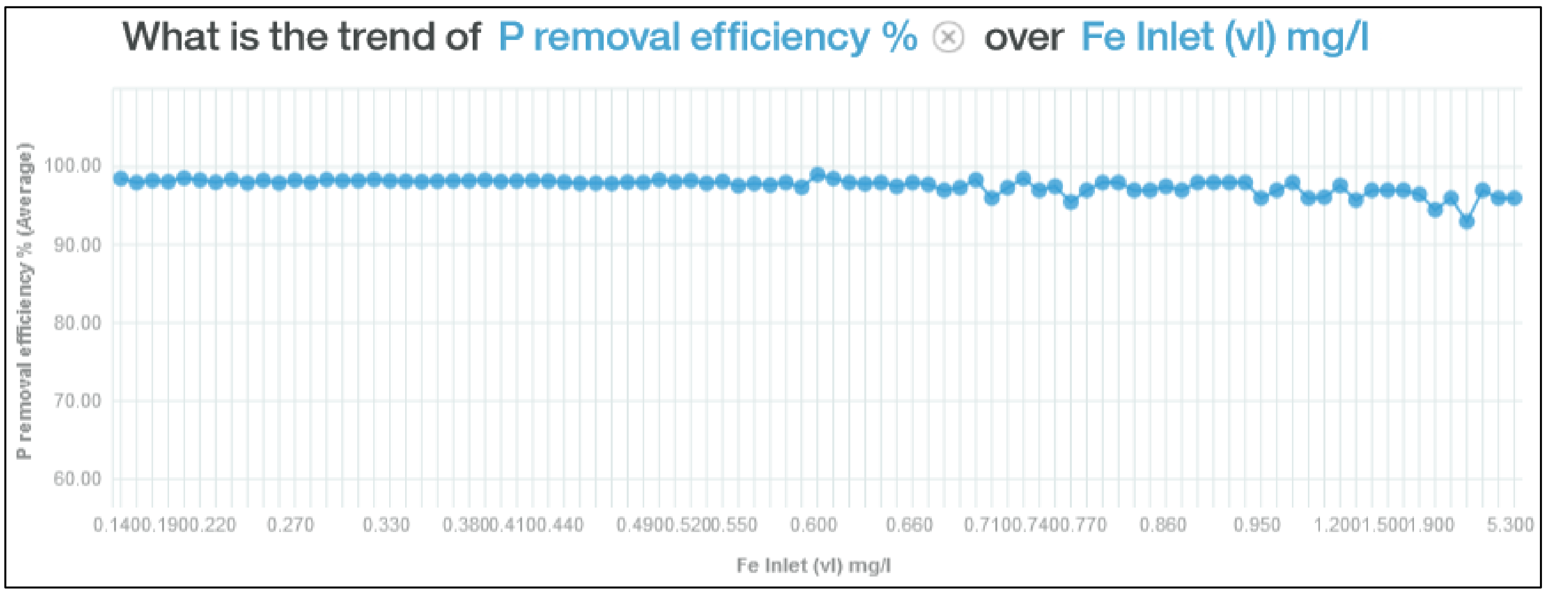
Appendix 6 – Iron concentration



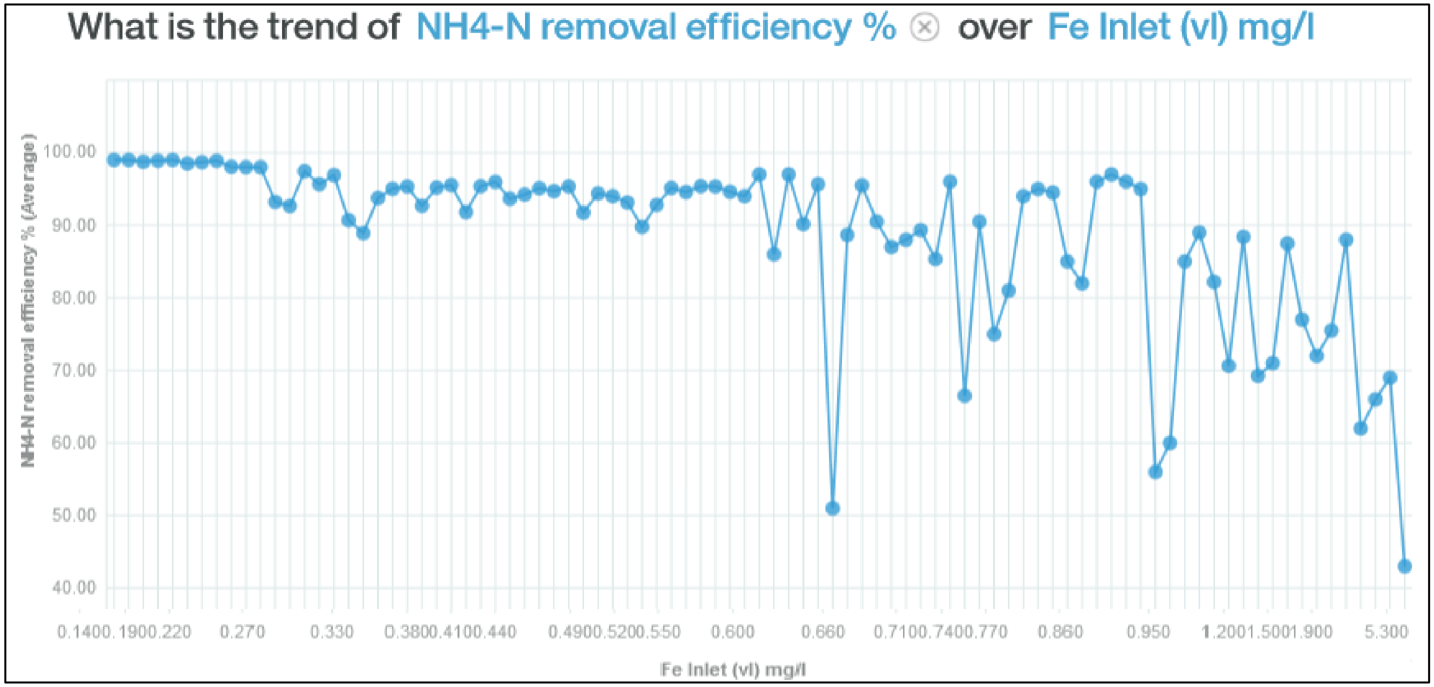


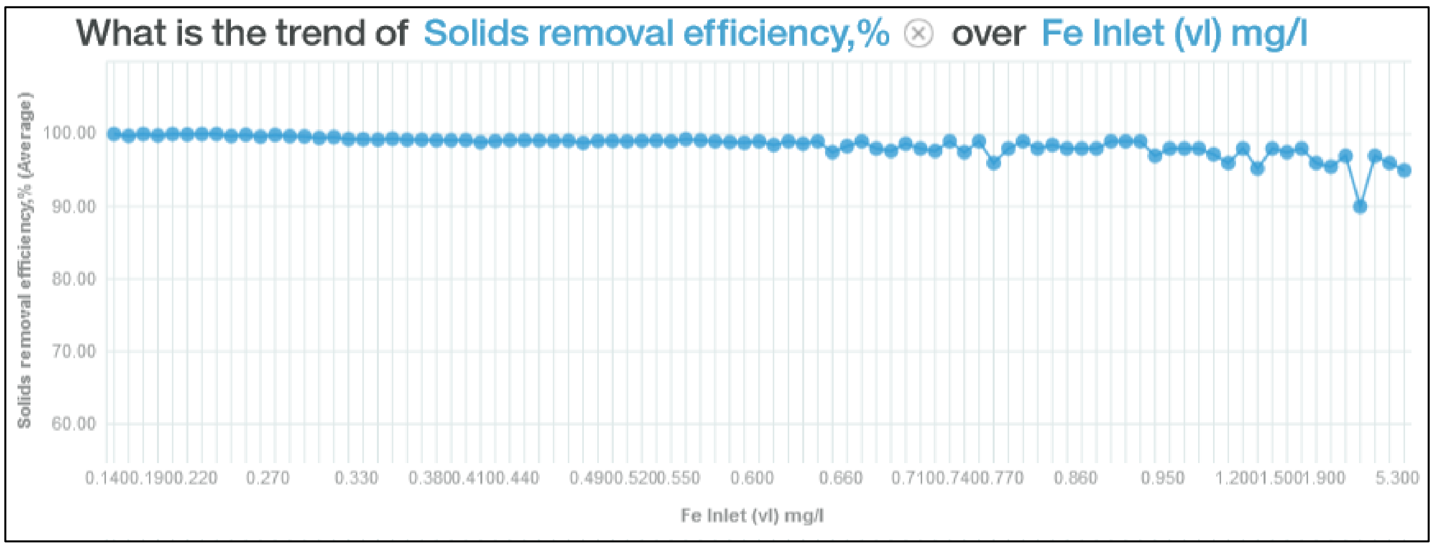


What is the trend of N removal efficiency % ⊗ over Fe Inlet (vl) mg/l
N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
40.00
0.1400.1900.220 0.270 0.330 0.3800.4100.440 0.4900.5200.550 0.600 0.660 0.7100.7400.770 0.860 0.950 1.2001.5001.900 5.300
Fe Inlet (vl) mg/l



What is the trend of P removal efficiency % ⊗ over Fe Inlet (vl) mg/l
P removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
0.1400.1900.220 0.270 0.330 0.3800.4100.440 0.4900.5200.550 0.600 0.660 0.7100.7400.770 0.860 0.950 1.2001.5001.900 5.300
Fe Inlet (vl) mg/l


What is the trend of NH4-N removal efficiency % over Fe Inlet (vl) mg/l
NH4-N removal efficiency % (Average)
100.00
90.00
80.00
70.00
60.00
50.00
40.00
0.1400.1900.220
0.270
0.330
0.3800.4100.440
0.4900.5200.550
0.600
0.660
0.7100.7400.770
0.860
0.950
1.2001.5001.900
5.300
Fe Inlet (vl) mg/l


What is the trend of Solids removal efficiency,% over Fe Inlet (vl) mg/l
Solids removal efficiency,% (Average)
100.00
90.00
80.00
70.00
60.00
0.1400.1900.220
0.270
0.330
0.3800.4100.440
0.4900.5200.550
0.600
0.660
0.7100.7400.770
0.860
0.950
1.2001.5001.900
5.300
Fe Inlet (vl) mg/l

Appendix 7 – Removal efficiency drivers

Removal efficiency	Drivers	Strength, %
Phosphorus	Total Nitrogen inlet and Total Phosphorus Outlet	89
	Inlet ammonia and Total Phosphorus Outlet	88
	Dissolved Iron and Inlet solids	79
	Treated NO_3^- concentration and Inlet flow	78
	Dissolved Iron and Inlet conductivity	77
	Inlet solids and Total Nitrogen outlet	76
	Dissolved Iron and Inlet flow	74
	Inlet COD and outlet alkali	73
	Nitrogen removal efficiency and Total Phosphorus Outlet	73
	Treated NO_3^- concentration and Total Nitrogen inlet	72
	Outlet COD and Inlet flow	71
	Inlet solids and Outlet BOD	71
	Concentration of Iron and Total Nitrogen inlet	70
	Inlet conductivity and Temperature	70
Biochemical Oxygen Demand (BOD)	Concentration of dissolved treated phosphorus and phosphorus removal efficiency	87
	Concentration of Iron and Concentration of dissolved treated phosphorus	84
	Concentration of dissolved treated phosphorus and Outlet conductivity	83
	Treated NO_3^- concentration and ammonia removal efficiency	82
	Iron concentration and Total Phosphorus Outlet	82
	Nitrification efficiency and Treated NO_3^- concentration	81
	Nitrification efficiency and Treated NO_2^- concentration	81
	Concentration of dissolved treated phosphorus and inlet conductivity	80
	Phosphorus removal efficiency and Outlet BOD	80
	Treated NO_3^- concentration and ammonia outlet concentration	79
	Nitrification efficiency and phosphorus removal efficiency	79
	Added polymers and Inlet alkali concentration	77
	Added polymers and Total Nitrogen outlet	77
	Added polymers and COD removal efficiency	77
	Iron concentration and Inlet COD concentration	77
	Inlet conductivity and Outlet Alkali concentration	77
	Iron concentration and Phosphorus removal efficiency	76
	Total Nitrogen outlet concentration and Inlet alkali concentration	76
	Added polymers and Inlet ammonia concentration	75
	Outlet conductivity and Outlet alkali concentration	74
	Outlet solids and Treated NO_2^- concentration	74
	Phosphorus removal efficiency and Total phosphorus outlet concentration	74
	Treated NO_2^- concentration and Inlet flow	72
	Inlet solids and Inlet pH	70
Chemical Oxygen demand (COD)	Treated NO_3^- concentration and Ammonia removal efficiency	87
	Outlet solids and Total Nitrogen inlet	85
	Nitrification efficiency and Inlet Conductivity	85
	Nitrification efficiency and Phosphorus removal efficiency	85
	Phosphorus removal efficiency and BOD removal efficiency	84
	Added polymers and Inlet alkali concentration	84
	Treated NO_3^- concentration and BOD removal efficiency	83
	Solids removal efficiency and Inlet conductivity	82

	Nitrogen removal efficiency and Outlet COD concentration	82
	Outlet COD concentration and Inlet conductivity	81
	Added polymers and Inlet ammonia concentration	81
	Treated NO_3^- concentration and Outlet ammonia concentration	81
	Dissolved Iron concentration and Inlet flow	81
	Concentration of dissolved treated phosphorus and BOD removal efficiency	81
	Solids removal efficiency and Outlet COD concentration	79
	Iron concentration and Concentration of total outlet phosphorus	77
	Outlet solid sand Outlet alkali concentration	76
	Outlet BOD concentration and Inlet conductivity	75
	Outlet solids concentration and Outlet COD concentration	73
	Inlet COD concentration	73
	Ammonia outlet concentration and Outlet COD concentration	73
	Phosphorus removal efficiency and Outlet BOD concentration	72
	Dissolved Iron and Total phosphorus outlet concentration	71
Nitrogen	Ammonia removal efficiency and Total Nitrogen outlet concentration	93
	Total outlet Nitrogen concentration and COD removal efficiency	91
	Total outlet Phosphorus concentration and outlet pH	90
	Total outlet Nitrogen concentration and Total outlet Phosphorus concentration	86
	Total outlet Nitrogen concentration and Outlet COD	86
	Ammonia outlet concentration and Total outlet Phosphorus concentration	85
	Total outlet Nitrogen concentration and Outlet alkali concentration	84
	Ammonia removal efficiency and Outlet conductivity	83
	Added polymers and Temperature	83
	Solids removal efficiency and Inlet conductivity	81
	Treated NO_2^- concentration and Inlet COD	81
	Nitrification efficiency and Iron concentration	80
	Nitrification efficiency and Total outlet Phosphorus concentration	80
Ammonia	Treated NO_3^- concentration and Nitrogen removal efficiency	90
	Added Fe sulfate and Solids removal efficiency	90
	BOD removal efficiency and Outlet pH	83
	BOD removal efficiency and Outlet pH	82
	Solids removal efficiency and Inlet conductivity	80
	COD removal efficiency and Temperature	79
	COD removal efficiency and Inlet conductivity	79
	COD removal efficiency and Inlet pH	78
	Inlet BOD and Inlet Conductivity	78
Nitrification	Treated NO_3^- concentration and Nitrogen removal efficiency	91
	Ammonia removal efficiency and phosphorus removal efficiency	91
	Ammonia removal efficiency and outlet conductivity	91
	Added Fe sulfate and Solids removal efficiency	90
	BOD removal efficiency and outlet pH	84
	Added polymers and Temperature	79